Comput Geosci (2012) 16:467–482
DOI 10.1007/s10596-011-9262-2

ORIGINAL PAPER

Comparing the adaptive Gaussian mixture filter with the ensemble Kalman filter on synthetic reservoir models

Andreas S. Stordal · Randi Valestrand · Hans Arnfinn Karlsen · Geir Nævdal · Hans Julius Skaug

Received: 26 January 2011 / Accepted: 11 October 2011 / Published online: 23 November 2011

Abstract Over the last years, the ensemble Kalman filter (EnKF) has become a very popular tool for history matching petroleum reservoirs. EnKF is an alternative to more traditional history matching techniques as it is computationally fast and easy to implement. Instead of seeking one best model estimate, EnKF is a Monte Carlo method that represents the solution with an ensemble of state vectors. Lately, several ensemble-based methods have been proposed to improve upon the solution produced by EnKF. In this paper, we compare EnKF with one of the most recently proposed methods, the adaptive Gaussian mixture filter (AGM), on a 2D synthetic reservoir and the Punq-S3 test case. AGM was introduced to loosen up the requirement of a Gaussian prior distribution as implicitly formulated in EnKF. By combining ideas from particle filters with EnKF, AGM extends the low-rank kernel particle Kalman filter. The simulation study shows that while both methods match the historical data well, AGM is better at preserving the geostatistics of the prior distribution. Further, AGM also produces estimated fields that have a higher empirical correlation with the reference field than the corresponding fields obtained with EnKF.

A. S. Stordal (✉) · R. Valestrand · G. Nævdal
International Research Institute of Stavanger,
Thormøhlensgate 55, 5008 Bergen, Norway
e-mail: Andreas.S.Stordal@iris.no

H. A. Karlsen · G. Nævdal · H. J. Skaug
Department of Mathematics, University of Bergen,
Johs. Brunsgt. 12, P.O. Box 7803, 5020 Bergen, Norway

Keywords History matching · EnKF · Gaussian mixture filters · Punq-S3

1 Introduction

From a statistical point of view, the optimal solution to the history matching problem is the posterior distribution of the states and parameters in the reservoir given all the measurements. For practical purposes, the posterior distribution is represented by a finite set of ensemble members (a sample) obtained from a Monte Carlo sampling approach. One of the most popular Monte Carlo methods for reservoir characterization is the ensemble Kalman filter [8] or improved versions of it [1]. The ensemble Kalman filter (EnKF) is a sequential Monte Carlo (SMC) method that uses the system dynamics to integrate the sample between time steps. When new measurements arrive, the Kalman equations are used to update each ensemble member. Since the Kalman equations update the variables by a linear integration using only first- and second-order moments, it suffers from severe bias when the model is far from linear and/or Gaussian. Asymptotically unbiased estimates may be obtained using SMC methods [7]; however, these methods suffer from the curse of dimensionality [5]; hence, they are too expensive for reservoirs above a certain size. Several methods have been proposed to combine the EnKF with SMC methods. Some recent references are Bengtsson et al. [4], Hoteit et al. [12], Mandel and Beezley [15], Lei and Bickel (submitted for publication), and Stordal et al. [21]. In this paper, we apply for the first time

to reservoir models the recently proposed adaptive Gaussian mixture filter (AGM) introduced by Stordal et al. [21] and compare it with EnKF. AGM loosens up the requirement of a linear/Gaussian model by making a smaller linear update than EnKF and take advantage of some of the information contained in the importance weights as is standard in SMC methods but with approximately the same computational cost as EnKF. AGM can be regarded as an approximation of the optimal SMC [6] when the model and observation errors are Gaussian and the measurement operator is linear.

Both EnKF and AGM aim at building up sequentially in time a finite ensemble approximation of the posterior distribution. EnKF can be viewed as a parametric method, where the forecast distribution at each time step is Gaussian with mean vector and covariance matrix estimated from the ensemble. When a new measurement arrives, also assumed to follow a Gaussian distribution, the estimated cross covariance of the state and observation is used to update the mean and covariance matrix of the Gaussian posterior distribution. AGM is more of a non-parametric method, where the prior distribution is approximated by a weighted sum of Gaussian distributions where each ensemble member represents the mean vector of a Gaussian density. The covariance matrix of each density is, as in EnKF, estimated from the ensemble but is smaller as it is multiplied by the square of a bandwidth h less than or equal to 1. That is if P is the covariance matrix, then $P_{\text{EnKF}} - P_{\text{AGM}}$ is positive semi definite.

In AGM, the covariance matrix of each Gaussian distribution represents the unknown model error. Then as the new observations arrive, the cross covariances between the model error and the observations are used to update the mean and covariance matrix of each of the Gaussian densities. The weights of the mixture are then updated according to Bayes' rule. The advantage of this approximation is that the linear update will move the ensemble members into areas of higher likelihood, contrary to standard SMC methods that only evaluates the likelihood of each ensemble member. However, to avoid linearizing the model too much, the bandwidth h should not be too large. A study of the Lorenz40 model [21] showed that to maintain both small linear updates and avoid a collapse of the weights, a tuning parameter, denoted by α, had to be introduced. This parameter interpolates the importance weights with the uniform weights used in EnKF. The tuning parameter α is chosen adaptively so that by construction, the effective sample size will always stay above 80% and a degeneracy of the weights will not occur [21]. For the Lorenz40 model, AGM outperformed both EnKF and the low-rank kernel particle Kalman filter with respect to RMSE and marginal Kullback–Leibler divergence criteria.

In this paper, we compare AGM with the standard EnKF on a synthetic 2D reservoir model and the Punq-S3 [17] test case. For the 2D model, AGM obtains a lower data mismatch than EnKF and is able to better preserve the prior geostatistics. For the Punq-S3 model, the data mismatch for AGM is slightly larger than the data mismatch from EnKF, but the value of the objective function that includes a prior mismatch (or regularization) term is smaller for AGM. AGM also maintains a higher ensemble spread which is good for prediction purposes and at the same time produces estimated parameter fields that are closer to the true parameter fields in mean squared error than EnKF. For the Punq-S3 model, the two methods are also validated by the consistency of the predicted total cumulative oil production where the cumulative distribution function (cdf) is estimated from repeated runs with different initial ensembles. The results clearly show that the results from AGM have less spread than the standard EnKF. That is, the results are less sensitive to the initial ensemble. In Section 2, we present the methodology before we describe the reservoirs and the results in Sections 3 and 4. A summary is given in Section 5.

2 Methodology

The model under consideration is the state-space model

$$\mathbf{X}_n = \mathcal{M}_n(\mathbf{X}_{n-1}) + \xi_n$$
$$\mathbf{Y}_n = \mathbb{H}_n \mathbf{X}_n + \epsilon_n,$$

where $\mathcal{M}_n$ is the nonlinear model operator, $\mathbb{H}$ is the linear measurement operator, $\mathbf{X}_n$ is the partially hidden state, $\mathbf{Y}_n$ are the observations and ξ_n, and ϵ_n are the stochastic model and observation error. The model error ξ_n is (usually) not accounted for in reservoir models, which from a filter perspective is undesirable (if two particles become identical, they will stay identical unless some artificial perturbation is added). The index $n = 0, 1 \ldots$ denotes the times at which observation $\mathbf{Y}_n$ arrives. The variable $\mathbf{X}_n$ usually consists of both dynamic variables and parameters. For a reservoir model, $\mathbf{X}_n$ typically contains parameters such as the permeability and porosity fields and dynamic variables such as the saturation and pressure fields and the simulated observations. The observations usually consist of production

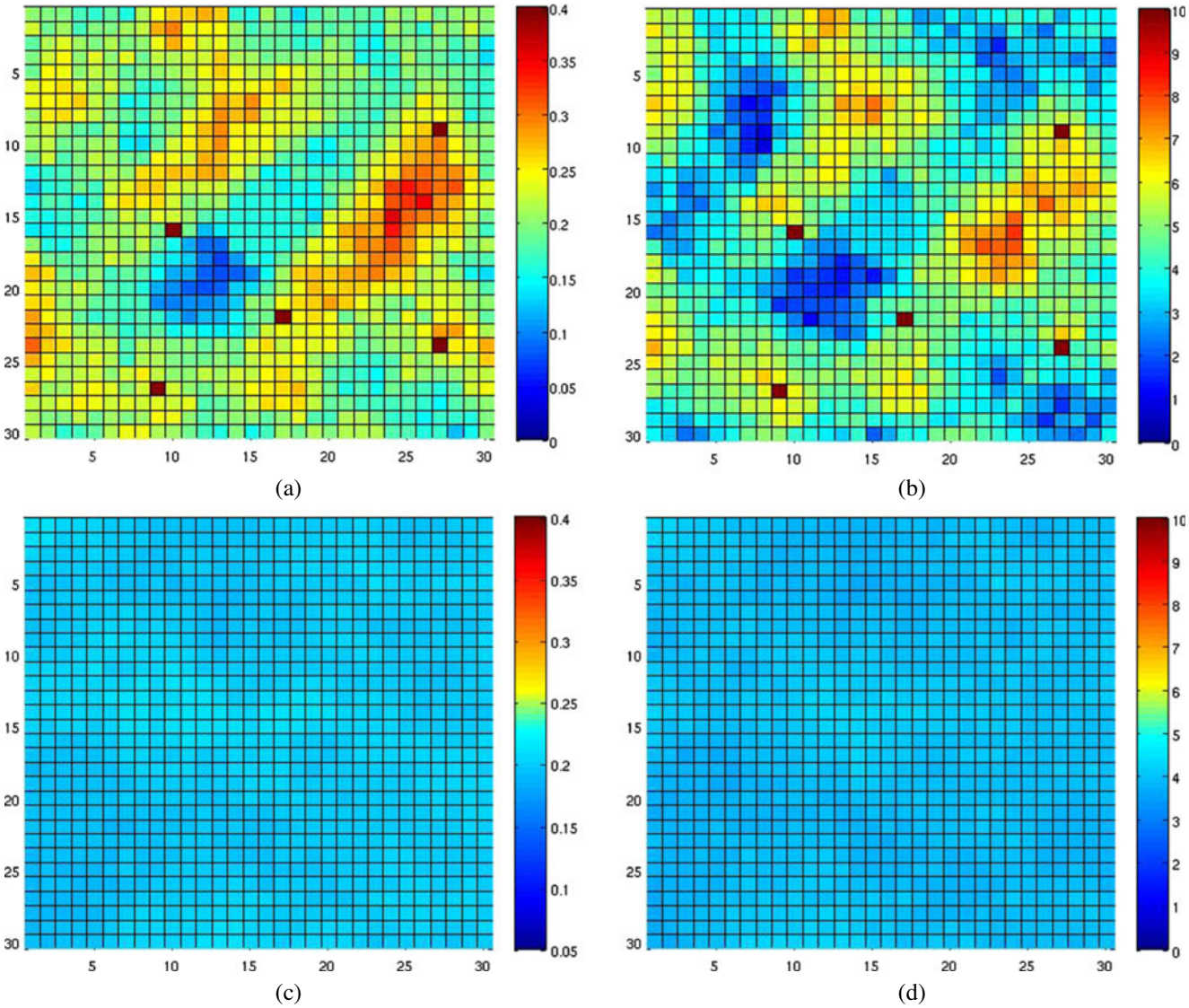

Fig. 1 True log permeability and porosity fields with well placements (*dark red*) and mean of initial ensemble. **a** True porosity. **b** True log permeability. **c** Initial mean of porosity. **d** Initial mean of log permeability

