RtEstim: Effective reproduction number estimation with trend filtering

Jiaping Liu

Department of Statistics

The University of British Columbia

Vancouver, BC V6T 1Z4

jiaping.liu@stat.ubc.ca

Zhenglun Cai

Centre for Health Evaluation and Outcome Sciences

The University of British Columbia

Vancouver, BC V6T 1Z4

ecai@cheos.ubc.ca

Paul Gustafson

Department of Statistics

The University of British Columbia

Vancouver, BC V6T 1Z4

gustaf@stat.ubc.ca

Daniel J. McDonald*

Department of Statistics

The University of British Columbia

Vancouver, BC V6T 1Z4

daniel@stat.ubc.ca

September 19, 2023

Abstract

---

*Corresponding author.

1   Introduction

Effective reproduction numbers (or instantaneous reproduction numbers) are a key to understand infectious disease dynamics including the potential sizes of a pandemic, the scale of epidemic prevention measures such as the proportion of a population that should be vaccinated, and the effectiveness of control effects. They are expected numbers of secondary infections caused by an infected individual in a population. It is a time series that can reflect the effect of time-varying factors, such as intervention policy and population immunity. Let $\mathcal{R}_t$ be the effective reproduction number at time t. A practical interpretation is that $\mathcal{R}_t < 1$ represents a circumstance when the infection dies out gradually and achieves a *disease-free equilibrium*, whereas when $\mathcal{R}_t > 1$, the infection is always present in a state of *endemic equilibrium*. Effective reproduction number reveals an unobservable biological reality. There exist a great number of models to uncover this reality relying on various domain-specific assumptions and using different types of observed data such as incidence data and death rates. Different estimation processes and results suggest different aspects and levels of policy making in epidemic control. For example, if effective reproduction numbers are assumed to be affected by human-human interaction over time in an epidemic model, controlling the parameters influencing human-human contact rate may eventually lead to a reduction of the effective reproduction numbers in the specific model. Since some assumptions cannot be verified in practice, it is critical for an estimator to be *robust* to model misspecification. The estimation relies heavily on the quality of the available data. The infectious data may contain poor information due to the limitations of data collection, which results in difficulties in reproduction number estimation.

Many existing approaches for effective reproduction number estimation are Bayesian approaches that estimate posterior distribution of $\mathcal{R}_t$. Cori et al. (2013) proposed a Bayesian approach, EpiEstim, to solve the posterior distribution of $\mathcal{R}_t$ given incidence data prior to time t. An advantage of EpiEstim is that it depends on limited assumptions (the Poisson distributed incidences and Gamma distribution effective reproduction number) and only requires incidence data that is easily obtainable. For these reasons, EpiEstim does not require much domain expertise in implementation. It is one of the earliest approaches that are both succinct and accurate in $\mathcal{R}_t$ estimation. They proposed EpiEstim(2.2) in Thompson et al. (2019), which distinguished imported cases from local transmission and directly estimated the serial interval. They further extended EpiEstim by using reconstructed daily incidence data to overcome the issue when incidences are not always daily records in Nash et al. (2023). Parag (2021) proposed an alternative Bayesian approach, EpiFilter, that is a recursive Bayesian smoother based on Kalman Filter. EpiFilter also solves the posteriori of $\mathcal{R}_t$ given a Gamma prior and Poisson distributed infection counts. Compared to EpiEstim, EpiFilter estimates $\mathcal{R}_t$ retrospectively using all available incidences and generates more robust estimation in low incidence periods and accurate one-step-ahead prediction. Abbott et al. (2020) proposed a Bayesian latent variable framework,

EpiNow2, which uses both incidence and death counts and provides precise $\mathcal{R}_t$ estimation and forecasts. They further proposed a generative Bayesian model to handle missing data by imputation followed by truncation adjustment in Lison et al. (2023). Gressani et al. (2022) proposed a Bayesian P-splines approach, EpiLPS, and argued that this approach was computationally efficient. They assumed the incidence cases follow a negative binomial distribution. Trevisin et al. (2023) proposed a Bayesian model based on particle filtering to estimate spatially explicit effective reproduction numbers. Bayesian approaches estimate the posterior distribution of the effective reproduction numbers with the advantage that the credible interval estimation can be easily computed. A limitation of many Bayesian approaches, however, is that they usually require heavy computational workload, especially when data sequences are long or hierarchical structures are complex.

