Observation and model error effects on parameter estimates in susceptible-infected-recovered epidemiological models

Tom L. Burr[1*] & Gerardo Chowell[2]

[1] Statistical Sciences (MS F600),

[2] Mathematical Modeling and Analysis (MS B284),

Los Alamos National Laboratory

Los Alamos, New Mexico 87545, USA.

[*]Corresponding author. Email: tburr@lanl.gov Fax: (505) 667-4470

Abstract

Recently, confidence intervals (CIs) associated with parameter estimates in the susceptible-infected-recovered epidemiological model have been developed. When model assumptions are met and the observation error is relatively small, these CIs are relatively short. This work describes the behavior of CIs for parameters as observation and/or equation or model error becomes larger, and includes a comparison of two estimation procedures. One procedure demonstrates significant bias as observation error increases; the other procedure demonstrates significant bias as model error increases.

Keywords: epidemiology, susceptible-infected-recovered model uncertainty, parameter estimates, bias, mean squared error.

1 Introduction

Recently, confidence intervals (CIs) associated with parameter estimates in the susceptible-infected-recovered (SIR) epidemiological model have been developed (Chowell et. al. [4]). When model assumptions are met and the observation error is relatively small, these CIs are relatively short, as we will illustrate. This work describes the behavior of CIs for parameters as observation and/or equation or model error becomes larger, and includes a comparison of two estimation procedures. The first procedure fits a simple linear regression relating the per-time-step response and predictors. This procedure demonstrates significant bias as observation error increases. In general, observation error in predictors leads to bias to varying degrees, as has been illustrated in the "errors-in-variables" literature (Carroll et al. [3]). For example, bias arising from using measurements of species abundance rather than true species abundance has recently

been reported in the context of a biological random walk extinction model (Buonaccorsi et al. [2]). The other procedure evaluated here solves the nonlinear differential equations to produce parameter estimates, thereby relying heavily on the shape of the observed epidemic curve, and mitigating the effects of any errors, such as observation errors, that do not distort the curve's shape. This method demonstrates significant bias if model error increases sufficiently to distort the curve's shape.

Following sections include a brief review of SIR models; associated parameter estimation, and performance results for several cases having observation or observation plus model errors.

2 Background

Mathematical modeling of infectious diseases can be traced to the work of Sir Ronald Ross (Ross et al. [21]) who discovered the vector mechanism of transmission of Malaria and explored the effects of controlling the mosquito population using simple mathematical models. The SIR model proposed by Kermack and MacKendrick [12] provides an established basis to model the transmission dynamics of infectious diseases. The SIR model classifies individuals as susceptible (S), infectious (I), and recovered (R). Susceptible individuals in contact with the virus enter the infectious class at the rate $\beta I(t)/N$ where β is the transmission rate, $I(t)$ is the number of infectious individuals at time t and $N(t) = S(t) + I(t) + R(t)$ is the total population at time t. This assumes that the disease latency period is negligible. The classical SIR model assumes homogeneous mixing between individuals and, therefore, the fraction $I(t)/N$ is the probability that a random contact would be with an infectious individual. For simplicity, we also assume that the time-scale of the epidemic is much faster than those of demographic

processes (natural birth and death). Moreover, recovered individuals are assumed to acquire immunity to the disease for at least the duration of the outbreak. The SIR transmission process (single outbreak) can then be modeled using the 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.

Alternatively, the differential equations can be viewed as difference equations

$$\begin{cases} S(t) - S(t-1) &= -\beta S(t-1) I(t-1)/N \\ I(t) - I(t-1) &= \beta S(t-1) I(t-1)/N - \gamma I(t-1) \\ R(t) - R(t-1) &= \gamma I(t-1) \end{cases} \tag{2}$$

which suggests that an estimate of β can be obtained by regressing the response $S(t) - S(t-1)$ versus the predictor $S(t-1)I(t-1)/N$, and an estimate of γ can be obtained by regressing the response $R(t) - R(t-1)$ versus the predictor $I(t-1)$. However, it is well known that errors in predictors leads to bias in the estimated parameters. The effects of errors in predictors is studied in an area of research that is often referred to as "errors-in-variables" (Carroll et al. [3]). Bias will be present to varying degrees in our examples to follow. Using the difference equations, if only the R or only the I series is observed, it is still possible to estimate β and γ using regression, although performance degrades.

2.1 The reproductive number and control measures

The basic reproductive number ($\mathcal{R}_0$) is the number of secondary cases generated by a primary infectious individual during its entire period of infectiousness in a completely susceptible population (Anderson and May [1]). For the SIR epidemic model, $\mathcal{R}_0$ is the product of the transmission rate β and the duration of the infectious period $1/\gamma$, so $\mathcal{R}_0 = \beta/\gamma$. Estimates of $\mathcal{R}_0$ can be obtained by first estimating β and γ from epidemic curve data, and then substituting the resulting estimates into the $\mathcal{R}_0$ expression. The estimation of the basic reproductive number by fitting epidemic models to epidemic curve data is probably the most widespread approach (examples of recent work include Gani et al. [9], Riley et al. [20] and Chowell et al. [4]). Another common approach to estimate $\mathcal{R}_0$ consists in estimating first the initial exponential growth rate characteristic of most human infectious diseases of rapid dissemination. Then $\mathcal{R}_0$ can be estimated by substituting the estimate of the initial exponential growth rate into a formula derived from the linearization of the deterministic epidemic model (e.g., Anderson and May [1], Nowak et al. [17], Lloyd [14], and Lipsitch et al. [13]). A recent review on the estimation of $\mathcal{R}_0$ from epidemiological data is given by Heffernan et al. [10].

