SCALE-CORRECTED ENSEMBLE KALMAN FILTER FOR OBSERVATIONS OF PRODUCTION AND TIME-LAPSE SEISMIC DATA

JON SÆTROM
Department of Mathematical Sciences, Norwegian University of Science and Technology

ABSTRACT

The Ensemble Kalman Filter (EnKF) is a Bayesian method for performing automatic and sequential history matching. High computational demands when performing reservoir fluid flow simulations often requires that an upscaling of the reservoir has to be carried out. This approximation is however known to introduce bias in production forecasts. The Scale-Corrected Ensemble Kalman Filter (SCEnKF) is a method that tries to account for this, which is an important feature when time-lapse seismic data is included, as this appear on a much finer scale than fluid flow simulation permits. Here we have tested the SCEnKF on a synthetic case study, where we have included both 4-D seismic and production data.

INTRODUCTION

Evaluation and prediction of the properties of a petroleum reservoir, such as permeability, hydrocarbon saturation and porosity, based on production history, well-logs and seismic data, implies solving a complex, ill-posed and non-linear inverse problem. The Ensemble Kalman Filter (EnKF) is a sequential Bayesian solution to this problem, which has shown promising results (Evensen, 2007).

Due to high computational demands when running fluid flow simulation on a very fine scale grid it is often necessary to reduce the dimension of the problem. This approximation however, is known to introduce bias, which should be accounted for (Omre and Lødøen, 2004). An extension to the EnKF, the Scale-Corrected EnKF (SCEnKF), was presented in (Lødøen and Omre, 2005), where the bias of coarse scale fluid flow simulation was corrected for. By including a so called calibration step, involving a small number of fluid flow simulations on both fine and coarse scale, they were able to correct for the loss in accuracy and precision when using an approximate

fluid flow simulator. This feature is especially important when including observations of 4-D seismic, as these usually appear on a much finer grid than the coarse scale grid used for fluid flow simulation.

NOTATION AND MODEL FORMULATION

Throughout this paper, the notation $\mathbf{x} \in \mathbb{R}^{n\times 1}$, will be used to denote both scalar, stochastic column vectors and realisations of dimension n, and $\mathbf{x}^T$ its transpose. Similarly, the notation $\Sigma \in \mathbb{R}^{m\times n}$, will denote matrices of dimension $m \times n$.

Consider a reservoir domain $\mathscr{D} \subset \mathbb{R}^3$, which has been discretised into a regular lattice $\mathscr{L}_{\mathscr{D}}$ of dimension n. The reservoir state properties within $\mathscr{L}_{\mathscr{D}}$ at time t will be denoted by $\mathbf{r}_t = [\boldsymbol{\kappa}^T, \boldsymbol{\phi}^T, \mathbf{s}_t^T, \mathbf{p}_t^T]^T \in \mathbb{R}^{n_r\times 1}$, where $\boldsymbol{\kappa}$ is log-permeability, $\boldsymbol{\phi}$ is porosity $\mathbf{s}_t$ is saturation and $\mathbf{p}_t$ is pressure at time t. Reservoir production properties at time t, such as gas-oil ratio (gor), bottom hole pressure (bhp) and oil production rates (opr), will be denoted by $\mathbf{q}_t \in \mathbb{R}^{n_q\times 1}$. For notational convenience we will let $\mathbf{x}_t = [\mathbf{r}_t^T, \mathbf{q}_t^T]^T$.

The state of the reservoir at time t, $\mathbf{x}_t$ is connected to the state of the reservoir at a previous time s through a possibly highly non-linear function $\boldsymbol{\omega}(\bullet)$, referred to as the fluid flow simulator. That is, $\mathbf{x}_t = \boldsymbol{\omega}_{t-s}(\mathbf{x}_s)$. Hence, if we consider evaluation of the reservoir at discrete timesteps t_k, $k \in \{0,\ldots,N\}$, this implies that the state of the reservoir at each timestep appears as a Markov process. The prior probability distribution function (pdf) for the time-varying reservoir characteristics $f(\mathbf{x}_0,\ldots,\mathbf{x}_N)$ can therefore be fully specified through the prior pdf at the initial timestep $f(\mathbf{x}_0)$, and $\boldsymbol{\omega}(\bullet)$. Here, and throughout the rest of this paper, $\mathbf{x}_k$ refers to the state of the reservoir at timestep t_k.

