Computers & Geosciences 55 (2013) 3–15

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Ensemble smoother with multiple data assimilation

Alexandre A. Emerick*, Albert C. Reynolds

University of Tulsa, 800 South Tucker Drive, Tulsa, OK 74104, United States

ARTICLE INFO

Article history:
Received 24 November 2011
Received in revised form
14 February 2012
Accepted 11 March 2012
Available online 18 March 2012

Keywords:
History matching
Ensemble smoother
Ensemble Kalman filter
Multiple data assimilation

ABSTRACT

In the last decade, ensemble-based methods have been widely investigated and applied for data assimilation of flow problems associated with atmospheric physics and petroleum reservoir history matching. This paper focuses entirely on the reservoir history-matching problem. Among the ensemble-based methods, the ensemble Kalman filter (EnKF) is the most popular for history-matching applications. However, the recurrent simulation restarts required in the EnKF sequential data assimilation process may prevent the use of EnKF when the objective is to incorporate the history matching in an integrated geo-modeling workflow. In this situation, the ensemble smoother (ES) is a viable alternative. However, because ES computes a single global update, it may not result in acceptable data matches; therefore, the development of efficient iterative forms of ES is highly desirable. In this paper, we propose to assimilate the same data multiple times with an inflated measurement error covariance matrix in order to improve the results obtained by ES. This method is motivated by the equivalence between single and multiple data assimilation for the linear-Gaussian case. We test the proposed method for three synthetic reservoir history-matching problems. Our results show that the proposed method provides better data matches than those obtained with standard ES and EnKF, with a computational cost comparable with the computational cost of EnKF.

1. Introduction

Since its introduction by Evensen (1994), the number of publications about EnKF has become extensive. EnKF has been applied in diverse research fields including, oceanography (Bertino et al., 2003; Keppenne and Rienecker, 2003), numerical weather prediction (Houtekamer and Mitchell, 2005; Szunyogh et al., 2005), hydrology (Reichle et al., 2002; Chen and Zhang, 2006; Liu et al., 2008) and petroleum reservoir history matching (Aanonsen et al., 2009; Oliver and Chen, 2010). Evensen (2007) presents a chronological list of applications of EnKF. The first reservoir application of EnKF was presented by Nævdal et al. (2002), where EnKF was applied to update permeability fields for near-well reservoir models. After this pioneer application, the interest and frequency of use of EnKF as a history-matching technique increased significantly. Some recent field applications of EnKF for history matching can be found in Skjervheim et al. (2007), Bianco et al. (2007), Evensen et al. (2007), Haugen et al. (2008), and Emerick and Reynolds (2011b). Two recent review papers (Aanonsen et al., 2009; Oliver and Chen, 2010) summarize the main developments and applications of EnKF in reservoir history-matching problems from 2001 to early 2010.

The sequential data assimilation with EnKF requires modification of the traditional history-matching problem from a parameter-estimation problem to a parameter-state-estimation problem. Specifically, when applying EnKF for history matching, we update a combined parameter-state vector, which includes the reservoir model parameters (uncertain reservoir rock properties) and the primary variables of the reservoir simulator (typically gridblock pressure, fluid saturations and dissolved gas–oil ratio in a standard black-oil reservoir simulator). The reason for including primary variables, which represent the state of the dynamical system, is to avoid the need to run the simulations from time zero every data assimilation time-step. However, it requires an additional assumption of statistical consistency between the updated vectors of model parameters and states, where "statistical consistency" is defined in Thulin et al. (2007). However, because the reservoir simulator equations are highly nonlinear, this assumption is often violated, which may result in substandard estimates of model parameters; see, e.g., Wang et al. (2010).

Sequential data assimilation is an attractive feature of EnKF when the objective is the closed-loop reservoir management (Jansen et al., 2005, 2009; Wang et al., 2009; Peters et al., 2010; Chen and Oliver, 2010; Chen et al., 2010). However, the simulation restarts required by the sequential data assimilation are inconvenient when the goal is to incorporate the history matching in workflows that integrate different parts of the reservoir modeling process, e.g., combining seismic, structural and geological modeling

* Corresponding author. Tel.: +1 918 691 9138.
E-mail addresses: aemerick@gmail.com (A.A. Emerick), reynolds@utulsa.edu (A.C. Reynolds).

0098-3004/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2012.03.011

with flow simulation; see, e.g., Zachariassen et al. (2011). These workflows typically require integrating different geo-modeling softwares and may include upscaling of the rock properties, which makes the simulation restarts required by EnKF very inconvenient, if not impossible.

The ensemble smoother (ES) was introduced by van Leeuwen and Evensen (1996). Unlike EnKF, ES does not assimilate data sequentially in time. Instead, ES computes a global update by simultaneously assimilating all data available. Other than that ES formulation is similar to EnKF. van Leeuwen and Evensen (1996) found that EnKF performed better than ES when applied to an ocean circulation model. Evensen and van Leeuwen (2000) compared ES and EnKF with Lorenz equations and concluded that EnKF outperforms ES because the recursive updates in the EnKF keep the ensemble of states "on track" and closer to the true solution. Recently, Skjervheim et al. (2011) compared ES and EnKF and concluded that both methods gave similar results for the reservoir history-matching problems considered in their paper. The major advantage of ES is that it avoids restarts of the reservoir simulator. This makes ES much faster and easier to implement than EnKF. For example, Skjervheim et al. (2011) reported that ES was 10 times faster than EnKF in the their reservoir history-matching applications. The elimination of simulation restarts also makes ES an attractive option for the previously mentioned geo-modeling workflows.

In fact, when we apply ES for reservoir history matching, we convert the parameter-state-estimation problem back to a parameter-estimation problem. Thus, ES removes the parameter-state inconsistency issue observed in the sequential data assimilation with EnKF. However, Reynolds et al. (2006) showed that EnKF is similar to applying, at each data assimilation time-step, one Gauss–Newton iteration with a full step and the sensitivity matrix replaced by an average sensitivity matrix estimated from the ensemble. Based on this fact, we conjecture that the process of assimilating data which are closely spaced in time approximately corresponds to applying several Gauss–Newton corrections to the state vector. If true, this conjecture would partially explain the reasonable performance of EnKF when applied to history matching production data. With ES, on the other hand, all data are assimilated simultaneously, which means that a single Gauss–Newton correction is applied to condition the ensemble to all data available. Hence, ES may not be able to provide acceptable data matches when applied to reservoir history-matching problems.

Inspired by the work of Rommelse (2009), in Emerick and Reynolds (2012), we introduced a method in which the same data are assimilated multiple times with the covariance of the measurement errors multiplied by the number of data assimilations. In Emerick and Reynolds (2012), we showed that single and multiple data assimilation (MDA) are equivalent for the linear-Gaussian case and presented computational evidence that MDA can improve EnKF estimates and data matches when assimilating time-lapse seismic data. Here, we extend the ideas presented in Emerick and Reynolds (2012) to history match production data with ES. In our previous paper, we considered only the case in which the covariance of the measurement errors was multiplied by the number of data assimilations. Here, we generalize this procedure by presenting the condition that the multiplication coefficients of the data covariance matrix must satisfy in order to guarantee the equivalence between single and multiple data assimilation for the linear-Gaussian case.

As pointed out by one of the reviewers, Oliver and Chen (2009) also considered an idea similar to the one from Rommelse (2009). In Oliver and Chen (2009), the authors present the assimilation of "pseudo-observation" as a procedure to improve the initial ensemble. Given an observation $d_{\mathrm{obs},1}$ with variance σ_d^2, Oliver and Chen (2009) considered the possibility of assimilating $d_{\mathrm{obs},1}$ with a measurement error variance $\sigma_{d,1}^2$ together with a pseudo-observation $d_{\mathrm{obs},2}$ with a measurement error variance $\sigma_{d,2}^2$. For the linear case, they worked out the variances that should be assigned to $d_{\mathrm{obs},1}$ and $d_{\mathrm{obs},2}$ to obtain the same posterior mean and variance that should be obtained by assimilating only $d_{\mathrm{obs},1}$ with the original measurement error variance σ_d^2. They focused on conditioning the initial ensemble to the pseudo-observation prior to applying EnKF to assimilate data and did not pursue that idea of assimilating the actual observation multiple times.

In a paper which appeared after our manuscript was originally submitted, Chen and Oliver (2012) proposed using the ensemble randomized maximum likelihood (EnRML) method (Gu and Oliver, 2007) as an iterative ensemble smoother. They compared EnKF, ES, sequential-EnRML and batch-EnRML for a small five-spot problem and the Brugge field case (Peters et al., 2010). They concluded that ES often requires iterations to achieve satisfactory data matches. For the Brugge case, the proposed iterative ensemble smoother (batch-EnRML) required 26 iterations to achieve an average data mismatch objective function which was slightly higher that the one obtained by the standard EnKF. Each iteration of batch-EnRML requires running reservoir simulations for the whole ensemble at least once. However, each iteration may require more than one ensemble run because the line search may require cutting the step-size and then running the whole ensemble again. Therefore, the computational cost of 26 iterations of batch-EnRML is definitely much higher than the cost of EnKF.

This paper is organized as follows: in the next three sections, we present EnKF, ES and EnRML equations. Next, we briefly review the equivalence between single and multiple data assimilation for the linear-Gaussian case and generalize the MDA method. Then, we explain the proposed ES-MDA procedure. After that we present three synthetic reservoir problems. The first problem is a 2D two-phase flow model. For this problem, we repeat the data assimilation with 10 different initial ensembles to obtain more statistically meaningful results and compare the performance of ES-MDA with EnKF, standard ES and batch-EnRML. The second problem is a reservoir with a single producing well where the observations correspond to layer-rates at a specific time, mimicking production-logging data. The third problem pertains to the Brugge field. The last section of the paper presents our conclusions.

2. Ensemble Kalman filter

In the EnKF, we define the N_y-dimensional state vector at the nth data assimilation time-step, y^n, as

$$y^n = \begin{bmatrix} m^n \\ p^n \end{bmatrix}, \tag{1}$$

where m^n is the N_m-dimensional vector of model parameters, and p^n is the N_p-dimensional vector representing the state of the dynamical system. For reservoir applications, p^n contains the primary variables of the reservoir simulator (typically gridblock pressures, saturations and dissolved gas–oil ratio). The superscript n indicates the nth data assimilation time-step. The EnKF analysis equation can be written as

$$y_j^{n,a} = y_j^{n,f} + C_{\mathrm{YD}}^{n,f}(C_{\mathrm{DD}}^{n,f} + C_{\mathrm{D}}^{n})^{-1}(d_{\mathrm{uc},j}^{n} - d_j^{n,f}) \tag{2}$$

for $j = 1, 2, \ldots, N_e$, with N_e denoting the number of ensemble members.