While the basic reproductive number is more relevant for the case of emerging infectious diseases, in the case of endemic or recurrent infectious diseases, the reproductive number ($\mathcal{R}$) is a more practical quantity because it accounts for the residual immunity in the population due to previous exposures or vaccination campaigns in the population. Hence, we can write $\mathcal{R} = s^*\mathcal{R}_0$ where s^* is the proportion of the population that is effectively susceptible. One also needs to account for the effect of control measures aiming at reducing the transmissibility of the disease. Once the effects of public health measures begin to take hold and the number

of susceptible individuals decreases, the reproductive number will decay. Hence, the goal of public health authorities is to bring the reproductive number to a number less than one as soon as possible. The types of available interventions will depend on the disease in question and availability of appropriate resources.

One main goal of SIR-type models is to predict the effect of candidate control measures on $\mathcal{R}$. However, until the relative importance of observational and model errors is well estimated, it is not clear how reliably SIR-type models can estimate the effect of possible control measures. For example, if the true population is structured but the method, such as the SIR model, assumes the population is homogeneously mixing, then parameter estimates and associated mitigation recommendations will be of questionable value.

Typically, the more mathematically tractable models make the least realistic assumptions but provide insights on how different factors affect the disease dynamics. For example, we mentioned that the standard SIR model assumes that infecteds mix homogeneously with susceptibles. Recent expansions of the SIR model allow varying degrees of heterogeneity. Chowell et. al. [5] fit a spatially mixing model for the spread of foot-and-mouth disease in which infected farms were more likely to infect neighboring farms. Demiris and O'Neil [6] used structured populations in which individuals were partitioned into households and the contact rate was higher between individuals in the same family. Still more detail is included in agent-based computational models such as EpiSims (Eubank et. al. [8]). It is generally believed that models that more closely resemble reality by including such detail can more accurately investigate the impact of factors that influence disease spread. If this is true, then the best recommendations concerning mitigation strategies (vaccination, quarantine, movement restrictions, etc.) should

tend to arise from the richer-model evaluations. It is also generally believed that richer models are more vulnerable to poorly specified inputs, such as the mixing/contact structure, rate of disease spread, etc. Therefore, richer models must be implemented in a manner that easily facilitates performing sensitivity studies.

3 Sensitivity Study

This section presents a sensitivity study concerning two types of errors that impact parameter estimation in the basic SIR model. The least harmful error type is pure observation error. More harmful error is model misspecification.

Chowell et. al. [4] modeled observation error in the R_{observed} time series. A Brownian Bridge approximation leads to approximate 95% confidence intervals (CI) for $\hat{\beta}$ and $\hat{\gamma}$. These CIs should be interpreted as containing 95% of future estimates when the same assumptions are made and the only noise source is observation error. It is tempting but incorrect to interpret these CIs as containing the "true" parameters with probability 0.95. We do not address Bayesian versus non Bayesian interpretations of CIs here. Instead, our focus is on the issue of repeatability versus accuracy. The Chowell et. al. [4] result is a repeatability result.

The notion of accuracy in epidemiological models requires careful interpretation, because the parameters are almost never believed to represent real phenomena such as, e.g., the speed of light in a vacuum. Because no real population obeys the homogeneous mixing assumption of the SIR model, it is not meaningful to estimate how close $\hat{\beta}$ is likely to be to β. The most meaningful such comparison is to compare the model-based estimate of the impact of candidate mitigation strategies to the true impact. Because such comparisons typically involve one or a

few realizations of the stochastic epidemiological situation, this can typically only be done by stochastically defining the true impact of the strategy. This is a large task that is beyond our scope here. Instead, we investigate the second type of error, model error, on the SIR model.

Model or equation error is typically the largest error source in epidemiology. Consider adding an equation error to the difference equations in Eq. (2). Then, for example, equation error that impacts $S(2)$ will impact $S(t)$ for $t > 2$. This means that equation error propagates in potentially harmful ways. For brevity, we will express all three examples of model error to follow by assuming a constant true β when the "true" β value varies with time. There are physical reasons for such models (such as unmodeled behavior changes), plus we have verified that this type of model error is capable of representing other model errors in which the assumed β equals the true β at each "epidemic day," which is constant over time, but an unobserved equation error is added to some subset of the S, I, and R difference equations.

3.1 Parameter estimation

For the SIR model, we estimated the transmission rate β and the inverse of the infectious period γ using two methods. Both methods used $\hat{\mathcal{R}}_0 = \hat{\beta}/\hat{\gamma}$ to estimate $\mathcal{R}_0$.

3.1.1 Method 1: nonlinear differential equation fitting

This method used least-squares fitting of the equation $R(t)$ in Eq. (1) to epidemic data of the cumulative number of observed recovered cases. For the least-squares fitting procedure, we used the Levenberg-Marquardt method with line-search implemented in MATLAB [15] in the built-in routine lsqcurvefit which is part of the optimization toolbox.

3.1.2 Method 2: simple regression fitting

This method used the difference equations in Eq. (2) applied to the time series of the observed R and of the observed I data. To estimate β the predictor was $S(t-1)I(t-1)/N$ and the response was $S(t) - S(t-1)$. To estimate γ the predictor was $I(t-1)$ and the response was $R(t) - R(t-1)$. Other choices are possible depending on what subset of S, I, and R are observed. Again least-squares fitting was used, as implemented in Splus (Splus 2004), but without knowledge of the underlying equations beyond the implicit knowledge used to define the responses and predictors. This method should demonstrate bias provided the observation error in the predictors is nonnegligible compared to the range of the predictors (Carroll et al. [3]).