Observed production properties will be denoted by $\mathbf{q}_k^o = \mathbf{q}_k + \boldsymbol{\varepsilon}_{\mathbf{q}_k}$, where $\boldsymbol{\varepsilon}_{\mathbf{y}_k} \sim \mathbf{Gauss}_{n_q}(\mathbf{0}, \Sigma_{\varepsilon_{q_k}})$, independent of $\mathbf{q}_k$ $\forall k$. Here the notation $\mathbf{y} \sim \mathbf{Gauss}_{n_y}(\mathbf{0}, \Sigma_{\mathbf{y}})$ means that $\mathbf{y}$ is a random vector following the multivariate Gaussian probability distribution, having zero mean, and covariance matrix $\Sigma_{\mathbf{y}}$ (Anderson, 2003).

Similarly we will denote seismic time-lapse AVO signals at timestep t_k by $\mathbf{d}_k \in \mathbb{R}^{n_d\times 1}$, and the observed seismic data by $\mathbf{d}_k^o = \mathbf{d}_k + \boldsymbol{\varepsilon}_{\mathbf{d}_k}$, where $\boldsymbol{\varepsilon}_{\mathbf{d}_k} \sim \mathbf{Gauss}_{n_d}(\mathbf{0}, \Sigma_{\varepsilon_{dk}})$ independent of $\mathbf{d}_k$, $\forall k$. Note that by assuming a linear Gaussian model between seismic AVO data and the elastic properties $\mathbf{m}$ (Buland and Omre, 2003), and further using the non-linear Gassmann's equations to connect $\mathbf{m}$ to the reservoir state properties (Bachrach, 2006; Batzel and Wang, 1992), seismic AVO data will be connected to the reservoir state properties through a non-linear function $\boldsymbol{\psi}_{\mathbf{d}}(\bullet)$. That is, $\mathbf{d}_k = \boldsymbol{\psi}_{\mathbf{d}}(\mathbf{r}_k)$.

Note that for notational convenience we let $\mathbf{x}_k^o = [\mathbf{d}_k^{oT}, \mathbf{q}_k^{oT}]^T$, in spite that either seismic or production can be missing data at certain timesteps. Moreover, the above entails that in general we have a non-linear Gaussian

likelihood model $\mathbf{x}_k^o = \psi_{\mathbf{x}}(\mathbf{x}_k) + \varepsilon_{\mathbf{x}_k^o}$, where $\varepsilon_{\mathbf{x}_k^o} \sim \text{Gauss}_{n_{x^o}}(\mathbf{0}, \Sigma_{\varepsilon_{x_k^o}})$.

Ensemble Kalman Filter

Bayesian inversion is well suited to solve the complex, ill-posed and non-linear inverse problem of evaluating and predicting the reservoir state, and production properties based on observed data. Due to the Markov properties of the model, evaluation of the posterior pdf of interest $f(\mathbf{x}_k|\mathbf{x}_k^o,\ldots,\mathbf{x}_0^o)$, $k = 0,\ldots,N$, can be done sequentially (Lødøen and Omre, 2005). However, due to the non-linear fluid simulator and likelihood model connecting seismic AVO data to the reservoir state properties, such a model is analytically intractable in a statistical sense.

