A Dual-Weighted Approach to Order Reduction in 4D-Var Data Assimilation

D. N. Daescu *
Department of Mathematics and Statistics
Portland State University, P.O. Box 751, Portland, OR 97207

I. M. Navon
Department of Mathematics and School of Computational Science
Florida State University, Tallahassee, FL 32306, U.S.A.

November 16, 2006

*Corresponding author address: Dr. Dacian N. Daescu, Department of Mathematics and Statistics, Portland State University, P.O. Box 751, Portland, OR 97207, U.S.A.; E-mail: daescu@pdx.edu

Abstract

Strategies to achieve order reduction in four dimensional variational data assimilation (4D-Var) search for an optimal low rank state subspace for the analysis update. A common feature of the reduction methods proposed in atmospheric and oceanographic studies is that the optimality criteria to compute the basis functions relies on the model dynamics only, without properly accounting for the specific details of the data assimilation system (DAS). In this study a general framework of the proper orthogonal decomposition (POD) method is considered and a cost-effective approach is proposed to incorporate DAS information into the order reduction procedure. The sensitivities of the cost functional in 4D-Var data assimilation with respect to the time varying model state are obtained from a backward integration of the adjoint model. This information is further used to define appropriate weights and to implement a dual-weighted proper orthogonal decomposition (DWPOD) method to order reduction. The use of a weighted ensemble data mean and weighted snapshots using the adjoint DAS is a novel element in reduced order 4D-Var data assimilation. Numerical results are presented with a global shallow-water model based on the Lin-Rood flux-form semi-Lagrangian scheme. A simplified 4D-Var DAS is considered in the twin-experiments framework with initial conditions specified from the ECMWF ERA-40 data sets. A comparative analysis with the standard POD method shows that the reduced DWPOD basis provides an increased efficiency in representing a specified model forecast aspect and as a tool to perform reduced order optimal control. This approach represents a first step toward the development of an order reduction methodology that combines in an optimal fashion the model dynamics and the characteristics of the 4D-Var DAS.

1. Introduction

Implementation of modern data assimilation techniques as formulated in the context of estimation theory (Jazwinski 1970, Lorenc 1986, Daley 1991, Bennett 1992, Cohn 1997, Kalnay 2002) is often hampered by the high computational cost to obtain the analysis state and to dynamically evolve the error statistics. A characteristic feature of the global ocean and atmospheric circulation models is the large dimensionality of the discrete state vector, typically in the range $10^6 - 10^7$. This dimension is likely to increase in the near future when climate models are envisaged to run at a horizontal resolution as high as 1/4 degree in forecast and data assimilation mode. To accommodate these requirements, computationally efficient techniques to assimilate an ever increasing amount of observational data into models must be developed.

Significant efforts have been dedicated to ease the computational burden of Kalman filter based algorithms through various simplifying assumptions. State reduction techniques and low-rank approximations of the error covariance matrix are described in the work of (Dee 1990), Todling and Cohn (1994), Cane et al. (1996), Pham et al. (1998), Hoteit and Pham (2003). Ensemble Kalman filter (EnKF) methods build on the original work of Evensen (1994) to provide the analysis state and error covariance using an ensemble of model forecasts (Molteni et al. 1996, Burgers et al. 1998, Anderson 2001). A review of the EnKF and low-rank filters can be found in the work of Evensen (2003) and Nerger et al. (2005) who emphasize that a common feature of these methods is that their analysis step operates in a low-dimensional subspace of the true error space.

In four dimensional variational data assimilation (4D-Var) the analysis state is obtained by solving a large-scale optimization problem (Le Dimet and Talagrand 1986) with the initial conditions of the discrete model as control parameters. The incremental approach (Courtier et al. 1994) is currently used at numerical weather prediction centers implementing 4D-Var (Rabier et al. 2000). Computational savings are further achieved by running a coarse resolution tangent linear and adjoint models in the inner loop of the minimization. Implementation issues and a study on the convergence of the incremental 4D-Var method are provided by Trémolet (2004, 2005).

Although running a coarse resolution model provides a certain state reduction, the issue of finding an optimal low-dimensional state subspace for the 4D-Var minimization problem is an open question where the current state of research is at an incipient stage. Mathematical foundations of approximation theory for large-scale dynamical systems and flow control are presented by Antoulas (2005) and Gunzburger (2003). A substantial amount of work was done in the climate research community to build reduced models of the atmospheric dynamics with a minimal number of degrees of freedom. The proper orthogonal decomposition (POD) method (also known as the method of empirical orthogonal functions - EOFs, Karhunen-Loève decomposition) has been widely used in fluid dynamics (Holmes, Lumley and Berkooz 1996, Sirovich 1987) and atmospheric flow modeling (Selten 1995, 1997, Achatz and Opsteegh 2003) to obtain basis functions for reduced order dynamics. Shortcomings of the POD/EOFs reduced models are discussed by Aubry et al. (1993) and in practice other choices should be also considered. In particular, principal interaction patterns (Hasselmann 1988) have shown the potential to achieve improved results when compared to EOFs (Achatz and Schmitz 1997, Kwasniok 2004, Crommelin and Majda 2004). While these studies were only concerned with the construction and analysis of reduced models to the atmospheric flow, the development and implementation of optimal order-reduction strategies in the context of 4D-Var atmospheric data assimilation is a far more difficult task.

For oceanic models, initial efforts on reduced order 4D-Var were put forward by Blayo et al. (1998) and Durbiano (2001). The use of EOFs to identify a low-rank control space has shown promising results in the studies of Robert et al. (2005), Hoteit and Köhl (2006), Cao et al. (2006). The potential use of the reduced order 4D-Var as a preconditioner to 4D-Var data assimilation was considered by Robert et al. (2006). A common feature of the reduction methods used in these studies is that the optimality criteria to compute the basis functions relies on the model dynamics only, without properly accounting for the specific details of the data assimilation system (DAS). As such, the efficiency of the reduced basis may be impaired by the lack of information on the optimization problem at hand.

Meyer and Matthies (2003) used adjoint modeling to improve the efficiency of the POD approach to model reduction when targeting a scalar aspect of the model dynamics. A method to achieve balanced model reduction of linear systems using POD and potential extensions to nonlinear dynamics are discussed by Willcox and Peraire (2002). A goal-oriented, model-constrained optimization framework to reduction of large-scale models is presented in the work of Bui-Thanh et al. (2007).

In this work we consider a novel method to incorporate DAS information into the order reduction procedure by implementing a dual-weighted proper orthogonal decomposition (DWPOD) method. The DWPOD method searches to provide an "enriched" set of basis functions that combine information from both model dynamics and DAS. The use of a weighted ensemble data mean and weighted snapshots using the adjoint DAS is a novel element in reduced order 4D-Var data assimilation. The traditional POD basis consists on the modes that capture most of the "energy" of the dynamical system whereas the DWPOD basis may include lower energy modes that are more significant to the representation of the 4D-Var cost functional. The DWPOD procedure is shown to be cost-effective since it provides a substantial qualitative improvement as compared to the standard POD approach at the additional computational expense of a single adjoint model integration.

Henceforth, the paper is organized as follows: in Section 2 the 4D-Var data assimilation problem is briefly revisited. A general POD framework to reduced order 4D-Var and the dual weighted POD approach are described in Section 3. Numerical experiments with a finite volume global shallow water model are provided in Section 4. Concluding remarks and further research directions are presented in Section 5.

2. The 4D-Var data assimilation problem

The 4D-Var data assimilation searches for an optimal estimate (analysis) $\mathbf{x}_0^a$ to the $m-$dimensional vector of the discrete model initial conditions by solving a large-scale optimization problem

$$\min_{\mathbf{x}_0 \in R^m} \mathcal{J}(\mathbf{x}_0); \quad \mathbf{x}_0^a = \text{ Arg } \min \mathcal{J} \tag{1}$$

The cost functional

$$\mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{k=1}^{N}(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k)^T \mathbf{R}_k^{-1}(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k) \tag{2}$$

includes the distance to a prior (background) estimate to initial conditions $\mathbf{x}_b$ and the distance of the model forecast $\mathbf{x}_k = \mathcal{M}(\mathbf{x}_0)$ to observations $\mathbf{y}_k$, $k = 1, 2, \ldots N$ time distributed over the analysis interval $[t_0, t_N]$. The model $\mathcal{M}$ is nonlinear and for simplicity we assume a linear representation of the observational operator $\mathbf{H}_k$ that maps the state space into the observation space at time t_k. Statistical information on the errors in the background and data is used to define appropriate weights: $\mathbf{B}$ is the covariance matrix of the background errors and $\mathbf{R}_k$ is the covariance matrix of the observational errors. An accurate estimation of the matrix $\mathbf{B}$ is difficult to provide and, given its huge dimensionality, simplifying approximations are required for the practical implementation (Lorenc et al. 2000). Information on the errors statistics may be obtained using differences between forecasts with different initialization times as in the "NMC" method (Parrish and Derber 1992) or ensemble methods based on a perturbed forecast-analysis system. Recent advances in modeling flow-dependent background error variances are discussed by

Kucukkaraca and Fisher (2006).

3. A general POD framework to reduced-order 4D-Var data assimilation

The specification of the basis functions lies at the core of the reduced-order 4D-Var procedure. The proper orthogonal decomposition (POD) method provides an optimal low-rank representation of an ensemble data set $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\},\ \mathbf{x}^{(i)} \in R^m$ that may be collected from observational data and/or the state evolution at various instants in time $t_1, t_2, \ldots, t_n$ (method of snapshots, Sirovich 1987). The use of data weighting as a tool to improve the performance of the POD method was previously considered in model reduction for dynamical systems. Graham and Kevrekidis (1996) proposed an ensemble average based on the arc-length in the phase space and emphasized that the choice of the ensemble average (weights) for the POD method can have a significant impact on the selection of the dominant modes. A weighted POD (w-POD) approach is discussed by Christensen et al. (2000) who consider including multiple copies of an "important" snapshot in the ensemble data set. Kunisch and Volkwein (2002) use the time distribution of the snapshots $\Delta t_i = t_{i+1} - t_i$ to specify weights and provide a detailed theoretical framework and error estimates with applications to Navier-Stokes equations.

