CODEE Journal

Volume 16 Article 4

10-10-2023

# Exploring Parameter Sensitivity Analysis in Mathematical Modeling with Ordinary Differential Equations

Viktoria Savatorova
*Central Connecticut State University*



## Recommended Citation

Savatorova, Viktoria (2023) "Exploring Parameter Sensitivity Analysis in Mathematical Modeling with Ordinary Differential Equations," *CODEE Journal*: Vol. 16, Article 4.
Available at: https://scholarship.claremont.edu/codee/vol16/iss1/4

# Exploring Parameter Sensitivity Analysis in Mathematical Modeling with Ordinary Differential Equations

## Cover Page Footnote (optional)

The author was supported by the 2023–2024 CSU-AAUP Faculty Research Grant.

# *Exploring Parameter Sensitivity Analysis in Mathematical Modeling with Ordinary Differential Equations*

Viktoria Savatorova
*Central Connecticut State University*

**Keywords:** parameter sensitivity, local and global sensitivity analysis, direct differential method, Pearson correlation coefficient, Spearman correlation coefficient, PRCC, Sobol indices, visualization techniques, fish population dynamics, harvesting.
Manuscript received on July 4, 2023; published on October 10, 2023.

**Abstract**: This paper presents an exploration into parameter sensitivity analysis in mathematical modeling using ordinary differential equations (ODEs). Taking the first steps in understanding local sensitivity analysis through the direct differential method and global sensitivity analysis using metrics like Pearson, Spearman, PRCC, and Sobol', we provide readers with a basic understanding of parameter sensitivity analysis for mathematical modeling using ODEs. As an illustrative application, the system of differential equations modeling population dynamics of several fish species with harvest considerations is utilized. The results of employing local and global sensitivity analysis are compared, shedding light on the strengths and limitations of each approach. The paper serves as a starting point for readers interested in exploring parameter sensitivity in their mathematical models.

## 1 Introduction

Mathematical modeling serves as a powerful tool for understanding and predicting the behavior of complex systems. Ordinary differential equations (ODEs) are widely employed to describe dynamic systems across various disciplines, ranging from physics, engineering, biology and ecology to sociology and economics ([1], [2], [3], [4]). The development of dynamic models requires accurate information on both the initial conditions and the system's parameter values. However, in many instances, the parameters characterizing the system are accompanied by uncertainty. This uncertainty arises due to limited data availability or challenges associated with experimental measurements of these parameters, which can be difficult or even infeasible. Furthermore, certain parameter values may exhibit large variations across different experimental or environmental conditions. As a result, our confidence in the model predictions is limited due to uncertainties and variations of model parameters.

Sensitivity analysis of model parameters plays a crucial role in the development of mathematical models, as it enables us to understand how fluctuations in the model outputs can be attributed to variations in the model inputs [5], [6]. This insight guides further research to reduce output variability. Sensitivity analysis also streamlines model reduction by identifying and removing parameters that have minimal impact on the model outputs. Furthermore, sensitivity analysis identifies parameters strongly correlated with the output capturing information on directional impact of input changes. This correlation information deepens our understanding of the driving factors of the model behavior. Sensitivity analysis is also important in identifying parameters that significantly contribute to output uncertainty, sometimes with asymmetric effects. For example, researches may confidently establish a precise lower bound on a specific output, while the upper bound remains more uncertain. Sensitivity analysis provides insights into the consequences that arise from altering specific input parameters, allowing for informed decision-making and scenario analysis. According to [7], sensitivity analysis is becoming an essential discipline for systems modeling and policy support, and it is now considered a requirement for good modeling practice.

In addition to its pivotal role in mathematical modeling, sensitivity analysis also serves as a powerful educational tool. It actively engages students in model exploration and hypothesis testing, enhancing their problem-solving skills by teaching them to identify crucial variables and make evidence-based decisions. This approach provides students with a practical context for applying mathematical concepts learned in calculus, differential equations, and statistics classes. Furthermore, it fosters interdisciplinary connections and encourages the application of mathematics to address real-world problems, equipping students with valuable skills for future career opportunities.

This paper aims to serve as an introductory guide for readers interested in exploring parameter sensitivity in mathematical modeling with ordinary differential equations (ODEs). While our intention is not to go into depth of the theoretical aspects or provide a comprehensive overview of available methods of sensitivity analysis, we will outline the key principles of both local and global sensitivity analysis approaches and introduce several well-established techniques to our readers. Our goal is to equip readers with the necessary tools to begin their own parameter sensitivity investigations. For those seeking a more in-depth understanding we suggest referring to such resources as [5], [6], [7].

To illustrate the methods of sensitivity analysis, we selected the Lotka-Volterra model as a representative example. We chose this model for its simplicity and its description through a system of $2 \times 2$ nonlinear differential equations. Also we were motivated by the work of Panayotova et al. [4], who applied the local sensitivity method to a similar system. In our study, we build upon their research by incorporating global sensitivity analysis methods and conducting a comparative analysis of the outcomes. Below, we present an application of parameter sensitivity analysis by examining a nonlinear system of differential equations that captures the population dynamics of two fish species, including the effects of harvesting. The model and parameter values utilized are adopted from the study by Panayotova et al. [4].

Let us consider the model describing the population dynamics of the prey, Atlantic menhaden, $x = x(t)$ and a the predator, striped bass, $y = y(t)$. In the system below, it is assumed that both the prey and the predator are harvested based on the catch per effort hypothesis.

$$\begin{cases} \frac{dx}{dt} = x(r - ax - by) - q_1 E_1 x \\ \frac{dy}{dt} = y(-e - cy + dx) - q_2 E_2 y \end{cases} \tag{1.1}$$

In (1.1), $r$ is the intrinsic growth rate of the prey, $e$ is the death rate of the predator in absence of prey, $a$ and $c$ are self-limitation parameters for the prey and the predator respectfully, $b$ is the predator effect of $y$ on $x$, and $d$ is the effect of prey consumption of $x$ on $y$.

| Parameter | r | a | b | e | c | d |
|---|---|---|---|---|---|---|
| Value | 0.513 | 0.026 | 1.765 | 1.213 | 0.520 | 9.999 |

Table 1: Table of the model's parameter values obtained from data fitting (see [4])

.

The values of parameters $r, a, b, e, c, d$ are given in Table 1. Parameters $E_1$ and $E_2$ are the harvesting efforts of the Atlantic menhaden and stripped bass respectively. Parameters $q_1$ and $q_2$ represent the catchability coefficients of each species. The range of catchability parameters is $[0, 1]$ with lower values used when the fish is relatively easy to catch, and higher values when the fish is more rare or difficult to catch. In computations, we assume $q_1 = 0.3$ and $q_2 = 0.9$. Parameters for the harvesting efforts were chosen to be $E_1 = 0.1$ and $E_2 = 0.1$. Chosing the initial conditions $x(0) = 0.16$ and $y(0) = 0.31$ billions, we get time-dependence of of quantities $x$ and $y$ presented in Figure 1.

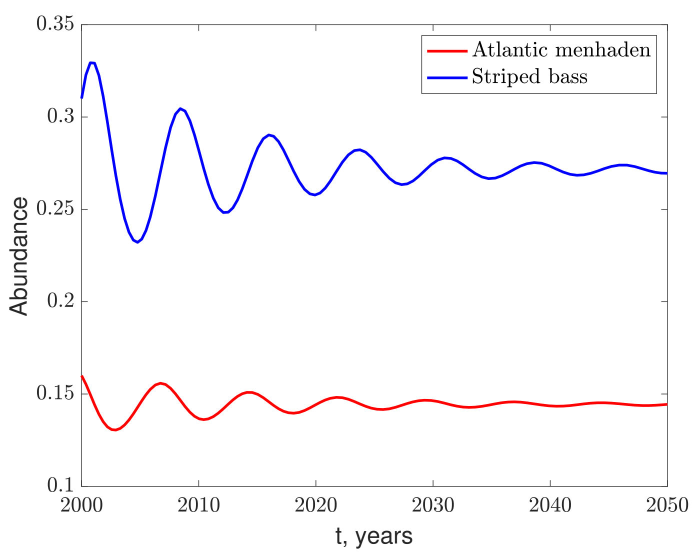



Figure 1: Population dynamics (in billions) up to 2050.

In the subsequent sections, we discuss the local and global sensitivity analysis methods and provide a comparative analysis of the results obtained from the fish population dynamics model. The structure of the paper is as follows: In Section 2, we provide an overview of the direct differential method. This local sensitivity analysis method examines individual parameters while holding all others constant, offering insights into parameter ranking and sensitivity indices. In Section 3, we explore global sensitivity analysis metrics including the Pearson correlation coefficient [8], Spearman correlation coefficient [9], Partial Rank Correlation Coefficient (PRCC) [10], and Sobol indices [11]. These methods enable the evaluation of the overall sensitivity of model outputs to multiple parameters simultaneously [6]. By comparing the outcomes of local and global sensitivity analyses, we can identify the respective strengths and limitations of each approach and gain deeper insights into how the model responds to parameter variations. Finally, we summarize the obtained results and provide readers with recommendations on using sensitivity analysis in their modeling endeavors. To assist instructors in incorporating parameter sensitivity analysis into their teaching, suggested tasks were included in the Appendix.

## 2 Local sensitivity analysis

Below we introduce the direct differential method, one of the local methods of parameter sensitivity analysis. This method is referred to as "local" because it examines how small changes in parameter can affect the model's behavior in the immediate vicinity of a particular parameter value.

