# Different Parameterizations of the Initial Ensemble for a Channelized Reservoir in an Assisted History Matching Context

Sebacher, Bogdan; Stordal, Andreas; Hanea, Gabriel

**DOI**
10.3997/2214-4609.201601817

**Publication date**
2016

**Document Version**
Final published version

**Published in**
Proceedings of the 15th European Conference on the Mathematics of Oil Recovery

Mo P080

# Different Parameterizations of the Initial Ensemble for a Channelized Reservoir in an Assisted History Matching Context

B. Sebacher (Delft University of Technology), A.S. Stordal (IRIS) & R.G. Hanea* (Statoil)

## SUMMARY

In this paper we present a comparison of three parameterizations of channelized reservoirs generated using multipoint geostatistics (MPS) in combination with a training image. In a previous study, we suggested estimating the facies probability fields from an ensemble generated with MPS and linked, marginally, the facies probability fields with the standard Gaussian variables by means of the normal score transform. We have parameterized the facies fields with random fields, marginally Gaussian, using the conditional mean of the Gaussian variables. This parameterization keeps a possible dependence structure inherited from the training image, but marginally the sampling from the Gaussian distribution is discrete and bi-modal. Here, we extend this parameterization in two directions. First, we do not take into account the dependence structure and parameterize by random sampling from the conditional distribution. The second idea is to draw samples from the conditional distribution, but using the same random seed for each grid cell within each ensemble member, but different random seeds across the ensemble members. This would preserve the dependence structure within each ensemble member while increasing the variability between the ensemble members. Both parameterizations have the property that, marginally, samples correctly from the standard Gaussian distribution. We compare the behavior of the parameterizations within a history matching process assimilating the production data. The comparison has two main directions: to prove the impact of the stochastic forcing on the history matching of geological properties and to prove the stochastic forcing on the predictive power of the models. We have used the iterative adaptive Gaussian mixture filter (IAGM) for history matching because the IAGM is suited for highly nonlinear problems and has a re-sampling step that allow us to use the already existing technique of re-sampling from the training image using updated probability fields. The re-sampling step is necessary to re-position the facies geometry, lost after a cycle of data assimilation.

## Introduction

The estimation of two facies channelized reservoirs conditioned to production data is a problem that was extensively studied in the reservoir engineering community. This problem poses difficulties from beginning when a reliable geological simulation model has to be chosen in order to simulate facies fields that exhibits channelized geometry. It is well known that the truncated Gaussian simulation model fails in this matter and, until now, two geological simulation models have been proved as suitable to simulate channelized reservoirs: the multi-point geostatistical simulation (MPS, Caers and Zhang (2004)) models and the object based simulation models (Deutsch and Wang, 1996). Once the geological simulation model is chosen, there are two options to estimate the position of the channels conditioned to production data: either the distribution of the permeability (or a parametrization of it) is estimated and afterwards, if necessary, the channel positions is inferred from it or a parametrization of the facies fields is defined, which permits the estimation of the channels directly. From the first category, (Sarma et al., 2008) proposed a method to project the permeability field into a multi-dimensional kernel space using the kernel-principal component analysis (K-PCA) and performing the parameterization in this space. The model parameters from the kernel space were coupled with the ensemble Kalman filter (Evensen, 2003) for history matching of the production data. Jafarpour and McLaughlin (2008) parameterize the permeability field using a basis obtained by applying of the discrete cosine transform (DCT) and the parameters that are updated in the EnKF framework are the coefficients of the DCT transformation. In Zhang et al. (2015) the permeability field is estimated from the production data using an iterative ensemble smoother and afterward it is projected onto a facies field using a post-processing technique. In Jafarpour and Khodabakhshi (2011) the authors proposed the probability conditioning method (PCM) in which a probability field of the channel is inferred from the estimated permeability field and, is further used to condition the MPS algorithm to it. As consequence, the updated channelized reservoir is obtained by a random sampling from a training image conditioned to the previous probability field. In Zhou et al. (2012) the authors have used the normal score transformation to project marginally the permeability field onto a standard Gaussian space where the ENKF is performed. The permeability field is re-constructed based on the inverse of the normal score transformation.
For the second category, in order to estimate the channel position, a parametrization of the facies fields must be defined first. In Lorentzen et al. (2012) the authors propose a parameterization of the facies fields using the distance from the current cell to the border of the channel. If the cell is inside of the channel the distance function is positive and if the cell is outside, the distance is negative. The distance function is further updated within an EnKF process, assimilating production data and conditioned to some statistical measures in order to keep the continuity of the channel bodies. In Hu et al. (2013) the parametrization is performed not on the facies fields but using a parameter inside of the MPS method. The authors considered the uniform numbers that helps the simulation of the facies field at each grid cell as the parameters to be updated in an EnKF process by assimilation of the production data. The principal component analysis (PCA) is also used to parameterize channelized reservoirs. In Vo and Durlofsky (2015) the authors implemented an optimization-PCA method to deal with non-Gaussian complex models in combination with the randomized maximum likelihood (RML) method.
In this study we are situated in the second category using as the geological simulation model the MPS in combination with a two-facies channelized training image. In Sebacher et al. (2015) and Sebacher et al. (2016) we introduced a parametrization bridging the MPS with the truncated Gaussian simulation model. The parameterization consist of defining of some random fields, marginally Gaussian, of which truncation with reliable thresholds simulate the same facies fields with those previously obtained from the training image with the MPS. Here we propose two parameterizations which are based on the same idea, but the random fields are constructed different. If in Sebacher et al. (2015), the marginally Gaussian random fields were defined by the means of the conditional mean of a standard normal variable, here we follow two directions: the first idea is to random sample from the conditional distribution using different random seeds for each grid cell within each ensemble member and the second idea is to random sample from the conditional distribution using constant random seed for the grid cells of an ensemble member but different seeds across ensemble members. We compare these parameterizations within an assisted history matching (AHM) process using the iterative adaptive Gaussian mixture filter (IAGM, Stordal and Lorentzen (2014)), assimilating production data. The comparison has two main directions: to quantify the impact of the stochastic forcing on the history matching of geological properties and to

prove the stochastic forcing on the predictive power of the models. The use of the IAGM is suitable here because the IAGM works well for highly nonlinear problems and has a re-sampling step which allows us to use the technique of re-sampling with MPS from the training image conditioned to the updated probability fields (Krishnan et al., 2005). The re-sampling is necessary because after an assimilation cycle the channelized geometry may be broken and the re-sampling is repositioning the channel structure in the field.
In the next sections we introduce the parameterizations, present the IAGM with its customized implementation for our problem and a case study where we perform the comparison. The papers ends with the conclusions.

### The parameterizations

The geological simulation model used for generation of the initial ensembles is the single normal equation simulation model (SNESIM, Strebelle (2002)) of which training image that represents the geological concept of channelized reservoir is presented in Figure 1. In this section we start by presenting the parameterization proposed in Sebacher et al. (2015) and based on this idea we extend it in two directions. Consider an ensemble of $n_e$ channelized reservoirs, simulated with SNESIM from the training image us-

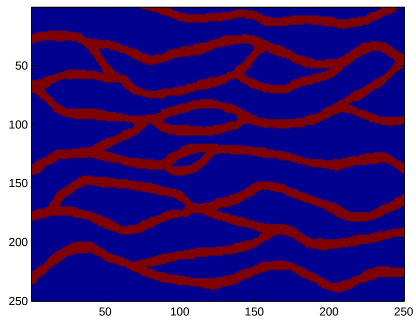



***Figure 1*** *The training image.*

ing as constraints only the global facies proportions on a squared reservoir domain with $100 \times 100$ grid cells. We calculate, from this ensemble, the probability field of the channel occurrence (Figure 2) which marginally means the calculation of a value $p^j$ representing the probability that the channel occurs in the grid cell $j$ of the reservoir domain.

