IRIS

International Research Institute of Stavanger

www.iris.no

Kristian Thulin, IRIS

Geir Nævdal, IRIS

Hans Julius Skaug, Department of Mathematics, University of Bergen

Sigurd Ivar Aanonsen, CIPR, University of Bergen

Quantifying Monte Carlo uncertainty in ensemble Kalman filter

Report IRIS - 2009/001

Project number:	7101502
Project title:	Continuous model updating using the ensemble Kalman filter with emphasis on complex reservoirs
Client(s):	Research Council of Norway
Distribution restriction:	Open

Bergen, 09.01.2009

Geir Nævdal — Sign.date
Project Manager

Kristin Flornes — Sign.date
Project Quality Assurance

Aina M. Berg — 19/1/09 Sign.date
Sr. Vice President

Preface

This report is presenting results obtained during Kristian Thulin PhD study, and is a slightly modified form of a paper submitted to SPE Journal. Kristian Thulin did most of his portion of the work while being a PhD student at CIPR, University of Bergen.

Résumé

The ensemble Kalman filter (EnKF) is currently considered one of the most promising methods for conditioning reservoir simulation models to production data. The EnKF is a sequential Monte Carlo method based on a low rank approximation of the system covariance matrix. The posterior probability distribution of model variables may be estimated from the updated ensemble, but because of the low rank covariance approximation, the updated ensemble members become correlated samples from the posterior distribution. We suggest using multiple EnKF runs, each with smaller ensemble size to obtain truly independent samples from the posterior distribution. This allows a point-wise confidence interval for the posterior cumulative distribution function (CDF) to be constructed. We present a methodology for finding an optimal combination of ensemble batch size (n) and number of EnKF runs (m) while keeping the total number of ensemble members ($m \times n$) constant. The optimal combination of n and m is found through minimizing the integrated mean square error (MSE) for the CDFs and we choose to define an EnKF run with 10.000 ensemble members as having zero Monte Carlo error. The methodology is tested on a simplistic, synthetic 2D model, but should be applicable also to larger, more realistic models.

Thanks to contributors.

The authors acknowledge the Research Council of Norway (PETROMAKS) for financial support.

Bergen, January 9, 2009.

Geir Nævdal, Project Manager

Contents

1 Introduction

The ensemble Kalman filter (EnKF) is a sequential Bayesian data assimilation method which estimates the posterior probability density function, or equivalently the posterior cumulative distribution function (CDF), of uncertain model variables and parameters. It has become increasingly popular in a large field of applications since its introduction in 1994 ([2]). In the limit of an infinite ensemble, the EnKF samples the posterior distribution exactly for a linear model with Gaussian statistics. For the linear and Gaussian problem this is the minimum variance solution. In practice, with a finite ensemble size, randomness is introduced in the drawing of the initial ensemble. When generating the initial ensemble, a random seed has to be chosen, on which the final estimates will be dependent. The final estimate will also depend on the randomness of the perturbations of the observed data at every analysis step. For a non-linear or non-Gaussian problem the posterior CDF will be approximate, but ideally the ensemble members from a finite EnKF run should be independent samples from the approximate solution from an infinite EnKF run. If this is satisfied, it follows that ensembles from repeated EnKF runs should be samples from the same distribution. For more details on the ensemble Kalman filter see e.g. [3, 4]

In previous studies we have observed that when repeating EnKF runs, using different random seeds to generate initial ensembles, the end results look visually different. To investigate this further we applied the two-sample Kolmogorov-Smirnov test [1, pp. 392-394] to pairs of EnKF runs, with the null hypothesis that the two samples are from the same distribution. Even in simplistic examples we found that the null hypothesis was generally rejected, leading to the conclusion that repeated EnKF runs might give estimates that are inconsistent with each other. We will show that this phenomenon is caused by strong positive correlations among ensemble members. Therefore, the independence assumption of the Kolmogorov-Smirnov test is not satisfied. Further, the lack of independence makes it difficult to quantify the simulation uncertainty in the estimate of the posterior distribution.

We propose to do multiple ensemble Kalman filter runs, each with a smaller ensemble size, to increase the amount of independent information. This also enables us to better quantify the Monte Carlo uncertainty in the final estimate of the posterior CDF. We further propose a methodology for choosing an optimally combination of the number of ensemble members (n) and the number of runs (m), under the constraint that the total resources ($n \times m$) is fixed. We know that if n is too small, the EnKF will collapse and there will be large biases in the final estimate. On the other hand, if m is too small then the variance on the estimated CDF will be very large. We seek a point which balances the two, and minimizes the mean square error (MSE). The MSE is defined relative to the infinite EnKF run, and not relative to the true posterior CDF.

