



An ensemble-based explicit four-dimensional variational assimilation method

Xiangjun Tian,¹ Zhenghui Xie,¹ and Aiguo Dai²

Received 3 May 2008; revised 23 August 2008; accepted 17 September 2008; published 15 November 2008.

[1] The adjoint and tangent linear models in the traditional four-dimensional variational data assimilation (4DVAR) are difficult to obtain if the forecast model is highly nonlinear or the model physics contains parameterized discontinuities. A new method (referred to as POD-E4DVAR) is proposed in this paper by merging the Monte Carlo method and the proper orthogonal decomposition (POD) technique into 4DVAR to transform an implicit optimization problem into an explicit one. The POD method is used to efficiently approximate a forecast ensemble produced by the Monte Carlo method in a 4-dimensional (4-D) space using a set of base vectors that span the ensemble and capture its spatial structure and temporal evolution. After the analysis variables are represented by a truncated expansion of the base vectors in the 4-D space, the control (state) variables in the cost function appear explicit so that the adjoint model, which is used to derive the gradient of the cost function with respect to the control variables in the traditional 4DVAR, is no longer needed. The application of this new technique significantly simplifies the data assimilation process and retains the two main advantages of the traditional 4DVAR method. Assimilation experiments show that this ensemble-based explicit 4DVAR method performs much better than the traditional 4DVAR and ensemble Kalman filter (EnKF) method. It is also superior to another explicit 4DVAR method, especially when the forecast model is imperfect and the forecast error comes from both the noise of the initial field and the uncertainty in the forecast model. Computational costs for the new POD-E4DVAR are about twice as the traditional 4DVAR method but 5% less than the other explicit 4DVAR and much lower than the EnKF method. Another assimilation experiment conducted within the Lorenz model indicates potential wider applications of this new POD-E4DVAR method.

Citation: Tian, X., Z. Xie, and A. Dai (2008), An ensemble-based explicit four-dimensional variational assimilation method, *J. Geophys. Res.*, 113, D21124, doi:10.1029/2008JD010358.

1. Introduction

[2] The four-dimensional variational data assimilation (4DVAR) method [Johnson *et al.*, 2006; Kalnay *et al.*, 2007; Tsuyuki and Miyoshi, 2007] has been a very successful technique used in operational numerical weather prediction (NWP) at many weather forecast centers [Bormann and Thépaut, 2004; Park and Zou, 2004; Caya *et al.*, 2005; Bauer *et al.*, 2006; Rosmond and Xu, 2006; Gauthier *et al.*, 2007]. The 4DVAR technique has two attractive features: (1) the physical (forecast) model provides a strong dynamical constraint, and (2) it has the ability to assimilate the observational data at multiple times. However, 4DVAR still faces numerous challenges in coding, maintaining and updating the adjoint model of the forecast model and it requires the linearization of the forecast model. Usually, the

control variables (or initial states) are expressed implicitly in the cost function. To compute the gradient of the cost function with respect to the control variables, one has to integrate the adjoint model, whose development and maintenance require significant resources, especially when the forecast model is highly nonlinear and the model physics contains parameterized discontinuities [Xu, 1996; Mu and Wang, 2003]. Many efforts have been devoted to avoid integrating the adjoint model or reduce the expensive computation costs [Courtier *et al.*, 1994; Kalnay *et al.*, 2000; Wang and Zhao, 2005]. Nevertheless, the tangent linear model of the forecast model is still required in all these methods. On the other hand, the usual ensemble Kalman Filter (EnKF) [e.g., Evensen, 1994, 2003; Kalnay *et al.*, 2007; Beezley and Mandel, 2008; also see Appendix A] has become an increasingly popular method because of its simple conceptual formulation and relative ease of implementation. For example, it requires no derivation of a tangent linear operator or adjoint equations, and no integrations backward in time. Furthermore, the computational costs are affordable and comparable with other popular and sophisticated assimilation methods such as the 4DVAR

¹ICCES, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China.

²CGD, National Center for Atmospheric Research, Boulder, Colorado, USA.

method. By forecasting the statistical characteristics, EnKF can provide flow-dependent error estimates of the background errors using the Monte Carlo method, but it lacks the dynamic constraint as in 4DVAR. *Heemink et al.* [2001] developed a variance reduced EnKF method by using a reduced-rank approximation technique to reduce the huge amount of computer costs. *Farrell and Ioannou* [2001] also proposed a reduced-order Kalman filter by the balanced truncation model-reduction technique. *Uzunoglu et al.* [2007] modified a maximum likelihood ensemble filter method [*Zupanski*, 2005] through an adaptive methodology. Generally, these three methods mentioned above belong to the Kalman filters. *Vermeulen and Heemink* [2006] have attempted to combine 4DVAR with EnKF; however, the tangent linear model is still needed in their method. How to retain the two primary advantages of the traditional 4DVAR while avoiding the need of an adjoint or tangent linear model of the forecast model has become a roadblock in advancing data assimilation. Recently, *Qiu et al.* [*Qiu and Chou*, 2006; *Qiu et al.*, 2007a, 2007b] proposed a new method for 4DVAR (more details below) using the singular value decomposition (SVD) technique based on the theory of the atmospheric attractors. *Cao et al.* [2007] have applied the proper orthogonal decomposition (POD) technique [*Ly and Tran*, 2001, 2002; *Volkwein*, 2008; also see Appendix C] to 4DVAR to reduce the forecast model orders while reducing the computational costs, but the adjoint integration is still necessary in their method. *Luo et al.* [2007] also applied the POD technique to the tropical ocean reduced gravity model.

[3] Here we resort to the idea of the Monte Carlo method and the POD technique. The basic idea of the POD technique is to start with an ensemble of data, called *snapshots*, collected from an experiment or a numerical procedure of a physical system. The POD technique is then used to produce a set of base vectors which span the snapshot collection. The goal is to represent the ensemble of the data in terms of an *optimal* coordinate system. That is, the snapshots can be generated by a smallest possible set of base vectors. On the basis of this approach, an explicit new 4DVAR method is proposed in this paper: it begins with a 4-D ensemble obtained from the forecast ensembles at all times in an assimilation time window produced using the Monte Carlo method. We then apply the POD technique to the 4-D forecast ensemble, so that the orthogonal base vectors cannot only capture the spatial structure of the state but also reflect its temporal evolution. After the model status is expressed by a truncated expansion of the base vectors obtained using the POD technique, the control variables in the cost function appear explicit, so that the adjoint or tangent linear model is no longer needed.

[4] Our new method was motivated by the need to merge the Monte Carlo method into the traditional 4DVAR to transform an implicit optimization problem into an explicit one. Our method not only simplifies the data assimilation procedure but also maintains the two main advantages of the traditional 4DVAR. This method is somewhat similar to *Qiu et al.*'s SVD-based method (referred to as SVD-E4DVAR hereafter, see Appendix B for details) because they both begin with a 4-D ensemble obtained from the forecast ensembles. However, they differ significantly in several aspects as discussed in section 2.2. *Hunt et al.* [2004], *John*

and Hunt [2007] and *Szunyogh et al.* [2008] also developed a 4-D ensemble Kalman filter that infers the tangent linear model dynamics from the ensemble instead of the tangent-linear map as done in the traditional 4DVAR, in which the model states are expressed by the linear combinations of the ensemble samples directly rather than some orthogonal base vectors of the ensemble space. This method is also largely Kalman filtering, with the generation of its ensemble space being different from our method.

[5] We conducted several numerical experiments using a one-dimensional (1-D) soil water equation and synthetic observations to evaluate our new method in land data assimilation. Comparisons were also made between our method, SVD-E4DVAR [*Qiu and Chou*, 2006; *Qiu et al.*, 2007a, 2007b], traditional 4DVAR, and EnKF. We found that our new ensemble-based explicit 4DVAR (referred to as POD-E4DVAR) performs much better than the usual EnKF method in terms of both increasing the assimilation precision and reducing the computational costs. It is also better than the traditional 4DVAR and SVD-E4DVAR, especially when the forecast model is not perfect and the forecast error comes from both the noise of the initial field and the uncertainty in the forecast model. We also evaluate this approach using the Lorenz model. The corresponding assimilation experiments show that POD-E4DVAR can adjust the forecast state to approach the true Lorenz curve rapidly only by assimilating the observations twice in an assimilation cycle, which indicates its potential applications in other fields.

