DeliveryMassager: a tool for propagating seismic inversion information into reservoir models

James Gunning [a,*], Michael E. Glinsky [b] Chris White [c]

[a] *CSIRO Division of Petroleum Resources, Bayview Ave, Clayton, Victoria, Australia, 3150, ph. +61 3 95458396, fax +61 3 95458380*

[b] *BHP Billiton Petroleum, 1360 Post Oak Boulevard Suite 150, Houston, Texas 77056, USA*

[c] *Dept. Petroleum Engineering, Louisiana State University, Baton Rouge, Louisiana, USA*

Abstract

We introduce a new open–source program for transforming inversion data from the open–source *Delivery* seismic inversion software to industry–standard cornerpoint grid formats suitable for reservoir modelling and flow simulations. The seismic inversion data produced by *Delivery* is an array of trace–local stochastic samples from a Bayesian posterior distribution of reservoir layer parameters, which contains complex correlations between layers boundaries, rock properties and fluid information, but no transverse correlations. This correlation structure is merged with lateral correlation requirements imposed by geological modelling inputs to the conversion process, thus producing cornerpoint grid models of the reservoir that honour seismic inversion, well data, and the desired lateral continuities. Realisations from the joint 'structural–stochastic', multi–property 3D correlated model can be drawn using a generalised p–field simulation algorithm. Distribution of volumetric quantities of commercial interest (e.g. net–hydrocarbon) can be directly generated. The software can produced both most–likely cornerpoint grids, and stochastic grids, which can be carried forward into production forecast studies for risking and uncertainty studies.

Key words: seismic, inversion, shared earth model, cornerpoint grid, Bayesian, stochastic, downscaling, geostatistics, open–source,

* Corresponding author.
Email address: James.Gunning@csiro.au (James Gunning).

1 Introduction

At the time an oil or gas field is being appraised or developed, the development of a reservoir model usually centers on the task of building computer models suitable for forward flow simulation. Prior to this, much of the work will have focused on data acquisition and interpretation, and the construction of models suitable for simple volumetric calculations or drilling decisions. In particular, much of the interpretative work will be based on surface seismic data, and this is routinely fed into various inversion routines which produce pointwise (trace–local) estimates of the properties of direct interest, such as surface positions, layer thicknesses, hydrocarbon content, net–to–gross etc.

In previous publications (Gunning and Glinsky, 2004, 2005; Gunning, 2003), we have introduced an open source tool *Delivery* that enables users to perform a fully probabilistic seismic inversion for a layer–based model of the reservoir. This is a trace–based inversion, so it produces an ensemble of realisations of the relevant reservoir parameters at each point in the imaged seismic grid over a field. The inversion data produced by this program is suitable for answering the simple kind of questions mentioned above, such as pointwise histograms of layer thickness, maps of hydrocarbon probability etc, but is not directly suitable for flow calculations per se.

For the calculation of volumetric (not pointwise) uncertainties, and the task of flow simulation, is is necessary to carry these inversion calculations over to grid formats that are more directly useful in 3D volumetric calculations and flow calculations. Various types of 3D grids are in common use, but perhaps the most ubiquitous is the *cornerpoint* grid (Ponting, 1989), which is used in the commercially dominant ECLIPSE family of flow simulators [1]. Moreover, the effect of intertrace correlations in the seismic data (which is deliberately neglected in *Delivery*) can be approximately modelled in this remapping calculation. The aim is to produce 3D models that capture both transverse correlations known from either well data or analogues, and the vertical inter–layer and inter–property correlations that seismic inversion can reveal. The joint process of remapping and merging of correlations is one we have dubbed 'massaging', hence the software name.

Since the inversion models are probabilistic, it is natural for the remapped or 'massaged' models to inherit this probabilistic character. Objects of interest will then naturally be the 'most likely' massaged model, as well as a suite of 'realisation' models, which enable stochastic forward flow simulations to be performed for risking or uncertainty evaluation purposes. Volumetric statistics of interest can of course be calculated on the fly as well.

[1] See Schlumberger ECLIPSE website, `http://www.slb.com/content/services/software/reseng/eclipse_simulators/index.asp`

The remainder of this paper is organised as follows. In section 2 we define the conceptual problem in closer detail . The smoothing and geostatistical simulation, containing most of the notation and requisite algorithms, is covered in sections 3 and 4. A brief description of the software follows in section 5. Some examples, both synthetic and real–life, are presented in section 6, and we offer some conclusions in section 7. Much of the tedious but necessary detail is relegated to Appendices, which will be helpful to readers wishing to obtain the software and run examples of their own.

We do not hope to offer exhaustive practical details and illustrations from examples in this paper. It is intended to serve chiefly as a technical statement of the generic motivation, algorithms and code operation for later reference, though we hope the examples of section 6 are a sufficient *hors d'oeuvre* for readers. We expect to publish several papers illustrating field examples and discussing practical issues at more length in the near future, along the lines of Glinsky et al. (2005).

2 Problem Definition

The trace–based Bayesian inversion model implemented by *Delivery* produces a ensemble of realisations from the posterior distribution of the model parameters at each common midpoint (CMP) gather, or trace, of the imaged seismic data. At each trace, the inversion model is quasi 1d, with a sequence of layers parametrised by times describing the (local) geometry, and each layer is characterised by a laminated mixture of permeable and impermeable rocks, with rock velocities, density, porosity and fluid content as additional model parameters. The depths of each layer are computed from relative traveltimes and velocities, hung from a nominated reference layer and supplied reference depth. For the purposes of reservoir simulation, the model parameters of interest are the layer depth `d`, the `thickness`, net–to–gross `NG` (or N_G), fluid content `net-hydrocarbon`, and `porosity` (to name the interesting quantities accessible from *Delivery* inversions in `typewriter` font). The Monte Carlo ensemble produced by the inversion encapsulates the coupling or correlations between these properties which is demanded by consistency with the seismic data and the prior model.

The imaged traces are typically spaced anywhere from 15m to 200m apart, in a regular array, and the inversion does not model the coupling that may occur between model parameters at different traces, chiefly in order to make the inversion problem tractable (a more detailed discussion of these issues can be found in Gunning and Glinsky (2004)). Very strong lateral correlations are induced in the mean (or most–likely) posterior models by the prior and seismic data, but the overall distribution describing the model fluctuations formed

by a naive resampling from the *Delivery* outputs is a product of trace–local distributions, and thus contains no lateral correlation. A necessary and strong qualification to this statement is that any spatial interpolations of the *Delivery* outputs will necessarily induce correlations via the interpolation algorithms, and indeed we often recommend running the inversion on relatively coarse spacings and using interpolation to smooth; this has the merit of reducing the inversion run–time considerably.

Geologists are accustomed to thinking of transverse correlations in terms of 'characteristic' body sizes and depositional directions. Large scale body shapes will usually be explicitly visible in the seismic, and thus will propagate into the mean posterior models. The residual fluctuations about these means, for many environments, are most simply characterised by two–point statistics and a distance metric which reflects depositional directions, a construct which is familiar to geologists as the conventional semivariogram. Continuity of surfaces is broken at fault locations, though many internal properties may be preserved across the fault. Faulting is usually modelled explicitly in reservoir modelling packages like *PETREL*, and these faults will be embedded in the cornerpoint grids we use as the receptacle for the massaging process. Additionally, hard data for various properties will be available at well locations that have been logged.

It is clear that the combined spatio–multi–property posterior distribution we are able to build, having neglected inter–trace correlations in the inversion, will be a *constructed* entity. The overall dimensions of the problem, for a large model, will be 10^6 parameters or more, so perhaps the only computationally feasible way to proceed is to merge the first and second order statistics from the inversion with the second–order statistics implicit in the variograms in a pseudo–Gaussian framework. In terms of second order statistics, it is natural to think of the correlation matrix of the overall structure as a block matrix, with the inversion ensemble furnishing the blocks for the inter–property correlations at each location, and the spatial variogram defining how off–diagonal blocks are coupled. The overall correlation matrix is then a direct (or Kronecker) product of correlation matrices, and the natural and efficient approximation to sampling from the Gaussian distribution attached to this correlation is a generalised p–field algorithm. The great advantage of this approach is that is is only necessary to form approximations to the first and second order statistics (mean and covariance), and then simulation is direct.

Some approximations are required in forming the statistics from the seismic inversion ensemble. All of the quantities of interest are non–negative, and will be approximately Gaussian if the forward seismic model is reasonably linear over the support of the prior (e.g. the thickness of a reservoir layer that is well above seismic resolution). But there are regions were approximate linearity does not hold, and model parameters are also often truncated at one end, so

their posterior univariate distribution is a mixture of a spike at the truncation and a continuous tail (see especially the first example in section 6, and Fig. 4). The median value is then a much better statistics for defining a most–likely model, since the mean of a skewed or truncated distribution is less likely than the median. Especially in the case of thicknesses, use of the mean statistic can often generate thin layers (sub seismic–resolution) over extensive areas of space. Obviously the multi–Gaussian model is a heavy approximation in this regime, so the means in the distribution have to be adjusted in this way to eliminate the worst biases. Output samples drawn from the 'massaged' distribution are obviously truncated in the same fashion. Some further interesting aspects of the thin–layers problem are discussed in Appendix 4.

