



**HAL**  
open science

# Calibration of a SEIR epidemic model to describe Zika virus outbreak in Brazil

Eber Dantas, Michel Tosin, Americo Cunha Jr

► **To cite this version:**

Eber Dantas, Michel Tosin, Americo Cunha Jr. Calibration of a SEIR epidemic model to describe Zika virus outbreak in Brazil. 2017. hal-01456776v2

**HAL Id: hal-01456776**

**<https://hal.science/hal-01456776v2>**

Preprint submitted on 15 Feb 2017 (v2), last revised 4 Jul 2018 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Calibration of a SEIR epidemic model to describe Zika virus outbreak in Brazil

Eber Dantas<sup>a</sup>, Michel Tosin<sup>a</sup>, Americo Cunha Jr<sup>a,\*</sup>

<sup>a</sup>*NUMERICO – Nucleus of Modeling and Experimentation with Computers,  
Universidade do Estado do Rio de Janeiro (UERJ), Rua São Francisco Xavier, 524, Rio  
de Janeiro - RJ, 20550-900, Brasil. Phone: +55 21 2334-0323 r: 208*

---

## Abstract

In the last two decades multiple instances of Zika virus epidemic have been reported around the world, turning the related disease into a international concern. In this context, the use of mathematical models for epidemics is of great importance, since they are useful tools to predict the outbreaks underlying numbers, and allow one to test the effectiveness of different strategies to combat associated diseases. This work deals with the development and calibration of a epidemic model to describe the 2016 outbreak of Zika virus in Brazil. A mathematical model with 8 differential equations and 7 parameters is employed. Nominal values for the model parameters are estimated from the literature. An inverse problem associated to the model identification is formulated and solved. The calibrated model obtained presents realistic parameters and returns reasonable predictions, with the curve shape similar to the outbreak evolution, and peak value close to the maximum number of infected people during 2016.

*Keywords:* Zika virus dynamics; nonlinear dynamics; mathematical biology; SEIR epidemic model; model calibration; system identification

---

---

\*Corresponding author.

*Email addresses:* eber.paiva@uerj.br (Eber Dantas), micehl.tosin@uerj.br (Michel Tosin), amero@ime.uerj.br (Americo Cunha Jr)

## 1. Introduction

The Zika virus is a flavivirus that upon infection in humans causes a illness, known as Zika fever, identified commonly with macular or papular rash, mild fever and arthritis [62, 7]. It is mainly a vector-borne disease, carried by the genus *Aedes* of mosquitoes [62, 26], although sexual transmission has been reported [16, 63] and contamination by blood transfusion is under investigation [4]. The Zika virus was first isolated in primates from the Zika forest in Uganda in 1947 [19]. Evidences of the virus in humans were found in Nigeria in 1968 [42]. An epidemic occurred in 2007 on Micronesia [21], followed by multiple outbreaks on several Pacific Islands between 2013 and 2014 [12, 58]. The first Zika virus autochthonous case in Brazil was reported around May, 2015 [66, 51], and nearly 30,000 cases of infection were already notified by January 30, 2016 [24], along with the Pan American Health Organization being informed in the same month about locally-transmitted cases on numerous continental and island territories of America [30]. The Brazilian Ministry of Health registered 211,770 probable cases of Zika fever (126,395 of which were confirmed) until the 49th epidemiological week (EW; 4–10 Dec, 2016) [51]. See in Figure 1 the empirical data of suspected and confirmed cases for the epidemiological weeks 1 to 34 of 2016 in Brazil. This epidemic caused concern in the international medical community, health authorities and population, specially due to a relation between Zika virus and other diseases — such as newborn microcephaly [59, 60] and Guillain-Barré syndrome [20], whose correlation to the Zika virus has recently been considered by the World Health Organization (WHO) a “*scientific consensus*” [63] — and the realization of the 2016 Summer Olympics in Rio de Janeiro, Brazil, which on account of the large flux of people could have spread the Zika virus to various parts of the world.

In this epidemic scenario, the development of control and prevention strategies for the disease is a critical issue. A mathematical model able to predict the number of infected people during the virus outbreak is an useful tool, which can be employed to identify effective and vulnerable aspects on disease control programs [11, 31, 37, 38, 48, 56, 67]. Nevertheless, for an epidemic model to be truly useful, it must undergo a judicious process of *validation* [17, 44], where model predictions are confronted with real data, in order to evaluate if they are realistic. In general, the first predictions of a model do not agree with the observations, which may occur due to possible inadequacies in the model hypotheses or a poor choice for the model nom-

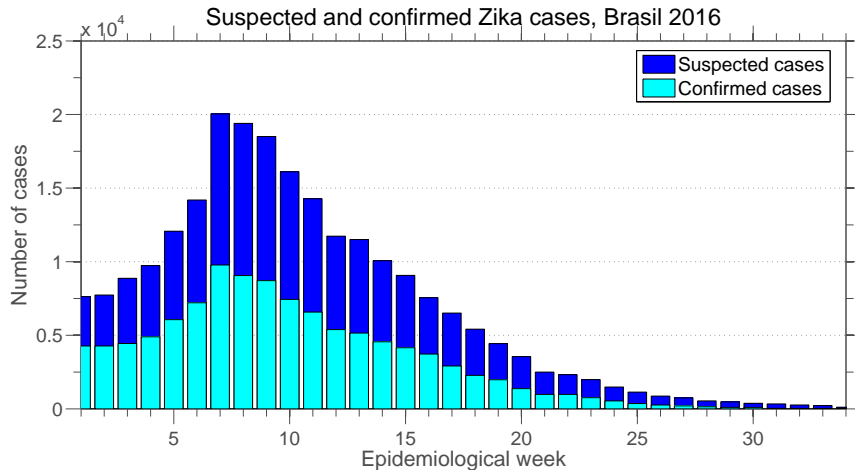


Figure 1: Confirmed and suspected cases of Zika infection in Brazil from 1-st to 34-th epidemiological week [47, 52].

inal parameters. The first case invalidates the model, but the second can be remedied through a procedure known as *model calibration*, where a set of parameters that promote a good agreement between predictions and observations is sought. After the calibration process, the model must be confronted again with new experimental data to ensure its validity.

