Geophysical Prospecting

EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Geophysical Prospecting, 2018, **66**, 904–919 doi: 10.1111/1365-2478.12625

Joint facies and rock properties Bayesian amplitude-versus-offset inversion using Markov random fields

James Gunning[1]* and Mark Sams[2]

[1]*CSIRO, and* [2]*Ikon Science*

Received March 2017, revision accepted January 2018

ABSTRACT

Seismic reflection pre-stack angle gathers can be simultaneously inverted within a joint facies and elastic inversion framework using a hierarchical Bayesian model of elastic properties and categorical classes of rock and fluid properties. The Bayesian prior implicitly supplies low frequency information via a set of multivariate compaction trends for each rock and fluid type, combined with a Markov random field model of lithotypes, which carries abundance and continuity preferences. For the likelihood, we use a simultaneous, multi-angle, convolutional model, which quantifies the data misfit probability using wavelets and noise levels inferred from well ties. Under Gaussian likelihood and facies-conditional prior models, the posterior has simple analytic form, and the maximum a-posteriori inversion problem boils down to a joint categorical/continuous non-convex optimisation problem. To solve this, a set of alternative, increasingly comprehensive optimisation strategies is described: (i) an expectation–maximisation algorithm using belief propagation, (ii) a globalisation of method (i) using homotopy, and (iii) a discrete space approach using simulated annealing. We find that good-quality inversion results depend on both sensible, elastically separable facies definitions, modest resolution ambitions, reasonably firm abundance and continuity parameters in the Markov random field, and suitable choice of algorithm. We suggest usually two to three, perhaps four, unknown facies per sample, and usage of the more expensive methods (homotopy or annealing) when the rock types are not strongly distinguished in acoustic impedance. Demonstrations of the technique on pre-stack depth-migrated field data from the Exmouth basin show promising agreements with lithological well data, including prediction accuracy improvements of 24% in v_p/v_s and twofold in density, in comparison to a standard simultaneous inversion. Much clearer and extensive recovery of the thin Pyxis gas field was evident using stronger coupling in the Markov random field model and use of the homotopy or annealing algorithms.

Key words: Inverse problem, Inversion, Reservoir geophysics, EM algorithm, AVO, Facies, Bayesian, belief propagation, Markov random field.

INTRODUCTION

The oil and gas business has traditionally used an industrialised seismic workflow that consists of distinct, commoditised steps of acquisition, processing, velocity model building, migration, and in recent decades quantitative interpretation (QI). This reflects a natural characteristic of the seismic inverse problem under standard acquisition, where low-frequency kinematic character (up to ≈ 2 Hz) is used to inform coarse but "stable" velocity models, and the

*E-mail: James.Gunning@csiro.au

mid/higher frequency character (> 10 Hz) is used to image bandpassed (but "stable") reflectivity. Usually, these imaging schemes have academic lineage to a scale-splitting assumption and a Born-like scattering model (Bleistein, Cohen and Stockwell 2001). For the migration to bandpass reflectivity, the Born approximation leads to linear models and convex optimisation, which is independent of starting models, and largely independent of any prior model information. Up to the QI step, the aim has been to ensure that each process produces stable outputs from its inputs (like most commodity sequential workflows) and avoids making predictions with significant model dependencies other than those derived directly from the acquired data.

This industrial workflow is a natural consequence of the missing, typically 2–10 Hz, frequency gap in conventional imaging: since the full inverse problem has a near null-space in this frequency gap, it makes most sense to try and image to a projected space, which is orthogonal in some sense to it, and "bandpass reflectivity" meets this requirement. Despite much hard work by the broadband community, there are strong physical constraints that apply on the acquisitional side, such as the low-frequency airgun response (Long and Reiser 2014), operational limits on long offset design, and the diminishing natural earth reflectivity at low frequencies (Walden and Hosken 1985), that suggest that weak or challenging signal-to-noise ratios over this missing gap are likely to remain a constant challenge in imaging. There are exciting new developments in ultralow vibrational sources (Dellinger *et al.* 2016), but it is not yet clear whether energy, environmental, or deployment constraints will make these sources a commodity technology. Thus, even though high-end imaging methods like full-wave inversion overcome the scale-splitting and Born-approximation limitations of conventional migrations (Virieux and Operto 2009), the challenges of missing spectral content and limited apertures from conventional streamer acquisition remain.

We observe that the most common industrial practice is for reservoir characterisation teams to work with "bandpass reflectivity" images, so the issue of the missing frequency gap is then effectively their "responsibility"; it is reasonable to suggest that this arrangement may endure for some time. Bandpass reflectivity images can be produced as true–amplitude, where amplitude effects unrelated to those intrinsic to the earth reflectivity have been removed. The migration velocity and bandpass reflectivity produced by standard workflows are then maximum–likelihood inferences under a suitable misfit model. The imaged amplitudes can then be regarded as "data" in a downstream amplitude-versus-offset (AVO) inversion workflow, where the forward model can now be approximated as a 1D convolutional problem with a wavelet that can be estimated from well ties (Gunning and Glinsky 2006). Under this approximation, which we maintain in the rest of this paper, 3D ray-propagation effects are assumed to have been effectively removed. The well tie provides misfit noise estimates which approximately embed the adequacy of this model.

Depending on your definition of "signal", the S:N ratios commonly observed at well ties based on the convolutional model may be around 2 to 4, and better than 10 is very rare. This makes the power-band of useful amplitude data often rather smaller than the spectral width which is sought in inversions. The missing spectral data are commonly supplied by an explicit model preference statement, usually as a low-frequency velocity/density model, and often stated as a regularisation term. We prefer to write this regularisation as a Bayesian prior, since Bayesian frameworks have a rich capacity for incorporating other kinds of data about velocities, e.g., regional models or well-logs, in a way that is conceptually separated from the seismic data. In many cases, the low-frequency background model is stated as a l_2 regularisation term, which corresponds to a Gaussian prior distribution.

The construction of this prior model of low-frequency velocity has become one of the major—and problematic—tasks in AVO-style inversion. Where there are abundant and stratigraphically correlatable well data, in clastic systems for example, the model is usually constructed from some kriged low-frequency weighted average of, for example, shale and "reservoir" velocities. However, most models do not enjoy such stratigraphic clarity, and spatially smoothed interpolations of net-to-gross are difficult and controversial models to build. What is usually much less difficult or contestable is the building of spatial and compaction models for individual rock types produced by log classification. In particular, non-seismic sources of information for shear velocity and density are crucial for making the connection to reflectivity. Sometimes, these compaction trends may be successfully applied over very significant lateral distances, reflecting perhaps the ubiquity of the typical rock textures and mineralogy within basins.

Most seismic energy comes from transitions between facies types, and low frequency models are typically a heavily smoothed approximation to the blocky facies character of well logs. Without facies parameters, the prior model has to form a kind of Gaussian covering "umbrella" over all facies types, which means that posterior inferences may generate predictions in transition regions corresponding to no existing rock type in regions where clean facies may be expected. Especially

away from seismic reflection events, these standard inversions tend to relax back to the background model quite quickly, and this may easily correspond to some kind of "mongrel" rock with respect to the desired mixture distribution. Posterior samples drawn from a distribution formed from this umbrella-prior model will certainly not capture the mixture character typical of sedimentary formations. The absence of mixture character follows from the mathematical fact that the product of a Gaussian prior and a Gaussian likelihood forms a Gaussian posterior, whereas mixtures have, in general, several "humps". This is most obvious at lower frequencies where the near-zero amplitude information (flat likelihood) by definition cannot force the Gaussian prior into a posterior mixture.

We hypothesise that a more geologically faithful way to supply this low-frequency character is to explicitly model the discrete or facies aspect of the subsurface and let a Bayesian probabilistic model of spatial variation in facies type, combined with compaction trends, implicitly supply the required low-frequency information. Other groups have tried similar models; for example, hierarchical discrete/continuous models for seismic inversion have been used by Eidsvik *et al.* (2004), Larsen *et al.* (2006), Ulvmoen and Omre (2010), Buland *et al.* (2008), and Kolbjørnsen *et al.* (2016). Interestingly, for the reasonably realistic examples discussed in Larsen *et al.* (2006) and Ulvmoen and Omre (2010), at the attempted resolution, the uncertainty in the posterior facies distribution does not appear to be very great, despite the legitimate concerns over this issue raised by Kolbjørnsen *et al.* (2016). This means that a maximum aposteriori (MAP) inference for facies and rock properties is likely to be very useful, furnishing a single "best guess" model for prediction purposes.

This paper describes an AVO inversion approach that jointly inverts for elastic parameters and facies labels, using facies-dependent compaction models, an explicit spatial probability model for facies, and with an emphasis on MAP estimation for the resulting models (Kemper and Gunning, 2014). The spatial model has similarities to those in the references cited above, but our computational approaches are quite different. For example, the approximate marginals used in Buland *et al.* (2008) are a good idea, but they only take second-order spatial interactions into account—the scheme we describe below has all higher order interactions. We aim to confine the scope of this paper to the MAP estimation problem, but the apparatus derived does give access to uncertainties, and examination of these for typical test problems with sensibly framed models and decent S:N ratios generally conveys the same impression about uncertainties observed by Larsen *et al.* (2006) and Ulvmoen and Omre (2010).

The material below is organised as follows. In the Theory section, we introduce the model and its assumptions and explain why the optimisation problem it introduces is very challenging. Next, we describe a sequence of increasingly comprehensive algorithms for solving this optimisation problem, trading quality against runtime and resources. A discussion of practical issues and preparatory modelling work is offered in a summary of the typical workflow. This is then illustrated by a short synthetic study, then a field-data case using an Australian offshore data set, followed by some recommendations and conclusions in the usual way.