Abry et al. (2020) proposed to regularize the smoothness of $\mathcal{R}_t$ regarding its temporal and spatial evolution. They considered a penalized regression with a second-order temporal regularization and a spatial regularization on $\mathcal{R}_t$ and with Poisson loss. They further extended it by introducing another penalty on outliers for robustness in Pascal et al. (2022). There are other spline-based approaches such as Azmon et al. (2014); Gressani et al. (2021); Pircalabelu (2023b). Jin et al. (2023) proposed EpiMix, which included exogenous factors other than incidence cases and introduced random effects in regression model. It is also robust against poor data condition, i.e., low incidence. There are many other practical considerations in effective reproduction number estimation. Hettinger et al. (2023) proposed a generalized autoregresive moving average (GARMA) model. It solved the bias introduced by practical concerns such as particular forms of measurement errors in incidence data. They also incorporates multiple serial interval distributions. Pircalabelu (2023a) is a frequentist method and based on splines. They assumed the incidence follow an exponential-family distribution. Ho et al. (2023) considered overdispersion by making a negative binomial assumption to estimate Covid-19 $\mathcal{R}_t$.

We propose a retrospective effective reproduction number estimation approach to filter out discrete splines. More specifically, it is an optimization problem with Poisson loss and ℓ_1 penalty on the temporal evolution of $\mathcal{R}_t$, which is known as the trend filtering penalty (Kim et al., 2009; Tibshirani, 2014). The estimated curves appear to be piecewise polynomial. The estimators have both local adaptivity (i.e., heterogeneous smoothness throughout the range of time) and allow the computational efficiency. Similar as the aforementioned approaches, our approach only requires limited assumptions on the distribution of incidences and serial intervals, and depends on the widely obtainable incidence data. Thus, it only requires limited domain expertise to apply our approach. We follow the common assumption of Poisson distributed incidence data, but the empirical study shows that the estimators are robust in negative binomial settings. Our effective reproduction number estimation is accurate and robust in model misspecification and computationally efficient compared

to the aforementioned competitors. In computation, our approach takes the advantage of convex optimization and is solved by Newton's method, which is known to converge rapidly. Moreover, the sparse structure of the divided difference matrix used in the temporal penalty allows further efficiency in computation.

2 Methods

2.1 Renewal model for incidence data

Effective reproduction number $\mathcal{R}_t$, the expected secondary infection by a primary infection in a population at time t, is inherently a ratio of new infections at t over the total primary infectious cases until t (excluding the new infections). Here, we assume an ideal scenario of homogenous population that individuals follow similar social behaviors, exposure risks and random mixing patterns such as having similar contact rates to each other or similar susceptibility and infectiousness. Denote the new infections at time t as N_t. Given an infectious period τ_t at t, we construct the total primary infectiousness as $\Lambda_t := \sum_{i=1}^{\tau_t} p_i N_{t-i}$, where p_i is the infectious probability that a secondary case is infected by a primary case which is infected i timepoints ago. The reproduction number can then be constructed based on the ratio $\mathcal{R}_t := N_t/\Lambda_t$. We assume a constant observable proportion of infections $c \in (0, 1)$, which is applied to all incidence counts and gets cancelled in $\mathcal{R}_t$ ratio. Rearranging terms of the ratio yields the widely used renewal equation

$$N_t = \sum_{i=1}^{\tau_t} p_i \mathcal{R}_t N_{t-i} \tag{1}$$

for the analysis of transmission dynamics of infectious diseases. Some examples are EpiEstim (Cori et al., 2013) and EpiFilter (Parag, 2021).

The sequence of probabilities $p_{1:\tau_t}$ gives the probabilities that a secondary infection at t is infected by a primary infection that is infected 1 to τ_t timepoints ago. The period between primary and secondary infections is called generation time. Theoretically, $p_{1:\tau_t}$ are the cumulative probabilities of generation time in discretized, contiguous time intervals, i.e., $(0, 1], (1, 2], \ldots, (\tau_{t-1}, \tau_t]$. We assume that the infectiousness disappears beyond τ_t timepoints, so that the sequence $p_{1:\tau_t}$ has a sum of 1.

