SURFACE WAVE DISPERSION INVERSION USING AN ENERGY LIKELIHOOD FUNCTION

ABSTRACT

Seismic surface wave dispersion inversion is widely used to study the subsurface structure of the Earth. The dispersion property is usually measured by using frequency-phase velocity (f-c) analysis and by picking phase velocities from the obtained f-c spectrum. However, because of potential contamination the f-c spectrum often has multimodalities at each frequency for each mode. These introduce uncertainty and errors in the picked phase velocities, and consequently the obtained shear velocity structure is biased. To overcome this issue, in this study we introduce a new method which directly uses the spectrum as data. We achieve this by solving the inverse problem in a Bayesian framework and define a new likelihood function, the energy likelihood function, which uses the spectrum energy to define data fit. We apply the new method to a land dataset recorded by a dense receiver array, and compare the results to those obtained using the traditional method. The results show that the new method produces more accurate results since they better match independent data from refraction tomography. This real-data application also shows that it can be applied efficiently since it removes the need to pick phase velocities, and with relatively little adjustment to current practice since it uses standard f-c panels to define the likelihood. We therefore recommend using the energy likelihood function rather than explicitly picking phase velocities in surface wave dispersion inversion.

1 Introduction

Seismic surface waves travel along the surface of the Earth while oscillating over depth ranges that depend on their frequency of oscillation [1]. This in turn makes surface waves dispersive – different frequencies travel at different speeds, and these speeds are sensitive to different parts of the Earth. By measuring speeds at different frequencies this dispersion property can be used to constrain subsurface structures over different depth ranges on global scale [43, 39, 27, 26], regional scale [57, 7, 42, 45] and industrial scale [32, 44, 8, 53].

Surface wave dispersion property (phase or group velocities at different frequencies) can be measured in different ways depending on different acquisition systems. In the case of a single station or a sparse receiver array, as is often the case in seismology, the dispersion property can be measured by using the frequency-time analysis (FTAN) method [10, 22, 21, 17, 38, 36, 20, 2]. In FTAN one constructs a frequency-time domain envelope image for each seismic trace by using a set of narrow bandpass Gaussian filters, and measures the group velocity using the arrival time of the maximum envelope at each frequency. The phase velocity can be derived using the phase of the signal at the time of the maximum envelope plus a phase ambiguity term (appropriate integer multiple of $2\pi$) and a source phase term, or by using an image transformation technique [45]. However, the obtained measurements are often uncertain because of the trade-off between the determination of frequency of a signal and the determination of arrival time of a signal, and because of the contamination contained in the envelope image caused by effects such as multipathing. In addition,
those methods cannot discriminate between different modes, which introduces errors to the measurements made for the fundamental or other modes and cannot provide measurements for higher modes.

In the case where a dense receiver array is deployed, the phase velocity can be determined using frequency-phase velocity (f-c) analysis, in which wavefields recorded by the array are slant stacked to obtain a f-c spectrum \([33, 44] \). The phase velocity is then determined as the velocity associated with the highest energy at each frequency in the spectrum. Because different modes are separated in the f-c domain, the method can also be applied to determine phase velocities for higher modes. However, the obtained spectrum may still be contaminated by multipathing effects or strong lateral heterogeneity \([18] \), and consequently the velocity associated with the highest energy does not necessarily represent the most appropriate phase velocity measurement. This issue can be overcome by imposing additional prior information to the phase velocity, for example, by forcing the picked phase velocity dispersion curve to be continuous. Unfortunately, on one hand such additional constrains are usually achieved by picking the phase velocity deliberately and manually for each spectrum, which cannot be applied to large datasets. On the other hand an automatic procedure can easily introduce errors to the picked phase velocities because of the complexity of the spectrum.

To overcome these issues in phase velocity estimation, in this study we introduce a new method which directly uses the spectrum as data rather than explicit picks of the phase velocities. The spectrum of data has been used in wave-equation dispersion inversion in the framework of full-waveform inversion \([23] \). However, that method is computationally expensive and the problem is solved using a deterministic method which cannot provide uncertainty estimates. To quantify uncertainty we solve the dispersion inversion problem using Bayesian inference. In Bayesian inference one constructs a so-called posterior probability density function (pdf) that describes the remaining uncertainty of models post inversion, by combining prior information with the new information contained in the data as represented by a pdf called the likelihood. The likelihood function describes the probability of observing data given a specific model, and is traditionally assumed to be a Gaussian distribution centred on the picked phase velocities. In this study we propose a new likelihood function, called the \textit{energy} likelihood function which directly uses the spectrum based on the intuition that higher energy in the spectrum reflects higher probability of observing the associated phase velocity.

We apply the new method to a land dataset recorded by a dense array and compare the results with those obtained using the traditional method. The dataset consists of raw shot records taken from a sub-area of a nodal land seismic survey that was conducted in a desert environment \([31] \). This dataset offers ultra-high trace density with over 180 million traces per km\(^2\) on a 12.5 m \(\times\) 12.5 m receiver grid and a 100 m \(\times\) 12.5 m source grid. The increased trace density greatly improves the spatial sampling of the wavefield, which in turn benefits the recording and analysis of surface waves. As a part of the depth model building process, refraction tomography was performed to yield a shallow P-wave velocity model \([5] \), which is then used here for qualitative comparison with the shallow S-wave velocity model obtained from our method.

