Parameter estimation, sensitivity and control strategies analysis in the spread of influenza in Mexico

K Prieto\textsuperscript{1}, and E Ibarguen-Mondragon\textsuperscript{2}
\textsuperscript{1} Instituto de Matemáticas, Universidad Nacional Autónoma de México, Cuernavaca, México
\textsuperscript{2} Departamento de Matemáticas y Estadística, Universidad de Nariño, Pasto, Colombia

E-mail: kernel@ciencias.unam.mx, edbargun@udena.edu.co

Abstract. In this paper we address a parameter estimation, sensitivity and control strategies analyses for influenza disease using a model the flows of people between four states: susceptible, exposed, infectious, recovered. We solved a curve-fitting mathematical model to Mexican influenza data using a nonlinear least-square method and the Landweber iteration. An optimal control problem is formulated and analyzed based on models between four states: susceptible, exposed, infectious, recovered; model considering educational campaign, vaccination and medical treatment as strategies for disease control. The sensitivity analysis is performed to determine which model parameters are the most important to disease transmission and prevalence. The numeric results suggest that an adequate implementation of these strategies during the outbreak of an epidemic could significantly mitigate the propagation of the disease.

1. Introduction
Effective Vaccination, early detection, proper treatment, isolation, quarantine, educational campaign are some control strategies to wane infectious diseases. With the aim to study the effect of vaccination and educational campaign in the propagation of an infectious disease, we formulate and analyse an optimal control problem with state equation governed by an SEIRS (susceptible (S), exposed (E), infected (I), and resistant (R)) type. Here, it is assumed that the disease transmits horizontally, i.e., vertical transmission is neglected. The horizontal transmission can occur either by direct contact, e.g., touching, licking, biting, or by indirect contact with no physical contact, e.g., vectors or fomites. The total population $N(t)$ is divided into subclasses of individuals who are susceptible, exposed (infected but not yet infectious), infectious and recovered, with sizes denoted by $S(t)$, $E(t)$, $I(t)$ and $R(t)$, respectively. The total population $N(t)$ has a homogeneous spatial distribution and is assumed to vary with time since the birth and natural death rates are assumed to be different. All new-borns are assumed to be susceptible. Individuals are susceptible, then exposed, then infectious, and then recovered with the possibility of becoming susceptible again with the rate constant of loss of immunity equal to $\delta$. The constants $\mu$ and $\kappa$ denote the birth and natural death rate, including immigration and emigration, respectively. Also, $\mu$ is named the recruitment rate of susceptible population. The death rate caused by the infectious disease is neglected. The nonnegative parameters $\beta$, $\alpha$ and $\delta$ denote the transfer rates between the compartments $E$, $I$, $R$ and $S$. The parameter $\beta$ is the effective per capita contact rate of infectious individuals, $\alpha$ is the rate the latent individuals become infectious and $\gamma$ can be regarded as the rate that infectious individuals become recovered and $\delta$ is the rate of loss of immunity, thus, $\alpha^{-1}$, $\gamma^{-1}$ and $\delta^{-1}$ represent the mean latent period, infectious period and immune period (average loss of immunity), respectively. The total
population \( N(t) \) is determined by \( N(t) = S(t) + E(t) + I(t) + R(t) \). Let \( s = S/N, e = E/N, i = I/N \) and \( r = R/N \) denote the fractions of the classes S, E, I and R in the population, respectively. It can be verified that \( s, e, i \) and \( r \) satisfy the system of differential equations, Equation (1).

\[
\begin{align*}
s' &= \mu - \mu s - \beta s i + \delta (1 - s - e - i) \\
e' &= \beta s i - (\alpha + \mu) e \\
i' &= \alpha e - (\gamma + \mu) i, \\
s + e + i + r &= 1
\end{align*}
\] (1)

subject to the restriction \( s + e + i + r = 1 \). We define the basic reproduction number, \( R_0 \), also called the modified contact number as the expected number of secondary cases produced, in a completely susceptible population, by a typical infected individual during its entire period of infectiousness, mathematically described by, \( R_0 = \frac{\alpha \beta}{(\beta + \mu)(\gamma + \mu)} \), which is the product of the transmission coefficient \( \beta \), the probability for the infected fraction becoming infective, \( \alpha / (\beta + \mu) \), and the average infectious period \( 1 / (\gamma + \mu) \). In order to find the steady state points of (1), we set the left side of the Equation (1) to zero, and solving for \( s, e, i \). There exist two equilibrium points, \( P_0 \) and \( P_\ast \) (Equation (2)) for Equation (1).