In the literature various ensemble sizes have been used. An ensemble size of 100 was typically used in the first reservoir applications (e.g. [7]), based on experience from other fields (see e.g. [6]). Joint estimation of parameters and state has been done with as few as 40 members with satisfactory results in [5]. In a previous study using a 2D five spot problem for single-phase flow it was concluded that an ensemble size of around 20 is sufficient to reproduce the observations, but a much larger ensemble is needed to have a good estimate for the uncertainty [11]. For another synthetic 2D five spot problem, the conclusion was that at fairly large ensemble is required to estimate the uncertainty in the estimates accurately [12].

The paper is organized as follows. We start by a description of the synthetic 2D model in Section 2 followed by an explanation of the two-sample Kolmogorov-Smirnov test in Section 3.

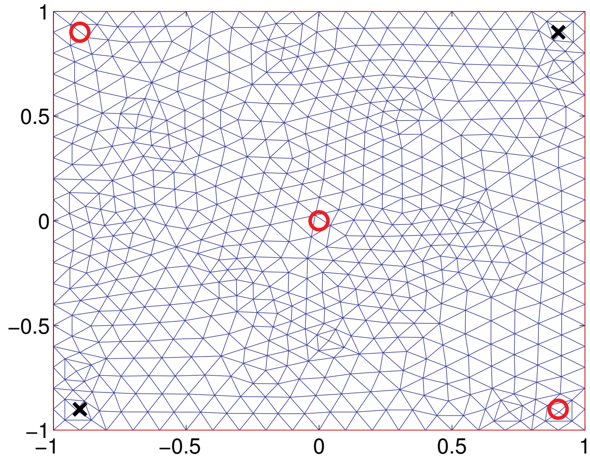

Figure 1: The triangular grid over the domain. Circles: sources, crosses: sinks.

The methodology of multiple EnKF runs, confidence interval and optimization of MSE is given in Section 4, and then results from the methodology applied to specific examples are presented in Section 5. Finally summary and conclusions are given in Section 6.

2 Description of the synthetic 2D-case

In this paper we consider a model for single-phase flow in a homogeneous medium. The single-phase reservoir equation for this case then becomes

$$\nabla \cdot \left(\frac{\kappa}{\mu} \nabla p \right) = \lambda \phi \frac{\partial p}{\partial t}, \tag{1}$$

where κ is permeability , μ is the viscosity, λ is the compressibility, ϕ is the porosity and p is pressure. In the combined parameter and state estimation problem, we would like to focus on a single dynamic variable and a single unknown parameter. For this reason further scaling assumptions are done to end up with the following model equation

$$\frac{\partial p}{\partial t} - \nabla \cdot (\kappa \nabla p) = f, \tag{2}$$

where p is the dynamic variable, and $\kappa = \kappa(x, y)$ is the unknown static parameter of the model. We assume that the unknown parameter κ can be represented by a two dimensional Gaussian random field. The right hand side has constant non-zero values at a five-spot pattern with three sources and two sinks (see Figure 1)

$$f(x,y) = \begin{cases} 1/2 & (x,y) = (0,0) \\ 1 & (x,y) = (-0.9, 0.9) \\ 2 & (x,y) = (0.9, -0.9) \\ -1 & (x,y) = (0.9, 0.9) \\ -1 & (x,y) = (-0.9, -0.9) \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The forward problem is solved by the finite element solver *parabolic* from the *Partial Differential Equation Toolbox* by MATLAB [9]. The square domain has a triangular discretization with 665 nodes and 1248 triangles (see Figure 1). The dynamic variables are calculated at the nodes while the static variable is calculated in the center of each triangle.

2.1 Reference data set

The unknown parameter $\kappa(x, y)$ has a prior distribution given by a Gaussian variogram and correlation length equal to half of the length of the reservoir. The mean is 4.0 over the whole domain. A reference Gaussian random field is drawn from this prior distribution. Artificial measurements are generated at uniform time steps by running the forward model on the reference field and extract values for the dynamic variables at the locations of the sinks and sources. The values are then perturbed with a noise with zero-mean and variance of 1.0×10^{-10} to generate synthetic observations. The same set of observations is used for all EnKF runs. The initial ensembles of realizations of $\kappa(x, y)$ are also drawn from the same prior distribution. For more details on the synthetic example, see [10, 11].

3 Kolmogorov-Smirnov test

In statistics the two-sample Kolmogorov-Smirnov test is used to determine whether two empirical CDFs differ, or equivalently if two samples can be from the same underlying distribution (see e.g. [1, pp. 392-394]). The empirical CDF for a sample $\{x_i\}_{i=1}^n$ is defined as

$$\mathrm{cdf}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I\left(x_i \leq x\right), \tag{4}$$

where $I(\cdot)$ is the indicator function. Given another sample $\{\widetilde{x}_i\}_{i=1}^n$ of same size, we want to test the null hypothesis that $\{x_i\}_{i=1}^n$ and $\{\widetilde{x}_i\}_{i=1}^n$ come from the same distribution. The test statistic of the two-sample Kolmogorov-Smirnov test is then given as

$$D_n = \sup_x \left| \mathrm{cdf}_n(x) - \widetilde{\mathrm{cdf}}_n(x) \right|, \tag{5}$$

where cdf_n and $\widetilde{\mathrm{cdf}}_n$ are the two respective empirical distribution functions. The null hypothesis is rejected when D_n is larger than some critical value, that can be looked up in a table. There are several equivalent formulations of the table of critical values for the Kolmogorov-Smirnov test, one that is often used can be found in [1].