The EnKF (Evensen, 2007) provides a solution to this problem by assuming that $\mathbf{x}_k^u \sim f(\mathbf{x}_k|\mathbf{x}_{k-1}^o,\ldots,\mathbf{x}_0^o)$, approximately follows a Gaussian distribution with unknown mean $\mu_{\mathbf{x}_k}$ and covariance matrix $\Sigma_{\mathbf{x}_k}$. Due to the Markov properties of the model, realisations can therefore be generated sequentially as $\mathbf{x}_k^u = \omega(\mathbf{x}_{k-1}^c)$, where $\mathbf{x}_k^c \sim f(\mathbf{x}_k|\mathbf{x}_k^o,\ldots,\mathbf{x}_0^o)$, $k = 0,\ldots,N$. By using an ensemble of n_s realisations, $\mathbf{x}_k^{u(1)},\ldots,\mathbf{x}_k^{u(n_s)}$ we can then estimate the mean and covariance using the standard unbiased estimators

$$\begin{aligned}\hat{\mu}_{\mathbf{x}_k} &= \frac{1}{n_s}\sum_{i=1}^{n_s}\mathbf{x}_k^{(i)} \\ \hat{\Sigma}_{\mathbf{x}_k} &= \frac{1}{n_s-1}\sum_{i=1}^{n_s}(\mathbf{x}_k^{(i)} - \hat{\mu}_{\mathbf{x}_k})(\mathbf{x}_k^{(i)} - \hat{\mu}_{\mathbf{x}_k})^T. \end{aligned} \tag{1}$$

If all likelihood models connecting observed data to the reservoir state are linear Gaussian, that is, $\mathbf{x}_k^o \sim \text{Gauss}_{n_{x^o}}(\mathbf{D}_{\mathbf{x}_k}\mathbf{x}_k, \Sigma_{\varepsilon_{x_k}})$, for some matrix $\mathbf{D}_{\mathbf{x}_k} \in \mathbb{R}^{n_{x^o}\times n_x}$, then the posterior model will also be Gaussian, and thus analytically obtainable (Anderson, 2003). However, as the likelihood for the seismic AVO data is non-linear another approximation has to be made (Evensen, 2007).

Similarly as above, assume that $\tilde{\mathbf{x}}_k^{o(i)} = \psi_{\mathbf{x}}(\mathbf{x}_k^{u(i)})$, $i = 1,\ldots,n_s$, are realisations from an approximate Gaussian pdf with unknown mean $\mu_{\mathbf{x}_k^o}$ and covariance $\Sigma_{\mathbf{x}_k^o}$. Further let $\Sigma_{\mathbf{x}\mathbf{x}_k^o}$ denote the covariance between $\mathbf{x}_k$ and $\mathbf{x}_k^o$. Using the standard unbiased estimators in Equation (1), the unknown covariance matrices can then be estimated based on the realisations $\tilde{\mathbf{x}}_k^{o(i)}$ and $\mathbf{x}_k^{u(i)}$.

Under the Gaussian assumptions of the EnKF, realisations from the posterior pdf can therefore be generate sequentially for $k = 1,\ldots,N$ as:

$$\begin{aligned}\mathbf{x}_k^u &= \omega(\mathbf{x}_{k-1}^c) \\ \mathbf{x}_k^c &= \mathbf{x}_k^u + \hat{\Sigma}_{\mathbf{x}\mathbf{x}_k^o}(\Sigma_{\varepsilon_{x_k^o}} + \hat{\Sigma}_{\mathbf{x}_k^o})^{-1}(\mathbf{x}_k^o + \varepsilon_{\mathbf{x}_k^o} - \tilde{\mathbf{x}}_k^o), \end{aligned} \tag{2}$$

where $\varepsilon_{\mathbf{x}_k^o} \sim \text{Gauss}_{n_{x_k^o}}(\mathbf{0}, \Sigma_{\varepsilon_{x_k^o}})$.

Approximate Fluid Flow Simulation

Fluid flow simulation on a very fine scale grid is a computationally demanding process, meaning that it could take weeks, or even months to advance one single realisation from the current timestep to the next. Approximations must therefore be made in order to use the fluid flow simulator on a multiple set of realisations as required by the EnKF algorithm. One solution is to reduce the dimension of the reservoir through the possibly non-linear process known as upscaling (Farmer, 2002). That is $\mathbf{x}_k^* = \nu(\mathbf{x}_k) + \varepsilon_{\mathbf{x}_k^*}$, where $\varepsilon_{\mathbf{x}_k^*}$ is a Gaussian variable reflecting the error introduced by moving the reservoir from a fine scale to a coarse scale (Deutsch, 2002). The dimensional reduction will make repeated simulation on a coarse scale possible through the use of a faster fluid flow simulator $\omega^*(\bullet)$.

