Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 7, pp. 1277–1299
DOI : 10.47974/JSMS-1019

Global sensitivity analysis in the SIHR epidemiological model with application to COVID-19

Liban Ismail *
Hacène Djellout †
Cédric Chauvière §
CNRS UMR 6620
Université Clermont Auvergne
Campus Universitaire des Cézeaux
3 place Varsarely
TSA 60026
CS 60026
63178 Aubière Cedex
France

Abstract

This paper is devoted to analyze the sensitivity of input parameters (basic reproduction number, cure rate, hospitalization rate) on the evolution of the SIHR epidemiological model. Lagrange polynomials affords a convenient way for computing Sobol indices. Those indices furnish important information on the relative weight of each uncertain input parameters. We use the COVID-19 disease as a numerical application to illustrate the study for two different basic reproduction numbers $\mathcal{R}_0$. Furthermore, we consider a sinusoidal and logistic infection rate and we give the Sobol indices when the entry parameters are assumed to follow the uniform law.

Subject Classification: *(2010) 62P10, 62J10, 92D30, 34D23.*

Keywords: *SIHR model, Lagrange polynomial, Stochastic collocation method, Sobol indices, Simulations, COVID-19.*

* *E-mail:* **liban_ismail_abdillahi@univ.edu.dj (Corresponding Author)**
† *E-mail:* **hacene.djellout@uca.fr**
§ *E-mail:* **cedric.chauviere@uca.fr**

1. Introduction

The most frequent way to model epidemic spread in populations consist in classifying people into various population groups. Those models are governed by a system of differential equations that follow the population over time, dividing them into many groups based on their infection status. The Kermack-Mckendric SIR model (susceptible-infectious-removed) is a very well established model and is widely used for various epidemics [15]. To do this, the SIR, SIRS and SEIR models have been developed which highlight the crucial role played by the $\mathcal{R}_0$ parameter, describing the average number of new infections due to a sick individual [21, 30]. We are interested in the SIHR model obtained from the classical SIR epidemic model from which a hospitalized compartment is added. The SIHR model appears in several research works [14, 27]. The course of a disease depends on several parameters (infection rate, cure rate, hospitalization rate, vaccination, etc.).

It is possible to group the sensitivity analysis methods into three classes: screening methods, which consist of a qualitative analysis sensitivity of the output variable to the input variables, local analysis methods [28], which quantitatively assess the impact of a small variation around a given value of the inputs and finally the global sensitivity analysis methods [1, 20, 26] which are interested in the variability of the output of the model in the whole of its range of variation. Global sensitivity analysis studies how the variability of inputs affects that of output, by determining how much of the variance of output is due to a given input or set of inputs.

Indeed, by studying how the response of the model reacts to variations in its input variables, the sensitivity analysis makes it possible to answer a number of questions: what are the variables that most contribute to the variability of the model response? What are the least influential variables? Which variables, or which groups of variables, interact with which others? To answer these questions, we use the Sobol indices. The sensitivity analysis is used in different mathematical models [1, 20, 32]. A global sensitivity analysis based on the chaos polynomial method is illustrated in several works [7, 24, 31] and a sensitivity analysis using the Taylor series method in [2]. Several articles [7, 14] have studied the sensitivity analysis of a nonautonomous system and with random parameters.

Among the significant number of papers dealing with the COVID-19 disease, a large proportion of them are dedicated to predict the number of infected people. For example, we can quote [25] that use a machine

learning model (namely a Support Vector Machine model) to accurately predict the spread of the disease. It is also worth noticing that in the descriptive analysis performed in [3] on Indians people, the authors show that the age of the infected people is not the most significant factor with respect to the fatality of the disease.

The objective of this paper is to develop the stochastic collocation method based on Lagrange polynomials to study the sensitivity analysis in the input parameters of the nonautonomous SIHR epidemiological model on the course of the disease with an infection rate depending on time. The main assumption is that these parameters are not known and they can therefore be modeled under the form of random variables having known laws. To compute the Sobol indices, applied to the numerical approximation of differential equations, we use the stochastic collocation method based on the Lagrange polynomials [6]. In [16], the authors used the stochastic collocation method based on Legendre polynomials. The solution of the differential equation is represented by means of orthogonal's polynomials. The coefficients of the polynomial basis are functions of time and can be calculated by solving a system of deterministic ordinary differential equations.

The paper is organized as follows: in section 2, we present the model and study the evolution of the disease with a nonautonomous infection rate. In section 3 and 4, we present the main equations of the sensitivity analysis together with the stochastic collocation method. In section 5, we compute Sobol' indices to measure the weight of the random input variables on the compartments. We simulate Sobol'indices for parameters with uniform law for two different nonautonomous infection rate.

2. The SIHR epidemiological model

In this paper, we consider the following SIHR compartment model:

$$\begin{cases} \frac{dS_t}{dt} = \tau - \mu S_t - \beta S_t I_t \\ \frac{dI_t}{dt} = \beta S_t I_t - (\nu + \mu + \gamma + \alpha) I_t \\ \frac{dH_t}{dt} = \alpha I_t - (\nu + \mu + \lambda) H_t \\ \frac{dR_t}{dt} = \gamma I_t + \lambda H_t - \mu R_t, \end{cases} \tag{1}$$

where S_t (respectively, I_t, H_t and R_t) is the proportion of susceptible (respectively, infected, hospitalized, recovered) in the population at time t and with initial condition $S_0 > 0$, $I_0 > 0$, $H_0 \geq 0$, $R_0 \geq 0$. We consider that

Table 1
Parameters and their descriptions.

Notations	Interpretations
τ	Birth rate of class S
μ	Natural death rate
ν	Death rate of infection
β	Transmission infection rate
γ	Recovery rate of class I
λ	Recovery rate of class H
α	Hospitalized rate

the mortality of infected people and hospitalized people is the same. The description of parameters is given in Table 1.

The system (1) is positive that means that the solution remains positive for any trajectory initialized at positive conditions.

The basic reproduction number $\mathcal{R}_0$ is an epidemiological metric used to describe the contagiousness or transmissibility of infectious agents. This number is the average number of susceptible an infected person will infect at the start of the epidemic. This parameter plays a crucial role. Following [5, 11, 29], $\mathcal{R}_0$ of the system (1) is given by

$$\mathcal{R}_0 = \frac{\tau}{\mu(\nu+\mu+\gamma+\alpha)}\beta.$$

When the parameter β don't varies over time, the system (1) is said to be autonomous and one can check readily that it admits two equilibrium points. The disease-free equilibrium $E_0 = \left(\frac{\tau}{\mu}, 0, 0, 0\right)$ which always exists and it is globally asymptotically stable if $\mathcal{R}_0 < 1$. If $\mathcal{R}_0 > 1$, there is an endemic equilibrium state $E^* = (S^*, I^*, H^*, R^*)$ which is locally asymptotically stable and given by (see [14]):

$$\begin{cases} S^* = S^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\tau}{\mu \times \mathcal{R}_0} \\ I^* = I^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\tau}{\mu+\nu+\gamma+\alpha}\left(1 - \frac{1}{\mathcal{R}_0}\right) \\ H^* = H^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\alpha\tau}{(\mu+\nu+\lambda)(\mu+\nu+\gamma+\alpha)}\left(1 - \frac{1}{\mathcal{R}_0}\right) \\ R^* = R^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\gamma\tau(\mu+\nu+\lambda)+\lambda\alpha\tau}{\mu(\mu+\nu+\lambda)(\mu+\nu+\gamma+\alpha)}\left(1 - \frac{1}{\mathcal{R}_0}\right). \end{cases} \tag{2}$$