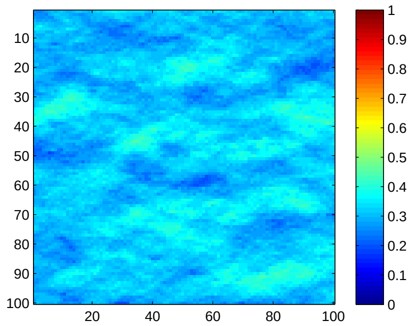



***Figure 2*** *The probability field of the channel.*

At the grid cell $j$ we consider the discrete random variable, denoted $facies\_distribution$ of which distribution can be described as in eq. 1.

$$facies\_distribution^j \sim \begin{pmatrix} Channel & Non-channel \\ p^j & 1-p^j \end{pmatrix} \tag{1}$$

The idea is to link this discrete random variable with a standard normal variable and the easiest way is by the means of the normal score transform. Consequently, we calculate the cumulative distribution of this random variable and we define the threshold $\alpha^j$ in the Gaussian space as $\alpha^j = \Phi^{-1}(p^j)$, where $\Phi$ is the standard Gaussian cumulative distribution function, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u2}{2}}\, du$. The threshold

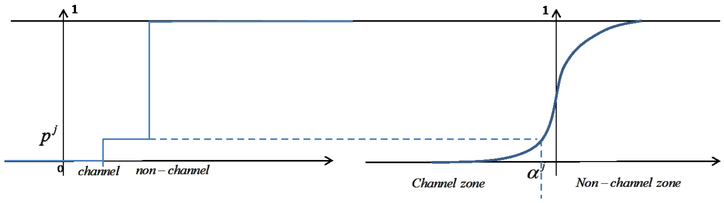



***Figure 3*** *The normal score transform*

$\alpha^j$ splits the real axis in two intervals $(-\infty, \alpha^j]$ and $(\alpha^j, \infty)$ corresponding to the channel and non-channel respectively (Figure3). In Figure 4 is shown the thresholds $\alpha$ on the reservoir domain where it can be seen a strong spatial structure similarly with the structure of the probability field of the channel (Figure 2) which is inherited from the training image. The parameterizations are performed in the multi-

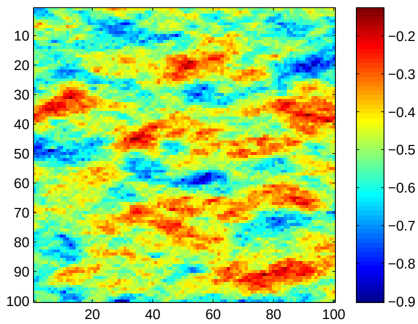



***Figure 4*** *The thresholds*

dimensional real space $\mathbf{R}^n$, where $n$ is the dimension of the reservoir domain (number of the grid cells, in our case 10000). For each ensemble member $i$ (two-facies channelized reservoir) we define a parameter field on the reservoir domain, denoted $\theta_i$, of which truncation with the thresholds $\alpha$ yields the ensemble member $i$. we consider the following situations:

*Parametrization with gravity centers (GC)*

Here we are using the idea from Sebacher et al. (2015) parameterizing the facies field using the gravity centers of the intervals $(-\infty, \alpha^j]$ and $(\alpha^j, \infty)$ with respect to the standard Gaussian density $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x2}{2}}$. Consequently, for ensemble member $i$ we define at grid cell $j$ the value of the parameter field $\theta_i^j$ as:

$$\theta_i^j = \begin{cases} \mathbf{E}(X|X \leq \alpha^j) = \frac{-\phi(\alpha^j)}{\Phi(\alpha^j)} & \text{if } j \in \text{channel}, \\ \mathbf{E}(X|X > \alpha^j) = \frac{\phi(\alpha^j)}{1-\Phi(\alpha^j)} & \text{if } j \in \text{non-channel}, \end{cases} \tag{2}$$

where, $X$ is a random variable having a standard normal distribution.

*Parametrization with random sampling (Random seed)*

The first new parametrization introduced is based on the idea to define the parameter field $\theta$ by a random sampling from a conditional Gaussian distribution. The sampling is done randomly for each ensemble member and for each grid cell of the reservoir domain. Consequently, for ensemble member $i$ we define at grid cell $j$ the value of the parameter field $\theta_i^j$ as:

$$\theta_i^j = \begin{cases} \text{random sampling from } (X|X \leq \alpha^j) & \text{if } j \in \text{channel}, \\ \text{random sampling from } (X|X > \alpha^j) & \text{if } j \in \text{non-channel}, \end{cases} \tag{3}$$

*Parameterization with random sampling using the same random seed for each grid cell within each ensemble member but different random seeds across the ensemble members (Fixed seed)*

This parametrization is defined so that for each ensemble member we are using a single random seed for each grid cell of the reservoir domain for sampling from the conditional distribution. In order to define this parametrization we firstly calculate the cumulative distribution function (**cdf**) of the conditional Gaussian distributions $(X|X \leq \alpha^j)$ and $(X|X > \alpha^j)$:

$$\mathbf{cdf}_{X|X\leq\alpha^j}(x) = \begin{cases} \frac{\Phi(x)}{\Phi(\alpha^j)} & \text{if } x \leq \alpha^j, \\ 0 & \text{if } x > \alpha^j, \end{cases} \tag{4}$$

$$\mathbf{cdf}_{X|X>\alpha^j}(x) = \begin{cases} 0 & \text{if } x \leq \alpha^j, \\ \frac{\Phi(x)-\Phi(\alpha^j)}{1-\Phi(\alpha^j)} & \text{if } x > \alpha^j, \end{cases} \tag{5}$$

For sampling from a random variable $Y$ with known cumulative distribution function $\mathbf{cdf}_Y$, we use the probability integral transform. Then, $\tilde{y}$ is a random sample from $Y$ if and only if $\mathbf{cdf}_Y(\tilde{y})$ is a random sample from the uniform distribution U(0,1). Consequently, for each ensemble member $i$ we draw a random seed y from the uniform distribution U(0,1) and after solving the equations: $\mathbf{cdf}_{X|X\leq\alpha^j}(x) = y$ and $\mathbf{cdf}_{X|X>\alpha^j}(x) = y$ we find the solutions:
$x = \Phi^{-1}(y * \Phi(\alpha^j))$ and respectively $x = \Phi^{-1}(\Phi(\alpha^j) + y * (1 - \Phi(\alpha^j))$ which are random samples from the conditional Gaussian distributions $(X|X \leq \alpha^j)$ and $(X|X > \alpha^j)$. Consequently, using the same random seed $y$ for each grid cell $j$ of the reservoir domain we define the parameter field $\theta$ associated with member $i$ as follows:

$$\theta_i^j = \begin{cases} \Phi^{-1}(y * \Phi(\alpha^j)) & \text{if } j \in \text{channel}, \\ \Phi^{-1}(\Phi(\alpha^j) + y * (1 - \Phi(\alpha^j)) & \text{if } j \in \text{non-channel}, \end{cases} \tag{6}$$