THEORY

We frame the inversion problem to solve for a suite of continuous rock properties $\mathbf{m} = \{v_p, v_s, \rho\}$ (P–velocity, S–velocity, density) together with facies labels $\mathcal{F}_i \in \mathcal{L} \equiv \{0, 1 \ldots, N_f - 1\}$ (e.g., shale, limestone, brine-sandstone, ...) at all the sample locations i in the inversion volume (usually the *xyt* lattice defined by the imaged seismic). The prior distribution of rock properties $P(\mathbf{m}|\boldsymbol{\mathcal{F}})$ is a conditional joint distribution, dependent on co-located facies labels, and is constructed from rock physics regressions developed from regional log data. The facies labels are modelled with a Markov random field (MRF) distribution $P(\boldsymbol{\mathcal{F}})$(Winkler 2003), whose underlying connectivity graph is aligned with the natural stratigraphy in the inversion domain. Thus, given suitable AVO angle (θ) stacks $\mathbf{y}$ (multiple samples and angles), a forward model $\mathbf{f}(\mathbf{m})$, and a suitable probability measure for the likelihood of the data given the model, $L(\mathbf{y}|\mathbf{m}) = L(\mathbf{y}|\mathbf{f}(\mathbf{m}))$, the joint posterior distribution of the model $\mathbf{m}$ and $\boldsymbol{\mathcal{F}}$ can be written as

$$P(\mathbf{m}, \boldsymbol{\mathcal{F}}|\mathbf{y}) \sim L(\mathbf{y}|\mathbf{m})P(\mathbf{m}|\boldsymbol{\mathcal{F}})P(\boldsymbol{\mathcal{F}}). \quad (1)$$

The forward model $\mathbf{f}(\mathbf{m})$ is usually taken to be a simple convolutional model with reflectivities taken in either linearised ($\mathbf{m} \equiv \{\log(\rho v_p), \log(\rho v_s), \log \rho\}$, (Fatti *et al.* 1994)) or Zoeppritz (Aki and Richards 1980) form, consistent with the angle stack and wavelet signatures and effective noise levels estimated from nearby well ties. Typically, for the Fatti approximation, this forward model may be written as $\mathbf{y} = \mathbf{f}(\mathbf{m}) = \mathbf{X}\mathbf{m}$ where the matrix $\mathbf{X}$ is a product of discrete time differentiation, Fatti-like reflectivity coefficients, and wavelet convolution. In this case, $\mathbf{X}$ is the Jacobian or Frechet derivative and has an unstable pseudo-inverse, which reflects both the frequency content loss mentioned above, but also weak ($O(\theta^2)$) sensitivity to some projections of the model vector $\mathbf{m}$, e.g., shear impedance, or density.

For convenience, we choose the conditional priors to be of multi-Gaussian form, i.e.,

$$P(\mathbf{m}|\boldsymbol{\mathcal{F}}) = N(\bar{\mathbf{m}}(\boldsymbol{\mathcal{F}}), \mathbf{C}_\mathrm{p}(\boldsymbol{\mathcal{F}})) \sim \exp\left[-\tfrac{1}{2}(\mathbf{m}-\bar{\mathbf{m}}(\boldsymbol{\mathcal{F}}))^T \times \mathbf{C}_\mathrm{p}{}^{-1}(\boldsymbol{\mathcal{F}})(\mathbf{m}-\bar{\mathbf{m}}(\boldsymbol{\mathcal{F}}))\right] / |\mathbf{C}_\mathrm{p}(\boldsymbol{\mathcal{F}})|^{1/2}, \quad (2)$$

where $\bar{\mathbf{m}}$ and $\mathbf{C}_\mathrm{p}$ are assembled to capture the rock physics trends and regressions for each rock type. We use a primary compaction model for v_p as a function of depth or effective stress, and v_p as a primary predictor for v_s and ρ as a regression structure. Specifically, for a known reference compaction surface t_ref, the mean properties are defined for each facies by a set of log-data-derived conditional regression models like $v_p \sim N(A_p + B_p(t - t_\mathrm{ref}), \sigma_p^2)$, $v_s|v_p \sim N(A_s + B_s v_p, \sigma_s^2)$, and $\rho|v_p \sim N(A_\rho + B_\rho v_p, \sigma_\rho^2)$, where the σ denotes regression error standard deviations. The suite of regression parameters $\{A_p, B_p, A_s, B_s, A_\rho, B_\rho, \sigma_p, \sigma_s, \sigma_\rho\}$ in this model then defines block entries in $\bar{\mathbf{m}}(\boldsymbol{\mathcal{F}})$ and the covariance matrix $\mathbf{C}_\mathrm{p}(\boldsymbol{\mathcal{F}})$ for each sample t. For simplicity, we suppress notation for the deterministic spatial dependence implicitly carried by t, t_ref, and possible transverse variations in the regression constants. The expected velocity increase with depth is thus implicitly carried by $\bar{\mathbf{m}}(\boldsymbol{\mathcal{F}})$, and $\mathbf{C}_\mathrm{p}(\boldsymbol{\mathcal{F}})$ embeds, for example, the characteristically strong correlation between v_p and v_s. This structure implies that the conditional prior is written as a product over all lattice sites ("blocks"). The likelihood is also conveniently chosen to be of Gaussian form, which we write as $L(\mathbf{y}|\mathbf{m}) \sim N(\mathbf{y} - \mathbf{f}(\mathbf{m}), \mathbf{C}_\mathrm{d})$ where $\mathbf{C}_\mathrm{d}$ is a suitable noise covariance matrix. The covariance of the residuals $\mathbf{y} - \mathbf{f}(\mathbf{m})$ implicitly absorbs external noise and a range of modelling and imaging errors.

The MRF for the facies labels is parameterised as an unsymmetric "Potts" model

$$P(\boldsymbol{\mathcal{F}}) \sim \exp\left(-\sum_i \left(\frac{1}{2}e_{i,\mathcal{F}_i} + \sum_{j\sim i} \beta^{(D_{ij})}_{\mathcal{F}_i,\mathcal{F}_j} I(\mathcal{F}_i \neq \mathcal{F}_j)\right)\right), \quad (3)$$

where i, j are sample/voxel locations, $j \sim i$ denotes voxel neighbours j of i, $I(\cdot)$ is an 0/1 indicator function, the $e_{i,l}$ ($l \in \mathcal{L}$) are pseudo-abundance parameters, and $\boldsymbol{\beta} = \{\beta^{(D_{ij})}_{\mathcal{F}_i,\mathcal{F}_j}\}$ are parameters coupling the neighbours in stratigraphic direction D_{ij} (a direction sense associated with the neighbouring voxels i, j: along-dip, vertical etc.). These $\boldsymbol{\beta}$ parameters are chosen to promote continuity for "permissible" transitions and are also used to preclude impermissible transitions, for example, vertical juxtapositions of brine above oil in communicating reservoir facies. For permissible transitions along stratigraphic directions, the $\boldsymbol{\beta}$ parameters are typically fixed at values not too far from the phase-transition of the equal-abundance case ($e_{i,l} = 0$) MRF, where the resultant spatial correlations of "representative" models drawn from the prior distribution are most geologically realistic: this is typically around $\beta \approx 0.7$ in 2D or $\beta \approx 0.4$ in 3D (Huang 1987). Prior proportions are encoded through the terms $e_{i,l}$, which are automatically adjusted in a calibration step such that the overall marginal distribution of each label, computed from equation (3), matches, voxelwise, a volume of geologist/interpreter-supplied marginal prior-proportions $p_{i,l}$. We assume henceforth that such a volume is available: its information content can be anything from strongly informative to relatively agnostic.

The step of calibrating the pseudo-abundance energies $e_{i,l}$ to the prior proportions $p_{i,l}$ is nontrivial: in the absence of correlations in the facies model (all $\beta = 0$), the marginal abundance $p_{i,l} = \langle I(\mathcal{F}_i = l)\rangle$ of facies l at pixel i relates to the $e_{i,l}$ by the trivial relation $p_{i,l} = \exp(-e_{i,l}/2)$. In the presence of correlations, the relationship is not analytically available, although it is smooth. Numerically, it is well approximated by the application of loopy Bayesian belief propagation (LBBP) to the prior distribution (equation (3)) (Wainwright and Jordan 2008), which computation we may notate as $\langle I(\mathcal{F}_i = l)\rangle = G_{il}(\boldsymbol{\beta}, \{e_{i,l}\})$. Since the LBBP calculation is fast, our setting of the parameters $e_{i,l}$ is performed by solving the mini-inverse matching problem:

$$\{e_{i,l}\} = \arg\min_{e_{i,l}} \sum_{i,l} (p_{i,l} - G_{il}(\boldsymbol{\beta}, \{e_{i,l}\}))^2 \quad (4)$$

for the $e_{i,l}$, which usually matches the marginal proportions within a percent or so. The details of this matching problem are not central to what follows and outside the scope of this paper. But the calibration step is important since the correlated MRF behaves very differently to a simple uncorrelated model, and the setting of the $e_{i,l}$ must take this into account: the mutual consistency of the $e_{i,l}$ and β parameters is important for the behaviour of the optimisation approaches below.

