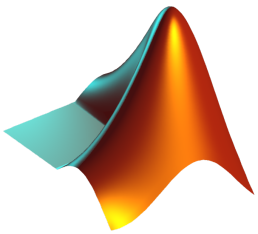


Kalman Filters
Matlab Scripts

Technical Note

–

Application of the Ensemble Kalman Filter to History Matching

---

Author:
Frederik Lehn

July 30, 2021

Introduction

The purpose of this note is to introduce the Ensemble Kalman Filter (EnKF) and its applications in history matching of petroleum reservoirs. Kalman filtering is a data assimilation technique which allows for estimating unobserved variables of a system, such as porosities and permeabilities. Data assimilation is the practice of combining a theoretical knowledge of a system with measurements, which is coupled together through a so called State Space Model. For history matching the theoretical knowledge is the reservoir model and the measurements are fluid rates.

In order to make the process of history matching by application of the EnKF more easily understandable, the theoretical background of the EnKF will be broken down into the following parts

1. Introduction of State Space Models
2. Introduction of Kalman Filtering Techniques
3. The EnKF as applied to history matching
4. Discussion on how to use the history matched ensemble in forecasting and development optimization

The introductions will begin with the basics of the models and filters and gradually increase in complexity, showing examples along the way.

State Space Models

A state space model is a time-dependent model, in which the next state of the time-series can be modelled solely by knowledge of the previous state and possible external forces acting upon the system. A good example is a reservoir simulation, which is an iterative process where the state of the system in the current time-step is solely dependent on the state of the system in the previous time-step.

Linear State Space Model

The linear state space model has the form

$$\begin{aligned} \text{System equation:} \quad & \mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t \\ \text{Observation equation:} \quad & \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t \end{aligned} \tag{1}$$

where

- $\mathbf{x}_t$ is the state vector at time t
- $\mathbf{F}_t$ is the state-transition matrix at time t
- $\mathbf{B}_t$ is the control-input matrix at time t
- $\mathbf{u}_t$ is the control vector at time t
- $\mathbf{z}_t$ is the measurements
- $\mathbf{H}_t$ is the observation matrix at time t

The system equation is the theoretical understanding of the system. It relates the state of the system at time t (for instance the pressure and saturations of a reservoir simulation) to the previous state at $t-1$, plus

some potential external input (typically not relevant in reservoir simulations). The observation equation is the measurement part of the data assimilation process. It relates the state of the system to the measured quantities (i.e. how fluid rates are calculated from pressures and saturations).

Linear Stochastic State Space Model

The linear stochastic state space model is very similar to the deterministic model

$$\begin{aligned} \text{System equation:} \quad & \mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \mathbf{w}_t \\ \text{Observation equation:} \quad & \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t \end{aligned} \tag{2}$$

where the added terms are noise terms

- $\mathbf{w}_t \sim \mathcal{N}(0, \mathbf{Q}_t)$ is the process noise following a multivariate normal distribution with mean zero and covariance $\mathbf{Q}_t$, which may vary in time
- $\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{R}_t)$ is the observation noise following a multivariate normal distribution with mean zero and covariance $\mathbf{R}_t$, which may vary in time

This model takes into account that there will be some noise in the physical system equation, $\mathbf{w}_t$, related to uncertainty in the model and some noise in the actual measurements of the system, $\mathbf{v}_t$.

Non-Linear Stochastic State Space Model

The non-linear stochastic state space model has the form

$$\begin{aligned} \text{System equation:} \quad & \mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{u}_t) + \mathbf{w}_t \\ \text{Observation equation:} \quad & \mathbf{z}_t = h(\mathbf{x}_t) + \mathbf{v}_t \end{aligned} \tag{3}$$

where

- $f(\mathbf{x}_{t-1})$ is some arbitrary function mapping of $\mathbf{x}_{t-1}$ and $\mathbf{u}_t$ to $\mathbf{x}_t$
- $h(\mathbf{x}_t)$ is some arbitrary function mapping of $\mathbf{x}_t$ to $\mathbf{z}_t$

the difference from previous is that $f(\mathbf{x}_{t-1}, \mathbf{u}_t)$ and $h(\mathbf{x}_t)$ can be non-linear functions. Given that the typical reservoir models (black-oil and compositional) are non-linear functions, the state space model used in history matching is the non-linear stochastic state space model.

Kalman Filter Techniques

The Kalman Filter (KF) works recursively, meaning that the current estimate is solely dependent on the estimate of the previous state and the current observation. It consists of a two step process, namely a prediction step based on a theoretical knowledge of the system (state space model) followed by an update step (adjusting the prediction to actual measurements).

The Kalman Filter is a minimum mean-square error estimator. This means that it is derived on the assumption that the estimate has to minimize the following objective function

$$\text{obj} = \min \quad \mathrm{E}\left[||\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}||_2^2 \right] = \min \quad \text{Trace}\left(\mathbf{P}_{t|t} \right) \tag{4}$$

where $\mathrm{E}[\cdot]$ is the mean operator, $||\cdot||_2$ is the Euclidean norm and $\mathrm{Trace}(\cdot)$ is the sum of the diagonal elements of a matrix. $\mathbf{P}_{t|t}$ is the covariance of the error between the true state vector and the estimated state-vector.

The Kalman Filter exploits a key property of uncorrelated Gaussian distributions, namely that the product of two Gaussian distributions is also Gaussian, and with a lower variance than either of the two original distributions, which is proven and very well demonstrated in [4]. Assume that $\mathbf{x}_{t|t-1}$ and $\mathbf{z}_t$ are one-dimensional (but the proof is easily extendable to multiple dimensions), and that their uncertainty is normally distributed, i.e.

