Geophysical Journal International

Geophys. J. Int. (2026) **244,** 1–34
Advance Access publication 2025 December 8
Research Paper

https://doi.org/10.1093/gji/ggaf498

Full wavefield surface wave analysis with Bayesian Evidential Learning

Anne-Sophie Mreyen, Hadrien Michel and Frédéric Nguyen

Urban and Environmental Engineering, University of Liege, 4000 *Liege, Belgium. E-mail: as.mreyen@uliege.be*

Accepted 2025 December 2. Received 2025 November 18; in original form 2025 June 11

SUMMARY

Surface waves such as Rayleigh, Love and Scholte waves can exhibit dispersion, that is, variations in phase velocity with wavelength as a function of frequency. This property enables the inversion of 1-D models of seismic velocity and density in the subsurface. Conventional deterministic and stochastic inversion schemes are widely applied to surface wave data but face two main challenges. The first is the identification of dispersion curves for fundamental and higher modes on wavefield-transformed images, which is often done manually. The second is the quantification of uncertainty, which can be computationally expensive in stochastic approaches or limited to data-propagated uncertainty in deterministic inversions. Our objectives are to (1) eliminate the need for manual or automatic dispersion curve picking, and (2) directly infer ensembles of 1-D velocity models—and their associated uncertainties—from the full velocity spectrum, that is, the complete dispersion image containing all modes. To this end, we employ Bayesian Evidential Learning, a predictive framework that reproduces experimental data from prior information while allowing prior falsification. In our application, ensembles of prior Earth models are sampled to predict 1-D subsurface structures in terms of seismic velocity and, where applicable, attenuation from near-surface seismic wave data. This approach bypasses traditional inversion schemes and provides a computationally efficient tool for uncertainty quantification.

Key words: Machine learning; Seismic attenuation; Surface waves and free oscillations; Interface waves; Waveform inversion; Machine learning.

1 INTRODUCTION

Interface seismic waves such as surface waves (Rayleigh and Love at the air–earth interface) or Scholte waves (at the water-earth interface) can be dispersive and can be inverted to obtain 1-D models of seismic *S*-wave velocity, *P*-wave velocity and density as a function of depth (Xia *et al.* 1999; Herrmann 2013). The conventional workflow first computes a dispersion image from seismic data and then extracts dispersion curves, which are subsequently inverted using deterministic or stochastic schemes (e.g. Herrmann & Ammon 2004; Wathelet *et al.* 2020). This approach faces two main challenges: (1) reliable identification of fundamental and higher modes in wavefield-transformed images (e.g. f–k), often requiring manual picking, and (2) quantification of uncertainty, which is either computationally demanding in stochastic frameworks or limited to propagating data errors in deterministic inversions. Dispersion curve picking therefore depends both on the chosen wavefield transform and on subjective interpretation of the resulting images. Socco *et al.* (2009) use the simultaneous extraction of multiple dispersion curves on seismic reflection ground roll with a search for spectral maxima using moving spatial windows from f–k (i.e., frequency–wavenumber) transformed common-shot-point gathers. While this latter can be considered as a pioneer work for automatized dispersion curve extraction, it necessitates a prior manual identification on representative f–k spectra to identify spectral regions that can be used for the dispersion curve search in the automatized process. A stochastic approach for near-surface modelling based on multimodal dispersion curves is introduced by Maraschini & Foti (2010) that minimizes the determinant for Monte Carlo generated dispersion curves, fitting the model given Haskell–Thomson matrix determinant being zero for the f–V (i.e., frequency–velocity) pairs of the experimental model curves. Even though computational efforts can be drastically minimized with the approach, the need for manual dispersion curve picking is still present and mode identification stays impossible in the case of modal superposition.

Automatic picking of dispersion curves by identifying the highest energy in the dispersion spectrum (non-machine learning) has been performed by various authors. Liu *et al.* (2024) suggest a methodology for automatic adaptive picking of dispersion curves that use a low- and high-frequency search strategy by identifying amplitude

minima and maxima in the phase velocity spectrum for each frequency. Miao *et al.* (2016) identify local maxima in Rayleigh and Scholte wave dispersion spectra assuring coherent mode order in the inversion procedure using the Eigen-determinant misfit function from Maraschini & Foti (2010). Zheng & Miao (2014) were able to extract multimodal dispersion curves by semi-automated picking after binerization and thinning of the f–k spectrum. Taipodia *et al.* (2020) applied threshold energy filtering on Surfseis (Park & Brohammer 2003) generated dispersion images to extract the fundamental mode dispersion curve of active land seismic surveys. Similarly, staging and algorithmic approaches have been proposed to handle higher-mode misidentification and mode aliasing, for example by using a combination of pattern search optimization and graph-theoretical assignment strategies (e.g. frequency-Bessel transform with Kuhn–Munkres algorithm; Yan *et al.* 2022).

Recent studies have advanced the field by introducing machine learning techniques for automatic dispersion curve picking. Wang *et al.* (2021) automatize the picking process of multimodal dispersion curves with the help of unsupervised machine learning, that is, the clustering methods Gaussian Mixture Model and density-based spatial clustering of applications with noise (DBSCAN); their workflow is also implemented by Xu *et al.* (2024). Similarly, unsupervised techniques are used by Rovetta *et al.* (2021) with DBSCAN, as well as by Masclet *et al.* (2019) and Yao *et al.* (2023) with K-means clustering. In contrast to unsupervised learning, supervised learning requires a training phase on large amounts of sample data. Such supervised deep learning approaches to pick multimodal dispersion curves are presented by Dai *et al.* (2021) and Alyousuf *et al.* (2018). Dong *et al.* (2021) developed a supervised neural network system (DisperNet) combined with an unsupervised dispersion curve classification; furthermore, a variety of authors implement convolutional neural networks (U-Net) to extract dispersion curves from dispersion images, such as Zhang *et al.* (2020a), Song *et al.* (2021, 2022), Dai *et al.* (2022), Yang *et al.* (2022a, b), Ren *et al.* (2023), Gan *et al.* (2024) and Cárdenas *et al.* (2025). The study of Chamorro *et al.* (2023) introduces a deep-learning approach that circumvents the need to compute dispersion images and directly predicts dispersion curves from seismic shot gathers. On a probabilistic basis, Ortega *et al.* (2019) use Bayesian inference and Markov chain Monte Carlo (McMC) sampling to automatically obtain a range of probable dispersion curve solutions given a prior model.

The approach to bypass dispersion curve picking as such by using the full phase velocity spectrum, alias dispersion image or dispersion spectrogram, is first presented by Forbriger (2003a, b) with a gradient-based inversion scheme; Ryden & Park (2006) and Dou & Ajo-Franklin (2014) followed with inversion procedures based on global and hybrid global-local search algorithms, respectively. Other deterministic methodologies using the entire dispersion spectrogram for surface wave inversion can be found in the studies of Dal Moro (2014, 2019), Dal Moro *et al.* (2015), Niu *et al.* (2022), Zhang *et al.* (2020b, 2024) and Liu *et al.* (2023). Cho *et al.* (2024) introduce a deep-learning algorithm in form of a convolutional neural network to directly invert from the dispersion spectrum for *S*-wave velocity by bypassing the definition of prior information and dispersion curve picking; the method does not allow to incorporate further information such as *P*-wave velocity. A first probabilistic framework using the full velocity spectrum by defining an energy likelihood function for Bayesian inference is presented by Zhang *et al.* (2023). The drawback of the latter study is that, while it allows to estimate uncertainty in a Bayesian framework, the posterior solution is sampled with a Markov chain Monte Carlo approach that may be computationally expensive.

Bayesian Evidential Learning (BEL; cf., Hermans *et al.* 2018; Scheidt *et al.* 2018) has recently been proposed to (hydro-)geological studies (Hermans *et al.* 2019; Yin *et al.* 2020; Tadjer & Bratvold 2021; Thibaut *et al.* 2021; Yang *et al.* 2023) as well as for different geophysical applications (Hermans *et al.* 2018; Michel *et al.* 2020; Pradhan & Mukerji 2020; Hermans *et al.* 2021; Michel *et al.* 2023; Ahmed *et al.* 2024; Aigner *et al.* 2025), with Michel *et al.* (2023) introducing probabilistic dispersion curve prediction through iterative prior resampling for improved posterior distributions. Unlike conventional deterministic or stochastic inversion approaches, BEL does not require explicit optimization of a misfit function. Instead, it learns statistical relationships between model parameters (e.g. layered Earth models) and data (e.g. velocity spectra) from a set of prior simulations. Once trained, this relationship allows new data to be rapidly mapped to a posterior distribution of possible models while providing an explicit measure of uncertainty. Conventional surface wave inversion approaches typically reduce the wavefield to manually or automatically picked dispersion curves, which discards a large amount of information. In contrast, in this work we adopt BEL in a 1-D framework (BEL1D) that exploits the full wavefield information contained in wavefield transformed, multimodal velocity spectrum. By using the entire dispersion image rather than extracted curves, the method directly predicts ensembles of 1-D near-surface profiles (velocity, density, attenuation) with quantified uncertainty similar to a stochastic inversion, while avoiding the need for dispersion curve picking and handling inverse problems. It relies on a simulation-based prior that through statistical relationships identifies the link to the data set of interest, for example, a field data set collected with a linear array of seismic receivers. Posterior sampling is computationally efficient once the prior is established, and the approach can naturally handle noise and prior falsification. This allows flexible application to both elastic and anelastic problems. All computations were performed on a workstation running Windows 10 with an AMD Ryzen CPU (16 physical cores, 32 logical cores) and 128 GB of RAM.

The paper is organized as follows: We first introduce the Bayesian Evidential Learning framework for surface wave modelling using the full velocity spectrum, illustrated by a workflow scheme. We then apply the methodology to synthetic benchmark models in the elastic domain, including noise-contaminated cases, and extend the tests to the anelastic domain to examine the influence of attenuation parameters. Applications to two real data sets from the literature follow, with attenuation considered in the second example. The paper concludes with a discussion and summary of the main findings.

2 BAYESIAN EVIDENTIAL LEARNING

The BEL1D framework has originally been developed by Michel *et al.* (2020) for low-dimensional problems (a few tens of parameters) and is available as the open-source package pyBEL1D (version 1.0.1; Michel *et al.* 2022). BEL is based on learning statistical relationships between (i) realistic models m—in this study, layered Earth models characterized by elastic and anelastic parameters, which define the prior, and (ii) synthetic data d generated from these models—here, the full velocity spectrum representing surface wave dispersion. Once this relationship is learned, BEL can predict an ensemble of Earth models (the posterior) from experimental data d^*, such as a dispersion image from a seismic shot gather. Contrary to classical inversion schemes, BEL does not try to match exactly

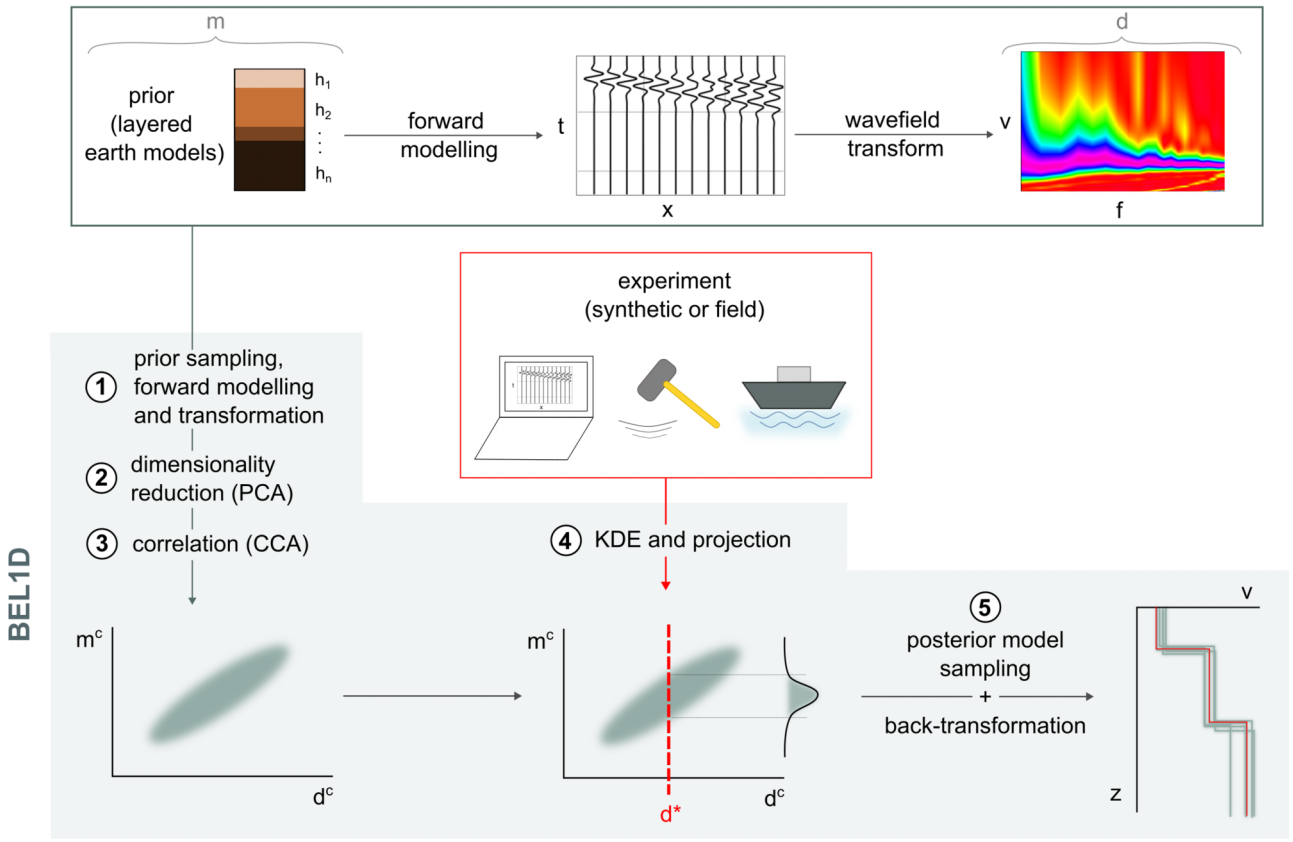

Figure 1. Workflow of BEL1D surface wave analysis.

the field data, but rather builds a set of posterior models that statistically relates to the data. The methodology is summarized in form of a workflow scheme in Fig. 1 and can be subdivided in five key steps: the prior sampling and modelling (1), the central learning steps using statistical relationships (2) and (3), the inference with experimental data (4) and the posterior sampling and modelling (5). In the following, each step will be explained in detail; complementary information can be found in Scheidt *et al.* (2018) and Michel *et al.* (2020).

2.1 Prior sampling and modelling

The first step in a Bayesian approach is the construction of a prior parameter space (step 1 in Fig. 1). This prior encodes the range of parameter values considered plausible before incorporating observational data. In our context, it defines a subsurface model space for 1-D surface wave modelling. Traditionally, the prior includes possible ranges of layer thickness h, S-wave velocity V_s, P-wave velocity V_p and bulk density ρ. From these distributions (often uniform), a large number of layered Earth models are sampled to form the model space m. In BEL, the adequacy of the prior is critical: it not only provides the model space for forward simulations but also allows for prior falsification, that is, testing whether the chosen parameter ranges are consistent with the experimental data and refining them if necessary. Through forward modelling (upper frame in Fig. 1), each sampled model generates an associated data realization d, namely a dispersion image that represents the velocity spectrum of the Earth model. Using the reflectivity method, we first compute synthetic seismograms in the offset–time (x–t) domain, which are then transformed into dispersion images in the frequency–velocity (f–V) domain using standard wavefield transforms. The implementation of these steps is described in detail below.

2.1.1 1-D synthetic seismograms: reflectivity method

The creation of synthetic seismograms is the intermediate step linking Earth models from the prior to their associated velocity spectra through forward modelling. For this, we use the enhanced reflectivity method implemented in the open-source Python package COMPOSTI (Niskanen & Lähivaara 2023). The algorithm incorporates prior-defined parameters and allows the inclusion of quality factors Q_p and Q_s for compressional and shear waves. Note, parameters can only vary vertically due to the 1-D restriction. The reflectivity method builds on the Thomson–Haskell propagator matrix formulation in the frequency–wavenumber (f–k) domain (Thomson 1950; Haskell 1953), later extended by Fuchs & Müller (1971), Kennett (1983) and others (Müller 1985). COMPOSTI evaluates highly oscillatory integrals with a Levin integrand (Levin 1982, 1996) and applies frequency-dependent sampling, which ensures stable solutions even at large offsets and across wide frequency ranges. A Ricker wavelet is used as the point source in the frequency domain, generating a band-limited, zero-phase pulse with a specified peak frequency, which drives the synthetic wavefield. The reflectivity method models a full 3-D wavefield generated by a point source by reducing the elastic wave equation to 1-D in cylindrical coordinates. It fully represents reflections, transmissions, attenuation and mode conversions of plane waves at layer interfaces, reproducing the complete wavefield (P, SV and SH), either coupled or decoupled (Ma *et al.* 2004; Sen 2020; Niskanen & Lähivaara 2023). This is particularly important for surface wave simulations, since Rayleigh waves are coupled P–SV interface waves. They can be modelled

over well-defined slowness ranges by summing near- and far-field internal multiples (O'Neill *et al.* 2003). In this study we only consider the vertical component of particle velocity at the receivers, as it contains the dominant Rayleigh-wave energy and is the standard choice for dispersion analysis. Love waves are therefore not explicitly analysed. Receiver response functions are not included, and synthetic traces are assumed to represent ground motion directly. Finally, the synthetic wavefield is transformed back to the time–offset (t–x) domain by Fourier transform, producing synthetic shot gathers that enhance the identification of phase velocity dispersion.

2.1.2 Wavefield transform

For both synthetic and experimental data, the wavefield transform converts the shot gather from the time–offset (t–x) domain to the frequency–velocity (f–V) domain, producing the dispersion image. We implement this step using the open-source library SW process for surface wave processing (Vantassel & Cox 2022). The package primarily employs a frequency-domain beamformer (FDBF) transform, which can be adjusted through the choice of steering vector and amplitude weighting. The FDBF with plane-wave steering and no amplitude weighting (cf. Zywicki 1999) is equivalent to the classical f–k transform, but avoids the intermediate wavenumber calculation and thereby reduces geometric limitations and aliasing effects (Vantassel & Cox 2022). Alternatively, the FDBF with a cylindrical-wave steering vector and square-root distance weighting (cf. Zywicki 1999; Zywicki & Rix 2005) best approximates the propagation of Rayleigh waves, capturing both their spreading geometry and amplitude decay.

In this study, we use the classical f–k transform for synthetic tests, while the real data applications employ the FDBF with cylindrical-wave steering, square-root distance weighting and absolute maximum normalization.

2.2 Training procedure

The training procedure forms the core of BEL1D, linking the model and data spaces through statistical relationships, with the observed data constraining posterior predictions.

2.2.1 Principal component analysis (PCA)

After the data space creation through froward modelling, we apply data dimensionality reduction using principal component analysis (PCA, step 2 in Fig. 1, implemented via the Scikit-learn Python library; Pedregosa *et al.* 2011). PCA transforms the original data into a set of orthogonal components, resulting in PCA scores ranked by the variance they explain (Tipping & Bishop 1999; Jolliffe 2011). The first components contain most of the variability in the data set, while higher-order components capture less information. These higher-order components can therefore be discarded, reducing memory usage and improving computational efficiency.

This reduction is particularly relevant in geophysical imaging problems, where the data space can be extremely high-dimensional. In our application, the model space m remains low-dimensional and physically interpretable; for example, a 4-layer model that integrates h, V_p and V_s has 11 free parameters. By contrast, the associated data space d may reach several hundred thousand dimensions, depending on the resolution of the velocity spectrum in the f–V domain. We therefore apply PCA only to the data space d, becoming the reduced data space d^f.

2.2.2 Canonical correlation analysis (CCA)

The core learning step in BEL1D is the canonical correlation analysis (CCA, step 3 in Fig. 1), implemented via the Scikit-learn Python library (Pedregosa *et al.* 2011). CCA identifies directions that maximize correlations between the reduced data space d^f and the model space m^f (here equal to m, since the model space is not reduced). Transformation matrices are then used to construct a joint correlated domain, yielding the CCA data space d^c and the CCA model space m^c, which are orthogonal to each other:

$$d^c = d^f A^T, \tag{1}$$

$$m^c = m^f B^T. \tag{2}$$

Here, A has dimensions $n \times k$ and B has dimensions $n \times q$, with n defined as the minimum of k and q. As a result, d^c and m^c both have dimensions $1 \times n$. Note, in order to allow for a back-transformation from the CCA to the original space, k must exceed q.

In essence, CCA provides a maximally correlated representation of the data and model spaces, highlighting the combinations of parameters and data features that most strongly relate to each other. This correlated space forms the foundation for posterior predictions in the BEL1D framework.

2.3 Posterior sampling and modelling

The experimental data d^* (central frame in Fig. 1), from which we aim to determine the S-wave velocity profile, are first transformed and projected into the bivariate and correlated CCA d–m space. This projection allows us to check whether the data lie within the prior, or if the prior needs to be adjusted. A pre-requisite for computing a posterior solution is that d^* falls within the prior; otherwise, the prior can be falsified.

