

Modeling the uncertainty in epidemiological models through interval analysis considering actual data from two municipalities in Colombia affected by dengue

Diana Paola Lizarralde-Bejarano^{a,*}, Hayriye Gulbudak^b, Ralph Baker Kearfott^b,
María Eugenia Puerta-Yepes^a

^a Departamento de Ciencias Matemáticas, Escuela de Ciencias Aplicadas e Ingeniería, Universidad EAFIT, Carrera 49 No. 7 Sur-50, Medellín 050022, Colombia

^b Department of Mathematics, University of Louisiana at Lafayette, 104 E. University Circle, Lafayette, LA 70504-1010, USA

ARTICLE INFO

Article history:

Received 16 September 2021

Revised 6 July 2022

Accepted 11 July 2022

Available online 14 July 2022

Keywords:

Uncertainty

Interval analysis

Dengue transmission model

ABSTRACT

Epidemiological models have become powerful tools for studying and understanding the characteristics and impact of transmitted diseases in a population. However, these models usually require specifying several values of input parameters obtained from experimental data, characterized by high uncertainty levels due to biological variation. This situation is evident for models that simulate the transmission of vector-borne diseases such as dengue, our case study. Therefore, treating and modeling this uncertainty is essential to ensure the robustness of designed models. For this, we propose to model the uncertainty through interval analysis by representing the input parameters and initial conditions by real closed intervals in the forward problem. This approach has the advantage of making a minimal number of assumptions concerning uncertainties, unlike the traditional methods (probabilistic and fuzzy). To illustrate the performance of this methodology, we consider a coupled ODE system of seven state variables and nine parameters, representing the transmission of Dengue between host-vector populations. Additionally, to enhance the use of the numerical method utilized for solving the system, the uncertain quantities (parameters and initial conditions) are determined based on the results of (i) the sensitivity analysis of R_0 , (ii) the structural identifiability analysis of the model, (iii) the characteristics of the available information about mosquito population, and (iv) dengue incidence data in two municipalities in Colombia, Itagüi and Neiva, during the outbreaks in 2016. We believe that the methodology proposed here to select and incorporate uncertainty in epidemiological models through interval analysis is widely applicable to other phenomena and models in science and engineering.

© 2022 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

* Corresponding author.

E-mail addresses: paola.lizarralde@gmail.com (D.P. Lizarralde-Bejarano), hayriye.gulbudak@louisiana.edu (H. Gulbudak), ralph.kearfott@louisiana.edu, rbk@lusfiber.net (R.B. Kearfott), mpuerta@eafit.edu.co (M.E. Puerta-Yepes).

1. Introduction

Uncertainty is present in any process of measuring and obtaining information that is required to explain a real phenomenon. One source of uncertainty may be the lack of knowledge about the phenomenon studied, to determine which characteristics will be considered and which to ignore within the modeling. Other sources include the impossibility of obtaining measurements of some relevant factors, collecting information over long time periods, etc. [1].

In the case of vector-borne diseases such as west nile virus, malaria, zika, and dengue, there is uncertainty due to the inability to accurately and reliably measure transmission rates, vector populations, and the recovery rate in humans. Usually, these characteristics are included in the modeling process as parameters or initial conditions. This information is necessary to build more reliable models that allow us to understand the dynamics of this type of disease and thus be able to propose appropriate control strategies. However, in contrast to some sciences where it is possible to carry out several experiments to obtain information and test hypotheses, such experiments are often impossible, unethical or expensive when modeling the spread of infectious diseases in human populations [2]. For instance, even when we perform experimental assays with vector populations, these experiments involve imprecision, some degree of approximation, or uncertainty to various degrees, since it is not possible to include in the laboratory all external aspects involved in the development process. As an example, consider an experiment where three replicas with vector population are carried out. Suppose that each experiment starts with 100 eggs, and we want to measure the percentage of eggs hatching for this vector population. This measure may be stated in different ways as follows: (a) between 86 and 92 percent, (b) about 89 percent, or (c) has a mean value of 89 and a standard deviation of 2 percent and follows a normal distribution. Depending on the nature of imprecision, the analysis of the system can be conducted using interval analysis, fuzzy theory, or a probabilistic approach [3,4].

According to the type of information obtained from these experimental assays, we consider that, for transmission of vector-borne diseases, an efficient and reliable way to account for uncertainty is through interval analysis. Unlike applications based on probability and fuzzy theory, interval analysis does not attempt to infer an uncertainty structure of the model-output based on an uncertainty structure assumed for model-input.

Here, we determined dengue transmission as our case study based on the availability of information on new dengue cases, reported weekly, and the results of experimental assays with the local population of mosquitoes, allowing us to establish the initial ranges for human initial conditions and parameters of development stages of the vector. Because of the high epidemiological, social, and economic impact of dengue transmission throughout diverse tropical countries [5,6], it is relevant to evaluate different levels of uncertainty in the parameters and the initial conditions that have the primary role in the production of new outbreaks.

To solve the ODEs that handle interval-valued uncertainties in the parameters and initial conditions, we use VSPODE (Validating Solver for Parametric ODEs), although other software is available, such as AWA, COSY INFINITY, and VNODE. However, to our best knowledge, VSPODE is the only one capable of explicitly handling the uncertainty in the parameters without increasing the dimensionality of the ODEs [7–9]. Nevertheless, in this work, we observe some computational limitations to obtain verified solutions when considering all the parameters and initial conditions for the proposed model to be uncertain. For this reason, it is necessary to perform additional analyses of the model to reduce the dimensionality of the number of uncertain quantities to be included. Specifically, in this study, the uncertain quantities (parameters and initial conditions) are selected based on the sensitivity analysis results on R_0 , the local structural identifiability analysis on the model, and the characteristics of the available information. Finally, to illustrate the execution of this methodology, we consider a dengue transmission model of seven states.

2. Background

This section presents the definitions and concepts necessary to establish the notation used in the document so that the paper is self-contained.

2.1. Interval analysis

Interval-arithmetic is largely attributed to Ramon Moore in the 1960s; he developed it to rigorously account for rounding errors linked to mathematical calculations: The object on which this theory is constructed is the set of closed intervals in $\mathbb{R}$.

$$\mathcal{I} = \{X = [\underline{x}, \bar{x}] \mid \underline{x} \leq \bar{x} \wedge \underline{x}, \bar{x} \in \mathbb{R}\}.$$

This definition can be extended in a natural way to n -dimensional real interval vectors, $\mathcal{I}^n$ as $\mathbf{X} = [X_1, \dots, X_n]^T$, where $X_i = [\underline{x}_i, \bar{x}_i]$ and $n \geq 1$. An n -dimensional interval vector can be interpreted geometrically as an n -dimensional rectangle or box. For X and $Y \in \mathcal{I}$ is possible to define the basic arithmetic operations according to $X \circ Y = \{x \circ y \mid x \in X, y \in Y\}$, $\circ \in \{+, -, *, \div\}$, where we require $0 \notin Y$ for division¹. Additionally, addition and multiplication in $\mathcal{I}$ are associative and commutative, but only subdistributive: $(X * (Y + Z)) \subseteq (X * Y) + (X * Z)$. The interval $[0, 0]$ plays the role of neutral element in addition, while the

¹ Division is extended in various ways to remove this restriction; see, for example [10].

interval $[1,1]$ has the same role for multiplication. In general, for an arbitrary interval X , there exists neither an additive nor multiplicative inverse, that is, $X - X = 0$ and $X * 1/X = 1$ are not satisfied. Furthermore, the result of evaluating an interval expression always *contains* the set of all values of the expression generated by allowing the arguments of the expression to range over all values in the specified intervals. Ways of rewriting the expression that are equivalent in real arithmetic are not equivalent in interval arithmetic; for example, matrix multiplication is not associative.

An interval-valued function F can be defined as an extension of a real valued function f , if for degenerate intervals, that is, intervals of the form $[a, a]$, $F([x, x]) = f(x)$. Moreover, for a real function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ we can use interval arithmetic to bound the range of f over an interval X , replacing all the occurrences of x by X , to obtain $f(X) = \{f(x) \mid x \in X\} \subseteq f(X)$. A challenge in particular applications is to choose the form of the expressions or computation order for f to obtain the narrowest possible interval extension $f(X)$. For details, consult introductions to interval computations, such as Moore et al. [11] or Ackleh et al. [12, Section 1.3].

Finally, if we consider the metric

$$d_H(X, Y) = \max\{|\underline{x} - \underline{y}|, |\bar{x} - \bar{y}|\},$$

where $X = [\underline{x}, \bar{x}]$ and $Y = [\underline{y}, \bar{y}] \in \mathcal{I}$, it is possible to define all the elements of local analysis, such as limits, sequences, continuity, convergence, weak differentiability, and integrability over $\mathcal{I}$. With this we have all the tools to formulate differential equations in $\mathcal{I}$ [11].

The main drawbacks when using interval analysis are the *dependency problem* and the *wrapping effect*. The dependency problem occurs when there is more than one occurrence of the same variable in the expression for a function. The wrapping effect appears when, in intermediate computation steps, the result is not an interval or box, and it is necessary to enclose the result in an interval or box [13].

2.2. Taylor models

One way to handle the overestimation caused by the dependency problem and the wrapping effect is through the application of the Taylor model methods developed by Berz et al. [14–16], combining interval arithmetic with symbolic computations. To apply these methods, we have to consider Taylor expansions and an enclosure for the remainder.

