# Non-linear forcing singular vector of a two-dimensional quasi-geostrophic model

*By* WANSUO DUAN[1]* and FEIFAN ZHOU[2], [1]*LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*; [2]*Laboratory of Cloud-Precipitation Physics and Severe Storms, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

### ABSTRACT

We propose a non-linear forcing singular vector (NFSV) approach to infer the effect of non-linearity on the predictability associated with model errors. The NFSV is a generalisation of the forcing singular vector (FSV) to non-linear fields and acts as a tendency perturbation that results in a significantly large perturbation growth. In predictability studies, the NFSV, as a tendency error, may provide useful information about model errors that cause severe prediction uncertainties. In this article, a two-dimensional quasi-geostrophic (QG) model is used to study NFSVs and make a comparison between NFSVs and FSVs. We choose two basic flows: the first is a zonal steady flow (Ref-1), and the second is a meridional steady flow (Ref-2). The results demonstrate that the corresponding NFSVs contain a phase where the stream function tends to be contracted around regions of strong velocity shear. Furthermore, the NFSVs for the Ref-1 tend to have a meridional asymmetric spatial structure. Due to the absence of non-linearity, FSVs tend to have a larger spatial extension than NFSVs; in particular, the FSVs for the Ref-1 are almost symmetric in the stream function component. The prediction errors caused by FSVs in the non-linear QG model are generally smaller than those caused by FSVs in the linearised QG model; therefore, the non-linearity in the QG model would significantly saturate the perturbation growth. Nevertheless, the prediction errors caused by NFSVs (especially for the Ref-1) in the non-linear QG model are larger than those caused by FSVs, which further implies that the tendency errors of NFSV structures tend to reduce the damping effect of the non-linearity on the perturbation growth and are more applicable than those of FSV structures to describing the optimal mode of the model errors. The differences between NFSVs and FSVs demonstrate the usefulness of NFSVs in revealing the effects of non-linearity on predictability. The NFSV may be a useful non-linear technique for exploring the predictability problems introduced by model errors.

*Keywords: weather, climate, predictability, model error, optimal perturbation*

## 1. Introduction

One of the central problems in atmospheric and oceanic sciences is weather and climate predictability, in which estimating the uncertainties associated with the predictions is very important. Tennekes (1991) proclaimed that no forecast was complete without an estimate of the prediction error. Since this perspective was first proposed by Thompson (1957), operational weather forecasting has progressed to making explicit attempts to quantify the evolution of the initial uncertainty during each forecast

(Palmer et al., 1992; Toth and Kalnay, 1997; Mu et al., 2003). Some understanding of the predictability for tropical ocean-atmosphere system has also been gained by studying the growth of errors and uncertainties during forecasting (Moore and Kleeman, 1996; Samelson and Tziperman, 2001; Mu et al., 2007a, 2007b).

Prediction uncertainties are generally caused by initial errors and model errors. To study the roles of initial errors and model errors in yielding prediction uncertainties, Lorenz (1975) classified two types of predictability problems: the first is related to initial errors and assumes a perfect model, whereas the second is associated with model errors and assumes a perfect initial field. The former has been largely explored, resulting in the proposal and introduction of many theories and methods in which

*Corresponding author.
email: duanws@lasg.iap.ac.cn

**1**

optimal methods are important for estimating the predictability limit of weather and climate (Lorenz, 1965; Toth and Kalnay, 1997; Mu et al., 2003; Riviere et al., 2008). In this context, the singular vector (SV) approach was first introduced in meteorology (Lorenz, 1965; Farrell, 1989). Considering the limitations of the linear theory of SV, Mu et al. (2003) proposed the conditional non-linear optimal perturbation approach (CNOP) to search for the optimal initial perturbation for a given constraint; the competing aspect of this approach is that it sufficiently considers the effect of non-linearity. CNOP has been applied to weather and climate predictability studies (Duan et al., 2004; Mu et al., 2004; Duan and Mu, 2006; Mu et al., 2007a, 2007b; Duan et al., 2008; Terwisscha and Dijkstra, 2008; Duan et al., 2009; Jiang and Mu, 2009; Mu et al., 2009; Wu and Mu, 2009). Recently, Riviere et al. (2009) demonstrated an extension of the CNOP approach and used it to estimate the predictability of atmospheric moisture processes in an attempt to reveal the effects of non-linear processes. Oortwijn and Barkmeijer (1995) and Barkmeijer (1996) also realised the limitations of the linear technique and considered non-linear effects using an iterative procedure. The bred vector (Toth and Kalnay, 1997) is another important non-linear optimal method that was generalised from the Lyapunov vector approach and has been used to investigate the first type of climate predictability problem (Cai et al., 2003). All of these theories and methods have played important roles in guiding scientists to develop and improve numerical models and propose innovative ideas to increase the accuracy of weather and climate forecasts (Houtekamer and Deroma, 1995; Xue et al., 1997; Thompson, 1998; Hamill et al., 2000; Mu and Zhang, 2006; Mu et al., 2009).

The existing numerical models cannot yet accurately describe atmospheric and oceanic motions. Uncertainties in the parameterisation of subgrid-scale physical processes, external forcing, and numerical discrete schemes, etc., may result in model errors. Substantial progress has been achieved in improving the quality of forecasts during the last few decades (Simmons and Hollingsworth, 2002); however, these factors will always limit forecast accuracy and should be considered during forecasting (Houtekamer et al., 1996; Buizza et al., 1999; Mylne et al., 2002). Studies of model errors are related to the second type of predictability problem (Lorenz, 1975). It is important to estimate the predictability limit caused by model errors. Mu et al. (2010) extended the CNOP approach to investigate the optimal mode of model parameter uncertainties and estimated the predictability limit introduced by model parameter errors. Roads (1987) used constant tendency errors to approximate model uncertainties. A procedure has been developed to compute the time-independent

forcing of model tendencies based on observed fields (Roads, 1987). By writing the model in the form

$$\frac{\partial \boldsymbol{U}}{\partial t} = F(\boldsymbol{U}(\boldsymbol{x}; t)) + \boldsymbol{f}, \qquad (1.1)$$

where $F(\boldsymbol{U}(\boldsymbol{x}; t))$ is the total model tendency, the forcing $\boldsymbol{f}$ is computed by inserting the observed fields for $x$. The forcing term $\boldsymbol{f}$ as defined by Roads' procedure can be interpreted as a crude method to account for processes that are not explicitly or correctly described by the model equations (Roads, 1987; Barkmeijer et al., 2003). Following this reasoning, Moore and Kleeman (1999) developed an approach, called the (linear) stochastic optimal approach, to study the ENSO predictability limit caused by model errors. However, Barkmeijer et al. (2003) thought that the stochastic optimal approach was not feasible in a realistic high-dimensional numerical model because of the explicit matrix computation of the linear model propagator and its adjoint. To compensate for this limitation of the stochastic optimal approach, Barkmeijer et al. (2003) proposed the concept of the (linear) forcing singular vector (FSV), which is constant in time but represents the tendency perturbation leading to a significantly large perturbation growth in a linearised model during a given forecast period (Barkmeijer et al., 2003). D'Andrea and Vautard (2000) studied similar structures as a way to reduce the systematic error in a quasi-geostrophic (QG) model. Farrell and Ioannou (2005) further determined the optimal set of the distributed deterministic and stochastic forcings in the forecast and observer systems over a chosen time interval, also based on a linear system.

The motions of the atmosphere and ocean are generally governed by complex non-linear systems. The FSV is derived by the linear approximation of a non-linear model (the linearised model), which raises concerns about the validity of the linearised model. That is, how can the linearised model approximation to its non-linear counterpart be validated in advance? Some papers have attempted to address this concern, but very few satisfying answers have been found (Lacarra and Talagrand, 1988; Tanguay et al., 1995). Therefore, it is desirable and often necessary to address non-linear models rather than their linear approximations in numerical weather and climate prediction. Therefore, this article investigates the FSV approach within the frame of a non-linear model.

This article is organised into five sections. In section 2, a non-linear forcing singular vector (NFSV) is proposed. The NFSV calculation is discussed in section 3. In section 4, we study the NFSVs of a two-dimensional QG model and investigate the differences between NFSVs and FSVs and illustrate the usefulness of NFSVs in revealing the effects of non-linearity. Finally, the results obtained in this study are summarised, and the

physical interpretations of the NFSV are discussed in section 5.

## 2. Non-linear forcing singular vector

We write the evolution equations for the state vector $U$, which may represent the surface current, thermocline depth, and sea surface temperature, etc., as follows:

$$\begin{cases} \frac{\partial U}{\partial t} = F(U(x,t)), \\ U|_{t=0} = U_0, \end{cases} \text{ in } \Omega \times [0,\tau] \qquad (2.1)$$

where $U(x,t) = (U_1(x,t), U_2(x,t), \cdots, U_n(x,t))$; $U_0$ is the initial state; $(x,t) \in \Omega \times [0,\tau]$, in which $\Omega$ is a domain in $R^n$, $x = (x_1, x_2, \cdots, x_n)$ and $t$ is the time; and $\tau < +\infty$ is the final time of the state variables evolution. $F$ is a non-linear operator. Assuming that the dynamical system equation [eq. (2.1)] and the initial state are known exactly, the future state can be determined by integrating eq. (2.1) using the appropriate initial conditions. The solution to eq. (2.1) for the state vector $U$ at time $\tau$ is given by

$$U(x,\tau) = M_\tau(U_0). \qquad (2.2)$$

Here, $M_\tau$ is the propagator of eq. (2.1) and 'propagates' the initial value to the prediction time $\tau$ as described by eq. (2.2).

In realistic predictions, forecast systems generally contain both initial errors and model errors. In predictability studies, initial errors are often measured by perturbations superimposed on the initial fields of the numerical model. For model errors, Roads (1987), Moore and Kleeman (1999), and Barkmeijer et al. (2003) used a tendency perturbation to investigate the effects of model errors on the prediction results. Furthermore, Barkmeijer et al. (2003) considered a constant tendency perturbation and developed the FSV approach to study the predictability problems associated with model errors.

We describe forecast models with initial perturbations and constant tendency perturbations as in eq. (2.3).

$$\begin{cases} \frac{\partial (U+u)}{\partial t} = F(U+u) + f(x), \\ U+u|_{t=0} = U_0 + u_0. \end{cases} \qquad (2.3)$$

We use $M_\tau(f)$ to denote the propagator of eq. (2.3). In fact, $M_\tau(f)$ with $f=0$ is the same as $M_\tau$ in eq. (2.2), i.e. the propagator of eq. (2.1). Then, we obtain

$$M_\tau(f)(U_0 + u_0) = U(x,\tau) + u_f^I(x,\tau), \qquad (2.4)$$

where $U(x,\tau) = M_\tau(0)(U_0) = M_\tau(U_0)$. From eqs. (2.2) and (2.4), it is easily derived that $u_f^I(x,\tau) = M_\tau(f)(U_0+u_0) - M_\tau(U_0)$, which describes the departure from the state $U(x,\tau)$ [i.e. $M_\tau(U_0)$] caused by the initial perturbation $u_0$ and constant tendency perturbation $f$. In the scenario describing the first type of predictability problem, the model is thought to be perfect. In this case, the constant tendency perturbations (as model systematic errors) are equal to zero (i.e. $f=0$), and the situation in which $M_\tau(0)(U_0+u_0) = U(x,\tau)+u^I(x,t)$, where $u^I(x,t)$ represents the evolution of initial errors $u_0$, is considered. For the second type of predictability problem, the initial fields are assumed perfect (i.e. the initial errors $u_0=0$), and $M_\tau(f)(U_0) = U(x,\tau)+u_f(x,\tau)$ is of interest. Then, $u_f(x,\tau)$ describes the departure from the state $U(x,\tau)$ caused by the constant tendency perturbation (i.e. constant tendency error) $f$, which may describe a type of model systematic error.

Barkmeijer et al. (2003) considered the linearised version of a non-linear model. The perturbation equation for the non-linear model eq. (2.1) can be obtained by subtracting eq. (2.1) from eq. (2.3). By omitting the non-linear term, a linearised perturbation equation is obtained:

$$\begin{cases} \frac{\partial u}{\partial t} = \mathbf{F}(U)u + f(x), \\ u|_{t=0} = u_0, \end{cases} \qquad (2.5)$$

where $u_0$, $u$, and $U$ have the same meanings as in eq. (2.3), and $\mathbf{F}(U)$ is the Jacobian of the non-linear operator $F$ with respect to the reference state $U(x,t)$. With the constant tendency perturbation $f(x)$, the FSV $f^*$ can be obtained by solving the linear optimisation problem of eq. (2.6).

$$\lambda(f^*) = \max_f \frac{\|\mathbf{M}_\tau(f)(0)\|}{\|f\|}, \qquad (2.6)$$

where the norm $\|\cdot\|$ is described by the inner product $< \cdot >$, and $\mathbf{M}_\tau$ is the propagator of eq. (2.5) and is equivalent to the tangent linear operator of the non-linear propagator $M_\tau(f)$ of eq. (2.3).