$$x_{t|t-1}(x;\mu_x,\sigma_x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \quad \wedge \quad z_t(x;\mu_z,\sigma_z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} e^{-\frac{(x-\mu_z)^2}{2\sigma_z^2}} \tag{5}$$

then the product of the two distributions becomes

$$x_{t|t}(x;\mu_x,\sigma_x,\mu_z,\sigma_z) = \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_z^2}} e^{-\left(\frac{(x-\mu_x)^2}{2\sigma_x^2}+\frac{(x-\mu_z)^2}{2\sigma_z^2}\right)} \tag{6}$$

which is a third Gaussian distribution given by

$$x_{t|t}(x;\mu_{xz},\sigma_{xz}) = \frac{1}{\sqrt{2\pi\sigma_{xz}^2}} e^{-\frac{(x-\mu_{xz})^2}{2\sigma_{xz}^2}} \tag{7}$$

where

$$\mu_{xz} = \frac{\mu_x\sigma_z^2+\mu_z\sigma_x^2}{\sigma_x^2+\sigma_z^2} \quad \Leftrightarrow \quad \mu_{xz} = \mu_x + \frac{\sigma_x^2(\mu_z-\mu_x)}{\sigma_x^2+\sigma_z^2} \quad \vee \quad \mu_{xz} = \mu_z + \frac{\sigma_z^2(\mu_x-\mu_z)}{\sigma_x^2+\sigma_z^2} \tag{8}$$

and similarly for the variance

$$\sigma_{xz}^2 = \frac{\sigma_x^2\sigma_z^2}{\sigma_x^2+\sigma_z^2} \quad \Leftrightarrow \quad \sigma_{xz}^2 = \sigma_x^2 - \frac{\sigma_x^4}{\sigma_x^2+\sigma_z^2} \quad \vee \quad \sigma_{xz}^2 = \sigma_z^2 - \frac{\sigma_z^4}{\sigma_x^2+\sigma_z^2} \tag{9}$$

especially the result for the combined variance is important. Given that $\sigma_x > 0$ and $\sigma_z > 0$ it means that the second term in the resulting variance is positive and thus $\sigma_{xz}^2 < \sigma_x^2$ and $\sigma_{xz}^2 < \sigma_z^2$ meaning that the resulting uncertainty is lower than both the prediction uncertainty and measurement uncertainty respectively. In [4] this proof is furthered, to show how the different variables in the Kalman Filter steps uses this result.

Kalman Filter

The original Kalman Filter is developed for estimating variables of the Linear Stochastic State Space Model assuming both model and measurement uncertainty is Gaussian distributed.

Prediction:

1. Predicted (*a priori*) state estimate
$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t\hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t\mathbf{u}_t \tag{10}$$
2. Predicted (*a priori*) error covariance
$$\mathbf{P}_{t|t-1} = \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^T + \mathbf{Q}_t \tag{11}$$

Update:

3. Calculate measurement residual
$$\tilde{\mathbf{y}}_t = \mathbf{z}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1} \tag{12}$$
4. Calculate measurement covariance
$$\mathbf{S}_t = \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R}_t \tag{13}$$
5. Calculate optimal Kalman gain
$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^T\mathbf{S}_t^{-1} \tag{14}$$
6. Update (*a posteriori*) state estimate
$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\tilde{\mathbf{y}}_t \tag{15}$$
7. Update (*a posteriori*) estimate covariance
$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \tag{16}$$

Breaking this down step for step:

1. The state of the system is progressed in time based on the theoretical knowledge (e.g. a time-step in a reservoir simulation)

2. The covariance of the error of the system at the current time-step is progressed in time (i.e. the uncertainty of the predicted state of the system)

3. The residual between the actual measurements and the forecasted quantities (i.e. fluid rates at time t) is calculated

4. The covariance of the error of the measurements at the current time-step is calculated

5. The optimal Kalman Gain is calculated. The formula for the optimal Kalman Gain is derived by differentiating (4) w.r.t. to the gain and solving the derivative equal to 0 (i.e. finding the optimum).

6. The predicted state of the system is corrected to account for the information in the measurement

7. The uncertainty of the state is also updated to account for the optimal gain

The process of assimilating the data recursively by adding the residual multiplied by an optimal step-direction (the Kalman Gain) can be compared to the inner Newton-loop of a time-step in a reservoir simulator, where the pressures and saturations of the next time-step is iteratively updated using the residual times the Jacobian.

The use of the state- and measurement covariances matrices, $\mathbf{P}_{t|t-1}$ and $\mathbf{S}_t$, in the calculation of the Kalman Gain means that the update of the estimate is based on a weighted average where uncertain parameters and measurements are given less influence than more certain ones.

The Kalman Filter was first used by NASA for better estimating their satellites positions. The following example will show what such an estimate looks like.

Assume that the satellite's distance from earth (radius of orbit) and the angle in radians, (r_t^m, θ_t^m) is measured using a telescope at a set of discrete times, t. Due to atmospheric disturbances, imperfections in the telescope, etc. the measured positions will be different from the true positions. The measurements can thus be related to the true positions by adding some noise in the system

$$\begin{aligned} r_t^m &= r_t^p + \epsilon_{rt} \\ \theta_t^m &= \theta_t^p + \epsilon_{\theta t} \end{aligned} \tag{17}$$

where the noise terms, ϵ_{rt} and $\epsilon_{\theta t}$ are independent and identically distributed (i.i.d.) and follows $\epsilon_{rt} \sim \mathcal{N}(0, 2000^2)$ and $\epsilon_{\theta t} \sim \mathcal{N}(0, 0.03^2)$.

In practice the trajectory of an orbiting satellite follows an ellipsis, however it will be assumed that it can be well approximated by a circle. Using this assumption the trajectory can be modelled in polar coordinates by the state vector $\mathbf{x}_t = \begin{bmatrix} r_t^p & \theta_t^p & v_{\theta,t}^p \end{bmatrix}^T$ where r_t^p, θ_t^p and $v_{\theta,t}^p$ are the true distance, angle and angular velocity at time t respectively. The evolution of the system from $t-1$ to t is described by the set of equations

$$\begin{aligned} r_t^p &= r_{t-1}^p + \epsilon_{rt}^p \\ \theta_t^p &= \theta_{t-1}^p + v_{\theta,t-1}^p + \epsilon_{\theta t}^p \\ v_{\theta,t}^p &= v_{\theta,t-1}^p + \epsilon_{v_\theta t}^p \end{aligned} \tag{18}$$

where the noise terms account for effects such as the orbit not being a perfect circle. The errors are again i.i.d and follows