2. Methodology

2.1. POD-E4DVAR Method

[6] In principle, the traditional, implicit 4DVAR (referred to as I4DVAR) analysis of \vec{x}_k is obtained through the minimization of a cost function J that measures the misfit between the model trajectory $H_k(\vec{x}_k)$ and the observation \vec{y}_k at a series of times t_k , $t = 1, 2, \dots, m$:

$$J(\vec{x}_0) = (\vec{x}_0 - \vec{x}_b)^T B^{-1} (\vec{x}_0 - \vec{x}_b) + \sum_{i=0}^m [\vec{y}_i - H_i(\vec{x}_i)]^T R_i^{-1} [\vec{y}_i - H_i(\vec{x}_i)], \quad (1)$$

with the forecast model $M_{0 \rightarrow k}$ imposed as strong constraints, defined by

$$\vec{x}_k = M_{0 \rightarrow k}(\vec{x}_0), \quad (2)$$

where the superscript T stands for a transpose, b is a background value, index k denotes the observational time, H_k is the observational operator, and matrices B and R are the background and observational error covariances, respectively. The control variables are the initial conditions \vec{x}_0 (at the start of the assimilation time window) of the model. In the cost function (1) the control variable \vec{x}_0 is connected with \vec{x}_k through forward integration of (2) and expressed implicitly, which makes it difficult to compute the gradient of the cost function with respect to \vec{x}_0 . Assuming there are S time steps within the assimilation time window $(0, T)$, generate N random perturbation fields using the Monte-Carlo method and add each perturbation field to the initial background field at $t = t_0$ to produce N initial fields

$\vec{x}_n(t_0)$, $n = 1, 2, \dots, N$. Integrate the forecast model $\vec{x}_n(t_i) = M_i(\vec{x}_n(t_{i-1}))$ with the initial fields $\vec{x}_n(t_0)$ ($n = 1, 2, \dots, N$) throughout the assimilation time window to obtain the state series $\vec{x}_n(t_i)$ ($i = 0, 1, \dots, S - 1$) and then construct the perturbed 4-D fields (snapshots) \vec{X}_n ($n = 1, 2, \dots, N$) over the assimilation time window:

$$\vec{X}_n = (\vec{x}_n(t_0), \vec{x}_n(t_1), \dots, \vec{x}_n(t_{S-1})), n = 1, 2, \dots, N. \quad (3)$$

It is obvious that such vectors can capture the spatial structure of the model state and its temporal evolution. All the perturbed 4-D fields \vec{X}_n ($n = 1, 2, \dots, N$) can expand a

finite dimensional space $\Omega(\vec{X}_1 \vec{X}_2 \dots \vec{X}_N)$. Similarly, the analysis field $\vec{x}_a(t_i)$ ($i = 0, 1, 2, \dots, S - 1$) over the same assimilation time window can also be stored into the following vector:

$$\vec{X}_a = (\vec{x}_a(t_0), \vec{x}_a(t_1), \dots, \vec{x}_a(t_{S-1})). \quad (4)$$

When the ensemble size N is increased by adding random samples, the ensemble space could cover the analysis vector \vec{X}_a , i.e., \vec{X}_a is approximately assumed to be embedded in the linear space $\Omega(\vec{X}_1 \vec{X}_2 \dots \vec{X}_N)$. Let \vec{X}_{bn} ($n = 1, 2, \dots, K$, $K \leq N$) be the base vectors of this linear space $\Omega(\vec{X}_1 \vec{X}_2 \dots \vec{X}_N)$, the analysis vector \vec{X}_a can be expressed by the linear combinations of this set of base vectors since it is in this space, i.e.

$$\vec{X}_a = \sum_{n=1}^K \beta_n \vec{X}_{bn}, \quad (5)$$

Substituting (4) and (5) into (1), the control variable becomes $\beta = (\beta_1 \dots \beta_K)^T$ instead of $\vec{x}(t_0)$, so the control variable is expressed explicitly in the cost function and the computation of the gradient is simplified greatly. The tangent linear model or adjoint model is no longer required. To minimize the cost function, equation (1) is transformed into an explicit optimization problem with the variable vector $\beta = (\beta_1 \dots \beta_K)^T$, which can be solved by the usual optimization algorithms, such as the quasi-Newton method. It is noted that, unlike EnKF, only one analyzed field is obtain in each analysis procedure in POD-E4DVAR and the initial condition should be perturbed at the start time of the assimilation in each cycle.

[7] How to obtain the appropriate base vectors remains the only task left. We found that the POD technique is a good choice for doing this. It can produce a set of base vectors spanning the ensemble of data in certain least squares optimal sense [Ly and Tran, 2001, 2002].

[8] The average of the ensemble of snapshots is given by

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N \vec{X}_n. \quad (6)$$

We form a new ensemble by focusing on deviations from the mean as follows

$$\delta X_n = \vec{X}_n - \bar{X}, n = 1, \dots, N, \quad (7)$$

which form the matrix $A(M \times N)$, where $M = M_g \times M_v \times S$, and M_g, M_v are the number of the model spatial grid points and the number of the model variables respectively. To compute the POD modes, one must solve an $M \times M$ eigenvalue problem:

$$(AA^T)_{M \times M} V = \lambda V$$

In practice, the direct solution of this eigenvalue problem is often not feasible if $M \gg N$, which occurs often in numerical models. We can transform it into an $N \times N$ eigenvalue problem through the following transformations:

$$A^T A V = \lambda V,$$

$$AA^T A V = A \lambda V,$$

$$AA^T A V = A \lambda V,$$

$$AA^T (A V) = \lambda (A V).$$

In the method of snapshots, one then solves the $N \times N$ eigenvalue problem

$$T V_k = \lambda_k V_k, k = 1, \dots, N, \quad (8)$$

where $T = (A^T A)_{N \times N}$, V_k is the k th column vector of V and is the k th row vector of λ . The nonzero eigenvectors λ_k ($1 \leq k \leq N$) may be chosen to be orthonormal, and the POD modes are given by $\phi_k = A V_k / \sqrt{\lambda_k}$, ($1 \leq k \leq N$).

[9] The truncated reconstruction of analysis variable in the four-dimensional space \vec{X}_a is given by

$$\vec{X}_a = \bar{X} + \sum_{j=1}^P \alpha_j \phi_j, \quad (9)$$

where P (the number of the POD modes) is defined as follows

$$P = \min \left\{ P, I(P) = \frac{\sum_{i=1}^P \lambda_i}{\sum_{i=1}^N \lambda_i} : I(P) \geq \gamma \right\}, 0 < \gamma \leq 1. \quad (10)$$

It is well known [Ly and Tran, 2001, 2002] that the expansion (9) is optimal. In particular, among all linear combinations, the POD is the most efficient, in the sense that, for a given number of modes P , the POD decomposition captures the most possible kinetic energy. The solution for the analysis problem is approximately expressed by a truncated expansion of the POD base vectors in the 4-D space. Substituting (9) and (4) into (1), the control variable becomes $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_P)^T$ instead of \vec{x}_0 , so the control variable is expressed explicitly in the cost function and the tangent linear model or adjoint model is not needed any more.

[10] In the above formulations, the usual optimization algorithms to find the solution of $\alpha = (\alpha_1, \dots, \alpha_P)^T$ still need an iterative procedure and likely result in higher computational costs. This issue is addressed as follows:

[11] Form the POD mode matrix

$$\Phi = (\phi_1, \phi_2, \dots, \phi_P), \quad (11)$$

where $\phi_j = (\phi_j(t_0), \phi_j(t_1), \dots, \phi_j(t_{S-1}))^T$, $j = 1, 2, \dots, P$. Transform (11) into the following format

$$\Phi = (\Phi_0, \Phi_1, \dots, \Phi_{S-1})^T, \quad (12)$$

where $\Phi_k = (\phi_1(t_k), \phi_2(t_k), \dots, \phi_P(t_k))$. Equation (9) is rewritten as follows:

$$\bar{X}_a = \bar{X} + \Phi\alpha, \quad (13)$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_P)^T$. The cost function (13) can be transformed into the following

$$\begin{aligned} J(\alpha) = & (\bar{x}(t_0) + \Phi_0\alpha - \bar{x}_b)^T B^{-1} (\bar{x}(t_0) + \Phi_0\alpha - \bar{x}_b) \\ & + \sum_{j=0}^m [\bar{y}_j - H_j\bar{x}(t_j) - H_j\Phi_j\alpha]^T \\ & \cdot R_j^{-1} [\bar{y}_j - H_j\bar{x}(t_j) - H_j\Phi_j\alpha], \end{aligned} \quad (14)$$

where H_j is the tangent linear observation operator.

