SCIENTIFIC REPORTS

OPEN

Understanding the survival of Zika virus in a vector interconnected sexual contact network

Tanvir Ferdousi[1], Lee W. Cohnstaedt[2], D. S. McVey[2] & Caterina M. Scoglio[1]

The recent outbreaks of the insect-vectored Zika virus have demonstrated its potential to be sexually transmitted, which complicates modeling and our understanding of disease dynamics. Autochthonous outbreaks in the US mainland may be a consequence of both modes of transmission, which affect the outbreak size, duration, and virus persistence. We propose a novel individual-based interconnected network model that incorporates both insect-vectored and sexual transmission of this pathogen. This model interconnects a homogeneous mosquito vector population with a heterogeneous human host contact network. The model incorporates the seasonal variation of mosquito abundance and characterizes host dynamics based on age group and gender in order to produce realistic projections. We use a sexual contact network which is generated on the basis of real world sexual behavior data. Our findings suggest that for a high relative transmissibility of asymptomatic hosts, Zika virus shows a high probability of sustaining in the human population for up to 3 months without the presence of mosquito vectors. Zika outbreaks are strongly affected by the large proportion of asymptomatic individuals and their relative transmissibility. The outbreak size is also affected by the time of the year when the pathogen is introduced. Although sexual transmission has a relatively low contribution in determining the epidemic size, it plays a role in sustaining the epidemic and creating potential endemic scenarios.

Zika virus (ZIKV) is predominantly vector-borne and the outbreaks are strongly dependent on the mosquito vector density which is influenced by the climatic conditions[1]. There have been substantial evidence[2] and reports[3] of sexual transmission of ZIKV. Although ZIKV epidemic models mostly focus on the mosquito-vectored transmission, multiple studies have quantified the contribution of sexual transmission[4–8]. Modeling analysis has helped to determine the relative importance of different transmission mechanisms. To enhance our understanding on how the combination of both vectored and sexual transmission can help sustain the pathogen during the vector free seasons (e.g. winter), we develop a novel epidemic model that interconnects a homogeneous mosquito vector population with a heterogeneous sexual contact network.

Zika virus is a positive-sense single-stranded RNA virus of the Flaviviridae family[9]. It is related to other flaviviruses such as dengue, yellow fever, West Nile, and Japanese encephalitis viruses. ZIKV infection symptoms include acute fever, maculopapular rash, arthralgia, and conjunctivitis. However, only about 20% of infected people suffer any significant symptoms[10]. A large proportion of asymptomatic cases poses a major problem in determining the outbreaks to initiate early response measures. ZIKV infection in pregnant women can result into congenital microcephaly, a severe birth defect[11]. The transmission of ZIKV occurs primarily by infected mosquito vectors such as *Aedes aegypti* and *Aedes albopictus*. However, it can also be spread via infected semen and blood[12]. The first autochthonous case of Zika in the mainland US was reported in July 2016, in Miami, Florida. As of April 2018, 5,676 cases have been reported in the US States[13]. In 2016, 224 autochthonous cases have been reported where the state of Florida itself is responsible for 97% of those cases, and the remaining 3% cases occurred in Texas[13]. Due to the complex nature and multiple pathways of pathogen transmission, several epidemic parameters are still unknown.

There have been numerous attempts in modeling the outbreaks of ZIKV. Several models only considered vectored transmission[1,14–16] and several others considered both sexual and vectored transmission routes[4,5]. Kuhlman *et al.* developed a hybrid agent-based model[16] to simulate Zika outbreaks in the central Miami region. While the work demonstrates interlinked host and vector populations, it does not address host-to-host sexual transmissions.

[1]Department of Electrical and Computer Engineering, Kansas State University, Manhattan, KS, 66506, USA. [2]Arthropod-Borne Animal Diseases Research Unit, Center for Grain and Animal Health Research, USDA, Manhattan, KS, 66502, USA. Correspondence and requests for materials should be addressed to T.F. (email: tanvirf@ksu.edu)

Vertical transmission has also been modeled in a work of Agusto et al.[17]. Olawoyin et al. analyzed the effects of multiple transmission pathways and concluded that, secondary transmision mechanisms increase the basic reproductive number and cause the outbreaks to occur sooner compared to the vector-only transmission[18]. The work of Gao et al. concluded that sexual transmission contributes about 3.044% of the overall transmission[4]. Moghadas et al. also found similar contribution from sexual transmission and they emphasized the importance of asymptomatic transmission[19]. Maxian et al. concluded that this contribution is too minor to sustain an outbreak, as they found that sexual transmission contributes about 4.8% to the basic reproductive number, R_0[6]. They also concluded that sexual transmission indirectly helps the vectored transmission by increasing the pool of infectious individuals. The relatively higher prevalence of infection in the female population compared to male population was analyzed by Pugliese et al.[20], who concluded that higher susceptibility of females compared to their male counterparts in the case of sexual transmission could be the reason. The work of Sasmal et al. estimated that sexual transmission contribution could be as high as 15.36% for sexual risk-stratified population[21]. Another study concludes that existing research could potentially underestimate the risk of sustained sexual transmission by an order of magnitude[22]. It has been found that Zika virus can sustain in semen for a long time after it disappears from blood[23]. Hence, even though it is clear that sexual transmission alone can't sustain an outbreak, we don't have clear information about the persistence of pathogen in a human host network without the help of vectors. Limited information on the reduction of transmissibility for both asymptomatically infected hosts and during the extended sexual transmission period adds to the complexity in understanding how the outcomes are affected.

In this paper, we develop an individual-based (i.e., node-based) network model, which is different from basic compartmental models as it features heterogeneous mixing. Each individual (i.e., a node) is connected to a specific number of individuals and this number can be very different. In the context of real world sexual networks, this is realistic. Using survey data[24] on sexual behavior, we develop a network generator tool that can produce networks with properties that closely match the given data. As the vector population is much larger than the host population and vectors don't transmit between them, it is unnecessary to use node-based models for vectors. We employ a homogeneous population model for the vectors. We interconnect these two populations (hosts and vectors) via the use of infection rate parameters where the infection of one population depends on the number of infected in the other population. There have been several works on epidemic spreading in interconnected networks[25]. However, due to the multi-host (one being the vector) scenario, our model is built differently from conventional interconnected networks. Our goal is to incorporate heterogeneity in the host population but reduce unnecessary computational overhead by interconnecting with a homogeneous vector population. Our model also differentiates between symptomatically and asymptomatically affected individuals by assigning them different states. We furthermore, incorporate the extended sexual transmissibility in our model. In the host network sub-model, each host node can be in either one of the seven states: $SEI_sI_aJ_sJ_aR$ (Susceptible, Exposed, Infectious (Symptomatic), Infectious (Asymptomatic), Convalescent (Symptomatic), Convalescent (Asymptomatic), and Recovered). The host population is also divided based on gender, sexual orientation, and age. In the vector population sub-model, the population is modeled into three compartments: $S_VE_VI_V$ (Susceptible, Exposed, and Infected). We only use birth-death demography for the vector model and the climatic variation is incorporated into the vector birth rate, which is one of the variable parameters. The simulation model is based on Gillespie's Stochastic Simulation Algorithm (SSA)[26]. However, we have a few variable rate parameters in our model, which prompted us to use the non-Markovian Gillespie Algorithm (n-MGA)[27] to simulate our model, which we implement using a modified version of the GEMF[28] tool.

The contributions of this paper are threefold: (i) we propose a novel interconnected model to evaluate host-to-host and host-to-vector-to-host pathogen transmission, (ii) we develop a sexual contact network generator based on aggregate data, and (iii) using our implementation of the non-Markovian Gillespie Algorithm (n-MGA) we examine survival of pathogen in the host network. We also perform sensitivity analyses on key model parameters to evaluate their role in disease outcomes. The interconnected model is presented in the Model Formulation section, three independent approaches in evaluating model outputs are presented in the Results section, and the Methods section contains the details on the simulation model.

Model Formulation

Interconnected population model. We propose an interconnected population model to investigate the spread of ZIKV within the human host and the mosquito vector population during an outbreak. A basic diagram of the model is shown in Fig. 1. It has two major components: the vector population marked by the cloud on top, and the host contact network marked by the solid circle nodes and the solid edges. The vector population is assumed to be in the vicinity of the host population such that each mosquito vector can bite any of the hosts (indicated by dashed lines).

