Biometrics WILEY

A marginal moment matching approach for fitting endemic-epidemic models to underreported disease surveillance counts

Johannes Bracher ⬤ | Leonhard Held

Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland

Correspondence Johannes Bracher, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland.
Email: johannes.bracher@uzh.ch

Abstract

Count data are often subject to underreporting, especially in infectious disease surveillance. We propose an approximate maximum likelihood method to fit count time series models from the endemic-epidemic class to underreported data. The approach is based on marginal moment matching where underreported processes are approximated through completely observed processes from the same class. Moreover, the form of the bias when underreporting is ignored or taken into account via multiplication factors is analyzed. Notably, we show that this leads to a downward bias in model-based estimates of the effective reproductive number. A marginal moment matching approach can also be used to account for reporting intervals which are longer than the mean serial interval of a disease. The good performance of the proposed methodology is demonstrated in simulation studies. An extension to time-varying parameters and reporting probabilities is discussed and applied in a case study on weekly rotavirus gastroenteritis counts in Berlin, Germany.

KEYWORDS
count time series analysis, effective reproductive number, infectious diseases, moment matching, temporal aggregation, underreporting

1 | INTRODUCTION

Data from routine surveillance of infectious diseases rarely reflect the entire burden in a population (Gibbons *et al.*, 2014; Noufaily, 2019). At the community level, not all infected individuals seek medical care, for example, due to absence or mildness of symptoms, which is referred to as underascertainment. At the health care level, part of those seeking health care are not diagnosed correctly, lack lab confirmation, or are not entered in the notification system. This is often referred to as underreporting, but we will use the term in a broader sense so that it also subsumes underascertainment. Various study types exist to estimate the degree of underreporting, including community-based survey studies (Woudenberg *et al.*, 2019), returning traveler studies, capture-recapture methods, and serological surveys (Gibbons *et al.*, 2014). For common and permanently immunizing childhood diseases, birth numbers can be used to infer reporting probabilities (Becker and Grenfell, 2017).

Underreporting has received a fair amount of attention in infectious disease modeling and statistics. Approaches of varying complexity have been suggested, starting from simple multiplication factor methods (eg, Jandarov *et al.* 2014, Stocks *et al.* 2018). More sophisticated techniques have been described for branching processes (Fraser *et al.*, 2009; White and Pagano, 2010; Azmon *et al.*, 2014) and compartmental models (Cauchemez and Ferguson, 2008;

Dorigatti *et al.*, 2012; Gamado *et al.*, 2014; Fintzi *et al.*, 2017; Magal and Webb, 2018). Moreover, disease mapping approaches for underreported spatial (Bailey *et al.*, 2005; Oliveira *et al.*, 2017) and spatio-temporal (Stoner *et al.*, 2019; Sharmin *et al.*, 2018b) count data have been suggested. Underreported count outcomes in regression have been considered by Dvorzak and Wagner (2016). In statistical count time series modeling, however, underreporting is rarely addressed. Exceptions are Fernández-Fontelo *et al.* (2016, 2019) who suggested a hidden integer valued autoregressive (INAR) model and Sharmin *et al.* (2018a) who use a hidden log-linear autoregressive count model. Here, we examine the impact of underreporting in the endemic-epidemic (EE) modeling framework (Held *et al.*, 2005; Held and Paul, 2012; Meyer and Held, 2014, 2017), which has links to integer-valued generalized autoregressive conditional heteroscedasticity (INGARCH) models (Ferland *et al.* 2006, Zhu 2011). This model class is of particular interest in epidemiological applications as it can be derived from a mechanistic model of disease transmission (Wakefield *et al.*, 2019) and serve to estimate effective reproductive numbers. The EE class has been used for this purpose by Wang *et al.* (2011) and Bauer and Wakefield (2018), but without accounting for underreporting of surveillance data.

Our main contribution consists in an approximate maximum likelihood method to fit EE models to underreported data. It is based on a marginal moment matching approach where underreported processes are approximated by completely observed processes from the same class. The reporting probability needs to be specified based on external knowledge, as it is generally not identifiable from time series data alone. The sensitivity of the estimates to the chosen value can subsequently be examined. The arguments underlying the likelihood approximation also have implications on the biases occurring when underreporting is ignored or taken into account using multiplication factors. Avoiding such biases is especially relevant when EE models are not just used for prediction, but to estimate effective reproductive numbers.

A marginal moment matching approach can also be used to fit models that are defined at a finer time scale than the observed data. Even if data are only available for weekly reporting intervals, one can then fit models defined at a half-weekly time scale, which is often desirable from an epidemiological perspective. An overview of similar temporal aggregation problems in classical time series analysis is given by Silvestrini and Veredas (2008) while Brännäs *et al.* (2002) address an application to count time series.

The article is structured as follows. In Section 2, we describe an EE model, extend it to account for underreporting, and derive the marginal moments of the underreported process. We also consider the impact of temporal aggregation of observations on the marginal moments. These steps form the basis of the approximate inference scheme introduced in Section 3 and assessed in simulation studies in Section 4. In Section 5, we extend the approach to models with time-varying parameters and apply it in a case study on rotavirus gastroenteritis in Berlin, Germany. Section 6 concludes with a discussion.

2 | METHODOLOGY

2.1 | Model definition

The EE modeling framework for time series of infectious disease counts was introduced by Held *et al.* (2005) and extended in a series of articles (Paul *et al.* 2008; Held and Paul 2012; Meyer and Held 2014, 2017). While much of the existing work treats multivariate models, we here focus on the univariate version. We start by assuming all model parameters to be time constant but consider time-varying parameters in Section 5.1 and Web Appendix A.

Conditional on the past, the number X_t of cases of a given disease in time period $t = 1, 2, ...$ is then assumed to follow a negative binomial distribution

$$X_t \mid \mathbf{X}_{<t} \sim \text{NegBin}(\lambda_t, \psi) \qquad (1)$$

with conditional mean λ_t and overdispersion parameter $\psi > 0$ so that $\text{Var}(X_t \mid \mathbf{X}_{<t}) = \lambda_t + \psi\lambda_t^2$. Here, we use $\mathbf{X}_{<t}$ to denote the history $X_{t-1}, ..., X_1$. The negative binomial distribution allows the model to adapt to the typically high dispersion in surveillance data. In the simplest case the conditional expectation λ_t is additively decomposed into

$$\lambda_t = \nu + \phi X_{t-1}. \qquad (2)$$

Here, the term $\nu > 0$ is interpreted as the *endemic component* of incidence, that is, the part that cannot be explained by previous cases. This can, for example, represent infection from sources outside the study region or indirect transmission from environmental sources. The autoregressive term ϕX_{t-1} with $\phi \geq 0$, also called the *epidemic component*, captures the dynamics of the process, see the next section for a more detailed discussion.

In Bracher and Held (2020), we recently relaxed the AR(1) assumption from (2) by regressing on a weighted sum of past observations. The weights represent the serial interval distribution, that is, the distribution of time between onset of clinical symptoms in a case and a secondary case infected by the first (Becker, 2015). A related approach is to include an autoregression on λ_{t-1}, as is done in INGARCH models (Ferland *et al.*, 2006;

Fokianos, 2016):

$$X_t \mid \mathbf{X}_{<t}, \lambda_1 \sim \text{NegBin}(\lambda_t, \psi) \tag{3}$$

$$\lambda_t = \nu + \phi X_{t-1} + \kappa \lambda_{t-1}, \tag{4}$$

where $\kappa \geq 0$. The initial value λ_1 is assumed to be random (a common assumption is that it follows the respective stationary distribution). This corresponds to a negative binomial INGARCH(1, 1) model (Zhu, 2011). We will work with this extended definition, which will turn out to be particularly useful in the context of underreporting.

In the following, we consider the case where the process $\{X_t\}$ is not fully observable. Instead, we observe an underreported process $\{\tilde{X}_t\}$, which depends on $\{X_t\}$ through independent binomial thinning,

$$\tilde{X}_t \mid X_t \overset{\text{ind}}{\sim} \text{Bin}(X_t, \pi), \tag{5}$$

where $0 < \pi \leq 1$ is the reporting probability.