3.2 Observation error

We simulated epidemic curves using the parametric bootstrap (Efron, 1986) following Chowell et. al. [4]. The observation error in the R_{observed} time series was generated by assuming each increment $\Delta R_{\text{observed}}$ has a Poisson distribution, $\Delta R_{\text{observed}} \sim \text{Poisson}(\mu = \Delta R_{\text{true}})$. The parameter estimation procedures (described above) were then applied for each of 200 simulated epidemic curves. Because the mean (μ) equals the variance (σ^2) for the Poisson distribution, we also considered more extreme cases with the variance being $\sigma^2 = k\mu$ where $k = 2,3,4,5$ by using the Gamma or Negative binomial probability distribution. The results reported here are for the Gamma distribution, and are essentially the same as for the Negative binomial distribution.

3.3 Model error

We considered three types of model error, each expressed as ignored time variation in the true β: (1) β varies around β_{avg} according to a uniform distribution. Specifically, each epidemic day has a unique β value and the time series of β values is independently and identically distributed according to $\beta \sim \text{Uniform}(\beta_{avg} - \Delta_\beta, \beta_{avg} + \Delta_\beta)$; (2) β varies as an autoregressive time series, with $\beta_t = a\beta_{t-1} + \epsilon_t$ (t in days). A lag-one autoregressive time series for β, $\beta_t = a\beta_{t-1} + \epsilon_t$, has variance $\sigma_\beta^2 = (1 + a^2) \times \sigma_e^2$, where σ_e^2 is the variance of the error sequence e_t. We chose $a = 0.99$, and let the error sequence e_t be independently distributed with a uniform distribution having the correct range to correspond to a σ_e^2 that gave the desired value of σ_β^2 (equal to $\frac{0.3^2}{12}$ for the $\Delta_\beta = 0.15$ case for example); and (3) β varies deterministically, equal to a constant until day 15, then exponentially decaying to a new value, so $\beta = \beta_1$ for $t \leq 15$ and $\beta = \beta_2 + (\beta_1 - \beta_2) \exp(-0.2 \times (t - 15))$ for $t > 15$.

4 Results

We experimented with several cases, four of which we will summarize here, divided into cases having observation error only, or observation error and model error. Model errors are potentially the most harmful because they tend to propagate. For example, $S(2)$ will impact $S(t)$ for $t > 2$, and similarly for $I(t)$ and $R(t)$. The cases summarized here are referred to as cases A, B, C, and D. The estimated (using Method 1) and observed number of recovered ($R(t)$) cases for each case are plotted in Figure 1.

All cases include N values on a grid from 100 to 5000, with the observation error having variance $\sigma^2 = k\mu$ from a gamma distribution, with $k = 2$ or $k = 5$. In addition, all cases used

120 days, $\beta = 0.5$, and $\gamma = 0.25$, which corresponds to $\mathcal{R}_0 = 2$. The initial number of infecteds was 5.

The results of 200 simulations for each subcase are summarized by reporting the root average squared scaled error, RASSE $= \frac{1}{\beta}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\beta}_i - \beta)^2}$, which measures the combined effect of bias and precision of the estimator $\hat{\beta}$, and the average scaled bias, $\frac{1}{\beta n}\sum_{i=1}^{n}(\hat{\beta}_i - \beta)$. These two performance measures are reported for $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\mathcal{R}}_0$.

4.1 Observation Error

4.1.1 Case A

Case A is observation error only. Results for Case A are summarized in Figure 2 and Table 1 for Method 1 and for Method 2. Most entries in Tables 1 to 4 are within approximately 0.01 of replicate results on the basis of a second set of 200 simulations. Method 1 outperforms Method 2 for $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\mathcal{R}}_0$. Method 2 demonstrates sizeable bias for the larger variance (smaller N, and $k = 5$) cases. Fortunately, both methods demonstrate a tendency for the bias in $\hat{\beta}$ to be in the same direction as that of $\hat{\gamma}$, resulting in a smaller bias in $\hat{\mathcal{R}}_0$ than if that tendency were not present. This same tendency is demonstrated when model error is included, although performance is generally worse. Both methods improve as N increases.

4.2 Observation Error and Model Error

Generally, results are significantly worse when model error is included, with Method 2 showing a potential advantage. As mentioned, Method 1 degrades in some cases with model error because it relies on the relation between the β and γ parameters and the shape of the $R(t)$ curve.

4.2.1 Case B

Results for Case B are summarized in Table 2 for Method 1 and for Method 2. Both methods do worse, but not much worse, than they did for Case A. Both methods improve as N increases.

4.2.2 Case C

Results for Case C are summarized in Table 3 for Method 1 and for Method 2. Both methods demonstrate large bias for small N, but this decreases for larger N. Especially for Method 2, and somewhat for Method 1, the RASSE for $\hat{\mathcal{R}}_0$ does not noticeably decrease as N increases.

4.2.3 Case D

Results for Case D are summarized in Figure 3 and Table 4 for Method 1 and for Method 2. This is the worst case for both methods. Method 1 does much worse than Method 2, and does worse as N increases.

5 Summary

By definition, a model is an intentional simplification of a complex process. Because any model is therefore "wrong," a key question is whether a particular model is useful. SIR-type models have a long history in mathematical biology, with many successful and useful applications documented (e.g., Kermack and MacKendrick [12], Ross [21], Hethcote and Yorke [11], Perelson et al. [19], May [16], Okubo and Levin [18]). An important potential application of disease outbreak modeling is the ability to evaluate candidate mitigation strategies such as quarantine, vaccine, school closures, etc.

