



Original Research Article

Epidemic spread on patch networks with community structure

Brandon Lieberthal^{a,*}, Aiman Soliman^b, Shaowen Wang^b, Sandra De Urioste-Stone^a,
Allison M. Gardner^a

^a University of Maine, Orono, ME, USA

^b University of Illinois, Urbana-Champaign, IL, USA



ARTICLE INFO

Keywords:

Community networks
Patch networks
Metapopulation SIR
Disease epidemics
Super-spreader events

ABSTRACT

Predicting and preparing for the trajectory of disease epidemics relies on a knowledge of environmental and socioeconomic factors that affect transmission rates on local and global spatial scales. This article discusses the simulation of epidemic outbreaks on human metapopulation networks with community structure, such as cities within national boundaries, for which infection rates vary both within and between communities. We demonstrate mathematically, through next-generation matrices, that the structures of these communities, setting aside all other considerations such as disease virulence and human decision-making, have a profound effect on the reproduction rate of the disease throughout the network. In high modularity networks, with high levels of separation between neighboring communities, disease epidemics tend to spread rapidly in high-risk communities and very slowly in others, whereas in low modularity networks, the epidemic spreads throughout the entire network as a steady pace, with little regard for variations in infection rate. The correlation between network modularity and effective reproduction number is stronger in population with high rates of human movement. This implies that the community structure, human diffusion rate, and disease reproduction number are all intertwined, and the relationships between them can be affected by mitigation strategies such as restricting movement between and within high-risk communities. We then test through numerical simulation the effectiveness of movement restriction and vaccination strategies in reducing the peak prevalence and spread area of outbreaks. Our results show that the effectiveness of these strategies depends on the structure of the network and the properties of the disease. For example, vaccination strategies are most effective in networks with high rates of diffusion, whereas movement restriction strategies are most effective in networks with high modularity and high infection rates. Finally, we offer guidance to epidemic modelers as to the ideal spatial resolution to balance accuracy and data collection costs.

Author summary

Predicting and preparing for the trajectory of disease epidemics relies on a knowledge of environmental and socioeconomic factors that affect transmission rates on local and global spatial scales. This article discusses the simulation of epidemic outbreaks on human metapopulation networks with community structure, such as cities within national boundaries, for which infection rates vary both within and between communities. We demonstrate mathematically, through next-generation matrices, that the structures of these communities, setting aside all other considerations such as disease virulence and human decision-making, have a profound effect on the reproduction rate of the disease throughout the network. This implies that the community structure, human diffusion rate, and disease reproduction number are all intertwined, and the relationships between them can be affected by mitigation strategies such as restricting movement between and within high-risk communities. We then test through numerical simulation

the effectiveness of movement restriction and vaccination strategies in reducing the peak prevalence and spread area of outbreaks. Finally, we offer guidance to epidemic modelers as to the ideal spatial resolution to balance accuracy and data collection costs.

1. Introduction

Predicting and explaining spatial patterns of infectious disease spread is a complex challenge in part due to the significant geographic, demographic, economic, and social heterogeneity that exists in human populations, which can alter the spatiotemporal trajectory of an epidemic [1–3]. In addition to physical factors that enhance or inhibit disease spread, such as climate, landscape configuration, and land use, the distribution of human host populations at local, national, and global scales may profoundly influence transmission dynamics [4,5]. Analysis of epidemics across multiple spatial scales is useful, for example, to

* Corresponding author.

E-mail address: brandon.lieberthal@maine.edu (B. Lieberthal).

assess whether the drivers of disease spread are the same at local versus national scales [6] or to compare the contributions of different modes of transportation to spreading disease [7].

Several modeling techniques exist to study these multi-scale problems. One such technique is the community metapopulation network, which consists of a graph of nodes, each representing a population center, that are connected by mobility pathways [8]. These nodes are grouped into communities, which are abstract structures defined by having significantly more connections among themselves compared to connections with neighboring communities, where connections can represent physical adjacency or transportation routes [9]. The community network has several available algorithms for design and numerous practical applications, including the simulation of cities divided by political boundaries, children and adults who attend the same schools and workplaces, or even users of a social network grouped abstractly by a common interest [10–12]. Due to the increasing urbanization of human population centers, the community network is an extremely important model for studying epidemics and pandemics, and considerable research has addressed the effects of community structure on disease spread [13–17]. Prior research has shown that a strong, well-connected community structure, contrary to common wisdom, actually reduces the danger of epidemics as the disease stays less isolated in high-risk communities [18].

Scientific research into community networks typically focuses on the number of bridges between connected communities, the rate of host movement between communities, or the modularity of the network, and how these factors affect the trajectory of an epidemic [19–21]. Recent research has also explored the disease epidemic threshold and optimal control strategies in community networks [22–25]. This research is more often theoretical than empirical, and has found no straightforward relationships between community structure and epidemic reproduction rates. Community networks are typically described mathematically using generating functions [26], but generating functions are only applicable to the specific case in which the disease infection rate is constant across the network. In general, prior research concerning community networks assumes that each community in the network has qualitatively similar graph structures and human populations, as well as the same infection rate, and there is considerable room to improve upon the applicability of these models.

Another modeling technique, the patch network, is useful in epidemiological cases in which the rate of infection varies either spatially or among different susceptible host groups in the same population center. Examples include diseases for which vector habitat suitability varies with climate or land cover [27–29], and diseases for which exposure and infection rates vary between demographic groups [30, 31]. Patch networks are also useful when different segments of the population experience different policies to mitigate the impact of an ongoing epidemic [32,33]. Numerous studies have attempted to predict the spatial distribution of infection rates for specific diseases as a product of environmental and socioeconomic factors, and patch networks are of considerable importance for diseases with multiple host species, since they can accurately model interactions within and between species [34,35]. The next-generation matrix, a mathematical construct that is computed by the infection rates of individual host populations and the contact rate between these populations, is commonly used to compute the reproduction rate of epidemics in patch networks [36].

An accurate and robust epidemiological model across a large spatial extent should ideally integrate elements of the community network and the patch network techniques, and a key question in designing these models is the scale at which host populations should be considered. There is a trade-off between the precision offered by fine scale data and the lower costs of obtaining coarse scale data, but additionally the values of some simulation parameters may fundamentally change depending on the spatial scale under consideration [37,38]. Climate data, for example, are typically collected at very fine scales and vary

continuously with space, so the scale at which the data are rendered is not critical to their precision [39]. Several other data types, such as population density, urbanicity, and socioeconomic factors, can vary significantly within a small region. If a given geographical region contains both substantial urban and rural populations, for example, the network graph of a human mobility model will look very different depending on the spatial granularity of the region [40]. This issue is related to the ecological fallacy, in which inferences about smaller population centers are made from the properties of the community at large [41], and to the modifiable areal unit problem, in which the arbitrary drawing of political boundaries can affect the statistical analysis of natural phenomena [42]. These issues compound with the additional complication of a patch network, because if the infection rate depends both on natural and human factors, as is the case for most diseases [43], then an epidemiological model can be expected to exhibit variation on multiple scales.

To this end, this manuscript builds upon previous research by combining these two modeling approaches, the patch network and the community network, and studies in tandem the effects of community network structure and heterogeneous infection rates on several metrics of epidemic severity, such as reproduction rate, peak infection rate, and the area of disease spread. We address two questions: (1) how does community structure in patch networks affect the outcome of epidemiological simulations? (2) what are the consequences of the choice to represent a host population region as either a community or a single node? We hypothesize that infection rates, host diffusion, and community structure are primary, secondary, and tertiary effects on the reproduction rate of the epidemic, respectively, and a similar order of importance applies to the peak prevalence, the area of spread, and super-spreader capacity [44]. We explore these questions through analytical models and through simulation models of randomly generated networks, and we discuss how the modularity of the network, as well as the properties of individual communities, affect the scale-dependence of reproduction rate (R), the number of individuals infected, and the spatial extent of the epidemic.

The first part of this manuscript features an analytical model that applies next-generation matrix algorithms to patch networks with community structure. This model, which assumes homogeneous mobility between nodes, proves mathematically that the community structure of the network has a direct effect on the basic reproduction number throughout the population. We show through equations and examples that with all other factors held constant, the modularity of the community network can nearly double the R value of the disease, or reduce it by about half. In the second part, we use Monte Carlo simulations to study an expanded model that considers asymptomatic and non-traveling infected individuals, as well as heterogeneous host mobility. We use these simulations to study how community structure affects the peak infection rate and super-spreader capacity of the network, and we demonstrate the level of error that may be introduced to the simulation as a result of neglecting community structure when designing an empirical epidemic model.

2. Methods

2.1. SEIIR model with traveler behavior and asymptomatic transmission

A metapopulation network is defined by a series of nodes, representing subpopulations, that are connected by edges, representing mobility routes. The most common metapopulation network used in epidemiology is a map of cities connected by roads, air routes, and sea routes, but a metapopulation network can be used to model a population of disease hosts at any spatial scale through any means of interpersonal connection. Mathematically speaking, a metapopulation network with n nodes is represented by an n -vector $\mathbf{N}$ and an $n \times n$ adjacency matrix $\mathbf{A}$. The value N_i represents the population of node i in the network, and the matrix $\mathbf{A}$ is a square symmetric matrix in which

$A_{ij} = 1$ if nodes i and j are connected, and $A_{ij} = 0$ if they are not. There is no mathematical significance to the order in which the nodes are listed, nor to their geometrical coordinates on a visual representation of the network. When modeling an empirical disease outbreak for a broad audience, nodes are typically placed on a map corresponding to their geographical locations.

A community network is mathematically identical in structure to a metapopulation network, except the nodes are characterized as being concentrated into subnetworks, or communities. Community structure is typically quantified by network modularity, a value ranging from -0.5 to 1 , where a higher modularity indicates a greater ratio of nodal connections within communities than between them [19]. Communities are connected by bridge connections, which connect two nodes in different communities. In this paper, nodes within the same community are sorted together, so the adjacency matrix has a pseudo-block diagonal structure, where non-zero values outside the blocks represent bridge connections.

The metapopulation SIR (Susceptible–Infected–Recovered) model is useful for next-generation matrix analysis in the early outbreak limit, when infection case growth is exponential, but to simulate the effects of traveler behavior and mitigation measures on reproduction rates, a more complex model is necessary. Therefore, we utilize an SEIIIR (Susceptible–Exposed–Infected (Asymptomatic, Traveling, Not Traveling), Recovered) model, in which the infected group is decomposed into asymptomatic individuals, symptomatic individuals who continue to travel, and symptomatic individuals who cease to travel [45]. The differential equations that govern this system are as follows:

$$\begin{aligned}\frac{dS_i}{dt} &= -\frac{\beta_i S_i}{N_i} (r_\beta I_i^a + I_i^t + I_i^m) - \sum_j M_{ij} S_i + \sum_j M_{ji} S_j \\ \frac{dE_i}{dt} &= \frac{\beta_i S_i}{N_i} (r_\beta I_i^a + I_i^t + I_i^m) - \epsilon E_i - \sum_j M_{ij} E_i + \sum_j M_{ji} E_j \\ \frac{dI_i^a}{dt} &= \epsilon E_i p_a - \mu I_i^a - \sum_j M_{ij} I_i^a + \sum_j M_{ji} I_j^a \\ \frac{dI_i^t}{dt} &= \epsilon E_i (1 - p_a) p_t - \mu I_i^t - \sum_j M_{ij} I_i^t + \sum_j M_{ji} I_j^t \\ \frac{dI_i^m}{dt} &= \epsilon E_i (1 - p_a) (1 - p_t) - \mu I_i^m \\ \frac{dR_i}{dt} &= \mu (I_i^a + I_i^t + I_i^m) - \sum_j M_{ij} R_i + \sum_j M_{ji} R_j\end{aligned}\quad (1)$$

where S_i, E_i, R_i, N_i represent the number of susceptible, exposed, recovered, and total population sizes, respectively, of node i . Infected individuals are grouped into asymptomatic individuals (I_i^a) with probability p_a , and symptomatic individuals are further divided into those who continue to travel (I_i^t) with probability p_t , or cease to travel (I_i^m). The variable β_i is the infection rate for node i , r_β is an infection rate adjustment factor for asymptomatic individuals, ϵ is the rate at which exposed individuals become infected, μ is the recovery rate, and M_{ij} is the mobility matrix for the network.

