

Research Article

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

Sergei Soldatenko,¹ Peter Steinle,¹ Chris Tingwell,¹ and Denis Chichkine²

¹Centre for Australian Climate and Weather Research, 700 Collins Street, Melbourne, VIC 3008, Australia

²University 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

Academic Editor: Guijun Han

Copyright © 2015 Sergei Soldatenko et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Variational data assimilation (VDA) remains one of the key issues arising in many fields of geosciences including the numerical weather prediction. While the theory of VDA is well established, there are a number of issues with practical implementation that require additional consideration and study. However, the exploration of VDA requires considerable computational resources. For simple enough low-order models, the computational cost is minor and therefore models of this class are used as simple test instruments to emulate more complex systems. In this paper, the sensitivity with respect to variations in the parameters of one of the main components of VDA, the nonlinear forecasting model, is considered. For chaotic atmospheric dynamics, conventional methods of sensitivity analysis provide uninformative results since the envelopes of sensitivity functions grow with time and sensitivity functions themselves demonstrate the oscillating behaviour. The use of sensitivity analysis method, developed on the basis of the theory of shadowing pseudoorbits in dynamical systems, allows us to calculate sensitivity functions correctly. Sensitivity estimates for a simple coupled dynamical system are calculated and presented in the paper. To estimate the influence of model parameter uncertainties on the forecast, the relative error in the energy norm is applied.

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 pseudo-orbit 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 ε (e.g., [33, 34]):

$$\begin{aligned}\dot{x} &= \sigma(y - x) - c(aX + k), \\ \dot{y} &= rx - y - xz + c(aY + k),\end{aligned}\quad (1a)$$

$$\dot{z} = xy - bz + c_z Z,$$

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

$$\dot{Y} = \varepsilon(rX - Y - aXZ) + c(y + k),\quad (1b)$$

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

where lower case letters represent the fast subsystem and capital letters the slow subsystem, σ , 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 $\mathbf{x} = (x, y, z, X, Y, Z)^T$ and the model parameter vector is

$\alpha = (\sigma, r, b, a, c, c_z, k, \varepsilon)^T$. Without loss of generality, we can assume that $a = 1$, $k = 0$, and $c = c_z$, then $\alpha = (\sigma, r, b, c, \varepsilon)^T$ and the system (1a) and (1b) can be rewritten in the operator form:

$$\frac{d\mathbf{x}}{dt} = (\mathcal{L} + \mathcal{Q})\mathbf{x}, \quad (2)$$

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

$$\mathcal{L} = \begin{bmatrix} -\sigma & \sigma & 0 & & & \\ r & -1 & -x & & & 0 \\ 0 & x & -b & & & \\ & & & -\varepsilon\sigma & \varepsilon\sigma & 0 \\ 0 & & & \varepsilon r & -\varepsilon & -\varepsilon X \\ & & & 0 & \varepsilon X & -\varepsilon b \end{bmatrix}, \quad (3)$$

$$\mathcal{Q} = \begin{bmatrix} & & -c & 0 & 0 \\ & 0 & 0 & c & 0 \\ & & 0 & 0 & c \\ -c & 0 & 0 & & \\ 0 & c & 0 & & 0 \\ 0 & 0 & -c & & \end{bmatrix}.$$

The unperturbed parameter values are taken as

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

Chosen values of σ , 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 $c \approx 0.8$ they approach the x -axis and at about

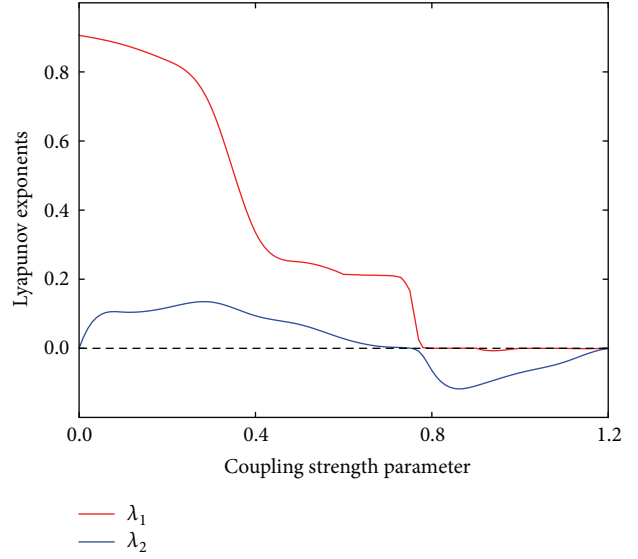


FIGURE 1: Two largest conditional Lyapunov exponents as functions of coupling strength parameter.

$c \approx 0.95$ 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

$$C(s) = \langle x_i x_{i+s} \rangle - \langle x_i \rangle \langle x_{i+s} \rangle, \quad (5)$$

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:

$$C(s) = \langle u_m u_{m+s} \rangle - \langle u \rangle^2. \quad (6)$$

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:

$$\frac{d\mathbf{x}(t)}{dt} = f(\mathbf{x}(t), \boldsymbol{\alpha}(t)), \quad t \in [0, \tau] = \mathcal{T}, \quad \mathbf{x}(0) = \mathbf{x}_0. \quad (7)$$

Here $\mathbf{x}(t)$ is the state vector belonging to a Hilbert space \mathcal{X} , $\boldsymbol{\alpha} \in \mathcal{P}$ 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 $f : \mathcal{X} \times \mathcal{P} \times \mathcal{T} \rightarrow \mathcal{X}$, and \mathbf{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

$$\mathbf{x}_{i+1} = \mathcal{M}_{0,i+1}(\mathbf{x}_i) + \boldsymbol{\varepsilon}_i^m, \quad (8)$$

where $\mathbf{x}_i \in \mathbb{R}^n$ 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 , $\mathcal{M}_{0,i+1}$ is the nonlinear operator that indirectly contains model parameters and propagates the state vector from time t_i to time t_{i+1} for $i = 0, \dots, N-1$, and $\boldsymbol{\varepsilon}_i^m$ is the model errors. It is usually assumed that model (8) is “perfect” ($\boldsymbol{\varepsilon}_i^m = 0$), that is the forecast has no errors if the initial condition is perfect. In this case, given the model operator and the initial condition \mathbf{x}_0 , (8) uniquely specifies the orbit of the dynamical system. Let $\mathbf{y}_i^0 \in \mathbb{R}^m$ be the m -dimensional vector of observations measured at a discrete time t_i , $i = 0, \dots, N$ that are linked to the system state via the following equation:

$$\mathbf{y}_i^0 = \mathcal{H}_i(\mathbf{x}_i) + \boldsymbol{\varepsilon}_i^0, \quad (9)$$

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 $\boldsymbol{\varepsilon}_i^0$ are unbiased, serially uncorrelated, and normally distributed with known covariance matrices $\mathbf{R}_i \in \mathbb{R}^{m \times m}$.

Suppose that at the initial time the prior (background) model state \mathbf{x}_0^b is known and represents the “best” estimate of the “true” state \mathbf{x}_0^t before any observations are taken. This background state is provided by a previous forecast. It is assumed that \mathbf{x}_0^b has unbiased and normally distributed errors $\boldsymbol{\varepsilon}^b$ with known covariance matrix $\mathbf{B}_0 \in \mathbb{R}^{n \times n}$:

$$\mathbf{x}_0^b = \mathbf{x}_0^t + \boldsymbol{\varepsilon}^b. \quad (10)$$

Given the observations \mathbf{y}_i^0 at time t_i , the corresponding observation error covariance matrices \mathbf{R}_i ($i = 0, \dots, N$), the background initial state \mathbf{x}_0^b , and the error covariance matrix \mathbf{B}_0 , the 4D-Var data assimilation seeks to minimize, with respect to \mathbf{x}_0 , a certain cost function $J(\mathbf{x})$ expressing the “distance” between observations and corresponding model state using the model equations as constraints:

$$\mathbf{x}_0^a = \arg \min J(\mathbf{x}) \quad (11)$$

subject to \mathbf{x} satisfying the set of the “ideal” ($\boldsymbol{\varepsilon}_i^m = 0$) model equations with initial state \mathbf{x}_0 :

$$\mathbf{x}_{i+1} = \mathcal{M}_{0,i+1}(\mathbf{x}_0), \quad i = 1, \dots, N. \quad (12)$$

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

$$J(\mathbf{x}_0) = \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{i=0}^N (\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i^0)^T \mathbf{R}_i^{-1} (\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i^0). \quad (13)$$

The optimization problem (11) 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:

$$\nabla_{\mathbf{x}_0} J(\mathbf{x}_0) = \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \sum_{i=0}^N \mathbf{M}_{0,i}^T \mathbf{H}_i^T \mathbf{R}_i^{-1} (\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i^0), \quad (14)$$

where $\mathbf{M}_{0,i}^T$ is the adjoint of the linearized model operator $\mathbf{M}_{0,i} = \mathcal{M}'_{0,i}(\mathbf{x}_i)$ and \mathbf{H}_i^T is the adjoint of the linearized observation operator $\mathbf{H}_i = \mathcal{H}'_i(\mathbf{x}_i)$. If the model is not “perfect” then we need to take into account the model errors $\boldsymbol{\varepsilon}^m$, which are sometimes taken as Gaussian noise:

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

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

$$J_m = \frac{1}{2} \sum_{i=1}^N (\mathbf{x}_i - \mathcal{M}_{i-1,i}(\mathbf{x}_{i-1}))^T \mathbf{Q}_i^{-1} (\mathbf{x}_i - \mathcal{M}_{i-1,i}(\mathbf{x}_{i-1})) \quad (16)$$

