

Bayes linear strategies for matching hydrocarbon reservoir history

Peter S. Craig, Michael Goldstein, Allan H. Seheult and James A. Smith

Department of Mathematical Sciences, University of Durham, Durham DH1 3LE, England

1 Introduction

Hydrocarbon reservoir models are widely used to predict future oil, gas and water production and to give guidance for sensible reservoir management. The models are implemented as simulators which take as input a description of the geology and yield as output a simulated production history for the reservoir.

History matching concerns the problem of finding an input reservoir geology for which the simulator production output matches the observed production history for the reservoir which is being modelled. History matching is an important and time consuming component in the creation of reservoir models. Usually, this matching is done by trial and error by reservoir engineers, running the simulator with various choices of geology, to attempt to reproduce the history of the reservoir. Problems with this approach are that input and output dimensions may be very high, and each run of the simulator may be very expensive in CPU time, so that often, even with a large investment of time and effort, the quality of the final match will be unsatisfactory. In a recent study of a large reservoir, it took two man years to obtain a less than satisfactory match, with each simulator run taking up to twenty hours. Mattax and Dalton (1990) say

History matching can be time consuming, expensive, and frustrating, primarily because reservoir performance can be complex, responding to numerous interactions that, as a whole, may be difficult to comprehend. As a consequence, there has been considerable research on programs for inverse simulation or automatic history matching. The programs are rarely used, however, because for most studies automatic history matching is less efficient than manual matching.

In this paper, we develop Bayes linear methods for history matching. Our intentions are twofold. Firstly, we hope to make a practical contribution, by showing how the expert judgments of the reservoir engineer, in combination with Bayes linear analysis, may lead to improved history matching. Secondly, history matching is an example of a wide class of hard problems, namely high dimensional, ill-posed, non-linear inverse problems for functions which are expensive to evaluate, and so offers an interesting case study in how we may structure and exploit expert judgments to help us solve such problems.

Our account offers a general strategy for history matching. We discuss the essential components of the strategy, and offer suggestions for implementing each component. We demonstrate our approach with an example which, while comparatively simple, illustrates many of the issues involved in practical history matching.

2 History matching

Efficient management and prediction of reservoir production in the oil industry is crucially dependent on having a good mathematical model of the reservoir. Reservoir engineers, in conjunction with petrophysicists, geologists

and mathematicians, build a mathematical model of an operating reservoir. The model takes as data the physical description of the reservoir, including the geology, geometry and other parameters, such as porosities, permeabilities and relative permeabilities, which affect fluid flow. The model is a complex system of time dependent, non-linear, partial differential equations that describe oil, gas and water flows through a structural description (including well placement, faults and other geological features) of the reservoir over time. The model cannot be solved exactly, but an approximate numerical solution can be obtained using a reservoir simulator.

The simulator is a computer code which solves the model by including certain simplifying assumptions about the data. For example, in the real reservoir rock permeabilities are likely to be a varying function of position. In the simulator the physical region is represented by a grid, and it is often assumed that porosities and permeabilities are constant over each grid block. Small grid blocks increase the complexity of the problem, but more closely approximate the original model. It is common to run simulators with several tens of thousands of grid blocks. A single run of the simulator will usually be expensive both in CPU time and in monetary terms. The output is several time series, including production and pressures at each of possibly several wells.

A major problem is that we often have a very imprecise geological description, so that we don't know the values of the physical parameters to any degree of accuracy. The aim of history matching is to solve the inverse problem where we know aspects of reservoir behaviour, such as a history of production, pressures and other variables, over a historical period of time, and we try to deduce the geology which will cause the simulator output to match this history. After running the simulator, the output will be compared to historical data usually in the form of several time series (oil/gas/water production and pressures) from each of the production wells. Some aspects of the history may be very detailed, while others are sparse, and possibly themselves estimated. Further, unlike the corresponding simulator output, these data may not be smooth, and this raises questions as to what constitutes a good match. The engineer will then attempt to change the input data for the simulator in order to make the simulator output match the reservoir behaviour more closely.

History matching has been and will continue to be a major economic and scientific concern in the oil industry. There have been various one-shot attempts at history matching using regression-type approaches, but these have generally yielded infeasible solutions. Therefore in practice, history matching is done by trial and error; see for example the comparative study in Anterion (1988). The reservoir engineer, using experience and knowledge, runs the simulator several times for different choices of the values of the parameters until a satisfactory match is obtained, if at all. During this process, the engineer may also decide to occasionally change the model structure, by modifying the location or extent of a fault, or the alignment or layering of rock types. The match is validated by the geologists, who may suggest further modifications to the model. For a recent overview of history matching, see Mattax and Dalton (1990).

There are various motivations for formalising this process. Firstly, the time of the expert is very costly, so that we may gain directly by partially automating the process. Secondly, history matching is often at least partially unsuccessful, so that we may hope to improve the quality of the history match. Thirdly, systematically quantifying and comparing expert judgments with the performance of the simulator should lead to useful qualitative insights about reservoir performance. Finally, history matching is an example of a wide class of hard problems, to which similar approaches may be applied. While we shall describe our approach in the context of history matching, the methodology that we shall develop is potentially applicable for quite general computer experiments and in particular for sequential searches to solve general high dimensional inverse problems.

2.1 History matching as a computer experiment

Mathematically, history matching can be thought of as finding a zero of a multivariate time-varying function, namely the reservoir simulator. Thus, the problem can be viewed as one of design and analysis of a computer experiment; for an overview of this field, see Sacks, Welch, Mitchell, and Wynn (1989). The particular features which make this problem extremely challenging are: the function often has very high dimensional input and output; the input dimension may greatly exceed the output dimension, so that there may be many solutions; the function is very costly to evaluate at any input; and the simulator is just a model of the reservoir, so that the best that we can reasonably hope for is a “close” approximate match, and there may be no exact solutions.

The design problem is to choose, sequentially, the collection of input values at which to evaluate the function. This may be viewed as a “design for control” problem, in which we learn about a surface in order to identify the inputs which are necessary to produce a desired output. Design for control problems are known to be hard even for relatively simple, well specified linear models of moderate dimension. Realistically, therefore, we should not expect to produce formal solutions to the much harder problem in which each function evaluation tells us about (i) the location of the zero; (ii) the qualitative shape of the surface; (iii) the lack of fit between the reservoir and the reservoir simulator. Further, given the possibility of no solutions or a very large class of solutions, estimation of the zero given evaluation of a relatively small number of values of the function requires solution of a very high dimensional, ill-posed, non-linear inverse problem for which expert knowledge is crucial in identifying solutions which are acceptable in engineering terms; see O’Sullivan (1986) for a statistical overview of such problems, where a much simplified version of the history matching problem is used as an illustration. See also O’Hagan (1992) and Smith (1993) for a discussion of Bayesian numerical analysis.

2.2 Notation

In line with the above view, we may consider that we have a function $y = f(x)$, where the vector x is termed the geology input to the simulator, for example permeabilities and porosities, expressing aspects of the reservoir geology, and the vector y is termed the production output from the simulator, for example pressures and oil/gas/water rates at the wells, expressing the reservoir history. We write $y(x)$ when we need to explicitly display the dependence on x . The observed reservoir production history is y_h , and we are seeking x_h defined by

$$y_h = y(x_h). \tag{1}$$

It is possible that either there are several solutions, or, as the simulator is only an approximation to the reservoir and the actual history of the reservoir is recorded with error, that there may be no exact solutions to (1) on the simulator. To simplify our notation, we shall suppose throughout that the observed history, y_h , of the reservoir has been transformed to zero, so that x_h is the zero of y , i.e. $y(x_h) = 0$.

2.3 Example

The example we will use to illustrate the paper is one which was created by our industrial collaborator Scientific-Software Intercomp (SSI) as a test for their Adaptive History Matching (AHM) software. We received reservoir simulator control files which, once values of uncertain reservoir properties are supplied, enable the simulator to calculate the corresponding production. A special aspect of the example is that the production history we are trying to match is actually the simulated production for a known set of values for the uncertain properties. While

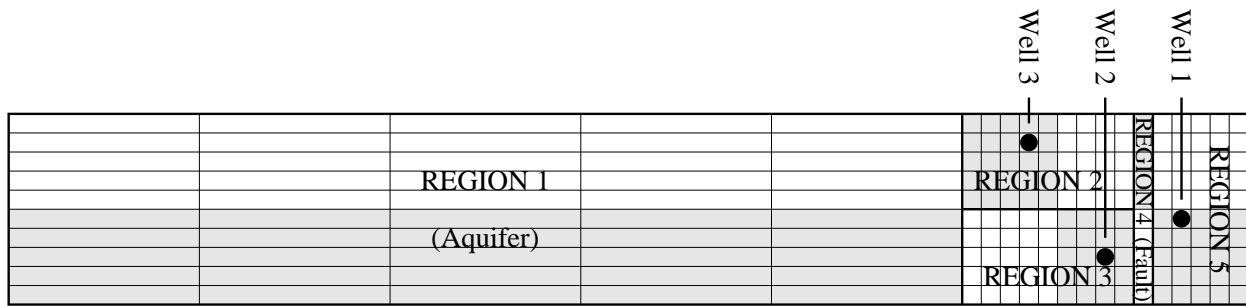


Figure 1: The reservoir geometry.

Region	Porosity			Permeability		
	lower	guess	upper	lower	guess	upper
1	4	6	8	10	100	1000
2	8	10	12	100	320	1000
3	8	10	12	100	320	1000
4	8	10	12	0.1	1	10
5	8	10	12	100	320	1000

Table 1: Best guesses and 95% limits for permeabilities and porosities.

one of the difficult components of real history matching has been removed, the example still embodies many of the important issues in history matching.

Figure 1 is a two-dimensional plan of the three-dimensional reservoir. It is a single layer dipping oil-water reservoir divided into five regions. The lowest part of the reservoir is an *aquifer* (*region 1*), which is the main pressure source to drive gas or oil through the reservoir to the three production wells, *well 1* in *region 5*, *well 2* in *region 3* and *well 3* in *region 2*. The narrow *region 4* represents a fault, which acts as a leaky barrier. The angle of dip of the reservoir, except in region 1, is in excess of 45° . The plan shows a 200 cell solver grid superimposed over the five regions that make up the reservoir: notice that the grid cells are not all the same size. The shading will be explained later.