2.2 | Effective reproductive number and serial interval distribution

Bauer and Wakefield (2018) recently derived the AR(1) version (1)–(2) of the EE model from a time-discrete SIR (susceptible-infected-removed) model with a large population and a relatively rare disease. They showed that if the serial interval corresponds to one time step, the parameter ϕ can be interpreted as the local effective reproductive number R_{eff}, that is, the average number of secondary cases an infected individual causes in the considered population, given all partial immunities and intervention measures. This approach had previously been used by Wang *et al.* (2011).

For the extended model (3)–(4), the conditional mean λ_t can be rewritten as (Fokianos 2016, p. 7)

$$\lambda_t = \sum_{d=0}^{t-2} \kappa^d \nu + \frac{\phi}{1-\kappa} \cdot \sum_{d=1}^{t-1} (1-\kappa)\kappa^{d-1} X_{t-d} + \kappa^{t-1}\lambda_1, \tag{6}$$

which has the form of an *epidemic renewal equation* (Fraser *et al.*, 2009, S17). The effective reproductive number is thus

$$R_{\text{eff}} = \frac{\phi}{1-\kappa},$$

while the serial interval distribution is assumed to be geometric with parameter $(1-\kappa)$. This corresponds to a mean serial interval of $1/(1-\kappa)$ at the time scale of the observation interval. While there is often no specific biological mechanism that would justify a geometric serial interval distribution, it represents as a decrease of infectiousness over time.

2.3 | Marginal moments of the underreported process

As shown by Zhu (2011), $\{X_t\}$ is second-order stationary if $(\phi + \kappa)^2 + \phi^2\psi < 1$, with marginal mean, variance, and autocorrelation function

$$\mu = \frac{\nu}{1-\xi}, \quad \sigma^2 = \frac{1-\xi^2+\phi^2}{1-\xi^2-\psi\phi^2} \cdot (\mu + \psi\mu^2),$$

$$\text{and} \quad \rho(d) = \eta\xi^{d-1}, \quad d = 1, 2, \ldots, \tag{7}$$

respectively. Here, we denote by

$$\eta = \phi \cdot \frac{1-\kappa\xi}{1-\xi^2+\phi^2} \quad \text{and} \quad \xi = \phi + \kappa \tag{8}$$

the first-order autocorrelation and the decay factor of the autocorrelation function, respectively.

The corresponding quantities for the underreported process $\{\tilde{X}_t\}$ are

$$\tilde{\mu} = \pi\mu, \quad \tilde{\sigma}^2 = \pi^2\sigma^2 + \pi(1-\pi)\mu,$$

$$\text{and} \quad \tilde{\rho}(d) = \tilde{\tau}\eta\xi^{d-1}, \quad d = 1, 2, \ldots, \tag{9}$$

where the *attenuation factor* $\tilde{\tau} = \tilde{\rho}(d)/\rho(d)$ can be written as

$$\tilde{\tau} = 1 - (1-\pi) \cdot \frac{\tilde{\mu}}{\tilde{\sigma}^2},$$

see Web Appendix B.1 for the derivation of (9). The attenuation of the autocorrelations gets stronger, that is, $\tilde{\tau}$ gets smaller, as π and the index of dispersion $\tilde{\sigma}^2/\tilde{\mu}$ decrease.

2.4 | Second-order equivalent models

An interesting implication of Equations (7)–(9) is that different combinations of reporting probabilities and parameters for the latent process can lead to the same marginal second-order properties of the underreported process. In fact, for any underreported process $\{\tilde{X}_t\}$ with underlying parameters ν, ϕ, κ, ψ, and reporting probability π, and any $\pi_Y \in (\pi, 1]$, one can find a process

$$Y_t \mid \mathbf{Y}_{<t}, \lambda_{Y,1} \sim \text{NegBin}(\lambda_{Y,t}, \psi_Y)$$

$$\lambda_{Y,t} = \nu_Y + \phi_Y Y_{t-1} + \kappa_Y \lambda_{Y,t-1} \tag{10}$$

so that the underreported process

$$\tilde{Y}_t \mid Y_t \overset{\text{ind}}{\sim} \text{Bin}(Y_t, \pi_Y)$$

has the same marginal second-order properties as $\{\tilde{X}_t\}$:

$$\tilde{\mu}_Y = \tilde{\mu}; \quad \tilde{\sigma}_Y^2 = \tilde{\sigma}^2; \quad \tilde{\rho}_Y(d) = \tilde{\rho}(d) \text{ for } d = 1, 2, \dots \quad (11)$$

We then call $\{\tilde{X}_t\}$ and $\{\tilde{Y}_t\}$ *second-order equivalent*. While they are not equal in distribution, they behave very similarly, see the simulation study in Web Appendix E.1. An important special case emerges for $\pi_Y = 1$ when $\{\tilde{X}_t\}$ and the fully observed process $\{Y_t\}$ are second-order equivalent.

The equality of marginal means, variances, and autocorrelation functions in (11) implies a system of four equations, which yields explicit formulas for the parameters of the process $\{Y_t\}$:

$$\nu_Y = \frac{\tilde{\mu}(1-\xi)}{\pi_Y} = \frac{\pi}{\pi_Y} \cdot \nu, \quad (12)$$

$$\phi_Y = \frac{\sqrt{\tilde{\tau}_Y^2(1-\xi^2)^2 + 4(\tilde{\tau}_Y\xi - \tilde{\tau}\eta)\tilde{\tau}\eta(1-\xi^2)} - \tilde{\tau}_Y(1-\xi^2)}{2(\tilde{\tau}_Y\xi - \tilde{\tau}\eta)}, \quad (13)$$

$$\kappa_Y = \xi - \phi_Y, \quad (14)$$

and

$$\psi_Y = \frac{(1-\xi^2)\tilde{\tau}_Y\tilde{\sigma}^2 - \pi_Y\tilde{\mu}\{1-\xi^2 + \phi_Y^2\}}{\phi_Y^2\tilde{\tau}_Y\tilde{\sigma}^2 + \tilde{\mu}^2\{1-\xi^2 + \phi_Y^2\}}, \quad (15)$$

where $\tilde{\tau}_Y = 1 - (1-\pi_Y)\tilde{\mu}_Y/\tilde{\sigma}_Y^2 = 1 - (1-\pi_Y)\tilde{\mu}/\tilde{\sigma}^2$. Derivations are given in Web Appendix B.2, where it is also shown that for $\pi_Y \in (\pi, 1]$, the parameters ν_Y, ϕ_Y, κ_Y, and ψ_Y are all nonnegative, that is, $\{Y_t\}$ is a well-defined process. Equations (12)–(15) simplify in the important special case $\pi_Y = 1$, see Equations (W20)–(W23) in Web Appendix B.2. It can also be shown that ν_Y and ϕ_Y are monotonically decreasing in π_Y, while κ_Y and ψ_Y are monotonically increasing (Web Appendix B.3). This has implications on the biases occurring when underreporting is ignored, see Section 3.2. The presented argument on second-order equivalence of processes with different reporting probabilities can even be extended to models with time-varying parameters (Section 5.1 and Web Appendix A).

2.5 | Marginal moments of temporally aggregated processes

An assumption underlying the arguments in Section 2.2 is that the serial interval distribution is well described by a geometric distribution at the time scale of the observation intervals, typically weeks. For diseases with serial intervals of a few days, however, it would be more reasonable to define the model $\{X_t\}$ at a half-weekly time scale, say. Ignoring underreporting for the moment, we can then only observe the *temporally aggregated* process $\{X_t^*\}$ where

$$X_t^* = X_{2t-1} + X_{2t}. \quad (16)$$

The mean, variance, and autocorrelation function of $\{X_t^*\}$ can be shown to be

$$\mu^* = 2\mu, \quad \sigma^{*2} = 2(1+\eta)\sigma^2,$$

$$\rho^*(d) = \eta \cdot \frac{(1+\xi)^2}{2(1+\eta)} \cdot \xi^{2(d-1)}, \quad (17)$$

where η and ξ are defined in (8). The autocorrelation function of $\{X_t^*\}$ is thus of the same general form as that of $\{X_t\}$ given in (7). Similar results for various classical time series models can be found in Silvestrini and Veredas (2008). Furthermore, if $\kappa \geq 2 - \sqrt{3} \approx 0.27$, which for half-weekly time steps corresponds to a serial interval of at least 4.8 days, one can find a process $\{Y_t\}$ from our class that is defined directly at the time scale of $\{X_t^*\}$ and shares its marginal second-order properties. Details and expressions for the parameters of $\{Y_t\}$ can be found in Web Appendix C.2.