We define the weighted ensemble average of the data as

$$\overline{\mathbf{x}} = \sum_{i=1}^{n} \omega_i \mathbf{x}^{(i)} \tag{3}$$

where the snapshot weights ω_i are such that $0 < \omega_i < 1, \sum_{i=1}^{n} \omega_i = 1$ and are used to assign a degree of importance to each member of the ensemble. Time weighting is usually considered and in the standard approach $\omega_i = 1/n$. A modified $m \times n$ dimensional matrix is obtained by subtracting the mean from each snapshot

$$\mathbf{X} = \left[\mathbf{x}^{(1)} - \overline{\mathbf{x}}, \mathbf{x}^{(2)} - \overline{\mathbf{x}}, \ldots \mathbf{x}^{(n)} - \overline{\mathbf{x}}\right] \tag{4}$$

and the weighted covariance matrix $\mathbf{C} \in R^{m \times m}$ is defined

$$\mathbf{C} = \mathbf{X}\mathbf{W}\mathbf{X}^T \tag{5}$$

where $\mathbf{W} = diag\{\omega_1, \ldots, \omega_n\}$ is the diagonal matrix of weights. Since the metric on the state space is often related to the physical properties of the system, we consider a general norm $\|\mathbf{x}\|_{\mathbf{A}}^2 = \langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{A}} = \mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{A} \in R^{m \times m}$ is a symmetric positive definite matrix. For the standard Euclidean norm $\mathbf{A}$ is the identity matrix and for the total energy metric $\mathbf{A}$ is a diagonal matrix.

The POD basis of order $k \leq n$ provides an optimal representation of the ensemble data in a k-dimensional state subspace by minimizing the averaged projection error

$$\min_{\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_k\}} \sum_{i=1}^{n} \omega_j \left\| (\mathbf{x}^{(i)} - \overline{\mathbf{x}}) - \mathcal{P}_{\boldsymbol{\psi},k}(\mathbf{x}^{(i)} - \overline{\mathbf{x}}) \right\|_{\mathbf{A}}^2 \tag{6}$$

subject to the $\mathbf{A}$-orthonormality constraint $\langle \boldsymbol{\psi}_i, \boldsymbol{\psi}_j \rangle_{\mathbf{A}} = \delta_{i,j}, 1 \leq i, j \leq k$, where $\mathcal{P}_{\boldsymbol{\psi},k}$ is the projection operator onto the k-dimensional space $Span\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_k\}$

$$\mathcal{P}_{\boldsymbol{\psi},k}(\mathbf{x}) = \sum_{i=1}^{k} \langle \mathbf{x}, \boldsymbol{\psi}_i \rangle_{\mathbf{A}} \boldsymbol{\psi}_i$$

The POD modes are m-dimensional eigenvectors to the eigenvalue problem

$$\mathbf{C}\mathbf{A}\boldsymbol{\psi}_i = \sigma_i^2 \boldsymbol{\psi}_i \tag{7}$$

and since in practice the number of snapshots is much smaller than the state dimension, $n \ll m$, an efficient way to compute the reduced basis is to solve the n-dimensional eigenvalue problem

$$\mathbf{W}^{\frac{1}{2}} \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{W}^{\frac{1}{2}} \boldsymbol{\mu}_i = \sigma_i^2 \boldsymbol{\mu}_i \tag{8}$$

to obtain the orthonormal (Euclidean) eigenvectors $\boldsymbol{\mu}_i \in R^n$ then to compute the POD modes as

$$\boldsymbol{\psi}_i = \frac{1}{\sigma_i} \mathbf{X} \mathbf{W}^{\frac{1}{2}} \boldsymbol{\mu}_i \tag{9}$$

From (8) and (9) the close relationship with the singular value decomposition (Golub and Van Loan 1996)

$$\mathbf{A}^{\frac{1}{2}}\mathbf{X}\mathbf{W}^{\frac{1}{2}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \tag{10}$$

is established: $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_n \geq 0$ are the singular values, $\boldsymbol{\mu}_i$ the right singular vectors and $\mathbf{A}^{\frac{1}{2}}\boldsymbol{\psi}_i$ the left singular vectors. The fraction of total information ("energy") captured by the dominant k modes is $I(k) = (\sum_{i=1}^{k}\sigma_i^2)/(\sum_{i=1}^{n}\sigma_i^2)$ and in practice, given a tolerance $0 < \gamma \leq 1$ in the vicinity of the unity, k is selected such that $I(k) \geq \gamma$.

a. The reduced order 4D-Var

The k-dimensional reduced order control problem is obtained by projecting $\mathbf{x}_0 - \overline{\mathbf{x}}$ onto the POD space

$$\mathcal{P}_{\boldsymbol{\psi},k}(\mathbf{x}_0 - \overline{\mathbf{x}}) = \boldsymbol{\Psi}\boldsymbol{\eta} = \sum_{i=1}^{k}\eta_i\boldsymbol{\psi}_i \tag{11}$$

where the matrix $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k] \in R^{m\times k}$ has the POD basis vectors as columns and $\boldsymbol{\eta} = (\eta_1, \ldots \eta_k)^T \in R^k$ is the coordinates vector in the reduced space

$$\eta_i = \boldsymbol{\psi}_i^T\mathbf{A}(\mathbf{x}_0 - \overline{\mathbf{x}}), \quad \boldsymbol{\eta} = \boldsymbol{\Psi}^T\mathbf{A}(\mathbf{x}_0 - \overline{\mathbf{x}}) \tag{12}$$

The large-scale 4D-Var optimization (1) is thus replaced by the reduced order 4D-Var problem of finding the optimal coefficients $\boldsymbol{\eta}$

$$\hat{\mathcal{J}}(\boldsymbol{\eta}) := \mathcal{J}(\overline{\mathbf{x}} + \boldsymbol{\Psi}\boldsymbol{\eta}); \quad \min_{\boldsymbol{\eta}\in R^k}\hat{\mathcal{J}}(\boldsymbol{\eta}) \tag{13}$$

If $\boldsymbol{\eta}^a$ denotes the solution to (13), an approximation to the analysis (1) is obtained as

$$\mathbf{x}_0^a \approx \overline{\mathbf{x}} + \boldsymbol{\Psi}\boldsymbol{\eta}^a \tag{14}$$

It should be noticed that in the reduced-order 4D-Var as formulated in (13) only the initial conditions are projected into the POD state subspace and the cost functional is computed using the full model dynamics. The gradient of the cost (13) is expressed as

$$\nabla_{\boldsymbol{\eta}} \hat{\mathcal{J}}(\boldsymbol{\eta}) = \boldsymbol{\Psi}^T \left(\nabla_{\mathbf{x}_0} \mathcal{J} \right) |_{\mathbf{x}_0 = \overline{\mathbf{x}} + \boldsymbol{\Psi}\boldsymbol{\eta}} \tag{15}$$

and its evaluation requires integration of the full adjoint model. Second order derivatives in the reduced space may be computed if a full second order adjoint model is available (Daescu and Navon 2006). Consequently, computational savings may be achieved only by a drastic reduction in the number of iterations due to the low dimension of the optimization problem (13).

Once the POD basis is selected, a *reduced model* approach to order reduction may be also considered by projecting the full model dynamics into the POD space. The projected state $\hat{\mathbf{x}}(t) = \overline{\mathbf{x}} + \boldsymbol{\Psi}\boldsymbol{\eta}(t)$ evolves in time according to the differential equations system

$$\hat{\mathbf{x}}'(t) = \boldsymbol{\Psi}\boldsymbol{\Psi}^T \mathbf{A} \mathcal{M}(\hat{\mathbf{x}}, t) \tag{16}$$

$$\hat{\mathbf{x}}(0) = \boldsymbol{\Psi}\boldsymbol{\Psi}^T \mathbf{A} (\mathbf{x}(0) - \overline{\mathbf{x}}) + \overline{\mathbf{x}} \tag{17}$$

and the coefficients $\boldsymbol{\eta}(t)$ may be obtained by integrating the reduced model equations

$$\boldsymbol{\eta}'(t) = \boldsymbol{\Psi}^T \mathbf{A} \mathcal{M}(\overline{\mathbf{x}} + \boldsymbol{\Psi}\boldsymbol{\eta}(t), t) \tag{18}$$

$$\boldsymbol{\eta}(0) = \boldsymbol{\Psi}^T \mathbf{A} (\mathbf{x}(0) - \overline{\mathbf{x}}) \tag{19}$$

Such approach may result in significant computational savings when Galerkin type numerical schemes are implemented (Ravindran 2002, Kunisch and Volkwein 1999) or an implicit time integration scheme to finite differences/finite volume semi-discretization is considered (van Doren et al. 2006). However, for finite difference and finite volume numerical methods with explicit time schemes, integration of the reduced model equations (18-19) will require in general an increased CPU time due to the cost of repeated projection operations (unless analytic simplifications can be made). An additional issue in the *reduced model* approach is that the projection (16-17) introduces a *model error* that is difficult to quantify (Rathinam and Petzold 2003) and thus to account for in the reduced 4D-Var data assimilation.

b. The dual-weighted POD basis

The specification of the weights ω_i to the snapshots may have a significant impact on which modes are selected as dominant and thus inserted into the POD basis. The dual-weighted approach we propose makes use of the time varying sensitivities of the 4D-Var cost functional with respect to perturbations in the state at the time instants $t_i, i = 1, \ldots, n$ when the snapshots are taken.

The use of the adjoint modeling to identify "target" regions where observational data is of most benefit to a forecast aspect $\mathcal{J}(\mathbf{x})$ is well established in the context of targeted observations for high impact weather events (Langland et al. 1999). By analogy, the dual-weighted approach may be taught as a targeting in time procedure (rather than targeting the state space at a given time) that assigns weights to time distributed snapshots data. For simplicity of the presentation, we assume a cost functional $\mathcal{J}(\mathbf{x}(t))$ defined in terms of the state at time t. The impact of small errors/perturbations $\delta\mathbf{x}_i$ in the state vector at a snapshot time $t_i \le t$ on $\mathcal{J}$ may be estimated using the tangent linear model $\mathbf{M}(t_i, t)$ and its adjoint model $\mathbf{M}^*(t, t_i)$

$$\begin{aligned} \delta\mathcal{J} \approx \langle \mathcal{J}'(\mathbf{x}(t)), \delta\mathbf{x}(t)\rangle &= \langle \mathcal{J}'(\mathbf{x}(t)), \mathbf{M}(t_i, t)\delta\mathbf{x}(t_i)\rangle = \langle \mathbf{M}^*(t, t_i)\mathcal{J}'(\mathbf{x}(t)), \delta\mathbf{x}(t_i)\rangle \\ &= \langle \mathbf{A}^{-1}\mathbf{M}^*(t, t_i)\mathcal{J}'(\mathbf{x}(t)), \delta\mathbf{x}(t_i)\rangle_{\mathbf{A}} \qquad (20) \end{aligned}$$

The dual-weights ω_i to the snapshots are obtained as normalized values

$$\alpha_i = \|\mathbf{A}^{-1}\mathbf{M}^*(t, t_i)\mathcal{J}'(\mathbf{x}(t))\|_{\mathbf{A}}, \quad \omega_i = \frac{\alpha_i}{\sum_{j=1}^{n} \alpha_j}, \quad i = 1, 2, \ldots n \qquad (21)$$

and provide a measure of the relative impact of the state errors $\|\delta\mathbf{x}(t_i)\|_{\mathbf{A}}$ on the cost functional. A large value of ω_i indicate that state errors at t_i play an important role in the representation of the cost functional and an increased weight is assigned to the fit to snapshot data $\mathbf{x}^{(i)}$ in the reduced-basis optimization problem (6). The weights (21) are determined by the 4D-Var data assimilation cost functional (2) such that *information from the DAS is incorporated directly into the optimality criteria* that identifies the reduced-space basis functions. The DWPOD basis is thus adjusted to the 4D-Var optimization problem at hand.