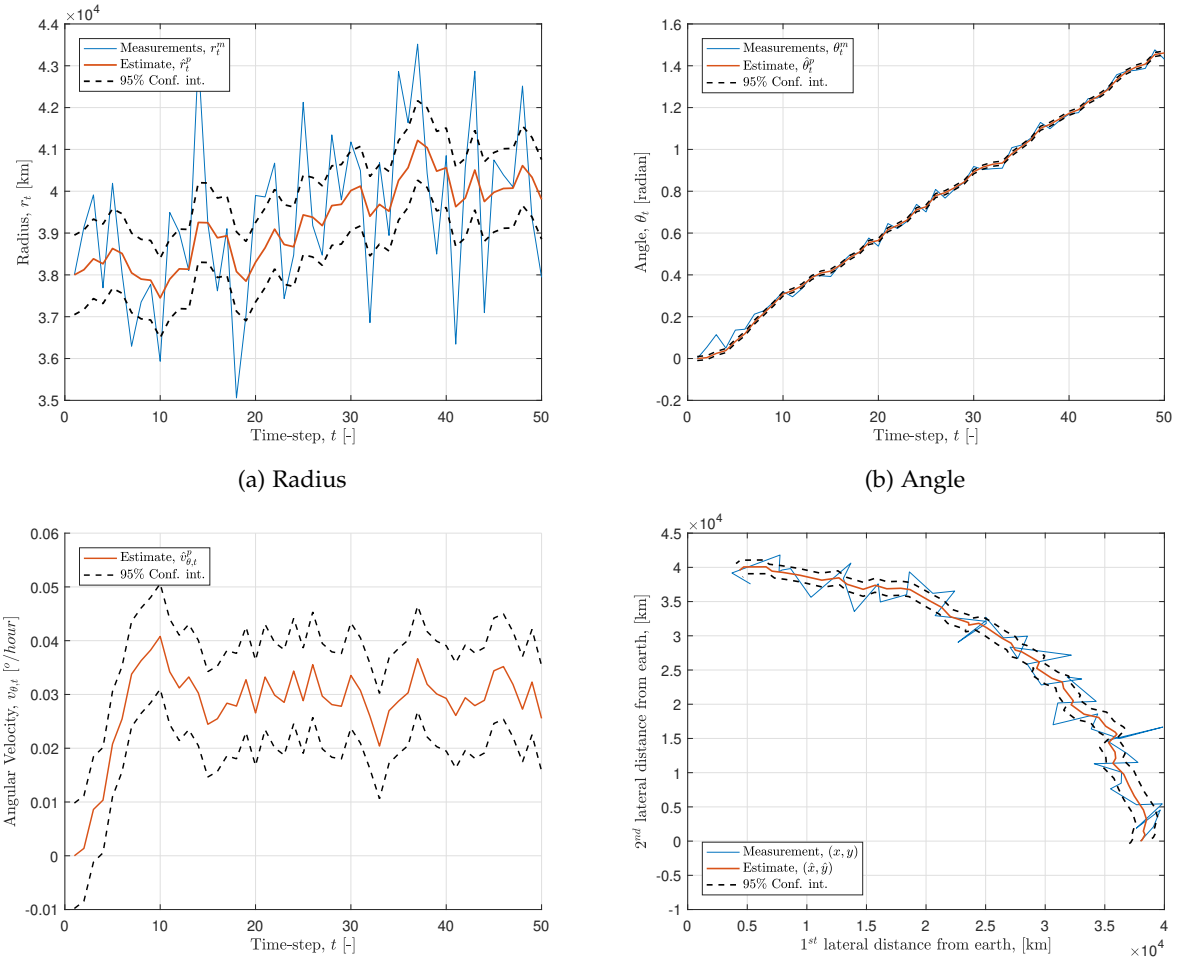
$$\begin{aligned} \epsilon_{rt}^{p} &\sim \mathcal{N}(0, 500^2) \\ \epsilon_{\theta t}^{p} &\sim \mathcal{N}(0, 0.005^2) \\ \epsilon_{v\theta t} &\sim \mathcal{N}(0, 0.005^2) \end{aligned} \tag{19}$$

The first thing to do is writing the problem as a state space model.

$$\begin{aligned} \text{System equation:} \quad \mathbf{x}_t &= \begin{bmatrix} r_t^p \\ \theta_t^p \\ v_{\theta,t}^p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{t-1}^p \\ \theta_{t-1}^p \\ v_{\theta,t-1}^p \end{bmatrix} + \begin{bmatrix} \epsilon_{rt}^p \\ \epsilon_{\theta t}^p \\ \epsilon_{v_\theta t}^p \end{bmatrix} \\ \text{Observation equation:} \quad \mathbf{z}_t &= \begin{bmatrix} r_t^m \\ \theta_t^m \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_t^p \\ \theta_t^p \\ v_{\theta,t}^p \end{bmatrix} + \begin{bmatrix} \epsilon_{rt} \\ \epsilon_{\theta t} \end{bmatrix} \end{aligned} \tag{20}$$

Notice here that r_t^p, θ_t^p are observed variables, whereas $v_{\theta,t}^p$ is an unobserved variable. Also there is no external input to the model, i.e. $\mathbf{B} = \mathbf{0}$ and $\mathbf{u} = \mathbf{0}$.

Given that the chosen state space model is linear, the KF may be used to estimate the true position of the satellite from the measured position.



(a) Radius

(b) Angle

(c) Angular Velocity

(d) Cartesian Position

Figure 1: Measurements of a satellite's geoposition and estimates of the true position of the satellite by use of the Kalman Filter.

As can be seen from Fig. 1 the estimated state variables exhibits substantially less fluctuations than the measurements, and further, the 95 % confidence intervals are quite narrow compared to the measurements, showing a high degree of certainty in the estimates, perfectly in line with the previously shown proof of variance reduction.

Extended Kalman Filter

The Extended Kalman Filter (EKF) was developed for estimating variables of the Non-Linear Stochastic State Space Model assuming both model and measurement uncertainty is Gaussian distributed.

Prediction:

Predicted (*a priori*) state estimate	$\hat{\mathbf{x}}_{t\|t-1} = f(\hat{\mathbf{x}}_{t-1\|t-1}, \mathbf{u}_t)$	(21)
Predicted (*a priori*) error covariance	$\mathbf{P}_{t\|t-1} = \mathbf{F}_t \mathbf{P}_{t-1\|t-1} \mathbf{F}_t^T + \mathbf{Q}_t$	(22)

Update:

Calculate measurement residual	$\tilde{\mathbf{y}}_t = \mathbf{z}_t - h(\hat{\mathbf{x}}_{t\|t-1})$	(23)
Calculate measurement covariance	$\mathbf{S}_t = \mathbf{H}_t \mathbf{P}_{t\|t-1} \mathbf{H}_t^T + \mathbf{R}_t$	(24)
Calculate optimal Kalman gain	$\mathbf{K}_t = \mathbf{P}_{t\|t-1} \mathbf{H}_t^T \mathbf{S}_t^{-1}$	(25)
Update (*a posteriori*) state estimate	$\hat{\mathbf{x}}_{t\|t} = \hat{\mathbf{x}}_{t\|t-1} + \mathbf{K}_t \tilde{\mathbf{y}}_t$	(26)
Update (*a posteriori*) estimate covariance	$\mathbf{P}_{t\|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t\|t-1}$	(27)

where the state transition- and observation matrices are defined by the Jacobians of f and h respectively

$$\begin{aligned} \mathbf{F}_t &= \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{t-1\|t-1}, \mathbf{u}_t} \\ \mathbf{H}_t &= \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{t-1\|t-1}} \end{aligned} \tag{28}$$

The EKF is a linearization of the non-linear state space model based on a 1^{st}-order Taylor expansion (i.e. the Jacobians) and consequently suffers from convergence issues for highly non-linear models. Notice that applying an EKF on a linear state space model yields the same result as using a normal KF, as the 1^{st}-order Taylor expansion of a linear model is equal to the model itself.