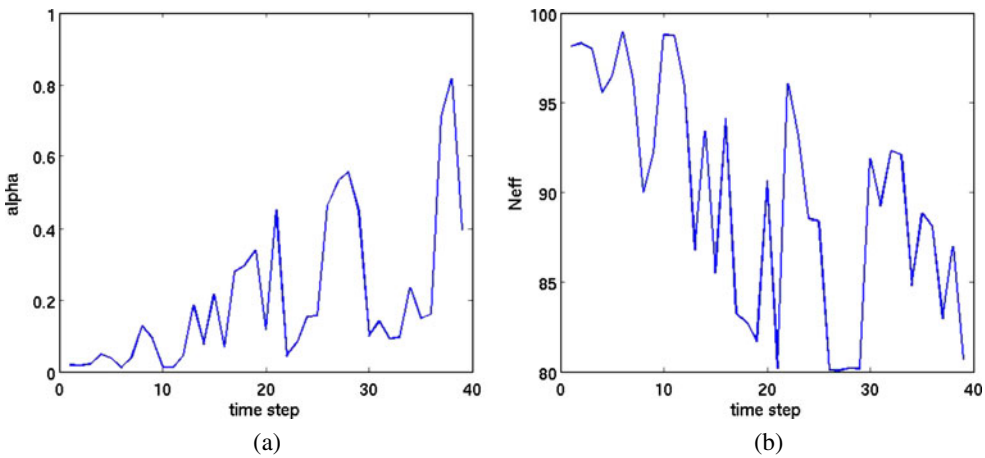

Fig. 2 Evolution of α and the effective sample size for $h = 0.3$ in the 2D model. **a** α. **b** Effective sample size

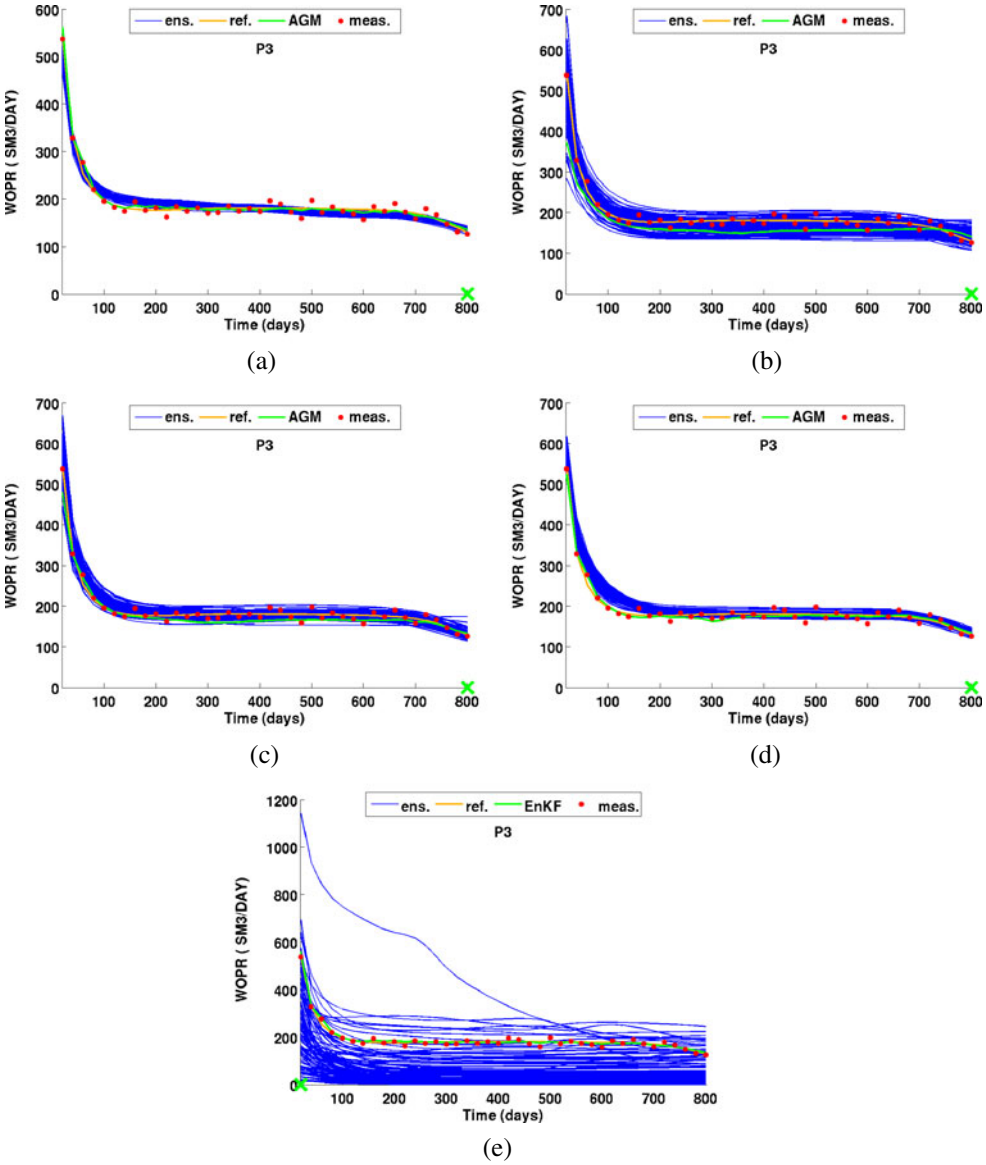

Fig. 3 Predicted oil production rate in P3, rerun from time zero. **a** EnKF. **b** AGM $h = 0.1$. **c** AGM $h = 0.2$. **d** AGM $h = 0.3$. **e** Initial

data. Our aim is to build a sample from the posterior distribution $p(\mathbf{x}_n|\mathbf{y}_{1:n})$, where $\mathbf{y}_{1:n} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ is the vector of all observations up to and including time n.

We now give a brief description of the EnKF and the AGM for completeness; for more details, see Stordal et al. [21]. Given a set of particles (an ensemble), $\{\hat{\mathbf{x}}_{n-1}^i\}_{i=1}^N$, from the posterior density $p(\mathbf{x}_{n-1}|\mathbf{y}_{1:n-1})$, the EnKF then uses the Markov property of the system to obtain a sample from the prior by sampling from the forward density $p(\mathbf{x}_n|\hat{\mathbf{x}}_{n-1}^i)$ according to

$$\mathbf{x}_n^i = \mathcal{M}_n(\hat{\mathbf{x}}_{n-1}^i) + \xi_n^i, \quad i = 1, \ldots, N.$$

With the additional assumption that $p(\mathbf{x}_n|\mathbf{y}_{1:n-1})$ is Gaussian with mean μ_n and covariance matrix $\mathbb{P}_n$, it follows from the Gaussian observation error that $p(\mathbf{x}_n|\mathbf{y}_{1:n})$ is Gaussian with mean $\tilde{\mu}_n$ and covariance matrix $\tilde{\mathbb{P}}_n$,

$$\begin{aligned}
\tilde{\mu}_n &= \mu_n + \mathbb{K}_n(\mathbf{y}_n - \mathbb{H}_n\mu_n),\\
\tilde{\mathbb{P}}_n &= (\mathbb{I} - \mathbb{K}_n\mathbb{H}_n)\mathbb{P}_n,\\
\mathbb{K}_n &= \mathbb{P}_n\mathbb{H}_n^T(\mathbb{H}_n\mathbb{P}_n\mathbb{H}_n^T + \mathbb{R}_n)^{-1}.
\end{aligned} \tag{1}$$

The EnKF takes $\mathbb{P}_n = \mathbb{S}_X$ to be the empirical covariance matrix of the N particles $\{\mathbf{x}_n^i\}_{i=1}^N$, and in updating particle i, it puts $\mu_n = \mathbf{x}_n^i$ and $\mathbf{y}_n = \mathbf{y}_n + \epsilon_n^i$. Thus, a sample from the posterior is provided by

$$\tilde{\mathbf{x}}_n^i = \mathbf{x}_n^i + \mathbb{K}_n(\mathbf{y}_n + \epsilon_n^i - \mathbb{H}_n\mathbf{x}_n^i).$$