Assuming birth and death rates are negligible over short time scales, the basic reproduction rate for an individual node is given by $R_0 = \frac{\beta_i}{\mu} [1 + (r_\beta - 1)p_a]$. We define the traffic-dependent mobility model as:

$$M_{ij} = \begin{cases} p \frac{(k_i k_j)^\theta}{C k_i^{1+\theta}} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}, \quad (2)$$

where p is the diffusion rate (a value between 0 and 1, indicating the fraction of the population that migrates outside their home node per unit time), and C is a calibration factor [46]. This model represents a system in which the rate of movement between two neighboring nodes is proportional to the product of their respective degrees of connectivity. The variable θ represents heterogeneity of movement, the shape of the curve that relates a node's degree of connectivity to its rate of movement. If $\theta = 0$ then movement is completely homogeneous, meaning all individuals in all nodes have the same movement rate. A

typical value of θ in empirical case studies is 0.5 [47]. More complex mobility models may account for different spatial and temporal scales of commuters [48], or between origin-driven and destination-driven contact [49], but for this analysis we choose a mobility model that is broadly applicable.

2.2. Community network generation

For our Monte Carlo simulations, we generate a series of random metapopulation community networks with the Girvan–Newman algorithm, which generates a random network with a specified size and degree distribution, then removes edges between nodes with a high betweenness factor until the desired community structure emerges [50, 51]. We specify a network of 30 communities, with 30 nodes within each community, and with the average number of edges, between and within communities, randomly selected from 1 to 5.

This algorithm creates community networks with modularities ranging from 0 to 0.9. The majority of nodes have degree of connectivity less than 5, with hub nodes ranging from 20 to 140. About one third of networks have an average clustering coefficient of nearly 0, the rest are distributed between 0 and 0.5, with most networks between 0.25 and 0.4. The average centrality among networks ranges from 0.35 to 0.65, with a peak at 0.5. This algorithm allows for the rapid generation of human mobility networks that bear a resemblance to real-world case studies and is frequently used in the study of infectious diseases [19, 52–54]. Mathematically speaking, the code that executes the SIR simulation treats mobility between nodes in the same community and between nodes identically. There is no mathematical distinction between them.

2.3. Monte Carlo method

After a network has been constructed, a random value of diffusion p is assigned, ranging from 0.1 to 0.5, and movement heterogeneity θ is assigned as 0.5. A population of about 500,000 individuals is assigned to the network, such that the population of each node N is proportional to $k_i^{1+\theta}$, where k_i is the node's degree of connectivity. All individuals are initially classified as Susceptible. The mobility matrix is then defined based on a traffic-dependent model. The recovery rate μ is set to 0.1, the disease latency time is set to 1/2, the fraction percent of individuals who are asymptomatic, or are symptomatic and continue to travel, are set to 0.2 and 0.5, respectively, and the infection rates of asymptomatic individuals were multiplied by 0.5. The infection rate β is assigned over a spatial field with an exponential distribution with mean 0.15, to mimic spatial variability in the reproduction rate of seasonal influenza [55]. In order to ensure that the spatial field is continuous, we apply smoothing and despiking algorithms over the domain to remove any discontinuities [56]. This results in a spatial field with the same exponential distribution, but with a degree of cross-correlation between adjacent points.

Infection rate spatial fields generated with this method tend to feature a few hot spots with multiple incidences of spatial clustering. Based on this spatial field, an infection rate value β_i is assigned to each node.

The epidemic is triggered by the appearance of one infected individual in a random node in the network, weighted by population and the course of the epidemic is simulated using the SEIIIR model described above, using the ODE45 function in MATLAB [57]. In this analysis we explore three different scenarios:

1. The base case, in which the epidemic proceeds unimpeded as per the parameters given in this section.
2. The mitigation case, in which the effective R score (R_t value) of all communities is measured using a Next-Generation matrix. Communities with an R score greater than 1.1 have all mobility within and between adjacent communities reduced by 90%.
3. The vaccination case, in which before the epidemic outbreak begins, a randomly selected 10% of the population, uniformly distributed, is moved from the Susceptible condition to the Recovered condition. A value of 10% is chosen because preliminary research shows that even a small number of vaccinated individuals can have profound effects on the dynamics of the epidemic outbreak under the conditions described above.

Numerical relationships between community structure and a variety of epidemic parameters, such as R_0 value, peak prevalence, and epidemic duration, are evaluated through the Random Forest package in R v4.2.1 [58]. Random Forest is a type of regression model in which decision trees are used to predict a continuous independent variable based on a set of permuted predictor variables. For this project, our Random Forest decision trees consisted of 500 decision branches, at which two variables were considered at each branch to minimize overfitting.

3. Results

3.1. Next-generation matrix for community networks

We begin with a basic metapopulation SIR model, on the assumption that infection rate β varies between nodes but recovery rate μ does not.

$$\begin{aligned}\frac{dS_i}{dt} &= -\beta_i \frac{S_i I_i}{N_i} + \sum_j M_{ji} S_j - \sum_j M_{ij} S_i \\ \frac{dI_i}{dt} &= \beta_i \frac{S_i I_i}{N_i} - \mu I_i + \sum_j M_{ji} I_j - \sum_j M_{ij} I_i \\ \frac{dR_i}{dt} &= \mu I_i + \sum_j M_{ji} R_j - \sum_j M_{ij} R_i\end{aligned}\quad (3)$$

where S_i, I_i, R_i, N_i , are the susceptible, infected, recovered, and total population of node i , β_i is the infection rate in node i , and M_{ij} is the mobility rate from node i to node j . Consider the early epidemic outbreak, assuming $\frac{S_i I_i}{N_i} \sim I_i$, and define the diffusion rate $p_i = \sum_j M_{ij}$. Then the differential equation for I_i can be linearly approximated as

$$\frac{dI_i}{dt} = (\beta_i - \mu) I_i + \sum_j M_{ji} I_j - p_i I_i \quad (4)$$

Next, consider a metapopulation network with community structure, consisting of n nodes divided into m communities. It is assumed, but not required, that each community contains roughly the same number of nodes. Assume that the network graph is strongly connected, meaning there exists a path connecting every pair of nodes, and assume that for every pair of nodes in a community, there exists a path within the community that connects those nodes. Also assume that each node has a non-zero population, mobility rate, and infection rate. Then M is a square block matrix of $m \times m$ blocks and $n \times n$ elements. Let the matrix M_{ij} represent movement within each community, and N_{ij} represent movement from community i to community j :

$$M = \begin{bmatrix} M_1 & N_{12} & N_{13} & \dots \\ N_{21} & M_2 & N_{23} & \dots \\ N_{31} & N_{32} & M_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (5)$$

The calculation of the next-generation matrix involves decomposing the infection rate into local infection growth and the transfer of infected individuals between groups [59]. Here we define each node as a group, and we write the infection rate equation as

$$\frac{dI_i}{dt} = F_i - V_i \quad (6)$$

where I_i is the vector of infected individuals by node, and F_i and V_i are vectors defined as $F_i = \text{diag}(\beta) I_i$ and $V_i = \text{diag}(\mu) I_i - \sum_j M_{ji} I_j + p_i I_i$. The Jacobian matrices of F and V , therefore, are

$$F_{ij} = \frac{\partial F_i}{\partial I_j} = \text{diag}(\beta) \quad (7)$$

$$V_{ij} = \frac{\partial V_i}{\partial I_j} = \text{diag}(\mu \mathbf{1} + \mathbf{p}) - M_{ji}, \quad (8)$$

where $\mathbf{p}$ is the vector of nodal diffusion rates in the network. Note that the matrix M has no diagonal elements, since it only represents movement between nodes. Therefore, we can decompose the matrix V into its diagonal and off-diagonal parts:

$$V = D - M^T, \text{ where } D = \text{diag}(\mu \mathbf{1} + \mathbf{p}) \quad (9)$$

Given a next-generation matrix, the reproduction number R for the entire network is defined as $\rho(FV^{-1})$, where ρ represents the spectral radius function. Assuming all nodes have a diffusion rate greater than 0, then D is invertible, with inverse matrix

$$D^{-1} = \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1}. \quad (10)$$

Let us assume that V , but not necessarily M^T , is invertible. Then according to matrix perturbation theory [60], V^{-1} can be calculated to second-order approximation as:

$$\begin{aligned}V^{-1} &= (D - M^T)^{-1} \\ V^{-1} &= D^{-1} + D^{-1} M^T D^{-1} + (D^{-1} M^T)^2 D^{-1} + \dots \\ V^{-1} &= \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} + \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} M^T \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} \\ &\quad + [\text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} M^T]^2 \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} + \dots\end{aligned}\quad (11)$$

Then the next-generation matrix is computed to second-order approximation as

$$\begin{aligned}FV^{-1} &= \text{diag}(\beta) \left[\text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} + \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} M^T \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} \right. \\ &\quad \left. + [\text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} M^T]^2 \text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} + \dots \right] \quad (12)\end{aligned}$$

If we assume that $\mathbf{p}$ is constant, then $\text{diag}(\mu \mathbf{1} + \mathbf{p})^{-1} = \frac{1}{\mu + p} \mathbf{I}$, and this formula simplifies considerably:

$$FV^{-1} = \text{diag}(\beta) \left[\frac{1}{\mu + p} \mathbf{I} + \frac{1}{(\mu + p)^2} M^T + \frac{1}{(\mu + p)^3} (M^T)^2 + \dots \right] \quad (13)$$

Consider a homogeneous mobility model, such that $M_{ij} = \frac{p}{k_i} A_{ij}$, or $M = p \text{diag}(\mathbf{k})^{-1} A$, where A is the adjacency matrix and k_i is the degree of connectivity of node i . Note that since M has a zero diagonal and all rows of M add to p , the spectral radius of M , and therefore M^T , has an upper limit of p . Also note that A is symmetric. From this we finally derive a formula for the next-generation matrix that is dependent only on the infection rates and connections of the network nodes.

$$\begin{aligned}FV^{-1} &= \text{diag}(\beta) \left[\frac{1}{\mu + p} \mathbf{I} + \frac{p}{(\mu + p)^2} [A \text{diag}(\mathbf{k})^{-1}] \right. \\ &\quad \left. + \frac{p^2}{(\mu + p)^3} [A \text{diag}(\mathbf{k})^{-1}]^2 + \dots \right] \quad (14)\end{aligned}$$

Unfortunately, the spectral radius of FV^{-1} cannot be directly computed. The diagonal matrices have the unit vectors as eigenvectors, but we cannot assume as such for A . We do know from the Perron-Frobenius theorem [61] that the spectral radius of A is bounded above by the maximum nodal degree of connectivity, and we know that the

eigenvector corresponding to the largest eigenvalue of A is a commonly used metric of nodal centrality [8].