The flows of gas, oil and water in this model are determined by *porosity* and *permeability*, a single value of each being assigned to each region. Thus, $x = (x_1, \dots, x_{10})^T$, where the first five components x_1, \dots, x_5 represent the porosities of the corresponding regions and the last five components x_6, \dots, x_{10} represent the logarithms of the corresponding permeabilities. Porosity is a measure of the percentage of volume that may be occupied by fluids, and permeability measures how easily fluid can flow through rock. Porosity (ϕ) is a scalar and is expressed as a percentage or a proportion, whereas permeability (κ), measured in millidarcies, is in general dependent on the direction of flow; in this problem it is taken to be the same in all directions. Both ϕ and κ are functions of position, but in this example they are taken to be constant within each region. In reality there is almost certainly spatial correlation in both ϕ and κ , and well core rock samples show a strong positive correlation between ϕ and κ ; we ignore these considerations here.

SSI provided us with 95% ranges and best guesses for the components of x_h as shown in table 1. Note that the porosity ranges are symmetric about the best guess and that the same is true for the permeability ranges after logarithms have been taken. We use this information to restrict the set of geologies in which we search for x_h . Each component is searched for over the range given in the table and we impose no further restriction on the simultaneous values of the components.

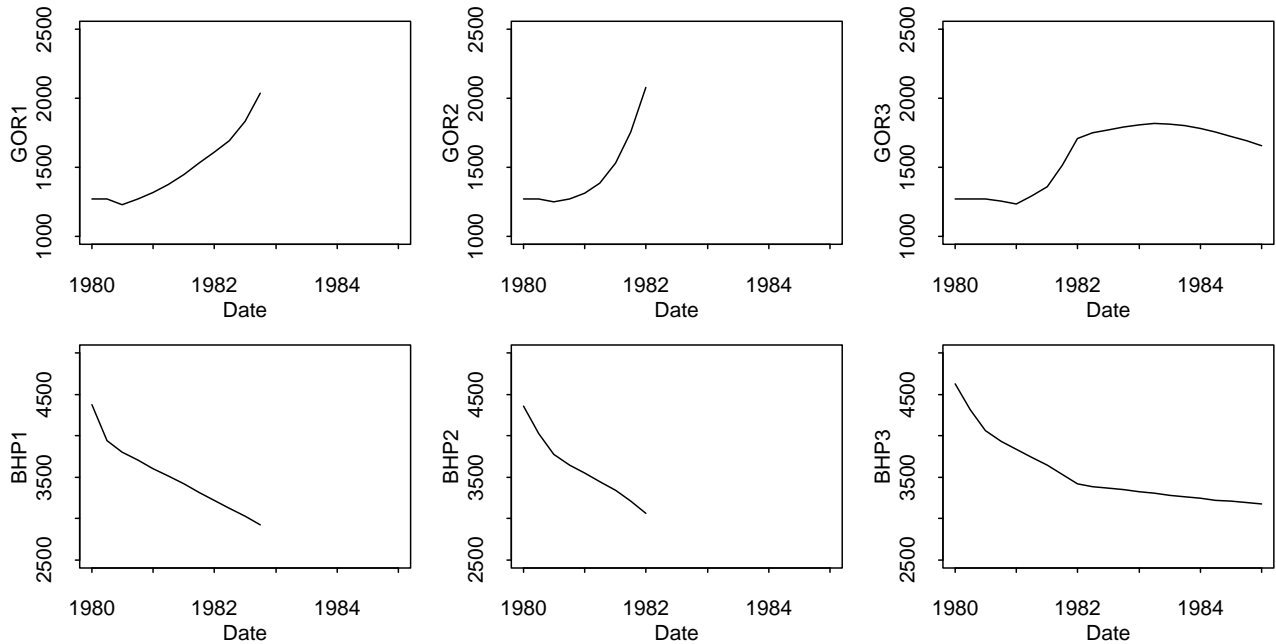


Figure 2: Production history time series.

In what follows, the porosities and logarithms of the permeabilities have been transformed to make the search range be $[-\frac{1}{2}, \frac{1}{2}]$ for each geology variable, so that we search over a 10-dimensional cube centered at the origin. The simulated production history which we are attempting to match was generated by taking each geology variable to be 0.25. Thus x_h is well within the region described in the table but is still well away from the best guess values which might be thought of as prior means for the components of x_h .

There are two history variables, *gas/oil ratio* (gor) and *bottom hole pressure* (bhp) which are given as output each quarter over a period of about five years for each of the three wells. The simulated production history for our chosen x_h is shown in figure 2. A recorded gas/oil ratio is simply the ratio of volume of gas (in cubic feet at standard conditions) to volume of oil (in standard barrels) produced over the previous three months, and we notice immediately from the history that production at both well 1 and well 2 was halted after a certain time because of overproduction of gas at these two wells. The simulator is instructed to automatically halt production at a well when gas production exceeds a certain threshold; the time when it happens depends on x . Bottom hole pressure is the pressure at the bottom of the well and physically can only be measured during production.

3 History matching as a Bayes decision problem

As the reservoir model is very complex, high dimensional and expensive to evaluate, it is essential to incorporate the expert judgments of the reservoir engineers into our procedures. These include both generalised knowledge about the reservoir and also beliefs about the response of the simulator to changes in the values of inputs.

The software development company Scientific Software Intercomp have developed a history matching approach based on a fully-specified Bayesian one-step-ahead sequential scheme to advise on possible choices for the values of the reservoir properties at which to make the next simulator run. This approach proceeds by assigning a joint prior distribution over the values of the unknown input quantities and the collection of possible deterministic outcomes of the simulator for given inputs. The approach has been implemented in the software system AHM

(Adaptive History Matching) now available for leasing from SSI; see (Watkins, Parish, and Modine 1992). The project was funded by several oil companies, and the software has been used with a degree of success.

Problems with any Bayes formulation for history matching are as follows:

Prior knowledge It is very difficult to translate the highly detailed but partial statements of prior knowledge from the engineers into the extremely high dimensional probability measures required for a formal Bayes analysis. In particular, for design tractability, we are usually restricted to very simple prior forms.

Numerical solutions The Bayes analysis involves very complicated numerical problems of integration and maximisation of extremely high dimensional multimodal distributions.

Design Bayesian design problems are notoriously computer intensive, even for moderate dimensions, simple models, fixed sample sizes and straightforward objectives. Bayes sequential design for a control problem for an ill-posed, non-linear, inverse problem in very high dimensions can require calculations which are more extensive than those for the original simulator!

The model All beliefs relate to the reservoir, but all experiments are performed on the simulator. It is difficult to incorporate into the analysis the discrepancy between model and reality.

The problems in Bayes history matching are typical of the difficulties arising in many fields of application. There is a complex design-type task which is carried out by experts with only partial success. A very complicated, high dimensional mathematical model partially underlies this task. Our only hope for approximating the solution to the mathematical problem is to incorporate expert prior knowledge. However, when this knowledge is formalised by a full Bayes specification, the resulting design problem appears to be about as hard as (or harder than!) the original task. To make progress, we should therefore consider using a simpler and more tractable approach to quantifying and analysing the beliefs of the expert. We now develop history matching using an approach which is similar in spirit to full Bayes analysis but easier to implement, in requiring both more modest prior inputs and also less complex numerical calculations.

Our formulation for history matching is therefore based on Bayes linear methodology, which uses limited prior moment specifications as input, and makes no distributional assumptions. We combine the prior inputs linearly with data to adjust the prior specifications. Also provided are a variety of interpretive and diagnostic tools to explore the implications of these beliefs. For an overview of Bayes linear methodology, see Farrow and Goldstein (1993). For a discussion of foundational issues raised by this approach, see Goldstein (1985, 1994).

Prior belief specification for Bayes linear methods is based on moment assessments. In our problem, we shall observe production output $y(x)$ for various choices of geology input x . Using these observations, we want to predict the value of y for further values of x , and in particular identify the history match x_h defined by (1). The minimal prior inputs that we require for this analysis are (i) a prior mean vector $E(y(x))$, for each geology x , and (ii) a prior covariance matrix $Cov(y(x), y(x'))$, for each pair of geologies x, x' . In this paper, we describe how to make these specifications, and their application to history matching. A full formalisation would incorporate a prior covariance structure over this variance structure; that is, specification of variances of variances, in order to learn about the covariance structure from the data. Such details are beyond the scope of this paper, however, so that we shall merely make informal comments as to how such second order learning might proceed.

4 Formulating prior beliefs for history matching

4.1 General approach to prior specification

Our strategy for prior specification is as follows.

1. Working with geologists familiar with the reservoir, we elicit a high prior probability (say 95%) region, \mathcal{G} , for the actual values of the geology inputs in the reservoir.
2. Working with reservoir engineers, we now elicit qualitative beliefs relating geology input x to production output y which we hope to be valid over \mathcal{G} . This qualitative structure is based partly on numerical/physical insights about simulator performance and partly on previous practical experience in history matching.
3. Qualitative prior beliefs about the simulator are tested, quantified and refined by constructing simple, fast approximations to the simulator; for example, reducing the number of grid blocks in the simulator to speed up the calculations.
4. The revised, quantified structure is checked by the engineer for plausibility and consistency in relation to beliefs about geology and production on the full simulator, and differences in beliefs between the full and the approximate simulator are quantified.

We now describe the qualitative and quantitative construction of such prior specifications. For brevity, we omit stage 1 of the process above, as this involves a relatively familiar type of elicitation problem. However, there are substantial questions of interest here, in particular concerning beliefs about the relationship between the actual geology of the reservoir and the formal history match x_h in the simulator.

4.2 Qualitative beliefs

Our intention is to create a prior specification which is simple enough to lead to a tractable design search methodology, while describing enough variation to generate sensible design points. Therefore, we hope to find, for each output variable, a relatively small set of input variables which govern variation over the range of interest.

While each problem will have distinctive features, certain considerations will be fairly common. For example, production at a given well will usually be strongly affected by geological features which are local to that well. The effect of non-local inputs may be expected to be more pronounced at later time points, so that our beliefs will tend to be simpler for early times. Further, we may speed up the running of the simulator by evaluating the production only at early time points for each well. This suggests a procedure where we try to identify geology inputs local to each well by matching on early production output. When we have a reasonable local match, we identify non-local features of geology by matching on later production output, treating local geology essentially as known. This illustrates the advantages of matching at each stage on carefully chosen subsets of the production variables. It is common engineering practice to match at an early and a late time point.

