Assisted seismic history matching of the Nelson field: managing large numbers of unknowns by divide and conquer.

Karl D. Stephen
Institute of Petroleum Engineering,
School of Energy, Geoscience, Infrastructure and Society,
Heriot-Watt University,
Riccarton,
Edinburgh,
Scotland, EH14 4AS,
UK
e-mail: k.d.stephen@hw.ac.uk

Abstract

Automatic history matching may be used to condition reservoir simulation models with time-lapse seismic data. Stochastic optimization algorithms are used to perform a good search of the parameter space to ensure effective determination of the best models. These approaches can require many thousands of simulations for large dimensional problems. Divide and conquer is an assisted history matching approach that enables deconvolution of the parameters so that they can be searched more efficiently and also leads to better uncertainty analysis.

We present an application of this approach on the Nelson field. Nine years of production history data are used along with baseline and monitor seismic surveys. Localised variations were made to permeability and net-to-gross ratio in the model. The reservoir was divided into separate parameter regions by combining experimental design and proxy model analysis. The former enabled insignificant parameters to be discarded while the latter showed that each region could be treated as a separate history matching sub-problem. Each sub-problem was then solved simultaneously using an adapted stochastic neighbourhood algorithm.

The results show that a forty-two dimensional problem could be reduced to a combination of three 9D problems and a 3D problem due to the spatial deconvolution of parameters and misfits. An improved match was obtained for the production and seismic data. Compared to a full stochastic search of the parameter space, the number of required models was several orders of magnitude smaller. Improved uncertainty analysis was made possible resulting in better understanding of the future behaviour of the reservoir.

An improved match to reservoir models leads to better confidence in their prediction and thus they can be used more effectively in reservoir management. The method presented here to improve the match retains the benefits of stochastic searching without the penalty of requiring an impractical number of simulations.

1. Introduction

Seismic history matching is the process where time-lapse (4D) seismic data are used to condition reservoir simulation models along with more conventional production data. 4D seismic data offers spatial information that is somewhat missing from production data and by integrating this data, history matched models will then give a more accurate representation for forecasting. Numerous studies have been presented previously and despite the success of this approach there are many known difficulties. Some of these arise from the overwhelming number of unknowns. This is often reduced by suitable choice of parameterization. Optimization routines such as gradient based methods (e.g. Lépine, et al., 1999, Dong and Oliver, 2003, Gosselin et al., 2003, Mezghani, et al., 2004) are not so troubled by the volume of the search space but they often produce a single best model unless combined with stochastic approaches. Probabilistic methods such as EnKF (e.g. Skjervheim et al. 2007, Aanonsen et al. 2009) work around this by finding geostatistical realizations so that the simulations match the data while maximizing probabilities in a Bayesian framework. However the
EnKF approach can be problematic with large seismic datasets and a number of specific techniques such as smoothing or localisation must be applied (Aanonsen et al. 2009).

In this paper we consider an approach to separate the problem into manageable sub-problems using an approach which we call Divide and Conquer (Sedighi and Stephen 2010). The philosophy is similarly applied in other areas such as calculation of Fast Fourier Transforms and the Merge Sort Routine. The aim is to break a problem into smaller problems which can be treated identically or in a similar manner using existing algorithms. This approach has been presented previously in seismic history matching and has been applied to the Schiehallion field with some success (Sedighi and Stephen 2010). It is based on the estimation of a proxy model which is then analysed to determine interactions between parameters and their effect on the misfit. Parameters showing no interaction can be modified independently, albeit as groups. This reduces the volume of the hypercube of the parameter space and makes the search more efficient.

Proxy surfaces have been used in reservoir simulation as a way of generating fast models that capture the majority of reservoir behaviour (e.g. Cullick et al. 2006, Zangl et al. 2006). In history matching they may be used to represent the misfit surface (Christie et al. 2012) and are an alternative to methods such as gradient based and adjoint methods (Wu et al. 1999) which calculate the response of the misfit to property changes. There are also various options for calculating the proxy function depending on their form. Kriging of the response variable can be carried out (Pan and Horne 1995; Goodwin 2015) or machine learning (Zangl, 2006) through neural nets have been used. We use a reasonably simple quadratic equation.

We apply a Divide and Conquer approach to the Nelson field for the first time in this paper. The problems are somewhat different compared to the Schiehallion field in that permeability and net-to-gross ratio are the main unknown quantities rather than barrier transmissibility, there are many more wells and the reservoir is generally better connected. However, we have applied seismic history matching previously to this field without Divide and Conquer being fully explored. The approach is attractive for seismic history matching where seismic sensitivities are quite difficult to calculate directly as a numerical form of the analytical solution.

2. Method

History matching consists of an iterative modification of simulation input data during which predictions of equivalent observed data are made followed by an assessment of the match via a misfit function which is used to guide updates (Fig. 1). Model updates are controlled through parameterisation using a control vector of parameters, $\theta$. Conventionally production data are predicted and may include oil and water rates and possibly reservoir or flowing pressures. Time-lapse seismic data may also be included in the process. This step requires that either seismic data is inverted to obtain reservoir saturation and/or pressure data or some form of seismic modelling is required. In this paper we predict maps of acoustic impedance using a petro-elastic model (for details see Stephen et al. 2009 and Kazemi et al. 2014) and compare to equivalent observed data obtained by inversion.

2.1 Petro-elastic model and seismic data

We first calculate the bulk density along with the saturated bulk and shear moduli for each simulation cell using output from the simulator and a petro-elastic model. The bulk density is then

$$\rho = \rho_{\text{sa}}NTG(1-\phi) + \rho_{\text{sh}}(1-\text{NTG}) + [\phi(\rho_{\text{w}}S_{\text{w}} + \rho_{\text{o}}(1-S_{\text{w}})]\text{NTG} \tag{1}$$

where $\rho_{\text{sa}}$, $\rho_{\text{sh}}$, $\rho_{\text{w}}$, $\rho_{\text{o}}$, are the densities of sand matrix, shale matrix, water and oil respectively, NTG is the net-to-gross ratio, $\phi$ is the sand porosity and $S_{\text{w}}$ is the water saturation. In the field studied here, laboratory measurements revealed that dry and shear bulk moduli follow quadratic equations in terms of porosity. The dry bulk modulus for sandstone in each cell is then:

$$\kappa_d = 32(1 - 2.07\phi + 2.38\phi^2) \tag{2}$$
and the shear modulus is:

$$\mu = 30.2(1 - 4.67\phi + 7.16\phi^2)$$  \hspace{1cm} (3)

The saturated bulk modulus was calculated using the Gassmann equation:

$$k_{sat} = k_d + \frac{(1-k_d/k_m)^2}{\phi + \frac{1-\phi}{k_f} - \frac{k_d}{k_m}}$$  \hspace{1cm} (4)

where $k_m$ is the bulk modulus of the sand grains and taken to be 37 GPa (Simmons and Wang 1971) and $k_f$ is the fluid modulus from (Domenico 1974):

$$\frac{1}{k_f} = \frac{1-S_w}{k_o} + \frac{S_w}{k_w}$$  \hspace{1cm} (5)

where $k_o$ and $k_w$ are the oil and water moduli respectively. From lab measurements, these are pressure dependent with

$$k_o = 10.15P + 2.44 \quad \text{and} \quad k_w = 10.15P + 0.306$$  \hspace{1cm} (6)

where P is pressure in GPa.

These equations are used along with grid cell pressures and porosity values to compute the dry bulk and shear moduli for each simulation cell. From this the p-wave modulus is calculated and up-scaled vertically using Backus (1962) to give a single value for each cell over the reservoir interval:

$$I_{mod} = \sqrt{\frac{\rho}{\kappa_{sat}} \frac{1}{\kappa_{sat} + 4\mu_{sat}}}$$  \hspace{1cm} (7)

where “<\kappa>” is the arithmetic mean over the depth weighted by cell volume. Each cell is divided into sand and shale when net-to-gross ratio is less than 1 and the appropriate shear and bulk moduli are used. Shale is dry and of constant bulk and shear modulus at 20 GPa and 6 GPa respectively. The above average is then weighted by rock type volume (Postma, 1955). This is directly equivalent to applying a harmonic average within the cell and then averaging over the depth. Bulk density is similarly calculated.

Observed seismic attributes are obtained as an equivalent pseudo-acoustic impedance property for each bin in the seismic cube. These data are compared to the downscaled model data (Stephen et al. 2006). A distance weighted interpolation method is used to create a representation of the model on the seismic grid. Then the data can be compared bin by bin.