Each host node can be in either one of the seven states: Susceptible (S), Exposed (E), Infectious (Symptomatic) (I_s), Infectious (Asymptomatic) (I_a), Convalescent (Symptomatic) (J_s), Convalescent (Asymptomatic) (J_a), and Recovered (R). The convalescent states were used to model the extended presence of ZIKV in semen as reported by previous works[23]. People in J_s or J_a states cannot transmit infections via vectors, they can only transmit sexually. The vector population is divided into three compartments: Susceptible (S_V), Exposed (E_V), and Infected (I_V). The unidirectional arrows indicate transitions from one state/compartment to another. The symbol adjacent to each arrow indicates the rate of that particular transition. The set of all these symbols constitute the model parameters, and they are listed in Table 1. This table also lists the nominal values and the range of values to be used for the simulation experiments. Some of these values are not readily available in the literature and we performed calibrations to determine acceptable ranges of values to conduct the experiments. We also performed sensitivity analyses with several critical parameters to evaluate model behavior. We have also used boolean condition functions marked by the symbol, $C_x(n)$ as shown in Fig. 1. The subscript 'x' indicates which condition a node n has to satisfy. If a condition is met, the condition function returns True (=1), otherwise it returns False (=0). For example, $C_{AM}(i)$
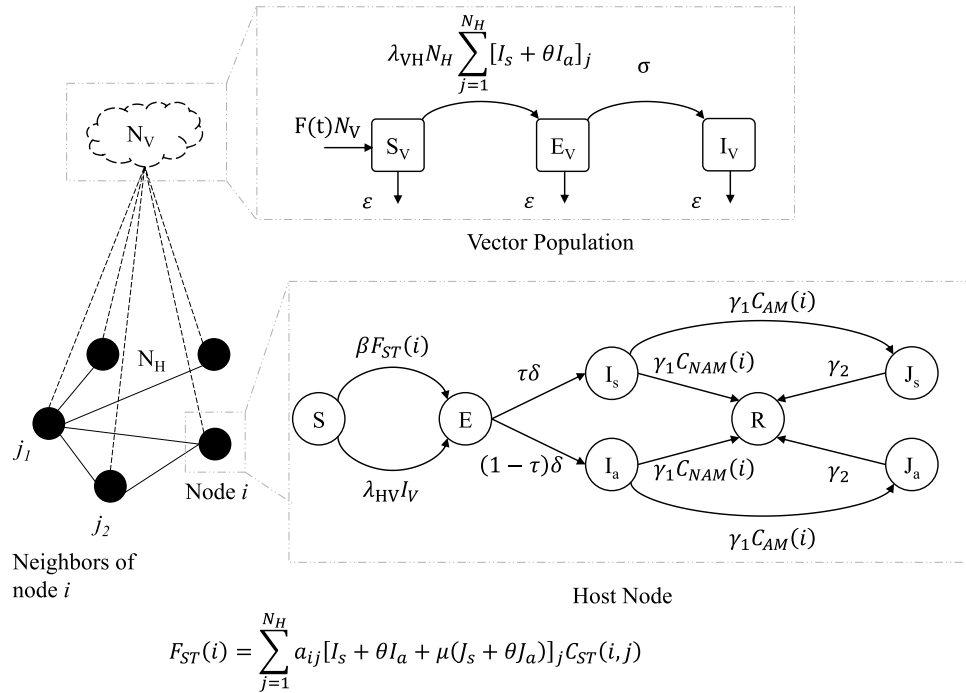
$$F_{ST}(i) = \sum_{j=1}^{N_H} a_{ij}[I_s + \theta I_a + \mu(J_s + \theta J_a)]_j C_{ST}(i, j)$$

Figure 1. Coupled population model for ZIKV. The black solid circles indicate host nodes (individuals) and the cloud shape above indicates the vector population. The solid edges connecting the nodes indicate the sexual contact network and the dashed edges indicate contacts of the vector population with the host nodes. A host node can be in any of the seven states ($SEI_sI_aJ_sJ_aR$) and the entire vector population is divided into three compartments ($S_VE_VI_V$).

will return 1 if node i is an adult male (AM), it will return 0 otherwise. The conditions that we have used are: $AM =$ Adult Male, and $NAM =$ Not Adult Male. It should be obvious that for any i, $C_{AM}(i)$ is the complement of $C_{NAM}(i)$. $F_{ST}(i)$ is the multiplier component function of the force of infection for sexual transmission from other nodes to node i. The symbols I_s, I_a, J_s, and J_a in the formulation of $F_{ST}(i)$ are host state indicators for neighbors (j) of node i. $C_{ST}(i, j)$ is the sexual transmission condition function. It only returns True (1) if i is an adult and j is an adult male. Here, j is considered the transmitter and i is consider the recipient of infection.

As indicated in Fig. 1, a susceptible host node can get infected ($S \rightarrow E$) via two mechanisms: (i) getting bitten by an infected mosquito (at the rate λ_{HV}) or (ii) having sexual contact with an infected or convalescent (I_s, I_a, J_s or, J_a) host (at the rate β). We use the adjacency matrix representation of a graph, where $\mathcal{A}$ is the adjacency matrix. An element $a_{ij} \in \mathcal{A}$ indicates the connection/link between node i and node j. If a link exists, then $a_{ij} = 1$, otherwise $a_{ij} = 0$. The sexual transmission rate, β was computed using the formula $\beta = 1 - (1 - \alpha)^n$, which was previously used in the work of Agusto[5]. Here, α is the probability of transmission per coital act with an infected partner. As data specific to Zika virus is not available, we assume a range of α between 0.001 and 0.1[29]. The exponent n is the average number of coital acts per unit time. Some estimated data indicate that people have sex an average of 60 times per year[24]. Hence, we estimate $n = 60/365 = 0.1644$ per day. The range of β for the range α is 0.000164 to 0.0172. Once a human host is infected, he/she can either develop symptoms or not show any sign of infection at all. Hence, there are two possible transitions from the exposed (E) state: (i) a proportion, τ of the people becomes symptomatically infected (I_s) and (ii) a proportion, $1 - \tau$ of the people becomes asymptomatically infected (I_a). The proportion of symptomatic ZIKV infections is claimed by previous works[10] to be about 20%. However, one recent work[30] claims that this proportion could be higher ranging from 27–50%. For our analysis, we take this into consideration as shown in Table 1. Once someone is infected, they will stay infected for an average duration of $1/\gamma_1$. However, once the virus disappears from blood, the adult male hosts will go through some extended period of infectiousness which is modeled by the convalescent states (J_s and J_a). An adult male in the convalescent state may only transmit sexually and would recover after an average duration of $1/\gamma_2$. The infectiousness of the 4 states (I_s, I_a, J_s, and J_a) are not considered the same. The relative transmissibility of the asymptomatic states is represented by θ. The relative transmissibility for the convalescent states is denoted by μ. Although, symptomatically infected people could be spreading less if they are seeking medical attention, we do not reduce the transmissibility for the symptomatic cases in this model. Historically, it has been seen that, possible negligence associated with ZIKV has contributed to sexual transmissions[31]. A healthy mosquito vector can get infected with some nonzero probability only if it bites an infectious human (I_s or I_a) (at the rate λ_{VH}). The time dependent per-capita mosquito recruitment rate is represented by $F(t)$ where, $F(t) = (1/12)A(t)$. Here, 1/12 is a fixed birth rate assumed to be the same as the mortality rate (ε). The mosquito abundance factor $A(t)$ is the seasonality parameter which incorporates seasonal variations into the model.

Symbol	Parameter Description	Range	Nominal Value	Reference
N_H	Total human host population	—	1,000	—
N_V	Total mosquito vector population	1,000–10,000	5,000	Calibrated
λ_{HV}	Vector to host pathogen transmission rate	—	$rTHV$	As defined
λ_{VH}	Host to vector pathogen transmission rate	—	$rTVH$	As defined
β	Host to host sexual transmission rate (day^{-1})	0.000164–0.0172	0.0087	Agusto et al.[5]
δ	Intrinsic incubation rate in hosts (day^{-1})	1/2–1/9	1/7	Zhang et al.[1]
τ	Proportion of symptomatically infected individuals	0.27–0.5	0.2	Mitchell et al.[30]
θ	Relative transmissibility of asymptomatic states	0.0–1.0	0.5	—
μ	Relative transmissibility of convalescent states	0.0–1.0	0.5	—
σ	Extrinsic incubation rate in vectors (day^{-1})	1/7–1/10	1/8	Zhang et al.[1]
γ_1	Host recovery rate from infectiousness (day^{-1})	1/3–1/7	1/7	Caminade et al.[43]
γ_2	Host recovery rate from convalescence (day^{-1})	—	1/30	Turmel et al.[23]
$A(t)$	Mosquito abundance factor	0–1	—	Reiskind et al.[34]
$F(t)$	Mosquito birth rate ($day^{-1}\ vector^{-1}$)	—	$(1/12) \times A(t)$	As defined
ε	Mosquito mortality rate (day^{-1})	1/4–1/35	1/12	Gao et al.[4]
T_{HV}	Vector to host pathogen transmission probability	0.214–0.8	0.634	Castro et al.[14]
T_{VH}	Host to vector pathogen transmission probability	0.6–0.95	0.770	Castro et al.[14]
r	Mosquito bite rate ($host^{-1}\ vector^{-1}\ day^{-1}$)	—	b/NH	As defined
b	Mosquito bite rate ($vector^{-1}\ day^{-1}$)	0.4–0.8	0.63	Castro et al.[14]
$C_{AM}(i)$	Condition: Host i is an adult male	{0,1}	—	—
$C_{NAM}(i)$	Condition: Host i is not an adult male	{0,1}	—	—
$C_{ST}(i,j)$	Condition: Host i is an adult and j is an adult male	{0,1}	—	—
T_{end}	Simulation termination time / Max run-time (day)	—	250	—
M_{start}	Simulation start month / Pathogen introduction month	1–12	4 and 11	Marini et al.[36]

Table 1. Epidemic model parameters and functions.

