Some Aspects of Sensitivity Analysis in Variational Data Assimilation for Coupled Dynamical Systems

Sergei Soldatenko, Peter Steinle, Chris Tingwell, and Denis Chichkine

1 Centre for Australian Climate and Weather Research, 700 Collins Street, Melbourne, VIC 3008, Australia
2 University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1

Correspondence should be addressed to Sergei Soldatenko; s.soldatenko@bom.gov.au

Received 18 December 2014; Revised 24 February 2015; Accepted 11 March 2015

1. Introduction

The earth system consists of several interactive dynamical subsystems and each of them covers a broad temporal and spatial spectrum of motions and physical processes. The components of the earth system have many differences in their physical properties, structure, and behavior but are linked together by fluxes of momentum and mass as well as sensible and latent heat. All of these subsystems interact with each other in different ways and can be strongly or weakly coupled. Prediction of future state of the earth system and its components is one of the most important problems of modern science. The most significant progress has been achieved in the forecasting of the atmosphere via numerical models, which describe the dynamical and physical processes in the earth’s gaseous envelope. It is clear that further improvement of forecasts can be pursued via the development of coupled modeling systems that primarily combine the atmosphere and the ocean and describe the interactions between these two systems. Since numerical weather prediction systems calculate a future state of the atmosphere and ocean by integrating a set of partial differential equations that describe the fluid dynamics and thermodynamics, initial conditions that accurately represent the state of the atmosphere and ocean at a certain initial time must be formulated. Numerical weather prediction systems use data assimilation procedures to estimate initial conditions for forecasting models from observations. Data assimilation remains one of the key issues not only in the numerical weather prediction (NWP) but also in other geophysical sciences.

One of the most advanced and effective data assimilation techniques is four-dimensional variational data assimilation (4D-Var). In particular, the weather forecasts produced by the ACCESS (Australian Community Climate and Earth System Simulator) at the Bureau of Meteorology use 4D-Var in the incremental formulation developed at the Met Office [1, 2].
A comprehensive historical review and current status of 4D-Var are presented, for example, in [3–7]. In our view point, some papers and books, such as [8–17], have contributed significantly to the development of mathematical foundation, theory, and practice of 4D-Var.

In general, the main objective of 4D-Var is to define, as perfectly as possible, the state of a dynamical system by combining, in statistically optimal manner, the observations of state variables of a real physical system together with certain prior information. In NWP this prior information is usually referred to as the background. Mathematically, 4D-Var procedures are formulated as an optimization problem, in which the initial condition plays the role of control vector and model equations are considered as constraints. While the theory of variational data assimilation is well established, there are a number of issues with practical implementation that require additional consideration and study. The performance of 4D-Var schemes depends on their key information components, such as the available observations, estimates of the observation, and background error covariances that are quantified by the corresponding matrices, as well as the background state. All of those information components strongly impact the accuracy of calculated initial conditions, thus influencing the forecast quality. Thus, it is important to estimate the sensitivity of a certain forecast aspect with respect to variations in the observational data, background information, and error statistics. This problem is usually formulated within the framework of the adjoint-based sensitivity analysis (e.g., [18–29]).

However, an adjoint model used to calculate sensitivity functions is derived from a linearized forward propagation model and contains numerous input parameters. Consequently, linearization of strongly nonlinear NWP models and also uncertainties in their numerous parameters generate errors in the initial conditions obtained by data assimilation systems. The influence of linearization and parameter uncertainties on the results of data assimilation can be studied, ideally for each particular NWP model, using sensitivity analysis [30, 31]. Even more problems arise when considering coupled 4D-Var data assimilation schemes since the atmosphere and ocean have very different physical properties and time-space spectrum of motions generating initialization shock. Several coupling strategies are being developed for use in NWP systems; however, all of them introduce issues that require additional detailed consideration. For example, the influence of coupling strength on the initial conditions obtained by 4D-Var procedures is one such issue that is important to study and analyze.

Good practice in the development of NWP models and data assimilation systems requires evaluating the confidence in the model. In this context, it is important to estimate the influence of parameter variations on system dynamics and to find those parameters that have the largest impact on system behaviour. Sensitivity analysis, which is an essential element of model building and quality assurance, addresses this very important issue.

The exploration of coupled 4D-Var systems, parameter estimation, and sensitivity analysis require considerable computational resources. For simple enough low-order coupled models, the computational cost is minor and, for that reason, models of this class are widely used as simple test instruments to emulate more complex systems. In this paper, we describe a coupled nonlinear dynamical system, which is composed of fast (the “atmosphere”) and slow (the “ocean”) versions of the well-known Lorenz [32] model. This low-order coupled system allows us to mimic the atmosphere-ocean system and therefore serves as a key element of a theoretical and computational framework for the study of various aspects of coupled 4D-Var procedures [33, 34]. Under certain conditions the Lorenz model exhibits a chaotic behaviour and using conventional methods of sensitivity analysis can be questionable in terms of interpretation of the obtained results [35–37]. The “shadowing” method [36, 37] for estimating the system sensitivity to variations in its parameters allows us to calculate the average along the trajectory sensitivities and therefore to make a clear conclusion with respect to the system sensitivity to its parameters. This method is based on the pseudoorbit shadowing in dynamical systems [38, 39]. Calculated sensitivity coefficients obtained via conventional methods and the “shadowing” approach are presented in the paper.

We also succinctly consider commonly used techniques for sensitivity analysis and parameter estimations of dynamical systems and study the influence of coupling strength parameter on the dynamical behaviour of the coupled system using Lyapunov characteristic exponent analysis. It was found that the coupling strength parameter strongly affects the system dynamics both quantitatively and qualitatively. This fact should be taken into consideration when choosing the coupling strategy in data assimilation systems.

### 2. Low-Order Coupled Dynamical System

In this section we consider a low-order coupled nonlinear dynamical system obtained by coupling of two versions of the original Lorenz model (L63) [32] with distinct time scales, which differ by a factor $\varepsilon$ (e.g., [33, 34]):

$$
\dot{x} = \sigma (y - x) - c (aX + k),
$$

$$
\dot{y} = r x - y - x z + c (aY + k),
$$

$$
\dot{z} = x y - b z + c_z Z,
$$

$$
\dot{X} = \varepsilon \sigma (Y - X) - c (x + k),
$$

$$
\dot{Y} = \varepsilon (rX - Y - aXZ) + c (y + k),
$$

$$
\dot{Z} = \varepsilon (aXY - bZ) - c_z z,
$$

where lower case letters represent the fast subsystem and capital letters the slow subsystem, $\sigma$, $r$, and $b$ are the parameters of L63 model, $c$ is a coupling strength parameter for the $x$ and $y$ variables, $c_z$ is a coupling strength parameter for $z$, $k$ is an “uncentering” parameter, and $a$ is a parameter representing the amplitude scale factor. The value $a = 1$ indicates that two systems have the same amplitude scale. Thus, the state vector of the coupled model (1a) and (1b) is $x = (x, y, z, X, Y, Z)^T$ and the model parameter vector is...
\( \alpha = (\sigma, r, b, a, c, k, e)^T \). Without loss of generality, we can assume that \( a = 1 \), \( k = 0 \), and \( c = c_e \), then \( \alpha = (\sigma, r, b, c, e)^T \) and the system (1a) and (1b) can be rewritten in the operator form:

\[
\frac{d\mathbf{x}}{dt} = (\mathbf{L} + \mathbf{Q}) \mathbf{x}, \tag{2}
\]

where the nonlinear uncoupled operator \( \mathbf{L} \) and linear coupled operator \( \mathbf{Q} \) are represented by the following matrices:

\[
\mathbf{L} = \begin{bmatrix}
-\sigma & \sigma & 0 \\
 r & -1 & -x \\
 0 & x & -b \\
\end{bmatrix},
\]

\[
\mathbf{Q} = \begin{bmatrix}
-\varepsilon \sigma & \varepsilon a & 0 \\
 0 & \varepsilon r & -\varepsilon X \\
 \varepsilon X & -\varepsilon b & 0 \\
\end{bmatrix}. \tag{3}
\]

