UNCERTAINTY ASSESSMENT IN HISTORY MATCHING AND FORECASTING[1]

BJØRN KÅRE HEGSTAD AND HENNING OMRE
Department of Mathematical Sciences
Norwegian University of Science and Technology
Gløshaugen
N-7034 Trondheim
Norway

Abstract. A formalism for history matching in a Bayesian framework is presented. It gives a formal way to assess production forecast uncertainty. It may also reduce the complexity of the history matching problem by defining it in a lower dimensional parameter space, and thereby make the task less resource requiring.

1. Introduction

The ultimate goal of reservoir characterization is to forecast fluid flow and assess forecast uncertainty of production in a reservoir. A good reservoir model should share the same crucial features for fluid flow as the real reservoir, and ignore the less important one. By doing this one avoids bias in fluid flow forecast and the validity of the model is kept if production strategy changes by e.g. producing from a new well or starting injection of chemicals. If the model lacks the necessary small scale heterogeneity, the fluid flow forecast may be severely biased, see Jones et al. (1992). By making models which are reasonable in a geological sense, it is more likely to extrapolate the results to a new production strategy.

In the present approach small scale heterogeneity and large scale properties are modeled in a Bayesian framework, see also Omre and Tjelmeland (1996). In Bayesian models the model parameters are regarded as stochas-

[1]Preprint Statistics NO. 9/1996, Norwegian University of Science and Technology.
[1]To appear in proceedings from the 5th. International Geostatistical Congress, Wollongong 1996.

tic. These parameters may be general or large scale properties as mean porosity, mean permeability, intensity for shale units, net to gross ratios, spatial dependencies for various petrophysical parameters etc. The importance of these are demonstrated in Kittridge et al. (1989).

There are several sources of information about the reservoir under study. This may be general geological knowledge as evolution of the geological formation and experience from analogues. Other sources are reservoir specific observations. It may be point observations in wells as core plugs or well logs or so-called dynamic data as transient pressure data and production history. Some of these sources of information are easily included into a reservoir model but others are harder to include. Especially transient pressure and production history are complicated to include without making strong assumptions as e.g. the reservoir being homogeneous or radially symmetric.

To retain small scale heterogeneity one may divide the reservoir into grid blocks and assign petrophysical values to each block. One may then reproduce production history by altering the values in each individual block. This defines an extremely high dimensional problem and often the resulting realizations look unrealistic in a geological sense. The Bayesian approach may reduce the dimensionality drastically and ensure the reproduction of certain geological features.

In the present paper a general formalism for stochastic reservoir modeling with conditioning on production data is presented. An approach which exploits the formalism to obtain forecast uncertainty will be presented. Although the examples uses lognormal fields as reservoir model, marked point fields, Markov random fields, hybrid models etc. fit well into the formalism.

2. Uncertainty modeling

There are several ways to find a history matched reservoir description, see Bissell et al. (1992), Sultan et al. (1993), Ewing and Lin (1991), Parish et al. (1993) and Craig et al. (1994). Forecasting of production performance may be obtained from this match. However, assessment of the forecast uncertainty is important. Even a good history match may give a poor forecast or two equally good matches may diverge in the forecast.

Let $Q = Q(t)$ be the future production performance as production rate, gas oil ratio, transient pressure in wells etc.. Since it is unknown, it may be represented as a stochastic variable with associated probability density function (pdf) $f_Q(q)$. This pdf is a stochastic model expressing the forecast uncertainty. Let O denote the information available about the reservoir under study. The uncertainty after all available information is included is denoted $f_{Q|O}(q|o)$, the posterior pdf for production forecast uncertainty. This uncertainty may be displayed by sampling the posterior pdf. This paper will demonstrate a formalism to do this correctly.

For practical reasons split the information O into a dynamic part O_p as production history, and a static part O_w as core plugs, well logs etc. called well observations. Hence $O = (O_p, O_w)$.

Let $v[\cdot, s]$ be a fluid flow simulator being such that if a reservoir was known exhaustively $v[\cdot, s]$ would give the true reservoir performance for recovery strategy s. This is of course a simplification.

