Revista de la Facultad de Ciencias Universidad Nacional de Colombia, Sede Medellín
V 11 N°2 julio-diciembre de 2022 • ISSN-e 2357-5749 • Research paper • Pages 74 a 99
DOI: https://doi.org/10.15446/rev.fac.cienc.v11n2.100986

FLAT LIKELIHOODS: SIR-POISSON MODEL CASE[a]

VEROSIMILITUDES PLANAS: CASO DEL MODELO SIR-POISSON

JOSÉ A. MONTOYA[b*], GUDELIA FIGUEROA PRECIADO[c], MAYRA TOCTO ERAZO[d]

Recibido 11-02-2022, aceptado 29-06-2022, versión final 30-06-2022
Research paper

ABSTRACT: Systems of differential equations are used as the basis to define mathematical structures for moments, like the mean and variance of probability distributions of random variables. Nevertheless, the integration of a deterministic model and a probabilistic one, in order to describe a random phenomenon, and take advantage of the observed data for making inferences on certain population dynamic characteristics, can lead to parameter identifiability problems for the observed sample. Furthermore, approaches to deal with those problems are usually inappropriate. In this paper, the shape of the likelihood function of a SIR-Poisson model is used to describe the relationship between flat likelihoods and the practical parameter identifiability problem. In particular, we show how a flattened shape for the profile likelihood of the basic reproductive number R_0, arises as the observed sample (over time) becomes smaller, causing ambiguity regarding the shape of the average model behavior. We conducted some simulation studies to analyze the flatness severity of the R_0 likelihood, and the coverage frequency of the likelihood-confidence regions for the model parameters. Finally, we describe some approaches to deal the practical identifiability problem, showing the impact that those can have on inferences. We believe this work can help to raise awareness on the way statistical inferences can be affected by a priori parameter assumptions and the underlying relationship between them, as well as those arising by model reparameterizations and incorrect model assumptions.
KEYWORDS: Flat likelihood function; ordinary differential equations; SIR model, basic reproductive number; likelihood contours; profile likelihood function.

RESUMEN: Sistemas de ecuaciones diferenciales son usados como base para definir estructuras matemáticas para momentos, como la media y la varianza de distribuciones de probabilidad de variables aleatorias. Sin embargo, integrar un modelo determinista con uno de probabilidad para representar un fenómeno aleatorio y aprovechar los datos observados de dicho fenómeno para hacer inferencia sobre características de la dinámica poblacional, puede producir problemas de identificabilidad de parámetros para la muestra observada. Más aún, dichos problemas suelen ser tratados de forma inapropiada. En este artículo se utiliza la forma de la función de verosimilitud del modelo SIR-Poisson para describir la relación entre funciones de verosimilitud planas y el problema de identificabilidad práctica de parámetros. En particular, se muestra cómo la forma aplanada de la función de verosimilitud perfil del número reproductivo básico R0 aparece a medida que la muestra observada (en el tiempo) se reduce y produce ambigüedad en la forma del comportamiento medio del modelo. Se efectuan diversos estudios de simulación con el fin de analizar

[a]Montoya, J. A., Figueroa-Preciado, G. & Tocto Erazo, Mayra (2022). Flat likelihoods: SIR-Poisson model case. *Rev. Fac. Cienc.*, 11 (**2**), 74–99. DOI: https://doi.org/10.15446/rev.fac.cienc.v11n2.100986

[b]PhD in Probability and Statistics. Statistics Professor. Department of Mathematics. University of Sonora, México.
[*]Corresponding author: arturo.montoya@unison.mx
[c]PhD in Probability and Statistics. Statistics Professor. Department of Mathematics. University of Sonora.
[d]PhD in Mathematics. Mathematics Professor. Department of Mathematics. University of Sonora, México.

la severidad de la planura de la verosimilitud de R_0 y la frecuencia de cobertura de las regiones de verosimilitud-confianza de los parámetros del modelo. Finalmente, se describen algunos tratamientos que se han utilizado para abordar el problema de identificabilidad práctica y se muestra el impacto que éstos pueden tener en las inferencias. Se considera que este trabajo ayudará a tomar conciencia sobre cómo las inferencias estadísticas se ven afectadas por suposiciones a priori sobre los parámetros y la relación subyacente entre ellos, así como por reparametrizaciones del modelo y por suposiciones incorrectas sobre el mismo.
PALABRAS CLAVE: Función de verosimilitud plana; ecuaciones diferenciales ordinarias; modelo SIR; número reproductivo básico; contornos de verosimilitud; función de verosimilitud perfil.

1. INTRODUCTION

Systems of differential equations are mathematical tools broadly used for modeling physical, chemical, biological and epidemiological processes, among others. In particular, mathematical models based on ordinary differential equations have been widely proposed to study different infectious diseases such as dengue, Zika, Ébola, COVID-19, cholera, among others. For this modeling approach, the SIR model (Kermack & McKendrick, 1927) has been the basis of many models (Arino & Van den Driessche, 2003; Pandey *et al.*, 2013; Khan *et al.*, 2014; Xiao & Zou, 2014; Cosner, 2015; Lee & Castillo-Chavez,, 2015; Hendron & Bonsall, 2016; Phaijoo & Gurung, 2016; Kim *et al.*, 2017; Ghosh *et al.*, 2018; Mishra *et al.*, 2018; Mishra & Gakkhar, 2018; Nguyen *et al.*, 2018; Qi *et al.*, 2018; Sasmal *et al.*, 2018; Tuncer *et al.*, 2018; Vinh *et al.*, 2018; Tocto-Erazo *et al.*, 2020; Tocto-Erazo *et al.*, 2021). These models usually provide a threshold value, known as the basic reproductive number (R_0), that indicates the severity of the disease, as well as the existence of endemic states. Estimation of parameter R_0 is relevant in public health decision making, being very helpful for implementing adequate controls for infectious diseases.

In the problem regarding the estimation of parameter R_0, as well as some other parameters of systems of differential equations, it is common to use a Poisson model or a Negative Binomial distribution, to describe the frequency distribution of random variables associated to different observable characteristics of the dynamic under study, like the number of infected people, number of deaths, and some others (Chowell *et al.*, 2007; Chowell *et al.*, 2008; Lloyd-Smith, 2007; Chowell *et al.*, 2012; Capistrán *et al.*, 2012; Acuña-Zegarra *et al.*, 2020; Acuña-Zegarra *et al.*, 2021; Núñez-López *et al.*, 2021). In these cases it is common to use systems of differential equations as the basis for defining mathematical structures for the mean and variance of these probability distributions. In this paper we consider the Poisson model, whose mean and variance are fully determined by a SIR model. We selected both models (the SIR and the Poisson one) due to: convenience regarding the number of unknown parameters and computational effort; they are representative in the statistical and ordinary differential equations literature. Despite the mathematical simplicity of these models, they enable us to illustrate the essence of the parameter identifiability problem for an observed sample, the main objective of this manuscript; allowing also to show not only how this problem have been handled, but also the impact it can have on resulting inferences.

Regarding the negative binomial case, even though it is also a counting model used in the differential equations parameter estimation, this will not be addressed here. This model has two unknown parameters, a probability of exit and an overdispersion parameter; both involved in the mathematical expression of its mean or expected value. Therefore, it could happen that the occurrence of a lack of parameter identifiability, based on the observed sample, could not only be caused by the relationship between the mean of the negative binomial distribution and the SIR model parameters, and their own relationship, but also could be due (to some extent) by the negative binomial model itself, caused by the relationship between its parameters. In fact, we consider that the analysis of this distribution, in the context studied in this manuscript, is a research topic in its own right.

The problem of structural and practical identifiability, in models involving systems of differential equations, has already been reported in literature by several authors (Rosenbaum *et al.*, 1999; Raue *et al.*, 2009; Zhan *et al.*, 2015; Tuncer *et al.*, 2016; Gábor *et al.*, 2017; Kao & Eisenberg, 2018; Marquis *et al.*, 2018; Saccomani & Thomaseth, 2018). In this paper we are interested in the problem of parameter identifiability for an observed sample (practical identifiability), which mainly occurs when relatively flat likelihood functions are observed, regarding to their maximum, yielding likelihood-confidence regions that cover a large part of the parameter space, and possibly extending infinitely (Raue *et al.*, 2009; Cole, 2020). In particular, for the case of the SIR-Poisson model we are interested in describing the relationship between flat likelihood functions and the practical parameter identifiability problem, considering not only a standard reparameterization of the SIR model, but also another one involving the basic reproductive number R_0. Moreover, we are also interested about quantifying the frequency of occurrence of flat likelihoods, and to understand the observed sample configurations that cause these kind of likelihoods. In this paper we also show some approaches that have been used to address the problem of parameter identifiability, pointing out the impact that these can have on inferences. This is done in order to understand the effect that our assumptions and reparameterizations have on inferences, and to become aware about the validity of the uncertainty quantification in the statistical estimation of the parameters of systems of differential equations.

In this paper we perform some simulations in order to explore the extent to which the SIR-Poisson model leads to flat likelihoods, analyzing the coverage frequency of the likelihood-confidence regions of the model parameters. In particular, we focus on the behavior of the relative profile likelihood function of R_0, when R_0 takes large values, showing that flat likelihoods frequently occur when the observed sample (over time) is reduced, producing ambiguity in the shape of the average model behavior. Thus, in real applications, the upper limit of R_0 confidence intervals would be practically infinite, a result that is not very useful for practical purposes, in a real problem.

However, the coverage frequencies of these intervals are still close to the nominal probability coverage. Finally, in this paper we describe three approaches that have been used to address the problem of practical

parameter identifiability, showing the impact these can have on inferences: one of them involves setting values to the unknown parameters; another approach relies on estimating the growth rate of the SIR model, by assuming an exponential growth model for the data; and the third one involves performing Monte Carlo simulations to estimate R_0, based on a prior knowledge for the distributions of the parameters of the SIR model. In either case, we illustrate how quantitative claims about parameters and uncertainty estimation are both completely determined by these estimation procedures.