Ensemble Kalman Filter

The Ensemble Kalman Filter was invented to handle problems with a large number of variables such as discretizations of partial differential equations in geophysical models (e.g. reservoir models). Even though the EKF can handle the potential non-linearity of the state space model, it struggles to deal with scaling issues (memory and computational efficiency) when the number of parameters increases. The EnKF solves this problem by not progressing the covariance matrice in time, but rather constructing it at each time-step by calculating the sample covariance of an ensemble of equally probable realisations. Rather than having just one state variable, $\mathbf{x}_t$, an ensemble of state variables is generated

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{x}_t^{(1)} & \dots & \mathbf{x}_t^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_t^{(i)} \end{bmatrix} \tag{29}$$

where $\mathbf{X}_t$ is an $n \times N$ matrix, n being the number of state variables and N the number of realisations in the ensemble. Similarly the measurements are represented as an ensemble

$$\mathbf{Z}_t = \begin{bmatrix} \mathbf{z}_t^{(1)} & \dots & \mathbf{z}_t^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{z}_t^{(i)} \end{bmatrix}, \quad \mathbf{z}_t^{(i)} = \mathbf{z}_t + \mathbf{v}_t^{(i)}, \quad \mathbf{v}_t^{(i)} \sim \mathcal{N}(0, \mathbf{R}_t) \tag{30}$$

There exists different implementations of the ensemble kalman filter. The two most common of these are the

- The Pertubed Observation Ensemble Kalman Filter (POEnKF)
- The Ensemble Square Root Kalman Filter

Pertubed Observation Ensemble Kalman Filter

The POEnKF works by including Monte Carlo sampled (pertubed) observations in the estimation process. The algorithm works as follows [1]:

Generate an ensemble of size N state vectors

$$\hat{\mathbf{x}}_{0|0}^{(i)} = \hat{\mathbf{x}}_{0|0} + \mathbf{d}^{(i)}, \quad i = 1,...,N \tag{31}$$

where $\mathbf{d}^{(i)} \sim \mathcal{N}(0, \mathbf{P}_0)$ and $\mathbf{P}_0$ is the initial guess for the state covariance.

Prediction:

Forward each state vector realisation in time

$$\hat{\mathbf{x}}_{t|t-1}^{(i)} = f\left(\hat{\mathbf{x}}_{t-1|t-1}^{(i)}, \mathbf{u}_t\right) + \mathbf{w}_t^{(i)}, \quad i = 1,...,N \tag{32}$$

Predicted (*a priori*) state estimate

$$\hat{\mathbf{x}}_{t|t-1} = \frac{1}{N}\sum_{i=1}^{N} \hat{\mathbf{x}}_{t|t-1}^{(i)} \tag{33}$$

Update:

Calculate prediction of each realisation

$$\hat{\mathbf{y}}_t^{(i)} = h\left(\hat{\mathbf{x}}_{t-1|t}^{(i)}\right), \quad i = 1,...,N$$

Calculate ensemble mean of prediction

$$\hat{\mathbf{y}}_t = \frac{1}{N}\sum_{i=1}^{N} \hat{\mathbf{y}}_t^{(i)}$$

Calculate measurement covariance

$$\mathbf{P}_{y_t} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\mathbf{y}}_t^{(i)} - \hat{\mathbf{y}}_t\right)\left(\hat{\mathbf{y}}_t^{(i)} - \hat{\mathbf{y}}_t\right)^T + \mathbf{R}_{e_t}$$

Calculate cross covariance

$$\mathbf{P}_{xy_t} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\mathbf{x}}_{t|t-1}^{(i)} - \hat{\mathbf{x}}_{t|t-1}\right)\left(\hat{\mathbf{y}}_t^{(i)} - \hat{\mathbf{y}}_t\right)^T$$

Calculate Kalman Gain

$$\mathbf{K}_t = \mathbf{P}_{xy_t}\mathbf{P}_{y_t}^{-1}$$

Update each state vector realisation

$$\hat{\mathbf{x}}_{t|t}^{(i)} = \hat{\mathbf{x}}_{t|t-1}^{(i)} + \mathbf{K}_t\left(\mathbf{z}_t + \mathbf{v}_t^{(i)} - \hat{\mathbf{y}}_t^{(i)}\right), \quad i = 1,...,N$$

Update (*a posteriori*) state estimate

$$\hat{\mathbf{x}}_{t|t} = \frac{1}{N}\sum_{i=1}^{N} \hat{\mathbf{x}}_{t|t}^{(i)}$$

where $\mathbf{w}_t^{(i)}$ is randomly sampled from the multivariate normal distribution $\mathbf{w}_t^{(i)} \sim \mathcal{N}(0, \mathbf{Q}_t)$ and $\mathbf{R}_{e_k}$ is the sample measurement noise covariance

$$\mathbf{R}_{e_t} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{v}_t^{(i)}\right)\left(\mathbf{v}_t^{(i)}\right)^T \tag{34}$$