In BEL1D, the posterior distribution is generated directly in the CCA space by constraining the model space m^c to the projected data d^*. This is achieved using kernel density estimation (KDE) with a Gaussian kernel (cf. Wand & Jones 1993; Michel *et al.* 2020), which approximates the conditional probability of the model parameters given the data. Posterior models are then sampled from this distribution via inverse transform sampling (Devroye 2006) and back-transformed to the original parameter space (assuming no PCA reduction on the model space):

$$m_{\text{post}} = m^f_{\text{post}} = B^{-1} m^c_{\text{post}}. \tag{3}$$

The resulting m_{post} can be visualized in the parameter space, for example showing the variation of shear wave velocity with depth. The precision and accuracy of the obtained posterior models have been benchmarked against a classical McMC approach by Michel *et al.* (2023), demonstrating that the obtained distributions reasonably estimate the benchmark method results. However, using the same forward modeller as for the prior, the posterior models are converted into velocity spectra, allowing evaluation of the fit to the experimental data d^*. It is important to note that this step is optional: the sampled models inherently represent the posterior uncertainty, and forward simulations are only required if one wishes to quantitatively assess the fit to the experimental data.

To quantify the misfit between experimental and posterior-predicted data d_{post}, we use a root-mean-square error (RMSE, eq. 4; cf. Hodson 2022) calculated on the normalized wavefield power of the dispersion images. For this, each sampled posterior model is first passed through the forward modelling algorithm to generate its corresponding velocity spectrum. The RMSE is then computed per

Table 1. Synthetic benchmarks models after Tokimatsu *et al.* (1992). h = thickness (m); V_s = *S*-wave velocity (m s^{-1}); V_p = *P*-wave velocity (m s^{-1}); ρ = density (kg m^{-3}).

Layer	h	V_s			V_p	ρ
		Model 1	Model 2	Model 3		
1	2	80	180	80	300	1800
2	4	120	120	180	1000	1800
3	8	180	180	120	1400	1800
4	–	360	360	360	1400	1800

pixel by comparing the normalized wavefield power of the posterior and experimental dispersion images:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i^* - d_{\mathrm{post},i})^2}, \tag{4}$$

where n is the total number of pixels in the dispersion image, d_i^* and $d_{\mathrm{post,i}}$ are the normalized wavefield power values of the experimental and predicted data at pixel i, respectively. This image-based RMSE provides a quantitative measure of how well the posterior models reproduce the power distribution of the experimental surface wave dispersion. Note that in the uncertainty figures based on RMSE (model ensembles), values are displayed on a linear numerical scale, while colours are assigned by quantiles, providing an illustration of the models distributions in the misfit space.

3 SYNTHETIC BENCHMARK

3.1 Elastic benchmark models

In the following, we introduce 1-D synthetic elastic benchmark models from Tokimatsu *et al.* (1992) (cf. O'Neill *et al.* 2003), which are tested with the BEL1D approach for shallow surface wave modelling. The three models (Table 1) consist of three layers over an infinite half-space, parametrized by layer thickness h, shear-wave velocity V_s, compressional-wave velocity V_p and density ρ. Model 1 can be characterized as normally dispersive, that is, *S*-wave velocity increasing with depth (Fig. 2a), whereas the other two are irregularly dispersive with a velocity inversion at the second layer for model 2 (Fig. 2d) and a velocity inversion at the third layer for model 3 (Fig. 2g). These configurations capture a range of modal behaviours, including higher modes and mode jumping.

In Figs 2(b), (e) and (h), we show the synthetic seismograms of models 1–3 generated with the reflectivity method composed by 48 receivers (spacing of 1 m) for an acquisition time of 1 s. The seismic source was simulated with an 80 Hz Ricker wavelet. Corresponding velocity spectra were obtained via f–k wavefield transformation, while theoretical modal curves (dashed lines) were computed with Dunkin's matrix (Dunkin 1965) as implemented in the Python library `disba` (Luu 2024). For the normally dispersive model 1 (Fig. 2c), the velocity spectrum shows a distinctive fundamental mode across the frequency range, while the theoretical modal curves reveal additional higher modes. In model 2, the velocity inversion in the second layer produces a fundamental mode with multiple overtones (Fig. 2f), leading to mode jumping. The mode distribution of model 3 (Fig. 2i) is even more complex: the spectrum suggests an apparent interruption of the fundamental mode at 10 Hz, but the Dunkin modal curves clarify that the first mode continues throughout the spectrum and that the energy between 10–18 Hz and 150–350 m s^{-1} belongs to the first and second overtones. In the following, we illustrate the workflow using model 2; results for models 1 and 3 are provided in the Appendix Figs A9–A12.

3.1.1 *Prior and dimensionalities*

In a first step, we define a common prior model space m to all three models 1–3 with a uniform distribution, designed to encompass the benchmark models (see Table 2 and red curves in Figs 2a, d and g). To limit the model space dimensionality to 11 parameters, we decided to fix the bulk density ρ to the true value of 1800 kg m^{-3} due to its low sensitivity in dispersion curve modelling (cf. Xia *et al.* 2003; Pan *et al.* 2019). From this prior, 1000 models were sampled to generate the corresponding full velocity spectra with an averaged computational time of 5.18 s per model. The generated spectra constitute the data space in the f–V domain, that is, pairs of frequency–velocity data points, that form the dispersion image of each sampled model.

To further lower the complexity of the problem, the high dimensionality of the created data space is reduced with the principal components analysis (PCA). For model 2, the dispersion-image data set (correlated f–V points) is reduced from 12 100 to 602 independent principal components (PCs), while still explaining 99 per cent of the variability (Fig. 3a). The projected experimental data (red points) in the lower PC dimensions thereby provide information on the quality of data projection, that is, few outliers in comparison to the modelled data (black points) show rather good conformity. Note, due to the small dimensionality of the model space, we exclusively apply PCA to the data space (which must remain larger than or equal to the model space to allow for back-transformation).

The parameter contribution to the 11 dimensions in the canonically correlated space for the models is shown in Fig. 3(b). Because the training is based on the same prior, the CCA space is identical for all three models. Most parameters are distributed across several dimensions, although some dominate individually (e.g. h_1 in dimension 4 or V_{s1} in dimension 7). Data-model correlations within the CCA space are illustrated by scatter plots for selected dimensions in Figs 3(c)–(f) (see Appendix Fig. A6 for all dimensions). It is visible that the higher the dimension, the less the data and model space are correlated.

3.1.2 *Data projection and posterior*

The trained and correlated bivariate space is then used to project the experimental data d^*, that is, the benchmark models of Table 1. The data are transformed into the reduced CCA data space for each dimension, illustrated by the red lines in Figs 3(c)–(h). An example of a data projection outside the prior is illustrated in Appendix Fig. A5. KDE is applied in this reduced space, and the resulting probability distributions are displayed left of each data-model correlation plot. From these distributions, a defined number of posterior models is sampled, here again 1000, and back-transformed to the original data space (see eq. 3). The ensemble of posterior solutions for V_s and V_p, coloured by the normalized RMS misfit of the dispersion images, is shown in Fig. 3(i). The best-fit posterior model (Fig. 3j) reproduces the full velocity spectrum with high fidelity and

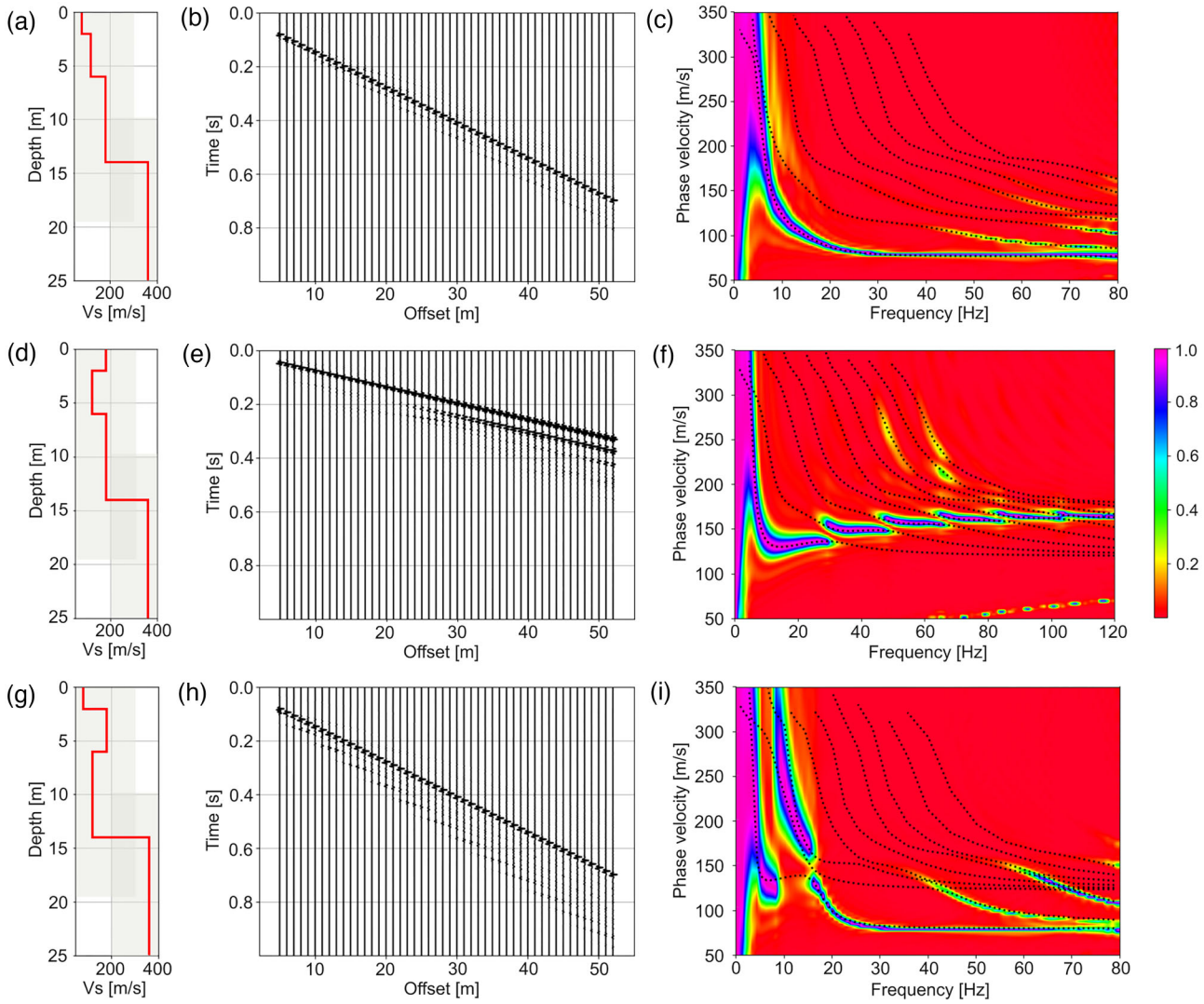

Figure 2. Synthetic benchmark models from Tokimatsu *et al.* (1992) (cf. O'Neill *et al.* 2003). Shown are shear-wave velocity V_s profiles (a, d, g), reflectivity method seismograms (b, e, h) and f–k velocity spectra with theoretical modal curves (c, f, i) for models 1–3, respectively, and normalized colour scale. Model parametrization is given in Table 1. The prior model space is illustrated by the light grey background in the V_s profiles (a, d, g) and by value ranges in Table 2.

Table 2. Uniform prior distribution for the synthetic benchmark modelling (cf., Table 1); the values in brackets represent the minimum and maximum values. h = thickness (m); V_s = S-wave velocity (m s^{-1}); V_p = P-wave velocity (m s^{-1}); ρ = bulk density (kg m^{-3}). Only for the anelastic modelling (cf. Fig. 6): Q_s = S-wave attenuation (–); Q_p = P-wave attenuation (–).

Layer	h	V_s	V_p	ρ	Q_s	Q_p
1	[1, 3]	[50, 250]	[100, 500]	1800	[5, 40]	[15, 65]
2	[2, 6]	[50, 250]	[500, 1500]	1800	[8, 60]	[50, 100]
3	[5, 10]	[50; 250]	[1000, 2000]	1800	[15, 80]	[75, 120]
4	–	[200, 500]	[1000, 2000]	1800	[40, 150]	[100, 200]

satisfactorily recovers the V_s profile, including the velocity inversion in the second layer. A slight bias is visible for parameter V_{s1}, for which the benchmark model is present in the obtained distribution, but located at its edge (Fig. 3i). This is explained by the relatively low sensitivity of BEL1D to this parameter as is visible in Fig. 3(b), (e) and (f).

Similar trends are observed for models 1 (Appendix Figs A1 and A2) and 3 (Appendix Figs A3 and A4). Model 1 shows consistent V_s and V_p posterior distributions, with the exception of a slight underestimation of the third-layer S-wave velocity V_{s3}. Model 3 successfully reproduces the scattered modes in the best posterior dispersion image and captures the V_s inversion in the third layer, demonstrating that the approach robustly handles different model complexities.

3.1.3 *Noise contamination*

The robustness of the method is evaluated on synthetic benchmark model 2 under two types of perturbations: first, disturbing the time

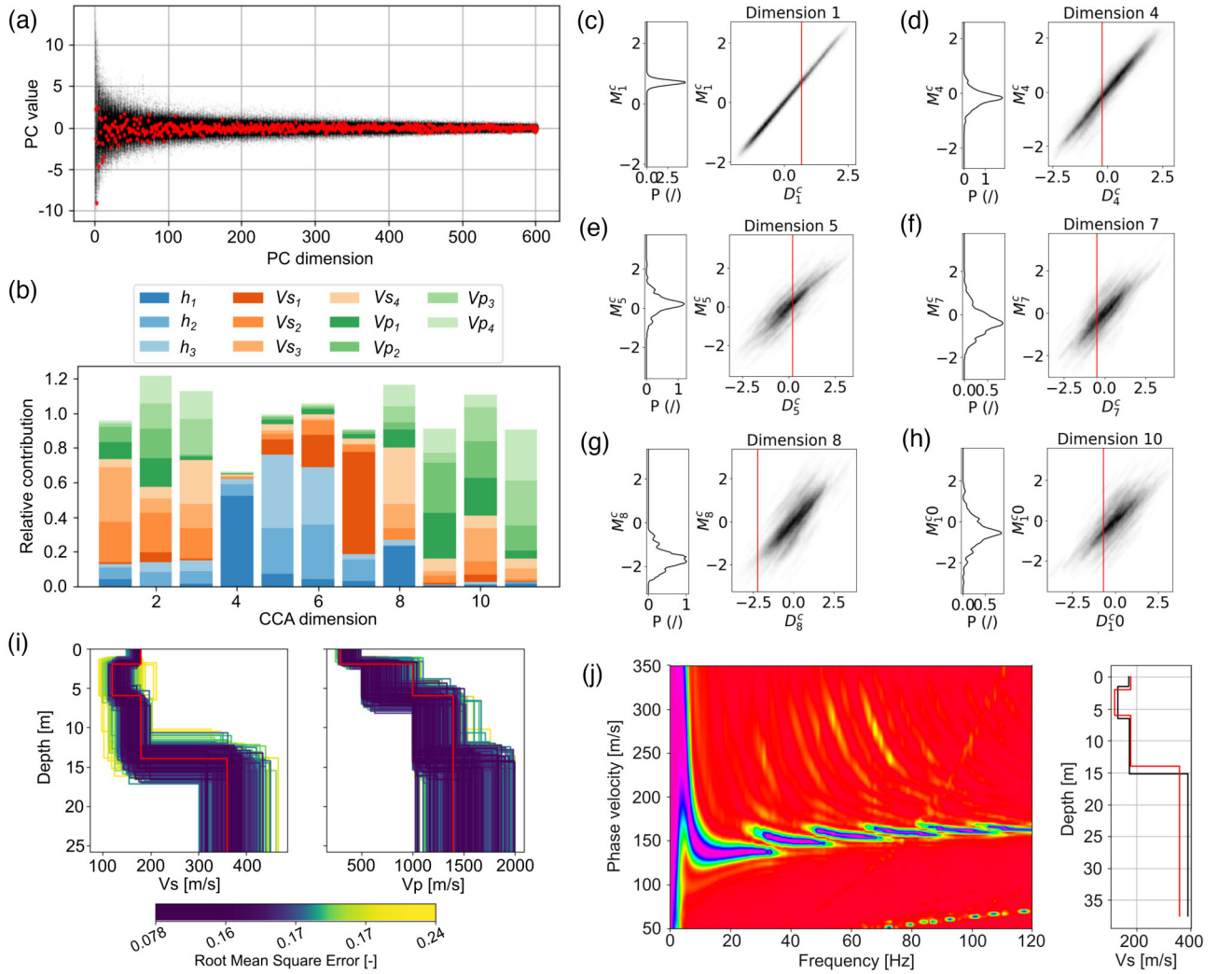

Figure 3. Model 2: (a) PCA space with black dots showing the modelled data and red dots the experimental data; (b) CCA space; (c–h) KDE and projections in CCA dimensions 1, 4, 5, 7, 8 and 10, with red lines marking the position of the experimental data; (i) ensemble of posterior solutions for V_s and V_p, coloured by normalized RMS error (red profiles = benchmark); (j) forward modelled velocity spectrum from the best posterior solution (cf. Fig. 2f), here shown for V_s and V_p over depth (red profiles = benchmark).

traces with Gaussian noise, and second, adding a sound wave (V_p = 330 m s^{-1}) to the seismic shot gather. For the first test, random Gaussian noise is added in the time domain with an amplitude multiplier of 10^{-9}, consistent with the order of magnitude of the seismic signal (S/N ratio: 21.9 dB). The altered traces are shown in Fig. 4(a) alongside the noise-free traces of model 2. Compared to the clean dispersion image (Fig. 2b), the noisy image is mainly affected at low frequencies below 20 Hz (Fig. 4b). The PCA space exhibits more outliers in the lower dimensions (Fig. 4c), while the CCA space remains largely consistent, apart from a reduced sensitivity to V_s in dimensions 7 and 8 (Fig. 4d). In the correlated data—model spaces, the benchmark projections are less central than in the noise-free case (Figs 4e–j; cf. Appendix Fig. A7). The posterior ensemble (Fig. 4k) reproduces V_s reliably in the first two layers but fails to capture the third layer and the half-space; V_p is also inaccurate in the half-space. These misfits can be explained by the disturbances in the lower frequency content of the noisy data, that affects the model solutions at higher depth. Note that this could be improved in the future by using a more robust wavefield transform (Vantassel & Cox 2022). The velocity spectrum of the best-fitting posterior (Fig. 4l) shows clear mode separation and satisfactorily reproduces the observed dispersion, though the associated V_s profile remains inaccurate for the deeper layers. The same bias appears for the same reason as for the noise-free case for V_{s1} (Fig. 4k).

For the second test, a synthetic sound wave of 330 m s^{-1} with a source frequency of 100 Hz was added to the time domain shot gather of model 2 (Fig. 5a). The wavefield transform produces an altered velocity spectrum, where the added noise appears in the 0–75 Hz bandwidth at phase velocities above 200 m s^{-1} (Fig. 5b). In this case, we examine the effect of noisy data when projected into a clean prior, that is, a prior trained on noise-free models and data, with only the projected data d^* being contaminated. The PCA space is shown in Fig. 5(c) with some outliers in the lower dimensions. The CCA data space and the data projections for the CCA dimensions 1, 4, 5, 7, 8 and 10 are shown in Figs 5(d) and (e)–(j), respectively (see Appendix Fig. A8 for all dimensions). The resulting posterior ensemble fails to reproduce the benchmark solution: in particular, the velocity inversion in the second V_s layer is not captured, and both layer depths and V_p show poor agreement (Fig. 5k). The dispersion image generated from the best posterior model

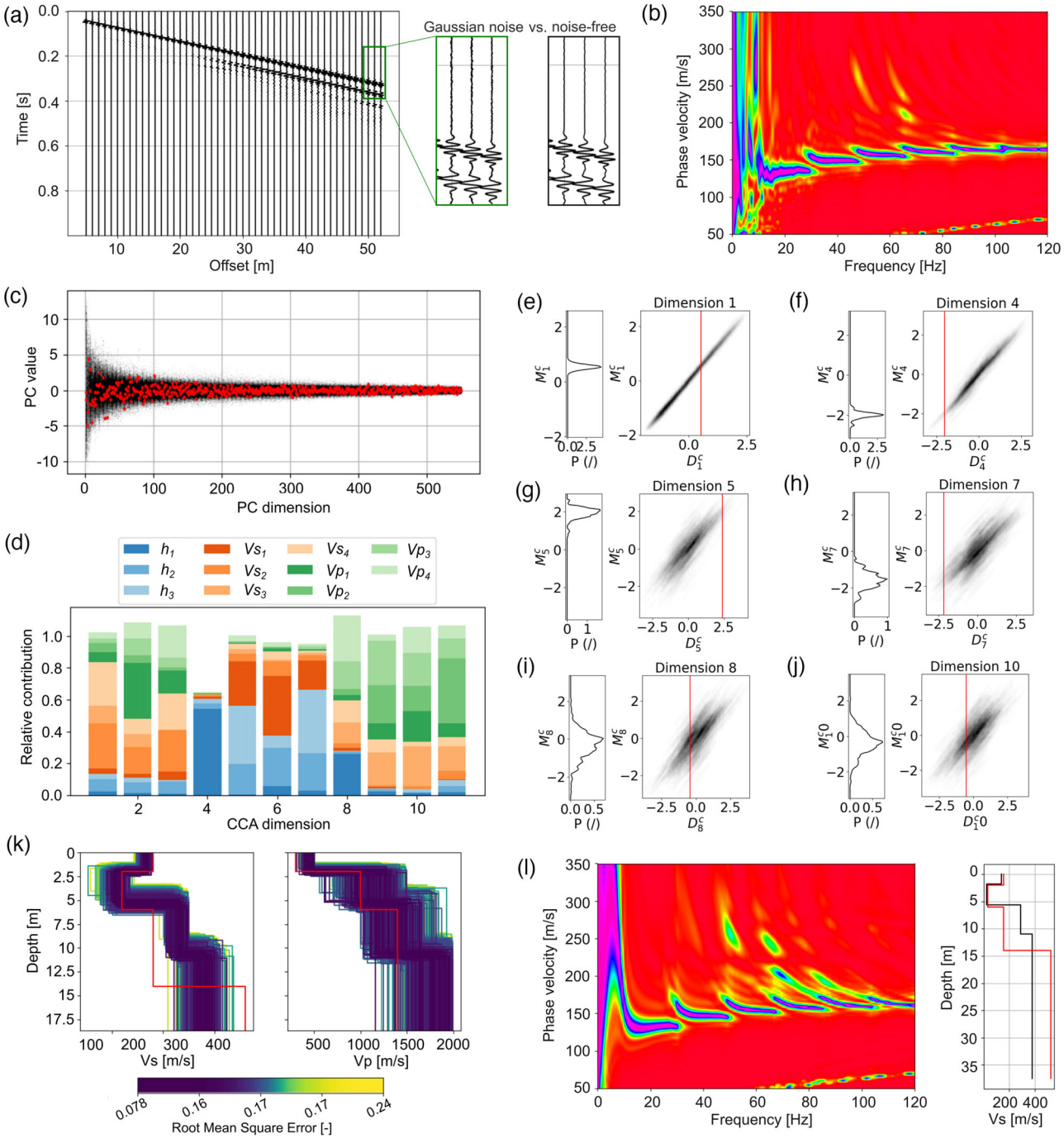

Figure 4. Model 2 with Gaussian noise: (a) shot gather with zoom illustrating the noise impact; (b) velocity spectrum (f–k transform, frequency-maximum normalization); (c) PCA space with black dots showing modelled data and red dots the experimental data; (d) CCA space; (e–j) KDE and projections in CCA dimensions 1, 4, 5, 7, 8 and 10, with red lines marking the position of the experimental data; (k) ensemble of posterior solutions for V_s and V_p, coloured by normalized RMS error (red profiles = benchmark); (l) forward modelled velocity spectrum from the best posterior solution, here shown for V_s over depth (red profile = benchmark).