2.1 Gridding considerations

The input–grid formed by the x, y locations of the sequence of seismic traces is usually regularly sampled transversely, but the inversion region may have been confined to some polygon of interest. This grid may or may not strictly contain the extremity of the desired output grid. See Fig 1. Typically from the inversion, we have available the distribution of (among others) layer–thickness, N_G, layer–top depth, and net–hydrocarbon for the sequence of model layers at each trace x, y location.

[Fig. 1 about here.]

Conversely, reservoir geological models are usually built with uneven spatial sampling, and often with less transverse resolution than that available in the seismic inversion. Various kinds of grid geometry are possible (Deutsch, 2002), but we have chosen the the widely used *cornerpoint* grid (see Fig. 2) format for exporting inversion information to reservoir modelling packages. These grids are used by the ECLIPSE reservoir simulator, so we use the adjectives 'cornerpoint' and 'Eclipse' as loosely interchangeable.

Cornerpoint grids specify lines for the 'vertical' corners of each column of gridblocks, and a set of 8 depth (z) points which define the top and bottom face of any particular gridblock. (The faces are not strictly planar, but flow calculations make suitable projections so as the conserve mass and represent flux correctly). Blocks are indexed by an i, j, k triple, and usually many blocks are tagged as 'inactive' if they represent an uninteresting region of space. The file formats consist of chunks denoting the grid size, local origin, block–corner lines (COORD), corner–depths (ZCORN), and block–centred properties, such as 'active' flags, and segment labels.

For the massaging calculation, we form, for each layer, a 2d output grid comprising the (x, y) projections of the midpoints of the edges of each corner point

block (see little circles in Fig. 2). Each such point has an associated 'segment' or 'fault–block' tag, and points along the fault plane are repeated according to the number of intersecting segments.

A typical plan view of the cornerpoint geometry is shown in Fig. 1.b). In the algorithms described below, all properties are interpolated or simulated at the corner points. Block centred properties like porosity are later formed by suitable 4–corner averages (e.g. thickness–weighted averages, or more elaborate weights based on integration of the Jacobian (Peaceman, 1996)).

[Fig. 2 about here.]

3 Smoothing and simulation algorithms

The overall simulation process is perhaps easier to describe in words than mathematically. Roughly speaking, we spatially smooth the means and covariances of the multi–property *Delivery* output ensembles to produce a smooth trend map for each property and a smooth between–properties trend covariance. The trend map may be allowed to have discontinuities at faults for certain properties. The property trends are then deformed by a kriging adjustment from well observations to produce a trend that passes through all hard observations. The kriging–variance maps produced in the kriging calculation are normalised and stored for each property, for later use. The maps may be truncated or clipped depending on consistency requirements for the properties (e.g. net–sand will be clipped to less than total thickness). The trend maps then constitute the most–likely maps of properties. The uncertainty maps are defined by the diagonal entries of the smooth trend covariance multiplied by the kriging variance maps. The square root of the latter map is then a local 'standard deviation' for each property, which honours the inversion uncertainties and the well data.

To generate realisations, a set of normalised, spatially correlated random fields (p–fields) are simulated for each property and layer. At each spatial location these fields are then mixed in a linear combination described by the Cholesky factor of the smoothed between–properties trend covariance. The final set of fields are then scaled by the normalised kriging–variance maps and added to the trend maps to produce a series of realisations.

The actual program execution follows the steps of section 3.2 and 3.3 reasonably literally, but we will need some notation first.

3.1 Notation

In general there exists a seismic grid on which the inversion is run, which we call the input grid (G_I). The properties are to be generated on a different grid, called the output grid (G_O). This is a treated as a sequence of 2D grids for each layer, $G_{O,l}$. The index l pertains to layers, p to properties, j to nodes on G_I or a particular $G_{O,l}$ (as arranged in convenient ordering - say a raster scan). The set of hard well observations (O) may be suffixed l or p with implied constraint to layer l or property p. A vector m of properties of interest may be suffixed $m_{(l,p)}$, which denotes a generic unrolled index of property p on layer l. Segment labels for node j are denoted S_j.

The local neighbours of node j are denoted by $j' \sim j$, or $j' \in \partial j$. Neighbours are defined by a Euclidean distance metric confined to the same layer, with azimuth angle and principal ranges inherited from a conventional variogram specified by the user. Neighbours used in the trend smoothing are defined by the trend–smoothing variogram, whereas observation–kriging and simulations are associated with a separate, layer–specific 'interpolation' variogram. Since the grids are large and irregular in general, a kd–tree algorithm is used for efficient nearest neighbour searching (Skiena, 1997). Nearest neighbour searches in the sequential simulation algorithms must be confined to previously visited nodes, and this is achieved by a naive dynamic kd–tree implementation.

3.2 Initialisations

Re–ordering of properties

Internally, the property vectors are re–ordered to the following sequence: {*depth*, *thickness*, *net–sand*, other properties. . . ,N_G}, if all the italicised properties are available. This enables truncation rules to be sensibly applied from known quantities later in the calculation.

Input grid segmentation

Fault–sensitive smoothing of the *Delivery* statistics requires a segment label to be attached to each input grid point. In corner point style grids, segments labels are associated with block centers. The input grid can then naturally inherit the segment label associated with the corner–point grid block in which it falls, as computed for the user–specified reference layer.

Kriging

Kriging is used for both integrating well observations and interpolating from the input to the output grid. Interpolation kriging calculations are performed with a fixed number of nearest neighbours, typically 8. For well observation kriging, all hard–data values are used, and for the sequential simulation routine described later, around 25 neighbours are used. Because the output grid may contain coincident points, some rank–deficient kriging systems can arise. Robust solution of these is performed using a adaptation of the Schnabel–Eskow modified Cholesky decomposition (Schnabel and Eskow, 1999).

Well Observations

For each layer l, for the observation q at location r_q, we construct and solve the ordinary kriging (OK) system for $\{w_{j'}, \gamma\}$,

$$\sum_{j'' \in \partial q} C(r_{j'}, r_{j''}) w_{j''} + \gamma = C(r_{j'}, r_q), \quad j' \in \partial q \subset G_O$$
$$\sum_{j'' \in \partial q} w_{j''} = 1 \tag{1}$$

where the covariance used is that proper to interpolation for the layer, with unitised sill. We store a data structure for the kriging weights and neighbours, viz. $\{w^{(\mathrm{O})}_{(l,q);j'}, j' \in \partial q \subset G_O\}$, writing all the bookkeeping.

Grid interpolation

For each layer l, at the output grid location $j \in G_O$, we construct and solve the OK system for interpolation from the input grid to the output grid

$$\sum_{j'' \in \partial j} C(r_{j'}, r_{j''}) w_{j''} + \gamma = C(r_{j'}, r_j), \quad j' \in \partial j \subset G_I$$
$$\sum_{j'' \in \partial j} w_{j''} = 1, \tag{2}$$

using, again, the interpolation covariance with unit sill. We save a data structure for the kriging weights and neighbours $\{w^{(G_I)}_{(l,j);j'}, j' \in \partial j \subset G_I\}$.

Smoothing

If $\bar{m}^{(G_I)}_{l,p,j}$ is the p50 statistic from the *Delivery* inversion for property p, layer l, location $l \in G_I$, we smooth this onto the output grid using a moving average

filter whose weights are based on the covariance $C_{\mathrm{sm}}()$ specified by the trend–smoothing variogram. Specifically,

$$\bar{m}^{(G_O)}_{l,p,j} = \frac{1}{\mathcal{N}} \sum_{j' \sim j} W_{l,p,j'} \bar{m}^{(G_I)}_{l,p,j'} \tag{3}$$

where the weights W are defined by

$$W_{l.p.j'} = C_{\mathrm{sm}}(r_{l,j'}, r_{l,j}) \times \begin{cases} 1 & p \text{ continuous across faults} \\ \delta(S_{j'} = S_j) & \text{otherwise} \end{cases} \tag{4}$$

and the normalisation constant $\mathcal{N} \equiv \sum_{j' \sim j} W_{l.p.j'}$. A larger number of nearest neighbours are used here, typically 50 or so.

For the covariances, the full covariance matrix (coupling all layers and properties) is smoothed back onto the input grid (for reasons which will become apparent later), without regard to segmentation. The segmentation is ignored, as uncertainties can be expected to be continuous across faults. In accord with the usual rules for the sum of independent random processes (and also, conveniently, to ensure positive definiteness), the smoothed covariance is

$$\bar{C}^{(G_I)}_{l,p,l',p',j} = \frac{1}{\mathcal{N}'} \sum_{j' \sim j} W^2_{l,p,j'} C^{(G_I)}_{l,p,l',p',j'} \tag{5}$$

here, $C^{(G_I)}_{l,p,l',p',j}$ is the *Delivery* covariance statistic for property p, layer l with property p', layer l', evaluated at location $j \in G_I$. The normalisation constant $\mathcal{N}'$ is the sum of the weights $W^2_{l.p.j'}$.

It is also useful to interpolate, for later use, the smoothed trend surface at the well observation points.

$$\bar{m}^{(\mathrm{O})}_{l,p,q} = \sum_{j' \sim q} w^{(\mathrm{O})}_{(l,q);j'} \bar{m}^{(G_O)}_{l,p,j'} \tag{6}$$

FFT based methods for smoothing are awkward to use for this problem, on account of the segmentation and the irregular grid.

