On inferring the noise in probabilistic seismic AVO inversion using hierarchical Bayes

Madsen, Rasmus Bødker; Zunino, Andrea; Hansen, Thomas Mejer

Published in:
SEG International Exposition and Annual Meeting, 24-29 September, Houston, Texas

Publication date:
2017

Document version
Publisher's PDF, also known as Version of record

On inferring the noise in probabilistic seismic AVO inversion using hierarchical Bayes

Rasmus Bødker Madsen, Andrea Zunino and Thomas Mejer Hansen*
Niels Bohr Institute, University of Copenhagen

SUMMARY

A realistic noise model is essential for trustworthy inversion of geophysical data. Sometimes, as in case of seismic data, quantification of the noise model is non-trivial. To remedy this, a hierarchical Bayes approach can be adopted in which properties of the noise model, such as the amplitude of an assumed uncorrelated Gaussian noise model, can be inferred as part of the inversion. Here we demonstrate how such an approach can lead to substantial overfitting of noise when inverting a 1D reflection seismic NMO data set. We then argue that usually the noise model is correlated, and suggest to infer the amplitude of a correlated Gaussian noise model. This provides better results than assuming an uncorrelated model. In general though, the results suggest that care should be taken using the hierarchical Bayes approach to infer the noise model.

INTRODUCTION

Data ($\mathbf{d} \in \mathbb{R}^N$) are the unique forward response from a physical model (g) given some model parameters ($\mathbf{m} \in \mathbb{R}^M$):

$$\mathbf{d} = g(\mathbf{m}) \tag{1}$$

In nature, observed data ($\mathbf{d}_{\text{obs}}$) are not noise-free, i.e. $\mathbf{d}_{\text{obs}} = g(\mathbf{m}) + \varepsilon$. The inverse problem of inferring values of the model parameters from the observed data are therefore non-unique (Sen and Stoffa, 1996; Tarantola, 2005).

The chosen likelihood function, which measures how well the model parameters match the observed data, is dependent on the distribution of noise (Box and Tiao, 1992). The likelihood, and hence the final posterior distribution of model parameters, can be very sensitive to noise-level and noise models (Thore, 2015). The noise model should ideally include information about measurement and experimental errors as well as account for imperfections in the forward model and/or simplifications due to the parametrization (Dosso and Holland, 2006). However, it is often argued that since no theory is exact, all features that are not captured by the theory are just observational errors, and no distinction should be made between different types of noise (Sen and Stoffa, 1996).

In the classical Bayesian stochastic inference paradigm information on the uncertainty distribution of the data should be established a priori as part of the likelihood function. If for instance a Gaussian noise model is assumed, the corresponding Gaussian likelihood takes the following form (Box and Tiao, 1992)

$$p(\mathbf{d}_{\text{obs}}|\mathbf{m}) = \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - g(\mathbf{m}))^{\mathrm{T}}\mathbf{C}_{\mathrm{D}}(\mathbf{d}_{\text{obs}} - g(\mathbf{m}))\right) \tag{2}$$

where $\mathbf{C}_{\mathrm{D}}$ is the data covariance model. The data covariance can be split into contributions from measurement error $\mathbf{C}_{\mathrm{d}}$ and theory error $\mathbf{C}_{\mathrm{T}}$, assuming independence of the two. (Mosegaard and Tarantola, 2002)

$$\mathbf{C}_{\mathrm{D}} = \mathbf{C}_{\mathrm{d}} + \mathbf{C}_{\mathrm{T}}\,. \tag{3}$$

An extension to the classical paradigm is offered through the use of the hierarchical Bayesian approach, where parameters of the noise (and of the prior) can be inferred from the observed data (Gelman et al., 2014). A hierarchical scheme has been used in various geophysical inverse problems in order to infer information about noise level from data (Buland and Omre, 2003b; Malinverno and Briggs, 2004; Malinverno and Parker, 2006; Bodin et al., 2012; Dettmer and Dosso, 2012; Ray et al., 2013).