Upscaling is however, known to introduce bias in the fluid flow simulation (Omre and Lødøen, 2004), which should be accounted for. Moreover, as seismic data often appears on a much finer grid than is computationally feasible for fluid flow simulation, it is important also get good predictions of the reservoir state on a fine scale.

The empirical statistical solution to these challenges is the Scale-Corrected Ensemble Kalman Filter (SCEnKF) (Lødøen and Omre, 2005). The main idea of this method to use two ensembles, where one is referred to as the calibration ensemble, and the other referred to as the simulation ensemble. The calibration ensemble contains n_c realisations obtained using both the fine and coarse fluid flow simulator, while the simulation ensemble contains n_s realisations obtained only using $\omega^*(\bullet)$. Due to computational demands it is thus assumed that $n_c << n_s$. Also note that updating the two ensembles is done using the EnKF algorithm.

The EnKF assumption entails that outputs from the non-linear fluid flow simulator, on both fine, and coarse follow approximate Gaussian distributions with unknown mean and covariance. As in the EnKF updating scheme, we can therefore find analytical expressions for the conditional distribution of the fine scale reservoir state properties $\mathbf{x}_k$ given the coarse $\mathbf{x}_k^*$ namely:

$$\mathbf{x}_k^u = \mu_{\mathbf{x}_k} + \hat{\Sigma}_{\mathbf{x}\mathbf{x}_k^*}\hat{\Sigma}_{\mathbf{x}_k^*}^{-1}(\mathbf{x}_k^{u*} - \mu_{\mathbf{x}_k^*}) + \tilde{\varepsilon}_{\mathbf{x}_k}, \tag{3}$$

where $\tilde{\varepsilon}_{\mathbf{x}_k} \sim \text{Gauss}_{n_x}(\mathbf{0}, \hat{\Sigma}_{\mathbf{x}|\mathbf{x}_k^*})$, with

$$\hat{\Sigma}_{\mathbf{x}|\mathbf{x}_k^*} = \hat{\Sigma}_{\mathbf{x}_k} - \hat{\Sigma}_{\mathbf{x}\mathbf{x}_k^*}\hat{\Sigma}_{\mathbf{x}_k^*}\hat{\Sigma}_{\mathbf{x}\mathbf{x}_k^*}^T. \tag{4}$$

Here all unknown mean vectors and covariance matrices are estimated based on the calibration ensemble (Lødøen and Omre, 2005). Note that in the equations above, we have assumed that upscaling is a deterministic process, meaning that $\mathbf{x}_k^* = \nu^*(\mathbf{x}_k)$. If the covariance structure of the error term in the upscaling, $\varepsilon_{\mathbf{x}_k^*}$ is known, this matrix should be added to the covariance matrix $\Sigma_{\mathbf{x}_k^*}$ as in Equation (2).

SOLUTION TO RANK ISSUES

In reservoir applications the dimension of the reservoir state properties are typically much larger than the number of ensemble members, both in the calibration, and simulation ensembles. This implies that most of the estimated covariance matrices will be rank deficient, which will have indirect consequences for the success of the EnKF, and both direct and indirect consequences for the success of the SCEnKF. The direct effect can easily be seen by looking at Equations (3) and (4). Here we can see that both expressions involve inverting the covariance matrix $\hat{\Sigma}_{\mathbf{x}_k^*}$, which will be singular as long as $n_c - 1 < n_{x^*}$.

The indirect consequences can for instance be seen by looking at the EnKF update scheme in Equation (2). Since

$$\text{rank}(\hat{\Sigma}_{\mathbf{x}\mathbf{x}_k^o}) = \min\{(n_s - 1), n_x, n_{x^o}\},$$

$\hat{\Sigma}_{\mathbf{x}\mathbf{x}_k^o}$ will in most practical applications also be rank deficient. This may influence the conditional realisation $\mathbf{x}_k^c$ to such a degree that artifacts are created (Skjervheim, Evensen, Aanonsen, Ruud and Johansen, 2005). An empirical statistical solution to these issues is in statistical literature known as Principal Component Analysis (PCA) (Anderson, 2003).