To solve the Bayesian inference problem, we use the reversible-jump Markov chain Monte Carlo (rj-McMC) method. The rj-McMC method is a generalized McMC method which allows a trans-dimensional inversion to be carried out, meaning that the dimensionality of parameter space (the number of parameters) can vary in the inversion \([15] \). Thus the parameterization itself can be dynamically adapted to the data and to the prior information. The method has been used to estimate phase or group velocity maps of the crustal structure \([3, 13, 56] \) and to estimate shear velocity structures of the crust and upper mantle using surface wave dispersion data \([4, 40, 46, 14, 53] \).

In the following section we first perform frequency-phase velocity analysis for the recorded data to obtain the f-c spectrum around each geographic location. In section 3 we introduce the new energy likelihood function and give an overview of the rj-McMC algorithm. We then apply the new likelihood function to the obtained spectra to estimate the shear velocity structure, and compare the results with those obtained using the traditional method. The results demonstrate that the new method can generate more accurate results than the traditional method, and can be applied efficiently to large data sets. We therefore conclude that the energy likelihood function provides a valuable tool for surface wave dispersion inversion.

## 2 Surface wave dispersion analysis

Figure [1] shows the locations of all 100,627 sensors which are deployed in a regular grid with a spacing of approximately 12.5 m in both directions, and record samples at 250 Hz. In total, 70261 active sources are fired with 100 m spacing in X direction and 12.5 m spacing in Y direction to generate seismic surface waves. Figure [1b] shows an example of a shot gather which mainly contains surface waves.

To analyse the surface wave dispersion, we performed frequency-phase velocity (f-c) analysis of the recorded data. For a given geographic location \(p\), the f-c spectrum \(U_p(c, w)\) can be computed using the data recorded by a receiver...
Surface wave dispersion inversion using an energy likelihood function

\[ U_p(c, w) = \int_{C_p} e^{-j \frac{2\pi}{c} x} u(x, w) |u(x, w)| dx \approx \sum_{i=1}^{N_p} e^{-j \frac{2\pi}{c} x_i} u(x_i, w) |u(x_i, w)| \Delta x \]  

where \( C_p \) denotes that the integration is performed around the location \( p \), \( x \) is the source-receiver distance, \( w \) is frequency in radian, \( c \) is phase velocity, \( j = \sqrt{-1} \), \( i \) is the index of records and \( N_p \) is the number of receivers around location \( p \); \( u(x, w) \) is the Fourier transform of the wavefield \( u(x, t) \). For a given receiver array a larger \( N_p \) improves resolution of phase velocity, but reduces spatial resolution. In this study for a given location \( p \) we stack all the records whose receiver and source locations are respectively within 300 m and 1500 m to the location \( p \). These threshold distances are selected such that the phase velocity dispersion curve can be clearly identified in the spectrum without increasing the number of records unnecessarily. This process is repeated for every geographic location on a regular grid with a 12.5 m spacing in both directions across the survey area.

Figure 2a shows an example spectrum obtained using the above method (displayed as phase velocity versus period) at one specific geographic location (red star in Figure 1a). The spectrum shows three modes. The phase velocity of the fundamental mode varies from 500 m/s to 800 m/s and contains two different branches at periods shorter than 0.25 s. This might be caused by effects such as multipathing and strong heterogeneity, or represents two different surface wave modes. To further understand this, we performed an inversion using one of the branches (black dots in Figure 2a) and modelled the first overtone using the obtained shear velocity profile (see Appendix A). The results show that the modelled first overtone is close to the mode with velocity higher than 600 m/s (see Figure A1). This therefore demonstrates that the two branches are probably not associated with different modes. The first overtone mainly appears in the period range from 0.1 s to 0.25 s with phase velocities varying from 620 m/s to 1000 m/s. The second overtone has much lower energy compared to the other two modes, and we discarded this mode in the inversion.

Traditionally for each mode those phase velocities associated with the peak energy are used as data to constrain the subsurface shear velocity. However, for those modes that have complex structures in the spectrum (such as the fundamental mode in Figure 2a), it becomes difficult to determine the correct phase velocity dispersion curve to use as these may have apparent jumps between neighbouring periods (e.g., the dashed black line in Figure 2a). One way to reduce this issue is to impose continuity or smoothness constrains on the dispersion curves. For example, in Figure 2 the black dots are determined by forcing the dispersion curve to be smooth. Unfortunately the picked dispersion curve is then forced to follow one of the branches at shorter periods (< 0.25 s) which may not be the true dispersion curve generated by the subsurface, and consequently the inverted velocity structure may be biased. In addition, for large datasets the dispersion curves need to be determined automatically which introduces difficulties to balance higher energy against the smoothness of dispersion curves, and therefore may cause errors in the estimated phase velocities.

In the next section, we propose a method which directly uses the f-c spectrum as data to constrain the subsurface velocity structure, thus avoiding these issues.

---

**Figure 1:** (a) The acquisition system. Blue dots denote receiver locations. The red star is referred to in the text. (b) shows an example of a shot gather, which displays the waveform data generated by a specific shot and recorded by receivers at different offsets.
Surface wave dispersion inversion using an energy likelihood function

3 Shear-wave velocity inversion

The surface wave dispersion information obtained as above can be used to constrain the subsurface shear velocity structure, which involves solving a nonlinear and nonunique inverse problem. In this study we use Bayesian inference to characterize the fully nonlinear uncertainty of the solution.