Formally, consider a function $f: D \subset \mathbb{R}^s \rightarrow \mathbb{R}$ that is $(n+1)$ times continuously partially differentiable. A *Taylor model* for a function f that is $(n+1)$ times continuously partially differentiable is given by $T = (P, e) = P + e$ where P denotes the n -th order Taylor polynomial of f around the expansion point $x_0 \in D$ and e is a small bounding set for the remainder of this approximation:

$$f(x) - P(x - x_0) \in e, \quad \forall x \in D \text{ where } x_0 \in D.$$

In this paper, P and e are obtained by a truncated Taylor series.

3. Materials and methods

This section introduces the tools that we use to address the objective of this paper. To avoid making assumptions about the probability distribution that the parameters and initial conditions follow, we included uncertainty in the parameters and initial conditions through interval analysis since it is consistent with the type of available information for the phenomenon that we consider here (see Tables 3 and 4). In this manner, the modeler can calculate model solutions for parameter ranges instead of a single parameter value.

To compute mathematically and computationally guaranteed enclosures on the possible trajectories solutions of the systems given by (1) when the uncertain quantities (parameters and initial conditions) are given by real closed intervals, we use the VSPODE software (see Section 3.1). However, because of the wrapping effect, we catch excessive points in the solution enclosure at each integration step. Thus, this enclosure may eventually explode [17]. Some possible causes of this effect are the number of uncertain quantities considered as intervals, the width of these intervals, and the integration interval. Therefore, we design a strategy to reduce the number of uncertain quantities based on sensitivity and identifiability analysis results to address these difficulties. In particular, to reduce the number of uncertain parameters included in model (4), we performed a local sensitivity analysis on R_0 and a locally structural identifiability analysis of model (4). These analyses allow us to determine which parameters had more influence in the occurrence of new dengue cases, and by fixing the less important parameters, we reduced the parameter dimension space.

3.1. Solution procedure

In this work, as in Lin and Stadtherr [8], Enszer and Stadtherr [18], we consider systems of ordinary differential equations given by the following formulation:

$$\dot{x}(t) = f(x, \theta), \quad x(t_0) = x_0, \quad (1)$$

where $t \in [t_0, t_k]$, $t_k > t_0$, and $\theta \in \Theta$ is the m -dimensional vector of parameters. The variables x and x_0 are n -dimensional vectors of state variables and initial conditions, respectively. In addition, Θ and X_0 are interval vectors that represent the

Table 1

Elasticity expressions for R_0 . In the third column, we find how we rewrite these expressions to avoid the dependency problem in computation with intervals.

Elasticity	Expressions with the dependency problem	Expressions for avoiding the dependency problem
$\varepsilon_{\beta_m}^{R_0}$	0.5	0.5
$\varepsilon_{\beta_h}^{R_0}$	0.5	0.5
$\varepsilon_{\theta_m}^{R_0}$	$\frac{\mu_m}{2(\theta_m + \mu_m)}$	$\frac{1}{2\left(1 + \frac{\theta_m}{\mu_m}\right)}$
$\varepsilon_{\theta_h}^{R_0}$	$\frac{\mu_h}{2(\theta_h + \mu_h)}$	$\frac{1}{2\left(1 + \frac{\theta_h}{\mu_h}\right)}$
$\varepsilon_{\mu_m}^{R_0}$	$-\frac{\mu_m}{2(\theta_m + \mu_m)}\left(2 + \frac{\theta_m}{\mu_m}\right)$	$-\frac{1}{\left(1 + \frac{\theta_m}{\mu_m}\right)} - \frac{1}{2\left(\frac{\mu_m}{\theta_m} + 1\right)}$
$\varepsilon_{\gamma_h}^{R_0}$	$-\frac{\gamma_h}{2(\gamma_h + \mu_h)}$	$-\frac{1}{2\left(1 + \frac{\mu_h}{\gamma_h}\right)}$

enclosures of the uncertainties in the parameters and initial conditions, respectively. Also, we assume that $f: \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ is $k-1$ times continuously differentiable with respect to x and $q+1$ times continuously differentiable with respect to θ , where q is the Taylor model order.

To solve (1), we applied the method proposed in Lin and Stadtherr [8] which was implemented by the authors in the VSPODE software (Validated solutions of initial value problems for parametric ODEs). We briefly describe the method here; for more detailed information, we refer to Lin and Stadtherr [8].

First, consider a sequence of values $t_0 < t_1 < \dots < t_m$ with step size $h_j = t_{j+1} - t_j$ at the $(j+1)$ th integration step, $j = 0, 1, \dots, m-1$. A solution to the IVP

$$\dot{x}(t) = f(x, \theta), \quad x(t_j) = x_j$$

is given by

$$x(t; t_j, X_j, \Theta) = \{x(t; t_j, x_j, \theta) \mid x_j \in X_j, \theta \in \Theta\}.$$

In algorithms to solve (1), each integration step is divided into two stages. The first stage consists of validating the existence and uniqueness of the solution, while the second stage consists of computing a tighter enclosure.

3.1.1. First stage

The goal in the first stage is to find a step size $h_j = t_{j+1} - t_j > 0$ and an a priori enclosure $\tilde{X}_j$ of the solution such that a unique solution $x(t; t_j, x_j, \theta)$ is guaranteed to exist for all $t \in [t_j, t_{j+1}]$, all $x_j \in X_j$ and all $\theta \in \Theta$. For this purpose, the algorithm uses Interval Taylor Series (ITS) with respect to time. The uniqueness of the solution $x(t; t_j, x_j, \theta)$ is proved by using the Picard–Lindelöf operator and the Banach fixed-point theorem [9].

To compute the enclosure $\tilde{X}_j$, VSPODE uses high-order enclosure methods based on using many terms in the Taylor series. In this way, it is possible to determine $h_j = t_{j+1} - t_j$ and $\tilde{X}_j$ such that

$$\tilde{X}_j = \sum_{i=0}^{k-1} [0, h_j]^i F^{[i]}(X_j, \Theta) + [0, h_j]^k F^{[k]}(\tilde{X}_j^0, \Theta) \subseteq \tilde{X}_j^0. \quad (2)$$

where $(X_i)_j = F^{[i]}(X_j, \Theta)$ are the interval extensions of the Taylor coefficients² for $x_j \in X_j$ and $\theta \in \Theta$. One of the advantages of considering more terms in the Taylor series is that it is possible to consider larger step sizes, unlike first-order enclosure methods (constant enclosure methods).

3.1.2. Second stage

The goal in the second stage is to compute a tighter enclosure X_{j+1} such that $X_{j+1} \subseteq \tilde{X}_j$. In VSPODE this is done by using ITS to compute a Taylor model $T_{f[i]} = f^{[i]}(T_{x_j}, T_\theta)$ which depends on the initial conditions (x_0) and parameters (θ) . For the

² The j th Taylor coefficient evaluated at t_i is denoted by $(x_i)_j = \frac{x^{(j)}(t_i)}{j!}$

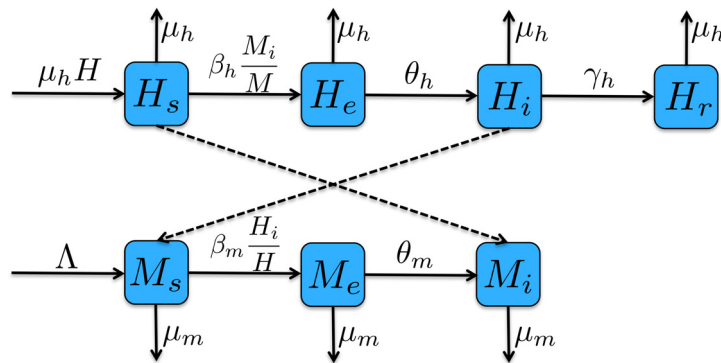


Fig. 1. Diagram of the dengue transmission model. Subscript s , e , i , and r indicate susceptible, exposed, infected, and recovered, respectively. H represents human and M represent female mosquitoes.

Taylor model computations, the interval initial states and parameters are represented by the Taylor models

$$\begin{aligned} x_{0_i} \in T_{x_{0_i}} &= m(X_{0_i}) + (x_0 - m(X_{0_i})) + [0, 0], \quad i = 1, \dots, n. \\ \theta_i \in T_{\theta_i} &= m(\Theta_i) + (\theta_i - m(\Theta_i)) + [0, 0], \quad i = 1, \dots, p. \end{aligned}$$

Then, it is possible to determine Taylor models $T_{f[i]}$ of the ITS coefficients $f^{[i]}(x_j, \theta)$ by using remainder differential algebra (RDA) [19] to compute $T_{f[i]} = f^{[i]}(T_{x_j}, T_{\theta})$. To reduce the overestimation produced due to interval dependency and the continuous growth of the remainder in each integration step, we use Taylor models $T_{f[i]}$ and the mean value theorem to compute the enclosure for each coefficient $f^{[i]}(x_j, \theta)$ for the ITS of x_{i+j} . Thus, we obtain the Taylor model $T_{x_{j+1}}$ for x_{j+1} in terms of the uncertain quantities θ and x_0 . Finally, to control the wrapping effect, the state enclosures are propagated using a new type of Taylor model. This new Taylor model consists of a polynomial and a remainder bound represented by an n -dimensional parallelepiped.