Trend adjustment from well observations

A final initialisation calculation is the adjustment of the trend surface via simple kriging (SK) so it passes through the well data. For each layer and

property, at each location $j \in G_O$ we compute the residual trend adjustment

$$\Delta\bar{m}_{l,p,j} = \sum_{q \in \mathrm{O}_{l,p}} w_q (m^{(\mathrm{O})}_{l,p,q} - \bar{m}^{(\mathrm{O})}_{l,p,q}) \quad j \in G_O \tag{7}$$

where $m^{(\mathrm{O})}_{l,p,q}$ is the qth observation of property p on layer l. The simple kriging weights w_q are the solution of the set of equations

$$\sum_{q' \in \mathrm{O}_{l,p}} C(r_q, r_{q'}) w_{q'} = C(r_q, r_j), \quad q \in \mathrm{O}_{l,p}, \quad j \in G_O. \tag{8}$$

Again, the covariance used in this SK step is normalised to unit sill. The new trend surface is then defined to be

$$\bar{m}'_{l,p,j} = \bar{m}_{l,p,j} + \Delta\bar{m}_{l,p,j} \quad j \in G_O \tag{9}$$

We store also the kriging variance

$$\sigma^2_{l,p,j} = 1 - \sum_{q \in \mathrm{O}_{l,p}} w_q C(r_q, r_j) \quad j \in G_O \tag{10}$$

which is used in the subsequent p–field simulation.

A final step in the trend adjustment is application of a set of truncation rules. The loop over properties p occurs innermost in the calculation, and the internal ordering of properties described earlier enables successive application of these rules: i) $\mathit{net\text{–}sand} = \min(\mathit{net\text{–}sand}, \mathit{thickness})$, ii) $N_G = \mathit{net\text{–}sand}/\mathit{thickness}$, iii) $p = \max(p, 0)$, iv) if p is normalised, $p = \min(p, 1)$.

3.3 Realisations

Realisations are generated using a generalised p–field technique (Deutsch and Journal, 1998), which requires a set of unconditional correlated realisations on the output grid. There are various ways to do this, but the sequential simulation technique is most easily adapted to the unstructured grid.

Some notation and apparatus is necessary. For each layer l, we construct a pseudo–multigrid path P_l, which is a visiting sequence for all the nodes in the layer. The sequence is pseudo–multigrid in the sense that the grid nodes are visited in a sequence derived from a breadth–first traversal of a binary tree representation of the nodes. This ensures that widely spaced points are visited

early in the path. For the visited node $\hat{j}$, we denote by $\partial\hat{j}$ the nearest neighbouring points of $\hat{j}$ that have already been visited, up to some maximum of N_n neighbours, and with the notion of distance derived from the layer variogram. We will generate and store $\alpha = 1 \dots N_R$ realisations at each grid point during the path traversal.

The conditional distribution for the p–field $\xi_{l,p,\hat{j},\alpha}$ is $N(\mu_{l,p,\hat{j},\alpha}, \sigma^2_{l,p,\hat{j}})$, where the conditional mean is

$$\mu_{l,p,\hat{j},\alpha} = \sum_{\hat{j}' \in \partial\hat{j}} w^{(\mathrm{SK})}_{\hat{j}'} \xi_{l,p,\hat{j}',\alpha}, \tag{11}$$

the conditional variance (geometry dependent only) is

$$\sigma^2_{l,p,\hat{j}} = 1 - \sum_{\hat{j}' \in \partial\hat{j}} w^{(\mathrm{SK})}_{\hat{j}'} C(\hat{j}', \hat{j}), \tag{12}$$

and the simple kriging weights in these last two relations are the solution of the SK system

$$\sum_{\hat{j}'' \in \partial\hat{j}} C(r_{\hat{j}'}, r_{\hat{j}''}).w^{(\mathrm{SK})}_{\hat{j}''} = C(r_{\hat{j}'}, r_{\hat{j}}) \quad \hat{j}' \in \partial\hat{j} \subset G_O. \tag{13}$$

Again, the variogram is normalised, the the p–fields ξ have univariate distribution $N(0,1)$. The fields $\xi_{l,p,\hat{j},\alpha}$ are stored by a fully nested loop on l, $\hat{j}$, p and α.

These p–fields are now unconditional correlated fields that contain the necessary spatial correlation on the output grid, but honour neither the inter–property/ inter–layer correlations from the seismic inversion nor the hard well data.

To introduce the inter–property/layer correlations, a local interpolated covariance is computed at each node $j \in G_{O,l}$ (in a conventional loop over the output grid l) as

$$\bar{C}^{(G_{O,l})}_{l,p,l',p',j} = \frac{1}{\mathcal{N}} \sum_{j \in \partial j} (w^{(G_I)}_{l,j,j'})^2 \bar{C}^{(G_I)}_{l,p,l',p',j'} \quad j' \in \partial j \subset G_I, \tag{14}$$

using the saved OK/interpolation weights from (2). The normalisation $\mathcal{N}$ is again defined as the sum of the squared OK/interpolation weights: squaring is used again to ensure positive definiteness. Define by $\bar{L}^{(G_{O,l})}_j$ the conventional (left) Cholesky factor of $\{\bar{C}^{(G_{O,l})}_{l,p,l',p',j}\}$ where we unroll the indices l, p in the

usual way. Define also the diagonal scaling matrix $S_j^{(\mathrm{WO})} = \mathrm{diag}\{\sigma_{l,p,j}\}$ using equation (10), with the same indexing. Realisations are them computed on the fly by 'mixing' the correlated p–fields and adding back the trend:

$$m^{(\mathrm{R})}_{l,p,j,\alpha} = S^{(\mathrm{WO})}_{l,p,j} \cdot \sum_{l',p'} \bar{L}^{(G_{O,l})}_{l,p,l',p',j} \xi_{l',p',j,\alpha} + \bar{m}'_{l,p,j} \tag{15}$$

This is the p–field mixing equation which imposes the inter property/layer correlations and well constraints on the correlated fields to produce the realisations. A final step consists in the imposing of the property truncation rules on $m^{(\mathrm{R})}_{l,p,j,\alpha}$, as per the mean trend calculation.

Under this construction, the covariance of the residuals $\Delta m^{(\mathrm{R})}_{l,p,j,\alpha}$ (the first term in equation (15)) can be shown to have these reasonable special cases:

$$\langle m^{(\mathrm{R})}_{l,p,j} m^{(\mathrm{R})}_{l',p',j'} \rangle = \begin{cases} S^{(\mathrm{WO})}_{l,p,j} S^{(\mathrm{WO})}_{l',p',j} \bar{C}^{(G_O)}_{l,p,l',p',j} & \text{Same location, } j = j' \\ S^{(\mathrm{WO})}_{l,p,j} S^{(\mathrm{WO})}_{l',p',j'} \bar{C}^{(G_O)}_{l,p,l',p',j'} C(r_j, r'_j) & \text{Smoothed inversion statistics match at } j, j' \end{cases} \tag{16}$$

4 Downscaling, or 'decoration' algorithms

In most realistic applications, fluid flow will be sensitive to the manner in which impermeable material (usually clay: we will use the placeholding name 'shale' hereon) is spatially dispersed within the 'meso–scale' reservoir layers used for the inversion. Capturing this effect will then require subdivision of the vertical gridding and suitable categorical simulation of the shales within a meso–scale layer.

This categorical simulation must be consistent with the net–to–gross N_G obtained from the seismic inversion, or, equivalently, a 'massaged' realisation. The inversion forward model uses an effective–medium approximation based on a separation of length scales between the vertical spatial scales characterising the shale distribution and the seismic wavelength. In this regime, the effect of the shale on the seismic response is then captured by an effective macroscopic parameter, the layer net–to–gross (N_G), via the Backus average. The model also assumes a laminated distribution of shale, which is a respectable assumption for reservoirs where internal shales are gently dipping.

A variety of categorical simulation techniques are in common use in petroleum geostatistics. Perhaps the best understood algorithms for binary simulation

are truncated Gaussian fields (le Loc'h and Galli, 1996), and we have chosen to adapt this method to the 'decoration' problem on account of the efficiency of simulation of the underlying continuous field. Users are expected to furnish a 3D variogram describing the spatial continuity of the underlying Gaussian field, which can be estimated in consultation with a geologist. This variogram is embedded in the (normalised) covariance function $C_{\mathrm{TG}}()$. The algorithm we describe is somewhat heuristic, but very efficient, and strikes a good compromise between the connectivity embedded in the variogram and the coarse–scale constraints. It may be loosely described as a greedy, optimising, sequential truncated Gaussian simulation.

Rigorous sampling of high dimensional categorical spatial distributions with tight likelihoods is notoriously difficult (Winkler, 2003). A formulation in terms of discrete Markov Random Fields (MRFs) would have been more satisfactory in terms of incorporating the net–to–gross likelihood constraint, but explicit control of correlation length scales is much more difficult with MRFs. We have also resisted the temptation to try and provide a 'most–likely' categorical map, since this object is highly (combinatorially) non–unique, and any of the most–likely models is very non–representative. An analogy with the celebrated Ising model of statistical mechanics is helpful (Winkler, 2003), since this represents by far the best understood MRF model in the literature. If we map Ising ± 1 spin states to rock categories, the temperature of the Ising model determines the correlation length of the realisations, but at any temperature, the 'most–likely' model is all one category or spin in an unconstrained model. For the case of smooth N_G constraints, the most likely model(s) will be a layered two–zone partitioning, with the zone boundary of minimal length: this will clearly violate the homogenisation assumptions used in the inversion.