It is clear from eq. (2.6) that the FSV is derived from a linear model. Next, we generalise the FSV to a non-linear field and propose a non-linear technique of FSV to determine the effects of non-linearity on the predictability limit caused by the model uncertainties. As mentioned in the introduction, Mu et al. (2010) extended the CNOP approach to investigate the model parameter perturbations (CNOP-P) that cause the largest prediction error. It is inferred that if we express the constant tendency perturbation as external parameters that display a certain spatial structure, we can use the CNOP-P technique to derive the non-linear FSV.

For a chosen measurement, a tendency perturbation $f_\delta$ is defined as a non-linear FSV (NFSV) if and only if

$$J_\delta(f_\delta) = \max_{\|f\|_a \leq \delta} J(f), \qquad (2.7)$$

where

$$J(f) = \|M_\tau(f)(U_0) - M_\tau(0)(U_0)\|_b, \qquad (2.8)$$

$M_\tau(\boldsymbol{f})$ is the propagator of the non-linear model, eq. (2.3), and $\|\boldsymbol{f}\|_a \leq \delta$, which is defined by the norm $\|\cdot\|_a$, is the constraint condition of the tendency perturbation $\boldsymbol{f}$. The objective function $J$ with the norm $\|\cdot\|_b$ measures the magnitude of the departure from the reference state $\mathrm{M}_\tau(0)(\boldsymbol{U}_0)$ caused by the tendency perturbation $\boldsymbol{f}$. We note that the norms $\|\cdot\|_a$ and $\|\cdot\|_b$ represent different norms. In some situations, the norm $\|\cdot\|_a$ can be the same as $\|\cdot\|_b$ depending on the physical problem being investigated. As mentioned in the introduction, Roads (1987) proposed a procedure to compute a time-independent forcing of model tendencies based on observation fields, and this motivates a similar approach to compute the allowable magnitudes of the constant tendency perturbation before determining the constraint condition amplitude (i.e. $\delta$). Because this article is focused on methodology, to illustrate the NFSV approach, we experimentally choose the amplitudes of the constraint conditions to calculate the NFSVs.

Eq. (2.7) is a constrained maximisation problem. By solving this optimisation problem, one can obtain the NFSV, $\boldsymbol{f}_\delta$. It is easily shown from eqs. (2.7) and (2.8) that the function $J$, which is associated with the NFSV, measures the magnitude of the departure of the reference state $M_\tau(0)(\boldsymbol{U}_0)$ caused by the constant tendency perturbation. Furthermore, the function $J_\delta$ requires us to find the constant tendency perturbation that causes the greatest departure from the reference state $M_\tau(0)(\boldsymbol{U}_0)$ at a future time $\tau$. Then, the constant tendency perturbation of the NFSV pattern introduces the greatest departure from the reference state $M_\tau(0)(\boldsymbol{U}_0)$ at a future time $\tau$ in terms of the chosen measurement norm. In other words, the NFSV induces the largest perturbation growth at the given future time $\tau$. In predictability studies, the NFSV $\boldsymbol{f}_\delta$, as a constant tendency error, may represent a type of model systematic error that causes a significantly large prediction error at the prediction time.

We notice that the NFSV defined by eq. (2.7) is established on the assumption that the tendency perturbation does not change with time. Therefore, in predictability studies, the tendency errors are assumed to be constant in time. However, realistic tendency errors may be time-dependent; thus, it is necessary to generalise the constant NFSV to be time-dependent. Actually, using the constant NFSV, we can derive a time-dependent NFSV. For the time steps $t_0$, $t_1$, $t_2$, ..., $t_k$, a time-dependent NFSV $\boldsymbol{f}_{\delta,t_k-t_0} = (\boldsymbol{f}_{\delta_0,t_0}, \boldsymbol{f}_{\delta_1,t_1}, \boldsymbol{f}_{\delta_2,t_2}, ..., \boldsymbol{f}_{\delta_k,t_k})$ can be obtained by the following optimisation problem:

$$J(\boldsymbol{f}_{\delta_i,t_i}) = \max_{\|\boldsymbol{f}_{t_i}\|_a \leq \delta_i} \left\| M_{t_{i+1}-t_i}(\boldsymbol{f}_{t_i})(\boldsymbol{U}_{t_i}) - M_{t_{i+1}-t_i}(0)(\boldsymbol{U}_{t_i}) \right\|_b,$$

(2.9)

where $M_{t_{i+1}-t_i}(\boldsymbol{f}_{t_i})$ is the propagator of eq. (2.3) from time $t_i$ to $t_{i+1}$ and $\boldsymbol{U}_{t_i} = M_{t_i-t_{i-1}}(\boldsymbol{f}_{\delta_{i-1},t_{i-1}})(\boldsymbol{U}_{t_{i-1}})$. From eq. (2.9), it is easily observed that the time-dependent NFSVs can be obtained by computing several constant NFSVs. Considering that the FSV is constant (see Barkmeijer et al., 2003) and that the goal of this article is to reveal the differences between FSV and NFSV, we only study the constant NFSVs and explore the effects of non-linearity on the optimal tendency perturbation.

## 3. Computing the non-linear forcing singular vectors

The NFSV is related to a constrained optimisation problem. It is very difficult to solve such optimisation problems analytically; therefore, we have to solve it numerically. Such optimisation problems are often solved using optimisation algorithms, where the gradient of the related objective function is very important. During large-scale optimisation, the gradient of the objective function with respect to the initial perturbations is often obtained using the adjoint method (Le Dimet and Talagrand, 1986). For example, Mu et al. (2003) proposed the CNOP approach, which is characterised by an optimisation problem, and computed the CNOP along the gradient direction of the objective function, where the gradient is derived using an adjoint method. Although the NFSV can also be solved using an adjoint method, computing the NFSV requires the gradient of the objective function with respect to the constant tendency perturbation $\boldsymbol{f}$. In fact, the gradient of the objective function with respect to the tendency perturbation $\boldsymbol{f}$ can be transferred to a particular case of the objective function with augmented initial perturbations. Next, we describe how to compute the NFSVs using a gradient.

Existing optimisation solvers are often used to compute minimisation problems; however, the NFSV is related to a constrained maximisation problem. To calculate the NFSVs, we convert the maximisation problem to a minimisation problem. In particular, we rewrite eq. (2.8) as follows:

$$\begin{aligned} J_1(\boldsymbol{f}) &= -J(\boldsymbol{f}) = -\frac{1}{2}\|M_\tau(\boldsymbol{f})(\boldsymbol{U}_0) - M_\tau(0)(\boldsymbol{U}_0)\|_b^2 \\ &= -\frac{1}{2} < \boldsymbol{u}_f(\tau), \boldsymbol{u}_f(\tau) >, \end{aligned}$$

(3.1)

where $\boldsymbol{u}_f(\tau)$ is a departure from the reference state $U(\boldsymbol{x},\tau) = M_\tau(0)(\boldsymbol{U}_0)$ caused by the constant tendency perturbation $\boldsymbol{f}$, and $< \cdot >$ is the inner product. To facilitate the description, we will use $\boldsymbol{u}(\tau)$ instead of $\boldsymbol{u}_f(\tau)$. The maximisation problem (2.7) related to the NFSVs then becomes a minimisation problem of the function $J_1(\boldsymbol{f})$.

The first-order variation of $J_1(\boldsymbol{f})$ is as follows:

$$-\delta J_1 = <\boldsymbol{u}(\tau), \delta\boldsymbol{u}(\tau)> = <\frac{-\partial J_1}{\partial u_0}, \delta u_0> + <\frac{-\partial J_1}{\partial f}, \delta f>.$$

$$(3.2)$$

Furthermore, $\delta\boldsymbol{f}$ is governed by the following tangent linear model:

$$\begin{cases} \frac{\partial \delta\boldsymbol{u}}{\partial t} = \frac{\partial F(\boldsymbol{U}(t)+\boldsymbol{u}(t))}{\partial \boldsymbol{u}}\delta\boldsymbol{u} + \delta\mathbf{f}(\boldsymbol{x}), \\ \frac{\partial \delta\boldsymbol{f}}{\partial t} = 0, \\ \delta\boldsymbol{u}|_{t=0} = \delta\boldsymbol{u}_0 = 0, \\ \delta\mathbf{f}|_{t=0} = \delta\boldsymbol{f}(x). \end{cases} \quad (3.3)$$

By introducing two Lagrangian multipliers, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, we obtain

$$\begin{aligned} -\delta J_1 = \ & <\boldsymbol{u}(\tau), \delta\boldsymbol{u}(\tau)> \\ & - \int_0^\tau <\lambda_1(t), \frac{\partial \delta\boldsymbol{u}}{\partial t} - \frac{\partial F(\boldsymbol{U}+\boldsymbol{u})}{\partial u}\delta\boldsymbol{u} - \delta\mathbf{f}> dt \\ & - \int_0^\tau <\lambda_2(t), \frac{\partial \delta\mathbf{f}}{\partial t}> dt. \end{aligned} \quad (3.4)$$

By solving eq. (3.4) (details are found in the Appendix) and comparing it with eq. (3.2), we obtain the following gradient:

$$\frac{\partial J_1}{\partial f} = -\frac{\partial J}{\partial f} = -\lambda_2(0). \quad (3.5)$$

$\lambda_1(t)$ and $\lambda_2(t)$ in eqs. (3.4) and (3.5) satisfy

$$\begin{cases} \frac{\partial \lambda_1}{\partial t} + \left[\frac{\partial F(\boldsymbol{U}(t)+\boldsymbol{u}(t))}{\partial \boldsymbol{u}}\right]^* \lambda_1 = 0, \\ \frac{\partial \lambda_2}{\partial t} + \lambda_1 = 0, \\ \lambda_1|_{t=\tau} = \boldsymbol{u}(\tau), \\ \lambda_2|_{t=\tau} = 0. \end{cases} \quad (3.6)$$

and eq. (3.6) is the adjoint equation of eq. (3.3). By integrating eq. (3.6), we obtain the gradient $\partial J_1/\partial f = -\lambda_2(0)$ and thereby allow the NFSV to be computed using optimisation solvers such as Spectral Projected Gradient 2 (SPG2; Birgin et al., 2000), Sequential Quadratic Programming (SQP; Powell et al., 1982), and Limited Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS; Liu and Nocedal, 1989).

From eq. (3.6), it is clear that $\lambda_1(0)$ is necessary to obtain the gradient $\lambda_2(0)$. From eq. (3.5), we know that $\lambda_1(t)$ satisfies the equation

$$\begin{cases} \frac{\partial \lambda_1}{\partial t} + \left[\frac{\partial F(\boldsymbol{U}(t)+\boldsymbol{u}(t))}{\partial \boldsymbol{u}}\right]^* \lambda_1 = 0, \\ \lambda_1|_{t=\tau} = \boldsymbol{u}(\tau). \end{cases} \quad (3.7)$$

It is clear that eq. (3.7) is composed of the adjoint model of the non-linear model, eq. (2.1), with respect to the initial perturbations. Therefore, the adjoint model in eq. (3.6), which is associated with the constant tendency perturbation $\boldsymbol{f}$, is based on eq. (3.7), which is associated with the

initial perturbations. Numerically, we do not alter the existing adjoint model in eq. (3.7) but only add a line of code showing the discretisation of the equation $\frac{\partial \lambda_2}{\partial t} + \lambda_1 = 0$ with $\lambda_2|_{t=\tau} = 0$ after the code of the adjoint model in eq. (3.7) to obtain the adjoint model in eq. (3.6). Several models have adjoint models with respect to the initial perturbations and have been used to solve the CNOP and linear singular vectors. For example, see the two-dimensional QG model used in this study, the three-layer QG model, the Zebiak-Cane model (Zebiak and Cane, 1987), and the Mesoscale Model 5 (Dudhia, 1993). Therefore, if one hopes to study the NFSVs of these models, one can easily modify the adjoint eq. (3.7) associated with the initial perturbations and then obtain the adjoint model in eq. (3.6), which is related to the NFSVs.

In this article, the gradient of the objective function $J_1$ with respect to the initial perturbations and model perturbations is derived by introducing two Lagrangian multipliers, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$. This approach clearly shows the relationship between the adjoint model associated with initial perturbations and the adjoint model associated with the model perturbations, and this relationship helps us to numerically obtain the adjoint model associated with the constant tendency perturbations by modifying the existing adjoint models associated with the initial perturbations. Of course, the gradient can also be obtained by differentiating the objective function with respect to the initial perturbations and model perturbations (i.e. using the definition of the derivative of a non-linear operator) (see Shutyaev et al., 2008).

