Scale-Corrected Ensemble Kalman Filtering Applied to Production History Conditioning in Reservoir Evaluation

Ole Petter Lødøen and Henning Omre

Department of Mathematical Sciences,
Norwegian University of Science and Technology,
Trondheim, Norway

Abstract

In reservoir evaluation problems, the reservoir properties are largely unknown. To infer these properties from observations of the reservoir production is referred to as history matching or production history conditioning. Traditionally, this is done by repeated fluid flow simulations, where all the available production data are used simultaneously to arrive at a set of history matched reservoir models. In recent years there has been an increase in the amount of data continuously collected from a reservoir under production. Hence, the need for automatic, continuous model updating is apparent. The ensemble Kalman filter has been shown to be suitable for this purpose. However, large reservoir evaluation problems require upscaling of the reservoir properties to be able to perform the required number of fluid flow simulations. Traditional ensemble Kalman filtering is shown to give bias in the production history conditioned reservoir representations. The loss in accuracy and precision introduced by performing fluid flow simulations on a coarser scale should be accounted for, but this is rarely or never done. We introduce the scale-corrected ensemble Kalman filter approach in order to quantify loss in accuracy and precision. A reference scale is defined and all uncertainty quantifications are made relative to this scale, although the fluid flow simulations are made on a coarser scale. The production history conditioned reservoir representation will be accurate with realistic precision measures on this reference scale. The methodology is demonstrated on a large case study inspired by the characteristics of the Troll field in the North Sea.

Keywords: Approximate fluid flow simulation; Bayesian statistics; Ensemble Kalman filter; Parameter estimation; Production history matching

1 Introduction

One of the objectives of reservoir evaluation is to find the optimal well configuration and well operating conditions for a given reservoir. Prediction of the reservoir properties, such as porosity, permeability, location of faults and fractures, and hydrocarbon saturation is important when deciding where to drill new wells and when predicting the fluid flow patterns for a given well configuration. Forecasts of hydrocarbon production for a given recovery strategy can be used to determine optimal well operating conditions. The evaluation must be based on both general reservoir experience and reservoir specific observations, such as seismic data, well logs and observed production history. Quantification of the uncertainty both in the prediction of the reservoir properties and in the forecast of the production properties should be an integral part of the evaluation process.

The assessment of the uncertainty in the production forecasts requires repeated fluid flow simulations. This is done by using a reservoir production simulator. The reservoir production simulator solves a set of partial differential equations, taking the reservoir conditions, well configuration and well operating conditions as input, and provides time-series containing pressures and produced rates of hydrocarbons as output. The reservoir conditions needed as input are in practice largely unknown, however. Therefore, the uncertainty in the reservoir properties must be described by a stochastic reservoir model, taking all the available data into account. Prior to production start, the available data consist of static data, such as well logs and seismic data, combined with general geological knowledge obtained through experience with similar reservoirs. After the reservoir has been in production for a while, however, an observed production history is also available. The observed production history carries information about the reservoir properties, hence it should be used to update the reservoir model, and thereby improve the production forecasts. In petroleum related literature this is referred to as 'history matching'.

Traditionally, production history conditioning is performed through repeated fluid flow simulations, where the reservoir properties are tuned to the production history, either manually or automatically by minimising an objective function involving the mismatch between simulated and observed production. There are mainly two problems with this methodology. The first problem is the computational cost of repeated fluid flow simulations, which severely restricts the size of the reservoir models to which the production history conditioning can be applied. The second problem is that the reservoir models are updated using all the available production data simultaneously. This means that when new production data become available, the entire production history conditioning process must be repeated. In recent years, the use of permanent sensors for monitoring dynamic production properties has increased, requiring more frequent updating of the reservoir models.

Ideally, the observations should be included in the model sequentially as they be-

come available. This approach requires continuous or sequential production history conditioning techniques. The Kalman filter has been widely used for these types of time series problems. However, the Kalman filter is most appropriate when the number of variables in the model is low and the observations are linearly related to the model. This is not the case in spatio-temporal reservoir evaluation problems, where the number of model parameters are typically very high, and the relation between the reservoir model and the observed production, represented by a fluid flow simulator, is highly non-linear.

Several extensions to the Kalman filter techniques have been suggested, among these the ensemble Kalman filter, developed by Evensen (1994). The ensemble Kalman filter is used to update both the reservoir properties and the production properties. The computations are based on an ensemble of realisations of the reservoir and production properties, from which relevant statistics concerning the model uncertainty can be estimated. At times where new observations become available, all ensemble members are updated to honour these observations. Consequently, the realisations are always kept up to date with the latest observations. The ensemble Kalman filter methodology has been applied to numerous case studied in various fields of application, such as weather forecasting (Evensen (1994), Houtekamer and Mitchell (1998)), ground water hydrology (Reichle et al. (2002)), and petroleum engineering (Nævdal et al. (2002, 2003), Gu and Oliver (2004), Wen and Chen (2005)). For a review of recent progress see Evensen (2003).

The ensemble Kalman filter is shown to perform well with ensemble sizes around 100 members. In practice, however, the computational demands by fluid flow simulation on large reservoir models prohibits ensembles of this size. This problem is typically overcome by performing approximate fluid flow simulations on a coarser scale representation of the reservoir model. This upscaling is known to introduce bias, however, which should be accounted for. In this paper we use the general ensemble Kalman filtering framework of Evensen (1994), and extend it to correct for the effect of using coarse scale fluid flow simulators, using the approach in Omre and Lødøen (2004). The basic idea of Omre and Lødøen (2004) is to use coarse scale fluid flow simulation results to predict fine scale fluid flow simulation results, and to assess the associated prediction uncertainty. The fine scale representation is termed the reference scale. This correction is feasible if the coarse scale fluid flow simulations capture the most important features of the fine scale fluid flow simulations. We coin our approach scale-corrected ensemble Kalman filter.

This paper proceeds as follows. In Section 2 we start by defining the notation and describing the ensemble Kalman filter methodology. Then we motivate and present our model extensions. Section 3 presents the case study, which is inspired by the characteristics of the Troll field in the North Sea. In Section 4 we present and discuss the results from our simulation studies, and finally in Section 5 we draw some conclusions.

2 Model formulation

This section contains a general formulation of the stochastic model. More problem specific details are presented, together with a case study, in Section 3.

2.1 Notation

When evaluating a reservoir, a stochastic model for the state of the reservoir should be defined. In reservoir production forecasting problems, including production history conditioning, the state of the reservoir contain two types of uncertain variables. First, the reservoir properties are unknown, meaning that porosity, permeability, saturations and pressures are unknown over the reservoir. These properties vary spatially over the reservoir, and are usually represented on a three dimensional lattice covering the reservoir domain. Saturations and pressures are also dynamic properties and change in time, leading to the reservoir properties being spatio-temporal. The evaluation of the reservoir properties takes place at times $t \in [t_0, t_e]$, where t_0 represents time for production start, at which the reservoir is assumed to be in equilibrium, and t_e represents a suitable time for production completion. At any given time $t \in [t_0, t_e]$, let $R_t \in \Re^{n_r}$ be a vector containing the reservoir properties in each lattice node at time t. Here, n_r is the number of nodes in the lattice times the number of reservoir properties. The second uncertain variable is the production properties of the reservoir. This variable is related to the wells and represents pressures and rates of water and hydrocarbon production. At any given time $t \in [t_0, t_e]$, let $Q_t \in \Re^{n_q}$ be a vector containing the production properties from the reservoir at time t. Here, n_q is the number of production properties considered times the number of wells. Hence, the state of the reservoir at any given time $t \in [t_0, t_e]$ is described by the vector

$$\begin{bmatrix} R_t \\ Q_t \end{bmatrix} \in \Re^{n_r+n_q}, \tag{1}$$

which is referred to as the reservoir state vector. For notational convenience $r_i q_i$ refers to the reservoir state vector at time t_i.

At the initial time step the reservoir is in equilibrium, and the state of the reservoir is represented by an initial prior model, $f(r_0 q_0)$. At a later time, with given reservoir state vector, the forecast only depends on the current state and not the previous history. This entails the reservoir state vector appears as a Markov process in time. The forecasts can be made by a fluid flow simulator $\omega : \Re^{n_r+n_q} \to \Re^{n_r+n_q}$. The simulator ω replicates the flow of fluids through a porous medium. The input given to ω is the reservoir state vector at a given time s, in addition to a set of reservoir parameters, the well configuration and the well operating conditions. The reservoir parameters (rock compressibility, mobility ratios, relative permeability curves etc.), well configuration and well operating conditions are considered known, and are left unchanged throughout this work. Therefore, they are all hidden in the

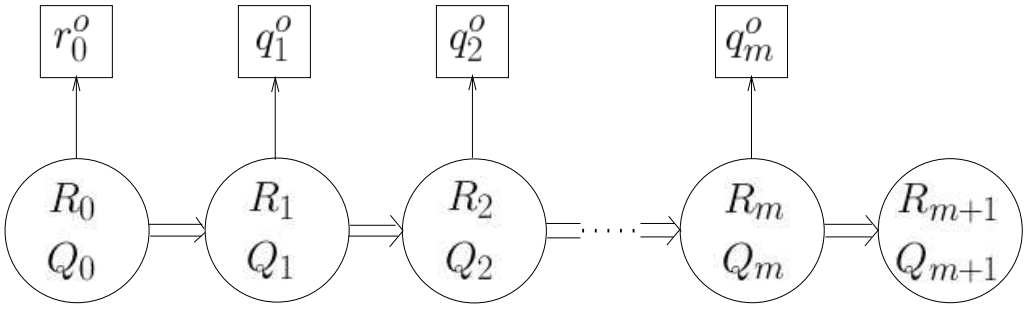

Figure 1: Graph of stochastic model with exact fluid flow simulator.

notation. The output from the simulator is the reservoir state vector at a given time $t \in [t_0, t_e]$, where $t > s$,

$$\begin{bmatrix} R_t \\ Q_t \end{bmatrix} = \omega_{t-s}\left(\begin{bmatrix} R_s \\ Q_s \end{bmatrix}\right). \tag{2}$$

Here subscript $t - s$ indicates the time for which the fluid flow simulation is performed. Expression (2) defines $f(r_t q_t | r_s q_s)$ to be a Dirac probability density function (pdf), where the fluid flow simulator is assumed to represent the true fluid flow process without error.

The evaluation takes place at a set of times $\mathcal{T} : \{t_0, t_1, \ldots, t_m, t_{m+1}\}$, where $t_0 < t_1 < \cdots < t_m < t_{m+1} < t_e$, as shown in Figure 1. The double-arrows between the reservoir state vectors at various time points indicate the deterministic relation between them. The prior model for all times evaluated then becomes:

$$f(r_0 q_0, r_1 q_1, \ldots, r_{m+1} q_{m+1}) = f(r_0 q_0) \prod_{i=1}^{m+1} f(r_i q_i | r_{i-1} q_{i-1}). \tag{3}$$

The observations from the reservoir consist of seismic data and well logs at time t_0, and measurements of the production properties at times $\mathcal{T}^o : \{t_1, t_2, \ldots, t_m\}$. Let r_0^o denote the static seismic data and well logs, which are related to the reservoir properties R_0 at time t_0. The likelihood model for the static data, r_0^o, given the initial reservoir state is defined by:

$$f(r_0^o | r_0) \sim \text{Gauss}(D_0 r_0, \Sigma_0^o), \tag{4}$$

where D_0 is a matrix tying the static data to the reservoir properties, and Σ_0^o is an observation error covariance matrix. Further, let $q_t^o \in \Re^{n_q}$ be a vector containing the observations of the production properties at times $t \in \mathcal{T}^o$. Each of the measurements has an associated measurement error. These measurement errors must be determined based on knowledge about the accuracy of the data acquisition equipment used. The likelihood model for the observed production is defined by:

$$f(q_t^o | q_t) \sim \text{Gauss}(D_t q_t, \Sigma_t^o), \tag{5}$$

where D_t is a matrix tying the observed production to the production part of the reservoir state vector, and Σ_t^o represents the uncertainty in the observations at time t. The observations are graphically displayed in Figure 1, and the single-arrows indicate the stochastic relations between reservoir state vectors and corresponding observations.