This paper is organized as follows. In Section 2, a SIR model is reparametrized in terms of an unknown parameter vector that involves the mean and variance of a Poisson model. The likelihood function, as well as the relative likelihood function are defined and used to construct likelihood regions for this parameter vector. Profile likelihood function is also defined and used to test hypothesis concerning a scalar parameter. All these functions allow to easy visualize plausible and implausible parameter values. Section 3 includes an example where a sample from a SIR-Poisson distribution is simulated, in order to show situations where different parameter values give rise to same likelihood, for the observed data; illustrating the practical parameter identifiability problem. On the other hand, in Section 4, a simulation study to analyze the shape of the profile likelihood function of parameter R_0, the basic reproductive number, is performed. The selected scenarios and experimental features allow to analyze how severe the flatness of this function can become. Section 5 includes three common approaches used to deal with flat likelihoods in the SIR-Poisson model case. Some situations that may lead to inappropriate or overestimated inferences are carefully discussed. Finally, some general conclusions are presented in Section 6.

2. SIR-POISSON LIKELIHOOD

Suppose that $X_1, X_2, \cdots, X_n$ are independent random variables, observed at times $t_1, \ldots, t_n$, respectively. Suppose furthermore that X_i is Poisson distributed with mean

$$\lambda_i = \int_{t_{i-1}}^{t_i} \beta \frac{I(t)}{N} S(t)\, dt, \tag{1}$$

where N is a constant population, $\beta > 0$, and $\gamma > 0$ are the transmission and recovery rates, $I(t)$ and $S(t)$ are solutions for the SIR model given by

$$\begin{aligned} \dot{S} &= -\beta \frac{I(t)}{N} S(t), \\ \dot{I} &= \beta \frac{I(t)}{N} S(t) - \gamma I(t), \\ \dot{R} &= \gamma I(t), \end{aligned} \tag{2}$$

$N = S + I + R$, and initial conditions at time $t_0 = 0$ are fixed and known: $S(t_0) = S_0$, $I(t_0) = I_0$, and $R(t_0) = N - S_0 - I_0$.

Note that λ_i in (1) relies on the SIR model reparameterization given in (2), so for a sake of clarity we will write it as $\lambda_i(\theta)$ instead of λ_i, in order to emphasize that SIR model is reparametrized in terms of the unknown parameters vector $\theta = (\beta, \gamma)$, and also that the unknown parameters of the Poisson model, governing its mean and variance, are β and γ. This kind of Poisson model will be called SIR-Poisson model. It is important here to note that $\theta = (\beta, R_0)$ indicates that the SIR-Poisson model is reparametrized in terms of the transmission rate β and the basic reproductive number (or basic reproductive ratio) $R_0 = \beta/\gamma$; that is, $(\beta, \gamma) \leftrightarrow (\beta, \beta/R_0)$. We consider this reparameterization since it involves R_0, a parameter of considerable interest in infectious disease control literature, for which some estimation techniques have been developed when there is a presence of parameter identifiability problems.

Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ be an observed sample from $X_1, X_2, \cdots, X_n$. The likelihood function of parameter vector $\theta = (\beta, \gamma)$, associated to the observed sample $\boldsymbol{x}$, can be written as

$$L(\theta; \boldsymbol{x}) = \prod_{i=1}^{n} \frac{1}{x_i!} [\lambda_i(\theta)]^{x_i} \exp[-\lambda_i(\theta)], \tag{3}$$

where $\theta = (\beta, \gamma) \in \mathbb{R}^+ \times \mathbb{R}^+$. Note that the likelihood function of (β, R_0) is obtained by replacing $\theta = (\beta, \beta/R_0)$ en (3).

The relative likelihood function of θ is a standardized version of (3). It takes the value of one at the maximum likelihood estimate of θ,

$$R(\theta; \boldsymbol{x}) = \frac{L(\theta; \boldsymbol{x})}{\max_{\theta} L(\theta; \boldsymbol{x})} = \frac{L(\theta; \boldsymbol{x})}{L\left(\widehat{\theta}; \boldsymbol{x}\right)}, \tag{4}$$

where $\widehat{\theta} = \left(\widehat{\beta}, \widehat{\gamma}\right)$ is the maximum likelihood estimator (MLE) of parameter $\theta = (\beta, \gamma)$. Using the invariance likelihood property, the MLE for R_0 results $\widehat{R}_0 = \widehat{\beta}/\widehat{\gamma}$. The relative likelihood in (4) measures the plausibility of any specified θ value, relative to that of $\widehat{\theta}$. This likelihood function ranks all possible θ values according their respective plausibilities, in light of the observed sample $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$. Note that $\theta \in \mathbb{R}^+ \times \mathbb{R}^+$, thus, given a plot of $R(\theta; \boldsymbol{x})$ we can easily distinguish plausible and implausible values for $\theta = (\beta, \gamma)$, or equivalently for $\theta = (\beta, R_0)$. Furthermore, a plot of $R(\theta; \boldsymbol{x})$ can be used also to visualize the flatness of the likelihood surface.

Another practical way to make inferences about θ, which also allows us to visualize and describe the likelihood surface, is through likelihood regions. The likelihood region for θ, at likelihood level c, where $c \in [0, 1]$, is given by

$$\mathcal{C}_\theta(c; \boldsymbol{x}) = \{\theta : R(\theta; \boldsymbol{x}) \geq c\}. \tag{5}$$

Any θ value within this region has a relative likelihood $R(\theta; \boldsymbol{x}) \geq c$, while $R(\theta; \boldsymbol{x}) < c$ for those outside this region; so just looking at the contour plot we can easily identify plausible and implausible θ values, at likelihood level c. Now, varying c from 0 to 1, we obtain a complete set of nested likelihoods regions

converging to the MLE $\widehat{\theta}$, as $c \to 1$; see Sprott (2000, pp. 14-15).

On the other hand, a $100(1-\alpha)$ % confidence interval can be assigned to $\mathcal{C}_\theta(c;\boldsymbol{x})$ in (5), considering the asymptotical distribution of the likelihood ratio $-2\ln R(\theta;\boldsymbol{X})$, under $H_0 : \theta = \theta_0$. This likelihood ratio has a Chi-squared distribution with two degrees of freedom, χ^2_2, (Kalbfleisch, 1985; Serfling, 2002). When $-2\ln c = q_{2,1-\alpha}$, where $q_{2,1-\alpha}$ is the quantile of probability $1-\alpha$ of a Chi-squared distribution with two degrees of freedom, the confidence level for $\mathcal{C}_\theta(c;\boldsymbol{x})$ results $100(1-\alpha)$ %. For example, when $c = 0.05$, the confidence level is approximately 95 %. Moreover,

$$P\text{-value} = P\left\{\chi^2_2 \geq -2\ln R(\theta_0;\boldsymbol{x})\right\} = R(\theta_0;\boldsymbol{x}),$$

and every specified θ value has an associated P-value greater or equal than $c = 0.05$ when lying within this relative likelihood contour, and less than $c = 0.05$ otherwise.

Now, assume that vector parameter θ can be partitioned as follows: an interest parameter ψ and a nuisance parameter λ. Then, the profile likelihood of ψ and its corresponding relative likelihood function are defined as

$$\begin{aligned} L_{\max}(\psi;\boldsymbol{x}) &= \max_{\lambda} L(\psi,\lambda;\boldsymbol{x}), \\ R_{\max}(\psi;\boldsymbol{x}) &= \frac{L_{\max}(\psi;\boldsymbol{x})}{\max\limits_{\psi,\lambda} L(\psi,\lambda;\boldsymbol{x})} = \frac{L_{\max}(\psi;\boldsymbol{x})}{L_{\max}(\widehat{\psi};\boldsymbol{x})} \end{aligned} \tag{6}$$

For example, we might be interested in making inferences about R_0, so the profile likelihood of $\psi = R_0$, and its corresponding relative likelihood function can be calculated as

$$\begin{aligned} L_{\max}(R_0;\boldsymbol{x}) &= \max_{\beta} L(\theta = (\beta, \beta/R_0);\boldsymbol{x}), \\ R_{\max}(R_0;\boldsymbol{x}) &= \frac{L_{\max}(R_0;\boldsymbol{x})}{\max\limits_{\theta} L(\theta;\boldsymbol{x})} = \frac{L_{\max}(R_0;\boldsymbol{x})}{L_{\max}\left(\widehat{R}_0;\boldsymbol{x}\right)}. \end{aligned} \tag{7}$$

The relative profile likelihood varies between 0 and 1, and ranks all possible R_0 values based on the observed sample $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ on times $t_1, \ldots, t_n$. Thus, a plot of $R_{\max}(R_0;\boldsymbol{x})$ allows to distinguish plausible and implausible R_0 values. Note that we might be interested in making inferences about β, so we can use $\theta = (\beta,\gamma)$ and consider $(\psi,\lambda) = (\beta,\gamma)$.

A level c profile likelihood region (commonly an interval) for the scalar parameter of interest ψ is given by

$$\mathcal{C}_\psi(c;\boldsymbol{x}) = \{\psi : R_{\max}(\psi;\boldsymbol{x}) \geq c\}, \tag{8}$$

where $0 \leq c \leq 1$. Considering that the profile likelihood ratio statistic $-2\ln R_{\max}(\psi;\boldsymbol{X})$ follows an asymptotically χ^2_1 distribution, if the null hypothesis $H_0 : \psi = \psi_0$ is true, then (8) is a confidence interval for

the scalar parameter ψ, with approximate confidence level determined from $-2\ln R_{\max}(\psi_0;\boldsymbol{X}) \sim \chi^2_1$. When $-2\ln c = q_{1,1-\alpha}$, where $q_{1,1-\alpha}$ is the quantile of probability $1-\alpha$ of a Chi-squared distribution with one degree of freedom, the confidence level for $\mathcal{C}_\psi(c;\boldsymbol{x})$ is $100(1-\alpha)$ %. For example, the confidence level will be approximately 90 %, 95 %, and 99 % if $c = 0.25$, 0.146, and 0.036, respectively; see Kalbfleisch (1985, pp. 115-116).

