atmosphere

MDPI

Article

Role of the Observability Gramian in Parameter Estimation: Application to Nonchaotic and Chaotic Systems via the Forward Sensitivity Method †

John M. Lewis [1,2,*] and Sivaramakrishnan Lakshmivarahan [3]

[1] National Severe Storms Laboratory, Norman, OK 73072, USA
[2] Desert Research Institute, Reno, NV 89512, USA
[3] School of Computer Science, University of Oklahoma, Norman, OK 73069, USA
* Correspondence: john.lewis@dri.edu
† This paper is dedicated to the memory of Sherry Lynn Lewis.

Abstract: Data assimilation in chaotic regimes is challenging, and among the challenging aspects is placement of observations to induce convexity of the cost function in the space of control. This problem is examined by using Saltzman's spectral model of convection that admits both chaotic and nonchaotic regimes and is controlled by two parameters—Rayleigh and Prandtl numbers. The problem is simplified by stripping the seven-variable constraint to a three-variable constraint. Since emphasis is placed on observation positioning to avoid cost-function flatness, forecast sensitivity to controls is needed. Four-dimensional variational assimilation (4D-Var) is silent on this issue of observation placement while Forecast Sensitivity Method (FSM) delivers sensitivities used in placement. With knowledge of the temporal forecast sensitivity matrix V, derivatives of the forecast variables to controls, the cost function can be expressed as a function of the observability Gramian V^TV using first-order Taylor series expansion. The goal is to locate observations at places that force the Gramian positive definite. Further, locations are chosen such that the condition number of V^TV is small and this guarantees convexity in the vicinity of the cost function minimum. Four numerical experiments are executed, and results are compared with the structure of the cost function independently determined though arduous computation over a wide range of the two nondimensional numbers. The results are especially good based on reduction in cost function value and comparison with cost function structure.

Keywords: Rayleigh–Bénard convection; data assimilation; observation placement; forecast sensitivity; chaotic regimes; low-order modeling

check for updates

Citation: Lewis, J.M.; Lakshmivarahan, S. Role of the Observability Gramian in Parameter Estimation: Application to Nonchaotic and Chaotic Systems via the Forward Sensitivity Method. *Atmosphere* **2022**, *13*, 1647. https://doi.org/10.3390/atmos13101647

Academic Editors: Bo-Wen Shen, Roger A. Pielke Sr. and Xubin Zeng

Received: 16 August 2022
Accepted: 28 September 2022
Published: 10 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

In his pioneering work, Saltzman [1] derived the well-known spectral model of Rayleigh–Bénard convection [2,3]. It was a 7-variable double Fourier series representation of the convection herein called Saltzman's low-order model and denoted by S-LOM (7). His model included a single parameter, the Rayleigh number R. In the governing equations, the ratio of the Rayleigh number to the critical Rayleigh number R_c, $\lambda = \frac{R}{R_c}$, becomes the key parameter. He assumed the other nondimensional number associated with this convection problem, the Prandtl number σ, was set equal to 10. Historically, Saltzman's [1] model achieved fame by providing Lorenz [4] with a nonlinear deterministic model that exhibits nonperiodic solution to explore the limitations of extended-range predictability. Within a decade, Lorenz's three-variable version of Saltzman's model became known as the chaotic-butterfly model [5], so-named because the phase-space depiction of model evolution resembled butterfly wings.

The goal of the research is to explore data assimilation [6] for the Rayleigh–Bénard convection in both chaotic and nonchaotic regimes. Two three-variable stripped-down

versions of S-LOM (7) are used as dynamical constraints: (1) S-LOM (3) based on Saltzman's scaling, and L-LOM (3) based on Lorenz's scaling. These truncated systems contain the ingredients central to defining observation placement that yield a cost function minimum. The strength of this methodology rests on expressing the cost function (to a first-order accuracy around the true but unknown minimum) as a quadratic function of the increment in the control variable with (the symmetric and positive semi-definite) observability Gramian as the matrix of this approximate quadratic function. It is shown that by choosing the observation sites that maximize the square of the forecast sensitivities to control, we can indeed force the observability Gramian to be a positive definite matrix. This in turn avoids the occurrence of flat patches of the cost function in the control space. Refer to Lakshmivarahan et al. [7,8] and Lewis et al. [9] for further details and applications. An alternate strategy would be to place observations at locations that would lead to a small condition number for the observability Gramian.

Following this Introduction, we summarize S-LOM (7) and its reduction to the three-variable form [10] in Section 2. In Section 2, we also give attention to the scaling of S-LOM (3) and L-LOM (3). Discussion of the three-variable spectral dynamics is presented alongside graphics for nonchaotic and chaotic regimes in Section 3. Mechanics for data assimilation follow in Section 4. Data assimilation experiments are executed in Section 5, and discussion along with conclusions end the paper in Section 6.

2. Three-Variable Forms of Saltzman's Model

2.1. Spectral Form of Solution

The laboratory experiments of Bénard [2] were conducted with a rigid surface (a metal plate) below and a free upper surface while later experimental investigations used rigid surfaces at top and below (Chandrasekhar 1961 [11], Ch. 2, sect. 18c; Turner 1973 [12], Ch. 7, sect. 7.2). Saltzman's [1] theoretical development assumed two free surfaces. The onset of convection has different critical Rayleigh numbers R_c associated with the various boundary conditions—roughly 600 for free surfaces, 1800 for rigid surfaces, and 1200 for one-free and one-rigid surface. We use the free-surfaces value of $R_c = 657.5$ to develop the data assimilation model.

In an earlier technical report by the authors (Lewis and Lakshmivarahan [13]), a re-derivation of S-LOM (7) led to functional forms of coefficients in the equations that governed the Fourier amplitudes. As it stood in Saltzman's paper [1], the coefficients were expressed in numerical values only. Since our goal was to use Saltzman's equations in FSM data assimilation with the two nondimensional numbers as parameters, we needed to know the coefficients in terms of the wavenumbers and variable nondimensional numbers.

The three-variable form of Saltzman's equations used in this study is a subset of his seven-variable form. By setting the initial condition on the 3-horizontal wave temperature departure to 1 while the other initial conditions are set to zero delivers the three-variable form involving the (horizontal wavenumber 3/vertical wavenumber 1) streamfunction, the (horizontal wavenumber 3/vertical wavenumber 1) temperature departure, and the (vertical wavenumber 2) temperature departure. Let us label the three amplitudes in S-LOM (3) as follows: x_1 for the streamfunction (corresponding to amplitude A in Saltzman's notation). Temperature-departure components x_4 and x_7 (corresponding to amplitudes D and G in Saltzman' notation). This revised notation is consistent with the ordering of coefficients from 1 to 7 in Lewis and Lakshmivarahan [13] where $x_1, x_2, \ldots, x_7$ replaced A, B, . . . , G.

The streamfunction ψ in S-LOM (3) is given by

$$\psi(x, z, t) = -4x_1(t)\sin(\pi a x)\sin(\pi z) \tag{1}$$

where a, the nondimensional wavenumber, is given by $a = \frac{1}{\sqrt{2}}$ which defines the 3-horizontal wavenumber wave, and the temperature departure θ is given by

$$\theta(x, z, t) = 4x_4(t)\cos(\pi a x)\sin(\pi z) + 4x_7(t)\sin(2\pi z) \tag{2}$$

2.2. Amplitude Equations

When Equations (1) and (2) are substituted into the partial differential equations that govern Rayleigh–Bénard convection, the two-dimensional (x, z) vorticity equation and the temperature departure equation, followed by imposition of integral orthogonality conditions for the double Fourier series, ordinary differential equations governing the time-dependent amplitudes follow where we have expressed coefficients to two decimal places (same accuracy as used by Saltzman [1]):

Saltzman's S-LOM (3) equations:

$$\begin{aligned} \dot{x_1} &= -\sigma\pi^2\left(1+a^2\right)x_1 - \sigma\tfrac{ax_4}{\pi(1+a^2)} \\ &= -14.80\sigma\, x_1 - 0.15\sigma x_4 \end{aligned} \tag{3}$$

$$\begin{aligned} \dot{x_4} &= -4\pi^2 a x_1 x_4 - \pi a R_c \lambda x_1 - \pi^2\left(1+a^2\right)x_4 \\ &= -27.92 x_1 x_7 - 1460.62\lambda x_1 - 14.80 x_4 \end{aligned} \tag{4}$$

$$\begin{aligned} \dot{x_7} &= 2\pi^2 a x_1 x_4 - 4\pi^2 x_7 \\ &= 13.96 x_1 x_4 - 39.48 x_7 \end{aligned} \tag{5}$$

where

$$\lambda = \frac{R}{R_c}\ (R : \textit{Rayleigh}\ \text{number}) \tag{6}$$

Lorenz [4] used another scaling for the derivation of his L-LOM (3) that took the form:

Lorenz's L-LOM (3) equations:

$$\dot{x_1} = -\sigma x_1 + \sigma x_4 \tag{7}$$

$$\dot{x_4} = -x_1 x_7 + \lambda x_1 - x_4 \tag{8}$$

$$\dot{x_7} = x_1 x_4 - b x_7 \tag{9}$$

where

$$b = \frac{4}{1+a^2} = \frac{8}{3} \tag{10}$$

Table 1 shows non-dimensional and scaling for S-LOM (3) and L-LOM (3). The different signs on the right-hand sides of S-LOM (3) and L-LOM (3) stem from the fact that Lorenz assumed double Fourier representation of temperature departure and streamfunction as follows:

$$\theta(x,z,t) = x_4(t)\cos(\pi a x)\sin(\pi z) - x_7(t)\sin(2\pi z) \tag{11}$$

$$\psi(x,z,t) = x_1(t)\sin(\pi a x)\sin(\pi z) \tag{12}$$

where the sign of the x_7-component term in Equation (11) is opposite the corresponding term in Equation (2) and the sign of the streamfunction in Equation (12) is opposite the sign in Equation (1). These sign differences lead to a shift of temperature departure structure relative to the streamfunction structure in the convective regime compared to Saltzman's results. However, temperature–motion structures in S-LOM (3) and L-LOM (3) are both faithful to Rayleigh–Bénard cellular convection.

$$\frac{1}{(1+a^2)\pi^2} = 0.068;\ \ \frac{1+a^2}{a} = 2.121;\ \ \frac{\left[\pi(1+a^2)\right]^3}{a^2} = 209.292 \tag{13}$$

If the physical parameters are taken to be those associated with water at 20 °C (Chandrasekhar [11] (1961, Ch. 2, Table V1), the physical parameters are given by

$$\begin{aligned} \nu &= 1.006{\cdot}10^{-2}\ \text{cm}^2\ \text{sec}^{-1} \\ \kappa &= 1.433{\cdot}10^{-3}\ \text{cm}^2\ \text{sec}^{-1} \\ \epsilon &= 2.0{\cdot}10^{-3}\ {}^{\circ}\text{C}^{-1} \end{aligned} \tag{14}$$

Table 1. Nondimensional forms of Saltzman and Lorenz variables. The superscript asterisk (*) denotes the nondimensional variables where physical parameters for the model follow: g, H, κ, ν, ϵ: acceleration of gravity, depth of the fluid, coefficient of thermal diffusivity, kinematic viscosity, and coefficient of volume expansion, respectively.

Saltzman	Lorenz
$t = \frac{H^2}{\kappa} t^*$	$t = \frac{H^2}{\kappa} t^* \frac{1}{(1+a^2)\pi^2}$
$x = Hx^*, z = Hz^*$	$x = Hx^*, z = Hz^*$
$\Psi = \kappa\psi^*$	$\Psi = \kappa\psi^* \frac{1+a^2}{a}$
$\Theta = \frac{\kappa\nu}{g\epsilon H^3}\theta^*$	$\Theta = \frac{\kappa\nu}{g\epsilon H^3}\theta^* \frac{[\pi(1+a^2)]^3}{a^2}$

In laboratory experiments, the fluid height falls between 2 mm and 1 cm (Chandrasekhar 1961 [11], Ch. 2, sect. 18). Let us assume $H = 2$ mm. Then, the relative values of Saltzman's and Lorenz's nondimensional variables follow:

- Lorenz's nondimensional t^* is roughly 15 times greater than Saltzman's nondimensional t^*,
- Lorenz's nondimensional x^* and z^* are equal to Saltzman's nondimensional x^* and z^*,
- Lorenz's nondimensional ψ^* is roughly 2 times less than Saltzman's nondimensional ψ^*, and
- Lorenz's nondimensional θ^* is roughly 200 times less than Saltzman's nondimensional θ^*

Thus, for example, if the nonchaotic convective regime comes to equilibrium at Lorenz's $t^* = 15$, it comes to equilibrium at Saltzman's $t^* = 1$. Similarly, a temperature departure of $\theta^* = 200$ in Saltzman's case is associated with Lorenz's value of $\theta^* = 1$.

The equations governing the evolution of forecast sensitivities are found by taking partial derivatives of S-LOM (3) and L-LOM (3) with respect to λ and σ separately. Thus, six equations are added to each set, S-LOM (3) and S-LOM (3).

We demonstrate the process of determining evolution of forecast sensitivities by deriving the equation governing for $\frac{\partial x_4}{\partial \lambda}$:

$$\begin{aligned} \dot{x_4} = \frac{dx_4}{dt} &= -x_1 x_7 + \lambda x_1 - x_4 \\ \frac{d}{dt}\left(\frac{\partial x_4}{\partial \lambda}\right) &= -x_1 \frac{\partial x_7}{\partial \lambda} - x_7 \frac{\partial x_1}{\partial \lambda} + x_1 + \lambda \frac{\partial x_1}{\partial \lambda} - \frac{\partial x_4}{\partial \lambda} \end{aligned} \tag{15}$$

Similar ordinary differential equations are derived for the other forecast sensitivities. The initial conditions for the six sensitivity equations are zero, i.e., change in controls λ and σ affect the sensitivity solutions for $t > 0$, but not at $t = 0$. The 3-variable amplitude solutions feed into the forecast sensitivity equations as shown in Equation (15), but the forecast sensitivity equations do not feedback into the equations governing x_1, x_4, and x_7. The set of nine differential equations are solved simultaneously as a coupled set.