With the reservoir engineer, we create qualitative beliefs as follows. We choose a representative collection of production variables y_u , which are judged appropriate to match. For each selected y_u , we identify the *active collection*, $A_u = \{x_{1u}, x_{2u}, \dots\}$, of geology inputs which are judged to have substantial effects on y_u , in the range \mathcal{G} . From A_u we construct the *linear collection*, $A_u^* = \{x_{1u}^*, x_{2u}^*, \dots\}$, where each x_{iu}^* is a deterministic function of the elements of A_u which we judge to have substantial linear effect on y_u , over \mathcal{G} . Typically, though not necessarily, each x_{iu}^* will be a linear or quadratic function in one, or a product of two, elements of A_u , so that we fit a locally quadratic surface, allowing for first order interactions, although we may choose to redefine some

of the variables, for example placing them on the log scale. Interactions may be anticipated between non-local variables, as the effect on a well of changes in the geology far from that well may be largely dependent on the intervening geology. We further elicit prior judgments as to the ‘signs’ and relative strengths of the effects of the various elements of each collection A_u^* on the corresponding production output y_u , and informal judgments as to the amount of variation in y_u which can be explained by A_u^* .

4.3 Example: qualitative beliefs

We simplify our treatment of the example somewhat by reducing the number of components of y to match on; specifically, we consider gor and bhp at time points 2 and 6 (6 and 18 months) at each of the three wells. In this reduced version y has twelve components. The reduced history y_h is given by

$$y_h = (1229, 1448, 3806, 3414, 1249, 1532, 3777, 3337, 1270, 1358, 4061, 3643)$$

where the components are in three successive groups of four in the order gor_2 , gor_6 , bhp_2 and bhp_6 for each of wells 1, 2 and 3. We subtract these values from the values produced by the simulator for other input geologies and so we seek to find the geology for which the outputs are all 0.

The simplification avoids over-burdening our account with too many layers of modelling. We do not have to model the time of well shut in, which will not occur earlier than the sixth quarter for the range of x under consideration and we avoid time series modelling of the production variables. Moreover, matching only on early time points means that simulated production is needed only for early time points, which implies a substantial saving of CPU time. Of course, as we match on these twelve variables we compare the simulator outputs at the other time points with the corresponding history values.

We now consider the first stage of modelling prior beliefs about the relationships between the 10 input variables, the 5 porosities and 5 permeabilities for the 5 regions, and the twelve output variables, gor and bhp at time points 2 and 6 for each of the three wells.

We had a brief elicitation discussion with a reservoir engineer at SSI. For each of the twelve output variables, he was asked to rank the 10 input variables in order of the strength of their influence on the output variable. His beliefs turned out to be very highly structured. At each well, he chose the same ranking for all four variables: gor and bhp at the two time points; in other words his beliefs were about influences on wells rather than about particular variables at wells. Moreover, in all cases, he ranked equal the porosity and permeability from each region; his beliefs about influence were in terms of regions rather than specific input variables. As a result, table 2, which shows his beliefs, is comparatively simple.

In the table, κ_j symbolises the permeability in region j and ϕ_j its porosity. The rankings displayed in the table indicate the relative importance, for each well, of the influences from the five regions. Thus, for example, the 4 in the top left hand corner indicates that the reservoir engineer believes that region 1 is the region having the least influence on well 1, which is situated in region 5. In fact, this well is believed to be most influenced by the permeability and porosity in its own region, the influence of the geology of the other regions declining with distance. Belief in strong dependence on geology local to a well is a feature for all 3 wells.

The engineer qualified his rankings with the following remarks, which we shall not use directly but which we may expect to confirm in the quantitative procedures described later: although the rank orderings for time point 6 are the same as those for time point 2 the “strength” of the contrasts is reduced for the later time point; bhp

Well	Region	κ_1 and ϕ_1	κ_2 and ϕ_2	κ_3 and ϕ_3	κ_4 and ϕ_4	κ_5 and ϕ_5
1	5	4	3	3	2	1
2	3	4	2	1	3	4
3	2	5	1	2	3	4

Table 2: Reservoir engineer’s beliefs about relative importance of regional parameters for different wells.

is an increasing function of each of permeability and porosity, whereas gor is a decreasing function of each of them, for both time points and at each of the three wells; the influence of permeability on both gor and bhp in a region where a well is located is greater than that of the porosity of the region; and the porosity of the aquifer (region 1) and the permeability of the fault (region 4) will be key geological variables, particularly for gor and bhp at later time points.

4.4 Quantifying prior beliefs

We now construct a simple description of second order beliefs consistent with the qualitative features that we have identified. For each production output y_u that we intend to match, we have a collection $A_u^* = \{x_{1u}^*, x_{2u}^*, \dots\}$ of functions of x which we judge to have substantial linear effects on y_u . We express this dependence as

$$y_u = \sum \beta_{iu} x_{iu}^* + \epsilon_u(x) + \delta_u(x). \quad (2)$$

In (2), each β_{iu} is a random quantity, for which we must specify a prior mean and variance, and possibly covariances with other such quantities. Each $\epsilon_u(x)$ expresses the difference between y_u and the linear form based on the collection A_u^* , keeping all other elements of the geology x fixed, and is taken to be a stationary random function of x , uncorrelated with all terms in β and δ , and having prior mean zero and prior covariance

$$Cov(\epsilon_u(x), \epsilon_u(z)) = \sigma_{\epsilon_u}^2 \exp\left(-\sum \theta_{iu} (x_{iu}^* - z_{iu}^*)^{p_i}\right) \quad (3)$$

where θ_{iu} , p_i are non-negative constants to be specified. Each $\delta_u(x)$ is a residual term, with prior mean zero, prior variance $\sigma_{\delta_u}^2$, for all x , and $Cov(\delta_u(x), \delta_u(z)) = 0$, unless $x = z$. The terms $\delta_u(x)$ express the additional variation (hopefully small) in our beliefs for y_u given that y_u does depend, to some extent, on aspects of the geology that we have excluded from A_u . The reason that we do not express a spatial correlation structure for δ_u is that, when history matching for y_u , for simplicity we usually treat those elements excluded from A_u as unknown. Technically, the effect of this form is to slightly reduce spatial correlations, which has the side benefit of compensating for possible local ill-conditioning and thus stabilising various of our subsequent calculations. However, if $\sigma_{\delta_u}^2$, forms a large component of total variation, then we must express a covariance structure over δ_u , of similar form to (3), as even if we only search for matches for y_u within A_u , covariances over δ_u will be relevant for updating beliefs.

In this formulation, spatial and temporal correlations between different production outputs are induced by the shared dependence on similar geology inputs. We may strengthen such relationships by introducing correlations between coefficients β_{iw}, β_{jv} if the terms x_{iw}^*, x_{jv}^* express similar types of variation in outputs y_w, y_v which are closely spatially or temporally related, and we may similarly induce correlations between terms ϵ_w, ϵ_v . Alternately, we may explicitly model the temporal development of the process, for example applying forms such as (2) to

appropriate transforms of the changes in the production outputs over time.

It is possible that, when we try to validate (2), as described below, we find that the various coefficients β_{iu} are very small, with most of the variation in y_u explained by ϵ_u . In such cases, we may replace the coefficients β_{iu} by corresponding stationary random (but slowly varying) functions of x_{iu}^* , namely $\beta_{iu}(x_{iu}^*)$ with stationary mean and covariance structure, for example using the localised regression forms of O’Hagan (1978). Such forms are more flexible and may provide better descriptions of variation, but are less tractable for the design problems arising in history matching.

There are two complementary approaches that we may follow in quantifying prior second order structure over (2). Firstly, we may elicit prior means, variances, and covariances for the quantities directly from the reservoir engineer. Secondly, we may assess these values by creating simple approximations to the reservoir simulator, for example versions which have the same formal geology and production variables, but which are much faster to run as they are based on larger, and thus fewer, grid blocks. As we can evaluate many runs of the fast simulator, for the cost of one run of the full simulator, we may quantify prior beliefs by data analytic fitting of (2) over a sample of values from the fast simulator, exploiting the qualitative judgments of the reservoir engineer.

The latter approach may be particularly appropriate for history matching, as one way in which the engineer will try to answer elicitation questions related to the full simulator is by trying informally to give approximate solutions for very simple versions of the simulator. Therefore, it is natural to view the actual behaviour of the fast simulator as providing extra information which refines any such prior assessments. Further, such fitting acts as a useful substitute if the expert is unwilling or unable to devote the considerable time and effort required for careful quantitative elicitation. While we prefer using expert judgment to suggest the qualitative form for our beliefs, it is feasible to quantify prior beliefs by fitting as described below. Note that the approximation of a complex system by a simpler version of that system offers a general approach to quantifying prior beliefs.

If we first quantify beliefs about the fast simulator, we must then transform our description from the fast to the full simulator. There are various ways to do this, and we currently express the relationship between the simulators as follows. Let $\tilde{y}_u(x)$ denote the value of production variable u calculated by the fast simulator for geology x . For each \tilde{y}_u , we fit two components, namely a systematic component $\tilde{s}_u(x) = \sum_u \tilde{\beta}_{iu} x_{iu}^*$ and a residual component $\tilde{r}_u(x) = \tilde{\epsilon}_u(x) + \tilde{\delta}_u(x)$. For each corresponding y_u , we fit corresponding uncorrelated systematic and residual components $s_u(x) = \sum_i \beta_{iu} x_{iu}^*$ and $r_u(x) = \epsilon_u(x) + \delta_u(x)$.

The forms $\tilde{s}_u(x)$ are checked for qualitative plausibility on the full simulator. We therefore view $\tilde{s}_u(x)$ as our prior mean for $s_u(x)$, given the observations on the fast simulator, and so judge the discrepancy $s_u(x) - \tilde{s}_u(x)$ to be uncorrelated with $\tilde{s}_u(x)$. Thus, we require, for each coefficient β_{iu} , that

$$\beta_{iu} = \tilde{\beta}_{iu} + b_{iu} \tag{4}$$

where all quantities b_{iu} have mean zero and are uncorrelated with all quantities $\tilde{\beta}_{jv}$. Thus, our prior mean for each β_{iu} is the mean value that we have assessed for $\tilde{\beta}_{iu}$ and the variance is the sum of $Var(\tilde{\beta}_{iu})$, the variance due to uncertainty about this mean value, as we only evaluate a sample of values on the fast simulator, and $Var(b_{iu})$, expressing beliefs about differences between the two simulators. The latter variance is elicited by expert judgment based on physical and numerical insights into differences between the simulators. If we have physical reasons to change the mean estimates, then these will be incorporated as prior means for b_{iu} . This procedure is usually a reasonable approximation to the fully coherent procedure involving joint prior second moment specification over

the two simulators, followed by adjusting beliefs over the fast simulator, and consequently the full simulator, by Bayes linear updating given the sample values on the fast simulator. However, if there are strong quantitative prior beliefs as to the values of the β coefficients, then the full prior specification may substantially reduce the number of runs needed on the fast simulator.