INVERSION

"Standard" inversion can now be regarded, statistically, as a model "point-estimate" obtained by maximising the posterior probability (equation (1)). This is equivalent to minimising, jointly with respect to $\mathbf{m}$ and $\boldsymbol{\mathcal{F}}$, the log-posterior energy:

$$E(\mathbf{m}, \boldsymbol{\mathcal{F}}) = (\mathbf{y} - \mathbf{f}(\mathbf{m}))^T \mathbf{C}_\mathrm{d}{}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{m})) + (\mathbf{m} - \bar{\mathbf{m}}(\boldsymbol{\mathcal{F}}))^T \mathbf{C}_\mathrm{p}{}^{-1}(\boldsymbol{\mathcal{F}}) (\mathbf{m} - \bar{\mathbf{m}}(\boldsymbol{\mathcal{F}})) + \log|\mathbf{C}_\mathrm{p}(\boldsymbol{\mathcal{F}})| + \sum_i \left(e_{i,\mathcal{F}_i} + \sum_{j\sim i} 2\beta_{\mathcal{F}_i,\mathcal{F}_j} I(\mathcal{F}_i \neq \mathcal{F}_j)\right). \quad (5)$$

In its general form, this optimisation problem is very challenging, as it is non-convex with respect to most "neighbour" definitions in the labelling space, although it is convex in $\mathbf{m}$ for any fixed labelling (though perhaps, weakly non-convex for the Zoeppritz model at large angles). It belongs to the class of NP-hard optimisation problems for which guaranteed exact solutions are not available from algorithms that scale acceptably with the dimension of the model. Among a large number of approaches we have tried, the most successful are described below, and the three methods we outline correspond to increasing resource requirements and globality claims. These should be regarded as a set of alternative approaches trading off solution time versus quality. The first is a local optimisation strategy based on the expectation-maximisation (EM) algorithm, the second, a globalised version of the first using a graduation strategy and the last, a fully globalised approach based on annealing.

(i) EM algorithm. The local optimisation strategy we have found to be most efficacious is a form of EM algorithm, where, alternately, the labels $\boldsymbol{\mathcal{F}}$ are estimated in an expectation "E-step", and the rock properties $\mathbf{m}$ then estimated in a maximisation "M-step", repeatedly until convergence. In the E-step, we form the $\boldsymbol{\mathcal{F}}$-expectation of the log-posterior with the properties $\mathbf{m}$ taken as fixed ("current values"). This is most simple to see in the case that the prior covariance factors over voxels, in which case it is essentially of the form $E(\boldsymbol{\mathcal{F}}) = \sum_i (u_{\mathcal{F}_i} + \sum_{j\sim i} \beta_{\mathcal{F}_i,\mathcal{F}_j} I(\mathcal{F}_i \neq \mathcal{F}_j))$, which is a standard "pairwise graphical model", whose marginal probabilities $z_{il} = \langle I(\mathcal{F}_i = l)\rangle$ $(l \in \mathcal{L})$ are estimated by loopy belief propagation (Wainwright and Jordan 2008). These marginal probabilities can then be regarded as soft estimates of the labels $\boldsymbol{\mathcal{F}}$. This algorithm both scales with the system size and parallelises over traces. In the M-step, these (current) marginal probabilities ("memberships") are fixed, and we minimise a membership-reweighted objective with respect to the elastic parameters $\mathbf{m}$. In the simple case above, this ends up as a reweighted least-squares objective $\mathbf{Q}(\mathbf{m}) = (\mathbf{y} - \mathbf{f}(\mathbf{m}))^T {\mathbf{C}_\mathrm{d}}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{m})) + \sum_{il} z_{il} (\mathbf{m}_i - \bar{\mathbf{m}}_i)^T {\mathbf{C}_\mathrm{p}}^{-1}(l)(\mathbf{m}_l - \bar{\mathbf{m}}_l)$. The latter is easily solved with a modified LSQR algorithm (Paige and Saunders 1982) and again scales well with the system size and enjoys parallelisability. Of a variety of local optimisation approaches, this method, combined with some step-length limiting on membership changes, has been observed to exhibit the widest basin of attraction around high-quality optima (Gunning and Kemper 2012). The starting configuration we use comprises setting initial memberships guesses from the supplied prior proportions and commencing with an M-step. EM algorithms produce parameter estimates at the mode of the marginal distribution obtained by integrating/summing out the discrete variables. As a surrogate for the MAP point, we form a pointwise facies-label estimate by greedy Bayesian classification of the final memberships z_{il}. The compute requirements here are typically 10–50 times those of simple classical least-squares inversion.

(ii) Homotopy. Given an efficient local method, non-convex problems usually benefit from some kind of graduation or homotopy apparatus to increase the probability of finding global optima from arbitrary starting points. In these graduation strategies, the objective function is gradually deformed from a suitable convex approximation of the original problem towards the original non-convex objective. The current local optimum is used as a starting point for the next problem in the graduation sequence. Since the non-convexity originates in the mixture distribution $P(\mathbf{m}|\boldsymbol{\mathcal{F}})$ used to model the rock physics, we define the homotopy schedule by gradually peeling this mixture away from a common distribution in the central part of the elastic parameter space. Specifically, for a schedule of values $\lambda = 0, \Delta\lambda, \ldots 1$, the facies means of the mixture (equation (2)) are mapped to $\bar{\mathbf{m}}'_l = \lambda \bar{\mathbf{m}}_l + (1-\lambda)\bar{\mathbf{m}}_*$, where $\bar{\mathbf{m}}_*$ is a prior facies-probability-weighted average of $\{\bar{\mathbf{m}}_l\}$. A similar graduation sequence is used for the covariances $\mathbf{C}_\mathrm{p}(l)$. The inner local optimisation routine here is the EM algorithm of item (i). The resource requirements here are usually 5–10 times that for the EM algorithm.

(iii) Global annealing. Here, we use a reasonable approximation of the prior rock physics model that leads to a more manageable purely discrete global optimisation problem. We approximate the covariances $\mathbf{C}_\mathrm{p}(l)$ by a common prior dispersion $\mathbf{C}_\mathrm{p}$ (a prior–proportions weighted average), use the fully linear (Fatti) forward model $\mathbf{f}(\mathbf{m}) = \mathbf{X}\mathbf{m}$, and set the means by a selection matrix $\mathbf{Z}$, such that $\bar{\mathbf{m}} = \mathbf{Z}\tilde{\boldsymbol{\mathcal{F}}}$, where $\tilde{\boldsymbol{\mathcal{F}}}$ is a binary-indicator equivalent representation of $\boldsymbol{\mathcal{F}}$. The minimum of equation (5) in the subspace of $\mathbf{m}$ is available analytically and can be substituted back in to the same equation to leave a discrete objective function in the $\tilde{\boldsymbol{\mathcal{F}}}$–space that is quadratic in the term $\bar{\mathbf{m}}(\tilde{\boldsymbol{\mathcal{F}}})$. This can be simplified in general to a constrained binary/integer quadratic programming (BQP/IQP) problem for $\tilde{\boldsymbol{\mathcal{F}}}$ with objective $E(\tilde{\boldsymbol{\mathcal{F}}}) = \tilde{\boldsymbol{\mathcal{F}}}^T \mathbf{Q} \tilde{\boldsymbol{\mathcal{F}}}$, for a huge but sparse matrix $\mathbf{Q}$. A very attractive property of this result is that precisely the same objective function holds also for the marginal distribution of $\tilde{\boldsymbol{\mathcal{F}}}$, i.e., where all the uncertainty from the elastic parameters has been integrated out. Further, the matrix $\mathbf{Q}$ has modest density in the vertical indexing direction corresponding to de-convolution, but there are only sparse fringes in the x, y stratigraphic directions

originating in the MRF. About half the entries are positive, expressing anti-correlation, and yield a property known as non-submodularity in discrete optimisation (Wainwright and Jordan 2008), which means the minimisation problem is NP-hard. In some IQP problems where **Q** has only a very narrow bandwidth k of nonzero entries, solutions are possible using dynamic programming (Rimstad 2012), but the complexity scales like N_f^{k+1}. The typical number of samples in wavelets ($k > 8$, say), plus the fringes from the MRF, means that this idea is not practicable. The upshot of this is that the IQP form of the optimisation problem is most practically solved using a highly customised simulated annealing scheme; our implementation is based on Katayama and Narihisa (2001) with many enhancements for speed and parallelised. The resource needs here are relatively heavy for memory and speed, since the linear algebra must be done in partially dense form to populate the **Q** explicitly. The annealing schedule is implemented in a typical geometric scheme, with the temperature lowered after a full pixel sweep of random label-flip proposals. We find that even with temperate cooling ratios as aggressive as 0.95, the bulk of the labelling inference settles in fairly robustly, and the total annealing time is comparable to the linear-algebra setup time. High-quality global optima are best generated with more cautious annealing rates (e.g., 0.995), so compute times will grow accordingly. This method has been validated by comparison runs with solutions using exact binary optimisation methods, on small problems where the latter option is feasible. A final benefit of the annealing scheme is that, since the theory of Markov Chain Monte Carlo (MCMC) is intimately connected with that of simulated annealing, approximate samples from the Bayesian posterior can be generated by arresting the annealing at temperature of $T = 1$, drawing a sample, then melting and re–annealing repeatedly. Though expensive, such a scheme ensures multiple modes are visited in sampling, in contrast to the strong tendency of cheap MCMC proposal schemes to get trapped in local minima.

The choice of algorithm for practical use depends on resource availability and the size of the inversion to be performed. It also depends on the level of difficulty or ambitiousness of the model. Keeping the level of ambitiousness in the model realistic is important. An old and reliable rule in discrete optimization is the principle of breaking of symmetries. In the context of discrete facies inversion this means formulating models such that label interchange near symmetries or hidden effective media symmetries (e.g., laminated hard/soft facies behaving as "intermediate") operate as little as possible. In the choice of the effective facies types and the resolution of the model sought, these principles need to be kept in mind. For example, it is clear that the inversion framework should operate decently when facies distinctions are clear in acoustic impedance (AI) log data. Conversely, the odds of recovery will be diminished when there is great overlap of AI character caused by, e.g., diagenetic alteration of material. Judicious choices of proportions and use of the MRF parameters to promote spatial continuity can help break symmetries. The cheapest algorithms tend to work well only if no near symmetries exist. For example, if the seismic is resampled and recovery at resolution below the $\lambda/4$ Rayleigh limit is pursued, the parameters in the MRF should break the symmetry (e.g., by setting unequal prior proportions, enforcing density ordering of fluids in permeable facies, etc...). Amongst facies sequences for which the seismic likelihood is very flat, the posterior preference (and implicit low frequency trend) is strongly driven by the prior abundances. We do not recommend working very far below the Rayleigh resolution limit, simply because of the large combinatorial explosion of effective media assemblies that can occur sub-resolution and the commensurate growth in model uncertainty.

