Article

On the Potential of 3D Transdimensional Surface Wave Tomography for Geothermal Prospecting of the Reykjanes Peninsula

Amin Rahimi Dalkhani [1,*], Xin Zhang [2] and Cornelis Weemstra [1,3]

[1] Department of Geoscience and Engineering, Delft University of Technology, 2628 CN Delft, The Netherlands; c.weemstra@tudelft.nl
[2] School of Geosciences, University of Edinburgh, Edinburgh EH8 9XP, UK; x.zhang2@ed.ac.uk
[3] Seismology and Acoustics, Royal Netherlands Meteorological Institute, 3731 GA De Bilt, The Netherlands
[*] Correspondence: a.rahimidalkhani@tudelft.nl

Abstract: Seismic travel time tomography using surface waves is an effective tool for three-dimensional crustal imaging. Historically, these surface waves are the result of active seismic sources or earthquakes. More recently, however, surface waves retrieved through the application of seismic interferometry have also been exploited. Conventionally, two-step inversion algorithms are employed to solve the tomographic inverse problem. That is, a first inversion results in frequency-dependent, two-dimensional maps of phase velocity, which then serve as input for a series of independent, one-dimensional frequency-to-depth inversions. As such, a set of localized depth-dependent velocity profiles are obtained at the surface points. Stitching these separate profiles together subsequently yields a three-dimensional velocity model. Relatively recently, a one-step three-dimensional non-linear tomographic algorithm has been proposed. The algorithm is rooted in a Bayesian framework using Markov chains with reversible jumps, and is referred to as transdimensional tomography. Specifically, the three-dimensional velocity field is parameterized by means of a polyhedral Voronoi tessellation. In this study, we investigate the potential of this algorithm for the purpose of recovering the three-dimensional surface-wave-velocity structure from ambient noise recorded on and around the Reykjanes Peninsula, southwest Iceland. To that end, we design a number of synthetic tests that take into account the station configuration of the Reykjanes seismic network. We find that the algorithm is able to recover the 3D velocity structure at various scales in areas where station density is high. In addition, we find that the standard deviation of the recovered velocities is low in those regions. At the same time, the velocity structure is less well recovered in parts of the peninsula sampled by fewer stations. This implies that the algorithm successfully adapts model resolution to the density of rays. It also adapts model resolution to the amount of noise in the travel times. Because the algorithm is computationally demanding, we modify the algorithm such that computational costs are reduced while sufficiently preserving non-linearity. We conclude that the algorithm can now be applied adequately to travel times extracted from station–station cross correlations by the Reykjanes seismic network.

Keywords: seismic interferometry; transdimensional tomography; surface wave dispersion; probabilistic inversion; Markov chain Monte Carlo

1. Introduction

The Reykjanes high temperature geothermal system is located at the tip of the Reykjanes peninsula, southwest Iceland. The Reykjanes peninsula can be considered a landward extension of the Reykjanes Ridge, making it an active volcanic zone. The heat from the Reykjanes geothermal system is currently harvested by two 50 MWe steam turbines, which draw from numerous wells. Most of these wells produce from depths of 2 to 2.5 km, where temperatures are estimated to be around 280° [1]. Relatively recently, the 4.5 km deep

exploratory IDDP-2 well was drilled to examine the economic potential of the production of super critical fluids [2].

Although the subsurface below the Reykjanes peninsula has been the subject of various geophysical surveys (e.g., [3,4]), resolution remains an issue. In addition, the uncertainty of the recovered (best) models is often not estimated. In this study, we therefore investigate the potential of three-dimensional probabilistic surface wave inversion for the purpose of imaging the Reykjanes peninsula in general, and the Reykjanes geothermal system in particular. It should be understood that this is a feasibility study: we limit ourselves to probabilistic inversions of synthetic travel times. In a follow-up study we will draw from the experience obtained in this study. This follow-up paper will hence involve probabilistic inversion of travel times extracted from field-data-derived surface wave responses.

Surface wave responses can be retrieved from recordings of ambient seismic noise (e.g., [5]). This process is often referred to as seismic interferometry, and hence the retrieved responses are usually referred to as 'interferometric responses' [6]. Specifically, seismic interferometry dictates that, by averaging cross correlations of ambient seismic noise over a sufficient amount of time, receivers can be turned into so-called virtual sources. This is a pairwise process: the Earth's response to one of the two receivers (the virtual source) is recorded by the other receiver. By applying this procedure, pairwise, to all stations constituting an array of seismic surface stations, surface wave responses between all seismic stations will be retrieved. In practice, not all station couples yield interferometric surface wave responses of sufficient quality (for details regarding interferometric surface wave retrieval from recordings of ambient seismic noise see, for example, Kästle et al. [7]).

Interferometric techniques may well play an important role in the future of geothermal exploration (e.g., [8]). In particular, this applies to tomographic travel time inversions: contrary to travel time tomography exploiting earthquake signals, the spatial sampling by the interferometric ray paths does not suffer from either a limited number of earthquakes or an irregular distribution of these earthquakes [3]. In addition, Rayleigh (and Love) waves have a depth-dependent sensitivity to subsurface velocity. This sensitivity varies as a function of frequency, with lower frequencies being more sensitive to the deeper subsurface and higher frequencies more sensitive to the shallower structure [9,10]. It is this frequency-dependent depth dependence that enables the three-dimensional approach proposed in this study.

The interferometric travel times between the different stations of a seismic array can serve as input to a tomographic inverse problem (e.g., [11,12]). Surface wave tomography is a well-known and popular method to obtain the Earth's seismic velocity distribution. Many of the existing tomographic algorithms and inversion methods rely on first-arrival travel times [13–15]. Ref. Martins et al. [3] uses a two-step linearized tomographic inversion method to recover the 3D surface wave velocity of the Reykjanes Peninsula. The first step involves recovering frequency-dependent 2D maps of phase velocities using a linearized tomographic inversion method assuming straight rays. In the second step, they run separate frequency-to-depth inversions [9,10].

Conventional linearized or gradient-based iterative tomographic inversion schemes (e.g., [3,16,17]) usually do not include a detailed assessment of the uncertainty [18]. In addition, such schemes require an (a-priori) prescribed cell size, which does not account for spatial variations in sampling (i.e., a non-uniform ray coverage). Moreover, these methods require a proper regularization to account for the ill-conditioned nature of the inverse problem [17,19]. The transdimensional hierarchical Bayesian method introduced by Bodin et al. [20] overcomes these two limitations. The transdimensional method is a Bayesian inference method that relies on reversible jump Markov chains (rj-McMC) [21], in combination with a Voronoi partitioning. The number of Voronoi cells is allowed to vary between steps of the Markov chain, and is in fact one of the parameters that is estimated. It is this specific feature of the algorithm that causes the solution to be "transdimensional". The ensemble of velocity models visited by the transdimensional Markov chain asymptotically approaches the posterior probability density function (henceforth simply 'posterior').

The method is purely data-driven and requires minimal assumptions regarding the model. Compared to the aforementioned linearized and iterative inversion schemes (which keep the spatial model parameterization fixed), the method is particularly flexible: it dynamically adapts to a non-uniform data coverage without requiring the use of any arbitrary regularization (e.g., damping or smoothing).

The transdimensional method was used successfully by Bodin and Sambridge [11] to obtain 2D Rayleigh-wave velocity maps of Australia. Galetti et al. [22] further generalized the method by making it fully non-linear; that is, ray paths are updated at each step of the Markov chain in their case. The method has recently been applied successfully to the British Isles [23]. Similar to the aforementioned study by Martins et al. [3], Young et al. [18] and Galetti et al. [23] use a two-step scheme to recover the 3D surface wave velocity structure. In the latter studies, however, the frequency-dependent 2D maps of phase velocity are obtained using a 2D transdimensional approach. Additionally, the obtained (laterally varying) dispersion curves are inverted using a 1D transdimensional approach. Still, however, the two-step procedure results in a collection of spatially varying 1D velocity models. These velocity models are subsequently interpolated to build a 3D velocity structure of the subsurface.

More recently, Zhang et al. [24] showed that the two-step transdimensional inversion scheme is not optimal in the sense that information is lost. Consequently, they proposed a one-step 3D transdimensional approach that uses a 3D discretization of the subsurface using a Voronoi polyhedral parameterization. The approach has a computational cost comparable to the two-step transdimensional approach, but preserves valuable information and results in a more accurate velocity structure and a better interpretive uncertainty estimation [24].

In this study, we investigate the potential of the one-step 3D probabilistic inversion method [24,25] to recover the 3D velocity structure beneath the Reykjanes peninsula, southwest Iceland. Specifically, we focus on the ability to use interferometric surface wave responses retrieved from ambient noise recorded by an array of stations on and around this Peninsula (the array is hereafter referred to as the RARR; [26]). The RARR was a dense seismic deployment consisting of 83 continuously recording stations (Figure 1). As can be seen in Figure 1, the station distribution of the RARR is non-uniform, that is, the station coverage is dense in one area while it is sparse in other areas. A non-uniform distribution of stations implies that the achievable velocity resolution can be expected to vary significantly across the region covered by the seismic array, that is, higher in the regions that are more densely covered by stations and lower where station density is low. The transdimensional method uses a model parameterization using Voronoi cells in conjunction with a reversible-jump Markov chain Monte Carlo algorithm to account for the variable station coverage. As such, the method dynamically adapts itself to both data density and underlying velocity structure [11].

Surface wave dispersion curves (i.e., phase velocities or frequency-dependent travel times) have been extracted from surface wave responses retrieved using seismic interferometry by cross-correlation (e.g., [26,27]). Although our final objective is to apply the 3D probabilistic inversion scheme to these dispersion curves, this study is limited to the synthetic dispersion curves. That is, we specifically aim to understand and showcase the potential of the proposed inversion scheme for the RARR's station configuration in combination with the frequencies for which surface wave responses could be retrieved (0.1–0.5 Hz). The dispersion curves extracted from the field data will be inverted in a follow-up paper, and we will benefit from the findings presented here.

As shown in Figure 1, the station's elevation varies between approximately −0.2 km and 0.2 km. Based on a previous study by Martins et al. [3], we can assume a minimum velocity of 2 km/s. For the range of frequencies considered (0.1–0.5 Hz), the minimum wavelength would then be 4 km. This implies that the minimum wavelength is more than ten times larger than the elevation variation. We therefore ignore the station's elevation in the rest of this study.

To investigate the potential of this one-step 3D transdimensional method, we generate synthetic frequency-dependent travel times between the station locations of the RARR. Specifically, we prescribe various 3D block model velocity distributions. The frequency-dependent travel times, derived from phase velocity data, are then used as input for the one-step 3D transdimensional algorithm. We find that the algorithm recovers the velocity structure reasonably well for the RARR configuration within the desired frequency range. However, the method is computationally demanding. Therefore, we modify the algorithm by updating ray paths occasionally, instead of doing this at each step of the Markov chain. A simple test is designed to decide on the update level efficiently. The effect of (the computational) grid size on forward modeling errors and how to choose an appropriate size are also discussed.
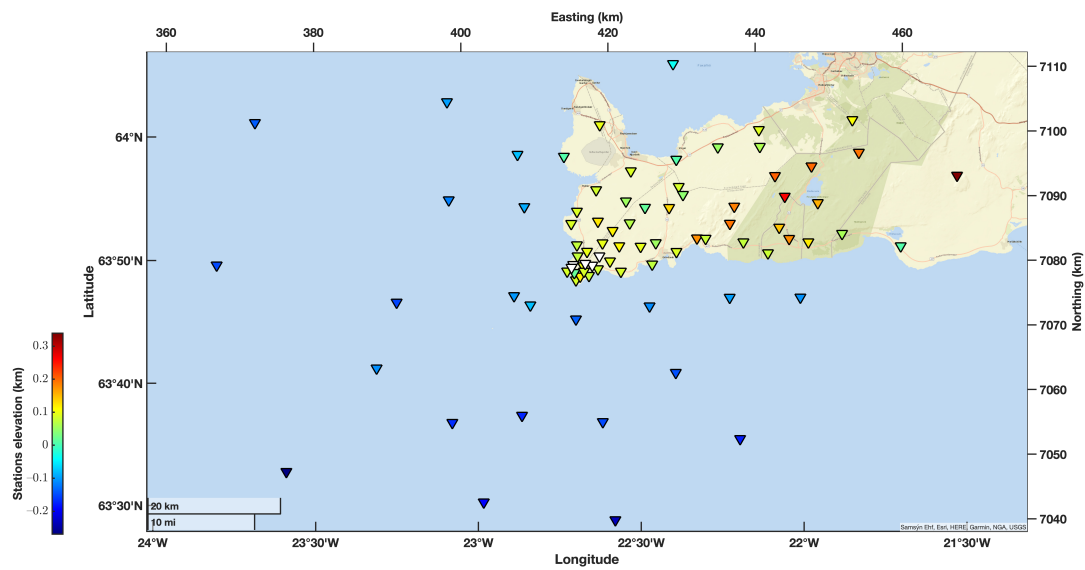

Figure 1. Geographical locations of 83 seismic stations of the RARR. The left and bottom axes indicate spherical coordinates, whereas the right and the top axes display UTM coordinates. Stations locations are depicted as colored triangles. Colors indicate the elevation of the corresponding station. Elevation information was lacking for five stations that are depicted with white triangles.

