Computers & Geosciences 102 (2017) 34–44

Contents lists available at ScienceDirect

Computers & Geosciences

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

Research paper

An auxiliary adaptive Gaussian mixture filter applied to flowrate allocation using real data from a multiphase producer

Rolf J. Lorentzen[a,*], Andreas S. Stordal[a], Neal Hewitt[b]

[a] IRIS Energy, Thormølensgate 55, 5008 Bergen, Norway
[b] ENGIE E&P Norge AS, Vestre Svanholmen 6, 4313 Sandnes, Norway

ARTICLE INFO

Keywords:
Rate allocation
Field case
Nonlinear filtering
Data assimilation

ABSTRACT

Flowrate allocation in production wells is a complicated task, especially for multiphase flow combined with several reservoir zones and/or branches. The result depends heavily on the available production data, and the accuracy of these. In the application we show here, downhole pressure and temperature data are available, in addition to the total flowrates at the wellhead. The developed methodology inverts these observations to the fluid flowrates (oil, water and gas) that enters two production branches in a real full-scale producer. A major challenge is accurate estimation of flowrates during rapid variations in the well, e.g. due to choke adjustments. The Auxiliary Sequential Importance Resampling (ASIR) filter was developed to handle such challenges, by introducing an auxiliary step, where the particle weights are recomputed (second weighting step) based on how well the particles reproduce the observations. However, the ASIR filter suffers from large computational time when the number of unknown parameters increase. The Gaussian Mixture (GM) filter combines a linear update, with the particle filters ability to capture non-Gaussian behavior. This makes it possible to achieve good performance with fewer model evaluations. In this work we present a new filter which combines the ASIR filter and the Gaussian Mixture filter (denoted ASGM), and demonstrate improved estimation (compared to ASIR and GM filters) in cases with rapid parameter variations, while maintaining reasonable computational cost.

1. Introduction

Knowledge about downhole production flowrates is important for making optimal decisions when operating production and injection wells in a petroleum reservoir. The information can be used for adjusting wellhead chokes, gas lift rates, downhole inflow control valves (ICVs), etc. Flowrate allocation becomes especially important if the wells are connected to multiple reservoir zones, or if they have multiple branches, or both. Also estimation of wellhead flowrates are important if the well is not equipped with flowmeters, or if the accuracy of the flowmeters is diminished. A theoretical foundation for rate allocation based on a transient well flow model and the auxiliary sequential importance resampling (ASIR, (M Pitt and Shephard, 1999)) filter, is found in Lorentzen et al. (2014).

The rate allocation methodology is tested on two real full-scale wells in Lorentzen et al. (2016). For the first well, two flowrates are estimated for a one-phase gas producer. The well is equipped with two traditional pressure/temperature gauges, and 100 particles are used for the ASIR filter. The second well is a water injector, equipped with a distributed temperature sensor. Here nine flowrates are estimated, and 500 particles are used for the ASIR filter. Satisfactory results were obtained for these wells, but there is room for improvement when it comes to computational time, especially when the number of unknown parameters increases.

For online estimation of hidden Markov models there exist a vast literature on sequential Monte Carlo methods (see e.g. Doucet et al. (2001)). Although asymptotic optimality is well established (see e.g. Künsch (2005); Doucet et al. (2006); Moral and Jacod (2004)) the analysis is usually carried out using the Markov transition of the system as the proposal density. In practice, however, significant improvement with a finite sample size can be achieved using an optimal proposal (Doucet et al., 2000). Unfortunately, the optimal proposal (minimizing weight variance) is in general difficult to obtain with the exception of a system described by a linear measurement operator and Gaussian model and observation noise (see e.g. Ristic et al. (2004)). Although several improvements exist in some scenarios (see e.g. Carpenter et al. (1999); Doucet et al. (2000); Gustafsson et al. (2002)) we focus on the ASIR (M. Pitt and Shephard, 1999), and introduce a general transformed version with focus on a Gaussian mixture approximation under the assumption of Gaussian measurement noise, although the latter is

* Corresponding author.
E-mail address: rjl@iris.no (R.J. Lorentzen).

http://dx.doi.org/10.1016/j.cageo.2017.02.003
Received 8 July 2016; Received in revised form 31 January 2017; Accepted 4 February 2017
Available online 05 February 2017

not necessary. The algorithm we propose combines the two stage analysis scheme of the ASIR filter with a transformation of the particles base on a Gaussian mixture, although in theory the transformation can be more general. If a transformation can transport the particles into regions of higher posterior likelihood the new method will improve the diversity of the resampled particles. To our knowledge the proposed methodology is the first to combine Gaussian mixture approximations and a two-stage auxiliary filter.

The ASIR filter was introduced to improve the diversity of the particles (unique particles instead of multiple copies) without regularization (Musso et al., 2001), making it more robust when the measurements change rapidly. Although the ASIR filter may not improve the results when the model error is large, it usually outperforms the standard sequential importance resampling (SIR) filter when the model is quite accurate (Arulampalam et al., 2002). The ASIR filter is a two stage particle filter where the particles are resampled at time k using measurements at time $k+1$ before they are propagated forward in time and new weights are computed. The idea is that by looking one step ahead in time, one can avoid situations where a few particles are copied many times due to e.g. rapid changes in the model or outliers in the measurements. Another advantage when the model is complex and/or the dimension is large, is that the first step of the ASIR can be viewed as improving the proposal density.

In recent years many applications and theoretical studies of Gaussian Mixture filters have occurred (see e.g. Frei and Künsch (2013); Chen and Liu (2000); Stordal et al. (2011); Bengtsson et al. (2003); Kotecha and Djurić (2003); Crisan and Li (2015)). Although the Gaussian mixture filter can be viewed as a regularization of the SIR filter when the measurement noise is Gaussian (Stordal et al., 2011), it has the advantage of a Kalman type update step and has been successfully implemented on large scale models where particle filters collapse due to the curse of dimensionality (Bengtsson et al., 2008). For some recent applications see e.g. Liu et al. (2016); Hoteit et al. (2008); Stordal (2011), and for an extension to iterative versions for bias reduction see (Stordal and Lorentzen, 2011). A recent publication (Stordal et al.,) propose a stable particle filter in high dimension (d) at a computational cost of $O(Nd^2)$ which is simply too time consuming for many high dimensional or complex models. The auxiliary filter proposed here has a computational cost of $O(2N)$. Other related work is the probability hypothesis density filter (Vo and Ma, 2006) and the extension to linear Markov jump models (Pasha et al., 2009).

In Papadakis et al. (2010) and Mandel et al. (2009) the ensemble Kalman filter (EnKF, Evensen (2007)) is used directly as a proposal density. However the EnKF will only be efficient if the posterior is close to a Gaussian. The Gaussian mixture filter could also be used directly as a proposal density, but that would require $O(N^2)$ operations, although it may be reduced to $O(2N)$ operations in some cases, see Elgammal et al. (2003).

In the next section we give a description of the transformed particle filter including the ASIR filter. Then we give a brief overview of the Gaussian mixture filter, before introducing the auxiliary sequential Gaussian mixture (ASGM) filter. This is followed by two synthetic examples: a one dimensional state space model, the three dimensional Lorenz63 model (Lorenz, 1963), and a transient well flow model. The last example is a full-scale application where data from a multiphase producer is used. This well has two branches with ICVs, and we estimate oil, water and gas flowrates entering each branch from the reservoir. Finally, we present the conclusions from the work.

2. Method and theory

2.1. Auxiliary transform particle filter

We consider a state space model $(X_k, Y_k) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$, $k = 0, 1, \ldots$, where the Markov model (X_k) is defined by

$$X_k = \boldsymbol{\mathcal{M}}(X_{k-1}) + \eta_k, \quad k > 0, \tag{1}$$

where $\eta_k \sim \boldsymbol{\mathcal{N}}(0, \mathbb{M}_k)$, and $\boldsymbol{\mathcal{N}}$ represents a multivariate Gaussian distribution with given mean and covariance matrix. In addition, we assume that the initial distribution for X_0 is known. The measurement process Y_k is defined by

$$Y_k = \boldsymbol{\mathcal{H}}(X_k) + \epsilon_k, \quad k > 0, \tag{2}$$

where $\epsilon_k \sim \boldsymbol{\mathcal{N}}(0, \mathbb{W}_k)$. We assume that the transition kernels $\kappa_k(x, dx')$ induced by Eq. (1), admits a density with respect to the Lebesgue measure on $\mathbb{R}^{n_x}$. That is $\kappa_k(x, dx') = \kappa_k(x, x')dx'$, where $\kappa_k(x, x')$ represents a transition from x to x'. Further, we assume that the likelihood function $g(y_k|x_k)$ induced by Eq. (2) is known. For fixed x, $g(\cdot|x)$ is a density on $\mathbb{R}^{n_y}$. The posterior density π_k is then well defined and a recursive version of it is given by

$$\pi_k(x_k) \overset{\text{def}}{=} p(x_k|y_{0:k}) = Cg(y_k|x_k) \int \kappa_k(x_{k-1}, x_k)\pi_{k-1}(x_{k-1})dx_{k-1}, \tag{3}$$

where C is the normalizing constant. Although both $\boldsymbol{\mathcal{M}}$ and $\boldsymbol{\mathcal{H}}$ might depend on k, we omit the time index for notational convenience. Since (3) is not available analytically, an alternative is to approximate it using a set of weighted particles