### 2.2 Misfit calculation

The seismic misfit function, which varies with the input parameter vector $\theta$, is defined:

$$M_S(\theta) = \sum_i \frac{(\Delta I_{obs}^i - \Delta I_{mod}^i)^2}{\sigma_{st}^2}$$  \hspace{1cm} (8)

The denominator, $\sigma_{st}$, is a measure of data and model errors which have been estimated to be uncorrelated and are obtained from the standard deviation of the time-lapse signature in the region outside of the oil flowing region where the signal should be zero. $\Delta I_{obs}$ is the observed estimated seismic impedance (Stephen and Kazemi 2014) of the $i,j$ bin and $\Delta I_{mod}$ is the down-scaled equivalent data for the bin. A similar misfit is used for production data summing over time series data of oil and water production rates. The production and seismic misfits are summed to give a total misfit:
\[
M(\theta) = \sum_{ij} \left( \frac{(\Delta I^i_{obs} - \Delta I^i_{mod})^2}{\sigma^2_{\Delta i}} \right) + \sum_{r=1}^{N_r} \sum_{q=1}^{n_w} \sum_{p=1}^{n_w} \left( \frac{(W^p_{obs} - W^p_{mod})^2}{\sigma^2_{w}} \right)
\]

where \(\sigma_w\) is the standard deviation of the data errors for production variable \(W^{pq}\) which represent the \(p^{th}\) well data variable in the \(q^{th}\) well at the \(r^{th}\) time step. The subscripts “obs” and “mod” indicate observed and model data respectively (and may include fluid rates and well pressures for producers and injectors). There are \(N_w\) production variables, \(n_w\) wells and \(n_{data}\) time steps.

In this paper we used the misfit in Eq. 9 having calibrated the data errors such that a single misfit is used. 36

where \(\sigma_w\) is the standard deviation of the data errors for production variable \(W^{pq}\) which represent the \(p^{th}\) well data variable in the \(q^{th}\) well at the \(r^{th}\) time step. The subscripts “obs” and “mod” indicate observed and model data respectively (and may include fluid rates and well pressures for producers and injectors). There are \(N_w\) production variables, \(n_w\) wells and \(n_{data}\) time steps.

We ignore regularisation in this context, which is an approach to reduce the ill-posed nature of the problem and help ensure that suitable models are obtained. The regularisation term penalises models that deviate from the expected values of the parameters in the model. The regularisation term is also related to the prior probability distribution of parameters, particularly when working in a Bayesian framework. In this work we use a uniform prior probability distribution for the parameter variables which is really the equivalent of a “top hat” function as parameters have a maximum variation from -1 to 1 (the parameters are the logarithm of multipliers on porosity and permeability which can therefore vary by a factor of 10 either way). This provides a limitation to the possibility of generating solutions that are not geologically acceptable. Such an approach is often taken when the simulation model has somewhat lost the connection to the fine scale geological model, perhaps through upscaling or because the original data is not present. In this work an upscaled model was used having been created in proprietary software which was unavailable to the history matching project.

Another approach that uses geostatistics can be configured for local updates retaining the globals statistics. Marteau et al. (2014) have shown that local changes can be blended with multiple realisations. Another approach that uses geostatistics is the Ensemble Kalman Filter (for a review see Aanonsen et al. 2009). The properties can be updated as data is sequentially assimilated in time providing they are Gaussian in distribution. These geostatistical methods are excellent for updating models in a Big Loop approach where the original geological model is changed and may require an upscaling step. Alternatives also include the Discrete

### 2.3 Parameterization

History matching using requires a method of controlling changes to the model known as parameterization. Early methods using gradient based search algorithms controlled changes using regional multipliers (e.g. Gosselin et al. 2001). While flexible, this approach is ungeological with regions mainly being represented as boxes. A facies based approach has been tried with seismic history matching (e.g. Kazemi and Stephen (2012) although insufficient control was available given the scale of the facies bodies. The pilot point method (de Marsily et al. 1984) was derived to improve the controls, smoothing changes out for petrophysical properties and enabling changes that resemble those of underlying geostatistical models. Local changes are made at the pilot points and these are propagated using Kriging. Geostatistical approaches have been applied including the gradual deformation (Hu, 2000) and the probability perturbation methods (Hoffman and Caers, 2003). The former blends geostatistical realisations while the latter blends properties in probability space. Both can be configured for local updates retaining the globals statistics. Marteau et al. (2014) have shown that local changes can be blended with multiple realisations. Another approach that uses geostatistics is the Ensemble Kalman Filter (for a review see Aanonsen et al. 2009). The properties can be updated as data is sequentially assimilated in time providing they are Gaussian in distribution. These geostatistical methods are excellent for updating models in a Big Loop approach where the original geological model is changed and may require an upscaling step. Alternatives also include the Discrete
Cosine Fourier Transform approach (Jaffarpour and McLauchlin, 2009) and other methods using basis functions (Sarma et al. 2008).

In some field studies, however, the geological model may be generated in software that cannot be accessed during the history matching routine, the modelling routine may be too complex or it may be desirable to avoid upscaling, which itself introduces uncertainty. Since this is the case in this study, we therefore reverted to a more flexible approach and use the Pilot Point method with Kriging as described later.

2.4 Stochastic search algorithms

We use a stochastic algorithm to control changes to the properties during the history matching process. In this paper we use the neighbourhood algorithm (NA) (Sambridge 1999a). This methods is ensemble based and can be considered generational, analogous to evolutionary and genetic algorithms. The problem requires that we can express the parameters that we change as an orthogonal hyper cube from which we sample. We have used the Neighbourhood Algorithm extensively elsewhere (e.g. Stephen et al. 2009). We refer the reader to Sambridge 1999a for mathematical details and summarise the approach as follows:

1. Define $n_i$ initial models by selecting a combination of parameters randomly from the parameter space, set up and run the simulations and then generate misfits;
2. Select the best $n_r$ models by ranking misfits from all models generated so far.
3. Generate $n_r$ sub-volumes of the parameter space around the best models from Step 2 from which new models will be generated. The shape of the sub-volume depends on the optimization algorithm (see below).
4. Generate $n_s$ new "offspring" models so that $n_s/n_r$ models are selected for each one of the sub-volumes. New models are selected by randomly picking parameter values within the sub-volume of the parameter space. The Neighbourhood Algorithm uses a random walk in each sub-volume starting from the model that defines it.
5. Calculate the misfits for the new generation of models.
6. If not sufficiently converged or insufficient models have been generated, go to step 2.

The NA creates subvolumes from which new models are sought, as defined in Step 3, using Voronoi cells. Each model represents a point in the multidimensional hypercube and the Voronoi cell defines the volume closest to the model in a Euclidian sense. The idea is that the sub-volumes created as Voronoi cells reduces the search space iteratively and the method is therefore self-adaptive. This approach is similar to the Genetic Algorithm (GA) but that it uses a single parent (equivalent to parthenogenesis). In a real number GA (e.g. Stephen and Arwini 2010), two models are selected as "parents" from the set of $n_i$ best models. The sub-volume of the hypercube of Step 2 is then chosen as a hypercuboid with size scaled to the separation of the parents. This is usually called the crossover step in the GA. Apart from these differences, the NA and GA are very similar. The GA is obviously controlled by the way in which parents are combined. There is usually a mutation step to increase exploration and avoid false convergence which can occur.

The GA and NA methods both have advantages and disadvantages. The NA clearly has fewer control variables. The larger we set $n_r$, the more exploratory is the search and convergence will take longer. If $n_i$ is smaller then information is exploited more but the search may get stuck in a local minimum. NA is attractive by its simplicity. We always generate $n_s/n_r$ models for each of the best $n_r$ models. It is possible to invent an endless set of possible additional parameters to make it more sophisticated. For example we could detect the size of Voronoi cells and use that to speed up convergence or even widen the search. The GA can be simplified by ignoring mutation. The size of the sub-volume that is sampled by GA is a control variable, however. The GA method can reach a false convergence if the difference in parents shrinks faster than the convergence of the objective function (Stephen and Arwini 2010). In this work we use the basic NA.

It is evident that the GA method may be better suited to searching simpler response surfaces with one or two solutions while the NA is better for cases with multiple solutions. In cases with multiple minima, the GA may choose two parents each close to a different minimum resulting in an expanded subsequent search volume. This can delay convergence but does help the search for alternative mimima. The NA was thought to be ideal for cases with many mimima as the sampling occurs in greater density close the solution. On the other hand, the NA is inefficient if there is only a small number of mimima.

We will now show that NA is also inefficient for large problems where the misfit space can be
deconvolved. We examine the efficiency of the NA using a simple quadratic response surface for an n dimensional problem with the response surface in place of a misfit:

\[ f(\theta) = C_o + \sum_i^n C_i \theta_i + \sum_{i=1}^n \sum_{j<i}^n C_{ij} \theta_i \theta_j + \sum_i^n D_i \theta_i^2 \] (10)

\( C_o \) is a constant background term, \( C_i \) reflect linear terms, \( C_{ij} \) interaction terms and \( D_i \) quadratic terms.