Buland and Omre (2003b) explicitly used a hierarchical Bayes formulation of the inverse problem, enabling noise-level estimation as well as wavelet estimation in a joint AVO (Amplitude Versus Offset) inversion scheme. They concluded that estimating the seismic noise model as part of the probabilistic inversion is viable. However, in their real data case only negligible improvements were gained on posterior variance of model parameters, compared to estimating the noise covariance and wavelet prior to AVO inversion. As opposed to Buland and Omre (2003b), we will analyze the posterior resolution (not the posterior variance) obtained using hierarchical Bayes with different assumptions about the noise model. Specifically we will investigate whether the noise is underestimated, which will result in an apparent smaller posterior variability. However, such reduced posterior variability might reflect non-existent features appearing as a result of fitting noise as data. We propose a set of synthetic tests similar to that of Buland and Omre (2003a). The synthetic tests are based on a reference model from which possible posterior biases in the model parameters can be assessed alongside the posterior variance.

THEORY AND METHOD

Forward model and prior information

The synthetic data (AVO gather) are created following the methodology of Buland and Omre (2003a). The problem here is to infer information about the three elastic parameters: P-wave (v_p), S-wave (v_s), and density (ρ) from a seismic AVO gather. The observed data can be expressed as a linear convolution forward problem with some added noise:

$$\mathbf{d}_{\text{obs}} = \mathbf{G}\mathbf{m} + \mathbf{e} = \mathbf{W}\mathbf{A}\mathbf{D}\mathbf{m} + \mathbf{e} \tag{4}$$

where $\mathbf{W}$ is the wavelet matrix which is convolved with the reflectivity series $\mathbf{AD}\mathbf{m}$. The prior distribution of model parameters $p(\mathbf{m})$ is Gaussian and the three elastic parameters have an internal correlation of 0.7. This prior corresponds to the "Well B" scenario of (Buland and Omre, 2003a). In order to allow

a least-squares solution, the prior distribution of the model parameters is assumed to be log-Gaussian, so:

$$\mathbf{m} = [\ln(v_p/v_p^0)^{\mathrm{T}}, \ln(v_s/v_s^0)^{\mathrm{T}}, \ln(\rho/\rho^0)^{\mathrm{T}}]^{\mathrm{T}} \sim \mathrm{Gauss}(\boldsymbol{\mu}, \mathbf{C}_{\mathrm{M}}) \tag{5}$$

where $\boldsymbol{\mu}$ is the prior expectation of the model parameters, v_p^0, v_s^0, ρ^0 are reference values, and $\mathbf{C}_{\mathrm{M}}$ describes the prior covariance. The noise on the observed data is $\mathbf{e} = \mathbf{e}_{\mathrm{uncor}} + \mathbf{e}_{\mathrm{cor}}$, where the noise model is Gaussian and split into two components as in equation 3. Here we let $\mathbf{C}_{\mathrm{d}}$ be the uncorrelated white noise component and $\mathbf{C}_{\mathrm{T}}$ be the correlated colored noise component, so that:

$$\mathbf{e}_{\mathrm{uncor}} \sim \mathrm{Gauss}(0, \mathbf{C}_{\mathrm{d}}),\ \mathbf{e}_{\mathrm{cor}} \sim \mathrm{Gauss}(0, \mathbf{C}_{\mathrm{T}}). \tag{6}$$

A tricky aspect of noise in seismic AVO data is the fact that the error is often systematic (Riedel et al., 2003). In order to test the hierarchical Bayesian approach of inferring noise, the colored noise is set to:

$$\mathbf{C}_{\mathrm{T}} = \sigma_{\mathrm{T}}^2 \mathbf{C}_{\mathrm{T,shape}} = \sigma_{\mathrm{T}}^2 \frac{\mathbf{WADC}_{\mathrm{M}}(\mathbf{WAD})^{\mathrm{T}}}{\max[\mathbf{WADC}_{\mathrm{M}}(\mathbf{WAD})^{\mathrm{T}}]}. \tag{7}$$

This covariance matrix gives the covariance of the "prior data distribution". Noise realizations from this distribution would tend to imitate data. By normalizing with the maximum value, the variance of the noise can be set according to σ_{T}^2. Colored noise with signal-to-noise ratio (SNR) = 1.25 is added to the synthetic data, which is in the poor end of what can be expected from seismic data. In practice this is achieved by having the standard deviation σ_{T} as the standard deviation of the reference model's forward response (signal) divided by 1.25. The covariance of the uncorrelated noise is simply the identity matrix $\mathbf{I}$ times the variance:

$$\mathbf{C}_{\mathrm{d}} = \sigma_{\mathrm{d}}^2 \mathbf{C}_{\mathrm{d,shape}} = \sigma_{\mathrm{d}}^2 \mathbf{I}. \tag{8}$$

A SNR = 30 is chosen for the uncorrelated noise. This number is perhaps slightly generous towards the filtering processes. On the other hand, white uncorrelated noise is different in wavelength from the seismic signal and can often easily be filtered out during processing of the raw data (Vecken and Da Silva, 2004). Most of the remaining noise would therefore tend to resemble the observed data in the frequency domain.