Consider realisations $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n_s)} \in \mathbb{R}^{n \times 1}$ drawn independently from a probability density function having mean $\boldsymbol{\mu}$ and covariance $\Sigma_{\mathbf{x}}$. The aim of PCA is to explain the covariance structure of the data through a small number of orthogonal linear combinations termed Principal Components (PC) $\mathbf{z}^{(j)} = [\mathbf{z}_1^{(j)}, \ldots, \mathbf{z}_r^{(j)}]^T$, where $\mathbf{z}_i^{(j)} = \mathbf{b}_i^T \mathbf{x}^{(j)}$, $i = 1, \ldots, r$, $j = 1, \ldots, n_s$, $r \leq n$, such that the variance of $\mathbf{z}_i^{(\bullet)}$ is maximised under the condition that $\mathbf{b}_i^T \mathbf{b}_i = 1 \; \forall i$. It can be shown (Anderson, 2003) that the ith sample PC is given as $\mathbf{z}_{i,j} = \mathbf{v}_i^T \mathbf{x}_j$, where $\mathbf{v}_i$ is the ith eigenvector of the $\hat{\Sigma}_{\mathbf{x}}$ defined in Equation (1). Moreover, the estimated proportion of total variance explained by the kth PC is given as

$$w_k = \frac{\hat{\lambda}_k}{\sum_{i=1}^{r} \hat{\lambda}_i}, \tag{5}$$

where $\hat{\lambda}_k$ is the kth eigenvalue of $\hat{\Sigma}_{\mathbf{x}}$ and $r = \text{rank}(\hat{\Sigma}_{\mathbf{x}})$.

Note that if we assume that $\mathbf{x}^{(i)} \sim \text{Gauss}_n(\mu_{\mathbf{x}}, \Sigma_{\mathbf{x}})$, $i = 1, \ldots, n_s$, then by standard linear Gaussian theory $\mathbf{z}^{(i)} \sim \text{Gauss}_r(\mathbf{V}\mu, \Lambda)$. Here $\mathbf{V} \in \mathbb{R}^{r \times n}$ contains the r eigenvectors of $\hat{\Sigma}_{\mathbf{x}}$, and $\Lambda \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing the r eigenvalues. The rank issue can therefore be avoided by instead conditioning $\mathbf{z}^{(\bullet)}$, which corresponds to the optimal linear predictor in the case when the empirical covariance matrices are rank deficient. Working with PC will also reduce the computational demands, since we only have to work with matrices of low dimension. Note further that in the implementation of both the EnKF and SCEnKF algorithms, the use of Singular Value Decomposition avoids computation, and storage of potential very high dimensional covariance matrices (Evensen, 2007).

CASE STUDY

The synthetic case study is identical to the one used in (Hegstad and Omre, 2001), with the same reference model as used in (Lødøen and Omre, 2005). The reference reservoir, which is inspired by the Troll field in the North Sea offshore Norway, has a homogeneous three layered structure with high permeable middle layer, as seen in Figure 1. The reservoir is initially fully saturated with oil, and contains two horizontal production wells, and one vertical gas injection well. The size of the reservoir is $10^4 \times 10^4 \times 10^2$ feet^3, and it has been discretised into a $n = 50 \times 50 \times 15$ grid. The reference model was run for 2 100 days, using Eclipse 100 (2007.1) (GeoQuest, 2007) as the fluid flow simulator $\omega(\bullet)$. Reference pressure, gas saturation and production $\mathbf{q}_t^o$ was collected after 150, 300, 450, 600, 750, 900, 1 050, 1 200, 1 350, 1 650 and 2 100 days. Synthetic seismic data $\mathbf{d}_t^o$ was also generated after 300, 900 and 2 100 days using the non-linear seismic forward function $\psi_{\mathbf{d}}(\bullet)$. The seismic model parameters considered were the same as the ones used in (Sætrom, 2007). Upscaling of the reservoir was done such that permeability was mapped by geometric averaging, while porosity, saturation and pressure was mapped by arithmetic averaging, reducing the dimension of the reservoir to $n^* = 10 \times 10 \times 15$. Thus $\mathbf{r}_k^* = \mathbf{A}\mathbf{r}_k$, where $\mathbf{A} \in \mathbb{R}^{n_{r^*} \times n_r}$ is a block diagonal matrix where each entry is zero, or $1/25$.

