Mathematics and Statistics 11(2): 421-433, 2023
DOI: 10.13189/ms.2023.110221 http://www.hrpub.org

Sensitivity Equation for Competitive Model: Derivation, Numerical Realization and Parameter Estimation

Julan HERNADI[1,*] , **Ceriawan H. SANTOSO**[2], **Iwan T. R. YANTO**[3]

[1]Department of Mathematics, Ahmad Dahlan University, Jalan Ringroad Selatan Banguntapan Bantul, Yogyakarta, Indonesia

[2]Indonesian Humanist Mathematics Foundation (MATHUMINDO), Tempel-Wirogunan UH 3/905 Yogyakarta, Indonesia

[3]Department of Information System, Ahmad Dahlan University, Jalan Ringroad Selatan Banguntapan Bantul, Yogyakarta, Indonesia

Received December 21, 2022; Revised February 17, 2023; Accepted March 12, 2023

Cite This Paper in the following Citation Styles

***(a)**: [1] Julan HERNADI, Ceriawan H. SANTOSO, Iwan T. R. YANTO, "Sensitivity Equation for Competitive Model: Derivation, Numerical Realization and Parameter Estimation," Mathematics and Statistics, Vol.11, No.2, pp. 421-433, 2023. DOI: 10.13189/ms.2023.110221*

***(b)**: Julan HERNADI, Ceriawan H. SANTOSO, Iwan T. R. YANTO, (2023). Sensitivity Equation for Competitive Model: Derivation, Numerical Realization and Parameter Estimation. Mathematics and Statistics, 11(2), 421-433. DOI: 10.13189/ms.2023.110221*

Abstract Ecological systems can be quite complex, consisting of an interconnected system of plants and animals, predators and prey, flowering plants, seed dispersers, insects, parasites, pollinators, and so on. In the case of the existence of a species affecting the survival of other species and vice versa, it can derive a competitive model in the form of a system of differential equations. A competitive model involves a number of parameters which grow in proportion to the number of interacting species. The resistance of a state variable to tiny disturbances of some parameter is referred to as sensitivity. The competitive model of size N consists of N parameters for intrinsic growth, N parameters for carrying capacity, $N^2 - N$ parameters for species interaction, and N parameters for initial conditions. As a result, there will be $N^2(N + 2)$ distinct values of sensitivity. The purpose of this paper is to derive a general formulation of the sensitivity equations of dynamical system and then apply it to the competitive model. This study also encompasses the formulation of some algorithms and the implementation for solving the sensitivity equation numerically. Finally, the sensitivity functions are employed as qualitative instruments in the optimal design of measurement for parameter estimation through a series of numerical experiments. The results of this study are the ordinary and the generalized sensitivity functions for interacting species. Based on numerical experiments, each group of data provides different information about the existing parameters.

Keywords Sensitivity Equation, Competitive Model, Sensitivity Function, Optimal Sample, Parameter Estimation

AMS Subject Classification: 34A55; 49M05; 65L05; 65S05; 92D25.

1 Introduction

Parameter estimation is the process of using observation data in a dynamic system to develop a mathematical model that describes the characteristics of that system. Mathematical modeling through parameter estimation is a way that leads to a deeper understanding of the system characteristics [19, 11]. The mathematical model is assumed to have a finite number of parameters that must be estimated by some measurement data. In general, the estimation technique is based on minimizing the error between the model response and the system response. The system response is obtained through some discrete time measurements, where each measured data is assumed to convey information to a certain degree of the parameter being estimated [2, 24]. For an illustration, suppose the discrete-time observation is given by

$$y(t_k) = f(t_k; \theta) + \epsilon(t_k), k = 1, 2, \cdots, M \tag{1}$$

where f is the model response and $\epsilon(t_k)$ is the measurement noise assumed to be independently identically distributed (i.i.d.) with zero mean and variance $\sigma^2(t_k)$. Suppose that $f = (f_1, f_2, \cdots, f_M)$ and $\theta = (\theta_1, \theta_2, \cdots, \theta_p)$ represent the vector of states and parameters, respectively, then the time course derivatives of f_i with respect to parameter θ_j, i.e. $\frac{\partial f_i}{\partial \theta_j}$, $i = 1, 2, \cdots, M$, $j = 1, 2, \cdots, p$ are called the sensitivity functions.

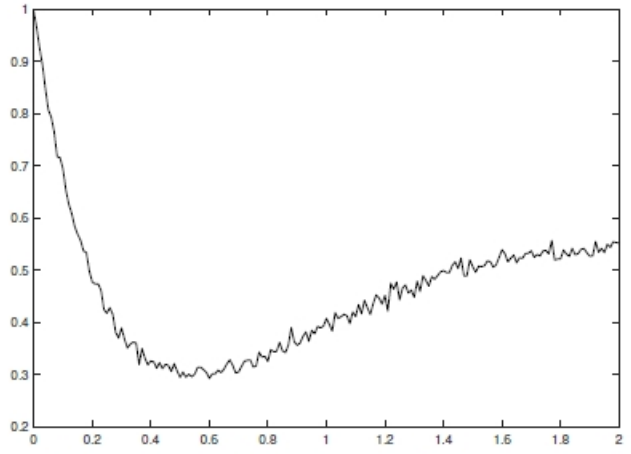


Figure 1. The measurement with noise

They describe how model output trajectories change in response to modest changes in model parameters [22, 3, 5]. In most cases, the model response f implicitly depends on the solution of some initial value problem (IVP).

To gain an awareness of the issues to be discussed, consider the explicit model of three parameters as follows.

$$y(t) = e^{-\alpha_1 t} + te^{-\alpha_2 t} + t^2 e^{-\alpha_3 t} + \epsilon(t). \tag{2}$$

In this case, the parameter vector is given by $\theta = (\alpha_1, \alpha_2, \alpha_3)$ and the noise $\epsilon(t)$ is assumed to be zero mean with constant variance.

This signal was generated at discrete time $t = 0 : 0.01 : 2$ with the true parameters $\alpha_1 = 5.12$, $\alpha_2 = 3.92$, $\alpha_3 = 0.96$, and noise $\sigma = 0.01$. The issue to be addressed is how to estimate these parameters using distorted data as shown in Figure 1. More specifically, how many and when measurements should be performed in order for the observation data to provide as much information about the parameter being estimated as possible. Let $\{t_1, t_2, \cdots, t_M\}$ be the measurement times, the performance index based on the mean square error (MSE):

$$MSE = \frac{1}{M}\sum_{k=1}^{M} |y(t_k) - f(t_k;\theta)|^2 . \tag{3}$$

Despite the fact that the signal was defined on $[0, 2]$, measurements are considered also separately in each subdomain $[0, 1]$ and $[1, 2]$. The following samples were purposefully chosen to assess the information content of parameters provided by the data set:

Sample#1 : $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$

Sampel#2 : $\{1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9\}$

Sample#3 : $\{0.25, 0.5, 0.75\}$

Sample4# :$\{1.25, 1.5, 1.75\}$

Sample#5 :$\{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8\}$

The estimate $\hat{\theta}$ is obtained by minimizing the MSE on specific measurement points. The simulation results presented in Table 1 were obtained by using the MATLAB toolbox "fminsearch" for multidimensional unconstrained nonlinear minimization based on Nelder-Mead method [10]. The results are presented in Table 1. Samples 2 and 4 were drawn from the interval $[1, 2]$ and they provide worse estimates than samples 1 and 3 which taken from the interval $[0, 1]$. For instance, sample 4 produces estimates $\hat{\alpha}_1 = 32.869$ and $\hat{\alpha}_2 = 34.171$ for the true parameter $\alpha_1 = 5.12$ and $\hat{\alpha}_2 = 3.92$, respectively, very worse even though their MSEs are small enough.

Table 1. Results of simulation with the initial iteration $\theta^{(0)} = (1, 1, 1)$ and the true parameter $\theta_0 = (5.12, 3.92, 0.96)$.

Measurement	Estimate: $\hat{\theta}$	$MSE(\hat{\theta})$
Sample#1	$(5.036, 4.089, 0.991)$	4.912×10^{-5}
Sample#2	$(3.804, 160.589, 0.961)$	1.643×10^{-4}
Sample#3	$(5.298, 3.778, 1.027)$	6.478×10^{-14}
Sample#4	$(32.869, 34.171, 0.956)$	1.441×10^{-5}
Sample#5	$(5.280, 3.866, 0.956)$	1.303×10^{-4}

Therefore, a reference is required in order to select the best sample in terms of providing the most information about the parameters to be estimated. The sensitivity functions, both traditional and generalized, were frequently used as a reference for optimal measurement. The sharp increase of generalized sensitivity functions indicate a high concentration of information about parameters, see [22, 5] and references therein. Figure 2 shows that fluctuations in both sensitivity functions occurred more frequently in the interval $[0, 1]$ than $[0, 2]$ for α_1 and α_2. On the other hand, the curves fluctuations relating to α_3 were not too often.

Sensitivity analysis refers to the procedures used in simulation studies to determine the impact of parameter variations on the time course of model output and also to identify which parameters in the model are the most sensitive or the least sensitive [3]. On this hand, a complete description of a physical system, i.e. the mathematical model and the parameter's values are supposed, then it can be used to predict the outcome of some measurements. The problem of predicting result of measurements is called the modelization problem, the simulation problem, or the forward problem [21]. On the other hand, the inverse problem consists of using the actual result of measurements to predict the value of the parameters that characterize the system. The forward problem has a unique solution or is well-conditioned, but the inverse problem has not or is ill-conditioned [21, 12, 1, 17, 27]. It means that multiple measurements could result in a single estimate. Another critical issue in the inverse problem is the possibility of different parameter sets which describe the same system. For instance, as shown in Table 1 that two estimates $E_1 = \{\hat{\alpha}_1 = 5.036, \hat{\alpha}_2 = 4.089, \hat{\alpha}_3 = 0.991\}$ and $E_2 = \{\hat{\alpha}_1 = 3.804, \hat{\alpha}_2 = 160.589, \hat{\alpha}_3 = 0.961\}$ are completely different for $\hat{\alpha}_1$ and $\hat{\alpha}_2$, but they produced similar outputs. As a result, the quality of estimates is qualified not only by MSE but also by their reliability with respect to disturbance.