In general, the profile or maximized likelihood is a powerful but simple inferential method for estimating a parameter of interest, separately from the remaining unknown parameters of the statistical model. More details about the profile likelihood approach are described in Kalbfleisch (1985, Section 10.3), Pawitan (2001, Section 3.4), Sprott (2000, p. 66), Barndorff-Nielsen & Cox (1994, pp. 89-91), Serfling (2002, pp. 155-160), and Murphy & Van Der Vaart (2000). An important and mostly unknown aspect about the profile likelihood function is that it can be used to study and visualize several features of the whole likelihood function, such as for instance those associated with the flatness of the likelihood surface. The example included in the following section illustrates how the graph of a profile likelihood, particularly the case of R_0, can reveal the flat nature of the likelihood function of a SIR-Poisson model.

3. EXAMPLE: SIR-POISSON LIKELIHOOD

The dataset in Table 1 has been generated using a SIR-Poisson distribution, with vector parameters $(\beta,\gamma) = (1.786946, 0.4761905)$, $N = 763$, and initial conditions $I_0 = 1$ and $S_0 = 762$. Figure 1 shows, in grayscale, the relative likelihood function for (β,γ) and its corresponding contour plot. In particular, the $c = 0.05$ contour level is indicated with a dashed line, and the MLE is marked with an asterisk ($\widehat{\beta} = 1.845084$ and $\widehat{\gamma} = 0.520862$). Just looking at the shape of the likelihood surface, and its elongated contours, there is evidence of the close relationship between parameters β and γ. Nevertheless, the region of plausible values for both parameters is clearly into the plots limits. In this case, profile likelihood functions for β and γ are well-behaved, in the sense that they grow, reach a maximum and then decrease; see Figure 2(a)-(b). The 95 % confidence intervals for β and γ are $\mathcal{C}_\beta(c = 0.146;\boldsymbol{x}) = [1.644828, 2.063218]$ and $\mathcal{C}_\gamma(c = 0.146;\boldsymbol{x}) = [0.3068966, 0.7609195]$, respectivamente.

It is important to note that in many applications, where parameters of differential equation models are estimated, only data with an increasing behavior are available or used, data that apparently have already reached a peak. In the following example we simulate this kind of scenario, assuming that only 40 % of the observations are available; that is, only the first four observations in Table 1, $\boldsymbol{x} = (2, 12, 59, 130)$ are known. Now, Figure 3 shows the relative likelihood function for (β,γ) and its corresponding contour plot; in this plot the $c = 0.05$ contour level is marked with a dashed line and the MLE is indicated with an asterisk ($\widehat{\beta} = 1.9047994$ and $\widehat{\gamma} = 0.5635199$). In this case, the resulting contours are even more elongated than those obtained when considering the 100 % of the simulated sample, emphasizing again the close relationship between both parameters, β and γ. Now, the region of plausible values for both parameters is truncated within

Table 1: Simulated sample from a SIR-Poisson distribution with $\beta = 1.786946$, $\gamma = 0.4761905$, $N = 763$, $I_0 = 1$, and $S_0 = 762$.

i	1	2	3	4	5	6	7	8	9	10
t_i	1	2	3	4	5	6	7	8	9	10
x_i	2	12	59	130	206	171	92	42	19	8

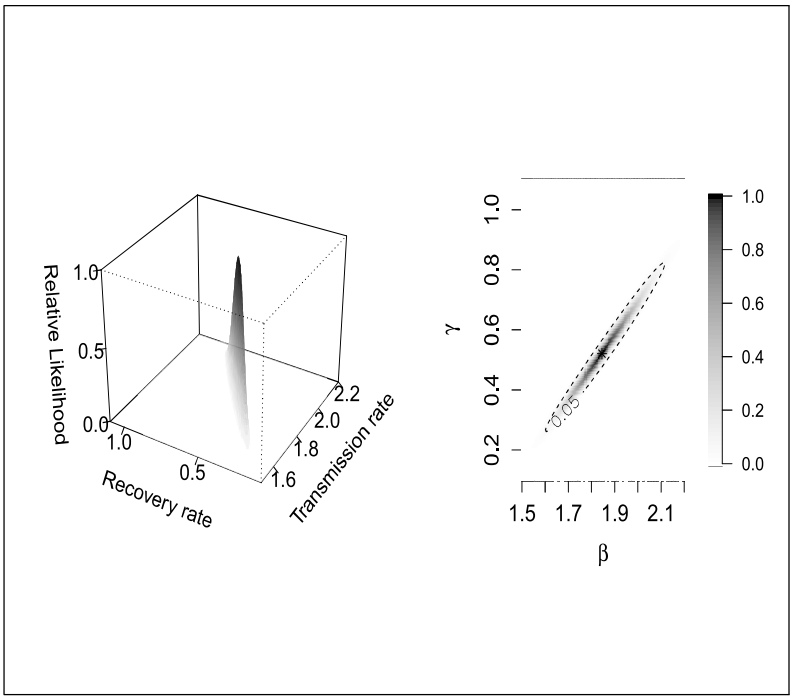

Figure 1: Likelihood surface and contour plot for the relative likelihood of (β, γ), corresponding to data presented in Table 1. The grayscale bar shows different values for the relative likelihood; the 95 % confidence region is plotted with a dashed line, at $c = 0.05$ contour level, and the MLE is marked with an asterisk. Source: Elaborated by the authors.

the limits of the plots. This can be better observed in the profile likelihood functions of both parameters, plotted in Figure 4(a)-(b); particularly, the γ profile likelihood function assigns a high plausibility to those γ values close to zero. Note that due to the intrinsic relationship exhibited by both parameters, the β profile likelihood function also presents a similar behavior. The upper limits of the 95 % confidence intervals for β and γ are 2.925301 and 1.656964, respectively.

Now, regarding the model reparameterization (β, R_0), the shape of the likelihood surface is much flatter than the observed in the other cases, with elongated and curved contours, as can be seen observed in Figure 5. This plot shows that the likelihood function has a ridge, with a relatively flat top.

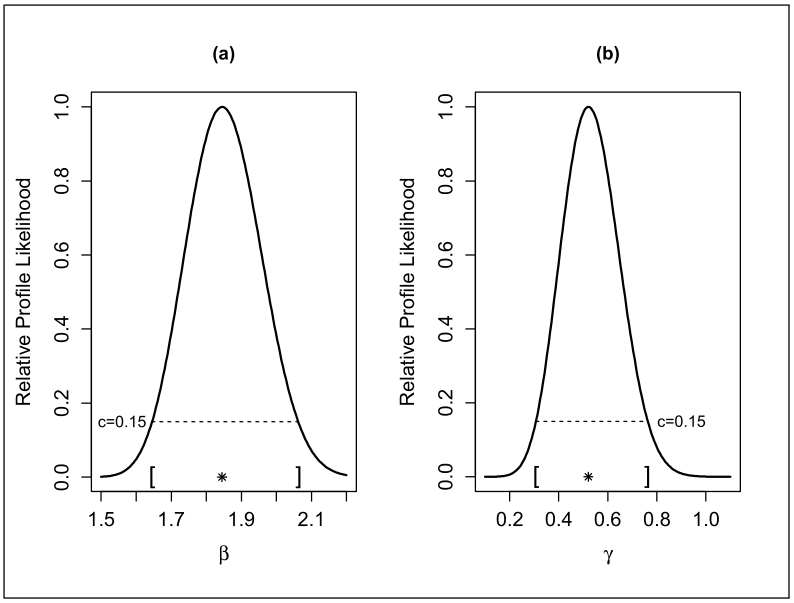

Figure 2: Relative profile likelihood functions for β and γ, based on the simulated data presented in Table 1. In both cases, (a) and (b), the 95% confidence intervals are indicated in brackets, and the MLE is marked with an asterisk. Source: Elaborated by the authors.

The profile relative likelihood of R_0, that is shown in Figure 6, illustrates this kind of flatness, where the likelihood increases until reaching its maximum, decreasing then slowly; actually, plausibility of $R_0 = 6$ and $R_0 = 10$ are 0.85 and 0.72, respectively. The lower limit of the 95% confidence interval for R_0 results 1.769076.

When for a fixed data set we observe two different parameter values giving rise to the same likelihood (probability of observed data), then it results practically impossible to distinguish between these two parameter candidates, based on the data alone; thus, it would be not feasible to identify the true parameter value. In that sense, when the relative profile likelihood function of R_0 is flat over a wide range of R_0 values, including the MLE such as in Figure 6, then the probability of the observed sample is the same over all this R_0 value set. Therefore, the lack of identifiability of this parameter occurs at least for the observed data. Now, since the profile likelihood function for the parameter of interest is invariant under reparameterizations of the nuisance parameters, then, the plot of the profile likelihood function of R_0 will not change under reparameterizations of (β, γ).

pagebreak

Sprott (2000, p. 11) stated that the MLE and the observed information exhibit two important features of the likelihood function, the MLE as a measure of its location relative to the parameter-axes, while the latter as a proxy of its curvature, in a neighborhood of the MLE. Thus, when the likelihood function of a scalar parameter is smooth and symmetric around the MLE, then the MLE and the observed information are usually

