

# A POD reduced order 4D-Var adaptive mesh ocean modelling approach

F. Fang<sup>\*1</sup>, C. C. Pain<sup>1</sup>, I. M. Navon<sup>2</sup>, M. D. Piggott<sup>1</sup>, G. J. Gorman<sup>1</sup>, P. Allison<sup>1</sup>,  
A. J. H. Goddard<sup>1</sup>

<sup>1</sup> *Applied Modelling and Computation Group,  
Department of Earth Science and Engineering,  
Imperial College, Prince Consort Road, London, SW7 2BP, U.K.*

*URL: <http://amcg.es.e.imperial.ac.uk>*

<sup>2</sup> *School of Computational Science  
and Department of Mathematics, Florida State University,  
Tallahassee, FL, 32306-4120, USA.*

*URL: <http://people.scs.fsu.edu/~navon/index.html>*

## SUMMARY

A novel Proper Orthogonal Decomposition (POD) inverse model, developed for a mesh adaptive ocean model (the Imperial College Ocean Model, ICOM) is presented here. The new POD model is validated using the Munk gyre flow test case, where it inverts for initial conditions. The optimised velocity fields exhibit overall good agreement with those generated by the full model. The correlation between the inverted modelled and the true velocity is 80%-98% over the majority of the domain. Error estimation (including the projection error, subspace integration error and the error introduced by the controls) was used to judge quality of reduced adaptive mesh models. The cost function (consisting of the misfit

between the inverted modelled and true velocity values spatially and temporally) is reduced by 50% of its original value, and further by 25% after the POD bases are updated.

In this study, the reduced adjoint model is derived directly from the discretised reduced forward model. The whole optimisation procedure is undertaken completely in reduced space. Computational cost for the 4-D Var data assimilation is significantly reduced (here a decrease of 70% in the test case) by decreasing the dimensional size of the control space, in both the forward and adjoint models. Computational efficiency is further enhanced (by a factor of  $\mathcal{N}$ , here,  $\mathcal{N}$  is the number of times to run the reduced models) since both the reduced forward and adjoint models are constructed by a series of time-independent sub-matrices. These sub-matrices are calculated prior to running the reduced models. The reduced forward and adjoint models can thus be run repeatedly with negligible computational costs.

An adaptive POD 4-D Var is employed to update the POD bases as minimisation advances and loses control, thus adaptive updating of the POD bases is necessary (here, when the value of cost function cannot be decreased by more than  $10^{-3}$  between the consecutive iterations). It is noted that the adaptive POD 4D-Var is not always effective. An appropriate choice of initial guess controls can help to achieve a reasonable minimum location during the optimisation procedure.

Previously developed numerical approaches [1] are employed to accurately represent the geostrophic balance and improve the efficiency of the POD simulation.

**keywords:** inverse, adjoint, POD; reduced-order modelling; ocean model; finite element; unstructured adaptive mesh Copyright © 2004 John Wiley & Sons, Ltd.

## 1. Introduction

The threat of impending climate change highlights the importance of improving the predictive capabilities of ocean models. Data assimilation techniques are a critical component of ocean modelling. By assimilating observations (such as *in situ* measurements and remote sensing) into models, unknown inputs such as initial and boundary conditions, bottom friction coefficients,



turbulent viscosity, and wind stress etc. can be optimised [2, 3, 4, 5, 6, 7, 8, 9]

A variety of approaches have in the past been used to facilitate data assimilation and include statistical interpolation methods, nudging data assimilation and variational methods along with sequential estimation such as Kalman filter (KF), extended Kalman filter (EKF) and Ensemble Kalman filter (EnKF). In particular the four-dimensional variational (4D-Var) method has proved an efficient means of assimilating observed data into simulations [10, 11]. the 4D-Var method is capable of producing a best estimate model solution by fitting a numerical simulation to observational data over both space and time. The technique also facilitates the estimation of the error sources caused by uncertainties (boundary conditions, initial condition and parameters) in the model. The solution is derived by minimising a cost function that contains the misfits between the data and dynamical model, as well as the covariances specifying spatial and temporal correlations of errors. 4D-Var data assimilation has been used widely in both atmospheric and oceanographic models over the past two decades [12, 13, 10, 14, 15, 16, 17, 18, 8, 9, 15, 19, 5, 7, 20, 21]. However, the major difficulty in the implementation of 4D-Var data assimilation in an ocean model is the large dimensionality of the control space (for a discrete realistic model, the size of the control variables is typically in the range  $10^6 - 10^8$ ), and hence the 4D-Var method incurs high memory and computational costs. However the computational cost can be reduced by decreasing the dimensions of the control space thus ensuring that the minimisation of the cost function (or error covariances) is carried out within a low-dimensional space. This can be achieved through an incremental



approach ([22]) whereby a succession of quadratic problems are generated over increasing time periods. The successive quadratic minimisation problems can then be solved by running tangent linear model and adjoint model approximations using a coarse resolution in the inner-

loop of the minimisation. Using this approach, the dimension of the minimisation problem can be decreased by one or two orders of magnitude. Although such an incremental approach is currently used in operational atmospheric models, the dimensions of the control space remain quite large in realistic applications [25].

The use of Empirical Orthogonal Functions (EOF) analysis has been advocated as an approach that can lead to reduced order ocean modelling ([26]). The implementation of this method additionally results in a drastic reduction of the dimension of the control space and thus the iterative minimisation process [25, 27]. Reduced order 4D-Var can also be used to precondition 4D-Var, and reduce computational cost [28]. It has been further proposed ([29]) that efficiencies can be enhanced if the adjoint model can be directly implemented in a subspace of the reduced model (determined by the leading EOFs) and then used to approximate the gradient of the cost function. The minimisation process can thus be solved completely in reduced space with negligible computational costs.

Proper Orthogonal Decomposition (POD) methodologies, in combination with the Galerkin projection procedure have additionally provide an efficient means of generating reduced order models [30, 31, 32]. This technique essentially identifies the most energetic modes in a time-dependent system thus providing a means of obtaining a low-dimensional description of the system's dynamics. POD has been widely and successfully applied to diverse disciplines, including signal analysis and pattern recognition [33], fluid dynamics and coherent structures [34, 30, 35] and image reconstruction [36]. To improve the accuracy of reduced models, a goal-oriented approach has been used to optimise the POD bases [37, 38]. The dual-weighted POD approach provides an 'enriched' set of basis functions combining information from both model dynamics and the data assimilation system. The practical utility of this approach has

been extended to include ocean and climate modelling and the solution of inverse problems [25, 27, 39, 31]. The POD-based 4D-Var not only reduces the dimension of control space, but also reduces the size of dynamical model, both in dramatic ways [39, 40].

Herein we describe a POD reduced order 4D-Var for an adaptive mesh ocean model. A POD-based reduced forward model [1] has been developed for the Imperial College Ocean Model (ICOM) that can simultaneously resolve both small and large scale ocean flows while smoothly varying mesh resolution and conforming to complex coastlines and bathymetry. In this work, a further step has been made to introduce the POD approach into an adaptive mesh refinement adjoint model. Using the POD and Galerkin projection approaches, the reduced forward model is derived in a subspace (details in [1]). Once the forward reduced model is available, the reduced order adjoint model can directly be obtained from the POD reduced forward model in the subspace, instead of the original forward model. The minimisation procedure is then carried out in the reduced space.

When adaptive meshes are employed in both the forward and adjoint models, the mesh resolution requirements for each model may be spatially and temporally different, as the meshes are adapted according to the flow features of each model. This poses additional challenges for the implementation of an inverse POD-based reduced adaptive model, which include snapshots of varying length at time levels; and the fact that the POD base of the forward model can differ from the POD base of the adjoint model. To overcome these difficulties, a standard reference fixed mesh is adopted for both the forward and adjoint reduced models. The solutions for both are interpolated from their own mesh onto the same reference fixed mesh at each time level. This allows the same number of base modes for both the reduced forward and adjoint models.

In this work, an adaptive POD procedure is employed to improve the reduced model by

updating the POD basis. The original reduced basis for inverse problems is calculated using a set of snapshots based on the results from the full forward model with the specified control variables. The re-calculation of the reduced basis is needed when the resulting control variables from the optimisation procedure are significantly different from those that the POD model is based on. Ravindran (2002, 2006) [41, 42] proposed an adaptive procedure that successively updates the reduced-order model being used via a Sequential Quadratic Programming (SQP) constrained optimisation algorithm. Cao et al. (2007) introduced an adaptive POD approach into POD inverse (adjoint) models. In this approach the reduced basis is recalculated using a refreshed set of snapshots based on the latest results obtained from the full forward model using a restart criterion of the adaptive POD procedure based on convergence of the minimisation process. One can also consider the Trust Region Method for restart criterion [43].

## 2. POD reduced model

POD is the most efficient choice among linear decompositions in the sense that it can capture the greatest possible kinetic energy. A 3D dynamical flow model is generally written as:

$$\frac{\partial \mathbf{u}}{\partial t} = f(\mathbf{u}, t, \mathbf{x}), \quad (1)$$

where  $f$  is a general function representing a 3D nonlinear flow dynamics (here, Navier-Stokes equation),  $\mathbf{u}$  is a vector containing all variables to be solved (e.g., velocities, pressure and temperature etc.),  $t$  is time and  $\mathbf{x} = (x, y, z)^T$  represents the Cartesian coordinate position.

