A mathematical analysis about Zika virus outbreak in Rio de Janeiro

Eber Dantas [1], Michel Tosin [2], João Peterson [3],
Vinicius Lopes [4] and Americo Cunha Jr [5]

NUMERICO - Nucleus of Modelling and Experimentation with Computers,
Universidade do Estado do Rio de Janeiro, Brazil

Abstract

This work deals with the calibration of SEIR epidemic model for Zika virus outbreak, in Rio de Janeiro, occurred in 2016, based on similar numerical analysis applied for disease's evolution in French Polynesia islands. Demographical increasing or decreasing is not considered. Vector is assumed to be infected for it's all lifespan. Model parameters fitting follows an empirical process, by comparison of obtained results with that in Polynesia case study, featuring as satisfactory the values that best reproduce model evolution in Rio de Janeiro scenario.

Keywords: nonlinear dynamics, mathematical biology, SEIR epidemic model, model calibration, Zika outbreak in Rio de Janeiro.

1 Introduction

Zika virus was first isolated in primates from Zika forest at Uganda about 1947 [1]. Evidences of the virus in humans were reported in Nigeria at 1968 [2]. It's first observation in south America occurred in Easter island, about 2014 [3]. First cases in Rio de Janeiro, Brazil, were reported in 2015 [4], quickly evolving to an outbreak in 2016. This scenario caused concern in medical community, health authorities and local population, specially due to a probable relation among Zika virus and other diseases, such as microcephaly [5, 6] and Guillain-Barré syndrome [7], and the realization of the 2016 Summer Olympics at Rio de Janeiro, which, in reason of the large flux of people, could spread Zika virus to several parts of the world.

[1] eber.paiva@uerj.br
[2] michel.tosin@uerj.br
[3] joao.peterson@uerj.br
[4] vinicius.lopes@uerj.br
[5] americo@ime.uerj.br

In this (near) epidemic scenario, developing strategies of control and prevention of the disease is a critical issue. In this sense, a mathematical model able to predict the number of infected people, during the virus outbreak, is an useful tool, that can be used to identify vulnerable aspects of the disease control strategies.

This paper deals with the study of Zika virus outbreak at Rio de Janeiro in 2016. For this purpose, a mathematical model that was used to describe the outbreak in French Polynesia [8] is adapted to Rio de Janeiro scenario. The nominal value of model parameters are estimated based on real data about the outbreak. Numerical simulation results are compared to the outbreak official numbers, in order to calibrate and validate the adapted model.

The rest of this paper is organized as follows. In section 2, the mathematical model is described, as well as its parameters. In section 3, the process of estimation of model parameters is discussed and numerical simulation results are reported, being followed by the comparison of numerical predictions and experimental data in section 4. Finally, in section 5, the main contributions of this work are emphasized and paths for future works are suggested.

2 Mathematical model

The transmission of Zika virus occurs, mainly, by *Aedes* genus of mosquitoes, but recently researches indicates that *Culex* genus is able to act as vector too [9]. Medical literature also relates sexual contact and blood transfusions as potential ways of transmission [10]. But, for purposes of modelling, in this work only the transmission by *Aedes* mosquitoes is considered.

To describe the Zika virus outbreak in Rio de Janeiro, this work employs an epidemic model of SEIR (susceptible, exposed, infected and recovered) type [11], adapted from the one used in French Polynesia [8]. In this model, the total population is distributed in four groups (susceptible, exposed, infected and recovered) according to the health condition of the individual at time t. The number of susceptible individuals at time t is denoted by $S(t)$, and comprehends healthy individuals which does not have any immunological resistance against Zika virus. The label of exposed individuals, denoted by $E(t)$, applies to the cases which are in an incubation period (latent period), i.e. a time gap while, despite of pathogen exposure, individuals are not infectious (asymptomatic period). The number of infected individuals, dubbed $I(t)$, means affected but not latent anymore. This group is, however, capable of transmitting the disease. The recovered or immune group,

