ASEG2024 HOBART

Towards Efficient Bayesian MASW Analysis using JAX

Jordan Dennis
Velseis
83 Jijaws St, Sumner Park, QLD
jdennis@velseis.com

Aaron Glanville
Velseis
83 Jijaws St, Sumner Park, QLD
aglanville@velseis.com

Shaun Strong
Velseis
83 Jijaws St, Sumner Park, QLD
sstrong@velseis.com

SUMMARY

Multichannel analysis of surface waves (MASW) offers a fast, low cost approach to characterising the shear-wave velocity profile of shallow sub-surface geology, emerging as a ubiquitous tool within a range of geotechnical applications. In this work, we explore the efficient evaluation of shallow sub-surface dispersion curves using the open source JAX library, and its optimisation for parallel systems. We also outline a framework for the Bayesian analysis of dispersion curves built around the JAX framework.

Key words: multichannel analysis of surface waves (MASW), dispersion analysis, shear wave velocity, bayesian statistics

INTRODUCTION

In most disciplines of exploratory geophysics, seismic signals from shallow sub-surface layering are typically considered a nuisance to be corrected for in our search for some deeper prize. Information on shallow geological compositions (most notably the stiffness of a given stratum and potential localised anomalies) are however of significant value across a range of geotechnical applications. Such data remains indispensable to the planning of safe and stable infrastructure. Methods to evaluate the stiffness profiles of shallow sub-surface layering can be broadly divided into direct borehole analyses, methods which directly measure soil penetration (such as the standard penetration test), and surface wave analysis methods such as MASW. Compared with direct penetration analyses, MASW offers a flexible and low cost inference methodology, however requires a deal of analytical care and computational complexity to recover accurate results.

Surface wave analysis techniques are broadly interested in recovering the shear wave velocity profile of a given region by constraining the dispersion curve (which relates the phase velocity and frequency of a travelling wave) recovered from seismic records. Traditionally the primary mode of this dispersion curve has represented the key observable of interest, however seismic datasets will naturally encode information beyond the first mode. A number of well-established techniques have been discussed in the literature to take advantage of the information contained within these higher order modes, most notably multi-modal inversion and full image inversion techniques (Forbriger, 2003a,b; Ryden et al., 2004). In addition to this shear velocity, the form of the dispersion curve is also modulated by parameters such as the number of strata, along with their associated thicknesses and densities. Higher order information from multi-modal analyses can also be used to improve constraints on the p-wave velocities, which may be of interest in characterising the make-up of strata.

Typically, measured dispersion curves are compared against a synthetic dispersion profile generated from some model geology (typically encoding information on the number of layers, along with p/s velocities, thicknesses, and densities for each layer). This is achieved using forward modelling, where a simulated seismic shot is propagated through an Earth model to generate a synthetic dispersion curve profile. The inversion approach has become ubiquitous across all disciplines in recent years. MASW presents an unique opportunity to apply this technique in exploration seismology because the Thompson-Haskell matrix (Haskell, 1953) method allows us to avoid a computationally expensive evaluation of the forwards model.

DISPERSION WAVE MODELS WITH JAX

To test our inversion methodology we aim to recover parameters from simulated data. By recovering known parameters we are able to test the sensitivity of our analysis to its priors. Inversions that require strict priors are less powerful than those that use relaxed priors. Suppose our inversion requires us to use ignorant priors over strict ranges for an assumed number of layers. We will need to perform a detailed meta-analysis to justify our choice of priors. Depending on precisely how constrained our priors must be, the meta-analysis may be sufficient by itself. Note that lower quality data usually require stricter priors.

We used SOFI2D (Bohlen, 2015) to simulate our earth models. In particular, we simulated visco-elastic models with between two and ten layers. Our models used absorbing boundary conditions with a single free surface. Our source was a 25Hz Ricker wavelet originating one grid-point below the free surface in the centre of the model. Similarly, we simulated geophones one grid-point below the free surface at every grid-point between the absorbing boundaries. Outside this basic design there were many hyper-parameters - e.g. time-step, grid-spacing, order, etc. - that we chose by trial and error to guarantee the stability of the simulation.