First, we revisit equation (1.1) and express it in vector form

$$\frac{dZ}{dt} = \boldsymbol{F}(\boldsymbol{Z}; \boldsymbol{p}), \quad \boldsymbol{p} = [r; a; b; e; c; d; E_1; E_2], \tag{2.1}$$

where $\boldsymbol{Z} = [x(t); y(t)]$ is two-component vector solutions of the system (1.1), and all parameters are represented as components of a vector $\boldsymbol{p}$. In the work by Panayotova et al. ([4]), parameters $q_1$ and $q_2$ were allowed to vary. For simplicity, we considered $q_1$ and $q_2$ to be constants and ignored their variability given our limited control over their values. In contrast, parameters $E_1$ and $E_2$ offers us the flexibility to manipulate and control harvesting. It worth noticing that if significant uncertainty were to arise regarding the values of parameters $q_1$ and $q_2$, we would parametrize to account for relative catchability. Consequently, our model is defined by a total of eight parameters: $\boldsymbol{p} = [r; a; b; e; c; d; E_1; E_2]$. We define sensitivity vector $\boldsymbol{s}_j$ for parameter $p_j$ as $\boldsymbol{s}_j = [s_{1,j}; s_{2,j}] = \left[\frac{\partial x}{\partial p_j}; \frac{\partial y}{\partial p_j}\right]$, where $j = 1, 2, ..., 8$. Each of its components describes how a variation in $p_j$ affects either $x$ or $y$. These sensitivity vectors form the matrix of sensitivity measures

$$\boldsymbol{S} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} & \frac{\partial x}{\partial e} & \frac{\partial x}{\partial c} & \frac{\partial x}{\partial d} & \frac{\partial x}{\partial E_1} & \frac{\partial x}{\partial E_2} \\ \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} & \frac{\partial y}{\partial e} & \frac{\partial y}{\partial c} & \frac{\partial y}{\partial d} & \frac{\partial y}{\partial E_1} & \frac{\partial y}{\partial E_2} \end{bmatrix}. \tag{2.2}$$

If the exact solutions for $x$ and $y$ would be known, we could find each of the components of the matrix $\boldsymbol{S}$ explicitly, and it would answer our questions regarding parameter sensitivity. Since it is not the case, we consider the time evolution of each of the sensitivity measures

$$\frac{ds_{i,j}}{dt} = \frac{d}{dt}\left(\frac{\partial Z_i}{\partial p_j}\right) = \frac{\partial}{\partial p_j}\left(\frac{dZ_i}{dt}\right) = \frac{\partial F_i}{\partial x} \cdot \frac{\partial x}{\partial p_j} + \frac{\partial F_i}{\partial y} \cdot \frac{\partial y}{\partial p_j} + \frac{\partial F_i}{\partial p_j}; \quad i = 1, 2; \; j = 1, 2, ..., 8.$$

Written in matrix form, the equation above will look like

$$\frac{dS}{dt} = \boldsymbol{J}\boldsymbol{S} + \boldsymbol{f}, \tag{2.3}$$

where $\boldsymbol{J}$ is the Jacobian defined by formula

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{bmatrix} = \begin{bmatrix} r - 2ax - by - q_1 E_1 & -bx \\ dy & -e - 2cy + dx - q_2 E_2 \end{bmatrix}, \tag{2.4}$$

and $\boldsymbol{f}$ is the matrix formed by taking the derivatives of the right-hand sides of the equations in the system (1.1) with respect to the parameters:

$$\boldsymbol{f} = \begin{bmatrix} x & -x^2 & -xy & 0 & 0 & 0 & -q_1 x & 0 \\ 0 & 0 & 0 & -y & -y^2 & xy & 0 & -q_2 y \end{bmatrix}. \tag{2.5}$$

System of differential equations (2.3) defines how the sensitivity measures change in time.

Now we can formulate the initial value problem consisting of differential equations (2.1) and (2.3) together with their initial conditions

$$\begin{cases} \frac{dZ}{dt} = \boldsymbol{F}(Z; \boldsymbol{p}), \\ \frac{dS}{dt} = \boldsymbol{J}\boldsymbol{S} + \boldsymbol{f}, \\ Z(0) = \boldsymbol{Z_0}, S(0) = \boldsymbol{S_0}. \end{cases} \tag{2.6}$$

It worth mentioning here that if no parameters appear in the initial condition of the state equations, then $\boldsymbol{S_0} = \boldsymbol{0}$ ([5]). Now we have a direct method for calculating sensitivity measures. For example, the initial value problem that investigates the rate of change of $x$ and $y$ with respect to parameter $b$ has the following form

$$\begin{cases} \frac{dx}{dt} = x(r - ax - by) - q_1 E_1 x \\ \frac{dy}{dt} = y(-e - cy + dx) - q_2 E_2 y \\ \frac{ds_{13}}{dt} = (r - 2ax - by - q_1 E_1)s_{13} - bxs_{23} - xy \\ \frac{ds_{23}}{dt} = dys_{13} - (e + 2cy - dx + q_2 E_2)s_{23} \\ x(0) = x_0, \quad y(0) = y_0 \\ s_{13}(0) = s_{23}(0) = 0. \end{cases} \tag{2.7}$$

Notice that the first two differential equations are our original differential equations (1.1) while the next two equations come from (2.3) using the Jacobian and the third column of the matrix $f$ since $b$ is the third of our eight parameters. Similar to (2.7) we can formulate initial value problems defining the rate of change of dependent variables with respect to any other parameter. To account for the different units of measurement and varying magnitudes of parameters, we introduce the dimensionless relative sensitivity $S_{ij}$, which is defined as $S_{ij} = \frac{\partial Z_i}{\partial p_j} \frac{p_j}{Z_i}$. For instance, we have $S_{13} = \frac{\partial x}{\partial b} \frac{b}{x}$ and $S_{23} = \frac{\partial y}{\partial b} \frac{b}{y}$. To capture the time-dependent nature of sensitivity measures, we define the sensitivity index $||S_{ij}||_2$ as the magnitude of the corresponding norm for each relative sensitivity:

$$||S_{ij}||_2 = \sqrt{\sum_{k=1}^{N} \left( \frac{\partial x_i}{\partial p_j}(t_k) \frac{p_j}{x_i(t_k)} \right)^2}. \tag{2.8}$$

For example, $||S_{13}||_2 = \left\| \frac{\partial x}{\partial b} \frac{b}{x} \right\|_2 = \sqrt{\sum_{k=1}^{N} \left( \frac{\partial x}{\partial b}(t_k) \frac{b}{x(t_k)} \right)^2}$, etc.

Computed values of relative sensitivities $S_{ij} = \frac{\partial Z_i}{\partial p_j} \frac{p_j}{Z_i}$ $(i = 1, 2; j = 1, ..., 8)$ and their associated sensitivity indices $||S_{ij}||_2$ are presented in Table 2. Computations were performed over a time period spanning from 2000 to 2080, with a duration of $t_f = 80$ years. Relative sensitivities provided in columns 2 and 4 of Table 2, exhibit very good agreement with similar results reported by Panayotova et al. [4]. Minor discrepancies can be attributed to the differences in the time range considered, with Panayotova et al. focusing on the years from 2050 to 2100, and it is noteworthy that we have provided the values for the relative sensitivities themselves, while Panayotova et al. presented their absolute values.

Table 2: Relative sensitivities & relative sensitivity indices

| Parameter $p$ | $\frac{\partial x}{\partial p} \frac{p}{x}$, ($t_f = 80$ years) | $\left\| \frac{\partial x}{\partial p} \frac{p}{x} \right\|_2$ | $\frac{\partial y}{\partial p} \frac{p}{y}$, ($t_f = 80$ years) | $\left\| \frac{\partial y}{\partial p} \frac{p}{y} \right\|_2$ |
|---|---|---|---|---|
| $r$ | 0.1176 | 3.0598 | 1.0647 | 18.1783 |
| $a$ | -0.0008 | 0.0202 | - 0.0077 | 0.1323 |
| $b$ | -0.1000 | 2.9838 | -1.0048 | 16.2097 |
| $e$ | 0.8502 | 15.4344 | - 0.0136 | 6.6165 |
| $c$ | 0.0999 | 1.8899 | -0.0014 | 1.1786 |
| $d$ | -1.0019 | 17.6650 | 0.0055 | 6.4200 |
| $E_1$ | -0.0069 | 0.1789 | -0.0623 | 1.0618 |
| $E_2$ | 0.0631 | 1.1466 | -0.0010 | 0.4910 |

Analyzing the results shown in Table 2, it becomes evident that the predator (striped bass) exhibits the highest sensitivity to variations in the intrinsic growth rate of the prey (the Atlantic menhaden) denoted as $r$, as well as changes in the predation rate $b$. On the other hand, the prey demonstrates greater sensitivity to alterations in the effect of its consumption on the predator $d$ and variations in the death rate of the predator $e$.

The direct differential method offers a computationally efficient approach to sensitivity analysis. It assesses local sensitivities, allowing researchers to gain detailed insights into how small parameter perturbations impact model outputs. The method is adaptable to different types of models and can be tailored to specific research questions or model characteristics. The local nature of the analysis allows for a more intuitive understanding of parameter sensitivities and their implications for model behavior.

At the same time, it is important to note that the direct differential method has limitations. For instance, it assumes that the model is differentiable and continuous, which holds true for our system but may not apply for more complicated models. Additionally, being a method of local sensitivity analysis, it primarily focuses on understanding how the model's behavior changes near a specific parameter configuration. Consequently, it may not reveal how the model behaves under different combinations of parameters. Local analysis may overlook phenomena like parameter interactions and other effects that occur when multiple parameters vary simultaneously. To achieve a more comprehensive understanding of how variations in multiple input parameters affect the model's output, we will proceed with global sensitivity analysis. As we will explore in the next section, global sensitivity analysis assesses the model's response to variations in all parameters across their entire range. Through systematic perturbations of multiple parameters, we gain insights into behavior of models exhibiting nonlinear relationships and intricate parameter dependencies.

## 3 Basic ideas of Global sensitivity analysis

This section introduces several methods of global sensitivity analysis and demonstrates their practical application through a specific case study discussed in Section 2.