RESULTS

Two different case studies have been carried out, which we will refer to as Case 1, and Case 2. Here we considered a calibration ensemble of size $n_c = 20$, and a simulation ensemble of size $n_s = 200$.

Case 1

Realisations from $f(\mathbf{x}_0|\mathbf{x}_0^o)$ were generated by adding Gaussian white noise with high variance to the reference reservoir state properties. Figure 1 shows the development of κ and ϕ for one of the downscaled simulation ensemble members, after 0, 300, 900 and 2 100 days of updating using the SCEnKF conditioning on both seismic and production data. As we can see from this figure, the ensemble member is drawn towards the reference case as more data becomes available. The same effect is present in the other ensemble members.

Finally we rerun the fine scale fluid flow simulator from the initial timestep, using the updated static reservoir properties after 0, 300, 900 and 2 100 days of updating as input, without conditioning on any observations, and predict for 3 000 days. Figure 2 shows the *opr* and *gor* for one of the two production wells. As we can see from this figure, the uncertainty in the initial predictions is reduced as more data becomes available. Further we see that there is no bias due to upscaling in the predictions.

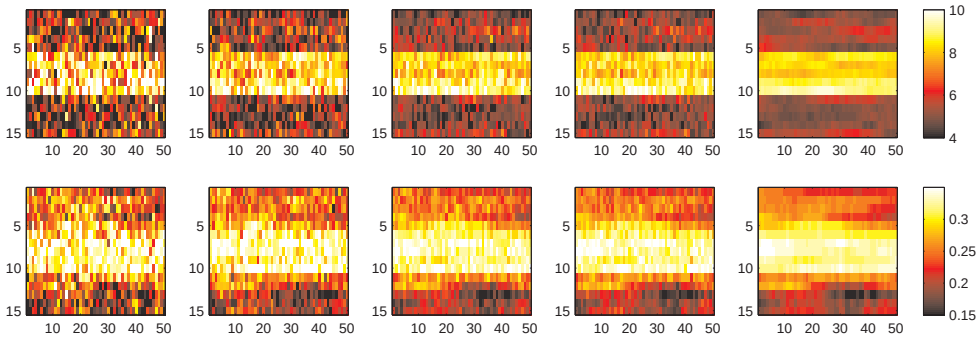

Figure 1: SCEnKF, Simulation Ensemble, Case 1: Vertical slice of log-permeability and porosity for one ensemble member. The first row shows from left to right; log-permeability after 0, 300, 900 and 2100 days of updating. The second row shows the porosity at the selected days. The rightmost plots shows, from top to bottom, reference log-permeability and porosity.

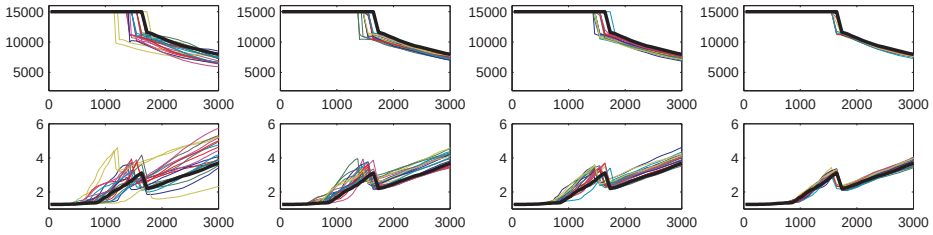

Figure 2: SCEnKF, Simulation Ensemble, Case 1: *opr* and *gor* for one production well based on reruns of the 20 first updated fine scale ensemble members compared to the fine scale reference production (thick line). The Figure shows from left to right; the initial ensemble, the ensemble after 300, 900 and 2100 days of updating. The first row shows the *opr*, and the second shows the *gor*.

