Climate system: A global sensitivity approach

Liban Ismail, Cédric Chauvière, Hacène Djellout

Climate system : A global sensitivity approach.

March 7, 2022

LIBAN ISMAIL[1] ; HACÈNE DJELLOUT[2]; CÉDRIC CHAUVIÈRE [3]

Laboratoire de Mathématiques Blaise Pascal (LMBP), CNRS UMR 6620, Université Clermont Auvergne, Campus Universitaire des Cézeaux, 3 place Varsarely, TSA 60026, CS 60026, 63178 Aubière Cedex France.

Abstract

This article is a first attempt to develop a numerical approach to solving differential equations based on Galerkin projections and extensions of polynomial chaos to analyze the sensitivity of input parameters in the Lorenz-Stenflo (LS) climate model [14] (chemical properties of the atmosphere, rotation, temperature gradient, convection motion). The sensitivity analysis was undertaken to identify the key parameters of the model.
Knowing that the climate system is chaotic, very sensitive to initial conditions, this article is also a first attempt to study the sensitivity of initial conditions using euclidean distance. In addition, we do simulations of the climate model in the non-chaotic case and in the chaotic case and we calculate the Sobol indices when the parameters follow the uniform law.

Keywords: Climate ; Chaotic system ; Sensitivity analysis ; Polynomials chaos ; Sobol Indice ; Lorenz-Stenflo ; Simulations.
2010 Mathematics Subject Classification:

1 Introduction to Chaos Theory

Lorenz is a pioneer of chaos theory. Long before him, several mathematicians and/or physicists such as Leonhard Euler, Pierre-Simon de Laplace, Alexis Claude Clairaut, then Henri Poincaré had foreseen that certain deterministic phenomena were sensitive to the initial conditions. This means that if you modify, even slightly, the initial conditions of the phenomenon, it can evolve in a radically different way. For example, if you throw two leaves side by side into a slightly choppy river, they will quickly follow along very different paths, even though they were originally very close. This discovery of sensitivity to the initial conditions was made by attempting to predict the

[1] liban_ismail_abdillahi@univ.edu.dj
[2] hacene.djellout@uca.fr
[3] cedric.chauviere@uca.fr

movements of the Moon, a celestial body subject to the pull of the Earth and the Sun (this problem is known as the three-body problem). But these precursors certainly lacked a valuable tool of data processing which became available only in the 1960s. Indeed, data processing allows to makes it possible to carry out the necessary calculations . Lorenz's [18] model had important repercussions in showing the possible limits on the ability to predict long-term climate and meteorological evolution. This model makes the starting point for the theory of dynamical systems serving as a source for new mathematical concepts.

The phenomena of chaos and chaotic systems have been studied by many researchers because of their various applications in the fields of atmospheric dynamics, population dynamics, electrical circuits, cryptology, fluid dynamics, lasers, engineering, stock exchanges, chemical reactions, etc. Most of the complex dynamic phenomena are characterized by chaotic and hyperchaotic systems of nonlinear ordinary differential equations [12, 15, 16, 18, 20, 36]. The LS model appears in several research works, in particular the articles [14, 22, 23, 26, 37]. Many dynamic behaviors such as the stability, bifurcation, periodic solutions and chaotic behaviors have been thoroughly studied for decades after Stenflo. The Lorenz system [18] has since become a paradigm in chaos theory. This describes the butterfly effect: a small disturbance, such as a butterfly plapping its wings, can lead to great changes is a chaotic system, e.g., a storm.

The sensitivity analysis is used with the chaotic model of Lorenz in several articles [1, 10, 13, 19]. An analysis based on the staggered methods is considered, in the article [1], to eliminate the least effective parameters. In [10], the author considered as an output function which corresponds to a ratio of temperature variations after a duration of 1 time unit. It also integrates the initial condition as uncertainty parameters. The article [13] addresses some fundamental methodological questions concerning the sensitivity analysis of chaotic geophysical systems. They show, using the Lorenz system [18] as an example, that an adjoint variational sensitivity analysis approach is of limited utility. Other papers [17, 27] have done sensitivity analysis by the chaos polynomial approach using the Lorenz system without ever calculating the Sobol indices. To our knowledge, this article is a first attempt to adapt the coefficients of polynomials chaos to calculate Sobol indices and illustrate the sensitivity of the parameters in a chaotic system.

The objective of this paper, firstly, is to study the sensitivity to initial conditions using the euclidean distance. Secondly, we use the polynomials chaos method to study and calculate the Sobol indices which give us the influence of several natural parameters (temperature, rotation, convection motion, fluid properties) that intervene in the (LS) climate systems and how they affect the evolution of the climate. The theoretical and computer aspects of the method are considered and illustrated. We assume that these parameters are not precisely known and therefore can be modeled as random variables with known laws. We calculate the Sobol indices of order 1 and 2 for each parameter and consider that the parameters follow the uniform law.

The paper is organized as follows. In section 2, we introduce the Lorenz-Stenflo model, describe different aspects of the climate system and study the sensitivity of the initial condition using the euclidean distance. In section 3, we introduce the theory of sensitivity analysis and the polynomial chaos method. In section 4, we compute the Sobol indices in the non-chaotic case and in the

chaotic case to understand the influence of the input random variables on the output variable. Finally, we draw some conslusions.

2 Model of a climate system

Today, with advanced technological tools and knowledge, scientists are able to model the climate system and predict how it will evolve over several decades. A model is a simplified representation of reality. A climate model thus aims to represent the climate and its evolution. As the climate is complex, climate models can take into account a fixed number of variables and their deviation from reality differs. But taking into account a large number of phenomena lengthens the calculation time. The climate models are then tested on known climates from the last decades to see if they are effective in terms of predictions. They are improved and then used to estimate future variations. There are several steps in building a climate model. Scientists first make observations (direct and indirect) and then apply the basic laws of physics, chemistry, biology and known mathematics. They study portions of planet Earth cut into a grid with a defined volume for simplicity. Many supercomputers make it possible to establish a model which is then compared to observations in order to be improved and therefore become more precise.

The LS system is a dynamic system for modeling atmospheric acoustic gravity waves in a rotating atmosphere. Knowledge of gravity acoustic waves is important as they can be responsible for both minor weather changes and large-scale phenomena. This kind of phenomenon occurs on any fluid subjected to a field of gravity, and can be made visible when there are several fluids arranged in several layers. In this case, the different layers in the atmosphere can give rise to visible gravity waves clouds.

The LS system [14] allowsone to show that disturbances of acoustic gravity at low frequency and short wavelength in the atmosphere can be described by a system of four generalized Lorenz equations. These coupled equations reduce to the usual three Lorenz equations [18] when the rotation of the earth is not considered:

$$\begin{cases} \dot{x} = \sigma(y - x) + sv \\ \dot{y} = rx - y - xz \\ \dot{z} = xy - bz \\ \dot{v} = -x - \sigma v \end{cases} \tag{1}$$

where the dot represents the derivative of the variable over time. The variable x which characterizes the intensity of the convection movement, y the horizontal temperature gradient, z the vertical temperature gradient and v which is proportional to the current function. The meaning of the parameters: σ depends on the properties of the fluid, b varies with the geometry of the convection cell, r varies according to the temperature gradient in the cell and s the rotation parameter of the system (depends on the angle of rotation and kinematic viscosity). Here, v and s are the new variable and parameter associated with rotational effects. More details of the LS system see [14, 22].

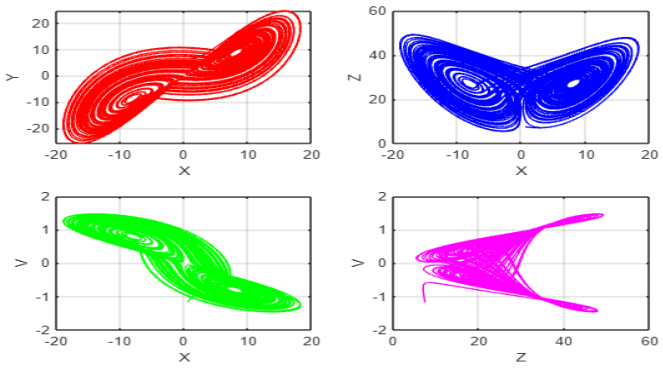

Figure 1: Chaotic attractors of the LS system in two-dimensional spaces.

Medium-scale meteorological phenomena such as thunderstorms, clouds, wind, intense precipitation at the origin of flash floods need to be understood and then modeled in order to be predicted by country level operational models. To study these phenomena, researchers are developing research models that reproduce their evolution. Intense initial scientific work and field experiments are necessary to validate the concepts envisaged and target poorly understood particularities of certain meteorological events. The research carried out on the prediction of rapidly developing meteorological phenomena constitutes a strong axis. Thanks to very fine mesh models (sometimes a few hundred meters), researchers represent rapidly developing dangerous phenomena (thunderstorms, fog, etc.). They develop digital tools to follow their trajectory or their development which will ultimately allow their forecasting just a few hours from the deadline. Research on the influence of relief on wind and clouds opens up interesting perspectives in terms of weather forecasting in the mountains. Past and future climate changes can be assessed using a global climate model. Such a model makes it possible to perform climate simulations from a few years to several millennia, and represents the actual planetary climate system digitally. The real "planet-climate" is in reality a very complex system, within which various environments are in perpetual interaction (the atmosphere, the ocean, sea ice, vegetation, rivers ...) exchange in particular water and energy all the time. These components of the climate system can be represented by numerical models, which are developed by specialized scientists and engineers. These models are based on the laws of physics and are regularly tested so that each of them correctly simulates the environment it represents.

Various dynamical behaviors such as stability, periodic and chaotic solutions, Lyapunov exponents spectra of the high-order LS equations (1) have been thoroughly studied [23, 26].
When we consider the values taken by Lorenz in his work [18], $\sigma = 10,\ b = 8/3,\ r = 28$, with $s = 10$ we have represented the chaotic system of LS (1), the evolutions of each state in Figure 2, the evolution in dimension 2 see Figure 1 and dimension 3 see Figure 3.

Climate models are also used to conduct seasonal forecasting research. Knowing the state of the atmosphere and the ocean on a given date, they can provide statistical information on seasonal climate trends a few months in advance in some parts of the world. It should be noted, however,

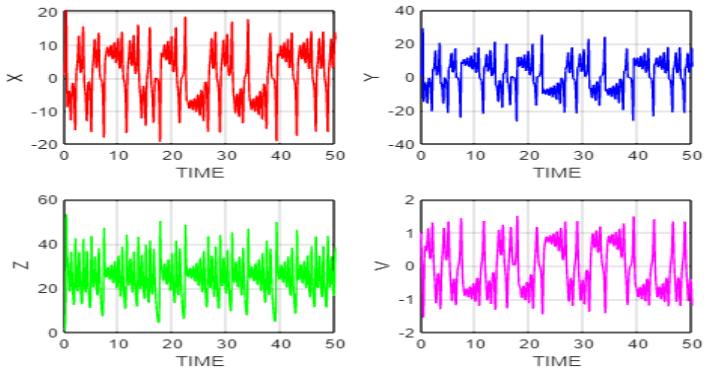

Figure 2: Evolution of each state on the chaotic case.

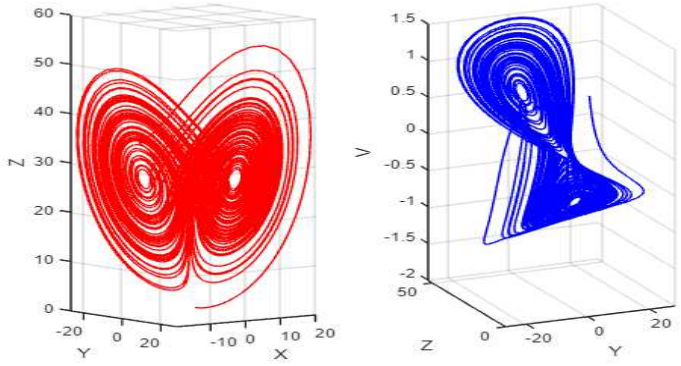

Figure 3: Chaotic attractors of the LS system in three-dimensional spaces.

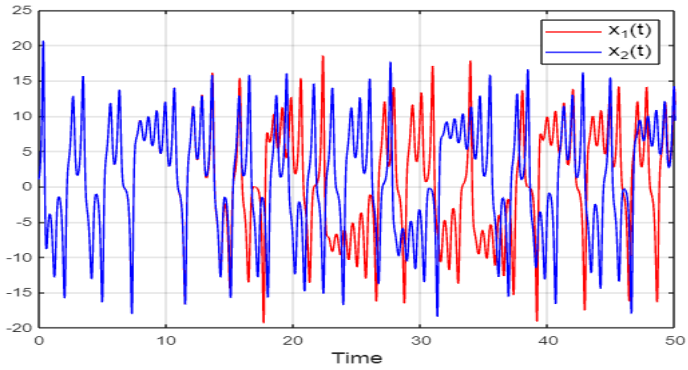

Figure 4: Representation of $x_1(t)$ and $x_2(t)$.

that these predictions are more reliable for the tropics than for mid-latitudes or polar regions, and temperatures are generally more predictable than precipitation.

In Figure 4, we have plotted the solution $x_1(t)$ with as initial condition $(1, 1, 1, 1)$ and also plotted the difference between $x_1(t)$ and another solution $x_2(t)$ with as initial condition $(1, 1, 1, 1.0001)$. We notice that, with very small modification on the initial condition, the two solutions separate after a certain time and will evolve in a totally different way.

The LS attractor is a fractal structure corresponding to the long-term behavior of the oscillator. The attractor shows how the different variables of the dynamic system evolve over time in a non-periodic trajectory. In 1963, meteorologist Edward Lorenz was the first to highlight the likely chaotic nature of meteorology. The LS model is a simplified modeling of meteorological phenomena based on fluid mechanics. This model is a dynamic four-dimensional system which generates chaotic behavior under certain conditions. LS model had important repercussions in showing the possible limits on the ability to predict long-term climate and meteorological evolution. This is an important part of the theory that the atmospheres of planets and stars can have a wide variety of quasi-periodic regimes and are subject to abrupt and random changes. It is also a useful example for the theory of dynamical systems serving as a source for new mathematical concepts.
The attractor resembles the two outstretched wings of a butterfly, see Figure 1. Each wing is formed by a series of concentric circles. The points describe several circles on one wing then switch to the other without any particular rhythm and without ever cutting their trajectories.