(lowest RMS error; Fig. 5l) also departs strongly from the benchmark, with an inaccurate V_s depth profile. In contrast, when the prior itself is trained with the same noise source (i.e. the 100 Hz sound wave added to all prior shot gathers), posterior predictions improve substantially. The ensemble of solutions (Fig. 5m) recovers both V_s and V_p in good agreement with the benchmark (again, with the bias for V_{s1} observed in the posterior distribution), and the best posterior model and its velocity spectrum (Fig. 5n) successfully match both the benchmark and the noisy dispersion image in Fig. 5(b). This highlights that training the prior with information that has a clear and quantifiable impact on the experimental data is crucial to ensure consistency between prior and experimental observations. Full results for the noise-trained prior are shown in Appendix Figs A9 and A10.

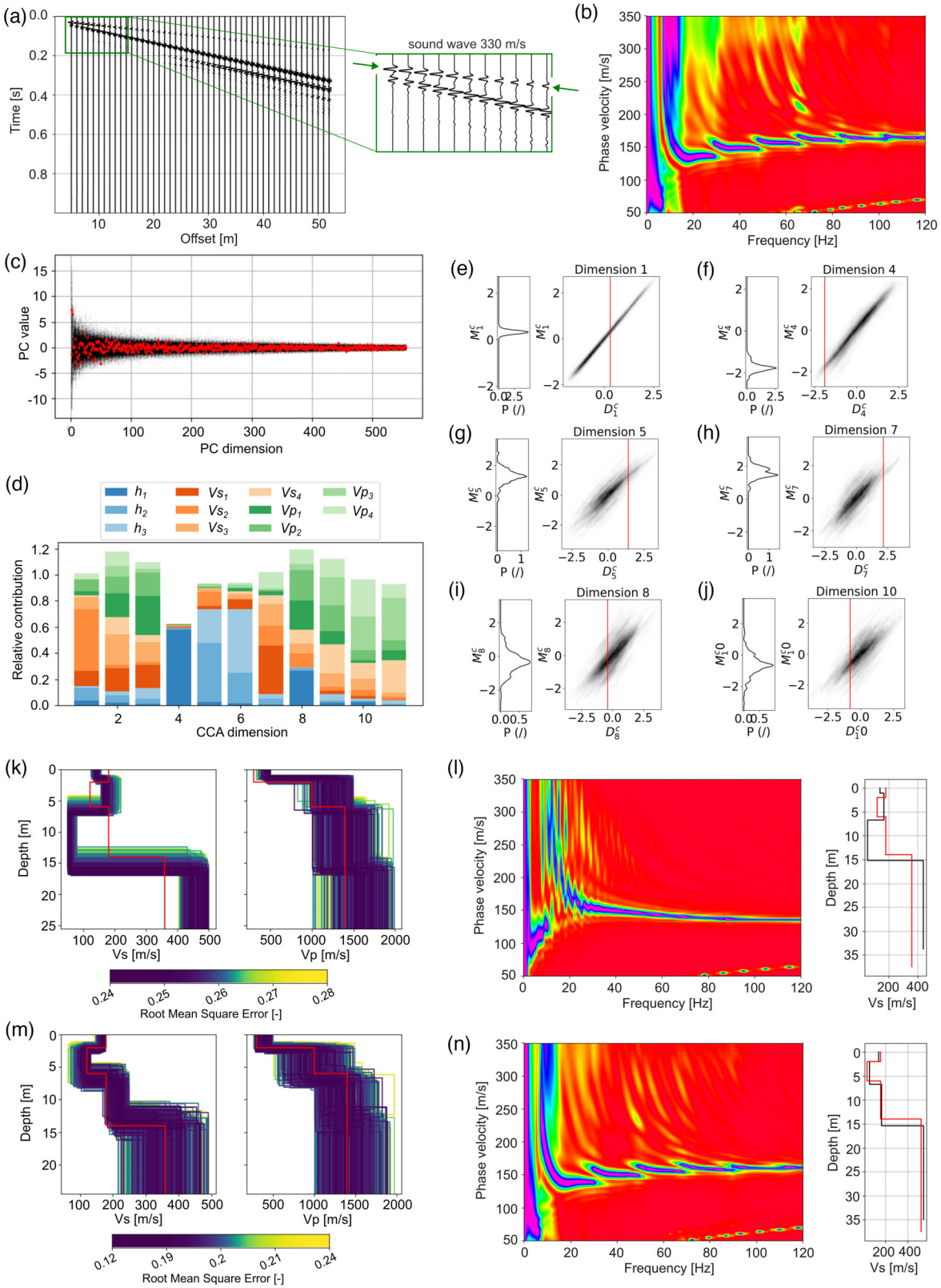

Figure 5. Model 2 with consistent noise. (a) Shot gather with zoom illustrating the integrated sound wave (V_p=330 m s^{-1}); (b) velocity spectrum (f–k transform, frequency-maximum normalization); (c) PCA space with black dots showing modelled data and red dots the experimental data; (d) CCA space; (e–j) KDE and projections in CCA dimensions 1, 4, 5, 7, 8 and 10, with red lines marking the position of the experimental data. Posterior ensembles and forward modelled velocity spectra from the best posterior solution are shown for (k–l) a clean prior and (m–n) a noise-trained prior. In each case, (k, m) show ensembles of V_s and V_p coloured by normalized RMS error and (l, n) show forward modelled velocity spectra from the best posterior solution, shown for V_s over depth (red profiles = benchmark).

3.2 Anelastic application

A factor neglected in the benchmark models above is the decay of seismic energy, that is, the attenuation of ground motion. This becomes evident when examining the absolute amplitudes of non-normalized traces in a shot gather. Unlike a synthetic, perfectly elastic medium, the amplitude of a real wavelet, and thus its stored energy, decreases as it propagates through the subsurface. Attenuation of seismic waves is governed by two main phenomena (cf. Barton 2006): (i) geometrical spreading, which redistributes seismic energy over larger volumes with distance from the source, and (ii) intrinsic (anelastic) attenuation, caused by mechanical energy loss due to internal friction, especially at grain or interface boundaries (e.g. water–soil). In addition, heterogeneous geology can produce apparent attenuation through scattering, without actual energy loss. The frequency-dependence in this matter is highly discussed whether to be subject to scattering alone, or also to intrinsic attenuation (e.g. Shapiro & Kneib 1993; Van Der Baan 2002; Morozov 2009), but it is generally accepted that high frequencies are subject to higher attenuation. The inelasticity of a medium can be described by the dimensionless quality factor Q (attenuation $= Q^{-1}$), which ranges from 1 to 5000 in seismic applications (Barton 2006). We distinguish between Q_p and Q_s for compressional and shear waves, respectively. A recent review of near-surface Q_s estimates from experiments and modelling is provided by Parolai *et al.* (2022).

For an anelastic test, we extend the synthetic model 2 with the attenuation parameters $Q_s = (20, 35, 45, 75)$ and $Q_p = (40, 70, 90, 150)$. Including these parameters expands the prior parameter space of model 2 from 11 to 19 dimensions (five free parameters over three layers and a half-space). The prior distributions of both attenuation parameters were kept broad (see values in Table 2). Although the benchmark models of Tokimatsu *et al.* (1992) are purely elastic, the selected Q values fall within the realistic ranges of $Q_s\text{min} = 0.1Q_p$ to $Q_s\text{max} = Q_p$ and Q_p being larger than Q_s by the factor of 2 (cf. Xia *et al.* 2012). Q_s values were chosen to increase with layer stiffness and depth, where we expect near-surface layers to be more dissipative. The trace-by-trace normalized anelastic shot gather exhibits a distinct signature in the first arrivals due to the integration of the quality factors in the layered Earth model (Fig. 6a). This signature is visible because the refracted wave has a relatively high amplitude, which becomes apparent here due to the reduced amplitude of the attenuated ground roll.

To maintain consistency with the previous implementations of model 2, attenuation is introduced sequentially. First, a prior of 1000 models (average computation time: 5.36 s/model) is generated using the same wavefield transform as before (f–k with frequency-maximum normalization). The normalization enhances the visibility of all surface wave modes across the full frequency range, including low frequencies that are often masked in amplitude-preserving transforms. It is therefore particularly suitable for constraining the elastic structure. Like before, a posterior is predicted based on the corresponding benchmark velocity spectrum. The results of this step are presented in Appendix Figs A11 and A12. Since the frequency-maximum normalization suppresses amplitude information by normalizing per frequency band, attenuation cannot be reliably predicted at this stage. To overcome this limitation, we perform a second step: the posterior ensemble from the first prediction is used as new prior. The sampled models from the new prior as well as the projected benchmark now employ an f–k wavefield transform with absolute-maximum normalization (Fig. 6b). In this way, the elastic structure is well-constrained in a pre-trained space for the parameters h, V_s and V_p, and attenuation relative information is introduced in a second step by preserving amplitude relationships across frequencies.

The BEL results show a PCA reduction from 12 100 to 73 dimensions (Fig. 6c) with minor outliers in the lowest components, similar to the elastic case. A distance-based generalized sensitivity analysis (DGSA; Fenwick *et al.* 2014), performed with the pyDGSA package (pypi.org/project/pyDGSA/1.0.4/), is compared against parameter contributions in the CCA space (Fig. 6d). The analysis highlights several dominant parameters: V_{s1}, V_{s2} and $Qs1$ (especially in the first eight CCA dimensions), h_1 (dominant in CCA dimension 4) and V_{p1} (dimensions 3 and 6). Lower in the ranking are the remaining P-wave velocities and the P-wave quality factors with Q_{p1} to Q_{p3} among the least sensitive. This reflects how CCA concentrates the most influential parameters in its leading axes (cf. Michel *et al.* 2020), while less sensitive parameters, here the P-wave attenuation factors, are present in higher dimensions. The correlated data—model projections show strong relationships in the lower CCA dimensions, with decreasing correlation at higher dimensions (Figs 6e–h; full projections in Appendix Fig. A13).

The posterior model with the lowest RMS error is able to reproduce the benchmark dispersion image and to resolve the V_s model with a deviation in the thickness h_3 of the third layer (Fig. 6i). The Q_s and Q_p models deviate from the benchmark but maintain consistent trends, apart from a localized inversion of Q_s in the third layer. Finally, Fig. 6(j) presents the full posterior ensembles for V_s, V_p, Q_s and Q_p, showing coherent predictions for all parameters. In this case, the parameter V_{s1} is within the obtained posterior as is expected by the much higher weight of this parameter in the lower dimensions of the CCA transform seen in Fig. 6(d). The sequential approach enabled a significant uncertainty reduction, with RMS values decreasing by approximately one order of magnitude between the first and second posterior ensembles (cf. Appendix Fig. A11).

The synthetic tests demonstrate the applicability of the proposed method and provide several key insights: (i) in out-of-prior projections, the experimental data fail to align with the CCA space in dimensions sensitive to parameters excluded from the prior; (ii) in the presence of systematic noise, a noise-trained prior is essential to obtain consistent posterior results; and (iii) in anelastic modelling, the inclusion of quality factors, although ranked as low-sensitivity parameters, can still exert a pronounced influence on modelled wavefields and posterior predictions, highlighting the importance of attenuation in amplitude preserving analyses.

4 REAL DATA APPLICATION

In the following, we illustrate the methodology on two field data sets from the literature collected with shallow refraction seismic methods. Both data sets differentiate primarily in the acquisition resolution in terms of receiver spacing and data quality.

The first field example is from the study of Eppinger *et al.* (2024), which aimed to characterize the critical zone with a full-waveform tomography in the Blair Wallis, Laramie Mountains, Wyoming. From 20 sledgehammer source points, we chose a seismic record close to a borehole location at 175 m profile length in order to use the logging data as reference model for V_s and V_p. Only elastic parameters are considered here, since no attenuation information is available for the near-surface. The selected shot gather is of high resolution (high-temporal sampling, dense receiver spacing of 1 m, see Fig. 7a) and contains little noise. Its velocity spectrum, obtained using a FDBF with absolute-maximum normalization, exhibits a clear fundamental mode and a weak first overtone (Fig. 7b).

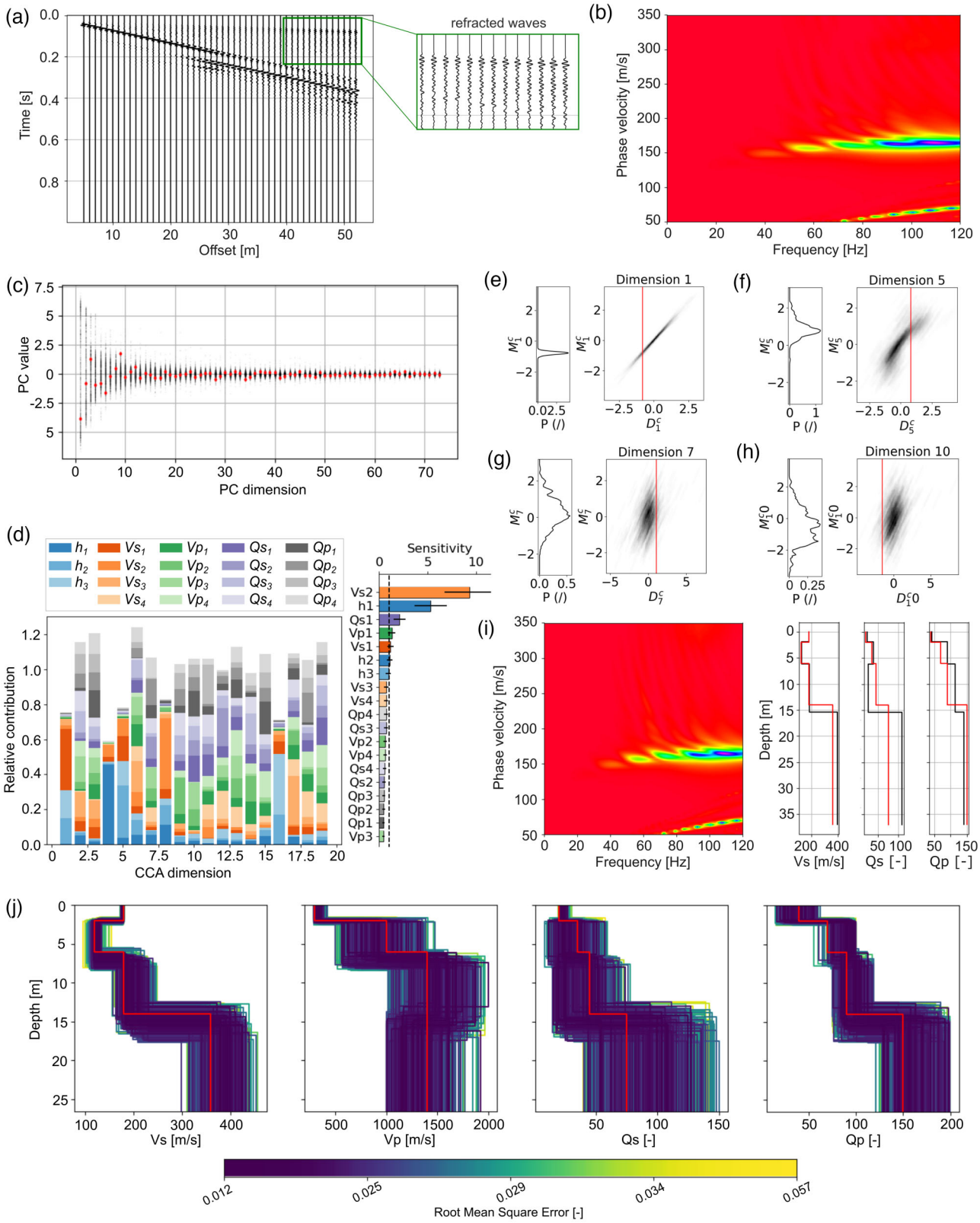

Figure 6. Model 2 in an anelastic milieu: (a) Shot gather with zoom to first arrivals showing anelasticity; (b) velocity spectrum (f–k transform, absolute-maximum normalization); (c) PCA space with black dots for the modelled data and red dots for the experimental data; (d) CCA space with parameter sensitivity analysis; (e–h) KDE and projection in the CCA dimensions 1, 5, 7 and 10, with red lines marking the position of the experimental data; (i) forward modelled velocity spectrum from the best posterior solution, here shown for V_s, Q_s and Q_p over depth (red profile = benchmark); (j) ensemble of posterior solutions for V_s, V_p, Q_s and Q_p, coloured by normalized RMS error (red profiles = benchmark).

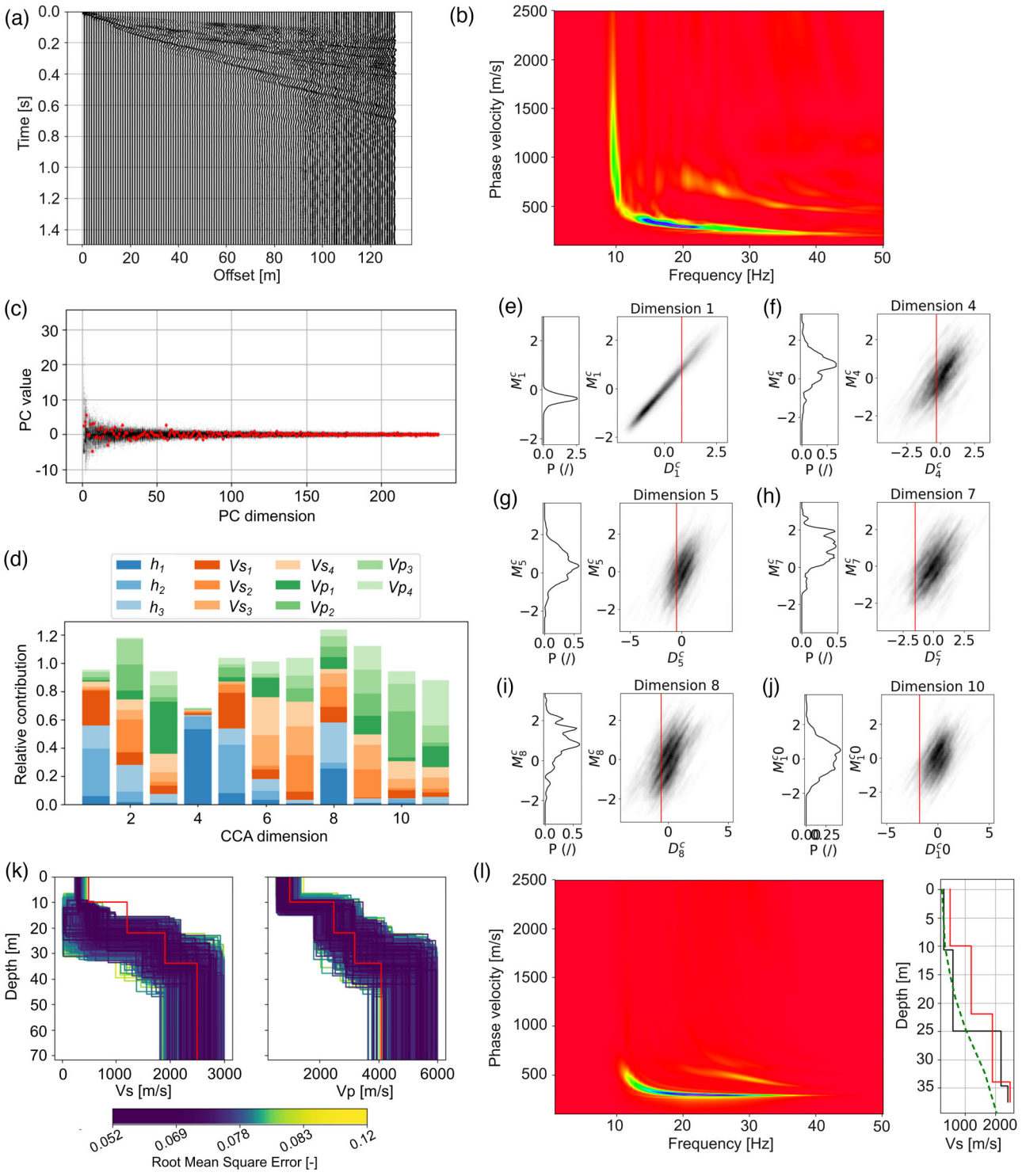

Figure 7. Real data application from Eppinger *et al.* (2024): (a) 240-channel seismic shot gather; (b) velocity spectrum (FDBF wavefield transform, absolute-maximum normalization); (c) PCA space with black dots showing modelled data and red dots the experimental data; (d) CCA space; (e–j) KDE and projection in the CCA dimensions 1, 4, 5, 7, 8 and 10, with red lines being the position of the experimental data; (k) ensemble of posterior solutions for V_s and V_p coloured by normalized RMS error (red profiles = benchmark); (l) forward modelled velocity spectrum from the best posterior solution here shown for V_s over depth (red profile = benchmark, green pointed line = final model of Eppinger *et al.* 2024).