The same second-order equivalence argument can be made for processes $\{\tilde{X}_t^*\}$, which are both underreported and temporally aggregated. Interestingly, a well-defined second-order equivalent process $\{Y_t\}$, which is fully observed and defined at the time scale of $\{\tilde{X}_t^*\}$ can then even be found in many cases where $\kappa < 0.27$. This is a consequence of the correlation attenuation caused by underreporting. As notation gets somewhat involved, details on processes with both underreporting and temporal aggregation have been moved to Web Appendix C.3.

3 | INFERENCE

3.1 | Need for external information on the reporting probability

An implication of the results from Section 2.4 is that it is generally not feasible to estimate the reporting probability from time series data alone. The reason is that practically identical fits can be achieved with quasi-equivalent models featuring different reporting probabilities, see also the simulation study in Section 4.1. In order to estimate π, one would need to fix one of the other parameters. Similar identifiability issues have been reported for count data

regression models (Dvorzak and Wagner, 2016) and latent INAR time series models (Bracher, 2019). For compartmental models, simulation studies have led to related conclusions. For example, Fraser *et al.* (2009, S18) state that the reporting probability and epidemic parameters cannot be estimated jointly.

3.2 | Naïve approaches and their biases

In practical analyses of surveillance data, underreporting is often ignored, but within the EE class, this will lead to biased parameter estimates. Assume that the data come from a process $\{\tilde{X}_t\}$ with reporting probability π and underlying parameters ν, ϕ, κ, ψ. If we wrongly assume full reporting, the parameter estimates can be expected to be around the parameters of a process $\{Y_t\}$, which is second-order equivalent to $\{\tilde{X}_t\}$, but has reporting probability $\pi_Y = 1$, see equations (W20)– (W23) in Web Appendix B.2. The derivatives $d\nu_Y/d\pi_Y$ and so forth given in Web Appendix B.3 imply that this leads to overestimation of κ and ψ and underestimation of ν, ϕ, and R_{eff}. Note that an upward bias in κ means we overestimate the average serial interval $1/(1 - \kappa)$.

A simple approach to account for underreporting is to inflate counts with a multiplication factor, given by the inverse of the assumed reporting probability, prior to analysis (eg, Jandarov *et al.* 2014, Stocks *et al.* 2018). This changes their overall level and variance, but leaves the autocorrelations unaffected. As ϕ and κ reflect the autocorrelation structure, we do not expect this approach to alleviate bias in these parameters. We can use (12)–(15) to find the parameters $\nu_Y, \phi_Y, \kappa_Y, \psi_Y$ of a fully observed process $\{Y_t\}$, which is second-order equivalent to the process $\{X_t/\pi\}$. This tells us which biases to expect, see Section 4.2 for a simulation study exploring this in more detail.

3.3 | Maximum likelihood estimation via marginal moment matching

The challenge in fitting the hierarchical model (3)–(5) to data is that the likelihood cannot be evaluated directly. There is no obvious way of recursively computing the likelihood contributions $\Pr(\tilde{X}_t = \tilde{x}_t \mid \tilde{X}_{t-1} = \tilde{x}_{t-1}, \dots \tilde{X}_1 = \tilde{x}_1)$ as for fully observed models. Methods for hidden Markov models cannot be applied either as the latent process $\{\tilde{X}_t\}$ lacks the Markov property, see Equation (6). Equations (12)–(15), however, offer two potential solutions. Both are based on marginal moment matching where underreported processes are approximated by second-order equivalent processes without underreporting, that is, a reporting probability of one. This is inspired by a method

for fitting Gaussian AR(p) models to mismeasured data by expressing them as ARMA(p, p) models (Staudenmayer and Buonaccorsi, 2005).

A simple strategy is to fit a model with $\pi = 1$, that is, ignoring underreporting, and subsequently debias the parameter estimates. The corrected estimates are given by the parameters $\nu_Y, \phi_Y, \kappa_Y, \psi_Y$ underlying a process $\{\tilde{Y}_t\}$, which is second-order equivalent to the fitted model $\{X_t\}$, but underreported with the assumed reporting probability π_Y. These can be obtained using Equations (12)–(15). However, as $\pi_Y > \pi = 1$ does not hold here, this does not ensure that the corrected estimates for κ and ψ are nonnegative. Especially if the true κ is small, this poses a problem, see the simulation study in Web Appendix E.7.

We therefore adopt a more elaborate approach where underreporting is accounted for during rather than after the likelihood optimization. In order to approximate the likelihood of a model $\{\tilde{X}_t\}$ with underlying parameters ν, ϕ, κ, ψ, and reporting probability π, we consider the second-order equivalent model $\{Y_t\}$ without underreporting ($\pi_Y = 1$) and parameters $\nu_Y, \phi_Y, \kappa_Y, \psi_Y$ as in (12)–(15). Such a process always exists and is well defined as $\pi_Y = 1 \geq \pi$. The likelihood of $\{\tilde{X}_t\}$ can then be approximated by that of $\{Y_t\}$, which is easier to evaluate. We also estimate the initial value λ_1 (Ferland *et al.*, 2006) and for simplicity set $\lambda_{Y,1} = \pi\lambda_1$. We can then compute

$$\mathcal{L}(\nu, \phi, \kappa, \psi, \lambda_1, \pi) \approx \mathcal{L}(\nu_Y, \phi_Y, \kappa_Y, \psi_Y, \lambda_{Y,1}, \pi_Y = 1)$$

$$= \Pr(Y_1 = \tilde{x}_1 \mid \lambda_{Y,1}, \psi_Y) \cdot \prod_{t=2}^{T} \Pr(Y_t = \tilde{x}_t \mid \nu_Y, \phi_Y,$$

$$\kappa_Y, \psi_Y, \lambda_{Y,1}, Y_{t-1} = \tilde{x}_{t-1}, \dots, Y_1 = \tilde{x}_1), \quad (18)$$

where $\tilde{x}_1, \dots \tilde{x}_T$ are the observed counts. The likelihood contribution of time t is a negative binomial probability mass function with mean $\lambda_{Y,t}$ and overdispersion parameter ψ_Y, see Equation (10). Maximization of the likelihood function for a given π can be done using standard numerical optimization and π can be varied to assess the sensitivity of the estimates to this choice. This procedure can be extended to models with time-varying parameters, see Web Appendix A.

3.4 | Marginal moment matching to account for temporal aggregation

Using the results from Section 2.5, a marginal moment matching approach can also be used to fit EE models to temporally aggregated data, that is, to fit a model defined at a half-weekly time scale to weekly counts. The likelihood of a model $\{X_t\}$ defined in half-weekly steps can be
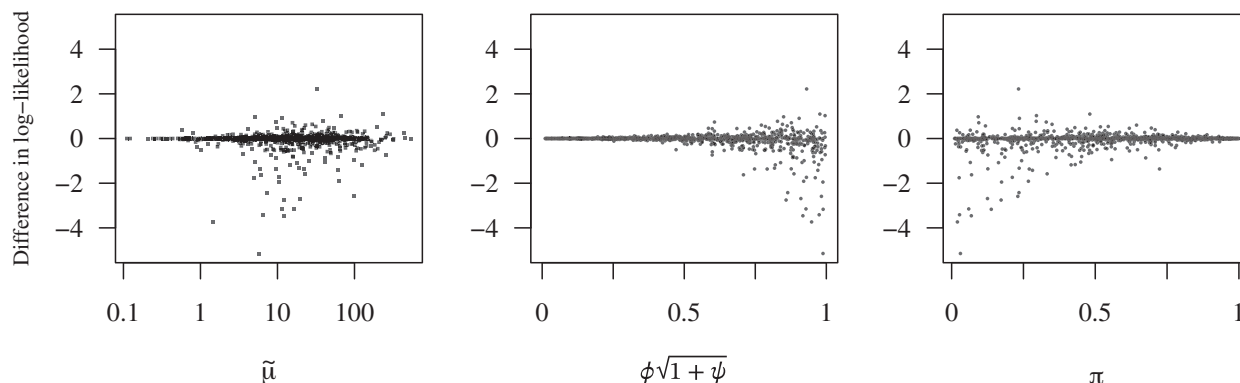
FIGURE 1 Difference between the log-likelihood values returned by the proposed approximation and the forward algorithm, evaluated at the true parameter values. The differences for 1000 randomly generated parameter combinations are plotted against the stationary mean $\tilde{\mu}$ of the underreported process $\{\tilde{X}_t\}$ (on a log scale, left); $\phi\sqrt{1+\psi}$, which describes how close the process is to nonstationarity (threshold at 1, middle); the reporting probability π (right)

approximated by that of a weekly model $\{Y_t\}$, which is second-order equivalent to the temporally aggregated process $\{X_t^* = X_{2t-1} + X_{2t}\}$, see Equation (17). If $\kappa > 0.27$, the process $\{Y_t\}$ is always well defined and the likelihood (18) can be evaluated.

