MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **6**, Number **2**, April **2009**

doi:10.3934/mbe.2009.6.239

pp. **239–259**

THE REPRODUCTION NUMBER R_t IN STRUCTURED AND NONSTRUCTURED POPULATIONS

Tom Burr

Statistical Sciences Group
Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Gerardo Chowell

School of Human Evolution and Social Change
Arizona State University, Box 872402, Tempe, AZ 85287, USA
and
Division of Epidemiology and Population Studies, Fogarty International Center
National Institutes of Health, Bethesda, MD, USA

Abstract. Using daily counts of newly infected individuals, Wallinga and Teunis (WT) introduced a conceptually simple method to estimate the number of secondary cases per primary case (R_t) for a given day. The method requires an estimate of the generation interval probability density function (pdf), which specifies the probabilities for the times between symptom onset in a primary case and symptom onset in a corresponding secondary case. Other methods to estimate R_t are based on explicit models such as the SIR model; therefore, one might expect the WT method to be more robust to departures from SIR-type behavior. This paper uses simulated data to compare the quality of daily R_t estimates based on a SIR model to those using the WT method for both structured (classical SIR assumptions are violated) and nonstructured (classical SIR assumptions hold) populations. By using detailed simulations that record the infection day of each new infection and the donor-recipient identities, the true R_t and the generation interval pdf is known with negligible error. We find that the generation interval pdf is time dependent in all cases, which agrees with recent results reported elsewhere. We also find that the WT method performs essentially the same in the structured populations (except for a spatial network) as it does in the nonstructured population. And, the WT method does as well or better than a SIR-model based method in three of the four structured populations. Therefore, even if the contact patterns are heterogeneous as in the structured populations evaluated here, the WT method provides reasonable estimates of R_t, as does the SIR method.

1. **Introduction.** A common assumption when modeling the transmission dynamics of infectious diseases is that the population is composed of a homogeneous group that mixes uniformly at random (*homogeneous mixing* in an unstructured population). That is, each infectious individual (regardless of age, geographic location, etc.) has the same probability of coming in contact with any susceptible individual in the population. A model that is closer to reality incorporates the fact that the population is probably structured, having many types of heterogeneities at the host

2000 *Mathematics Subject Classification.* Primary: 58F15, 58F17; Secondary: 53C35.
Key words and phrases. reproduction number; generation interval; structured population.

level such as differences in contact rates, susceptibility, and transmissibility (see e.g., [31, 1, 4, 38, 16, 5, 32, 21, 24]).

A key quantity in epidemiology is the basic reproduction number (R_0) defined as the number of secondary cases generated by an infectious case in an entirely susceptible population [1, 14, 4]. In traditional large-population approximations, if $R_0 > 1$, an epidemic can occur whereas if $R_0 < 1$ an epidemic cannot sustain itself. Later in the epidemic, as the number of susceptibles decreases, the number of secondary cases per primary case (R_t) is a useful measure of epidemic strength and will be smaller than R_0 for a given day t.

An estimate of R_0 is useful because it can predict the final epidemic size in the absence of interventions and can therefore be used to assess the impact of interventions. In an unstructured, homogeneous population, an expression for the expected final epidemic size of the simple susceptible-infected-recovered (SIR) epidemic model, the non-linear dynamics of which are described by depletion of susceptible individuals, was first derived by Kermack and Mckendrick (1927) [22], expressed as $Z = 1 - \exp^{-R_0 Z}$, where Z is the cumulative fraction of infecteds ("final size"). This final size formula for Z is robust with respect to a number of changes in the epidemic process including having arbitrarily distributed latency periods, having a distribution of different infectious stages, and heterogeneity in the transmission rate (including superspreading events). In situations having adequate robustness of the final size formula, one approach to estimate the effectiveness of interventions is to compare the final size estimate $\hat{Z}$ obtained from the final size formula using the estimate $\hat{R}_0$ to the observed final size of the epidemic. However, the final size is not robust to most contact heterogeneity, except for a special class of spatial contact structures that assumes the transmission rate is the same within all groups [28, 29, 30]. In situations where the final size formula is not a good approximation, other methods are required to assess the impact of interventions.

A real-time estimate of R_t is also useful in order to assess the impact of interventions. This paper focuses on two estimation schemes to estimate the effective reproduction number R_t at each time step t, and assumes that the number of new infections is recorded each day. Both estimation schemes are applied to four population structures (one unstructured and three structured).

Significant reductions in R_t can result from the implementation of public health interventions that can reduce the transmission rate over time. Public health measures include quarantine of suspected cases, isolation of infectious individuals, pre- and post-outbreak vaccination, prophylaxis with antivirals, and for animal infectious diseases such as foot-and-mouth disease, culling of premises is the control policy of many countries. The goal of public health interventions is to promptly reduce R_t to a value less than one. Estimation of R_t at the beginning of an epidemic is essential to implement interventions with the required intensity for epidemic control. Either too strong or too mild interventions can have negative consequences. Obtaining reliable estimates of R_t as soon as possible is of high priority for public health management, especially when dealing with emerging infectious pathogens.

Following sections include additional background, a description of the two schemes to estimate R_t, a description of the structured and unstructured populations that were simulated, estimation results, and conclusions. Qualitatively the two main conclusions are: (1) the estimation scheme from [41] (Wallinga and Teunis, WT) is surprisingly effective (robust) in most of the structured populations, and (2) the generation interval (GI) required in the WT scheme is not constant over time, but

the impact on the WT scheme appears minimal, except for one of the structured populations evaluated.

2. **Background.** Despite its unrealistic assumptions, the classical SIR (susceptible-infectious-removed) type models have a long history of providing useful insights into the dynamics of infectious diseases [22, 1, 4]. Simple models are usually tractable for parameterization and estimation from epidemic data that usually includes time series of cases or deaths. On the other hand, complex models that account for host heterogeneities are typically difficult to calibrate due to lack of appropriate data or become technically intractable for parameter estimation from epidemic data.

The classical SIR model was introduced by Kermack and McKendrick [22] in 1927. This simple epidemic model has provided an established basis to model the transmission dynamics of infectious diseases. The SIR model classifies individuals as susceptible (S), infectious (I), and recovered (R). From the homogeneous mixing assumption, it follows that the force of infection is given by $\beta I(t)/N$ where the transmission rate β is given by the product of the contact rate (c) and the probability of transmission per contact (p). The fraction $I(t)/N$ is the probability that a random contact would be with an infectious individual. This model also assumes no deaths or migration and hence the population size N is constant throughout the epidemic. This model also assumes that the disease latency period is short (and hence negligible), and that the time-scale of the epidemic is much faster than those of demographic processes (natural birth and death). Finally, recovered individuals from infection $R(t)$ are assumed to acquire immunity to the disease for at least the duration of the outbreak. This SIR transmission process (single outbreak) can then be modeled using the following system of nonlinear differential equations:

$$\begin{cases} \dot{S}(t) &= -\beta S(t)I(t)/N \\ \dot{I}(t) &= \beta S(t)I(t)/N - \gamma I(t) \\ \dot{R}(t) &= \gamma I(t) \end{cases} \tag{1}$$

where the dot denotes the time derivatives.

Assuming daily counts of newly infected individuals, Wallinga and Teunis (2004) [41] introduced the conceptually simple WT method to estimate the number of secondary cases per primary case for a given day. The method requires an estimate of the GI probability density function (pdf).

Other methods to estimate R_t are based on explicit models such as the susceptible-infected-recovered (SIR) model; therefore, one might expect the WT method to be more robust to departures from SIR-type behavior. Although some tests of the quality of the WT method were performed using simulated data, the simulations were limited to simple cases in which implicit modeling assumptions were nearly met or were met exactly. For example, Wallinga and Teunis (2004) [41] perform limited tests of the estimation procedure in simulated data using $R_t = 3$ for t prior to a specified date and $R_t = 0.7$ after that date. Effects of nonreporting and of a one-time change in the pdf of the GI were evaluated. However, the number of secondary infections was simply drawn from a negative binomial distribution with mean R_t and a variance determined by evaluation of data from a real SARS outbreak in Singapore. Therefore, model departures such as nonhomogeneous mixing of the population were not considered.