Let $r = r(x)$, with x being a spatial reference, be a full specification of the reservoir characteristics necessary for the simulator $v[r, s]$ to run. It contains the spatial distribution of e.g. permeability, porosity, initial saturation, initial pressure etc.. Since r is unknown, it may be represented as a stochastic variable with associated pdf $f_R(r)$. This pdf is called the prior pdf or the prior stochastic model for the reservoir characteristics. The reservoir specific observations are linked to the reservoir characteristics through the likelihood function $f_{O|R}(o|r)$. This function expresses how likely it is that a reservoir r could give the observations o. Its form may be determined in cooperation with reservoir engineers, petrophysicists and geologists. Combining this function with the prior pdf $f_R(r)$ yields the posterior pdf $f_{R|O}(r|o)$ by using Bayes rule:

$$f_{R|O}(r|o) = \left[\int f_{O|R}(o|r) f_R(r)\, dr\right]^{-1} f_{O|R}(o|r) f_R(r)$$

If r is a realization from $f_{R|O}(\cdot)$ one may obtain a realization q from $f_{Q|O}(\cdot)$ by $q = v[r, s]$. Hence the pdf $f_{Q|O}(\cdot)$ is uniquely determined by the pdf $f_{R|O}(\cdot)$ through the fluid flow simulator $v[\cdot, s]$.

3. The prior stochastic reservoir model

To fully specify r on a grid system in a full scale case, often $10^6 - 10^8$ values needs to be determined. This makes the reservoir characterization an extremely high dimensional problem. In order to make inference about R, a parametric model is convenient. One often determines the kind of model and the parameters to use on the basis of general geological knowledge of the reservoir under study. Given a set of parameters θ the parametric model fits well into a mathematical statistical framework, making it possible to construct adequate, flexible and complex stochastic models.

In general the parameters θ cannot be considered as completely known. There is uncertainty related to them as well, hence θ or parts of θ, will be considered as a stochastic vector Θ. The pdf for R can then be written as

$$f_R(r) = \int f_{R,\Theta}(r, \theta)\, d\theta = \int f_{R|\Theta}(r|\theta) f_\Theta(\theta)\, d\theta$$

with the pdf $f_{R|\Theta}(\cdot)$ being associated with heterogeneity uncertainty while the parameter prior pdf $f_{\Theta}(\cdot)$ is associated with the parameter uncertainty. They must be determined in cooperation with experienced geologists.

Conditioning on the observations O to obtain the posterior pdf $f_{R|O}(\cdot)$ can be done in several ways in this framework. This will be discussed below. It is important to note that a sample (r_i, θ_i), $i = 1, \ldots, n$ from the pdf $f_{R,\Theta}(\cdot)$ entails directly a sample from $f_R(\cdot)$ by only considering the r_i - entries.

4. History matching example

The prior model is a modified SPE test case. The reservoir dimensions are $10.000 \times 10.000 \times 100$ ft^3 and it contains three layers being 20 ft., 30 ft. and 50 ft. thick. The layers are numbered from top to base. See also figure 1. Each layer has constant properties with exception of the permeability $K(x) = (K_1(x), K_2(x), K_3(x))$ which is considered to be stochastic. The porosity is set to 0.3 and the transmissibility multipliers are 0.64 between layer 1 and 2 and 0.26 between layer 2 and 3. The transmissibility multipliers are the factors by which vertical permeability between the geological layers are reduced. The value 0.0 indicates a non-permeable horizon as e.g. a layer of shales. Initially the reservoir is saturated with oil and immobile water with a pressure of 5800 psi measured at a depth of 8400 ft. It is assumed that the oil-water capillary pressure is zero. The model is defined on a $50 \times 50 \times 15$ grid with 5 grid nodes vertically in each layer.

Permeability is frequently modeled as a lognormal field since this provides a reasonable trade off between reality and mathematical tractability. Hence the stochastic permeability model is lognormal and defined as follows:

$$
\begin{aligned}
K_1(x)|\mu_1 &\sim \text{lognorm } (\mu_1, \sigma_1^2) \\
& \qquad Y_1(x) = \log(K_1(x)|\mu_1) \sim \text{N } (\nu_1, 0.53^2) \\
& \qquad \mu_1 \sim \text{N}(350, 100^2) \\
K_2(x) &\sim \text{lognorm}(50, \sigma_2^2) \\
& \qquad Y_2(x) = \log(K_2(x)) \sim \text{N}(\nu_2, 0.78^2) \\
K_3(x)|\mu_3 &\sim \text{lognorm } (\mu_3, \sigma_3^2) \\
& \qquad Y_3(x) = \log(K_3(x)|\mu_3) \sim \text{N } (\nu_3, 0.62^2) \\
& \qquad \mu_3 \sim \text{N } (350, 100^2) \\
\text{Corr } \{Y_i(x'), Y_j(x'')\} &= \begin{cases} \text{Spherical } (3000, 3000, \frac{3000}{4}) & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}
\end{aligned}
$$