$$\hat{\pi}_k(x_k) = \sum_{i=1}^{N} w_k^i \delta_{X_k^i}(x_k), \tag{4}$$

where δ is the Dirac delta function and $\sum_{i=1}^{N} w_k^i = 1$. If some of the weights in (4) become very large, it is natural to include a resampling step where particles are sampled with probability equal to the weight. Although this is an efficient way of removing poor particles (low weight), it could lead to problems of having multiple copies of just a few particles (low diversity of the particles). The ASIR filter was introduced to improve diversity of particles. The idea is to perform resampling at time $k-1$ using information from the measurement at time k, as is also the case for the optimal proposal distribution in the standard SIR filter (Doucet et al., 2006). The main idea is to 'look ahead' with each particle by computing some characteristic of next measurement, typically the mean or a sample. From these characteristics, it is possible to compute weights as in the standard particle filter. A resampling of the particles (not the characteristics) can then be performed based on these weights. The particles can then be propagated forward in time with the model noise, resulting in a more diverse set of particles. New weights can be computed based on the actual particles. Unless the model noise is very large, these particles should match the data well so that the weights should not vary too much and resampling should not be required. Mathematically speaking this is done by augmenting the state space with an integer variable, $i = 1 \ldots N$ representing the particle indices. The distribution of the new integer variables is simply defined by the particle weights. Hence $P(i = j) = w_k^j$ at time k. The distribution of the augmented state vector (X_k^i, i) can be updated by applying Bayes' rule to the augmented state

$$p(x_k, i|y_{0:k}) \propto g(y_k|x_k)\kappa_k(X_{k-1}^i, x_k)w_{k-1}^i, \tag{5}$$

where w_{k-1}^i is the importance weight of particle i at time $k-1$. By defining a characteristic, ν_k^i, of X_k^i given X_{k-1}^i (typically the mean or a sample), a proposal density q is constructed on $\mathbb{R}^{n_x} \times \mathbb{I}_N$, where $\mathbb{I}_N$ is the set $\{1, \ldots, N\}$, as

$$q(x_k, i|y_{0:k}) = q(x_k|i, y_{0:k})q(i|y_{0:k}) \overset{\text{def}}{=} \kappa_k(X_{k-1}^i, x_k)g(y_k|\nu_k^i)w_{k-1}^i. \tag{6}$$

A sample X_k^j is constructed by sampling an index i^j with probability proportional to $g(y_k|\nu_k^i)w_{k-1}^i$ and then sample from $\kappa_k(X_{k-1}^{i^j}, x_k)$. The importance weights are then given by the ratio of Eqs. (5) and (6) as

$$w_k^j \propto \frac{g(y_k|X_k^j)}{g(y_k|\nu_k^{i^j})}. \tag{7}$$

For more details see (M. Pitt and Shephard, 1999; Arulampalam et al., 2002). The original algorithm was proposed with an additional

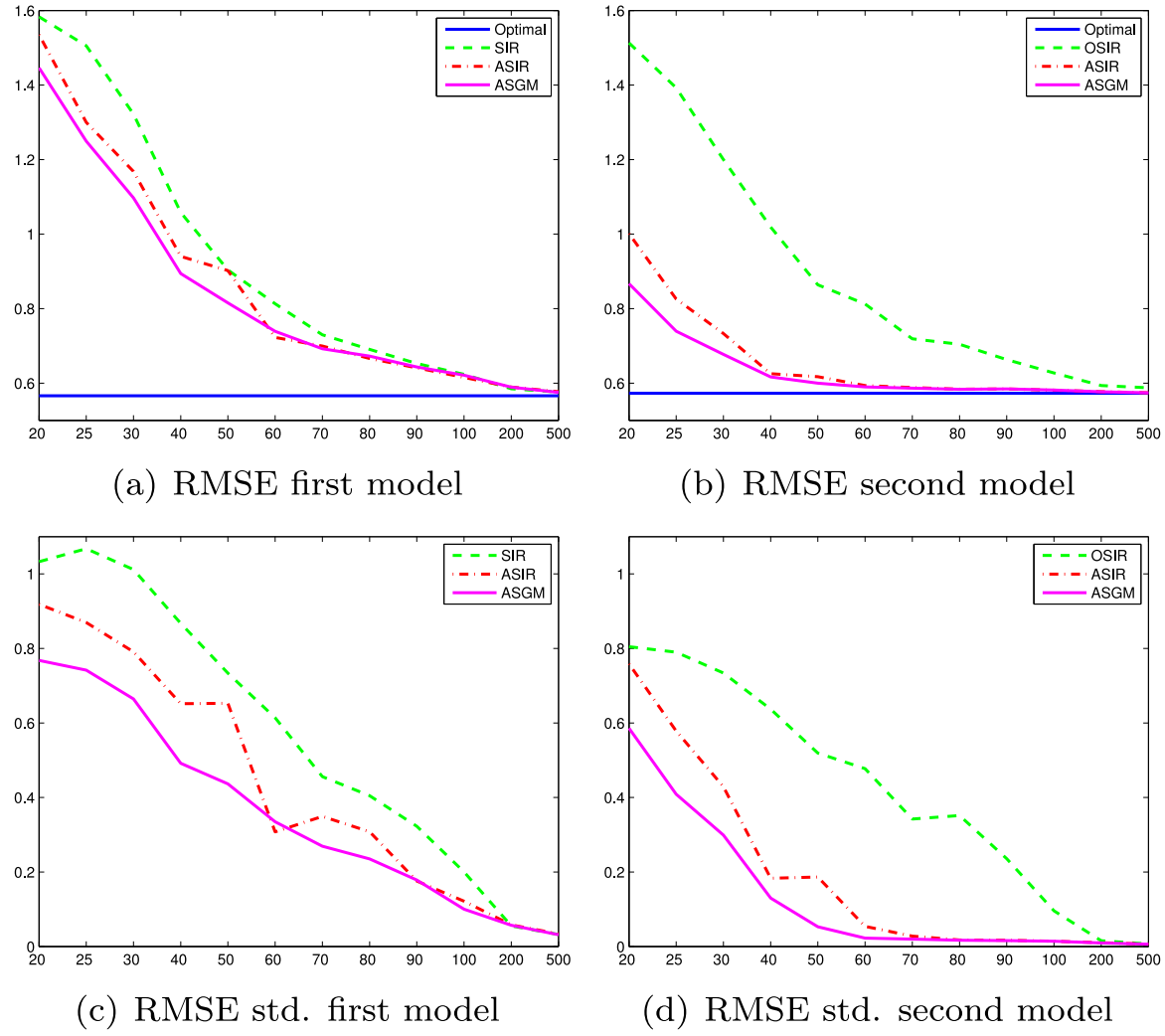

Fig. 1. Mean and standard deviation from 500 runs of the RMSE computed over 100 time steps. The sample size is given on the x-axis, and the root mean square error is given on the Y-axis, (a) RMSE second model, (b) RMSE std. first model, (c) RMSE std. second model.

resampling which can be disregarded (Fearnhead, 1998). Similar to optimal transport (see e.g. Reich (2013) and references therein) and implicit filtering (Chorin et al., 2010), we propose to add a transformation T: $\mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$. Since the particles in the SIR and ASIR filters are sampled or characterized by κ_k, it is the likelihood term g that causes a degeneracy or collapse of the filter. The idea behind the transformed filter is to balance between the two terms. Since the weights of the transformed filter contains both g and κ, it is possible to choose a transform that increase the value of g while reducing the value of κ. This is a key point that separates the new filter from SIR and ASIR. Instead of being a pure rejection algorithm (SIR, ASIR), it allows the particles to move closer to the observations without violating Bayes' law. The importance function on $\mathbb{R}^{n_x} \times \mathbb{I}_N$ is defined by

$$q_T(x_k, i|y_{0:k}) = q_T(x_k|i, y_{0:k}) q_T(i|y_{0:k}) \overset{\text{def}}{=} \kappa_k(T(X_{k-1}^i), x_k) g_T(y_k|\nu_k^i) w_{k-1}^i, \tag{8}$$

and the weights of the transformed ASIR filter are updated as

$$w_k^j \propto \frac{g(y_k|X_k^j)\kappa_k(X_{k-1}^{i^j}, X_k^j)}{g_T(y_k|\nu_k^{i^j})\kappa_k(T(X_{k-1}^{i^j}), X_k^j)}. \tag{9}$$

The likelihood function under the transformation, g_T, can in principle be chosen arbitrary but should reflect the transformation T. If T is the identity operator then $g_T = g$ and we regain the ASIR filter. The size of the transformation, $\| T(x) - x \|$, should depend on $\mathrm{Var}_{\kappa_k}(X)$ so that if the variance is small then $\| T(x) - x \|$ should be small to avoid a weight collapse. Ideally T should be chosen to minimize the variance of the weights in Eq. (9). Just like in the optimal SIR, one could look for a transformation T that minimizes the variance of the weights conditioned on X_{k-1}, or one could even try to minimize the unconditional variance for optimal performance. However, this would be a case specific and practically an impossible task (except in the linear Gaussian case). We therefore propose to define the transformation T using the theory of Gaussian mixtures since the Gaussian mixture filters can be viewed as a robust approximation of the SIR filter (Stordal et al., 2011).

2.2. Optimal importance function

The above transformed ASIR filter can be combined with the optimal SIR (OSIR) filter if $q_{\text{opt}} = p(x_k|x_{k-1}, y_k)$ is known. A particular case is when the model and observation noise is additive and the observation operator is linear, see (Doucet et al., 2000). After resampling and transformation, we sample from $p(x_k|T(x_{k-1}), y_k)$. The weights can then be evaluated as

$$w_k^j \propto \frac{g(y_k|T(X_{k-1}^{i^j}))\kappa_k(X_{k-1}^{i^j}, X_k^j)}{g_T(y_k|\nu_k^{i^j})\kappa_k(T(X_{k-1}^{i^j}), X_k^j)}. \tag{10}$$

However, since the optimal importance function is unknown in general, we propose to combine the transformed filter with the standard ASIR in the following.