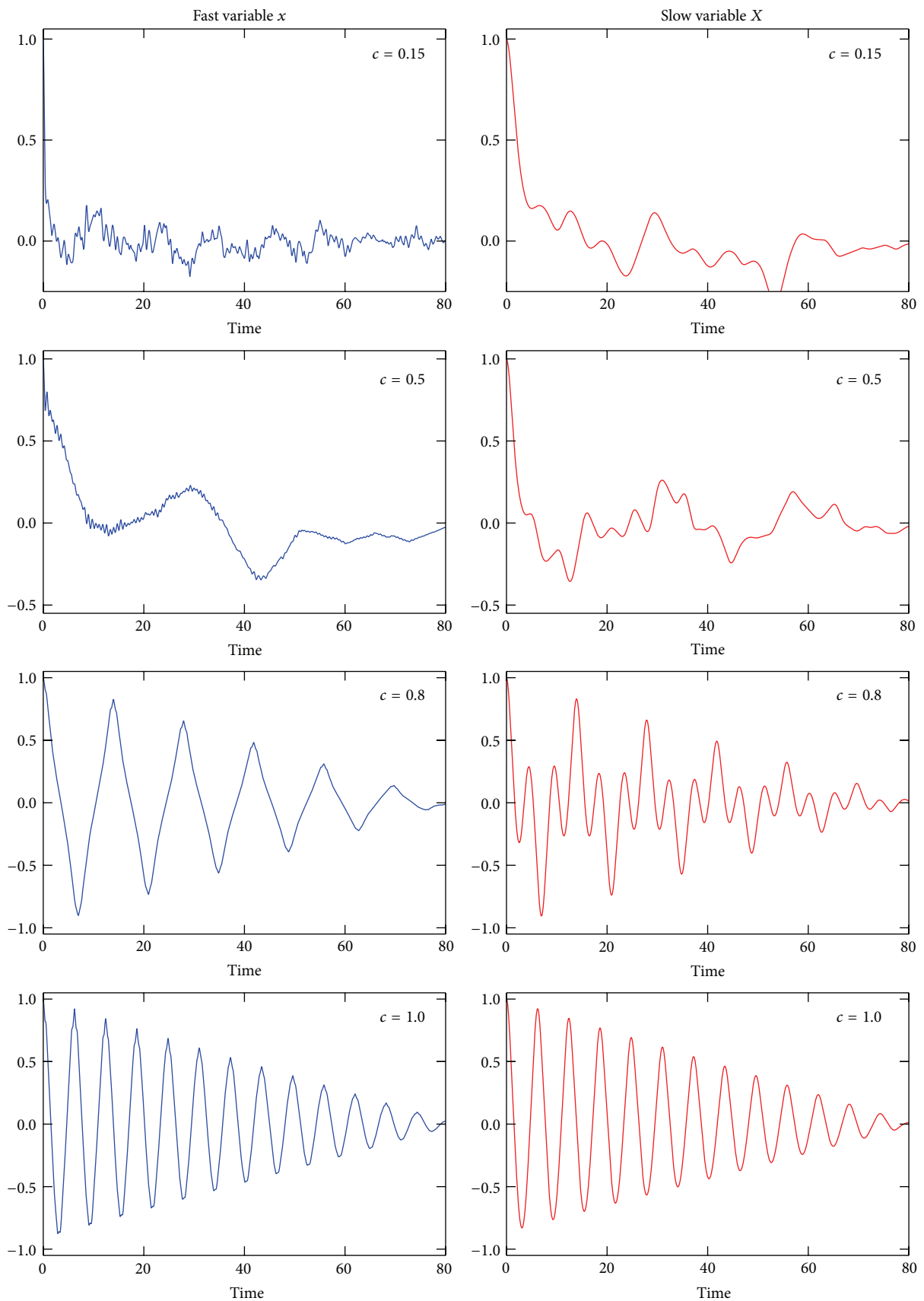


FIGURE 2: Autocorrelation functions for dynamic variables x and X for different parameter c .

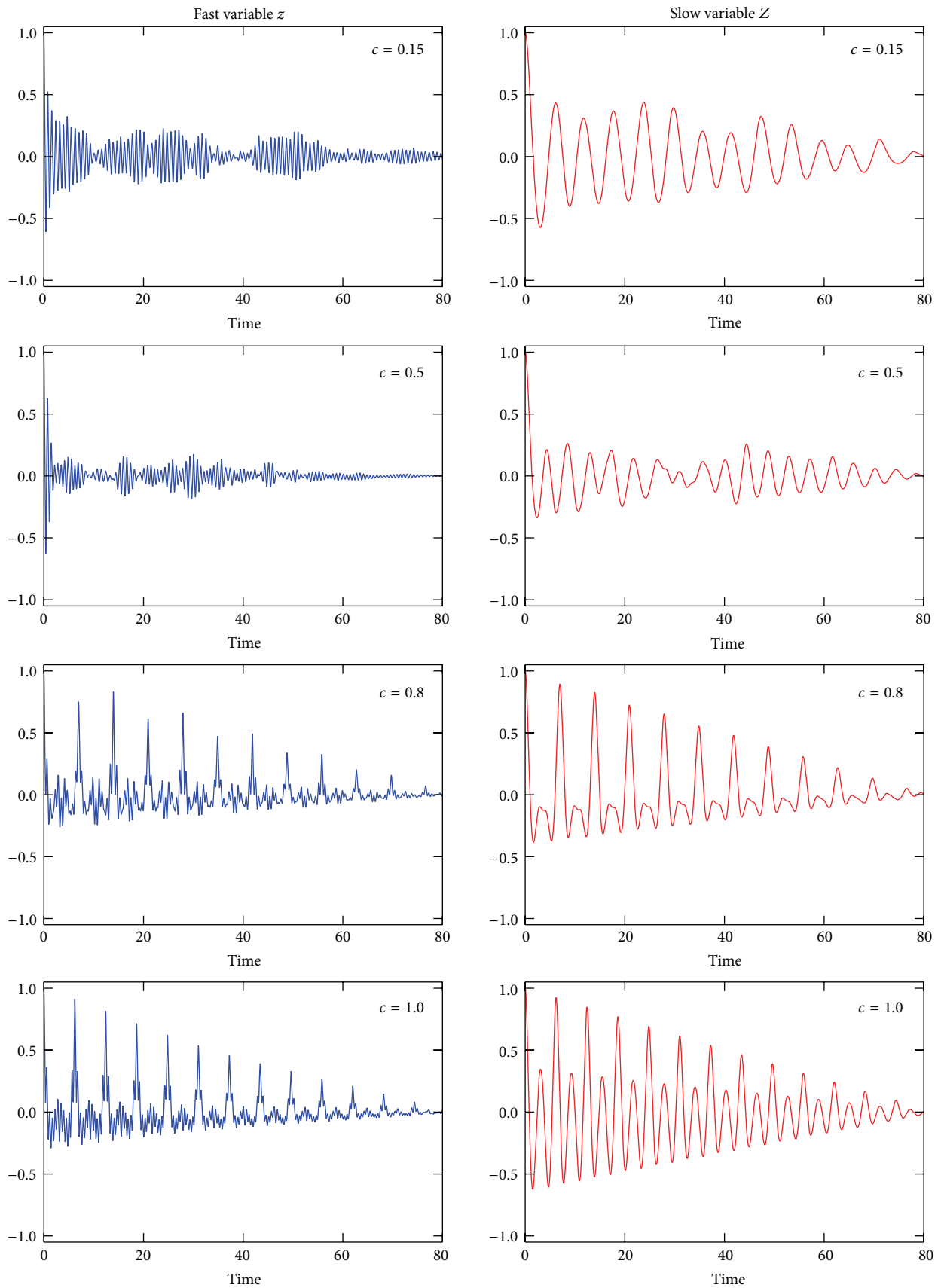


FIGURE 3: Autocorrelation functions for dynamic variables z and Z for different parameter c .

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 climate 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.

4. Sensitivity Analysis: Essentials of Conventional Approaches

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 x_i with respect to some model parameter α_j [30, 31, 44]:

$$S_{ij} \equiv \left. \frac{\partial x_i}{\partial \alpha_j} \right|_{\alpha_j = \alpha_j^0} = \lim_{\delta \alpha \rightarrow 0} \left[\frac{x_i(\alpha_j^0 + \delta \alpha_j) - x_i(\alpha_j^0)}{\delta \alpha_j} \right], \quad (17)$$

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

$$\frac{dS_j}{dt} = \mathbf{M} \cdot S_j + \mathbf{D}_j \quad j = 1, \dots, m, \quad (18)$$

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

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= f(\mathbf{x}, \alpha), \quad \mathbf{x}(0) = \mathbf{x}_0, \\ \frac{dS_j}{dt} &= \mathbf{M} \cdot S_j + \mathbf{D}_j, \quad S_j(0) = S_{j0}. \end{aligned} \quad (19)$$

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 $\mathbf{x}(t)$, $t \in [t_0, \tau]$.
- (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):

$$\mathcal{F}(\mathbf{x}, \alpha) = \int_0^\tau \Phi(t; \mathbf{x}, \alpha) dt, \quad (20)$$

where Φ is a nonlinear function of the state variables \mathbf{x} and model parameters α . The gradient of the functional \mathcal{F} with respect to the parameters α around the unperturbed state vector \mathbf{x}^0

$$\nabla_{\alpha} \mathcal{F}(\mathbf{x}^0, \alpha^0) = \left(\frac{d\mathcal{F}}{d\alpha_1}, \dots, \frac{d\mathcal{F}}{d\alpha_m} \right)^T \Big|_{\mathbf{x}^0, \alpha^0} \quad (21)$$

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

$$\begin{aligned} \left. \frac{d\mathcal{F}}{d\alpha_j} \right|_{\alpha_j^0} &= \frac{\mathcal{F}(\mathbf{x}^0 + \delta \mathbf{x}^0; \alpha_1^0, \dots, \alpha_j^0 + \delta \alpha_j, \dots, \alpha_m^0) - \mathcal{F}(\mathbf{x}^0, \alpha^0)}{\delta \alpha_j}, \\ &\approx \frac{\mathcal{F}(\mathbf{x}^0 + \delta \mathbf{x}^0; \alpha_1^0, \dots, \alpha_j^0 + \delta \alpha_j, \dots, \alpha_m^0) - \mathcal{F}(\mathbf{x}^0, \alpha^0)}{\delta \alpha_j}, \end{aligned} \quad (22)$$

where $\delta \alpha_j$ is the variation in parameter α_j^0 . Note that

$$\frac{d\mathcal{F}}{d\alpha_j} = \sum_{i=1}^n \frac{\partial \mathcal{F}}{\partial x_i} \frac{\partial x_i}{\partial \alpha_j} + \frac{\partial \mathcal{F}}{\partial \alpha_j} = \sum_{i=1}^n S_{ij} \frac{\partial \mathcal{F}}{\partial x_i} + \frac{\partial \mathcal{F}}{\partial \alpha_j}. \quad (23)$$

This approach is acceptable for low-order models. However, the accuracy of sensitivity estimates strongly depends on choice of the perturbation $\delta\alpha_i$. By introducing the Gâteaux differential, the sensitivity analysis problem can be considered in the differential formulation eliminating the need to set the value of $\delta\alpha_i$ [30, 31]. The Gâteaux differential for the objective function (20) has the following form:

$$\delta\mathcal{F}(\mathbf{x}^0, \boldsymbol{\alpha}^0; \delta\mathbf{x}, \delta\boldsymbol{\alpha}) = \int_0^\tau \left(\frac{\partial\Phi}{\partial\mathbf{x}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} \cdot \delta\mathbf{x} + \frac{\partial\Phi}{\partial\boldsymbol{\alpha}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} \cdot \delta\boldsymbol{\alpha} \right) dt. \quad (24)$$