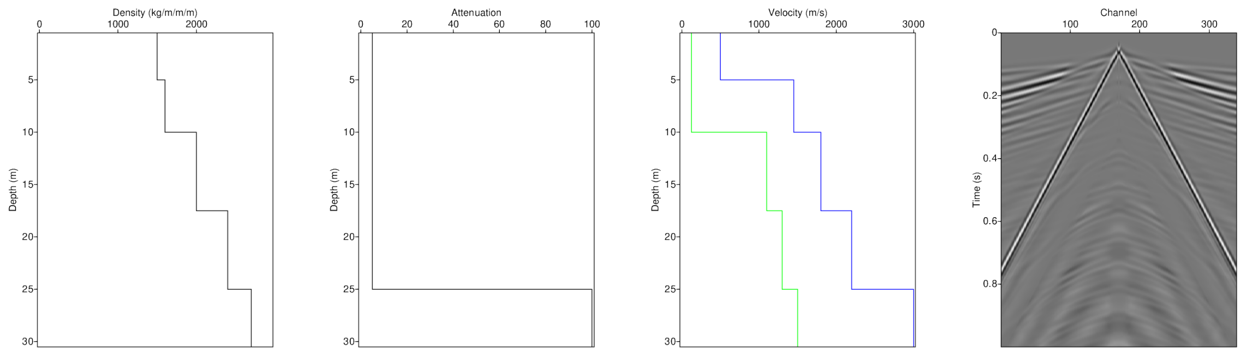

Figure 1. A model. In the panel labelled "Velocity" the blue curve represents the p-wave velocity and the green curve represents the s-wave velocity.

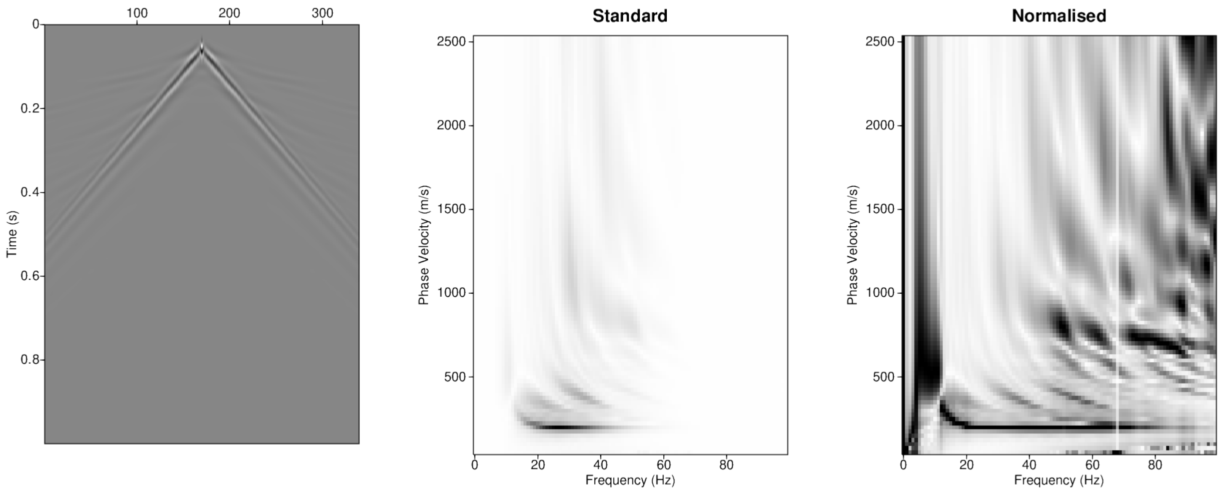

Figure 2. Our processing workflow. The panel labelled "Standard" represents the dispersion curve calculated without processing the shot record shown in the leftmost panel. On the other hand, the panel labelled "Normalised" represents the processed dispersion curve.