The Kolmogorov-Smirnov test assume that the x's that constitute the sample $\{x_i\}_{i=1}^n$ are independent, and similarly that the members of $\{\widetilde{x}_i\}_{i=1}^n$ are independent. In particular, the calculation of the critical value of the test relies on this assumption. Failure of this assumption may lead to a significant p-value (less than 0.05) even if the null hypothesis is true (i.e. x_i and $\widetilde{x}_i$ come from the same distribution).

Examples of prior and updated EnKF samples for a static parameter in the 2D case from Chapter 2 are shown in Figure 2. The spread among the CDFs to the right is clearly larger than the spread among those to the left. To investigate this further we apply pair-wise Kolmogorov-Smirnov tests (significance level 5%) separately to the example data in the left and right panel of Figure 2. The number of rejected pair-wise comparisons are 2 out of 45 times for the unconditioned ensembles (left panel), and 28 out of 45 times for the conditioned

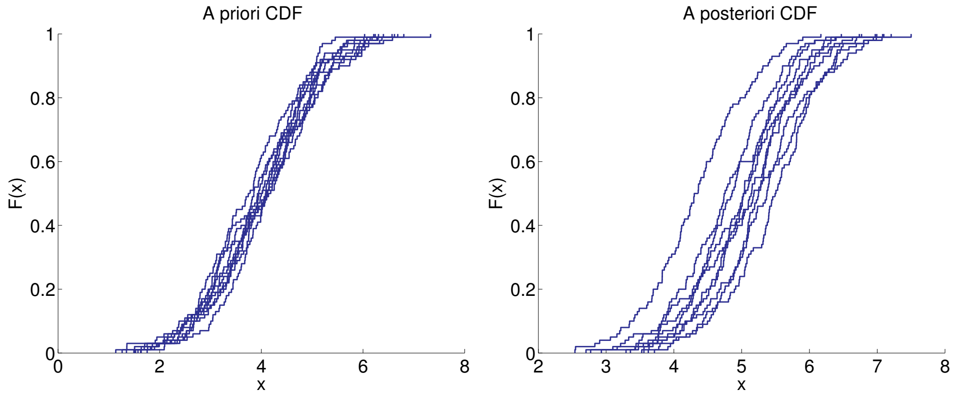

Figure 2: Ten CDF estimates before and after conditioning on data, for the synthetic 2D case. The posterior CDF estimates (right) are updated using the EnKF with 100 ensemble members.

ensemble members (right panel). A rejection of 2 out of 45 pairs is to be expected with a 5% significance level. The higher rejection rate for the right hand panel is due to the lack of independence between ensemble members within each of the CDF curves. As mentioned above, the lack of independence may cause rejection by the Kolmogorov-Smirnov test even if each individual ensemble member are actually drawn from the target (infinite EnKF run) posterior distribution.

To confirm that the 28 rejections out of 45 tests can not occur by chance we use a Bonferroni correction, which is a multiple-comparison correction used when several statistical tests are being performed simultaneously. The use of the Bonferroni method guaranties a simultaneous significance level of 5%, i.e. the probability of having at least one rejection is 0.05 when all the null hypotheses are true. In our situation the Bonferroni correction amounts to dividing the significance level 0.05 by the number of tests performed. For the case in Figure 2 we have $10(10-1)/2 = 45$ pair-wise comparisons in each panel, which means that we should apply a significance level of $0.05/45 = 0.0011$ to each individual test. For the CDFs to the left, the minimum p-value is 0.0082, well above the critical value of 0.0011. For the CDFs to the right, many p-values were below the critical value and the minimum p-value was as low as 3.96×10^{-16}. This clearly shows that they can not all come from the same distribution by chance.

4 Methodology

To identify the extent of the inconsistencies, ten EnKF runs are done starting with different independently sampled initial ensembles. The posterior variables after the second data assimilation step are then examined. For each of the ten runs, the 45 pairings for all 665 dynamic and 1248 static variables are checked using the Kolmogorov-Smirnov test with the Bonferroni correction. For the smaller ensemble sizes (20 and 50) the Kolmogorov-Smirnov test fails for almost all variables, both static and dynamic. Even for an ensemble size as large as 1000, a large part of the variables have the observed inconsistencies. The extent of failures for the Kolmogorov-Smirnov test for 200 ensemble members can be seen in Figure 3. This figure shows the *lowest* from the 45 p-values for each variable, plotted spatially. From the