The prior and likelihood models define the posterior model of interest:

$$f(r_0q_0, \ldots, r_{m+1}q_{m+1}|r_0^o, q_1^o, \ldots, q_m^o)$$

$$= const \times f(r_0^o|r_0)f(r_0q_0)f(r_{m+1}q_{m+1}|r_mq_m)\prod_{i=1}^{m} f(q_i^o|q_i)f(r_iq_i|r_{i-1}q_{i-1}), \quad (6)$$

where *const* is a normalising constant. The marginal posterior at time t_{m+1} is of particular interest since it constitutes a forecast of the reservoir state vector, and it can be expressed as:

$$f(r_{m+1}q_{m+1}|r_0^o, q_1^o, \ldots, q_m^o)$$

$$= \int \cdots \int f(r_1q_1, \ldots, r_mq_m, r_{m+1}q_{m+1}|r_0^o, q_1^o, \ldots, q_m^o)dr_1q_1 \cdots dr_mq_m. \quad (7)$$

This marginal posterior pdf can be determined recursively with initial pdf at time t_0 where equilibrium is assumed:

$$f(r_0q_0|r_0^o) = const \times f(r_0^o|r_0)f(r_0q_0) \quad (8)$$

where *const* is a normalising constant. The recursive expression for r_iq_i conditioned on $(r_0^o, q_1^o, \ldots, q_{i-1}^o)$ is:

$$f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o) = \int f(r_iq_i|r_{i-1}q_{i-1})f(r_{i-1}q_{i-1}|r_0^o, q_1^o, \ldots, q_{i-1}^o)dr_{i-1}q_{i-1}. \quad (9)$$

Later in the paper realisations from this pdf will be denoted by:

$$\begin{bmatrix} R_i \\ Q_i \end{bmatrix} \sim f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o). \quad (10)$$

The expression for r_iq_i conditioned on $(r_0^o, q_1^o, \ldots, q_i^o)$, i.e. on all production observations prior to and including time t_i, is:

$$f(r_iq_i|r_0^o, q_1^o, \ldots, q_i^o) = const \times f(q_i^o|q_i)f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o)$$

$$= const \times f(q_i^o|q_i)\int f(r_iq_i|r_{i-1}q_{i-1})f(r_{i-1}q_{i-1}|r_0^o, q_1^o, \ldots, q_{i-1}^o)dr_{i-1}q_{i-1} \quad (11)$$

where *const* is a normalising constant. Later in the paper realisations from this pdf will be denoted by:

$$\begin{bmatrix} R_i^c \\ Q_i^c \end{bmatrix} \sim f(r_iq_i|r_0^o, q_1^o, \ldots, q_i^o). \quad (12)$$

We obtain the forecast pdf $f(r_{m+1}q_{m+1}|r_0^o, q_1^o, \ldots, q_m^o)$ in expression (7), which is the primary objective of this study, by performing the recursion in expression (11) up to $i = m$ and extending it to $i = m + 1$ through expression (9). Note that this recursive formulation makes it possible to condition on the production observations sequentially as they appear.

A sequential simulation algorithm can be defined as follows:

- Generate r_0q_0 from $f(r_0q_0|r_0^o)$
- For $i = 1, \ldots, m$
  - Generate r_iq_i from $f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o)$
  - Generate $r_i^cq_i^c$ from $f(r_iq_i|r_0^o, q_1^o, \ldots, q_i^o)$
- Generate $r_{m+1}q_{m+1}$ from $f(r_{m+1}q_{m+1}|r_0^o, q_1^o, \ldots, q_m^o)$

The problematic part of this algorithm is simulation from $f(r_iq_i|r_0^o, q_1^o, \ldots, q_i^o)$ defined in expression (11). The unknown normalising constant is computer demanding to determine. Several simulation algorithms are available for this step, for example Markov chain Monte Carlo (MCMC) algorithms, rejection sampling and sampling importance resampling (SIR) algorithms, see Omre (2000), but they all are computer demanding. Note, however, that if $f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o)$, $i = 1, \ldots, m$ in expression (11) are Gaussian and $f(q_i^o|q_i)$, $i = 1, \ldots, m$ are Gaussian as defined in expression (5), then $f(r_iq_i|r_0^o, q_1^o, \ldots, q_i^o)$, $i = 1, \ldots, m$ are also Gaussian and analytically tractable. The pdfs $f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o)$, $i = 1, \ldots, m$ as defined in expression (9) will not be Gaussian, however, due to the complexity of the fluid flow simulator ω.

In the current study $f(r_iq_i|r_0^o, q_1^o, \ldots, q_{i-1}^o)$, $i = 1, \ldots, m$ are approximated by Gaussian pdfs with empirically estimated parameters, hence the simulation from $f(r_iq_i|r_0^o, q_1^o, \ldots, q_i^o)$, $i = 1, \ldots, m$ are very computer efficient. This approximate approach corresponds to the ensemble Kalman filter, see Evensen (1994), which is explained in greater detail in the next section.

The computational demands of fluid flow simulators often prohibits solutions requiring repeated simulations on a fine scale representation, hence approximate fluid flow simulators are frequently used. Traditionally, the fluid flow simulator ω is only replaced by an approximation without accounting for the fact that this may introduce biases and change the error structure in the forecasts. The focus of the current study is to account for these effects. In the model we use, the dynamic part of the prior model, $f(r_tq_t|r_sq_s)$, is defined by:

$$\begin{bmatrix} R_t \\ Q_t \end{bmatrix} = \omega_{t-s}^* \left(\begin{bmatrix} R_s \\ Q_s \end{bmatrix} \right) + \varepsilon_{t,s}^*, \tag{13}$$

where $\omega_{t-s}^* : \Re^{n_r+n_q} \rightarrow \Re^{n_r+n_q}$ is the approximate model and $\varepsilon_{t,s}^*$ is a centred Gaussian approximation error with covariance matrix $\Sigma_{t,s}^*$. The associated graph is presented in Figure 2, and it is worth noting that the graphs in Figure 1 and 2 are

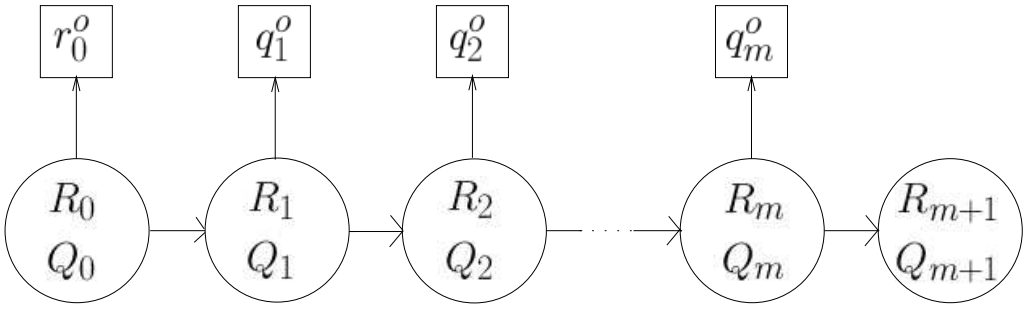

Figure 2: Graph of stochastic model with approximate fluid flow simulator.

identical except for the stochastic relations in the dynamic part of the prior model of the latter.

The approximate model ω^*_{t-s} will normally consist of three parts: upscaling, coarse scale fluid flow simulation, and downscaling/correction. Upscaling of reservoir properties entails:

$$\begin{bmatrix} R^*_s \\ Q^*_s \end{bmatrix} = \begin{bmatrix} \nu^*(R_s) \\ Q_s \end{bmatrix}, \tag{14}$$

where $\nu^* : \Re^{n_r} \rightarrow \Re^{n_{r^*}}$ is an upscaling operator which provides a coarser representation of the reservoir properties, see Farmer (2000) and Durlofsky (2003). Here, n_{r^*} is the number of nodes on the coarse lattice times the number of reservoir properties. Coarse scale fluid flow simulation entails:

$$\begin{bmatrix} R^*_t \\ Q^*_t \end{bmatrix} = \bar{\omega}_{t-s}\left(\begin{bmatrix} R^*_s \\ Q^*_s \end{bmatrix}\right), \tag{15}$$

where $\bar{\omega}_{t-s} : \Re^{n_{r^*}+n_q} \rightarrow \Re^{n_{r^*}+n_q}$ is a fluid flow simulator similar to ω_{t-s} but on a coarser scale. Downscaling of reservoir properties and correction of production properties entails:

$$\omega^*_{t-s}\left(\begin{bmatrix} R_s \\ Q_s \end{bmatrix}\right) = \begin{bmatrix} \mu_{R_t} + A_{R_t} R^*_t \\ \mu_{Q_t} + A_{Q_t} Q^*_t \end{bmatrix}, \tag{16}$$

where μ_{R_t} and μ_{Q_t} are shift vectors, and A_{R_t} and A_{Q_t} are correction matrices of dimension $(n_r \times n_{r^*})$ and $(n_q \times n_q)$ respectively. The covariance matrix $\Sigma^*_{t,s}$ in expression (13) contains $\mathrm{Var}\{R_t|R^*_t\}$ and $\mathrm{Var}\{Q_t|Q^*_t\}$. The assessment of the various parts of the approximate model ω^* will be described in greater detail in Section 2.3.

2.2 Traditional ensemble Kalman filtering

The ensemble Kalman filter has been successfully applied to relatively small production history conditioning problems. Nævdal et al. (2003) and Wen and Chen (2005) use the ensemble Kalman filter on two dimensional problems of size 39×55 and 50×50, respectively. Gu and Oliver (2004) consider a three dimensional problem of size 19×28×5.

The ensemble Kalman filter uses a set of realisations of reservoir state vectors, referred to as an ensemble, which assimilate the observations as they become available with time. The reservoir state covariance matrix at any given time is estimated from the ensemble of realisations. Hence, there is no need to neither integrate any covariance or sensitivity matrix forward in time, nor to assume that the forecasting step is linear as is done in regular Kalman filtering.

In practice, the ensemble Kalman filtering proceeds as follows: At the initial time-step, t_0, n initial reservoir state vectors are generated from the prior model, $f(r_0 q_0)$, updated by the static observations available at t_0:

$$\left[\begin{array}{c} R^c_{0,i} \\ Q^c_{0,i} \end{array} \right] \; ; \; i \in \Gamma = \{1, 2, \ldots, n\} \;\; \text{iid} \;\; \text{from} \;\; f(r_0 q_0 | r^o_0), \tag{17}$$

with index i referring to ensemble member number i.

By using the fluid flow simulator, each member of the ensemble is advanced from time t_0 to time t_1, to get a forecast of the production:

$$\left[\begin{array}{c} R_{1,i} \\ Q_{1,i} \end{array} \right] = \omega_1 \left(\left[\begin{array}{c} R^c_{0,i} \\ Q^c_{0,i} \end{array} \right] \right) \; ; \; i \in \Gamma. \tag{18}$$

These variables are realisations from $f(r_1 q_1 | r^o_0)$.

The ensemble at time t_1 conditioned on observations at t_0 is used to estimate the reservoir state vector covariance matrix Σ_1 at that time,

$$\Sigma_1 = \left[\begin{array}{cc} \Sigma_{R_1 R_1} & \Sigma_{R_1 Q_1} \\ \Sigma_{Q_1 R_1} & \Sigma_{Q_1 Q_1} \end{array} \right]. \tag{19}$$

The covariance matrix Σ_1 represents the uncertainty in the forecasts of the reservoir state vector at time t_1 given r^o_0 and the estimate is found through standard statistical estimation.

By using the estimated covariance matrices, each reservoir state vector in the ensemble is updated to honour the observed production at time t_1:

$$\left[\begin{array}{c} R^c_{1,i} \\ Q^c_{1,i} \end{array} \right] = \left[\begin{array}{c} R_{1,i} + \Sigma_{R_1 Q_1} \left[\Sigma_{Q_1 Q_1} + \Sigma^o_1 \right]^{-1} \left(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i} \right) \\ Q_{1,i} + \Sigma_{Q_1 Q_1} \left[\Sigma_{Q_1 Q_1} + \Sigma^o_1 \right]^{-1} \left(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i} \right) \end{array} \right] \; ; \; i \in \Gamma, \tag{20}$$

where

$$\varepsilon^o_{1,i} \; ; \; i \in \Gamma \text{ iid from } Gauss(0, \Sigma^o_1). \tag{21}$$

Here, $\varepsilon^o_{1,i}$ represents the observation error, which is assumed to be independent of all other variables. In literature the update step is referred to as the analysis step. It is easy to show that this corresponds to sampling from the approximate posterior distribution conditioned on the production history up to time t_1, $f(r_1 q_1 | r^o_0, q^o_1)$. Thus the correct variability is reproduced. Note that expression (20) only requires estimates of $\Sigma_{R_1 Q_1}$ and $\Sigma_{Q_1 Q_1}$, not $\Sigma_{R_1 R_1}$. The latter can be cumbersome to estimate if n_r is huge.

The forecasting and analysis steps are repeated for $\{t_2, \ldots, t_m\}$ when new observations are available. Forecasts of reservoir states at t_{m+1} are obtained by:

$$\begin{bmatrix} R_{m+1,i} \\ Q_{m+1,i} \end{bmatrix} = \omega_1 \left(\begin{bmatrix} R^c_{m,i} \\ Q^c_{m,i} \end{bmatrix} \right) \; ; \; i \in \Gamma. \tag{22}$$

Experience from previous work suggests using ensemble sizes of 40 (Gu and Oliver (2004)), 100 (Nævdal et al. (2003)) or 200 members (Wen and Chen (2005)). If the reservoir representation is on a fine scale, simulation of fluid flow on ensembles of these sizes is not possible, considering that one fluid flow simulation could take days or even weeks to complete. This problem is normally solved by upscaling the reservoir state vector to a coarser representation. The forecast and analysis steps follow expressions (17) through (22), by upscaling the initial ensemble using ν^*, see expression (14), and replacing R, Q and ω by R^*, Q^* and ω^*, respectively. In the following, we refer to this as traditional coarse scale ensemble Kalman filtering.