3.1 Bayesian inference

In Bayesian inference one constructs a so-called posterior pdf $p(m|d_{obs})$ of velocity model $m$ given the observed data $d_{obs}$, by combining prior information with new information contained in the data. According to Bayes’ theorem,

$$p(m|d_{obs}) = \frac{p(d_{obs}|m)p(m)}{p(d_{obs})} \tag{2}$$

where $p(m)$ describes the prior information of model $m$ that is independent of the current data, $p(d_{obs}|m)$ is called the likelihood which describes the probability of observing $d_{obs}$ if model $m$ was true, and $p(d_{obs})$ is a normalization factor called the evidence.

The prior information is critical for Bayesian inference. To construct a more informative prior distribution than the commonly-used Uniform distribution with little or no depth dependence, we first conduct a set of inversions at multiple geographic locations using a Uniform prior pdf from 300 m/s to 1500 m/s, which spans the range of shear velocities in the upper 500 m according to a variety of similar studies [19, 29, 6, 53]. The average of the mean models from these inversions are then used as the mean of the prior pdf, and we construct a Uniform distribution with a width of 1000 m/s (larger than four standard deviations obtained from the previous inversions) at each depth (Figure 2b). This prior information improves the depth resolution and constrains the subsurface structure better than an identical Uniform distribution across the depth ranges [47].

3.2 Energy likelihood function

In the f-c spectrum higher intensity (energy) means higher probability that the associated phase velocity represents the true phase velocity of a surface wave mode. Based on this assumption, we can directly use the spectrum as data and write a new likelihood function. Define $E$ as the matrix representing the f-c spectrum and $E(T, c)$ as the energy

Figure 2: (a) An example spectrum obtained using the f-c analysis at a specific location (red star in Figure 1b). Black dots denote the picked phase velocities, black pluses show the phase velocities associated with the second branch for the fundamental mode and the black dashed line shows the phase velocity dispersion curve associated with the maximum energy at each period. The white line is used to separate the two modes. (b) shows the prior information, which is a Uniform distribution with an interval up to 1000 m/s at each depth. The red line shows the mean of the prior pdf.
at period $T = 1/f$ and phase velocity $c$, and assuming that the energy at each value of $T$ and $c$ has exponentially decaying probability away from the maximum energy value at that period, the likelihood can be expressed as:

$$p(E|\mathbf{m}) = \frac{1}{Z} \exp\left[-\sum_i \frac{\max(E(T_i, \cdot)) - E(T_i, c_i(\mathbf{m}))}{\sigma_i^2}\right]$$

(3)

where $T_i$ is the $i^{th}$ period, $\max(E(T_i, \cdot))$ is the maximum energy at period $T_i$ which guarantees that the exponent is negative, $c_i(\mathbf{m})$ is the phase velocity at period $T_i$ predicted using model $\mathbf{m}$, $\sigma_i$ is a scaling factor and $Z$ is a normalization factor. The scaling factor $\sigma_i$ is generally unknown, so we treat it as an additional parameter and estimate it hierarchically [25].

For multimodal inversion the energy of the fundamental mode may dominate the likelihood function in equation 3 (for example, see Figure 2b), and consequently the inverted results can be biased because models may have apparently larger likelihood values if their predicted higher modes also fit the fundamental mode energy. We therefore separate each mode by windowing out other modes. For example, define $E_j$ to be the spectrum of the $j^{th}$ mode after other modes have been windowed out. Then the likelihood function becomes:

$$p(E|\mathbf{m}) = \frac{1}{Z} \exp\left[-\sum_{ij} \frac{\max(E_j(T_i, \cdot)) - E_j(T_i, c_i(\mathbf{m}))}{\sigma_{ij}^2}\right]$$

(4)

Although this requires that we define a window function for each mode, this process is usually straightforward. For example, the white dashed line in Figure 2a shows the boundary used to separate the first two modes; this same line is used for all other spectra across the survey area since it appeared appropriate for a large number of spectra examined manually.

### 3.3 Reversible-jump Markov chain Monte Carlo (rj-McMC)

We use reversible-jump Markov chain Monte Carlo (rj-McMC) to generate samples from the posterior pdf. The rj-McMC method is a generalised version of the Metropolis-Hastings algorithm [28, 16], which allows the number of model parameters to be variable in the inversion [15]. Thus the parameterization of the seismic velocity model can itself be determined by the data and prior information. The method has been applied in a range of geophysical applications [24, 3, 46, 13, 54, 14, 52, 53, 54]. In this study we use the method to solve the surface wave dispersion inversion problem.

In rj-McMC one constructs a (Markov) chain of samples by perturbing the current model $\mathbf{m}$ using a proposal distribution $q(\mathbf{m}'|\mathbf{m})$ to generate a new model $\mathbf{m}'$, and by accepting or rejecting this new model with a probability $\alpha(\mathbf{m}'|\mathbf{m})$ called the acceptance ratio:

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

(5)

where $J$ is the Jacobian matrix of transforming $\mathbf{m}$ to $\mathbf{m}'$ and is used to account for any volume changes of parameter space during jumps between dimensionalities. In our case the Jacobian matrix is an identity matrix [3].

For surface wave dispersion inversion beneath each geographical location we use a set of layers to parameterize the subsurface, which can be changed in different ways within the rj-McMC algorithm: adding a new layer, removing a layer, changing layer positions and changing layer velocities. There is also another perturbation related to the hypoparameters of the likelihood function: changing the scaling factor $\sigma_{ij}$ in equation 4 or standard deviation in the Gaussian likelihood function. After each perturbation of the current model $\mathbf{m}$, the acceptance ratio $\alpha$ is computed using equation 5 and is compared with a random number $\gamma$ generated from the Uniform distribution on [0,1]. If $\gamma < \alpha$ the new model is accepted; otherwise the new model is rejected and the current model is repeated as a new sample in the chain. This process guarantees that the generated samples are distributed according to the posterior pdf if the number of samples tends to infinity [15].