## 4. The NFSVs of a two-dimensional quasi-geostrophic model

We consider a two-dimensional QG model

$$\begin{cases} \frac{\partial P}{\partial t} + \partial(\Phi, P) = 0, \\ P = \nabla^2\Phi - F\Phi + f_0 + \frac{f_0}{H}h_s, \text{ in } \Omega \times [0, T] \\ \Phi|_{t=0} = \Phi_0. \end{cases} \quad (4.1)$$

where $P$ is the potential vorticity, $\Phi$ is the stream function, $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplacian operator, $x$ and $y$ are the zonal and meridional coordinates, $t$ is the time, $F$ is the planetary Froude number, $f_0$ is the Coriolis parameter, $H$ is the characteristic depth, and $h_s$ is the topography. The horizontal Jacobian operator is shown as $\partial(\Phi, P) = \Phi_x P_y - \Phi_y P_x$, and $\Omega = [0, X] \times [0, Y]$ with a double periodical boundary condition. For any fixed $T > 0$ and initial condition $\Phi_0$, we can solve the initial value problem (4.1) to obtain $\Phi(x, y, T)$. Hence, the propagator $M_\tau$ is well defined; that is, $\Phi(x, y, T) = M_\tau(\Phi_0)$ is the solution to eq. (4.1) at a time $T$ (see Mu and Zeng, 1991).

For eq. (4.1), we use a constant tendency error $f(x, y)$ to disturb the model. The result is as follows:

$$\begin{cases} \frac{\partial P}{\partial t} + \partial(\Phi, P) = f(x, y), \\ P = \nabla^2\Phi - F\Phi + f_0 + \frac{f_0}{H}h_s, \text{ in } \Omega \times [0, T] \\ \Phi|_{t=0} = \Phi_0. \end{cases} \quad (4.2)$$

The propagator of the perturbed QG model eq. (4.2) is denoted by $\boldsymbol{M}_\tau(f)$, where $f$ is the tendency error of the model. Using this notation, the propagator $\boldsymbol{M}_\tau$ in eq. (4.1) can be rewritten as $\boldsymbol{M}_t(0)$, which indicates that model eq. (4.1) is perfect. Following eq. (2.8), we write the objective function related to the NFSV as follows:

$$J(f_\delta) = \max_{\|f\|_a \leq \delta} \|\boldsymbol{M}_\tau(f)(\Phi_0) - \boldsymbol{M}_\tau(0)(\Phi_0)\|_b, \quad (4.3)$$

where $\Phi_0$ is the initial value of a basic flow. The norm $\|\Phi^2\|_b = \int_\Omega (|\nabla\Phi|^2 + F|\Phi|^2)dxdy$ is used to measure the effect of prediction errors caused by constant tendency errors $f(x, y)$, where $\nabla = (\partial/\partial x, \partial/\partial y)$. The constraint conditions of the tendency errors $f(x, y)$ is measured by the $L_2$ norm, i.e. $\|f(x, y)\|_a = \sqrt{\sum_{i,j} f_{i,j}^2}$, where $f_{i,j}$ represents the tendency errors at each space grid $(i, j)$. The objective function $J$ measures the magnitudes of the prediction errors caused by the constant tendency errors $f(x, y)$ at time $\tau$. The NFSVs obtained using eq. (4.3) represent the constant tendency errors that cause the largest prediction errors at the prediction time and describe a type of model systematic errors that induce severe prediction uncertainties.

To obtain the NFSV numerically, we use the discretisation versions of the operators $M_\tau(0)$ and $M_\tau(f)$. The Jacobian operator was discretised through the Arakawa finite difference scheme (cf. Arakawa, 1966). The temporal discretisation was carried out using the Adams-Bashforth scheme. The five-point difference scheme was employed to discretise the Laplacian operator. The stream function $\Phi$ is treated as an unknown term, and the potential vorticity $P$ is calculated using the second equation from eq. (4.1). For an optimisation algorithm, we use the SPG2 method, which can be used to calculate the least value of a function of a large number of variables, subject to equality and inequality constraints. A detailed description of the algorithms can be found in Birgin et al. (2000).

To calculate the NFSVs using the SPG2 method, we made at least 30 initial random guesses; if several initial guesses for the SPG2 converge at a point in the phase space, this point can be considered as the maximum in a neighbourhood. Several such points are obtained. The point that maximises the objective function in eq. (2.8) is regarded as the NFSV. We note that the NFSV is the global maximum of the objective function $J$. However, it is possible that the objective function $J$ attains its local maximum in the small neighbourhood of a point in the phase space. Such a tendency perturbation is called a local NFSV. In this study, we

attempt to investigate the constant tendency error that causes the largest prediction error at the prediction time; therefore, we do not consider local NFSVs.

Based on the QG model, we study the NFSVs. Several non-linearly stable reference states (or basic flows) were used to investigate the NFSVs. The results demonstrate that the NFSVs depend on the given basic flows. In this article, we use two of the basic flows to describe the results associated with the NFSVs: one basic flow possesses the initial structure with a zonal flow (Ref-1) as in Fig. 1a and the other has a structure with an almost meridional flow (Ref-2) as in Fig. 1c. These two basic flows, as in Mu and Zhang (2006), are chosen based on Arnold's non-linear stability criteria (see Mu and Shepherd, 1994). A detailed description of these basic flows can be found in the Appendix to Mu and Zhang (2006). As a comparison, the FSVs of Ref-1 and Ref-2 are also computed by considering the tangent linear approximations $\mathbf{M}_t(\boldsymbol{f})$ and $\mathbf{M}_\tau(0)$ of $M_t(\boldsymbol{f})$ and $M_\tau(0)$ with respect to the given basic flows. To show the NFSVs and demonstrate the differences between the NFSVs and FSVs, two groups of comparisons are made. The first group of comparisons is concerned with the differences between the patterns of the NFSVs and FSVs, and the second is related to the differences between the prediction errors caused by the NFSVs and FSVs.

We take the spatial domain of $\Omega = [0,6.4] \times [0,3.2]$, which corresponds to the dimensional case $[0,6400 \text{ km}] \times [0,3200 \text{ km}]$. The QG model parameters are chosen as follows: $F = 0.102, f_0 = 10$. The grid spacing $d = 0.2$ corresponds to a dimensional length of 200 km, and the time step $dt = 0.006$ corresponds to 10 min.

The NFSVs are a generalisation of the FSVs in a non-linear field and are required for a comparison with FSVs. The FSVs are derived from a linearised model as in eq. (2.4). It is easily shown that if $\boldsymbol{f}_L$ is an FSV, then a scaled vector $\boldsymbol{cf}_L$ ($c$ is a constant) is also a FSV. We define a scaled FSV $\hat{\boldsymbol{f}}_L$ as follows:

$$\hat{f}_L = \frac{\|f_\delta\|}{\|f_L\|} f_L,$$

thus,

$$\|\hat{f}_L\| = \|f_\delta\| = \delta.$$

That is, the scaled FSV $\hat{f}_L$ possesses the same magnitude as that of the NFSV $f_\delta$. Note that we are comparing NFSVs and FSVs.

### 4.1. A steady zonal basic flow

In this section, we choose a stable basic flow, which is obtained by integrating eq. (4.1) using the initial state $\Phi_0 = 0.2724 \sin(2\pi y/3.2) + 27.993$. The relevant
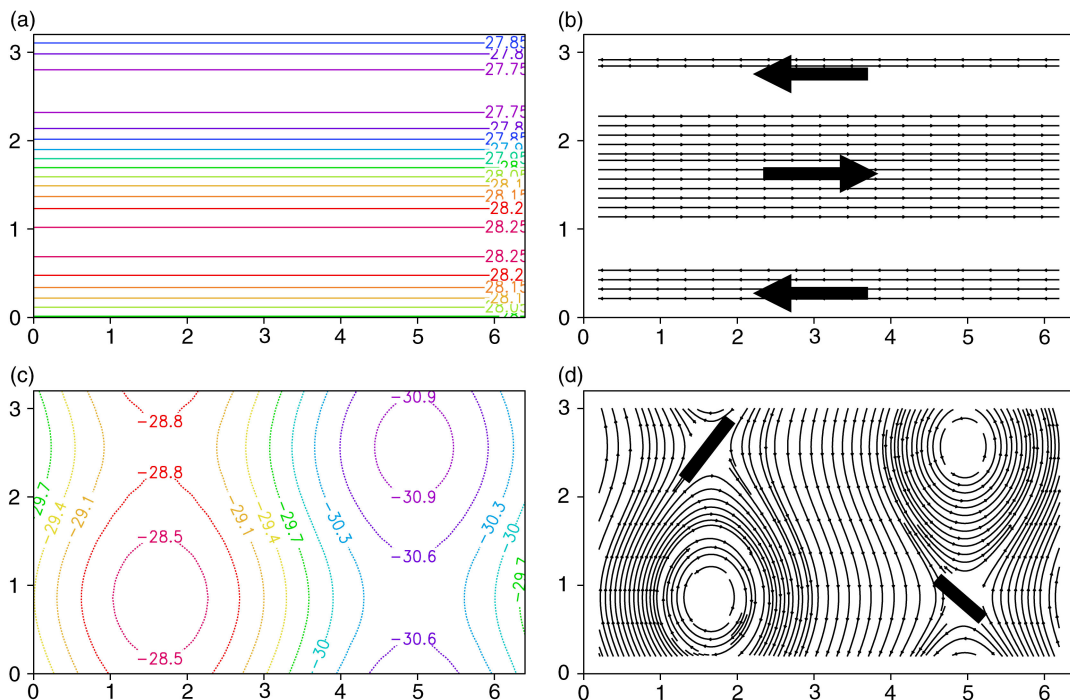
*Fig. 1.* The stream function and the corresponding quasi-geostrophic winds of the Ref-1 and Ref-2 basic flows adopted in Experiments I and II, respectively. (a) and (b) correspond to the Ref-1 basic flow; (c) and (d) correspond to the Ref-2 basic flow.

topography is $h_s = \sin(2\pi y/3.2) + 5$, and $1/H = 0.1$. This basic flow is an equilibrium solution to eq. (4.1) and is denoted by 'Ref-1' (see Fig. 1). In the numerical experiments, the NFSVs are calculated for time steps 288, 720, 1008, and 1296, which correspond to the optimisation times (i.e. forecast periods) $\tau = 2$, 5, 7 and 9 d, respectively. The tendency perturbations $f$ are experimentally constrained in $\|f\|_a \leq \delta$ with $\delta \in [0.8, 3.2]$. In particular, we take $\delta = 0.8$, 1.6, 2.4 and 3.2 to describe the results.

*4.1.1. Comparing the NFSVs and FSVs.* The computations demonstrate that for each optimisation time, there exists one NFSV of the Ref-1 for a given constraint magnitude. Figure 2 shows the stream function component of the NFSVs $f_\delta$ and the corresponding FSVs $\hat{f}_L$ for the optimisation times of 2, 5, 7, and 9 d and a tendency perturbation magnitude (value of $\delta$) of 1.6. The results show that the FSVs have a typical zonal wave number of 5 for all optimisation times and always show a symmetric V-shaped structure with concentrations around $y = 0.8$ and $y = 2.4$.

We note that the NFSVs have patterns similar to the FSVs only for short optimisation times (for example, the optimisation time $\tau = 2$ d) but are significantly different from the FSVs for long optimisation times (see Fig. 2). More precisely, NFSVs with long optimisation times compared to the FSVs have a very asymmetric V-shaped

structure and cover a much smaller region, although they also tend to be concentrated around $y = 0.8$ and $y = 2.4$. The differences between the NFSV and FSV patterns depend not only on the optimisation times but also on the perturbation magnitudes. In Fig. 3, we plot the NFSV and FSV patterns with different magnitudes for the optimisation time of 7 d. The results show that the NFSVs are more similar to the FSVs for small tendency perturbations than for large tendency perturbations. That is, when the tendency perturbations are sufficiently small, the NFSVs tend to have a symmetric V-shaped structure similar to that of the FSVs, but when the tendency perturbations are large, the NFSVs become much more asymmetric. In summary, when the optimisation times are long and the tendency perturbations are large, the NFSVs tend to have a much more asymmetric region than the FSVs and cover a much smaller region.

