Hybrid algorithm of ensemble transform and importance sampling for assimilation of non-Gaussian observations

By AUTHOR[*], *The Institute of Statistical Mathematics, Tachikawa, Tokyo, Japan*

(Manuscript received xx xxxx xx; in final form xx xxxx xx)

ABSTRACT

However, this problem could be resolved by monitoring the effective sample size and tuning the factor for covariance inflation. In this paper, the proposed hybrid algorithm is introduced, and its performance is evaluated through experiments with non-Gaussian observations.

Keywords: in this paper, in this paper, in this paper.

1. Introduction

The ensemble-based approach is now recognized as a valuable tool for data assimilation in nonlinear systems. In particular, the ensemble Kalman filter (EnKF) (Evensen, 1994; 2003) and the ensemble square root filters (Tippett et al., 2003; Livings et al., 2008) are widely used in various practical applications. However, since these algorithms basically assume a linear Gaussian observation model like the Kalman filter (KF), they could give biased or incorrect estimates when the observation is nonlinear or non-Gaussian.

1.1. Experiment with a linear Gaussian observation

The particle filter (PF) (Gordon et al., 1993; Kitagawa, 1996; van Leeuwen, 2009) is an ensemble-based algorithm that is applicable even in cases with nonlinear or non-Gaussian observations. However, the PF tends to be computationally expensive in comparison with other ensemble-based algorithms. One source of the high computational cost is the degeneracy of the ensemble. In the PF, ensemble members are weighted in the manner of the importance sampling (e.g., Liu, 2001; Candy, 2009), and are then resampled with probabilities equal to the weights. After resampling, many of the ensemble members are replaced by duplicates of some particular members with large weights. Consequently, the diversity of the ensemble is rapidly lost by repeating the resampling process. In order to achieve sufficient diversity, the PF usually requires a huge ensemble size, which results in high computational cost.

Many studies have investigated maintaining the ensemble diversity. One approach is based on kernel density estimation, in which a new ensemble is generated by resampling from a smoothed empirical density function (e.g., Musso et al., 2001). The Gaussian resampling approach (Kotecha and Djurić, 2003; Xiong et al., 2006) and the merging particle filter (Nakano et al., 2007) have also been devised to maintain the ensemble diversity, although these methods consider only the first and second moments rather than the shape of the PDF. Another method by which to maintain the ensemble diversity is to improve the distribution for sampling. If we can draw samples from a distribution similar to a posterior PDF, the weights would be well balanced among the samples which would enable us to improve the computational efficiency. Accordingly, several studies have attempted to improve the distribution for sampling. Pitt and Shephard (1999) proposed the auxiliary particle filter, in which the temporal evolution is calculated for each ensemble member after the ensemble is resampled according to the expected score of the prediction for each ensemble member. Recently, van Leeuwen (2010; 2011) proposed another algorithm that refers to the observed data in order to obtain the distribution for sampling. Chorin et al. (2010) and Morzfeld et al. (2012) also took a similar approach. Papadakis et al. (2010) proposed the weighted ensemble Kalman filter (WEnKF), in which the distribution for sampling is obtained by the EnKF, and the samples drawn from the distribution are weighted and resampled. Beyou et al. (2013) took the same approach but they used the the ensemble transform Kalman filter (ETKF) (Bishop et al., 2001; Wang et al., 2004), which is one of ensemble square root filters, instead of the EnKF.

However, even if the ensemble diversity is well maintained by improving the distribution for sampling, the ensemble size

[*] Corresponding author.
e-mail: xxx

© 0000 Tellus

that we can use is not necessarily sufficient to represent non-Gaussian features of the PDF. In practical applications, the ensemble size is usually limited by the available computational resources because a model run for each ensemble member is costly in the forecast step. Indeed, it is not unusual that the allowed ensemble size is much smaller than the system dimension. If the ensemble size N is smaller than the effective system dimension, the ensemble would form a simplex in $(N-1)$-dimensional space (Julier, 2003; Wang et al., 2004), which is obviously insufficient to represent the third or higher-order moments. In such a situation, the non-Gaussian features can not be represented even using the importance sampling. In addition, after weighting the ensemble members, some of the members no longer effectively contribute to the estimation. This means that the probability distribution estimation would be based on a substantially smaller sample size than the original sample size. Therefore, for the case in which the ensemble size is limited, the weighting of the ensemble would not necessarily provide a good approximation of the posterior PDF.