It is important to note that there is some inconsistency in the literature in the definition of GI and related concepts such as the serial interval [15, 40]. The GI

is defined as the time between the infection of a primary case i and one of its secondary cases j while the serial interval is defined as the time between occurrence of observable events in the progress of an infectious disease (e.g., the onset of clinical symptoms) [40]. We use the term "primary" case here to mean any donor case during the epidemic. Therefore, a primary case could also be among the "index" cases, which are the initially infected individuals.

Generally, it is important to distinguish among time of infection, time of symptom onset, and time of infectiousness. However, infection and infectiousness are assumed to occur at the same time in the simple SIR model used here. And, the time of symptom onset is not a factor because our simulations do not consider time of symptom onset. This means that the GI as defined here can also be regarded as the time between primary and secondary infection. Although it is of interest to extend this study to include modified SIR-type models such as the SEIR (susceptible-exposed-infectious-recovered) model, we do not believe it will alter our basic conclusion that the WT method is fairly robust to structure-type violations of SIR assumptions.

3. **Methods.**

3.1. **WT method.**

Wallinga and Teunis (2004) [41] introduced a simple way to estimate the effective reproduction number, R_t. Let $f(i-j)$ denote the probability density function (pdf) of the GI (e.g., 1 day, 2 days, ...) between the infection of a primary case i_1 infected on day i and one of its secondary cases i_2 infected on day j. The time of infection, onset of symptoms, and infectiousness are all assumed to occur at the same time in our SIR model. Then the relative probability that case i_2 initially infected on day j is infected by case i_1 is calculated as $p_{ij} = f(j-i)/\sum_{k\leq j} f(j-k)$. The effective reproduction number R_t on day i is then the average over all individuals who are infected on day i of $\sum_j p_{ij}$. Because all individuals are probabilistically equal, this average over all individuals infected on day i is the same as $\sum_j p_{ij}$.

Cauchemez et al. (2006) [8] extended Wallinga and Teunis (2004) [41] to provide a method to estimate confidence limits when the GI pdf is constant over time and assumed known. Cauchemez et al. (2006) [9] assumed the GI pdf was a constant-in-time Weibull density with an unknown shape and scale parameter to be estimated from the outbreak data. There could be many explanations for why the average GI could change with time in real populations. Because it is unrealistic to expect a highly accurate estimate of the GI pdf for each outbreak for various times over the course of a given outbreak, it is important to study the effects of the time dependence of the GI pdf on R_t.

3.2. **SIR method.**

Denote the per-day recovery probability p, where $p = \frac{1}{\gamma}$. It has been shown that on day i, $R_i = \beta \sum_{j=0}^{\infty} \frac{S_{i+j}}{N}(1-p)^j$ [3, 4], and this result is the basis for our second method to estimate R_t on day i. The observed S, I, and R time series can be used as described in [6] to estimate β and γ via least squares (LS) fitting, and then using $\hat{R}_t = \hat{\beta} \sum_{j=0}^{\infty} \frac{S_{i+j}}{N}(1-\hat{p})^j$, where $\hat{p} = \frac{1}{\hat{\gamma}}$.

Note that in practice, $\frac{S_t}{N}$ must be estimated, so performance of this method will be somewhat worse than illustrated in the Results section, where it is assumed that $\frac{S_t}{N}$ is known exactly.

The LS method applied to the SIR equations obviously assumes a homogeneous contact structure. The classic SIR model in which depletion of susceptibles fully

captures the observed dynamics is very unlikely to accurately describe any real population, or even the idealized structured populations described below. This makes the SIR method effective as a basis for comparison to the WT method, which makes weaker assumptions about the contact structure.

4. **Epidemics on more realistic population structures.** When the population contact structure is expected to depart from the homogeneous mixing assumption, networks provide a natural alternative framework, in terms of nodes and relationships between them. In the last few years there has been an increasing amount of work in the context of epidemic spread in networks (see for example [32, 12]). The substrate network on which the disease spreads can change depending on the disease in question.

For the case of influenza and SARS, social networks (e.g., [10, 35]) provide an appropriate description of the contact network whereas for the case of sexually transmitted disease, the network where the disease disseminates is quite different with a few individuals having a remarkably high number of contacts and most individuals having just a few contacts (the so-called core group [17] or scale-free structure [25]). Network models can help characterize and understand the underlying complex system of disease transmission. In particular, network properties can help to distinguish and compare systems. Common network characteristics include the degree distribution [2], the average distance between nodes (or characteristic path length [43]), and the clustering coefficient [43].

We simulated several types of population structures using R [37], as described next. For each structure, transmission rates were either homogeneous (constant for all individuals) or heterogeneous (vary with individual). Results are reported for the case of heterogeneous transmission rates for all the non-SIR models, but results for the homogeneous rates are similar. We chose homogeneous rates for the SIR model because we wanted perfect agreement with the SIR assumptions in the base case. The population structures were static (do not change over the course of the epidemic) and the population size was $N = 10,000$.

For this study, we chose parameters that are believed to correspond to various influenza outbreaks. For influenza, the mean infectious period is 4.1 days [11], and hence the reproduction number for the SIR model is $R_0 = \frac{\beta}{\gamma} = .37/(1/4.1)) = 1.5$, which corresponds via the "final size" formula to $Z = 0.58$ [22].

All simulations are discrete-event simulations, recording the status (S, I, or R) for each individual at each time step (day). All simulations were carried out in the R statistical programming language [37] which is freely available, and our custom R functions are available upon request. The initial number of infecteds (randomly selected) was chosen to be 15, which was empirically determined to be large enough that each realization would result in an outbreak (except for the spatial network, where we found that 25 was large enough for each realization to result in an outbreak). The index case was set to "0" for the initial group of infected individuals; the index case for all other individuals who experience infection is some individual described by an integer in $1, 2, \ldots, 10,000$.

A susceptible individual i_1 transitioned to infected with probability $\beta \times I_{i_1,t}/N$, at each time step (day), where $I_{i_1,t}$ is the number of infected contacts of individual i_1 on day t. In the network structures described below, each individual has a unique set of contacts. In the SIR model, $I_{i_k,t} = I_t$ for all individuals i_k. In general, $I_{i_k,t}$ characterizes the network structure. Similarly, an infected individual

transitioned to recovered with probability equal to $p = 1/(\text{mean infectious period}) = 1/\gamma = 1/4.1$. Note that this implies that the infected period was sampled from an geometric distribution (approximately an exponential distribution), and sampling directly from a geometric distribution is sometimes used to simulate outbreaks. Also note that the expression for the per-day probability of transitioning from infected to recovered would need to be modified if $\gamma \leq 1$.

The index case was set to "0" for the initial group of infected individuals; the index case for all other individuals who experience infection is some individual described by an integer in $1, 2, \ldots, 10,000$.

4.1. **Random network.** There are many ways to generate a random network. We generated random numbers of neighbors for each individual, sampling the number of neighbors of individual i_1 from a uniform distribution from 1 to 20 neighbors. This implies that the average number of neighbors is 10.5. In this context a "neighbor" of individual i_1 can infect individual i_1 if that neighbor is itself infected. However, individual i_1 is not necessarily eligible to infect a "neighbor" unless symmetry is forced. Results reported here are for the asymmetric case which means that if individual i_1 can infect i_2 (so i_2 is in the neighborhood of i_1), it does not follow that i_2 can infect i_1, except by the chance assignment of i_1 to the neighborhood of i_2. That is, we simply assign a random set of neighbors (potential infectors) to each individual and do not enforce symmetry in the results reported below.

4.2. **Lattice network.** A lattice network has been described in, for example [21]. We used a 100-by-100 grid with the individuals at edge nodes having fewer neighbors. In this network, infected individuals can only infect their neighbors. Each node has a north, south, east, and west neighbor (4 neighbors) except for nodes along edges, which have 2 or 3 neighbors.

For the lattice network, we also simulated a subcase in which a few individuals (10 to 20 of 10,000) were long-ranging, as a way to allow a small amount of random mixing. Burr and Chowell [7] report that small amounts of random mixing makes some otherwise structured populations behave in some aspects like unstructured (SIR-type) populations. Although beyond our scope, this is a topic that deserves further study, perhaps as a special case of current "small world" models.