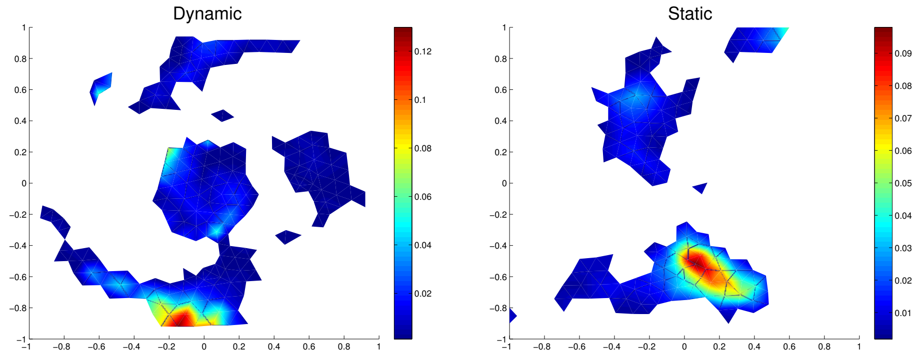

Figure 3: Spatial variation of p-values for the Kolmogorov-Smirnov test with Bonferroni correction in a case with 200 ensemble members. Regions lower than the critical value, are blank in the plot. Dynamic variables are shown to the left and static variables are shown to the right.

Bonferroni correction we know that the null hypothesis can be rejected with a 95% confidence if the lowest p-value is below $0.05/45 = 0.0011$. For the static variable the areas furthest away from the measurement locations stays consistent, while the variables closest become inconsistent. The reason for this is that the variables far away from the well is not very sensitive to the observed data, and hence is not updated. Because the initial ensembles are sampled independently from the prior, the independence assumption made in the Kolmogorov-Smirnov test is fulfilled approximately. From the dynamic plot the same behavior is observed. In addition we have a region with non-significant p-values around the center well. This can be explained by the fact that the dynamic variables in this area are described up to the measurement error. The measurement error is also independently drawn, and should produce consistent estimates according to the Kolmogorov-Smirnov test. These plots show the extent of the inconsistencies after only two data assimilation steps, as more data gets assimilated larger areas will have p-values below the critical value.

4.1 Point-wise confidence band on the posterior CDF

We have observed severe inconsistencies even for the relatively simple problem described in Section 2. Motivated by this, and the fact that people have reported satisfying results that match the history with as few as 40 ensemble members, we suggest doing several EnKF runs, each with fewer ensemble members. This strategy provides independent samples from the posterior. After a single run with the ensemble Kalman filter, with ensemble size $n = 200$, one has only a single estimate of the posterior CDF. Using a different random seed when generating each initial ensemble, the estimated posterior CDFs will differ. With only a single run one has no estimate of precision for the estimated posterior CDF.

Our suggestion is then, instead of doing one large run with $n = 200$ members, to do e.g. $m = 10$ runs with $n = 20$ members each, which has the same computational cost. In the following we fix a variable (static or dynamic), and let x denote the different values of the variable. The mean of the ten runs is then the new estimate of the posterior CDF:

$$\overline{\mathrm{cdf}_n}(x) = \frac{1}{m}\sum_{j=1}^{m} \mathrm{cdf}^{(j)}(x), \tag{6}$$

where $\text{cdf}^{(j)}(x)$ is the version of (4) calculated from ensemble number j. Similarly, we calculate the empirical standard deviation $s(x)$ from the m ensembles. A standard $100(1-\alpha)\%$ confidence interval for the true $\text{cdf}(x)$ is then

$$\left[\overline{\text{cdf}}(x) - t^* \frac{s(x)}{\sqrt{m}} \quad \overline{\text{cdf}}(x) + t^* \frac{s(x)}{\sqrt{m}}\right] \tag{7}$$

where t^* is the upper $\alpha/2$-quantile of the t-distribution with $m-1$ degrees of freedom. When calculated for different values of x, (7) constitute a point-wise confidence band for the function $\text{cdf}(x)$. The next step is, given the total number of resources $(n \times m)$, to select a reasonable number of ensemble members (n) and number of runs (m).

4.2 Optimal combination of ensemble size (n) and number of runs (m)

We want to find an optimal combination of n and m under the constraint that the computational effort $(n \times m)$ is fixed. With too few ensemble Kalman filter runs the number of samples are very low, and this results in a very large uncertainty on your estimated CDF. On the other hand, with a too small ensemble size the EnKF will not give satisfying results and there will be a large bias in the end estimate. The goal when doing repeated runs is to balance the two terms and minimize the *mean square error* (MSE) defined generally as

$$\text{MSE}(x) = \text{Variance}(x) + \text{Bias}(x)^2. \tag{8}$$

The bias for a given x-value is the absolute distance to the CDF for infinite run. For a given x-value and with m EnKF runs, the bias term is

$$\text{Bias}(x) = |\text{cdf}_\infty(x) - \overline{\text{cdf}_n}(x)|, \tag{9}$$

where $\text{cdf}_\infty(x)$ is the CDF for the infinite ensemble run and $\overline{\text{cdf}}(x)$ is given by (6).