Utilizing the properties of the spectral radius, matrix 2-norm, and diagonal matrices [62], we can set an upper limit for the spectral radius of the next-generation matrix:

$$\begin{aligned} \rho(FV^{-1}) &\leq \|FV^{-1}\| \\ &\leq \|F\| \|V^{-1}\| \\ &= \|\text{diag}(\beta)\| \left\| \frac{1}{\mu+p} I + \frac{p}{(\mu+p)^2} [A \text{diag}(\mathbf{k})^{-1}] \right. \\ &\quad \left. + \frac{p^2}{(\mu+p)^3} [A \text{diag}(\mathbf{k})^{-1}]^2 + \dots \right\| \\ &\leq \|\text{diag}(\beta)\| \left\| \frac{1}{\mu+p} \|I\| + \frac{p}{(\mu+p)^2} \|A \text{diag}(\mathbf{k})^{-1}\| \right. \\ &\quad \left. + \frac{p^2}{(\mu+p)^3} \|A \text{diag}(\mathbf{k})^{-1}\|^2 + \dots \right\| \end{aligned} \quad (15)$$

For this manuscript we are specifically interested in the effects of community structure on the epidemic reproduction rate of individual communities. To that end, for a block-diagonally dominant matrix A , define the notation $[A]_i$ as the specific submatrix of A whose rows and columns correspond to community i [63]. We similarly define the notation $[\mathbf{v}]_i$ for a vector $\mathbf{v}$. More precisely, we define

$$[A]_i = E^T A E \text{ and } [\mathbf{v}]_i = E^T \mathbf{v}, \text{ where } E = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad (16)$$

where the rows of the identity matrix in E corresponds to the nodes within the community of interest.

Finally, we define R_i as the reproduction number for that specific community, or

$$R_i = \rho([FV^{-1}]_i) \quad (17)$$

Utilizing the properties of diagonal matrices, we can evaluate this reproduction number as:

$$\begin{aligned} R_i &= \rho([FV^{-1}]_i) \\ &= \rho\left(\text{diag}([\beta]_i) \left[\frac{1}{\mu+p} I + \frac{p}{(\mu+p)^2} [A \text{diag}(\mathbf{k})^{-1}]_i \right. \right. \\ &\quad \left. \left. + \frac{p^2}{(\mu+p)^3} [(A \text{diag}(\mathbf{k})^{-1})^2]_i + \dots \right] \right) \end{aligned} \quad (18)$$

and the upper limit of the reproduction number for a specific community is given by

$$\begin{aligned} R_i &\leq \max([\beta]_i) \left\| \frac{1}{\mu+p} I + \frac{p}{(\mu+p)^2} [A \text{diag}(\mathbf{k})^{-1}]_i \right\| \\ &\quad + \frac{p^2}{(\mu+p)^3} \left\| [(A \text{diag}(\mathbf{k})^{-1})^2]_i \right\| + \dots \end{aligned} \quad (19)$$

The question of how community structure affects the reproduction number for a specific community is reduced to how community structure affects the matrix norm of $[A \text{diag}(\mathbf{k})^{-1}]_i^k$ for each integer k .

The submatrix $[A]_i$ is just the adjacency matrix of the nodes within community i , therefore the community structure of the broader network has no effect on it. Multiplying A by the matrix $\text{diag}(\mathbf{k})^{-1}$ has the effect of dividing each row of A by the degree of the corresponding node and does not affect which elements of the matrix are zero or non-zero.

The square of the adjacency matrix A^2 indicates the number of paths of length 2 connecting each pair of nodes [64]. The submatrix $[A^2]_i$ is

generally not equal to the submatrix $[A]_i^2$, but they are equal for a specific community if there are no paths of length 2 that connect two nodes within that community but pass through a different community. The same is true for $[A^3]_i$, and so on. The matrix $[(A \text{diag}(\mathbf{k})^{-1})^n]_i$ may have miniscule differences on the rows and columns that correspond to the bridge nodes of the community, since their degrees of connectivity are reduced when the communities are severed from each other. However, numerical tests show that for a sufficiently large community network, this matrix does exhibit the property $[(A \text{diag}(\mathbf{k})^{-1})^n]_i = ([A]_i \text{diag}([\mathbf{k}]_i)^{-1})^n$ within numerical error if adjacent communities have only one bridge connection between them.

This brings us to our main result. We quantify the effect of community structure on the reproduction rate of a community by comparing the difference between the actual value of R_i and the theoretical value of R_i if all links to neighboring communities were severed. Then for a specific community, let n be the length of the shortest path that connects two nodes within the same community but passes through an adjacent community. Then the community structure of the network causes a perturbation of $\max([\beta]_i) \left(\frac{p^n}{(\mu+p)^{n+1}} \right)$ on the upper limit of the reproduction rate of that community. It is important to note that this conclusion only applies in the short-term after the introduction of the disease outbreak. If any node in the network has an R_0 value greater than 1, infection cases will eventually rise everywhere due solely to the migration of infected individuals. However, in highly modular community networks this may not occur until herd immunity has taken affect in some regions of the network.

Adapting this formula to the SEIIR model is straightforward. It is sufficient to replace β in the above equation with $\beta(1 + (r_\beta - 1)p_a)$ and any instances of p with $p[1 - (1 - p_a)(1 - p_i)]$.

3.2. Numerical demonstration of community structure and spectral radius

To demonstrate the effect of network modularity on the spectral radius of the next-generation matrix, we start with the scale-free network of 30 communities, each community consisting of a scale-free network of 30 nodes, for a total of 900 nodes [65]. An example of this type of network is shown in Fig. 1a. Infection rates are assigned to each node independently as a Weibull(1.2,1) distribution, to mimic spatial variability in the reproduction rate of seasonal influenza [55]. Recovery rates are set to 1, diffusion rates are set to either 0.05, 0.1, 0.5, or 1.0, and all nodal pairs demonstrate homogeneous mobility. The next-generation matrix and its spectral radius R_i is computed for each individual community, which for a weakly connected, highly modular network we expect to be equal to the highest β/μ value in each community. After computing these values, we subsequently add one bridge connection to each adjacent community pair and recompute the R_i values (Fig. 1b). We repeat this until the modularity of the network converges to a value close to zero. In addition, we ran multiple metapopulation SIR simulations on these networks and used an exponential curve fitting method to verify that the actual reproduction rate of the epidemic is equal to the computed R_i values, within numerical error [66]. We also validated the main result of the previous section by comparing the R_i values of the original community networks to the modified networks with additional bridge connections. We ran a sensitivity analysis to ensure that our choices of parameters did not significantly impact the outcome of these results.

Fig. 2 shows the results of this analysis. Without modifying the infection rates of any nodes, or the connections within communities, the number of bridge connections between communities has a strong inverse correlation to the median and maximum reproduction rates among the communities. In fact, the addition of bridge connections, or the reduction of community modularity, can decrease the median R_i value from nearly 1.5 to below 1, bringing several communities below the threshold of an epidemic outbreak. The relationship between

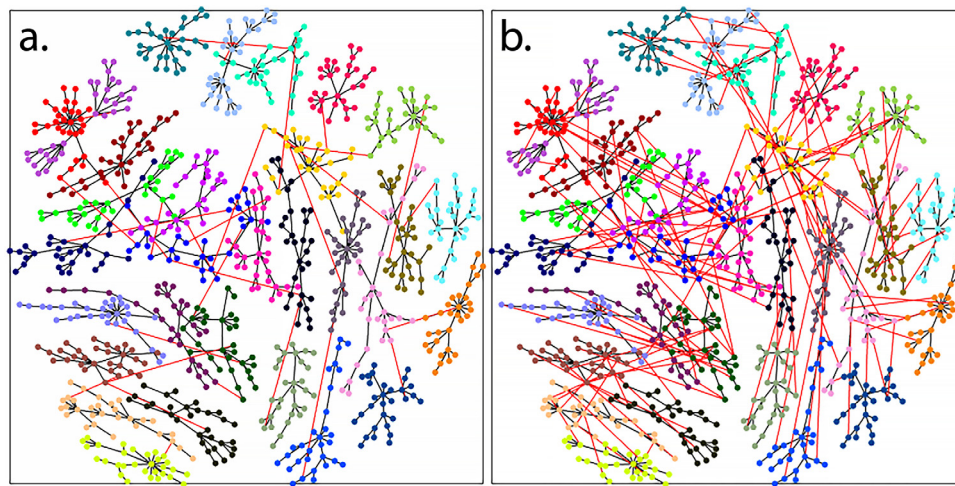


Fig. 1. (a) A scale-free community of scale-free networks, with each adjacent community connected by a single bridge. (b) The same community network but with five bridges per adjacent community. Each community is highlighted in a different color, connections within communities are indicated by black edges, and connections between communities are indicated by red edges. Note that a metapopulation SIR simulation treats each node as its own unit of population, and communities arise naturally from how the nodes are connected.

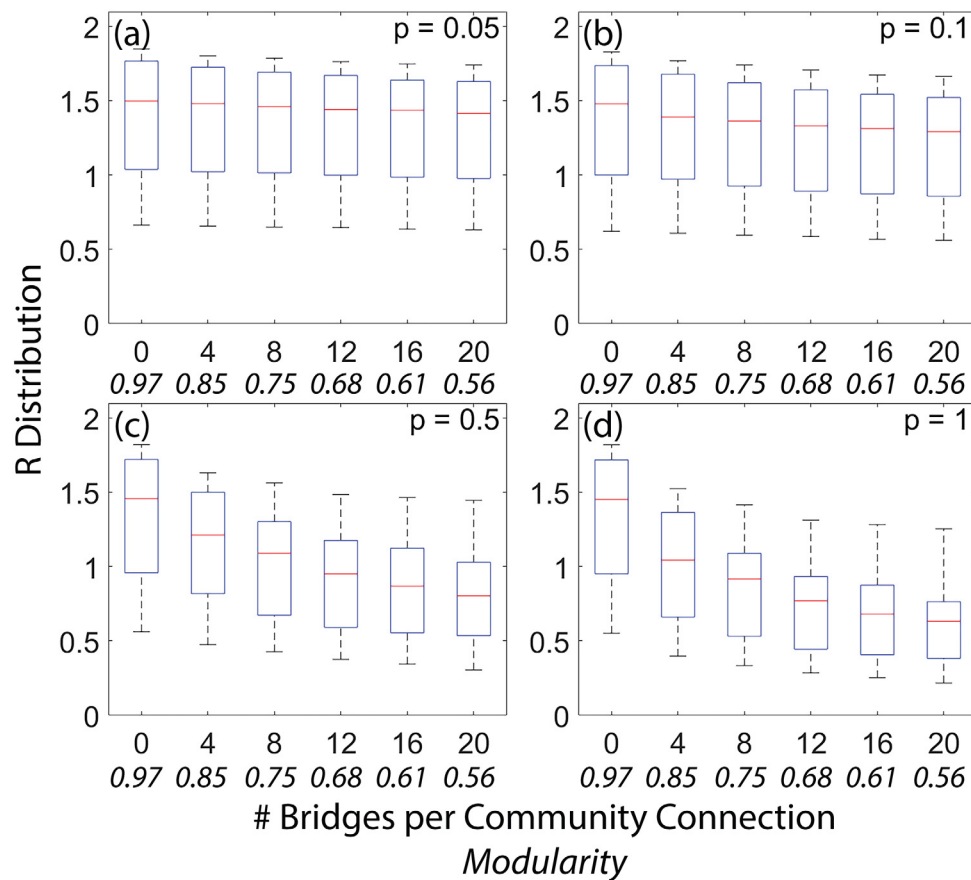


Fig. 2. A box plot of reproduction rate R for all 30 communities in the network. Infection rates are assigned as a Weibull distribution with mean 0.12 and shape factor 2, and recovery rates are set at 0.1. The x-axis indicates the number of bridge connections per community pair, and each subplot represents a different host diffusion rate. Note that a community with more bridge connections tends to have a lower modularity, and vice versa.

network modularity and reproduction rate has a higher correlation when diffusion is high. Conversely, when diffusion is less than the recovery rate, there is almost no relationship between modularity and reproduction rate. The reproduction rate of the entire network is equal to the highest reproduction rate of any individual community, and it is likewise decreased by the addition of bridge connections and by increased diffusion.