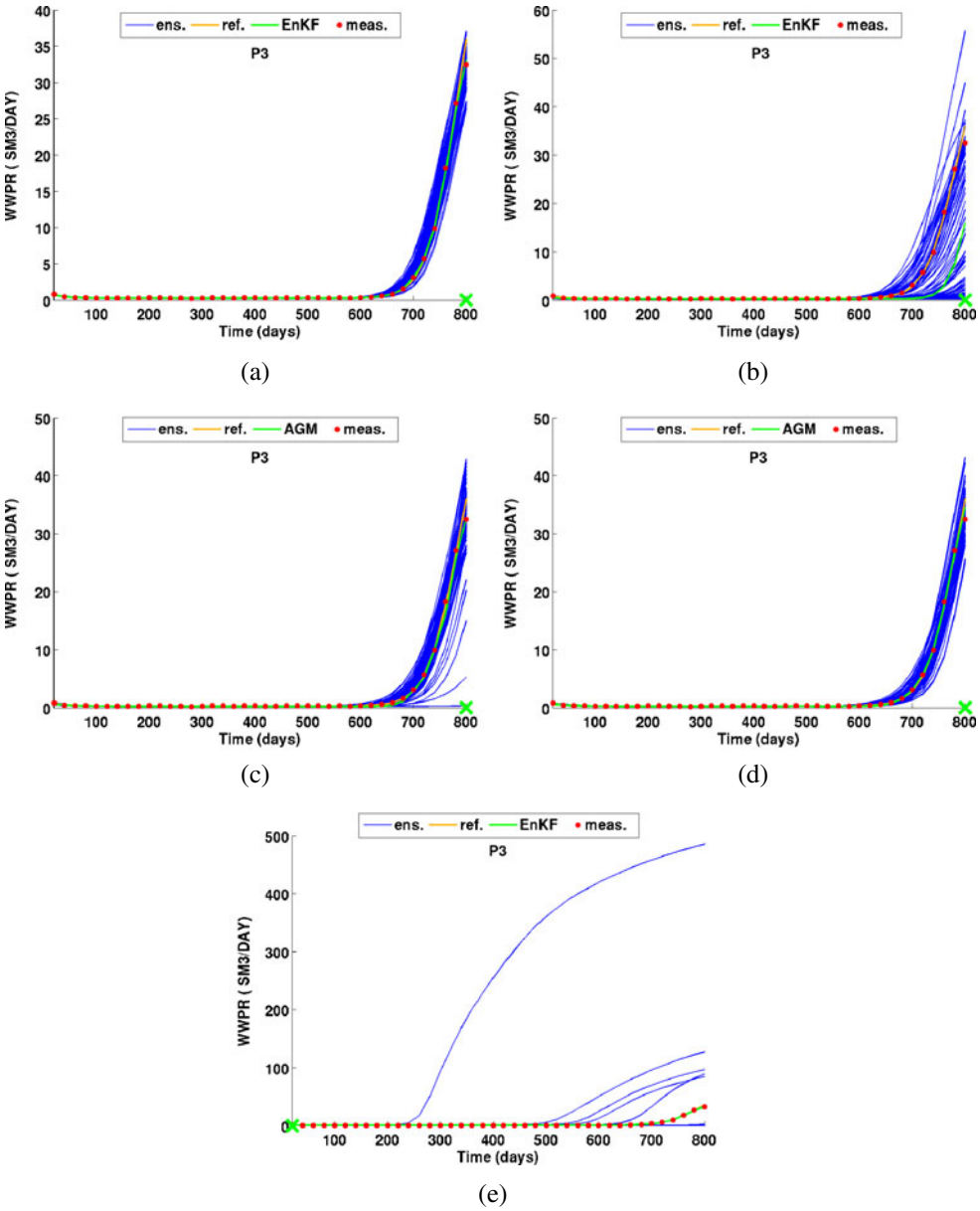

Fig. 4 Predicted water production rate in P3, rerun from time zero. **a** EnKF. **b** AGM $h = 0.1$. **c** AGM $h = 0.2$. **d** AGM $h = 0.3$. **e** Initial

Due to sampling error, the relationship $\tilde{\mathbb{P}}_n = (\mathbb{I} - \mathbb{K}_n\mathbb{H}_n)\mathbb{P}_n$ is not satisfied in the EnKF since $\tilde{\mathbb{P}}_n$ is approximated from the updated ensemble. Due to the Gaussian assumption, the EnKF is biased for a general nonlinear system. However, it is robust in the sense that it does not suffer from the dimensionality problem in the same way as other SMC methods since the importance weights attached to the ensemble members are uniform by construction. Also it is more computationally efficient than other rigorous approaches such as Markov chain Monte Carlo [18] and Langevin sampling [3].

Instead of approximating the prior with a Gaussian density at each time step n, the prior density $p(\mathbf{x}_n|\mathbf{y}_{1:n-1})$ in Gaussian mixture filters is approximated by a sum of Gaussian kernels [19], where each particle represents the mean of a Gaussian density with a covariance matrix $\mathbb{P}_n$, i.e.,

$$P_N(\mathbf{x}_n|\mathbf{y}_{1:n-1}) = \sum_{i=1}^{N} w_{n-1}^i \Phi(\mathbf{x}_n - \mathbf{x}_n^i, \mathbb{P}_n), \tag{2}$$

where $\Phi(\mathbf{x}, \mathbb{C})$ denotes a multivariate Gaussian density with zero mean, covariance matrix $\mathbb{C}$, and N is the

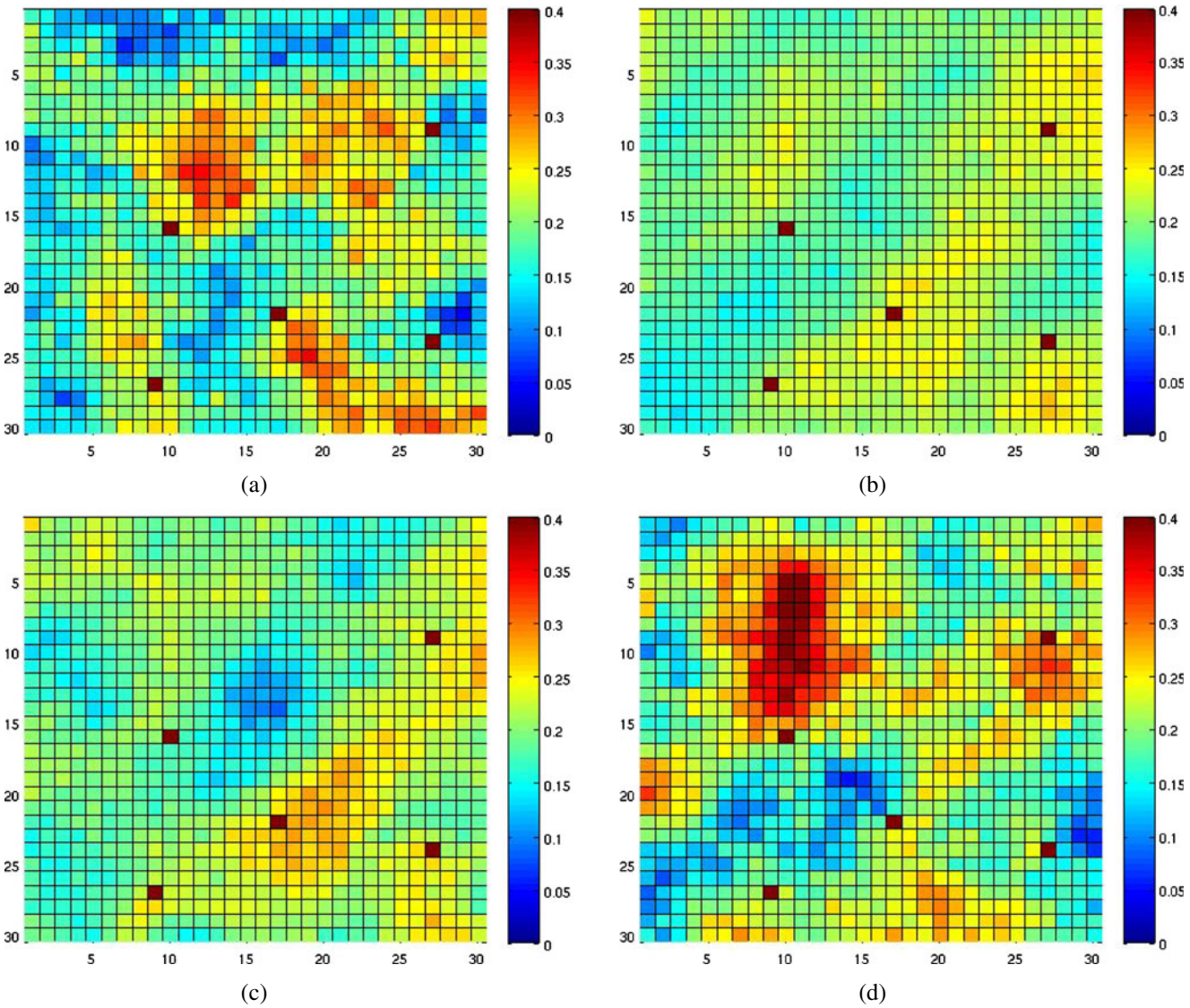

Fig. 5 Mean porosity of the ensemble members. Final time step. **a** EnKF. **b** AGM $h = 0.1$. **c** AGM $h = 0.2$. **d** AGM $h = 0.3$

number of ensemble members. Since we assume that the measurement noise follow a Gaussian distribution, the posterior density approximation is also a Gaussian mixture known up to a normalizing constant via Bayes' theorem,

$$P_N(\mathbf{x}_n|\mathbf{y}_{1:n}) \propto P_N(\mathbf{x}_n|\mathbf{y}_{1:n-1})\Phi(\mathbf{y}_n - \mathbb{H}_n\mathbf{x}_n, \mathbb{R}_n). \tag{3}$$

By adding and subtracting $\mathbb{H}_n\mathbf{x}_n^i$ in the first argument of $\Phi(\mathbf{y}_n - \mathbb{H}_n\mathbf{x}_n, \mathbb{R}_n)$ and normalizing the weights, the posterior mixture is given by

$$P_N(\mathbf{x}_n|\mathbf{y}_{1:n}) = \sum_{i=1}^{N} w_n^i \Phi(\mathbf{x}_n - \tilde{\mathbf{x}}_n^i, \tilde{\mathbb{P}}_n),$$

where

$$\begin{aligned}
\tilde{\mathbf{x}}_n^i &= \mathbf{x}_n^i + \mathbb{K}_n(\mathbf{y}_n - \mathbb{H}_n\mathbf{x}_n^i),\\
\mathbb{K}_n &= \mathbb{P}_n\mathbb{H}_n^T\left(\mathbb{H}_n\mathbb{P}_n\mathbb{H}_n^T + \mathbb{R}_n\right)^{-1},\\
\tilde{\mathbb{P}}_n &= (\mathbb{I} - \mathbb{K}_n\mathbb{H}_n)\mathbb{P}_n.
\end{aligned} \tag{4}$$