If $\kappa < 2 - \sqrt{3} < 0.27$, the process $\{Y_t\}$ may not be well defined as κ_Y can get negative. However, it can be shown that κ_Y cannot get smaller than $3 - \sqrt{8} \approx -0.17$ (Web Appendix C.2). In many practical settings, such a modest negative value of κ_Y does not hamper evaluation of the likelihood (18). If one of the negative binomial likelihood contributions cannot be evaluated due to a negative conditional expectation $\lambda_{Y,t}$, we pragmatically set $\lambda_{Y,t} = \nu_{Y,t}$ in order to evaluate the likelihood.

The approximation steps to account for underreporting and temporal aggregation are straightforward to combine so that we can also fit models accounting both for underreporting and temporal aggregation. As mentioned before, the condition $\kappa > 0.27$ can even be relaxed if we assume $\pi < 1$. Details have been moved to Web Appendix C.3.

4 | SIMULATION STUDIES

In the simulation studies presented here, we focus on underreported, but not temporally aggregated processes. Simulation studies on accounting simultaneously for both can be found in Web Appendix E.5 and Web Appendix E.6.

4.1 | Comparison to forward algorithm

If $\kappa = 0$, the latent process $\{X_t\}$ is a first-order Markov chain and the exact likelihood of an underreported EE model can be evaluated with the forward algorithm

(Zucchini and MacDonald, 2009). This is now used to assess the quality of proposed approximation method.

We consider the agreement over a broad range of parameter combinations, independently sampling $N = 1000$ sets of parameters from uniform distributions $\nu \sim \text{Unif}(3, 30)$, $\phi \sim \text{Unif}(0.01, 0.99)$, $\psi \sim \text{Unif}(0.001, 0.2)$, $\pi \sim \text{Unif}(0.01, 1)$, $\lambda_1 \sim \text{Unif}(0.5\mu, 2\mu)$, where $\mu = \nu/(1 - \phi)$. As in the estimation method, we treat the initial value λ_1 as an additional unknown parameter. Sets of parameters with $\phi^2(1 + \psi) \geq 1$, that is, which imply nonstationarity, were discarded. For each set of parameters, we generated a time series of length 100 from the respective underreported process $\{\tilde{X}_t\}$ and evaluated the log-likelihood at the true parameter values using the forward algorithm and the proposed approximation method.

Figure 1 shows the differences between the two methods, plotted against characteristics of the generating model. The agreement between both methods is very good, with absolute differences below 0.1 in 73% and below 1 in 97% of the cases. The approximation works somewhat less well as $\phi\sqrt{1+\psi}$ approaches 1, that is, close to nonstationarity (middle panel), and when the reporting probability π is low (right panel). The computation time required for the forward algorithm was two orders of magnitude longer than for the approximation method.

The presented results also support our assertion that the reporting probability cannot be estimated from time series data. After all, the approximating processes with $\pi_Y = 1$ led to equally good fits to the data as the true generating models with reporting probability π.

We have also conducted a simulation study comparing the maximum likelihood estimates obtained via the forward algorithm and the moment matching approach in Web Appendix E.8. Again, agreement between the two is very good.

4.2 | Bias and loss of precision due to underreporting

In a second simulation study, we examine whether the arguments from Section 3.2 describe the biases in naïve analyses well and whether the proposed estimation method can avoid them. We simulated 1000 sample paths of length 416 (ie, 8 years of weekly data, as in the case study in Section 5) from a model with $\nu = 15, \phi = 0.4, \kappa = 0.3$, and $\psi = 0.1$. This implies a marginal mean of $\mu = 50$ and a standard deviation of $\sigma \approx 20$, while the effective reproductive number is $R_{\text{eff}} = 0.57$. We generated underreported versions of these time series with five different reporting probabilities $\pi \in \{0.1, 0.25, 0.5, 0.75, 1\}$, where $\pi = 1$ corresponds to full reporting. Selected sample paths can be found in Web Appendix E.3. The underreported time series are analyzed (a) using the proposed estimation method and the correct π, (b) ignoring underreporting, and (c) inflating the counts with a factor of $1/\pi$ (with subsequent rounding) prior to fitting the model. The results are shown in Figure 2.

The proposed method avoids bias with increasing variability of the estimates as π—and thus the data quality—decreases. Ignoring underreporting leads to pronounced biases for low values of π. These are well described by Equations (12)–(15) with $\pi_Y = 1$, compare Section 3.2 (dashed red lines in the plots; this figure appears in color in the electronic version of this article, and any mention of color refers to that version). The multiplication factor method avoids bias in ν, but not the other parameters. We reran the analysis with time series of length 208 (4 years of weekly data) and 832 (16 years), which gave similar results (Web Appendix E.3).

We also assessed the coverage of Wald confidence intervals for the different parameters in Web Appendix E.4. Empirical coverage was close to the nominal level given sufficient training data. For shorter time series, there was some undercoverage especially when the reporting probability was small.

We moreover performed a simulation study featuring both underreporting and temporal aggregation, using the methodology described in Sections 3.4 and Web Appendix C.3 for fitting. The results, shown in Web Appendix E.5 indicate that the proposed methodology works well also in this case. This was confirmed in a more realistic setting where we used the fitted model from Section 5 as the data generating process, see Web Appendix E.6.

5 | CASE STUDY: ROTAVIRUS GASTROENTERITIS IN BERLIN

Rotavirus causes gastroenteritis with symptoms including diarrhoea, vomiting, and fever and affects mainly young children (Heymann, 2015). It is transmitted predominantly via the faecal oral route. Grimwood *et al.* (1983) report a mean serial interval of 4.9 days for children and 6.4 days for adults. Previous infection has an immunizing effect, but immunity is imperfect and waning over time. Since 2006, a vaccine has been available in Germany and since 2013, it has been recommended by the Standing Committee on Vaccination.

Rotavirus is a notifiable disease in Germany. We analyze weekly counts of laboratory-confirmed cases in Berlin, provided by the Robert Koch Institute via its web platform (2020). The data cover the years 2001–2008 and are displayed in Figure 3. Incidence in this period has previously been assumed to be largely unaffected by vaccination, while from 2009 onward, effects of increased coverage could be observed (Weidemann *et al.* 2014). To avoid this extra level of complexity, we restrict our analysis to the prevaccination period as was done in Stocks *et al.* (2018).

Rotavirus surveillance counts are known to be strongly underreported (Tam *et al.*, 2012). For the period 2001–2008, the reporting probability has been estimated as 4.3% (95% credibility interval: 3.9% to 4.7%) in the Western Federal States of Germany, including Berlin (Weidemann *et al.*, 2014).

5.1 | Model specification

We apply an EE model accounting for both underreporting and temporal aggregation to the rotavirus data. The latter is considered because the average serial interval of rotavirus is reported to be less than 1 week (Grimwood *et al.*, 1983). Rotavirus is often modeled using more complex mechanistic approaches describing the effects of births, partial immunity acquired by infection, and the waning thereof (Atkins *et al.* 2012; Weidemann *et al.* 2014; Stocks *et al.* 2018). Such an approach is beneficial to inform intervention strategies, but a considerable number of parameter values must be prespecified to enable identifiability of the remaining ones (Weidemann *et al.*, 2014). Here we take a different approach, where the above factors are subsumed in the effective (rather than basic) reproductive number and its variation over the course of a season. A similar approach has been taken by van Gaalen *et al.* (2017). The simplicity of the EE model formulation then enables us to explicitly assess the impact of underreporting.