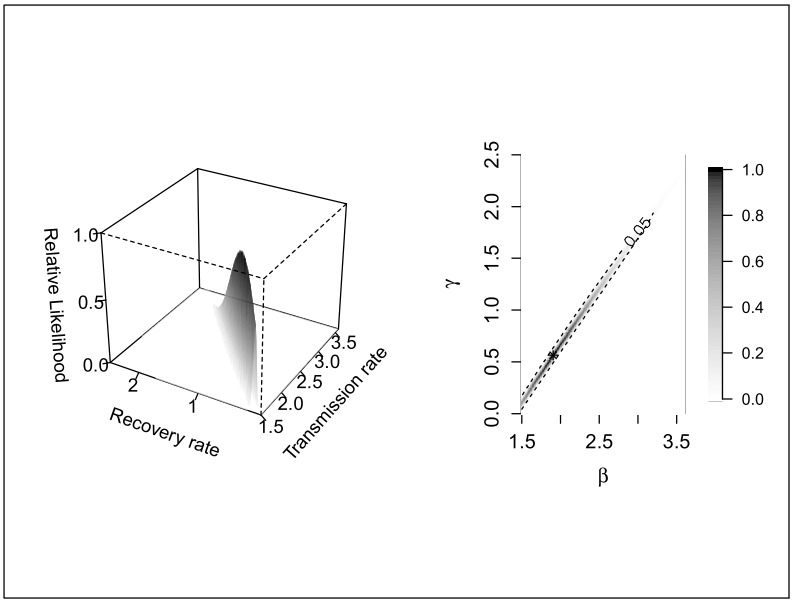

Figure 3: Likelihood surface and contour plots of the relative likelihood of (β,γ) corresponding to $\boldsymbol{x} = (2,12,59,130)$ data in Table 1. Grayscale bar shows different relative likelihood values; the 95 % confidence region is denoted with a dashed line at $c = 0.05$ contour level, whereas the MLE is marked with an asterisk. Source: Elaborated by the authors.

the main features that determine the shape of the likelihood function. However, it should be stressed that flat likelihoods can not be determined by these quantities, without loss of information.

Intuitively, the sensitivity of the MLE depends on the shape of the likelihood function. If the R_0 likelihood is very curved around MLE, then, a stable estimate of R_0 is obtained. On the other hand, if the likelihood is flat near the MLE, this estimate becomes highly unstable. Beside these facts, it is important to mention that a flat likelihood can lead to R_0 likelihood-confidence intervals whose upper limit becomes infinity.

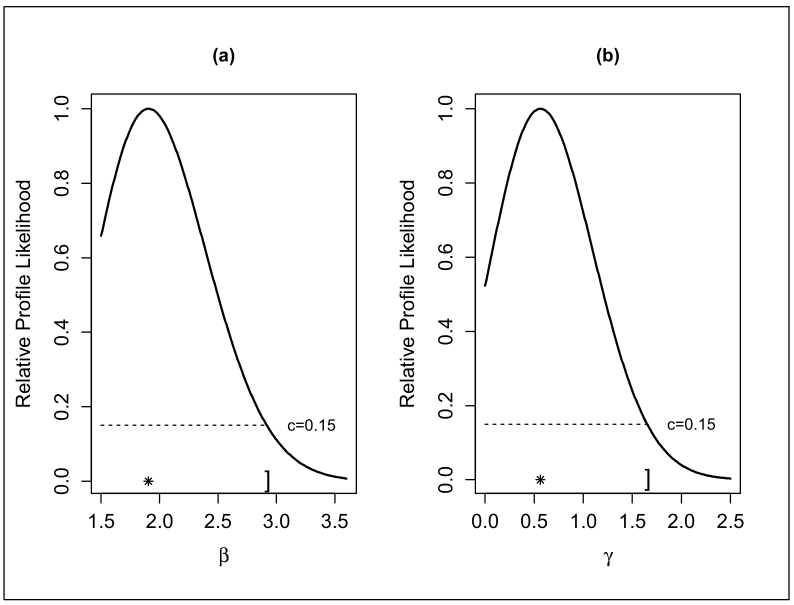

Figure 4: The relative profile likelihood function of β and γ for the simulated data $\boldsymbol{x} = (2, 12, 59, 130)$ in Table 1. In both cases, (a) and (b), the upper limit of the 95 % confidence interval is denoted with a bracket, whereas the MLE is marked with an asterisk. Source: Elaborated by the authors.

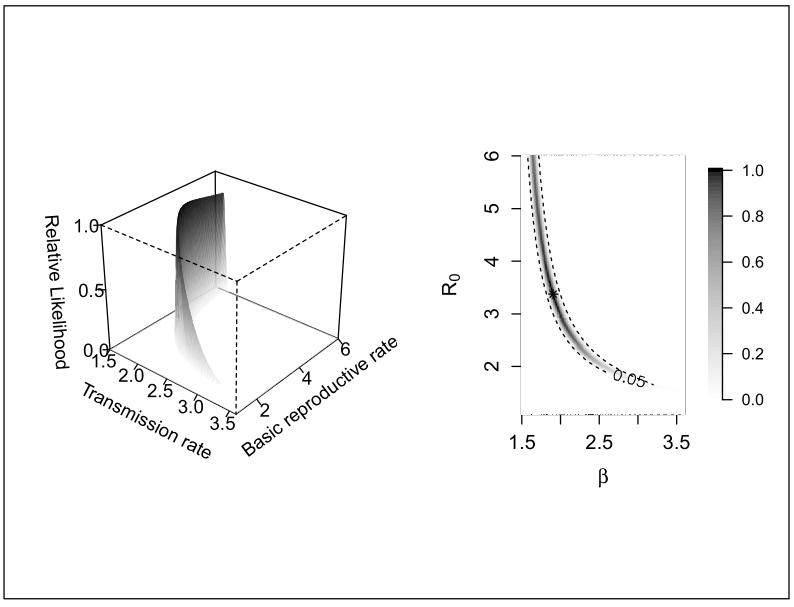

Figure 5: Likelihood surface and contour plots of the relative likelihood of (β, R_0), corresponding to $\boldsymbol{x} = (2, 12, 59, 130)$ data in Table 1. Grayscale bar shows different relative likelihood values; the 95 % confidence region is denoted with a dashed line at $c = 0.05$ contour level, whereas the MLE is marked with an asterisk. Source: Elaborated by the authors.

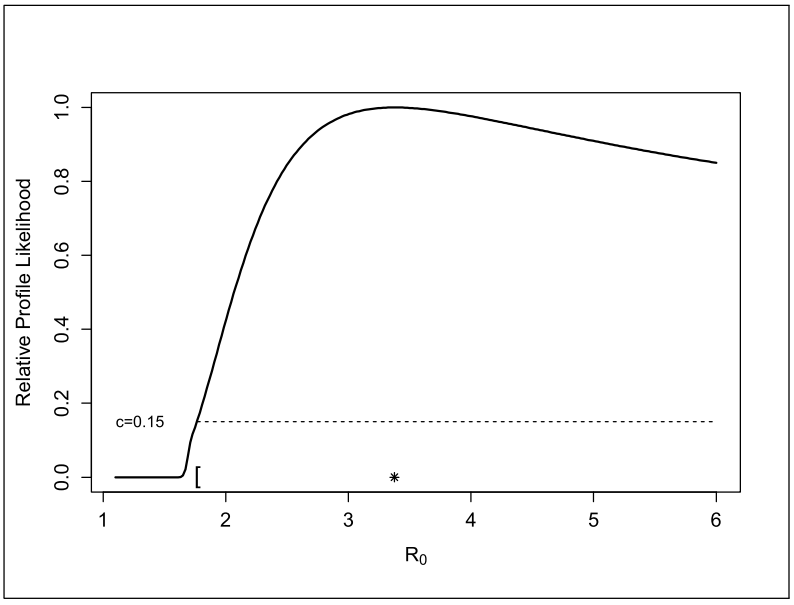

Figure 6: Relative profile likelihood function of R_0 for simulated $\boldsymbol{x} = (2, 12, 59, 130, 206)$ data in Table 1. The upper limit of the 95 % confidence interval is indicated with a bracket and the MLE is marked with an asterisk. Source: Elaborated by the authors.

4. FLATNESS OF THE RELATIVE PROFILE LIKELIHOOD OF R_0

In this section we will conduct some experiments, in order to analyze how the shape of the R_0 profile likelihood function changes, as the observed sample size varies over time. For that purpose, we will examine the distribution of the relative profile likelihood function of R_0, evaluated at different values, all of them larger than the nominal ones used in the simulation scenarios. In particular, we are interested in analyzing the severity of the flatness of the profile likelihood of R_0, when the observed sample (over time) produces ambiguity regarding the shape of the average behavior of the SIR-Poisson model. Furthermore, in all experiments considered here, we will compute the coverage frequency of the likelihood-confidence regions for (β, R_0), β and R_0, in order to study the effect that flat likelihoods could have on the coverage probabilities of these intervals.

In this simulation study we consider 5000 SIR-Poisson samples of sizes $n_1 = 40\,\%n$, $n_2 = 60\,\%n$, $n_3 = 80\,\%n$, and $n_4 = 100\,\%n$, simulated according the following scenarios for n and R_0:

- Case 1, $n = 23$ and $(\beta, \gamma) = (0.893473, 0.4761905)$, that yields $n_1 = 9$, $n_2 = 14$, $n_3 = 18$, and $n_4 = 23$, and $R_0 = 1.876293$.
- Case 2, $n = 10$ and $(\beta, \gamma) = (1.786946, 0.4761905)$ that yields $n_1 = 4$, $n_2 = 6$, $n_3 = 8$, and $n_4 = 10$, and $R_0 = 3.752587$.
- Case 3, $n = 7$ and $(\beta, \gamma) = (2.680419, 0.4761905)$ that yields $n_1 = 3$, $n_2 = 4$, $n_3 = 6$, and $n_4 = 7$, and $R_0 = 5.62888$.

In our simulation scenarios, samples of size n were selected in such a way that maximum is already reached and $I(n) / \max\{I(t)\} \approx 0.31$, where $I(t)$ is the solution of the SIR model. Figure 7 shows the SIR model solution $I(t)$, for each of the R_0 values considered in our simulation cases. It is noteworthy that the n selected values in our simulation experiment allow that λ_i, the mean of the SIR-Poisson model given in (1), grows until reaching its maximum, decreasing later to values close to zero. In this way, $n_1 = 40\,\%n$, $n_2 = 60\,\%n$, $n_3 = 80\,\%n$, and $n_4 = 100\,\%n$, represent observed sample sizes that produce a gradient of information associated to different ambiguity levels regarding the shape of the average behavior of the model.

