

Assessment of the impact of noise magnitude and bandwidth variations on a probabilistic inversion of seismic data

Hamed Heidari¹, Rasmus Bødker Madsen², Hamed Amini³, Thomas Mejer Hansen⁴,
Mohammad Emami Niri⁵, and Navid Amini⁶

¹University of Tehran

²Geological Survey of Denmark and Greenland

³AkerBP

⁴Aarhus University, University of Aarhus

⁵Institute of Petroleum Engineering, College of Engineering, University of Tehran

⁶CoCoLink, subsidiary of Seoul National University, Republic of Korea

November 28, 2022

Abstract

Accounting for an accurate noise model is essential when dealing with real data which are noisy due to the effect of environmental noise, failures and limitations in data acquisition and processing. Quantifying the noise model is a challenge for practitioners in formulating an inverse problem and usually, a simple Gaussian noise model is assumed as a white noise model. Here we propose a pragmatic approach to using an estimated seismic wavelet to capture the correlated noise model (coloured noise) for the processed reflection seismic data. We test the method for a probabilistic sampling-based inversion where post-stack seismic data, associated with a hard carbonate reservoir in southwest Iran, is inverted directly to porosity. We assume eight different scenarios for the bandwidth and the magnitude of the noise. The investigation of the corresponding posterior statistics shows that ignoring the correlation of the noise samples in the noise covariance matrix generates unrealistic features in porosity realisations while underestimating the noise magnitude leads to overfitting the data and generating a biased model with low uncertainty. Furthermore, by considering an imperfect bandwidth for the noise model, the error is propagated to the posterior realisations. These issues are resolved considerably when the correlated noise is considered in the inversion. Therefore, in real data applications where the estimation of the magnitude and correlations of the noise is not trivial, the estimated seismic wavelet provides a good proxy for describing the correlation of the noise samples or equivalently the bandwidth of the noise model. In addition, it might be better to overestimate the noise magnitude than to underestimate it. This is true especially for an uncorrelated noise model and to a lesser degree also for the correlated noise model.

ABSTRACT

Accounting for an accurate noise model is essential when dealing with real data which are noisy due to the effect of environmental noise, failures and limitations in data acquisition and processing. Quantifying the noise model is a challenge for practitioners in formulating an inverse problem and usually, a simple Gaussian noise model is assumed as a white noise model. Here we propose a pragmatic approach to using an estimated seismic wavelet to capture the correlated noise model (coloured noise) for the processed reflection seismic data. We test the method for a probabilistic sampling-based inversion where post-stack seismic data, associated with a hard carbonate reservoir in southwest Iran, is inverted directly to porosity. We assume eight different scenarios for the bandwidth and the magnitude of the noise. The investigation of the corresponding posterior statistics shows that ignoring the correlation of the noise samples in the noise covariance matrix generates unrealistic features in porosity realisations while underestimating the noise magnitude leads to overfitting the data and generating a biased model with low uncertainty. Furthermore, by considering an imperfect bandwidth for the noise model, the error is propagated to the posterior realisations. These issues are resolved considerably when the correlated noise is considered in the inversion. Therefore, in real data applications where the estimation of the magnitude and correlations of the noise is not trivial, the estimated seismic wavelet provides a good proxy for describing the correlation of the noise samples or equivalently the bandwidth of the noise model. In addition, it might be better to overestimate the noise magnitude than to underestimate it. This is true especially for an uncorrelated noise model and to a lesser degree also for the correlated noise model.

Key words: Probabilistic inverse problem, Noise model, Reservoir Geophysics, Uncertainty quantification

INTRODUCTION

Recorded seismic data are indispensably associated with uncertainty. From a theoretical perspective, the solution to the seismic inversion problem is non-unique, in that infinitely many models exist that allow fitting observed seismic data within the data uncertainty bounds. If the acquired data would be absolutely reproducible, i.e. noise-free, then all models whose responses did not match the data perfectly would be incorrect (Scales and Snider, 1998). The uncertainty in the estimated model can be attributed to diverse sources, including data uncertainties, data processing flow, and approximate or even faulty physical relationships between data and model parameters (Oliver *et al.* 2021). The noise in seismic data is often considered to be additive, such that the noise can be considered as the residual between the observed (recorded) and predicted (modelled) measurements, i.e. all that cannot be reconstructed by the forward modelling. The assumption of additive noise is common and well formulated in a standard framework in the inverse problem theory (Sen and Stoffa 1996; Mosegaard and Tarantola 2002; Trantola 2005; Ulrych and Sacchi 2005; Aster, Borchers, and Thurber 2018). By acknowledging that the seismic data is contaminated with the noise of some form or another, a realistic noise model must be considered in an inversion of seismic data in order to account for the data uncertainty precisely.

In a probabilistic inverse problem framework, the noise is assumed to consist of a combination of measurement and modelling errors (Tarantola 2005). Modelling errors stem from a variety of sources such as imperfect model parameter space parameterization, uncertainty related to the geometry of measurement, the use of approximate forward models, etc. In some cases, the magnitude of the modelling errors may be much smaller than the measurement errors, in which case they can be ignored. However, as will be demonstrated, for many practical cases, the modelling errors have a significant portion, and should not be disregarded.

It is common in seismic data inversion to assume that the data uncertainty is uncorrelated and stationary (Aleari, Ciabbari, and Gukov 2018; Gunning and Sams 2018; de Figueiredo *et al.* 2019; Grana 2020; Sengupta *et al.* 2021, Grana, de Figueiredo, and Mosegaard 2022; Li *et al.* 2022; Aleari 2021; Aleari and Salusti 2020). Cordua *et al.* (2009) assess the impact of correlated noise on cross-borehole ground-penetrating radar data. They argue that as the correlation of the data errors is accounted for, the higher resolution images are obtained. Hansen *et al.* (2014) argue that ignoring the modelling error can lead to severe artefacts, which erroneously appear to be well resolved in the solution of the inverse problem, which has been exemplified for seismic data (Madsen and Hansen 2018; Madsen, Nørmark, and Hansen 2018). x. Among few publications on this subject, Jakobsen and Hansen (2019) assess the effect of correlated and uncorrelated noise in a direct Bayesian inversion of 1D synthetic seismic data to lithofacies. They show that considering correlated noise improves the resolution of the model parameters compared to the case when using a white uncorrelated noise model. Madsen *et al.* (2017) also adopt a hierarchical Bayesian approach to infer the properties of the noise model as a part of inversion. Malinverno and Parker (2006) propose two approaches to invert geophysical measurements and estimate subsurface properties and their uncertainties when little is known a priori about the size of the errors associated with the data. They use an empirical Bayes approach, where hyperparameters (e.g., the variance of data uncertainty) are not well known in advance to the inversion and are estimated from the most probable value, given the observed data. Estimation of the variance of data uncertainty in each iteration of the sampling algorithm is practically demanding as the data covariance needs to be inverted at each iteration. However, Hansen *et al.* (2013) developed a toolbox through which it is straightforward to estimate the noise magnitude of even large-scale seismic data using a local subset of seismic data.

In this study, the uncorrelated noise assumption is compared with a range of scenarios for the correlated noise model with varying magnitude (variance) and shape (wavelet bandwidth). The performance of the probabilistic inversion under different noise assumptions is examined over a realistic synthetic case study where seismic data is inverted directly to a posterior distribution of porosity in a hard-rock carbonate reservoir. By analysing the posterior distribution of porosity and its statistics in different noise scenarios, a pragmatic approach is proposed to account for the noise model in the probabilistic petrophysical seismic inversion.

In the following, we first outline the inversion methodology and associated parameterisations. Then we discuss the results of probabilistic inversion of 2D realistic synthetic seismic data in presence of different noise models. As showcased in Heidari *et al.* (2022), the results of this work can be directly applied in industry-led studies, where the noise model is a challenging aspect in setting up the probabilistic inversion of real seismic data.

THEORY

Suppose \mathbf{m} represents some model parameters, which describes physical properties, and \mathbf{d} refers to the noise-free physical response of the model \mathbf{m} through the forward operator $\mathbf{d} = g(\mathbf{m})$. Assuming additive noise, the observed data \mathbf{d}_{obs} can be represented by the following equation,

$$\mathbf{d}_{\text{obs}} = g(\mathbf{m}) + \mathbf{e} = \mathbf{d} + \mathbf{e}, \quad (1)$$

where \mathbf{e} symbolizes the noise, due to both the measurement and theoretical (modelling) errors (Tarantola, 2005). The inverse problem then consists of inferring information about the model parameters \mathbf{m} , given information about the observed data, forward operator, and noise model.