4.3. **Spatial network.** To construct this network, N random 2-dimensional locations are distributed uniformly in a 1-by-1 unit square and then rescaled to mean 0 and variance 1, each representing an individual. Euclidean distance between each pair of points determined the neighborhood structure (e.g, also known as random geometric graphs [13]). The neighbors (potential infectors) of individual i_1 are any individual within a specified distance (we used 0.7 for points in the unit square transformed to mean 0 and variance 1). To limit the neighborhood sizes, we allowed no more than 20 neighbors in any individual's neighborhood. If more than 20 points fell within 0.7 in distance of an individual,then 20 neighbors were randomly selected from these points. In cases where random selection of 20 points was done, all points within 0.7 in distance were equally likely to be selected as a neighbor, without regard to distance rankings. As with the Random network, symmetry was not enforced, so in many cases, individual i_2 was in the neighborhood of individual i_1 and so could infect i_1, but not vice versa. However, there is a tendency toward symmetry simply because if i_2 is among the 20 nearest transformed points to i_1 that were less than 0.7 in distance, then i_1 tends to also be among the 20 nearest points to i_2.

For the spatial network, as with the lattice network, we also simulated subcases in which a few or many individuals mixed randomly with the entire population. Surprisingly, these "spatial plus some random" cases behaved much more like "spatial only" than like random (SIR) populations.

4.4. **Results.** All results reported here use a population size of $N = 10,000$. To quantify estimation performance, a simple but effective performance summary is the mean absolute error (MAE) in the estimation errors $e_t = \hat{R}_t - R_t$ on day t. Although the root mean squared error is often used in analytical comparisons, the MAE is less sensitive to outliers so in simulation studies is often preferred. Table 1 lists the MAE for the four population structures. Results for the lattice population structure are given both with and without long-ranging individuals. Exploratory data analysis suggest that the estimation errors are symmetric, and approximately normal in distribution. Therefore, approximate 95% confidence intervals for the true R_t can be based on using $\hat{R}_t \pm 2 \times MAE$ for each case.

Recall that R_t is defined here as the expected number of secondary cases per primary, where the expectation is taken over the population of possible outbreak realizations. We found empirically that using $N = 10,000$ gives somewhat erratic daily estimates of R_t in a given realization. We also found that using $N = 10,000$ gives larger estimation errors on average than using larger N, such as $N = 50,000$ or $N = 100,000$. This is partly because when more individuals are involved, more quantities are averaged in the estimation procedures at each time step and partly because the true R_t is relatively poorly estimated by a single realization in a population of size $N = 10,000$. However, the average MAE results over multiple outbreak realizations given in Table 1 are reproducible to small relative uncertainties, and the conclusions are qualitatively the same for $N = 10,000$ as for any larger value of N. Therefore, for practical purposes here, in calculating $e_t = \hat{R}_t - R_t$ and the associated MAEs, we defined R_t as the average number of cases infected by a case with symptom onset occurring on day t in a given outbreak realization [9] without averaging over the population of possible outbreak realizations.

As an aside, we also evaluated MAE results that used a smooth fit to the R_t values over time to define the "true" R_t. Not surprisingly, these MAEs are smaller than those reported in Table 1, because errors in the calculated "true" R_t values are reduced by curve smoothing. However, our main interest is in comparing the WT method to the SIR method so the WT MAEs relative to the SIR MAEs are most important here. Therefore, we report in Table 1 the MAE results for only the unsmoothed R_t values, where R_t is defined as the average number of newly infected cases arising over days $i, i+1, \ldots$, per newly infected case on day i in a given outbreak realization.

We used 100 realizations from a homogeneous (nonstructured) SIR model with transmission rate homogeneity as a basis for comparison, and 100 realizations from each of the structured populations described above. In the structured populations, to model heterogeneous transmission rates, we sampled the transmission rates from a beta distribution having a specified mean and variance. The variance was set to 0.01 (so the standard deviation was 0.1); the mean transmission rate was 0.37 for the homogeneous SIR model, and was selected in the other population models in order to have approximately the same final outbreak size of 60% of the total population. Again, results are reported for the heterogeneous transmission rates

having standard deviation of 0.1 for all the non-SIR cases (networks) and of 0 (homogeneous transmission rates) for the SIR case; however, results for homogeneous transmission rates in the non-SIR cases are similar.

Figure 1 (top) is an example realization of the number of new infected individuals in a homogeneous unstructured SIR case simulation at each time step (day). We will denote the daily number of newly infected individuals as $\dot{C}_t$, where C_t is the cumulative number of individuals who are infected by day i. Because we use discrete-event simulation with a 1-day time step, $\dot{C}$ is the first difference in the cumulative number individuals who are ever infected. The top right, bottom left, and bottom right plots in Figure 1 are the same as the top left, but for the random, lattice with long ranging individuals, and spatial networks, respectively.

The plots in Figure 2 are the observed average GI at each day, averaged over 100 realizations, and the theoretical average GI based on the SIR model assumptions (see below). The top left plot in Figure 2 is for the SIR case and the theoretical average GI is within the 95% confidence limit of the observed average GI. The top right, bottom left, and bottom right plots in Figure 2 are the same as the top left, but for the random, lattice with long ranging individuals, and spatial networks, respectively. For these three networks, the theoretical average GI is not within the 95% confidence limit of the average observed GI.

As the number of new infecteds increases (up to approximately day 40 in the top plot), there is a relatively short average GI. Then as the number of new infecteds decreases (after approximately day 40 in the top left plot), there is a shift toward a longer GI. This trend is consistent with results in [23], although [23] mostly evaluated cases having considerably larger values of R_0 than the $R_0 = 1.57$ case considered here. Larger values of R_0 will have a larger time-dependence in the GI pdf.

More quantitatively, the theoretical average GI μ_i at day i can be computed as follows. Recall that the discrete-time (one-day time steps) GI is the number of days elapsed from the infection day of the donor individual to the infection day of the corresponding infected individual. To compute the GI at day i, we identify all individuals who are first infected on day i, and all individuals who are infected by these. This gives one GI per donor-infectee pair that is used to estimate the GI pdf via repeated simulation.

At each simulated time step (day) new infections (if any) occur first, followed by recoveries (if any), and new infections on day $i + j$ are not yet eligible to recover until the next simulated day. Therefore, results in [14] using a renewal equaion with integral kernel can be modified in the discrete time step case to obtain $\mu_i = \frac{\sum_{j=0}^{\infty} j S_{i+j}(1-p)^j}{\sum_{j=0}^{\infty} S_{i+j}(1-p)^j}$, where $p = 1/(\text{mean infectious period}) = 1/\gamma = 1/4.1$ in our case.

This theoretical average GI μ_i applies for the "forward" definition of the generation interval [40], which means that an infected individual is identified, and we trace forward in time to find the day that any new case was infected by this individual. Alternatively, we could trace backward in time to identify the day on which the corresponding primary case of this infected individual became infected [40, 23]. Although the WT method is expressed using the "forward" definition, generally we found very little difference in MAE results using the forward calculation compared to results using the backward calculation. All results shown in Table 1 are for the forward GI definition.

The change in the average GI over time in our simulated examples occurs because of the time dependent nature of $\dot{C}_j$, as described above. A simple explanation of the relation between the average GI and $\dot{C}_j$ is that as $\dot{C}_j$ increases, there is a higher probability of a one-day GI, which is the shortest possible GI in our discrete-time simulations. Therefore, the average GI is a minimum when $\dot{C}_j$ is a maximum, as noted in comparing the top left plots in Figures 1 and 2 for the SIR (unstructured) population. Some references either explicitly state the assumption that the GI pdf is constant over time (e.g., [8, 44, 9, 42]), or, for example, Hugh-Jones and Tinline (1976) [19] empirically observed time dependence of the average serial interval (which is the same as the GI in our SIR model) in real cattle populations infected with foot and mouth disease, and Lipsitch et al. (2003) [26] reports a decrease in the average GI after controls were implemented in the Singapore SARS outbreak in humans. We also note that [9] extended WT by developing a method to estimate unseen new infecteds during the last stages of an outbreak, thus allowing estimation of R_t in real time, each day in an epidemic.