[12] Given the vector of measurements $\bar{y}_j = (y_{j1}, \dots, y_{jm_j})$, where m_j is the size of \bar{y}_j , we can define the N vectors with perturbed observations as

$$Y_{ij} = \bar{y}_j + \varepsilon_i, i = 1, \dots, N,$$

where $\varepsilon_i = (\varepsilon_{i,1}, \varepsilon_{i,2}, \dots, \varepsilon_{i,m_j})^T$. The ensemble of perturbations, with ensemble mean equal to zero, can be stored in the matrix $E_j = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)$. The measurement error covariance matrix can be estimated by

$$R_j = \frac{E_j E_j^T}{N-1}, j = 0, \dots, m. \quad (15)$$

Because R_j^{-1} is symmetrical, the gradient of the cost function is obtained through simple calculations

$$\begin{aligned} \nabla J(\alpha) = & (\Phi_0)^T B^{-1} (\bar{x}(t_0) - \bar{x}_b + \Phi_0\alpha) \\ & + \sum_{j=0}^m -[H_j\Phi_j]^T R_j^{-1} [\bar{y}_j - H_j\bar{x}(t_j) - H_j\Phi_j\alpha], \end{aligned} \quad (16)$$

One can solve the optimization problem as follows

$$\nabla J(\alpha) = 0, \quad (17)$$

and

$$\begin{aligned} & \left((\Phi_0)^T B^{-1} \Phi_0 + \sum_{j=0}^m [H_j\Phi_j]^T R_j^{-1} [H_j\Phi_j] \right) \alpha \\ & = \sum_{j=0}^m [H_j\Phi_j]^T R_j^{-1} [\bar{y}_j - H_j\bar{x}(t_j)] - (\Phi_0)^T B^{-1} (\bar{x}(t_0) - \bar{x}_b). \end{aligned} \quad (18)$$

Equation (18) can be solved directly without an iterative procedure (see Figure 1 for a flowchart of the outlined method).

2.2. Difference Between the POD- and SVD-Based Methods

[13] The two explicit (i.e., the POD- and SVD-based) methods are essential for constructing a forecast ensemble $\Omega(\bar{X}_1 \bar{X}_2 \dots \bar{X}_N)$ in a 4-D space with a view to cover the analysis state over each assimilation time window. Since the 4-D analysis vector is assumed to be in the linear space, it can be expressed by a set of base vectors of this space. How to extract the base vectors becomes the key to the two methods: neither of the two methods decomposes the ensemble $A_{M \times N} = (X_1, X_2, \dots, X_N)$ directly (i.e., $\delta X_n = \bar{X}_n - 0$, $n = 1, \dots, N$, referred to as Full-E4DVAR hereafter) instead of forming another ensemble by focusing on deviations from the mean or the background vector separately. The two different vector transformations ($\delta X_n = \bar{X}_n - \bar{X}(\bar{X}_b)$) differentiate them clearly and affect their assimilation performance considerably even though POD and SVD techniques are closely related (see <http://www.uni-graz.at/imawww/volkwein/svd.ps>):

[14] Given a forecast ensemble matrix $A_{M \times N} = (X_1, X_2, \dots, X_N)$, $X_n \in R^M$, one can define the following L^2 norm $\|\cdot\|$ subjected to

$$\|X_n\|^2 = \sum_{i=1}^M X_{ni}^2. \quad (19)$$

As in (X. Tian and Z. Xie, Effects of sample density on the performance of an explicit four-dimensional variational data assimilation method, submitted to *Science in China (D)*, 2008), we define the mean norm \bar{R} of the ensemble A

$$\bar{R} = \sqrt{\frac{1}{N} \sum_{n=1}^N \|X_n\|^2}, \quad (20)$$

For a given sample number N , we propose a concept of sample density as follows:

$$\rho = \frac{N}{\frac{\pi^{\frac{M}{2}}}{\Gamma(\frac{M}{2} + 1)} \bar{R}^M}, \quad (21)$$

where Γ is the Gamma function. We want to look for a vector $X^* \in R^M$ to minimize

$$\bar{R}(X) = \sqrt{\frac{1}{N} \sum_{n=1}^N \|X_n - X\|^2}, \quad (22)$$

which leads to the maximum sample density

$\rho_{\max} = \frac{N}{\frac{\pi^{\frac{M}{2}}}{\Gamma(\frac{M}{2} + 1)} (\bar{R}(X^*))^M}$. The assimilation performance would be most efficient if $\rho = \frac{N}{\frac{\pi^{\frac{M}{2}}}{\Gamma(\frac{M}{2} + 1)} \bar{R}^M}$ reaches its

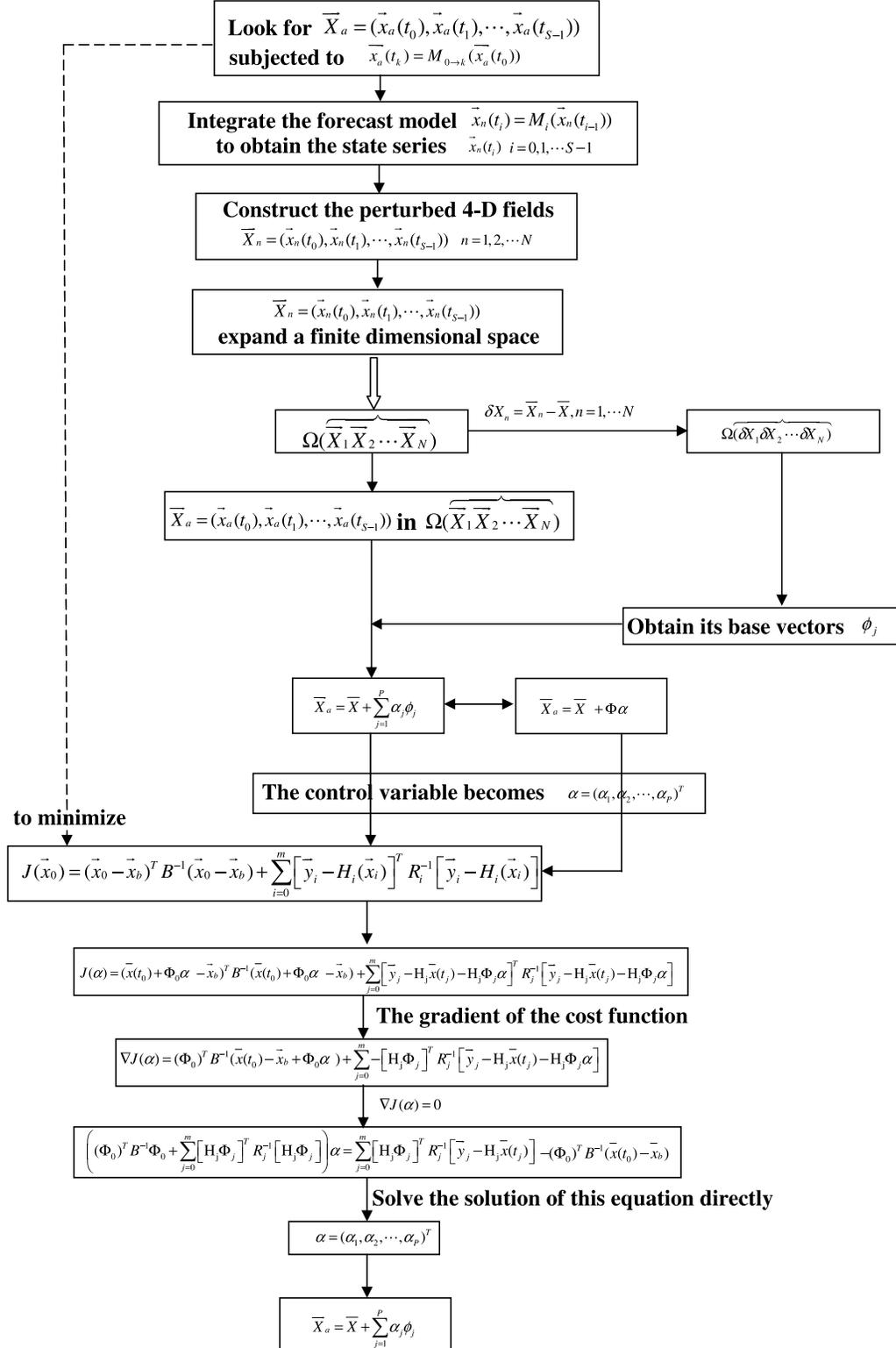


Figure 1. The algorithmic flowchart of the POD-E4DVAR method.

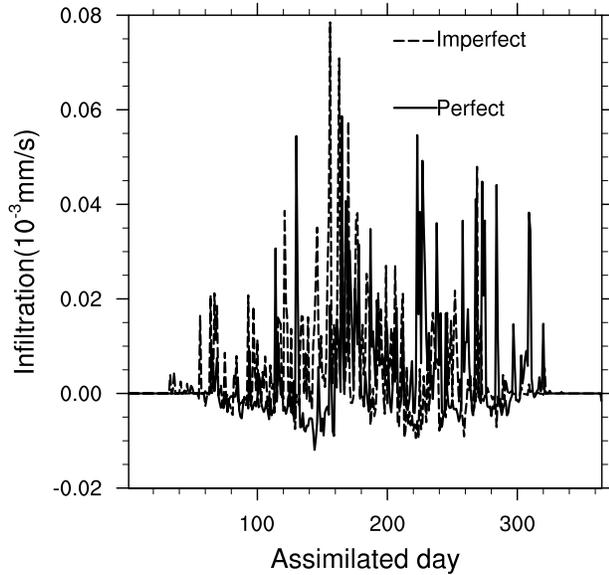


Figure 2. The “perfect” (solid line) and “imperfect” (dashed line) infiltration time series used in the assimilation experiments.

maximum value ρ_{\max} for the given sample ensemble, which means one single sample vector can represent the least “true space” and then gain the best assimilation effect. Straightforward calculations show that the function (22) is minimized when $X^* = \bar{X} = \frac{1}{N} \sum_{n=1}^N X_n$. It should be noted that even minor deviations, which always exist, of the background vector \bar{X}_b from the mean vector \bar{X} would result in the sample density difference being magnified greatly by M orders ($\sim 10^6$) in the commonly used climate model:

$$\frac{\rho_{\bar{X}}}{\rho_{X_b}} = \frac{\frac{N}{\pi^2 \bar{R}_{\bar{X}}^M / \Gamma(\frac{M}{2} + 1)}}{\frac{M}{\pi^2 \bar{R}_{X_b}^M / \Gamma(\frac{M}{2} + 1)}} = \left(\frac{\bar{R}_{X_b}}{\bar{R}_{\bar{X}}} \right)^M. \quad (23)$$

The above POD-E4DVAR method also differs from SVD-E4DVAR significantly in two technical aspects:

[15] 1. the 4-D sample in the SVD-E4DVAR method is only composed of the state vectors at the observational times over the assimilation time window, while it is composed of the state vectors at all the time steps over the assimilation time window in the POD-EDVAR method. The latter contains the most possible forecast information in the assimilation time window.

