Society for Mathematical Biology

CrossMark

Structural and Practical Identifiability Analysis of Zika Epidemiological Models

Necibe Tuncer[1] · **Maia Marctheva**[2] ·
Brian LaBarre[1] · **Sabrina Payoute**[1]

Abstract The Zika virus (ZIKV) epidemic has caused an ongoing threat to global health security and spurred new investigations of the virus. Use of epidemiological models for arbovirus diseases can be a powerful tool to assist in prevention and control of the emerging disease. In this article, we introduce six models of ZIKV, beginning with a general vector-borne model and gradually including different transmission routes of ZIKV. These epidemiological models use various combinations of disease transmission (vector and direct) and infectious classes (asymptomatic and pregnant), with addition to loss of immunity being included. The disease-induced death rate is omitted from the models. We test the structural and practical identifiability of the models to find whether unknown model parameters can uniquely be determined. The models were fit to obtain time-series data of cumulative incidences and pregnant infections from the Florida Department of Health Daily Zika Update Reports. The average relative estimation errors (AREs) were computed from the Monte Carlo simulations to further analyze the identifiability of the models. We show that direct transmission rates are not practically identifiable; however, fixed recovery rates improve identifia-

✉ Necibe Tuncer
  ntuncer@fau.edu

  Maia Marctheva
  maia@ufl.edu

  Brian LaBarre
  blabarre2013@fau.edu

  Sabrina Payoute
  spayoute@my.fau.edu

[1] Department of Mathematical Sciences, Florida Atlantic University, Science Building, Room 234 777 Glades Road, Boca Raton, FL 33431, USA

[2] Department of Mathematics, University of Florida, 358 Little Hall, Gainesville, FL 32611, USA

🖄 Springer

bility overall. We found ARE is low for each model (only slightly higher for those that account for a pregnant class) and help to confirm a reproduction number greater than one at the start of the Florida epidemic. Basic reproduction number, $\mathcal{R}_0$, is an epidemiologically important threshold value which gives the number of secondary cases generated by one infected individual in a totally susceptible population in duration of infectiousness. Elasticity of the reproduction numbers suggests that the mosquito-to-human ratio, mosquito life span and biting rate have the greatest potential for reducing the reproduction number of Zika, and therefore, corresponding control measures need to be focused on.

Keywords Zika virus · Structural and practical identifiability analysis · Parameter estimation · Arbovirus diseases

Mathematics Subject Classification 92D30 · 92D40

1 Introduction

Zika virus (ZIKV) was first isolated in Zika Forest of Uganda in a rhesus monkey in 1947 and then in humans in 1952 (Dick et al. 1952). Following the isolation of the first human case, several incidences occurred in a number of countries in Africa and Asia (Faye and Freire 2014) in the 1970s and 1980s. The first major outbreak of Zika was recorded in the Island of Yap with 185 suspected cases in 2007 (Kindhauser et al. 2016). Since 2015 the geographical distribution of the Zika virus has continued with reported cases from Brazil, Puerto Rico and most recently in Miami, Florida. It appears that the nature of the Zika virus infections has been changing as the virus moves from Africa to the Americas (Kindhauser et al. 2016). It was initially classified as obscure mosquito-borne infection causing mild illness across equatorial Africa and Asia, but, since 2007 it has been causing large outbreaks and has become global health emergency WHO http://www.who.int/emergencies/zika-virus/articles/one-year-outbreak/en/index1.html). Zika outbreaks have been linked to neurological disorders including Guillain–Barre syndrome and microcephaly in newborns born to mothers infected with Zika across the Pacific region and the Americas (WHO http://www.who.int/emergencies/zika-virus/articles/one-year-outbreak/en/index1.html; Perkins et al. 2016).

The main route of transmission for Zika virus is through the bite of an infected mosquito mainly from the Aedes species (Ae. aegypti and Ae. albopictus). These are the same vectors that transmit dengue, chikungunya and yellow fever. CDC estimates that these mosquitoes species can be found in about half of the area of the continental US, mostly the southern states but also in the northeast (CDC 2016). Local transmission of Zika in the continental US has been detected only in Florida and Texas.[1] Though very similar to dengue and chikungunya, Zika has other routes of transmission: through sexual contact, vertical transmission and blood infusion (see footnote 1). The virus can survive in semen longer than in blood and can sporadically be found in the vaginal

---

[1] http://www.cdc.gov/zika/geo/united-states.html.

fluids (Osuna and Lim 2016). A pregnant woman can pass Zika virus to her fetus during pregnancy. Vertical transmission of the Zika virus causes microcephaly and other severe fetal brain defects (Mlakar et al. 2016).

As of March 2017, the USA have had symptomatic cases of Zika, 215 of them locally acquired in Florida (see footnote 1). The local transmission of Zika in Florida began in June 2016 and was declared resolved by the governor in December 2016. Zika has been endemic in Puerto Rico. These incidences show that Zika is a significant public health problem in the USA as well as the remaining Americas.

Zika has been investigated through mathematical models. Kucharski et al. (2016) uses a standard vector–host ODE model to understand Zika transmission during the 2013–2014 outbreak of Zika in French Polynesia. The authors estimate a reproduction number of 2.6–4.8 with an estimated 11.5% of the cases reported. The article further estimates that 94% of the population was infected, mostly asymptomatically. Sexual transmission alongside vector-borne transmission for Zika was first modeled in Gao et al. (2016). Gao et al use data from Brazil, Colombia and El Salvador to estimate the reproduction number at 2.055. The large confidence interval for $\mathcal{R}_0$ (0.523–6.3) suggests that some of the parameters that comprise the reproduction number are not identifiable. The article further estimates that about 3% of the transmissions are sexual; however, the large CI suggests again that the parameters related to sexual transmission are not identifiable. We address this question explicitly and show that without data on sexual transmission, parameters related to sexual transmission are not identifiable. On the other hand, obtaining data for sexual transmission in the context of local transmission is very difficult since it is hard to distinguish which case is mosquito generated and which has resulted from sexual transmission (personal communication with Florida Department of Health). Sexual and vector-borne transmissions are also investigated in Baca-Carrasco and Velasco-Hernandez (2016) where the authors evaluate the impact of sexual transmission as well as importation of cases and find that sexual transmission impacts the magnitude of the outbreak, while migration generates outbreaks over time, possibly with lower magnitude. Chowell et al use the generalized Richards model to project new cases and estimate the burden of Zika, using data from Antioquia, Colombia (Chowell et al. 2016). One of the most serious impacts of Zika is on newborn babies to women infected with Zika. Perkins et al. (2016) uses mathematical models to project Zika virus infections in childbearing women in the Americas.

The goal of this article is twofold: (1) use a number of models to estimate $\mathcal{R}_0$ of the local cases in Florida of Zika outbreak in 2016 and (2) develop identifiable models of Zika, including models that explicitly account for pregnant women. In the next section, we introduce six models of Zika, starting from the very generic vector–host model and incorporating one by one distinct features of Zika, such as asymptomatic infections, sexual transmission and separate class for pregnant women. In Sect. 3, we discuss the structural identifiability of the models. In Sect. 4, we fit the models to the data, estimate $\mathcal{R}_0$ and discuss the practical identifiability of the models. In Sect. 5, we derive basic analytical results for the Zika models. Section 6 summarizes our conclusions. We have added all the MATLAB code used in this study to https://github.com/NecibeTuncer/ZikaODEModels.

Table 1 Definition of the variables in the between host models

Variables	Meaning
S_v	Number of susceptible vectors
I_v	Number of infected vectors
S	Number of susceptible individuals
S_p	Number of susceptible pregnant individuals
I	Number of infected individuals
I_p	Number of infected pregnant individuals
A	Number of asymptomatic individuals
R	Number of recovered individuals
N	Total human population
N_v	Total vector population

2 Epidemiological Models of the Zika Virus Infection

Zika is an epidemiologically complex disease. Though in many respects similar to other arboviral diseases, such as dengue and Chikungunya (Lanciotti et al. 2016), it also has many distinct features. In this section, we introduce a number of epidemiological models, starting from very simple and generic vector-borne model and including gradually the more distinctive features of Zika. Since Zika rarely leads to death (Petersen et al. 2016), we neglect the disease-induced death rate in all models. Nonetheless, we use standard incidence. Prior research has used outbreak models for the human population (Gao et al. 2016). We use endemic models as the Zika epidemic has continued for nearly 2 years; estimates of reproduction numbers suggest that the virus is endemic and WHO has put the disease on the list of continued threat diseases.[2] Model 1 is a general vector-borne model, much as the ones developed by Ross and McDonald for Malaria (Smith et al. 2012). Although Zika infection with a given strain (lineage) is believed to offer life-long protection (Dudley et al. 2016), infections with other strains may be possible. To capture that possibility, we include loss of immunity in the models. The dependent variables in the models are listed in Table 1. The parameter meanings of the various models are listed in Table 2.

Zika model with vector transmission only:

$$\text{Model 1 (M}_1) : \begin{cases} \dfrac{dS_v}{dt} = \Lambda_v - \dfrac{\beta I S_v}{N} - \mu_v S_v, \\[2mm] \dfrac{dI_v}{dt} = \dfrac{\beta I S_v}{N} - \mu_v I_v, \\[2mm] \dfrac{dS}{dt} = \Lambda - \dfrac{\beta_v I_v S}{N} - \mu S + \omega R, \\[2mm] \dfrac{dI}{dt} = \dfrac{\beta_v I_v S}{N} - (\mu + \gamma)I, \\[2mm] \dfrac{dR}{dt} = \gamma I - (\mu + \omega)R. \end{cases} \tag{1}$$

---

[2] https://newsline.com/zika-virus-no-longer-international-emergency-still-threat/.

The total human population $N(t) = S(t) + I(t) + R(t)$ satisfies the following differential equation,

$$N' = \Lambda - \mu N, \quad N(0) = S(0) + I(0) + R(0). \tag{2}$$

Similarly, the total mosquito population $N_v(t) = S_v(t) + I_v(t)$ can be determined from the following differential equation,

$$N_v' = \Lambda_v - \mu_v N_v, \quad N_v(0) = S_v(0) + I_v(0). \tag{3}$$

Zika infections are often asymptomatic with an estimated 80% of the cases being without symptoms (Petersen et al. 2016). In symptomatic individuals, clinical manifestation are mild. Symptoms in non-pregnant individuals last from several days to a week.[3] To account for the asymptomatic infections, we consider a version of the above model with asymptomatic class A:

Zika model with vector transmission and asymptomatic class:

$$\text{Model 2 (M}_2) : \begin{cases} \dfrac{dS_v}{dt} &= \Lambda_v - \dfrac{\beta(I + qA)S_v}{N} - \mu_v S_v, \\[2mm] \dfrac{dI_v}{dt} &= \dfrac{\beta(I + qA)S_v}{N} - \mu_v I_v, \\[2mm] \dfrac{dS}{dt} &= \Lambda - \dfrac{\beta_v I_v S}{N} - \mu S + \omega R, \\[2mm] \dfrac{dA}{dt} &= \dfrac{(1 - \phi)\beta_v I_v S}{N} - (\mu + \gamma_A)A, \\[2mm] \dfrac{dI}{dt} &= \dfrac{\phi \beta_v I_v S}{N} - (\mu + \gamma)I, \\[2mm] \dfrac{dR}{dt} &= \gamma I + \gamma_A A - (\mu + \omega)R, \end{cases} \tag{4}$$

where q is the reduction in infectivity of asymptomatic individuals and ϕ is the fraction of the new infections that are symptomatic. $N = S + A + I + R$. Models with asymptomatic class have also been investigated before (Chitnis et al. 2013).

As a vector-borne disease, Zika is transmitted predominantly by mosquitos. A distinctive feature of Zika is that it can be transmitted through sexual contact.[4] More recent data suggest that sexual transmissions may not be as rare as originally thought. To account for sexual transmission, we include a direct transmission term in Model 1:

---

[3] http://www.cdc.gov/zika/hc-providers/preparing-for-zika/clinicalevaluationdisease.html.