3.2. Mathematical model: dengue transmission

The model developed here is based on the one given in Lizarralde-Bejarano et al. [20], which can be interpreted as a reduction of the model introduced in Lizarralde-Bejarano et al. [21]. Here, the female mosquito population M is divided into three compartments: susceptible (M_s), exposed (M_e), and infected (M_i). Moreover, we allowed the size of the mosquito population to change over time. Also, we captured the behavior of the aquatic phase of the vector population (A) in one parameter, Λ , which is interpreted as the recruitment rate. To establish an appropriate biological range for this parameter, we define

$$\Lambda = f\gamma_m A^* \quad (3)$$

with $A^* = C(1 - 1/R_m)$ and $R_m = \rho f\gamma_m/(\mu_m(\gamma_m + \mu_a))$, where C represents the carrying capacity of the environment, γ_m represents the transition rate from the aquatic phase to the adult phase, ρ represents the effective per capita oviposition rate, f represents the fraction of female mosquitoes hatched from all eggs, and μ_a and μ_m represent the mortality rates of the aquatic and adult phases, respectively. Moreover, R_m is interpreted as the number of secondary females produced by only one female (the offspring), and A^* is the equilibrium value of the aquatic phase in which mosquitoes are present. By this definition of Λ , we take into account parameters that describe the development stages of the vector.

The size of human population H is considered constant with respect to the per capita mortality rate (μ_h), and is divided into four compartments: susceptible (H_s), exposed (H_e), infected (H_i), and recovered (H_r).

In both populations, the flow from the susceptible to exposed compartment depends on the proportion of infected in each population (H_i/H and M_i/M) and the transmission coefficients (β_h and β_m). Here, we assumed the transmission coefficients to be the product of the mosquito's biting rate and the transmission probabilities. Once extrinsic and intrinsic incubation periods are completed, the exposed mosquitoes and humans become infected at a rate of θ_m and θ_h , respectively. Finally, infected humans recover at a rate of γ_h , while mosquitoes remain infected for the rest of their lives [22]. Fig. 1 shows all transitions described above.

Based on the above assumptions, the flow of individuals from one compartment to another is described by the following system of differential equations:

$$\begin{aligned}
 \frac{dM_s}{dt} &= \Lambda - \beta_m \frac{H_i}{H} M_s - \mu_m M_s \\
 \frac{dM_e}{dt} &= \beta_m \frac{H_i}{H} M_s - (\theta_m + \mu_m) M_e \\
 \frac{dM_i}{dt} &= \theta_m M_e - \mu_m M_i \\
 \frac{dH_s}{dt} &= \mu_h H - \beta_h \frac{M_i}{M} H_s - \mu_h H_s \\
 \frac{dH_e}{dt} &= \beta_h \frac{M_i}{M} H_s - (\theta_h + \mu_h) H_e \\
 \frac{dH_i}{dt} &= \theta_h H_e - (\gamma_h + \mu_h) H_i \\
 \frac{dH_r}{dt} &= \gamma_h H_i - \mu_h H_r
 \end{aligned} \tag{4}$$

3.3. Basic reproductive number

The Basic Reproductive Number (R_0), is defined as the expected number of new cases of an infection produced by a typical infected individual in a wholly susceptible population over the full course of the infectious period [23]. In mathematical epidemiology, this number is one of the most important concepts, since it is a threshold parameter that helps us to determine if the disease dies out ($R_0 < 1$) or if the disease persists ($R_0 > 1$).

Our outcome of interest is to evaluate the most relevant parameters in the production of new dengue cases. To do so, we derived an expression for R_0 applying the *Next Generation Matrix* (NGM) [24] to model (4) around the disease-free equilibrium point, $M_s = M$, $H_s = H$, and $M_e = M_i = H_e = H_i = H_r = 0$.

$$R_0 = \sqrt{\frac{\beta_m \theta_m}{\mu_m (\theta_m + \mu_m)}} \sqrt{\frac{\beta_h \theta_h}{(\theta_h + \mu_h) (\gamma_h + \mu_h)}} \tag{5}$$

This expression gives us the geometric mean between the number of secondary infections of all sub-populations considered in each model, where the two components represent the number of infected mosquitoes and humans in the next generation, respectively.

To determine which parameters have more influence in the occurrence of new dengue cases, we calculate the *elasticity* of R_0 with respect to each parameter $\alpha \in \{\beta_m, \beta_h, \theta_m, \theta_h, \mu_m, \gamma_h\}$. The elasticity is given by

$$\varepsilon_{\alpha}^{R_0} = \frac{\partial R_0}{\partial \alpha} \frac{\alpha}{R_0} \approx \frac{\% \Delta R_0}{\% \Delta \alpha}. \tag{6}$$

The elasticities give the percentage change in R_0 in response to 1% increase in the parameter α . When $\varepsilon_{\alpha}^{R_0} > 0$, that means that R_0 increases with α ; when $\varepsilon_{\alpha}^{R_0} < 0$ that means that R_0 decreases when α increases [25]. For instance, in Table 1, the fact that $\varepsilon_{\beta_m}^{R_0} = 0.5$ means that 1% increase in β_m will produce 0.5% increase in R_0 . We summarize all the elasticity expressions for R_0 in Table 1:

3.4. Structural identifiability analysis

Structural identifiability analysis of a model can be interpreted as a way to determine if it is possible to uniquely recover the best model parameters if the data is assumed to be noise-free [26]. This analysis is only based on the model structure, and is independent of the accuracy of experimental data.

Formally, we say that a system (1) is *globally identifiable* if for any two parameter vectors θ_1 and θ_2 in the parameter space,

$$f(x(t), \theta_1) = f(x(t), \theta_2) \tag{7}$$

holds only if $\theta_1 = \theta_2$, where $f(x(t), \theta_1)$ and $f(x(t), \theta_2)$ are the solution trajectories for θ_1 and θ_2 , respectively. If the Eq. (7) is only satisfied for any θ_1 and θ_2 within an open neighborhood of some point θ^* in the parameter space, we say the system (1) is *locally identifiable* (definitions taken from Miao et al. [27]).

Different approaches have been proposed to test if a model is structurally identifiable; among these are the *direct test*, *differential algebra*, *Laplace transform*, *implicit function theorem*, the application of *Taylor series*, *profile likelihood*, and *output sensitivities*. These approaches are reviewed in more detail in Miao et al. [27], Chis et al. [28].

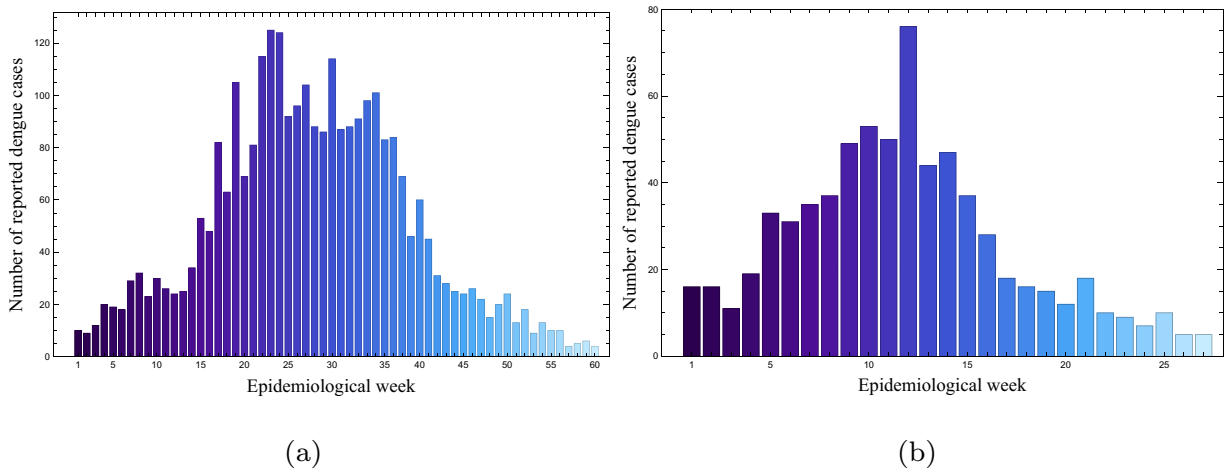


Fig. 2. (a) shows the reported dengue cases from epidemiological week 51 of 2015 to epidemiological week 6 of 2017 for Itagüí. (b) shows the reported dengue cases from epidemiological week 38 of 2016 to epidemiological week 9 of 2017 for Neiva.

Table 2

Parameters used to define the ranges of the recruitment rate (Λ) in (3) for Itagüí and Neiva, their biological description, their range of values per day and their range of values per week.