2. Transdimensional Surface Wave Tomography

Transdimensional tomography is a Bayesian inference method that uses Voronoi cells in conjunction with a rj-McMC algorithm to allow a variable parameterization of the model space. The objective of Bayesian inversion is to recover the posterior probability density of the model parameters given the observed data, $p(\mathbf{m}|\mathbf{d})$. Based on Bayes' rule [28]:

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \tag{1}$$

where the vector $\mathbf{d}$ contains the frequency-dependent travel times. The vector $\mathbf{m}$ contains the model parameters and depends on the parameterization of the region of interest; we discuss this vector in more detail later. The likelihood function $p(\mathbf{d}|\mathbf{m})$ plays a fundamental role in the inference of the model space as it provides the probability of the observed travel times given a specific velocity model. The prior probability distribution, $p(\mathbf{m})$, depends on the (a priori) known information about the model space. Finally, $p(\mathbf{d})$ is the so-called evidence, which here can be considered a constant because it is not a function of any particular model $\mathbf{m}$. Consequently, we have:

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m}). \tag{2}$$

2.1. Model Parameterization

In conventional linearized or gradient-based seismic tomography methods, the region of interest is usually parameterized using fixed cells or a grid of points with predefined shapes and sizes (usually regular rectangular grids). Determining the shape and size of the grid is always a challenge and, ideally, is a function of the data uncertainty (i.e., level of noise available in data) and data resolution (i.e., the ability of the data to resolve different scales or features in the model). Data uncertainty is usually not known and should therefore be estimated before the inversion. In addition to unknown data uncertainty, differences in data resolution due to nonuniform spatial sampling render a uniform grid size undesirable. Consequently, irregular grids or meshes have been proposed by some authors [11,29–32].

Because of the challenge in defining cell geometries, Bodin et al. [30] and Bodin and Sambridge [11] proposed to invert for cell geometries and seismic properties (e.g., velocities) simultaneously. This means that the data directly determine the parameterization. To allow the size and shape of cells to be variable and unknown in the tomographic algorithm, different kinds of methods have been proposed, including Voronoi cells in 2D [11,20,23,30], Voronoi polyhedra (or simply also cells) in 3D [24,25], Delaunay triangulation in 2D [32,33], Johnson–Mehl tessellation in 2D [31], and Gaussian kernels for a 2D model [31]. Despite the abilities of Voronoi cells to accommodate sharp changes and discontinuities, some authors questioned their ability in the case of smooth velocity models or models with a gradient in velocity changes (e.g., [25,32,33]).

In this study, we parameterize the 3D velocity field using polyhedral Voronoi cells to recover the 3D shear wave velocity structure of the subsurface [25]. By assigning a constant shear wave velocity (V_s) to each Voronoi cell, the Voronoi tessellation turns into a shear wave velocity model. Figure 2 shows the region of interest with a coarse parameterization using 33 cells and a finer one with 704 cells; both generated randomly. The number of Voronoi cells, their position, and the shear-wave velocity in each cell are treated as unknowns in the tomographic inversion. Importantly, p-wave velocity (V_p) is considered to be a linear function of V_s, using the relation $V_p = 1.73V_s$ [24]. Density (ρ) is considered to be a function of p-wave velocity as $\rho = 2.35 + 0.036(V_p - 3)^2$, where V_p and V_s are in km/s and ρ is in g/cm^3 [24,34]. It is useful to note that the parameterization and relations in this paragraph imply that the medium is assumed to be isotropic.
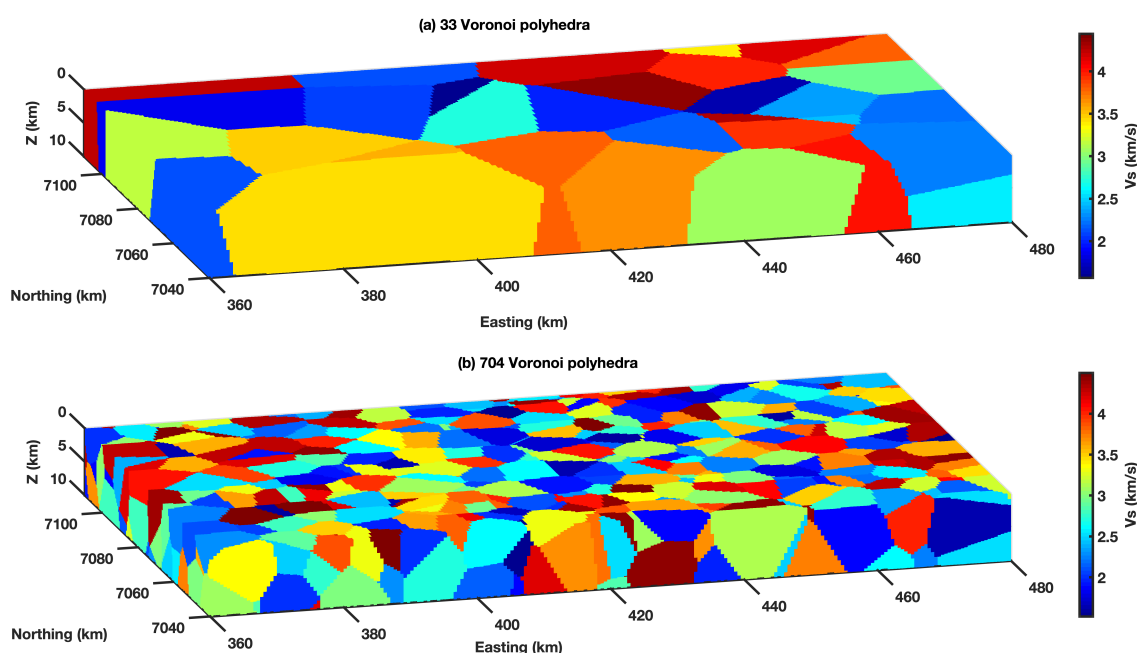

Figure 2. Two random Voronoi parameterizations. The first uses 33 Voronoi cells (**a**), and the second uses 704 Voronoi cells (**b**). Both cell locations and the assigned velocities are generated randomly.

2.2. The Likelihood

The dependence of the posterior on the input data is encoded in the likelihood function $p(\mathbf{d}|\mathbf{m})$. The likelihood function can be considered as a measure of misfit between observed and predicted data. The normalized Gaussian likelihood function reads:

$$p(\mathbf{d}|\mathbf{m}) = \left(\prod_{l=1}^{n_f} \prod_{k=1}^{n_r} \frac{1}{\sqrt{2\pi}\sigma_{kl}} \right) \exp\left(-\sum_{j=1}^{n_f} \sum_{i=1}^{n_r} \frac{(g_{ij}(\mathbf{m}) - d_{ij})^2}{2\sigma_{ij}^2} \right), \tag{3}$$

where n_f and n_r are the number of discrete frequencies and ray paths, respectively. d_{ij} is the travel time along ray path i at discrete frequency j. The vector $\mathbf{m}$ contains the parameters describing the proposed subsurface model. g_{ij} is the computed travel time along ray path i at discrete frequency j in that subsurface model. σ_{ij} is the data uncertainty or noise level for the travel time associated with discrete frequency j and ray path i (and hence σ_{kl} is associated with discrete frequency l and ray path k) . This data uncertainty includes both observational errors and modeling errors [20]. The data uncertainty controls the level of fit and directly affects the complexity of the models in the posterior distribution. That is, the number of Voronoi cells needed to explain the data is highly dependent on the estimated data noise [20]. Consequently, σ_{ij} plays an important role in the transdimensional algorithm. Here, we consider it to be a linear, frequency-dependent function:

$$\sigma_{ij} = a_j * t_{ij} + b_j, \tag{4}$$

where t_{ij} is the computed travel time along ray path i at discrete frequency j. a_j and b_j are noise hyper-parameters for the discrete frequency j. It is worth noting that this linear relation for the noise parameters is a common assumption and is demonstrated in several previous studies (e.g., [20,23,25]). Moreover, it is straightforward to implement it in the context of the transdimensional algorithm.

2.3. Forward Modeling

To evaluate the likelihood of a proposed model given the observed data and to compare this likelihood with the likelihood of the current model in the chain, we need to compute the frequency-dependent travel times in the proposed model ($g_{ij}(\mathbf{m})$ in Equation (3)). This is achieved by employing a two-step approach detailed in [24]. First, at each point on the Earth's surface, the local frequency-dependent dispersion curves are computed using a modal approximation method [10] in the 1D vertical depth profile of surface wave velocity beneath each of the surface points. As such, we obtain frequency-dependent 2D maps of surface wave phase velocity. Figure 3 depicts this first step. Figure 3a depicts a randomly generated 3D surface wave velocity model using Voronoi polyhedra. Figure 3b depicts the corresponding surface wave phase velocity as a function of frequency (vertical axis). We refer to this first step as the depth to frequency step.

In the second step, we use the frequency-dependent phase velocities to compute frequency-dependent travel times by solving the Eikonal equation in two dimensions:

$$|\nabla T(\mathbf{x}, \omega)| = \frac{1}{c(\mathbf{x}, \omega)}, \tag{5}$$

where $T(\mathbf{x}, \omega)$ is the frequency-dependent travel time of the wave-front at surface location $\mathbf{x}$ and angular frequency ω, and $c(\mathbf{x}, \omega)$ is the phase velocity of the model at that surface location and frequency. By solving this equation for each source at each frequency, the travel time field corresponding to first arrivals is obtained at all points of the (2D) model, including the receiver locations. A variety of methods has been proposed for the solution of this equation, including finite-difference [35,36] and the fast marching method (FMM) [37,38]. The fast-marching method is unconditionally stable, meaning that we can use relatively coarse grids to solve Equation (5). A coarser grid, however, does introduce additional travel time errors, which we refer to as modeling errors.