[4] http://www.cdc.gov/zika/transmission/sexual-transmission.html.

Zika model with vector and direct transmissions:

$$\text{Model 3 (M}_3): \begin{cases} \dfrac{dS_v}{dt} = \Lambda_v - \dfrac{\beta I S_v}{N} - \mu_v S_v, \\ \dfrac{dI_v}{dt} = \dfrac{\beta I S_v}{N} - \mu_v I_v, \\ \dfrac{dS}{dt} = \Lambda - \dfrac{\beta_v I_v S + \beta_d I S}{N} - \mu S + \omega R, \\ \dfrac{dI}{dt} = \dfrac{\beta_v I_v S + \beta_d S I}{N} - (\mu + \gamma) I, \\ \dfrac{dR}{dt} = \gamma I - (\mu + \omega) R. \end{cases} \tag{5}$$

Here, $N = S + I + R$. Models of vector-borne diseases with direct transmission are not new and have been considered before (Velasco-Hernandez 1994; Wei et al. 2008).

We include the asymptomatic infectious class in the model with vector and direct transmission.

Zika model with vector/direct transmission and asymptomatic class:

$$\text{Model 4 (M}_4): \begin{cases} \dfrac{dS_v}{dt} = \Lambda_v - \dfrac{\beta(I + qA)S_v}{N} - \mu_v S_v, \\ \dfrac{dI_v}{dt} = \dfrac{\beta(I + qA)S_v}{N} - \mu_v I_v, \\ \dfrac{dS}{dt} = \Lambda - \dfrac{\beta_v I_v S + \beta_d S(I + q_A A)}{N} - \mu S + \omega R, \\ \dfrac{dA}{dt} = \dfrac{(1 - \phi)\beta_v I_v S + (1 - \phi)\beta_d S(I + q_A A)}{N} - (\mu + \gamma_A) A, \\ \dfrac{dI}{dt} = \dfrac{\phi \beta_v I_v S + \phi \beta_d S(I + q_A A)}{N} - (\mu + \gamma) I, \\ \dfrac{dR}{dt} = \gamma I + \gamma_A A - (\mu + \omega) R, \end{cases} \tag{6}$$

where $N = S + I + A + R$.

Although Zika is a mild infection in individuals, it may be serious for pregnant women and their unborn children. Now it is well determined that Zika can cross the placenta and infect the fetus, particularly the brain, causing birth defects (Wu and Zuo 2016; Mlakar et al. 2016).[5] CDC reports the number of pregnant women infected with Zika,[6] allowing for modeling of this class separately. We include a model of Zika tracking separately pregnant women below. We do not include infected births because we assume that the contribution to new infections from the newly born infected babies is minimal.

---

[5] https://www.cdc.gov/zika/pregnancy/.

[6] http://www.cdc.gov/zika/geo/pregwomen-uscases.html.

Table 2 Definition of the parameters in models (1), (4), (5), (6), (7)

Parameter	Meaning
β	Transmission rate of infection from an infected human to a susceptible vector
β_v	Transmission rate of infection from an infected mosquito to a susceptible human
β_{v_p}	Transmission rate of infection from an infected mosquito to a susceptible pregnant human
β_d	Direct transmission rate of infection from an infected human to a susceptible human
β_{d_p}	Direct transmission rate of infection from an infected human to a susceptible pregnant human
Λ	Recruitment rate for human population
Λ_v	Recruitment rate for mosquito population
μ	Natural death rate for human population
μ_v	Natural death rate for mosquito population
ω	The rate at which recovered individuals lose immunity
ξ	The rate of becoming pregnant
$1/\xi_p$	Duration of pregnancy
γ	Recovery rate for infected human population
γ_p	Recovery rate for infected pregnant human population
γ_A	Recovery rate for asymptomatic human population
ϕ	Proportion of symptomatic human population
q	Reduction of infectivity to vectors of asymptomatic human population
q_A	Reduction of infectivity to humans of asymptomatic human population
m	Ratio of mosquito to human population at disease-free equilibrium

Zika model with vector/direct transmission and pregnant women:

$$
\text{Model 5 (M}_5\text{)} : \begin{cases}
\dfrac{\mathrm{d}S_v}{\mathrm{d}t} = \Lambda_v - \dfrac{\beta(I + I_p)S_v}{N} - \mu_v S_v, \\[2mm]
\dfrac{\mathrm{d}I_v}{\mathrm{d}t} = \dfrac{\beta(I + I_p)S_v}{N} - \mu_v I_v, \\[2mm]
\dfrac{\mathrm{d}S}{\mathrm{d}t} = \Lambda + \xi_p S_p - \dfrac{\beta_v I_v S + \beta_d(I + I_p)S}{N} - (\mu + \xi)S + \omega R, \\[2mm]
\dfrac{\mathrm{d}S_p}{\mathrm{d}t} = \xi S - \dfrac{\beta_{v_p} I_v S_p + \beta_{d_p} I S_p}{N} - (\mu + \xi_p)S_p, \\[2mm]
\dfrac{\mathrm{d}I}{\mathrm{d}t} = \dfrac{\beta_v I_v S + \beta_d(I + I_p)S}{N} - (\mu + \gamma)I, \\[2mm]
\dfrac{\mathrm{d}I_p}{\mathrm{d}t} = \dfrac{\beta_{v_p} I_v S_p + \beta_{d_p} I S_p}{N} - (\mu + \gamma_p)I_p, \\[2mm]
\dfrac{\mathrm{d}R}{\mathrm{d}t} = \gamma I + \gamma_p I_p - (\mu + \omega)R.
\end{cases}
\tag{7}
$$

Here, $N = S + S_p + I + I_p + R$. Research (Dudley et al. 2016) in monkeys suggests that pregnant individuals are infected for longer but the infection persists at lower level in the serum making them on average perhaps less invective to the mosquitos. To

Table 3 Summary of Zika models presented in the paper

Model	Brief Summary
M_1	Vector transmission, no direct transmission, no asymptomatic class, no pregnant class
M_2	Vector transmission, no direct transmission, asymptomatic class, no pregnant class
M_3	Vector transmission, direct transmission, no asymptomatic class, no pregnant class
M_4	Vector transmission, direct transmission, asymptomatic class, no pregnant class
M_5	Vector transmission, direct transmission, no asymptomatic class, pregnant class
M_6	Vector transmission, no direct transmission, no asymptomatic class, pregnant class

account for that difference, we assume distinct transmission rates for pregnant women to mosquitoes. However, for direct transmission we assume the same transmission rate based on the fact that direct transmission is more rare and the difference in the transmission rates will have a small impact on the disease dynamics.

Zika model with vector transmission and pregnant women class:

$$\text{Model 6 (M}_6\text{)} : \begin{cases} \dfrac{dS_v}{dt} &= \Lambda_v - \dfrac{\beta(I + I_p)S_v}{N} - \mu_v S_v, \\[2mm] \dfrac{dI_v}{dt} &= \dfrac{\beta(I + I_p)S_v}{N} - \mu_v I_v, \\[2mm] \dfrac{dS}{dt} &= \Lambda + \xi_p S_p - \dfrac{\beta_v I_v S}{N} - (\mu + \xi)S + \omega R, \\[2mm] \dfrac{dS_p}{dt} &= \xi S - \dfrac{\beta_{v_p} I_v S_p}{N} - (\mu + \xi_p)S_p, \\[2mm] \dfrac{dI}{dt} &= \dfrac{\beta_v I_v S}{N} - (\mu + \gamma)I, \\[2mm] \dfrac{dI_p}{dt} &= \dfrac{\beta_{v_p} I_v S_p}{N} - (\mu + \gamma_p)I_p, \\[2mm] \dfrac{dR}{dt} &= \gamma I + \gamma_p I_p - (\mu + \omega)R. \end{cases} \qquad (8)$$

Here, $N = S + S_p + I + I_p + R$. All the models and their characteristic differences are summarized in Table 3.

3 Structural Identifiability Analysis of the Epidemiological Models of Zika Virus Infection

As in many applications, in this study as well the parameters of the models cannot be directly measured by clinical studies, but can only be determined by indirect approaches such as parameter estimation methods using time-varying incidence reports provided by the health organizations. However, it is necessary to answer the fundamental question of whether the mathematical model is structured to identify its parameters from the given observations. For a well-posed parameter estimation prob-

lem, we need to know whether there exists a unique set of parameters that had produced the data. We first study the well posedness of the parameter estimation problem for the given observations such as cumulative incidences and infected pregnant women. A model is said to be structurally identifiable if for given large enough data sets, free of errors, it is theoretically possible to uniquely determine the parameter values from the observations that generated these data points. If infinite number of parameter sets leads to the same observations, then the model is called non-identifiable. If two or more (finite and isolated) parameter sets lead to the same observational output, the model is called locally identifiable.

If the model is not structurally identifiable, then we cannot estimate the "true values" of the parameters. As a first step in determining the parameters, we investigate the identifiability of the candidate models M_1-M_6. There are multiple ways to test for the identifiability of a model: Taylor's or generating series approaches, identifiability tableaus, differential algebra approach, direct methods, implicit function approach, profile likelihood, output sensitivities and differential geometry approaches (Chis et al. 2011; Miao et al. 2011; Raue et al. 2009; Stigter and Molenaar 2015; Villaverde et al. 2016; Meshkat et al. 2014).[7] Among these methods, there is not a single method that is applicable to all mathematical models. For comparison of these methods in terms of applicability, computational complexity and information provided, we refer the reader to Chis et al. (2011).

To set up the problem, without loss of generality we express the models M_1 through M_6 in the following compact form

$$x' = f(x(t), p) \quad x(0) = x_0, \tag{9}$$

where p denotes the parameters of the system, $x(t)$ denotes the state variables and x_0 is the initial values. The observations, cumulative number of incidences and pregnant women infected with Zika are given by the output function $g(x(t), p)$. The definition of the structural identifiability in the literature is given as in Miao et al. (2011).

Definition 3.1 A parameter set p is called *structurally globally (or uniquely) identifiable* if for every q in the parameter space, the equation

$$g(x(t), p) = g(x(t), q) \iff p = q .$$

That is, if $p \neq q$, then $g(x(t), p) \neq g(x(t), q)$ and hence the corresponding noise-free data are as well distinct. In other words, if any observation of the mathematical model can only be determined by a unique set of parameters, then the model is said to be globally (structurally) identifiable. The definition of local identifiability is given in Miao et al. (2011) as the following.

Definition 3.2 Let $\mathcal{N}(p)$ denote the neighborhood of the parameter p. The parameter set p is called *locally identifiable* if for every p there exists an open neighborhood $\mathcal{N}(p)$, such that for every $q \in \mathcal{N}(p)$ the equation

---

[7] http://biocyb1.cs.ucla.edu/combos/.

$$g(\boldsymbol{x}(t), \boldsymbol{p}) = g(\boldsymbol{x}(t), \boldsymbol{q}) \iff \boldsymbol{p} = \boldsymbol{q} \,.$$

Among all the methods to test for the structural identifiability, the differential algebra approach stands out because not only it distinguishes between local and global identifiability, but also it reveals the parameter correlations that lead to un-identifiability. So, if the model is not identifiable, using the parameter combinations obtained by the differential algebra approach, it is possible to scale the model to obtain a structurally identifiable model (Tuncer et al. 2016). Another advantage of the differential algebra approach is that there exists a software package "Differential Algebra for Identifiability of SYstems (DAISY)" implemented in REDUCE introduced by Bellu et al. (2007). However, the software does not perform well for large dynamical systems as pointed out in Chis et al. (2011). We ran DAISY for the model M_1 with mass action incidence term. DAISY could not finish the computations to produce an input–output equation where the output is the cumulative number of cases. DAISY reported computational errors due to the lack of memory. Computation of input–output equation does not depend on which parameters are fixed. We believe that the model M_1, with only 5 state variables, should not be considered a large dynamical system. In the era of connecting epidemiological models with time-series data, it is essential to develop computer packages (implemented in MATHEMATICA or MAPLE) to obtain structural identifiability analysis of epidemiological models using differential algebra approach.