According to previous simulation, it is clear that both types of sensitivity functions play an important role in the optimal design of measurement. The sensitivity functions are generally not explicitly available; nevertheless, they must be revealed by solving

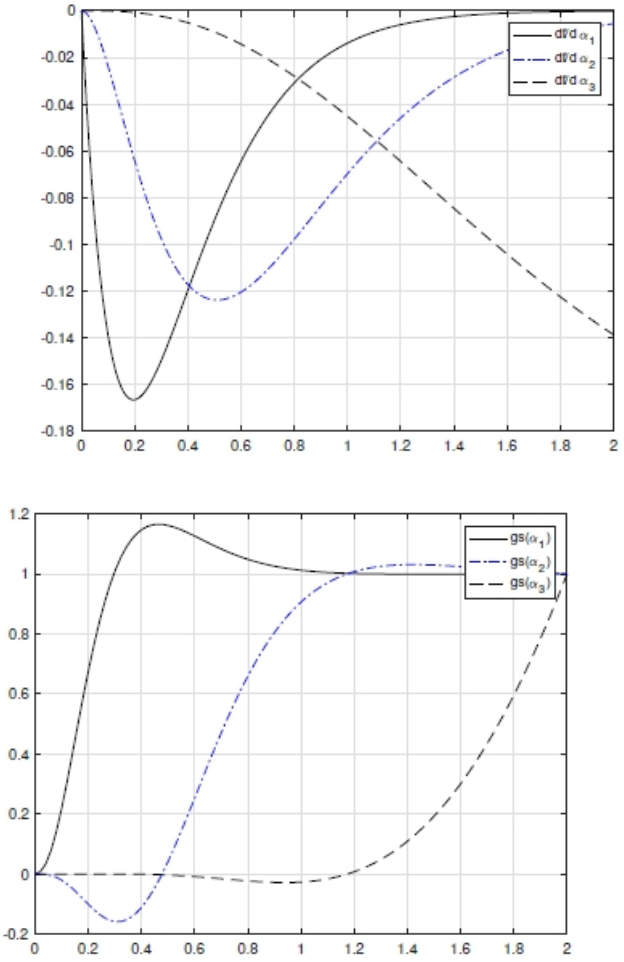


Figure 2. The sensitivity functions of system (1.2): traditional (up) and generalized (bottom).

the sensitivity equation, which is a system of differential equations equipped with some initial condition. This paper will address issues related to the derivation and numerical implementation of the sensitivity equation, and to the application in the competitive models. The paper is structured as follows. Section 2 will describe an abstract formulation of the sensitivity equation. A construction of the sensitivity equation for competitive model will be discussed in section 3. Section 4 will focus on the algorithms composition for numerical implementation. Finally, some numerical experiments concerned with the use of TSF and GSF in selecting the best sample for parameter estimation will be demonstrated in section 5. Various mean square errors will be employed in some numerical experiments to assess the information content of the parameters conveyed by each existing state.

2 Abstract Formulation of Sensitivity Equation

Let $X(t) := [x_1(t), \cdots, x_N(t)]^T$ be the state function involving a parameter vector $\theta = [\theta_1, \theta_2, \cdots, \theta_p]^T$ and $F : [t_0, t_f] \times \mathbf{R}^N \to \mathbf{R}^N$ be a vector function of the form

$$F(t, X; \theta) = [f_1(t, X; \theta), f_2(t, X; \theta), \cdots, f_N(t, X; \theta)]^T$$

where $f_i : [t_0, t_f] \times \mathbf{R}^N \to \mathbf{R}$, $i = 1, 2, \cdots, N$. Consider the general nonlinear dynamical system

$$\frac{dX}{dt} = F(t, X; \theta), X(t_0) = X_0. \tag{4}$$

In this case $X_0 := [x_1(t_0), x_2(t_0), \cdots, x_N(t_0)]^T = [x_{01}, x_{02}, \cdots, x_{0N}]^T$ is the initial value vector. The sensitivity equation is derived by taking the derivative of the state vector X with respect to the vector parameter θ on both sides of (4). By employing the interchange property of derivative, i.e. $\frac{\partial}{\partial\theta}\left(\frac{dX}{dt}\right) = \frac{d}{dt}\left(\frac{\partial X}{\partial\theta}\right)$, the following sensitivity equation is obtained.

$$\frac{d}{dt}\left(\frac{\partial X}{\partial\theta}\right) = \frac{\partial F}{\partial X}\frac{\partial X}{\partial\theta} + \frac{\partial F}{\partial\theta}. \tag{5}$$

It is a system of differential equations with state matrix $\left[\frac{\partial X}{\partial\theta}\right]$. The elements of matrices in the equation (5) are explicitly given as

$$\frac{\partial X}{\partial\theta} = \begin{bmatrix} \frac{\partial x_1}{\partial\theta_1} & \cdots & \frac{\partial x_1}{\partial\theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_N}{\partial\theta_1} & \cdots & \frac{\partial x_N}{\partial\theta_p} \end{bmatrix}, \frac{\partial F}{\partial X} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1} & \cdots & \frac{\partial f_N}{\partial x_N} \end{bmatrix},$$

$$\frac{\partial F}{\partial\theta} = \begin{bmatrix} \frac{\partial f_1}{\partial\theta_1} & \cdots & \frac{\partial f_1}{\partial\theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial\theta_1} & \cdots & \frac{\partial f_N}{\partial\theta_p} \end{bmatrix}.$$

Since $x_i(t_0)$ does not contain parameter in θ then $\frac{\partial x_i(t_0)}{\partial\theta_j} = 0$, for $i = 1, 2, \cdots, N$ and $j = 1, 2, \cdots, p$. Therefore, the initial condition of system (5) can be written as

$$\frac{\partial X(t_0)}{\partial\theta} = zeros(N, p), \tag{6}$$

i.e. the null matrix of size $N \times p$. The problem now is how to determine the sensitivity matrix $\frac{\partial X}{\partial\theta}$ that satisfies (5) with initial condition (6). There are $N \times p$ elements of sensitivity matrix where for i dan j fixed, $\frac{\partial x_i}{\partial\theta_j}$ represents the trajectory change of x_i with respect to the small change of parameter θ_j. The solution of state equation (4) is required for defining matrices $\frac{\partial F}{\partial X}$ and $\frac{\partial F}{\partial\theta}$ in (5). For simplicity, the equation (5) is decomposed into individual parameters. As an illustration, for $i = 1, 2, \cdots, p$, this equation can be written as

$$\frac{d}{dt}\begin{bmatrix} \frac{\partial x_1}{\partial\theta_i} \\ \vdots \\ \frac{\partial x_N}{\partial\theta_i} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1} & \cdots & \frac{\partial f_N}{\partial x_N} \end{bmatrix}\begin{bmatrix} \frac{\partial x_1}{\partial\theta_i} \\ \vdots \\ \frac{\partial x_N}{\partial\theta_i} \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial\theta_i} \\ \vdots \\ \frac{\partial f_N}{\partial\theta_i} \end{bmatrix}. \tag{7}$$

This is a linear system of the form $y' = Ay + f$ and the theoretical studies of these types of equations can be found in some differential equation textbooks, e.g. Borelli & Colleman [8]. In this paper, the sensitivity equations (5) - (6) and the state equation (4) are solved simultaneously by some numerical approaches in

which the initial condition is treated as a parameter. As a consequence, the quantities $\frac{\partial x_i}{\partial x_{0j}}$ must be considered as a state variable too. Since F does not involve X_0 it fulfills $\frac{\partial F}{\partial \theta} = 0$, and thus the sensitivity equation for X_0 is obtained as follows:

$$\frac{d}{dt}\left(\frac{\partial X}{\partial X_0}\right) = \frac{\partial F}{\partial X}\frac{\partial X}{\partial X_0}. \tag{8}$$

The state variable $\frac{\partial X}{\partial X_0}$ is a matrix function of size $N \times N$, i.e. $\frac{\partial X}{\partial X_0} : [t_0, t_f] \rightarrow \mathbf{R}^{N\times N}$ where for each $t \in [t_0, t_f]$, it can be written explicitly as

$$\frac{\partial X}{\partial X_0}(t) = \begin{bmatrix} \frac{\partial x_1(t)}{\partial x_{01}} & \cdots & \frac{\partial x_1(t)}{\partial x_{0N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_N(t)}{\partial x_{01}} & \cdots & \frac{\partial x_N(t)}{\partial x_{0N}} \end{bmatrix}.$$

The differential equation (8) is equipped by the initial condition

$$\frac{\partial X(t_0)}{\partial X_0} = I_{N\times N}, \tag{9}$$

where $I_{N\times N}$ denotes an indentity matrix of size N. This matrix is easily obtained by remembering that $\frac{\partial x_i(t_0)}{\partial x_{0j}} = 1$ if $i = j$ and 0 if $i \neq j$.

It can be understood that a system of size N with p parameters will govern the sensitivity equation of size $N \times p$. In case the initial condition is also considered as parameter, then there will be one additional equation of size $N \times N$. The large sizes of system, the matrix function as unknown, and the dependence on the state equation are some issues that give rise to challenges in solving the sensitivity equations.