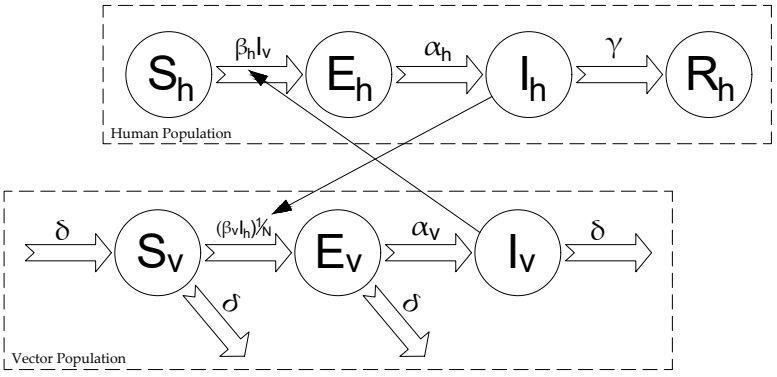


Figure 1: Schematic representation of SEIR epidemic model used to describe Zika virus outbreak.

represented by $R(t)$, contains the individuals that were exposed and are recovered.

The evolution of the diseased vector population obeys almost the same rules, in exception that there is no recovered group and it's lifetime is considered, since it is small compared to human lifetime. The individuals stream trough the model groups is depicted in Figure 2, both for affected and vector populations.

The stream (evolution) of individuals trough the groups follows the rules defined by the following set of differential equations

$$\frac{\mathrm{d}S_h}{\mathrm{d}t} = -\beta_h \, S_h \, I_v, \tag{1}$$

$$\frac{\mathrm{d}E_h}{\mathrm{d}t} = \beta_h \, S_h \, I_v - \alpha_h \, E_h, \tag{2}$$

$$\frac{\mathrm{d}I_h}{\mathrm{d}t} = \alpha_h \, E_h - \gamma \, I_h, \tag{3}$$

$$\frac{\mathrm{d}R_h}{\mathrm{d}t} = \gamma \, I_h, \tag{4}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \alpha_h \, E_h, \tag{5}$$

$$\frac{\mathrm{d}S_v}{\mathrm{d}t} = \delta - \beta_v\, S_v\, \frac{I_h}{N} - \delta\, S_v, \tag{6}$$

$$\frac{\mathrm{d}E_v}{\mathrm{d}t} = \beta_v\, S_v\, \frac{I_h}{N} - (\delta + \alpha_v)\, E_v, \tag{7}$$

$$\frac{\mathrm{d}I_v}{\mathrm{d}t} = \alpha_v\, E_v - \delta\, I_v, \tag{8}$$

where β represents the transmission rates and γ, the recovery rate, defined as the inverse of infection period; α means disease's incubation ratio, calculated as the inverse of incubation period and δ, the vector lifespan, and N represents human total population. The h subscripts refers to the human parameters, while v, the mosquitoes (vector).

3 Numerical experiments

3.1 Model parameters

The following values are used, at first, for the model parameters: $N = 16.5 \times 10^6$ humans, a recent estimation for Rio de Janeiro population [12]; extrinsic incubation ratio $\alpha_v = 1/14\,\mathrm{days}^{-1}$ [13]; intrinsic incubation ratio $\alpha_h = 1/7.5\,\mathrm{days}^{-1}$ [14]; human infectious ratio $\gamma_h = 1/7\,\mathrm{days}^{-1}$ [15]; mosquito lifespan ratio $\delta_v = 1/25\,\mathrm{days}^{-1}$ [16]; while the vector-to-human transmission rate $\beta_h = 400\,\mathrm{days}^{-1}$, and human-to-vector transmission rate $\beta_v = 40\,\mathrm{days}^{-1}$ were arbitrarily chosen through model experimentation, starting from the Tahiti values proposed for the French Polynesia outbreak [8].