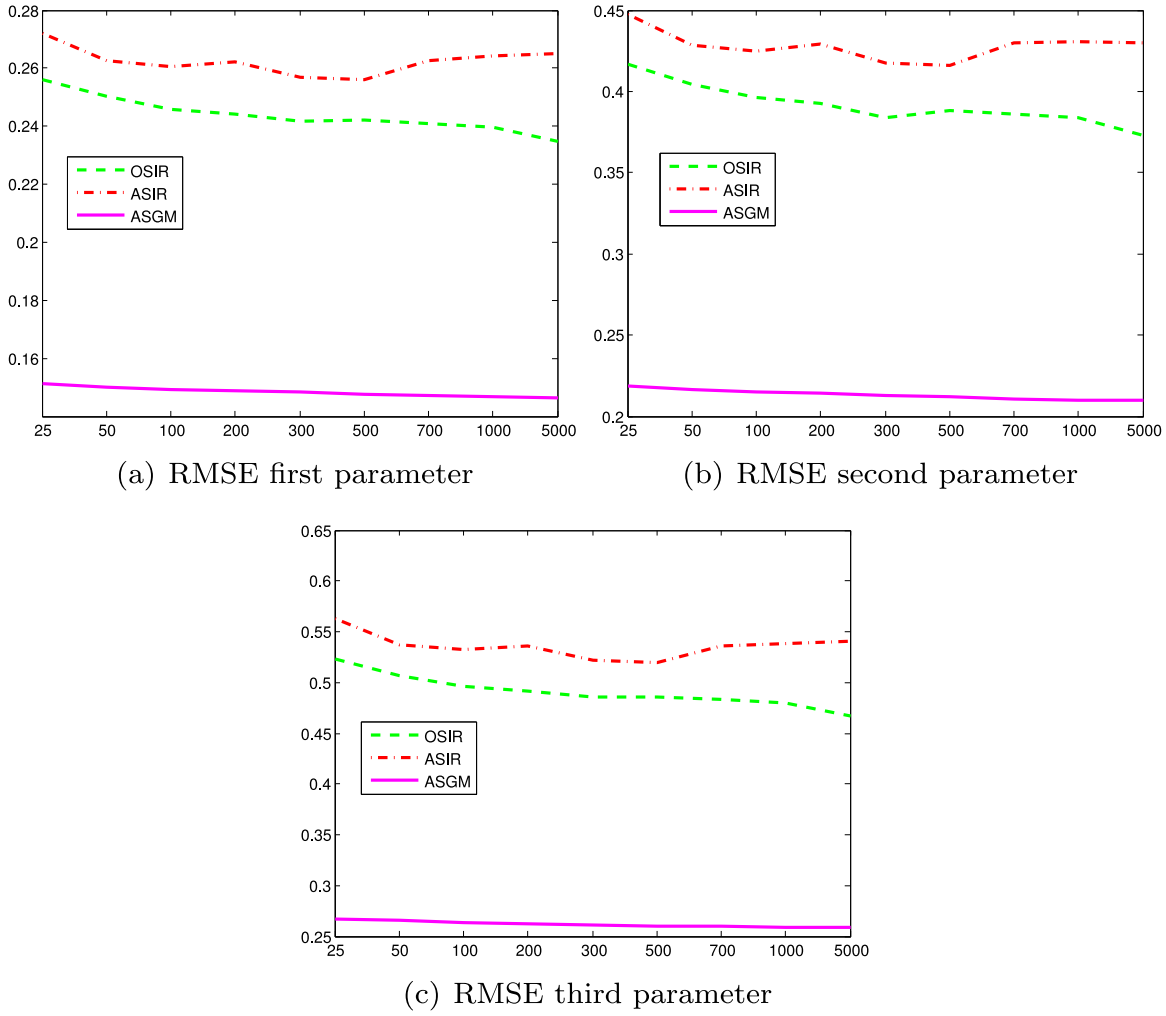

Fig. 2. Mean RMSE for 500 runs and 1000 timesteps. The sample size is given on the x-axis, (a) RMSE first parameter, (b) RMSE second parameter, (c) RMSE third parameter.

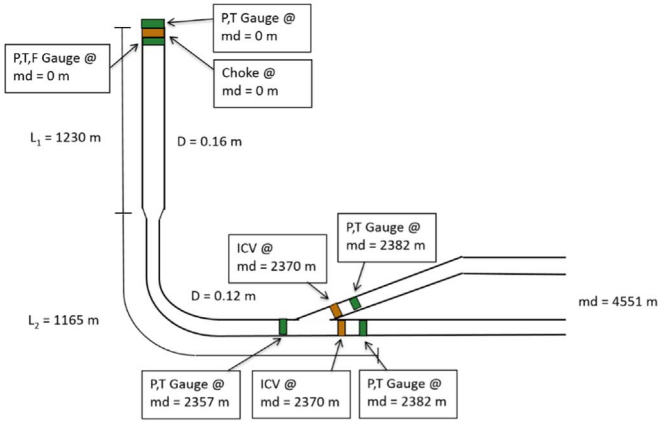

Fig. 3. Sketch of producer (not to scale). Note that measured depth is zero at seabed, approximately 380 m below the sea level.

2.3. Gaussian mixtures

A Gaussian mixture filter may, as the ASIR, be viewed as an approximation of the optimal SIR filter (Stordal et al., 2011) using a regularization prior to Bayes' theorem, rather than after as in the regularized SIR filter. The additional assumption is that the measurement error is Gaussian and additive. Assume that we have, at time $k-1$, a weighted sample $\{X_{k-1}^i, w_{k-1}^i\}_{i=1}^N$ approximating π_{k-1}. By sampling $X_k^i \sim \kappa_k(X_{k-1}^i, x_k)$ and augmenting the state vector with $\mathcal{H}(X_k^i)$ we may estimate the joint posterior density of $Z_k \overset{\text{def}}{=} [X_k, \mathcal{H}(X_k)]$ using a Gaussian kernel density estimator (Silverman, 1986) and Bayes' theorem (assuming a Gaussian likelihood function)

$$\pi_k^N(z_k) \propto \sum_{i=1}^N w_{k-1}^i \phi(z_k - Z_k^i | \mathbb{P}_k^i) \phi(y_k - \mathbb{H} z_k | \mathbb{W}), \tag{11}$$

where $\phi(\cdot|\Sigma)$ denotes a zero mean Gaussian density with covariance matrix Σ, $\mathbb{W}$ is the covariance of the measurement error and $\mathbb{H}$ is a binary matrix selecting $\mathcal{H}(X_k^i)$ from $Z_k{}^i$ and $\mathbb{P}_k^i$ is a positive definite matrix. Typically it is the sample covariance, which is independent of i. The expression in Eq. (11) may be rewritten as

$$\pi_k^N(z_k) \propto \sum_{i=1}^N \phi(z_k - \widetilde{Z}_k^i | \widetilde{\mathbb{P}}_k^i) w_{k-1}^i \phi(y_k - \mathbb{H} Z_k^i | \mathbb{H} \mathbb{P}_k^i \mathbb{H}^T + \mathbb{W}), \tag{12}$$

where

$$\widetilde{Z}_k^i = Z_k^i + \mathbb{K}_k^i (y_k - \mathbb{H} Z_k^i), \ \widetilde{\mathbb{P}}_k^i = (\mathbb{I} - \mathbb{K}_k^i \mathbb{H}) \mathbb{P}_k^i, \ \mathbb{K}_k^i = \mathbb{P}_k^i \mathbb{H}^T (\mathbb{H} \mathbb{P}_k^i \mathbb{H}^T + \mathbb{W})^{-1}. \tag{13}$$

For more details on Gaussian Mixture filters see e.g. Chen and Liu (2000) and Kotecha and Djurić (2003).

2.4. Transform via Gaussian mixture

The ASGM algorithm combines the ASIR with a Gaussian mixture as follows. A key point is that we want to apply the Gaussian mixture to