PARAMETERISATION OF THE PROBABILISTIC INVERSE PROBLEM

Tarantola and Valette (1982) propose a probabilistic approach to solve the inverse problem, in which all information is quantified through a prior probability distribution $\rho(\mathbf{m})$ and a likelihood function $L(\mathbf{m})$. The prior information, quantified through $\rho(\mathbf{m})$, can come from various sources that are independent of the observed data, such as previous surveys, geoscientists' knowledge, outcrops, and so forth. In this study, we use the porosity logs at four boreholes as the prior information. The geophysical data information is quantified through the likelihood function $L(\mathbf{m})$, which describes the expected distribution of the data residual for the model \mathbf{m} as $L(\mathbf{m}) = f(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}))$. Once the available information is quantified, the combined state of information, prior distribution, and the likelihood is obtained using the concept of *conjunction of states of information* (Tarantola and Valette, 1982). This leads to the solution of the inverse problem as the posterior probability distribution,

$$\sigma(\mathbf{m}) = \kappa \rho(\mathbf{m}) L(\mathbf{m}), \quad (2)$$

where,

$$\kappa^{-1} = \int \rho(\mathbf{m}) L(\mathbf{m}) d\mathbf{m} \quad (3)$$

is a normalization constant, which ensures that $\int \sigma(\mathbf{m}) d\mathbf{m} = 1$.

Assuming a Gaussian distribution for the data uncertainty, the likelihood function is also Gaussian and is represented as (Mosegaard and Tarantola, 2002),

$$L(\mathbf{m}) = \left((2\pi)^N |C_D| \right)^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}))^T C_D^{-1} (\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m})) \right), \quad (4)$$

where N is the length of data and C_D is the data covariance, which can be split into

contributions from the measurement error C_d and modelling error C_t such that $C_D = C_d + C_t$, assuming independence of the two (Tarantola, 2005).

To solve the probabilistic inverse problem, it is necessary to describe the posterior distribution $\sigma(\mathbf{m})$ either analytically, or numerically. In this study, we use the extended Metropolis-Hastings algorithm (Mosegaard and Tarantola, 1995), through the “Stochastic Inverse Problem with informed Prior Information” (SIPPI) toolbox developed by Hansen *et al.* (2013) to sample the posterior distribution (See Appendix A for more details). To run the extended Metropolis-Hastings algorithm, we need to consider at least three main components: 1) sampling the prior model through a random walk, 2) solving the forward problem $\mathbf{d} = g(\mathbf{m})$, and 3) evaluating the likelihood $L(\mathbf{m})$. The third component is equivalent to evaluating the data residual as a realisation of the noise model. To represent the prior model numerically, we use the fast Fourier transform moving average (FFT-MA) method proposed by Le Ravalec *et al.* (2000), which is an efficient approach for generating independent realisations from a stationary multivariate Gaussian realisation. We also apply the inverse normal score transform (See, e.g., Goovaerts, 1997) to ensure that the distribution of the model realisations follows the target distribution, i.e., the distribution of the porosity logs (Hansen *et al.*, 2013). This leads to a more realistic distribution of the porosity realisations than a simple Gaussian distribution. In order to implement a random walk in the prior model space, we perturb the Gaussian random deviates used by the FFT-MA method at each iteration following Hansen et al (2012).

To solve the forward modelling problem and generate a synthetic trace associated with each model parameter, we use the widely used one-dimensional convolution approach (Yilmaz, 2001). We calibrate the Nur critical porosity rock physics model through the methodology proposed by Amini (2018) to obtain the optimized minerals’ elastic moduli and

pointwise critical porosity using the petrophysical logs, i.e., the volume of shale and water saturation, and density and sonic logs for four pseudo-logs. The petrophysical logs and porosity realisations are converted to elastic parameters using rock physics modelling. In equations (1) to (4), \mathbf{m} refers to porosity and the prior model refers to the assumption of the porosity distribution.

NOISE MODEL

In probabilistic seismic data inversion, it is a common practice to assume uncorrelated white noise, where the noise covariance matrix in equation (4) is described by the noise variance σ^2 and the identity matrix \mathbf{I} in the form of $\mathbf{C}_d = \sigma^2 \mathbf{I}$. Finding a representative covariance matrix is a key aspect of a correlated noise setup. It is acknowledged that the summation of the earth impulse response and the additive noise is convolved with the receiver array, geophone and recording system responses (Yilmaz 2018). Furthermore, during processing, all effort goes to attenuation of the noise components that do not overlap with the seismic signal. Consequently, the remnants of noise in seismic data after processing must be correlated (Figure 1). Therefore, we suggest modelling the correlated noise using an estimated source wavelet that is a good representative of both the seismic signal and the noise. To address the correlation of the noise samples and simulate realistic correlated noise, we generate a set of realisations from a Gaussian random reflectivity $\mathcal{N}(0,0.0324)$ and convolve them with a seismic wavelet extracted from a 3D real seismic mid-angle gather. Figure 2a shows some 1D band-limited Gaussian noise realisations. The corresponding covariance matrix of 531 correlated noise can be seen in Figure 2b. The mean of the nonzero diagonals of the covariance matrix, the seismic wavelet as well as their corresponding frequency spectra are shown in Figures 3a and b, respectively. The mean of the nonzero diagonals of the covariance matrix could be a reasonable approximation of the seismic wavelet, as their main lobes are

similar and their frequency spectrums are in approximately the same range. Therefore, we suggest that the seismic wavelet can potentially be assumed as a suitable proxy for capturing the coupling of data samples for both signal and noise within the seismic bandwidth. Consequently, in contrast to the routine practice where the main diagonal of the noise covariance matrix is a delta function, it is sensible to insert the seismic wavelet along the main diagonal of the covariance matrix $C_{t,shape}$ to account for the correlation of the noise samples more realistically.

THE SYNTHETIC SEISMIC DATA SET

A 2D geological model of porosity from a hard-rock carbonate reservoir in southwest Iran is used to build the synthetic seismic data (Figure 4a). Porosity has the dominant petrophysical control on the seismic response (Mavko *et al.* 2020). The rock-physics model proposed by Heidari *et al.* (2020) in conjunction with an estimated wavelet extracted from the real seismic data is used to generate the synthetic seismic data (Figure 4b). To simulate a realistic correlated (bandlimited) Gaussian noise, we convolve the wavelet with a 2D Gaussian random reflectivity field with zero mean and standard deviation (STD) of 0.18, providing a signal-to-noise (SNR) of 3.3 (Figure 4c). This 2D noise section was added to the noise-free seismic data to generate the noisy seismic section (Figure 4d).

NOISE SCENARIOS IN THE INVERSION ALGORITHM

In order to test the hypothesis that the correlated part of the noise model can be determined based on the seismic wavelet, we consider eight different scenarios with various noise magnitudes and bandwidths. The first, second, and third scenarios assume uncorrelated noise models $C_{\mathbf{d}} = C_{\mathbf{d}}$ with different levels of noise variance σ^2 . The other five scenarios assume correlated noise models $C_{\mathbf{d}} = C_{\mathbf{t}}$ (Table 1). Similar to the uncorrelated noise $C_{\mathbf{d}}$, the covariance of the correlated noise model is scaled such that $C_{\mathbf{t}} = \sigma^2 C_{t,shape}$. In the fourth

scenario, to capture the optimal coupling of data samples for both signal and noise within the seismic bandwidth, we use a C_t equivalent to the noise added to the data. To address the uncertainty with the wavelet estimation and its effect on the statistics of the posterior realisations, scenarios five and six make use of shorter and longer wavelets than the reference wavelet to construct $C_{t,shape}$ with the correct level of noise variance σ^2 given (Figures 5a and b). Estimation of the magnitude of the noise in real data is not straightforward. Scenarios seven and eight aim at examining the performance of the inversion algorithm where the magnitude of the noise variance σ^2 is over- and under-estimated respectively but the correct $C_{t,shape}$ is provided.

