8th International IFAC Symposium on
Dynamics and Control of Process Systems
Preprints Vol.2, June 6-8, 2007, Cancún, Mexico

STATE ESTIMATION OF A LARGE-SCALE SYSTEM IN THE PETROLEUM INDUSTRY: THE ENSEMBLE KALMAN FILTER FOR UPDATING RESERVOIR MODELS

Geir Nævdal [*] Alberto Bianco [**]
Alberto Cominelli [**] Laura Dovera [**]
Rolf Johan Lorentzen [*] Brice Vallès [*]

[*] *IRIS Petroleum, Thormøhlensgate 55, N-5008 Bergen, Norway*
[**] *Eni E&P, Milan, Italy*

Abstract: The ensemble Kalman filter is presented and it is shown how it is applied for updating models for fluid flow in oil reservoirs. A set of updated models are produced, which fit better to the available measurements and will be useful for further decision making.

Keywords: ensemble Kalman filter, reservoir simulation model, large-scale estimation

1. INTRODUCTION

For maximizing the production of an oil reservoir, improved modelling of the oil reservoir is important. Modelling of the oil reservoir is a difficult task, and obviously there is a problem with identification as the available observations are limited and the model is large. Today reservoir models that describe the fluid flow in the oil reservoir are used for decision support by evaluating different future prediction scenarios. One could for instance evaluate the effect of drilling a new well at a certain location, or evaluate different strategies for flooding the reservoir.

In the future, one envisions that oil production will be improved by introduction of methodologies from other industries, as process control, see e.g. (Jansen *et al.*, 2005). One interesting methodology that can be used for this purpose is closed-loop control. This will require online system identification. The Kalman filter is a likely choice for solving the system identification problem, but one has to take into account that the oil reservoirs are modelled using a large-scale non-linear system. Moreover, modelling of oil reservoir is a very complicated task due to the fact that the knowledge of important properties of the reservoir is very limited. For instance, important model parameters as permeability and porosity are not known. To acknowledge for this fact, the model must be improved from the available measurements by updating these model parameters.

The ensemble Kalman filter (EnKF) has shown to be a promising alternative for large-scale non-linear models. The EnKF was introduced for state and parameter estimation of reservoir models in a smart-well setting (Nævdal *et al.*, 2002). The use of this technique has subsequently attracted a lot of attention within petroleum sciences. However, there are only a very modest number of smart wells drilled yet. Therefore most of the research on using EnKF for updating reservoir models has focused on estimating certain parameters (permeability and porosity) in the reservoir model and not on providing an updated model in real-time. Here we present the EnKF as a tool for updating reservoir simulation models, and show some results from a recent study.

A demonstration of the potential using "smart-wells" (wells equipped with increased possibilities of measuring and control) combined with the EnKF and optimal control for improved oil recovery is presented in (Nævdal *et al.*, 2006), by showing its performance in a synthetic study. The current paper demonstrates the feasibility of the EnKF part, demonstrating that it does indeed improve the reservoir models when it is applied on a real data set.

The EnKF was introduced (Evensen, 1994) as an alternative to the extended Kalman filter for solving the problems encountered with large-scale non-linear models. The EnKF has been applied in many large-scale applications, in particular within earth sciences (Evensen, 2003), but the applications within control theory are scarce. Within the control sciences, other alternatives to the extended Kalman filter have been introduced for solving the challenges posed by non-linear models, as the unscented Kalman filter (UKF) (Julier and Uhlmann, 1995) and other related filters.While introducing the EnKF in this context, it is natural to compare the ideas behind the EnKF with the UKF.

The paper is organized as follows: first we introduce the EnKF. Thereafter we discuss reservoir modeling in more details, including the use of EnKF for state estimation of the reservoir model. Finally, we illustrate this practise by showing some results from a recent study.