The unperturbed parameter values are taken as

\[
\sigma^0 = 10, \quad r^0 = 28, \quad b^0 = \frac{8}{3}, \quad \varepsilon^0 = 0.1, \quad c^0 \in [0.1; 1.2]. \tag{4}
\]

Chosen values of \( \sigma, r, \) and \( b \) correspond to chaotic behaviour of the L63 model. For \( \sigma = 10 \) and \( b = 8/3 \) the critical value of parameter \( r \) is 24.74, which means that any value of \( r \) larger than 24.74 induces chaotic behaviour [32]. The parameter \( \varepsilon = 0.1 \) indicates that the slow system is 10 times slower than the fast system.

The coupling strength parameter \( c \) plays a very important role in qualitative changes in the system dynamics since this parameter controls the interactions between fast and slow subsystems. Qualitative changes in the dynamical properties of a system can be detected by determining and analyzing the system’s spectrum of Lyapunov exponents, which characterize the average rate of exponential divergence (or convergence) of nearby trajectories in the phase space. In the analysis of coupled dynamical systems we are dealing with conditional Lyapunov exponents that are normally used to characterize the synchronization with coupled systems.

System (2) has six distinct exponents. If the parameter \( c \) tends to zero, then system (2) has two positive, two zero, and two negative Lyapunov exponents. The influence of coupling strength parameter \( c \) on the two largest conditional Lyapunov exponents is illustrated in Figure 1. The numerical experiments demonstrated that those, initially positive, exponents decrease monotonically with an increase in the parameter \( c \). At about \( c = 0.8 \) they approach the x-axis and at about

\( c \approx 0.95 \) negative values. Thus, for \( c > 0.95 \), the dynamics of both fast and slow subsystems become phase-synchronous [40]. For \( c > 1.0 \), a limit circle dynamical regime is observed since all six exponents become negative.

Apart from Lyapunov exponents, autocorrelation functions (ACFs) enable one to distinguish between regular and chaotic processes and to detect transition from order to chaos. In particular, for chaotic motions, ACF decreases in time, in many cases exponentially, while for regular motions, ACF is unchanged or oscillating. In general, however, the behaviour of ACFs of chaotic oscillations is frequently very complicated and depends on many factors (e.g., [41, 42]). Autocorrelation functions can also be used to define the so-called typical time memory (typical timescale) of a process [43]. If it is positive, ACF is considered to have some degree of persistence: a tendency for a system to remain in the same state from one moment in time to the next. The ACF for a given discrete dynamic variable \( \{x_i\}_{i=0}^{N-1} \) is defined as

\[
C(s) = \langle x_i x_{i+s} \rangle - \langle x_i \rangle \langle x_{i+s} \rangle, \tag{5}
\]

where the angular brackets denote ensemble averaging. Assuming time series originates from a stationary and ergodic process, ensemble averaging can be replaced by time averaging over a single normal realization:

\[
C(s) = \langle u_i u_{i+s} \rangle - \langle u_i \rangle^2. \tag{6}
\]

Signal analysis commonly uses the normalized ACF, defined as \( R(s) = C(s)/C(0) \). ACF plots for realizations of dynamic variables \( x \) and \( X \), and \( z \) and \( Z \) calculated for different values of the coupling strength parameter \( c \) are presented in Figures 2 and 3, respectively. For relatively small parameter \( c \) (\( c < 0.4 \)), the ACFs for both \( x \) and \( X \) variables decrease fairly rapidly to zero, consistently with the chaotic behaviour of the coupled system. However, as expected, the rate of decay of
the ACF of the slow variable $X$ is less than that of the fast variable $x$. The ACF envelopes for variables $z$ and $Z$ also decay almost exponentially from the maximum to zero. For coupling strength parameter on the interval $0.4 < c < 0.6$ the ACF of the fast variable $x$ becomes smooth and converges to zero. At the same time, the envelopes of the ACFs of variables $X$, $z$, and $Z$ demonstrate a fairly rapid fall, indicating the chaotic behaviour. As the parameter $c$ increases, the ACFs become periodic and their envelopes decay slowly with time, indicating transition to regularity. For $c > 0.8$ calculated ACFs show periodic signal components.

3. The Basics of Four-Dimensional Variational Data Assimilation

Atmospheric models used for operational NWP are mainly deterministic and derived from a set of multidimensional nonlinear differential equations in partial derivatives, which are the equations of fluid dynamics and thermodynamics that describe atmospheric processes and atmosphere-underlying surface interactions. A generic atmospheric model can be represented by the following continuous autonomous dynamical system:

$$\frac{dx(t)}{dt} = f(x(t), a(t)), \quad t \in [0, \tau] = \mathcal{F}, \quad x(0) = x_0. \quad (7)$$

Here $x(t)$ is the state vector belonging to a Hilbert space $\mathcal{X}$, $a \in \mathcal{P}$ is a parameter vector, where $\mathcal{P}$ is a Hilbert space (the space of the model parameters), $f$ is a nonlinear vector-valued function, such that $f: \mathcal{X} \times \mathcal{P} \times \mathcal{F} \rightarrow \mathcal{X}$, and $x_0$ is a given vector function. This infinite-dimensional model has to be truncated by some means to finite-dimensional approximate model, for which a solution can be sought numerically. Applying either a projection onto a finite set of basic functions or a discretization in time and space, one can derive the discrete atmospheric model which can be represented as $s$ discrete nonlinear dynamical system given by the equation

$$x_{i+1} = M_{i+1}(x_i) + \epsilon_i^m, \quad (8)$$

where $x_i \in \mathbb{R}^n$ is the $n$-dimensional state vector representing the complete set of the model variables that determine the internal state of the atmospheric model at time $t_i$, $M_{i+1}$ is the nonlinear operator that indirectly contains model parameters and propagates the state vector from time $t_i$ to time $t_{i+1}$ for $i = 0, \ldots, N - 1$, and $\epsilon_i^m$ is the model errors. It is usually assumed that model (8) is “perfect” ($\epsilon_i^m = 0$), that is the forecast has no errors if the initial condition is perfect. In this case, given the model operator and the initial condition $x_0$, (8) uniquely specifies the orbit of the dynamical system. Let $y_i^0 \in \mathbb{R}^m$ be the $m$-dimensional vector of observations measured at a discrete time $t_i, i = 0, \ldots, N$ that are linked to the system state via the following equation:

$$y_i^0 = \mathcal{H}_i(x_i) + \epsilon_i^0, \quad (9)$$

where $\mathcal{H}_i : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the nonlinear observation operator that maps the state vector to observation space. It is usually assumed that the observation errors $\epsilon_i^0$ are unbiased, serially uncorrelated, and normally distributed with known covariance matrices $R_i \in \mathbb{R}^{m \times m}$.