\[
P_0 = (1, 0, 0) \\
P_\ast = \left( \frac{1}{R_0}, \frac{(\gamma + \mu)}{\alpha k_1} \left( 1 - \frac{1}{R_0} \right), \frac{1}{k_1} \left( 1 - \frac{1}{R_0} \right) \right)
\] (2)

With, Equation (3).

\[
k_1 = \frac{(\gamma + \mu)(\delta + \mu) + \alpha (\gamma + \delta + \mu)}{\alpha (\delta + \mu)}, \tag{3}
\]

which were calculated using Wolfram Mathematica Software. The remainder of this paper is organized as follows. In section 2, we formulate an optimal control problem based on vaccination and prevention as control strategies of an epidemic. Discussion and conclusions are presented in the last section.

2. Optimal control problem

In this section, we present the setting of the optimal control problem in SEIRS epidemic model discussed in section 1. In order to mitigate the disease, we investigate the control policy based on educational campaign and vaccination and a combination of either (prevention or reduction of transmission). Denoting by \( x(t) = (s(t), e(t), i(t))^T \in \mathbb{R}^3, t > 0 \), the solution of Equation (1), also called the state of the system at time \( t \), we can rewrite Equation (1) with respective initial conditions \( s(0) = s_0, e(0) = e_0, i(0) = i_0 \) as Equation (4).

\[
\dot{x} = \varphi(x) \\
x(0) = x_0
\] (4)

Next, we define a control variable \( u(t) \), which is a piecewise continuous function that takes values in a positive bounded set \( U = [0, u_{\text{max}}] \), i.e., \( u : \rightarrow U \). We apply three different control policies by adding a linear term in the control variable \( u \) to model (1), namely considering the system \( \dot{x} = \varphi(x(t)) + u(t) h(x(t)) \), where \( \varphi \) defined by Equation (1), is a vector function of class \( C^1 \), and \( h(x(t)) \) is a function which depends on the chosen control policy. Our target is eradicating the disease outbreak in a fixed time \( T \), i.e., reduce the fractional infected population. We can then write the optimal control problem as Equation (5).
minimize: \( J(u) = \int_0^T (b_1 e + b_2 i + b_3 u^2) \, dt \),
subject: \( \dot{x} = q(x(t)) + u(t) h(x(t)), t \geq 0 \)  \( \theta \)
with initial condition: \( x(0) = x_0 \land u: [0, \infty) \rightarrow U \),

where \( J(u) \) is the functional cost, also called the objective functional, and \( b_1, b_2 \) represent social costs which depend on the total infected population, \( i.e., \) the susceptible and exposed populations \((e + i)\), and \( b_3 \) depend on the relative weight associated with implementing the control \( u(t) \). For details of control strategies proposed; educational campaign and vaccination, see [1]. The control variable will be in the rate term of recovery of infected people in the medical treatment strategy.

3. Sensitivity analysis

In determining how best to reduce the number of susceptible population due to the disease outbreak, it is necessary to know the relative importance of the different factors responsible for its transmission and prevalence. Initial disease transmission is directly related to \( R_0 \), and disease prevalence is directly related to the endemic equilibrium point, \( P_\ast \), specifically to the magnitudes of \( s, e, i \). The proportion of infectious people, \( i \), is especially important because it represents the people who may be clinically. We calculate the sensitivity indices of the reproductive number, \( R_0 \), and the endemic equilibrium point, \( P_\ast \), to the parameters in the model. These indices tell us how crucial each parameter is to disease transmission and prevalence. Sensitivity analysis is commonly used to determine the robustness of model predictions to parameter values since there are usually errors in data collection and presumed parameter values [2,3]. Here we use it to discover parameters that have a high impact on \( R_0 \) and \( P_\ast \), and should be targeted by intervention strategies.

- **Definition.** The normalized forward sensitivity index of a variable, \( Q \), which depends differentiable on a parameter, \( p \), is defined as Equation (6).

\[
\gamma_p^Q := \frac{\partial Q}{\partial p} \times \frac{p}{Q}
\]

As we have an explicit formula for \( R_0 \) Equation (4) and \( P_\ast \) Equation (6), we derive an analytical expression for the sensitivity of both \( R_0 \) and \( P_\ast \), to each of the seven different parameters described in Table 1.