It may seem unintuitive that more highly connected, highly mobile communities tend to have lower epidemic reproduction rates, but it is consistent with the conclusions of the analytical model shown above and with previous research [67]. The cause of this phenomenon is that in high-risk, weakly connected community networks, infected people tend to stay within their communities, allowing the number of infected individuals to grow quickly. In highly connected community networks,

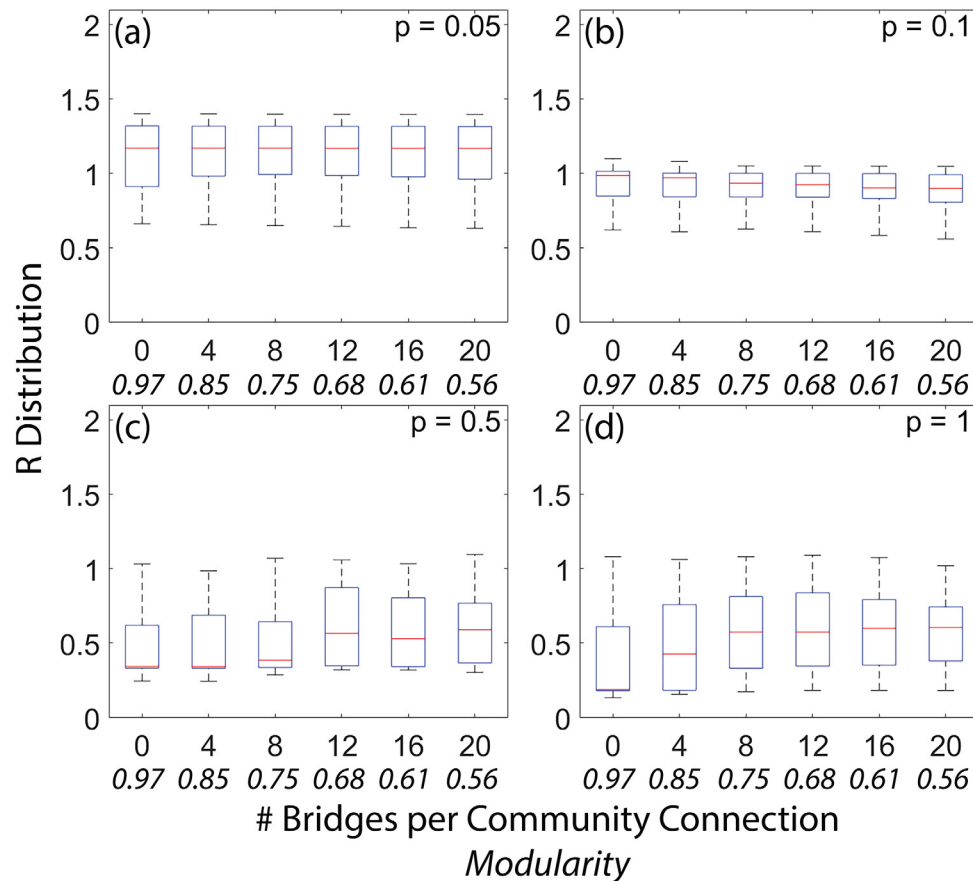


Fig. 3. A box plot of reproduction rate R for all 30 communities in the network. These data represent the same simulation as in Fig. 2, but with a mitigation step included. All communities with a next-generation matrix R score of 1.1 or greater have movement within the community and between adjacent communities reduced to 10%.

infected individuals tend to distribute themselves evenly across all communities, and the growth rate for the network is reduced. Note that the minimum reproduction rate is not significantly affected by the community structure or by mobility rates. Communities that are not at risk of an epidemic outbreak will not sustain an outbreak regardless of connections to neighboring communities.

These results do not imply that closing connections between communities (for example, an interstate travel ban) actually increases the risk of outbreaks, since these results are only based on the presence or absence of bridge connections, not how heavily they are trafficked, and does not consider other preventative behaviors within those communities. Reducing the rate of travel is mathematically analogous to reducing the diffusion rate out of individual nodes or communities, not increasing their modularity. A targeted policy that reduces travel within and out of high-risk communities is still expected to reduce the reproduction rate of the network.

To demonstrate, we run the same analysis on the same network with the same allocation of infection and mobility rates, but we now impose a mitigation strategy. Specifically, communities with an R_i value greater than 1.1 have their mobility within the community and among adjacent communities reduced to 10%. The results of this analysis are shown in Fig. 3. In all cases, the median and maximum R_i values of communities are reduced, but in the low diffusion models, the communities with low R_i values are not affected. This implies that in low diffusion networks, mitigation techniques are only effective for the communities in which they are applied. In the high diffusion networks, on the other hand, all communities show a reduction in R_i values, even the communities to which no mitigation was applied.

Mitigation is most effective in networks with few bridge nodes and high modularity, up to a 66% reduction in R_i values in the highest modularity networks. However, mitigation has a negligible

effect on networks with several bridge nodes and low modularity. As a result, there is a positive correlation between the number of bridge connections and the distribution of R_i values, a reverse from the case with no mitigation. The implication of this analysis is that mitigation techniques, even if they are only targeted at selected communities, can have wide-reaching benefits over the entire network, but they are most effective in networks with high diffusion and high modularity.

3.3. Monte Carlo simulations

Next-generation matrices are useful for observing how network community structure affects the epidemic reproduction rate in individual communities. Other metrics of interest, including peak prevalence, the area of the disease spread, and the time for the epidemic to establish itself in communities, require the use of Monte Carlo simulations for further study.

We ran three separate scenarios, as described in the Methods section, with a range of parameters broad enough to encompass many different infection and human movement scenarios. We also ran a sensitivity analysis, not shown, to ensure that the values we chose for these scenarios did not disproportionately affect the results of this study. Table 1 gives a description of the variables we tested in these simulations. Fig. 4 shows the response curves, computed using the randomForest package in R, of the peak fraction of individuals infected and the fraction of nodes that become infected for the entire network, in response to the network's maximum R value, variance in R value, modularity, diffusion, mean clustering coefficient among communities, and mean centrality among communities [68]. The Random Forest model processed 500 decision trees with 2 variables randomly sampled at each branch split, and we configured the algorithm to calculate proximity on all rows of the tree to ensure that the response curves

Table 1

A table of network variables used in the Random Forest models shown in Figs. 4–6.

Variable	Description
Peak infected	Fraction of individuals who become infected over the course of the epidemic
Area of infection	Fraction of nodes with at least one local infection case
Maximum R variance of R	R calculated as $\frac{\mu}{\mu} (1 + (r_p - 1) p_a)$
Modularity	A measure of the level of separation between communities, ranging from $-1/2$ to 1. Equal to the fraction of the edges that fall within each community minus the expected fraction if edges were assigned at random among all nodes in the network without regard to community. Inversely related to the average number of bridge connections between neighboring communities [50].
Diffusion	Calculated as $p[1 - (1 - p_a)(1 - p_t)]$
Degree	Number of adjacent communities
Clustering	Fraction of connected communities that are also connected to each other, forming a triangle
Centrality	Reciprocal of communities' mean distance to all other communities in the network
# Bridges	Number of bridge connections between each pair of connected communities
Inner degree	Average degree of connectivity among all nodes in the community
Inner clustering	Average clustering coefficient among all nodes in the community
Inner centrality	Average centrality among all nodes in the community

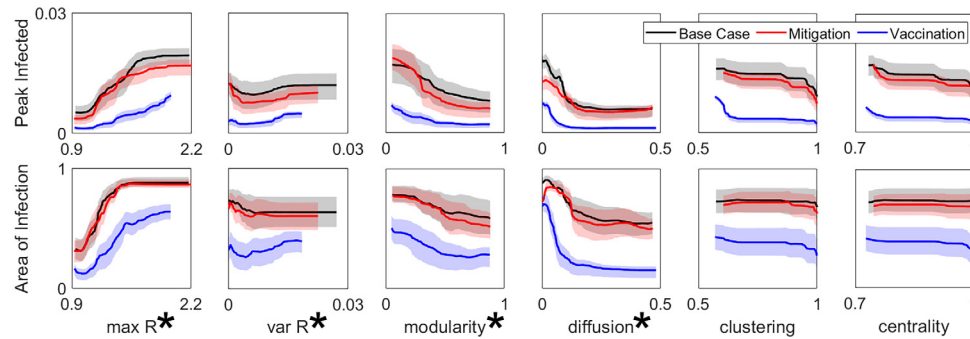


Fig. 4. Response curves of Random Forest model comparing the peak number of individuals infected and the fraction of nodes with at least one local infection case among community networks. The three scenarios shown are the base case, the mitigation case, and the vaccination case as described in the Methods. Statistically significant predictor variables (% increase mean-square-error > 40%) are noted with an asterisk. Additional details are included in the Supplement.

were continuous. Goodness-of-fit data, variable importance plots, and ANOVA data are included in the Supplement, but in general the Random Forest diagnostics are very favorable, with mean squared errors of about 0.05. Fig. 5 shows similar response curves of peak infection prevalence, and the area of epidemic spread compared to the properties of individual communities.

For both peak infected and area of infection response variables, maximum R , variance of R , modularity, and diffusion were all found to be statistically significant variables. The ranking of variable importance, in terms of % Increase in mean-square-error (%IncMSE), depended on the specific epidemic scenario. In the base case and vaccination cases, diffusion was the most important variable for peak prevalence, whereas in the mitigation case maximum R and modularity were tied for most important. For the area of infection, maximum R was the most important variable for the base case and mitigation cases, whereas diffusion was the most important in the vaccination case. The results for the individual community analysis are similar, except that inner degree of connectivity is also an important factor. In general, communities with less inner-connectivity tend to have higher peak prevalence, although the difference is minuscule. In all cases, clustering and centrality had no importance in the random forest models.