One aspect of model validation is the ability to estimate model parameters from real or simulated outbreaks. We have evaluated simulated outbreaks that either exactly followed the assumed model or that arose from a different model than the estimation procedure assumed. The former situation included observation error while the latter included both model and observation error. Two estimation methods were empirically compared on simulated data sets. Both methods demonstrated bias due to observation error. Method 1, which relies on solving the SIR model's differential equations, tended to demonstrate less bias in the presence of observation error alone, but for some of the model error examples, it demonstrated much worse bias and total error. This is because Method 1 relies strongly on the shape of the observed number of recovered cases each time step. Results are summarized in the figures and tables. Generally, parameter estimates had acceptably small bias and total error in most, but not all cases.

It should be mentioned that if it could be known that a particular type of model departure (such as a change in the β parameter over time) was known to be in effect, then alternative estimation strategies could be developed. This was not our goal here. Instead, we used time dependence in β to generate model error. As mentioned, alternatively, errors could be added at each "epidemic day" to some subset of the S, I, and R difference equations while keeping β constant in time. Similarly, if the observation error variance is known, then there are several known methods for making bias adjustments to estimators. However, this was not our goal here, and these methods do not necessarily reduce the average root squared error in the estimator.

References

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

[2] J. Buonaccorsi, J. Staudenmayer, M. Carreras, Modeling observation error and its effects in a random walk/extinction model, Theoretical Population Biology (in press).

[3] R. Carroll, D. Ruppert, and L. Stefanski, Measurement Error in Nonlinear Models, Chapman & Hall: London, 1996.

[4] G. Chowell, E. Shim, F. Brauer, P. Díaz-Dueñas, J.M. Hyman, C. Castillo-Chavez, Modeling the transmission dynamics of acute hemorrhagic conjunctivitis: Application to the 2003 outbreak in Mexico, Stat. Med. (in press).

[5] G. Chowell, A. Rivas, N. W. Hengartner, J. M. Hyman, and C. Castillo-Chavez, The role of spatial mixing in the spread of foot-and-mouth disease, Prev. Vet. Med. 73 (2006), 297-314.

[6] N. Dimiris and P. O'Neill, Bayesian inference for stochastic multitype epidemics in structured populations via random graphs, Journal of the Royal Statistical Society B 67 (2005), 731-745.

[7] B. Efron and R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, Statistical Science 1 (1986), 54-75.

[8] S. Eubank, H. Guclu, V. Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang, Modeling disease outbreaks in realistic urban social networks, Nature 429 (2004), 180-184.

[9] R. Gani, S. Leach, Transmission potential of smallpox in contemporary populations, Nature 414 (2001), 748-751.

[10] J.M. Heffernan, R.J. Smith, L.M. Wahl. Perspectives on the basic reproductive ratio. Journal of The Royal Society Interface 2 (2005), 281-293.

[11] H.W. Hethcote and J.A. Yorke, Gonorrhea Transmission Dynamics and Control, Lecture Notes in Biomathematics 56, Springer, Berlin, 1984, 105 pages, ISBN 0-387-13870-6.

[12] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. Roy. Soc. A 115 (1927), 700-721.

[13] M. Lipsitch, T. Cohen, B. Cooper, et al, Transmission dynamics and control of severe acute respiratory syndrome, Science 300 (2003), 1966-1970.

[14] A.L. Lloyd, The dependence of viral parameter estimates on the assumed viral life cycle: limitations of studies of viral load data, Proc. R. Soc. B 268 (2001), 847-854.

[15] Matlab, Mathworks Inc., Natick, Mass.

[16] R. M. May, Stability and Complexity in Model Ecosystems, Princeton University Press, Princeton NJ, 1974.

[17] M.A. Nowak, A.L. Lloyd, G.M. Vasquez, et al, Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection, J. Virol. 71(1997), 7518-7525.

[18] A. Okubo, S.A. Levin, A theoretical framework for data analysis of wind dispersal of seeds and pollen, Ecology 70 (1989), 329-338.

[19] A.S. Perelson, A.U. Neumann, M. Markowitz, J.M. Leonard, D.D. Ho, HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time, Science 271 (1996), 1582-1586.

[20] S. Riley, C. Fraser, C.A. Donnelly, et al, Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions, Science 300 (2003), 1961-1966.

[21] R. Ross, The Prevention of Malaria, John Murray, London, 1911.

[22] Splus6.1 for Unix, Insightful Corp, 2004.

Table 1: Case A: Performance results for SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) with observation error only through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of k, and a comparison of two different estimation methods.