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

$$\begin{aligned} \frac{\partial\delta\mathbf{x}}{\partial t} &= \frac{\partial f}{\partial\mathbf{x}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} \cdot \delta\mathbf{x} + \frac{\partial f}{\partial\boldsymbol{\alpha}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} \cdot \delta\boldsymbol{\alpha}, \\ t \in [0, \tau], \quad \delta\mathbf{x}(0) &= \delta\mathbf{x}_0. \end{aligned} \quad (25)$$

Then using (24) we can calculate the variation $\delta\mathcal{F}$. Since $\delta\mathcal{F}(\mathbf{x}^0, \boldsymbol{\alpha}^0; \delta\mathbf{x}, \delta\boldsymbol{\alpha}) = \langle \nabla_{\boldsymbol{\alpha}}\mathcal{F}, \delta\boldsymbol{\alpha} \rangle$, 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 $\nabla_{\boldsymbol{\alpha}}\mathcal{F}$. 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 $\nabla_{\boldsymbol{\alpha}}\mathcal{F}$ can be calculated by the following equation:

$$\nabla_{\boldsymbol{\alpha}}\mathcal{F}(\mathbf{x}^0, \boldsymbol{\alpha}^0) = \int_0^\tau \left[\frac{\partial\Phi}{\partial\boldsymbol{\alpha}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} - \left(\frac{\partial f}{\partial\boldsymbol{\alpha}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} \right)^T \cdot \mathbf{x}^* \right] dt, \quad (26)$$

where the vector function \mathbf{x}^* is the solution of adjoint model

$$\begin{aligned} -\frac{\partial\mathbf{x}^*}{\partial t} - \left(\frac{\partial f}{\partial\mathbf{x}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0} \right)^T \mathbf{x}^* &= -\frac{\partial\Phi}{\partial\mathbf{x}} \Big|_{\mathbf{x}^0, \boldsymbol{\alpha}^0}, \\ t \in [0, \tau], \quad \mathbf{x}^*(\tau) &= 0. \end{aligned} \quad (27)$$

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 $\mathbf{x}(t)$, $t \in [t_0, \tau]$.
- (3) Calculate the right-hand side of (27) and then integrate numerically this equation in the inverse time direction with the initial conditions $\mathbf{x}^*(\tau) = 0$.
- (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 $t_D = -20$ with the initial conditions

$$\mathbf{x}(t_D) = (0.01, 0.01, 0.01, 0.02, 0.02, 0.02)^T \quad (28)$$

and finished at $t_0 = 0$ to guarantee that the calculated model state vector $\mathbf{x}_0 = \mathbf{x}(0)$ 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 $\mathbf{x}^t(t_i)$ and the “true” state $\mathbf{x}^t(t_f)$ 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 $\boldsymbol{\alpha}^0$ and initial conditions \mathbf{x}_0 . Then, the forecast trajectory $\mathbf{x}^f(t_k)$ and the forecast state $\mathbf{x}^f(t_f)$ at t_f are obtained by integration of the “forecast” model (2) with initial conditions \mathbf{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:

$$e(\mathbf{x}^f) = \left[\frac{(\mathbf{x}^f - \mathbf{x}^t)^T (\mathbf{x}^f - \mathbf{x}^t)}{(\mathbf{x}^t)^T \mathbf{x}^t} \right]^{1/2}. \quad (29)$$

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 $e(\mathbf{x}^f)$, 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 $\mathbf{x}^t(t_i)$, the background (first guess) state $\mathbf{x}^b(t_i)$, and observations $\mathbf{y}^0(t_i)$ inside the assimilation window $[t_0, t_N]$ as well as error covariance matrices of the prior guess \mathbf{B}_0 and observations \mathbf{R}_i . 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 $\mathbf{x}^t(t_i)$ and $\mathbf{x}^b(t_i)$, 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 $\sigma_0^2 = 0.2$) to the “true” state. Since observation grid and model grid are the same, observation

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

	Parameter variations in percent of unperturbed value					
	1%	-1%	5%	-5%	10%	-10%
$r_{2.5}$	0.1273	0.1266	0.2871	0.2796	0.4081	0.3890
$c_{2.5}$	0.0584	0.0583	0.1312	0.1300	0.1866	0.1833
$r_{5.0}$	0.1663	0.1689	0.3617	0.3202	0.5004	0.5742
$c_{5.0}$	0.0705	0.0704	0.1583	0.1569	0.2250	0.2210

operator \mathcal{H} is simply an identity mapping. To take into consideration the background covariances, for simplicity, the assumption $\mathbf{B}_0 = \sigma_b^2 \mathbf{I}$ can be used, where $\sigma_b^2 = 0.2$ is the variance of background error and \mathbf{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 $\mathbf{R}_i = \mathbf{R} = \sigma_0^2 \mathbf{I}$.

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

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

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

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

should be calculated. The TL model is valid if $e_R \rightarrow 0$ when $\zeta \rightarrow 0$. 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

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

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 $\nabla J(\mathbf{x}_0)$. A linear Taylor approximation of the cost function can be written as

$$J(\mathbf{x}_0 + \zeta \delta \mathbf{x}) \approx J(\mathbf{x}_0) + \zeta (\delta \mathbf{x})^T \nabla J(\mathbf{x}_0). \quad (33)$$

Let us introduce the following function:

$$\Psi(\zeta) = \frac{J(\mathbf{x}_0 + \zeta \delta \mathbf{x}) - J(\mathbf{x}_0)}{\zeta (\delta \mathbf{x})^T \nabla J(\mathbf{x}_0)}. \quad (34)$$

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

$$\delta \mathbf{x} = \frac{\nabla J(\mathbf{x}_0)}{\|\nabla J(\mathbf{x}_0)\|}, \quad (35)$$

TABLE 2: Results of verification of tangent linear model for $c = 0.8$ and $\delta \mathbf{x} = 10^{-2} \times \mathbf{x}_0$.

ζ	e_R
1	0.9182066027544249
10^{-1}	0.9997279965782743
10^{-2}	0.9999925468155463
10^{-3}	0.9999991929611531
10^{-4}	0.9999999534965012
10^{-5}	0.9999999911217883
10^{-6}	0.9999999914427087
10^{-7}	0.9999997435447022

TABLE 3: Results of verification of 4D-Var cost function gradient for $c = 0.8$ and $\delta \mathbf{x} = 10^{-2} \times \mathbf{x}_0$.

ζ	$\Psi(\zeta)$	$\log_{10}(\Psi(\zeta) - 1)$
10^{-4}	0.8727731981461396	-0.8954213897009056
10^{-5}	0.9975483775420343	-2.6105464089686840
10^{-6}	0.9998765512756632	-3.9085133935113787
10^{-7}	0.9999884562844123	-4.9376543818657845
10^{-8}	0.9999979865432855	-5.6960577024708600
10^{-9}	0.999998912431426	-6.9635433500893210
10^{-10}	0.999999244103234	-7.1215375125562080

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

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 $\|\delta \mathbf{x}(t)\| \approx \|\delta \mathbf{x}(0)\| e^{\lambda t}$, 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:

$$\begin{aligned} S_{1c} &= \frac{\partial x}{\partial c}, & S_{2c} &= \frac{\partial y}{\partial c}, & S_{3c} &= \frac{\partial z}{\partial c}, \\ S_{4c} &= \frac{\partial X}{\partial c}, & S_{5c} &= \frac{\partial Y}{\partial c}, & S_{6c} &= \frac{\partial Z}{\partial c}. \end{aligned} \quad (36)$$

The corresponding sensitivity equations can be written as

$$\begin{aligned} \dot{S}_{1c} &= \sigma(S_{2c} - S_{1c}) - cS_{4c} - X, \\ \dot{S}_{2c} &= rS_{1c} - S_{2c} - xS_{3c} - zS_{1c} + cS_{5c} + Y, \\ \dot{S}_{3c} &= xS_{2c} + yS_{1c} - bS_{3c} + cS_{6c} + Z, \\ \dot{S}_{4c} &= \varepsilon\sigma(S_{5c} - S_{4c}) - cS_{1c} - x, \\ \dot{S}_{5c} &= \varepsilon(rS_{4c} - S_{5c} - XS_{6c} - ZS_{4c}) + cS_{2c} + y, \\ \dot{S}_{6c} &= \varepsilon(XS_{5c} + YS_{4c} - bS_{6c}) - cS_{3c} - z. \end{aligned} \quad (37)$$

Sensitivity coefficients can be introduced for any particular model parameter. Since the parameter vector α consists of five components (σ, r, b, c , and ε), 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 Figure 4. Envelopes of these coefficients grow over time while sensitivity coefficients themselves exhibit oscillating behavior. Figure 5 provides more detailed information on the evolution of sensitivity coefficients (36) calculated on the time interval $[0, 5]$. It is known that sensitivity coefficient is a measure of the change in state variable due to the variation of the estimated parameter. Unfortunately, obtained sensitivity coefficients are inherently uninformative and misleading. We cannot make a clear conclusion from them about system sensitivity to variations in the parameter c . In this regard, the average values of sensitivity functions $\nabla_{\alpha} \langle \mathcal{F}(\alpha) \rangle$ over a certain period of time can be considered as one of the most important measures of sensitivity, where \mathcal{F} is a generic objective function (20). However, the gradient $\nabla_{\alpha} \langle \mathcal{F}(\alpha) \rangle$ cannot be correctly estimated within the framework of conventional methods of sensitivity analysis since for chaotic systems it is observed that [35, 36]

$$\nabla_{\alpha} \langle \mathcal{F}(\alpha) \rangle \neq \langle \nabla_{\alpha} \mathcal{F}(\alpha) \rangle. \quad (38)$$

This is because the integral

$$\mathcal{F} = \lim_{T \rightarrow \infty} \int_0^T \lim_{\delta\alpha \rightarrow 0} \left[\frac{\mathcal{F}(\alpha + \delta\alpha) - \mathcal{F}(\alpha)}{\delta\alpha} \right] dt \quad (39)$$

does not possess uniform convergence and two limits ($T \rightarrow \infty$ и $\delta\alpha \rightarrow 0$) would not commute.