Suppose that at the initial time the prior (background) model state $x_i^0$ is known and represents the “best” estimate of the “true” state $x_i$, before any observations are taken. This background state is provided by a previous forecast. It is assumed that $x_i^0$ has unbiased and normally distributed errors $\epsilon_i^b$ with known covariance matrix $B_i \in \mathbb{R}^{m \times m}$:

$$x_i^0 = x_0 + \epsilon_i^b. \quad (10)$$

Given the observations $y_i^0$ at time $t_i$, the corresponding observation error covariance matrices $R_i$ ($i = 0, \ldots, N$), the background initial state $x_i^0$, and the error covariance matrix $B_i$, the 4D-Var data assimilation seeks to minimize, with respect to $x_i$, a certain cost function $J(x)$ expressing the “distance” between observations and corresponding model state using the model equations as constraints:

$$x_i^0 = \arg \min J(x) \quad (11)$$

subject to $x$ satisfying the set of the “ideal” ($\epsilon_i^m = 0$) model equations with initial state $x_0$:

$$x_{i+1} = M_{i+1}(x_i), \quad i = 1, \ldots, N. \quad (12)$$

The 4D-Var cost function is usually written as (e.g., [17, 18]):

$$J(x_i) = \frac{1}{2}(x_i - x_i^0)^T B_i^{-1}(x_i - x_i^0) + \frac{1}{2} \sum_{i=0}^N (\mathcal{H}_i(x_i) - y_i^0)^T R_i^{-1}(\mathcal{H}_i(x_i) - y_i^0). \quad (13)$$

The optimization problem (12) is nonlinear with strong constraints and an iterative minimization algorithm (e.g., gradient-based technique) is required to obtain the solution. The gradient of the cost function (13) is as follows:

$$\nabla_x J(x_i) = B_i^{-1}(x_i - x_i^0) + \sum_{i=0}^N M_i^T H_i^T R_i^{-1}(\mathcal{H}_i(x_i) - y_i^0), \quad (14)$$

where $M_i^T$ is the adjoint of the linearized model operator $M_i(x_i)$ and $H_i^T$ is the adjoint of the linearized observation operator $H_i(x_i)$. If the model is not “perfect” then we need to take into account the model errors $\epsilon_i^m$, which are sometimes taken as Gaussian noise:

$$\epsilon_i^m \in \mathcal{N}(0, Q), \quad (15)$$

where $Q$ is a model error covariance matrix. Thus, we obtain the weakly constrained 4D-Var data assimilation and the following term:

$$J_m = \frac{1}{2} \sum_{i=1}^N (x_i - M_{i-1,j}(x_i - 1)^T Q_i^{-1}(x_i - M_{i-1,j}(x_i - 1)^T) \quad (16)$$
Figure 2: Autocorrelation functions for dynamic variables $x$ and $X$ for different parameter $c$. 
Figure 3: Autocorrelation functions for dynamic variables $z$ and $Z$ for different parameter $c$. 
should be added to the right-hand side of the cost function (13) [6, 19]. It is important to make the following comments. While NWP has significantly improved over the last several decades, weather forecasts are still intrinsically uncertain. This is because mathematical models used in NWP have several sources of uncertainty and therefore a number of sources of errors. The first one is an intrinsic uncertainty due to chaotic nature of the system. The second one is a structural uncertainty, which is how the model itself represents the physical processes incorporated into it. It is important to underline that our knowledge of the earth system is always imperfect and, therefore, we can only theoretically design the “ideal” model. However, improving model physics demonstrated that even theoretically well-posed models failed to accurately simulate and predict the dynamics of real atmospheric processes. This is because numerical models have a parametric uncertainty (how accurate model parameters are) initial and boundary uncertainty (are initial and boundary conditions known precisely) and, in addition, numerical errors. All of those uncertainties (errors) impair the weather forecast accuracy and limit the time horizon of accurate NWP. Current time horizon of synoptic-scale NWP is several days.

The climate study has significantly longer time horizon: several decades. The sensitivity analysis of the climate system is associated with stability of characteristics of climatic model attractors with respect to perturbations in model parameters. A priori estimation of the behaviour of state vector, when the perturbations in the climatic model parameters tend to zero, is generally an unresolved problem since it is not known whether or not the invariant measure of climate modeling system is continuous with respect to the small perturbations in the differential matrix operator of the numerical model. Indeed, at certain model parameter values, different bifurcation can occur in the system phase space. Therefore, dynamics on the attractor generated by the model may change considerably even for small parameter perturbations.


One of the commonly used measures for estimating the influence of model parameter variations on the state variables is the sensitivity coefficient, which is the derivative of a certain component of a model state vector with respect to some model parameter. Let

\[ S_j = \frac{\partial x_i}{\partial \alpha_j} |_{\alpha_j=\alpha_j^0} = \lim_{\delta \alpha_j \to 0} \left[ \frac{x_i(\alpha_j^0 + \delta \alpha_j) - x_i(\alpha_j^0)}{\delta \alpha_j} \right] \]

where \( \delta \alpha_j \) is the infinitesimal perturbation of parameter \( \alpha_j \) around some fixed point \( \alpha_j^0 \). Differentiating (8) with respect to \( \alpha \), we obtain the set of nonhomogeneous ODEs, the so-called sensitivity equations, which can be written as

\[ \frac{dS_j}{dt} = M \cdot S_j + D_j \quad j = 1, \ldots, m \]

where \( S_j = (\partial x_i/\partial \alpha_j) = (S_{1,j}, S_{2,j}, \ldots, S_{n,j})^T \) is the sensitivity vector with respect to parameter \( \alpha_j \), \( D_j = (\partial f_1/\partial \alpha_j, \partial f_2/\partial \alpha_j, \ldots, \partial f_n/\partial \alpha_j)^T \), and \( M \) is a Jacobian matrix. Thus, to analyze the sensitivity of system (7) with respect to parameter \( \alpha_j \) one can solve the following set of differential equations with given initial conditions:

\[ \frac{dx}{dt} = f(x, \alpha), \quad x(0) = x_0 \]

\[ \frac{dS_j}{dt} = M \cdot S_j + D_j, \quad S_j(0) = S_{j0} \]

Sensitivity equations describe the evolution of sensitivity coefficients along a given trajectory and therefore allow tracing the sensitivity dynamics in time. The procedure for computing sensitivity coefficients includes the following steps:

1. Obtain initial conditions on the system attractor at time \( t = t_0 \) by integrating the nonlinear model equations (7) for a long enough time range \([t_0, \tau]\), starting from random initial conditions.
2. Solve the nonlinear model equations (7) to calculate a trajectory \( x(t), t \in [t_0, \tau] \).
3. Calculate a model Jacobian matrix and a parametric Jacobian matrix.
4. Solve the sensitivity equations (18) with given initial conditions to obtain the desired sensitivity coefficients.