*Comparison of the parameterizations before history matching*

In Figure 5 we present two ensemble members of the ensemble of channelized reservoirs (first column) and the associated parameter fields as follows: In the second column is shown the parameter fields calculated with the parameterization with gravity center, in the second column the parameter fields calculated with the random seed and in the fourth column the parameter fields calculated with the fixed seed. From this figure it can be seen that the random seed parameterization destroys any possible two-points correlation that may be inherited from the training image; inside of the channels or outside of the channels we find values with no spatial correlation. This is not happening for the other two parameterizations, inside of each body of facies the values of the parameter field are spatially correlated. However, for the case of the last parameterization (last column) we observe a higher variability than the one observed for the first parameterization (second column) because the ranges of the parameter field are different for different ensemble members. In Figure 6 are shown the mean of the parameter field calculated from the ensemble for all parameterizations. If for the first parameterization the mean is a zero uniform field (Sebacher et al., 2015) the other two means are not zeros, but their values are close to zero. The random seed parameterization has generated a mean field with values randomly distributed over domain whilst in the mean field generated with the fixed seed parameterization one can clearly see a spatial correlation similarly with those seen in the threshold field (Figure 4). In Figure 7 is shown the histograms of

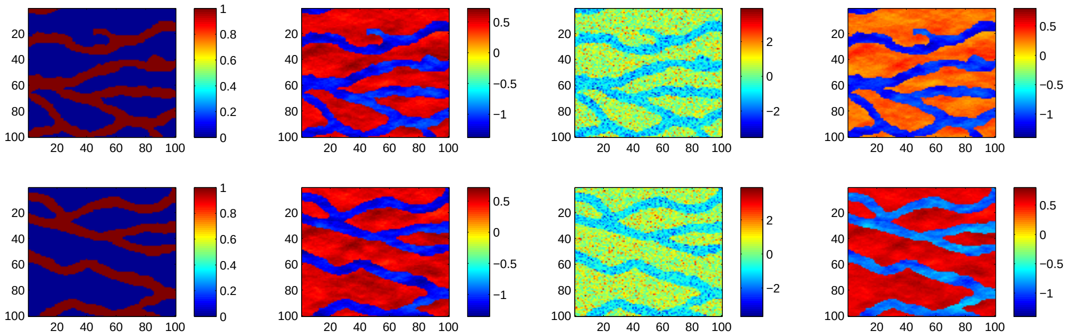
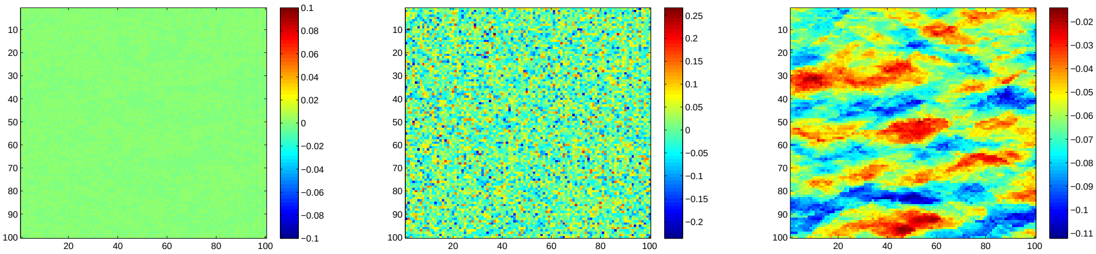


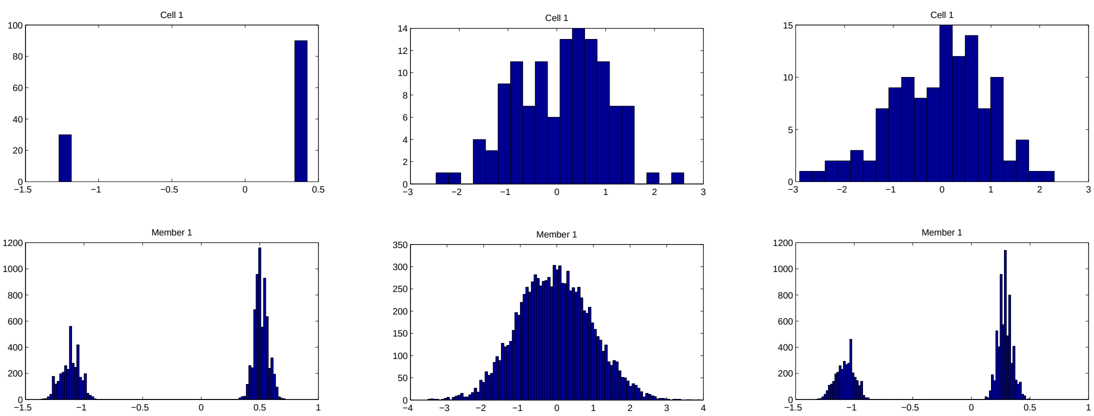
***Figure 5*** *The facies fields and the associated parameter fields.*



(a) Gravity center (b) Random seed (c) Fixed seed

***Figure 6*** *The mean of the parameter field for all parameterizations.*

the parameter field values as follows: in the first line the histogram is calculated from the ensemble of values but for the first cell of the reservoir domain and in the second line we present the histogram of all values for the first ensemble member. For the parameterization with gravity centers, we approximate (marginally) the standard Gaussian distribution with a bi-modal distribution (Figure 7, position (1,1)) which is a bad approximation. The last two parameterizations samples much better from the Gaussian distribution (last two positions of the first row of Figure 7). If we calculate the histogram of the parameter field from Member 1 one can observe a bi-modal behavior of the parameterization with gravity centers and with fixed seeds whilst the random sampling yields a Gaussian distribution (second row of Figure 7).




(a) Gravity center (b) Random seed (c) Fixed seed

***Figure 7*** *The histograms of the values of the parameter fields in all parameterizations*

**The iterative adaptive Gaussian mixture filter (IAGM)**

The iterative adaptive Gaussian mixture filter (IAGM) is the iterative version of the adaptive Gaussian mixture filter (AGM, Stordal et al. (2011)) developed for improving the AGM for nonlinear models (Stordal and Lorentzen, 2014). In the AGM the uncertainty of the system is represented by an ensemble of possible states and each ensemble member (particle) has an associated likelihood weight. Initially all ensemble members have the same weight (equally weighted), $\widehat{W}_j = 1/n_e$, $j = 1,\dots,n_e$, as in the ensemble Kalman filter (EnKF). In our case the ensemble of parameter fields consists of $\{\theta^i\}_{i=1}^{n_e}$ and we denote by $\mathscr{G}_t$, the function with output the simulated observations given the parameters. We augment our state vector with these variables. Hence our ensemble state vector $X_t^i$, $i = 1,\dots,n_e$, is given by

$$X_t^i = [\theta_i^T \quad \mathscr{G}_t(\theta_i)]^T, i = 1,\dots,n_e. \tag{7}$$

This augmentation allows us to construct, for each assimilation time $t$, a binary matrix $H_t$ with the relation $Y_t = H_t X_t + \varepsilon_t$ where $\varepsilon_t$ follows a Gaussian distribution with mean 0 and known covariance matrix $R_t$. That is, the measurement is a linear function of our augmented state vector with additive Gaussian white noise. The measurements used in this study are the bottom hole pressures taken at the injection wells, the oil and water rates taken at the production wells. The dynamical variables (the pressure and saturation fields) are not included in the state vector as we have chosen to rerun the simulator from time zero.
Let be $C_t$ the sample covariance matrix of $\{X_t^i\}_{i=1}^{n_e}$ calculated based on the weighted ensemble mean:

$$\overline{X_t} = \sum_{i=1}^{n_e} \widehat{W}_t^i X_t^i \tag{8}$$

.