From the implementation point of view, the evaluation of all dual-weights requires only one adjoint model integration to obtain the backward trajectories of the adjoint variables (influence functions) $\boldsymbol{\lambda}(\tau) = \mathbf{M}^*(t,\tau)\mathcal{J}'(\mathbf{x}(t))$, $t_0 \le \tau \le t$. Since in the 4D-Var data assimilation context the adjoint model is already available, little additional software development is required and the increased computational cost of implementing DWPOD over the standard POD method is modest. In the numerical experiments section we compare the performance of these two methods first as tools to provide a reduced order representation of a forecast output, then as tools to perform reduced order 4D-Var data assimilation.

4. Numerical Experiments

The numerical experiments are performed with a two-dimensional global shallow-water (SW) model using the explicit flux-form semi-Lagrangian (FFSL) scheme of Lin and Rood (1997). The finite volume FFSL scheme is of particular importance since it provides the horizontal discretization to the finite-volume dynamical core of NCAR Community Atmosphere Model (CAM) and NASA GEOS-5 data assimilation and forecasting system (Lin 2004). The adjoint model to the SW-FFSL scheme used in this study was developed in the work of Akella and Navon (2006) with the aid of TAMC software (Giering and Kaminski 1998).

Input data obtained from the ECMWF ERA-40 atmospheric data sets is used to specify the SW model state variables at the initial time: geopotential height h and the zonal and meridional wind velocities (u, v). We consider a $2.5° \times 2.5°$ resolution $(144 \times 72$ grid cells) such that the dimension of the discrete state vector $\mathbf{x} = (h, u, v)$ is $\sim 3 \times 10^4$. The time integration is performed with a constant time step $\Delta t = 450\,\mathrm{s}$ using a staggered 'CD-grid' system with the prognostic variables updated on the D-grid (Lin and Rood 1997). Point values of the model output are obtained by converting the winds from the D-grid to an unstaggered A-grid.

As a reference initial state $\mathbf{x}_0^{ref}$ we consider the 500mb ECMWF ERA-40 data valid for 06h UTC 15 March 2002. The configuration of the geopotential height at the initial time and a 24h SW model forecast is displayed in Fig. 1. On the discrete state space we consider a total energy norm

$$\|\mathbf{x}\|_{\mathbf{A}}^2 = \frac{1}{2}\left[\|u\|^2 + \|v\|^2 + \frac{g}{\overline{h}}\|h\|^2\right] \tag{22}$$

where $\|\cdot\|$ denotes the Euclidean norm, g is the gravitational constant and $\overline{h}$ is the mean height of the reference data at the initial time, such that $\mathbf{A}$ is a diagonal matrix with block constant entries $g/2\overline{h}, 1/2, 1/2$.

To generate the set of snapshots we introduced small random perturbations $\delta\mathbf{x}_0$ in the reference initial conditions and performed a full model integration starting with $\mathbf{x}_0^{ref} + \delta\mathbf{x}_0$. The state evolution $\mathbf{x}(t_i) = \mathcal{M}_{t_0,t_i}(\mathbf{x}_0^{ref} + \delta\mathbf{x}_0)$ was stored at each time step and used to define the ensemble data set $\mathbf{x}^{(i)} = \mathbf{x}(t_i)$, $i = 1, 2, \ldots n$. This data set is then used by the POD and DWPOD methods to identify an appropriate reduced order state subspace. In the standard POD approach, all the weights are set $\omega_i = 1/n$ and the POD basis of order $k < n$ is determined by the data only. In the DWPOD approach the weights are determined according to (21) such that the DWPOD reduced basis of order $k < n$ depends on the problem at hand.

a. Reduced-order representation of a forecast aspect

In the first set of experiments we consider the POD and DWPOD methods as tools to provide a reduced-order representation of a scalar aspect of the model forecast. The target functional is taken as a measure of the time integrated energy of the system for a 24h forecast initiated from $\mathbf{x}_0^{ref}$, $\mathcal{J}(\mathbf{x}) = \sum_{i=1}^{n} \|\mathbf{x}_i\|_{\mathbf{A}}^2$. For the 24h period, the ensemble data set includes 193 snapshots. The variance ("energy") $I(k)$ captured by the leading POD and DWPOD modes from the ensemble data as a function of the dimension k of the reduced space is displayed in Fig. 2, and selected numerical values are provided in Table 1. It is noticed that for the same dimension k of the reduced space a similar amount of variance is captured by the POD and DWPOD from the data set and weighted data set, respectively. In each case the dominant mode provides $\sim 78\%$ of the information, first ten modes $\sim 99\%$, and up to a small fraction, most of the information is contained in the leading 25 modes. However, the $k-$dimensional bases $\mathbf{\Psi}_{pod}$ and $\mathbf{\Psi}_{dwpod}$ identified by the POD and respectively DWPOD are distinct and in particular, higher modes of same rank may differ significantly from the POD basis to the DWPOD basis. In Fig. 3 isopleths of the POD and DWPOD modes are displayed using the energy norm to provide point values. A close resemblance is noticed between the dominant POD and DWPOD modes, whereas higher POD and DWPOD modes of same rank have a completely different structure.

In the POD approach the reduced order representation of the initial state is

$$\hat{\mathbf{x}}_0 = \overline{\mathbf{x}} + \mathbf{\Psi}_{pod}\mathbf{\Psi}_{pod}^T\mathbf{A}(\mathbf{x}_0^{ref} - \overline{\mathbf{x}})$$

with $\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}^{(i)}$ and in the DWPOD approach the initial state is represented as

$$\hat{\mathbf{x}}_0 = \overline{\mathbf{x}}^{\omega} + \mathbf{\Psi}_{dwpod}\mathbf{\Psi}_{dwpod}^{T}\mathbf{A}(\mathbf{x}_0^{ref} - \overline{\mathbf{x}}^{\omega})$$

with the weighted mean $\overline{\mathbf{x}}^{\omega}$ computed according to (3), (21). Since in practice the dimension k of the reduced space is determined by specifying a threshold value $0 < \gamma < 1$ such that $I(k) \geq \gamma$, it is of interest to analyze the error in the reduced-order representation of the target functional $|\mathcal{J}(\mathbf{x}) - \mathcal{J}(\hat{\mathbf{x}})|$ as the dimension of the reduced space varies. The numerical results using POD and DWPOD bases of dimension $k = 5, 10, 15, 20, 25$ are displayed in Fig. 4 and it is noticed that the DWPOD basis provided a significantly improved accuracy as compared to the POD basis. For example, projection of the initial conditions in the 10-dimensional DWPOD space provided qualitative results similar to the 15-dimensional POD space, whereas the representation in the 15-dimensional DWPOD space provided one order of magnitude gain in accuracy over the 15-dimensional POD space.

The reduced DWPOD space provided not only an improved representation of $\mathcal{J}(\hat{\mathbf{x}})$ but also a more accurate state forecast trajectory, as measured in the total energy norm. The forecast error $\|\mathbf{x}_i^{ref} - \hat{\mathbf{x}}_i\|_{\mathbf{A}}$ is displayed in Fig. 4 for each time step during the model integration. It is noticed the increased efficiency of the DWPOD basis that provided qualitative results similar to the POD basis while requiring fewer basis vectors. In particular, the errors in the 5-dimensional DWPOD space are close to the errors in the 10-dimensional POD space, the 10-dimensional DWPOD space provided forecast errors nearly identical to the errors in the 15-dimensional POD space, and the 15-dimensional DWPOD provided one order of magnitude gain in forecast accuracy over the 15-dimensional POD space.

As the dimension of the reduced space increases, each basis captures practically all of the information from the ensemble data. Little improvement in forecast accuracy may be achieved by increasing the DWPOD dimension from 20 to 25 and the state forecast error using a 25-dimensional DWPOD and a 25-dimensional POD basis provided overlapping graphs (visually indistinguishable) in Fig. 4.

While for both POD and DWPOD methods the state reduction from $\sim 3 \times 10^4$ to ~ 20 is remarkable, in practical applications it is important to obtain accurate reduced order representations using a small (the smallest) number of basis vectors. The enhanced efficiency of the DWPOD over POD in providing accurate results for small dimensional bases is thus a desirable property that may become increasingly significant as the dimension of the full model state increases. The dimension of the reduced space is also crucial in the efficiency of the reduced order 4D-Var data assimilation that aims to perform a minimal number of iterations to achieve a certain accuracy gain in the analysis.

b. Data assimilation experiments

To analyze the potential computational savings of the reduced order procedure, 4D-Var data assimilation experiments are setup in a twin experiments framework. As a background estimate $\mathbf{x}_b$ to the initial conditions we consider 500mb ECMWF ERA-40 data valid for 00h UTC 15 March 2002, six hours prior to the reference state $\mathbf{x}_0^{ref}$. The errors in the background term, averaged over the longitudinal coordinate, are displayed in Fig. 5. A data assimilation time interval $[t_0, t_0 + 24\text{h}]$ is considered with four data sets at 6h, 12h, 18h, and 24h provided by a model integration initiated from $\mathbf{x}_0^{ref}$. Two data assimilation experiments are setup: the first experiment, hereafter referred to as DAS-I, is a model inversion problem where data is provided for all discrete state components and no background term is included in the cost functional (2); in the second experiment, hereafter referred to as DAS-II, the background term is included in the cost and data is provided at every 4^{th} grid point on the longitudinal and latitudinal directions ($\sim 6\%$ of the state is "observed" every six hours). The distance to the background and observations is measured in the $\mathbf{A}$-norm that corresponds to diagonal matrices $\mathbf{B}$ and $\mathbf{R}$. To emphasize the fit to data, a weight factor of 0.01 is assigned to the distance to background in DAS-II.