4. Parameter estimation of the epidemic model

Even for the correct set of parameters, there is intrinsic discrepancy between the data and the model prediction because of noise measurements. The model may be incomplete, \( i.e., \) it may not consider all the relevant variables of the phenomena, or there could be a model error, where the model’s assumptions may not be fulfilled. For the parameter estimation, we used inverse problems theory in this manuscript. Denoting the state variable \( x(t) \) as in section, and the parameters \( \theta \), as \( \theta = (\alpha, \beta, \gamma, \mu, \delta) \in \mathbb{R}^5 \) in model, we can write the Equation (1) as the Cauchy problem; Equation (7).

\[
\dot{x} = q(x, 0) \\
x(0) = x_0
\]

Equation (7) define a mapping \( \Phi(\theta) = x \) from parameters \( \theta \) to state variables \( x \), where \( \Phi: \mathbb{R}^m_+ \rightarrow (L^2[0, T])^n \) denotes the nonnegative real numbers and \( m \) is the dimension number of parameters to estimate, in this paper we will estimate the parameters \( \theta = (\alpha, \gamma) \), the ones most sensitive regarding Table 2 and Table 3, thus in this case \( m = 2 \), and \( n \) is the number of state variables. We assume that \( \Phi \) has a Fréchet derivative, denoted by \( \Phi \), and is injective in this paper, thus the forward problem has a unique solution \( x \) for a given \( \theta \). The Fréchet derivative of \( \Phi \) is a mapping \( \Phi'(\theta): \mathbb{R}^m_+ \rightarrow (L^2[0, T])^n \),
resulting as the usual derivative since the domain and range of $\Phi'$ are finite dimensional spaces for the Equation (1).

Typically, the data in epidemics consists of measurements of the state variable at a discrete set of point $t_1, ..., t_k$, and only a subset of the state variables are measurable. In case of influenza, data consists of counting the number of infected and dead people caused by a certain disease. This defines a linear observation mapping from state variables to data $\Psi(x): (L^2[0,T])^n \rightarrow \mathbb{R}^{p \times k}$, where $p \leq n$ is the number of observed variables and $k$ is the number of sample points. Considering infectious disease such as influenza, $\Psi(x) = ([i(t_1), ..., i(t_k)] \in \mathbb{R}^{1 \times k}$, i.e., in this case, the infected population $i(t)$ is the state variable measurable, thus $p = 1$. We want to point out that considering influenza disease, a forced model would be more suitable than model than Equation (1) since a forced model takes into account the seasonal nature of transmission of the illness, i.e., it takes into account the periodicity of the illness. This task will be part of our future work. Because of this observation, we only used data of one season. Also, we will work on a more appropriate model which takes into account the mortality of rate due to the disease. Let $F: \mathbb{R}^n \rightarrow \mathbb{R}^{p \times k}$ be defined by $F(\theta) = \Psi(\Phi(\theta))$, the inverse problem parameter estimation of ordinary differential equation systems is to find the parameter $\theta$ which minimize the cost functional $J$, Equation (8).

$$\min_{\theta \in \mathbb{R}^n} J(\theta),$$

subject: $\dot{x} = \Phi(x, \theta), x(0) = x_0$;

The cost functional, $J$, is typically written as a Gaussian distribution or the following least-squares form Equation (9).

$$J(\theta) = \|\Psi(x) - z_\eta\|^2,$$

where $\eta \in \mathbb{R}^{p \times k}$ is the data which has error measurements of size $\eta$. Since the data in epidemics consists of counting the number of infected and dead people, a Poisson distribution instead of a Gaussian distribution may be used to model the cost functional $J(\theta)$. Equation (7) may be solved using numerical tools to deal with a nonlinear least-squares problem or the Landweber method or the combination of both. The Landweber method has been used in [4]. We point out that faster methods as the Levenberg-Marquardt or Conjugate Gradient ones than Landweber method, may be used to obtain a faster convergence. The Landweber iteration is stated in the theorem [5].