In the above equation, $C_{\mathrm{YD}}^{n,f}$ is the $N_y \times N_n$ cross-covariance matrix between the forecast state vector and predicted data; $C_{\mathrm{DD}}^{n,f}$ is the $N_n \times N_n$ auto-covariance matrix of predicted data; C_{D}^{n} is the $N_n \times N_n$ covariance matrix of observed data measurement errors;

N_n is the number of data points assimilated at the nth data assimilation time-step; d_{uc}^n is the vector of "perturbed" observations, i.e., $d_{\mathrm{uc}}^n \sim \mathcal{N}(d_{\mathrm{obs}}^n, C_{\mathrm{D}}^n)$ with d_{obs}^n denoting the N_n-dimensional vector of observed data at the nth data assimilation time-step; $d^{n,f}$ is the corresponding N_n-dimensional vector of predicted data. The superscripts a and f denote analysis and forecast, respectively.

3. Ensemble smoother

Unlike oceanic and atmospheric models, which present chaotic and unstable dynamics (Evensen, 2007, Chapter 6) reservoir simulation models are typically stable functions of the rock property fields. If we also neglect model uncertainty, which is a common assumption in reservoir history-matching problems, we only need to consider the parameter-estimation problem when applying ES. In this case, we write the analyzed vector of model parameters, m^a, as

$$m_j^a = m_j^f + C_{\mathrm{MD}}^f (C_{\mathrm{DD}}^f + C_{\mathrm{D}})^{-1} (d_{\mathrm{uc},j} - d_j^f) \tag{3}$$

for $j = 1, 2, \ldots, N_e$. The notation is similar to the one presented for EnKF. C_{MD}^f is the cross-covariance matrix between the prior vector of model parameters, m^f, and the vector of predicted data, d^f; C_{DD}^f is the $N_d \times N_d$ auto-covariance matrix of predicted data. Note that we use N_d to denote the total number of measurements assimilated, which is different from our previous notation for EnKF, where we use N_n to denote the number of measurements at the nth data assimilation time-step. Similar to the EnKF case, $d_{\mathrm{uc}} \sim \mathcal{N}(d_{\mathrm{obs}}, C_{\mathrm{D}})$ with d_{obs} denoting the N_d-dimensional vector of observed data and C_{D} denoting the $N_d \times N_d$ covariance matrix of observed data measurement errors.

4. Ensemble randomized maximum likelihood

EnRML was introduced by Gu and Oliver (2007) as an iterative EnKF. In a recent paper, Chen and Oliver (2012) propose using EnRML as an iterative ensemble smoother. They call this procedure batch-EnRML. EnRML uses a Gauss–Newton update equation, i.e.

$$m_j^{\ell+1} = \beta_\ell m_j^f + (1-\beta_\ell) m_j^\ell - \beta_\ell C_{\mathrm{M}}^f \overline{G}_\ell^{\mathrm{T}} (\overline{G}_\ell C_{\mathrm{M}}^f \overline{G}_\ell^{\mathrm{T}} + C_{\mathrm{D}})^{-1} \times [d_j^\ell - d_{\mathrm{uc},j} - \overline{G}_\ell (m_j^\ell - m_j^f)] \tag{4}$$

for $j = 1, 2, \ldots, N_e$. In the above equation, ℓ denotes the iteration index and β_ℓ denotes the step-size. During the iterative process, C_{M}^f is fixed and estimated based on the forecast ensemble. $\overline{G}$ is the average sensitivity matrix, which is estimated based on the ensemble by "solving"

$$\Delta D^\ell = \overline{G}_\ell \Delta M^\ell \tag{5}$$

for $\overline{G}_\ell$ using singular value decomposition (SVD). In Eq. (5), $\Delta D^\ell = D^\ell - \overline{D}^\ell$, where D^ℓ is the matrix with the ensemble of predicted data at the ℓth iteration, i.e., the jth column of D^ℓ corresponds to the predicted data from the jth ensemble member. $\overline{D}^\ell$ is the matrix with all columns equal to $\overline{d}^\ell$, which represents the average of all columns of D^ℓ. Similarly, $\Delta M^\ell = M^\ell - \overline{M}^\ell$, where the jth column of M^ℓ contains the vector of model parameters corresponding to the jth ensemble member at the ℓth iteration, and $\overline{M}^\ell$ is the matrix with all columns equal to $\overline{m}^\ell$, which represents the average of all columns of M^ℓ.

At each iteration of EnRML, if the updated ensemble obtained with Eq. (4) does not result in a decrease in the average data mismatch objective function, it is necessary to perform a line search. Gu and Oliver (2007) and Chen and Oliver (2012) do not provide details about the line-search procedure they use. Here, we use the following batch-EnRML implementation:

1. Run the ensemble from time zero until the end of the history.
2. Initialize: $\ell = 0$, $N_{\mathrm{cuts}} = 0$, β_0 and $m_j^0 = m_j^f$ for $j = 1, 2, \ldots, N_e$.
3. Compute $\overline{G}_\ell$ using Eq. (5).
4. For $j = 1$ to N_e:
   (a) Compute $m_j^{\ell+1}$ using Eq. (4).
   (b) Rerun the ensemble from time zero.
   (c) Compute $O_{d,j}^{\ell+1} = \frac{1}{2}(d_j^{\ell+1} - d_{\mathrm{uc},j})^{\mathrm{T}} C_{\mathrm{D}}^{-1} (d_j^{\ell+1} - d_{\mathrm{uc},j})$.
   end (for).
5. Compute $\overline{O}_d^{\ell+1} = (1/N_e) \sum_{j=1}^{N_e} O_{d,j}^{\ell+1}$.
6. If $\overline{O}_d^{\ell+1} < \overline{O}_d^\ell$ then:
   (a) Accept the step and increase the step-size for the next iteration, $\beta_{\ell+1} = 2\beta_\ell$.
   (b) If $\beta_{\ell+1} > \beta_0$, then set $\beta_{\ell+1} = \beta_0$.
   (c) Set $N_{\mathrm{cuts}} = 0$.
   (d) Increase the iteration index, $\ell = \ell + 1$.
   Else:
   (a) Reduce the step size, $\beta_\ell = \beta_\ell / 2$.
   (b) Set $N_{\mathrm{cuts}} = N_{\mathrm{cuts}} + 1$.
   (c) If $N_{\mathrm{cuts}} \leq 5$, then return to step 4.
   end (if).
7. Check termination criteria.
8. If any one of the termination criteria is satisfied, then stop the data assimilation. Otherwise, return to step 3.

We use the following termination criteria:

- $\max |m_{i,j}^{\ell+1} - m_{i,j}^\ell| < 10^{-3}$ for $i = 1, 2, \ldots, N_m$ and $j = 1, 2, \ldots, N_e$.
- $|(\overline{O}_d^{\ell+1} - \overline{O}_d^\ell)/\overline{O}_d^\ell| < 10^{-2}$.
- Maximum number of iterations = 10.
- Maximum number of consecutive cuts in the step-size = 5.

5. Multiple data assimilation for the linear-Gaussian case

In Emerick and Reynolds (2012), we proved the equivalence between single and multiple data assimilation for the linear-Gaussian case using two procedures. In the first procedure, we started with the Kalman filter equations and showed that the updated mean and covariance obtained with MDA corresponds to the same posterior mean and covariance obtained assimilating data only once with the actual covariance of the measurement errors, C_{D}. In the second procedure, we started with a sample of the prior probability density function (pdf) and showed that after data assimilation with MDA, the analyzed state vector is a sample of the correct posterior pdf. Although these two proofs are formally equivalent, the second proof explicitly shows that we also need to perturb the observations based on the inflated C_{D} to obtain the correct posterior pdf for the linear-Gaussian case. In this section, we generalize the second proof. Specifically, instead of assuming that C_{D} is multiplied by the number of data assimilations, here, we consider a more general case, in which C_{D} is increased by a different coefficient, α_i, each time we assimilate data. Then, we derive the condition for the coefficients α_i's required to guarantee the equivalence between single and multiple data assimilation. For the sake of completeness, we repeat some of the steps of the proof presented in Emerick and Reynolds (2012).

The linear problem refers to the case in which the relation between the vector of predicted data, d, and the vector of model parameters, m, is expressed in the form

$$d = Gm, \tag{6}$$

where G is the $N_d \times N_m$ sensitivity matrix. To assimilate data N_a times, we first define

$$\tilde{d}_{\text{obs}} \equiv \begin{bmatrix} d_{\text{obs}} \\ \vdots \\ d_{\text{obs}} \end{bmatrix}, \tag{7}$$

$$\tilde{G} \equiv \begin{bmatrix} G \\ \vdots \\ G \end{bmatrix} \tag{8}$$

and

$$\tilde{C}_{\text{D}} \equiv \begin{bmatrix} \alpha_1 C_{\text{D}} & 0 & \cdots & 0 \\ 0 & \alpha_2 C_{\text{D}} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \alpha_{N_a} C_{\text{D}} \end{bmatrix}. \tag{9}$$

In the above definitions, we simply repeated the vector d_{obs} and the matrices G and C_{D} N_a times. Hence, for the MDA case, the linear relation (Eq. (6)) becomes

$$\tilde{d} = \tilde{G} m. \tag{10}$$

Because we focus on ES, in this paper, we write the equations in terms of the vector of model parameters, m. This is different from the notation used in Emerick and Reynolds (2012), where the equations were presented for a combined parameter-state vector y.

For sampling the posterior pdf with MDA, we start with a sample from the prior pdf, m^f, i.e., $m^f \sim \mathcal{N}(m_{\text{pr}}, C_{\text{M}})$, where m_{pr} is the prior mean and C_{M} is the prior model covariance matrix. In addition, we define the modified vector of perturbed observations as

$$\tilde{d}_{\text{uc}} \equiv \begin{bmatrix} d_{\text{uc}}^1 \\ \vdots \\ d_{\text{uc}}^{N_a} \end{bmatrix}, \tag{11}$$

where $d_{\text{uc}}^i \sim \mathcal{N}(d_{\text{obs}}, \alpha_i C_{\text{D}})$, for $i = 1, 2, \ldots, N_a$. It is important to note that the superscript i for d_{uc}^i does not refer to time; i simply refers to the ith assimilation of the same data.

Following the procedure presented in Emerick and Reynolds (2012), we define the random vector $\tilde{m}^a$ as the minimizer of the RML objective function (Reynolds et al., 1999) modified for the MDA case, i.e.

$$\tilde{m}^a = \arg\min_m \tilde{O}(m), \tag{12}$$

where

$$\tilde{O}(m) = \tfrac{1}{2}(m - m^f)^{\text{T}} C_{\text{M}}^{-1}(m - m^f) + \tfrac{1}{2}(\tilde{G}m - \tilde{d}_{\text{uc}})^{\text{T}} \tilde{C}_{\text{D}}^{-1}(\tilde{G}m - \tilde{d}_{\text{uc}}). \tag{13}$$

In Eq. (13), equivalent definitions of the random variables, $\tilde{d}_{\text{uc}}$ and m^f, respectively, are given by

$$\tilde{d}_{\text{uc}} = \tilde{d}_{\text{obs}} + \tilde{C}_{\text{D}}^{1/2} \tilde{z}_d, \tag{14}$$

where $\tilde{z}_d \sim \mathcal{N}(0, I_{N_a N_d})$ and

$$m^f = m_{\text{pr}} + C_{\text{M}}^{1/2} z_m, \tag{15}$$

where $z_m \sim \mathcal{N}(0, I_{N_m})$. Note that, as in Reynolds et al. (1999), the model $\tilde{m}^a$ that minimizes Eq. (13) is also a random variable.