At each assimilation time, the augmented state vector is updated in the AGM (and IAGM) for each $i = 1,\dots,n_e$ as

$$\widehat{X}_t^i = X_t^i + C_t H_t^T (H_t C_t H_t^T + h^{-2} R_t)^{-1} \left(y_t - H_t X_t^i + \varepsilon_t^i\right) \tag{9}$$

here $\varepsilon_t^i$ is a sample from the Gaussian measurement error distribution $N(0,R)$. The update is similar with the standard EnKF, the only difference is the scaling $h^{-2}$ of the measurement error covariance matrix $R_t$. In other words the linear update is dampened, where the dampening factor $h^{-2}$ ($h \in [0,1]$) is the bandwidth of the Gaussian mixture (Stordal et al., 2011). For h equal to 1 we get the standard EnKF update. In addition to a linear update, importance weights are derived from the Gaussian mixture and updated sequentially as

$$\begin{aligned} \overline{W}_t^i &= \widehat{W}_{t-1}^i \Phi(y_t - H_t X_t^i, H_t C_t H_t^T + R_t), \\ W_t^i &= \frac{\overline{W}_t^i}{\sum_{i=1}^N \overline{W}_t^i}, \end{aligned} \tag{10}$$

Here, the function $\Phi(x-\mu,\mathbf{C})$ represents multivariate Gaussian density with mean $\mu$ and covariance matrix $\mathbf{C}$. To avoid filter degeneracy that occurs in high dimension and complex systems a weight interpolation is introduced

$$\widehat{W}_t^i = \alpha_t W_t^i + (1-\alpha_t) n_e^{-1}, \quad \alpha_t \in [0,1], \tag{11}$$

where

$$\alpha_t = n_e^{-1} (\sum_{i=1}^{n_e} (W_t^i)^2)^{-1}. \tag{12}$$

For details of the AGM we refer to Stordal et al. (2011). A re-sampling and re-weighting step in the algorithm before rerunning AGM is discussed in Stordal and Lorentzen (2014).

*Reconstruction of the facies fields*

After the assimilation of the data the values of the parameter field $\theta$ is changing according to eq. 9. Using the new values we updated the facies fields on a cell by cell basis:
*For each ensemble member $i \in 1,2,\ldots,n_e$ and for each grid cell $j \in 1,2,\ldots,n$ if $\theta_i^{update,j} \leq \alpha^j$ we assign in the grid cell $j$ the channel and if $\theta_i^{update,j} > \alpha^j$ we assign the non-channel.*
At this step what we have to do is to truncate the values of the updated parameter field with the thresholds calculated at the initial time. This truncation will project a continuous field into a discrete field (as in the truncated Gaussian method).

*Re-sampling in IAGM*

The SNESIM has incorporated the tau model (Journel (2002), Krishnan et al. (2005)) which enables integration of the probabilities coming from soft data and the training image. Usually, the probability maps used as soft data are coming from seismic interpretations, but here, we adopt the procedure from Jafarpour and Khodabakhshi (2011) but using the probability fields of the facies calculated after an iteration of IAGM. The tau model depends on two weights, denoted $\tau_1$ and $\tau_2$, that calibrates the geological concept (the training image) with the probability fields. In this study the weights used are both equal to 1, which correspond to the equal importance assigned to the training image and weighted probability map respectively.
After one complete assimilation cycle, new weighted probability fields (of the channel and non-channel) are constructed from the updated ensemble of channelized reservoirs:

$$p_k^j = \sum_{i=1}^{n_e} \widehat{W}_t^i Ind_k^{i,j}, \quad k = channel, non-channel \tag{13}$$

where

$$Ind_k^{i,j} = \begin{cases} 1 & \text{if cell j} \in \text{facies type k} \\ 0 & \text{if cell j} \notin \text{facies type k} \end{cases} \tag{14}$$

Where $i$ represents the ensemble member and $j$ is the indicative of the grid cell and $t$ denotes the last time step where we have measurements. Using the tau model we generate a new ensemble of channelized reservoirs conditioning the training image to the weighted probability fields of the facies types. This is equivalent to re-sampling marginal facies variables from the empirical distribution obtained with AGM (re-sampling) with the constraint that the dependence structure is given by the training image. The re-sampling step of the IAGM repositions the facies geometry and this is one of the reason for using this AHM method. After each iteration we evaluate the ensemble from the geometry and data match perspective and decide if it is necessary to proceed with another iteration. After a finite number of iterations the updated ensemble of facies fields remain geologically realistic and provide, in addition, a good data match.

*Algorithm*

The algorithm of one complete assimilation cycle with the AGM can be summarized as follows:

1. Generate an ensemble of $n_e$ channelized reservoirs from the training image.
2. Calculate the parameter fields $\theta_i$ for $i \in \overline{1,n_e}$.
3. Construct the state vector in (7) associating equal weights of $1/n_e$ to each ensemble member and set the dampening parameter $h$.
4. For each assimilation time do
    - Apply the forward model from time $t$ to the next assimilation time, populating first the facies fields with reliable petrophysical properties.

- At the assimilation time, assimilate the data, updating the state vector according to (9) and the weights according to (11).
- Reconstruct the updated channelized reservoirs using the updated parameters $\theta_i^{updated}$.

**Case study**

The parameterizations presented above are tested using a reservoir model of which domain has a square shape of dimension is $100 \times 100$ grid cells.The dimension of each grid cell is set as $30 \times 30 \times 20$ ft. We design the reservoir as a 13-spot water flooding black oil model, having four injection wells at the center of reservoir domain and nine production wells surrounding the injection wells. The reservoir is initially filled with oil at a constant uniform saturation of 0.8 (the connate water saturation is 0.2) and with a uniform pressure of 5000 psi in every grid cell. The producers works under constant bottom hole pressure (BHP) with a value of 3000 psi and the injectors operate at 3500 STB/D constrained by a maximum BHP of 100000 psi. For each experiment the measurements were obtained through forward simulation of a synthetic model presented as the "reference" which was randomly sampled from the same training image using SNESIM. The reference model with the position of the wells is presented in Figure 8. The measurement errors of the production data are assumed Gaussian with 0 mean and