In summary, we find that problems with two or three facies that separate well in AI are often cleanly recovered using the cheapest option above, but the initial **m** guess has to be not too far off since there are spurious secondary minima into which the EM algorithm can be attracted. Use of the prior-weighted mean as an initial guess seems to achieve this. Separating facies along directions that have nearly constant AI is usually rather more delicate, as it depends on weaker information from the larger angle stacks. Such problems benefit greatly from the globalisation strategies of methods 2 or 3. For about 10-fold the runtime, the homotopy algorithm (equation (2)) offers about the best level of globality that can be achieved with the modest memory resources required by a local method. We recommend it as a basic workhorse. For cases with more facies, especially those not well distinguished in AI, the annealing method may be recommended if memory resources are adequate: the overall runtime may then be another 10-fold slower or so for respectable cooling rates like 0.99.

Inversion codes always make prediction errors, and the foregoing material emphasises the importance of adequate optimisation in minimising these. However, predictive errors in Bayesian inversion can occur for at least three reasons: (i) external noise in the data, over which no inversion scheme has any real power if the noise overlays the signal in space and spectral character: we discuss this briefly in the Workflow section below; (ii) inadequacies in optimisation methodologies, such as premature terminations in optimisation schemes, or

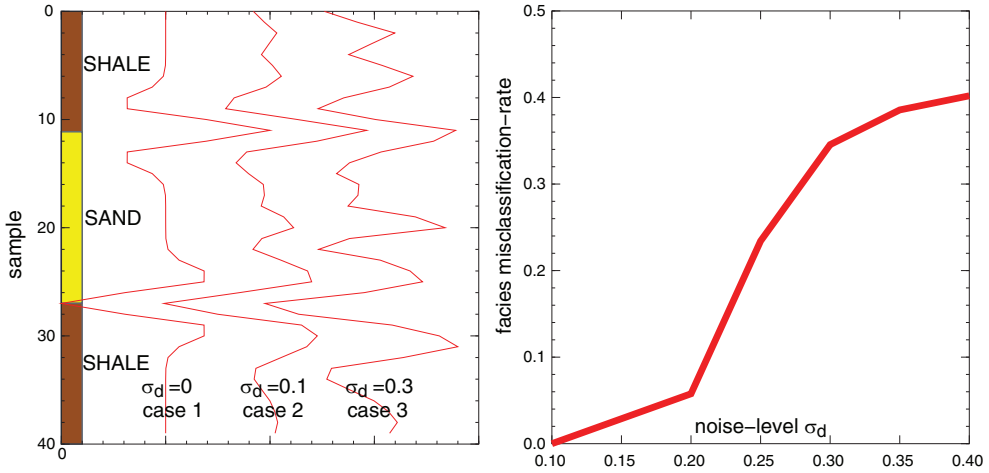


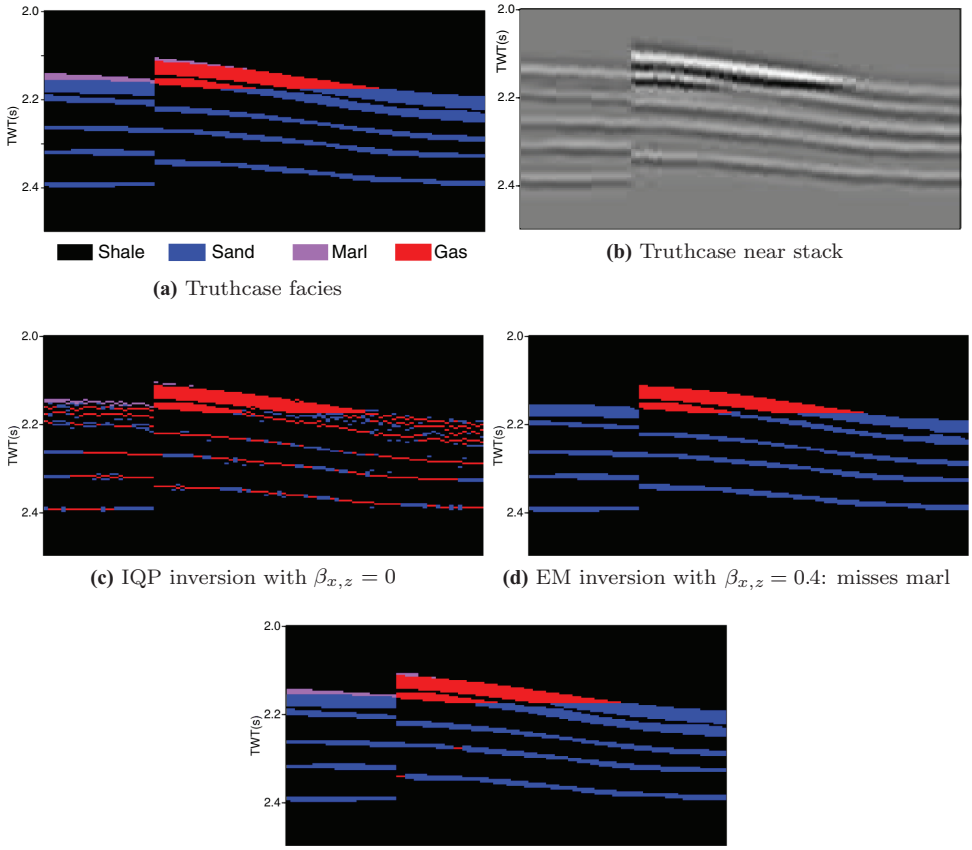
Figure 1 Left, "truth" sand-shale model with seismic data corrupted by Gaussian noise, of RMS power σ_d relative to the peak signal of 1. Three cases with $\sigma_d = 0, 0.1,$ and 0.3 are shown. Under the Monte-Carlo experiment of recovering facies labels, using the rigorous IQP algorithm, from multiples instances of noisy seismic traces, the misclassification rate as a function of noise level is shown on the right, rising steeply beyond $\sigma_d = 0.2$.

trapping in local minima, as discussed above; and (iii) inadequacies in prior model formulations, such that even a noise-free signal and perfect global optimisation produces incorrect predictions. Researchers in computer vision who use MRFs have observed that the last issue can sometimes be a dominant consideration, since it is hard for natural scenes or geology to be well modelled by simple first-order MRFs (Szeliski *et al.* 2006). It is thus sometimes the case that the extra effort required to achieve high-quality global optima via annealing will not improve the prediction accuracy. The synthetic example below illustrates the importance of point (ii) above, whereas the field case illustrates aspects of points (i) and (iii).

WORKFLOW

The basic workflow requirements in setting up an inversion comprise (i) wavelet and noise estimation from well ties (wavelets enter the forward model $\mathbf{f}(\mathbf{m})$, and noise levels the covariance $\mathbf{C}_\mathrm{d}$), (ii) rough (auto)picking of zonations and basic stratigraphic frameworks, and (iii) petrophysical log classification (to "effective elastic facies") and derivation of per facies compaction trends, relative to sensible reference surfaces. The latter step usually requires some sensible pooling of facies types, as elastic inversions require some separability of rock types in impedance space. The primary sensitivity is of course to AI, and far-offsets yield increasing capacity to exploit v_p/v_s ratio contrast. Further, most regions will have some known major horizons, where it is known that particular facies only occur in related subzones, e.g., above a major epoch marker. This regional knowledge is represented in the prior-proportions $\{p_{il}\}$ specification, which is set by geological judgement and automatically assembled with respect to horizons in a simple workflow. Typically, fixed proportions are declared for facies within a small number of subzones, with some graded transitions over a few samples.

The prior proportions specification is usually generated, such that only (say) three to four facies are possible at any given depth. This is usually a sensible upper limit on the number of categories that can be distinguished by AI and v_s/v_p ratio. The argument in support of this limit runs roughly as follows. It is clear that even a perfect model and algorithm will produce misclassifications if the level of external noise is high enough. To illustrate this, a simple synthetic Monte Carlo experiment to recover sand/shale labels from a simple single-trace, two-facies model with a sand slab buried in shale is depicted in Fig. 1. The Monte Carlo part entails adding noise of matching spectrum to the truth-case data, of RMS amplitude σ_d relative to a peak truth-case amplitude of 1 (i.e. S:N= $1/\sigma_d$). For the model, the MRF parameters are set at equal proportions, with $\beta = 0$ (vertically), and a rigorous (exact) inversion algorithm is used (i.e., the global optimum of equation (5) is found). The average misclassification rate, shown in Fig. 1, shows a steep rise at around S:N= 5 ($\sigma_d = 0.2$), with obvious asymptotic limit of 0.5 for random guessing. For more than two facies, the asymptotic limit will increase, and one should expect even steeper deterioration with noise, so three or four facies seems like a sensible



(a) Truthcase facies

(b) Truthcase near stack

(c) IQP inversion with $\beta_{x,z} = 0$

(d) EM inversion with $\beta_{x,z} = 0.4$: misses marl

(e) IQP inversion with $\beta_{x,z} = 0.4$

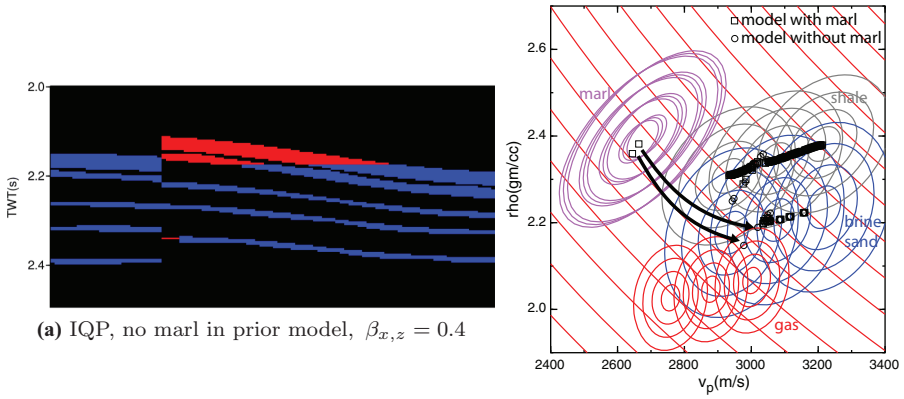
Figure 2 Synthetic inversion model; effects of spatial coupling (β values) and EM (local) versus IQP (global) optimisation: global methods are needed to find the marl.

pragmatic upper limit if the misclassification rate is expected to be low at typical S:N ratios.