For the synthetic forward simulations, we applied a 20 Hz Ricker source wavelet, consistent with the dominant frequency content of the field data (cf. Eppinger *et al.* 2024). A layered reference model (further called benchmark) was derived from the borehole log BW4 presented in the original study, located at two meters from the hammer source. Because the borehole data only cover depths from 10 m downwards, the shallow layers are unconstrained. Therefore, we approximated a first layer with $V_s = 500$ m s^{-1} and $V_p = 1000$ m s^{-1}

Table 3. Uniform prior distribution for the real data application of Eppinger *et al.* (2024); the values in brackets represent the minimum and maximum values. The suffix *bench* refers to the benchmark model set up from the borehole log. h = thickness (m); V_s = *S*-wave velocity (m s^{-1}); V_p = *P*-wave velocity (m s^{-1}); Q_s = *S*-wave quality factor. The bulk density ρ was fixed to 2500 kg m^{-3}.

Layer	h_{bench}	h	Vs_{bench}	V_s	Vp_{bench}	V_p
1	10	[5, 15]	500	[150, 1000]	1000	[300, 1500]
2	12	[5, 20]	1200	[150, 1500]	2500	[1800, 3500]
3	12	[5, 20]	1900	[750, 2200]	3200	[2000, 4500]
4	–	–	2500	[1800, 3000]	4100	[3000, 6000]

over 10 m thickness; these values should not be considered direct measurements but reasonable estimates. Accordingly, the prior was built as a four-layer model (Table 3) keeping value intervals for the first layer large enough to ensure coherent posterior predictions, that is, $V_{s1} \in [150, 1000]$, $V_{p1} \in [300, 1500]$ and $h_1 \in [5, 15]$.

In total, 1000 models were sampled from the prior, with an average computational time of 44.7 s per forward simulation. Note, to enhance efficiency for such high-resolution data, computations can be parallelized since the sampled models are independent of one another. The PCA could reduce data dimensions from the original 37 500 to 238 (Fig. 7c), reproducing a variability of 99 per cent. The CCA space and the relative contribution of each parameter in a 11-dimensional space is shown in Fig. 7(d). The data projection in the correlated CCA space is shown in Figs 7(e)–(j) for the dimensions 1, 4, 5, 7, 8 and 10, respectively (see Appendix Fig. A14 for all dimensions).

The ensemble of posterior BEL models is shown in Fig. 7(k) for V_s and V_p. Especially for the first and second layer, V_s predictions are lower than the benchmark, given the restricted borehole information. The best posterior V_s model and the corresponding velocity spectrum are shown in Fig. 7(l), with lower V_s predictions also for the second layer. A comparison between our best posterior model and the final model of the original study (marked by the green pointed line), reveals a similar V_s distribution. A reason for the deviation between both modelled results and the logged groundtruth are possibly due to the offset between the seismic experiment and the borehole location, as well as the difference between sampled rock volumes.

The second field example originates from the study of Mreyen *et al.* (2021) characterizing a landslide site in the Carpathian flysch mountains, Romania, with active sledgehammer shot seismics. We use the time-domain stacked real shot gather S2 (duration 1 s, source position 25 m offset from first geophone, 24 geophones with 5 m spacing and 4.5 Hz eigenfrequency) that is shown in Fig. 8(a). In contrast to the first field example, the data set can be described as very noisy with several dead traces. The FDBF transformed velocity spectrum with absolute-maximum normalization (Fig. 8b) presents a multimode pattern with a presumably clear fundamental mode in the 10–25 Hz frequency range. The Fourier transform of three time-traces of the shot gather are shown in Fig. 8(c) and illustrate the frequency content of the 1st, 5th and 15th geophones with amplitudes decreasing, respectively. The first geophone shows a pronounced peak at 40 Hz, while the peak frequency for the following receivers is shifting to lower frequencies with distance from the source. For our simulations, we decided for a source frequency of 20 Hz, which is the secondary peak in the freuqency spectrum of the 1st geophone and the overall averaged frequency response of the array, producing coherent wave phases in the synthetic seismograms. As benchmark model, we use the best fit result from Mreyen *et al.* (2021) that manually picked dispersion curves and ran inversions with the Geopsy dinver module (global search surface wave inversion; Wathelet *et al.* 2020). From the latter, a 4-layered prior was built (three layers and a subjacent half-space; see Table 4) for the parameters h, V_s and V_p. For this field example, we directly adopted an anelastic approach, since forward simulations of the elastic benchmark model could not reproduce waveforms similar to the experimental shot gather without including attenuation, as illustrated in the Appendix Figs A15(b) and (c), respectively. In an anelastic setting, the parameter space is expanded to 15 dimensions including the *S*-wave quality factor Q_s for the four layers. Lacking prior information on attenuation but considering the loose geological setting, we adopted a rather low prior interval $Q_s \in [1, 50]$ for all layers. The parameter of *P*-wave attenuation, that has shown to be less sensitive in our synthetic attenuation analysis, was kept at the constant value of $Q_p = 50$ to reduce the dimensionality of the model space. Unlike the synthetic attenuation test, a sequential approach is not required here since the absolute-maximum normalization of the wavefield transform preserves relative amplitude information across frequencies.

In total, 1000 models were sampled from the prior with an average runtime of 2.14 s per forward simulation.The PCA could reduce data dimensions from the original 15 000 to 172 (Fig. 8d), reproducing a variability of 99 per cent. The CCA space and the relative contribution of each parameter in a 15-D space is shown in Fig. 8(e). We also include the parameters sensitivity distribution (DGSA; Fenwick *et al.* 2014), where the attenuation parameters Q_{s2} and Q_{s4} rank fourth and fifth in sensitivity, respectively. The data projection in the correlated CCA space is shown in Figs 8(f)–(k) for the dimensions 1, 4, 5, 7, 8 and 10, respectively (see Appendix Fig. A16 for all dimensions).

The posterior ensemble of models is shown in Fig. 8(l) for V_s, V_p and Q_s evaluated by RMS calculations between the true and the modelled velocity spectrum. For Q_s, a clear trend is observed in the form of an inversion within the second subsurface layer. Fig. 8(m) shows the best posterior model (V_s, V_p and Q_s) together with its forward modelled dispersion image. The fundamental mode is well reproduced in the frequency range 10–25 Hz and the velocity range 200–500 m s^{-1}. The modelled dispersion image exhibits higher energies but does not fully capture the supposed upper modes seen in Fig. 8(b). When comparing the benchmark of the best-fit Geopsy models for V_s and V_p (red curves), we observe a similar velocity trend but lower layer thicknesses. For Q_s, no benchmark is available, as attenuation parameters were not studied in parallel to V_s in Mreyen *et al.* (2021) and are not invertible with the employed Geopsy Dinver module. Nevertheless, an inversion of Q_s is apparent in the second layer, consistent with the pattern observed in the posterior ensemble, as well as the half-space. The forward modelled shot gather from the best posterior model (black traces in Fig. 8n), approaches the experimental shot gather in terms of arrivals and spreading rather efficiently compared to the suggested benchmark solution (see forward modelled waveforms in Appendix Figs A15b and d).

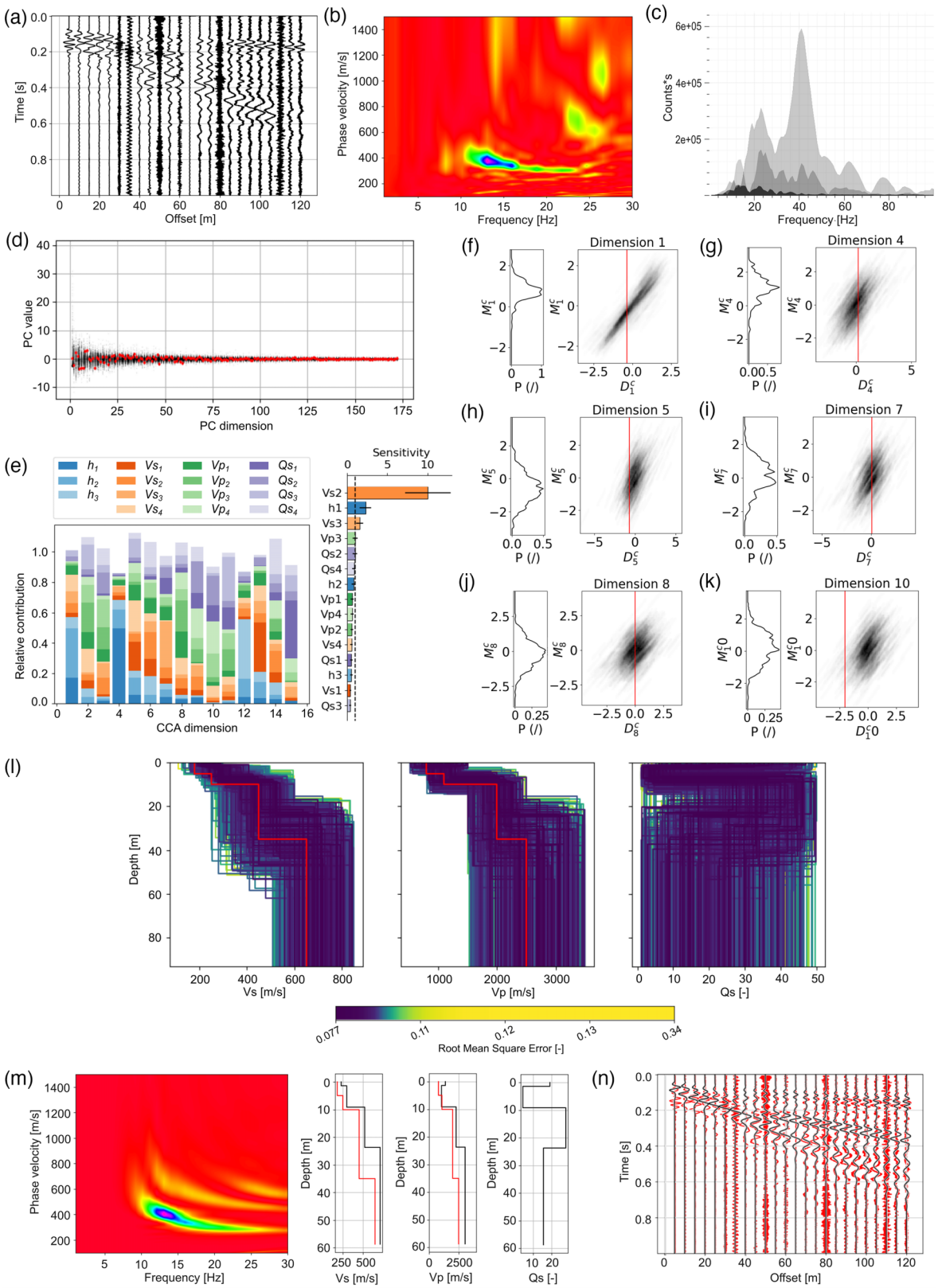

Figure 8. Real data application from Mreyen *et al.* (2021): (a) 24-channel seismic shot gather; (b) velocity spectrum (FDBF wavefield transform, absolute-maximum normalization); (c) frequency content of the 1st, 5th and 15th (biggest to lowest amplitude) geophones; (d) PCA space with black dots showing the modelled and red dots the experimental data; (e) CCA space with sensitivity analysis; (f–k) KDE and projection in the CCA dimensions 1, 4, 5, 7, 8 and 10, with red lines being the position of the experimental data; (l) ensemble of posterior solutions for V_s, V_p and Q_s coloured by normalized RMS error (red profiles = benchmark); (m) forward modelled velocity spectrum from the best posterior solution, here shown for V_s, V_p and Q_s over depth (red profiles = benchmark); (n) full waveform of the best posterior model (black) projected on the experimental shot gather (red).

Table 4. Uniform prior distribution for the real data application of Mreyen *et al.* (2021); the values in brackets represent the minimum and maximum values. The suffix *bench* refers to the benchmark model. h = thickness (m); V_s = *S*-wave velocity (m s^{-1}); V_p = *P*-wave velocity (m s^{-1}); Q_s = *S*-wave quality factor. The bulk density ρ was fixed to 2000 kg m^{-3}.

Layer	h_{bench}	h	Vs_{bench}	V_s	Vp_{bench}	V_p	Q_s
1	5	[1, 10]	180	[100, 300]	800	[500, 1000]	[1, 50]
2	5	[1, 10]	250	[200, 500]	1100	[800, 1500]	[1, 50]
3	25	[10, 50]	450	[250, 600]	2000	[1500, 2500]	[1, 50]
4	–	–	650	[500, 850]	2500	[1500, 3500]	[1, 50]

5 DISCUSSION

In the following, we address common challenges in Bayesian inversion frameworks, with an emphasis on those that are particularly relevant for surface wave applications.

As illustrated in Fig. 1, the prior model space, that is defined by the number of layers and distributions of (an-)elastic parameters, is the starting point of BEL. In general, the definition of the prior is a key element in stochastic and Bayesian approaches; in our case, a data set solution can only be predicted if it falls within the dimensional range of the trained prior, due to the inability of CCA to extrapolate beyond the training space. This becomes particularly problematic for real data applications when no information on local subsurface parameter distributions is available. However, geophysical data sets rarely exist in isolation and should ideally be accompanied by complementary information, either geophysical or geotechnical. A common method of surface wave acquisition is the multichannel analysis of surface waves (MASW), which can provide *P*-wave velocity estimates from first arrivals along the receiver array. These can then be used to define physically plausible parameter ranges through empirical relationships. The data from our field applications originate from active seismic measurements, that were interpreted in terms of *P*-wave velocity (seismic refraction tomography) and surface waves (MASW) providing the benchmark solutions validating our method. The field applications demonstrate that an accurate selection of the prior parameter space is essential for the BEL framework. The prior can be flexibly adapted, and we recommend beginning with relatively broad parameter intervals to avoid biasing predictions. To reduce dimensionality, parameters that are weakly sensitive to the problem, as in our case, density ρ or compressional quality factor Q_p, may be fixed.

When moving from prior definition to data realism, an important aspect is the impact of noise. For the synthetic benchmarks, we tested both random Gaussian and systematic noise sources. Gaussian noise affects especially the low-frequency components of the dispersion image, leading to reduced sensitivity to deeper layers. Posterior solutions remain stable when noise levels are moderate, but we advise caution and, if possible, pre-processing (e.g. bandpass filtering, eliminating dead traces) to enhance the seismic signal. Systematic noise, such as acoustic waves travelling through air and recorded by the receivers, can strongly distort posterior predictions when not represented in the prior. Since such noise is often easily identifiable, we suggest explicitly including a simulated noise component in the training data, as demonstrated in our application using a 330 m s^{-1} acoustic wave. This ensures that data are projected into a coherent prior space reflecting the field conditions and thus improves posterior accuracy.

A challenge in anelastic data prediction, inevitably encountered in real data, is the inclusion of the quality factors Q_p and Q_s for compressional and shear waves, respectively, that characterize the mediums attenuation. As discussed earlier, the net attenuation of seismic energy represents a complex combination of spatial, intrinsic and scattering effects. In the literature, Q values are often fixed empirically for surface wave inversion, for example, $Q_s = V_s/8$ and $Q_s = Q_p$ (Dal Moro 2014), or $0 < Q_s < 100$ and $Q_p = 2Q_s$ (Xia *et al.* 2012). Other studies have proposed standardized or depth-dependent Q values, for example, $Q_s = 20$ at 5 m depth (Badsar *et al.* 2010). Alternative approaches invert for Q_s through coupled or uncoupled analysis of phase velocity and attenuation curves (e.g. Lai 1998; Rix *et al.* 1998; Xia *et al.* 2012). In the context of full-waveform inversion (FWI), Groos *et al.* (2014) showed that source wavelet properties can partly compensate for the absence of attenuation modelling when elastic FWI is applied to anelastic data.

In our study, we addressed the challenge of simultaneously constraining elastic and anelastic parameters by adopting a sequential approach in the synthetic tests. For the landslide field data application, which inherently includes amplitude information, and thus attenuation effects, we opted for a simplified workflow. Here, a direct waveform transform preserving relative amplitudes was sufficient to constrain attenuation without the need for sequential training. Therefore, while the sequential approach can serve as an efficient strategy to improve model stability and prediction accuracy, its necessity depends on the data quality, model configuration and the level of required attenuation sensitivity. This also highlights the importance of selecting an appropriate wavefield transform and normalization strategy to optimally extract model features from the data.

In addition to attenuation, the source wavelet plays a critical role in reproducing field wavefields. For synthetic tests, we used an 80 Hz Ricker wavelet, whereas in the field case, lower frequency excitations, more representative of hammer sources, were required. We simulated this using 20 Hz Ricker wavelets, and future extensions could include the source parameters (frequency and amplitude) directly in the BEL parameter space, enabling a joint characterization of source and subsurface. This would, however, require a higher number of model realizations to adequately cover the expanded parameter domain.

The computational efficiency of BEL1D has been benchmarked against conventional stochastic sampling methods, such as McMC, by Michel *et al.* (2020). Their analysis shows that the total runtime is primarily governed by the performance of the forward modelling routine, in this study implemented using the reflectivity method and the wavefield transform, and by the number of prior realizations used to span the model space. Unlike traditional stochastic approaches that iteratively sample the posterior by minimizing a misfit function and require thousands of iterations to converge, BEL1D benefits from the independence of its prior samples, allowing the forward simulations to be fully parallelized and thus substantially reduce computation times. Once a well-trained prior has been established for a given site, BEL1D allows near-instantaneous prediction of new data sets, when not requiring new forward runs to analyse

the results misfits. This reusability of the trained prior makes the method especially suitable for applications involving repeated or long-term monitoring, such as landslides, where it enables rapid and automated imaging with minimal additional computational cost.

Finally, we note that in field applications, true benchmark solutions are typically unavailable (since even borehole information does not rely on the same sampled volume of rocks or sediments). For controlled experiments, a distance-based misfit between the benchmark V_s profile and the posterior prediction can serve as an alternative to image-based error measures. In field applications, however, image-based misfit evaluations between observed and simulated dispersion spectra remain more practical and can be strengthened by waveform comparisons. Future developments could integrate advanced image comparison metrics or machine learning based similarity measures to improve posterior validation.

The here presented methodology has been developed from the pyBEL1D version 1.0.1 maintaining its modular structure and open-source compatibility. In future work, we aim to incorporate advanced image-analysis tools such as 2-D functional principal component analysis (e.g. Shi *et al.* 2022) and extend the method to (pseudo-)2-D domains. This would open new opportunities for coupling BEL with full-waveform inversion, where BEL-derived posteriors could serve as robust and data-consistent starting models (e.g. Aleardi & Mazzotti 2017; Adriaens *et al.* 2023).

6 CONCLUSIONS

We present an application of BEL to predict 1-D subsurface models on the basis of multimodal velocity spectra from wavefield-transformed seismic shot gathers and prior information. This approach captures surface wave dispersion in its entirety, thereby eliminating the need for manual dispersion curve picking, reducing subjectivity and enabling uncertainty quantification through the prediction of posterior ensembles. We further expanded the parameter space to account for attenuation and performed a corresponding sensitivity analysis. For the anelastic synthetic test, a sequential workflow was introduced: elastic parameters were first constrained, followed by the inclusion of amplitude-relevant information to capture attenuation effects. Field applications included a high-resolution data set from a critical zone site and a noisy data set from a landslide, both demonstrating that with adequately defined priors, BEL produces posterior ensembles that yield coherent dispersion images consistent with experimental observations. A challenging issue remains the definition of the source wavelet and attenuation parameters; future studies should focus on incorporating these parameters—such as source frequency and amplitude—into the prior model space, allowing their posterior prediction and compensating for potential elastic–anelastic mismatches in waveform simulations.

ACKNOWLEDGMENTS

We gratefully acknowledge the *Service Public de la Wallonie* (SPW, Belgium) for supporting this research through the Win2Wal funding programme (*EXPANSION*, convention number 20101). We also thank the editors, Ludovic Métivier, Fern Storey and Louise Alexander, for their guidance, and the two anonymous reviewers for their constructive comments that helped improve the manuscript.