To generate a realistic sexual network, we develop a network generator tool based on the 'configuration model[32]'. Our tool takes as input, data of sexual behavior, age, gender, and produces representative random networks. Details are featured in the Methods section. An example network generated by our tool is given in Fig. 2. In our model, we also intend to incorporate the temporal variations of the climate in different seasons and the consequences of such variations on the mosquito vector population. To simplify this task, we use a vector abundance factor ($A(t)$) expressed as a fraction in the range [0,1]. This parameter was derived from data points used in the work of Monaghan et al.[33], which they originally extracted from the works of Reiskind et al.[34]. For our simulations, we use vector abundance data from Miami, Florida. The mosquito abundance factor over the course of 12 months is plotted in Fig. 3. The data originally had 12 sample points (monthly values). As the model requires the abundance factor on a daily basis, we use linear interpolations between data points of adjacent months.

Simulation tool. If the transition rate parameters are constant, the overall stochastic model is a Markov process with Poisson arrivals and exponential inter-event times. A stochastic algorithm such as the Gillespie SSA[26] can be used to perform simulations in most of these cases. The well-developed GEMFsim[28] has been used previously[35] to solve such stochastic spreading processes in the networks. However, our model requires a few changes before such simulations can be performed.

In our model, the transition rate parameters are not constant. The mosquito birth rate is a seasonally dependent parameter that varies according to the given vector abundance input data. The vector population compartments are changing with time during the outbreak simulation, which change the host infection rate $\lambda_{HV}I_V$. Similarly, the host infected populations are also changing with time which modify the vector infection rate $\lambda_{VH}N_H\sum_{j=1}^{N_H}[I_s + \theta I_a]_j$. These constitute a set of exogenously varying parameters (i.e., varying due to external forces/catalysts). As a consequence, the processes are no longer Poisson, rather these are called non-homogeneous Poisson. To cope with this issue, we use the non-Markovian generalized Gillespie Algorithm (nMGA)[27] which assumes general renewal processes (which is more generic) that allow exogenous variations of parameters. We modified the existing GEMFsim[28] accordingly, and added a vector compartmental model solver on top of it. The vector compartmental solver is effectively coupled to the modified network-based nMGA solver via the infection rate parameters ($\lambda_{HV}I_V$ and $\lambda_{VH}N_H\sum_{j=1}^{N_H}[I_s + \theta I_a]_j$), where a parameter in one population depends on some quantity in the other population.

Results

The simulation tool uses the parameters listed in Table 1, the initial conditions, the seasonal variation data, the population data, and the contact network representation. The initial condition is a single infected vector ($I_V(0) = 1$) while keeping the remaining vector population and the entire host population susceptible (S). We specify the pathogen introduction month (M_{start}) and the max run time (T_{end}). For example, if we set $M_{start} = 5$ and $T_{end} = 100$, the simulation will start from 1st of May and run up to the 2nd week of August. A simulation can
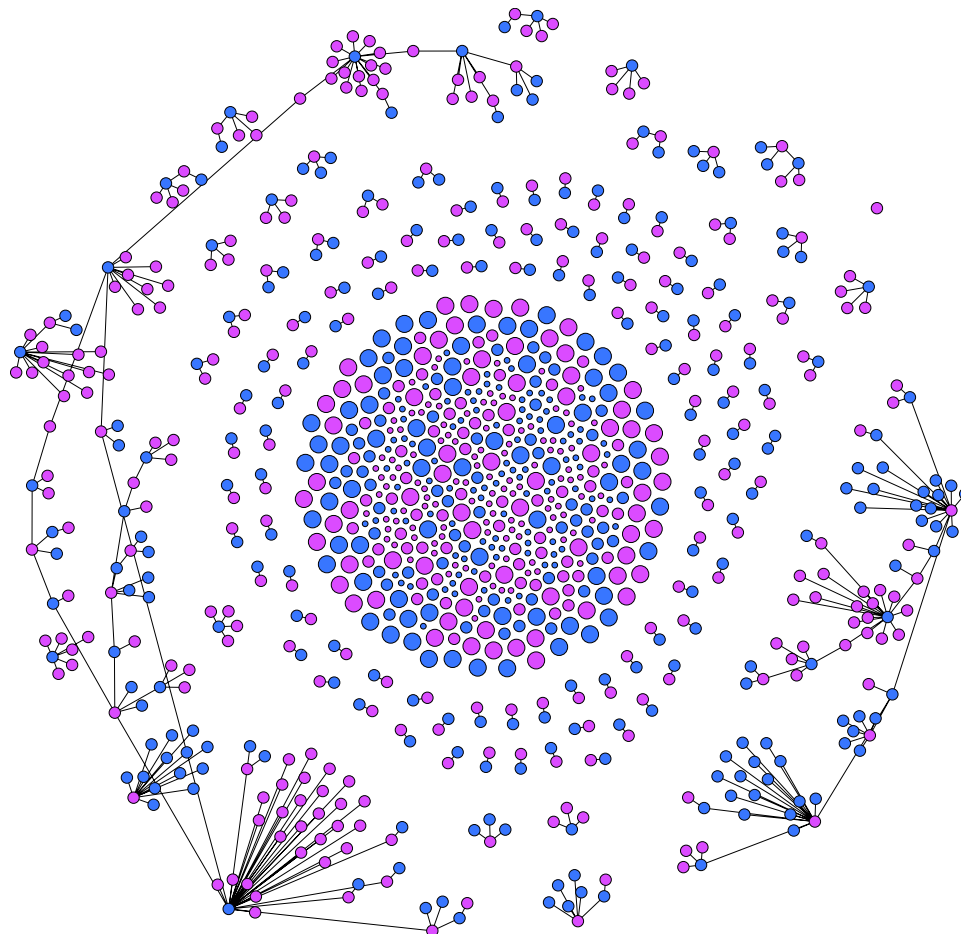
Figure 2. Generated sexual contact network based on sexual behavior[24]. The given network has $N = 1000$ nodes. The Methods section contains the details on how it is generated. The nodes are presented in two distinct colors and three distinct sizes. The males are marked as blue and the females are marked as pink. The three node sizes correspond to Children (small), Adults (medium), and Elderly (large) people. This particular instance of the network has an average node degree of 0.803 whereas the 64% adult population have an average degree of 1.26. This particular instance of network has 463 male and 537 female host nodes. The edge density of this network is 0.0008038. There are 600 connected components and the largest one consists of 117 nodes.
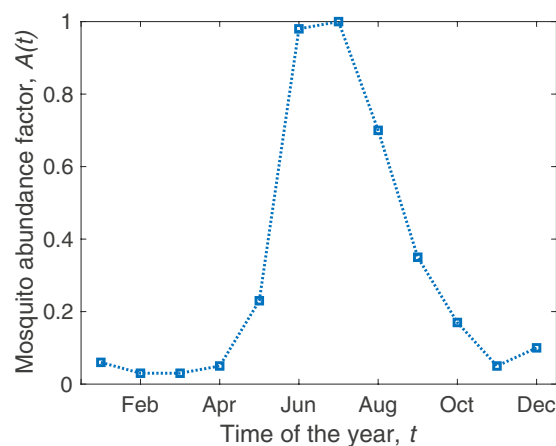
Figure 3. Variation in mosquito vector abundance with time. Due to changing temperature, rainfall, and other climatic factors, mosquito abundance varies throughout the year. The squares show the data points and the dotted lines are linear interpolations. The data represent Palm Beach, Florida where the vector abundance peaks during June-July. The plot was constructed from data observed in Palm Beach, FL over 27 four-week periods from 2006–2008[33,34].

terminate before T_{end} if the outbreak dies out. As we are running stochastic simulations, we smooth out our results by averaging over 500 repetitions. The population data contains the distribution of genders, sexual orientations and age groups, used to generate the contact network. For the survival analysis, we introduce the pathogen in November ($M_{start} = 11$) to indicate the start of winter. For the sensitivity analysis, we introduce the pathogen into the population in April ($M_{start} = 4$), which is the most likely scenario for the 2016 Miami-Dade outbreak[36].

Seasonal analysis.   First, we explore the effect of seasonal variations on the epidemic progression. We use the nominal parameter values defined in Table 1 and run the simulations for three distinct pathogen introduction months, M_{start}. Taking some cues from the seasonal patterns of Miami, FL shown in the Fig. 3, we choose to run independent scenarios starting on 1st of April ($M_{start} = 4$), 1st of August ($M_{start} = 8$), and 1st of October ($M_{start} = 10$). The averaged results of 500 simulations are shown in Fig. 4.