2.1 Stability analysis

The LS system (1) is symmetrical with respect to the transformation $(x, y, z, v) \mapsto (-x, -y, z, -v)$. It is also a dissipative system because the divergence of the vector field $\mathcal{V}(x, y, z, v)$ is equal to:

$$Div\big(\mathcal{V}(x, y, z, v)\big) = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} + \frac{\partial \dot{v}}{\partial v} = -\big(2\sigma + b + 1\big),$$

The dissipativity of this system depends only on σ and b, which is identical to the classical Lorenz system [18].
The LS system (1) have three fixed points which are: the origin $E_0 = (0, 0, 0, 0)$ and two others points symmetrical with respect to the transformation : $E_1 = \big(-\beta_1, -\beta_2, \beta_3, -\beta_4\big)$ and $E_2 = \big(\beta_1, \beta_2, \beta_3, \beta_4\big)$ where:

$$\beta_1 = \left[b\beta_3/\big(1 + \frac{s}{\sigma^2}\big)\right]^{1/2}, \quad \beta_2 = \left[b\beta_3\big(1 + \frac{s}{\sigma^2}\big)\right]^{1/2}, \quad \beta_3 = r - 1 - \frac{s}{\sigma^2}, \quad \beta_4 = -\frac{\beta_1}{\sigma}.$$

The linear instability of LS equations around stationary points has been studied in [23]. It has been found that in general the presence of the rotation parameter reduces the chaotic regime. Several periodic and chaotic solutions were also presented. For the value studied, the solutions were qualitatively similar to that of the Lorenz system [18]. For appropriate parameters, the origin E_0 and (E_1 and E_2) are saddle points. The point E_0 has a one-dimensional unstable manifold and three-dimensional stable manifold. The two points (E_1 and E_2) have one-dimensional stable

manifold and three-dimensional unstable manifold where the orbits have an outwards spiralling motio.

2.2 Sensitivity to initial condition

Chaos theory deals with rigorously deterministic dynamic systems, but which present a fundamental phenomenon of instability called sensitivity to initial conditions which, modulating an additional property of recurrence, makes them unpredictable in practice over the long term. A dynamic system is said to be chaotic if a "significant" portion of its phase space simultaneously exhibits the following two characteristics: the phenomenon of sensitivity to initial conditions and the strong recurrence. The presence of these two properties results in an extremely disordered behavior which is rightly described as chaotic. An excellent summary of the Lorenz system [18] has been written by Viana [32].
In this section, we illustrate this sensitivity by the euclidean distance which makes it possible to describe the evolution of the difference between trajectories which have slightly different initial conditions.
We consider (x, y, z, v) the solution of the equation (1) with initial condition $p_0 = (x_0, y_0, z_0, v_0)$ and we recall the expression of euclidean distance :

$$d_2\big(x, y, z, v\big) = \sqrt{x^2 + y^2 + z^2 + v^2} \tag{2}$$

2.2.1 Non-chaotic case

The notion of stability aims to formalize the property of a dynamic system such that the system remains close to a state called equilibrium. We choose the parameters $\sigma = 3.5,\ b = 2.5,\ r = 11.5,$ and $s = 1$ to have a non-chaotic regime. The two points of equilibrium are $E_1 = (-4.9055;\ -5.3065;\ 10.4187;\ 1.4028)$ and $E_2 = (4.9055;\ 5.3065;\ 10.4187;\ -1.4028)$. In the phase of stability, the trajectory will evolve and converge towards a fixed point. If we take $p_0 = (1; 1; 1; 1)$ as the initial condition, we notice, according to the Figures 5 and 6, that the solution converges to the point of equilibrium E_2.

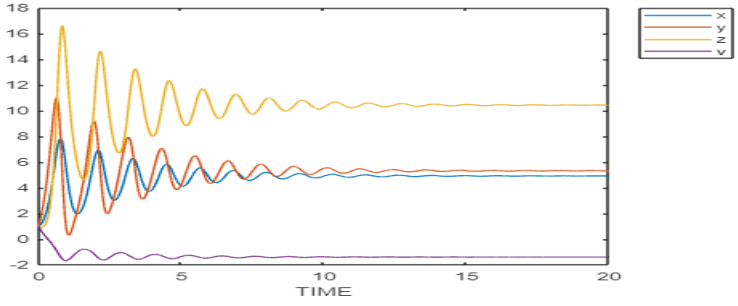

Figure 5: Evolution of states in the non-chaotic case with $\sigma = 3.5,\ b = 2.5,\ r = 11.5,\ s = 1$.

If we consider the evolution of (x, y), Figure 6 in red color, we can clearly see the convergence of the system towards the point of equilibrium $(4.9055; 5.3065)$.

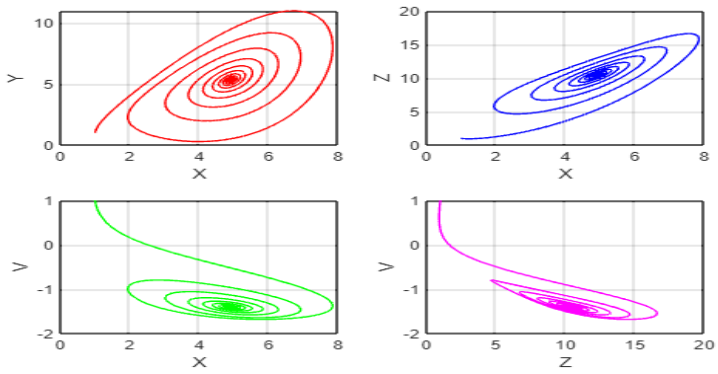

Figure 6: Evolution of states in the non-chaotic case $\sigma = 3.5,\ b = 2.5,\ r = 11.5,\ s = 1.$

Now we consider (x_1, y_1, z_1, v_1) the solution of the equation (1) with initial condition $p_1 = E_1$ and (x_2, y_2, z_2, v_2) another solution of the (1) with initial condition $p_2 = E_2$.
Now we are going to look at the euclidean distance around the equilibrium points and see if we have a convergence.

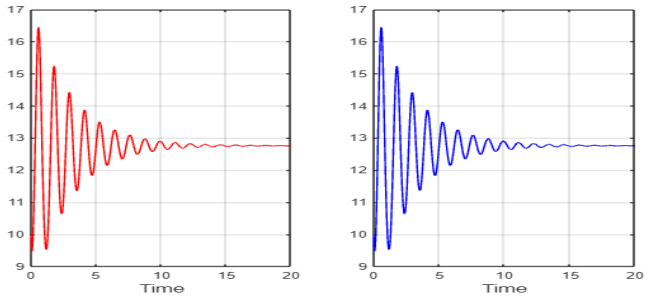

Figure 7: Evolution of the euclidean distance: in red for the initial condition p_1 and in blue for the initial condition p_2.

Around the points of equilibrium, we notice in Figure 7 that the distance is exactly the same and that it converges towards the same value this is due to the symmetry of two solutions.

Now we are going to plot the euclidean distance, see Figure 8, for the difference of two solutions (x_1, y_1, z_1, v_1) and (x_2, y_2, z_2, v_2) with their respective initial conditions $p_1 = E_1$ and $p_2 = E_2$.
We notice that the distance between the two solutions stabilizes, this allows us to say that the two solutions each converge towards its respective point of equilibrium.

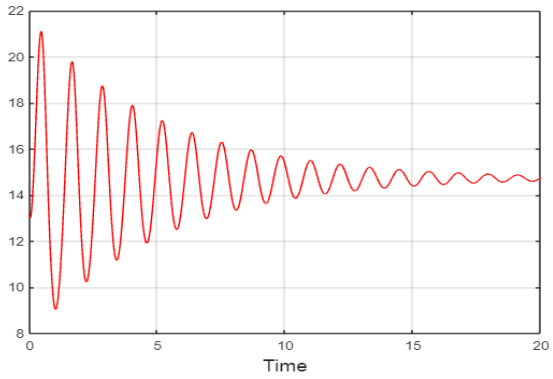

Figure 8: Evolution of the euclidean distance for the difference of two solutions.

2.2.2 Chaotic case

In the chaotic phase, we have the appearance of two strange attractors which are the two fixed points. We choose the parameters $\sigma = 0.5,\ b = 0.5,\ r = 18.5,$ and $s = 2.5$ to have a chaotic effect. The two points of equilibrium are $E_1 = (-0.5838;\ -6.4226;\ 7.5;\ 1.1677)$ and $E_2 = (0.5838;\ 6.4226;\ 7.5;\ -1.1677)$. If we take $p_0 = (1; 1; 1; 1)$ as the initial condition, we notice, according to the Figures 9 and 10, where we have represented the evolution of the transient regime that the solution evolves periodically indefinitely.

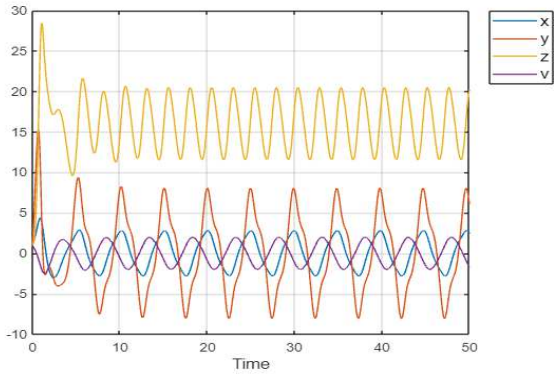

Figure 9: Evolution of states in the chaotic case with $\sigma = 0.5,\ b = 0.5,\ r = 18.5,\ s = 2.5$.

All the orbits of the system behave in the same way and tend towards an attractor set. This set is made up of orbits which revolve around the two points (E_1 and E_2), the same scenario is observed in the Lorenz system [18]. After studying the attractor set more closely, it seems that the Lorenz-Stenflo system has a strange attractor. Stange attractor because the solution will evolve around a point of equilibrium then will evolve in the same way around the other symmetrical point of equilibrium and will do the same thing over time. We can clearly see in Figure 9 that in the trajectories of the state variables, there is an ascent and a descent phase, this represents the return passage

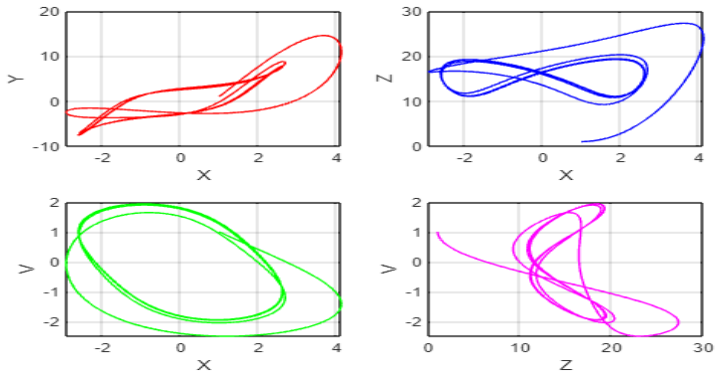

Figure 10: Evolution of states in the chaotic case $\sigma = 0.5,\ b = 0.5,\ r = 18.5,\ s = 2.5.$

from one point of equilibrium to another. This phase (ascent / descent) will continue indefinitely.

Indeed when the parameters σ, b, r and s take the following values: $\sigma = 0.5$, $b = 0.5$, $r = 18.5$ and $s = 2.5$, the differential dynamic system of LS presents a strange attractor represented in Figure 10. For almost all initial conditions, the orbit of the system is rapidly approaching the attractor, with the trajectory starting by winding up on one wing, then jumping from one wing to the other to start rolling up on the last one and erratically. The trajectory tends to rotate these two attractors in a random manner until forming two orbits around the attractors, see Figure 10. We therefore notice that whatever the starting point, the system is irresistibly attracted by the two buckles of the butterfly wings. We therefore qualify this kind of figure as a "strange attractor". Strange because the trajectories never intersect and yet seem to evolve at random.

We consider (x_1, y_1, z_1, v_1) the solution of the equation (1) with initial condition $p_1 = E_1 = (-0.5838;\ -6.4226;\ 7.5;\ 1.1677)$ and (x_2, y_2, z_2, v_2) another solution of the (1) with initial condition $p_2 = E_2 = (0.5838;\ 6.4226;\ 7.5;\ -1.1677)$. In Figure 11 , we have represented the evolution of the two solutions (x_1, y_1, z_1, v_1) and (x_2, y_2, z_2, v_2) aroud the equilibrium points.
We notice the symmetry of two solutions. For a period of time, one notices a pseudo-convergence of the solution towards its point of equilibrium and suddenly the solution crosses the origin and begins to evolve in the other point of equilibrium without converging and repeats the same thing indefinitely. Thanks to the symmetry, the two solutions will evolve in an identical way around the points of equilibrium and cross as many times at the point of origin. We also see that the two solutions have the same solution for the variable z and that the two respective solutions z_1 and z_2 are identical.
Around their respective equilibrium points, in the Figure 12, it seems that the distance converges at the beginning and suddenly the distance begins to evolve in an eratic way before stabilizing and evolving in a periodic way without converge.
In Figure 13, we see that the distance between the two solutions evolves and does not converge. We notice that we have several phases of (ascent / descent) which indicates to us that the two solutions once they are close (the ascent period)) and another time they are far apart (the descent period).

