INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2007; **53**:1571–1583
Published online 10 October 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1365

A reduced-order approach to four-dimensional variational data assimilation using proper orthogonal decomposition

Yanhua Cao[1], Jiang Zhu[1,*,†], I. M. Navon[2] and Zhendong Luo[3]

[1]*Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China*
[2]*School of Computational Science and Department of Mathematics, Florida State University, Tallahassee, FL 32306-4120, U.S.A.*
[3]*College of Fundamental Sciences, North China University of Technology, Beijing 100041, China*

SUMMARY

Four-dimensional variational data assimilation (4DVAR) is a powerful tool for data assimilation in meteorology and oceanography. However, a major hurdle in use of 4DVAR for realistic general circulation models is the dimension of the control space (generally equal to the size of the model state variable and typically of order 10^7–10^8) and the high computational cost in computing the cost function and its gradient that require integration model and its adjoint model.

In this paper, we propose a 4DVAR approach based on proper orthogonal decomposition (POD). POD is an efficient way to carry out reduced order modelling by identifying the few most energetic modes in a sequence of snapshots from a time-dependent system, and providing a means of obtaining a low-dimensional description of the system's dynamics. The POD-based 4DVAR not only reduces the dimension of control space, but also reduces the size of dynamical model, both in dramatic ways. The novelty of our approach also consists in the inclusion of adaptability, applied when in the process of iterative control the new control variables depart significantly from the ones on which the POD model was based upon. In addition, these approaches also allow to conveniently constructing the adjoint model.

The proposed POD-based 4DVAR methods are tested and demonstrated using a reduced gravity wave ocean model in Pacific domain in the context of identical twin data assimilation experiments. A comparison with data assimilation experiments in the full model space shows that with an appropriate selection of the basis functions the optimization in the POD space is able to provide accurate results at a reduced computational cost. The POD-based 4DVAR methods have the potential to approximate the performance of full order 4DVAR with less than 1/100 computer time of the full order 4DVAR. The HFTN (Hessian-free truncated-Newton)algorithm benefits most from the order reduction (see (*Int. J. Numer. Meth. Fluids*, in press)) since computational savings are achieved both in the outer and inner iterations of this method.

Received 22 December 2005; Revised 5 July 2006; Accepted 6 July 2006

*Correspondence to: Jiang Zhu, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China.
†E-mail: jzhu@mail.iap.ac.cn

Contract/grant sponsor: Natural Science Foundation of China; contract/grant numbers: 40437017, 40225015
Contract/grant sponsor: NSF; contract/grant number: ATM-9731472

KEY WORDS: proper orthogonal decomposition; variational data assimilation; reduced order; ocean modelling

1. INTRODUCTION

Four-dimensional variational data assimilation (4DVAR) is a powerful tool to obtain dynamically consistent atmospheric and oceanic flows that optimally fit observations. Since its introduction (see, Reference [1]), 4DVAR has been applied to numerical weather prediction (NWP) (e.g. Reference [2]), ocean general circulation estimation (e.g. Reference [3]) and atmosphere–ocean–land coupled modelling [3]. However, a major hurdle in use of 4DVAR for realistic general circulation models is the dimension of the control space, generally equal to the size of the model state variable and typically of order 10^7–10^8. Current ways to conquer these difficulties in using 4DVAR consist mainly of the incremental method [2, 4], checkpointing method (e.g. Reference [5]) and parallelization. The incremental method is characterized by the fact that the dimension of the control space remains very large (see, References [6, 7]). Memory storage requirements impose a severe limitation on the size of assimilation studies, even on the largest computers. Checkpointing strategies [5, 8] have been developed to address the explosive growth in both on-line computer memory and remote storage requirements for computing the gradient by the forward/adjoint technique that characterizes large-scale assimilation studies. Parallelization using message-passing interface (MPI) is currently used to implement 4DVAR (ECMWF, NCEP and WRF). But these three methods can only alleviate the difficulties in realistic applications to some extent. It was shown that the trade-off between the storage requirements and the computational time might be optimized such that the storage and computational time grow only logarithmically [9].

In order to reduce the computational cost of 4DVAR data assimilation we can consider carrying out the minimization of the cost functional in a space whose dimension is much smaller than that of the original one. A way to drastically decrease the dimension of the control space without significantly compromising the quality of the final solution but sizably decreasing the cost in memory and CPU time of 4DVAR motivates us to choose to project the control variable on a basis of characteristic vectors capturing most of the energy and the main directions of variability of the of the model, i.e. SVD, EOF, Lyapunov or bred vectors. One would then attempt to control the vector of initial conditions in the reduced space model.

Up to now, most efforts of model reduction have centred on Kalman and extended Kalman filter/smoother data assimilation techniques [10–16]. In particular, Cane *et al.* [12] employed a reduced order method in which the state space is reduced through the projection onto a linear subspace spanned by a small set of basis functions, using an empirical orthogonal function (EOF) analysis. This filter is referred to as the reduced order extended Kalman (ROEK) filter. Similar works were also done by Kaplan *et al.* [17–19] and Canizares *et al.* [20] for analysis long-term ocean observations. Lermusiaux and Robinson [21] and Lermusiaux [22] also discussed the space reduction in sequential data assimilation by projection on low-dimensional error space. The time integration of the error space is based on Monte Carlo ensemble forecasts [23]. See also [24, 25].