Based on the simulated samples, we compute the coverage frequency of the likelihood confidence regions $C_{(\beta,R_0)}(0.05;\boldsymbol{X})$, $C_{(\beta)}(0.146;\boldsymbol{X})$ and $C_{(R_0)}(0.146;\boldsymbol{X})$, whose nominal coverage probability is 95 %. Furthermore, for each simulation experiment we also computed the empirical distribution for $R_{\max}\left(R_0 = R_0^U;\boldsymbol{X}\right)$, where $R_0^U = 10$, $R_0^U = 15$, and $R_0^U = 20$. It is worth mentioning that MLEs were estimated using the constrOptim function of the R 4.0.3 version software, and that all results were obtained considering only those samples where the constrOptim function yielded a successful performance of the Nelder and Mead optimization method. For example, in Case 1 $(\beta, \gamma) = (0.893473, 0.4761905)$ with $n = 9$ and Case 2 $(\beta, \gamma) = (1.786946, 0.4761905)$ with $n = 4$ convergence of the method was obtained for the 95 % and 99 % of all 5000 simulated samples. In all remaining scenarios, 100 % of simulated samples successfully converged.

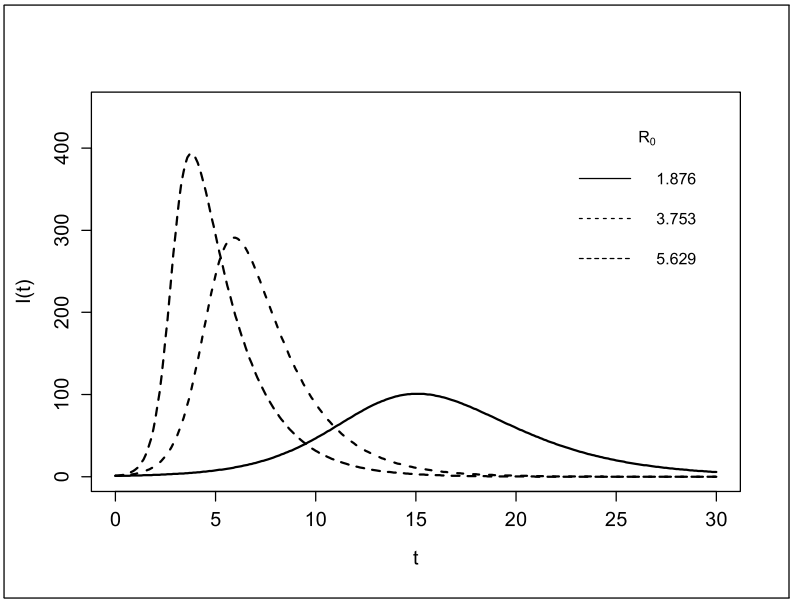

Figure 7: $I(t)$ solutions for the SIR model for different R_0 values considered in simulation scenarios. Source: Elaborated by the authors.

As can be observed in Table 2, in all cases the coverage frequency of the likelihood-confidence regions for the SIR-Poisson model parameters is close to the 95 % nominal probability coverage. On the other hand, it can be observed in Figures 8 to 10, that those scenarios where sample size yield ambiguity regarding the shape of the average behavior of the SIR-Poisson model, the median of the distribution of $R_{\max}\left(R_0 = R_0^U; \boldsymbol{X}\right)$ in $R_0^U = 10$, $R_0^U = 15$, and $R_0^U = 20$ takes positive values that are far from 0, with a downward trend thereafter, though slightly leaning, allowing to infer that the shape of the R_0 relative profile likelihood function becomes fundamentally flat for $R_0 \in [10, 20]$.

For instance, in all cases (Case 1-3) where sample size is $n_1 = 40\,\%n$, the 0.35 level quantile of the R_0^U empirical distributions was larger than the relative likelihood level at $c = 0.146$. That is, each one of the corresponding samples yielded 95 % confidence intervals for R_0, where the interval upper limit was greater or equal to $R_0^U = 20$. This result can be considered as uninformative, given the fact that real R_0 values were 1.876293, 3.752587, 5.62888, and almost absurd in the practical context of an epidemic. On the other hand, for all scenarios in cases with $100\,\%n$ and $R_0^U = 20$, is clear that at least 91 % of the times, the 95 % confidence intervals for R_0 have a finite upper limit, less than $R_0^U = 20$; thus, $R_{\max}\left(R_0 = 20; \boldsymbol{x}\right)$ was less or equal than the relative likelihood at $c = 0.146$ level.

Table 2: Coverage frequency of the likelihood-confidence regions for the SIR-Poisson model parameters.

$(\beta,\gamma) \to R_0 = \beta/\gamma$	n	$\mathcal{C}_{(\beta,R_0)}(0.05;\boldsymbol{X})$	$\mathcal{C}_{(\beta)}(0.146;\boldsymbol{X})$	$\mathcal{C}_{(R_0)}(0.146;\boldsymbol{X})$
$(0.893473, 0.4761905) \to R_0 = 1.876293$	9	96.58	97.36	97.30
	14	95.34	95.24	95.52
	18	95.58	95.40	95.40
	23	95.40	95.32	95.78
$(1.786946, 0.4761905) \to R_0 = 3.752587$	4	96.88	97.89	97.75
	6	94.78	94.74	95.32
	8	94.86	94.72	94.72
	10	94.92	94.82	94.60
$(2.680419, 0.4761905) \to R_0 = 5.62888$	3	95.74	96.62	96.38
	4	94.80	95.18	95.00
	6	94.78	94.88	95.14
	7	94.36	94.24	94.52

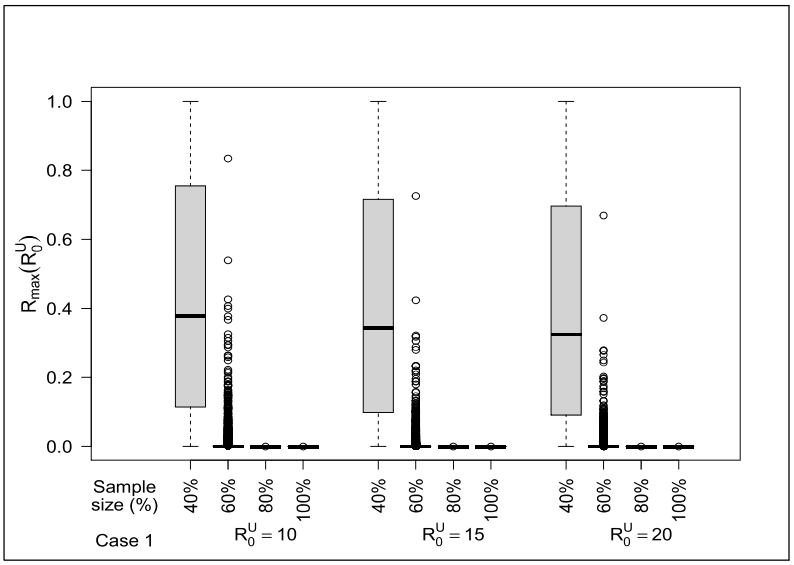

Figure 8: Box plots for $R_{\max}\left(R_0 = R_0^U\right)$ from simulated data under Case 1 parameter values. Source: Elaborated by the authors.

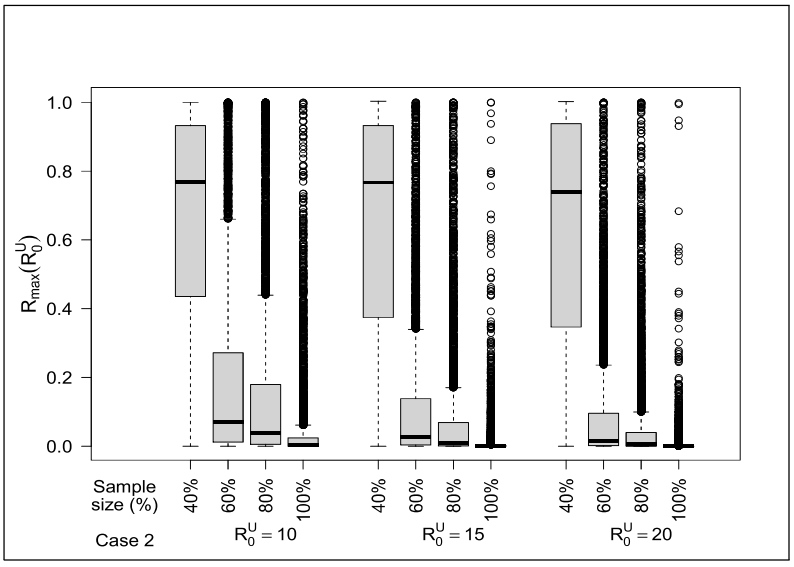

Figure 9: Box plots for $R_{\max}\left(R_0 = R_0^U\right)$ from simulated data under Case 2 parameter values. Source: Elaborated by the authors.