Data assimilation experiments performed in the full model space resulted in a slow convergence for the large scale optimization problem (1). The minimization process using a high-performance limited memory quasi-Newton L-BFGS algorithm (Liu and Nocedal 1989) is displayed in Fig. 6 and it is noticed that a large number of iterations is required to approach the optimal point. A slower convergence rate is observed for DAS-I versus DAS-II due to the increased number of data constraints and the absence of the regularization provided by the background term.

c. Reduced-order 4D-Var data assimilation

Twin reduced-order 4D-Var data assimilation experiments were implemented using the POD and DWPOD bases respectively. It should be noticed that while the POD basis vectors remain unchanged for both DAS-I and DAS-II experiments, in the dual-weighted approach the reduced basis is adjusted to the optimization problem at hand. As shown in Fig. 7, the dual weights to the snapshot data are distinct from DAS-I to DAS-II and, as an illustrative example, in Fig. 8 isopleths of the DWPOD mode of rank 10 reveal a different configuration in DAS-I than in DAS-II.

The low dimensionality of the reduced spaced allowed the implementation of a full quasi-Newton BFGS algorithm to solve the optimization problem (13). The minimization process is displayed in Fig. 9 and it is noticed that only few iterations were required to reach the optimal point for each of the DAS-I and DAS-II experiments. For example, in DAS-II experiments 3 to 4 iterations are practically sufficient to reach a close vicinity of the optimal point and the computational savings of the reduced-order 4D-Var are thus significant. To facilitate the qualitative analysis, the total energy errors in the retrieved initial conditions, averaged over the longitudinal direction, are displayed in Fig. 10 and

Fig. 11 for DAS-I and respectively, DAS-II. By comparison to Fig. 5, the reduced 4D-Var data assimilation is able to provide analysis errors that are lowered by an order of magnitude as compared to the errors in the background estimate. For the 5- and 10-dimensional spaces, the analysis errors corresponding to the DWPOD space have much lower values as compared to the analysis errors for the POD space showing that the dual weighted approach to order reduction is of significant benefit. In particular, for the 10-dimensional spaces, in the DAS-I experiments one notices an error reduction by as much as a factor of three in the DWPOD space as compared to the POD space, whereas in the DAS-II experiments, the analysis error is reduced by as much as a factor of two in the DWPOD space as compared to the POD space. Increasing the dimension of the reduced space from 10 to 15 proves to be of little benefit to the analysis thus indicating that further improvements are constrained by the limited information provided by the snapshot data.

5. Conclusions and further research

The computational burden of the large-scale 4D-Var optimization problem may be significantly reduced by performing the optimization in a low order control space. An optimal order reduction approach to 4D-Var data assimilation must capture accurately the properties of the full dynamical model that are most relevant to a specific data assimilation system. To date, studies on reduced order 4D-Var have considered low order state subspaces based on the properties of the flow only, without properly taking into account the characteristics of the DAS. In this work an adjoint-model approach is proposed to directly incorporate information from the DAS into the optimality criteria that defines the reduced space basis. The dual weighted POD method is novel in reduced order 4D-Var data assimilation and relies on a weighted ensemble data mean and weighted snapshots with weights determined by the adjoint DAS. The numerical experiments presented with a finite volume global shallow water model indicate that the DWPOD approach may significantly improve the efficiency of the reduced basis as compared to the standard POD method. The DWPOD space was found to increase the accuracy in the representation of a forecast aspect by as much as an order of magnitude versus the POD space representation. In 4D-Var data assimilation twin experiments, optimization in the DWPOD space provided a reduction in the analysis errors by as much as a factor of two when compared to the POD-based optimization. The dual-weighted approach is thus cost-effective since the additional computational requirements to identify the DWPOD basis consist of a single adjoint model integration to evaluate the dual weights to the snapshot data.

This work represents a first step toward the development of an order reduction methodology that combines in an optimal fashion the model dynamics and the characteristics of the 4D-Var DAS. The mathematical formulation of the dual-weighted POD approach to model reduction is sound however, taking into account the simplicity of the shallow water model used in this study, the enhanced efficiency of the DWPOD method remains to be validated for numerical weather prediction and general circulation models in an operational data assimilation environment.

Strategies to implement an adaptive update of the reduced basis functions as the minimization algorithm advances toward the optimal point are at an incipient stage and this is an area where future research is much needed. Evaluation of the Hessian matrix of the 4D-Var cost functional in the reduced space is feasible using a second order adjoint model (Daescu and Navon 2006) and may be used to provide statistical information on the analysis errors. The reduced order 4D-Var approach is highly dependent on the quality of the snapshot data and the issue of generating a "good" set of snapshots is crucial for the reduced order procedure to be effective. The twin experiments setup used in this study facilitated the selection of the snapshot data in a close vicinity of the reference state trajectory. For practical applications, an ensemble of model forecasts may be used to generate snapshots taken from multiple state calculations with perturbations in the initial conditions that capture the main directions of variability of the model such as the bred vectors and singular vectors of the tangent linear model (Kalnay 2002). In this context the reduced order procedure will result in a hybrid approach that combines in an optimal fashion features of the ensemble and variational methods in data assimilation.

Acknowledgments. This research was supported by NASA Modeling, Analysis and Prediction Program under award NNG06GC67G.

References

Achatz, U., and G. Schmitz, 1997: On the closure problem in the reduction of complex atmospheric models by PIPs and EOFs: a comparison for the case of a two-layer model with zonally symmetric forcing. *J. Atmos. Sci.*, **54**, 2452–2474.

Achatz, U., and J.D. Opsteegh, 2003: Primitive-equation-based low-order models with seasonal cycle. Part I: Model construction. *J. Atmos. Sci.*, **60** (3), 465–477.

Akella, S., and I.M. Navon, 2006: A comparative study of the performance of high resolution advection schemes in the context of data assimilation. *Int. J. Numer. Meth. Fluids*, **51**, 719–748.

Anderson, J.L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.

Antoulas, A.C., 2005: *Approximation of Large-Scale Dynamical Systems.* Advances in Design and Control, SIAM, Philadelphia, PA, 481pp.

Aubry, N., W.-Y. Lian, and E.S. Titi, 1993: Preserving symmetries in the proper orthogonal decomposition. *SIAM J. Sci. Comput.*, **14**, 483–505.

Bennett, A.F., 1992: *Inverse Methods in Physical Oceanography.* Cambridge University Press, 346pp.

Blayo, E., J. Blum, and J. Verron, 1998: Assimilation variationnelle de donées en océanographie et réduction de la dimension de l'espace de contrôle. In *Equations aux Dérivées Partielles et Applications*, Elsevier, New-York, 199–219.

Bui-Thanh, T., Willcox, K., Ghattas, O., and B. van Bloemen Waanders, 2007: Goal-oriented, model-constrained optimization for reduction of large-scale systems. *J. Comput. Phys.*, accepted.

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

Cane, M.A., Kaplan, A., Miller, R.N., Tang, B., Hackert, E.C., and A.J. Busalacchi, 1996:

Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter. *J. Geophys. Res.*, **101**, 22599–22617.

Cao, Y., Zhu, J., Navon, I.M., and Z. Luo, 2006: A reduced order approach to four-dimensional variational data assimilation using proper orthogonal decomposition. *Int. J. Numer. Meth. Fluids*, in press.

Christensen, E.A., Brøns, M. and J.N. Sørensen, 2000: Evaluation of Proper Orthogonal Decomposition–Based Decomposition Techniques Applied to Parameter-Dependent Nonturbulent Flows. *SIAM J. Sci. Comput.* **21** (4), 1419–1434.

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteorol. Soc. Japan*, **75**, No. 1B, 257-288.

Courtier, P., Thepaud, J.N., and A. Hollingsworth, 1994: A strategy of operational implementation of 4D-Var using an incremental approach. *Q.J.R. Meteorol. Soc.* **120**, 1367–1388.

Crommelin, D.T., and A.J. Majda, 2004: Strategies for model reduction: comparing different optimal bases. *J. Atmos. Sci.*, **1**, 2206–2217.

Daescu, D.N., and I.M. Navon, 2006: Efficiency of a POD-based reduced second-order adjoint model in 4D-Var data assimilation. *Int. J. Numer. Meth. Fluids*, in press.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457pp.

Dee, D.P., 1990: Simplification of the Kalman filter for meteorological data assimilation. *Quart. J. Roy. Meteor. Soc.*, **117**, 365–384.

Durbiano, S., 2001: *Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données*. Ph.D. Thesis, Université Joseph Fourier, Grenoble, France.

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), 10143–10162.

Evensen, G., 2003: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367.

Giering, R., and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Trans. Math. Soft.* **24**(4), 437–474.

Golub G.H., and C.F. Van Loan, 1996: *Matrix Computations* (3rd edn). John Hopkins Univ. Press, Baltimore, Maryland, 642pp.

Graham, M.D., and I.G. Kevrekidis, 1996: Alternative approaches to the Karhunen-Loève decomposition for model reduction and data analysis. *Comput. Chem. Eng.*, **20** (5), 495–506.

Gunzburger, M.D., 2003: *Perspectives in Flow Control and Optimization.* Advances in Design and Control, SIAM, Philadelphia, PA, 261pp

Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11015–11021.

Holmes, P., Lumley, J.L., and G. Berkooz, 1996: *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge University Press.

Hoteit, I., and D.T. Pham, 2003: Evolution of the reduced state space and data assimilation schemes based on the Kalman filter. *J. Meteor. Soc. Japan*, **81**(1), 21–39.

Hoteit, I., and A. Köhl, 2006: Efficiency of reduced-order, time-dependent adjoint data assimilation approaches. *J. Ocean.*, **62**, 539–550.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376pp.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 364pp.

Kucukkaraca, E. and M. Fisher, 2006: Use of analysis ensembles in estimating flow-dependent background error variances. *ECMWF Technical Memorandum* **492**. European Centre for Medium-range Weather Forecasts (ECMWF), Reading, UK.

Kunisch, K., and S. Volkwein, 1999: Control of the Burgers equation by a reduced-order approach using Proper Orthogonal Decomposition. J. Optim. Theory Appl., **102** (2), 345-371.

Kunisch, K., and S. Volkwein, 2002: Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.*, **40**, 492–515.

Kwasniok, F., 2004: Empirical low-order models of barotropic flow. *J. Atmos. Sci.*, **61**, 235–245.

Langland, R.H., Gelaro, R., Rohaly, G.D., and M.A. Shapiro, 1999: Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP-17 and IOP-18. *Q.J.R. Meteorol. Soc.*, **125**, 3241–3270.

Le Dimet, F.X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* **38A**, 97–110.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293–2307.

Lin, S.-J., and R.B. Rood, 1997: An explicit flux-form semi-Lagrangian shallow-water model on the sphere. *Q.J.R. Meteorol. Soc.*, **123**, 2477–2498.

Liu, D.C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale minimization. *Math. Prog.,* **45**, 503–528.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Q.J.R. Meteorol. Soc.*, **112**, 1177–1194.

Lorenc, A.C., and Coauthors, 2000: The Met. Office global three-dimensional variational data assimilation scheme. *Q.J.R. Meteorol. Soc.*, **126**, 2991–3012.