The generation time (period between the infections), however, is unobservable and tricky to estimate. We take a common strategy — approximate it by serial interval (period between the onsets of symptoms). When the infectiousness profile after symptoms is independent of the incubation period, the serial interval is regarded identical as the generation time (Cori et al., 2013). We assume the distribution of generation time, and correspondingly serial interval, to be independent to time t, i.e., the probability p_i only corresponds to the relative time range i between a primary and

secondary infections. The sequence $p_{1:\tau_t}$ only depends on t by a chosen length τ_t; if τ_t is given, the probability sequence will be right-truncated at τ_t and rescaled to have a sum of 1. We choose Gamma distribution to estimate the serial interval, which is demonstrated to be a reasonable choice in practice.

The renewal equation in Equation (1) quantifies the transmission dynamic that primary incidences result in new incidences by effective reproduction numbers given serial interval. This dynamic is straightforward and efficient in data usage as it only depends on the observed incidence counts, which can be easily obtainable and usually sufficient and of good quality, and the specification of serial interval distribution.

2.2 Poisson trend filtering estimator

We use the daily confirmed cases y_t on day t to estimate the observed infectious cases by assuming a consistent incubation period (from the time of infections to the time of symptom onsets when the cases are confirmed) and further assume y_t to be Poisson distributed with mean I_t, i.e.,

$$y_t \sim \mathrm{Poisson}(I_t) = \mathrm{Poisson}(\Lambda_t \mathcal{R}_t).$$

Considering a fixed period of days n, our interest is to estimate the Poisson parameter $\mathcal{R}_t$ given observations $y_{1:n} := \{y_1, \ldots, y_n\}$ and the total infectiousness based on confirmed cases $\Lambda_t^* := \sum_{i=1}^{\tau_t} p_i y_{t-i}$. A natural approach is to solve the maximum likelihood estimates (MLEs), i.e.,

$$\hat{\mathcal{R}}_t := \operatorname*{argmax}_{\mathcal{R}_t} \mathbb{P}(\mathcal{R}_t \mid y_{1:n}, p_{1:\tau_t}) = \prod_{t=1,\ldots,n} \frac{e^{-\mathcal{R}_t \Lambda_t^*} (\mathcal{R}_t \Lambda_t^*)^{y_t}}{y_t!}. \tag{2}$$

This maximization problem, however, yields a one-to-one correspondence between the confirmed case (or the total infectiousness) and the effective reproduction number at each day, so that the estimated curves have no significant graphical smoothness.

Smoothness of the effective reproduction numbers is a key to understand the trend of transmissibility of infectious diseases in retrospective studies. Smoother estimated curves give more high-level ideas with less changing points and hide the information of minor importance, and vice versa. We assume the effective reproduction numbers to be piecewise polynomials with multiple knots (i.e., changing points) with varying degrees. We specifically consider discrete splines with various degrees of continuity. For instance, the 0th degree discrete splines are piecewise constant, the 1st degree curves are piecewise linear, and the 2nd degree curves are piecewise quadratic. For $k \geq 2$, the kth degree discrete splines are continuous and have continuous discrete differences up to degree $k - 1$ at the knots.

To achieve such smoothness, we regularize the distance between adjacent effective reproduction numbers. Since $\mathcal{R}_t > 0$, penalizing the distance between $\mathcal{R}_t$s directly may cause numerical issues such that there may be negative estimates in computation. Therefore, we equivalently penalize the distance between natural logarithm between neighboring $\mathcal{R}_t$s through divided differences (i.e., discrete derivatives) with various orders. Compared to splines, discrete splines introduce computational efficiency. We penalize ℓ_1 norm of the distance, which introduces sparsity into the curvature, so that the estimates have heterogeneous smoothness in different subregions of the entire domain. It is a more realistic setting compared to homogeneous smoothness in squared ℓ_2 norm. The divided differences (i.e., discrete derivatives) with various orders to represent the temporal evolution of reproduction numbers with different degrees.