### 3.4 A 1D example

We first apply the above method to the dispersion data in Figure 2a using both the energy likelihood function and the Gaussian likelihood function and compare their results. We use both the fundamental mode and the first overtone with 21 equally spaced periods from 0.0835 s to 0.425 s. The prior pdf of shear velocity is shown in Figure 2b and the prior pdf of the number of layers is set to be a discrete Uniform distribution between 2 and 25. For the energy likelihood function the prior pdf of the scaling factor $\sigma_{ij}$ is chosen to be a Uniform distribution between 0.5 and 20, and for the Gaussian likelihood function in the traditional method the prior pdf of the data noise (standard deviation of the
Figure 3: (a) The shear velocity and (b) the phase velocity marginal posterior distributions obtained using the posterior samples of the energy likelihood function for the spectrum in Figure 2. The phase velocity distribution is obtained by modelling the phase velocity for all posterior samples. (c) and (d) show the associated shear velocity and phase velocity marginal distributions obtained by the traditional method of using the picked phase velocities shown by black dots in Figure 2. Red dashed lines show the mean velocity profile. White dots and pluses are associated with the black dots and pluses in Figure 2 respectively.
Gaussian distribution) is a Uniform distribution between 0.05 and 50. We use Gaussian distributions as the proposal distribution: the width of the Gaussian distribution for the fixed-dimensional steps (changing layer positions, changing layer velocities and changing likelihood hyperparameters) is chosen by trial and error to ensure an acceptance ratio between 20 and 50 percent; the width of the transdimensional steps (adding or removing a layer) is selected to produce the maximum acceptance ratio. For each inversion we run six chains, and each chain contains 800,000 iterations with a burn-in period of 300,000 during which all samples are discarded. After the iteration each chain is thinned (decimated) by a factor of 50, and the remaining samples are used for the subsequent inference.

Figure 3 and 4 show the marginal distributions of shear velocity obtained using the energy likelihood function and the Gaussian likelihood function respectively. The corresponding phase velocity distributions generated by those posterior samples are shown in Figure 5 and 6 respectively. The shear velocity marginal distribution obtained using the energy likelihood function shows clearly multimodal distributions in the near surface (< 100 m), which are associated with the two branches in the f-c spectrum (Figure 2a and Figure 3b). In comparison the marginal distribution obtained using the Gaussian likelihood function shows a unimodal distribution and the predicted data only fit the single branch on which the phase velocities were picked. Since we do not know which branch reflects the most appropriate phase velocity a priori, the shear velocity obtained using the Gaussian likelihood function is biased and the estimated uncertainty failed to take account of the full, multi-modal uncertainty in the data. In contrast, by directly using the spectrum as data one can embed all data uncertainty in the likelihood function and therefore obtain a less biased result. In addition, the phase velocity distribution of the first overtone obtained using the energy likelihood function (Figure 3b) fits the picked phase velocity better than that obtained using the Gaussian likelihood function. This suggests that there is inconsistency between the picked phase velocities of the fundamental mode and the first overtone, and therefore further demonstrates the necessity of including the second branch in the likelihood function. In the deeper part (> 100 m) the velocity obtained using the energy likelihood function decreases from 600 m/s at 100 m depth to 1500 m/s around 220 m and stays almost constant down to 500 m; whereas the velocity obtained using the Gaussian likelihood function is around 1200 m/s across the whole depth from 100 m to 500 m. This is probably because the picked phase velocities (black dots in Figure 2b) are not sensitive to the deeper part (> 200m), and consequently the shear velocity in the deeper part are dominated by the prior pdf. In comparison the energy likelihood function uses all the information contained in the spectrum and constrains deeper structure better. For example, at periods longer than 0.22s phase velocities are not determined for the first overtone because of its lower energy, whereas the information is still used in the energy likelihood function to constrain deeper structures. This can also be observed in the phase velocity distributions predicted from posterior samples (Figure 5a and d), where the first overtone distribution obtained using the energy likelihood function at longer periods (> 0.22 s) is more similar to the spectrum than that obtained using the Gaussian likelihood function. Thus, by directly using the spectrum as data the new energy likelihood function can use more information in the data and can obtain more accurate results than the traditional method.

4 3D results

To obtain 3D shear velocity models we perform 1D inversions to all of the spectra across the survey area. For each inversion at each geographic location the inversion is conducted in the same way as described in the previous section with the same prior pdf and proposal pdf. To compare the results we use both the energy likelihood function and the Gaussian likelihood function around picked phase velocities. For each spectrum we automatically determine the phase velocity from longer periods to shorter periods: at each period the phase velocity is determined as the local maximum energy point whose phase velocity is closest to the already picked phase velocity at the neighbouring period, such that the picked dispersion curve is as continuous as possible. To ensure the quality of picked phase velocities, we only determine the phase velocity at frequencies with sufficiently high energy, such that the picked energy is at least 1.8 times higher than the average energy at each frequency.