Kenah et al. (2007) [23] provide recent discussion on the implications of the time-dependent aspect of the GI pdf (and hence also of the average GI). A method involving hazard rates rather than pdfs is suggested as a way to improve the WT method, which is described as "ingenious, but only approximately correct because the distribution of serial intervals varies systematically over the course of an epidemic." Although numerical results comparing the use of pdfs to hazard rates in the WT method are not provided, presumably this is future work. Because our main focus here is in evaluating the impact of ignored population structure, our results here use pdfs as in the original WT method. We believe that the impact of ignored population structure will be similar in the modified WT method that uses hazard rates. This is because the hazard rate approach also ignores population structure, yet Figure 2 suggests that the time behavior of the average GI and therefore also the GI pdf is very different for for the nonstructured (SIR) case than for any of the random, lattice, or spatial networks.

The top plot in Figures 3-6 shows a smooth fit to R_t, and to the WT-based and SIR-based estimates of R_t. Because the MAE values are erratic and large at the very beginning and ending of most outbreaks, we averaged the MAEs from day 10 to day 70 for all outbreaks except the spatial outbreak. For the spatial outbreak, days 10 to days 25 were used to compute the average. All 100 outbreak realizations for each case resulted in a meaningful $\hat{R}_t$ estimate for each day within these restricted day ranges.

Our main quantitative interests are to: (1) compare the estimates for the SIR method to the estimates for the WT method; (2) evaluate to what extent any of the three structured networks have different MAE values than the SIR case, and (3) check for trend in the estimation errors for the SIR and WT methods. All MAE values are repeatable to within ± 0.01 or less. Therefore, the MAE results in Table 1 suggest in terms of the MAE, the WT-based method is better than the SIR-based method for the lattice network, worse than the SIR-based method for the spatial network, and approximately the same as the SIR-based method for the unstructured population and for the random network. Using the WT method, the MAE is substantially higher for the spatial network than for other networks and the MAE is slightly lower for the lattice network than for the SIR case or the random network.

Figure 3 (middle plot) is the average estimation error over all 100 outbreak realizations from time step 10 to 70 for the WT method; the bottom plot is for the SIR method. Approximate 95% confidence intervals (CI) for the average estimation error are shown. These CIs were obtained by bootstrap sampling with replacement from the 100 outbreak realizations, computing the average estimation error for each bootstrap sample, and repeating for 1000 bootstrap samples.

Figures 4-6 are the same as Figure 3, but for the random, lattice, and spatial networks, respectively. Recall from Figure 2 that the average GI is poorly estimated for the structured networks using the result for μ_i given above, although each with a fairly consistent offset from the true average GI. Because the structured networks have non-SIR contact structure, the SIR-based calculation of μ_i is evidently flawed in important ways.

Regarding the spatial network, notice in Figure 6 that R_t is less than one. However, as we discussed the final outbreak size (the cumulative number of infecteds) is approximately 60% in all cases considered here. Therefore, the outbreak "took off" in this spatial network despite having $R_t < 1$ for each day during the early stages. This is a surprising finding that we verified in several re-initializations of the spatial population structure. Conventional wisdom regarding interpretation of R_t suggests that if $R_t < 1$ the outbreak will quickly die out. However, provided we used 25 as the initial number of infecteds, this spatial network with the chosen parameters did have a final size of at least 50% cumulative number of infected in each of 100 realizations. This finding must involve the fact that the network involves unusual overlapping of partially overlapping cliques (groups) of individuals [18]. It also raises a question regarding what are the appropriate measures of epidemic strength in structured networks.

A manuscript under review [34] extends results from network theory to show how networks having $R_0 < 1$ can "take off" (certainly a final epidemic size of 50% or more is an example of "taking off") if one or more initial cases are in highly connected nodes. This deserves additional study that is beyond our scope here. As a preliminary test of whether adding various amounts of random mixing to the spatial network would make the WT method performance closer to its performance in the SIR model, we found that even with a very large portion of random mixing (such as 50%) included with the spatial network, these "spatial plus some random" cases behaved much more like "spatial only" than like random (SIR) populations. This finding also deserves additional study and appears to provide a counter example to a finding presented in [7] in which small amounts of random mixing did in some aspects make a spatial network behave like an unstructured population.

We are focusing on simulated outbreaks in structured or unstructured populations, and the time dependence of the GI pdf implies that the estimation errors $e_t = \hat{R}_t - R_t$ could have a trend, exhibiting bias in opposite direction in the early epidemic stage compared to the late epidemic stage. This is because the average GI is overestimated during the early stages and underestimated during the late stages if the average GI over the entire outbreak is used throughout the outbreak.

Visually, we found evidence of serial structure in the e_t values using simple smooth (using smooth.spline in R) spline fits to the estimation errors, as illustrated in Figures 3-6. The trend was modest or small in all cases except for the spatial network where the trend was strong.

As a simple test for trend in the estimation errors, we applied a t-test to the 100 average differences (over all 100 independent realizations) in the errors between

the first half and second half. The anticipated tendency for a trend in e_t values is apparent in nearly all cases, except the SIR case, with p-values from the t-test all near 0 in the structured network cases. This strong evidence for trend in e_t values is present regardless of whether R_t is smoothed or unsmoothed. All Table 1 results are for the unsmoothed case.

Generally, except for the spatial network, although the trend in estimation errors is statistically significant, the magnitude of the trend is not large (column 2 of table 1) and this is modest good news for the WT method in that the change over time of the generation interval pdf appears mild enough that the procedure still performs adequately. Of course it would be possible to estimate the GI pdf during each of several stages (early, middle, late) of a real outbreak as a way to mitigate the impact on the WT method. The strong trend is obvious in the spatial network in Figures 2 and 6.

For the MAE results in Table 1, the GI pdf is well estimated, because all individuals (among the $N = 10,000$ individuals) who infect one or more others contribute to the estimate. More realistically, far fewer individuals would typically be used to estimate the GI pdf. To study the effect of, in practice, using only a few individuals to estimate the GI pdf, we subsampled far fewer individuals. Somewhat surprisingly, using a small sample of either 100 or 10 GIs to estimate the GI pdf, there was almost no impact on the MAE. For example, in the SIR case, the MAE for the WT method increased from 0.13 (table 1 entry, using all available observed GIs) to 0.16 (using 10 observed GIs). This suggests some robustness with respect to the estimation of the GI's pdf. We believe this robustness arises primarily because the WT method involves a ratio of relative likelihoods, and so can tolerate some degree of misspecification in these likelihoods. In addition, other auxiliary simulations deliberately used the "wrong" GI pdf. For example, we used the generation interval pdf estimate based on the spatial network for a second set of 100 simulations of the SIR (unstructured) population. The resulting MAE increased from 0.13 to 0.17 for the WT method, and there was of course no impact on the SIR method because it does not use the estimated GI pdf. Again, this small to modest increase in the MAE suggests some robustness to using the "wrong" GI pdf. In practice, methods described in [15] including contact tracing, particularly for family members, is a common strategy for estimating the GI pdf. However, we have illustrated a time trend in the GI pdf (Fig. 2) and recall that the WT method as implemented here ignores that fact.