We define a penalized regression problem to solve the MLE problem in Equation (2) and the smoothness regularization simultaneously in . It is a minimization problem with Poisson loss and the trend filtering penalty (Kim et al., 2009; Tibshirani, 2014). Let $\theta := \log(\mathcal{R}) \in \mathbb{R}^n$, and then $\Lambda \circ \mathcal{R} = \Lambda \circ e^{\theta}$, $\log(\Lambda \circ \mathcal{R}) = \log(\Lambda) + \theta$, where $\circ$ is elementwise product, $e^a, \log(a)$ apply to a vector a elementwise. The problem solves a Poisson trend filtering (PTF) estimator on univariate cases. For evenly spaced observed incidences, it is defined as:

$$
\hat{\theta} := \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} -y_i \theta_i + \Lambda_i e^{\theta_i} + \lambda \left\| D^{(k+1)} \theta \right\|_1,
$$
$$
\hat{\mathcal{R}} := e^{\hat{\theta}},
$$

(3)

where $D^{k+1} \in \mathbb{Z}^{(n-k-1) \times n}$ is a k-th order divided difference matrix with $k = 0, 1, 2, \ldots$. Define $D^{(k+1)}$ recursively as $D^{(k+1)} := D^{(1)} D^{(k)}$, where $D^{(1)} \in \mathbb{N}^{(n-k-1) \times (n-k)}$ is a banded matrix defined as

$$
D^{(1)} := \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & & \vdots & \vdots \\ & & & -1 & 1 \end{pmatrix}.
$$

Define $D^{(0)} := I_n$, which is an identity matrix with size n. An exponential transformation is applied to the PTF estimator to get the estimated reproduction numbers.

The tuning parameter λ balances the contributions between data fidelity and smoothness. When $\lambda = 0$, the problem in Equation (3) reduces to regular least squares problem. A larger tuning parameter gives a higher importance on the regularization term and yields a smoother curve until the divided differences are all zeros, i.e., all parameters are projected onto the null space of the corresponding divided difference matrix.

For unevenly spaced observations, the distances between neighboring parameters vary by the periods between observation times, and thus, the divided differences should be adjusted by observation days (or data locations). Given the data locations $x_{1:n} = \{x_1, \ldots, x_n\}$, define a kth order diagonal matrix

$$X^k := \text{diag}\left(\frac{k}{x_{k+1} - x_1}, \frac{k}{x_{k+2} - x_2}, \cdots, \frac{k}{x_n - x_{n-k}}\right)$$

for $k \geq 1$. Let $D^{(x,1)} := D^{(1)}$ and define $D^{(x,k+1)}$ for $k \geq 1$ recursively as

$$D^{(x,k+1)} := D^{(1)} \cdot X^k \cdot D^{(x,k)}.$$

Pascal et al. (2022) considered the second-order divided different of effective reproduction number. In comparison to their study, our estimator is more flexible in the degree of temporal evolution of the effective reproduction numbers and also avoid the potential numerical issues of penalizing/estimating positive numbers. Our estimator is locally adaptive so that it captures the local changes such as the initiation of effective control measures. More specifically, it regularizes the similarity among reproduction numbers across a chosen number of neighboring time points and segments the curvature of the reproduction numbers such that there are more jumpiness in some subregions and more smoothness in others.

2.3   Proximal Newton solver

The proximal Newton method is a second-order algorithm solving a proximal Newton optimization iteratively followed by a line search algorithm adjusting the step size at each iteration for faster convergence. The proximal Newton method for Poisson trend filtering in Equation (3) takes a second-order Taylor expansion of the Poisson loss, which results in a proximal optimization — Gaussian trend filtering with dynamic weights during iteration.