This work constitutes a first step in a long and rigorous process of identification and validation of representative model to describe Zika virus 2016 outbreak in Brazil. For this purpose, a mathematical model is adapted to the Brazilian scenario based on similar SEIR descriptions used successfully for other Zika outbreaks that happened around the world [28, 33] and on studies performed over SEIR dynamic systems capable of modeling epidemics [41, 61, 64]. The nominal values of the model parameters belong to Zika infection's and vector's characteristics, quantitatively estimated in the literature or published by health organizations. Predictions are obtained from numerical simulation and further heuristic manipulation of the parameters, and are compared to real data of the outbreak, in a first step to validate the model. In sequence, a rigorous process of model calibration is performed through the formulation and solution of an inverse problem.

The rest of this paper is organized as follows. In section 2, the mathematical model is described and estimation of nominal values for the model

parameters is discussed. In section 3, the reader is exposed to an attempt to validate the model, where results are presented with subsequent comparison between model predictions and experimental data. Section 4 presents the formulation and solution of an inverse problem to calibrate the model. Finally, in section 5, the main contributions of this work are emphasized and a path for future works is suggested.

## 2. Epidemic model for Zika virus dynamics

### 2.1. Model hypotheses

This work utilizes a variant of the Ross-McDonald model [50, 53] for epidemic predictions, separating the populations into a SEIR framework (susceptible, exposed, infectious, recovered) [8, 9, 22, 40]. Each category represents the health condition of a individual inside such group at time  $t$ , with respect to the considered infection. The susceptible group, denoted by  $S(t)$ , represents those who are uncontaminated and are able to become infected. The exposed portion of the population,  $E(t)$ , comprehends anyone that is carrying the pathogen but is still incapable of transmitting the disease. While the infectious individuals,  $I(t)$ , can spread the pathogen and may display symptoms associated with the illness. Finally, the recovered group,  $R(t)$ , contains those who are no longer infected. The populations of humans and vectors are segmented into the aforesaid classes (excepting the group of recovered vectors), as Figure 2 depicts schematically with the accordingly subscripts.  $S_h, E_h, I_h$  and  $R_h$  amass the number of people at each stage of the model description, and  $S_v, E_v, I_v$  signifies proportion of vectors. Note that  $0 \leq S_v, E_v, I_v \leq 1$ .

Demographical changes in the number of humans are not considered, and the total vector population is maintained constant during the analysis, although variations of the proportion of vectors on the particular SEIR compartments are introduced via birth and death rates. The vector in question is regarded as a hypothetical mosquito apt to being infected or infectious throughout all its lifetime — which means the model accounts only for the adult stage of their life cycle — and also unable to recover.

The time elapsed by a individual while he/she is on the aforementioned exposed group is known as the latent period of a organism and, in this work, is adopted as equivalent to the so-called incubation period (the time elapsed between being infected and exhibiting symptoms), since data is extremely sparse on the latter for humans [13] (namely, the intrinsic incubation period).

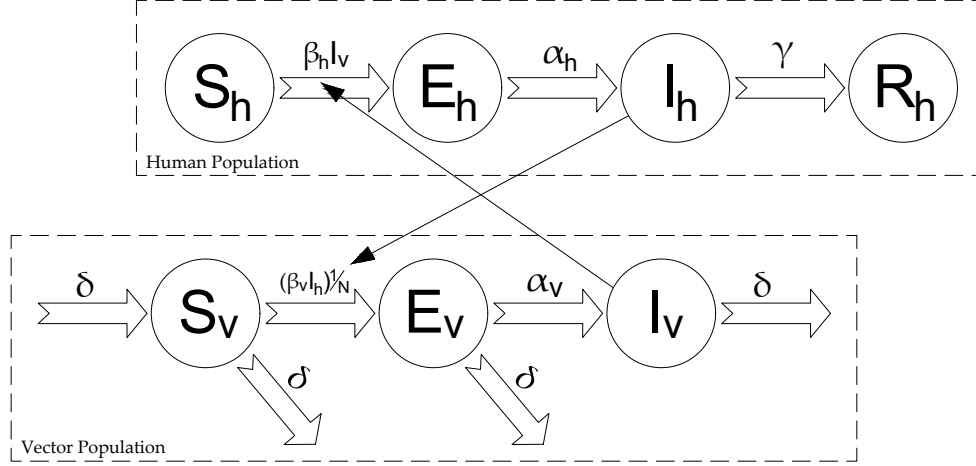


Figure 2: Schematic representation of SEIR model for Zika virus outbreak description.

Both terms are used interchangeably hereafter, and the concepts do not differ on the mosquitoes' case (the extrinsic incubation period) [34]. In addition, all the members of the susceptible group are treated as equally capable of being infected and the recovered ones as completely immunized.

## 2.2. Model equations

The evolution of individuals through the SEIR groups is governed by the following (nonlinear) autonomous system of ordinary differential equations (ODE)

$$\frac{dS_h}{dt} = -\beta_h S_h I_v, \quad (1)$$

$$\frac{dE_h}{dt} = \beta_h S_h I_v - \alpha_h E_h, \quad (2)$$

$$\frac{dI_h}{dt} = \alpha_h E_h - \gamma I_h, \quad (3)$$

$$\frac{dR_h}{dt} = \gamma I_h, \quad (4)$$

$$\frac{dS_v}{dt} = \delta - \beta_v S_v \frac{I_h}{N} - \delta S_v, \quad (5)$$

$$\frac{dE_v}{dt} = \beta_v S_v \frac{I_h}{N} - (\delta + \alpha_v) E_v, \quad (6)$$

$$\frac{dI_v}{dt} = \alpha_v E_v - \delta I_v, \quad (7)$$

$$\frac{dC}{dt} = \alpha_h E_h, \quad (8)$$

where  $N$  represents the total human population,  $1/\alpha$  is the disease's incubation period (each with the corresponding subscript of  $h$  for human's and  $v$  for vector's),  $1/\delta$  means the vector lifespan,  $1/\gamma$  is the human infection period — which is defined in this work as the interval of time that a human is infectious — and  $\beta$  identifies the transmission rate, specifically  $\beta_h$  is the mosquito-to-human rate and  $\beta_v$  the human-to-mosquito rate. The first delta on the right hand side of Eq. (5) sustains the hypothesis that the total vector population is unchanged during the analysis. Additionally, Eq. (8) allows evaluation of the cumulative number  $C(t)$  of infected people until the time  $t$ , that is, the amount of humans so far that contracted the disease and have passed through or are in the infectious group at the given time.