We use the FMM to solve Equation (5) to obtain frequency-dependent travel times at the location of each receiver. As one can see, the relation between the model parameters, $c(\mathbf{x}, \omega)$, and travel time, $T(\mathbf{x}, \omega)$, in Equation (5) is nonlinear, implying the solution of the forward function involves relatively high computational costs (compared to other computational steps in the rj-McMC algorithm). In order to reduce overall computational costs, it is common practice to linearize the problem using a fixed ray path geometry (e.g., straight ray [3]) or update the ray geometry only occasionally to account for the non-linearity of the problem [11,39]. Ray paths are perpendicular to the travel time field. Hence, after computing the travel time field with the fast-marching method, the ray path geometry can be calculated by starting at each receiver location in the travel time field and following the travel time gradient (∇T) back to the source location [38]. Once the ray paths are determined, integration of the slowness along each ray path is straightforward and travel times can be computed at relatively little computational expense.
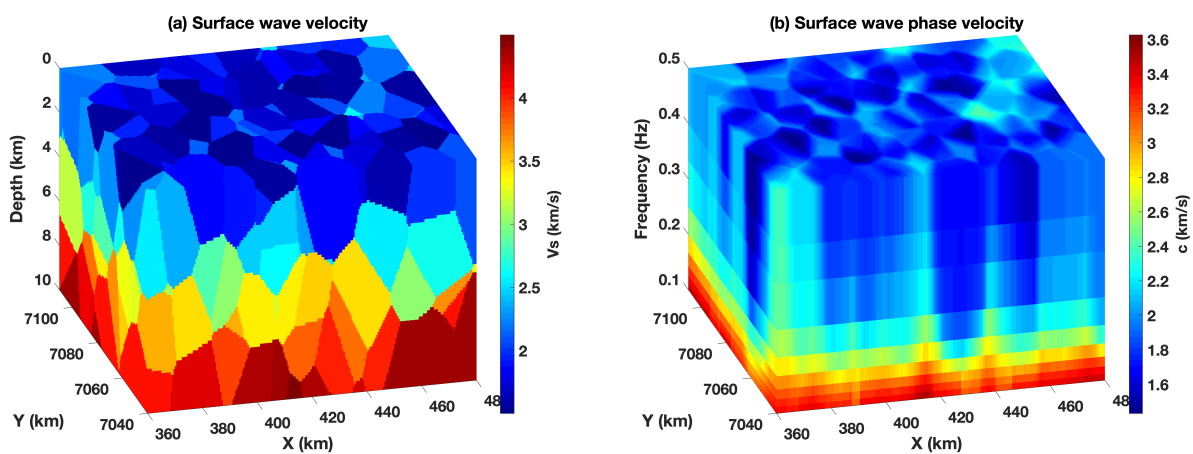

Figure 3. Depth to frequency conversion of surface wave velocities using the modal approximation method. (**a**) Surface wave velocity model generated randomly using Voronoi cells. (**b**) Computed phase velocities corresponding to each depth profile using the modal approximation method.

2.4. The Prior

Since all inferences on the posterior are relative to the prior distribution, priors have great importance in Bayesian inversion schemes. The final result may suffer heavily from an incorrect prior. In order to prevent the introduction of prior-related biases into the solution, we choose uniform prior distributions with wide bounds for all model parameters. Assuming independent model parameters, the prior can be written as:

$$p(\mathbf{m}) = p(n)p(\mathbf{c}|n)p(\mathbf{v}|n)p(h), \tag{6}$$

where n is the number of Voronoi cells, $p(n)$ is the prior on the number of cells, $p(\mathbf{c}|n)$ is the prior on cell nuclei location, $p(\mathbf{v}|n)$ is the prior on cell velocity, and $p(h)$ is the prior on noise hyper-parameters or data uncertainty (see [11,20] for details). The modal approximation method for computing phase velocities fails to compute the right surface wave mode when the model contains a layer whose velocity is lower than the velocity of the top layer [23,25]. Hence, this issue has to be taken into account in the prior by discarding models for which the top layer does not have the lowest velocity value.

2.5. Reversible Jump McMC

The reversible jump Markov chain draws samples from the posterior distribution employing an iterative stochastic process (such as the Metropolis–Hasting algorithm). The reverse jumps allow for a variable number of Voronoi cells, hence a variable number of parameters. Jumping between different model dimensions allow the rj-McMC algorithm to perform a global search [40]. The algorithm proposes different types of model perturbations.

Specifically, we use four different perturbation types to sample the posterior distribution efficiently [11]: a velocity update, Voronoi cell move, cell death and cell birth. In addition, we perturb the amplitude of the noise to infer the posterior probability of the errors on the measured surface wave travel times (this is introduced by Bodin et al. [20]). These perturbation types allow the model to dynamically adapt itself to data density, underlying velocity structure, and travel time noise.

The flowchart for the algorithm of the transdimensional McMC inversion is depicted in Figure 4. The process starts with a random initial model **m**. Then, the algorithm draws the next sample of the chain by proposing a new model, **m**′, based on a known proposal probability function, $q(\mathbf{m}'|\mathbf{m})$, which only depends on the previous state of the model **m**. To propose a new model in a velocity update step and a Voronoi cell move, a Voronoi cell is selected randomly, then the velocity or the location of that cell is perturbed using a Gaussian proposal distribution. In a noise perturbation step, one of the two noise parameters at one frequency (i.e., a_j or b_j) is chosen randomly, then it is perturbed using a Gaussian proposal distribution. The proposed sample, then, will be accepted or rejected based on an acceptance probability for the proposed model, **m**′ [30]. A new sample is drawn at each step of the Markov chain by perturbing the 3D velocity structure or the noise hyperparameters (using one of the five perturbation types). The surface wave dispersion data (i.e., the frequency-dependent travel times) are then calculated to evaluate the following acceptance probability [11]:

$$\alpha(\mathbf{m}'|\mathbf{m}) = min[1, \frac{p(\mathbf{m}')}{p(\mathbf{m})}\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})} \times |\mathbf{J}|], \tag{7}$$

where $\alpha(\mathbf{m}'|\mathbf{m})$ is the probability of accepting the proposed model **m**′ given the current model **m**, $\frac{p(\mathbf{m}')}{p(\mathbf{m})}$ the prior probability ratio of two models, **d** the observed data (here these are the frequency-dependent travel times), $\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}$ the likelihood ratio of the two models, $\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}$ the proposal ratio, and **J** is the Jacobian matrix, which accounts for (potential) differences in dimensionality between **m** and **m**′ (resulting from a different number of Voronoi cells in **m** and **m**′). For the birth and death steps used here, the determinant of the Jacobian matrix is unity [11,24].

When sufficient samples are drawn from the posterior, we can compute the mean, standard deviation, and other statistical measures of the posterior. The samples from a certain initial period of the chain are discarded. This initial sampling period is usually referred to as the burn-in period, which is the period that the algorithm needs to remove the effects of the initial model and reach sufficient mixing of the posterior samples. Since each sample is drawn based on a small perturbation of the previous model, adjacent samples are correlated or similar. To ensure that drawn samples are uncorrelated, we retain a sample every so many iterations; for example, every 200th model is retained. This process is usually referred to as 'thinning' of the chain.
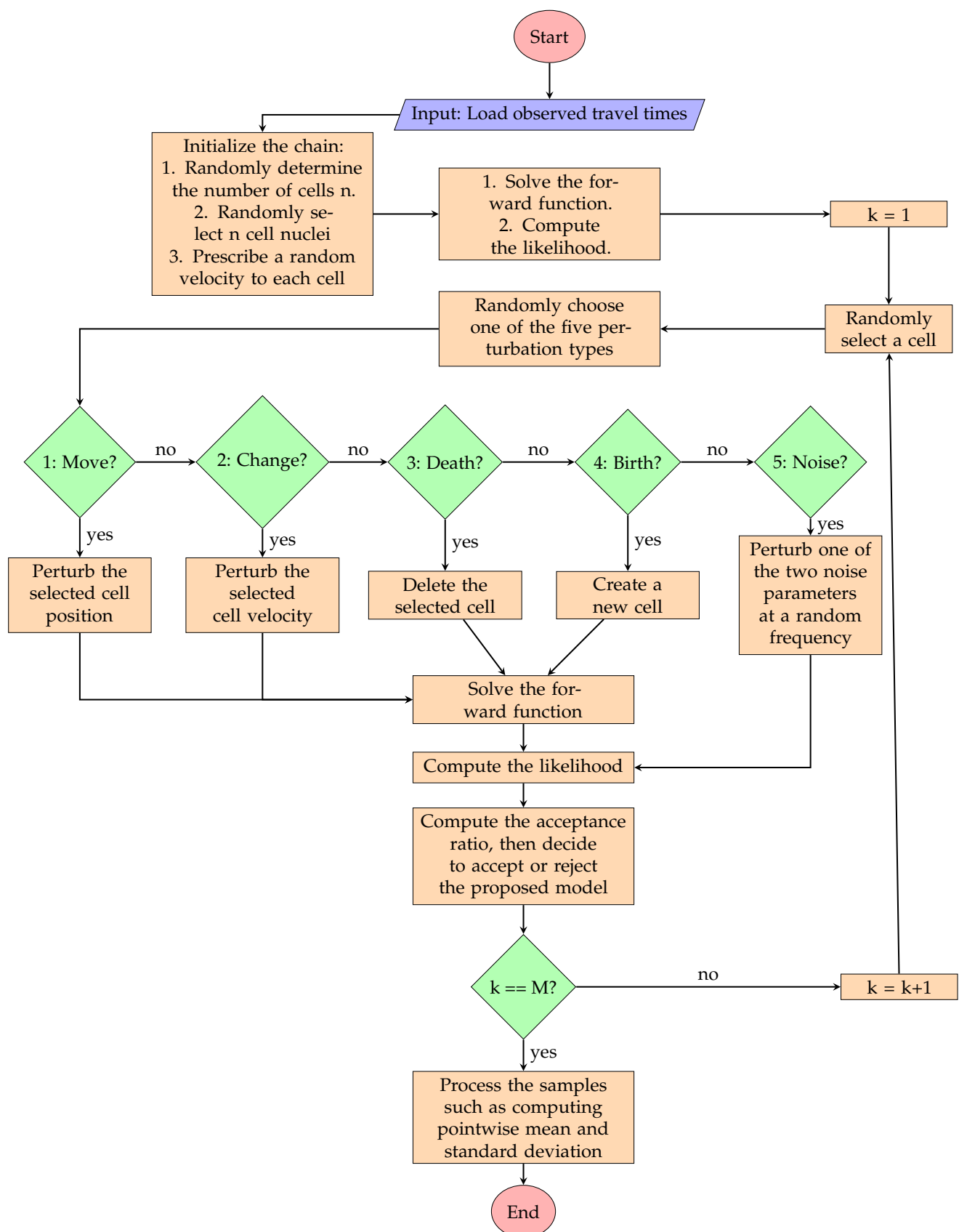
Figure 4. Flowchart of the transdimensional Markov chain Monte Carlo algorithm used in this work. Here, a total of M samples is drawn from the posterior.

3. Experiment Setup and Computational Performance Tests

To investigate the potential of the transdimensional algorithm for the interferometric surface wave responses of the RARR, we design three distinct 3D (block model) tests. With these tests, we evaluate the ability of the algorithm to recover subsurface models with different resolutions. In this section we introduce the three block models, but additionally conduct a number of experiments to determine acceptable values for the different numerical parameters (e.g., grid size). The results obtained by applying the transdimensional algorithm to travel times in the three block models are presented in the next section.

3.1. Ray Path Update and Computational Cost

The described transdimensional Markov chain naturally adapts model resolution to data resolution, implying it is self-regularized and self-smooth and hence does not require the regularization and/or smoothing needed in deterministic inversions (e.g., [3]). In addition, model uncertainty is naturally captured in the posterior. Despite these benefits, the high computational cost is still a drawback of the method. To sample the posterior sufficiently well, we need to run multiple chains for at least a million iterations.

As we mentioned earlier, computing frequency-dependent travel times using the fast-marching method contributes most to the computational cost. To make the algorithm computationally less demanding, Bodin et al. [20] used fixed ray paths to compute travel times and updated the ray paths only occasionally (three times for three million samples). In this way, they linearized the algorithm partially. Galetti et al. [23] argued that this might introduce artifacts and bias into the solution. Therefore, Galetti et al. [23] and Zhang et al. [25] updated the ray paths at each step of the Markov chain.

