RtEstim: Effective reproduction number estimation with trend filtering

Jiaping Liu[a,1], Zhenglun Cai[b], Paul Gustafson[a], and Daniel J. McDonald[a]

[a]Department of Statistics, The University of British Columbia
[b]Centre for Health Evaluation and Outcome Sciences, The University of British Columbia

October 27, 2023

Abstract

We need an abstract.

Keywords: kwd1 | kwd2

Contents

1 Introduction **3**

2 Methods **5**

2.1 Renewal model for incidence data 5
2.2 Poisson trend filtering estimator 7
2.3 Numerical optimization of the $\mathcal{R}_t$ estimator 9
2.4 Solving over a sequence of tuning parameters 11
2.5 Choosing a final λ 11
2.6 Approximate confidence bands 11
2.7 Bayesian perspective 11

[1]To whom correspondence should be addressed. E-mail: jiaping.liu@stat.ubc.ca

3 Results **11**
3.1 Experimental settings 11
3.2 Experimental results 15
3.3 Covid-19 incidences in British Columbia 17
3.4 Pandemic influenza in Baltimore, Maryland, 1918 19

4 Discussion **19**

References **21**

1 Introduction

The effective reproduction number (also called the instantaneous reproduction number) is a key quantity for understanding infectious disease dynamics including the potential size of a pandemic, the required stringency of prevention measures, and the efficacy of other controls. The effective reproduction number is defined to be the average number of secondary infections caused by a new primary infection that occurs at a specific time. Tracking the time series of this quantity is therefore useful for understanding whether or not future infections are likely to increase or decrease from the current state. Let $\mathcal{R}(t)$ denote the effective reproduction number at time t. Practically, as long as $\mathcal{R}(t) < 1$, infections will decline gradually, eventually resulting in a disease-free equilibrium, whereas when $\mathcal{R}(t) > 1$, infections will continue to increase, resulting in endemic equilibrium. While $\mathcal{R}(t)$ is fundamentally a continuous time quantity, it can be related to data only at discrete points in time $t = 1, \ldots, n$. This sequence of effective reproduction numbers over time is not observable, but nonetheless is easily interpretable and retrospectively describes the course of an epidemic. Therefore, a number of procedures exist to estimate $\mathcal{R}_t$ from different types of observed incidence data such as cases, deaths, or hospitalizations while relying on various domain-specific assumptions. Importangly, accurate estimation of effective reproduction numbers relies heavily on the quality of the available data, and, due to the limitations of data collection, such as underreporting and lack of standardization, epidemiological models, estimation methodologies rely on various assumptions to compensate. Because model assumptions may not be easily verifiable from data alone, it is also critical for any estimation procedure to be robust to model misspecification.

Many existing approaches for effective reproduction number estimation are Bayesian: they estimate the posterior distribution of $\mathcal{R}_t$ conditional on the observations. One of the first such approaches is the software `EpiEstim` (Cori et al., 2020), described by Cori et al. (2013). This method is prospective, in that it uses only information available at time t in order to estimate $\mathcal{R}_t$ for $i = 1, \ldots, t$. An advantage of `EpiEstim` is its straightforward statistical model: incidence data follows the Poisson distribution conditional on past incidence and $\mathcal{R}_t$, the conjugate prior distribution for $\mathcal{R}_t$ is Gamma with fixed hyperparameters, and the serial interval distribution is fixed and known. For this reason, `EpiEstim` requires little domain expertise for use, and it is computationally fast. Thompson et al. (2019) modified this method to distinguish imported cases from local transmission and simultaneously estimate the serial interval distribution. Nash et al. (2023) further extended `EpiEstim` by using "reconstructed" daily incidence data to handle cases when the data is not regularly spaced. Abbott et al. (2020b) proposed a Bayesian latent variable framework, `EpiNow2` (Abbott et al., 2020a), which leverages incident cases, deaths or other available streams simultaneously along with allowing

additional delay distributions (incubation period and onset to report delay) in modelling. Lison et al. (2023) proposed an extension that handles missing data by imputation followed by a truncation adjustment. Parag (2021) also proposed a Bayesian approach, `EpiFilter` based the (discretized) on Kalman Filter and Smoother. `EpiFilter` also estimates the posterior of $\mathcal{R}_t$ given a Gamma prior and Poisson distributed incident cases. Compared to `EpiEstim`, `EpiFilter` estimates $\mathcal{R}_t$ retrospectively using all available incidence data both prior and subsequent to time t, and provides robust estimation in low incidence cases. Gressani et al. (2022) proposed a Bayesian P-splines approach, `EpiLPS`, that assumes negative Binomial distributed incidences. 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, and possess the advantage that the credible intervals can be easily computed. A limitation of many Bayesian approaches is that they usually require heavy computational workload, especially when observed data sequences are long or hierarchical structures are complex. Below, we compare our method to two computationally efficient Bayesian models, EpiEstim and EpiLPS.

There are also frequentist approaches for $\mathcal{R}_t$ estimation. 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. Pascal et al. (2022) extended this procedure by introducing another penalty on outliers for robustness in. Pircalabelu (2023a) is a spline-based model relying on the assumption of exponential-family distributed incidences. Ho et al. (2023) estimates $\mathcal{R}_t$ while monitoring the time-varying level of overdispersion. There are other spline-based approaches such as Azmon et al. (2014); Gressani et al. (2021); Pircalabelu (2023b), autoregressive models with random effects (Jin et al., 2023) that are robust to low incidence cases, and generalized autoregressive moving average (GARMA) model (Hettinger et al., 2023) that are robust to measurement errors in incidence data.

We propose a retrospective effective reproduction number estimator called `RtEstim`that requires only incidence data. Our model makes the conditional Poisson assumption, similar to much of the prior work described above, but is empirically more robust to misspecification. This estimator is a convex optimization problem with Poisson loss and ℓ_1 penalty on the temporal evolution of $\log(\mathcal{R}_t)$ to impose smoothness over time. As a result `RtEstim` generates discrete splines, and the estimated curves appear to be piecewise polynomials of an order selected by the user. Importantly, The estimates are locally adaptive, meaning that different time ranges may posses heterogeneous smoothness.

ATTN: Insert "killer" figure here.

Our approach is straightforward and requires little expertise in domain knowledge for

implementation. We use a proximal Newton method to solve the convex optimization problem along with a number of computational tricks to produce estimates efficiently, typically in a matter of seconds even for long sequences of data. In a number of simulated experiments, we show empirically that our approach is more accurate than existing methods at estimating the true effective reproduction numbers.

The manuscript unfolds as follows. We first introduce the methodology of `RtEstim` including the usage of renewal equation, the development of Poisson trend filtering estimator, and the proximal Newton algorithm. We explain how this method can be interpreted from Bayesian perspective, connecting it to previous work in this context. We provide illustrative experiments comparing our estimator to `EpiEstim` and `EpiLPS`. We then apply our `RtEstim` on the Covid-19 pandemic incidence in British Columbia and the 1918 influenza pandemic in the United States. Finally, we conclude with a discussion on the advantages and limitations of our approach and as well as describe practical considerations for the effective reproduction number estimation.