Sensitivity analysis allows us also to explore the sensitivity of a generic objective function (performance measure), which characterizes the dynamical system (7):

\[ J(x, \alpha) = \int_0^\tau \Phi(t; x, \alpha) dt \]

where \( \Phi \) is a nonlinear function of the state variables \( x \) and model parameters \( \alpha \). The gradient of the functional \( J \) with respect to the parameters \( \alpha \) around the unperturbed state vector \( x^0 \)

\[ \nabla_{\alpha} J(x^0, \alpha^0) = \left( \frac{dJ}{dx_{\alpha_1}}, \ldots, \frac{dJ}{dx_{\alpha_m}} \right)^T \]

quantifies the influence of parameters on the model output results. In particular, the effect of the \( j \)th parameter can be estimated as follows:

\[ \frac{dJ}{dx_{\alpha_j}} = \frac{\partial J(x^0, \alpha^0)}{\delta \alpha_j} = \frac{J(x^0 + \delta x^0, \alpha^0 + \delta \alpha_1, \ldots, \alpha^0 + \delta \alpha_j, \ldots, \alpha^0 + \delta \alpha_m) - J(x^0, \alpha^0)}{\delta \alpha_j} \]

where \( \delta \alpha_j \) is the variation in parameter \( \alpha^0_j \). Note that

\[ \frac{dJ}{dx_{\alpha_j}} = \sum_{i=1}^n \frac{\partial J}{\partial x_i} \frac{\partial x_i}{\partial \alpha_j} + \frac{\partial J}{\partial \alpha_j} = \sum_{i=1}^n S_{ij} \frac{\partial J}{\partial x_i} + \frac{\partial J}{\partial \alpha_j}. \]
This approach is acceptable for low-order models. However, the accuracy of sensitivity estimates strongly depends on choice of the perturbation \( \delta \alpha \). By introducing the Gâteaux differential, the sensitivity analysis problem can be considered in the differential formulation eliminating the need to set the value of \( \delta \alpha \). The Gâteaux differential for the objective function (20) has the following form:

\[
\delta J \left( x^0, \alpha^0; \delta x, \delta \alpha \right) = \int_0^T \left( \frac{\partial \Phi}{\partial x} x^{\alpha^0} \cdot \delta x + \frac{\partial \Phi}{\partial \alpha} x^{x, \alpha^0} \cdot \delta \alpha \right) dt.
\] (24)

Here \( \delta x \) is the state vector perturbation due to the variation in the parameter vector in the direction \( \delta \alpha \). Linearizing the nonlinear model (7) around an unperturbed trajectory \( x^0(t) \), we obtain the following system of differential equations, the so-called tangent linear model, for calculating \( \delta x \):

\[
\frac{\partial \delta x}{\partial t} = \frac{\partial f}{\partial x} x^{\alpha^0} \cdot \delta x + \frac{\partial f}{\partial \alpha} x^{x, \alpha^0} \cdot \delta \alpha,
\] (25)

\[ t \in [0, \tau], \quad \delta x (0) = \delta x_0. \]

Then using (24) we can calculate the variation \( \delta J \). Since \( \delta J \left( x^0, \alpha^0; \delta x, \delta \alpha \right) = \langle \nabla_x J, \delta x \rangle \), where \( \langle \cdot, \cdot \rangle \) is a scalar product, the model sensitivity with respect to parameter variations can be estimated by calculating the components of the gradient \( \nabla_x J \). However, this method is computationally ineffective if the number of model parameters \( m \) is large. The use of adjoint equations allows obtaining the required sensitivity estimates within a single computational experiment (e.g., [6, 30, 31]) since the gradient \( \nabla_x J \) can be calculated by the following equation:

\[
\nabla_x J \left( x^0, \alpha^0 \right) = \int_0^T \left[ \frac{\partial \Phi}{\partial x} x^{\alpha^0} - \left( \frac{\partial f}{\partial \alpha} x^{x, \alpha^0} \right)^T \cdot x^* \right] dt,
\] (26)

where the vector function \( x^* \) is the solution of adjoint model

\[
-\frac{\partial x^*}{\partial t} = \left( \frac{\partial f}{\partial x} x^{x, \alpha^0} \right)^T x^* = -\frac{\partial \Phi}{\partial x} x^{x, \alpha^0},
\] (27)

\[ t \in [0, \tau], \quad x^* (\tau) = 0. \]

This equation is numerically integrated in the inverse time direction. Thus, the algorithm for computing sensitivity functions is as follows.

1. Obtain initial conditions on the system attractor at time \( t_0 \) by integrating the nonlinear model equations (7) for a long enough time range \([t_0, \tau]\), starting from random initial conditions.
2. Solve the nonlinear model equations (7) to calculate a trajectory \( x(t), t \in [t_0, \tau] \).
3. Calculate the right-hand side of (27) and then integrate numerically this equation in the inverse time direction with the initial conditions \( x^* (\tau) = 0 \).
4. Calculate the gradient (26).

### 5. Testing the Variational Data Assimilation

We will consider the dynamics of system (2) on its attractor. In order to obtain the model attractor, the numerical integration of (2) is started at \( t_2 = -20 \) with the initial conditions

\[
x (t_0) = (0.01, 0.01, 0.01, 0.02, 0.02, 0.02)^T
\] (28)

and finished at \( t_0 = 0 \) to guarantee that the calculated model state vector \( x_0 = x(0) \) is on the model attractor. The forecast, obtained by NWP models, is substantially determined by initial conditions, which calculated via 4D-Var. The accuracy of initial conditions strongly depends on numerous parameters of 4D-Var systems. Some of these parameters are uncertain. To estimate the influence of parameter variations on the forecast obtained by model (2), the "true" and "forecast" trajectories were calculated. The "true" trajectory \( x'(t_f) \) and the "true" state \( x'(t_f) \) at the verification time \( t_f \) are obtained by integrating the model equations over the time interval \([t_0, t_f]\) with unperturbed parameter vector \( \alpha^0 \) and initial conditions \( x_0 \). Then, the forecast trajectory \( \tilde{x}'(t_f) \) and the forecast state \( \tilde{x}'(t_f) \) at \( t_f \) are obtained by integration of the "forecast" model (2) with initial conditions \( x_0 \) and a certain perturbed model parameter. Thus, the "forecast" model has the same set of equations as the "true" model; however, some model parameter is slightly changed (i.e., this parameter is known with uncertainty). In order to measure forecast errors the relative error in energy norm is used:

\[
e (x') = \left[ (x' - x)^T (x' - x) \right]^{1/2}.
\] (29)

As an example, variations in parameters \( r \) and \( c \) are considered and forecast is made for two time periods \( t_f \): 2.5 and 5.0 of nondimensional time units. The forecast error (29) is estimated at time \( t_f \). Table 1 shows the results of forecast verifications. It is obvious that the less the forecast error measure \( e(x') \), the higher the forecast skill. Qualitatively, calculated results are consistent with real numerical weather forecasts obtained with complex state-of-the-art NWP models: longer \( t_f \) leads to lower forecast accuracy and smaller parameter variations (difference between the "true" parameter value and the value used in the "real" model) lead to better forecast accuracy. It can also be observed that parameter \( r \) influences the forecast accuracy almost twice as much as parameter \( c \).