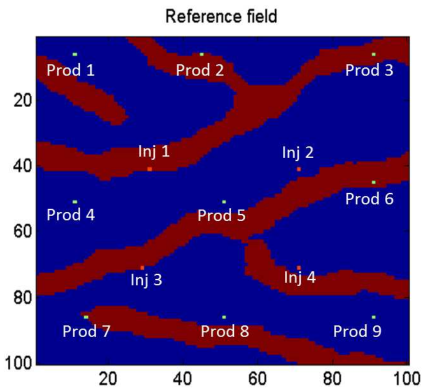


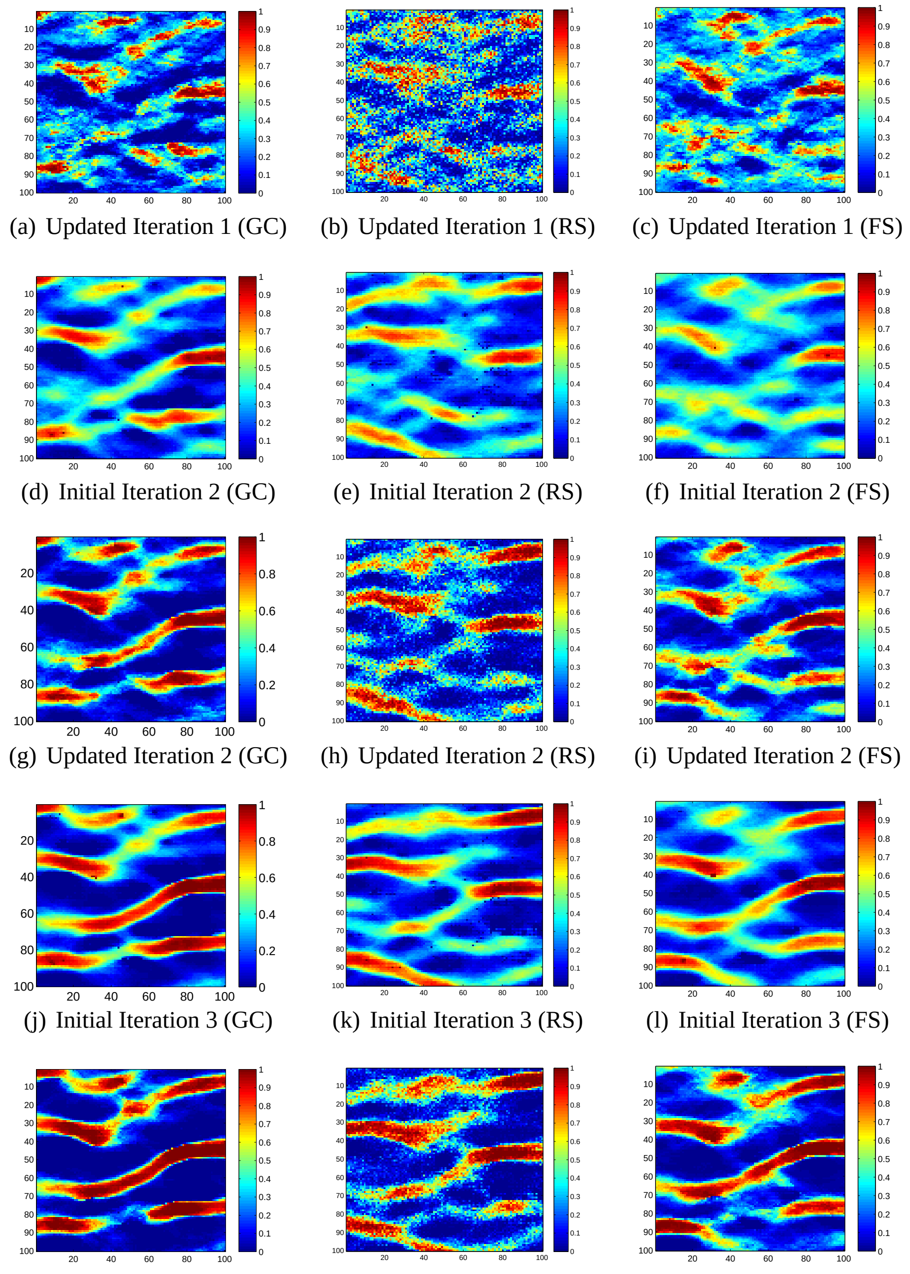

***Figure 8*** *The reference field.*

standard deviations of 70 STB/D for water rates (WR) and oil rates (OR) at the producers, and 200 psi for BHP at the injectors. These values are used to generate noisy observations from the reference model in addition predicted observations of the production data used in the analysis step of the HM processes. Water injection starts from the first day and continue thereafter a period of 351 days of production. We assimilate data at a 60-day interval resulting in a total of 6 assimilation steps. The permeability values were set at 9 mD and 1 mD for the channel facies type and for the non-channel facies type, respectively, while the porosity of both facies types is set to 0.2 and considered as known. The dampening parameter, $h$, is set to 0.25 in the AGM runs based on previous experience and we have performed three iterations with the IAGM. The ensemble size is set to 120.

**Results**

In Figure 9 are shown the probability fields of the channel in initial (re-sampled , in even rows) and updated ensembles (odd rows) in three iterations with IAGM, for all parameterizations (GC represents the parameterization with gravity centers, RS the parameterization with random seeds and FS represents the parameterization with fixed seed). In all experiments the initial ensemble of channelized reservoirs is the same and is shown by Figure 2. The results obtained for parameterization with gravity centers are taken from Sebacher et al. (2015). Analyzing the last row of the Figure 2 where the probability fields of the channel in the last iteration is presented, one can observe that all parameterization capture the same trend of the possible position of the channels into the field. The channel estimation is good enough if compared it, visually, with the reference (Figure 8). However, there are some particularities:

- The parameterization with gravity centers yields a smaller variability than other two parameterizations (Figure 9, sub-figure (n) vs (o) and (p)) .
- The lack of the spatial correlation of the parameter fields of the parameterization with random seeds affects the probability fields of the channel in the sense that the channel position is diffusely estimated (Figure 9, sub-figures (c), (i), (o)).
- The spatial dependence structure introduced by the parameterizations and with gravity centers and fixed seeds really help the estimation; the probability fields of the channel for both parameterizations exhibit much shaped zones than the probability field of the channel of the parameterization with random seed.



(a) Updated Iteration 1 (GC) (b) Updated Iteration 1 (RS) (c) Updated Iteration 1 (FS)

(d) Initial Iteration 2 (GC) (e) Initial Iteration 2 (RS) (f) Initial Iteration 2 (FS)

(g) Updated Iteration 2 (GC) (h) Updated Iteration 2 (RS) (i) Updated Iteration 2 (FS)

(j) Initial Iteration 3 (GC) (k) Initial Iteration 3 (RS) (l) Initial Iteration 3 (FS)

(m) Updated Iteration 3 (GC) (n) Updated Iteration 3 (RS) (o) Updated Iteration 3 (FS)

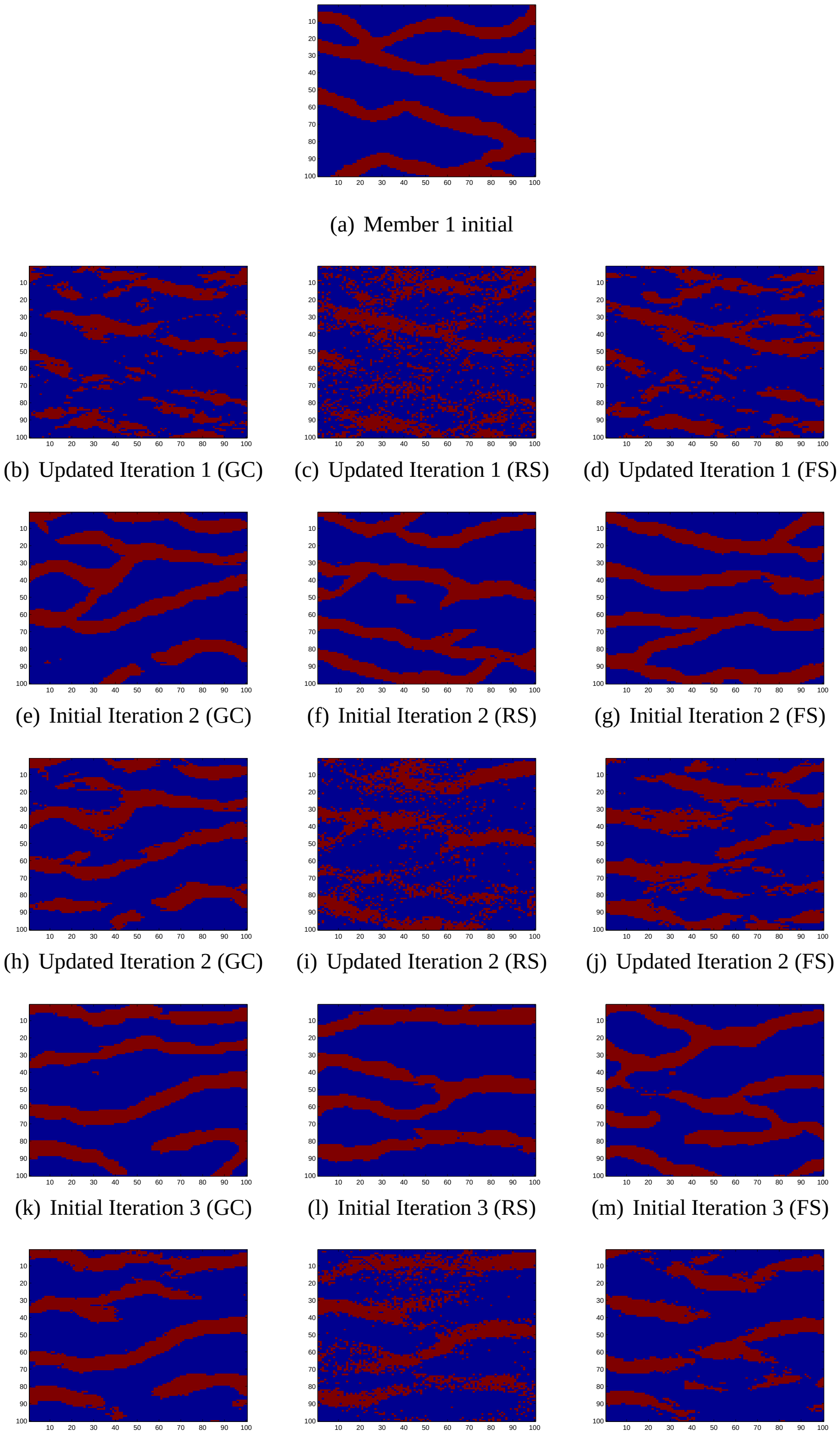
***Figure 9*** *The probability fields of the channel for all parameterizations (GC-gravity center, RS-random seed, FS-fixed seed)*