5. Numerical solution of the optimal control problem

We solve numerically the optimal control Equation (8) via the forward-backward scheme [6,7] using Matlab R2015a, license number 1081117; starting with an initial guess for the optimal control $u$, the state variables are solved forward in time form the dynamics, Equation (9), using a Runge-Kutta method of the fourth order. Then, those state variables and initial guess for $u$ is used to solve the adjoint backward in time with given final conditions, again employing a fourth order Runge-Kutta method. The control $u$ is updated and used to solve the state and then the adjoint system. This iterative process terminates when current state, adjoint, and control values converge sufficiently. We implemented this algorithm using MatLab and using parameters in Table 1. The units of all parameters values are day$^{-1}$. All control strategies consider an example based on a case of influenza epidemic in Brazil in 1918 [1]. We assume the following initial conditions for the fractional populations: $s_0 = 0.9516, e_0 = 0.0312, i_0 = 0.0156$.

Figure 1(a) shows the control rate $u$ must be implemented using an educational campaign (dashed blue line) and/or vaccination (dashed red line) once at a time, with respect to use both control strategies simultaneously, i.e., both controls $u$ and $v$, educational campaign (solid blue line) and vaccination (solid red line). Figure 1(b) shows the comparison control rate variable $u$ with respect to each strategy proposed; viz: educational campaign, vaccination and medical treatment.
Table 2 shows the sensitivity indices of the basic reproductive number $R_0$. We observe in this table that the parameters $\beta$ and $\gamma$ are the most sensitive with respect to the initial transmission of the disease. Table 3 shows the sensitive indices of parameters with respect to the disease's prevalence at a high baseline Figure 1(a) and a low baseline Figure 1(b). In Table 3, we observe that the parameters $\alpha$ and $\gamma$ are the most sensitive with respect to the disease's persistence.

**Table 1.** Parameters' values used to solve (1)

| Parameter          | Description                                  | Baseline high | Baseline low |
|--------------------|----------------------------------------------|---------------|--------------|
| $\mu$              | Natural birth rate (day$^{-1}$)              | 0.0440        | 0.0044       |
| $\beta$            | Infection transmission rate (day$^{-1}$)     | 0.5000        | 0.5000       |
| $\kappa$           | Natural death rate (day$^{-1}$)              | 0.0042        | 0.0042       |
| $\sigma$           | Efficiency of vaccination                    | 0.8000        | 0.8000       |
| $\delta$           | Immunity loss rate (day$^{-1}$)              | 0.2000        | 0.1000       |
| $\alpha$           | Exp. to infectious rate (day$^{-1}$)         | 0.0700        | 0.0700       |
| $\gamma$           | Recovery rate (day$^{-1}$)                   | 0.2000        | 0.1000       |
| $\omega$           | Efficiency of educational campaign           | 0.8000        | 0.8000       |

**Figure 1.** Comparison of control variable rate $u$ using educational campaign and vaccination and the control variables' rates $u; v$ using a combination of educational campaign and vaccination (a); and comparison of control variable's rate $u$ using educational campaign, vaccination and medical treatment (b).

**Table 2.** Sensitivity indices of $R_0$ to parameters for the SEIRS model, evaluated at the baselines parameter values given in Table 1.

| Parameter | Baseline high (BH) | Baseline low (BL) |
|-----------|--------------------|-------------------|
| $\mu$     | -0.566             | -0.101            |
| $B$       | 1.000              | 1.000             |
| $A$       | 0.386              | 0.059             |
| $\Gamma$  | -0.820             | -0.956            |

**Table 3.** Sensitivity indices of $i_1$ to parameters for the SEIRS model, evaluated at the high baseline (up) and at the low baseline (down) parameter values given in Table 1.

| Parameter | Baseline high | Baseline low |
|-----------|---------------|--------------|
| $\mu$     | -0.464        | -0.035       |
| $B$       | 0.659         | 0.286        |
| $A$       | 0.127         | 0.266        |
| $\Gamma$  | 0.911         | 0.445        |
|           | -1.233        | -0.965       |
Figure 2(a) shows the curve fitting of data set of “Dirección General de Epidemiología (DGE)” [8] using the Nelder-Mead method. Figure 2(b) shows the curve fitting of data set of DGE [8] using the Landweber iteration formula.

![Figure 2. Curve-fitting of data set using the least-squares method (a); and curve fit of data set using the Landweber iteration (b).](image)

For the estimation of parameters θ, we have chosen to reconstruct the parameters θ = (α, γ) and θ = (β, γ) since they are the most sensitive parameters regarding Table 3. The last combination taken as a pair would be θ = (α, β). In order to estimate the parameter θ, i.e., to solve the minimization problem, we used two approaches, viz: in the first approach, we have used the Nelder-Mead method [9,10], more precisely, we used the fminsearch MATLAB function, in the second approach. A third approach may be to use a combination of the previous two approaches, i.e., the initial guess θ₀ for the Landweber iteration is calculated using the Nelder-Mead method. We point out since the third approach can consume a lot of time in the first stage of the algorithm; finding a good initial guess calculated with the Nelder-Mead method, we decided to analyze it in a future work. Data corresponds to the influenza infected people registers of Mexico in 2016-2017. The data sets are taken from the DGE [8].