Similar results were obtained when we considered the influence of variations in the parameter r on the system dynamics. This parameter plays an important role in the formation of system's dynamical structure and transition to chaotic behavior [32]. Figure 6 shows the differences between

components of state vector $\mathbf{x}(r^0)$ obtained with $r = r^0 = 28$ and $\mathbf{x}(r^0 + \delta r)$ obtained with $r = r^0 + \delta r$, where $\delta r = 0.01r^0 = 0.28$. Even a small perturbation in the parameter r generates a tangible difference between corresponding state variables. Let us define the following sensitivity coefficients:

$$\begin{aligned} S_{1r} &= \frac{\partial x}{\partial r}, & S_{2r} &= \frac{\partial y}{\partial r}, & S_{3r} &= \frac{\partial z}{\partial r}, \\ S_{4r} &= \frac{\partial X}{\partial r}, & S_{5r} &= \frac{\partial Y}{\partial r}, & S_{6r} &= \frac{\partial Z}{\partial r}. \end{aligned} \quad (40)$$

The associated system of sensitivity equations can be written as

$$\begin{aligned} \dot{S}_{1r} &= \sigma(S_{2r} - S_{1r}) - cS_{4r}, \\ \dot{S}_{2r} &= x + rS_{1r} - S_{2r} - (xS_{3r} + zS_{1r}) + cS_{5r}, \\ \dot{S}_{3r} &= (xS_{2r} + yS_{1r}) - bS_{3r} + cS_{6r}, \\ \dot{S}_{4r} &= \varepsilon\sigma(S_{5r} - S_{4r}) - cS_{1r}, \\ \dot{S}_{5r} &= \varepsilon[X + rS_{4r} - S_{5r} - (XS_{6r} + ZS_{4r})] + cS_{2r}, \\ \dot{S}_{6r} &= \varepsilon[(XS_{5r} + YS_{4r}) - bS_{6r}] - cS_{3r}. \end{aligned} \quad (41)$$

Envelopes of calculated sensitivity coefficients (40) grow over time and sensitivity coefficients demonstrate the oscillating behavior (Figures 7 and 8). Obtained sensitivity coefficients are also uninformative and inconclusive. The ‘‘shadowing’’ approach for estimating the system sensitivity to variations in its parameters [36, 37] allows us to calculate the average sensitivities $\langle \nabla_{\alpha} J(\alpha) \rangle$ and therefore to make a clear conclusion with respect to the system sensitivity to its parameters. A detailed description of two variants of the ‘‘shadowing’’ approach is provided in an appendix, and some results of numerical experiments are presented below.

The main problem arising in the ‘‘shadowing’’ method is to calculate the pseudotrajectory. We consider two sets of numerical experiments: weak coupling ($c = 0.01$) and strong coupling ($c = 0.8$) between fast and slow systems. Fast and slow variables that correspond to the original and pseudoorbits are shown in Figures 9 and 10 when the coupling strength parameter $c = 0.01$. The Least Square Shadowing variant of the ‘‘shadowing’’ approach was used to calculate pseudotrajectories. The differences between state variables corresponding to the original and pseudotrajectories of the fast and slow systems are plotted in Figure 11. These figures show that the calculated pseudoorbits are close to corresponding true trajectories over a specified time interval, demonstrating the shadowability. The strong coupling does not introduce significant qualitative and quantitative changes in the behavior of pseudotrajectories with respect to the true orbits. The original and pseudo fast and slow variables for $c = 0.8$ are shown in Figures 12 and 13, and the differences between these state variables are presented in Figure 14. Sensitivity estimates with respect to the parameter r calculated over the time interval $[0, 20]$ for different values of coupling strength parameter are shown in Table 4. The most sensitive variables are z and Z . The sensitivity of variables x ,

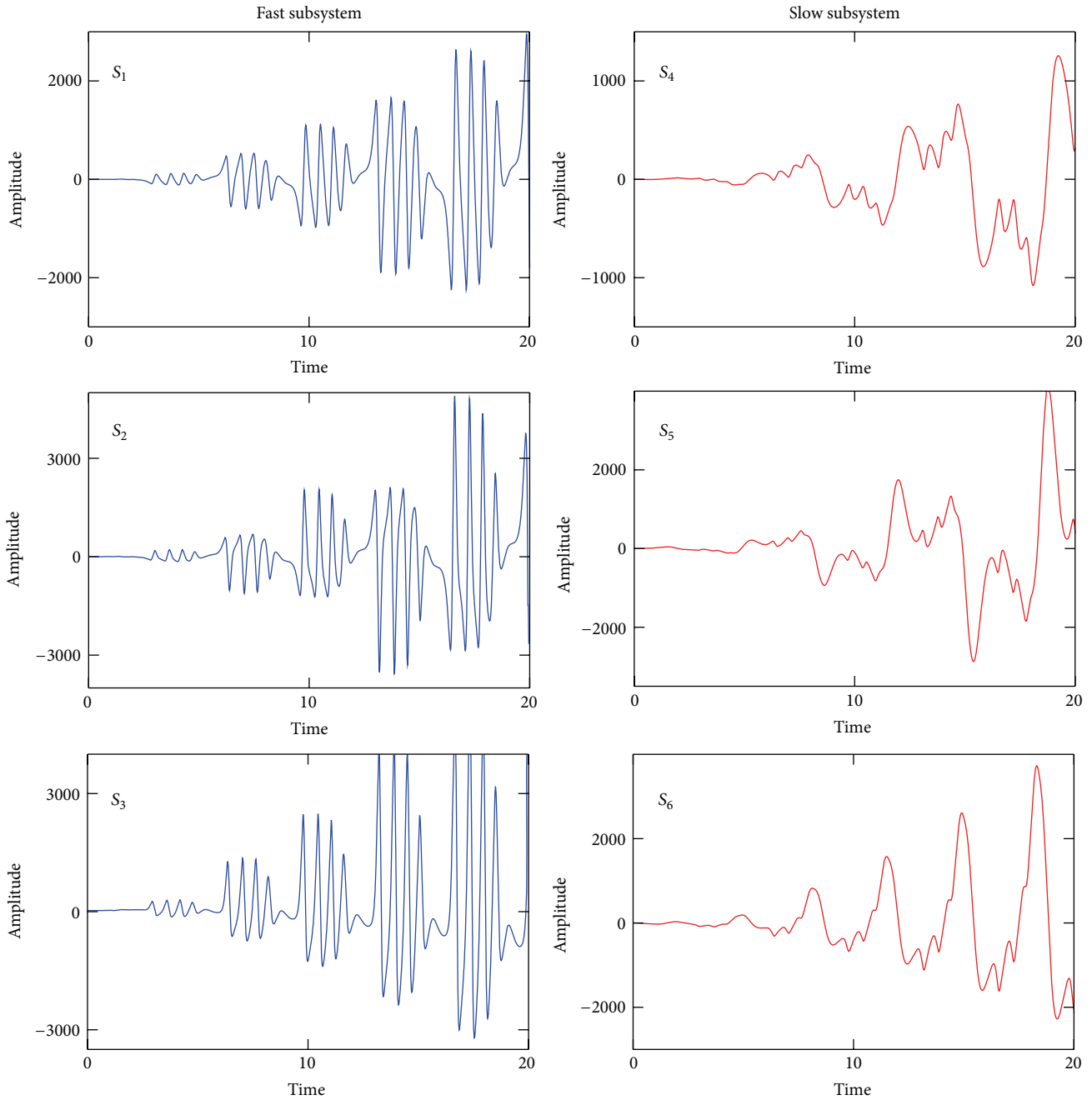


FIGURE 4: Time dynamics of sensitivity functions with respect to parameter c on the time interval $[0, 20]$ for $c^0 = 0.9$.

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

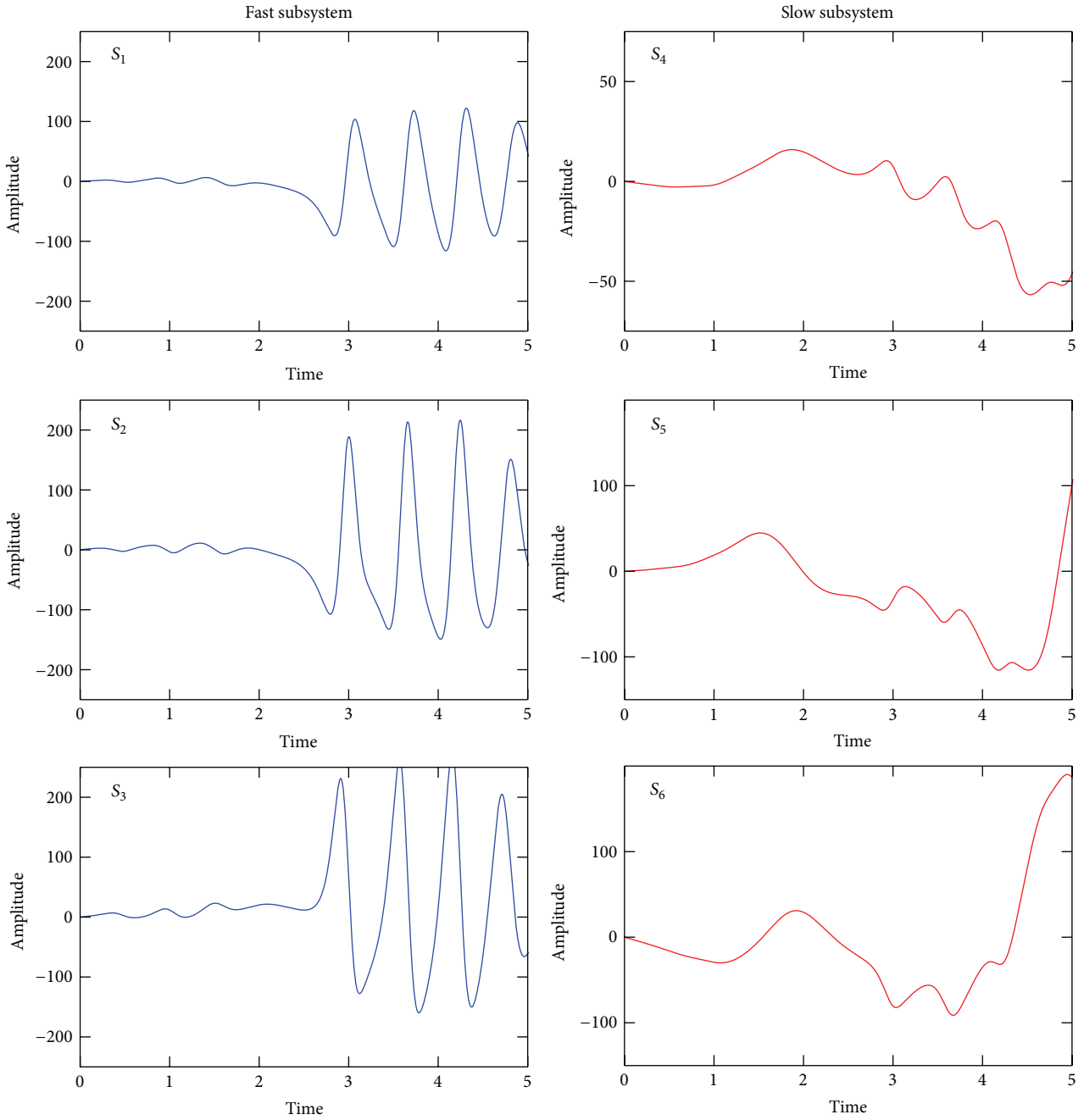


FIGURE 5: Time dynamics of sensitivity functions with respect to parameter c on the time interval $[0, 5]$ for $c^0 = 0.9$.