Some initial efforts aiming at the reduction of the dimension of the control variable—referred to as reduced order strategy for 4DVAR ocean data assimilation were put forward initially by Blayo *et al.* [26], Durbiano [27] and more recently by Robert *et al.* [28]. They used a low-dimensional space based on the first few EOF's or empirical orthogonal functions, which can be computed

DOI: 10.1002/fld

from a sampling of the model trajectory. Hoteit and Dinh-Tuan Pham [29] used the reduced order space approach for part of the 4DVAR assimilation then switched to the full model in a manner similar to earlier work of Peterson [30].

The proper orthogonal decomposition (POD) is an efficient way to reduced order modelling by identifying the few most energetic modes in a time-dependent system, thus providing a means of obtaining a low-dimensional description of the system's dynamics. It was successfully used in a variety of fields including signal analysis and pattern recognition (see, Reference [31]), fluid dynamics and coherent structures (see, References [32, 33]) and more recently in control theory (see, References [34–38]) and inverse problems (see, Reference [37]). Moreover, Atwell *et al.* [38] had successfully utilized POD to compute reduced order controllers. For a comprehensive description of POD theory and state of the art research, see References [39–41].

In this paper, we apply POD to 4DVAR, our first aim being to explore the feasibility of significant reduction in the computational cost of 4DVAR. Our basic approach will build on the POD-based adaptive control of Arian *et al.* [35]. The main difference in method between this work and Blayo *et al.* [26], Hoteit and Pham [29] and Robert *et al.* [28] is that we not only work in reduced order space for control variables, but also constructed and used the reduced order model as constraint in 4DVAR and the adjoint model of the reduced order model for calculation of the gradient of the cost function while they used the full order model as constraint and its adjoint model for calculation of the gradient of the cost function. The novelty of our approach resides also in the inclusion of adaptivity, applied when in the process of iterative control the new initial condition departs significantly from the one on which the POD model was based upon. Though the SEEK filter used a similar idea to update reduced order basis during data assimilation, in our approach the adaptivity criterion is based on the norm of the gradient of the cost function in the context of 4DVAR. See also work of [42].

The paper is arranged as follows. A brief review of the numerical model and POD used in this study is given in Section 2. A 4DVAR formulation based on POD and an adaptive POD 4DVAR are proposed in Section 3. Section 4 contains results from identical twin data assimilation experiments using 4DVAR, POD 4DVAR and adaptive POD 4DVAR, respectively. Finally, Section 5 provides main conclusions and discussions of some related issues of this study.

2. NUMERICAL MODEL AND POD

2.1. Model of upper tropic Pacific

The numerical model used in this paper is a reduced-gravity model. The equations for the depth-averaged currents are

$$\begin{aligned}
\frac{\partial u}{\partial t} - fv &= -g'\frac{\partial h}{\partial x} + \frac{\tau^x}{\rho_0 H} + A\nabla^2 u - \alpha u \\
\frac{\partial v}{\partial t} + fu &= -g'\frac{\partial h}{\partial y} + \frac{\tau^y}{\rho_0 H} + A\nabla^2 v - \alpha v \qquad (1)\\
\frac{\partial h}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) &= 0
\end{aligned}$$

DOI: 10.1002/fld

Table I. The values of the model parameters used in the model.

Parameter	Value	Remarks
g'	3.7×10^{-2}	Reduced gravity
C_D	1.5×10^{-3}	Wind stress drag coefficient
H	1.5×10^{2} m	Mean depth of upper layer
ρ_a	$1.2\,\mathrm{kg\,m^{-3}}$	Density of air
ρ_0	$1.025 \times 10^{3}\,\mathrm{kg\,m^{-3}}$	Density of seawater
A	$7.5 \times 10^{2}\,\mathrm{m^2\,s^{-1}}$	Coefficient of horizontal viscosity
α	2.5×10^{-5}	Coefficient of bottom friction

where (u, v) are the horizontal velocity components of the depth-averaged currents; h the total layer thickness; f the Coriolis force; H the mean depth of the layer; ρ_0 the density of water; and A the horizontal eddy viscosity coefficient and α the friction coefficient; (τ^x, τ^y) is the wind stress.

In this study, we applied the model to the tropic Pacific Ocean domain (29°S–29°N, 120°E–70°W). This chosen model domain allows all possible equatorially trapped waves, which can be excited, for example, by the applied wind forcing [43]. The model is discretized on the Arakawa C-grid, and all the model boundaries are closed. The no-normal flow and no-slip conditions are applied at these solid boundaries. The time integration uses a leapfrog scheme, with a forward scheme applied every 10th time step to eliminate the computational mode. We choose the spatial interval for the dynamical model to be $\Delta x = \Delta y = 0.5°$ and the time step to be $\Delta t = 100$ s. This temporal–spatial resolution will allow to resolve all possible waves and to make the model integration numerically stable. The model is driven by the Florida State University (FSU) climatology monthly mean winds [44]. The data are projected into each time step by a linear interpolation and into each grid point by a bilinear interpolation. The values of numerical parameters used in the model integration are listed in Table I. It takes about 20 years for the model to reach a periodic constant seasonal cycle; at that time, the main seasonal variability of dynamical fields has been successfully captured. The currents and the upper layer thickness of the 21st year are saved for POD reduced model and data assimilation experiments as described below.