Regarding the initial conditions, a fully susceptible population is considered, meaning $S_h(0) = N$ for humans and $S_v(0) = 1$ for the proportion of mosquitoes. The hypothesis made for the French-Polynesian model, that initial number of infected (humans or vector) was equal to their respective incubating group is adopted: $I_h(0) = E_h(0)$ and $I_v(0) = E_v(0)$ [8]. As such, the value of $I_v(0) = 0.014$ from the Tahiti outbreak is also used [8]. The initial number of infected humans considered is $I_h(0) = 412$, such number is a sum of sparse reported cases of Zika infection through some Rio de Janeiro cities in 2015 [17]. Of course, their sum is not a accurately representation of the quantity I_h at such time (beginning of 2016), but a arbitrary choice to condensate the 2015 information due to the lack of more precise data.

Lastly, $R(0)$ is arbitrarily chosen to be zero, signifying no recovered individual at the beginning of the outbreak, and $C(0) = I_h(0)$ because of the system dynamics implications.

3.2 Model predictions

In order to verify the representativeness of the proposed model, it is necessary to compare its predictions with field data. In this sense, real data about Zika outbreak, extracted from the epidemiological reports of Brazil's Ministry of Health [18] is used as reference. The report catalog data about the zika infection in the form of suspected accumulated cases since the thirteenth epidemiological week (03/27/2016 until 04/02/2016). The precise number of confirmed cases is actually unknown by the health authorities. Thus, this work took this value as 35%, as a first estimate.

Unfortunately, only six points of significant data are available through these reports so far (depicted on Table 1), and Zika cases along with the associated neurological and congenital syndrome related were not given notifiable conditions by the Brazil Ministry of Health until February 17th of 2016 [19], thus leaving 2015 data unreliable.

week	number of cases
13	25930
16	32312
18	38196
20	43516
21	46027
32	60176

Table 1: Number of suspected cases of infected humans in 2016 [18].

The comparison between the number of new infected humans per week, predicted by the model, and the real data is shown in Figure 2. Clearly, the reader can see that model response (blue curve) is not in good agreement with the real data (red circles).

Being 2015 the year of the official first confirmed case of Zika in Brazil and the probable epidemic peek that happened in mid-July of the same year [17], the lack of information about the spread of the disease during that year proposed a major difficulty on judging the consistency of the model. The choosing of parameters by ways of physical and biological context was not enough to establish the system on a basis of reasonable comparison with empirical data, and the lack of such data also hindered proper selection of

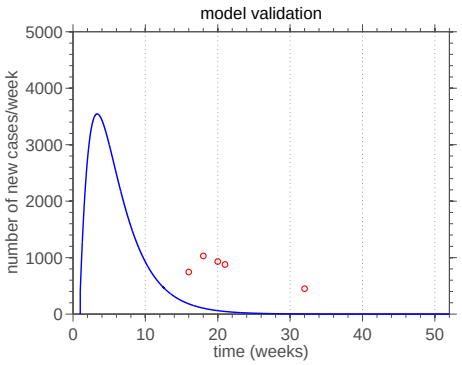


Figure 2: Comparison between the model prediction for the number of new infected humans per week (blue curve), and real data (red circles) obtained from [18].

parameters as β_h, β_v and initial conditions as $IV(0)$. Therefore, a calibration process is required to analyze the necessary parameters to fit the system.

4 Model calibration

To calibrate the model, a non-linear curve fitting procedure, based on a trust-region-reflective algorithm [20], is employed. The algorithm is performed first varying a single parameter, while the others remained fixed. This procedure allows to identify the parameters that most affect the model response. The parameters α_h and α_v show up to be unreliable, since they required very high initial guesses and usually would deform the quality of the fitting curve. Besides that, the model is very sensible to variations on β_h and β_v, being more affected by the values of β_h. The other parameters do not bring remarkable conclusions.

Next, a two varying parameters attempt is conducted through the method. Most combinations do not bring satisfactory results, to the extent that some pairs of parameters could not even be computed in the adopted tolerance levels, e.g (α_h, α_v), probably because of inefficient initial guesses or the low number of data points available. Furthermore, it sounds that some parameters control the quality of the fitting curve. Such control parameters are β_h and β_v and, unsurprisingly, the pair (β_h, β_v) proved to be the best one for fitting purposes. Hence, the pair (β_h, β_v) is chosen to be focus of the process of model calibration.