In this study, we use the Identifiability Analysis package in MATHEMATICA to test for the local identifiability of the epidemiological models of Zika, M_1 through M_6. This implementation is based on a probabilistic numerical method of computing the rank of the identifiability (Jacobian) matrix (11) where the matrix parameters and initial state variables are specialized to random integers. We briefly describe the method here; for more detailed information, we refer to Karlson et al. (2012). Let $y(t) = g(\boldsymbol{x}(t), \boldsymbol{p})$ denote the observations which had generated the data and $\boldsymbol{p}$ be set of model parameters. The power series expansion of the observations $y(t)$ at the initial time $t = 0$ is given as

$$y(t) = y(0) + y'(0)t + y''(0)\frac{t^2}{2} + \ldots + y^{\nu}(0)\frac{t^{\nu}}{\nu!} + \mathcal{O}(t^{\nu+1}),$$

where

$$
\begin{aligned}
y(0) &= g(\boldsymbol{x}(0), \boldsymbol{p}) \\
y'(0) &= \frac{\partial g(\boldsymbol{x}(0), \boldsymbol{p})}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial t} = \mathcal{L}_f^1 g(\boldsymbol{x}(0), \boldsymbol{p}) \\
y''(0) &= \mathcal{L}_f^2 g(\boldsymbol{x}(0), \boldsymbol{p}) \\
&\quad\vdots \\
y^{\nu}(0) &= \mathcal{L}_f^{\nu} g(\boldsymbol{x}(0), \boldsymbol{p})
\end{aligned}
\tag{10}
$$

with $\mathcal{L}_f$ denoting the Lie-derivative along the vector field f. Setting $\mathcal{Y} = (y(0), y'(0), y''(0), \cdots, y^{\nu}(0))^T$, (10) can be written in compact form, $\mathcal{Y} =$

$\mathcal{Y}(\mathbf{x}(0), \mathbf{p})$. The inverse function theorem states that the equation $\mathcal{Y} = \mathcal{Y}(\mathbf{x}(0), \mathbf{p})$ can be uniquely solved for $\mathbf{x}(0)$ and $\mathbf{p}$ if and only if the Jacobian matrix

$$J(\mathbf{x}(0), \mathbf{p}) = \frac{\partial \mathcal{Y}(\mathbf{x}(0), \mathbf{p})}{\partial(\mathbf{x}(0), \mathbf{p})} \tag{11}$$

has full rank. The `Identifiability Analysis` package determines the rank of the matrix (11) by assigning random integers to model parameters and initial state variables. The method is based on two assumptions (observations). First, this method assumes that if a model is identifiable locally at time close to 0, then the identifiability carries over to all times. Second, probability of obtaining the actual rank of the matrix (11) by assigning random integers to initial state variables and model parameters is high. We are mainly interested in estimating epidemiologically important parameter values such as transmission and recovery rates in models M_1 through M_6. Thus, the parameter such as recruitment rate or natural death rates is fixed when estimating the other parameters. Full list of fixed parameters is given in Table 4. `Identifiability Analysis` states that the parameters of the models M_1 through M_4 are locally identifiable from the cumulative incidence observations, but parameters of the model M_5 cannot be obtained from the cumulative incidence observations only. For model M_5, we use two data sets, cumulative incidences and infected pregnant women, and then, the `Identifiability Analysis` states that the models M_5 and M_6 are locally identifiable. Structural identifiability analysis is necessary but not sufficient in concluding the identifiability of the parameter estimation problem. A model that is structurally identifiable may not be identifiable in practice when real data with noise are considered. On the other hand, the `IdentifiabilityAnalysis` tool is a numerical algorithm which relies on the determining the identifiability at the initial time and by determining the rank of the Jacobian matrix randomly. Hence, we further investigate the identifiability of the models M_1 through M_6 by Monte Carlo simulations.

4 Fitting the Epidemiological Models of Zika to Data

The observations (cumulative incidences and infected pregnant women) $\{y_i\}_{i=1}^{n}$ are obtained at discrete time points $t_1, t_2, \ldots t_n$ of the output function $g(\mathbf{x}(t), \mathbf{p})$. We define the statistical model by following the definition in Banks et al. (2014) as,

$$y_i = g(\mathbf{x}(t_i), \hat{\mathbf{p}}) + E_i, \tag{12}$$

where $\hat{\mathbf{p}}$ denotes the true parameters that generate the observations $\{y_i\}_{i=1}^{n}$ and E_i are the random variables that represent the observation or measurement error which cause the observations not fall exactly on the points $g(\mathbf{x}(t_i), \hat{\mathbf{p}})$ of the smooth path $g(\mathbf{x}(t), \hat{\mathbf{p}})$. In a general setting, the measurements errors are assumed to have the following form,

$$E_i = g(\mathbf{x}(t_i), \hat{\mathbf{p}})^{\xi} \epsilon_i, \tag{13}$$

where $\xi \geq 0$ and ϵ_i are independent and identically distributed with mean zero and constant variance σ_0^2. The random variables y_i have mean $\mathbb{E}(y_i) = g(\boldsymbol{x}(t_i), \hat{\boldsymbol{p}})$ and variances $Var(y_i) = g(\boldsymbol{x}(t_i), \hat{\boldsymbol{p}})^{2\xi}\sigma_0^2$. Varying ξ allows for varying error scales in the measurements. We use the *relative error* model, that is, $\xi = 1$ in (13), and use *ordinary least squares* in the parameter estimation problem. For the parameter estimation problem, we suppose that the Zika outbreak in Florida is exactly described by one of the deterministic models M_1 through M_6, that is, there is no modeling error and the expected value of the random variables ϵ_i is zero, hence $\mathbb{E}(\epsilon_i) = 0$.

Parameter estimation problem in the sense of least squares is to find the "true" parameter $\hat{\boldsymbol{p}}$ by solving the following optimization problem

$$\hat{\boldsymbol{p}} = \min_{\boldsymbol{p}} \sum_{i=1}^{n} (y_i - g(\boldsymbol{x}(t_i), \boldsymbol{p}))^2 \ . \tag{14}$$

4.1 Data and Parameter Values

As of November 23, 2016, there are total 4444 cases of Zika in the US, 182 of which locally acquired. All of the locally acquired cases have been acquired in Florida, where the vectors transmitting Zika, Aedes aegypti and Aedes albopictus can be found (see footnote 1). We obtained time-series data of cumulative incidences from the Florida Department of Health Daily Zika Update Reports.[8] The first locally acquired Zika case is observed on July 19, 2016 (see footnote 8). We use the time-series data of locally inquired Zika cases from July 19, 2016, to September 29, 2016. The Daily Zika Update Reports do not include locally acquired pregnant women cases; reports only consider the travel-related pregnant cases. Through email communications with Florida Department of Health, we obtained the infected Zika pregnant cases acquired in Florida. Based on the identifiability results of the previous section, we fit the models in the framework to the data with the following goals: (1) to select the best model representing the data in Florida; (2) to estimate the reproduction number of Zika in Florida; (3) to make short-term projections about the epidemic in Florida and its impact on pregnant women.

Florida's population currently is about 20 million people with life expectancy in the USA at 79 years. We take $\mu = 1/(79*365)$ days^{-1} and $\Lambda = 20000000*\mu$ people per day. Female mosquitoes, which bite and transmit the disease, live in captivity up to 30 days, but in the wild they often do not survive longer than 2 weeks.[9] We take $\mu_v = 1/10$ days^{-1} and we set $\Lambda_v = \mu_v$ to work with proportions of mosquitoes, rather than mosquito numbers. There are 2% pregnant women in the entire population on average at any time in the USA[10] which gives a value for $\xi = \dfrac{0.02}{0.98}\mu$.[11] Various sources

---

[8] http://www.floridahealth.gov/newsroom/all-articles.html.

[9] https://en.wikipedia.org/wiki/Mosquito.

[10] http://www.cdc.gov/nchs/data/databriefs/db136.htm.

[11] http://www.cdc.gov/zika/hc-providers/pregnant-woman.html.

Table 4 Fixed parameters in models (1), (4), (5), (6), (7), (8)

Parameter	Value	Range	Reference	Comment
Λ	$20,000,000\mu$	$1,950,000\mu$–$2,500,000\mu$	Population (2016)	Fixed all models
Λ_v	1/10	1/30–1/4	See footnote 9	Fixed all models
μ	$1/(78.8 \times 365)$	$1/(78.5 \times 365)$ –$1/(79 \times 365)$	a	Fixed all models
μ_v	1/10	1/30–1/4	See footnote 9	Fixed all models
ξ	7×10^{-7}	10^{-7}–10^{-6}	See footnote 10	Used for reference
$1/\xi_p$	280 days			Fixed in all models
γ	1/5	1/7–1/2	See footnote 12	Used for reference
γ_p	1/45	1/70–1/7	Driggers et al. (2016)	Used for reference
γ_A	1/5	1/7–1/2	See footnote 12	Used for reference
ϕ	0.2	0.2–0.25	Petersen et al. (2016)	Fixed in all models
q	1	0–1	Dudley et al. (2016)	Fixed in all models
q_A	1	0–2	Dudley et al. (2016)	Fixed in all models
ω	0	0–1/100	Dudley et al. (2016)	Fixed in all models

[a] http://www.cdc.gov/nchs/fastats/life-expectancy.htm

give information about the duration of Zika symptoms which last for 2–7 days.[12] That is also the typical duration of viremia (Dudley et al. 2016) in non-pregnant individuals. Duration of viremia in pregnant individuals can last 40–50 days and up to 10 weeks (Driggers et al. 2016).

Viremia studies in monkeys suggest that viremia levels of symptomatic and asymptomatic individuals are not that different than the asymptomatic so we assume that $q \approx 1$. We also surmise that in direct transmission asymptomatic individuals may be more infectious than symptomatic as they may not know that they are sick.

Table 4 shows a list of fixed parameters and their ranges.

4.2 Fitting the Models to the Data

We fit models (1), (4), (5), (6), the models with no pregnant classes, to the cumulative local Zika infections in Florida, starting from July 19 to September 29, 2016. For the optimization, we use MATLABs fminsearchbnd with both lower and upper bound on the fitted parameters. We fit repeatedly until the error does not decrease and the algorithm terminates because optimization tolerances have been reached. We observe that the recovery rates typically fit at the lower bound. We fit model (7), the model with pregnant classes, to cumulative number of local cases and cumulative number of local pregnant Zika cases. We use similar fitting approach as with other models; however, we only fit with lower bound (that we assume the upper bound for all parameters is infinity). In this case, the recovery rates also fit at the lower bound. To avoid fitting γ_p

---

[12] http://www.who.int/mediacentre/factsheets/zika/en/.

Table 5 Fitted parameters in models (1), (4), (5), (6), (7), (8) models M_2 and M_4 have asymptomatic classes, whereas the models M_1, M_3, M_5 and M_6 do not have asymptomatic classes

Parameter	M_1	M_2	M_3	M_4	M_5	M_6
β	2825.0	547.6	2825.0	565.0	2617.7	2617.7
β_v	100.4	519.3	100.4	501.9	91.3	91.4
β_d			9.4×10^{-7}	3.6×10^{-4}	6.3×10^{-9}	
β_{vp}					399.0	399.0
β_{dp}					2.5×10^{-5}	
γ	0.1	0.1	0.1	0.1	0.1	0.1
γ_A		0.1		0.1		
γ_p					0.02	0.02
$\mathcal{R}_{0i}$	1.46	1.47	1.46	1.46	1.24	1.24

at a very low value, we take a lower bound 1/50. Only epidemiological parameters such as transmission and recovery rates in all models are estimated, and the list of fitted parameters with their values is given in Table 5. The fitted parameters of the models are within the same magnitude when models are compared according to having an asymptomatic class or not.