Let $g(\theta) := \frac{1}{n}\sum_{i=1}^{n} -y_i\theta_i + \Lambda_i e^{\theta_i}$ be the Poisson regression loss and $h(\theta) := \lambda \left\|D^{(k+1)}\theta\right\|_1$ be the regularization. At iterate $t + 1$, consider the following approximation of $g(\theta)$ using a second-order Taylor expansion around θ^t:

$$g(\theta) = g(\theta^t) + (\theta - \theta^t)^\top \nabla_\theta^{(1)} g(\theta^t) + \frac{1}{2}(\theta - \theta^t)^\top \nabla_\theta^{(2)} g(\theta^t)(\theta - \theta^t),$$

where $\nabla_\theta^{(1)} g(\theta^t) = \frac{1}{n}\left(-y + \Lambda \circ e^{\theta^t}\right) \in \mathbb{R}^n$ is the gradient of $g(\theta)$ at θ^t and $\nabla_\theta^{(2)} g(\theta^t) = \frac{1}{n}\text{diag}\left(\Lambda \circ e^{\theta^t}\right) \in \mathbb{R}^{n \times n}$ is the Hessian matrix of $g(\theta)$ at θ^t.

Define the proximal operator as $\text{prox}_{W,D}(x) := \underset{z \in \mathbb{R}^n}{\text{argmin}} \frac{1}{2n} \|z - x\|_W^2 + \lambda \|D\theta\|_1$, where $\|a\|_W^2 := a^\top W a$. The proximal optimization problem at iterate $t + 1$ can be further written as, given θ^t,

$$
\begin{aligned}
\theta^{t+} :&= \underset{\theta \in \mathbb{R}^n}{\text{argmin}} \; (\theta - \theta^t)^\top \nabla_\theta^{(1)} g(\theta^t) + \frac{1}{2}(\theta - \theta^t)^\top \nabla_\theta^{(2)} g(\theta^t)(\theta - \theta^t) + h(\theta), \\
&= \underset{\theta \in \mathbb{R}^n}{\text{argmin}} \; \frac{1}{2n} \|\theta - c^t\|_{W^t}^2 + \lambda \|D^{(k+1)}\theta\|_1, \\
&= \text{prox}_{W^t, D^{(k+1)}}(c^t),
\end{aligned}
\tag{4}
$$

where $W^t := \text{diag}\left(\Lambda \circ e^{\theta^t}\right)$ is the weighted (Hessian) matrix multiplied by n and $c^t := \theta^t - (W^t)^{-1}\left(n\nabla_\theta^{(1)} g(\theta^t)\right) = y \circ \Lambda^{-1} \circ e^{-\theta^t} - \mathbf{1} + \theta^t \circ \Lambda^{-1}$, where $\{e^{\theta^t}\}_{i \in [n]} > 0$, $[n] := 1, 2, \ldots, n$. This is just univariate *Gaussian trend filtering* with weights W^t (Tibshirani, 2014).

We solve the trend filtering problem in Equation (4) using the specialized ADMM, proposed by Ramdas and Tibshirani (2016), with the primal θ step solved in closed-form and the auxiliary step solved by the dynamic programming algorithm for fused lasso proposed by Johnson (2013). Let the auxiliary variable $z := D^{(k)}\theta$. The scaled augmented Lagrangian is

$$
\mathcal{L}_{\lambda,\rho}(\theta, z, u) = \frac{1}{2n} \|\theta - c^t\|_{W^t}^2 + \lambda \|D^{(1)}z\|_1 + \frac{\rho}{2} \|D^{(k)}\theta - z + u\|^2 - \frac{\rho}{2} \|u\|^2,
$$

where ρ is a scaled dual parameter and u is a dual variable. The specialized ADMM solves the following subproblems, at ADMM iteration $l + 1$:

$$
\begin{aligned}
\theta^{l+1} &:= \underset{\theta}{\text{argmin}} \; \frac{1}{2n} \|\theta - c^t\|_{W^t}^2 + \frac{\rho}{2} \|D^{(k+1)}\theta - z^l + u^l\|_2^2, \\
z^{l+1} &:= \underset{z}{\text{argmin}} \; \frac{\lambda}{\rho} \|D^{(1)}z\|_1 + \frac{1}{2} \|D^{(k+1)}\theta^{l+1} - z + u^l\|_2^2, \\
u^{l+1} &\leftarrow u^l + D^{(k+1)}\theta^{l+1} - z^{l+1}.
\end{aligned}
\tag{5}
$$

We further adjust the step size $\gamma^t \in (0, 1]$ at iterate t by a backtracking line search algorithm:

$$
\theta^{t+1} \leftarrow \theta^t + \gamma^{t+1}(\theta^{t+} - \theta^t).
$$

The proximal Newton algorithm iterates until the convergence of objective.