3. Dynamics of 3-Mode Systems: S-LOM (3) and L-LOM (3)

3.1. Overview

The 3-mode truncated systems that describe Rayleigh–Bénard convection have similar strength (they capture the interaction between temperature and vorticity that produce the physically meaningful convective motions) and similar weakness (inability to depict turbulent cellular motion with limited wave structures in the presence of large Rayleigh numbers). However, it is the difference in scaling and nondimensional forms that highlights the dynamical difference in the equation sets. It would be interesting to ask Lorenz why he refused to follow Saltzman's scaling and associated coefficients for the differential equations plainly set forth in Saltzman's Table 2 [1] One can guess that he wanted differential equations where the coefficients exhibited comparable magnitudes that lead to ease in understanding interaction between terms in the equation set. However, beyond

this difference, the model forecast sensitivities to controls are drastically different between the two sets—relatively large model forecast sensitivity to controls in S-LOM (3)'s case compared to L-LOM (3)'s case. These differences lead to poor data assimilation results when slight changes in ideal observation times are used in the S-LOM (3) experiments. On the other hand, L-LOM (3) tolerated these slight changes very well. Results of these experiments are discussed in Section 5.1 (L-LOM (3) dynamics) and Section 5.2 (S-LOM (3) dynamics). All calculations in Sections 3.2 and 3.3 use L-LOM (3).

3.2. Nonchaotic Regime: $\lambda = 12$, $\sigma = 7$ ($\sigma = \frac{\nu}{\kappa} = 7.03$ *for Water at* 20 °C)

When $\lambda = 12$, the Rayleigh number R is given by

$$R = \lambda R_c = 12\ (657.5) = 8.106{\cdot}10^3 \tag{16}$$

With this value of R and $H = 2$ mm for water at 20 °C, Rayleigh's [3] formula for the temperature difference ΔT between the upper and lower boundaries of the fluid is given by

$$\Delta T = \frac{\kappa \nu R}{g \epsilon H^3} = 7.11\ ^\circ\mathrm{C} \tag{17}$$

In the first three panels of Figure 1, amplitudes are depicted in pairs between $t = 0$ and $t = 10$ (5 min dimensionally): $x_1 x_4$ (first panel, top), $x_4 x_7$(second panel), and $x_1 x_7$ (third panel). In each of these panels, amplitudes are oscillatory and exhibit convergence to $x_1 \cong 5.50$, $x_4 = 5.75$, and $x_7 = 12.0$. In the bottom panel of Figure 1, the 3D trajectory of the amplitudes in phase space—(λ, σ) space—is shown over the period $t = 0$ to $t = 10$. This plot indicates convergence of L-LOM (3) to the x_1, x_4, x_7 values stated above.

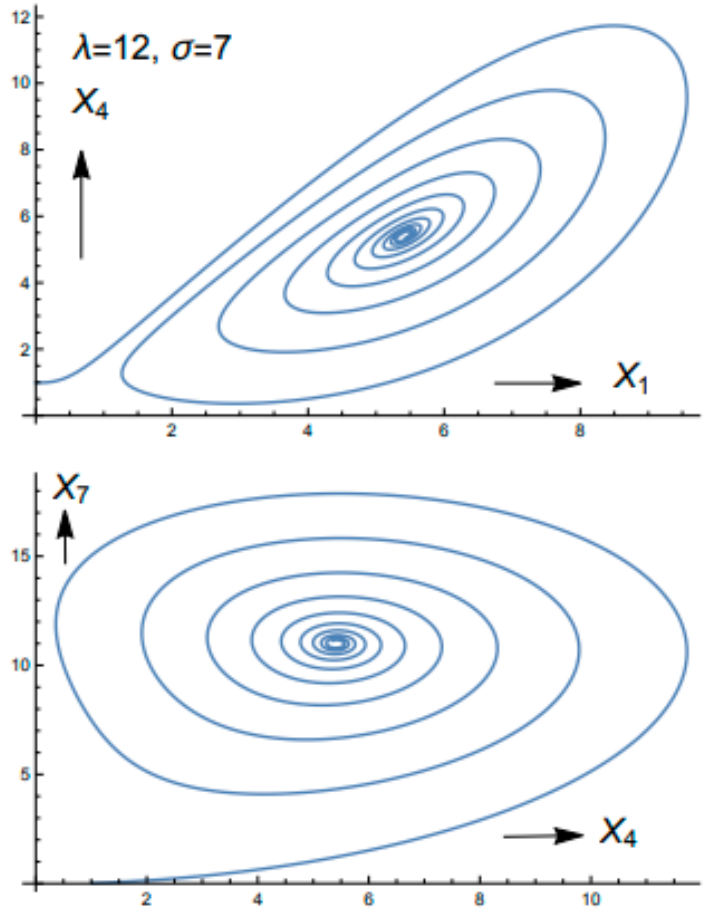


Figure 1. *Cont.*

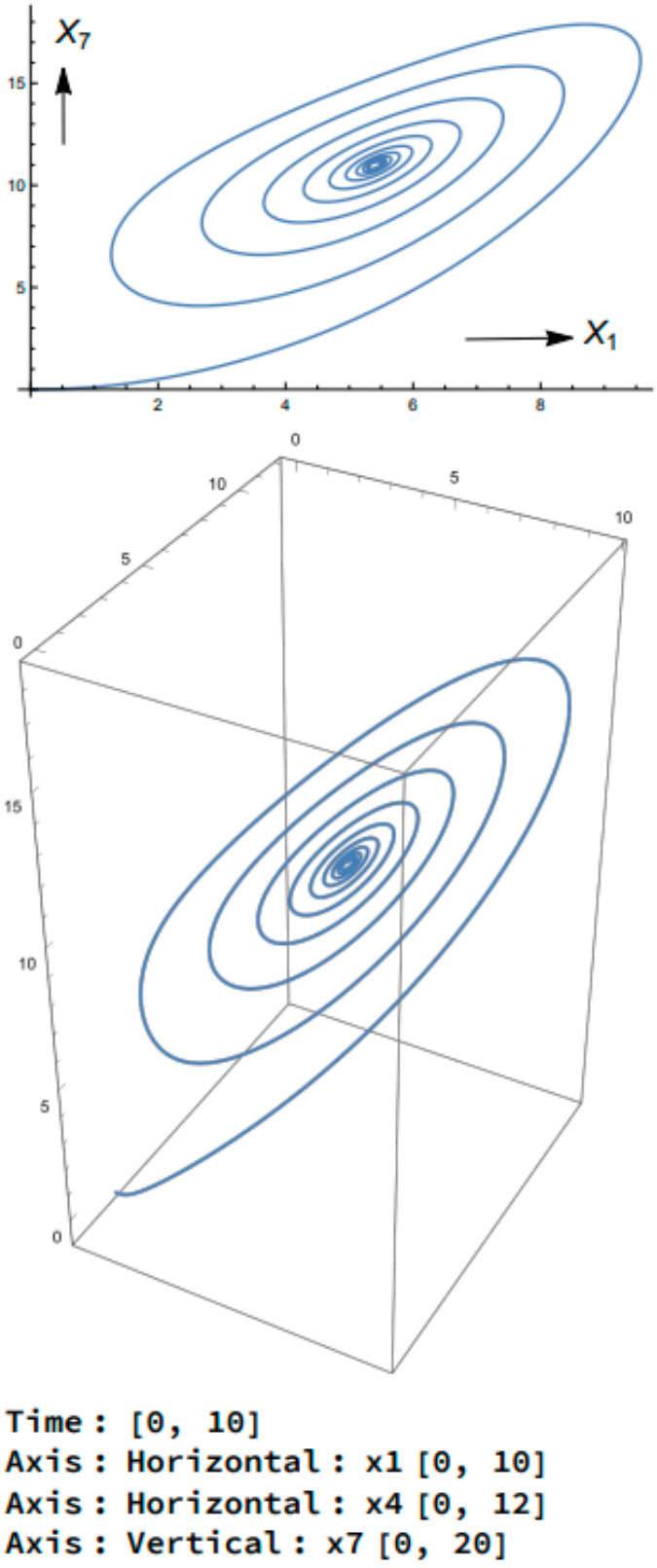


Figure 1. Phase diagrams for the nonchaotic regime with controls $\lambda = 12$, $\sigma = 7$ over time span [0, 10]: Panel 1 $(x_1,\ x_4)$, Panel 2 $(x_4,\ x_7)$, Panel 3 (x_1, x_7), and Panel 4 (x_1, x_4, x_7) with axes: horizontal $x_1[0,\ 10]$, $x_4[0,\ 12]$, $x_7[0,\ 20]$.

The forecast sensitivities of x_4 and x_7 with respect to λ and σ are shown in Figure 2. These two amplitude components are used to represent the temperature departure, and since only temperature departure observations will be used in the data assimilation experiments, we confine our forecast sensitivity analysis to these two components. These sensitivities exhibit oscillatory decay as steady state is approached.

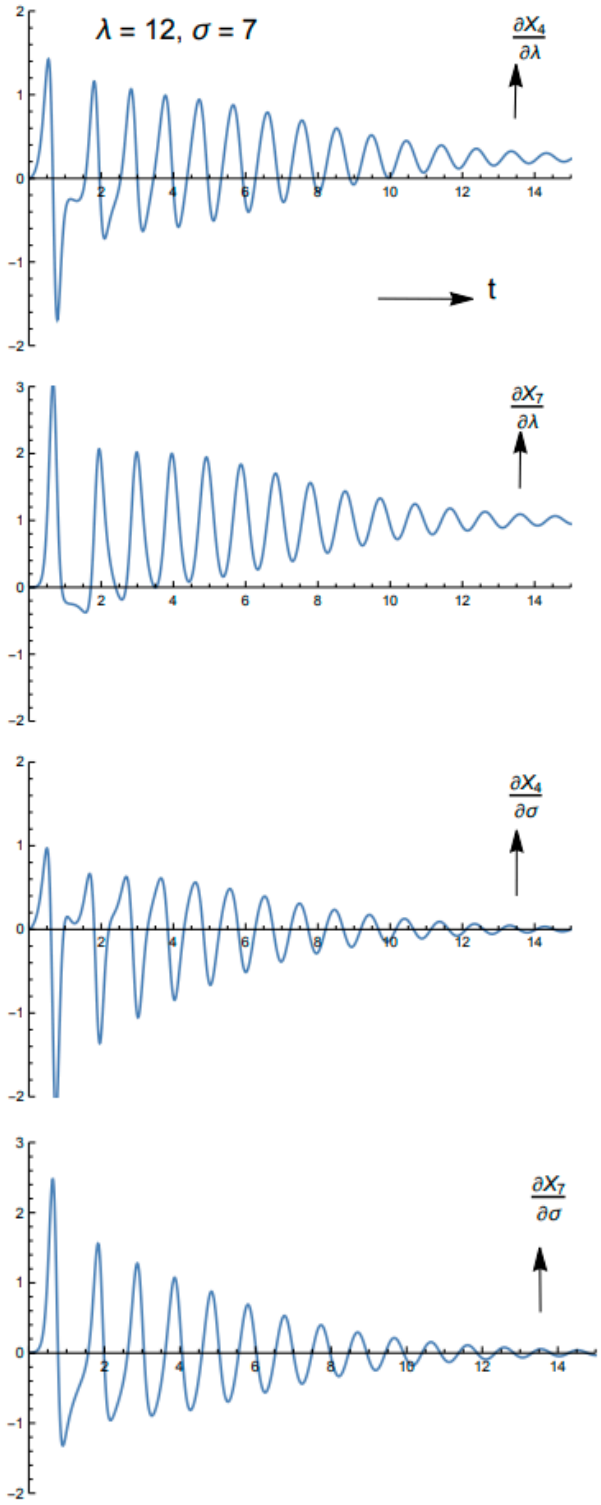
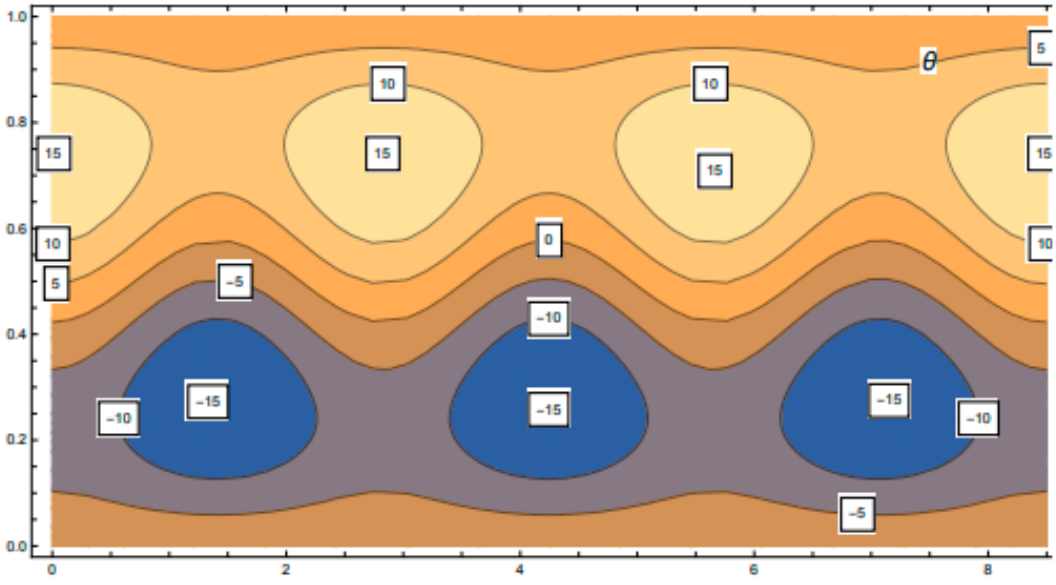


Figure 2. Forecast sensitivities of x_4 and x_7 with respect to (λ, σ) over time span [0, 14] for the non-chaotic regime with controls $\lambda = 12$, $\sigma = 7$.