2 Methods

2.1 Renewal model for incidence data

The effective reproduction number $\mathcal{R}(t)$ is defined to be the expected number of secondary infections at time t produced by a primary infection sometime in the past. To make this precise, denote the number of new infections at time t as $y(t)$. Then the total primary infectiousness can be written as $\eta(t) := \int_0^\infty p(i)y(t-i)\mathsf{d}i$, where $p(i)$ is the probability that a new secondary infection is the result primary infection which occurred i time units in the past. The reproduction number is then given as the value that equates

$$\mathbb{E}[y(t) \mid y(j),\ j < t] = \mathcal{R}(t)\eta(t) = \mathcal{R}(t) \int_0^\infty p(i)y(t-i)\mathsf{d}i, \tag{1}$$

otherwise known as the renewal equation. The period between primary and secondary infections is exactly the generation time of the disease, but given real data, observed at discrete times (say, daily) this delay distribution must be discretized into contiguous time intervals, say, $(0, 1], (1, 2], \ldots$, which results in the sequence $\{p_i\}_1^\infty$ corresponding to observations y_t and resulting in the discretized version of Equation (1),

$$\mathbb{E}[y_t \mid y_1, \ldots, y_{t-1}] = \mathcal{R}_t\eta_t = \mathcal{R}_t \sum_{i=0}^{\infty} p_i y_{t-i}. \tag{2}$$

Many approaches to estimating $\mathcal{R}_t$ rely on Equation (2) as motivation for their procedures, among them, `EpiEstim` (Cori et al., 2013) and `EpiFilter` (Parag, 2021).

In most cases, it is safe to assume that infectiousness disappears beyond τ timepoints ($p(i) = 0$ for $i > \tau$) so that the truncated integral of the generation interval distribution $\int_0^\tau p(i)\mathtt{d}i = 1$. Generation time, however, is usually unobservable and tricky to estimate, so common practice is to approximate it by the serial interval: the period between the symptom onsets of primary and secondary infections. If the infectiousness profile after symptom onset is independent of the incubation period (the period from the time of infection to the time of symptom onset), then this approximation is justifiable: the serial interval distribution and the generation interval distribution share the same mean. However, other properties may not be similarly shared, and, in general, the generation interval distribution is a convolution of the serial interval distribution with the distribution of the differance between independent draws from the delay distribution from infection to symptom onset. See, for example, Gostic et al. (2020) for a fuller discussion of the dangers of this approximation. Nonetheless, treating these as interchangable is common (Cori et al., 2013) and beyond the scope of this work. Additionally, we assume that the generation interval (and, therefore, the serial interval), is constant over time t. That is, the probability $p(i)$ depends only on the gap between primary and secondary infections and not on the time t when the secondary infection occurs. For our methods, we will assume that the serial interval can be accurately estimated from auxilliary data (say by contact tracing, or previous epidemics) and we will take it as fixed, as is common in existing studies, e.g., Abry et al. (2020); Cori et al. (2013); Pascal et al. (2022).

The renewal equation in Equation (2) relates observable data streams (incident cases) occurring at different time points to the reproduction number given the serial interval. The fact that it depends only on the observed incidence counts makes it reasonable to estimate $\mathcal{R}_t$. However, this relationship obscures some difficulties in data collection. Diagnostic testing targets symptomatic individuals, omitting asymptomatic primary infections which can lead to future secondary infections. Testing practices, availability, and uptake can vary across space and time (Hitchings et al., 2021; Pitzer et al., 2021). Finally, incident cases as reported to public health are subject to delays due to laboratory confirmation, test turnaround times, and eventual submission to public health (Pellis et al., 2021). For these reasons, reported cases are lagging indicators of the course of the pandemic. Furthermore, they do not represent the actual number of new infections that occur on a given day, as indicated by exposure to the pathogen. The assumptions described above (constant serial interval distribution, homogenous mixing, similar susceptibility and social behaviours, etc.) are therefore consequential. That said, Equation (2) also provides some comfort about deviations from these assumptions. If y_t is scaled by a constant (in time) describing the reporting ratio, then it will cancel from

both sides. Similar arguments mean that even if such a scaling varies in time, as long as it varies slowly relative to the set of p_i that are larger than 0, Equation (2) will be a reasonably accurate approximation, so that $\mathcal{R}_t$ can still be estimated well from reported incidence data. Finally, even a sudden change, say from c_1 for $i = 1, \ldots, t_1$ to c_2 for $i > t_1$ would only result in large errors for t in the neighbourhood of t_1 (where the size of this neighbourhood is again determined by the effective support of $\{p_i\}$). This robustness to certain types of data reporting issues provides some degree of comfort when depending on Equation (2) to calculate $\mathcal{R}_t$.

2.2 Poisson trend filtering estimator

We use the daily confirmed incident cases y_t on day t to estimate the observed infectious cases under the model that y_t given previous incident cases $y_{t-1}, \ldots, y_1$ and a constant serial interval distribution follows a Poisson distribution with mean Λ_t. That is,

$$y_t \mid y_1, \ldots, y_{t-1} \sim \text{Poisson}(\Lambda_t), \text{ where } \Lambda_t = \mathcal{R}_t \sum_{i=0}^{t-1} p_i y_{t-i} = \mathcal{R}_t \eta_t.$$

Given a history of n confirmed incidence counts $\mathbf{y} = (y_1, \ldots, y_n)^\mathsf{T}$, our interest is to estimate $\mathcal{R}_t$. A natural approach is to maximize the likelihood, producing the MLE:

$$\begin{aligned} \widehat{\mathcal{R}} &= \underset{\mathcal{R} \in \mathbb{R}_+^n}{\text{argmax}}\ \mathbb{P}(\mathcal{R} \mid \mathbf{y},\ \mathbf{p}) = \underset{\mathcal{R} \in \mathbb{R}_+^n}{\text{argmax}} \prod_{t=1,\ldots,n} \frac{e^{-\mathcal{R}_t \eta_t} (\mathcal{R}_t \eta_t)^{y_t}}{y_t!} \\ &= \underset{\mathcal{R} \in \mathbb{R}_+^n}{\text{argmin}}\ \frac{1}{n} \sum_{t=1}^{n} \mathcal{R}_t \eta_t - y_t \log(\mathcal{R}_t \eta_t). \end{aligned} \tag{3}$$

This optimization problem, however, is easily seen to yield a one-to-one correspondence between the confirmed cases and the effective reproduction, i.e., $\widehat{\mathcal{R}}_t = y_t / \eta_t$, so that the estimated sequence $\widehat{\mathcal{R}}$ will have no significant graphical smoothness.

The MLE is an unbiased estimator of the true parameter $\mathcal{R}_t$, but unfortunately has high variance: changes in y_t result in proportional changes in $\widehat{\mathcal{R}}_t$. To avoid this behaviour, and to match the intuition that $\mathcal{R}_t \approx \mathcal{R}_{t-1}$, we advocate enforcing smoothness of the effective reproduction numbers. This requirement will decrease the variance, and hopefully lead to more accurate estimation of $\mathcal{R}$, as long as the smoothness assumption is reasonable. Smoothness assumptions are common (see e.g., Parag (2021) or Gostic et al. (2020)), but the *type* of smoothness assumed is critical. Cori et al. (2020) imposes smoothness indirectly by estimating $\mathcal{R}_t$ with moving windows of of past observations. The Kalman filter procedure of Parag (2021) would result in ℓ_2-smoothness ($\int_0^n (\widehat{\mathcal{R}}''(t))^2 \mathsf{d}t < C$ for some C), although the algorithm results

in $\widehat{\mathcal{R}}$ taking values over a discrete grid. Pascal et al. (2022) produces piecewise-linear $\widehat{\mathcal{R}}_t$, which turns out to be a special case of our methodology. Smoother estimated curves will provide high-level information about the entire epidemic, obscuring small local changes in $\mathcal{R}(t)$, but may also remove the ability to detect large sudden changes, such as those resulting from lockdowns or other major containment policies.