### 2.1. Proper Orthogonal Decomposition

The model variables  $\mathbf{u}$  are sampled at defined checkpoints during the simulation period  $[t_1, \dots, t_K]$ , also referred to as snapshots  $U = (U_1, \dots, U_K)^T$  where  $K$  is the number of

snapshots. The snapshots can be obtained either from a numerical model of the phenomenon or from experiments/observations. The sampled values of variables at the snapshot  $k$  are stored in a vector  $U_k$  with  $\mathcal{N}$  entries,  $U_k = (U_{k,1}, \dots, U_{k,i}, \dots, U_{k,\mathcal{N}})$ , where  $\mathcal{N}$  is the number of nodes and  $U$  represents one of variables  $u, v, w, p$ . The average of the ensemble of snapshots is defined as:

$$\bar{U}_i = \frac{1}{K} \sum_{k=1}^K U_{k,i}, \quad 1 \leq i \leq \mathcal{N}, \quad (2)$$

where  $U_{k,i}$  is the model variable value at the snapshot  $k$  and node  $i$ . Taking the deviation from the mean of variables yields

$$V_{k,i} = U_{k,i} - \bar{U}_i, \quad 1 \leq i \leq \mathcal{N}. \quad (3)$$

A collection of all  $V_{k,i}$  constructs a rectangular  $\mathcal{N}$  by  $K$  matrix  $A$ . The aim of POD is to find a set of orthogonal basis functions which can represent the most dynamic energy in the original flow system. The  $N \times N$  eigenvalue problem is established

$$AA^T x_k = \lambda_k x_k; \quad 1 \leq k \leq K. \quad (4)$$

The order  $\mathcal{N}$  for matrix  $AA^T$  is far larger than the order  $K$  for matrix  $A^T A$  in realistic ocean cases. Therefore the  $K \times K$  eigenvalue problem is solved

$$A^T A y_k = \lambda_k y_k; \quad 1 \leq k \leq K. \quad (5)$$

This procedure is equivalent to a Singular Value Decomposition (SVD). The eigenvalues  $\lambda_k$  are real and positive and for reasons which will be apparent, should be sorted in an descending order. The POD basis vectors  $\Phi_k$  associated with the eigenvalues  $\lambda_k$  are orthogonal and expressed as follows:

$$\Phi_k = A y_k / \sigma_k = A y_k / \sqrt{\lambda_k}, \quad (6)$$

where the  $k^{th}$  eigenvalue is a measure of the kinetic energy transferred within the  $k^{th}$  basis mode. If the POD spectrum (energy) decays fast enough, practically all the support of the invariant measure is contained in a compact set. Roughly speaking, all the likely realisations in the ensemble can be found in a relatively small set of bounded extent. By neglecting modes corresponding to the small eigenvalues, the following formula is therefore defined to choose a low-dimensional basis of size  $M$ , where  $M \ll K$ ,

$$I(M) = \frac{\sum_{i=1}^M \lambda_i}{\sum_{i=1}^K \lambda_i}, \quad (7)$$

where  $I(M)$  represents the percentage of energy which is captured by the POD basis  $\Phi_1, \dots, \Phi_m, \dots, \Phi_M$ .

## 2.2. Reduced model

The variables in (1) can be expressed as an expansion of the POD basis functions  $\{\Phi_1, \dots, \Phi_M\}$ :

$$\mathbf{u}(t, x, y, z) = \bar{\mathbf{u}} + \sum_{m=1}^M \alpha_m(t) \Phi_m(\mathbf{x}), \quad (8)$$

where  $\bar{\mathbf{u}}$  is the mean of the ensemble of snapshots for the variables  $\mathbf{u}(t)$ ,  $\alpha_m$  ( $1 \leq m \leq M$ ) are the time-dependent coefficients to be determined, and  $\alpha_m(0)$  are the coefficients at the initial time level. Substituting (8) into (1) and taking the POD basis function as the test function, then integrating over the computational domain  $\Omega$ ,

$$\int_{\Omega} \Phi_m f \left( (\bar{\mathbf{u}} + \sum_{m=1}^M \alpha_m \Phi_m(\mathbf{x})), t, \mathbf{x} \right) d\Omega. \quad (9)$$

The POD reduced model is then obtained:

$$\frac{\partial \alpha_m}{\partial t} = \left\langle f \left( (\bar{\mathbf{u}} + \sum_{m=1}^M \alpha_m(t) \Phi_m(\mathbf{x})), t, \mathbf{x} \right), \Phi_m \right\rangle, \quad (10)$$

subject to the initial condition

$$\alpha_m(0) = ((\mathbf{u}(0, \mathbf{x}) - \bar{\mathbf{u}}(\mathbf{x})), \Phi_m). \quad (11)$$



In the finite element method, the POD basis  $\Phi_m(\mathbf{x}) = \sum_{i=1}^{\mathcal{N}} N_i \Phi_{m,i}$ , equation (10) can therefore be written as:

$$\frac{\partial \alpha_m}{\partial t} = \left\langle f \left( (\bar{\mathbf{u}} + \sum_{m=1}^M \alpha_m(t) \sum_{i=1}^{\mathcal{N}} N_i \Phi_{m,i}), t, \mathbf{x} \right), \Phi_m \right\rangle, \quad (12)$$

where  $N_i$  is the basis function in the finite element and  $\mathcal{N}$  is the number of nodes in the computational domain.

### 3. Reduced order 4D-Var

The aim of 4D-Var is to determine optimal control variables (e.g., initial conditions). Optimal solution for (1) is obtained by minimising the functional  $\mathfrak{J}(U^0)$ :

$$\mathfrak{J}(U^0) = \frac{1}{2} (U^0 - U_b)^T \mathbf{B}^{-1} (U^0 - U_b) + \frac{1}{2} \sum_{n=1}^{N_t} (\mathbf{H}U^n - y_o^n)^T W_o (\mathbf{H}U^n - y_o^n), \quad (13)$$

where  $\mathbf{B}$  is the background error covariance matrix,  $W_o$  is the observation error covariance matrix,  $\mathbf{H}$  is the observation operator,  $U^0$  is a vector containing the control variables (here, initial conditions),  $U^n$  is a vector containing the solution of variables from the model (the reduced order model) at the time level  $n$  (here,  $N_t$  is the number of time levels), and  $y_o^n$  is the observation at time level  $n$ . In a POD reduced model, the initial value  $U^0$  and the reduced order solution  $U^n$  are expressed as:

$$U^0 = \bar{U} + \sum_{m=1}^M \alpha_m(0) \Phi_m(\mathbf{x}), \quad (14)$$

$$U^n = \bar{U} + \sum_{m=1}^M \alpha_m(t^n) \Phi_m(\mathbf{x}). \quad (15)$$

Substituting (14) and (15) into (13) yields:

$$\begin{aligned} \mathfrak{S}(\alpha(0)) = & \frac{1}{2} \left( \left( \bar{U} + \sum_{m=1}^M \alpha_m(0) \Phi_m(\mathbf{x}) \right) - U_b \right)^T \mathbf{B}^{-1} \left( \left( \bar{U} + \sum_{m=1}^M \alpha_m(0) \Phi_m(\mathbf{x}) \right) - U_b \right) \\ & + \frac{1}{2} \sum_{n=1}^{N_t} \left( \mathbf{H} \left( \bar{U} + \sum_{m=1}^M \alpha_m(t^n) \Phi_m(\mathbf{x}) \right) - y_o^n \right)^T W^o \left( \mathbf{H} \left( \bar{U} + \sum_{m=1}^M \alpha_m(t^n) \Phi_m(\mathbf{x}) \right) - y_o^n \right). \end{aligned} \quad (16)$$

where the snapshots are chosen at time intervals with a constant time interval between them during the simulation period.

### 3.1. Discrete reduced order adjoint equations

The discrete forward model of (10) at the time level  $n$  can be written in a subspace:

$$\mathbf{A}^n \alpha^n = s^n, \quad (17)$$

where,

$$s^n = \mathbf{B}^n \alpha^{n-1} + f_s, \quad (18)$$

$\mathbf{A}^n$  and  $\mathbf{B}^n$  ( $\mathbf{A}^n, \mathbf{B}^n \in R^{M \times M}$ , where,  $M$  is the number of POD bases) are the matrices at the time level  $n$  which include all the discretization of (10),  $\alpha^n = (\alpha_1^n, \dots, \alpha_M^n)$  and  $\alpha^{n-1} = (\alpha_1^{n-1}, \dots, \alpha_M^{n-1})$  are the vectors of variables to be solved at the time levels  $n$  and  $n-1$  respectively, here including the coefficients related to the POD basis functions for state variables  $\mathbf{u}$  in (1),  $n$  is the time index,  $s^n$  is a discretised source term at the time level  $n$  and  $f_s$  is a source term including the forcing terms on the boundaries.