The structure of the convection at $t = 10$ is shown in Figure 3—two panels, the temperature departure and streamfunction. The maximum/minimum values of temperature departure are $\sim \pm 15$ (± 3 °C dimensionally for water at 20 °C) where the maximum values are in the top portion of the water and the minimum values in the bottom portion. This leads to a spurious stable stratification first noted by Kuo [14] when convection is controlled by a truncated low-order spectral system. Within the cells, the maximum ($\frac{\partial \psi}{\partial x} > 0$) and minimum ($\frac{\partial \psi}{\partial x} < 0$) upward/downward vertical velocities, respectively, are the order of 1 mm·s^{-1} so roughly 50 circuits of a fluid particle around the cell takes place between the onset of convective motion ($t = 0$) and steady state ($t = 10$).



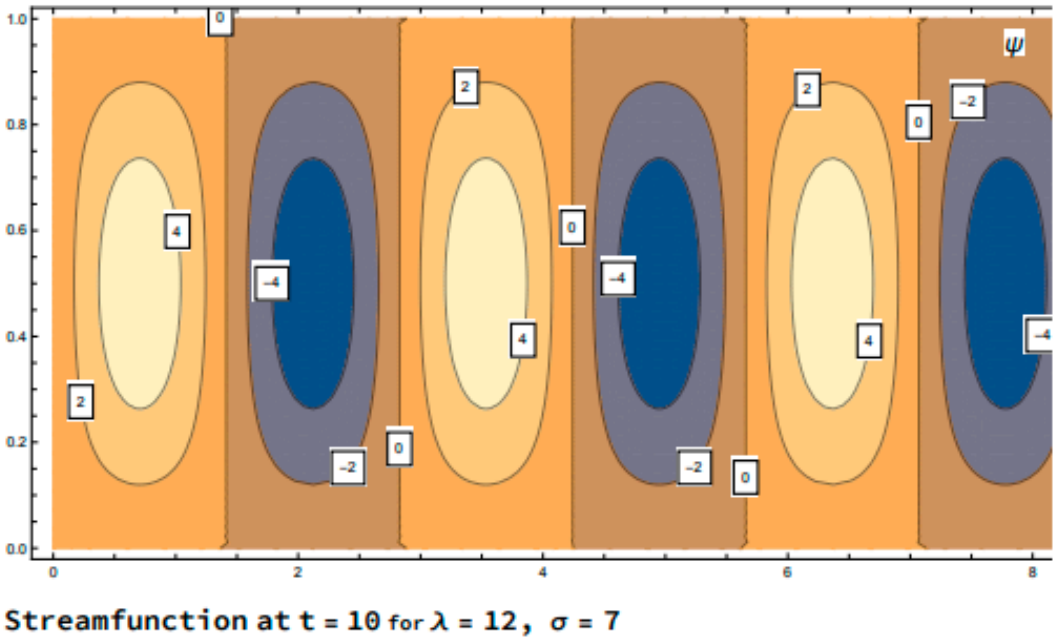


Figure 3. Display of convection pattern in the $x - z$ plane at time $= 10$ for the nonchaotic regime with controls $\lambda = 12$, $\sigma = 7$.

3.3. Chaotic Regime: $\lambda = 28$, $\sigma = 10$

The Prandtl number $\sigma = 10$ was assumed by Saltzman [1] in nonchaotic regimes and by Lorenz [4] in chaotic regimes. With these nondimensional numbers, Lorenz [4] found that instability of steady convection occurred when $\lambda = 24.74$ and beyond. The slightly supercritical value of $\lambda = 28$ was used by Lorenz in his numerical experiments [4]. The states of steady convection are then represented by amplitude points $(x_1,\ x_4,\ x_7)$: $(6\sqrt{2},\ 6\sqrt{2},\ 27)$ and $(-6\sqrt{2},\ -6\sqrt{2},\ 27)$, while the state of no convection corresponds to the origin in phase space (0, 0, 0).

In the first three panels of Figure 4, amplitudes are depicted in pairs between $t = 0$ and $t = 30$: x_1x_4, x_4x_7, and x_1x_7. The well-known butterfly pattern appears for each pair over the time period $t = 0$ to $t = 30$. The convergence to $(x_1,\ x_4,\ x_7)$: $(6\sqrt{2},\ 6\sqrt{2},\ 27)$ and $(-6\sqrt{2},\ -6\sqrt{2},\ 27)$ does not occur in finite time, only as $t \longrightarrow \infty$. The 3-D phase plot of amplitudes is shown in the lower right corner of Figure 4 where the two steady-state solutions are located at the centroids of the circular/elliptical trajectories.

The time evolution of sensitivities $\frac{\partial x_7}{\partial \lambda}$, $\frac{\partial x_7}{\partial \sigma}$ is shown in Figure 5. Here, we have broken the plots into two time zones: t [0, 5] and t [5, 10] for $\frac{\partial x_7}{\partial \lambda}$ and t [0, 3] and t [3, 10] for $\frac{\partial x_7}{\partial \sigma}$. Note that the coordinate axes are scaled differently in these plots. Plots for $\frac{\partial x_4}{\partial \lambda}$, $\frac{\partial x_4}{\partial \sigma}$ exhibit similar structures. The common thread between sensitivity plots is that the magnitudes oscillate between + and − values and increase ad infinitum.

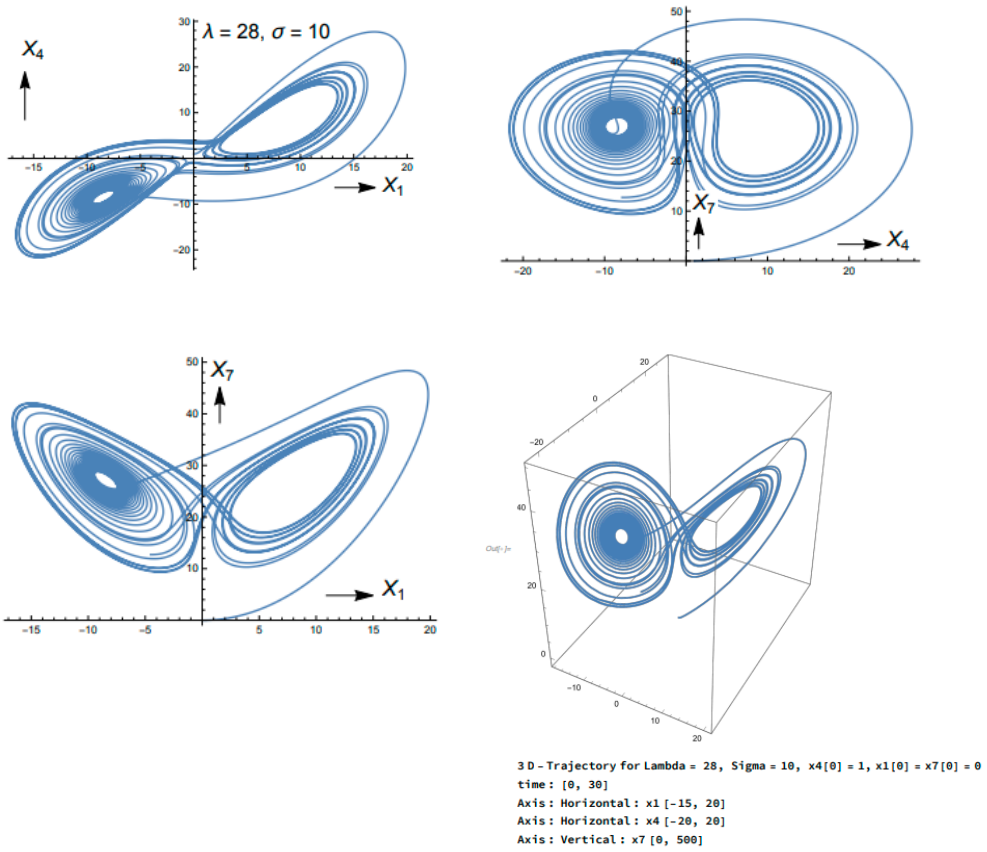


Figure 4. Phase plots of amplitudes for the chaotic regime with controls $\lambda = 28$, $\sigma = 10$.

The sensitivity of the $x_1(t)$ solution to a small change in the initial condition is shown in Figure 6. In this figure, a 5% change in the initial condition leads to significant change in $x_1(t)$ after t∼20.

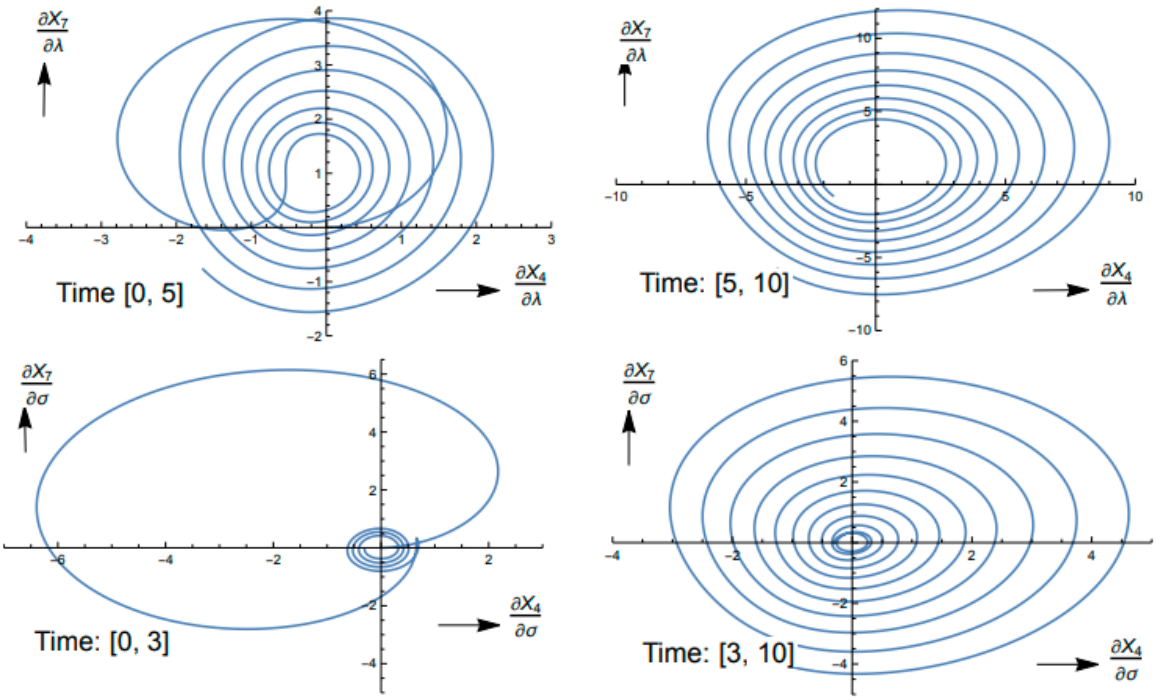


Figure 5. Phase plots of x_4 and x_7 forecast sensitivity over time span [0, 10] in the chaotic regime with controls $\lambda = 28$, $\sigma = 10$.

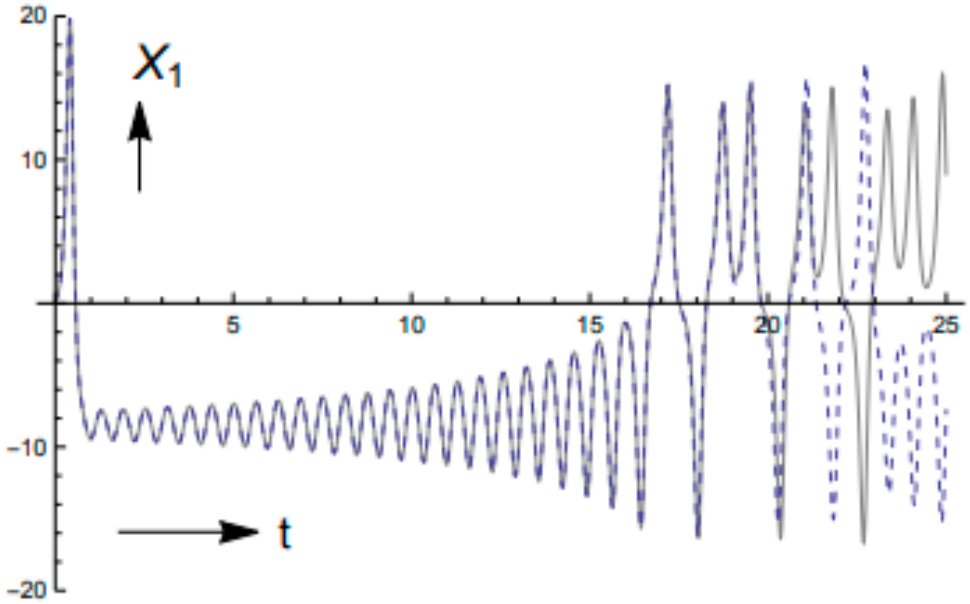


Figure 6. Extreme sensitivity of streamfunction amplitude x_1 to slight change in $x_4(0)$ from 1.00 to 1.05 for the chaotic regime—solid line ($x_4(0) = 1.00$) and dashed line ($x_4(0) = 1.05$).

4. Design of Observation Network

4.1. Model and Forecast Sensitivities

Let $\boldsymbol{X}(k) = (x_1(k),\ x_4(k),\ x_7(k))\ \epsilon R^3$ be the state of the S-LOM (3) or L-LOM (3) model at times $k = 0,\ 1,\ 2,\ 3\ldots$ given by

$$\boldsymbol{X}(k+1) = M(\boldsymbol{X}(k), \alpha)\ \epsilon\ R^3 \tag{18}$$

with $\boldsymbol{X}(0)\ \epsilon R^3$ as initial condition, $\alpha = (\alpha_1,\ \alpha_2) = (\lambda,\ \sigma)^T \epsilon\ R^2$, the parameter vector, and $\boldsymbol{X}(k) = \boldsymbol{X}(k, X(0), \alpha)$ the solution.