We consider a 32 dimensional problem with three versions of Eq. 10. In all cases \( C_o = C_i = 0 \) and \( D_i = 1 \). We consider a fully interacting version with \( C_{ij} = 1 \) for all \( i \) and \( j \), a partially interacting case such that \( C_{ij} = 1 \) if \((i-1) \text{ div } 8 = (j-1) \text{ div } 8\) and \( C_{ij} = 0 \) otherwise (in other words, the 32D parameter space is divided into 4 sub-spaces and there is interaction of parameters that lie within each sub-space but not between) and a non-interacting case where \( C_{ij} = 0 \) for all \( i \) and \( j \). The solution is simply that \( \theta_i = 0 \) for all \( i \).

We apply the NA to all three cases. We also apply the NA to one of the 8D sub-spaces of the partially interacting case. This illustrates the potential for speed up of the search mechanism in the Divide and Conquer approach that we propose.

In all cases we choose an initial set of models, \( n_i = 1024 \) and then set \( n_s = 32 \) and \( n_r = 16 \). Fig. 2 shows the results for the four cases showing the misfit reduction. The three 32D cases (Fig. 2a-f) all have very similar behaviour. The interaction terms in Eq. 10 add to the range of the response function. However, the convergence is pretty much independent of the interaction. For the 8D case, representing a subvolume of the partially interacting case, Fig 2g-h show that convergence is much faster. Much less time is spent searching away from the optimal solution due to the mixing of parameters between the four subspaces of \( \theta \).

The reason that NA is slow here is because the parameter space is too large. Sambridge (1999a) reported that we should generate at least \( 2^n \) models for an n-dimensional problem to saturate the space. If we don't build this number of models then the Voronoi cells will span the parameter space in one or several dimensions and there is no adaptation taking place. NA is therefore very resource hungry in large dimensional problems. However, it is very useful if we can subdivide the problem as in Fig. 2g-h.

2.5 Divide and Conquer

In many cases, we may consider that the misfit and associated parameter space can be subdivided. The misfit itself is already a summation of misfits from different regions (both as production and seismic misfits). It is intuitive that the control of those regional misfits may be obtained by localised changes to the flow properties. Further, the misfit of a single well or seismic region could be further subdivided depending on parameter interactions. For example, if a misfit component can be split into the sum of two separate misfits, each the function of distinct groups of parameters, the volume of the parameter space can be subdivided and searched separately. This is what we did in Fig. 2g-h. This then avoids situations where we generate redundant parameter combinations and the search is much more efficient. We call this approach Divide and Conquer and discuss it in detail in Sedighi and Stephen (2010). The misfit can then be thought of as:

\[ M(\theta) = M_1(\theta_1, ..., \theta_d) + M_2(\theta_{d+1}, ..., \theta_e) + \cdots + M_k(\theta_{d+1}, ..., \theta_e) \] (11)

where \( \theta_k \) is the vector of parameters, \( (\theta_{d+1}, ..., \theta_e) \) and represents a subset of \( \theta \). \( M_k \) is the sub-misfit that depends on it.

The key to this approach is the identification of the separate parameter spaces. This is the “Divide” step and can be made in a simple way where misfits can be separated regionally. A sensitivity matrix is calculated by examining the variation in misfit at various locations (either by well or by seismic misfit) as parameters are changed one at a time. Large changes can be seen in some locations in the field with small ones elsewhere. If a parameter has no effect on a particular misfit then we separate out the misfits. This approach is similar to the simplest of experimental design methods such as Plackett-Burman and is relatively cheap.

An alternative approach may be used that is more detailed using a proxy model to interrogate
the misfit for functional parameter interactions (Fig. 1). A quadratic proxy model was used previously (Sedighi and Stephen 2010) which is equivalent to Equation 10 and we use this here. This approach can become somewhat expensive to compute even in cases where only the main coefficients are required. We use this approach as a secondary analysis once the main subdivisions have been found and also remove parameters from the search completely when coefficients are small overall. Sensitivities can be calculated to identify parameter effects, as is often done with gradient based methods (e.g. Lépine, et al., 1999, Dong and Oliver, 2003, Gosselin et al., 2003, Mezghani, et al., 2004) or the Adjoint method (Wu et al., 1999). More recently, decomposition of the misfit has been proposed to help calculate the sensitivities to great effect (Ding and McKee, 2013). This is somewhat harder with seismic data however and we adopt a pragmatic approach here.

The second part of the problem is to “Conquer” by applying a “parallel” stochastic algorithm such that the optimization algorithm searches the sub-volume based on its distinct misfit. That is we calculate the individual misfits identified in Equation 10 once the simulations have been run. These misfits are used in steps 2 and 3 of Section 2.3 such that only $\theta_k$ is updated when $M_k$ is used. When all the parameters have been updated they are combined and new models can be run. We note that we could simply update the parameter values to reduce the misfit in one sub-space at a time in a serial approach. The disadvantage of this is that more simulations are required. Our aim is to simultaneously update the whole model.

For the NA, we adapt the algorithm summarised in Section 2.4. Step 1 is carried out as before. After the division step, the subdivisions of the parameter space given in Equation 10 are assumed to be known. After Step 1 we identify the sub-volumes of the parameter space as in Equation 11. The first sub-volume is assessed and its misfits are used to rank models as in Step 2. Steps 3 and 4 to select a sub-set of new parameter values. Steps 2 to 4 are repeated for each remaining sub-volume of the parameter space until new values of all parameters are created for each model in the new generation. Steps 5 is carried out to complete the iteration and then Step 6 closes the loop if necessary.

Fig. 3 illustrates how the search space is reduced in a 2D case. We can estimate the average rate of reduction of the parameter space in the case of Equation 10 or any other response surface that has a single solution. At the first iteration, $n_0$ models are selected stochastically. The area of the circle in Fig. 3 is on average $n_0/n_1$ times the initial search area as we select only the best $n_0$ models. After the second generation of models have been generated, the circle contains $n_0+n_s$ models so the subsequent search space will be reduced by a factor of $n_0/(n_0+n_s)$. This ratio of reduction is true on average for any dimensionality of problem. If we divide the 2D parameter space into $\theta_1$ and $\theta_2$ and then rank models based on those individual sub-spaces, each subspace reduces in size by $n_0/n_0$ initially and then $n_0/(n_0+n_s)$ for each subsequent iteration. However the total search area at the next iteration in the 2D space is the intersect of the separated sub-spaces (this is a rectangle in Fig. 3, due to randomisation) and reduces by $(n_0/(n_0+n_s))^2$ compared to the initial search space size. Thus for a p dimensional search space overall with q subspaces the rate of reduction in the search space volume from dividing is on average $(n_0/(n_0+n_s))^{p-1}$ per iteration compared to the conventional approach of searching the whole parameter space. We would therefore expect the divide and conquer approach to be $((n_0+n_s)/n_0)^{p-1}$ faster.

2.6 Posterior analysis

Finally, one of the advantages of generating multiple models is that the parameter distributions can be obtained based on the sample of models generated. Sambridge (1999b) showed that the sampling density of models within the parameter space is proportional to the probability distribution. This enables the Neighbourhood Algorithm with Bayes to be used to generate 1D marginals, the parameter distributions.

In a Bayesian framework and for a uniform prior probability $p(m)$ of the model data, the posterior probability density (PPD) of a model, $m$, given the data, $d$, is:

$$p(m|d) \propto e^{-M/2}$$

where $M$ is the misfit in Equation 9 and the exponential term is the likelihood. The 1D marginal distribution function of parameter $m$ is
\[ MDF(m_i) = \int_{-\infty}^{\infty} \prod_{j \neq m_i} p(m|d) dm_j = \int_{-\infty}^{\infty} e^{-\frac{M_2}{2}} \prod_{j \neq i} dm_j / \int_{-\infty}^{\infty} e^{-\frac{M_1}{2}} \prod_{j \neq i} dm_j \]  

This integral may be calculated using Markov Chain Monte Carlo methods and the results of history matching using the NA sampling algorithm, which we call NAB (Sambridge 1999b). The algorithm works very efficiently because the NA samples models in the parameter space with a density proportional to probability.