2.2. POD reduced model

POD has been shown an efficient way to reduced order modelling by identifying the few most energetic modes in a time-dependent system, thus providing a means of obtaining a low-dimensional description of the system's dynamics. For successful POD 4DVAR, it is crucial to construct an accurate POD reduced model. The construction of the above reduced-gravity model (referred as full model thereafter) and the accuracy of POD reduced model had been discussed in detail (see, Reference [45]). Here we only briefly review this procedure.

For a complex temporal–spatial flow $U(t, x)$, we denoted by $U^1, \ldots, U^n$ a set adequately chosen in a time interval $[0, T_N]$, that is $U_i = U(t_i, x)$. Define the mean:

$$\bar{U} = \frac{1}{n} \sum_{i=1}^{n} U^i \tag{2}$$

DOI: 10.1002/fld

We expand $U(t, x)$ as

$$U^{\mathrm{POD}}(t, x) = \bar{U}(x) + \sum_{i=1}^{M} c_i(t)\Phi_i(x) \tag{3}$$

where the POD basis vector $\Phi_i(x)$ and M are judiciously chosen to capture the dynamics of the flow as follows.

1. Compute the mean $\bar{u} = \frac{1}{n}\sum_{i=1}^{n} u^i$;
2. Build the correlation matrix $K = k_{ij}$, $k_{ij} = \int_\Omega (u^i - \bar{u})(u^j - \bar{u})\,\mathrm{d}x$;
3. Compute eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$ and the corresponding orthogonal eigenvectors $v_1, v_2, \ldots, v_n$ of K;
4. Set $\Phi_i := \sum_{j=1}^{n} v_j^i (u^j - \bar{u})$.

One can define a relative information content to choose a low-dimensional basis of size $M << n$ by neglecting modes corresponding to the small eigenvalues. We define

$$I(k) = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{4}$$

and choose M such that

$$M = \mathrm{argmin}\{I(m) : I(m) \geqslant \gamma\}$$

where $0 \leqslant \gamma \leqslant 1$ is the percentage of total information captured by the reduced space $D^M = \mathrm{span}\{\Phi_1, \Phi_2, \ldots, \Phi_M\}$. The tolerance γ must be chosen to be near the unity in order to capture most of the energy of the snapshot basis. The reduced order model is then obtained by expanding the solution as in (3).

For an atmospheric or oceanic flow $U(t, x)$, it is usually governed by a dynamic model

$$\begin{aligned} \frac{\mathrm{d}U}{\mathrm{d}t} &= F(t, U) \\ U(0, x) &= U_0(x) \end{aligned} \tag{5}$$

To obtain a reduced model of (5), we can first solve (5) for a set of snapshots and follow above procedures, then use a Galerkin projection of the model equations onto the space spanned by the POD basis elements (replacing U in (5) by (3), then multiplying Φ_i and integrating over spatial domain):

$$\begin{aligned} \frac{\mathrm{d}c_i}{\mathrm{d}t} &= \left\langle F\left(t, \bar{U} + \sum_{i=1}^{M} c_i \Phi_i\right), \Phi_i \right\rangle \\ c_i(t = 0) &= c_i(0) \end{aligned} \tag{6}$$

Equation (6) defines a reduced model of (5). In the following sections we will discuss applying this model reduction to 4DVAR in which the forward model and the adjoint model for computing the cost function and its gradient are the reduced model and its corresponding adjoint, respectively.

DOI: 10.1002/fld

3. POD 4DVAR

3.1. POD 4DVAR

At the analysis time interval $[0, T_N]$, strong constraint 4DVAR looks for an optimal solution of (5) to minimize the cost function

$$J(U_0) = (U_0 - U_\mathrm{b})^\mathrm{T} B^{-1}(U_0 - U_\mathrm{b}) + (HU - y^\mathrm{o})^\mathrm{T} O^{-1}(HU - y^\mathrm{o}) \tag{7}$$

In POD 4DVAR, we look for an optimal solution of (5) to minimize the cost function

$$J(c_1(0), \ldots, c_M(0)) = (U_0^\mathrm{POD} - U_\mathrm{b}) B^{-1}(U_0^\mathrm{POD} - U_\mathrm{b}) + (HU^\mathrm{POD} - y^\mathrm{o}) O^{-1}(HU^\mathrm{POD} - y^\mathrm{o}) \tag{8}$$

where U_0^POD is the control vector, H is an observation operator, B is the background error covariance matrix and O is the observation error covariance matrix.