The variance term is defined as an inflated variance

$$\text{Variance}(x) = \frac{(z^*)^2 s^2(x)}{(t^*)^2 \times m}, \tag{10}$$

where z^* is the upper $\alpha/2$-quantile of the normal distribution. The inflated variance measure is used here to account for the increased uncertainty when there are very few samples (m is low). For m higher than 10, $t^* \approx z^*$. See Fig. 4 for a plot of the upper critical values of the student's t-distribution as a function of degrees of freedom.

Finally, as scalar measure of approximation error we use the integrated mean square error (IMSE),

$$\text{IMSE} = \int_{\infty}^{-\infty} \text{MSE}(x)\, dx. \tag{11}$$

If this is done for several combinations of n and m, where $n \times m$ is kept constant, we could find an optimal combination that minimizes the mean square error for the given example.

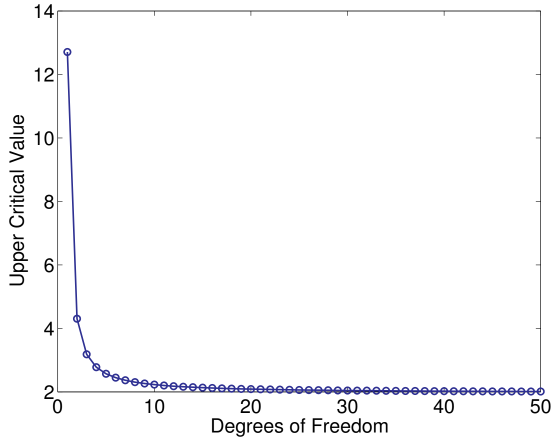

Figure 4: Critical values for the student's t-distribution as a function of degrees of freedom. The degrees of freedom in m samples are $m-1$.

Table 1: ensemble sizes and # runs	
m (runs)	n (members)
1	1000
2	500
3	333
4	250
5	200
10	100
20	50
50	20
100	10
200	5

5 Examples

The simple single-phase flow 2D model with 665 dynamic and 1248 static variables described in Section 2 is run for two assimilation steps. Rather than comparing the estimated fields after the two assimilations with the reference fields, we want to compare the estimated CDFs with the corresponding estimated CDF from an infinite solution. In this case we define a single large run with 10.000 ensemble members is defined as the *infinite run*, and the estimated CDFs from the infinite run as the *infinite solution*. In the two following examples the combination of ensemble size n and the number of runs m are varied while $n \times m$ is kept equal to 1000 and 200, respectively. Out of the 1913 variables we have chosen to look at only a single static parameter located at the center of the domain (middle circle in Fig. 1). Since this point is also a data-point, there is information in the data to update the static variable during the data assimilation steps.

5.1 $n \times m = 1000$

The different ensemble sizes employed are given in Table 1. Plots of the estimated CDFs for

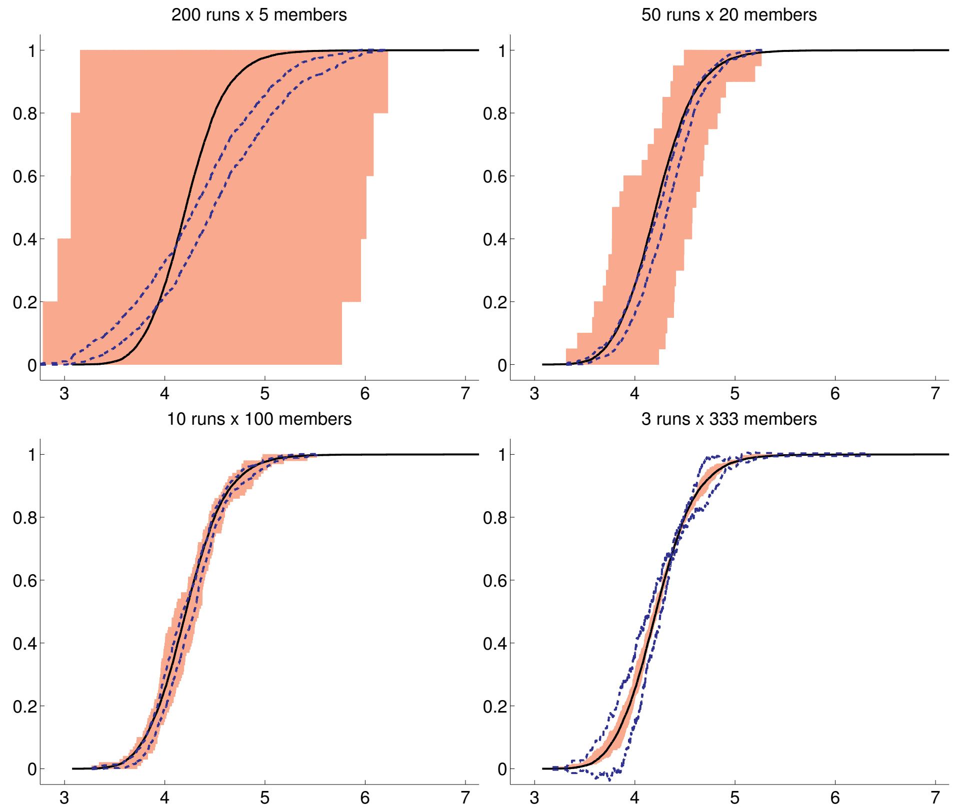

Figure 5: CDF for the static variable for the reference data set. The dashed lines outline the 95-% confidence interval. The lighter background indicates the span of all the runs, and the solid line is the CDF of the run with the infinite ensemble (10.000 members).