We choose to implement smoothness by assuming that $\mathcal{R}(t)$ is a piecewise polynomial of arbitrary degree. We specifically consider discrete splines with various degrees of continuity. For example, 0^{th}-degree discrete splines are piecewise constant, the 1^{st}-degree curves are piecewise linear, and 2^{nd}-degree curves are piecewise quadratic. For $k \geq 1$, k^{th}-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 size of changes between adjacent effective reproduction numbers. Because $\mathcal{R}_t > 0$, we explicitly penalize the divided differences (discrete derivatives) of neighbouring values of $\log(\mathcal{R})_t$. To achieve this, we penalize the ℓ_1 norm of the divided differences, 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 created by the squared ℓ_2 norm. Taking different orders of divided differences result in estimated effective reproduction numbers with different smoothness assumptions.

To enforce smoothness of $\hat{\mathcal{R}}_t$, we add a trend filtering penalty to Equation (4) (Kim et al., 2009; Sadhanala et al., 2022; Tibshirani, 2014; Tibshirani et al., 2022). Let $\theta := \log(\mathcal{R}) \in \mathbb{R}^n$, so that $\Lambda_t = \eta_t \exp(\theta_t)$, and $\log(\eta_t \mathcal{R}_t) = \log(\eta_t) + \theta_t$. For evenly spaced incident case data, we write our estimator as the solution to the optimization problem

$$\widehat{\mathcal{R}} = \exp(\widehat{\theta}) \quad \text{where} \quad \widehat{\theta} = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} \; \eta^{\mathsf{T}} \exp(\theta) - \mathbf{y}^{\mathsf{T}} \theta + \lambda \left\| D^{(k+1)} \theta \right\|_1, \tag{4}$$

where $\exp(\cdot)$ applies elementwise. Here, $D^{(k+1)} \in \mathbb{Z}^{(n-k-1)\times n}$ is the $(k+1)^{\text{th}}$ order divided difference matrix for any $0 \leq k < n-1$. $D^{(k+1)}$ is defined recursively as $D^{(k+1)} = D^{(1)} D^{(k)}$, where $D^{(1)} \in \{-1, 0, 1\}^{(n-k-1)\times(n-k)}$ is a banded matrix with diagonal band:

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

The tuning parameter λ balances data fidelity with desired smoothness. When $\lambda = 0$, the problem in Equation (4) reduces to the MLE in Equation (3). Larger tuning parameters privilege the regularization term and yield smoother estimates. Finally, there exists $\lambda_{\max}$ such

that any $\lambda \geq \lambda_{\max}$ will result in $D^{(k+1)}\widehat{\theta} = 0$ and $\widehat{\theta}$ will be the Kullback-Leibler projection of $\mathbf{y}$ onto the null space of $D^{(k+1)}$.

For unevenly spaced incidence data, the spacing between neighboring parameters varies by the time between observations, and thus, the divided differences must be adjusted by the times that the observations occur. Given observation times $\mathbf{x} = (x_1, \ldots, x_n)^\mathsf{T}$, for $k \geq 1$, define a k^{th}-order diagonal matrix

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

Let $D^{(\mathbf{x},1)} := D^{(1)}$. Then for $k \geq 1$, the $(k+1)^{\text{th}}$-order divided difference matrix for unevenly spaced data can be created recursively by

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

Importantly, due to the penalty structure, this estimator is locally adaptive, meaning that it can potentially capture local changes such as the initiation of control measures. Abry et al. (2020); Pascal et al. (2022) considered only the 2^{nd}-order divided difference of $\mathcal{R}_t$ rather than its logarithm. In comparison to their work, our estimator (1) allows for arbitrary degrees of temporal smoothness and (2) avoids the potential numerical issues of penalizing/estimating positive real values. Furthermore, as we will describe below, our procedure is computationally efficient for estimation over an entire sequence of penalty strengths λ and provides methods for choosing how smooth the final estimate should be.

2.3 Numerical optimization of the $\mathcal{R}_t$ estimator

ATTN: To appendix? Almost certainly.

The proximal Newton method is a second-order algorithm solving a proximal 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 (4) solves an approximate problem iteratively — specifically, it takes a second-order Taylor expansion of the Poisson loss, which results in a proximal optimization, i.e., trend filtering with squared ℓ_2 loss, with dynamic weights during iteration, and solves it iteratively until convergence to the objective.

Let $g(\theta) = \eta^\mathsf{T} \exp(\theta) - \mathbf{y}^\mathsf{T}\theta$ be the Poisson loss and $h(\theta) = \lambda \|D^{(k+1)}\theta\|_1$ be the regularization in Equation (4). At iterate $j+1$, consider the following approximation of $g(\theta)$ using the

second-order Taylor expansion around θ^j,

$$g(\theta) = g(\theta^j) + (\theta - \theta^j)^\mathsf{T}\nabla_\theta^{(1)} g(\theta^j) + \frac{1}{2}(\theta - \theta^j)^\mathsf{T}\nabla_\theta^{(2)} g(\theta^j)(\theta - \theta^j),$$

where $\nabla_\theta^{(1)} g(\theta^j) = \frac{1}{n}\eta^\mathsf{T}\exp(\theta^j) - y \in \mathbb{R}^n$ is the gradient of $g(\theta)$ evaluated at θ^j, and $\nabla_\theta^{(2)} g(\theta^j) = \frac{1}{n}\operatorname{diag}(\eta \circ \exp(\theta^j)) \in \mathbb{R}^{n\times n}$ is the Hessian matrix of $g(\theta)$ evaluated at θ^j and $\circ$ means elementwise product.

Define the proximal operator as $\operatorname{prox}_{W,D}(\mathbf{x}) := \underset{\mathbf{z}\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2n}\|\mathbf{z} - \mathbf{x}\|_W^2 + \lambda\|D\theta\|_1$, where $\|\mathbf{a}\|_W^2 := \mathbf{a}^\mathsf{T} W\mathbf{a}$. The proximal optimization problem at iterate $j+1$ given θ^j becomes

$$\begin{aligned}
\theta^{j+} :&= \underset{\theta\in\mathbb{R}^n}{\operatorname{argmin}} \ (\theta - \theta^j)^\mathsf{T}\nabla_\theta^{(1)} g(\theta^j) + \frac{1}{2}(\theta - \theta^j)^\mathsf{T}\nabla_\theta^{(2)} g(\theta^j)(\theta - \theta^j) + h(\theta), \\
&= \underset{\theta\in\mathbb{R}^n}{\operatorname{argmin}} \ \frac{1}{2n}\left\|\theta - \mathbf{c}^j\right\|_{W^j}^2 + \lambda\left\|D^{(k+1)}\theta\right\|_1, \\
&= \operatorname{prox}_{W^j, D^{(k+1)}}(\mathbf{c}^j),
\end{aligned} \tag{5}$$

where $W^j := \operatorname{diag}(\eta \circ \exp(\theta^j))$ is the weighted (Hessian) matrix multiplied by n and

$$\mathbf{c}^j := \theta^j - n\left(W^j\right)^{-1}\nabla_\theta^{(1)} g(\theta^j) = \mathbf{y} \circ \eta^{-1} \circ \exp(-\theta^j) - \mathbf{1} + \theta^j \circ \eta^{-1}.$$

This is just univariate Gaussian trend filtering with weights W^t (Tibshirani, 2014).