In (8),

$$U_0^\mathrm{POD}(x) = U_0^\mathrm{POD}(0, x) = \bar{U}(x) + \sum_{i=1}^{M} c_i(0)\Phi_i(x)$$

$$U^\mathrm{POD}(x) = U^\mathrm{POD}(t, x) = \bar{U}(x) + \sum_{i=1}^{M} c_i(t)\Phi_i(x)$$

In POD 4DVAR, the control variables are $c_1(0), \ldots, c_M(0)$. As shown later, the dimension of the POD reduced space could be much smaller than that the original space. In addition, the forward model is the reduced model (6) which can be very efficiently solved. The adjoint model of (6) is used to calculate the gradient of the cost function (8) and that will significantly reduce both the computational cost and the coding effort.

It is important to notice that the initial value of the cost function in the full model space is distinct from the initial value of the cost function in the POD space. Starting with the initial guess given by the background estimate $U_0 = U_\mathrm{b}$, the value of the cost in the full model space is $J(U_\mathrm{b})$. The corresponding initial guess in the reduced space is obtained by projecting the background on the POD space $\eta_0 = X_k^\mathrm{T}(U_\mathrm{b} - \bar{U})$, thus providing an initial cost value $\hat{J}(\eta_0) = J(\bar{U} + X_k\eta_0)$. This situation arises also in practical applications where the background is the best prior estimate to the initial conditions and the projection/retrieval operations on/from the POD space may deteriorate the quality of the first guess estimate. For comparison with the POD procedure, the relative reduction in cost as a metric to assess the efficiency may not be well suited. To account for both computational and qualitative aspects of the reduced/full order optimization, it is of interest to assess the CPU time required to attain the same value of the cost functional. To facilitate the analysis, all the results are shown in terms of the value of the cost function versus the number of iterations and versus CPU time.

To establish the POD model in POD 4DVAR, we need first to obtain a set of snapshots, which is taken from the background trajectory, or integrate original model (5) with background initial conditions.

3.2. Adaptive POD 4DVAR

Since the POD model is based on the solution of the original model for a specified initial condition, it might be a poor model when the new initial condition is significantly different from the one on

DOI: 10.1002/fld

which the POD model is based upon. Therefore, we propose an adaptive POD 4DVAR procedure as follows:

(i) establish POD model using background initial conditions and then perform optimization iterations to approximate the optimal solution of the cost function (8);
(ii) if after a preset number of iterations, the cost function cannot be reduced significantly as measured by a preset convergence criterion, we generate a new set of snapshots by integrating the original model using the newest initial conditions;
(iii) establish a new POD model using a new set of snapshots and continue the optimization iteration; and
(iv) check if the optimality conditions are reached, if yes, then stop; if no, go to step (ii).

4. POD 4DVAR EXPERIMENTS

4.1. Assimilation experiments

In this section, we present results of identical twin data assimilation experiments to examine the performances of POD 4DVAR and adaptive POD 4DVAR by comparing them with the full 4DVAR. The 'true' seasonal cycle of tropic Pacific is generated by forcing the model using FSU climatology monthly wind fields as described in the previous section. From the 12-month's truth, we generate a set of observations of h that have uncorrelated Gaussian observational errors of zero mean and $0.06m$ of variances. Observations are sampled at the one by one degree resolution and a 10-day temporal resolution. This observation network and error characteristics imitate the Topex/POSEIDON/JASON-1 satellite sea surface height observations.

The control variables in these experiments are the initial conditions only. The cost function consists of the observation and the background terms. The observation error covariance matrix is taken to be a diagonal one with 0.06^2 as diagonal elements. The background field is taken from the true state, but on the 100th day. The background covariance matrix is assumed to be diagonal and the variances are determined, based on truth-minus-background.

The first experiment is the standard 4DVAR. The dimension of optimization problem exceeds 10^4. In the 4DVAR experiment, we apply a preconditioning by the inverse of square root of the background error covariance matrix. The second experiment is POD 4DVAR. The POD model is constructed in the way described in Section 4, but the snapshots are taken from the background model results. The number of the snapshots is 60. Six, nine and ten POD basis functions for h and u, v are used respectively, this being sufficient to capture more than 99% energy of variability of the snapshots. The dimension of the optimization problem is 25. The third experiment is the adaptive POD 4DVAR that can update the POD model during the optimization procedures. In the adaptive POD 4DVAR experiments, the optimization comprises several outer iterations. In each outer iteration, the POD model is updated from a new set of snapshots taken from the full model results based on the result of the previous outer iteration. We stop the present outer iteration and switch to a new outer iteration following a criterion that the gradient should decrease by at least three orders of magnitude from the initial gradient value in the outer iteration minimization or a certain number of iterations are performed, whichever occurs first. A frequent updating POD model is not necessary, as we have found experimentally. However, when the old POD model cannot substantially reduce the cost function, it is necessary to update the POD model. In future research we plan to test various strategies for updating in POD 4DVAR such as trust region POD,

DOI: 10.1002/fld

etc. The results in the 4DVAR and POD 4DVAR experiments both show that the cost functions decrease substantially during the first several iteration times. Then the cost functions decrease in a rather modest fashion. The number of iterations as well as the extent of the gradient reduction attained in every inner iteration to justify an updating procedure are drawn from experience.