Figure 4 shows the mean and standard deviation models at the depth of 20 m, 80 m, 140 m and 200 m obtained using the energy likelihood function. Among various structures we have highlighted several velocity anomalies using boxes which we will refer to below. Overall the standard deviation model shows lower uncertainty at the shallower part (20 m and 80 m) and higher uncertainty at the deeper part (140 m and 200 m) due to the fact that seismic surface waves are more sensitive to near surface structure. At all the depths the standard deviation shows similar features to the mean model, but notice that at 20 m depth higher velocity anomalies are associated with lower uncertainties, whereas at 140 m depth higher velocity anomalies correspond to higher uncertainties. This phenomenon has also been observed previously [48, 49, 50], and the different correlation between velocity anomalies and uncertainties at different depths suggests that there is a complex, nonlinear relationship between seismic velocity and phase velocity data.

Figure 5 shows the mean and standard deviation models obtained using the Gaussian likelihood function at the same depths. Overall the mean model obtained using the Gaussian likelihood function shows more small scale variations than that obtained using the energy likelihood function. This is probably caused by errors in the picked phase velocities,
which is inevitable when the phase velocities are estimated automatically. At 20 m depth the mean model shows similar structures to those observed in the previous results in Figure 4, for example the northern high velocity anomaly denoted by black dashed boxes and the low velocity anomalies in the southease and northwest. However, the low velocity anomalies that are denoted by black and red solid line boxes in Figure 4 are not present in Figure 5. Similar to the previous results, at 80 m depth there are small scale variations in the east, but the low velocity anomaly denoted by red dashed line box and the high velocity anomaly next to this low anomaly are not visible. At greater depths (140 m and 200 m) the mean velocity model is significantly different from the previous results and contains many small scale structures which probably do not reflect the true geologic structure as the scale of these structures is smaller than the scale (300 m) used for stacking which implicitly imposes smoothness to the velocity structure. This suggests that the traditional method can cause bias in the inverted velocity structure because of errors in the picked phase velocity and loss of useful information in the data. Similarly to the previous results, the standard deviation model shows similar features to the mean model, and the uncertainty is lower in shallower parts.

To further understand the results we compare the above shear velocity models with P-wave velocity model (Figure 6) obtained using refraction tomography from the same dataset [5]. At 20 m depth the Vp model shows similar features to the shear velocity models. For example, there is a high velocity anomaly in the east between X = 6000 m and 8000 m, which may be related to the high velocity anomaly (black dashed line box) observed in the shear velocity models even though they are not at exactly the same location. Similarly to the shear velocity model obtained using the energy likelihood function, there is a low velocity anomaly between X = 2000 m and 5000 m in the southeastern direction (black solid line box) and a low velocity anomaly in the southwestern corner (red solid line box). This strongly suggests that the new energy likelihood method is effective since those low velocity anomalies are not visible in the results obtained using the traditional method (Figure 5). Even though at depths of 80 m and 140 m the Vp model is different from both shear velocity models, there are still similarities between the Vp model and the shear velocity model obtained using the energy likelihood function. For example, at 80 m depth there is a low velocity channel in the southeastern direction (denoted by red dashed line boxes) in the west of both models. At 200 m depth the Vp
model shows more similarities to the shear velocity model obtained using the energy likelihood function: both models show a high velocity anomaly between X = 2000 m and 5000 m in the southeastern direction across the area (the same location as denoted by the black solid line box at 20 m depth) and a low velocity channel to the east of this high velocity anomaly. In the east (X > 4000 m) there are similar small scale high velocity anomalies in both models. In comparison the shear velocity model obtained using the traditional method is very different from the Vp model. This again shows that the new energy likelihood function generates more accurate shear velocity models than the traditional method.

Figure 7 shows vertical sections through the different models at Y = 900 m. To further compare these models, in Figure 7f we show the average velocity profile across the section of each model with the P-wave velocity converted to shear velocity using a factor of 1.73 [35]. In the near surface (< 100m) all three models show a lower velocity layer (Figure 7f), although the two shear velocity models show more complex structures than the P-wave velocity. Between depths of 100 m and 230 m the shear velocity obtained using the energy likelihood function increases from 600 m/s to 1500 m/s which is consistent with the trend of P-wave velocity, and both models show a high velocity layer below 200 m. In comparison the shear velocity obtained using the traditional method is around 1300 m/s below 100 m. A similar phenomenon was observed in section 3 in the 1D example, which is caused by the fact that the picked phase velocities are not sensitive to the deeper structure. For example, the standard deviation model obtained using the traditional method shows higher uncertainties (> 300 m/s) below 200 m, whereas that obtained using the energy likelihood function shows a much lower uncertainty (< 100 m/s) between 200 m and 400 m. In addition, the shear velocity model obtained using the traditional method shows higher lateral spatial variations in deeper parts of the section (> 200 m) than the other two models. This further demonstrates the importance of using all information contained in the spectrum, rather than using only picked phase velocities. Note that both shear velocity standard deviation models show higher uncertainties at the layer boundary, which has been observed previously [13] and reflects the uncertainty of boundary positions.
Figure 6: The P-wave velocity model obtained using refraction tomography across horizontal slices at depths 20 m, 80 m, 140 m and 200 m. Boxes are the same as in Figure 4.
Figure 7: The mean and standard deviation model obtained using the energy likelihood function (a and b) and the traditional method (c and d) along a vertical section at Y = 875 m. (e) shows the vertical section of the P-wave velocity model obtained using refraction tomography. (f) shows the average shear velocity profiles across the section obtained using the suite of methods. The P-wave velocity is converted to shear velocity using a Vp/Vs ratio of 1.73.