Without any intervention, we observe a large outbreak (more than 50% of the population affected) if pathogen is introduced on 1st of August. For the remaining two cases, the outbreak sizes are comparatively much smaller. This can be explained using the seasonal variations of Fig. 3. In Miami, the mosquito abundance is high during the beginning of August. This contributes to a boost in the infected vector population compared to the other two dates.

We also observe some interesting behaviors in the outbreak dynamics. Despite the vector abundance having a large positive slope during the month of April and a large negative slope during the month of August, Miami suffers a larger outbreak in the August introduction compared to the April introduction. Figure 4b shows that the infected vectors die out soon in the April outbreak although the upcoming summer (positive slope in abundance) sustains healthy vectors for a long time. In the case of the August introduction, the infected vectors rise rapidly during the first 50 days (Fig. 4d). Although the vector population declines rapidly after 50 days due to unsuitable climates of the upcoming winter, the initial surge in infected mosquitoes causes a large outbreak. The results indicate that *the suitability of the climate during pathogen introduction plays a major role in determining the size of the outbreak*. The impacts of climatic changes in the following months are minimal.

Survival analysis.   As a part of the survival analysis we will be observing several measured quantities: the epidemic/infection lengths, the pathogen survival periods, and the epidemic attack rates. Here, we define them first before discussing the results.

Host Infection Length (T_{HL}).   The host infection length is defined as,

$$T_{HL} = \text{The last time instant between } 0 \text{ and } T_{end} \text{ when there is at least one infected host left} \qquad (1)$$

It is the last day on the outbreak time-line an infected host can be found. For this property, the infected hosts who are asymptomatic or in the convalescent state are also considered infected. In this article, when we use the term "*epidemic length*", we imply host infection length T_{HL} unless otherwise mentioned.

Vector Infection Length (T_{VL}).   The vector infection length is defined as,

$$T_{VL} = \text{The last time instant between } 0 \text{ and } T_{end} \text{ when there is at least one infected vector left} \qquad (2)$$

It is the last day on the outbreak time-line an infected vector can be found. In typical outbreak situations, infected vectors die out before all the infected hosts recover.

Pathogen Survival Period, (T_{PS}).   The vector free pathogen survival period is defined as,

$$T_{PS} = T_{HL} - T_{VL} \qquad (3)$$

It measures how long pathogen can survive in the host population without the presence of vectors.

Epidemic Attack Rate (AR).   The epidemic attack rate (*AR*) is defined as,

$$AR = \frac{\text{Number of hosts who experience infection throughout the outbreak}}{\text{Total number of hosts, } N_H} \qquad (4)$$

The value of *AR* is in the range [0, 1]. We sometimes express this quantity as a percentage instead of a fraction. For example, an *AR* of 0.3 means 30% of the population were infected during the outbreak and the remaining 70% never experienced any infection.

To see how some of the above mentioned quantities relate to sexual transmission, we vary the sexual transmission rate, β in the range mentioned in Table 1. We find that varying β does not noticeably increase or decrease the epidemic length (Fig. 5a), which remains between 2–4 days for the range of β in consideration. Outbreaks in most of these cases last for about 158 days. The effect of increased sexual transmission in modifying pathogen survival is also minor (Fig. 5b), being within 74–78 days with an average of 76.75 days. A similar situation arises for the attack rate as well (Fig. 5c), with the mean attack rate being 15.12%. If there is no sexual transmission ($\beta = 0$), the attack rate is 14.48%. With a sexual transmission rate of 0.0175 (corresponding to about 10% probability of transmission per coital act), the attack rate increases to 15.97%. This accounts to about 10.29% increase in the epidemic size that can be caused by a high rate of sexual transmission. We also have an unknown factor, the relative sexual transmissibility (μ) when recovering hosts are in the convalescent phases. For all other simulations, we assume the infectiousness will reduce by 50% in convalescent stages. However, we also vary this parameter to determine how it affects the outcomes. Just like the effects we saw for β, the epidemic lengths remain within 1–2 days of the
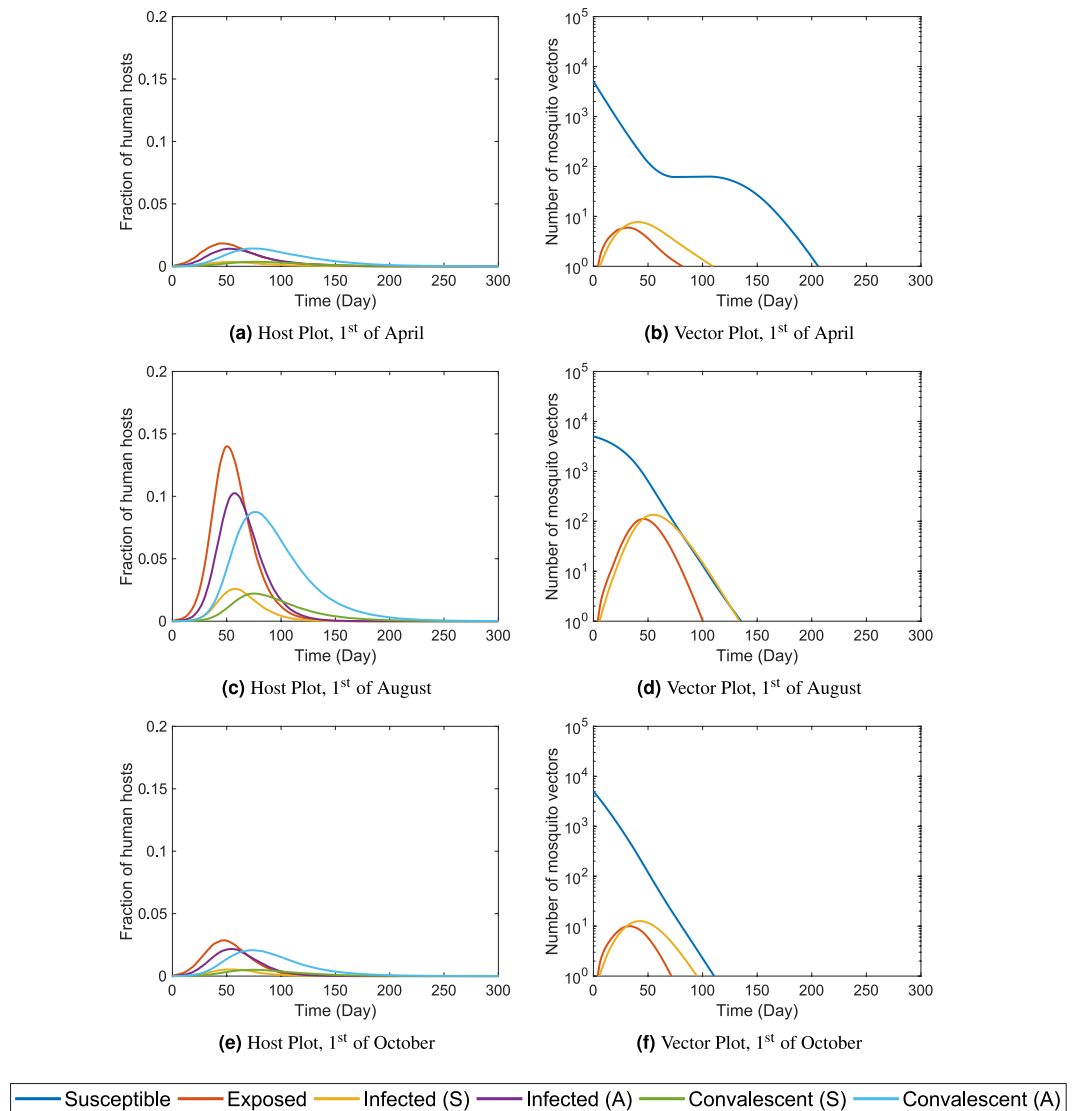
Figure 4. The time series plots of ZIKV outbreaks for three different pathogen introduction months, M_{start} in Miami, FL. The left column contains the host plots and the right column contains the corresponding vector plots. The three rows indicate three different pathogen introduction times: 1st of April, 1st of August, and 1st of October respectively. The hosts in different states are expressed as fractions of the total population. The number of vectors in different compartments are expressed in log scaled axes. Here, we present 5 out of the 7 host states (E, I_s, I_a, J_s, and J_a) and the all three vector compartments (S_V, E_V, and I_V) which are marked with distinct colors described by the legends at the bottom. All plots are averages of 500 independent stochastic simulations.

mean of 157.6 days (Fig. 5d). The pathogen survival also show similar characteristics as before and stays within 73–76 days (Fig. 5e). The effect on the epidemic size is barely noticeable (Fig. 5f), with small fluctuations.