Param.	Meaning	Itagüí ranges per		Neiva ranges per	
		day	week	day	week
ρ	Effective per capita oviposition rate	[12,60]	[12,240]	[14,29]	[14,128]
C	Carrying capacity of the environment	[6400, 95, 000]	[6400, 95, 000]	[6400, 95, 000]	[6400, 95, 000]
γ_m	Transition rate from the aquatic to the adult phase	[0.11, 0.13]	[0.77, 0.88]	[0.11, 0.13]	[0.77, 0.88]
μ_a	Mortality rate in the aquatic phase	[0.001, 0.027]	[0.008, 0.19]	[0.015, 0.028]	[0.11, 0.19]
f	Fraction of female mosquitoes hatched from all eggs	[0.39, 0.51]	[0.39, 0.51]	[0.32, 0.45]	[0.32, 0.45]
μ_m	Mortality rate in the adult phase	[0.011, 0.016]	[0.008, 0.25]	[0.02, 0.027]	[0.14, 0.45]
Λ	Recruitment rate	[273, 6297]	[1779, 42, 612]	[223, 5550]	[1454, 37, 529]

3.5. Data and parameter values

We consider data from the 2016 dengue outbreaks in the municipalities of Itagüí (Antioquia, Colombia) and Neiva (Huila-Colombia). The outbreak in Itagüí lasted 60 epidemiological weeks,³ beginning in epidemiological week 51 of 2015 (with 10 reported cases) and ending in epidemiological week 6 of 2017 (with 4 reported cases). The total number of dengue cases reported during this period was 2915 (see Fig. 2(a)). The outbreak in Neiva lasted 27 epidemiological weeks, beginning in the epidemiological week 38 of 2016 (with 16 reported cases) and ending in the epidemiological week 12 of 2017 (with 5 reported cases). The total number of dengue cases reported during this period was 687 (see Fig. 2(b)). The information on the reported dengue cases was obtained from the *National Public Health Surveillance System* (SIVIGILA by its Spanish initials) (http://portalsivigila.ins.gov.co/sivigila/documentos/Docs_1.php).

To define the biological ranges for the transition rate from the aquatic phase to the adult phase (γ_m), the mortality rates of the aquatic and adult phases (μ_a and μ_m , respectively), the effective per capita oviposition rate (ρ), and the fraction of female mosquitoes hatched from all eggs (f), we consider the maximum and minimum values from the results of life tables created from experiments performed in the BCEI laboratory (Grupo de Biología y Control de Enfermedades infecciosas de la Universidad de Antioquia) between 2017 and 2019, with mosquito populations of *Aedes aegypti* of Itagüí and Neiva. For a deeper description of the experimental protocol, we refer the reader to Lizarralde-Bejarano et al. [21]. Then, we applied interval arithmetic [11] to compute the range of the recruitment rate (Λ) in (3) for each municipality. However, we extend intervals for mortality rate (μ_m) since these ranges were calculated under experimental conditions and did not consider external factors (as fumigation) that can increase it. The biological interpretations of these parameters and their ranges for each municipality are summarized in Table 2.

Ranges of values for intrinsic incubation period (θ_h), extrinsic incubation period (θ_m), and recovery rate (γ_h) were calculated considering the average duration of each period ($1/\theta_h$, $1/\theta_m$), and ($1/\gamma_h$) according to Organization et al. [30]. Ranges for parameters are summarized in Table 3.

³ An epidemiological week is simply a standardized method of counting weeks to allow for the comparison of data year after year. The first epidemiological week of the year ends, by definition, on the first Saturday of January, as long as it falls at least four days into the month. The epidemiological weeks change year by year [29].

Table 3

Parameters used in the simulations of model (4) for Itagüí and Neiva, their biological descriptions, their range of values per day and their range of values per week.

Param.	Meaning	Itagüí ranges per		Neiva ranges per	
		day	week	day	week
Λ	Recruitment rate	[273, 6297]	[1779, 42612]	[223, 5550]	[1454, 37529]
H	Size of human population	248, 036	248, 036	324, 466	324, 466
μ_m	Mortality rate in the adult phase	[0.011, 0.016]	[0.008, 0.25]	[0.02, 0.027]	[0.14, 0.45]
μ_h	Birth and death rate for human population	0.000032	0.00023	0.000015	0.00011
β_m	Transmission rate from human to mosquito	[0, 4]	[0, 4]	[0, 4]	[0, 4]
β_h	Transmission rate from mosquito to human	[0, 4]	[0, 4]	[0, 4]	[0, 4]
θ_m	Transition rate from exposed to infected mosquito	[0.08, 0.13]	[0.58, 0.88]	[0.08, 0.13]	[0.58, 0.88]
θ_h	Transition rate from exposed to infected human	[0.1, 0.25]	[0.7, 1.75]	[0.1, 0.25]	[0.7, 1.75]
γ_h	Recovery rate	[0.07, 0.25]	[0.5, 1.75]	[0.07, 0.25]	[0.5, 1.75]

Table 4

Initial conditions used in the simulations of model (4) for Neiva and Itagüí, their biological descriptions, and their range of values.

Initial condition	Meaning	Itagüí	Neiva
$M_s(0)$	For susceptible mosquitoes	[0, 5, 000, 000]	[0, 5, 000, 000]
$M_e(0)$	For exposed mosquitoes	[0, 200]	[0, 200]
$M_i(0)$	For infectious mosquitoes	[0, 200]	[0, 200]
$H_s(0)$	For susceptible humans	[198, 429, 247, 912]	[259, 573, 324, 294]
$H_e(0)$	For exposed humans	[21, 84]	[27, 108]
$H_i(0)$	For infectious humans	[10, 40]	[16, 64]
$H_r(0)$	For recovered humans	[0, 49, 576]	[0, 64, 850]

The information about the size of human population for each municipality was taken from the National Administrative Department of Statistics (DANE by its Spanish initials). We establish the following initial conditions for the model defined in (4) for the total human population; we used sizes of 248, 036 and 324, 466 as recorded for the urban area of Itagüí and Neiva in 2016, respectively. The range for the susceptible human population at the epidemic's beginning was between 80% and 100% of the total population. The initial condition for infected human populations was defined as the number of cases reported at the epidemic's beginning. At the same time, the lower bound for the initial condition for the exposed human population was defined as the totality of reported cases in the second and third epidemiological weeks by taking into account the time between the onset of symptoms of dengue disease and the infected mosquito bite. The upper bound for these initial conditions was established by considering the under-reporting, which can affect up to 75% of the total number of cases occurring anywhere dengue transmission occurs [31].

For the mosquitoes population, we assumed up to 20 susceptible mosquitoes per person for Itagüí and 15 susceptible mosquitoes per person for Neiva based on entomological findings for these municipalities [32]. The ranges for exposed and infected mosquitoes were considered higher than the total number of exposed and infected humans at the epidemic's beginning because of the unavailability of more specific information. Ranges for the initial conditions of Itagüí and Neiva are summarized in Table 4.

4. Results

Dengue transmission modeling is subject to multiple sources of uncertainty. Here we focus on: (i) including uncertain parameters and uncertain initial conditions to obtain robust solutions that consider the possible errors in the information measurement and digitization processes, and (ii) managing the numerical errors (truncation errors and rounding errors) associated with computing the numerical solutions of system (4).

4.1. Local sensitivity analysis of R_0

The elasticity ranges for each municipality are summarized in Table 5. These ranges were computed using interval arithmetic, as explained in Moore et al. [11]. Note that $\varepsilon_{\beta_m}^{R_0}$ and $\varepsilon_{\beta_h}^{R_0}$ always are equal to 0.5, while the other elasticities depend on parameter values. It is worth mentioning that these ranges enclose all the values produced by any combination of parameter values. For example, the ranges obtained for both expressions of $\varepsilon_{\mu_m}^{R_0}$ for Itagüí contain the values of any combination of $\theta_m \in [0.08, 0.88]$ and $\mu_m \in [0.008, 0.25]$.

To illustrate how the dependency problem can affect the width of the function range we are computing, we use both expressions presented in Table 1. For instance, the difference between the ranges calculated for $\varepsilon_{\mu_m}^{R_0}$ and $\varepsilon_{\gamma_h}^{R_0}$ using both expressions are rather noticeable and show the necessity of rewriting these to reduce the number of occurrences of the

Table 5
Elasticity ranges per week of R_0 for Itagüí and Neiva.

	Itagüí ranges		Neiva ranges	
	with dependency	avoiding dependency	with dependency	avoiding dependency
$\varepsilon_{\beta_m}^{R_0}$	0.5	0.5	0.5	0.5
$\varepsilon_{\beta_h}^{R_0}$	0.5	0.5	0.5	0.5
$\varepsilon_{\theta_h}^{R_0}$	[0.0035, 0.2211]	[0.0045, 0.151]	[0.0526, 0.3125]	[0.0686, 0.2184]
$\varepsilon_{\theta_h}^{R_0}$	$[6.5 \times 10^{-5}, 1.7 \times 10^{-4}]$	$[6.5 \times 10^{-5}, 1.7 \times 10^{-4}]$	$[3.1 \times 10^{-5}, 7.9 \times 10^{-5}]$	$[3.1 \times 10^{-5}, 7.9 \times 10^{-5}]$
$\varepsilon_{\mu_m}^{R_0}$	[-24.762, -0.015]	[-0.79, -0.36]	[-2.589, -0.173]	[-0.86, -0.42]
$\varepsilon_{\gamma_h}^{R_0}$	[-1.75, -0.143]	[-0.5, -0.49]	[-1.75, -0.143]	[-0.5, -0.49]

parameters μ_m and γ_h , respectively. However, for $\varepsilon_{\theta_h}^{R_0}$, we obtain the same range. This is because the parameter μ_h is constant, and θ_h only occurs once.

Finally, from the ranges shown in Table 5, we conclude that parameters β_m , β_h , μ_m , and γ_h are the most influential in the occurrence of secondary dengue cases.

4.2. Structural identifiability analysis

To carry out the local identifiability analysis, we use the *Identifiability Analysis* package in the *Mathematica* software provided by the authors of Karlsson et al. [33]. This implementation is based on a probabilistic numerical method of computing the rank of the identifiability (Jacobian) matrix, where the matrix parameters and initial state variables are assigned random integers.