N	k	β				γ				$\mathcal{R}_0$			
		ASB		RASSE		ASB		RASSE		ASB		RASSE	
		M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
100	2	0.01	0.06	0.12	0.42	0.02	0.03	0.22	0.51	0.06	0.11	0.38	0.34
500	2	0.00	0.02	0.04	0.17	0.00	0.00	0.13	0.23	0.02	0.04	0.13	0.18
1000	2	0.00	0.00	0.03	0.12	0.00	-0.01	0.11	0.17	0.01	0.03	0.09	0.14
1500	2	0.00	0.01	0.03	0.10	-0.01	0.00	0.09	0.14	0.01	0.02	0.07	0.10
3000	2	0.00	0.00	0.02	0.07	0.00	-0.01	0.06	0.09	0.00	0.02	0.04	0.07
5000	2	0.00	0.00	0.02	0.05	0.00	-0.01	0.05	0.07	0.00	0.01	0.03	0.05
100	5	0.01	0.21	0.18	0.85	0.08	0.23	0.30	1.01	-0.01	0.19	0.32	0.80
500	5	0.02	0.01	0.07	0.30	0.04	-0.03	0.22	0.38	0.02	0.10	0.23	0.28
1000	5	0.01	0.01	0.05	0.18	0.01	-0.01	0.16	0.24	0.02	0.05	0.15	0.19
1500	5	0.00	0.00	0.04	0.16	0.00	-0.03	0.12	0.21	0.01	0.05	0.10	0.18
3000	5	0.00	0.02	0.03	0.11	-0.01	0.02	0.09	0.16	0.01	0.02	0.07	0.13
5000	5	0.00	0.02	0.03	0.09	0.00	0.00	0.08	0.12	0.01	0.02	0.05	0.08

N = Population size; β = Transmission rate; γ = Recovery rate; $\mathcal{R}_0$ = Basic reproductive number; ASB = Average scaled bias; RASSE= Root average squared scaled error; M_1 = Nonlinear differential equation fitting; M_2 = Simple regression fitting.

Table 2: Case B: Performance results for SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) with observation error through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of k plus model error in the transmission rate β such that a different value of β for each epidemic day is independently and identically distributed according to $\sim$ Uniform($\beta_{avg} - \Delta_\beta, \beta_{avg} + \Delta_\beta$) with $\Delta_\beta = 0.15$.

N	k	β				γ				$\mathcal{R}_0$			
		ASB		RASSE		ASB		RASSE		ASB		RASSE	
		M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
100	2	0.01	0.06	0.14	0.41	0.01	0.56	0.23	0.32	0.06	0.34	0.31	0.81
500	2	0.00	0.00	0.06	0.07	-0.01	-0.01	0.14	0.09	0.03	0.01	0.14	0.08
1000	2	0.01	0.01	0.07	0.13	0.01	-0.01	0.13	0.17	0.00	0.03	0.09	0.13
1500	2	0.00	0.00	0.07	0.09	-0.01	-0.01	0.10	0.10	0.01	0.02	0.07	0.08
3000	2	0.00	0.01	0.06	0.07	0.00	0.01	0.10	0.07	0.01	0.00	0.06	0.06
5000	2	0.00	-0.01	0.05	0.07	0.00	0.00	0.08	0.08	0.00	0.00	0.05	0.07
100	5	0.05	0.31	0.56	1.15	0.08	0.28	0.30	1.13	0.14	0.23	1.96	0.74
500	5	0.00	0.03	0.09	0.30	0.00	-0.03	0.23	0.35	0.05	0.13	0.26	0.36
1000	5	0.01	-0.01	0.08	0.20	0.00	-0.05	0.18	0.25	0.03	0.07	0.16	0.23
1500	5	0.01	0.00	0.07	0.16	0.01	0.00	0.15	0.23	0.01	0.04	0.10	0.20
3000	5	0.00	0.00	0.07	0.12	0.00	-0.01	0.12	0.16	0.01	0.03	0.09	0.12
5000	5	0.00	0.02	0.07	0.09	-0.01	0.00	0.10	0.12	0.01	0.02	0.06	0.08

N = Population size; β = Transmission rate; γ = Recovery rate; $\mathcal{R}_0$ = Basic reproductive number; ASB = Average scaled bias; RASSE= Root average squared scaled error; M_1 = Nonlinear differential equation fitting; M_2 = Simple regression fitting.

Table 3: Case C: Performance results for SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) with observation error through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of k plus model error in the transmission rate β and the recovery rate γ such that a different value of β and γ for each epidemic day is generated via an autoregressive time series, with $\beta_t = a\beta_{t-1} + \epsilon$ (t in days), and similarly for γ.

N	k	β				γ				$\mathcal{R}_0$			
		ASB		RASSE		ASB		RASSE		ASB		RASSE	
		M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
100	2	-0.02	0.14	0.23	1.18	0.07	0.22	0.37	2.27	0.03	0.28	0.65	0.67
500	2	-0.05	0.03	0.26	0.21	0.00	0.01	0.41	0.44	0.05	0.16	0.39	0.48
1000	2	-0.02	0.00	0.25	0.16	-0.01	-0.05	0.40	0.37	0.08	0.23	0.38	0.60
1500	2	0.00	0.00	0.26	0.15	0.00	-0.03	0.45	0.39	0.10	0.22	0.37	0.53
3000	2	0.01	0.02	0.28	0.13	0.01	-0.04	0.53	0.34	0.14	0.28	0.43	0.80
5000	2	-0.03	0.01	0.26	0.12	-0.06	-0.07	0.43	0.32	0.13	0.27	0.41	0.64
100	5	-0.05	0.45	0.27	2.64	0.06	0.62	0.38	3.63	-0.03	0.23	0.37	0.76
500	5	-0.02	0.11	0.26	0.89	0.04	0.15	0.48	1.71	0.05	0.26	0.35	0.68
1000	5	-0.02	0.06	0.24	1.31	-0.04	0.08	0.38	2.75	0.13	0.34	0.45	0.78
1500	5	-0.03	-0.01	0.26	0.22	-0.01	-0.07	0.37	0.40	0.08	0.24	0.38	0.59
3000	5	-0.01	0.01	0.29	0.18	-0.02	-0.04	0.53	0.41	0.18	0.28	0.50	0.73
5000	5	-0.05	0.01	0.24	0.17	-0.07	-0.05	0.40	0.39	0.13	0.25	0.39	0.62

N = Population size; β = Transmission rate; γ = Recovery rate; $\mathcal{R}_0$ = Basic reproductive number; ASB = Average scaled bias; RASSE= Root average squared scaled error; M_1 = Nonlinear differential equation fitting; M_2 = Simple regression fitting.

