Comment on "Data Assimilation Using an Ensemble Kalman Filter Technique"

PETER JAN VAN LEEUWEN

IMAU, Utrecht University, Utrecht, the Netherlands

23 November 1997 and 8 June 1998

1. Introduction

In an interesting paper Houtekamer and Mitchell (1998; hereafter HM98) introduce a variant of the ensemble Kalman filter (EnKF) as proposed by Evensen (1994). HM98 point to the hitherto unacknowledged problem that the EnKF has an "inbreeding" problem: in the analysis step the ensemble is updated with a gain calculated from that same ensemble. In their new approach a double ensemble (DEnKF) is used and the gain of each ensemble is used to update the other ensemble in the analysis step. The new approach is argued to be much less sensitive to this inbreeding. They strengthen their argument by a specific example, in which the EnKF shows a variance that is too low for small ensemble sizes (≤ 100), but they fail to give a rigorous justification. The purposes of this comment are twofold. First, a theoretical justification of the inbreeding effect is given, and it is shown that the DEnKF has similar, but smaller, problems. Second, a serious concern about the use of small ensemble sizes is expressed, thus bringing into question the use of the DEnKF over the EnKF for real applications.

2. Covariances in ensemble Kalman filters

To analyze the effect of approximate knowledge of the error covariances before the analysis step in ensemble Kalman filters we follow Burgers et al. (1998). The prior and posterior error covariances are given by

$$\mathsf{P}^f = \overline{(\boldsymbol{\psi}^f - \boldsymbol{\psi}^t)(\boldsymbol{\psi}^f - \boldsymbol{\psi}^t)^{\mathrm{T}}} \tag{1}$$

$$\mathsf{P}^a = \overline{(\boldsymbol{\psi}^a - \boldsymbol{\psi}^t)(\boldsymbol{\psi}^a - \boldsymbol{\psi}^t)^{\mathrm{T}}}, \tag{2}$$

where the overline indicates an expectation value; $\boldsymbol{\psi}$ is the model state at a particular time; and the superscripts f, a, and t denote forecast, analyzed, and true state, respectively.

Because the truth is not known a convenient estimate of it must be defined. In ensemble Kalman filters the ensemble covariance is used, so the mean state is taken as an estimate for the truth, leading to

$$\mathsf{P}_e^f = \overline{(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f})(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f})^{\mathrm{T}}} \tag{3}$$

$$\mathsf{P}_e^a = \overline{(\boldsymbol{\psi}^a - \overline{\boldsymbol{\psi}^a})(\boldsymbol{\psi}^a - \overline{\boldsymbol{\psi}^a})^{\mathrm{T}}}. \tag{4}$$

Evensen [1994; but see also Burgers et al. (1998)] showed that the analyzed ensemble can be obtained from the old ensemble by updating each member i according to

$$\boldsymbol{\psi}_i^a = \boldsymbol{\psi}_i^f + \mathsf{K}_e(\mathbf{d}_i - \mathsf{H}\boldsymbol{\psi}_i^f), \tag{5}$$

in which K_e is the gain determined from the ensemble, defined as

$$\mathsf{K}_e = \mathsf{P}_e^f\mathsf{H}^{\mathrm{T}}(\mathsf{H}\mathsf{P}_e^f\mathsf{H}^{\mathrm{T}} + \mathsf{W})^{-1}, \tag{6}$$

and $\mathbf{d}_i$ is a vector obtained by adding a random vector to the data vector. This random vector is chosen from a Gaussian distribution with zero mean and data covariance W. (In passing, I note that the analysis step used in the Kalman filter is not optimal for a nonlinear model. The assumption made is that the error covariances are Gaussian distributed or that the model is linear. However, this issue is not the focus of this comment.)