The upscaling changes the model error. It introduces bias, and the variability are most likely much more significant on the coarse scale. The coarse scale model error is ignored in the traditional ensemble Kalman filtering, however. It is widely known, see Omre and Lødøen (2004), that such simplifications inevitably leads to bias problems in the production history conditioning and underestimation of the variability in the ensemble. Hence, by not considering the bias in the coarse scale production forecasts, biased production conditioned reservoir state vectors on the fine scale are identified. Moreover, by ignoring the modelling error, the coarse scale and fine scale fluid flow simulators are treated as equally reliable. The real loss in reliability caused by using the coarse scale simulator is not quantified.

2.3 Scale-corrected ensemble Kalman filtering

Estimating the relation between the production properties from a coarse scale fluid flow simulator and from a fine scale fluid flow simulator have been suggested earlier, see Omre and Lødøen (2004) and Lødøen et al. (2005). This idea can easily be extended to an ensemble Kalman filter setting, where we estimate the relation between the reservoir state vectors at different scales. To determine this relation from general understanding of the physical processes involved is difficult. Therefore, as in Omre and Lødøen (2004), we propose an empirical, statistical approach.

First, we generate a small set of realisations of the initial reservoir state vector from the prior model, $f(r_0 q_0)$, updated by the static observations available at t_0, to get an ensemble of initial fine scale reservoir state vectors:

$$\begin{bmatrix} R^c_{0,i} \\ Q^c_{0,i} \end{bmatrix} \; ; \; i \in \Gamma = \{1, \ldots, n_c\} \;\; \text{iid} \;\; \text{from} \;\; f(r_0 q_0 | r^o_0). \tag{23}$$

This scale is termed the reference scale, and the set is termed the calibration set. The size of this ensemble should be such that we can afford the computational cost

of running all the fine scale fluid flow simulations. Further, we upscale the initial ensemble of fine scale reservoir state vectors to get the corresponding coarse scale ensemble:

$$\begin{bmatrix} R^{*c}_{0,i} \\ Q^{*c}_{0,i} \end{bmatrix} = \begin{bmatrix} \nu^*(R_{0,i}) \\ Q_{0,i} \end{bmatrix} \; ; \; i \in \Gamma. \tag{24}$$

The forecasts of the reservoir states at the next time-step are found using the fine scale and coarse scale fluid flow simulators, where the fine scale simulator, operating on the reference scale, is assumed to simulate without error, and the coarse scale simulator is assumed to introduce a coarse scale model error:

$$\begin{aligned} \begin{bmatrix} R_{1,i} \\ Q_{1,i} \end{bmatrix} &= \omega_1 \left(\begin{bmatrix} R^{c}_{0,i} \\ Q^{c}_{0,i} \end{bmatrix} \right) \\ \begin{bmatrix} R^{*}_{1,i} \\ Q^{*}_{1,i} \end{bmatrix} &= \bar{\omega}_1 \left(\begin{bmatrix} R^{*c}_{0,i} \\ Q^{*c}_{0,i} \end{bmatrix} \right) \end{aligned} \; ; \; i \in \Gamma. \tag{25}$$

The uncertainty in and dependence of the forecasts of the fine and coarse scale reservoir state vectors are described by a joint pdf. The fine and coarse scale ensemble members are used to find an estimate of this joint pdf, described by a mean vector and a covariance matrix:

$$\begin{bmatrix} \mu_{R_1} \\ \mu_{Q_1} \\ \mu_{R^*_1} \\ \mu_{Q^*_1} \end{bmatrix} \begin{bmatrix} \Sigma_{R_1R_1} & \Sigma_{R_1Q_1} & \Sigma_{R_1R^*_1} & \Sigma_{R_1Q^*_1} \\ \Sigma_{Q_1R_1} & \Sigma_{Q_1Q_1} & \Sigma_{Q_1R^*_1} & \Sigma_{Q_1Q^*_1} \\ \Sigma_{R^*_1R_1} & \Sigma_{R^*_1Q_1} & \Sigma_{R^*_1R^*_1} & \Sigma_{R^*_1Q^*_1} \\ \Sigma_{Q^*_1R_1} & \Sigma_{Q^*_1Q_1} & \Sigma_{Q^*_1R^*_1} & \Sigma_{Q^*_1Q^*_1} \end{bmatrix} \tag{26}$$

Under linearity assumptions, we later find conditional distributions that describe the ability of the coarse scale reservoir vectors to predict the fine scale reservoir state vector from the joint distribution above. In practice, as will be shown later, not all parts of the covariance matrix needs to be estimated at each time-step.

The ensemble of forecasted reservoir states is then updated to honour the production history using the covariance matrix estimates:

$$\begin{bmatrix} R^{c}_{1,i} \\ Q^{c}_{1,i} \end{bmatrix} = \begin{bmatrix} R_{1,i} + \Sigma_{R_1Q_1}[\Sigma_{Q_1Q_1} + \Sigma^o_1]^{-1}(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i}) \\ Q_{1,i} + \Sigma_{Q_1Q_1}[\Sigma_{Q_1Q_1} + \Sigma^o_1]^{-1}(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i}) \end{bmatrix} \; ; \; i \in \Gamma. \tag{27}$$

These variables are realisations from $f(r_1q_1|r^o_0, q^o_1)$, based on the fine scale reference model. These updates follow regular ensemble Kalman filtering theory, where we consider the reference, fine scale model error to be negligible. As the conditioning is thought to be more accurate on the fine scale than on the coarse scale, the coarse scale conditioning is based on the fine scale forecasts:

$$\begin{bmatrix} R^{*c}_{1,i} \\ Q^{*c}_{1,i} \end{bmatrix} = \begin{bmatrix} R^{*}_{1,i} + \Sigma_{R^*_1Q_1}[\Sigma_{Q_1Q_1} + \Sigma^o_1]^{-1}(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i})) \\ Q^{*}_{1,i} + \Sigma_{Q^*_1Q_1}[\Sigma_{Q_1Q_1} + \Sigma^o_1]^{-1}(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i}) \end{bmatrix} \; ; \; i \in \Gamma. \tag{28}$$

These variables are realisations of the coarse scale fluid flow simulations conditioned to the available observations prior to and including time t_1. Note that the updates of $R^*_{1,i}$ and $Q^*_{1,i}$ can be rewritten as:

$$\begin{aligned} R^{*c}_{1,i} &= R^*_{1,i} + \Sigma_{R^*_1 Q_1}\Sigma^{-1}_{Q_1Q_1}\Sigma_{Q_1Q_1}[\Sigma_{Q_1Q_1} + \Sigma^o_1]^{-1}(q^o_1 + \varepsilon^o_{1,i} - Q_{1,i}) \\ &= R^*_{1,i} + \Sigma_{R^*_1 Q_1}\Sigma^{-1}_{Q_1Q_1}[Q^c_{1,i} - Q_{1,i}] \qquad ;\ i \in \Gamma, \quad (29) \\ Q^{*c}_{1,i} &= Q^*_{1,i} + \Sigma_{Q^*_1 Q_1}\Sigma^{-1}_{Q_1Q_1}[Q^c_{1,i} - Q_{1,i}] \end{aligned}$$

such that the coarse scale updates are conditioned on the difference between the forecasted and updated production vector on the fine scale.

The forecast and analysis steps are repeated for $\{t_2, \ldots, t_m\}$ where new observations are available. Forecasts of reservoir states at t_{m+1} are obtained by:

$$\begin{aligned} \begin{bmatrix} R_{m+1,i} \\ Q_{m+1,i} \end{bmatrix} &= \omega_1 \left(\begin{bmatrix} R^c_{m,i} \\ Q^c_{m,i} \end{bmatrix} \right) \\ \begin{bmatrix} R^*_{m+1,i} \\ Q^*_{m+1,i} \end{bmatrix} &= \bar{\omega}_1 \left(\begin{bmatrix} R^{*c}_{m,i} \\ Q^{*c}_{m,i} \end{bmatrix} \right) \end{aligned} \quad ;\ i \in \Gamma. \qquad (30)$$

At this point we have an empirical estimate of the relation between the reservoir state vectors on the fine and coarse scales for all time-steps $t \in \mathcal{T}$ where we evaluate the reservoir. We refer to this as the calibration part of scale-corrected ensemble Kalman filtering.

Simulation of reservoir state vectors on a coarse scale and the empirical estimates of expression (26) are then used to generate realisations of the reservoir state vectors and production forecasts on a fine scale. We refer to the following as the simulation part of scale-corrected ensemble Kalman filtering. First, we generate a large set of realisations of the initial reservoir state vector from the prior model, $f(r_0 q_0)$, updated by the static observations available at t_0, and upscale the realisations to get an initial ensemble:

$$\begin{aligned} \begin{bmatrix} R^c_{0,j} \\ Q^c_{0,j} \end{bmatrix} &\text{ iid from } f(r_0 q_0 | r^o_0) \\ \begin{bmatrix} R^{*c}_{0,j} \\ Q^{*c}_{0,j} \end{bmatrix} &= \begin{bmatrix} \nu^*(R^c_{0,j}) \\ Q^c_{0,j} \end{bmatrix} \end{aligned} \quad ;\ j \in \Gamma^* = \{1, \ldots, n_s\}. \qquad (31)$$

Note that since the fluid flow simulations are now only performed on the coarse scale, n_s can be much larger than n_c. The forecasts of the coarse scale reservoir state vectors at the next time-step are found using the coarse scale fluid flow simulator:

$$\begin{bmatrix} R^*_{1,j} \\ Q^*_{1,j} \end{bmatrix} = \bar{\omega}_1 \left(\begin{bmatrix} R^{*c}_{0,j} \\ Q^{*c}_{0,j} \end{bmatrix} \right) \ ;\ j \in \Gamma^*. \qquad (32)$$

At each time-step the downscaled forecast of the reservoir properties can be generated using the mean and covariance matrix estimates, together with the set of coarse scale forecasts,

$$[R_1|R_1^*] \sim Gauss\left(\mu_{R_1} + \Sigma_{R_1R_1^*}\Sigma^{-1}_{R_1^*R_1^*}(R_1^* - \mu_{R_1^*}), \Sigma_{R_1R_1} - \Sigma_{R_1R_1^*}\Sigma^{-1}_{R_1^*R_1^*}\Sigma_{R_1^*R_1}\right), \tag{33}$$

hence a realisation is:

$$R_{1,j} = \mu_{R_1} + \Sigma_{R_1R_1^*}\Sigma^{-1}_{R_1^*R_1^*}(R^*_{1,j} - \mu_{R_1^*}) + \tilde{\varepsilon}_{1,j} \ ; \ j \in \Gamma^*, \tag{34}$$

where

$$\tilde{\varepsilon}_{1,j} \ ; \ j \in \Gamma^* \ \text{ iid from } \ Gauss(0, \Sigma_{R_1R_1} - \Sigma_{R_1R_1^*}\Sigma^{-1}_{R_1^*R_1^*}\Sigma_{R_1^*R_1}). \tag{35}$$

Note that sampling $R_{1,j}$ is not straightforward, since $\Sigma_{R_1R_1}$ and $\Sigma_{R_1^*R_1^*}$ do not have full rank, but this is solved through Bayesian regularisation, see Appendix A.

The forecasts of the production properties on the coarse scale are used to generate a set of corresponding production properties on the reference, fine scale, by using the mean and covariance estimates from the calibration. The production property correction is:

$$[Q_1|Q_1^*] \sim Gauss\left(\mu_{Q_1} + \Sigma_{Q_1Q_1^*}\Sigma^{-1}_{Q_1^*Q_1^*}(Q_1^* - \mu_{Q_1^*}), \Sigma_{Q_1Q_1} - \Sigma_{Q_1Q_1^*}\Sigma^{-1}_{Q_1^*Q_1^*}\Sigma_{Q_1^*Q_1}\right), \tag{36}$$

hence

$$Q_{1,j} = \mu_{Q_1} + \Sigma_{Q_1Q_1^*}\Sigma^{-1}_{Q_1^*Q_1^*}(Q^*_{1,j} - \mu_{Q_1^*}) + \varepsilon^*_{1,j} \ ; \ j \in \Gamma^*, \tag{37}$$

where

$$\varepsilon^*_{1,j} \ ; \ j \in \Gamma^* \ \text{ iid from } \ Gauss(0, \Sigma_{Q_1Q_1} - \Sigma_{Q_1Q_1^*}\Sigma^{-1}_{Q_1^*Q_1^*}\Sigma_{Q_1^*Q_1}). \tag{38}$$

Under linearity assumptions, this ensures that the set has the correct mean and variance. The updates are performed using the set of predicted fine scale production properties:

$$\begin{bmatrix} R^{*c}_{1,j} \\ Q^{*c}_{1,j} \end{bmatrix} = \begin{bmatrix} R^*_{1,j} + \Sigma_{R_1^*Q_1}[\Sigma_{Q_1Q_1} + \Sigma_1^o]^{-1}(q_1^o + \varepsilon^o_{1,j} - Q_{1,j}) \\ Q^*_{1,j} + \Sigma_{Q_1^*Q_1}[\Sigma_{Q_1Q_1} + \Sigma_1^o]^{-1}(q_1^o + \varepsilon^o_{1,j} - Q_{1,j}) \end{bmatrix} \ ; \ j \in \Gamma^*. \tag{39}$$

This updating directly on R_1^* is an alternative to downscaling R_1^* to R_1, updating R_1 and then upscale R_1^c to R_1^{*c}. This bypass makes it possible to avoid downscaling for times where only production characteristics are of interest.