Using the template described in the previous paragraph, we explored how the dispersion curve related to the model. Our approach is similar to that of Mills et al. (2016). Our analysis used three steps: constructing the model, simulating the shot, and generating the dispersion curve. Each layer in our theoretical model was described using a tuple of five parameters: density, p-wave velocity, s-wave velocity, depth and attenuation. We fixed the attenuation parameter (Q) to be 10 in the weathering layer and 100 in the sub-weathering. Moreover, SOFI2D allows for both p and s-wave attenuation models, but we used the same attenuation model for both (see Figure 1).

To generate our models we used seismic unix - specifically suplane and suzero. These were formatted as blobs that were ingested by SOFI2D, which was then run to generate shot records in SU/XDR format. We transformed the shot record into a dispersion curve using the suphasevel command from seismic unix. We created dispersion curves up to a frequency of 80Hz (see Figure 2), to avoid aliasing. Furthermore, frequencies lower than 10Hz weren't typically useful because the curve broadened (see Figure 2).

We experimented with a variety of techniques - e.g. F-K filtering and mutes - to enhance the dispersion curves. However, we have currently settled on a system of normalisation that involves normalising each trace in the dispersion curve independently. We chose this approach because the Haskell output does not inform the relative amplitude of the modes, only their position (Park et al., 1999). Using our workflow we investigated how the Earth responded to perturbations in the model. Our early analysis showed that the dispersion curves were not very sensitive to the attenuation (Liner, 2012). Furthermore, we found that more thinner layers in a shallow weathering layer produced stronger ground roll and therefore less noisy dispersion curves.

We implemented the Haskell equations in Python, which we compiled and vectorized using JAX (Bradbury et al., 2018). The primary advantage of our implementation is that it will make good use of most hardware. As a result we get GPU acceleration and parallel processing for free just by writing Python code. Using our implementation we were able to evaluate the model of a 300 by 300 grid in the frequency phase-velocity domain in about 1s on a typical laptop. We are planning to do rigorous performance testing, including on GPUs, at a later date.

BAYESIAN ANALYSIS OF DISPERSION WAVES

The use of JAX to efficiently generate synthetic dispersion curves allows us to evaluate the feasibility of characterising shallow sub-surface properties via Bayesian statistical frameworks. The most notable advantage of a Bayesian approach for MASW (as opposed to alternative statistical approaches) is the explicit incorporation of prior information in the estimate of our probability distribution, allowing us to leverage basic geological information to substantially improve both the accuracy and convergence rate of our analysis. In addition, growing inter-disciplinary interest in Bayesian approaches within recent decades have yielded a wide array of open source tools for practical and efficient statistical analysis. These key advantages have generated significant discussion on the application of Bayesian frameworks within MASW inversion analysis, with a number of novel approaches being proposed in this space (Killingbeck et al., 2018; Olafsdottir et al., 2020; Aimar et al., 2024; Niu et al., 2024).

Despite these advantages, a number of clear challenges arise in the application of Bayesian principles to the analysis of shallow dispersion curves. Central to the confidence of Bayesian parameter estimates is the use of an accurate likelihood function, which relies on the generation of robust model data, and the accurate characterisation of both systematic and statistical errors associated with a set of observations. In the case of MASW, while the form of the dispersion curve represents the observable of interest, characterising this data vector and its associated errors remain ambiguous. We explore an approach in which the peaks of the Thomson-Haskell matrix represent the primary data vector of interest (as illustrated in Figure 3), with the statistical measurement uncertainty informed by the width of our peaks across both frequency and phase velocity slices. We will use the following sections to briefly outline the methodology behind our statistical approach, and potential avenues of application.