Synthetic data assimilation requires the "true" state \( x'(t_f) \), the background (first guess) state \( x^b(t_f) \), and observations \( y^o(t_f) \) inside the assimilation window \([t_0, t_N]\) as well as error covariance matrices of the prior guess \( B_0 \) and observations \( R \). The length of data assimilation window should be defined as well. The "true" and background trajectories on the data assimilation interval \([t_0, t_N]\) represent some portions of \( x'(t_f) \) and \( x^b(t_f) \), respectively. Observations should be provided every 5–10 time steps inside the assimilation window and can be generated by adding the Gaussian random noise (with zero mean and variance \( \sigma^2 = 0.2 \)) to the "true" state. Since observation grid and model grid are the same, observation
operator $\mathcal{H}$ is simply an identity mapping. To take into consideration the background covariances, for simplicity, the assumption $B_0 = \sigma_b^2 I$ can be used, where $\sigma_b^2 = 0.2$ is the variance of background error and $I$ is the identity matrix. Under assumption that the observation quality is the same for all variables, the observation covariance matrix can be defined as $R = R = \sigma_a^2 I$.

Testing the TL model and its adjoint is required to ensure the convergence of the minimization algorithm in data assimilation procedures. If $\zeta \delta x$ is a small perturbation of the model state, then

$$\mathcal{H} (x + \zeta \delta x) - \mathcal{H} (x) \approx M(x) \zeta \delta x. \quad (30)$$

To verify the applicability of the TL model on the time interval $[t_0, t_N]$, the relative error

$$e_R = \frac{\mathcal{H} (x + \zeta \delta x) - \mathcal{H} (x)}{M(x) \zeta \delta x} \quad (31)$$

should be calculated. The TL model is valid if $e_R \to 0$ when $\zeta \to 0$. The results of numerical experiments showed that the TL model passed this test with $e_R$ tending towards zero (Table 2). The TL adjoint correctness can be tested by verification of the inner product identity

$$\langle M \delta x, M \delta x \rangle = \langle \delta x, M^T M \delta x \rangle. \quad (32)$$

It was found that this equality is essentially correct: the difference was observed only in the 7th digit, which is consistent with a round-off error. The second test to verify the adjoint model is the so-called gradient test [45], which aims to compare a finite difference representation of the gradient of 4D-Var cost function (13) with the gradient obtained via adjoint model $\nabla J(x_0)$. A linear Taylor approximation of the cost function can be written as

$$J (x_0 + \zeta \delta x) \approx J (x_0) + \zeta (\delta x)^T \nabla J (x_0). \quad (33)$$

Let us introduce the following function:

$$\Psi (\zeta) = \frac{J (x_0 + \zeta \delta x) - J (x_0)}{\zeta (\delta x)^T \nabla J (x_0)}. \quad (34)$$

If the gradient is estimated correctly then the function $\Psi (\zeta) \to 1$ as $\zeta \to 0$. The perturbation vector $\delta x$ is taken to be [45]

$$\delta x = \frac{\nabla J (x_0)}{\|\nabla J (x_0)\|}, \quad (35)$$

where $\| \cdot \|$ is the $L_2$ norm. Table 3 manifests the success of the gradient test.

### Table 1: Relative errors in energy norm (subscript at parameters denotes the nondimensional time unit).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{2.5}$</td>
<td>1%</td>
<td>0.1273</td>
</tr>
<tr>
<td></td>
<td>-1%</td>
<td>0.1266</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.2871</td>
</tr>
<tr>
<td></td>
<td>-5%</td>
<td>0.2796</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>0.4081</td>
</tr>
<tr>
<td></td>
<td>-10%</td>
<td>0.3890</td>
</tr>
<tr>
<td>$c_{1.5}$</td>
<td>1%</td>
<td>0.0584</td>
</tr>
<tr>
<td></td>
<td>-1%</td>
<td>0.0583</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.1312</td>
</tr>
<tr>
<td></td>
<td>-5%</td>
<td>0.1300</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>0.1866</td>
</tr>
<tr>
<td></td>
<td>-10%</td>
<td>0.1833</td>
</tr>
<tr>
<td>$r_{5.0}$</td>
<td>1%</td>
<td>0.1663</td>
</tr>
<tr>
<td></td>
<td>-1%</td>
<td>0.1689</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.3617</td>
</tr>
<tr>
<td></td>
<td>-5%</td>
<td>0.3202</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>0.5004</td>
</tr>
<tr>
<td></td>
<td>-10%</td>
<td>0.5742</td>
</tr>
<tr>
<td>$c_{5.0}$</td>
<td>1%</td>
<td>0.0705</td>
</tr>
<tr>
<td></td>
<td>-1%</td>
<td>0.0704</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.1583</td>
</tr>
<tr>
<td></td>
<td>-5%</td>
<td>0.1569</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>0.2250</td>
</tr>
<tr>
<td></td>
<td>-10%</td>
<td>0.2210</td>
</tr>
</tbody>
</table>

### Table 2: Results of verification of tangent linear model for $c = 0.8$ and $\delta x = 10^{-3} x_c$.  

<table>
<thead>
<tr>
<th>$\zeta$</th>
<th>$e_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9182066027544249</td>
</tr>
<tr>
<td>10^{-1}</td>
<td>0.9997279965782743</td>
</tr>
<tr>
<td>10^{-2}</td>
<td>0.9999925468155463</td>
</tr>
<tr>
<td>10^{-3}</td>
<td>0.99999929961531</td>
</tr>
<tr>
<td>10^{-4}</td>
<td>0.99999934965012</td>
</tr>
<tr>
<td>10^{-5}</td>
<td>0.99999991217883</td>
</tr>
<tr>
<td>10^{-6}</td>
<td>0.99999994427087</td>
</tr>
<tr>
<td>10^{-7}</td>
<td>0.9999997435447022</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\zeta$</th>
<th>$\Psi (\zeta)$</th>
<th>$\log_{10} (|\Psi (\zeta) - 1|)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^{-4}</td>
<td>0.872773198461396</td>
<td>0.8954213897009506</td>
</tr>
<tr>
<td>10^{-5}</td>
<td>0.9975483775420343</td>
<td>-2.6105464089686840</td>
</tr>
<tr>
<td>10^{-6}</td>
<td>0.9998765512756632</td>
<td>-3.90851395113787</td>
</tr>
<tr>
<td>10^{-7}</td>
<td>0.9999845628441233</td>
<td>-4.937654381675845</td>
</tr>
<tr>
<td>10^{-8}</td>
<td>0.9999979865432885</td>
<td>-5.696057724708600</td>
</tr>
<tr>
<td>10^{-9}</td>
<td>0.9999998912431426</td>
<td>-6.963543350893210</td>
</tr>
<tr>
<td>10^{-10}</td>
<td>0.9999999244103234</td>
<td>-7.1215375125562080</td>
</tr>
</tbody>
</table>

6. Sensitivity of the System with respect to Parameters

According to the sensitivity theory [44], general solutions of sensitivity equations for oscillatory nonlinear dynamical systems grow unbounded as time tends to infinity; therefore, sensitivity functions calculated by conventional approaches have a high degree of uncertainty. The reason is that nonlinear dynamical systems that exhibit chaotic behavior are very sensitive to its initial conditions. Thus, the solutions to the linearized Cauchy problem (7) grow exponentially as $\|\delta x(t)\| = \|\delta x(0)\| e^{\lambda t}$, where $\lambda > 0$ is the leading Lyapunov exponent. As a result, calculated sensitivity coefficients contain a fairly large error [35–37]. To illustrate this point, let us explore the sensitivity of model output to changes in the coupling...
strength parameter. Let us introduce the following sensitivity coefficients:

\[
S_{1c} = \frac{\partial x}{\partial c}, \quad S_{2c} = \frac{\partial y}{\partial c}, \quad S_{3c} = \frac{\partial z}{\partial c}, \quad S_{4c} = \frac{\partial \dot{x}}{\partial c}, \quad S_{5c} = \frac{\partial \dot{y}}{\partial c}, \quad S_{6c} = \frac{\partial \dot{z}}{\partial c}.
\] (36)