The algorithm runs roughly as follows, conditional on some known realisation of N_G on the coarser grid:

- Vertically subdivide the layer l into a more finely meshed grid G_{OD}, with whatever style of gridding is appropriate, preserving the lateral spacing. Suppose there are N_Z fine gridblocks in a subgrid column.
- Form a multigrid traversal path for G_{OD}, comprising an outer 2D multigrid path P_{col} for 'column' visiting, and a inner 1D path P_z for visiting gridblocks within the chosen column.
- At each column in the outer path, loop over blocks $\hat{j} \in P_z$ in the column, forming and solving the set of SK equation for each block, conditioned on the nearest known blocks from previously visited columns and previously visited blocks in the current column. The SK systems use the covariance $C_{\mathrm{TG}}()$. We store the weights $w^{(\mathrm{SK})}_{(\hat{j});\hat{j}'}, \hat{j}' \in \partial\hat{j}$ for all N_n nearest neighbours $\hat{j}'$ of all blocks $\hat{j}$ in the column.
- From the coarse grid simulation or map of N_G at this column, use the local threshold $y = G^{-1}(N_G)$ to form the truncated Gaussian field, where $G^{-1}()$ is

the inverse Normal CDF. If the underlying coarse grid simulation is spatially smooth (expected), the threshold y will be as well.

- Form a set of, say, $\alpha' = 1 \ldots 50$ realisations η of the underlying Gaussian field for the current column, by repeatedly running the sequential sampler over the column path, using the pre–stored weights and the conditional cdf's $N(\mu_{\hat{j},\alpha'}, \sigma^2_{\hat{j}})$, with conditional mean

$$\mu_{\hat{j},\alpha'} = \sum_{\hat{j}' \in \partial\hat{j}} w^{(\mathrm{SK})}_{\hat{j}'} \eta_{\hat{j}',\alpha'}, \tag{17}$$

and conditional variance

$$\sigma^2_{\hat{j}} = 1 - \sum_{\hat{j}' \in \partial\hat{j}} w^{(\mathrm{SK})}_{\hat{j}'} C(\hat{j}', \hat{j}). \tag{18}$$

This is very fast, requiring only $O(N_z N_n)$ flops per simulation. We greedily accept the simulation whose associated truncated field $\eta^{\mathrm{TG}}_{\hat{j},\alpha'} = H(\eta_{\hat{j},\alpha'} - y)$ best matches the column N_G (a columnwise sum), and proceed to the next column in the 2D multigrid column path.

The complexity of this algorithm is only a small multiple of the workload of a conventional sequential simulation, typically $O(10^3)$ flops per node.

The field–case example of section 6.2 illustrates some realisations drawn using this algorithm, especially Fig. 10.

5 The Software

The software is written in ANSI C, and distributed along with the open–source *Delivery* and *WaveletExtractor* codes (Gunning and Glinsky, 2004, 2005; Gunning, 2003). Library dependencies are all open source. It contains an *ECLIO* library for handling ECLIPSE–style cornerpoint grids, and links to the high performance ATLAS library for the intensive linear algebra work (Whaley et al., 2001). The supplied kd–tree range–search library is based on *Ranger* from Stonybrook (Skiena, 1997). For large models, the grid smoothing is intensive as well as the sequential simulation, and RAM sizes over 1Gb may be required.

Compilation instructions are provided in the README file at the top level of the source tree. Installation of the third party ATLAS, glib, and libxml libraries is straightforward on any variety of Unix/Linux, but should also be possible on other architectures with an ANSI compiler. The binaries supplied will be valid for current intel Linux architectures.

6 Examples

6.1 Simple Wedge With Graben–like fault

This very simple synthetic example has been constructed to illustrate some of the main considerations in constructing a workflow involving the massaging software. All the files referred to are available in the distribution tarball (see the `3LayerFaultedWedge` example directory), so are not listed here.

This model is constructed as a 2d case for visualisation simplicity. Fig. 3 shows a 'truth–case' cross–section of the 40x1 cornerpoint grid with seismic.

[Fig. 3 about here.]

It is common to fix a reference depth layer to the strongest reflection –often the top of a reservoir, which is chosen as the top of layer 2 (see also Appendix 3). This reference is invisible when the sand pinches out, so is extrapolated horizontally for simplicity. A depth uncertainty of $\sigma_z = 5$m is attached to this surface in the *Delivery* prior; all other depths are referenced to this and computed from layer velocities and times. After the delivery inversion is run, typically with
`% delivery -PD -v 3 -RWS -N 100 -m prior_traces.su ModelDescription.xml`,
the summary file of statistics is generated for use in the massaging step:
```
% deliveryAnalyser -i realisations.su --massage-analyse ep,cdp,gx,gy \
1,2,3 d,thickness,net-sand,NG massage_analysis.mab
```
which produces a set of median and covariance statistics for layers 1,2,3 and properties depth, thickness, net–sand, and N_G in the file `massage_analysis.mab`. The seismic header words `ep,cdp,gx,gy` are reproduced in the file for spatial locations. This file will be very large for big models, so special binary compression techniques are used.

The detail of what happens with the statistics near the pinchout is interesting. Fig. 4 shows the histogram of the layer–2 thickness computed from exhaustive (5000 realisations) samples at four of the traces where the sand is thin. Note the evolution towards symmetric Gaussian–like distributions as the layer thickens.

[Fig. 4 about here.]

Producing corner–point style grids is then straightforward. An XML file *Massager.xml* is created with suitable entries for the required properties, smoothing variograms, residuals' variograms, and hard well data etc. A typical runtime command is then
`% deliveryMassager Massager.xml -v 3 -a -N 10 --ecl`
which would produce 10 realisations of the cornerpoint grid, plus the p50

model, in files with obvious names like *MassagedEclPropertiesLayer** and suitable suffixes. See Appendices 1 and 2 for more details.

The layer depths extracted from the p50 statistics at each seismic trace lack spatial continuity for two reasons: 1) noise in the seismic traces feeding the inversion, and 2) sampling error in the MCMC ensemble, which will usually scale like $N^{-1/2}$ if the ensemble has N samples, but may central limit more slowly if the posterior contains many modes and/or the modes have eccentric shapes (like pinchouts, which are 'half–Gaussians'). The example above, with only 100 realisations, can be expected to have substantial sampling noise (in practice, more realisations would be generated to reduce this). Fig. 5 shows what the median model extracted from these samples will look like in both time and depth.

[Fig. 5 about here.]

If we impose smoothing with an isotropic Gaussian variogram of range 600m, and form a most–likely model and several realisations, the cornerpoint grids look like those depicted in Fig. 6. The XML property attribute `smooth_across_faults` controlling continuity across faults have been altered to produce the rather non–geological realisations of Fig. 6.d, for demonstrative purposes.

[Fig. 6 about here.]

6.2 A Field Example: Stybarrow

The Stybarrow field off Western Australia has been subjected to the full gamut of *Delivery* style workflow. A more comprehensive overview is given in Glinsky et al. (2005). The field is an early Cretaceous turbidite sandstone, whose structure comprises a narrow, wedge–like NE–to–SW tilted fault block, with normal faults providing closure to the SW. Cross–section and elevation views are in Fig. 7.

[Fig. 7 about here.]

Four wells were used for simultaneous wavelet extraction, using the software of Gunning and Glinsky (2005). The coarse layer–based model constructed for *Delivery* inversion comprised 6 layers in the sequence shale/thin–sand/thin–shale/main–sand/hard–shale/shale. The seismic inversion was run only on traces within the fault block/hydrocarbon trap region. The asset geological team built an ECLIPSE model of the reservoir using the same layering, identifying about a dozen internal faults and associated segments.

Since one of the wells (Stybarrow 4) penetrated the lower fluid contact, the uncertainty of chief interest was that of the net–sand volume within closure above the known contact. Fig. 8 shows grayscale maps of the p50 net–sand of the main sand, with obvious short–scale transverse noise caused by sampling error and various other sources of noise. The asset geologist suggested transverse correlation lengths in the km range for the main bodies in this field, and the smoothing effect of this on the p50 'massaged' map is evident in Fig. 8.b.

[Fig. 8 about here.]

The uncertainties in the main sand net–sand volume are strongly influenced by the correlation lengths of the allowable body fluctuations, as the extent to which the stochastic volumes will central–limit (within the fault-block integration area) to a sharply defined average is strongly controlled by these lengths. The distribution of this volume was estimated by drawing an ensemble of realisations of net–sand, conditioned on well data, and integrating over the maps above the contact level. The code produces ascii files with summary statistics of 200 realisations by using the typical runtime command

```
% deliveryMassager StybarrowMassager.xml -v 4 -a -N 200 --stats
```

Fig. 9 shows typical realisations of the reservoir main sand body generated in this way. The inset Fig.9.b shows the CDF of the net–sand volume computed from these realisations, which the effect of shortening the transverse correlation lengths demonstrated with a second curve. With shorter fluctuation lengths, the overall integrated volume is naturally less uncertain, with a sharper rise in the CDF. The variogram structures do not affect the median volumetric statistics, as expected, since this is only a function of the smoothing algorithm.