3 Derivation of Sensitivity Equation for Competitive Model

A detailed discussion of models for population interactions has been written by Murray in [18], particularly on two-species system. There are three main types of interaction, i.e. predatorprey situation if the growth rate of one population is decreased meanwhile the other increased, competition if the growth of each population is decreased, and mutualism or symbiosis when both the growth rates are enhanced. In this paper, the three cases are not distinguished and they are regarded as competitive models. In some references, the predator-prey equations are also referred as the kind of Lotka-Volterra equations.

The competitive model can be described as

$$\frac{dx_i}{dt} = r_i x_i \left(1 - \frac{\sum_{j=1}^{N}\alpha_{ij}x_j}{K_i}\right), x_i(t_0) = x_{i0}, \tag{10}$$

where $i = 1, 2, \cdots, N$. In this model, N species where $N \geq 2$ are having interactions with one another. For each species $i = 1, 2, \cdots, N$, x_i indicates the size of species, r_i signifies the intrinsic growth, K_i represents the carrying capacity, and $\alpha_{i,j}$ describes the impact of i-th species on j-th species. It is assumed that $\alpha_{ii} = 1, i = 1, 2, \cdots, N$ and $\alpha_{ij} > 0$ for $i \neq j$. There are no special constraints on the interaction parameters as long as they do not generate the chaotic behavior, as explained in [26]. The system (10) is regarded as the standard model of Lotka-Volterra equations. The more complicated part of Lotka-Volterra model has been discussed in [9] where the states emerged with the feedback control and deviating argument.

Clearly that the competitive equation (10) is the special case of the abstract formulation (4), for which $f_i(t, X; \theta) = r_i x_i \left(1 - \frac{\sum_{j=1}^{N}\alpha_{ij}x_j}{K_i}\right)$, $i = 1, 2, \cdots, N$. Observe that $\alpha_{ii} := 1$, and thus the competitive equation for N species involves $N(N+1)$ parameters for equation, viz. $\{r_i, K_i, \alpha_{ij} : i, j = 1, 2, \cdots, N\} \setminus \{\alpha_{ii} : i = 1, 2, \cdots, N\}$ and N additional parameters for initial values, i.e. $x_i(t_0) = x_{0i}$, $i = 1, 2, \cdots, N$. The special case of (10) is when $N = 2$ which is well-known as the Lotka-Volterrra equation [13, 16, 18].

$$\frac{dx_1(t)}{dt} = r_1 x_1(t)\left(1 - \frac{x_1(t) + \alpha_{12}x_2(t)}{K_1}\right), t \geq t_0 \tag{11}$$

$$\frac{dx_2(t)}{dt} = r_2 x_2(t)\left(1 - \frac{\alpha_{21}x_1(t) + x_2(t)}{K_2}\right), t \geq t_0 \tag{12}$$

These equations are equipped with the initial conditions $x_1(t_0) = x_{10}$ and $x_2(t_0) = x_{20}$. The logistic growth model of Verhulst-Perl is this type for $N = 1$ [2].

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right), x(t_0) = x_0 \tag{13}$$

where r and K denote growth rate and carrying capacity, respectively. Interestingly, the Verhulst-Perl equation has a simple explicit solution provided by

$$x(t) = \frac{K}{1 + \left(\frac{K}{x_0} - 1\right)e^{-rt}}, \tag{14}$$

where $x_0 = x(0)$ denotes the initial population size. The corresponding parameters vector is $\theta = (K, r, x_0)$.

The elements of matrices in the sensitivity equations (5) must be revealed in order to solve the sensitivity equation numerically. The results of some algebraic manipulations for defining the elements of those matrices are shown below. The elements of the Jacobian matrix $\frac{\partial F}{\partial X} = \left(\frac{\partial f_i}{\partial x_j}\right)$, $i, j = 1, 2, \cdots, N$ are determined by

$$\frac{\partial f_i}{\partial x_j} = \begin{cases} r_i\left(1 - \frac{\sum_{k=1}^{N}\alpha_{ik}x_k + x_i}{K_i}\right), & j = i \\ -\frac{r_i\alpha_{ij}x_i}{K_i}, & j \neq i. \end{cases} \tag{15}$$

Take note of the parameters groups attached to the competitive equation (10), i.e. $\{r_i : i = 1, 2, \cdots, N\}$, $\{K_i : i = 1, 2, \cdots, N\}$, and $\{\alpha_{i,j} : i, j = 1, 2, \cdots, N, i \neq j\}$. In the numerical implementation, it will be advantageous for matrix $\frac{\partial F}{\partial \theta}$ to be decomposed into parameter groups $\frac{\partial F}{\partial r}$, $\frac{\partial F}{\partial K}$, and $\frac{\partial F}{\partial \alpha}$. The intrinsic growth rate and carrying capacity matrices are calculated as follows:

$$\left(\frac{\partial F}{\partial r}\right) = \text{diag}\left[x_i\left(1 - \frac{\sum_{k=1}^{N}\alpha_{ik}x_k}{K_i}\right), i = 1, 2, \cdots, N\right]. \tag{16}$$

$$\left(\frac{\partial F}{\partial K}\right) = \text{diag}\left[r_i x_i \left(\frac{\sum_{k=1}^{N} \alpha_{ik} x_k}{K_i^2}\right), i = 1, 2, \cdots, N\right]. \tag{17}$$

Furthermore, the interaction parameters $\{\alpha_{ij}, i, j = 1, 2, \cdots, N, j \neq i\}$ must be rearranged, for instance, using the lexicographic technique. For example, for $N = 3$, the new indices are given by $\alpha_{12} := \tau_1, \alpha_{13} := \tau_2, \alpha_{21} := \tau_3, \alpha_{23} := \tau_4, \alpha_{31} = \tau_5, \alpha_{32} = \tau_6$. Accordingly, the following matrix is obtained. $\left(\frac{\partial F}{\partial \alpha}\right)_{3\times 6} =$

$$\begin{bmatrix} \frac{-r_1 x_1 x_2}{K_1} & \frac{-r_1 x_1 x_3}{K_1} & 0 & 0 & & 0 \\ 0 & 0 & \frac{-r_2 x_2 x_1}{K_2} & \frac{-r_2 x_2 x_3}{K_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-r_3 x_3 x_1}{K_3} & \frac{-r_3 x_3 x_2}{K_3} \end{bmatrix} \tag{18}$$

The simple case is when $N = 2$, i.e. $\alpha_{12} := \tau_1$ and $\alpha_{21} := \tau_2$, so that

$$\left(\frac{\partial F}{\partial \alpha}\right)_{2\times 2} = \left(\frac{\partial f_i}{\partial \tau_j}\right)_{i,j=1,2} = \begin{bmatrix} \frac{-r_1 x_1 x_2}{K_1} & 0 \\ 0 & \frac{-r_2 x_2 x_1}{K_2} \end{bmatrix}.$$

As a result, there are as many as four groups of sensitivity equations, three for parameters attached to the state equation and the rest for initial value. They are summarized as follows:

- The intrinsic parameter $r_i, i = 1, 2, \cdots, N$, let $X_r := \frac{\partial X}{\partial r} = \left(\frac{\partial x_i}{\partial r_j}\right), i, j = 1, 2, \cdots, N$.

$$\frac{dX_r}{dt} = \frac{\partial F}{\partial X} X_r + \frac{\partial F}{\partial r}, X_r(t_0) = \text{zeros}(N, N). \tag{19}$$

- The carrying capacity $K_i, i = 1, 2, \cdots, N$, let $X_K := \frac{\partial X}{\partial K} = \left(\frac{\partial x_i}{\partial K_j}\right), i, j = 1, 2, \cdots, N$.

$$\frac{dX_K}{dt} = \frac{\partial F}{\partial X} X_K + \frac{\partial F}{\partial K}, X_K(t_0) = \text{zeros}(N, N). \tag{20}$$

- The interaction parameter $\alpha_{ij}, i \neq j, i, j = 1, 2, \cdots, N$, let $X_\alpha := \frac{\partial X}{\partial \alpha} = \left(\frac{\partial x_i}{\partial \alpha_{j,k}}\right) = \left(\frac{\partial x_i}{\partial \tau_j}\right), i = 1, 2, \cdots, N, j = 1, 2, \cdots, N \times (N-1)$.

$$\frac{dX_\alpha}{dt} = \frac{\partial F}{\partial X} X_\alpha + \frac{\partial F}{\partial \alpha}, X_\alpha(t_0) = \text{zeros}(N, N \times (N-1)). \tag{21}$$

- The initial value x_{0i}, $i = 1, 2, \cdots, N$, let $X_{x_0} := \frac{\partial X}{\partial x_0} = \left(\frac{\partial x_i}{\partial x_{0j}}\right), i, j = 1, 2, \cdots, N$.

$$\frac{dX_{x_0}}{dt} = \frac{\partial F}{\partial X} X_{x_0}, X_{x_0}(t_0) = I_{N\times N}. \tag{22}$$

It can be seen that the coefficient matrices in the sensitivity equations are almost diagonal so that it will be worthwhile in terms of computational complexity. In the following section, some numerical algorithms will be implemented using the 4-th order Runge-Kutta method. The other approach to parameter estimation related to the predator-prey equations was proposed in [14] by introducing a smoothing stage before the Runge-Kutta method applied.