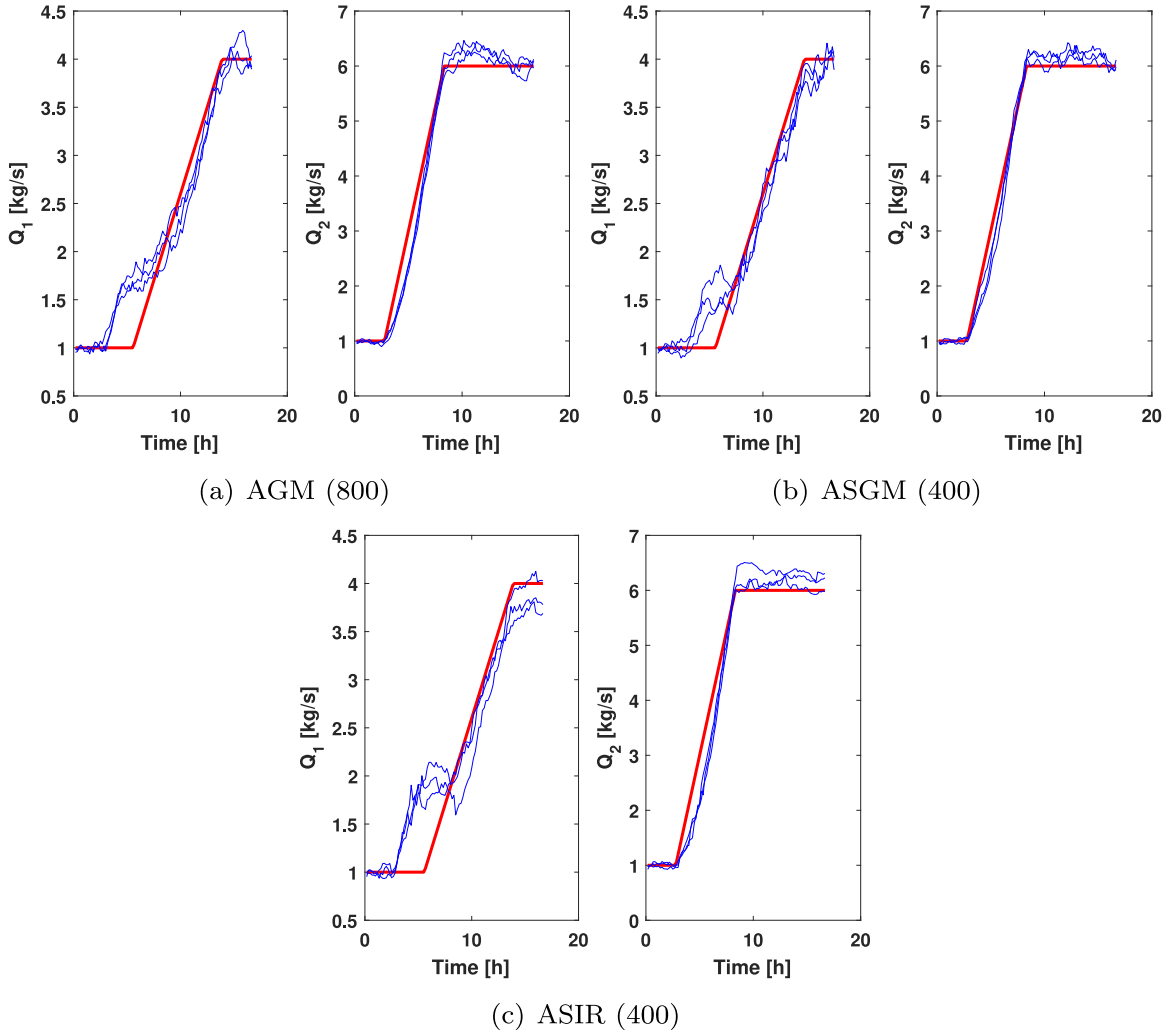

Fig. 4. Estimated flowrates (Q_1 and Q_2) for each simulation run as function of time. The true gas flowrates are shown in red, and the estimated flowrates are shown as blue curves, (a) ASGM (400), (b) ASIR (400). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

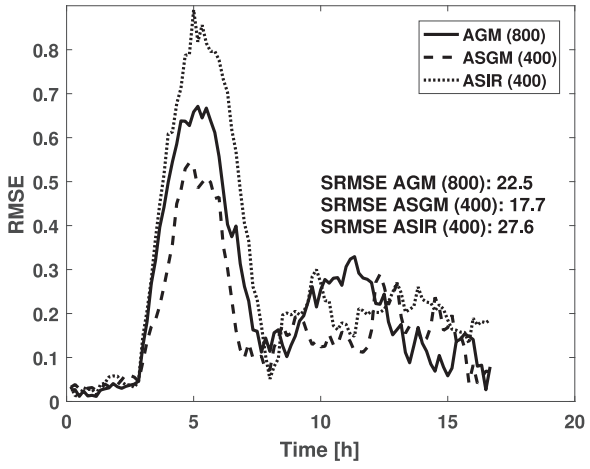

Fig. 5. RMSE for AGM (800), ASGM (400), and ASIR (400).

X_{k-1} instead of X_k in the first step of the algorithm and then compute the weights as described in (9) for the second part of the algorithm. First the characteristic ν_k^i are computed/sampled given X_{k-1}^i. Here we use the model prediction $\nu_k = \mathcal{M}(x_{k-1})$. That is, we propagate the particles forward in time using the model, but without adding the model noise. The next step is to use the update equations (13) with $Z_k^i = [X_{k-1}^i, \mathcal{H}(\nu_k^i)]$ as a rough approximation of $p(z_{k-1}|y_{1:k})$. Note the difference with (11) where we approximated $\pi(z_k) = p(z_k|y_{1:k})$. The first part of the updated vector, $\widetilde{Z}_{k-1}^i$ ((13) with k replaced by $k-1$), is given by $\widetilde{X}_{k-1}^i$, which defines our transformation $T(X_{k-1})$. Finally, we perform resampling (here we use multinomial sampling, see e.g. Arulampalam et al. (2002)) at time $k-1$ using information from the measurement one step ahead. A set of integers i^j is sampled with probabilities $P(i^j = i) \propto w_{k-1}^i \phi(y_k - \mathbb{H}Z_{k-1}^i | \mathbb{H}\mathbb{P}_{k-1}^i \mathbb{H}^T + \mathbb{W})$, where the last part is obtained from (13) and corresponds to g_T in the transformed ASIR algorithm. For each i^j, $j = 1,\ldots,N$ a new sample X_k^j is sampled from $\kappa_k(\widetilde{X}_{k-1}^{i^j}, x_k)$ and the ASGM weights are computed as

$$w_k^j \propto \frac{g(y_k|X_k^j)\kappa_k(X_{k-1}^{i^j}, X_k^j)}{\phi(y_k - \mathbb{H}Z_{k-1}^{i^j}|\mathbb{H}\mathbb{P}_{k-1}^i\mathbb{H}^T + \mathbb{W})\kappa_k(\widetilde{X}_{k-1}^{i^j}, X_k^j)}. \tag{14}$$

Here, $\widetilde{X}_{k-1}^{i^j}$ is the first part of the updated state vector $\widetilde{Z}^{i^j}$ after resampling. This filter can be formulated as the transform filter defined above with the affine transformation $T(x)$ given by (13). In the following examples, $\mathbb{P}_k$ in (11), is defined as the sample covariance matrix of $\{Z_k^i\}_{i=1}^N$.

The ASGM distinguishes itself from the ASIR filter in two ways. First, the linear update step that aims at moving the particles closer to the measurements, and second, the weights that are computed after the auxiliary step contains the transition density κ_k, contrary to the ASIR. The first may improve upon the ASIR filter if the measurement noise is small. The second step might deteriorate the behavior of ASGM if the

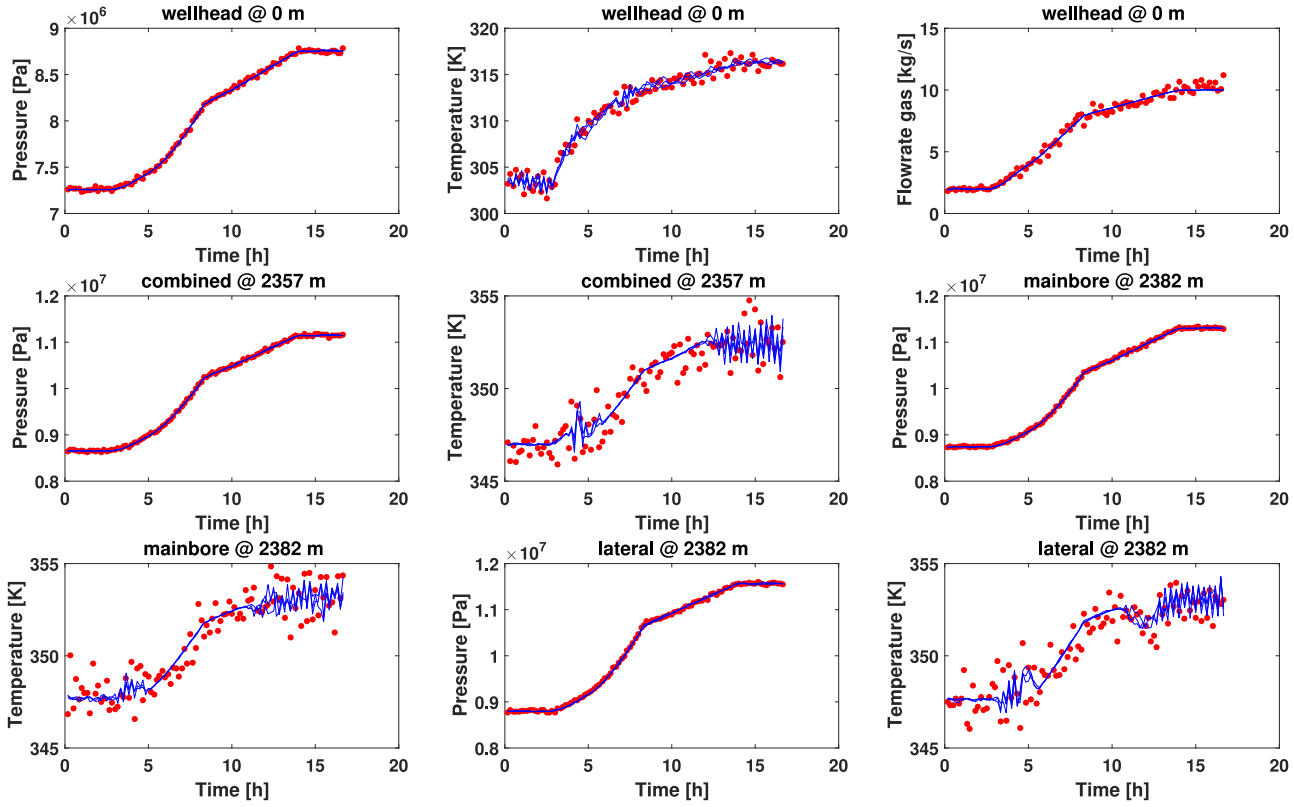

Fig. 6. Estimated (blue) and synthetic observations (red) using AGM (800). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