[16] 2. The application of the matrix transformation technique in the POD-E4DVAR greatly lowers the computational costs by reducing the decomposition into an $N \times N$ eigenvalue problem ($N \ll M$).

3. Evaluations in a 1-D Soil Water Model

[17] In this section, the applicability of this new method is evaluated through several assimilation experiments with a simple 1-D soil water equation model used in the NCAR

Community Land Model (CLM) [Oleson *et al.*, 2004]. In addition, we also compare assimilation results using the Full-E4DVAR, SVD-E4DVAR, I4DVAR, and EnKF methods.

3.1. Setup of Experiments

[18] The volumetric soil moisture (θ) for 1-D vertical water flow in a soil column in the CLM is expressed as

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - E - R_{fm}, \quad (24)$$

where q is the vertical soil water flux, E is the evapotranspiration rate, and R_{fm} is the melting (negative) or freezing (positive) rate, (for simplicity, E, R_{fm} are taken as zero in the experiments), and z is the depth from the soil surface. Both q and z are positive downward.

[19] The soil water flux q is described by Darcy’s law [Darcy, 1856]:

$$q = -k \frac{\partial(\varphi + z)}{\partial z}, \quad (25a)$$

where $k = k_s \left(\frac{\theta}{\theta_s} \right)^{2b+3}$ is the hydraulic conductivity, and $\varphi = \varphi_s \left(\frac{\theta}{\theta_s} \right)^{-b}$ is the soil matrix potential, k_s, φ_s, θ_s and b are constants. The CLM computes soil water content in the 10 soil layers through ((24)–(25)) (see Oleson *et al.* [2004] for details). The upper boundary condition is

$$q_0(t) = -k \frac{\partial(\varphi + z)}{\partial z} \Big|_{z=0}, \quad (25b)$$

where $q_0(t)$ is the water flux at the land surface (referred to as infiltration), and the lower boundary condition is $q_l = 0$. The time step Δt is 1800 s (0.5 hour).

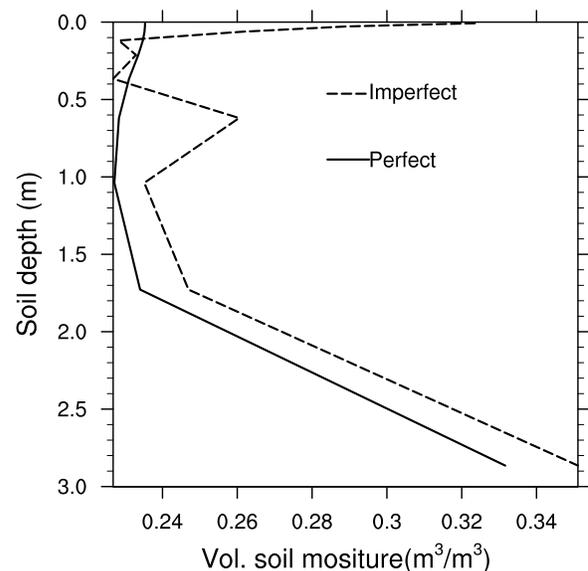


Figure 3. The “perfect” (solid line) and “imperfect” (dashed line) initial soil moisture profiles used in the assimilation experiments.

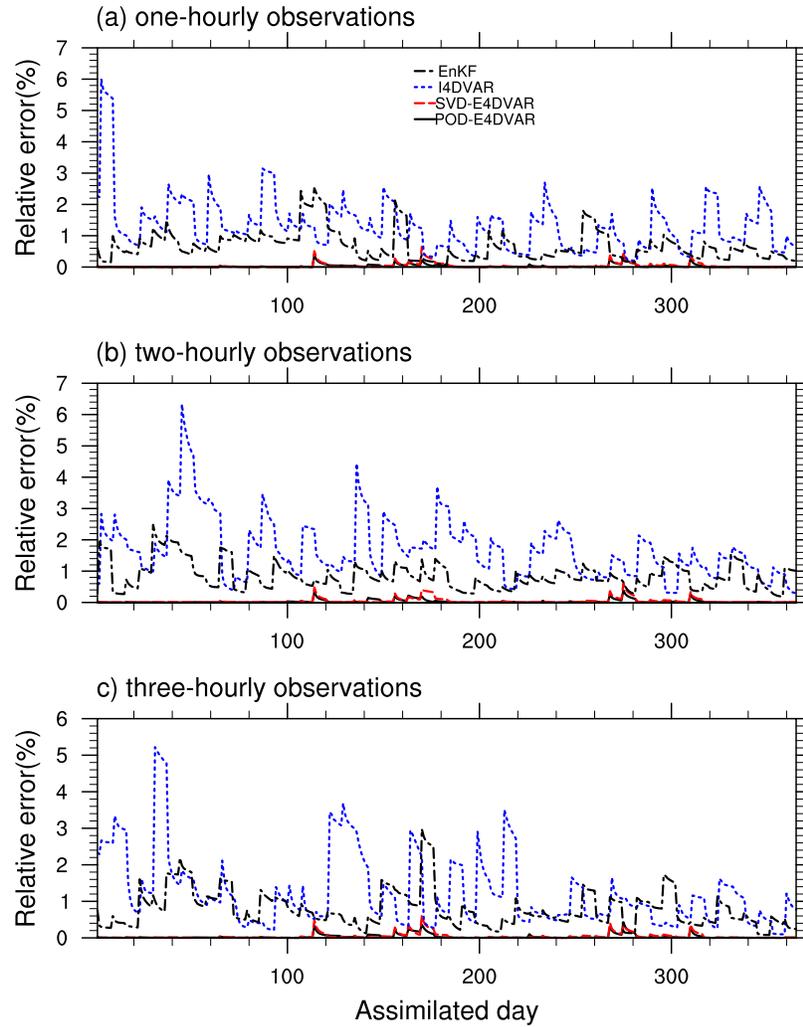


Figure 4. Relative error (E_n) for analyzed soil moisture in the assimilation experiments by the perfect model with the “imperfect” initial field.

[20] We took a site at (47.43°N, 126.97°E) as the experimental site. The soil parameters k_s , ϕ_s , θ_s and b at this site were calculated by the CLM using the high-resolution soil texture data released with the CLM by NCAR: $\theta_s = 0.46\text{m}^3/\text{m}^3$, $k_s = 2.07263\text{E-}6$ m/s, $b = 8.634$, $\phi_s = -3.6779$ m. We then ran the model at the site forced with observation-based 3-hourly forcing data [Qian *et al.*, 2006; Tian *et al.*, 2007] from 1 January 1992 to 31 December 1993 after ten-year spinning-up to obtain a two-year time series of simulated infiltration (i.e., the water flux q at the surface, c.f., equation (25b)) for driving the soil water hydrodynamic equation (24). We used the first year (1 January 1992 to 31 December 1992) data of CLM-simulated infiltration as the “perfect” infiltration series, and took the second year data as the “imperfect” infiltration series (Figure 2). In our experiments, we integrated the soil water hydrodynamic equation (24) forced by the two infiltration time series for 365 days separately: equation (24) forced by the “perfect” infiltration series represents the perfect *forecast model*, whose

forecast error comes only from the noise in the initial (soil moisture) field; on the contrary, equation (24) forced by the “imperfect” infiltration series acts as the “imperfect” forecast model, whose forecast error comes from not only the noise of the initial field but also the uncertainty in the forecast model itself.