The probability fields of the channel are calculated from ensembles of facies fields either generated with SNESIM or the result of estimation with the IAGM. They only present possible positions of the channel into the field, do not offer any information of the geological plausibility of the channel fields. In order to quantify the uncertainty of the channel distribution into the field we have to look at the facies fields. In Figure 10 is shown the evolution of the first member of the ensembles during the experiments. Initially, it presents a nice channelized structure (sub-figure (a)) because is the result of a simulation with SNESIM from the training image, structure regained after each re-sampling for all parameterizations (odd rows). The channelized structure is broken differently after the iterations of the history matching. For the parameterizations with gravity centers and fixed seed the updated member is still presenting continuous bodies, whilst for the parameterization with random seed the updated member is far from the geological concept presented by the training image, even after the third iteration. We believe the reason is the lack of the spatial dependence of the parameter field. However, the facies fields obtained with this parameterization look better with iterations and are close to the estimation with gravity centers (sub-figure (o) vs sub-figure (n)). After the third iteration the parametrization that is able to keep better the channel continuity is surprisingly the one with gravity centers (sub-figure (n)) while the parametrization with fixed seed seems to needs extra iterations. This is the reason that we have stop after three iterations, the parameterization with gravity centers was able to keep the channel continuity in the updates and we production data match and the predictions are also very good (Figure 11).
We continuing the comparison presenting the water rates at all the production wells in all the experiments for 651 days which is split in two periods: the assimilation period (from 0 to 351) and the prediction period from (351 to 651 days). Figure 11 shows the evolution of the water production rate for the gravity center parametrization and is taken from Sebacher et al. (2015). In Figure 12 and Figure 13 is shown the production water rates for the parameterization with random seeds and fixed seeds respectively. Analyzing with a visual inspection these evolutions we are not able to rank them, all present a good data match and predictions. This is not a surprise because all parameterizations match the data modifying the channel position into the field on a grid cell basis and all have a high degree of freedom to do it. Here the AGM helps because of the small value of the dampening factor (h=0.25) which yields small correction in the updates while matching the data.

In order to better analyze the comparison of the proposed parameterizations we present the evolution over iterations of the Root Mean Square Error calculated as weighted euclidian distance between the reference, (Ref) and each ensemble member, $X^i$ (Eq. 15), calculated for permeability and for total production rates for a period of 651 days.

$$RMSE^2 = (n_d)^{-1} \sum_{i=1}^{n_e} \sum_{j=1}^{n_d} \widehat{W}_i (V_j(Ref) - V_j(X^i))^2, \tag{15}$$

where $n_d$ is the number of data, $X^i$ denotes ensemble member number $i$, $\widehat{W}_i$ the corresponding likelihood weight and $V_j$ is the variable of interests (i.e. the permeability and the total production rates). Table 2 presents the RMSE calculated for permeability field and Table 1 the RMSE calculated for total production rates for a period of 651 days. Both results indicates that, after three iterations, the smaller values of RMSE arw obtained by the parameterization with gravity centers and the higher values are obtained by the parameterization with random seeds even though all values are within a small bandwidth. In addition, after the third iteration, the parameterization with fixed seeds is very close to the parametrization with gravity centers and observing its higher variability in the updated ensemble one can try another iteration with fixed seed. However, this iteration will not help in the estimation of the channel position because, after the third iteration, the probability fields of the channel show similar characteristics (Figure 9, sub-figures (m) and (o)). A new iteration for the parameterization with fixed seed will constrain the re-sampled ensemble so that it yield a probability field and predictions comparable with those obtained by the parameterization with gravity centers after the third iteration.

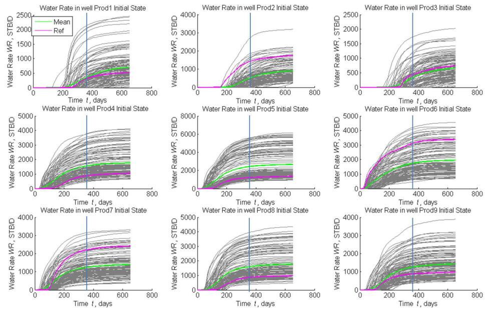


(a) Member 1 initial
(b) Updated Iteration 1 (GC) (c) Updated Iteration 1 (RS) (d) Updated Iteration 1 (FS)
(e) Initial Iteration 2 (GC) (f) Initial Iteration 2 (RS) (g) Initial Iteration 2 (FS)
(h) Updated Iteration 2 (GC) (i) Updated Iteration 2 (RS) (j) Updated Iteration 2 (FS)
(k) Initial Iteration 3 (GC) (l) Initial Iteration 3 (RS) (m) Initial Iteration 3 (FS)
(n) Updated Iteration 3 (GC) (o) Updated Iteration 3 (RS) (p) Updated Iteration 3 (FS)

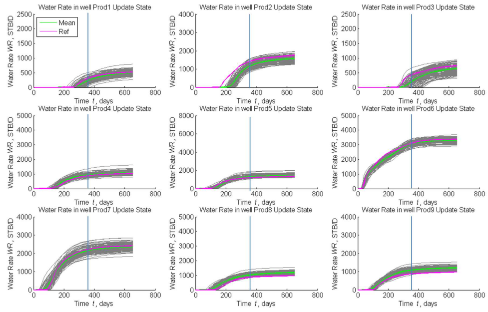
***Figure 10*** *Member 1 in initial and updated ensembles for all parameterizations (GC-gravity center, RS-random seed, FS-fixed seed)*