Define forecast sensitivity to parameters as

$$\begin{array}{rcl} \boldsymbol{V}(k) & = & \frac{\partial \boldsymbol{X}(k)}{\partial \alpha}\ \epsilon\ R^{3x2} \\ & = & \begin{bmatrix} \frac{\partial x_1}{\partial \lambda} & \frac{\partial x_1}{\partial \sigma} \\ \frac{\partial x_4}{\partial \lambda} & \frac{\partial x_4}{\partial \sigma} \\ \frac{\partial x_7}{\partial \lambda} & \frac{\partial x_7}{\partial \sigma} \end{bmatrix} \end{array} \tag{19}$$

It can be shown that $\boldsymbol{V}(k)$ evolves according to the discrete time dynamics given by

$$\begin{array}{c} \boldsymbol{V}(k+1) = D_M(k)\boldsymbol{V}(k) + D_M^{\alpha}(k) \\ with\ \boldsymbol{V}(0) = 0 \end{array} \tag{20}$$

where

$$D_M(k) = \frac{\partial}{\partial x_j(0)}[M_i(\boldsymbol{X}(k),\ \alpha)]\ \epsilon\ R^{3x3} \tag{21}$$

$$D_M^{\alpha}(k) = \frac{\partial}{\partial \alpha_j}[M_i(\boldsymbol{X}(k),\ \alpha)]\ \epsilon\ R^{3x2} \tag{22}$$

with $\alpha = (\alpha_1,\ \alpha_2) = (\lambda\ ,\ \sigma\ )$ [15,16].

An example of forecast sensitivity to parameter λ has been shown earlier in Equation (15).

4.2. Forecast Error

Let $\overline{c} = \overline{\alpha}$ be the true control and $\overline{\boldsymbol{X}}(k)$ be the solution starting from $\overline{c}$. Let $c = \alpha$ be the arbitrary control and let $\boldsymbol{X}(k)$ be the model solution starting from c.

Define

$$\delta c = \overline{c} - c = \delta\alpha = \overline{\alpha} - \alpha \tag{23}$$

where $\overline{\alpha} - \alpha = (\overline{\lambda} - \lambda,\ \overline{\sigma} - \sigma)^T = \begin{bmatrix} \delta\lambda \\ \delta\sigma \end{bmatrix} = \delta\alpha$.

Let $\delta\boldsymbol{X}(k)$ be the change in the state $\boldsymbol{X}(k)$ resulting from change δc. Then, from Equation (19) and using first-order Taylor expansion, we obtain

$$\begin{array}{rcl} \delta\boldsymbol{X}(k) & = & \overline{\boldsymbol{X}}(k) - \boldsymbol{X}(k) \\ & = & \boldsymbol{X}(k,\ \boldsymbol{X}(0),\ \overline{\alpha}) - \boldsymbol{X}(k,\ \boldsymbol{X}(0), \alpha) \\ & = & \frac{\partial \boldsymbol{X}(k)}{\partial \alpha}\delta\alpha \\ & = & \boldsymbol{V}(k)\delta\alpha \end{array} \tag{24}$$

4.3. Observations and Cost Function

We work with L-LOM (3) in this subsection.

Consider the physical domain defined by the rectangle

$$D = \left\{\eta = (x,\ z)^T\right\}, \quad 0 \le x \le 6\sqrt{2}, \quad 0 \le z \le 1 \tag{25}$$

where x denotes the lengthwise coordinate and z the breadthwise coordinate of a point η in D. Convection is confined to this domain.

Define a matrix

$$H(\eta) = [f_4(\eta),\ f_7(\eta)]\epsilon\ R^{1x2} \tag{26}$$

where

$$\begin{array}{l} f_4(\eta) = \cos(\pi a x)\sin(\pi z) \\ f_7(\eta) = -\sin(2\pi z) \end{array} \tag{27}$$

Let

$$\overline{\theta}(k,\eta) = H(\eta)\begin{bmatrix} \overline{x}_4 \\ \overline{x}_7 \end{bmatrix} = H(\eta)\hat{\overline{X}} + \xi(k) \tag{28}$$

be the actual observation where $\hat{\overline{X}}$ is the projection of $\overline{X}$ onto its last two elements and $\xi(k)$ is the observation noise with variance σ^2.

Let

$$\theta^M(k,\ \eta) = H(\eta)\begin{bmatrix} x_4 \\ x_7 \end{bmatrix} = H(\eta)\hat{X} \tag{29}$$

where $\theta^M(k,\eta)$ is the model counterpart to the observation based on perturbed control.

The difference is called the forecast error or innovation (*innov*).

$$e(k) = \overline{\theta}(k,\eta) - \theta^M(k,\ \eta) \tag{30}$$

Using Equations (28) and (29), we rewrite Equation (30) as

$$\begin{aligned} e(k) &= H(\eta)\left(\hat{\overline{X}} - \hat{X}\right) \\ &= H(\eta)\delta\hat{X} \end{aligned} \tag{31}$$

For $\alpha\ \epsilon\ R^2$ and $\eta\ \epsilon\ R^2$, we define the cost function for a given $(k,\ \eta)$ as

$$J(\alpha, k,\ \eta) = \frac{1}{2\sigma^2}\left(\overline{\theta}(k,\eta) - \theta^M(k,\ \eta)\right)^2 \tag{32}$$

The squared difference between the observation $\overline{\theta}(k,\eta)$ and its model counterpart $\theta^M(k,\ \eta)$. This cost function in Equation (32) is additive in the number of observations. Using Equations (31) and (32), we get

$$J(\alpha,\ \eta) = \frac{1}{2\sigma^2}(H(\eta)(\delta\hat{X})^T H(\eta)(\delta\hat{X}) = \frac{1}{2\sigma^2}(\delta\hat{X})^T H^T(\eta)H(\eta)\ (\delta\hat{X}) \tag{33}$$

Using first-order Taylor expansion, the vector of temperature departure error components, $\begin{bmatrix} \delta x_4 \\ \delta x_7 \end{bmatrix}$, can be represented by

$$\delta\hat{X} = \begin{bmatrix} \delta x_4 \\ \delta x_7 \end{bmatrix} = \begin{bmatrix} \frac{\partial x_4(k)}{\partial\lambda} & \frac{\partial x_4(k)}{\partial\sigma} \\ \frac{\partial x_7(k)}{\partial\lambda} & \frac{\partial x_7(k)}{\partial\sigma} \end{bmatrix}\begin{bmatrix} \delta\lambda \\ \delta\sigma \end{bmatrix} = \hat{V}(k)\delta\alpha, \tag{34}$$

where $\hat{V}\ (k)$ is a submatrix of the forecast sensitivity matrix $V(k)$; namely, the last two rows of that matrix. Arguments related to the derivation of Equation (34) are found in Lakshmivarahan et al. [7].

Substituting Equation (34) into Equation (33), we get

$$J(\alpha,\ \eta) = \frac{1}{2\sigma^2}(\delta\alpha)^T\left[\hat{V}^T(k)\ \left(\ H^T(\eta)H(\eta)\right)\hat{V}(k)\right]\delta\alpha = \frac{1}{2\sigma^2}(\delta\alpha)^T G\ (\delta\alpha) = \frac{1}{2\sigma^2}(\delta\alpha)^T G\ (\delta\alpha) \tag{35}$$

where

$$G(k) \equiv \hat{V}^T(k)\ \left(\ H^T(\eta)H(\eta)\right)\hat{V}(k)] = \hat{V}^T(k)\overline{H}(\eta)\hat{V}(k)\Big] \tag{36}$$

and

$$\overline{H}(\eta) \equiv H^T(\eta)H(\eta) \tag{37}$$

Remark 1. *It is important to note that $G(k)$ applies to a single time k. Since the cost function is additive in the number of observations, i.e., if there is more than one observation, a separate matrix exists for each time. Take the case when observations are available at three observation times with G matrices G_1, G_2, and G_3, then the appropriate combined matrix will be $G = G_1 + G_2 + G_3$.*

In the presence of two forms of the cost function, namely Equation (32) and (35), we use Equation (35) to determine observation placement. We use Equation (32) to display the structure of the cost function in the space of controls.

From Equation (35), the gradient of $J(\alpha,\ \eta)$, $\nabla J(\alpha,\eta)$, is given by

$$\nabla J(\alpha,\eta) = \begin{bmatrix} \frac{\partial J}{\partial \lambda} \\ \frac{\partial J}{\partial \sigma} \end{bmatrix} = G\ (\delta\alpha) \tag{38}$$

To determine the optimal adjustment to control through FSM, we write the cost function as

$$J(\alpha,\eta) = \frac{1}{2\sigma^2}(H\hat{X} + \frac{\partial H(\eta)\hat{X}}{\partial \alpha}\delta\alpha - \overline{\theta})^T\ (H\hat{X} + \frac{\partial H(\eta)\hat{X}}{\partial \alpha}\delta\alpha - \overline{\theta}) \tag{39}$$

where

$$\frac{\partial H(\eta)\hat{X}}{\partial \alpha} = H(\eta)\ \hat{V} \tag{40}$$

Expanding Equation (40) and setting its gradient to zero yields

$$\delta\alpha = (G^T G)^{-1} G^T innov \tag{41}$$

where the innovation *innov* is given by

$$innov = H\hat{X} - \overline{\theta} \tag{42}$$

4.4. Observation Placement

In this paper we consider two structures for the observations: (1) time only, and (2) time and space. Our strategy for observation placement is to render G positive definite. However, not only render it positive definite, but also reduce its condition number so that a "bowl-like" cost function's structure is present in the vicinity of its true minimum located at $\overline{\alpha} = (\overline{\lambda},\ \overline{\sigma})$ for error-free observations. Two types of observations are considered: (a) observations in time alone (Section 4.4.1) and (b) observations in space and time (Section 4.4.2).

4.4.1. Observations in Time Alone

When we consider observations of spectral components $\overline{X}(k)$ as a function of time alone in our data assimilation problem, the $H\ (\eta)$ matrices in Equation (36) are replaced by identity matrices and the cost function then becomes

$$J(\alpha) = \frac{1}{2\sigma^2}(\delta\alpha)^T\left[\hat{V}^T(k)\ \hat{V}(k)\right]\delta\alpha \tag{43}$$

where $G = \hat{V}^T(k)\hat{V}(k)$ is positive definite as shown below. In matrix form

$$\hat{V}^T(k)\hat{V}(k) = \begin{bmatrix} \left(\frac{\partial x_4(k)}{\partial \lambda}\right)^2 + \left(\frac{\partial x_7(k)}{\partial \lambda}\right)^2 & \left(\frac{\partial x_4(k)}{\partial \lambda}\frac{\partial x_4(k)}{\partial \sigma} + \frac{\partial x_7(k)}{\partial \lambda}\frac{\partial x_7(k)}{\partial \sigma}\right) \\ \left(\frac{\partial x_4(k)}{\partial \lambda}\frac{\partial x_4(k)}{\partial \sigma} + \frac{\partial x_7(k)}{\partial \lambda}\frac{\partial x_7(k)}{\partial \sigma}\right) & \left(\frac{\partial x_4(k)}{\partial \sigma}\right)^2 + \left(\frac{\partial x_7(k)}{\partial \sigma}\right)^2 \end{bmatrix} \tag{44}$$

or alternatively represented by

$$\hat{V}(k) = (\varepsilon_1(k),\ \varepsilon_2(k)) \qquad \hat{V}^T = \begin{bmatrix} \varepsilon_1^T(k) \\ \varepsilon_2^T(k) \end{bmatrix} \tag{45}$$

$$\hat{V}^T(k)\hat{V}(k) = \begin{bmatrix} \varepsilon_1^T(k)\varepsilon_1(k) & \varepsilon_1^T(k)\varepsilon_2(k) \\ \varepsilon_2^T(k)\varepsilon_1(k) & \varepsilon_2^T(k)\varepsilon_2(k) \end{bmatrix} \tag{46}$$

From this point onwards, we simplify notation by omitting the parenthetical expression "(k)" following forecast sensitivities.

From the physics of our problem, $V(k)$ is sensitive to λ and σ. The dot products "$\langle a,\ b\rangle$" are expressed in matrix-vector notation where $\varepsilon_1(k)$ and $\varepsilon_2(k)$ are the first and second columns of $\hat{V}(k)$.

$$\langle \varepsilon_1,\ \varepsilon_2\rangle = \varepsilon_1^T\varepsilon_2 = \varepsilon_2^T\varepsilon_1 = \langle \varepsilon_2,\ \varepsilon_1\rangle \tag{47}$$

$$\left\langle \varepsilon_1^T,\ \varepsilon_1 \right\rangle = \| \varepsilon_1 \|^2,\ \left\langle \varepsilon_2^T,\ \varepsilon_2 \right\rangle = \| \varepsilon_2 \|^2$$

Thus,

$$\hat{V}^T(k)\hat{V}(k) = \begin{bmatrix} \| \varepsilon_1 \|^2 & \langle \varepsilon_1,\ \varepsilon_2 \rangle \\ \langle \varepsilon_2,\ \varepsilon_1 \rangle & \| \varepsilon_2 \|^2 \end{bmatrix} \tag{48}$$

Further, $\| \varepsilon_1 \|^2 > 0$ and $\| \varepsilon_2 \|^2 > 0$, and the determinant of $\hat{V}^T(k)\hat{V}(k)$ is given by

$$\det\left(\hat{V}^T\hat{V}\right) = \| \varepsilon_1 \|^2 \| \varepsilon_2 \|^2 - < \varepsilon_1,\ \varepsilon_2 >^2 \tag{49}$$

and

$$\langle \varepsilon_1,\ \varepsilon_2 \rangle = \| \varepsilon_1 \| \| \varepsilon_2 \| \cos\theta \tag{50}$$

so

$$< \varepsilon_1,\ \varepsilon_2 >^2 = \| \varepsilon_1 \|^2 \| \varepsilon_2 \|^2\ \cos^2\theta \tag{51}$$

The determinant of $\hat{V}^T\hat{V}$ can then be written as

$$\det\left(\hat{V}^T\hat{V}\right) = \| \varepsilon_1 \|^2 \| \varepsilon_2 \|^2 - \| \varepsilon_1 \|^2 \| \varepsilon_2 \|^2\ \cos^2(\theta) = \| \varepsilon_1 \|^2 \| \varepsilon_2 \|^2 \left(1 - \cos^2\theta\right) > 0\ (if\ \theta \neq 0) \tag{52}$$

So long as ε_1 and ε_2 are not colinear or $\theta = 0$ or 180°, then

$$\cos\theta \neq 0,\ \text{and} \det\left(\hat{V}^T\hat{V}\right) > 0 \tag{53}$$

and the following inequalities hold

$$\| \varepsilon_1 \|^2 > 0,\ \| \varepsilon_2 \|^2 > 0,\ \text{and}\ \det\left(\hat{V}^T\hat{V}\right) > 0 \tag{54}$$

These inequalities imply that $\hat{V}^T\hat{V}$ is a positive definite matrix.