The depth-trend and rock physics curves are simply derived by guided semi-automatic regression routines applied to upscaled log data on a per-facies basis; this informs the modelling structure in the trend $\bar{\mathbf{m}}(\mathcal{F})$ and correlation $\mathbf{C}_{\mathrm{p}}(\mathcal{F})$. Typically, given a reference depth/time surface t_{ref} suitable for effective stress (e.g., the mudline or "base Cretaceous"), this consists in fitting the free parameters and regression errors in the set of fitting/regression models $v_p \sim N(A_p + B_p(t - t_{\mathrm{ref}}), \sigma_p^2)$, $v_s|v_p \sim N(A_s + B_s v_p, \sigma_s^2)$, $\rho|v_p \sim N(A_\rho + B_\rho v_\rho, \sigma_\rho^2)$, where the σ are regression error standard deviations. These parameters then define subblocks of the covariance matrix $\mathbf{C}_{\mathrm{p}}$ at each t, t_{ref}.

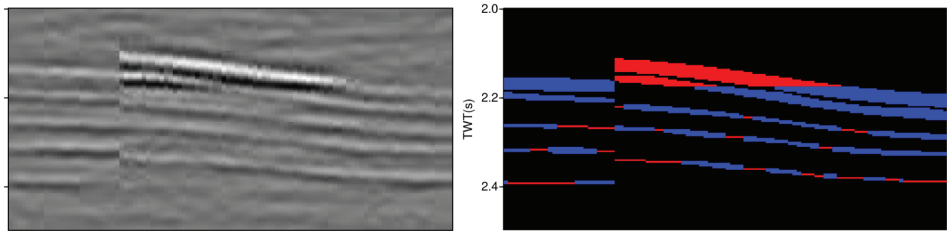
ILLUSTRATIVE SYNTHETIC MODEL

Before proceeding to the actual field example, we show here some salient behavioural characteristic of the inversion on synthetic data to illustrate some of the modelling and algorithmic issues discussed in previous sections. The rock physics characteristics of this model have modestly close similarities to those of the following field data study. Our synthetic model and study has inverse-crime characteristics and begins with zero-noise cases, but this simplifies the presentation of the salient issues. Figure 2(a) shows a synthetic 2D model with facies for shale(0), marl (1), brine–sandstone (2), and gas (3), with seismic data at 10°, 40° used for inversion (Fig 2(b)). The marginal pdfs of $\{v_p, \rho\}$ for each facies are shown as sets of ellipses in Fig 3(b), for the top, mid, and bottom of the model, with diagonal contours for AI. Figure 2(c) shows an inversion with no spatial continuity imposed ($\beta_{x,z} = 0$): the MAP model here finds an effective-medium degeneracy in the brine leg and settles on a salt-and-pepper like model. Imposing spatial continuity ($\beta_{x,z} = 0.4$) in the EM inversion of Fig. 2(d) breaks this symmetry and recovers the brine successfully, but misses the marl: the impedance overlap between marl and brine is significant, and the local optimisation finds a



(a) IQP, no marl in prior model, $\beta_{x,z} = 0.4$

(b) shift of classification under omitted marl facies



(c) near–stack plus noise

(d) IQP inversion of noisy data, $\beta_{x,z} = 0.4$

Figure 3 Synthetic inversion model; (a) and (b) illustrate the effect of omitting a facies (marl) from the model. See main text for explanation of contours. The data points in (b) are final elastic values for a trace a little left of the fault. Insets (c) and (d) show the likely effects of noise on facies recovery for this model.

brine solution. Figure 2(e) uses the global annealing approach and manages to recover the marl as well. Figure 3(a) shows an inversion where the marl is a "surprise facies" and omitted in the prior: here, the inversion unsurprisingly classifies that region of space as the facies with nearest required impedance (here, brine), see Fig. 3(b). Though not shown, in this marl region, the brine impedance is pulled down. Figure 3(c) and (d) show representative likely effects of noise on the recovery—here the seismic is contaminated by correlated, bandlimited noise of RMS reflectivity about 0.01. The general recovery is quite good, but thinner bodies are rather vulnerable to the noise.

AUSTRALIAN NORTH–WEST SHELF EXAMPLE

Here, we show a practical application of the joint facies and rock properties inversion to highlight some of the features of the different solution algorithms and to compare with a standard deterministic, model-based simultaneous inversion. The inversions were applied to pre-stack depth-migrated (PSDM) images (courtesy of Searcher Seismic and Spectrum Multi–Client), shot in the Exmouth basin near the Pluto and Wheatstone Fields (Sams *et al.* 2016). Well ties were performed at nearby Bellatrix–1 and Urania–1 wells, possessing a complete set of elastic logs, and a total of four wells were used to help define facies models. Wavelets were estimated using the parametric constant-phase approach of Naeini, Gunning and White (2017). The facies model comprised a transition from marls above the upper Jurassic to sandstones below this, with shales, sands, siltstones, and limestones present over the section. In the following examples, the focus is on the clastic section, with the marl section above acting as a short lead in. The facies model was simplified to six classes: shale, limestone, two types of marl, brine–sandstone, and gas–sandstone, with the variability of elastic properties within these effective facies shown in Fig. 4. The regression curves underlying this figure are built relative to the (near) base-Cretaceous K horizon picked in the seismic and displayed as a thick line in the section displays of Fig. 5. These data are spread over a range of depths, so the true separability of facies at a given depth is greater than superficially apparent in Fig. 4. In terms of AVO response, the gas–sands are relatively bright, roughly doubling the expected reflection

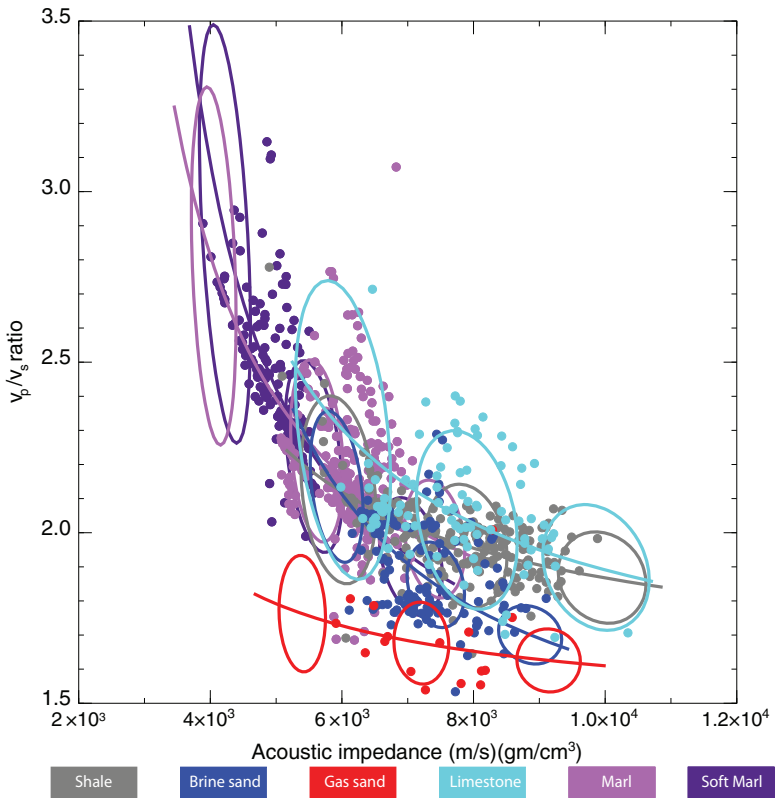


Figure 4 Variability of key elastic parameters in chosen simplified geological facies. The regression curves for each rock type follow from the statistical model in the main text. The ellipses show "one standard deviation" uncertainties in the marginal prior distribution of $\{\rho v_p, v_p/v_s\}$ at a shallow, middle, and deeper position in the inversion window.

coefficient against shale compared with brine–sand. These coefficients vary from about −0.04 to −0.18 from the shallower to deeper part of the inversion zone, a range over which the linearised (Fatti) reflectivity approximation is very good. Angle stacks at 12°, 22°, 32° and 42° degrees were used, with RMS–signal:RMS–noise ratios estimated from the well ties to be around 3:1. In the following tests, the input parameters remain the same unless otherwise indicated and were selected to optimise the results at the wells based on the homotopy algorithm. Improvements to all predictions are possible by further refinement of the prior information by interpretation of more stratigraphic intervals. The noise levels for each of the wavelets, as per estimates from well ties, are set to 20%, 25%, 30%, and 30%. Priors are assigned separately to the intervals above and below the K horizon (see Table 1) and are then smoothed across the K–horizon using an 11-sample moving-average filter. Figure 5(a) shows the facies prediction from the application of the EM algorithm with minimal beta values ($\beta_{xy} = 0.1$, $\beta_z = 0.05$), and Fig. 5(b) shows the same algorithm but with beta factors increased ($\beta_{xy} = 0.2$, $\beta_z = 0.1$) to enhance the 3D continuity of the facies predictions. In this case, the continuity is imposed with respect to the K horizon only. The degree of continuity is improved, and some blockiness has been removed. The result of homotopy is shown in Fig. 5(c). Further improvement in continuity can be observed and, in particular for very thin sands. There is an increase in the amount of gas-sand predicted. It is worth noting in particular the prediction of a thin gas-sand just below the K–horizon to the right of the Pluto well, which is the recently discovered Pyxis field. There are hints of gas at this location from the EM algorithm, but the homotopy algorithm is better suited to separating gas-sands from brine sands and shales when the starting model lies between the shale and brine sand. The IQP algorithm shows a slightly more extensive gas-sand at this location with gas predicted all the way up dip as the sand thins (Fig. 5(d)). The homotopy result shows a switch to brine up dip, which seems possible only if there is some barrier present but unresolved from the seismic. Interestingly, despite the differing morphological character of the inversion results produced by these three inversion schemes, the prediction success rate at the wells is relatively constant at ≈ 66%. This seems about commensurate with the noise levels and impedance separability in the cluster plots. By comparison, a basic simultaneous inversion with greedy posterior classification direct from the mixture distribution generates prediction success rates of ≈ 60%.