The analyzed ensemble covariance can now be evaluated as

$$\begin{aligned}
\mathsf{P}_e^a &= \overline{(\boldsymbol{\psi}^a - \overline{\boldsymbol{\psi}^a})(\boldsymbol{\psi}^a - \overline{\boldsymbol{\psi}^a})^{\mathrm{T}}} \\
&= \overline{(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f} - \mathsf{K}_e\mathsf{H}(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f}) + \mathsf{K}_e(\mathbf{d} - \overline{\mathbf{d}}))(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f} - \mathsf{K}_e\mathsf{H}(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f}) + \mathsf{K}_e(\mathbf{d} - \overline{\mathbf{d}}))^{\mathrm{T}}} \\
&= \mathsf{P}_e^f - \mathsf{K}_e\mathsf{H}\mathsf{P}_e^f - \mathsf{P}_e^f\mathsf{H}^{\mathrm{T}}\mathsf{K}_e^{\mathrm{T}} + \mathsf{K}_e(\mathsf{H}\mathsf{P}^f\mathsf{H}^{\mathrm{T}} + \mathsf{W})\mathsf{K}_e^{\mathrm{T}} + (\mathsf{I} - \mathsf{K}_e\mathsf{H})\mathsf{D}_e + \mathsf{D}_e^{\mathrm{T}}(\mathsf{I} - \mathsf{K}_e\mathsf{H})^{\mathrm{T}} \\
&= (\mathsf{I} - \mathsf{K}_e\mathsf{H})\mathsf{P}_e^f + (\mathsf{I} - \mathsf{K}_e\mathsf{H})\mathsf{D}_e + \mathsf{D}_e^{\mathrm{T}}(\mathsf{I} - \mathsf{K}_e\mathsf{H})^{\mathrm{T}},
\end{aligned} \tag{7}$$

where the forecast–observation covariance is given by

$$\mathsf{D}_e = \overline{(\boldsymbol{\psi}^f - \overline{\boldsymbol{\psi}^f})(\mathbf{d} - \overline{\mathbf{d}})^{\mathrm{T}}\mathsf{K}_e^{\mathrm{T}}}. \tag{8}$$

Note that it is essential that each ensemble member

Corresponding author address: Peter Jan van Leeuwen, IMAU, Utrecht University, P.O. Box 80005, 3508 TA Utrecht, the Netherlands.
E-mail: leeuwen@fys.ruu.nl

is updated with different data to avoid an analyzed variance that is too low (see Burgers et al. 1998). At first sight the inbreeding effect might be compensated for by the fact that each ensemble member is updated with a different data vector. As we will see below, and anticipated by HM98, this is not the case. When the ensemble size increases, $\mathbf{K}_e$ will approach the true gain $\mathbf{K}$. Also, the model–data covariance will become negligible; hence $\mathbf{P}_e^a$ will tend to the true $\mathbf{P}^a$.

In the DEnKF the situation is as follows. Each ensemble member i of ensemble 1 is updated as

$$\boldsymbol{\psi}_{i1}^a = \boldsymbol{\psi}_{i1}^f + \mathbf{K}_2(\mathbf{d}_{i1} - \mathbf{H}\psi_{i1}^f) \tag{9}$$

and vice versa, in which $\mathbf{K}_2$ is the gain obtained from the other ensemble:

$$\mathbf{K}_2 = \mathbf{P}_2^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_2^f\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}. \tag{10}$$

The analyzed ensemble covariance of the first ensemble now becomes

$$\begin{aligned}\mathbf{P}_1^a &= \overline{(\boldsymbol{\psi}_1^a - \overline{\boldsymbol{\psi}_1^a})(\boldsymbol{\psi}_1^a - \overline{\boldsymbol{\psi}_1^a})^{\mathrm{T}}} = \ldots \\ &= (\mathbf{I} - \mathbf{K}_2\mathbf{H})\mathbf{P}_1^f - \mathbf{P}_1^f\mathbf{H}^{\mathrm{T}}\mathbf{K}_2^{\mathrm{T}} + \mathbf{K}_2(\mathbf{H}\mathbf{P}_1^f\mathbf{H}^{\mathrm{T}} + \mathbf{W})\mathbf{K}_2^{\mathrm{T}} \\ &\quad + (\mathbf{I} - \mathbf{K}_2\mathbf{H})\mathbf{D}_1 + \mathbf{D}_1^{\mathrm{T}}(\mathbf{I} - \mathbf{K}_2\mathbf{H})^{\mathrm{T}}\end{aligned} \tag{11}$$

and the forecast-observation covariance is given by

$$\mathbf{D}_1 = \overline{(\boldsymbol{\psi}_1^f - \overline{\boldsymbol{\psi}_1^f})(\mathbf{d}_1 - \overline{\mathbf{d}_1})^{\mathrm{T}}\mathbf{K}_2^{\mathrm{T}}}. \tag{12}$$