Bayesian linearized AVO inversion

Since the inverse problem is linear Gaussian, the Gaussian likelihood in equation 2 can be used with the linear operator $\mathbf{G}$. The posterior distribution $p(\mathbf{d}_{\mathrm{obs}}|\mathbf{m}) \sim \mathrm{Gauss}(\hat{\boldsymbol{\mu}}, \hat{\mathbf{C}}_{\mathrm{M}})$ is then also a multivariate Gaussian distribution (Tarantola and Valette, 1982), where the expectation and covariance are given by:

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu} + (\mathbf{WADC}_{\mathrm{M}})^{\mathrm{T}} \mathbf{C}_{\mathrm{D}}^{-1} (\mathbf{d}_{\mathrm{obs}} - \mathbf{WAD}\boldsymbol{\mu}) \tag{9}$$

$$\hat{\mathbf{C}}_{\mathrm{M}} = \mathbf{C}_{\mathrm{M}} - (\mathbf{WADC}_{\mathrm{M}})^{\mathrm{T}} \mathbf{C}_{\mathrm{D}}^{-1} \mathbf{WADC}_{\mathrm{M}} \tag{10}$$

This analytical solution of the Bayesian linear inverse problem, which is also referred to as the least-squares solution, describes the full posterior distribution of model parameters with uncertainty under a Gaussian assumption.

Hierarchical Bayes

In a hierarchical model, the conditional parameters (e.g. model parameters or noise model) for the observed data are themselves given a probabilistic specification. Consequently, the conditional parameters are then also dependent on another set of parameters (Gelman et al., 2014). These additional parameters are typically known as hyperparameters $\mathbf{h} = [h_1, h_2, \ldots]$. Uncertainty now includes both model parameters and hyperparameters. Therefore a prior distribution $p(\mathbf{h})$ should be set for the hyperparameters (hyperprior) that reflects the initial uncertainty on these. The posterior distribution of hyperparameters (hyperposterior) is then determined by inversion of the observed data. For linear inverse Gaussian problems, as the one outlined above, Malinverno and Briggs (2004) propose a computationally efficient approach of hieracrchical Bayes, that we adopt here. The marginal likelihood of the observed data conditional on the hyperparameters, in case a linear Gaussian solution to the problem exists, is given by

$$p(\mathbf{d}_{\mathrm{obs}}|\mathbf{h}) = \left[\frac{\det \hat{\mathbf{C}}_{\mathrm{M}}}{\det \mathbf{C}_{\mathrm{M}}}\right]^{\frac{1}{2}} \exp\left[-\frac{1}{2}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{C}_{\mathrm{M}}^{-1} (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})\right] \frac{1}{[(2\pi)^N \det \mathbf{C}_{\mathrm{D}}]^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{d}_{\mathrm{obs}} - \mathbf{G}\hat{\boldsymbol{\mu}})^{\mathrm{T}} \mathbf{C}_{\mathrm{D}}^{-1} (\mathbf{d}_{\mathrm{obs}} - \mathbf{G}\hat{\boldsymbol{\mu}})\right]. \tag{11}$$

Using the sampling strategy suggested by Malinverno and Briggs (2004), the posterior probability distribution of the hyperparameters are essentially sampled using a Metropolis-Hastings algorithm (Mosegaard and Tarantola, 1995). At each step a random walk goes through an exploration phase where a candidate value of the hyperparameters is proposed in vicinity of the current. Thereafter, an exploitation phase either accepts or rejects the candidate based on the marginal likelihood in equation 11. The posterior for the model parameters is sampled with a Gibbs sampler (Sen and Stoffa, 1996). The model parameters are at each step conditioned on the current accepted set of hyperparameters. For our proposed sampling strategy we let the step-length of the random walk in the hyperparameters be dynamic for the first 1000 iterations, with a target acceptance rate of 30%. This is a commonly used practice in Metropolis-Hastings algorithms in order to secure a more efficient burn-in period and subsequent sampling algorithm (Gelman et al., 1996).