AUTHOR CONTRIBUTIONS

Anne-Sophie Mreyen (Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing—original draft), Hadrien Michel (Formal analysis, Methodology, Software, Writing—review & editing) and Frédéric Nguyen (Conceptualization, Funding acquisition, Methodology, Supervision, Writing—review & editing).

DATA AVAILABILITY

The data and code that support the findings of this study are openly available at Mreyen et al. 2025 (https://doi.org/10.5281/zenodo.17807710).

REFERENCES

Adriaens, X., Métivier, L. & Geuzaine, C., 2023. Inner product preconditioned trust-region methods for frequency-domain full waveform inversion, *J. Comput. Phys.,* **493,** 112 469.

Ahmed, A., Aigner, L., Michel, H., Deleersnyder, W., Dudal, D., Flores Orozco, A. & Hermans, T., 2024. Assessing and improving the robustness of Bayesian evidential learning in one dimension for inverting time-domain electromagnetic data: introducing a new threshold procedure, *Water,* **16**(7), 1056.

Aigner, L., Michel, H., Hermans, T. & Orozco, A.F., 2025. Stochastic inversion of transient electromagnetic data to derive aquifer geometry and associated uncertainties, *Geophys. J. Int.,* **242**(2), ggaf236..

Aleardi, M. & Mazzotti, A., 2017. 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm–Gibbs sampler approach, *Geophys. Prospect.,* **65**(1), 64–85..

Alyousuf, T., Colombo, D., Rovetta, D. & Sandoval-Curiel, E., 2018. Near-surface velocity analysis for single-sensor data: an integrated workflow using surface waves, AI, and structure-regularized inversion, *SEG Technical Program Expanded Abstracts 2018,* pp. 2342–2346, Society of Exploration Geophysicists.

Badsar, S., Schevenels, M., Haegeman, W. & Degrande, G., 2010. Determination of the material damping ratio in the soil from SASW tests using the half-power bandwidth method, *Geophys. J. Int.,* **182**(3), 1493–1508.

Barton, N., 2006. *Rock Quality, Seismic Velocity, Attenuation and Anisotropy,* CRC Press.

Cárdenas, J., Burzawa, A., Radic, N., Bodet, L., Vidal, R., Diop, K., Dangeard, M. & Dhemaied, A., 2025. Deep learning contribution to the automatic picking of surface-wave dispersion for the characterization of railway earthworks, *Comput. Geosci.,* **198,** 105 883.

Chamorro, D., Zhao, J., Birnie, C., Staring, M., Moritz, F. & Ravasi, M., 2023. Deep Learning-based extraction of surface wave dispersion curves from seismic shot gathers, *Near Surface Geophysics,* **22(4),** 421–437

Cho, S., Pyun, S., Choi, B., Lee, G., Jang, S. & Choi, Y., 2024. Prediction of S-wave velocity models from surface waves using deep learning, *Near Surf. Geophys.,* **22**(3), 281–297.

Dai, R.T., Shao, G.Z., Liu, X.Y., Ren, Z.M., Heng, X.T. & Ren, X.D., 2022. Study of an automatic picking method for multimode dispersion curves of surface waves based on an improved U-net, *IEEE Trans. Geosci. Remote Sens.,* **60,** 1–14.

Dai, T., Xia, J., Ning, L., Xi, C., Liu, Y. & Xing, H., 2021. Deep learning for extracting dispersion curves, *Surv. Geophys.,* **42,** 69–95.

Dal Moro, G., 2014. *Surface Wave Analysis for Near Surface Applications,* Elsevier.

Dal Moro, G., 2019. Surface wave analysis: improving the accuracy of the shear-wave velocity profile through the efficient joint acquisition and Full Velocity Spectrum (FVS) analysis of Rayleigh and Love waves, *Explor. Geophys.,* **50**(4), 408–419.

Dal Moro, G., Coviello, V. & Carlo, G.D., 2015. Shear-wave velocity reconstruction via unconventional joint analysis of surface waves: a case study in the light of some theoretical aspects. *Engineering Geology for*

Society and Territory, Volume 5: Urban Geology, Sustainable Planning and Landscape Exploitation, pp. 1177–1182, Springer.

Devroye, L., 2006. Nonuniform random variate generation, *Handbooks in Operations Research Management and Science,* Vol. **13,** pp. 83–121. Elsevier

Dong, S., Li, Z., Chen, X. & Fu, L., 2021. DisperNet: an effective method of extracting and classifying the dispersion curves in the frequency–Bessel dispersion spectrum, *Bull. seism. Soc. Am.,* **111**(6), 3420–3431.

Dou, S. & Ajo-Franklin, J.B., 2014. Full-wavefield inversion of surface waves for mapping embedded low-velocity zones in permafrost, *Geophysics,* **79**(6), EN107–EN124.

Dunkin, J.W., 1965. Computation of modal solutions in layered, elastic media at high frequencies, *Bull. seism. Soc. Am.,* **55**(2), 335–358.

Eppinger, B., Holbrook, W., Liu, Z., Flinchum, B. & Tromp, J., 2024. 2D near-surface full-waveform tomography reveals bedrock controls on critical zone architecture, *Earth Space Sci.,* **11**(2), e2023EA003248.

Fenwick, D., Scheidt, C. & Caers, J., 2014. Quantifying asymmetric parameter interactions in sensitivity analysis: application to reservoir modeling, *Math. Geosci.,* **46,** 493–511.

Forbriger, T., 2003a. Inversion of shallow-seismic wavefields: I. Wavefield transformation, *Geophys. J. Int.,* **153**(3), 719–734.

Forbriger, T., 2003b. Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms, *Geophys. J. Int.,* **153**(3), 735–752.

Fuchs, K. & Müller, G., 1971. Computation of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. Int.,* **23**(4), 417–433.

Gan, Y., Yang, Z., Pan, L., Sun, Y.C., Zhang, D., Gao, Y. & Chen, X., 2024. Deep learning-based dispersion spectrum inversion for surface wave exploration, *IEEE Trans. Geosci. Remote Sens,* **62,** 1–11.

Groos, L., Schäfer, M., Forbriger, T. & Bohlen, T., 2014. The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves, *Geophysics,* **79**(6), R247–R261.

Haskell, N.A., 1953. The dispersion of surface waves on multilayered media, *Bull. seism. Soc. Am.,* **43,** 17–34.

Hermans, T., Compaire, N., Thibaut, R. & Lesparre, N., 2021. Bayesian evidential learning: An alternative to hydrogeophysical coupled inversion, *First International Meeting for Applied Geoscience and Energy,* pp. 3125–3129, Society of Exploration Geophysicists.

Hermans, T., Lesparre, N., De Schepper, G. & Robert, T., 2019. Bayesian evidential learning: a field validation using push-pull tests, *Hydrogeol. J.,* **27**(5), 1661–1672.

Hermans, T., Nguyen, F., Klepikova, M., Dassargues, A. & Caers, J., 2018. Uncertainty quantification of medium-term heat storage from short-term geophysical experiments using Bayesian evidential learning, *Water Resour. Res.,* **54**(4), 2931–2948.

Herrmann, R. & Ammon, C., 2004. Surface waves, receiver functions and crustal structure, Computer Programs in seismology: version 3.30. Saint Louis University, Saint Louis.

Herrmann, R.B., 2013. Computer programs in seismology: an evolving tool for instruction and research, *Seism. Res. Lett.,* **84**(6), 1081–1088.

Hodson, T.O., 2022. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not, *Geosci. Model Dev. Discuss.,* **2022,** 1–10.

Jolliffe, I., 2011. Principal component analysis, *International Encyclopedia of Statistical Science,* pp. 1094–1096, Springer.

Kennett, B., 1983. *Seismic Wave Propagation in Stratified Media,* Cambridge University Press.

Lai, C.G., 1998. Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization, *PhD Thesis,* Georgia Institute of Technology.

Levin, D., 1982. Procedures for computing one-and two-dimensional integrals of functions with rapid irregular oscillations, *Math. Comput.,* **38**(158), 531–538.

Levin, D., 1996. Fast integration of rapidly oscillatory functions, *J. Comput. Appl. Math.,* **67**(1), 95–101.

Liu, H., Li, J. & Hu, R., 2024. Automatic and adaptive picking of surface-wave dispersion curves for near-surface application, *J. Appl. Geophys.,* **221,** 105 282.

Liu, Q., Chen, X., Gao, L., Yu, Z. & Chen, J., 2023. Direct image dissimilarity inversion of ambient noise multimodal dispersion spectrograms, *Bull. seism. Soc. Am.,* **113**(5), 1960–1981.

Luu, K., 2024. disba: Numba-accelerated computation of surface wave dispersion,https://github.com/keurfonluu/disba.

It Ma, Y., Loures, L. & Margrave, G.F., 2004. Seismic modeling with the reflectivity method, *CREWES Res. Rep.,* **15,** 1–7.

Maraschini, M. & Foti, S., 2010. A Monte Carlo multimodal inversion of surface waves, *Geophys. J. Int.,* **182**(3), 1557–1566.

Masclet, S., Bardainne, T., Massart, V. & Prigent, H., 2019. Near surface characterization in Southern Oman: multi-wave inversion guided by machine learning. *81st EAGE Conference and Exhibition 2019, Vol. 2019,* pp. 1–5, European Association of Geoscientists and Engineers.

Miao, X., Zheng, D., Zi, L., Zhou, Z. & Gao, M., 2016. Robust multimodal surface-wave inversion for shallow velocity and shear statics, *SEG Technical Program Expanded Abstracts 2016,* pp. 4956–4960, Society of Exploration Geophysicists.

Michel, H., Hermans, T. & Nguyen, F., 2023. Iterative prior resampling and rejection sampling to improve 1-D geophysical imaging based on Bayesian evidential learning (BEL1D), *Geophys. J. Int.,* **232**(2), 958–974.

Michel, H., Nguyen, F. & Aigner, L., 2022. hadrienmichel/pyBEL1D: Latest version of pyBEL1D, Accessed: 2022-08-04.

Michel, H., Nguyen, F., Kremer, T., Elen, A. & Hermans, T., 2020. 1D geological imaging of the subsurface from geophysical data with Bayesian evidential learning, *Comput. Geosci.,* **138,** 104 456.

Morozov, I.B., 2009. Thirty years of confusion around "scattering Q"?, *Seism. Res. Lett.,* **80**(1), 5–7.

Mreyen, A.-S., Michel, H. & Nguyen, F., 2025. Full wavefield surface wave analysis with Bayesian Evidential Learning [Data set], *Geophys. J. Int.,* Zenodo, https://doi.org/10.5281/zenodo.17807710.

Mreyen, A.S., Cauchie, L., Micu, M., Onaca, A. & Havenith, H.B., 2021. Multiple geophysical investigations to characterize massive slope failure deposits: application to the Balta rockslide, Carpathians, *Geophys. J. Int.,* **225**(2), 1032–1047.

Müller, G., 1985. The reflectivity method: a tutorial, *J. Geophys.,* **58**(1), 153–174.

Niskanen, M. & Lähivaara, T., 2023. COMPOSTI: A Python-based program for seismic trans-dimensional inversion, *SoftwareX,* **21,** 101 298.

Niu, Y., Elita Li, Y. & Chen Chian, S., 2022. Full phase-velocity spectrum inversion of multimode Rayleigh waves. *Second International Meeting for Applied Geoscience and Energy,* pp. 2066–2070, Society of Exploration Geophysicists and American Association of Petroleum.

O'Neill, A., Dentith, M. & List, R., 2003. Full-waveform P-SV reflectivity inversion of surface waves for shallow engineering applications, *Explor. Geophys.,* **34**(3), 158–173.

Ortega, R., Carciumaru, D., Huesca, E. & Gutierrez, E., 2019. Automatic selection of dispersion curves based on a weighted probability scheme, *Seism. Res. Lett.,* **90**(2A), 623–632.

Pan, L., Chen, X., Wang, J., Yang, Z. & Zhang, D., 2019. Sensitivity analysis of dispersion curves of Rayleigh waves with fundamental and higher modes, *Geophys. J. Int.,* **216**(2), 1276–1303.

Park, C. & Brohammer, M., 2003. SurfSeis 1.5 V software for multichannel analysis of surface waves, Kansas Geological Survey, Kansas.

Parolai, S., Lai, C.G., Dreossi, I., Ktenidou, O.J. & Yong, A., 2022. A review of near-surface QS estimation methods using active and passive sources, *J. Seismol.,* **26**(4), 823–862.

Pedregosa, F. *et al.*, 2011. Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res.,* **12,** 2825–2830.

Pradhan, A. & Mukerji, T., 2020. Seismic Bayesian evidential learning: estimation and uncertainty quantification of sub-resolution reservoir properties, *Comput. Geosci.,* **24**(3), 1121–1140.

Ren, L., Gao, F., Wu, Y., Williamson, P., McMechan, G.A. & Wang, W., 2023. Automated dispersion curve picking using multi-attribute convolutional-neural-network based machine learning, *Geophys. J. Int.,* **232**(2), 1173–1208.

Rix, G.J. & Lai, C.G. & 1998. Simultaneous inversion of surface wave velocity and attenuation, *Geotechnical Site Characterization, Vol. 1,* pp. 503–508, Balkema.

Rovetta, D., Kontakis, A. & Colombo, D., 2021. Application of a density-based spatial clustering algorithm for fully automatic picking of surface-wave dispersion curves, *Leading Edge,* **40**(9), 678–685.

Ryden, N. & Park, C.B., 2006. Fast simulated annealing inversion of surface waves on pavement using phase-velocity spectra, *Geophysics,* **71**(4), R49–R58.

Scheidt, C., Li, L. & Caers, J., 2018. *Quantifying Uncertainty in Subsurface Systems,* Vol. **236,** John Wiley and Sons.

Sen, M.K., 2020. Seismic, reflectivity method, *Encyclopedia of Solid Earth Geophysics,* pp. 1–12, Springer.

Shapiro, S. & Kneib, G., 1993. Seismic attenuation by scattering: theory and numerical results, *Geophys. J. Int.,* **114**(2), 373–391.

Shi, H., Yang, Y., Wang, L., Ma, D., Beg, M.F., Pei, J. & Cao, J., 2022. Two-dimensional functional principal component analysis for image feature extraction, *J. Comput. Graph. Stat.,* **31**(4), 1127–1140.

Socco, L.V., Boiero, D., Foti, S. & Wisén, R., 2009. Laterally constrained inversion of ground roll from seismic reflection records, *Geophysics,* **74**(6), G35–G45.

Song, W., Feng, X., Wu, G., Zhang, G., Liu, Y. & Chen, X., 2021. Convolutional neural network, Res-Unet++,-based dispersion curve picking from noise cross-correlations, *J. Geophys. Res.: Solid Earth,* **126**(11), e2021JB022027.

Song, W., Feng, X., Zhang, G., Gao, L., Yan, B. & Chen, X., 2022. Domain adaptation in automatic picking of phase velocity dispersions based on deep learning, *J. Geophys. Res.: Solid Earth,* **127**(6), e2021JB023389.

Tadjer, A. & Bratvold, R.B., 2021. Managing uncertainty in geological CO2 storage using Bayesian evidential learning, *Energies,* **14**(6), 1557.

Taipodia, J., Dey, A., Gaj, S. & Baglari, D., 2020. Quantification of the resolution of dispersion image in active MASW survey and automated extraction of dispersion curve, *Comput. Geosci.,* **135,** 104 360.

Thibaut, R., Laloy, E. & Hermans, T., 2021. A new framework for experimental design using Bayesian Evidential Learning: the case of wellhead protection area, *J. Hydrol.,* **603,** 126 903.

Thomson, W.T., 1950. Transmission of elastic waves through a stratified solid medium, *J. Appl. Phys.,* **21**(2), 89–93.

Tipping, M.E. & Bishop, C.M., 1999. Probabilistic principal component analysis, *J. R. Stat. Soc. Ser. B: Stat. Method.,* **61**(3), 611–622.

Tokimatsu, K., Tamura, S. & Kojima, H., 1992. Effects of multiple modes on Rayleigh wave dispersion characteristics, *J. Geotech. Eng.,* **118**(10), 1529–1543.

Van Der Baan, M., 2002. Constant Q and a fractal, stratified earth, *Seismic Waves in Laterally Inhomogeneous Media,* pp. 1707–1718, Springer Nature.

Vantassel, J.P. & Cox, B.R., 2022. SWprocess: a workflow for developing robust estimates of surface wave dispersion uncertainty, *J. Seismol.,* **26**(4), 731–756.

Wand, M.P. & Jones, M.C., 1993. Comparison of smoothing parameterizations in bivariate kernel density estimation, *J. Am. Stat. Assoc.,* **88**(422), 520–528.

Wang, Z., Sun, C. & Wu, D., 2021. Automatic picking of multi-mode surface-wave dispersion curves based on machine learning clustering methods, *Comput. Geosci.,* **153,** 104 809.

Wathelet, M., Chatelain, J.L., Cornou, C., Giulio, G.D., Guillier, B., Ohrnberger, M. & Savvaidis, A., 2020. Geopsy: a user-friendly open-source tool set for ambient vibration processing, *Seism. Res. Lett.,* **91**(3), 1878–1889.

Xia, J., Miller, R.D. & Park, C.B., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics,* **64**(3), 691–700.

Xia, J., Miller, R.D., Park, C.B. & Tian, G., 2003. Inversion of high frequency surface waves with fundamental and higher modes, *J. Appl. Geophys.,* **52**(1), 45–57.

Xia, J., Xu, Y., Miller, R.D. & Ivanov, J., 2012. Estimation of near-surface quality factors by constrained inversion of Rayleigh-wave attenuation coefficients, *J. Appl. Geophys.,* **82,** 137–144.

Xu, X., Tian, Y., Wu, D., Xie, J., Wang, Z. & Zhang, T., 2024. Near-surface velocity inversion and modeling method based on surface waves in petroleum exploration: a case study from Qaidam Basin, China, *Front. Earth Sci.,* **12,** 1379 668.

Yan, Y., Chen, X., Huai, N. & Guan, J., 2022. Modern inversion workflow of the multimodal surface wave dispersion curves: staging strategy and pattern search with embedded Kuhn–Munkres algorithm, *Geophys. J. Int.,* **231**(1), 47–71.

Yang, H.Q., Chu, J., Qi, X., Wu, S. & Chiam, K., 2023. Bayesian evidential learning of soil-rock interface identification using boreholes, *Comput. Geotech.,* **162,** 105 638.

Yang, S., Zhang, H., Gu, N., Gao, J., Xu, J., Jin, J., Li, J. & Yao, H., 2022a. Automatically extracting surface-wave group and phase velocity dispersion curves from dispersion spectrograms using a convolutional neural network, *Seismol. Res. Lett.,* **93**(3), 1549–1563.

Yang, T., Xu, Y., Cao, D., Nan, F., Du, N. & Hou, Z., 2022b. SDCnet: An Unet with residual blocks for extracting dispersion curves from seismic data, *Comput. Geosci.,* **166,** 105 183.

Yao, H., Cao, W., Huang, X., Li, L. & Wu, B., 2023. Automatic extraction of surface wave dispersion curves using unsupervised learning and high-resolution tau-p transform, *Earth Space Sci.,* **10**(12), e2023EA003198.

Yin, Z., Strebelle, S. & Caers, J., 2020. Automated Monte Carlo-based quantification and updating of geological uncertainty with borehole data (AutoBEL v1. 0), *Geosci. Model Develop.,* **13**(2), 651–672.

Zhang, X., Jia, Z., Ross, Z.E. & Clayton, R.W., 2020a. Extracting dispersion curves from ambient noise correlations using deep learning, *IEEE Trans. Geosci. Remote Sens*., **58**(12), 8932–8939.

Zhang, X., Zheng, Y. & Curtis, A., 2023. Surface wave dispersion inversion using an energy likelihood function, *Geophys. J. Int.,* **232**(1), 523–536.

Zhang, Z., Alkhalifah, T. & Liu, Y., 2024. Full dispersion-spectrum inversion of surface waves, *J. Geophys. Res.: Solid Earth,* **129**(6), e2023JB028469.

Zhang, Z.d., Alajami, M. & Alkhalifah, T., 2020b. Wave-equation dispersion spectrum inversion for near-surface characterization using fibre-optics acquisition, *Geophys. J. Int.,* **222**(2), 907–918.

Zheng, D. & Miao, X., 2014. Multimodal Rayleigh wave dispersion curve picking and inversion to build near surface shear wave velocity models, *76th EAGE Conference and Exhibition-Workshops,* pp. cp–401, European Association of Geoscientists and Engineers.

Zywicki, D.J. & Rix, G.J., 2005. Mitigation of near-field effects for seismic surface wave velocity estimation with cylindrical beamformers, *J. Geotech. Geoenviron. Eng.,* **131**(8), 970–977.

Zywicki, D.J., 1999. Advanced signal processing methods applied to engineering analysis of seismic surface waves, *PhD thesis,* Georgia Institute of Technology.