3.4. Super-spreader capacity and spatial scaling

In [44], a formula is derived to determine the risk that a given node in a metapopulation network may become a super-spreader event. A node's super-spreader capacity is based on the expected number of neighboring nodes to which it would spread the infection, should an outbreak occur there. Suppose a certain node, representing a subpopulation, is experiencing an epidemic outbreak. Given the R value of this node, we can estimate the maximum fraction of individuals who will

become infected, the peak infection rate, with the function $\alpha(R)$. This function is equal to 0 if R is less than 1 and equal to $1 + \frac{W(-Re^{-R})}{R}$ if R is greater than or equal to 1, where $W()$ is the Lambert product log function [69]. Also, given mobility rates between this node and all of its neighboring nodes, we can compute the probability that any infected individuals will migrate to a neighboring node, potentially spreading the epidemic.

Among these neighboring nodes, some of them will have R values significantly less than 1, in which case they have no chance of spreading the epidemic. Some of these nodes will have R values significantly greater than 1, in which case they will almost definitely spread the pandemic. For neighboring nodes with R values close to 1, the probability that at least one infected individual will migrate to that node and spread the infection can be computed. With this in mind, we can estimate if a given node were to become infected, how many neighboring nodes will also become infected.

Consider a node with a given reproduction number R and degree of connectivity k . This node has k neighboring nodes, indicated by an index $j = 1, 2, \dots, k$, each of which also have reproduction numbers R_j and degrees of connectivity k_j . Assuming a traffic-dependent mobility model (Eq. (2)) and constant diffusion, the super-spreader capacity for the population node in question is calculated as:

$$SSC = \sum_{j, R_j \geq 1} \left[1 - (R_j)^{-p \frac{\langle k \rangle}{\langle k^2 \rangle} \frac{\bar{N}}{\mu} (k_i k_j)^\theta \alpha(R_i)} \right] \quad (20)$$

where R_i and k_i are the reproduction rate and degree of connectivity for the node in question, R_j and k_j are the reproduction rate and degree of connectivity of the neighboring nodes, p is diffusion, $\bar{N}$ is the average nodal population, μ is the recovery rate, and θ is the heterogeneity factor (typically equal to 0.5). The series is summed over

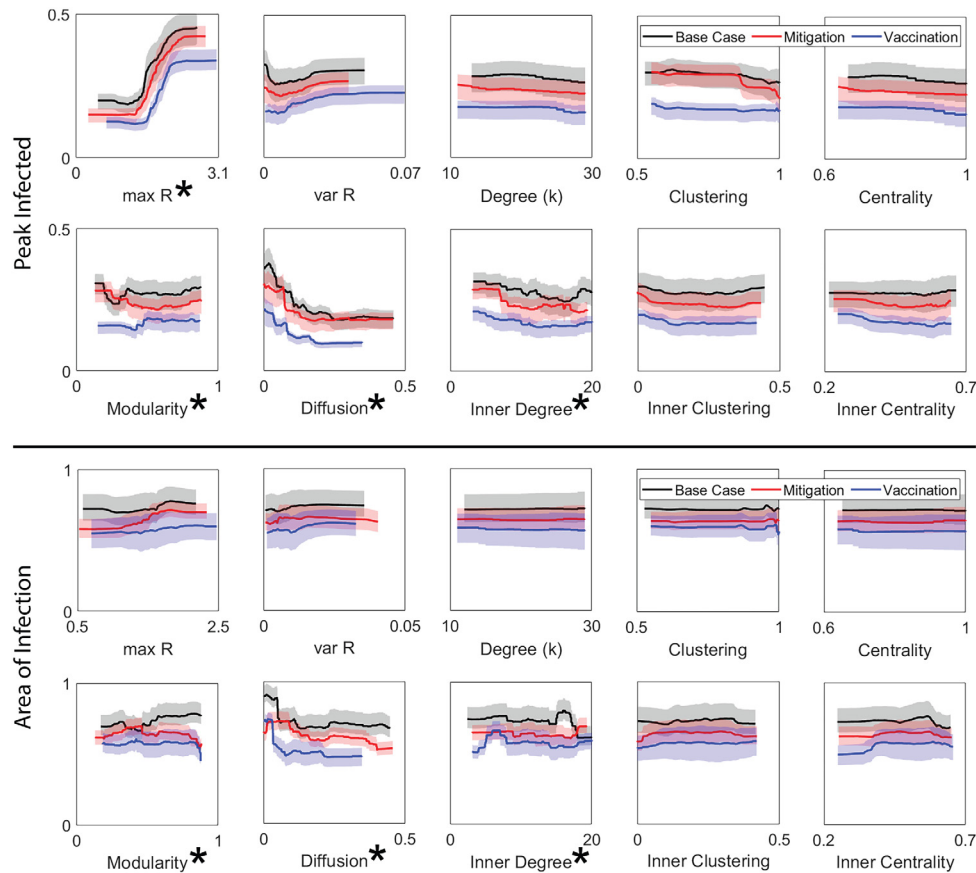


Fig. 5. Response curves of Random Forest model comparing the peak number of individuals infected and the fraction of nodes with at least one local infection case among individual communities. The three scenarios shown are the base case, the mitigation case, and the vaccination case as described in the Methods. Statistically significant predictor variables (% increase mean-square-error > 40%) are noted with an asterisk. Additional details are included in the Supplement.

all neighboring nodes j whose R_j value is greater than or equal to 1. The summation is computed over all neighboring nodes for which R values are at least 1.

The advantage of this formula is that because it is dependent on the average population of each node, it should scale easily with the spatial resolution of the metapopulation network. In order to test the precision of this algorithm across spatial scales, we compute the super-spreader capacity of each node in a community patch network and assign them a risk index based on percentile, where 1 is the highest risk of a super-spreader event and 0 is the lowest risk. An example of this risk map is shown in Fig. 6a. We then run the same algorithm on the community-scale network, computing their reproduction rates using a next-generation matrix, as shown in Fig. 6b.

Since the next-generation matrix is based on the maximum infection rate of each node in the community, the community-scale risk index is almost always an overestimate of the risk associated with each node in the network. Typically, the risk index for the community is on par with the highest risk index node in that community. On average, the community-scale model overestimated super-spreader capacity by about 8%, and the overestimation error was highest in communities with a reproduction rate between 1 and 2. Communities with high degrees of connectivity and high centrality, that is communities with a small-world structure, tended to have a higher rate of error [70]. There was a weak correlation between the network modularity and the overestimation of super-spreader capacity, but error tended to be greater in networks with a small number of bridge connections and high modularity.

4. Discussion

The analytical and computational results of this study support our hypothesis, that in order of importance, infection rates are a primary effect on the magnitude of a disease epidemic, human diffusion is a secondary effect [71,72], and community structure, measured as modularity or the number of bridge connections between communities, is a tertiary effect [73]. According to our results, community structure is not the most statistically important variable for determining epidemic reproduction rates, but it can make a very significant difference in networks with highly variable infection rates and high diffusion, even determining whether or not an outbreak occurs [21,74]. These results build on previous research that compared the heterogeneity of community networks to the rate of epidemic spread [63,75]. Most interestingly, the modularity of the network may determine the effectiveness of different mitigation strategies [67,76]. A pre-outbreak vaccination strategy may be equally effective regardless of the network structure, but a travel restriction strategy tends to be more effective in high modularity, low diffusion networks [77,78].

In the base case scenario, the most important variables in determining peak prevalence were maximum R , followed by diffusion and modularity, and the variance of R [79]. Note that in the Random Forest model, the effects of variance of R are computed independently of maximum R . Therefore, if two networks have the same maximum R but different variance in R , the network with higher variance of R will result in more infected individuals, which concurs with our analytical results. On the other hand, clustering and centrality were relatively insignificant for determining peak prevalence, consistent with previous research that the structure of individual communities may not affect peak infection rates [80,81]. The ranking of variable importance was

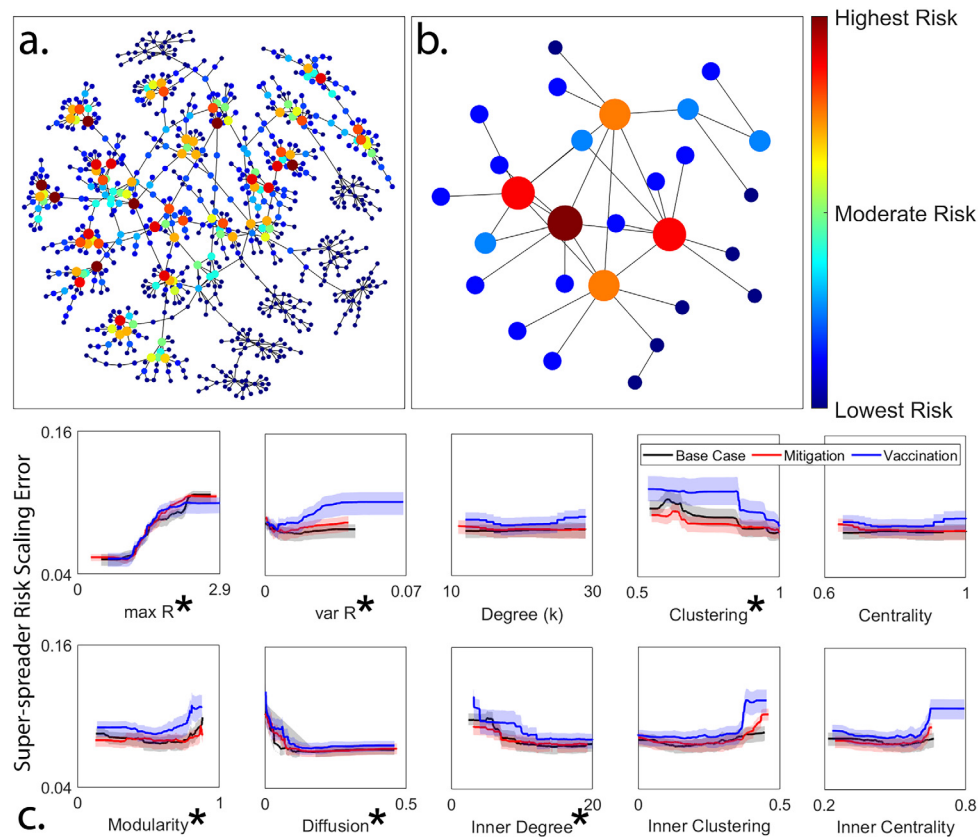


Fig. 6. Super-spreader capacity map for a randomly generated community network, computed on (a) the nodal scale, (b) the community scale. The bottom figure shows the response curves for the root-mean-square error in risk estimate between (a) and (b). Statistically significant predictor variables (% increase mean-square-error > 40%) are noted with an asterisk. Additional details are included in the Supplement.

the same in the mitigation scenario, but in the vaccination scenario diffusion was by far the most important predictor variable, followed by maximum R and modularity.

In terms of the area affected by the epidemic, in the base case and mitigation scenarios maximum R and diffusion are still the most important factors, followed by modularity and variance of R . In the vaccination scenario, however, diffusion is significantly more important than maximum R . In other words, diffusion, arguably the only factor that the host population can completely control, is very important for determining the spatial extent of the epidemic but not as important as maximum R for the number of individuals who become infected [46, 71,82].