For a nonlinear simulation, the matrices  $\mathbf{A}^n$  and  $\mathbf{B}^n$  can be written as:

$$\mathbf{A}^n = \hat{\mathbf{A}}_0^n + \sum_{m=1}^M \alpha_m^{n-1} \hat{\mathbf{A}}_m^n, \quad (19)$$

$$\mathbf{B}^n = \hat{\mathbf{B}}_0^n + \sum_{m=1}^M \alpha_m^{n-1} \hat{\mathbf{B}}_m^n, \quad (20)$$

where,  $\mathbf{A}^n, \mathbf{B}^n, \hat{\mathbf{A}}_0^n, \hat{\mathbf{B}}_0^n, \hat{\mathbf{A}}_m^n, \hat{\mathbf{B}}_m^n \in R^{M \times M}$  and the components can be expressed as:

$$\hat{\mathbf{A}}_{0,i,j}^n = \int_{\Omega} \Phi_i f(\theta \bar{\mathbf{u}} \Phi_j) d\Omega, \quad 1 \leq i, j \leq M, \quad (21)$$

$$\hat{\mathbf{A}}_{m,i,j}^n = \int_{\Omega} \Phi_i f(\theta \Phi_j \Phi_m) d\Omega, \quad 1 \leq i, j \leq M, \quad (22)$$

$$\hat{\mathbf{B}}_{0,i,j}^n = \int_{\Omega} \Phi_i f((\theta - 1) \bar{\mathbf{u}} \Phi_j) d\Omega, \quad 1 \leq i, j \leq M, \quad (23)$$

$$\hat{\mathbf{B}}_{m,i,j}^n = \int_{\Omega} \Phi_i f((\theta - 1) \Phi_j \Phi_m) d\Omega, \quad 1 \leq i, j \leq M. \quad (24)$$

(19) and (20) can be rewritten as:

$$\mathbf{A}^n = \hat{\mathbf{A}}_0^n + \hat{\mathbf{A}}^n \alpha^{n-1}, \quad (25)$$

$$\mathbf{B}^n = \hat{\mathbf{B}}_0^n + \hat{\mathbf{B}}^n \alpha^{n-1}, \quad (26)$$

where  $\hat{\mathbf{A}}^n = (\hat{\mathbf{A}}_1^n, \dots, \hat{\mathbf{A}}_M^n)$  and  $\hat{\mathbf{B}} = (\hat{\mathbf{B}}_1^n, \dots, \hat{\mathbf{B}}_M^n)$  [1].

Taking into account (17), the discrete forward equation during the simulation period  $[t_1, \dots, t_{N_t}]$  can be written:

$$\mathbf{A} \alpha = s \quad (27)$$

where,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^1 & & & & \\ -\mathbf{B}^2 & \mathbf{A}^2 & & & \\ & \ddots & \ddots & & \\ & & & -\mathbf{B}^{N_t} & \mathbf{A}^{N_t} \end{pmatrix} \quad (28)$$

and

$$\alpha = (\alpha^1, \alpha^2, \dots, \alpha^{N_t})^T \quad (29)$$

$$s = (\mathbf{B}^1 \alpha^0, 0, \dots, 0)^T + f_s \quad (30)$$

Differentiating (27) with respect to the control variables to be optimised (i.e., the initial coefficient,  $\alpha^0 = \alpha(0)$ ), the tangent linear model is obtained

$$\bar{\mathbf{A}}\alpha + \mathbf{A}\bar{\alpha} = \bar{s}, \quad (31)$$

where the overbar is defined as the differentiation with respect to the control variables  $\alpha^0$

$$\bar{\mathbf{A}} = \frac{\partial \mathbf{A}}{\partial \alpha^0} = \begin{pmatrix} \bar{\mathbf{A}}^1 & & & & \\ -\bar{\mathbf{B}}^2 & \bar{\mathbf{A}}^2 & & & \\ & \ddots & \ddots & & \\ & & & -\bar{\mathbf{B}}^{N_t} & \bar{\mathbf{A}}^{N_t} \end{pmatrix}, \quad (32)$$

and

$$\bar{\alpha} = \frac{\partial \alpha}{\partial \alpha^0} = (\bar{\alpha}^1, \bar{\alpha}^2, \dots, \bar{\alpha}^{N_t})^T \quad (33)$$

$$\bar{s} = \frac{\partial s}{\partial \alpha^0} = (\bar{B}^1 \alpha^0, 0, \dots, 0)^T + (B^1 \bar{\alpha}^0, 0, \dots, 0)^T \quad (34)$$

where taking into account (25) and (26)

$$\bar{\mathbf{A}}^n = \frac{\partial \mathbf{A}^n}{\partial \alpha^0} = \hat{\mathbf{A}} \bar{\alpha}^{n-1}, \quad (35)$$

$$\bar{\mathbf{B}}^n = \frac{\partial \mathbf{B}^n}{\partial \alpha^0} = \hat{\mathbf{B}} \bar{\alpha}^{n-1}, \quad (36)$$

where  $n$  describes time levels,  $1 \leq n \leq N_t$ . (31) can therefore be rewritten as

$$\begin{pmatrix} \bar{\mathbf{A}}^1 & & & & \\ -\bar{\mathbf{B}}^2 & \bar{\mathbf{A}}^2 & & & \\ & \ddots & \ddots & & \\ & & & -\bar{\mathbf{B}}^{N_t} & \bar{\mathbf{A}}^{N_t} \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^{N_t} \end{pmatrix} + \begin{pmatrix} \mathbf{A}^1 & & & & \\ -\mathbf{B}^2 & \mathbf{A}^2 & & & \\ & \ddots & \ddots & & \\ & & & -\mathbf{B}^{N_t} & \mathbf{A}^{N_t} \end{pmatrix} \begin{pmatrix} \bar{\alpha}^1 \\ \bar{\alpha}^2 \\ \vdots \\ \bar{\alpha}^{N_t} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{B}}^1 \alpha^0 \bar{\alpha}^0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{B}^1 \bar{\alpha}^0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (37)$$

The tangent linear model is then derived

$$(\mathbf{A} + \mathbf{A}_{extra}) \frac{\partial \alpha}{\partial \alpha^0} = \frac{\partial s}{\partial \alpha^0}, \quad (38)$$

where  $\mathbf{A}$  is calculated in [27] and  $\mathbf{A}_{extra}$  originates from the nonlinear terms and is expressed as:

$$\mathbf{A}_{extra} = \begin{pmatrix} 0 & & & & \\ \hat{\mathbf{A}}\alpha^2 - \hat{\mathbf{B}}\alpha^1 & 0 & & & \\ & \ddots & \ddots & & \\ & & & \hat{\mathbf{A}}\alpha^{N_t} - \hat{\mathbf{B}}\alpha^{N_t-1} & 0 \end{pmatrix} \quad (39)$$

The variation of the objective function (13) with respect to the control variables  $\alpha^0$  is

$$\frac{\partial \mathfrak{J}}{\partial \alpha^0} = \left( \frac{\partial \alpha}{\partial \alpha^0} \right)^T \frac{\partial \mathfrak{J}}{\partial \alpha}. \quad (40)$$

Taking into account equation (38), yields

$$\frac{\partial \mathfrak{J}}{\partial \alpha^0} = \left( \frac{\partial s}{\partial \alpha^0} \right)^T (\mathbf{A} + \mathbf{A}_{extra})^{-T} \frac{\partial \mathfrak{J}}{\partial \alpha}. \quad (41)$$

The gradient of the objective function can then be written

$$\frac{\partial \mathfrak{J}}{\partial \alpha^0} = \left( \frac{\partial s}{\partial \alpha^0} \right)^T \alpha^*, \quad (42)$$

where,  $\alpha^*$  is the adjoint variable and can be calculated by solving the following adjoint equation:

$$(\mathbf{A} + \mathbf{A}_{extra})^T \alpha^* = \frac{\partial \mathfrak{J}}{\partial \alpha}. \quad (43)$$

### 3.2. Adaptive POD

In this work, the POD model is based on the solution of the original model for specified control variables (e.g., initial and boundary conditions, etc). It is therefore necessary to reconstruct the POD model when the resulting control variables from the latest optimisation iteration are significantly different from the ones upon which the POD model is based. An adaptive POD 4D-Var procedure is used to periodically update the POD basis, and the reduced direct and inverse models (for detail see in [41, 42]. The inversion procedure starts with the initial estimation of the control variables. An initial set of snapshots is obtained by running the full forward model, and the corresponding POD subspace and reduced model are constructed. The adaptive POD procedure proposed is as follows:

1. Set the POD iteration level  $it = 1$  and the initial guess controls  $c_{it}$ ;
2. Set up the snapshots  $U_{it}$  from the solution of the full forward model with the controls  $c_{it}$ ;
3. Calculate the POD bases (the number of POD bases is chosen to capture a prescribed energy level);
4. Project the controls  $c_{it}$  on the reduced space  $\alpha_{it,jt}$  ( $jt = 1$ );
5. Optimise the initial controls  $\alpha_{it,jt}$  on the reduced space (note: the optimisation procedure is carried out completely on the reduced space. The Polak-Ribiere nonlinear conjugate gradient (CG) technique is employed here and  $jt$  is the Nonlinear CG iteration level );
6. (a) check the value of cost function (13). If  $|\mathfrak{S}_{jt}| < \epsilon$  (where,  $\epsilon$  is the tolerance for the optimisation), then go to step 7;
- (b) if  $|\mathfrak{S}_{jt}| > \epsilon$  and  $|\mathfrak{S}_{jt} - \mathfrak{S}_{jt-1}| > 10^{-3}$  (where,  $jt - 1$  and  $jt$  are the consecutive optimisation iteration levels), then set  $jt = jt + 1$  and go back step 5;

- (c) if  $|\mathfrak{S}_{jt}| > \epsilon$  and  $|\mathfrak{S}_{jt} - \mathfrak{S}_{jt-1}| < 10^{-3}$ , then update the POD bases:
- i. find the new controls  $c_{it+1}$  by projecting the optimisation controls  $\alpha_{it,j}$  onto the original flow domain, and
  - ii. set  $it = it + 1$  and go back step 2;
7. The adaptive POD optimisation procedure is completed.