For model (4) we found that the parameter Λ and the initial conditions for susceptible, exposed and infected mosquitoes are not locally identifiable from the weekly number of reported dengue cases when we fix: (i) the values of human mortality rate (μ_h), (ii) the size of human population (H), (iii) the initial condition of infected humans ($H_i(0)$ as the lower bound), and (4) the initial condition of recovered humans as $H_r(0) = H - H_s(0) - H_e(0) - H_i(0)$. However, for model (4) to be locally structurally identifiable, we need to assume or obtain information about at least one of the unidentifiable parameters. This information corresponds to the minimum necessary for the identifiability matrix to be of full rank.

4.3. Numerical simulations

Henceforth we say that the enclosures obtained in this study are 100% reliable up to the correctness of the input intervals. The simulations aimed to assess the uncertainty levels that can be considered in the parameters and initial conditions of model (4) to obtain verified solutions for Itagüí and Neiva. To do so, we applied VSPODE, with its default ITS (Interval Taylor Series) order $k = 17$ and a default Taylor model order $q = 5$. We defined the interval of integration for each municipality according to the duration of the outbreak. For Itagüí the interval of integration was from $t = 0$ to $t = 60$ epidemiological weeks, while the interval for Neiva was from $t = 0$ to $t = 27$ epidemiological weeks. For simulations, we chose some parameter values within the biological ranges defined in Tables 3 and 4. These parameter ranges were not estimated, but we selected them such that the infected human trajectories could be contrasted with the reported dengue cases.

According to Table 5, for model (4), the occurrence of new dengue cases was more sensitive to the transmission rate from human to mosquito (β_m), the transmission rate from mosquito to human (β_h), the recovery rate in humans (γ_h), and the mortality rate in mosquitoes (μ_m). Fig. 3 shows the enclosures computed using VSPODE to solve the system (4) for the infected humans considering uncertainty in these parameters. The trajectory associated with any combination of parameter values β_m , β_h , μ_m , and γ_h in their respective intervals always stays within the found enclosure. Moreover, it does not require additional computations unlike Monte Carlo simulations, where multiple trials are necessary to define upper and lower bounds, and the solutions are not 100% reliable, as was shown by the authors in Enszer and Stadtherr [18] for other epidemiological models. The R_0 range for the parameters values of Fig. 3 was between 0.75 and 0.79 for Itagüí, and between 0.58 and 0.68 for Neiva.

Fig. 4 shows the guaranteed enclosures for the possible trajectories for susceptible and infected humans per week considering uncertainty in the non-identifiable parameter, Λ , while, Fig. 5 shows the guaranteed enclosures for the possible trajectories for infected humans per week considering uncertainty in the non-identifiable initial conditions ($M_s(0)$, $M_e(0)$, $M_i(0)$). Note that in these figures, both the recruitment rate values and the mosquito initial condition values were bigger for Neiva than for Itagüí. These values were consistent with each municipality's climatic characteristics since, in Neiva, there is a more significant presence of the *Aedes aegypti* population [32]. In Fig. 4, we observe that at the end of each municipality's integration steps, the enclosures become significantly wider. This occurs when the solution blows up (i.e. that at some point of the integration process the solution is not bounded due to the wrapping effect). For Itagüí VSPODE breaks down at $t = 56$, and for Neiva at $t = 25$. To avoid that the solution blows up, we can divide the parameter intervals into equal-sized sub-boxes and then use VSPODE to determine the solution for each sub-box. The final solution enclosure is then the union of all

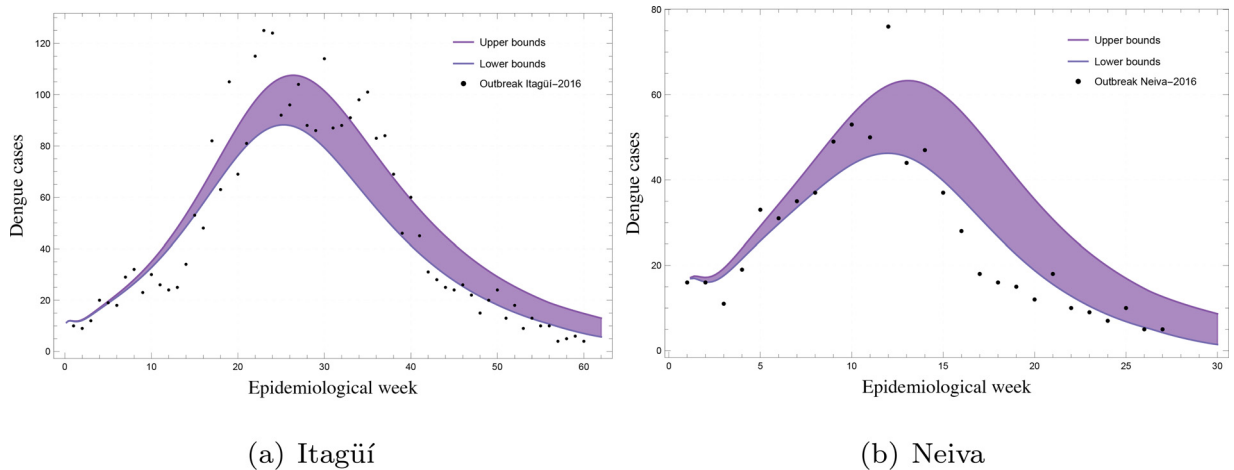


Fig. 3. The figure shows in purple the enclosure obtained with VSPODE for the infected humans for Itagüí and Neiva when β_m , β_h , μ_m , and γ_h are given by intervals. These enclosures contain all the trajectories produced by any parameter combination in the given ranges for each municipality. The black points represent the number of reported dengue cases per epidemiological week. We overlaid the data on the enclosures to show the possibility of simulating them when the parameters are intervals. The parameter and initial condition values used for: **(a)** $\beta_m = [0.12, 0.125]$, $\beta_h = [2.5, 2.55]$, $\mu_m = [0.217, 0.22]$, $\gamma_h = [1.74, 1.75]$, $M_s(0) = 1,800,000$, $M_e(0) = 50$, $M_i(0) = 40$, $H_s(0) = 223,000$, $H_e(0) = 21$, $H_i(0) = 10$, $H_r(0) = 25,005$, $H = 248,036$, $\mu_h = 0.00023$, $\theta_m = 0.6$, $\theta_h = 1.3$, and $\Lambda = 2000$. **(b)** $\beta_m = [0.14, 0.16]$, $\beta_h = [2.5, 2.7]$, $\mu_m = [0.39, 0.4]$, $\gamma_h = [1.6, 1.7]$, $M_s(0) = 3,000,000$, $M_e(0) = 100$, $M_i(0) = 50$, $H_s(0) = 315,952$, $H_e(0) = 27$, $H_i(0) = 16$, $H_r(0) = 8471$, $H = 324,466$, $\mu_h = 0.00011$, $\theta_m = 0.8$, $\theta_h = 1.3$, $\Lambda = 15,000$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the enclosures resulting from each parameter sub-box. We illustrate this strategy in more detail in [Appendix A](#) where we consider uncertainty in one parameter and two initial conditions.

Lastly, [Fig. 6](#) shows mathematically and computationally guaranteed upper and lower bounds on the possible trajectories of infected humans for Itagüí and Neiva when we consider human initial conditions as intervals. This simulation makes sense since the Colombian surveillance system (SIVIGILA) collects incidence data based on the notification of patients treated medically. For this reason, the official information does not include asymptomatic cases and other cases of infection that did not require a visit to health care facilities. Thus, it is impossible to determine the number of susceptible, exposed, and infected humans in a specific region accurately.

In general, for all the simulations carried out, broader intervals were considered for the parameters and initial conditions of Neiva since the integration time was shorter than for Itagüí. To consider wider intervals for Itagüí, we can proceed in the same way that we have mentioned before, to prevent that the solution explodes (see [Appendix A](#)).

5. Discussion

This study presents a strategy to include uncertainty in modeling based on ODEs, through the application of interval arithmetic, structural identifiability analysis, and local sensitivity analysis. To illustrate the performance of these analyses jointly, we considered as an example a model of seven state variables and nine model parameters that simulates the transmission of dengue diseases (see [Eq. \(4\)](#)).

To define initial intervals for parameters and initial conditions for model [\(4\)](#) with biological meaning, we have: (i) results from experimental assays with local mosquito populations for each municipality; (ii) the average time for transition from exposed to infected (mosquitoes and humans); (iii) the average time of recovery rate in humans; (iv) the official information of new dengue cases per week; and (v) the size of human population for each municipality (see [Tables 3 and 4](#)).

For model [\(4\)](#), it was possible to obtain the trajectories with VSPODE since its vector field is sufficiently continuously differentiable with respect to the state variables and the parameters. Nevertheless, we can just say that the enclosures obtained with VSPODE are 100% reliable up to the correctness of the input intervals since these were constructed from measurements. A possible way to address this limitation is by considering wider intervals that do not require so much precision in the initial assessment. Additionally, it was not possible to consider uncertainty in a larger set of parameters and initial conditions simultaneously due to overestimation caused by the dependency problem, the wrapping effect, and the curse of dimensionality [\[9\]](#).