APPENDIX A:

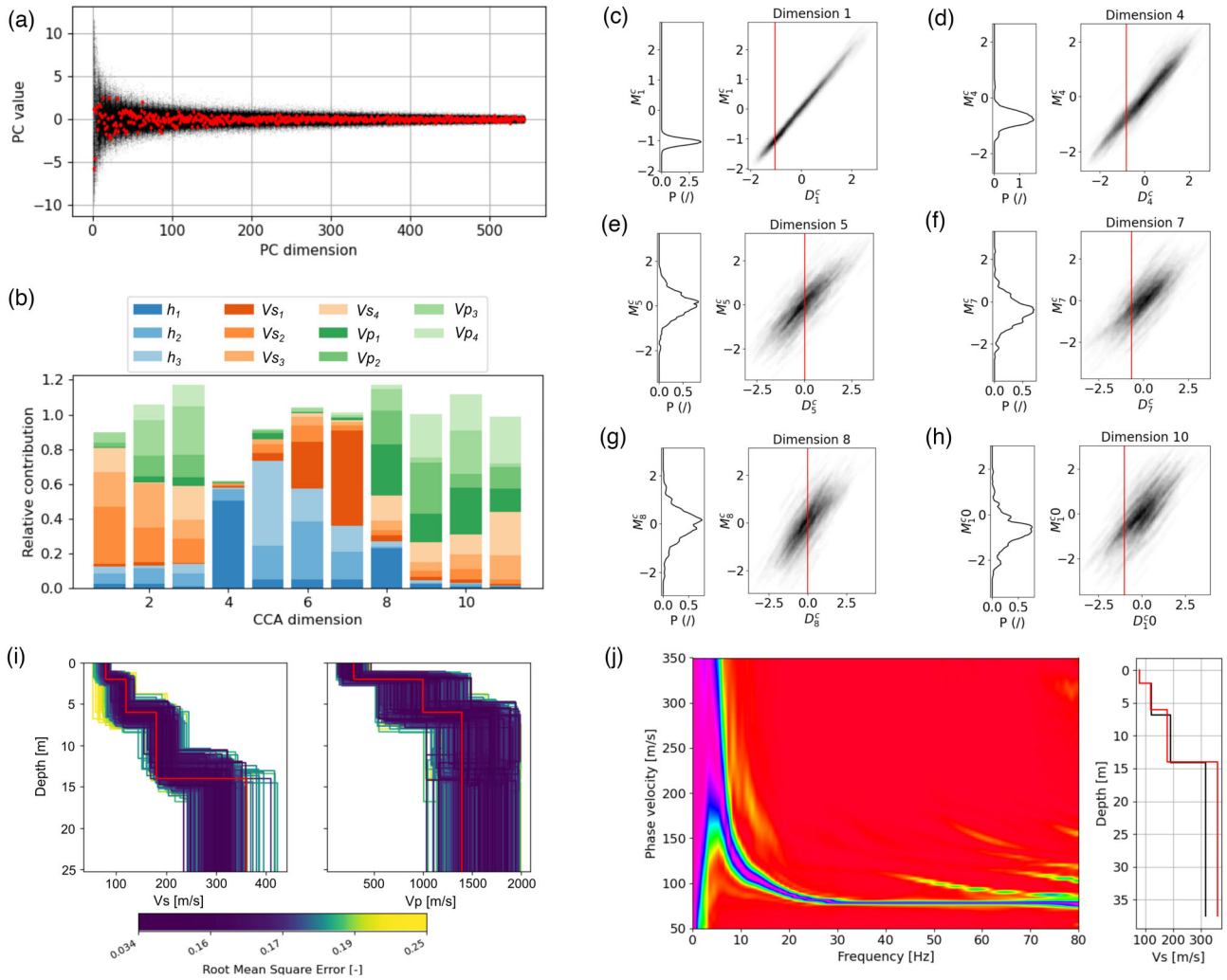

Figure A1. Model 1: (a) PCA space with black dots showing the modelled data and red dots the experimental data; (b) CCA space; (c–h) KDE and projections in CCA dimensions 1, 4, 5, 7, 8 and 10 (see Fig. A2 for all dimensions), with red lines marking the position of the experimental data; (i) ensemble of posterior solutions for V_s and V_p, coloured by normalized RMS error (red profiles = benchmark); (j) forward modelled velocity spectrum from the best posterior solution, here shown for V_s over depth (red profile = benchmark).

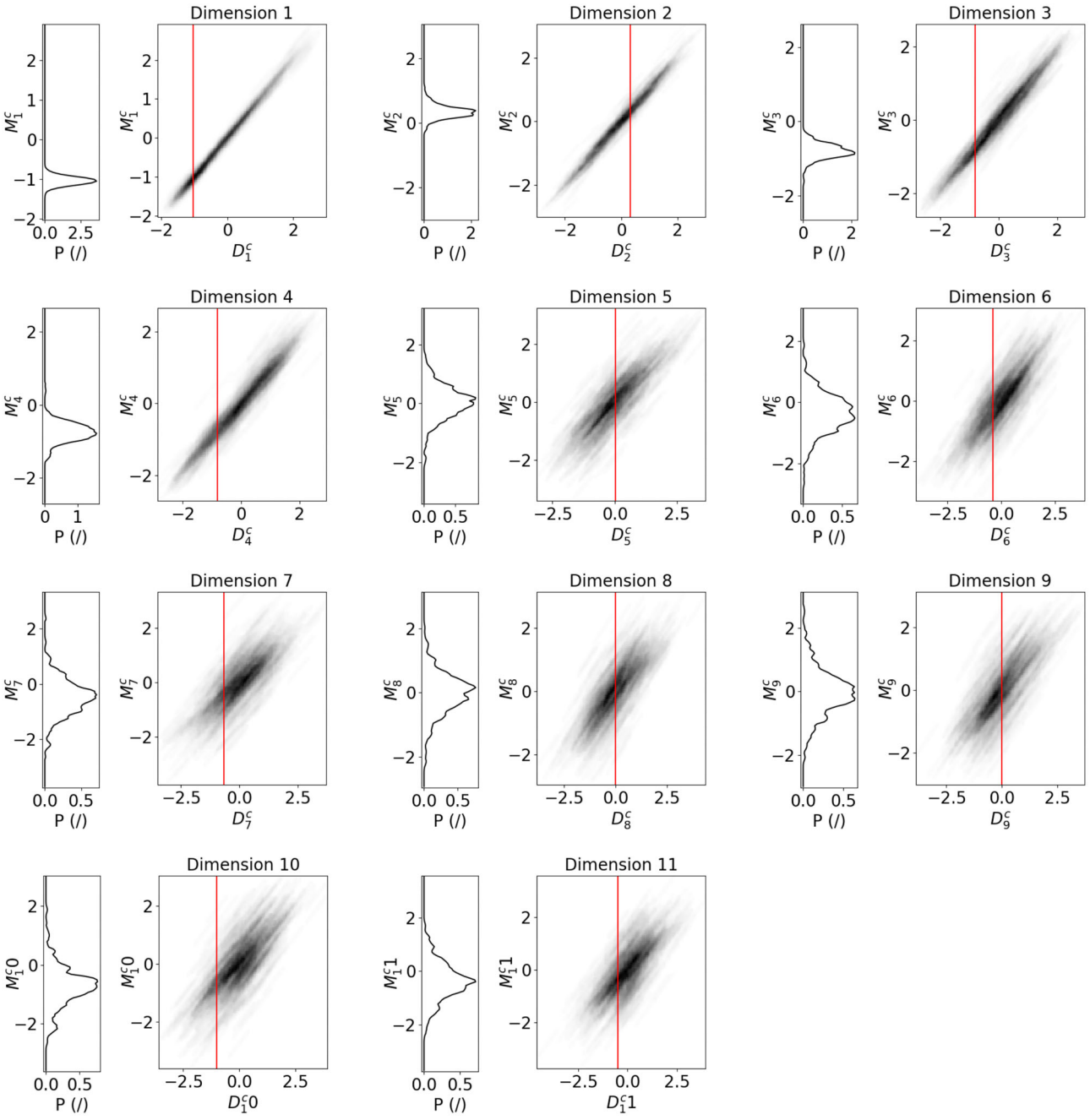

Figure A2. CCA dimensions 1 to 11 of model 1 (cf. Fig. A1).

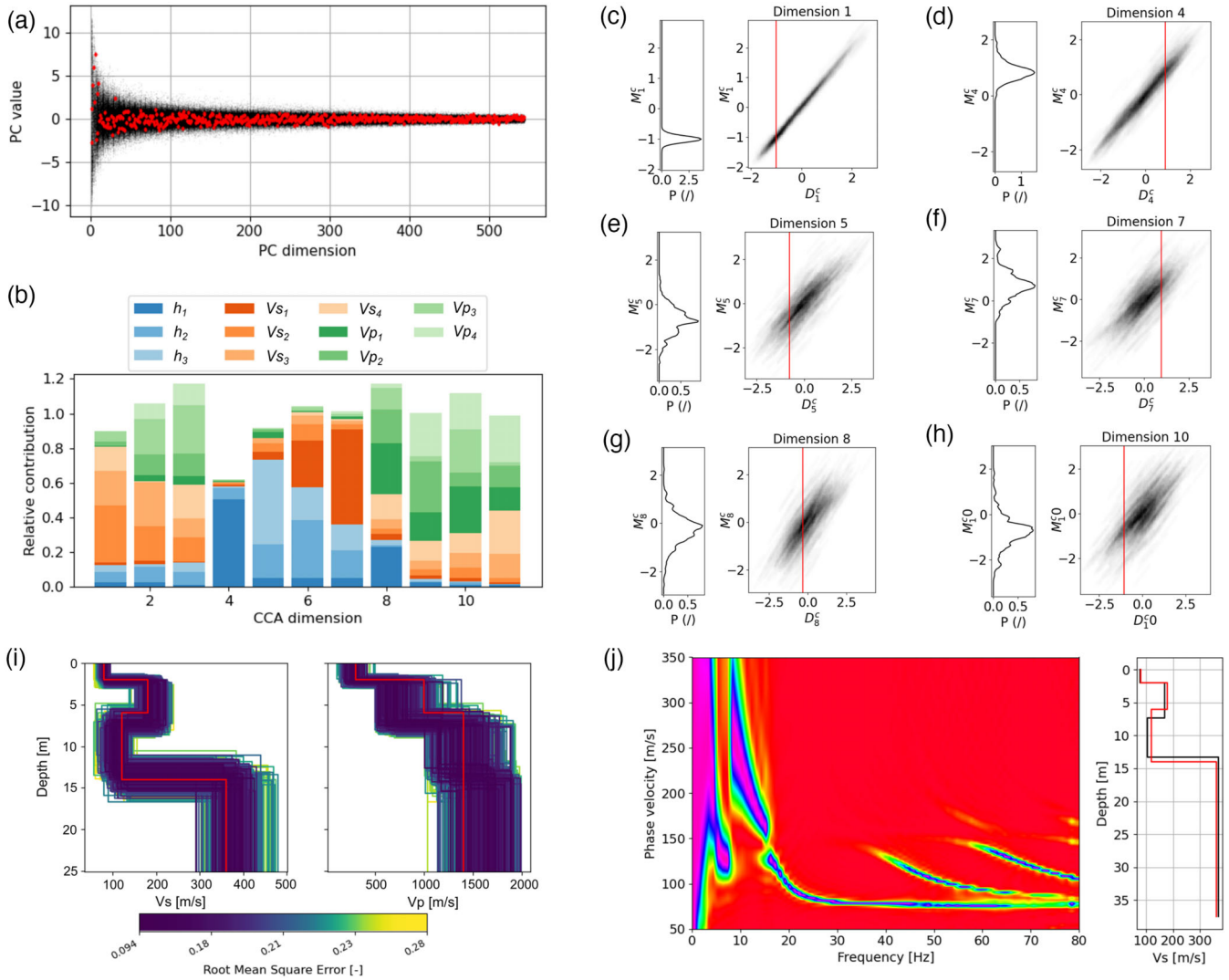

Figure A3. Model 3: (a) PCA space with black dots showing the modelled data and red dots the experimental data; (b) CCA space; (c–h) KDE and projections in CCA dimensions 1, 4, 5, 7, 8 and 10 (see Fig. A4 for all dimensions), with red lines marking the position of the experimental data; (i) ensemble of posterior solutions for V_s and V_p, coloured by normalized RMS error (red profiles = benchmark); (j) forward modelled velocity spectrum from the best posterior solution, here shown for V_s over depth (red profile = benchmark).

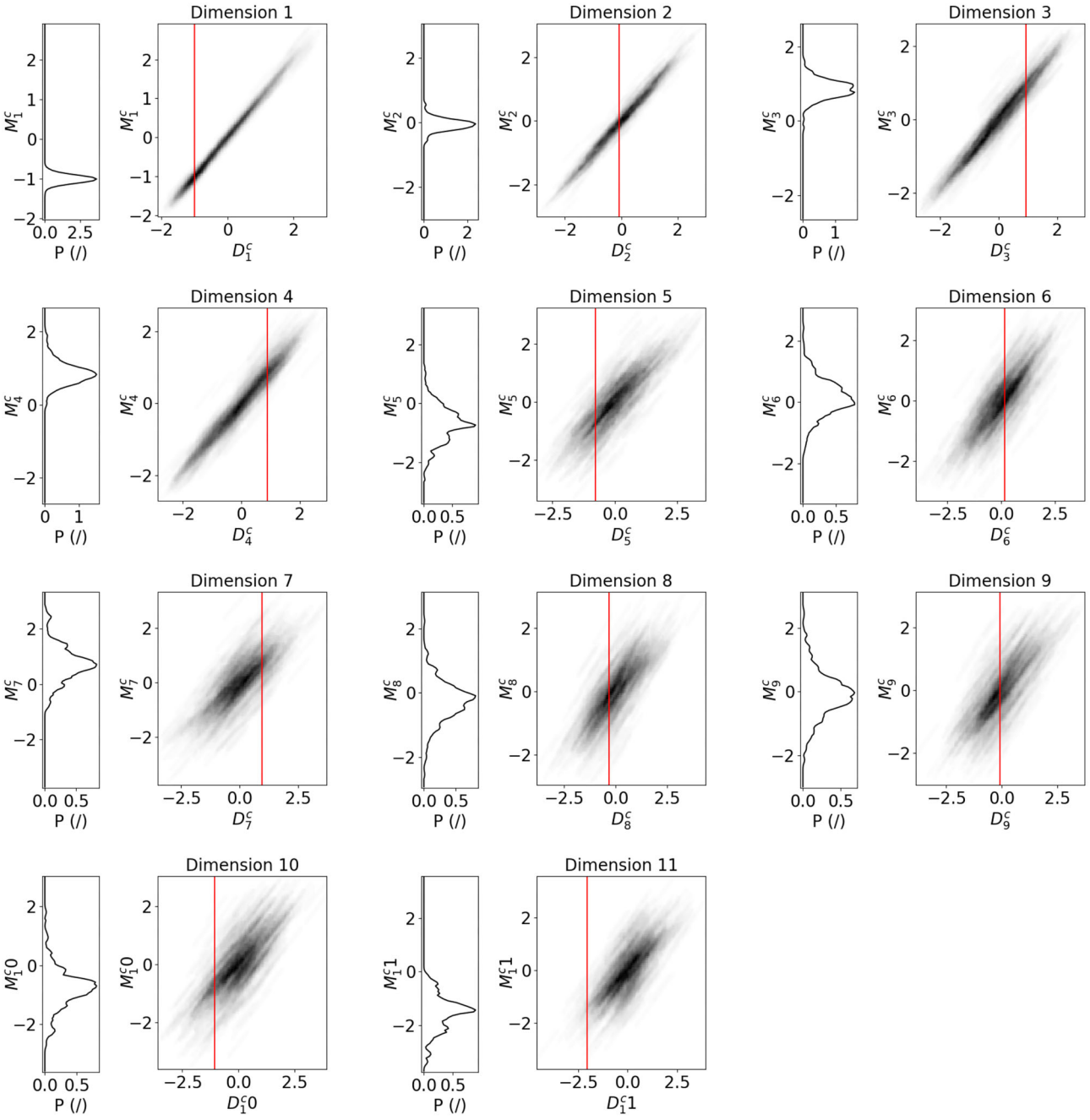

Figure A4. CCA dimensions 1 to 11 of model 3 (cf. Fig. A3).

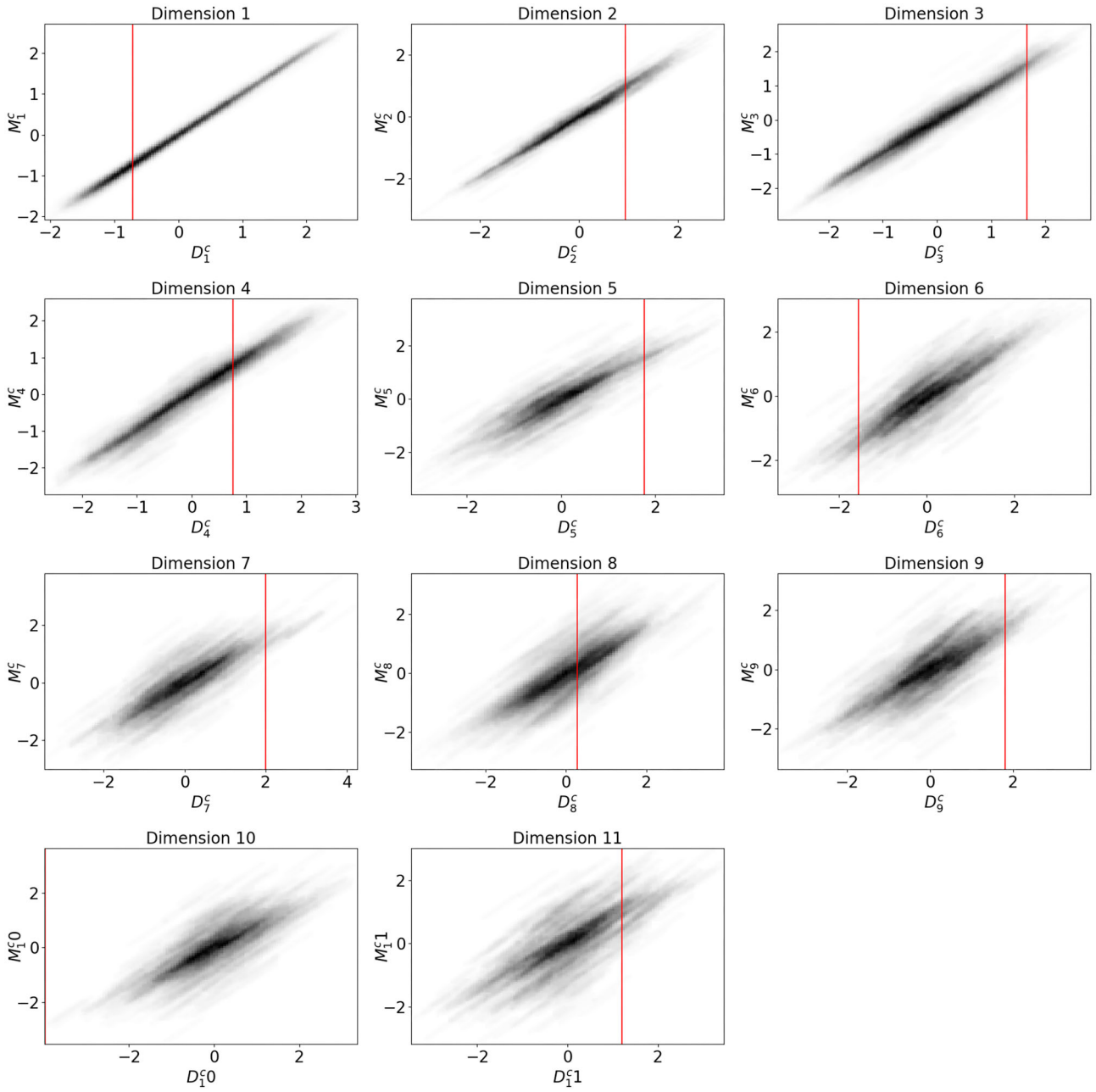

Figure A5. CCA dimensions 1 to 11 of an out-of-prior model ($Vs_1 = 310$ m s^{-1}, Vs_2 to $Vs_4 = 300$ m s^{-1}, Vp_1 to $Vp_4 = 2500$ m s^{-1}; thicknesses and density unaltered) projected in the correlated CCA space trained for model 2. We can see, that the data projection is not able to match the correlated CCA space for dimension 10; as a result no posterior model ensemble can be extracted. Note, dimension 10 is dominated by a mixture of Vs_1 to Vs_4 and Vp_1 to Vp_3 (see CCA space in Fig. 3) that are overall sensitive parameters to the model 2.