The strategy for choosing observations is therefore:

Pick observation $z(k) = \hat{X}(k) + v(k)$ when $z(k)\ \epsilon\ R^2$ and noise $v(k)\ \epsilon\ R^2$ such that $z(k_1)$ at time k_1 makes $\| \varepsilon_1 \|^2$ a maximum and $z(k_2)$ makes $\| \varepsilon_2 \|^2$ a maximum.

Then, at

$$k_1 : V(k_1) = [\varepsilon_1(k_1),\ \varepsilon_2(k_1)]$$

and at

$$k_2 : V(k_2) = [\varepsilon_1(k_2),\ \varepsilon_2(k_2)]$$

and

$$G = V^T(k_1)V(k_1) + V^T(k_2)V(k_2) = G(k_1) + G(k_2) \tag{55}$$

4.4.2. Observations Taken in Space and Time

When observation locations are to be determined in space and time, the G matrix takes the form

$$G(k,\eta) = \hat{V}^T(k)\left(H^T(\eta)H(\eta)\right)\hat{V}(k)] = \hat{V}^T(k)\overline{H}(\eta)\hat{V}(k)\Big]. \tag{56}$$

The space element $\overline{H}(\eta)$ is separable from the time elements $\hat{V}^T(k)$ and $\hat{V}^T(k)$ or in combination $\hat{V}^T(k)\ \hat{V}(k)$. In the presence of this separability, the strategy reduces to first determining desired times as explained above. This is followed by determining desired positions in $(x,\ z)$ space.

Recall that $\hat{x} = (x_4,\ x_7)^T$ and $\alpha = (\lambda,\ \sigma)^T$ and represent $\hat{V}$ as a column partition

$$\hat{V} = \begin{bmatrix} \frac{\partial x_4}{\partial \lambda} & \frac{\partial x_4}{\partial \sigma} \\ \frac{\partial x_7}{\partial \lambda} & \frac{\partial x_7}{\partial \sigma} \end{bmatrix} = \left[\frac{\partial \hat{x}}{\partial \lambda},\ \frac{\partial \hat{x}}{\partial \sigma}\right] = [V_1,\ V_2] \tag{57}$$

Define vectors

$$g_4 = \left[\frac{\partial x_4}{\partial \alpha}\right] = \left(\frac{\partial x_4}{\partial \lambda},\ \frac{\partial x_4}{\partial \sigma}\right)^T \epsilon\ R^2 \tag{58}$$

and

$$g_7 = \left[\frac{\partial x_7}{\partial \alpha}\right] = \left(\frac{\partial x_7}{\partial \lambda}, \frac{\partial x_7}{\partial \sigma}\right)^T \epsilon R^2 \tag{59}$$

It can be verified that the row partition of $\hat{V}$ is given by

$$\hat{V} = \begin{bmatrix} g_4^T \\ g_7^T \end{bmatrix} \tag{60}$$

Recall that H is given by

$$H = [f_4, f_7] \epsilon R^{1x2} \tag{61}$$

where $f_4 = \cos(\pi a x)\sin(\pi z)$ and $f_7 = -\sin(2\pi z)$.

Using Equations (60) and (61), we obtain a row vector $w(k, \eta) = (w_1(k, \eta), w_2(k, \eta))$ given by

$$w(k, \eta) = H\hat{V} = [f_4, f_7]\begin{bmatrix} g_4^T \\ g_7^T \end{bmatrix} = (f_4 g_4^T + f_7 g_7^T) \epsilon R^{1x2} \tag{62}$$

Consequently, we can express the Gramian $G(k, \eta)$ as a rank-one symmetric product matrix

$$G(k, \eta) = w^T(k, \eta) w(k, \eta) = \begin{bmatrix} w_1^2(k, \eta) & w_1(k, \eta) w_2(k, \eta) \\ w_2(k, \eta) w_1(k, \eta) & w_2^2(k, \eta) \end{bmatrix} \tag{63}$$

Strategy for determining observation placement:

(1) Pick $\eta(x, z)$ such that $\| f_4 \|$ and $\| f_7 \|$ are maximum

(2) Let k_1, k_2, k_3, k_4 be four time instances where the squares of $\frac{\partial x_4}{\partial \lambda}$, $\frac{\partial x_4}{\partial \sigma}$, $\frac{\partial x_7}{\partial \lambda}$, $\frac{\partial x_7}{\partial \sigma}$ are maximum. Determine the respective $G(k_i, \eta)$'s. Then, the final Gramian is given by

$$G = G(k_1, \eta) + G(k_2, \eta) + G(k_3, \eta) + G(k_4, \eta) \tag{64}$$

Remark 2. *In principle we can use just two observations at times k_1 and k_2 where $\left(\frac{\partial x_4}{\partial \lambda}\right)^2$ and $\left(\frac{\partial x_7}{\partial \sigma}\right)^2$ attain their maxima.*

5. Data Assimilation Experiments

Once the observations and their locations have been determined, the FSM method of data assimilation is used to determine optimal values of controls. Development of the FSM was first shown in Lakshmivarahan and Lewis [15] and explored with numerous examples in the textbook Lakshmivarahan et al. [16]. In our case, the optimal values of control are found through the iterative methodology captured by Equation (41). Starting from guess control, solution to Equation (41) delivers the adjustment to guess control. The adjustment is added to the guess and the second operating point in space of control is determined. New forecasts and new forecast sensitivities are found from the updated controls. The process continues until the cost function's change falls within a specified tolerance.

5.1. Experiment I: Nonchaotic Data Assimilation Process (Lorenz Scaling)

True controls are taken to be $\lambda = 12.0$, $\sigma = 7.0$. The forecasts of x_1, x_4, x_7 based on guess control ($\lambda = 12.6$, $\sigma = 4.0$) are shown in Figure 7. The amplitudes reach steady state near t = 10. Graphs of the diagonal elements of $V^T V$ that dictate observations to be made at t = 1 are shown in Figure 8. The observations at this time are found by adding $\approx 10\%$ random error to the forecast based on true controls. The observations of x_4 (Obx_4) and x_7 (Obx_7) at t = 1 are given by

$$(Obx_4, Obx_7) = (0.371, 11.378). \tag{65}$$

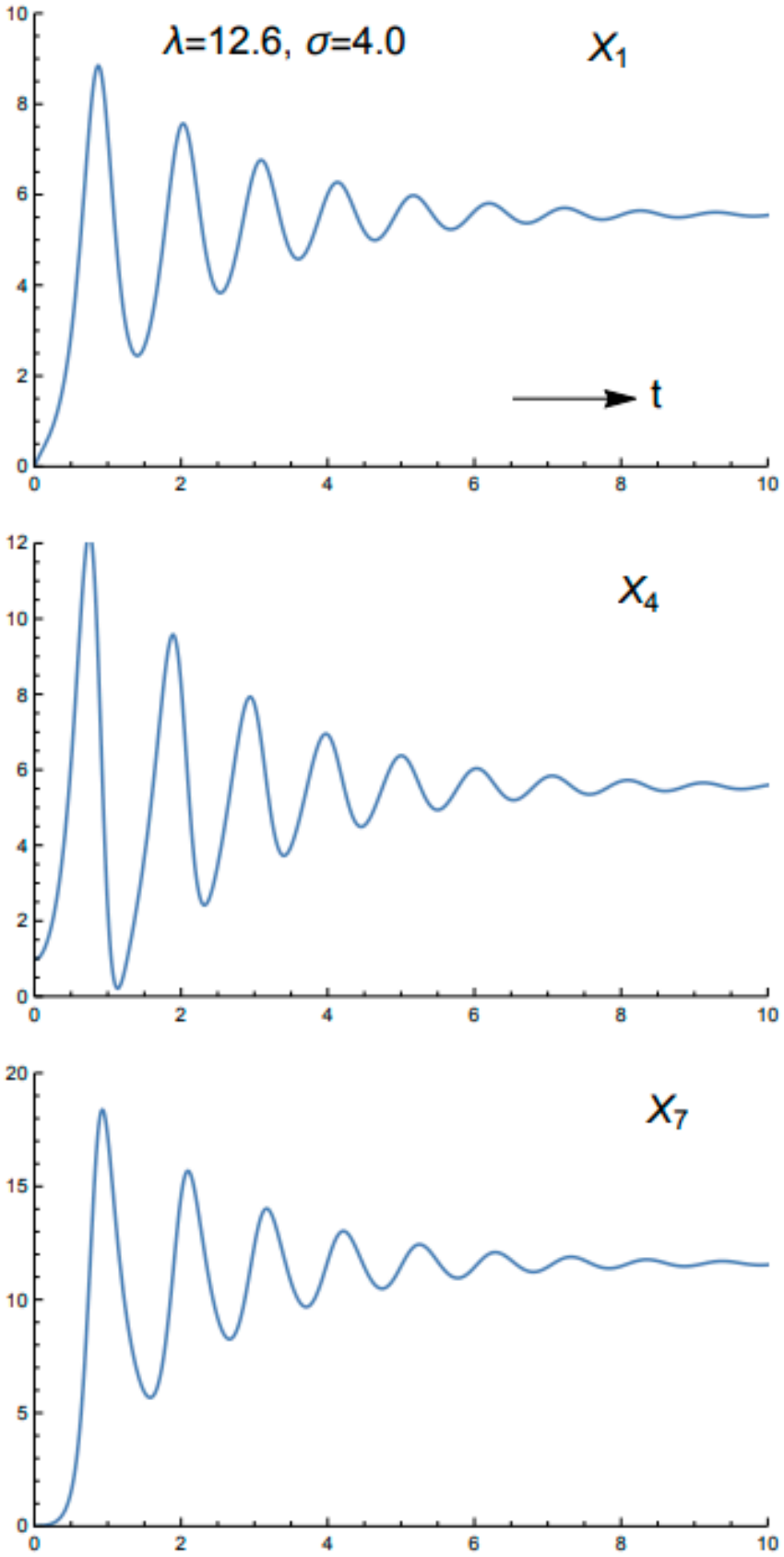


Figure 7. Forecast of the three amplitudes x_1, x_4, x_7 for the nonchaotic regime using L-LOM (3) over the time span [0, 10] with guess controls $\lambda = 12.6$, $\sigma = 4.0$.

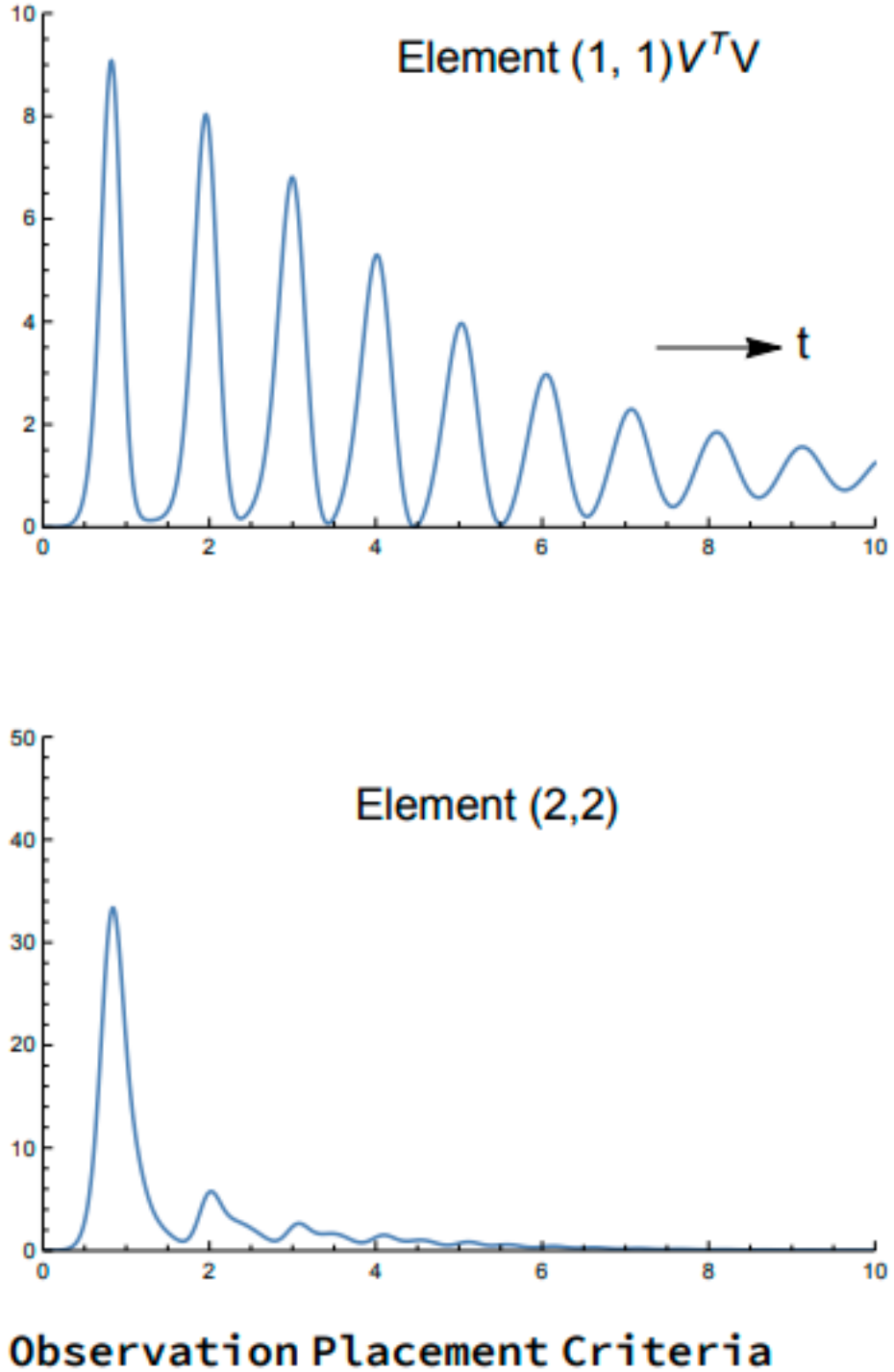


Figure 8. Plots of the diagonal elements of the product sensitivity matrix $V^T V$—elements (1,1) and (2,2)—for the nonchaotic regime using L-LOM (3) when guess controls are given by $\lambda = 12.6$, $\sigma = 4.0$.