Let us denote the quantity of interest (QoI) as $z$ and consider a model with $m$ parameters $p_1, p_2, ..., p_m$. In our case study, the QoI can be either abundance of the prey $x$ or abundance of the predator $y$. Referring to equation 2.1, the number of parameters in our example is $m = 8$.

Global sensitivity analysis involves several typical steps:

1. **Define the parameter ranges and distributions**. This step initiates global sensitivity analysis by defining parameter ranges based on existing knowledge or experimental data. Selecting an appropriate probability distribution for each parameter is also important. A uniform distribution is common in the absence of prior data, while a normal distribution is suitable for capturing bell-shaped distributions, symmetric behavior, or known constrains on parameter values.

2. **Generate a sample space of parameter sets**. In the next step, we create a sample space of parameter sets. When dealing with large models, traditional analytic techniques, such as linear stability analysis, often prove impractical due to complexity of interactions and nonlinear relationships among parameters. While linearization and other approximations can shed light on individual parameter effects, understanding their collective impact often demands a probabilistic approach. Random sampling is preferable to systematic parameter variation since it eliminates potential bias influenced by modeler's prior expectations and ensures equal exploration of all parametric regions. Selecting the sample size depends on factors like model complexity, desired accuracy, and computational resources, aiming for a balance between model intricacy and computational efficiency. Two common techniques for generating random parameter sets include Monte Carlo sampling and Latin hypercube sampling (LHS) [5]. In Monte Carlo sampling, we randomly assign values to our model's parameters based on their probability distributions. We do this repeatedly to create a collection of different parameter sets. This random exploration helps us understand how our model behaves under various conditions. LHS, introduced by McKay et al. [15] , is a Monte Carlo-based sampling method that divides the parameter ranges into equal intervals, ensuring that each interval contributes a sample. The procedure

for LHS sampling, which selects $n$ different values for each of the $m$ parameters $p_1, p_2, ..., p_m$, can be summarized as follows: First, the range $[p_i^{min}, p_i^{max}]$ of each parameter $p_i$ (where $i = 1, 2, ..., m$) is divided into $n$ equal intervals. Then, a value is randomly selected from each interval based on the probability density within the interval. The $n$ values obtained for $p_1$ are randomly paired with $n$ values of $p_2$. These $n$ pairs are combined randomly with the $m$ values of $p_3$ to form $n$ triplets, and so on, until an $(n \times m)$ matrix $P$ is obtained. The matrix $P$ consists of $n$ rows for the number of samples and $m$ columns for the number of parameters. The advantage of LHS is its efficiency. Unlike standard Monte Carlo sampling, which may lead to over-sampling in some regions and under-sampling in others, LHS ensures more even and efficient exploration of the parameter space. This approach results in a better understanding of the model's behavior with fewer samples, saving time and computational resources. LHS implementation is supported by functions available in different programming languages. For instance, in MATLAB, the *lhsdesign* and *lhsnorm* functions can be utilized for uniform and normal LHS sampling respectively (`https://www.mathworks.com/help/stats/lhsdesign.html`).

3. **Simulate the mathematical model for each parameter set to collect the corresponding output data**. In this step, $n$ model solutions will be simulated for each combination of parameter values, i.e., each row of the matrix $P$.

4. **Apply global sensitivity analysis method**. Collected data will be used to apply global sensitivity analysis method and quantify the sensitivity of the model outputs to variations in the parameters. The selection of a method is guided by considerations such as the type of input and output variables, the assumed relationships among variables and data availability.

5. **Interpret the results** of global sensitivity analysis to identify significant parameters and understand their impacts on the model outputs.

Below, we provide several examples that demonstrate practical application of global sensitivity analysis methods and rank system parameters based on the obtained results.

## 3.1 Pearson correlation coefficients

Before proceeding further, let us establish some statistical definitions that will be useful in the context of global sensitivity analysis. These definitions include variance, covariance, and standard deviation. Let $p$ be a random variable representing a quantity that can take on $n$ values $p_1, p_2, ..., p_n$, governed by a probability density function. The mean of $p$ is denoted as $\bar{p}$.

The variance of a random variable quantifies the measure of the spread of the values of the random variable around the mean. In statistical analysis, there exists a distinction between biased and unbiased variance estimation. Biased variance is computed using the formula $Var_b(p) = \frac{1}{n}\sum_{j=1}^{n}(p_j - \bar{p})^2$, where $n$ signifies the number of elements in the sample. However, when working with data samples, it is preferable to use the unbiased variance defined as $Var(p) = \frac{1}{n-1}\sum_{j=1}^{n}(p_j - \bar{p})^2$. The preference arises from the need to avoid the possibility of underestimation (see [13], [14] for more details).

The covariance between two random variables $p$ and $z$ measures the degree to which the variables vary together. Mathematically, biased covariance is defined as $Cov_b(p,z) = \frac{1}{n}\sum_{j=1}^{n}(p_j - \bar{p})(z_j - \bar{z})$, where $p_j$ and $z_j$ represent the individual observations of variables $p$ and $z$ respectively, and $\bar{p}$ and $\bar{z}$ are their respective means. However, when working with samples of data, the use of unbiased covariance, defined as $Cov(p,z) = \frac{1}{n-1}\sum_{j=1}^{n}(p_j - \bar{p})(z_j - \bar{z})$, is advisable to avoid potential underestimation.

The standard deviation of a random variable $p$, denoted as $\sigma_p$, is the square root of its variance: $\sigma_p = \sqrt{Var(p)}$.

In the context of global sensitivity analysis, Pearson's coefficients, commonly referred to as Pearson correlation coefficients, allow us to assess the strength and direction of the relationship between model inputs and outputs [5]. Mathematically, Pearson's coefficient $\rho$ is defined as the covariance between two variables $p$ and $z$ divided by the product of their standard deviations: $\rho = \frac{Cov(p,z)}{\sigma_p \sigma_z}$. The coefficient $\rho$ takes values between -1 and +1, where -1 indicates a perfect negative linear relationship, +1 indicates a perfect positive linear relationship, and 0 suggests no linear relationship between the variables. A positive value of $\rho$ indicates that as one variable increases, the other tends to increase, while a negative value indicates an inverse relationship.

In our case study, we are interested in the correlation between the quantity of interest $z$ and each of the $m$ parameters $p_1, p_2, ..., p_m$. So, we define the Pearson correlation coefficients as

$$\rho_{p_{i,z}} = \frac{\sum_{j=1}^{n}(p_{ij} - \bar{p}_i)(z_j - \bar{z})}{\sqrt{\sum_{j=1}^{n}(p_{ij} - \bar{p}_i)^2 \sum_{j=1}^{n}(z_j - \bar{z})^2}}, \quad i = 1, 2, ..., m \tag{3.1}$$

In formula (3.1) $\bar{p}_i$ and $\bar{z}$ represent the mean of $p_i$ and $z$ respectively. The value of $\rho_{p_{i,z}}$ can change between -1 and 1.

Below, we will compute and compare Pearson correlation coefficients for each parameter in system (1.1) to rank their importance and understand their influence on the system's behavior. Various software packages offer the functionality to compute Pearson correlation coefficients. For instance, MATLAB provides a function called *corrcoef* that can be used for this purpose (`https://www.mathworks.com/help/matlab/ref/corrcoef.html`). The $p$-value associated with each correlation coefficient indicates whether the observed correlation is statistically significant. A low $p$-value, typically below 0.05, indicates a strong correlation and provides evidence to reject the null hypothesis of no correlation between the variables. For a more in-depth understanding of hypothesis testing and $p$-values, interested readers are encouraged to explore relevant statistical textbooks, such as [13] or [14] .

As mentioned earlier, one should start with generating a sample space of parameter sets. For our analysis, we employ Latin hypercube sampling, which ensures that each parameter $p_i (i = 1, 2, ..., 8)$ is sampled within an interval of $\pm 5\%$ of its nominal value provided in Table 1. This interval is defined as $[0.95p_i, 1.05p_i]$. Since we lack information about the probability distribution of the parameters, we opt for uniform sampling. We conducted sensitivity analyses with sample sizes of $n = 50$ and $n = 150$ for this simple model. The results did not show noticeable variation across these different sample sizes.

Let us assume that the quantity of interest (QoI) is the predicted population of the Atlantic menhaden $x$ in 2050 ($t_f = 50$years) and compare the Pearson correlation coefficients for parameters $r, a, b, e, c, d, E_1, E_2$. Notice that catchabilities $q_1$ and $q_2$ we assume to be fixed.

Analyzing the results presented in Figure 2, we find that the most crucial parameters for the Atlantic menhaden are $e$ and $d$. These findings are supported by small $p$-values associated with correlation coefficients which indicate statistical significance. Specifically, the abundance of the prey, the Atlantic menhaden, and the death rate $e$ of its predator exhibit a positive correlation. Additionally, there exists a negative correlation between the abundance of the prey and parameter $d$, which is responsible for the effect of the prey consumption on the predator. The effect of variations of all other parameters does not seem to be very noticeable.

Do the results shown in Figure 2 imply that the other parameters are less significant? It is likely, but we should approach this with caution and consider the limitations of the Pearson correlation coefficient method.

While Pearson's coefficients provide insightful information, it is crucial to recognize their limitations when applied to global sensitivity analysis. It is important to note that Pearson's coefficients only capture linear relationships and may not accurately represent non-linear dependencies between input parameters and the quantity of interest (QoI). To address this limitation, examining scatterplots of each parameter against the QoI can provide additional valuable information.

A scatter plot is a graphical representation that allows us to visually observe patterns, trends, and potential non-linear associations between two variables by plotting their data points in a coordinate system.The horizontal axis represents the values of the input parameter $p_i (i = 1, 2, ..., m = 8)$, and the vertical axis represents the corresponding values of the QoI.