Taking policy interventions into account, it makes sense to assume that the transmission rate varies over time in the SIHR model, in such case the system is said nonautonomous. We assume that the parameter β varies over time and would decrease with the interventions. For example, we assume that $\beta_t \in \mathcal{C}(\mathbb{R}, [\beta^-, \beta^+])$, where $0 < \beta^- \leqslant \beta^+ < \infty$. When the system (1) is nonautonomous [17, 23], there is no endemic equilibrium state but a nonautonomous equilibrium solution $E_t^* = (S_t^*, I_t^*, H_t^*, R_t^*)$ which is globally asymptotically stable when $\beta^- > \frac{\mu(\nu+\mu+\gamma+\alpha)}{\tau}$. For $\beta_t = \beta$ is a constant, we have $E_t^* = E^*$. The corresponding $\mathcal{R}_t$ was defined as

$$\mathcal{R}_t = \frac{\tau}{\mu(\nu+\mu+\gamma+\alpha)}\beta_t = \frac{\mathcal{R}_0}{\beta_0}\beta_t$$

which varies in the interval $\left[\frac{\beta^-\tau}{\mu(\nu+\mu+\gamma+\alpha)}; \frac{\beta^+\tau}{\mu(\nu+\mu+\gamma+\alpha)}\right]$.

We will be interested in two infection rate. The logistic function (see [27]) given by

$$\beta_t = \frac{\beta_0}{1+\exp(-kt)} \tag{3}$$

where k is the intensity of beta convergence towards β_0. And the sinusoidal function (see [4]) given by

$$\beta_t = \beta_0\left[1+\varepsilon\sin\left(\frac{2\pi}{T}t\right)\right], \tag{4}$$

where ε is the amplitude and T the period.

2.1 *Parameter estimation and simulations*

The COVID-19 virus is a new virus linked to the same family of viruses as Severe Acute Respiratory Syndrome (SARS) and some types of common cold [8, 18, 19].

This epidemic is deadlier than the 21st century seasonal flu epidemics in France. You have to go back to 1957-1958 and 1968-1969 to find epidemics of seasonal that claimed more lives in France than COVID-19. As of 19 November 2020, the reports related to COVID-19 in medico-social establishments published daily by the National Public Health Agency show a total of 47,127 deaths and 2,086,288 confirmed positive cases by polymerase chain reaction (PCR) test.

Table 2

Parameters estimations.

Parameter	value: 01 October 15 November
τ	714000/64821954=0.011
μ	599000/64821954=0.00924
ν	0,0002
γ	0.423254705
λ	0.05732
α	0.0336454691

The total population in France is 64821957. According to the report cases about COVID-19 infection at $01th$ October 2020, we can see that:

$$S_0 = 64806884/64821957, I_0 = 13970/64821957, H_0 = 626/64821957,$$
$$R_0 = 474/64821957.$$

We have estimated the main model parameters using the French cases of infection going from 01th october to 20*th* November 2020 as shown in Table 2. For the estimation of the parameters, we considered the average of each parameter over the number of days.

$\mathcal{R}_0$ is computed with a population that is totally susceptible to be infected. It takes into account three factors: the risk of contracting the virus by contact, the number of contacts in a unit of time and the number of days during which an infected person stays contagious. For a highly contagious virus such as the coronavirus, $\mathcal{R}_0$ can reach a value of three or more. $\mathcal{R}_0$ can also varies according to population density, susceptibility and other factors.

In Figures 1 and 2, we consider the sinusoidal transmission rate (4) for two different basic reproduction rate $\mathcal{R}_0^{(1)} = 3$ in Figure 1 and $\mathcal{R}_0^{(2)} = 0.95$

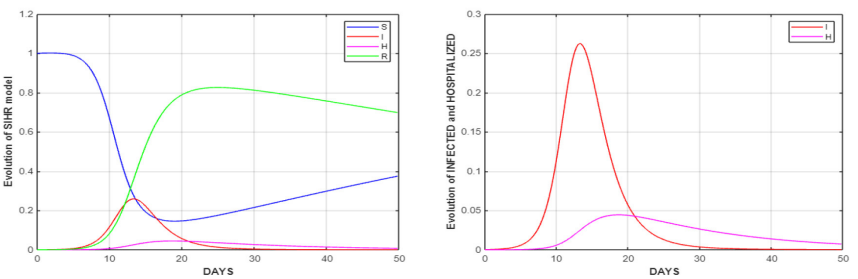

Figure 1

Evolution of the system with sinusoidal rate: $\mathcal{R}_0^{(1)} = 3, \ \varepsilon = 0.1.$

in Figure 2. The logistic transmission rate (3) is considered in Figures 3 and 4.

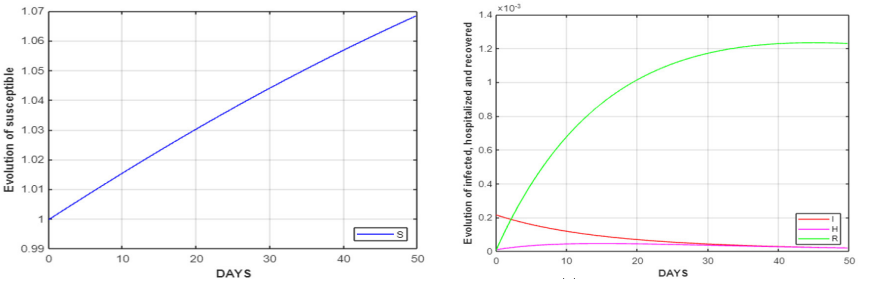

Figure 2

Evolution of the system with sinusoidal rate: $\mathcal{R}_0^{(2)} = 0.95, \quad \varepsilon = 0.1.$

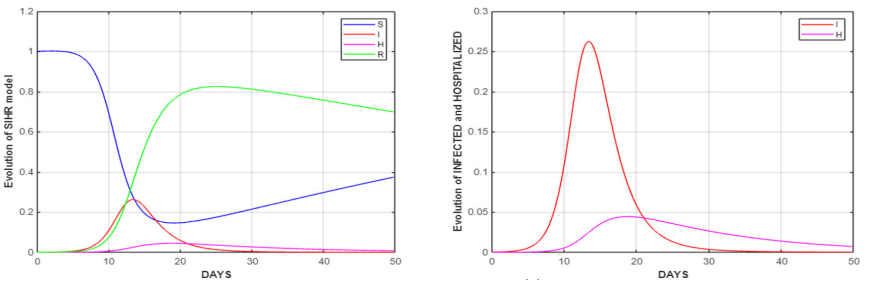

Figure 3

Evolution of the system with logistic rate: $\mathcal{R}_0^{(1)} = 3$ **and** $k = 1.5.$

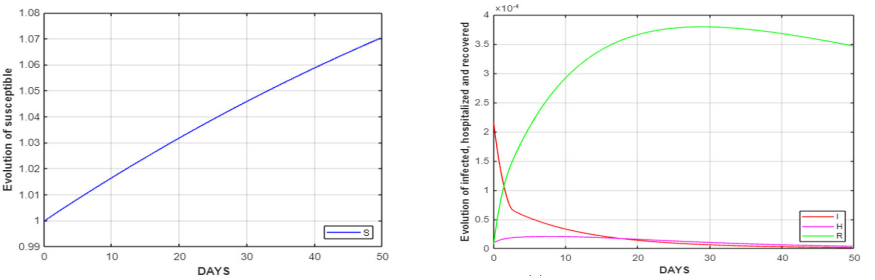

Figure 4

Evolution of the system with logistic rate: $\mathcal{R}_0^{(2)} = 0.95$ **and** $k = 1.5.$

In Figures 1 and 3, we see exactly the same evolution of the system with two different transmission rates. We notice that there will be one wave after 13 days. In these Figures, we observe that the susceptible population is decreasing with time and more people are getting exposed.