Other methodologies such as Monte Carlo could offer advantages with respect to the number of parameters that can be considered uncertain. However, they do not guarantee the solutions, as in the case of interval analysis, since there may be regions of the parameter space not being evaluated in the sampling process. A more detailed comparison was made in Enszer and Stadtherr [\[18\]](#), where those authors found that the enclosures obtained by using VSPODE are just as narrow as those obtained by the Monte Carlo analysis. In addition, it was shown that the computational time was significantly lower in the case of interval analysis than in the case of Monte Carlo simulations for the cases considered in that work.

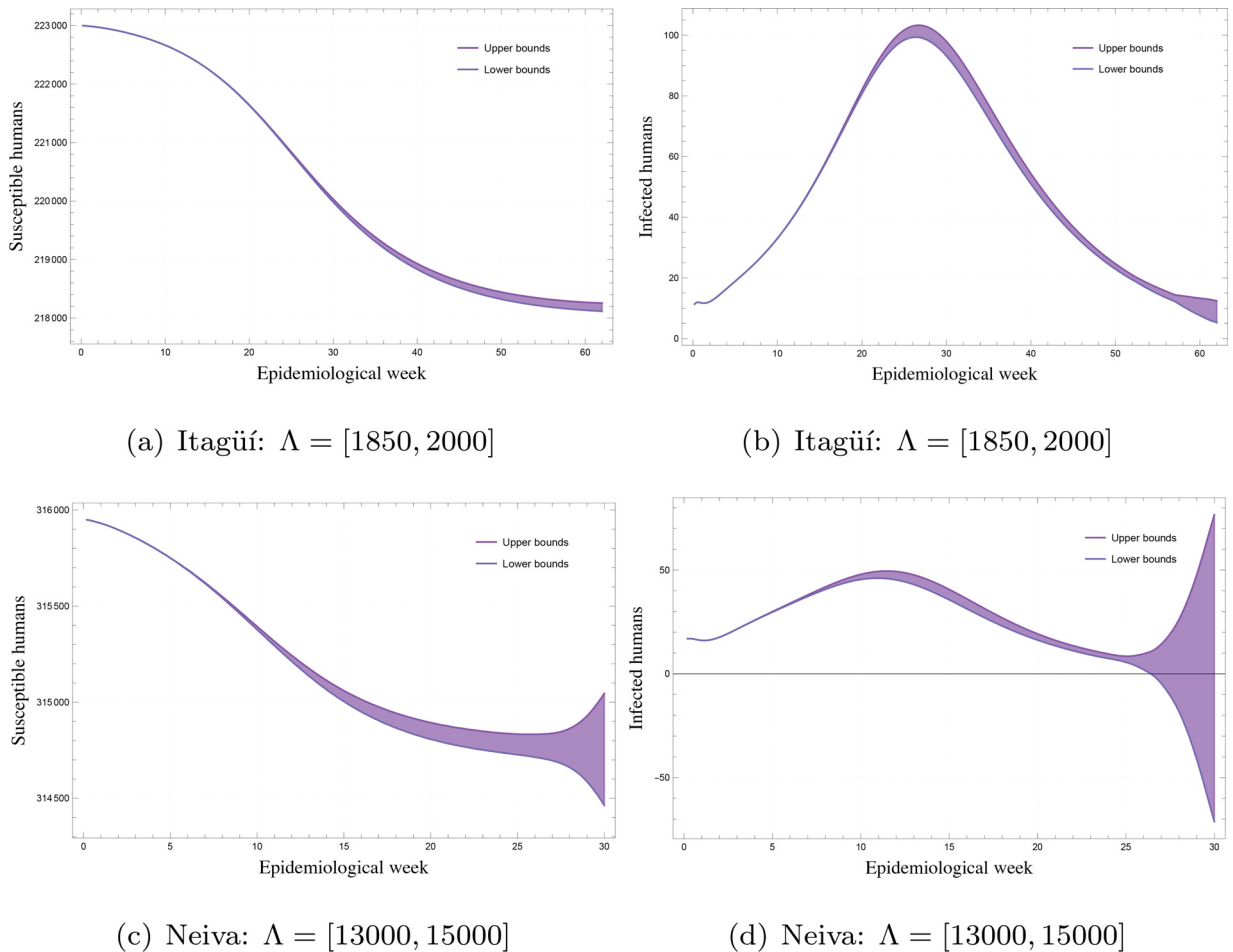


Fig. 4. The figure shows the enclosures obtained with VSPODE for susceptible and infected humans for Itagüí and Neiva when the recruitment rate (Λ) is given by an interval. These enclosures contain all the trajectories for any value of Λ in the given ranges. The parameter and initial condition values used are: For Itagüí, figures (a) and (b) $M_s(0) = 1,800,000$, $M_e(0) = 50$, $M_i(0) = 40$, $H_s(0) = 223,000$, $H_e(0) = 21$, $H_i(0) = 10$, $H_r(0) = 25,005$, $H = 248,036$, $\mu_h = 0.00023$, $\theta_m = 0.6$, $\theta_h = 1.3$, $\mu_m = 0.22$, $\beta_m = 0.12$, $\beta_h = 2.5$, $\gamma_h = 1.75$, $M_s(0) = 3,000,000$, $M_e(0) = 100$, $M_i(0) = 50$, $H_s(0) = 315,952$, $H_e(0) = 27$, $H_i(0) = 16$, and $H_r(0) = 8471$. For Neiva, figures (c) and (d) $M_s(0) = 3,000,000$, $M_e(0) = 100$, $M_i(0) = 50$, $H_s(0) = 315,952$, $H_e(0) = 27$, $H_i(0) = 16$, $H_r(0) = 8471$, $H = 324,466$, $\mu_h = 0.00011$, $\theta_m = 0.8$, $\theta_h = 1.3$, $\mu_m = 0.4$, $\beta_h = 2.5$, $\beta_m = 0.14$, and $\gamma_h = 1.7$.

Additionally, it was shown that the enclosure of the state variables obtained with the VSPODE at the equilibrium point contains all the values obtained analytically. At the same time, some solutions are lost with the Montecarlo simulations.

Here, it was necessary to examine alternative strategies to select the uncertain quantities that should be considered. In this way, we can reduce the problem's dimension and successfully apply interval methods to find guaranteed bounds for model solutions. In particular, for our case study, we performed a local sensitivity analysis on R_0 and a locally structural identifiability analysis on the model to select the uncertain parameters and uncertain initial conditions to include. From the biological point of view, it is relevant to consider uncertainty in parameters measured under laboratory conditions, since these results do not always correspond to the vector's life in the wild. Additionally, the local structurally identifiability analysis results suggest that it is necessary to obtain more information about the mosquito population for model (4) to be structurally identifiable. However, collecting this information for long periods can be expensive and unreliable. This suggests that considering uncertainty in the initial conditions of mosquito population through the definition of lower and upper bounds is an excellent way to determine how an outbreak would be in the presence of larger populations (see Fig. 5). Under this assumption, it is possible to define the frequency, intensity, and duration of more efficient and robust control strategies. This result is significant since, at present, the only way to mitigate dengue outbreaks efficiently is by controlling the vector population [34].

These findings highlight the potential usefulness of verified methods in mathematical epidemiology as an alternative to manage uncertainty in actual phenomenon modeling. However, it is worth noting that although VSPODE attempts to handle overestimation at each integration step, further research could explore to create verified solvers for specific model characteristics. Thus, it is possible to exploit those characteristics to design better strategies to reduce the dependency problem and

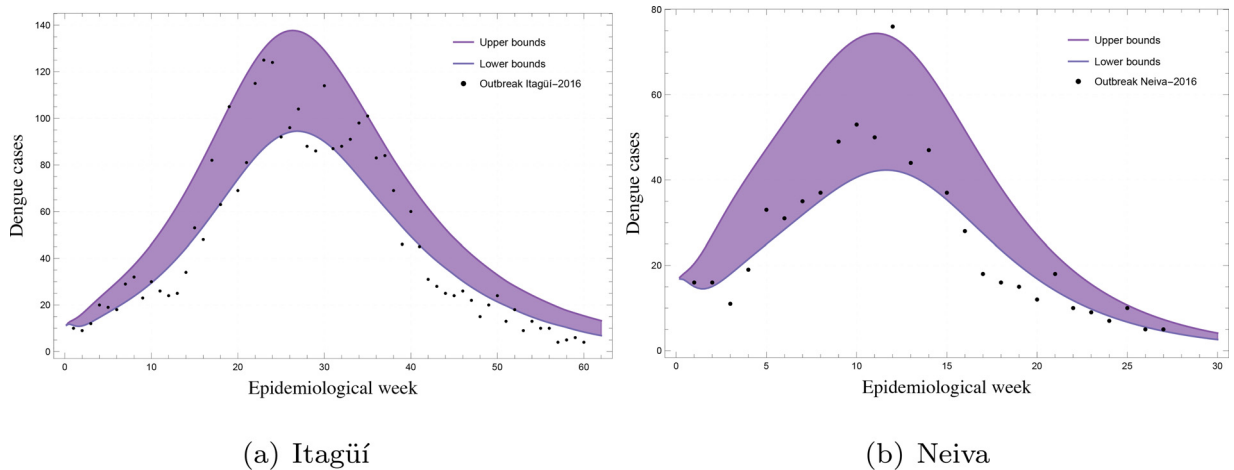


Fig. 5. The figure shows in purple the enclosure obtained with VSPODE for the infected humans for Itagüi and Neiva when the mosquito initial conditions ($M_s(0)$, $M_e(0)$, $M_i(0)$) are given by intervals. These enclosures contain all the trajectories produced by any combination of these initial conditions in the given ranges for each municipality. The black points represent the number of reported dengue cases per epidemiological week. We overlaid the data on the enclosures to show the possibility of simulating them when the mosquito initial conditions are intervals. Parameter and initial condition values used for: **(a)** $M_s(0) = [1, 800, 000, 2, 000, 000]$, $M_e(0) = [50, 70]$, $M_i(0) = [40, 60]$, $H = 248, 036$, $\mu_h = 0.00023$, $\theta_m = 0.6$, $\theta_h = 1.3$, $\Lambda = 2000$, $\gamma_h = 1.75$, $\mu_m = 0.22$, $\beta_m = 0.12$, $\beta_h = 2.5$, $H_s(0) = 223, 000$, $H_e(0) = 21$, $H_i(0) = 10$, and $H_r(0) = 25, 005$. **(b)** $M_s(0) = [3, 000, 000, 3, 500, 000]$, $M_e(0) = [100, 150]$, $M_i(0) = [50, 100]$, $H = 324, 466$, $\mu_h = 0.00011$, $\theta_m = 0.8$, $\theta_h = 1.3$, $\Lambda = 15, 000$, $\gamma_h = 1.7$, $\mu_m = 0.4$, $\beta_h = 2.5$, $\beta_m = 0.14$, $H_s(0) = 315, 952$, $H_e(0) = 27$, $H_i(0) = 16$, and $H_r(0) = 8471$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