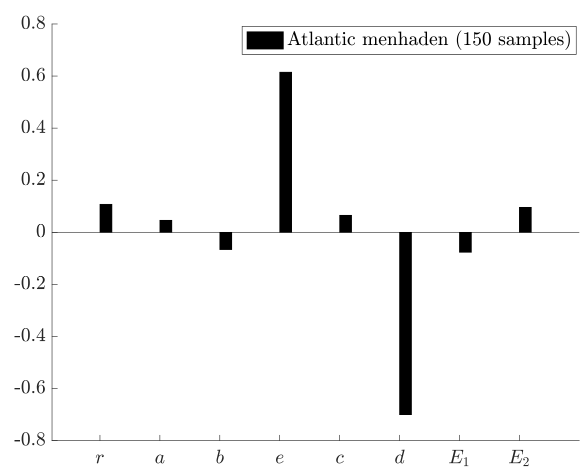



Figure 2: Pearson's sensitivity coefficients for the Atlantic menhaden (n=150 samples).

Figure 3 exhibits scatterplots for parameters $e$, $d$, $r$, $E_1$, and $E_2$, revealing a noticeable correlation between the QoI and parameters $e$ and $d$. At the same time, the remaining parameters exhibit significant variations with no clear correlation pattern. The scatterplots for parameters $a$, $b$, and $c$ were not included, as they would also demonstrate substantial variations without any clear pattern.

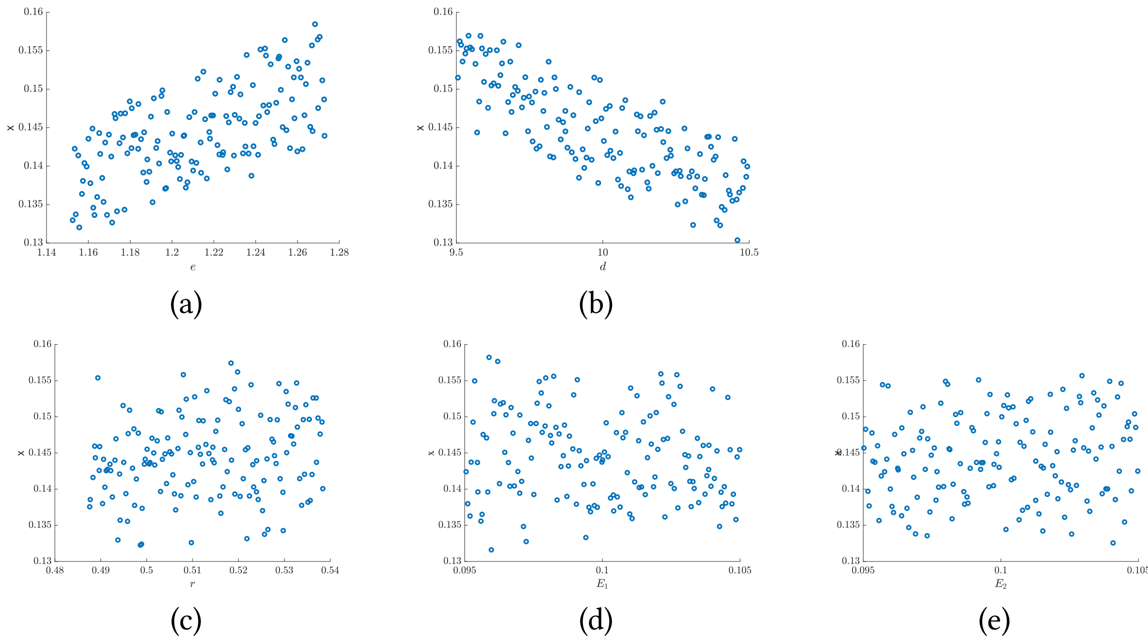



Figure 3: Scatterplots for parameters $e$, $d$, $r$, $E_1$ and $E_2$ against the QoI which is the value of $x$ in 2050.

Now let us change the quantity of interest. This time it will be the predicted population of striped bass $y$ in 2050. We again compute and compare the Pearson's sensitivity coefficients for parameters $r, a, b, e, c, d, E_1, E_2$ keeping catchabilities $q_1$ and $q_2$ unchanged. From the results presented in Figure 4, it follows that the most important parameters for striped bass are $r$ and $b$. These conclusions are further supported by the computation of p-values associated with correlation coefficients, which attest to their statistical significance ($p$-values are less than 0.05). One can see that the abundance of the predator, striped bass, and the growth rate $e$ of its prey are positively correlated, and there exists a negative correlation between the abundance of the predator and parameter $b$ responsible for the effect of the predator on the prey. The effect of variations of other parameters seems to be less pronounced but we still want to examine the corresponding scatterplots before making judgements.

Figure 5 displays scatterplots for parameters $r$, $b$, $d$, $E_1$, and $E_2$, against the QoI which is now the value of $y$ at $t_f = 50$ years. It reveals a clear correlation between the QoI and parameters $r$ and $b$. At the same time, the remaining parameters exhibit significant variations with no clear correlation pattern. The scatterplots for parameters $a$, $c$, and $e$ were not included, as they would also demonstrate substantial variations without noticeable pattern.

Interpreting Pearson's correlation becomes challenging in the presence of nonlinear

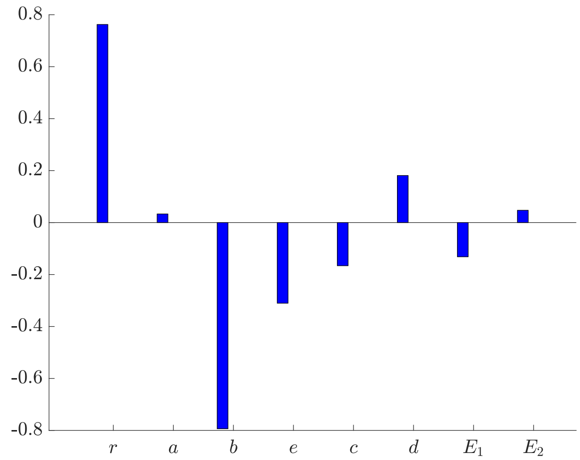



Figure 4: Pearson's sensitivity coefficients for striped bass (n=50 samples).

relationships between the parameter and the QoI due to its linear-regressive nature. In such cases, the slope lacks meaningful interpretation. To overcome this limitation, a rank transform can be employed by replacing actual values with ordinal rankings. This transformation allows for the observation of linear relationships as long as the relationships remain monotonic. In the following section we will discuss Spearman correlation that addresses the need for an alternative correlation measure in the context of rank-transformed data.

## 3.2 Spearman correlation

Interpreting Pearson correlation coefficients becomes challenging with non-linear parameter-model relationships. To address this, rank transformation is used to reduce the effects of nonlinearity. This method works well when the dependence between parameters and model outcomes is monotonic.

Let us explore the idea of rank transformation. Rank transformation is a technique used to convert the actual values of a dataset into their corresponding ranks. Let's say we have a dataset with numerical values. The rank transformation involves ordering the values from smallest to largest and assigning each value its ordinal rank. The smallest value receives a rank of 1, the second smallest value receives a rank of 2, and so on. If there are ties in the data (i.e., multiple values with the same magnitude), the average rank is assigned to those tied values. For example, if we had a dataset $\{10, 5, 8, 8, 3\}$, after applying the rank transformation, he first value, 3, receives a rank of 1, the second value, 5, receives a rank of 2, the third value, 8, and the fourth value, 8, both receive an average rank of (3+4)/2 = 3.5, and the largest value, 10, receives a rank of 5.

If the input/output relationships are monotonic, then rank transformations of the input and output values (i.e. replacing the values with their ranks) results in linear relationships and the rank coefficient indicates the degree of monotonicity between the input and output values. Thus, the Spearman correlation coefficient quantifies the strength and direction of

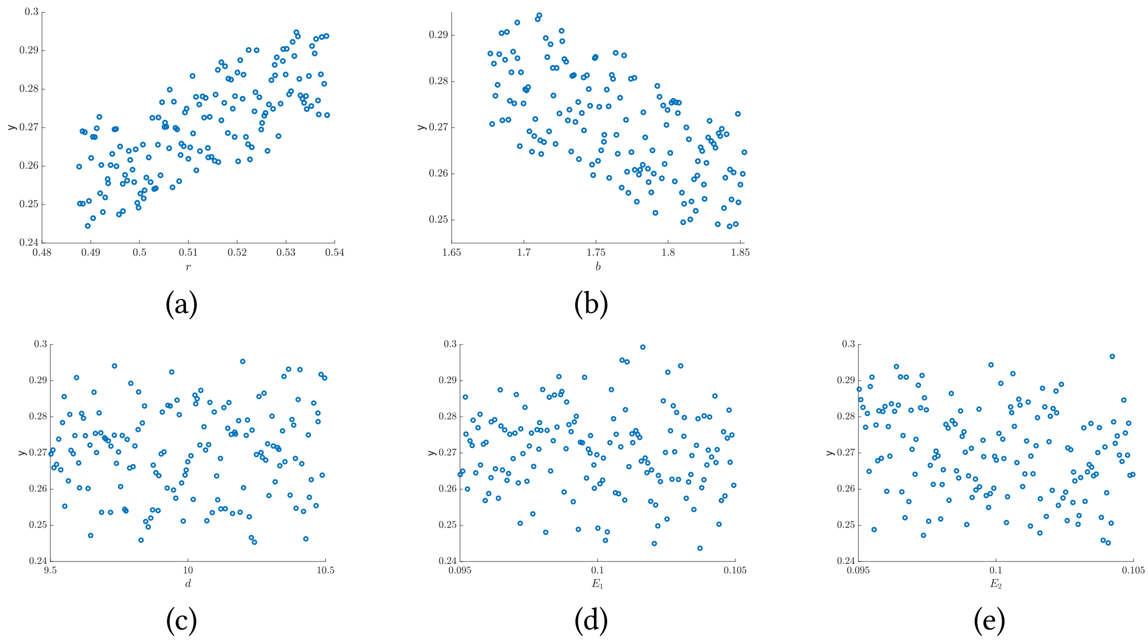



Figure 5: Scatterplots for parameters $r$, $b$, $d$, $E_1$ and $E_2$ against the QoI which is the value of $y$ in 2050.