### 2.3. Nominal parameters

The preliminary values for the parameters of Eqs. (1)–(8) come from the related literature concerning the Zika infection, the *Aedes aegypti* mosquito (which is the vector for Zika, Dengue and Yellow fever in Brazil), vector-borne epidemic models and reports provided by health organizations and government agencies. The Brazilian Institute of Geography and Statistics (IBGE) reports that Brazil had approximately  $N = 206 \times 10^6$  people by July, 2016 [27]. The adopted extrinsic incubation period (EIP) is  $1/\alpha_v = 15$  days, as is proposed in [6], showing it as the average time required for detection of significant levels of Zika virus in adult *Ae. aegypti* mosquitoes blood. This value agrees with other statistical confidence intervals (CI) that are presented for the parameter in another works (95% CI: 4.4–17 days [28]). A systematic review and pooled analysis of the literature and case studies available in [35] suggests that 95% of people infected by the Zika virus who develop symptoms will do so within 11.2 days of infection (95% CI: 7.6–18.0). This value is selected in this work for the intrinsic incubation period (IIP)  $1/\alpha_h$  and is in agreement with the range of 3 – 12 days that the European Centre

for Disease Prevention and Control (ECDC) recommends [23], also formerly used in other works [32, 60]. The aforementioned literature analysis in [35] also concludes that 95% of the cases will have no detectable virus in the blood by 18.9 days after infection (95% CI: 13.6–79.4). Considering the assumption that the infectiousness in Zika infection ends 1.5 – 2 days before the virus becomes undetectable [25, 28], the chosen value for the human infectious period is  $1/\gamma = 18.9 - 1.5 = 17.4$  days. As for the vector lifespan  $1/\delta$ , “*the adult stage of the mosquito is considered to last an average of eleven days in the urban environment*” according to [45]. This is the assumed value for the parameter in this work, which is also consistent with biological studies about the species [43] and agrees with the usual life expectancy for the mosquito in Rio de Janeiro, Brazil [18]. Lastly, the time between a mosquito being infected and it infecting a human,  $1/\beta_h$ , and the time between a human infection and a mosquito taking an infectious blood meal,  $1/\beta_v$ , is estimated in [25] as an average of 11.3 days (95% CI: 8.0–16.3) and 8.6 days (95% CI: 6.2–11.6), respectively. A summary of the nominal values adopted for the SEIR model parameters is shown in Table 1.

Table 1: Nominal parameters of SEIR model.

parameter	value	unit
$\alpha_h$	1/11.2	days <sup>-1</sup>
$\alpha_v$	1/15	days <sup>-1</sup>
$\gamma$	1/17.4	days <sup>-1</sup>
$\delta$	1/11	days <sup>-1</sup>
$\beta_h$	1/11.3	days <sup>-1</sup>
$\beta_v$	1/8.6	days <sup>-1</sup>
$N$	$206 \times 10^6$	individuals

#### 2.4. Initial conditions

In order to properly define the dynamic system underlying SEIR epimidec model presented in section 2.2, it is necessary to prescribe the value of its variables at the initial time  $t = 0$ , established as the first epidemiological week of 2016.

In this date, both populations are assumed completely susceptible at first, meaning  $S_h(0) = N$  and  $S_v(0) = 1$ . The initial values for the exposed and



infectious groups are considered the same,  $I_h(0) = E_h(0)$  and  $I_v(0) = E_v(0)$ . Likewise, the initial value of the cumulative number must be  $C(0) = I_h(0)$ , given its definition. The value of  $I_h(0)$  is taken as the number of confirmed Zika cases in Brazil on the first EW of 2016 (see Figure 1) [47, 52]. The recovered group is assumed null at the beginning of infection,  $R_h(0) = 0$ . As for the proportion of infectious vectors in the first week, repetitive manual estimations were tried until the resulted time series of  $I_h$  presented reasonable values compared to the real data. It became clear that the system response is very sensible to  $I_v(0)$ , more than the other initial conditions, as slight variations in its value are required to achieve feasible results. In the process of choosing its value, the matching of the  $I_h$  curve's peak value to the amplitude of infection is also taken as a priority, since this is the main interest region for evaluation of the outbreak. The nominal values of the parameters exhibited viable  $I_h$  curves around  $I_v(0) = 6.5 \times 10^{-5}$ . A summary of the nominal values adopted for the SEIR model parameters is shown in Table 2.

Table 2: Nominal initial conditions of SEIR model.

parameter	value	unit
$S_h(0)$	$206 \times 10^6$	individuals
$E_h(0)$	4 272	individuals
$I_h(0)$	4 272	individuals
$R_h(0)$	0	individuals
$C(0)$	4 272	individuals
$S_v(0)$	1	—
$E_v(0)$	$13.5 \times 10^{-4}$	—
$I_v(0)$	$1.5 \times 10^{-4}$	—

### 3. Attempt to validate the epidemic model

The SEIR model response, for the set of parameters defined in sections 2.3 and 2.4, is presented in Figures 3 and 4 on a epidemiological week temporal domain consisting of one to fifty-two weeks (7 to 365 days). It is possible to see an outbreak around the fifth epidemic week, once time series of infected and exposed people show a sudden increase.