At the same time, we know that each new sample involves only a small perturbation of the model. Consequently, ray paths do not change too much in one thinning "period" (in our case, every 200 iterations). Hence, we update ray paths every 200 iterations based on the average velocity of the previous 200 samples. In this way, we reduce the computational cost significantly while still retaining most of the non-linearity. Figure 5 shows the speed-up associated with different ray path update steps (with respect to updating ray paths at every chain step). The dimension of our assumed model is 120 km by 70 km. We used a constant grid spacing of 0.5 km in the depth direction (used for the modal approximation method, [10]). Then, we compared the time needed to take three million Markov chain steps using grid spacing of 0.5 km, 1 km, and 2 km in the two horizontal directions. Figure 5 demonstrates that updating ray paths every 200 iterations significantly decreases computational costs. Updating the ray paths less frequently (i.e., using higher update step), however, does not reduce computational cost much more. Hence, we consider updating ray paths every 200 iterations optimal in terms of both computational cost and honoring the non-linearity of the problem.

3.2. Block Models

The three considered synthetic models are depicted in Figure 6. Figure 6a depicts a block model with blocks of size 20 km × 20 km × 10 km in the northing, easting, and depth directions, respectively. We will henceforth refer to this subsurface model as the coarse block model. Figure 6b depicts a block model with blocks of size 10 km × 10 km × 5 km in the northing, easting, and depth directions, respectively. We will henceforth refer to this subsurface model as the intermediate block model. Figure 6c depicts a block model with blocks of size 5 km × 5 km × 2.5 km in the northing, easting, and depth directions, respectively. We will henceforth refer to this subsurface model as the fine block model. In addition to these resolution tests, we study the impact of additive noise on the ability of the algorithm to adapt the model to (and to estimate) the noise level.
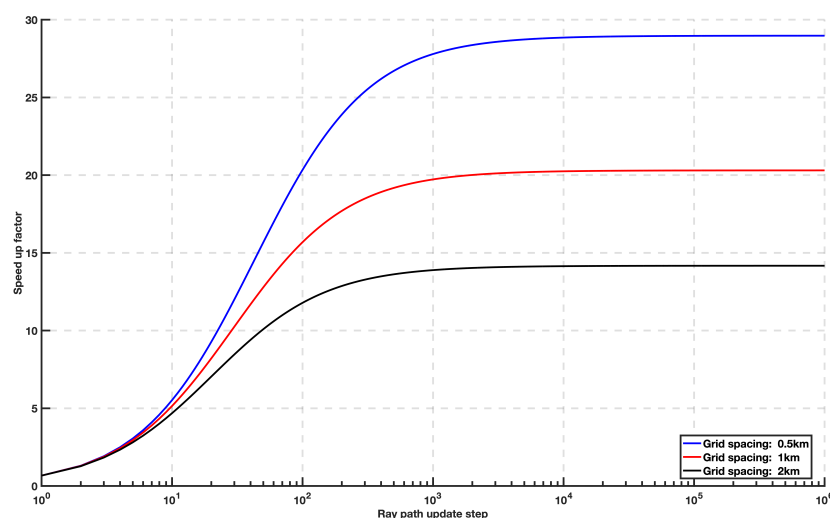
Figure 5. Decrease in computation time as a function of the frequency at which ray paths are updated. The speed up is depicted for three different FMM resolutions.
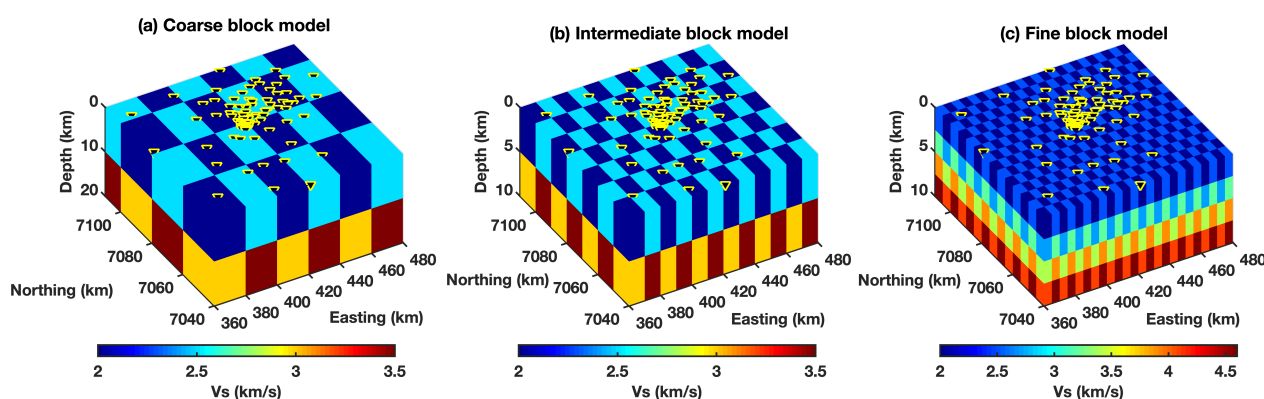

Figure 6. Three different synthetic block models to test the 3D transdimensional algorithm. (**a**) Coarse block model, (**b**) Intermediate block model, (**c**) Fine block model. Inverted yellow triangles indicate the locations of the stations of the RARR. (See also Figure 1).

3.3. Sensitivity Kernels

Each of the three block models in Figure 6 has only two distinct velocity profiles, of which one is shown in the left panel of Figure 7a,d,g. The other velocity profiles (not shown) have coinciding step sizes and depths, but their velocities are simply 500 m/s higher for each layer. The middle panel of this figure (b, e, h) depicts the associated sensitivity kernels of the Rayleigh waves computed using the modal approximation method [10]. For different frequencies, these kernels give the Rayleigh waves' sensitivity to the shear wave velocity as a function of depth.

Figure 7b shows that Rayleigh waves in the frequency range of 0.1–0.5 Hz are predominantly sensitive to the shear wave velocities of the shallowest 10 km. Hence, we do not expect the algorithm to recover the shear wave velocity below that depth very well in the coarse block model. Figure 7e,h show that Rayleigh waves in the frequency range 0.1–0.5 Hz are sensitive to the shear wave velocities of all blocks of the intermediate and fine block models down to 10 km, and should, in principle, be able to recover these blocks. We, therefore, expect to recover the velocity structure in areas traversed by sufficient ray paths. A decrease in resolution with increasing depth is expected. The right panel of Figure 7 is the phase velocity for the corresponding depth profile of shear wave velocity at ten different discrete frequencies in the frequency range of 0.1–0.5 Hz using the modal approximation method [10].
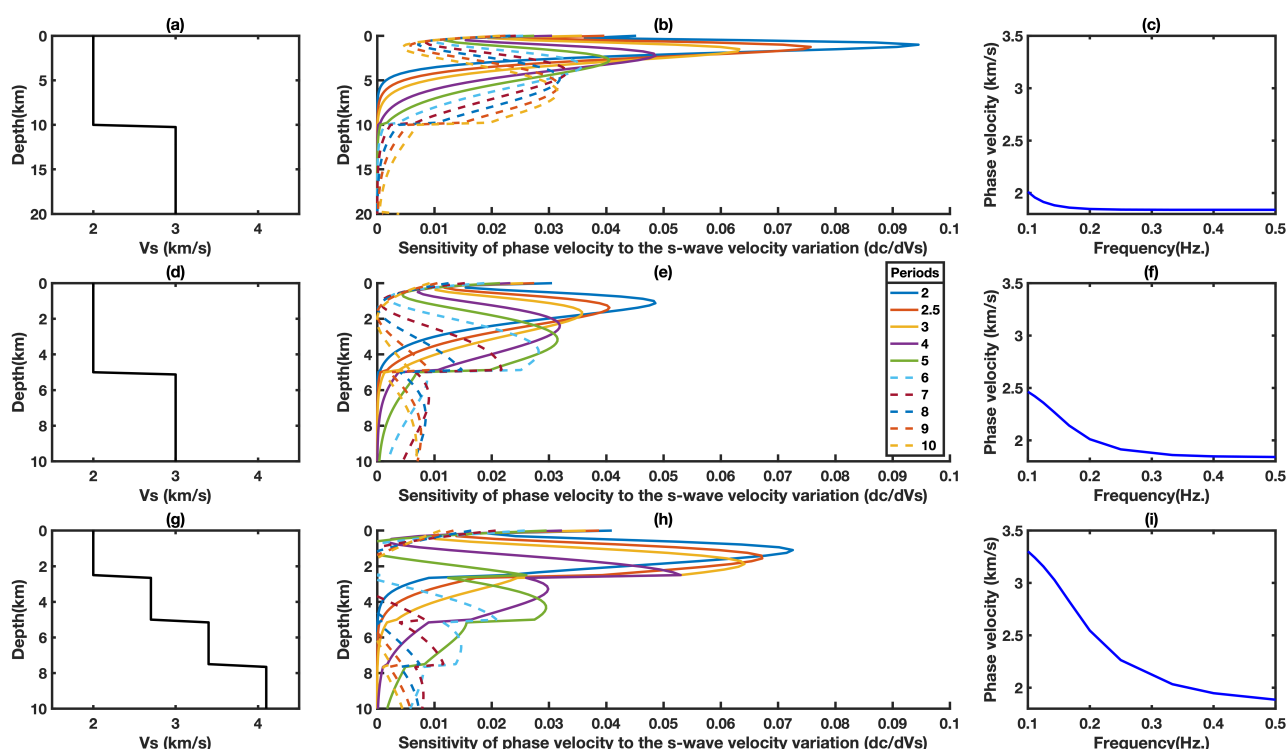
Figure 7. Surface wave sensitivity analysis of one of the depth profiles of each block model. The left column displays shear wave velocity as a function of depth for (**a**) the coarse block model, (**d**) the intermediate block model, and (**g**) the fine block model. The middle column (**b**,**e**,**h**) displays the corresponding sensitivity of the Rayleigh waves for different periods (1/f) as a function of depth. On the right (**c**,**f**,**i**), the corresponding phase velocity dispersion is shown as a function of frequency.

3.4. Additive Noise and Modeling Errors

As we mentioned before, noise parameters are assumed to be unknown and are estimated by the transdimensional algorithm. To evaluate the ability of the algorithm to recover the noise level, we designed a noise-free experiment and an experiment with additive Gaussian random noise. A noise-free experiment means that we did not add any additive noise to the modeled synthetic data. Modeling errors (effectively resulting in noise) are, however, still present. These are relatively small, and we therefore nevertheless refer to this experiment as a noise-free experiment. In the second experiment, after computing frequency-dependent travel times through the true synthetic block models of Figure 6, we have added random Gaussian noise with a variance based on Equation (4) with $a_j = 0.04$ and $b_j = 0.1$.

To determine the number of computational grid points while solving the forward function of the transdimensional algorithm and a reasonable level of additive noise, we designed two tests. First, we considered the fine block model of Figure 6c. Then, we computed the travel times between the RARR stations at a frequency of 0.1 Hz, which implies, on average, a wavelength of 20 km. Travel times were computed for four different grid sizes ($61 \times 36 \times 21$, $121 \times 71 \times 41$, $241 \times 141 \times 81$, and $481 \times 281 \times 161$). The corresponding grid separations are 2 km $\times$ 2 km $\times$ 0.5 km, 1 km $\times$ 1 km $\times$ 0.25 km, 0.5 km $\times$ 0.5 km $\times$ 0.125 km, and 0.25 km $\times$ 0.25 km $\times$ 0.0625 km, respectively. Assuming the obtained travel times with the finer grid points ($481 \times 281 \times 161$) as a reference, we calculated the other three grids' relative errors. Errors are then sorted based on interstation distances and are presented in Figure 8a. As such, we get an impression of the modeling errors associated with different grid sizes. In the second test, we used the randomly generated model of Figure 3a and followed the same procedure as the first test. Relative errors are presented in Figure 8b. The relative error due to a random Gaussian noise based on Equation (4) with $a_j = 0.04$ and $b_j = 0.1$ is also presented in this figure by the green solid line. The relative error is around 5% at longer distances and reaches more

than 30% at shorter distances. We believe this represents the available noise in a real case study. Additionally, it seems that the modeling error is compatible with the assumed linear relation for the noise variance, Equation (4).