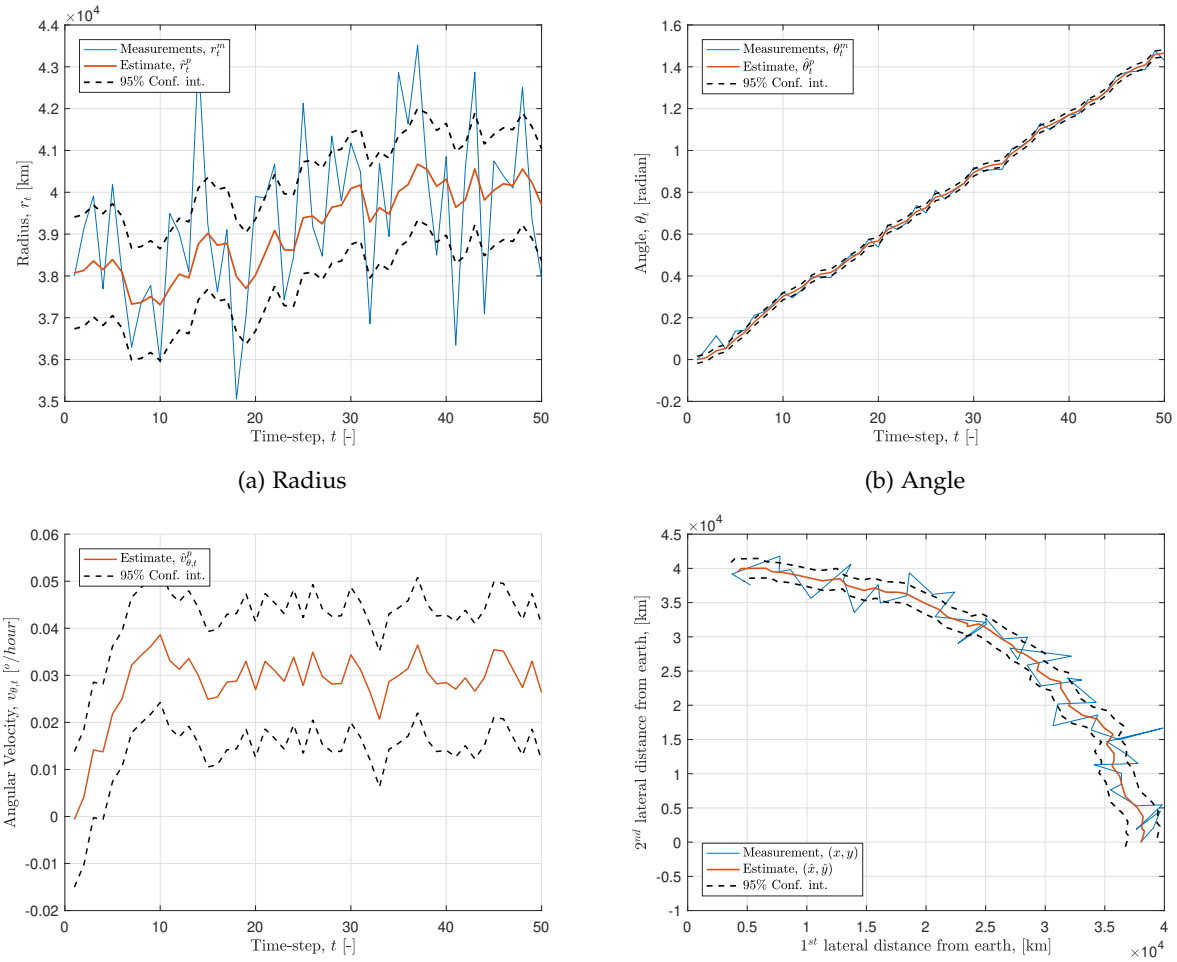
here, $\mathbf{v}_t^{(i)}$ is is randomly sampled from the multivariate normal distribution $\mathbf{v}_t^{(i)} \sim \mathcal{N}(0, \mathbf{R}_t)$. Additionally

the error covariance can be calculated

$$\mathbf{P}_{t|t} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\hat{\mathbf{x}}_{t|t-1}^{(i)} - \hat{\mathbf{x}}_{t|t-1} \right) \left(\hat{\mathbf{x}}_{t|t-1}^{(i)} - \hat{\mathbf{x}}_{t|t-1} \right)^T \tag{35}$$

if confidence intervals are required.

The POEnKF is tested on the same geopositioning problem as used in the section "Kalman Filter"