4.3 Practical Identifiability Analysis of Zika Models

To further analyze the identifiability of the models, we perform Monte Carlo simulations which have been widely used for practical identifiability of ODE models (Miao et al. 2011). We generate 1000 synthetic data sets using the true parameter set $\hat{p}$ and adding noise at increasing levels. The true parameter set $\hat{p}$ for each model is obtained through fitting, and the results are given in Table 5. We outline the Monte Carlo simulations in the following steps.

(1.) Solve the epidemiological model (M_1 through M_6) numerically with the true parameters $\hat{p}$ and obtain the output vector $g(x(t), \hat{p})$ at the discrete data time points $\{t_i\}_{i=1}^n$.

(2.) Generate $M = 1000$ data sets from the statistical model (12) with a given measurement error. Data sets are drawn from a normal distribution whose mean is the output vector obtained in step (1.) and standard deviation is the $\sigma_0\%$ of the mean. That is, we set $\xi = 1$ in the error structure given in (12)

$$y_i = g(x(t_i), \hat{p}) + g(x(t_i), \hat{p})\epsilon_i \quad i = 1, 2, \ldots, n,$$

where $\mathbb{E}(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \sigma_0^2$. Hence, the random variables y_i have mean $\mathbb{E}(y_i) = g(x(t_i), \hat{p})$ and variances $\text{Var}(y_i) = g(x(t_i), \hat{p})^2\sigma_0^2$.

(3.) Fit the epidemiological model $\boldsymbol{x}' = f(\boldsymbol{x}, t, \boldsymbol{p})$, $\boldsymbol{x}(0) = \boldsymbol{x}_0$ to each of the M simulated data sets to estimate the parameter set $\boldsymbol{p}_j$ for $j = 1, 2, \ldots, M$. That is,

$$\boldsymbol{p}_j = \min_{\boldsymbol{p}} \sum_{i=1}^{n} (y_i - g(\boldsymbol{x}(t_i), \boldsymbol{p}))^2 , \quad j = 1, 2, \ldots, M .$$

(4.) Calculate the average relative estimation error for each parameter in the set $\boldsymbol{p}$ by Miao et al. (2011)

$$ARE(p^{(k)}) = 100\% \frac{1}{M} \sum_{j=1}^{M} \frac{|\hat{p}^{(k)} - p_j^{(k)}|}{\hat{p}^{(k)}},$$

where $p^{(k)}$ is the kth parameter in the set $\boldsymbol{p}$, $\hat{p}^{(k)}$ is the kth parameter in the true parameter set $\hat{\boldsymbol{p}}$ and $p_j^{(k)}$ is the kth parameter in the set $\boldsymbol{p}_j$.

(5.) Repeat steps 1 through 5 with increasing level of noise, that is, take $\sigma_0 = 0, 1, 5, 10, 20, 30\%$.

We perform Monte Carlo simulations by generating 1000 random data sets for each measurement error level and fitting each data set to the epidemiological model. We then compute the relative estimation errors (ARE) for each parameter in the epidemiological model which gives an insight about the practical identifiability of the parameters. When $\sigma_0 = 0$, that is, when there is no noise in the data, the ARE of the parameters of a structurally (globally) identifiable model should be 0 or very close to 0. As the noise level in the data increases, the ARE of the model parameters increases as well. If a parameter is not practically identifiable, then the ARE of that parameter will be significantly high even for a reasonable level of measurement error. Some of the parameters will be very sensitive to the noise in the data, and increasing the measurement errors will result in significantly high AREs, and then, we claim that the parameter is practically unidentifiable. To be specific, if the ARE of the parameter is higher than the measurement error σ_0, then we say that the parameter is practically unidentifiable.

The average relative errors computed from the first Monte Carlo simulations are presented in Table 6. As we see from Table 6, only the transmission rate from an infected mosquito to a susceptible human (β_v) is practically identifiable in models with vector transmission only, that is, in models M_1 and M_2. When direct transmission is added to models with vector transmission, we observe that the identifiability of β_v is lost, and none of the parameters in models M_3 and M_4 are identifiable. The average relative errors of the direct transmission rate (β_d) in models with direct transmission (M_3, M_4 and M_5) are significantly high compared with all other parameters. That is, we conclude that the direct transmission is not practically identifiable from time-series data of cumulative incidences. Hence, the uncertainties in the estimates of direct transmission rate are very high in models M_3, M_4 and M_5 (see Table 6). The uncertainties in the parameter estimation decrease if data related to other state variables are used in the fitting. Even though model M_5 includes direct transmission, since we use both

Table 6 Monte Carlo simulations: average relative estimation error (ARE) for parameters of the models M_1 through M_6

Model M_1				Model M_2			
σ_0 (%)	β (%)	β_v (%)	γ (%)	β (%)	β_v (%)	γ (%)	γ_A (%)
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	7.7	2.2	6.3	9.3	2.2	36.0	5.9
5	67.5	10.9	49.4	74.4	10.2	79.7	55.4
10	97.9	14.8	72.6	137.4	15.3	106.9	103.7
20	115.6	18.6	88.9	181.1	19.7	128.8	140.9
30	126.4	21.7	99.0	200.3	22.6	146.3	156.3

Model M_3					Model M_4				
σ_0 (%)	β (%)	β_v (%)	β_d (%)	γ (%)	β (%)	β_v (%)	β_d (%)	γ (%)	γ_A (%)
0	0.0	0.0	1.5e−5	0.0	0.0	0.0	0.0	0.0	0.0
1	4.7	3.7	4.4e5	4.8	17.1	4.5	2.9e3	58.0	11.7
5	27.2	21.2	3.9e6	29.0	73.2	17.4	4.8e4	58.0	61.5
10	38.4	32.1	6.4e6	42.7	131.5	21.7	5.6e4	85.3	115.2
20	43.1	44.2	9.2e6	53.4	155.7	34.2	1.7e5	135.7	138.5
30	45.4	52.6	1.1e7	59.3	200.9	36.4	1.8e5	192.7	150.7

Model M_5							
σ_0 (%)	β (%)	β_v (%)	β_d (%)	β_{vp} (%)	β_{dp} (%)	γ (%)	γ_p (%)
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	1.1	0.9	3.1e4	1.2	9.0e3	1.2e−4	0.008
5	5.7	4.5	2.7e4	5.8	4.0e4	2.7e−4	0.01
10	11.6	9.0	6.4e4	11.1	7.0e4	7.8e−4	0.01
20	25.1	18.1	5.2e4	20.8	8.0e4	0.001	0.03
30	42.0	26.9	2.9e5	30.8	9.8e4	0.002	0.03

Model M_6					
σ_0 (%)	β (%)	β_v (%)	β_{vp} (%)	γ (%)	γ_p (%)
0	0.0	0.0	0.0	0.0	0.0
1	1.1	0.9	0.8	1.4e−10	3.0e−9
5	5.7	4.5	4.2	3.4e−10	6.4e−9
10	11.7	9.1	8.7	5.1e−10	1.1e−8
20	25.1	18.1	18.2	8.0e−10	1.3e−8
30	42.1	26.9	28.7	10.0e−10	2.1e−8

Models M_1, M_2, M_3 and M_4 are fitted to cumulative local Zika cases. Models M_5 and M_6 are fitted to both cumulative local cases and pregnant infected cases

cumulative incidences and Zika infected pregnant cases while fitting this model, the AREs of the model M_5 parameters are less compared with model M_3. (Both models have direct and vector transmission; the models differ only at the pregnant classes.) The direct transmission rates β_d and β_{dp} in model M_5 are not identifiable. Comparing

Table 7 Monte Carlo simulations: average relative estimation error (ARE) for parameters of the models M_1 through M_6

M_1			M_2		M_3		
σ_0 (%)	β (%)	β_v (%)	β (%)	β_v (%)	β (%)	β_v (%)	β_d (%)
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	1.1	0.8	1.1	0.7	1.8	1.3	6.2
5	5.5	4.0	5.3	3.8	9.0	6.6	20.9
10	11.2	7.9	10.8	7.5	17.9	13.2	33.5
20	23.6	15.7	22.7	15.0	36.8	26.2	40.1
30	38.8	23.5	37.1	22.5	53.2	43.0	1.1e6

M_4				M_5				
σ_0 (%)	β (%)	β_v (%)	β_d (%)	β (%)	β_v (%)	β_d (%)	β_{vp} (%)	β_{dp} (%)
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	2.2	2.0	7.2e3	1.1	0.9	6.4e4	1.3	1.1e4
5	13.6	9.2	3.4e4	5.7	4.5	1.2e4	4.8	1.5e4
10	23.2	14.6	5.0e4	11.6	9.0	2.9e4	10.2	3.7e4
20	34.9	23.1	7.8e4	25.1	18.0	9.0e4	20.5	7.5e4
30	62.2	31.0	1.1e5	42.0	26.9	3.7e4	30.7	7.8e4

M_6			
σ_0 (%)	β (%)	β_v (%)	β_{vp} (%)
0	0.0	0.0	0.0
1	1.1	0.9	0.8
5	5.7	4.5	4.3
10	11.6	9.1	8.6
20	24.1	18.1	18.3
30	42.1	26.9	28.7

Models M_1, M_2, M_3 and M_4 are fitted to cumulative local Zika cases. Models M_5 and M_6 are fitted to both cumulative local cases and pregnant infected cases. The recovery rates (γ and γ_A) are fixed to the fitted values given in Table 5

all the models, we see that the model M_6 has the least ARE parameters. For model M_6, β has AREs up to 50% higher than σ_0, but in comparison with an unidentifiable parameter such as β_d in model M_5 which has ARE at 2.9×10^5, it is not that high. We conclude that ARE of β is in reasonable range and all other parameters of the model M_6 are practically identifiable from cumulative incidence and infected pregnant cases (see Table 6).

A standard approach to increasing the identifiability of parameters, when the parameters of the model are not identifiable, is to fix some parameters to previously known values, especially in the case that no other state variables are measurable. Since the recovery rate of Zika infections can be obtained from other sources, we perform second Monte Carlo simulations by fixing the recovery rates to the fitted values. The AREs

of the parameters computed for the second Monte Carlo experiment are presented in
Table 7. It is clear that fixing recovery rates decreased the AREs of all parameters in
fitting models M_1, M_2, M_3 and M_4, in the fittings where we could only fit cumulative
incidences. This process has also decreased the ARE of the unidentifiable parameter,
i.e., the direct transmission rates in all models. But the direct transmission rates remain
unidentifiable even after fixing recovery rates.

The ultimate goal in estimating the parameters of an epidemiological model is to
estimate the basic reproduction number of the infection. Since, we observe that fixing
recovery rates decreases the uncertainties in parameter estimation, we would like to
see whether this is still true in estimating the basic reproduction number. So, next we
perform the following Monte Carlo simulations.

(1.) Generate $M = 2000$ recovery rates (γ, γ_A) from a normal distribution whose
mean is the fitted value of the recovery rate, that is, $\mu_\gamma = 0.1$, and standard
deviation is $\sigma_\gamma = 0.1$, that is,

$$\gamma_j = \mathcal{N}(\mu_\gamma, \sigma_\gamma) = \mathcal{N}(0.1, 0.1) \quad j = 1, 2, \ldots, M.$$

Since recovery rate ranges between 2 and 15 days, we move to next step only if
$\gamma_j \geq 0.07$ for each j. If randomly chosen recovery rate is less than 0.07, then
another recovery rate is chosen from the normal distribution $\mathcal{N}(0.1, 0.1)$.
Similarly, generate $M = 2000$ recovery rates (γ_p) from a normal distribution
whose mean is the fitted value of the recovery rate, that is, $\mu_{\gamma_p} = 0.02$, and
standard deviation is $\sigma_{\gamma_p} = 0.005$, that is,

$$\gamma_{pj} = \mathcal{N}(\mu_{\gamma_p}, \sigma_{\gamma_p}) = \mathcal{N}(0.02, 0.005) \quad j = 1, 2, \ldots, M.$$

Randomly chosen recovery rate for pregnant women from a normal distribution
with $\mathcal{N}(0.02, 0.005)$ puts the recovery rate in the range of 40–70 days.
(2.) Fix recovery rate(s) in the epidemiological model

$$x' = f(x, t, p) \quad x(0) = x_0$$

to the randomly chosen value γ_j in step (1.) $j = 1, 2, \ldots, M$
(3.) For models M_1, M_2, M_3 and M_4 fit the epidemiological model $x' = f(x, t, p)$,
$x(0) = x_0$ to the observed cumulative Florida cases to estimate the rest of the
parameters in p_j. That is,

$$p_j = \min_p \sum_{i=1}^{n} (y_i - g(x(t_i), p))^2 \quad j = 1, 2, \ldots, M.$$

Estimate the parameters of the models M_5 and M_6 by fitting to cumulative Florida
cases and pregnant infections.
(4.) Compute the basic reproduction number, $\mathcal{R}_{0i}^j$ using γ_j and p_j for each j.