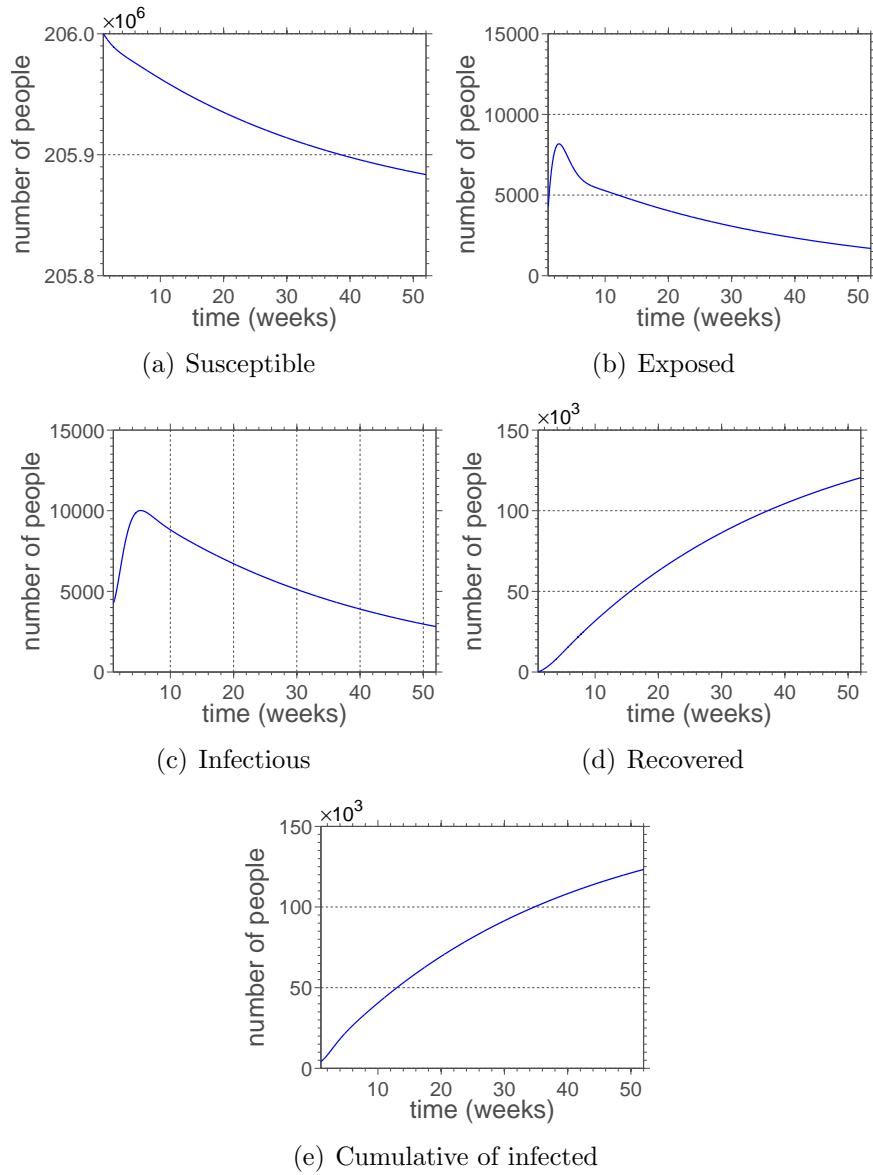


Figure 3: Model response for human-related quantities using the nominal parameters.

Unfortunately, from the quantitative point of view, the model predictions shown in Figures 3 and 4 are not realistic. The veracity of such an assertion is proved by Figure 5, which compares the time series of  $I_h$  with the real data for the number of confirmed Zika cases in each EW [52]. Clearly, the

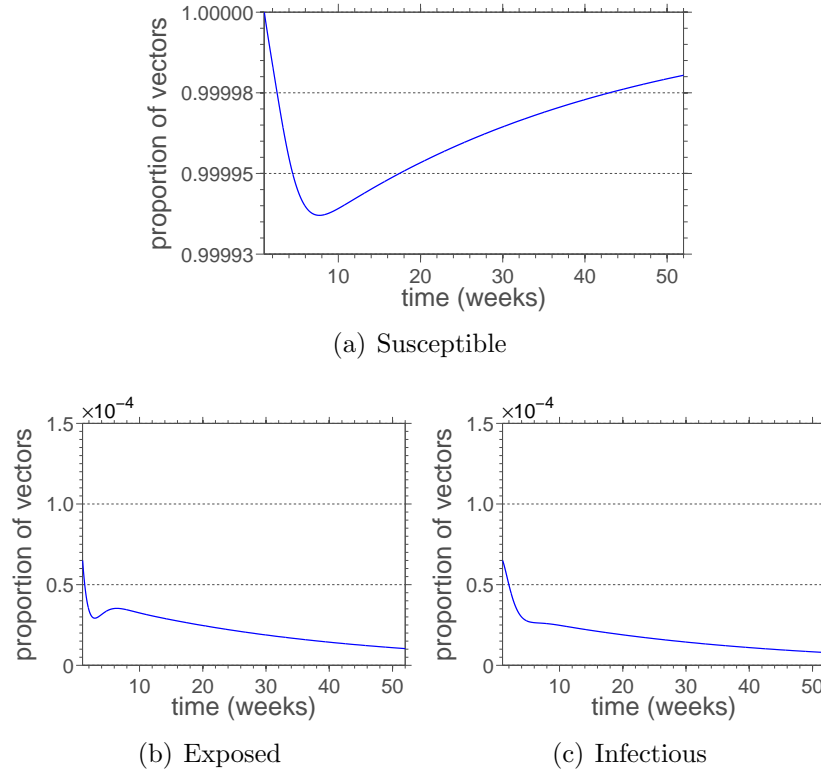


Figure 4: Model response for vector-related quantities using the nominal parameters.

given  $I_h$  curve overpredicts the infection numbers. This result proves that SEIR model of section 2, provided with the parameters of sections 2.3 and 2.4, is not validated, i.e., it does not faithfully represent Zika virus dynamics in Brazil throughout 2016.

Meanwhile, the general shape of  $I_h$  do provide qualitative information about the evolution of the infection, as well as predictions for the peak's value in the same order of magnitude than that of the empirical data and its time of occurrence with a less-than-two-weeks error. This qualitative agreement, between simulation and epidemic data, induces the hypothesis that model predictions may differ from the real data due to the use of unsuitable values for model parameters. In other words, model predictions would be (possibly) closer to the reference values if more realistic parameters were used.

The search for parameter values that make the simulations fit well to observed data is known in the literature as model calibration (or system

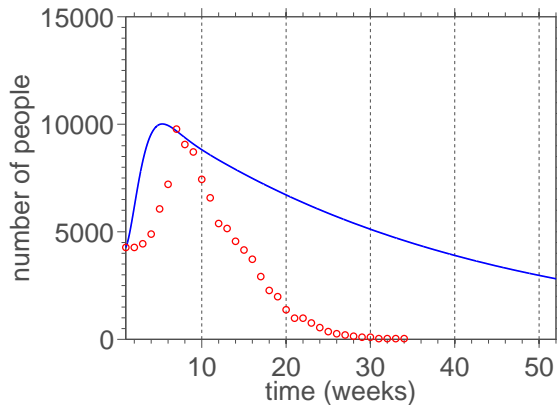


Figure 5: Comparison between the number of infected humans predicted by the model (curve) and the number of confirmed cases in each EW [52] (dots).

identification) [3, 5, 10, 55], being the object of interest of the next section.

#### 4. Calibration of the epidemic model

In general, the calibration of a mathematical model is the first step in the so called *validation process*, procedure where simulations and experiments (data) are compared, aiming to prove (or disprove) the model representativeness [17, 44]. Once calibration is done, the model response must be compared again with new reference data, independently obtained from the data used in the calibration, to prove (or disprove) the model quality [29, 54].