(a) Radius

(b) Angle

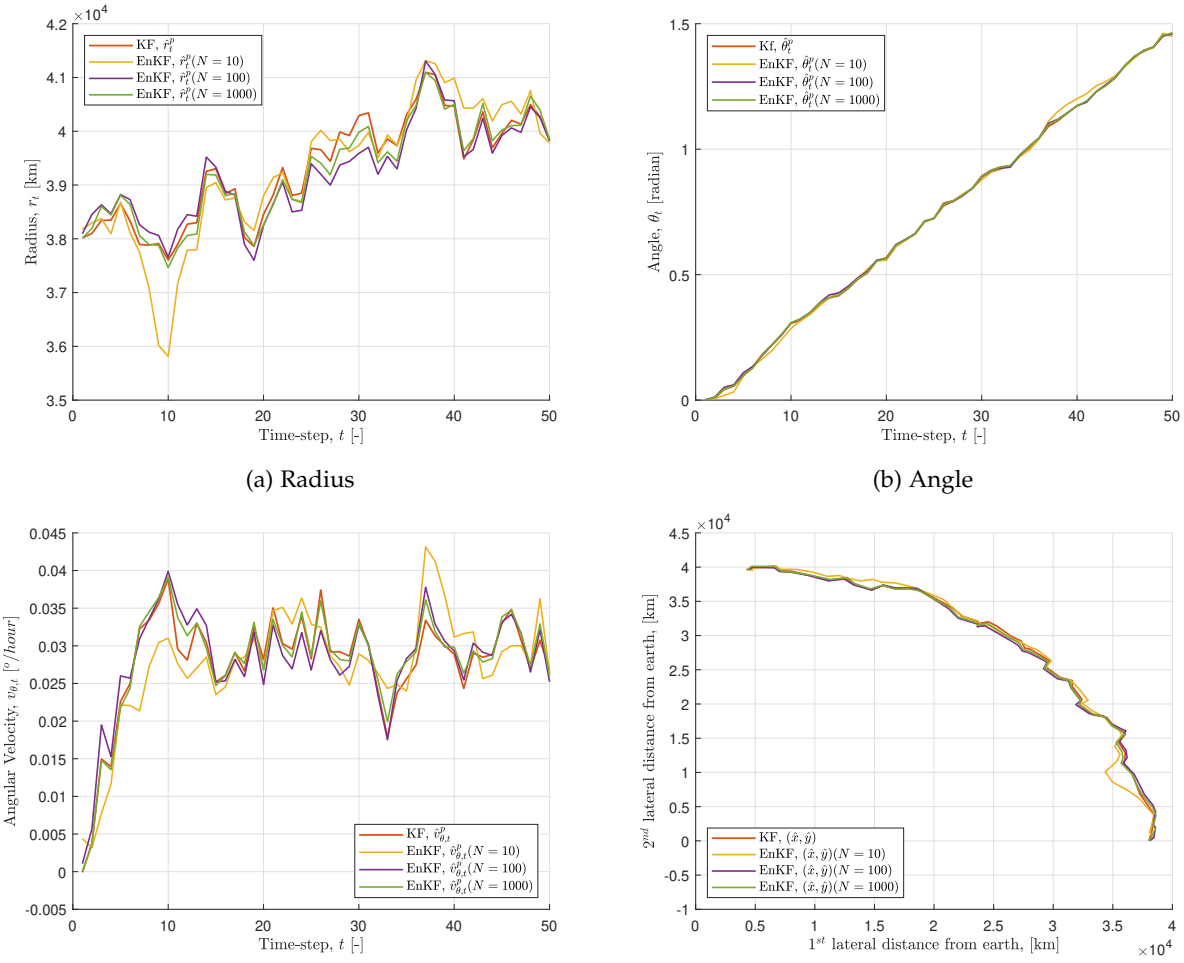
(c) Angular Velocity

(d) Cartesian Position

Figure 2: Measurements of a satellite's geoposition and estimates of the true position of the satellite by use of the POEnKF with an ensemble size of $N = 100$.

As can be seen the result is practically the same as for the normal KF, as expected.

The EnKF is logically dependent on the size of the ensemble, namely $\hat{\mathbf{x}}_{t|t} \rightarrow \mathbf{x}_{t|t}$ as $N \rightarrow \infty$, so an investigation of the effect of the ensemble size is shown



(a) Radius

(b) Angle

(c) Angular Velocity

(d) Cartesian Position

Figure 3: Measurements of a satellite's geoposition and estimates of the true position of the satellite by use of the KF and the POEnKF with an ensemble size of $N = [10, 100, 1000]$.

When the ensemble is of sufficient size it should show the same result as the normal KF (the red line). However, when the ensemble is small it will suffer from sample issues, due to the inherent randomness of sampling pertubed observations as part of the algorithm. This is exactly what is seen on Fig. 3 where the small ensemble fluctuates substantially more from the red line, than the larger ensembles.

Ensemble Square Root Kalman Filter

The Ensemble Square Root Kalman Filter is a method that avoids the use of Monte Carlo sampling (the pertubations) and thus does not suffer from sampling errors at small ensemble sizes. This however introduces other issues, and this filter will not be further discussed here.

Applications in History Matching

History matching reservoir models through a data assimilation process is very suitable as reservoir simulators provide a tool for solving the system equation and actual production measurements from the wells provide data to assimilate.
Consider a dynamic model of non-linear partial differential equations discretized in space, with a set of known boundary conditions. Let $\hat{\mathbf{u}}_t \approx \mathbf{u}(t) \in \mathbb{R}^{n_u}$ denote the discretized approximation to the solution at time t. It is further assumed that the solution is dependent on a set of highly uncertain parameter, $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$, which are the unobservable variables to be estimated

$$\frac{\partial \mathbf{u}(t)}{\partial t} = f(\mathbf{u}; \boldsymbol{\theta}) + \mathbf{w}_t \tag{36}$$

and some initial conditions

$$\mathbf{u}(t_0) = \mathbf{u}_0 \tag{37}$$

which may or may not be uncertain. The error term $\mathbf{w}_t$ denotes the model error. [3] suggests the model uncertainty should be modelled using spatial correlations on the distance from the wells such that static parameters in near-wellbore cells will be more affected than boundary cells.

$\hat{\mathbf{u}}$ typically includes all primary variables from the simulation (e.g. the pressure and saturation for a Black-Oil Model).

It is further assumed that the model is time constrained by a set of discrete measurements at times t_1,...,t_{n_t} and that the state variables can be related to these measurements through the non-linear relation

$$\mathbf{z}_t = h(\mathbf{u}(t),\boldsymbol{\theta}) + \mathbf{v}_t \in \mathbb{R}^{n_z} \tag{38}$$

where $\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{R}_t)$ is the measurements error with covariance matrix $\mathbf{R}_t$ at time t.

In practice, when applying the EnKF to non-linear problems such as history matching, it is common to define an augmented state variable, containing both the parameters to be estimated, the state variables and the simulated results

$$\mathbf{x}_t = \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{u}_t \\ \mathbf{z}_t \end{bmatrix} \in \mathbb{R}^{n_\theta + n_u + n_z} \tag{39}$$

Notice that, by applying the augmented state vector the relationship between the measurements and the predicted state variables becomes linear

$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{v}_t \tag{40}$$

where the observation matrix, $\mathbf{H}$, is now simply a matrix of ones and zeros.

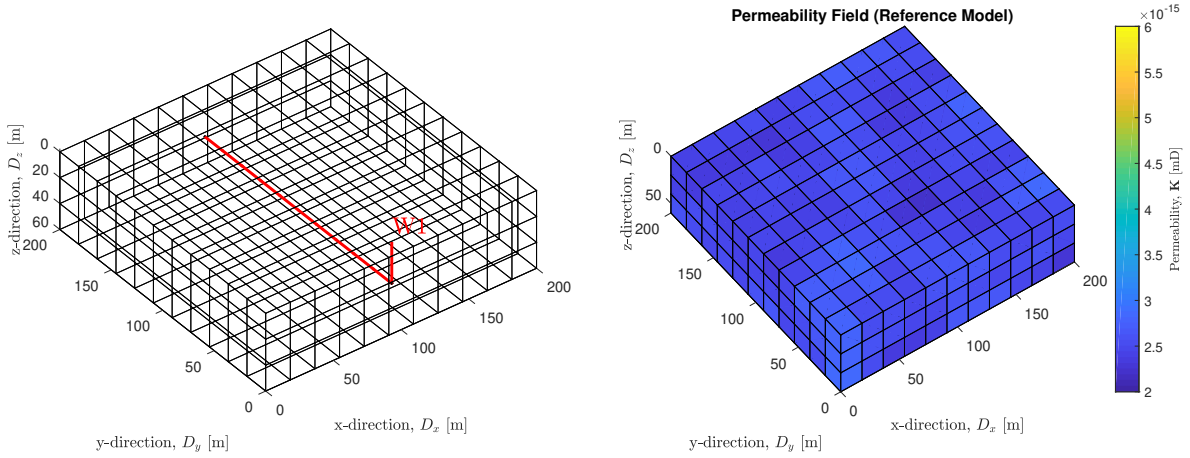
Logically the EnKF should be run recursively in time, where the state variable at time t is updated from the previous state $(t-1)$ by only running the reservoir simulator from the previous state (i.e. using restart files). However, [2] notes in their review, that the best results are obtained when the reservoir simulation is rerun from $t = 0$ with the updated parameters, $\boldsymbol{\theta}$, from $\hat{\mathbf{x}}_{t-1}$. In any case the parameters, $\boldsymbol{\theta}$, obtained after the last data assimilation step, are the history matched parameters.