The simulation part of the scale-corrected ensemble Kalman filtering is repeated for $\{t_2, \ldots, t_m\}$ when new observations are available. In practice, we do the calibration part for one time step first and then the simulation part. Then we move on to the next time-step. Hence, the calibration and simulation ensembles are always kept up to date with the latest conditioning. Forecasts of reservoir states in the simulation ensemble at t_{m+1} are obtained by forecasting the coarse scale reservoir states:

$$\begin{bmatrix} R^*_{m+1,j} \\ Q^*_{m+1,j} \end{bmatrix} = \bar{\omega}_1\left(\begin{bmatrix} R^{*c}_{m,j} \\ Q^{*c}_{m,j} \end{bmatrix}\right) \ ; \ j \in \Gamma, \tag{40}$$

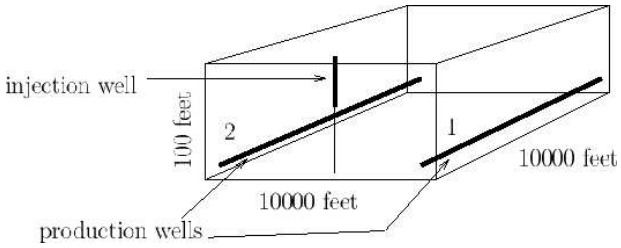

Figure 3: Outline of the reservoir. The thick lines indicate where the wells are perforated.

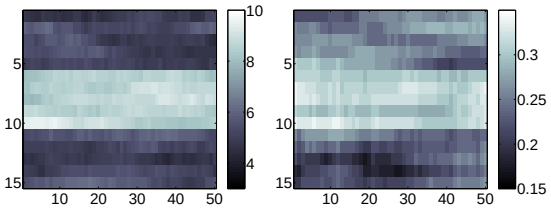

Figure 4: Vertical slice of true log-permeability (left) and true porosity (right).

and then generating the corresponding reference, fine scale reservoir states:

$$\begin{bmatrix} R_{m+1,j} \\ Q_{m+1,j} \end{bmatrix} = \begin{bmatrix} \mu_{R_{m+1}} + \Sigma_{R_{m+1}R^*_{m+1}} \Sigma^{-1}_{R^*_{m+1}R^*_{m+1}} (R^*_{m+1,j} - \mu_{R^*_{m+1}}) + \tilde{\varepsilon}_{m+1,j} \\ \mu_{Q_{m+1}} + \Sigma_{Q_{m+1}Q^*_{m+1}} \Sigma^{-1}_{Q^*_{m+1}Q^*_{m+1}} (Q^*_{m+1,j} - \mu_{Q^*_{m+1}}) + \varepsilon^*_{m+1,j} \end{bmatrix}; j \in \Gamma^* \tag{41}$$

where $\tilde{\varepsilon}_{m+1,j}$ and $\varepsilon^*_{m+1,j}$ are defined as in expressions (35) and (38), respectively.

3 Case study

The case study is identical to the one in Omre and Lødøen (2004). The study is inspired by the Troll field in the North Sea offshore Norway. The reservoir under study covers a domain of size $10^4 \times 10^4 \times 10^2$ feet^3 and is discretised onto a lattice of size $50 \times 50 \times 15$. Initially, the reservoir is fully saturated with oil. The reservoir has one vertical injection well, perforated only in the upper part of the reservoir, and two horizontal production wells. The well configuration is shown in Figure 3.

A reference reservoir is created by geoscientists familiar with the Troll field. This is done so that the results we get can be compared to true values, to verify the strengths and weaknesses of the methodology. Specific details on the construction can be found in Hegstad and Omre (2001). Vertical cross sections of the static properties (porosity and log-permeability) in the reference reservoir are displayed in Figure 4. The reference reservoir has a clearly layered structure where the middle part of the reference reservoir has higher permeability and porosity than the upper and lower parts, but it also contain considerable heterogeneity.

By using the geology in the reference reservoir as input to a fluid flow simulator, we find the reference production. In this work we use Eclipse 300 version 2004a

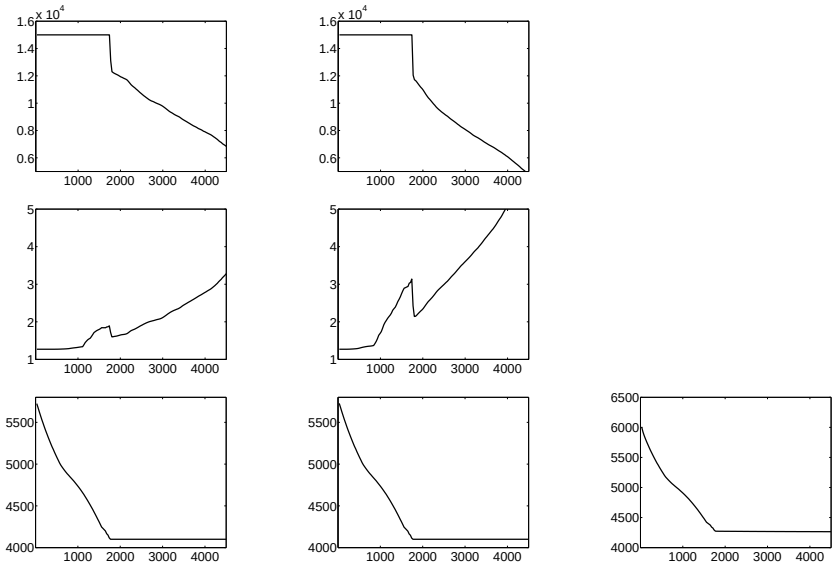

Figure 5: The reference production, shown row-wise from top to bottom, for: $opr1$, $opr2$, $gor1$, $gor2$, $bhp1$, $bhp2$, and $bhpi$.

simulator, see GeoQuest (2004), as the fluid flow simulators ω and ω^*, where ω simulates on the full $50 \times 50 \times 15$ lattice, and ω^* simulates on a coarser $10 \times 10 \times 15$ lattice. The upscaling function $\nu^* : \Re^{150,000} \rightarrow \Re^{6,000}$ is defined in the following way: The pressure, fluid saturations and porosity are mapped to the coarse grid by arithmetic averaging, while the permeability is upscaled by harmonic averaging.

The reference production is found using ω, and it is shown in Figure 5. The production properties we are considering are the oil production rates ($opr1$ and $opr2$), gas/oil ratios ($gor1$ and $gor2$), and bottom hole pressures ($bhp1$ and $bhp2$) in both production wells, and the bottom hole pressure in the injection well ($bhpi$). The reference production clearly demonstrates how the well operating conditions are defined. Initially, each production well produces at a target oil production rate of 15,000 stb/day. When the bottom hole pressure in a production well drops to 4,100 psi, the operating condition in this well switches to a target bottom hole pressure of 4,100 psi. The injected fluid is gas, and the injection happens at a target rate of 65,000 mscf/day.

The observed well logs and seismic data are generated from the reference reservoir. Fairly exact observations in the wells and convolved post-stack seismic data are made, and this constitute the observations r_0^o at initial time. For more details see Hegstad and Omre (2001).

The observed production history is found from the reference production. We consider the reference production up to 1,800 days, at 30 day intervals, to be the observed production history. Hence, the observed production history contains measurements of seven production properties at sixty times,

$$q_t^o \in \Re^7, \ t \in \{t_1, t_2, \ldots, t_{60}\}. \tag{42}$$

The standard deviations in the measurement errors are defined to be one percent of the observed value for $opr1$, $opr2$, $bhp1$, $bhp2$ and $bhpi$, and 25 percent of the observed value for $gor1$ and $gor2$. This defines the measurement error covariance matrix Σ_t^o at each time-step t.

3.1 Stochastic model for initial reservoir state vectors

The initial reservoir state vector is found from an initial model, $f(r_0 q_0 | r_0^o)$. This is done by generating samples from a prior model, $f(r_0 q_0)$, updated by static observations, r_0^o, at the initial time step. The porosity and permeability are generated from the model in Hegstad and Omre (2001), using the same model parameters as Omre and Lødøen (2004). Specific details about how this is done are not given here. Note, however, that the distribution is conditioned on seismic data and well logs, and that it is easy to sample from, in the sense that samples can be generated relatively fast. The initial pressure and saturations over the reservoir are equal in all realisations since the reservoir is assumed to be in equilibrium. The oil-saturation is one everywhere, and the pressure is equal to the steady state pressure in each grid block, defined by setting a constant initial pressure of 5,780 psi at the top of the reservoir. The initial production rates and production ratios are zero, while the bottom hole pressures in the wells are equal to the steady state pressure in the grid block where the bottom hole pressure is measured. This defines the model $f(r_0 q_0 | r_0^o)$ from which we generate the initial reservoir state vectors.

4 Simulation studies

As mentioned earlier, conditioning the forecasted coarse scale reservoir state vectors to the observed production history leads to bias in the production history conditioned fine scale reservoir properties. We will quantify this effect by comparing the results we get from performing the traditional coarse scale ensemble Kalman filtering, described in Section 2.2, to the results we get using the scale-corrected ensemble Kalman filtering methodology, outlined in Section 2.3, on the case study described in Section 3.

4.1 Traditional coarse scale ensemble Kalman filtering

An initial ensemble of $n = 100$ reservoir state vectors is generated from the initial model $f(r_0 q_0 | r_0^o)$. From this initial fine scale ensemble of reservoir state vectors, the corresponding initial coarse scale ensemble is found, using the upscaling function ν^*. The traditional coarse scale ensemble Kalman filtering is performed for 1,800 days, conditioning each ensemble member to the observed production history every 30 days.

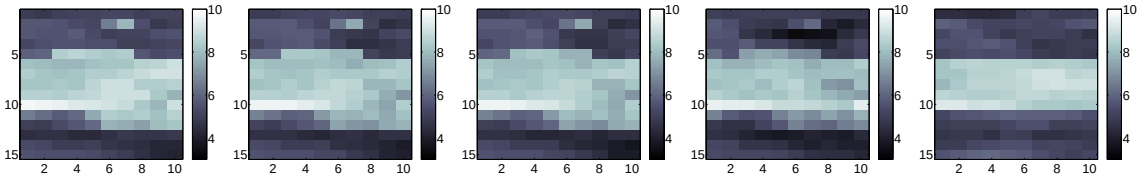

Figure 6: Traditional coarse scale ensemble Kalman filtering: Vertical slice of the coarse scale log-permeability in one of the ensemble members at selected times. The times considered are, from left to right; the initial time-step, and after 300 days, 900 days and 1,800 days of updating. The rightmost plot shows the respective slice of the upscaled reference reservoir.

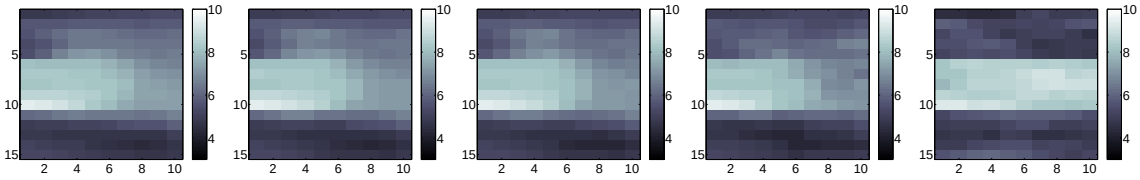

Figure 7: Traditional coarse scale ensemble Kalman filtering: Vertical slice of the mean coarse scale log-permeability at selected times. The times considered are, from left to right; the initial time-step, and after 300 days, 900 days and 1,800 days of updating. The rightmost plot shows the respective slice of the upscaled reference reservoir.

Figure 6 shows a vertical slice of the log-permeability for one of the ensemble members at the initial time-step, and after 300 days, 900 days and 1,800 days of traditional coarse scale ensemble Kalman filtering, compared to the upscaled reference log-permeability.

It is clear that the high permeable middle layer gets more and more distinct as more production data are assimilated into the model. Vertical slices of the mean log-permeability in the ensemble at the same time steps are shown in Figure 7. The initial mean log-permeability is not far from the true values, but the interfaces between the high and low permeable layers are not sharp. This is due to the fact that the location of the high permeable layer is uncertain in the initial model. As more and more production data are incorporated into the model, these interfaces become sharper, especially in the areas close to the injection well, which is located in the leftmost area of the vertical slices. The area to the right in the vertical slices are further away from the wells. The conditioning to the well logs does not have a large impact there. Also, because of the distance to both the injection well and the production wells, this region experiences little or no changes in saturations during the time window we are considering. Hence, the area has less impact on the fluid flow calculations, and thus we should not expect to accurately reproduce the true values here. Figure 8 shows the corresponding vertical slices of the standard deviation for the log-permeability. This is the square root of the respective diagonal

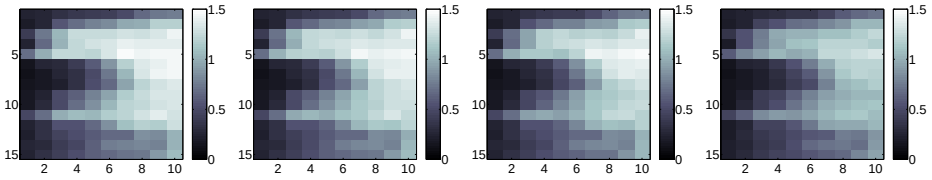

Figure 8: Traditional coarse scale ensemble Kalman filtering: Vertical slice of the standard deviation for the coarse scale log-permeability at selected times. The times considered are, from left to right; the initial time-step, and after 300 days, 900 days and 1,800 days of updating.

elements of $\Sigma_{R_t^* R_t^*}$. The initial uncertainty is low in the leftmost and lower areas of the vertical slice. These areas are close to the injection well and the production wells, respectively. Since we condition on well logs in the initial model, the initial uncertainty is lower in these areas of the reservoir. The standard deviation estimates get lower and lower as more and more observations are assimilated.