Requiring the gradient of $\tilde{O}(m)$ to vanish, we can solve the resulting expression for m and denote the result as $\tilde{m}^a$. This procedure leads to

$$\tilde{m}^a = C_{\text{MAP}}\{C_{\text{M}}^{-1}(m_{\text{pr}} + C_{\text{M}}^{1/2} z_m) + \tilde{G}^{\text{T}} \tilde{C}_{\text{D}}^{-1}(\tilde{d}_{\text{obs}} + \tilde{C}_{\text{D}}^{1/2} \tilde{z}_d)\}. \tag{16}$$

The steps required to obtain Eq. (16) are presented in Emerick and Reynolds (2012). In Eq. (16), C_{MAP} is the correct posterior covariance matrix, which is given by Tarantola (2005, p. 66)

$$C_{\text{MAP}} = (C_{\text{M}}^{-1} + G^{\text{T}} C_{\text{D}}^{-1} G)^{-1} = C_{\text{M}} - C_{\text{M}} G^{\text{T}} (C_{\text{D}} + G C_{\text{M}} G^{\text{T}})^{-1} G C_{\text{M}}. \tag{17}$$

In order to prove the equivalence between single and multiple data assimilation, it is necessary to show that $\tilde{m}^a$ obtained by Eq. (16) is a sample of the correct pdf, i.e., that $\tilde{m}^a \sim \mathcal{N}(m_{\text{MAP}}, C_{\text{MAP}})$, where m_{MAP} denotes the maximum a posteriori (MAP) estimate, which is given by Tarantola (2005, p. 66)

$$\begin{aligned} m_{\text{MAP}} &= C_{\text{MAP}}(C_{\text{M}}^{-1} m_{\text{pr}} + G^{\text{T}} C_{\text{D}}^{-1} d_{\text{obs}}) \\ &= m_{\text{pr}} + C_{\text{M}} G^{\text{T}} (C_{\text{D}} + G C_{\text{M}} G^{\text{T}})^{-1} (d_{\text{obs}} - G m_{\text{pr}}). \end{aligned} \tag{18}$$

Because the posterior pdf is Gaussian, $\mathcal{N}(m_{\text{MAP}}, C_{\text{MAP}})$, we only need to show that

$$E[\tilde{m}^a] = m_{\text{MAP}} \tag{19}$$

and

$$\text{cov}[\tilde{m}^a] = E[(\tilde{m}^a - m_{\text{MAP}})(\tilde{m}^a - m_{\text{MAP}})^{\text{T}}] = C_{\text{MAP}}. \tag{20}$$

Writing the expression for $E[\tilde{m}^a]$, we obtain

$$\begin{aligned} E[\tilde{m}^a] &= C_{\text{MAP}}\{C_{\text{M}}^{-1}(m_{\text{pr}} + C_{\text{M}}^{1/2} E[z_m]) + \tilde{G}^{\text{T}} \tilde{C}_{\text{D}}^{-1}(\tilde{d}_{\text{obs}} + \tilde{C}_{\text{D}}^{1/2} E[\tilde{z}_d])\} \\ &= C_{\text{MAP}}\{C_{\text{M}}^{-1} m_{\text{pr}} + \tilde{G}^{\text{T}} \tilde{C}_{\text{D}}^{-1} \tilde{d}_{\text{obs}}\} \\ &= C_{\text{MAP}}\left\{ C_{\text{M}}^{-1} m_{\text{pr}} + [G^{\text{T}} \ \cdots \ G^{\text{T}}] \begin{bmatrix} \frac{1}{\alpha_1} C_{\text{D}}^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\alpha_{N_a}} C_{\text{D}}^{-1} \end{bmatrix} \begin{bmatrix} d_{\text{obs}} \\ \vdots \\ d_{\text{obs}} \end{bmatrix} \right\} \\ &= C_{\text{MAP}}\left\{ C_{\text{M}}^{-1} m_{\text{pr}} + [G^{\text{T}} \ \cdots \ G^{\text{T}}] \begin{bmatrix} \frac{1}{\alpha_1} C_{\text{D}}^{-1} d_{\text{obs}} \\ \vdots \\ \frac{1}{\alpha_{N_a}} C_{\text{D}}^{-1} d_{\text{obs}} \end{bmatrix} \right\} \\ &= C_{\text{MAP}}\left\{ C_{\text{M}}^{-1} m_{\text{pr}} + \left(\sum_{i=1}^{N_a} \frac{1}{\alpha_i} \right) G^{\text{T}} C_{\text{D}}^{-1} d_{\text{obs}} \right\}. \end{aligned} \tag{21}$$

The last equality of Eq. (21) is equivalent to Eq. (18) if, and only if

$$\sum_{i=1}^{N_a} \frac{1}{\alpha_i} = 1. \tag{22}$$

Note that the choice $\alpha_i = N_a$ for $i = 1, \ldots, N_a$ satisfies the condition of Eq. (22), but there are infinitely many other possible choices for the α_i's. By choosing a set of α_i's that satisfies Eq. (22), MDA yields the correct posterior mean for the linear-Gaussian case, i.e.

$$E[\tilde{m}^a] = m_{\text{MAP}}. \tag{23}$$

Following the same procedure presented in Emerick and Reynolds (2012), it can be shown that

$$\text{cov}[\tilde{m}^a] = C_{\text{MAP}}(C_{\text{M}}^{-1} + \tilde{G}^{\text{T}} \tilde{C}_{\text{D}}^{-1} \tilde{G}) C_{\text{MAP}}. \tag{24}$$

Eq. (24) corresponds to Eq. (40) of Emerick and Reynolds (2012) and it was obtained by using the fact that the vectors z_m and $\tilde{z}_d$ are independent realizations of $\mathcal{N}(0, I_{N_m})$ and $\mathcal{N}(0, I_{N_a N_d})$, respectively. From Eqs. (24), (22), (17), it follows that

$$\begin{aligned} \text{cov}[\tilde{m}^a] &= C_{\text{MAP}}\left\{ C_{\text{M}}^{-1} + [G^{\text{T}} \ \cdots \ G^{\text{T}}] \begin{bmatrix} \frac{1}{\alpha_1} C_{\text{D}}^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\alpha_{N_a}} C_{\text{D}}^{-1} \end{bmatrix} \begin{bmatrix} G \\ \vdots \\ G \end{bmatrix} \right\} C_{\text{MAP}} \\ &= C_{\text{MAP}}\left\{ C_{\text{M}}^{-1} + \left(\sum_{i=1}^{N_a} \frac{1}{\alpha_i} \right) G^{\text{T}} C_{\text{D}}^{-1} G \right\} C_{\text{MAP}} \\ &= C_{\text{MAP}}\{C_{\text{MAP}}^{-1}\} C_{\text{MAP}} \\ &= C_{\text{MAP}}, \end{aligned} \tag{25}$$

which completes the proof. This proof assumes the correct full-rank prior covariance matrix, C_{M}, which will be the case for ES (or EnKF) only when the size of the ensemble goes to infinity.

6. ES-MDA

The proof presented in the previous section applies for the linear-Gaussian case. For the nonlinear case, this equivalence does not hold. The motivation for using MDA in the nonlinear case comes from the fact that ES is equivalent to a single Gauss–Newton iteration with a full step and an average sensitivity estimated from the prior ensemble (Reynolds et al., 2006). In this sense, MDA can be interpreted as an "iterative" ensemble smoother (with a predefined number of iterations), where instead of a single and potentially large Gauss–Newton correction, we perform multiple smaller corrections in the ensemble. Note that inflating the covariance matrix associated with the measurement errors to damp the changes in the model at early iterations of Newton-like methods is not new in the reservoir history-matching literature; see, e.g., Wu et al. (1999) and Gao and Reynolds (2006).

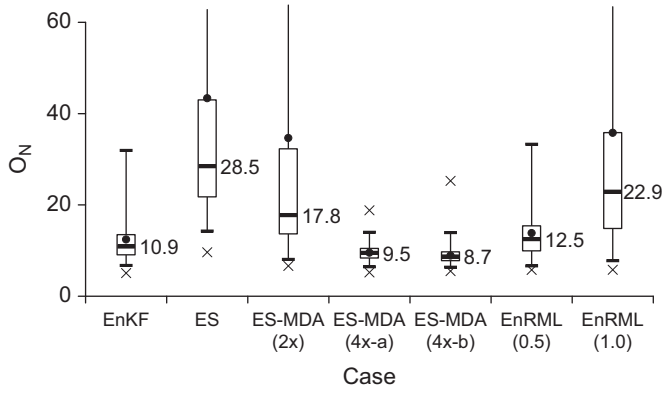

Fig. 1. Box plots of the normalized objective function based on 10 ensembles: Example 1.

The ES-MDA algorithm follows:

1. Choose the number of data assimilations, N_a, and the coefficients α_i for $i=1,\ldots,N_a$.
2. For $i=1$ to N_a:
   (a) Run the ensemble from time zero.
   (b) For each ensemble member, perturb the observation vector using $d_{uc}=d_{obs}+\sqrt{\alpha_i}C_D^{1/2}z_d$, where $z_d\sim\mathcal{N}(0,I_{N_d})$.
   (c) Update the ensemble using Eq. (3) with C_D replaced by $\alpha_i C_D$.