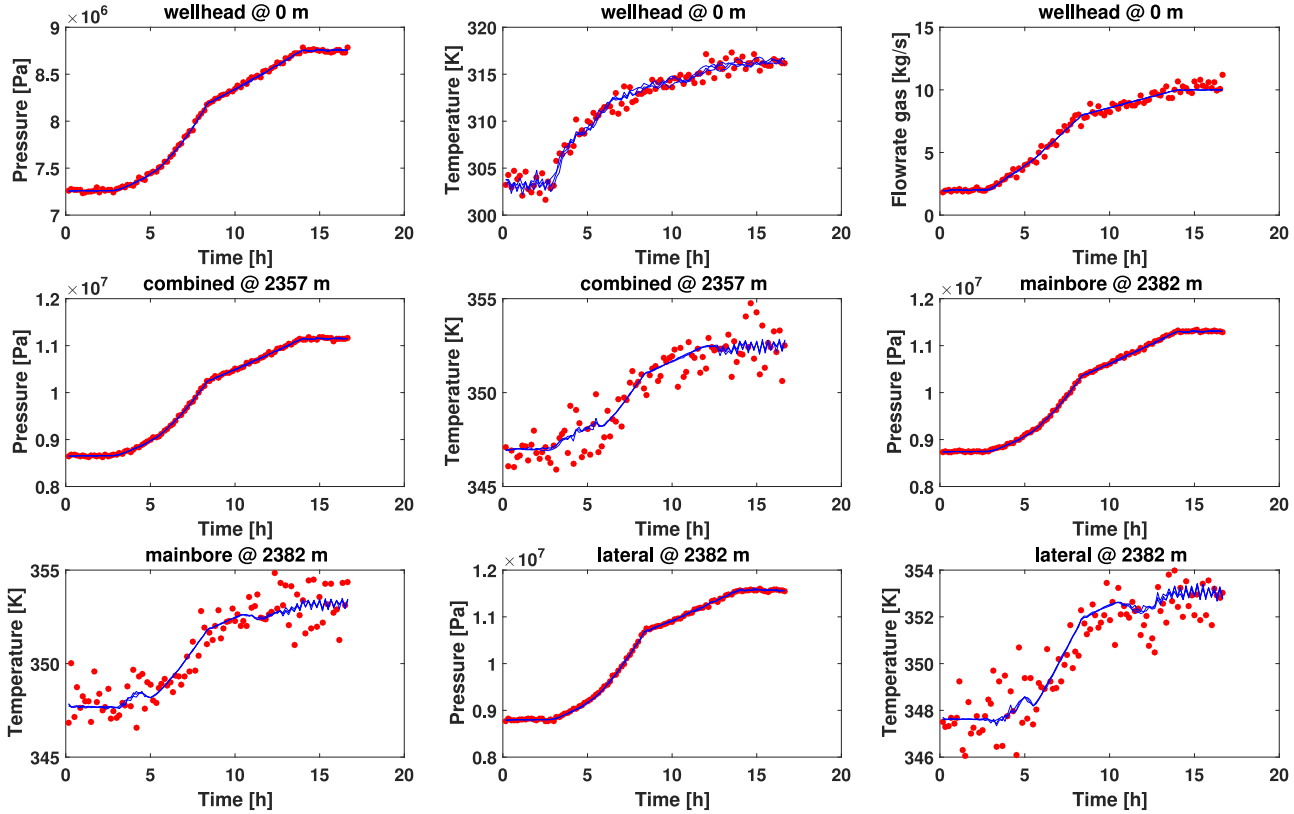

Fig. 7. Estimated (blue) and real observations (red) using ASGM (400). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

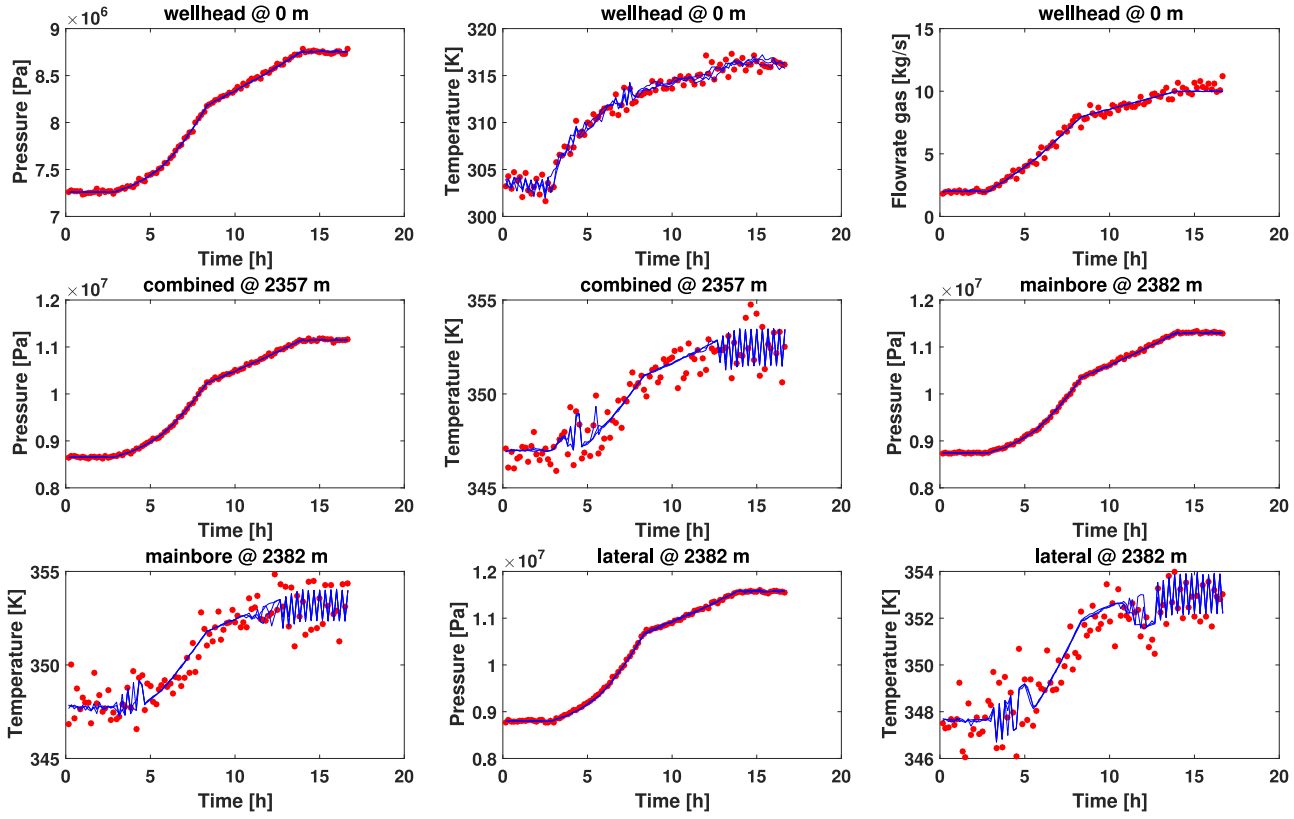

Fig. 8. Estimated (blue) and real observations (red) using ASIR (400). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

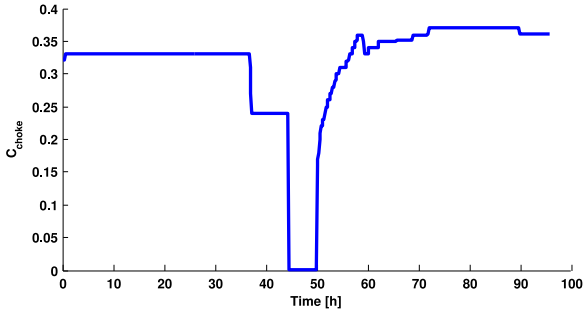

Fig. 9. Choke opening shown as fraction of maximum diameter.

model error is small, and it is therefore important to balance between the size of the update step (using the bandwidth parameter h described in Stordal et al. (2011)), and the model error. Further, there should also be a balance between the auxiliary weights given by the characteristic, ν_k, and the size of the update. This is automatically achieved by the Gaussian mixture filter as it balances the size of the update and the sharpness of the weight function via the bandwidth parameter h. In the examples presented below, the bandwidth parameter h is selected based on experience. A thorough analysis of the filter performance for different values of the bandwidth parameter, and the relation to different values for the measurement error and model error, is an interesting and important task (see e.g Li et al. (2016)). A further analysis is however beyond the scope of this paper, and we refer to future work for this task.

Further improvements in nonlinear systems can be achieved with local Kalman updates (Stordal et al., 2012) which, however, is more time consuming as the Kalman gain has to be calculated locally for each particle.

3. Results

In this section we present two synthetic examples, and one example using real observations. First we study a simple one dimensional nonlinear state space model, and compare ASGM, ASIR (M Pitt and Shephard, 1999), SIR (Gordon, 1993), and OSIR (Doucet et al., 2000). Then we move to the Lorenz63 model and compare ASGM, ASIR, and OSIR. Next we consider a full-scale production well, and compare ASGM, ASIR, and the adaptive Gaussian mixture (AGM) filter, using synthetic generated measurements. The AGM filter is introduced in Stordal et al. (2011) as a hybrid between the ensemble Kalman filter (EnKF, Evensen (2004)) and traditional particle filters. Finally, we run the ASGM using real observations, and present estimated flowrates and compare real and simulated measurements. The results are also compared to the estimated flowrates using the AGM filter, even though we do not have the actual (true) flowrates available. The experiment demonstrates a solution to a real rate allocation problem, and shows that the ASGM computes stable and reliable results using sensor data from complex systems.