Table 2 summarizes results from FSM. The assimilation process converges to a solution in 4 steps. The last column of Table 2 displays the sequential reduction in cost function values from ~19 to $3{\cdot}10^{-5}$, reduction by a factor of 10^{-6}. FSM determines values of $\Delta\lambda$ and $\Delta\sigma$ at each of the 4 iterations. The plot of these values is shown in Figure 9. The contours are always elliptical since the FSM cost function at each iteration is quadradic in $\Delta\lambda$ and $\Delta\sigma$. Display of cost function Equation (32) is shown in Figure 10 and the minimum is found to be located at $\lambda = 12$, $\sigma = 7$.

In a test that changed observation time from t_{obs} = 1.00 to t = 1.05, the data assimilation algorithm converged to $(\lambda,\ \sigma) = (11.995,\ 6.968)$ in 4 iterations.

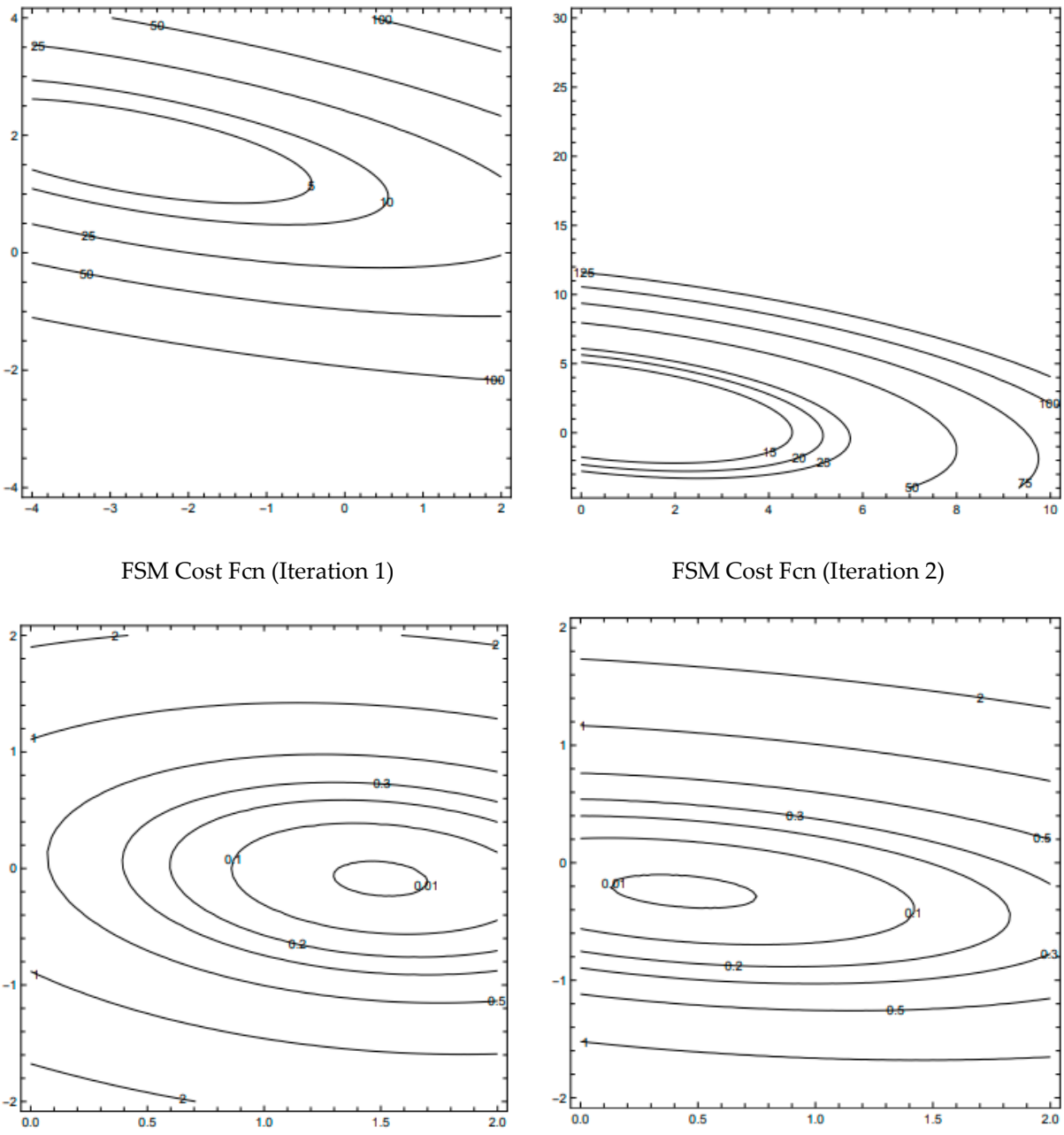


Figure 9. Contour plots of the FSM cost function in the space of control increments ($\delta\lambda$ along the horizontal axis and $\delta\sigma$ along the vertical axis). The values of $\delta\lambda$ and $\delta\sigma$ at the cost function minimum are the incremental adjustments to controls at the operating point.

Table 2. FSM Iterates for λ.

i	$\Delta\lambda^{(i)}$	$\Delta\sigma^{(i)}$	$\lambda^{(i)}$	$\sigma^{(i)}$	$Inov\ (x_4)$	$Inov\ (x_7)$	Cost Fcn
0	—	—	12.600	4.000	−1.987	−5.831	18.957
1	−2.787	1.735	9.813	5.735	−2.245	−1.456	3.580
2	0.252	1.579	10.065	7.309	−1.048	0.113	0.555
3	1.499	−0.087	11.559	7.213	−0.210	0.172	0.037
4	0.439	−0.243	11.997	6.969	−0.002	−0.005	$3{\cdot}10^{-5}$

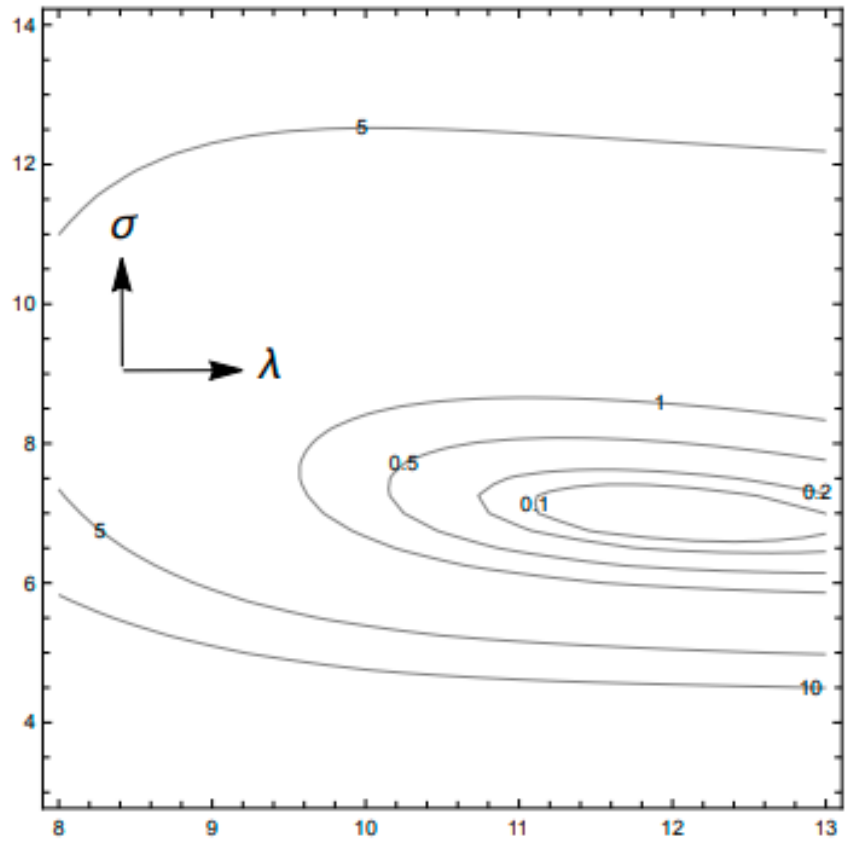


Figure 10. Plot of the cost function given by Equation (32). The minimum is located at $\lambda = 12.0$, $\sigma = 7.0$ (λ the horizontal axis and σ the vertical axis).

5.2. Experiment II: Nonchaotic Data Assimilation Process (Saltzman Scaling)

This data assimilation experiment follows the exact steps for the nonchaotic data assimilation process as shown in Section 5.1 except that the governing 3-mode equations are scaled according to Saltzman's form of the equations as found in Equations (3)–(5).

Figure 11 displays the time-dependent solutions to the 3-mode model. The solid curves are derived from the true controls ($\lambda = 12,\ \sigma = 7$) and the dashed curves from forecast controls ($12.6,\ 4.0$). The primary difference between this forecast and the one shown earlier for Lorenz's scaling is the much greater magnitudes of the variables and the shorter time required to reach equilibrium. Otherwise, the time-dependent oscillatory nature of the components is the same. The forecast sensitivity of x_4 and x_7 to λ and σ are shown in Figure 12—again the sensitivities are much greater than those associated with Lorenz's scaling. Based on the elements of the sensitivity product matrix V^TV, $t_{obs} = 0.16$. The observations at this time are found by adding $\approx 10\%$ random error to the forecast based on true controls. The observations of x_4 (Obx_4) and x_7 (Obx_7) at t = 0.16 are given by

$$(Obx_4,\ Obx_7) = (427.5 - 260.6) \tag{66}$$

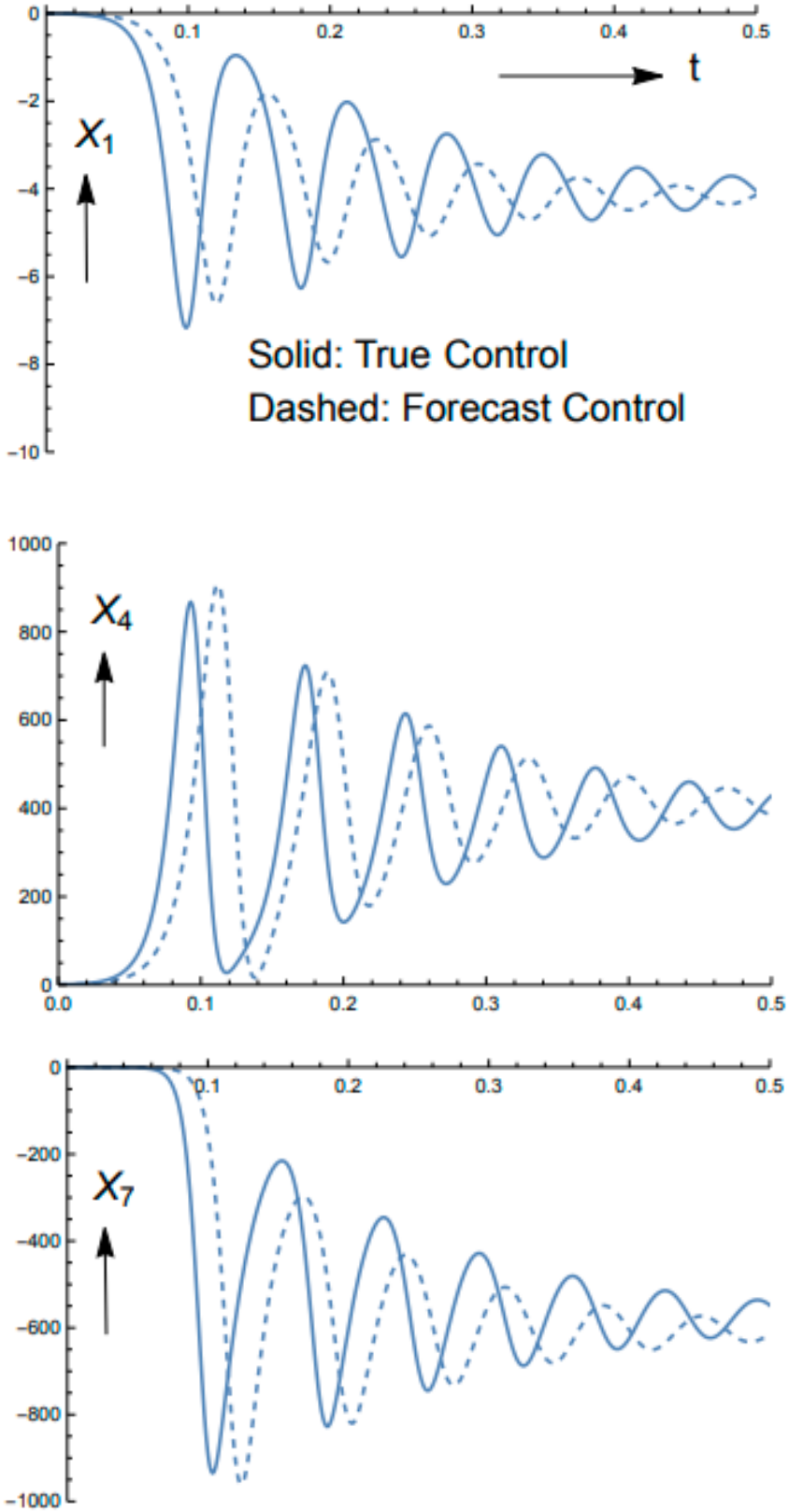


Figure 11. Non-chaotic regime solutions $x_1,\ x_4,\ x_7$ for S-LOM (3) using true control (solid curves: $\lambda = 12.0,\ \sigma = 7.0$) and perturbed/forecast control (dashed curves: $\lambda = 12.6,\ \sigma = 4.00$).