The mitigation strategy was most effective at reducing peak infection in networks with a high maximum R and high modularity. Conversely, the vaccination strategy was most effective in networks with low modularity and high diffusion. This may imply that the ideal strategy depends on the nature of the epidemic, and whether its propagation is due to its infectiousness or host mobility rate. Highly infectious epidemics are better mitigated by restricting movement, and high movement communities (or diseases with a high rate of asymptomatic cases) show a greater benefit from vaccination [15,83–86]. It is important to reiterate that the vaccination model used in this study is highly idealized, and it shows a remarkably strong effect by vaccinating 10% of the population before the initial epidemic outbreak. Additional simulation tests show that if vaccinations were specifically targeted to 10% of the population in nodes with the highest infection rates, the results would be qualitatively similar, but with about an additional 15% reduction in infection cases. We also observe similar rates of effectiveness in non-idealized cases, such as when vaccination occurs later in the epidemic outbreak.

Observing individual communities, maximum R and diffusion were again the most significant variables in determining peak prevalence

and the area of the infection. Community network structure, in terms of modularity, clustering, and degree of connectivity were somewhat significant, and the properties of the individual communities, such as their place in the network and the arrangement of their internal nodes, had very little effect on the magnitude of the epidemic. The implication of these results is that the arrangement of communities and the number of connections between them plays an important role in the dynamics of the epidemic across the entire network, but the network structures of individual communities, besides their individual infection rates, have little bearing on the magnitude of the epidemic within those communities [87,88].

In terms of measuring super-spreader capacity, the overestimation of risk at the community-scale resolution is typically no more than 15%, compared to the measure of risk at the node-scale resolution. The level of error is highest in communities that have an R value greater than 1, as well as communities with low diffusion, low clustering, and low internal connectivity. This implies that on networks in which the communities form a small-world network structure, there may be some merit in dividing the communities into individual nodes. Surprisingly, super-spreader prediction error did not increase significantly when there was large variance in R among nodes. This means, for example, that a community that consists of a single high-risk population center and large low risk surrounding area may still be represented as a single population node without a significant degree of prediction error [89].

It is often but not always necessary to use the finest possible spatial-scale resolution when modeling disease epidemics on patch networks [90–92]. The required resolution depends on the properties of the human mobility network as well as the variables researchers are interested in studying, and disparities in resolution can be resolved through ecological regression models [93]. A network model in which communities are represented as single, homogeneous nodes is analogous to an extremely high modularity community network, so fine-scale

models are most necessary in networks with loosely separated communities and a large number of bridge connections [8,94]. When predicting values of R , the most used metric of disease magnitude in epidemiology [95], the community structure is a very important factor in networks with high rates of mobility and high variance in infection rates among individual nodes, and in extreme cases the spatial resolution may affect estimates of R by about a factor of two. This only occurs in very idealized models that do not consider any effects of human movement mitigation. In more realistic epidemic models, however, the failure to not consider community structure and heterogeneous infection rates may cause the researcher to significantly underestimate the R value of their study site.

The definition of a “high diffusion network” is contingent on the type of mobility involved [45,96]. This study distinguishes between local and inter-community travel, as well as mobility by uninfected individuals and asymptomatic and symptomatic infected individuals. In general, the most effective method of reducing the spread of epidemics is to reduce their infection rates, either through vaccinations or environmental control, but if that is not possible, reducing the contribution to diffusion by symptomatic individuals is the next most effective intervention [97,98]. This is especially true in high modularity communities, where blocking a single route of transportation may significantly reduce diffusion into and out of a given community.

It is generally not feasible for local government officials to change the modularity of their network, besides significantly reducing air travel and interstate commuting, but network modularity should be an important indicator to epidemic modelers [99,100]. An important epidemiological conclusion that can be derived from this study is that higher spatial resolutions are necessary to accurately model low modularity communities. However, individual community properties such as clustering and centrality do not generally affect whether it is necessary to model them as homogeneous nodes or as communities. For example, in an epidemic model of the United States, the ideal spatial resolution for every metropolitan area would be the same, and an individual node in the Chicago and New York City community networks should represent the same number of people. This also implies that the ecological fallacy and modifiable areal unit problem need not be major design considerations, if the spatial resolution of the epidemic model is kept consistent throughout the course of the scientific study.

The traffic-dependent mobility model used in this analysis is simplistic, and the consideration of various types of movement, human behaviors, and mitigation strategies may affect the presented results [101]. This study also assumes a disease with a basic R_0 value close to 1 (such as influenza, ebola, and tuberculosis), for which there is a statistically strong correlation between infection rates, infection prevalence, and epidemic spread area, despite their very different modes of transmission and virulence [102]. For a highly infectious disease such as measles or mumps, with R_0 values greater than 10, we would paradoxically expect diffusion and modularity to play a more important role than infection rate for determining the magnitude of the outbreak [103]. The same is also true if R_0 is much less than 1, although in those cases an outbreak is very unlikely.

5. Conclusion

This study explores the relationships between community networks and patch networks in modeling disease epidemics. Through an analytical and computational framework, we explore how community structure affects the epidemic reproduction rate and magnitude, as well as the effectiveness of various mitigation strategies. As a result, we inform the circumstances in which modeling a community network adds explanatory power. A key finding of this study is that community structure does matter, although generally less so than infection rates and human mobility, and it matters most in networks that have low modularity and a small-world configuration, meaning that communities are isolated and decentralized. In addition, the decision whether to

model populations as nodes or as communities, in terms of balancing numerical accuracy and data collection costs, depends almost entirely on the properties of the entire community network, rather than the properties of individual communities. There are no perfect simulations of real-world phenomena, and researchers will always be required to make subjective choices when designing a simulation model. The analysis presented in this article demonstrates that some disease modeling studies may be improved by gathering fine-scale information during data collection, and it helps inform under what circumstances this extra effort would improve epidemic forecasting.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgments

This project was funded by NSF CNH, USA award #1824961 and USDA National Institute of Food and Agriculture, USA, Hatch Project Number ME021826 through the Maine Agricultural and Forest Experiment Station. Thank you to our collaborators Brian Allan, Cole Butler, and Andrew Mackay for their advice and guidance throughout the course of this project.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.mbs.2023.108996>.

References

- [1] K.L. Gage, T.R. Burkot, R.J. Eisen, E.B. Hayes, Climate and vectorborne diseases, *Am. J. Prev. Med.* 35 (5) (2008) 436–450, <http://dx.doi.org/10.1016/j.amepre.2008.08.030>.
- [2] Z. Jin, G. Sun, H. Zhu, Epidemic models for complex networks with demographics, *Math. Biosci. Eng.* 11 (6) (2014) 1295–1317, <http://dx.doi.org/10.3934/mbe.2014.11.1295>.
- [3] N. Afshordi, B. P. Holder, M. Bahrami, D. Lichtblau, Diverse local epidemics reveal the distinct effects of population density, demographics, climate, depletion of susceptibles, and intervention in the first wave of COVID-19 in the United States, *Fields Inst. Commun.* 85 (2022) 1–23, http://dx.doi.org/10.1007/978-3-030-85053-1_1.
- [4] L. Hufnagel, D. Brockmann, T. Geisel, Forecast and control of epidemics in a globalized world, *Proc. Natl. Acad. Sci. USA* 101 (42) (2004) 15124–15129, <http://dx.doi.org/10.1073/pnas.0308344101>.
- [5] D. Guo, K.C. Li, T.R. Peters, B.M. Snively, K.A. Poehling, X. Zhou, Multi-scale modeling for the transmission of influenza and the evaluation of interventions toward it, *Sci. Rep.* 5 (2015) 1–9, <http://dx.doi.org/10.1038/srep08980>.
- [6] A. Gikas, J. Padiaditis, J.A. Papadakis, J. Starakis, S. Levidiotou, P. Nikolaidis, G. Kioumis, E. Maltezos, M. Lazanas, E. Anevlavis, M. Roubelaki, Y. Tselentis, M. Maragos, V. Nikiforou, V. Barba, S. Metalidis, P. Kolaras, E. Koumentaki, I. Kokkinou, C. Nikolopoulou, G. Chrysos, S. Vasilogiannaki, S. Katanakis, E. Archontidou, D. Glaros, P. Peichaberis, I. Liouris, V. Karagianni, P. Drandakis, G. Troulakis, A. Drandaki, N. Apidianaki, Prevalence study of hospital-acquired infections in 14 Greek hospitals: Planning from the local to the national surveillance level, *J. Hosp. Infect.* 50 (4) (2002) 269–275, <http://dx.doi.org/10.1053/jhin.2002.1181>.
- [7] X. Qian, S.V. Ukkusuri, Connecting urban transportation systems with the spread of infectious diseases: A Trans-SEIR modeling approach, *Transp. Res. B* 145 (2021) 185–211, <http://dx.doi.org/10.1016/j.trb.2021.01.008>.
- [8] M.E. Newman, Modularity and community structure in networks, *Proc. Natl. Acad. Sci. USA* 103 (23) (2006) 8577–8582, <http://dx.doi.org/10.1073/pnas.0601602103>.
- [9] S. Fortunato, Community detection in graphs, *Phys. Rep.* 486 (3–5) (2010) 75–174, <http://dx.doi.org/10.1016/j.physrep.2009.11.002>.