If we assume the misfit can be partitioned into two sub-misfits, each dependent on different parameters:

\[ M(\theta) = M_1(\theta_1) + M_2(\theta_2) \]  

where \( \theta_1 \) contains \( m_i \) for \( 1 < i \leq q \) and \( \theta_2 \) contains \( m_i \) for \( q+1 < i \leq n \) and \( n \) is the dimension of \( m \) then for \( i \leq q \) the marginal becomes:

\[ MDF(m_i) = \int_{-\infty}^{\infty} e^{-\frac{M_1}{2}} \prod_{j \neq i} dm_j / \int_{-\infty}^{\infty} e^{-\frac{M_1}{2}} \prod_{j \neq i} dm_j \]  

such that the second misfit terms cancel. Thus the integral can be simplified following misfit and parameter space partitioning. This calculation only requires posterior information from the parameter subspace containing the parameter.

3. Field application

In the field application below, history matching was split into:

- Analysis to determine which regions of the model were matching worst according to production data
- Analysis of interactions between the regions for obvious division
- Experimental design analysis to determine whether any parameters could be removed from analysis or whether further division is possible
- Multi-dimensional stochastic history matching sampling divided parameter spaces compared to full sampling. Two cases were considered (a) the sampling of the full dimensional space and (b) sampling of the partitioned ("divided") parameter space. We also consider two sampling rates with \( n_s = 64 \) and \( n_s = 16 \).
- NAB analysis

3.1 Field summary

The Nelson field is an undersaturated oil field located in blocks 22/11, 22/6a and 22/12a in the UK Central North Sea. The first exploration well was drilled in 1967, 3D seismic data first acquired in 1985 and led to discovery in 1988. The first production began in 1994 and 27 production wells were drilled up to 2000. The original oil in place was estimated at 126 million cubic metres and up to the end of 2000, 46 million cubic metres had been produced from the field (UK DTI, 2009). The production drive is aquifer supported coupled with water injection from 4 injection wells in the edge of the reservoir. The reservoir sands in Nelson are turbidities with excellent reservoir quality with average net-to-gross ratio of 70%, average porosity of 23% and permeability ranging from 200 to 1700 miliarcies. Geologically there are three distinct layers in Nelson separated mainly by shale. Each layer has a channelized characterization and the amount of shale and distribution within and between the channels are the main uncertainty (Gill et al. 2012) and so horizontal and vertical permeability as well as net-to-gross ratio have been modified. The combination of these parameters in the reservoir will define shale amount and distribution and control water movement in reservoir (for details see Kazemi 2011).

3D seismic monitor surveys of Nelson have been carried out on a three year basis since production began. With a baseline survey in 1990, 4D seismic attribute maps have been derived on
the reservoir level (Boyd-Gorst et al. 2001; McInally et al. 2003). We use amplitude data which has been phase shifted by 90 degrees to improve the tie between seismic polarity and lithology. This will modify the seismic trace in such a way that a peak or trough in a zero phase trace is changed to a zero-crossing to produce a pseudo-impedance equivalent to coloured inversion (Lancaster and Whitcombe 2000). This 3D cube is then converted to attribute maps using root mean square (RMS) averages of the time trace calculated between the horizons of the reservoir intervals. The difference of these attributes is used as the seismic 4D map.

History matching of Nelson in this paper follows the work by Kazemi et al. (2011) and Stephen and Kazemi, (2014). The same base case simulation model and data set have been used in this work. Predictions of water and oil rates from the 27 production wells were analysed and the 9 worst matched wells were selected. We used the same parameterization scheme (Fig. 4) detailed in Kazemi et al. (2011) and Stephen and Kazemi (2014) such that pilot points with Kriging (de Marsily et al. 1984) were used to control changes to horizontal and vertical permeability as well as net-to-gross ratio around the wells. The pilot points represent localized control points, where changes are made explicitly using multipliers. These changes are propagated laterally using Kriging. The range of the variogram used in the Kriging is equivalent to the pilot point separation. We group pilot points around each well so that the same changes are made at that location. This avoids a “target” response, spreading the change further. Net-to-gross ratio and the horizontal and vertical permeabilities are varied independently. Several of the wells were completed through two reservoir intervals and in that case separate pilot point groups are used for each interval. For those wells that complete in a single interval, changes are only made in that interval. Each well represents a 3D or a 6D optimization problem for history matching. In total there are 42 free variables. Please see the references for full details.

We have studied the Nelson field for history matching with time-lapse seismic previously. Stephen et al. 2009 investigated the precision of streamline simulation in the context of history matching and reported that better results were obtained if the model errors were calibrated in the misfit. Kazemi et al. 2011 used repeatability data to assess the data errors in 4D seismic and Stephen and Kazemi (2014) removed areas of high uncertainty to improve the matching process. Several options for controlling changes to the model during history matching were explored in Kazemi and Stephen, 2012. A pilot point approach was examined and contrasted with a facies based approach. The latter was based on seismically derived facies with fixed location but variable flow and storage properties. The pilot point approach gave more flexibility and better results were obtained despite having a larger dimension of the parameter space. The location of the pilot points were investigated using streamlines (Kazemi and Stephen, 2013). Understanding where the fluid flows greatly helps target those regions with most influence and a similar approach has been taken elsewhere (Gervais and Le Ravalec, 2018).

In previous studies, all parameters were updated simultaneously assuming that they could not be deconvolved. We next determine whether or not each well region can be treated separately and whether or not parameters from one region affect others.

### 3.2 Dividing the parameter space

The first step of history matching consisted of a heuristic analysis to determine which parameters affected the misfits of the 9 targeted wells. Each parameter was varied individually by an order of magnitude either way so that permeability or net-to-gross ratio was increased and decreased by a factor of 10. The effect of changing that parameter was then quantified by the range of misfit variation for each well. The results are tabulated as a similarity matrix in Table 1. Large numbers in each box mean that the parameter indicated in the first three columns of each row had a significant effect. The cells are colour coded so that red indicates large effect and white negligible effect.

Misfit variations are non-zero because a streamline simulator (Schlumberger, 2014) was used which was optimized for speed. The streamlines were generated randomly and changes to properties anywhere can generate slight alterations to flow throughout the field. The effect of these random variations on flow and on history matching have been discussed previously in terms of a model error (Stephen et al. 2009). We did not treat the model error explicitly though we did could calibrate it. Fig. 5 shows a cross plot of oil production rate misfits versus corresponding misfits for water production rate for each case in Table 1. At high values of misfit the two properties correlate and lie on a straight line of unit slope. This line is slightly offset from y=x because of differences in the magnitude of the data errors in oil and water rates. Some misfits diverge from this straight line at high values and this divergence is relatively greater at lower values. In the simulations we controlled flow by liquid rate.
which means that this should be perfectly matched especially given that none of the wells switch to pressure control. Differences between model and observed values for the production rate of oil should have been identical to differences for water rates. Fig. 5 shows that this is not the case due to model errors. The total liquid rate is the same for each well but individual oil and water rates do not sum to this value. We found that the model errors cause severe divergence from $y=x$ in the cross plot for misfits of around 300 for the water rate and 800 for the oil rate misfits. We deduced that this is the cutoff at which we can be sure that the influence of changing a parameter has negligible effect on a misfit. We have colour coded the cells in Table 1 to indicate those near this cutoff in yellow.

Table 1 therefore indicates that several of the parameters affect misfits outside of the geographic region to which they are applied (indicated with solid black borders for cells). We can therefore group those parameters that affect multiple wells. We use solid red borders on the table cells to indicate groups of misfits. The misfit space is now four dimensional consisting of Misfit Group 1: a 12D sub-volume to the north; Misfit Group 2 an 18D sub-volume in the centre; Misfit Group 3 a 3D sub-volume to the south, and Misfit Group 4 a 9D sub-volume to the west. The dimensionality can be inferred from Table 1.

The second step of history matching consisted of an experimental design analysis to determine whether any of the 42 parameters could be removed from the parameter set to further reduce the problem dimensionality. A Central Composite Design was applied to maximize the amount of information that can be gained from the minimum number of models. For Misfit Group 2 only 549 models were required whereas for Misfit Groups 1 and 4 it was 281 and 149 models respectively. The 3D case was not analysed here as this problem is small enough already. The seismic misfit for that region was also calculated and added to the production misfit. Seismic regions are shown in Fig. 4.

Fig. 6 shows the coefficients of the response function indicating the most important parameters. In each case a number of parameters were found to make negligible difference to the response surface such that the correlation coefficient ($R^2$) with the true misfits was 0.95. These were then removed from the parameter space where possible. The parameters that were removed could be inferred from Table 1. For Group 1, the pilot points in the 2nd interval near P1 had little effect in Table 1 and the coefficients were very small in the proxy model. Similarly, the pilot points in the 2nd interval near P4 and also the 2nd interval near P5 could be ignored in Misfit Group 2. In Misfit Group 4, parameters 1, 4 and 9 were removed.