Table 4: Case D: Performance results for SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) with observation error through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of k plus model error such that β varies deterministically, equal to a constant until day 15, then exponentially decaying to a new value, so $\beta = \beta_1$ for $t \leq 15$ and $\beta = \beta_2 + (\beta_1 - \beta_2)\exp(-0.2 \times (t - 15))$ for $t > 15$.

N	k	β				γ				$\mathcal{R}_0$			
		ASB		RASSE		ASB		RASSE		ASB		RASSE	
		M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
100	2	0.83	0.80	0.85	1.04	0.12	0.06	0.26	0.49	0.71	0.07	0.86	0.34
500	2	1.04	0.45	1.04	0.54	0.49	-0.02	0.51	0.24	0.38	-0.12	0.39	0.17
1000	2	1.21	0.31	1.21	0.37	0.73	-0.02	0.74	0.18	0.28	-0.22	0.29	0.23
1500	2	1.34	0.22	1.35	0.29	0.91	-0.01	0.92	0.18	0.23	-0.28	0.23	0.28
3000	2	1.63	0.11	1.63	0.18	1.32	-0.01	1.32	0.15	0.14	-0.35	0.14	0.35
5000	2	1.90	0.05	1.91	0.14	1.69	-0.01	1.69	0.14	0.08	-0.38	0.08	0.38
100	5	0.86	1.08	0.91	1.77	0.20	0.24	0.36	0.98	0.66	0.17	0.88	0.80
500	5	1.03	0.48	1.04	0.66	0.48	0.01	0.53	0.39	0.40	-0.11	0.43	0.22
1000	5	1.23	0.30	1.24	0.46	0.76	-0.03	0.78	0.32	0.28	-0.20	0.29	0.23
1500	5	1.36	0.24	1.36	0.41	0.93	-0.01	0.94	0.32	0.23	-0.26	0.23	0.27
3000	5	1.63	0.13	1.64	0.28	1.32	0.01	1.33	0.26	0.14	-0.34	0.14	0.34
5000	5	1.91	0.02	1.91	0.19	1.70	-0.03	1.70	0.21	0.08	-0.38	0.08	0.38

N = Population size; β = Transmission rate; γ = Recovery rate; $\mathcal{R}_0$ = Basic reproductive number; ASB = Average scaled bias; RASSE= Root average squared scaled error; M_1 = Nonlinear differential equation fitting; M_2 = Simple regression fitting.

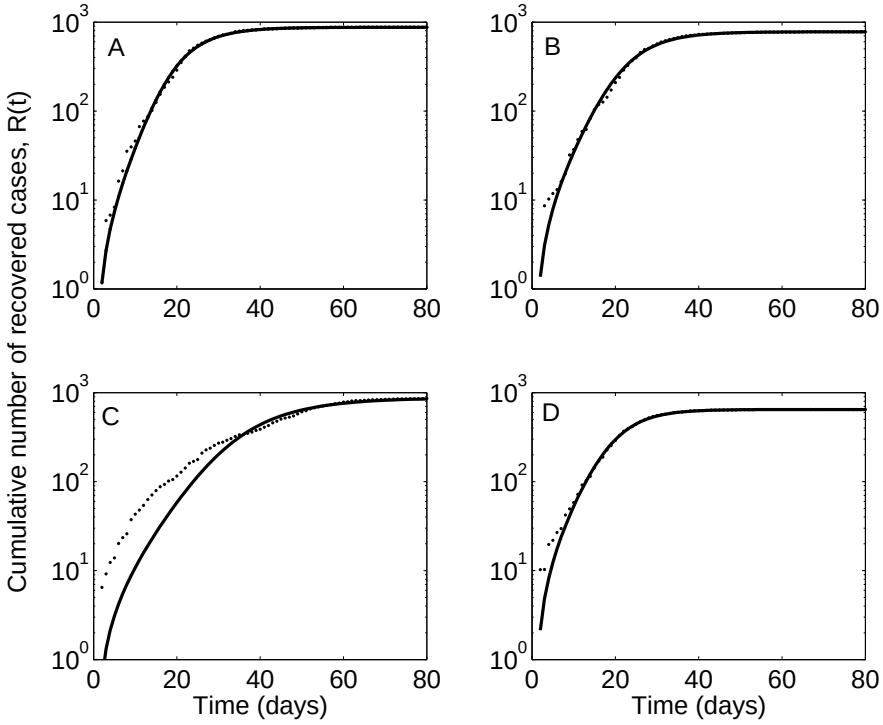

Figure 1: Representative simulated epidemic curves (dots) and their corresponding fit (solid line) using the standard SIR model with nonlinear differential equation fitting to the cumulative number of recovered cases $R(t)$ in log scale with $N = 1000$, $\beta = 1/2$, $\gamma = 1/4$, and $I(0) = 5$. A) SIR epidemic curve with observation error generated using a Gamma error structure with variance $\sigma^2 = 2\mu$; B) Case A) plus model error in the transmission rate β such that a different value of β for each epidemic day is independently and identically distributed according to $\sim$ Uniform$(\beta_{avg} - \Delta_\beta, \beta_{avg} + \Delta_\beta)$ with $\Delta_\beta = 0.15$; C) Case A) plus model error such that a different value of β and γ for each epidemic day is generated via an autoregressive time series, with $\beta_t = a\beta_{t-1} + \epsilon$ (t in days), and similarly for γ; D) Case A plus model error such that β varies deterministically, equal to a constant until day 15, then exponentially decaying to a new value, so $\beta = \beta_1$ for $t \leq 15$ and $\beta = \beta_2 + (\beta_1 - \beta_2)\exp(-0.2 \times (t - 15))$ for $t > 15$.