SYNTHETIC TESTS

In the following, we present three case studies: the first assuming uncorrelated data noise, the second assuming a known shape of the noise, and the third assuming only an approximate knowledge of the shape of the noise. Regarding the simulated data, the standard deviations of the realization of the error are $\sigma_{\mathrm{d}} = 0.0039$ and $\sigma_{\mathrm{T}} = 0.0949$, respectively, so the noise is mainly correlated. The standard deviation on the final combined noise realization is $\sigma_{\mathrm{d+T}} = 0.0947$.

Hierarchical Bayes inversion - Case 1

In the first inversion example we assume that all noise on the data is uncorrelated. The only hyperparameter that is needed in the hierarchical model is then $h_1 = \sigma_{\mathrm{d}}$ in equation 8. We want

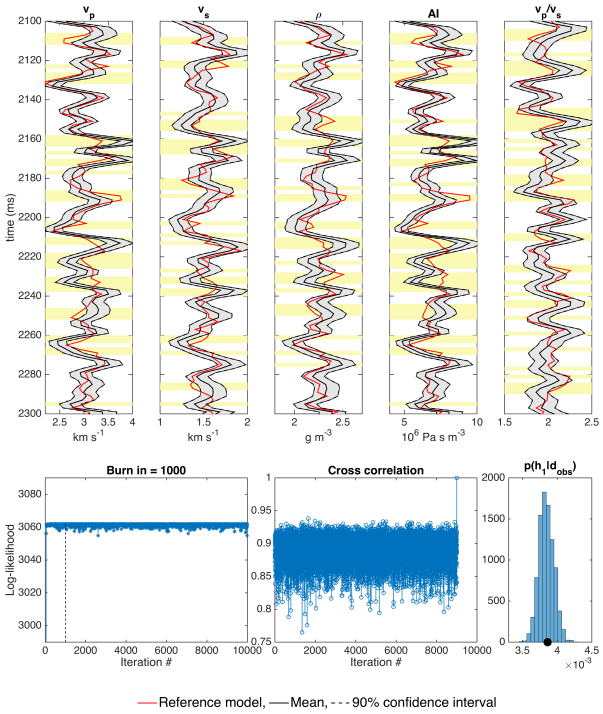

Figure 1: Posterior distribution of $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ and $p(\mathbf{h}|\mathbf{d}_{\text{obs}})$, log-likelihood and cross-correlation for case 1 of hierarchical Bayes inversion using $\mathbf{C}_{\text{D}} = h_1^2\mathbf{C}_{\text{d}}$

to replicate a setting where little prior knowledge of the noise level exists, but the noise is assumed to be uncorrelated (e.g. Buland and Omre (2003b); Malinverno and Briggs (2004); Bodin et al. (2012)). In these circumstances it seems appropriate to have a relative non-informative hyperprior distribution. We therefore set a rather wide uniform hyperprior distribution for the standard deviation of the uncorrelated noise: $p(h_1) \sim \text{Unif}(0.0001, 1)$. Realizations and statistics of the resulting posterior distribution of model parameters $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ and hyperparameters $p(\mathbf{h}|\mathbf{d}_{\text{obs}})$ are shown in Figure 1, as well as analytically calculated values for acoustic impedance (AI) and the ratio between P-wave and S-wave velocity (v_p/v_s). Areas where the reference model (red line) are not found within the 90% confidence interval is marked with yellow. Ideally, 90% of the profile should be non-yellow for the posterior resolution to be trustworthy for this interval. However, this is not the case as the reference model is found inside the confidence interval considerably less than 90%. This is especially true for v_p/v_s where the reference model is only found within the 90% confidence interval 46% of the time. This indicates a severe case of overfitting the data. The hyperparameter h_1, instead, is actually well determined when comparing the true standard deviation (black dot) to the histogram in Figure 1. However, the red dot only indicates the level of uncorrelated noise on the data σ_{d}. The total standard deviation of the error realization $\sigma_{\text{d+T}}$ (blue dot) is heavily underestimated as seen in Figure 2. Since the noise on the data is apparently too similar in its structure to the actual signal, the colored noise is then considered as part of the signal by the algorithm. This explains both the