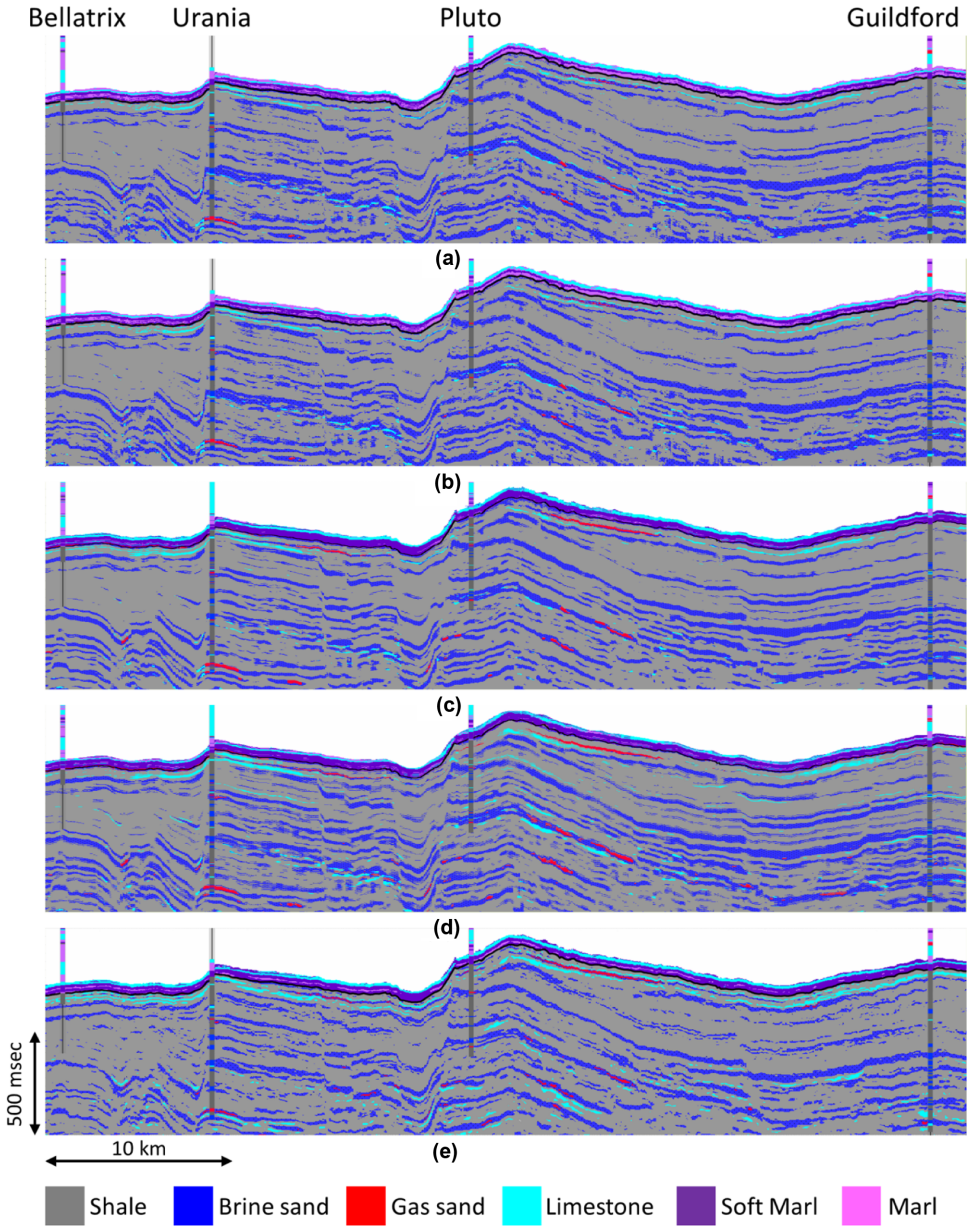


Figure 5 Facies predicted along an arbitrary line through the wells from various algorithms (a) EM with minimal betas, (b) EM with moderate betas, (c) homotopy, (d) IQP, and (e) model-based simultaneous inversion and polygon analysis.

Table 1 Prior probability model

Layer	Shale	Limestone	Soft Marl	Marl	Brine sand	Gas
above K	0	.4	.4	.2	0	0
below K	.59	.1	0	0	.3	.01

As described earlier, the EM algorithm begins with an M–step whose starting point is a weighted trend based on the assigned priors to each interval. For comparison, these initial weighted trends are used as the background model for a standard simultaneous inversion, whose elastic inversion results are classified deterministically through application of tailored polygons to the absolute AI and v_p/v_s results, as shown in Fig. 5(e); we denote this "standard simultaneous inversion" as SSI henceforth. Algorithmically, this is closely related to

terminating the EM algorithm after just one cycle, with additional parametric freedom allowed in the choice of polygonal boundaries so as to maximise the correct classification rate at calibration wells. The choice of polygon design is rather empirical due to the lower quality of elastic property prediction from the standard method. This can be observed in Fig. 6. Figure 6(a) shows a cross-plot of the AI versus v_p/v_s for the Urania well. The data are the measured log data re-sampled to the seismic sampling of 2 ms, and coloured by the log interpreted facies. Figure 6(b) shows the same plot using the elastic properties and facies predicted from the homotopy algorithm. The tight clustering along trends of the data points for each predicted facies is apparent and consistent with the input model. Figure 6(c) shows the elastic properties from the SSI and the facies from the polygon analysis. The separation of facies is apparently poor, and no gas sand is predicted despite the empirical nature of the polygons. There are two main reasons for this: the model-based inversion has poor constraints on eliminating noise, and in order to achieve reasonable elastic properties, the results are forced to sit close to a single, general trend; there is no means to bridge the missing frequency gap between the supplied trend and the lowest frequencies in the seismic (the correlation between bandpass-filtered (BPF) well logs and inversion results for the frequencies 1–2–4–8 Hz is 0.05 for v_p/v_s from SSI compared with 0.49 for homotopy). The lack of consistency between the standard simultaneous results and the expected distribution from the rock physics suggests that application of a rock physics template for interpretation of inversion results would not be appropriate without adjustment for these limitations. In the current case, no hydrocarbons are predicted as all known gas-sands lie closer to the brine–sand properties. The use of well log interpolation to provide the missing frequencies is unlikely to be successful in many, if not most, situations (Sams and Carter 2017). The empirical design of the interrelation polygons has been made to optimise the facies predictions at the wells. Comparison with the homotopy and IQP results suggests that the facies predictions for the thicker sands is reasonable, though noisier; the prediction of the thinner sands is very poor. Forcing the polygons to predict gas at the lower level in the Urania well would result in too much gas throughout the whole section.

The facies-based approach clearly has the capacity to predict within the frequency gap as indicated by the correlations. This is illustrated further through the inversion results at two wells: (i) Urania, one of the wells used to establish the rock physics trends, and (ii) Guildford, which is a completely blind test. The results are shown in the detailed Figs. 7 and 8.

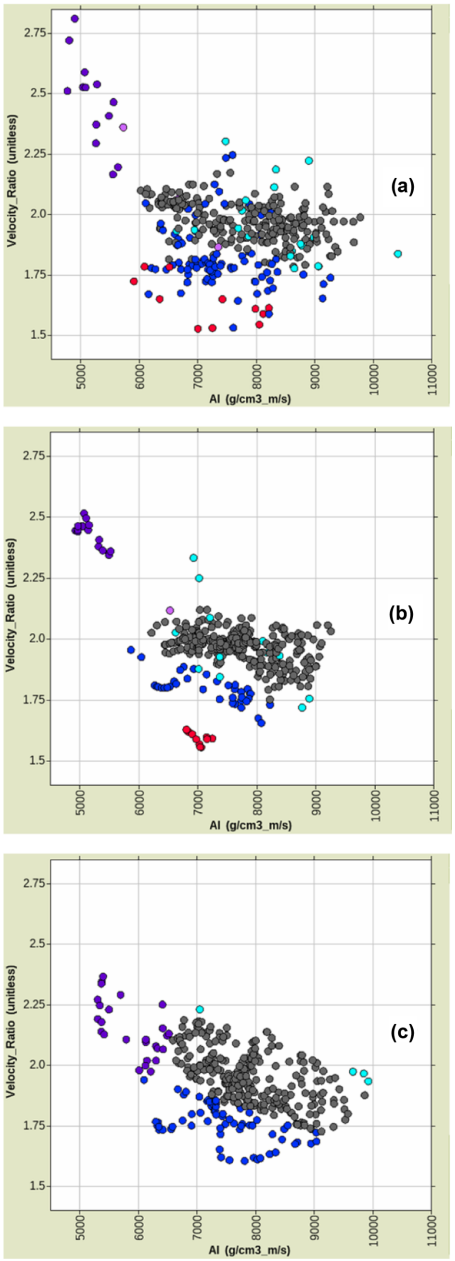


Figure 6 Acoustic impedance versus v_p/v_s for the Urania well sampled at 2-ms (a) measured data and log-interpreted facies (b) homotopy results and (c) model-based simultaneous inversion results and polygon analysis.