As the number of infected people is increasing, the disease spreads to a large part of the population in a very short period of time. This is confirmed by analyzing the infected population dynamics which shows that it grows faster in the first twenty days.

In Figures 2 and 4, for $\mathcal{R}_0^{(2)} = 0.95 < 1$, we note that the population increases almost linearly and that the number of infected and hospitalized decreases until approaching to zero. We see exactly the same evolution of the system with two different transmission rates. We note that with a reproduction rate $\mathcal{R}_0^{(2)}$ the epidemic will not last over time.

For an $\mathcal{R}_0^{(1)} = 3$ we have a peak of infectivity in the population after 13 days, see Figures 1 and 3. In contrary, there is not a peak of infectivity in the population for an $\mathcal{R}_0 = 0.95$ and an epidemic that will not last over time, see Figures 2 and 4.

3. Sensitivity analysis

In this section, we introduce the main mathematical components that lead to the decomposition of the variance, from which the Sobol indices are derived.

We consider the mathematical model $Y = Y(\mathbf{X})$, where the input parameter $\mathbf{X} = (X_1, \ldots, X_n)$ consists of n independent random variables which satisfy :

$$\mathbb{E}(X_i^2) = \int x_i^2 p_i(x_i)\, dx_i < \infty, \tag{5}$$

where p_i is the marginal probability density of X_i and $\mathbb{E}$ stands for the mathematical expectation.

Similarly, Y is assumed to have a finite second order moment.

The article [26] proposed an indicator of the influence of the input parameter X_i defined by

$$S_i = \frac{\mathbb{V}(\mathbb{E}(Y|X_i))}{\mathbb{V}(Y)}. \tag{6}$$

commonly termed `first order Sobol index'. The S_i indices lie between 0 and 1, and they take values close to one when the influence of X_i is important. $\mathbb{V}_i = \mathbb{V}(\mathbb{E}(Y / X_i))$ is the conditional variance of Y with respect to X_i and $\mathbb{V} = \mathbb{V}(Y)$ is the total variance of Y. Similarly, sensitivity indices of higher order can be defined by first introducing the following decomposition of the total variance

$$\mathbb{V}(Y) = \sum_{i_1=1}^{n} \mathbb{V}_i + \sum_{1 \le i_1 < i_2 \le n} \mathbb{V}_{i_1 i_2} + \sum_{1 \le i_1 < i_2 < i_3 \le n} \mathbb{V}_{i_1 i_2 i_3} + \cdots + \mathbb{V}_{i_1 \ldots i_n}, \tag{7}$$

where

$$\begin{aligned}
\mathbb{V}_{i_1} &= \mathbb{V}(\mathbb{E}(Y / X_{i_1})),\\
\mathbb{V}_{i_1 i_2} &= \mathbb{V}(\mathbb{E}(Y / X_{i_1}, X_{i_2})) - \mathbb{V}_{i_1} - \mathbb{V}_{i_2},\\
\mathbb{V}_{i_1 i_2 i_3} &= \mathbb{V}(\mathbb{E}(Y / X_{i_1}, X_{i_2}, X_{i_3})) - \mathbb{V}_{i_1 i_2} - \mathbb{V}_{i_1 i_3} - \mathbb{V}_{i_2 i_3} - \mathbb{V}_{i_1} - \mathbb{V}_{i_2} - \mathbb{V}_{i_3},\\
&\cdots,\\
\mathbb{V}_{i_1 \ldots i_n} &= \mathbb{V} - \sum_{i=1}^{n} \mathbb{V}_i - \sum_{1 \le i_1 < i_2 \le n} \mathbb{V}_{i_1 i_2} - \cdots - \sum_{1 \le i_1 < i_2 \ldots < i_{n-1} \le n} \mathbb{V}_{i_1 \ldots i_{n-1}}.
\end{aligned} \tag{8}$$

By equations (6) and (7) we can derive the formulas for all Sobol' indices of orders greater than 1

$$\begin{cases}
S_{i,j} = \dfrac{\mathbb{V}(\mathbb{E}(Y|X_i, X_j))}{\mathbb{V}(Y)} - S_i - S_j,\\
S_{i,j,k} = \dfrac{\mathbb{V}(\mathbb{E}(Y|X_i, X_j, X_k))}{\mathbb{V}(Y)} - S_i - S_j - S_k - S_{i,j} - S_{i,k} - S_{j,k},\\
\ldots.\\
S_{1,\ldots,n} = \dfrac{\mathbb{V}(\mathbb{E}(Y|X_1, \ldots, X_n))}{\mathbb{V}(Y)} - \sum_{U \subseteq \{1,\ldots,n\}} S_U.
\end{cases} \tag{9}$$

When the model contains n entry random variables, the number of Sobol' indices is $(2^n - 1)$. This number grows exponentially fast as n increases. This is why Sobol' indices can only be computed for problems with moderate input sizes.

In order to make the computation efficient, one should avoid to estimate the integrals appearing into equations (8). Alternatively, modeling the problem under the form of Lagrange polynomials turns out to be a worthy option and it is embraced in this work.

4. The Stochastic collocation method (SCM)

In this section, we describe a method called the stochastic collocation method (SCM). We will present briefly this method when Y depends only on one random variable X and then apply it to the case of many random variables. This method makes it possible to represent a random variable Y of unknown distribution as a function of a random variable X of known distribution in the form

$$Y = \sum_{i=1}^{n} y_i L_i(X) \tag{10}$$

where $\{L_i(x)\}_{i=1}^n$ are Lagrange polynomial basis and $\{y_i\}_{i=1}^n$ are real numbers. If a relation $Y = f(X)$ exists then $y_i = f(x_i)$ and relation (10) becomes the projection of f on the basis of Lagrange polynomials.

Before describing the method, we will recall two concepts that are important. These two notions are the interpolation polynomials and the quadrature rule.