The numerical solution of the optimal control problem is obtained using the M1QN3 large-scale unconstrained minimization routine, which is based on a limited memory quasi-Newton method. In the optimization software M1QN3 used in this study, the user can set a parameter m that determines the total memory used by the optimization. With larger m the optimization will use more information from the approximated Hessian matrix. Generally, $3 \leqslant m \leqslant 7$. In our case we use $m = 5$.

4.2. Results

Here we present the numerical results for the three experiments. The assimilation window is one year. The true state is taken from the 21st year of 21-year simulation. The background field for initial condition in the three experiments is taken from the true state on the 100th day. The number of snapshots used in POD 4DVAR and adaptive POD 4DVAR is 60 and the energy captured is more than 99%.

Figure 1 shows the history of the cost function and its gradient during the full 4DVAR experiment. The cost function was reduced by more than three orders of magnitude. The gradient of the cost function is also sufficiently reduced that indicates that 4DVAR can successfully approximate the minimum after 150 iterations.

Figure 2 shows the history of the minimization of the cost function and its gradient in the POD 4DVAR experiment. The reduction of the cost function is less than that obtained in the full 4DVAR experiment. The gradient of the cost functional is reduced by more than three orders in magnitude. The POD 4DVAR has the limitation that the optimal solution can only be sought within the space spanned by POD basis of background fields. When observations lay outside of the POD space, the POD 4DVAR solution may fail to provide a sufficient fit to observations. This limitation can be improved by adaptively updating POD bases during the optimization.

Figure 3 shows the history of the minimization of the cost function and its gradient in the adaptive POD 4DVAR experiment. The cost function is reduced much more than in POD 4DVAR experiment, and is closer to the result obtained in 4DVAR experiments. The final value of the cost function obtained is about 1/20 of that of the first guess.

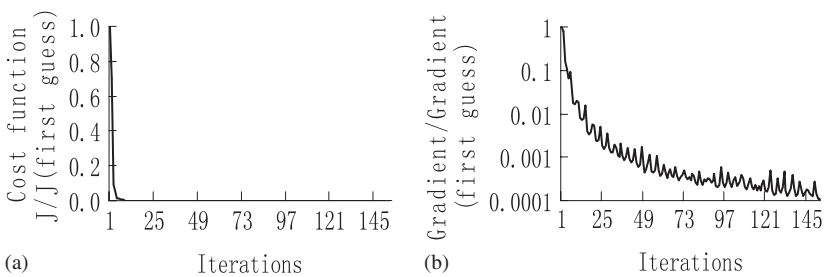

Figure 1. Evolution of the cost function and gradient in 4DVAR experiment: (a) cost function; and (b) gradient as a function of the number of minimization iterations.

DOI: 10.1002/fld

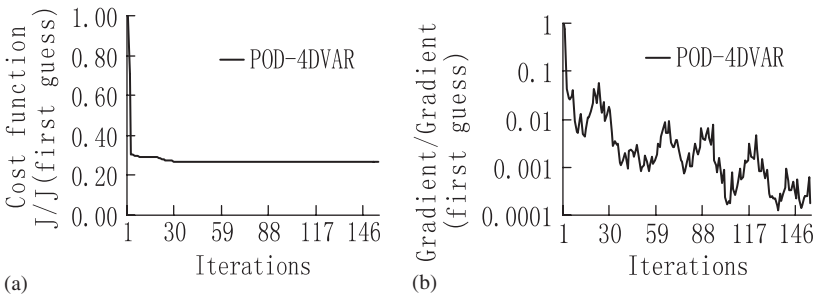

Figure 2. Evolution of the cost function and gradient in POD 4DVAR experiment: (a) cost function; and (b) gradient as a function of the number of minimization iterations.

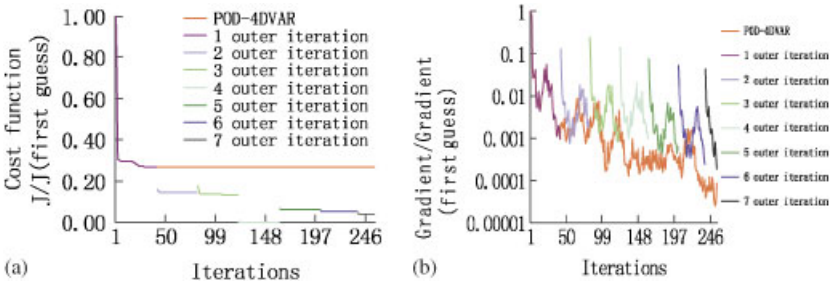

Figure 3. Evolution of the cost function and gradient in adaptive POD 4DVAR experiment: (a) cost function; and (b) gradient as a function of the number of minimization iterations.

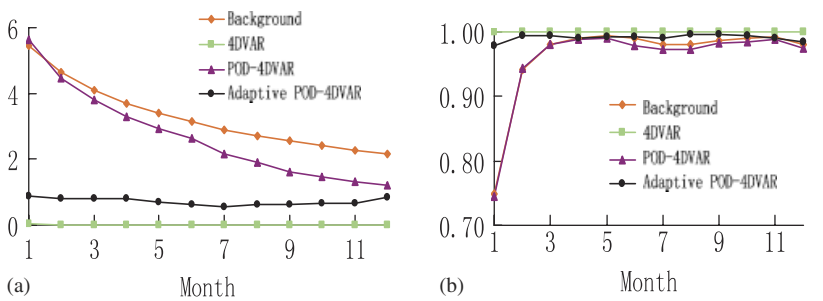

Figure 4. RMSE and correlation of the results compared to the true state for upper layer thickness.