The data are presented showing the predicted facies versus the log facies in the first two columns. The AI, v_p/v_s, and density are shown in three columns each: first, the predicted (red) log versus the measured log (black); second, the predicted log (red) versus the trend for each of the facies where predicted (blue); and third, the BPF-predicted

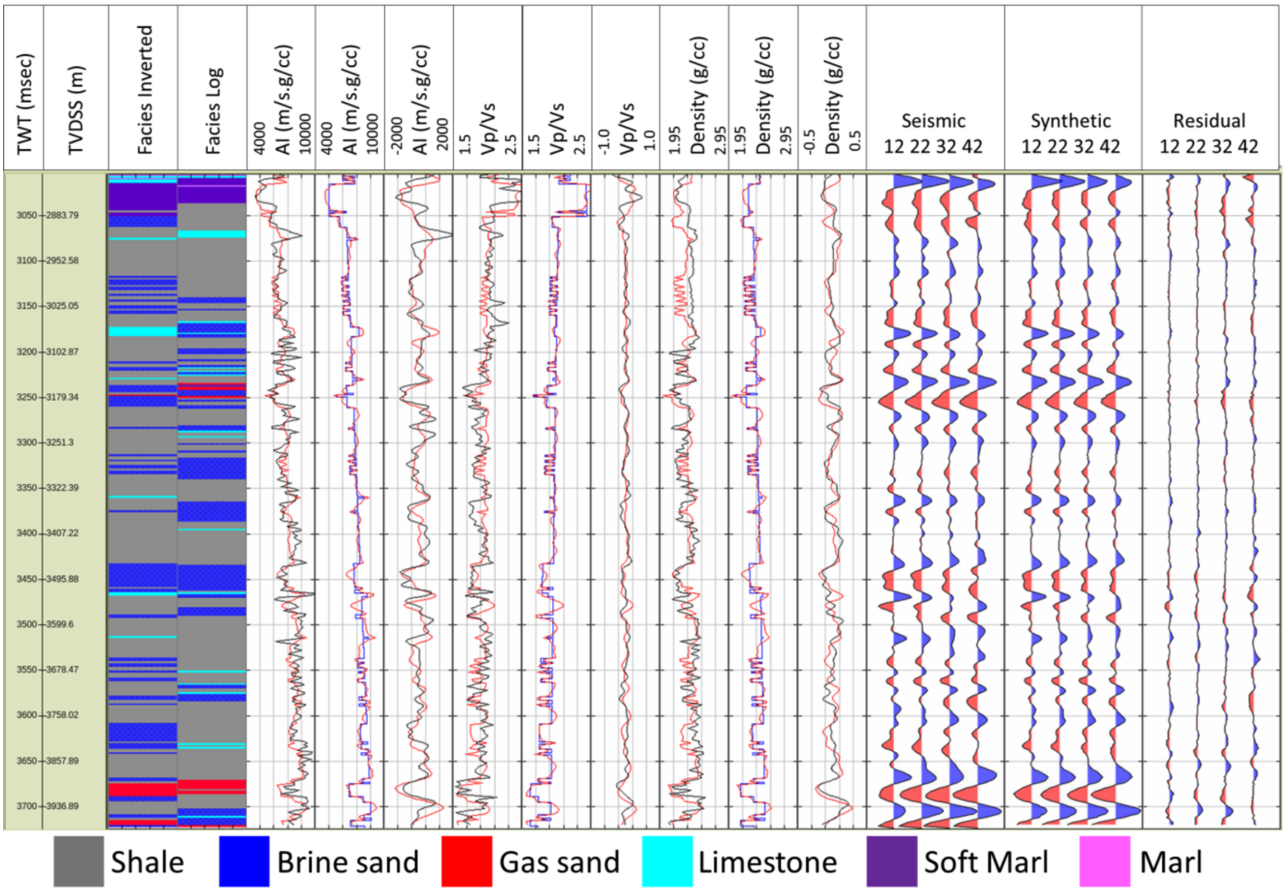


Figure 7 IQP inversion results at the calibration well Urania. Red logs are from the inversion, black logs are measured data, and blue logs show the input-per-facies trends.

(red) and measured (black) logs (a {1, 2, 70, 80}Hz trapezoidal filter). The final panels in these images are the seismic angle stacks, synthetics, and residuals from the inversion. A numerical comparison of the prediction accuracy of the joint categorical inversion versus SSI is shown in Table 2, pooling the Urania, Pluto, and Guildford wells. The cross-correlation coefficient of two quantities x_1 and x_2 is computed as $\mathrm{XCC} = \langle x_1 x_2 \rangle / (\bar{x}_{1,\mathrm{RMS}} \bar{x}_{2,\mathrm{RMS}})$, $\bar{x}_{\mathrm{RMS}} = \sqrt{\langle x^2 \rangle}$.

The presence of sharp contrasts at the boundaries between facies provides information within a very broad frequency range. The low frequency component below the seismic bandwidth is therefore a result of the predicted distribution of the facies and the trends within those facies. The results at the validation well (Guildford) suggest a number of limitations. First, the general AI trend is slightly low in this well particularly in the mid-section. It is possible that the choice of datum for the trends, chosen here as the K–horizon is not correct. The K–horizon is close to a significant unconformity, and this would place a question over whether it is appropriate to use for a datum. Alternatively, there may be a facies that is not captured by the calibration wells. Obviously, this does not invalidate the method but is more of a reflection of the application. Second, there is a mismatch in the dynamic range of the bandpass AI in the deeper interval. This might be a result of lateral or vertical variations in seismic scaling not accounted for and may explain the increased presence of limestone in the deeper section of the simultaneous inversion result (Fig. 5(e)). It is worth noting that the quality of the facies prediction at the wells is not necessarily a good quality control of the inversion. For example, there are two significant sands in the Urania well at 3300–3400 ms that are not predicted by the inversion at the well. Observation of the results on the cross-sections in Fig. 5 indicates that these two sands are present but are not predicted to intercept the well. This is thought to be due to data quality issues as the well lies close to a fault and the seismic imaging might not be optimal.