In this paper, each kernel has the same covariance matrix $\mathbb{P}_n = h^2\mathbb{S}_X$. This particular choice is from density estimation [19] and is convenient for filtering as we see from Eq. 4 that each ensemble member is updated linearly proportional to the EnKF update. However, there are certain scenarios where a different choice

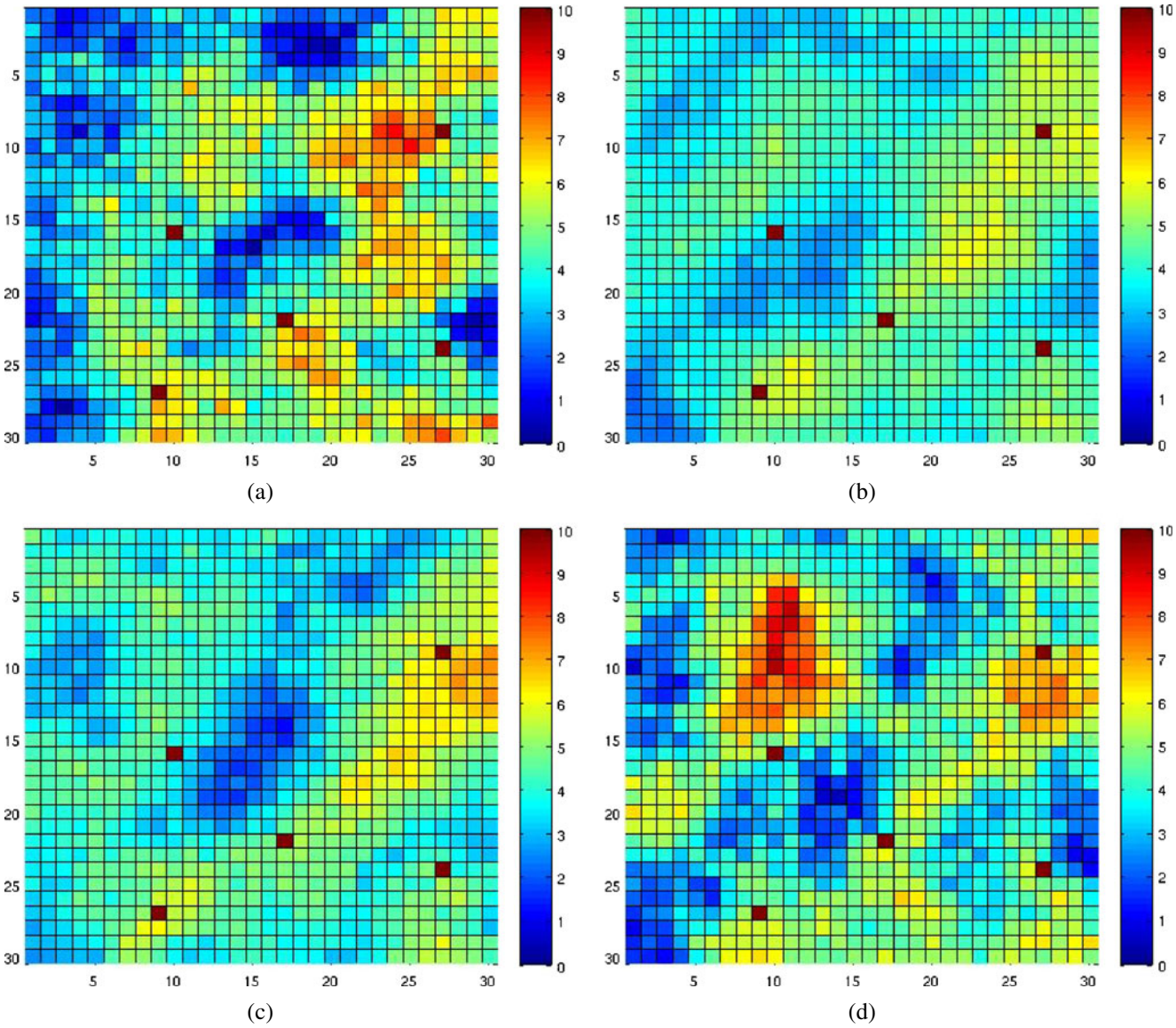

Fig. 6 Mean log permeability of the ensemble members. Final time step. **a** EnKF. **b** AGM $h = 0.1$. **c** AGM $h = 0.2$. **d** AGM $h = 0.3$

of $\mathbb{P}_n$ is more suitable (Stordal [22]). The weights are updated as

$$\bar{w}_n^i = w_{n-1}^i \Phi(\mathbf{y}_n - \mathbb{H}_n \mathbf{x}_n^i, \mathbb{H}_n \mathbb{P}_n \mathbb{H}_n^T + \mathbb{R}_n),$$

$$w_n^i = \frac{\bar{w}_n^i}{\sum_{j=1}^N \bar{w}_n^j},$$

$$w_{n,\alpha}^i = \alpha w_n^i + (1-\alpha) N^{-1},$$

$$\alpha = \frac{1}{\sum_{i=1}^N \left(w_n^i\right)^2} N^{-1},$$

where the next to last equation above is introduced to avoid a weight collapse, extending the methodology in Hoteit et al. [12]. With the new weights and the adaptive selection for α, the effective sample size N_{eff} [13] stays above 80% and a weight collapse is impossible by construction due to the identity

$$N_{\text{eff}}^{\alpha} = \frac{N^3}{N_{\text{eff}}(N - N_{\text{eff}}) + N^2}.$$

Table 1 Empirical correlation between true and estimated fields

Filter	Empirical correlation with true field
Initial	−0.0109
EnKF	0.0838
$h = 0.3$	0.2424
$h = 0.2$	0.3088
$h = 0.1$	0.1826

Table 2 History match for the different filters

Filter	History match
EnKF	13.2639
$h = 0.3$	6.9112
$h = 0.2$	8.3481
$h = 0.1$	38.3100

3 2D reservoir

In our first example, we consider a synthetic 2D reservoir model. The reservoir is a 30×30 grid cell oil water model where each grid block is $100 \times 200 \times 50$ m. There are four oil producers, P1, P2, P3, and P4, located at grid blocks (10, 16), (9, 27), (27, 9), and (27, 24) and one water injector, I1, at (17, 22). The unknown parameters are the log-permeability (k) and porosity (ϕ) in each grid cell giving us a total of 1,800 unknown parameters. The pressure and saturation fields are also updated during assimilation so the total number of unknowns is 3,600. The true log permeability and porosity fields and all the initial ensembles are constructed as Gaussian random fields with an isotropic spherical variogram. The correlation length is equal to ten grid blocks (corresponding to 2,000 m). The following set of parameters were used to generate the reference fields and the initial ensemble: $\mu_\phi = 0.2$, $\mu_{\ln k} = 4$, $\sigma^2_\phi = 0.05^2$, $\sigma^2_{\ln k} = 2$ and $\rho_{\ln k,\Phi} = 0.8.$, where μ is the mean, σ^2 is the variance, and ρ is the correlation. The realizations were not conditioned to any well data. In Fig. 1, the reference fields ((a) and (b)) along with the mean of the initial ensembles ((c) and (d)) are shown.

Included in the state vector is the permeability, porosity, saturation, and pressure fields; the water and

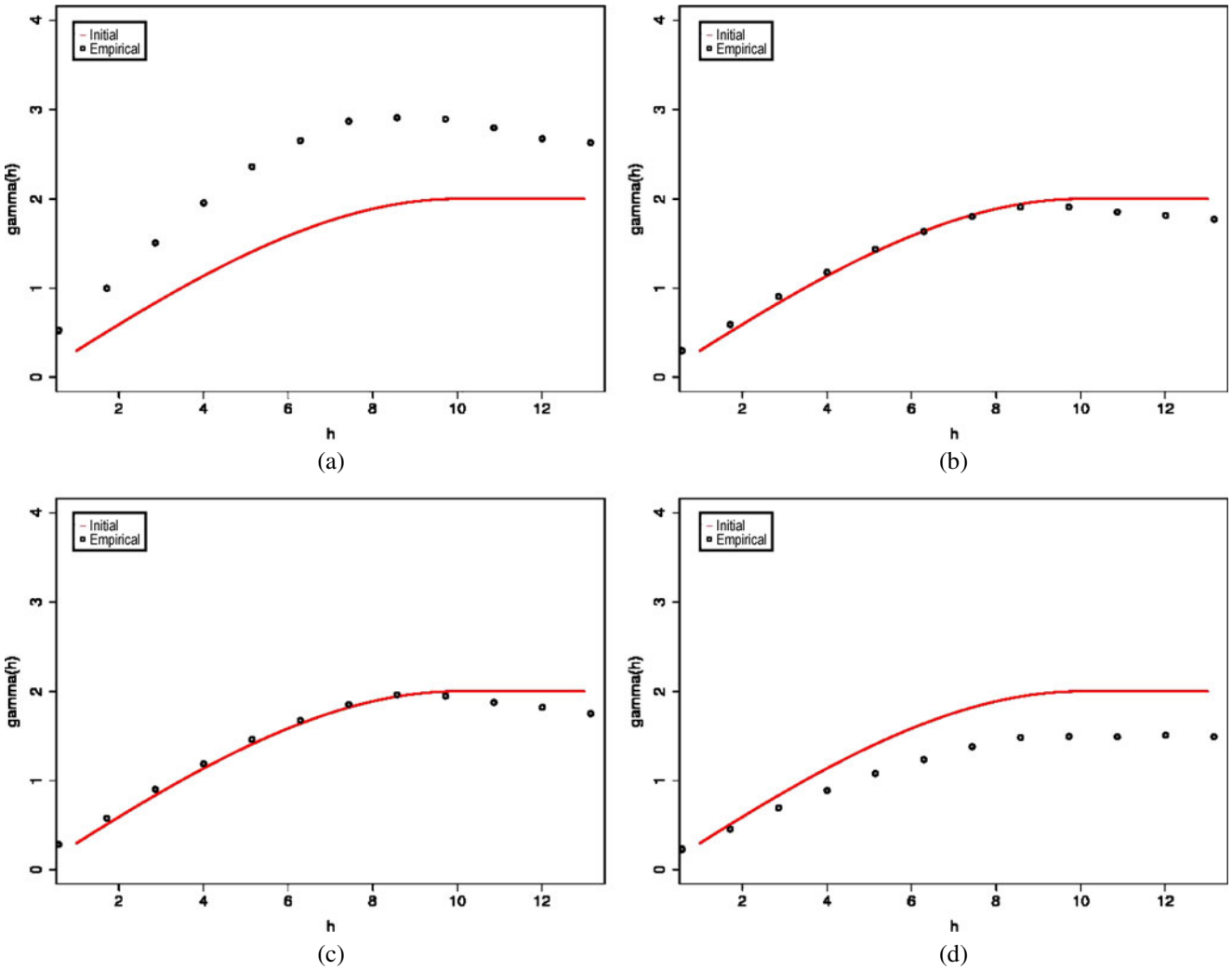

Fig. 7 Empirical semivariograms (*blue dots*) and initial semivariogram (*red line*). **a** EnKF. **b** AGM $h = 0.1$. **c** AGM $h = 0.2$. **d** AGM $h = 0.3$