We deduce from Figure 8 that the modeling errors at short distances are much higher than at long distances. In particular, we consider the modeling errors using a grid of $61 \times 36 \times 21$ unacceptably high at short distances. We consider modeling errors using a grid of $121 \times 71 \times 41$ and $241 \times 141 \times 81$ points acceptable. They are low enough to be used in our modeling and inversion. That is why we select the grid of $241 \times 141 \times 81$ points to compute the synthetic travel times (i.e., observed travel times). Solving the forward problem using a grid of $241 \times 141 \times 81$ points, however, renders the transdimensional Markov chain computationally very expensive (given the size of our computational cluster). Consequently, we use a grid of $121 \times 71 \times 41$ points while running the Markov chain.
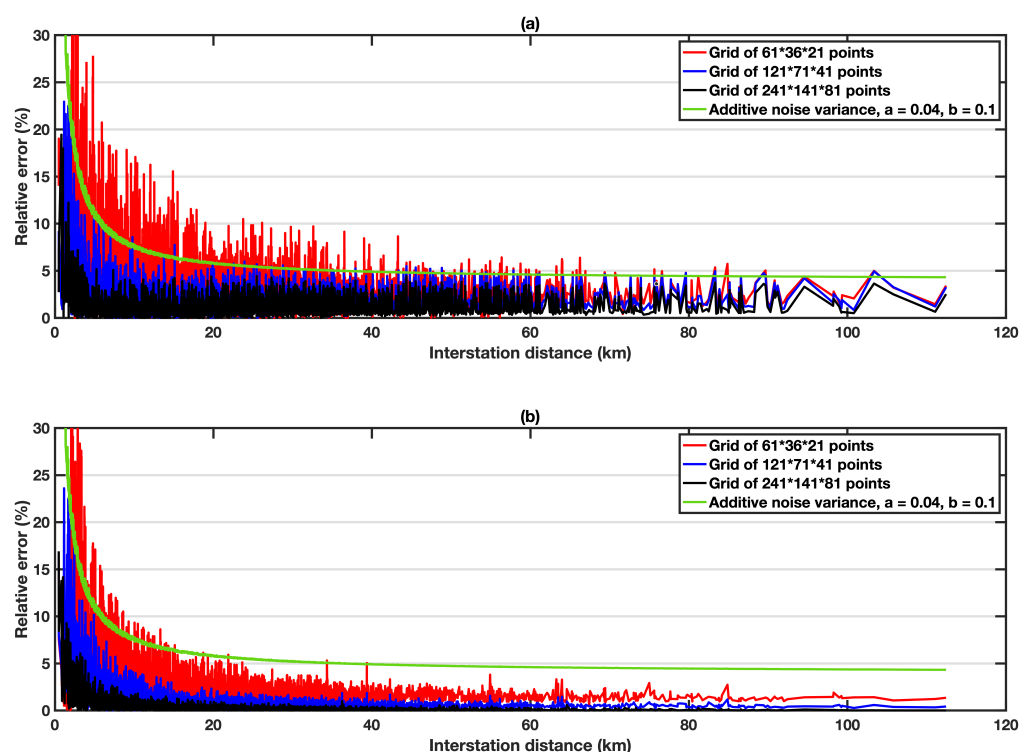

Figure 8. Two tests to analyze modeling errors. (**a**) Relative modeling errors while computing travel times through the fine block model of Figure 6c using three different grid resolutions. (**b**) Modeling errors for computing travel times through the randomly generated model of Figure 3a using three different grid resolutions. Relative error due to the additive random Gaussian noise based on Equation (4) with $a_j = 0.04$ and $b_j = 0.1$ is also included in green line on both (**a**,**b**).

It is useful to note that, in application to field data, we will discard travel times associated with interstation distances shorter than 1.5 wavelength. In practice, given (i) a maximum frequency of 0.5 Hz for which reliable interferometric surface wave responses can be retrieved and (ii) an average phase velocity of about 2 km/s at that frequency, no travel times associated with interstation distances shorter than approximately 6 km will serve as input.

3.5. Modeling and Inversion Parameters

The modeling and inversion parameters used in the synthetic experiments are presented in Table 1. We used the same inversion parameters for the noise-free experiments and the experiments with additive noise. We started the transdimensional sampling of the posterior with a random initial model in all experiments, implying that the number of Voronoi cells, their position and their velocities were chosen randomly. However, to take

into account the necessity of having the lowest velocity at the surface, the initial model was generated, increasing in depth, yet randomly. As can be seen in Table 1, we generated synthetic frequency-dependent travel times ("observed" data) on a relatively fine grid for all three block models. While computing the frequency-dependent travel times (forward modeling) during McMC sampling, however, we used a coarse grid. Finally, we used a finer grid for calculating the post-burn-in pointwise average and standard deviation of the sampled models.

For drawing new velocity values as well as new nuclei locations, we used Gaussian proposal probability distributions. The width of the proposal distributions control the chance of a proposed model being accepted and, consequently, the transdimensional algorithm's convergence rate. On the one hand, narrow proposal distributions lead to higher acceptance ratios, more correlated samples and a lower convergence rate. On the other hand, wide proposal distributions lead to lower acceptance ratios and consequently lower convergence rates. Previous studies suggest that an efficient proposal distribution width results in an acceptance ratio of 25–50% [30,41]. These values were determined, a priori, from the data. We determined the efficient proposal distribution width through a pilot run of the algorithm and looked at samples' acceptance ratios of different perturbation types. This means that one does (of course) not need to know the true velocity model to define the proposal widths. The obtained proposal widths for the velocity update, moving a cell and noise perturbation are presented in Table 1. The valid range of parameters presented in the table defines the bounds of the uniform prior on each parameter.

To reduce computational cost while still preserving the non-linearity of the problem, we updated ray paths at every 200 iterations as we discussed earlier. Thinning was also achieved by retaining every 200th model.

Table 1. Modeling and inversion parameters. Parameters are the same for the noise-free experiment and the experiment with synthetic additive noise.

Modeling and Inversion Parameters	Coarse Block Model	Intermediate Block Model	Fine Block Model
Model dimension (km)	$120 \times 70 \times 20$	$120 \times 70 \times 10$	$120 \times 70 \times 10$
Number of grid points in generating synthetic data	$241 \times 141 \times 81$	$241 \times 141 \times 81$	$241 \times 141 \times 81$
Number of grid points in McMC sampling	$121 \times 71 \times 41$	$121 \times 71 \times 41$	$121 \times 71 \times 41$
Number of grid points for calculating post-burn-in pointwise average	$241 \times 141 \times 41$	$241 \times 141 \times 41$	$241 \times 141 \times 41$
Valid range of shear wave velocity (km/s)	1.5–6	1.5–6	1.5–6
Valid range of noise hyper parameter a	10^{-5}–1	10^{-5}–1	10^{-5}–1
Valid range of noise hyper parameter b	0-2	0–2	0–2
Proposal width for a move step. Md is the model dimension.	$0.07 * Md$	$0.06 * Md$	$0.05 * Md$
Velocity proposal width (km/s)	0.4	0.4	0.3
Proposal width for a	10^{-3}	10^{-3}	10^{-3}
Proposal width for b	10^{-2}	10^{-2}	10^{-2}
Thinning level	200	200	200
Ray path update step	200	200	200

4. Results and Discussion

In this section, we present the results obtained by applying the transdimensional Markov sampler to the synthetic travel times through the block models introduced in Figure 6. That is, the travel times between the RARR stations (computed using a grid of $241 \times 141 \times 81$ points) serve as input, that is, as the observed data.

4.1. Coarse Block Model

Figure 9 shows the results of the one-step 3D transdimensional method for the coarse block model of Figure 6a in the noise-free experiment. A horizontal slice at the surface of the model and two vertical cross-sections of the true velocity model are depicted in Figure 9a–c. Figures 9d–f depict the (pointwise) average of all post-burn-in models (after thinning).

The bottom row of this figure (Figure 9g–i) provides an estimate of the model uncertainty, that is, the post-burn-in (pointwise) standard deviation. Retained models are obtained by thinning ten separate transdimensional Markov chains, each with 1×10^6 iterations of which 4×10^5 are considered burn-in steps. Since thinning involves retaining one in every 200 iterations in our case, 30,000 samples are used for the computation of the pointwise average and standard deviation.

Figure 9 shows that the algorithm recovered the velocity distribution at the surface very well. In line with the low sensitivity of the sensitivity kernels to greater depth (Figure 7b), the velocity structure below 10 km is not recovered very well, and reveals a higher uncertainty than at shallower depth. Figure 10 is similar to Figure 9, but presents the results in the experiment with additive noise. Comparing Figure 9d–f and Figure 10d–f shows that the additive Gaussian noise results in a smoother (pointwise averaged) model. The block interfaces are clearly visible in the uncertainty maps of Figure 10g–i. Though the velocity models are not recovered very well in deeper parts of Figure 10d–f, the deeper subsurface structure is reflected in the uncertainty maps of Figure 10g–i. This exemplifies why the uncertainties aid the interpretation of the results.

Comparing the average maps and the uncertainty maps of Figures 9 and 10, we can conclude that the rj-McMC converges faster in case of higher noise levels. The results presented in Figure 9 suggest that more iterations are needed to sample the posterior sufficiently well (this will be discussed in more detail in Section 4.4). We explain this by the fact that stronger noise (σ in Equation (3), which we also estimate by means of the hyper-parameters) results in a flatter likelihood. This makes it easier to explore the posterior compared to a narrower likelihood associated with less noise (i.e., a smaller σ). Mathematically speaking, a higher σ simply increases the acceptance ratio of the Markov chain and hence results in more unique samples. Consequently, the pointwise average of all sampled models (after burn-in) is likely to be smoother and closer to the true posterior. In the case of a noise-free experiment, we could use a fixed (higher) noise estimate to achieve a smoother result. However, we would lose resolution unnecessarily by imposing such a higher value. Of course, were one to have ample computational power, one could simply run the Markov chains for the noise-free experiment for a longer time to better approach the posterior.

4.2. Intermediate Block Model

Figure 11 shows the results of the one-step 3D transdimensional method for the intermediate block model of Figure 6b in the noise-free experiment. Figure 11a–c depicts the horizontal slice at the surface of the model and two vertical cross-sections of the true velocity model. Figure 11d–f depicts the post-burn-in pointwise average of the retained velocity models. Figure 11g–i depicts the uncertainty, which is the post-burn-in standard deviation of the retained samples.

Figure 11 reveals that the algorithm recovered the true velocity model well in both horizontal and vertical directions in densely sampled regions. Regions that are not sampled that well by the data (i.e., regions traversed by few ray paths) are recovered less well; for example, the eastern part of the velocity model. These regions also exhibit larger uncertainties estimated from the pointwise averages. The fact that the algorithm results in smoother pointwise averages in regions sampled by fewer rays, is the reason that the transdimensional algorithm is often referred to as a self-smooth, self-regularized algorithm.