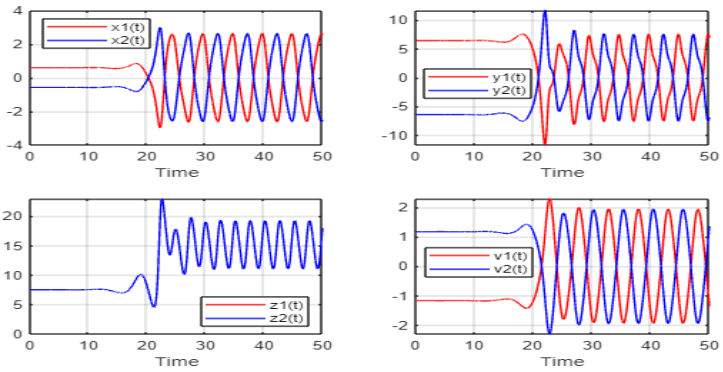

Figure 11: Evolution of the solutions (x_1, y_1, z_1, v_1) (resp. (x_2, y_2, z_2, v_2)) with the initial conditions p_1 (resp. p_2).

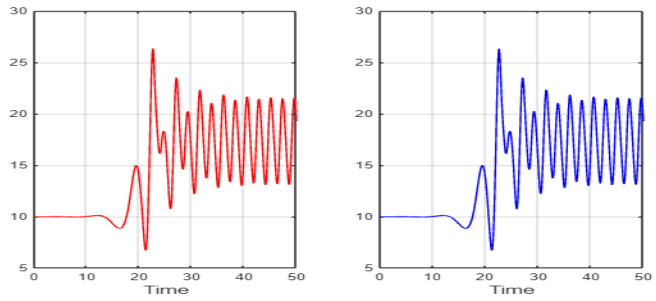

Figure 12: Evolution of the euclidean distance: in red for the initial condition p_1 and in blue for the initial condition p_2.

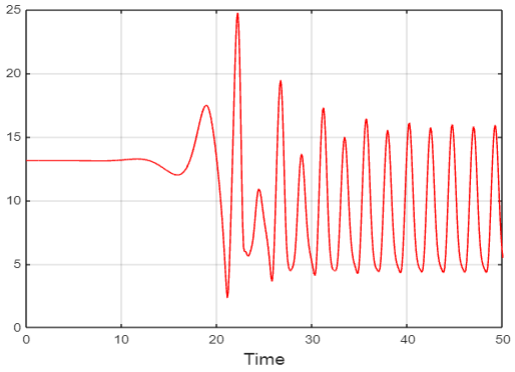

Figure 13: Evolution of the euclidean distance for the difference of two solutions.

Remark 2.1. *In the non-chaotic case, we can clearly see that small disturbances around the equilibrium points have no effect on the evolution of the system and on the euclidean distance. On the other hand, in the chaotic case, we notice that this changes with a great impact on the*

evolution of the solution and the euclidean distance.

3 Sensitivity analysis

3.1 Sobol indices

The sensitivity analysis studies how disturbances on the input variables of the model generate disturbances on the output variable. Indeed, by studying how the response of the model reacts to variations in its input variables, the sensitivity analysis makes it possible to answer a number of questions: what are the variables that most contribute to the variability of the model response? What are the least influential variables? Which variables, or which groups of variables, interact with which others? To answer these questions, we use the Sobol indices. These indices make it possible to predict consequences on the evolution of the climate according to the variation of the input parameters of the model in order to prevent or predict the consequences of climate change. To calculate the Sobol indices, applied to the numerical approximation of differential equations, we use the polynomials chaos method [2, 5, 11, 24]. The solution of the differential equation is represented by means of orthogonals polynomials. The coefficients of the polynomial basis are functions of time and can be calculated by solving a system of deterministic ordinary differential equations. Numerical examples are presented to illustrate the accuracy and efficiency of the proposed method.

During the development, construction or use of a mathematical model, sensitivity analysis can prove to be a valuable tool. It is possible to group the sensitivity analysis methods into three classes: screening methods, which consist of a qualitative analysis sensitivity of the output variable to the input variables, local analysis methods [30] , which quantitatively assess the impact of a small variation around a given value of the inputs and finally the global sensitivity analysis methods which are interested in the variability of the output of the model in the whole of its range of variation. Global sensitivity analysis studies how the variability of inputs affects that of output, by determining how much of the variance of output is due to a given input or set of inputs. While the local sensitivity analysis is more concerned with the value of the response variable, the global sensitivity analysis is concerned with its variability. In this work, we use the global sensitivity namely the Sobol indices.

We denote by $(\Omega, \mathcal{A}, \mathbb{P})$ the probability space. We consider a random variable X with values in $\mathbb{R}$ which describes the input uncertainties. The probability law of X can be defined by the probability density $f(x)$ over $(\Omega, \mathcal{A}, \mathbb{P})$. We also denote by $\mathbb{L}_2(\Omega, \mathcal{A}, \mathbb{P})$ the space of real random variables with finite second order moments. $\mathbb{L}_2(\Omega, \mathcal{A}, \mathbb{P})$ is a Hilbert space which can be provided with a complete orthogonal basis $\{\Psi_j(x)\}_{j\geqslant 0}$ with respect to the density of X.
Sobol's sensitivity indices are used when considering the following model :

$$Y = g(\mathbf{X}) = g(X_1, ..., X_p), \tag{3}$$

with $\mathbf{X} \in I_{\mathbf{X}} \subset \mathbb{R}^p$ the input random vector and $Y \in I_Y \subset \mathbb{R}$ the output variable. The input variables X_i, with $i = 1, ..., p$, are independent and the function g of this model may not be known

explicitly.

We recall the following formulas which are essential for the calculations of the Sobol indices:

$$\mathbb{E}(Y|X=x) = \int_{I_Y} y f_{Y|X=x}(y)\mathrm{d}y = \int_{I_Y} y \frac{f_{X,Y}(x,y)}{f_X(x)}\mathrm{d}y, \tag{4}$$

$$\mathbb{V}(Y|X=x) = \mathbb{E}(Y^2|X=x) - \big(\mathbb{E}(Y|X=x)\big)^2. \tag{5}$$

For such a model, there are 2^{p-1} indices, one for each variable and each interaction of variables. These indices belong to the category of global indices, since they represent the proportion of the variance explained by the input variable concerned. They allow us to measure the influence of the uncertainties of the input variables on the output variable Y. More specifically, they can give us information about: the input variables that generate the most variability on the output variable Y, variables which have little influence on the Y output, the interactions between input variables which have an impact on the output Y.
For the calculation of the Sobol indices, we need the following theorem:

Theorem 3.1. *Let X and Y be two random variables, continuous or discrete, such that $\mathbb{E}(|Y|) < \infty$, then the expectation of Y satisfies:*

$$\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y|X)\big).$$

For all random variables X and Y, the variance of Y satifies:

$$\mathbb{V}(Y) = \mathbb{E}\big(\mathbb{V}(Y|X)\big) + \mathbb{V}\big(\mathbb{E}(Y|X)\big). \tag{6}$$

In formula (6), we notice that the variance $\mathbb{V}(\mathbb{E}(Y|X))$ increases with the influence of X. Indeed, if X has a strong influence on Y, it also has one on its variance . By fixing X, we then have the variance of $Y|X$ which is on average smaller than Y. And the more this influence increases, the more the expectation $\mathbb{E}(\mathbb{V}(Y|X))$ decreases. According to formula (6), we thus have the variance $\mathbb{V}(\mathbb{E}(Y|X))$ which increases with the decrease in $\mathbb{E}(\mathbb{V}(Y|X))$. It is precisely this variance which is used in the calculation of the so-called first order Sobol indices defined in the following definition:

Definition 3.1. *We denote by S_i the part of the variance of Y to the variable X_i:*

$$S_i = \frac{\mathbb{V}\big(\mathbb{E}(Y|X_i)\big)}{\mathbb{V}(Y)} \qquad i = 1, ..., p. \tag{7}$$

This first-order indice therefore gives us an evaluation of the influence of an input variable X_i on the output variable Y. The S_i indices, between 0 and 1, is high when the influence of X_i is large.

Then, higher order Sobol indices, assess the importance of the combined effects of several input variables on the output variable. We find their expression from a decomposition of the variance, in the same way as for the first order indices. This variance decomposition is based on the Hoeffding-Sobol decomposition of a function [28]. This is generally presented with input variables of law $\mathcal{U}([0,1])$, but it is also true in the general case [6, 31].

Definition 3.2. ***Hoeffding-Sobol decomposition***
Let $g : I = I_1 \times ... \times I_p \subset \mathbb{R}^p \to \mathbb{R}$ a squared integrable function with respect to the joint density function $f : I \to \mathbb{R}^+$ of p variables $(X_1, ..., X_p)$. It is assumed that these variables are independent. Then we have g in (3) *which admits a unique decomposition of the form:*

$$Y = g(X_1, ..., X_p) = h_0 + \sum_{i=1}^{p} h_i(X_i) + \sum_{1 \leqslant i \leqslant j \leqslant p} h_{i,j}(X_i, X_j) + ... + h_{1,...,p}(X_1, ..., X_p), \quad (8)$$

where

$$\begin{cases} h_0 = \mathbb{E}(Y), \\ h_i(X_i) = \mathbb{E}\big(Y|X_i\big) - h_0, \\ h_{i,j}(X_i, X_j) = \mathbb{E}\big(Y|X_i, X_j\big) - h_i(X_i) - h_j(X_j) - h_0, \\ .... \end{cases} \quad (9)$$

for all $k = 1, ..., s$

$$\int_{\mathbb{R}} h_{i_1,...,i_s}(x_{i_1}, ..., x_{i_s}) f_{X_{i_k}}(x_{i_k}) \mathrm{d}x_{i_k} = 0, \quad (10)$$

$f_{X_{i_k}}$ is density of X_{i_k} .

We deduce from the conditions (10) and the independence of the variables $\big(X_i\big)_{1 \leqslant i \leqslant p}$ that the functions h_U (with $U \subseteq \{1, ..., p\}$) are orthogonal :

$$\int_{I_X} h_U(x_u) h_V(x_V) f_{\mathbf{X}}(x) \mathrm{d}x = 0,$$

if $U \neq V$, with $U, V \subseteq \{1, ..., p\}$ and $f_{\mathbf{X}}(x)$ the density function of the vector $\mathbf{X} = (X_1, ..., X_p)$. To obtain sensitivity indices for several variables of order greater than 1, we use the decomposition of the variance:

$$\mathbb{V}(Y) = \left(\sum_{i=1}^{p} V_i \right) + \sum_{1 \leqslant i < j \leqslant p} V_{i,j} + ... + V_{1,...,p} \ , \quad (11)$$

where

$$\begin{cases} V_i = \mathbb{V}\big(\mathbb{E}(Y|X_i)\big), \\ V_{i,j} = \mathbb{V}\big(\mathbb{E}(Y|X_i, X_j)\big) - V_i - V_j, \\ ..... \\ V_{1,...,p} = \mathbb{V}(Y) - \sum_{i=1}^{p} V_i - \sum_{1 \leqslant i < j \leqslant p} V_{i,j} - ... - \sum_{1 \leqslant i_1 < i_2 < ... < i_{p_1} \leqslant p} V_{i_1,...,i_{p-1}}. \end{cases}$$

Definition 3.3. *By equations* (7)*,* (10) *and* (11) *we obtain the formulas of the various Sobol indices of orders greater than 1:*

$$\begin{cases} S_{i,j} = \dfrac{\mathbb{V}(h_{i,j}(X_i, X_j))}{\mathbb{V}(Y)} = \dfrac{\mathbb{V}\big(\mathbb{E}(Y|X_i, X_j)\big)}{\mathbb{V}(Y)} - S_i - S_j, \\ S_{i,j,k} = \dfrac{\mathbb{V}(h_{i,j,k}(X_i, X_j, X_k))}{\mathbb{V}(Y)} = \dfrac{\mathbb{V}\big(\mathbb{E}(Y|X_i, X_j, X_k)\big)}{\mathbb{V}(Y)} - S_i - S_j - S_k - S_{i,j} - S_{i,k} - S_{j,k}, \\ .... \\ S_{1,...,p} = \dfrac{\mathbb{V}(h_{i,j,k}(X_1, ..., X_p))}{\mathbb{V}(Y)} = \dfrac{\mathbb{V}\big(\mathbb{E}(Y|X_1, ..., X_p)\big)}{\mathbb{V}(Y)} - \displaystyle\sum_{U \subseteq \{1,...,p\}} S_U. \end{cases} \tag{12}$$

Corollary 3.1. *By decomposing the variance, equation* (11)*, we have the sum of the sensitivity indices equal to 1.*

These sensitivity indices can easily be interpreted: the more indice S_i (respectively $S_{i,j,k}$) will be close to 1, the greater the influence of variable X_i (respectively of variables X_i, X_j, X_k).
We can also study the total sensitivity indice for the variable X_i. This represents the sum of the sensitivity indices involving the variable X_i. We note $S_{T_i} = \sum_{j \in I_i} S_j$, where I_i represents all the sets of indices containing the indice i.
Sobol indices are easier to compute from a meta-model such as Chaos's polynomial chaos. This is why we introduce Chaos's polynomial in the next section.

3.2 Polynomials chaos

In the work of Wiener in 1938 [33] and the book by Ghanem and Spanos in the 1990s [8] chaos polynomials (PC) have successfully solved a wide variety of random problems in different domains. The decomposition in PC is an efficient way to build a model allowing to study the propagation of the randomness in a complex system. It consists in writing the random variable of interest Y as a function of a random variable taken as input X under the form:

$$Y(X) = \sum_{i=0}^{p} \alpha_i \Psi_i(X),$$

with $\{\Psi_i\}_{0 \leqslant i \leqslant p}$ the orthonormal polynomials and form a basis of degree p. The α_i must be determined. There are several techniques for this: intrusive techniques, see [5, 24] and non-intrusive techniques, see [11, 21].
After obtaining the coefficients α_i, the mean, the variance or the Sobol indices of the output Y can be easily calculated. Indeed, when one wishes to model a data sample, an a priori distribution law is chosen and its parameters are estimated, in particular by using the maximum likelihood estimator or the method of moments.
First we explain in more detail the chaos polynomial expansion of a function. Secondly, we describe how one can calculate the coefficients of the chaos polynomials.