4 Numerical Implementation

Consider the initial value problem (IVP) of single dimension $x' = f(t, x)$ dan $x(t_0) = x_0$ where $x : [t_0, \infty] \to \mathbf{R}$ and $f : [t_0, \infty] \times \mathbf{R} \to \mathbf{R}$. Assume the IVP has a unique, non-explicit solution, which is approximated by the fourth-order Runge-Kutta method. This method is widely used in practice because it is simple to implement and has a high enough convergence order. Instead of the infinite time horizon $[t_0, \infty]$, only the finite part $[t_0, t_f]$ is considered in the implementation. The method is carried out by discretizing the time domain $[t_0, t_f]$ as $t_j = t_{j-1} + h$, $j = 1, 2, 3, \cdots$, where h represents the mesh size, probably not uniform. For the sake of simplicity, assume that the discretization of time is given by $t_0, t_1, t_2, \cdots, t_M$ where $t_j - t_{j-1} := h_j$ is the lenght of $j^{th}-$subinterval. Let w_j be the approximation of x at $t = t_j$, i.e. $w_j \approx x(t_j)$. Specifically, $w_0 = x_0$. The fourth-order of Runge-Kutta method adheres to the following algorithm stages.

Algortihm 4.1: Scalar variable

- start: $w_0 = x_0$ (given)
- for $j = 1, 2, 3, \cdots, M$ calculate k_1, k_2, k_3, and k_4:
  - $k_1 = f(t_{j-1}, w_{j-1})$
  - $k_2 = f(t_{j-1} + \frac{h}{2}, w_{j-1} + \frac{h}{2} k_1)$
  - $k_3 = f(t_{j-1} + \frac{h}{2}, w_{j-1} + \frac{h}{2} k_2)$
  - $k_4 = f(t_{j-1} + h, w_{j-1} + h k_3)$
  - $w_j = w_{j-1} + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$.
- The collection of points $\{(t_j, w_j) : j = 0, 1, 2, 3, \cdots, M\}$ visualizes the trajectory that approximates the solution of IVP.

Consider the case $x(t) = [x_1(t), x_2(t), \cdots, x_N(t)]$ of $N-$dim and $f := [f_1, f_2, \cdots, f_N]$ where $x_i : [t_0, \infty] \to \mathbf{R}$ and $f_i : [t_0, \infty] \times \mathbf{R}^N \to \mathbf{R}^N$, $i = 1, 2, \cdots, N$. Let $w_j = [w_{1j}, w_{2j}, \cdots, w_{Nj}]$ be the approximation of $x(t_j) = [x_1(t_j), x_2(t_j), \cdots, x_N(t_j)]$, i.e. $w_{ij} \approx x_i(t_j)$, $i = 1, 2, \cdots, N$, $j = 0, 1, 2, 3, \cdots, M$. The Runge-Kutta method is applied to determine the vectors w_j where $w_j \approx x(t_j)$, $j = 0, 1, 2, \cdots, M$.

Algorithm 4.2: Vector variable

- start: $w_0 = [x_{10}, x_{20}, \cdots, x_{N0}]$ (given)
- for $j = 1, 2, 3, \cdots, M$ calculate vectors $k_1 = [k_{11}, k_{12}, \cdots, k_{1n}]$, $k_2 = [k_{21}, k_{22}, \cdots, k_{2n}]$, $k_3 = [k_{31}, k_{32}, \cdots, k_{3n}]$, and $k_4 = [k_{41}, k_{42}, \cdots, k_{4n}]$ as follows:
  - for $i = 1, 2, \cdots, N$:
    - $k_{1i} = f_i(t_{j-1}, w_{j-1})$
    - $k_{2i} = f_i(t_{j-1} + \frac{h}{2}, w_{j-1} + \frac{h}{2} k_1)$
    - $k_{3i} = f_i(t_{j-1} + \frac{h}{2}, w_{j-1} + \frac{h}{2} k_2)$
    - $k_{4i} = f_i(t_{j-1} + h, w_{j-1} + h k_3)$

- $w_j = w_{j-1} + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$.
- collect the output in some matrixes, e.g.

$$W_1 := \begin{bmatrix} w_{01} & w_{11} & \cdots & w_{M1} \\ w_{02} & w_{12} & \cdots & w_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{0N} & w_{1N} & \cdots & w_{MN} \end{bmatrix}$$

  where the rows correspond to coordinate of the state x and the columns indicate the time grid.

Let Z where $Z : [t_0, t_f] \to \mathbf{R}^{N\times p}$ be the unknown and $G : [t_0, t_f] \times \mathbf{R}^{N\times p} \to \mathbf{R}^{N\times p}$ is given. For a scalar matrix Z_0 of size $N \times p$, consider the following initial value problem:

$$\frac{dZ}{dt} = G(t, Z), t \in (t_0, t_f], Z(t_0) = Z_0. \tag{23}$$

This formulation incorporates the sensitivity equation introduced in the previous section. To begin constructing the algorithm, create a 3-dimensional array W of size $N \times p \times M$, i.e. $W = (w_{ijk})$, where $w_{ijk} \approx z_{ij}(t_k)$, $i = 1, 2, \cdots, N$, $j = 1, 2, \cdots, p$, and $k = 1, 2, \cdots, M$.

Algorithm 4.3: Matrix variable

- start: $W_0 = Z_0$ (given)
- for $k = 1, 2, 3, \cdots, M$ calculate matrices K_1, K_2, K_3, dan K_4 as follows:
  - $K_1 = G(t_{k-1}, W_{k-1})$
  - $K_2 = G(t_{k-1} + \frac{h}{2}, W_{k-1} + \frac{h}{2}K_1)$
  - $K_3 = G(t_{k-1} + \frac{h}{2}, W_{k-1} + \frac{h}{2}K_2)$
  - $K_4 = G(t_{k-1} + h, W_{k-1} + hK_3)$
  - $W_k = W_{k-1} + \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4)$.
- Extract $Z(t_k) \approx W_k$.

Surprisingly, the numerical procedure of the Runge-Kutta method for matrix variables is as simple as that for scalar variables. The sensitivity equations (3.10)-(3.13) must be solved simultaneously with the state equation (10), so there are four pairs of equations. On this occasion, the only intrinsic parameters r_i are derived in detail, others are treated similarly.

$$\frac{dx_i}{dt} = r_i x_i \left(1 - \frac{\sum_{j=1}^{N} \alpha_{ij} x_j}{K_i}\right), x_i(t_0) = x_{i0}, i = 1, 2, \cdots, N, \tag{24}$$

$$\frac{dX_r}{dt} = \frac{\partial F}{\partial X} X_r + \frac{\partial F}{\partial r}, X_r(t_0) = \text{zeros}(N, N), \tag{25}$$

where $\frac{\partial F}{\partial X}$ and $\frac{\partial F}{\partial r}$ are matrices defined by (15) and (16), respectively.

Algorithm 4.4

- for each $i = 1, 2, \cdots, N$, define $f_i(t, x) := r_i x_i \left(1 - \frac{\sum_{j=1}^{N} \alpha_{ij} x_j}{K_i}\right)$.
- apply Algorithm 4.2 to solve (24).
- store the elements matrix $W_1 := (w_{ij})$ where $w_{ij} \approx x_i(t_j)$.
- define matrix functions $G(t, Z)$ as $G(t, Z) = \left[\frac{\partial F}{\partial X}(t)\right] Z + \left[\frac{\partial F}{\partial r}(t)\right]$
- apply Algorithm 4.2 to solve (16).
- $X_r(t_k) = \frac{\partial X}{\partial r}(t_k) \approx Z(t_k)$, $k = 1, 2, \cdots, M$.

The elements of Jacobian $\frac{\partial F}{\partial X} = \left(\frac{\partial f_i}{\partial x_j}(t_k)\right)$, $i, j = 1, 2, \cdots, N$ at $t = t_k$, $k = 1, 2, \cdots, M$ are defined by evaluating (15) at $t = t_k$, i.e.

$$\frac{\partial f_i}{\partial x_j}(t_k) = \begin{cases} r_i \left(1 - \frac{\sum_{\ell=1}^{N} \alpha_{i\ell} x_\ell(t_k) + x_i(t_k)}{K_i}\right), & j = i \\ -\frac{r_i \alpha_{ij} x_i(t_k)}{K_i}. & j \neq i. \end{cases}$$

Likewise, the elements of $\frac{\partial F}{\partial r}$ are obtained according to

$$\frac{\partial F}{\partial r}(t_k) = \text{diag}\left[x_i(t_k)\left(1 - \frac{\sum_{\ell=1}^{N} \alpha_{i\ell} x_\ell(t_k)}{K_i}\right)\right].$$

where $i = 1, 2, \cdots, N$. In the numerical realization, the values $x_i(t_k)$ are substituted by w_{ik} obtained from (24).