The corresponding sensitivity equations can be written as

\[
S_{1c} = \sigma (S_{2c} - S_{1c}) - c S_{4c} - x,
\]

\[
S_{2c} = r S_{1c} - S_{2c} - x S_{5c} - z S_{1c} + c S_{6c} + Y,
\]

\[
S_{3c} = x S_{2c} + y S_{1c} - b S_{3c} + c S_{6c} + Z,
\]

\[
S_{4c} = \varepsilon (S_{5c} - S_{4c}) - c S_{4c} - x,
\]

\[
S_{5c} = \varepsilon (r S_{4c} - S_{5c} - X S_{6c} - Z S_{5c}) + c S_{3c} + y,
\]

\[
S_{6c} = \varepsilon (X S_{5c} + Y S_{4c} - b S_{6c}) - c S_{2c} - z.
\] (37)

Sensitivity coefficients can be introduced for any particular model parameter. Since the parameter vector \(\alpha\) consists of five components \((\alpha, r, b, c, \varepsilon)\), five sets of sensitivity equations can be derived from the model equations (2). The dynamics of sensitivity coefficients (36) can be traced by solving the sensitivity equations (37) along with the nonlinear model (2).

Sensitivity coefficients (36), calculated on the time interval \([0, 20]\), are shown in Table 4. The most sensitive variables are \(z\) and \(Z\). The sensitivity of variables \(x\),...
y, X, and Y with respect to $r$ is significantly less than variables $z$ and $Z$.

7. Concluding Remarks

We considered a coupled nonlinear dynamical system, which is composed of fast (the "atmosphere") and slow (the "ocean") versions of the well-known Lorenz model. This low-order mathematical tool allows us to mimic the atmosphere-ocean system and therefore serves as a key part of a theoretical and computational framework for the study of various aspects of coupled 4D-Var procedures. Numerical models used to predict the weather are highly nonlinear but tangent linear approximations and their adjoints are used in VDA algorithms. Linear approximation of strongly nonlinear NWP models and also uncertainties in their numerous parameters generate errors in the initial conditions obtained via data assimilation systems. The influence of parameter uncertainties on the results of data assimilation can be studied using sensitivity analysis.

We discussed conventional methods of sensitivity analysis and their inefficiency with respect to calculating sensitivity
Figure 5: Time dynamics of sensitivity functions with respect to parameter $c$ on the time interval $[0, 5]$ for $c^0 = 0.9$.

Table 4: Sensitivity estimates of fast and slow variables with respect to parameter $r$.

<table>
<thead>
<tr>
<th>$c$</th>
<th>$\partial Z / \partial r$</th>
<th>$\partial Y / \partial r$</th>
<th>$\partial X / \partial r$</th>
<th>$\partial z / \partial r$</th>
<th>$\partial y / \partial r$</th>
<th>$\partial x / \partial r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.10</td>
<td>0.05</td>
<td>0.01</td>
<td>1.08</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>0.8</td>
<td>0.69</td>
<td>0.08</td>
<td>0.03</td>
<td>1.02</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>0.4</td>
<td>0.95</td>
<td>0.03</td>
<td>-0.01</td>
<td>1.03</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>0.15</td>
<td>0.91</td>
<td>-0.08</td>
<td>-0.09</td>
<td>1.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>1.04</td>
<td>-0.02</td>
<td>-0.03</td>
<td>1.02</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
</tbody>
</table>
coefficients for chaotic dynamics. To calculate sensitivity coefficients with acceptable accuracy, the sensitivity analysis method [36, 37], developed on the basis of theory of shadowing of pseudoorbits in dynamical systems [38, 39], was applied. Previously, this method was used to analyze the sensitivity of the periodic van der Pol oscillator, the original Lorenz system, and simplified aeroelastic model that exhibit both periodic and chaotic regimes.

Calculated sensitivity coefficients obtained via conventional methods and the "shadowing" approach are presented and discussed. It was shown that envelopes of sensitivity coefficients obtained by conventional methods grow over time and the coefficients themselves exhibit the oscillating behaviour. Using the "shadowing" method allows us to calculate the average sensitivity functions (coefficients) that can be easily interpreted.

In conclusion, two comments should be highlighted.

(1) The shadowing property of dynamical systems is a fundamental feature of hyperbolic systems that was first discovered by Anosov [46] and Bowen [47]. However, most physical systems are nonhyperbolic. Despite the fact that much of shadowing theory has been developed for hyperbolic systems, there is evidence that nonhyperbolic attractors also have the shadowing property (e.g., [48–51]). In theory this property should be verified for each particular dynamical system, but this is more easily said than done.

(2) The applicability of the shadowing method for sensitivity analysis of modern atmospheric and climate models is a rather complicated problem since these models are quite complex and they contain numerous input parameters. Thus, further research and computational experiments are required. However, we are confident that, by using the basic ideas of the shadowing method, it is possible to better understand the sensitivity analysis of atmospheric models of various levels of complexity.

**Appendix**

The novel sensitivity analysis method for chaotic dynamical systems developed in [36, 37] is based on the theory of pseudoorbit shadowing in dynamical systems [38, 39], which is one of the most rapidly developing components of the global theory of dynamical systems and classical theory of structural stability [52]. Naturally, pseudo- (or approximate-) trajectories arise due to the presence of round-off errors, method errors, and other errors in computer simulation of dynamical systems. Consequently, we will not get an exact trajectory of a system, but we can come very close to an exact solution and the resulting approximate solution will be a pseudotrajectory. The shadowing property (or pseudoorbit tracing property) means that, near an approximate trajectory, there exists the exact trajectory of the system considered, such that it lies uniformly close to a pseudotrajectory. The shadowing theory is well developed for the hyperbolic dynamics, which is characterized by the presence of expanding and contracting directions for derivatives. The study of shadowing problem was originated by Anosov [46] and Bowen [47].

Let $(M, \text{dist})$ be a compact metric space and let $f : M \rightarrow M$ be a homeomorphism (a discrete dynamical system on $M$). A set of points $X = \{ x_k : k \in \mathbb{Z} \}$ is a $d$-pseudotrajectory ($d > 0$) of $f$ if

$$\text{dist}(x_{k+1}, f(x_k)) < d, \quad k \in \mathbb{Z}. \quad (A.1)$$

Here the notation $\text{dist}(\cdot, \cdot)$ denotes the distance in the phase space between two geometric objects within the brackets.

We say that $f$ has the shadowing property if given $\varepsilon > 0$ there is $d > 0$ such that for any $d$-pseudotrajectory $X = \{ x_k : k \in \mathbb{Z} \}$ there exists a corresponding trajectory $Y = \{ y_k : k \in \mathbb{Z} \}$, which $\varepsilon$-traces $X$; that is

$$\text{dist}(x_k, y_k) < \varepsilon, \quad k \in \mathbb{Z}. \quad (A.2)$$

The shadowing lemma for discrete dynamical systems [53] states that, for each $\varepsilon > 0$, there exists $d > 0$ such that each $d$-pseudotrajectory can be $\varepsilon$-shadowed.