3.2.1 Development in polynomials chaos

The expansion into chaos polynomials is based on orthogonal polynomials, that is why, in this section we start by making some reminders on this topic. Let E be the real vector space of all polynomials with a single variable and real coefficients, and the scalar product of E defined by:

$$< u, v >= \int_I u(x)v(x)f(x)\mathrm{d}x \qquad \forall u, v \in E, \tag{13}$$

where $f : I \subset \mathbb{R} \to \mathbb{R}^+$ is a nonnegative integrable function of x. The set of polynomials $\{\Psi_n\}_{n\geqslant 0}$ are said to be orthogonal with respect to the function f,

$$< \Psi_n, \Psi_m >= \int_I \Psi_n(x)\Psi_m(x)f(x)\mathrm{d}x = h_n^2\delta_{n,m} \qquad n, m \in \mathbb{N}, \tag{14}$$

where δ is the Kronecker's delta functions and h_n are non-zero constants. We recall that, for orthogonal polynomials of degree $d = 0$, $\Psi_0 = 1$. The system (14) is called orthonormal if $h_n = 1$. We use the following recurrence relation to construct these polynomials:

$$\begin{cases} \Psi_{n+1}(x) = (x - a_n)\Psi_n(x) - b_n\Psi_{n-1}(x) \\ \Psi_0(x) = 1, \qquad \Psi_{-1}(x) = 0 \end{cases} \tag{15}$$

with

$$\begin{cases} a_n = \dfrac{< x\Psi_n, \Psi_n >}{< \Psi_n, \Psi_n >} \qquad n \in \mathbb{N} \\ b_n = \dfrac{< \Psi_n, \Psi_n >}{< \Psi_{n-1}, \Psi_{n-1} >} \qquad n \in \mathbb{N}^* \end{cases} \tag{16}$$

Example 3.1. *Legendre polynomials form a complete basis of polynomials orthogonal with respect to the density of a uniform law $\mathcal{U}([-1, 1])$. The first polynomials of this basis, defined for all $x \in [-1, 1]$, are constructed as: $Q_{-1}(x) = 0$ and $Q_0(x) = 1$*

$$a_0 = \frac{< xQ_0(x), Q_0(x) >}{< Q_0(x), Q_0(x) >} = \frac{\int_{-1}^1 x\mathrm{d}x}{\int_{-1}^1 1\mathrm{d}x} = 0,$$

and $b_0 = 0$, so we have $Q_1(x) = (x - a_0)Q_0(x) - b_0Q_{-1}(x) = x$,

$$a_1 = \frac{< xQ_1(x), Q_1(x) >}{< Q_1(x), Q_1(x) >} = \frac{\int_{-1}^1 x^3\mathrm{d}x}{\int_{-1}^1 x^2\mathrm{d}x} = 0,$$

$$b_1 = \frac{< Q_1(x), Q_1(x) >}{< Q_0(x), Q_0(x) >} = \frac{2/3}{2} = \frac{1}{3},$$

$Q_2(x) = (x - a_1)Q_1(x) - b_1Q_0(x) = x^2 - 1/3$,

$$a_2 = \frac{< xQ_2(x), Q_2(x) >}{< Q_2(x), Q_2(x) >} = \frac{\int_{-1}^1 x(x^2 - 1/3)^2\mathrm{d}x}{\int_{-1}^1 (x^2 - 1/3)^2\mathrm{d}x} = 0,$$

$$b_2 = \frac{< Q_2(x), Q_2(x) >}{< Q_1(x), Q_1(x) >} = \frac{8/45}{2/3} = \frac{4}{15},$$

$Q_3(x) = (x - a_2)Q_2(x) - b_2Q_1(x) = x(x^2 - 1/3) - 4/15x = x^3 - 3/5x.$

Remark 3.1. *The polynomials constructed according to the formulas* (15) *and* (16) *are only orthogonal. For all $j \in \mathbb{N}$, we can divide the polynomial $Q_j(x)$ by the root of its norm $< Q_j(x), Q_j(x) >$ so that it becomes orthonormal.*

Example 3.2. *Hermite polynomials form a basis of polynomials orthogonal to the density of a normal distribution $\mathcal{N}(0, 1)$. For all $n \in \mathbb{N}$, corresponding to the degree of the polynomial, and for all $x \in \mathbb{R}$, they are defined by:*

$$Q_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n}\left(e^{-x^2/2}\right). \tag{17}$$

We can prove by recurrence, for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$,

$$\frac{d^n}{dx^n}\left(e^{-x^2/2}\right) = e^{-x^2/2} \sum_{k=0}^{[n/2]} (-1)^{n+k} \frac{n!}{2^k k!(n-2k)!} x^{n-2k}, \tag{18}$$

$$Q_n(x) = \sum_{k=0}^{[n/2]} (-1)^k \frac{n!}{2^k k!(n-2k)!} x^{n-2k},$$

with $[n/2]$ the whole part of $n/2$.

The first polynomials of the Hermite base are therefore the following:

$$Q_0(x) = 1,\ Q_1(x) = x,\ Q_2(x) = x^2 - 1,\ Q_3(x) = x^3 - 3x,\ Q_4(x) = x^4 - 6x^2 + 3.$$

All these polynomials are orthogonal with respect to the density of the reduced centered normal distribution.

When analytical calculations are not possible a base of chaos polynomials can be constructed according to (15) and (16) using quadrature formula.
We consider $Y = Y(\mathbf{X}) = Y(X_1, ..., X_p)$ a given output function which is assumed to belong $\mathbb{L}_2(\Omega, \mathcal{A}, \mathbb{P})$ can be represented by [8, 34]:

$$Y(\mathbf{X}) = \sum_{i=0}^{\infty} \alpha_i \Psi_i(\mathbf{X}) \tag{19}$$

In numerical simulations, the sum (19) is truncated (see [34]) by keeping the terms less than one degree P:

$$Y(\mathbf{X}) \approx \tilde{Y}(\mathbf{X}) = \sum_{i=0}^{P} \alpha_i \Psi_i(\mathbf{X}) \tag{20}$$

where $P+1 = \dfrac{(p+n)!}{n!p!}$ with p the number of independent random variables and n the degree of the polynomials chaos.

Once this approximation of the system response has been found, in the form of a decomposition into polynomials chaos, it is easy to calculate the different macroscopic quantities. For example, the mean of $\tilde{Y}$ is simply the first coefficient of the decomposition into polynomials chaos,

$$\mathbb{E}(\tilde{Y}) = \sum_{i=0}^{P} \alpha_i \mathbb{E}(\Psi_i(X)) = \sum_{i=0}^{P} \alpha_i \int_\Omega \Psi_i(x) f_X(x) \mathrm{d}x = \sum_{i=0}^{P} \alpha_i \int_\Omega \Psi_0(x) \Psi_i(x) f_X(x) \mathrm{d}x,$$

where f_X is the density function of X. By orthonormality of $\{\Psi_i(x)\}_{i \geqslant 0}$, we have $\mathbb{E}(\tilde{Y}) = \alpha_0$. In the same way, we find

$$\mathbb{V}(\tilde{Y}) = \sum_{i=1}^{P} \alpha_i^2.$$

In the next section, we detail a method to calculate the coefficients α_i when a functional relation of the form $Y = g(X)$ is known: the projection by Galerkin.

3.2.2 Coefficients of polynomials chaos

Let $\tilde{Y}(\mathbf{X})$ be a response of the form (20), with $\{\Psi_i\}_{0 \leqslant i \leqslant P}$ a family of orthonormal polynomials, which models the system $Y = g(\mathbf{X})$. We can represent g as a black box representing our model of such that for a given input value x_i the response of the system $y_i = g(x_i)$ is computable. To calculate the coefficients α_i of (20), we operate as follows:

$$\tilde{Y}(\mathbf{X}) = g(\mathbf{X}) = \sum_{i=0}^{P} \alpha_i \Psi_i(\mathbf{X}),$$

multiply by $\Psi_j(\mathbf{X})$,

$$\tilde{Y}(\mathbf{X}) \Psi_j(\mathbf{X}) = \sum_{i=0}^{P} \alpha_i \Psi_i(\mathbf{X}) \Psi_j(\mathbf{X}),$$

taking the expectation value, we have

$$\mathbb{E}(\tilde{Y}(\mathbf{X}) \Psi_j(\mathbf{X})) = \mathbb{E}\left(\sum_{i=0}^{P} \alpha_i \Psi_i(\mathbf{X}) \Psi_j(\mathbf{X}) \right).$$

Then, using the linearity of the expectation and the orthonormality of the polynomials $\Psi_i(\mathbf{X})$, we obtain

$$\alpha_i = \mathbb{E}\left(\tilde{Y}(\mathbf{X}) \Psi_i(\mathbf{X}) \right) = \int_\Omega g(x) \Psi_i(x) f(x) \mathrm{d}x \tag{21}$$

with $f(x)$ the density of the random variable $\mathbf{X} \in \Omega$.

This procedure is known as the Galerkin projection in the deterministic community and it also

has a strong connection with the linear square mean technique used in the statistical community (the calculation of the minimum of $\mathbb{E}\big[(g(\mathbf{X}) - \sum_{i=0}^{P} \alpha_i \Psi_i(\mathbf{X}))^2\big]$). The next step is to evaluate the above integral, using the known points of the function g. When the problem contains a low number of input random variables, Gaussian quadrature rules are generally used (see [9]). For a single variable, they take the form

$$\int_I h(x) f(x) \mathrm{d}x \approx \sum_{k=1}^{m} \omega_k h(x_k), \tag{22}$$

for any function h integrable over I. The construction of the quadrature rules is usually carried out in tandem with the construction of polynomials orthogonal with respect to the weight function $f(x)$ (see [7]). Note that the points and weights are such that the formula (22) is exact when h is a polynomial of lower degree or equal to $2^m - 1$. By combining (21) and (22), we obtain the expression for the computation of the coefficients of the following polynomial chaos:

$$\alpha_i = \sum_{k=0}^{m} \omega_k g(x_k) \Psi_i(x_k) \tag{23}$$

It is important to stress that the computation of α_i requires the evaluation of $g(x_k)$ at the specified points x_k, see [3, 4, 25] for more explanation.
For more details on this polynomial chaos method see also [29, 34, 35] which applies the method with random variables and stochastic processes.

We break down the state variables that satisfy the system (1) in the form of polynomials of degree n and we fix the time $t = t_k$. We consider that the parameters $\mathbf{X} = (\sigma, b, r, s)$ are random and we decompose the state variables as follows :

$$\begin{cases} x(t_k, \mathbf{X}) = x(t_k, \sigma, b, r, s) = \sum_{i=0}^{P} x_i(t_k) \Psi_i(\sigma, b, r, s), \\ y(t_k, \mathbf{X}) = y(t_k, \sigma, b, r, s) = \sum_{i=0}^{P} y_i(t_k) \Psi_i(\sigma, b, r, s), \\ z(t_k, \mathbf{X}) = z(t_k, \sigma, b, r, s) = \sum_{i=0}^{P} z_i(t_k) \Psi_i(\sigma, b, r, s), \\ v(t_k, \mathbf{X}) = v(t_k, \sigma, b, r, s) = \sum_{i=0}^{P} v_i(t_k) \Psi_i(\sigma, b, r, s), \end{cases} \tag{24}$$

with $P+1 = \dfrac{(4+n)!}{n!4!}$, where $\{\Psi_i\}$ is a complete basis of orthonormal polynomials of degree n.

We illustrate the method on the component $x(t_k, \sigma, b, r, s)$. We multiply by Ψ_j, use orthogonality and then we integrate: we find

$$\mathbb{E}\bigg(x(t_k,\sigma,b,r,s)\Psi_j(\sigma,b,r,s)\bigg)=\sum_{i=0}^{P}x_i(t_k)\mathbb{E}\bigg(\Psi_i(\sigma,b,r,s)\Psi_j(\sigma,b,r,s)\bigg)=x_j(t_k),\quad(25)$$

in particular, with $\Psi_0(\sigma,b,r,s)=1$, therefore $\mathbb{E}\bigg(x(t_k,\sigma,b,r,s)\bigg)=x_0(t_k)$.

We fix the points $(\sigma_{q_1},b_{q_2},r_{q_3},s_{q_4})$ and we determine the coefficients $x_j(t_k),\ y_j(t_k),\ z_j(t_k),\ v_j(t_k)$ by solving the system:

$$\begin{cases}\dfrac{dx(t,\sigma_{q_1},b_{q_2},r_{q_3},s_{q_4})}{dt}=\sigma_{q_1}(y-x)+s_{q_4}v,\\ \dfrac{dy(t,\sigma_{q_1},b_{q_2},r_{q_3},s_{q_4})}{dt}=r_{q_3}x-y-xz,\\ \dfrac{dz(t,\sigma_{q_1},b_{q_2},r_{q_3},s_{q_4})}{dt}=xy-b_{q_2}z,\\ \dfrac{dv(t,\sigma_{q_1},b_{q_2},r_{q_3},s_{q_4})}{dt}=-x-\sigma_{q_1}z,\end{cases}\quad(26)$$

with as initial condition (x_0,y_0,z_0,v_0).