The extra terms compared to the EnKF are due to the fact that the gains of the two ensembles are not the same. Also, in this case the true $\mathbf{P}^a$ will be approached for increasing ensemble size because the two ensembles will obtain identical gains. At this point, I introduce

$$\delta = \mathbf{P}_1^f\mathbf{H}^{\mathrm{T}} - \mathbf{P}_2^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_2^f\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}(\mathbf{H}\mathbf{P}_1^f\mathbf{H}^{\mathrm{T}} + \mathbf{W}) \tag{13}$$

to simplify (11) to

$$\begin{aligned}\mathbf{P}_1^a &= (\mathbf{I} - \mathbf{K}_2\mathbf{H})\mathbf{P}_1^f - \delta\mathbf{K}_2^{\mathrm{T}} + (1 - \mathbf{K}_2\mathbf{H})\mathbf{D}_1 \\ &\quad + \mathbf{D}_1^{\mathrm{T}}(1 - \mathbf{K}_2\mathbf{H})^{\mathrm{T}}.\end{aligned} \tag{14}$$

3. Analysis of finite ensemble size effects

Let us neglect the forecast-model covariances $\mathbf{D}$ for the moment to clarify the discussion. We will come back to them later.

The prior covariances will differ from the true prior covariance due to finite ensemble effects. The mean state is taken for granted; we are interested in the error covariance of that mean state. The differences between the estimated and the true covariance around this mean state are denoted by ϵ_e, ϵ_1, and ϵ_2 for the three covariances in question. So,

$$\mathbf{P}_e^f = \mathbf{P} + \epsilon_e; \qquad \mathbf{P}_1^f = \mathbf{P}_1 + \epsilon_1; \qquad \mathbf{P}_2^f = \mathbf{P}_2 + \epsilon_2. \tag{15}$$

If we assume that the ensemble estimates are not too far off, $\|\epsilon\|$ is small (in some sense, not too important here). An ensemble Kalman gain can now be written as

$$\begin{aligned}\mathbf{K}_e &= \mathbf{P}_e^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1} \\ &= (\mathbf{P} + \epsilon_e)\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{H}\epsilon_e\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1} \\ &= \mathbf{K}[\mathbf{I} - \mathbf{H}\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1} \\ &\qquad + \mathbf{H}\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}] \\ &\quad + \epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}[\mathbf{I} - \mathbf{H}\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}] \\ &\quad + \mathbf{O}(\|\epsilon_e\|^3).\end{aligned} \tag{16}$$

Similar equations arise for $\mathbf{K}_1$ and $\mathbf{K}_2$. These resulting expressions are used in the analyzed ensemble covariances (7) and (11). We then obtain for the EnKF (neglecting the forecast–observation covariances)

$$\begin{aligned}\mathbf{P}_e^a &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P} + (\mathbf{I} - \mathbf{K}\mathbf{H})\epsilon_e(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}) \\ &\quad - (\mathbf{I} - \mathbf{K}\mathbf{H})\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_e(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}) \\ &\quad + \mathbf{O}(\|\epsilon_e\|^3).\end{aligned} \tag{17}$$

The first term on the rhs is the optimal posterior covariance in the Kalman filter framework. The second term is proportional to ϵ_e. Since this term can have any sign it will not lead to a bias initially. This error term is downweighted with a factor $(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}})$ compared to the covariance itself. The third term, however, is always negative, leading to a negative bias in the total variance in the standard ensemble Kalman filter. It is illustrative to see where these terms come from. The second term on the rhs can be written as

$$(\mathbf{I} - \mathbf{K}\mathbf{H})\epsilon_e - (\mathbf{I} - \mathbf{K}\mathbf{H})\epsilon_e\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}. \tag{18}$$