2. THE ENSEMBLE KALMAN FILTER

As pointed out in the introduction, the extended Kalman filter faces difficulties if the non-linearity of the model becomes to severe. For instance the assumption on local linearity may fail. The extended Kalman filter also requires partial derivatives of all the state variables, and in some cases these may be difficult to obtain. For large-scale systems the storage of the covariance matrix for the state may also cause problems.

The EnKF is a Monte Carlo approach, where a set of sampling points is used to estimate the statistics of the estimated state of the model. In the presentation of the EnKF we start with a simple non-linear model, and will then point out certain modifications needed to cover the specifications of our problem afterwards. These modifications do not change anything essential in the filter equations.

Consider the non-linear model, f, which propagates the state x from time $k-1$ to time k according to

$$x_k = f(x_{k-1}) + w_k, \tag{1}$$

where w_k represent process noise assumed to follow a normal distribution with mean zero and covariance matrix Q, $w_k \sim N(O, Q)$. The subscripts denote the timesteps.

At time k an observation is available, and it is related to the state according to

$$z_k = Hx_k + v_k, \tag{2}$$

where H is the observation matrix and $v \sim N(0, R)$ is the measurement noise.

As is well known, the Kalman filter provides the estimated mean $\hat{x}_k$ and error covariance matrix, P_k, of the state at time k. For the EnKF, an approximation of the error statistics are available in form of an ensemble of realizations of the states of the model

$$\begin{bmatrix} x_{k,1} & x_{k,2} & \dots & x_{k,N} \end{bmatrix}, \tag{3}$$

where N is the number of ensemble members. This means that the mean $\hat{x}_k$ and covariance matrix P_k of the state can be approximated at a certain time step by

$$\begin{aligned} \hat{x}_k &\approx \frac{1}{N}\sum_{i=1}^{N} x_{k,i}, \\ P_k &\approx \frac{1}{N-1}\sum_{i=1}^{N}(x_{k,i} - \hat{x}_k)(x_{k,i} - \hat{x}_k)^T. \end{aligned} \tag{4}$$

The initialization of the EnKF is done by drawing a set of ensemble members from the distribution that represents the initial belief of the value of the state at time zero. The initial state is most commonly drawn from a normal distribution.

The EnKF is now evolving, as the Kalman filter, with a *prediction step* and an *update step*. In the prediction step the model is evaluated for each ensemble member by

$$\begin{aligned} x_{k,i}^f &= f(x_{k-1,i}^a) + w_{k,i} \\ w_{k,i} &\sim N(0, Q), \end{aligned} \tag{5}$$

where the superscript f denotes the quantities computed after the prediction step, and a the quantities after the update step and subscript i denotes the ensemble member. In the update step the states are adjusted to take into account an observation z_k assumed to have measurement noise following a normal distribution with mean zero and covariance matrix R. The equations for the update step are computed as follows:

$$\begin{aligned} K_k &= P_k^f H^T (H P_k^f H^T + R)^{-1}, \\ z_{k,i} &= z_k + v_{k,i}, \\ v_{k,i} &\sim N(0, R), \\ x_{k,i}^a &= x_{k,i}^f + K_k(z_{k,i} - Hx_{k,i}^f)), \end{aligned} \tag{6}$$

where P_k is defined as in Eq. 4. The Kalman gain K_k must be computed with care to reduce internal storage requirements during the computation. The similarity with the Kalman filter is obvious, but note that in the EnKF one avoids the computation both of the apriori and the aposteriori covariance matrices.

More details about the EnKF, with references to different variations in the implementation and applications and discussions about its theoretical and numerical properties, can be found in (Evensen, 2003; Evensen, 2006).

In the above model description (Eq. 1 and 2) a linear relationship between the state and the observed quantities was assumed. For the reservoir model that we will discuss later, this assumption does not hold. To be able to use the EnKF in this setting, we extend the state vector to include the observed quantities (that do not depend linearly on the states). We denote these quantities at step k by d_k. This construction allows us to consider cases when the observations are related to the state vector through a non-linear function as $d_k = h(x_k)$ for an arbitrary function h.