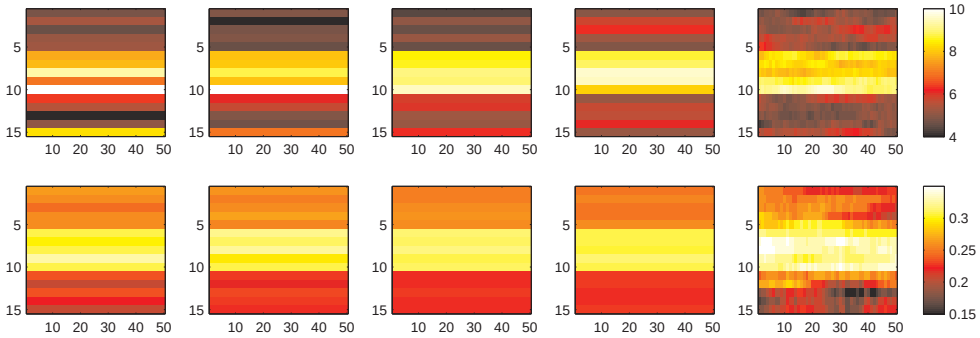

Figure 3: SCEnKF, Simulation Ensemble, Case 2: Vertical slice of log-permeability and porosity for one ensemble member. The first row shows from left to right; log-permeability after 0, 300, 900 and 2100 days of updating. The second row shows the porosity at the selected days. The rightmost plots shows, from top to bottom, reference log-permeability and porosity.

Case 2

The purpose of this case was to investigate how sensitive the SCEnKF is to the selection of a biased initial ensemble. Recall that in Case 1, the mean of the initial ensemble was identical to the reference model. The initial ensemble was generated by first computing the overall mean within the three separating layers of the reference reservoir, and adding layers constant value, generated as independent Gaussian variables with high variance. Figure 3 again shows the development in time for $\boldsymbol{\kappa}$ and $\boldsymbol{\phi}$ for one realisation. Here we can see that the SCEnKF, is only able to correct the level of each layer, and the realisations are drawn towards the mean value of the ensemble. This can be explained by the seismic data $\mathbf{d}_t$, which only carries information concerning contrasts, and not the actual level of the variables. Again the uncertainty in the production forecasts is reduced, as seen in Figure 4. Note however, that the bias is not entirely corrected for, and this is caused by the incorrect centring of the prior. It should also be noted that the production forecasts from the SCEnKF, is less biased than the forecasts obtained running the EnKF only on a coarse scale with the same initial ensemble, as shown in Figure 5.

CONCLUSIONS

Sequential data assimilation techniques for reservoir evaluation, has in recent years been given much attention, due to the increasing amount of time dependent data available. The EnKF is a Bayesian approach to this, which has shown promising results. When considering high dimensional reservoir models, approximations often have to be made, due to the high computational demands in the fluid flow simulation. These approximations

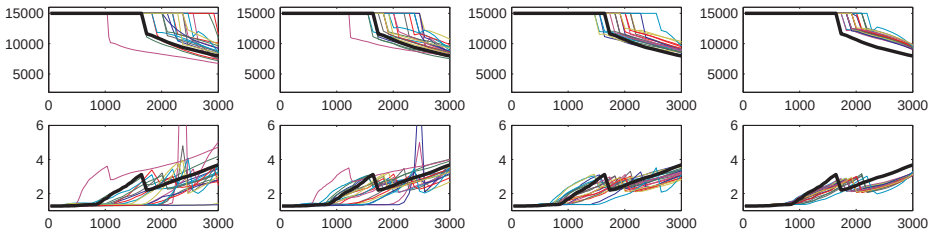

Figure 4: SCEnKF, Simulation Ensemble, Case 2: *opr* and *gor* for one production well based on reruns of the 20 first updated fine scale ensemble members compared to the fine scale reference production (thick line). The Figure shows from left to right; the initial ensemble, the ensemble after 300, 900 and 2 100 days of updating. The first row shows the *opr*, and the second shows the *gor*.