Figure 4 shows the patterns of the stream function components of the prediction errors caused by the NFSV and FSV obtained using an optimisation time of 7 d and a tendency perturbation magnitude of 1.6. The results show that the stream function component of the prediction error caused by the NFSV in the non-linear QG model is significantly different from that caused by the FSV. It is clear that the differences between the NFSV and FSV patterns also induce different responses in terms of the prediction errors at the prediction time. In particular,
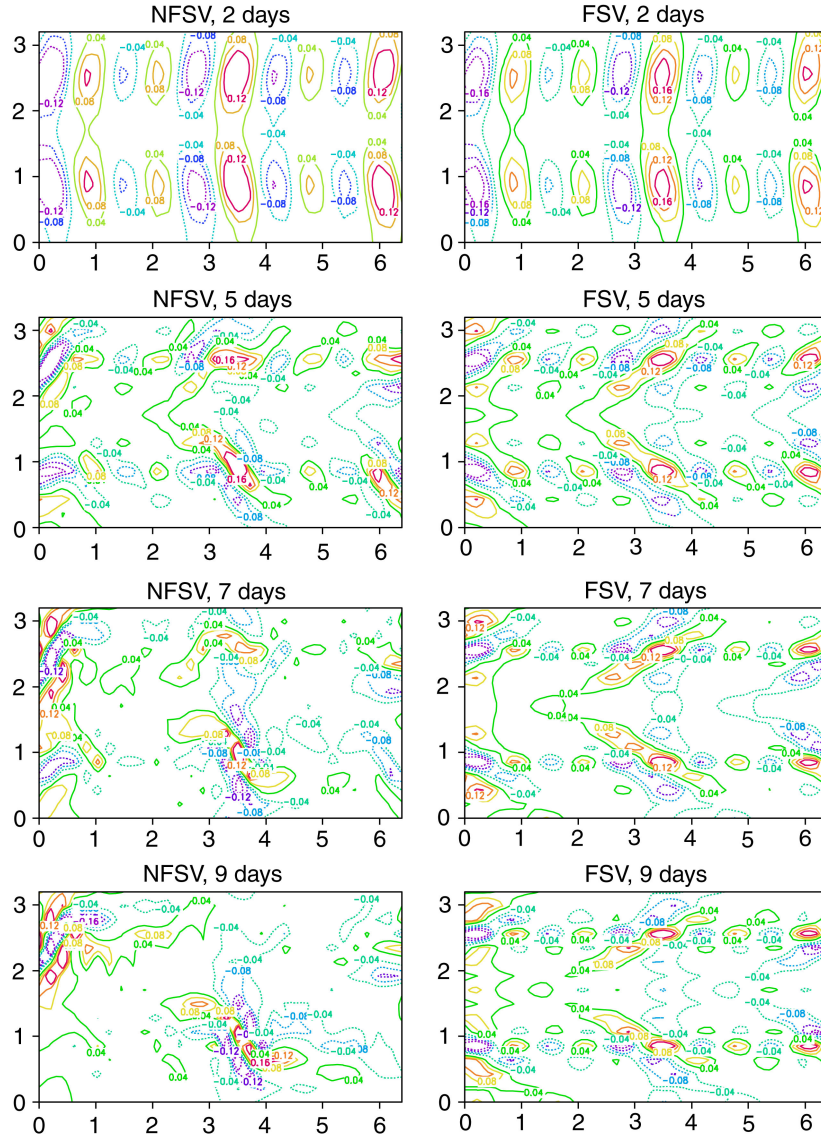
Fig. 2.   The stream function components of the NFSVs and FSVs of the Ref-1 basic flow with a perturbation magnitude of 1.6. The left (right) column lists the NFSVs (FSVs) with optimisation times of 2, 5, 7, and 9 d.

we find that the stream function components of the prediction errors caused by the NFSVs and FSVs in the non-linear QG model exhibit smaller-scale spatial patterns than those caused by the FSVs using the linearised model. Furthermore, some small-scale vortices can be observed in the stream function patterns of the prediction errors caused by the FSVs. All of these suggest that the small scales of the patterns result from the effects of the non-linearities.

In addition, we also investigate the differences in the magnitudes of the prediction errors caused by the NFSVs and FSVs and find that the magnitudes of the prediction errors caused by the FSVs in the linearised QG model are larger than those caused by the FSVs in the non-linear QG model, especially for long optimisation times and/or large

tendency perturbations (see Fig. 5). Thus, the non-linearity suppresses the linear growth of the prediction errors, i.e. the non-linearity has a damping effect on the growth of the prediction errors. Furthermore, we find that the prediction errors caused by the NFSVs in the non-linear QG model are larger than those caused by the FSVs in the non-linear QG model, indicating that the NFSVs of the asymmetric structure are inclined to reduce the non-linear damping effect on the error growth.

*4.1.2. Interpretation.* The Ref-1 basic flow is zonal, and there is a strong velocity shear near $y = 0.8$ and $y = 2.4$, as the wind direction shifts from east to west near $y = 0.8$ and
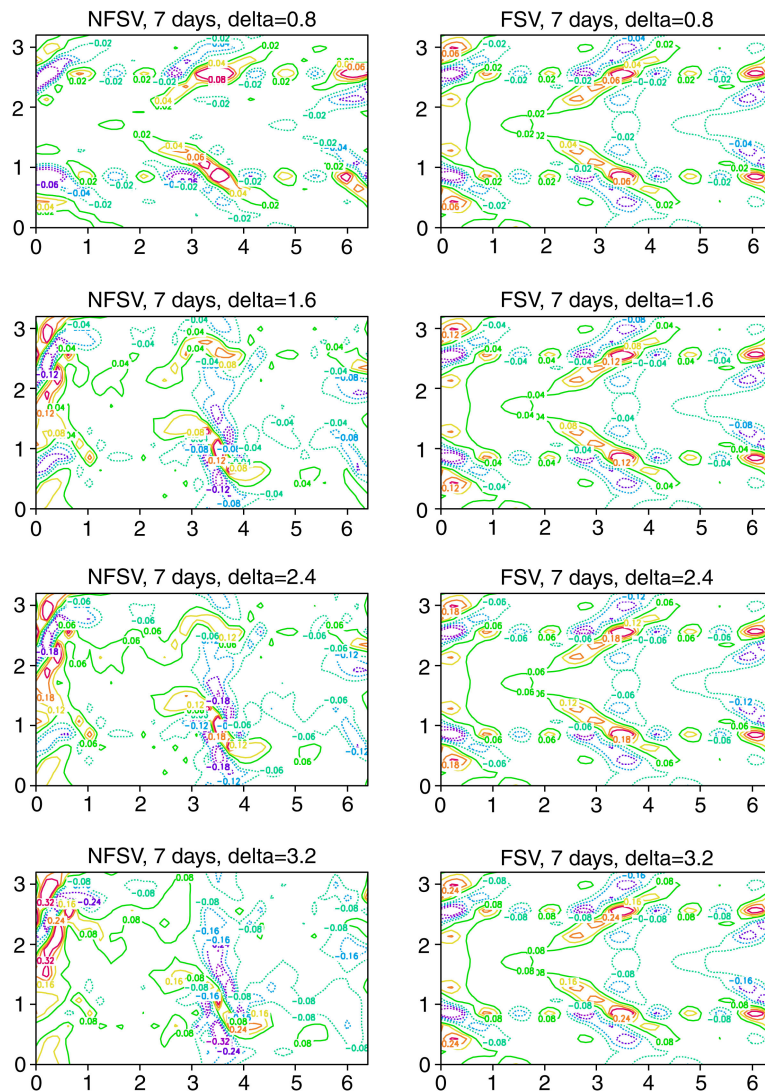
*Fig. 3.* The stream function components of the Ref-1 NFSVs and FSVs with an optimisation time of 7 d. The left (right) column lists the NFSVs (FSVs) with perturbation magnitudes of 0.8, 1.6, 2.4, and 3.2.

from west to east near $y = 2.4$ [see Fig. 1b]. Barotropic instabilities usually occur in regions with a strong velocity shear, and this condition may be favourable for the perturbations induced by the initial perturbations to extract energy from the basic flows (Pedlosky, 1979; Tung, 1981; Lacarra and Talagrand, 1988). We have shown that the FSVs and NFSVs that are expected to become large are concentrated around $y = 0.8$ and $y = 2.4$; therefore, the perturbations induced by the NFSVs and FSVs also tend to extract energy from the basic flow in regions with strong shear. In addition, Mak and Cai (1989) have demonstrated that a perturbation that has a phase tilted against the basic shear would optimally extract energy from the basic flow; this finding lends support to the assertion that the NFSVs and FSVs have V-shaped structures. Nevertheless, the

NFSVs have much more asymmetric structures than the FSVs; furthermore, the NFSVs yield much larger prediction errors than the FSVs. Next, we will explain why the NFSVs of asymmetric structures cause larger prediction errors than the FSVs of symmetric structures.

We follow the Riviere et al. (2008), which shows that a shear opposite to the basic shear develops through time, to study the zonal mean of the stream function components of the prediction errors caused by the NFSVs and FSVs. In response, the growth rate of the instability (given by the total zonal-mean shear) should diminish (Gutowski, 1985; Nakamura, 1999; Riviere et al., 2008), and the perturbation will extract less energy from the basic flow. As a result, the growth of the prediction errors induced by the tendency errors will be limited. In Fig. 6, we illustrate the zonal-mean
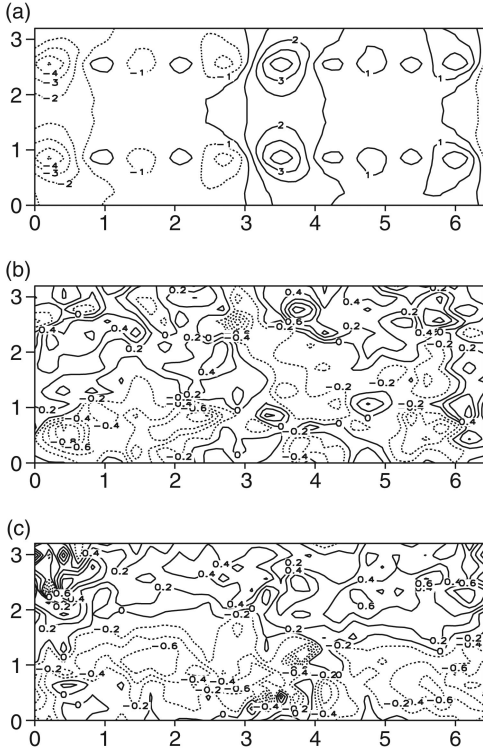
*Fig. 4.* Patterns of the stream functions of the prediction errors caused by (a) the FSV in the linearised QG model, (b) the FSV in the non-linear QG model, and (c) the NFSV in the non-linear QG model, where the optimisation is 7 d and the model perturbation magnitude is 1.6 in terms of the chosen norm.

of the stream function component of the prediction errors caused by the NFSVs and FSVs with an optimisation time of 7 d and a tendency perturbation magnitude of 1.6. The figure shows that the zonal-mean shear induced by the NFSVs is weaker than that induced by the FSVs at the initial stage (e.g. at 2–5 d). It is inferred that the total zonal shear in the NFSV case is stronger than that in the FSV case, finally favouring the NFSVs to induce much larger prediction errors than the FSVs at the prediction time.

We have demonstrated that the non-linearity has a damping effect on the growth of the prediction errors caused by the FSVs, and the tendency errors of the NFSV structures favour reducing the non-linear damping effect on the prediction errors and inducing larger prediction errors than the FSVs (see the last paragraph in section 4.1.1). In fact, the non-linearity here originates from the effect of perturbation advection in the non-linear QG model.