As noted in Eq. (3), ES requires the inversion of the $N_d\times N_d$ matrix C given by

$$C=C_{DD}^{f}+C_D. \tag{26}$$

Because C_{DD}^{f} is a real-symmetric positive semi-definite matrix, C given by Eq. (26) will be real-symmetric positive-definite as long as we choose C_D positive-definite. However, C may be poorly conditioned (Evensen, 2007, Chapter 14). Hence, EnKF and ES implementations typically use a pseudo-inverse of C computed using a truncated singular value decomposition (TSVD). For reservoir applications, because C is typically constructed based on data with different magnitudes, e.g., water-cut and pressure data, C may be poorly scaled and we may lose the information necessary to match data when truncating small singular values; see, e.g., Wang et al. (2010). Therefore, in our implementation, we rescale the components of the matrix C before calculating the

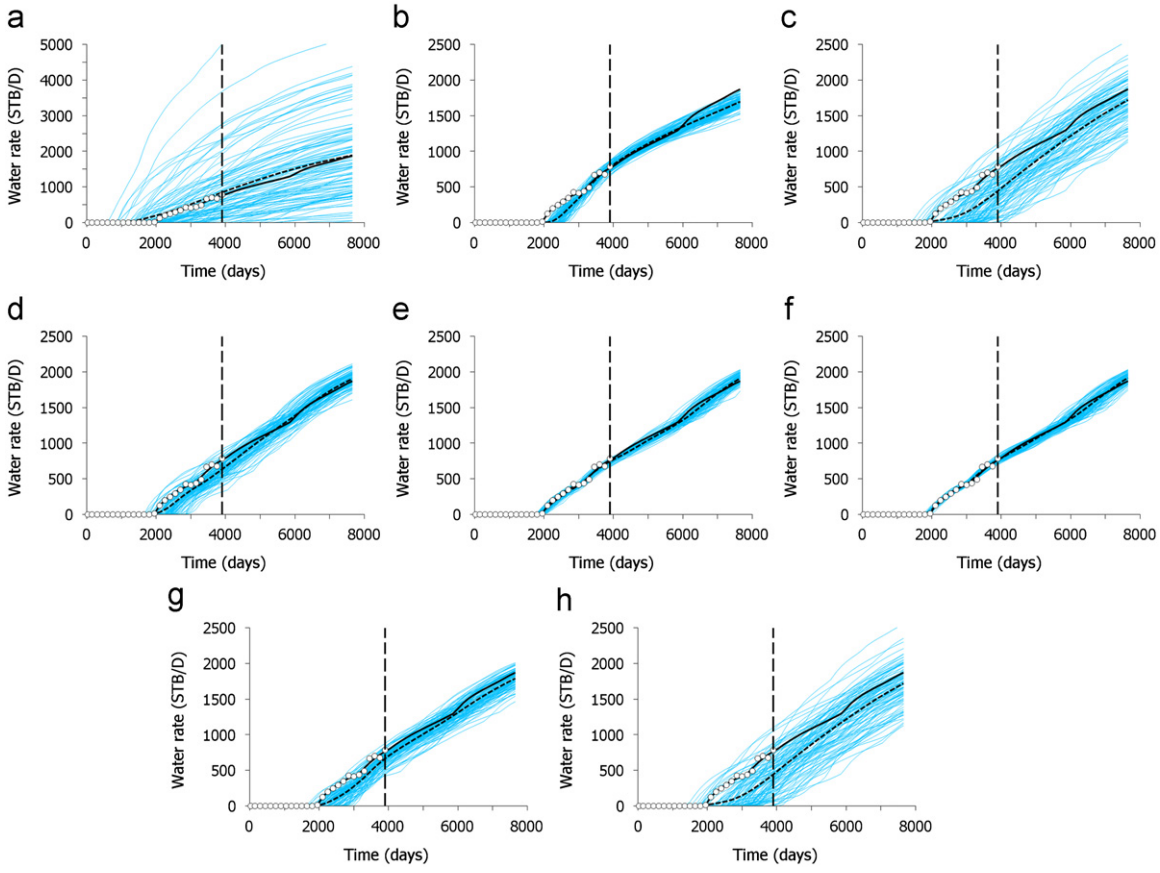

Fig. 2. Field water production rate for the first ensemble: Example 1. The vertical dashed line indicates the end of the historical period. White dots are the history; solid curve is the prediction from the true model; dashed curve is the ensemble mean prediction. The remaining curves are the predictions from the ensemble members. (a) Prior, (b) EnKF, (c) ES, (d) ES-MDA (2 ×), (e) ES-MDA (4 × -a), (f) ES-MDA (4 × -b), (g) batch-EnRML ($\beta_0=0.5$), and (h) batch-EnRML ($\beta_0=1$).

TSVD. For ES, the number of measurements assimilated at the analysis step can be very high, in which case, the TSVD of C becomes computationally expensive. In this situation, the subspace inversion introduced by Evensen (2004) is an attractive alternative. In our implementation of ES, if the number of measurements is no greater than the size of the ensemble, we compute the pseudo-inverse of C using TSVD. If the number of measurements is larger than the size of the ensemble, we use the subspace inversion. In both cases, we use rescaling as described in Appendix A of Emerick and Reynolds (2012). For the three examples presented in this paper, we kept 99.9% of the total energy of the singular values when applying SVD in both, pseudo-inverse or subspace inversion procedures.

As noted in Emerick and Reynolds (2012), we proved the equivalence between single and multiple data assimilation for the linear case by assimilating data N_a times simultaneously. However, in the MDA procedure presented above, we assimilate data N_a times consecutively and, before each data assimilation, we rerun the ensemble. For the linear-Gaussian case, these two approaches are equivalent. For the nonlinear case, the reruns effectively serve to update the "average sensitivity" before the next data assimilation. Also note that in the ES-MDA algorithm, every time we repeat the data assimilation, we resample the vector of perturbed observations, i.e., we recompute $d_{uc} \sim \mathcal{N}(d_{obs}, \alpha_i C_D)$. This procedure tends to reduce sampling problems caused by matching outliers that may be generated when perturbing the observations.

One potential difficultly with the proposed MDA procedure is that N_a and the coefficients α_i's need to be selected prior to the data assimilation. The simplest choice for α is $\alpha_i = N_a$ for all i. However, intuitively we expect that choosing α_i in a decreasing order can improve the performance of the method. In this case, we start assimilating data with a large value of α, which reduces the magnitude of the initial updates; then, we gradually decrease α.

Even though we focus on ES, conceptually, this procedure can also be applied to EnKF as described in Emerick and Reynolds (2012). Nevertheless, the moderate set of tests we have done indicate that MDA gives only small improvements when applied to EnKF, which makes ES-MDA a much more effective option in terms of computational cost than EnKF-MDA. This is consistent with our conjecture that when successive data have similar information content, the sequential data assimilation in EnKF is somewhat similar to accumulating several Gauss–Newton iterations.

7. Example 1: waterflooding

The first example is a two-phase (oil and water) synthetic reservoir model on a 2D uniform grid with 60×60 gridblocks. The dimensions of the gridblocks are 150 ft $\times$ 150 ft $\times$ 25 ft. The model parameters are gridblock log-permeabilities, $\ln(k)$. The true permeability field was generated from an anisotropic exponential correlation function with a major correlation length of 3750 ft

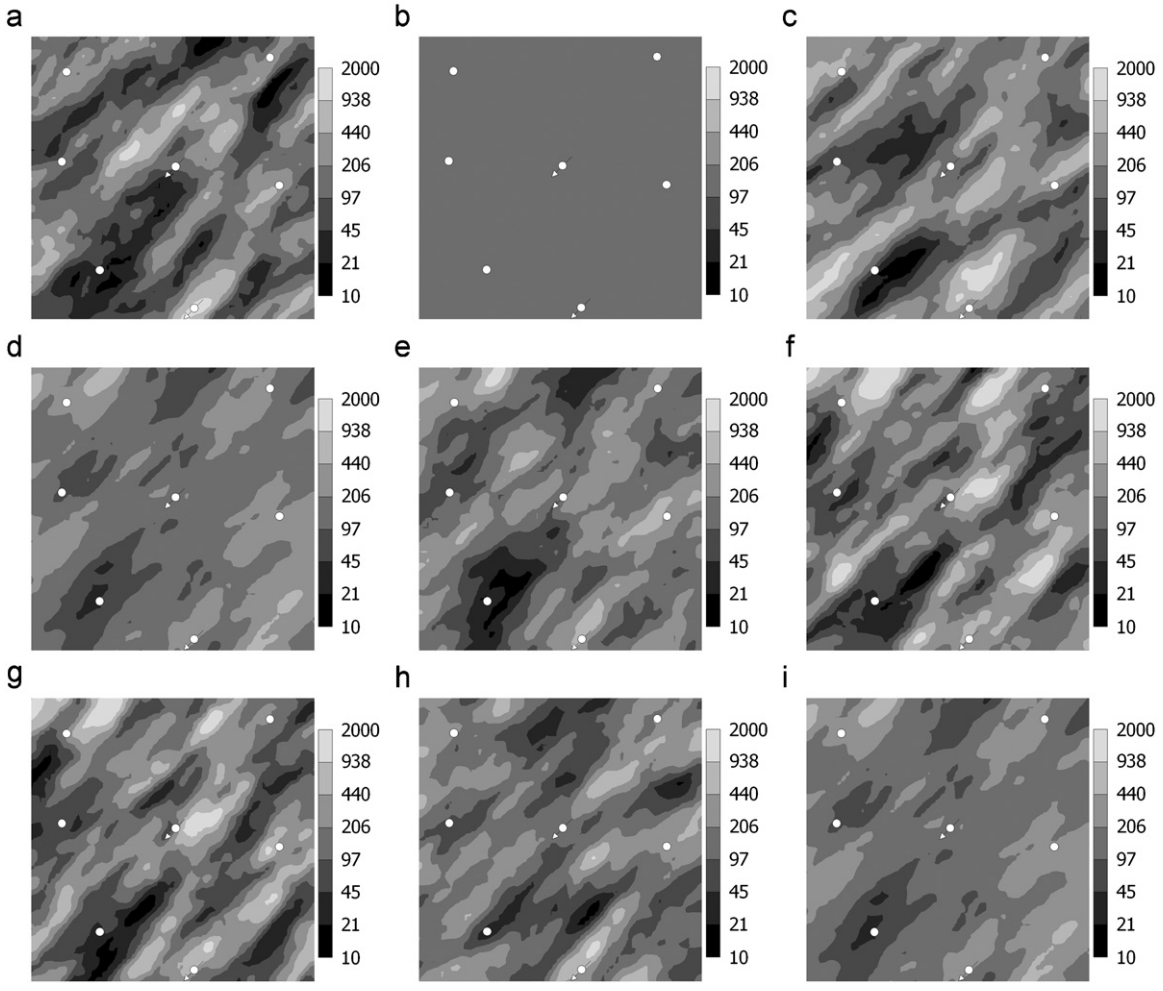

Fig. 3. Mean permeability fields (mD) for the first ensemble: Example 1. (a) True, (b) prior, (c) EnKF, (d) ES, (e) ES-MDA (2 $\times$), (f) ES-MDA (4 $\times$ -a), (g) ES-MDA (4 $\times$ -b), (h) batch-EnRML ($\beta_0 = 0.5$), and (i) batch-EnRML ($\beta_0 = 1$).

(which corresponds to the width of 25 gridblocks) and a minor correlation length of 1050 ft (i.e., seven gridblocks) oriented at 45°. The prior mean of ln(k) is 5.0 and the prior variance is 1.0 for all gridblocks. In this model, there are five producing wells and two water injectors. All producing wells operate at a fixed bottomhole pressure of 1000 psi, and the injectors operate at a fixed bottomhole pressure of 3000 psi. The observed data correspond to oil and water production rate and water injection rate every 150 days for a period of 3900 days. The observations were obtained by adding random noise to the data generated by the true model. The noise level was selected as 5% of the data.