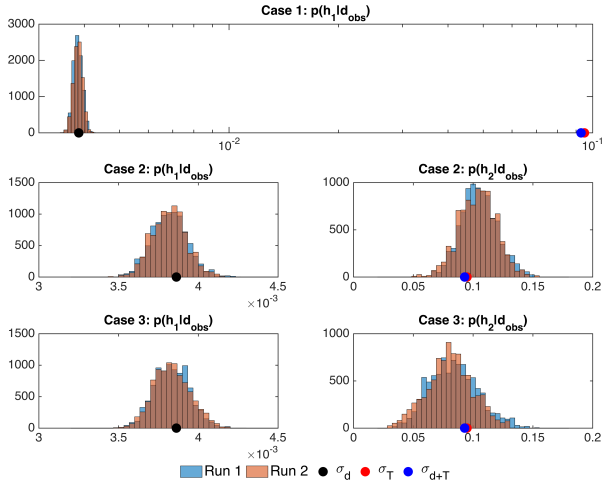

Figure 2: Histograms of the hyperposterior distributions for all MCMC runs $p(\mathbf{h}|\mathbf{d}_{\text{obs}})$. Notice the logarithmic scale used for Case 1 (top plot).

overfitting of the posterior for the model parameters and the algorithms ability to correctly determine the variance of uncorrelated noise on the data while underestimating the actual total noise level.

The correlation between adjacent samples is not high, as the correlation decreases to the average level quickly form the last sample. The pattern of log-likelihood and cross-correlation is similar for all the following MCMC runs in general and are therefore not shown.

The overfitting of the hierarchical Bayes inversion with the uncorrelated noise model becomes even more apparent in Figure 3. Here the percentage of the reference model being inside the confidence interval is plotted as a function of the size of the confidence interval. It is clearly visible that the posterior distribution of the elastic variables is not capturing the reference model for all confidence intervals. The apparently small posterior variability is actually reflecting non-existent features, i.e. noise being fitted as data. The result of the hierarchical inversion is similar to just applying a linear AVO inversion with $\sigma_{\text{d}} = 0.0039$. To evaluate the consistency of the results from case 1, the algorithm is run an additional time. The histogram of the hyperposterior of h_1 in Figure 2 for the secondary run shows the same pattern as for the first run. This indicates a certain level of consistency in the MCMC results.

Hierarchical Bayes inversion - Case 2

In the second test case we assume a known shape of colored noise $\mathbf{C}_{\text{T,shape}}$, i.e. we test whether it is possible at all to infer the variance of the noise knowing the reference color. The standard deviation of the colored noise is set as an additional hyperparameter $\sigma_{\text{T}} = h_2$. The noise model then takes the following form: $\mathbf{C}_{\text{D}} = h_1^2\mathbf{C}_{\text{d,shape}} + h_2^2\mathbf{C}_{\text{T,shape}}$. Again, we set a rather broad uniform hyperprior distribution for the standard deviation of the correlated noise: $p(h_2) \sim \text{Unif}(0.0001, 1)$. The result for two runs are summarized in Figure 2 and 3. Both MCMC runs show approximately the same hyperposterior dis-

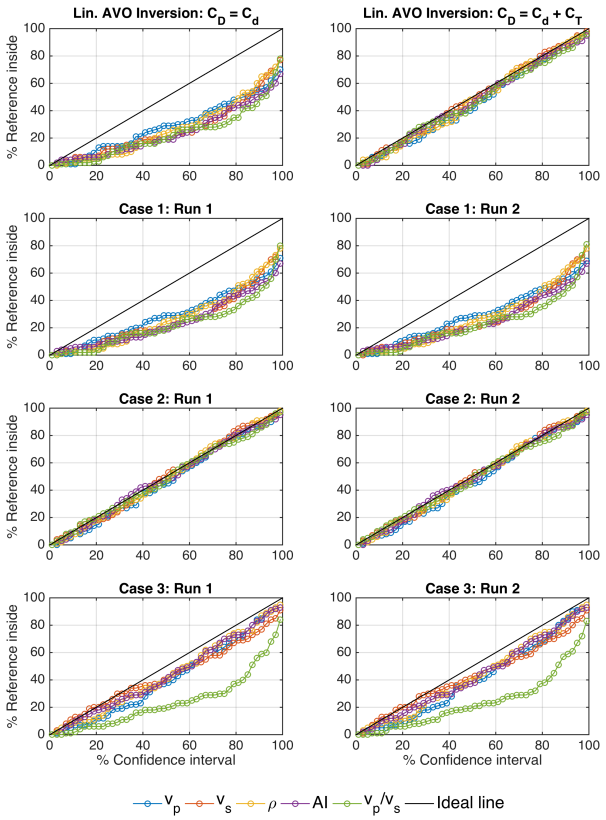

Figure 3: Percentage of reference model being inside the confidence interval, i.e. non-yellow areas in Figure 1), as a function of confidence interval. The uppermost figures are calculated using Bayesian linearized AVO inversion for reference.