Assuming that the potential vorticity $P$ with the corresponding stream function $\Phi$ is the basic flow of the QG model as mentioned in section 4.1.1, and the perturbation $p$ with the corresponding stream function $\phi$ describes its prediction error caused by the tendency error $f(x, y)$, then the evolution of the prediction error $(p, \phi)$ is described by the non-linear perturbation eq. (4.4).

$$\frac{\partial p}{\partial t} + \frac{\partial \Phi}{\partial x}\frac{\partial p}{\partial y} + \frac{\partial \phi}{\partial x}\frac{\partial P}{\partial y} - \frac{\partial P}{\partial x}\frac{\partial \phi}{\partial y} - \frac{\partial p}{\partial x}\frac{\partial \Phi}{\partial y} + \frac{\partial \phi}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial p}{\partial x}\frac{\partial \phi}{\partial y}$$
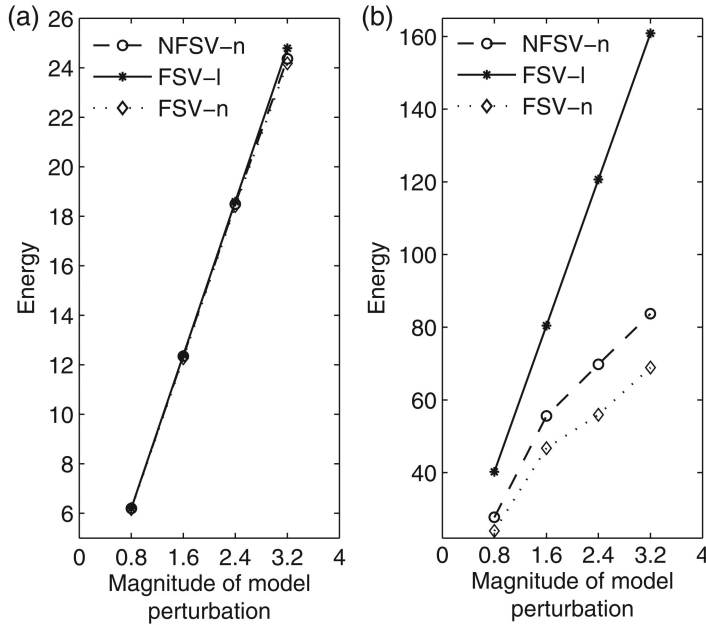$$= f(x, y).$$
$$(4.4)$$



*Fig. 5.* The prediction errors caused by the NFSVs and the FSVs in the QG model (denoted by NFSV-n and FSV-n) and those caused by the FSVs in the linearised QG model (denoted by FSV-l). The horizontal axis represents the perturbation magnitudes, and the vertical axis describes the prediction errors in terms of the energy. (a) uses an optimisation time of 2 d, and (b) uses the optimisation time of 7 d.

The linearised version of eq. (4.4) is given by eq. (4.5).

$$\frac{\partial p}{\partial t} + \frac{\partial \Phi}{\partial x}\frac{\partial p}{\partial y} + \frac{\partial \phi}{\partial x}\frac{\partial P}{\partial y} - \frac{\partial P}{\partial x}\frac{\partial \phi}{\partial y} - \frac{\partial p}{\partial x}\frac{\partial \Phi}{\partial y} = f(x,y). \quad (4.5)$$

It is noticed that eq. (4.5), compared to eq. (4.4), lacks the perturbation advection term $\frac{\partial \phi}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial p}{\partial x}\frac{\partial \phi}{\partial y}$, which is the non-linear term of the eq. (4.4). If $p_N$ and $p_L$ are the solutions of eqs. (4.4) and (4.5), respectively, $p_N - p_L$ represents the effect of perturbation advection (or non-linear advection) on $p_N$. In this article, we are concerned with their corresponding stream function components $\varphi_N$ and $\varphi_L$. Similarly, $\varphi_N - \varphi_L$ describes the effect of non-linear advection on the evolution of the stream function $\varphi_N$.

In Fig. 7b, we show the effect of the non-linear advection, which is indicated by $\varphi_N - \varphi_L$ and associated with the tendency error FSV for an optimisation time of 7 d and a tendency perturbation magnitude of 1.6. It is

shown that the vortices induced by the non-linear advection are concentrated around $y = 0.8$ and $y = 2.4$ and have vorticities of nearly opposite signs to those in the stream function components of the prediction errors caused by the FSVs in the linearised QG model [see Fig. 7a and 7b]. It is therefore inferred that, when the non-linear advection term is superimposed onto the linearised QG model, the vorticities associated with the prediction errors in the linearised QG model will be offset and the linear growth of the prediction errors induced by the FSV will then be suppressed by the non-linear advection. It follows that the non-linear advection has a damping effect on the growth of the prediction errors. In addition, we notice that the non-linear advection associated with the FSV shows an asymmetric effect in meridional direction [see Fig. 7b]. Therefore, despite the fact that the FSV is symmetric, the related non-linear advection has the ability
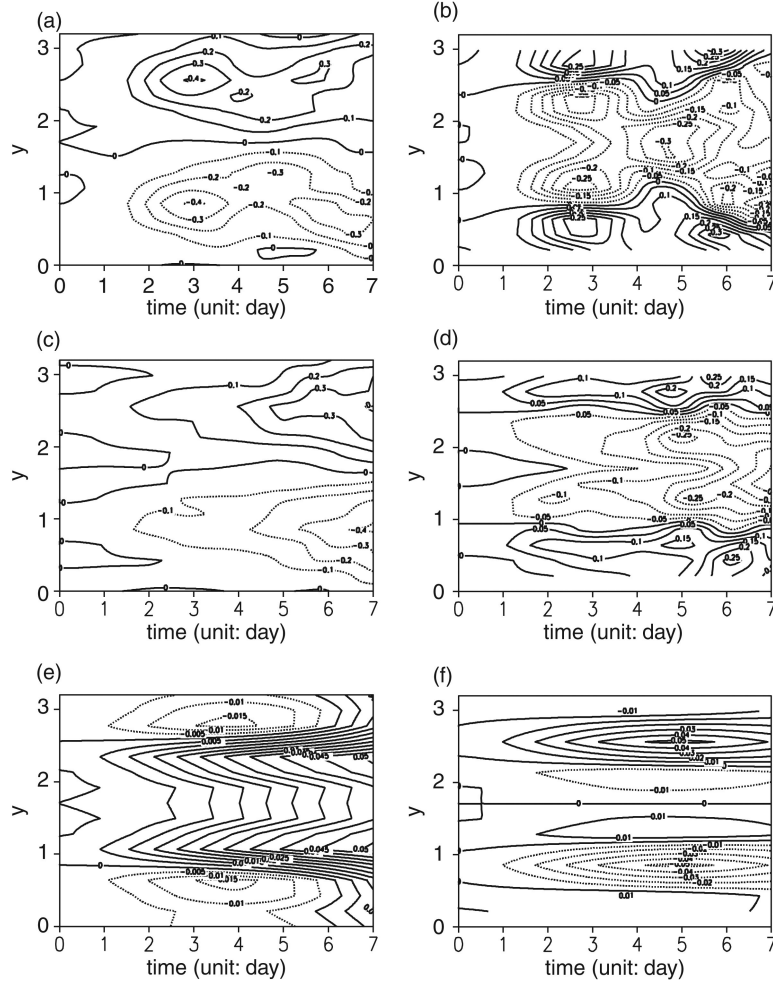


*Fig. 6.* The zonal-means of the stream function of prediction errors (left column) and the corresponding zonal wind (right column) caused by (a, b) the FSV in the linearised QG model, (c, d) the FSV in the non-linear QG model, and (e, f) the NFSV in the non-linear QG model. The optimisation time is 7 d, and the model perturbation magnitude is 1.6.
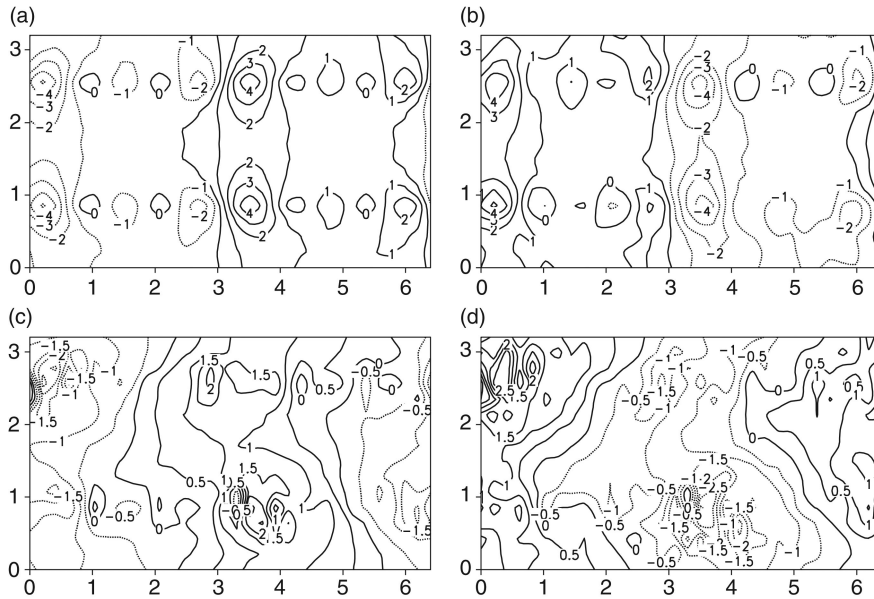
*Fig. 7.*    The stream function patterns of the prediction error caused by the tendency perturbation in the linearised QG model (left column) and the differences between the stream function induced by the tendency perturbation in the non-linear QG model and that in the linearised QG model (right column). The latter indicates the effect of non-linear advection on the perturbation growth. (a) and (b) are for the FSV, and (c) and (d) are for the NFSV. Both FSV and NFSV are computed with the optimisation time 7 d and the tendency perturbation magnitude 1.6. The figures in (a, b, c, and d) correspond to the patterns at the final time of optimisation time.

to induce an asymmetric effect in the meridional direction; thus, the NFSVs resulting from the non-linear QG model exhibit an asymmetric mode. When the tendency error of the NFSV structure are forced to the non-linear QG model, the vortices induced by the non-linear advection also tend to have vorticities with opposite signs to those in the stream function components of the prediction errors caused by the NFSV in the linearised QG model and therefore suppress the perturbation growth [see Fig. 7c and 7d]. Furthermore, the NFSV, due to the effect of its spatial structure, induces the related non-linear advection to have smaller amplitude than the FSV case to suppress the perturbation growth. In other words, the tendency errors of the NFSV structure favour reducing the damping effect of non-linearity on the growth of the prediction errors and lead to larger prediction errors than the FSV. Another non-linear effect is that the non-linear advection tends to break the waves shown in the stream function patterns of the prediction errors caused by the NFSVs and FSVs and saturate them into much smaller scales (see Fig. 4; also see Riviere et al., 2008).

### 4.2.  A steady meridional basic flow

In this experiment, we adopt the basic flow using the initial state $\Phi_0 = 1.097 \sin\frac{2\pi x}{6.4} + 0.2629 \sin\frac{2\pi y}{3.2} - 29.674$.

The corresponding topography is $h_s = \sin(2\pi x/6.4) + \sin(2\pi y/3.2) + 1$, and $1/H = 0.1$. This basic flow, denoted by 'Ref-2', is also non-linearly stable but is almost meridional [Fig. 1c] and has two deformation fields, which are located at approximately $x = 1.8$, $y = 2.5$ and $x = 5$, $y = 0.9$ [Fig. 1d], respectively. Strong velocity shears exist in these two deformation fields. In the former field, the axis of contraction is northeast–southwest, whereas in the latter field, the axis of contraction is northwest–southeast. According to the explanation provided by experiment I, the related NFSVs should be concentrated around the regions of strong shear in the basic flow. Our numerical results demonstrate this point.

For each optimisation time, one NFSV of the Ref-2 exists for a given constraint magnitude. Figures 8 and 9 show the stream function components of the NFSVs $f_\delta$ and FSVs $\hat{f}_L$ of the Ref-2 basic flow with different optimisation times and magnitudes, respectively. In these figures, it can be observed that the NFSVs are indeed concentrated around the deformation field regions that exhibit strong velocity shear. Furthermore, the NFSVs have elongated local phases along the deformation field axes of contraction; these local phases are most favourable for the perturbations that optimally extract energy from the basic flow (Lacarra and Talagrand, 1988; Mak and Cai, 1989). The FSVs show patterns similar to the NFSVs. Even for a given
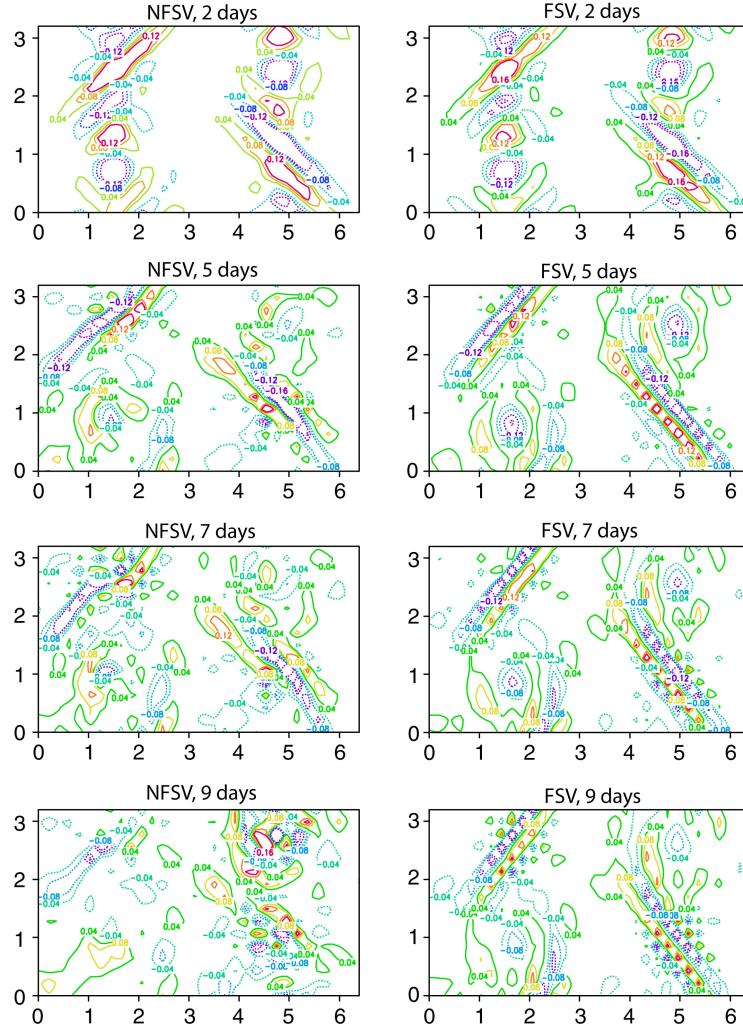
*Fig*. *8*.    The stream function components of the Ref-2 NFSVs and FSVs with a perturbation magnitude of 1.6. The left (right) column lists the NFSVs (FSVs) with optimisation times of 2, 5, 7, and 9 d.