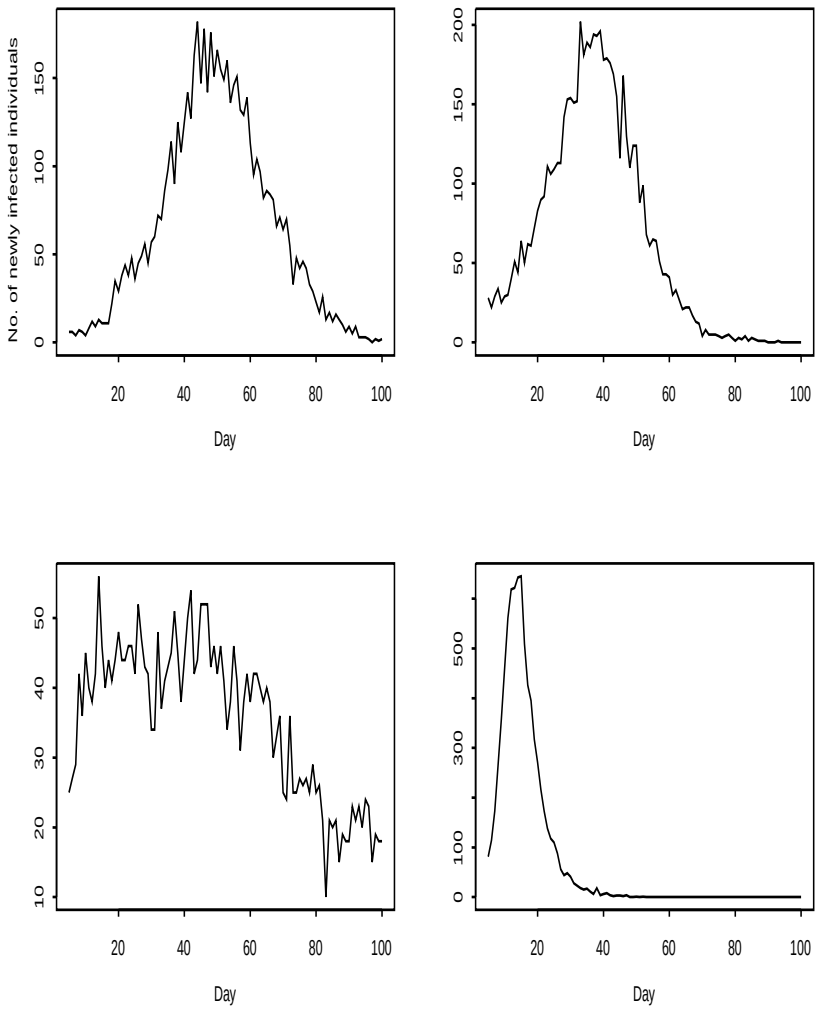

FIGURE 1. A stochastic realization of the daily fraction of newly infected individuals for (top left) unstructured population case generated using the SIR model; (top right) random network; (bottom left) lattice network with long ranging individuals; (bottom right) spatial network.

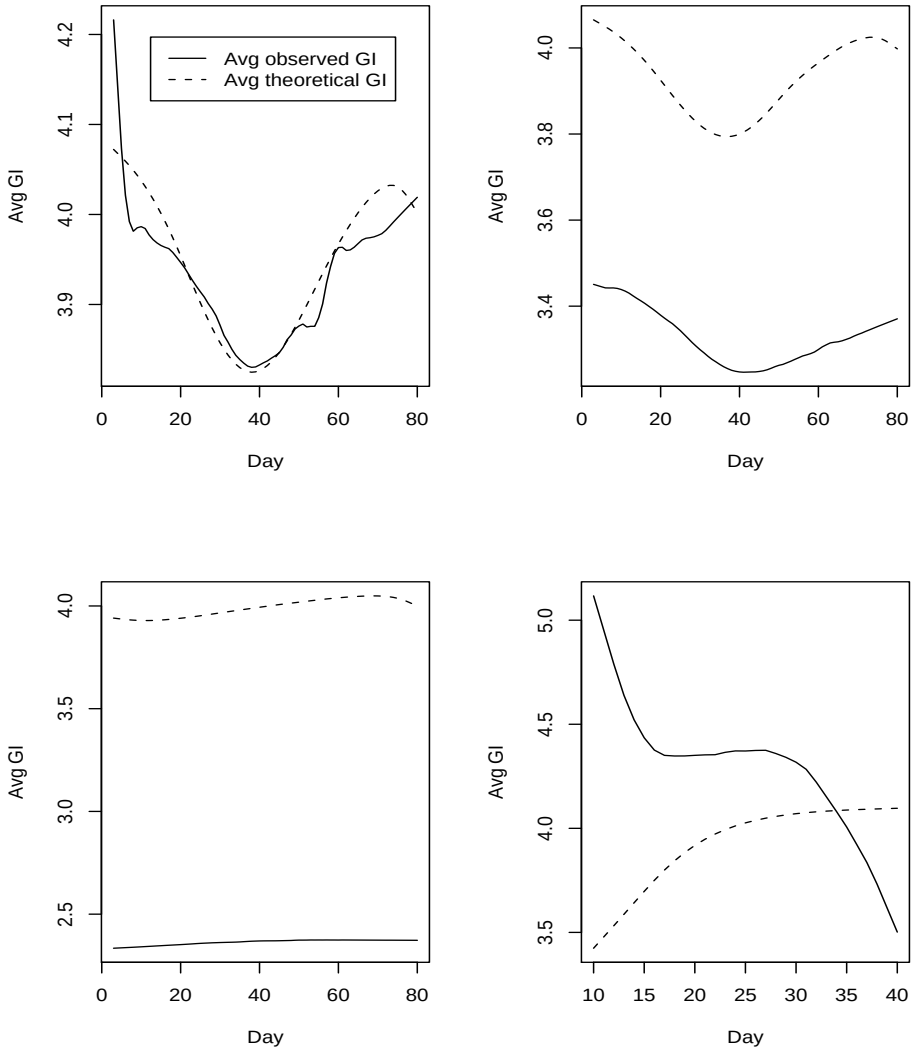

FIGURE 2. Observed and estimated generation intervals for (top left): SIR; (top right): random network; (bottom left): lattice network with long rangers; (bottom right): spatial network.

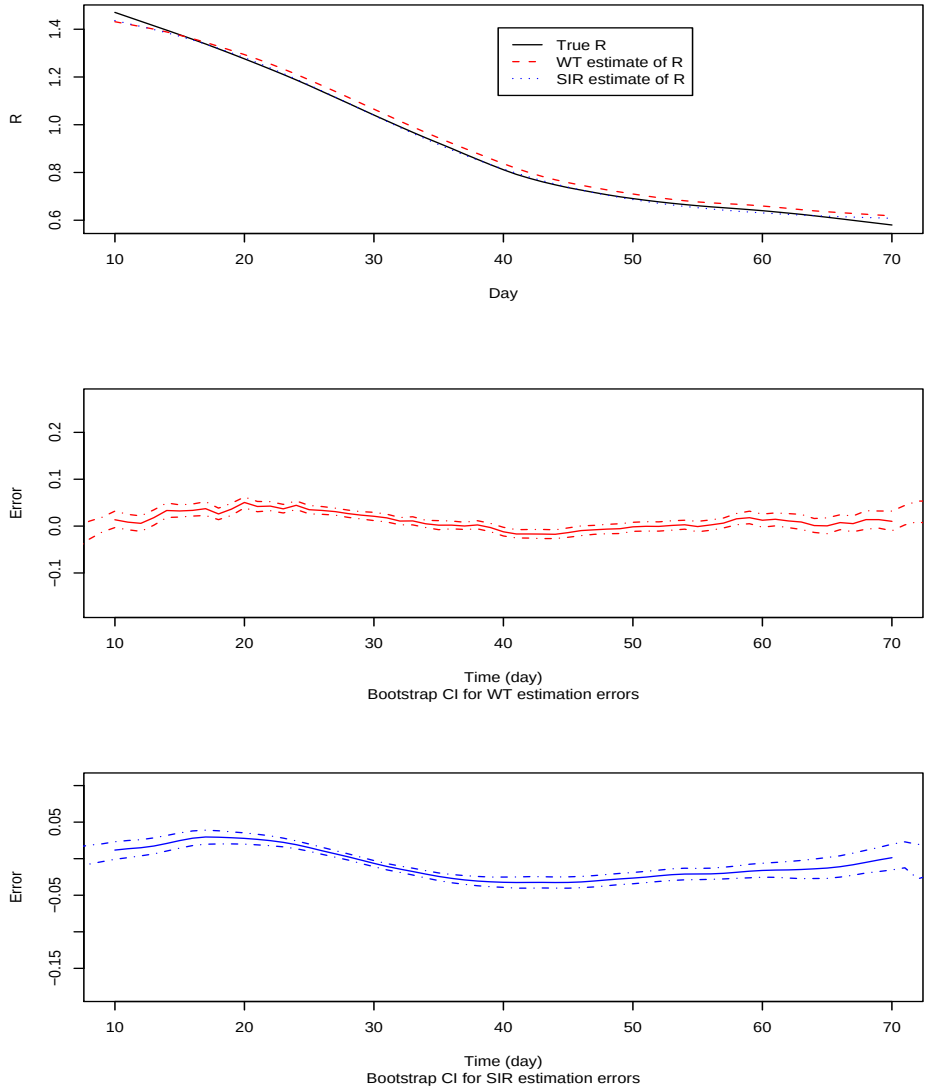

FIGURE 3. (top); True R_t and estimated R_t using the WT and SIR methods for the SIR model; (middle): Estimation errors from WT method and 95% bootstrap CIs; (bottom): Estimation errors from SIR method and 95% bootstrap CIs.