Building a Data Vector

The first step in our approach is to build a data vector (listing the locations of dispersion curve peaks within phase velocity space), to facilitate a robust comparison between observed dispersion curves and their modelled equivalents. Beginning from the raw form of a forward-modelled wave (generated using the Thomson-Haskell matrix), we apply a mild Gaussian filter to smooth and reduce fine discontinuities in this image. Taking constant frequency slices of this image, we extract local maxima and assign peak picks at these points subject to a minimum spacing condition. We then reject local maxima with a normalised amplitude below a pre-defined threshold to create our suite of high-confidence peak picks, as illustrated in Figure 3

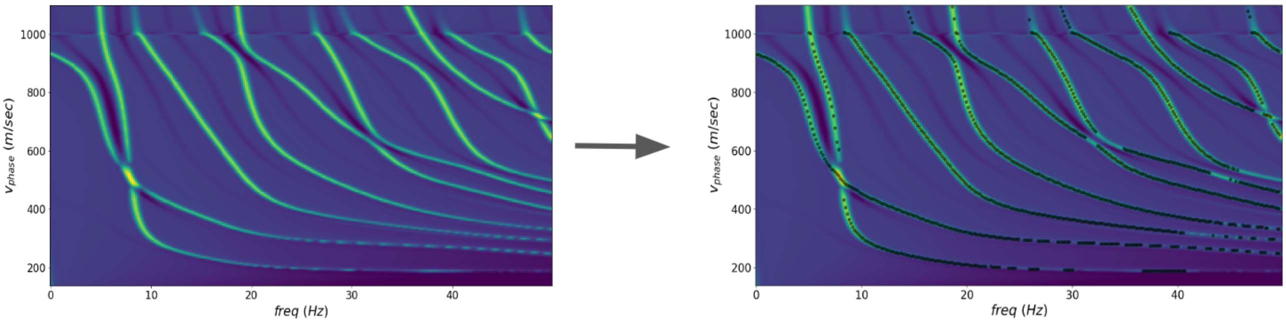

Figure 3. Extracting peaks from dispersion curves generated using the Thomson-Haskell matrix for a simple multi-layer system. The position of these dispersion curve peaks (in frequency/phase velocity space) represent the data vector of interest in our approach.

Loss function/error estimate

In order to meaningfully extract statistical information from these observations we must quantify their corresponding uncertainties, i.e. the expected variance in peak positions across successive observations. With only a single "true" observation to work with, the question of how to assign these uncertainties becomes highly non-trivial. In this work, we make the simplifying assumption that successive measurements of the dispersion curve peaks follow the form of the dispersion curve itself, i.e. that the statistical uncertainties of our simulated dataset are correlated with the width of the dispersion curve. It is important to note that this approach necessarily neglects the potential impact of systematic uncertainties (i.e. measurement errors that could offset the aggregate dispersion curve, and by extension all corresponding peaks through a bulk shift). This is almost certainly likely to represent a major component of measurement uncertainty in realistic observation, and requires further study to quantify. This allows us to independently assign uncertainties to the frequency and phase velocity measurements of a given peak pick, using the width of the Thomson-Haskell matrix across their corresponding array slice.

In order to assign independent uncertainties across both parameters in our matrix, we first estimate the "active width" of slices along both axes for each peak pick (i.e, the region which encompasses the wings of our 1D distribution before asymptotically approaching some baseline amplitude). This is done using a slice-wise approach, placing the extents of our active width at indices where the array-wise difference crosses a tunable difference threshold. This process is repeated across both v_{phase} and f slices for all picked peaks, independently assigning their active widths. Finally, the 1D distribution corresponding to these active slices is fit to a 1D Gaussian model, with the resulting standard deviations used as our statistical uncertainties, to be combined in quadrature for joint constraints.