This work has at its disposal a single set of data for Zika virus outbreak in Brazil. Thus, only the calibration process can be conducted, as a first step in the direction of validating the model. This calibrated model can be tested evaluated in the future, when new observations become available.

##### 4.1. Forward and inverse problem

From the mathematical point of view, the epidemic model of section 2 is a continuous-time dynamical system of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{p}), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (9)$$

where  $\mathbf{x}_0 \in \mathbb{R}^8$  is a prescribed initial condition vector, the vector  $\mathbf{p} \in \mathbb{R}^7$  lumps the model parameters,  $\mathbf{f} : \mathbb{R}^8 \times \mathbb{R}^7 \rightarrow \mathbb{R}^8$  is a nonlinear map which

defines the laws of evolution of this dynamics, and  $\mathbf{x}(t) \in \mathbb{R}^8$  the vector of states at time  $t$ .

Given initial conditions and a set of parameters, represented by the pair  $(\mathbf{x}_0, \mathbf{p})$ , it is possible to compute, by means of numerical integration, the model response  $\mathbf{x}(t)$  from which a scalar observable  $\phi(\mathbf{x}_0, \mathbf{p}, t)$  is obtained. This is called the *forward problem* associated to the epidemic model.

The model calibration problem seeks to find a set of parameters that, in a certain sense, makes the response of the model as close as possible to the system observations (reference data). In formal terms, given a data set with  $M$  scalar observation  $y_1, y_2, \dots, y_M$ , and for a prescribed initial condition  $\mathbf{x}_0$ , find  $\mathbf{p}^*$  such that

$$\mathbf{p}^* = \arg \min_{\mathbf{p}} \left\{ \sum_{n=1}^M \left| y_n - \phi(\mathbf{x}_0, \mathbf{p}, t_n) \right|^2 \right\}, \quad (10)$$

where  $\{t_n\}_{n=1}^M$  corresponds to a discrete set of time-instants where the  $M$  scalar observation are obtained.

This is called an *inverse problem* associated to the epidemic model. In general this type of problem is extremely nonlinear, with none or low regularity, multiple solutions (or even none), being much more complicated to attacked in comparison with the forward problem [2, 57, 65]. A schematic representation of the forward and the inverse problem associated to the epidemic model is shown in Figure 6.

#### 4.2. Method of Levenberg-Marquardt

The method of Levenberg-Marquardt (LM) [36, 39, 49], is employed here to numerically approximate a solution for inverse problem of Eq.(10). The choice for this method is substantiated on its ability of rapidly converge, once a suitable neighborhood has been reached, with the rather bigger convergence capacity present in the steepest descent methods [14].

The basis of the LM method is the linear approximation to  $\phi$  in the neighborhood of  $\mathbf{p}$

$$\phi(\mathbf{x}_0, \mathbf{p} + \mathbf{h}, t) \approx \phi(\mathbf{x}_0, \mathbf{p}, t) + \mathbf{J} \mathbf{h}, \quad (11)$$

where  $\mathbf{J} = \partial\phi/\partial\mathbf{p}$  is the Jacobian matrix at  $\mathbf{p}$ , and  $\mathbf{h}$  is small perturbation of  $\mathbf{p}$ , i.e.,  $\|\mathbf{h}\| \ll 1$ .

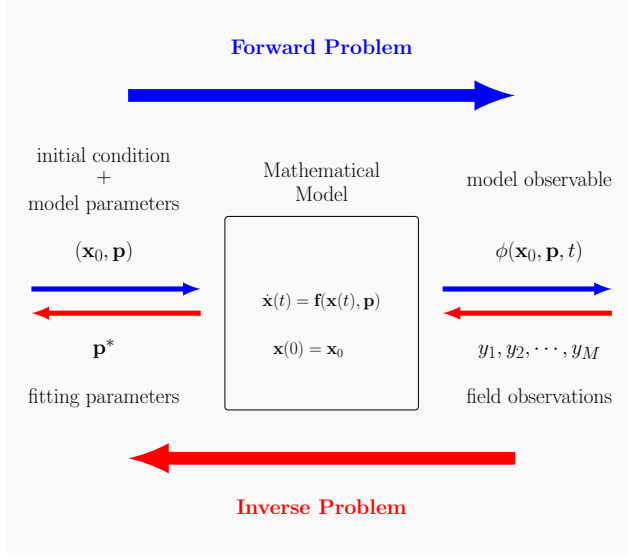


Figure 6: Schematic representation of the forward and inverse problems associated to the epidemic model.

Starting at  $\mathbf{p}_0$  — an initial guess — the method produces a series of vectors  $\mathbf{p}_1, \mathbf{p}_2, \dots$ , that converge towards a local minimizer  $\mathbf{p}^*$  for  $\phi$ . The  $r$ th iteration is defined by

$$\left( \mathbf{A}^{(r)} + \lambda^{(r)} \mathbf{I} \right) \mathbf{h}^{(r)} = \mathbf{g}^{(r)}, \quad (12)$$

where  $\mathbf{A} = \mathbf{J}^T \mathbf{J}$  and  $\mathbf{g} = \mathbf{J}^T (\mathbf{y} - \phi(\mathbf{x}_0, \mathbf{p}, t) - \mathbf{J} \mathbf{h})$  [49]. The scalar  $\lambda$  is a Lagrange multiplier and controls both the magnitude and direction of the step size correction  $\mathbf{h}$ . When  $\lambda$  is zero, the direction of  $\mathbf{h}$  is identical to that of the Gauss-Newton method, and as  $\lambda \rightarrow \infty$ ,  $\mathbf{h}$  tends towards the steepest-descent direction, with its magnitude tending to zero. Thus,  $\lambda$  can be controlled to ensure descent even when second-order terms are encountered, which would otherwise restrict the efficiency of the Gauss-Newton method. The parameter vector at each step,  $\mathbf{p}^{(r+1)} = \mathbf{p}^{(r)} + \mathbf{h}^{(r)}$ , will lead to a new sum of squares in Eq.(10), and it is essential to select  $\lambda$  such that this sum is strictly smaller than the previous one. A sufficiently large  $\lambda$  always exists such this condition is met, unless  $\mathbf{p}^{(r)}$  is already at a minimum [39].

### 4.3. Numerical experiments for calibration

Initially a sensitivity analysis process was conducted. The first analysis performed was a variation of a single parameter while the others remained with the same nominal values presented in section 2.3. Such process allowed the observation of which parameters the system was more sensible to. The parameters  $\alpha_h$  and  $\alpha_v$  were proved unreliable, since they required very high initial guesses and usually would deform the quality of the curve. On the other hand, the model was very sensible to variations on  $\beta_h$  and  $\beta_v$ , being largely more affected by the values of  $\beta_h$ . The other parameters did not bring remarkable conclusions.