The derivation of the EnKF presented above considers the most common application where only the state of the model is modified in the update step. In petroleum science many of the models have poorly known model parameters. To improve the predictions it is necessary to tune these model parameters. This can be done using augmentation of the state variable. This idea is well known for the extended Kalman filter. Let p_k denote the parameters that are modified in the update step, where the index k refers to the parameter values obtained at time step k.

Now we introduce an extended state vector that combines both the additional observed quantities, d_k, and the parameters, p_k. The extended state becomes

$$\tilde{x}_k = \begin{bmatrix} x_k \\ d_k \\ p_k \end{bmatrix}.$$

We can now replace the ensemble of realizations of the states of the model with an ensemble of extended states. If we further ignore the process noise, as we will do while working with the reservoir model, then the prediction step (5) can be written as

$$\begin{bmatrix} x^f_{k,i} \\ d^f_{k,i} \end{bmatrix} = f(x^a_{k-1,i}, p^a_{k-1,i}),$$
$$p^f_{k,i} = p^a_{k-1,i}$$

The update step (6) is computed as before, but now on the ensemble of extended states. Further, H is replaced in an obvious manner by an extended matrix $\tilde{H} = [0 \quad I \quad 0]$. This means that both the states x_k and the parameters p_k are updated based on how they are correlated to the measurements $z_k = d_k$.

Let us now relate the EnKF to the UKF. The UKF was introduced to meet the challenges posed by the non-linearity of the model, and was first introduced in (Julier and Uhlmann, 1995). The basic idea behind the UKF is to get a better approximation of the covariance matrix of the state after the prediction step. The prediction step of the EnKF and of the UKF are quite similar, although the stochastic component of Eq. 5 is replaced by using sample points for the UKF. The ensemble members are, however, obtained differently. For the UKF, a set of sample points that represent the uncertainty in state estimate provided by the aposteriori mean and covariance matrix P^a_k is generated. They are selected such that they reproduce the mean, $\hat{x}^a_k$, and covariance matrix, P^a_k, exactly. To achieve this, one needs $2n + 1$ ensemble members, where n is the number of states, i.e. the length of the vector x. For large-systems, such a construction becomes infeasible.

The update step of the EnKF and UKF is different. As shown above, in the EnKF the Kalman gain is applied on each ensemble member individually. In the UKF, the Kalman gain is used to compute the updated mean and covariance matrix, and thereafter a new set of sampling points (corresponding to the ensemble members) are generated. This means that the UKF is based on computing the first and second order moments (mean and covariance matrix) to describe the probability distribution of the state vector both in the update and prediction step. This is relaxed in the EnKF since the mean and the covariances are only used to compute the Kalman gain, but all the ensemble members are treated individually both in the prediction and the update step. Previously, many studies have compared either EnKF or UKF with the extended Kalman filter. As far as we know very few studies compare UKF and EnKF, with (Nygaard *et al.*, 2006) as one exception, where the performance was compared on a small size model. Such comparisons are probably most interesting on medium size problems, where the number of simulations for the two approaches will be of the same order.

3. RESERVOIR MODELS AND UPDATING THEM USING ENSEMBLE KALMAN FILTER

As pointed out in the previous section, the type of models we can consider with the EnKF is of a very general type. Here, we discuss the application of the EnKF on reservoir models. For readers without background in this area, we give a first simplified exposition of the modelling behind fluid flow in oil reservoirs. For our practical implementation, we have used a commercial reservoir simulator, that may be viewed as a black-box model in the extreme, fulfilling the structure of Eq. 1.

One common approach of modelling fluid flow in an oil reservoir is to build a model based on Darcy's law, mass conservation, the petrophysical behavior of the system and certain empirical relations (as relative permeability, capillary pressure, etc.). The boundary conditions are given by the wellflow model.