[Fig. 9 about here.]

As an example of the 'decoration' algorithm, Fig. 10 shows typical 'decorated' layer images generated using the algorithms of section 4, along a particular cross section where the net–to–gross is quite high. The code produces a user–specified number of decorations for each NG realisation of the coarser layer. In this case, the subgridding uses proportional gridding, with variogram distances measured in the proportional sense within the layer.

[Fig. 10 about here.]

7 Conclusions

The *deliveryMassager* program we have introduced in this paper is an essential tool for coercing the stochastic seismic outputs from the *Delivery* seismic

inversion tool to formats suitable for flow simulation or further 3D modelling and analysis. It performs a merging of expert–prescribed lateral correlations with the vertical correlations inferred in the inversion, which is essential for the generation of both realistic most–likely–case models and for uncertainty studies using stochastic realisations. Hard observations, faulting information, and segmentation requirement are honoured. The massaging process generates industry–standard cornerpoint grid formats usable directly by common 3D modelling tools and flow simulators.

8 Acknowledgements

James Gunning and Chris White gratefully acknowledges generous funding from the BHP Billiton technology program. Discussions with Didier Renard from l'Ecole des Mines de Paris were greatly appreciated.

References

Besag, J., 1986. On the Statistical Analysis of Dirty Pictures. Journal of the Royal Statistical Society, Series B 48 (3), 259–302.

Deutsch, C. V., 2002. Geostatistical Reservoir Modelling. Oxford University Press.

Deutsch, C. V., Journal, A., 1998. GSLIB Geostatistical Software Library and User's Guide, 2nd edition. Oxford University Press.

Glinsky, M. E., et al., 2005. Integration of Uncertain Subsurface Information into Multiple Reservoir Simulation Models. The Leading Edge 24, 990–999.

Greig, D., Porteous, B., Seheult., A., 1989. Exact maximum a posteriori estimation for binary images. Journal of the Royal Statistical Society, Series B 51 (2), 271–279.

Gunning, J., 2003. **Delivery** website: follow links from `http://www.petroleum.csiro.au`.

Gunning, J., Glinsky, M., 2004. Delivery: an open-source model-based Bayesian seismic inversion program. Computers and Geosciences 30 (6), 619–636.

Gunning, J., Glinsky, M., 2005. WaveletExtractor: A Bayesian well–tie and wavelet extraction program, submitted to *Computers and Geosciences*.

le Loc'h, G., Galli, A., 1996. Truncated PluriGaussian method: theoretical and practical points of view. In: Baafi, E., Schofield, N. (Eds.), Geostatistics Wollongong '96. Vol. 1. Kluwer, Dordrecht, pp. 211–222.

Peaceman, D. W., 1996. Calculation of transmissibilities of gridblocks defined by arbitrary corner point geometry. SPE eLibrary, unsolicited paper 37306.

Ponting, D. K., 1989. Corner point geometry in reservoir simulation. In: Joint

IMA/SPE European Conference on the Mathematics of Oil Recovery, Cambridge. pp. 45–65.

Schnabel, R. B., Eskow, E., 1999. A revised modified Cholesky factorization algorithm. SIAM Journal on Optimization 9 (4), 1135–1148.

Skiena, S., 1997. The Algorithm Design Manual. Springer, New York, online software at `http://www.cs.sunysb.edu/~algorith`.

Thompson, C., 1972. Mathematical Statistical Mechanics. Princeton University Press, New Jersey.

Whaley, R. C., Petitet, A., Dongarra, J. J., 2001. Automated empirical optimizations of software and the ATLAS project. Parallel Computing 27 (1–2), 3–35, See the netlib repository: www.netlib.org/atlas.

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

Appendices

Appendix 1: Input files

When running the *deliveryMassager* code, very frequently changed runtime options are reserved for the commandline: the executable *deliveryMassager* self–documents if no arguments are supplied, for those wishing to peruse these options. Otherwise, all input parameters are specified in an XML file (see 1.1 below), but this will in turn reference other files that may be required:

- A *mandatory* summary file of the inversion statistics, which contains trace locations and median and covariance statistics for each of the inversion quantities of interest in all the salient layers. Once the inversion is run, with the realisations in the file (say) `realisations.su`, this summary statistics file is generated using
  ```
  % deliveryAnalyser -i realisations.su --massage-analyse \
  headerwords layer-numbers properties outputfile
  ```
  Here, each of `headerwords/layer-numbers/properties` is a comma separated list, as per the examples.
- If (*optional*) output grids in ECLIPSE style corner point grids are required, a 'receptacle' corner point grid must be supplied, which has the same layering as the inversion model. The critical information gleaned from this file is the geometry and fault blocking, so the fields PINCH, MAPAXES, GRIDUNIT, SPECGRID, COORDSYS, COORD, ZCORN, ACTNUM, and SEGMENTS will be required (SEGMENTS may be skipped, with the consequence that all blocks are assumed to be in the same segment). These fields are routinely supplied in the ECLIPSE files exported from commercial modelling packages like *PETREL*. Readers are referred to ECLIPSE documentation for further detail on these file formats.

If no corner point grid is supplied, the massaging code can produce most–likely maps and realisation files on the same grid as the seismic inversion ('duplicate' mode), which is often very useful. No fault block information is available when using this mode.

1.1 XML formats and Schema

The XML format used to control the massaging process has a meta–description in the associated `Massager.xsd` XML schema file, which can be used in the XML editor supplied with the *Delivery* distribution to produce strictly legal XML files. The format of the XML is largely self-explanatory, but a few explanations may be helpful.

- <delivery_massage_analysis_filename> tags the median and covariance statistics file formed from the inversion analysis step (section 1).
- <x-coordinate_header_word> and <y-coordinate_header_word> define the seismic header words, passed through in the *deliveryAnalyser --massage-analyse* process, which set the transverse coordinates in local UTM coordinates. They must correspond to the UTM coordinates used in the input ECLIPSE files. If these seismic header words are not available, local x,y coordinates can be computed from inline/crossline header words (always available) using the <utm-conversions> block available through the schema. The x,y values read from the named header fields <x-coordinate_header_word> and <y-coordinate_header_word> are then taken to be inline and crossline.
- If the <utm-conversions> block is supplied, this will force the code to compute x,y's from the linear mapping from inline/crossline to x,y implied by the three non–collinear {x,y,inline,crossline} 4–tuples supplied in this section.
- <property_name normalised="false" smooth_across_faults="true"> is a typical property specification; normalised forces truncation to $[0, 1]$, and smooth_across_faults will allow kriging/simulation neighbours to come from different segments.
- Variograms: <azimuth> specifications on all variograms are GSLIB 2D convention: degrees clockwise, with x as Easting, y Northing. Lengths are in UTM units.
- Eclipse *input* grids. Currently, a common input grid <ECLIPSE_grid_filename> is expected for all layers. <ECLIPSE_segments_keyword> is a block centred integer property that labels the fault blocks. Positive numbers are expected. <ECLIPSE_reference_depth_layer_number> is the number of the layer in the Eclipse model that corresponds to the depth surface from which the model will be hung (see also Appendix 3). It should correspond to the same physical layer tagged by <master_depth_layer_number> in the XML file controlling the *Delivery* inversion. The actual reference layer number might differ if the Eclipse model contains extra layers above the inversion model. <ECLIPSE_layer_number> identifies the <inversion_layer_number> within the <layer> block with the correct layer of the Eclipse model. The <ECLIPSE_layer_number> entries will form a simple unit–ascending sequence in the XML file.
- Eclipse *output* grids. The runtime option *–ecl* will write massaged properties (most–likely, realisations) into an output Eclipse grid whose file name(s) are prefixed by the content of <ECLIPSE_output_grid_filename>.
- <well_observation> entries in each layer are optional. Positions are in UTM's again. Entries for properties that are not 'massaged' will be ignored.
- <stats> blocks within a layer define a depth window within which contributions to the volumetric statistics can be made. For example, if a fluid contact depth is known, suitable entries here will permit the distribution of net-sand above a contact depth to be formed.

Appendix 2: Output files

The code produces a variety of output files, with names constructed from relevant entries in the master XML file. The simpler files are in naive geoEAS ascii format used by GSLIB (Deutsch and Journal, 1998) for ease of parsing. Stochastic outputs ('realisations') are generated if the `-N number` flag is supplied.

If ascii–mode is used (runtime flag `-a`)

- `post_massaged_properties_layer_*.txt` is a geoEAS file containing columns for the gridpoint locations, segments, and most–likely values and standard deviations for all the properties in the layer implied by the filename.
- `post_massaged_realisations_layer_*.txt` is geoEAS, with columns for the gridpoint locations, segments, and stochastic samples for each property and realisation in the layer implied ¡by the filename.