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

c	$\partial Z/\partial r$	$\partial Y/\partial r$	$\partial X/\partial r$	$\partial z/\partial r$	$\partial y/\partial r$	$\partial x/\partial r$
1.0	1.10	0.05	0.01	1.08	0.04	0.03
0.8	0.69	0.08	0.03	1.02	0.07	0.07
0.4	0.95	0.03	-0.01	1.03	0.09	0.09
0.15	0.91	-0.08	-0.09	1.01	-0.01	-0.01
10^{-4}	1.04	-0.02	-0.03	1.02	-0.01	-0.01

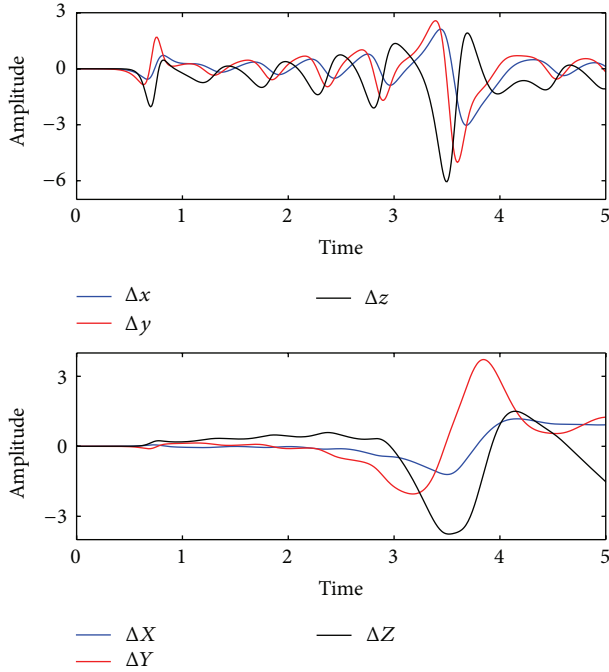


FIGURE 6: Difference between $\mathbf{x}(r^0)$ and $\mathbf{x}(r^0 + 0.01r^0)$ for fast and slow subsystems.

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 pseudo-orbits 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 pseudo-orbit 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 pseudo-orbit 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, 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

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

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 ε -traces X ; that is

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

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 ε -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 Φ^t if the inequalities

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

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 Φ^t , the shadowing property holds in a small neighborhood of a compact hyperbolic set for dynamical system Φ^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

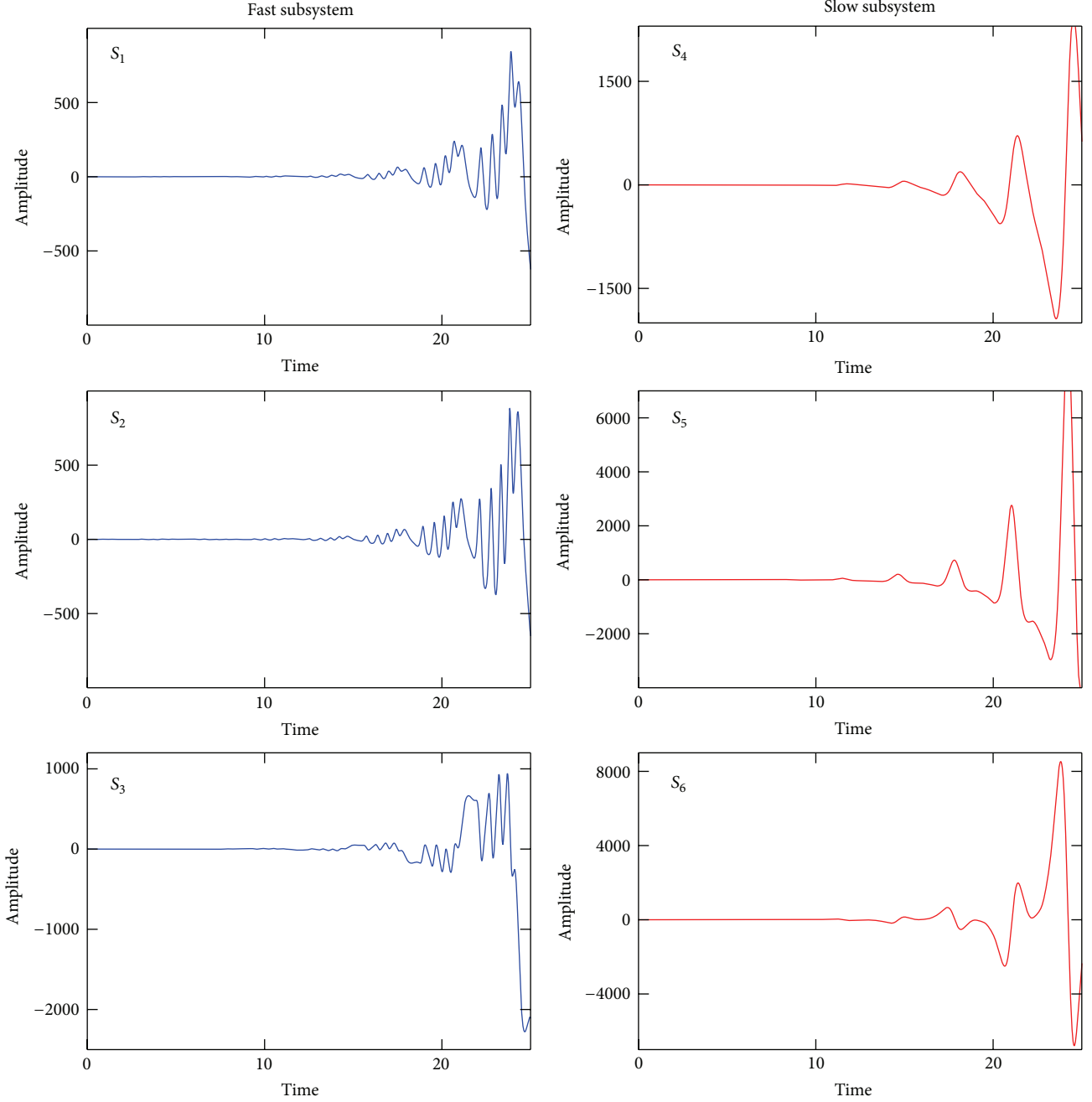


FIGURE 7: Time dynamics of sensitivity functions with respect to parameter r on the interval $[0, 25]$ for $c^0 = 0.9$.

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 $\varepsilon > 0$, $\text{Rep}(\varepsilon)$ 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 α :

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

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 $T \rightarrow \infty$, and an averaged

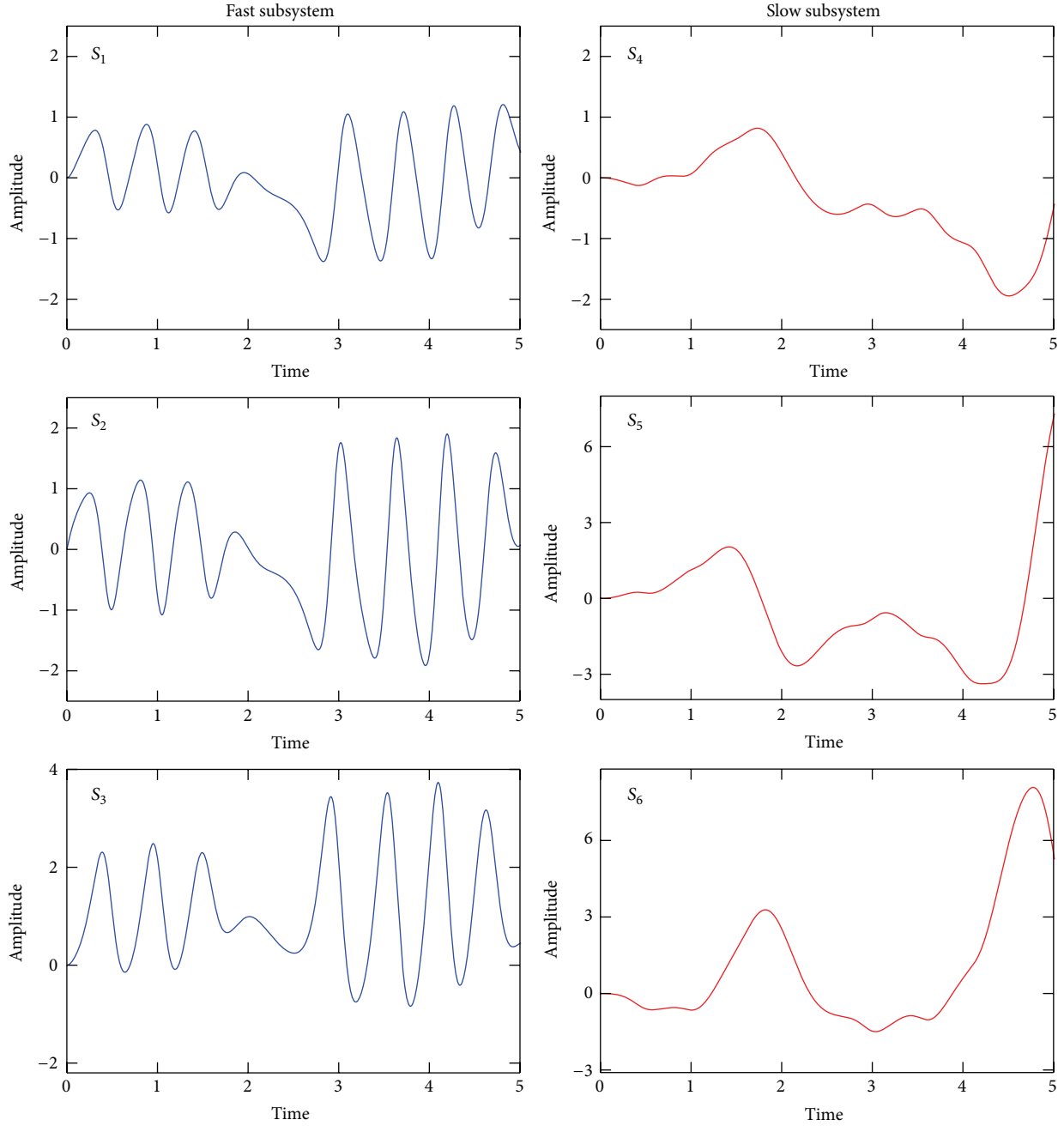


FIGURE 8: Time dynamics of sensitivity functions with respect to parameter r on the time interval $[0, 5]$ for $c^0 = 0.9$.