Darcys's law

$$\mathbf{u}_p = -\frac{kk_{rp}}{\mu_p}\left(\nabla p_p - \gamma_p \nabla z\right)$$

gives a relationship between flow rate $\mathbf{u}_p$ and pressure gradient ∇p_p for each phase (denoted by subscript p). In our case, we consider a three phase system, with the phases oil, o, gas, g and water, w. In the above equation, k denotes the permeability, k_{rp} the relative permeability, μ_p the viscosity, z the depth and $\gamma_p = \rho_p g$, where ρ_p is the density of phase p and g the gravitational constant.

Combining Darcy's law with the mass conservation law for each phase, for a three phase black-oil system, yields the equations that should be solved by the reservoir simulator. To simplify the exposition, we present the equations for a two phase system with oil and water. Further, we also ignore the gravity terms. For details about modelling of reservoir flow, the reader could consult (Aziz and Settari, 1979).

The equations that need to be solved for a two-phase system without gravity effects are

$$\nabla \cdot \left(\frac{k_{ro}}{\mu_o B_o} k \nabla p_o\right) = \frac{\partial}{\partial t}\left(\frac{\Phi S_o}{B_o}\right) + q_o \qquad (7)$$

and

$$\nabla \cdot \left(\frac{k_{rw}}{\mu_w B_w} k \nabla p_w\right) = \frac{\partial}{\partial t}\left(\frac{\Phi S_w}{B_w}\right) + q_w. \qquad (8)$$

Here Φ denotes the porosity, S_p the saturation of phase p, B_p is the formation factor of phase p, and q_p the inflow or production of each phase. The terms q_p are further modelled using a wellflow model, which we will not go into any details about. Both the porosity Φ and permeability k are functions of space.

To get a closed system of equations we add a closure relation for the saturation

$$S_o + S_w = 1$$

and one by the capillary pressure between oil and water

$$P_{cow} = p_o - p_w = f(S_w, S_o).$$

The model is then discretized, and the system typically solved by a finite difference approach.

The modelling of an oil reservoir is a complex task where one needs to combine information from many sources. The largest structures of the reservoir are modelled using information from seismic surveys. After drilling exploration wells further information can be obtained by different tests taken during drilling. The petrophysical properties of the fluids in the reservoir, as the relative permeabilities and capillary pressures, can be determined using core samples from the reservoir. The core samples may also be used for determination of the permeability and porosity at these locations. Pressure tests will also be available. Still there will be lots of uncertainties in the model.

After the production is started from the reservoir, production data become available, and these can be used for further tuning of the model. Among the quantities that have large influence on the flow are porosity, Φ, and permeability, k, and it is natural to use the production data to get a better description of them. The spatial variability of the porosity Φ and permeability k, are typically assumed to vary randomly, and are described using geostatistical methods (Chilès and Delfiner, 1999).

The use of EnKF for updating reservoir models according to the information from production data has been a topic for extensive research during the past years. Some recent papers include (Gu and Oliver, 2006; Haugen *et al.*, 2006).

4. AN EXAMPLE

In this section we present some results from a recent study. A reservoir simulation model is built based on seismic and prior geological knowledge of the depositional effects. The simulation model is based on a grid of size $156 \times 77 \times 10$. Many of the grid blocks are inactive (to represent the complex geometrical layout of the reservoir), such that the total number of active grid cells is $25,669$. Each grid cell has a horizontal size of 125×125 meter. The height of each grid cell is chosen to produce layers containing the same type of geological structure. On the average, the height of the cells is three feet. The horizontal layers consist of sand and sandstones. Although they may have the same type of geological properties, a lot of spatial variability in porosity and permeability remains. The original model has seven different geological layers, but for better numerical modelling three of them are splitted thereby giving ten horizontal layers in the simulation model.

Prior to production, 10 appraisal wells have been drilled. These wells provided information about initial saturations and pressures in the reservoir. From the appraisal wells core samples have been obtained and analyzed. These core samples provided information about the porosity in certain points of the reservoir and also the petrophysical parameters of the oil reservoir.