Furthermore, it can be shown that

$$\big(x(t_k,\sigma,b,r,s)\big)^2=\sum_{i,j=0}^{P}x_i(t_k)x_j(t_k)\Psi_i(\sigma,b,r,s)\Psi_j(\sigma,b,r,s).$$

By using orthonormality of the $\{\Psi_i\}$, we have:

$$\begin{aligned}\mathbb{E}\bigg(\big(x(t_k,\sigma,b,r,s)\big)^2\bigg)&=\sum_{i,j=0}^{P}x_i(t_k)x_j(t_k)\mathbb{E}\big(\Psi_i(\sigma,b,r,s)\Psi_j(\sigma,b,r,s)\big)\\&=\sum_{i=0}^{P}\big(x_i(t_k)\big)^2\mathbb{E}\bigg[\Psi_i^2(\sigma,b,r,s)\bigg]=\sum_{i=0}^{P}\big(x_i(t_k)\big)^2.\end{aligned}$$

After simplication, we have:

$$\mathbb{V}\big(x(t_k,\sigma,b,r,s)\big)=\sum_{i=0}^{P}\big(x_i(t_k)\big)^2-\big(x_0(t_k)\big)^2=\sum_{i=1}^{P}\big(x_i(t_k)\big)^2.$$

For the calculation of the Sobol indices, we use the method presented by B. Sudret in [29], which consists in using a base of polynomials chaos. By construction, the base $\{\Psi_j(\mathbf{X})\}_{j\geqslant 0}=\{\Psi_j(X_1,X_2,X_3,X_4)\}_{j\geqslant 0}=\{\Psi_j(\sigma,b,r,s)\}_{j\geqslant 0}$ satisfied two points.
Firstly, for all $(\sigma,b,r,s)\in\mathbb{R}^4$,

$$\Psi_0(\sigma,b,r,s)=1.$$

The second point, to each $j \in \{0, 1, ..., P\}$ corresponds a multi-index $(\gamma_1, \gamma_2, \gamma_3, \gamma_4) \in \{0, 1, ..., n\}^4$ such that

$$\begin{aligned}\Psi_j(\mathbf{X}) &= \Psi_j(X_1, X_2, X_3, X_4) = \Psi_j(\sigma, b, r, s) \\ &= \Psi_{\gamma_1,\gamma_2,\gamma_3,\gamma_4}(\sigma, b, r, s) = \Psi^1_{\gamma_1}(\sigma)\Psi^2_{\gamma_2}(b)\Psi^3_{\gamma_3}(r)\Psi^4_{\gamma_4}(s),\end{aligned} \tag{27}$$

with for all $(h, i) \in \{1, 2, 3, 4\}^2$, $\Psi^h_0(X_h) = 1$, $\Psi^h_{\gamma_i}$ which is a polynomial of degree γ_i, and $\{\Psi^h_{\gamma_i}\}$ which forms a one-dimensional basis of polynomials orthogonal with respect to the density f_{X_h} of X_h, that is to say: $\forall (l, m) \in \{0, ..., n\}^2$ with $l \neq m$,

$$\int_{\mathbb{R}} \big(\Psi^h_l(x_h)\big)^2 f_{X_h}(x_h)\mathrm{d}x_h = 1 \text{ and } \int_{\mathbb{R}} \Psi^h_l(x_h)\Psi^h_m(x_h)f_{X_h}(x_h)\mathrm{d}x_h = 0. \tag{28}$$

According to the formula (27), we can then rewrite the equation (24) for the component x as follows:

$$x(t_k, \sigma, b, r, s) = \sum_{(\gamma_1,...,\gamma_4)\in\{0,...,n\}^4} x_{\gamma_1,...,\gamma_4}(t_k)\Psi^1_{\gamma_1}(\sigma)\Psi^2_{\gamma_2}(b)\Psi^3_{\gamma_3}(r)\Psi^4_{\gamma_4}(s), \tag{29}$$

with $(\gamma_1, ..., \gamma_4) \in \{0, ..., n\}^4$ such that $\sum_{i=1}^4 \gamma_i \leqslant n$.

Now we calculate the Sobol indice S_{x_r} for x with respect to the influence of the parameter r :

$$S_{x_r}(t_k) = \frac{\mathbb{V}\Big(\mathbb{E}\big(x(t_k, \sigma, b, r, s)\big|r\big)\Big)}{\mathbb{V}\big(x(t_k, \sigma, b, r, s)\big)} = \frac{\mathbb{E}\Big[\Big(\mathbb{E}\big(x(t_k, \sigma, b, r, s)\big|r\big)\Big)^2\Big] - \Big[\mathbb{E}\Big(\mathbb{E}\big(x(t_k, \sigma, b, r, s)\big|r\big)\Big)\Big]^2}{\mathbb{V}\big(x(t_k, \sigma, b, r, s)\big)}. \tag{30}$$

The Sobol indices of the first order (30) can then be expressed as a function of the coefficients of the decomposition into polynomials chaos (29). It only remains to calculate the following conditional expectation:

Firstly, we calculate the expectation $\mathbb{E}\big(x(t_k, \sigma, b, r, s)\big|r\big)$:

$$\begin{aligned}\mathbb{E}\Big(x(t_k,\sigma,b,r,s)\big|r\Big) &= \sum_{j=(\gamma_1,\dots,\gamma_4)\in\{0,\dots,n\}^4} x_j(t_k)\Psi^3_{\gamma_3}(r)\mathbb{E}\Big(\prod_{h\leqslant 4;h\neq 3}\Psi^h_{\gamma_h}(X_h)\Big)\\ &= \sum_{j=(\gamma_1,\dots,\gamma_4)\in\{0,\dots,n\}^4} x_j(t_k)\Psi^i_{\gamma_3}(r)\\ &\times \prod_{h\leqslant 4;h\neq 3}\Big(\int_{\mathbb{R}}\Psi^h_0(x_h)\Psi^h_{\gamma_h}(x_h)f_{X_h}(x_h)\mathrm{d}x_h\Big)\\ &= \sum_{\gamma_3=0}^{n} x_{0,0,\gamma_3,0}(t_k)\Psi^3_{\gamma_3}(r), \end{aligned} \tag{31}$$

which gives the variance,

$$\mathbb{V}\Big(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|r\big)\Big) = \sum_{\gamma_3=0}^{n}\big(x_{0,0,\gamma_3,0}(t_k)\big)^2\mathbb{V}\big(\Psi^3_{\gamma_3}(r)\big) = \sum_{\gamma_3=1}^{n}\big(x_{0,0,\gamma_3,0}(t_k)\big)^2, \tag{32}$$

according to the equation (28).

We then have the Sobol indice S_{x_r}:

$$S_{x_r}(t_k) = \frac{\mathbb{V}\Big(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|r\big)\Big)}{\mathbb{V}\big(x(t_k,\sigma,b,r,s)\big)} = \frac{\sum_{\gamma_3=1}^{n}\big(x_{0,0,\gamma_3,0}(t_k)\big)^2}{\sum_{i=1}^{P}\big(x_i(t_k)\big)^2}. \tag{33}$$

The other Sobol indices of order 1 (S_{x_σ}, S_{x_b} and S_{x_s}) are obtained in the same way.

Now we calculate the Sobol indice $S_{x_{b,r}}$ for x with respect to the influence of the parameters b and r

$$S_{x_{b,r}} = \frac{\mathbb{V}\Big(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|b,r\big)\Big)}{\mathbb{V}\big(x(t_k,\sigma,b,r,s)\big)} = \frac{\mathbb{E}\Big[\Big(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|b,r\big)\Big)^2\Big] - \Big[\mathbb{E}\Big(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|b,r\big)\Big)\Big]^2}{\mathbb{V}\big(x(t_k,\sigma,b,r,s)\big)}.$$

Firstly we calculate the expectation $\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|b,r\big)$:

$$\begin{aligned}\mathbb{E}\bigg(x(t_k,\sigma,b,r,s)\big|b,r\bigg) &= \sum_{j=(\gamma_1,\ldots,\gamma_4)\in\{0,\ldots,n\}^4} x_j(t_k)\Psi^2_{\gamma_2}(b)\Psi^3_{\gamma_3}(r)\mathbb{E}\bigg(\prod_{h\leqslant 4;h\neq\{2,3\}}\Psi^h_{\gamma_h}(X_h)\bigg)\\ &= \sum_{j=(\gamma_1,\ldots,\gamma_4)\in\{0,\ldots,n\}^4} x_j(t_k)\Psi^2_{\gamma_2}(b)\Psi^i_{\gamma_3}(r)\\ &\times \prod_{h\leqslant 4;h\neq\{2,3\}}\bigg(\int_{\mathbb{R}}\Psi^h_0(x_h)\Psi^h_{\gamma_h}(x_h)f_{X_h}(x_h)\mathrm{d}x_h\bigg)\\ &= \sum_{\gamma_2,\gamma_3=0}^{n} x_{0,\gamma_2,\gamma_3,0}(t_k)\Psi^2_{\gamma_2}(b)\Psi^3_{\gamma_3}(r), \end{aligned} \tag{34}$$

which gives the variance,

$$\mathbb{V}\bigg(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|b,r\big)\bigg) = \sum_{\gamma_2,\gamma_3=0}^{n}\big(x_{0,\gamma_2,\gamma_3,0}(t_k)\big)^2\mathbb{V}\big(\Psi^2_{\gamma_2}(b)\Psi^3_{\gamma_3}(r)\big) = \sum_{\gamma_2,\gamma_3=1}^{n}\big(x_{0,\gamma_2,\gamma_3,0}(t_k)\big)^2,$$

We then have the Sobol indice $S_{x_{b,r}}$:

$$S_{x_{b,r}}(t_k) = \frac{\mathbb{V}\bigg(\mathbb{E}\big(x(t_k,\sigma,b,r,s)\big|b,r\big)\bigg)}{\mathbb{V}\big(x(t_k,\sigma,b,r,s)\big)} = \frac{\sum_{\gamma_2,\gamma_3=1}^{n}\big(x_{0,\gamma_2,\gamma_3,0}(t_k)\big)^2}{\sum_{i=1}^{P}\big(x_i(t_k)\big)^2}. \tag{35}$$

The other Sobol indices of order 2 ($S_{x_{\sigma,b}}$, $S_{x_{\sigma,r}}$, $S_{x_{\sigma,s}}$, $S_{x_{b,s}}$ and $S_{x_{r,s}}$) are obtained in the same way.

4 Numerical application:

Global Sensitivity Analysis (GSA) is used to quantify the influence of uncertain variables in a mathematical model. Before performing the GSA, it is necessary to specify a probability distribution to model the uncertainty and possibly the statistical dependencies of the variables. Quantifying this distribution is difficult in practice because we have limited and imprecise knowledge of uncertain variables.

In this section, we consider the PC expansion to propagate the uncertainty in the performance measures of the LS model, due to the epistemic uncertainties in the input parameter of the model. We consider that the parameters σ, b, r, s follow the uniform distribution (in the stable case and the chaotic case). We calculate the Sobol indices of order 1 and 2 using the coefficients of the polynomials chaos as detailed in section 3 and finally we identify the most influential parameters.

4.1 Sobol indice in non-chaotic regime

In this subsection, we consider the stable case where there is no chaotic effect. We set the final time $T_f = 20$, the quadrature points $Nq = 8$, the degree of the polynomials chaos $n = 10$ and the

initial condition $p_0 = (1, 1, 1, 1)$. In Table 1, the 95% confidence intervals for the parameters are given.

Parameters	confidence interval at 95%
σ	[3.32 ; 3.68]
b	[2.37 ; 2.63]
r	[10.92 ; 12.08]
s	[0.95 ; 1.05]

Table 1: Parameters estimations.

Before starting the study of Sobol indices, we calculate the different statistical measures (expectation, variance, skewness and non-centered kurtosis). In Table 2, we give the different statistical measures (expectation, variance, skewness, kurtosis) which make it possible to describe a state variable in the stable case. We notice in the Figure 14 that the variance take finite value and therefore allows us to compute the Sobol indices.

Statistical quantities	x	y	z	v
Expectations	4.9054	5.3068	10.4181	-1.4027
Variances	0.0118	0.0138	0.1121	0.0024
Skewness$\times 10^4$	0.0118	0.0149	0.1131	-0.0003
Kurtosis$\times 10^5$	0.0058	0.0079	0.1178	0.0000

Table 2: Statistical quantities of the stationary regime for the non-chaotic case.

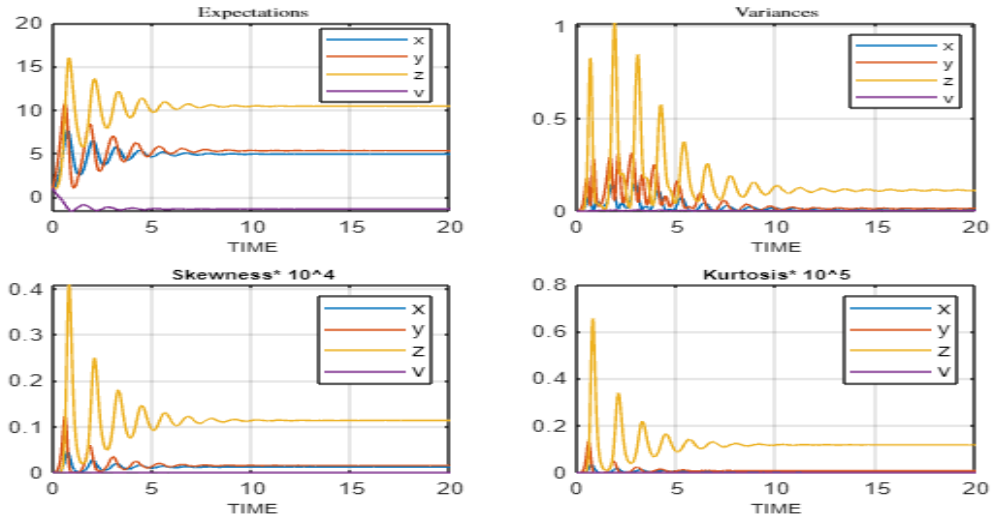

Figure 14: Statistical quantities of the transient regime for the non-chaotic case.

In the non-chaotic case, the factor that is often the most impacting is z which describes the temperature gradient. The factor z have the greatest intensity in all statistical quantities (expectation, variance, skewness, kurtosis), see Table 2 and Figure 14. In Table 2, we see that the expectation of the state variables converges towards the point of equilibrium $E_2 = (4.9055;\ 5.3065;\ 10.4187;\ -1.4028)$. The transient regime admits a stationary equilibrium that is reached at $T_f = 20$. The Sobol indices of the steady state for each parameters at that time is given in Table 3.

Sobol Indices	$w = x$	$w = y$	$w = z$	$w = v$
S_{w_σ}	0.0130	0.0087	0.0003	0.6029
S_{w_b}	0.4575	0.4603	0.0000	0.1841
S_{w_r}	0.5253	0.5272	0.9992	0.2113
S_{w_s}	0.0035	0.0023	0.0002	0.0013
Sum of indices	0.9973	0.9985	0.9997	0.9996

Table 3: Sobol indices at final time $T_f = 20$.