## Conclusions

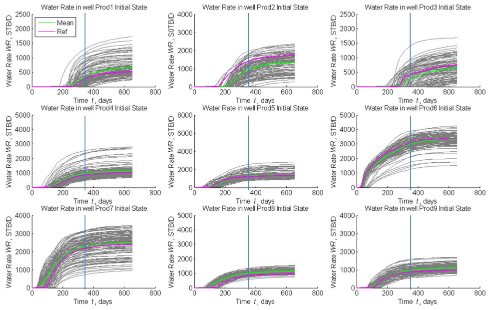
In this study we have presented a comparison between three parameterizations of channelized reservoirs generated with a multi-point geostatistical simulation model before and after they are coupled with a



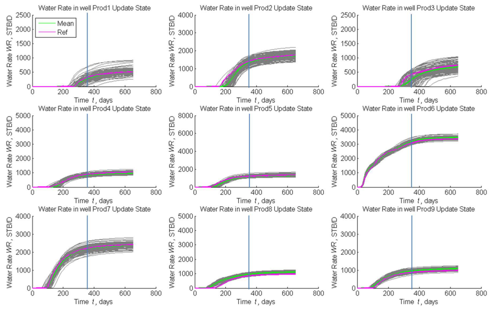
(a) Production water rates from initial ensemble



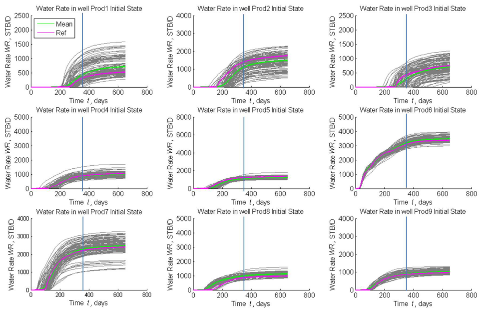
(b) Production water rates after first iteration



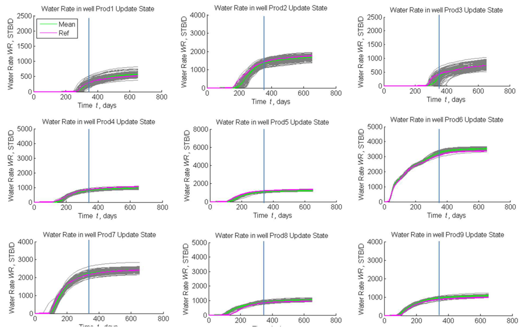
(c) Production water rates after first re-sampling



(d) Production water rates after second iteration



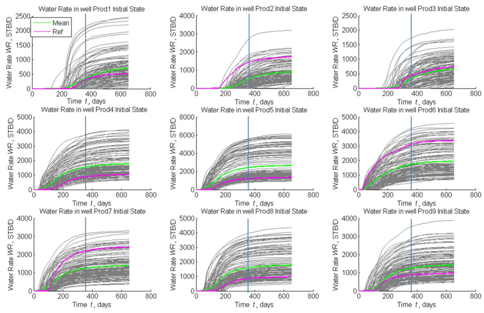
(e) Production water rates after second re-sampling



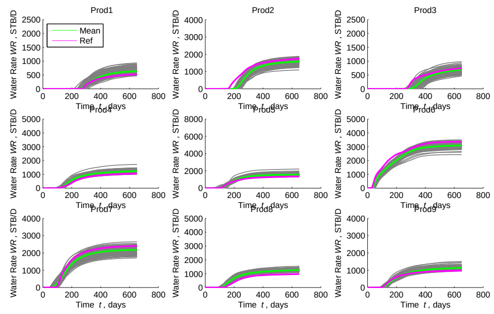
(f) Production water rates after third iteration

***Figure 11*** *Water production rates for the parameterization with gravity centers.*

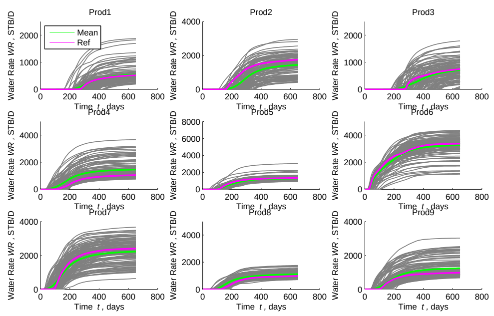
history matching method. Two of parameterization were newly introduced, based on extensions of a parameterization designed in a previous article. If in the older parameterization the randomness was ensured only by the initial ensemble of channelized reservoir, here we enhanced the randomness by the means of a stochastic forcing involving the use of some random seeds. In the first parameterization introduced, we use different random seeds when sampling from a conditional distribution for each grid cell of the domain and for each ensemble member. In the second parameterization introduced, we have used the same random seed when parameterizing an ensemble member but changing the seeds when parameterizing different members. In this way this parameterization keeps a spatial dependence structure inherited from the training image whilst the first parameterization does not. For the history matching part we have used the iterative adaptive Gaussian mixture filter in three iterations for its ability to deal with highly non-linear systems and for the re-sampling step between iterations. The results show that the dependence spatial structure is crucial in keeping the geological plausibility during iterations while the channel distribution into the fields was equally well estimated by all parameterizations. This means that the uncontrolled use of the stochastic forcing does not help in uncertainty quantification of the channelized reservoirs. In addition, the stochastic forcing used controlled, even though enhances the variability in the initial ensemble of the parameter field and in the updated ensemble of facies fields,



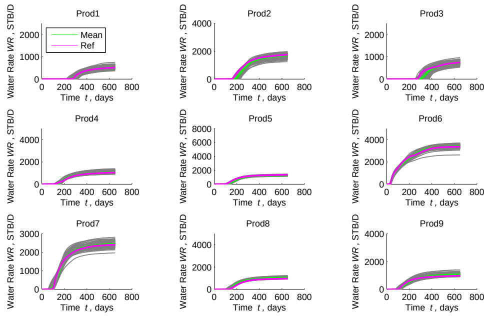
(a) Production water rates from initial ensemble



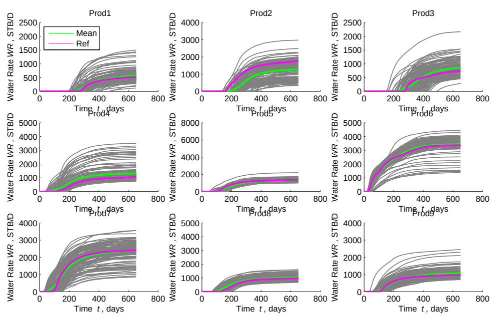
(b) Production water rates after first iteration



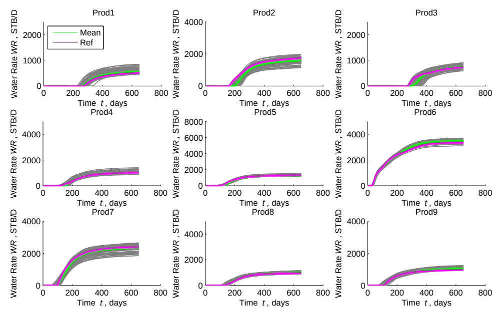
(c) Production water rates after first re-sampling



(d) Production water rates after second iteration



(e) Production water rates after second re-sampling



(f) Production water rates after third iteration

***Figure 12*** *Water production rates for the parameterization with random seeds.*

seems to not improve the channel estimation and the data match and predictions, showing the same trend as the older parameterization. All of these proves the robustness of the parameterizations that account for the spatial dependence structure inherited from the training image.

## References

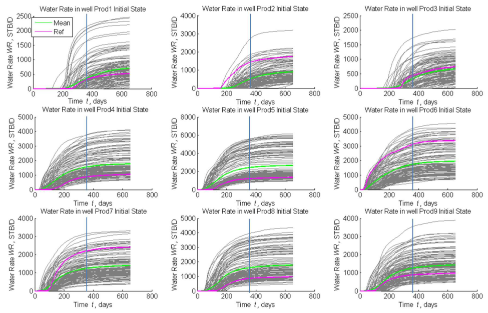
Caers, J. and Zhang, T. [2004] Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models.

