Extended Abstract

History Matching by Ensemble Kalman Filter

Gonçalo José de Carvalho Rodrigues
Instituto Superior Técnico, Portugal

Abstract

Bayesian statistics provides an adequate framework to incorporate field observations in reservoir simulation models in a way that allows one to describe uncertainty in the reservoir parameters and the prediction about fluid flow simulations. The Ensemble Kalman Filter (EnKF) is a data assimilation method, which can be derived from Bayesian statistics. It represents an attractive method for reservoir history matching because it is easy to implement and computationally efficient. The EnKF is a sequential Monte Carlo method that provides an alternative to the traditional Kalman Filter, and is used extensively in many different fields including ocean and atmospheric sciences, oil reservoir simulations, and hydrological modeling. Due to limit the size of the initial set and of linear Gaussian assumption in the analysis stage some problems remain: excessive reduction in the variability of the ensemble, limited capacity to absorb a large quantity of independent data, loss of geological realism in case of complex geology. Many of these problems become prominent when the number of data is large, which implies that more techniques are required to obtain satisfactory results using EnKF. In this context we decided to incorporate in the initial ensemble of EnKF, nonparametric permeability models and analyse the results obtained. For that purpose we used direct sequential simulation to generate the permeability models. The results obtained in the characterization of petro physical properties were not satisfactory, although, in terms of dynamic response of the results were promising.

KEYWORDS: Data Assimilation; History Matching; Ensemble Kalman Filter; Direct Sequential Simulation;

Introduction

The ensemble Kalman filter is a powerful method for sequentially updating estimates of model variables and for assimilating various types of data. One of the problems with the traditional Kalman filter is the difficulty of the computation of the covariance of the model parameters, which is necessary for ensuring that each adjustment to current model parameters does not destroy the match to previous observation. The updated covariance
matrix needs to be constructed and stored each time a new set of data are assimilated. A second problem with the traditional Kalman filter is that it is necessary to compute the sensitivity of data to model variables, as in many history matching algorithms (Bissell et al., 1994; Oliver, 1994; Chu et al., 1995; Omre et al., 1996; He et al., 1997; Gosselin et al., 2001; Li et al., 2003). This computation makes the traditional Kalman filter impractical for even moderate-sized reservoir problems. Eisenmann et al. (1994) and Corser et al. (2000) have attempted to apply the traditional Kalman filter to reservoir characterization problems, but the applications have been restricted to problems with small number of parameters and the relationships between observation and model parameters were nearly linear. The ensemble Kalman filter method was first introduced by Evensen (1994). One way to explain the algorithm is to describe the steps in its application. It begins with the generation of an ensemble of initial models (typically 40–100) consistent with prior knowledge of the initial state and its probability distribution (uncertainty). For flow problems, each of the reservoir models is advanced to the time of the next observation using a reservoir simulator. The covariance of model variables, which is needed for the Kalman filter, could be estimated directly from the ensemble of states. It is never necessary, however, to compute the covariance explicitly as only a few columns are needed for updating of the model variables. Anderson and Anderson (1999) described the model updating step of EnKF within the Bayesian framework of matching data as well as honoring prior model probability distributions. When the number of ensemble states is large enough and the problem is nearly linear, the ensemble of vectors in the EnKF method is able to correctly assess uncertainty in the distribution of model parameters. Even when the initial ensemble members are from the correct distribution, the resulting conditional realizations can still be reasonable if the dimension of the subspace is large enough (Wen and Chen, 2005). The ensemble Kalman filter has been developed and successfully applied mainly in the fields of physical oceanography and meteorology. Anderson (2001) demonstrated the ability of a modified ensemble Kalman filter to assimilate data in a problem with a state vector much larger than the number of ensembles. Evensen (2003) recently provided a comprehensive review of the progress on the application of the ensemble Kalman filter since its introduction by Evensen (1994). This method is now beginning to be applied in other fields, including groundwater hydrology (Reichle et al., 2002) and petroleum engineering. N&vdal et al. (2003) applied ensemble Kalman filter techniques for continuous model updating on two 2-dimensional 3-phase reservoir problems. One was a synthetic model with two producers and one injector, the other model was a simplified model of a North Sea oilfield. The measurements in both cases included well pressure, oil rates, GORs, and water cut. The reservoir models were updated by assimilating production data at least once a month and also when a well began production or was shut in. They found that the predication of future performance from the ensembles improved with more data assimilated; the permeability estimate, however, became worse. Gu and Oliver (2005) applied the EnKF to the PUNQ-S3 problem (Floris et al., 2001). They found that the method was quite efficient compared to the
gradient based methods and provided a reasonable estimate of uncertainty.