Figure 12 is similar to Figure 11 but presents the results in the case of additive noise. Similar to our findings for the coarse block model, the additive Gaussian noise results in a smoother (pointwise averaged) model. Block interfaces and areas traversed by no (or few) ray paths show higher uncertainties, as one would expect.
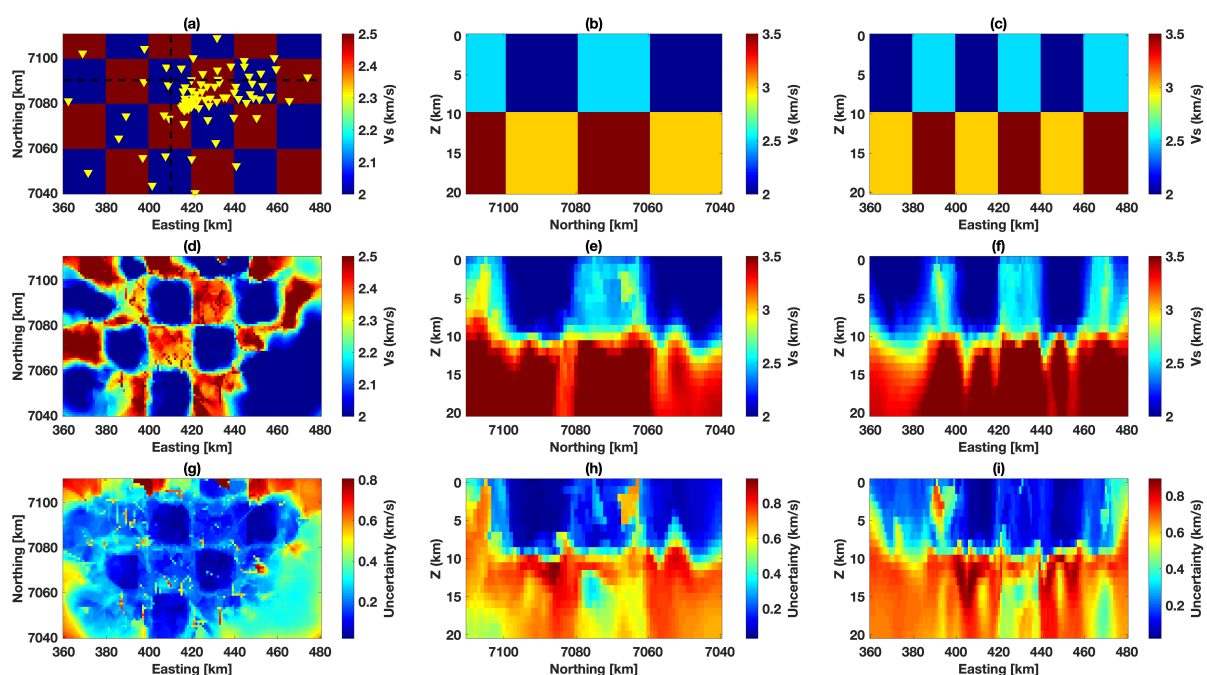
Figure 9. Transdimensional tomographic models estimated from noise-free synthetic travel times through the coarse block model in Figure 6a. The true block model: (**a**) horizontal slice at the surface of the model including station locations, (**b**) vertical cross-section at easting of 410 km, and (**c**) vertical cross-section at northing of 7090 km. The vertical cross-sections are indicated with black dashed lines in (**a**). (**d**–**f**) Pointwise averaged velocities calculated from post-burn-in retained samples. (**g**–**i**) Standard deviation (model uncertainty) calculated from post-burn-in retained samples.
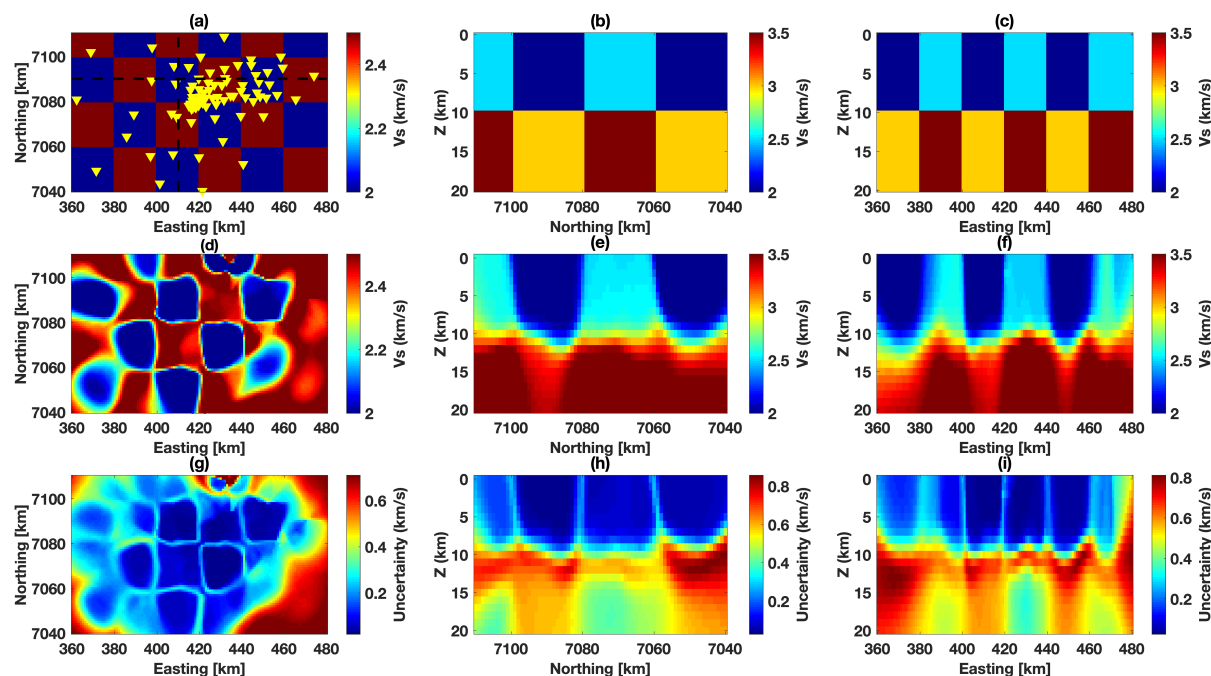

Figure 10. Transdimensional tomographic models estimated from synthetic travel times containing Gaussian random noise for the coarse block model in Figure 6a. The true block model: (**a**) horizontal slice at the surface of the model including station locations, (**b**) vertical cross-section at easting of 410 km, and (**c**) vertical cross-section at northing of 7090 km. The vertical cross-sections are indicated with black dashed lines in (**a**). (**d**–**f**) Pointwise averaged velocities calculated from post-burn-in retained samples. (**g**–**i**) Standard deviation (model uncertainty) calculated from post-burn-in retained samples.
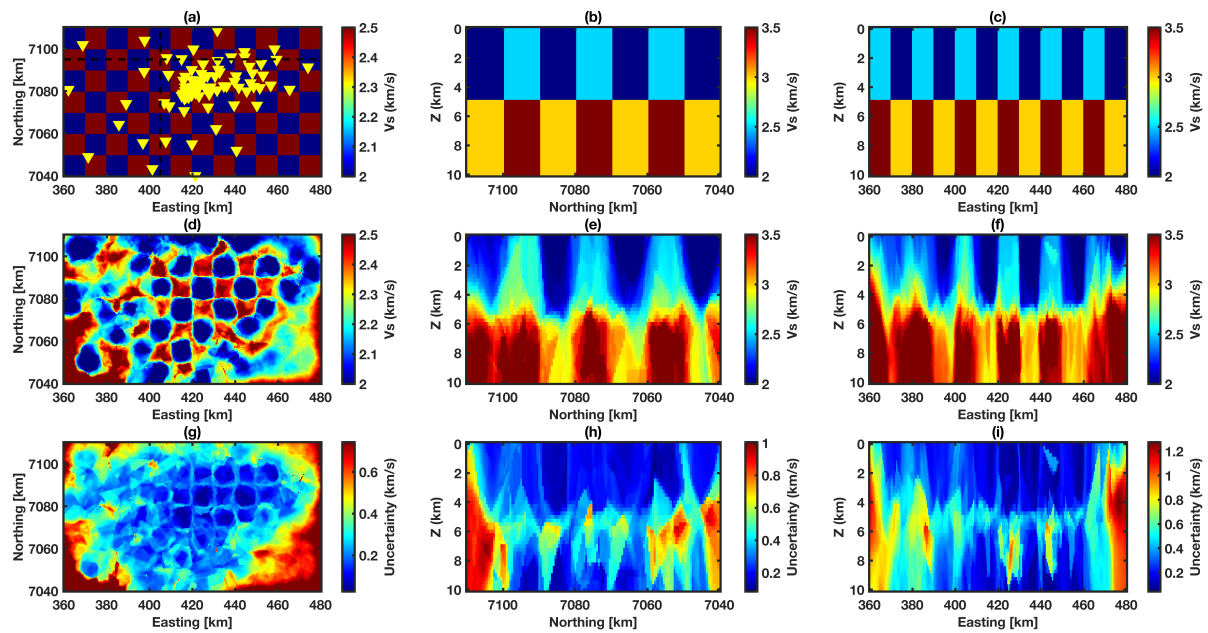
Figure 11. Transdimensional tomographic models for the noise-free synthetic travel times through the intermediate block model in Figure 6b. The true block model: (**a**) horizontal slice at the surface of the model including station locations, (**b**) vertical cross-section at easting of 405 km, and (**c**) vertical cross-section at northing of 7095 km. The vertical cross-sections are indicated with black dashed lines in (**a**). (**d**–**f**) Pointwise averaged velocities calculated from post-burn-in retained samples. (**g**–**i**) Standard deviation (model uncertainty) calculated from post-burn-in retained samples.
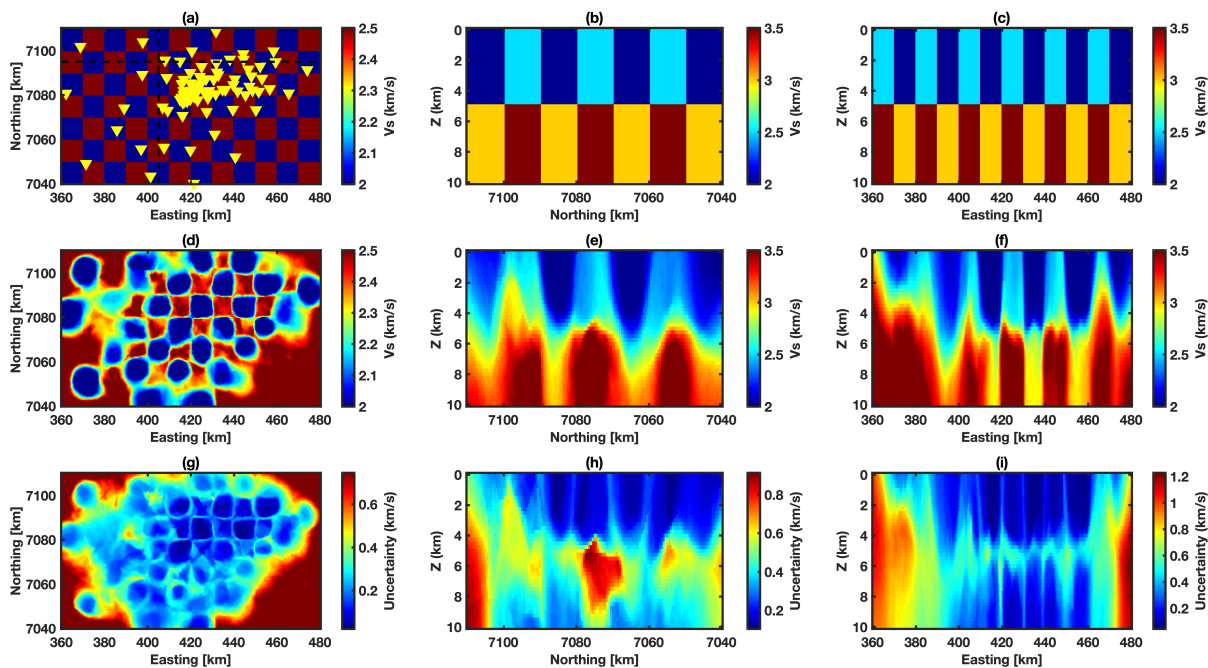

Figure 12. Transdimensional tomographic models for the synthetic travel times containing Gaussian random noise for the intermediate block model in Figure 6b. The true block model: (**a**) horizontal slice at the surface of the model including station locations, (**b**) vertical cross-section at easting of 405 km, and (**c**) vertical cross-section at northing of 7095 km. The vertical cross-sections are indicated with black dashed lines in the horizontal slice. (**d**–**f**) Pointwise averaged velocities calculated from post-burn-in retained samples. (**g**–**i**) Standard deviation (model uncertainty) calculated from post-burn-in retained samples.