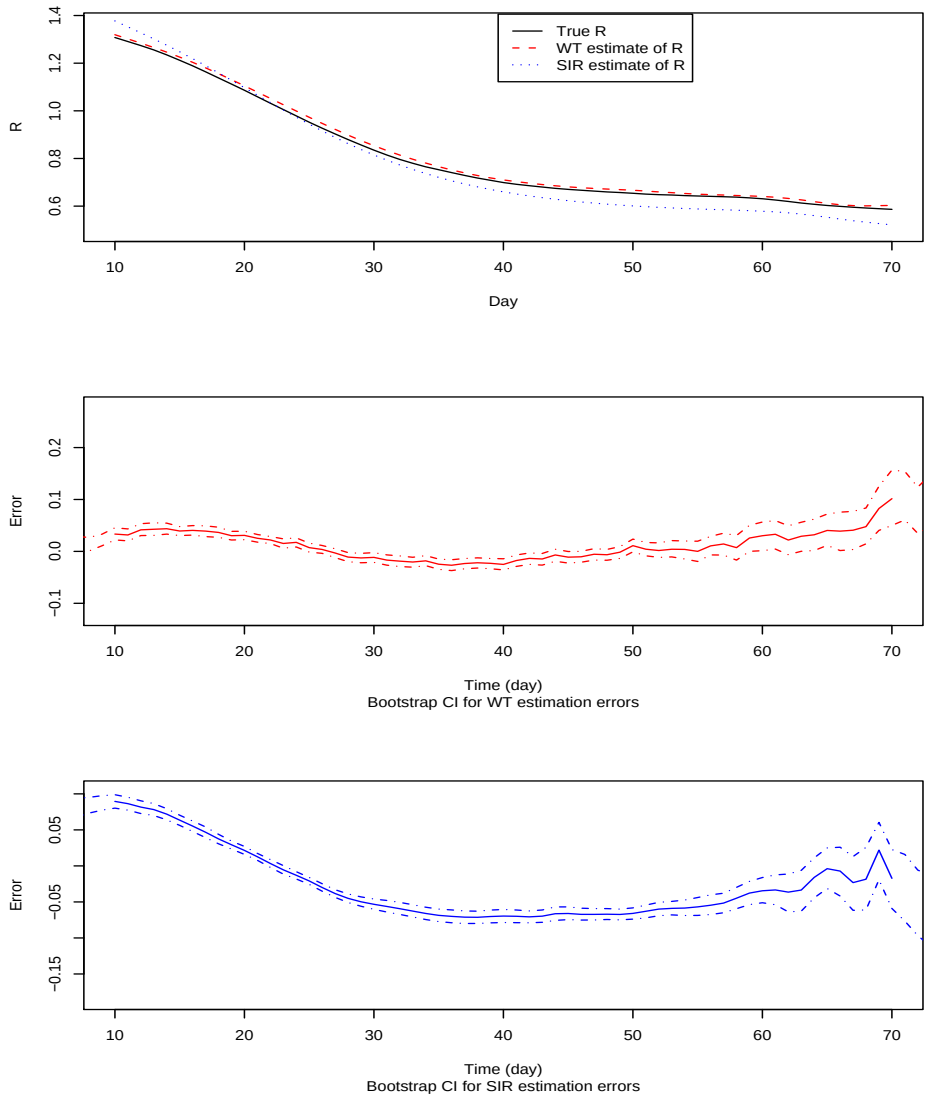

FIGURE 4. (top): True R_t and estimated R_t using the WT and SIR methods for the random network; (middle): Estimation errors from WT method and 95% bootstrap CIs; (bottom): Estimation errors from SIR method and 95% bootstrap CIs.

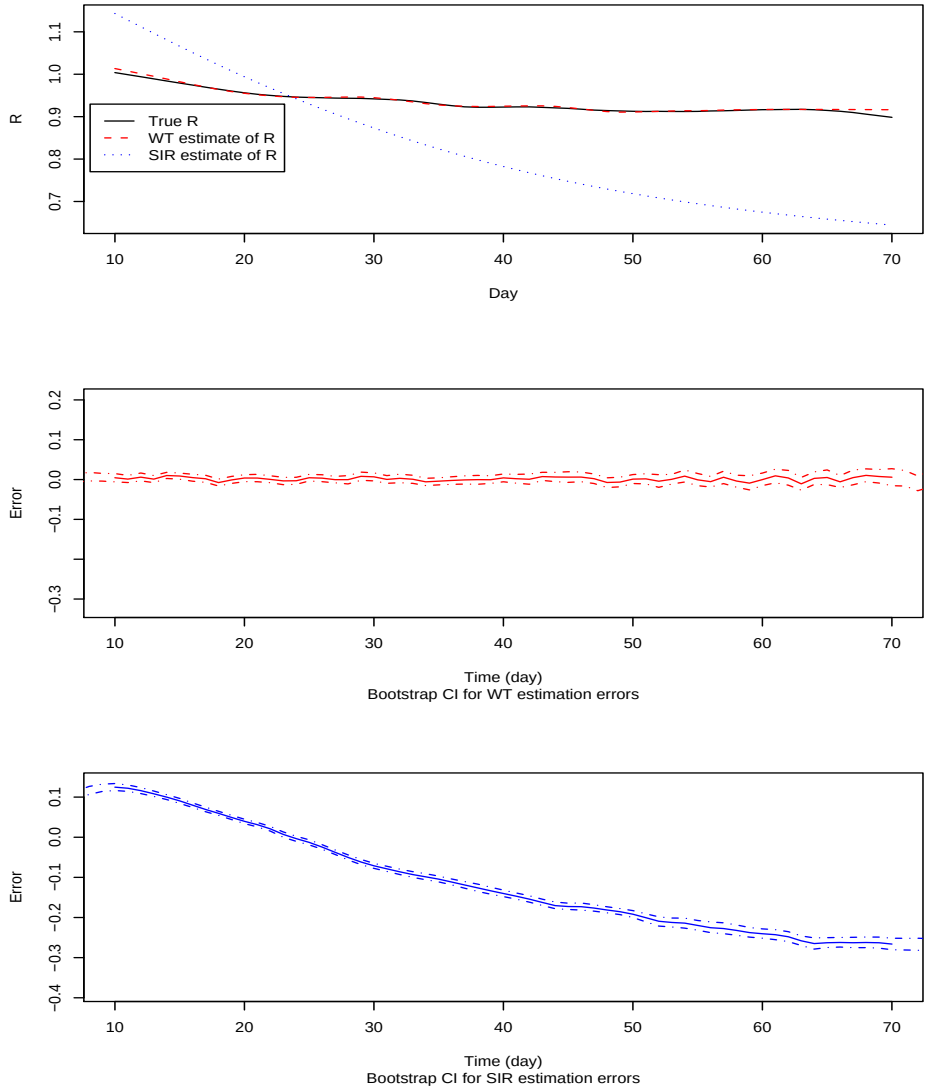

FIGURE 5. (top): True R_t and estimated R_t using the WT and SIR methods for the lattice network with long rangers; (middle): Estimation errors from WT method and 95% bootstrap CIs; (bottom): Estimation errors from SIR method and 95% bootstrap CIs.