Objective

The main impetus for the development of this master thesis is to evaluate the limitations of EnKF methodology, and give some contribution to this methodology may have success in the oil and gas industry weaving and continues to have in other areas of science. The EnKF was introduced in the oil industry recently and has attracted attention as a promising method for solving the historical setting problem. In recent years, especially in the last decade, EnKF methodology has been the subject of numerous studies to try to mitigate some limitations to the methodology presented in nonlinear cases, cases related to non-parametric distributions. These same limitations have removed this promising methodology of the oil and gas industry.

This thesis presents an application of EnKF methodology, a method for data assimilation, which enables the integration simultaneously of: well data, seismic data production and its starting point a set (or ensemble) achievements permeability. An advantage of this method is the low computational cost when compared to similar methods and little time spent.

Next is presented the objectives of their work:

- Familiarization with the simulation software MRST fluids;
- Familiarization with the toolbox EnKF-MRST;
- Implementation of an adjustment methodology of production history using the Ensemble Kalman filter;
- Deviation Optimization of the historical setting and forecasts for a synthetic reservoir;
- Analysis of the sensitivity of input parameters of a traditional methodology EnKF;

Methodology

The basic idea of the ensemble Kalman filter method is that it is possible to propagate a group of state models along time using the full nonlinear model dynamics while adjusting the paths to assimilate to data; the statistical information among the group of states is used for model updating. In the following introduction, we denote Y as a group of ensemble states:

\[ Y = \{ y_1, y_2, \ldots, N_e \} \]

Where \( N_e \) is the number of state vectors in the ensemble. The ensemble Kalman filter for assimilating data consists of two sequential steps. One is the forecast forward in time based on solution of the dynamical equations for flow and transport in the reservoir. The other is data assimilation to update the model by correcting the variables describing the state of the system to honor the observations. The state vector in the ensemble Kalman filter contains all the uncertain and dynamic variables that define the state of the system. At a certain time step \( i \), the state vector for the reservoir model is expressed as:

\[ y^i = [(m^i)^T, d(m^i)^T]^T \] (1)

Where \( m_i \) consists of variables for rock properties and flow system in every grid block, \( d(m) \) is the simulated data from the previous simulation run. The number of simulated data in the vector \( d(m) \) does not have to be the same at every assimilation step since it depends on the number of observation data at time step \( i \).

Initial state vectors, which are sampled from the prior probability density function of the
state vector before any data assimilation. The update to each ensemble member is made using the Kalman update formula:

\[ y_j^u = y_j^p + K_e (d_j - H y_j^p), \]

for \( j = 1, \ldots, N_e \)  

(2)

Where the superscript \( p \) denotes predicted in contrast to \( u \), which means updated, \( N_e \) is the number of ensemble members, \( K_e \) is the ensemble Kalman gain, and \( H \) is the measurement operator that extracts the simulated data from the state vector \( y \). If the state vector is constructed as in Eq. (1), then has 1's in locations corresponding to data and 0's elsewhere. \( d_j \) is the observation data at current time plus noise from the same distribution as the measurement error:

\[ d_j = d_{\text{obs}} + \epsilon_j, \quad \text{for } j = 1, \ldots, N_e \]  

(3)

The ensemble Kalman gain is computed as:

\[ K_e = C_{y,x} H^T (HC_{y,x} H^T + C_D)^{-1} \]  

(4)

Where the covariance matrix of the state vectors at any time can be obtained from the ensemble members by the standard estimator:

\[ C_{y,x}^p = \frac{1}{N_e - 1} \sum_{i,j=1}^{N_e} (y_i^p - \bar{y}^p) (y_j^p - \bar{y}^p)^T \]  

(5)

Where \( \bar{y} \) is the mean of the \( N_e \) ensemble members at the current data assimilation step. The subscript \( Y \) represents the ensemble of state vectors.