5 Discussion

The McMC method is generally computationally expensive. For the above inversion each chain took 0.173 CPU hours using one core of an Intel Xeon CPU, which results in a total of 1.04 CPU hours for each geographic location and 107,015.4 CPU hours for all 102,817 inversions across the survey area. However, the inversions are fully parallelizable since the inversion at each geographical location is independent of other inversions. For example, the above inversion across the whole survey area took 15 hours using 7,200 CPU cores. We also note that other methods can be used to further improve the computational efficiency, for example, Hamiltonian Monte Carlo [9, 12], Langevin Monte Carlo [37, 41], variational inference [30, 48, 49, 55, 50] and neural network inversion [27, 26, 11, 51].

In the above results the phase velocities used for the traditional method are determined automatically from the f-c spectrum, and these may contain errors due to the complexity of the f-c spectrum. We therefore note that the results obtained using the traditional method may be improved if the phase velocities can be determined more accurately. However, this usually requires deliberately and manually picking of a dispersion curve for each spectrum, which restricts its application to a relatively small number of spectra. By contrast the new energy likelihood function directly using the spectrum as data and can easily be applied to larger datasets.

We performed the surface wave dispersion inversion independently at each geographic location, which loses lateral spatial correlations between neighbouring velocities and can cause biases in the final results [52]. For example, there are horizontal discontinuities in the shear velocity mean and standard deviation models (see Figure 7), which probably do not reflect the real subsurface structure. To include the lateral spatial correlation and to further improve the results a 3D parameterization may be used instead of the 1D parameterization, for example, 3D Voronoi tessellation can be used in surface wave tomography [52, 53].

In this study we used the spectrum obtained by stacking wavefields in the f-c domain, which requires a dense receiver array. In cases where only a sparse array is available, the energy likelihood can still be used to perform the inversion. For example, one could directly use the spectrum obtained between each source-receiver or receiver-receiver as the data to conduct phase or group velocity tomography using earthquake-generated surface waves or ambient noise data.
6 Conclusion

In this study we introduced a new likelihood function for seismic surface wave dispersion inversion, called the energy likelihood function which directly uses the spectrum as data. We applied the new likelihood function to image the subsurface shear velocity structure using surface wave data recorded by a dense array, and compared the results with those obtained using the traditional method. The results showed that the new likelihood function can take account of all information contained in the spectrum and produce a less biased result than that obtained using the traditional method. In addition the velocity model obtained using the new likelihood function is more similar to the P-wave velocity structure than that obtained using the traditional method. Because the new likelihood function directly uses the spectrum as data, it requires less effort to apply to large datasets than the traditional method, since the latter requires that we determine the phase velocity for each spectrum either manually or automatically and cannot be applied easily to large datasets. We thus conclude that the new energy likelihood function provides a powerful way to conduct surface wave dispersion inversion.

Acknowledgments

The authors thank the Edinburgh Imaging Project sponsors (BP and Total) for supporting this research. We would like thank BP and ADNOC for providing the seismic data and CGG for the refraction tomography model.

References

[1] Keiiti Aki and Paul G Richards. Quantitative seismology. 1980.
[2] GD Bensen, MH Ritzwoller, MP Barmin, A Lin Levshin, Feifan Lin, MP Moschetti, NM Shapiro, and Yanyan Yang. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophysical Journal International, 169(3):1239–1260, 2007.
[3] Thomas Bodin and Malcolm Sambridge. Seismic tomography with the reversible jump algorithm. Geophysical Journal International, 178(3):1411–1436, 2009.
[4] Thomas Bodin, Malcolm Sambridge, H Tkalcic, Pierre Arroucau, Kerry Gallagher, and Nicholas Rawlinson. Transdimensional inversion of receiver functions and surface wave dispersion. Journal of Geophysical Research: Solid Earth, 117(B2), 2012.
[5] F. Buriola, K. Mills, S. Cooper, S. Hollingworth, A. Crosby, and A. Ourabah. Ultra-high density land nodal seismic – processing challenges and rewards. In 82nd EAGE Annual Conference and Exhibition, 2021.
[6] M Chmiele, A Mordret, P Boué, F Brenguier, T Lecocq, R Courbis, D Hollis, X Campman, R Romijn, and W Van der Veen. Ambient noise multimode rayleigh and love wave tomography to determine the shear velocity structure above the groningen gas field. Geophysical Journal International, 218(3):1781–1795, 2019.
[7] Andrew Curtis, Jeannot Trampert, Roel Snieder, and Bernard Dost. Eurasian fundamental mode surface wave phase velocities and their relationship with tectonic structures. Journal of Geophysical Research: Solid Earth, 103(B11):26919–26947, 1998.
[8] Sjoerd de Ridder and Joe Dellinger. Ambient seismic noise Eikonal tomography for near-surface imaging at Valhall. The Leading Edge, 30(5):506–512, 2011.
[9] Simon Duane, Anthony D Kennedy, Brian J Pendleton, and Duncan Roweth. Hybrid monte carlo. Physics letters B, 195(2):216–222, 1987.
[10] A Dziewonski, S Bloch, and M Landisman. A technique for the analysis of transient seismic signals. Bulletin of the seismological Society of America, 59(1):427–444, 1969.
[11] Stephanie Earp, Andrew Curtis, Xin Zhang, and Fredrik Hansteen. Probabilistic neural network tomography across grane field (north sea) from surface wave dispersion data. Geophysical Journal International, 223(3):1741–1757, 2020.
[12] Andreas Fichtner, Andrea Zunino, and Lars G Aberad. Hamiltonian monte carlo solution of tomographic inverse problems. Geophysical Journal International, 216(2):1344–1363, 2018.
[13] Erica Galetti, Andrew Curtis, Giovanni Angelo Meles, and Brian Baptie. Uncertainty loops in travel-time tomography from nonlinear wave physics. Physical review letters, 114(14):148501, 2015.
[14] Erica Galetti, Andrew Curtis, Brian Baptie, David Jenkins, and Heather Nicolson. Transdimensional love-wave tomography of the British Isles and shear-velocity structure of the east Irish Sea Basin from ambient-noise interferometry. Geophysical Journal International, 208(1):36–58, 2017.
[15] Peter J Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, pages 711–732, 1995.