RESULTS AND DISCUSSIONS

Figure 6 represents the results of the probabilistic inversion as the mean of the porosity posterior realisations for all of the noise model scenarios. To better investigate the inversion performance, we show the log-likelihood curves of the sampling procedure in Figure 7. Figure 6a represents the result of the inversion for case 1, where the inversion resolves the true porosities even for the thin layers. However, there are some unrealistic features (stripes) on the porosity section. These artefacts can be attributed to ignoring the correlation of the noise samples. Figures 6b and 6c show the results of over- and underestimation of the noise magnitude in the uncorrelated noise setup (cases 2 and 3). Overestimation of the noise variance generates a smooth posterior mean model with low resolution. It should be noted that the individual realisations may not be necessarily smooth. The results also clearly indicate that ignoring the correlation in the noise setup as well as underestimation of the magnitude generate apparently well-resolved features with lower uncertainty, exemplified by white arrows in Figure 6c, which are simply biases from overfitting the data. We discuss the uncertainty assessment further in Figure 9. In addition, assuming uncorrelated noise leads to a harder sampling problem. In other words, the sampling algorithm is not able to search the

high probability regions and sample them sufficiently. This is due to the failure of the algorithm to reach the “burn-in” phase, where the correct posterior distribution is sampled, and instead, the data is overfitted through an optimization procedure rather than a sampling procedure. The behaviour of the log-likelihood curve of this case in Figure 7c also confirms the failure of the algorithm to converge to a steady-state and search the full posterior distribution. In this case, the true porosity is found in a 95% CI only for 17.52% of data samples. Also, the root-mean-square error (RMSE) of the residual of the true and estimated porosity is 13.20. This indicates that although the resolution of the posterior realisations is high, the algorithm fails to sufficiently capture the posterior distribution and generate independent realisations. Here, there is a trade-off between the number of iterations and the resolution. The inability of the sampling algorithm to fully sample the posterior and generate enough independent realisations may be resolved if the number of iterations is increased. However, it is not feasible to run the algorithm for numerous iterations, especially in real data applications. Therefore, considering just the posterior realisations to conclude the sampling algorithm performance in presence of different noise models is not enough. In other words, not only the behaviour of the log-likelihood curve but the statistical features of the posterior realisations should be analysed as well.

Figure 6d shows the results of case 4 for the correlated noise assumption using the seismic wavelet and true STD. We expect this case to be the best-case scenario as the assumed noise is the same as the added noise and the distribution of the posterior realisations follows the distribution of the realistic reference model due to applying the normal score transformation. Therefore, in this case, the algorithm should obtain the optimal inversion resolution for the problem. Furthermore, the true porosity is found in a 95% CI for 92.23% of data points and the RMSE of the estimated porosity is 10.70. This shows that the sampled realisations describe the true porosity reasonably well. The results of cases 5 and 6 are shown

in Figures 6e and 6f. These figures qualitatively show that the error propagated into the posterior realisations by over- and underestimating the wavelet bandwidth in noise covariance construction is insignificant as 92% and 93.23% of the true porosity is found in a 95% CI and the RMSE is 10.73 and 10.49, respectively. It indicates that in practice and for real data applications, we would not be penalized a lot by choosing an imperfect wavelet for constructing the covariance matrix of a correlated noise model. Also, the associated log-likelihood curve shown in Figure 7 levels out, which shows that the burn-in phase is finished normally. The results of cases 7 and 8 are depicted in Figures 6g and 6h, respectively. The results are consistent with cases 2 and 3 (Figures 6b and 6c). Therefore, it is sensible to be conservative and overestimate rather than underestimate the noise magnitude. This is consistent with findings related to correlated noise in travel-time inversion where only little resolution is lost assuming a too high magnitude of the correlated noise (Hansen et al. 2014). A comparison of the results of these cases with the results of cases 2 and 3 highlights the key contribution of including the correlation of the noise samples (the non-diagonal elements) in the covariance matrix. In addition to fewer stripes on the mean porosity section, a larger number of data points are found within 95% CI, in the correlated noise setup. However, comparing figures 6b and 6g (i.e., black arrows), denotes that although ignoring the correlation of the noise samples generates some unrealistic features in case 2, this case leads to a better estimation of the porosity magnitude for the mean model. The first three columns in Table 2 summarise the percentage of the true porosity found in a 95% CI and the RMSE of the estimated porosity for all the noise scenarios.

The log-likelihood curves in Figure 7 might be misleading for the assessment of the sampling performance as they seem to be similar for different cases of the correlated and uncorrelated noise scenarios. Therefore, to quantify the efficiency of the different Markov chains in the sampling procedure, we obtain the number of iterations between independent

realisations and the number of independent realisations obtained from the posterior distribution for all the cases using the correlation between the sampled realisations from the posterior distribution (See Table 2). The numbers in the last two columns of Table 2 show that the noise models in cases 3 and 8 lead to the hardest sampling as the algorithm fails to sample sufficiently and the number of independent realisations is notably lower in comparison with the other noise models. In addition, in Table 2, the difference between the statistics of the porosity realisations and the RMSE of the estimated porosity for cases 3 and 8 is more notable compared to the other cases. For better illustration of the statistical analysis criteria mentioned in Table 2, we normalised their associated values in a range between 0 and 1 and depicted them as a spider plot shown in Figure 8. Consistent with Table 2, the highest value of the number of iterations between independent realisations and consequently the lowest value of the number of independent realisations corresponds to Case 3.

Figure 9 represents a binary map, where all the regions where the reference (true) porosity is located outside the 95% confidence interval (CI) are shown in white colour. In an ideal inversion, 5% of this map should show white colours scattered evenly. Thus, this figure visually gives a good insight into the regions where the inversion could not resolve the true porosity properly. For instance, in regions where the white points are predominant, the posterior uncertainty is underestimated. This showcases features in the mean model that appear well-resolved but are not, such as the mean model in Figure 6c (case 3), which is clearly not resolved. In these binary maps, some recognizable lateral patterns (which are highlighted by red curves in Figure 9a) are visible for all cases except case 3. Judging from figure 6, this is probably attributed to some geological boundaries, where the current prior model is incapable of fully describing the spatial variability of the reference model. This may, in part, also be attributed to some edge effect in the inversion algorithm (Proakis, 2001). Cases 4 to 7 with the correlated noise models are superior to the rest of the inversion results

in alleviating the effect of the imperfect prior model, resolving the true porosity, and hence providing trustworthy uncertainty estimates. This is confirmed when describing the overall performance as a single number as done in Table 2, where the CI of cases 4 to 7 are all close to the target value of 95%.

In these maps, some features such as the ones highlighted by the red ellipses are striking in cases where the noise model is uncorrelated. These regions are more prone to incorrect uncertainty estimates in cases where the noise bandwidth is not considered and the sampling algorithm is unable to resolve the porosity correctly. This notable difference indicates the importance of assuming correlation in the noise model, such that the risk of overfitting the data and consequently generating artefacts is minimized.

CONCLUSIONS

Due to the data acquisition and data processing flow, the remnant noise in the seismic data is more likely to be correlated noise. It has been demonstrated that ignoring such correlated noise and considering only uncorrelated noise will lead to artefacts in the posterior distribution as un-accounted correlated noise will be propagated into the posterior distribution. The statistical analysis of the posterior realisations for all of the noise model scenarios indicates the indispensable role of the shape of the noise model such that, ignoring the correlation of the noise samples in the covariance matrix leads to the generation of unrealistic features on the posterior realisations and issues related to the sampling convergence. As a practical solution, we suggest using the scaled seismic wavelet based on the estimated noise magnitude to build the noise covariance matrix. The results indicate that the inversion is robust to variations in the shape (wavelet bandwidth) of the correlated noise model. In other words, even if an approximate wavelet is used, the results suggest only a minor effect on the posterior distribution. However, the effect of noise magnitude is more pronounced, such that assuming too small magnitude for the correlated noise leads to fitting noise as resolved features and also leads to a hard sampling problem. On the other hand, considering a higher magnitude than the true magnitude leads to a minor effect on the resolution of the posterior distribution. Thus, in order to provide a robust and practically useful noise model when inverting reflection seismic data, a conservative choice for the magnitude of the noise is to choose a magnitude at least as high as the true magnitude. Adopting the frequency spectrum of the wavelet as a proxy for the correlated noise model, while remaining conservative about noise levels is hence the most pragmatic takeaway of the inversion tests as it can potentially guide the practitioners to account for the precise noise model in a probabilistic inversion of real seismic data.

REFERENCES

- Aleardi, M., 2021. A gradient-based Markov chain Monte Carlo algorithm for elastic pre-stack inversion with data and model space reduction. *Geophysical Prospecting*, 69(5), pp.926-948.
- Aleardi, M., Ciabbarri, F. and Gukov, T., 2018. A two-step inversion approach for seismic-reservoir characterization and a comparison with a single-loop Markov-chain Monte Carlo algorithm. *Geophysics*, 83(3), pp.R227-R244.
- Aleardi, M. and Salusti, A., 2020. Markov chain Monte Carlo algorithms for target-oriented and interval-oriented amplitude versus angle inversions with non-parametric priors and non-linear forward modellings. *Geophysical Prospecting*, 68(3), pp.735-760.
- Amini, H., 2018, June. Calibration of minerals' and dry rock elastic moduli in sand-shale mixtures. *In 80th EAGE Conference and Exhibition 2018 (Vol. 2018, No. 1, pp. 1-5)*. European Association of Geoscientists & Engineers.
- Aster, R.C., Borchers, B. and Thurber, C.H., 2018. *Parameter estimation and inverse problems*. Elsevier.
- Cordua, K.S., Nielsen, L., Looms, M.C., Hansen, T.M. and Binley, A., 2009. Quantifying the influence of static-like errors in least-squares-based inversion and sequential simulation of cross-borehole ground penetrating radar data. *Journal of Applied Geophysics*, 68(1), pp.71-84.
- de Figueiredo, L.P., Grana, D., Roisenberg, M. and Rodrigues, B.B., 2019. Gaussian mixture Markov chain Monte Carlo method for linear seismic inversion. *Geophysics*, 84(3), pp.R463-R476.

- Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. Oxford University Press on Demand.
- Grana, D., 2020. Bayesian petroelastic inversion with multiple prior models. *Geophysics*, 85(5), pp.M57-M71.
- Grana, D., de Figueiredo, L. and Mosegaard, K., 2022. Markov chain Monte Carlo for petrophysical inversion. *Geophysics*, 87(1), pp.M13-M24.
- Li, K., Yin, X., Zong, Z. and Grana, D., 2022. Estimation of porosity, fluid bulk modulus, and stiff-pore volume fraction using a multitrace Bayesian amplitude-variation-with-offset petrophysics inversion in multiporosity reservoirs. *Geophysics*, 87(1), pp.M25-M41.
- Geman, S. and Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6), pp.721-741.
- Gunning, J. and Sams, M., 2018. Joint facies and rock properties Bayesian amplitude-versus-offset inversion using Markov random fields. *Geophysical Prospecting*, 66(5), pp.904-919.
- Hansen, T.M., Cordua, K.S., Jacobsen, B.H. and Mosegaard, K., 2014. Accounting for imperfect forward modeling in geophysical inverse problems—exemplified for crosshole tomography. *Geophysics*, 79(3), pp.H1-H21.
- Hansen, T.M., Cordua, K.S., Looms, M.C. and Mosegaard, K., 2013. SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 2—Application to crosshole GPR tomography. *Computers & Geosciences*, 52, pp.481-492.

- Hansen, T.M., Cordua, K.S. and Mosegaard, K., 2012. Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling. *Computational Geosciences*, 16(3), pp.593-611.
- Heidari, H., Hansen, T.M., Amini, H., Niri, M.E., Madsen, R.B. and Amini, N., 2022. Single-step probabilistic inversion of 3D seismic data of a carbonate reservoir in Southwest Iran. *Geophysics*, 87(2), pp.B159-B177.
- Heidari, A., Amini, N., Amini, H., Niri, M.E., Zunino, A. and Hansen, T.M., 2020. Calibration of two rock-frame models using deterministic and probabilistic approaches: Application to a carbonate reservoir in south-west Iran. *Journal of Petroleum Science and Engineering*, 192, p.107266.
- Jakobsen, A.F. and Hansen, H.J., 2019, June. Exploring Noise Models in Approximate Bayesian Inversion for Facies. In 81st EAGE Conference and Exhibition 2019 (Vol. 2019, No. 1, pp. 1-5). European Association of Geoscientists & Engineers.
- Li, K., Yin, X., Zong, Z. and Grana, D., 2022. Estimation of porosity, fluid bulk modulus, and stiff-pore volume fraction using a multitrace Bayesian amplitude-variation-with-offset petrophysics inversion in multiporosity reservoirs. *Geophysics*, 87(1), pp.M25-M41.
- Madsen, R.B. and Hansen, T.M., 2018. Estimation and accounting for the modeling error in probabilistic linearized amplitude variation with offset inversion. *Geophysics*, 83(2), pp.N15-N30.
- Madsen, R.B., Nørmark, E. and Hansen, T.M., 2018, June. Accounting for Processing Errors in AVO/AVA Data. In 80th EAGE Conference and Exhibition 2018 (Vol. 2018, No. 1, pp. 1-5). European Association of Geoscientists & Engineers.

- Madsen, R.B., Zunino, A. and Hansen, T.M., 2017. On inferring the noise in probabilistic seismic AVO inversion using hierarchical Bayes. In SEG Technical Program Expanded Abstracts 2017 (pp. 601-606). Society of Exploration Geophysicists.
- Malinverno, A. and Parker, R.L., 2006. Two ways to quantify uncertainty in geophysical inverse problems. *Geophysics*, 71(3), pp.W15-W27.
- Mavko, G., Mukerji, T. and Dvorkin, J., 2020. *The rock physics handbook*. Cambridge university press.
- Mosegaard, K. and Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research: Solid Earth*, 100(B7), pp.12431-12447.
- Mosegaard, K. and Tarantola, A., 2002. Probabilistic approach to inverse problems. *International Geophysics Series*, 81(A), pp.237-268.
- Oliver, D.S., 2022. Information content in 4D seismic data: Effect of correlated noise. *Journal of Petroleum Science and Engineering*, 208, p.109728.
- Oliver, D.S., Fossum, K., Bhakta, T., Sandø, I., Nævdal, G. and Lorentzen, R.J., 2021. 4D seismic history matching. *Journal of Petroleum Science and Engineering*, 207, pp.109119.
- Proakis, J.G., 2001. *Digital signal processing: principles algorithms and applications*. Pearson Education India.
- Ravalec, M.L., Noetinger, B. and Hu, L.Y., 2000. The FFT moving average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations. *Mathematical Geology*, 32(6), pp.701-723.
- Scales, J.A. and Snieder, R., 1998. What is noise?. *Geophysics*, 63(4), pp.1122-1124.

- Sen, M.K. and Stoffa, P.L., 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion. *Geophysical Prospecting*, 44(2), pp.313-350.
- Sengupta, M., Zhang, H., Zhao, Y., Jervis, M. and Grana, D., 2021. Direct depth-domain Bayesian amplitude-variation-with-offset inversion. *Geophysics*, 86(5), pp.M167-M176.
- Tarantola, A., 2005. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics.
- Tarantola, A. and Valette, B., 1982. Inverse problems= quest for information. *Journal of geophysics*, 50(1), pp.159-170.
- Ulrych, T.J. and Sacchi, M.D., 2005. *Information-based inversion and processing with applications*. Elsevier.
- Yilmaz, Ö., 2001. *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists.

APPENDIX A: THE EXTENDED METROPOLIS-HASTINGS ALGORITHM

The extended Metropolis algorithm (Mosegaard and Tarantola, 1995) can be used to sample the product of two probability distributions, such as the product between a likelihood and a prior distribution. It requires that one can sample the prior using a random walk, and evaluate the likelihood for any proposed model. If we have a way of evaluating the values of the likelihood $L(\mathbf{m})$ and an algorithm A that can perform a random walk in the prior model $\rho(\mathbf{m})$ directly (without necessarily evaluating $\rho(\mathbf{m})$ anywhere), the following algorithm will sample the posterior $\sigma(\mathbf{m})$:

- Starting in the current model \mathbf{m}_c , perform one step in the random walk with the prior sampler A .
- Accept the new point \mathbf{m}_t only with probability $\mathbf{P}_{accept} = \min(1, L(\mathbf{m}_t) / L(\mathbf{m}_c))$.
- If \mathbf{m}_t is rejected here, re-use \mathbf{m}_c in the next step.
- If \mathbf{m}_t is accepted, let $\mathbf{m}_t = \mathbf{m}_c$ in the next step.

The above algorithm thus, through \mathbf{P}_{accept} , allows for the acceptance of less likely models than the currently accepted and the full likelihood to be explored. In this study, we use the sequential Gibbs sampler (algorithm A) to perform the random walk in the prior distribution. Considering a known realisation \mathbf{m} of the random field described by the probability distribution $\rho(m_1, \dots, m_N)$, if we randomly select a model parameter $m_i = m(\mathbf{u}_i)$ and compute the local conditional pdf $\mathbf{f}(m_i | m_1, m_2, \dots, m_{i-1}, m_{i+1}, \dots, m_N)$, and draw a value from it, we get a new realisation of the random field defined by $\rho(m_1, \dots, m_N)$. If this is repeated iteratively, it will be an application of the Gibbs sampler (Geman and Geman, 1984). By combining the Gibbs sampling and sequential simulation (Hansen et al., 2012) it is possible to sample realisations of the probability distribution $\rho(\mathbf{m})$ instead of computing the full conditional pdf analytically.