For the rotavirus data, we extend model (3)–(5) by allowing ν_t and ϕ_t to vary over time and adjust for temporal aggregation of half-weekly data, that is,

$$X_t \mid \mathbf{X}_{<t}, \lambda_1 \sim \text{NegBin}(\lambda_t, \psi)$$

$$\lambda_t = \nu_t + \phi_t X_{t-1} + \kappa \lambda_{t-1}, \tag{19}$$

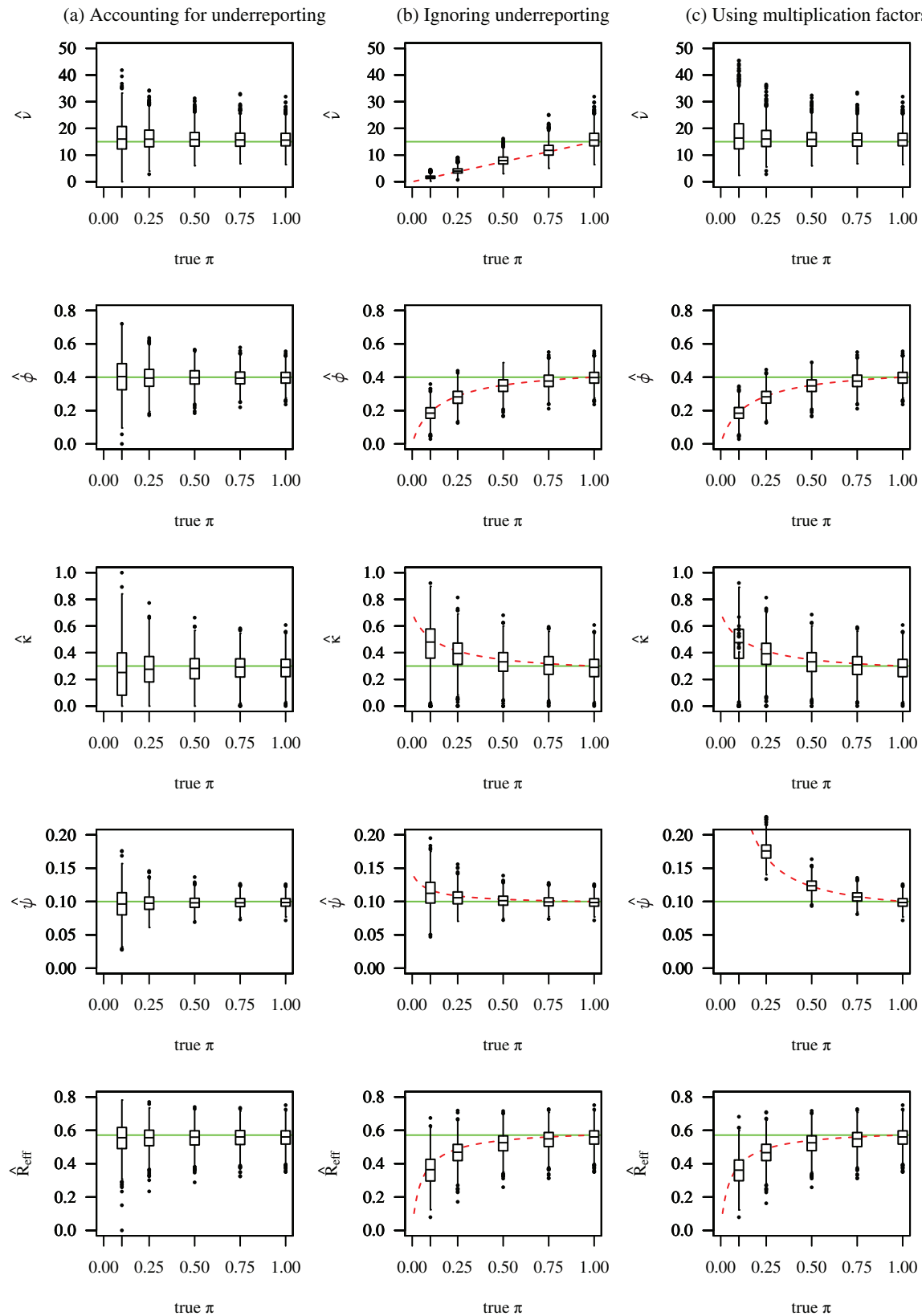**FIGURE 2** Empirical distribution of parameter estimates under different degrees of underreporting. (a) With the proposed method and the correct π. (b) ignoring underreporting, that is, assuming $\pi = 1$; (c) applying the multiplication factor $1/\pi$ prior to analysis. Solid green lines show true parameter values, dashed red lines expected biases. This figure appears in color in the electronic version of this article, and any mention of color refers to that version
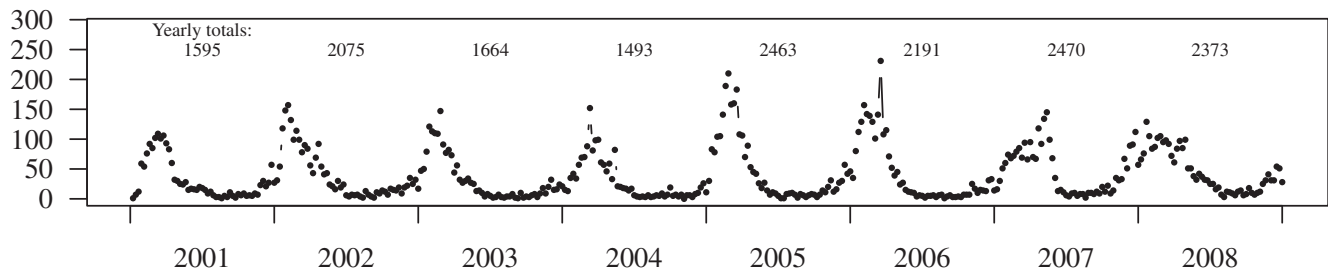
FIGURE 3 Weekly numbers of laboratory-confirmed rotavirus cases in Berlin, 2001–2008

$$\tilde{X}_t \,|\, X_t \overset{\text{ind}}{\sim} \text{Bin}(X_t, \pi)$$

$$\tilde{X}_t^* = \tilde{X}_{2t-1} + \tilde{X}_{2t}. \tag{20}$$

As the latent $\{X_t\}$ is defined for half-weekly time steps, we set (Held and Paul, 2012)

$$\log(\nu_t) = \alpha^{(\nu)} + \gamma^{(\nu)} \sin(2\pi t/104) + \delta^{(\nu)} \cos(2\pi t/104),$$

$$\log(\phi_t) = \alpha^{(\phi)} + \gamma^{(\phi)} \sin(2\pi t/104) + \delta^{(\phi)} \cos(2\pi t/104)$$

to account for seasonal variation with a period length of 104 half-weeks. The parameter κ (and thus the serial interval distribution) is assumed to be time-constant. Exploratory analyses indicated that including more than one pair of sine/cosine waves or linear time trends into $\log(\nu_t)$ or $\log(\phi_t)$ can lead to slight improvements in Akaike's information criterion (AIC), but consistently worse Bayesian information criterion (BIC). As the respective parameter estimates were not significantly different from zero (based on Wald confidence intervals), we did not include these additional terms.

The models are fitted using the approach described in Web Appendix A, which extends that from Section 3.3 to time-varying parameters and temporally aggregated observations. Given the initial value λ_1, marginal means, variances, and autocovariance functions of $\{\tilde{X}_t^*\}$ can be recursively computed even if ν_t and ϕ_t vary over time. The parameters of a second-order equivalent process $\{Y_t\}$ without underreporting, that is, $\pi_Y = 1$, and defined directly at the weekly time scale, can be obtained using a second recursive procedure. As in Equation (18), the likelihood of the process $\{\tilde{X}_t^*\}$ can then be approximated by that of $\{Y_t\}$, and likelihood optimization can be done using numerical optimization.

