Ensemble Data Assimilation with the Maximum Likelihood Ensemble Filter

Milija Zupanski and Kenneth Eis
DoD Center for Geosciences/Atmospheric Research, CIRA/Colorado State University, Fort Collins CO

Introduction

The Ensemble Data Assimilation (EnsDA) methods are new generation data assimilation methods which unify the data assimilation and ensemble forecasting. Based on the Kalman Filtering theory, as well as on the nonlinear dynamics (chaos) theory, they represent a methodology used to make an optimal estimate of the atmosphere/ocean state, and the associated uncertainty of the estimate.

Supported by the development of parallel (multi-processor) computing, for the first time the EnsDA methods can be applied in challenging, realistic data assimilation/forecasting environment., Only nonlinear models are used (e.g., ensembles), thus there is no need for an adjoint model. The analysis solution is searched in the subspace spanned by the ensemble forecasts.

Most of the EnsDA algorithms developed so far are designed to estimate the conditional mean of Probability Density Function (PDF). The Maximum Likelihood Ensemble Filter (MLEF), presented here, is designed to estimate the conditional mode of PDF.

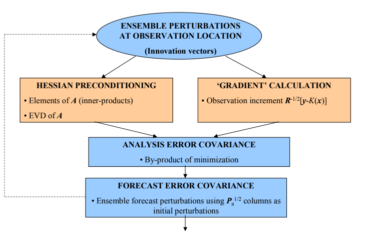
Ensemble Data Assimilation Approaches

Monte-Carlo Ensemble Kalman Filter
- Extended Kalman Filter solution form + ensembles
- Observation perturbation
- Analysis perturbation (with covariance inflation)
- Ensemble mean as first-guess

Reduced-rank Ensemble Kalman Filter
- Extended Kalman Filter solution form + ensembles
- Singular-value decomposition of forecast error covariance
- Analysis perturbation
- Short-range forecast as first-guess

Maximum Likelihood Ensemble Filter
- Nonlinear solution by cost-function minimization
- Analysis perturbation
- Short-range forecast as first-guess

MLEF Algorithmic Aspects



Conceptually and algorithmically simple, theoretically advanced!

Methodology

The MLEF algorithm has an embedded minimization algorithm (nonlinear conjugate-gradient, similar to variational data assimilation methods). The characteristics of the MLEF algorithm are:

(i) only nonlinear model is used (ensembles, no adjoint model),

(ii) superior Hessian preconditioning is automatically employed, and

(iii) uncertainty of the analysis is obtained as a by-product of minimization algorithm.

MLEF is equivalent to minimizing a cost function

- Maximize posterior probability to obtain the conditional mode (maximum likelihood estimate)
- General non-quadratic cost function

$$J=\frac{1}{2}(x-x_b)^T P_f^{-1}(x-x_b)+\frac{1}{2}[y-K(x)]^T R^{-1}[y-K(x)]$$

No assumption regarding the linearity of observation operator K !

Change of variable (Hessian preconditioning)

$$x - x_B = P_a^{1/2}\zeta = P_f^{1/2}(I+A)^{-T/2}\zeta$$

$$P_f^{1/2}=\begin{pmatrix} b_1 & b_2 & \cdot & b_{N\_ens}\end{pmatrix} \qquad b_i=\begin{pmatrix} b_1^i \\ b_2^i \\ \cdot \\ b_{N\_mod}^i\end{pmatrix}$$

$$A=\begin{pmatrix} z_1^T z_1 & z_1^T z_2 & \cdot & z_1^T z_{N\_ens} \\ z_2^T z_1 & z_2^T z_2 & \cdot & z_2^T z_{N\_ens} \\ \cdot & \cdot & \cdot & \cdot \\ z_{N\_ens}^T z_1 & z_{N\_ens}^T z_2 & \cdot & z_{N\_ens}^T z_{N\_ens}\end{pmatrix}$$

$$z_i=(R^{-1/2}KP_f^{1/2})_i \approx R^{-1/2}K(x+b_i)-R^{-1/2}K(x)$$

Forecast Error Covariance

$$P_f \approx MP_aM^T=[M(P_a)^{1/2}][M(P_a)^{1/2}]^T=P_f^{1/2}P_f^{T/2}$$

$$P_f^{1/2}=(MP_a^{1/2})_i \approx M(x+p_i)-M(x)$$

Analysis Error Covariance (In MLEF, a by-product of minimization)

$$P_a=P_a^{1/2}P_a^{T/2}=H^{-1} \qquad P_a^{1/2}=P_f^{1/2}(I+A)^{-T/2}$$

$$P_a^{1/2}=\begin{pmatrix} p_1 & p_2 & \cdot & p_{N\_ens}\end{pmatrix} \qquad p_i=\begin{pmatrix} p_1^i \\ p_2^i \\ \cdot \\ p_{N\_mod}^i\end{pmatrix}$$

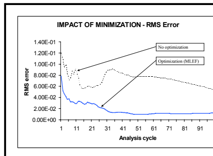
Results

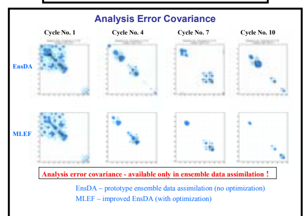
The results of MLEF are presented in applications to the one-dimensional Korteweg-de Vries-Burgers (KdVB) equation. The KdVB equation includes nonlinear advection, dispersion, and diffusion.

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+6\frac{\partial^3 u}{\partial x^3}=\nu\frac{\partial^2 u}{\partial x^2}$$

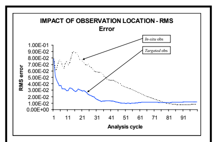
Experimental set-up:
- 101 degress of freedom
- 10 ensemble members
- 3 minimization iterations in each analysis cycle
- 10 observations (model forecast + perturbation)
- quadratic observation operator: $K(x)=x^2$

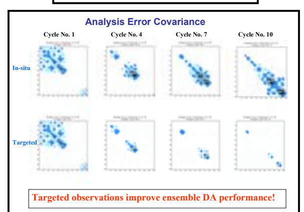
Impact of cost function minimization





Impact of observation location





Relevance to DoD

- Army-scale meteorology:
  - microscale data assimilation
  - model error detection and correction
  - uncertainty transfer between mesoscale and microscale models
  - high-performance computing (HPC)
- Probabilistic methodology – probability density function
  - optimal link with Tactical Decision Aids (TDAs)
- Optimal observation network design
  - self-automated optimized system
- Homeland security:
  - probabilistic sensitivity to detect the source of contamination (chem-bio)
  - estimated likelihood of future contamination, evacuation measures

Publications

Zupanski, M., 2004: Ensemble Data Assimilation with Hessian Preconditioning. Submitted to *J. Num. Lin. Alg. with Appl.*

Zupanski, M., 2004: Maximum Likelihood Ensemble Filter. Part I: Theoretical Aspects. Submitted to *Mon. Wea. Rev.* [Available at ftp://ftp.cira.colostate.edu/milija/papers/MLEF_mwr.pdf]

Zupanski, D., and M. Zupanski, 2004: Maximum Likelihood Ensemble Filter. Part II: Model Error Estimation. Submitted to *Mon. Wea. Rev.* [Available at ftp://ftp.cira.colostate.edu/milija/papers/MLEF_model_err.pdf]