- [10] M. Girvan, M.E. Newman, Community structure in social and biological networks, *Proc. Natl. Acad. Sci. USA* 99 (12) (2002) 7821–7826, <http://dx.doi.org/10.1073/pnas.122653799>.
- [11] Y.Y. Ahn, J.P. Bagrow, S. Lehmann, Link communities reveal multiscale complexity in networks, *Nature* 466 (7307) (2010) 761–764, <http://dx.doi.org/10.1038/nature09182>.
- [12] M. Hamdaqa, L. Tahvildari, N. LaChapelle, B. Campbell, Cultural scene detection using reverse louvain optimization, *Sci. Comput. Program.* 95 (P1) (2014) 44–72, <http://dx.doi.org/10.1016/j.scico.2014.01.006>.
- [13] G. Witten, G. Poulter, Simulations of infectious diseases on networks, *Comput. Biol. Med.* 37 (2) (2007) 195–205, <http://dx.doi.org/10.1016/j.combiomed.2005.12.002>.
- [14] A. Gautreau, A. Barrat, M. Barthélemy, Global disease spread: Statistics and estimation of arrival times, *J. Theoret. Biol.* 251 (3) (2008) 509–522, <http://dx.doi.org/10.1016/j.jtbi.2007.12.001>.
- [15] K. Gong, M. Tang, P.M. Hui, H.F. Zhang, Y. Do, Y.C. Lai, An efficient immunization strategy for community networks, *PLoS One* 8 (12) (2013) 83489, <http://dx.doi.org/10.1371/journal.pone.0083489>.
- [16] D. Soriano-Pa, G. Ghoshal, A. Arenas, J. Gómez-Gardees, Impact of temporal scales and recurrent mobility patterns on the unfolding of epidemics, *J. Stat. Mech. Theory Exp.* 2020 (2) (2020) <http://dx.doi.org/10.1088/1742-5468/ab6a04>.
- [17] L.D. Valdez, L.A. Braunstein, S. Havlin, Epidemic spreading on modular networks: The fear to declare a pandemic, *Phys. Rev. E* 101 (3) (2020) <http://dx.doi.org/10.1103/PhysRevE.101.032309>.
- [18] W. Huang, C. Li, Epidemic spreading in scale-free networks with community structure, *J. Stat. Mech. Theory Exp.* 2007 (1) (2007) <http://dx.doi.org/10.1088/1742-5468/2007/01/P01014>.
- [19] M. Salathé, J.H. Jones, Dynamics and control of diseases in networks with community structure, in: C. Fraser (Ed.), *PLoS Comput. Biol.* 6 (4) (2010) e1000736, <http://dx.doi.org/10.1371/journal.pcbi.1000736>.
- [20] L.X. Yang, Y. Deng, J.R.C. Piqueira, Epidemic processes on complex networks, *Discrete Dyn. Nat. Soc.* 2017 (2017) 1, <http://dx.doi.org/10.1155/2017/9873678>.
- [21] D. Vaknin, B. Gross, S.V. Buldyrev, S. Havlin, Spreading of localized attacks on spatial multiplex networks with a community structure, *Phys. Rev. Res.* 2 (4) (2020) <http://dx.doi.org/10.1103/PhysRevResearch.2.043005>.
- [22] R. Cohen, S. Havlin, D. Ben-Avraham, Efficient immunization strategies for computer networks and populations, *Phys. Rev. Lett.* 91 (24) (2003) 2–5, <http://dx.doi.org/10.1103/PhysRevLett.91.247901>.
- [23] Z. Liu, H.U. Bambi, Epidemic spreading in community networks, *Europhys. Lett.* 72 (2) (2005) 315–321, <http://dx.doi.org/10.1209/epl/i2004-10550-5>.
- [24] N. Gupta, A. Singh, H. Cherifi, Community-based immunization strategies for epidemic control, 2015 7th International Conference on Communication Systems and Networks, COMSNETS 2015 - Proceedings, IEEE, 2015, <http://dx.doi.org/10.1109/COMSNETS.2015.7098709>.
- [25] B. Gross, S. Havlin, Epidemic spreading and control strategies in spatial modular network, *Appl. Netw. Sci.* 5 (1) (2020) <http://dx.doi.org/10.1007/s41109-020-00337-4>.
- [26] G. Dong, J. Fan, L.M. Shekhtman, S. Shai, R. Du, L. Tian, X. Chen, H. Eugene Stanley, S. Havlin, Resilience of networks with community structure behaves as if under an external field, *Proc. Natl. Acad. Sci. USA* 115 (27) (2018) 6911–6915, <http://dx.doi.org/10.1073/pnas.1801588115>.
- [27] A.J. Tatem, S. Adamo, N. Bharti, C.R. Burgert, M. Castro, A. Dorelien, G. Fink, C. Linard, M. John, L. Montana, M.R. Montgomery, A. Nelson, A.M. Noor, D. Pindolia, G. Yetman, D. Balk, Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation, *Popul. Health Metr.* 10 (1) (2012) 8, <http://dx.doi.org/10.1186/1478-7954-10-8>.
- [28] S.I. Hay, K.E. Battle, D.M. Pigott, D.L. Smith, C.L. Moyes, S. Bhatt, J.S. Brownstein, N. Collier, M.F. Myers, D.B. George, P.W. Gething, Global mapping of infectious disease, *Philos. Trans. R. Soc. B* 368 (1614) (2013) <http://dx.doi.org/10.1098/rstb.2012.0250>.
- [29] C. Caminade, J. Turner, S. Metelmann, J.C. Hesson, M.S.C. Blagrove, T. Solomon, A.P. Morse, M. Baylis, Erratum: Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015 (Proceedings of the National Academy of Sciences of the United States of America (2016) 114:1 (119-12419) DOI: 10.1073/pnas.1614303114), *Proc. Natl. Acad. Sci. USA* 114 (7) (2017) E1301–E1302, <http://dx.doi.org/10.1073/pnas.1700746114>.
- [30] K. Honjo, Social epidemiology: Definition, history, and research examples, *Environ. Health Prev. Med.* 9 (5) (2004) 193–199, <http://dx.doi.org/10.1007/bf02898100>.
- [31] K. Fox, Social epidemiology: How socioeconomic risk factors become health realities, *Virtual Mentor* 8 (11) (2006) 737–743, <http://dx.doi.org/10.1001/virtualmentor.2006.8.11.jdscl-0611>.
- [32] M.E. Halloran, N.M. Ferguson, S. Eubank, I.M. Longini, D.A. Cummings, B. Lewis, S. Xu, C. Fraser, A. Vullikanti, T.C. Germann, D. Wagnier, R. Beckman, K. Kadau, C. Barrett, C.A. Macken, D.S. Burke, P. Cooley, Modeling targeted layered containment of an influenza pandemic in the United States, *Proc. Natl. Acad. Sci. USA* 105 (12) (2008) 4639–4644, <http://dx.doi.org/10.1073/pnas.0706849105>.
- [33] P. Deb, D. Furceri, J. Ostry, N. Tawk, The Effect of Containment Measures on the COVID-19 Pandemic, IMF Working Papers 20 (159), 2020, <http://dx.doi.org/10.5089/9781513550268.001>.
- [34] C.A. Manore, K.S. Hickmann, J.M. Hyman, I.M. Foppa, J.K. Davis, D.M. Wesson, C.N. More, A network-patch methodology for adapting agent-based models for directly transmitted disease to mosquito-borne disease, *J. Biol. Dyn.* 9 (1) (2014) 52–72, <http://dx.doi.org/10.1080/17513758.2015.1005698>.
- [35] S. Chen, Y. Owolabi, A. Li, E. Lo, P. Robinson, D. Janies, C. Lee, M. Dulin, Patch dynamics modeling framework from pathogens' perspective: Unified and standardized approach for complicated epidemic systems, *PLoS One* 15 (10 October) (2020) <http://dx.doi.org/10.1371/journal.pone.0238186>.
- [36] O. Diekmann, J.A. Heesterbeek, M.G. Roberts, The construction of next-generation matrices for compartmental epidemic models, *J. R. Soc. Interface* 7 (47) (2010) 873–885, <http://dx.doi.org/10.1098/rsif.2009.0386>.
- [37] H.L. Mills, S. Riley, The spatial resolution of epidemic peaks, *PLoS Comput. Biol.* 10 (4) (2014) e1003561, <http://dx.doi.org/10.1371/journal.pcbi.1003561>.
- [38] S. Riley, K. Eames, V. Isham, D. Mollison, P. Trapman, Five challenges for spatial epidemic models, *Epidemics* 10 (2015) (2015) 68–71, <http://dx.doi.org/10.1016/j.epidem.2014.07.001>.
- [39] S.E. Fick, R.J. Hijmans, WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, *Int. J. Climatol.* 37 (12) (2017) 4302–4315, <http://dx.doi.org/10.1002/joc.5086>.
- [40] X.Y. Yan, W.X. Wang, Z.Y. Gao, Y.C. Lai, Universal model of individual and population mobility on diverse spatial scales, *Nature Commun.* 8 (1) (2017) 1–9, <http://dx.doi.org/10.1038/s41467-017-01892-8>.
- [41] W.S. Robinson, Ecological correlations and the behavior of individuals, *Am. Sociol. Rev.* 15 (3) (1950) 351, <http://dx.doi.org/10.2307/2087176>.
- [42] S. Openshaw, Ecological fallacies and the analysis of areal census data (UK, Italy), *Environ. Plan. A* 16 (1) (1984) 17–31, <http://dx.doi.org/10.1068/A160017>.
- [43] P.E. Parham, J. Waldo, G.K. Christophides, D. Hemming, F. Agosto, K.J. Evans, N. Fefferman, H. Gaff, A. Gumel, S. Ladeau, S. Lenhart, R.E. Mickens, E.N. Naumova, R.S. Ostfeld, P.D. Ready, M.B. Thomas, J. Velasco-Hernandez, E. Michael, Climate, environmental and socio-economic change: Weighing up the balance in vector-borne disease transmission, *Philos. Trans. R. Soc. B* 370 (1665) (2015) 1–17, <http://dx.doi.org/10.1098/rstb.2013.0551>.
- [44] B. Lieberthal, A.M. Gardner, Connectivity, reproduction number, and mobility interact to determine communities' epidemiological superspreader potential in a metapopulation network, *PLoS Comput. Biol.* 17 (3) (2021) e1008674, <http://dx.doi.org/10.1371/JOURNAL.PCB.1008674>.
- [45] D. Balcan, V. Colizza, B. Gonçalves, H. Hud, J.J. Ramasco, A. Vespignani, Multiscale mobility networks and the spatial spreading of infectious diseases, *Proc. Natl. Acad. Sci. USA* 106 (51) (2009) 21484–21489, <http://dx.doi.org/10.1073/pnas.0906910106>.
- [46] V. Colizza, A. Vespignani, Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations, *J. Theoret. Biol.* 251 (3) (2008) 450–467, <http://dx.doi.org/10.1016/j.jtbi.2007.11.028>.
- [47] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks, *Proc. Natl. Acad. Sci. USA* 101 (11) (2004) 3747–3752, <http://dx.doi.org/10.1073/pnas.0400087101>.
- [48] T. Louail, M. Lenormand, M. Picornell, O.G. Cantú, R. Herranz, E. Frias-Martinez, J.J. Ramasco, M. Barthélemy, Uncovering the spatial structure of mobility networks, *Nature Commun.* 6 (1) (2015) 1–8, <http://dx.doi.org/10.1038/ncomms7007>.
- [49] L. Wang, X. Li, Spatial epidemiology of networked metapopulation: an overview, *Chin. Sci. Bull.* 59 (28) (2014) 3511–3522, <http://dx.doi.org/10.1007/s11434-014-0499-8>.
- [50] M.E. Newman, M. Girvan, Finding and evaluating community structure in networks, *Phys. Rev. E* 69 (2) (2004) 026113, <http://dx.doi.org/10.1103/PhysRevE.69.026113>.
- [51] A. Kehagias, Community detection toolbox, File Exch. - MATLAB Cent. (2020).
- [52] L. Danon, A.P. Ford, T. House, C.P. Jewell, M.J. Keeling, G.O. Roberts, J.V. Ross, M.C. Vernon, Networks and the epidemiology of infectious disease, *Interdiscip. Perspect. Infect. Dis.* 2011 (2011) 28, <http://dx.doi.org/10.1155/2011/284909>.
- [53] C. Stegehuis, R. Van Der Hofstad, J.S. Van Leeuwen, Epidemic spreading on complex networks with community structures, *Sci. Rep.* 6 (1) (2016) 1–7, <http://dx.doi.org/10.1038/srep29748>, 2016 6:1.
- [54] C.L. Staudt, M. Hamann, A. Gutfraind, I. Safto, H. Meyerhenke, Generating realistic scaled complex networks, *Appl. Netw. Sci.* 2 (1) (2017) 36, <http://dx.doi.org/10.1007/s41109-017-0054-z>.
- [55] M. Biggerstaff, S. Cauchemez, C. Reed, M. Gambhir, L. Finelli, Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature, *BMC Infect. Dis.* 14 (1) (2014) <http://dx.doi.org/10.1186/1471-2334-14-480>.
- [56] N. Mori, Despiking, MATLAB Cent. File Exch. (2022).
- [57] L.F. Shampine, M.W. Reichelt, The MATLAB ode suite, *SIAM J. Sci. Comput.* 18 (1) (1997) 1–22, <http://dx.doi.org/10.1137/S1064827594276424>.
- [58] M.K. Andy Bunn, A language and environment for statistical computing, R Found. Statist. Comput. 10 (1) (2017) 11–18.