We used the following parameters in for section 4, i.e., the estimation of (α, γ). The initial values for (s₀, e₀, i₀) = (0.999999913.85 × 10⁻⁹, 5.12 × 10⁻⁹). The rates of the natural birth, infectious transmission, immunity loss is: (μ, β, δ) = (1.2 × 10⁻⁴, 0.5465 × 10⁻⁵). Solving the problem (8) using the Nelder-Mead algorithm, we obtain (α, γ) = (12.46,0.44). Substituting these values in the basic reproductive number formula, we obtain R₀ = 1.24056, that is, R₀ > 1. Since for , R₀ > 1, the endemic point Pₑ is locally asymptotically stable in Ω, thus the point Pₑ = (0.999999913.85 × 10⁻⁹, 5.12 × 10⁻⁹) will asymptotically tend to Pₑ, thus the disease will persist. Solving the Equation (8) using the Landweber iteration [11-13], we obtain we obtain (α, γ) = (18.00,0.44). Substituting these values in the basic reproductive number Equation (4), we obtain R₀ = 1.24056, that is, R₀ > 1. Analogously, the point Pₑ will asymptotically tend to Pₑ, thus the disease will persist. In Figure 2. In this case simply justify the caption so that it is as the same width as the graphic.

6. Conclusions
In this work, we formulate and analyze an optimal control problem based on a SEIRS type epidemiological model via vaccination and educational campaign. Also, we implemented the Nelder-Mead method and the Landweber iteration to solve the problem of estimating the parameters in a SEIRS type epidemiological model. The predictive capabilities of the both curve-fitting methods using noisy true data are demonstrated through numerical simulations.
Acknowledgments
E Ibarguen-Mondragon acknowledge support from project “Acuerdo No 106-25/9/2018 (VIPRI-UDENAR)”.

References
[1] Prieto K, Ibargüen-Mondragon E 2019 Modelación matemática III biomatemáticas e ingeniería: control óptimo de un modelo epidemiológico tipo SEIRS ed Barragan F (México: Universidad Tecnológica de la Mixteca)

[2] De la Sen M, Ibeas A, Alonso-Quesada S 2012 On vaccination controls for the SEIR epidemic model Commun Nonlinear Sci Numer Simulat 17(6) 2637

[3] Zü Z 2011 Sensitivity analysis approaches applied to systems biology models JET Sys. Biol. 5(6) 336

[4] Chitnis N, Hyman J, Cushing J 2008 Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model Bulletin of Mathematical Biology 70(5) 1272

[5] Prieto K, Dorn O 2016 Sparsity and level set regularization for diffuse optical tomography using a transport model in 2D Inverse Problems 33(1) 1

[6] Capistrán M, Moreles M, Lara B 2009 Parameter estimation of some epidemic models. The case of recurrent epidemics caused by respiratory syncytial virus Bulletin of Mathematical Biology 71(8) 1890

[7] Tchuenche J, Khamis S, Agusto F 2011 Optimal control and sensitivity analysis of an influenza model with treatment and vaccination Acta Biother 59(1) 1

[8] Dirección General de Epidemiología (EDG) 2018 Informes semanales para la vigilancia epidemiológica de influenza (México: Dirección General Adjunta de Epidemiología Gobierno de México)

[9] Kornienko I, Paiva L, De Pinho M 2014 Introducing state constraints in optimal control for health problems Procedia Technology 17(1) 415

[10] Alonso-Quesada S, De la Sen M, Ibeas A 2017 On the discretization and control of an SEIR epidemic model with a periodic impulsive vaccination Communication in Nonlinear Science and Numerical Simulation 42(1) 247

[11] Castillo C 2006 Optimal control of an epidemic through educational campaigns Electronic Journal of Differential Equations 125 1

[12] Bashier E, Patidar C 2017 Optimal control of an epidemiological model with multiple time delays Applied Mathematics and Computation 292(1) 47

[13] Lenhart S, Workman J 2007 Optimal control applied to biological models (Boca Raton: Chapman & Hall/CRC)