The key uncertainty in the model is believed to be the porosity and permeability. Using the EnKF, we update the porosity by including it as the vector p_k in the extended state vector $\tilde{x}$. The porosity is varying in the reservoir, and p_k is a vector of length $25,669$, each representing the porosity in one grid block of the reservoir. The permeability is updated using a deterministic relationship with porosity, i.e. $k = f(\phi)$. This means that an update porosity values also leads to updated permeability values. In other studies, stochastic relationships between the permeability and porosity have been used (see (Haugen *et al.*, 2006)). Then the permeability values will be included in the state vector together with the porosity values.

The initial ensemble was generated by assuming an initial probability distribution of the porosity values. The mean and uncertainty in the porosity were computed for each layer using the available data from the appraisal wells. The distribution was further conditioned to fit the observed porosities in the appraisal wells. Such a conditioning was not done in the previous study (Haugen *et al.*, 2006). The porosity values for different grid cells are further assumed to be correlated using a variogram structure with horizontal range 4000 meter. This means that the porosity values in grid blocks close to each other are strongly correlated. The correlation is decreasing as the distance between the grid blocks increases, and are zero for grid blocks that have a distance of 4000 meters. The initial ensemble is generated by drawing 100 ensemble members independently from the described probability distribution.

The dynamic variables updated using the EnKF are the pressure, the water and gas saturations and the solution gas-oil ratio. The initial conditions for the dynamic variables are computed from gas-oil and water-oil contacts, that we assume to know exactly. Monthly production data are available from two production wells for three years. We are measuring oil production rate, water cut (fraction of water produced by the well), and gas-oil-ratio (GOR). This leads to a model with a state vector of length $128,350$. The well flow is modelled using fixed oil production rate which means that this is considered as a boundary condition of the model, and is not useful as a measurement.

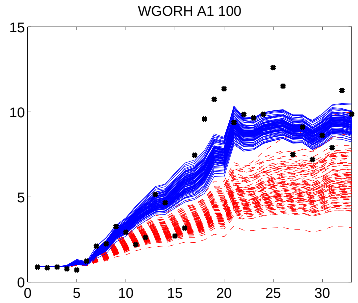

Fig. 1. GOR at well A1. The red dashed lines are predictions done based on the initial ensemble. The blue lines are the simulations using the updated porosity and permeability values obtained after assimilation of the production data. The black bullets are the measurements. The unit of the x-axis is months, of the y-axis is MScf/STB.

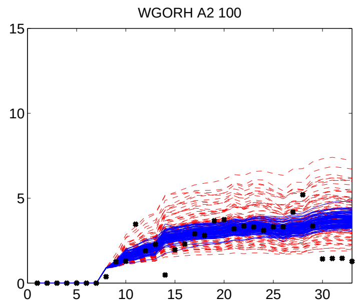

Fig. 2. GOR at well A2. The same notation as in Fig. 1.

As there are many sources of possible errors in obtaining the measurements, engineering judgement is used to assign the measurement noise used in the filter. This also means that while evaluating the results, one is mostly looking for a correct qualitatively behavior of the model after applying the EnKF.

For the GOR measurements, the uncertainty used for well A1 was 0.4 MScf/STB for GOR values larger than 6, 0.2 MScf/STB for lower values. For well A2 the uncertainty was 0.6 MScf/STB except for the last five measurements. The last five GOR measurements for well A2 were considered very unreliable, and an uncertainty of 600% was used. For the water cut measurements the uncertainty used for well A1 was 0.05, except for five measurements that were considered very unreliable and given an uncertainty of 600%. For well A2 the uncertainty of the water cut measurements was 0.0025 for all the measurements. There was no water breakthrough during the three year of production, leaving the GOR as the most interesting measurements.

The primary objective of this study is to verify that the EnKF is working properly for this problem, and in particular that the estimated porosity values are reasonable. The usual way of verifying reservoir models tuned to production data, is to show that they, in the end, have a reasonable match to the data. Here, we present the predicted values obtained from the 100 initial porosity fields.