3.1. One dimensional state space model

The model is described by

$$X_0 \sim \mathcal{N}(1,\,1),\; X_k = \frac{1}{2}X_{k-1} + \frac{25X_{k-1}}{1+X_{k-1}^2} + 8\cos(1.2k) + \eta_k, \quad k > 0, \tag{15}$$

where η_k is a zero mean Gaussian variable with standard deviation of 0.5 for all $k > 0$. We consider two models for the measurement process

$$Y_k = \mathcal{H}_j(X_k) + \epsilon_k, \quad k > 0,\; j = 1,\,2, \tag{16}$$

where

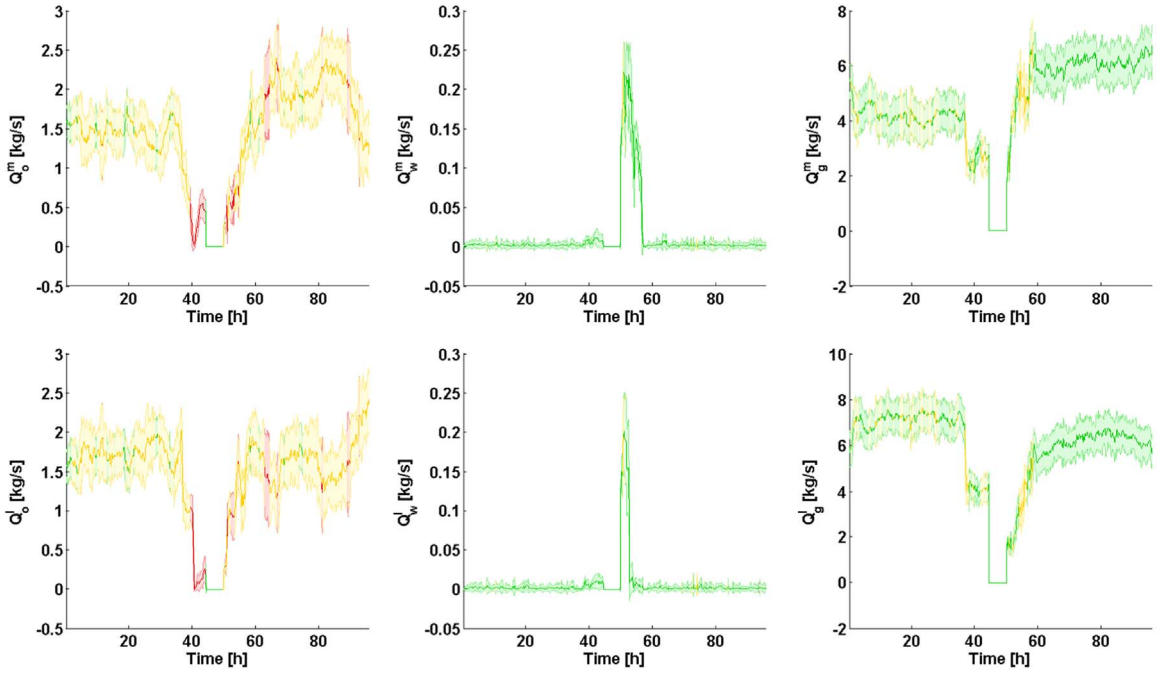

Fig. 10. Estimated flowrates in mainbore (m) and lateral (l). The color code re present reliability of the estimated parameters (green=Good, orange=Fair, red=Bad). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

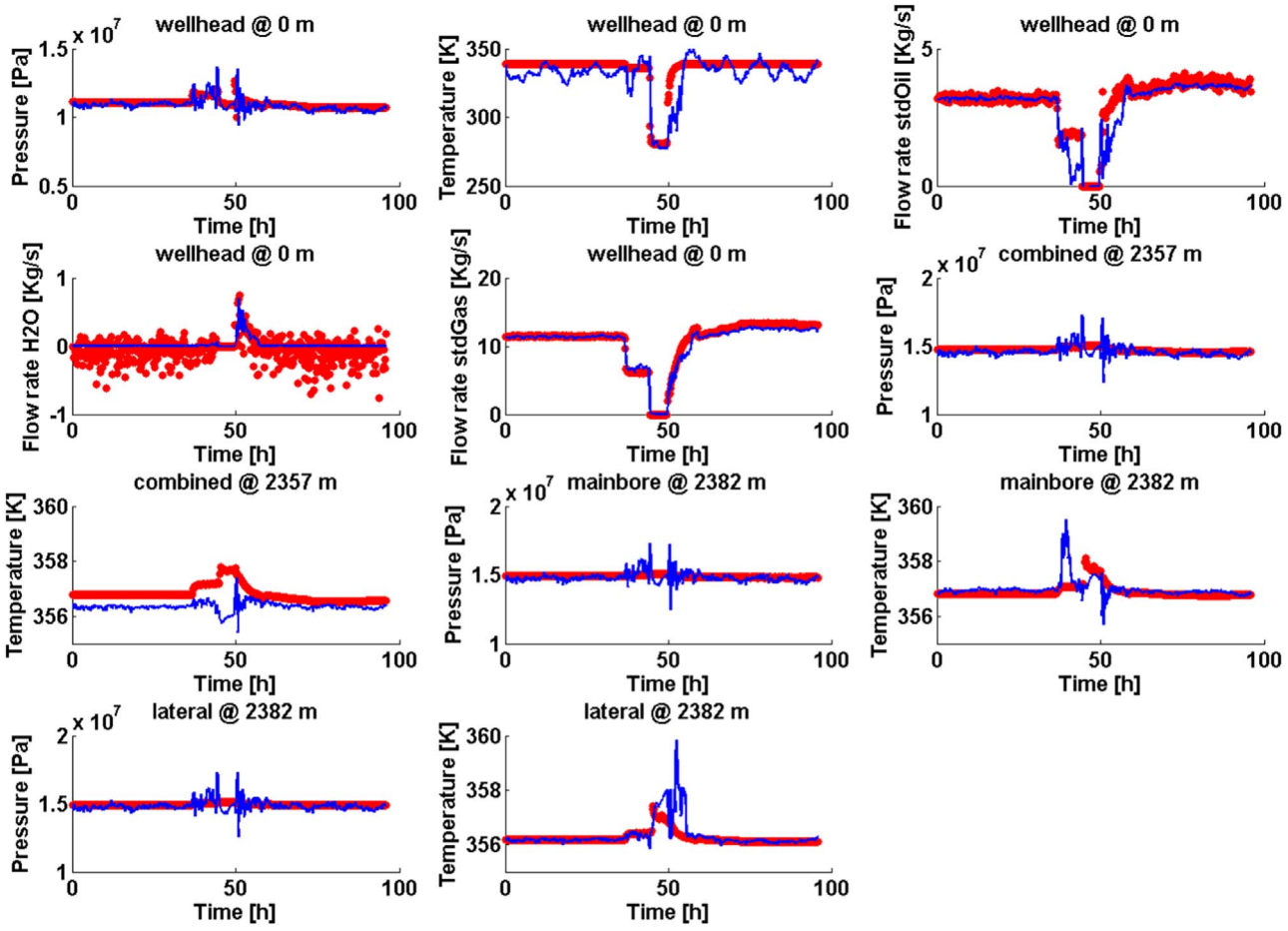

Fig. 11. Estimated (blue) and real observations (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

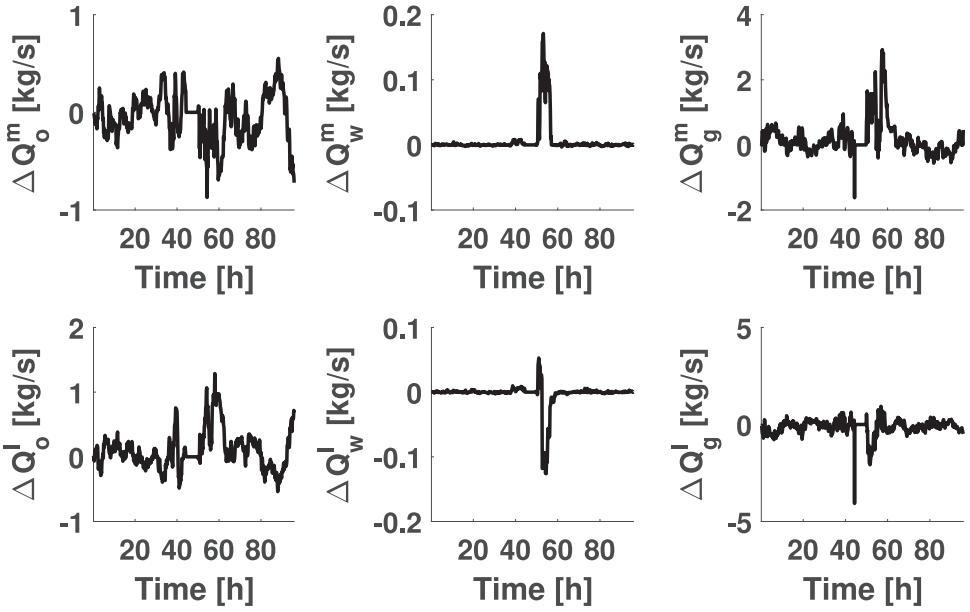

Fig. 12. Estimated flowrates using ASGM minus estimated flowrates using AGM.

$$\mathcal{H}_1(x) = \frac{x^3}{90}, \quad \mathcal{H}_2(x) = x, \tag{17}$$

and ϵ_k is a standard Gaussian variable for all $k > 0$. We run the experiment with bandwidth parameter h fixed at 0.1 and varying sample size. For the first model we compare it with the ASIR and SIR filter while in the second model we compare it with the ASIR and the optimal SIR filter since the measurement operator is linear in this case. For both models a SIR filter with 5×10^5 particles is regarded as the true optimal solution to the filter problem. The experiments are repeated five hundred times and the mean root square error over 100 time steps and 500 runs are reported together with the standard deviation over the five hundred runs. We clearly see from Fig. 1 the improvement of the new proposed ASGM method where the x-axis indicates the sample size. We stress that when the sample size of ASIR and ASGM is N, the sample size of SIR and OSIR is $2N$ since the ASIR and ASGM require twice evaluation of the likelihood function and may require twice evaluation of the model for each time step. For this experiment N ranges from 20 to 500 and the RMSE for the mean and standard deviation compared with the optimal solution, is plotted is a function of the sample size in Fig. 1.