5.2 | Results

We fitted model (19)–(20) for different reporting probabilities $\pi \in \{0.02, 0.04, \ldots, 0.98, 1\}$. Plots showing fitted and observed values as well as the residual autocorrelation function are provided in Web Appendix F.1. The parameter

estimates obtained under different assumptions on the reporting probability are displayed in Figure 4. The estimated effective reproductive number $R_{\text{eff},t} = \phi_t/(1-\kappa)$ shows considerable seasonal variation, with higher values during winter and early spring. With $\pi = 0.043$ as in Weidemann *et al.* (2014), $R_{\text{eff},t}$ ranges from 0.72 in calendar week 24 to 1.02 in calendar week 50 (highlighted in red). On average we have $R_{\text{eff},t} < 1$, so that the endemic component ν_t, representing infections imported from other regions, is necessary to maintain rotavirus transmission. Comparing the estimates of the endemic terms ν_t and marginal expectations μ_t, we find that the endemic component accounts for roughly one in 10 cases. The estimated $\kappa = 0.41$ implies a mean serial interval of $1/(1-0.41) \times 3.5 \text{ days} = 6.0 \text{ days}$, which is slightly larger than the estimate by Grimwood *et al.* (1983).

Under the unrealistic assumption of no underreporting ($\pi = 1$), we obtain an estimate of $\kappa = 0.54$ and thus a longer average serial interval of 7.5 days. The seasonal range for R_{eff} is now from 0.56 to 1 in this case (darkest purple line in Figure 4), lower than for $\pi = 0.043$. The sensitivity of the estimates to the assumed reporting probability is thus stronger during the off-season.

Weidemann *et al.* (2014) state that the reporting probability may have increased in 2005 due to changed reimbursement schemes, and this is in line with the somewhat increased reported incidence in the seasons 2005–2008. We therefore also fitted a model where $\pi_t = 0.043$ in the years 2001 through 2004, then increases linearly to 0.063 over the course of the year 2005, and stays there until 2008. However, as Figure 5 and Web Table W5 show, the resulting parameter estimates are similar to those obtained with a time-constant $\pi = 0.043$.

To gain an understanding of the uncertainty in the estimates, Figure 5 also shows 90% Wald confidence intervals. As $R_{\text{eff},t}$ is a complicated nonlinear function of the estimated parameters, rather than using the delta method, we computed the corresponding confidence intervals by repeatedly sampling the parameters from a multivariate normal distribution with the maximum likelihood estimates as the mean and the inverse Fisher information as the covariance matrix. This is somewhat heuristic as
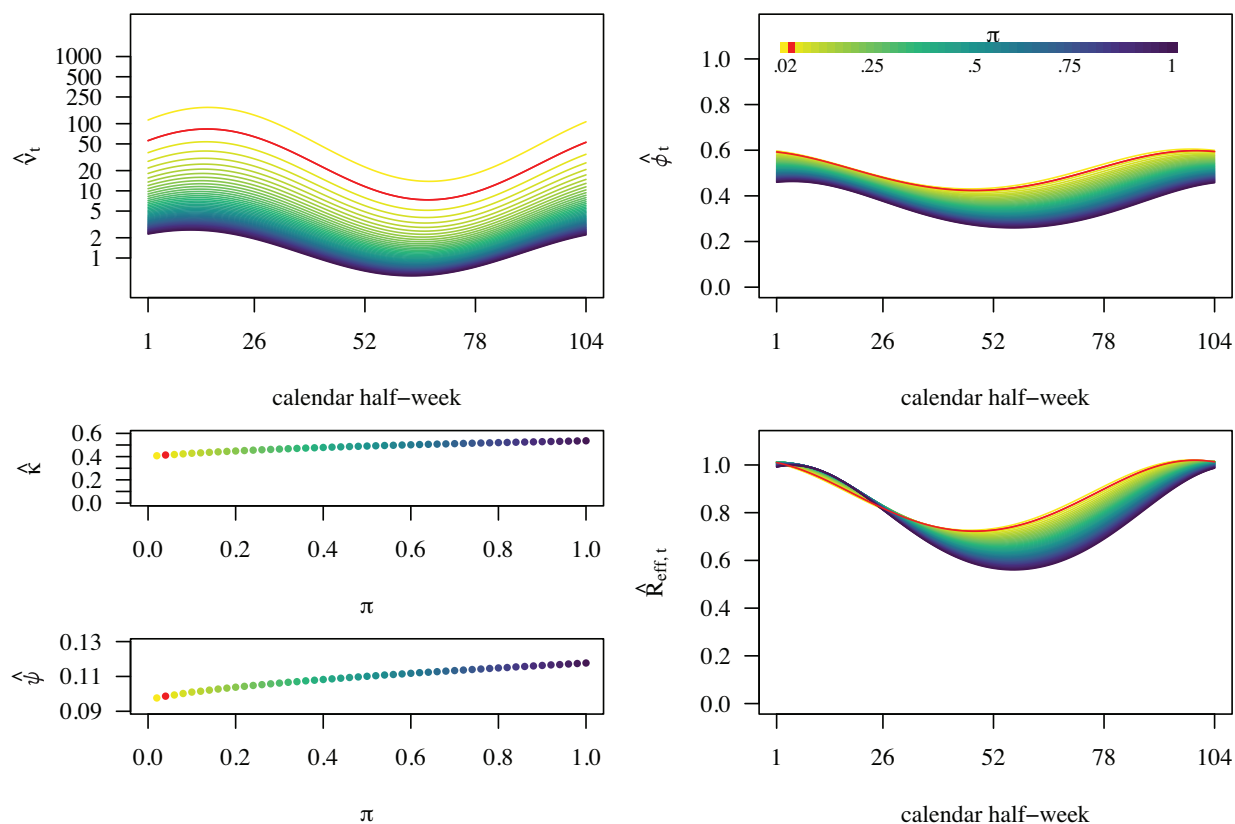
FIGURE 4 Parameter estimates of the underreported EE model applied to weekly rotavirus counts in Berlin, as functions of the assumed reporting probability π. The bottom right panel shows the time-varying local effective reproductive number $R_{\text{eff},t} = \phi_t/(1-\kappa)$. The fit with $\pi = 0.043$ as in Weidemann *et al.* (2014) is highlighted in red. This figure appears in color in the electronic version of this article, and any mention of color refers to that version
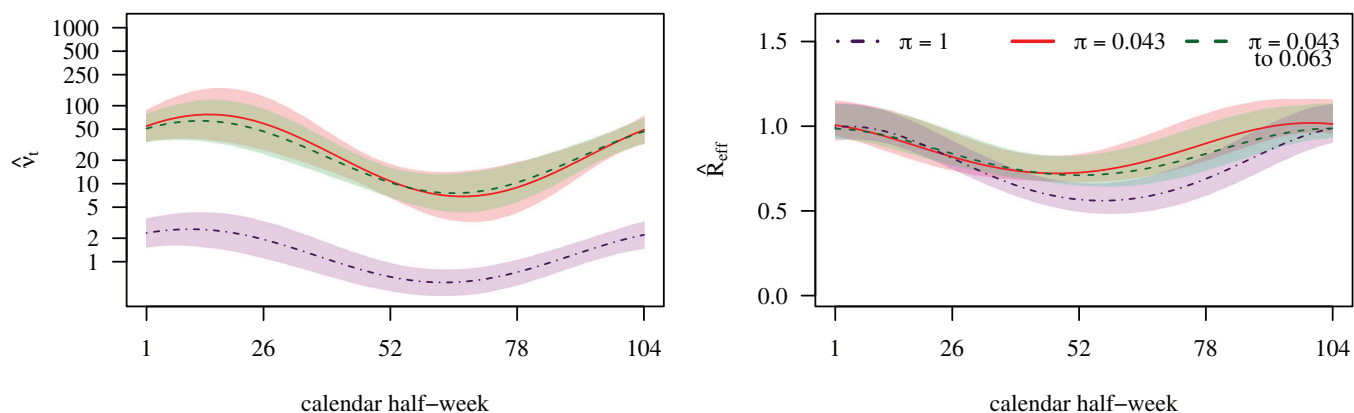

FIGURE 5 Estimated time-varying endemic parameters ν_t and local effective reproductive numbers $R_{\text{eff},t}$ when assuming (a) $\pi = 1$ (purple, dash-dotted line), (b) $\pi = 0.043$ (red, solid line) and (c) time-varying reporting probability increasing from $\pi_t = 0.043$ to $\pi_t = 0.063$ over the course of 2005 (green, dashed line). Point-wise 90% confidence intervals are shown as shaded areas. This figure appears in color in the electronic version of this article, and any mention of color refers to that version

even for completely observed models with time-constant parameters, asymptotic normality of the conditional maximum likelihood estimators is only established if ψ is known (Cui and Zheng, 2017). However, simulation studies show that these intervals have good coverage (Web Appendix E.6).