TABLES

Table 1 The noise scenarios considered in probabilistic seismic inversion

| Noise scenarios | Description |
|-----------------|--|
| Case 1 | Uncorrelated (STD of 0.18) |
| Case 2 | Uncorrelated – overestimated noise (STD of 0.36) |
| Case 3 | Uncorrelated – underestimated noise (STD of 0.02) |
| Case 4 | Correlated – bandwidth similar to seismic bandwidth (STD of 0.18) |
| Case 5 | Correlated – bandwidth more than the seismic bandwidth (STD of 0.18) |
| Case 6 | Correlated – bandwidth less than the seismic bandwidth (STD of 0.18) |
| Case 7 | Correlated – bandwidth similar to seismic bandwidth (STD of 0.36) |
| Case 8 | Correlated – bandwidth similar to seismic bandwidth (STD of 0.02) |

Table 2 The statistical analysis of the posterior distribution. The first three columns: the percentage of the true porosity inside a 95% CI and the RMSE of the estimated porosity. The last two columns: the number of iterations between independent realisations and the number of independent realisations estimated through posterior static analysis for all the eight noise scenarios in probabilistic seismic inversion

| Noise scenarios | The true porosity inside a 95% CI (%) | The RMSE | The number of iterations between independent realisations | The number of independent realisations |
|------------------------|--|-----------------|--|---|
| Case 1 | 81.18 | 11.13 | 4000 | 100 |
| Case 2 | 90.11 | 11.03 | 3000 | 133 |
| Case 3 | 17.52 | 13.20 | 82000 | 3 |
| Case 4 | 92.23 | 10.70 | 1000 | 400 |
| Case 5 | 92.00 | 10.73 | 1000 | 400 |
| Case 6 | 93.23 | 10.49 | 1000 | 400 |
| Case 7 | 94.45 | 11.13 | 1818 | 220 |
| Case 8 | 80.00 | 10.10 | 8000 | 50 |

FIGURES

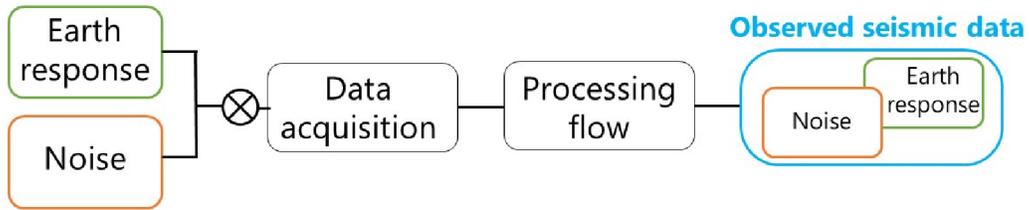


Figure 1 A diagram describing the effect of data acquisition and the processing flow on the earth impulse response and additive noise. The final observed seismic data contains the noise in the bandwidth of the seismic wavelet, which cannot be discriminated from the seismic signal.

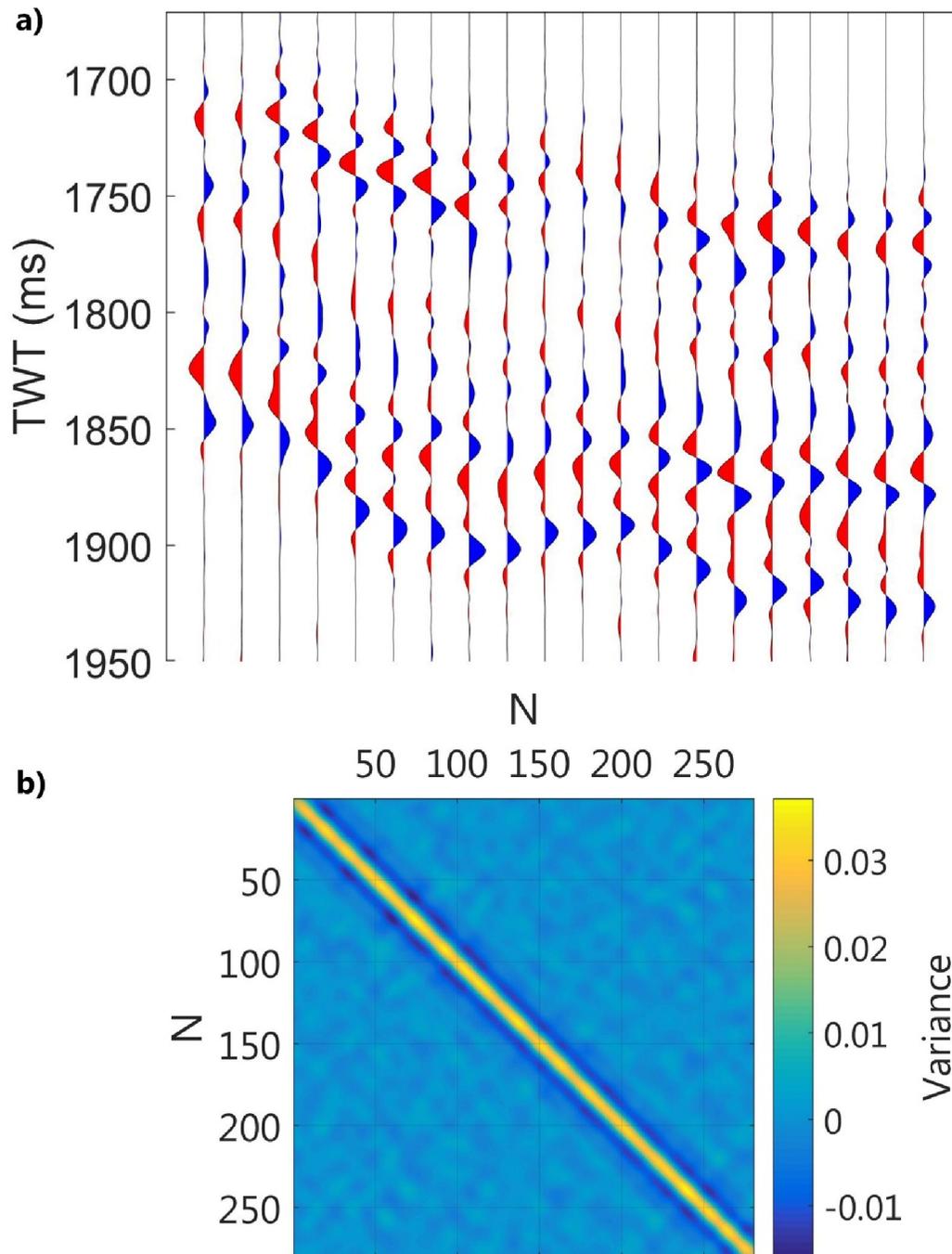


Figure 2 (a) Some correlated (bandlimited) noise traces constructed through convolution of the extracted seismic wavelet with 1D Gaussian random reflectivity series with mean zero and variance of 0.0324 (Gaussian noise), (b) the covariance matrix of a 2D correlated noise section with 531 traces, some of which are exemplified in (a).