Table 8 Monte Carlo simulations: average relative estimation error (ARE) for basic reproduction numbers and the parameters of the models M_1-M_6 when recovery rates are randomly chosen from the range given in the literature

Model	$\mathcal{R}_{0i}$ (%)	β (%)	β_v (%)	β_d (%)	β_{vp} (%)	β_{dp} (%)
M_1	6.7	92.8	16.0	–	–	–
M_2	6.0	77.2	13.5	–	–	–
M_3	6.8	93.5	16.3	1.2e5	–	–
M_4	6.1	78.4	15.4	1.1e4	–	–
M_5	10.2	52.6	11.3	98.0	4.3	97.5
M_6	10.2	52.6	11.3		4.3	

(5.) Compute the average relative error in basic reproduction number,

$$ARE(\mathcal{R}_{0i}) = 100\% \frac{1}{M} \sum_{j=1}^{M} \frac{|\mathcal{R}_{0i}^{j} - \mathcal{R}_{0i}|}{\mathcal{R}_{0i}},$$

where $\mathcal{R}_{0i}$ is the fitted value obtained in Table 5.

Performing this third Monte Carlo simulations for model M_1, we obtain the following average relative errors (see Table 8),

$$ARE(\mathcal{R}_{01}) = 6.7\% \quad ARE(\beta) = 92.8\% \quad ARE(\beta_v) = 16.0\%.$$

That is, even though the average relative error of the transmission rate β is very high (the estimates for β ranges from 2000 to 20000), the average relative error in the computation of the basic reproduction number $\mathcal{R}_{01}$ is low. The computed basic reproduction number $\mathcal{R}_{01}$ ranges between 1.25 and 1.6. Based on this Monte Carlo simulation results, we conclude that by fixing recovery rate to any value in the range 2–15 days will result in large variations in the estimates of the transmission rate β from an infected individual to an infected mosquito, but the computation of the reproduction rate will not have huge variations.

4.4 Elasticity of the Reproduction Numbers

In this section, we investigate the elasticities of the reproduction numbers. The formulas for the reproduction numbers are computed in the next section. Elasticities of $\mathcal{R}_{0i}$ are shown in Fig. 1.

The elasticity of quantity Q with respect to parameter p is given by

$$\mathcal{E}_p^Q = \frac{\partial Q}{\partial p} \frac{p}{Q}.$$

The elasticities give the percentage change in the quantity Q in response to 1% increase in the parameter p. When $\mathcal{E}_p^Q > 0$ that means that Q increases with p; when $\mathcal{E}_p^Q < 0$

that means that Q decreases when p increases. The elasticities are relative to the size of the quantity and the parameter and allow us to compare the sensitivity of the quantity to different in size parameters.

Looking through the panels in Fig. 1, we can observe common trends. (1) All reproduction numbers are most sensitive to the parameters β (transmission from infected humans to susceptible mosquitoes), μ_v (death rate of mosquitos) and m (ratio of mosquitos to human). The elasticities $\mathcal{R}_{0i}$ to these parameters are approximately 1%, that is, 1% change in the parameter results in 1% change in $\mathcal{R}_{0i}$. This observation suggests that control measures targeted toward decreasing the mosquito/human ratio, decreasing the mosquito life span and decreasing the biting rate are most effective in reducing the reproduction number of Zika. (2) All reproduction numbers that depend on direct transmission β_d (β_{d_p}) show very small sensitivities to direct transmission parameters. For instance, the elasticity of $\mathcal{R}_{03}$ with respect to β_d is $6.4*10^{-6}$%, which is negligent. We surmise that the low sensitivity of the reproduction numbers to the direct transmission parameters is due to the very small values of these parameters. On the other hand, these low elasticities explain why even if we cannot identify the direct transmission parameters from the given data, the estimates of the reproduction are still quite reliable. The small sensitivities of the reproduction numbers with respect to the direct transmission parameters imply that control measures targeted at direct transmission have little population-level effect.

Figure 1 also suggests that the elasticities of $\mathcal{R}_{01}$ and $\mathcal{R}_{03}$ are quite similar; the elasticities of $\mathcal{R}_{02}$ and $\mathcal{R}_{04}$ are quite similar and the elasticities of $\mathcal{R}_{05}$ and $\mathcal{R}_{06}$ are quite similar. We continue by discussing more carefully the elasticities of $\mathcal{R}_{04}$ and $\mathcal{R}_{05}$.

Figure 1 panel (d) shows the elasticities of $\mathcal{R}_{04}$. One immediate observation that can be made from Figure 1 is that $\mathcal{R}_{04}$ is most sensitive to β, β_v and μ_v, m and q. These are all parameters that govern the vector-borne transmission of Zika. Reducing β or β_v with 1% will reduce $\mathcal{R}_{04}$ with approximately 0.9975%. The effect of vector mortality rate and the quotient of mosquitos to humans is similar. Increasing μ_v or reducing m with 1% will decrease $\mathcal{R}_{04}$ with 0.9975%. This suggests that control measures that may reduce the vector life span or reduce the vector to human ratio are some of the most efficient. On the other hand, the reproduction number depends very little on the direct transmission coefficient β_d. Reducing β_d with 1% will reduce $\mathcal{R}_{04}$ with only 0.00245%. One surprising observation is that the proportion of symptomatic/asymptomatic individuals ϕ has little impact on the reproduction number $\mathcal{R}_{04}$. That effect may be a result of the fact that $\gamma \approx \gamma_A$ causing the terms multiplied by ϕ to cancel out. Reciprocally, since $(1 - \phi) >> \phi$, the sensitivity to γ_A is much larger than the sensitivity to γ. This suggests that when potential treatment is available, treating asymptomatic individuals is at least as important as treating symptomatic individuals.

Figure 1 panel (e) shows the elasticities of $\mathcal{R}_{05}$. $\mathcal{R}_{05}$ is most sensitive to β, μ_v and m where 1% increase in the parameter will lead to 1% change in $\mathcal{R}_{05}$. As before, parameters related to the vector-borne transmission are most influential, suggesting control measures targeted toward reducing mosquito life span and the ratio of mosquitos to humans. Using personal protection to prevent bites is also important control strategy. The elasticities of $\mathcal{R}_{05}$ with respect to transmission parameters from vector to human are 0.29% with respect to β_{v_p} and 0.706% with respect β_v. The elasticities of $\mathcal{R}_{05}$ with respect to recovery rates are 0.29% with respect to γ_p and 0.705% with respect γ.
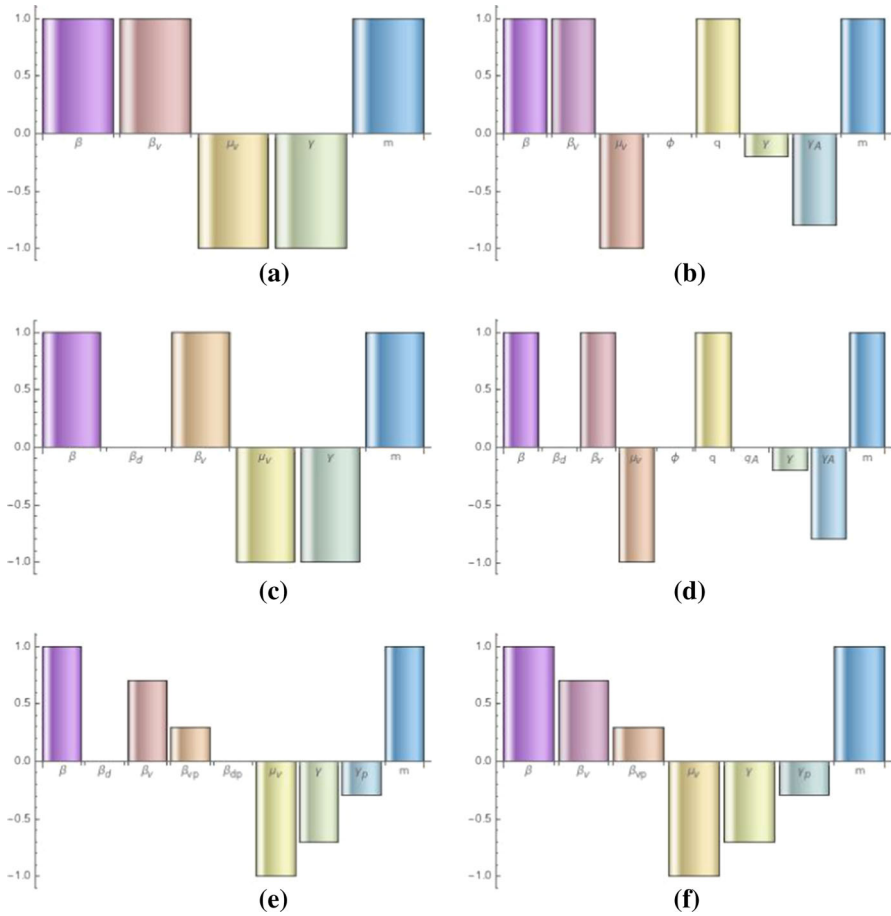
Fig. 1 (Color figure online) Elasticities of basic reproduction numbers of models M_1 through M_6. **a** Elasticities of $\mathcal{R}_{01}$, **b** elasticities of $\mathcal{R}_{02}$, **c** elasticities of $\mathcal{R}_{03}$, **d** elasticities of $\mathcal{R}_{04}$, **e** elasticities of $\mathcal{R}_{05}$, **f** elasticities of $\mathcal{R}_{06}$

Even though pregnant women are fewer than the general population, the transmission to them has a significant impact on the reproduction number. Moreover, the duration of infectiousness of pregnant women γ_p is a very notable factor influencing the reproduction number. This suggests that there is an urgent need of treatment strategies that will reduce the duration of infectiousness in pregnant women, and these will not only protect unborn children but also contribute to the reduction in population-level transmission of Zika. Regarding sexual transmission, the elasticity of $\mathcal{R}_{05}$ with respect to β_d is $1.9 * 10^{-10}\%$ and the elasticity of $\mathcal{R}_{05}$ with respect to β_{d_p} is $1.5 * 10^{-7}\%$. The direct transmission parameters impact on $\mathcal{R}_{05}$ is minimal, perhaps because although sexual transmission can occur, it is much more rare.[13]

---

[13] https://www.cdc.gov/zika/hc-providers/clinical-guidance/sexualtransmission.html.

5 Analysis of the Epidemiological Models of Zika Virus Infection

In this section, we provide basic analysis of the models in Sect. 2.