The first term is due to the error in $\mathbf{P}_e^f$ directly while the second is due to the error in $\mathbf{P}_e^f$ via the gain matrix. Note that these two terms have opposing tendency. The third term on the rhs can be decomposed as

$$\begin{aligned}&-(\mathbf{I} - \mathbf{K}\mathbf{H})\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_e \\ &\quad + (\mathbf{I} - \mathbf{K}\mathbf{H})\epsilon_e\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_e\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}.\end{aligned} \tag{19}$$

The first term in this expression is due to the interaction of the errors in $\mathbf{P}_e^f$ and the gain, while the second term is due to the nonlinearity of the gain. Again, these terms have opposing tendency. HM98 only argued about the first effect, while the nonlinearity of the gain tends to compensate for this inbreeding.

The effect of the finite ensemble size will be as follows. Initially the second-order error term, the third term on the rhs of (17), will tend to decrease the total variance compared to the Kalman filter analysis step. As soon as a negative bias is created that is not compensated for by the spread of the ensemble between analyses, the linear term will take over and decrease the bias even further, until the error becomes so large that our analysis fails.

The analysis in the DEnKF case is more complicated, but following the same approach we arrive at

$$\begin{aligned}\mathbf{P}_1^a = {} & (\mathbf{I} - \mathbf{K}_2\mathbf{H})\mathbf{P}_1 - \delta\mathbf{K}_2^{\mathrm{T}} + (\mathbf{I} - \mathbf{K}_2\mathbf{H})\epsilon_1(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}_2^{\mathrm{T}}) \\ & - \delta(\mathbf{HP}_2\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_2(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}_2^{\mathrm{T}}) - (\text{transpose}) \\ & - (\mathbf{I} - \mathbf{K}_2\mathbf{H})\epsilon_2\mathbf{H}^{\mathrm{T}}(\mathbf{HP}_2\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_1(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}_2^{\mathrm{T}}) \\ & - (\text{transpose}) + (\mathbf{I} - \mathbf{K}_2\mathbf{H})\epsilon_2\mathbf{H}^{\mathrm{T}}(\mathbf{HP}_2\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1} \\ & \times \mathbf{H}\epsilon_2(\mathbf{HP}_2\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\delta + (\text{transpose}) \\ & + (\mathbf{I} - \mathbf{K}_2\mathbf{H})\epsilon_2\mathbf{H}^{\mathrm{T}}(\mathbf{HP}_2\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}(\mathbf{HP}_1\mathbf{H}^{\mathrm{T}} + \mathbf{W}) \\ & \times (\mathbf{HP}_2\mathbf{H}^{\mathrm{T}} + \mathbf{W})^{-1}\mathbf{H}\epsilon_2(\mathbf{I} - \mathbf{H}^{\mathrm{T}}\mathbf{K}_1^{\mathrm{T}}) + \mathbf{O}(\|\epsilon_i\|^3).\end{aligned} \tag{20}$$

The first two terms on the rhs form the optimal posterior covariance in the Kalman filter framework. The third term is proportional to ϵ_1 and can have any sign. The fourth term (and its transpose) is proportional to $\delta\epsilon_2$, so it can also have any sign, being smaller than the third. An interesting term is the fifth. This term describes the imbreeding effect in the EnKF. In the DEnKF the term can have any sign because it is proportional to $\epsilon_1\epsilon_2$. The remaining two terms are due to the nonlinearity of the gain. The last will be positive and larger than the other term. This term will lead to a positive bias, and so to an overestimation of the optimal error variance in the Kalman filter framework. So, also in the first analysis step, inaccuracies in the prior error covariances due to a small ensemble size lead to a bias in the posterior error covariances. This effect is due to the nonlinearity of the gain. HM98 do not seem to have realized this.

At a later analysis step the fifth term, proportional to $\epsilon_1\epsilon_2$, will become important due to inbreeding, as HM98 also mention. In the experiment by HM98 this is an important effect for small ensemble sizes: a bias develops and the linear term takes over.