tributions $p(\mathbf{h}|\mathbf{d}_{\text{obs}})$. The variance of the correlated noise h_2 (red dot) is slightly overestimated, whereas the variance of the uncorrelated noise h_1 is well-determined. The nearly-correct estimation of variance level, results in the reference model being within the confidence interval following the ideal line in Figure 3. The model therefore appears to neither over- nor underfit the data. The results are similar to doing a linear AVO inversion with the reference noise model.

Hierarchical Bayes inversion - Case 3

Since the correct shape of the noise is never readily available in a real-world scenario, we propose to estimate or assume some correlation of the noise prior to inversion. We assume that the noise is showing smoothness comparable with the wavelet. Furthermore, the noise is correlated between the individual angle-stacks. The correlation between angle stacks is believed to vary slightly as a function of angle. This is set up in the following manner

$$\mathbf{W}_{\text{corr}} = \begin{bmatrix} \beta_1\mathbf{I} & \frac{\beta_1+\beta_2}{2}\mathbf{I} & \dots & \frac{\beta_1+\beta_n}{2}\mathbf{I} \\ \frac{\beta_2+\beta_1}{2}\mathbf{I} & \beta_2\mathbf{I} & \dots & \frac{\beta_2+\beta_n}{2}\mathbf{I} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_n+\beta_1}{2}\mathbf{I} & \frac{\beta_n+\beta_2}{2}\mathbf{I} & \dots & \beta_n\mathbf{I} \end{bmatrix}. \quad (12)$$

where $\mathbf{I}$ is the identity matrix with the size of one individual angle stack. A simple model can easily be derived by setting $\boldsymbol{\beta} = \beta_1, \beta_2, \dots, \beta_n = 1$. Using the variance of each angle stack on the observed data offers an estimate for $\boldsymbol{\beta}$. These beta values could possibly also be obtained using optimization of the marginal likelihood in equation 11. In our case we use $\boldsymbol{\beta} = [1, 0.99, 0.98, 0.95, 0.92, 0.90, 0.88]$. The final approximate shape of the colored noise is then:

$$\mathbf{C}_{\text{W,shape}} = \frac{\mathbf{W}\mathbf{W}_{\text{corr}}\mathbf{W}^{\text{T}}}{\max[\mathbf{W}\mathbf{W}_{\text{corr}}\mathbf{W}^{\text{T}}]} \quad (13)$$

Using the same approach as for case 2, the approximate noise model takes the following form: $\mathbf{C}_{\text{D}} = h_1^2\mathbf{C}_{\text{d,shape}} + h_2^2\mathbf{C}_{\text{W,shape}}$. The variance of the uncorrelated noise is as for case 1 and 2 correctly estimated as both histograms are centered around the correct value (black dot) in Figure 2. However, the variance of the uncorrelated noise is underestimated when comparing the hyperposterior distribution of $p(h_2|\mathbf{d}_{\text{obs}})$ with the correct value (blue dot). As for both case 1 and case 2, there is consistency between the results from the two MCMC runs. Figure 3 shows improvements of case 3 compared to case 1 for all elastic parameters but the v_p/v_s ratio, which is still not captured by the posterior distribution. This indicates that overfitting of the data is still present using the approximate shape, but is nevertheless reduced compared to simply assuming an uncorrelated shape of the noise.