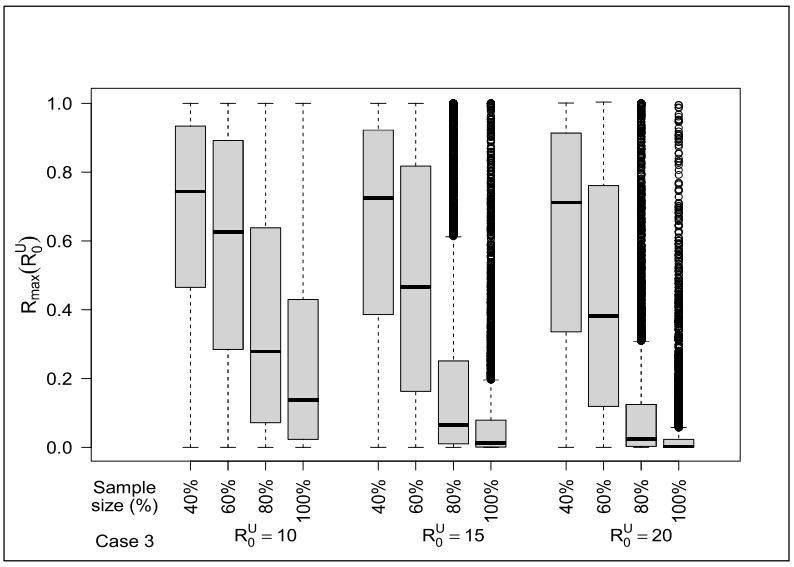

Figure 10: Box plots for $R_{\max}\left(R_0 = R_0^U\right)$ from simulated data under Case 3 parameter values. Source: Elaborated by the authors.

5. SOME ESTIMATION PROCEDURES IN THE PRESENCE OF FLAT LIKELIHOODS

In this section we describe three common parameter estimation approaches, basically applied to the SIR-Poisson model in the presence of flat likelihoods, an indication of a lack of parameter identifiability situation, based on the sample. One approach involves setting values to unknown parameters; another one concerns with estimating the growth rate $\delta = \beta - \gamma$ of the SIR model, assuming a data exponential growth. The third one uses Monte Carlo simulations, based on prior knowledge of the parameter distributions of the SIR model. In addition, we emphasize the impact that these practices can have on R_0 inferences.

5.1. Setting values to unknown parameters

In literature concerning parameter estimation in differential equations, the practice of setting values to an unknown parameter stands out (Khan *et al.*, 2014; Funk *et al.*, 2016; Gui-Quan *et al.*, 2017; Ghosh *et al.*, 2017; Vinh *et al.*, 2018; Guanghu *et al.*, 2019). In general, assigning values to an unknown parameter in a model, in order to estimate some other parameters of that model, causes statistical inferences (estimators, confidence intervals, precision, coverage probability, etc.) to depend on the selected parameter values, that are considered to be known. Our model under study, a SIR-Poisson distribution, is not an exception. To clarify this situation, we will just consider the observed SIR-Poisson model sample $\boldsymbol{x} = (2, 12, 59, 130)$, which represents 40 % of a simulated sample from a SIR-Poisson distribution with vector parameters $(\beta, R_0) = (1.786946, 3.752587)$, $N = 763$, and initial conditions $I_0 = 1$ and $S_0 = 762$. The likelihood functions linked to this sample were shown and described in Section 3. In particular, it was stated that the likelihood of (β, R_0), shown in Figure 3, displays a flat surface with elongated and curved contours, indicating the strong relationship between these parameters, based on the observed sample. Moreover, the relative profile likelihood of R_0 grows until achieving its maximum, decreasing slowly thereafter and assigning high plausibility values (close to 1) not only to true R_0 value, but also to some larger values like $R_0 = 6$ and $R_0 = 10$ (with plausibility 0.85 and 0.72, respectively).

In this particular case, setting parameter β in a fixed and known value β^* yields a likelihood $L(R_0; \boldsymbol{x}) = L(\beta = \beta^*, R_0; \boldsymbol{x})$ that only depends on R_0, the parameter of interest. Figure 11 shows the corresponding relative likelihood, standardized to one at its maximum, for different β values (included $\beta_0 = 1.786946$, the true value used in data simulation). Now, just for the sake of comparison, R_0 relative profile likelihood function and the true R_0 value are marked in this figure.

From this plot becomes clear that inferences are almost completely determined by the value that is specified for β, and only the likelihood corresponding to $\beta = \beta_0$ assigns a non-negligible plausibility to the true R_0 value. Moreover, the three β specified values give rise to likelihoods with different amplitudes, yielding different precision in the likelihood-confidence intervals for R_0, that are constructed considering the same relative likelihood level $0 < c < 1$. In fact, these amplitudes are determined by the relationship between β

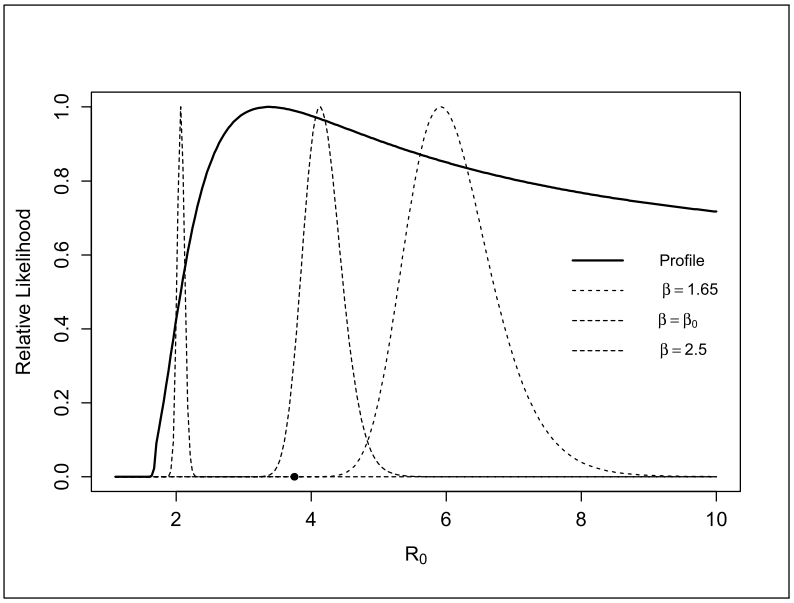

Figure 11: Relative likelihood function of R_0 for different β values ($\beta = 1.85$, $\beta = \beta_0$, and $\beta = 2.5$), and relative profile likelihood function of R_0. True R_0 value is marked with a black dot. Source: Elaborated by the authors.

and R_0, displayed in the contour plot of Figure 5. It is worth mentioning that simulated samples from previous section (Case 2, with $n = 4$) were used to estimate the coverage probability of the 95 % confidence intervals for R_0 that these likelihoods yield. Coverage frequencies associated to $\beta^* = 1.65$, 1.786946, and 2.5 were 2.8 %, 95.38 % and 0 %, respectively.

5.2. Estimation of $\delta = \beta - \gamma$ assuming an exponential growth model

Another approach for estimating parameters of differential equations is to estimate the parameters of a general model, but considering a simpler model for that data; one containing fewer parameters that are apparently involved in the general model. For example, many authors propose compartment models for the dynamics under study using an exponential growth model, where the growth rate δ is estimated independently from the general model (Chowell *et al.*, 2007; Chowell *et al.*, 2012; Towers *et al.*, 2016). Once δ estimates are obtained, they are used to set the parameter values of the general compartment model. However, a problem with this practice is that estimating δ, based on an exponential growth model, does not consider the intrinsic relationship between the parameters of the general differential equation model, originally proposed to model the dynamics under study. Consequently, inferences about these parameters could be very different from those that would be obtained if δ were estimated with the general compartment model.

We show this estimation approach considering the SIR model presented in (2), as a general model, and using a Poisson distribution with mean $f(t;\delta)$ to fit the observed data $\{(t_i;x_i)\}_{i=0}^{n}$, where $\ln(f(t;\delta)) = \ln(I_0) + \delta t$,

δ is the parameter that describes the growth rate in this exponential model, I_0 is the number of initially observed cases, and x_i is the number of new cases during the time interval $(t_{i-1}, t_i]$. In this case, the likelihood function of δ, that yields this exponential growth model is

$$L_{Exp}(\delta) \propto \exp\left(\delta \sum_{i=1}^{n} t_i x_i\right) \exp\left[-I_0 \sum_{i=1}^{n} \exp(\delta t_i)\right]. \tag{9}$$

Moreover, we will use the data set $\boldsymbol{x} = (2, 12, 59, 130)$, presented in previos subsection, in order to compare δ inferences that arise from $L_{Exp}(\delta)$, given in (9), against those inferences arising from the profile likelihood function of the exponential growth rate $\beta - \gamma$ (Ma, 2020), of the SIR model given in (2). In particular, and with certain notation abuse, this profile can be obtained by simply reparameterizing the SIR model with parameters β and γ, in terms of new parameters β and $\delta = \beta - \gamma$ (that is, $\gamma = \beta - \delta$), and then maximizing on β for each fixed δ value.

Figure 12 shows the relative version of (9), standardized to be one at its maximum. Moreover, and just for a comparison purpose, the δ profile likelihood function based on the SIR-Poisson model is also included in this figure, where the true δ value, obtained as the difference between β and γ true values, and used for data simulation, is marked with a black dot. In this figure we can observe that the relative likelihood of δ, yielded by the exponential model $R_{Exp}(\delta)$, is narrower than the profile likelihood of δ, and is located to its left. Moreover, $R_{Exp}(\delta)$ assigns negligible plausibility to many δ values that are plausible under the profile likelihood, included the true δ value. It is important to note that Case 2 simulated samples with $n = 4$ were also used here, in order to estimate the coverage probability of the 95 % confidence intervals of δ, yielded by $R_{Exp}(\delta)$ and δ profile likelihood functions (this last one computed under the basis of the SIR-Poisson general model). The resulting coverage frequencies associated to $R_{Exp}(\delta)$ and the profile likelihood were 0.04 % and 95.38 %, respectively.

5.3. Monte Carlo Simulation

Another approach to estimate the basic reproductive number R_0, of a system of differential equations with parameter θ, involves obtaining a mathematical expression for $R_0 = h(\theta)$, where $h(\cdot)$ is a function of θ, and assume a joint distribution $F(\cdot)$ for this parameter, that allows to generate $\theta_1, \theta_2, \cdots, \theta_M$ values, that are evaluated in $h(\cdot)$ in order to obtain R_0, $R_{01}, R_{02}, \cdots, R_{0M}$ values of the parameter of interest. These values are used to estimate the distribution of the basic reproductive number and to construct confidence intervals for R_0.