The definition of pseudotrajectory and shadowing lemma for flows (continuous dynamical systems) [38] are more complicated than for discrete dynamical systems. Let $\Phi^t : \mathbb{R} \times M \rightarrow M$ be a flow of a vector field $X$ on $M$. A function $g : \mathbb{R} \rightarrow M$ is a $d$-pseudotrajectory of the dynamical system $\Phi^t$ if the inequalities

$$\text{dist}(\Phi^t (t, g(t)), g(t + \tau)) < d \quad (A.3)$$

hold for any $t \in [-1, 1]$ and $\tau \in \mathbb{R}$. The "continuous" shadowing lemma ensures that, for the vector field $X$ generating the flow $\Phi^t$, the shadowing property holds in a small neighborhood of a compact hyperbolic set for dynamical system $\Phi^t$.

It is very important to note that the shadowing problem for continuous dynamical systems requires reparameterization of shadowing trajectories. This is the case because for
Figure 7: Time dynamics of sensitivity functions with respect to parameter $r$ on the interval $[0, 25]$ for $\varepsilon^0 = 0.9$.

Continuous dynamical systems close points of pseudotrajectory and true trajectory do not correspond to the same moments of time. A monotonically increasing homeomorphism $h : \mathbb{R} \rightarrow \mathbb{R}$ such that $h(0) = 0$ is called a reparameterization and denoted by Rep. For $\varepsilon > 0$, Rep($\varepsilon$) is defined as follows [38]:

$$\text{Rep}(\varepsilon) = \left\{ h \in \text{Rep} : \left| \frac{h(t_1) - h(t_2)}{t_1 - t_2} - 1 \right| \leq \varepsilon \right\} \quad \text{(A.4)}$$

For any different $t_1, t_2 \in \mathbb{R}$.

For simplicity, we will consider a generic autonomous dynamical system with one parameter $\alpha$:

$$\frac{dx}{dt} = f(x, \alpha), \quad x \in \mathbb{R}^n. \quad \text{(A.5)}$$

The new sensitivity analysis method [36, 37] is based on the “continuous” shadowing lemma and the following two basic assumptions.

(a) Model state variables are considered over long time interval $t \in [0, T]$, where $T \to \infty$, and an averaged
Figure 8: Time dynamics of sensitivity functions with respect to parameter $r$ on the time interval $[0, 5]$ for $c^0 = 0.9$.

The performance measure $\langle J(\alpha) \rangle$ is of the most interest for us:

$$
\langle J(\alpha) \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T J(x(t, \alpha), \alpha) \, dt. \tag{A.6}
$$

(b) The dynamical system under consideration is ergodic. With these assumptions, we can use the arbitrarily chosen trajectory of the system to trace the state variables along the orbit and calculate the performance measure $J(\alpha)$. For example, the arbitrary trajectory $x(t)$ can be chosen as a solution of the model equation, such that it is located nearby a certain reference trajectory $x_r(t)$. Taking into account the shadowing lemma, the closest orbit $x(t)$ to $x_r(t)$ satisfies the following constrained minimization problem [37]:

$$
\min_{x, \tau} \frac{1}{T} \int_0^T \left[ \| x(\tau(t)) - x_r(t) \|^2 + \eta^2 \left( \frac{d\tau}{dt} - 1 \right)^2 \right] \, dt,
$$

such that $\frac{dx}{dt} = f(x, \alpha), \tag{A.7}$
where $\eta$ is the parameter that provides the same order of magnitude of the two terms in the integrand and $\tau(t)$ is a time transformation. The second term in the integrand describes reparameterization. Problem (A.7) is called the nonlinear Least Square Shadowing (LSS) problem, and its solution, denoted by $x_{s}^{(T)}(t, \alpha)$ and $\tau_{s}^{(T)}(t, \alpha)$, is a solution of the model equation and time transformation that provides the trajectory $x_{s}^{(T)}(t, \alpha)$ to be close to $x_{r}(t)$. The performance measure (A.6) averaged over the time interval $t \in [0, T]$ can be then approximated as

$$
\langle J(\alpha) \rangle = \langle J_{s}^{(T)}(\alpha) \rangle = \frac{1}{\tau(T) - \tau(0)} \int_{\tau(0)}^{\tau(T)} J\left(x_{s}^{(T)}(t, \alpha), \alpha\right) dt
$$

(A.8)
since $x_s^{(T)}(t, \alpha)$ satisfies the model equation at a different $\alpha$.
The desired sensitivity estimate $\nabla_\alpha (f_s^{(T)}(\alpha))$ can be computed by solving the following linearized LLS problem [37]:

$$\min \frac{1}{T} \int_0^T [\|S\|^2 + \eta^2 \mu^2] \, dt,$$

(A.9)

such that

$$\frac{dS}{dt} = \frac{\partial f}{\partial x} S + \frac{\partial f}{\partial \alpha} \mu (x, \alpha).$$

The solutions of this problem $S(t)$ and $\mu(t)$ relate to the solutions of the nonlinear LSS problem (A.7) as follows:

$$S(t) = \frac{d}{d\alpha} \left( x_s^{(T)} \left( \nu_s^{(T)}(t, \alpha), \alpha \right) \right),$$

(A.10)

$$\mu(t) = \frac{d}{d\alpha} \frac{dS}{dt_{\nu_s^{(T)}}(t, \alpha)}.$$

The time-dependent parameter $\mu$ is called a time-dilation variable and corresponds to the time transformation from the shadowing lemma. Using $S$ and $\mu$, we can estimate the desired sensitivity (the derivative of the linearized performance measure (A.8) with respect to the parameter $\alpha$):

$$\nabla_\alpha \langle f_s^{(T)}(\alpha) \rangle \approx \frac{1}{T} \int_0^T \left( \frac{\partial f}{\partial x} S + \frac{\partial f}{\partial \alpha} \mu (J - \bar{J}) \right) \, dt,$$

(A.11)

where $\bar{J} = \frac{1}{T} \int_0^T J \, dt$.