Using the same prior mean and covariance function of the true model, we generated 10 ensembles of 100 models each. For each ensemble, we performed data assimilation using EnKF, ES, batch-EnRML and ES-MDA. For batch-EnRML, we tried two step-sizes, $\beta_0 = 0.5$ and $\beta_0 = 1$. For ES-MDA, we considered the following cases:

- ES-MDA (2 ×): ES assimilating data twice ($\alpha_1 = \alpha_2 = 2.0$).
- ES-MDA (4 × -a): ES assimilating data four times ($\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 4.0$).
- ES-MDA (4 × -b): ES assimilating data four times ($\alpha_1 = 9.333$, $\alpha_2 = 7.0$, $\alpha_3 = 4.0$ and $\alpha_4 = 2.0$).

Fig. 1 presents the box plot of the normalized objective function (O_N) obtained from each case. These box plots were computed based on combining the results of the 10 different initial ensembles. In the box plots, the horizontal line within each box corresponds to the median, the bottom and top of each box correspond to the percentiles P25 and P75. The ending points of the box plots correspond to the percentiles P2 and P98. The dots correspond to the mean and the crosses correspond to the minimum and maximum. The numbers next to the boxes correspond to the value of the median. The normalized objective function was computed using

$$O_N = \frac{O(m)}{N_d}, \tag{27}$$

where

$$O(m) = \tfrac{1}{2}(m - m_{\mathrm{pr}})^{\mathrm{T}} C_{\mathrm{M}}^{-1}(m - m_{\mathrm{pr}}) + \tfrac{1}{2}(d - d_{\mathrm{obs}})^{\mathrm{T}} C_{\mathrm{D}}^{-1}(d - d_{\mathrm{obs}}). \tag{28}$$

According to the results in Fig. 1, ES gave relatively high values of O_N compared to EnKF. This fact supports our conjecture that

Table 1
Average computational costs: Example 1.

Method	Number of equivalent simulation runs
EnKF	458
ES	101
ES-MDA (2 ×)	203
ES-MDA (4 ×)	405
Batch-EnRML ($\beta_0 = 0.5$)	1360
Batch-EnRML ($\beta_0 = 1$)	1270

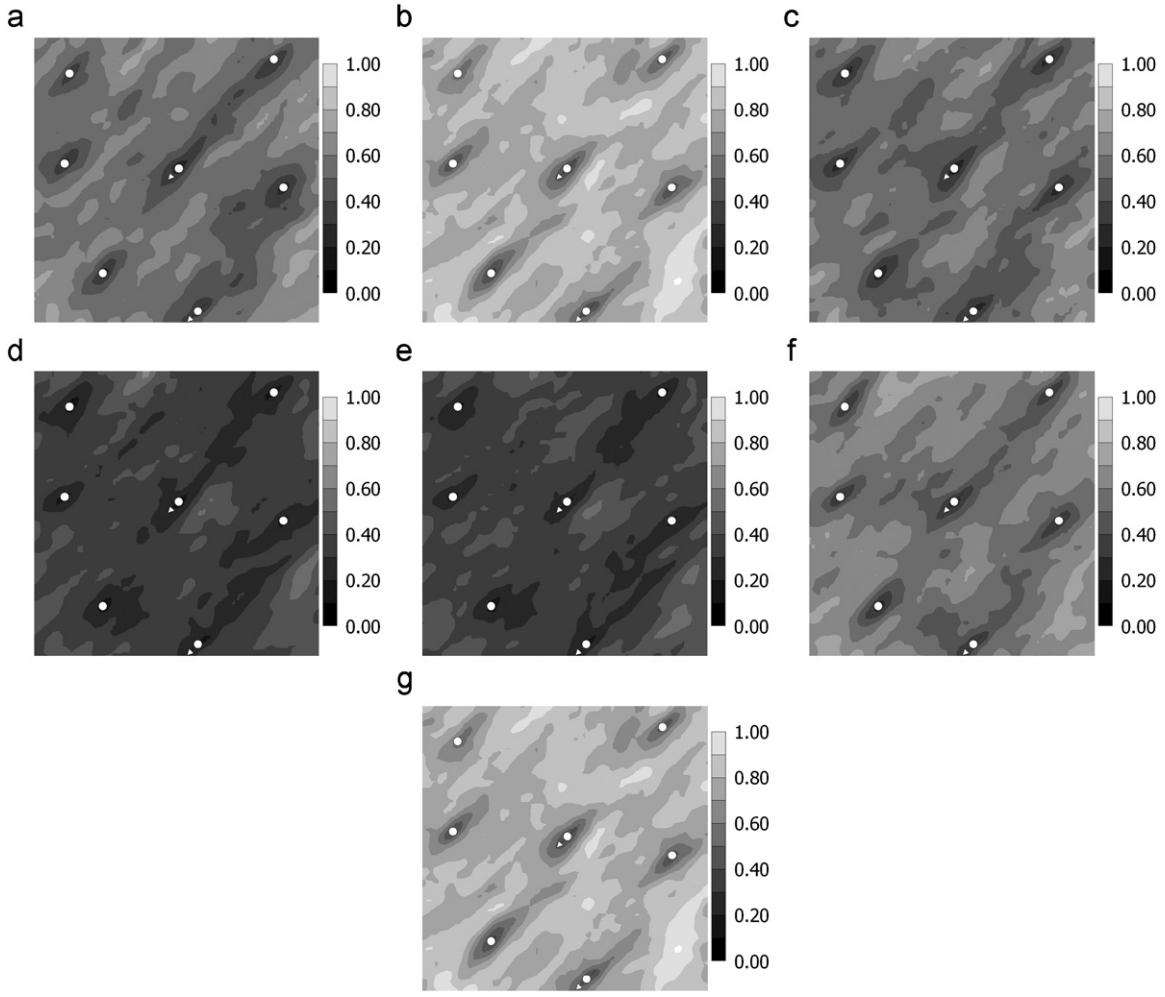

Fig. 4. Standard deviation of log-permeability for the first ensemble after data assimilation: Example 1. (a) EnKF, (b) ES, (c) ES-MDA (2 ×), (d) ES-MDA (4 × -a), (e) ES-MDA (4 × -b), (f) batch-EnRML ($\beta_0 = 0.5$), and (g) batch-EnRML ($\beta_0 = 1$).

EnKF performs reasonably well when history matching production data because the sequential assimilation approximately corresponds to several Gauss–Newton corrections. Increasing the number of data assimilations improved the data matches and, consequently, reduced the values of O_N obtained by ES. After four data assimilations, the values of O_N are slightly lower than the ones obtained by EnKF. Selecting the coefficients α_i's in a decreasing order resulted in further reduction in the values of O_N, although the differences are not significant. We tried other combinations of the coefficients α_i's, but no significant differences were obtained. The final values of O_N obtained by batch-EnRML were lower than the ones obtained by ES, but higher than EnKF and ES-MDA ($4\times$). The step-size $\beta_0 = 0.5$ resulted in better data matches than $\beta_0 = 1$. In fact, for 4 of the 10 initial ensembles, batch-EnRML with $\beta_0 = 1$ obtained results similar to standard ES. For these ensembles, the iterative process stopped at the second iteration because of five consecutive cuts in the step-size without reduction in the average data mismatch objective function. For $\beta_0 = 0.5$, on the other hand, batch-EnRML was able to improve the results of ES for the 10 ensembles. These results illustrate one problematic aspect of EnRML, namely, the performance of the method may be strongly dependent on the choice of the initial step-size.

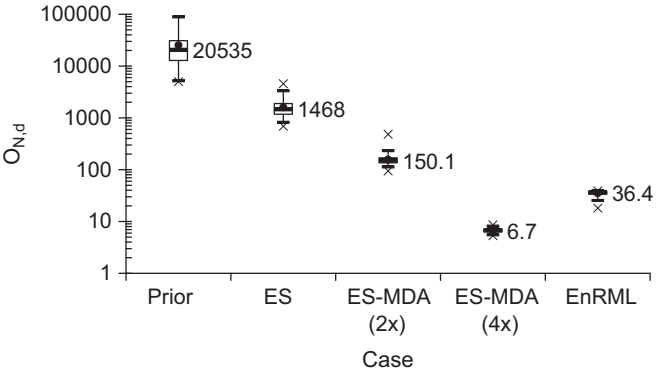

Fig. 5. Box plots of the normalized objective function: Example 2.

Fig. 2 presents the field water production rate for the first of the 10 ensembles. Besides the historical period (3900 days), we also include 3750 days of forecast. For comparison, we also show the prediction obtained by the true model in each plot. The results

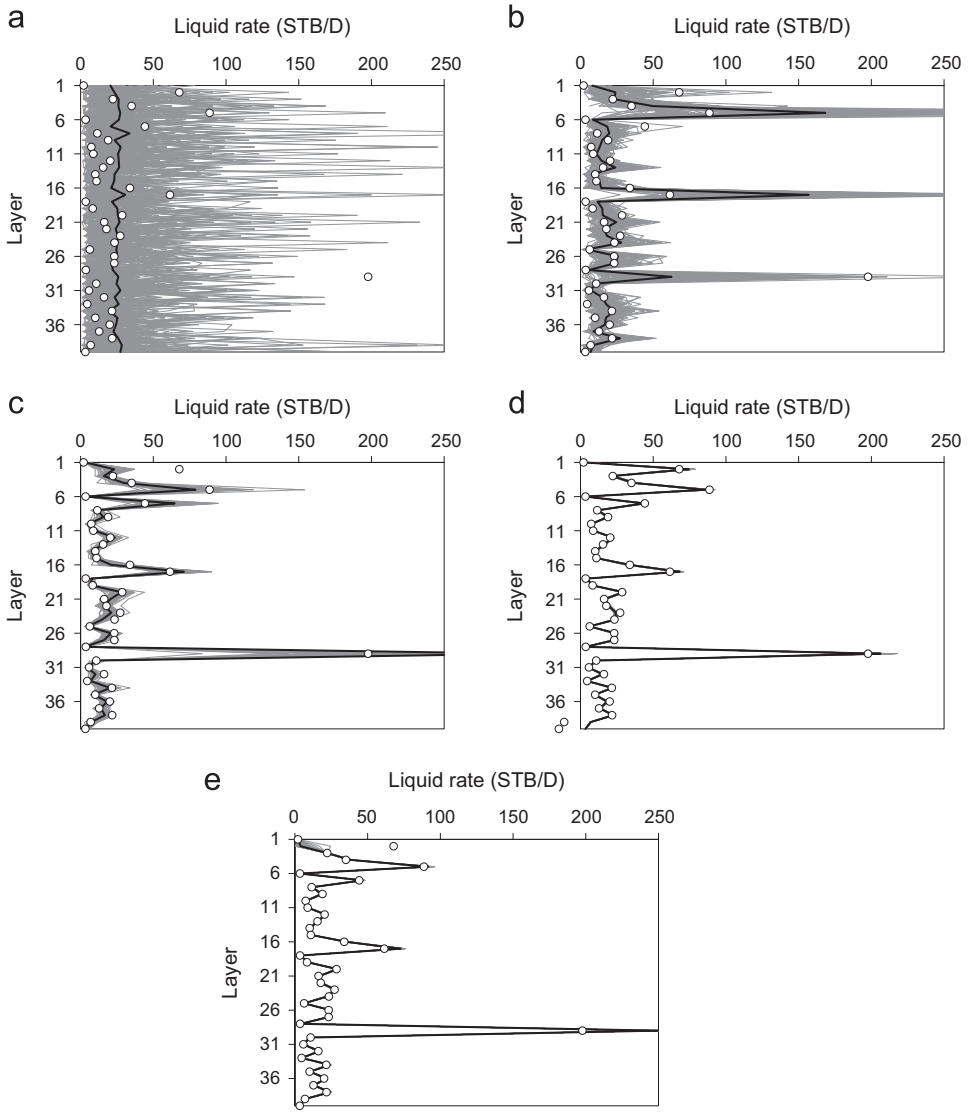

Fig. 6. Predicted liquid rate for each reservoir layer: Example 2. White dots are the history; solid curve is the ensemble mean prediction and gray curves are the prediction from the ensemble members. (a) Prior, (b) ES, (c) ES-MDA ($2\times$), (d) ES-MDA ($4\times$), and (e) batch-EnRML.