In order to verify how all this works in practice an ensemble of 50 simple reservoir models has been constructed. All models consists of an $11 \times 11 \times 3$ box model of dimensions $200 \times 200 \times 60$ meters with uniform porosity and net-to-gross ratio of $\phi = 30$ % and NTG $= 1$ respectively. Uncertainty is added to the ensembles by randomly sampling the permeability fields of the models using a Gaussian distribution with

a target value range of $\mathbf{K} \in [0.1, 5]$ mD and a standard deviation of 0.65 mD. The model is assumed to be completely saturated with oil and developed with a horizontal production well in the center of the reservoir with a BHP of 10^5 Pa and a wellbore radius of 0.1 m.

The governing equation of the simulation is a compressible single-phase flow equation solved using a fully implicit scheme.

There is assumed to be no model uncertainty (which is an over simplification) and the uncertainty in the measurements are $R = 10 \left[\frac{\text{m}^3}{\text{day}}\right]$

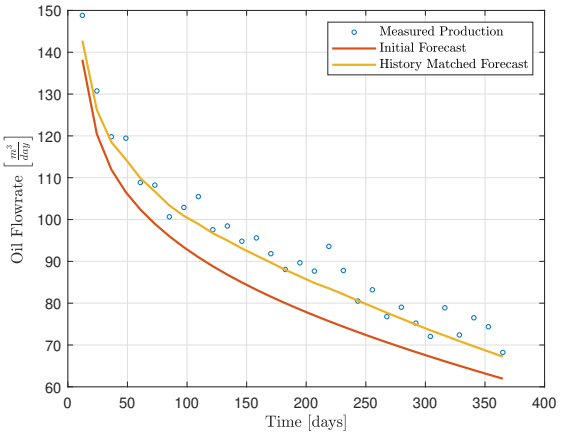


(a) Well Trajectory

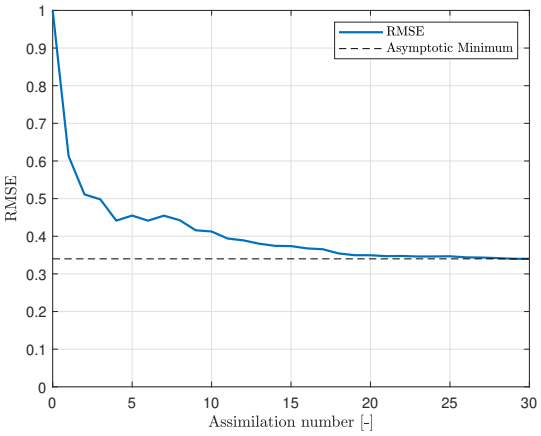
(b) Permeability Field

Figure 4: Well trajectory of the horizontal oil producer and the reference permeability field (given by the mean of the initial ensemble permeability fields).

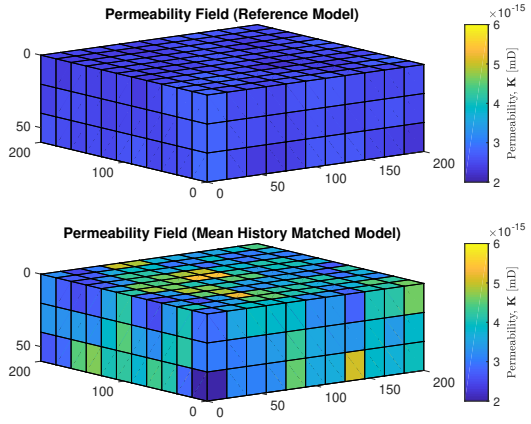
The parameters, $\boldsymbol{\theta}$, to be optimized by the history match is the permeability field and the measurements are oil flow rates. The resulting output from the history match using 30 measurements of oil flow rates looks as follows



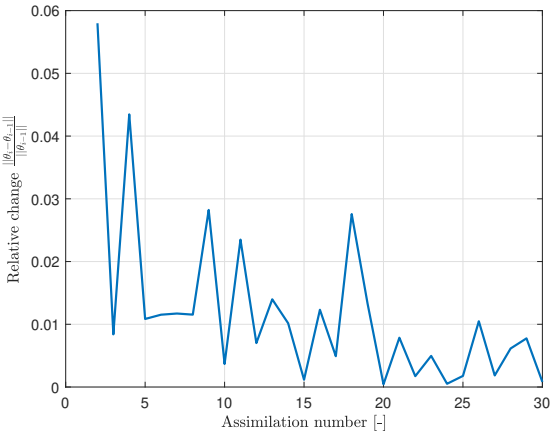
(a) Production Forecast



(b) RMSE



(c) Permeability Fields



(d) Relative Change in Permeability

Figure 5: Output from the history match using an EnKF with an ensemble size of $N = 50$ and $n_t = 30$ measurements. The RMSE is normalized by the initial RMSE at t_0.