As the observed GOR for the two wells are the most interesting observations, we restrict ourselves in providing results only for these two observations. The initial predictions, obtained by running the reservoir model from time zero through three years without applying the EnKF are shown in Figure 1 and 2. The porosity values (and through them, the permeability values) were then updated using monthly observations of water-cut and GOR from the two production wells over a three year period, using the EnKF with 100 ensemble members. This results in 100 updated porosity fields. The quality of these updated fields are evaluated by running simulations from time zero using each of the updated fields, again without applying the EnKF. The results are shown in Figure 1 and 2. We immediately observe that the variation in the simulated values are considerably reduced. This is the major effect obtained for well A2 (Fig. 2). For well A1 we also have achieved a better match to the measurements, in particular at later times. Moreover, the new simulations are much more in agreement with the observed values. This means that the new porosity fields are more suitable for improved reservoir management and better decision making.

A more thorough presentation of this study is available in (Bianco, 2006), and will also be presented in (Bianco *et al.*, 2007).

5. CONCLUSIONS

The EnKF is a promising method for updating properties as porosity and permeability of large-scale reservoir simulation models. The theory behind the EnKF has been outlined and the similarities and differences with another non-linear filter, the UKF has been discussed.

6. ACKNOWLEDGEMENTS

The authors affiliated at IRIS acknowledge financial support for this work from the Norwegian Research Council (projects 163376/S30 and 169315/S30) and industrial sponsors through the project “ROAW Phase - II: Continuous model updating of reservoir simulation models and improved reservoir management”.

Eni E & P are acknowledged for providing the necessary data for this study and giving permission to present the results.

REFERENCES

Aziz, K. and A. Settari (1979). *Petroleum Reservoir Simulation*. Applied Science. London.

Bianco, A., A. Cominelli, L. Dovera, G. Naevdal and B. Valles (2007). History matching and production forecast uncertainty by means of the ensemble Kalman filter: a real field application. In: *SPE Europec/EAGE Annual Conference and Exhibition*. Society of Petroleum Engineers. London, UK. SPE107161.

Bianco, A. (2006). Application of ensemble Kalman filter (EnKF). methodology for history matching and risk analysis.. Master’s thesis. Imperial College. London.

Chilès, J.-P. and P. Delfiner (1999). *Geostatistics. Modeling spatial uncertainty*. Wiley.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**(C5), 10,143–10,162.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics* **53**, 343–367.

Evensen, G. (2006). *Data Assimilation: The Ensemble Kalman Filter*. Springer.

Gu, Y. and D. S. Oliver (2006). The ensemble Kalman filter for continuous updating of reservoir simulation models. *Journal of energy resources technology - Transactions of the ASME* **128**(1), 79–87.

Haugen, V., L. J. Natvik, G. Evensen, A. Berg, K. Flornes and G. Nævdal (2006). History matching using the ensemble Kalman filter on a North Sea field case. In: *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers. San Antonio, Texas, USA. SPE102430.

Jansen, J. D., D. R. Brouwer, G. Nævdal and C. P. J. W. van Kruijsdijk (2005). Closed-loop reservoir management. *First Break* **23**, 43–48.

Julier, S. J. and J. K. Uhlmann (1995). A new approach for filtering nonlinear systems. In: *Proceedings of the American Control Conference, pages 1628-1632*.

Nævdal, G., D. R. Brouwer and J. D. Jansen (2006). Closed-loop waterflooding. *Computational Geosciences* **10**(1), 37–60.

Nævdal, G., T. Mannseth and E. H. Vefring (2002). Near-well reservoir monitoring through ensemble Kalman filter. In: *SPE/DOE Improved Oil Recovery Symposium*. Tulsa, Oklahoma. SPE75235.

Nygaard, G., G. Nævdal and S. Mylvaganam (2006). Evaluating nonlinear Kalman filters for parameters estimation in reservoirs during petroleum well drilling. In: *IEEE Conference on Control Applications*. München.