The approach proposed in the present paper considers such a situation in which the forecast PDF is represented by an ensemble of limited size less than the system dimension. Since non-Gaussian features are not represented by the limited-sized ensemble, the proposed approach assumes that the forecast PDF is Gaussian. On the other hand, in the analysis step, we use a sufficiently large number of samples to represent non-Gaussian features of the posterior PDF. These non-Gaussian features are represented by the importance sampling technique (e.g., Liu, 2001; Candy, 2009). An outline of the proposed approach is illustrated in Figure 1. Before the analysis step, the forecast PDF is represented by an ensemble of limited size that forms a simplex. From this forecast PDF, we obtain a Gaussian proposal PDF, which is similar to but not necessarily identical to the posterior PDF. This proposal PDF is represented by a small ensemble obtained using the ETKF. We then generate a large number of samples from the proposal PDF, and these samples are weighted so as to approximate the posterior PDF. If we use the proposal PDF obtained by the ETKF, we can efficiently generate a large number of samples, which allows us to represent non-Gaussian features of the posterior PDF using the importance sampling technique. For the next forecast step, the approximation of the posterior PDF with a large number of samples is converted into a new approximation with a small ensemble. This small ensemble is constructed under the assumption of a Gaussian distribution, but is obtained after considering the nonlinearity or non-Gaussianity of the observation. We can therefore reduce the effects of the biases due to nonlinear or non-Gaussian observation on the next forecast.

Various algorithms have been proposed that combine the PF algorithm with a Gaussian-based algorithm, such as the KF and EnKF. For example, several studies considered a Gaussian mixture model and used the KF or EnKF to update each Gaussian component of the Gaussian mixture model (e.g., Smith, 2007; Hoteit et al., 2012). However, a Gaussian mixture model requires a large number of parameters to represent the covariance matrices of each of the Gaussian components for high-dimensional systems and would tend to require too much memory and computational resources. Although Hoteit et al. (2008) proposed another approach that uses a mixture of Gaussian components with the same covariance matrix, their approach assumes a Gaussian observation model. Lei and Bickel (2011) considered another method by which to combine the PF algorithm and the EnKF algorithm, in which the ensemble is adjusted so as to represent the mean and covariance estimated by weighting the members of the forecast ensemble. This approach did not consider an asymmetric probability distribution.

Table 1. Results of experiments with the linear Gaussian observation model.

	RMSE ($\sigma = 0.5$)		RMSE ($\sigma = 1.0$)	
	Hybrid filter	ETKF	Hybrid filter	ETKF
N=16	0.84	0.62	4.31	3.44
N=18	0.18	0.19	0.71	0.50
N=20	0.18	0.19	0.39	0.41
N=24	0.19	0.20	0.39	0.41
N=28	0.19	0.20	0.40	0.42
N=32	0.19	0.20	0.40	0.43
N=36	0.19	0.21	0.40	0.43

The ETKF algorithms and the importance sampling technique, on which the proposed method is based, are explained in Sections **??** and **??**, respectively. Section **??** discusses how to use the ETKF output as a proposal PDF and how to represent the posterior PDF using samples drawn from the proposal PDF. Section **??** discusses how to approximate the posterior PDF with a small-sized ensemble, which allows us to achieve high computational efficiency in the forecast step. We experimentally evaluate the proposed algorithm in Section **??**, and provide a summary in Section **??**.