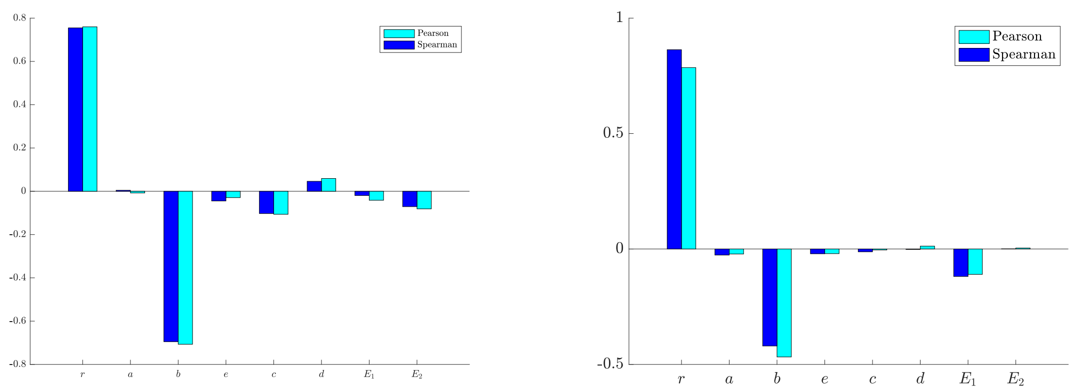
the monotonic relationship between the parameter values and the corresponding model outputs or quality-of-interest (QoI) measures. Unlike Pearson's correlation, which assumes linearity, the Spearman correlation is based on the ranks of the data rather than the actual values. It assesses the extent to which the variables tend to change together in a consistent, monotonic manner, regardless of the specific functional form of the relationship. Positive Spearman correlation indicates that higher parameter values tend to be associated with higher QoI values, while negative Spearman correlation indicates an inverse relationship.

Spearman's rank correlation coefficient can be calculated using the equation similar to equation (3.1) with the exception of operating on the rank transformed data. Several software packages offer the functionality to compute Spearman correlation coefficients. For example, MATLAB's *corr* command can compute the Spearman correlation coefficient by specifying the option *type*, *Spearman* (`https://www.mathworks.com/help/stats/corr.html`). Spearman correlation coefficients, similar to Pearson coefficients are associated with $p$-values. Low $p$-values (usually below 0.05) signify a strong correlation and offers evidence to reject the null hypothesis of no correlation between the variables.

In the context of our case study, let's consider the task of determining the parameters that have the greatest impact on the quantity of interest (QoI), which is defined as the population of the striped bass $y$ at time $t_f = 50$ years. We have a total of eight parameters to rank, resulting in an 8-dimensional sample space. To sample these parameters, we employ a Latin hypercube centered at the nominal parameter values provided in Table 1, sampling each parameter $p_i (i = 1, 2, ..., 8)$ within the interval $[0.95p_i, 1.05p_i]$.

Given the unknown probability distribution of the parameters, we explore both uniform and normal sampling approaches and compare their results for potential differences. We conducted calculations using sample sizes of $n = 150$ and $n = 4000$, with a larger sample size used for the case of normal parameter distribution. The results for the uniform probability density function are depicted in Figure 6(a) and Figure 7(a), while the corresponding results for the normal probability density function are shown in Figure 6(b) and 7(b). Notably, both sets of results exhibit good qualitative agreement, indicating that the most important parameters remain consistent regardless of the chosen probability density function.

In Figure 6 we provide the comparison of Pearson and Spearman correlation for the case of the uniform and normal distribution of the parameters. The results show qualitative agreement between the two methods. It is consistently observed that parameters $e$ and $d$ have the highest impact on prey abundance, whereas parameters $r$ and $b$ exhibit the greatest influence on predator abundance. Although the correlation coefficients may vary, the signs remain consistent, with negative correlations indicating an inverse relationship between the quantity of interest (QoI) and a parameter, and positive correlations indicating a positive relationship.




(a) Uniform distribution of the parameters

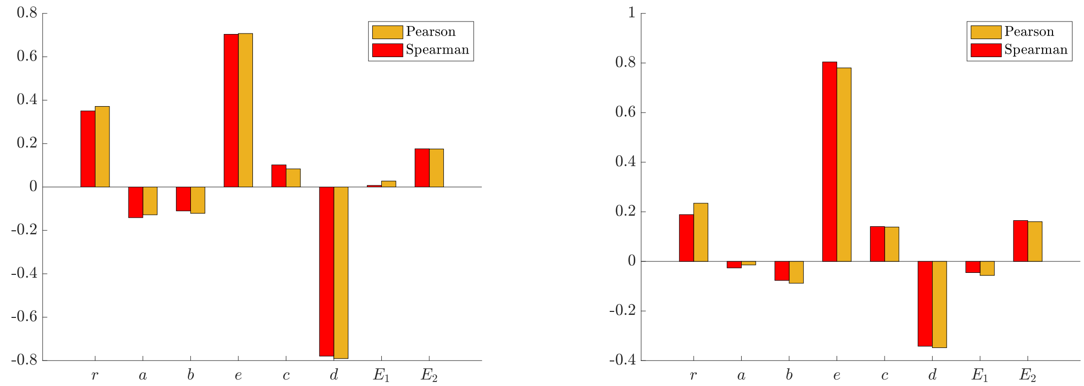
(b) Normal parameters' distribution

Figure 6: Comparison of Pearson and Spearman correlation (the QoI is the value of $y$ in 2050).

Figure 7 illustrates the computation and comparison of Pearson and Spearman correlation coefficients, with the QoI being the population of the Atlantic menhaden $x$ at time $t_f = 50$ years. The results demonstrate qualitative similarity between the two methods.

Notice that while the quantitative values of the correlation coefficients may vary between the uniform and normal distributions of the parameters, we consistently observe that parameters $e$ and $d$ remain the most influential for prey abundance, while parameters $r$ and $b$ are the most influential for predator abundance. Importantly, the signs of the correlation coefficients remain consistent, with negative correlations continuing to indicate an inverse relationship between the QoI and a parameter and positive correlations indicating a positive relationship, especially for the more influential parameters.

It is important to emphasize that monotonicity plays a significant role in interpreting correlation coefficients. For rank transformation to yield a linear-like relationship, the original data should exhibit monotonicity.




(a) Uniform distribution of the parameters (b) Normal parameters' distribution

Figure 7: Comparison of Pearson and Spearman correlation (the QoI is the value of $x$ in 2050).

## 3.3 Partial Rank Correlation Coefficient (PRCC)

While, Pearson's correlation coefficient captures linear or near-linear relationships between a parameter and the QoI, Spearman's correlation coefficient incorporates rank transformation to handle nonlinearity in monotonous parameter-QoI relationships, but it also considers the variability in the QoI influenced by other parameters. To address this limitation, Kendall ([16]) introduced partial correlation coefficients, which effectively discount the influences of other parameters and focus solely on the relationship between a single parameter and the QoI. Note that MATLAB has a function called *partialcorr* that can be used to compute partial correlation coefficients (`https://www.mathworks.com/help/stats/partialcorr.html`).

Here we introduce the Partial Rank Correlation Coefficient (PRCC) as a robust sensitivity analysis method that combines ranked correlation and partial correlation.

Let us consider parameters $p_1, p_2, \ldots, p_n$ as input and $z$, the QoI, as the output of our mathematical model. For example, given parameters $p_1$ and $p_2$ as input and the question of interest $z$ as the output, a partial rank correlation coefficient $PRCC_{p_1}$ is a measure of the correlation between $p_1$ and $z$, while eliminating indirect correlations due to relationships that may exist between $p_1$ and $p_2$ or $p_2$ and $z$.

The $PRCC_{p_1}$ is defined as

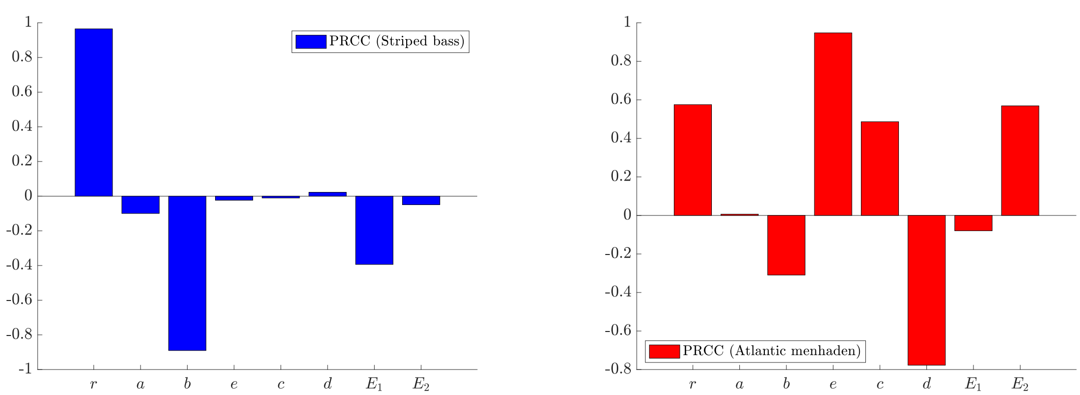
$$PRCC_{p_1} = \frac{\rho_{p_1,z} - \rho_{p_1,p_2}\rho_{p_2,z}}{\sqrt{(1-\rho^2_{p_1,p_2})(1-\rho^2_{p_2,z})}} \tag{3.2}$$

In general, to discount all the interaction we subtract the sum of all covariances between $p_1$ and $p_j$, $j \neq 1$. Notice that the parameters in formula (3.2) were rank-transformed.