For the states x and y, see Figures 15 and 16, the two parameters that influence the most are b and r. The variable b influences x at 45% and 46% at y. The variable r influences x at 52% and y at 52%. For state z, the only parameter that influences 99 is r, see Figure 17. The parameters σ, b and r affect the state v at 60%, 18% and 21% respectively, see Figure 18. For the parameter s, see Figures 15-18, which represent the rotation there is no influence on the state, its influence on the stability of the system is negligible. The parameter r is the only one that influences all four states (x, y, z, v) with the greatest intensity on z at 99%.
The sum of the indices of order 1 being almost equal to 1 then the Sobol indices of order 2 also tell us that the combined effects on the state variables remain weak (see Table 3).

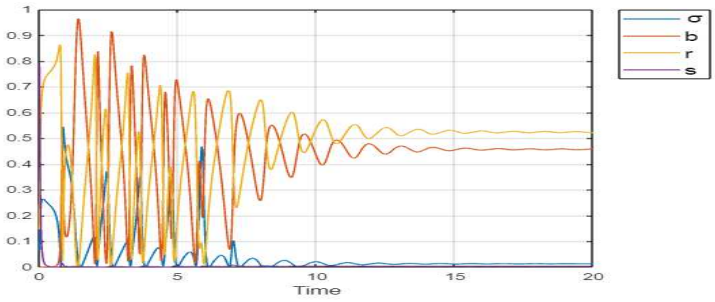

Figure 15: Influence all parameters on x in the non-chaotic case.

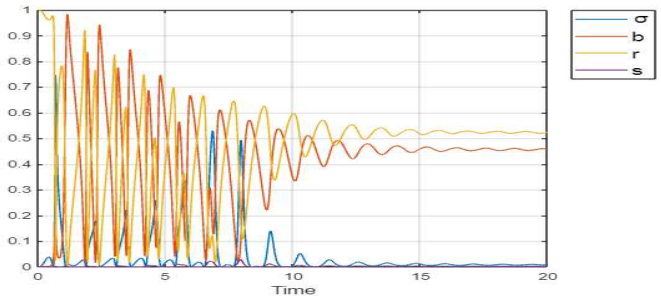

Figure 16: Influence all parameters on y in the non-chaotic case.

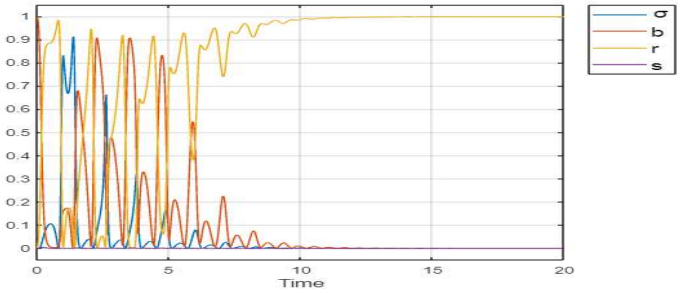

Figure 17: Influence all parameters on z in the the non-chaotic case.

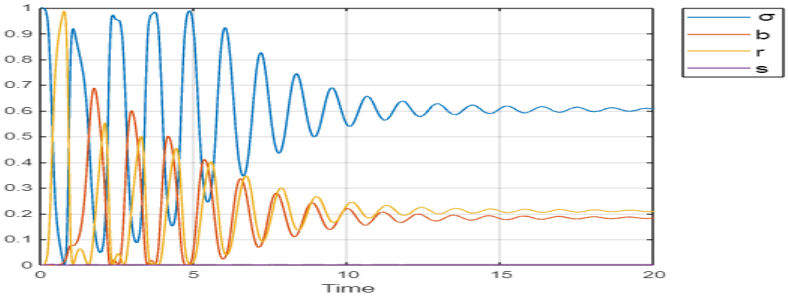

Figure 18: Influence all parameters on v in the the non-chaotic case.

4.2 Sobol indice in the chaotic regime

In this subsection, we consider the chaotic case. In Table 4, the 95% confidence intervals for the parameters are given. We set the different parameters $n = 10$, $T_f = 100$, $Nq = 8$ and $p_0 = (1, 1, 1, 1)$. In Table 4, the 95% confidence intervals for the parameters are given.

Parameters	confidence interval at 95%
σ	[0.47 ; 0.53]
b	[0.47 ; 0.53]
r	[17.57 ; 19.43]
s	[2.37 ; 2.63]

Table 4: Parameters estimations.

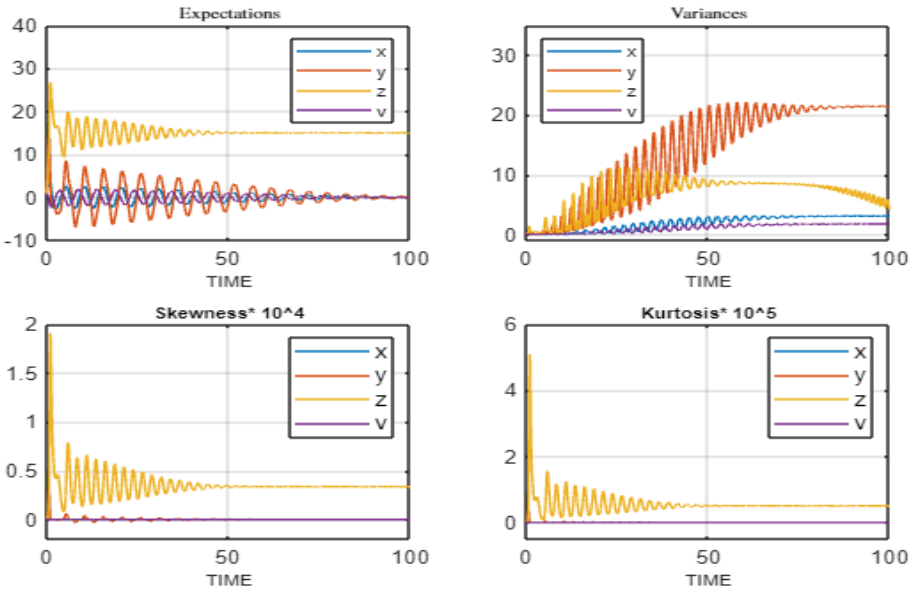

Figure 19: Statistical quantities of the transient regime for the chaotic case.

Figure 19 shows the expectation, variance, skewness and kurtosis for the four state a function of time for $t \in [0; 100]$. In contrast to the non-chaotic case, these statistical quantities do not always reach a steady-state.

Instead, we observe a pseudo-periodic regime that will also be present in Sobol's indices since Sobol's indices are obtained from a decomposition of the variance (see equation (11)).
This is confirmed in Figures 20-23 that show first order Sobol's indices for the four states. On top of the oscillations a downward trend can be seen. This indicates that higher order Sobol's indices should exhibit an opposit trend (since the sum of all the indices of any order should be equal to 1).

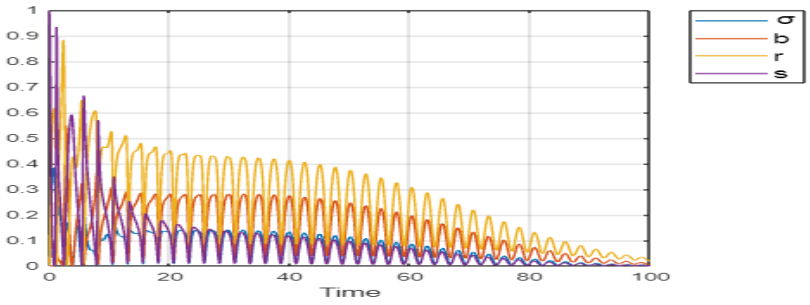

Figure 20: Influence all parameters on x in the chaotic case.

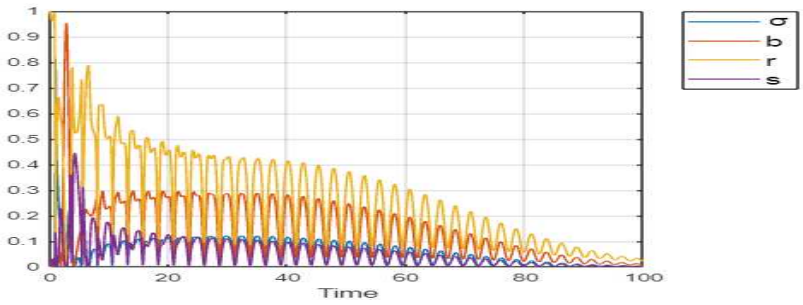

Figure 21: Influence all parameters on y in the chaotic case.

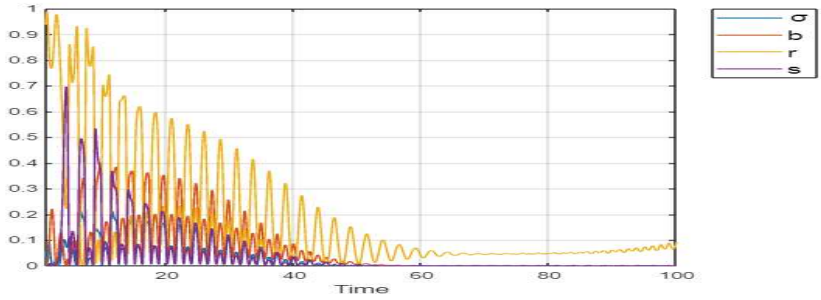

Figure 22: Influence all parameters on z in the chaotic case.

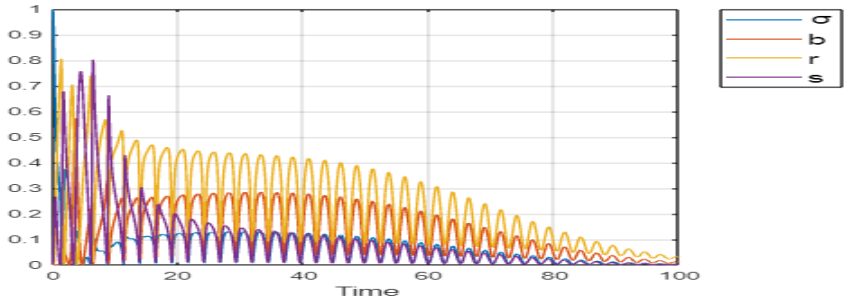

Figure 23: Influence all parameters on v in the chaotic case.

In other word, over time the combined effect of the parameters should increase. This trend is well observed in Figures 24-27, that show the second order Sobol's indices.

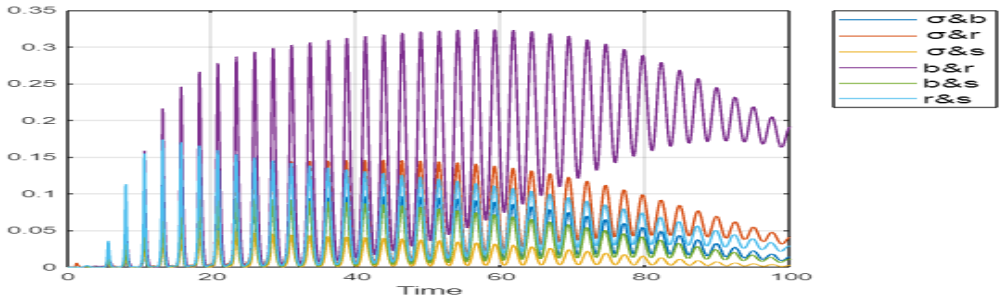

Figure 24: Combined influence all parameters on x in the chaotic case.

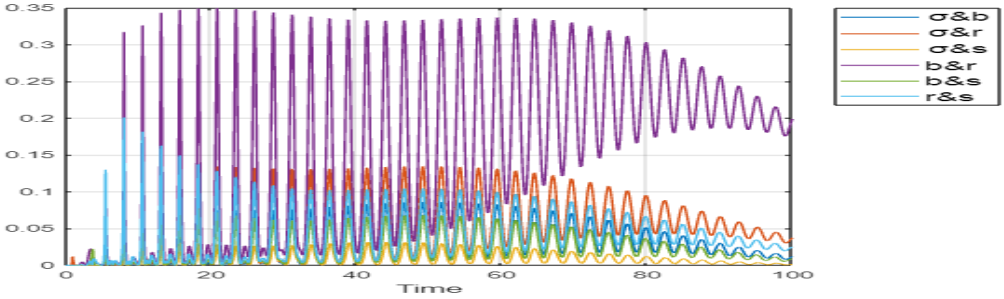

Figure 25: Combined influence all parameters on y in the chaotic case.

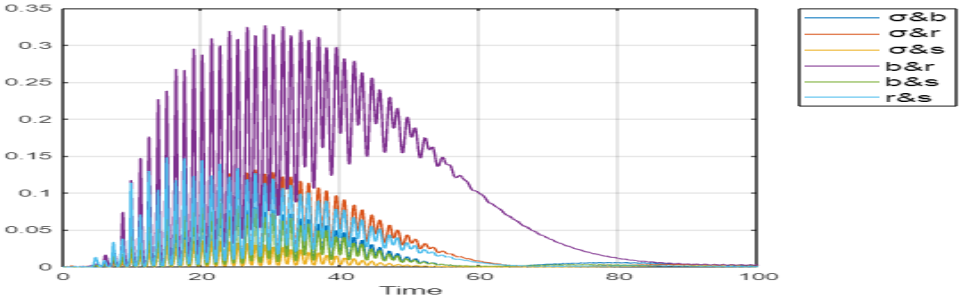

Figure 26: Combined influence all parameters on z in the chaotic case.

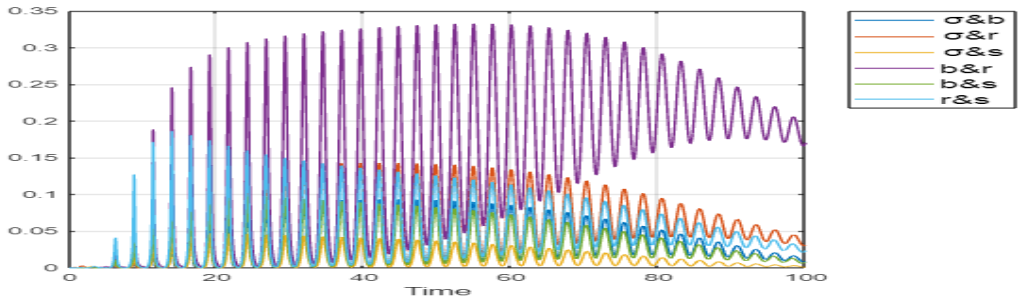

Figure 27: Combined influence all parameters on v in the chaotic case.