While we view the systematic components for the fast simulator as predictive for the corresponding components on the full simulator, our current judgment is that residual variation on the two simulators is far more dependent on the individual characteristics of the two simulators, so that we view $\tilde{r}_u(x)$, $r_u(x)$ as uncorrelated. Therefore, we do not use the residual values on the fast simulator to predict the corresponding prior mean values on the full simulator. Instead, we choose the prior means for terms $\epsilon_u(x)$, $\delta_u(x)$ to be zero, and reassess the constants in the covariance structure (3) using the values we have assessed for $\sigma_{\tilde{\epsilon}_u}^2$, $\sigma_{\tilde{\delta}_u}^2$. In principle, this involves a Bayes linear variance analysis which is outside the scope of this paper, but informally we express beliefs relating the values $\sigma_{\tilde{\epsilon}_u}^2$, $\sigma_{\tilde{\delta}_u}^2$, to $\sigma_{\epsilon_u}^2$, $\sigma_{\delta_u}^2$. The simplest such beliefs result in assessing roughly similar values for $\sigma_{\tilde{\epsilon}_u}^2$, $\sigma_{\tilde{\delta}_u}^2$, to those fitted on the fast simulator. We will typically use similar values for each correlation parameter θ_{i_u} as were fitted on the fast simulator, at least initially, unless there is some external reason to change these values. However, we may alter these values subsequently, if this is indicated by our diagnostic monitoring

4.5 Example: quantifying prior beliefs

In the example, we used a coarsening of the simulation grid from 200 to 10 cells for our fast simulator. The shading in figure 1 shows the coarsened grid. A rule of thumb used by SSI suggests that for real problems, run time of the simulator is proportional to the number of cells raised to the power 1.2. For the coarsening used in the example, that implies 36 runs of the the fast simulator in the time taken for a single run on the full simulator. We have used 360 runs of the fast simulator to develop a description for the fast simulator of the relationship between input geologies and output productions.

We took a 360 point latin hypercube design in the input variables (for details see McKay, Conover, and Beckman (1979)) and ran the fast simulator on each geology. Starting from the engineer’s qualitative description and using ordinary least squares, we then built for each output variable a linear model of the form (2) with linear, quadratic and first order interaction terms in the geology variables. The goal was to choose models which explained a large proportion (75% or more) of the output variance while involving a small collection of geology variables. We also preferred to split the output variables into groups so that the model for each output within each group involves the same active geology variables and so that different groups mostly involve different geology variables. Some overlap turned out to be unavoidable. Having chosen the model for each output variable, we then estimated the the variances $\tilde{\sigma}_{\tilde{\epsilon}_u}^2$ and $\tilde{\sigma}_{\tilde{\delta}_u}^2$ and the correlation parameter $\tilde{\theta}_u$ for the fast simulator using the residuals from the least squares fits.

The fast simulator description is shown in table 3. We see that, broadly speaking, the engineer’s assessments are correct. The active variables are mostly those local to the wells, increasing permeability or porosity increases bhp and decreases gor, and permeabilities have larger coefficients than porosities. However, his belief that aquifer porosity (x_1) and fault permeability (x_9) are key variables is not supported at this stage; in fact it is aquifer permeability (x_6) which is a driving force in the problem and is an active variable for both wells 2 and 3.

As described in section 4.4, we take the estimated coefficients from the equations fitted on the fast simulator, shown in table 3, as our prior means for those coefficients for the full simulator. As prior standard deviation for each coefficient, for simplicity we take the magnitude of the prior mean, which allows considerable uncertainty

Equation	$\sigma_{\tilde{\delta}}$	$\sigma_{\tilde{\epsilon}}$	$\tilde{\theta}$
Well 1 in region 5			
$\widehat{bhp}_2 = -219 + 96x_5 + 588x_{10} - 209Q(x_{10})$	4	10	21
$\widehat{bhp}_6 = -397 + 353x_5 + 868x_{10} - 336Q(x_{10})$	20	32	7
$\widehat{gor}_2 = 26 - 58x_5 - 40x_{10} + 11Q(x_5) + 10Q(x_{10})$	2	2	5
$\widehat{gor}_6 = 521 - 599x_5 - 254x_{10} + 177Q(x_5) + 104Q(x_{10})$	34	29	6
Well 2 in region 3			
$\widehat{bhp}_2 = -338 + 63x_3 + 962x_8 + 172x_6 - 348Q(x_8)$	24	18	4
$\widehat{bhp}_6 = -499 + 177x_3 + 289x_6 + 1380x_8 - 548Q(x_8)$	72	61	2
$\widehat{gor}_2 = 9 - 20x_3 - 57x_6 - 39x_8 + 57x_3x_6 + 93x_6x_8$	10	9	6
$\widehat{gor}_6 = 315 - 215x_3 - 442x_6 - 171x_8 + 180Q(x_6) + 138Q(x_8)$	99	57	1
Well 3 in region 2			
$\widehat{bhp}_2 = -102 + 43x_2 + 210x_6 + 150x_7 - 58Q(x_7)$	21	19	6
$\widehat{bhp}_6 = -123 + 133x_2 + 206x_6 + 155x_7 - 48Q(x_6) - 54Q(x_7)$	45	21	1
$\widehat{gor}_2 = -5 + 7x_2 + 19x_6 - 9Q(x_6) - 21x_2x_6$	1	3	2
$\widehat{gor}_6 = 97 - 165x_2 - 263x_6 + 85Q(x_6) + 231x_2x_6$	52	38	1

Table 3: Equations approximating the coarse grid simulator; $Q(x) = 3x^2 - \frac{1}{4}$ is chosen to be orthonormal to x .

about the coefficient, and is consistent with subsequently discovering that the sign of a coefficient should be reversed. All prior correlations between coefficients are 0, again largely for simplicity, although least squares estimation for the latin hypercube design produces uncorrelated estimates. The parameters σ_{ϵ_u} , σ_{δ_u} and θ_u are all taken to be the same as the estimates of their fast simulator equivalents.

Because the example has a small number of grid cells, we were able to make a detailed examination of the relationships between the fast and full simulators. It was possible to run the full simulator on each of the geologies in the latin hypercube used to estimate equations for the fast simulator, a luxury that would be unavailable in real history matching problems. We observed that the A_u^* provide good fits for y_u on the full simulator, but often with noticeably different coefficients. However, the magnitude of the change in a coefficient from fast to full models is generally, but not uniformly, much smaller than suggested by our simple choice of prior standard deviation. We also observed that generally the covariance structures on the two grids differ little.

5 Bayes linear history matching

High dimensional Bayes design is extremely computer intensive. Bayes linear design calculations are more tractable, and in particular much of the numerical processing for choosing the next design point may be carried out while the simulator is running using the current choice. The process of history matching involves the following iterative cycle.

1. Choose one (or several) design points; that is, values of the geology input x at which to evaluate the corresponding production output, y .
2. Having observed simulator production, update beliefs.
3. Assess diagnostics based on observed outputs and where necessary reformulate qualitative as well as quantitative modelling, either because beliefs are discrepant with observations, or because we need to refine our belief description for efficient history matching.
4. Use updated beliefs to reassess plausible locations for the history match.

We now describe this process, beginning with the updating of beliefs.

5.1 Bayes linear belief adjustment

As we only assess second order structure, the appropriate belief adjustments are by linear fitting; see the overview in Farrow and Goldstein (1993). The adjusted expectation of the random vector W given Z , denoted as $E_Z(W)$, is the best linear fit in Z for W , under quadratic loss. In this case

$$E_Z(W) = E(W) + Cov(W, Z)[Var(Z)]^{-1}(Z - E(Z)). \quad (5)$$

(If $Var(Z)$ is of less than full rank, then we first reduce Z to any maximal subcollection of full rank.) Linear fitting of W by Z leaves an adjusted variance of $Var_Z(W)$, where

$$Var_Z(W) = Var(W) - Cov(W, Z)[Var(Z)]^{-1}Cov(Z, W) = Var([W/Z]) \quad (6)$$

and $[W/Z] = W - E_Z(W)$ is the adjusted version of W by Z . The amount of variance resolved by the adjustment is $RVar_Z(W)$, where

$$RVar_Z(W) = Var(W) - Var_Z(W) = Var(E_Z(W)).$$

When we adjust W firstly by U and secondly by V , then the change in adjusted expectation for W is equal to the centred adjusted expectation of W given $[V/U]$, i.e.

$$E_{\{U,V\}}(W) = E_U(W) + E_{[V/U]}(W - E(W)). \quad (7)$$

The terms on the right of (7) are uncorrelated, so that

$$RVar_{\{U,V\}}(W) = RVar_U(W) + RVar_{[V/U]}(W). \quad (8)$$

We call $E_{[V/U]}(W - E(W))$ and $RVar_{[V/U]}(W)$, the partial adjusted expectation and partial resolved variance of W by V given U .

5.2 Adjusted beliefs for history matching

At stage m , we have evaluated a collection $\langle y \rangle_{[m]} = \{y_{[1]}, \dots, y_{[m]}\}$, of production output vectors corresponding to input geologies $x_{[1]}, \dots, x_{[m]}$. We now adjust beliefs about the value of the production y , corresponding to general geology x , using $\langle y \rangle_{[m]}$. We simplify notation by writing

$$E_{[m]}(y) = E_{\langle y \rangle_{[m]}}(y), \quad Var_{[m]}(y) = Var_{[\langle y \rangle_{[m]}}(y), \quad RVar_{[m]}(y) = RVar_{\langle y \rangle_{[m]}}(y).$$

At stage $m - 1$, having observed $\langle y \rangle_{[m-1]}$, we choose a further geology $x_{[m]}$, and wait for the simulator to evaluate the production $y_{[m]}$. We denote the adjusted version of $y_{[m]}$, given $\langle y \rangle_{[m-1]}$, and the partial adjustment for a general y given $y_{[m]}$ as

$$y_{[m/]} = y_{[m]} - E_{[m-1]}(y_{[m]}), \quad E_{[m/]}(y) = E_{y_{[m/]}(y - E(y)),$$

with corresponding partial resolved variance

$$RVar_{[m/]}(y) = RVar_{y_{[m/]}}(y).$$

From (7) and (8) we have the stepwise updating rules

$$E_{[m]}(y) = E_{[m-1]}(y) + E_{[m/]}(y) \quad (9)$$

and

$$RVar_{[m]}(y) = RVar_{[m-1]}(y) + RVar_{[m/]}(y). \quad (10)$$

As our prior beliefs for each component of production, y_u , that we are currently matching are expressed by (2) except at design points, our adjusted expectations are given by

$$E_{[m]}(y_u) = \sum_i (E_{[m-1]}(\beta_{iu}) + E_{[m/]}(\beta_{iu}))x_{iu}^* + E_{[m-1]}(\epsilon_u(x)) + E_{[m/]}(\epsilon_u(x)) \quad (11)$$

so that we stepwise update the regression coefficients and the ϵ terms.