If files are being written to the cornerpoint grids (`--ecl`), we get

- A *most–likely* cornerpoint grid (tagged `<ECLIPSE_output_grid_filename>`) with the geometry altered in accordance with the massaged layer thicknesses from the inversion, and hung on the nominated reference depth layer. Other properties are written into this file as block–centred, with entries for both the most–likely and standard deviations of each massaged property.
- A separate realisation cornerpoint grid file for each realisation, with block–centred values for the properties. Filenames are suffixed versions of the most likely grid.

If volumetric statistics of certain properties are requested (runtime flag `--stats`), simple ascii files (`Realisation_Summary_Stats*`) with the cumulative distribution of a requested ordering property (e.g. net–sand) are generated.

Appendix 3: Treatment of reference depths

A variety of processes control the reference depth used to 'hang' the ECLIPSE models. The *delivery* inversion xml file requires an entry tagged `<master_depth_layer_number>` which is a layer whose top depth is used to hang all other depths in the inversion model, via effective–velocity (`vp_eff`) $\times$ time conversions. If the quantity `<sigma_depth>` for this layer is set at non–zero, the reference depth and other layers will acquire a normally–distributed stochastic component of the specified standard deviation in `<sigma_depth>`. The *deliveryAnalyser --massage-analyse ...* step will detect and model this un-

certainty and the cross correlations with other layer–depths it induces. The mean and p50 reference layer depth from the *massage-analyse* step will be the reference depth supplied to the inversion, within sampling error.

In the massaging code, 3 possibilities arise.

- Massaging of depth `d` is specified, typically by `<property_name normalised="false" smooth_across_faults="false">d`. Here, the reference depth surface will be smoothed (respecting faults if specified), and the reference depth residual uncertainty will be a Gaussian random field of the same standard deviation as that of the smoothed reference–depth massage-analyse'd standard deviations, but transversely correlated according to the supplied layer variogram. All ECLIPSE gridblock tops and bottoms will be set relative to this depth and the massaged thicknesses, if the latter if available. Thicknesses of ECLIPSE layers that are not massaged are preserved. The tag `<ECLIPSE_reference_depth_layer_number>` identifies the ECLIPSE grid layer corresponding to the `<master_depth_layer_number>` in the *delivery* inversion model. If thicknesses are not massaged, all the in-situ ECLIPSE gridblock thicknesses are preserved even though the model is re–hung according to the massaged reference depth.
- Massaging of depth `d` is not specified. In this case, the in–situ ECLIPSE reference depth is preserved. ECLIPSE layer thicknesses will be set by the massaged thicknesses if thickness massaging is specified.
- Massaging of depth `d` is specified, but the original inversion did not include a stochastic component corresponding to the reference depth uncertainty, so the reference depth standard deviation is zero. A global reference depth uncertainty can be re–introduced at the massaging step here by supplying the flag *--systematic-depth-uncert standard-deviation* to the commandline of *deliveryMassager*.

Appendix 4: The thin–layer detection problem

The strong nonlinearity of the forward model in the regime of thin–layers makes the correct introduction of trace correlations difficult. Thin layers are always difficult to detect (or reject) with strong probability in single traces, as a layer introduced between identical bounding layers will introduce equal and opposite reflectivities, which will nearly cancel each other in the convolution, and thus produce very weak (i.e. within–noise level) amplitudes. But the inversion at each trace can still provide an (perhaps weakly) *updated* estimate of the probability p that the layer is present. The case $p \approx 1/2$ is most interesting.

There is a simple mapping of this problem to a Bayesian Markov Random Field (MRF) model (Winkler, 2003; Besag, 1986) which offers considerable insight.

If we think of an array of traces i characterised by an integer $x_i = \pm 1$ denoting 'layer present/absent' at each trace, then the product of (independent) updated likelihoods p_i over all traces in the set can be written in the form

$$p(\{x_i\}) = \exp(\sum_i B_i(x_i - a_i)), \tag{19}$$

where $B_i = (1/2)\log(p_i/(1-p_i))$ and $a_i = (1 - (1/B_i)\log p_i)$ are constants that come from setting the odds ratio

$$p_i/(1-p_i) \equiv \exp(B_i(+1 - a_i))/\exp(B_i(-1 - a_i)).$$

We may think of the exponent in equation (19) as a (-ve) 'likelihood' Hamiltonian for the problem, which needs to be added to a Hamiltonian expressing the prior mean and correlations between the states x_i, as they might plausibly be related in a model prior to any observations (i.e. inversion results). If we write the prior for the model $\{x_i\}$ as a MRF with coupling over nearest neighbours given by the Hamiltonian

$$H = -\beta \sum_{j \sim i} x_i x_j,$$

then the prior model corresponds to an Ising model with inverse temperature β, mean state $\langle x_i \rangle = 0$ (i.e. agnostic view of layer presence/absence), and, in 1d, an exactly derivable correlation function (in the large system limit):

$$\rho_{i,j} \equiv \langle x_i x_j \rangle \sim (\tanh \beta)^{|i-j|}.$$

Clearly the correlation decays geometrically/exponentially between traces, so we define a correlation length λ_c by

$$\rho_{i,j} \sim (\tanh \beta)^{|i-j|} \equiv \exp(-|i-j|/\lambda_c).$$

Clearly, longer correlations (large λ_c) correspond to 'colder' temperatures (large β).

When we add the likelihood Hamiltonian to the prior Hamiltonian, the overall system is

$$H = -\beta \sum_{j \sim i} x_i x_j - \sum_i B_i(x_i - a_i),$$

which corresponds exactly to the Bayesian image models discussed byWinkler (2003), in the binary case. Exact MAP estimated of the most probable state

can be computed by annealing or the Ford–Fulkerson algorithm (Greig et al., 1989).

Some insight into the effect of the correlations can gleaned by considering the one dimensional case with a common update probability $p = p_i$. This then corresponds to the Ising model in an external magnetic field B. The question of interest is then, given a set of (identical) likelihood updates at each trace corresponding to $B = (1/2)\log(p/(1-p))$, what is the *expected* state of the system. This corresponds precisely to the mean Ising magnetisation, which is known (Thompson, 1972) for the 1D case to be

$$\langle x \rangle = \frac{\sinh B}{\sqrt{e^{-4\beta} + \sinh^2 B}} \tag{20}$$

$$= \frac{(1/2)(\sqrt{p/(1-p)} - \sqrt{(1-p)/p})}{\sqrt{\tanh^2(1/2\lambda_c) + (\sqrt{p/(1-p)} - \sqrt{(1-p)/p})^2/4}} \tag{21}$$

Graphs of this curve show that the correlation in the prior strongly 'corroborates' any weak inclinations in the likelihood p. E.g. for a correlation length $\lambda_c = 10$ and $p = 0.6$, the expected state is *almost certainly* 'layer present'.

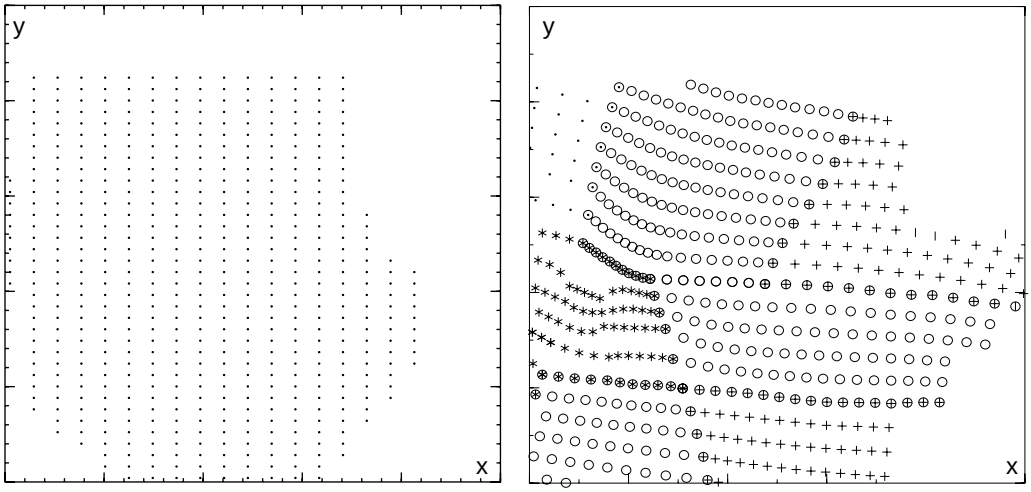
[Fig. 11 about here.]

This behaviour is reasonable: we expect a particular observation to be repeated many times if the correlation lengths are long, and if the observations are truly independent, the multiplication of probabilities forces the 'suspected' state to be very much more likely. In the inversion context, we would have to be very careful with asserting true independence of observations, since the imaged amplitudes may well have systematic effects from the processing or other geological effects in the overburden.

For the case where the trace likelihoods p_i vary, a typical example of the effect of the correlations (via β) on the MAP estimate is shown in Fig. 12, as calculated by dynamic programming methods.

[Fig. 12 about here.]

List of Figures



a) Seismic trace array

b) Cornerpoint grid corners, segment -coded

Fig. 1. (a) Typical layout of input–grid trace array for seismic inversion: regular spaced points in a polygon. (B) Typical output grid for some layer, derived from ECLIPSE–style corner point grids with replication of grid points at fault planes. Different symbols identify segments or fault blocks.