ZIKV disease outcomes are also affected by a large proportion of asymptomatic hosts. We vary the two parameters (τ and θ) that we use to model the asymptomatically infected individuals. Although in most works we found that about 20% of the cases were symptomatic, one recent work estimated that this proportion could be higher[30] (27–50%). To evaluate the effect of such variations, we vary the symptomatic proportion parameter, τ from 10% to 60%. The epidemic length increases gradually with τ starting from 155 days for 10% symptomatic to 172 days for 60% symptomatic population (Fig. 5g). A 50% increase in the symptomatic proportion causes a 11% increase in the length of the outbreak. The pathogen survival is affected much less compared to epidemic length. However, we find a small fluctuating decrease followed by a gradual increase (Fig. 5h). The vector free pathogen survival remains within 2–3 days of the average length of 75.6 days. The epidemic size has a clear increasing trend as shown in Fig. 5i. This is expected because the symptomatically infected individuals are subject to higher transmission rates. As an example, the outbreaks will be about 45% larger if 50% of the people are symptomatic compared to the usual proportion of 20%. The relative transmissibility has significantly larger effect in the disease outcomes and all three properties increase (Fig. 5 bottom row). In most other simulations, it is assumed by default that infectiousness reduces to half ($\theta = 0.5$) when someone is asymptomatic. Compared to our default situation,
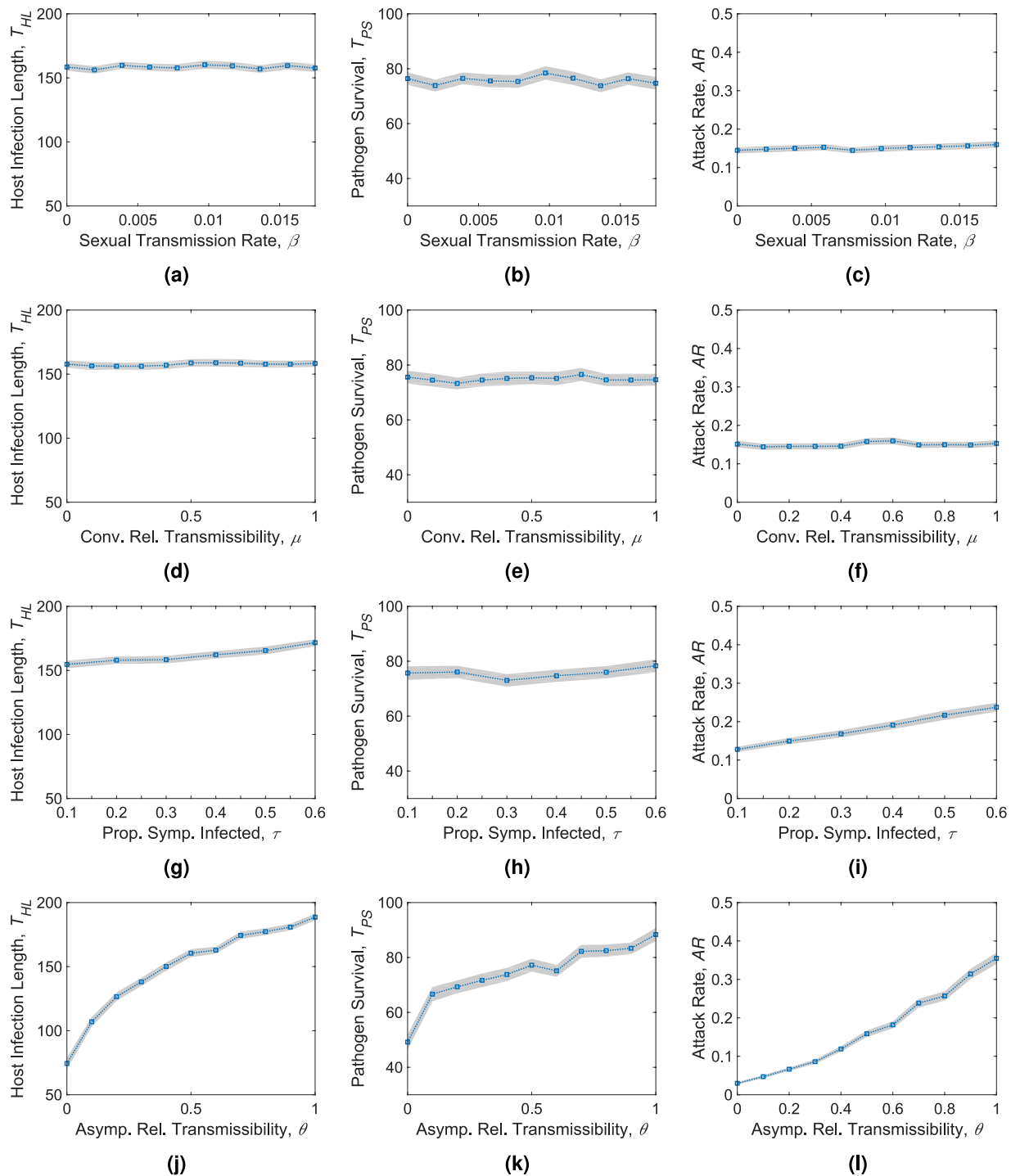
Figure 5. The plots depicting the results of pathogen survival analysis. The results were obtained by varying the sexual transmission rate β (1st row), the relative transmissibility of the convalescent states μ (2nd row), the proportion of symptomatically infected individuals τ (3rd row), and the relative transmissibility of the asymptomatic states θ (4th row). The plots demonstrate the host infection length T_{HL} (1st col), the pathogen survival T_{PS} (2nd col), and the attack rate AR (3rd col). Each data point (blue square) in the above plots is an average of 500 independent stochastic simulations. The shaded regions indicate 95% confidence intervals.

if we make infectiousness of both symptomatic and asymptomatic individuals the same, we see 17.5% increase in the epidemic length (Fig. 5j), 14.45% increase in the pathogen survival period (Fig. 5k), and a massive 123.5% increase in the epidemic size (Fig. 5l).

Sensitivity analysis.   A sensitivity analysis is performed in order to evaluate the model response with respect to several key parameters. For a vector borne disease such as ZIKV, the ratio of vector and host population is

expected to play a prominent role in the disease outcomes. In our previous analysis, we have also found that the proportion of asymptomatic individuals and their relative transmission rate affect the disease outcomes. To evaluate how these parameters interplay, we perform a pairwise sensitivity analysis of the three parameters: the vector-to-host ratio (N_V/N_H), the proportion of asymptomatic hosts (τ), and the relative transmissibility of asymptomatic hosts (θ). The results are depicted in Fig. 6.

The epidemic length (T_{HL}) is affected by the three parameters in question. However, we can see that, the effect of symptomatic proportion depends on the relative transmissibility. The symptomatic proportion can regulate epidemic length if relative transmissibility is low as seen in Fig. 6b. If θ is above 0.4, the capability of τ is greatly reduced. The vector-to-host ratio (N_V/N_H) always increases the epidemic length (Fig. 6a,c), which is also true for the relative transmissibility of asymptomatic individuals (θ) (Fig. 6b,c).

We obtain some interesting results when we analyze the sensitivity of pathogen survival, T_{PS}. Pathogens can survive longer for an intermediate range of host to vector ratio if the proportion of symptomatic individuals is low as shown in Fig. 6d. On the other hand, despite having some large epidemics on the upper right corner of Fig. 6g, we see that pathogen survival is relatively lower (as low as 67 days) (Fig. 6d). In Fig. 6e, we see that depending on the value of τ, the survival can be higher for certain intermediate values of θ. For example, When $\tau < 0.2$ and $0.3 < \theta < 0.7$, there is a region where pathogen can survive more compared to the other situations. Survival reduces for both high and low values of asymptomatic relative transmissibility (θ). When comparing vector-to-host ratio along with the asymptomatic relative transmissibility, we see that a thick band or a region appears, where the pathogen survival is high (Fig. 6f). There are slight fluctuations in that peak region, however, they are relatively minor to indicate any particular phenomenon. If we compare this plot with the attack rate plot (Fig. 6i), we see that pathogen survival is relatively long for intermediate values of attack rates and relatively short for small or large attack rates. Small outbreaks naturally end sooner. Very large outbreaks also can end soon due to faster spreading dynamics and herd immunity of the recovered population.

The epidemic attack rate (AR), as expected, is highly sensitive to vector-to-host ratio, N_V/N_H for the entire range of asymptomatic proportion, τ. For the nominal value of $\tau = 0.2$, doubling the N_V/N_H ratio increases epidemic size by 249.78% (Fig. 6g). However, if the asymptomatic individuals have very low transmission capabilities, it can limit the effect of vectors on epidemic size as seen in Fig. 6i. The fact that majority (80%) of the hosts are asymptomatic in ZIKV infections, indicate that they hold a critical role in spreading. The parameters τ and θ both demonstrate importance in determining epidemic size (Fig. 6h) and among them, the effect of θ is more radical.