Note that $RVar_{[m/]}(y)$ depends on $x_{[m]}$ but not on $y_{[m]}$, which allows us to program much of the design choice for $x_{[m+1]}$ while $y_{[m]}$ is being evaluated by the simulator. Further from (5), (6) and (9), the partial change in adjusted expectation for a general production y , given $y_{[m]}$, is

$$E_{[m]}(y) - E_{[m-1]}(y) = E_{[m/]}(y) = C_{[m/]}(y)^T d_{[m/]} \quad (12)$$

where

$$C_{[m/]}(y) = Cov(y, y_{[m/]}), \quad d_{[m/]} = (Var(y_{[m/]}))^{-1} y_{[m/]}. \quad (13)$$

Thus, updating the expectation surface for y is fast as, from the data, we only evaluate the single vector $d_{[m/]}$, which we multiply by the adjusted covariance matrix between each element of interest and the current adjusted observation $y_{[m/]}$. The adjusted covariance does not depend on the observed value of $y_{[m]}$ and so may be computed while the simulator is evaluating $y_{[m]}$.

Further, there is a natural diagnostic which is suitable for tracking all changes in our expectations over the whole production surface, $y(x)$, namely the comparison of the observed with the predicted value for $d_{[m/]}$, or equivalently for $y_{[m/]}$. For example, we will usually be interested in the magnitude of the largest predictive discrepancy over all possible linear combinations $c^T y$ of components of production for all geologies x , which from (12) and (13) is bounded as

$$\max_{c,x} \frac{(E_{[m]}(c^T y) - E_{[m-1]}(c^T y))^2}{Var(E_{[m]}(c^T y) - E_{[m-1]}(c^T y))} \leq S_{[m/]} = \max_c \frac{(c^T y_{[m/]})^2}{Var(c^T y_{[m/]})}. \quad (14)$$

$S_{[m/]}$ is straightforward to compute; for example, it is equal to the sum of squares of any linear transform of $y_{[m/]}$ whose components are uncorrelated and have unit variance. Therefore

$$E(S_{[m/]}) = Rank(Var(y_{[m/]})) = r_{[m/]}, \quad (15)$$

so that we evaluate the ratio $S_{[m/]}/r_{[m/]}$ after each observation. We expect that, possibly after a burning in

run	adjusted expectation					adjusted standard deviation				
	1	x_5	x_{10}	$Q(x_5)$	$Q(x_{10})$	1	x_5	x_{10}	$Q(x_5)$	$Q(x_{10})$
0	521	-600	-255	177	104	521	600	255	177	104
5	214	-165	-659	129	30	93	88	101	86	28
10	183	-379	-459	179	-56	58	58	94	82	21

Table 4: Adjusted expectations and standard deviations of well 1 gor at time 6.

period, the ratio will settle down to a stable value. If that value is much larger/smaller than one, then we should inflate/deflate the residual variance terms σ_c^2 , σ_d^2 . If the ratio does not settle down, then this may suggest that there is something qualitatively wrong with our prior description, for example that we have omitted an important term in some collection A_u .

Note that we have supposed in the above account that we use all of the production output at each stage to predict all of the production output at the next stage. However, usually we structure our prior description so that various aspects of the production, for example uncertainties related to different wells, are uncorrelated, and we may therefore analyse each such sub-problem in isolation.

5.3 Example: adjusted expectations

For the specified description, adjustments for each output variable can be performed separately. Here we illustrate the adjusted expectation and standard deviation of the coefficients and the expectation and standard deviation surfaces for the gas/oil ratio at time 6 for well 1 in region 5, after five and after ten runs on the full simulator. The actual choice of geology at each run was determined by a design criterion to be described in section 6.2. We do not give adjusted correlations between the coefficients, but note that in magnitude they are all less 0.5, and initially they were all taken to be *zero*. The results are summarised in table 4. Notice that apart from the quadratic term in x_{10} , the signs of the coefficients do not change, but there are large changes in magnitude. The eventual ratios of expectation to standard deviation emphasise the dominance of the constant and linear terms in the description for this production variable, at least for the range of variation of x_5 and x_{10} considered so far.

Table 5 shows for each of the 10 runs the (x_5, x_{10}) coordinates of the design points chosen according to the design criterion described in section 6.2, the signed difference $y_{[m]}$ between the observed value and the predicted value of the gas/oil ratio, and the adjusted standard deviation $s_{y_{[m]}}$ of this difference, and their ratio. Notice that runs 4, 5 and 6 are near three of the four ‘‘corner points’’ of the (x_5, x_{10}) region. For this production variable, the first three runs are estimating the overall mean, the next three the other parameters, and the last four have little effect.

The diagnostic ratios indicate close agreement between observed and predicted gor, with the exception of design point six. We experimented by taking the same 10 design points but moving the point 6 to the end of the sequence. We found that the coefficient expectations changed little from run 5 to run 9 and very substantially from run 9 to run 10. Design run 6 is both influential and surprising, which is true also for the other well 1 variables, especially the other gor variable. In this illustration, design points have been generated automatically, but this diagnostic would signal the desirability of expert intervention to consider changes in the model.

Figure 3 shows a contour plot of the expected gor surface after 10 runs. Notice the position of the zero line in the expectation surface, and that the standard deviation surface appears very ‘‘flat’’.

m	x_5	x_{10}	$y_{[m]}$	$s_{y_{[m]}}$	$\frac{y_{[m]}}{s_{y_{[m]}}}$
1	0.11	0.12	-279	530	-0.526
2	-0.22	-0.15	-44	216	-0.204
3	-0.26	0.44	-342	169	-2.024
4	0.50	0.48	-30	195	-0.154
5	-0.50	-0.50	-14	138	-0.101
6	0.48	-0.42	-292	94	-3.106
7	0.05	0.15	38	44	0.864
8	0.08	0.14	28	41	0.683
9	0.09	0.15	12	39	0.308
10	0.10	0.14	11	38	0.289

Table 5: Diagnostics for a sequence of ten design points for well 1 gor at time 6.

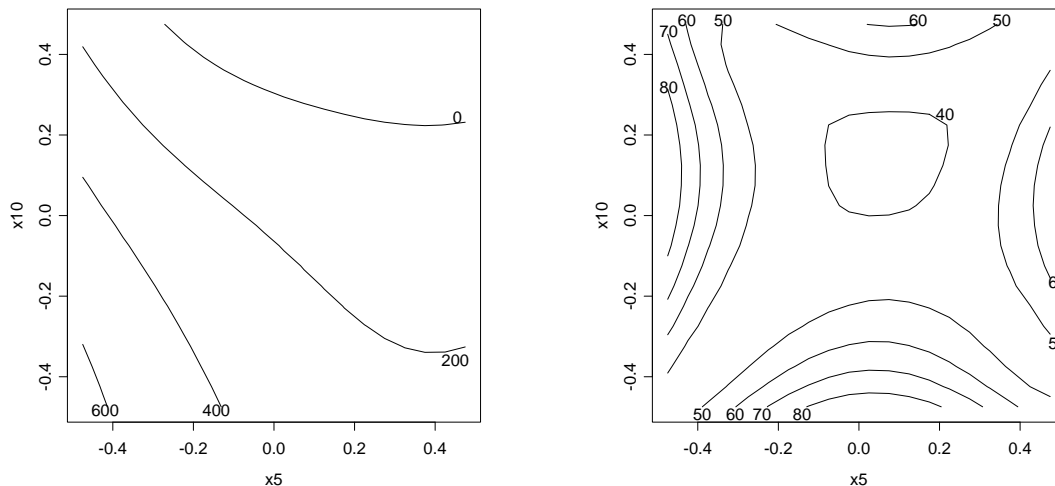


Figure 3: Contours of adjusted expectation and adjusted standard deviation of well 1 gor at time 6 as functions of x_5 and x_{10} after 10 design points.

5.4 Identifying potential history matches

At any stage, we may assess for which values of x it is ‘plausible’ that $y(x) = y_h$. We have transformed the observed production history, y_h to zero, and we must first decide whether to seek the set G of values of x for which it is plausible that $y_j = 0$, for each j , or to seek for each j the set G_j of x values for which it is plausible that $y_j = 0$, and then look for history matches in the intersection. Seeking G directly is most efficient if we have carefully formulated joint beliefs and have reliable history data from the reservoir. In favour of seeking each G_j individually, however, is that if we have described part of the system badly, or if some aspects of the history that we are trying to match are in error (through faulty data measurement or misrepresentation of our beliefs), then overall matching may be impossible, whereas matching on most individual components may still be straightforward. Further, the number of quantities to be calculated and stored is substantially reduced by single variable matching. In general, we consider it appropriate to take an intermediate position, namely to group together ‘similar’ outputs into exclusive subsets, and match on a subset by subset basis. As we have already observed, practical considerations usually lead us to match only on a subset of output variables at any stage.

Any methods to identify plausible history matches involve searches and evaluations over high dimensional x spaces, repeated at many stages. Therefore, we use fast simple methods for identifying regions where it is

possible that we may find the match, and within these regions, where appropriate, we carry out more complex assessments to quantify relative plausibility. To simplify our account, we describe history matching separately for individual scalar components y_j of y . Extensions to matching for collections of production variables are straightforward. Our overall procedure is as follows.