The mean vector $\overline{\boldsymbol{x}}_{k|k-1}$ is represented by the ensemble mean of all of the members:

$$\overline{\boldsymbol{x}}_{k|k-1} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}^{(i)}_{k|k-1}. \tag{1}$$

The ETKF then considers the deviation from the mean vector as

$$\Delta\boldsymbol{x}^{(i)}_{k|k-1} = \boldsymbol{x}^{(i)}_{k|k-1} - \overline{\boldsymbol{x}}_{k|k-1}, \tag{2}$$

$$\Delta\boldsymbol{y}^{(i)}_{k|k-1} = \boldsymbol{h}_k(\boldsymbol{x}^{(i)}_{k|k-1}) - \overline{\boldsymbol{h}_k(\boldsymbol{x}_{k|k-1})}, \tag{3}$$

where the function $\boldsymbol{h}_k$ is a nonlinear predictive observation given a state $\boldsymbol{x}_k$, and $\overline{\boldsymbol{h}_k(\boldsymbol{x}_{k|k-1})}$ denotes the ensemble mean of the predictive observation $\{\boldsymbol{h}_k(\boldsymbol{x}^{(i)}_{k|k-1})\}_{i=1}^{N}$ as

$$\overline{\boldsymbol{h}_k(\boldsymbol{x}_{k|k-1})} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{h}_k(\boldsymbol{x}^{(i)}_{k|k-1}). \tag{4}$$

Fig. 1. Outline of the hybrid approach proposed in the present paper.

2. Sequential data assimilation problem

We describe the state transition of a dynamical system by the following probability density function (PDF):

$$\boldsymbol{x}_k \sim p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \tag{5}$$

where the vector $\boldsymbol{x}_k$ denotes the state of the system at time t_k ($k = 1, 2, \ldots$). We then consider the following observation model to describe the relationship between the system state and the observation:

$$\boldsymbol{y}_k \sim p(\boldsymbol{y}_k|\boldsymbol{x}_k), \tag{6}$$

where $\boldsymbol{y}_k$ is the observed data at time t_k.

Sequential data assimilation is regarded as a problem that estimates the conditional PDF of the system state $\boldsymbol{x}_k$ from the sequence of observations $\boldsymbol{y}_{1:k} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\}$ according to the following recursive procedure. Suppose that the conditional PDF at the time step t_{k-1}, $p(\boldsymbol{x}_{k-1}|\boldsymbol{y}_{1:k-1})$, is given. Then, the forecast PDF $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})$ can be obtained by the following equation:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1}) = \int p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\, p(\boldsymbol{x}_{k-1}|\boldsymbol{y}_{1:k-1}) d\boldsymbol{x}_{k-1}. \tag{7}$$

The observation $\boldsymbol{y}_k$ is then assimilated using Bayes' theorem to obtain the filtered (analysis) PDF $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)\, p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})}. \tag{8}$$

Filtering algorithms for sequential data assimilation estimate the system states based on this filtered PDF. In the following, we discuss how to obtain a good approximation of the filtered PDF.

References

Ades, M. and van Leeuwen, P. J. 2012. An exploration of the equivalent weights particle filter. *Q. J. R. Meteorol. Soc.* **139**, 10.1002/qj.1995, 820–840.

Anderson, J. L. and Anderson, S. L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **127**, 2741.

Beyou, S., Cuzol, A., Gorthi, S. S. and Mémin, E. 2013. Weighted ensemble transform Kalman filter for image assimilation. *Tellus* **65A**, 1–17.

Bishop, C. H., Etherton, R. J. and Majumdar, S. J. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.* **129**, 420–436.

Bishop, C. M. 2006. Pattern recognition and machine learning Springer, New York.

Candy, J. V. 2009. Bayesian signal processing–Classical, modern, and particle filtering methods John Wiley & Sons, Inc.

Chorin, A. J., Morzfeld, M. and Tu, X. 2010. Implicit particle filters for data assimilation. *Commun. Appl. Math. Comput.* **5**, 221–240.

Doucet, A., Godsill, S. and Andrieu, C. 2000. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statist. Comput.* **10**, 197–208.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99(C5)**, 10143.

Evensen, G. 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics* **53**, 343.

Gordon, N. J., Salmond, D. J. and Smith, A. F. M. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F* **140**, 107.

Hoteit, I., Luo, X. and Pham, D.-T. 2012. Particle Kalman filtering: a nonlinear Bayesian framework for ensemble Kalman filters. *Mon. Wea. Rev.* **140**, 528–542.

Hoteit, I., Pham, D.-T., Triantafyllou, G. and Korres, G. 2008. A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography. *Mon. Wea. Rev.* **136**, 317–334.

Hunt, B. R., Kostelich, E. J. and Szunyogh, I. 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D* **230**, 112–126.