The same tendencies are seen in the development of the porosity field, when considering a single ensemble member, the mean and the standard deviation of the coarse scale ensemble. This is not shown here.

Since the Kalman updates are linear and the fluid flow equations are non-linear, there might be inconsistencies between the corrected static properties (permeability and porosity) and the corrected dynamic properties (pressure, saturation and production properties). Therefore, we use the updated static reservoir properties at selected times, as input to the coarse scale reservoir simulator ω^* and re-run the simulations from the initial time-step without conditioning to the observations. Figure 9 shows the simulation results for all the production properties considered for the twenty first members of the ensemble when using the updated ensemble at selected time-steps as input to ω^*, compared to the fine scale reference production. Initially the uncertainty in the forecasts is large. As more and more production data are assimilated into the reservoir models, the uncertainty is reduced. At intermediate times, especially after 900 days of coarse scale ensemble Kalman filtering, biases are observed. The gas-breakthrough times and the times where the well controls switch from target rate to target bottom hole pressure happen to late. At 1,800 days the bias is corrected. This can be explained by the fact that at 900 days neither gas-breakthrough nor change of production controls have been observed in the production history. The production forecasts after 1,800 days of coarse scale ensemble Kalman filtering are close to the fine scale reference production. Note especially that the gas-breakthrough times and the times where the well controls switch from target rate to target bottom hole pressure are well reproduced.

However, this does not mean that the corresponding fine scale simulation runs match the observed production history. Figure 10 shows the simulation results for all production properties considered for the twenty first members of the ensemble using

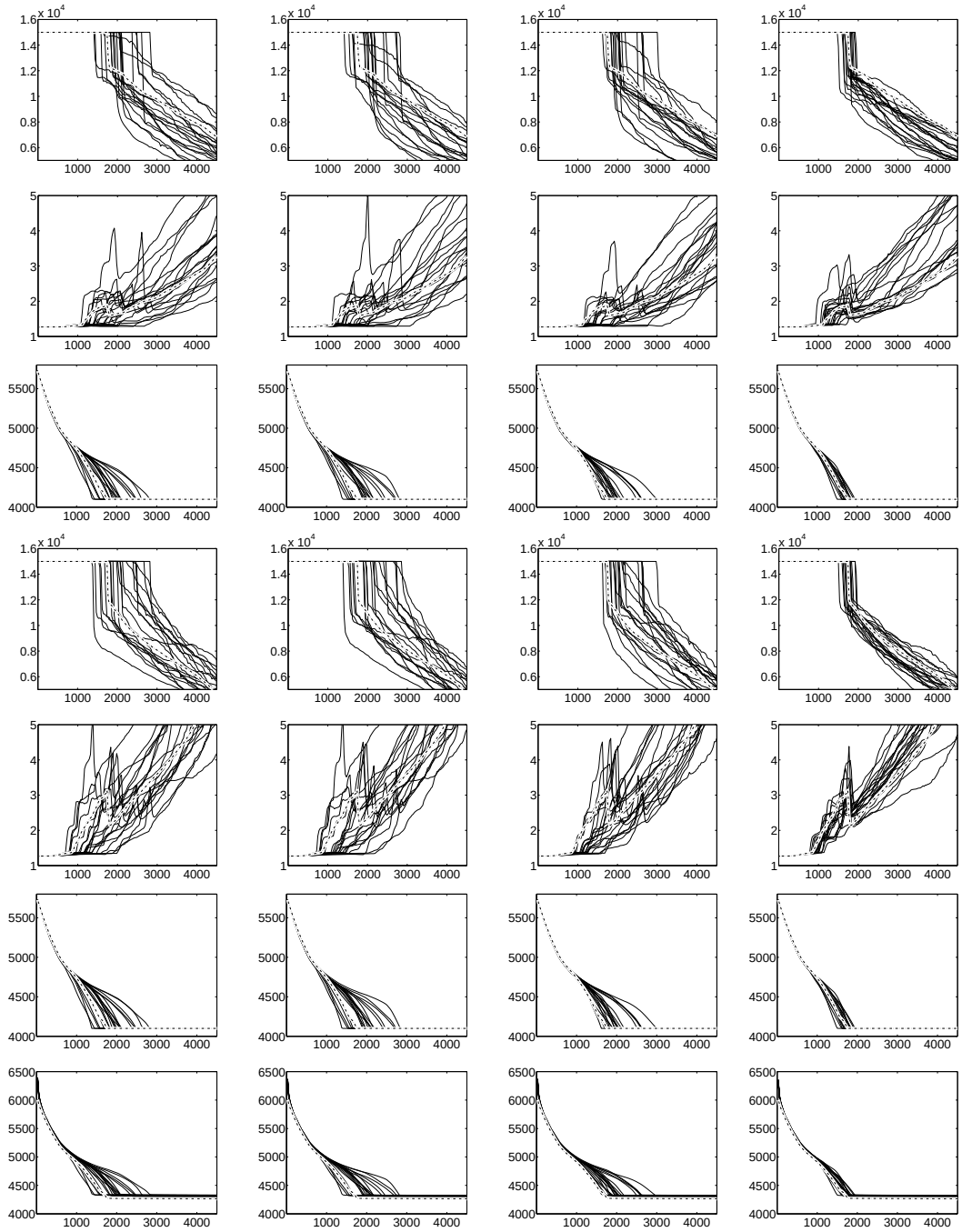

Figure 9: Traditional coarse scale ensemble Kalman filtering: Production properties from re-runs of the updated coarse scale ensemble at selected times, compared to the corresponding fine scale reference production properties (–·–). The columns represent fluid flow simulations with different inputs, while the rows represent the different production properties. The inputs used are, from left to right; the initial ensemble, the ensemble after 300 days, 900 days, and 1,800 days of updating. The production properties considered are, from top to bottom; $opr1$, $gor1$, $bhp1$, $opr2$, $gor2$, $bhp2$ and $bhpi$.

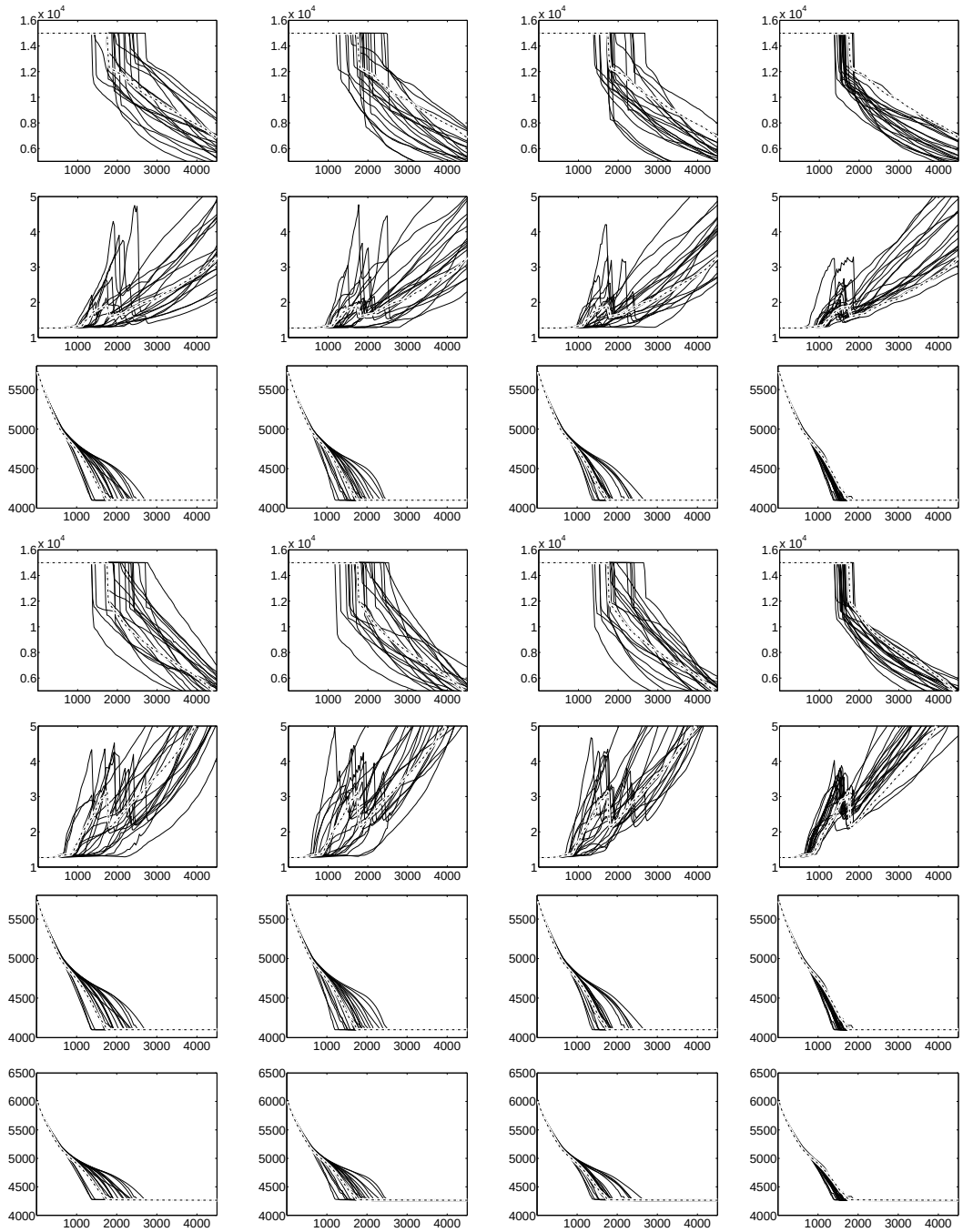

Figure 10: Traditional coarse scale ensemble Kalman filtering: Production properties from re-runs of the downscaled coarse scale ensemble at selected times, compared to the corresponding fine scale reference production properties (–·–). The inputs used are, from left to right; the initial fine scale ensemble, the downscaled ensemble after 300 days, 900 days, and 1,800 days of updating. The production properties considered are, from top to bottom; $opr1$, $gor1$, $bhp1$, $opr2$, $gor2$, $bhp2$ and $bhpi$.

downscaled versions of the coarse scale updates at selected time-steps as input to ω, compared to the fine scale reference production. Here also, the initial uncertainty in the production forecasts is quite large. As more and more production data are assimilated into the reservoir models, the uncertainty is reduced. At the latest times, the uncertainty is low, but we observe biases in the forecasts, especially for the gas-breakthrough time in production well number two, and for the time when the production controls switch from target rate to target bottom hole pressure in both production wells. Moreover, the variability in the ensemble is too small since we consider the coarse scale fluid flow simulator to be as reliable as the fine scale one. A modelling error should be added to the former. This shows that relying on traditional coarse scale ensemble Kalman filtering, without modelling the effects introduced by the coarse scale fluid flow simulator, leads to biases and underestimation of variance in the ensemble of production history conditioned reservoir state vectors. The production conditioning is accurate and precise on the coarse scale, but not on the fine scale.

4.2 Scale-corrected ensemble Kalman filtering

An initial ensemble of $n_c = 20$ reservoir state vectors is generated from the initial model $f(r_0 q_0 | r_0^o)$ at the reference scale. This ensemble is used to estimate the means and covariance matrices in expression (26) at each time-step. We refer to this as the calibration ensemble. In addition, we generate an ensemble of $n_s = 100$ reservoir state vectors from the initial model. These realisations are generated on the reference, fine scale, but fluid flow simulation is only performed on the upscaled, coarse reservoir representation. We refer to this as the simulation ensemble, and it is used to perform the simulation part of the scale-corrected ensemble Kalman filtering in expressions (31)-(41). The scale-corrected ensemble Kalman filtering is performed for 1,800 days, with updates every 30 days.

Note that considering only the fine scale results in the calibration process corresponds to performing traditional ensemble Kalman filtering on the fine scale, which is the optimal approach, with only $n = 20$ members in the ensemble. Recall that the reservoir state vector is of dimension 150,007. We want the ensemble to be representative of the pdf of this high-dimensional space, hence 20 realisations appear as a very small set. Even a set of 100 realisations, as we have in our coarse scale ensemble, appears as small, but this set can be expanded without dramatically increasing the computer demands. One may use an even coarser scale to afford larger ensembles.