The arguments in the correlation are the ranges in the two horizontal and the vertical directions respectively. It is assumed that $(\mu_i, Y_i)^T$, $i = 1, 3$ are multinormal distributed. This joint distribution is fully defined from the relations above with $\sigma_i^2,\ i = 1, 2, 3$ and $\nu_i,\ i = 1, 2, 3$ implicitly defined. Hence the permeabilities in layer 1 and 3 are defined in a Bayesian setting with the same prior on the expected permeabilities. The model parameters $\theta = (\mu_1, \mu_3)$ are believed to have considerable uncertainty and large influence on the production performance. Hence these are not held constant but regarded as stochastic. The expected permeability in layer 2 is found to have small impact on production performance and hence kept constant.

This stochastic model now fully specifies the pdf $f_{R|\Theta}(\cdot)$ representing the heterogeneity uncertainty and $f_\Theta(\cdot)$ representing the model parameter uncertainty. A realization of the permeability field $K(x)$ together with all the predetermined reservoir characteristics gives a realization of the stochastic variable R described previously.

The producer and injector are situated in diagonally opposite corners of the reservoir, see figure 1. The injector well is perforated in layer 1 and injects gas, which is not allowed to dissolve into the oil. The producer well is perforated in layer 3 as indicated in figure 1. Three different configurations for the observation wells are considered with zero, five and thirteen wells respectively. In the five well case a traditional five spot pattern is used and in the thirteen wells case the additional eight wells are placed symmetrically in between the first five wells, see figure 2.

The fluid flow simulation is performed on a $10 \times 10 \times 15$ grid and the permeability upscaling is done by geometrical averaging. The program system ECLIPSE 100 is used as fluid flow simulator and together with upscaling it defines the function $v[\cdot, s]$ described previously.

If the well observations O_w are exact observations of permeability in the corresponding grid nodes, a good approximation can be found for the parameter well posterior $f_{\Theta|O_w}(\cdot)$. A factorization of the posterior $f_{R,\Theta|O}(\cdot)$ is then:

$$\begin{aligned} f_{R,\Theta|O}(r, \theta|o) &= f_{R,\Theta|O_w,O_p}(r, \theta|o_w, o_p) \\ &= \text{const} \times f_{O_p|R,\Theta,O_w}(o_p|r, \theta, o_w) \times f_{R|\Theta,O_w}(r|\theta, o_w) \times f_{\Theta|O_w}(\theta|o_w) \end{aligned} \tag{1}$$

with the parameter well posterior:

$$f_{\Theta|O_w}(\theta|o_w) = \text{const} \times f_{O_w|\Theta}(o_w|\theta) \times f_\Theta(\theta)$$

Here $f_\Theta(\cdot)$ is the parameter prior as previously defined and the well-likelihood function $f_{O_w|\Theta}(\cdot)$ is the likelihood of observing o_w given that θ is the parameter. The pdf $f_{R|\Theta,O_w}(\cdot)$ represents the parametric model exactly reproducing o_w, while the well conditioned production likelihood $f_{O_p|R,\Theta,O_w}(\cdot)$ is the likelihood of observing o_p given that r is the true reservoir and that o_w is reproduced.

In general one does not expect to reproduce O_p exactly. Firstly observation errors may occur in the production characteristics, moreover neither the fluid flow simulator $v[\cdot, s]$ nor the reservoir model is exact. Hence the corresponding production likelihood allow some variance around O_p typically proportional to $\exp(-\Delta^2)$, where Δ^2 is the sum of squared residuals. More complicated and sophisticated likelihood functions may be chosen without additional cost. In this example the likelihood is of the traditional form with Δ^2 being the sum of squares of production rate and bottom hole pressure residuals after 31, 455 and 911 days. Also the difference in breakthrough time during the matching period is included in Δ^2.

One may now sample from (1) by first sampling a θ_i from $f_{\Theta|O_w}(\cdot)$, then generating a reservoir realization r_i from the parametric model $f_{R|\Theta,O_w}(\cdot)$. Finally this realization is accepted with a probability proportional to $f_{O_p|R,\Theta,O_w}(\cdot)$, see e.g. Smith and Gelfand (1992). The accepted sample (r_i, θ_i) will then be from the pdf $f_{R,\Theta|O}(\cdot)$. Note that one does not need to have an analytical expression for the parameter well posterior. One just has to be able to sample it by e.g. Markov chain Monte Carlo techniques

The evaluation is based on the ability to match and forecast results from the "true" reservoir generated from the model. Hence problems related to mismatch between model and reality are avoided. The "true" reservoir $r^o(x)$ is generated from the prior model $f_R(\cdot)$. The model parameters turned out to be $\theta = (200, 500)$. The production history for the first ten years is computed by using $v[r^o(x), s]$. See figure 4 where $v[r^o(x), s]$ is displayed as black circles superposed with a white line. Only the first 2.5 years (911 days) of production is regarded as known and hence contained in o_p. Breakthrough was not observed within this period.