different combinations of n and m for the static variable are shown in Figure 5, the plot for the dynamic variable is similar. In these plots the light background outlines the span of all the CDFs from the m runs. For clarity, the average is not plotted, but a 95% confidence interval band is shown with dashed lines. The solid line shows the infinite run. For the upper left plot with 200 runs of only 5 ensemble members, we see that the span of the estimated CDFs are very wide, and that the covariance interval band does not at all cover the infinite run. For 50 runs with 20 ensemble members the estimates have improved a lot and the span of the different runs are narrower. The infinite run is just on the boundary of the confidence interval band. In the case of 10 runs with 100 ensemble members, the interval has narrowed even further and it now contains the infinite run perfectly. With only 3 runs with 333 members it is evident that the confidence interval band becomes much wider because of the low number of runs.

With only 5 ensemble members we observe a systematic bias in the confidence interval, relative to the infinite solution CDF. This means that an ensemble size of 5 is too small for the EnKF. To investigate this closer we have plotted the average CDF over 50 runs for each combinations of n and m in Figure 6. By averaging over 50 runs we eliminate most of the randomness, and a systematic trend in the bias is clearly visible in the figure. For our simplistic example the bias is clear for 5 and 10 members, and ensemble sizes above 20 seems enough that the EnKF can give satisfying results on average.

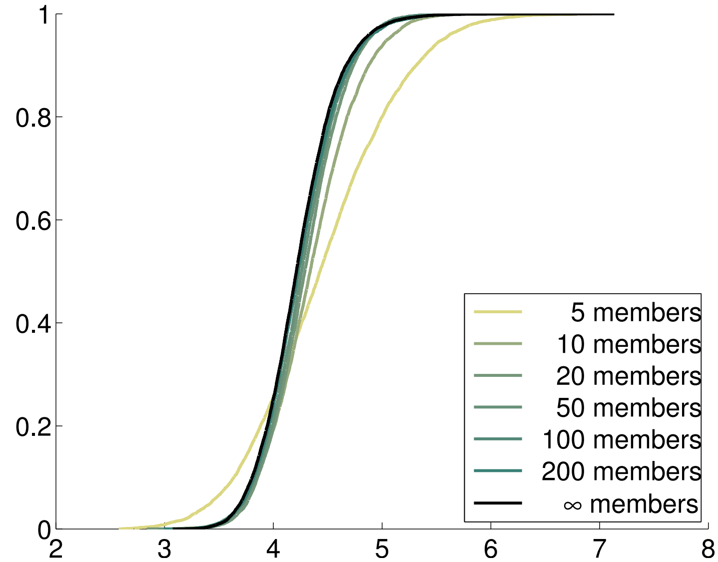

Figure 6: Means of CDF for different ensemble sizes (static variable, reference data).

Table 2: Ensemble sizes and runs	
m (runs)	n (members)
1	200
2	100
4	50
10	20
20	10
40	5

The IMSE values (see Section 4.2) are plotted in Figure 7. It clearly shows that the bias grows when n is decreasing, and that the variance is growing when n increases (and therefore m decreases). For our case, with $n \times m = 1000$, it looks like n should be at least somewhere between 50 and 100 before the bias is at an acceptable level. From the variance in Figure 7, and from the plot of the critical values for the student's t-distribution in Figure 4, we can further conclude that m should be larger than approximately 5 in this case.

From the monotonic trends of both the bias and the variance curves plot in Figure 7 it looks like there should be a minimum value of the IMSE function for a specific choice of n, but with only this few points on the x-axis it is difficult to conclude any further. Also, the IMSE curve is almost constant at the *optimal interval* and the ultimate result will not be sensitive to small changes in the choice of n.

5.2 $n \times m = 200$

The optimal value of the IMSE will be dependent on $n \times m$, therefor it is informative to repeat the same exercise with a different restriction on the resources. One might also argue that 1000 ensemble members are too much for a real reservoir model. For these reason, we have done the same experiment as in the previous example, only this time with $n \times m = 200$, closer to what is typically being used. The selected combination of n and m are shown in Table 2.