Let $\mathcal{P}_n$ denote the linear space of polynomials of degree less or equal to n. Let $(n+1)$ distinct points $x_0, x_1, ..., x_n$ and corresponding values $y_0, y_1, ..., y_n$, then there exists a unique polynomial $P \in \mathcal{P}_n$ such that $P(x_i) = y_i$. Lagrange establishes a representation of such polynomials under the form

$$P(x) = \sum_{i=0}^{n} y_i L_i(x),$$

where

$$L_i(x) = \prod_{j=0 | j \neq i}^{n} \frac{x - x_j}{x_i - x_j}, \qquad i = 0, ..., n. \tag{11}$$

This equality shows that the quality of the approximation depends only on the choice of the points x_i, for more details see [13].

Designing an optimal set of interpolating points is a hard task that is problem-dependent. This is why people choose x_i that are the same as the quadrature points. In practice, it is observed that it usually leads to good approximation properties. Quadrature rules can be written as a weighted sum

$$\int_I f(x) p_X(x) dx \approx \sum_{i=0}^{n} \omega_i f(x_i) \tag{12}$$

where ω_i are the quadrature weight and x_i are the quadrature points and p_X is a positive function that corresponds to a probability density of X. We choose x_i and ω_i so that the relation (12) is exact when f is a polynomial of degree less than or equal to $2n-1$, for more details see [9, 10, 12, 22]. For more information on the stochastic collocation method see [6] which applies the method with random variables and stochastic processes.

We will illustrate the method by computing the Sobol indices for the SHIR model (1). We consider the compartment susceptible S. We suppose that the input variable

$$\mathbf{X} := (X_1, X_2, X_3, X_4) := (\beta, \alpha, \gamma, \lambda)$$

are independent random variables with density function $p_{\mathbf{X}}(\mathbf{x}) = \prod_{k=1}^{4} p_{X_k}(x_k)$.

We have the following decomposition of S in Lagrange polynomials

$$S_t(\mathbf{X}) = \sum_{i,j,k,l=0}^{n} S_{ijkl}(t)\mathbf{L}_{ijkl}(\mathbf{X}), \tag{13}$$

$$I_t(\mathbf{X}) = \sum_{i^*,j^*,k^*,l^*=0}^{n} I_{i^*j^*k^*l^*}(t)\mathbf{L}_{i^*j^*k^*l^*}(\mathbf{X}), \tag{14}$$

where we used the notation $\mathbf{L}_{i,j,k,l}(\mathbf{X}) := L_i(X_1)L_j(X_2)L_k(X_3)L_l(X_4)$, and L_h is given in (11).

Replacing (13) and (14) and in the differential equation (1), we obtain

$$\begin{aligned}\sum_{i,j,k,l=0}^{n} \frac{dS_{ijkl}(t)}{dt}\mathbf{L}_{ijkl}(\mathbf{X}) &= \tau - \beta\Big(\sum_{i,j,k,l=0}^{n} S_{ijkl}(t)\mathbf{L}_{ijkl}(\mathbf{X})\Big) \\ &\times\Big(\sum_{i^*,j^*,k^*,l^*=0}^{n} I_{i^*j^*k^*l^*}(t)\mathbf{L}_{i^*j^*k^*l^*}(\mathbf{X})\Big) - \mu\sum_{i,j,k,l=0}^{n} S_{ijkl}(t)\mathbf{L}_{ijkl}(\mathbf{X}).\end{aligned} \tag{15}$$

For the particular value $\mathbf{X}' = (x_m,\ x_n,\ x_p,\ x_q)$, we obtain

$$\mathbf{L}_{i,j,k,l}(\mathbf{X}') = L_i(x_m)L_j(x_n)L_k(x_p)L_l(x_q) = \delta_{im}\times\delta_{jn}\times\delta_{kp}\times\delta_{lq}$$

where δ_{im} is the Kronecker symbol.

So, the last equation (15) becomes

$$\frac{dS_{mnpq}}{dt} = \tau - \beta_m S_{mnpq} I_{mnpq} - \mu S_{mnpq}. \tag{16}$$

Similar equations can be written for the other components of the differential equation.

Furthermore, using (12), it can be shown that

$$\begin{aligned}\mathbb{E}[S_t(\mathbf{X})] &= \sum_{i,j,k,l=0}^{n} S_{ijkl}(t)\int \mathbf{L}_{ijkl}(\mathbf{x})p_{\mathbf{X}}(\mathbf{x})d\mathbf{x} \\ &= \sum_{i,j,k,l=0}^{n} S_{ijkl}(t)\omega_i\omega_j\omega_k\omega_l.\end{aligned} \tag{17}$$

Taking the square of the decomposition (13), we obtain

$$(S_t(\mathbf{X}))^2 = \sum_{i,j,k,l,i^*,j^*,k^*,l^*=0}^{n} S_{ijkl}(t)S_{i^*j^*k^*l^*}(t)\mathbf{L}_{ijkl}(\mathbf{X})\mathbf{L}_{i^*j^*k^*l^*}(\mathbf{X}).$$

In a similar way as before and using (12), we obtain:

$$\mathbb{E}[(S_t(\mathbf{X}))^2] = \sum_{i,j,k,l=0}^{n} (S_{ijkl}(t))^2 \omega_i \omega_j \omega_k \omega_l. \tag{18}$$

The variance of $S_t(\mathbf{X})$ is given by

$$\mathbb{V}(S_t(\mathbf{X})) = \mathbb{E}[(S_t(\mathbf{X}))^2] - [\mathbb{E}[S_t(\mathbf{X})]]^2.$$

Now, we will calculate the Sobol indices for S. We recall the Sobol indice S_{X_1}

$$S_{X_1}(t) = \frac{\mathbb{V}(\mathbb{E}(S_t(\mathbf{X})|X_1))}{\mathbb{V}(S_t(\mathbf{X}))} = \frac{\mathbb{E}[(\mathbb{E}(S_t(\mathbf{X})|X_1))^2] - [\mathbb{E}(\mathbb{E}(S_t(\mathbf{X})|X_1))]^2}{\mathbb{V}(S_t(\mathbf{X}))}. \tag{19}$$

We have

$$\begin{aligned}\mathbb{E}(S_t(\mathbf{X}) \mid X_1) = & \sum_{i,j,k,l} S_{ijkl}(t) L_i(X_1) \int L_j(x_2) L_k(x_3) L_l(x_4) p_{X_2}(x_2) p_{X_3}(x_3) p_{X_4}(x_4) \\ & \mathrm{d}x_2 \mathrm{d}x_3 \mathrm{d}x_4 \\ = & \sum_{i,j,k,l} S_{ijkl}(t) L_i(X_1) \omega_j \omega_k \omega_l.\end{aligned}$$

Taking its square, we obtain

$$(\mathbb{E}(S_t(\mathbf{X}) \mid X_1))^2 = \sum_{i,j,k,l,i^*,j^*,k^*,l^*} S_{ijkl}(t) S_{i^*j^*k^*l^*}(t) \omega_j \omega_k \omega_l L_i(X_1) \omega_{j^*} \omega_{k^*} \omega_{l^*} L_{i^*}(X_1).$$

Then, the mean of the squared term is

$$\begin{aligned}\mathbb{E}[(\mathbb{E}(S_t(\mathbf{X}) \mid X_1))^2] = & \sum_{i,j,k,l,i^*,j^*,k^*,l^*} S_{ijkl}(t) S_{i^*j^*k^*l^*}(t) \omega_j \omega_k \omega_l \omega_{j^*} \omega_{k^*} \omega_{l^*} \\ & \int L_i(x_1) L_{i^*}(x_1) p_{X_1}(x_1) \mathrm{d}x_1 \\ = & \sum_{i,j,k,l,j^*,k^*,l^*} S_{ijkl}(t) S_{ij^*k^*l^*}(t) \omega_i \omega_j \omega_k \omega_l \omega_{j^*} \omega_{k^*} \omega_{l^*}.\end{aligned}$$

And the term $\mathbb{E}(\mathbb{E}(S_t(\mathbf{X}) \mid X_1))$ appearing in (19) is equal to $\mathbb{E}(S_t(\mathbf{X}))$ which is given by (17).