2.4 Bayesian perspective

Since many current approaches are Bayesian methods, we here provide an alternative interpretation of our approach from the Bayesian perspective. Our approach can be interpreted as a state-space model of Poisson observational noises and Laplace transition noises with certain degree $k \geq 0$, e.g., $\theta_{t+1} = 2\theta_t - \theta_{t-1} + \varepsilon_{t+1}$ with $\varepsilon_{t+1} \sim \text{Laplace}(0, 1/\lambda)$ for $k = 1$. Compared to EpiFilter (Parag, 2021), another retrospective study of $\mathcal{R}_t$, we share same observational assumptions, but

our approach has a different transition noises. EpiFilter estimates the posterior distribution of $\mathcal{R}_t$, and thus it can provide the credible interval estimation with various credible levels. Our approach solves the point estimation using optimization problem, which has the advantage of computational efficiency.

3 Results

Implementation of the proximal Newton method is provided in the R package `rtestim`.

3.1 Covid-19 cases

We implement the proposed model on the Covid-19 confirmed cases in British Columbia (B.C.) as of May 18, 2023 reported by B.C. Conservation Data Centre. We choose the gamma distribution with shape 2.5 and scale 2.5 to approximate the serial interval function.

Considering the temporal evolutions of neighboring $3, 4, 5$ reproduction numbers, the estimated reproduction numbers of Covid-19 in British Columbia (displayed in the top right, bottom left, and bottom right panels in Fig 1 respectively) are always lower than 2.5, which means that two distinct infected individuals can on average infect less than five other individuals in the population. The three degrees of the temporal evolution (across all regularization levels λ) all yield similar results that $\hat{\mathcal{R}}_t$ achieves the highest peak around the end of 2021 and reaches the lowest trough shortly thereafter. Throughout the estimated curves, the peaks and troughs of the reproduction numbers roughly come prior to the following growths and decays of confirmed cases respectively.

The reproduction numbers are relatively unstable before April 1st, 2022. The highest peak coincides with the emergence and globally spread of the Omicron variant. The estimated reproduction numbers are apparently below the threshold 1 during two time periods – roughly from April 1st, 2021 to July 1st, 2021 and from January 1st, 2022 to April 1st, 2022. The first trough of $\hat{\mathcal{R}}_t$ coincides with the first authorization for use of Covid-19 vaccines in British Columbia. The second trough shortly after the greatest peak may credit to many aspects, including self-isolation of the infected individuals and application of the second shot of Covid-19 vaccines. Since around April 1st, 2022, the reproduction numbers stay stable (at around 1) and the infected cases stay low.

Greater regularization levels (i.e., larger λs) result in smoother estimated curves. Smoother curves (e.g., the yellow curve in the top right panel in Fig 1) suggest that the estimated reproduction numbers are around 1 during most time periods; however, they may not be appropriate to interpret the reality. More wiggly curves better reflect the fluctuation of $\mathcal{R}_t$, but sometimes fail to highlight the significant peaks or troughs. The tuning parameter λ needs to be chosen corresponding to the information in practice for a better interpretation.

4 Discussion

Our approach provides a locally adaptive estimator using Poisson trend filtering on univariate data. It captures the heterogeneous smoothness of effective reproduction numbers given the observed incidence time series in a certain region. This is a nonparametric regression model which can be written as convex optimization(minimization) problem. Minimizing the distance (averaged KL divergence per coordinate) between the estimators and (functions of) observations guarantees the data fidelity; minimizing a certain order of divided differences between each pair of neighboring parameters regularizes the smoothness. The ℓ_1 regularization introduces sparsity to the divided differences, which leads to heterogeneous smoothness within certain periods of time. The homogeneous smoothness within a time period can be either performed by a constant reproduction number, or a constant rate of changes, or a constant graphical curvature depending on the prescribed degree ($k = 0, 1, 2$ respectively).