long optimisation time and/or large tendency perturbation magnitude, the spatial structures of the FSVs are only slightly different from those of the NFSVs. More precisely, the FSVs have only a negligible extension in their patterns compared with the NFSVs. As a result, the prediction errors they cause are approximately equivalent in either their magnitudes or their corresponding spatial patterns (see Figures 10 and 11). It is inferred that the non-linear advection trivially affects the spatial structure of the FSVs in the case of Ref-2. Physically, the meridional-mean shear of the stream function of the prediction errors induced by the FSVs in the non-linear QG model is opposite to the basic shear and has almost the same intensity as that induced by the NFSVs (Fig. 12); therefore, the intensity of the total meridional-mean shear in the non-linear QG model is nearly the same for the NFSV case as for the FSV case;

therefore, the prediction errors caused by the FSVs in the non-linear QG model are trivially smaller than those caused by the NFSVs in the non-linear QG model. However, the non-linear advection has a significant influence on the amplitudes of the prediction errors caused by the FSVs in the non-linear QG model. As we show in Fig. 10, the non-linearity for the Ref-2 basic flow is also obviously suppressing the perturbation growth, and the non-linearity has a damping effect on the perturbation growth (also see Fig. 13). Nevertheless, in terms of the energies of prediction errors, the NFSVs slightly reduce the damping effect of the non-linearity on the perturbation growth (see Fig. 10). In addition, we note that the non-linear advection also tends to break the waves shown in the stream function pattern of the prediction errors and saturate them to small scales (see Fig. 11).
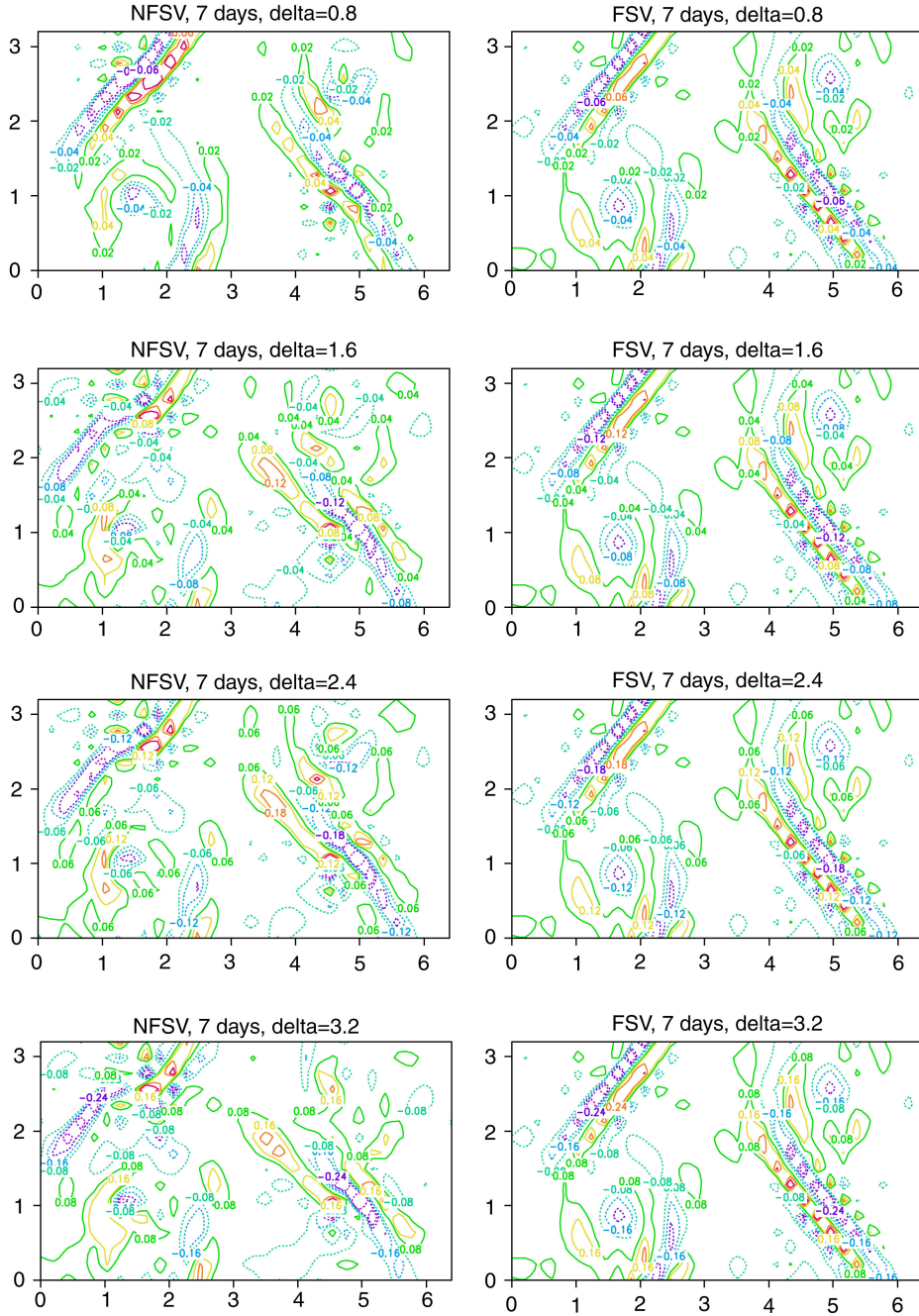
*Fig. 9.*    The stream function components of the Ref-2 NFSVs and FSVs with an optimisation time of 7 d. The left (right) column lists the NFSVs (FSVs) with model perturbation magnitudes of 0.8, 1.6, 2.4, and 3.2.

## 5.  Summary and discussion

In this article, we extend the FSV approach to the non-linear field and propose the concept of a NFSV. The advantage of the NFSV approach is that it sufficiently considers the effects of non-linearity. The NFSVs can determine the time-dependent tendency error of a model. In this article, because the FSV is constant in time and the goal is to

compare FSV and NFSV, we only study the constant NFSV, which may describe a type of model systematic error that causes significantly large prediction errors at the prediction time. A two-dimensional QG model is adopted to study the NFSV model errors. Two non-linearly stable states (denoted by Ref-1 and Ref-2) of the QG model are chosen as reference states (or basic flows) to describe the results. Furthermore, the Ref-1 basic flow is zonal and
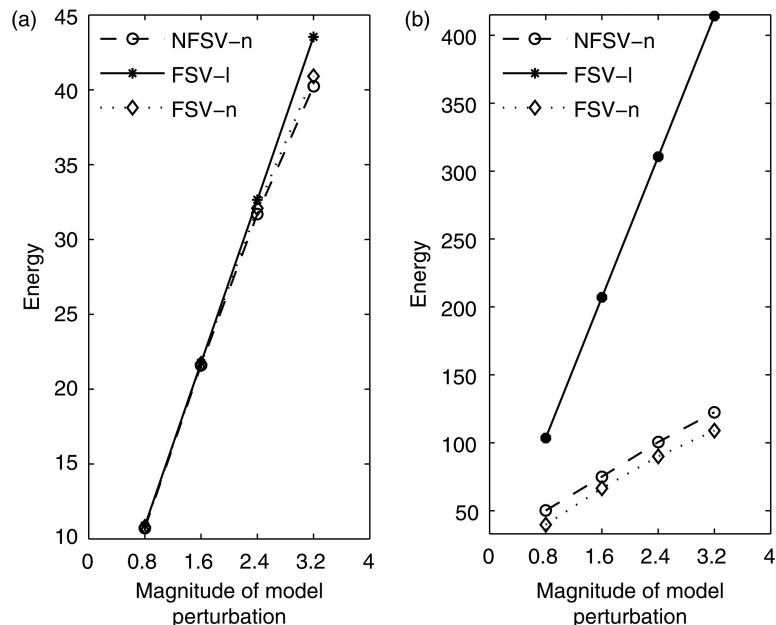
*Fig. 10.* The prediction errors caused by the NFSVs and the FSVs in the QG model (denoted by NFSV-n and FSV-n) and those caused by the FSVs in the linearised QG model (denoted by FSV-l). The horizontal axis represents the perturbation magnitudes, and the vertical axis describes the prediction errors in terms of the energy. (a) uses an optimisation time of 2 d, and (b) uses an optimisation time of 7 d.

includes regions of strong velocity shear, whereas the Ref-2 basic flow is approximately meridional with two deformation fields containing strong velocity shears. After computing the NFSVs and the corresponding linear FSVs,
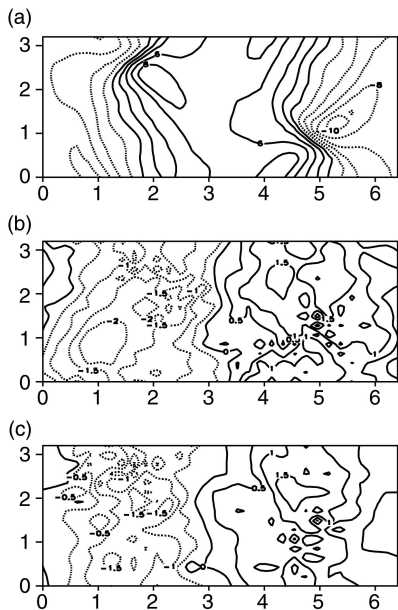


*Fig. 11.* Patterns of the stream functions of the prediction errors caused by (a) the FSV in the linearised QG model, (b) the FSV in the non-linear QG model, and (c) the NFSV in the non-linear QG model, where the optimisation is 7 d and the model perturbation magnitude is 1.6 in terms of the chosen norm.