Discussion

In this study we have proposed an individual based interconnected network model for ZIKV that can also be used to simulate any vector-borne disease that features contact based direct transmissions. We employ heterogeneous mixing based on the host contact network which is generated based on real world data on human sexual behavior, sexual orientation, gender, and age structure. We utilize the approximations of the non-Markovian Gillespie algorithm to run stochastic simulations. In the beginning, we explore how the seasons can affect this predominantly vector-borne disease. Later, we focus on the survival of pathogen in the climates similar to Florida if an outbreak starts prior to the colder months. In this step, we examine how some of the important model parameters affect the outcomes. Finally, we perform a sensitivity analysis to evaluate our model behavior in response to changes in some key parameters.

Our seasonal analysis results indicate that outbreak size is strongly related to the environmental conditions during the pathogen introduction. The first few weeks are crucial in determining how much the pathogen would spread. After that period, environmental variations have a much weaker effect in reducing or increasing the outbreak size. If the pathogen is introduced during the peak mosquito season, there is a high probability that we will see a large outbreak. Even if climatic suitability of mosquito vectors decline rapidly after the first few weeks, the vectors manage to spread the pathogen in the host population rapidly and cause large outbreaks. This suggests that a ZIKV outbreak can spread rapidly out of control if it is not effectively contained during the initial stage. Early interventions are crucial even though it may be difficult to identify outbreaks due to large asymptomatic group of hosts. The ratio of vector to host, as expected, has shown its prominence in determining outbreak size and the length (Fig. 6).

We have analyzed how sexual transmission affect the outbreaks. Our results in Fig. 5c,f align with the conclusions drawn by some of the earlier works[4–6] that sexual transmission is a small component in the overall force of infection, which is dominated by the vectored-transmission. However, pathogen still survives up to 75 days in the host network without the help of infected vectors. This long survival can be attributed to small amount of sexual transmission but it is mainly due to the extended infectiousness (convalescence) of the hosts as it is assumed that pathogen can survive up to a month in the semen of recovering males. The use of certain birth controls (e.g. condoms) could be effective in combating the spread of pathogen during this period. These conclusions should inspire further clinical studies in this area to test the efficacy of control measures.

A large number of asymptomatically infected individuals play an important role in outbreak dynamics. The proportion of symptomatic individuals do not affect the survival of pathogen noticeably. However, it is positively correlated to outbreak size. The relative transmissibility of asymptomatic states is one of the key factors in determining disease outcomes which can extend the vector-free pathogen survival up to 3 months. It is one of the most important parameters that needs to be properly estimated in order to obtain informative outbreak predictions.

The conclusions drawn from our results can be useful in evaluating potential endemic scenarios for Zika virus disease in temperate regions. Although the contribution of sexual transmission is small, the ability of the pathogen to survive in a human sexual contact network for extended periods can have consequences in sustaining a Zika outbreak in a region as well as spreading to other regions due to long distance travels. In addition to that, a large asymptomatic proportion could be the most important hurdle in controlling ZIKV outbreaks. Although
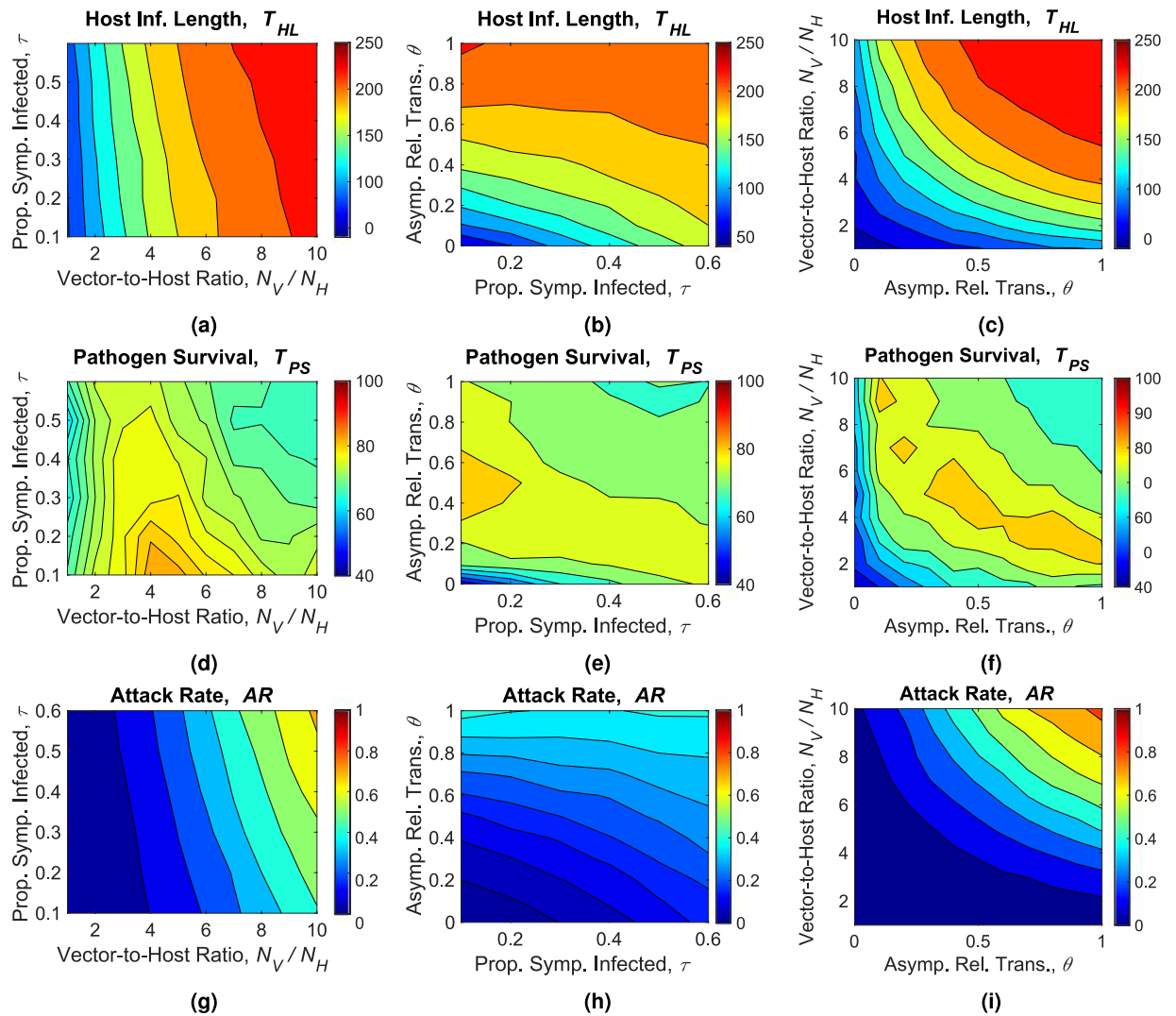
Figure 6. Contour plots depicting the results of the sensitivity analysis. The parameters varied in each plot are marked as axes labels. The quantity for which the contours are being displayed is mentioned on top of each plot. The contours are color mapped and a color-bar on the side of each plot indicates the range of values represented by the plot. The 1st row shows the sensitivity of the epidemic length (T_{HL}) on the parameters, the 2nd row shows the sensitivity of the pathogen survival (T_{PS}) on the parameters, and the 3rd row shows the sensitivity of the epidemic attack rate (AR) on the parameters. The parameters that were varied here are: the vector-to-host ratio (N_V/N_H), the proportion of symptomatically infected individuals (τ), and the relative transmissibility of the asymptomatic states (θ). Each data point shown in the above plots is an average of 500 independent stochastic simulations.

we provide conclusions on the relative importance of key parameters, the unavailability of data on some of those warrants for further efforts towards estimations. This model should be applicable to other vector-borne disease which have the potential of being transmitted sexually. Our model is also mesoscale (medium sized population) in terms of hosts, due to the limitations imposed by computational complexities. This work can be extended in the future by the use of activity driven networks (ADN), vertical transmission, and larger populations. Those studies would provide more insights into the study of Zika virus epidemics.

Methods

Host network characterization.
For the host population, we assume equal number of males and females. We divide the population into three age groups: Children (0–14), Adults (15–64), and Elderly (65+). Based on the World Development Indicators (WDI) data published for 2017[37], we find that for the USA population, the three age groups are distributed as 19% Children, 66% adults, and 15% elderly people. The adults are the only ones assumed to be capable of transmitting the disease sexually, hence the remaining population are only affected by vectors. Based on the data provided in a sexual behavior study[24], we find that sexual orientation of men consisted of approximately 97.2% heterosexuals, 2.5% homosexuals, and 0.3% bisexuals. For women, the study revealed 98.9% heterosexuals, 0.9% homosexuals, and 0.2% bisexuals. The same work[24] also compiles a distribution of the

No. of Partners	% of population
0	15.7
1	71.8
2	5.5
3	2.8
4	1.6
5–9	1.5
10–19	0.4
20+	0.2

Table 2. Average number of sexual partners in the last 12 months[24].

population based on the number of sexual partners which is listed in Table 2. The data indicate that majority of the population have a single partner in a timespan of 12 months. The average number of partners for the adult population was 1.28.