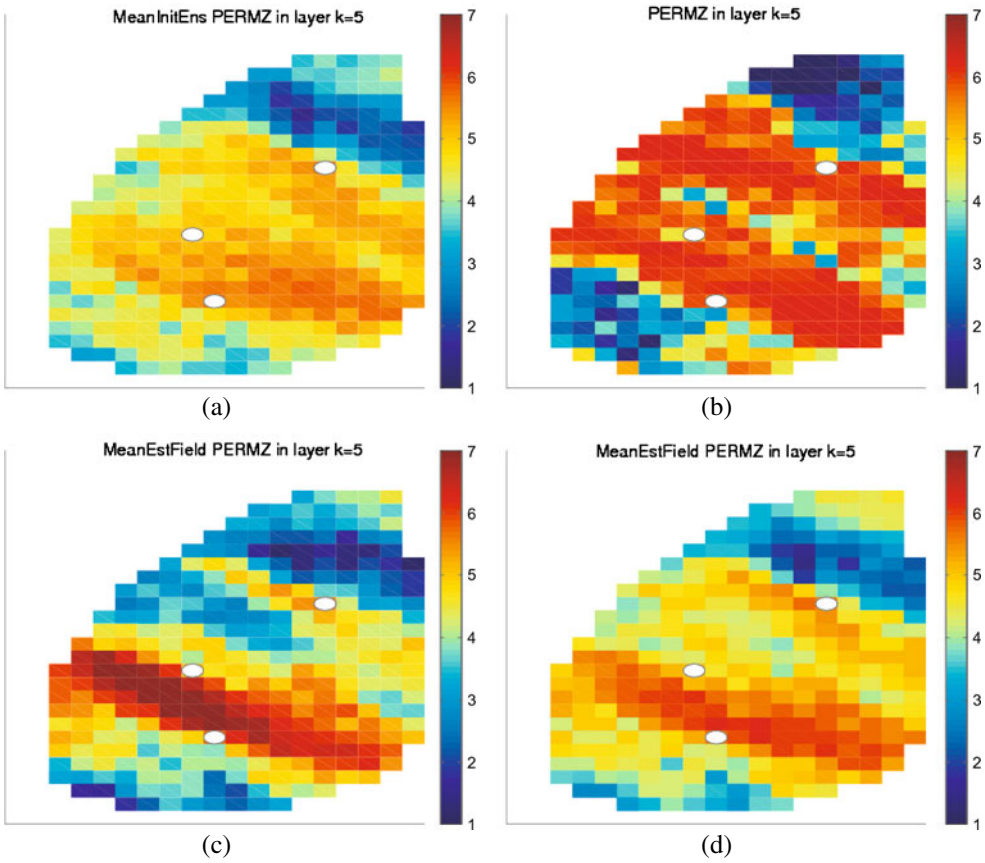

Fig. 8 Permeability fields for layer 5 in Punq-S3. **a** Initial mean of permZ layer 5. **b** True permZ layer 5. **c** EnKF mean of permZ layer 5. **d** AGM mean of permZ layer 5

oil production rate from the four producers; and the water injection rate for the injector. From the reference fields, we simulate measurements at every 20th day for 800 days by adding Gaussian noise to the production and injection rates with a relative error of 5% and regard this as the production history. We then apply EnKF and AGM for $h = 0.1, 0.2, 0.3$ to this set of measurements.

An important feature of AGM is the use of importance weights. Since the α parameter determines how much the weights influences the results, we plot the evolution of α and the new effective sample size for $h = 0.3$ in Fig. 2. Although α varies a lot, the main trend is that it increases with time which is a good sign as the higher value for α, the lower the bias of the filter. We also see that the new estimated effective sample size always stays above 80%.

The reruns of the final ensembles are shown in Figs. 3 and 4 where we plot the oil and water production rate in well P3. Although it is difficult to discuss the ensemble spread, not knowing what it looks like for a sample from the true posterior, it seems that ensemble collapse is at least not an issue for AGM with h equal to 0.1 and 0.2. We also remind the reader that a larger ensemble spread could be obtained for both AGM and EnKF using covariance localization/Kalman gain screening [2, 11, 16, 23].

The mean fields are plotted in Figs. 5 and 6. From these figures, it seems that AGM better captures the channel going from northeast to southwest in the reservoir. In order to quantify this, we compute the empirical correlation coefficient between each estimated permeability field and the true permeability field [20]. The final coefficient is taken as the weighted mean of the 100 coefficients for each of the four methods. The results are shown in Table 1, and we clearly see how AGM improves upon the EnKF with the highest correlation coefficient of 0.3 for $h = 0.2$. The filters are also compared based on the mean squared error mismatch between predicted data from a rerun and the measurements. A measure for the data mismatch, D, is given as the normalized mean squared difference between simulated data computed from a rerun of the final permeability and porosity fields and the observed data. That is

$$D = \sum_{i=1}^{N} w^i \parallel \mathbf{Y} - \mathbb{H}(\hat{\mathbf{x}}_t^i) \parallel_{\mathbb{R}}^2,$$

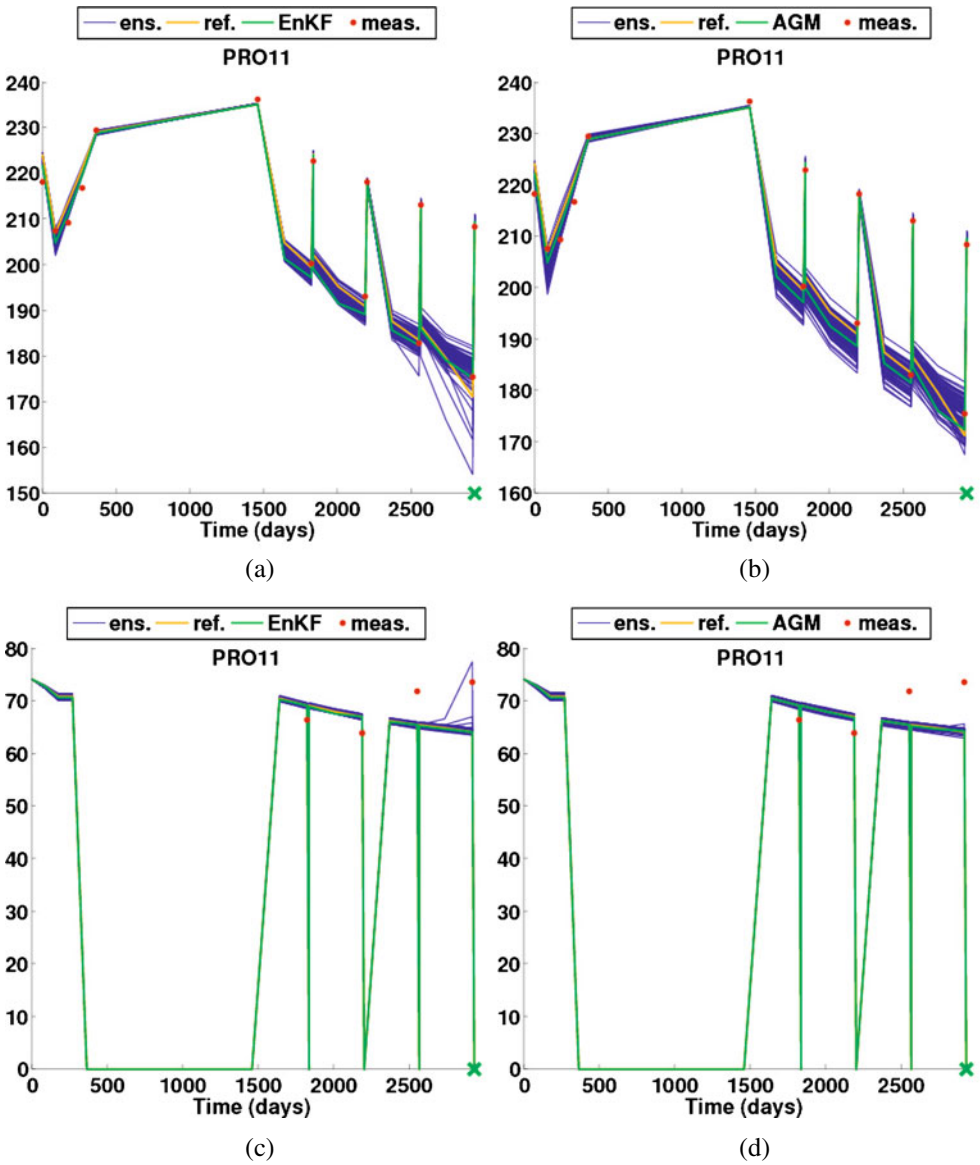

Fig. 9 History match after reruns from time zero bottomhole pressure and gas/oil ratio for producer 11 in the Punq-S3 model. The *green line* is the mean filter solution during assimilation, and the *blue lines* are the ensemble predictions from a rerun. **a** EnKF BHP. **b** AGM BHP. **c** EnKF GOR. **d** AGM GOR

$\mathbb{R}$ is the covariance of the measurement errors, N is the ensemble size, and $\mathbf{Y}$ is the vector of all observations. The squared norm $||\mathbf{u}||_{\mathbb{V}}^2$ in $\mathbf{R}^k$ is given by $\mathbf{u}^T\mathbb{V}^{-1}\mathbf{u}$. The distribution of D is not known in general, and it is therefore difficult to say what values of D are satisfactory. If a single best estimate was computed, we would expect the deviations to behave approximately as the measurement noise and D should be approximately one. However, if we want a sample from the posterior, we would also like samples in the tail of the distribution, which obviously would increase the average value of D. For prediction purposes, we claim that it is better to overestimate the posterior uncertainty than to under estimate it. The data mismatch value D is calculated for each filter, and the lowest result was obtained for the AGM with $h = 0.3$ with a value of 6.91. For the AGM with $h = 0.2$ and $h = 0.1$, the data mismatch was 8.35 and 38.31, respectively. The EnKF produced a data mismatch equal to 13.2639. The history match values are shown in Table 2.