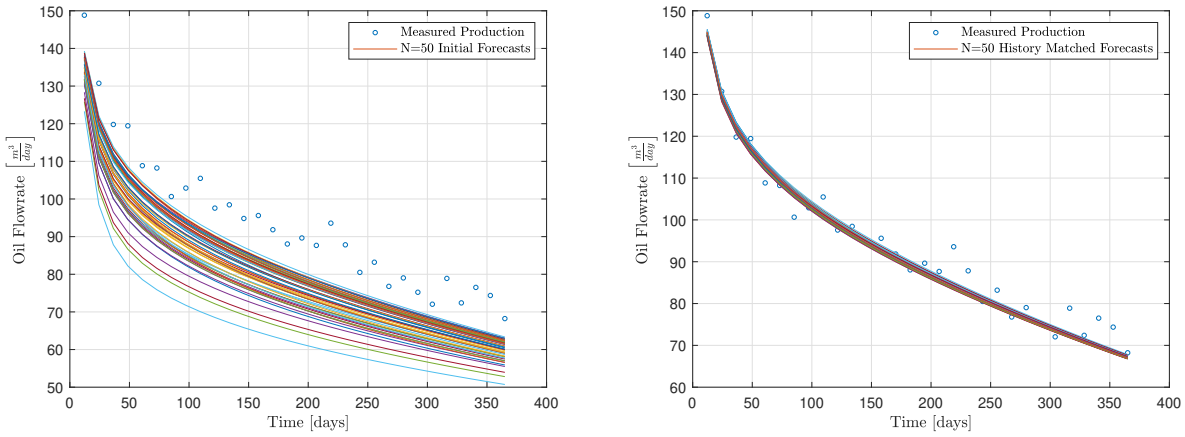
The effectiveness of the history match is evaluated by the root mean-squared error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}\left(\mathbf{z}_{mes,t} - \mathbf{z}_{mean,t}\right)^2} \tag{41}$$

where $\mathbf{z}_{mes,t}$ and $\mathbf{z}_{mean,t}$ is the measurement at time t and the mean of the ensemble prediction respectively. The RMSE as a function of assimilation steps is seen on Fig. 5b, which shows the expected asymptotic convergence towards a minimum error. Once the minimum error has been reached, it will require a more accurate model if the error is to be lowered further.

Fig. 5d shows that, while the change per assimilation step is sporadic, the optimized parameters do become more and more stable as more data is assimilated.

It is not only the mean model that is history matched, but each individual model in the ensemble



(a) Initial Ensemble Forecasts

(b) History Matched Ensemble Forecasts

Figure 6: Production profiles for each of the realisations in the ensemble, before and after history matching.

The value of this is that the entire ensemble is history matched simultaneously and thus ready for future use in other applications such as uncertainty modelling and optimization under uncertainty. As can be seen from Fig. 6 the spread of the history matched ensemble is almost negligible. This is called ensemble collapse and is a common issue, to which the most well known solution is introducing modelling uncertainty by means of spatial correlation. It should also be noted that the example used here is very simple, and the collapse will most likely not be as extreme for a more complex model.

As with any parameter tuning exercise, overparameterization is a concern, and as such the provided case is an example of bad practise as the EnKF is used to fit 363 parameters to 30 measurements. Given the enormous convergence power of an algorithm like the EnKF, it is almost guaranteed to manage to find a solution which matches the observed production. As can be seen from Fig. 5c the permeabilities have changed substantially, both in terms of absolute values (wider spread), but also in terms of which regions of the reservoir that are high or low permeable. Typically the initial ensemble will be provided by a team of geologists, geophysicists and petrophysicists and based on a suite of available data from seismic measurements, logs, etc. If the number of parameters that the EnKF is allowed to optimize on is large and unconstrained there is a chance the resulting model will differ significantly from the initial model, thus providing a good fit to the production measurements, but ignoring the ranges estimated by the remaining subsurface team.

This leads to the conclusion that the EnKF should not be used as a black-box tool that is given free reign over which parameters to optimize on, but rather as a tool for assisted history matching. The key in using an EnKF will be to choose the parameters it is allowed to optimize on, how to constrain those parameters and how to manage the model uncertainty so near-wellbore and boundary cells are not given equal weight.

Discussion

History matching by means of an EnKF is not in itself particularly useful, unless there is a well defined use for the resulting ensemble. The logical answer for this is using the ensemble for uncertainty modelling either in production forecasting or optimization under uncertainty. The problem with that, is the computational requirements are proportional to the number of realisations in the ensemble and that the ensemble usually has to be rather large to appropriately represent the posterior distribution, as seen on Fig. 3.

Two potential options for overcoming the high computational requirements are

1. Use a small ensemble for the history matching and subsequent forecasting/optimization, but make sure the realisations are intelligently sampled using for instance an experimental design method
2. Use a large ensemble for the history matching and then identify a sub-ensemble based on a some degree of non-similarity. This identification can be done using tracer/ToF analysis. Subsequent forecast/optimization can then be done using the sub-ensemble

Both solutions suffer from the same problem, namely that the small resulting ensemble does not appropriately sample the posterior distribution which leads to too high perceived uncertainty in the forecasts.

[5] mentions a comprehensive list of methods which can be used to overcome the issue with a small sample size. Essentially they all employ some sort of interpolation technique, such as kriging, regression, artificial neural network, etc. which are used to sample additional points without needing to perform a reservoir simulation. This is comparable to how the the Kriging Genetic Algorithm (KRGA) used internally in 6X works for optimization under uncertainty. Essentially N reservoir simulations are run and then a large number of additional points are sampled using kriging. All of these approaches all come with different pros and cons, but common for all of them are that they introduce additional modelling uncertainty.

Unfortunately we do not currently have sufficient understanding to make a conclusion.

References

Used References

[1] W. V. Bossche (2013). *Data assimilation toolbox for Matlab*. Thesis from KU Leuven.

[2] S. I. Aanonsen, et al. (2009). *The Ensemble Kalman Filter in Reservoir Engineering - a Review*. Society of Petroleum Engineers, SPE Journal.

[3] G. Nævdal, et al. (2005). *Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter*. Society of Petroleum Engineers (SPE 84372).

[4] R. Faragher (2012) *Understanding the Basis of the Kalman Filter Via a Simple and Intuitive Derivation*. IEEE Signal Processing Magazine.

[5] R. Patel, et al. (2018). *Performance Assessment of Ensemble Klamn Filter and Markov Chain Monte Carlo under Forecast Model Uncertainty*. ECMOR XVI 2018.

Relevant References

[6] L. Heidari, et. al. (2011). *History Matching of Reservoir Models by Ensemble Kalman Filtering: The State of the Art and a Sensitivity Study*. The American Association of Petroleum Geologists.

Further Reading

[7] A. Fowler (2012). *Data Assimilation Tutorial on the Kalman filter*. University of Reading.

[8] J. Lei, P. Bickel (2009). *Ensemble Filtering for High Dimensional Non-linear State Space Models*. UC Berkeley

[9] G. Evensen (1994). *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*. American Geophysical Union (94JC00572).

[10] G. Evensen (2003). *The Ensemble Kalman Filter: theoretical formulation and practical implementation*. Springer-Verlag.

[11] G. Evensen (2004). *Sampling strategies and square root analysis schemes for the EnKF*. Springer-Verlag.