The uncertainty in the prior model $Q(t) = v[R(x), s]$ is explored with $f_R(r) = \int f_{R,\Theta}(r, \theta)\, d\theta$. In figure 3a realizations are displayed. There is a large variation in both production level and breakthrough time. The influence of the model parameters $\theta = (\mu_1, \mu_3)$ relative to heterogeneity uncertainty is explored by displaying the production rate $[Q(t)|\Theta = \theta^i] = v[(R(x)|\Theta = \theta^i), s]$, $i = 1, 2, 3$ in figures 3b , 3c and 3d. In each figure the expected permeability θ in layer 1 and 3 are kept constant, hence all variation is due to heterogeneity uncertainty. The figures demonstrates the importance of the model parameters on both breakthrough time and production level in this example. The relative contribution to total uncertainty from heterogeneity uncertainty seems to be around 25%, hence parameter uncertainty contributes with approximate 75% in accordance with the results in Lia et al. (1996).

In figure 4a, 4b, and 4c realizations from the posterior model $f_{Q|O}(\cdot)$

are displayed. The reservoir realizations are conditioned on zero, five and thirteen wells respectively. The uncertainty is drastically reduced compared to the prior model. All realizations gives a reasonable good match up to 2.5 years for which history is matched, but forecast seems somewhat skewed in the zero well and five well cases. This is mainly due to influence of the parameter prior which is not centered around the true value of the parameter θ.

The results indicate a fairly large influence of the model parameters on the production performance. Hence further exploration of parameter uncertainty is interesting. An alternative expression to (1) is the factorization

$$f_{R|\Theta,O}(r,\theta|o) = \text{const} \times f_{R|\Theta,O_w,O_p}(r|\theta,o_w,o_p) \times f_{\Theta|O_w,O_p}(\theta|o_w,o_p) \tag{2}$$

Notice the change in the last term compared to (1). The model parameters are conditioned on both well observations as previously and the production history. This pdf can be factorized as

$$\begin{aligned}
&f_{\Theta|O_w,O_p}(\theta|o_w,o_p) = \text{const} \times f_{O_p|\Theta,O_w}(o_p|\theta,o_w) \times f_{\Theta|O_w}(\theta|o_w) \\
&= \text{const} \times \int f_{O_p|R,\Theta,O_w}(o_p|r,\theta,o_w) \times f_{R|\Theta,O_w}(r|\theta,o_w)\, dr \times f_{\Theta|O_w}(\theta|o_w) \\
&= \text{const} \times E_{R|\Theta,O_w}[f_{O_p|R,\Theta,O_w}(o_p|R,\theta,o_w)] \times f_{\Theta|O_w}(\theta|o_w)
\end{aligned}$$

where $f_{\Theta|O_w}(\cdot)$ and $f_{O_p|R,\Theta,O_w}(\cdot)$ are as previously defined and $E_{R|\Theta,O_w}[\cdot]$ is the expectation taken with respect to $f_{R|\Theta,O_w}(\cdot)$. In the lognormal case the pdf $f_{\Theta|O_w}(\cdot)$ could be calculated. This is not possible with the parameter production and well posterior $f_{\Theta|O_w,O_p}(\cdot)$. It may be estimated by sampling, a procedure requiring several fluid flow simulation runs.

In figures 5a through 5c, the well posterior $f_{\Theta|O_w}(\cdot)$, the expected well conditioned production likelihood $E_{R|\Theta,O_w}[f_{O_p|R,\Theta,O_w}(o_p|R,\theta,o_w)]$ and the production and well posterior $f_{\Theta|O_w,O_p}(\cdot)$ are displayed in the zero well case. Since there is no well $f_{\Theta|O_w}(\theta|o_w) = f_\Theta(\theta)$ here. Figure 5b demonstrates the influence of O_p on the parameter uncertainty. The posterior pdf $f_{\Theta|O}(\cdot)$ displayed in figure 5c could be compared to figure 5a displaying $f_\Theta(\cdot)$, and it is obvious that the production data bring information about the model parameters Θ.