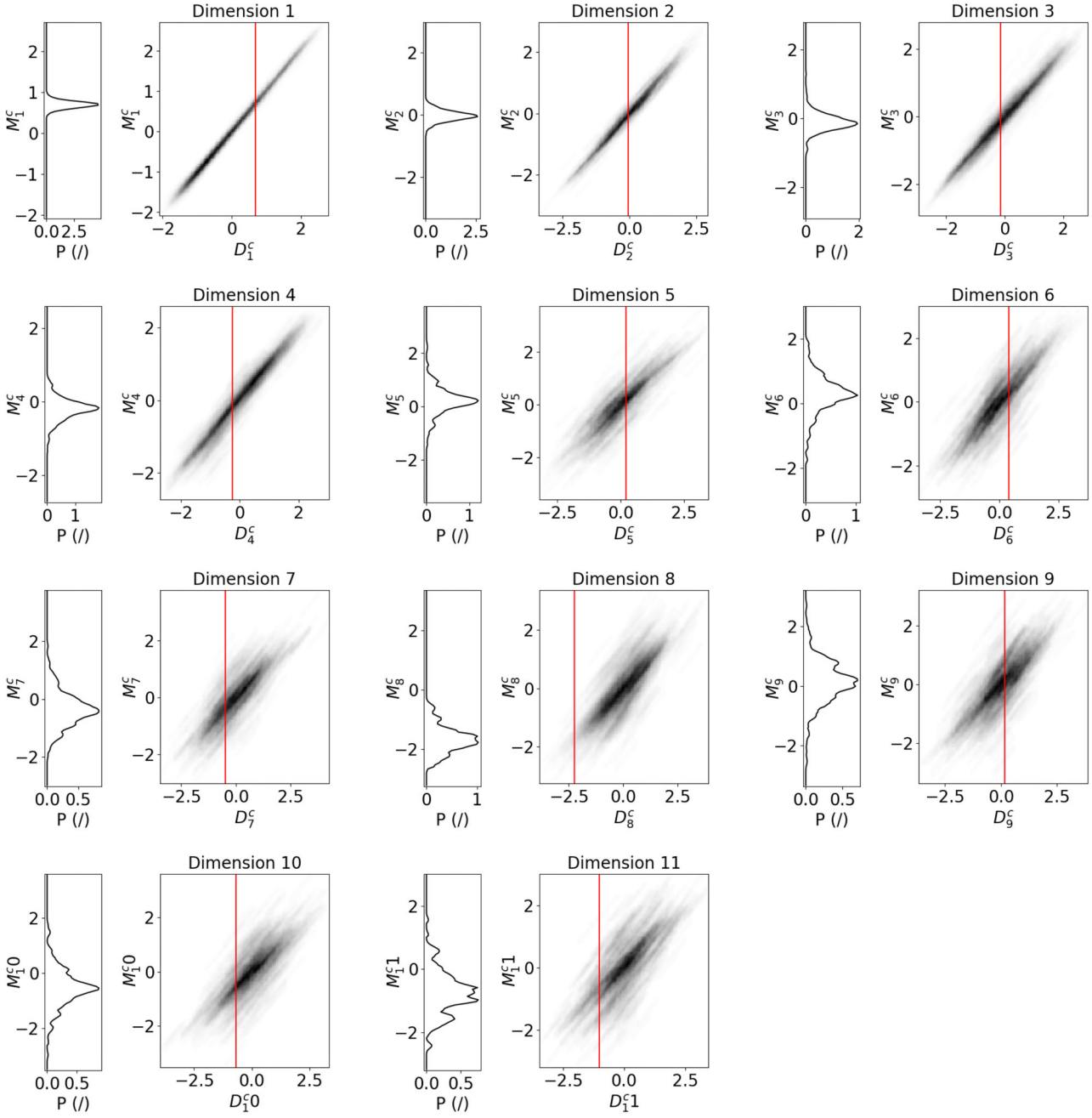

Figure A6. CCA dimensions 1 to 11 of model 2 (cf. Fig. 3).

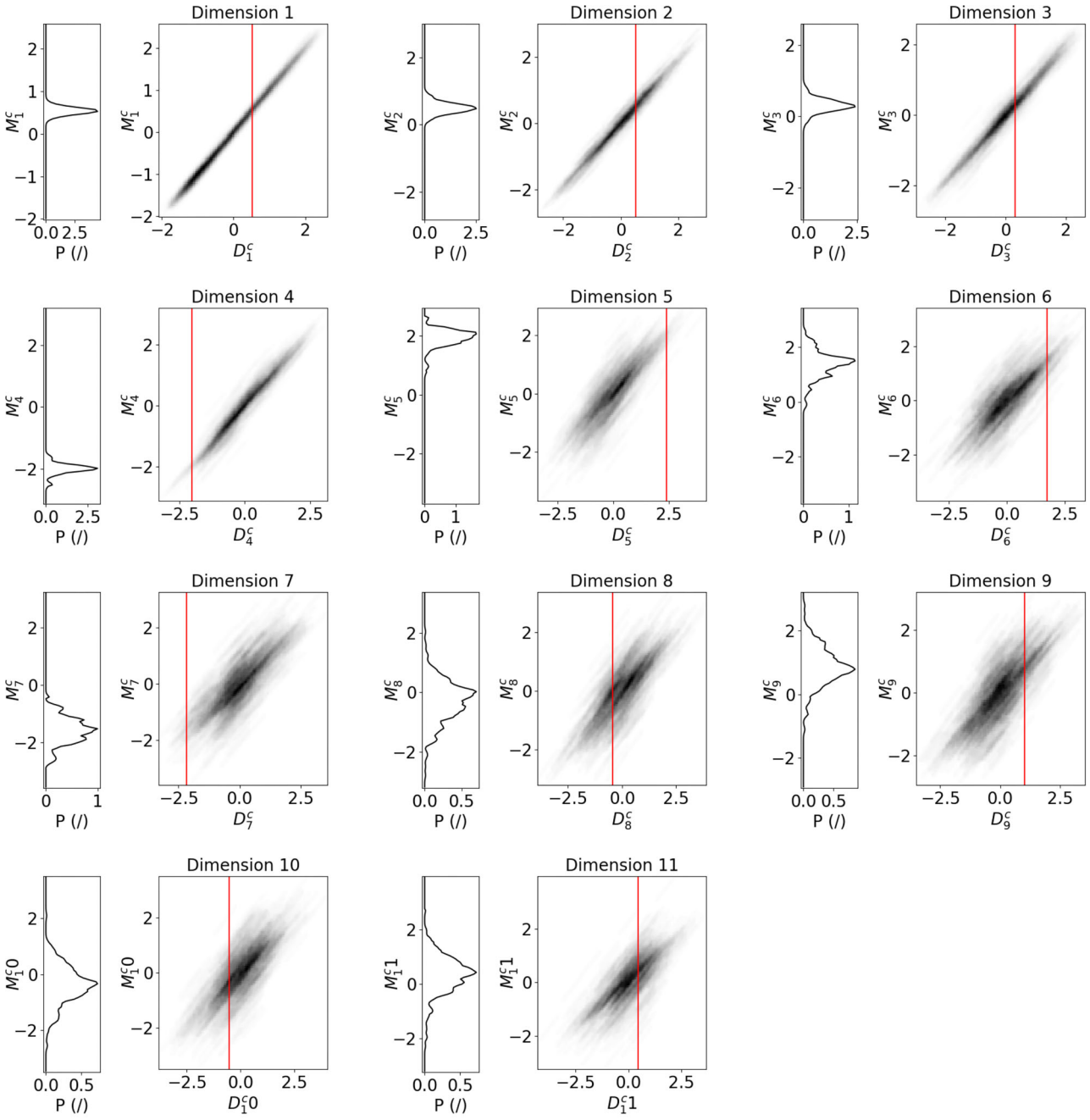

Figure A7. CCA dimensions 1 to 11 of model 2 with Gaussian noise (cf. Fig. 4).

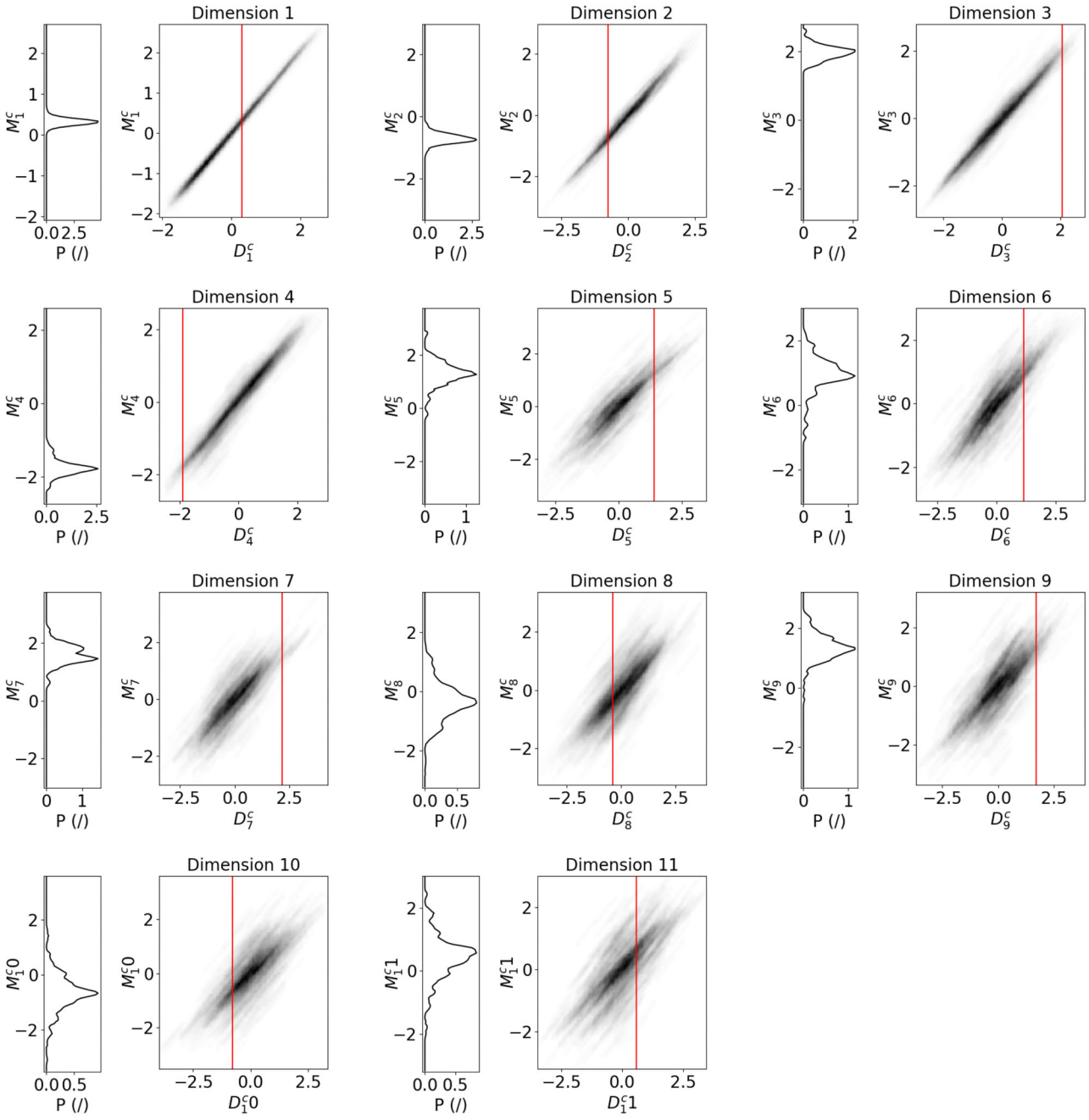

Figure A8. CCA dimensions 1 to 11 of model 2 with consistent noise projected in a clean prior (sound wave of 330 m s^{-1}; cf. Fig. 5).

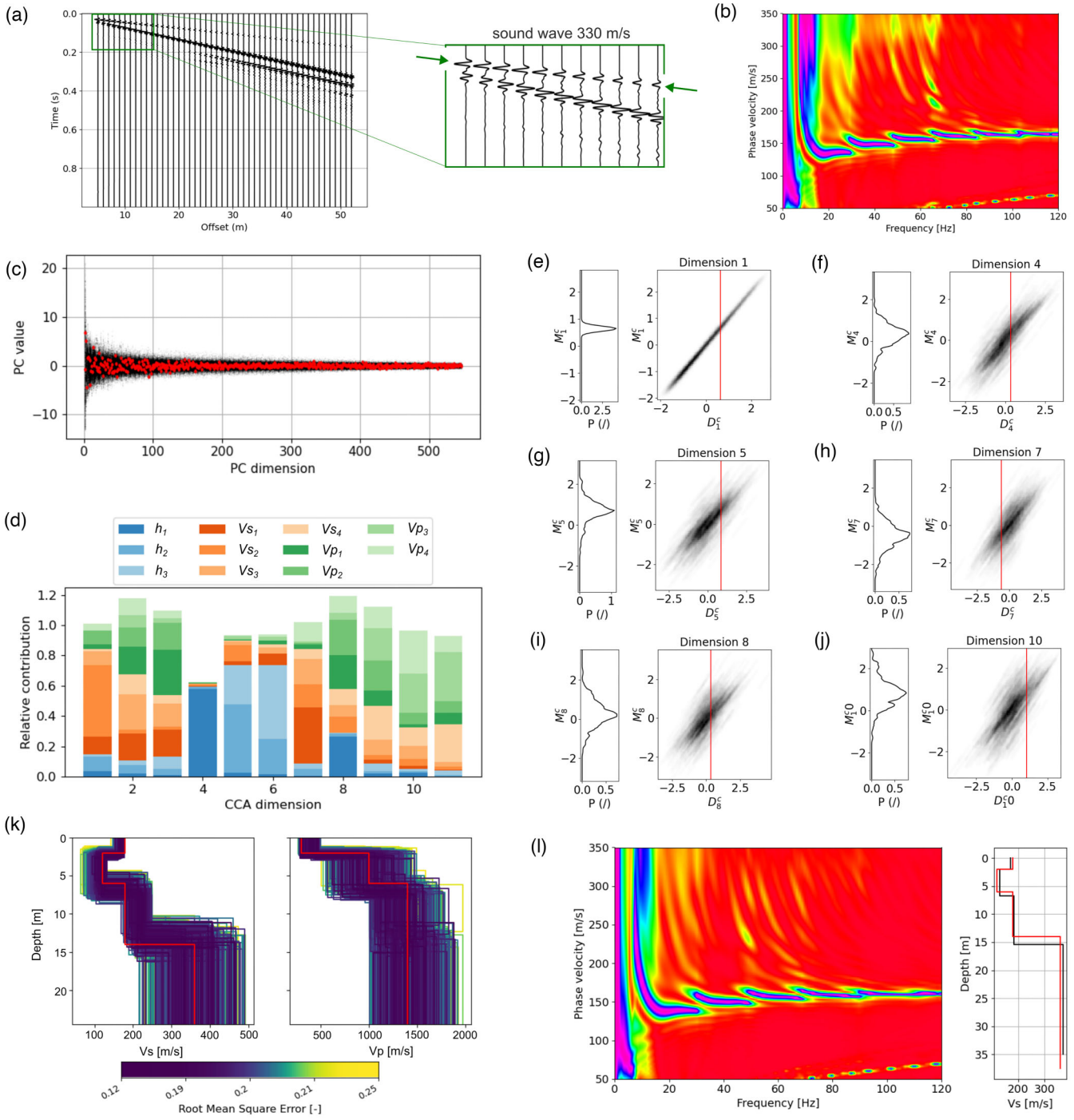

Figure A9. Model 2 with consistent noise projected in a noise trained prior: (a) shot gather with zoom to illustrate the integrated sound wave (Vp=330 m s^{-1}); (b) velocity spectrum from noise disturbed shot gather; (c) PCA space with black dots showing the modelled data and red dots the experimental data; (d) CCA space; (e–j) KDE and projection in the CCA dimensions 1, 4, 5, 7, 8 and 10 (see Fig. A10 for all dimensions), with red lines marking the position of the experimental data; (k) ensemble of posterior solutions for Vs and Vp coloured by normalized RMS error (red profiles = benchmark); (l) forward modelled velocity spectrum from the best posterior solution, here shown for Vs over depth (red profile = benchmark).

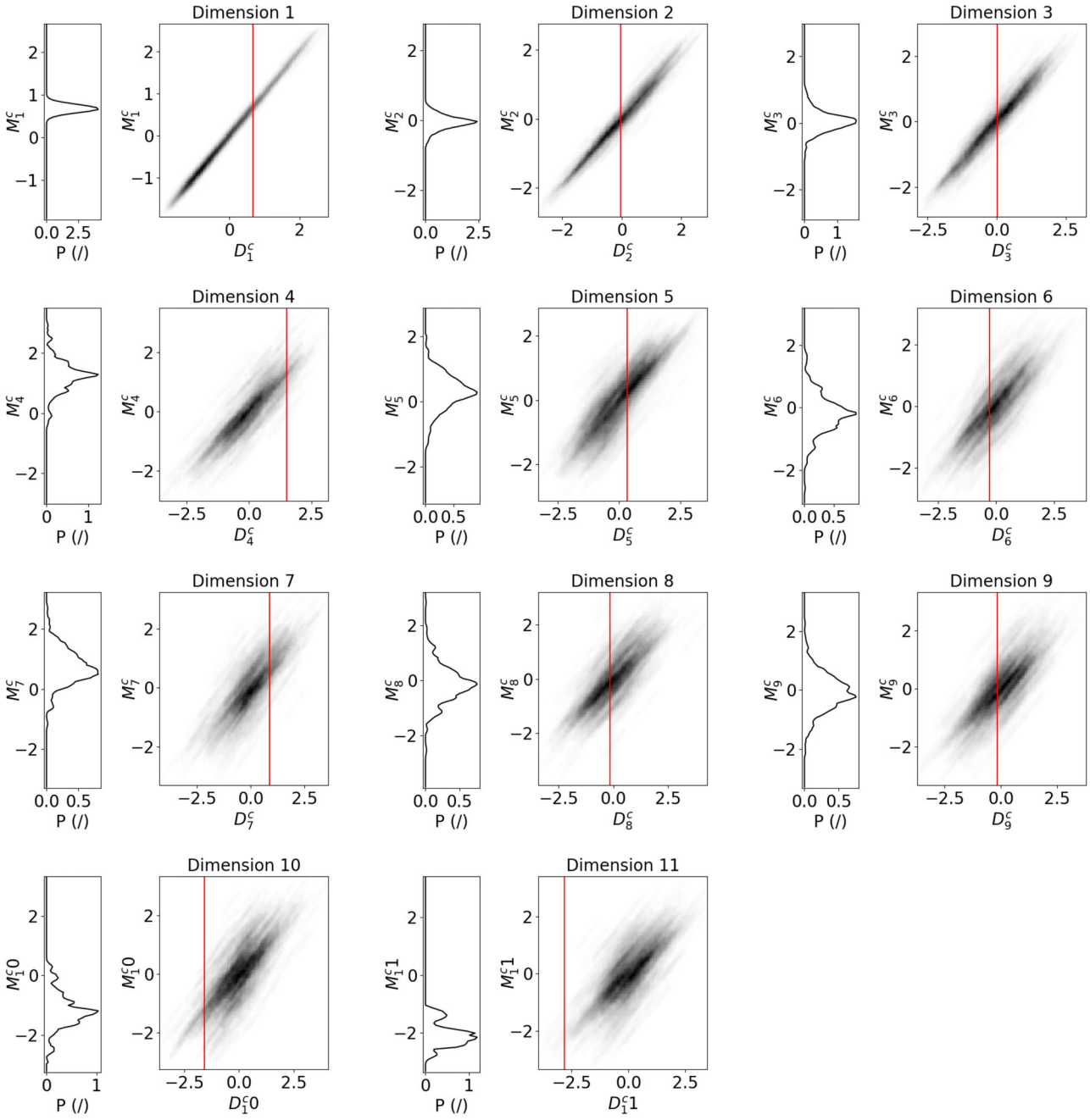

Figure A10. CCA dimensions 1 to 11 of model 2 with consistent noise in a noise trained prior (sound wave of 330 m s^{-1}; cf. Fig. 5).

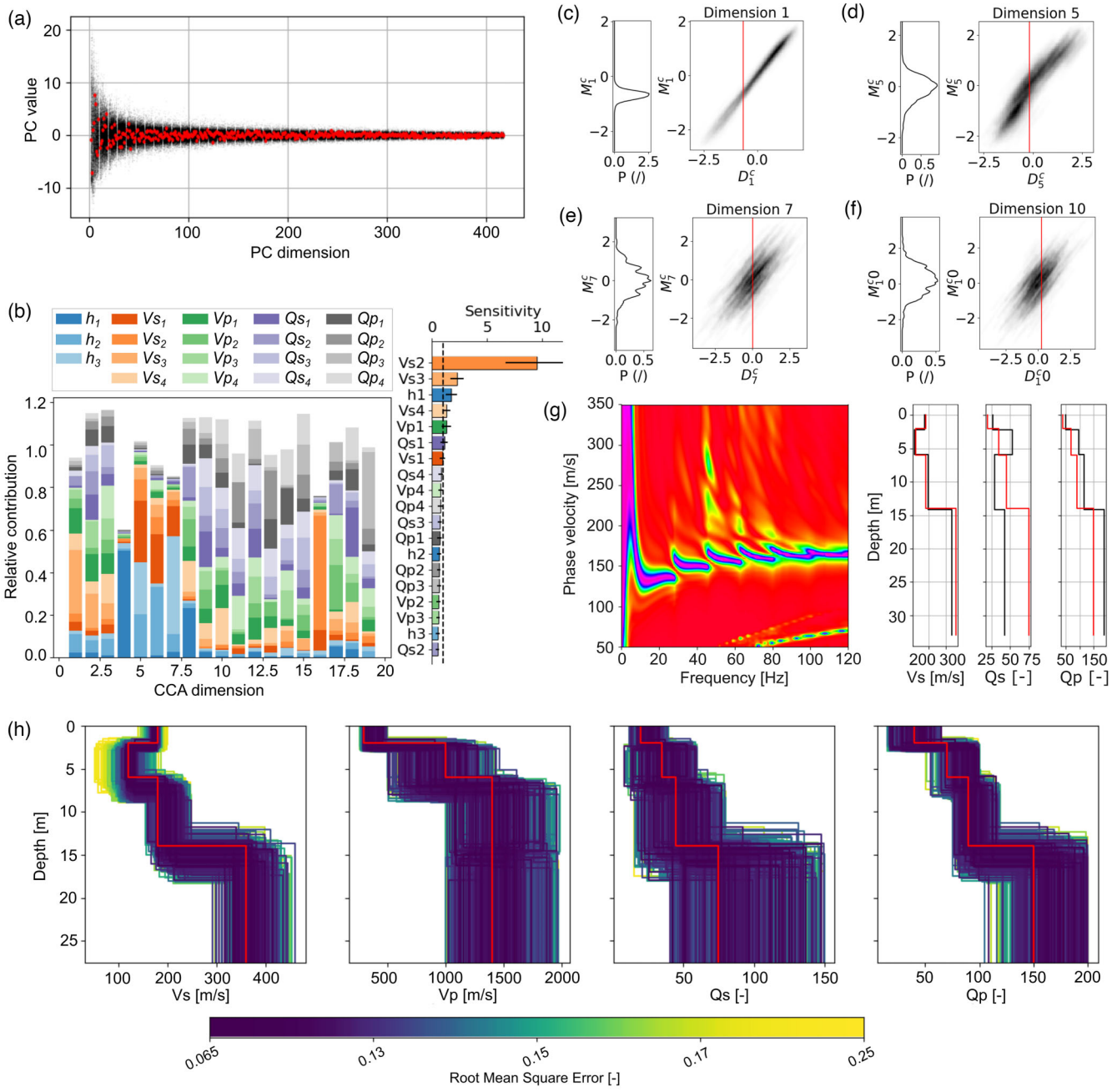

Figure A11. Model 2 in an anelastic milieu with frequency-maximum normalization (used as prior for the anelastic model, cf. Fig. 6): (a) PCA space with black dots for the modelled and red dots for the experimental data; (b) CCA space with parameter sensitvity analysis; (c–f) KDE and projection in the CCA dimensions 1, 5, 7 and 10, with red lines marking the position of the experimental data; (g) forward modelled velocity spectrum from the best posterior solution, here shown for Vs, Qs and Qp over depth (red profile = benchmark); (h) ensemble of posterior solutions for Vs, Vp, Qs and Qp, coloured by normalized RMS error (red profiles = benchmark).