4.3. Fine Block Model

Figure 13 shows the results of the one-step 3D transdimensional method for the fine block model of Figure 6c with no additive noise on the travel times. Because of the greater complexity of the fine block model, we used 20 chains (instead of 10) to sample the posterior. Each chain generated 2×10^6 samples, of which 1×10^6 samples are considered part of the burn-in period. Figure 13a–c depicts the true velocity model by means of a horizontal slice at the surface of the model and two vertical cross-sections. Figure 13d–f represent the recovered velocity models, which are the post-burn-in pointwise averages of the retained samples. Retaining again every 200 iterations, the pointwise averages and standard deviation are computed from a total of 100,000 samples. As one can see, the algorithm recovered the blocks fairly well in the densely sampled area of the model. Figure 13g–i estimates the model uncertainty through the computation of the post-burn-in standard deviation of the retained samples.

Figure 14 is the same as Figure 13 but presents the results in the case of noisy travel times. Similar to our observations for the other two block models, the additive Gaussian noise results in a smoother (pointwise averaged) model. Both the noise free and the noisy travel times result in pointwise averaged velocity models exhibiting large uncertainty compared to the velocity models recovered for the intermediate and coarse block models.
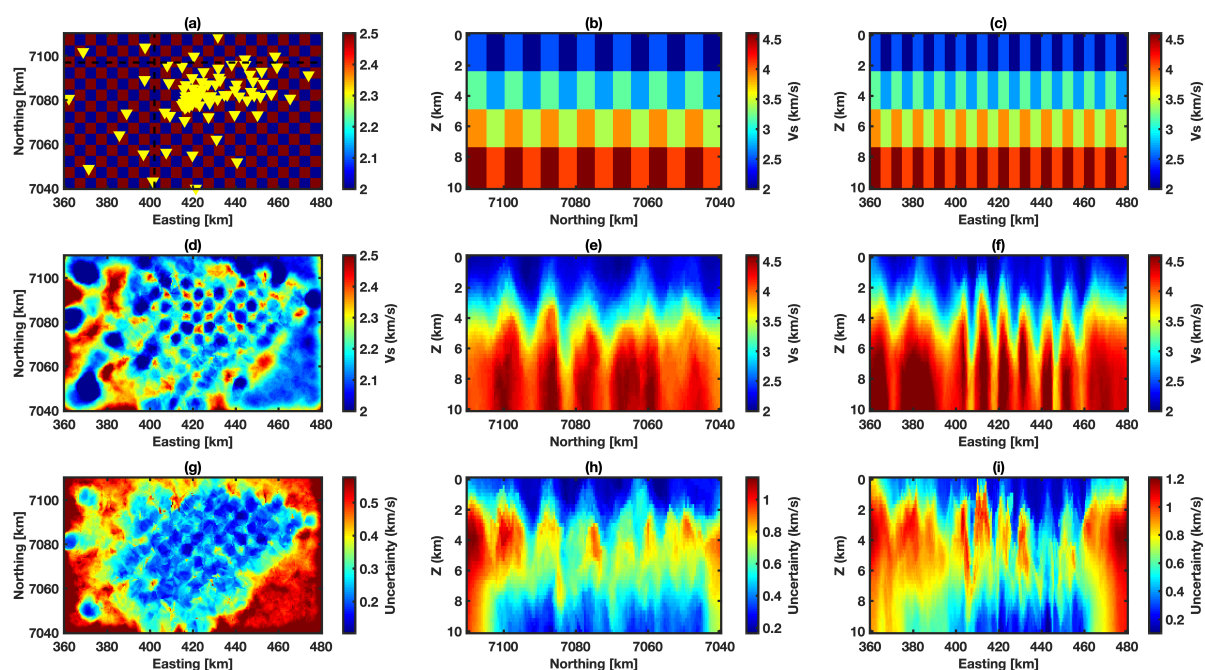

Figure 13. Transdimensional tomography results of the noise-free synthetic travel times for the fine block model in Figure 6c. The true block model: (**a**) horizontal slice at the surface of the model including station locations, (**b**) vertical cross-section at easting of 402 km, and (**c**) vertical cross-section at northing of 7097 km. The vertical cross-sections are indicated with black dashed lines in the horizontal slice of (**a**). (**d**–**f**) Pointwise averaged velocities calculated from post-burn-in retained samples. (**g**–**i**) Standard deviation (model uncertainty) calculated from post-burn-in retained samples.
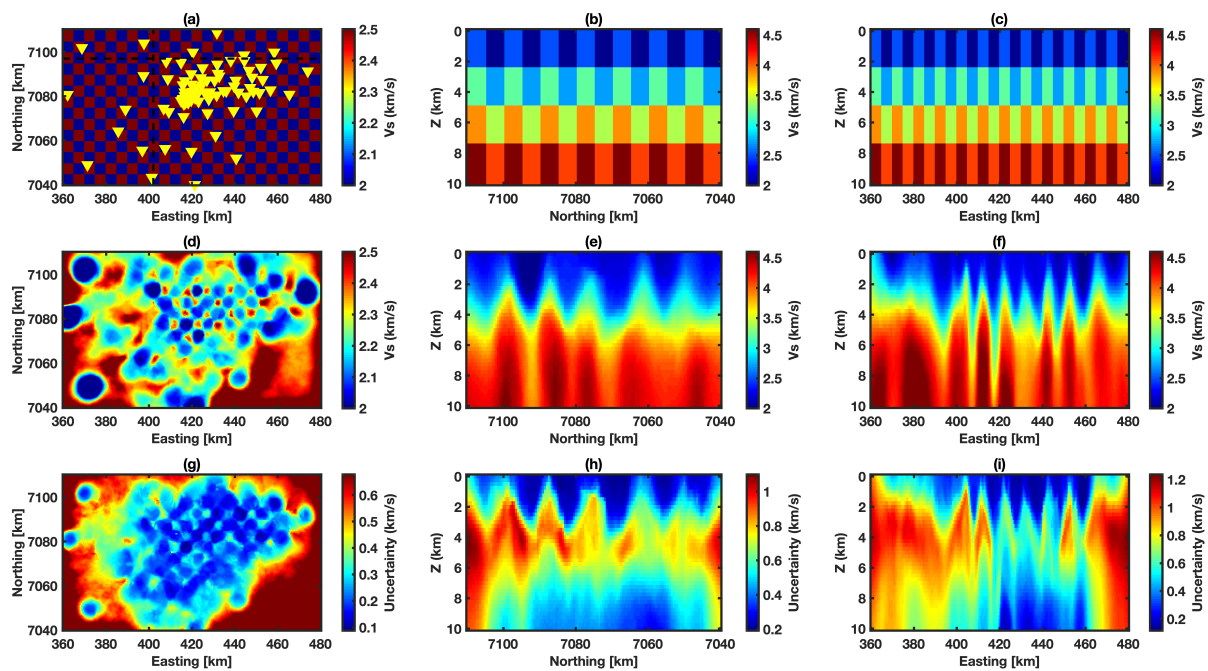
Figure 14. Transdimensional tomography results of the synthetic travel times containing Gaussian random noise for the fine block model in Figure 6c. The true block model: (**a**) horizontal slice at the surface of the model including station locations, (**b**) vertical cross-section at easting of 402 km, and (**c**) vertical cross-section at northing of 7097 km. The vertical cross-sections are indicated with black dashed lines in the horizontal slice of (**a**). (**d–f**) Pointwise averaged velocities calculated from post-burn-in retained samples. (**g–i**) Standard deviation (model uncertainty) calculated from post-burn-in retained samples.

4.4. Chain Statistics and Convergence

To evaluate the ability of the algorithm to infer the noise level and complexity of the model space, we show here the chain statistics for the intermediate block model. Results are approximately the same for the other two block models (i.e., the coarse block model and the fine block model). Figure 15 shows how misfit, number of cells (model dimension), and noise hyper-parameters change during McMC sampling. Figures 15a–d are for the noise-free experiment, and Figures 15e–h are for the experiment with additive noise. Different colors represent different chains. Noise hyper-parameters are presented for a single frequency as they vary by frequency. Looking at this figure, the number of cells converged to a small number of cells (around 300) in the experiment with additive noise. In the noise-free experiment, however, the number of cells approaches higher values and also the variation between different chains is larger. For the inferred noise level, in the experiment with additive Gaussian random noise, the recovered hyper-parameter *a* is fairly close to the true value. However, the recovered hyper-parameter *b* slightly underestimates the true value.

The results presented in Figure 15 can also be used as a means to determine chain convergence and the length of the burn-in period. As one can see, for the experiment with additive Gaussian noise, all parameters at all chains stabilize around the same value. This stationarity and convergence can be interpreted as a fully mixed Markov chain. For the noise-free experiment, however, it seems as if the separate chains still need more time to stabilize entirely; the misfit and the noise hyper-parameter are still decreasing, and the number of cells is still increasing. It means that if we continue sampling from the posterior for the noise-free experiment, the quality of results will improve and the Markov chain will reach a steady state. In other words, the experiment with additive Gaussian noise converged faster and hence better explores the posterior distribution for the same number of samples.
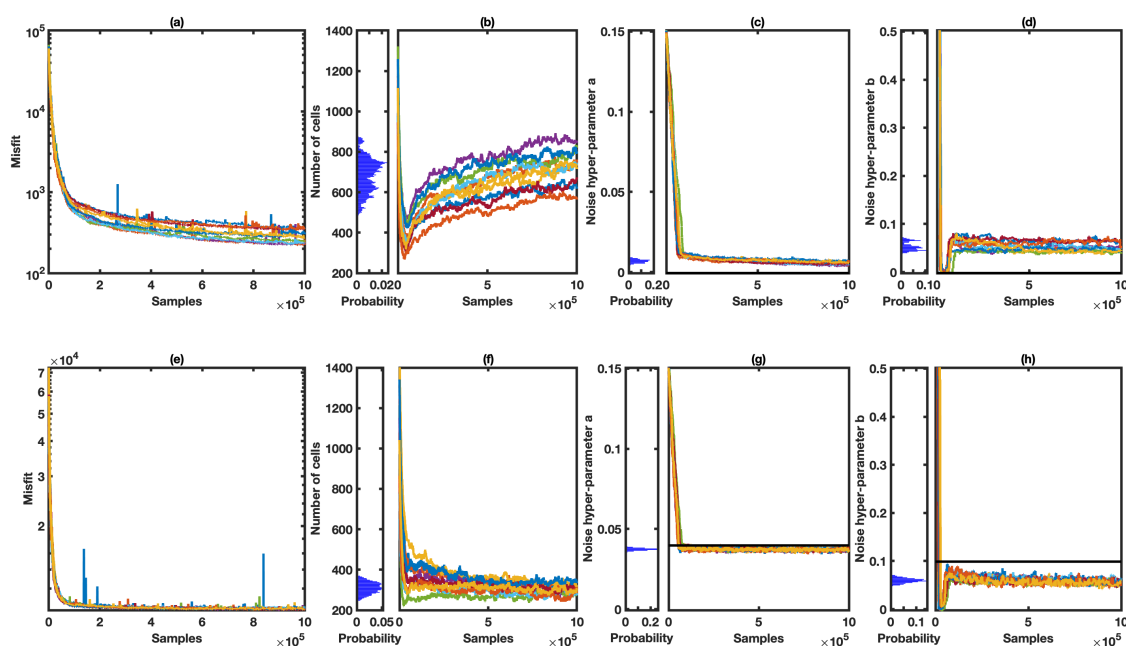
Figure 15. Variation of the misfit, the number of cells (dimensionality of model space), and the noise hyper-parameters during ten chains of McMC sampling for the intermediate block model in the noise-free experiment (**a**–**d**) and the experiment with additive Gaussian noise (**e**–**h**). Each color represents a sampling chain. Noise hyper-parameters are plotted for a single frequency. The black lines in (**g**–**h**) represent the actual values. Histograms of the posterior distribution for the retained models for each parameter are shown as insets of each panel.