5.1 Analysis of Model (1)

Model M_1 has disease-free equilibrium $\epsilon_{M_1}^0 = (S_v^0, 0, S^0, 0, 0) = (\dfrac{\Lambda_v}{\mu_v}, 0, \dfrac{\Lambda}{\mu}, 0, 0)$ which is locally and globally asymptotically stable when $\mathcal{R}_{01} < 1$, where $\mathcal{R}_{01}$ is the reproduction number given by

$$\mathcal{R}_{01} = \frac{\beta \beta_v m}{\mu_v(\mu + \gamma)},$$

where $m = N_v/N$ is the ratio of mosquitos to human and $N_v = \dfrac{\Lambda_v}{\mu_v}$, $N = \dfrac{\Lambda}{\mu}$. This result is not hard to establish and we omit the proof. When $\mathcal{R}_{01} > 1$, the model M_1 has a unique locally stable endemic equilibrium $\epsilon_{M_1}^* = (S_v^*, I_v^*, S^*, I^*, R^*)$ given by,

$$
\begin{aligned}
I_v^* &= \frac{\beta N_v i^*}{(\beta i^* + \mu_v)} \\
S^* &= \frac{\Lambda}{\mu}\left(1 - \frac{(\mu + \omega + \gamma)i^*}{\mu + \omega}\right) \\
i^* &= \frac{\mu_v(\mathcal{R}_{01} - 1)}{\beta K} \qquad K = 1 + \frac{\beta_v m}{(\mu + \gamma)}\frac{\mu + \omega + \gamma}{\mu + \omega},
\end{aligned}
\tag{15}
$$

where $i^* = I^*/N$. Substituting I^* in S^* and simplifying show that $S^* > 0$. This equilibrium is also locally and globally stable (Yang et al. 2010), at least in the case $\omega = 0$.

5.2 Analysis of Model (4)

Model M_2 has a disease-free equilibrium $\epsilon_{M_2}^0 = (S_v^0, 0, S^0, 0, 0) = (\dfrac{\Lambda_v}{\mu_v}, 0, \dfrac{\Lambda}{\mu}, 0, 0, 0)$ which is locally and globally asymptotically stable when $\mathcal{R}_{02} < 1$, where $\mathcal{R}_{02}$ is the reproduction number given by

$$\mathcal{R}_{02} = \frac{\phi \beta \beta_v m}{\mu_v(\mu + \gamma)} + \frac{(1 - \phi)q\beta\beta_v m}{\mu_v(\mu + \gamma_A)}.$$

The reproduction number consists of sum of two terms, the first one giving the secondary infections of symptomatic individuals and the second one—of asymptomatic individuals. This result is also not hard to establish using the Jacobian approach. In the case $\mathcal{R}_{02} > 1$, the model M_2 has a unique endemic equilibrium

$\epsilon_{M_2}^* = (S_v^*, I_v^*, S^*, A^*, I^*, R^*)$. The components of this equilibrium or their fractions in the total human population s, i, a and r, respectively, are given by:

$$A^* = \frac{(1-\phi)(\mu+\gamma)}{(\mu+\gamma_A)\phi}I^* = \kappa I^*,$$

$$I_v^* = \frac{\beta K_1 N_v i^*}{(\beta K_1 i^* + \mu_v)},$$

$$S^* = \frac{\Lambda}{\mu}\left(1 - K_2 i^*\right), \tag{16}$$

$$i^* = \frac{\mu_v(\mathcal{R}_{02} - 1)}{\beta K_1 + \mu_v \mathcal{R}_{02} K_2},$$

$$R^* = \frac{\gamma I^* + \gamma_A A^*}{\mu+\omega},$$

where the constants K_1 and K_2 are defined as follows:

$$K_1 = 1 + \kappa q,$$

$$K_2 = \frac{\mu+\omega+\gamma}{\mu+\omega} + \kappa \frac{\mu+\omega+\gamma_A}{\mu+\omega}. \tag{17}$$

Substituting I^* and A^* in S^* and simplifying show that $S^* > 0$.

5.3 Analysis of Model (5)

Models with vector-borne and direct transmission have been investigated before (Wei et al. 2008). Model M_3 has disease-free equilibrium $\epsilon_{M_3}^0 = (S_v^0, 0, S^0, 0, 0) = (\frac{\Lambda_v}{\mu_v}, 0, \frac{\Lambda}{\mu}, 0, 0)$ which is locally and globally asymptotically stable when $\mathcal{R}_{03} < 1$, where

$$\mathcal{R}_{03} = \frac{\beta_d}{\mu+\gamma} + \frac{\beta\beta_v m}{\mu_v(\mu+\gamma)}.$$

When $\mathcal{R}_{03} > 1$, the model M_3 has a unique locally stable (Wei et al. 2008) endemic equilibrium $\epsilon_{M_3}^* = (S_v^*, I_v^*, S^*, I^*, R^*)$ given by,

$$I_v^* = \frac{\beta\Lambda_v i^*}{\mu_v(\beta i^* + \mu_v)},$$

$$S^* = \frac{\Lambda}{\mu} - \frac{(\mu+\omega+\gamma)I^*}{\mu+\omega}.$$

The equilibrium i^* can be determined uniquely from

$$\left(1 - \frac{(\mu+\omega+\gamma)i^*}{\mu+\omega}\right)\left(\frac{\beta_d}{\mu+\gamma} + \frac{\beta_v \beta \Lambda_v}{N\mu_v(\mu+\gamma)(\beta i^* + \mu_v)}\right) - 1 = 0.$$

Note that when $i^* = 0$, we have $\mathcal{R}_{03} - 1 > 0$. Once i^* is obtained, we have $I^* = \frac{\Lambda}{\mu} i^*$.

5.4 Analysis of Model (6)

Model M_4 has a disease-free equilibrium $\epsilon_{M_4}^0 = (S_v^0, 0, S^0, 0, 0, 0) = (\frac{\Lambda_v}{\mu_v}, 0, \frac{\Lambda}{\mu}, 0, 0, 0)$. The system has three transmitting infectious classes which leads to a full next-generation matrix. Alternatively, we use the Jacobian. The Jacobian has one eigenvalue $\lambda_1 = -(\mu + \omega)$. The remaining eigenvalues are eigenvalues of the characteristic polynomial:

$$
\begin{aligned}
0 = {}& (\mu_v + \lambda)(\lambda + \mu + \gamma)(\lambda + \mu + \gamma_A) - \phi\beta_d(\mu_v + \lambda)(\lambda + \mu + \gamma_A) \\
& - (1 - \phi)q_A\beta_d(\mu_v + \lambda)(\lambda + \mu + \gamma) - \beta m \phi \beta_v(\lambda + \mu + \gamma_A) \\
& - \beta m (1 - \phi)q\beta_v(\lambda + \mu + \gamma).
\end{aligned}
\tag{18}
$$

This polynomial has a positive leading term; whence, if the constant term $c_0 < 0$, the equation has a positive root and the disease-free equilibrium is unstable. The condition $c_0 < 0$ is satisfied if and only if the reproduction number $\mathcal{R}_{04} > 1$, where

$$
\mathcal{R}_{04} = \frac{\beta m \phi \beta_v}{(\mu + \gamma)\mu_v} + \frac{\beta m (1 - \phi)q\beta_v}{(\mu + \gamma_A)\mu_v} + \frac{\phi\beta_d}{\mu + \gamma} + \frac{(1 - \phi)q_A\beta_d}{\mu + \gamma_A}.
$$

That implies that if $\mathcal{R}_{04} > 1$ the disease-free equilibrium is unstable. If $\mathcal{R}_{04} < 1$, then we can rewrite the characteristic equation in the form $\mathcal{H}(\lambda) = 1$ where

$$
\begin{aligned}
\mathcal{H}(\lambda) = {}& \frac{\beta m \phi \beta_v}{(\mu + \gamma + \lambda)(\mu_v + \lambda)} + \frac{\beta m (1 - \phi)q\beta_v}{(\mu + \gamma_A + \lambda)(\mu_v + \lambda)} \\
& + \frac{\phi\beta_d}{\mu + \gamma + \lambda} + \frac{(1 - \phi)q_A\beta_d}{\mu + \gamma_A + \lambda}.
\end{aligned}
$$

Then, for λ with $\Re\lambda \geq 0$ we have $|\mathcal{H}(\lambda)| \leq \mathcal{H}(0) = \mathcal{R}_{04} < 1$. We conclude that if $\mathcal{R}_{04} < 1$ the disease-free equilibrium is locally asymptotically stable. In interpreting the reproduction number, we notice that the first term gives the secondary infections obtained by vector transmission generated by a single symptomatic individual; the second term gives the secondary infections obtained by vector transmission generated by a single asymptomatic individual; the third term gives the secondary infections obtained by direct transmission generated by a single symptomatic individual; and the last term gives the secondary infections obtained by direct transmission generated by a single asymptomatic individual.

When $\mathcal{R}_{04} > 1$, the model M_4 has a unique endemic equilibrium $\epsilon_{M_4}^* = (S_v^*, I_v^*, S^*, A^*, I^*, R^*)$. Dividing the last two equations, we obtain A^* in terms of I^*. From the first equation, we obtain I_v in terms of i^* where $i^* = I^*/N$ and we express S in terms of I^* from the equation for the total populations size:

$$A^* = \kappa I^* = \frac{(1-\phi)(\mu+\gamma)}{\phi(\mu+\gamma_A)} I^*,$$

$$I_v^* = \frac{\beta(1+q\kappa)i^* N_v}{\beta(1+q\kappa)i^* + \mu_v},$$

$$S^* = \frac{\Lambda}{\mu} - K_3 I^*,$$

(19)

where

$$K_3 = \frac{\omega+\mu+\gamma}{\omega+\mu} + \kappa\frac{\omega+\mu+\gamma_A}{\omega+\mu}.$$

I^* is then obtained from the following equation for i^*:

$$\frac{\phi}{\mu+\gamma}\left(\frac{\beta m \beta_v(1+q\kappa)}{\beta(1+q\kappa)i^* + \mu_v} + \beta_d(1+q_A\kappa)\right)\left(1-K_3 i^*\right) = 1. \qquad (20)$$

It is not hard to see that if $F(i^*)$ is the left-hand side of the above equation, $F(0) = \mathcal{R}_{04} > 1$. That says that the equation has a positive solution. On the other hand, it is clear that $F(i^*)$ is a decreasing function, and therefore, if a solution exists, it must be unique.

5.5 Analysis of Model (7)

We recast the model in the following form, which has the same dynamics:

Model 5 (M_5) :
$$\begin{cases} \dfrac{dS}{dt} = \Lambda - \dfrac{\beta_v I_v S + \beta_d(I+I_p)S}{N} - (\mu+\xi)S + \omega R + \xi_p S_p, \\[2mm] \dfrac{dS_p}{dt} = \xi S - \dfrac{\beta_{v_p} I_v S_p + \beta_{d_p} I S_p}{N} - (\mu+\xi_p)S_p, \\[2mm] \dfrac{dI}{dt} = \dfrac{\beta_v I_v S + \beta_d(I+I_p)S}{N} - (\mu+\gamma)I, \\[2mm] \dfrac{dI_p}{dt} = \dfrac{\beta_{v_p} I_v S_p + \beta_{d_p} I S_p}{N} - (\mu+\gamma_p)I_p, \\[2mm] \dfrac{dR}{dt} = \gamma I + \gamma_p I_p - (\mu+\omega)R, \\[2mm] \dfrac{dI_v}{dt} = \dfrac{\beta(I+I_p)}{N}(\dfrac{\Lambda_v}{\mu_v} - I_v) - \mu_v I_v. \end{cases}$$

(21)

This model also has a disease-free equilibrium which always exists

$$\epsilon_{M_5}^0 = (S_v^0, 0, S^0, S_p^0, 0, 0, 0) = \left(\frac{\Lambda_v}{\mu_v}, 0, \frac{\Lambda(\mu+\xi_p)}{\mu(\mu+\xi+\xi_p)}, \frac{\xi}{\mu+\xi_p}S^0, 0, 0, 0\right).$$

We will denote by s^0 and s_p^0 the fractions S^0/N and S_p^0/N given by

$$s^0 = \frac{\mu + \xi_p}{\mu + \xi + \xi_p} \qquad s_p^0 = \frac{\xi}{\mu + \xi + \xi_p}.$$