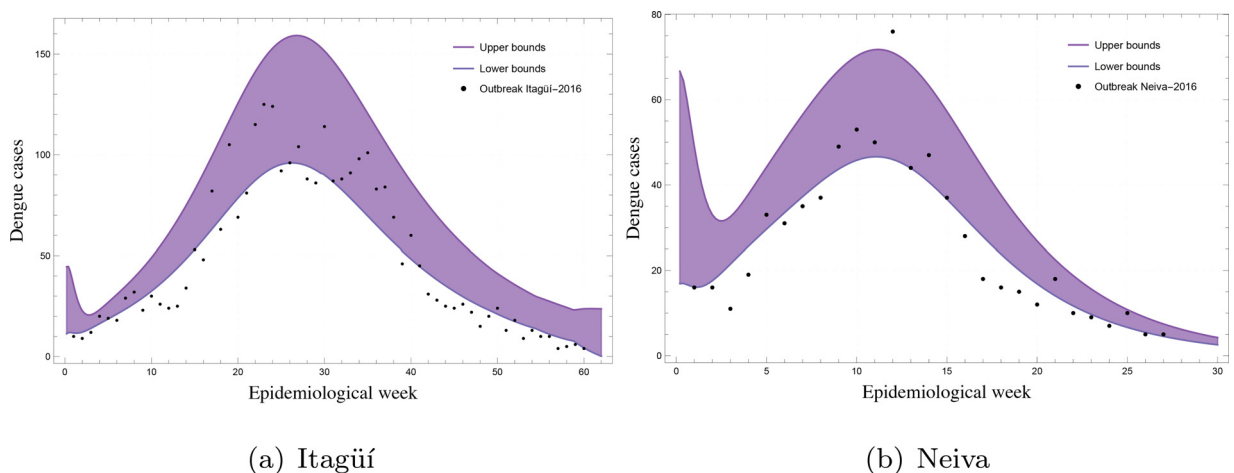


Fig. 6. The figure shows in purple the enclosure obtained with VSPODE for the infected humans for Itagüi and Neiva when the human initial conditions are given by intervals. These enclosures contain all the trajectories produced by any combination of these initial conditions in the given ranges for each municipality. The black points represent the number of reported dengue cases per epidemiological week. We overlaid the data on the enclosures to show the possibility of simulating them when the parameters are intervals. Parameter and initial condition values used for: **(a)** $H_s(0) = [223, 000, 235, 634]$, $H_e(0) = [21, 84]$, $H_i(0) = [10, 40]$, $H_r(0) = [12, 340, 25, 005]$, $M_s(0) = 1, 800, 000$, $M_e(0) = 50$, $M_i(0) = 40$, $H = 248, 036$, $\mu_h = 0.00023$, $\theta_m = 0.6$, $\theta_h = 1.3$, $\Lambda = 2000$, $\gamma_h = 1.75$, $\mu_m = 0.22$, $\beta_m = 0.12$, and $\beta_h = 2.5$. **(b)** $H_s(0) = [315, 952, 324, 294]$, $H_e(0) = [27, 108]$, $H_i(0) = [16, 64]$, $H_r(0) = [0, 8471]$, $M_s(0) = 3, 000, 000$, $M_e(0) = 100$, $M_i(0) = 50$, $H = 324, 466$, $\mu_h = 0.00011$, $\theta_m = 0.8$, $\theta_h = 1.3$, $\Lambda = 15, 000$, $\gamma_h = 1.7$, $\mu_m = 0.4$, $\beta_h = 2.5$, and $\beta_m = 0.14$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the wrapping effect. In this way, it may be possible to include more uncertain quantities (parameters and initial conditions) at the same time for more long integration periods.

A final aspect that should be mentioned is that the strategy presented here to select and subsequently incorporate uncertainty can be extrapolated to models that simulate other phenomena of different application areas, and models that can incorporate uncertainty in other ways.

Summarizing, this paper focuses on solving the forward problem by including uncertain quantities (parameters and initial conditions) via interval analysis. The reported dengue cases were included in Figs. 3, 5, and 6 to motivate the parameter estimation problem and show the potential of interval analysis for solving it. We consider this work a necessary step before addressing the parameter estimation problem since we observed limitations in the number of uncertain quantities that can be included in the model and the width of these intervals to obtain verified solutions.

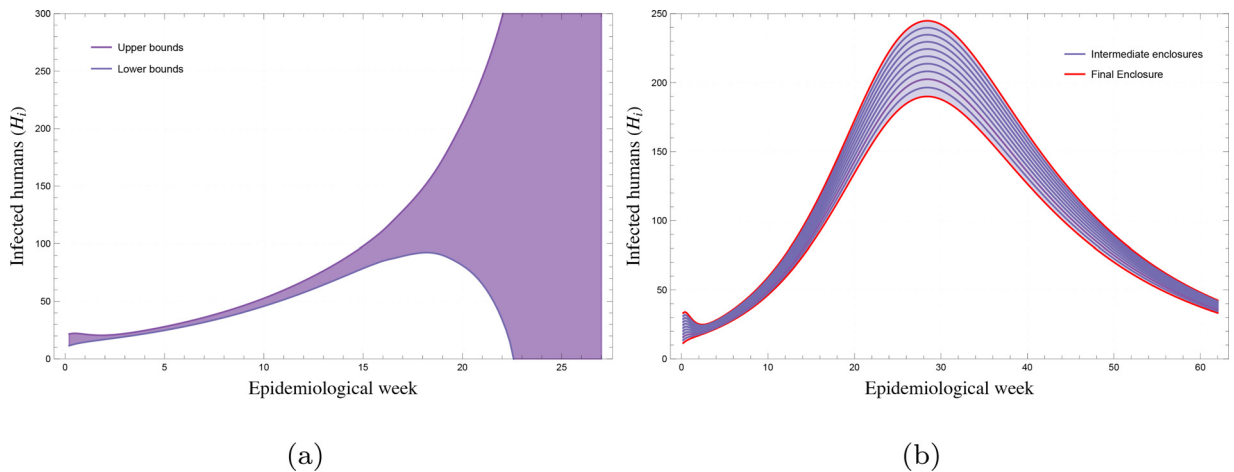


Fig. 7. (a) Shows how the trajectory for infected humans (H_i) starts blowing up at $t = 15$ when $H_e(0) = [21, 37]$, $H_i(0) = [10, 20]$, and $\theta_m = [0.58, 0.88]$. (b) Shows in purple the enclosures obtained when we split $H_e(0) = [21, 61]$, and $H_i(0) = [10, 30]$ into 10 sub-boxes. In red, it is shown the union of all the enclosures resulting from each sub-box. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Acknowledgments

Funding for the research was provided jointly by Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología Francisco José de Caldas, COLCIENCIAS, Universidad EAFIT and Universidad de Antioquia (Project 111572553478). We give special thanks to the authors of VSPODE software [8] for sharing it and answering some questions. The authors thank two anonymous reviewers for their helpful comments and feedback on the manuscript. Hayriye Gulbudak was supported by NSF grant (DMS-1951759) and Simons Foundation/SFARI(638193).

Appendix A. Considering uncertainty in one parameter and two initial conditions at the same time

To illustrate how we can proceed when the model solution blows up, we consider as uncertain quantities θ_m , $H_e(0)$, and $H_i(0)$ (transition rate from an exposed to an infected mosquito, the initial conditions of exposed human and infected human, respectively). Here, we applied VSPODE, with ITS (Interval Taylor Series) order $k = 12$ and a Taylor model order $q = 9$. For Fig 7(a) and (b) the other parameters and initial conditions are $M_s(0) = 1,800,000$, $M_e(0) = 50$, $M_i(0) = 40$, $H_s(0) = 223,000$, $H_r(0) = 25,005$, $H = 248,036$, $\mu_h = 0.00023$, $\beta_m = 0.12$, $\beta_h = 2.5$, $\mu_m = 0.22$, $\gamma_h = 1.75$, $\Lambda = 2000$, and $\theta_h = 1.3$. Fig. 7(a) shows that VSPODE breaks down at $t = 15$, due to the rapid growth of the enclosure. On the other hand, to obtain guaranteed enclosures, when $\theta_m = [0.58, 0.88]$, $H_e(0) = [21, 61]$, and $H_i(0) = [10, 30]$, we split the intervals of $H_e(0)$ and $H_i(0)$ into 10 equal-sized sub-boxes. Then, we use VSPODE to determine the solution for each sub-box. The final solution enclosure is the union of all the enclosures resulting from each sub-box. Fig. 7(b) shows these enclosures.