presented in Fig. 2 were obtained by running the final ensembles after data assimilation from time zero. Fig. 2 shows that ES failed in matching the water breakthrough time for most of the ensemble members. Two data assimilations were not enough to obtain good data matches of the water production rate. Four data assimilations, on the other hand, resulted in very good water data matches, which are better than the ones obtained by EnKF and batch-EnRML. For batch-EnRML with $\beta_0 = 1$, the predictions of water production rate are very similar to the ones obtained by ES because, for the first ensemble, the iterative process stopped at the second iteration because of five consecutive cuts in the step-size without reduction in the value of the average objective function. Fig. 3 presents the final mean permeability fields for the first ensemble. For visual comparison, we also include the true permeability field in this figure. Fig. 3 shows that the mean permeability field obtained by ES is much smoother than the ones obtained by EnKF, ES-MDA and batch-EnRML ($\beta_0 = 0.5$), which is another indication that the single linear update of ES is not enough for fully conditioning the ensemble of permeability fields to the observations. Fig. 4 presents the standard deviation of log-permeability after data assimilation for the first of the 10 ensembles considered in this section. This figure shows that improving the quality of the data matches results in reduction of

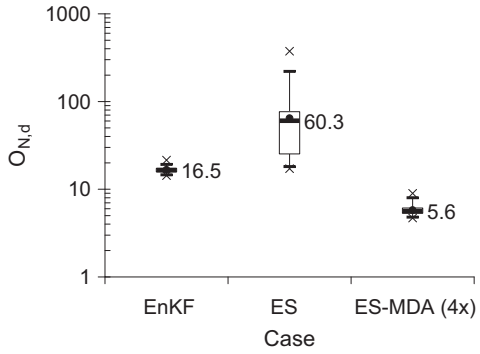

Fig. 8. Box plots of the normalized objective function: Brugge case.

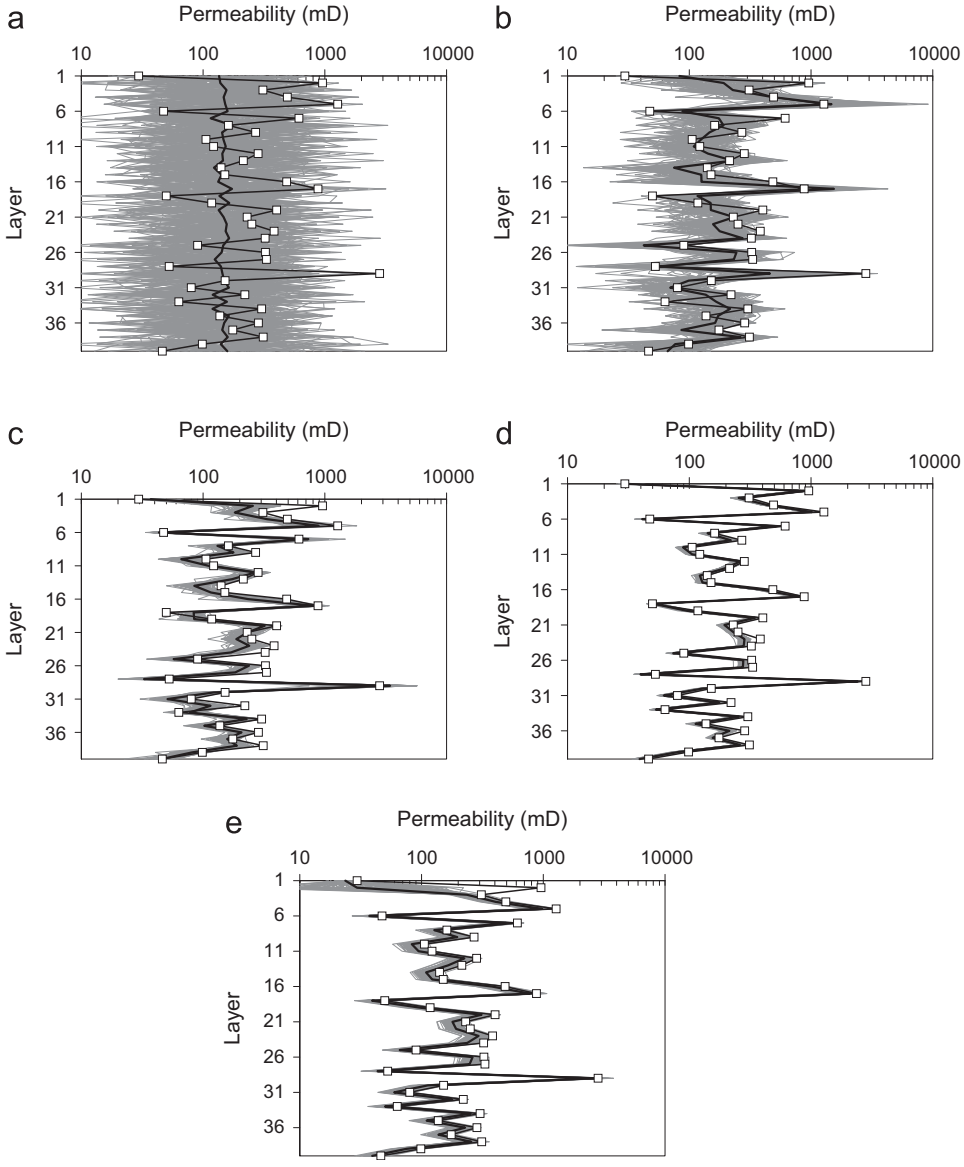

Fig. 7. Layer-permeabilities: Example 2. Curve with squares is the truth; solid curve is the ensemble mean and gray curves are the ensemble members. (a) Prior, (b) ES, (c) ES-MDA (2 ×), (d) ES-MDA (4 ×), and (e) batch-EnRML.

the ensemble variability. This is expectable because ensemble-based methods "search for solutions" in a subspace spanned by the initial ensemble. When we use an iterative method, we become more restrictive by searching for "solutions" with better data matches. Thus, it is reasonable to expect that there will be fewer independent "solutions" in this subspace which give a good data match, and this results in a reduction in the ensemble variance.

We estimated the average computational cost of each procedure in terms of the number of equivalent simulation runs, and the results are presented in Table 1. The CPU time required by the data assimilation with EnKF is 4.58 times larger than the CPU time used to run an initial ensemble from time zero without data assimilation. This difference occurs mainly because of the additional CPU time required by EnKF to restart reservoir simulations. The CPU time required by ES is practically the same as the CPU time used to run an ensemble without data assimilation. MDA increases the CPU time of ES by the factor equal to the number of data assimilations. For this example, the CPU time of ES-MDA ($4\times$) is slightly lower than the CPU time used by EnKF. For batch-EnRML, the iterative process often requires cutting the step-size and rerunning the ensemble, which results in a CPU time more than three times larger than the CPU time of ES-MDA ($4\times$). For this example, we used observations every 150 days during the data assimilations. Note that increasing the frequency of data would increase the computational cost of EnKF (because we would have more simulation restarts), but would have a small impact on the computational cost of ES, ES-MDA and batch-EnRML.

8. Example 2: production-logging data

The second example is a single-phase synthetic reservoir model on a 3D uniform grid with $11\times 11\times 40$ gridblocks. The dimensions of the gridblocks are 200 ft $\times$ 200 ft $\times$ 25 ft. In this model, there is a single producing well at the center of the reservoir completed in all 40 layers and operating at a constant rate of 1000 stb/day. The observed data correspond to a single measurement of liquid rate at each of the 40 layers after 30 days of production, mimicking the data obtained from a production-logging acquisition. The noise level added to each datum corresponds to 2% of the layer-rate predicted by the true model. We chose a small noise level to make the problem more challenging for data assimilation. The model parameters are the log-permeabilities for each of the 40 layers. Each layer of the reservoir has homogeneous and isotropic permeabilities. The log-permeability of each layer of the true model and the initial ensemble was obtained by independently sampling $\mathcal{N}(5.0,1.0)$. Hence, there is no correlation in the permeability between layers. This problem was designed to test our conjecture that sequential assimilation of data with overlapping information content is one of the reasons for the good performance of EnKF when history matching production data. For the problem considered in this section, the production-logging data are available at only one time. In this situation, EnKF becomes equivalent to ES.

We assimilated the production-logging data using ES, batch-EnRML and ES-MDA with two and four data assimilations. For batch-EnRML, we used an initial step-size $\beta_0 = 0.5$. For ES-MDA ($2\times$), we used $\alpha_1 = \alpha_2 = 2.0$. For ES-MDA ($4\times$), we used $\alpha_1 = 9.333$,

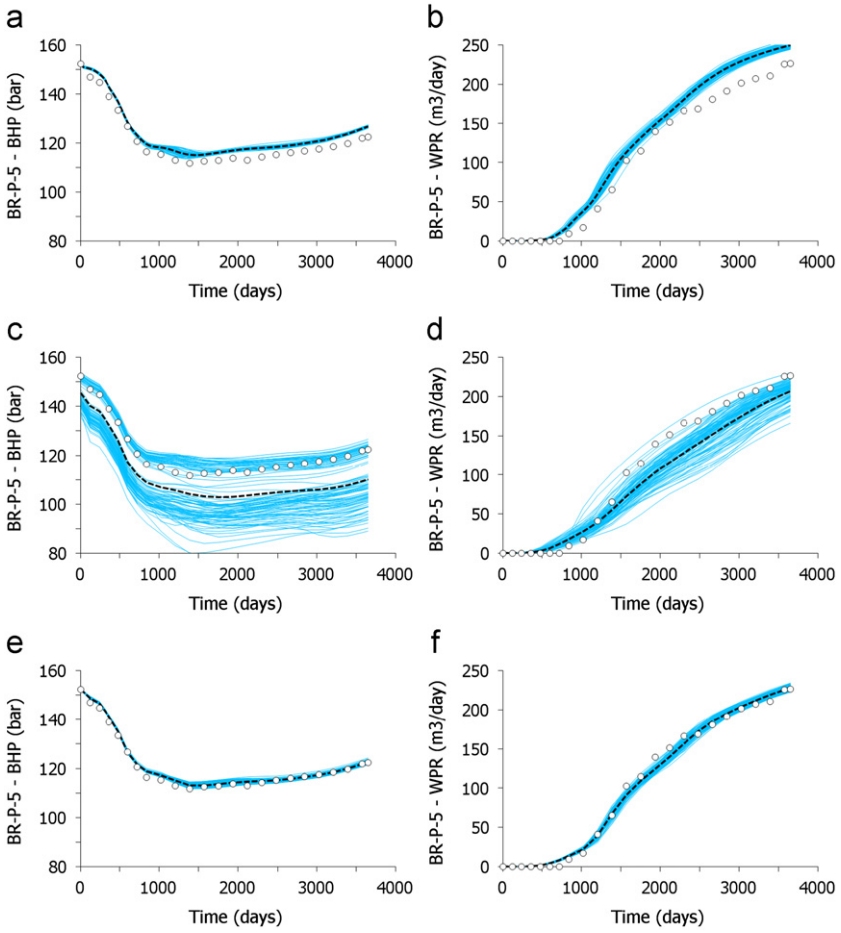

Fig. 9. Data matches for well P5: Brugge case. White dots are the history; dashed curve is the ensemble mean prediction. The remaining curves are the predictions from the ensemble members. (a) EnKF (bottomhole pressure), (b) EnKF (water rate), (c) ES (bottomhole pressure), (d) ES (water rate), (e) ES-MDA ($4\times$) (bottomhole pressure), and (f) ES-MDA ($4\times$) (water rate).