In figures 6a through 6c the same distributions are displayed but now conditioned on five wells. The contribution of O_w is seen from the reduction of uncertainty from figure 5a to 6a. The influence of O_p in figure 6b is still important, but smaller than in figure 5b. Adding more wells will reduce the influence of O_p. The posterior pdf $f_{\Theta|O_w,O_p}(\cdot)$ will of course be narrower with five wells than without wells since more information is available.

The influence of the model parameters of production performance may be further explored by considering an approximate expression for the pdf $f_{R,\Theta|O}(r,\theta|o)$:

$$f_{R,\Theta|O}(r,\theta|o) \approx \tilde{f}_{R,\Theta|O}(r,\theta|o) = \text{const} \times f_{R|\Theta,O_w}(r|\theta,o_w) \times f_{\Theta|O_w,O_p}(\theta|o_w,o_p)$$

Note that the realizations from $\tilde{f}_{R,\Theta|O}(r,\theta|o)$ is conditioned on production history through the parameters only, not through realizations if R.

In figures 7a through 7c realizations from $\tilde{f}_{Q|O}(q|o)$ associated with $\tilde{f}_{R,\Theta|O}(r,\theta|o)$ are displayed. One clearly see the same qualitative behavior as in the exact sampling approach in figure 4 but with larger variance. Note in particular that the forecast uncertainty for time after 911 days are very comparable. Although production history is included only in the parameter posterior, a significant reduction of forecast uncertainty compared to prior uncertainty displayed in figure 3a is observed. Hence the impact of the low dimensional model parameters, θ, on production forecast is considerable. This may be used to reduce the dimensionality of the history matching problem in order to obtain predictions more efficiently.

Exploration of normal models have also been performed, see Hegstad and Omre (1996), but it is not documented here. It shows a behavior more linear in the parameter θ, hence making the uncertainty assessment easier, and forecast uncertainties smaller.

5. Closing remarks

A Bayesian formalism for stochastic reservoir modeling has been presented. It gives a formal way to assess production forecast uncertainty. It may also reduce the complexity of the history matching problem by defining it in a lower dimensional parameter space. One may find a posterior distribution of the model parameters conditioned on both well observations and production data, making it easier to get realizations giving a good history match.

Acknowledgements

The research is done as a part of the EU-project Joule/ Reservoir Engineering/ Topic 5. The funds are provided by the Norwegian Research Council. We are greatful to Geomatic AS, Stavanger, Norway who kindly gave us their source code for fast simulation of 3D Gaussian fields from the STORM (STochastic Reservoir Modeling) system.

References

Bissell, R., L. Killough, and Y. Sharma (1992). Reservoir History Matching Using the Methods of Gradients on a Workstation. SPE 24265. Presented at the SPE European Petroleum Computer Conference, Stavanger, Norway, May 1992.

Craig, P., M. Goldstein, A. Seheult, and J. Smith (1994). Bayes Linear Strategies for History Matching of Hydrocarbon Reservoirs. In *5. International Meeting on Bayesian Statistics, Alicante*, pp. 531–563.

Eclipse 100, Reference manual, 94a release. Intera Information Technologies Ltd.

Ewing, R. and T. Lin (1991). A class of parameter estimation techniques for fluid flow in porous media. *Adv. Water Resources 14*(2), pp. 89–97.

Hegstad, B.K. and H. Omre (1996). History matching in stochastic reservoir models. results from a simulation study. Technical report, Norwegian University of Science and Technology.

Jones, A., G. Verly, and J. Williams (1992). What reservoir characterisation is required for predicting waterflood performance in a high net-to-gross fluvial enviroment? In J. Aasen, E. Berg, A. Buller, O. Hjelmeland, R. Holt, J. Kleppe, and O. Torsæter (Eds.), *North Sea Oil & Gas Reservoirs- III*, pp. 223–232. The Norwegian Institute of Technology: Kluwer Academic Publishers.

Kittridge, M., L. Lake, F. Lucia, and G. Fogg (1989). Outcrop-Subsurface Comparisons of Heterogeneity in the San Andreas Formation. *Society of Petroleum Engineers*. SPE 19596. Presented at the 64th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, San Antonio, TX, October 1989.

Lia, O., H. Omre, H. Tjelmeland, L. Holden, and T. Egeland (1996). Uncertainties in reservoir production forecasts. Accepted for publication in AAPG Bulletin.

Omre, H. and H. Tjelmeland (1996). Petroleum geostatistics. Invited lecture at the Fifth International Geostatistics Congress, Wollongong, Australia, 1996.