The PRCC method is widely adopted in sensitivity analysis due to its ease of implementation and robustness in capturing the individual contributions of parameters to the model outcomes. One of its advantages is an interpretation of the sign of the correlation even in nonlinear models. Similar to the Pearson and Spearman methods, PRCC values close to 1 indicate that a small increase in a parameter leads to a relatively larger increase in the QoI. Similarly, PRCC values close to -1 indicate that a small increase in the parameter leads to a relatively larger decrease in the QoI. The significance of PRCC values can be assessed through hypothesis testing and the calculation of associated p-values. A low p-value (typically below 0.05) indicates that the observed correlation is statistically significant, providing evidence to reject the null hypothesis of no correlation between the variables.

In Figure 8 we present the results obtained using the PRCC method for our case study. We conducted the analysis in an 8-dimensional sample space using a Latin hypercube centered at the nominal parameter values provided in Table 1. Each parameter $p_i (i = 1, 2, ..., 8)$ was sampled normally within the interval $[0.95p_i, 1.05p_i]$. To determine the optimal number of samples, we explored sample sizes $n$ ranging from $10^2$ to $10^3$.

The PRCC values were calculated for each of the eight input parameters with respect to two outcome variables: (a) the predicted abundance of striped bass and (b) the predicted abundance of Atlantic menhaden in 2050 ($t_f = 50$ years). By employing the PRCC method, we effectively identified the key parameters that significantly influenced the abundances of both striped bass and Atlantic menhaden. In the case of striped bass abundance, we observed a robust positive correlation with parameter $r$, signifying its positive influence on the QoI, while parameter $b$ exhibited a clear negative correlation, implying its adverse impact. Notably, parameter $E_1$ showed a $p$-value less than 0.05, and its PRCC-value was approximately -0.4, indicating a moderate yet statistically significant negative correlation between the harvesting effort of Atlantic menhaden and the abundance of striped bass. When considering the QoI as the abundance of Atlantic menhaden, we observed a distinct positive correlation with parameter $e$ and a prominent negative correlation with parameter $d$. Parameters $r$, $c$, and $E_2$ demonstrated moderate yet statistically significant positive correlations with the abundance of Atlantic menhaden. Comparing these PRCC results with the findings presented in Figures 2, 4, 6, and 7, we noted a strong qualitative agreement regarding the parameters that had a substantial and highly statistically significant impact on the QoI. This reaffirms the consistency of our sensitivity analysis.




(a) Striped bass (b) The Atlantic menhaden

Figure 8: PRCC for striped bass the Atlantic menhaden ($t_f = 50$ years).

To examine sensitivity over time, we conduct a series of sensitivity analyses at specific time points, ranking the parameters for various values of $t_f$. This allows us to explore how the system's evolution towards the steady state is influenced by the parameters. By performing sensitivity analysis at discrete times $(t_1, t_2, ..., t_k)$ and observing the changes in partial rank correlation coefficients over time, we can gain insights into the dominant processes and their time-scale. For example, rapidly changing sensitivities at the initial stages of the process may require closely spaced observations to capture crucial information. Identifying the most sensitive parameters offers efficient control strategies when aiming to manipulate system behavior. If the sensitivity rankings change over time, it can even indicate the need for time-dependent adjustments in control strategies.

Figures 9 and 10 depict the temporal changes in PRCC values for the abundance of striped bass and the Atlantic menhaden, respectively, serving as the QoI. One can see that the sensitivity rankings at the initial stages of the process differ from those observed over a sufficiently extended period. For example, initially, parameters $d$, $e$, and $E_2$ demonstrate comparable importance to parameters $r$ and $b$ for the abundance of striped bass, but their sensitivities diminish significantly over time. Similarly, during the initial stage of the process for the abundance of the Atlantic menhaden, parameters $r$ and $b$ were important, but their sensitivity ranking diminished significantly compared to parameters $e$ and $d$ as time progressed.

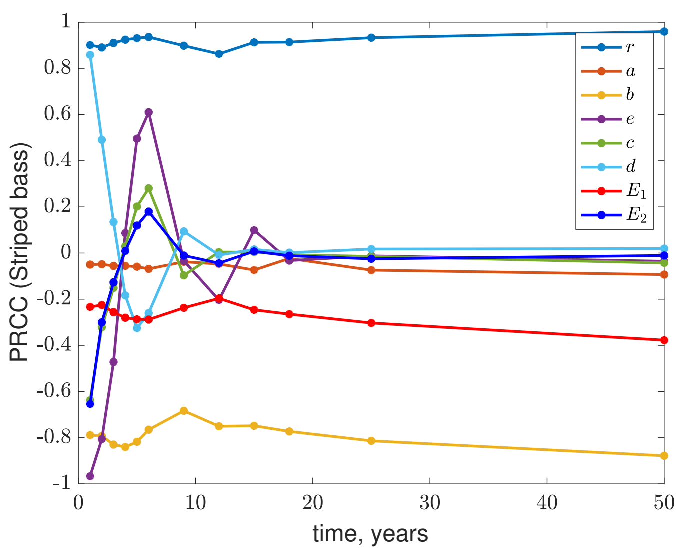



Figure 9: PRCC values in time for the case when the QoI is the value of $y$.

PRCC is a widely used straightforward ranking method that offers robustness and ease of implementation. However, it is important to note that the effectiveness of its correlation coefficients relies on the presence of monotonic relationships between parameters and the QoI, ensuring a linear relationship through rank transformation. This can require visual inspection of scatterplots and sometimes reduction of the parameter space. Also, it is a good practice to vary the number of samples in order to decide how many is enough.

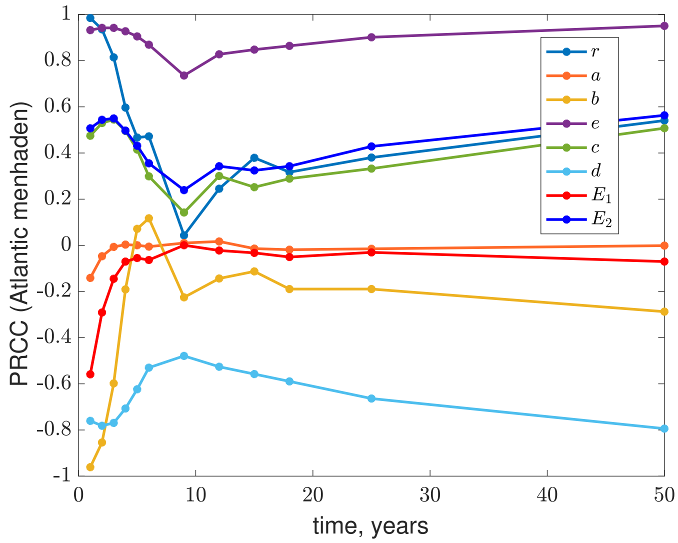



Figure 10: PRCC values in time for the case when the QoI is the value of $x$.

## 3.4 Method of Sobol'

We will now explore the implementation of the sensitivity analysis method proposed by I.M. Sobol' in 1993 [11]. The method is based on the ANOVA (Analysis of Variance) and it partitions the variance between the input parameters. There are at least two distinguishing features of Sobol'. First, this method applies to nonlinear relationships unlike the other methods we have discussed, which typically assume some degree of linearity. Second, Sobol' provides insights about the interactions between parameters. In contrast, PRCC, which we previously discussed as the most robust sensitivity method, disregards parameter interactions and focuses solely on the primary relationship between the input parameter and the QoI. interactions between parameters and only considers the primary relationship between the input parameter and the QoI.

The fundamental methodology of Sobol' estimates differs from the approaches discussed above. Sobol' sensitivity starts with the variance of an input/output relationship. Relating the sensitivity to the variance is different from relating it to the correlation. A QoI is considered correlated with a parameter if changes in the parameter result in changes in the QoI. On the other hand, a parameter induces high variance in the QoI if changes in the parameter cause larger variations relative to the mean of the QoI [17].

ANOVA, initially proposed by Fisher ([18]), is a widely used statistical method that assesses variations in means among different groups. It partitions the total variation into the variances within each group and the variances between groups, enabling the identification of significant variations that affect the measurements under consideration.

Sobol' ([11], [12]) generalized this concept and demonstrated that any general functional relationship $f(p_1, p_2, ..., p_m)$ between the QoI and parameters can be decomposed into the sum of specific orthogonal functions that allow deriving a different decomposition of the total variance.

The basis of the Sobol' method is to decompose the function $f(p_1, p_2, ..., p_m)$ into summands in increasing dimensionality, namely

$$f(p_1, p_2, ..., p_m) = f_0 + \sum_{i=1}^{m} f_i(p_i) + \sum_{1\le i<j\le m} f_{ij}(p_i, p_j) + ....f_{1,...,m}(p_1, ..., p_m). \tag{3.3}$$

If the input parameters are mutually independent, then there exist a unique decomposition (3.3) such that all the summands are mutually orthogonal i.e. if $(i_1, ..., i_s) \neq (j_1, ..., j_q)$, then

$$\int_{[0,1]^m} f_{i_1,...,i_s}(p_{i_1}, ..., p_{i_s}) f_{j_1,...,j_q}(p_{j_1}, ..., p_{j_q}) d\mathbf{p} = \mathbf{0} \tag{3.4}$$

From equations (3.3) and (3.4) we can find a relationship between the total variance ($Var$) and variance due to parameters

$$Var = \sum_{i=1}^{m} V_i + \sum_{i<j} V_{ij} + \sum_{i<j<l} V_{i,j,l} + ... + V_{1,2,3,...,m}, \tag{3.5}$$

where $V_{i_1,i_2,...,i_s} = \int_{[0,1]^m} f^2_{i_1,i_2,...,i_s} dp_{i_1}...dp_{i_s}$ is the variance due to specific parameter combinations $(p_{i_1}, ..., p_{i_s})$.