$\alpha_2 = 7.0$, $\alpha_3 = 4.0$ and $\alpha_4 = 2.0$. The ensemble size is 100. Fig. 5 presents the box plot of the normalized data mismatch objective function ($O_{N,d}$) obtained for each case. Here, $O_{N,d}$ was computed using

$$O_{N,d}(m) = \frac{1}{2N_d}(d - d_{obs})^T C_D^{-1}(d - d_{obs}). \quad (29)$$

According to the results in Fig. 5, ES with single data assimilation resulted in an unreasonably high value of $O_{N,d}$. ES-MDA significantly improved the data matches. After four data assimilations, the median value of $O_{N,d}$ was reduced to 6.7, which is 219 times smaller than the value obtained with the standard ES. Batch-EnRML obtained a median value of $O_{N,d}$ equal to 36.4, which is 5.4 times larger than the value obtained with ES-MDA ($4\times$). In terms of the computational cost, batch-EnRML required nine iterations which corresponded to a total of 1100 reservoir simulations, while ES-MDA ($4\times$) required 400 reservoir simulations. Fig. 6 presents the predicted layer-rates for each case. Fig. 6b indicates that ES resulted in possible overcorrection when matching the data from layers 5 and 17. ES-MDA ($2\times$) and batch-EnRML were not able to match data from layer 2 and overestimated the liquid rate from layer 29. ES-MDA ($4\times$), on the other hand, gave excellent data matches for all layers of the model. Fig. 7 presents the corresponding values of permeability for each reservoir layer showing that ES-MDA ($4\times$) captured the correct permeability field.

9. Example 3: Brugge field

The third example is the Brugge field (Peters et al., 2010). The Brugge field is a synthetic reservoir designed as a benchmark problem for evaluating methods for waterflooding optimization combined with history matching in a closed-loop workflow. A description of the case can be found in Peters et al. (2010). In the original Brugge dataset, there are 104 realizations of rock properties (porosity, horizontal and vertical permeabilities and net-to-gross ratio), 10 years of production history and a synthetic time-lapse seismic. Here, we history match only the production data, which correspond to "measurements" of the oil and water rates at the producing wells and bottomhole pressure at the producing and water injection wells. We assumed that the noise level corresponds to 3% for oil rate data and 5% for water rate data. For bottomhole pressure, we assumed a constant measurement error of 0.5 bars (7.25 psi). Even though data are available with a frequency of 30 days, we assimilate data every 120 days using EnKF, ES and ES-MDA ($4\times$). For this last case, the coefficients α_i's are the same as in the case ES-MDA ($4\times$-b) of the example 1. Because the number of wells in the Brugge field is large (20 producers and 10 water injectors), we used localization to reduce problems related to sampling errors and limited degrees of freedom. We defined the localization regions using the procedure described in Emerick and Reynolds (2011a,b).

Fig. 8 presents the box plot of $O_{N,d}$ obtained from each case. According to the results presented in this figure, ES resulted in values of $O_{N,d}$ about four times larger than those based on results

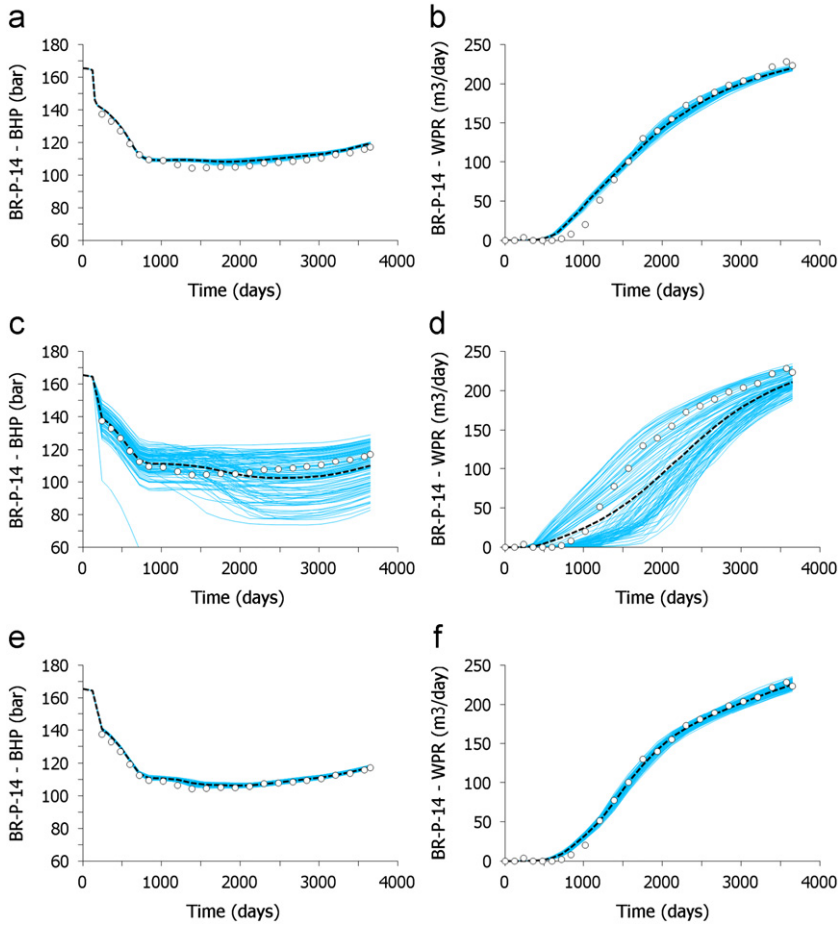

Fig. 10. Data matches for well P14: Brugge case. White dots are the history; dashed curve is the ensemble mean prediction. The remaining curves are the predictions from the ensemble members. (a) EnKF (bottomhole pressure), (b) EnKF (water rate), (c) ES (bottomhole pressure), (d) ES (water rate), (e) ES-MDA ($4\times$) (bottomhole pressure), and (f) ES-MDA ($4\times$) (water rate).

from EnKF. ES-MDA ($4\times$) improved significantly the final data matches and the values of $O_{N,d}$ are about three times lower than EnKF. Figs. 9 and 10 present the data matches obtained for two producing wells (P5 and P14). These figures show that ES was not able to achieve reasonable data matches. Also, the predictions from the ES-MDA ($4\times$) are in better agreement with the historical data than the ones obtained by EnKF; see, e.g., Fig. 9a and b, which shows that EnKF results in slightly biased predictions compared to the historical data. Although we show the plots of only two wells, these results are representative of what is observed in most of the wells. Fig. 11 presents the first layer of the prior and final permeability fields for one of the 104 realizations of the ensemble. Fig. 11 shows that EnKF, ES and ES-MDA ($4\times$) increased the permeability of the prior realization. The three methods resulted in visually similar permeability fields. Fig. 12 shows the ensemble standard deviation of log-permeability, $\ln(k)$, before and after data assimilation. Because EnKF and ES-MDA ($4\times$) resulted better matches of data, both methods resulted in lower values of standard deviation of $\ln(k)$ than were obtained from the standard ES.

Table 2 presents the computational cost in terms of the equivalent number of reservoir simulation runs for EnKF, ES and ES-MDA ($4\times$). The results in this table were obtained by measuring the actual CPU time required during the data assimilations divided by the average CPU time required for one reservoir simulation run. Our covariance localization implementation requires running the ensemble mean model and computing pseudo-tracer concentrations, as described in Emerick and Reynolds (2011a,b). For this reason, the CPU time required by ES was higher than the CPU time required to run the 104 prior realizations. The CPU time of ES-MDA ($4\times$) was 17% higher than the CPU time of EnKF, but resulted in far better data matches than were obtained with EnKF.

Table 2
Computational costs: Brugge case.

Case	Number of equivalent simulation runs
EnKF	367
ES	106
ES-MDA ($4\times$)	430

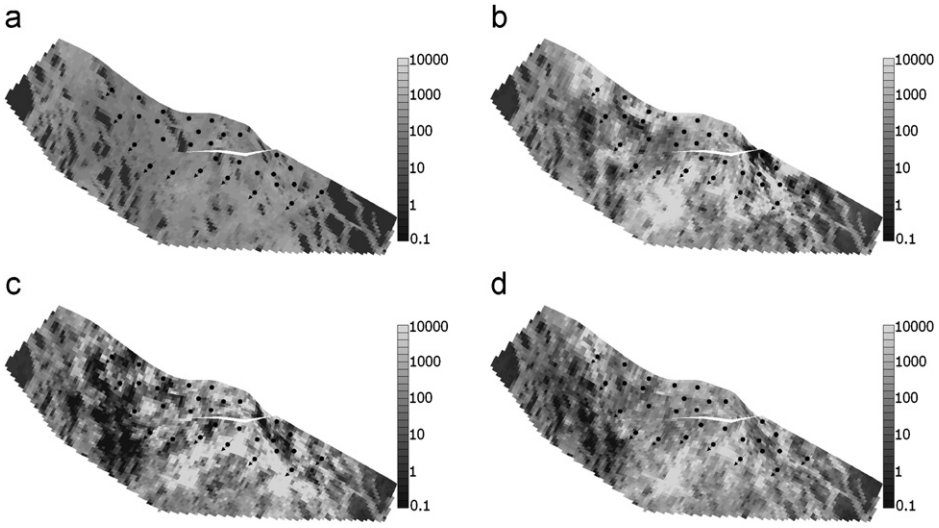

Fig. 11. Permeability (mD) of the first layer of one of the 104 realizations: Brugge case. (a) Prior, (b) EnKF, (c) ES, and (d) ES-MDA ($4\times$).