Parish, R., V. Calderbank, A. Watkins, A. Muggeridge, A. Goode, and P. Robinson (1993). Effective History Matching: The Application of Advanced Software Techiques to the History-Matching Process. *Society of Petroleum Engineers*. SPE 25250. Presented at the 12th Symposium on Reservoir Simulation, New Orelans, LA, 1993

Smith, A. and A. Gelfand (1992, May). Bayesian Statistics Without Tears: A Sampling-Resampling Perspective. *The American Statistician 46*(2), pp. 84–88.

Sultan, A., A. Ouenes, and W. Weiss (1993). Reservoir Description by Inverse Modeling: Application to EVGSAU Field. *Society of Petroleum Engineers*. SPE 26478. Presented at the 68th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, Huston, Texas, October 1993.

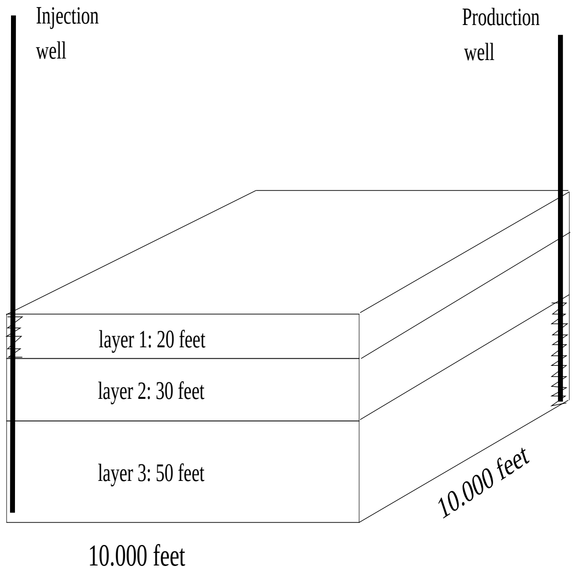


Figure 1. Outline of the reservoir model. Perforations of injection well and production well are indicated in layer 1 and 3 respectively.

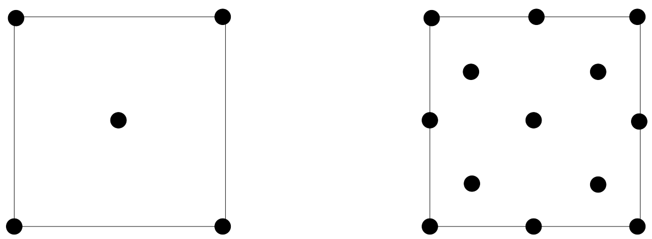


Figure 2. The reservoir seen from above: the configurations for the five wells and thirteen wells cases.