References

- [1] W. Maziak, Is uncertainty in complex disease epidemiology resolvable? *Emerg. Themes Epidemiol.* 12 (1) (2015) 7, doi:[10.1186/s12982-015-0028-5](https://doi.org/10.1186/s12982-015-0028-5).
- [2] H.W. Hethcote, The basic epidemiology models: models, expressions for R_0 , parameter estimation, and applications, in: *Mathematical Understanding of Infectious Disease Dynamics*, 2009, pp. 1–61, doi:[10.1142/9789812834836_0001](https://doi.org/10.1142/9789812834836_0001).
- [3] H.-J. Zimmermann, An application-oriented view of modeling uncertainty, *Eur. J. Oper. Res.* 122 (2) (2000) 190–198, doi:[10.1016/S0377-2217\(99\)00228-3](https://doi.org/10.1016/S0377-2217(99)00228-3).
- [4] J. Helton, J. Johnson, W. Oberkampf, An exploration of alternative approaches to the representation of uncertainty in model predictions, *Reliab. Eng. Syst. Saf.* 85 (1) (2004) 39–71, doi:[10.1016/j.ress.2004.03.025](https://doi.org/10.1016/j.ress.2004.03.025).
- [5] D. Gulber, Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century, *Trends Microbiol.* 10 (2) (2002) 100–103, doi:[10.1016/S0966-842X\(01\)02288-0](https://doi.org/10.1016/S0966-842X(01)02288-0).
- [6] R.C. Rodríguez, G. Carrasquilla, A. Porras, K. Galera-Gelvez, J.G.L. Yescas, J.A. Rueda-Gallardo, The burden of dengue and the financial cost to Colombia, 2010–2012, *Am. J. Trop. Med. Hyg.* 94 (5) (2016) 1065–1072, doi:[10.4269/ajtmh.15-0280](https://doi.org/10.4269/ajtmh.15-0280).
- [7] N.S. Nedialkov, Interval tools for ODEs and DAEs, in: *12th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN 2006)*, IEEE, 2006, p. 4.
- [8] Y. Lin, M.A. Stadtherr, Validated solutions of initial value problems for parametric ODEs, *Appl. Numer. Math.* 57 (10) (2007) 1145–1162, doi:[10.1016/j.apnum.2006.10.006](https://doi.org/10.1016/j.apnum.2006.10.006).
- [9] N. Nedialkov, K. Jackson, G. Corliss, Validated solutions of initial value problems for ordinary differential equations, *Appl. Math. Comput.* 105 (1) (1999) 21–68, doi:[10.1016/S0096-3003\(98\)10083-8](https://doi.org/10.1016/S0096-3003(98)10083-8).
- [10] IEEE Standard for Interval Arithmetic, 2015. 10.1109/IEEESTD.2015.7140721.
- [11] R.E. Moore, R.B. Kearfott, M.J. Cloud, *Introduction to Interval Analysis*, SIAM, 2009.
- [12] A.S. Ackleh, E.J. Allen, R.B. Kearfott, P. Seshaiyer, *Classical and Modern Numerical Analysis. Theory, Methods and Practice*, Chapman & Hall/CRC Numerical Analysis and Scientific Computing, 2010.
- [13] M. Neher, From interval analysis to Taylor models-an overview, in: *Proc. IMACS*, 2005.

- [14] K. Makino, M. Berz, Taylor models and other validated functional inclusion methods, *Int. J. Pure Appl. Math.* 6 (2003) 239–316.
- [15] K. Makino, M. Berz, Verified computations using Taylor models and their applications, in: *International Workshop on Numerical Software Verification*, Springer, 2017, pp. 3–13.
- [16] M. Neher, K.R. Jackson, N.S. Nedialkov, On Taylor model based integration of odes, *SIAM J. Numer. Anal.* 45 (1) (2007) 236–262.
- [17] R.J. Lohner, On the Ubiquity of the Wrapping Effect in the Computation of Error Bounds, Springer Vienna, 2001, pp. 201–216, doi:[10.1007/978-3-7091-6282-8_12](https://doi.org/10.1007/978-3-7091-6282-8_12).
- [18] J. Enszer, M. Stadtherr, Verified solution method for population epidemiology models with uncertainty, *Int. J. Appl. Math. Comput. Sci.* 19 (3) (2009) 501–512, doi:[10.2478/v10006-009-0040-4](https://doi.org/10.2478/v10006-009-0040-4).
- [19] K. Makino, M. Berz, Remainder differential algebras and their applications, in: *Computational Differentiation: Techniques, Applications and Tools*, 1996, pp. 63–74.
- [20] D.P. Lizarralde-Bejarano, D. Rojas-Díaz, S. Arboleda-Sánchez, M.E. Puerta-Yepes, Sensitivity, uncertainty and identifiability analyses to define a dengue transmission model with real data of an endemic municipality of Colombia, *PLoS One* 15 (3) (2020) e0229668, doi:[10.1371/journal.pone.0229668](https://doi.org/10.1371/journal.pone.0229668).
- [21] D.P. Lizarralde-Bejarano, S. Arboleda-Sánchez, M.E. Puerta-Yepes, Understanding epidemics from mathematical models: details of the 2010 dengue epidemic in Bello (Antioquia, Colombia), *Appl. Math. Model.* 43 (2017) 566–578, doi:[10.1016/j.apm.2016.11.022](https://doi.org/10.1016/j.apm.2016.11.022).
- [22] L. Esteve, C. Vargas, A model for dengue disease with variable human population, *J. Math. Biol.* 38 (3) (1999) 220–240, doi:[10.1007/s002850050147](https://doi.org/10.1007/s002850050147).
- [23] F. Brauer, Mathematical epidemiology: past, present, and future, *Infect. Dis. Model.* 2 (2) (2017) 113–127, doi:[10.1016/j.idm.2017.02.001](https://doi.org/10.1016/j.idm.2017.02.001).
- [24] O. Diekmann, J.A. Heesterbeek, J.A. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.* 28 (4) (1990) 365–382, doi:[10.1007/BF00178324](https://doi.org/10.1007/BF00178324).
- [25] M. Martcheva, Ch. 6: Fitting Models to Data, *An Introduction to Mathematical Epidemiology*, Springer US, 2015, doi:[10.1007/978-1-4899-7612-3_6](https://doi.org/10.1007/978-1-4899-7612-3_6).
- [26] M.P. Saccomani, Structural vs. practical identifiability in system biology, *IWBBIO 2013 Proceedings* (2013) 305–313.
- [27] H. Miao, X. Xia, A.S. Perelson, H. Wu, On identifiability of nonlinear ODE models and applications in viral dynamics, *SIAM Rev.* 53 (1) (2011) 3–39, doi:[10.1137/090757009](https://doi.org/10.1137/090757009).
- [28] O.-T. Chis, J.R. Banga, E. Balsa-Canto, Structural identifiability of systems biology models: a critical comparison of methods, *PLoS One* 6 (11) (2011) e27755, doi:[10.1371/journal.pone.0027755](https://doi.org/10.1371/journal.pone.0027755).
- [29] A.D. Langmuir, William Farr: founder of modern concepts of surveillance, *Int. J. Epidemiol.* 5 (1) (1976) 13–18.
- [30] W.H. Organization, et al., *Dengue: Guías para el Diagnóstico, Tratamiento, Prevención y Control: Nueva Edición*, Technical Report, Ginebra: Organización Mundial de la Salud, 2009.
- [31] S. Bhatt, P.W. Gething, O.J. Brady, J.P. Messina, A.W. Farlow, C.L. Moyes, J.M. Drake, J.S. Brownstein, A.G. Hoen, O. Sankoh, M.F. Myers, D.B. George, T. Jaenisch, G.R.W. Wint, C.P. Simmons, T.W. Scott, J.J. Farrar, S.I. Hay, The global distribution and burden of dengue, *Nature* 496 (7446) (2013) 504–507, doi:[10.1038/nature12060](https://doi.org/10.1038/nature12060).
- [32] J.C. Padilla, D.P. Rojas, R.S. Gómez, *Dengue en Colombia: Epidemiología de la Reemergencia a la Hiperendemia*, Guías de Impresión Ltda., 2012.
- [33] J. Karlsson, M. Anguelova, M. Jirstrand, An efficient method for structural identifiability analysis of large dynamic systems, *IFAC Proc. Vol.* 45 (16) (2012) 941–946, doi:[10.3182/20120711-3-BE-2027.00381](https://doi.org/10.3182/20120711-3-BE-2027.00381). 16th IFAC Symposium on System Identification
- [34] I.A. Rather, H.A. Parray, J.B. Lone, W.K. Paek, J. Lim, V.K. Bajpai, Y.-H. Park, Prevention and control strategies to counter dengue virus infection, *Front. Cell. Infect. Microbiol.* 7 (2017) 336, doi:[10.3389/fcimb.2017.00336](https://doi.org/10.3389/fcimb.2017.00336).