[21] Figure 3 shows the “imperfect” and the “perfect” initial soil moisture profiles, which were obtained by randomly taking two arbitrary CLM-simulated soil moisture profiles in the process of the infiltration series producing. These profiles represent the initial fields with and without noise. The “perfect” (or “true”) state was produced by integrating the “perfect” model with the “perfect” initial soil moisture profile for 365 days. The “observations” were generated by adding 3% random error perturbations to the time series of the “perfect” state (i.e., “observation” = $(1 + \varepsilon) \times$ “perfect”, where ε is a real random number varying from -3% to 3%), and these “observations” were assimilated using the various methods in the assimilation experiments (but not in the forecast experiments). In addition, two separate forecast

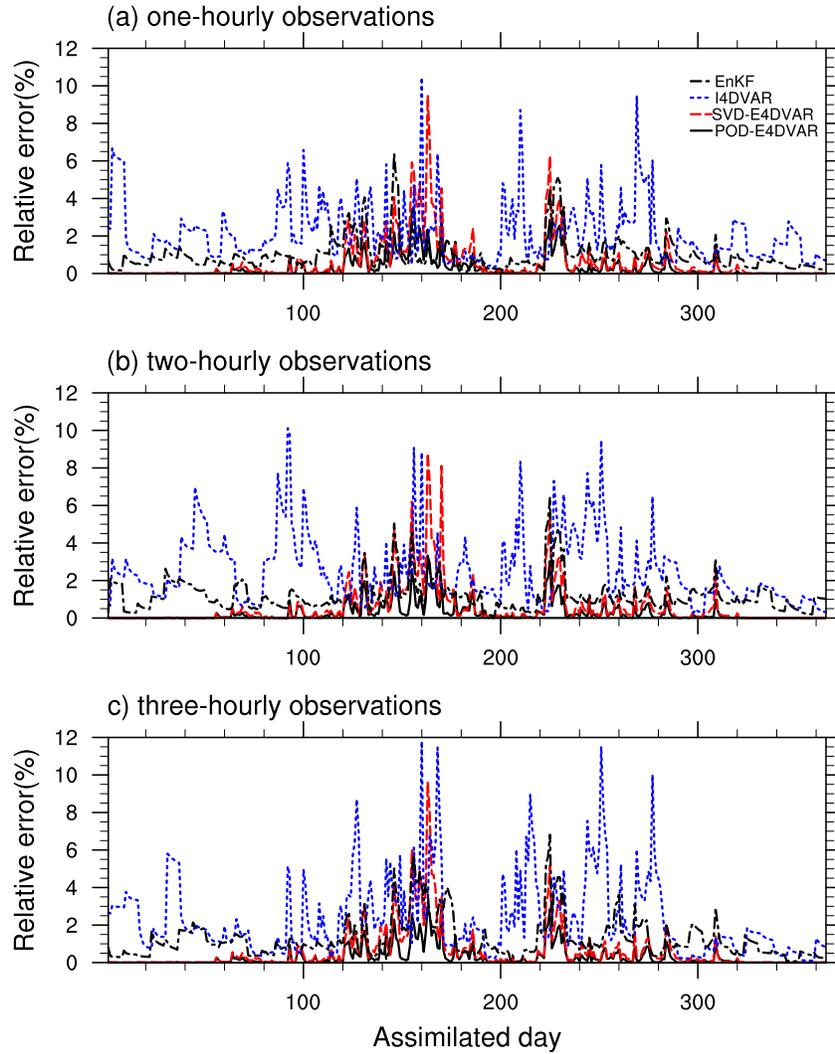


Figure 5. Relative error (E_n) for analyzed soil moisture in the assimilation experiments by the imperfect model with the “imperfect” initial field.

states were produced by integrating the perfect and imperfect models with the “imperfect” initial soil moisture separately: for the former case, the forecast error comes only from the noise in the initial field, but in the latter case it comes from both the noise and the uncertainty in the forecast model.

[22] The length of an assimilation time window in our experiments is one day (48 time steps), i.e., $S = 48$. The size of $X_n = (x_n(t_1), x_n(t_2), \dots, x_n(t_{S-1}))$ in our method is 480, where $x_n(t_i) = (\theta_{n1}(t_i), \theta_{n2}(t_i), \dots, \theta_{n10}(t_i))$ and $M_g = 10$, $M_v = 1$. The background and observational error covariance matrices used in the E/I4DVAR methods can be obtained by using the ensemble covariance matrices defined by equations (A4) and (A8), respectively. We used $\gamma = 0.90$ in our experiments.

[23] Two groups of experiments were done: The perfect model with the “imperfect” initial field as Group 1 and the imperfect model with the “imperfect” initial field as Group 2. Three observation sampling frequencies (hourly, 2-hourly, and 3-hourly) were tested in each group’s experiments. For simplicity, the Full-E4DVAR is only tested in Group 2. The ensemble size used in the Full-, POD- and SVD-E4DVAR and EnKF methods was 60 in this study (the impact of the ensemble size on the assimilation results will be discussed

in another study). The linearization of the soil moisture equation (24) follows the format of *Zhang et al.* [2006].

3.2. Experimental Results

[24] To evaluate the performance of the five algorithms (Full-E4DVAR, SVD-E4DVAR, POD-E4DVAR, I4DVAR and EnKF), a relative error is defined as follows

$$E_{t_0 \rightarrow s-1} = \frac{\sum_{i=0}^{S-1} \sum_{j=1}^{M_g \times M_v} (\bar{x}_j^a(t_i) - \bar{x}_j^t(t_i))^2}{\sum_{i=0}^{S-1} \sum_{j=1}^{M_g \times M_v} (\bar{x}_j^f(t_i) - \bar{x}_j^t(t_i))^2}, \quad (26)$$

where the index $t_{0 \rightarrow s-1}$ denotes an assimilation time window (one day in our experiments), S is the length of an assimilation window ($S = 48$ in our experiments), f and a denote the forecast state (without assimilation of the “observations”) and the analysis state, respectively, t represents the “true” (“perfect”) state. Thus a relative error of 1% for a given assimilation method would mean

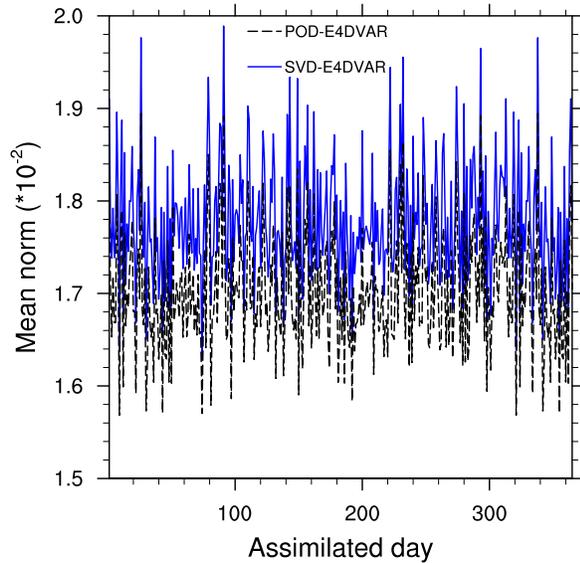


Figure 6. Time series of mean norm for the Full-E4DVAR (solid line), SVD-E4DVAR (long dashed line), and POD-E4DVAR (short dashed line) methods in Group 2 experiments.

that the mean error of the analyzed soil moisture is only 1% of that in the forecast case.

[25] Figures 4–5 show that the POD/SVD-E4DVAR methods perform much better than the EnKF and I4DVAR methods in both groups of experiments. The two explicit 4DVAR methods perform almost same in Group 1 experiments. Their relative errors for analyzed soil moisture are very small (<1%) in the case that the forecast model is perfect, in which the forecast error comes only from the noise of the initial field (Figure 4). However, the relative errors of the EnKF method are many times higher than those of POD/SVD-E4DVAR, around 1~2% or so. The traditional 4DVAR method performs even worse than EnKF, which is consistent with the results of Reichle *et al.* [Reichle and Entekhabi, 2001; Reichle *et al.*, 2002a, 2002b]. This is expected because the soil water hydrodynamic equation (24) is a highly nonlinear system and the tangent linearization operator used in the usual 4DVAR can only propagate analytically with the first-order precision, which introduces large errors in variable estimation and leads to sub-optimal performance.

[26] When the forecast model is imperfect, its forecast error comes from both the noise of the initial field and the uncertainty in the model itself. The relative errors of the four methods all become larger in this case (Figure 5), presumably due to the reduced effect of data assimilation under a poorly constrained model. Nevertheless, the relative errors for POD-E4DAVR are substantially smaller than those of the other methods, including SVD-E4DVAR which performs similarly with the EnKF in this case. The relative errors of POD-E4DVAR are still in the range of 0–6%; however, the relative errors of the I4DVAR and SVD-E4DVAR methods are higher than 6%, and some are even up to 10%. It is also a bit surprising that the SVD-based method is apparently inferior to POD-E4DVAR in some assimilation time windows and even worse than the EnKF method (Figure 4). The relative errors of the Full-E4DVAR

method fluctuate in magnitude between 20 and 60% (not shown). Direct comparisons between the Full-E4DVAR assimilations, the simulations and the truth show that Full-E4DAVR performs even worse than the pure simulation sometimes. This can be explained by the difference between the mean norms of Full-, SVD-, and POD-E4DVAR: the time-averaged mean norms of the SVD- and POD- based methods are around 1.8×10^{-2} and 1.7×10^{-2} respectively in Group 2 experiments (Figure 6). On the contrary, the mean norms of Full-E4DVAR are all higher than 2.06 (not shown). The huge difference between their (Full- and POD/SVD-E4DVAR) mean norms leads to a very small sample density ratio $\left(\left(\frac{1.8 \times 10^{-2}}{2.06} \right)^M \text{ or } \left(\frac{1.7 \times 10^{-2}}{2.06} \right)^M, M = 240 \right)$.