Meyer, M., and H.G. Matthies, 2003: Efficient model reduction in non-linear dynamics using the Karhunen-Loève expansion and dual-weighted-residual methods. *Computational Mechanics* **31**, 179–191.

Molteni, F., Buizza, R., Palmer, T. N., and T. Petroliagis, 1996: The new ECMWF ensemble prediction system: methodology and validation. Q. J. R. Meteorol. Soc.,

122, 73–119.

Nerger, L., Hiller, W., and J. Schröter, 2005: A comparison of error subspace Kalman filters. *Tellus* **57A**, 715–735.

Parrish, D.F. and J.C. Derber, 1992: National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763.

Pham, D.T., Verron, J., and M.C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.* **16**, 323–340.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.F., and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational data assimilation. Part I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**, 1143–1170.

Rathinam M., and L. Petzold, 2003: A new look at proper orthogonal decomposition. *SIAM J. Numer Anal.* 2003; **41**(5):1893–1925.

Ravindran S.S., 2002: Adaptive reduced-order controllers for a thermal flow system using proper orthogonal decomposition. *SIAM J. Sci. Comp.* **23** (6), 1924–1942.

Robert, C., S. Durbiano, E. Blayo, J. Verron, J. Blum, and F.-X. LeDimet, 2005: A reduced order strategy for 4D-Var data assimilation. *J. Mar. Syst.*, **57**, 70–82.

Robert, C., E. Blayo, and J. Verron, 2006: Reduced-order 4D-Var: a preconditioner for the incremental 4D-Var data assimilation method. *Geophys. Res. Let.* **33** L18609, doi:10.1029/2006GL026555

Selten, F.M., 1995: An efficient description of the dynamics of barotropic flow. *J. Atmos. Sci.*, **52**, 915–936.

Selten, F.M., 1997: Baroclinic empirical orthogonal functions functions as basis functions in an atmospheric model. *J. Atmos. Sci.*, **54**, 2099–2114.

Sirovich, L., 1987: Turbulence and the dynamics of coherent structures, Part I-III. *Quart. Appl. Math.* **45**, 561–590.

Todling R., and S.E. Cohn, 1994: Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Mon. Wea. Rev.*, **122**, 2530–2557.

Trémolet Y., 2004: Diagnostics of linear and incremental approximations in 4D-Var. *Q. J. R. Meteorol. Soc.* **130**, 2233–2251.

Trémolet Y., 2005: Incremental 4D-Var Convergence Study. *ECMWF Technical Memorandum* **469**. European Centre for Medium-range Weather Forecasts (ECMWF), Reading, UK.

van Doren, J.F.M., Markovinović, R., and J.D. Jansen, 2006: Reduced-order optimal control of water flooding using proper orthogonal decomposition. *Computat. Geosci.* **10**, 137–158.

Willcox, K., and J. Peraire, 2002: Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*, **40** (11), 2323–2330.

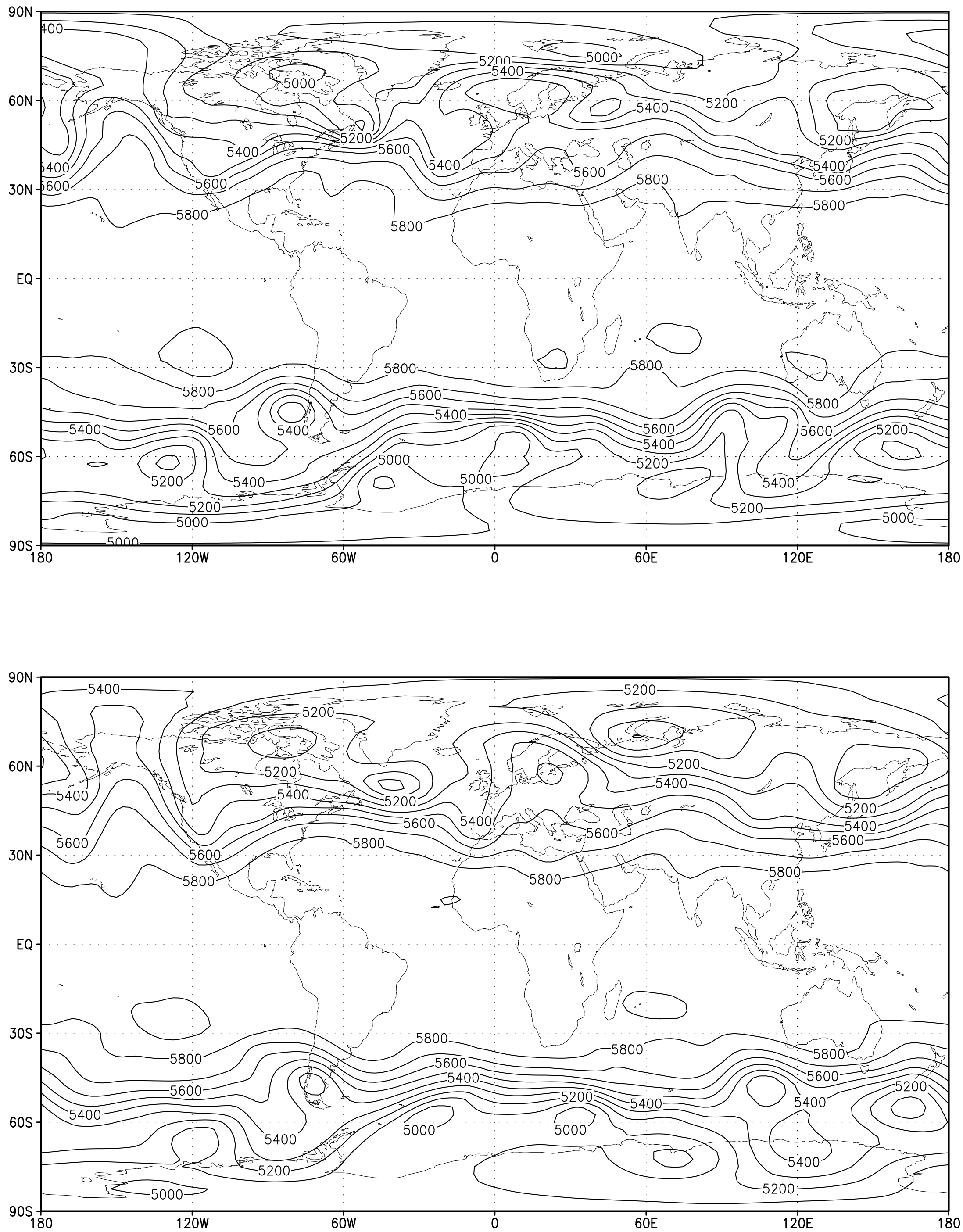


Figure 1: Isopleths of the geopotential height (m) for the reference run: top figure - configuration at the initial time specified from ECMWF ERA-40 data sets; bottom figure - the 24h forecast of the shallow water model.

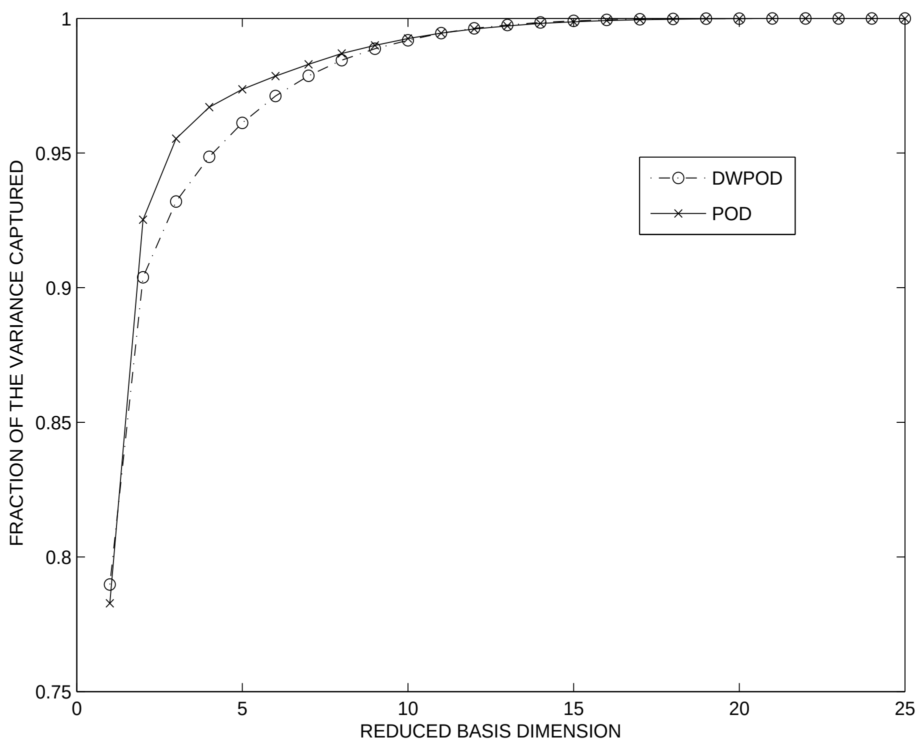


Figure 2: The fraction of the variance captured by the POD and DWPOD modes from the snapshot data as a function of the dimension of the reduced space.