We solve the trend filtering problem in Equation (5) 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 $\mathbf{z} := D^{(k)}\theta$. The scaled augmented Lagrangian is

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

where ρ is a scaled dual parameter and $\mathbf{u}$ is the dual variable. At the $(j+1)^{\text{th}}$ Newton step, the specialized ADMM solves the following subproblems, at ADMM iteration $l+1$:

$$\begin{aligned}
\theta^{l+1} &:= \underset{\theta}{\operatorname{argmin}} \ \frac{1}{2n}\left\|\theta - \mathbf{c}^l\right\|_{W^j}^2 + \frac{\rho}{2}\left\|D^{(k+1)}\theta - \mathbf{z}^l + \mathbf{u}^l\right\|_2^2, \\
\mathbf{z}^{l+1} &:= \underset{\mathbf{z}}{\operatorname{argmin}} \ \frac{\lambda}{\rho}\left\|D^{(1)}\mathbf{z}\right\|_1 + \frac{1}{2}\left\|D^{(k+1)}\theta^{l+1} - \mathbf{z} + \mathbf{u}^l\right\|_2^2, \\
\mathbf{u}^{l+1} &\leftarrow \mathbf{u}^l + D^{(k+1)}\theta^{l+1} - \mathbf{z}^{l+1}.
\end{aligned} \tag{6}$$

Finally, the step size $\gamma^{j+1} \in (0, 1]$ at iterate $j+1$ is adjusted by a backtracking line search

algorithm to solve for θ^{j+1}, i.e.,

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

The proximal Newton algorithm iterates until convergence of the objective.

2.4 Solving over a sequence of tuning parameters

ATTN: Need a section about the solving for a sequence of λ.

2.5 Choosing a final λ

ATTN: And one about CV. Mention that other procedures don't choose this.

2.6 Approximate confidence bands

ATTN: And one about confidence bands.

2.7 Bayesian perspective

Unlike many other methods for $\mathcal{R}_t$ estimation, our apporach is frequentist rather than Bayesian. Nonetheless, it can has a corresponding Bayesian interpretation: as a state-space model with Poisson observational noise, autoregressive transition equation of degree $k \geq 0$, e.g., $\theta_{t+1} = 2\theta_t - \theta_{t-1} + \varepsilon_{t+1}$ for $k = 1$, and Laplace transition noise $\varepsilon_{t+1} \sim \text{Laplace}(0,\ 1/\lambda)$. Compared to `EpiFilter` (Parag, 2021), another retrospective study of $\mathcal{R}_t$, we share same observational assumptions, but our approach has a different transition noise. `EpiFilter` estimates the posterior distribution of $\mathcal{R}_t$, and thus it can provide credible interval estimates as well. Our approach produces the maximum *a posteriori* estimate via an efficient convex optimization, omitting the need for MCMC sampling. But the associated confidence bands are approximated differently.

3 Results

Implementation of the our approach is provided in the `R` package `rtestim`.

3.1 Experimental settings

We consider four scenarios of the time-varying effective reproduction numbers to simulate different epidemics. The first two scenarios are simple cases that are rapidly controlled by

intervention, where the graphical curves consist of one knot and two segments. Scenario 1 is instantaneous prior and post-intervention, and Scenario 2 is exponentially grow and decay. The last two scenarios are more complicated, where more waves in the epidemics are involved. Scenario 3 has four linear segments with three knots, which reflect the effect of intervention, the resurgence to large epidemics, and the suppression of pandemic respectively. Scenario 4 involves more complicated waves and curvatures of the epidemic. Effective reproduction numbers across all scenarios are evenly spaced. The first three scenarios and the last scenario are motivated by Gressani et al. (2022); Parag (2021) respectively. We name the four scenarios as (1) 2-segment constant line, (2) 2-segment exponential curve, (3) 4-segment linear line, and (4) periodic curve respectively.

We consider epidemics of length $n = 300$. Specifically, in Scenario 1, $\mathcal{R}_t = 2, 0.8$ before and after $t = 70$. In Scenario 2, $\mathcal{R}_t$ increases and decreases exponentially with rates $0.015, 0.005$ pre and post $t = 50$. In Scenario 3, $\mathcal{R}_t$ reduces from 2.5 to 2 linearly between $t \in [1, 60]$, falls to 0.8 at $t = 61$ and goes linearly down to 0.6 until $t = 110$, resurges to 1.7 at $t = 111$ and grows linearly back to 2 until $t = 150$, and then drops to 0.9 at $t = 151$ and descends to 0.5 until the end. In Scenario 4, $\mathcal{R}_t$ is a continuous, periodic curve generated by the function

$$f(x) = 0.2\left(\left(\sin(\frac{\pi x}{12}) + 1\right) + \left(2\sin\left(\frac{\pi x}{6}\right) + 2\right) + \left(3\sin(\frac{\pi x}{1.2}) + 3\right)\right)$$

at equally spaced points $x \in [0, 10]$.

We assume that the serial interval follows Gamma distribution with fixed shapes and scales $(3, 3)$, $(2.5, 2.5)$, $(3.5, 3.5)$ and $(3.5, 3.5)$ for Scenarios $1 - 4$ respectively. We consider all epidemics starting from $N_1 = 2$ incidences and generating until timepoints $t = 300$. We compute the expected incidence N_t use renewal equation, and generate the incidence samples from the Poisson distribution $y_t \sim Pois(N_t)$. To verify the performance of our model under the violation of distributional assumption of incidence, we generate incidence samples using negative Binomial distribution with dispersion size 5, i.e., $y_t \sim NB(N_t, size = 5)$. We generate 50 random samples for each setting of experiments. It results in 4 $\mathcal{R}_t$ scenarios $\times 2$ Incidences distributions $\times 50$ random samples $= 400$ problems in total. Examples of each effective reproduction number scenario with corresponding Poisson and negative Binomial incidences are displayed in Figure 1.

We compare our `RtEstim` to `EpiEstim` and `EpiLPS`. `EpiEstim` is a widely used Bayesian method that estimates the posterior distribution of effective reproduction numbers given the Gamma prior and Poisson distributed incidences. They estimate the reproduction number over a sliding window by assuming the reproduction number is constant during the specific time window. A longer sliding window averages out more fluctuations and noises, and leads