Next, a two-varying-parameters attempt was conducted. Most combinations did not bring satisfactory results, to the extent that some pairs of parameters could not even be computed in the time dedicated to the analysis, e.g.  $(\alpha_h, \alpha_v)$ , probably because of inefficient initial guesses or high computational cost. Besides, some parameters seemed to control the quality changing of the considered curve, meaning the accompanying parameter would vary relatively less. Such control parameters were  $\beta_h$  and  $\beta_v$  and, unsurprisingly, this pair proved to be the best one for fitting purposes. Three-way-varying attempts were made, but usually would result not computable or giving in to the referred control parameters. Thus,  $\beta_h$  and  $\beta_v$  were chosen to be the focus of the calibration process.

Figure 7 shows the best result obtained for the  $I_h$  curve fitting problem using the parameters from section 2.3, and is found by singly varying  $\beta_v$  via Levenberg-Marquardt algorithm. The initial guess used for  $\beta_v$  in this case and the values of the parameters held constant are the same as described in section 2.3, as well as the initial conditions used for the numerical integration are the ones mentioned in section 3. The resulting  $1/\beta_v$  that graphs the curve is 15.5 days.

It is clear in Figure 7 that the system response  $I_h$  is a reasonable prediction of the outbreak: the general shape of the infection evolution is attained, the curve peak and empirical data maximum value differ only by a couple hundred of individuals, and all parameters and initial conditions are within realistic possibilities.

Another result to the inverse problem is presented, considering again the fitting of the  $I_h$  curve. The chosen values for this second set of parameters are the product of comparing the empirical data with multiple iterations of the numerical strategy, but lacking the caution for the possibility of unrealistic measures for the parameters. The initial conditions for the numerical

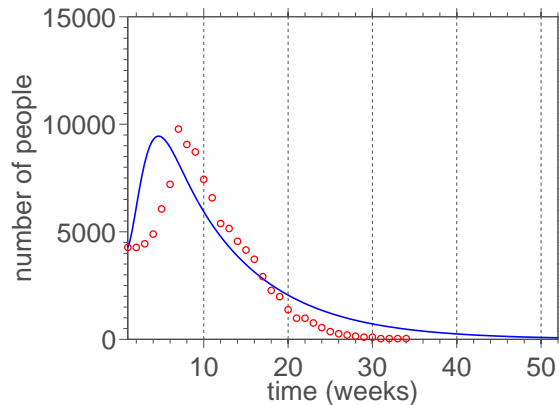


Figure 7: Calibrated time series of infectious number of humans using the nominal parameters.

Table 3: Second set of parameters used on the calibration process that provided Figure 8. The  $\beta_h$  value is the initial guess for the LM algorithm.

parameter	value	unit
$\alpha_h$	1/12	days <sup>-1</sup>
$\alpha_v$	1/17	days <sup>-1</sup>
$\gamma$	1/8.8	days <sup>-1</sup>
$\delta$	1/25	days <sup>-1</sup>
$\beta_h$	1/16.3	days <sup>-1</sup>
$\beta_v$	1/11.6	days <sup>-1</sup>
$N$	$206 \times 10^6$	individuals

integration in this analysis follow the same assumptions presented in section 3, excepting the  $E_v(0) = I_v(0)$  hypothesis, since the values of these are the result of an additional heuristic process of calibration: successive applications of the LM algorithm were performed while manually changing the values of  $I_v(0)$  and  $E_v(0)$  at each application, searching for the best fit of the curve's peak to the field data. Figure 8 presents the most satisfactory result obtained through this analysis, utilizing a  $\beta_h$  variation in the LM method. The parameters maintained constant and the  $\beta_h$  initial guess used in this calibration process to graph Figure 8 are summarized in Table 3. The resulting  $\beta_h$  after the LM algorithm is 0.0127 days<sup>-1</sup>.



The curve in Figure 8 presents a better calibration of the model according to the empirical data, since the peak's time of the  $I_h$  curve is significant closer to the epidemiological week that registered the maximum number of infected people. However, this result comes at the cost of physical meaning in the parameters, because  $1/\beta_h = 78.7$  days for the time between a mosquito being infected and it infecting a human is certainly unrealistic.

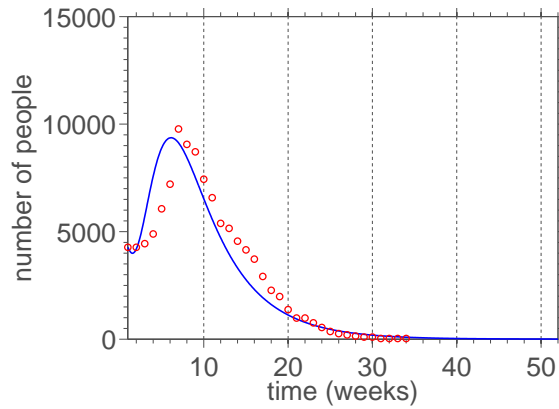


Figure 8: Calibrated time series of infectious number of humans using the second set of parameters.

## 5. Concluding remarks

An epidemic model of SEIR type to describe the dynamics of Zika virus outbreak in Brazil, occurred in 2016, is developed and calibrated in this work. Nominal quantities for the parameters are selected from the related literature concerning the Zika infection, the *Aedes aegypti* genus of mosquitoes, vector-born epidemic models and information provided by health organizations. The calibration process is done through the solution of an inverse problem with the aid of Levenberg-Marquardt method, used to pick the best parameter values that would fit the curve “number of infectious people per week” into the disease’s empirical data, thus calibrating the model. Results within realistic values for the parameters are presented, stating reasonable predictions with the curve shape similar to the outbreak evolution and proximity between the estimated peak value and data for maximum number of infected during 2016. Improved fitting is also achieved via convenient choice of the parameters and

initial conditions during the numerical and heuristic process regarded in the analysis, but at the expense of physical meaning of such parameters.

This work is only the first step in a long project of modeling and prediction of epidemics related to the Zika virus in the Brazilian context. In this way, only the SEIR model was tested here, but in an upcoming study the authors intend to analyze the efficacy of other of epidemic models (e.g. SIR, MSIR, etc). Furthermore, in order to make the calibration process more robust, by taking into account the uncertainties underlying the model and its parameters, the authors intend to use a Bayesian updating rule to attack the inverse problem [46]. The use of statistical methods to choose the most appropriate model, within a set of validated models, is also part of the plans [1, 15].