5. Conclusions

In this paper, we investigated the ability of 3D transdimensional Markov chain Monte Carlo to recover the 3D surface wave velocity structure of the Reykjanes peninsula. Anticipating future application of this algorithm to interferometric travel times extracted from ambient seismic noise, we specifically considered travel times between stations of the Reykjanes seismic array. We find that updating ray paths every 200 iterations does not significantly affect the performance of the algorithm (i.e., it honors the non-linearity of the problem sufficiently), while at the same time significantly reducing computational costs. Our results show that the algorithm successfully adapts model resolution to ray density and hence yields a higher resolution and lower uncertainty in more densely sampled areas. Similarly, the uncertainty is larger in regions where the station density is lower. In addition, the algorithm successfully adapts to the noise level of the observed travel times; smoother models are obtained for higher levels of additive random noise. The algorithm converges (i.e., stabilizes) faster with a higher noise level because noise flattens the posterior and, consequently, the posterior is easier to explore. The uncertainty maps aid the interpretation of the results. The block interfaces are visible in uncertainty maps with higher uncertainties.

Author Contributions: Conceptualization, C.W.; Data curation, A.R.D.; Formal analysis, A.R.D.; Funding acquisition, A.R.D. and C.W.; Investigation, A.R.D. and C.W.; Methodology, A.R.D. and C.W.; Project administration, C.W.; Resources, C.W.; Software, A.R.D. and X.Z.; Supervision, C.W.; Validation, A.R.D. and C.W.; Visualization, A.R.D.; Writing—original draft, A.R.D.; Writing—review & editing, A.R.D. and C.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Friðleifsson, G.; Albertsson, A.; Elders, W.; Sigurdsson, Ó.; Karlsdottir, R.; Pálsson, B. The Iceland Deep Drilling Project (IDDP): Planning for the second deep well at Reykjanes. *Geotherm. Resour. Counc. Trans* **2011**, *35*, 347–354.
2. Friðleifsson, G.Ó.; Elders, W.A.; Zierenberg, R.A.; Fowler, A.P.; Weisenberger, T.B.; Mesfin, K.G.; Sigurðsson, Ó.; Níelsson, S.; Einarsson, G.; Óskarsson, F.; et al. The Iceland Deep Drilling Project at Reykjanes: Drilling into the root zone of a black smoker analog. *J. Volcanol. Geotherm. Res.* **2020**, *391*, 106435. [CrossRef]
3. Martins, J.E.; Weemstra, C.; Ruigrok, E.; Verdel, A.; Jousset, P.; Hersir, G. 3D S-wave velocity imaging of Reykjanes Peninsula high-enthalpy geothermal fields with ambient-noise tomography. *J. Volcanol. Geotherm. Res.* **2020**, *391*, 106685. [CrossRef]
4. Darnet, M.; Wawrzyniak, P.; Coppo, N.; Nielsson, S.; Schill, E.; Fridleifsson, G. Monitoring geothermal reservoir developments with the Controlled-Source Electro-Magnetic method—A calibration study on the Reykjanes geothermal field. *J. Volcanol. Geotherm. Res.* **2020**, *391*, 106437. [CrossRef]
5. Weemstra, C.; Boschi, L.; Goertz, A.; Artman, B. Seismic attenuation from recordings of ambient noise. *Geophysics* **2013**, *78*, Q1–Q14. [CrossRef]
6. Wapenaar, K.; Fokkema, J. Green's function representations for seismic interferometry. *Geophysics* **2006**, *71*, SI33–SI46. [CrossRef]
7. Kästle, E.D.; Soomro, R.; Weemstra, C.; Boschi, L.; Meier, T. Two-receiver measurements of phase velocity: Cross-validation of ambient-noise and earthquake-based observations. *Geophys. J. Int.* **2016**, *207*, 1493–1512. [CrossRef]
8. Sánchez-Pastor, P.; Obermann, A.; Schimmel, M.; Weemstra, C.; Verdel, A.; Jousset, P. Short- and Long-Term Variations in the Reykjanes Geothermal Reservoir From Seismic Noise Interferometry. *Geophys. Res. Lett.* **2019**, *46*, 5788–5798. [CrossRef]
9. Zhou, Y.; Dahlen, F.; Nolet, G. Three-dimensional sensitivity kernels for surface wave observables. *Geophys. J. Int.* **2004**, *158*, 142–168. [CrossRef]
10. Herrmann, R.B. Computer programs in seismology: An evolving tool for instruction and research. *Seismol. Res. Lett.* **2013**, *84*, 1081–1088. [CrossRef]
11. Bodin, T.; Sambridge, M. Seismic tomography with the reversible jump algorithm. *Geophys. J. Int.* **2009**, *178*, 1411–1436. [CrossRef]
12. Korostelev, F.; Weemstra, C.; Leroy, S.; Boschi, L.; Keir, D.; Ren, Y.; Molinari, I.; Ahmed, A.; Stuart, G.W.; Rolandone, F.; et al. Magmatism on rift flanks: Insights from ambient noise phase velocity in Afar region. *Geophys. Res. Lett.* **2015**, *42*, 2179–2188. [CrossRef]
13. Rawlinson, N.; Sambridge, M. Seismic traveltime tomography of the crust and lithosphere. *Adv. Geophys.* **2003**, *46*, 81–199.
14. Nolet, G. *A Breviary of Seismic Tomography*; Cambridge University Press: Cambridge, UK, 2008. [CrossRef]
15. Rawlinson, N.; Pozgay, S.; Fishwick, S. Seismic tomography: A window into deep Earth. *Phys. Earth Planet. Inter.* **2010**, *178*, 101–135. [CrossRef]
16. Bensen, G.; Ritzwoller, M.; Yang, Y. A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise. *Geophys. J. Int.* **2009**, *177*, 1177–1196. [CrossRef]
17. Martins, J.E.; Ruigrok, E.; Draganov, D.; Hooper, A.; Hanssen, R.; White, R.; Soosalu, H. Imaging Torfajökull's magmatic plumbing system with seismic interferometry and phase velocity surface wave tomography. *J. Geophys. Res. Solid Earth* **2019**, *124*, 2920–2940. [CrossRef]
18. Young, M.K.; Rawlinson, N.; Bodin, T. Transdimensional inversion of ambient seismic noise for 3D shear velocity structure of the Tasmanian crustTransdimensional shear velocity inversion. *Geophysics* **2013**, *78*, WB49–WB62. [CrossRef]
19. Tikhonov, A.N. On the solution of ill-posed problems and the method of regularization. In *Doklady Akademii Nauk*; Russian Academy of Sciences: Jerusalem, Russia, 1963; Volume 151, pp. 501–504.
20. Bodin, T.; Sambridge, M.; Rawlinson, N.; Arroucau, P. Transdimensional tomography with unknown data noise. *Geophys. J. Int.* **2012**, *189*, 1536–1556. [CrossRef]
21. Green, P.J.; Hastie, D.I. Reversible jump MCMC. *Genetics* **2009**, *155*, 1391–1403.
22. Galetti, E.; Curtis, A.; Meles, G.A.; Baptie, B. Uncertainty loops in travel-time tomography from nonlinear wave physics. *Phys. Rev. Lett.* **2015**, *114*, 148501. [CrossRef] [PubMed]
23. Galetti, E.; Curtis, A.; Baptie, B.; Jenkins, D.; Nicolson, H. Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry. *Geophys. J. Int.* **2017**, *208*, 36–58. [CrossRef]
24. Zhang, X.; Curtis, A.; Galetti, E.; De Ridder, S. 3-D Monte Carlo surface wave tomography. *Geophys. J. Int.* **2018**, *215*, 1644–1658. [CrossRef]
25. Zhang, X.; Hansteen, F.; Curtis, A.; de Ridder, S. 1-D, 2-D, and 3-D Monte Carlo Ambient Noise Tomography Using a Dense Passive Seismic Array Installed on the North Sea Seabed. *J. Geophys. Res. Solid Earth* **2020**, *125*, e2019JB018552. [CrossRef]
26. Weemstra, C.; De Laat, J.I.; Verdel, A.; Smets, P. Systematic recovery of instrumental timing and phase errors using interferometric surface-waves retrieved from large-N seismic arrays. *Geophys. J. Int.* **2021**, *224*, 1028–1055. [CrossRef]
27. Weemstra, C.; Obermann, A.; Verdel, A.; Paap, B.; Blanck, H.; Guðnason, E.Á.; Hersir, G.; Jousset, P.; Sigurðsson, A. Time-lapse seismic imaging of the Reykjanes geothermal reservoir. In Proceedings of the European Geothermal Congress, Strasbourg, France, 19–24 September 2016.
28. Bayes, T. Essay towards solving a problem in the doctrine of chances. *Biometrika* **1958**, *45*, 293–315. [CrossRef]

29. Sambridge, M.; Rawlinson, N.; Levander, A.; Nolet, G. Seismic tomography with irregular meshes. *Geophys.-Monogr.-Am. Geophys. Union* **2005**, *157*, 49.
30. Bodin, T.; Sambridge, M.; Gallagher, K. A self-parametrizing partition model approach to tomographic inverse problems. *Inverse Probl.* **2009**, *25*, 055009. [CrossRef]
31. Belhadj, J.; Romary, T.; Gesret, A.; Noble, M.; Figliuzzi, B. New parameterizations for Bayesian seismic tomography. *Inverse Probl.* **2018**, *34*, 065007. [CrossRef]
32. Hawkins, R.; Bodin, T.; Sambridge, M.; Choblet, G.; Husson, L. Trans-dimensional surface reconstruction with different classes of parameterization. *Geochem. Geophys. Geosyst.* **2019**, *20*, 505–529. [CrossRef]
33. Ryberg, T.; Haberland, C. Bayesian inversion of refraction seismic traveltime data. *Geophys. J. Int.* **2017**, *212*, 1645–1656. [CrossRef]
34. Kurita, T. Regional variations in the structure of the crust in the central United States from P-wave spectra. *Bull. Seismol. Soc. Am.* **1973**, *63*, 1663–1687. [CrossRef]
35. Vidale, J.E. Finite-difference calculation of traveltimes in three dimensions. *Geophysics* **1990**, *55*, 521–526. [CrossRef]
36. Qin, F.; Luo, Y.; Olsen, K.B.; Cai, W.; Schuster, G.T. Finite-difference solution of the eikonal equation along expanding wavefronts. *Geophysics* **1992**, *57*, 478–487. [CrossRef]
37. Sethian, J.A. A fast marching level set method for monotonically advancing fronts. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1591–1595. [CrossRef] [PubMed]
38. Rawlinson, N.; Sambridge, M. Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophys. J. Int.* **2004**, *156*, 631–647. [CrossRef]
39. Gorbatov, A.; Fukao, Y.; Widiyantoro, S.; Gordeev, E. Seismic evidence for a mantle plume oceanwards of the Kamchatka—Aleutian trench junction. *Geophys. J. Int.* **2001**, *146*, 282–288. [CrossRef]
40. Andrieu, C.; de Freitas, N.; Doucet, A. Robust Full Bayesian Learning for Neural Networks. Citeseer. 1999. Available online: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.36.5243 (accessed on 6 October 2021).
41. Gelman, A.; Roberts, G.O.; Gilks, W.R. Efficient Metropolis jumping rules. *Bayesian Stat.* **1996**, *5*, 42.