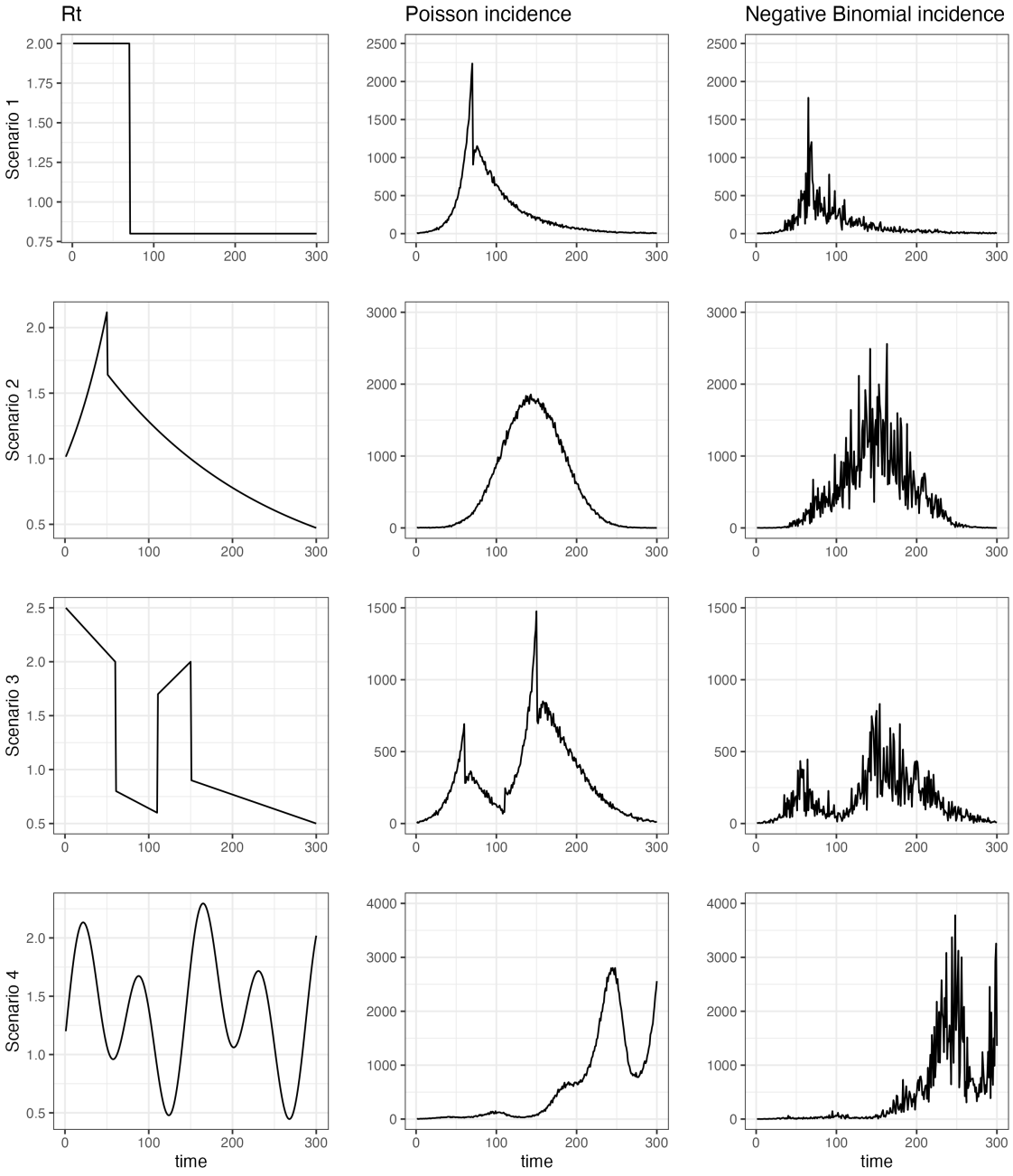

Figure 1: An example of effective reproduction numbers and corresponding incidences following Poisson or negative Binomial distribution. Three columns illustrate four $\mathcal{R}_t$ cases, Poisson incidences, and negative Binomial distributed incidences for each $\mathcal{R}_t$ case respectively. Four rows correspond to four $\mathcal{R}_t$ cases respectively. *Sample ranges vary significantly from less than* $1,000$ *to more than* $10,000$ *by random seeds.*

to smoother estimation over time; whereas, a shorter sliding window is more responsive to sudden spikes or declines in a shorter period. We use the default weekly sliding window for the experimental study. Monthly sliding window is also considered in simulation. Since neither of the two sliding windows considerably outperforms the other across all scenarios, we defer the estimation of monthly sliding window to the supplementary document. `EpiLPS` is another Bayesian approach that estimates P-splines coupled with Laplace approximations of the conditional posterior of the spline vector based on negative Binomial distributed incidences. We tune the model over the candidate set of size 50 using cross validation. For four scenarios, we estimate piecewise constant $k = 0$, piecewise linear & cubic $k = 1, 3$, piecewise linear $k = 1$ and piecewise cubic polynomials $k = 3$ respectively. We assume the serial intervals are known and use same serial intervals across all models for each problem.

Kullback-Leibler (KL) divergence is a metric that measures the distance between two probability distributions. Since $\mathcal{R}_t$ can be regarded as the expectation of Poisson distribution, we use the mean KL divergence for Poisson distributions (averaged across all coordinates) to measure the accuracy of the $\mathcal{R}_t$ estimates:

$$\frac{1}{n} D_{KL}(\hat{\mathcal{R}} || \mathcal{R}) = \frac{1}{n} \sum_{t=1}^{n} \hat{\mathcal{R}}_t \log \left(\frac{\hat{\mathcal{R}}_t}{\mathcal{R}_t} \right) + \mathcal{R}_t - \hat{\mathcal{R}}_t,$$

where $\mathcal{R} := \{\mathcal{R}_t\}_{t=1}^{n}$. In comparison of the accuracy across methods, we drop the estimates during the first week as the $\mathcal{R}_t$ estimates of `EpiEstim` starts at $t = 8$. The Euclidean (ℓ_2) norm is less appropriate here since we assume Poisson distributed incidences, but the Euclidean norm corresponds to Gaussian data in terms of the log-likelihood functions. Other details of the experimental settings are deferred to the supplementary document.

We run leave-third-out cross validation (CV) to choose the best tuning parameter from the candidate set of size 50, i.e., $\boldsymbol{\lambda} = \{\lambda_1, \cdots, \lambda_{50}\}$. Specifically, we divide the all samples into three folds and build models on each sample set which excludes one fold of the samples across all hyperparameters. Every third samples are placed into the same fold by excluding the first and last samples. We select the tuning parameter that gives the lowest averaged mean squared errors (MSEs) of the estimated reproduction numbers from the observed samples across all folds. The experiments are run in `R` with version 4.3.1 on Cedar cluster provided by Compute Canada. The `R` packages used for simulation and real-data application are of versions `EpiEstim_2.2-4`, `EpiLPS_1.2.0`, and `rtestim_0.0.3`.

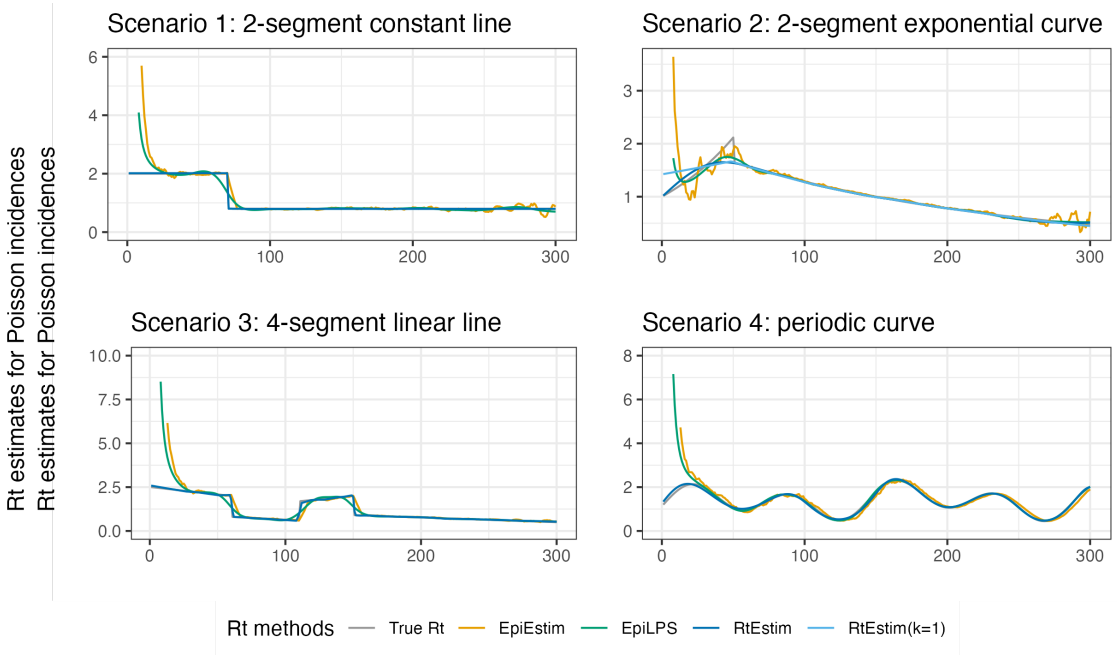

Figure 2: Effective reproduction number estimation for Poisson incidences. `RtEstim(k=1)` is the alternative estimation using $k = 1$ in Scenario 2.