We found that all three sub-volumes (Groups 1, 2 and 4) could be treated as an independent 9D problem so that the total parameter space became 30D. We chose not to subdivide these volumes further using the proxy model approach as described in Sedighi and Stephen (2010). There was only a small number of remaining parameters which were not part of the main group in this case.

### 3.3 Conquering the unknowns

History matching was carried out using the neighbourhood and genetic algorithms. The 30 parameters were sampled randomly to initialize the population with $n_d$ models. For each iteration of the full parameter space (“Full NA”) the total misfit was used to rank models and in the process of generating new models the whole parameter space was divided into Voronoi cells. With Divide and Conquer we treated one sub-volume of parameter space at a time within the iteration. Models were ranked according the regional misfit (which included the local well and seismic misfit) and then the new models were generated. In this case the voronoi cells of these models were constructed in the sub-volume only and new parameters were chosen. This was repeated for the other regions of the parameter space and the complete set of parameters was recombined to generate a new set of models. These were then forward propagated to obtain misfits for that iteration and the process was repeated.

In this study we considered two sampling strategies for both the Full NA and for Divide and Conquer. For moderate sampling we set $n_d=1024$, $n_s=64$ and $n_i=32$. We also considered a low sampling case where $n_d=256$, $n_s=16$ and $n_i=8$. The total number of models used is summarized in Table 2. The results of history matching are shown in Fig. 7 and Fig. 8. Fig. 7 shows a plot of misfit convergence as we generated more models. Convergence is obvious for the Divide and Conquer cases. Both moderate and low sampling decreased the misfit to a minimum, generally. Convergence was not observed in the Full NA case. The reason for this was that too many similar models were generated from the 30D parameter space and there was redundancy with many models providing very little new information. In Fig. 8, individual parameter convergence is shown for the Region 1 of Fig. 4. A 3 by 3 grid of parameters for each of the 4 history matching runs. Again both Divide and Conquer runs show good convergence.
The convergence for Divide and Conquer is compared to a history matching run where only
the parameters associated with Region 1 were varied (Fig. 9). The overall pattern is very similar and
the Divide and Conquer approach gave the same results compared to breaking the problem down and
running models separately on each region. To further analyse the problem we set up a version of
Equation 10 with that was used in the Fig. 2 for the 9 dimensional case except that the solutions have
been shifted from zero to the values of the parameters estimated for the Nelson case (Fig. 8 and 9).
We also added a constant term to the response surface so that zero response is not possible. The
same NA parameters ($n_1$, $n_2$, $n_3$) were used as in the Nelson cases). This is directly equivalent to what
we see in the misfit response in history matching as we rarely removed all errors due to limitations of
the modelling approach and through data errors. The purpose here was to show the kind of evolution
that occurred in a very simple response surface and compare it to the field case. Obviously the
absolute values of the response surface were different but the evolution was very similar (Fig. 10).
Most of the parameters evolved in a similar way in the two cases. This suggests a relatively simple
response surface for the field case. There are some differences possibly because some of the
parameters have less effect on the misfit in the field case. In the synthetic example they all had the
same influence. Thus several iterations were required before all parameters take effect in the field
case.

We compare the resulting production data for water rates for four of the main wells in Fig. 11.
It is clear that the convergence lead to better matched models in most of the cases. Wells P1 and P4
are both significantly better matched. P7 is quite well matched although the model predicts an early
build up of water production. P8 is better matched but there remains some room for improvement. The
Divide and Conquer cases are all better than the full history matching case except for P8 when they
are equivalent. The improvement to the 4D impedance map can be seen in Fig. 12. There are
improvements to both the full NA run and the Divide and Conquer cases. Overall the change is better
in the latter case.

We applied a posterior analysis (NAB) in this case to determine whether the uncertainty of the
parameters had decreased significantly as a function of the choice of parameterization and
optimization scheme. The posterior analysis (Sambridge 1999b) estimates 1D marginals, (i.e.
probability distributions), by converting misfits into model probabilities in a Bayesian framework and
using Markov Chain Monte Carlo (MCMC) methods for integration to estimate the posterior probability
density function. Fig. 13 shows the results for the four most important parameters. In each case we
performed three analyses. NAB was applied such that the full 3D parameter space was resampled
following the full NA history matching run. We find that the uncertainty barely changes from the prior
uncertainty which was a "top hat function" (uniform probability within a range, zero elsewhere) for two
of the parameters. The remaining two show reduction in uncertainty. During the Full NA run,
convergence was not obtained and so it was little better than a random sample. There was very little
improvement to the field model and little reduction of uncertainty.

On the other hand, if we know that the misfit can be partitioned by sub-volumes of the
parameter space then the MCMC resampling of NAB can be confined to sampling each sub-volume
separately based on that region’s misfits. This can be used to eliminate a lot of noise in the resampling
procedure along with undersampling and duplication. We did this using the results of the full NA but
resampled using the regional misfits to get the green lines in Fig. 13. All four parameters appear to
have reduced uncertainty with two showing a Gaussian-like shape, one is bimodal and one suggests
that the lower limit of the parameter is optimal.

The same partitioned NAB analysis can also be applied to the results of the Divide and
Conquer history matching run. The black lines in Fig. 13 also show reduced uncertainty and reflect
more closely the convergence observed in Figures 7 and 8. The plots are different from the analysis of
the full NA run however. Fig. 13a suggests a different solution is probable, closer to the upper limit on
the parameter. Fig. 13b suggests that bi-modality is less likely. Fig. 13c shows a shift of the peak. Only
Fig. 13d suggests a similar result to the analysis of the full run with a partitioned parameter space. The
differences in the three NAB analyses are not surprising. The Full NA run barely sees any
convergence. The sampling is almost random. Partitioning the search space in the analysis removes
many of the parameter values with models that simply don’t give a good match. However, there is still
insufficient sampling for the NAB to approximate the true PPD. Apart from that, the non-uniqueness of
the problem means that some models may appear to be better matching in the early stage of the Full
NA run. Later these models may be rejected as convergence towards the global solution begins. Thus
the peaks in the distribution curves are incorrect compared to the NA with Divide and Conquer.
4. Discussion

History matching using stochastic methods can become excessively demanding as we increase the number of parameters that we might want to include in the process. Assisted methods of history matching are therefore essential to determine which of those parameters should be included. Often some parameters can be ignored. It is therefore impossible to carry out fully automated history matching without some understanding of the model behaviour and the impact that properties have on improving the match.

We used a streamline simulator (Schlumberger, 2014), which could be run in a few minutes, rather than a finite difference equivalent which ran in just under one hour. The consequence was that we experienced model error as reported in (Stephen et al. 2009). We could estimate the degree of model error using Fig. 5, for example, and also deduce that the seismic misfits were also affected to a similar degree. A better calibration involves the running of the finite difference simulator although that adds computational costs. In this case we verified that the levels of noise and that the same responses were seen for dominant parameters as shown in Table 1. A double advantage of the streamline simulator is that we can visualize the regions affecting the wells and the seismic misfit.

Experimental design is a good way of identifying parameter importance and can be used carefully even in large dimensional cases. However, it is often better to reduce the parameter space of a given analysis by sub-dividing it first. Then we can also improve the way we carry out history matching using stochastic methods.

History matching using stochastic methods can be dramatically improved if we can identify that the problem can be divided into sub-volumes of the parameter space with associated misfits. In that case sub-volumes can be searched much more efficiently. In some cases it is intuitively obvious where the separations can be made. For example the parameters in Misfit Region 3 (Fig. 4) are unlikely to influence the misfits in Misfit Region 1. However, the separation of Misfit Regions 1 and 2 require checking and verification.

In total we used 1896 models to obtain convergence in this apparently 42D problem. We used one model to identify which wells had the worst misfit, 84 models to perform one at a time parameter changes, 979 models during the experimental design to reduce the parameter space further and the 982 models for history matching. We avoid running additional models by performing the experimental design on the three sub-volumes of the parameter space in parallel. We have also found that the experimental design models actually make a good starting point for NA thus saving the n models. These savings could result in a good convergence and analysis using just 1260 models.

We compare the above results with the case where we sampled the full 42D parameter space with an initial set of 1024 models followed by 982 for history matching and observe that this case performed hardly any better than a random search.

In the method section we wrote that the NA method was perhaps better for cases where there are multiple minima (of equivalent importance) whereas the GA algorithm is better suited to find convergence. We do not report it here but we find that the GA is equally amenable to “Divide and Conquer” and in fact other similar methods such as Particle Swarm Optimisation would benefit. We may also ask whether convergence does in fact matter at all. For NAB analysis the aim is to develop sufficient models to detect the distribution about the best model or models. If the misfit response surface is very complex then we may simply have to make do with a good sample set.