Computing the reproduction number via the next-generation approach does not result in a compact closed-form expression because the next-generation matrix is a full three-dimensional matrix. We compute the reproduction number using the Jacobian. Arranging the variables as (S, S_p, I, I_p, R, I_v), the Jacobian of the system has two eigenvalues, λ_1, λ_2 with negative real parts and $\lambda_3 = -(\mu + \omega)$. The remaining three eigenvalues satisfy the following characteristic equation:

$$\begin{aligned}
0 = {} & (\lambda + \mu_v)(\lambda + \mu + \gamma_p)(\lambda + \mu + \gamma) - \beta_d s^0 (\lambda + \mu_v)(\lambda + \mu + \gamma_p) \\
& - \beta_{d_p} s_p^0 \beta m \beta_v s^0 - \beta_{d_p} s_p^0 \beta_p s^0 (\lambda + \mu_v) \\
& - \beta_{v_p} s_p^0 \beta m (\lambda + \mu + \gamma) - \beta_v s^0 \beta m (\lambda + \mu + \gamma_p).
\end{aligned} \tag{22}$$

Since the leading term of this equation is positive, the equation has a positive eigenvalue if the constant term $c_0 < 0$. This inequality holds if the reproduction number $\mathcal{R}_{05} > 1$ where

$$\mathcal{R}_{05} = \frac{\beta_d s^0}{\mu + \gamma} + \frac{\beta_v s^0 \beta m}{\mu_v(\mu + \gamma)} + \frac{\beta_{v_p} s_p^0 \beta m}{\mu_v(\mu + \gamma_p)} + \frac{\beta_{d_p} s_p^0}{\mu + \gamma} \left(\frac{\beta_d^0 s^0}{\mu + \gamma_p} + \frac{\beta_v s^0 \beta m}{\mu_v(\mu + \gamma_p)} \right)$$

with $m = N_v/N$. On the other hand, one can show that if $\mathcal{R}_{05} < 1$, then the characteristic equation (22) does not have roots with nonnegative real part (Martcheva 2015). In interpreting $\mathcal{R}_{05}$, notice that the first two terms give secondary infections generated by one infected individual. The first term accounts for the secondary infections generated by direct transmission, while the second accounts for secondary infections generated through the vector transmission pathway. The third and fourth terms are the secondary infections of pregnant individuals generated by one pregnant individual. The third term accounts for secondary infections generated through vector transmission by pregnant women. The last term is most difficult to understand. We obtain this term because pregnant individuals do not infect directly other pregnant individuals. Thus, for one infected pregnant woman to generate a secondary infected pregnant woman through direct transmission one of two routes must be taken: (1) the pregnant woman infects a non-pregnant individual though direct transmission who in turn infects another pregnant individual; or (2) the pregnant woman infects a vector which in turn infects a non-pregnant individual who transmits through direct transmission to a pregnant woman. The number of secondary cases generated through the first scenario is given by the first term in the parenthesis, while the second scenario is given by the second term in the parenthesis. More specifically, one pregnant individual infects $\beta m/(\mu + \gamma_p)$ vectors during her life span as infectious, which in turn infect $\beta_v s^0/\mu_v$ individuals, who in turn infect $\beta_{d_p} s_p^0/(\mu + \gamma)$ pregnant women through direct transmission.
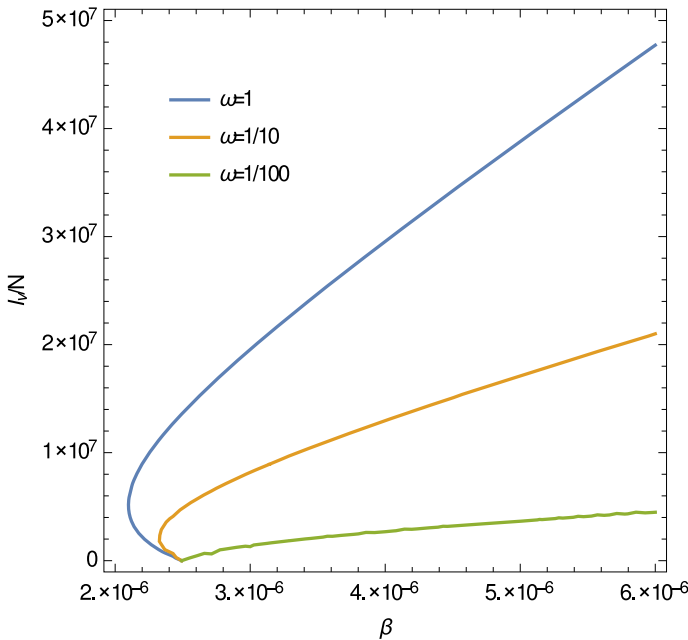
Fig. 2 (Color figure online) Backward bifurcation in model (7) in the case of no direct transmission. Parameter values are as follows: $\mu = 1/(75 * 365)$, $\Lambda = \mu/3$, $\Lambda_v = 100000000000$, $\mu_v = 1/10$, $\xi = 0.5$, $\gamma_p = 1/40$, $\gamma = 1/5$, ω is variable, $\beta_v = 0.0000005$, $\beta_{v_p} = 0.000000001$, $\xi_p = 0$

Theorem 1 *Assume $\xi_p = 0$ and $\mathcal{R}_{05} > 1$. Then, model (7) has an endemic equilibrium. If $\omega = 0$, then that equilibrium is unique.*

The proof of the theorem is delegated to Appendix.

Theorem 2 *Assume there is no direct transmission, that is, $\beta_d = \beta_{d_p} = 0$ and $\xi_p = 0$. Then, model (7) exhibits backward bifurcation if and only if*

$$
\left(\frac{\beta_v m}{\mu + \gamma} + \frac{\beta_{v_p} m}{\mu + \gamma_p} \frac{\xi}{\mu} \right) \frac{\omega}{\mu + \omega} \left(\frac{\gamma}{\mu + \gamma} \frac{\beta_v s^0}{\mu + \xi} + \frac{\gamma_p}{\mu + \gamma_p} \frac{\beta_{v_p} s_p^0}{\mu + \xi} \right)
$$

$$
- \left(\frac{\beta_v m}{\mu + \gamma} + \frac{\beta_{v_p} m}{\mu + \gamma_p} \frac{\xi}{\mu} \right) \frac{\beta_v s^0}{\mu + \xi} - \frac{\beta_{v_p} m}{\mu + \gamma_p} \frac{\beta_{v_p} s_p^0}{\mu} - \frac{\beta_{v_p} s_p^0}{\mu + \gamma_p} - \frac{\beta_v s^0}{\mu + \gamma} > 0,
$$

(23)

where $m = N_v/N$ and $N_v = \frac{\Lambda_v}{\mu_v}$, $N = \frac{\Lambda}{\mu}$.

The proof of the theorem is also delegated to Appendix. Figure 2 shows the backward bifurcation. There are several conclusions that can be drawn from the figure and condition (23). As Theorem 1 suggests if $\omega = 0$ there is no backward bifurcation or multiple equilibria. More careful examination of (23) reveals that $\gamma_p = 0$ or $\beta_{v_p} = 0$ there is no backward bifurcation. Figure 2 suggests that ω controls the depth of the backward bifurcation—the larger ω, the deeper the backward bifurcation. Furthermore, we seem to need a very large value of m to produce an example of the backward

bifurcation. We must note that if $\beta_d, \beta_{d_p} \neq 0$ backward bifurcation may still occur, just the necessary and sufficient condition for this to happen is more cumbersome to derive.

6 Discussion

Investigating the necessary public health measures needs mathematical models that are developed with available data in mind and have identifiable parameters. In this study, we develop six ODE models of Zika, encompassing various features of the disease, and we test them against two data sets: the cumulative number of *local* cases in Florida and cumulative number of *local* cases in pregnant women in Florida. Our main objectives were: (1) to compute the reproduction number of Zika in Florida, (2) to develop identifiable models of Zika and to estimate some of the critical parameters associated with Zika transmission.

We found that M_1–M_4 are structurally locally identifiable from data on cumulative number of *local* cases only, while M_5 and M_6 are not. We found that M_5 and M_6 are structurally locally identifiable from data on cumulative number of local cases and cumulative number of local cases of pregnant women in Florida (data that we obtained courtesy of the Florida Department of Public Health). We fit all the models to the relevant data sets using MATLAB's `fminsearchbnd` routine and obtained the fitted parameters. Models M_1 through M_4 give a value of $\mathcal{R}_{0i} = 1.46$. Models M_5 and M_6 give reproduction number $\mathcal{R}_{0i} = 1.24$, $i = 5, 6$. We conclude that the reproduction number depends on the data used but not on the model (given the same methodology of computation).

After fitting the models and estimating the parameters, we performed practical identifiability analysis using Monte Carlo simulations. We fit the transmission rates and the recovery rates. We find that for model M_1 through M_4 only the transmission rate from humans to mosquitos is practically identifiable, while the direct transmission rates have average relative errors (AREs) of 10^6 for all noise levels except the zero noise. Practical identifiability of parameters improves dramatically in M_5 and M_6 where we use two data sets. In model M_5 only the direct transmission rates are not practically identifiable with AREs of 10^5 for noise level above 20%. Model M_6 parameters are all practically identifiable.

In this study, we have not considered a model with all possible cases, that is, a model which includes vector and direct transmissions and asymptotic and pregnant classes. This was intentional. The goal of this paper is to study the identifiability analysis of Zika models. We have shown that adding asymptotic class definitely increases the ARE of the model parameters. Same is true for direct transmission. So, if we had a model with all possible cases, then we know for sure that the model would be unidentifiable with the current available data. If we had data available for the asymptotic and direct transmission cases, then we would have studied the model with all possible cases.

Next, we fixed the recovery rates at their fitted values and performed Monte Carlo simulation to understand the identifiability of the transmission rates. In this case, the identifiability of all transmission rates in all models has improved, but the direct transmission rates are still not identifiable in M_3 and M_4. In model M_5, the AREs of

the direct transmission coefficients are only $2 * 10^2$ at noise level of 30%. Since data on direct transmission are hard to obtain in a place of local transmission of Zika, such as Florida, inference about the direct transmission can best be made from model M_5 with fixed recovery rates.

Finally, we address the impact of how we fix the recovery rates in their plausible intervals on the value of the reproduction number and the estimates of the transmission rates. We perform Monte Carlo simulations by randomly choosing the recovery rates from their biologically realistic intervals, fitting the transmission rates and computing the reproduction number of each model and its AREs. Notably, the AREs of $\mathcal{R}_{0i}$ for $i = 1, \ldots, 4$ are around 6.5%, while AREs of $\mathcal{R}_{0i}$ for $i = 5, 6$ are 10%. This leads to a very small range for the reproduction number. For instance, $\mathcal{R}_{01}$ ranges between 1.25 and 1.6. There seems to be little doubt that the reproduction number of Zika in Florida at the start of the epidemic was above one.

Elasticities of the reproduction numbers suggest that the reproduction numbers are most sensitive to the transmission from humans to vectors, the ratio of vectors to humans and the vector mortality rate. This means that public health measures should focus in two main directions: (1) insecticide spraying to reduce mosquito life span and the mosquito population; (2) education of the public how to protect themselves against mosquito bites. Furthermore, the reproduction numbers are not sensitive at all to the direct transmission rate. That means that control measures targeting direct transmission have little population-level impact. The low sensitivity of the reproduction number to the direct transmission rates explains why its value has not been impacted by the presence of direct transmission in the model.

Appendix

Here, we prove Theorem 1.