For $t \in [0; 20]$, Sobol's indices exhibit complex patterns this is why in Figures 28-31 and 32-35, they have been plotted for that range of time.

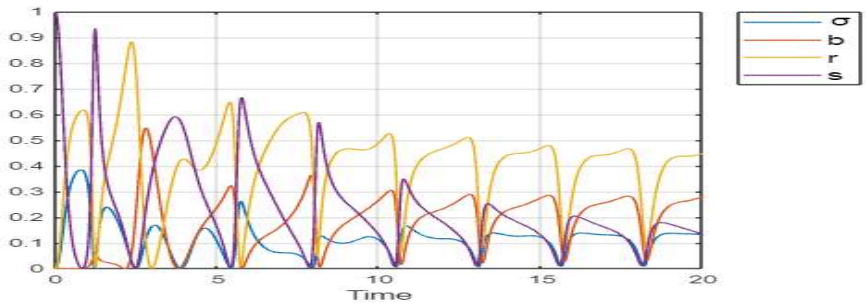

Figure 28: Influence all parameters on x in the chaotic case on $[0; 20]$.

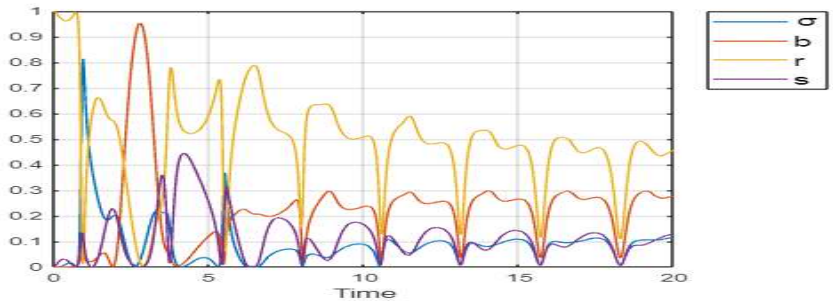

Figure 29: Influence all parameters on y in the chaotic case on $[0; 20]$.

There again we can observe complex pseudo-periodic profiles with a downward trend for the first order indices and an upward trend for the second order indices. Another noticeable observation is the amplitude of the oscillations that damper over time.

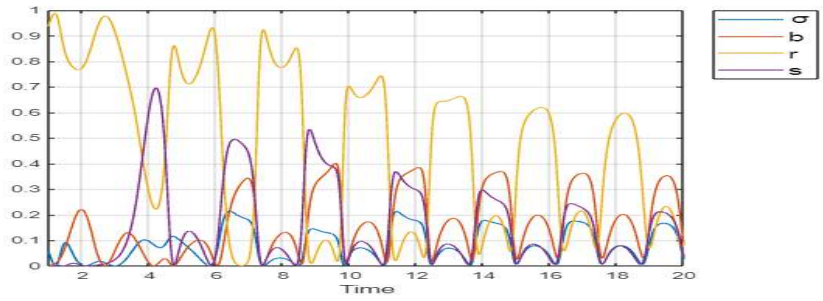

Figure 30: Influence all parameters on z in the chaotic case on $[0; 20]$.

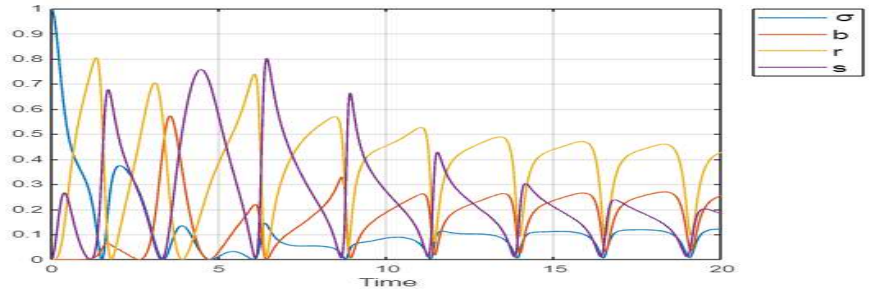

Figure 31: Influence all parameters on v in the chaotic case on $[0; 20]$.

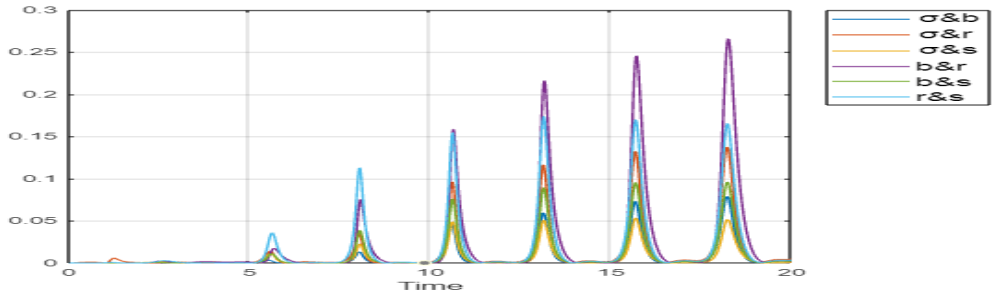

Figure 32: Combined influence all parameters on x in the chaotic case on $[0; 20]$.

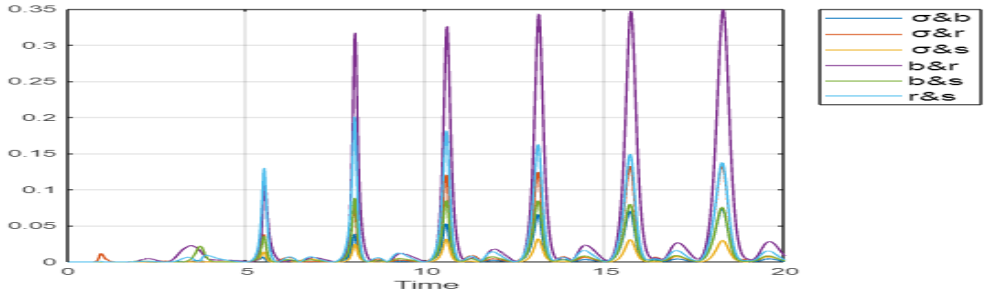

Figure 33: Combined influence all parameters on y in the chaotic case on $[0; 20]$.

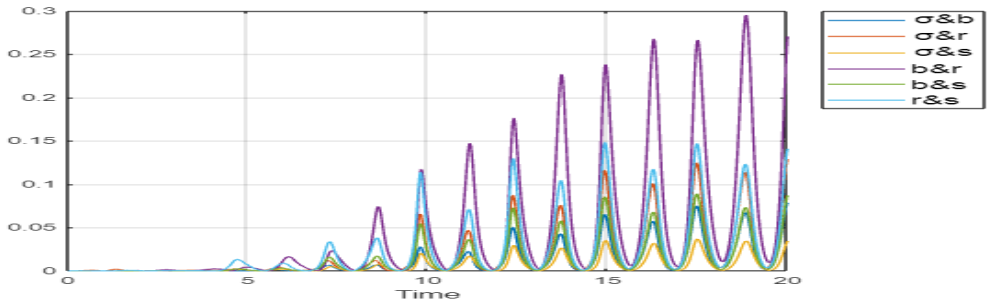

Figure 34: Combined influence all parameters on z in the chaotic case on $[0; 20]$.

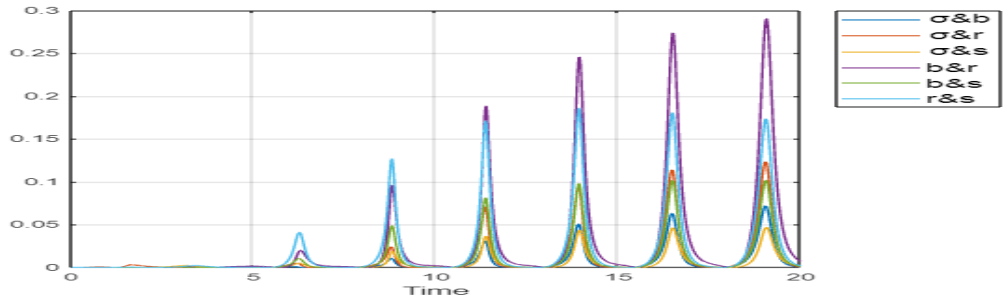

Figure 35: Combined influence all parameters on v in the chaotic case on $[0; 20]$.

5 Conclusion

When qualitative estimates of sensitivity are desired, a mathematical model of the phenomena is desirable. However, such a model poses questions of stability, optimality and sensitivity. In this research work, we have developed a numerical approach based on chaos polynomials to calculate Sobol indices in order to study the sensitivity of the factors that intervene in the climate system LS (σ; b; r; s). We observed in the calculation of the Sobol indices, in the chaotic case, that the climate system is extremely sensitive to the slightest change, regardless of the parameters that intervene in the climate system (chemical properties of the atmosphere, rotation, movement of convection, temperature gradient) and also the parameters involved in the method undertaken to model the climate system (the final time, the degree of polynomials, the number of quadrature points, the initial condition of the start). The chaotic effect seems to disturb the statistical measurements.
This paper presents a new framework in which the robustness of the Sobol indices with respect to the distribution of the input variables was evaluated. The proposed method makes it possible to carry out such an analysis and numerical examples have been presented to show its effectiveness. Chaos polynomials can be extended to more complex stochastic systems governed by partial differential equations without any fundamental difficulty.
We also illustrated the sensitivity to the initial condition using the Euclidean distance and we found that the distance evolves without converging in the chaotic case. Because of this sensitivity involved in the modeling of the climate system (the natural parameters and the parameters that intervene in the digital part) it is difficult to prevent or predict climate change.

References

[1] M. A. Akinlar. A new method for parameter sensitivity analysis of lorenz equations. *Math. Probl. Eng.*, 2013.

[2] K. Bachi, C. Chauvière, H. Djellout, and K. Abbas. Propagation of epistemic uncertainty in queueing models with unreliable server using chaos expansions. *Comm. Statist. Simulation Comput.*, 50(4):1019–1041, 2021.

[3] G. Dahlquist and A. Björck. *Numerical methods in scientific computing. Vol. I.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[4] P. J. Davis and P. Rabinowitz. *Methods of numerical integration*. Dover Publications, Inc., Mineola, NY, 2007.

[5] B. Debusschere. *Intrusive polynomial chaos methods for forward uncertainty propagation*. Springer, Cham, 2017.

[6] B. Efron and C. Stein. The jackknife estimate of variance. *Ann. Statist.*, 9(3):586–596, 1981.

[7] W. Gautschi. *Orthogonal polynomials: computation and approximation*. Num. Math. Sci. Comput. Oxford University Press, N.Y., 2004.

[8] R. G. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, N.Y., 1991.

[9] G. H. Golub and J. H. Welsch. Calculation of Gauss quadrature rules. *Math. Comp.*, 23:A1–A10, 1969.

[10] J. Hart and P. Gremaud. Robustness of the sobol indices to distribution uncertainty. *Int. J. Uncertain. Quantif.*, 9(5):453–469, 2019.

[11] S. Hosder, R. Walters, and R. Perez. A non-intrusive polynomial chaos method for uncertainty propagation in CFD simulations. *A.I.A.A.*, pages 1–20, 2006.

[12] N. V. Kuznetsov, T. N. Mokaev, and P. A. Vasilyev. Numerical justification of Leonov conjecture on Lyapunov dimension of Rossler attractor. *Commun. Nonlinear Sci. Numer. Simul.*, 19(4):1027–1034, 2014.

[13] D. J. Lea, T. Haine, M. R. Allen, and J. A. Hansen. Sensitivity analysis of the climate of a chaotic ocean circulation model. *Q. J. R. Meteorol*, 128(586):2587–2605, 2002.

[14] G. A. Leonov. Generalized Lorenz equations for acoustic-gravity waves in the atmosphere. Attractors dimension, convergence and homoclinic trajectories. *Commun. Pure Appl. Anal.*, 16(6):2253–2267, 2017.

[15] G. A. Leonov and N. V. Kuznetsov. Hidden attractors in dynamical systems. From hidden oscillations in Hilbert-Kolmogorov, Aizerman, and Kalman problems to hidden chaotic attractor in Chua circuits. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 23(1):1330002, 69, 2013.

[16] G.A. Leonov. Bounds for attractors and the existence of homoclinic orbits in the lorenz system. *J. Appl. Math. Mech.*, 65(1):19–32, 2001.

[17] J. Li and D. Xiu. A generalized polynomial chaos based ensemble Kalman filter with high accuracy. *J. Comput. Phys.*, 228(15):5454–5469, 2009.

[18] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmospheric Sci.*, 20(2):130–141, 1963.

[19] C. Marzban. Variance-based sensitivity analysis: An illustration on the lorenz'63 model. *Mon. Weather Rev.*, 141(11):4069–4079, 2013.

[20] A. E. Matouk. Chaos synchronization of a fractional-order modified van der Pol-Duffing system via new linear control, backstepping control and Takagi-Sugeno fuzzy approaches. *Complexity*, 21(S1):116–124, 2016.

[21] G. J. McRae, J. W. Tilden, and J. H. Seinfeld. Global sensitivity analysis : A computational implementation of the fourier amplitude sensitivity test (fast). *Comput. Chem. Eng.*, 6(1):15–25, 1982.

[22] S. Moon, J. M. Seo, B.S. Han, J. Park, and J.J. Baik. A physically extended Lorenz system. *Chaos*, 29(6):063129, 12, 2019.

[23] J. Park, B.S. Han, H. Lee, Y.L. Jeon, and J.J. Baik. Stability and periodicity of high-order lorenz–stenflo equations. *Phys. Scr.*, 91(6):065202, 2016.