The situation where there are field wide parameters should also be addressed or where there is effectively a chain of connectivity between potential regions that cannot be broken down. The former might include properties such as relative permeability. Often jone relative permeability curve only is available for the whole field or perhaps a small set are used for each rock type or facies but these are also applied throughout the field. It may be tempting to localise these relative permeability curves this is possible in our approach. However, this adds complexity to the problem and the changes to relative permeability may become somewhat arbitrary and merely reflect the non-uniqueness. It may be better to attempt to find the relative permeability curves first with a global update to the properties and this is in keeping with the traditional approach of the stratigraphic method (Williams et al. 1998) and has been applied with a Genetic Algorithm previously (Cheng et al. 2008). In the second of these cases, the wells may be so close that the parameters around each well also affect at least one neighbour. In this case a Divide step not may be possible. At this point care is needed and it may be most efficient to adopt an iterative approach which may speed up convergence compared to a full parameter space search.

At the start of the project we stated that we would update the simulation model rather than carry out Big Loop history matching, which requires upscaling in this case. As in all history matching
cases, there is always the possibility that the model we start with can never be updated in a sensible
or useful manner. In this work we set limitations to the potential changes that the model may see by
restricting the multipliers of the pilot points. For example permeability should not increase by a factor
of 10 (a factor of two or three may be more appropriate in some cases). This can then eliminate the
possibility of geologically inconsistent models being generated. If a match cannot be found then an
obvious course of action is to revisit the original model or attempt to include changes in a Big Loop
manner. Similarly we could relax the restrictions on model parameters somewhat to detect what will
give a good match. If we are unhappy with the results it is most likely that a completely new model is
required.

The method presented here can be tried in a general setting. We recognize, however that all
models are unique and it may not be possible to apply the Divide and Conquer approach therefore, in
a number of cases. It requires that the problem can be broken down spatially both in terms of the
misfits and the parameterization. If we pick any location in a reservoir model, we could assign a misfit
to it (either as a well or as a point on a 4D seismic data map, for example). That misfit has a zone of
influence in terms of how it can be affected by changes to the flow and storage properties of the grid
cells nearby. This zone is like a drainage zone for a well but is larger in extent for actual wells. The
zones of influence overlap for neighbouring locations, especially for seismic data misfits. The key is to
be able to identify these zones of influence. Some work is required to assess how we break down the
problem to assess whether the misfits and the parameter space can be deconvolved therefore. Often
this step may be carried out by sensitivity studies by use of experimental design so the step does not
always require extra work. Such a step is recommended for both manual and so called automatic
history matching to reduce the parameter space. In the work presented here we were fortunate to be
able to separate out the zones of influence spatially based on the need to improve the match of a sub-
volume of the reservoir.

5. Conclusions

• Large dimensional problems require numerous simulations and stochastic methods can be
resource hungry while deconvolution of the misfits and parameters into independent subspaces
can vastly improve the efficiency of stochastic search algorithms.
• Sensitivity analysis is a useful process for understanding the relationship between parameters and
various misfits and ultimately informs on which ones should be changed to improve the history
match. The process also informs on how parameters and misfits can be deconvolved.
• With suitable regionalization of the misfits and sub-division of the parameters, the history matching
problem was reduced to four independent problems, one was three dimensional and the other
three nine dimensional. The problem could then be solved with many fewer simulation runs,
equivalent to a single nine dimensional problem.
• We have shown that a 42 dimensional problem can be reduced to 30 dimensions using
experimental design.
• The Neighbourhood Algorithm was amenable to “Divide and Conquer”.

6. Acknowledgments

We would like to thank the Nelson Field partners (Esso Exploration and Production UK Ltd.,
Idemitsu E&P UK Ltd., Premier Oil ONS Ltd., Shell UK Ltd., and Total E&P UK Ltd.) for providing the
data and permission to publish the results. BP, ConocoPhilips, Shell and StatoilHydro are thanked for
funding of this work as part of Phase III of the Seismic History Matching Project. We thank
Schlumberger Geoquest for use of their software. Malcolm Sambridge is thanked for use of the
Neighbourhood Algorithm.

References

Kalman Filter in Reservoir Engineering—a Review.” SPE 117274, SPE Journal 14(3), pp. 393-
412.


Table 1. Matrix of interactions between changes made to parameters at specific locations and corresponding misfits for wells. The first column indicates the pilot point group associated with each well in Figure 4. The second column indicates the interval and the third column is the petrophysical property for the pilot points. The fourth column indicates the misfit group that each parameter is ultimately assigned to and the fifth column indicates an index for later analysis. One parameter was changed at a time on each row and the corresponding misfit for each well is shown under columns P1-P9. The numbers indicate the magnitude of the seismic and production misfit and boxes are colour coded so that red is hot and white cold. The outer red boxes indicate groups of parameters that cannot be subdivided from these results. The inner black boxes indicate parameters local to the wells.

<table>
<thead>
<tr>
<th>Well</th>
<th>Interval</th>
<th>Property</th>
<th>Misfit group</th>
<th>Index by group</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Kx</td>
<td>1</td>
<td>5359</td>
<td>11267 2844</td>
</tr>
<tr>
<td>P1</td>
<td>Kx</td>
<td>1</td>
<td>2355</td>
<td>11372 509</td>
</tr>
<tr>
<td>P1</td>
<td>NTG</td>
<td>1</td>
<td>4032</td>
<td>566 14</td>
</tr>
<tr>
<td>P1</td>
<td>Kx</td>
<td>2</td>
<td>62</td>
<td>620 507</td>
</tr>
<tr>
<td>P1</td>
<td>Kz</td>
<td>1</td>
<td>48</td>
<td>625 147</td>
</tr>
<tr>
<td>P1</td>
<td>NTG</td>
<td>2</td>
<td>367</td>
<td>135 381</td>
</tr>
<tr>
<td>P2</td>
<td>Kx</td>
<td>1</td>
<td>5111</td>
<td>934 1790</td>
</tr>
<tr>
<td>P2</td>
<td>Kz</td>
<td>1</td>
<td>2795</td>
<td>24800 529</td>
</tr>
<tr>
<td>P2</td>
<td>NTG</td>
<td>1</td>
<td>827</td>
<td>10185 621</td>
</tr>
<tr>
<td>P3</td>
<td>Kx</td>
<td>1</td>
<td>930</td>
<td>245 7416</td>
</tr>
<tr>
<td>P3</td>
<td>Kx</td>
<td>1</td>
<td>756</td>
<td>94 12257</td>
</tr>
<tr>
<td>P3</td>
<td>NTG</td>
<td>1</td>
<td>2569</td>
<td>538 5403</td>
</tr>
<tr>
<td>P4</td>
<td>Kx</td>
<td>2</td>
<td>167</td>
<td>255 186</td>
</tr>
<tr>
<td>P4</td>
<td>Kx</td>
<td>2</td>
<td>409</td>
<td>69 550</td>
</tr>
<tr>
<td>P4</td>
<td>NTG</td>
<td>2</td>
<td>260</td>
<td>45 139</td>
</tr>
<tr>
<td>P4</td>
<td>Kx</td>
<td>2</td>
<td>335</td>
<td>515 31</td>
</tr>
<tr>
<td>P4</td>
<td>Kz</td>
<td>2</td>
<td>33</td>
<td>309 8</td>
</tr>
<tr>
<td>P4</td>
<td>NTG</td>
<td>2</td>
<td>55</td>
<td>11 145</td>
</tr>
<tr>
<td>P5</td>
<td>Kx</td>
<td>2</td>
<td>332</td>
<td>481 219</td>
</tr>
<tr>
<td>P5</td>
<td>Kz</td>
<td>2</td>
<td>394</td>
<td>565 142</td>
</tr>
<tr>
<td>P5</td>
<td>NTG</td>
<td>2</td>
<td>483</td>
<td>37 29</td>
</tr>
<tr>
<td>P5</td>
<td>Kx</td>
<td>2</td>
<td>32</td>
<td>648 218</td>
</tr>
<tr>
<td>P5</td>
<td>Kz</td>
<td>2</td>
<td>33</td>
<td>309 8</td>
</tr>
<tr>
<td>P5</td>
<td>NTG</td>
<td>2</td>
<td>55</td>
<td>11 145</td>
</tr>
<tr>
<td>P6</td>
<td>Kx</td>
<td>2</td>
<td>167</td>
<td>255 186</td>
</tr>
<tr>
<td>P6</td>
<td>Kx</td>
<td>2</td>
<td>409</td>
<td>69 550</td>
</tr>
<tr>
<td>P6</td>
<td>NTG</td>
<td>2</td>
<td>260</td>
<td>45 139</td>
</tr>
<tr>
<td>P6</td>
<td>Kx</td>
<td>2</td>
<td>335</td>
<td>515 31</td>
</tr>
<tr>
<td>P6</td>
<td>Kz</td>
<td>2</td>
<td>33</td>
<td>309 8</td>
</tr>
<tr>
<td>P6</td>
<td>NTG</td>
<td>2</td>
<td>55</td>
<td>11 145</td>
</tr>
<tr>
<td>P7</td>
<td>Kx</td>
<td>3</td>
<td>47</td>
<td>192 316</td>
</tr>
<tr>
<td>P7</td>
<td>Kx</td>
<td>3</td>
<td>130</td>
<td>8 80</td>
</tr>
<tr>
<td>P7</td>
<td>NTG</td>
<td>3</td>
<td>94</td>
<td>110 302</td>
</tr>
<tr>
<td>P8</td>
<td>Kx</td>
<td>4</td>
<td>208</td>
<td>609 131</td>
</tr>
<tr>
<td>P8</td>
<td>Kx</td>
<td>4</td>
<td>130</td>
<td>100 71</td>
</tr>
<tr>
<td>P8</td>
<td>NTG</td>
<td>4</td>
<td>97</td>
<td>254 134</td>
</tr>
<tr>
<td>P9</td>
<td>Kx</td>
<td>4</td>
<td>341</td>
<td>34 107</td>
</tr>
<tr>
<td>P9</td>
<td>Kx</td>
<td>4</td>
<td>140</td>
<td>23 73</td>
</tr>
<tr>
<td>P9</td>
<td>NTG</td>
<td>4</td>
<td>356</td>
<td>148 395</td>
</tr>
<tr>
<td>P9</td>
<td>Kz</td>
<td>4</td>
<td>34</td>
<td>417 32</td>
</tr>
<tr>
<td>P9</td>
<td>Kz</td>
<td>4</td>
<td>267</td>
<td>208 76</td>
</tr>
<tr>
<td>P9</td>
<td>NTG</td>
<td>4</td>
<td>111</td>
<td>117 239</td>
</tr>
</tbody>
</table>
Table 2: Summary of the number of models used at each stage