The parameter N is held fixed in all attempts. A change in the initial

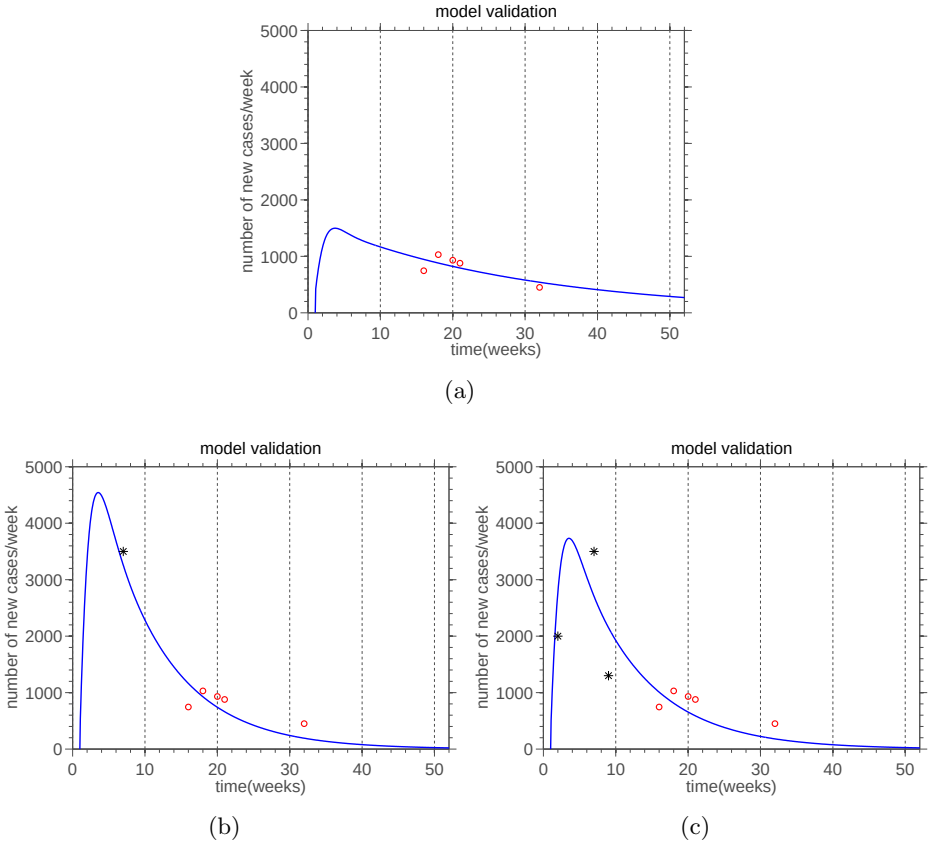
conditions is also performed, considering in particular I_h and I_v. The initial number of infected humans proved unresponsive, threatening significant changes on the curve only when varying his order of magnitude, what is beyond of our assumptions. But I_v is noteworthy. The value of $I_v = 0.008$ is chosen when testing the (β_h, β_v) pair due to compatible general results.

Some considerations on this process of investigation to find a good fitting curve regard the model curve peak position (local maxima). Brazil's 2016 Zika peak infection happened on February [19], but precise data for the state of Rio de Janeiro can only be estimated thus far, as discussed. This implicates that the six data points of Table 1 give very few information on the general shape the curve should have around the peek, because of their distance from such event. Additionally, the lack of reliable data about the infection in the beginning of year 2016 and end of 2015 allows a multitude of different shape-wise curves to fit the midyear data provided by the Ministry of Health. These problems manifested themselves while trying the parameter testing method, as a great number of low peak or wrong peak-positioned curves would fit the precise empirical numbers for the region (Table 1) but, of course, contradict the other information about the general distribution of the infection along the year.