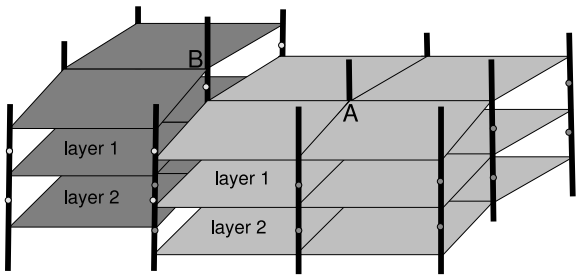


Fig. 2. Architecture of corner point grids. Typical vertical fault shown. Each layer's output grid is formed from the x, y projections of the edge midpoints (little circles). Thickness and depth information can be shared by blocks that meet at a common points (A), but this information must be replicated at corners along the fault plane (B).

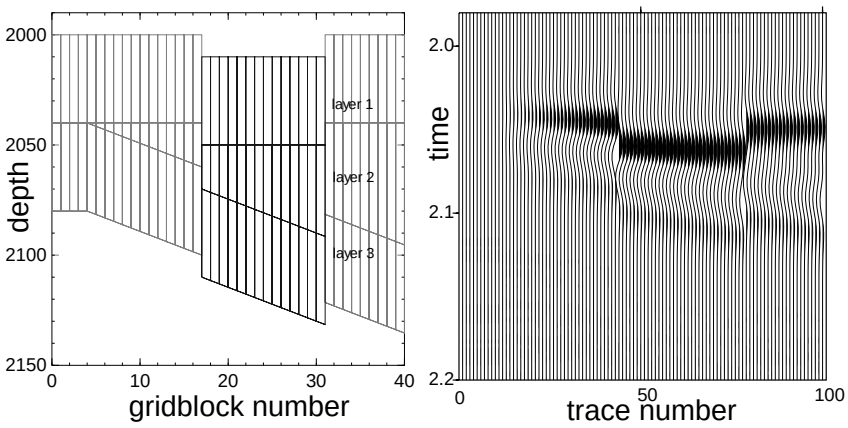


Fig. 3. Truth–case faulted 3–layer shale/sand/shale wedge model, with the sand layer of about 70% N_G. Fault–block segments (based on seismic 'interpretation' in the corner point grid are shown grayscale coded. The cornerpoint grid is based on a 100×100m grid, but the seismic is sampled at 40m spacing, i.e. denser traces than grid blocks, which is a common situation.

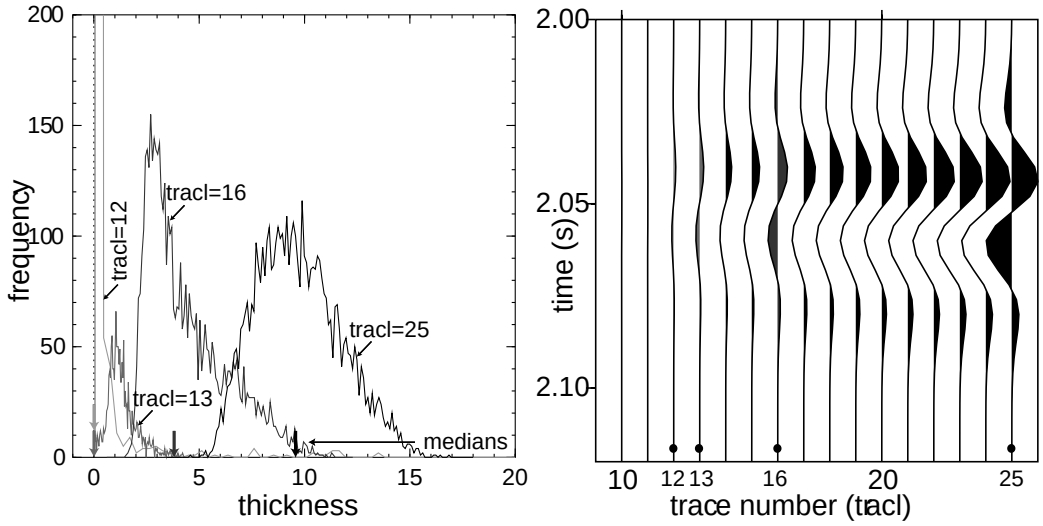


Fig. 4. Histogram of the reservoir (layer 2) thickness at the four traces tracl=12,13,16 and 25 (shown at right). The distribution gradually evolves from a pure spike at thickness=0, through a mixed spike and minor mode (tracl=13), to a nearly Gaussian single mode when the amplitudes improve. The median (50% quantile) statistic used in the analysis to define the most–likely value at each trace is highlighted on the axis.

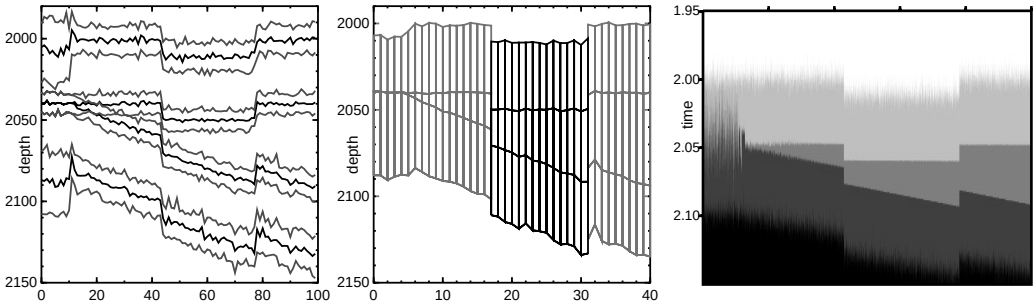


Fig. 5. Issues in forming a cornerpoint model from undersampled inversion with no smoothing. a) Layer depth p10,p50,p90 surfaces on seismic grid. b) Eclipse grid interpolated from seismic with very weak smoothing. c) Inversion realisations on time scale, with obvious increase in uncertainty in pinchout region. Notice how depth uncertainties increase where the amplitudes disappear, since both event times and reflection coefficients (and therefore velocities) cannot be inverted for.

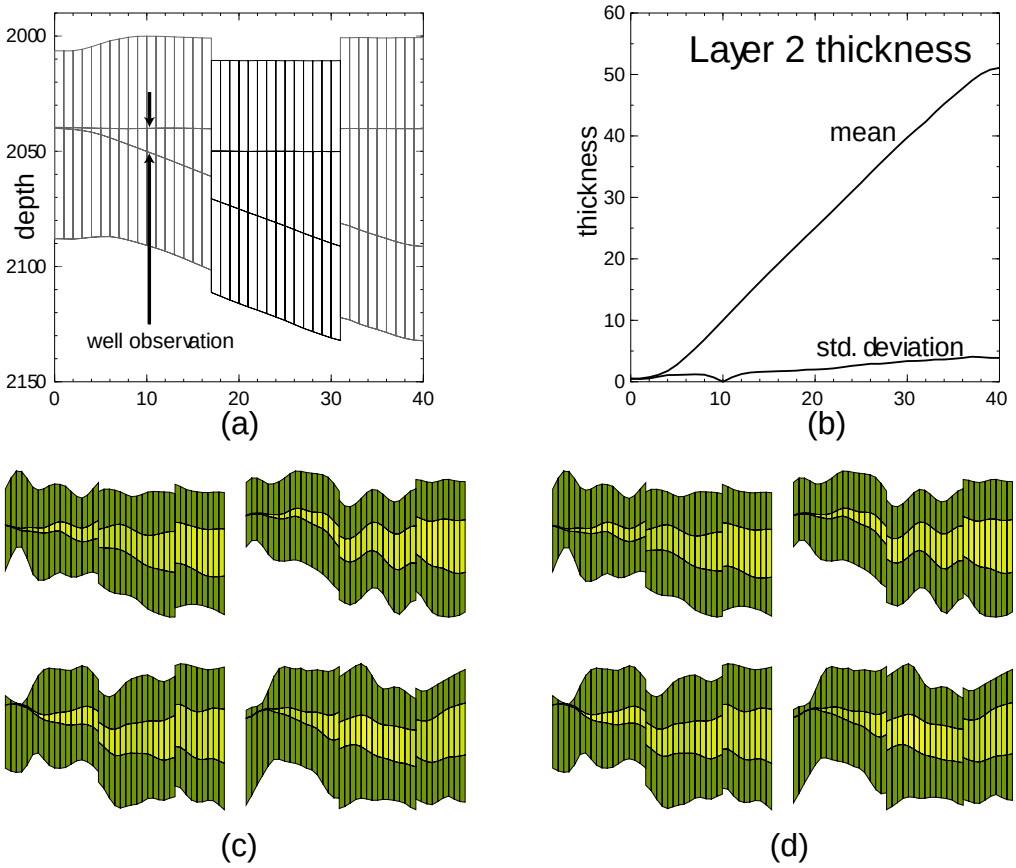


Fig. 6. a) Cross section of most likely smoothed corner point model, with segments grayscale–coded b) p50 thickness and thickness uncertainty of middle layer, showing effect of single well observation in the 10th gridblock c) Four typical model realisations, with N_G grayscale scale between dark (shale) and sand (light). Thickness and net–sand are continuous across faults d) Same as c), but with continuity of layer thicknesses etc not enforced across the faults.