- [59] P. Van Den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.* 180 (1–2) (2002) 29–48, [http://dx.doi.org/10.1016/S0025-5564\(02\)00108-6](http://dx.doi.org/10.1016/S0025-5564(02)00108-6).
- [60] H.V. Henderson, S.R. Searle, On deriving the inverse of a sum of matrices, *SIAM Rev.* 23 (1) (1981) 53–60, <http://dx.doi.org/10.1137/1023004>.
- [61] F.G. Frobenius, Ueber Matrizen aus nicht negativen Elementen, *Sitzungsberichte Preussische Akademie Der Wissenschaft, Berlin*, 1912, pp. 456–477.
- [62] B. MacWhinney, Functional analysis, *Companion Cogn. Sci.* (2008) 402–412, <http://dx.doi.org/10.1002/9781405164535.ch31>.
- [63] D. Bichara, A. Iggidr, Multi-patch and multi-group epidemic models: A new framework, *J. Math. Biol.* 77 (1) (2018) 107–134, <http://dx.doi.org/10.1007/s00285-017-1191-9>.
- [64] D. Kranda, The square of adjacency matrices, 2012, Available from: <http://arxiv.org/abs/1207.3122>.
- [65] R. Albert, A.L. Barabási, Statistical mechanics of complex networks, *Rev. Modern Phys.* 74 (1) (2002) 47–97, <http://dx.doi.org/10.1103/RevModPhys.74.47>.
- [66] J. Ma, Estimating epidemic exponential growth rate and basic reproduction number, *Infect. Dis. Model.* 5 (2020) 129–141, <http://dx.doi.org/10.1016/j.idm.2019.12.009>.
- [67] D. Knipf, A new approach for designing disease intervention strategies in metapopulation models, *J. Biol. Dyn.* 10 (1) (2016) 71–94, <http://dx.doi.org/10.1080/17513758.2015.1107140>.
- [68] A. Liaw, M. Wiener, Classification and regression by randomforest, *R News* 2 (3) (2002) 18–22.
- [69] J.D. Murray, *Mathematical Biology I an Introduction*, 3rd editio, Springer, 2007, Springer Verlag.Pdf.
- [70] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* 393 (6684) (1998) 440–442, <http://dx.doi.org/10.1038/30918>.
- [71] N. Masuda, Effects of diffusion rates on epidemic spreads in metapopulation networks, *New J. Phys.* 12 (2010) <http://dx.doi.org/10.1088/1367-2630/12/9/093009>.
- [72] L. Chang, M. Duan, G. Sun, Z. Jin, Cross-diffusion-induced patterns in an SIR epidemic model on complex networks, *Chaos* 30 (1) (2020) 13147, <http://dx.doi.org/10.1063/1.5135069>.
- [73] A. Nematzadeh, E. Ferrara, A. Flammini, Y.Y. Ahn, Optimal network modularity for information diffusion, *Phys. Rev. Lett.* 113 (8) (2014) 088701, <http://dx.doi.org/10.1103/PhysRevLett.113.088701>.
- [74] L.D. Valdez, H.H. Araújo Régo, H.E. Stanley, S. Havlin, L.A. Braunstein, The role of bridge nodes between layers on epidemic spreading, *New J. Phys.* 20 (12) (2018) <http://dx.doi.org/10.1088/1367-2630/aaf3ea>.
- [75] L. Xue, C. Scoglio, The network level reproduction number for infectious diseases with both vertical and horizontal transmission, *Math. Biosci.* 243 (1) (2013) 67–80, <http://dx.doi.org/10.1016/j.mbs.2013.02.004>.
- [76] L.G. Alvarez-Zuzek, M.A. Di Muro, S. Havlin, L.A. Braunstein, Dynamic vaccination in partially overlapped multiplex network, *Phys. Rev. E* 99 (1) (2019) <http://dx.doi.org/10.1103/PhysRevE.99.012302>.
- [77] A.H. Dekker, Network centrality and super-spreaders in infectious disease epidemiology, *Proceedings - 20th International Congress on Modelling and Simulation, MODSIM 2013 (December 2013)* (2013) 331–337, <http://dx.doi.org/10.36334/modsim.2013.a5.dekker>.
- [78] K.A. Kabir, K. Kuga, J. Tanimoto, The impact of information spreading on epidemic vaccination game dynamics in a heterogeneous complex network: A theoretical approach, *Chaos Solitons Fractals* 132 (2020) 109548, <http://dx.doi.org/10.1016/j.chaos.2019.109548>.
- [79] E. Vergu, H. Busson, P. Ezanno, Impact of the infection period distribution on the epidemic spread in a metapopulation model, *PLoS One* 5 (2) (2010) 9371, <http://dx.doi.org/10.1371/journal.pone.0009371>.
- [80] J. Badham, R. Stocker, The impact of network clustering and assortativity on epidemic behaviour, *Theor. Popul. Biol.* 77 (1) (2010) 71–75, <http://dx.doi.org/10.1016/j.tpb.2009.11.003>.
- [81] H. Rahmandad, J. Sterman, Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models, *Manage. Sci.* 54 (5) (2008) 998–1014, <http://dx.doi.org/10.1287/mnsc.1070.0787>.
- [82] N.T. Bailey, Macro-modelling and prediction of epidemic spread at community level, *Math. Model.* 7 (5–8) (1986) 689–717, [http://dx.doi.org/10.1016/0270-0255\(86\)90128-4](http://dx.doi.org/10.1016/0270-0255(86)90128-4).
- [83] C. Scoglio, W. Schumm, P. Schumm, T. Easton, S.R. Chowdhury, A. Sydney, M. Youssef, Efficient mitigation strategies for epidemics in rural regions, *PLoS One* 5 (7) (2010) 11569, <http://dx.doi.org/10.1371/journal.pone.0011569>.
- [84] J. Hadidjojo, S.A. Cheong, Equal graph partitioning on estimated infection network as an effective epidemic mitigation measure, *PLoS One* 6 (7) (2011) 22124, <http://dx.doi.org/10.1371/journal.pone.0022124>.
- [85] T.D. Hollingsworth, D. Klinkenberg, H. Heesterbeek, R.M. Anderson, Mitigation strategies for pandemic influenza a: Balancing conflicting policy objectives, *PLoS Comput. Biol.* 7 (2) (2011) <http://dx.doi.org/10.1371/journal.pcbi.1001076>.
- [86] M. Kafsi, E. Kazemi, L. Maystre, L. Yartseva, M. Grossglauer, P. Thiran, Mitigating epidemics through mobile micro-measures, 2013, Available from: <http://arxiv.org/abs/1307.2084>.
- [87] C. Li, G. ping Jiang, Y. Song, L. Xia, Y. Li, B. Song, Modeling and analysis of epidemic spreading on community networks with heterogeneity, *J. Parallel Distrib. Comput.* 119 (2018) 136–145, <http://dx.doi.org/10.1016/j.jpdc.2018.04.009>.
- [88] M. Doostmohammadian, H.R. Rabiee, U.A. Khan, Centrality-based epidemic control in complex social networks, *Soc. Netw. Anal. Min.* 10 (1) (2020) 32, <http://dx.doi.org/10.1007/s13278-020-00638-7>.
- [89] W.C.B. Chin, R. Bouffanais, Spatial super-spreaders and super-susceptibles in human movement networks, *Sci. Rep.* 10 (1) (2020) 1–19, <http://dx.doi.org/10.1038/s41598-020-75697-z>.
- [90] R.S. Ostfeld, G.E. Glass, F. Keesing, Spatial epidemiology: An emerging (or re-emerging) discipline, *Trends Ecol. Evol.* 20 (6 SPEC. ISS.) (2005) 328–336, <http://dx.doi.org/10.1016/j.tree.2005.03.009>.
- [91] M.G. Garner, S.A. Hamilton, Principles of epidemiological modelling, *OIE Rev. Sci. Tech.* 30 (2) (2011) 407–416, <http://dx.doi.org/10.20506/rst.30.2.2045>.
- [92] Y. Cheng, N.B. Tjaden, A. Jaeschke, S.M. Thomas, C. Beierkuhnlein, Deriving risk maps from epidemiological models of vector borne diseases: State-of-the-art and suggestions for best practice, *Epidemics* 33 (2020) 100411, <http://dx.doi.org/10.1016/j.epidem.2020.100411>.
- [93] L. Beale, J.J. Abellan, S. Hodgson, L. Jarup, Methodologic issues and approaches to spatial epidemiology, *Environ. Health Perspect.* 116 (8) (2008) 1105–1110, <http://dx.doi.org/10.1289/ehp.10816>.
- [94] L.O. Prokhorenkova, P. Prałat, A. Raigorodskii, Modularity of complex networks models, in: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, in: LNCS, vol. 10088, Springer Verlag, 2016, pp. 115–126.
- [95] P.L. Delamater, E.J. Street, T.F. Leslie, Y.T. Yang, K.H. Jacobsen, Complexity of the basic reproduction number (R₀), *Emerg. Infect. Diseases* 25 (1) (2019) 1–4, <http://dx.doi.org/10.3201/eid2501.171901>.
- [96] X. Boulet, M. Zargayouna, G. Scemama, F. Leurent, A middleware-based approach for multi-scale mobility simulation, *Future Internet* 13 (2) (2021) 1–21, <http://dx.doi.org/10.3390/fi13020022>.
- [97] J.C. Miller, J.M. Hyman, Effective vaccination strategies for realistic social networks, *Physica A* 386 (2) (2007) 780–785, <http://dx.doi.org/10.1016/j.physa.2007.08.054>.
- [98] V.J. Davey, R.J. Glass, Rescinding community mitigation strategies in an influenza pandemic, *Emerg. Infect. Diseases* 14 (3) (2008) 365–372, <http://dx.doi.org/10.3201/eid1403.070673>.
- [99] Y. Min, X. Jin, Y. Ge, J. Chang, The role of community mixing styles in shaping epidemic behaviors in weighted networks, *PLoS One* 8 (2) (2013) 57100, <http://dx.doi.org/10.1371/journal.pone.0057100>.
- [100] M. Nadini, K. Sun, E. Ubaldi, M. Starnini, A. Rizzo, N. Perra, Epidemic spreading in modular time-varying networks, *Sci. Rep.* 8 (1) (2018) 1–11, <http://dx.doi.org/10.1038/s41598-018-20908-x>.
- [101] P.D. Fajgelbaum, A. Khandelwal, W. Kim, C. Mantovani, E. Schaal, Optimal lockdown in a commuting network, *Am. Econ. Rev.: Insights* 3 (4) (2021) 503–522, <http://dx.doi.org/10.1257/aeri.20200401>.
- [102] J.B. Wang, C. Li, X. Li, Predicting spatial transmission at the early stage of epidemics on a networked metapopulation, in: *IEEE International Conference on Control and Automation, ICCA, Vol. 2016-July*, IEEE Computer Society, 2016, pp. 116–121, <http://dx.doi.org/10.1109/ICCA.2016.7505262>.
- [103] Y. Xia, O.N. Bjørnstad, B.T. Grenfell, Measles metapopulation dynamics: A gravity model for epidemiological coupling and dynamics, *Am. Nat.* 164 (2) (2004) 267–281, <http://dx.doi.org/10.1086/422341>.