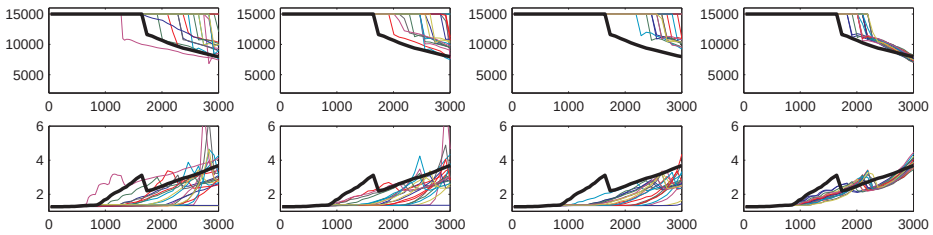

Figure 5: EnKF, Coarse Scale, Case 2: *opr* and *gor* for one production well based on reruns of the 20 first updated coarse scale ensemble members compared to the fine scale reference production (thick line). The Figure shows from left to right; the initial ensemble, the ensemble after 300, 900 and 2 100 days of updating. The first row shows the *opr*, and the second shows the *gor*.

are known to introduce bias, and the SCEnKF was therefore proposed as an extension to the EnKF in order to correct for this.

Testing the SCEnKF on two different case studies including both time-lapse seismic and production data revealed that the success of this method depends highly on the centring of the initial ensemble. This is due to lack of information about the level of the variables in seismic data. It is therefore of utmost importance that the prior model is selected such that it reflects the true reservoir. This may be a complicated task and should be performed in collaboration with both geologists, and using available well observations. It appears however, that the scale-correction in the SCEnKF partly corrects for bias in production forecasts, and that lack of precise observations concerning the level of reservoir variables is the major problem.

The problem of low rank in the estimated covariance matrices, both in the EnKF and the SCEnKF, can be handled using Principal Component Analysis. However, such an approach tend to underestimate the uncertainty in the ensemble, and we could experience an ensemble collapsing towards the centre of the ensemble. Moreover, the uncertainty in the empirical estimates is not possible to assess.

REFERENCES

Anderson, TW (2003). *An introduction to multivariate statistical analysis.* Wiley, 3 ed.

Bachrach, R (2006). *Joint estimation of porosity and saturation using stochastic rock-physics modelling.* In Geophysics, vol. 71, no. 5, pp. 53–63.

Batzel, M and Wang, Z (1992). *Seismic properties of pore fluids.* In Geophysics, vol. 57, no. 11.

Buland, A and Omre, H (2003). *Bayesian linearized AVO inversion.* In Geophysics, vol. 68, no. 1, pp. 185–198.

Deutsch, CV (2002). *Geostatistical Reservoir Modeling.* Oxford University Press.

Evensen, G (2007). *Data Assimilation. The Ensemble Kalman Filter.* Springer.

Farmer, CL (2002). *Upscaling: a review.* In International Journal for Numerical Methods in Fluids.

GeoQuest (2007). *ECLIPSE Reference Manual 2007.1.* Schlumberger GeoQuest.

Hegstad, BK and Omre, H (2001). *Uncertainty in production forecasts based on well observations, seismic data and production history.* In Society of Petroleum Engineers Journal, pp. 409–425.

Lødøen, OP and Omre, H (2005). *Scale-corrected ensemble kalman filtering applied to production history conditioning in reservoir evaluation.* Paper submitted for publication.

Omre, H and Lødøen, OP (2004). *Improved production forecasts and history matching using approximate fluid-flow simulators.* In Society of Petroleum Engineers Journal, vol. 9, pp. 339–351.

Skjervheim, JA, Evensen, G, Aanonsen, S, Ruud, B and Johansen, T (2005). *Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter.* In SPE.

Sætrom, J (2007). *Hierarchical Ensemble Kalman Filter for Observations of Production and 4-D Seismic Data.* Master's thesis, Norwegian University of Science and Technology.