3.2 Experimental results

Figure 2 illustrates the estimated reproduction numbers by three models for the Poisson incidence cases. Compared to `EpiEstim` and `EpiLPS`, which have an edge problem at the beginning of the time series, our `RtEstim` estimates are more accurate — almost overlap with the true values — without suffering from the edge problem. Scenario 2 is a difficult problem for all methods; the sharp increase at the end of the stage of exponential growth is hard to capture for all models. Although the cubic curve is a more desirable choice since the true curvature is exponential, the graphical curvature is so gentle that linear estimation also seems a reasonable choice. We, therefore, fit piecewise linear and cubic $\hat{\mathcal{R}}_t$ curves using `RtEstim` for Scenario 2 in case of model misspecification. However, both estimations fail to recover the acute rise in the growing stage. An explanation of such failure is that the estimated curves are both continuous at the knot, which hinders the estimates from fitting two discontinuous phases of steep increase followed by a sudden decline. Scenario 1 is the simplest case with only one knot and two constant segments. Besides the edge problem, `EpiEstim` and `EpiLPS` produce "smooth" estimated curves that are continuous at the knot, which results in divergence from the true values in the first segment in Scenario 1. Since the piecewise constant `RtEstim` estimator does not require the smoothness in $\mathcal{R}_t$, it captures the sharp decrease in Scenario 1.

To investigate the performance under the violation of distributional assumption of in-

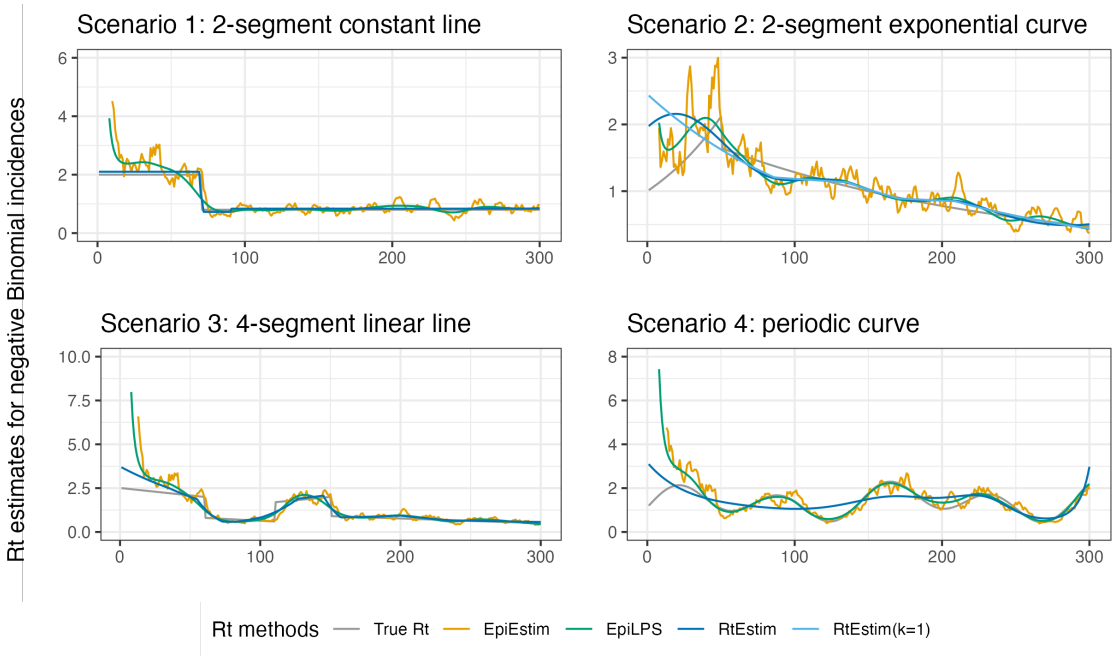

Figure 3: Effective reproduction number estimation for negative Binomial incidences. `RtEstim(k=1)` is the alternative estimation using $k = 1$ in Scenario 2.

cidences (of `RtEstim` and `EpiEstim`), we estimate $\mathcal{R}_t$s using negative Binomial incidences. Figure 3 displays the estimates across all methods. `RtEstim` estimates overall do not perform as remarkably accurate as in the Poisson incidence cases in the beginning. Especially for Scenario 4, `RtEstim` fails to recover the wiggly curvature except the last wave. In Scenario 2, `RtEstim` succeeds to capture the knot, but suffers from the same problem as in the Poisson cases. In Scenario 3, piecewise linear `RtEstim` estimates overall recover the curvature of $\mathcal{R}_t$, but is not as accurate as in the Poisson incidence cases.

`RtEstim` overall outperforms `EpiEstim` and `EpiLPS` in the experimental study. Figure 4 visualizes the KL divergences across three models in boxplots. Given both Poisson and negative Binomial incidences, `RtEstim` is more accurate than `EpiEstim` and `EpiLPS` across Scenarios 1, 3, 4 in terms of the medians of KL scores and the boxes are always lower than counterparts of the other two methods. The advantage is narrowed for negative Binomial cases compared to the Poisson cases. `EpiLPS` wins in Scenario 2 as the boxes of KL scores are lower than counterparts of the other two methods for both incidence distributions, but `EpiLPS` has large outliers for negative Binomial incidences cases.

We also compare the running times of three models across 8 experimental settings. We find that almost all models across all experiments takes within 10 seconds to converge. Our model runs longer generally, which is likely due to a relatively large candidate set (of size 50), other models only run a single time for a fixed setting of hyperparameters per experiment.