Proof Let $\xi_p = 0$. We assume that $\gamma \geq \gamma_p$. The case $\gamma < \gamma_p$ is similar. We express all variables in terms of $i + i_p$ where $i = I/N$ and $i_p = I_p/N$

$$I_v = (i + i_p)\frac{\beta N_v}{\beta(i + i_p) + \mu_v} =: (i + i_p)f_v(i + i_p)N_v.$$

We solve for $s = S/N$:

$$s = \frac{\mu + \frac{\omega\gamma_p}{\mu+\omega}(i + i_p) + \frac{\omega}{\mu+\omega}(\gamma - \gamma_p)i}{\beta_v I_v/N + \beta_d(i + i_p) + \mu + \xi}.$$

Substituting in the equation for i and solving for i, we have

$$i = (i + i_p)\frac{(\beta_v f_v m + \beta_d)(\mu + \frac{\omega\gamma_p}{\mu+\omega}(i + i_p))}{(\mu + \gamma)(\beta_v I_v/N + \beta_d(i + i_p) + \mu + \xi)(1 - g(i + i_p))}$$
$$=: (i + i_p)G(i + i_p),$$

where

$$g(i + i_p) = \frac{\frac{\omega}{\mu+\omega}(\gamma - \gamma_p)(\beta_v I_v/N + \beta_d(i + i_p))}{(\beta_v I_v/N + \beta_d(i + i_p) + \mu + \xi)(\mu + \gamma)}.$$

It is not hard to show that $i > 0$. Replacing i from the above expression in s, we can obtain s in terms of $i + i_p$.

$$s = \frac{\mu + \frac{\omega\gamma_p}{\mu+\omega}(i + i_p) + \frac{\omega}{\mu+\omega}(\gamma - \gamma_p)(i + i_p)G(i + i_p)}{\beta_v I_v/N + \beta_d(i + i_p) + \mu + \xi}.$$

We express s_p in terms of $i + i_p$:

$$s_p = \frac{\xi s}{\beta_{v_p} I_v/N + \beta_{d_p}(i + i_p)G(i + i_p) + \mu}.$$

From here, we obtain i_p in terms of $i + i_p$:

$$i_p = (i + i_p)\frac{\beta_{v_p} f_v m + \beta_{d_p} G(i + i_p)}{\mu + \gamma_p} s_p =: (i + i_p)\mathcal{K}(i + i_p).$$

Substituting in $i + i_p$ and canceling $i + i_p$, we obtain the following equation for $i + i_p$:

$$1 = G(i + i_p) + \mathcal{K}(i + i_p).$$

It is not hard to see that if $\omega = 0$, the right-hand side of the above equality is a decreasing function of $i + i_p$ that does not depend of $\gamma - \gamma_p$. Hence, in the case $\omega = 0$, if the above equation has a solution, it must be unique. In the general case to see that the above equation has a solution, notice that $G(0) + \mathcal{K}(0) = \mathcal{R}_{05} > 1$. Next, we show that $G(1) + \mathcal{K}(1) < 1$. Some computation shows that

$$G(1) \leq \frac{(\mu + \gamma_p)(\beta_v f_v m + \beta_d)}{(\mu + \gamma_p)(\beta_v f_v m + \beta_d) + (\mu + \gamma)(\mu + \xi)}.$$

Furthermore,

$$s(1) \leq \frac{(\mu + \gamma_p)(\mu + \gamma)}{(\mu + \gamma_p)(\beta_v f_v m + \beta_d) + (\mu + \gamma)(\mu + \xi)}.$$

This gives the estimate

$$\mathcal{K}(1) \leq \frac{\xi s}{\mu + \gamma_p} = \frac{\xi(\mu + \gamma)}{(\mu + \gamma_p)(\beta_v f_v m + \beta_d) + (\mu + \gamma)(\mu + \xi)}.$$

Hence,

$$G(1) + \mathcal{K}(1) \leq \frac{(\mu + \gamma_p)(\beta_v f_v m + \beta_d) + \xi(\mu + \gamma)}{(\mu + \gamma_p)(\beta_v f_v m + \beta_d) + (\mu + \gamma)(\mu + \xi)} < 1.$$

We conclude that in the case $\mathcal{R}_{05} > 1$ there must be at least one solution. □

Now we prove Theorem 2.

Proof We apply Theorem 4.1 in Castillo-Chavez and Song (2004). We denote by $f_1, \ldots, f_6$ the right-hand sides of Eq. (7) with variables ordered the same way as the equations. We use β as a bifurcation parameter in place of ϕ in Theorem 4.1. It is not hard to show that if $\beta = \beta^*$ where β^* is the value that makes $\mathcal{R}_{05} = 1$, the Jacobian of the system has a unique zero eigenvalue and all other eigenvalues have negative real parts. We compute the right eigenvector to obtain:

$$
\begin{array}{lll}
w_6 = 1 & w_3 = \dfrac{\beta_v s^0}{\mu + \gamma} & w_4 = \dfrac{\beta_{v_p} s_p^0}{\mu + \gamma_p} \\[3mm]
w_5 = \dfrac{\gamma w_3 + \gamma_p w_4}{\mu = \omega} & w_1 = \dfrac{\omega w_5 - \beta_v s^0}{\mu + \xi} & w_2 = \dfrac{\xi w_1 - \beta_{v_p} s_p^0}{\mu}.
\end{array}
\tag{24}
$$

Next, we compute the left eigenvector which has the following nonzero components:

$$v_6 = 1 \qquad v_3 = \frac{\beta^* m}{\mu + \gamma} \qquad v_4 = \frac{\beta^* m}{\mu + \gamma_p}. \tag{25}$$

Since only v_2, v_3, v_6 are nonzero, we need the partial derivatives of f_3, f_4 and f_6. They are not hard to compute:

$$\frac{\partial^2 f_3}{\partial S \partial I_v} = \frac{\beta_v}{N} \qquad \frac{\partial^2 f_4}{\partial S_p \partial I_v} = \frac{\beta_{v_p}}{N} \qquad \frac{\partial^2 f_6}{\partial I \partial I_v} = -\frac{\beta^*}{N} \qquad \frac{\partial^2 f_6}{\partial I_p \partial I_v} = -\frac{\beta^*}{N}. \tag{26}$$

In addition, the only right-hand side in (21) that has nonzero derivative with respect to β is f_6. Hence, we have

$$\frac{\partial^2 f_6}{\partial I \partial \beta} = m \qquad \frac{\partial^2 f_6}{\partial I_p \partial \beta} = m. \tag{27}$$

From here, $b = v_6 m (w_3 + w_4) > 0$. Thus, the direction of the bifurcation is controlled by a and is a backward bifurcation if and only if $a > 0$ where

$$a = 2\beta^* / N w_6 \left(\frac{\beta_v m}{\mu + \gamma} w_1 + \frac{\beta_{v_p} m}{\mu + \gamma_p} w_2 - w_2 - w_3 \right),$$

which gives the condition in the statement of the theorem. This concludes the proof. □

References

Baca-Carrasco D, Velasco-Hernandez JX (2016) Sex, mosquitoes and epidemics: an evaluation of Zika disease dynamics. Bull Math Biol 78(11):2228–2242

Banks HT, Hu S, Thompson WC (2014) Modeling and inverse problems in the presence of uncertainty. CRC Press, Boca Raton

Bellu G, Saccomani MP, Audoly S, D'Angio L (2007) DAISY: a new software tool to test global identifiability of biological and physiological systems. Comput Methods Programs Biomed 88(1):52–61

Castillo-Chavez C, Song B (2004) Dynamical models of tuberculosis and their applications. Math Biosci Eng 1(2):361–404

CDC (2016) Estimated range of Aedes albopictus and Aedes aegypti in the United States. https://www.cdc.gov/zika/vector/range.html

Chis OT, Banga JR, Balsa-Canto E (2011) Structural identifiability of systems biology models: a critical comparison of methods. PLoS ONE 6(11):e27755

Chitnis N, Hyman JM, Manore CA (2013) Modelling vertical transmission in vector-borne diseases with applications to Rift Valley fever. J Biol Dyn 7(1):11–40

Chowell G, Hincapie-Palacio D, Ospina J et al (2016) Using phenomenological models to characterize transmissibility and forecast patterns and final burden of Zika epidemics. PLoS Curr Outbreaks. https://doi.org/10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583

Dick GW, Kitchen SF, Haddow AJ (1952) Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg 46:509520

Driggers RW, Ho CY, Korhonen EM et al (2016) Zika virus infection with prolonged maternal viremia and fetal brain abnormalities. N Engl J Med 374:2142

Dudley DM, Aliota MT et al (2016) A rhesus macaque model of Asian-lineage Zika virus infection. Nat Commun 7:12204

Faye O, Freire CCM et al (2014) Molecular evolution of Zika Virus during its emergence in the 20th century. PLOS Negl Trop Dis 8(1):e2636

Gao D, Lou Y et al (2016) Prevention and control of Zika as a Mosquito-Borne and sexually transmitted disease: a mathematical modeling analysis. Sci Rep 6, Article number: 28070

Karlson J, Anguelova M, Jirstrand M (2012) An efficient method for structural identifiability analysis of large dynamic systems. In: 16th IFAC Symposium on System Identification, The international Federation of Automatic Control, Brussels, Belgium, 11–13 July 2012

Kindhauser MK, Allen T, Frank V, Santhana R, Dye C (2016) Zika: the origin and spread of a mosquito-borne virus. Bull World Health Organ. https://doi.org/10.2471/BLT.16.171082

Kucharski AJ et al (2016) Transmission dynamics of Zika virus in island populations: a modelling analysis of the 2013–14 French Polynesia outbreak. PLoS Negl Trop Dis. https://doi.org/10.1371/journal.pntd.0004726

Lanciotti RS, Lambert AJ, Holodniy M, Saavedra S, del Carmen L, Signor C (2016) Phylogeny of Zika virus in western hemisphere, 2015. Emerg Infect Dis 22(5):933935

Martcheva M (2015) An introduction to mathematical epidemiology. Springer, New York

Meshkat N, Kuo CE, DiStefano J III (2014) On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: a novel web implementation. PLoS ONE 9(10):e110261

Miao H, Xia X, Perelson AS, Wu H (2011) On identifiability of nonlinear ODE models and applications in viral dynamics. SIAM Rev 53(1):3–39

Mlakar J, Korva M, Tul N et al (2016) Zika virus associated with microcephaly. N Engl J Med 374:951

Osuna CE, Lim S-Y (2016) Zika viral dynamics and shedding in rhesus and cynomolgus macaques. Nat Med 22(12):1448–1455

Perkins TA, Siraj AS, Ruktanonchai CW, Kraemer MUG, Tatem AJ (2016) Model-based projections of Zika virus infections in childbearing women in the Americas. Nat Microbiol 1(9):16126

Petersen EE, Staples JE, Meaney-Delman D, Fischer M, Ellington SR, Callaghan WM, Jamieson DJ (2016) Interim guidelines for pregnant women during a Zika virus outbreak United States, 2016, CDC, MMWR. http://www.cdc.gov/mmwr/volumes/65/wr/mm6502e1.htm

Population (2016) http://population2016.com/population-of-florida-in-2016.html

Raue A et al (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25(15):1923–1929

Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE (2012) Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. PLoS Pathogens 8(4):e1002588

Stigter JD, Molenaar J (2015) A fast algorithm to assess local structural identifiability. Automatica 58:118–124

Tuncer N, Gulbudak H, Cannataro VL, Martcheva M (2016) Structural and practical identifiability issues of immuno-epidemiological vector–host models with application to Rift Valley Fever. Bull Math Biol 78:1796–1827

Velasco-Hernandez JX (1994) A models for Chagas disease involving transmission by vectors and blood transfusions. Theor Popul Biol 46:1–31

Villaverde AF et al (2016) Structural identifiability of dynamic systems biology models. PLOS Comput Biol 12(10):e1005153

Wei HM, Li XZ, Martcheva M (2008) An epidemic model of a vector-borne disease with direct transmission and time delay. J Math Anal Appl 342(2):895–908

WHO, One year into the Zika outbreak: how an obscure disease became a global health emergency. http://www.who.int/emergencies/zika-virus/articles/one-year-outbreak/en/index1.html

Wu KY, Zuo GL et al (2016) Vertical transmission of Zika virus targeting the radial glial cells affects cortex development of offspring mice. Cell Res 26:645654. https://doi.org/10.1038/cr.2016.58

Yang H, Wei H, Li X (2010) Global stability of an epidemic model for vector-borne disease. J Syst Sci Complex 23:279–292