For the other Sobol indices of order 1, we do exactly the same.

Now, we are going to calculate the Sobol indice of order 2. As for the indices of order 1, we will just calculate a single indice of order 2 and the other indices are done in exactly the same way.

We recall the Sobol indice $S_{X_1,X_2}(t)$:

$$S_{X_1,X_2}(t) = \frac{\mathbb{E}[(\mathbb{E}(S_t(\mathbf{X})|X_1,X_2))^2] - [\mathbb{E}(\mathbb{E}(S_t(\mathbf{X})|X_1,X_2))]^2}{\mathbb{V}(S_t(\mathbf{X}))} - S_{X_1}(t) - S_{X_2}(t). \tag{20}$$

Using (12), we have

$$\mathbb{E}(S_t(\mathbf{X})\,|\,X_1,X_2) = \sum_{i,j,k,l} S_{ijkl}(t)L_i(X_1)L_j(X_2)\int L_k(x_3)L_l(x_4)p_{X_3}(x_3)\,p_{X_4}(x_4)\mathrm{d}x_3\mathrm{d}x_4 = \sum_{i,j,k,l} S_{ijkl}(t)L_i(X_1)L_j(X_2)\omega_k\omega_l.$$

Taking the square of the last expression, we obtain:

$$(\mathbb{E}(S_t(\mathbf{X})\,|\,X_1,X_2))^2 = \sum_{i,j,k,l,i^*,j^*,k^*,l^*} S_{ijkl}(t)S_{i^*j^*k^*l^*}(t)L_i(X_1)L_j(X_2)L_{i^*}(X_1)\,L_{j^*}(X_2)\omega_k\omega_l\omega_{k^*}\omega_{l^*}.$$

The expectation of this expression is given by:

$$\begin{aligned}&\mathbb{E}[(\mathbb{E}(S_t(\mathbf{X})\,|\,X_1,X_2))^2]\\ &= \sum_{i,j,k,l,i^*,j^*,k^*,l^*} S_{ijkl}(t)S_{i^*j^*k^*l^*}(t)\omega_k\omega_l\omega_{k^*}\omega_{l^*}\int L_i(x_1)L_{i^*}(x_1)p_{X_1}(x_1)\mathrm{d}x_1\\ &\quad \int L_j(x_2)L_{j^*}(x_2)p_{X_2}(x_2)\mathrm{d}x_2\\ &= \sum_{i,j,k,l,k^*,l^*} S_{ijkl}(t)S_{ijk^*l^*}(t)\omega_k\omega_l\omega_{k^*}\omega_{l^*}\omega_i\omega_j.\end{aligned}$$

And finally the term $\mathbb{E}(\mathbb{E}(S_t(\mathbf{X})\,|\,X_1,X_2))$ appearing in (20) is equal to $\mathbb{E}(S_t(\mathbf{X}))$ which is given by (17).

For the Sobol indices of order 1 and 2 for *S*, *I*, *H* and *R*, we do exactly the same.

5. Numerical application

In this section, we use the stochastic collocation method to spread the uncertainty in epidemiological model (1). We make the hypothesis that the parameters $\mathcal{R}_0, \alpha, \gamma, \lambda$ are independent random variables and we compute the Sobol indices of order 1 and 2 using the Lagrange polynomial with two different infection rate and two different basic reproduction number. Finally, we stand out the most influential parameters for each compartment.

We consider that the random input variables $(X_1, X_2, X_3, X_4) = (\mathcal{R}_0, \alpha, \gamma, \lambda)$ of the system (1) follow the uniform law $\mathcal{U}[a_{X_i}, b_{X_i}]$.

In Table 3, we give the values of the different parameters with their confidence intervals. We have distinguished two different values of $\mathcal{R}_0$ which are noted $\mathcal{R}_0^{(1)}$ (before the lockdown) and $\mathcal{R}_0^{(2)}$ (after the lockdown).

Table 3

Specification of the law of the random entry parameters.

Random entry parameters	Uniform law on $[a_{X_i}, b_{X_i}]$
$\mathcal{R}_0^{(1)}$	[2.85 ; 3.15]
$\mathcal{R}_0^{(2)}$	[0.7 ; 1.2]
α	[0.030 ; 0.037]
γ	[0.38 ; 0.46]
λ	[0.0515 ; 0.0630]

5.1 *Sobol indices of the stationary regime*

We recall the endemic stationary point of the system (1) given in (2):

$$\begin{cases} S^* = S^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\tau}{\mu \times \mathcal{R}_0} \\ I^* = I^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\tau}{\mu+\nu+\gamma+\alpha}\left(1 - \frac{1}{\mathcal{R}_0}\right) \\ H^* = H^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\alpha\tau}{(\mu+\nu+\lambda)(\mu+\nu+\gamma+\alpha)}\left(1 - \frac{1}{\mathcal{R}_0}\right) \\ R^* = R^*(\mathcal{R}_0, \alpha, \gamma, \lambda) = \frac{\gamma\tau(\mu+\nu+\lambda)+\lambda\alpha\tau}{\mu(\mu+\nu+\lambda)(\mu+\nu+\gamma+\alpha)}\left(1 - \frac{1}{\mathcal{R}_0}\right), \end{cases}$$

The analytical calculation of the Sobol indices for the endemic equilibrium point are done explicitly.

In Table 4 (resp. Table 5) we give the Sobol indices of the stationary endemic state with $\mathcal{R}_0^{(1)}$ (resp. $\mathcal{R}_0^{(2)}$) that we calculated in an analytical way.

Table 4

Sobol indices of the stationary regime calculated analytically with $\mathcal{R}_0^{(1)}$.

Sobol Indices	$X^* = S^*$	$X^* = I^*$	$X^* = H^*$	$X^* = R^*$
$X^*_{\mathcal{R}_0^{(1)}}$	1	0.069946	0.025106	0.98471
X^*_{α}	0	0.005864	0.34415	0.0012
X^*_{γ}	0	0.92393	0.33147	0.010134
X^*_{λ}	0	0	0.29659	0.0013

Table 5

Sobol indices of the stationary regime calculated analytically with $\mathcal{R}_0^{(2)}$.

Sobol Indices	$X^* = S^*$	$X^* = I^*$	$X^* = H^*$	$X^* = R^*$
$X^*_{\mathcal{R}_0^{(2)}}$	1	0.99663	0.99023	0.9999

We note that after the lockdown, the states (S^*, I^*, H^*, R^*) are almost exclusively impacted by the parameter $\mathcal{R}_0^{(2)}$. And before the lockdown, only S^* and R^* are mostly impacted by $\mathcal{R}_0^{(2)}$.