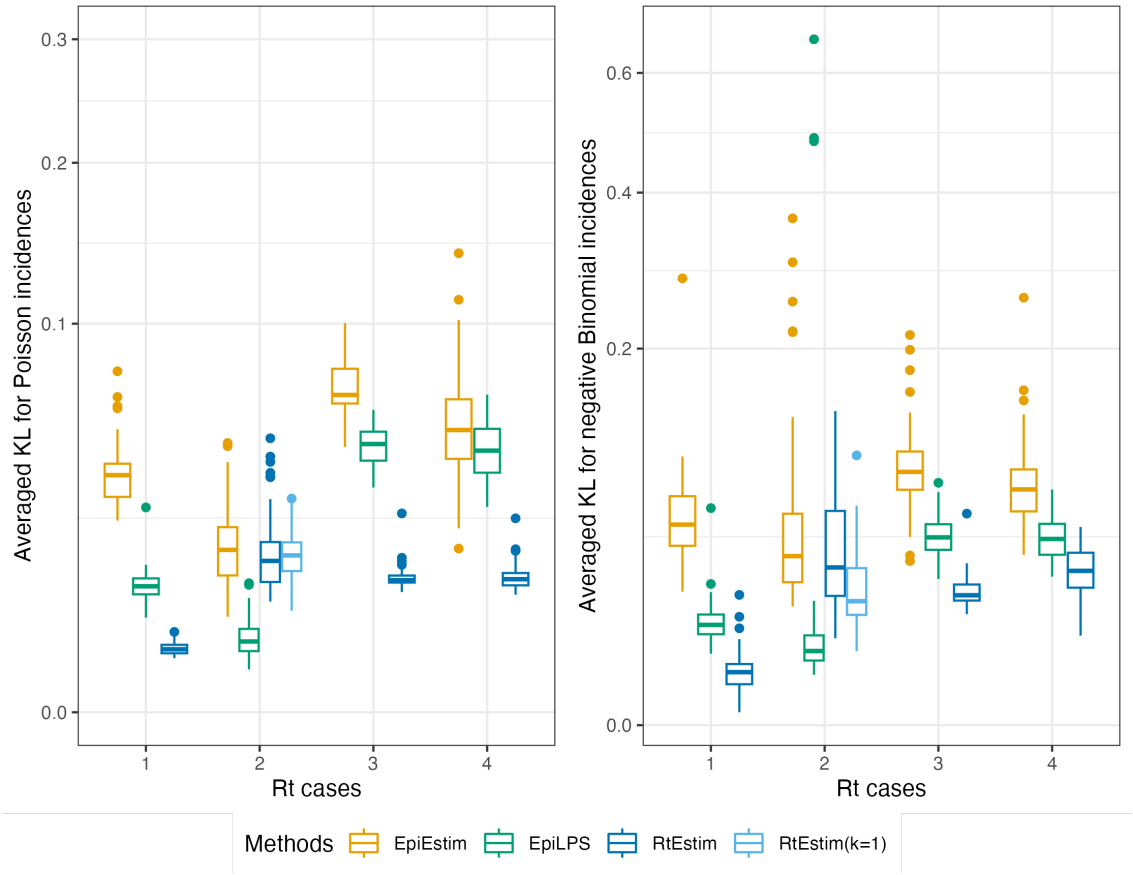

Figure 4: Boxplot of Kullback-Leibler divergence between the estimated effective reproduction numbers and the true ones across all methods given Poisson incidences and negative Binomial incidences across 50 samples. Left panel visualizes the KL divergences for the Poisson incidence cases. Right panel displays the KL divergences for the negative Binomial incidence cases.

More experimental results on time comparisons are deferred to the supplementary document.

3.3 Covid-19 incidences in British Columbia

We implement our `RtEstim` on the Covid-19 confirmed cases in British Columbia (B.C.) as of May 18, 2023 (visualized in Figure 5) reported by B.C. Centre for Disease Control. The reported incidence cases provide a snapshot of how testing recommendations and practices have changed over the three years. We choose the gamma distribution with shape 2.5 and scale 2.5 to approximate the serial interval function, which is empirically found to be a reasonable choice.

Considering the temporal evolutions of effective reproduction numbers across $3, 4, 5$ days, the estimated reproduction numbers of Covid-19 in British Columbia (illustrated in Figure 6) are less than 3 during most of the time, which means that one distinct infected individuals

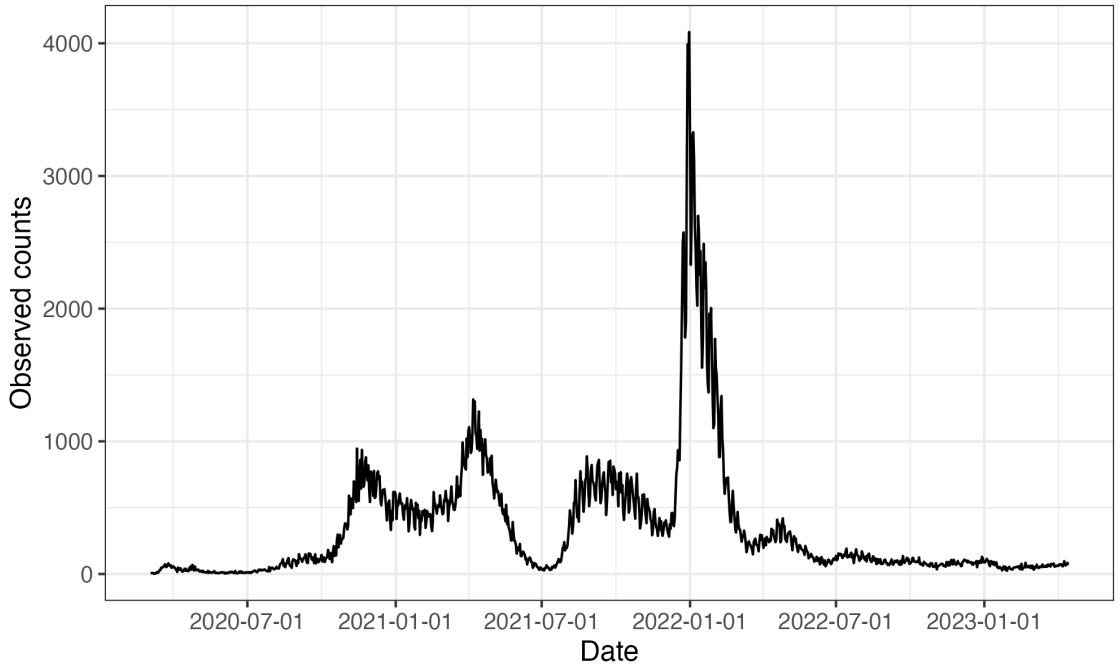

Figure 5: Covid19 daily confirmed incidence counts between March 1st, 2020 and April 15th, 2023 in British Columbia, Canada.

can on average infect less than three other individuals in the population. The three degrees of the temporal evolution (across all regularization levels λ) all yield similar results that $\hat{\mathcal{R}}_t$ comes to the highest peak around the end of 2021 and then drops down to 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. We also visualize the 95% empirical confidence bands of the point estimates for the "best" tuning parameter (in terms of MSEs).

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 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 (fluctuating around 1) and the infected cases stay low.

Greater regularization levels (by using larger λs) result in smoother estimated curves. Smoother curves suggest that the estimated reproduction numbers are around 1 during most

time periods; however, it may miss to capture some outbreaks of the pandemic. 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. Here, we provide the CV-chosen $\mathcal{R}_t$ estimates with confidence bands.

3.4 Pandemic influenza in Baltimore, Maryland, 1918

We then apply `RtEstim` on the pandemic influenza in Baltimore, Maryland, 1918. Dataset in Figure 7 is obtained from the `R` package `EpiEstim`. The 1918 influenza, caused by H1N1 influenza A virus, was an unprecedentedly deadly influenza that infected almost one-third of the population across the world (Taubenberger and Morens, 2006). In the estimation displayed in Figure 8, the CV-tuned piecewise cubic estimates better capture the growing tendency at the beginning of the pandemic. It suggests that the pandemic has yielded a decrease after around 20 days and reached 1 when the pandemic has lasted for nearly 50 days. However, it also suggests an increase at the end of the period, while a steady decline (as in CV-tuned piecewise constant and linear estimates) is more reasonable. The smoothness of $\mathcal{R}_t$ curves should be chosen based on the purpose of the study in practice, e.g., epidemic forecasting may require a more wiggly curve that contains more fluctuation information, while retrospective studies that solely target on understanding of the pandemic may prefer a smoother curve with less important information smoothed out.

4 Discussion

`RtEstim` 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 a 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 method `RtEstim` 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 $S \to I \to S$. 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 later. 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 or without a specified level of overdispersion. We also

assume the confirmed incidence is observable to a smoothly time-varying proportion over all incidences. If the observable proportion is constant over time, our effective reproduction numbers will be able to reflect the "true" $\mathcal{R}_t$ values; however, if the proportion varies by time (which is almost true in reality due to the change of surveillance resources and testing recommendations), the estimates may deviate from the "true" values.