Figure 4 shows RMSE and correlation coefficients of the outcomes of the three experiments compared to the true state. All the three experiments have smaller errors than the background. The 4DVAR yields upper layer thickness results that turn out to have the smallest errors. The adaptive POD 4DVAR result exhibits smaller errors than those of POD-4DVAR in term of upper layer thickness. Similar results hold in the terms of the correlation coefficients.

DOI: 10.1002/fld

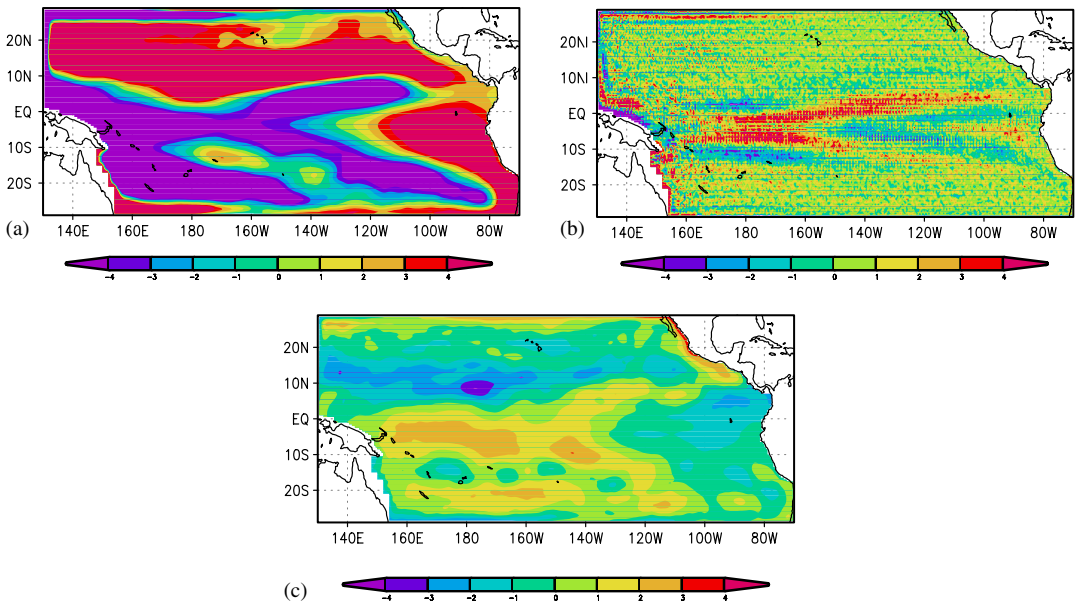

Figure 5. Errors between the true state and the numerical approximations for upper layer thickness h (m) in the initial time: (a) the error between the true state and background state; (b) the error between the true state and 4DVAR; and (c) the error between the true state and adaptive POD 4DVAR.

The POD 4DVAR as a reduced order approach with much less computation cost may not achieve exactly the same cost function reduction as the high-resolution 4DVAR does. There are several potential approaches to yield more cost function reduction using POD 4DVAR. One is using more snapshots and more POD bases. We tried this approach and found that there are limitations to further reduce the cost function. This may be related to the way we choose the snapshots. In our approach, we chose snapshots at evenly distributed instances in the assimilation window. The snap shots selected this way maybe not a good representation of the continuous flow in the assimilation window. In fact, how to select snapshots is an important issue in recent POD-related research.

Figure 5 provides an error comparison between the true state and results obtained from full 4DVAR and adaptive POD 4DVAR at the initial time. Both 4DVAR and adaptive POD 4DVAR improved initial field significantly. 4DVAR result is basically true field plus some white noise originating from noise in observations. The adaptive POD 4DVAR result does not fit observations as close as that in 4DVAR. The patch structure in Figure 5(c) indicates that indeed some small-scale features in observations cannot be resolved by POD basis.

4.3. Comparison of computational costs

The computation cost can be calculated as follows: POD model requires about $1/100$ of the full model; adaptive POD 4DVAR takes N full model integrations (N is the number of the outer loops) plus $M/100$ of the full model integrations (M is the number of the inner iterations) plus the time for determining POD basis and building POD model. Since the time required for determining POD basis and building POD model is much less than that of running the full model (about $1/50$). See the Table II for detailed computation time of each item in our experiments.

DOI: 10.1002/fld

Table II. The CPU time required in the three experiments (computer platform: IBM p. 690).