#### 4. Mesh adaptivity in reduced models

##### 4.1. Description of Imperial College Ocean Model

The POD-based reduced model presented was implemented for ICOM. This unstructure adaptive model can simultaneously resolve both small and large scale ocean flows while smoothly varying resolution and conforming to complex coastlines and bathymetry. With more appropriate focused numerical resolution (e.g. adaptive and anisotropic resolution of fronts and boundary layers, and optimal representation of vertical structures in the ocean) the ocean dynamics resulting from climate change will be predicted with greater accuracy. To accurately represent local flow around steep topography the hydrostatic assumption is not made. The pressure is split into the non-geostrophic and geostrophic parts which are solved separately. This allows the accurate representation of hydrostatic/geostrophic balance [44]. A dynamically adapting anisotropic mesh in 3-D is used here [45, 46]. Mesh adaptivity or optimisation relies on the derivation of appropriate error measures, which dictate how the mesh is to be modified [45]

#### 4.2. Anisotropic mesh adaptivity

The mesh adaptivity technique used in this work first constructs a metric tensor which encodes local error estimates using information from the current solution fields [45]. The metric tensor is used to calculate the required edge lengths and orientation to control solution errors. It is constructed so that an ideal edge length is unity when measured in metric space. This gives a guide to those elements which have edges that are too large or too small. Since the metric is dependent on both location and direction it is able to reflect locally anisotropic information within the solution. Thus inhomogeneous and anisotropic meshes result from this approach. By defining an objective functional which is based on the element quality in this metric space, an optimisation technique is used to improve the overall quality of the mesh. That is, local operations on the mesh connectivity and node positioning are performed which aim to improve the local quality of the mesh. The operations are performed on a three-dimensional tetrahedral mesh by the adaptivity method and include: edge collapsing/splitting; face to edge and edge to face swapping; edge to edge swapping; and local node movement or mesh smoothing in a fashion similar to Freitag and Ollivier- Gooch [47] and Buscaglia and Dari [48]. Constraints are imposed on these operations so as to preserve the integrity of non-planar geometrical boundaries [45, 46]. The interpolation error resulting from a piecewise linear approximation of a smooth function  $\psi$  may be written in terms of the functions second derivatives or Hessian  $H \equiv \nabla^T \nabla \phi$ . In particular, over a tetrahedral element the quantity

$$\epsilon = \mathbf{v}^T |H(\mathbf{x})| \mathbf{v}, \quad (44)$$

gives a guide to this interpolation error at location  $\mathbf{x}$  and direction  $\mathbf{v}$ . Note that here the notation  $|\cdot|$  corresponds to the absolute value of the matrix, defined by diagonalising the matrix and taking the absolute value of the eigenvalues, rather than the determinant. Taking



the maximum of (44) over all locations and all vectors  $\mathbf{v}$  that lie within the element gives a bound on the magnitude of the interpolation error on this element. For more details see [46] and the references therein.

Given a user defined error tolerance,  $\hat{\epsilon}$ , a metric tensor may be defined by

$$\hat{M} = \frac{1}{\hat{\epsilon}} |H|. \quad (45)$$

To obtain this error tolerance everywhere, length scales in the mesh should therefore take the value unity when measured in metric space. The mesh optimisation technique described above is used to achieve this aim as closely as possible. In practise the Hessian must be approximated and here reconstructed from the piecewise linear solution fields. To simultaneously satisfy separate error bounds for separate solution fields a metric is obtained for each and these are then superimposed to yield a combined error metric. Finally, by suitably altering the definition of this metric, computational constraints on the total number of nodes and maximum/minimum allowed element sizes and aspect ratios may be imposed [45]. An alternative approach of error measures (goal-based) is described in [49].

#### 4.3. Mesh adaptive technique in POD and error estimate

When adaptive meshes are employed in ocean models, the mesh resolution requirements vary spatially and temporally, as the meshes are adapted according to the flow features through the whole simulation. The dimensional size of the variable vectors is different at each time level since the number of nodes varies during the simulation. Snapshots can therefore be of different length at different time levels. This unavoidably brings difficulties in the implementation of a POD-based reduced model for an adaptive model. To overcome these difficulties, a standard reference fixed mesh is adopted for the reduced model. The solutions from the original full

model are interpolated from their own mesh onto the same reference fixed mesh at each time level, and then stored in the snapshots. The information at the snapshots is used to find the optimal POD basis. This allows the same length of base modes to be obtained at each time level. The resolution of the reference mesh and the piecewise linear interpolation errors between the two meshes (the adaptive mesh and the fixed reference mesh) may affect the accuracy of the POD simulation. To reduce the interpolation error, a high order interpolation approach is suggested in future work.

## 5. Application and discussion

The new model has been applied to 2D gyre flows. The initial conditions are optimised using the POD reduced adjoint model. The accuracy and validation of the reduced order POD adjoint model have been evaluated. Error estimation is undertaken through the comparison of the results obtained from the original (full) and POD reduced models. Furthermore, the adaptive POD approach is employed to update the POD bases (when the value of the cost function cannot be decreased by more than  $10^{-3}$  between the consecutive iterations), and its effectiveness is discussed. The Polak-Ribiere nonlinear conjugate gradient technique is employed in the implementation of the inversion.

### 5.1. Description of the case: Gyre

The POD reduced adjoint model is tested in a computational domain, 1000 km by 1000 km with a depth of  $H = 500$  m. The wind forcing on the free surface is given

$$\tau_y = \tau_0 \cos(\pi y/L), \quad \tau_x = 0.0, \quad (46)$$

where  $\tau_x$  and  $\tau_y$  are the wind stresses on the free surface along the  $x$  and  $y$  directions respectively, and  $L = 1000 \text{ km}$ . A maximum zonal wind stress of  $\tau_0 = 0.1 \text{ Nm}^{-1}$  is applied in the latitude ( $y$ ) direction. The Coriolis terms are taken into account with the beta-plane approximation ( $f = \beta y$ ) where  $\beta = 1.8 \times 10^{-11}$  and the reference density  $\rho_0 = 1000 \text{ kgm}^{-3}$ .

The problem is non-dimensionalised with the maximum Sverdrup balance velocity

$$\beta H \rho_0 v = \frac{\partial \tau}{\partial y} \leq \frac{\tau_0 \pi}{L} \Rightarrow v \leq 3.5 \times 10^{-2} \text{ m/s}^{-1} \quad (47)$$

(and so the velocity scale  $U = 3.5 \times 10^{-2} \text{ m/s}^{-1}$  is used here), and the length scale  $L = 1000 \text{ km}$ . Time is non-dimensionalised with  $T = \frac{L}{U}$ . Incorporating the beta-plane approximation gives a non-dimensional  $\beta^* = \frac{L^2 \beta}{U} = 514.286$ . The non-dimensional wind stress (applied as a body force here averaged over the depth of the domain) takes the same cosine of latitude profile with  $\tau_0^* = \frac{\tau_0 L}{(U^2 \rho_0 H)} = 163.2653$ . The Reynolds number is defined as  $Re = \frac{UL}{\nu} = 250$  (here the kinematic viscosity is  $140 \text{ m}^2 \text{ s}^{-1}$ ).

No-slip boundary conditions are applied to the lateral boundaries. The simulation starts from 'rest' and is driven by the wind stresses (46) on the free surface. The spin-up period is 0.3024 (equivalent to 100 *days*). The simulation period is [0.3024, 0.6048] (equivalent to [100, 200] *days*). The time step is  $3.78 \times 10^{-4}$  (or 3 *hrs*).

## 5.2. Optimisation of initial conditions

The control variables in the case tested are the initial conditions on the 100<sup>th</sup> day. The data assimilation experiments are designed by an identical twin technique. The 'true' flow state is generated by running the full forward model (ICOM) with a zonal wind forcing on the free surface. The pseudo-observational data is taken on days 125, 150 and 175 over the computational domain. The guess values of the initial conditions are given by the background

flow (here, taken from the 'true' flow fields on days either 112.5 or 137.5 in the experiments).

The POD bases are constructed by the snapshots which are obtained from the numerical solutions by forcing the full forward model with the background flow. 40 snapshots with 35 POD bases for the velocity field  $u$ ,  $v$ ,  $w$  and pressure are chosen which capture more than 99.5% of energy (calculated by the first 35 leading eigenvalues (7)).

To accurately represent geostrophic pressure its basis functions are split into two sets:  $\Phi_{pgu}$  and  $\Phi_{pgv}$  which are associated with the  $u$ - and  $v$ -velocity components. Furthermore the geostrophic pressure can be represented by a summation of the two sets of geostrophic basis functions. The geostrophic basis functions are calculated by solving the elliptic equations (the geostrophic balance equations) using a conjugate gradient iterative method. The geostrophic pressure has a quadratic finite element representation whilst linear finite element representations are used for the velocity components.

An adaptive mesh is adopted in the full model. The mesh for the full model adapts every 19 time steps with maximum and minimum mesh size of 0.2 and 0.001 (non-dimensional) respectively. To allow the same length of POD bases at the snapshots for both the reduced forward and adjoint models, a reference fixed mesh is chosen for the POD inversion (right panel in figure 7). To build up the snapshots, the solutions from the full forward model are interpolated from the adaptive mesh (left panel in figure 7) onto the reference fixed mesh.


### 5.3. Issue on adaptive reduced 4D-Var

As discussed in section 3.2, the POD reduced model based on the background flow can be improved by updating the POD bases. Here the POD bases are re-calculated when the value of cost function cannot be decreased by more than  $10^{-3}$  between the consecutive optimisation

iterations. The resulting control variables from the latest optimisation iteration are applied back to the full model to generate the new POD bases. The new POD bases then replace the previous ones to derive a new POD reduced order model [41].