We expect to recognise similar behaviours in the development of the fine scale and coarse scale calibration ensembles over time. Figure 11 shows the development over time for a vertical slice of the log-permeability in one of the calibration ensemble members, on both scales, compared to the corresponding vertical slices in the

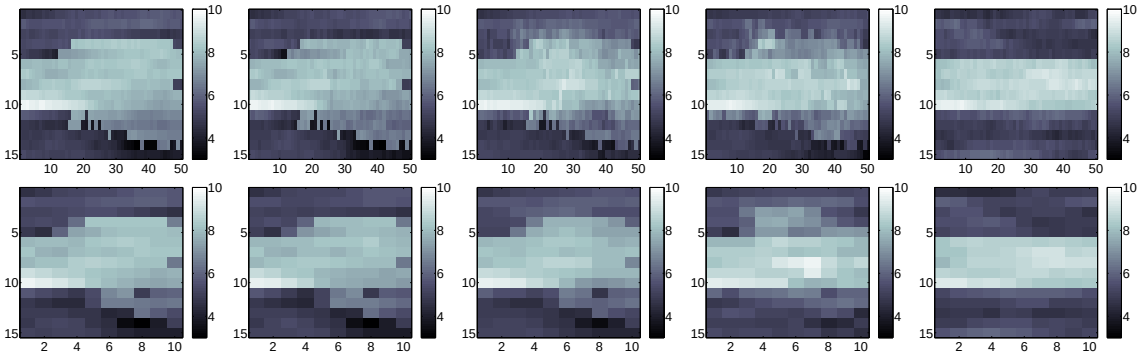

Figure 11: Scale-corrected ensemble Kalman filtering: Vertical slice of the development of the fine scale log-permeability in one of the calibration ensemble members. The first row shows the development on the fine scale, while the second row shows the development on the coarse scale. The columns represent, from left to right; the initial time-step, after 300 days, 900 days, and 1,800 days of updating, and the slice of the reference reservoir at the respective scale.

reference reservoir at the respective scales. We clearly see that the size and location of the high permeable region gets closer and closer to the reference reservoir as more and more production data are assimilated into the reservoir models. As expected, we recognise the same trends on both scales.

Figure 12 shows the development over time for the same vertical slice of the mean log-permeability in the calibration ensemble, also on both scales. Close to the injection well, to the left in the figures, the initial mean of the calibration ensemble is close to the true values of the reference reservoir because of the conditioning to the well logs in the initial model. Further away from the injection well, the interfaces between the high and low permeable areas are not as clearly defined. As more production data are assimilated into the reservoir models, the mean of the calibration ensemble gets more and more similar to the respective values in the reference reservoir.

Figure 13 shows the development over time for the estimated standard deviation for the log-permeability in the calibration ensemble at both scales. Close to the injection well, where the initial estimate of the mean log-permeability is in good accordance with the values in the reference reservoir, the initial estimate of the standard deviation is also low. In other areas, the initial estimate of the standard deviation is high, which is due to the uncertainty in the size and location of the high permeable region in the prior model. The initial estimate of the standard deviation in the coarse scale calibration ensemble is lower, due to the fact that upscaling is an averaging process. Very high and very low values on the fine scale are averaged out, making the the estimate of the standard deviation on the coarse scale lower. By conditioning the reservoir models to more and more production data, the size and location of the high permeable layer gets more certain, reflected in the reduction in

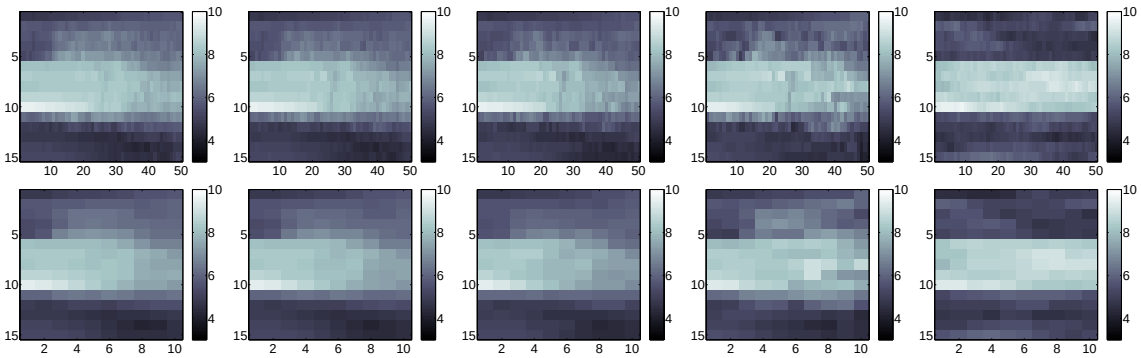

Figure 12: Scale-corrected ensemble Kalman filtering: Vertical slice of the development of the mean log-permeability in the calibration ensemble. The first row shows the development on the fine scale, while the second row shows the development on the coarse scale. The columns represent, from left to right; the initial time-step, after 300 days, 900 days, and 1,800 days of updating, and the slice of the reference reservoir at the respective scale.

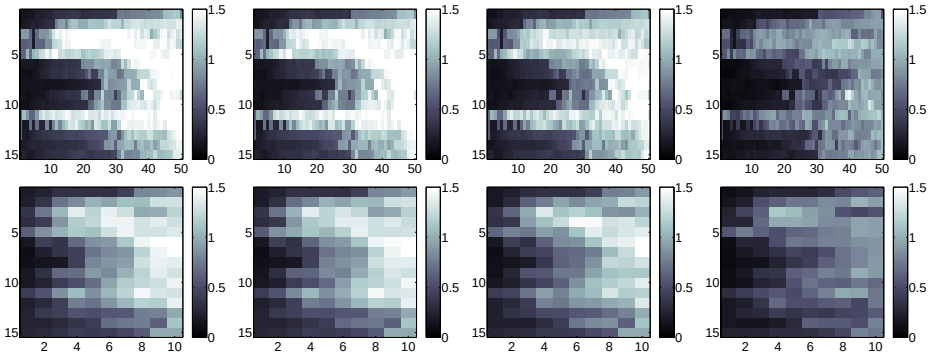

Figure 13: Scale-corrected ensemble Kalman filtering: Vertical slice of the development of the standard deviation of the log-permeability in the calibration ensemble. The first row shows the development on the fine scale, while the second row shows the development on the coarse scale. The columns represent, from left to right; the initial time-step, after 300 days, 900 days, and 1,800 days of updating.

the estimates of the standard deviation at later times.

Similar trends are seen for the development of the porosity in a single calibration ensemble member over time, and also for the mean and standard deviation estimates. This is not shown here.

Because of the possible inconsistencies between static and dynamic properties in the Kalman updates, we again use the corrected static properties at selected times and re-run the fluid flow simulations from the initial time-step using these values. Figure 14 shows the fluid flow simulation results for all production properties considered for the fine scale calibration ensemble when using the initial fine scale ensemble and the updated fine scale ensemble after 300 days, 900 days and 1,800 days of updates as input to ω, compared to the fine scale reference production. We see that by using the initial fine scale calibration ensemble we get large uncertainties in both gas-breakthrough times and the times where the production wells switch from target rate to target pressure control. When using the updated ensemble after 1,800 days as input to ω, we get fluid flow simulation results that almost exactly reproduce the production history up to 2,000 days for all the production properties considered. These 20 realisations are from the posterior pdf based on fine scale fluid flow simulation and are as representative as they can be. The problem is that the low number of realisations give low coverage in the space of reservoir state vectors, and that increasing the number require dramatic computer resources.

Figure 15 shows the corresponding results when using the coarse scale calibration ensemble members as input to ω^*, compared to the fine scale reference production. Initially, the uncertainty is quite large. As more and more production data are assimilated into the reservoir model, the uncertainty is reduced. At later times, we observe bias compared to the reference production. This observation means that in the calibration ensemble, the production history conditioning is accurate on the fine scale, but biased and less precise on the coarse scale. This fact is used to estimate the model error introduced by the coarse scale reservoir simulator. If we can find a reliable estimate, the model error can be used to simulate a large ensemble on the coarse scale, where the bias and precision are corrected for.

In the simulation ensemble we can predict the fine scale reservoir properties from the coarse scale ensemble members whenever needed. We do the prediction after 300 days, 900 days and 1,800 days of updating. We re-run the fluid flow simulations by using the static reservoir properties from the initial fine scale ensemble, and the static reservoir properties from the scale-corrected fine scale ensemble at the aforementioned time steps as input to the fine scale fluid flow simulator. Figure 16 shows the results obtained compared to the reference production. We see that the large initial uncertainties are reduced as more and more production data are assimilated. The predicted fine scale ensemble after 1,800 days of updating provides apparently unbiased forecasts of the production. The variability in the forecasts is, however, higher than in the production forecasts from the fine scale calibration

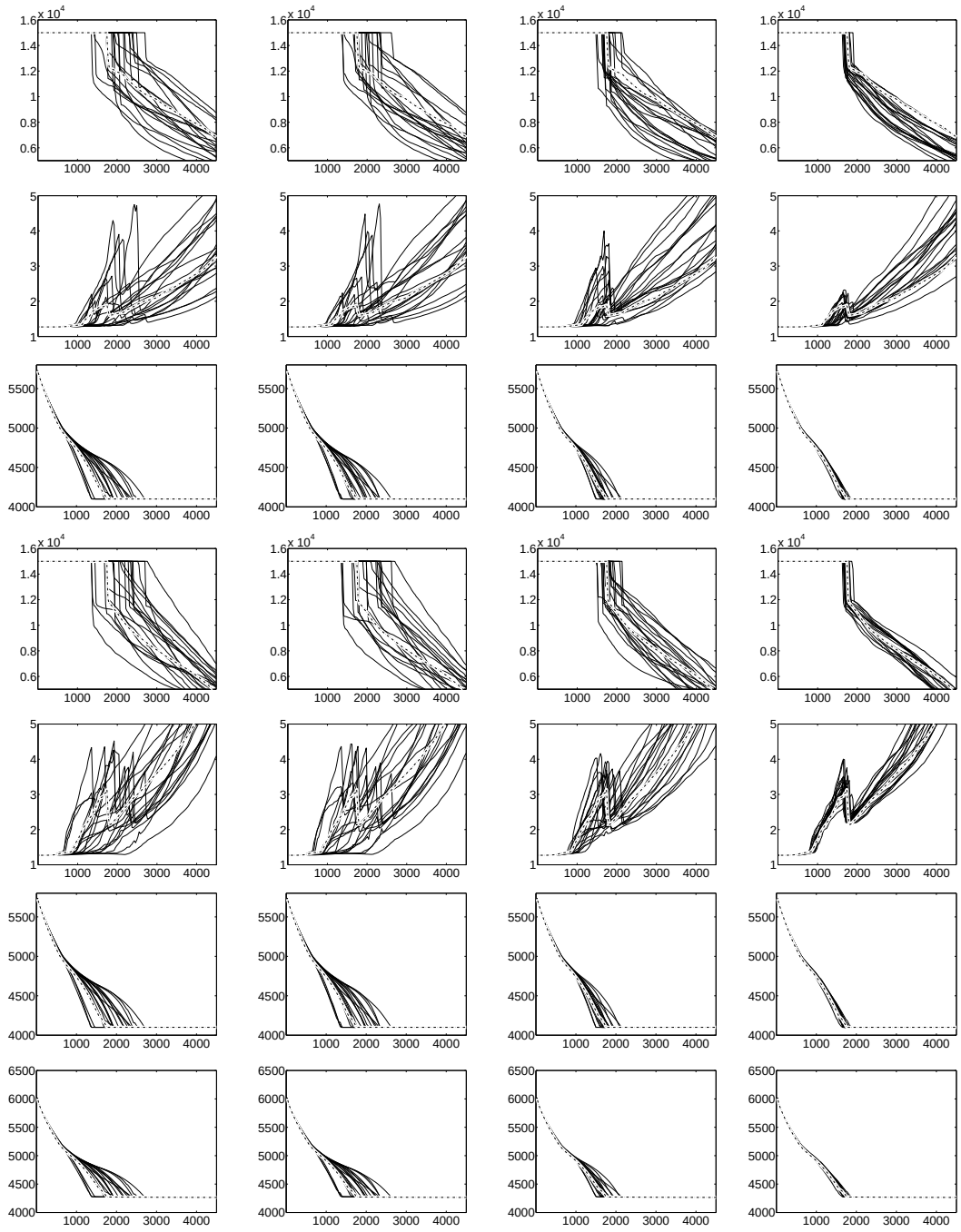

Figure 14: Scale-corrected ensemble Kalman filtering: Production properties from re-runs of the fine scale calibration ensemble at selected times, compared to the corresponding fine scale reference production properties (–·–). The inputs used are, from left to right; the initial fine scale calibration ensemble, the fine scale calibration ensemble after 300 days, 900 days, and 1,800 days of updating. The production properties considered are, from top to bottom; $opr1$, $opr2$, $gor1$, $gor2$, $bhp1$, $bhp2$ and $bhpi$.