The network generator tool is designed using the configuration model[32]. It takes all the above mentioned population and sexual behavior properties as inputs, and generates a graph whose statistical properties closely match the given partner distribution of Table 2 and the distribution of sexual orientation. Studies on sexual networks also indicate that sexual contact networks have node degree distributions that follow the scale-free structure[38]. There are large variations in the number of sexual partners while a small group of highly active people form the core[38]. Core groups are important in sustaining pathogen transmission specially if the duration of infection is short[39]. To maintain these features, the network generator is designed in such a way that, high degree nodes have higher probabilities of connecting to low degree nodes. An example of network generated by our generation tool is shown in Fig. 2.

Vector characterization. The mosquito vectors are modeled as homogeneous population and we use the classical Ross-Macdonald approach used by Keeling *et al.* in their book[40]. We divide the vector population into three compartments, Susceptible (S_V), Exposed (E_V), and Infected (I_V). The transitions between the three compartments are showed in Fig. 1. Table 1 describes the different parameters that were used for the model. The equations for disease dynamics in mosquito vectors are given below,

$$
\begin{aligned}
\dot{S}_V &= F(t)N_V - \lambda_{VH}N_H\left(\sum_{j=1}^{N_H}[I_s + \theta I_a]_j\right)S_V - \varepsilon S_V \\
\dot{E}_V &= \lambda_{VH}N_H\left(\sum_{j=1}^{N_H}[I_s + \theta I_a]_j\right)S_V - \sigma E_V - \varepsilon E_V \\
\dot{I}_V &= \sigma E_V - \varepsilon I_V
\end{aligned}
\tag{5}
$$

We incorporate seasonality into this model by using a time dependent mosquito recruitment rate, $F(t)$. This rate depends on the time (day) of the year. The transmission parameters, the λ's are computed from the mosquito bite rate, r and transmission probability, T. The formula are given in Table 1.

Non-Markovian Gillespie Algorithm. The nMGA (Non-Markovian Gillespie Algorithm) was proposed by Boguná *et al.*[27]. The following derivation was also described in the work of Masuda *et al.*[41].

We first consider N renewal processes running in parallel. Let t_i be The time elapsed since the last event of the ith process. We denote $\psi_i(\tau)$ as the probability density function of inter-event times for the ith process. The survival function of the ith process (i.e., the probability that the inter-event time is larger than t_i) is,

$$
\Psi_i(t_i) = \int_{t_i}^{\infty} \psi_i(\tau)d\tau
\tag{6}
$$

Now, the probability that no process generates an event for time Δt is,

$$
\Phi(\Delta t|\{t_j\}) = \prod_{j=1}^{N} \frac{\Psi_j(t_j + \Delta t)}{\Psi_j(t_j)}
\tag{7}
$$

To determine the time until the next event, Δt, we take a sample u from uniform distribution over [0, 1] and solve $\Phi(\Delta t|\{t_j\}) = u$. This step is computationally expensive when N is large. To improve performance, we approximate this step as proposed by Boguná *et al.*[27]. This approximation is exact as $N \to \infty$. When Δt is small (N is large), equation (7) becomes[41],

$$
\Phi(\Delta t|\{t_j\}) \approx exp\left[-\Delta t\left(\sum_{j=1}^{N}\lambda_j(t_j)\right)\right]
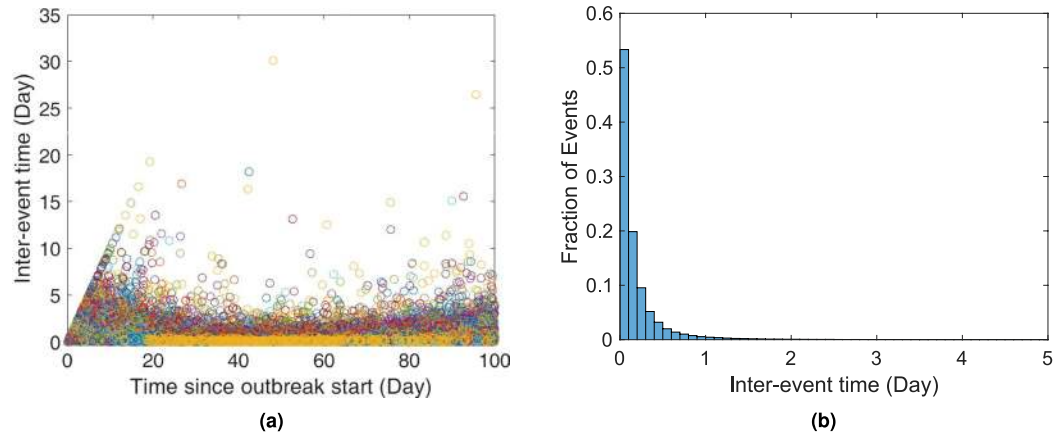\tag{8}
$$

Figure 7. Inter-event times in different stages for all of the 500 simulations during the first 100 days (left) and their distribution (truncated inter-event times >5) (right). For this case, the average inter-event time is 0.2074 day (95% CI [0.2056 to 0.2092]). About 96.87% of the events that occurred had intervals shorter than a day. There are a few outliers though, which are mostly due to slowing down of the events at the end of the epidemics (when pathogen is low in the host population or has been wiped out).

Now, the instantaneous (hazard) rate of the ith process, which is generally assumed to be a function of time since the last event is determined by,

$$\lambda_i(t_i) \equiv \frac{\psi_i(t_i)}{\Psi_i(t_i)} \tag{9}$$

With the above equations in hand, we can run the Non-Markovian Gillespie algorithm as follows,

1. Initialize t_j for all $(1 \leq j \leq N)$.
2. Determine the time until next event from,

$$\Delta t = \frac{-\ln u}{\sum_{j=1}^{N} \lambda_j(t_j)} \tag{10}$$

3. Select the process i that has generated the event with probability,

$$\Pi_i \equiv \frac{\lambda_i(t_i)}{\sum_{j=1}^{N} \lambda_j(t_j)} \tag{11}$$

4. Update the time since the last event, $t_j = t_j + \Delta t$ for all $j \neq i$. Set $t_i = 0$.
5. Repeat steps 2–4.

The original Gillespie Algorithm can be recovered from this nMGA using $\lambda_i(t_i) = \lambda_i$. This is the case for all the parameters that are constant.

Numerical simulation. We use the non-Markovian Gillespie Algorithm (nMGA) to simulate the processes related to host nodes. The vector population is simulated using deterministic ordinary differential equation (ODE) solver. We combine the hosts and the vector simulations together by calling the ODE solver at the end of each event. The GEMFsim[28] tool, which already supports the Gillespie algorithm was modified in order to include the vector population model and time varying parameters. We use the variables X and Y to describe the host and vector populations in different states/compartments respectively. Here, X contains information about the state of each host node and Y contains the population count of each vector compartment. X_0 and Y_0 are the initial states/compartments at the start of the simulation. $\mathcal{A}$ is the host network representation. The term *PARAM* is used to denote the set of model parameters listed in Table 1. T_H and T_V in the output are the host and vector indexing arrays that contain information about time where the data points (X and Y) were generated. We denote $x_n \in X$ as the state of node n. $\lambda_n(x_n \rightarrow j)$ is the transition rate of node n from its current state x_n to state j. We use M to denote the total number of host states ($=7$ in our case). For a node n, Λ_n denotes the sum of transition rates from its current state to all other possible states. The *VECTSOLVER* is an ordinary differential equation (ODE) solver that takes as input the current situation of the vector population Y and solves them from the current instance to the time δt. The ODE equations that are being solved are given in equation (5). This solver is invoked once after each event (with the updated parameters) and the time indexing vectors are updated with respect to the global time t. The simplified pseudo-code of the modified GEMFsim is given below,

```
    input  : X_0, Y_0, PARAM, T_end, 𝒜
    output : X, Y, T_H, T_V
 1  X ← X_0;
 2  Y ← Y_0;
 3  while t ≤ T_end do
 4      for n ← 1 to N do
 5          Λ_n ← ∑_{j=1}^M λ_n(x_n → j);
 6      end
 7      Λ_tot ← ∑_{n=1}^N Λ_n;
        // sample a value from uniform distribution
 8      u ~ Uniform(0,1);
        // sample time until the next event
 9      δt ← − ln(u)/Λ_tot;
10      P_1(n) ← Λ_n/Λ_tot;
        // sample a node from activity rate distribution
11      k ~ P_1;
        // read the current state of the selected node k
12      i_k ← x_k;
13      P_2(j) ← λ_k(i_k → j)/Λ_k;
        // sample the future state of the selected node k
14      m_k ~ P_2;
        // the event quartet is defined as (time until the event, participating node,
            old state, new state)
15      event ← (δt, k, i_k, m_k);
16      x_k ← m_k;
17      I_H ← count the infected hosts;
        // invoke the ODE solver for vector eqns for a duration of δt
18      [Y, t_v] ← VECTSOLVER (Y, PARAM, [0, δt]);
19      t_v ← t_v + t;
20      t ← t + δt;
21      T_H ← [T_H, t];
22      T_V ← [T_V, t_v];
23      I_V ← count the infected vectors;
24  end
```

As the simulation is event-based and parameters are updated following each event, the inter-event times should be small to keep the variable parameters up to date in both populations. Longer inter-event times in a simulation can increase parameter discrepancy between the host and the vector sub-models. These sub-models are coupled by inter-dependent parameters that vary with time and update after each event. Hence, it is important that events occur at shorter intervals to keep the model outcomes accurate. Fortunately, due to the nature of the Gillespie algorithm, it is indeed the case when a pathogen is present in the host population. We have demonstrated this fact in Fig. 7 that ensures parameter accuracy. The inter-event times are also exponentially distributed which is shown in Fig. 7b. We have a small number of outliers that indicate long inter-event times. Most of these occur when the infection is either very low in the population or it is about to die out. Hence, the simulation results remain unaffected.