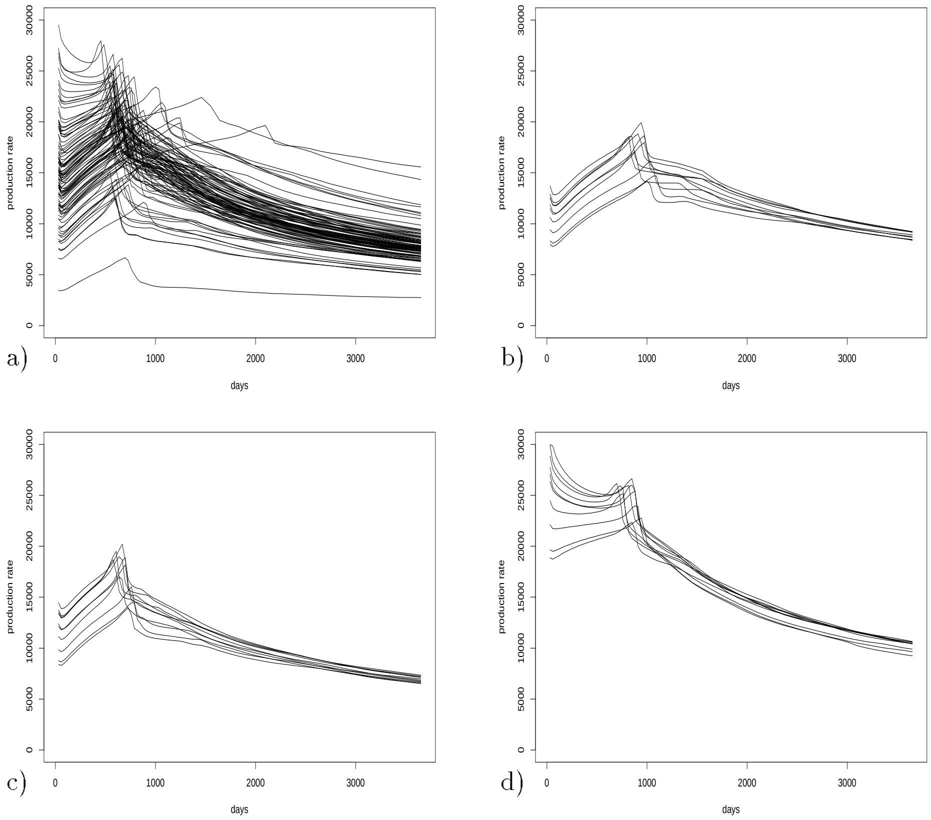


Figure 3. a) Hundred realization from the prior model $f_Q(q)$. b), c), d) Ten realizations conditioned on $\theta = (200, 275), (350, 275), (350, 725)$ respectively

a)

b)

c)

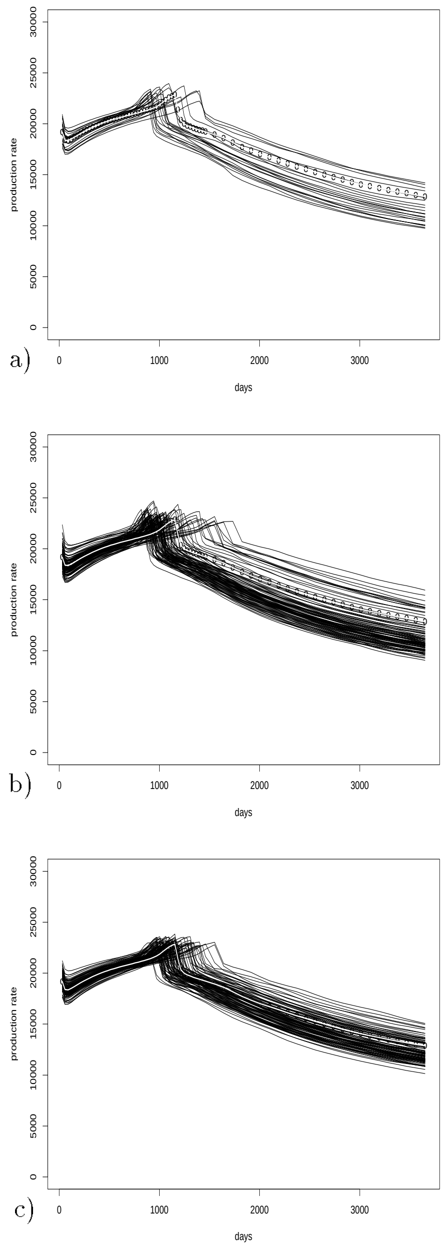


Figure 4. a) Realizations from the posterior pdf $f_{Q|O}(\cdot)$ conditioned on zero, five and thirteen wells in a), b) and c) respectively. The black circles superposed by a white line represents the true production rate.