Example 1

The special case of competitive model is when $N = 2$ which is described by the Lotka-Volterra equation.

$$\frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1 + \alpha_{12} x_2}{K_1}\right), t \in (t_0, t_f]\, x_1(t_0) = x_{10}$$
$$\frac{dx_2}{dt} = r_2 x_2 \left(1 - \frac{\alpha_{21} x_1 + x_2}{K_2}\right), t \in (t_0, t_f]\, x_2(t_0) = x_{20}.$$

Since $N = 2$, then there will be $N^2(N + 2) = 16$ sensitivity functions corresponding to the sensitivity matrix represented in the matrix $\frac{\partial X}{\partial \theta}$ given by

$$\begin{bmatrix} \frac{\partial x_1}{\partial r_1} & \frac{\partial x_1}{\partial r_2} & \frac{\partial x_1}{\partial \alpha_{12}} & \frac{\partial x_1}{\partial \alpha_{21}} & \frac{\partial x_1}{\partial K_1} & \frac{\partial x_1}{\partial K_2} & \frac{\partial x_1}{\partial x_{10}} & \frac{\partial x_1}{\partial x_{20}} \\ \frac{\partial x_2}{\partial r_1} & \frac{\partial x_2}{\partial r_2} & \frac{\partial x_2}{\partial \alpha_{12}} & \frac{\partial x_2}{\partial \alpha_{21}} & \frac{\partial x_2}{\partial K_1} & \frac{\partial x_2}{\partial K_2} & \frac{\partial x_1}{\partial x_{10}} & \frac{\partial x_1}{\partial x_{20}} \end{bmatrix}.$$

It will be beneficial to partition the sensitivity matrix $\frac{\partial X}{\partial \theta}$ into parameter groups.

$$\frac{\partial X}{\partial r} = \begin{bmatrix} \frac{\partial x_1}{\partial r_1} & \frac{\partial x_1}{\partial r_2} \\ \frac{\partial x_2}{\partial r_1} & \frac{\partial x_2}{\partial r_2} \end{bmatrix}, \frac{\partial X}{\partial \alpha} = \begin{bmatrix} \frac{\partial x_1}{\partial \alpha_{12}} & \frac{\partial x_1}{\partial \alpha_{21}} \\ \frac{\partial x_2}{\partial \alpha_{12}} & \frac{\partial x_2}{\partial \alpha_{21}} \end{bmatrix},$$

$$\frac{\partial X}{\partial K} = \begin{bmatrix} \frac{\partial x_1}{\partial K_1} & \frac{\partial x_1}{\partial K_2} \\ \frac{\partial x_2}{\partial K_1} & \frac{\partial x_2}{\partial K_2} \end{bmatrix}, \frac{\partial X}{\partial x_0} = \begin{bmatrix} \frac{\partial x_1}{\partial x_{10}} & \frac{\partial x_1}{\partial x_{20}} \\ \frac{\partial x_2}{\partial x_{10}} & \frac{\partial x_2}{\partial x_{20}} \end{bmatrix}.$$

The following results were obtained using simulation data $t_0 = 0, t_f = 150, x_{10} = 10; x_{20} = 30, r_1 = 0.1, r_2 = 0.3, a_{12} = 0.3,$

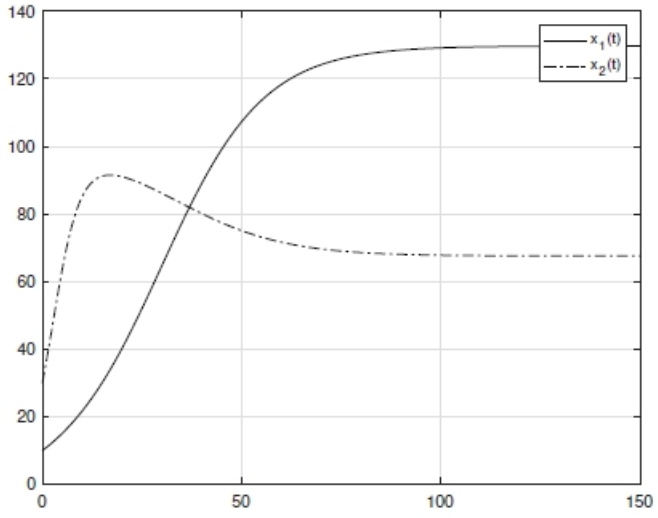
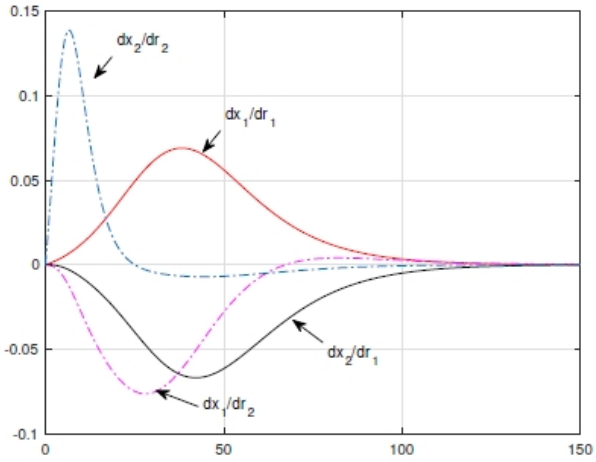


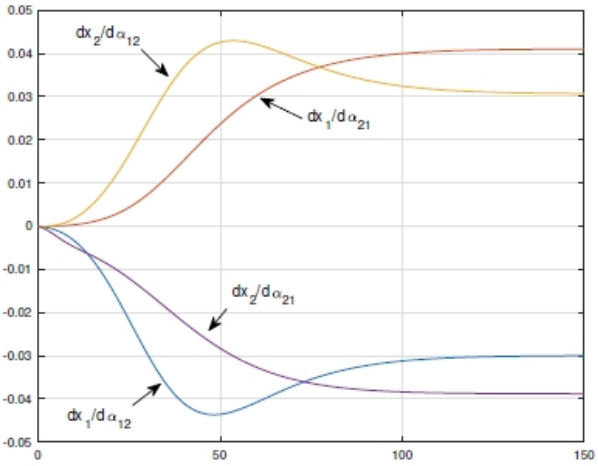
Figure 3. States trajectory

$a_{21} = 0.25$, $K_1 = 150$, and $K_2 = 100$. The trajectories of states are presented in Figure 3. Initially, the number of species 1 ($x_{10} = 10$) is less than the number of species 2 ($x_{20} = 30$), however species 1 gradually outnumbers species 2. This is primarily due to the fact that species 1 has a bigger carrying capacity than species 2, i.e. $K_1 = 150$ and $K_2 = 100$. In fact, neither species' maximum capacities is ever realized. Obviously, the growth of species 1 directly depends on the parameters r_1, α_{12}, and K_1 and implicitly depends on the parameters involved in species 2, specifically r_2, α_{21}, and K_2.

The graph of states is presented in Figure 3 and the normalized sensitivity functions are displayed in Figure 4. They describe how these parameters affect each species. The sensitivity of state x_i with respect to parameter θ_j is denoted by $\frac{\partial x_i}{\partial \theta_j}$. The greater the value of $\frac{\partial x_i}{\partial \theta_j}$, the more impact of parameter θ_j to x_i. The value $\frac{\partial x_i}{\partial \theta_j}$ close to zero within an interval indicates that the parameter θ_j has little effect on state x_i. For example, Figure 4(a) shows that each parameter influences both states x_1 and x_2, though their timing and duration may differ. For instance, based on the curve of $\frac{dx_1}{dr_1}$, the effect of parameter r_1 with respect to x_1 occurs significantly in $[0, 100]$. Similar patterns for the parameter r_1 to x_2 and r_2 to x_1. On the other hand, the parameter r_2 has a significant effect on x_2 only on a small part of interval at beginning, around $[0, 20]$, otherwise it is close to zero. Comparing Figure 4(a) and 4(b), it can be seen that the patterns of $\frac{\partial X}{\partial r}$ and $\frac{\partial X}{\partial x_0}$ are similar and so are $\frac{\partial X}{\partial K}$ and $\frac{\partial X}{\partial \alpha}$. Furthermore, the pattern of those curves will be used as a qualitative tool in the optimal design of experiment.



(a) Sensitivity functions $X_r := \frac{\partial X}{\partial r}$.

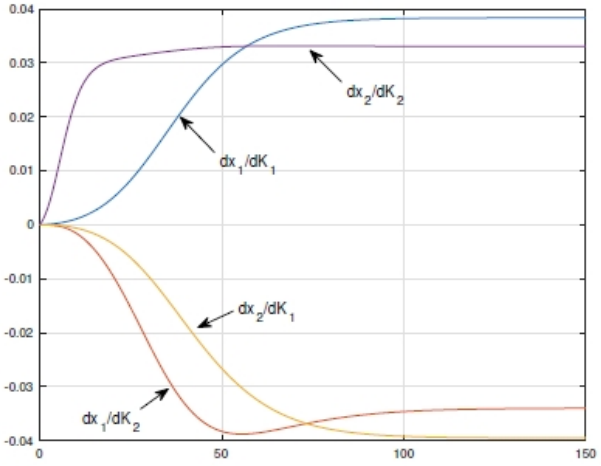


(b) Sensitivity functions $X_\alpha := \frac{\partial X}{\partial \alpha}$.

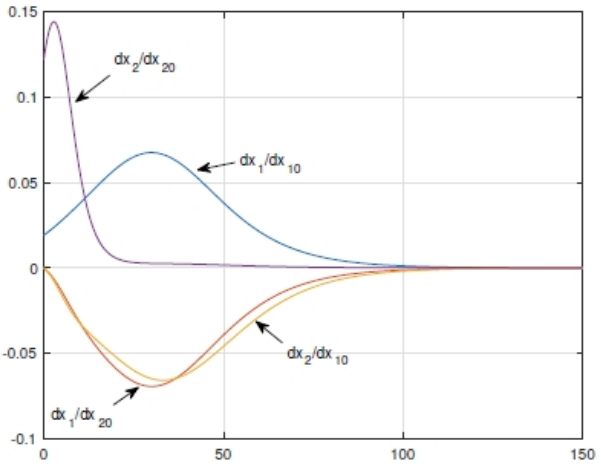
Figure 4. Graph of sensitivity functions with respect to parameter r and α.

5 Application to Parameter Estimation

The problem to be addressed in this section is how to select a sample with a high information content about the parameters to be estimated. This is known as the optimal design of experiment, and it is a common challenge in many domains of applied science, for instance [6, 7, 15, 20, 28, 29]. Some optimal crite-



(a) Sensitivity functions $X_K := \frac{\partial X}{\partial K}$.



(b) Sensitivity functions $X_{x_{x_0}} := \frac{\partial X}{\partial x_{x_0}}$.