4. Conclusions and discussion

The EnKF leads to systematically underestimated error variances for small ensemble sizes. This is not only due to the effect that the ensemble is updated with a gain calculated from that same ensemble, as HM98 claim, but also due to finite ensemble effects in the gain itself. This last effect comes about because the gain is nonlinear in the prior covariance. It tends to partly compensate for the first effect. The combined effect will decrease with increasing ensemble size.

The remedy proposed by HM98 in which a double ensemble is used (the DEnKF) has the same kind of problem, but to a lesser degree. In the first analysis step the inbreeding is absent, but the nonlinearity of the gain gives rise to an overestimation of the optimal error in the Kalman filter framework.

Clearly, the size of the problems will be dependent on the application, that is, the size and structure of the gain matrix, the random generator used, the size of the ensemble, etc.

The question then becomes which method to use. The above analysis favors the DEnKF for small ensemble sizes, when the errors in the estimations of the prior error covariances, ϵ, will be relatively large. However, for the same total number of ensemble members, the EnKF will give a better estimate of the mean. (See, for instance, Fig. 3 in HM98, the 32 vs 2 $\times$ 16 members case.) It must be said that a worse estimate of the mean with a proper error estimate is probably worth more than a better estimate of the mean with a poor error estimate.

The question then becomes, how relevant are small-ensemble-size calculations? Apart from the biases described above, a new problem arises: the forecast-observation covariances. HM98 do not mention this problem, but clearly they are nonnegligible for small ensemble sizes. It is unclear at what ensemble size the effect will come into play, that will be application dependent, but simple experiments with a random generator show that the correlation between two independently distributed Guassian random variables becomes less than 5% at a sample size of about 250.

Finally, an ensemble size that is too small will cause problems with statistics due to the fact that the sample is just too small. Using the Tchebycheff inequality the probability that the error in the estimated mean value $\overline{\boldsymbol{\psi}}$ is smaller than $\sigma/\sqrt{N(1-\gamma)}$ is given by (see, e.g., Papoulis 1991)

$$\Pr\left\{|\overline{\boldsymbol{\psi}} - \boldsymbol{\psi}^t| \le \frac{\sigma}{\sqrt{N(1-\gamma)}}\right\} > \gamma, \tag{21}$$

in which γ is the confidence coefficient, σ^2 is the variance, N is the size of the ensemble, and $\boldsymbol{\psi}^t$ is the true mean. So, if a confidence coefficient is chosen, and the variance is approximated by the estimated variance, a minimum value for the ensemble size will result. Suppose we want the mean with a confidence of 95% to lie in the interval $\boldsymbol{\psi}^t - 0.25\sigma < \overline{\boldsymbol{\psi}} < \boldsymbol{\psi}^t + 0.25\sigma$; then the ensemble size should be $\mathbf{O}(300)$! (Note that this argument is related to the above.)

To conclude, for small ensemble sizes ($N \le 100$) the DEnKF has smaller biases, but the smallness of the sample and the fact that forecast-observation covariances are neglected puts serious doubts on the use of ensemble methods of this size at all. Note that up to now we only discussed the variances. To estimate covariances correctly, many more ensemble members are needed. In view of this, the amount of skill shown in the experiments performed by HM98 is surprising. An ensemble size of 2 $\times$ 16 members gives a global error estimate comparable to the true error. A factor might be the perfect model assumption that they adopt, but that is difficult to study in a general setting;

it will be model dependent. Experiments with the EnKF by Evensen (1994) and Evensen and van Leeuwen (1996) in an oceanographic context show that ensemble sizes of at least 100 members are necessary to obtain reliable converged error variances (not global rms errors). The experiments by HM98 show that the difference between the EnKF and the DEnKF become negligible at those sizes. However, we might be comparing apples with pears; clearly, more research is needed on the size of ensembles.

Acknowledgments. This research was sponsored by the Dutch National Research Program on Global Change (NRP II) Grant 01301237.10.

REFERENCES

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

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

——, and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas Current using the ensemble Kalman filter with a quasigeostrophic model. *Mon. Wea. Rev.,* **124,** 85–96.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

Pappoulis, A., 1991: *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 666 pp.