## Acknowledgments

The authors are indebted to the Brazilian agencies CNPq (National Council for Scientific and Technological Development), CAPES (Coordination for the Improvement of Higher Education Personnel) and FAPERJ (Research Support Foundation of the State of Rio de Janeiro) for the financial support given to this research. They are also grateful to João Peterson and Vinícius Lopes, both engineering students at UERJ, for the collaboration in the initial stages of this work.

## References

- [1] Ando, T. (Ed.), 2010. Bayesian Model Selection and Statistical Modeling. CRC Press.
- [2] Aster, R., Borchers, B., Thurber, C. (Eds.), 2012. Parameter Estimation and Inverse Problems, 2nd Edition. Elsevier.
- [3] Bao, N., Wang, C., 2015. A Monte Carlo simulation based inverse propagation method for stochastic model updating. *Mechanical Systems and Signal Processing* 60–61, 928–944.
- [4] Barjas-Castro, M. L., et al., 2016. Probable transfusion-transmitted Zika virus in Brazil. *Transfusion* 56 (7), 1684–1688.
- [5] Billings, S. A. (Ed.), 2013. Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains. Wiley.

- [6] Boorman, J. P. T., Porterfield, J. S., 1956. A simple technique for infection of mosquitoes with viruses transmission of Zika virus. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 50 (3), 238–242.
- [7] Brasil, P., et al., 2016. Zika virus outbreak in Rio de Janeiro, Brasil: clinical characterization, epidemiological and virological aspects. *PLOS Neglected Tropical Diseases* 10 (4).
- [8] Brauer, F., Castillo-Chavez, C. (Eds.), 2012. *Mathematical Models in Population Biology and Epidemiology*. Springer.
- [9] Brauer, F., van den Driessche, P., Wu, J. (Eds.), 2008. *Mathematical Epidemiology*. Springer.
- [10] Brewick, P. T., Masri, S. F., 2016. An evaluation of data-driven identification strategies for complex nonlinear dynamic systems. *Nonlinear Dynamics* 85 (2), 1297–1318.
- [11] Britton, T., 2010. Stochastic epidemic models: A survey. *Mathematical Biosciences* 225, 24–35.
- [12] Cao-Lormeau, V.-M., et al., 2014. Zika virus, French Polynesia, South Pacific, 2013. *Emerging Infectious Diseases* 20 (6), 1085–1086.
- [13] Chan, M., Johansson, M. A., 2012. The incubation periods of Dengue viruses. *PLOS ONE* 7 (11).
- [14] Chavent, G., 2010. *Nonlinear Least Squares for Inverse Problems: Theoretical Foundations and Step-by-Step Guide for Applications*. Springer.
- [15] Claeskens, G., 2016. Statistical model choice. *Annual Review of Statistics and Its Application* 3, 233—256.
- [16] Coelho, F. C., Durovni, B., Saraceni, V., Lemos, C., Codeco, C. T., Camargo, S., Carvalho, L. M., Bastos, L., Arduini, D., Villela, D. A., Armstrong, M., 2016. Higher incidence of Zika in adult women than adult men in Rio de Janeiro suggests a significant contribution of sexual transmission from men to women. *International Journal of Infectious Diseases* 51, 128–132.

- [17] Cunha Jr, A., 2017. Modeling and Quantification of Physical Systems Uncertainties in a Probabilistic Framework. In: Ekwaro-Osire, S., Gonçalves, A. C., Alemayehu, F. M. (Eds.), Probabilistic Prognostics and Health Management of Energy Systems. Springer, New York, (accepted for publication).
- [18] de Freitas, R. M., Codeço, C. T., de Oliveira, R. L., 2007. Daily survival rates and dispersal of *Aedes aegypti* females in Rio de Janeiro, Brazil. *The American Journal of Tropical Medicine and Hygiene* 76 (4), 659–665.
- [19] Dick, G. W. A., Kitchen, S. F., Haddow, A. J., 1952. Zika virus (I). Isolations and serological specificity. *Transactions of the Royal Society of Tropical Medicine & Hygiene* 46 (5), 509–520.
- [20] dos Santos, T., et al., 2016. Zika virus and the Guillain–Barré syndrome — Case series from seven countries. *New England Journal of Medicine* 375 (16), 1598–1601.
- [21] Duffy, M. R., et al., 2009. Zika virus outbreak on Yap Island, Federated States of Micronesia. *New England Journal of Medicine* 360 (24), 2536–2543.
- [22] Ellner, S. P., Guckenheimer, J. (Eds.), 2006. *Dynamic Models in Biology*. Princeton University Press.
- [23] European Centre for Disease Prevention and Control, November 2015. Rapid risk assessment: Microcephaly in Brazil potentially linked to the Zika virus epidemic. Retrieved from <https://goo.gl/UqJkUi>, Accessed: 02/01/2017.
- [24] Faria, N. R., et al., 2016. Zika virus in the Americas: early epidemiological and genetic findings. *Science* 352 (6283), 345–349.
- [25] Ferguson, N. M., et al., 2016. Countering the Zika epidemic in Latin America. *Science* 353 (6297), 353–354.
- [26] Fernandes, R. S., et al., 2016. *Culex quinquefasciatus* from Rio de Janeiro is not competent to transmit the local Zika virus. *PLOS Neglected Tropical Diseases* 10 (9).