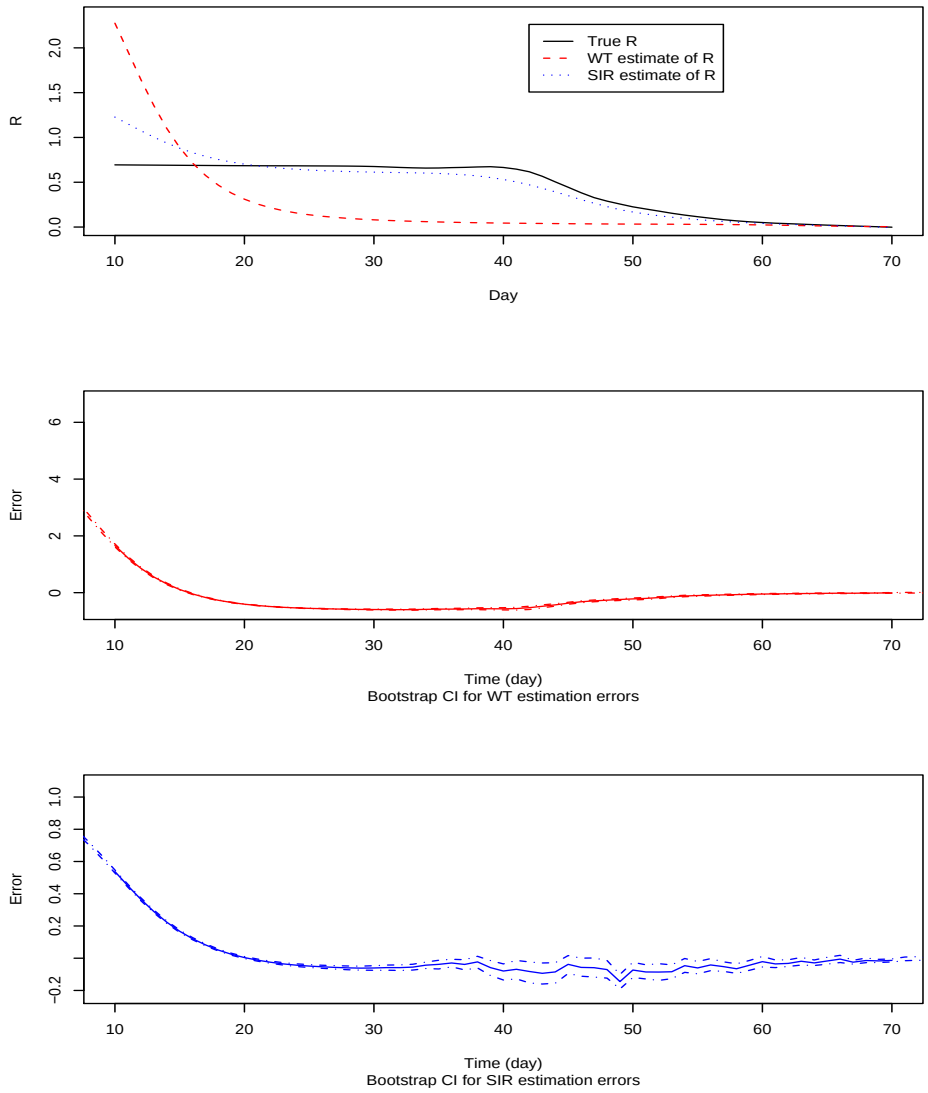

FIGURE 6. (top): True R_t and estimated R_t using the WT and SIR methods for the spatial network; (middle): Estimation errors from WT method and 95% bootstrap CIs; (bottom): Estimation errors from SIR method and 95% bootstrap CIs.

TABLE 1. Results summary for the WT and SIR methods to estimate R_t. All MAE values are repeatable to within ± 0.01 or less. The average difference column is the average estimation error in the first half of the outbreak minus the average estimation error in the second half. The p-value column is the corresponding p-value resulting from a t-test applied to the average differences over all simulations. The % positive difference column is the percent of 100 simulations in which the second half average error was larger than the first half average error. Table entries are the WT results followed by the SIR results. Zero entries are $< 10^{-10}$. All MAE results are for the approach in which the "true" R_t is calculated separately in each realization without curve smoothing or averaging over simulations.

Structure	MAE (WT,SIR)	Avg diff	p-value	% pos diff
a. SIR	0.13, 0.14	0.002, -0.001	0.60, 0.76	44, 52
b. Random	0.13, 0.14	0.005, 0.05	0.16, 0	36, 13
c. Lattice- rangers	0.10, 0.17	-0.002, 0.21	0.08, 0	59, 0
Lattice- no rangers	0.10, 0.17	0.001, 0.21	0.40, 0	41, 0
d. Spatial	0.45, 0.18	3.78, 0.72	0, 0	0, 0

5. **Discussion.** The Results section illustrates that the GI pdf is poorly estimated in all three of the non-SIR models. However, except for the spatial network, the WT method appears to be somewhat robust to poor estimation of GI pdf. The robustness extent is not yet fully understood, although deliberate use of the wrong GI pdf did not significantly increase the MAE.

A common theme in any inference setting is the well-known bias-variance trade off. There are many biased methods that outperform unbiased methods in terms of mean squared error, where informally, the mean absolute error includes the effects of both random variance and bias. This is because allowing moderate bias can sometimes greatly reduce random variance. In this setting, to implement the WT method, one could use a wide range of days to estimate the generation interval pdf to reduce random variance, but because the GI pdf changes with time, this also introduces bias. Choosing the time window to minimize the MSE would be useful in future work.

Although we evaluated the WT and the SIR-based methods, our emphasis has not been on fully characterizing these methods, but rather to compare their performance as population structure changes. For that goal, the MAE results are adequate. We remarked in the Results section that confidence limits for R_t can be based on using $\hat{R}_t \pm 2 \times MAE$ for each case. However, this requires the MAE results in Table 1. In practice, if confidence limits are desired, then [8] provided approximate confidence limits for the WT method, assuming the GI pdf is known and constant (so our results and [23] suggest that further work is needed here).

6. **Conclusion and summary.** We demonstrated surprisingly good robustness to structure-based violations of the SIR model of both the WT and SIR-based methods to estimate R_t, with the SIR-based method doing surprisingly well in all cases and the WT method doing slightly better except in the spatial network.

The WT method and the SIR-based method require somewhat different data sources in order to be implemented so might not both be available in all situations. An attractive feature of the WT method is that it can be implemented using only the number of new infecteds each time step provided an estimate of the recovery rate γ is available so that the GI pdf can be estimated. To date, either an auxiliary data source such as contact tracing has been used to estimate the GI pdf using an approximating Weibull distribution, or the hazard rate approach ("race to infect") as in [23] has been suggested. The hazard rate approach properly adjusts for the time dependence of the GI pdf. In contrast, the SIR approach must estimate $\frac{S_t}{N}$ in addition to the transmission rate β and the recovery rate γ.

Fortunately, the MAE of the WT method when using only 10 generation times to estimate the GI pdf was essentially the same as when using hundreds of generation times in the estimate. And, although the GI pdf is quite different among the structured networks compared to the GI pdf in the unstructured population (Fig. 2), it appears not to be necessary to estimate the GI pdf for each outbreak, because the WT method is fairly robust to using the "wrong" GI pdf estimate, even if the estimate arises from the "wrong" network structure. As further study, other noise sources could be investigated, such as reporting delays or underreporting of cases [41].

We also demonstrated that the GI pdf is not constant over time, as is also shown in [23] in continuous time. Figure 2 showed time-dependent behavior in the average GI, μ_i. A simple explanation involving the time dependent behavior of the number of new infecteds each day was provided, leading to an accurate calculation of μ_i on day i for the SIR-case. This same calculation is not accurate for the structured populations studies. It would also be useful to explore why the WT method, which assumes μ_i is constant over time, does not lead to significant patterns in the daily estimation errors of μ_i in the unstructured population cases (except in the spatial network). Perhaps such robustness arises primarily because the WT method involves a ratio of relative likelihoods, and so can tolerate some degree of misspecification in the GI pdf. Stronger time patterns in the GI pdf could also be investigated by considering cases having larger R_0.