Figure 5. Graph of sensitivity functions with respect to parameter K and x_0.

ria were based on the functional of FIM [24, 25, 5]. Here, the sensitivity functions behavior will be utilized to make the selection [2, 3, 22]. The prior sensitivity function is known as the traditional sensitivity function (TSF). Another criterion required for this purpose is known as the generalized sensitivity function (GSF).

In the nonlinear dynamical system

$$\frac{dX}{dt} = F(t, X; \theta), X(t_0) = X_0, \tag{26}$$

where θ is the parameter system, suppose that a discrete time observation is represented as

$$Y_j = PX(t_j; \theta) + \epsilon_j, j = 1, 2, \cdots, n, \tag{27}$$

where P is the matrix of measurement process and the observation errors ϵ_j are assumed to be independently identically distributed (iid) Gaussian noise with zero mean. The problems of finding the optimal sensor location and parameter estimation with non-Gaussian model noise had been considered in [23].

Assume that there are M observation coordinates, and since the state X consists of coordinates $x_1, x_2, \cdots, x_N$, the matrix P should be of size $M \times N$. Hence, the measurement could be expressed explicitly as

$$\begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{Mj} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N} \\ c_{21} & c_{22} & \cdots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M1} & c_{M2} & \cdots & c_{MN} \end{bmatrix} \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{Nj} \end{bmatrix} + \begin{bmatrix} \epsilon_{1j} \\ \epsilon_{2j} \\ \vdots \\ \epsilon_{Mj} \end{bmatrix}. \tag{28}$$

For example, the first measurement coordinate at time t_j given by $y_{1j} = \sum_{\ell=1}^{N} c_{1\ell} x_\ell + \epsilon_{1j}$, i.e. the observation data are represented in terms of the state variable X. A special case of (28) is when $M = N$ and P is an identity matrix, it reduces the standard form

$$y_{kj} = x_{kj} + \epsilon_{kj},$$

where $k = 1, 2, \cdots, N$ denotes the coordinate of state, $j = 1, 2, \cdots, n$ stands for the points of measurement, and $x_{kj} := x_k(t_j; \theta)$. Let σ_i^2 be the variance associated with the coordinate of the error ϵ_i, $i = 1, 2, \cdots, M$, i.e. $\epsilon_j \sim \mathcal{N}(0, V)$ where $V = \text{diag}(\sigma_i^2, i = 1, 2, \cdots, M)$.

The fundamental assumption underlying parameter estimation is the existence of a nominal (true) parameter θ_0 that characterizes the system. The least squares approach to estimate θ_0 is performed by seeking the minimizer of the cost functional

$$J(\theta) := \sum_{j=1}^{n} \left[Y_j - PX(t_j; \theta)\right]^T V^{-1} \left[Y_j - PX(t_j; \theta)\right]. \tag{29}$$

For $M = N = 2$, the cost functional becomes

$$J(\theta) := \sum_{j=1}^{n} \left[\frac{1}{\sigma_1^2(t_j)}(y_{1j} - x_{1j})^2 + \frac{1}{\sigma_2^2(t_j)}(y_{2j} - x_{2j})^2\right]. \tag{30}$$

The simulation assumes that the observation variances for each measurement are similar, i.e. $\sigma_1^2(t_j) = \sigma_2^2(t_j)$ for all $j = 1, 2, \cdots, n$, therefore the cost functional now is simply to

$$J(\theta) := \sum_{j=1}^{n} \left[(y_{1j} - x_{1j})^2 + (y_{2j} - x_{2j})^2\right] = \sum_{j=1}^{n} \epsilon_j^T \epsilon_j, \tag{31}$$

where $\epsilon_j = [y_{1j} - x_{1j}, y_{2j} - x_{2j}]^T$ is the error vectors at $t = t_j$. The simplest form of cost functional when the state consists of the single coordinate.

$$J(\theta) := \sum_{j=1}^{n} |y_j - x_j|^2 , \tag{32}$$

where $x_j = x(t_j; \theta)$. Furthermore, the Fisher information matrix (FIM) is defined by

$$F = \sum_{j=1}^{n} \frac{1}{\sigma^2(t_j)} \left[\nabla_\theta X(t_j; \theta_0)\right] \left[\nabla_\theta X(t_j; \theta_0)\right]^T , \tag{33}$$

where

$$\nabla_\theta X(t_j; \theta_0) = \left[\begin{array}{cccc} \frac{\partial X(t_j;\theta_0)}{\partial \theta_1} & \frac{\partial X(t_j;\theta_0)}{\partial \theta_2} & \cdots & \frac{\partial X(t_j;\theta_0)}{\partial \theta_p} \end{array} \right]^T , \tag{34}$$

$$\frac{\partial X(t_j; \theta_0)}{\partial \theta_k} = \left[\begin{array}{cccc} \frac{\partial x_1(t_j;\theta_0)}{\partial \theta_k} & \frac{\partial x_2(t_j;\theta_0)}{\partial \theta_k} & \cdots & \frac{\partial x_N(t_j;\theta_0)}{\partial \theta_k} \end{array} \right]^T . \tag{35}$$

As a consequence of the term $\nabla_\theta X(t_j; \theta_0)$ being a column vector of dimension $N + p$, the FIM is a symmetric squared matrix of dimension $N + p$. The FIM is assumed to be invertible. The FIM measures the information content of the data corresponding to the model parameters [5]. The generalized sensitivity function (GSF) is defined as

$$G(t_\ell) = \sum_{j=1}^{\ell} \frac{1}{\sigma^2(t_j)} \left[F^{-1} \times \nabla_\theta(t_j; \theta_0)\right] \bullet \left[\nabla_\theta(t_j; \theta_0)\right], \tag{36}$$

where the notation "$\bullet$" stands for element-wise vector multiplication. The actual information is associated with the rate of change of GSF, and thus sharp increases of GSF indicate a high information about parameters [22, 5]. For the sake of simplicity, and assuming that the variances are uniform across the observations, the following step-wise algorithms are required to define the GSF.

Algorithm to define the vector $\nabla_\theta X(t_j \theta_0)$

- for $k = 1 : p$
  - for $\ell = 1 : N$

$D_j(k + \ell - 1) = \frac{\partial x_\ell(t_j;\theta_0)}{\partial \theta_k}$

Algorithm to define the FIM F

- prepare a squared matrix of size $(N + p)$, i.e. $F = zeros(N + p)$.
- for $j = 1 : n$
  - $F_0 = D_j D_j^T$,
  - $F = F + F_0$.

Algorithm to define the GSF G

- prepare a column vector of size $N+p$, i.e. $G = zeros(N + p, 1)$.
- for $j = 1 : \ell$
  - define the vector H_j of size $N + p$ by $H_j = F^{-1} D_j$,
  - $G_0 = H_j \bullet D_j := [H_j(k) D_j(k), k = 1, 2, \cdots, N + p]^T$,
  - $G = G + G_0$.

It is readily seen from the definition that the GSFs are the vector-valued functions of dimension $N+p$ where the $k-$th component g_k of G represents the generalized sensitivity function of specific state coordinate with respect to certain parameter. For instance, for $N = 2$ and $p = 3$, there will be 6 sensitivity functions as shown in the following vector.

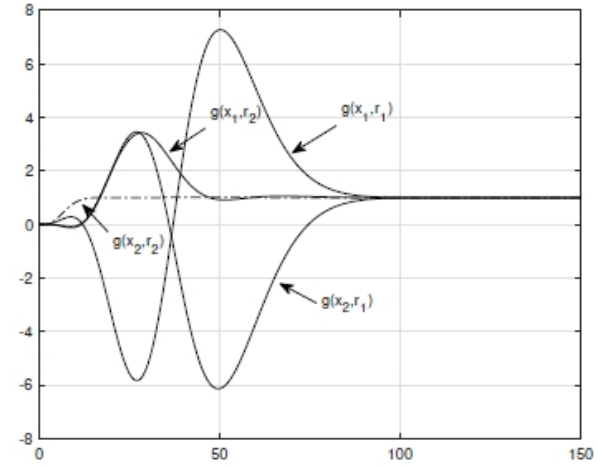
$$G = [g(x_1, \theta_1), g(x_2, \theta_1), g(x_1, \theta_2), g(x_2, \theta_2), g(x_3, \theta_3), g(x_3, \theta_3)]^T$$

where $g(x_i, \theta_j)$ specifies the generalized sensitivity function of x_i with respect to θ_j.

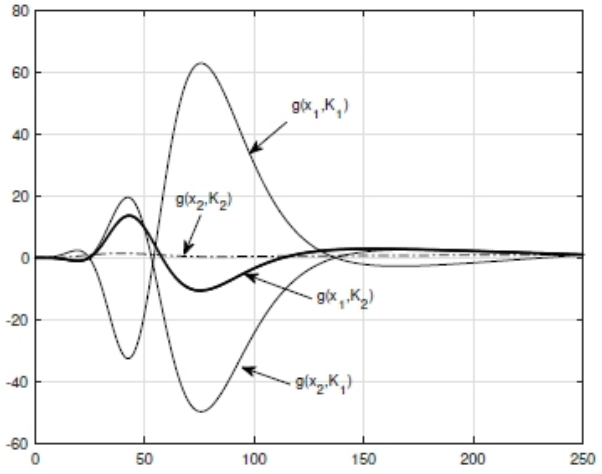
Example 2

Consider back to the competitive model with $N = 2$ as given in Example 4.1. The eight parameters contributing in the model are $\theta_1 := r_1$, $\theta_2 := r_2$, $\theta_3 := \alpha_{12}$, $\theta_4 := \alpha_{21}$, $\theta_5 := K_1$, $\theta_6 := K_2$, $\theta_7 := x_{10}$, and $\theta_7 := x_{20}$, and therefore there will be 16 generalized sensitivity functions $g(x_i, \theta_j)$, $i = 1, 2$; $j = 1, 2, \cdots, 8$. The following simulation uses data $x_{10} = 10$; $x_{20} = 30$, $r_1 = 0.1$, $r_2 = 0.3$, $a_{12} = 0.3$, $a_{21} = 0.25$, $K_1 = 150$, and $K_2 = 100$. The trajectories of GSFs are presented in Figure 6 and 7. The time horizon $[0, t_f]$ is adaptively extended until the steady state is reached.