- [27] Fundação Instituto Brasileiro de Geografia e Estatística, August 2016. Diário Oficial da União, N. 167- Seção 1. Imprensa Nacional, p. 47–65. Brasília.
- [28] Funk, S., et al., 2016. Comparative analysis of Dengue and Zika outbreaks reveals differences by setting and virus. *PLOS Neglected Tropical Diseases* 10 (12).
- [29] Ghanem, R., Higdon, D., Owhadi, H., 2017. *Handbook of Uncertainty Quantification*. Springer.
- [30] Hennessey, M., Fischer, M., Staples, J. E., 2016. Zika virus spreads to new areas — Region of the Americas, May 2015–January 2016. *Morbidity and Mortality Weekly Report* 65 (3), 55–58.
- [31] Huppert, A., Katriel, G., 2013. Mathematical modelling and prediction in infectious disease epidemiology. *Clinical Microbiology and Infection* 19, 999–1005.
- [32] Ioos, S., Mallet, H.-P., Goffart, I. L., Gauthier, V., Cardoso, T., Herida, M., 2014. Current Zika virus epidemiology and recent epidemics. *Médecine et Maladies Infectieuses* 44 (7), 302–307.
- [33] Kucharski, A. J., Funk, S., Eggo, R. M., Mallet, H. P., Edmunds, W. J., Nilles, E. J., 2016. Transmission dynamics of Zika virus in island populations: a modelling analysis of the 2013–14 French Polynesia outbreak. *PLOS Neglected Tropical Diseases* 10 (5).
- [34] Lessler, J., et al., 2016. Assessing the global threat from Zika virus. *Science* 353 (6300).
- [35] Lessler, J. T., et al., 2016. Times to key events in Zika virus infection and implications for blood donation: a systematic review. *Bulletin of the World Health Organization* 94, 841–849.
- [36] Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics* 2, 164–168.
- [37] Lizarralde-Bejarano, D. P., Arboleda-Sánchez, S., Puerta-Yepes, M. E., 2017. Understanding epidemics from mathematical models: Details of

- the 2010 dengue epidemic in Bello (Antioquia, Colombia). *Applied Mathematical Modelling* 43, 566–578.
- [38] Malik, H. A. M., Mahesar, A. W., Abid, F., Waqas, A., Wahiddin, M. R., 2017. Two-mode network modeling and analysis of dengue epidemic behavior in Gombak, Malaysia. *Applied Mathematical Modelling* 43, 207–220.
- [39] Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11 (2), 431–441.
- [40] Martcheva, M., 2015. *An Introduction to Mathematical Epidemiology*, 1st Edition. Springer.
- [41] Meng, X., Li, Z., Wang, X., 2010. Dynamics of a novel nonlinear SIR model with double epidemic hypothesis and impulsive effects. *Nonlinear Dynamics* 59 (3), 503–513.
- [42] Moore, D. L., et al., 1975. Arthropod-borne viral infections of man in Nigeria, 1964–1970. *Annals of Tropical Medicine & Parasitology* 69 (1), 49–64.
- [43] Nelson, M. J., 1986. *Aedes Aegypti: Biology and Ecology*. Pan American Health Organization, Washington, D.C.
- [44] Oberkampf, W. L., Roy, C. J., 2010. *Verification and Validation in Scientific Computing*. Cambridge University Press.
- [45] Otero, M., Solari, H. G., Schweigmann, N., 2006. A stochastic population dynamics model for *Aedes Aegypti*: formulation and application to a city with temperate climate. *Bulletin of Mathematical Biology* 68 (8), 1945–1974.
- [46] O’Neill, P. D., 2002. A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods. *Mathematical Biosciences* 180, 103–114.
- [47] Pan American Health Organization, World Health Organization, October 2016. Zika – Epidemiological Report Brazil. Retrieved from <https://goo.gl/ph1eM4>, Accessed: 11/11/2016.

- [48] Pinto, C. M. A., Machado, J. A. T., 2014. Fractional dynamics of computer virus propagation. *Mathematical Problems in Engineering* 2014, 476502.
- [49] Pujol, J., 2007. The solution of nonlinear inverse problems and the Levenberg-Marquardt method. *Geophysics* 72 (4).
- [50] Reiner, R. C., et al., 2013. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970–2010. *Journal of The Royal Society Interface* 10 (81).
- [51] Secretaria de Vigilância em Saúde, 2016. Boletim Epidemiológico. monitoramento dos casos de dengue, febre de chikungunya e febre pelo vírus Zika até a Semana Epidemiológica 49, Volume 47. Retrieved from <https://goo.gl/33ulfg>, Accessed: 02/01/2017.
- [52] Secretaria de Vigilância em Saúde, 2016. Situação epidemiológica de Dengue, Chikungunya e Zika até a SE 46. Retrieved from <https://goo.gl/y6pezM>, Accessed: 02/01/2017.
- [53] Smith, D. L., et al., 2012. Ross, MacDonald, and a theory for the dynamics and control of mosquito - Transmitted pathogens. *PLOS Pathogens* 8 (4).
- [54] Smith, R. C., 2014. *Uncertainty Quantification: Theory, Implementation, and Applications*. SIAM.
- [55] Steyerberg, E. (Ed.), 2009. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer.
- [56] Sun, C., Hsieh, Y., 2010. Global analysis of an {SEIR} model with varying population size and vaccination. *Applied Mathematical Modelling* 34, 2685–2697.
- [57] Tarantola, A. (Ed.), 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.
- [58] Tognarelli, J., et al., 2016. A report on the outbreak of Zika virus on Easter Island, South Pacific, 2014. *Archives of Virology* 161 (3), 665–668.

- [59] U.S. Department of Health and Human Services, April 2016. CDC Concludes Zika Causes Microcephaly and Other Birth Defects. Retrieved from <https://goo.gl/wDRe0d>, Accessed: 02/01/2017.
- [60] Valentine, G., Marquez, L., Pammi, M., 2016. Zika virus-associated microcephaly and eye lesions in the newborn. *Journal of the Pediatric Infectious Diseases Society* 5 (3), 323–328.
- [61] Wang, Y., Cao, J., Alsaedi, A., Ahmade, B., 2017. Edge-based SEIR dynamics with or without infectious force in latent period on random networks. *Communications in Nonlinear Science and Numerical Simulation* 45, 35—54.
- [62] World Health Organization, September 2016. Zika Virus. Retrieved from <https://goo.gl/17fk63>, Accessed: 02/01/2017.
- [63] World Health Organization, April 2016. Zika virus, Microcephaly and Guillain-Barré Syndrome - Situation Report. Retrieved from <https://goo.gl/uiqbKA>, Accessed: 02/01/2017.
- [64] Xu, R., Ma, Z., 2010. Global stability of a delayed SEIRS epidemic model with saturation incidence rate. *Nonlinear Dynamics* 61 (1), 229–239.
- [65] Yaman, F., Yakhno, V. G., Potthast, R., 2013. A survey on inverse problems for applied sciences. *Mathematical Problems in Engineering* 2013, 976837.
- [66] Zanluca, C., et al., 2015. First report of autochthonous transmission of Zika virus in Brazil. *Memórias do Instituto Oswaldo Cruz* 110 (4), 569–572.
- [67] Zhang, T., Teng, Z., 2008. Pulse vaccination delayed {SEIRS} epidemic model with saturation incidence. *Applied Mathematical Modelling* 32, 1403–1416.