It is more useful to compare the partial variances to the total variance dividing equation (3.5) by $Var$ and rewriting it as

$$1 = \frac{\sum_{i=1}^{m} V_i}{Var} + \frac{\sum_{i<j} V_{ij}}{Var} + ... + \frac{\sum_{i<j<l} V_{i,j,l}}{Var} + \frac{V_{1,2,...,m}}{Var} \tag{3.6}$$

The first-order Sobol' sensitivity index, often denoted as $S_i$, quantifies the contribution of the individual input parameter $p_i$ to the total variance of the model output:

$$S_i = \frac{V_i}{Var}. \tag{3.7}$$

$S_i$ quantifies the portion of the total output variance that can be directly attributed to variations in parameter $p_i$, regardless of its interactions with other parameters.

The higher-order Sobol' sensitivity indices $S_{i_1,...,i_s} = \frac{V_{i_1,i_2,...,i_s}}{Var} V, (s > 1)$ extend the analysis beyond the individual input parameters to capture the interactions between multiple parameters. These indexes provide insights into the combined effects of parameter interactions on the output variance.

The total Sobol' sensitivity index, often denoted as $S_{T_i}$, represents the total contribution of parameter $p_i$ to the output variance, considering both its individual effect and its interactions with other parameters. It quantifies the proportion of the total output variance that can be attributed to the variations in parameter $p_i$. The total Sobol' index is defined as

$$S_{T_i} = \frac{\sum_{i=1}^{m} V_i}{Var} + \frac{\sum_{i\neq j} V_{ij}}{Var} + ... + \frac{\sum_{i\neq j\neq l} V_{i,j,l}}{Var} + \frac{V_{1,2,...,m}}{Var} \tag{3.8}$$

Below, we present the computed results for the total Sobol' indices, which we use to determine rankings of the parameter sensitivities in our case study model. Using the Monte

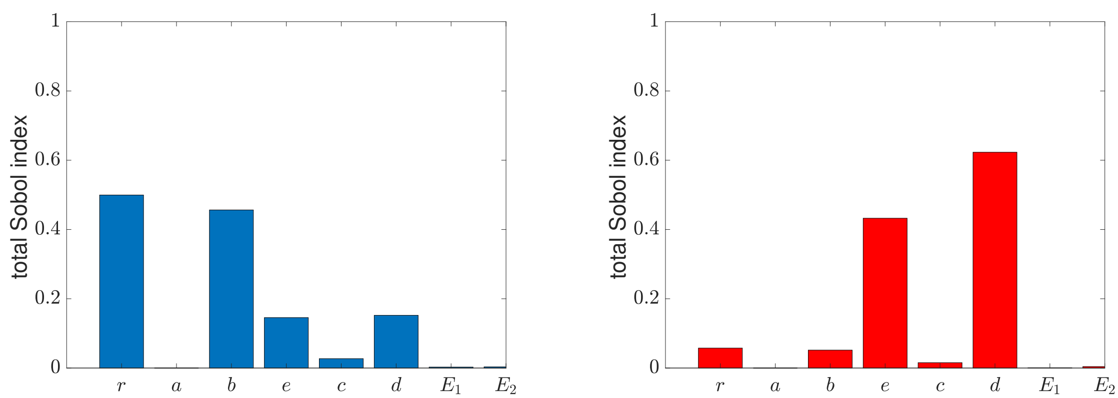
Carlo method, we generated $n = 5000$ samples from the parameter space and evaluated the model for each sample to obtain the corresponding output. The sensitivity indices were then calculated by decomposing the variance of the model outputs, considering both the main effects and interactions of the parameters.

For those interested in computing Sobol' indices, freely available software packages such as the SAFE toolbox for MATLAB ([19]) and SALib for Python (`https://salib.readthedocs.io/en/latest/`) offer convenient options.

In Figure 11 we present the ranking of Sobol' total sensitivity indices for stripped bass and the Atlantic menhaden. It makes sense to compare the results presented in Figure 11 with those presented in Figure 8. Notice that the total Sobol' indices ($S_{T_i}$ ) displayed in Figure 11 are calculated using the formula (3.8), and they can only take non-negative values. These indices indicate the importance of each parameter $p_i$ but do not provide information on whether the variation will result in an increase or decrease in the QoI.

Considering these factors, we observe qualitative agreement between Figures 8 and 11. Consistently, parameters $r$ and $b$ exhibit the highest influence on the abundance of the predator, while parameters $e$ and $d$ have the most pronounced effect on the abundance of the prey.

One can observe some disparity in the rankings of the subsequent parameters between Figures 8 and 11 . We attribute this to the limitations of the PRCC method related to possible non-linearity and non-monotonicity in the relationship between the QoI and some of the parameters (see scatterplots in Figures 3 and 5).



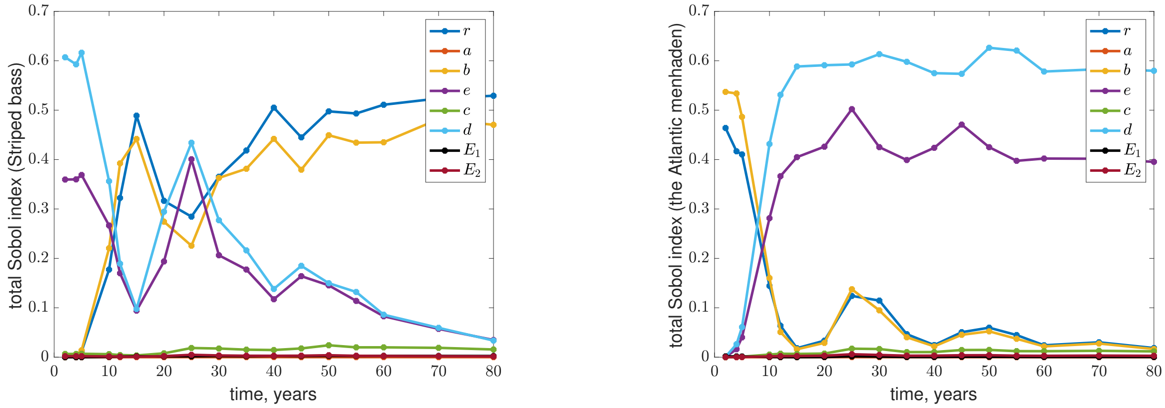

(a) Striped bass

(b) The Atlantic menhaden

Figure 11: Sobol' sensitivity ranking for striped bass the Atlantic menhaden ($t_f = 50$ years).

To examine time dependencies of parameter sensitivities, we conducted a series of sensitivity analyses at specific time points spanning from 2000 to 2080 ($0 \le t \le t_f = 80$ years). By performing sensitivity analysis at discrete time intervals, we gain insights into the varying significance of different parameters throughout the process and better understand their influence on the system's evolution towards a steady state. Figure 12 presents the time-varying Sobol' total sensitivity indices over an interval of 0-80 years.

In order to further investigate the time dependencies of parameter sensitivities, we compare the results obtained using the Sobol' and PRCC methods. This comparison allows




(a) The QoI is the value of $y$ (Striped bass) (b) The QoI is the value of $x$ (the Atlantic menhaden)

Figure 12: Sobol' total sensitivity indices changing in time.

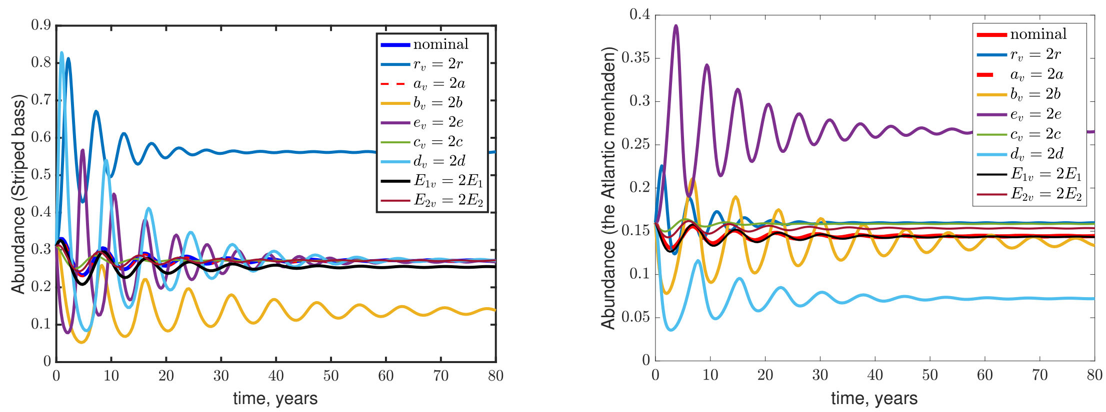
us to assess the significance of parameters for the predator (striped bass) as illustrated in Figure 9 (PRCC) and Figure 12(a) (Sobol). Similarly, we compare the results obtained for the prey (the Atlantic menhaden) as presented in Figure 10 (PRCC) and Figure 12(b) (Sobol). When comparing, we want to remember that the total Sobol' indices calculated using the formula (3.8) can only take non-negative values.

Both Figure 9 and Figure 12(a) illustrate the high sensitivity of striped bass to variations in the intrinsic growth rate of Atlantic menhaden $r$ and the impact of striped bass on Atlantic menhaden $b$. The comparison between Figure 10 and Figure 12(b) highlights the significant roles played by the death rate of striped bass $e$ and the effect of Atlantic menhaden on striped bass $d$ in determining the abundance of Atlantic menhaden. The findings align with the outcomes of the local sensitivity analysis conducted by Panayotova et al. [4]. When it comes to the next important parameters, parameters $e$ and $d$ for striped bass, and parameters $r$ and $b$ for the Atlantic menhaden may deserve consideration. While parameters $d$ and $e$ initially demonstrate importance for striped bass, their significance diminishes over time as indicated by the declining PRCC and Sobol' total sensitivity indices. Similarly, parameters $r$ and $b$ for the Atlantic menhaden exhibit importance but primarily during the early stages of the process.