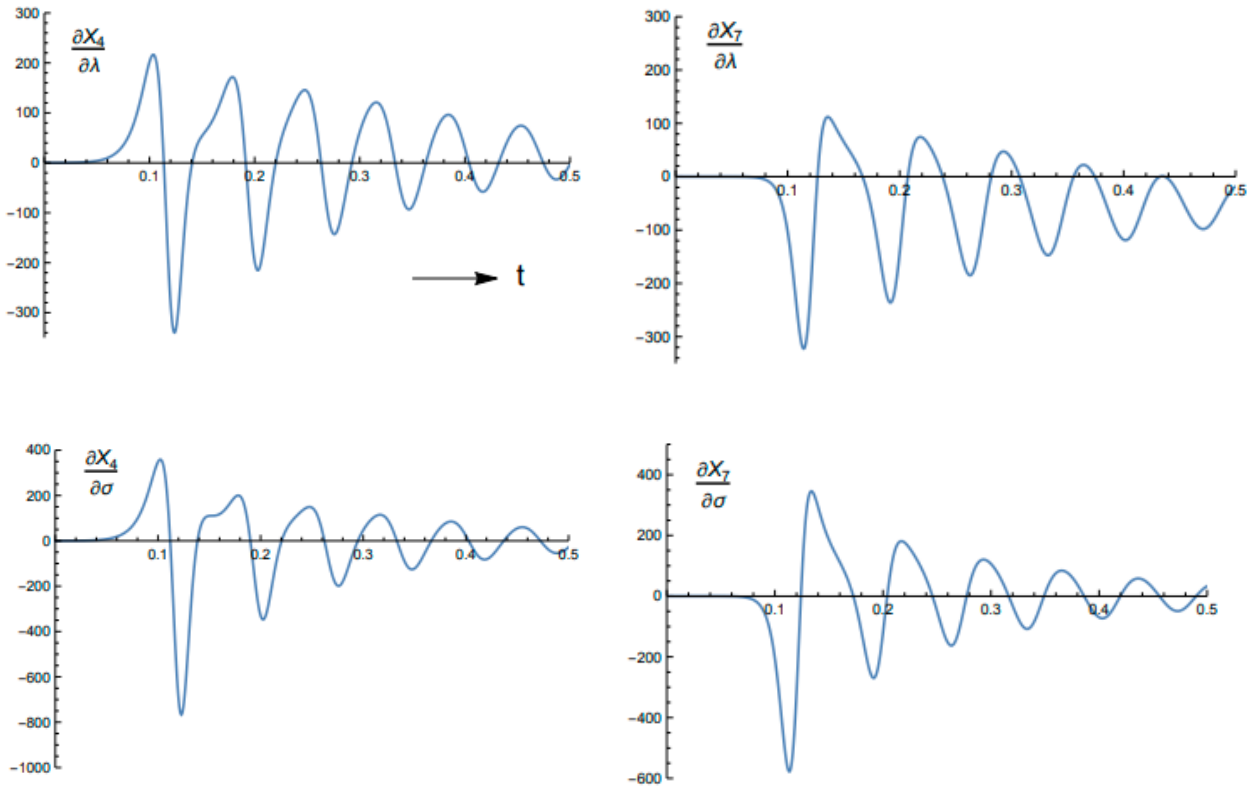


Figure 12. S-LOM (3) forecast sensitivity of x_4 and x_7 to controls for the non-chaotic regime.

Results from the data assimilation experiment are shown in Table 3. The data assimilation process converges to $(\lambda,\ \sigma) = (13.8,\ 4.8)$ in 4 iterations. The cost function is reduced 9 orders of magnitude. The cost function for this case is shown in Figure 13. The converged values of control shown in Table 3 are consistent with this cost function's minimum.

In a test that changed observation time from $t_{obs} = 0.16$ to $t = 0.10$, FSM did not converge to an optimal state. Instead, after a single iteration the result was $(\lambda,\ \sigma) = (-46.3,\ 26.8)$, unreasonable physically.

Table 3. FSM Iterates for $\lambda = 12$, Saltzman Scaling, Observations taken at $t = 0.16$.

i	$\Delta\lambda^{(i)}$	$\Delta\sigma^{(i)}$	$\lambda^{(i)}$	$\sigma^{(i)}$	*Inov* (x_4)	*Inov* (x_7)	Cost Fcn
0	—	—	12.600	4.000	191.857	63.905	$2{\cdot}10^{4}$
1	2.397	−0.102	15.000	3.900	−52.480	58.896	$3{\cdot}10^{3}$
2	−1.054	0.638	13.943	4.536	4.627	10.944	$7{\cdot}10^{1}$
3	−0.135	0.218	13.807	4.754	0.750	0.742	$6{\cdot}10^{-1}$
4	−0.006	0.018	13.801	4.771	0.006	0.004	$3{\cdot}10^{-5}$

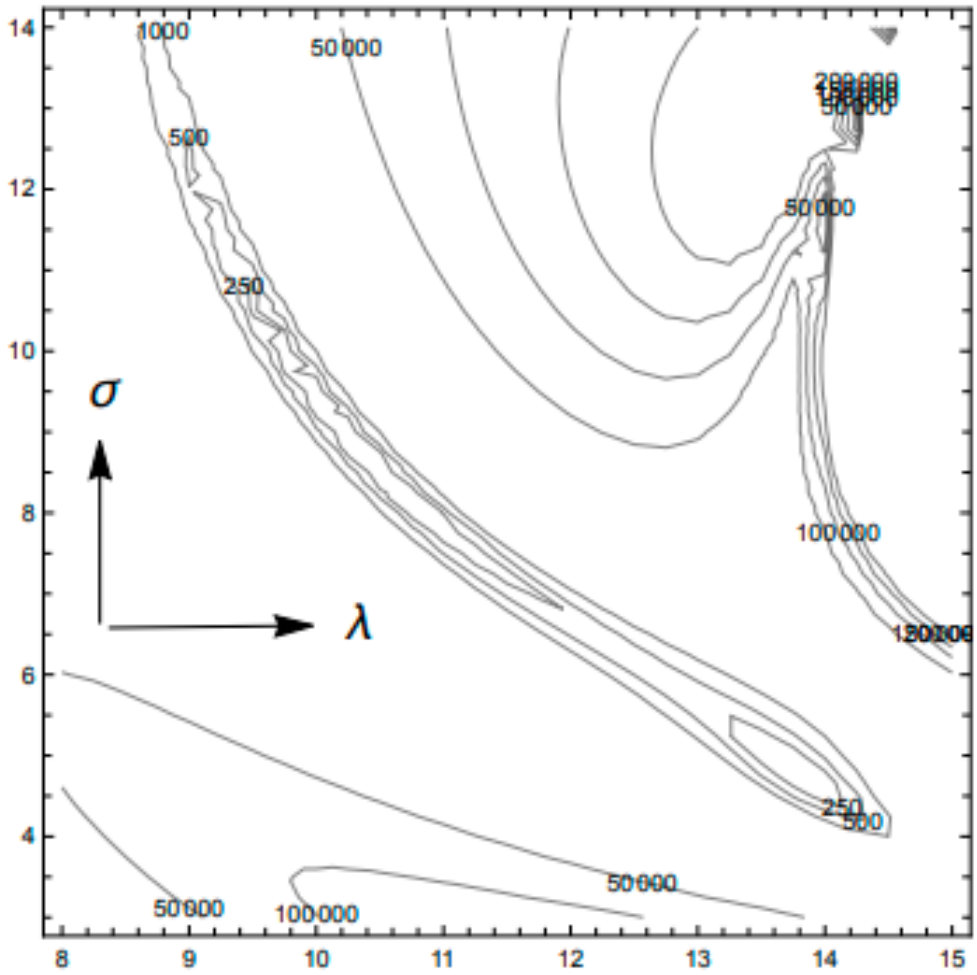


Figure 13. Plot of the cost function based on Equation (32) for the nonchaotic regime using S-LOM (3). The minimum is located at $\lambda = 13.8$, $\sigma = 4.8$ (λ the horizontal axis and σ the vertical axis).

5.3. Experiment III: Chaotic Data Assimilation Process (Lorenz Scaling)

True controls are taken to be $\lambda = 28.0$, $\sigma = 10.0$. The forecasts of x_1, x_4, x_7 based on guess control ($\lambda = 29.0$, $\sigma = 8.0$) are shown in Figure 14. Based on the elements of the product sensitivity matrix $V^T \cdot V$, observations are collected at $t = 0.4$. The observations at this time are found by adding $\approx 10\%$ random error to the forecast based on true controls. The observations of x_4 (Obx_4) and x_7 (Obx_7) at t = 0.4 are given by

$$(Obx_4,\ Obx_7) = (22.832,\ 40.194). \tag{67}$$

Table 4 summarizes results from FSM. Assimilation process converges to a solution in 4 steps. The last column of the table displays the sequential reduction in cost function values from ~74 to ~10^{-4}, reduction by a factor of 10^6. The cost function is shown in Figure 15 and the minimum is found to be located at $\lambda = 28$, $\sigma = 10$, nearly identical to the solution found by FSM.

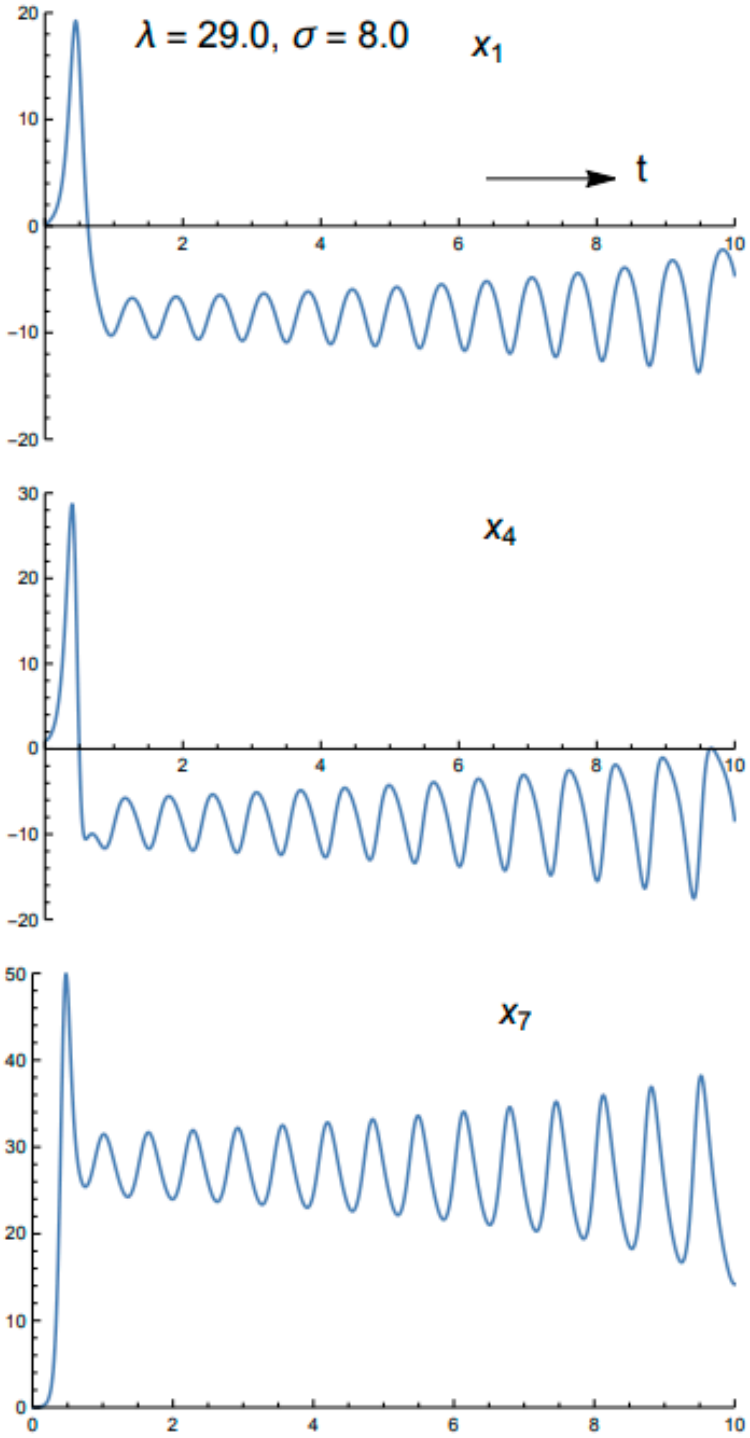


Figure 14. Solution to the amplitude equations for the chaotic regime using L-LOM (3) with guess control given by $\lambda = 29.0$, $\sigma = 8.0$.