Firstly, we identify, for each component of production, those regions of x space in which we judge that the match might possibly occur. Our belief revisions are limited to the current adjusted mean and standard deviation functions $\mu_{mjx} = E_{[m]}(y_j(x))$, $\sigma_{mjx} = \sqrt{\text{Var}_{[m]}(y_j(x))}$ for the production surface. A simple assessment is to identify those values of x for which $c_{mjx} = \mu_{mjx}/\sigma_{mjx}$ is near zero, so that, informally, we consider it possible that such an x could be a history match for component y_j . We are most interested in values of x which lie in such ‘possible regions’ for most components of production. If we can’t find such common values, then this may lead us to question our formulation.

Within regions where, for most j , c_{mjx} is relatively small, we also want σ_{mjx} to be small so that x is not only a possible, but also a relatively likely history match. We may formalise such judgments in various ways. For example, we may suppose that the corresponding y_j has probability distribution of form $(1/\sigma_{mjx})f((y_j - \mu_{mjx})/\sigma_{mjx})$ for some general density function, $f(\cdot)$, symmetric and unimodal about zero. We may evaluate this function at $y_j = 0$, giving $l(x) = (1/\sigma_{mjx})f((-c_{mjx})$ as analogous to a likelihood that x is really a history match for y_j . Different choices of density f suggest different ways of combining c_{mjx} and σ_{mjx} into a relative plausibility measure. For example, we might change from a comparatively thick tailed distribution early in our search to a more centrally concentrated form when we have more confidence in the quality of our fit.

More in the spirit of the Bayes linear approach is to consider whether a value $y_j(x) = 0$ would be consistent with the rest of our judgments and data. For example, we may consider, were we to make our next observation at x and to observe that $y_j = 0$, how our beliefs about other values of y_j would change, for example on a grid of geology values neighbouring to x . We can now measure our surprise at the magnitude of these changes, given such an observation, using Bayes linear diagnostics, such as the ‘bearing length’ described in Goldstein (1988). This approach conveys more information than the likelihood type approach, as it pays attention not just to the local mean and variance at x , but also to the covariances with neighbouring values. The approach requires more computation and so is most practicable when applied to carefully selected subsets of values after we have substantially reduced the search space of possible history matches.

5.5 Example: plausibility maps

As suggested, we have used the ratio of the adjusted expectation to adjusted standard deviation of the output variable as a “possibility function”. Figure 4 is based on the adjusted expectation, and standard deviation of well 1 production variables at time 6, after observing the values at the 10 design points shown in section 5. The left panel of figure 4 shows the ratio of adjusted expectation to adjusted standard deviation for well 1 gor at time 6. We only plot contours for the ratio up to a magnitude of 5. The only ‘impossible’ points are those where both x_5 and x_{10} are negative. The centre panel shows the maximum, over the four well 1 production variables, of the absolute values of the ratios of adjusted expectation to adjusted standard deviation. Note that the region where the ratio is small is much more concentrated than for the single variable in the left panel.

The right panel of figure 4 shows the gaussian ‘likelihood’ after 10 design points that values of x_5 and x_{10} are the values in x_h , based on well 1 gor at time 6. We see points with high likelihood are confined to a strip which determines x_{10} fairly well as a function of x_5 and where x_5 cannot take values much less than -0.2 . The

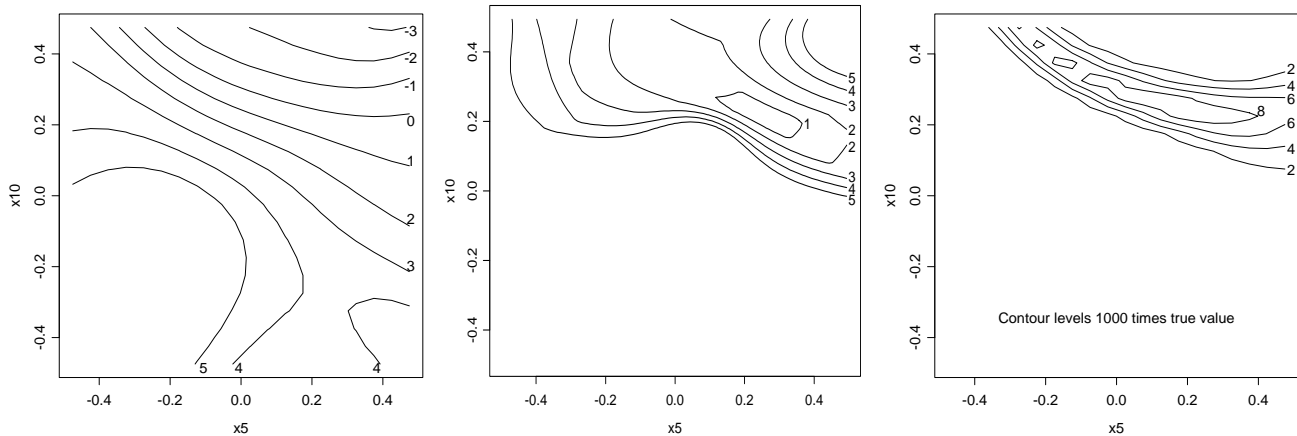


Figure 4: Contours of three functions of the adjusted expectations and standard deviations of well 1 variables as functions of x_5 and x_{10} : ratio of expectation to standard deviation for time 6 gor, the maximum magnitude over all four well 1 variables of the ratio and the ‘likelihood’ based on time 6 gor. The adjusted beliefs used are those after 10 design points.

middle of the strip corresponds closely to the curve of zero adjusted expectation which is the zero contour in the left panel of the figure.

6 Sequential design for history matching

A simple way to choose design points for history matching is, at each stage, to evaluate the simulator for the geology x which we believe will most closely satisfy $y(x) = y_n$. Given a reasonably accurate location for the match, this may be a sensible procedure, but, in general high dimensional searches, it is very inefficient. Therefore, we develop instead an approach which begins by identifying and reducing uncertainty over regions of plausible history matches, and only gradually moves to seeking the match point. While the design criterion approach does suggest automatic design choices, it is also intended to offer guidance to an expert carrying out the match, by summarising the strengths and weaknesses of competing design choices.

6.1 Criteria for uncertainty reduction

Each design point at which we evaluate the simulated production is expected to reduce uncertainty about each component y_u of the production for each possible input value x . We want to combine the various variance reductions in a way that is tractable and interpretable. There is a close analogy here with multicriterion utility for design (see Farrow and Goldstein (1992)), but space limitations allow us to give only a simple restricted development in terms of variance reduction: a fuller account will be given elsewhere.

We develop a tractable and flexible criterion for judging the potential value of any design choice. This criterion separates our calculations into two parts, formal evaluations of expected reductions in variance and informal assessments of the relative importance that we give to such reductions, based on any information that we feel relevant in guiding our choice. We may elicit such weightings directly from the reservoir engineer, by asking for preference rankings over alternative combinations of variance reductions. However, as such weightings change at each stage, typically we will model the choice of weightings, aiming to correspond informally with expert judgment. The design criterion may be used to generate automatic design choices, or alternately, to support decisions of the reservoir engineer, allowing design choice to proceed with interventions at key stages.

Our variance based design criterion is constructed as follows. Let y_u and $y_{[m]u}$ be values of the production component u corresponding to a general geology x and a potential design choice $x_{[m]}$. At stage m , the expected reduction in variance of y_u at x by evaluating the simulator next at $x_{[m]}$ is

$$(Var_{[m-1]}(y_u(x)))^{-1} RVar_{[m]}(y_u(x)). \quad (16)$$

The overall weighted reduction in variance of y_u by evaluating $y(x_{[m]})$ is given by

$$U_{y_u}(x_{[m]}) = \int w(x)(Var_{[m-1]}(y_u(x)))^{-1} RVar_{[m]}(y_u(x))dx \quad (17)$$

where the weight $w(x)$ is a function of the current plausibility evaluation of x as a potential history match.

Our design criterion is to choose $x_{[m]}$ to maximise

$$U(x_{[m]}) = \sum_u l_u U_{y_u}(x_{[m]}) \quad (18)$$

where the weights l_u reflect the relative importance of matching the various production variables y_u .

Thus, we use formal calculations to identify design choices which reduce the variation of various output components in various subregions, and exploit, where available, the engineer's expertise to choose which such reductions are most important at the current stage. For example, early in the process, we may weight roughly equally all points that we judge as possible history matches, whereas when we feel we have narrowed down our search sufficiently, we may only weight highly on points which we feel are very likely to be the match.

6.2 Choosing design points

It is a substantial computational task to find good design points when searching over a high dimensional input space. We have created prior descriptions for which as much of the variation as possible is attributed to a linear surface. Therefore, we begin by exploiting this approximate linearity. The description of our beliefs for production component y_u is of the form

$$y_u = \beta_u^T x_u^* + r_u(x) \quad (19)$$

where x_u^* is the vector of elements of A_u^* . When choosing $x_{[m]}$, we have current assessments $E_{[m-1]}(\beta_u) = \bar{\beta}_u$ and $Var_{[m-1]}(\beta_u) = V_u$. Also, $E_{[m-1]}(r_u(x))$ is near zero, except perhaps for x near preceding design points. Thus

$$Var_{[m-1]}(r_u(x)) \leq \sigma_{\delta_u}^2 + \sigma_{\epsilon_u}^2 = \sigma_u^2 \quad \text{and} \quad |Cov_{[m-1]}(r_u(x_1), r_u(x_2))| \leq \sigma_{\epsilon_u}^2.$$

Then, except close to the preceding design points, to a first order approximation,

$$E_{[m-1]}(y_u) \simeq x_u^{*T} \bar{\beta}_u, \quad Var_{[m-1]}(y_u) \simeq x_u^{*T} V_u x_u^* + \sigma_u^2 \quad (20)$$

and similarly for $y_{[m]u}$. If there are no covariances across components of production, then variance reduction for y_u by observing $y_{[m]u}$ depends only on the adjusted squared correlation, $\rho_{[m-1]}^2$, between y_u and $y_{[m]u}$, and the criterion (17) becomes

$$U_{y_u}(x_{[m]}) = \int \rho_{[m-1]}^2(y_{[m]u}, y_u(x))w(x)dx \quad (21)$$

where $w(\cdot)$ incorporates plausibility calculations, for geology x , based on the values for the adjusted expectation and variance of each component of production given x . Therefore, we may, if we need to further simplify our calculations, base plausibility evaluations on (20). For example, a simple approximation would be to make $w(\cdot)$ a decreasing function of $c(\cdot)$ defined as

$$c(x) = \max_u \frac{(E_{[m-1]}(y_u(x)))^2}{Var_{[m-1]}(y_u(x))} \simeq \max_u \frac{x_u^{*T} \bar{\beta}_u \bar{\beta}_u^T x_u^*}{x_u^{*T} V_u x_u^* + \sigma_u^2} \quad (22)$$

though, usually, we would refine c , taking account of the various plausibility measures that we have introduced.