The adopted solution for this lack of information is to perform (educated) guesses for the the peak region values, around the month of February. An imprecise number of 3500 new cases per week on February for Rio de Janeiro that circulated around brazilian media [21] is used on the iterative algorithm in search of a β_h and β_v that would better represent a peak like curve around such month. Figure 3.a summarizes the result of only using the data from Table 1, where the results of the estimation algorithm are $\beta_h = 0.0017\,\text{days}^{-1}$ and $\beta_v = 4.0599\,\text{days}^{-1}$. Figure 3.b represents the curve when a single unofficial data point (marked by a asterisk) of 3500 new cases per week is assumed on the epidemiological week 7. For this case the fitting algorithm returned in $\beta_h = 0.0055\,\text{days}^{-1}$ and $\beta_v = 0.6479\,\text{days}^{-1}$. Finally, Figure 3.c depicts the best fitted curved with respect to the empirical data at disposal, considering three extra unofficial data points: in addition to the 3500 new cases per week on the seventh week, a 2000 new cases/week on the second and 1300 weekly new cases on the third are included. The resulting estimated parameters are $\beta_h = 0.0045\,\text{days}^{-1}$ and $\beta_v = 0.8296\,\text{days}^{-1}$.

Figure 3.a indicates a general shape of the curve by only considering the data of Table 1: a early peak on the year 2016 and a decay like appearance around the data points. The introduction of an arbitrary point in Figure 3.b moved the peak closer to the desired week and defined better the shape. However, without comparison data around the maximum region, the

algorithm could not direct the parameters to values that would make 3500 be about the highest new infected per week around the seventh week. Using three arbitrary points as disposed on Figure 3.c, intentionally chosen to do the shape of the peak, the curve better distributed itself through the data and the magnitude stayed closer to the expected maximum value.



(a)

(b) (c)

Figure 3: Camparision between the model prediction (blue curve) and data from epidemiological bulletin (red circles). In the top only bulletin data is presented, while in the bottom 'ghost data' (black stars) are also used.

Although graphical results above are satisfactory, the estimated parameters are not good representations of reality. A $\beta_h = 0.0039$ translates to a vector only contacting a human being in a period of 256 days, which surpass extensively the lifespan of the mosquito. This may indicate this SEIR model is unrealistic to represent Rio de Janeiro outbreak, once it fits the observation data only when unreal parameters are used.

5 Final remarks

This work adapted a SEIR epidemic model, used to predict French Polynesia Zika spread, to describe the 2016 Zika virus outbreak in Rio de Janeiro, Brazil. Nominal quantities for the parameters were select through available experimental data in conjunction with arbitrary adopted values. A model calibration procedure, that uses a nonlinear curve fitting method, was employed to pick the best parameter values that would fit the model response into the empirical data regarding the evolution of the infection. Due to the low number of real data about the Zika outbreak, the curve fitting procedure was not effective at first. The effects of the lack of information have been solved by incorporating the reference data ansatz for the outbreak peak values. In this second step, the calibration procedure could get a good fit to describe the outbreak. Nevertheless, the adjust parameters obtained this way are unrealistic, which may suggest that this SEIR model is not a realistic representation of the Rio de Janeiro outbreak scenario. In future work the authors intend to construct a probabilistic model of uncertainties, to better describe the model parameters variability, and employ a Bayesian procedure for model calibration, in order to to verify if the model calibration can only be done via unrealistic data.

Acknowledgments

The authors are indebted to the Brazilian agencies CNPq, CAPES, and FAPERJ for the financial support given to this research.

References

[1] Dick, G. W. A., Kitchen, S. F. and Haddow, A. J., Zika virus(I). Isolations and Serological Specificity. Transactions of the Royal Society of Tropical Medicine and Hygiene, 46:509-520, 1952. http://dx.doi.org/10.1016/0035-9203(52)90042-4

[2] Moore, D. L. et al., Arthropod-borne viral infections of man in Nigeria, 19641970. Annals of Tropical Medicine & Parasitology, 69:49-64, 1975. http://dx.doi.org/10.1080/00034983.1975.11686983