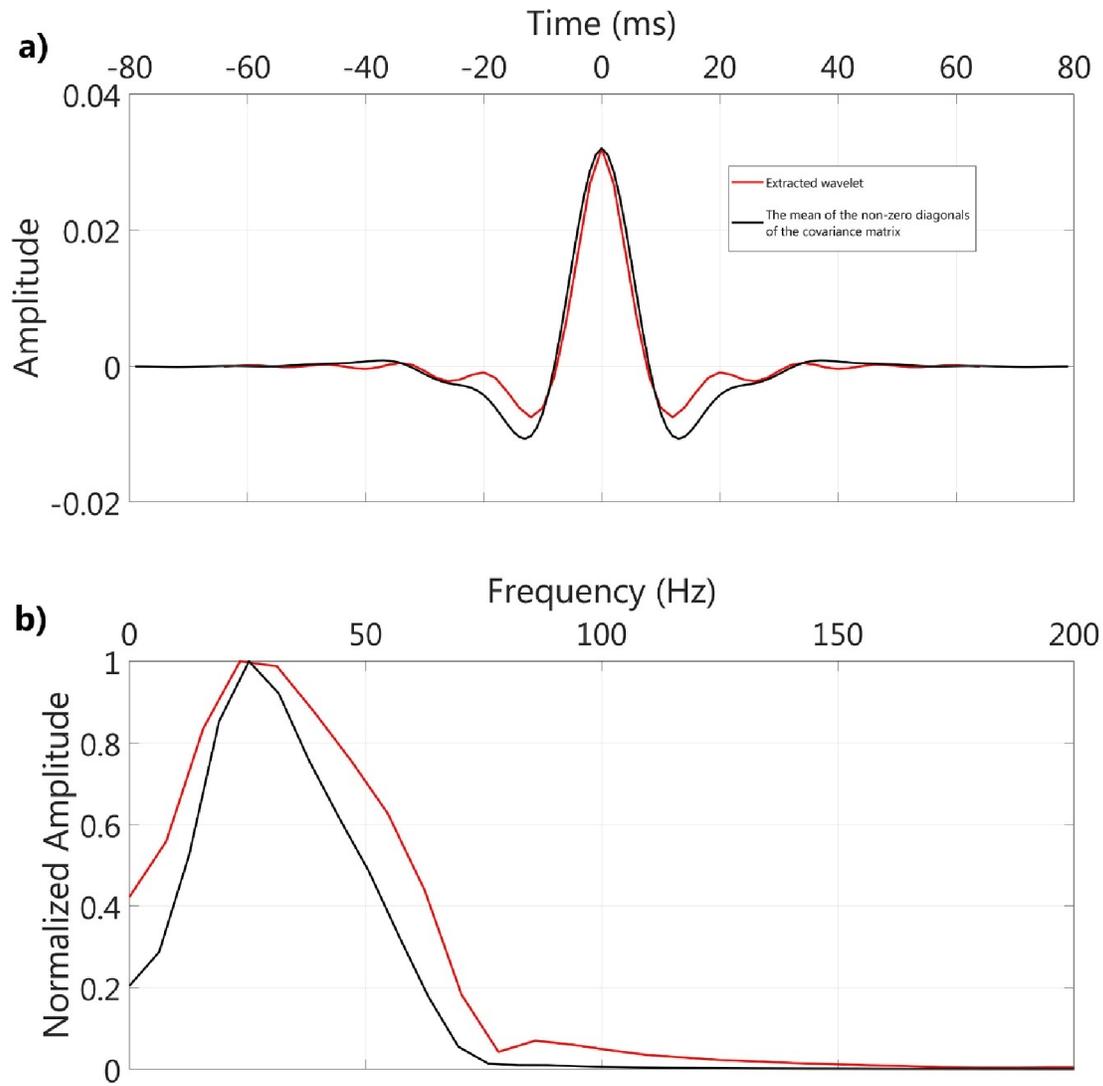


Figure 3 (a) The statistical seismic wavelet (red) with the mean of the nonzero diagonals of the covariance matrix (black) in Figure 2b, and (b) the frequency spectrum of the wavelets in (a).

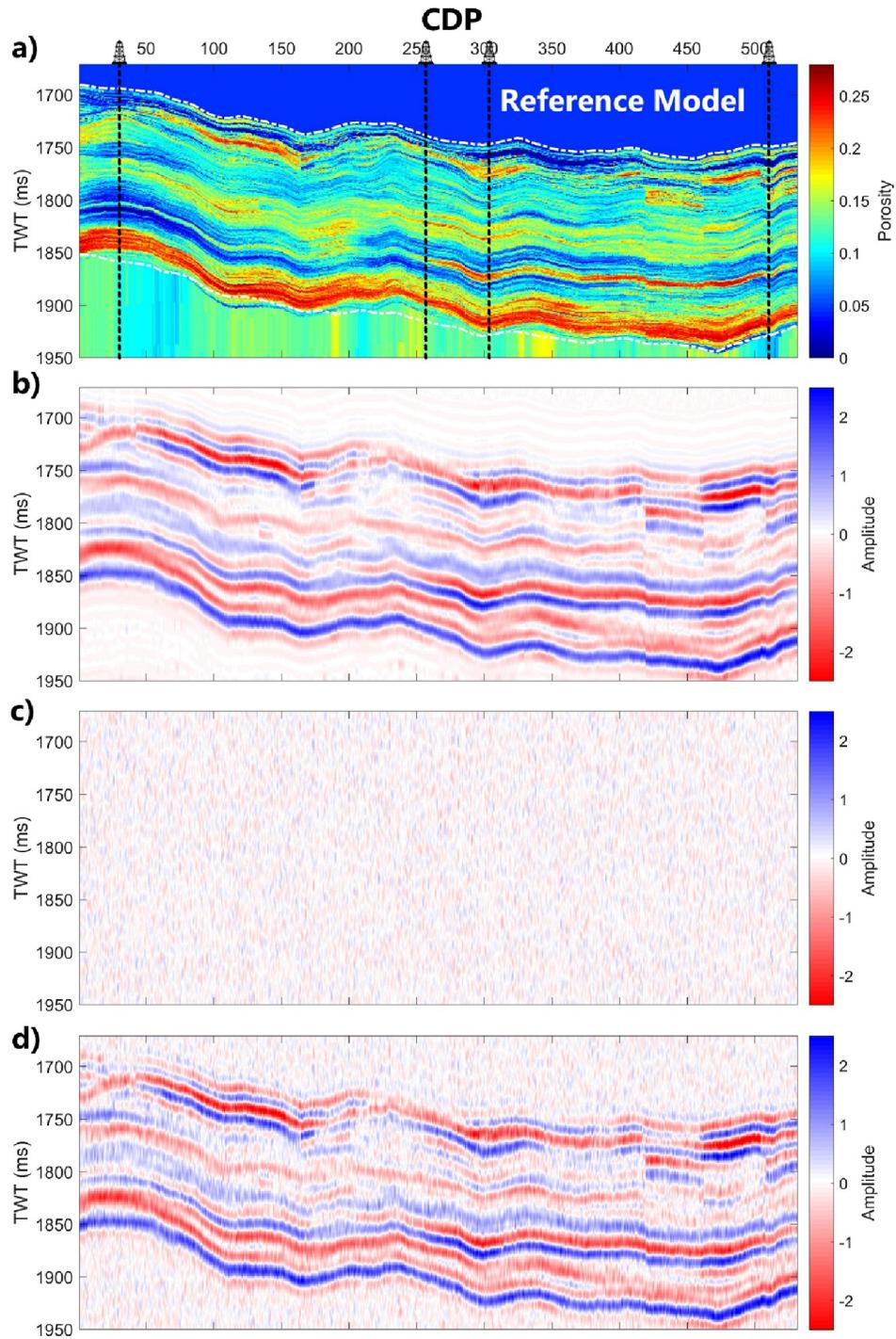


Figure 4 (a) The reference porosity model superimposed by the pseudo-log location (black dashed lines), (b) the noise-free seismic data, (c) 2D band-limited Gaussian random field constructed through convolution of the statistical seismic wavelet with the 2D Gaussian random reflectivity series with mean zero and variance of 0.0324 (Gaussian noise), (d) the noisy seismic section.