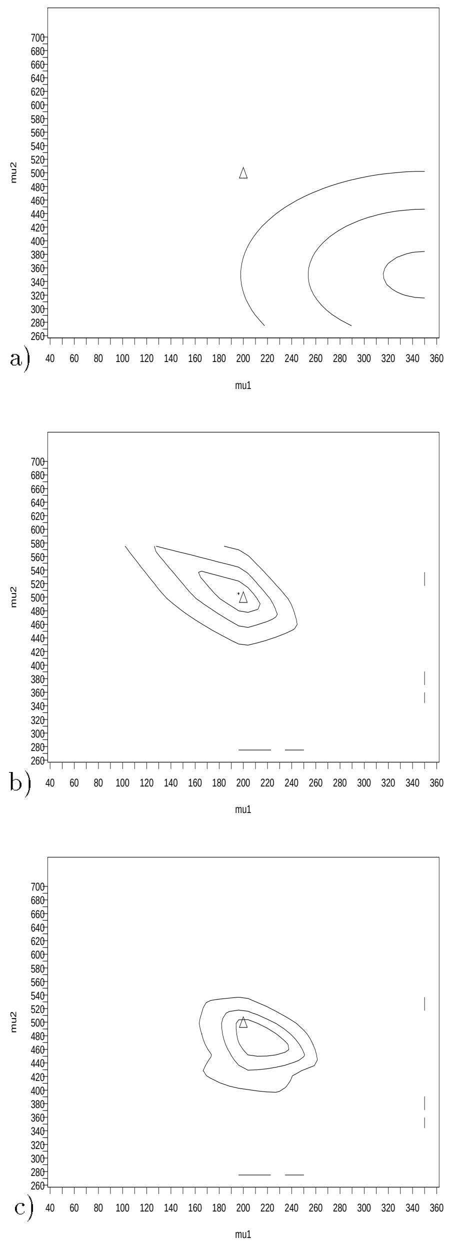
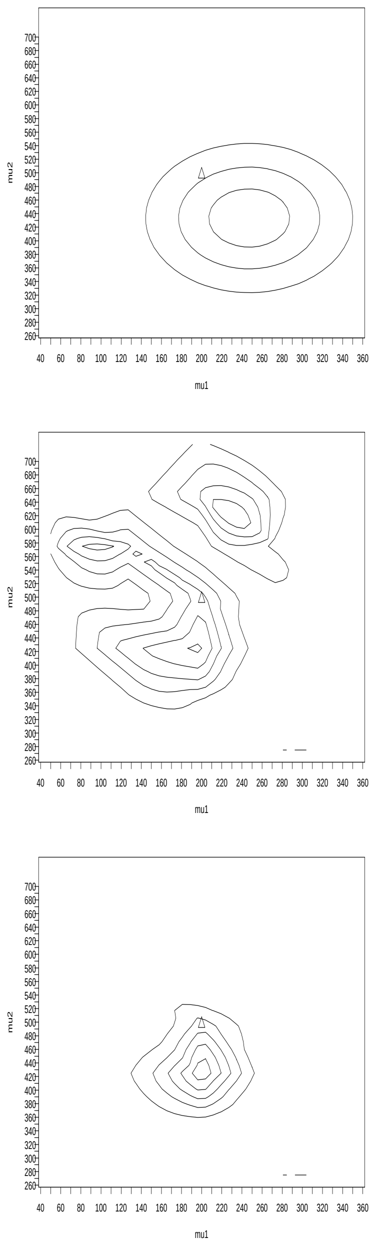


Figure 5. Parameter well posterior $f_{\Theta|O_w}(\cdot)$, well conditioned production likelihood $f_{O_p|\Theta,O_w}(\cdot)$ and parameter posterior $f_{\Theta|O_w,O_p}(\cdot)$ in a), b) and c) respectively for the zero well case. The triangle marks the true parameter $\theta = (200, 500)$.

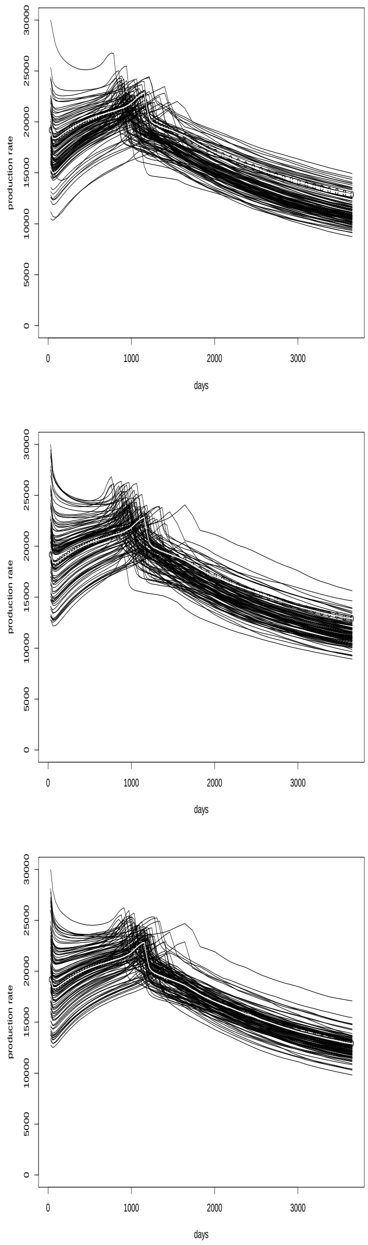


a)

b)

c)

Figure 6. Parameter well posterior $f_{\Theta|O_w}(\cdot)$, well conditioned production likelihood $f_{O_p|\Theta,O_w}(\cdot)$ and parameter posterior $f_{\Theta|O_w,O_p}(\cdot)$ in a), b) and c) respectively for the five well case. The triangle marks the true parameter $\theta = (200, 500)$



a)

b)

c)

Figure 7. Realizations from the approximate posterior model $\tilde{f}_{Q|O}(\cdot)$ conditioned on zero, five and thirteen wells in a), b) and c) respectively. The white line represents the true production rate.