our results show that the NFSVs have a phase in which the perturbation stream function tends to be contracted around the strong velocity shear region and tilted against the basic shear, presenting an asymmetric structure, especially for large tendency perturbations. Due to the absence of non-linearity, the FSVs for the Ref-1 tend to be more symmetric and cover a larger region than the NFSVs; in contrast, the FSVs for the Ref-2 only cover a slightly larger region compared to the NFSVs. NFSVs and FSVs also cause different prediction errors. For the Ref-1 basic flow, the magnitudes of the prediction errors caused by the FSVs in the non-linear QG model are significantly smaller than those caused by the FSVs in the linearised QG model for large tendency perturbations and/or long optimisation times. Non-linearities in the QG model then have a damping effect on the linear evolution of the tendency perturbations. Furthermore, we demonstrate that the prediction errors caused by the NFSVs in the non-linear QG model are obviously larger than those caused by the FSVs; therefore, the NFSVs are more applicable than the FSVs in numerically describing the optimal tendency perturbations. Physically, the constant tendency errors of the NFSV structure tend to reduce the non-linear damping effect and have the most potential for extracting energy from the basic flows and exhibiting optimal growth. For the Ref-2 basic flow, the non-linearity still suppresses the growth of the perturbation. Nevertheless, the NFSVs only slightly reduce this damping effect of the non-linearity. In this case, the evolutions of the prediction errors caused by the
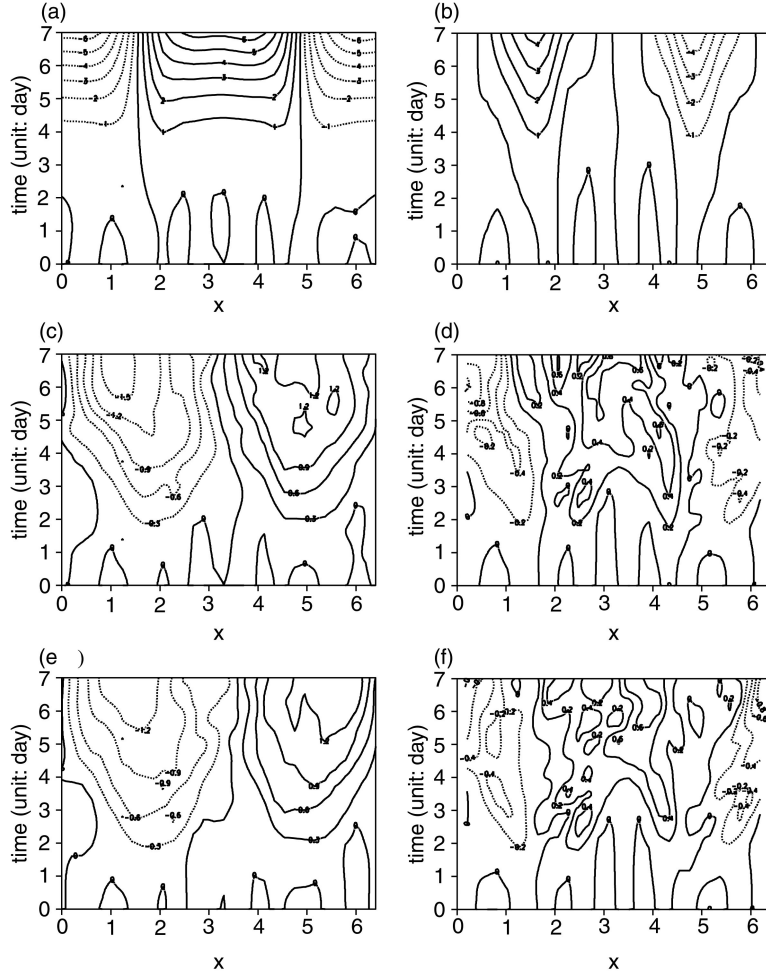
*Fig. 12.*    The meridional-mean of the stream function of prediction errors (left column) and the corresponding meridional wind (right column) caused by (a, b) the FSV in the linearised QG model, (c, d) the FSV in the non-linear QG model, and (e, f) the NFSV in the non-linear QG model. The optimisation time is 7 d, and the model perturbation magnitude is 1.6.

NFSVs and FSVs tend to have nearly the same amplitudes. It is therefore clear that the differences between the NFSVs and FSVs depend not only on the tendency perturbation magnitudes and/or optimisation times but also on the properties of the basic flows. Although we demonstrate that the Ref-2 NFSVs can be approximated by the corresponding FSVs, it is difficult for us to validate this argument in advance. Therefore, for a non-linear system, we would use the NFSVs rather than the FSVs to study the predictability of weather and climate. In particular, the differences between the NFSVs and FSVs for the Ref-1 basic flow show the usefulness of the NFSVs in revealing the effects of non-linearity on the predictability. The NFSV may therefore be a useful non-linear technique for exploring the predictability problems introduced by model errors.

In this study, we regard the constant tendency perturbations as constant tendency errors and compute the NFSVs, which may describe the model systematic errors that

cause progressively large prediction errors at a prediction time. In this case, the patterns of the NFSVs may allow us to find the regions in which the predictions are most sensitive to model systematic errors; therefore, the corresponding physical processes should be better described by the models. For example, we demonstrate in this article that the NFSVs have a phase in which the perturbation stream function tends to be contracted around the region of strong velocity shear; thus, the predictions generated by the two-dimensional QG model may be highly sensitive to the systematic errors of the model in the strong velocity shear region and the physical processes associated with this region should be better described by the models. In numerical predictions, we can determine an optimal external forcing to offset the model uncertainties by assimilating observations in this sensitive region. In addition, if we take the constant tendency perturbation as an external forcing with a particular
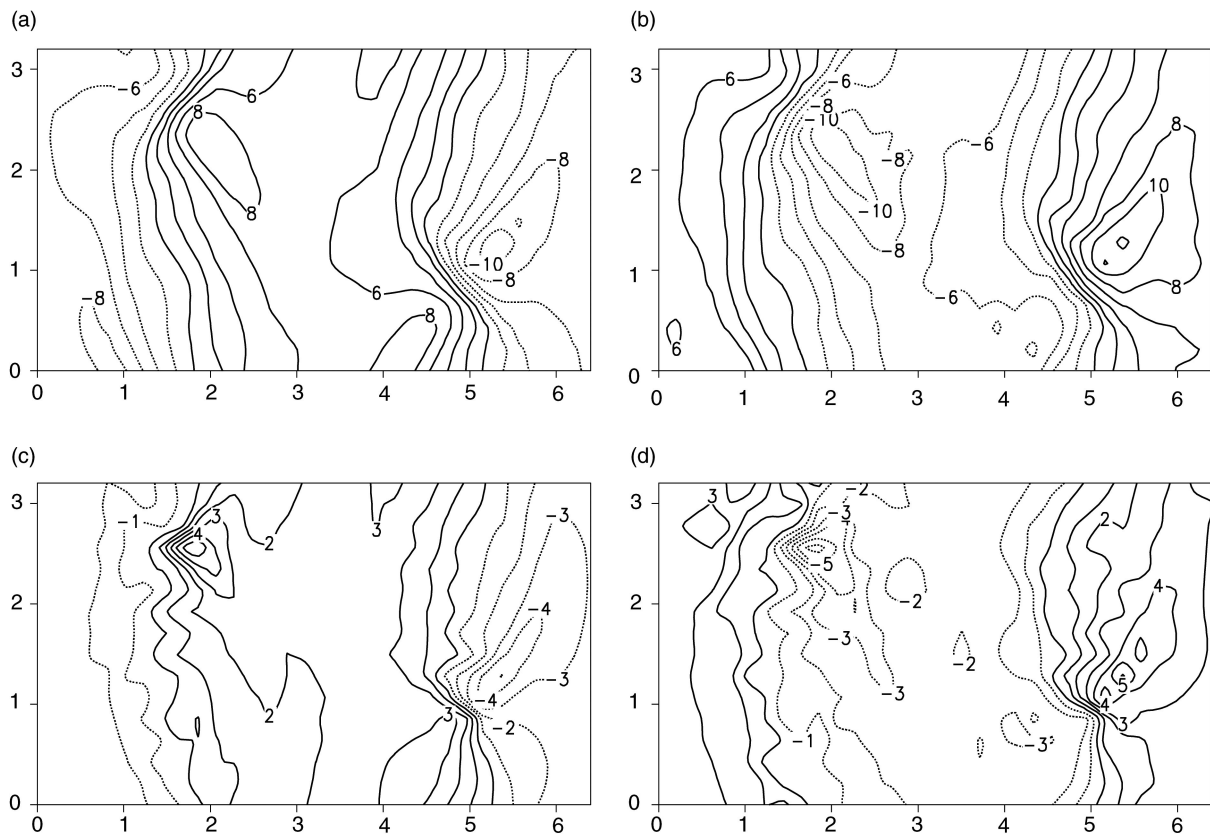
*Fig. 13.* The patterns of the stream function of the prediction error caused by the optimal forcing in the linearised QG model (left column) and the differences between the perturbation stream function induced by the optimal forcing in the non-linear QG model and in the linearised QG model (right column). The latter indicates the non-linear effects of perturbation advection on the perturbation growth. (a) and (b) are for the FSV, and (c) and (d) are for the NFSV. Both the FSV and NFSV are computed with an optimisation time of 7 d and a model perturbation magnitude of 1.6.

physical meaning, the NFSVs may describe the external forcings to which the weather and climate are significantly sensitive. Furthermore, if the NFSVs are superimposed on an external forcing with a particular physical meaning, they can be used to investigate the effects of the external forcing uncertainties on the prediction results. Of course, these theories should be realised by applying them to physical problems of interest. It is anticipated that the NFSV approach may play an important role in addressing the second type of predictability problem.

Riviere et al. (2008) demonstrated that the non-linear singular vectors (non-linear optimal initial perturbations) of a two-layer QG model tend to have larger meridional extensions than the singular vectors (linear optimal initial perturbations). However, we demonstrated in this article that the FSVs tend to have larger extensions than the NFSVs. It is clear that differences may exist in the effects of non-linearity on the tendency perturbations and initial perturbations. Despite the fact that the NFSVs provide useful information about the model systematic errors,

as demonstrated above, it is unclear why the effects of non-linearity on the model perturbations and initial perturbations are different. It may be challenging to address this question, which requires further investigation.

In numerical experiments, we examine the norms such as the $L_2$ norm, the potential enstrophy norm and the energy norm in measuring the magnitudes of the constant tendency perturbations and/or the prediction errors caused by them. However, in some cases, the algorithm adopted here may not capture the maximum of the cost function associated with the NFSVs and allow the resultant NFSVs to comprise many small scales; alternatively, the resultant NFSVs do not possess clear physical meanings. In this article, after several attempts, we finally chose the $L_2$ norm to measure the tendency perturbation magnitudes and the energy norm to measure the evolution of the prediction error caused by constant tendency errors and reveal the effects of non-linearity on the NFSV structures. Furthermore, we are almost able to provide a physical explanation for the structure of the NFSVs and why the NFSVs cause a

larger perturbation growth than the FSVs. Therefore, the norm adopted in this study may be physically relevant.

To study the effects of the model parameter errors on predictability, Mu et al. (2010) proposed the CNOP-P approach to solve the optimal model parameter perturbation. Although the NFSV used to describe the model systematic errors can be derived using the CNOP-P approach, the constant tendency perturbation requires a different reasoning than that applied to a specific model parameter perturbation. In fact, the perturbation superimposed on a model parameter represents a parameter error, whereas the constant tendency perturbation may describe the combined systematic errors of physical processes that are not explicitly or not entirely correctly described by the model equations. Furthermore, the NFSVs allow us to investigate the spatial structures of the model systematic errors. Additionally, the NFSVs can be generalised to be time-dependent and can then describe much realistic model errors. Therefore, the NFSVs are much more general than the optimal parameter perturbations associated with the CNOP-P approach.

Predictability studies are challenging due to the non-linearity and complexity of atmospheric and oceanic motions. In particular, the predictability problems associated with model errors pose great difficulties due to the lack of effective approaches. In this article, we propose the NFSV approach to address certain model errors. It is expected that this approach will play an important role in studying the predictability problems associated with model errors and provide ideas for improving weather and climate predictabilities.

## 6. Acknowledgements

## 7. Appendix

**The first-order variational of $J_1(f)$**

$$
\begin{aligned}
-\delta J_1(\boldsymbol{f}) =& <\boldsymbol{u}(\tau), \delta \boldsymbol{u}(\tau)> \\
&- \int_0^\tau <\lambda_1(t), \frac{\partial \delta \boldsymbol{u}}{\partial t} - \frac{\partial F(\boldsymbol{U}+\boldsymbol{u})}{\partial u}\delta \boldsymbol{u} - \delta \mathbf{f}> dt \\
&- \int_0^\tau <\lambda_2(t), \frac{\partial \delta \mathbf{f}}{\partial t}> dt.
\end{aligned}
$$
(A1)

With integration by parts, we can obtain

$$
\begin{aligned}
\int_0^\tau <\lambda_1(t), \frac{\partial \delta \boldsymbol{u}}{\partial t}> dt =& \int_0^\tau \frac{\partial}{\partial t} <\lambda_1(t), \delta \boldsymbol{u}> dt \\
&- \int_0^\tau <\frac{\partial \lambda_1(t)}{\partial t}, \delta \boldsymbol{u}> dt \\
=& <\lambda_1(\tau), \delta \boldsymbol{u}(\tau)> - <\lambda_1(0), \delta \boldsymbol{u}(0)> \\
&- \int_0^\tau <\frac{\partial \lambda_1(t)}{\partial t}, \delta \boldsymbol{u}> dt
\end{aligned}
$$

and

$$
\begin{aligned}
\int_0^\tau <\lambda_2(t), \frac{\partial \delta \mathbf{f}}{\partial t}> dt =& \int_0^\tau \frac{\partial}{\partial t} <\lambda_2(t), \delta \mathbf{f}> dt \\
&- \int_0^\tau <\frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f}> dt = <\lambda_2(\tau), \delta \boldsymbol{f}(\tau)> \\
&- <\lambda_2(0), \delta \boldsymbol{f}(0)> \\
&- \int_0^\tau <\frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f}> dt = <\lambda_2(\tau), \delta \boldsymbol{f}> \\
&- <\lambda_2(0), \delta \boldsymbol{f}> - \int_0^\tau <\frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f}> dt
\end{aligned}
$$