Approximating $Cov_{[m]}(y_u, y_{[m]u})$ by $x_u^{*T} V_u x_{[m]u}^*$, (21) reduces to the approximation

$$U_{y_u}(x_{[m]}) \simeq \frac{x_{[m]u}^{*T} V_u W_u V_u x_{[m]u}^*}{x_{[m]u}^{*T} V_u x_{[m]u}^* + \sigma_u^2} \quad (23)$$

where the matrix W_u is given by

$$W_u = \int x_u^* x_u^{*T} \frac{w(x)}{Var_{[m-1]}(y(x))} dx.$$

Hence, our overall design criterion in (18) is given by the approximation

$$U(x_{[m]}) \simeq \sum_u l_u \frac{x_{[m]u}^{*T} V_u W_u V_u x_{[m]u}^*}{x_{[m]u}^{*T} V_u x_{[m]u}^* + \sigma_u^2}. \quad (24)$$

The value of approximation (24) is that each matrix W_u is only calculated once for each design search, reducing the design problem to optimisation of a relatively simple function.

Thus, at each stage we may maximise (24) to select our next design point as the choice which maximises, approximately, the overall reduction in uncertainty about the surface, as weighted by the relative plausibilities of the various points as potential history matches. Alternately, we may use (24) for preliminary screening, to identify a variety of potential design choices, for each of which we may then evaluate a more precise form of the design criterion.

The approximation works as long as the variances in terms $\beta_u^T x^*$ are large compared to variances in $r(x)$, and the design points are fairly well spaced, so that the covariances that we have ignored are of second order. When we have largely reduced the systematic variation, then variation around the fitted surface will be as important as variation due to the fitted surface. Therefore, at each design choice, we compare the magnitudes of the various variation terms. When we decide the approximation is no longer valid, then we have two choices.

Firstly, we may judge that we have concentrated our beliefs to a sufficiently small region of highly plausible matches for the components of production that it is feasible to search for the match directly within this region, either using direct numerical methods or by refining our design criterion. Secondly, we may decide that either we still have a comparatively large region to search or that we have not identified the match closely enough at this stage. In this case, we may refit the prior description by resampling on the fast simulator, over the sub-region of highest plausibility. This leads to a revised choice of active variables and a reassessment of regression coefficients and residual variances. We now continue our search on the full simulator, as before, but with the refitted description. We illustrate these considerations using our example.

6.3 Example: choosing design points

For simplicity, we take $l_u = 1$ for all u in (18). The fast simulator prior description equations, shown in table 3, suggest a natural grouping of the output variables into three groups H_1 , H_2 and H_3 where H_i is the collection of production variables at well i . For all output variables in the same group, we use the same weight function $w(x)$ which is the inverse of the maximum magnitude of the possibility ratio (adjusted expectation divided by adjusted standard deviation), the maximum being taken over the variables in the group.

We maximise the overall design criterion (24) with respect to the input variables using the **constr** function in MATLAB (see MathWorks Inc. 1992), which is a gradient method with line search. We start the optimisation at the origin which is the centre of the geology space being searched and accept the point found by **constr** as the next design point. We acknowledge that this may well not be the global maximum of $U(x)$ but note that many of the points found have been on the boundary.

We found 10 design points sequentially and then halted because there seemed to be little further change in the possibility plots we examined; in fact there is not much change after 6. Figure 5 shows the weight function $w(x)$ for each group of outputs, as a function of the active variables for the group. We do not show points for which $w(x)$ is less than 2. From versions of these plots, we concluded that we had narrowed the likely region for x_h to the following sub-region:

Well 1	$x_5 \in [-.1, .4]$	$x_{10} \in [.1, .5]$	
Well 2	$x_3 \in [0, .5]$	$x_6 \in [-.1, .4]$	$x_8 \in [.2, .4]$
Well 3	$x_2 \in [.2, .5]$	$x_6 \in [.2, .5]$	$x_7 \in [-.2, .5]$

and combining the last two gives us a tighter region $x_6 \in [.2, .4]$.

To see if we would make further progress in locating x_h we applied the entire procedure to the sub-region. The first stage was to quantify knowledge about the fast simulator in the sub-region using a new latin hypercube design. The results are shown in table 6 which corresponds to table 3. Time 2 gor output variables for both wells 2 and 3 have been dropped. At well 3 the fast simulator showed no variation in time 2 gor over the sub-region and at well 2 time 2 gor would have required a different set of active variables to any other output variable and it seemed simpler just to ignore it. The structure has changed considerably by moving to the sub-region. There are now four groups of output variables, all well 1 variables, bhp at well 2, bhp at well 3, and gor at time 6 at wells 2 and 3. The corresponding collections of active variables have also changed considerably. The residual variance components are also substantially reduced from those for the original region.

Working in the smaller region is similar to working in the full region; we do not see much change after about six or seven design points. Examination of plots like those in figure 5 and consideration of the possible sets for the four groups and their intersections suggests that x_h lies in the following region:

Well 1	$x_5 \in [.2, .35]$	$x_{10} \in [.175, .275]$	
Well 2 bhp	$x_1 \in [-.5, .5]$	$x_6 \in [.2, .3]$	$x_8 \in [.2, .4]$
Well 3 bhp	$x_1 \in [.2, .5]$	$x_6 \in [.2, .5]$	$x_7 \in [-.05, .5]$
Wells 2 and 3 gor at time 6	$x_1 \in [.2, .35]$	$x_3 \in [.1, .5]$	$x_9 \in [-.4, .5]$

The ranges of x_8 and x_7 are strongly dependent on x_1 . Thus, combining the last three lines of the table gives us a tighter region on x_8 and x_7 : $x_8 \in [.2, .3]$ and $x_7 \in [0, .45]$. The new region has a volume about a millionth of the volume of the original region. Each iteration has reduced the volume of the region to about a thousandth

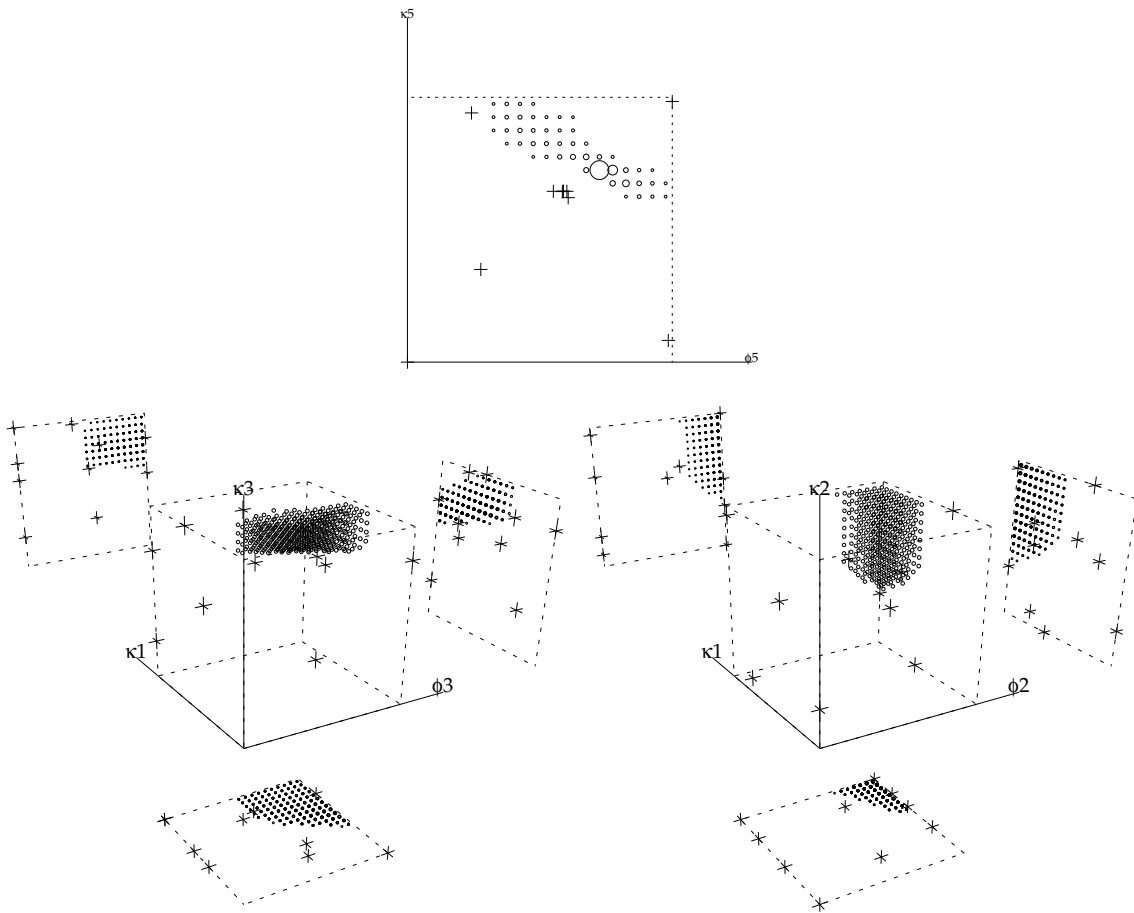


Figure 5: Plots (one for each well) showing location of possible history matches in the active variables for the well after ten design points (marked with +). Size of circles in the two-dimensional projections indicates the weight used for the point when computing the utility function for the next design point.

of its previous volume. Note that there is no reduction for x_4 , which appeared in none of our fast simulator equations, and little for x_9 which only appeared at the second iteration and with small coefficients.

Finally we show the quality of the matches for the production variables for *all* time points, not only the two used for the belief specifications and design calculations. We calculate the simulated production for the midpoint of our final region: $x = (.275, .35, .3, .0, .275, .25, .225, .25, .05, .225)$. The differences between history and simulated production for the midpoint of our original region, for the midpoint of our intermediate region and for the midpoint of our final region are shown in figure 6. Note that the final match is excellent over almost all time points which strongly supports our judgment that we are reducing to the right region.