Julier, S. J.: 2003, The spherical simplex unscented transformation, *Proc. of the American Control Conference*, pp. 2430–2434.

Kitagawa, G. 1996. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Comp. Graph. Statist.* **5**, 1.

Kotecha, J. H. and Djurić, P. M. 2003. Gaussian particle filtering. *IEEE Trans. Signal Processing* **51**, 2592.

Lawson, W. G. and Hansen, J. A. 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.* **132**, 1966–1981.

Lei, J. and Bickel, P. 2011. A moment matching ensemble filter for nonlinear non-Gaussian data assimilation. *Mon. Wea. Rev.* **139**, 3964–3973.

Liu, J. S. 2001. Monte Carlo strategies in scientific computing Springer-Verlag, New York.

Liu, J. S. and Chen, R. 1995. Blind deconvolution via sequential imputations. *J. Amer. Statist. Assoc.* **90**, 567–576.

Livings, D. M., Dance, S. L. and Nichols, N. K. 2008. Unbiased ensemble square root filters. *Physica D* **237**, 1021–1028.

Lorenz, E. N. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations: Simulations with a small model. *J. Atmos. Sci.* **55**, 399.

Morzfeld, M., Tu, X., Atkins, E. and Chorin, A. J. 2012. A random map implementation of implicit filters. *J. Comput. Phys.* **231**, 2049–2066.

Musso, C., Oudjane, N. and Le Gland, F.: 2001, Improving regularized particle filters, *in* A. Doucet, N. de Freitas and N. Gordon (eds), *Sequential Monte Carlo methods in practice*, Springer-Verlag, New York, chapter 12, p. 247.

Nakano, S., Ueno, G. and Higuchi, T. 2007. Merging particle filter for sequential data assimilation. *Nonlin. Process. Geophys.* **14**, 395–408.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J. and Yorke, J. A. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **56A**, 415.

Papadakis, N., Mémin, E., Cuzol, A. and Gengembre, N. 2010. Data assimilation with the weighted ensemble Kalman filter. *Tellus* **62A**, 673–697.

Pham, D. T. 2001. Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.* **129**, 1194–1207.

Pitt, M. K. and Shephard, N. 1999. Filtering via simulation: Auxiliary particle filter. *Journal of the American Statistical Association* **94**, 590.

Rao, C. R.: 1973, Linear statistical inference and its applications, 2nd ed., John Wiley & Sons, Inc., chapter 8.

Roweis, S. and Ghahramani, Z. 1999. A unifying review of linear Gaussian models. *Neural Computation* **11**, 305–345.

Sakov, P. and Oke, P. R. 2008. Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.* **136**, 1042.

Smith, K. W. 2007. Cluster ensemble Kalman filter. *Tellus* **59A**, 749–757.

Snyder, C., Bengtsson, T., Bickel, P. and Anderson, J. 2008. Obstacles to high-dimensional particle filtering. *Mon. Wea. Rev.* **136**, 4629.

Song, H., Hoteit, I., Cornuelle, B. D. and Subramanian, A. C. 2010. An adaptive approach to mitigate background covariance limitations in the ensemble Kalman filter. *Mon. Wea. Rev.* **138**, 2825.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003. Ensemble square root filters. *Mon. Wea. Rev.* **131**, 1485–1490.

Tipping, M. E. and Bishop, C. M. 1999. Probabilistic principal component analysis. *J. Roy. Statist. Soc. B* **61**, 611–622.

van Leeuwen, P. J. 2009. Particle filtering in geophysical systems. *Mon. Wea. Rev.* **137**, 4089–4114.

van Leeuwen, P. J. 2010. Nonlinear data assimilation in geosciences: an extremely efficient particle filter. *Q. J. R. Meteorol. Soc.* **136**, 1991–1999.

van Leeuwen, P. J. 2011. Efficient nonlinear data-assimilation in geophysical fluid dynamics. *Computers Fluids* **46**, 52–58.

Wang, X., Bishop, C. H. and Julier, S. J. 2004. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble?. *Mon. Wea. Rev.* **132**, 1590–1605.

Xiong, X., Navon, I. M. and Uzunoglu, B. 2006. A note on particle filter with posterior Gaussian resampling. *Tellus* **58A**, 456.