[16] W Keith Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[17] Eugene Herrin and Tom Goforth. Phase-matched filters: application to the study of Rayleigh waves. *Bulletin of the Seismological Society of America*, 67(5):1259–1275, 1977.

[18] S Hou, D Zheng, XG Miao, and RR Haacke. Multi-modal surface wave inversion and application to north sea obn data. In *78th EAGE Conference and Exhibition 2016*, volume 2016, pages 1–5. European Association of Geoscientists & Engineers, 2016.

[19] MW Lee and TS Collett. Integrated analysis of well logs and seismic data to estimate gas hydrate concentrations at keathley canyon, gulf of mexico. *Marine and Petroleum Geology*, 25(9):924–931, 2008.

[20] AL Levshin and MH Ritzwoller. Automated detection, extraction, and measurement of regional surface waves. In *Monitoring the Comprehensive Nuclear-Test-Ban Treaty: Surface Waves*, pages 1531–1545. Springer, 2001.

[21] Anatoli Levshin, Ludmila Ratnikova, and JON Berger. Peculiarities of surface-wave propagation across central Eurasia. *Bulletin of the Seismological Society of America*, 82(6):2464–2493, 1992.

[22] Anatoli L Levshin, VF Pisarenko, and GA Pogrebinsky. On a frequency-time analysis of oscillations. In *Annales de Geophysique*, volume 28, pages 211–218. Centre National de la Recherche Scientifique, 1972.

[23] Jing Li, Zongcai Feng, and Gerard Schuster. Wave-equation dispersion inversion. *Geophysical Journal International*, 208(3):1567–1578, 2017.

[24] Alberto Malinverno. Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, 151(3):675–688, 2002.

[25] Alberto Malinverno and Victoria A Briggs. Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes. *Geophysics*, 69(4):1005–1016, 2004.

[26] U Meier, A Curtis, and J Trampert. A global crustal model constrained by nonlinearised inversion of fundamental mode surface waves. *Geophysical Research Letters*, 34:L16304, 2007.

[27] Ueli Meier, Andrew Curtis, and Jeannot Trampert. Global crustal thickness from neural network inversion of surface wave data. *Geophysical Journal International*, 169(2):706–722, 2007.

[28] Nicholas Metropolis and Stanislaw Ulam. The Monte Carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949.

[29] A Mordret, M Landès, NM Shapiro, SC Singh, and P Roux. Ambient noise surface wave tomography to determine the shallow shear velocity structure at valhall: depth inversion with a neighbourhood algorithm. *Geophysical Journal International*, 198(3):1514–1525, 2014.

[30] Muhammad Atif Nawaz and Andrew Curtis. Variational Bayesian inversion (VBI) of quasi-localized seismic attributes for the spatial distribution of geological facies. *Geophysical Journal International*, 214(2):845–875, 2018.

[31] Amine Ourabah and Alistair Crosby. A 184 million traces per km2 seismic survey with nodes-acquisition and processing. In *90th SEG International Exposition and Annual Meeting*, pages 41–45. OnePetro, 2020.

[32] Choon B Park, Richard D Miller, and Jianghai Xia. Multichannel analysis of surface waves. *Geophysics*, 64(3):800–808, 1999.

[33] Choon Byong Park, Richard D Miller, and Jianghai Xia. Imaging dispersion curves of surface waves on multi-channel record. In *SEG Technical Program Expanded Abstracts 1998*, pages 1377–1380. Society of Exploration Geophysicists, 1998.

[34] Nicola Piana Agostinetti, Genny Giacomuzzi, and Alberto Malinverno. Local three-dimensional earthquake tomography by trans-dimensional Monte Carlo sampling. *Geophysical Journal International*, 201(3):1598–1617, 2015.

[35] George R Pickett. Acoustic character logs and their applications in formation evaluation. *Journal of Petroleum technology*, 15(06):659–667, 1963.

[36] Michael H Ritzwoller and Anatoli L Levshin. Eurasian surface wave tomography: Group velocities. *Journal of Geophysical Research: Solid Earth*, 103(B3):4839–4878, 1998.

[37] Gareth O Roberts, Richard L Tweedie, et al. Exponential convergence of langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.
Surface wave dispersion inversion using an energy likelihood function

[38] David R. Russell, Robert B. Herrmann, and Horng-Jye Hwang. Application of frequency variable filters to surface-wave amplitude analysis. *Bulletin of the Seismological Society of America*, 78(1):339–354, 1988.

[39] NM Shapiro and MH Ritzwoller. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle. *Geophysical Journal International*, 151(1):88–105, 2002.

[40] Weisen Shen, Michael H. Ritzwoller, Vera Schulte-Pelkum, and Fan-Chi Lin. Joint inversion of surface wave dispersion and receiver functions: a Bayesian Monte-Carlo approach. *Geophysical Journal International*, 192(2):807–836, 2012.