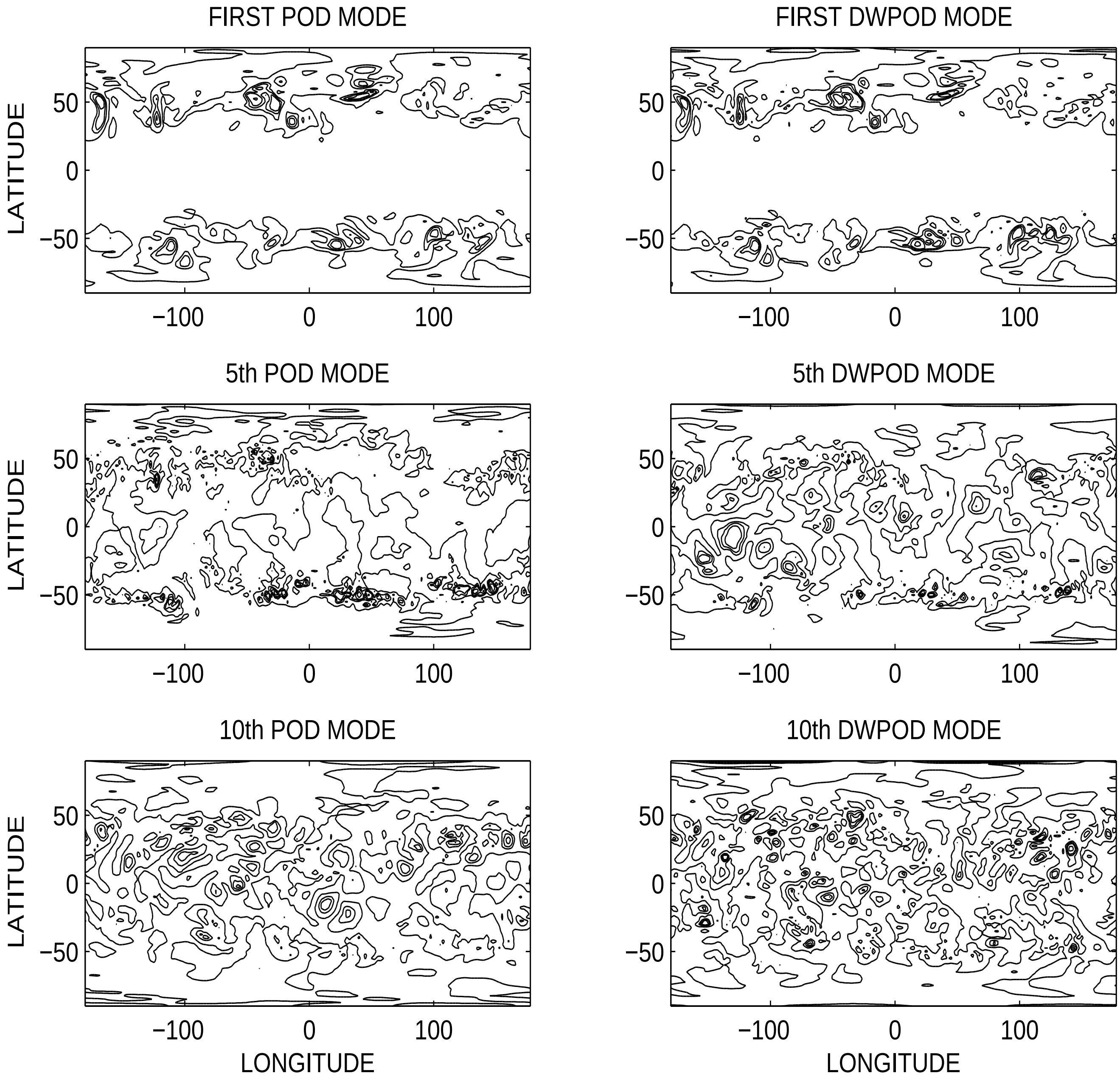
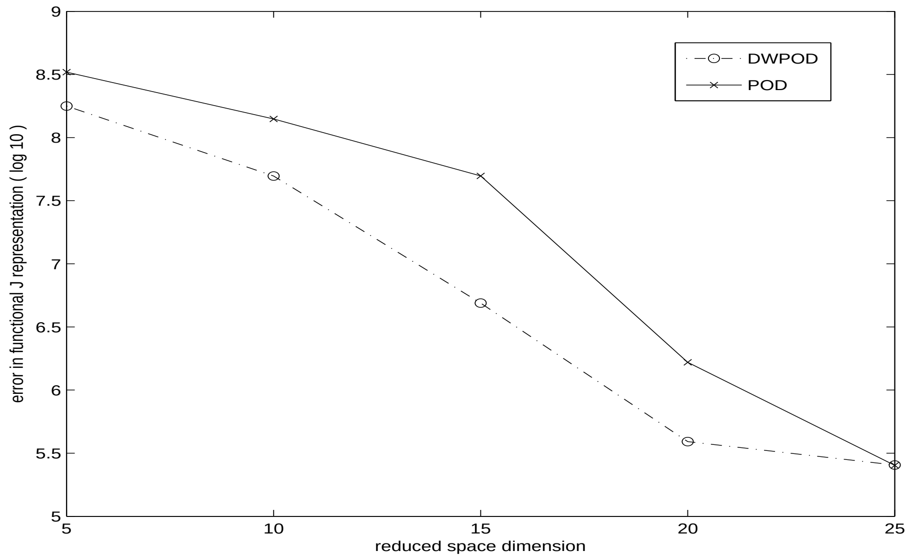


Figure 3: Isopleths of the POD and DWPOD modes of rank 1, 5, and 10. A total energy norm is used to provide point values.



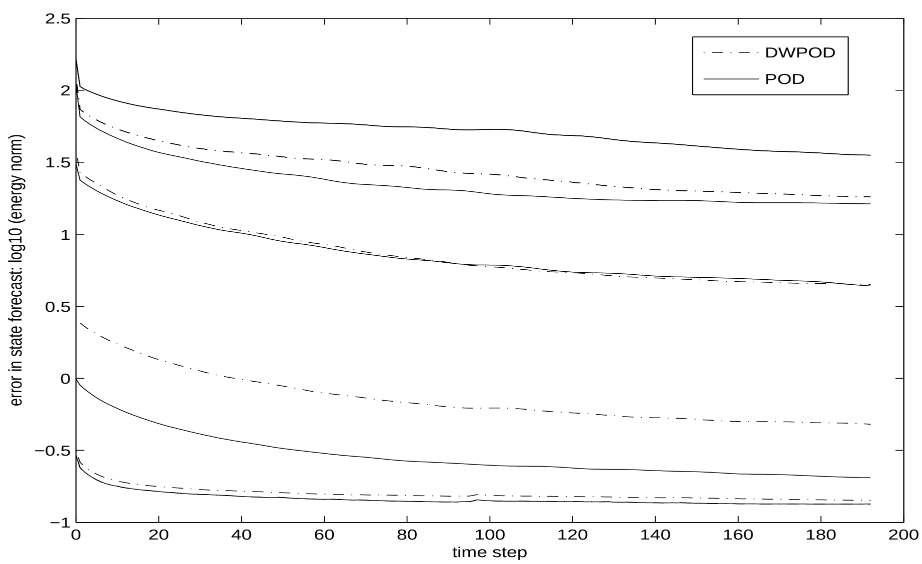


Figure 4: Comparative results for the reduced-order POD and DWPOD forecasts as the dimension of the reduced space varies, $k = 5, 10, 15, 20, 25$. Top figure: error (log 10) in the reduced-order representation of the time integrated total energy of the system. Bottom figure: total energy error (log 10) $\|\mathbf{x}_i^{ref} - \hat{\mathbf{x}}_i\|_{\mathbf{A}}$ of the reduced-order representation of the forecast at each time step in the interval 0-24h. The errors decrease as the dimension of the reduced space increases.

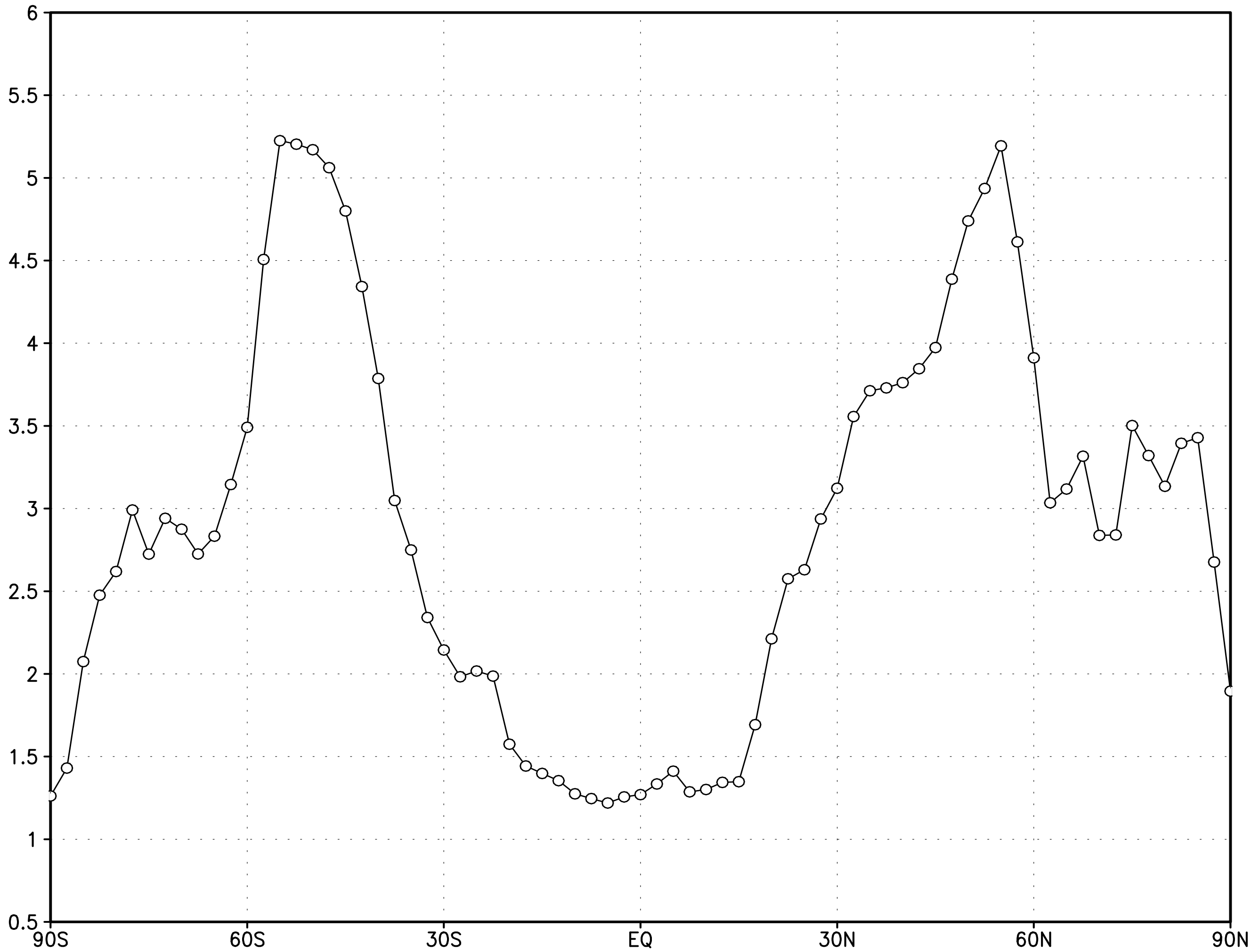


Figure 5: Zonal averaged errors in the background estimate to the initial conditions.