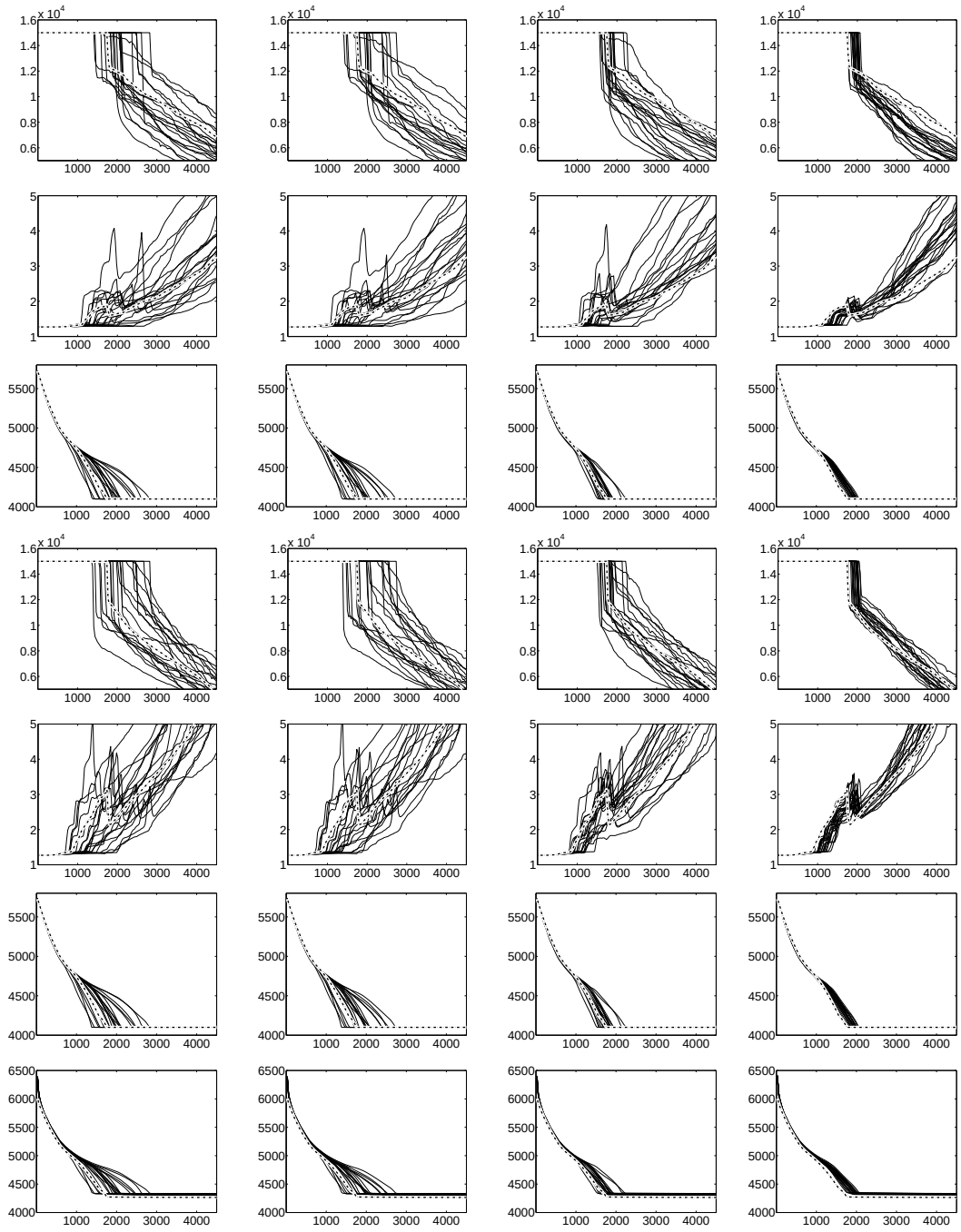

Figure 15: Scale-corrected ensemble Kalman filtering: Production properties from re-runs of the coarse scale calibration ensemble at selected times, compared to the corresponding fine scale reference production properties (–·–). The inputs used are, from left to right; the initial coarse scale calibration ensemble, the coarse scale calibration ensemble after 300 days, 900 days, and 1,800 days of updating. The production properties considered are, from top to bottom; $opr1$, $gor1$, $bhp1$, $opr2$, $gor2$, $bhp2$ and $bhpi$.

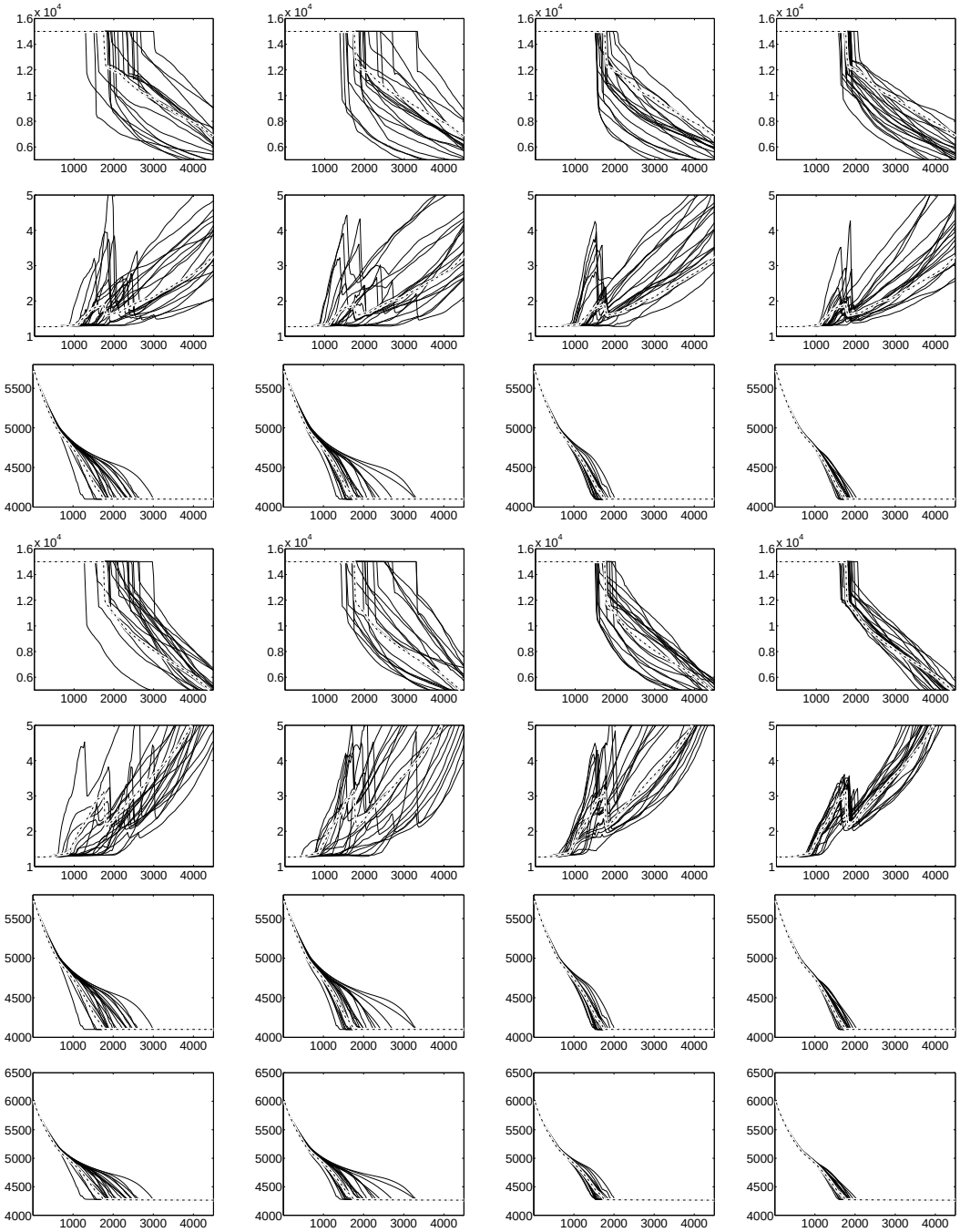

Figure 16: Scale-corrected ensemble Kalman filtering: Production properties from re-runs of the predicted fine scale simulation ensemble at selected times, compared to the reference production properties (–·–). The inputs used are, from left to right; the initial fine scale simulation ensemble, and the predicted simulation ensemble after 300 days, 900 days, and 1,800 days of updating. The production properties considered are, from top to bottom; $opr1$, $gor1$, $bhp1$, $opr2$, $gor2$, $bhp2$ and $bhpi$.

ensemble, shown in Figure 14. This loss of precision is the cost of using the coarse scale fluid flow simulator.

In Figure 17 the results in this section are summarised. The figure shows a column-wise comparison between the production properties obtained from conditioning to static data only, from the fine scale calibration ensemble, from the scale-corrected ensemble, and from traditional coarse scale ensemble Kalman filtering. The three latter after 1,800 days of updating. Note that all these fluid flow simulations are performed on the fine scale. By comparing production forecasts with no production history conditioning, presented in the first column, with the production history conditioned ones in the three other columns, we easily see that conditioning on the production history has considerable impact. Recall that the results from the fine scale calibration ensemble, presented in the second column, are the optimal ones, but the problem is low coverage of the high dimensional posterior pdf for the reservoir properties due to very large computational requirements when increasing the ensemble size. The results from the scale-corrected fine scale simulation ensemble, presented in the third column, are apparently unbiased, but compared to the fine scale calibration ensemble they appear with more variability. This loss of precision represents the cost of using the coarse scale fluid flow simulator. Increasing the size of the simulation ensemble has a low computational cost, however, since all fluid flow simulations are performed on the coarse scale. Hence, we can generate a high number of realisations of the fine scale reservoir state vector in the scale-corrected ensemble, providing much better coverage of the high dimensional posterior pdf for the reservoir properties. In the fourth column, we observe the bias in fine scale reservoir representations associated with traditional coarse scale ensemble Kalman filtering.

In the simulation examples above we perform approximately 200 coarse scale fluid flow simulations in the same amount of time it takes to perform a single fine scale fluid flow simulation. With this ratio, we can perform scale-corrected ensemble Kalman filtering with calibration ensemble of size $n_c = 20$ and simulation ensemble of size $n_s \approx 5,000$ in the same amount of time it would take to perform ensemble Kalman filtering on the fine scale with $n = 50$ ensemble members. An ensemble with 5,000 members will considerably improve the representation of the high dimensional posterior pdf for the reservoir properties, compared to an ensemble of only 50 members.

5 Conclusions

When considering very large reservoir evaluation problems, the reservoir properties are usually represented on a very-fine scale, which may contain up to 10^9 grid blocks. The conditioning on well logs and seismic data can be performed on this very-fine scale, but even a single fluid flow simulation will in most cases be computationally

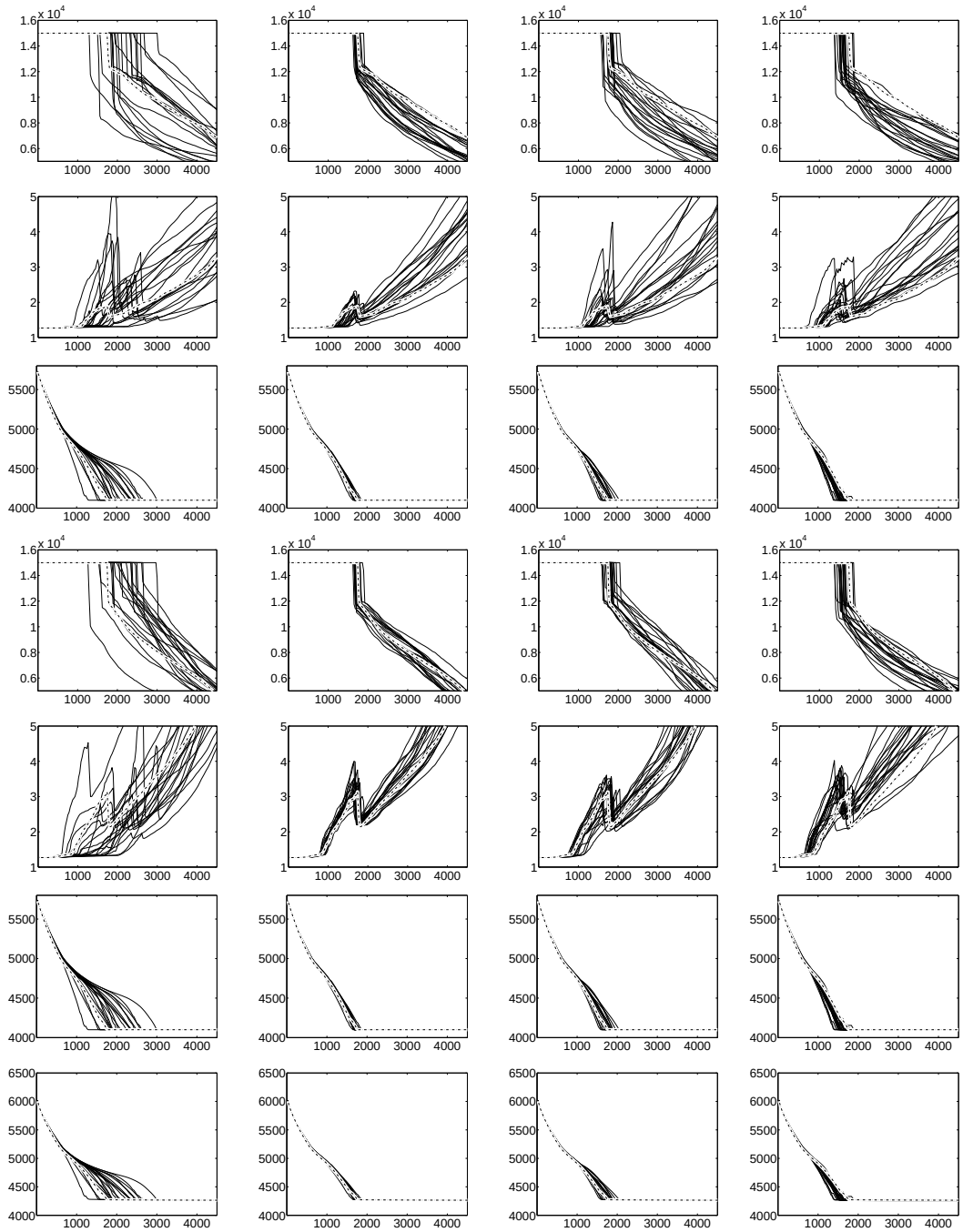

Figure 17: Comparison of results: The columns, from left to right, represents production properties obtained from fine scale fluid flow simulations: conditioning on static data only, the fine scale calibration ensemble, the scale-corrected ensemble, and traditional coarse scale ensemble Kalman filtering. The three latter after 1,800 days of updating. The production properties considered are, from top to bottom; $opr1$, $gor1$, $bhp1$, $opr2$, $gor2$, $bhp2$ and $bhpi$.

prohibitive. In order to allow for fluid flow simulation, the reservoir properties must be represented on a fine scale, which may contain 10^6 grid blocks. The fine scale representation of the reservoir properties is found from the very-fine scale representation, usually through upscaling. At this fine scale, we can afford to run a low number of fluid flow simulations. For production history conditioning problems, requiring a large number of repeated flow simulations, however, the representation of the reservoir properties must be further coarsened. The coarse scale representation will usually contain around 10^4 grid blocks.