<table>
<thead>
<tr>
<th>WORKFLOW STAGE</th>
<th>LOW SAMPLING</th>
<th>MODERATE SAMPLING</th>
</tr>
</thead>
<tbody>
<tr>
<td>INITIALIZATION (RANDOM)</td>
<td>256</td>
<td>1024</td>
</tr>
<tr>
<td>PER ITERATION</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>PER NEIGHBOURHOOD</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>896</td>
<td>2304</td>
</tr>
</tbody>
</table>

=256+32 x 20  =1024 + 64 x 20

Figure 1. Divide and Conquer workflow within the history matching process (after Sedighi and Stephen, 2010). A proxy model is fitted to the misfits from the initialization step. Interaction terms in the misfit then identify which parameters combine and these are used then to identify sub-volumes. The parallel history matching loop computes a misfit for each parameter sub-volume and selects new parameter values based on the NA or GA approach. The simulation model is then updated using the whole parameter space and new predictions are made.
Figure 2. Evolution of $f(\theta)$ and the first parameter, $\theta_1$, as the NA evolves the search towards the minimum of Equation 10 for various dimensionalities and interactions in the equation: (a) and (b) 32D problem with full interaction, (c) and (d) 32D problem with interaction in 4 sub-volumes, (e) and (f) 32D problem with no interaction and (g) and (h) 8D problem with full interaction (equivalent to one sub-volume of (c) and (d)).
Figure 3. Example of parameter selection when we separate out the ranking of models according to the individual misfits. In this case Equation 10 is used with no interaction for a 2D case. 64 models are generated randomly (blue diamonds). The best 10 are selected based on Equation 10 (red squares). New models will be located within the circle. On the other hand if the best 10 are selected according to $\theta_1^2$ (green triangles), new values of $\theta_1$ will lie between the vertical dashed lines. Similarly new values of $\theta_2$ will be selected based on the lowest value of $\theta_2^2$ (purple dots) then the search space lies between the horizontal dashed lines. New models at the next iteration will be located within the red rectangle.
Figure 4. Map of the Nelson field showing initial oil saturation (yellow indicates the oil and blue the aquifer). Large cylinders indicate the wells used in history matching (P1-P9). Smaller cylinders indicate pilot points for modification of properties and these have the same colour as the associated well. Note that some pilot points are offset from the well to focus on flow regions identified by streamlines (Kazemi and Stephen, 2012). The saturation colours are semi-transparent outside of the regions identified for the seismic misfit. These regions also correlate with the identified Misfit Groups 1-4.
Figure 5. Cross plot of the absolute change in oil and water production rate misfits for the wells and parameter changes indicated in Table 1. The colours indicate the misfits for different wells (P1-P9).
Figure 6. Pareto charts showing major coefficients obtained for the polynomial proxy model following experimental design for each of the three biggest parameter groups (a) Group 1, (b) Group 2 and (c) Group 3. Red bars indicate the linear coefficients, Brown bars the quadratic terms and yellow the interaction terms. The labels on the y-axis indicate the parameter index for each group (as defined in Table 1, column 5).
Figure 7. Evolution of the total misfits (y-axis) for both Divide and Conquer and the full NA for both moderate and low level sampling. The x-axis shows the model number indicating the order of creation of the model. Models were generated in groups or generations for each iteration.
Figure 8. Evolution of the parameter values ($\log_{10}$ of multipliers of permeability and net-to-gross ratio) for Divide & Conquer and Full sampling for both moderate and low level sampling. This shows parameter group 1 only. Other parameter groups show a similar response. The response is very similar for the GA method.
Figure 9. Misfit and parameter $\log_{10}$ of multipliers of permeability and net-to-gross ratio as they evolve for the Nelson case where only the 10 parameters of Region 1 in Fig. 4 were updated. The top left figure shows total misfit and the other plots parameters 2 to 9.
Figure 10 shows the evolution of the response function calculated from Eq. 10 (top left) and the variation of parameters 2-9 (other plots) assuming a 2D problem. The solutions of this equation are estimated and set equivalent to those anticipated from Figures 8 and 9 for Region 1.
Figure 11. Water production matches for the History case, basecase and best models after history matching with the Divide and Conquer method and the full parameter space NA. The well locations are shown in Fig. 4. The numbers on the rate axis are deliberately hidden.
Figure 12. Maps of elastic impedance change from 1993 to 2000 for the observed data, base case and for the results of history matching with the Divide and Conquer method as well as Full NA. Red indicates water saturation increase, blue no change. The observed map contains noise.
Figure 13. Probability distribution curves following posterior analysis (NAB) for wells (a) P1, (b) P4, (c) P7 and (d) P8. Each graph shows the distribution for the four most important parameters as a function of the log\textsubscript{10} of the parameter multiplier. We compare NAB applied to Full NA (red lines), NAB applied to the separate parameter space misfits following full NA (green lines) and then NAB applied to the Divide and Conquer run (black lines).
<table>
<thead>
<tr>
<th>Well</th>
<th>Interval</th>
<th>Property</th>
<th>Misfit group</th>
<th>Property by group</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1</td>
<td>Kx</td>
<td>1</td>
<td>5359</td>
<td>11267</td>
<td>2844</td>
<td>25</td>
<td>167</td>
<td>98</td>
<td>335</td>
<td>88</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>1</td>
<td>Kz</td>
<td>2</td>
<td>2535</td>
<td>11372</td>
<td>509</td>
<td>126</td>
<td>242</td>
<td>190</td>
<td>78</td>
<td>42</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>1</td>
<td>NTG</td>
<td>3</td>
<td>4032</td>
<td>566</td>
<td>14</td>
<td>29</td>
<td>104</td>
<td>201</td>
<td>314</td>
<td>150</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>2</td>
<td>Kx</td>
<td>4</td>
<td>62</td>
<td>620</td>
<td>507</td>
<td>109</td>
<td>61</td>
<td>112</td>
<td>346</td>
<td>314</td>
<td>221</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>2</td>
<td>Kz</td>
<td>5</td>
<td>48</td>
<td>625</td>
<td>147</td>
<td>54</td>
<td>115</td>
<td>223</td>
<td>64</td>
<td>196</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>2</td>
<td>NTG</td>
<td>6</td>
<td>367</td>
<td>135</td>
<td>381</td>
<td>39</td>
<td>147</td>
<td>6</td>
<td>265</td>
<td>100</td>
<td>413</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>1</td>
<td>Kx</td>
<td>7</td>
<td>5111</td>
<td>934</td>
<td>1790</td>
<td>58</td>
<td>583</td>
<td>20</td>
<td>41</td>
<td>89</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>1</td>
<td>Kz</td>
<td>8</td>
<td>2793</td>
<td>24800</td>
<td>529</td>
<td>128</td>
<td>61</td>
<td>154</td>
<td>716</td>
<td>217</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>1</td>