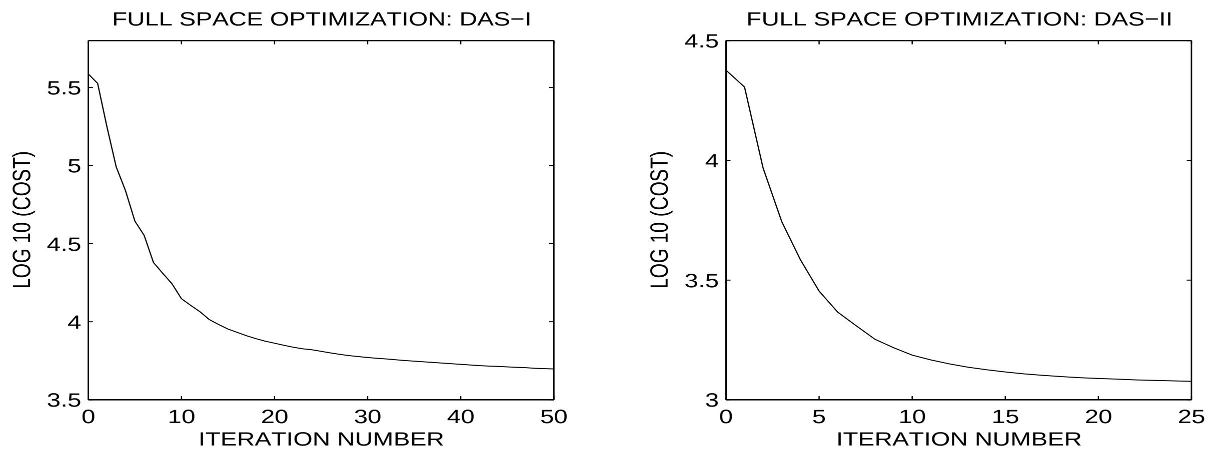


Figure 6: The iterative minimization process in the full state space for DAS-I (left) and DAS-II (right).

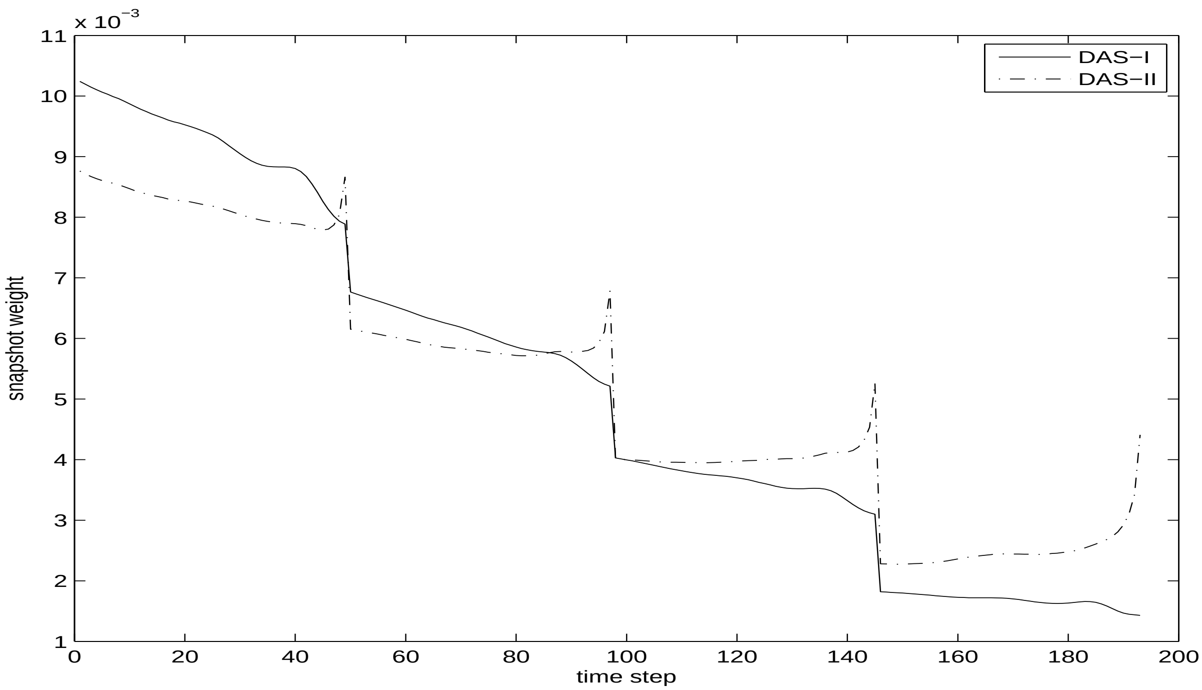
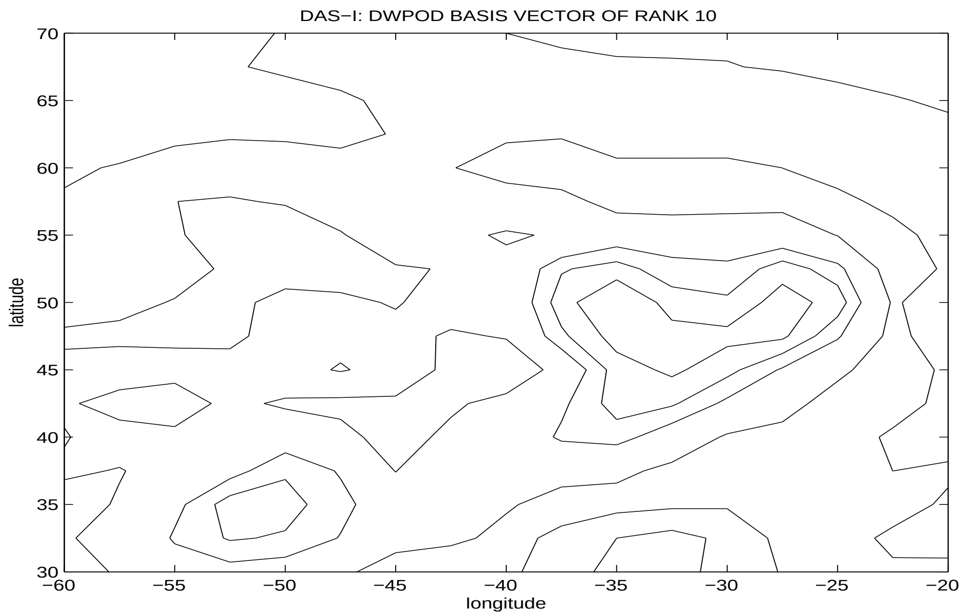


Figure 7: The dual weights to the snapshot data determined by the adjoint model in DAS-I and in DAS-II



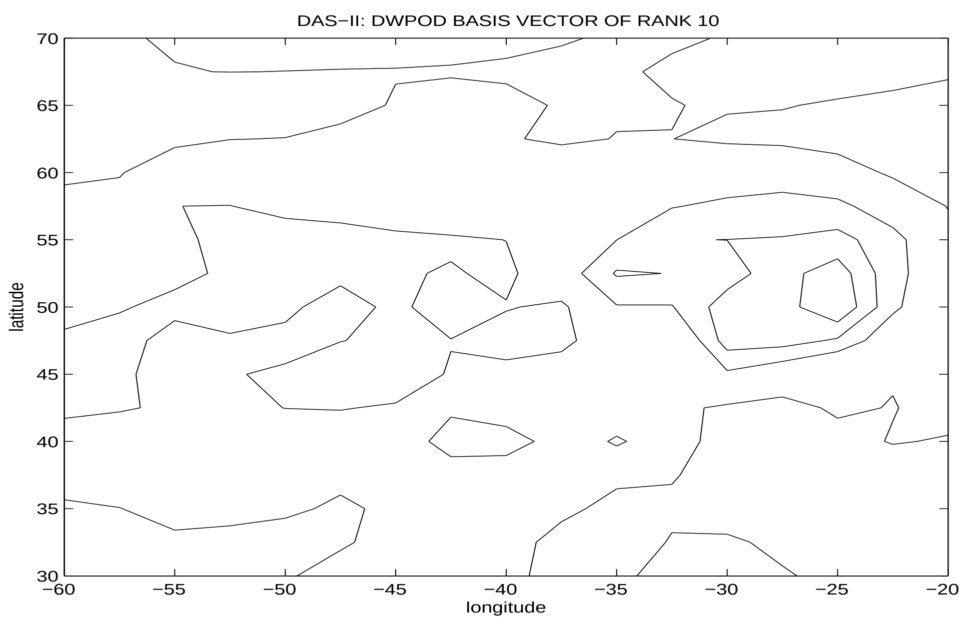
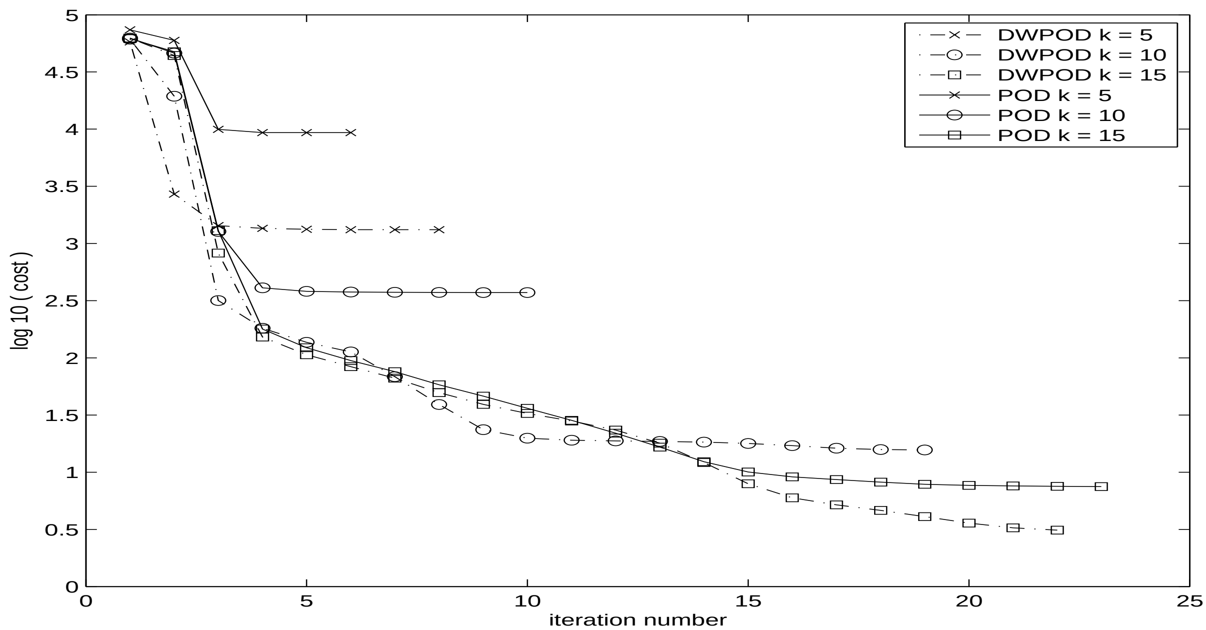


Figure 8: Isopleths of the 10^{th} mode in the DWPOD basis for DAS-I (top figure) and DAS-II (bottom figure). A distinct configuration it is noticed since the DWPOD basis is adjusted to the optimization problem at hand.