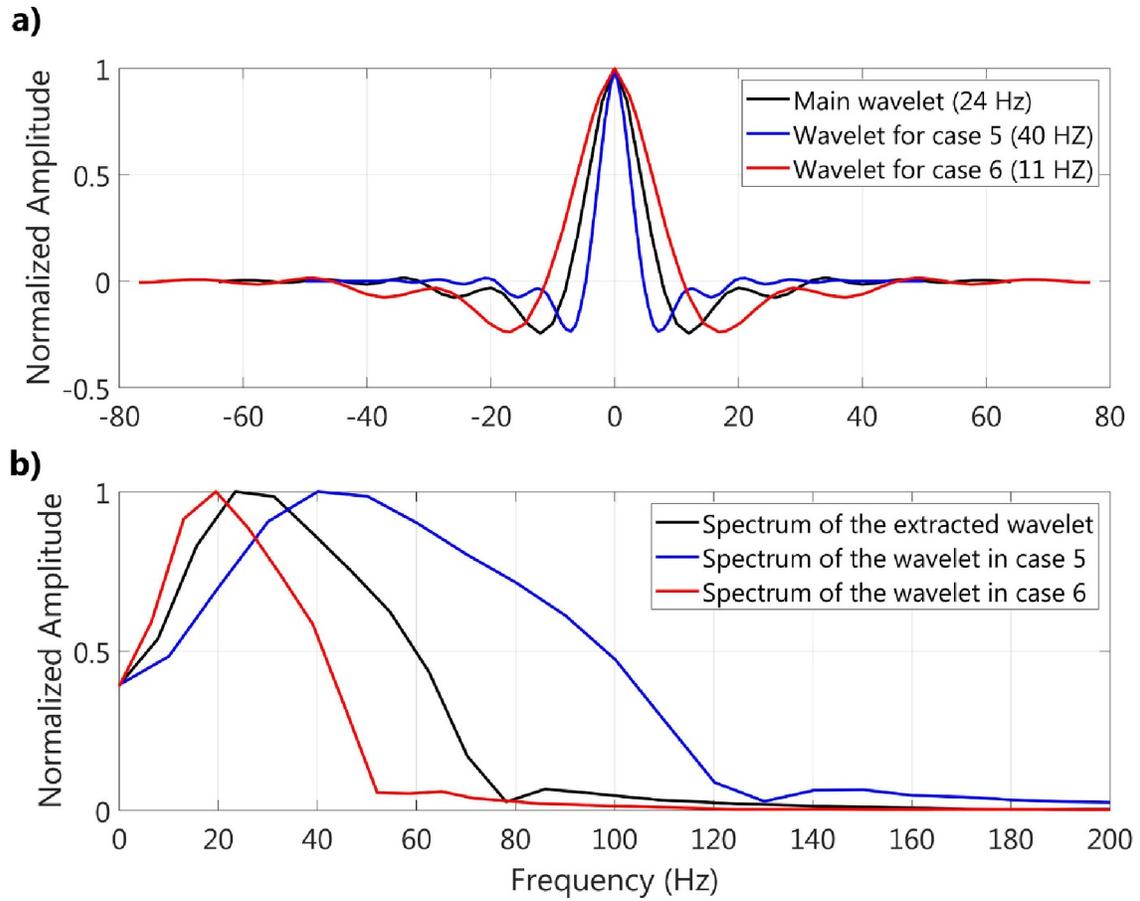


Figure 5 (a) seismic wavelets and (b) their frequency spectrum that are used for noise covariance matrix construction, black for cases 4, 7, and 8, blue for case 5, and red for case 6.

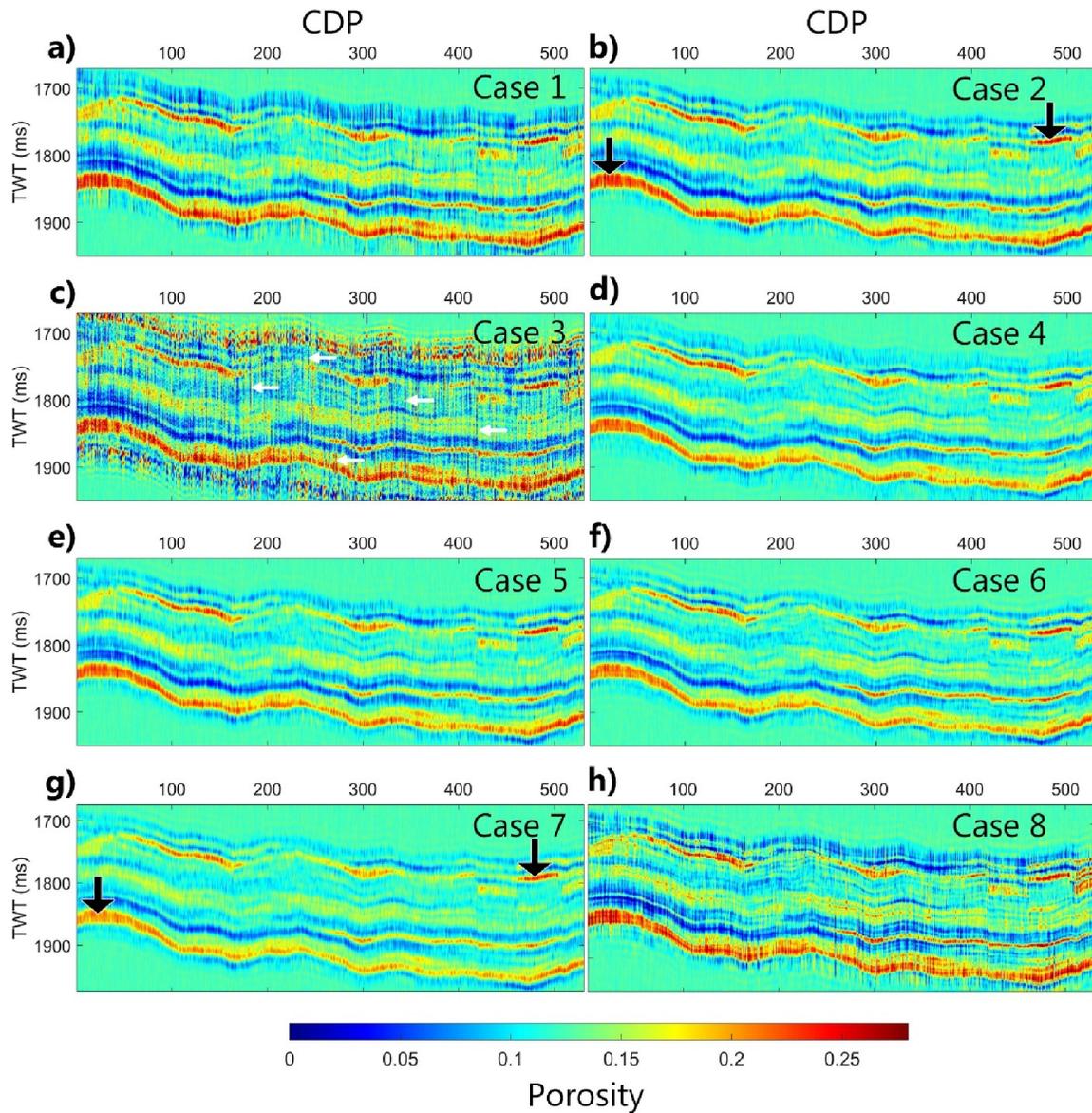


Figure 6 The mean of the porosity posterior realisations for eight noise scenarios in probabilistic seismic inversion, which are described in Table 1 (See the text for detailed description).

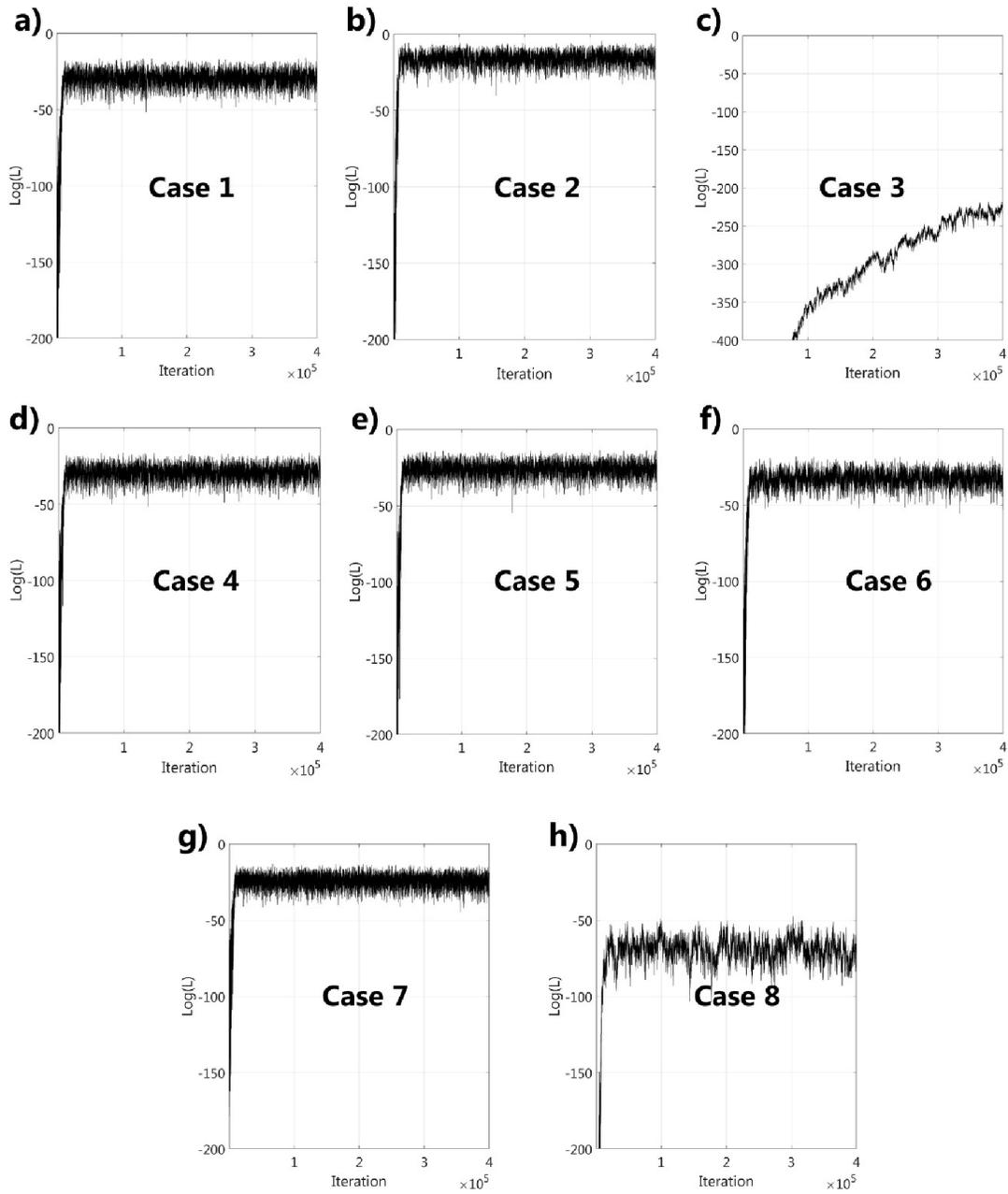


Figure 7 The log-likelihood curves for different noise scenarios in seismic probabilistic inversion. The failure of the sampling algorithm to reach the burn-in phase is notable in case 3.