However in the experiments the updated POD bases may be unacceptable if the control variables optimised during the latest POD iteration are far from true values. As an example, the initial guess controls (background flow) are taken from the true flow field on the 137.5<sup>th</sup> day. By fitting the numerical solution to observational data, the errors of the inverted modelled velocity field at the time levels ( $t = 125, 150, 175$  days) are reduced to small values (less than 8 in figure 1). The cost function (13) is reduced by 73 % of its original value during the optimisation procedure. However, the optimised initial controls during the current POD inversion iteration are far from what it is expected (inverted modelled velocity on the top panel in figure 1 and the true value in figure 5a). It is assumed that a local minimum was attained in this case. Obviously the optimised initial conditions cannot be used to update the POD bases. An appropriate choice of initial guess controls is needed to improve the adaptive reduced 4D-Var solution (here, taken from the true flow field on the 112.5<sup>th</sup> day, see section 5.4.2).

#### 5.4. Error estimation and POD results

The total error of the inverted modelled results comprises: (a) the integration error of the POD reduced model and the projection error (including the piecewise linear interpolation error when adaptive meshes are adopted); and (b) the error introduced by the optimised controls. 

*5.4.1. Error estimation for the POD reduced model* The error of the POD reduced model is split into the projection error and the error from the integration in the subspace (see [50]).

The two norm of the projection error can be calculated by [50]

$$\|e_{proj}\|_{L_2} = \sqrt{\sum_{i=M+1}^K \lambda_i} \quad (48)$$

where  $K$  is the number of snapshots and  $M$  is the number of POD bases (i.e. the subspace size). Here  $M = 35$ , the two norm of the projection error is therefore  $\sqrt{\sum_{i=M+1}^K \lambda_i} = 1.22, 1.3$  and 1.37 for the velocity components  $u, v$  and pressure  $p$  respectively.

To isolate the error of the reduced model, the POD reduced model is driven by the true controls rather than the optimised ones. The results from the POD reduced model are compared with those from the full model. The root mean square error (RMSE) between the POD velocity solution and the true one at the time level  $n$  is used to estimate the error of the POD model:

$$RMSE^n = \sqrt{\frac{\sum_{i=1}^{\mathcal{N}} (U_i^n - U_{0,i}^n)^2}{\mathcal{N}}} \quad (49)$$

where,  $U_i^n$  and  $U_{0,i}^n$  are the vectors containing the POD velocity components and true ones at the node  $i$  respectively,  $\mathcal{N}$  is the total number of nodes over the domain. The Root Mean Square Error (RMSE) during the simulation period is provided in figure 2. The dashed line represents the total reduced model error (the integration error plus project error) when 41 snapshots with 35 POD bases are chosen, while the solid line shows the integration error isolated by eliminating the projection error (the number of POD bases being the same as that of snapshots). The total reduced model error (RMSE) remains small (less than 2) when  $t \leq 150$  days, and increases to 4 at  $t = 175$  days. Since adaptive meshes are employed in the full model, the RSME shown in figure 2 includes the (piecewise linear) interpolation error as well. The absolute error between the POD solution and true flow state over the domain at the different time levels ( $t = 125, 150, 175$  days) is shown in figure 3. The maximum error is less than 8 (non-dimensional) during the first half simulation period and increases as the simulation time accrues.

5.4.2. *Optimised results and error estimation* In this case, the guess values of the initial conditions (background flows) are taken from the true flow state on the 112.5<sup>th</sup> day. The POD reduced forward and adjoint models are used to optimise the initial conditions. The POD bases are updated once during the optimisation procedure. The error between the inverted modelled velocity and the true value (figure 4) decreases by 10 – 50% in the larger part of flow after updating the snapshots (right panel). The cost function (taking into account only the error introduced by the optimised controls) is reduced by 50 % of its original value at the first POD inversion iteration. It is further reduced by 25 % at the second adaptive POD iteration, i.e., after the snapshots are updated.

The correlation *cor* defined below is also used to evaluate the quantity of the inversion simulation:

$$cor = \frac{cov_{12}}{\sigma_1\sigma_2},$$

where,

$$\begin{aligned} \sigma_1(\mathbf{x}) &= \sum_{n=1}^{N_t} (U^n(\mathbf{x}) - \bar{U}(\mathbf{x}))^2, \quad \sigma_2(\mathbf{x}) = \sum_{n=1}^{N_t} (U_o^n(\mathbf{x}) - \bar{U}_o(\mathbf{x}))^2, \\ cov_{12}(\mathbf{x}) &= \sum_{n=1}^{N_t} (U^n(\mathbf{x}) - \bar{U}(\mathbf{x}))(U_o^n(\mathbf{x}) - \bar{U}_o(\mathbf{x})), \end{aligned} \quad (50)$$

where,  $U^n$  and  $U_o^n$  are the vectors containing the optimal and true velocity components ( $u, v$ ) at the time level  $n$  over the domain respectively, their respective means over the simulation period are  $\bar{U}$  and  $\bar{U}_o$ ,  $n$  is the time level, and  $N_t$  is the total number of time levels,  $\mathbf{x} = (x, y, z)$ . The correlation between the true and modelled velocity in the case of running the POD model with the initial guess control (background flow) is low, mostly less than 0.5 over the domain (figure 6(b)). It is improved after the initial conditions are optimised (figure 6(c)), especially after updating the POD bases (at the second adaptive POD iteration, figure 6(d)). The correlation

between the inverted modelled and true velocity varies between 0.80-0.98 over the domain. The optimised velocity fields are drawn in figure 7 and exhibit an overall good agreement with the true ones while the POD bases are updated only once for the entire optimisation process.

### 5.5. Computational efficiency of reduced 4D-Var

In this test case, it takes 10 hours to run the full model, and 3 hours to run the reduced model. Thus running the reduced model results in a decrease of 70% of CPU time. As a consequence, the computational time required for reduced 4D-Var is reduced by a factor of  $3N$  (where  $N$  is the number of times needed to run the models until the optimality is satisfied). POD preconditions the minimisation process which results in less minimisation iterations.

It is also noted that 99% of CPU required for the reduced model is used to calculate the discretised matrix. The numerical technique developed in [1] is adopted to accelerate the POD inversion, that is, the matrices in the discretised POD forward and adjoint equations can be constructed by sets of time-independent sub-matrices (see equations (19) and (20)) prior to running the reduced forward and adjoint models. These sub-matrices remain the same until the POD bases are updated.

In total, the computer time required for the inverse simulation in this test case is given below:

- Running the full forward model to setup the snapshots and calculate the POD bases ( 10 hours);
- Calculating the time-independent sub-matrices in preparation for running both the reduced forward and adjoint models ( 2.5 hours);
- Running the reduced forward and adjoint models during the optimisation procedure ( 10 minutes, where the Polak-Ribiere nonlinear conjugate gradient approach for large-scale



unconstrained minimisation is employed).

## 6. Conclusion

The development of the POD reduced 4D-Var model for an adaptive mesh, non-hydrostatic finite element ocean model is presented here. Using the POD and Galerkin projection approaches, the reduced forward model is derived in a reduced subspace (details in [38]). Once the forward reduced model is available, the reduced order adjoint model can be directly obtained from the POD reduced forward model in the subspace, instead of the original forward model. The minimisation procedure is then carried out in the reduced space. The matrix for the discretised forward and adjoint models is constructed by a series of time-independent submatrices which remain unchanged until the POD bases are updated[1]. The reduced forward and adjoint models can thus be run repeatedly with negligible computational cost.

The performance of POD 4D-Var model is demonstrated by inverting for the initial conditions of a wind driven gyre in an idealised geometry. The correlation between the inverted modelled and true velocity is of the order of 80%- 98% over the majority of the domain. The cost function (taking into account only the error introduced by the optimised controls) is reduced by 50 % of its original value at the first POD inversion iteration. It is further reduced by 25 % after the snapshots are updated. The projection error is less than 1.22 for  $u$ , 1.3 for  $v$  and 1.37 for  $p$  (non-dimensional unit).

In general, the advantages of the POD reduced inverse model developed here over existing POD approaches are the ability:

- to implement the reduced adjoint model from the discretised forward model easily; ;

- to use dynamically adaptive meshes in the reduced 4D-Var; ;
- to significantly reduce the computation cost for 4D-Var by carrying out the optimisation in the reduced space since POD reduces condition number of Hessian of cost function [51];
- to accurately represent the geostrophic balance by two sets of POD bases for the velocity components  $u$  and  $v$ .

#### Acknowledgements

This work was carried out under funding from NE/C52101X/1 and from the UK EPSRC GR/60898. Prof. I. M. Navon would like to acknowledge support of NSF grants ATM-0201808 and CCF-0635162 and NSF Award ID 0635162.

#### REFERENCES

1. F. Fang, C. C. Pain, I. M. Navon, M. D. Piggott, G. J. Gorman, P. Allison, and A. J. H. Goddard. Reduced order modelling of an adaptive mesh ocean model. *International Journal for Numerical Methods in Fluids* 2007; accepted with revisions.
2. R. W. Lardner. Optimal control of open boundary conditions for a numerical tidal model. *Computer Methods in Applied Mechanics and Engineering* 1993; 102:367–87.
3. A. F. Bennett, Chua B. S., D. E. Harrison, and M. J. McPhaden. Generalized inversion of Tropical Atmosphere-Ocean (TAO) data and a coupled model of the tropical Pacific. Part II: The 1995-96 La Nina and 1997-98 El Nino. *J. Climate* 2000; 13(15):2770–2785.
4. Y. Leredde, J. L. Devenon, and I. Dekeyser. Turbulent viscosity optimized by data assimilation. *Ann. Geophysicae* 1999; 17:1463–1477.
5. A. M. Moore. Data assimilation in a quasi-geostrophic open-ocean model of the Gulf Stream region using the adjoint method. *J. Phys. Oceanogr.* 1991; 21(3):398–427.
6. W. C. Thacker and R. B. Long. Fitting dynamics to data. *J. Geophys. Res.* 1988; 93(C2):1227–1240.