A difficulty that arises when trying to properly apply this practice, is obtaining a joint distribution $F(\theta)$ that rescues the intrinsic relationship between the parameters of the differential equation model, that is proposed to model the dynamics of the phenomenon under study. Typically, according to literature on differential equation parameter estimation, it is common to make literature reviews to collect information about an estimated range for the parameters of the differential equation model, and use this information to propose

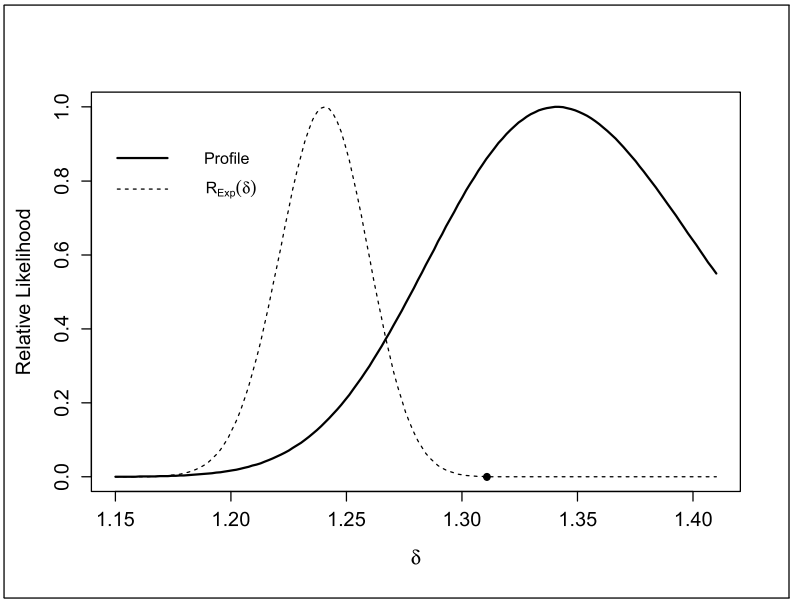

Figure 12: Relative likelihood and relative profile likelihood functions of δ, obtained from a model with an exponential growth. True δ value is marked with a black dot. Elaborated by the authors.

independent distributions for each of the model parameters (Pandey *et al.*, 2013; Camacho *et al.*, 2014; Funk *et al.*, 2016; Saldaña *et al.*, 2020).
It is also common to combine the approach described in the previous section and use the asymptotic distribution of the δ MLE, the parameter representing the growth rate of the exponential model, instead of a uniform distribution (Chowell *et al.*, 2007; Towers *et al.*, 2016).

Intuitively, setting aside the relationship between model parameters and assuming independence in order to construct $F(\theta)$, causes that implausible θ values become considered in the calculation of R_0, and that occurs given the intrinsic relationship between model parameters. In addition, misspecification of any of the parameter distributions may lead also to inaccurate inferences. To exemplify this practice and the kind of results it can yield, we will consider the SIR model in (2), but reparametrized in terms of $\theta = (\beta, \delta)$. To specify $\theta = (\beta, \delta)$ distribution, we assume independence and propose a uniform distribution $(1.2, 2)$ for β, in such a way that this interval includes true $\beta = 1.786946$ value. On the other hand, we use a normal distribution $N\left(\widehat{\delta}, \widehat{\sigma}^2\right)$ for δ, whose mean $\widehat{\delta} = 1.240564$ and variance $\widehat{\sigma}^2 = 1.240564$ are the MLE of δ and the Fisher information inverse, respectively; both obtained through the likelihood given in (9). Considering the proposed distributions, $M = 10000$ random samples for β and δ were simulated; in all cases it occurred that $\beta > \delta$, obtaining then 10000 computations of $R_0 = \beta/(\beta - \delta)$, which were used to build the histogram shown in Figure 13, where true R_0 value is marked with a black dot. It is interesting to observe that this frequency distribution is mainly concentrated into the interval $(2.604502, 3.480875)$, whose lower and upper limits are the 2.5 % and 97.5 % probability quantiles, respectively, and this interval does not include true $R_0 = 3.752587$ value.

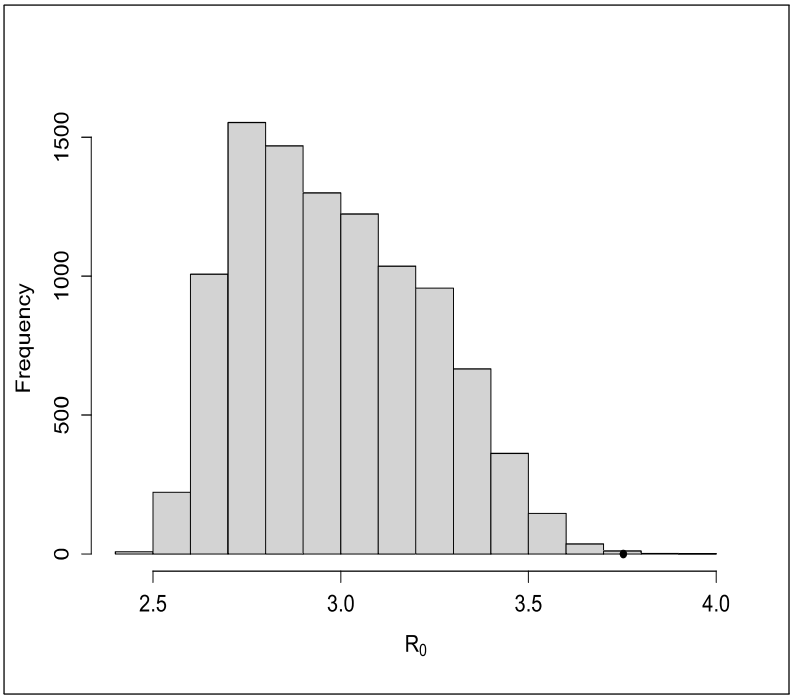

Figure 13: Histogram of R_0 values obtained from the simulation of 1000 random samples of β and δ. True R_0 value is marked with a black dot. Elaborated by the authors.

6. CONCLUSIONS

The shape of the likelihood function helps to detect practical parameter identifiability problems, revealing the intrinsic relationship of model parameters, in light of the observed sample. Moreover, the profile likelihood function allows to visualize and analyze the severity of the identifiability problem for a scalar parameter of interest, associated to the model under study, such as R_0, the basic reproductive number.
In the cases studied here, narrow likelihood surfaces with elongated or curved contours exhibit the problem of practical identifiability, whenever one of their tails does not stop decreasing and the surface plot appears to be truncated or cutoff within the parameter space. In general, much of the corresponding trajectory of the ridge or peak of the likelihood surface is nearly flat or constant, assigning to these points a plausibility as high as that of the MLE. Thus, fixing one parameter and maximizing the likelihood surface regarding some other one, is the way how the profile likelihood provides information about the ridge likelihood surface flatness.

In our simulation study, occurrence of the practical identifiability problem did not affect the coverage frequencies of the likelihood-confidence regions, that were always close to the nominal coverage probability.

On the other hand, the vast majority of practical parameter identifiability problems occurred when the observed sample of the SIR-Poisson model came from the initial and increasing part of the mean of the model, over time. In these cases, flat likelihoods can be interpreted in the sense that data do not provide sufficient information to discern about the shape of the average model behavior, over time. Thus, there are many parameter combinations yielding average behaviors that make the observed sample as likely as the average behavior associated with the MLE.

The methods discussed here for dealing with the practical identifiability problem should be cautiously used. Setting parameters can lead to false inference precision, also called overestimation or spurious accuracy. On the other hand, parameter estimation in systems of differential equations, with lack of consideration on the mathematical structure and the underlying parameters relationship involved in such a model, in light of the data, may lead to inappropriate inferences.

Perhaps, a strategy for obtaining prior information concerning the parameters of the model, in order to rescue their relationship, would be to consider general restrictions regarding the characteristics of the phenomenon under study, this could allow to generate feasible solutions of the system of differential equations, and from these to determine the values of the parameters that generated them, and use that information to propose a joint prior distribution. There is also the possibility of using this joint prior distribution in a Bayesian inference framework.

Extreme care should be taken when applying the estimation practices identified in this article, as they may lead to inappropriate conclusions about the basic reproductive number value, as well as the scenarios projections of an epidemic, among others. The latter becomes highly relevant when resulting estimates are used in public health decision making.
Perhaps the support and knowledge of experts in the field of the phenomenon being modeled can help to make better decisions to identify the relationship between the parameters of the model, and hence achieving the formulation of more appropriate hypotheses regarding the phenomenon under study.

References

Acuña-Zegarra, M. A., Díaz-Infante, S., Baca-Carrasco & D., Olmos-Liceaga, D. (2021). COVID-19 optimal vaccination policies: a modeling study on efficacy, natural and vaccine-induced immunity responses. *Mathematical Biosciences*, 6337, 108614.

Acuña-Zegarra, M. A., Santana-Cibrian, M. & Velasco-Hernández, J. X. (2020). Modeling behavioral change and COVID-19 containment in Mexico: A trade-off between lockdown and compliance. *Mathematical Biosciences*, 325, 108370.

Arino, J. & Van den Driessche, P. (2003). A multi-city epidemic model. *Mathematical Population Studies*, 10(**3**), 175-193.

Barndorff-Nielsen, O.E. & Cox, D.R. (1994). Inference and asymptotics. Chapman & Hall/CRC. Boca Raton.

Camacho, A., Kucharski, A. J., Funk, S., Breman, J., Piot, P. & Edmunds, W. J. (2014). Potential for large outbreaks of Ebola virus disease. *Epidemics*, 9, 70-78.

Capistrán, M.A., Christen, J. A. & Velasco-Hernández, J. X. (2012). Towards uncertainty quantification and inference in the stochastic SIR epidemic model. *Mathematical Biosciences*, 240(**2**), 250-259.