The property of local adaptivity is useful to distinguish, for example, the seasonal outbreaks from the un-seasonal outbreak periods. Given a properly chosen degree of polynomials, for example, the growth rate of un-seasonal outbreak periods can suggest a potential upcoming outbreak, which alerts epidemiologists to propose sanitary policies to prevent the progressing outbreak ahead of the infection surge. The effective reproduction numbers can be estimated afterwards to check the efficiency of the sanitary policies referring to whether they are below the threshold, their tendencies of reduction, or their graphical curvatures.

Our proposed model provides a natural way to deal with missing data, e.g., on weekends and holidays. We linearly impute the missing data in the computation of total primary infectiousness by assuming these values are missing at random. While solving the convex optimization problem, the edge lengths of the line graphs can be adjusted, so we can manually increase the length between two observations while penalizing the distance between them. It is remarkable that our focus is to provide a mathematical model for epidemiologists to use, rather than to focus on a specific disease. In addition, more specialized methodologies are needed for the diseases with relatively long incubation periods (e.g., HIV and HBV).

A group of epidemiological models are compartmental models. They establish the epidemic transmission process by creating compartments with labels and connecting them by directed edges. A simple compartmental model – for example, *Susceptible-Infectious-Susceptible* (SIS) model – divides the population (N) into two compartments for susceptible cases (S) and infectious cases (I) respectively and connects them in serial as $S \rightarrow I \rightarrow S$. It only focuses on susceptible individuals. Each directed edge corresponds to a ratio of transmission (say, α, β respectively). In such models, reproduction numbers are defined as functions of the estimated transmission parameters and the

numbers of compartments or population, e.g., $\hat{\mathcal{R}}_0 = \hat{\beta}N/\hat{\alpha}$ in the SIS models Brauer et al. (2019), as by-products. Compartmental models usually solve ordinary differential equations (ODE) systems for transmission numbers (e.g., α, β in the SIS model). A disadvantage of such parametric models is that they are less flexible than nonparametric models and the number of parameters to be estimated grows along with the increase of compartments in practice, which results in a growing computational complexity. Since the epidemic mechanism depends highly on the contexts, e.g., if a latency period exists or not, such models are lack of generalizability. Moreover, data of high quality are not always available for all compartments especially when there is a pandemic outbreak that results in a sudden shortage of resources in collecting daily new infections.

There are more practical considerations that may influence the quality of $\mathcal{R}_t$ estimation to be considered late. In our approach, we consider a homogeneous population without distinguishing the imported cases from the local cases. Poisson distribution is frequently used to model non-negative count data with heteroskedasticity. Another common alternative is negative binomial distribution with a specified level of overdispersion. We consider a fixed serial interval throughout the transmission dynamics, but as the factors such as population immunity vary, the serial interval may vary as well. A common argument is that the distribution density of serial intervals is generally wider than the correspondence of generation intervals as serial interval includes both generation time and incubation time. If we assume generation time and incubation time both follow gamma distributions, the serial interval is likely to perform as a bimodal density.

References

Sam Abbott, Joel Hellewell, Robin N Thompson, Katharine Sherratt, Hamish P Gibbs, Nikos I Bosse, James D Munday, Sophie Meakin, Emma L Doughty, June Young Chun, et al. Estimating the time-varying reproduction number of sars-cov-2 using national and subnational case counts. *Wellcome Open Research*, 5(112):112, 2020.

Patrice Abry, Nelly Pustelnik, Stéphane Roux, Pablo Jensen, Patrick Flandrin, Rémi Gribonval, Charles-Gérard Lucas, Éric Guichard, Pierre Borgnat, and Nicolas Garnier. Spatial and temporal regularization to estimate covid-19 reproduction number r (t): Promoting piecewise smoothness via convex optimization. *Plos one*, 15(8):e0237901, 2020.

Amin Azmon, Christel Faes, and Niel Hens. On the estimation of the reproduction number based on misreported epidemic data. *Statistics in medicine*, 33(7):1176–1192, 2014.

Fred Brauer, Carlos Castillo-Chavez, and Zhilan Feng. *Mathematical models in epidemiology*, volume 32. Springer, 2019.