Table 4. FSM Iterates for λ = 28.

i	$\Delta\lambda^{(i)}$	$\Delta\sigma^{(i)}$	$\lambda^{(i)}$	$\sigma^{(i)}$	*In* (x_4)	*In* (x_7)	Cost Fcn
0	—	—	29.000	8.000	−5.732	10.712	73.801
1	−4.419	3.130	24.581	11.130	0.773	7.839	31.024
2	3.003	−0.815	27.584	10.315	0.294	0.258	$8{\cdot}10^{-2}$
3	0.251	−0.136	27.835	10.178	−0.008	0.011	$9{\cdot}10^{-5}$
4	0.000	0.002	27.835	10.180	$3{\cdot}10^{-6}$	$3{\cdot}10^{-6}$	$9{\cdot}\ 10^{-12}$

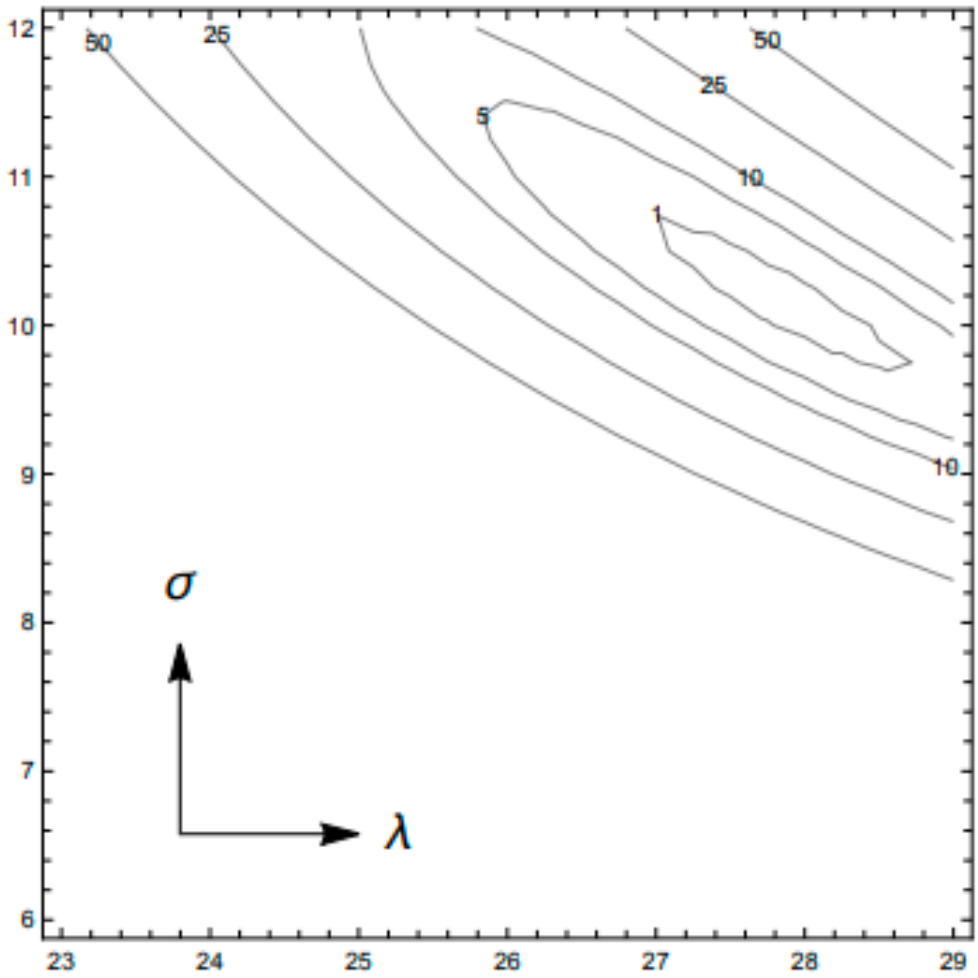


Figure 15. Plot of the cost function based on Equation (32) for the chaotic regime using L-LOM (3). The minimum is located at $\lambda = 27.6$, $\sigma = 10.3$ (λ the horizontal axis and σ the vertical axis).

It is interesting to consider cases where the optimal observation time was greater than t = 0.4, times where the model forecast sensitivities are extremely large as shown earlier in Figure 5. There is no reason to believe that observations at these times would not give a reasonable result. The reasonableness of the result, however, would be dependent on the smallness of the condition number.

5.4. Experiment IV: Data Assimilation in Time and Space (x, z)

We confine our attention to a single numerical experiment that combines space and time. Our experiment focuses on true controls $\lambda = 2$ and $\sigma = 7$, a nonchaotic regime. The perturbed controls are $\lambda = 2.5$ and $\sigma = 3.0$. The time of observation is determined from the structure of the sensitivity product matrix V^TV. Based on the trace of V^TV, observations

are taken at t_{obs} = 2. The spatial observation sites are determined from the trace of $H^T H$ at t = 2. These spatial locations are

$$\begin{array}{l} x_1 = 0.10 \;\; z_1 = 0.25 \\ x_2 = 0.10 \;\; z_2 = 0.75 \\ x_3 = 2.55 \;\; z_3 = 0.23 \\ x_4 = 3.11 \;\; z_4 = 0.68 \end{array} \tag{68}$$

In this case, the condition number for the G matrix is 3.8 which builds convexity into the cost function as shown in Figure 16.

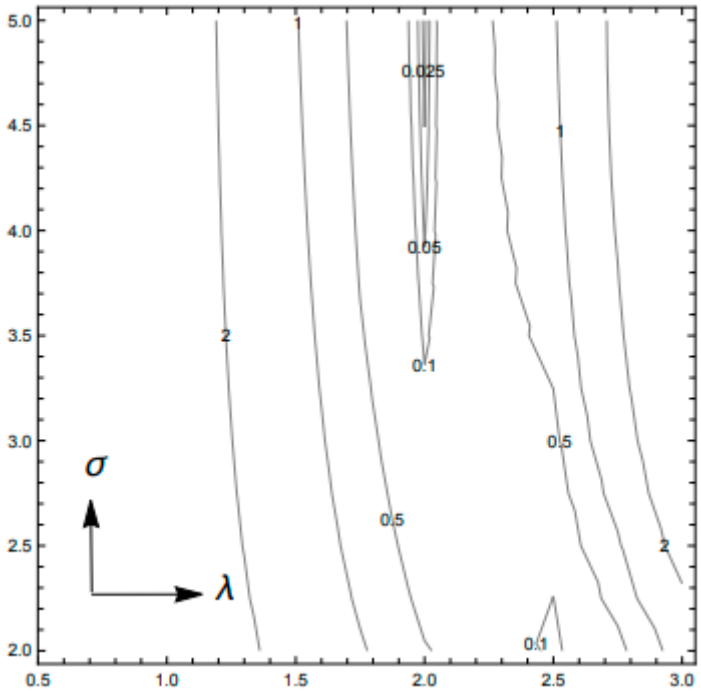


Figure 16. Plot of the cost function based on Equation (32) for the nonchaotic regime using L-LOM (3) with observations in space-time. The minimum is located at $\lambda = 2.0$, $\sigma = 4.4$ (λ the horizontal axis and σ the vertical axis).

Table 5 summarizes to convergence to optimal solution. The innovations at the observation locations exhibit a reduction of 10^7 in the 3 iterations and the optimal minimum is very close to the minimum shown in Figure 16.

Table 5. FSM Iterates for the space-time constraint.

i	$\Delta\lambda^{(i)}$	$\Delta\sigma^{(i)}$	$\lambda^{(i)}$	$\sigma^{(i)}$	*Innov* (x_1,z_1)	*Innov* (x_2,z_2)	*Innov* (x_3,z_3)	*Innov* (x_4,z_4)
0	—	—	2.500	3.000	0.076	−0.638	0.138	−0.599
1	−0.412	0.824	2.087	3.824	−0.032	0.056	−0.034	0.051
2	0.004	0.334	2.091	4.158	−0.002	0.006	−0.002	0.005
3	0.002	0.009	2.093	4.167	$5{\cdot}10^{-6}$	$7{\cdot}10^{-7}$	$4{\cdot}10^{-6}$	$9{\cdot}10^{-7}$

6. Discussion and Conclusions

S-LOM (3) and L-LOM (3), low-order forms of spectral equations that govern Rayleigh–Bénard convection developed by Saltzman [1] and used by Lorenz [4], respectively, have been used to study advantages of observation placement in dynamic data assimilation. Observation placement is controlled by forecast sensitivity to model control; in this case the control has two elements, Rayleigh and Prandtl numbers. We have used FSM data

assimilation since it delivers forecast sensitivities to control that are used to determine observation placement.

By examination of the cost function in terms of forecast sensitivities, a Gramian matrix G becomes the centerpiece to determine observation placement. For time-only observations, the Gramian matrix reduces to the product of $V^T V$ where V is the time-dependent forecast sensitivity to control. This matrix is symmetric positive definite and the maximum values of the trace or norm of this matrix determine the best observation locations. Observation locations that yield the matrix's smallest condition number build convexity into the cost function's minimum. For our case study, the contours of the cost function in the space of the two controls are ellipses. As the condition number approaches the number 1, the major and minor axes of the ellipses approach the same number and the ellipses approach circles. In this case, the path to the minimum is along a straight line. For time-space observations, the Gramian matrix is symmetric semi positive definite in the form $V^T H^T H V$ where H is the spatial-dependent matrix that accounts for the double Fourier series representation of the observations. This product exhibits separability in space and time and therefore permits determination of observation-time placement as a first step to be followed by spatial placement of observations.

We tested the observation placement strategy for both nonchaotic and chaotic regimes. Results from the numerical experiments with FSM were especially good based on (1) iterative reduction of the cost functions' values, and (2) comparison of the optimal parameters with the cost function minimum in the space of controls.

We also wanted to test the role of equation scaling in the data assimilation process. The two models have different scaling where S-LOM (3) follows the standard fluid dynamical scaling based on the fluid's physical parameters while L-LOM (3) incorporates various forms of the non-dimensional wavenumber along with the physical parameters in its scaling. The scaling for L-LOM (3) leads to coefficients of comparable magnitude in the governing amplitude differential equations while scaling for S-LOM (3) leads to order of magnitude differences in the coefficients. In turn, the forecast sensitivity equations for S-LOM (3) exhibit large differences in their coefficients. Nevertheless, the FSM data assimilation method delivered good results for both models. On the other hand, when the ideal observation time is changed slightly, S-LOM (3) results were poor while L-LOM (3) were excellent. It thus appears that scaling plays an important role in data assimilation, at least for our limited number of experiments.

Author Contributions: Formal analysis, J.M.L. and S.L.; Writing—original draft, J.M.L. and S.L.; Writing—review & editing, J.M.L. and S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: We thank Bo-Wen Shen for encouraging us to explore application of FSM to chaotic systems. The authors also thank: (1) Kevin Fisher-Mack and Junjun Hu for writing the computer program that delivered the cost function display in the space of controls, and (2) the three anonymous reviewers who made cogent suggestions that were followed and this led to clarification of discussions in the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Saltzman, B. Finite Amplitude Free Convection as an Initial Value Problem-I. *J. Atmos. Sci.* **1962**, *19*, 329–341. [CrossRef]
2. Bénard, M. Les Tourbillions cellulaires dans une nape liquide. *Rev. Générale Des Sci. Pures Et Appl.* **1990**, *11*, 1261–1271, 1309–1328.
3. Rayleigh, L. Convection currents in a horizontal layer of fluid, when higher temperature is on the underside. *Philos. Mag.* **1916**, *32*, 529–546. [CrossRef]
4. Lorenz, E.N. Deterministic Nonperiodic Flow. *J. Atmos. Sci.* **1963**, *20*, 220–241. [CrossRef]

5. Saravanan, R. *The Climate Demon: Past, Present, and Future of Climate Change*; Cambridge University Press: Cambridge, UK, 2022; 379p.
6. Lewis, J.M.; Lakshmivarahan, S.; Dhall, S.K. *Dynamic Data Assimilation: A Least Squares Approach*; Cambridge University Press: Cambridge, UK, 2006; 654p.
7. Lakshmivarahan, S.; Lewis, J.M.; Hu, J. On Controlling the Shape of the Cost Functional in Dynamic Data Assimilation: Guidelines for Placement of Observations and Application to Saltzman's Model of Convection. *J. Atmos. Sci.* **2020**, *77*, 2969–2989.
8. Lakshmivarahan, S.; Lewis, J.M.; Maryada, S.K.R. Observability Gramian and its Role in the Placement of Observations in Dynamic Data Assimilation. In *Data Assimilation for Atmospheric, Oceanic and Hydrographic Applications*; Park, S.K., Xu, L., Eds.; Springer: Cham, Switzerland, 2022; Volume IV, pp. 215–257.
9. Lewis, J.M.; Lakshmivarahan, S.; Maryada, S.K.R. Placement of observations for variational data assimilation: Application to Burgers' Equation and seiche phenomenon. In *Data Assimilation for Atmospheric, Oceanic and Hydrographic Applications*; Park, S.K., Xu, L., Eds.; Springer: Cham, Switzerland, 2022; Volume IV, pp. 215–257.
10. Lakshmivarahan, S.; Lewis, J.M.; Hu, J. Saltzman's Model: Complete Characterization of Solution Properties. *J. Atmos. Sci.* **2019**, *76*, 1587–1608. [CrossRef]
11. Chandrasekhar, S. *Hydrodynamic and Hydromagnetic Stability*; Oxford at the Clarendon Press: Oxford, UK, 1961; 654p.
12. Turner, J.S. *Buoyancy Effects in Fluids*; Cambridge University Press: Cambridge, UK, 1973; 368p.
13. Lewis, J.M.; Lakshmivarahan, S. Derivation of Saltzman's spectral equations for Rayleigh-Bénard convection in terms of two variable parameters, Rayleigh and Prandtl numbers. *Tech. Rep.* 2019; *Unpublished*.
14. Kuo, H.-L. Solution of the nonlinear equations of cellular convection and heat transport. *J. Fluid Mech.* **1961**, *10*, 611–634. [CrossRef]
15. Lakshmivarahan, S.; Lewis, J.M. Forward Sensitivity Based Approach to Dynamic Data Assimilation. *Adv. Meteorol.* **2010**, *2010*, 375615. [CrossRef]
16. Lakshmivarahan, S.; Lewis, J.M.; Jabrzemski, R. *Forecast Error Correction Using Dynamic Data Assimilation*; Springer: Berlin/Heidelberg, Germany, 2017; 270p.