We also want to see how much of the prior geostatistical structure that has been preserved in the posterior ensemble. In order to do so, we compute the empirical semivariogram for each of the estimated permeability fields and compare it with the initial semivariogram [10]. In Fig. 7, the empirical semivariograms are plotted against the theoretical semivariogram from which the true field and initial ensemble was generated. We clearly see the advantage with a small h since then AGM to a large extent preservers the prior

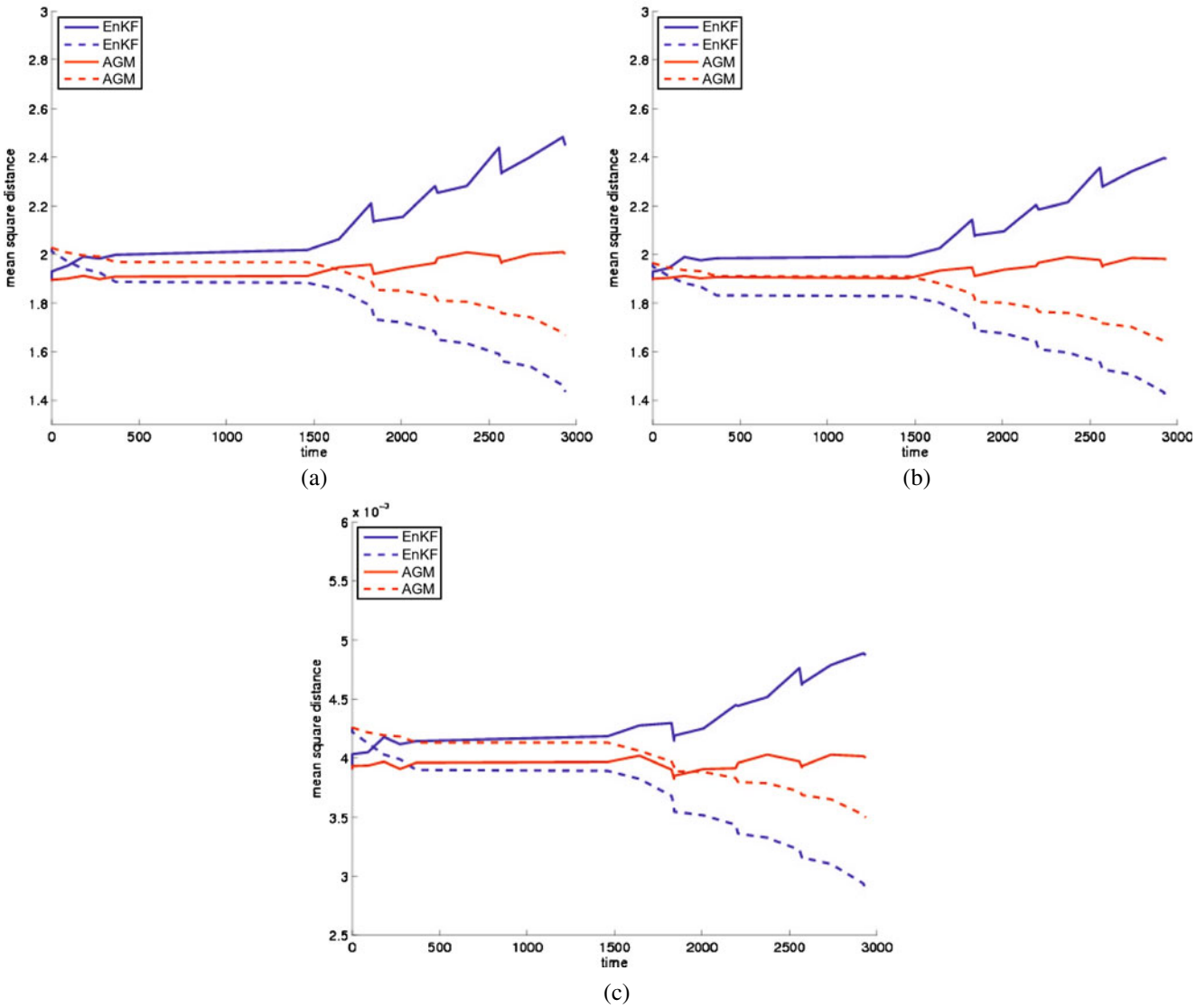

Fig. 10 Ensemble spread (*dashed line*) and distance from ensemble to reference fields. Mean over 10 runs with 100 members. **a** PermX 100 ensemble members. **b** PermZ 100 ensemble members. **c** Poro 100 ensemble members

geostatistics. Thus, the AGM is able to match the historical data without changing the prior geostatistical model as much as the EnKF.

4 The Punq-S3 test case

Our next test case is a synthetic 3D model known as the Punq-S3 [9, 17]. The model is based on a real field with five layers and a total of 1,761 active grid cells. There are six production wells but no injectors; thus, the oil production comes from primary depletion. The production schedule consists of two periods: First, there is an 8-year history match period consisting of well testing, a shut in period and production. Next, an 8.5-year prediction period follows consisting of production with shut in tests every year. Our state vector consists of porosity, permeability in x- and z-direction (PERMX and PERMZ), pressure, water and gas saturation, and the initial solution gas/oil ratio (RS). Twenty initial ensembles, 10 with 100 members and 10 with 200 members, are generated using the programs Sgsim and Sgcosim from the GSLIB. Each filter is then applied with all 20 different initial ensembles. In Fig. 8, the true permeability field (a) and the mean of one of the initial ensembles (b) are shown.

Since the filters are run several times, AGM is only used with one value of h. Theoretically, we want the

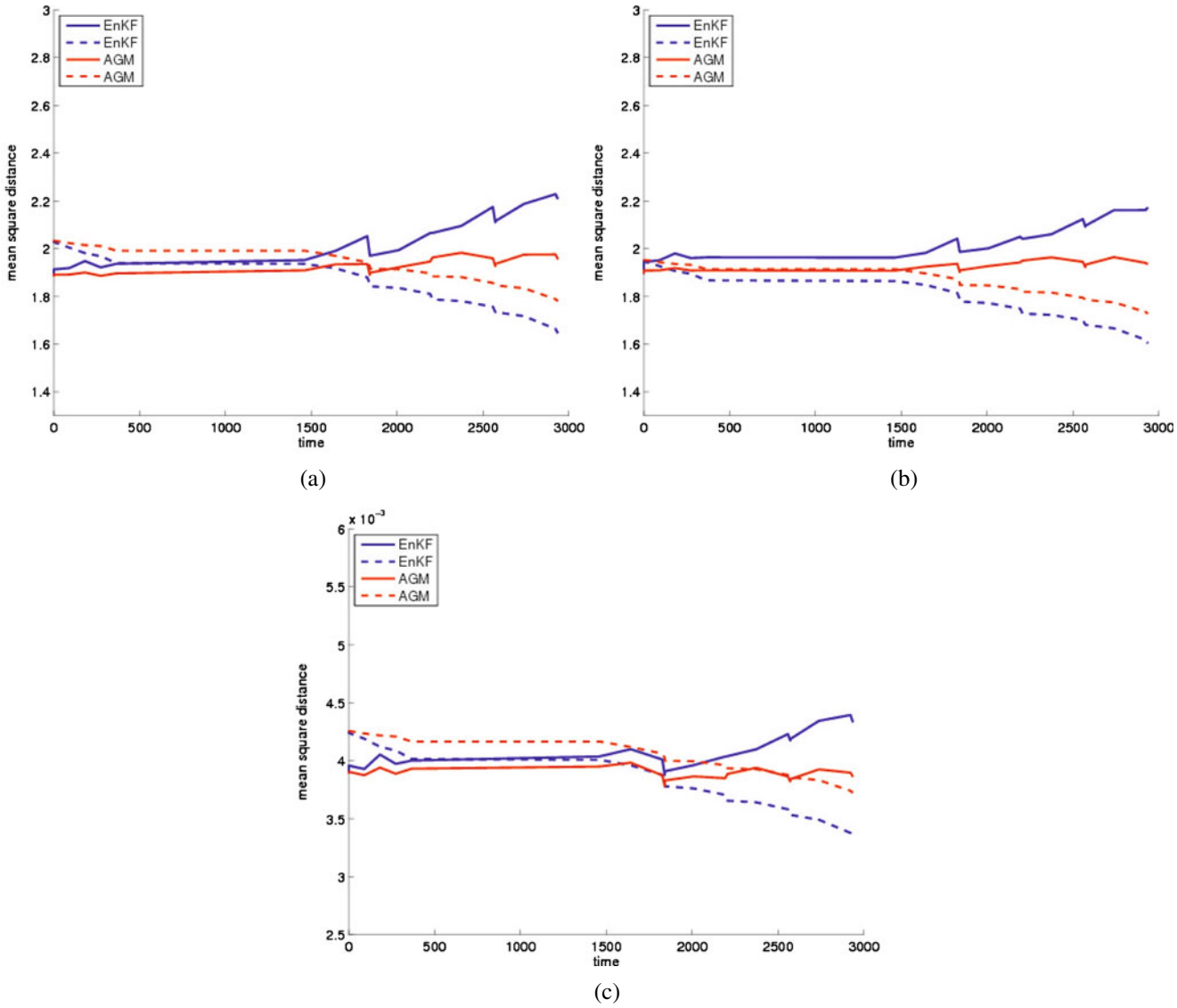

Fig. 11 Ensemble spread (*dashed line*) and distance from ensemble to reference fields. Mean over 10 runs with 200 members. **a** PermX 200 ensemble members. **b** PermZ 200 ensemble members. **c** Poro 200 ensemble members

value of h to be as small as possible, but at the same time, we would like the history match (data mismatch) to be approximately the same as for EnKF. After a trial-and-error approach, $h = 0.4$ was selected. For the 2D model, the value of 0.2 gives a better history match than EnKF, and it is not all clear to us why h has to be increased for the Punq-S3 model, but it might be a result of smaller prior uncertainty which makes the problem more linear. As the size of the Kalman gain is decided by h and the uncertainty in the ensemble, a smaller ensemble spread may force a larger h value in order to make significant updates that lead to a reasonable history match. Secondly, if the prior uncertainty is small, the nonlinearity of the model is not as transparent as if the prior uncertainty was large (as seen in the 2D model).