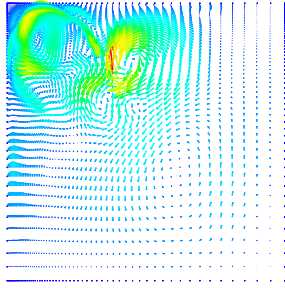
7. E. Tziperman, W. C. Thacker, R. B. Long, S. M. Hwang, and S. R. Rintoul. Oceanic data analysis using a general circulation model. Part II: A North Atlantic model. *J. Phys. Oceanogr.* 1992; 22(12):1458–1485.
8. J. Zhu, M. Kamachi, and D. Wang. Estimation of air-sea heat flux from ocean measurements: An ill-posed problem. *Journal of Geophysical Research*, 107(10):3159–3159, 2002.
9. X. Zou and Y. H. Kuo. Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Weath. Rev.* 1996 124(12):2859–2882.
10. M. D. Gunzburger. *Perspectives in Flow Control and Optimization*. SIAM, 2003.
11. O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation (I). Theory. *Q. J. Roy. Meteor. Soc.* 1987; 113(478):1311–1328.
12. B. F. Farrell and A. M. Moore. An adjoint method for obtaining the most rapidly growing perturbation to oceanic flows. *J. Phys. Oceanogr.* 1992; 22(4):338–349.
13. D. G. Cacuci, M. I. Bujor, and I. M. Navon. *Sensitivity and Uncertainty Analysis: Applications to Large-scale Systems Vol 2*. CRC, 2005.
14. A. M. Moore, N. S. Cooper, and D. L. T. Anderson. Initialization and data assimilation in models of the Indian Ocean. *J. Phys. Oceanogr.* 1987; 17:1965–1977.
15. D. N. Daescu and I. M. Navon. Adaptive observations in the context of 4D-Var data assimilation. *Meteorol. Atmos. Phys.* 2004; 85(4):205–226.
16. F. X. Le Dimet and I. M. Navon. Variational and optimization methods in meteorology: A review, 1988. Technical Report: Early review on variational data assimilation, SCRI report No 144.
17. Y. Q. Zhu and I. M. Navon. Impact of parameter estimation on the performance of the FSU global spectral model using its full-physics adjoint. *Mon. Weath. Rev.*, 127(7):1497–1517, 1999.
18. P. Courtier, J. N. Thepaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteor. Soc.* 1994; 120(519):1367–1387.
19. J. Derber and A. Rosati. A global oceanic data assimilation system. *Meteorol. Atmos. Phys.* 2003; 85(4):205–226.
20. A. T. Weaver, J. Vialard, D. L. T. Anderson, and P. Delecluse. Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. Part II: physical validation. *Mon. Wea. Rev.*, 131:1379–1395, 2003.
21. M. Wenzel, J. Schröter, and D. Olbers. The annual cycle of the global ocean circulation as determined by 4D VAR data assimilation. *Prog. Oceanogr.* 2001; 48(1):73–119.
22. P. Courtier, J. N. Thepaut, and A. Hollingsworth. A strategy for operational implementation of 4d-var,

- using an incremental approach. *Quarterly Journal of the Royal Meteorological Society* 1994; 120:1367–138.
23. Y. Tremolet. Incremental 4D-Var convergence study. *Tellus Series A Dynamic Meteorology and Oceanography* 2007; 59(5): 706-718.
  24. Y. Tremolet. Diagnostics of linear and incremental approximations in 4D-Var. *Quarterly Journal of the Royal Meteorological Society* 2004 PartB; 130(601): 2233-2251.
  25. C. Robert, S. Durbiano, E. Blayo, J. Verron, J. Blum, and F. X. le Dimet. A reduced-order strategy for 4D-VAR data assimilation. *Journal of Marine Systems* 2005; 57(1-2):70–82.
  26. E. Blayo, J. Blum, and J. Verron. *Assimilation Variationnelle de Donnees en Oceanographie et reduction de la dimension de l'espace de controle In Equations aux Derivees partielles et Applications*. Gauthiers Villars, 1998.
  27. I. Hoteit and A. Kohl. Efficiency of reduced-order, time-dependent adjoint data assimilation approaches. *J. of Oceanography* 2006; 62(4):539–550.
  28. C. Robert, E. Blayo, J. Verron, J. Blum, and F. X. le Dimet. Reduced-order 4d-var: A preconditioner for the incremental 4d-var data assimilation method. *Geophysical Research Letters* 2006; 33, L18609:1–4.
  29. P. T. M. Vermeulen and A. W. Heemink. Model-reduced variational data assimilation. *Monthly Weather Review* 2006; 134:2888–2899.
  30. P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge, UK: Cambridge University Press, 1998.
  31. Z. Luo, J. Chen, J. Zhu, R. Wang, and I. M. Navon. An optimizing reduced order FDS for the tropical Pacific ocean reduced gravity model. *International Journal for Numerical Methods in Fluids*, 55, 143-161, 2007.
  32. Z. Luo, J. Zhu, R. Wang, and I. M. Navon. Proper orthogonal decomposition approach and error estimation of mixed finite element methods for the tropical Pacific ocean reduced gravity model. *Computer methods in Applied Mechanics and Engineering* 2007; 196/41-44:4184–4195.
  33. K. Fukunaga. *Introduction to Statistical Recognition Second Edition (Computer Science and Scientific Computing Series)*. Boston: Academic Press, 1990.
  34. N. Aubry, P. Holmes, and J. L. Lumley. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Dynamics* 1988; 192:115–173.
  35. K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal* 2002; 40(11):2323–2330.
  36. M. Kirby and L. Sirovich. Application of the Karhunen-Loève procedure for the characterization of human

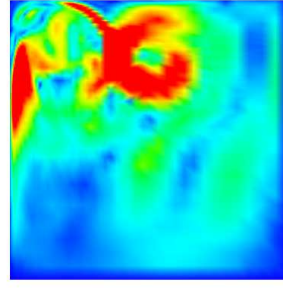
- faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1990; 12(1):103–108.
37. K. Willcox, O. Ghattas, B. van Bloemen Waanders, and B. Bader. An optimization framework for goal-oriented, model-based reduction of large-scale systems. *44th IEEE Conference on Decision and Control and European Control Conference*, 2005.
  38. T. Bui-Thanh, K. Willcox, O. Ghattas, and B. van Bloemen Waanders. Goal-oriented, model-constrained optimization for reduction of large-scale systems. *J. of Computational Physics* 2007; 224(2):880–896.
  39. Y. Cao, J. Zhu, I. M. Navon, and Z. Luo. A reduced order approach to four-dimensional variational data assimilation using proper orthogonal decomposition. *International Journal for Numerical Methods in Fluids* 2006; 53(10):1571–1583.
  40. D. N. Daescu and I. M. Navon. A dual-weighted approach to order reduction in 4d-var data assimilation. *In press, Monthly Weather Review* 2007.
  41. S. S. Ravindran. Adaptive reduced-order controllers for a thermal flow system using proper orthogonal decomposition. *SIAM Journal on Scientific Computing* 2002; 23(6):1924–1942.
  42. S. S. Ravindran. Reduced-order controllers for control of flow past an airfoil. *International Journal for Numerical Methods in Fluids* 2002; 50(5):531–554.
  43. M. Fahl. Trust-region method for flow control based on reduced order modelling. *Thesis dem Fachbereich IV der Universitat Trier*, 2000; 50(5):531–554.
  44. R. Ford, C. C. Pain, M. D. Piggott, A. J. H. Goddard, C. R. E. de Oliveira, and A. P. Umbleby. A nonhydrostatic finite-element model for three-dimensional stratified oceanic flows. Part I: Model formulation. *Mon. Weath. Rev.* 2004; 132(12):2816–2831.
  45. C. C. Pain, A. P. Umpleby, C. R. E. de Oliveira, and A. J. H. Goddard. Tetrahedral mesh optimisation and adaptivity for steady-state and transient finite element calculations. *Comput. Methods Appl. Mech. Engrg.* 2001; 190:3771–3796.
  46. M. D. Piggott, C. C. Pain, G. J. Gorman, P. W. Power, and A. J. H. Goddard.  $h$ ,  $r$ , and  $hr$  adaptivity with applications in numerical ocean modelling. *Ocean Modell.* 2005; 10(1–2):95–113.
  47. L. A. Freitag and C. Ollivier-Gooch. Tetrahedral mesh improvement using swapping and smoothing. *Internat. J. Numer. Methods Engrg.* 1997; 40:3979–4002.
  48. G. C. Buscaglia and E. A. Dari. Anisotropic mesh optimization and its application in adaptivity. *J. Numer. Methods Engrg.* 1997; 40:4119–4136.
  49. P. W. Power, C. C. Pain, M. D. Piggott, G. J. Gorman, F. Fang, D. P. Marshall, and A. J. H. Goddard. Adjoint goal-based error norms for adaptive mesh ocean modelling. *Ocean Modell.* 2006; 15:3 – 38,

doi:10.1016/j.ocemod.2006.05.001.