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

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

(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 $\frac{dx}{dt} = f(x, \alpha),$

(A.7)

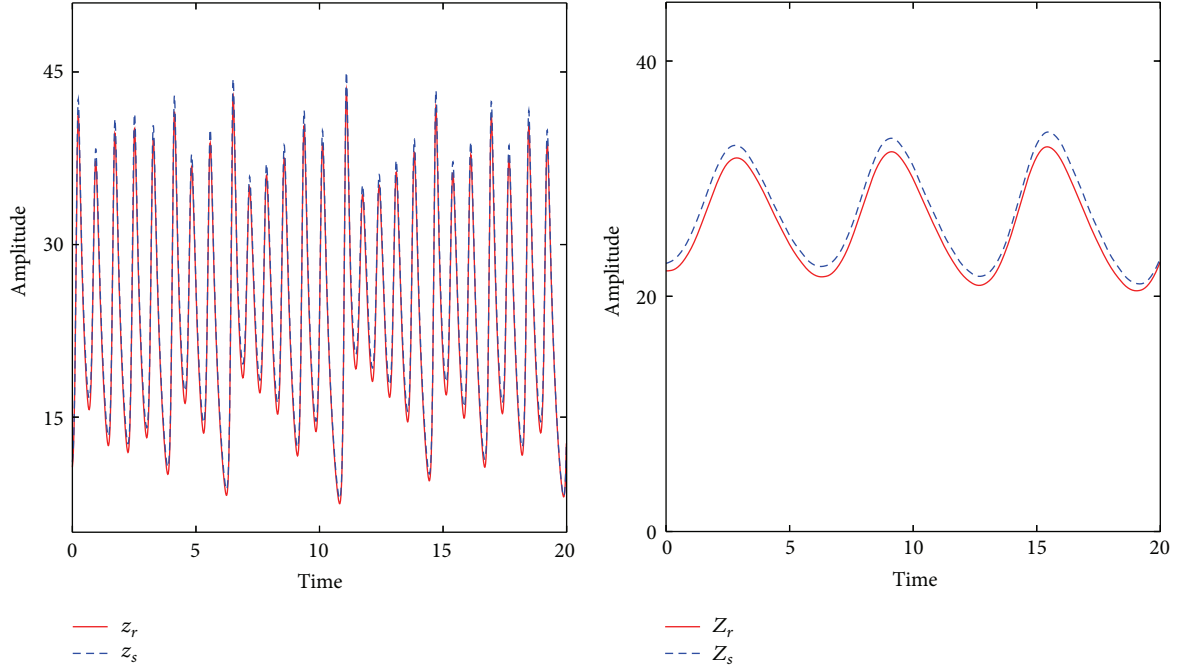


FIGURE 9: Original trajectory (in red) and pseudoorbit (in blue) for the fast z and slow Z variables for $c^0 = 0.01$.

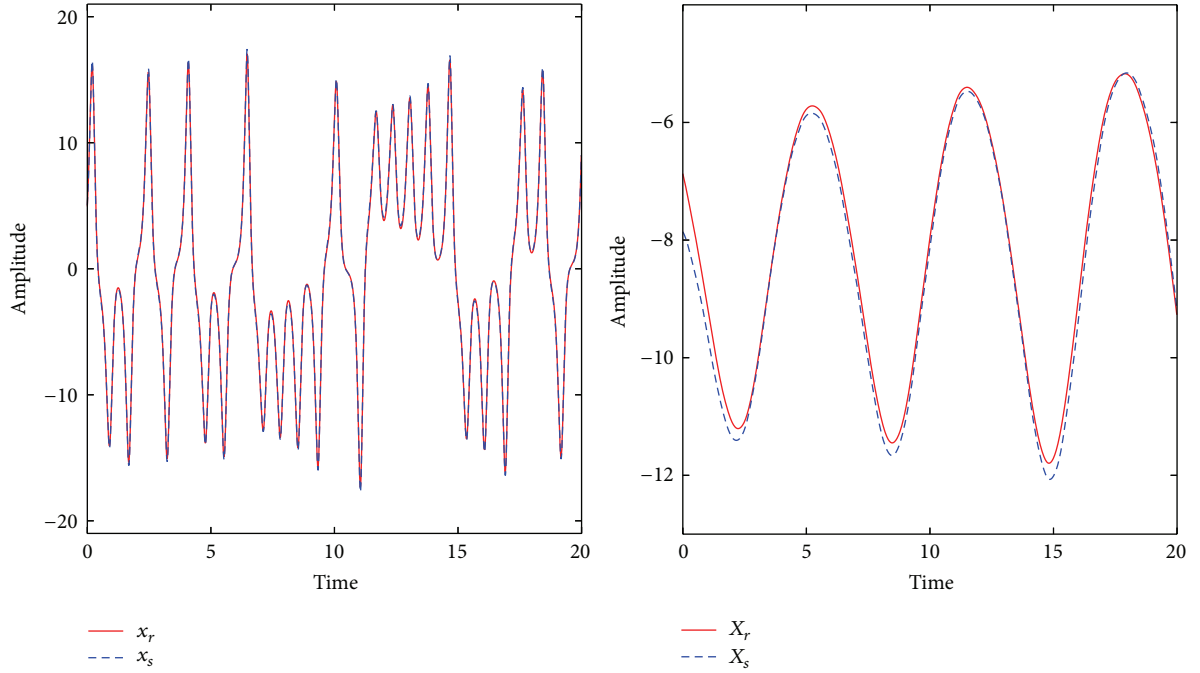


FIGURE 10: Original trajectory (in red) and pseudoorbit (in blue) for fast x and slow X variables for $c^0 = 0.01$.

where η 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

$$\langle J(\alpha) \rangle \approx \langle J_s^{(T)}(\alpha) \rangle = \frac{1}{\tau(T) - \tau(0)} \int_{\tau(0)}^{\tau(T)} J(x_s^{(T)}(t, \alpha), \alpha), \quad (\text{A.8})$$

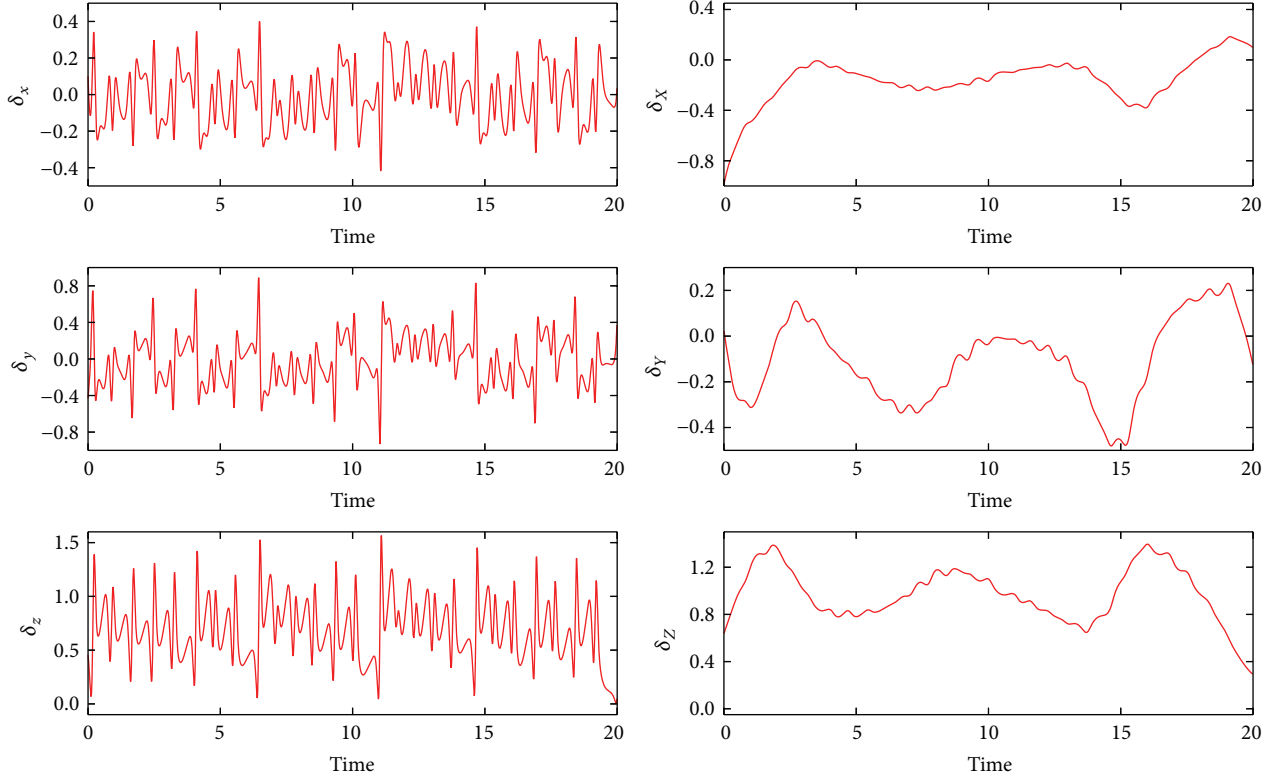


FIGURE 11: Difference between variables that correspond to the original trajectory and pseudorbit for $c^0 = 0.01$.