CONCLUSION

Our results in general indicate that caution should be taken when inferring the noise as an additional parameter in inversion. It seems that the assumption of uncorrelated noise in case 1 is not good for inferring the correct noise level on data with correlated noise. The hierarchical Bayes approach was in all cases able to accurately estimate the variance of the uncorrelated noise on the data. However, using the approach of case 1 the total variance of the noise is not recovered and significant overfitting of the data was demonstrated. Choosing the correct shape of the noise, as in case 2, eliminates the overfitting, and a correct variance is estimated even for a wide hyperprior. The results from case 2 indicate that it is possible to estimate a correct variance of the noise model using the data. Finally, for case 3 with an approximate shape of the noise, the variance estimate of the noise is improved compared to case 1. The model is however still overfitting the data. Especially v_p/v_s shows an apparent smaller posterior variability that is not capturing the true model. In a real world case it is probably reasonable to assume that substantial knowledge about the general noise-level is available to constrain the wide hyperprior distribution (Gelman et al., 2014). This could potentially improve the result for an approximate noise model. For our synthetic tests, using an approximate shape of the correlated noise offers an improvement on the noise-level estimate compared with using an uncorrelated noise model. Further work could nonetheless be done to obtain more reliable estimates for an approximate shape of the noise, which could further improve the posterior resolution of the hierarchical Bayes approach in general.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2017 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Bodin, T., M. Sambridge, N. Rawlinson, and P. Arroucau, 2012, Transdimensional tomography with unknown data noise: Geophysical Journal International, **189**, 1536–1556, http://doi.org/10.1111/j.1365-246X.2012.05414.x.

Box, G. E., and G. C. Tiao, 1992, Bayesian inference in statistical analysis: A Wiley-Interscience Publication, 608.

Buland, A., and H. Omre, 2003a, Bayesian linearized AVO inversion: Geophysics, **68**, 185–198, http://doi.org/10.1190/1.1543206.

Buland, A., and H. Omre, 2003b, Joint AVO inversion, wavelet estimation and noise-level estimation using a spatially coupled hierarchical Bayesian model: Geophysical Prospecting, **51**, 531–550, http://doi.org/10.1046/j.1365-2478.2003.00390.x.

Dettmer, J., and S. E. Dosso, 2012, Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains: Journal of the Acoustical Society of America, **132**, 2239–2250, http://doi.org/10.1121/1.4746016.

Dosso, S. E., and C. W. Holland, 2006, Geoacoustic uncertainties from viscoelastic inversion of seabed reflection data: IEEE Journal of Oceanic Engineering, **31**, 657–671, http://doi.org/10.1109/JOE.2005.858358.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2014, Bayesian data analysis, 3rd ed: Chapman and Hall/CRC.

Gelman, A., G. O. Roberts, and W. R. Gilks, 1996, Efficient metropolis jumping rules: Bayesian Statistics, **5**, 42.

Malinverno, A., and V. A. Briggs, 2004, Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes: Geophysics, **69**, 1005–1016, http://doi.org/10.1190/1.1778243.

Malinverno, A., and R. L. Parker, 2006, Two ways to quantify uncertainty in geophysical inverse problems: Geophysics, **71**, no. 3, W15–W27, http://doi.org/10.1190/1.2194516.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431–12447, http://doi.org/10.1029/94JB03097.

Mosegaard, K., and A. Tarantola, 2002, Probabilistic approach to inverse problems, International handbook of earthquake and engineering seismology, part A: Academic Press, 237–265.

Ray, A., K. Key, and T. Bodin, 2013, Hierarchical Bayesian inversion of marine CSEM data over the Scarborough gas field A lesson in correlated noise: 83rd Annual International Meeting, SEG, Expanded Abstracts, 723–727, http://doi.org/10.1190/segam2013-1439.1.

Riedel, M., S. E. Dosso, and L. Beran, 2003, Uncertainty estimation for amplitude variation with offset (AVO) inversion: Geophysics, **68**, 1485–1496, http://doi.org/10.1190/1.1620621.

Sen, M. K., and P. L. Stoffa, 1996, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion: Geophysical Prospecting, **44**, 313–350, http://doi.org/10.1111/j.1365-2478.1996.tb00152.x.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation, 1st ed.: SIAM.

Tarantola, A., and B. Valette, 1982, Inverse problems = quest for information: Journal of Geophysics, **50**, 159–170.

Thore, P., 2015, Uncertainty in seismic inversion: What really matters?: The Leading Edge, **34**, 1000–1004, http://doi.org/10.1190/tle34091000.1.
Vecken, P. C. H., and M. Da Silva, 2004, Seismic inversion methods and some of their constraints: First Break, **22**, 47–70, http://doi.org/10.3997/1365-2397.2004011.