Chowell, G., Diaz-Dueñas, P., Miller, J. C., Alcazar-Velazco, A., Hyman, J. M., Fenimore, P. W.& Castillo-Chavez, C. (2007). Estimation of the reproduction number of dengue fever from spatial epidemic data. *Mathematical Biosciences*, 208(**2**), 571-589.

Chowell, G., Torre, C. A., Munayco-Escate, C., Suarez-Ognio, L., Lopez-Cruz, R., Hyman, J. M. & Castillo-Chavez, C. (2008). Spatial and temporal dynamics of dengue fever in Peru: 1994?2006. *Epidemiology & Infection*, 136(**12**), 1667–1677.

Chowell, G., Towers, S., Viboud, C., Fuentes, R., Sotomayor, V., Simonsen, L.,Miller, M. A., Lima, M., Villarroel, C., Chiu, M., Villarroel, J. E. & Olea, A. (2012). The influence of climatic conditions on the transmission dynamics of the 2009 A/H1N1 influenza pandemic in Chile. *BMC Infectious Diseases*, 12(**1**), 1-12.

Cole, D. J. (2020). Parameter redundancy and identifiability. Chapman & Hall/CRC. Boca Raton.

Cosner, C. (2015). Models for the effects of host movement in vector-borne disease systems. *Mathematical Biosciences*, 270, 192-197.

Funk, S., Kucharski, A. J., Camacho, A., Eggo, R. M., Yakob, L., Murray, L. M. & Edmunds, W. J. (2016). Comparative Analysis of Dengue and Zika Outbreaks Reveals Differences by Setting and Virus. *PLoS Neglected Tropical Diseases*, 10(**12**), e0005173.

Gábor, A., Villaverde, A. F.& Banga, J. R. (2017). Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems. *BMC Systems Biology*, 11(**1**), 1-16.

Ghosh, I., Sardar, T.& Chattopadhyay, J. (2017). A Mathematical Study to Control Visceral Leishmaniasis: An Application to South Sudan. *Bulletin of Mathematical Biology*, 79(**5**), 1100-1134.

Ghosh, I., Tiwari, P. K., Samanta, S., Elmojtaba, I. M., Al-Salti, N. & Chattopadhyay, J. (2018). A simple SI-type model for HIV/AIDS with media and self-imposed psychological fear. *Mathematical Biosciences*, 306, 160-169.

Guanghu, Z., Tao, L., Jianpeng, X., Bing, Z., Tie, S., Yonghui, Z., Lifeng, L., Zhiqiang, P., Aiping, D., Wenjun, M. & Yuantao, H. (2019). Effects of human mobility, temperature and mosquito control on the spatiotemporal transmission of dengue. *Science of The Total Environment*, 651, 969-978.

Gui-Quan, S., Jun-Hui, X., Sheng-He, H., Zhen, J,; Ming-Tao, L. & Liqun, L. (2017). Transmission dynamics of cholera: Mathematical modeling and control strategies. *Communications in Nonlinear Science and Numerical Simulation*, 45, 235-244.

Hendron, R. W. S. & Bonsall, M. B. (2016). The interplay of vaccination and vector control on small dengue networks. *Journal of Theoretical Biology*, 407, 349-361.

Kalbfleisch, J. G. (1985). Probability and Statistical Inference, Vol. 2. Springer-Verlag. New York.

Kao, Y. H. & Eisenberg, M. C. (2018). Practical unidentifiability of a simple vector-borne disease model: Implications for parameter estimation and intervention assessment. *Epidemics*, 25, 89-100.

Kermack, W. O. & McKendrick, A. G. (1927). Contribution to the mathematical theory of epidemics. *Proccedings of the Royal Society A*, 115(**772**), 700–721.

Khan, A., Hassan, M. & Imran, M. (2014). Estimating the basic reproduction number for single-strain dengue fever epidemics. *Infectious Diseases of Poverty*, 3(**1**), 1-17.

Kim, J. E., Lee, H., Lee, C. H. & Lee, S. (2017). Assessment of optimal strategies in a two-patch dengue transmission model with seasonality. *PLoS ONE*, 12(**3**), e0173673.

Lee, S. & Castillo-Chavez, C. (2015). The role of residence times in two-patch dengue transmission dynamics and optimal strategies. *Journal of Theoretical Biology*, 374, 152-164.

Lloyd-Smith, J. O. (2007). Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS ONE*, 2(**2**), e180.

Ma, J. (2020). Estimating epidemic exponential growth rate and basic reproduction number. *Infectious Disease Modelling*, 5, 129-141.

Marquis, A.D., Arnold, A., Dean-Bernhoft, C., Carlson, B.E. & Olufsen, M. S. (2018). Practical identifiability and uncertainty quantification of a pulsatile cardiovascular model. *Mathematical Biosciences*, 304, 9-24.

Mishra, A., Ambrosio, B., Gakkhar, S. & Aziz-Alaoui, M. A. (2018). A network model for control of dengue epidemic using sterile insect technique. *Mathematical Biosciences & Engineering*, 15(**2**), 441-460.

Mishra, A. & Gakkhar, S. (2018). Non-linear dynamics of two-patch model incorporating secondary dengue infection. *International Journal of Applied and Computational Mathematics*, 4(**19**), 1-22.

Murphy, S. A. & Van Der Vaart, A. W. (2000). On profile likelihood. *Journal of the American Statistical Association*, 95(**450**), 449-465.

Nguyen, V. K., Parra-Rojas, C. & Hernandez-Vargas, E. A. (2018). The 2017 plague outbreak in Madagascar: Data descriptions and epidemic modelling. *Epidemics*, 25, 20-25.

Núñez-López, M., Ramos, L. A. & Velasco-Hernández, J. X. (2021). Migration rate estimation in an epidemic network. *Applied Mathematical Modelling*, 89, 1949-1964.

Pandey, A., Mubayi, A. & Medlock, J. (2013). Comparing vector-host and SIR models for dengue transmission. *Mathematical Biosciences*, 246(**2**), 252-259.

Pawitan, Y. (2001). In All Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford University Press. New York.

Phaijoo, G. R. & Gurung, D. B. (2016). Mathematical study of dengue disease transmission in multi-patch environment. *Applied Mathematics*, 7(**14**), 1521-1533.

Qi, L., Xue, M., Cui, J.A., Wang, Q. & Wang, T. (2018). Schistosomiasis transmission model and its control in Anhui province. *Bulletin of Mathematical Biology*, 80(**9**), 2435-2451.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U. & Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(**15**), 1923-1929.

Rosenbaum, E. A., Pechen De D'angelo, A. M., Bergoc, R. M. & Venturino, A. (1999). Modelling acetylcholinesterase kinetics: The identifiability problem in parameter estimation. *Journal of Biological Systems*, 7(**01**), 95-111.

Saccomani, M. P. & Thomaseth, K. (2018). The union between structural and practical identifiability makes strength in reducing oncological model complexity: a case study. *Complexity*, 2018.

Saldaña, F., Flores-Arguedas, H., Camacho-Gutiérrez, J. A. & Barradas, I. (2020). Modeling the transmission dynamics and the impact of the control interventions for the COVID-19 epidemic outbreak. *Mathematical Biosciences and Engineering*, 17(**4**), 4165-4183.

Sasmal, S. K., Ghosh, I., Huppert, A. & Chattopadhyay, J. (2018). Modeling the spread of Zika virus in a stage-structured population: effect of sexual transmission. *Bulletin of Mathematical Biology*, 80(**11**), 3038-3067.

Serfling, R. J. (2002). Approximation Theorems of Mathematical Statistics. John Wiley & Sons. New York.

Sprott, D. A. (2000), Statistical inference in science. Springer-Verlag. New York.

Tocto-Erazo, M. R., Espíndola-Zepeda, J. A., Montoya-Laos, J. A., Acuña-Zegarra, M. A., Olmos-Liceaga, D., Reyes-Castro, P. A. & Figueroa-Preciado, G. (2020), Lockdown, relaxation, and acme period in COVID-19: A study of disease dynamics in Hermosillo, Sonora, Mexico. *PLoS ONE*, 15(**12**), e0242957.

Tocto-Erazo, M. R., Olmos-Liceaga, D. & Montoya, J. A. (2021). Effect of daily periodic human movement on dengue dynamics: The case of the 2010 outbreak in Hermosillo, Mexico. *Applied Mathematical Modelling*, 97, 559-567.

Towers, S., Brauer, F., Castillo-Chavez, C., Falconar, A.K., Mubayi, A. & Romero-Vivas, C. M. (2016). Estimate of the reproduction number of the 2015 Zika virus outbreak in Barranquilla, Colombia, and estimation of the relative role of sexual transmission. *Epidemics*, 17, 50-55.

Tuncer, N., Gulbudak, H., Cannataro, V. L. & Martcheva, M. (2016). Structural and practical identifiability issues of immuno-epidemiological vector-host models with application to rift valley fever. *Bulletin of Mathematical Biology*, 78(**9**), 1796-1827.

Tuncer, N., Mohanakumar, C., Swanson, S. & Martcheva, M. (2018). Efficacy of control measures in the control of Ebola, Liberia 2014-2015. *Journal of Biological Dynamics*, 12(**1**), 913-937.

Vinh, D. N., Ha, D.T.M., Hanh, N.T., Thwaites, G., Boni, M. F., Clapham, H. E. & Thuong, N. T. T. (2018). Modeling tuberculosis dynamics with the presence of hyper-susceptible individuals for Ho Chi Minh City from 1996 to 2015. *BMC Infectious Diseases*, 18(**1**), 1-13.

Xiao, Y. & Zou, X. (2014). Transmission dynamics for vector-borne diseases in a patchy environment. *Journal of Mathematical Biology*, 69(**1**), 113-146.

Zhan, C., Li, B.Y.S. & Yeung, L.F. (2015). Structural and practical identifiability analysis of S-system. *IET Systems Biology*, 9(**6**), 285-293.