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

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

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 δf . To solve this problem, functions δx and δf are decomposed into their constituent Lyapunov covariant vectors $v_1(x), \dots, v_n(x)$:

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

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

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

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

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

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

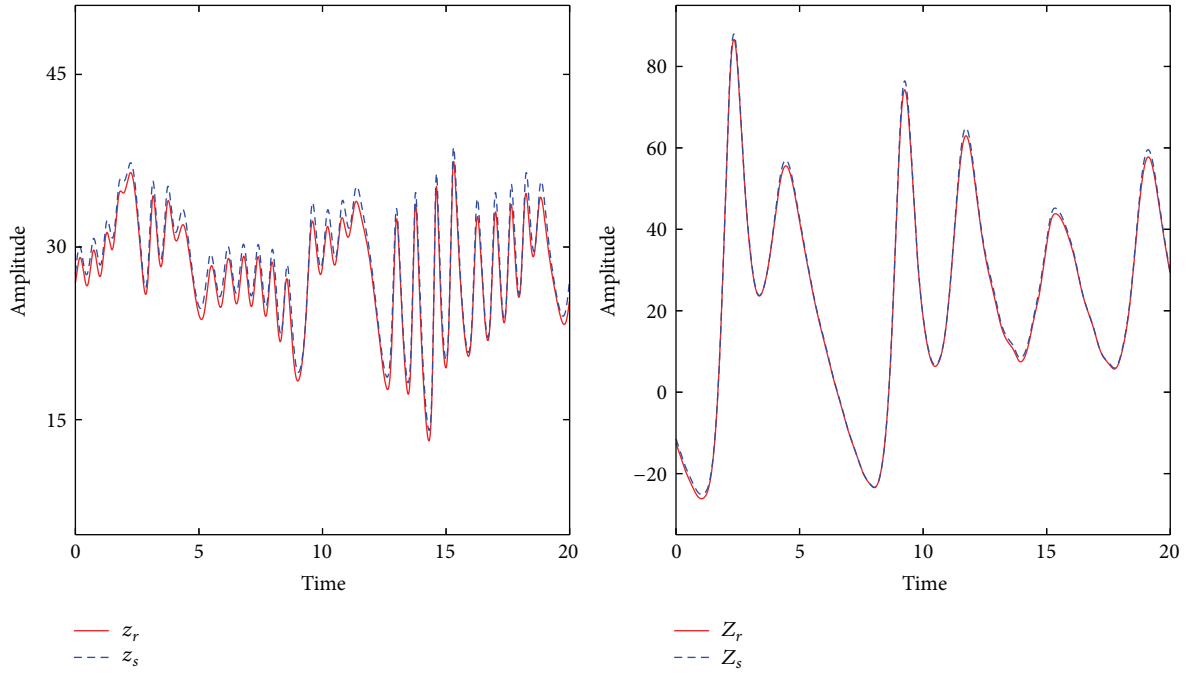
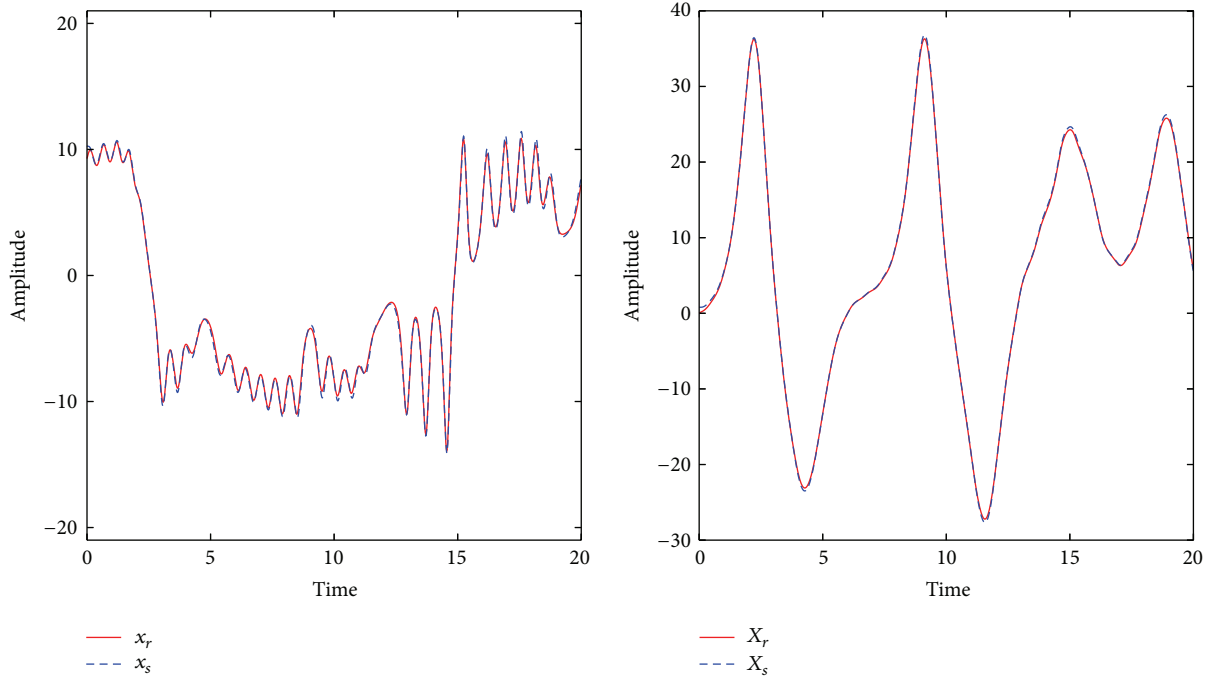
From (A.15a) and (A.15b) and (A.18) and the relation $\delta f(x) = A\delta x(x)$, we get

$$\delta f(x) = \sum_{i=1}^n A(\psi_i v_i) = \sum_{i=1}^n \underbrace{\left(\frac{d\psi_i}{dt} - \lambda_i \psi_i \right)}_{\varphi_i} v_i. \quad (\text{A.19})$$

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

$$\frac{d\psi_i(x)}{dt} = \varphi_i(x) + \lambda_i \psi_i(x). \quad (\text{A.20})$$

Thus, we can calculate $\psi_i(x)$ using (A.20) by first decomposing δf as a sum (A.15b), and then the desired δx can be obtained from (A.15a). However, if dynamical system has a zero Lyapunov exponent, $\lambda_i^0 = 0$, then the algorithm described above fails to compute δx [36]. The problem can


 FIGURE 12: Original trajectory (in red) and pseudoorbit (in blue) for the fast z and slow Z variables for $c^0 = 0.8$.

 FIGURE 13: Original trajectory (in red) and pseudoorbit (in blue) for fast x and slow X variables for $c^0 = 0.8$.

be resolved by introducing a time-dilation variable μ that satisfies the following equation:

$$\mu + \langle \varphi_i^0 \rangle = 0, \quad (\text{A.21})$$

where

$$\langle \varphi_i^0 \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} [\psi_i^0(x(T)) - \psi_i^0(x(0))]. \quad (\text{A.22})$$

In the presence of the variable μ , the expression for calculating δx takes the following form: $\delta x = A^{-1}(\delta f + \mu f)$. The supplement μf affects (A.20) only for $\lambda_i^0 = 0$:

$$\frac{d\psi_i^0(x)}{dt} = \varphi_i^0(x) + \mu. \quad (\text{A.23})$$

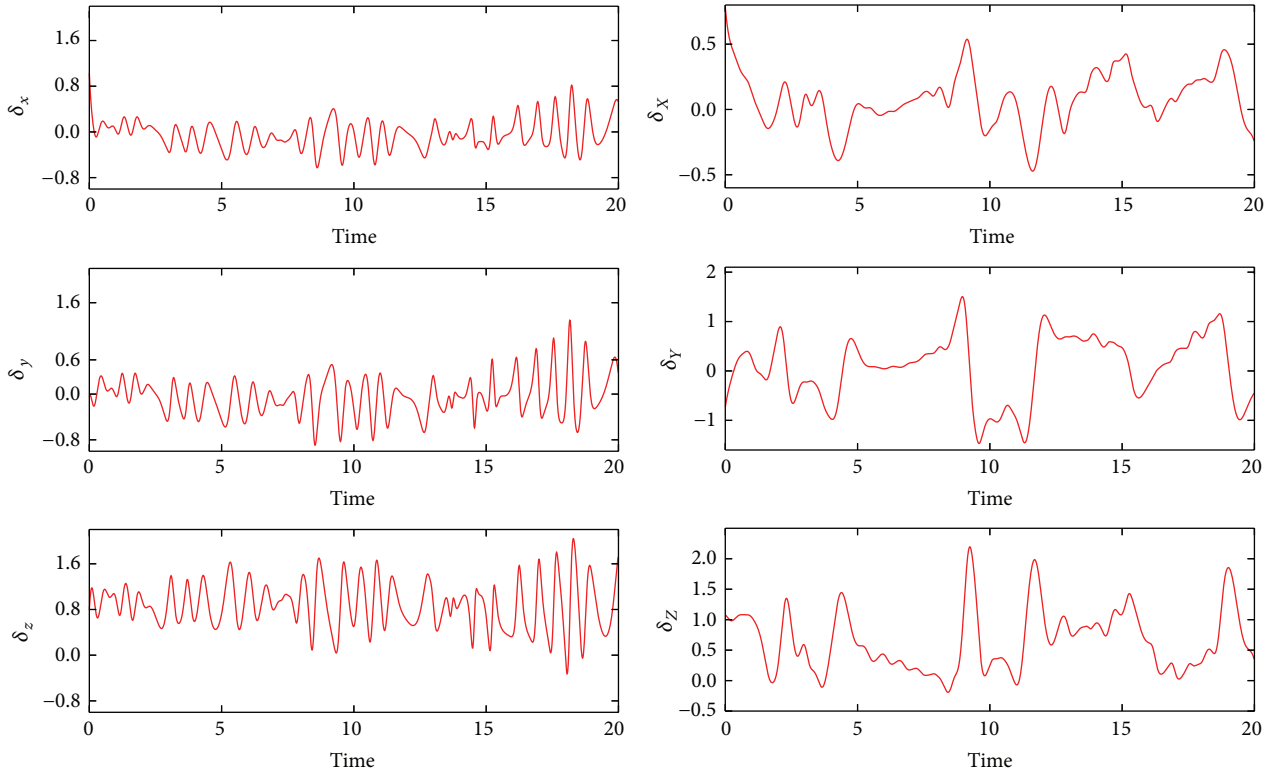


FIGURE 14: Difference between variables that correspond to the original trajectory and pseudoorbit for $c^0 = 0.8$.