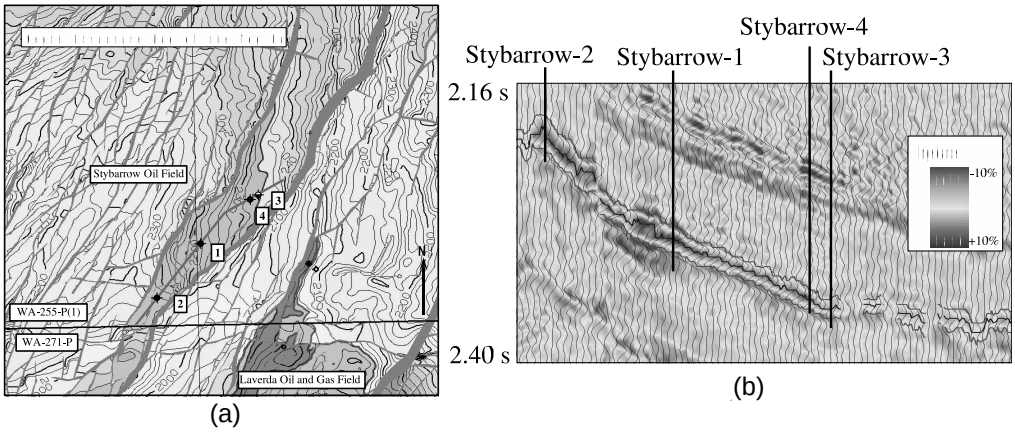


Fig. 7. (a) Plan and (b) elevation views of the Stybarrow Field, with the elevation in the down–dip direction along the line passing through the four wells.

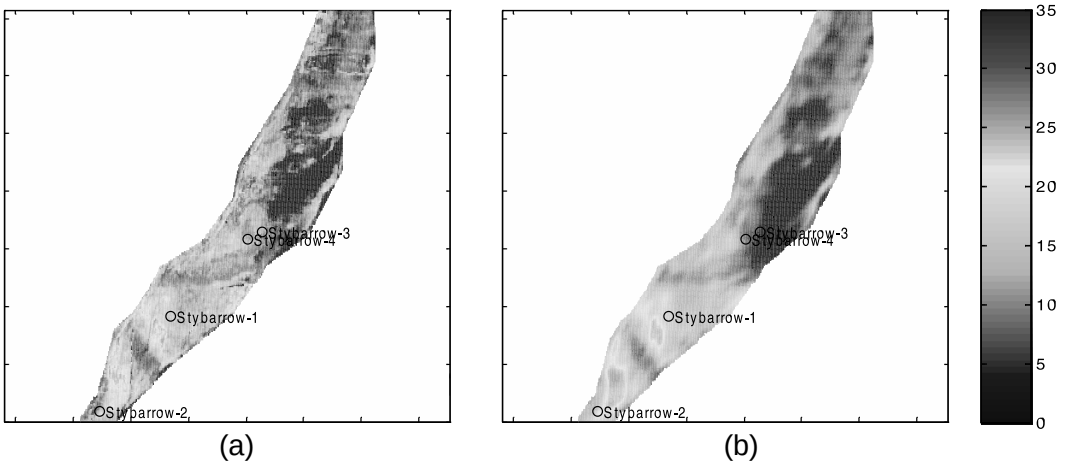


Fig. 8. Median maps of net–sand for the Stybarrow main sand, (a) directly from *Delivery* inversion, (b) after massaging with long–scale smoothing variograms. The maps smooths across segments since the post–depositional faulting is treated as normal.

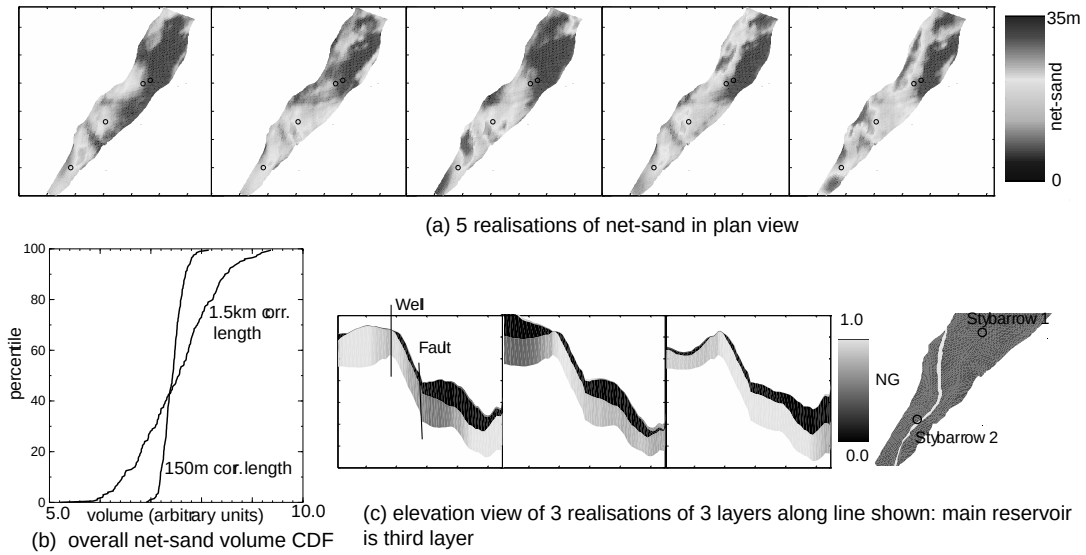


Fig. 9. Realisations of the main sand net–sand map conditional on the well data and the seismic inversion constraints, (a) in plan view, (c) in elevation. (b) The volumetric CDF is shown inset, with arbitrary units. Correlation lengths used in the actual model were on the km scale, leading to much greater volumetric uncertainty than that obtained with short correlation scales.

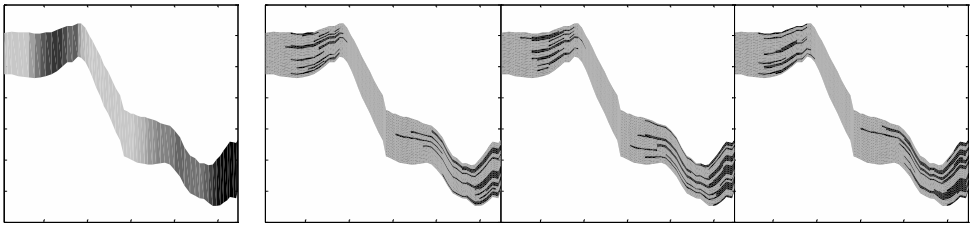


Fig. 10. Downscaled, or 'decorated' realisations of the main sand net–sand, for a particular NG realisation, shown in elevation along the same transect shown in Fig. 9.c.

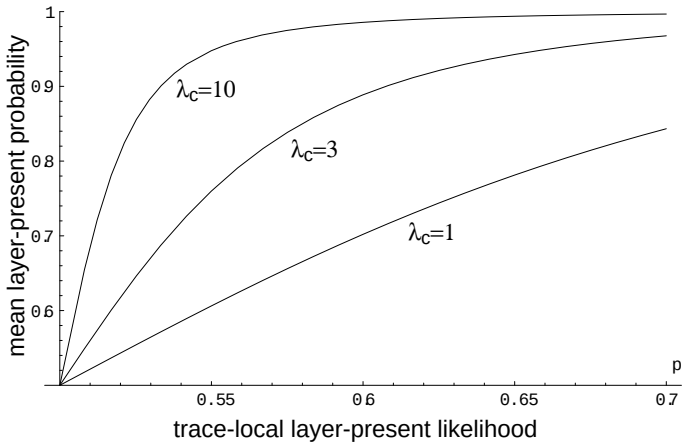


Fig. 11. Mean 'layer present' probability for Ising system with various correlation lengths λ, versus trace–update probability p.

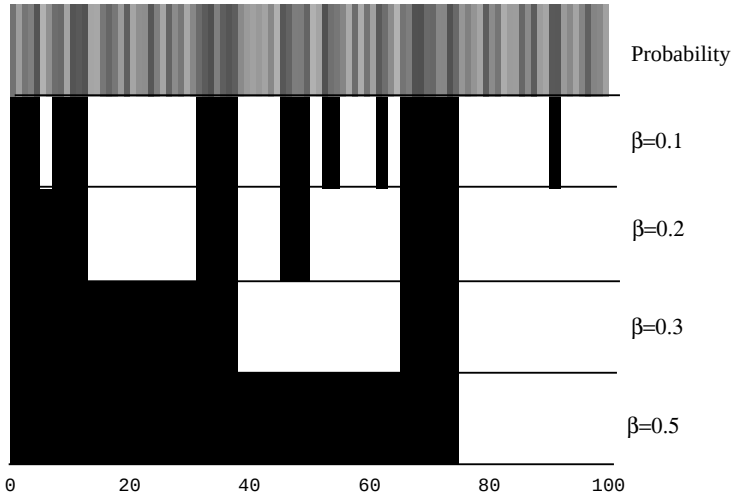


Fig. 12. One dimensional 'layer–present' probability sequence (on grayscale 0=black, 1=white) for $i = 1 \ldots 100$ traces, with the updated MAP layer model for various values of the correlation/smoothing parameter β. Higher β values generate more aggressive smoothing/segmentation.