In the first study, we primarily focused the comparison on the history match and the empirical semivariograms within the final ensemble. In this study, in addition to history match, we focus on the spread within the ensemble, the distance from the prior ensemble to the posterior ensemble (in lack of a variogram), and the consistency between the predicted total cumulative oil production from the different initial ensembles. If no resampling is performed in AGM, we can calculate the distance from the initial to the final ensemble. That is how much each ensemble member has changed during the assimilation. Thus, AGM is run without resampling in this experiment.

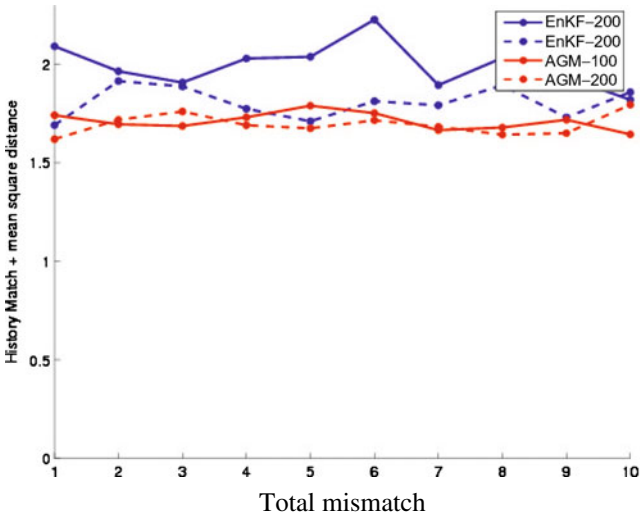

Fig. 12 Value of objective function (Eq. 5) for all runs

After the filters are used for estimation, the final ensemble of the static parameters is rerun throughout the history match period, and the traditional data mismatch is calculated. In addition, the norm between the initial ensemble and final ensemble for the parameters is calculated. Thus, for each of the filters, we calculate the objective function,

$$O = \sum_{i=1}^{N} \left(w^i (||\hat{\mathbf{x}}_i - \mathbf{x}_i^p||^2_{\mathbb{C}_p} + ||\mathbf{Y} - \mathbb{H}(\hat{\mathbf{x}}_i)||^2_{\mathbb{R}} \right), \tag{5}$$

where $\mathbb{C}_p$ is the prior covariance matrix, $\mathbb{R}$ is the covariance matrix of the measurement error, $\mathbf{Y}$ is the vector of all measurements, $\hat{\mathbf{x}}_i$ is the posterior ensemble member i, and $\mathbf{x}_i^p$ is the prior realization of ensemble i. If we were only interested in a single best estimate, we would minimize Eq. 5 where the last term is the data mismatch term and the first term can be viewed as a regularization. Figure 8c, d shows the estimated permeability in the z direction for layer 5 for one of the 100 membered ensemble. In Fig. 8, we see that the EnKF update has a higher variance than the AGM update, which is expected. The results shown in Fig. 8 is representative for all the 40 estimations performed. Figure 9 shows the history match, second part of Eq. 5, of the bottom hole pressure and the gas/oil ratio for one of the ensembles with 100 members, and the match with the initial ensemble. Reaching conclusions based on visualizing all the history match plots are hard, especially for studies like this where many ensembles are run for each method.

The performance of EnKF and AGM for the history match period, for all the 40 estimations, is summarized and compared in Figs. 10, 11, 12, and 13. The mean square difference between the true and estimated states is calculated as a function of time. Likewise is the ensemble spread, which is the mean square distance between the ensemble mean and the ensemble as a function of time. The results are shown in Figs. 10 and 11 where we plot the mean over the filter runs. The figures show that AGM has a larger ensemble spread than EnKF, and at the same time, it produces

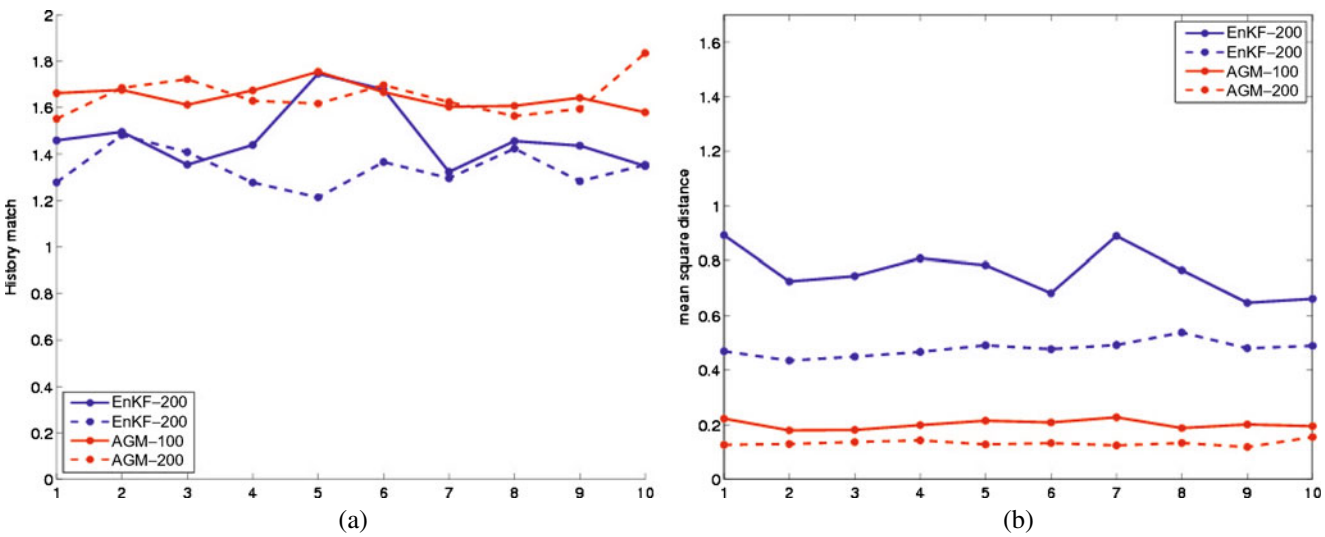

Fig. 13 Value of the two terms in the objective function for all runs. **a** Data mismatch. **b** Distance from prior to posterior