Deutsch, C.V. and Wang, L. [1996] Hierarchical object-based stochastic modeling of fluvial reservoirs. *Mathematical Geology*, **28**(7), 857–880.

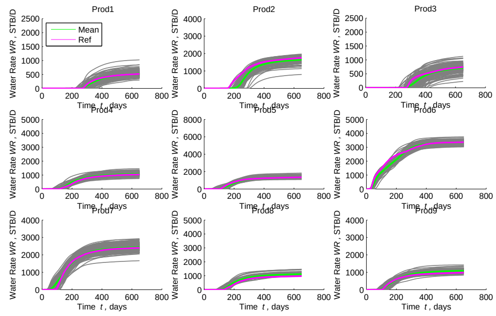
Evensen, G. [2003] The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, **53**(4), 343–367.

Hu, L., Zhao, Y., Liu, Y., Scheepens, C. and Bouchard, A. [2013] Updating multipoint simulations using the ensemble Kalman filter. *Computers & Geosciences*, **51**, 7–15.

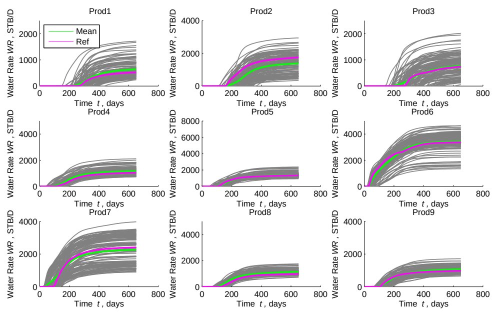
Jafarpour, B. and Khodabakhshi, M. [2011] A probability conditioning method (PCM) for nonlinear flow data integration into multipoint statistical facies simulation. *Mathematical Geosciences*, **43**(2),



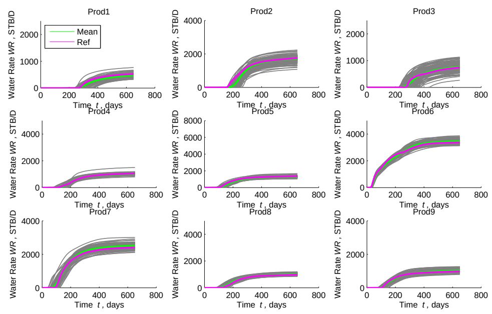
(a) Production water rates from initial ensemble



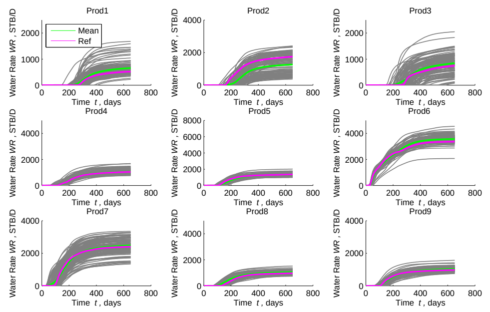
(b) Production water rates after first iteration



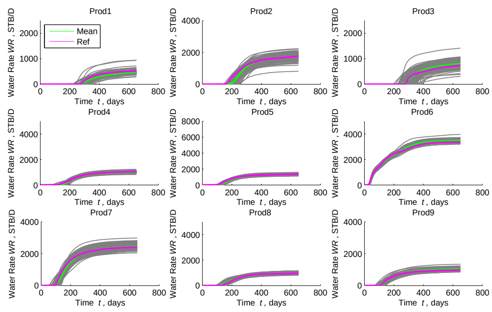
(c) Production water rates after first re-sampling



(d) Production water rates after second iteration



(e) Production water rates after second re-sampling



(f) Production water rates after third iteration

***Figure 13*** *Water production rates for the parameterization with fixed seeds.*

133–164.

Jafarpour, B. and McLaughlin, D.B. [2008] History matching with an ensemble Kalman filter and discrete cosine parameterization. *Computational Geosciences*, **12**(2), 227–244.

Journel, A. [2002] Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses. *Mathematical geology*, **34**(5), 573–596.

Krishnan, S., Boucher, A. and Journel, A.G. [2005] Evaluating information redundancy through the tau model. In: *Geostatistics Banff 2004*, Springer, 1037–1046.

Lorentzen, R.J., Flornes, K.M. and Nævdal, G. [2012] History Matching Channelized Reservoirs Using the Ensemble Kalman Filter. *SPE Journal*, **17**(1), 137–151.

Sarma, P., Durlofsky, L.J. and Aziz, K. [2008] Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. *Mathematical Geosciences*, **40**(1), 3–32.

Sebacher, B., Stordal, A. and Hanea, R. [2015] Bridging Multi Point Statistics and Truncated Gaussian Fields for Improved Estimation of Channelized Reservoirs with Ensemble Methods. *Computational Geosciences, DOI:10.1007/s10596-014-9466-3*, **19**(2), 341–369.

Sebacher, B., Stordal, A. and Hanea, R. [2016] Complex geology estimation using the iterative adaptive Gaussian mixture filter. *Computational Geosciences, DOI 10.1007/s10596-015-9553-0*, **20**(1), 133–

*__Table 1__ The RMSE calculated for total rates*

| | Gravity center | | Random seed | | Fixed seed | |
|---|---|---|---|---|---|---|
| | Initial | Updated | Initial | Updated | Initial | Updated |
| Iteration 1 | 1147.32 | 204.16 | 1147.32 | 267.44 | 1147.32 | 192.76 |
| Iteration 2 | 410.51 | 151.58 | 512.52 | 154.30 | 488.68 | 152.22 |
| Iteration 3 | 287.05 | 131.25 | 446.25 | 142.07 | 364.32 | 139.52 |

*__Table 2__ The RMSE calculated for permeability*

| | Gravity center | | Random seed | | Fixed seed | |
|---|---|---|---|---|---|---|
| | Initial | Updated | Initial | Updated | Initial | Updated |
| Iteration 1 | 5.201 | 4.781 | 5.201 | 5.059 | 5.201 | 5.059 |
| Iteration 2 | 4.745 | 4.444 | 4.982 | 4.794 | 4.984 | 4.655 |
| Iteration 3 | 4.475 | 4.283 | 4.794 | 4.566 | 4.614 | 4.371 |

148.

Stordal, A. and Lorentzen, R. [2014] An Iterative Version of the Adaptive Gaussian Mixture Filter. *Computational Geosciences*, **18**(3), 579–595.

Stordal, A.S., Karlsen, H.A., Nævdal, G., Skaug, H.J. and Vallès, B. [2011] Bridging the ensemble Kalman filter and particle filters: the adaptive Gaussian mixture filter. *Computational Geosciences*, **15**(2), 293–305.

Strebelle, S. [2002] Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, **34**(1), 1–21.

Vo, H.X. and Durlofsky, L.J. [2015] Data assimilation and uncertainty assessment for complex geological models using a new PCA-based parameterization. *Computational Geosciences*, **19**(4), 747–767.

Zhang, Y., Oliver, D., Chen, Y. and Skaug, H. [2015] Data Assimilation by Use of the Iterative Ensemble Smoother for 2D Facies Models. *SPE Journal*, **20**(1), 169–185.

Zhou, H., Li, L., Franssen, H.J.H. and Gómez-Hernández, J.J. [2012] Pattern recognition in a bimodal aquifer using the normal-score ensemble Kalman filter. *Mathematical Geosciences*, **44**(2), 169–185.