[41] Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann. Uncertainty quantification in imaging and automatic horizon tracking—a bayesian deep-prior based approach. In *SEG Technical Program Expanded Abstracts 2020*, pages 1636–1640. Society of Exploration Geophysicists, 2020.

[42] Frederik J. Simons, Rob D. Van Der Hilst, Jean-Paul Montagner, and Alet Zielhuis. Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle. *Geophysical Journal International*, 151(3):738–754, 2002.

[43] Jeannot Trampert and John H. Woodhouse. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds. *Geophysical Journal International*, 122(2):675–690, 1995.

[44] Jianghai Xia, Richard D. Miller, Choon B. Park, and Gang Tian. Inversion of high frequency surface waves with fundamental and higher modes. *Journal of Applied Geophysics*, 52(1):45–57, 2003.

[45] Huajian Yao, Robert D van Der Hilst, and Maarten V De Hoop. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—i. phase velocity maps. *Geophysical Journal International*, 166(2):732–744, 2006.

[46] Mallory K. Young, Nicholas Rawlinson, and Thomas Bodin. Transdimensional inversion of ambient seismic noise for 3D shear velocity structure of the Tasmanian crust. *Geophysics*, 78(3):WB49–WB62, 2013.

[47] Huaiyu Yuan and Thomas Bodin. A probabilistic shear wave velocity model of the crust in the central west Australian craton constrained by transdimensional inversion of ambient noise dispersion. *Tectonics*, 37(7):1994–2012, 2018.

[48] Xin Zhang and Andrew Curtis. Seismic tomography using variational inference methods. *Journal of Geophysical Research: Solid Earth*, 125(4):e2019JB018589, 2020.

[49] Xin Zhang and Andrew Curtis. Variational full-waveform inversion. *Geophysical Journal International*, 222(1):406–411, 2020.

[50] Xin Zhang and Andrew Curtis. Bayesian full-waveform inversion with realistic priors. *Geophysics*, 86(5):1–20, 2021.

[51] Xin Zhang and Andrew Curtis. Bayesian geophysical inversion using invertible neural networks. *Journal of Geophysical Research: Solid Earth*, page e2021JB022320, 2021.

[52] Xin Zhang, Andrew Curtis, Erica Galetti, and Sjoerd de Ridder. 3-D Monte Carlo surface wave tomography. *Geophysical Journal International*, 215(3):1644–1658, 2018.

[53] Xin Zhang, Fredrik Hansteen, Andrew Curtis, and Sjoerd de Ridder. 1D, 2D and 3D Monte Carlo ambient noise tomography using a dense passive seismic array installed on the north sea seabed. *Journal of Geophysical Research: Solid Earth*, 125(2):e2019JB018552, 2020. doi: 10.1029/2019JB018552.

[54] Xin Zhang, Corinna Roy, Andrew Curtis, Andy Nowacki, and Brian Baptie. Imaging the subsurface using induced seismicity and ambient noise: 3-D tomographic Monte Carlo joint inversion of earthquake body wave traveltimes and surface wave dispersion. *Geophysical Journal International*, 222(3):1639–1655, 05 2020. ISSN 0956-540X. doi: 10.1093/gji/ggaa230.

[55] Xuebin Zhao, Andrew Curtis, and Xin Zhang. Bayesian seismic tomography using normalizing flows. *Geophysical Journal International*, 228(1):213–239, 2021.

[56] DingChang Zheng, Erdinc Saygin, Phil Cummins, Zengxi Ge, Zhaoxu Min, Athanasius Cipta, and Runhai Yang. Transdimensional Bayesian seismic ambient noise tomography across SE Tibet. *Journal of Asian Earth Sciences*, 134:86–93, 2017.

[57] Alet Zielhuis and Guust Nolet. Deep seismic expression of an ancient plate boundary in Europe. *Science*, 265 (5168):79–81, 1994.
Surface wave dispersion inversion using an energy likelihood function

Figure 8: (a) The spectrum obtained using f-c analysis for the location in Figure 1b (red star). Black dots show the picked phase velocities. (b) The marginal posterior distribution of shear velocities obtained using only the picked phase velocities of the fundamental mode. The red line shows the mean shear velocity profile. The black dashed line and the white line in (a) show the predicted phase velocities for the fundamental mode and the first overtone respectively, modelled using the posterior mean model.

A mode validation

To understand the two branches that appear at short periods (<0.25 s) in the f-c spectrum in Figure 2 and in particular to test whether the two branches represent two different modes, we first perform an inversion using the rj-McMC algorithm using one of the branches as data (black dots in Figure 8). The prior pdf of the shear velocity is set to be a Uniform distribution between 300 m/s and 1500 m/s. For the likelihood function we use the traditional Gaussian distribution. The inversion is then conducted in the same way as described in section 3.4 with the same prior pdf for the number of layers and noise hyperparameters and the same proposal pdf. Figure 8b shows the obtained mean and the marginal distribution of the shear velocity. We then use the mean shear velocity profile to simulate phase velocities of the fundamental mode and the first overtone. While the modelled fundament mode phase velocity (black dashed line in Figure 8a) fits the data used, the modelled phase velocity of the first overtone (while line in Figure 8a) is significantly closer to the mode with velocity higher than 600 m/s than to the other branch appearing in the fundamental mode. This clearly demonstrates that the two branches are unlikely to represent two different modes, and instead represent an effect such as the multipathing of the seismic energy of the fundamental model or the strong lateral heterogeneity.