The Sobol' method is recognized as one of the most commonly used approaches for parameter sensitivity analysis. Along with PRCC has found widespread acceptance in the field. Its advantage is that it does not assume a monotonic relationship between the quantity of interest (QoI) and the parameters. However, it may require a larger number of samples to achieve convergence compared to other methods. Additionally, while the assumption of parameter independence holds true in most cases, we should notice that Sobol' estimates may be less reliable when working with interdependent parameters. When parameters are highly correlated or interdependent, Sobol' indices may incorrectly attribute sensitivity to one parameter instead of recognizing its combined effect with others, preventing effective discrimination between parameters. Exploratory analysis of available data needs to be done to assess parameter interdependencies, identify Sobol's limitations, and guide result interpretation. Techniques like principal component analysis

(PCA) can mitigate the impact of correlated parameters [20].

We conclude our exploration with Figure 13, which presents a comparison of solutions obtained using nominal and varied parameter values for both striped bass and the Atlantic menhaden over a time period spanning from 2000 to 2080 ($0 \leq t \leq 80$ years). These results align with the findings of the discussed sensitivity analysis methods and show reasonable agreement with the graphs by Panayotova et al. [4], depicting the rate of change of each species' population over time with respect to the system parameters. Notably, we observe that harvesting the prey has a more negative impact on the predator, while harvesting the predator has a relatively minor effect on its own steady-state but leads to a modest increase in the prey population's steady-state. This finding is consistent with the results reported by Panayotova et al. [4] and supports our previous observation that the abundance of the predator is primarily determined by the availability of prey. Furthermore, the abundance of the prey is significantly influenced by the effect of its consumption on the predator and the decline in the predator's population due to death or harvest.




(a) Striped bass (all parameters)

(b) The Atlantic menhaden (all parameters)

Figure 13: Comparison between the solutions using the nominal and varied parameter values ($t_f = 80$ years).

# Conclusion

By providing an exploration of parameter sensitivity analysis in mathematical modeling, this paper has the potential to serve as a resource for both professors teaching differential equations or mathematical modeling and students interested in incorporating parameter sensitivity analysis into their modeling endeavors.

Through the application of local sensitivity analysis using the direct differential method and global sensitivity analysis using metrics such as Pearson, Spearman, PRCC, and Sobol, we have provided readers with a basics of parameter sensitivity analysis in ODE-based models. In our illustrative application, we have examined the population dynamics of two fish species with harvest considerations, uncovering important insights about the dependence of predator abundance on available prey and the significant influence of prey consumption on predator abundance.

By comparing the results obtained from local and global sensitivity analyses, we have highlighted the strengths and limitations of each approach.

Local sensitivity analysis is advantageous in its simplicity and interpretability, making it well-suited for preliminary investigations and quick assessments of parameter importance. It can help identify key parameters for further exploration and guide model refinement. However, it has limitations in capturing the combined effects of multiple parameters and the potential interactions among them. Local analysis may overlook important system dynamics and fail to capture the overall sensitivity landscape accurately.

In contrast, global sensitivity analysis offers a more comprehensive understanding of the model by quantifying the contributions of all parameters simultaneously. It captures non-linear relationships and interactions, enabling the identification of synergistic or antagonistic effects among parameters. Global sensitivity analysis is particularly useful for complex models where parameter interactions play a significant role.

However, global sensitivity analysis methods require more computational resources and can be computationally demanding, especially for high-dimensional models. Additionally, the interpretation of global sensitivity measures can be challenging, as they involve statistical correlations and complex mathematical calculations.

The choice between local and global parameter sensitivity analysis depends on the specific goals and characteristics of the modeling. Local analysis is suitable for initial investigations, identifying critical parameters, and gaining a basic understanding of the system. Global analysis provides a more comprehensive view, capturing interactions and non-linear effects, but at the cost of increased computational complexity. The two approaches can be complementary. A general suggestion would be to start with local sensitivity analysis to identify influential parameters and gain initial insights into the system's behavior. If the model exhibits complex dynamics or parameter interactions are suspected, conducting global sensitivity analysis can provide a deeper understanding of the system's sensitivity.

The choice of sensitivity indices is critical, as they vary in suitability for different models and research objectives. Researchers should align their index selection with their analysis goals. Furthermore, it is worth noting that the field of sensitivity analysis has evolved to incorporate statistical methods, such as hypothesis testing and confidence intervals, which enable us to draw more robust inferences from sensitivity results [7]. These methods help determine whether sensitivity indices are statistically significant and provide probabilistic bounds on their values.

Parameter sensitivity analysis enriches pedagogy by actively engaging students in real-world problem-solving, fostering critical thinking, and bridging interdisciplinary connections. It equips learners with data analysis skills, enhances the application of mathematical concepts, and prepares them for data-driven careers.

By examining the interplay between model parameters and system dynamics, researchers can refine their models, validate their assumptions, and make more informed predictions.

# Appendix: Sensitivity Analysis Tasks for Instructors and Students

In this appendix, we provide some tasks that instructors can adapt and use in their classrooms to help students gain hands-on experience with parameter sensitivity analysis in ODE-based models.

## Task 1: Local Sensitivity Analysis

**Objective**: Understand the concept of local sensitivity analysis and its applications.

1. Provide students with an ODE-based model.
2. Ask students to identify dependent variables, independent variables and parameters of the model.
3. Instruct students to perform local sensitivity analysis using the direct differential method.
4. Ask them to calculate sensitivity indices for each parameter and rank them in terms of their influence on the model's output.
5. Encourage students to interpret the results and discuss how changes in each parameter affect the model's behavior.

## Task 2: Global Sensitivity Analysis

**Objective**: Explore global sensitivity analysis methods and use scatterplots to select suitable methods.

1. Present an ODE-based model with multiple parameters.
2. Instruct students to identify model's parameters and the Quantity of Interest (QoI), representing the aspect they want to investigate (e.g., fish abundance).
3. Guide students in choosing appropriate parameter ranges and distributions.
4. Explore the generation of a sample space, emphasizing the importance of the number of samples and their distribution.
5. Ask students to create scatterplots by varying one parameter at a time while keeping others constant, plotting the QoI on the y-axis against the parameter on the x-axis.
6. Encourage visual analysis of scatterplots to identify parameter-QoI patterns.
7. Based on scatterplot observations, guide students in selecting an appropriate global sensitivity analysis method (e.g., Pearson correlation coefficient, Spearman correlation coefficient, or Sobol indices) that aligns with the observed patterns.

8. Have students perform the selected global sensitivity analysis and interpret results within the QoI context.

9. Facilitate a discussion where students compare their selected method's outcomes with initial scatterplot observations, reflecting on method advantages.

10. Discuss how the chosen method reveals parameter interactions, quantifying their impact on the QoI, aiding model behavior understanding.

Note: Students with prior exposure to statistics can also be assigned the task of conducting hypothesis testing to assess the statistical significance of observed results (e.g., by calculating p-values).

## Acknowledgment

The author was supported by the 2023–2024 CSU-AAUP Faculty Research Grant.

## References

[1] Blanchard, P., Devaney, R. L., Hall, G.R. (2012). *Differential Equations*, Brooks/Cole, Cengage Learning, Boston, MA.

[2] Bryan, K. (2022). *Differential Equations. A Toolbox for Modeling the World.* SIMIODE, Chardon OH USA.

[3] Hallare, M., Panayotova I. (2023) Maintaining Ecosystem and Economic Structure in a Three-Species Dynamical System in Chesapeake Bay, *CODEE Journal*, vol. 15, Article 3.

[4] Panayotova I.N., Herrmann, J., Kolling, N. (2023) Bioeconomic analysis of harvesting within a predator-prey system: A case study in the Chesapeake Bay fisheries. *Ecological Modeling*, 480, 110330.

[5] Saltelli, A., Chan, K., Scott, EM. (2000) *Sensitivity analysis*, Wiley series in probability and statistics. Wiley, Chichester; New York.

[6] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, J., Tarantola, S. (2008) *Global Sensitivity analysis: The primer*, Wiley, LTD., Chichester.

[7] Razavi, S., Jakeman, A., Saltelli, A., Prieur,C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J.H.A. , Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh,M., Arnald, P., Kucherenko, S., Maier, H.R. (2021) The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling and Software* 137: 104954

[8] Pearson, K. (1895) Vii. note on regression and inheritance in the case of two parents. *Proceedings of the royal society of London*, 58 (347-352):240-242.

[9] Spearman, C. (1904)The proof and measurement of association between two things. *The American Journal of Psychology*, v.15(1): 72-101.

[10] Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E. (2008) A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* 254, 178-196.

[11] Sobol', IM. (1993) Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput.Exp.*, 1(4):407-414.

[12] Sobol, I.M. (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55, p. 271-280.

[13] Larsen, R.J., Marx, M.L. (2018) *An Introduction to Mathematical Statistics and Its Applications*, Pearson.

[14] James, G., Witten, D., Hastie, T. and Tibshirani, R. (2017) *An Introduction to Statistical Learning with Applications in R*, Springer, Berlin.

[15] McKay, MD., Beckman, RJ., Conover, WJ. (1979) Comparison of 3 Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21(2): 239-245.

[16] Kendall, MG. (1942) Partial rank correlation. *Biometrica*, 32(3/4): 277-283.

[17] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer physics communications*, 181 (2):259-270.

[18] Fisher, R. A., & Mackenzie, W.A. (1923) Studies in crop variation. ii.the manurial response of different potato varieties. *The Journal of Agricultural Science*, 13(3): 311-320.

[19] Pianosi, F., Sarrazin, F., Wagener, T. (2015) A Matlab toolbox for Global Sensitivity Analysis. *Environmental Modeling & Software*, 70: 80-85.

[20] Jolliffe,I.T. (2002) *Principal component analysis for special types of data*, Springer, New York.