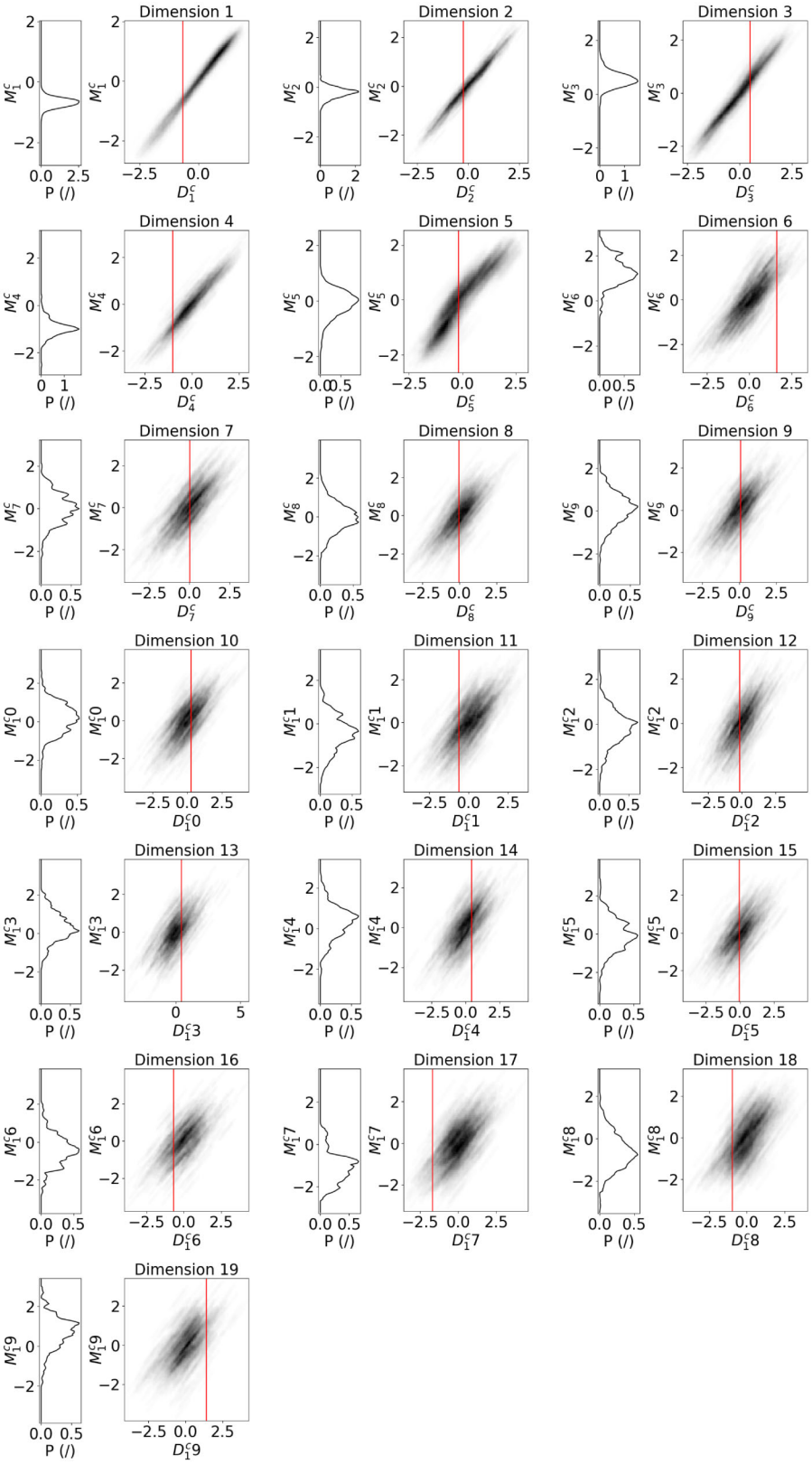

Figure A12. CCA dimensions 1 to 19 of the anelastic model 2 with frequency-maximum normalization (cf. Fig. A11).

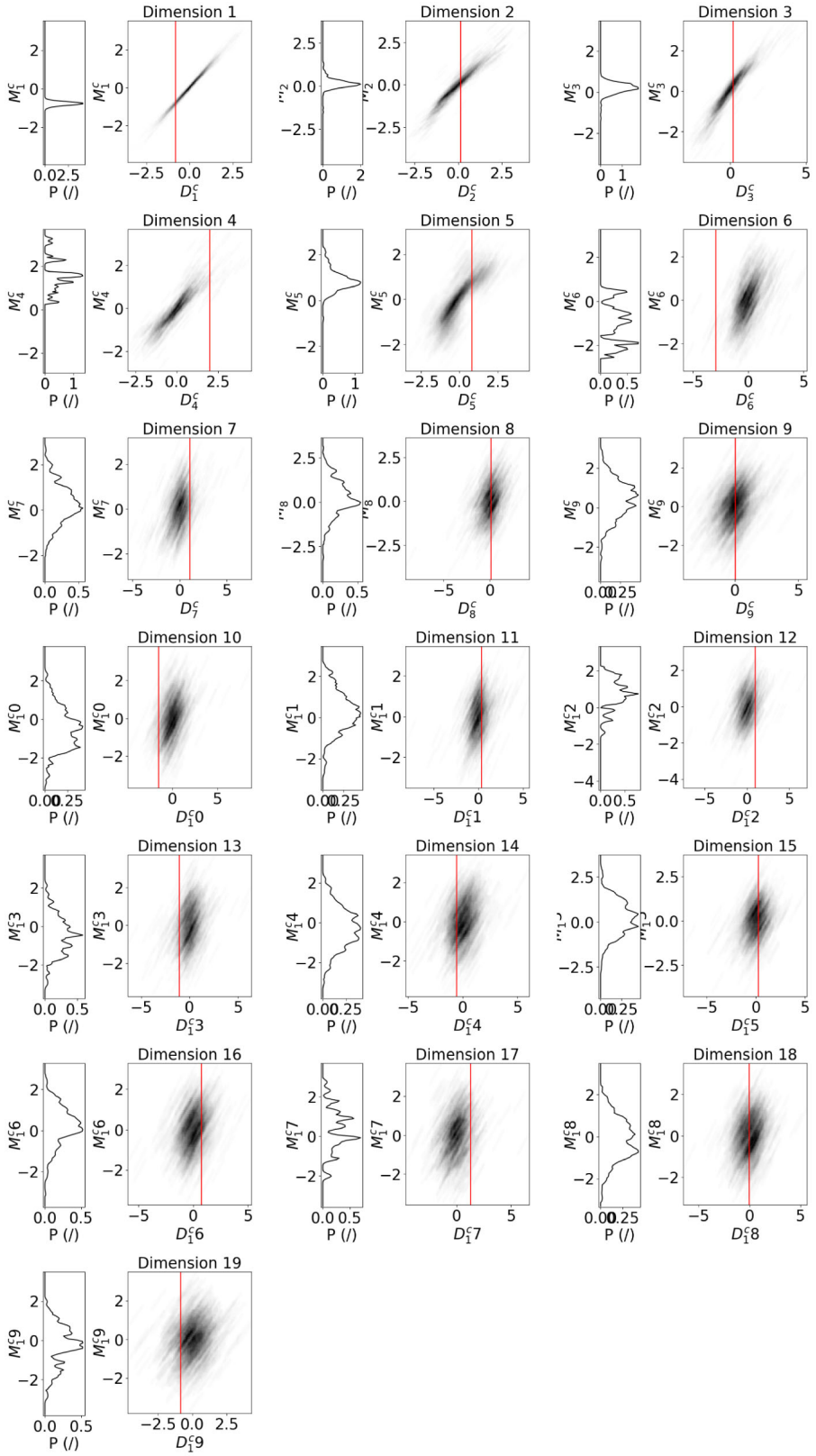

Figure A13. CCA dimensions 1 to 19 of the anelastic model 2 with absolute-maximum normalization (cf. Fig. 6).

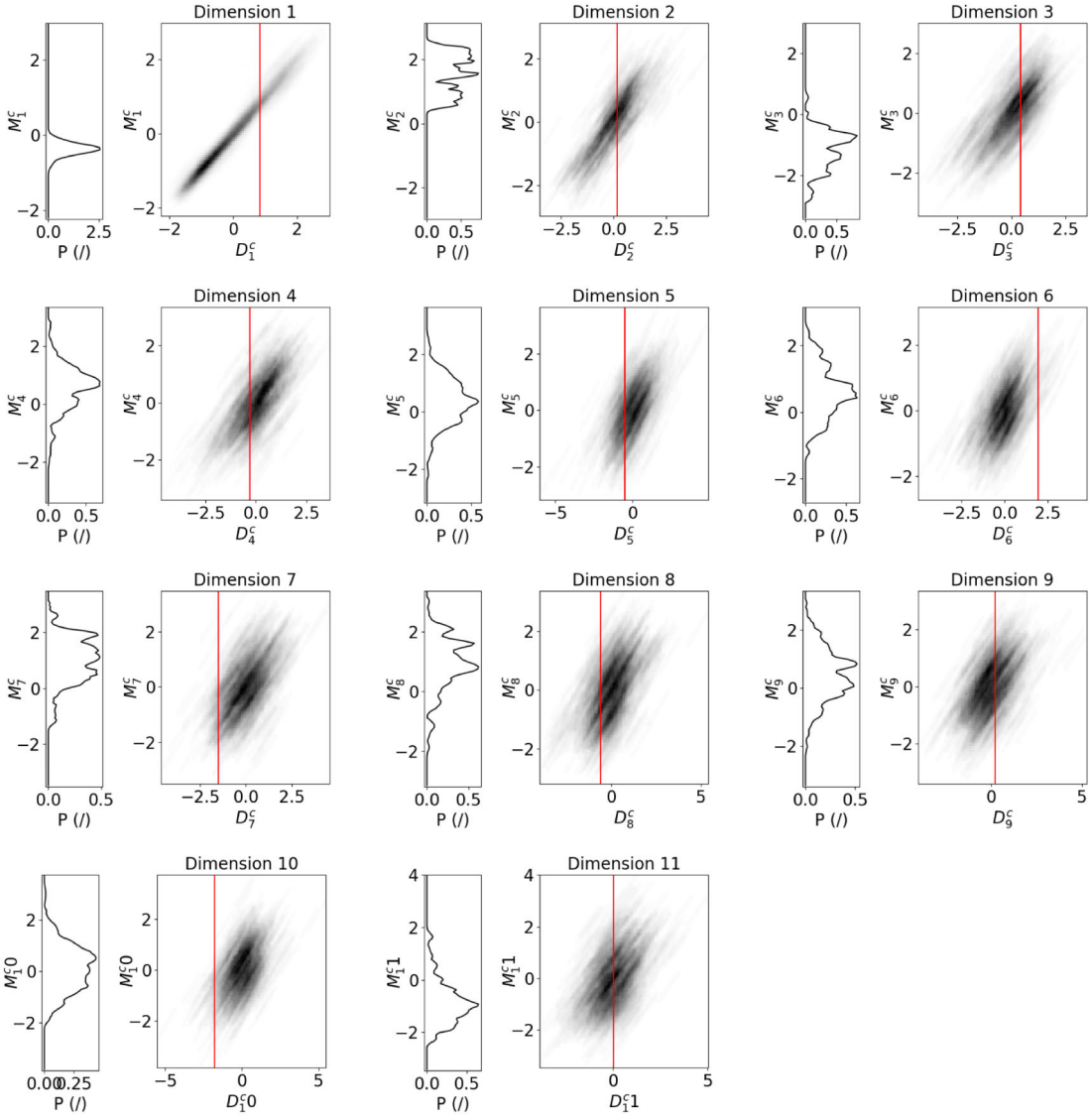

Figure A14. CCA dimensions 1 to 11 of the real data application from Eppinger *et al.* (2024) (cf. Fig. 7).

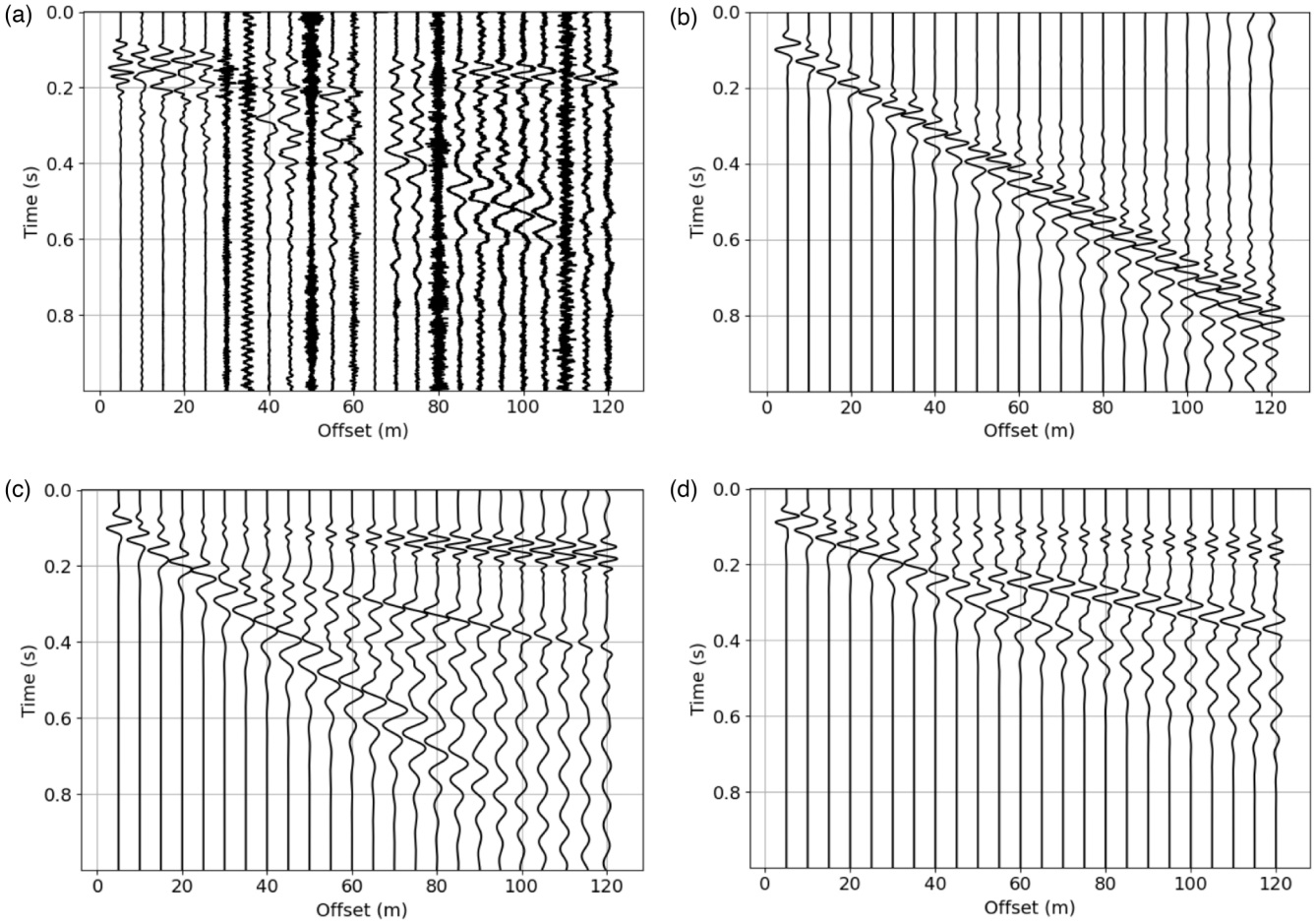

Figure A15. Comparison of wavefields of real data application from Mreyen *et al.* (2021): (a) real shot gather; (b) elastic forward simulation of benchmark model; (c) anelastic forward simulation of benchmark model with $Qs = [5, 10, 20, 30]$; (d) forward simulation of best posterior model with BEL1D.

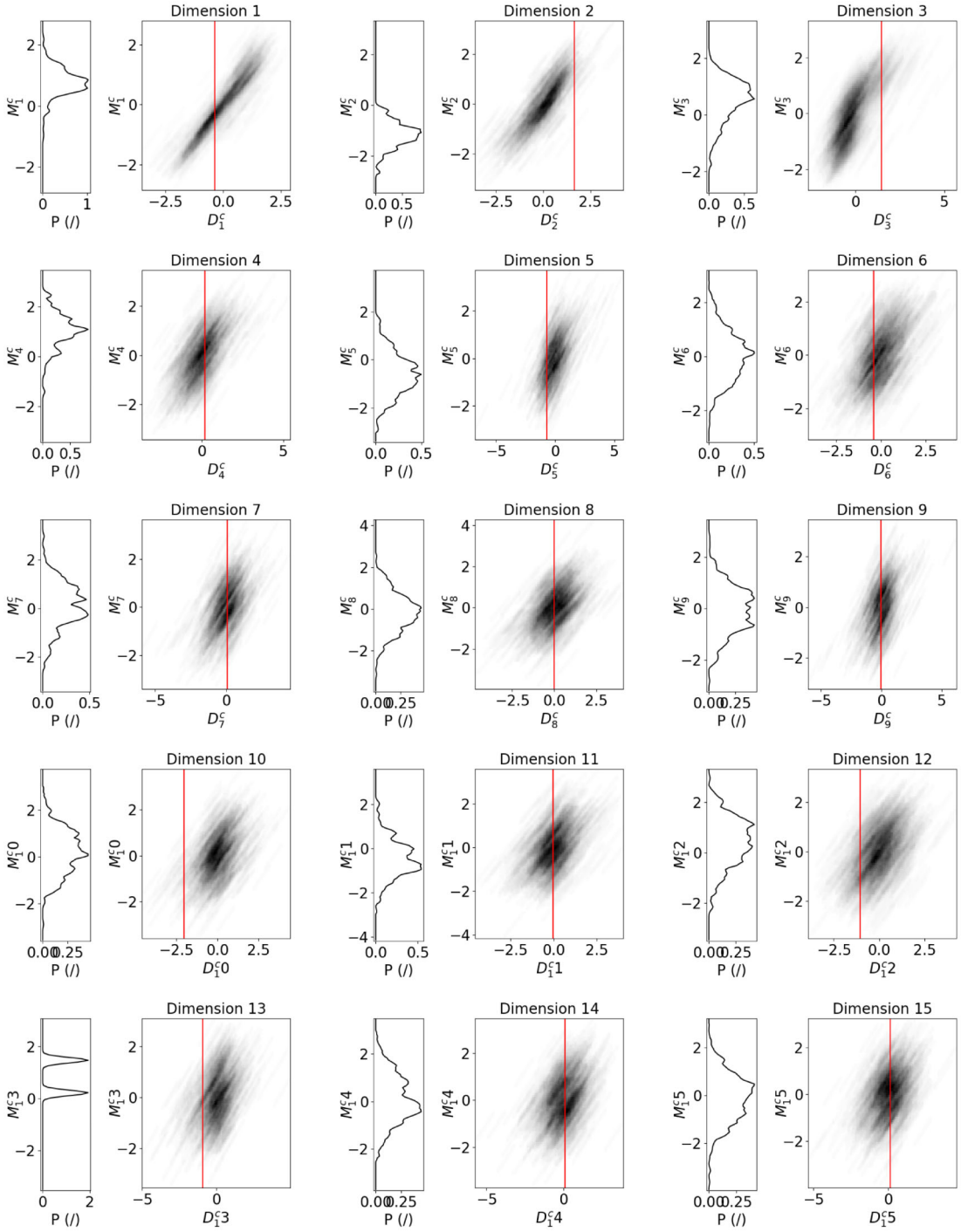

Figure A16. CCA dimensions 1 to 15 of the real data application from Mreyen *et al.* (2021).