Lastly, we fitted the same model without the aggregation step (20), that is, defined at a weekly time scale. This

led to somewhat lower values of R_{eff} between 0.65 and 1.00. The results are shown in Supplementary Figure W17. The average serial interval then was 7.9 days (corresponding to $\hat{\kappa} = 0.11$), while the endemic component accounts for roughly one in six cases.

Our estimated effective reproductive numbers are lower than the estimated basic (rather than effective) reproductive numbers from Stocks *et al.* (2018) who obtain a seasonal range of 0.9 to 1.2 for R_0. In van Gaalen *et al.* (2017), a study based on Dutch data, the effective reproductive number shows a similar seasonal pattern as in Figure 4, but is higher (between 0.85 and 1.15, fig. 3 in van Gaalen *et al.* 2017). A possible explanation is that van Gaalen *et al.* 2017 do not allow for any infections from outside of their study region. Note also that both Stocks *et al.* (2018) and van Gaalen *et al.* (2017) account for underreporting via multiplication factors.

6 | DISCUSSION

We extended the EE modeling framework to account for underreporting. In simulation studies, we showed that the proposed likelihood approximation is highly accurate and that, combined with the correct reporting probability, our inference method avoids bias. The case study on rotavirus gastroenteritis showed that estimates of effective reproductive numbers from EE models increase when underreporting is taken into account. Accounting for temporal aggregation increases the estimated reproductive numbers further.

We modeled underreporting through binomial thinning, which is a commonly used, but simple mechanism (Azmon *et al.*, 2014; Fintzi *et al.*, 2017; Stoner *et al.*, 2019). It assumes that reporting of different cases is independent, which may not always be the case. For instance, certain general practitioners or hospitals may systematically fail to report their cases. More sophisticated reporting processes could in principle be added to the model, but its properties would then get more complex.

A number of alternatives to the proposed approximate maximum likelihood method exist. Preliminary analyses with a full Bayesian approach have been hampered by two difficulties. First, mixing of standard Markov chain Monte Carlo (MCMC) samplers was poor. To circumvent this, more sophisticated algorithms seem to be needed. Second, prior choice is challenging in EE models (Bauer and Wakefield, 2018). Advantages, however, are that the uncertainty about the reporting probability π could be included more easily and that the hidden process $\{X_t\}$ could be reconstructed. Alternatively it may be worthwhile to combine the proposed likelihood approximation with a hybrid Bayesian approach, allowing for uncertainty in the reporting probability π but preventing feedback using the cut algorithm (Plummer, 2015). An alternative maximum likelihood estimation approach could be based on particle filtering, as implemented in the R package pomp (King *et al.*, 2016). While this represents a powerful general tool, it is computationally demanding and requires careful choice of starting values to obtain reliable results (Stocks *et al.*, 2018).

We focused on univariate EE models, whereas much of the previous work on the class covered spatio-temporal settings (most recently Meyer and Held 2017, Bracher and Held 2020). It is not entirely straightforward to extend our approach to this case and Bayesian methods as applied in a similar context, for example, by Gómez-Rubio *et al.* (2019) may then be required for inference. However, some of our qualitative findings will still be relevant. They indicate that underreported processes are well approximated by processes with an autoregression on λ_{t-1} or, as in Equation (6), higher order lags. If the goal is prediction (Held *et al.*, 2017) rather than parameter estimation, there should be little benefit in explicitly accounting for underreporting. Also, we observed that the biases were weaker when counts were high. They may thus be negligible if counts are in the hundreds or even thousands (eg, Bauer and Wakefield 2018). An additional difficulty in multivariate settings is to account for differences in reporting probabilities between strata, where reliable information is rarely available.

In our application to rotavirus, we used marginal moment matching to fit a model defined for half-weekly time steps to weekly data. This approach is appropriate for an average serial interval of roughly 5 days as reported by Grimwood *et al.* (1983), and indeed we obtained a quite similar estimate of 6 days. In this context, note that other studies have assumed shorter serial intervals for rotavirus (eg, 2.1 days in van Gaalen *et al.* 2017). Dynamics within a half-week then play a role, and more sophisticated data augmentation methods would likely be required to fit models under this assumption to weekly data.

Our fitted model implies that on average the effective reproductive number within Berlin is below one and infections from external sources, which account for around one in 10 cases, are necessary to maintain transmission. Given that slightly more than half a million persons commute to or from Berlin (Bundesagentur für Arbeit, 2020), a city with 3.6 million inhabitants, this does not seem excessive. However, as rotavirus is predominantly spread between children, this aspect may be less relevant than for other diseases. It is possible that other attenuation phenomena are at work, which cause an underestimation of effective reproductive numbers. For example, reporting delays, where cases may be registered one or several weeks after their actual occurrence (Azmon *et al.* 2014, Noufaily 2019), may contribute to additional attenuation.

ORCID

Johannes Bracher https://orcid.org/0000-0002-3777-1410

REFERENCES

Atkins, K.E., Shim, E., Pitzer, V.E. and Galvani, A.P. (2012) Impact of rotavirus vaccination on epidemiological dynamics in England and Wales. *Vaccine*, 30(3), 552–564.

Azmon, A., Faes, C. and Hens, N. (2014) On the estimation of the reproduction number based on misreported epidemic data. *Statistics in Medicine*, 33(7), 1176–1192.

Bailey, T., Carvalho, M., Lapa, T., Souza, W. and Brewer, M. (2005) Modeling of under-detection of cases in disease surveillance. *Annals of Epidemiology*, 15(5), 335–343.

Bauer, C. and Wakefield, J. (2018) Stratified space–time infectious disease modelling, with an application to hand, foot and mouth disease in China. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(5), 1379–1398.

Becker, A.D. and Grenfell, B.T. (2017) tsir: an R package for time-series susceptible-infected-recovered models of epidemics. *PLOS One*, 12(9), 1–10.

Becker, N. (2015) *Modeling to Inform Infectious Disease Control*. Boca Raton, FL: CRC Press.

Bracher, J. (2019) Comment on "Under-reported data analysis with INAR-hidden Markov chains." *Statistics in Medicine*, 38(5), 893–898.

Bracher, J. and Held, L. (2020) Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction. *International Journal of Forecasting*. Preprint available at https://arxiv.org/abs/1901.03090.

Brännäs, K., Hellström, J. and Nordström, J. (2002) A new approach to modelling and forecasting monthly guest nights in hotels. *International Journal of Forecasting*, 18(1), 19–30.

Bundesagentur für Arbeit (2020) Pendleratlas. Available at https://statistik.arbeitsagentur.de/DE/Navigation/Statistiken/Interaktive-Angebote/Pendleratlas/Pendleratlas-Nav.html. Accessed September 22, 2020.

Cauchemez, S. and Ferguson, N.M. (2008) Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *Journal of the Royal Society Interface*, 5(25), 885–897.

Cui, Y. and Zheng, Q. (2017) Conditional maximum likelihood estimation for a class of observation-driven time series models for count data. *Statistics and Probability Letters*, 123, 193–201.

Dorigatti, I., Cauchemez, S., Pugliese, A. and Ferguson, N.M. (2012) A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: application to the Italian 2009–2010 A/H1N1 influenza pandemic. *Epidemics*, 4(1), 9–21.

Dvorzak, M. and Wagner, H. (2016) Sparse Bayesian modelling of underreported count data. *Statistical Modelling*, 16(1), 24–46.

Ferland, R., Latour, A. and Oraichi, D. (2006) Integer-valued GARCH process. *Journal of Time Series Analysis*, 27(6), 923–942.