7 Concluding comments

In this paper we have developed and applied a general approach for history matching. This approach extends readily to many similar inverse problems involving sequential search for expensive, high dimensional functions. Each stage in implementing this methodology raises interesting practical and technical questions, which cross the boundaries between many areas of statistical investigation. Much work remains in order to clarify these questions. However, we hope that we have demonstrated our basic thesis, namely that expert qualitative and quantitative judgments, in combination with careful analysis on fast versions of the simulator, form a natural

Equation	$\sigma_{\tilde{\delta}}$	$\sigma_{\tilde{\epsilon}}$	$\tilde{\theta}$
Well 1 in region 5			
$\widetilde{\text{bhp}}_2 = -181 + 80x_5 + 458x_{10} - 105Q(x_{10}) + 7x_9$	1	2	8
$\widetilde{\text{bhp}}_6 = -310 + 256x_5 + 625x_{10} - 154Q(x_{10}) + 14x_9 - 26Q(x_9)$	3	6	11
$\widetilde{\text{gor}}_2 = 23 - 84x_5 - 27x_{10} + 24Q(x_5) + 84x_5x_{10}$	1	1	26
$\widetilde{\text{gor}}_6 = 414 - 446x_5 - 83x_{10} - 15x_9 + 56Q(x_9)$	7	12	12
Bhp at well 2 in region 3			
$\widetilde{\text{bhp}}_2 = -170 + 439x_8 + 61x_1 + 179x_6$	9	0	
$\widetilde{\text{bhp}}_6 = -139 + 109x_1 + 543x_8$	23	0	
Bhp at well 3 in region 2			
$\widetilde{\text{bhp}}_2 = -73 + 75x_1 + 89x_7 + 1948x_6$	7	5	25
$\widetilde{\text{bhp}}_6 = -6 + 98x_1 + 101x_7$	17	3	1
Time 6 gor at wells 2 and 3			
$\widetilde{\text{gor}}_2 = 64 - 164x_1 - 75x_9 - 136x_3$	16	3	3
$\widetilde{\text{gor}}_3 = -4 - 72x_1 - 31x_9 - 60x_3$	6	4	1

Table 6: Equations approximating the fast simulator for the sub-region.

basis for history matching, which may be tractably handled by Bayes linear search strategies.

Acknowledgments

We are especially grateful to Victor Alcobia, David Brindle, Al Modine, Ann Muggeridge, Bob Parish and Alan Watkins at Scientific Software-Intercomp (UK), our industrial collaborators, for their support and for making available their reservoir simulator and adaptive history matching software. We also wish to acknowledge help from colleagues: Alan Craig, a co-worker on our Science and Engineering Research Council grant, for many insights into fluid dynamics and simulator technology; David Wooff, for assistance in using the subjectivist package $[B/D]$ to verify some calculations; and James Blowey for discussions on parts of section 6.2. This work has been supported by a United Kingdom Science and Engineering Research Council grant under the Complex Stochastic Systems Initiative.

References

- Anterion, F. (1988). History matching, a one day long competition: Classical approaches versus the gradient method. Presented at the first international forum on reservoir simulation, Alpbach, Austria.
- Bernardo, J.-M. et al. (Eds.) (1992). *Bayesian Statistics 4*. Oxford University Press.
- Farrow, M. and M. Goldstein (1992). Reconciling costs and benefits in experimental design. See Bernardo et al. (1992).
- Farrow, M. and M. Goldstein (1993). Bayes linear methods for grouped multivariate repeated measurement studies with application to crossover trials. *Biometrika* 80(1), 39–59.
- Goldstein, M. (1985). Temporal coherence. In J.-M. Bernardo et al. (Eds.), *Bayesian Statistics 2*, pp. 231–248. North Holland.
- Goldstein, M. (1988). The data trajectory. In J.-M. Bernardo et al. (Eds.), *Bayesian Statistics 3*, pp. 189–209. Oxford University Press.

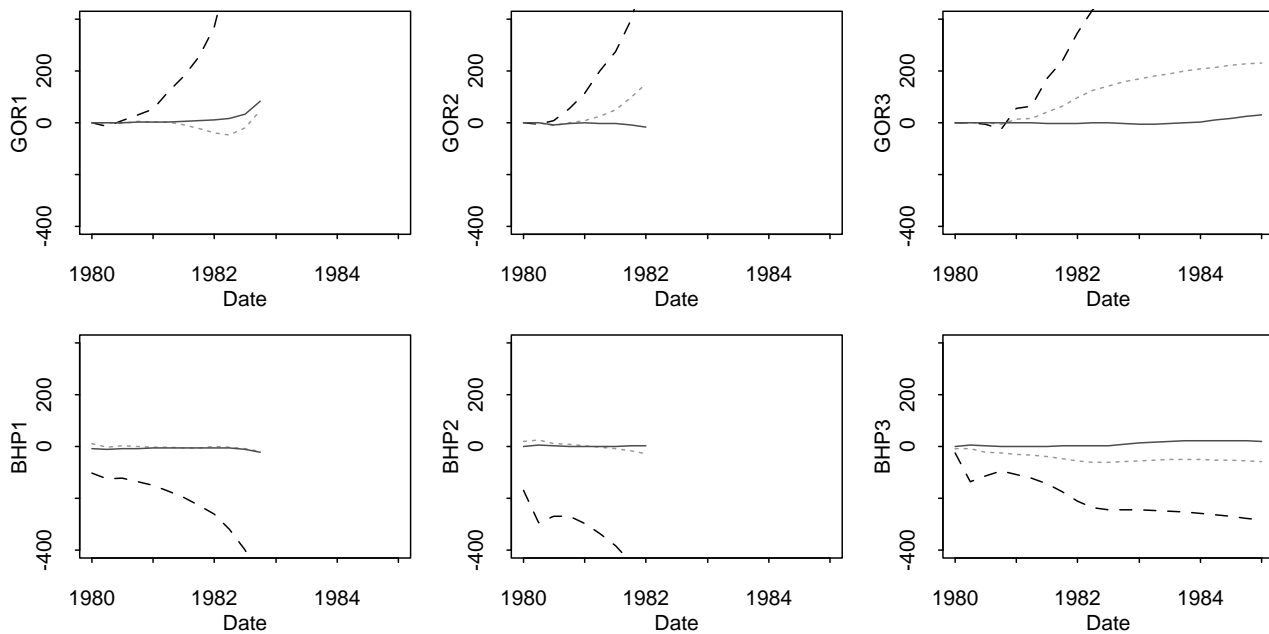


Figure 6: Errors in simulated production for midpoints of original (dashed lines), intermediate (dotted lines) and final (solid lines) geology regions.

Goldstein, M. (1994). Revising exchangeable beliefs: subjectivist foundations for the inductive argument. In P. Freeman and A. Smith (Eds.), *Aspects of Uncertainty: A Tribute to D. V. Lindley*. Wiley.

MathWorks Inc. (1992). *MATLAB: High-Performance Numeric Computation and Visualisation Software Reference Guide*. MathWorks Inc.

Mattax, C. C. and R. L. Dalton (Eds.) (1990). *Reservoir Simulation*. Society of Petroleum Engineers.

McKay, M., W. Conover, and R. Beckman (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239–245.

O’Hagan, A. (1978). Curve fitting and optimal design for prediction (with Discussion). *Journal of the Royal Statistical Society* 40(1), 1–42.

O’Hagan, A. (1992). Some Bayesian numerical analysis. See Bernardo et al. (1992), pp. 345–363.

O’Sullivan, F. (1986). A statistical perspective on ill-posed problems (with Discussion). *Statistical Science* 1, 502–527.

Sacks, J., W. Welch, T. Mitchell, and H. Wynn (1989). Design and analysis of computer experiments. *Statistical Science* 4, 409–435.

Smith, J. A. (1993). *Looking for Nothing*. Ph. D. thesis, University of Durham.

Watkins, A. J., R. G. Parish, and A. D. Modine (1992). A stochastic role for engineering input to reservoir history matching. Presented at Society of Petroleum Engineers 2nd Latin American Petroleum Exploration Conference, Caracas, Venezuela.

Rejoinder to discussion

We thank the discussants for their comments. Briefly, our response is as follows. As to the general issue of benefits of the Bayes linear approach, the basic considerations are that we have a simpler task of prior specification, and simpler calculations are required to revise our prior judgements, making possible analyses that might easily have been intractable under a full Bayes analysis.

History matching is a very complex problem, so that any comments that we make must be somewhat provisional. However, our current views are as follows. We fit linear terms rather than only a “nonparametric smoother” as the problem of searching in very high dimensions for the history match is computationally very hard, and the linear form suggests fast, tractable approximations which may substantially reduce the volume of the space that we have to search over. Further, nonparametric descriptions for data in high dimensions are very sensitive to mis-specifications, so that it appears to be simpler and safer to fit linear terms which account for substantial amounts of the variation, in ways which roughly accord with the judgements of the reservoir engineer, and use the nonparametric description for the residual variation. As we refit the prior description on reduced search regions, the balance between the parametric and nonparametric components may change, but often the more important feature of the refit is that various input variables which were not predictive in the original description become important when fitting on the reduced space, as happens for the example in the paper. For this reason, the random error component, which expresses the variation in the output quantities due to input quantities which are not explicitly included in our prior description, need not grow in importance as we continue our search.

As to our general objectives, in the paper, for simplicity, we considered an example where there was a precise solution that we could match. In the general problem, however, we must introduce a further level of modelling, to express the distance between the true history and the setting of the input variables which gives the closest approximation to this history. If the distance is small, then we may view the problem as finding the orthogonal projection of the true history into the space of simulator histories. In this formulation, the plausibility surface expresses the potential for each given input collection to identify the projection of the true history, so that the plausibility surface will only converge to zero if our judgements as to the magnitude of the distance between the true history and the output surface are at fault. Note that this would serve as an important diagnostic warning: if the closest approximation is poor, then this casts doubts on the reliability of the simulator model for predicting future reservoir performance, so that we should consider refining the simulator.

Finally, the suggestion to use a Bayesian synthesis approach to the problem is clearly inapplicable currently on grounds of computer time. Further, as computing power increases, so undoubtedly will the complexity of the reservoir models, so that increases in computer speed may not lead to shorter run times. However, a large part of our analysis is involved in substantially reducing the initially enormous search space for good solutions, and identifying relative plausibilities of values within the possible solution space. How we make use of the resulting collection of plausibilities for sensible reservoir prediction is an open question, which may not best be solved by seeking a single good match, and it may well be that when the search region has been sufficiently reduced, some form of full Bayes analysis might be helpful for identifying a sensible choice, or collection of choices, of inputs to consider for future predictive use.