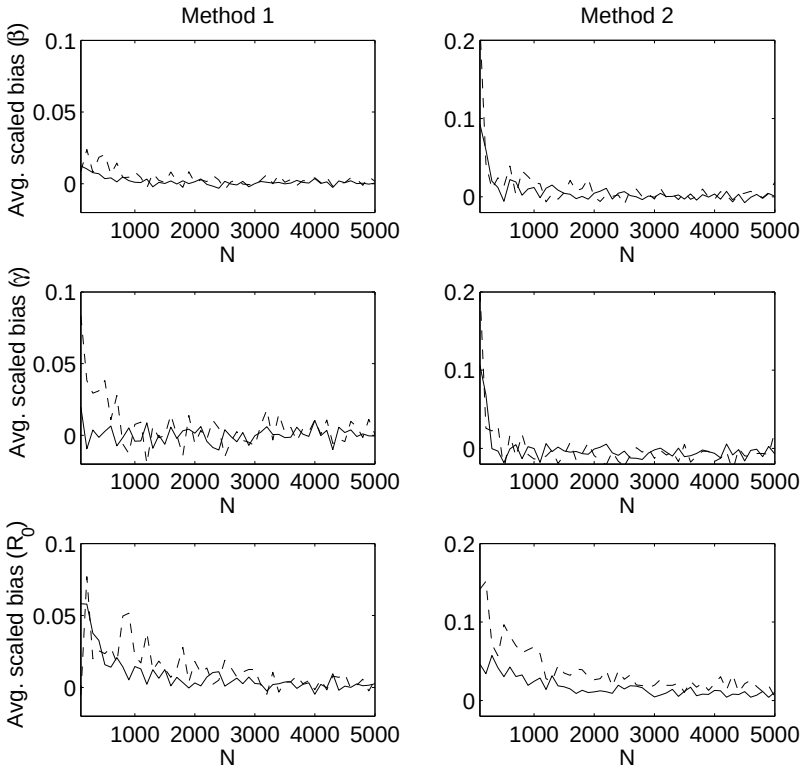

Figure 2: The average scaled bias for Case A for estimates of β, γ, and $\mathcal{R}_0 = \beta/\gamma$ from SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) using two different estimation methods with observation error only through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of $k = 2$ (solid) and 5 (dashed) as a function of the population size N.

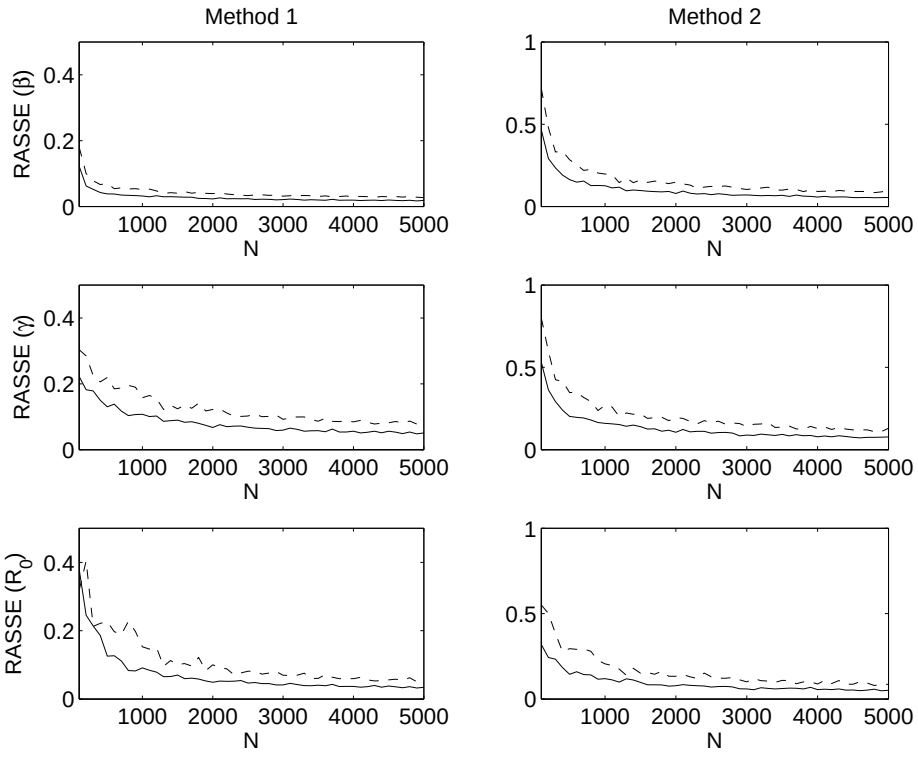

Figure 3: The root average squared scaled error (RASSE) for case A for estimates of β, γ, and $\mathcal{R}_0 = \beta/\gamma$ from SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) using two different estimation methods with observation error only through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of $k = 2$ (solid) and 5 (dashed) as a function of the population size N.