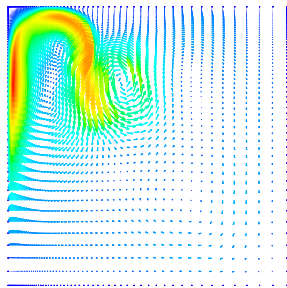
50. C. Homescu, L.R. Petzold, and R. Serban. Error estimation for reduced-order models of dynamical systems. *Siam Review* 2007; 49(2):277–299.
51. D. N. Daescu and I. M. Navon. Efficiency of a POD-based reduced second order adjoint model in 4-D VAR data assimilation. *International Journal for Numerical Methods in Fluids* 2007; 53,985-1004.



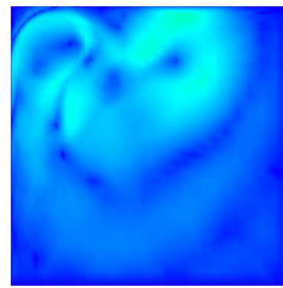
(a)



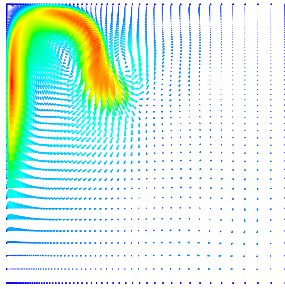
(b)



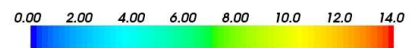
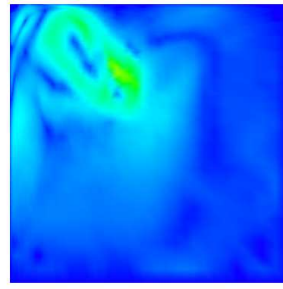
(c)



(d)



(e)



(f)

Figure 1. Optimisation problem (local minimum) introduced by an inappropriate choice of initial guess controls (the initial guess controls is taken from the true flow field on the 137.5<sup>th</sup> day). Velocity results: (a) (b) at the initial time level; (c)(d) at time level  $t = 125$  days; (e) (f) at time level  $t = 175$  days.

Left panel: inverted modelled velocity field; right panel: error for the velocity field.

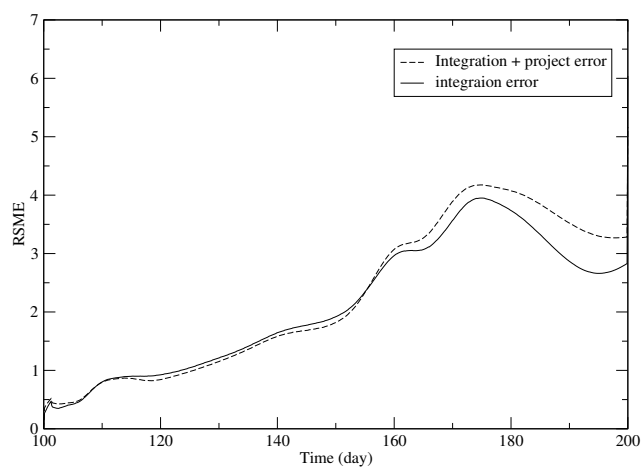


Figure 2. Error of the POD reduced forward model. Solid line: integration error (41 snapshots and 41 POD bases); dashed line: total reduced model error (integration error plus project error, 41 snapshots and 35 POD bases).



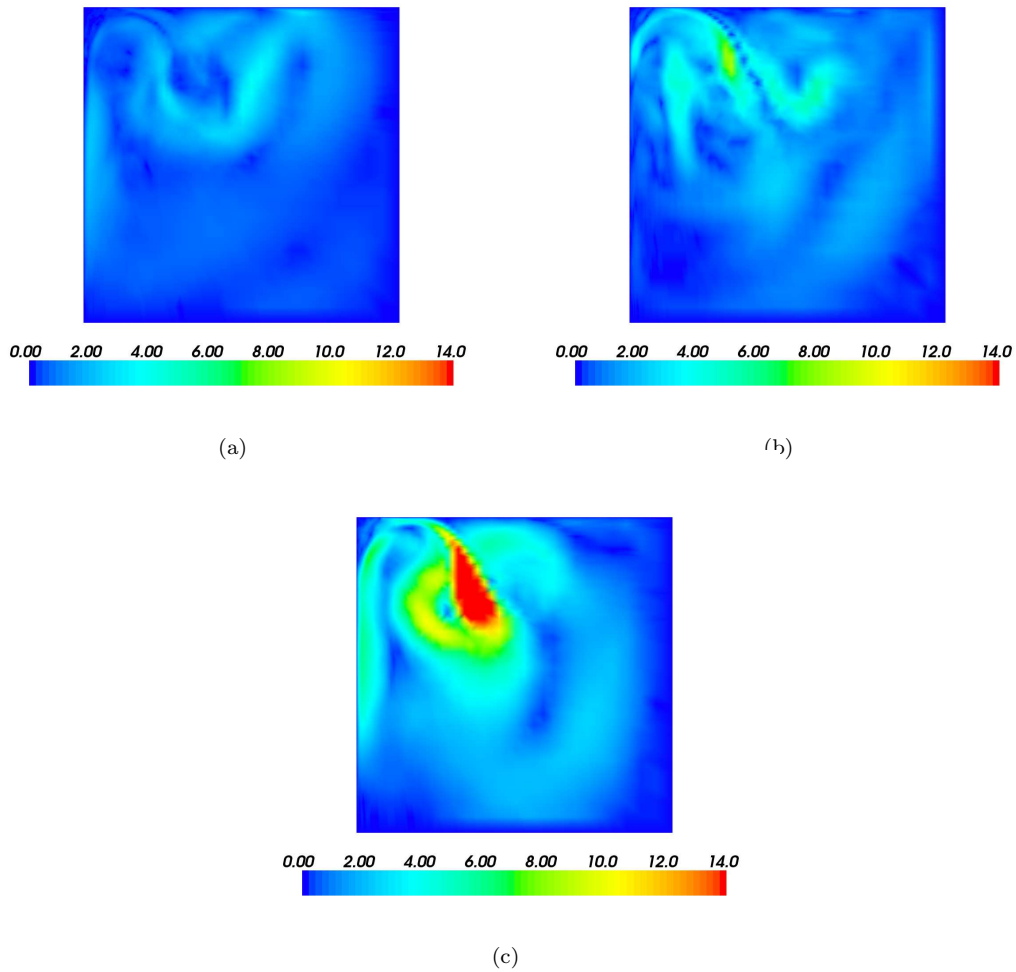


Figure 3. Error in the velocity field from the POD reduced model with the true initial conditions at time levels (a)  $t = 125$  days (top panel) (b)  $t = 150$  days (middle panel) (c)  $t = 175$  days (bottom panel).

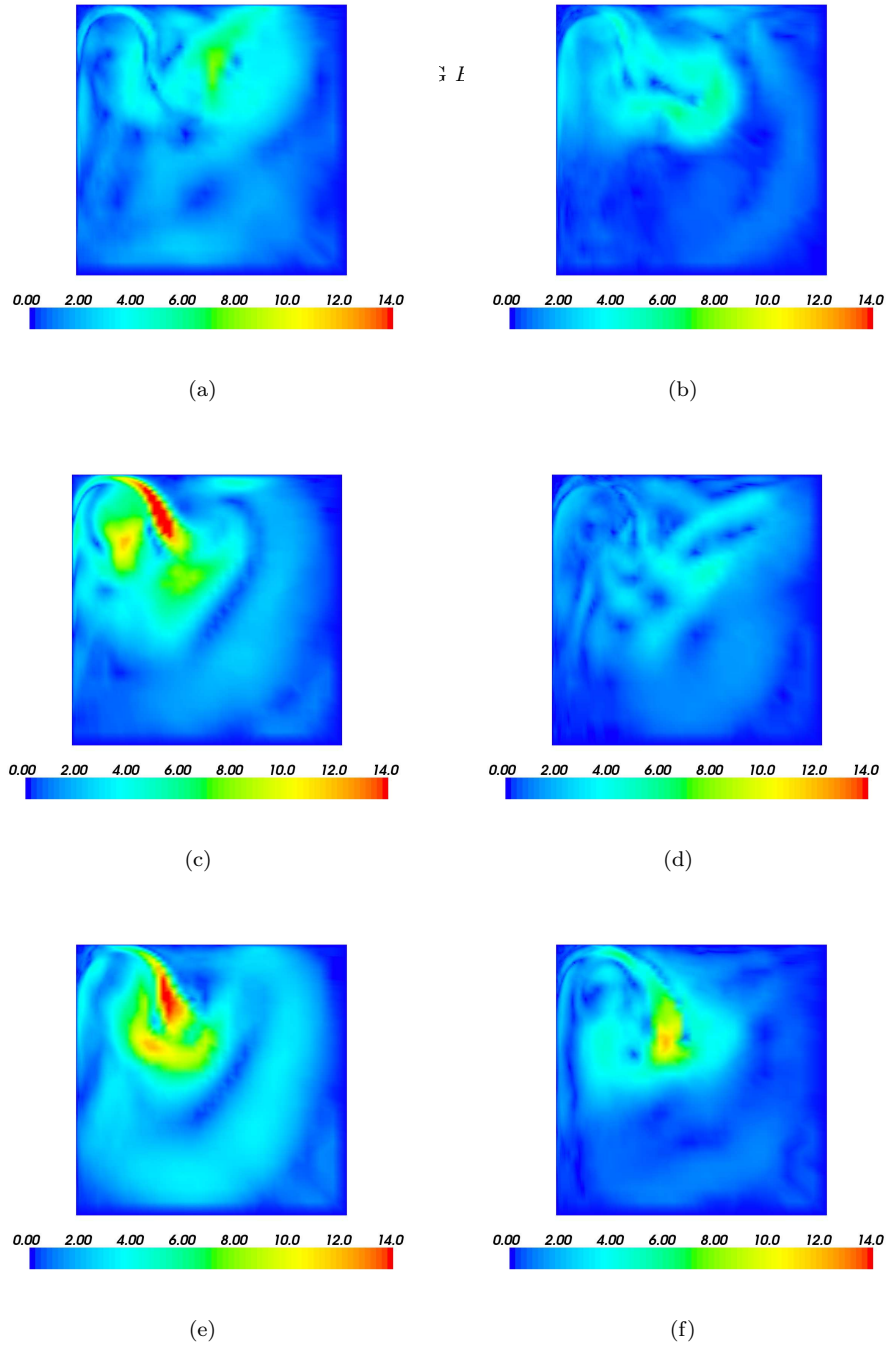


Figure 4. Error in the velocity of the POD forward model driven by the optimised initial conditions at time levels: (a)(b)  $t = 125$  days (the top panel); (c)(d)  $t = 150$  days (the middle panel); (e)(f)  $t = 175$  days (the bottom panel). Left panel: the first POD iteration; right panel: the second POD iteration (after updating the snapshots).