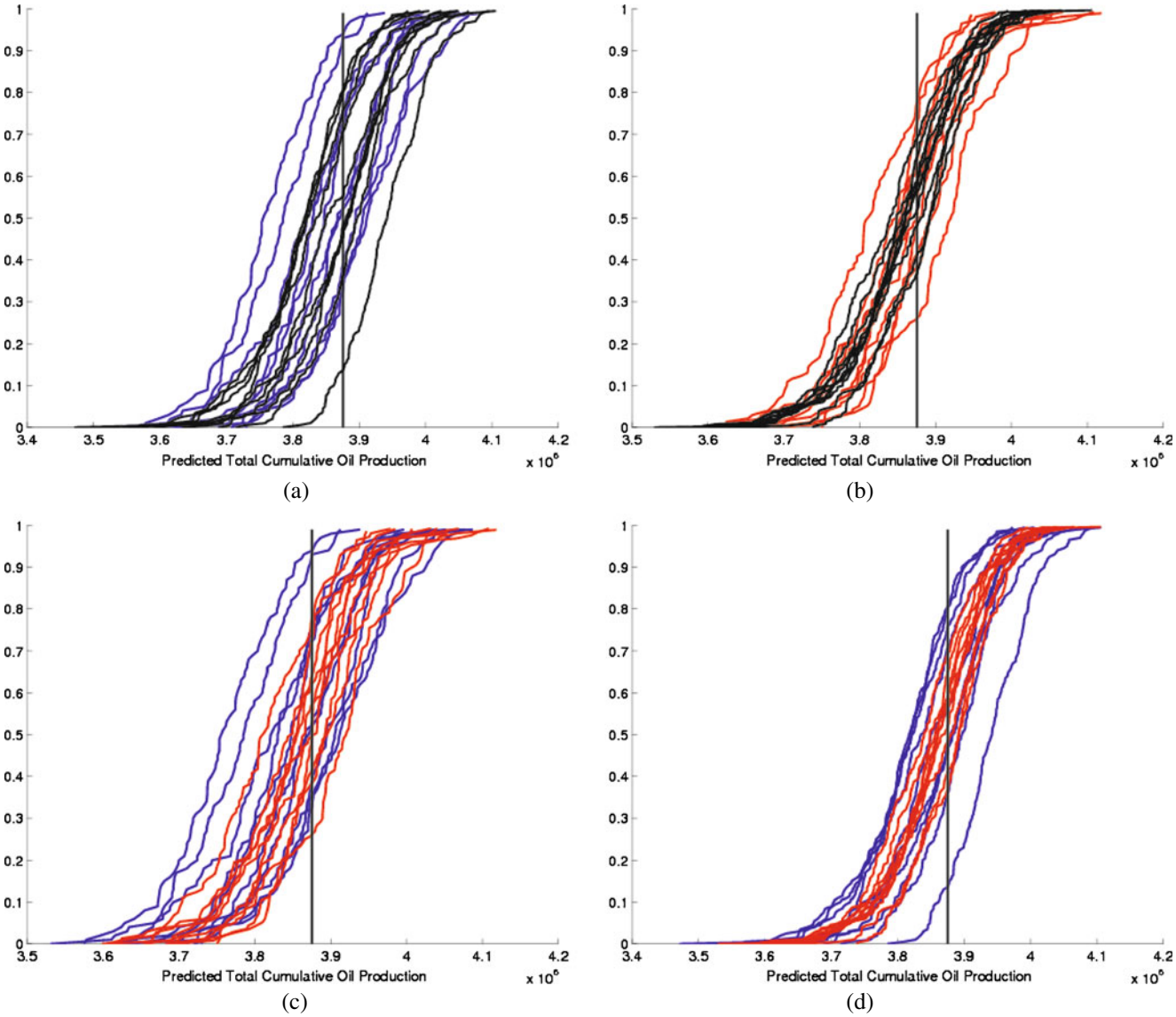

Fig. 14 Estimated cdfs of the total cumulative oil production. The *vertical black line* is the total cumulative oil production from the reference run. **a** Estimated cdfs from EnKF. *Blue lines* have 100 ensemble members; *black lines* have 200 ensemble members. **b** Estimated cdfs from AGM. *Red lines* have 100 ensemble members; *black lines* have 200 ensemble members. **c** Estimated cdfs with 100 ensemble members. *Blue lines* are EnKF; *red lines* are AGM. **d** Estimated cdfs with 200 ensemble members. *Blue lines* are EnKF; *red lines* are AGM

estimated fields that are closer to the true fields than EnKF. The results show that both EnKF and AGM improve when the ensemble size is increased from 100 to 200. Although EnKF gives a slightly lower history match, Fig. 12 shows that AGM gives lower values of the objective function both for 100 and 200 ensemble members. Thus, the ensemble members from AGM produce good history match without changing the prior ensemble as much as EnKF. We also see that there is less variation in the objective function values from AGM. The history match and the mean square distance from prior to posterior are shown in Fig. 13.

We now consider the second period of the Punq-S3 history, the prediction period. All the estimated fields obtained in the history matching period are run through both periods, and the total cumulative oil production is calculated and compared. Previous studies have shown that there seems to be some inconsistency in the estimated cdf with different initial ensembles in EnKF [14]. This is clearly undesirable as it indicates that we cannot trust the predictions if we run the model with only one initial ensemble. We investigate the spread of the cdf estimates from the filter solution obtained with different initial ensembles and also the effect of

increasing the ensemble size from 100 to 200. The results can be seen in Fig. 14. The spread of the estimated cdfs is smaller for AGM. Also, it seems that EnKF produce several clusters so that one may suspect that the estimated cdfs are not realizations from the same underlying cdf [14]. This surely implies that AGM is more consistent and that the solution is more reliable than the solution obtained from EnKF. Once again we see a clear improvement in the performances for both filters when the ensemble size is increased from 100 to 200.

5 Summary and conclusions

In this paper, we have applied the adaptive Gaussian mixture filter to reservoir models for comparison with the standard ensemble Kalman filter. AGM is designed to have better statistical properties than EnKF but at the same time avoid the problems with dimensionality that occur in standard SMC methods.

The first test case is a 2D model with oil and water. We run AGM three times with different values of h each time. For the two largest values of h, AGM produces a lower data mismatch than the standard EnKF, and the empirical variogram of the fields produced by AGM was close to the theoretical variogram of the prior; hence, the geostatistical structure of the prior is better preserved.

The second test case is the Punq-S3 3D model with oil, water, and gas. AGM is run with one value of h after a trial-and-error selection. Both EnKF and AGM are run with 20 different initial ensembles. The results from AGM has a smaller spread, especially the estimated cdfs of the total cumulative oil production. This indicates that the AGM is both more consistent and less sensitive to the initial ensemble compared to the standard EnKF. The estimated fields from AGM are closer to the true fields in mean square error while maintaining a larger spread within the ensemble. Although EnKF produces slightly smaller data mismatch, the values of the objective function are larger when we include the prior mismatch term.

Based on the comparisons we have made on the 2D model and the Punq-S3 model, we conclude that AGM performs better than the standard EnKF in terms of estimating the posterior distribution of the permeability and porosity fields given the production data. However, the drawback of AGM is that h is case sensitive. So far a trial-and-error approach has been used to select h. An automatic approach or ideally an adaptive h should be investigated in future work.

In both examples, the bandwidth h was selected ad hoc based on trial and error. Selecting h is a difficult task as the parameter depends on the number of ensemble members, the non-linearity/non-Gaussianity of the problem, the dimension of the problem, and the prior uncertainty/predictive quality of the initial ensemble among others. Ideally a different h at each time step should be sought, depending on the linearity of the measurement at the time and on the ensemble spread, maybe even different for each ensemble member, depending on whether or not it predicts the measurements well at that given time. The authors have made several attempts to find an automatic bandwidth selection. The problem is that those methods who have worked well for a given case did not work for a different case. The bandwidth selection therefor remains an open problem.

Neither EnKF nor AGM sample from the correct posterior distribution, however, we claim that AGM is a step toward a more rigorous approach for high-dimensional data assimilation.

Acknowledgements Stordal, Nævdal, and Valestrand acknowledge financial support from the Research Council of Norway (PETROMAKS) and industrial sponsors through the project "Reservoir characterization using ensemble Kalman filter." Stordal, Nævdal and Valestrand also acknowledge Schlumberger for academic eclipse lisences.

References

1. Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C., Vallès, B.: Ensemble Kalman filter in reservoir engineering—a review. SPE J. **14**(3), 393–412 (2009)
2. Anderson, J.L.: Exploring the need for localization in ensemble data assimilation using hierarchical ensemble filter. Physica D **230**, 99–111 (2007)
3. Apte, A., Hairer, M., Stuart, A.M., Voss, J.: Sampling the posterior: an approach to non-Gaussian data assimilation. Physica D **230**, 50–64 (2007)
4. Bengtsson, T., Snyder, C., Nychka, D.: Toward a nonlinear ensemble filter for high-dimensional systems. J. Geophys. Res. **108**, 35–45 (2003)
5. Bengtsson, T., Bickel, P., Li, B.: Curse-of-dimensionality revisited: collapse of particle filter in very large scale systems. Probab. Stat. **2**, 316–334 (2008)
6. Doucet, A., Godsill, S., Andrieu, C.: On sequential Monte Carlo sampling methods for Bayesian filtering. Stat. Comput. **10**, 197–208 (2000)
7. Doucet, A., de Freitas, N., Gordon, N.: Sequential Monte Carlo Methods in Practice. Springer, New York (2001)
8. Evensen, G.: Data Assimilation: The Ensemble Kalman Filter. Springer, New York (2007)
9. Floris, F.J.T., Bush, M.D., Cuypers, M., Roggero, F., Syversveen, A.R.: Methods for quantifying the uncertainty of production forecasts: a comparative study. Pet. Geosci. **7**, 87–96 (2001)
10. Fu, J.: A Markov Chain Monte Carlo Method for Inverse Stochastic Modeling and Uncertainty Assessment. Ph.D. thesis, Universidad Politecninca de Valencia, Spain (2007)
11. Furrer, R., Bengtsson, T.: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter vari-

ants. J. Multivar. Anal. **98**(2), 227–255 (2007). ISSN 0047-259X. doi:10.1016/j.jmva.2006.08.003
12. Hoteit, I., Pham, D.-T., Triantafyllou, G., Korres, G.: A new approximative solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography. Mon. Weather Rev. **136**, 317–334 (2008)
13. Kong, A., Liu, J., Wong, W.: Sequential imputations and Bayesian missing data problems. J. Am. Stat. Assoc. **89**(425), 278–288 (1994)
14. Lorentzen, R.J., Nævdal, G., Vallès, B., Berg, A.M., Grimstad, A.-A.: Analysis of the ensemble Kalman filter for estimation of permeability and porosity in reservoir models. In: SPE Annual Technical Conference and Exhibition, Dallas, Texas, 9–12 October 2005. Paper SPE96375 (2005)
15. Mandel, J., Beezley, J.D.: An ensemble Kalman-particle predictor–corrector filter for non-Gaussian data assimilation. In: ICCS 2009: Proceedings of the 9th International Conference on Computational Science, pp. 470–478. Springer, Berlin. ISBN 978-3-642-01972-2 (2009)
16. Myrseth, I., Omre, H.: Hierarchical ensemble Kalman filter. Spe J. **2** 569–580. doi:10.2118/125851-PA (2010)
17. PunqS3—website: http://www3.imperial.ac.uk/earthscienceandengineering/research/perm/punq-s3model (2011)
18. Robert, C.P., Casella, G.: Monte Carlo Statistical Methods. Springer, New York (2004)
19. Silverman, B.W.: Density Estimation for Statistics and Data Analysis. Chapman and Hall, London (1986)
20. Skjervheim, J.-A., Aanonsen, S.I., Ruud, B.O., Evensen, G., Johansen, T.A.: Estimating lithology and other reservoir properties from 4D seismic data using the ensemble Kalman filter. Included as a part of J.-A. Skjervheim, Continuous updating of a coupled reservoir-seismic model using an ensemble Kalman filter technique. Ph.D. thesis, University of Bergen, Norway (2007)
21. Stordal, A., Karlsen, H., Nævdal, G., Skaug, H., Vallès, B.: Bridging the ensemble Kalman filter and particle filters. Comput. Geosci. **15**, 293–305 (2011)
22. Stordal, A.S.: Sequential Data Assimilation in High Dimensional Nonlinear Systems. PhD thesis, University of Bergen (2011)
23. Zhang, Y., Oliver, D.S.: Improving the ensemble estimate of the Kalman gain by bootstrap sampling. Math. Geosci. **42**(3), 327–345. doi:10.1007/s11004-010-9267-8 (2010)