	4DVAR	POD 4DVAR	Adaptive POD 4DVAR
The full model single integration	2.5 h		2.5 h
The full adjoint single integration	3.75 h		3.75 h
The POD model construction and single integration		0.03 h	0.03 h
The adjoint POD model construction and single integration		0.04 h	0.04 h
Number of outer iterations	1	1	7
Number of inner iterations	160	160	30
Total time	1000 h	12 h	91 h

5. CONCLUSIONS

In this paper, we proposed a reduced order approach to 4DVAR using POD. The approach not only reduces the dimension of the control space, but also reduces the size of the dynamical model, both in dramatic ways. This approach also entails a convenient way of constructing the adjoint model. Further, an adaptive POD 4DVAR is also proposed. To test the POD approach to 4DVAR, a reduced-gravity tropical Pacific model is used to perform identical twin experiments in which conventional 4DVAR, POD 4DVAR and adaptive POD 4DVAR are tested and compared to each other. The main conclusions drawn from this study are:

- The POD model can accurately approximate the full order model with a much smaller size;
- The POD 4DVAR has the limitation that the optimal solution can only be sought within the space spanned by POD basis of background fields. When observations lay outside of the POD space, the POD 4DVAR solution may fail to fit observations sufficiently;
- The above limitation of POD 4DVAR can be improved by implementing adaptive POD 4DVAR. The computational time is as several times as that in the POD 4DVAR without adaptivity. But the computational cost is still much cheaper than that in 4DVAR;
- The adaptive POD 4DVAR is capable of delivering comparable results as full order 4DVAR with much less computational cost.

As an initial effort to dramatically reduce computational cost of 4DVAR, the testing assimilation experiments in this study are not as realistic as those in realistic applications. We use a simple model and model simulated data. The control variables consist of initial condition only. However, the results are very promising and show that further research efforts in this direction are worth pursuing and may lead ultimately to a practical implementation of POD 4DVAR in operational NWP and ocean forecasts. In future study, real data and more realistic ocean general circulation models should be tested. For ocean models, the atmospheric forcing fields should also be included in control variables.

ACKNOWLEDGEMENTS

This study is supported by Natural Science Foundation of China (40437017, 40225015). Prof. I. M. Navon acknowledges the support from the NSF grant number ATM-9731472.

DOI: 10.1002/fld

REFERENCES

1. LeDimet FX, Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* 1986; **38A**:97–110.
2. Courtier P, Thepaut J-N, Hollingsworth A. A strategy for operational implementation of 4D-VAR, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society* 1994; **120**:1367–1388.
3. Awaji T, Masuda S, Ishikawa Y, Sugiura N, Toyoda T, Nakamura T. State estimation of the North Pacific Ocean by a four-dimensional variational data assimilation experiment. *Journal of Oceanography* 2003; **59**(6):931–943.
4. Lawless AS, Gratton S, Nichols NK. An investigation of incremental 4D-VAR using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society* 2005; **131**:459–476.
5. Griewank A, Walther A. Revolve: an implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Transactions on Mathematical Software* 2000; **26**:19–45.
6. Li Z, Navon IM, Yanqiu Zhu Y. Performance of 4D-VAR with different strategies for the use of adjoint physics with the FSU global spectral model. *Monthly Weather Review* 2000; **128**(3):668–688.
7. Tremolet Y. Diagnostics of linear and incremental approximations in 4D-VAR. *Quarterly Journal of the Royal Meteorological Society* 2004; **130**(601):2233–2251.
8. Restrepo J, Leaf G, Griewank A. Circumventing storage limitations in variational data assimilation studies. *SIAM Journal on Scientific Computing* 1998; **19**:1586–1605.
9. Griewank A. Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optimization Methods and Software* 1992; **1**:35–54.
10. Todling R, Cohn SE. Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Monthly Weather Review* 1994; **122**:2530–2557.
11. Todling R, Cohn SE, Sivakumaran NS. Suboptimal schemes for retrospective data assimilation based on the fixed-lag Kalman smoother. *Monthly Weather Review* 1998; **126**:2274–2286.
12. Cane MA, Kaplan A, Miller RN, Tang B, Hackert EC, Busalacchi AJ. Mapping tropical Pacific sea level: data assimilation via a reduced state space Kalman filter. *Journal of Geophysical Research* 1996; **101**:22599–22617.
13. Evensen G. Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research* 1992; **97**(C11):17905–17924.
14. Farrell BF, Ioannou PJ. State estimation using a reduced order Kalman filter. *Journal of the Atmospheric Sciences* 2001; **58**:3666–3680.
15. Beck A, Ehrendorfer M. Singular-vector-based covariance propagation in a quasigeostrophic assimilation system. *Monthly Weather Review* 2005; **133**:1295–1310.
16. Hoang S, Baraille R, Talagrand O, Carton X, De Mey P. Adaptive filtering: application to satellite data assimilation in oceanography. *Dynamics of Atmospheres and Oceans* 1997; **27**(1–4):257–281.
17. Kaplan A, Kushnir Y, Cane MA, Blumenthal MB. Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *Journal of Geophysical Research-Oceans* 1997; **102**(C13):27835–27860.
18. Kaplan A, Kushnir Y, Cane MA. Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992. *Journal of Climate* 2000; **13**(16):2987–3002.
19. Kaplan A, Cane MA, Kushnir Y. Reduced space approach to the optimal analysis interpolation of historical marine observations: accomplishments, difficulties, and prospects. *Advances in the Applications of Marine Climatology*: *The Dynamic Part of the WMO Guide to the Applications of Marine Climatology*, *WMO/TD-1081*. World Meteorological Organization, Geneva, Switzerland, 2003; 199–216.
20. Canizares R, Kaplan A, Cane MA, Chen D, El Nino S. Use of data assimilation via linear low-order models for the initialization of El Nino Southern Oscillation predictions. *Journal of Geophysical Research-Oceans* 2001; **106**(C12):30947–30959.
21. Lermusiaux PFJ, Robinson AR. Data assimilation via error subspace statistical estimation. Part I: theory and schemes. *Monthly Weather Review* 1999; **127**(7):1385–1407.
22. Lermusiaux PFJ. Data assimilation via error subspace statistical estimation. Part II: Middle Atlantic Bight shelfbreak front simulations and ESSE validation. *Monthly Weather Review* 1999; **127**(7):1408–1432.
23. Evensen G. Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research* 1994; **99**(C5):10143–10162.
24. Fukumori I, Malanotte-Rizzoli P. An approximate Kalman filter for ocean data assimilation—an example with an idealized gulf-stream model. *Journal of Geophysical Research-Oceans* 1995; **100**(C4):6777–6793.
25. Fukumori I. Assimilation of TOPEX sea level measurements with a reduced-gravity, shallow water model of the tropical Pacific Ocean. *Journal of Geophysical Research-Oceans* 1995; **100**(C12):25027–25039.