<td>NTG</td>
<td>9</td>
<td>827</td>
<td>10185</td>
<td>631</td>
<td>59</td>
<td>130</td>
<td>172</td>
<td>725</td>
<td>75</td>
<td>347</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>1</td>
<td>Kx</td>
<td>10</td>
<td>930</td>
<td>246</td>
<td>7416</td>
<td>145</td>
<td>20</td>
<td>6</td>
<td>126</td>
<td>191</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>1</td>
<td>Kz</td>
<td>11</td>
<td>756</td>
<td>94</td>
<td>12257</td>
<td>80</td>
<td>6</td>
<td>9</td>
<td>169</td>
<td>162</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>1</td>
<td>NTG</td>
<td>12</td>
<td>2569</td>
<td>538</td>
<td>5403</td>
<td>238</td>
<td>133</td>
<td>196</td>
<td>529</td>
<td>97</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>1</td>
<td>Kx</td>
<td>2</td>
<td>167</td>
<td>255</td>
<td>188</td>
<td>239</td>
<td>14</td>
<td>889</td>
<td>252</td>
<td>423</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>1</td>
<td>Kz</td>
<td>2</td>
<td>409</td>
<td>69</td>
<td>550</td>
<td>15</td>
<td>158</td>
<td>2372</td>
<td>876</td>
<td>50</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>1</td>
<td>NTG</td>
<td>2</td>
<td>260</td>
<td>45</td>
<td>139</td>
<td>11</td>
<td>319</td>
<td>354</td>
<td>306</td>
<td>166</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>2</td>
<td>Kx</td>
<td>2</td>
<td>335</td>
<td>515</td>
<td>31</td>
<td>1491</td>
<td>1115</td>
<td>514</td>
<td>333</td>
<td>148</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>2</td>
<td>Kz</td>
<td>2</td>
<td>33</td>
<td>309</td>
<td>8</td>
<td>24444</td>
<td>374</td>
<td>262</td>
<td>608</td>
<td>147</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>2</td>
<td>NTG</td>
<td>2</td>
<td>55</td>
<td>11</td>
<td>145</td>
<td>14</td>
<td>104</td>
<td>386</td>
<td>295</td>
<td>37</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>1</td>
<td>Kx</td>
<td>2</td>
<td>332</td>
<td>481</td>
<td>219</td>
<td>102</td>
<td>510</td>
<td>167</td>
<td>483</td>
<td>196</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>1</td>
<td>Kz</td>
<td>2</td>
<td>394</td>
<td>565</td>
<td>142</td>
<td>541</td>
<td>353</td>
<td>160</td>
<td>32</td>
<td>89</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>1</td>
<td>NTG</td>
<td>2</td>
<td>483</td>
<td>37</td>
<td>29</td>
<td>155</td>
<td>149</td>
<td>19</td>
<td>170</td>
<td>227</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>2</td>
<td>Kx</td>
<td>2</td>
<td>32</td>
<td>648</td>
<td>218</td>
<td>67</td>
<td>306</td>
<td>7188</td>
<td>165</td>
<td>184</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>2</td>
<td>Kz</td>
<td>2</td>
<td>205</td>
<td>613</td>
<td>476</td>
<td>24</td>
<td>1991</td>
<td>8539</td>
<td>454</td>
<td>95</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>2</td>
<td>NTG</td>
<td>2</td>
<td>328</td>
<td>532</td>
<td>320</td>
<td>166</td>
<td>2387</td>
<td>176</td>
<td>443</td>
<td>17</td>
<td>383</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>1</td>
<td>Kx</td>
<td>2</td>
<td>72</td>
<td>386</td>
<td>32</td>
<td>192</td>
<td>314</td>
<td>303</td>
<td>26</td>
<td>251</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>1</td>
<td>Kz</td>
<td>2</td>
<td>314</td>
<td>287</td>
<td>200</td>
<td>79</td>
<td>1580</td>
<td>292</td>
<td>702</td>
<td>48</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>1</td>
<td>NTG</td>
<td>2</td>
<td>180</td>
<td>78</td>
<td>32</td>
<td>2</td>
<td>169</td>
<td>226</td>
<td>368</td>
<td>68</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>2</td>
<td>Kx</td>
<td>2</td>
<td>718</td>
<td>161</td>
<td>616</td>
<td>216</td>
<td>1641</td>
<td>2885</td>
<td>507</td>
<td>86</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>2</td>
<td>Kz</td>
<td>2</td>
<td>89</td>
<td>316</td>
<td>156</td>
<td>84</td>
<td>1464</td>
<td>162</td>
<td>25</td>
<td>27</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>2</td>
<td>NTG</td>
<td>2</td>
<td>117</td>
<td>257</td>
<td>263</td>
<td>5</td>
<td>365</td>
<td>657</td>
<td>932</td>
<td>40</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>1</td>
<td>Kx</td>
<td>3</td>
<td>47</td>
<td>192</td>
<td>316</td>
<td>189</td>
<td>149</td>
<td>16</td>
<td>17046</td>
<td>109</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>1</td>
<td>Kz</td>
<td>3</td>
<td>130</td>
<td>8</td>
<td>80</td>
<td>244</td>
<td>57</td>
<td>63</td>
<td>24062</td>
<td>277</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>1</td>
<td>NTG</td>
<td>3</td>
<td>94</td>
<td>110</td>
<td>302</td>
<td>43</td>
<td>148</td>
<td>298</td>
<td>1771</td>
<td>42</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>1</td>
<td>Kx</td>
<td>4</td>
<td>208</td>
<td>609</td>
<td>131</td>
<td>114</td>
<td>33</td>
<td>348</td>
<td>857</td>
<td>1104</td>
<td>444</td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>1</td>
<td>Kz</td>
<td>4</td>
<td>139</td>
<td>100</td>
<td>71</td>
<td>77</td>
<td>2</td>
<td>6</td>
<td>848</td>
<td>7123</td>
<td>1026</td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>1</td>
<td>NTG</td>
<td>4</td>
<td>97</td>
<td>254</td>
<td>134</td>
<td>25</td>
<td>84</td>
<td>22</td>
<td>186</td>
<td>2868</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>1</td>
<td>Kx</td>
<td>4</td>
<td>341</td>
<td>34</td>
<td>107</td>
<td>47</td>
<td>74</td>
<td>619</td>
<td>47</td>
<td>6268</td>
<td>2140</td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>1</td>
<td>Kz</td>
<td>4</td>
<td>140</td>
<td>23</td>
<td>73</td>
<td>53</td>
<td>163</td>
<td>684</td>
<td>76</td>
<td>4289</td>
<td>11584</td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>1</td>
<td>NTG</td>
<td>4</td>
<td>356</td>
<td>148</td>
<td>395</td>
<td>80</td>
<td>41</td>
<td>459</td>
<td>361</td>
<td>719</td>
<td>932</td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>2</td>
<td>Kx</td>
<td>4</td>
<td>34</td>
<td>417</td>
<td>32</td>
<td>48</td>
<td>1</td>
<td>134</td>
<td>336</td>
<td>2962</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>2</td>
<td>Kz</td>
<td>4</td>
<td>267</td>
<td>208</td>
<td>76</td>
<td>149</td>
<td>70</td>
<td>413</td>
<td>147</td>
<td>1550</td>
<td>232</td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>2</td>
<td>NTG</td>
<td>4</td>
<td>111</td>
<td>117</td>
<td>239</td>
<td>104</td>
<td>121</td>
<td>100</td>
<td>292</td>
<td>238</td>
<td>73</td>
<td></td>
</tr>
</tbody>
</table>
Assisted seismic history matching of the Nelson field: managing large numbers of unknowns by divide and conquer.

Karl D. Stephen  
Institute of Petroleum Engineering,  
School of Energy, Geoscience, Infrastructure and Society,  
(Heriot-Watt University,  
Riccarton,  
Edinburgh,  
Scotland, EH14 4AS,  
UK  
email: k.d.stephen@hw.ac.uk

**Highlights**

In this paper we show an application of a stochastic history matching approach applied to the Nelson field. The problem contains 42 parameters initially which can be very resource hungry computationally. We have used assisted history matching approaches to assessing how various parameters interact and used that to reduce the problem in an orderly fashion. The problem is shown to be reduced from something very difficult to solve to one that is manageable using a Divide and Conquer approach. The value of the approach is illustrated by applying it to a field case.