Several numerical algorithms can be used to solve the linearized LSS problem (A.9). One such method is based on variational calculus, which is used to derive optimality conditions representing a system of linear differential equations that are then discretized and solved numerically to calculate variables $S$ and $\mu$ [37]. Let $\Delta t = T/m$ be a uniform discretization time step, then denoting $x_{i+1/2} = x_i ((i + 1/2)\Delta t)$, $x_{i-1/2} = x_i ((i - 1/2)\Delta t)$, $i = 0, \ldots, m - 1$, and using the trapezoidal rule to approximate the time derivative of $S$ and $\mu$, we can obtain the following discrete Karush-Kuhn-Tucker (KKT) system [37, 54]:

$$\begin{bmatrix}
I & \eta^2 \\
\eta^2 & I \\
& \ddots & \ddots \\
& & \ddots & \eta^2 \\
& & & I \\
A_1 & B_1 & C_1 \\
A_2 & B_2 & C_2 \\
& \ddots & \ddots \\
& & A_{m-1} & B_{m-1} & C_{m-1}
\end{bmatrix} \begin{bmatrix}
S_{1/2} \\
\mu_1 \\
S_{1+1/2} \\
\mu_2 \\
S_{2+1/2} \\
\mu_3 \\
\vdots \\
\vdots \\
\mu_{m-1} \\
S_{m-1/2} \\
-a_1 \\
-a_2 \\
\vdots \\
-a_{m-1}
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
\vdots \\
\vdots \\
0 \\
0 \\
-d_1 \\
-d_2 \\
\vdots \\
-d_{m-1}
\end{bmatrix},$$

(A.12)

where

$$A_i = -\frac{1}{\Delta t} - \frac{\partial f}{\partial x} (x_{i-1/2}, \alpha), \quad B_i = \frac{x_{i+1/2} - x_{i-1/2}}{\Delta t},$$

$$C_i = \frac{1}{\Delta t} - \frac{\partial f}{\partial x} (x_{i+1/2}, \alpha),$$

$$d_i = \frac{1}{2} \left[ \frac{\partial f}{\partial x} (x_{i-1/2}, \alpha) + \frac{\partial f}{\partial x} (x_{i+1/2}, \alpha) \right].$$

(A.13)

The KKT system can be solved using, for example, iterative methods. Gauss elimination approach for solving the KKT system was considered in [37]. The convergence of the LSS method was proved in [55]. Calculated variables $S$ and $\mu$ are then used to determine the sensitivity estimate (A.11). The LSS algorithm is summarized as follows.

1. Define a temporal grid $t_i = i \Delta t, i = 0, \ldots, m$, and then discretize the model equation (A.5) on this grid.
2. Calculate a solution of (A.5) $x_\alpha$ on the time interval $[0, T]$.
3. Compute the vectors $A_i, B_i, C_i$, and $d_i$.
4. Solve the KKT system to obtain variables $S_i$ and $\mu_i$.
5. Compute the gradient components (sensitivity functions) (A.11).

Another method that also uses the concept of pseudoorbit shadowing in dynamical systems for sensitivity analysis of chaotic oscillations was developed in [36]. This method
is based on inverting the so-called “shadowing operator” that requires calculating the Lyapunov characteristic (covariant) vectors. However, the computational cost needed for the Lyapunov eigenvector decomposition is high when the dynamical system has many positive Lyapunov exponents. To illustrate this approach, suppose that the sensitivity analysis of system (A.5) aims to estimate the following sensitivity coefficient:

\[ S_\alpha = \frac{\partial x}{\partial \alpha} . \]

Let us introduce the following transform: \( x' (x) = x + \delta x (x) \), where \( x \) and \( x' \) are true trajectory and pseudoorbit, respectively. The orbit \( x' \) is generated due to the variation in parameter \( \alpha \). It can be shown [36] that \( \delta f (x) = A \delta x (x) \), where

\[
A = \left[ -\frac{\partial f}{\partial x} + \left( \frac{d}{dt} \right) \right] \quad \text{(A.14)}
\]

is a “shadow” operator. Thus, to find a pseudotrajectory, we need to solve the equation \( \delta x = A^{-1} \delta f \); that is, we must numerically invert the operator \( A \) for a given \( \delta f \). To solve this problem, functions \( \delta x \) and \( \delta f \) are decomposed into their constituent Lyapunov covariant vectors \( v_1 (x), \ldots, v_n (x) \):

\[
\delta x (x) = \sum_{i=1}^{n} \psi_i (x) v_i (x) , \quad \text{(A.15a)}
\]

\[
\delta f (x) = \sum_{i=1}^{n} \phi_i (x) v_i (x) . \quad \text{(A.15b)}
\]

Note that each \( v_i (x) \) satisfies the following equation:

\[
\frac{dv_i (x(t))}{dt} = \frac{\partial f}{\partial x} v_i (x(t)) - \lambda_i v_i (x(t)) , \quad \text{(A.16)}
\]

where \( \lambda_1, \ldots, \lambda_n \) are the Lyapunov exponents. From (A.14) one can obtain

\[
A \left( \psi_i v_i \right) = \left[ -\psi_i (x) \frac{\partial f}{\partial x} + \left( \frac{d}{dt} \right) \right] v_i (x) + \psi_i (x) \frac{dv_i (x)}{dt} . \quad \text{(A.17)}
\]

Substitution of (A.16) into the last term of (A.17) gives

\[
A \left( \psi_i v_i \right) = \left[ \frac{dv_i (x)}{dt} - \lambda_i \psi_i (x) \right] v_i (x) . \quad \text{(A.18)}
\]

From (A.15a) and (A.15b) and (A.18) and the relation \( \delta f (x) = A \delta x(x) \), we get

\[
\delta f (x) = \sum_{i=1}^{n} A \left( \psi_i v_i \right) = \sum_{i=1}^{n} \left( \frac{dv_i (x)}{dt} - \lambda_i \psi_i (x) \right) v_i . \quad \text{(A.19)}
\]

Equation (A.19) gives the following relationship between \( \psi_i (x) \) and \( \phi_i (x) \) along the orbit:

\[
\frac{dv_i (x)}{dt} = \phi_i (x) + \lambda_i \psi_i (x) . \quad \text{(A.20)}
\]

Thus, we can calculate \( \psi_i (x) \) using (A.20) by first decomposing \( \delta f \) as a sum (A.15b), and then the desired \( \delta x \) can be obtained from (A.15a). However, if dynamical system has a zero Lyapunov exponent, \( \lambda_0^* = 0 \), then the algorithm described above fails to compute \( \delta x \) [36]. The problem can
be resolved by introducing a time-dilation variable $\mu$ that satisfies the following equation:

$$\mu + \langle \phi_i^0 \rangle = 0,$$

(A.21)

where

$$\langle \phi_i^0 \rangle = \lim_{T \to \infty} \frac{1}{T} \left[ \psi_i^0 (x(T)) - \psi_i^0 (x(0)) \right].$$

(A.22)

In the presence of the variable $\mu$, the expression for calculating $\delta x$ takes the following form: $\delta x = A^{-1}(\delta f + \mu f)$.

The supplement $\mu f$ affects (A.20) only for $\lambda_i^0 = 0$:

$$\frac{d\psi_i^0 (x)}{dt} = \psi_i^0 (x) + \mu.$$

(A.23)
In general, the procedure for solving a sensitivity analysis problem represents the following set of steps.

1. Obtain initial conditions of the system attractor by integrating the model equations (A.5) from $t_D = -20$ to $t_0 = 0$, starting from random initial conditions.

2. Solve (A.5) to obtain a trajectory $x(t)$, $t \in [0, 20]$, on the attractor.

3. Compute the Lyapunov exponents $\lambda_i$ and the Lyapunov covariant vectors $v_i(x(t))$, $i = 1, \ldots, n$.

4. Define $\delta f = (\partial f / \partial \alpha) \delta \alpha$ and execute the Lyapunov spectrum decomposition of $\delta f$ along the trajectory $x(t)$ to obtain $\varphi_i(x)$, $i = 1, \ldots, n$.

5. Calculate the time-dilation variable $\mu$ using (A.21).

6. Compute $\delta x$ along the trajectory $x(t)$.

7. Estimate the sensitivity $S_\alpha = \delta x / \delta \alpha$ by averaging over the time interval $t \in [0, 20]$.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References