The higher mean norm of Full-E4DVAR results in its lower sample density and poor assimilation performance. Similarly, the ratio between the mean norms of the POD- and SVD-E4DVAR methods is about 0.94, which also affects their sample densities ($\rho_{\text{POD}}/\rho_{\text{SVD}} = \left(\frac{1}{0.94} \right)^M, M = 240$) and makes POD-E4DVAR outperforming SVD-E4DVAR. Figures 4–5 also show that the observation frequency has larger impacts in the I4DVAR method than in the POD-E4DVAR method.

[27] For the two groups of experiments, the ratio of the computational costs for the four methods (POD-E4DVAR, SVD-E4DVAR, I4DVAR, and EnKF) is about 1:1.05:0.5:30. The high computational cost in the EnKF method is mainly due to the fact the analysis process consists of huge matrix and the computation has to be conducted repeatedly when there are observations in the assimilation time window, while in POD-E4DVAR the computation is performed only once in each cycle. The 5% reduction in POD-E4DVAR compared with SVD-E4DVAR results from the application of the matrix transformation technique described in section 2. We also implemented a usual iterative method for optimization [Liu and Nocedal, 1989] into POD-E4DVAR in the same framework to investigate how the direction solution method proposed in this paper reduces its computational costs. The experiments show that the ratio of their (the direct solution method and the iterative method) computational costs varies around 1:10 to 1:5 or so. Of course, this conclusion is not absolute and case-dependent because the scale of the minimization of cost functional and the iteration times during each assimilation cycle vary greatly within different numerical models. The main computational costs of POD-E4DVAR come from the ensemble integrations over the assimilation time window, which can be done on parallel computers. Thus the additional costs of POD-E4DVAR compared with the traditional 4DVAR should not result in real difficulties, and it still costs only one thirtieth of that of the EnKF method in our experiments.

4. Evaluations Within the Lorenz Model

[28] In this section, our approach (POD-E4DVAR) is further evaluated within the Lorenz model for investigating its wider applications. The Lorenz model is widely used to test the new proposed methods in data assimilation community: e.g., Xiong *et al.* [2006] used it to test the performances of the EnKF and PF (particle filter) methods. Their results show that the PFGR (PF with Gaussian resampling) method possesses good stability and accuracy

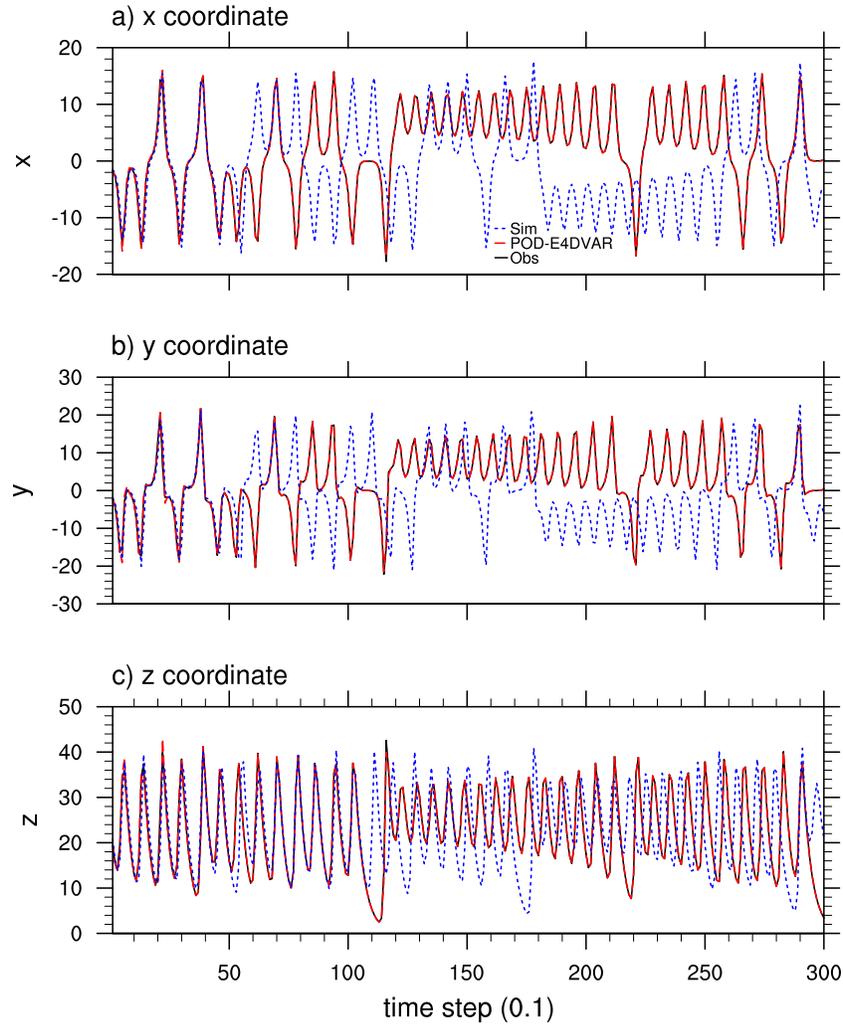


Figure 7. Time series of the Lorenz curve coordinates (x , y , z) from observations (solid line), assimilations (long dashed line), and background forecasts (short dashed line).

and is potentially applicable to large-scale data assimilation problems.

4.1. Setup of Experiments

[29] The Lorenz system under chaotic regime is used as a test problem, which is given by equation (e.g., see <http://www.taygeta.com/perturb/node2.html>):

$$\frac{dx}{dt} = -s(x - y), \quad (27a)$$

$$\frac{dy}{dt} = rx - y - xz, \quad (27b)$$

$$\frac{dz}{dt} = xy - bz, \quad (27c)$$

For numerical experiments the Lorenz system with parameters $s = 10$, $r = 28$, $b = \frac{8}{3}$ was integrated using a second-order Runge Kuatta's method, with $\Delta t = 0.1$, and initial conditions $x(0) = -1.5$, $y(0) = -1.5$, $z(0) = 25$ or the true solution and $x(0) = -1.52$, $y(0) = -1.3$, $z(0) = 27$ for

background solution (a-priori forecast). The observation insertion is done at each 12 time-step. The length of each assimilation time window is 24 time-step.

4.2. Experimental Results

[30] Figure 7 shows time series of the Lorenz curve coordinates (x , y , z) from observations, assimilations and background forecasts: the assimilated Lorenz curve is adjusted to approach the true curve rapidly at the end of the first assimilation cycle, even though there are only two observations in each assimilation time window. On the contrary, the pure forecast state without assimilations begins to deviate from the true solution after 60 time-step or so (Figure 7), even though the noise of the initial filed (x , y , z) only results in small departures from the true state in the first 48 time steps or two assimilation time windows.

5. Summary and Concluding Remarks

[31] To retain the main strength of traditional 4DVAR while avoiding the need of an adjoint or tangent linear model of the forecast model in data assimilation, we have developed an ensemble-based explicit 4DVAR method in

this paper (called POD-E4DVAR). This new method merges the Monte Carlo method and the proper orthogonal decomposition (POD) technique into the 4DVAR to transform an implicit optimization problem into an explicit one. The POD method efficiently approximates a forecast ensemble produced by the Monte Carlo method in a 4-D space using a set of base vectors that span this ensemble and capture its spatial structure and temporal evolution. After the analysis variables are represented by a truncated expansion of the base vectors in the 4-D space, the control (state) variables in the cost function appear explicit, so that the adjoint model, which is used to derive the gradient of the cost function with respect to the control variables in traditional 4DVAR, is no longer needed. This new method significantly simplifies the data assimilation process and retains the two main advantages of the traditional 4DVAR (i.e., dynamic constraint and assimilation of multiple time observations).

[32] Several numerical experiments performed with a simple 1-D soil water equation show that the new POD-E4DVAR method performs much better than the traditional 4DVAR and EnKF method with assimilation errors being reduced to a fraction of the latter two. It is also superior to the SVD-E4DVAR, another explicit 4DVAR method developed by *Qiu et al.* [2007a, 2007b], especially when the forecast model is imperfect and the error comes from both the noise of the initial field and the uncertainty in the forecast model. In our experiments, the traditional (implicit) 4DVAR method performs worst, which is due to errors associated with the tangent linearization operator used in the usual 4DVAR that only propagates analytically with the first-order precision. Another assimilation experiment conducted within the Lorenz model also shows its potential applications in numerical weather or climate models. The results show that the POD-E4DVAR method provides a promising new tool for data assimilation.