Fernández-Fontelo, A., Cabaña, A., Puig, P. and Moriña, D. (2016) Under-reported data analysis with INAR-hidden Markov chains. *Statistics in Medicine*, 35(26), 4875–4890.

Fernández-Fontelo, A., Cabaña, A., Joe, H., Puig, P. and Moriña, D. (2019) Untangling serially dependent underreported count data for gender-based violence. *Statistics in Medicine*, 38(22), 4404–4422.

Fintzi, J., Cui, X., Wakefield, J. and Minin, V.N. (2017) Efficient data augmentation for fitting stochastic epidemic models to prevalence data. *Journal of Computational and Graphical Statistics*, 26(4), 918–929.

Fokianos, K. (2016) Statistical analysis of count time series models: a GLM perspective. In: Davis, R., Holan, S., Lund, R. and Ravishanker, N. (Eds.) *Handbook of Discrete-Valued Time Series*. Boca Raton, FL: Wiley, pp. 3–28.

Fraser, C., Donnelly, C.A., Cauchemez, S., Hanage, W.P. and Van Kerkhove, M. D. (2009) Pandemic potential of a strain of Influenza A (H1N1): early findings. *Science*, 324(5934), 1557–1561.

Gamado, K.M., Streftaris, G. and Zachary, S. (2014) Modelling under-reporting in epidemics. *Journal of Mathematical Biology*, 69(3), 737–765.

Gibbons, C.L., Mangen, M.-J.J., Plass, D., Havelaar, A.H., Brooke, R.J. and Kramarz, P. (2014) Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health*, 14(1), 1–17.

Gómez-Rubio, V., Palmí-Perales, F., López-Abente, G., Ramis-Prieto, R. and Fernández-Navarro, P. (2019) Bayesian joint spatio-temporal analysis of multiple diseases. *SORT-Statistics and Operations Research Transactions*, 43(1), 51–74.

Grimwood, K., Abbott, G.D., Fergusson, D.M., Jennings, L.C. and Allan, J.M. (1983) Spread of rotavirus within families: a community based study. *British Medical Journal*, 287(6392), 575–577.

Held, L., Höhle, M. and Hofmann, M. (2005) A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling*, 5(3), 187–199.

Held, L., Meyer, S. and Bracher, J. (2017) Probabilistic forecasting in infectious disease epidemiology: the 13th Armitage lecture. *Statistics in Medicine*, 36(22), 3443–3460.

Held, L. and Paul, M. (2012) Modeling seasonality in space-time infectious disease surveillance data. *Biometrical Journal*, 54(6), 824–843.

Heymann, D. (Ed.) (2015) *Control of Communicable Diseases Manual*, 20th edition. Washington, DC: APHA Press.

Jandarov, R., Haran, M., Bjørnstad, O. and Grenfell, B. (2014) Emulating a gravity model to infer the spatiotemporal dynamics of an infectious disease. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 63(3), 423–444.

King, A., Nguyen, D. and Ionides, E. (2016) Statistical inference for partially observed Markov processes via the R package pomp. *Journal of Statistical Software*, 69(12), 1–43.

Magal, P. and Webb, G. (2018) The parameter identification problem for SIR epidemic models: identifying unreported cases. *Journal of Mathematical Biology*, 77(6), 1629–1648.

Meyer, S. and Held, L. (2014) Power-law models for infectious disease spread. *Annals of Applied Statistics*, 8(3), 1612–1639.

Meyer, S. and Held, L. (2017) Incorporating social contact data in spatio-temporal models for infectious disease spread. *Biostatistics*, 18(2), 338–351.

Noufaily, A. (2019) Underreporting and reporting delays. In: Held, L., Hens, N., O'Neill, P. and Wallinga, J. (Eds.) *Handbook of Infectious Disease Data Analysis*. London, UK: Chapman and Hall/CRC, pp. 437–454.

Oliveira, G.L., Loschi, R.H. and Assunção, R.M. (2017) A random-censoring Poisson model for underreported data. *Statistics in Medicine*, 36(30), 4873–4892.

Paul, M., Held, L. and Toschke, A.M. (2008) Multivariate modelling of infectious disease surveillance data. *Statistics in Medicine*, 27(29), 6250–6267.

Plummer, M. (2015) Cuts in Bayesian graphical models. *Statistics and Computing*, 25(1), 37–43.

Robert Koch Institute (2020) Survstat@rki 2.0. Web-based query on data reported under the German Protection against Infection Act. https://survstat.rki.de/ [Accessed 03 September 2020] .

Sharmin, S., Glass, K., Viennet, E. and Harley, D. (2018a) A Bayesian approach for estimating under-reported dengue incidence with a focus on non-linear associations between climate and dengue in Dhaka, Bangladesh. *Statistical Methods in Medical Research*, 27(4), 991–1000.

Sharmin, S., Glass, K., Viennet, E. and Harley, D. (2018b) Geostatistical mapping of the seasonal spread of under-reported dengue cases in Bangladesh. *PLOS Neglected Tropical Diseases*, 12(11), 1–13.

Silvestrini, A. and Veredas, D. (2008) Temporal aggregation of univariate and multivariate time series models: a survey. *Journal of Economic Surveys*, 22(3), 458–497.

Staudenmayer, J. and Buonaccorsi, J.P. (2005) Measurement error in linear autoregressive models. *Journal of the American Statistical Association*, 100(471), 841–852.

Stocks, T., Britton, T. and Höhle, M. (2018) Model selection and parameter estimation for dynamic epidemic models via iterated filtering: application to rotavirus in Germany. *Biostatistics*, 21(3), 400–416.

Stoner, O., Economou, T. and Drummond, G. (2019) A hierarchical framework for correcting under-reporting in count data. *Journal of the American Statistical Association*, 114, 1481–1492.

Tam, C.C., Rodrigues, L.C., Viviani, L., Dodds, J.P., Evans, M.R., Hunter, P.R., et al. (2012) Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut*, 61(1), 69–77.

van Gaalen, R., van de Kassteele, J., Hahné, S., Bruijning-Verhagen, P. and Wallinga, J. (2017) Determinants of rotavirus transmission. A lag nonlinear time series analysis. *Epidemiology*, 28(4), 503–513.

Wakefield, J., Dong, Q. and Minin, V.N. (2019) Spatio-temporal analysis of surveillance data. In: Held, L., Hens, N., O'Neill, P. and Wallinga, J. (Eds.) *Handbook of Infectious Disease Data Analysis*. London, UK: Chapman and Hall/CRC, pp. 455–476.

Wang, Y., Feng, Z., Yang, Y., Self, S., Gao, Y., Longini, I.M., et al., (2011) Hand, foot, and mouth disease in China: patterns of spread and transmissibility. *Epidemiology*, 22(6), 781–792.

Weidemann, F., Dehnert, M., Koch, J., Wichmann, O. and Höhle, M. (2014) Bayesian parameter inference for dynamic infectious disease modelling: rotavirus in Germany. *Statistics in Medicine*, 33(9), 1580–1599.

White, L.F. and Pagano, M. (2010) Reporting errors in infectious disease outbreaks, with an application to pandemic influenza A/H1N1. *Epidemiologic Perspectives and Innovations*, 7(1), 1–12.

Woudenberg, T., Woonink, F., Kerkhof, J., Cox, K., Ruijs, W., van Binnendijk, R., et al. (2019) The tip of the iceberg: incompleteness of measles reporting during a large outbreak in The Netherlands in 2013–2014. *Epidemiology and Infection*, 147, e23.

Zhu, F. (2011) A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, 32(1), 54–67.

Zucchini, W. and MacDonald, I. (2009) *Hidden Markov Models for Time Series*. New York, NY: Chapman and Hall/CRC.

SUPPORTING INFORMATION

Web Appendices, Tables, and Figures referenced in Sections 2–5 of this article are available at the Biometrics website on Wiley Online Library. The suggested methods have been implemented in R in the package hhh4underreporting. The package can be installed from the repository https://github.com/jbracher/hhh4underreporting which also contains data and and code to replicate the presented application. Along with simulation codes, these are also available in the Supporting Information of this paper.

How to cite this article: Bracher J, Held L. A marginal moment matching approach for fitting endemic-epidemic models to underreported disease surveillance counts. *Biometrics*. 2021;77:1202–1214. https://doi.org/10.1111/biom.13371