This simulation suggests that the information content about parameters may differ among state coordinates and along measurements. Take a look at Figure 6(a), both x_1 and x_2 provide a comparable degree of information about r_1 and the most information given by observations when $t < 90$. On the other hand, the most information about r_2 is given by x_1 when $t < 50$ and by x_2 when $t < 10$. Similar phenomena are given by parameters K and x_0 as shown in Figure 6(b). The pattern of the parameter α is considerably different. Even after extending the time span to $t = 300$, the GSFs corresponding to α still do not reach the steady state as displayed in Figure 7(a). This means that information about the α parameter is dispersed throughout the observation. However, the sharp increases of GSFs were occurred in the range $[50, 100]$. This indicates that the state variables contain the most information about the parameters in this interval. Figure 7(b) shows the sensitivity of state with respect to carrying capacity K. The patterns are similar to sensitivity of state with respect to intrinsic growths r's, i.e. the most information is provided in $[0, 90]$. To verify this claim, the following example examines the observation data for parameter estimation which are chosen from the range for which GSFs increase sharply as well as the range where steady state had been reached.

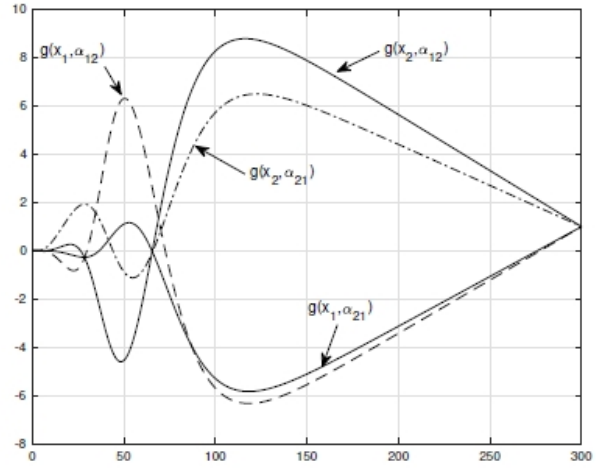


(a) Generalized sensitivity functions $g(X, r)$.

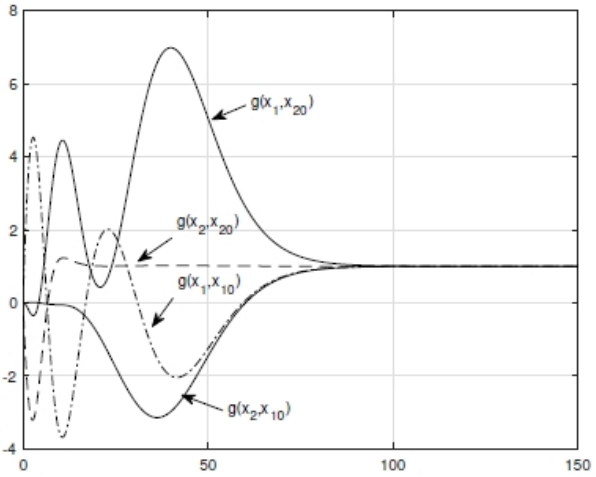


(b) Generalized sensitivity functions $g(X, K)$.

Figure 6. Graph of generalized sensitivity functions with respect to parameter r and K.



(a) Generalized sensitivity functions $g(X, \alpha)$.



(b) Generalized sensitivity functions $g(X, x_0)$.

Figure 7. Graph of generalized sensitivity functions with respect to parameter α and x_{x_0}.

Example 3

This example illustrates a number of issues concerning the information content of the parameters included in the observation data. The pertinent questions are when the measurements should be taken so that the observation data contains as much information about the parameters to be estimated as possible, how much minimum data required to estimate these parameters, how parameter content of one coordinate to be contained in another, and how to select the measurement moments using TSF and GSF. The simulation here uses data: $x_{10} = 10$; $x_{20} = 30$, $r_1 = 0.1$, $r_2 = 0.3$, $a_{12} = 0.3$, $a_{21} = 0.25$, $K_1 = 150$, and $K_2 = 100$.

Experiment 1

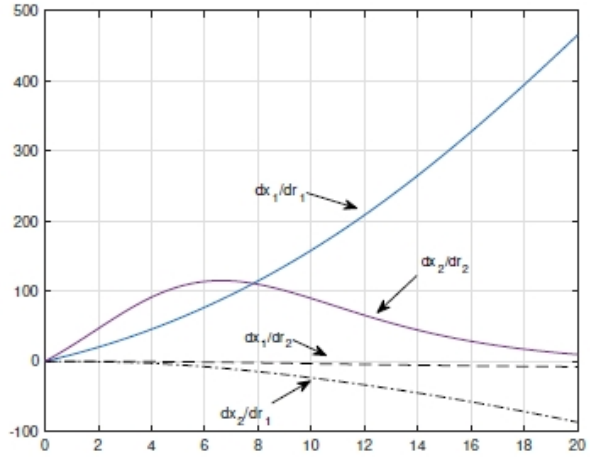
The parameters to be estimated here are r_1 and r_2 using data taken from intervals $[0, 20]$ and $[90, 120]$. Six evenly distributed data points from each interval were selected for estimation. The least squares error approach applies three criteria functions LS_1, LS_2, and LS_3 defined by $LS_1(\theta) = \sum_{j=1}^{n} |x_{1j} - y_{1j}|^2$, $LS_2(\theta) = \sum_{j=1}^{n} |x_{2j} - y_{2j}|^2$, and $LS_3(\theta) = LS_1(\theta) + LS_2(\theta) = \sum_{j=1}^{n} \left(|x_{1j} - y_{1j}|^2 + |x_{2j} - y_{2j}|^2\right)$ where for each $k = 1, 2$; $j = 1, 2, \cdots, n$, the x_{kj} and y_{kj} are the model outputs and observations including disturbance terms, respectively. The experiment results are summarized in the following table.

Table 2. Result of experiment 1: the nominal parameters were set to be $r_1 = 0.1$ and $r_2 = 0.3$.

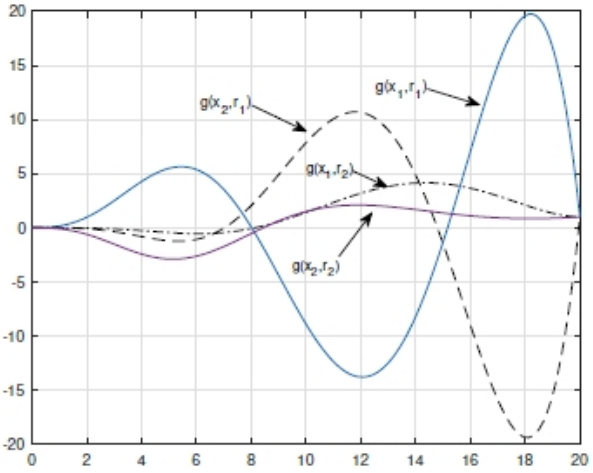
Domain	LS_1	LS_2	LS_3
$[90, 120]$	$(0.100, 0.100)$	$(3.540, 18.343)$	$(0.0012, 0.095)$
$[0, 20]$	$(0.0986, 0.2344)$	$(0.0999, 0.3002)$	$(0.1000, 0.3007)$

First, consider the estimates arising from $[90, 120]$. None of the estimates given by the three criteria are satisfactory. This indicates that the observation data obtained from this interval contains insufficient information about the parameters r_1 and r_2. This corresponds to the graph of GSF shown in Figure 6(a), which shows that the GSF curve no longer changes after $t > 90$. Next, consider the estimates obtained from the interval $[0, 20]$. The first criterion, LS_1, provides a good estimate for r_1 but a poor estimate for r_2. The second criterion, LS_2, on the other hand, yields very good estimates for both r_1 and r_2. This means

that the data obtained through observation of x_1 contains less information about parameter r_2 than x_2 about parameter r_1. The best result is given by LS_3. It means that the observation from interval $[0, 20]$ contains a lot of information about the parameters r_1 and r_2. The traditional and generalized sensitivity function curves shown in Figure 8 indicate the similar behavior, i.e. $\left|\frac{dx_1}{dr_2}\right| \ll \left|\frac{dx_2}{dr_1}\right|$ (left) and $|g(x_1, r_2)| \ll |g(x_2, r_1)|$ (right). These circumstances confirmed that the information content of parameters delivered by observation data of states could be identified through their sensitivity functions.

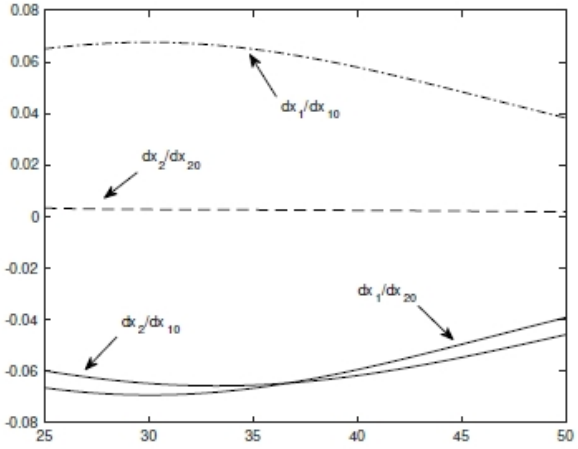


(a) Traditional sensitivity functions $X_r := \frac{\partial X}{\partial r}$.