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 λ_i and the Lyapunov covariant vectors $v_i(x(t))$, $i = 1, \dots, n$.
- (4) Define $\delta f = (\partial f / \partial \alpha) \delta \alpha$ and execute the Lyapunov spectrum decomposition of δf along the trajectory $x(t)$ to obtain $\varphi_i(x)$, $i = 1, \dots, n$.
- (5) Calculate the time-dilation variable μ using (A.21).
- (6) Compute δx along the trajectory $x(t)$.
- (7) Estimate the sensitivity $S_\alpha = \partial x / \partial \alpha$ by averaging over the time interval $t \in [0, 20]$.

Conflict of Interests

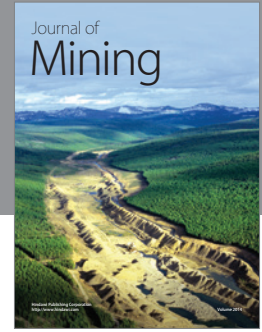
The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] F. Rawlins, S. P. Ballard, K. J. Bovis et al., "The met office global four-dimensional variational data assimilation scheme," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. 623, pp. 347–362, 2007.
- [2] K. Puri, G. Dietachmayer, P. Steinle et al., "Implementation of the initial ACCESS numerical weather prediction system," *Australian Meteorological and Oceanographic Journal*, vol. 63, no. 2, pp. 265–284, 2013.
- [3] R. Daley, *Atmospheric Data Analysis*, Cambridge University Press, New York, NY, USA, 1991.
- [4] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, New York, NY, USA, 2003.
- [5] F. Rabier, "Overview of global data assimilation developments in numerical weather-prediction centres," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3215–3233, 2005.
- [6] I. M. Navon, "Data assimilation for numerical weather prediction: a review," in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, S. K. Park and L. Xu, Eds., Springer, New York, NY, USA, 2009.
- [7] D. G. Cacuci, I. M. Navon, and M. Ionescu-Bujor, *Computational Methods for Data Evaluation and Assimilation*, CRC Press, Boca Raton, Fla, USA, 2013.
- [8] I. Sasaki, "An objective analysis based on variational method," *Journal of the Meteorological Society of Japan*, vol. 36, no. 3, pp. 29–30, 1958.
- [9] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, Germany, 1971.
- [10] V. V. Penenko and N. N. Obraztsov, "Variational method for assimilation of meteorological fields," *Meteorology and Hydrology*, no. 11, pp. 3–16, 1976.

- [11] F.-X. Le Dimet, "Une application des méthodes de contrôle optimal à l'analyse variationnelle," Rapport Scientifique LAMP 63170, Université Blaise Pascal, Clermont-Ferrand, France, 1981.
- [12] O. Talagrand, "A study of the dynamics of four-dimensional data assimilation," *Tellus*, vol. 33, no. 1, pp. 43–60, 1981.
- [13] D. G. Cacuci, "Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach," *Journal of Mathematical Physics*, vol. 22, no. 12, pp. 2794–2802, 1981.
- [14] D. G. Cacuci, "Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses," *Journal of Mathematical Physics*, vol. 22, no. 12, pp. 2803–2812, 1981.
- [15] J. M. Lewis and J. C. Derber, "The use of adjoint equations to solve a variational adjustment problem with advective constraints," *Tellus A*, vol. 37, no. 4, pp. 309–322, 1985.
- [16] F.-X. Le Dimet and O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects," *Tellus Series A*, vol. 38, no. 2, pp. 97–110, 1986.
- [17] O. Talagrand and P. Courtier, "Variational assimilation of meteorological observations with the adjoint vorticity equation. Part 1. Theory," *Quarterly Journal of the Royal Meteorological Society*, vol. 113, pp. 1311–1328, 1987.
- [18] P. Courtier and O. Talagrand, "Variational assimilation of meteorological observations with the adjoint vorticity equation. II: numerical results," *Quarterly Journal of the Royal Meteorological Society*, vol. 113, no. 478, pp. 1329–1347, 1987.
- [19] F.-X. le Dimet, H.-E. Ngodock, B. Luong, and J. Verron, "Sensitivity analysis in variational data assimilation," *Journal of the Meteorological Society of Japan*, vol. 75, no. 1, pp. 245–255, 1997.
- [20] N. L. Baker and R. Daley, "Observation and background adjoint sensitivity in the adaptive observation-targeting problem," *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 565, pp. 1431–1454, 2000.
- [21] G. Hello and F. Bouttier, "Using adjoint sensitivity as a local structure function in variational data assimilation," *Nonlinear Processes in Geophysics*, vol. 8, no. 6, pp. 347–355, 2001.
- [22] R. H. Langland and N. L. Baker, "Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system," *Tellus Series A*, vol. 56, no. 3, pp. 189–203, 2004.
- [23] R. M. Errico, "Interpretations of an adjoint-derived observational impact measure," *Tellus, Series A: Dynamic Meteorology and Oceanography*, vol. 59, no. 2, pp. 273–276, 2007.
- [24] R. Gelaro, Y. Zhu, and R. M. Errico, "Examination of various-order adjoint-based approximations of observation impact," *Meteorologische Zeitschrift*, vol. 16, no. 6, pp. 685–692, 2007.
- [25] D. N. Daescu, "On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation," *Monthly Weather Review*, vol. 136, no. 8, pp. 3050–3065, 2008.
- [26] Y. Trémolet, "Computation of observation sensitivity and observation impact in incremental variational data assimilation," *Tellus, Series A: Dynamic Meteorology and Oceanography*, vol. 60, no. 5, pp. 964–978, 2008.
- [27] D. N. Daescu and R. Todling, "Adjoint estimation of the variation in model functional output due to the assimilation of data," *Monthly Weather Review*, vol. 137, no. 5, pp. 1705–1716, 2009.
- [28] R. Gelaro and Y. Zhu, "Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models," *Tellus, Series A: Dynamic Meteorology and Oceanography*, vol. 61, no. 2, pp. 179–193, 2009.
- [29] D. N. Daescu and I. M. Navon, "Sensitivity analysis in nonlinear data assimilation: theoretical aspects and applications," in *Advances Numerical Methods for Complex Environmental Models: Needs and Availability*, I. Farago and Z. Zlatev, Eds., pp. 276–300, Benham Science Publishers, 2013.
- [30] D. G. Cacuci, *Sensitivity and Uncertainty Analysis, Volume I: Theory*, CRC, Boca Raton, Fla, USA, 2003.
- [31] D. G. Cacuci, M. Ionescu-Bujor, and I. M. Navon, *Sensitivity and Uncertainty Analysis*, CRC, Boca Raton, Fla, USA, 2005.
- [32] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [33] M. Peña and E. Kalnay, "Separating fast and slow modes in coupled chaotic systems," *Nonlinear Processes in Geophysics*, vol. 11, no. 3, pp. 319–327, 2004.
- [34] L. Siqueira and B. Kirtman, "Predictability of a low-order interactive ensemble," *Nonlinear Processes in Geophysics*, vol. 19, no. 2, pp. 273–282, 2012.
- [35] D. J. Lea, M. R. Allen, and T. W. N. Haine, "Sensitivity analysis of the climate of a chaotic system," *Tellus*, vol. 52, no. 5, pp. 523–532, 2000.
- [36] Q. Wang, "Forward and adjoint sensitivity computation of chaotic dynamical systems," *Journal of Computational Physics*, vol. 235, no. 15, pp. 1–13, 2013.
- [37] Q. Wang, R. Hu, and P. Blonigan, "Least squares shadowing sensitivity analysis of chaotic limit cycle oscillations," *Journal of Computational Physics*, vol. 267, pp. 210–224, 2014.
- [38] S. Y. Pilyugin, *Shadowing in Dynamical Systems*, vol. 1706 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1999.
- [39] K. J. Palmer, *Shadowing in Dynamical Systems. Theory and Applications*, Kluwer Academic, Dordrecht, The Netherlands, 2000.
- [40] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, "Phase synchronization of chaotic oscillators," *Physical Review Letters*, vol. 76, no. 11, pp. 1804–1807, 1996.
- [41] F. Christiansen, G. Paladin, and H. H. Rugh, "Determination of correlation spectra in chaotic systems," *Physical Review Letters*, vol. 65, no. 17, pp. 2087–2090, 1990.
- [42] C. Liverani, "Decay of correlations," *Annals of Mathematics*, vol. 142, no. 2, pp. 239–301, 1995.
- [43] S. Panchev and T. Spassova, "Simple general atmospheric circulation and climate models with memory," *Advances in Atmospheric Sciences*, vol. 22, no. 5, pp. 765–769, 2005.
- [44] E. Rosenwasser and R. Yusupov, *Sensitivity of Automatic Control Systems*, CRC Press, Boca Raton, Fla, USA, 2000.
- [45] I. M. Navon, X. Zou, J. Derber, and J. Sela, "Variational data assimilation with an adiabatic version of the NMC spectral model," *Monthly Weather Review*, vol. 120, no. 7, pp. 1433–1446, 1992.
- [46] D. V. Anosov, "On class of invariant sets of smooth dynamical systems," in *Proceedings of the 5th International Conference on Nonlinear Oscillations*, vol. 2, pp. 39–45, Kiev, Ukraine, 1975.
- [47] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, vol. 470 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1975.
- [48] A. Koropeccki and E. R. Pujals, "Some consequences of the shadowing property in low dimensions," *Ergodic Theory and Dynamical Systems*, vol. 34, no. 4, pp. 1273–1309, 2014.
- [49] D. Richeson and J. Wiseman, "Symbolic dynamics for nonhyperbolic systems," *Proceedings of the American Mathematical Society*, vol. 138, no. 12, pp. 4373–4385, 2010.

- [50] A. A. Petrov and S. Y. Pilyugin, "Shadowing near nonhyperbolic fixed points," *Discrete and Continuous Dynamical Systems*, vol. 34, no. 9, pp. 3761–3772, 2013.
- [51] S. Y. Pilyugin, *Spaces of Dynamical Systems*, De Gruyter, Berlin, Germany, 2012.
- [52] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, vol. 54 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, New York, NY, USA, 1995.
- [53] D. V. Anosov, *Geodesic Flows on Closed Riemann Manifolds with Negative Curvature*, vol. 90 of *Proceedings of the Steklov Institute of Mathematics*, 1967.
- [54] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.
- [55] Q. Wang, "Convergence of the least squares shadowing method for computing derivative of ergodic averages," *SIAM Journal on Numerical Analysis*, vol. 52, no. 1, pp. 156–170, 2014.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