5.2 *Sobol indices of the transient regime with* $\mathcal{R}_0^{(1)}$

The Sobol indices for each parameters depend on time and admit a stationary equilibrium that is reached at $T_f = 600$. We give in Table 6 (resp. Table 7) the steady state Sobol indices using the stochastic collocation method defined in section (4) and the influences of all variable on each compartment with $\mathcal{R}_0^{(1)}$ and sinusoidal transmission rate (resp. logistic transmission rate).

Table 6

Sobol indices at final time $T_f = 600$ calculated numerically with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

Sobol Indices	$X^* = S^*$	$X^* = I^*$	$X^* = H^*$	$X^* = R^*$
$X^*_{\mathcal{R}_0^{(1)}}$	0.9970	0.0700	0.0252	0.9882
X^*_{α}	0.0000	0.0059	0.3407	0.0012
X^*_{γ}	0.0012	0.9237	0.3347	0.0074
X^*_{λ}	0.0000	0.0000	0.2967	0.0014

Table 7

Sobol indices at final time $T_f = 600$ calculated numerically with $\mathcal{R}_0^{(1)}$ and logistic rate.

Sobol Indices	$X^* = S^*$	$X^* = I^*$	$X^* = H^*$	$X^* = R^*$
$X^*_{\mathcal{R}_0^{(1)}}$	0.9999	0.0710	0.0252	0.9846
X^*_{α}	0.0000	0.0058	0.3429	0.0011
X^*_{γ}	0.0000	0.9223	0.3306	0.0129
X^*_{λ}	0.0000	0.0000	0.2985	0.0013

In Tables 6 and 7, we notice the variable $\mathcal{R}_0^{(1)}$ strongly influences the compartments of susceptible S and recovered R at 99%, it influences the compartments of infected I at 7% and hospitalized H at 2%. The variable γ influences the compartments of infected I at more than 92% and hospitalized H people at more than 33%.

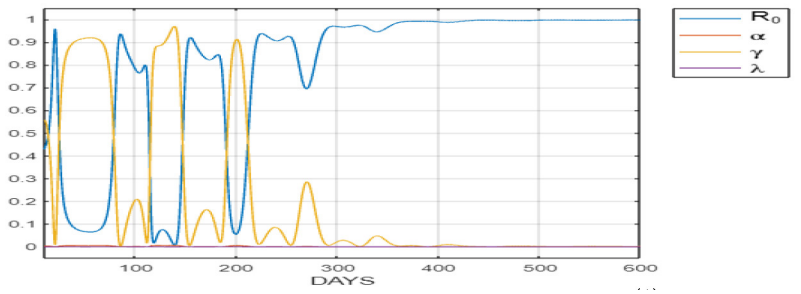

Figure 5

Influence of each parameter on compartment S with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

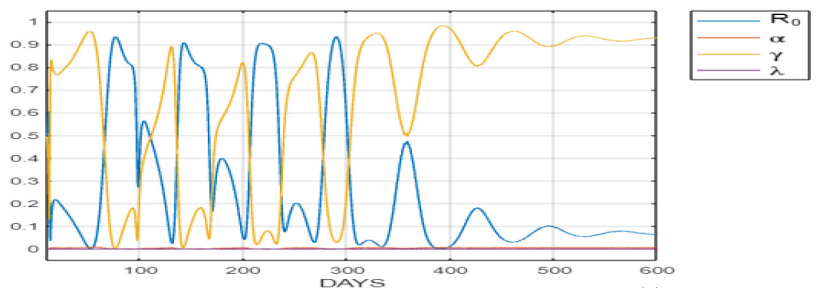

Figure 6

Influence of each parameters on compartment I with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

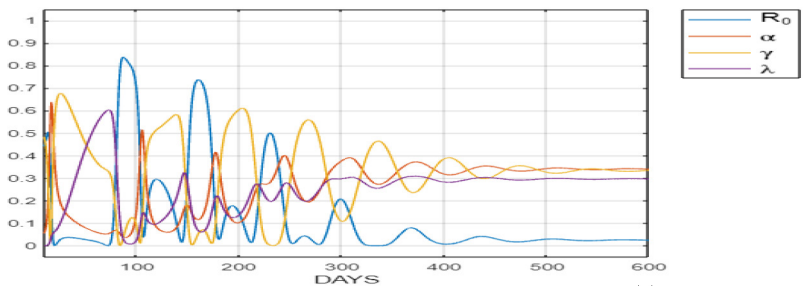

Figure 7

Influence of each parameters on compartment H with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

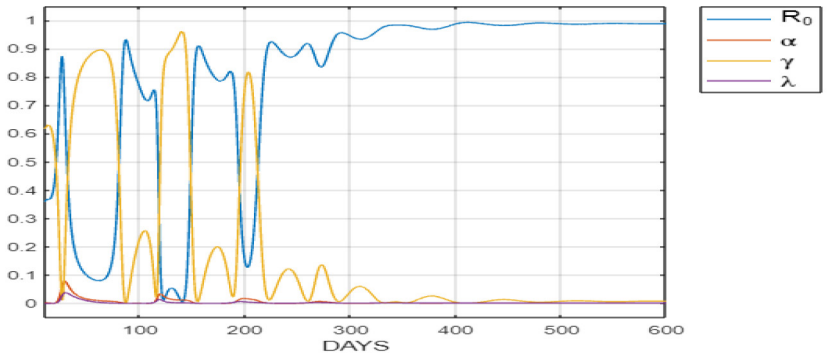

Figure 8

Influence of each parameters on compartment R with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

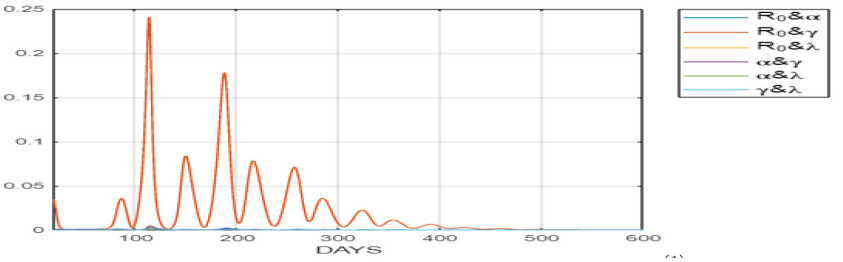

Figure 9

Combined influence of two parameters on compartment S with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

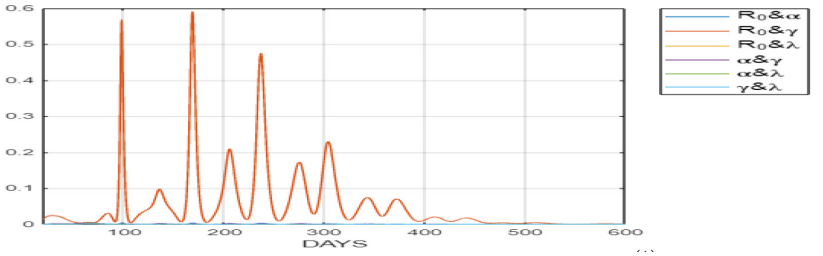

Figure 10

Combined influence two parameters on compartment I with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