Anne Cori, Neil M Ferguson, Christophe Fraser, and Simon Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American journal of epidemiology*, 178(9):1505–1512, 2013.

Oswaldo Gressani, Christel Faes, and Niel Hens. An approximate bayesian approach for estimation of the reproduction number under misreported epidemic data. *MedRxiv*, pages 2021–05, 2021.

Oswaldo Gressani, Jacco Wallinga, Christian L Althaus, Niel Hens, and Christel Faes. Epilps: A fast and flexible bayesian tool for estimation of the time-varying reproduction number. *PLoS computational biology*, 18(10):e1010618, 2022.

Gary Hettinger, David Rubin, and Jing Huang. Estimating the instantaneous reproduction number with imperfect data: A method to account for case-reporting variation and serial interval uncertainty. *arXiv preprint arXiv:2302.12078*, 2023.

Faith Ho, Kris V Parag, Dillon C Adam, Eric HY Lau, Benjamin J Cowling, and Tim K Tsang. Accounting for the potential of overdispersion in estimation of the time-varying reproduction number. *Epidemiology*, 34(2):201–205, 2023.

Shihui Jin, Borame Lee Dickens, Jue Tao Lim, and Alex R Cook. Epimix: A novel method to estimate effective reproduction number. *Infectious Disease Modelling*, 2023.

Nicholas A Johnson. A dynamic programming algorithm for the fused lasso and ℓ_0-segmentation. *Journal of Computational and Graphical Statistics*, 22(2):246–260, 2013.

Seung-Jean Kim, Kwangmoo Koh, Stephen Boyd, and Dimitry Gorinevsky. ℓ_1 trend filtering. *SIAM review*, 51(2):339–360, 2009.

Adrian Lison, Sam Abbott, Jana Huisman, and Tanja Stadler. Generative bayesian modeling to nowcast the effective reproduction number from line list data with missing symptom onset dates. *arXiv preprint arXiv:2308.13262*, 2023.

Rebecca K Nash, Samir Bhatt, Anne Cori, and Pierre Nouvellet. Estimating the epidemic reproduction number from temporally aggregated incidence data: A statistical modelling approach and software tool. *PLOS Computational Biology*, 19(8):e1011439, 2023.

Kris V Parag. Improved estimation of time-varying reproduction numbers at low case incidence and between epidemic waves. *PLoS Computational Biology*, 17(9):e1009347, 2021.

Barbara Pascal, Patrice Abry, Nelly Pustelnik, Stéphane Roux, Rémi Gribonval, and Patrick Flandrin. Nonsmooth convex optimization to estimate the covid-19 reproduction number space-time evolution with robustness against low quality data. *IEEE Transactions on Signal Processing*, 70:2859–2868, 2022.

Eugen Pircalabelu. A spline-based time-varying reproduction number for modelling epidemiological outbreaks. *Journal of the Royal Statistical Society Series C: Applied Statistics*, page qlad027, 2023a.

Eugen Pircalabelu. A spline-based time-varying reproduction number for modelling epidemiological outbreaks. *Journal of the Royal Statistical Society Series C: Applied Statistics*, page qlad027, 2023b.

Aaditya Ramdas and Ryan J Tibshirani. Fast and flexible admm algorithms for trend filtering. *Journal of Computational and Graphical Statistics*, 25(3):839–858, 2016.

Robin N Thompson, Jake E Stockwin, Rolina D van Gaalen, Jonny A Polonsky, Zhian N Kamvar, P Alex Demarsh, Elisabeth Dahlqwist, Siyang Li, Eve Miguel, Thibaut Jombart, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics*, 29:100356, 2019.

Ryan J Tibshirani. Adaptive piecewise polynomial estimation via trend filtering. *The Annals of Statistics*, 42(1):285–323, 2014.

Cristiano Trevisin, Enrico Bertuzzo, Damiano Pasetto, Lorenzo Mari, Stefano Miccoli, Renato Casagrandi, Marino Gatto, and Andrea Rinaldo. Spatially explicit effective reproduction numbers from incidence and mobility data. *Proceedings of the National Academy of Sciences*, 120(20): e2219816120, 2023.