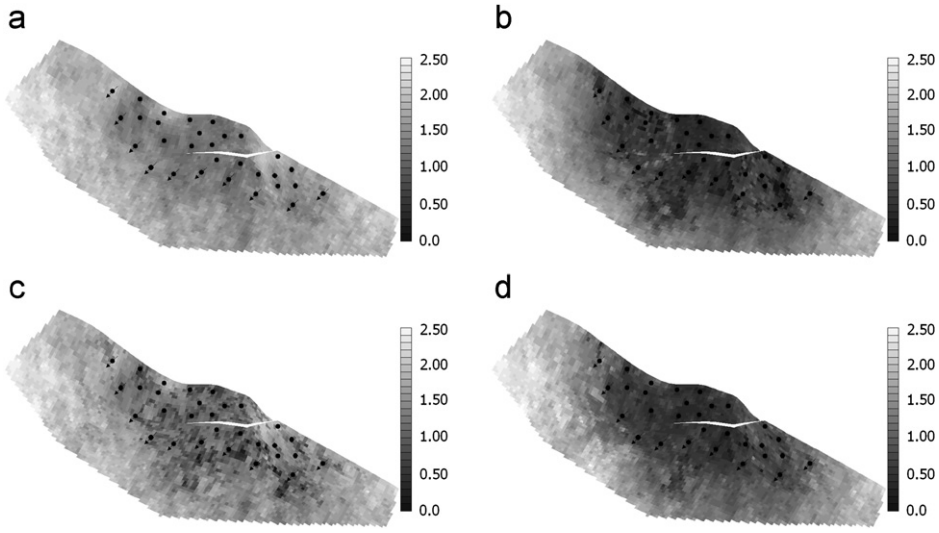

Fig. 12. Standard deviation of log-permeability of the first layer: Brugge case. (a) Prior, (b) EnKF, (c) ES, and (d) ES-MDA ($4\times$).

10. Important comments

This work was motivated by our conjecture that when sequential data have similar information content, the sequential assimilation of data employed by EnKF provides an approximation of multiple Gauss–Newton iterations. This conjecture yields a second conjecture, namely, EnKF will typically give a better data match than ES because the smoother effectively represents a single Gauss–Newton iteration with an average sensitivity matrix (Reynolds et al., 2006). Although we have not provided a proof of these conjectures, all results we have presented here are consistent with these conjectures.

For ES-MDA, we only consider the parameter-estimation problem. Thus, unlike EnKF, the parameters and states are always consistent (Thulin et al., 2007). This fact helps to explain the better data matches obtained by ES-MDA compared to EnKF.

11. Conclusions

In this paper, we introduced ES-MDA as an efficient iterative form of ES. Even though we presented ES-MDA in the context of reservoir history matching, the method is very general and can be applied in other research areas. Moreover, ES-MDA requires very few modifications to a standard ES implementation. Based on the results from the test problems, the following conclusions are warranted:

- ES performed poorly compared to standard EnKF and ES-MDA.
- ES-MDA resulted in better data matches than EnKF with comparable computational cost.
- The results of the first and second examples indicate that ES-MDA is computationally less expensive than batch-EnRML and results in better data matches. Based on the results presented in Chen and Oliver (2012), this conclusion also applies to the Brugge case.
- For the three examples considered in the paper, four data assimilations were enough for providing good data matches.
- The use of the inflation coefficients, α_i's, in a decreasing order resulted in only small improvements compared to using all α_i equal to the number of data assimilations.

Acknowledgments

The support of the member companies of TUPREP is gratefully acknowledged. The first author acknowledges the financial support from Petrobras.

References

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C., Vallés, B., 2009. Review of ensemble Kalman filter in petroleum engineering. SPE Journal 14 (3), 393–412.
Bertino, L., Evensen, G., Wackernagel, H., 2003. Sequential data assimilation techniques in oceanography. International Statistical Review 71 (2), 223–241.
Bianco, A., Cominelli, A., Dovera, L., Nævdal, G., Vallès, B., 2007. History matching and production forecast uncertainty by means of the ensemble Kalman filter: a real field application. In: Proceedings of the EAGE/EUROPEC Conference and Exhibition, London, U.K., 11–14 June. No. SPE 107161.
Chen, C., Li, G., Reynolds, A.C., 2010. Closed-loop reservoir management on the Brugge test case. Computational Geosciences 14 (4), 691–703.
Chen, Y., Oliver, D., 2010. Ensemble-based closed-loop optimization applied to Brugge field. SPE Reservoir Evaluation & Engineering 13 (1), 56–71.
Chen, Y., Oliver, D., 2012. Ensemble randomized maximum likelihood method as an iterative ensemble smoother. Mathematical Geosciences 44 (1), 1–26.
Chen, Y., Zhang, D., 2006. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. Advances in Water Resources 29 (8), 1107–1122.
Emerick, A.A., Reynolds, A.C., 2011a. Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. Computational Geosciences 15 (2), 251–269.
Emerick, A.A., Reynolds, A.C., 2011b. History matching a field case using the ensemble Kalman filter with covariance localization. SPE Reservoir Evaluation & Engineering 14 (4), 423–432.
Emerick, A.A., Reynolds, A.C., 2012. History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. Computational Geosciences 16 (3), 639–659.
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), 10143–10162.
Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics 54 (6), 539–560.
Evensen, G., 2007. Data Assimilation: The Ensemble Kalman Filter. Springer, Berlin.
Evensen, G., Hove, J., Meisingset, H.C., Reiso, E., Seim, K.S., Espelid, O., 2007. Using the EnKF for assisted history matching of a North Sea reservoir model. In: Proceedings of the SPE Reservoir Simulation Symposium, Houston, TX, 26–28 February. No. SPE 106184.
Evensen, G., van Leeuwen, P.J., 2000. An ensemble Kalman smoother for nonlinear dynamics. Monthly Weather Review 128 (6), 1852–1867.
Gao, G., Reynolds, A.C., 2006. An improved implementation of the LBFGS algorithm for automatic history matching. SPE Journal 11 (1), 5–17.
Gu, Y., Oliver, D.S., 2007. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. SPE Journal 12 (4), 438–446.
Haugen, V., Nævdal, G., Natvik, L.-J., Evensen, G., Berg, A.M., Flornes, K.M., 2008. History matching using the ensemble Kalman filter on a North Sea field case. SPE Journal 13 (4), 382–391.
Houtekamer, P.L., Mitchell, H.L., 2005. Ensemble Kalman filtering. Quarterly Journal of the Royal Meteorological Society 131, 3269–3289.
Jansen, J., Brouwer, D., Naevdal, G., van Kruijsdijk, C., 2005. Closed-loop reservoir management. First Break 23, 43–48.
Jansen, J.D., Douma, S.D., Brouwer, D.R., den Hof, P.M.J.V., Heemink, A.W., 2009. Closed-loop reservoir management. In: Proceedings of the SPE Reservoir Simulation Symposium, The Woodlands, TX, 2–4 February. No. SPE 119098.
Keppenne, C.L., Rienecker, M.M., 2003. Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble Kalman filter. Journal of Marine Systems 40–41, 363–380.
Liu, G., Chen, Y., Zhang, D., 2008. Investigation of flow and transport processes at the MADE site using ensemble Kalman filter. Advances in Water Resources 31 (7), 975–986.
Nævdal, G., Mannseth, T., Vefring, E.H., 2002. Near-well reservoir monitoring through ensemble Kalman filter. In: Proceedings of the SPE/DOE Improved Oil Recovery Symposium, 13–17 April. No. SPE 75235.
Oliver, D.S., Chen, Y., 2009. Improved initial sampling for the ensemble Kalman filter. Computational Geosciences 13 (1), 13–27.
Oliver, D.S., Chen, Y., 2010. Recent progress on reservoir history matching: a review. Computational Geosciences 15 (1), 185–221.
Peters, L., Arts, R., Brouwer, G., Geel, C., Cullick, S., Lorentzen, R., Chen, Y., Dunlop, K., Vossepoel, F., Xu, R., Sarma, P., Alhuthali, A., Reynolds, A., 2010. Results of the Brugge benchmark study for flooding optimisation and history matching. SPE Reservoir Evaluation & Engineering 13 (3), 391–405.
Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic data assimilation with the ensemble Kalman filter. Monthly Weather Review 130 (1), 103–114.
Reynolds, A.C., He, N., Oliver, D.S., 1999. Reducing uncertainty in geostatistical description with well testing pressure data. In: Schatzinger, R.A., Jordan, J.F. (Eds.), Reservoir Characterization—Recent Advances. American Association of Petroleum Geologists, pp. 149–162.
Reynolds, A.C., Zafari, M., Li, G., 2006. Iterative forms of the ensemble Kalman filter. In: Proceedings of 10th European Conference on the Mathematics of Oil Recovery, Amsterdam, 4–7 September.
Rommelse, J., 2009. Data Assimilation in Reservoir Management. Ph.D. Thesis, Technical University of Delft, Delft, The Netherlands.
Skjervheim, J.-A., Evensen, G., Aanonsen, S.I., Ruud, B.O., Johansen, T.A., 2007. Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. SPE Journal 12 (3), 282–292.
Skjervheim, J.-A., Evensen, G., Hove, J., Vabø, J.G., 2011. An ensemble smoother for assisted history matching. In: Proceedings of the SPE Reservoir Simulation Symposium, The Woodlands, TX, USA, 21–23 February. No. SPE 141929.
Szunyogh, I., Kostelich, E.J., Gyarmati, G., Patil, D.J., Hunt, B.R., Kalnay, E., Ott, E., Yorke, J.A., 2005. Assessing a local ensemble Kalman filter: perfect model experiments with the national centers for environmental prediction global model. Tellus A 57, 528–545.
Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM, Philadelphia, USA.
Thulin, K., Li, G., Aanonsen, S.I., Reynolds, A.C., 2007. Estimation of initial fluid contacts by assimilation of production data with EnKF. In: Proceedings of the SPE Annual Technical Conference and Exhibition, Anaheim, CA, 11–14 November. No. SPE 109975.
van Leeuwen, P.J., Evensen, G., 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. Monthly Weather Review 124, 2898–2913.
Wang, C., Li, G., Reynolds, A.C., 2009. Production optimization in closed-loop reservoir management. SPE Journal 14 (3), 506–523.
Wang, Y., Li, G., Reynolds, A.C., 2010. Estimation of depths of fluid contacts by history matching using iterative ensemble-Kalman smoothers. SPE Journal 15 (2).
Wu, Z., Reynolds, A.C., Oliver, D.S., 1999. Conditioning geostatistical models to two-phase production data. SPE Journal 3 (2), 142–155.
Zachariassen, E., Skjervheim, J., Vabø, J., Lunt, I., Hove, J., Evensen, G., 2011. Integrated work flow for model update using geophysical monitoring data. In: Proceedings of the 73rd EAGE Conference & Exhibition, Vienna, Austria, 23–26 May.