This thesis presents an application of Ensemble Kalman Filter method in a case of the production history matching. First introduced DSS technique in generating permeability models to appear in the initial ensemble of permeability realizations in EnKF methodology. With this introduction we plan to investigate the impact inherent in this change, knowing beforehand that are traditionally used Gaussian distribution of permeability models in the initial ensemble.
Case Study

The EnKF methodology for history matching petroleum reservoirs was tested and implemented a set of synthetic data. However, to support the application of this methodology, we chose Matlab Reservoir Simulation Toolbox (MRST). The MRST was developed by a group of Geoscientists Department of applied mathematics SINTEF ICT. This toolbox was developed in order to help students from various areas to develop their research projects without requiring subject to commercial software licenses. Along with MRST was used EnKF-MRST module that serves as the basis for practical part of the dissertation. This module was developed by Olwijn Leeuwenburgh (TNO), includes EnKF and EnRML schemes, location, inflation, provides production data (Bottom hole pressure, Water cut and Total liquid rate) and also seismic data (water saturation).

The EnKF-MRST module is programmed with a history matching methodology of production, production data information over twenty years and that allows you to forecast up to thirty years. The reservoir used for the application of this methodology is a 2D synthetic reservoir, the grid of the reservoir is defined by blocks 441, wherein [21x21] cells. The reservoir has four production wells and an injector configured in 5-spot system (Figure 2).
These two histograms show the distributions of each ensemble of permeability realizations:

To make a comprehensive analysis of these two ensembles of realizations, we run the reservoir simulator with different parameterization.

This application was tested in a historical setting methodology production prepared to historical setting up to 20 years (7300 days) and future forecasts up to 30 years (10950 days). The different parameterizations which was carried out consisted of variation of the initial ensemble size permeability of realizations (50, 100, 150, 200, 500 models). Then we show the results obtained with models the initial ensemble taking into account that there was a significant difference in the variation of the ensemble size.
Results

**Figura 3** - a) Real map of permeability b) Real map of porosity c) Map of pressure d) Map of water saturation in timestep 7300 days

**Figura 4** - a) Real map of permeability b) Real map of porosity c) Map of pressure d) Map of water saturation in timestep 10950 days

**Figura 5** - a) Model of permeability b) Model of porosity c) Model of pressure d) Model of water saturation [timestep 7300 dias] [RandomGauss]

**Figura 6** - a) Model of permeability b) Model of porosity c) Model of pressure d) Model of water saturation [timestep 10950 dias] [RandomGauss]

**Figura 7** - a) Model of permeability b) Model of porosity c) Model of pressure d) Model of water saturation [timestep 7300 dias] [DSS]

**Figura 8** - a) Model of permeability b) Model of porosity c) Model of pressure d) Model of water saturation [timestep 7300 dias] [DSS]
The introduction of non-parametric set of realizations given rise to porosity and permeability maps with values greater variability than with Gaussian distribution models. It is natural that we observe this result because we had already seen this situation in statistical distributions associated with the two ensembles. Regarding the pressure and water saturation values are relatively identical. This indicates that the EnKF has difficulty dealing with a non-parametric distribution in the characterization of petro-physical properties. As can be seen following images maps obtained through SSD permeability and porosity are far from the real maps.

**Bottom hole pressure**

![Figure 9 - BHP injector [timestep 10950 days] [Gaussian ensemble]](image)

![Figure 10 – BHP injector [timestep 10950 days] [DSS]](image)

Regarding the bottom hole pressure, the non-parametric ensemble realizations obtained a historical setting with less deviation, while the red curve represents the average of the values of all the closest models of blue curve represents the actual values. With this result it can be seen that the EnKF got a good result in the historic setting of the bottom hole pressure. As shown in Figure 10 the information provided (models) of BHP (gray curves) are closer to the actual values.
**Conclusions**

The introduction of non-parametric permeability realizations in an EnKF methodology showed no efficiency in characterizing the petrophysical properties of the reservoir, rather, they obtained more distant values of reality. However in relation to the dynamic response of the results were not very different from the results of the simulations with the ensemble of Gaussian models. Begins a precedent never before tested using the DSS to generate permeability realizations to appear in the initial ensemble of EnKF methodology.

The number of models to be included in the initial ensemble (ensemble size) did not show a marked difference results we attribute to the small scale of the reservoir where it was applied this methodology because the variation in the size of the initial ensemble is often crucial in getting good results with this methodology.

**References**