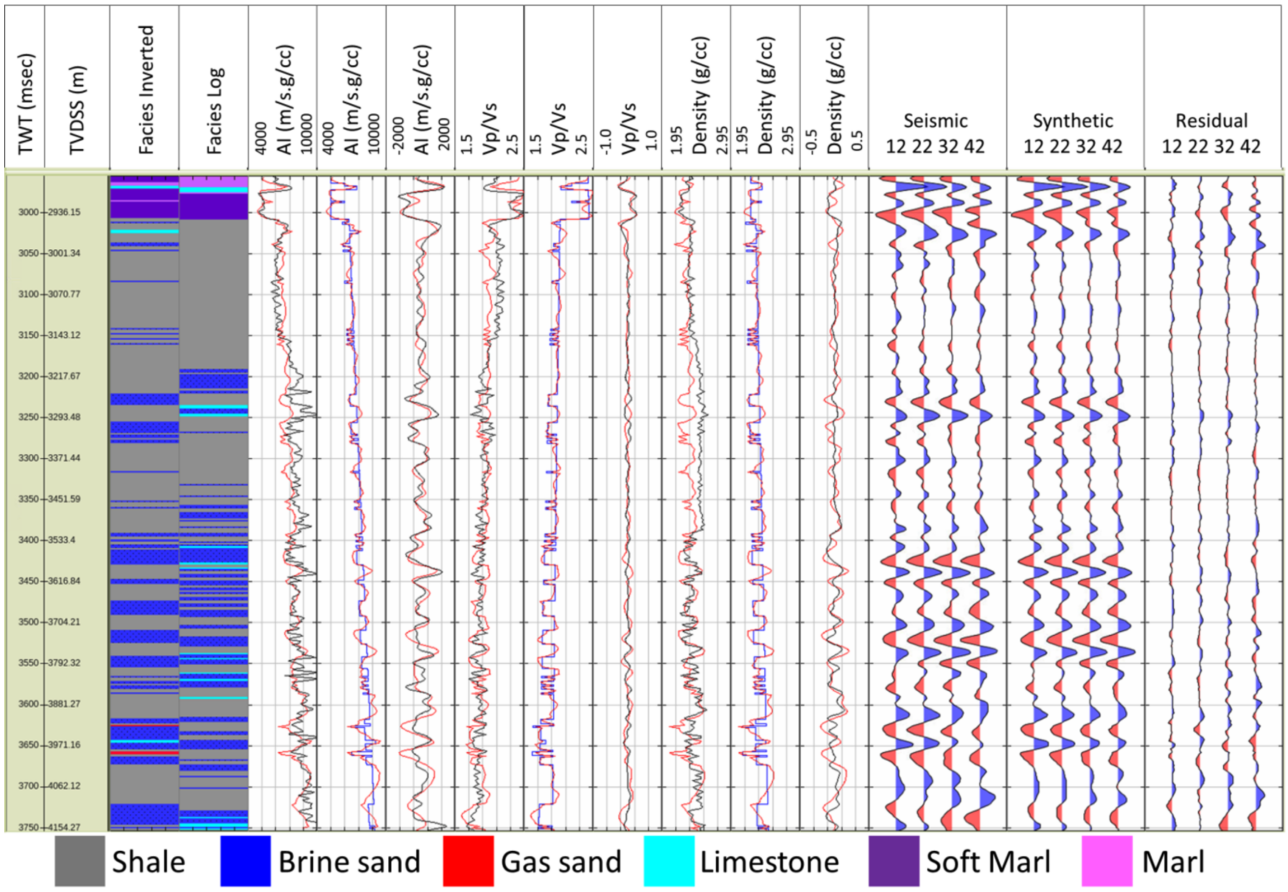


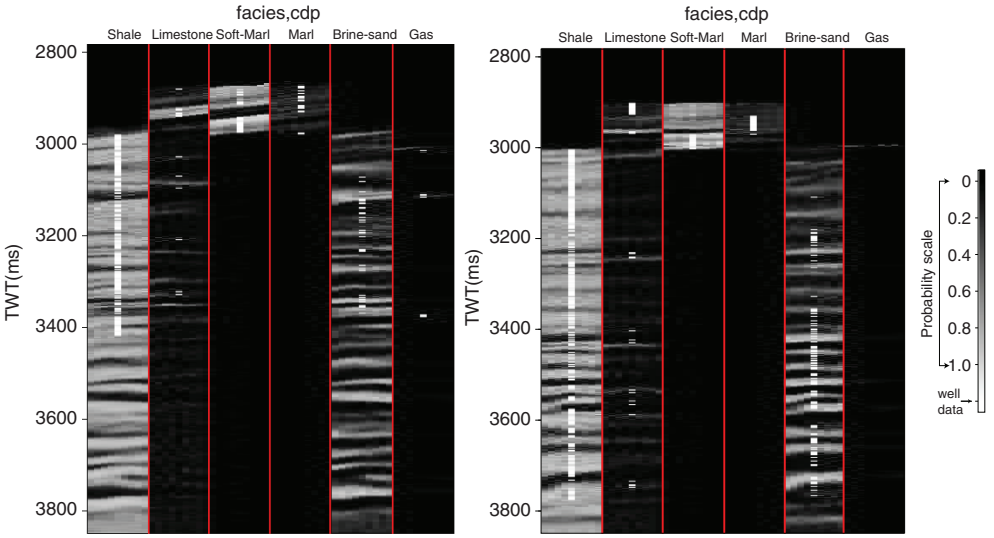
Figure 8 IQP inversion results at the validation well Guildford. Red logs are from the inversion, black logs are measured/synthesised data, and blue logs show the input-per-facies trends. Note that no measure shear sonic data were available for this well and that the shear velocity log was synthesised from a rock physics model.

Table 2 Prediction accuracy measures: XCC = correlation coefficient. AI(RMS) units are (m/s)(gm/cm^3), ρ in gm/cm^3. The joint inversion yields a 24% increase in XCC for v_p/v_s and more than two-fold for density

Quantity	AI	AI (BPF)	v_p/v_s	v_p/v_s (BPF)	ρ	ρ (BPF)
Joint inversion XCC	0.78	0.56	0.67	0.38	0.44	0.42
SSI XCC	0.79	0.56	0.54	0.19	0.21	0.16
Joint inversion RMS error	691.8	548.5	0.139	0.130	0.110	0.095
SSI RMS error	727.6	550.0	0.156	0.150	0.127	0.113

A final point of interest is the position of the well log data on the prediction uncertainties available from the annealing methods. Figure 9 shows an example of this at the two wells, Pluto and Guildford, where the well data are painted over the probability intensity images computed by sampling over a small region of nine traces centred on the well. With the exception of a small number of locations, virtually all log data land on significant probability locations. The exceptions may be variously due to edge effects, upscaling classification effects, or underestimation of systematic noise power associated with modelling or imaging issues: the latter can often be crudely compensated for by inflating the effective RMS variances in the covariance matrix $\mathbf{C}_d$.

This field study has highlighted a number of characteristics anticipated earlier. Errors in prediction of facies and properties can be due to (i) external or imaging noise, including wavelet calibrations and forward modelling approximations, for example, the missing sands along a fault; (ii) inadequacies in optimisation, for example the failure of the basic EM algorithm to predict known gas where the more expensive algorithms do; and (iii) mis-specification of the prior model, for example, the mismatch in trends at the validation Guildford well.



(a) Pluto **(b)** Guildford

Figure 9 Prediction at Pluto and Guildford wells (9-trace slice centred on wells), showing pointwise marginal facies probabilities (on scale [0, 1]) for the six facies (in order: shale, limestone, soft–marl, marl, brine sand, gas), with well data overlaid (white values).

The moderate success of facies prediction at the wells with a 66% prediction success attests to these errors, and it is significant that a contrived interpretation of standard model-based inversion results can achieve no better. The strong qualitative match seen by placing the wells on the sections might also suggest that further refinement of the well ties could markedly improve the quantitative match. In addition, in terms of quantitative prediction, the proposed method has overcome the limitations of supplying an inadequate prior low-frequency model to bridge the missing frequency gap. The result of this is that more stable estimates of both thin and thick layers have been made. In this example, the homotopy algorithm would be the most suitable given the very minor improvements achieved through simulated annealing, and further improvements in the predictions could be made through refinement of the input model.

CONCLUSIONS

The inversion described in this paper is designed expressly to promote consistency of petrophysical models through the interpretation and inversion workflows. The hierarchical model framework enables a tighter data integration and consistent interpretation workflow, and the explicit facies model makes integration with regional or analogue information easier. The hierarchical model framework provides a systematic and consistent way of addressing problems associated with low-frequency model building: the "background model" specification is a relatively simple assembly of per-facies compaction trends, global facies abundances declarations, and specification of continuity via a small number of parameters. We have implemented a comprehensive range of approaches to the optimisation problem, ranging from local EM algorithms, which perform well with reasonable initial guesses, globalised methods via homotopy, and strict global methods, which are starting-model independent. We expect that the hierarchical framework described here will serve well not only for standard AVO workflows but also for many other related problems where facies are the key latent variables connecting petrophysical variables and seismic responses.

REFERENCES

Aki K. and Richards P.G. 1980. *Quantitative Seismology: Theory and Methods*. San Francisco: W.H. Freeman and Co.

Bleistein N., Cohen J.K. and Stockwell J.W. 2001. *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*. Springer.

Buland A., Kolbjørnsen O., Hauge R., Skjæveland Ø. and Duffaut K. 2008. Bayesian lithology and fluid prediction from seismic prestack data. *Geophysics* **73**(3), C13–C21.

Dellinger J., Ross A., Meaux D., Brenders A., Gesoff G., Etgen J. et al. 2016. Wolfspar, an "FWI-friendly" ultralow-frequency marine seismic source. *Society of Exploration Geophysicists*, 4891–4895.

Eidsvik J., Avseth P., Omre H., Mukerji T. and Mavko G. 2004. Stochastic reservoir characterization using prestack seismic data. *Geophysics* **69**(4), 978–993.

Fatti J.L., Smith G.C., Vail P.J., Strauss P.J. and Levitt P.R. 1994. Detection of gas in sandstone reservoirs using AVO analysis; a 3-d seismic case history using the geostack technique. *Geophysics* **59**(9), 1362–1376.

Gunning J. and Glinsky M. 2006. WaveletExtractor: a Bayesian well–tie and wavelet extraction program. *Computers and Geosciences* **32**, 681–695.

Gunning J. and Kemper M. 2012. Some newer algorithms in joint categorical and continuous inversion problems around seismic data. In *Geostatistics Oslo 2012*, Vol. 17, pp. 263–273. Springer.

Huang K. 1987. *Statistical Mechanics*. Wiley.

Katayama K. and Narihisa H. 2001. Performance of simulated annealing-based heuristic for the unconstrained binary quadratic programming problem. *European Journal of Operational Research*, **134**(1), 103–119.

Kemper M. and Gunning J. 2014. Joint impedance and facies inversion–seismic inversion redefined. *First Break* **32**, 89–95.

Kolbjørnsen O., Buland A., Hauge R., Røe P., Jullum M., Metcalfe R.W. et al. 2016. Bayesian AVO inversion to rock properties using a local neighborhood in a spatial prior model. *The Leading Edge* **35**(5), 431–436.

Larsen A.L., Ulvmoen M., Omre H. and Buland A. 2006. Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model. *Geophysics* **71**(5), R69–R78.

Long A.S. and Reiser C. Ultra-low frequency seismic: benefits and solutions. Proceedings of the International Petroleum Technology Conference, IPTC, 2014.

Naeini E.Z., Gunning J. and White R. 2017. Well tie for broadband seismic data. *Geophysical Prospecting* **65**(2), 503–522.

Paige C.C. and Saunders M.A. 1982. LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software* **8**, 43–71.

Rimstad K. 2012. *Spatial mixture modeling based on latent random fields applied to seismic inversion*. PhD thesis, NTNU.

Sams M. and Carter D. 2017. Stuck between a rock and a reflection: a tutorial on low-frequency models for seismic inversion. *Interpretation* **5**(2), B17–B27.

Sams M., Westlake S., Thorp J. and Zadeh E. 2016. Willem 3D: reprocessed, inverted, revitalized. *The Leading Edge* **35**(1), 22–26.

Szeliski R., Zabih R., Scharstein D., Veksler O., Kolmogorov V., Agarwala A. et al. 2006. A comparative study of energy minimization methods for markov random fields. *Computer Vision–ECCV* **2006**, 16–29.

Ulvmoen M. and Omre H. 2010. Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: part 1 — methodology. *Geophysics* **75**(2), R21–R35.

Virieux J. and Operto S. 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics* **74**(6), WCC1–WCC26.

Wainwright M.J. and Jordan M.I. 2008. Graphical models, exponential families, and variational inference. *Foundation and Trends in Machine Learning* **1**, 1–305.

Walden A.T. and Hosken J.W.J. 1985. An investigation of the spectral properties of primary reflection coefficients. *Geophysical Prospecting* **33**(3), 400–435.

Winkler G. 2003. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction, 2nd edn*. Springer.