[33] Several issues, such as the impacts of the ensemble size and the initial perturbation fields on the assimilated results and the actual performance of this new method in real numerical forecast models, still need to be addressed. It also should be pointed out that since this method begins with a 4-D ensemble obtained from the perturbed ensembles, the quality of the results relies heavily on the perturbation method. How to generate a reasonable perturbed field is a critical step in using this method. This aspect also requires further investigation.

Appendix A: Ensemble Kalman Filter (EnKF) Method

A1. Ensemble Representation for Covariance Matrix

[34] One can define the matrix holding the ensemble members $\vec{x}_i \in R^n$ as

$$A = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \in R^{n \times N}, \quad (A1)$$

where N is the number of ensemble members and n is the size of the model state vector.

[35] The ensemble mean is stored in each column of \bar{A} which can be defined as

$$\bar{A} = A1_N, \quad (A2)$$

where $1_N \in R^{N \times N}$ is a matrix in which each element is equal to $1/N$. One can then define the ensemble perturbation matrix as

$$A' = A - \bar{A} = A(I - 1_N), \quad (A3)$$

The ensemble covariance matrix $P_e \in R^{n \times n}$ can be defined as

$$P_e = \frac{A'(A')^T}{N - 1}. \quad (A4)$$

A2. Measurement Perturbations

[36] Given a vector of measurements $y \in R^m$, with m being the number of measurements, one can define N vectors of perturbed observations as

$$\vec{y}_j = \vec{y} + \varepsilon_j, j = 1, 2, \dots, N, \quad (A5)$$

which can be stored in the columns of a matrix

$$Y = (\vec{y}_1, \vec{y}_2, \dots, \vec{y}_N) \in R^{m \times N}, \quad (A6)$$

while the ensemble of perturbations, with ensemble mean equal to zero, can be stored in the matrix

$$E = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N) \in R^{m \times N}, \quad (A7)$$

from which we can construct the ensemble representation of the measurement error covariance matrix

$$R_e = \frac{EE^T}{N - 1}, \quad (A8)$$

A3. Analysis Equation

[37] The analysis equation, expressed in terms of the ensemble covariance matrices, is

$$A^a = A + P_e H^T (HP_e H^T + R_e)^{-1} (D - HA). \quad (A9)$$

Using the ensemble of innovation vectors defined as

$$D' = D - HA, \quad (A10)$$

and the definitions of the ensemble error covariance matrices in equations (A4) and (A8), the analysis can be expressed as

$$A^a = A + A'A'^T H^T (HA'A'^T H^T + EE^T)^{-1} D'. \quad (A11)$$

When the ensemble size, N , is increased by adding random samples, the analysis computed from this equation will converge toward the exact solution of equation (A9) with P_e and R_e replaced by the exact covariance matrices P and R .

Appendix B: SVD-E4DVAR Method

[38] Assuming there are m observations \vec{y}_i ($i = 0, 1, \dots, m - 1$) at time $t = t_0, \dots, t_i, \dots, t_{m-1}$ during the

assimilation time window. Generate N random perturbation fields and add each to the initial background field and integrate the model to produce a perturbed 4-D field over the analysis time window. The i th difference field is then given by $\delta \vec{x}_i = \vec{x}_i - \vec{x}_b$ at time $t = t_0, \dots, t_i, \dots, t_m - 1$, where \vec{x}_b, \vec{x}_i denote the background and the perturbed fields, respectively. Consider an ensemble of column vectors represented by matrix $A = (\delta \vec{X}_1, \delta \vec{X}_2, \dots, \delta \vec{X}_N)$, where the i th column vector $\delta \vec{X}_i$ represents the i th sampled data field in a discrete four-dimensional analysis space. The length of vector $\delta \vec{X}_i$ is $M_g \times M_v \times m$, where M_g, M_v are the number of the model spatial grid points and the number of the model variables, respectively. The SVD of A yields

$$A = B\Lambda V^T, \quad (B1)$$

where Λ is a diagonal matrix composed of the singular values of A with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ and $\lambda_{r+1} = \lambda_{r+2} = \dots = 0$, $r \leq \min(M_g \times M_v \times m, N)$, is the rank of A , B and V are orthogonal matrices composed of the left and right singular vectors of A , respectively. The SVD in (B1) gives $C = A^T A = V\Lambda^2 V^T$ and $Q = AA^T = B\Lambda^2 B^T$. Thus the i th column vector of V , denoted by V_i , is the i th eigenvector of C , while the j th column vector of B , denoted by b_j , is the j th column vector of Q and is called the singular vector of A .

[39] The truncated reconstruction of analysis variable \vec{X}_a in 4-D space is given by

$$\vec{X}_a = \vec{X}_b + \sum_{i=1}^P \alpha_i b_i, \quad (B2)$$

where $P(\leq r)$ is the truncation number, which can be obtained through equation (10) in section 2, $\vec{X}_b = (\vec{x}_b, \vec{x}_b, \dots, \vec{x}_b)$ is composed of m vectors (\vec{x}_b) .

[40] Substituting (B2) into equation (1) in section 2, the control variable becomes α instead of \vec{x}_0 , so the control variable is expressed explicitly in the cost function.

Appendix C: Proper Orthogonal Decomposition

C1. Continuous Case

[41] Let $U_i(\vec{x})$, $i = 1, 2, \dots, N$ denote the set of N observations or simulations (also called *snapshots*) of some physical process taken at position $\vec{x} = (x, y)$. The average of the ensemble snapshots is given by

$$\bar{U} = \frac{1}{N} \sum_{i=1}^N U_i(\vec{x}), \quad (C1)$$

[42] We form new ensemble by focusing on deviation from mean as follows:

$$V_i = U_i - \bar{U}, \quad (C2)$$

[43] We wish to find an optimal compressed description of the sequence of data (C2). One description of the process

is a series expansion in terms of a set of base functions. Intuitively, the base functions should in some sense be representative of the members of the ensemble. Such a coordinate system, is provided by the Karhunan Loève expansion, where the base functions Φ are, in fact, admixtures of the snapshots and are given by:

$$\Phi = \sum_{i=1}^N a_i V_i(\vec{x}), \quad (C3)$$

Here, the coefficients a_i are to be determined so that Φ given by (C3) will resemble the ensemble $\{V_i(\vec{x})\}_{i=1}^N$ most closely. More specifically, we look for a function Φ to maximize

$$\frac{1}{N} \sum_{i=1}^N |(V_i, \Phi)|^2, \quad (C4)$$

subjected to $(\Phi, \Phi) = \|\Phi\|^2 = 1$, where (\cdot, \cdot) and $\|\cdot\|$ denote the usual L^2 inner product and L^2 -norm, respectively.

[44] It follows that the base functions are the eigenfunctions of the integral equation

$$\int C(\vec{x}, \vec{x}') \Phi(\vec{x}') d\vec{x}' = \lambda \Phi(\vec{x}), \quad (C5)$$

Substituting (C3) into (C5) yields the eigenvalue problem:

$$\sum_{j=1}^N L_{ij} a_j = \lambda a_i, \quad (C6)$$

where $L_{ij} = \frac{1}{N} (V_i, V_j)$ is a symmetric and nonnegative matrix. Thus our problem amounts to solving for the eigenvectors of an $N \times N$ matrix, where N is the ensemble size of the snapshots. Straightforward calculation shows that the cost function

$$\frac{1}{N} \sum_{i=1}^N |(V_i, \Phi)|^2 = (\lambda \Phi, \Phi) = \lambda, \quad (C7)$$

is maximized when the coefficients a_i 's of (C3) are the elements of the eigenvector corresponding to the largest eigenvalue of L .

C2. Discrete Case

[45] We consider the discrete Karhunan Loève expansion to find an optimal representation of the ensemble of snapshots. In the two-dimensional case, each sample of snapshots $U_i(x, y)$ (defined on a set of $n \times n$ nodal points (x, y)) can be expressed as an n^2 dimensional vector \vec{u}_i as follows:

$$\vec{u}_i = \begin{bmatrix} \vec{u}_{i1} \\ \vdots \\ \vec{u}_{ij} \\ \vdots \\ \vec{u}_{in^2} \end{bmatrix}, \quad (C8)$$

where \vec{u}_{ij} denotes the j th component of the vector \vec{u}_i . Here the discrete covariance matrix of the ensemble \vec{u} is defined as

$$C_{\vec{u}} = E \left\{ \left(\vec{u} - \vec{m}_{\vec{u}} \right) \left(\vec{u} - \vec{m}_{\vec{u}} \right)^T \right\}, \quad (\text{C9})$$

where

$$\vec{m}_{\vec{u}} = E \{ \vec{u} \} \quad (\text{C10})$$

is the mean vector, E is the expected value. Equations (C9) and (C10) can be replaced by

$$C_{\vec{u}} = \frac{1}{N} \left[\sum_{i=1}^N \vec{u}_i \vec{u}_i^T \right] - \vec{m}_{\vec{u}} \vec{m}_{\vec{u}}^T$$

and $\vec{m}_{\vec{u}} = \frac{1}{N} \sum_{i=1}^N \vec{u}_i$ respectively.