DOI: 10.1002/fld

26. Blayo E, Blum J, Verron J. Assimilation Variationnelle de Donnees en Ocenaographie et reduction de la dimension de l'espace de controle. *Equations aux Derivees Partielles et Applications* 1998; 199–218.
27. Durbiano S. Vecteurs caracteristiques de modeles oceaniques pour la reduction d'ordre en assimilation de donnees. *Ph.D. Thesis*, Universite Joseph Fourier, Laboratoire de Modelisation et calcul. Grenoble, France, 2001, 214.
28. Robert C, Durbiano S, Blayo E, Verron J, Blum J, Le Dimet FX. A reduced-order strategy for 4D-VAR data assimilation. *Journal of Marine Systems* 2005; **57**(1–2):70–82.
29. Hoteit I, Pham DT. Evolution of the reduced state space and data assimilation schemes based on the Kalman filter. *Journal of the Meteorological Society of Japan* 2003; **81**(1):21–39.
30. Peterson JS. The reduced basis method for incompressible viscous flow calculations. *SIAM Journal on Scientific and Statistical Computing* 1989; **10**(4):777–786.
31. Fukunaga K. *Introduction to Statistical Recognition*. Academic Press: New York, 1990.
32. Aubry N, Holmes P, Lumley JL, Stone E. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Mechanics* 1998; **192**:115–173.
33. Bansch E. An adaptive finite-element-strategy for the three-dimensional time-dependent Navier–Stokes equations. *Journal of Computational Mathematics* 1991; **36**:3–28.
34. Afanasiev *et al.* 2001.
35. Arian E, Fahl M, Sachs EW. Trust-region proper orthogonal decomposition for flow control. NASA/CR-. 2000-210124, *ICASE Report No. 2000-25*, 2000.
36. Ly HV, Tran HT. Proper orthogonal decomposition for flow calculations and optimal control in a horizontal CVD reactor. *Quarterly of Applied Mathematics* 2002; **60**(4):631–656.
37. Banks HT, Joyner ML, Winchesky B, Winfree WP. Nondestructive evaluation using a reduced-order computational methodology. *Inverse Problems* 2000; **16**:1–17.
38. Atwell JA, Borggaard JT, King BB. Reduced order controllers for Burgers' equation with a nonlinear observer. *International Journal of Applied Mathematics and Computer Science* 2001; **11**(6):1311–1330.
39. Gunzburger MD. Perspectives in flow control and optimization. *Society for Industrial and Applied Mathematics* 2003; 261.
40. Gunzburger MD. Reduced-order modeling, data compression, and the design of experiments. *Second DOE Workshop on Multiscale Mathematics*. Broomfield, CO, 20–22 July 2004.
41. Daescu DN, Navon IM. Efficiency of a Pod-based reduced second order adjoint model in 4-D VAR data assimilation. *International Journal for Numerical Methods in Fluids* 2006, in press.
42. Ravindran SS. A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *International Journal for Numerical Methods in Fluids* 2000; **34**(5):425–448.
43. Moore DW, Philander SGH. Modeling the equatorial oceanic circulation. In *The Sea*, Vol. VI. Wiley: New York, 1977; 319–361.
44. Stricherz, James N, O'Brien JJ, Legler DM. *Atlas of Florida State University Tropical Pacific Winds for TOGA 1966–1985*, Florida State University, Tallahassee, FL, 1992; 250.
45. Cao Y, Zhu J, Luo Z, Navon IM. Reduced order modeling of the upper tropical Pacific ocean model using proper orthogonal decomposition. *Computers and Mathematics with Applications* 2005, in press.

DOI: 10.1002/fld