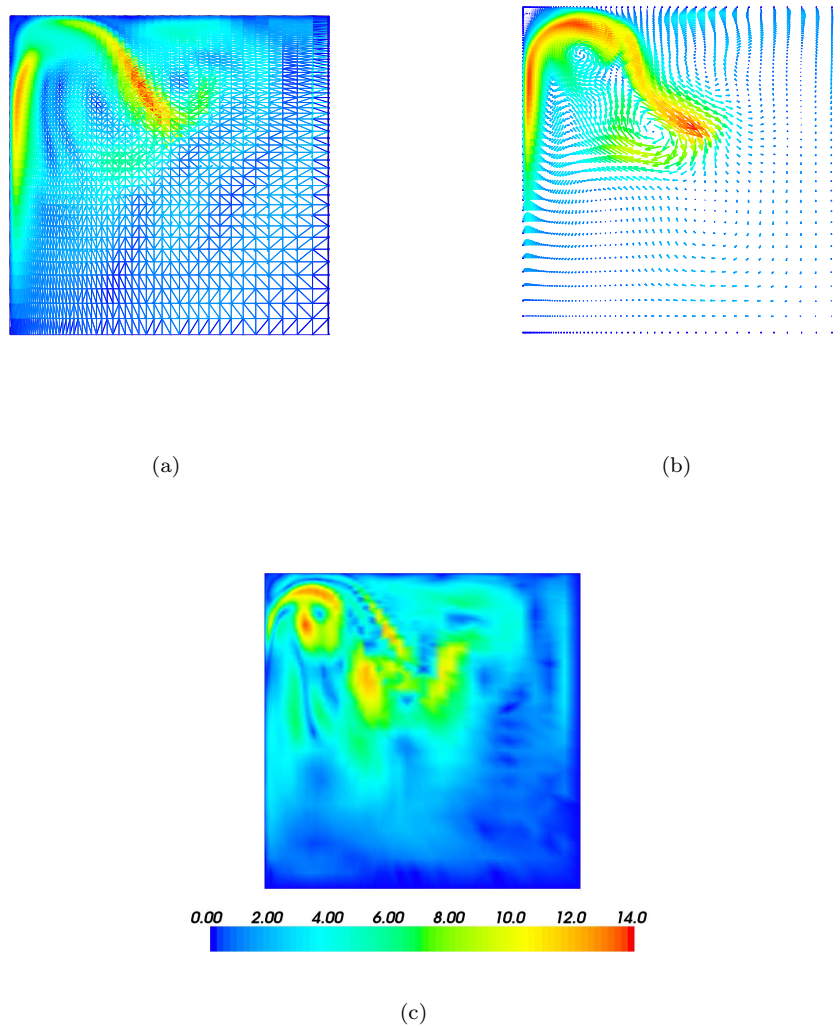


Figure 5. The comparison between the optimised and true initial velocity conditions. (a) the true initial conditions; (b) the optimised initial conditions; (c) the error between the optimised and true initial conditions.

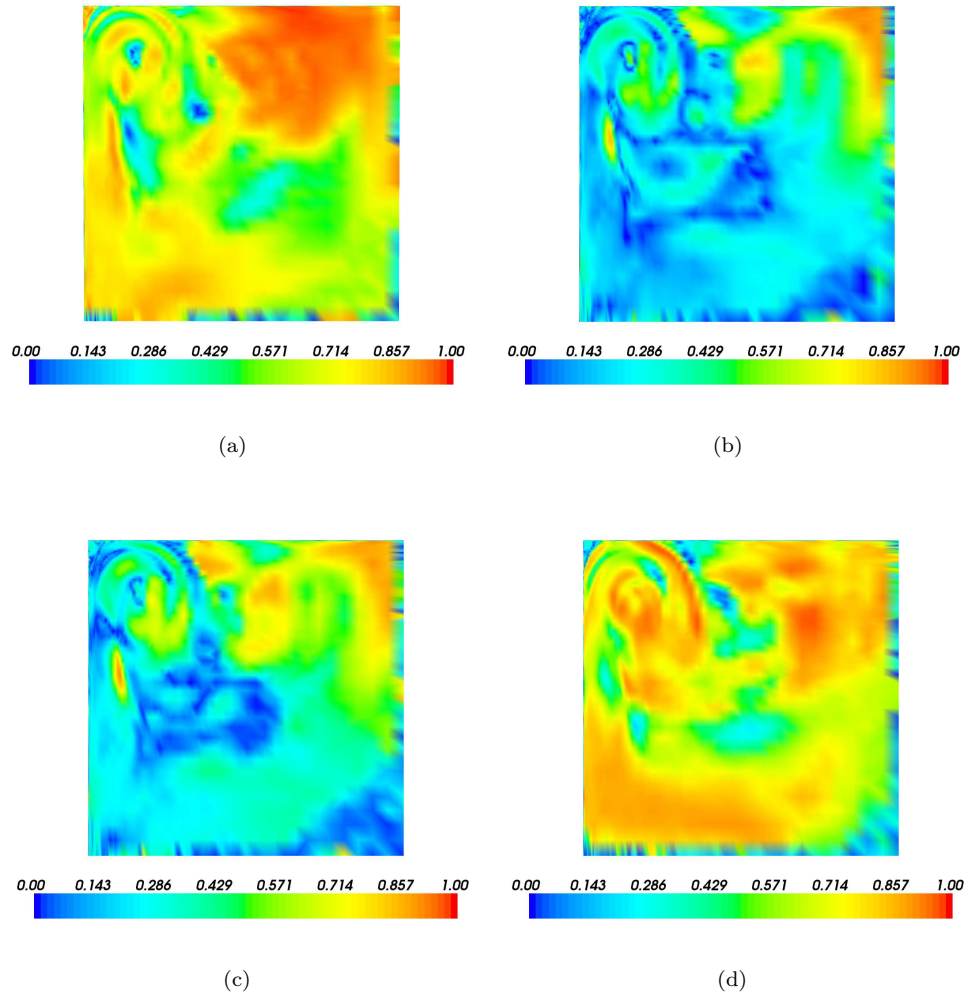
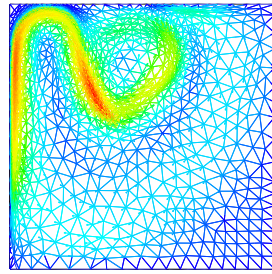
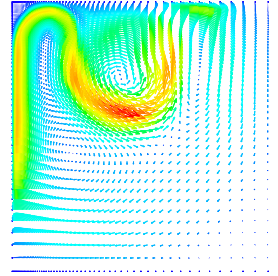


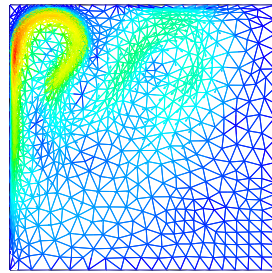
Figure 6. Correlation between the inverted modelled velocity and true one. The inverted modelled velocity is obtained from the POD model driven by (a) the true control; (b) the initial guess controls; (c) the optimised controls at the first POD iteration; (d) the optimised controls after the second POD iteration, i.e., after updating the POD bases.



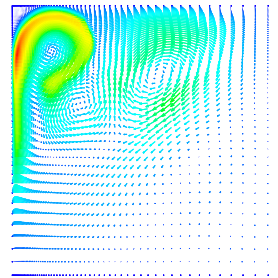
(a)



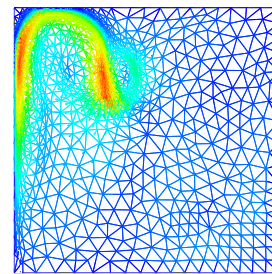
(b)



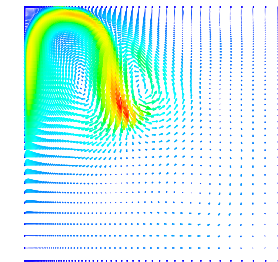
(c)



(d)



(e)



(f)

Figure 7. Comparison between the true velocity field and that from the POD reduced model (driven by the optimised initial conditions) at the time levels: (a)(b)  $t = 125$  days; (c)(d)  $t = 150$  days; (e)(f)  $t = 175$  days. Left panel: the true velocity field; right panel: the optimised velocity field)