[24] M. P. Pettersson, G. Iaccarino, and J. Nordström. *Polynomial chaos methods for hyperbolic partial differential equations*. Math. Eng. Springer, Cham, 2015.

[25] S. Rahman. Extended polynomial dimensional decomposition for arbitrary probability distributions. *J. Eng. Mech.*, 135(12):1439–1451, 2009.

[26] P. C. Rech. On the dynamics of a modified Lorenz-Stenflo system. *Internat. J. Modern Phys. C*, 31(7):2050104, 7, 2020.

[27] C. Y. Shen, T. E. Evans, and S. Finette. Polynomial chaos quantification of the growth of uncertainty investigated with a lorenz model. *J. Atmos. Ocean. Technol.*, 27(6):1059–1071, 2010.

[28] I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Math. Comput. Simulation*, 55(1-3):271–280, 2001.

[29] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Syst. Saf.*, 93(7):964–979, 2008.

[30] T. Turányi. Sensitivity analysis of complex kinetic systems. Tools and applications. *J. Math. Chem.*, 5(3):203–248, 1990.

[31] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[32] Marcelo Viana. What's new on Lorenz strange attractors? *Math. Intelligencer*, 22(3):6–19, 2000.

[33] N. Wiener. The homogeneous chaos. *Am. J. Math.*, 60(4):897–936, 1938.

[34] D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.

[35] D. Xiu and G. E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phys.*, 187(1):137–167, 2003.

[36] F. Zhang and G. Zhang. Further results on ultimate bound on the trajectories of the Lorenz system. *Qual. Theory Dyn. Syst.*, 15(1):221–235, 2016.

[37] C. Zhou, C. H. Lai, and M. Y. Yu. Bifurcation behavior of the generalized Lorenz equations at large rotation numbers. *J. Math. Phys.*, 38(10):5225–5239, 1997.

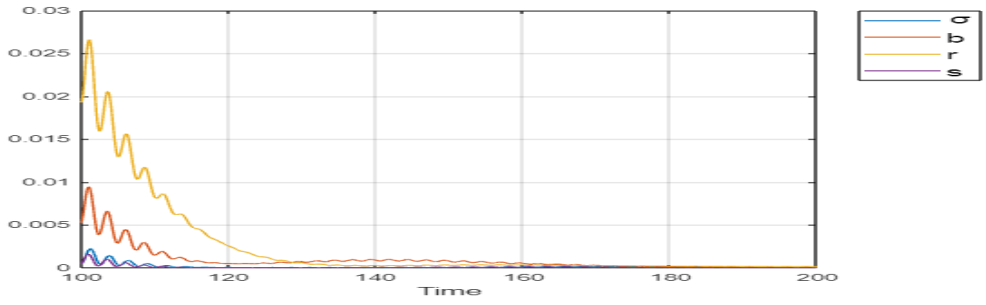

Figure 36: Combined influence all parameters on x in the chaotic case.

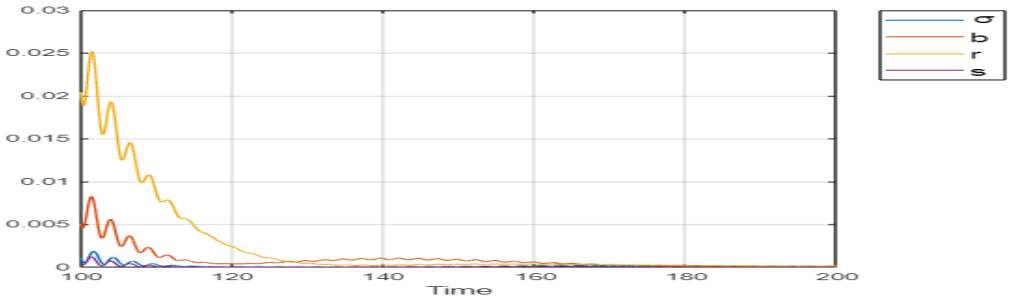

Figure 37: Combined influence all parameters on y in the chaotic case.

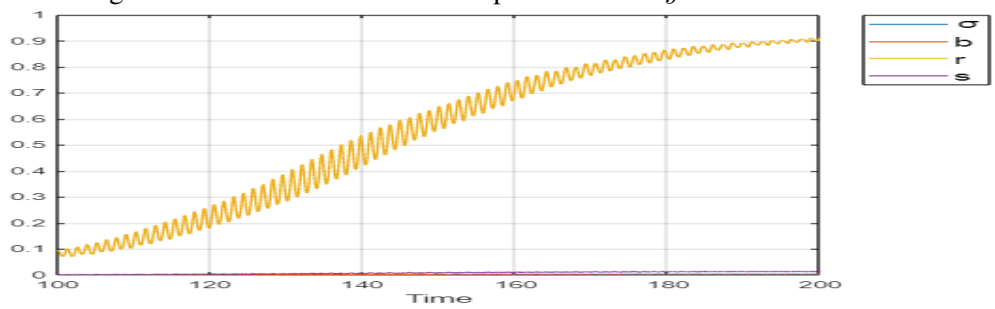

Figure 38: Combined influence all parameters on z in the chaotic case.

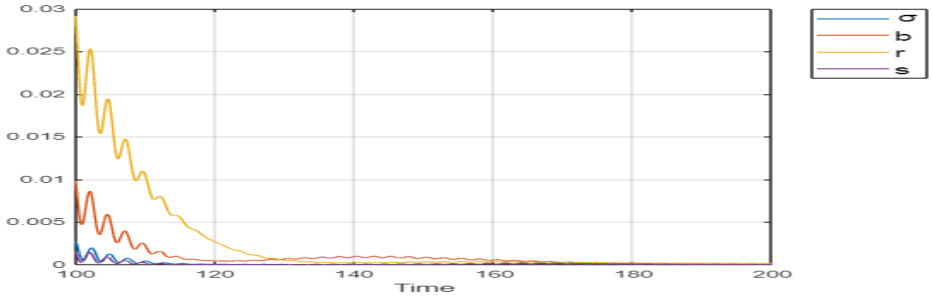

Figure 39: Combined influence all parameters on v in the chaotic case.

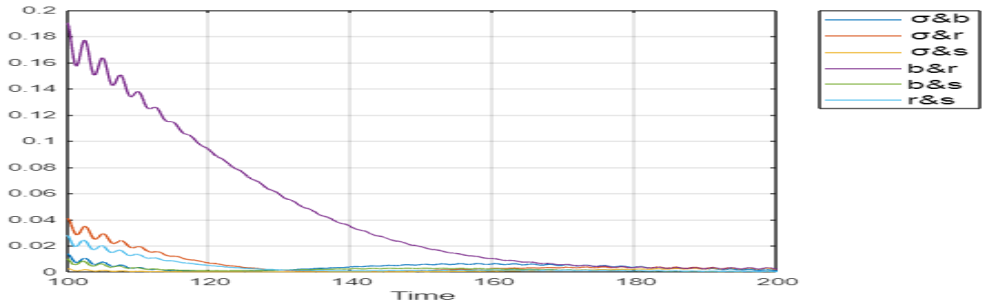

Figure 40: Combined influence all parameters on x in the chaotic case.

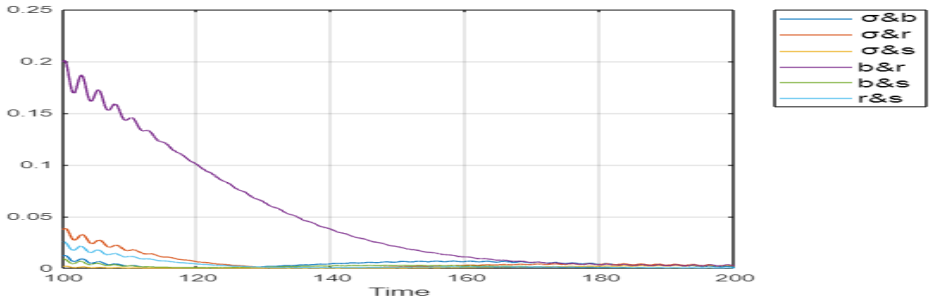

Figure 41: Combined influence all parameters on y in the chaotic case.

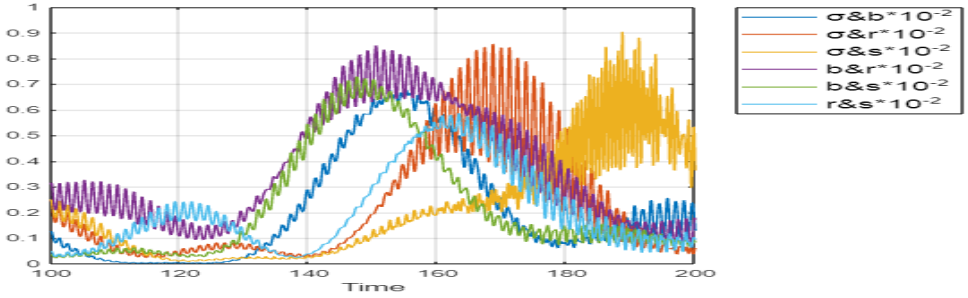

Figure 42: Combined influence all parameters on z in the chaotic case.

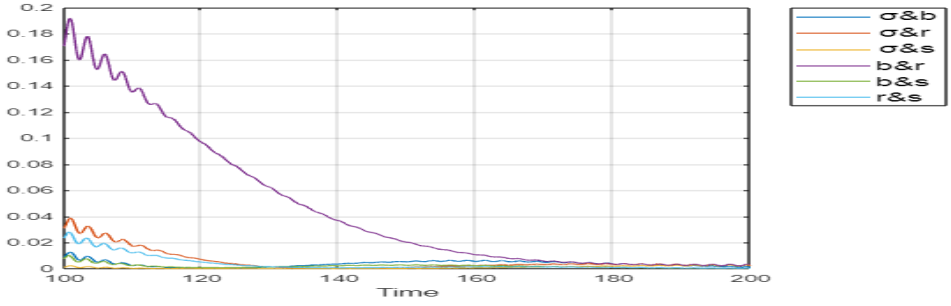

Figure 43: Combined influence all parameters on v in the chaotic case.

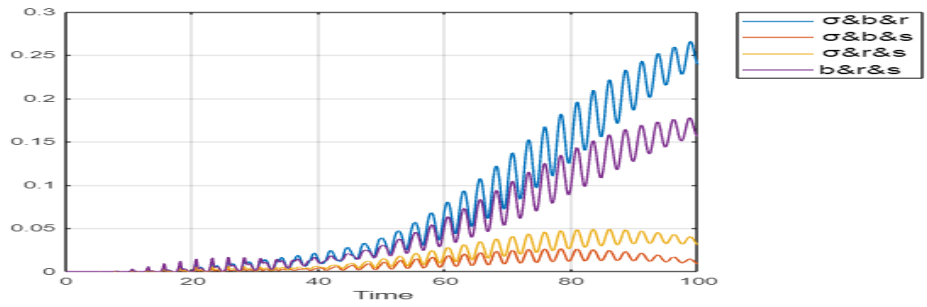

Figure 44: Combined influence all parameters on x in the chaotic case.

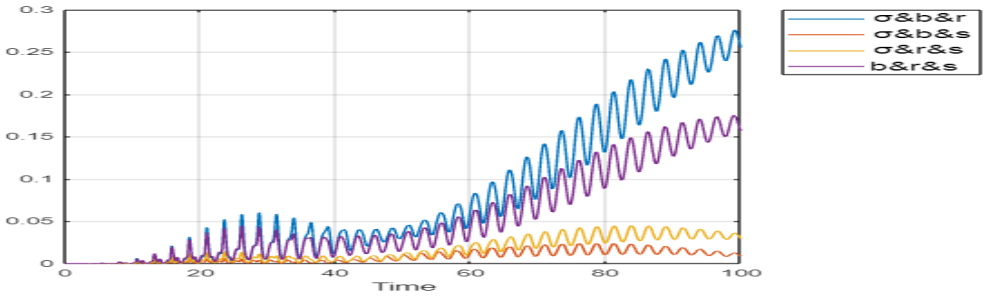

Figure 45: Combined influence all parameters on y in the chaotic case.

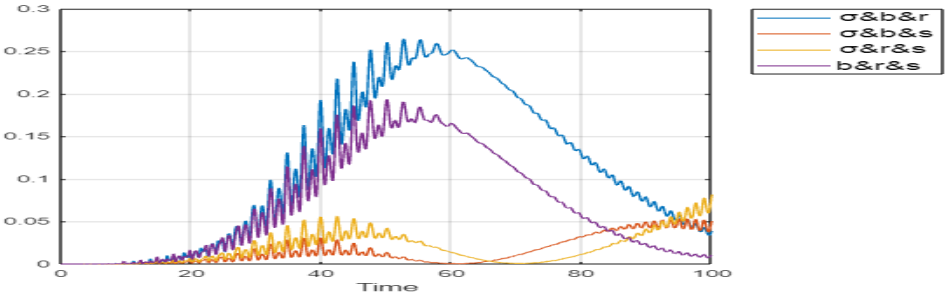

Figure 46: Combined influence all parameters on z in the chaotic case.

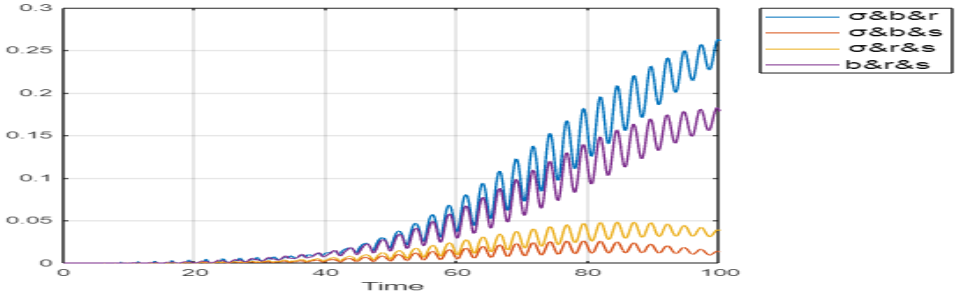

Figure 47: Combined influence all parameters on v in the chaotic case.