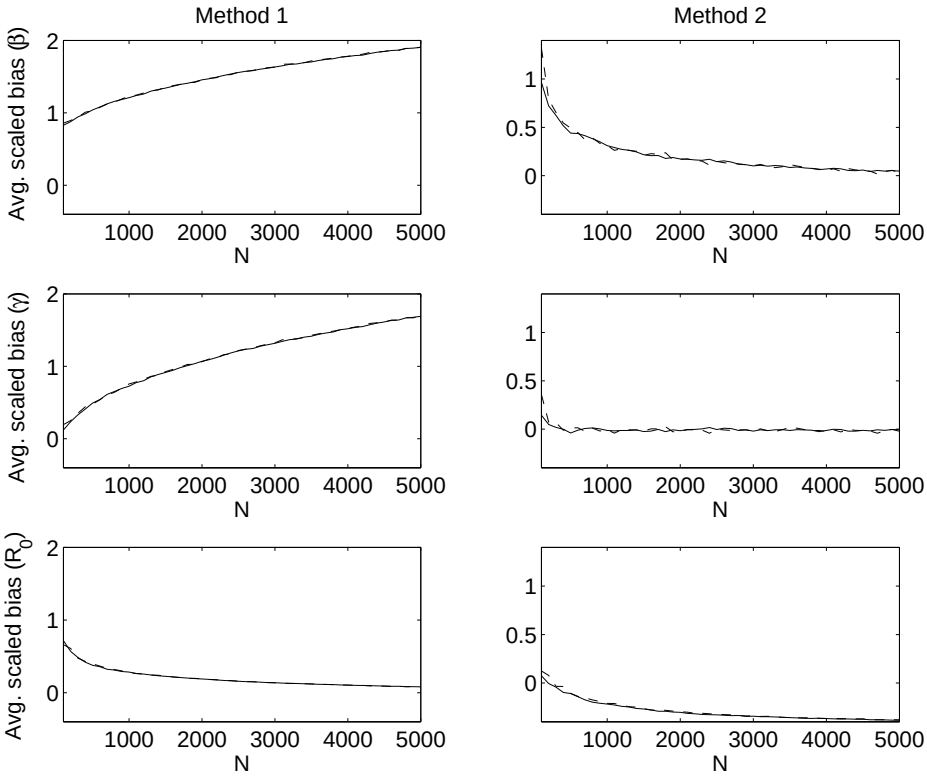

Figure 4: The average scaled bias for case D for estimates of β, γ, and $\mathcal{R}_0 = \beta/\gamma$ as a function of the population size N from SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) using two different estimation methods with observation error through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of $k = 2$ (solid) and 5(dashed) plus model error such that β varies deterministically, equal to a constant until day 15, then exponentially decaying to a new value, so $\beta = \beta_1$ for $t \leq 15$ and $\beta = \beta_2 + (\beta_1 - \beta_2)\exp(-0.2 \times (t - 15))$ for $t > 15$.

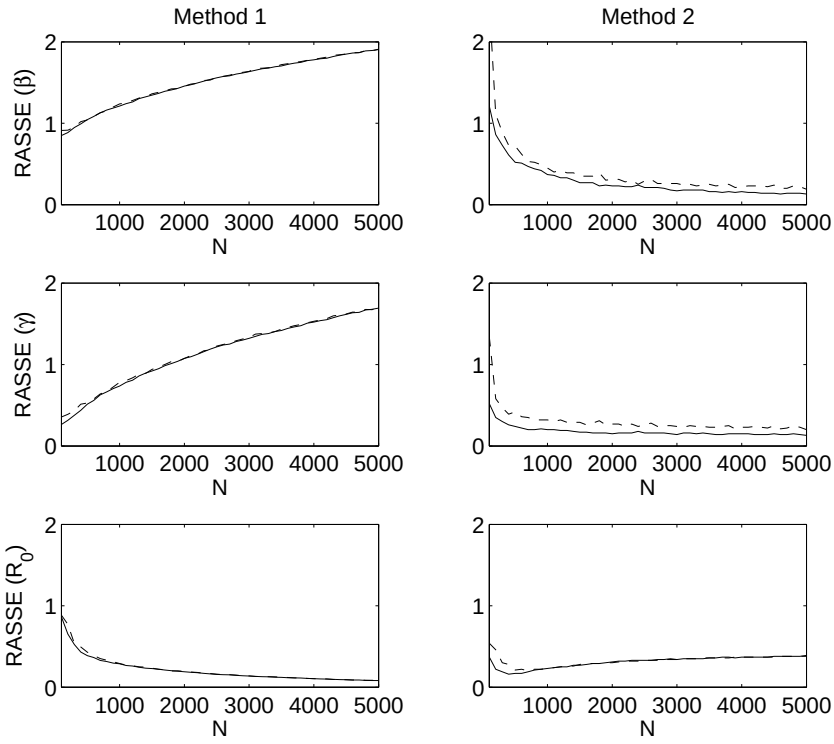

Figure 5: The root average squared scaled error (RASSE) for case D for estimates of β, γ, and $\mathcal{R}_0 = \beta/\gamma$ as a function of the population size N from SIR epidemics ($\beta = 1/2$, $\gamma = 1/4$, $N = 1000$, and $I(0) = 5$) using two different estimation methods with observation error through a Gamma error structure with variance $\sigma^2 = k\mu$ for two values of $k = 2$ (solid) and 5 (dashed) plus model error such that β varies deterministically, equal to a constant until day 15, then exponentially decaying to a new value, so $\beta = \beta_1$ for $t \leq 15$ and $\beta = \beta_2 + (\beta_1 - \beta_2)\exp(-0.2 \times (t - 15))$ for $t > 15$.