Code Availability
The MATLAB implementation of our model which was used to generate host networks, run outbreak simulations, and process simulation results is available for the interested reader at the Network Science and Engineering (NetSE) group website[42] of Kansas State University.

References
1. Zhang, Q. *et al*. Spread of Zika virus in the Americas. *Proceedings of the National Academy of Sciences* **114**, E4334–E4343 (2017).
2. D'Ortenzio, E. *et al*. Evidence of Sexual Transmission of Zika Virus. *New England Journal of Medicine* **374**, 2195–2198 (2016).
3. Musso, D. *et al*. Potential sexual transmission of Zika virus. *Emerging Infectious Diseases* **21**, 359 (2015).
4. Gao, D. *et al*. Prevention and Control of Zika as a Mosquito-Borne and Sexually Transmitted Disease: A Mathematical Modeling Analysis. *Scientific Reports* **6**, 28070 (2016).
5. Agusto, F., Bewick, S. & Fagan, W. Mathematical model for Zika virus dynamics with sexual transmission route. *Ecological Complexity* **29**, 61–81 (2017).
6. Maxian, O., Neufeld, A., Talis, E. J., Childs, L. M. & Blackwood, J. C. Zika virus dynamics: When does sexual transmission matter? *Epidemics* **21**, 48–55 (2017).
7. Saad-Roy, C., Ma, J. & van den Driessche, P. The effect of sexual transmission on Zika virus dynamics. *Journal of mathematical biology* 1–25 (2018).

8. Terefe, Y., Gaff, H., Kamga, M. & van der Mescht, L. Mathematics of a model for Zika transmission dynamics. *Theory in Biosciences* 1–10 (2018).
9. Sirohi, D. *et al*. The 3.8 å resolution cryo-EM structure of Zika virus. *Science* **352**, 467–470 (2016).
10. Sikka, V. *et al*. The emergence of zika virus as a global health security threat: A review and a consensus statement of the INDUSEM Joint working Group (JWG). *Journal of Global Infectious Diseases* **8**, 3 (2016).
11. Mlakar, J. *et al*. Zika Virus Associated with Microcephaly. *New England Journal of Medicine* **2016**, 951–958 (2016).
12. Malone, R. W. *et al*. Zika Virus: Medical Countermeasure Development Challenges. *PLoS Neglected Tropical Diseases* **10**, e0004530 (2016).
13. Zika Cases in the United States. URL, https://www.cdc.gov/zika/reporting/case-counts.html. [Accessed: 03-May-2018] (2018).
14. Castro, L. A. *et al*. Assessing real-time Zika risk in the United States. *BMC Infectious Diseases* **17**, 284 (2017).
15. Kucharski, A. J. *et al*. Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 2013–14 French Polynesia Outbreak. *PLoS Neglected Tropical Diseases* **10**, e0004726 (2016).
16. Kuhlman, C. J., Ren, Y., Lewis, B. & Schlitt, J. Hybrid Agent-based modeling of Zika in the United States. In *2017 Winter Simulation Conference (WSC)*, 1085–1096 (IEEE, 2017).
17. Agusto, F., Bewick, S. & Fagan, W. Mathematical model of Zika virus with vertical transmission. *Infectious Disease Modelling* (2017).
18. Olawoyin, O. & Kribs, C. Effects of multiple transmission pathways on Zika dynamics. *Infectious Disease Modelling* (2018).
19. Moghadas, S. M. *et al*. Asymptomatic Transmission and the Dynamics of Zika Infection. *Scientific Reports* **7**, 5829 (2017).
20. Pugliese, A., Gumel, A. B., Milner, F. A. & Velasco-Hernandez, J. X. Sex-biased prevalence in infections with heterosexual, direct, and vector-mediated transmission: a theoretical analysis. *Mathematical Biosciences & Engineering* **15**, 125–140 (2018).
21. Sasmal, S. K., Ghosh, I., Huppert, A. & Chattopadhyay, J. Modeling the Spread of Zika Virus in a Stage-Structured Population: Effect of Sexual Transmission. *Bulletin of mathematical biology* 1–30 (2018).
22. Allard, A., Althouse, B. M., Hébert-Dufresne, L. & Scarpino, S. V. The risk of sustained sexual transmission of Zika is underestimated. *PLoS pathogens* **13**, e1006633 (2017).
23. Turmel, J. M. *et al*. Late sexual transmission of Zika virus related to persistence in the semen. *The Lancet* **387**, 2501 (2016).
24. Smith, T. W. *American Sexual Behavior: Trends, Socio-Demographic Differences, and Risk Behavior* (National Opinion Research Center Chicago, 1998).
25. Sahneh, F. D., Scoglio, C. & Chowdhury, F. N. Effect of coupling on the epidemic threshold in interconnected complex networks: A spectral analysis. In *American Control Conference (ACC), 2013*, 2307–2312 (IEEE, 2013).
26. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81**, 2340–2361 (1977).
27. Boguná, M., Lafuerza, L. F., Toral, R. & Serrano, M. Á. Simulating non-Markovian stochastic processes. *Physical Review E* **90**, 042108 (2014).
28. Sahneh, F. D., Vajdi, A., Shakeri, H., Fan, F. & Scoglio, C. GEMFsim: a stochastic simulator for the generalized epidemic modeling framework. *Journal of Computational Science* **22**, 36–44 (2017).
29. Gray, R. H. *et al*. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. *The Lancet* **357**, 1149–1153 (2001).
30. Mitchell, P. K. *et al*. Reassessing Serosurvey-Based Estimates of the Zika Symptomatic Proportion. *American journal of epidemiology* (2018).
31. Zika has been sexually transmitted in Texas, CDC confirms. URL, https://www.cnn.com/2016/02/02/health/zika-virus-sexual-contact-texas/index.html. [Accessed: 01-November-2018] (2016).
32. Newman, M. *Networks* (Oxford University Press, 2018).
33. Monaghan, A. J. *et al*. On the Seasonal Occurrence and Abundance of the Zika Virus Vector Mosquito *Aedes aegypti* in the Contiguous United States. *PLoS Currents* **8** (2016).
34. Reiskind, M. & Lounibos, L. Spatial and temporal patterns of abundance of *Aedes aegypti* L.(*Stegomyia aegypti*) and *Aedes albopictus* (Skuse)[*Stegomyia albopictus* (Skuse)] in southern Florida. *Medical and Veterinary Entomology* **27**, 421–429 (2013).
35. Moon, S. A., Cohnstaedt, L. W., McVey, D. S. & Scoglio, C. M. A spatio-temporal individual-based network framework for West Nile virus in the USA: spreading pattern of West Nile virus. *PLoS Computational Biology* **15**, e1006875 (2019).
36. Marini, G., Guzzetta, G., Rosà, R. & Merler, S. First outbreak of Zika virus in the continental United States: a modelling analysis. *Eurosurveillance* **22** (2017).
37. World Development Indicators: Population dynamics. URL, http://wdi.worldbank.org/table/2.1. [Accessed: 16-October-2018] (2017).
38. Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E. & Åberg, Y. The web of human sexual contacts. *Nature* **411**, 907–908 (2001).
39. Robinson, K., Cohen, T. & Colijn, C. The dynamics of sexual contact networks: Effects on disease spread and control. *Theoretical Population Biology* **81**, 89–96 (2012).
40. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2011).
41. Masuda, N. & Rocha, L. E. A Gillespie Algorithm for Non-Markovian Stochastic Processes. *SIAM Review* **60**, 95–115 (2018).
42. GEMFsim vector-borne. URL, https://www.ece.k-state.edu/netse/software/ (2019).
43. Caminade, C. *et al*. Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015. *Proceedings of the National Academy of Sciences* **114**, 119–124 (2017).

Acknowledgements

Author Contributions

T.F., C.S. and L.C. formulated the model; T.F., L.C. and D.M. analyzed the data; T.F. and C.S. developed the methodology; T.F. developed the simulation tool; T.F., D.M. and L.C. analyzed the results and contributed to the discussions; T.F., L.C., D.M. and C.S. contributed to the writing. All authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.