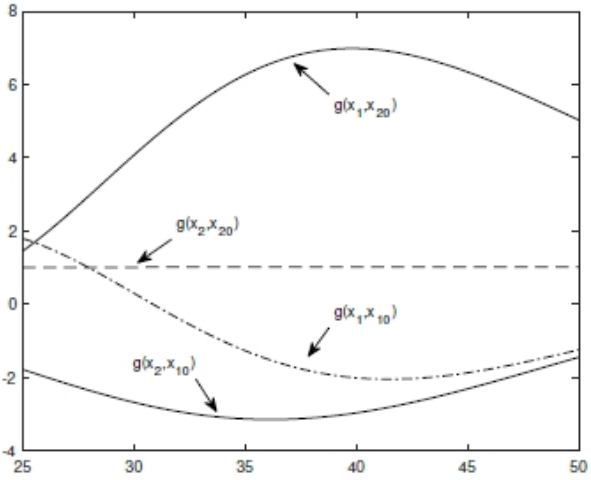


(b) Generalized sensitivity functions $g(X, r)$.

Figure 8. Graph of two kinds sensitivity functions with respect to parameter r in the same domain.



(a) Traditional sensitivity functions $X_{x_0} := \frac{\partial X}{\partial x_{x_0}}$.



(b) Generalized sensitivity functions $g(X, x_0)$.

Figure 9. Graph of two kinds sensitivity functions with respect to parameter x_0 in the same domain.

collected in Table 3. Recall that criteria LS_1 relates to information given by state x_1, LS_2 by x_2, and LS_3 by both x_1 and x_2. According to this table, state 1 contains much information for parameter x_{10} and x_{20}, state 2 contains a lot of information for x_{20} but less for x_{10}.

Table 3. Result of experiment 2: Estimates of parameter x_{x_0} given by three criteria, the true values $x_{10} = 10$, $x_{20} = 30$.

Domain	LS_1	LS_2	LS_3
$[25, 50]$	$(10.07, 30.32)$	$(12.54, 30.12)$	$(10.11, 29.96)$

Experiment 2

The initial conditions $x_0 = (x_{10}, x_{20})$ will be considered as parameters in this experiment where data are collected from interval $[25, 50]$. Figure 9 shows that $|\frac{dx_2}{dx_{20}}|$ and $|g(x_2, x_{20})|$ are significantly smaller than others. This fact does not mean that the state x_2 contains less information about parameter x_{20}. Intuitively, x_2 should contain much information about parameter x_{20} because it directly affects the state x_2. Results of experiment are

6 Concluding Remarks

The derivation of the sensitivity equations of a nonlinear parameterized dynamic system generates a system of linear differential equations in the form of a matrix whose size equals to the product of the number of state variables and parameters involved, plus a set of equations relating to the initial values. The sensitivity equations must be solved simultaneously with the original

state equations, and its solution yields the so-called traditional sensitivity functions (TSF). The traditional sensitivity functions (TSF) are then utilized to create the Fisher information matrix (FIM), which is subsequently used to define the generalized sensitivity function (GSF). These two kinds of sensitivity functions are employed in an optimal measurement design for selecting the sample that provides the most information about the parameters to be estimated.

Numerical experiments on the competitive model were conducted employing two types of sensitivity functions to perform data measurement for parameter estimation. According to those numerical experiments, the information content of the parameters varies along the measurement samples. The information content is determined by state coordinates employed and interval where measurements are taken place. Both kinds of sensitivity functions TSF and GSF have been shown to be the effective benchmark for determining the measurement interval, ensuring that the data acquired contains the most information of parameters. In parameter estimation, it is recommended to use data that provides a lot of information about these parameters.

Acknowledgements

This work was supported and financed by Human Resources Agency, Ahmad Dahlan University under Grant U7/152.67/B.6/III/2022 (email: sdm@uad.ac.id)

REFERENCES

[1] R. C. Aster, B. Borchers, C. H. Thurber. Parameter Estimation and Inverse Problems (second edition), Elsevier Academic Press, 2012.

[2] H.T., Banks, S. Dediu, S. L. Ernstberger. Sensitivity functions and their uses in inverse problems. *Journal of Inverse and Ill-Posed Problems*, **15 (7)**, 683–708, 2007.

[3] H. T. Banks, S. Dediu, S. L. Ernstberger, F. Kappel, F. Generalized sensitivities and optimal experimental design. *Journal of Inverse and Ill-Posed Problems*, **18 (1)**, 25–83, 2010.

[4] H. T. Banks, A. Cintron-Arias. Parameter Selection Methods in Inverse Problem Formulation. *arXiv: Quantitative Methods*, 43–73, 2020.

[5] H. T. Banks, A. Holm, F. Kappel. Comparison of Optimal Design Methods in Inverse Problems. *Inverse Problem* **27 (7)**, 1–46, 2011.

[6] J. Berger, D. Dutykh, N. Mendes. On the optimal experiment design for heat and moisture parameter estimation. *Experimental Thermal and Fluid Science*, **81**, 109–122, 2017.

[7] S. Bhonsale, P. Nimmegeers, S. Akkermans, D. Telen, I. Stamati, F. Logist, J. F. M. Van Impe. Chapter 10 - Optimal experiment design for dynamic processes, Editor(s): Michael Bortz, Norbert Asprion, Simulation and Optimization in Process Engineering, Elsevier, 243–271, 2022.

[8] R. L. Borrelli, C. S. Coleman. Differential Equation: A Modeling Perspective (2nd Edition), Willey, New York, 2004.

[9] M. Fan, P. J. Y. Wong, R. P. Agarwal. Periodicity and stability in periodic n-species Lotka-Voltera competition system with feedback control and deviating arguments. *Acta Mathematica Sinica, English Series*, **19(4)**, 801–822, 2003.

[10] F. Gao, L. Han. Implementing the Nelder-Mead simplex algorithm with adaptive parameters, *Computational Optimization and Applications*, **51**, 259–277, 2012.

[11] Goodwin, G. C., and Payne, R. L.: Dynamic system identification: experiment design and data analysis, Academic Press, New York, 1977.

[12] M. Grasselli. An inverse problem in population dynamics, *Numerical Functional Analysis and Optimization*, **18**: 311-321,1997.

[13] Y. Hao, M. Guo. Parameter Estimation of the Lotka–Volterra Model with Fractional Order Based on the Modulation Function and Its Application, *Mathematical Problems in Engineering*, 2021. Available at https://doi.org/10.1155/2021/6645059.

[14] T. Hu, Y. P. Qiu, H. J. Cui, L. H, Chen. Numerical discretization-based kernel type estimation methods for ordinary differential equation models. *Acta Mathematica Sinica, English Series* **31**, 1233–1254, 2015.

[15] Y. Jung, I. Lee. Optimal design of experiments for optimization-based model calibration using Fisher information matrix, *Reliability Engineering & System Safety*, **216**, 2021. Available at https://doi.org/10.1016/j.ress.2021.107968.

[16] A. D. Martins, E. Gjini. Modeling Competitive Mixtures With the Lotka–Volterra Framework for More Complex Fitness Assessment Between Strains. *Front. Microbiology, Sec. Systems Microbiology*, **22 September 2020**. Available at https://doi.org/10.3389/fmicb.2020.572487.

[17] P. Mathe, S. V. Pereverzev. Complexity of linear ill-posed problems in Hilbert space. *Journal of Complexity*, **38**: 50–67, 2017.

[18] J. D. Murray. Mathematical Biology: I. An Introduction, Third Edition, Springer-Verlag, Berlin Heidelberg, 2002.

[19] J. R. Raol, G. Girija, J. Singh. Modeling and Parameter Estimation Dynamic System, The Institution of Engineering and Technology, London, United Kingdom, 2004.

[20] A. Sinkoe, J. Hahn. Optimal experimental design for parameter estimation of an IL-6 signaling model. *Processes* **5 (3)**, 1–12, 2017.

[21] A. Tarantola. Inverse Problem Theory and Methods for Model Parameter Estimation, Philadelphia: SIAM, 2005.

[22] K. Thomaseth, C. Cobelli. Generalized Sensitivity Functions in Physiological System Identification. *Annals of Biomedical Engineering*, **21 (5)**, 607–616, 1999.

[23] F. E. Uilhoorn. An algorithm for finding optimal sensor placement and model uncertainty in data assimilation applied to Riemann problems. *Applied Mathematical Modeling*, **103** 649–673, 2022.

[24] D. Ucinski. Optimal Measurement Methods for Distributed Parameter System Identification, Boca Raton, Florida: CRC Press, 2005.

[25] D. Ucinski. E-optimum sensor selection for estimation of subsets of parameters. *Measurement*, **187**, 110286, 2022.

[26] J. A. Vano, J.C. Wildenberg, M. B. Anderson, J. K. Noel, J. C. Sprott. Chaos in low-dimensional Lotka–Volterra models of competition. *Nonlinearity IOP Publishing*, **19 (10)**, 2391–2404, 2006.

[27] C. R. Vogel. Computational Methods for Inverse Problems, Philadelphia: SIAM, 2002.

[28] B. Yu, X. Jiang, H. Qi. Numerical method for the estimation of the fractional parameters in the fractional mobile/immobile advection–diffusion model. *International Journal of Computer Mathematics*, **95**, 1131–1150, 2019.

[29] Y. Zhang, S. Chai, S., C. An, F. Lim, M. Duan. A new method for optimal sensor placement considering multiple factors and its application to deepwater riser monitoring systems. *Ocean Engineering*, **244**, 110403, 2021.