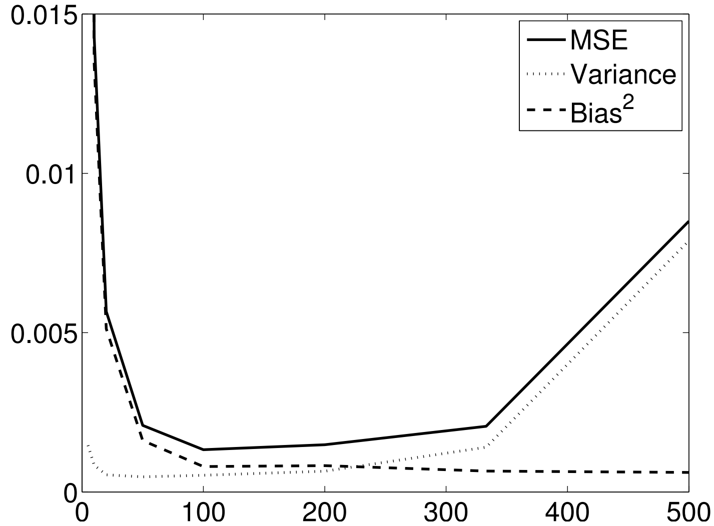

Figure 7: The integrated values for the bias, variance and mean square error as a function of n.

Plots of the CDFs for the different combinations of n and m are shown in Figure 8. Similarly to Example I, we see that with only 5 ensemble members the mean estimate is not close to the run with the infinite (10.000 members) EnKF run. With 10 ensemble members the results are considerably better. The reference infinite case is just on the border of the confidence interval as in the case for 1000 ensemble members, the confidence interval is also narrower than for 5 members. The case with 5 runs of 40 ensemble members shows further improvement in the results. The confidence interval is marginally narrower, but more importantly it now captures the infinite case. With only 3 runs of 66 members, the confidence interval increases because of the few number of runs. The behavior of the biases in the means are the same as shown in Figure 6.

The integrated mean square error plot are shown in Figure 9 and the tendency that the bias increases with low n (or high m) and that the variance increases with high n (or low m) is clear also here. For this example we can see that n should be at least 20, and m should be at least 3-4, with in this interval it is difficult to conclude any further. Looking at the two MSE plots, it is worth noticing that the value of the MSE is approximately a factor 5 higher using a total of 200 ensemble compared to using 1000 ensemble members.

6 Summary and Conclusions

The ensemble Kalman filter is a sequential Monte Carlo method that uses a finite ensemble to approximate the covariance matrix during each data assimilation update. These approximations will yield a different posterior estimate depending on the initial ensemble. It has been observed that the amount of variation is substantial and significant by the standards of a Kolmogorov-Smirnov test. The excess variation has been identified to be caused by ensemble members being correlated samples from the posterior distribution. Markov Chain Monte Carlo ([8]), which is another popular Monte Carlo method, is also plagued with the problem that the generated samples are correlated.

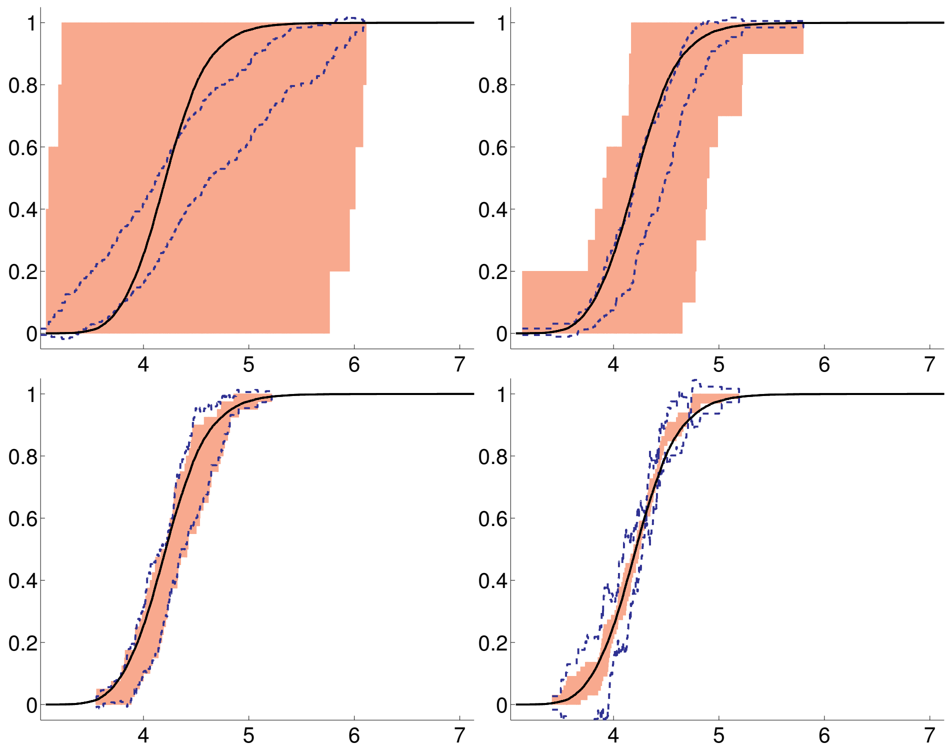

Figure 8: CDF plots for the static variable. The dashed lines outline the 95-% confidence interval. The lighter background indicates the span of all the runs, and the solid line is the CDF of the run with the infinite ensemble (10.000 members).