The traditional ensemble Kalman filtering operates on the coarse scale. As shown in the simulation studies, downscaling the reservoir properties from the coarse scale to the fine scale after the final update step provides a wrongly centred ensemble with too little variability. Scale-corrected ensemble Kalman filtering, however, provides an apparently unbiased representation of the reservoir properties on the fine scale, with realistic variability in the ensemble. Contrary to traditional ensemble Kalman filtering, the scale-corrected ensemble Kalman filtering accounts for the loss in reliability coming from further coarsening the representation of the reservoir properties from the fine scale.

Our proposed procedure to solving very large reservoir evaluation problems involving production history conditioning can be summarised in the following steps:

- Start with a very-fine scale representation of the reservoir properties (10^9 cells), conditioned on seismic data and well logs.
- Upscale the reservoir properties to a fine scale representation (10^6 cells), where 10-20 fluid flow simulations can be afforded. Term this scale the reference scale for production history conditioning.
- Perform scale-corrected ensemble Kalman filtering on a coarse scale representation (10^4 cells), such that 1000-2000 fluid flow simulations can be afforded.

The scale-corrected ensemble of size 1000-2000 on the fine scale representation is now correctly centred and contains realistic variability accounting for the approximations made by performing fluid flow simulations on the coarse scale. Note that the uncertainty associated with further downscaling to the very-fine scale representation of the reservoir properties is not addressed in this study. Assessing this uncertainty should be a topic for further research.

Acknowledgements

The study is partly supported by the Uncertainty in Reservoir Evaluation (URE) program at the Norwegian University of Science and Technology.

References

Durlofsky, L. (2003). Upscaling of geocellular models for reservoir flow simulation: A review of recent progress. presented at 7th International Forum on Reservoir Simulation, Bühl/Baden-Baden, Germany, June 23-27, 2003.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics, *Journal of Geophysical Research* **99**(C5): 10,143–10,162.

Evensen, G. (2003). The ensemble kalman filter: theoretical formulation and practical implementation, *Ocean Dynamics* **53**(4): 343–367.

Farmer, C. (2000). Upscaling: A review, *International Journal for Numerical Methods in Fluids* **40**: 63–78.

GeoQuest (2004). *ECLIPSE Reference Manual 2004A*, Schlumberger GeoQuest.

Gu, Y. and Oliver, D. S. (2004). History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter. Presented at the Society of Petroleum Engineers Annual Technical Conference and Exhibition, Houston, Texas, USA, 26–29 September 2004, SPE 89942.

Hegstad, B. and Omre, H. (2001). Uncertainty in production forecasts based on well observations, seismic data, and production history, *Society of Petroleum Engineers Journal* pp. 409–424. SPE 74699.

Houtekamer, P. and Mitchell, H. (1998). A sequential ensemble kalman filter for atmospheric data assimilation, *Monthly Weather Review* **126**(3): 796–811.

Johnson, R. and Wichern, D. (1998). *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey.

Lødøen, O., Omre, H., Durlofsky, L. and Chen, Y. (2005). Assessment of uncertainty in reservoir production forecasts using upscaled flow models, *Geostatistics 2004, Banff*. Proceedings of the Seventh International Geostatistical Congress held in Banff, Canada, in September 2004.

Nævdal, G., Johnsen, L., Aanonsen, S. and Vefring, E. (2003). Reservoir monitoring and continuous model updating using ensemble Kalman filter. Presented at the Society of Petroleum Engineers Annual Technical Conference and Exhibition, Denver, Colorado, USA, 5–8 October 2003, SPE 84372.

Nævdal, G., Mannseth, H. and Vefring, E. (2002). Near-well reservoir monitoring through ensemble kalman filter. Presented at the SPE/DOE Thirteenth Symposium of Improved Oil Recovery, Tulsa, Oklahoma, USA, 13–17 April 2002, SPE 75235.

Omre, H. (2000). Stochastic reservoir models conditioned to non-linear production history observations, *in* W. Kleingeld and D. Krige (eds), *Geostatistics 2000, Cape Town*, pp. 166–175. Proceedings of the Sixth International Geostatistical Congress held in Cape Town, South Africa, in April 2000.

Omre, H. and Lødøen, O. (2004). Improved production forecasts and history matching using approximate fluid-flow simulators, *Society of Petroleum Engineers Journal* **9**(3): 339–351. SPE 74691.

Reichle, R., McLaughlin, D. and Entekhabi, D. (2002). Hydrologic data assimilation with the ensemble kalman filter, *Monthly Weather Review* **130**(1): 103–113.

Wen, X.-H. and Chen, W. (2005). Real-time reservoir model updating using the ensemble Kalman filter. Presented at the 18th SPE Reservoir Simulation Symposium, The Woodlands, Texas, 31 January – 2 February, 2005, SPE 92991.

Appenix A - Downscaling Procedure

Let $R_i^* \in \Re^m$, $i \in \Omega = \{1, \ldots, n\}$, be Gaussian random variables,

$$R_i^* \text{ iid } f(r^*) \sim Gauss(\mu_R^*, \Sigma_{R^*R^*}),$$

where μ_R^* and $\Sigma_{R^*R^*}$ are unknown, but can be estimated from R_i^*. Since the estimates are based on n observations, the covariance matrix estimate $\Sigma_{R^*R^*}$ has a maximum rank of $n-1$.

Let $(e_{R^*,1}, \lambda_{R^*,1}), (e_{R^*,2}, \lambda_{R^*,2}), \ldots, (e_{R^*,n}, \lambda_{R^*,n})$ be the eigenvector-eigenvalue pairs of $\Sigma_{R^*R^*}$, where $\lambda_{R^*,1} \geq \lambda_{R^*,2} \geq \cdots \geq \lambda_{R^*,n} \geq 0$. If $\Sigma_{R^*R^*}$ has rank $n-1$, only the $n-1$ first eigenvalues will be larger than zero. Let $A_{R^*} \in \Re^{n \times n}$ be a matrix containing the eigenvectors $e_{R^*,1}, \ldots, e_{R^*,n}$ in rows $1, \ldots, n$. Since the eigenvectors define an orthonormal basis, $A_{R^*} A_{R^*}{}' = A_{R^*}{}' A_{R^*} = I$.

Let Y_i^*, $i \in \Omega$, be transformations of R_i^*, $i \in \Omega$, using the orthonormal basis A_{R^*},

$$Y_i^* = A_{R^*} R_i^* \in \Re^n, i \in \Omega.$$

It can be shown (see e.g. Johnson and Wichern (1998)) that

$$(\Sigma_{Y^*Y^*})_{ij} = e'_{R^*,i} \Sigma_{R^*R^*} e_{R^*,j} = \begin{cases} \lambda_{R^*,i} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Hence, $\Sigma_{Y^*Y^*}$ will be a diagonal matrix with the eigenvalues $\lambda_{R^*,i}$, $i \in \Omega$, along the diagonal. The covariance structure of R_i^*, $i \in \Omega$, can now be expressed as

$$\Sigma_{R^*R^*} = A'^*_R \Sigma_{Y^*Y^*} A_{R^*} = A'_{R^*} \begin{bmatrix} \Sigma_{Z^*Z^*} & 0 \\ 0 & 0 \end{bmatrix} A_{R^*} = A'_{Y^*} \Sigma_{Z^*Z^*} A_{Y^*},$$

where $A_{Y^*} \in \Re^{(n-1)\times(n-1)}$ contains the eigenvectors $e_{R^*,1}, \ldots, e_{R^*,n-1}$ as rows, $Z_i^* = A_{Y^*}R_i^*$, $i \in \Omega$, and $\Sigma_{Z^*Z^*}$ contains $\lambda_{R^*,1}, \ldots, \lambda_{R^*,n-1}$ along the diagonal.

This means that there will be as much information about the covariance structure between the m elements in R_i^*, $i \in \Omega$, in $n-1$ linear combinations, $Z_i^* = A_{Y^*}R_i^*$, of the original m elements.

Further, let $R_i \in \Re^n$, $i \in \Omega$ be Gaussian random variables

$$R_i \text{ iid } f(r) \sim Gauss(\mu_R, \Sigma_{RR}),$$

where μ_R and Σ_{RR} are unknown.

Consider the Gaussian random variable $R|R^*$. Since the covariance estimate $\Sigma_{R^*R^*}$ does not have full rank, conditioning on the linear combinations $Z^* = A_{Y^*}R^*$ is the same as conditioning on R^*. Hence, $R|R^* = R|Z^*$. The probability distribution $f(r|z^*)$ is fully described by the moments

$$\mu_{R|Z^*} = \mu_R + \Sigma_{RZ^*}\Sigma_{Z^*Z^*}^{-1}(Z^* - \mu_{Z^*})$$

and

$$\Sigma_{R|Z^*} = \Sigma_{RR} - \Sigma_{RZ^*}\Sigma_{Z^*Z^*}^{-1}\Sigma_{Z^*R}. \tag{43}$$

Let u and v be zero-mean Gaussian random variables, each with covariance matrix I. If expression (43) can be expressed as

$$\Sigma_{R|Z^*} = AA' - BB',$$

realisations of $R|Z^*$ can be generated through

$$\mu_{R|Z^*} + \varepsilon_{R|Z^*} = \mu_{R|Z^*} + Au - Bv,$$

if the right covariance Σ_{uv} between u and v can be found. To get the right covariance Σ_{uv}, the requirement is that

$$Var(Au - Bv) = AIA' - BIB' - 2A\Sigma_{uv}B' = \Sigma_{R|Z^*} = AA' - BB'.$$

Solving the equation above, yields

$$\Sigma_{uv} = (A'A)^{-1}A'B.$$

For $A'A$ to be invertible, the requirement is that $A \in \Re^{n\times(n-1)}$, since A has rank $n-1$. The matrix A can be found by looking at how Σ_{RR} is estimated. This is done through

$$\Sigma_{RR} = \frac{\sum_{i=1}^n (R_i - \bar{R})(R_i - \bar{R})'}{n-1} = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{bmatrix} \begin{bmatrix} \gamma_1' \\ \gamma_2' \\ \vdots \\ \gamma_n' \end{bmatrix},$$

where $\gamma_i = \frac{R_i - \bar{R}}{\sqrt{n-1}}$.

Let $W = [w_1\ w_2\ \cdots\ w_n]'$ be an orthonormal basis, where the last row vector is defined by $w_n = [1/\sqrt{n}\ \ 1/\sqrt{n}\ \ \cdots\ \ 1/\sqrt{n}]$. Since the rows of W are orthonormal, $WW' = I$. Eq. (44) can be rewritten using W in the following way

$$\Sigma_{RR} = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{bmatrix} \begin{bmatrix} \gamma_1' \\ \gamma_2' \\ \vdots \\ \gamma_n' \end{bmatrix} = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{bmatrix} WW' \begin{bmatrix} \gamma_1' \\ \gamma_2' \\ \vdots \\ \gamma_n' \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_n' \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} \end{bmatrix} \begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_{n-1}' \end{bmatrix} = AA',$$

since

$$\alpha_n = \sum_{i=1}^{n} \frac{\gamma_i}{\sqrt{n}} = \frac{\sum_{i=1}^{n}(R_i - \bar{R})}{\sqrt{n}\sqrt{n-1}} = 0$$

Since $\Sigma_{Z^*Z^*}$ is a diagonal matrix, B can be found directly from expression (43) as

$$B = \Sigma_{RZ^*}\Sigma_{Z^*Z^*}^{-\frac{1}{2}}.$$