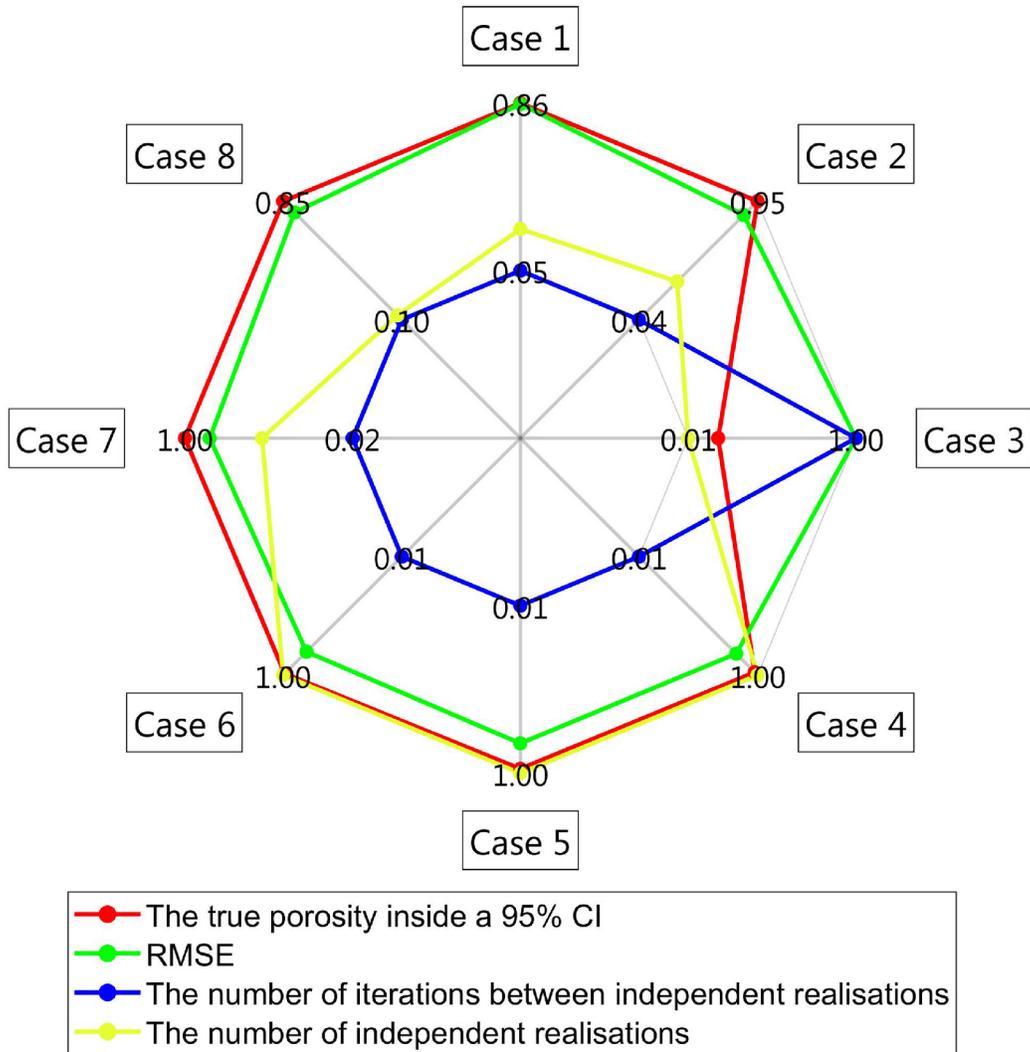


Figure 8 The spider plot of the statistical analysis criteria and their associated values, which are shown in Table 2. To facilitate the comparison of the values of each criterion between different noise scenarios, the values of all four criteria are normalised in a range between 0 and 1.

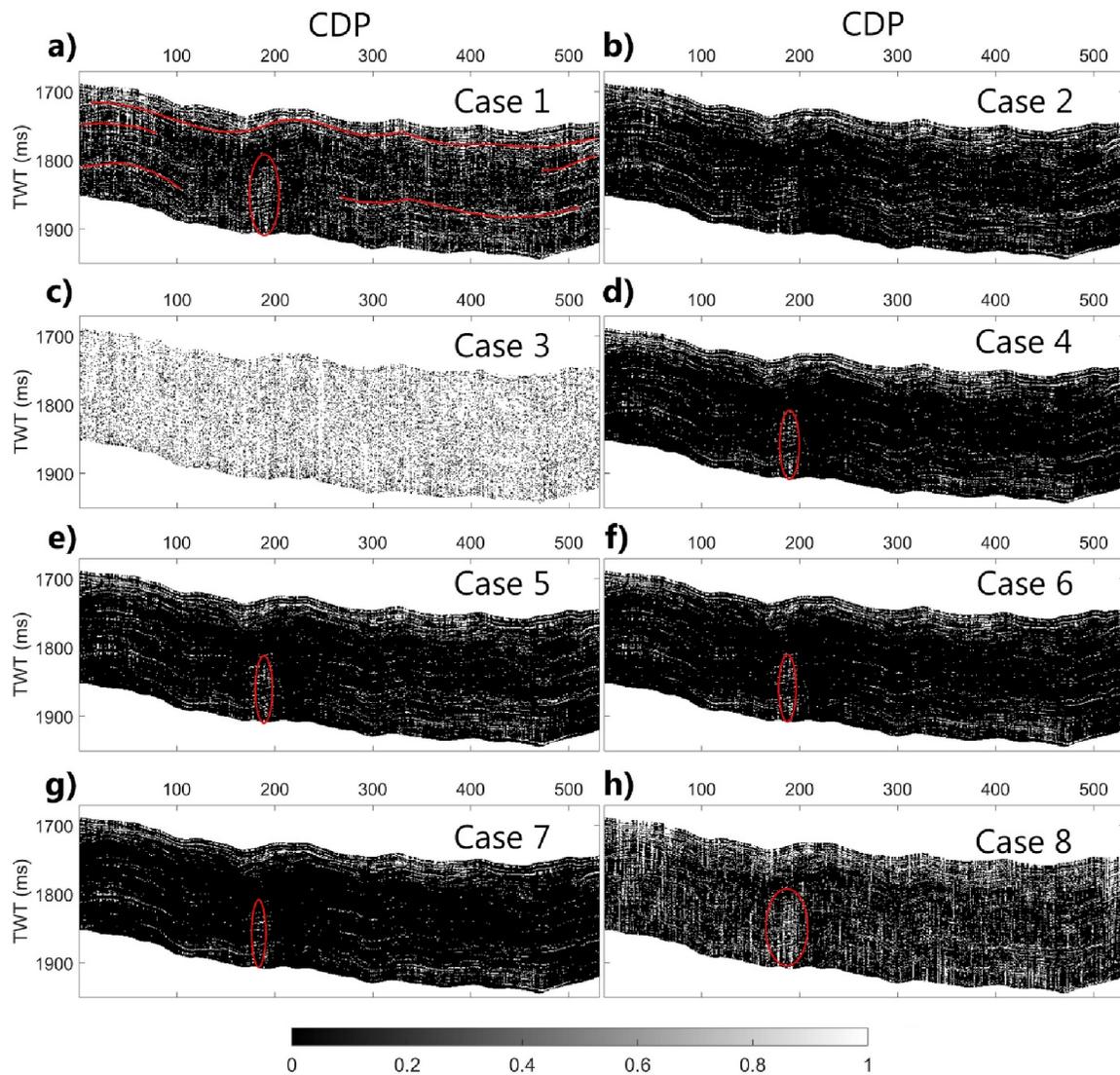


Figure 9 The binary map describing the regions where the true porosity is outside (white points) and inside a 95% CI (black points). Red curves highlight spatial patterns of regions where the true porosity is outside a 95% CI. The red ellipses exemplify a feature describing the ability of the sampling algorithm with different noise models in resolving the true porosity.