As a solution we have suggested to do several EnKF runs, each with fewer ensemble members, rather that one large run. In this way more independent information about the posterior CDF is obtained. Also we have presented a methodology and some general motivations for finding an optimal choice for the combination of the number of ensemble members (n) and the number of EnKF runs (m), keeping their product ($n \times m$) constant. This is done by minimizing the integrated mean square error.

One of the main points to be made here is that with only a single EnKF run, there is no control over the Monte Carlo error in the estimated CDF. This means that making decisions based on a single ensemble might be difficult, as a new EnKF run with a different seed may give a completely different result.

The methodology is tested on a synthetic single-phase 2D example with a total of 1913 variables with two different resource restrictions, $n \times m = 1000$ and $n \times m = 200$. In both cases we noticed that using too few ensemble members resulted in a very large bias and therefore also a large MSE. The other extreme is using too few EnKF runs (low m); this results in a large variance and therefor also in a large MSE. The optimal combination is then using enough runs to have an acceptable variance, and also keep the ensemble size large enough so that the bias is also at an acceptable level. For the first case, with $m \times n = 1000$ it looked like an ensemble size of approximately 100 was sufficient, where are for the case where $m \times n = 200$ an ensemble size of around 20 was optimal. For both the minimum value of the integrated MSE and the minimum tolerated ensemble size, there is approximately a factor 5 difference between using a total of 1000 members and 200 members. The minimum number of runs

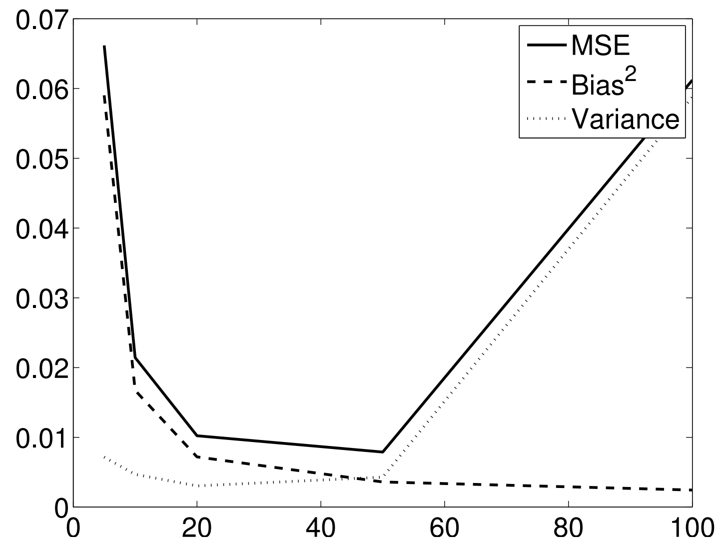

Figure 9: The integrated values for the bias, variance and mean square error as a function of n.

needed, on the other hand, is similar for the two cases, around 4 or 5. We believe that this lower bound can be used as a general guideline for more complicated problems, whereas the minimum number of ensemble members will be dependent on both the specific case and the total resources available.

References

[1] Chakravarti, Laha, and Roy. *Handbook of Methods of Applied Statistics, Volume I.* John Wiley & Sons, Inc., 1967.

[2] G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5):10,143–10,162, 1994.

[3] G. Evensen. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367, 2003.

[4] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag New York, Inc., 2006.

[5] Y. Gu and D. S. Oliver. History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter. *SPE Journal*, 10(2):217–224, 2005. 89942-PA.

[6] P. L. Houtekamer and H. L. Mitchell. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129:123–137, January 2001.

[7] G. Nævdal, T. Mannseth, and E. H. Vefring. Instrumented wells and near-well reservoir monitoring through ensemble Kalman filter. In *Proceedings of 8th European Conference on the Mathematics of Oil Recovery*, Freiberg, Germany, 3 - 6 September 2002.

[8] D. Oliver, L. B. Cunha, and A. C. Reynolds. Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology*, 29(1):61–91, 1997.

[9] http://www.mathworks.com.

[10] K. Thulin. Ensemble Kalman filter for field enstimation. Master thesis, Department of Mathematics, University of Bergen, 2005.

[11] K. Thulin and G. Nævdal. Ensemble Kalman filter for field estimation - investigation on the effect of the ensemble size. In *Proceedings of 10th European Conference on the Mathematics of Oil Recovery*, Amsterdam, The Netherlands, 4-7 September 2006.

[12] X.-H. Wen and W. H. Chen. Real-time reservoir model updating using ensemble Kalman filter with confirming option. *SPE Journal*, 11(4):431–442, December 2006. SPE92991-PA.