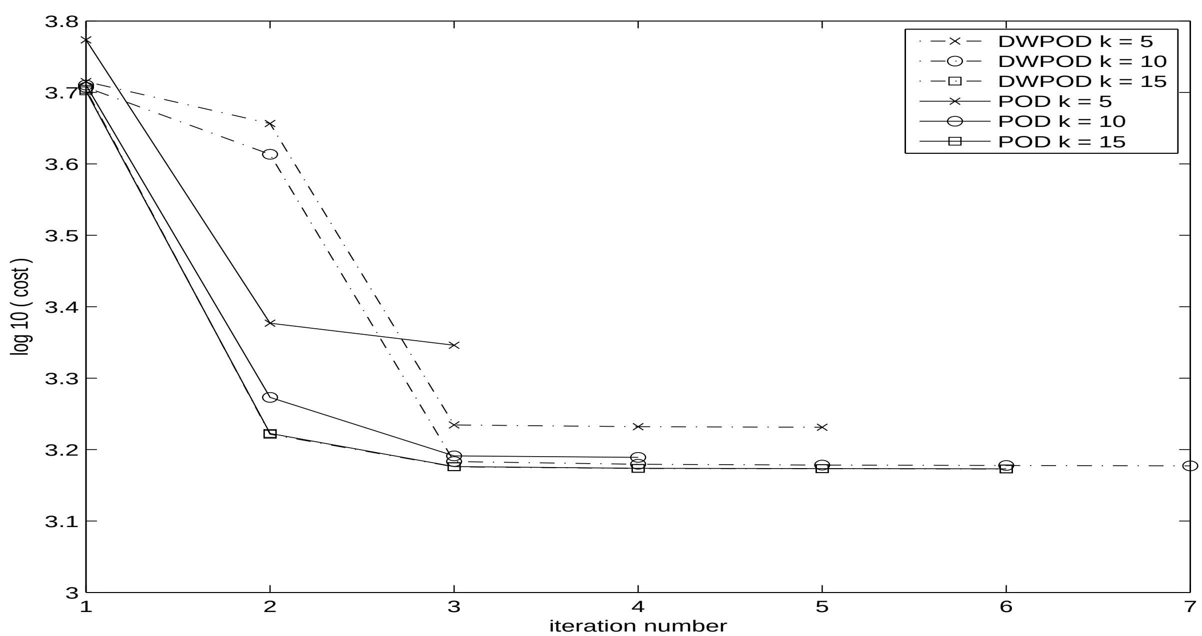


Figure 9: The iterative minimization process in the reduced space for the POD and DWPOD spaces of dimension 5, 10, and 15. Top figure - optimization without background term and dense observations, corresponding to DAS-I; bottom figure - optimization with background term and sparse observations, corresponding to DAS-II

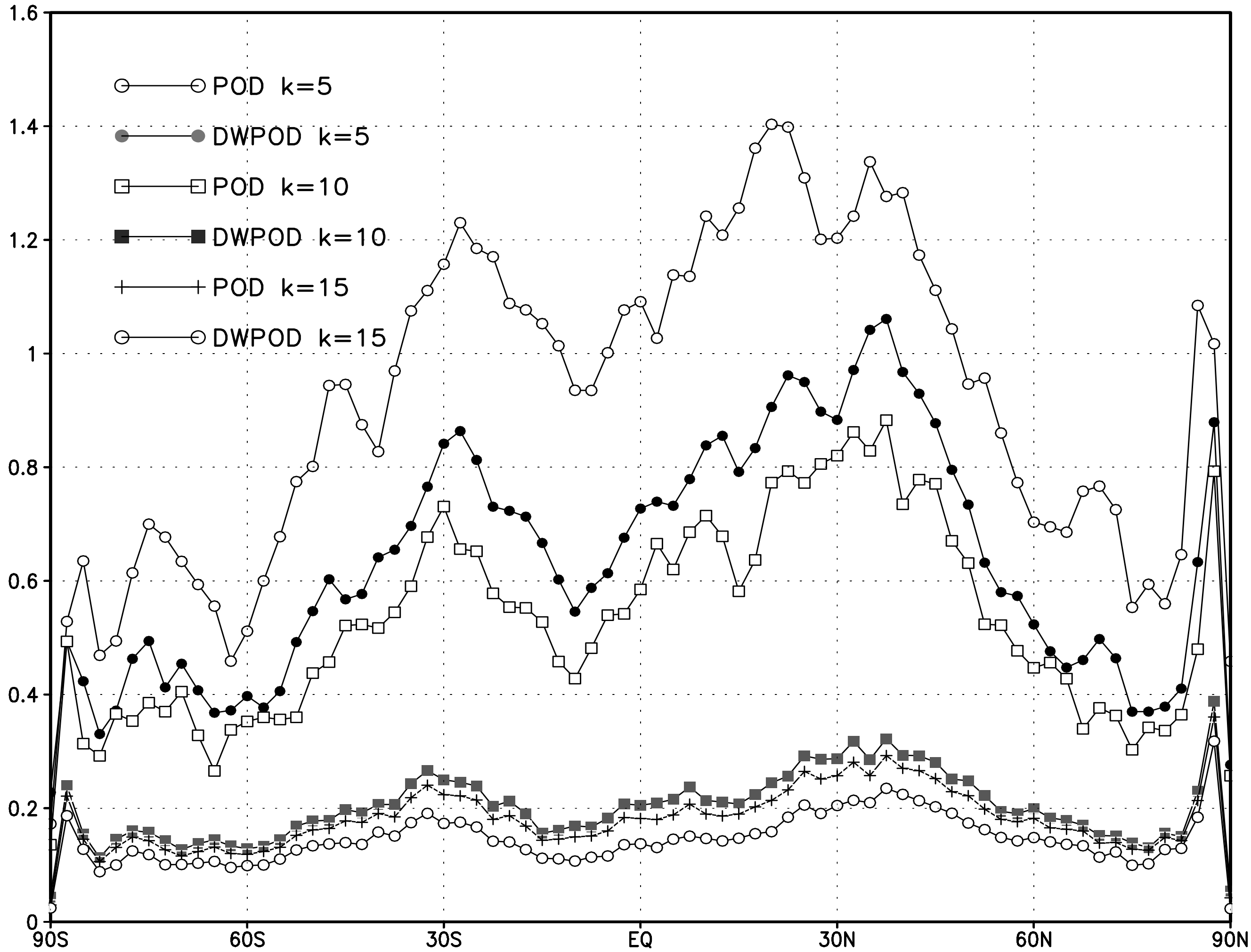


Figure 10: Zonal averaged errors in the analysis provided by the reduced order 4D-Var data assimilation. Results for the DAS-I experiments with POD and DWPOD spaces of dimension 5, 10, and 15.

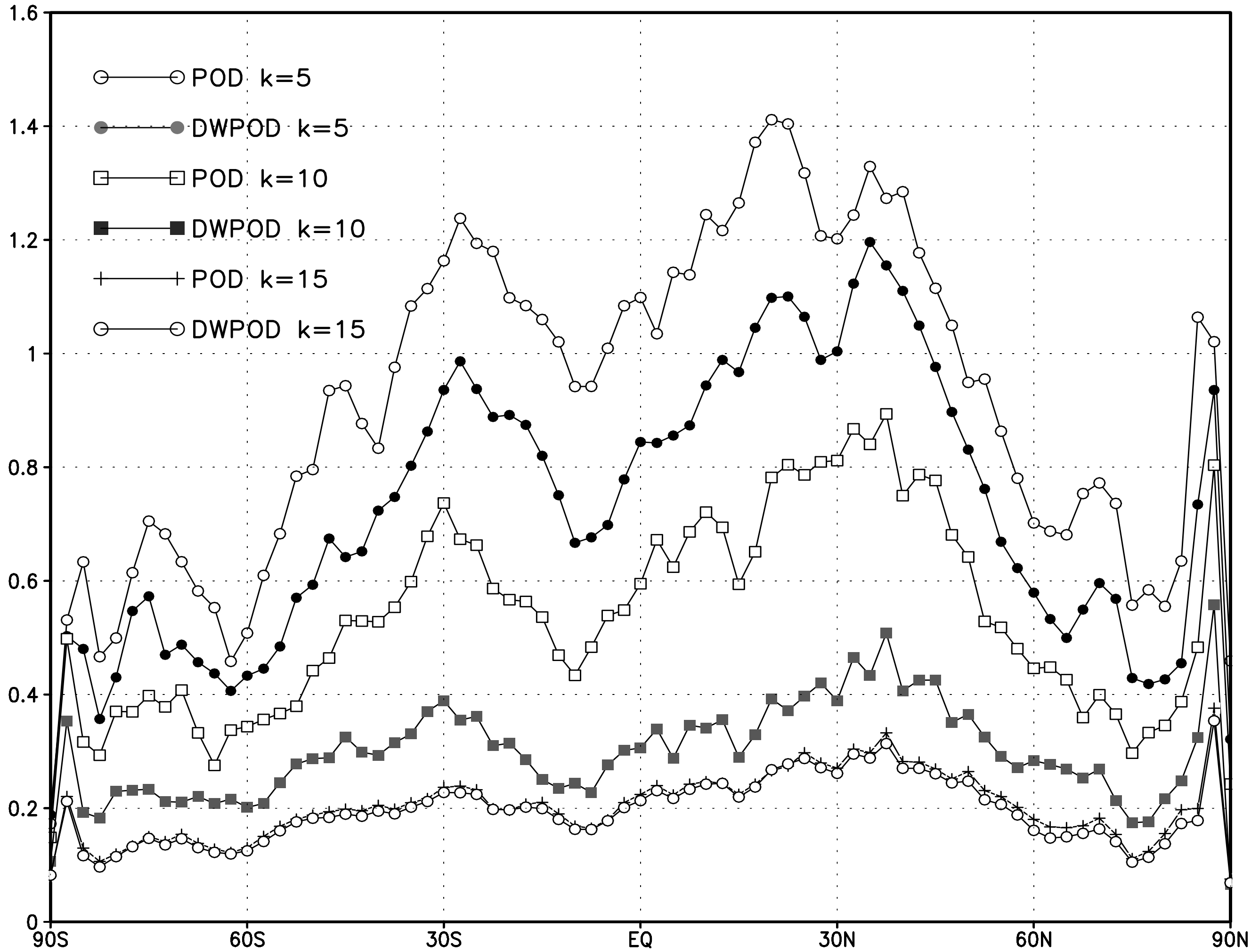


Figure 11: Zonal averaged errors in the analysis provided by the reduced order 4D-Var data assimilation. Results for the DAS-II experiments with POD and DWPOD spaces of dimension 5, 10, and 15.

Table 1: Fraction of the variance captured by the leading POD and DWPOD vectors

Basis Dimension	1	5	10	15	20	25
POD	0.7827	0.9736	0.9924	0.9987	0.9998	0.9999
DWPOD	0.7897	0.9612	0.9918	0.9990	0.9999	0.9999