Then, we derive $\delta J_1$ as follows:

$$
\begin{aligned}
-\delta J_1 =& \int_0^\tau <\frac{\partial \lambda_1}{\partial t}, \delta \boldsymbol{u}> dt + <\boldsymbol{u}(\tau) - \lambda_1(\tau), \delta \boldsymbol{u}(\tau)> \\
&+ <\lambda_1(0), \delta \boldsymbol{u}(0)> \\
&+ \int_0^\tau <\lambda_1(t), \left[\frac{\partial F(\boldsymbol{U}(t)+\boldsymbol{u}(t))}{\partial \boldsymbol{u}}\right]\delta \boldsymbol{u}> dt \\
&+ \int_0^\tau <\lambda_1(t), \delta \mathbf{f}> dt + \int_0^\tau <\frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f}> dt \\
&+ <-\lambda_2(\tau), \delta \boldsymbol{f}> + <\lambda_2(0), \delta \boldsymbol{f}> \\
=& \int_0^\tau <\frac{\partial \lambda_1}{\partial t}, \delta \boldsymbol{u}> dt + <\boldsymbol{u}(\tau) - \lambda_1(\tau), \delta \boldsymbol{u}(\tau)> \\
&+ <\lambda_1(0), \delta \boldsymbol{u}(0)> \\
&+ \int_0^\tau <\left[\frac{\partial F(\boldsymbol{U}(t)+\boldsymbol{u}(t))}{\partial \boldsymbol{u}}\right]^* \lambda_1(t), \delta \boldsymbol{u}> dt \\
&+ \int_0^\tau <\lambda_1(t), \delta \mathbf{f}> dt \\
&+ \int_0^\tau <\frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f}> dt + <0 - \lambda_2(\tau), \delta \boldsymbol{f}> \\
&+ <\lambda_2(0), \delta \boldsymbol{f}> .
\end{aligned}
$$
(A2)

where $[\cdot]^*$ denotes an adjoint operator. Then, we obtain

$$
\begin{aligned}
-\delta J_1 =& \int_0^\tau <\frac{\partial \lambda_1}{\partial t} + \left[\frac{\partial F(\boldsymbol{U}(t)+\boldsymbol{u}(t))}{\partial \boldsymbol{u}}\right]^* \lambda_1(t), \delta \boldsymbol{u}> dt \\
&+ <\boldsymbol{u}(\tau) - \lambda_1(\tau), \delta \boldsymbol{u}(\tau)> + <\lambda_1(0), \delta \boldsymbol{u}(0)> \\
&+ \int_0^\tau <\frac{\partial \lambda_2(t)}{\partial t} + \lambda_1(t), \delta \mathbf{f}> dt + <0 - \lambda_2(\tau), \delta \boldsymbol{f}> \\
&+ <\lambda_2(0), \delta \boldsymbol{f}> .
\end{aligned}
$$
(A3)

## References

Arakawa, A. 1966. Computation design for long-term numerical integrations for the equations of atmospheric motions. *J. Comput. Phys.* **1**, 119–143.

Barkmeijer, J. 1996. Constructing fast-growing perturbations for the nonlinear regime. *J. Atmos. Sci.* **53**, 2838–2851.

Barkmeijer, J., Iversen, T. and Palmer, T. N. 2003. Forcing singular vector and other sensitivity model structures. *Q. J. Roy. Meteor. Soc.* **129**, 2401–2423.

Birgin, E. G., Martinez, J. M. and Raydan, M. 2000. Nonmonotone spectral projected gradient methods on convex sets, SIAM. *J. Optim.* **10**, 1196–1211.

Buizza, R., Miller, M. and Palmer, T. N. 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. Roy. Meteor. Soc.* **125**, 2887–2908.

Cai, M., Kalnay, E. and Toth, Z. 2003. Bred vectors of the Zebiak-Cane model and their potential application to ENSO predictions. *J. Clim.* **16**, 40–55.

D'Andrea, F. and Vautard, R. 2000. Reducing systematic errors by empirically correcting model errors. *Tellus.* **52A**, 21–41.

Duan, W. S., Liu, X. C., Zhu, K. Y. and Mu, M. 2009. Exploring the initial errors that cause a significant "spring predictability barrier" for El Nino events. *J. Geophys. Res.* **114**, C04022. DOI: 10.1029/2008JC004925.

Duan, W. S. and Mu, M. 2006. Investigating decadal variability of El Nino-Southern Oscillation asymmetry by conditional nonlinear optimal perturbation. *J. Geophys. Res.* **111**, C07015. DOI: 10.1029/2005JC003458.

Duan, W. S., Mu, M. and Wang, B. 2004. Conditional nonlinear optimal perturbation as the optimal precursors for El Nino-Southern Oscillation events. *J. Geophys. Res.* **109**, D23105.

Duan, W. S., Xu, H.and Mu, M. 2008. Decisive role of nonlinear temperature advection in El Nino and La Nina amplitude asymmetry. *J. Geophys. Res.* **113**, C01014. DOI: 10.1029/2006JC003974.

Dudhia, J. 1993. A nonhydrostatic version of the Penn State-NCAR mesoscale model: validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.* **121**, 1493–1513.

Farrell, B. 1989. Optimal excitation of baroclinic waves. *J. Atmos. Sci.* **46**, 1193–1206.

Farrell, B. and Ioannou, P. J. 2005. Distributed forcing of forecast and assimilation systems. *J. Atmos. Sci.* **62**, 460–475.

Gutowski, W. J. Jr. 1985. Baroclinic adjustment and midlatitude temperature profiles. *J. Atmos. Sci.* **42**, 1733–1745.

Hamill, T. M., Snyder, C. and Morss, R. E. 2000. A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Mon. Wea. Rev.* **128**, 1835–1851.

Houtekamer, P. L. and Derome, J. 1995. Methods for ensemble prediction. *Mon. Wea. Rev.* **123**, 2181–2196.

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H. L. 1996. A system simulation approach to ensemble prediction. *Mon. Wea. Rev.* **124**, 1225–1242.

Jiang, Z. N. and Mu, M. 2009. A comparisons study of the methods of conditional nonlinear optimal perturbations and singular vectors in ensemble prediction. *Adv. Atmos. Sci.* **26**(3), 465–470.

Lacarra, J. and Talagrand, O. 1988. Short-range evolution of small perturbations in a barotropic model. *Tellus.* **40A**, 81–95.

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

Liu, D. C. and Nocedal, J. 1989. On the limited memory method for large scale optimization. *Math. Program. B.* **45**(3), 503–528.

Lorenz, E. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus.* **17**, 321–333.

Lorenz, E. N. 1975. *Climate Predictability: The Physical Basis of Climate Modelling.* Global Atmospheric Research Programme Publication Series, No. 16, World Meteorology Organization, Geneva, pp. 132–136.

Mak, M. and Cai, M. 1989. Local barotropic instability. *J. Atmos. Sci.* **46**, 3289–3311.

Moore, A. M. and Kleeman, R. 1996. The dynamics of error growth and predictability in a coupled model of ENSO. *Q. J. Roy. Meteor. Soc.* **122**, 1405–1446.

Moore, A. M. and Kleeman, R. 1999. Stochastic forcing of ENSO by the intraseasonal oscillation. *J. Clim.* **12**, 1199–1220.

Mu, M., Duan, W. S. and Wang, B. 2003. Conditional nonlinear optimal perturbation and its applications. *Nonlin. Process. Geophys.* **10**, 493–501.

Mu, M., Duan, W. S. and Wang, B. 2007a. Season-dependent dynamics of nonlinear optimal error growth and El Nino-Southern Oscillation predictability in a theoretical model. *J. Geophys. Res.* **112**, D10113. DOI: 10.1029/2005JD006981.

Mu, M., Duan, W., Wang, Q. and Zhang, R. 2010. An extention of conditional nonlinear optimal perturbation approach and its applications. *Nonlin. Process. Geophys.* **17**, 211–220.

Mu, M. and Shepherd, T. G. 1994. Nonlinear stability of multi-layer quasi-geostrophic flow. *Geophys. Astrophys. Fluid. Dyn.* **75**, 21–37.

Mu, M., Sun, L. and Dijkstra, H. 2004. The sensitivity and stability of the ocean's thermocline circulation to finite amplitude freshwater perturbations. *J. Phys. Oceanogr.* **34**, 2305–2315.

Mu, M., Xu, H. and Duan, W. 2007b. A kind of initial errors related to "spring predictability barrier" for El Nino events in Zebiak-Cane model. *Geophys. Res. Lett.* **34**, L03709. DOI: 10.1029/2006GL027412.

Mu, M. and Zeng, Q. C. 1991. New development on existence and uniqueness of solutions to some models in atmospheric dynamics. *Adv. Atmos. Sci.* **8**, 383–398.

Mu, M., Zhou, F. and Wang, H. 2009. A method for identifying the sensitive areas in targeted observations for tropical cyclone prediction: conditional nonlinear optimal perturbation. *Mon. Wea. Rev.* **137**(5), 1623–1639.

Mylne, K. R., Evans, R. E. and Clark, R. T. 2002. Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Q. J. Roy. Meteor. Soc.* **128**, 361–384.

Nakamura, N. 1999. Baroclinic–barotropic adjustments in a meridionally wide domain. *J. Atmos. Sci.* **56**, 2246–2260.

Oortwijn, J. and Barkmeijer, J. 1995. Perturbations that optimally trigger weather regimes. *J. Atmos. Sci.* **52**, 3932–3944.

Palmer, T. N., Molteni, F., Mureau, R., Buizza, R., Chapelet, P. and co-authors. 1992. Ensemble prediction. ECMWF Research

Department Tech. Memo. 188, ECMWF, Shinfield Park, Reading, England. 45 pp.

Pedlosky, J. 1979. *Geophysical Fluid Dynamics*. Springer-Verlag, New York.

Powell, M. J. D. 1982. *VMCWD: A FORTRAN Subroutine for Constrained Optimization*. DAMTP Report 1982/NA4, University of Cambridge, England.

Riviere, O., Lapeyre, G. and Talagrand, O. 2008. Nonlinear generalization of singular vectors: behavior in a baroclinic unstable flow. *J. Atmos. Sci.* **65**, 1896–1911.

Riviere, O., Lapeyre, G. and Talagrand, O. 2009. A novel technique for nonlinear sensitivity analysis: application to moist predictability. *Q. J. Roy. Meteor. Soc.* **135**, 1520–1537.

Roads, J. O. 1987. Predictability in the extended range. *J. Atmos. Sci.* **44**, 1228–1251.

Samelson, R. G. and Tziperman, E. 2001. Instability of the chaotic ENSO: the growth-phase predictability barrier. *J. Atmos. Sci.* **58**, 3613–3625.

Shutyaev, V., Le Dimet, V. F.-X. and Gejadze, I. 2008. On optimal solution error covariances in variational data assimilation. *Russ. J. Numer. Anal. Math. Modelling*. **23**, 185–206.

Simmons, A. and Hollingsworth, A. 2002. Some aspects of the improvement in skill of numerical weather prediction. *Q. J. Roy. Meteor. Soc.* **128**, 647–677.

Tanguay, M., Bartello, P. and Gauthier, P. 1995. Four dimensional data assimilation with a wide range of scales. *Tellus*. **47A**, 974–997.

Tennekes, H. 1991. *Karl Popper and the Accountability of Numerical Forecasting. In New Developments in Predictability*. ECMWF Workshop Proceedings, ECMWF, Shinfield Park, Reading, UK.

Terwisscha, A. D. and Dijkstra, H. 2008. Conditional nonlinear optimal perturbations of the double-gyre ocean circulation. *Nonlin. Process. Geophys.* **15**, 727–734.

Thompson, C. J. 1998. Initial conditions for optimal growth in a coupled ocean-atmosphere model of ENSO. *J. Atmos. Sci.* **55**, 537–557.

Thompson, P. 1957. Uncertainty of the initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*. **9**, 275–295.

Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.* **127**, 3297–3318.

Tung, K. K. 1981. Barotropic instability of zonal flows. *J. Atmos. Sci.* **38**, 308–321.

Wu, X. G. and Mu, M. 2009. Impact of horizontal diffusion on the nonlinear stability of thermohaline circulation in a modified box model. *J. Phys. Oceanogr.* **39**, 798–805.

Xue, Y., Cane, M. A. and Zebaik, S. E. 1997. Predictability of a coupled model of ENSO using singular vector analysis. Part I: optimal growth in seasonal background and ENSO cycles. *Mon. Wea. Rev.* **125**, 2043–2056.

Zebiak, S. E. and Cane, A. 1987. A model El Nino-Southern Oscillation. *Mon. Wea. Rev.* **115**, 2262–2278.