3.2. The Lorenz63 model

Our second numerical example is a modified version of the Lorenz 63 model (Lorenz, 1963). It is a chaotic model and requires a lot more particles than the previous example in order to successfully track the hidden states. Our state vector is $x = [x_1\ x_2\ x_3]^T \in \mathbb{R}^3$ and we denote by $\mathcal{M}(x)$ the solution to the differential equation

$$\frac{dx_1}{dt} = a(x_2 - x_1), \quad \frac{dx_2}{dt} = x_1(b - x_3) - x_2, \quad \frac{dx_3}{dt} = x_1x_2 - cx_3, \tag{18}$$

which we solve using a fourth order Runge-Kutta scheme. For our experiment $[a,\ b,\ c] = [10,\ 28,\ 8/3]$ and $\delta t = 0.01$. The hidden Markov model evolves as

$$X_0 \sim \mathcal{N}(\mathcal{O}, \mathbb{I}),\ X_k = \mathcal{M}(X_{k-1}) + \eta_k,\ \ k = 1,\ldots,K, \tag{19}$$

where the time between X_k and X_{k-1} is δt, K=1000 and η_k is a zero mean Gaussian variable with covariance matrix $0.001^2\mathbb{I}$ independent of everything else. The measurement process, $Y_k \in \mathbb{R}^2$, is defined by

$$Y_k = \mathbb{H}X_k + \epsilon_k,\ \ k = 1,\ldots,K,\ \mathbb{H} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{20}$$

where ϵ_k is a zero mean Gaussian variable with covariance matrix $0.25^2\mathbb{I}$. As for the previous example, we run the experiment 500 times, and compute the RMSE over all runs and timesteps. The results are summarized in Fig. 2 for different sample sizes ranging from 25 to 5000. We clearly see that ASGM converges faster than both ASIR and the OSIR for this particular experiment.

3.3. Full scale well model – synthetic data

Below is a sketch (Fig. 3) of a branched producer, including the sensors, wellhead choke and downhole ICVs. The outlet pressure and temperature are 70 bar and 4 °C, respectively. The wellhead choke has 4 in. (maximum) diameter, and the actual opening is 35% of the maximum. The geothermal temperature gradient is 0.04 degrees Celsius per meter. The mainbore ICV has 3.5 in. (maximum) diameter, and is operating with 75% opening. The lateral ICV has 5.5 in. (maximum) diameter, and is operating with 35% opening. The uncertainty (error standard deviation) for the measurements are 5% for flowrate measurements (gas), and 0.2% for pressure and temperature measurements. There is one sensor above both ICVs measuring the combined pressure and temperature, and one sensor below each ICV in the mainbore and lateral.

For this synthetic case we estimate the flowrates of gas entering the mainbore and the lateral. We limit this study to one phase flow in order to isolate the filer performance and reduce the effect of spurious updates (due to limited ensemble size). We compare the AGM filter and the ASIR filter with the ASGM filter. We use 400 ensemble members for the ASGM and ASIR, and 800 ensemble members for the AGM. This ensures that the computational time is equal for the three methods. The bandwidth parameter h (see (Stordal et al., 2011)) is 0.8. The well reaches a measured depth at approximately 4551 m, but the numerical well flow model ends right after the downhole ICVs, and in the model inflow from the reservoirs also occur directly after the ICVs. We update the flowrates every 10 min for approximately 17 h. We also run the example three times using different random seeds.

In this context, the flowrates correspond to X_k in Eq. (1), and the stochastic model $\mathcal{M}$ is given by a multiple switching model characterized by a few possible multiplication values:

$$X_{k+1,i} = \mathcal{M}(X_{k,i},\ T^{pr},\ U_i,\ \eta_i),\ \ i = 1,\ldots,r. \tag{21}$$

Assuming r parameters, $X_k = (X_{k,1},\ \ldots X_{k,r})^T$ is the vector of the unknown parameters (flowrates). Further, the multiplication values are

given by the vector U_i, and T^{pr} is a probability matrix used to select a given value from U_i. The well flow model corresponds to the operator $\mathcal{H}$ in Eq. (2). For additional details regarding both the switching model and the well flow model, we refer the papers (Lorentzen et al., 2014, 2016). For the rest of Section 3 we represent unknown parameters by Q_k, indicating flowrates.

We use three multiplication values for the flowrates, given by

$$U_i = (0.95,\ \ 1,\ \ 1.05),\ \ i = 1, 2. \tag{22}$$

We use the same three multiplication values for both rates. The transition probabilities is a 3×3 matrix with equal rows, where each row is given by (0.1, 0.8, 0.1). The standard deviation for the (constant) Gaussian error term is set to $\sigma = 0.04$.

Initially, the gas flowrates are 1 kg/s in both the mainbore and the lateral. The flowrates are then ramped up to 4 kg/s in the mainbore and 6 kg/s in the lateral. The ramping starts first in the lateral, and this branch has the steepest gas rate inclination. The results for the three runs are shown on Fig. 4. The difference between AGM, ASIR and ASGM is not large, but the most important improvement can be seen during the gas rate increase, where the AGM and ASIR underestimates the true value. This is clearly seen by comparing the average (over simulation runs) root mean square errors (RMSE) for the estimated flowrates. Fig. 5 shows the RMSE for the filters as function of time. The sum of the error (denoted SRMSE) decreases from 27.6 and 22.5 for the ASIR and AGM respectively, to 17.7 for the ASGM. Most of the difference is due to the gas inclination period, but the ASGM has better performance also after the ramping period.

The synthetic observations and the estimated measurements are shown on Figs. 6–8. The increasing gas flowrate leads to increased pressure drop over the wellhead choke, resulting in higher wellhead pressure (upper left plot). Higher wellhead pressure, and increased pipe wall friction, leads to higher downhole pressure and higher temperatures. The estimated measurements are shown as blue lines. The two filters produce almost the same results, but the AGM and ASIR returns simulated temperature observations with more oscillations.

3.4. Full scale well model – real data

The real data are collected from the multiphase producer shown in Fig. 3. The specification of the well and sensors are the same as used in the previous example, except that the outlet pressure is approximately 53 bar in this case, and the actual choke and ICV openings are different. The choke opening is shown on Fig. 9, and the well is shut in during a period of approximately 6 h. The ICV in the mainbore is constantly set at 75% of maximum diameter, and the ICV in the lateral is 50% before the shut-in period, and 41% after the shut-in period. The data are cleaned for erroneous outliers identified by zero pressure or zero temperature, and the flowrate measurements are averaged over 10 min. The data frequency for the flowrate measurements is approximately 0.008 s^{-1}. The data frequency for pressure and temperature is approximately 0.03 s^{-1}. The pressure and temperature data are interpolated linearly in time if sampling points do not coincide with the data assimilation points. The measurement uncertainty is the same as for the previous example. In this example we estimate oil, water and gas entering the mainbore and the lateral. Except for a small amount of water directly after the shut-in period, the water measurements are approximately zero. It is however important to continuously estimate the water flowrates to detect water breakthrough at an early stage (before it is seen from the wellhead measurements). Initially we split the measured wellhead flowrates equally, and use the values as initial flowrates in the mainbore and the lateral. We then spin up the ensemble for 10,000 s to generate an initial state.

We use three multiplication values for the flowrates, given by

$$U_i = (0.95,\ \ 1,\ \ 1.05),\ \ i = 1, \ldots, 6. \tag{23}$$

We use the same three multiplication values for all rates. The transition probabilities is a 3×3 matrix with equal rows, where each row is given by (0.1, 0.8, 0.1). The Gaussian error terms are in this case set to 1% of the corresponding measured total flowrate at the outlet. E.g. if the measured oil rate is q_k, then $\sigma_{k,i} = q_k/100$, for i corresponding to variables representing oil rates in the mainbore and lateral. The number of particles is 80 in this case, and the bandwidth parameter h is 0.8. We update the flowrates every 10 min for approximately 96 h. Note that relatively frequent updates are selected here to capture most of the dynamic behavior. In this case only the ASGM is run using the real observations, as we do not have the true solution in this case, and are not able to compute the root mean square errors.

The filter methodology provides a standard deviations for the estimates, but as the parameters have different magnitude it is difficult to use these measures alone to determine the reliability for each parameter. Instead, we use the coefficient of variation (c_v), which is a relative standard deviation, given by

$$c_{v,i} = \frac{\mathrm{Std}(Q_i)}{(\mathrm{Mean}(Q_i) + 1e - 3)}. \tag{24}$$

We consider three levels of reliability for the flowrates: Good if $c_v \leq 0.075$; Fair if $0.075 < c_v \leq 0.15$; and Bad otherwise. In the following we use the colors green, orange and red to indicate these categories.

Fig. 10 shows oil, water and gas flowrates entering the mainbore (upper row) and the lateral (lower row). A particularly difficult dataset is selected, due to the variations in the choke (see Fig. 9) and the shut-in period. It is possible to distinguish flowrates in the mainbore and the lateral due to the pressure and temperature difference before and after the downhole ICVs. The gas flowrates have higher influence on the pressure drop over the ICV than the oil flowrates, and the estimated gas flowrates are therefore generally more reliable (green areas on Fig. 10). It is also possible to see the effect of the lateral ICV variation that shift from 50% opening to 41%. The gas flowrate is higher in the lateral than the mainbore before the shut-in period. After the shut-in period, and the change in the lateral ICV opening, the situation is opposite.