[3] Daniele, F., Fabrizio, M. and Mauro, P., Zika Virus: Implications for Public Health. Infectious Diseases Society of America, 63: 227-233, 2016. http://dx.doi.org/10.1093/cid/ciw210

[4] Platonow, V., Estado do Rio Registra Primeiro Caso de Zika Virus. EBC Agência Brail, June 02, 2015. https://goo.gl/50fHse

[5] Schuler-Faccini, L. et al., Possible Association Between Zika Virus Infection and Microcephaly Brazil, 2015. MMWR Morb Mortal Wkly Rep, 65:5962, 2016. http://dx.doi.org/10.15585/mmwr.mm6503e2

[6] European Centre for Disease Prevention and Control. Rapid risk assessment: microcephaly in Brazil potentially linked to the Zika virus epidemic. Stockholm: ECDC, 2015. https://goo.gl/kpsFId

[7] Cao-Lormeau, Van-Mai et al., Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a Case-Control Study. The Lancet, 387:1531-1539, 2016. http://dx.doi.org/10.1016/S0140-6736(16)00562-6

[8] Kucharski, A. J., Funk, S., Eggo, R. M., Mallet, H., Edmunds, W. J., Nilles, E. J. Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 201314 French Polynesia Outbreak. PLOS Neglected Tropical Diseases, 2016. http://dx.doi.org/10.1371/journal.pntd.0004726

[9] Fiocruz Identifica Culex no Recife com Potencial para Transmitir o Virus Zika. Agência Fiocruz de Notícias, July 21, 2016. https://goo.gl/phQtFd

[10] Musso, D., Roche, C., Robin, E., Nhan, T., Teissier, A., Cao-Lormeau, V.M., Potential sexual transmission of Zika virus. Emerging Infectious Diseases, 21:359-361, 2015. http://dx.doi.org/10.3201/eid2102.141363

[11] Keeling, M. J., and Rohani, P., Modeling Infectious Diseases in Humans and Animals. Princeton University Press, 2007.

[12] IBGE. Estimativas da População Residente no Brasil e Unidades da Federação com Data de Referência em 1^o de Julho de 2015. Sep 09, 2015. https://goo.gl/NVbLfX

[13] Edward, B. H., Zika Virus Outside Africa. Emerging Infectious Diseases, 15, 2009. http://dx.doi.org/10.3201/eid1509.090442

[14] Lessler, J. T. et al., Times to Key Events in The Course of Zika Infection and Their Implications: a Systematic Review and Pooled Analysis [Submitted]. Bull World Health Organ E-pub. April, 2016. http://dx.doi.org/10.2471/BLT.16.174540

[15] Gregory, V., Lucila, M. and Mohan, P., Zika Virus-Associated Microcephaly and Eye Lesions in the Newborn. Journal of the Pediatric Infectious Diseases Society, 5:323-328, 2016. http://dx.doi.org/10.1093/jpids/piw037

[16] Zettel, C. and Kaufman, P., Yellow Fever Mosquito Aedes Aegypti (Linnaeus)(Insecta: Diptera: Culicidae). University of Florida IFAS. EENY-434, February, 2009. http://edis.ifas.ufl.edu/pdffiles/IN/IN79200.pdf

[17] N. R. Faria et al., Zika Virus in the Americas: Early Epidemiological and Genetic Gindings. Science, 352:345-349, 2016. http://dx.doi.org/10.1126/science.aaf5036

[18] Brazilian Ministry of Health, Monitoramento dos casos de dengue, febre de chikungunya e febre pelo vírus Zika. Boletim Epidemiológico, Vol.47, Number 18-33, 2016. https://goo.gl/cim1TX

[19] Pan American Health Organization / World Health Organization, Zika - Epidemiological Report, September 2016.

[20] Coleman, T.F., and Li, Y., An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds. SIAM Journal on Optimization, 6:418445, 1996. http://dx.doi.org/10.1137/0806023

[21] Prada, P., Redução de Casos de Zika Deve Diminuir Medo de Quem Vem para Olimpíadas. UOL Notcias, June, 2016. https://goo.gl/Yvr4Gq