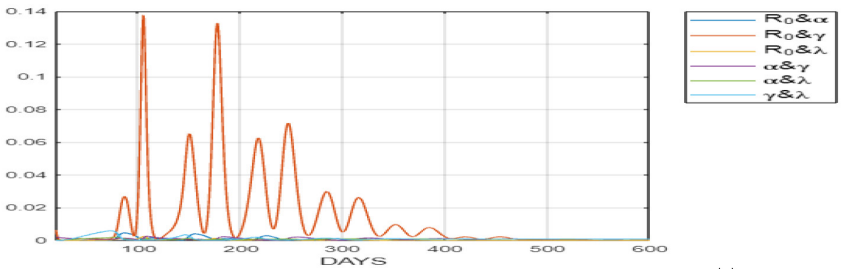

Figure 11

Combined influence of two parameters on compartment H with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

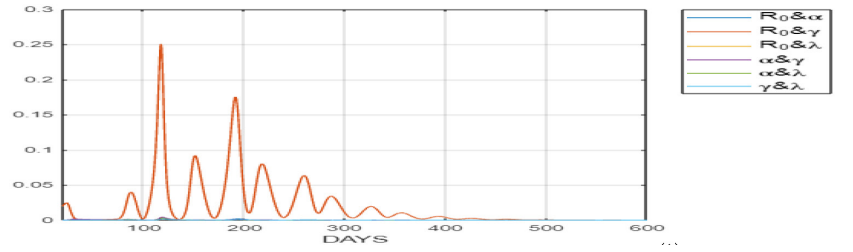

Figure 12

Combined influence of two parameters on compartment R with $\mathcal{R}_0^{(1)}$ and sinusoidal rate.

In Figures 5-8, we have presented the evolution over time of Sobol indices of each parameter on each compartment for a sinusoidal transmission rate.

Figures showing the same results but with a logistic transmission rate are not included since the plots are almost identical to the one of Figures 5-8.

In Figure 6, we can see that the parameters that most influence the compartment of the infected are γ and $\mathcal{R}_0^{(1)}$.

In Figure 7, we observe that the variables α and λ only influence the compartment of hospitalized H people.

The Sobol indices $S_\alpha, S_\lambda, I_\alpha, I_\lambda, R_\alpha$ and R_λ that remain close to zero over time have not been plotted.

All the compartment simulations for the 2nd order indices are given in Figures 9-12. Here again, we only show results with a sinusoidal transmission rate since the ones with a logistic transmission rate are very similar.

The Sobol indices of order 2 inform us that the combined effects on the compartments remains weak. We notice that there is an interaction between $\mathcal{R}_0^{(1)}$ and γ in the four compartments with different intensity. This interaction is strong at the beginning of the propagation but suddenly disappears and tends towards zero.

5.3 *Sobol indices of the transient regime with* $\mathcal{R}_0^{(2)}$

In Table 8 (resp. Table 9) we compute the Sobol indices of the stationary regime $(T_f = 600)$ with a sinusoidal transmission rate (4) (resp. with a logistic transmission rate 3) with $\mathcal{R}_0^{(2)}$.

Table 8

Sobol indices at final time $T_f = 600$ calculated numerically with $\mathcal{R}_0^{(2)}$ and sinusoidal rate.

Sobol Indices	$X^* = S^*$	$X^* = I^*$	$X^* = H^*$	$X^* = R^*$
$X^*_{\mathcal{R}_0^{(2)}}$	0.9996	0.9944	0.9873	0.9997

Table 9

Sobol indices at final time $T_f = 600$ calculated numerically with $\mathcal{R}_0^{(2)}$ and logistic rate.

Sobol Indices	$X^* = S^*$	$X^* = I^*$	$X^* = H^*$	$X^* = R^*$
$X^*_{\mathcal{R}_0^{(2)}}$	0.9999	0.9938	0.9868	0.9999

In Tables 8 and 9, we notice the variable $\mathcal{R}_0^{(2)}$ strongly influences all compartments S, I, H and R at 99%.

In Figure 13, we have shown the compartment results for each parameter when the transmission rate is sinusoidal. Only $S_{\mathcal{R}_0^{(2)}}$, $I_{\mathcal{R}_0^{(2)}}$, $H_{\mathcal{R}_0^{(2)}}$ and $R_{\mathcal{R}_0^{(2)}}$ are represented. The others indices of order 1 and the indices of order 2 remain close to zero over time. There is no interaction between the variables.

Figure 14 shows the compartment results for each parameter when the transmission rate is a logistic function. Indices that remain close to zero over time have not been represented. It should be noted that before the lockdown $(\mathcal{R}_0^{(1)} = 3)$, whether the transmission rate is logistic or sinusoidal has no effect on Sobol indices. On the other hand, after the lockdown $(\mathcal{R}_0^{(2)} = 0.95)$ Sobol indices are different when considering a sinusoidal or a logistic transmission rate.

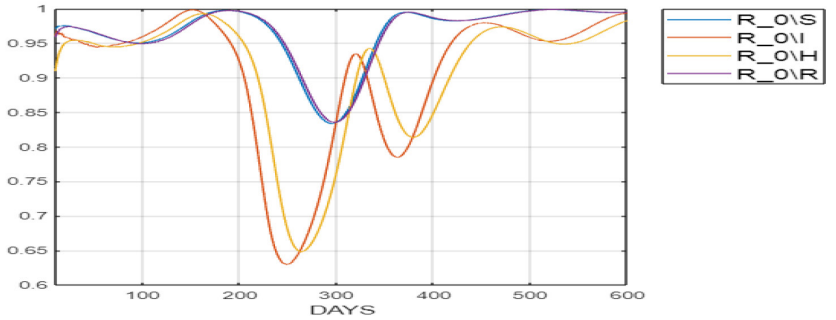

Figure 13

Influence of $\mathcal{R}_0^{(2)}$ on all compartment with sinusoidal rate.

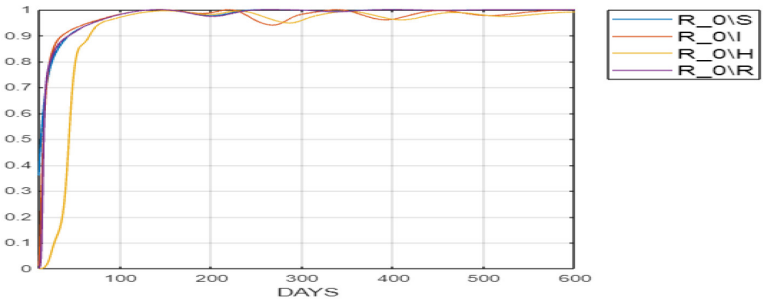

Figure 14

Influence of $\mathcal{R}_0^{(2)}$ on all compartment with logistic rate.

6. Conclusion

From all models for epidemics, the SIR and SIRS are among the simplest. However, the parameters involved cannot be determined exactly and it is necessary to introduce randomness. This paper was dedicated to conduct a sensitivity analysis on an SIHR epidemiological model with random parameters $(\mathcal{R}_0, \alpha, \gamma, \lambda)$. To study their impact on the evolution of the COVID-19 disease, we considered a nonautonomous infection rate of transmission.