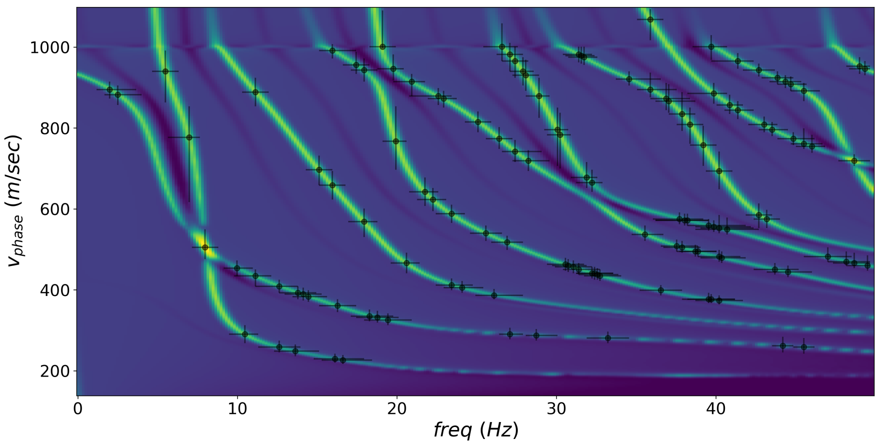

Figure 4. An example of the active widths extracted for a series of randomly subsampled peak picks from our synthetic dispersion curve. These widths are not used to allocate uncertainties directly, but the amplitude curve within these slices is used to estimate an uncertainty via comparison with the Gaussian distribution. Note that in practice, these uncertainties should be assigned exclusively to observed dispersion curves (i.e. the distribution of Figure 2), with this example on our Thomson-Haskell model remaining purely illustrative.

It should be noted this proposed likelihood methodology builds on similar work conducted to develop a Bayesian framework for MASW inversion (in particular Killingbeck et al. (2018)), with a few caveats. Notably, our method of assigning uncertainties has some subtle differences, with Killingbeck et al. (2018) assigning uncertainties to only the phase velocity using the half-width of the image, whereas we assign uncertainties to both phase velocity and frequency via a direct Gaussian fit. Most importantly however, our efforts principally deviate in the method used to generate synthetic spectra, and in our proposed applications (as discussed in the following section).

Bayesian Model Analysis

With our data vector and errors defined, we may now turn our attention to constraining physical parameters from observed dispersion curves. A standard application of Bayes theorem is to estimate the posterior probability distribution, giving constraints on the parameters underpinning a model (in this case, v_{phase} and f), in light of some observation (the form of the dispersion curve), and modulated by prior physical information on the system. This posterior distribution is typically evaluated using a numerical sampling routine (such as the Monte Carlo Markov Chain, or MCMC), with a number of well-established open-source tools available to achieve this (Hoffman and Gelman, 2011; Foreman-Mackey et al., 2013; Salvatier et al., 2015). While MASW inversion techniques are typically restricted to identifying the model best fit, we aim to use the improved model efficiency of JAX in combination with our likelihood model to compliment optimisation searches with a complete estimate of the full posterior distribution associated with a given model. This posterior distribution naturally encodes the uncertainty assigned to constrained parameters like shear velocity, shedding light on potentially overlooked nuances in complex sub-surface systems.

Beyond standard parameter estimation, an often overlooked application of Bayes theorem is to compare different models by quantifying how well they characterise some set of observations, through measurement of the Bayesian evidence ratio. In the case of MASW, where a priori assumptions about the nature of some geological system (e.g. the number of layers) can significantly constrain the space of viable solutions, Bayesian model comparison can play a clear role in supporting data-driven analysis. Measurements of the Bayesian evidence ratio typically rely on evaluating the integral of the likelihood function over the posterior distribution, requiring the use of a nested sampling algorithim. A number of robust, open-source nested sampling algorithms have recently gained attention within the broader scientific literature (Skilling, 2004; Handley et al., 2015; Speagle, 2020), including at least one recent library which is JAX compatible (Albert, 2020). Open-source implementation of a robust dispersion wave likelihood within an efficiently distributed Nested Sampling routine represents a potentially valuable future application of our work, and is a clear direction for future study.