[46] **Acknowledgments.** This work was supported by the National Natural Science Foundation of China under grant 40705035, the Knowledge Innovation Project of Chinese Academy of Sciences under grants KZCX2-YW-217 and KZCX2-YW-126-2, the National Basic Research Program under grant 2005CB321704, and the Chinese COPES project (GYHY200706005). The National Center for Atmospheric Research is sponsored by the U.S. National Science Foundation.

References

- Bauer, P., P. Lopez, A. Benedetti, D. Salmond, S. Saarinen, and M. Bonazzola (2006), Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var, *Q. J. R. Meteorol. Soc.*, *132*, 2307–2332.
- Beezley, J. D., and J. Mandel (2008), Morphing ensemble Kalman filters, *Tellus, Ser. A*, *60*(1), 131–140.
- Bormann, N., and J. N. Thepaut (2004), Impact of MODIS polar winds in ECMWF's 4DVAR data assimilation system, *Mon. Weather Rev.*, *132*, 929–940.
- Cao, Y., J. Zhu, I. M. Navon, and Z. D. Luo (2007), A reduced order approach to four-dimensional variational data assimilation using proper orthogonal decomposition, *Int. J. Numer. Methods Fluids*, *53*, 1571–1583, doi:10.1002/fld.1365.
- Caya, A., J. Sun, and C. Snyder (2005), A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation, *Mon. Weather Rev.*, *133*, 3081–3094.
- Courtier, P., J. N. Thepaut, and A. Hollingsworth (1994), A strategy for operational implementation of 4DVAR, using an incremental approach, *Q. J. R. Meteorol. Soc.*, *120*, 1367–1387.
- Darcy, H. (1856), *Les Fontaines Publiques de la Ville de Dijon*, Dalmont, Paris, France.
- Liu, D. C., and J. Nocedal (1989), On the limited memory BFGS method for large scale optimization, *Math. Program.*, *45*, 503–528.
- Evensen, G. (1994), Sequential data assimilation with a non-linear geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162.
- Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, *53*, 343–367.
- Farrell, B. F., and P. J. Ioannou (2001), State estimation using a reduced order Kalman filter, *J. Atmos. Sci.*, *58*, 3666–3680.
- Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau (2007), Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada, *Mon. Weather Rev.*, *135*(6), 2339–2354.
- Heemink, A. W., M. Verlaan, and A. J. Segers (2001), Variance reduced ensemble Kalman filtering, *Mon. Weather Rev.*, *129*(7), 1718–1728.
- Hunt, B. R., et al. (2004), Four-dimensional ensemble Kalman filtering, *Tellus, Ser. A*, *56*, 273–277.
- John, H., and B. R. Hunt (2007), Four-dimensional local ensemble transform Kalman filter: Numerical experiments with a global circulation model, *Tellus, Ser. A*, *59*, 731–748.
- Johnson, C., B. J. Hoskins, N. K. Nichols, and S. P. Ballard (2006), A singular vector perspective of 4DVAR: The spatial structure and evolution of baroclinic weather systems, *Mon. Weather Rev.*, *134*(11), 3436–3455.
- Kalnay, E., et al. (2000), Application of the quasi-inverse method to data assimilation, *Mon. Weather Rev.*, *128*, 864–875.
- Kalnay, E., H. Li, T. Miyoshi, S. C. Yang, and J. Ballabrera-Poy (2007), 4-D-Var or ensemble Kalman filter?, *Tellus, Ser. A*, *59*(5), 758–773.
- Luo, Z., Z. Jiang, W. Ruiwen, and I. Michael Navon (2007), Proper orthogonal decomposition approach and error estimation of mixed finite element methods for the tropical Pacific Ocean reduced gravity model, *Comput. Methods Appl. Mech. Eng.*, *196*(41–44), 4184–4195.
- Ly, H. V., and H. T. Tran (2001), Modeling and control of physical processes using proper orthogonal decomposition, *Math. Comput. Modeling*, *33*, 223–236.
- Ly, H. V., and H. T. Tran (2002), Proper orthogonal decomposition for flow calculations and optimal control in a horizontal CVD reactor, *Q. Appl. Math.*, *60*(3), 631–656.
- Mu, M., and J. Wang (2003), A method for adjoint variational data assimilation with physical “on-off” processes, *J. Atmos. Sci.*, *60*, 2010–2018.
- Oleson, K. W., et al. (2004), *Technical Description of the Community Land Model (CLM)*, 186 pp., NCAR/TN-461+STR, available as a PDF-file. (Available at http://www.cgd.ucar.edu/tss/clm/distribution/clm3.0/TechNote/CLM_Tech_Note.pdf)
- Park, K., and X. Zou (2004), Toward developing an objective 4DVAR BDA scheme for hurricane initialization based on TPC observed parameters, *Mon. Weather Rev.*, *132*(8), 2054–2069.
- Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2004. part I: Forcing data and evaluations, *J. Hydrometeorol.*, *7*, 953–975.
- Qiu, C. J., and J. F. Chou (2006), Four-dimensional data assimilation method based on SVD: Theoretical aspect, *Theoretical Appl. Climatol.*, *83*, 51–57.
- Qiu, C. J., L. Zhang, and A. M. Shao (2007a), An explicit four-dimensional variational data assimilation method, *Sci. China (D)*, *50*(8), 1232–1240.
- Qiu, C. J., A. M. Shao, Q. Xu, and L. Wei (2007b), Fitting model fields to observations by using singular value decomposition: An ensemble-based 4DVar approach, *J. Geophys. Res.*, *112*, D11105, doi:10.1029/2006JD007994.
- Reichle, R. H., and D. Entekhabi (2001), Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, *Water Resour. Res.*, *37*(9), 2353–2364.
- Reichle, R. H., et al. (2002a), Extended versus ensemble filtering for land data assimilation, *J. Hydrometeorol.*, *3*, 728–740.
- Reichle, R. H., D. B. McLaughlin, and D. Entekhabi (2002b), Hydrologic data assimilation with the ensemble Kalman filter, *Mon. Weather Rev.*, *130*, 103–114.
- Rosmond, T., and L. Xu (2006), Development of NAVDAS-AR: Non-linear formulation and outer loop tests, *Tellus, Ser. A*, *58*(1), 45–58.
- Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield, and J. A. Yorke (2008), A local ensemble transform Kalman filter data assimilation system for the NCEP global model, *Tellus, Ser. A*, *60*(1), 113–130.
- Tian, X., A. Dai, D. Yang, and Z. Xie (2007), Effects of precipitation-bias corrections on surface hydrology over northern latitudes, *J. Geophys. Res.*, *112*, D14101, doi:10.1029/2007JD008420.
- Tsuyuki, T., and T. Miyoshi (2007), Recent progress of data assimilation methods in meteorology, *J. Meteorol. Soc. Jpn.*, *85B*, 331–361.
- Vermeulen, P. T. M., and A. W. Heemink (2006), Model-reduced variational data assimilation, *Mon. Weather Rev.*, *134*(10), 2888–2899.
- Wang, B., and Y. Zhao (2005), A new data assimilation approach, *Acta Meteorol. Sin.*, *63*, 694–700, (in Chinese).
- Uzunoglu, B., S. J. Fletcher, M. Zupanski, and I. M. Navon (2007), Adaptive ensemble reduction and inflation, *Q. J. R. Meteorol. Soc.*, *133*(626), 1281–1294, Part A.
- Volkwein, S. (2008), Model reduction using proper orthogonal decomposition (script in English language), 41 pp., available as a PDF-file. (Available at <http://www.uni-graz.at/imawww/volkwein/publist.html>)
- Xiong, X., I. M. Navon, and B. Uzunoglu (2006), A note on the particle filter with posterior Gaussian resampling, *Tellus, Ser. A*, *58A*, 456–460.
- Xu, Q. (1996), Generalized adjoint for physical processes with parameterized discontinuities. part I: Basis issues and heuristic examples, *J. Atmos. Sci.*, *53*(8), 1123–1142.
- Zhang, S. L., et al. (2006), A soil moisture assimilation scheme with an unsaturated soil water flow model and in-site observation, *Adv. Earth Sci.*, *12*, 1350–1362, (in Chinese).
- Zupanski, M. (2005), Maximum likelihood ensemble filter: Theoretical aspects, *Mon. Weather Rev.*, *133*, 1710–1726.

A. Dai, CGD, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, USA. (adai@ucar.edu)

X. Tian and Z. Xie, ICCES, Institute of Atmospheric Physics, Chinese Academy of Sciences, P.O. Box 9804, Beijing 100029, China. (tianxj@mail.iap.ac.cn; zxie@lasg.iap.ac.cn)