Then, we applied the stochastic collocation method using Lagrange polynomials to compute the Sobol indices of order 1 and 2. This has allowed us to study the sensitivity of these factors which intervene in the nonautonomous SIHR epidemiological model. We also have analytically calculated the Sobol indices of the stationary state.

The SIHR model was studied for two cases: for $\mathcal{R}_0 = 3$ (before lockdown) and for $\mathcal{R}_0 = 0,95$ (after lockdown). In both cases, the influence

of the parameters on the compartments reach a steady state. We have shown that for $\mathcal{R}_0 = 0.95$, the basic reproduction rate $\mathcal{R}_0$ is the only parameter that acts on the compartments as well as on the spread of the disease. When $\mathcal{R}_0 = 3$, the influence of the parameters $(\alpha, \gamma, \lambda)$ is considerable on the infected and recovered compartments. We have also shown that, in both cases, the cumulative impact of the parameters is insignificant.

References

[1] M. Arnst, C. Soize, and K. Bulthuis, "Computation of Sobol indices in global sensitivity analysis from small data sets by probabilistic learning on manifolds," *Int. J. Uncertain. Quantif.*, vol. 11, no. 2, pp. 131-145 (2021).

[2] R. Barrio, "Sensitivity analysis of ODEs/DAEs using the Taylor series method," *SIAM,* vol. 27, no. 6, pp. 1929–1947 (2006).

[3] V. Bhatnagar, R. C. Poonia, P. Nagar, S. Kumar, V. Singh, L. Raja, and P. Dass, "Descriptive analysis of COVID-19 patients in the context of India," *J. Interdiscip. Math.,* vol. 24, no. 3, pp. 489–504 (2021).

[4] S. Boatto, C. Bonnet, B. Cazelles, and F. Mazenc, "SIR model with time dependent infectivity parameter: approximating the epidemic attractor and the importance of the initial phase," (2018). [Online]. Available: https://hal-01677886.

[5] C. Castillo-Chavez and B. Song, "Dynamical models of tuberculosis and their applications," *Math. Biosci. Eng.*, vol. 1, no. 2, pp. 361 (2004).

[6] C. Chauvière and H. Djellout, "An efficient spectral method for the numerical solution to some classes of stochastic differential equations," *Math. Methods Appl. Sci.*, vol. 44, no. 7, pp. 5888–5907 (2021).

[7] B. M. Chen-Charpentier and D. Stanescu, "Epidemic models with random coefficients," *Math. Model. Comput.*, vol. 52, no. 7-8, pp. 1004–1010 (2010).

[8] G. Chowell, P. W. Fenimore, M. A. Castillo-Garsow, and C. Castillo-Chavez, "SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism," *J. Theor. Biol.,* vol. 224, no. 1, pp. 1–8 (2003).

[9] G. Dahlquist and A. Björck, Numerical methods in scientific computing, vol. I, Philadelphia, PA: SIAM (1974).

[10] P. J. Davis and P. Rabinowitz, Methods of numerical integration, Mineola, NY: Dover Publications, Inc. (2007).

[11] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts, "The construction of next-generation matrices for compartmental epidemic models," *J. R. Soc. Interface,* vol. 7, no. 47, pp. 873–885 (2010).

[12] G. H. Golub and J. H. Welsch, "Calculation of Gauss quadrature rules," *Math. Comp.,* vol. 23, pp. A1–A10 (1969).

[13] D. Gottlieb and S. A. Orszag, Numerical analysis of spectral methods: theory and applications, Philadelphia, PA: SIAM (1977).

[14] S. Jiao and M. Huang, "An SIHR epidemic model of the COVID-19 with general population size dependent contact rate," *Aims Mathematics,* vol. 5, no. 6, pp. 6714–6725 (2020).

[15] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proc. Math. Phys. Eng.,* vol. 115, no. 772, pp. 700–721 (1927).

[16] S. U. Khan and I. Ali, "Numerical analysis of stochastic SIR model by Legendre spectral collocation method," *Adv. Mech. Eng.*, vol. 11, no. 7, pp. 1687814019862918 (2019).

[17] P. E. Kloeden and V. Kozyakin, "The dynamics of epidemiological systems with nonautonomous and random coefficients," MESA, vol. 2, no. 2, pp. 159–172 (2011).

[18] Z. Ma, Dynamical modeling and analysis of epidemics, World Scientific (2009).

[19] M. Martcheva, An introduction to mathematical epidemiology, vol. 61, Springer (2015).

[20] J. Morio, "Global and local sensitivity analysis methods for a physical system," *Eur. J. Phys.,* vol. 32, no. 6, pp. 1577 (2011).

[21] S. T. Pinho, C. P. Ferreira, L. Esteva, F. R. Barreto, V. C. Morato e Silva, and M. G. L. Teixeira, "Modelling the dynamics of dengue real epidemics," *Philos. Trans. Royal Soc. A.*, vol. 368, no. 1933, pp. 5679–5693 (2010).

[22] S. Rahman, "Extended polynomial dimensional decomposition for arbitrary probability distributions," *J. Eng. Mech.*, vol. 135, no. 12, pp. 1439–1451 (2009).

[23] M. A. Safi, M. Imran, and A. B. Gumel, "Threshold dynamics of a non-autonomous SEIRS model with quarantine and isolation," *Theory in Biosciences,* vol. 131, no. 1, pp. 19–30 (2012).

[24] F. Santonja and B. Chen-Charpentier, "Uncertainty quantification in simulations of epidemics using polynomial chaos," *Comput. Math. Methods Med.* (2012).

[25] V. Singh, R. C. Poonia, S. Kumar, P. Dass, P. Agarwal, V. Bhatnagar, and L. Raja, "Prediction of COVID-19 corona virus pandemic based on time series data using support vector machine," *J. Discret. Math. Sci. Cryptogr.,* vol. 23, no. 8, pp. 1583–1597 (2020).

[26] I. M. Sobol, "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," *Math. Comput. Simulation,* vol. 55, no. 1-3, pp. 271–280 (2001).

[27] T. Tian, J. Zhang, S. Lin, Y. Jiang, J. Tan, Z. Li, and X. Wang, "Data-driven analysis of the simulations of the spread of COVID-19 under different interventions of China," *J. Appl. Stat.*, pp. 1–14 (2021).

[28] T. Turányi, "Sensitivity analysis of complex kinetic systems. Tools and applications," *J. Math. Chem.*, vol. 5, no. 3, pp. 203–248 (1990).

[29] P. Van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Math. Biosci.,* vol. 180, no. 1-2, pp. 29–48 (2002).

[30] J. Wallinga and M. Lipsitch, "How generation intervals shape the relationship between growth rates and reproductive numbers," *Proc. R. Soc. B: Biol. Sci.,* vol. 274, no. 1609, pp. 599–604 (2007).

[31] F. Xiong, S. Chen, and Y. Xiong, "Dynamic system uncertainty propagation using polynomial chaos," *Chin. J. Aeronaut.,* vol. 27, no. 5, pp. 1156–1170 (2014).

[32] Z. Zhang, R. Gul, and A. Zeb, "Global sensitivity analysis of COVID-19 mathematical model," *Alex. Eng. J.,* vol. 60, no. 1, pp. 565–572, 2021.

Received June, 2022

Revised September, 2022