DISCUSSION AND CONCLUSIONS

MASW analyses remain the gold standard for determining the subsurface structure in civil applications. In this work, we explore a potential framework for the efficient analysis of MASW data, complementing new approaches in the geotechnical literature with interdisciplinary advances in Bayesian methodologies. Notably, by using JAX as a back-end we enjoy native performance and GPU acceleration from a high level language. The recent development of several open source python libraries for Bayesian analysis allows us to leverage this design to explore novel techniques in Bayesian model analysis (in particular the evidence ratio for data driven model comparison).

REFERENCES

Aimar, M., Foti, S., and Cox, B. R., 2024, Novel techniques for in situ estimation of shear-wave velocity and damping ratio through MASW testing part II: a Monte Carlo algorithm for the joint inversion of phase velocity and phase attenuation: Geophysical Journal International, **237**, no. 1, 525–539.

Albert, J. G., 2020, JAXNS: a high-performance nested sampling package based on JAX: arXiv e-prints, page arXiv:2012.15286.

Bohlen, T. Sofi2d:, 2015.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., and Zhang, Q. JAX: composable transformations of Python+NumPy programs:, 2018.

Forbriger, T., 06 2003, Inversion of shallow-seismic wavefields: I. Wavefield transformation: Geophysical Journal International, **153**, no. 3, 719–734.

——— 06 2003, Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms: Geophysical Journal International, **153**, no. 3, 735–752.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J., 2013, emcee: The MCMC Hammer: , **125**, no. 925, 306.

Handley, W. J., Hobson, M. P., and Lasenby, A. N., 2015, POLYCHORD: next-generation nested sampling: , **453**, no. 4, 4384–4398.

Haskell, N. A., 1953, The dispersion of surface waves on multilayered media.: Bulletin of the seismological society of america, **43**, 17–34.

Hoffman, M. D., and Gelman, A., 2011, The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo: arXiv e-prints, page arXiv:1111.4246.

Killingbeck, S., Livermore, P., Booth, A., and West, L., 11 2018, Multimodal layered transdimensional inversion of seismic dispersion curves with depth constraints: Geochemistry, Geophysics, Geosystems, **19**.

Liner, C. L., 2012, Elements of seismic dispersion: A somewhat practical guide to frequency dependent phenomena.: Society of exploration geophysics.

Mills, A., Cova, R., and Innanen, K., 2016, Surface wave modelling and near surface characterization background: CREWES Research Report, **28**.

Niu, Y., Fang, G., Li, Y. E., Chian, S. C., and Nilot, E., 2024, Multimode Rayleigh wave dispersion spectrum inversion using Wasserstein distance coupled with Bayesian optimization: Geophysics, **89**, no. 3, KS83–KS94.

Olafsdottir, E. A., Erlingsson, S., and Bessason, B., 2020, Open-source masw inversion tool aimed at shear wave velocity profiling for soil site explorations: Geosciences, **10**, no. 8.

Park, C., Miller, R., and Xia, J., 1999, Multichannel analysis of surface waves: Geophysics, **64**.

Ryden, N., Park, C., Ulriksen, P., and Miller, R., 06 2004, Multimodal approach to seismic pavement testing: Journal of Geotechnical and Geoenvironmental Engineering - J GEOTECH GEOENVIRON ENG, **130**.

Salvatier, J., Wiecki, T., and Fonnesbeck, C., 2015, Probabilistic Programming in Python using PyMC: arXiv e-prints, page arXiv:1507.08050.

Skilling, J., 2004, Nested Sampling Nested Sampling, AIP, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, 395–405.

Speagle, J. S., 2020, DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences: , **493**, no. 3, 3132–3158.