Although it is also of interest to extend this study to include modified SIR-type models such as the SEIR (susceptible-exposed-infectious-recovered) model, we do not believe it will alter our basic conclusion that the WT method is surprisingly robust to structure-type violations of SIR assumptions.

Finally, in small-population stochastic models, R_0 is not sufficient information to predict whether an epidemic can sustain itself [33]. For example, the simulated spatial network has $R_0 < 1$, yet the epidemic's final size is approximately 60% of the population size N, indicating that we need a better understanding of small-population stochastic models.

Acknowledgments. We thank Joel Miller and the two reviewers for their critical comments that helped improve this manuscript. We also extend our thanks and congratulations to Fred Brauer and Karl Hadeler for their many contributions to the field of mathematical biology.

REFERENCES

[1] R. Anderson and R. May, "Infectious Diseases of Humans," Oxford University Press, Oxford, 1991.

[2] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science, **286** (1999), 509–512.
[3] F. Brauer, *Some simple epidemic models*, Mathematical Biosciences and Engineering, **3** (2006), 1–15.
[4] F. Brauer and C. Castillo-Chavez, "Mathematical Models in Population Biology and Epidemiology," Springer-Verlag, New York, 2000.
[5] T. Britton, *A test of homogeneity versus a specified heterogeneity in an epidemic model*, Mathematical Biosciences, **141** (2007), 79–99.
[6] T. Burr and G. Chowell, *Observation and model error effects on parameter estimates in susceptible-infected-recovered epidemic models*, Far East Journal of Theoretical Statistics, **19** (2006), 163–183.
[7] T. Burr and G. Chowell, *Signatures of non-homogeneous mixing in disease outbreaks*, Mathematical and Computer Modeling, **48** (2008), 122–140.
[8] S. Cauchemez, P. Boelle, C. Donnelly, N. Ferguson, G. Thomas, G. Leung, A. Hedley, R. Anderson and A. Valleon, *Real time estimates in early detection of SARS*, Emerging Infectious Diseases, **12** (2006), 110–113.
[9] S. Cauchemez, P. Boelle, G. Thomas and A. Valleron, *Estimating in real time the efficacy of measures to control emerging communicable diseases*, American Journal of Epidemiology, **164** (2006), 591–597.
[10] G. Chowell, J. Hyman, S. Eubank and C. Castillo-Chavez, *Scaling laws for the movement of people between locations in a large city*, Physical Review E, **68** (2003), 066102.
[11] G. Chowell, C. Ammon, N. Hengartner and J. Hyman, *Estimating the reproduction number from the initial phase of the Spanish flu pandemic waves in Geneva, Switzerland*, Mathematical Biosciences and Engineering, **4** (2007), 457–470.
[12] V. Colizza, A. Barrat, M. Barthelemy and A. Vespignani, *The modeling of global epidemics: Stochastic dynamics and predictability*, Bulletin of Mathematical Biology, **68** (2006), 1893–1921.
[13] J. Dall and M. Christensen, *Random geometric graphs*, Physical Review E, **66** (2002), 016121.
[14] O. Diekmann and J. Heesterbeek, "Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation," John Wiley & Sons, New York, 2000.
[15] P. Fine, *The interval between successive cases of an infectious disease*, American Journal of Epidemiology, **158** (2003), 1039–1047.
[16] K. Hadeler and C. Castillo-Chavez, *A core group model for disease transmission*, Mathematical Biosciences and Engineering, **128** (1995), 41–55.
[17] H. Hethcote and J. Yorke, "Gonorrhea Transmission Dynamics and Control," Lecture Notes in Biomathematics, **56**, Springer, Berlin, 1984.
[18] E. Holmes, "Basic Epidemiological Concepts in a Spatial Context," pp. 111. In: D. Tilman, P. Kareiva, Eds. 1997. "Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions," Monographs in Population Biology, Princeton University Press.
[19] M. Hugh-Jones and R. Tinline, *Studies on the 1967–68 foot and mouth disease epidemic: incubation period and herd serial interval*, Journal of Hygiene, **77** (1976), 141–153.
[20] J. Hyman and J. Li, *Disease transmission models with biased partnership selection*, Applied Numerical Mathematics, **24** (1997), 1–14.
[21] M. Keeling and K. Eames, *Networks and epidemic models*, Journal of the Royal Society Interface, **2** (2005), 295–307.
[22] W. Kermack and A. McKendrick, *A contribution to the mathematical theory of epidemics*, Proceedings of the Royal Society A, **115** (1927), 700–721.
[23] E. Kenah, M. Lipsitch and J. Robins, *Generation interval contraction and epidemic data analysis*, Mathematical Biosciences, **213** (2007), 71–79.
[24] I. Kiss, D. Green and R. Kao, *The effect of contact heterogeneity and multiple routes of transmission on final epidemic size*, Mathematical Biosciences, **203** (2006), 124–136.
[25] F. Liljeros, C. Edling, L. Nunes Amaral, H. Stanley and Y. Aberg, *The web of human sexual contacts*, Nature, **411** (2001), 907–908.
[26] M. Lipsitch, T. Cohen, B. Cooper, J. Robins, S. Ma, L. James, G. Gopalakrishna, S. Chew, C. Tan, M. Samore, D. Fisman and M. Murray, *Transmission dynamics and control of severe acute respiratory syndrome*, Science, **300** (2003), 1966–1970.
[27] D. Ludwig and D. Jones, *Qualitative analysis of insect outbreak systems: The spruce budworm and forest*, The Journal of Animal Ecology, **47** (1978), 315–322.

[28] D. Ludwig, *Final size distributions for epidemics*, Mathematical Biosciences, **23** (1975), 33–46.
[29] S. Marion and C. Greenwood, *Computation of the transient and final size distribution for an epidemic in a finite heterogeneous population*, Abstract in: Conference on highly structured stochastic systems, Pavia, Italy. September, 1999.
[30] J. Ma and D. Earn, *Generality of the final size formula for an epidemic of a newly invading infectious disease*, Bulletin of Mathematical Biology, **68** (2006), 679–702.
[31] R. May and R. Anderson, *Spatial heterogeneity and the design of immunization programs*, Mathematical Biosciences, **72** (1984), 83-111.
[32] L. Meyers, B. Pourbohloul, M. Newman, D. Skowronski and R. Brunham, *Network theory and SARS: Predicting outbreak diversity*, Journal of Theoretical Biology, **232** (2005), 71–81.
[33] R. Miller, "Beyond ANOVA, Basics of Applied Statistics," John Wiley & Sons, New York, 1986.
[34] J. C. Miller, *The spread of infectious diseases through clustered populations*, J. Royal Society Interface, In press.
[35] M. Newman, *Coauthorship networks and patterns of scientific collaboration*, Proc. Natl. Acad. Sci. USA, **98** (2003), 404–409.
[36] H. Nishiura, *Time variations in the transmissibility of pandemic influenza in Prussia, Germany, from 1918-19*, Theor Biol Med Model, **4** (2007), 4–20.
[37] The R Project for Statistical Computing. http://www.r-project.org.
[38] L. Sattenspiel and C. Simon, *The spread and persistence of infectious diseases in structured population*, Mathematical Biosciences, **90** (1988), 341–366.
[39] Splus for Windows version 7.0, 2005, Insightful Corporation.
[40] A. Svensson, *A note on generation times in epidemic models*, Mathematical Biosciences, **208** (2007), 300–311.
[41] J. Wallinga and P. Teunis, *Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures*, American Journal of Epidemiology, **160** (2004), 509–516.
[42] J. Wallinga and M. Lipsitch, *How generation intervals shape the relationship between growth rates and reproduction numbers*, Proceedings Biological Sciences, **274** (2007), 599–604.
[43] D. Watts and S. Strogatz, *Collective dynamics of "small-world" networks*, Nature, **393** (1998), 440.
[44] L. White and M. Pagano, *A likelihood-based method for real-time esimation of the serial interval and reproduction number of an epidemic*, Statistics in Medicine, **27** (2007), 2999–3016.

Received November 5, 2007. Accepted September 12, 2008.

E-mail address: tburr@lanl.gov
E-mail address: gchowell@asu.edu