The estimated and real observations are shown on Fig. 11. It is generally small differences, but deviations are seen for pressure and temperature during the shut-in period. These are mainly due to numerical inaccuracies associated with the transition from a pressure boundary condition to zero-flow boundary condition at the outlet, when the choke is closed, and the opposite transition of boundary conditions when the choke opens.

For comparison, we have run the standard AGM filter on the same dataset. Fig. 12 shows the difference between the estimated flowrates, e.g. $Q_{\mathrm{ASGM}} - Q_{\mathrm{AGM}}$. The results prior to the shut-in period (approximately 40 h) show that the filters have almost equal performance, and the differences are probably due to stochastic and numerical effects. However, during and after the shut-in period the differences for the gas flowrates are larger, and at some point almost 5 kg/s. Towards the end of the assimilation period we can also identify trends in the difference for the estimated oil flowrates.

4. Conclusions

Motivated by the auxiliary sequential importance sampling (ASIR) filter, we have deduced the theory for a transformed particle filter. Further, we have applied a specific transformation given by the sequential Gaussian mixture filter, resulting in a new estimator denoted by auxiliary sequential Gaussian mixture (ASGM) filter. The theory for the transformed particle filter is general, and any transformation T aiming at minimizing the variance of the weights can be applied. In the same framework, we have also shown the formula for computing the weights for the optimal sequential importance sampling (OSIR) filter. The filters have been compared for several synthetic cases of varying complexity. The ASGM has superior performance. The final example shows the results using real measurements from a multiphase

producer, during a period of large choke variations. The estimated flowrates are credible, and the data match is good. Given proper selection of the bandwidth parameter h, the suggested approach improves performance when estimating parameters in cases with large variations of the observed variables.

Acknowledgments

The authors from IRIS acknowledge ENGIE E & P Norge AS for financial support through the Project Advanced Transient Flowrate Allocation (Grant No. 710.4116).

References

Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T., 2002. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. IEEE Trans. Signal Process. 50 (2), 174–188.

Bengtsson, T., Snyder, C., Nychka, D., 2003. Toward a nonlinear ensemble filter for high-dimensional systems. J. Geophys. Res.—Atmos. 108 (D24), ARTN 8775.

Bengtsson, T., Bickel, P., Li, B., 2008. Curse-of-dimensionality revisited: collapse of particle filter in very large scale systems. Probab. Stat. 2, 316–334.

Beskos, A., Crisan, D., Jasra, A., On the stability of sequential Monte Carlo methods in high dimensions. arXiv:1103.3965.

Carpenter, J., Clifford, P., Fearnhead, P., 1999. Improved particle filter for nonlinear problems. IEEE Proc. Radar Sonar Navig. 146 (1), 2–7.

Chen, R., Liu, J.S., 2000. Mixture kalman filters. J. R. Stat. Soc. Ser. B Stat. Methodol. 60, 493–508.

Chorin, A., Morzfeld, M., Tu, X., 2010. Implicit particle filters for data assimilation. Commun. Appl. Math. Comput. Sci. 5 (2), 221–240.

Crisan, D., Li, K., 2015. Generalised particle filters with gaussian mixtures. Stoch. Process. Appl. 125 (7), 2643–2673.

Doucet, A., Godsill, S., Andrieu, C., 2000. On sequential monte carlo sampling methods for bayesian filtering. Stat. Comput. 10, 197–208.

Doucet, A., de Freitas, N., Gordon, N. (Eds.), 2001. Sequential Monte-Carlo Methods in Practice. Springer-Verlag, New York.

Doucet, A., del Moral, P., Jasra, A., 2006. Sequential monte carlo samplers. J. R. Stat. Soc. B 68 (3), 411–436.

Doucet, A., De Freitas, N., Murphy, K., Russell, S., 2000. Rao-Blackwellised particle filtering for dynamic Bayesian networks. In: Proceedings of the Sixteenth conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers Inc., pp. 176–183.

Elgammal, A., Duraiswami, R., Davis, L.S., 2003. Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking. IEEE Trans. Pattern Anal. Mach. Intell. 25 (11), 1499–1504.

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn. 54 (6), 539–560.

Evensen, G., 2007. Data Assimilation: The Ensemble Kalman Filter. Springer-verlag, Berlin, Heidelberg, (ISBN 978-3-642-03711-5).

Fearnhead, P., 1998. Sequential Monte Carlo Methods in Filter Theory (Ph.D. thesis). , University of Oxford.

Frei, M., Künsch, H.R., 2013. Mixture ensemble Kalman filters, Computational Statistics and Data Analysis http://dx.doi.org/10.1016/j.csda.2011.04.013.

Gordon, N., 1993. Bayesian Methods for Tracking (Ph.D. thesis). . University of London.

Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R., Nordlund, P.-J., 2002. Particle filters for positioning, navigation, and tracking. IEEE Trans. Signal Process. 50 (2), 425–437.

Hoteit, I., Pham, D.T., Triantafyllou, G., Korres, G., 2008. A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography. Mon. Weather Rev. 136 (1), 317–334. http://dx.doi.org/10.1175/2007MWR1927.1, (ISSN 0027-0644).

Kotecha, J.H., Djurić, P.M., 2003. Gaussian sum particle filtering. IEEE 51 (10), 2602–2612.

Künsch, H.R., 2005. Recursive monte carlo filters: algorithms and theoretical analysis. Ann. Stat. 33 (5), 1983–2021.

Li, T., Corchado, J.M., Bajo, J., Sun, S., Paz, J.F.D., 2016. Effectiveness of bayesian filters: an information fusion perspective. Inf. Sci. 329, 670–689. http://dx.doi.org/10.1016/j.ins.2015.09.041, (ISSN 0020-0255).

Liu, B., Ait-El-Fquih, B., Hoteit, I., 2016. Efficient Kernel-Based Ensemble Gaussian Mixture Filtering. Monthly Weather Review, Published online February 2016. http://dx.doi.org/10.1175/MWR-D-14-00292.1.

Lorentzen, R.J., Stordal, A.S., Luo, X., Nævdal, G., 2016. Estimation of production rates using transient well flow modeling and the auxiliary particle filter - full-scale applications. SPE Prod. Oper. 31 (02), 163–175. http://dx.doi.org/10.2118/176033-PA.

Lorentzen, R.J., Stordal, A.S., Nævdal, G., Karlsen, H.A., Skaug, H.J., 2014. Estimation of production rates using transient well flow modeling and the auxiliary particle filter. SPE J. 19 (1), 172–180. http://dx.doi.org/10.2118/165582-PA, (SPE-165582-PA).

Lorenz, E., 1963. Deterministic nonperiodic flow. J. Atmos. Sci. 20 (2), 130–141.

Mandel, J., Beezley, J., 2009. An ensemble kalman-particle predictor-corrector filter for non-Gaussian data assimilation. In: Allen, G., Nabrzyski, J., Seidel, E., van Albada, G., Dongarra, J., Sloot, P., (Eds.), ICCS 2009: Computational Science. vol. 5545 of Lecture Notes in Computer Science. Springer Berlin/Heidelberg, pp. 470–478.

Moral, P.D., Jacod, J., 2004. The Monte-Carlo Method for filtering with discrete time observations: central limit theorems. In: Lyons T.J., Salisbury T.S., (Eds.), The Fields Institute Communications, Numerical Methods and Stochastics. American Mathematical Society.

Musso, C., Oudjane, N., Le Gland, F., 2001. Improving regularised particle filters. In: Sequential Monte Carlo Methods in Practice. Springer, pp. 247–271.

Papadakis, N., Mèmin, E., Cuzol, A., Gengembre, N., 2010. Data assimilation with the weighted ensemble Kalman filter. Tellus Ser. A Dyn. Meteorol. Oceanogr. 62 (5), 673–697. http://dx.doi.org/10.1111/j.1600-0870.2010.00461.x.

Pasha, S.A., Vo, B.-N., Tuan, H.D., Ma, W.-K., Gaussian, A., 2009. Mixture PHD filter for jump Markov system models. IEEE Trans. Aerosp. Electron. Syst. 45 (3), 919–936.

Pitt, M., Shephard, N., 1999. Filtering via simulation based auxiliary particle filters. J. Am. Stat. Assoc. 94, 590–599. http://dx.doi.org/10.1080/01621459.1999.10474153.

Reich, S., 2013. A nonparametric ensemble transform method for Bayesian inference. SIAM J. Sci. Comput. 35 (4), A2013–A2024.

Ristic, B., Arulampalam, S., Gordon, N., 2004. Beyond the Kalman filter. Artech House, 45–47, (ISBN 978-1580536318).

Silverman, B.W., 1986. Density Estimation for Statistics and Data analysis. Chapman and Hall, (ISBN 978-0412246203).

Stordal, A., Karlsen, H., Nævdal, G., Oliver, D., Skaug, H., 2012. Filtering with state space localized Kalman gain. Physica D. 241 (13), 1123–1135.

Stordal, A.S., Karlsen, H.A., Nævdal, G., Skaug, H.J., Vallès, B., 2011. Bridging the ensemble kalman filter and particle filters: the adaptive gaussian mixture filter. Comput. Geosci. 15, 293–305.

Stordal A.S., Lorentzen R.J., 2012. An iterative version of the adaptive Gaussian mixture filter. In: Proceedings of the EAGE European Conference on the Mathematics of Oil Recovery. Biarritz, France, 2012.

Stordal, A.S., 2011. Sequential Data Assimilation in High Dimensional Nonlinear Systems (Ph.D. thesis). . University of Bergen.

Vo, B.-N., Ma, W.-K., 2006. The gaussian mixture probability hypothesis density filter. IEEE Trans. Signal Process. 54 (11), 4091–4104.