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 due to the change of population factors such as herd immunity. Another common statement 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. A future adjustment on our model is to consider the delay distributions. We approximate the generation time (period from primary infection to secondary infection) using serial interval (period from primary onset to secondary onset) directly without measuring the uncertainty of the report delays. As a further adjustment, we may add another step from infection trajectory to reported case counts to measure the distribution of incubation period (period from infection to onset) and the report delay (period from onset to case report). Besides this, instead of assuming a fixed distribution, a more realistic setting may be to use time-varying parameters of these distributions. We may also consider other case counts for effective reproduction number estimation if the incidence data is unavailable or not of a good quality. We may use death counts instead, for example, with a corresponding distribution to measure the period between primary and secondary onsets of deaths to displace the current serial interval distribution.

References

Sam Abbott, Joel Hellewell, Katharine Sherratt, Katelyn Gostic, Joe Hickson, Hamada S. Badr, Michael DeWitt, Robin Thompson, EpiForecasts, and Sebastian Funk. *EpiNow2: Estimate Real-Time Case Counts and Time-Varying Epidemiological Parameters*, 2020a.

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, 2020b.

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.

Anne Cori, Simon Cauchemez, Neil M Ferguson, Christophe Fraser, Elisabeth Dahlqwist, P Alex Demarsh, Thibaut Jombart, Zhian N Kamvar, Justin Lessler, Shikun Li, et al. Package 'epiestim'. *CRAN: Vienna Austria*, 2020.

Katelyn M Gostic, Lauren McGough, Edward B Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A Hay, Pablo M De Salazar, et al. Practical considerations for measuring the effective reproductive number, r_t. *PLoS computational biology*, 16(12):e1008409, 2020.

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.

Matt DT Hitchings, Natalie E Dean, Bernardo García-Carreras, Thomas J Hladish, Angkana T Huang, Bingyi Yang, and Derek AT Cummings. The usefulness of the test-positive proportion of severe acute respiratory syndrome coronavirus 2 as a surveillance tool. *American journal of epidemiology*, 190(7):1396–1405, 2021.

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.

Lorenzo Pellis, Francesca Scarabel, Helena B Stage, Christopher E Overton, Lauren HK Chappell, Elizabeth Fearon, Emma Bennett, Katrina A Lythgoe, Thomas A House, Ian Hall, et al. Challenges in control of covid-19: short doubling time and long delay to effect of interventions. *Philosophical Transactions of the Royal Society B*, 376(1829):20200264, 2021.

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.

Virginia E Pitzer, Melanie Chitwood, Joshua Havumaki, Nicolas A Menzies, Stephanie Perniciaro, Joshua L Warren, Daniel M Weinberger, and Ted Cohen. The impact of changes

in diagnostic testing practices on estimates of covid-19 transmission in the united states. *American Journal of Epidemiology*, 190(9):1908–1917, 2021.

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.

Veeranjaneyulu Sadhanala, Robert Bassett, James Sharpnack, and Daniel J McDonald. Exponential family trend filtering on lattices. *arXiv preprint arXiv:2209.09175*, 2022.

Jeffery K Taubenberger and David M Morens. 1918 influenza: the mother of all pandemics. *Revista Biomedica*, 17(1):69–79, 2006.

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.

Ryan J Tibshirani et al. Divided differences, falling factorials, and discrete splines: Another look at trend filtering and related problems. *Foundations and Trends® in Machine Learning*, 15(6):694–846, 2022.

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.

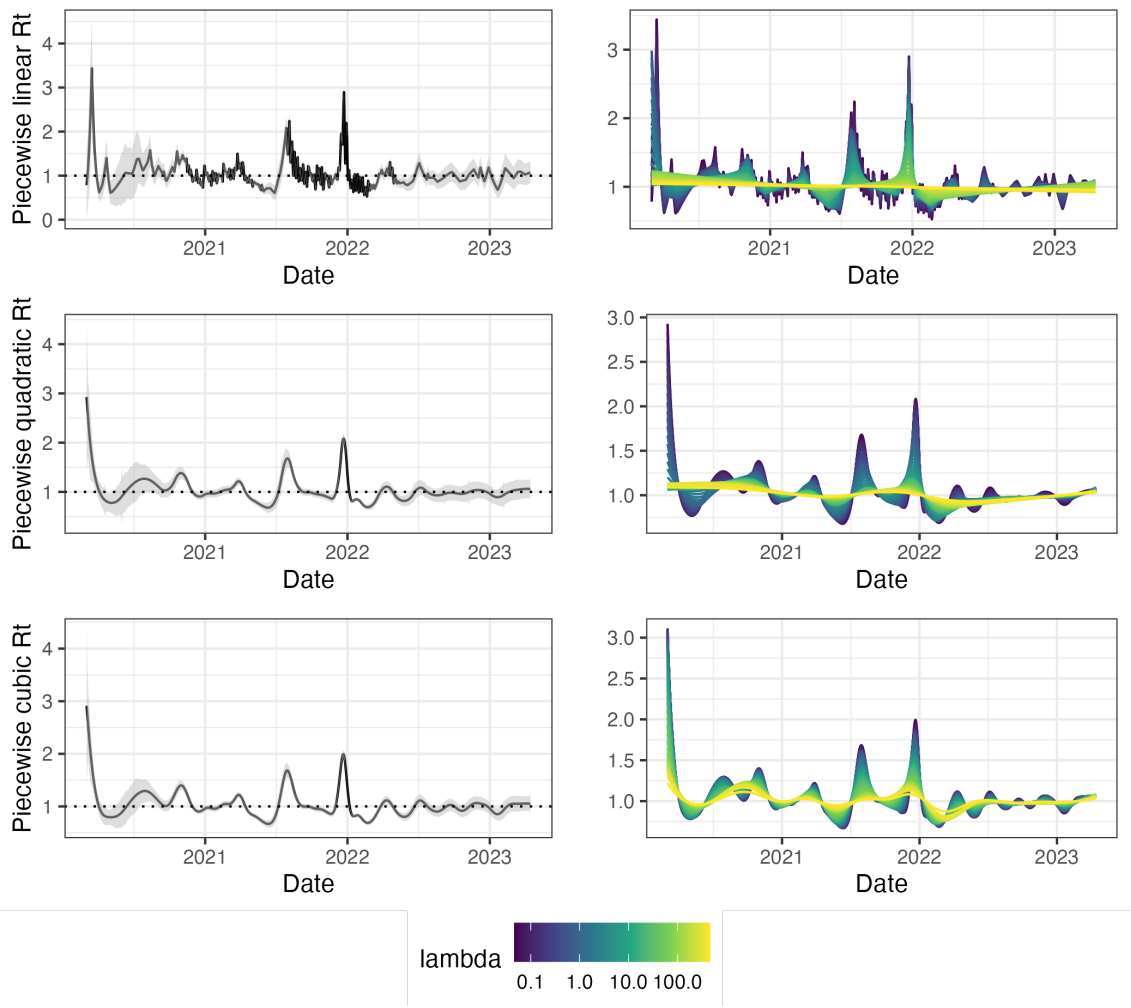

Figure 6: Estimated effective reproduction numbers for Covid19 daily confirmed counts between March 1st, 2020 and April 15th, 2023 in British Columbia, Canada. The left panels display the CV-tuned estimates with 95% confidence intervals. The right panels demonstrate estimates corresponding to 50 tuning parameters. The top, medium and bottom panels illustrate the estimated reproduction numbers ($\mathcal{R}_t$) using the Poisson trend filtering (in Equation (4)) with degrees $k = 1, 2, 3$ respectively.

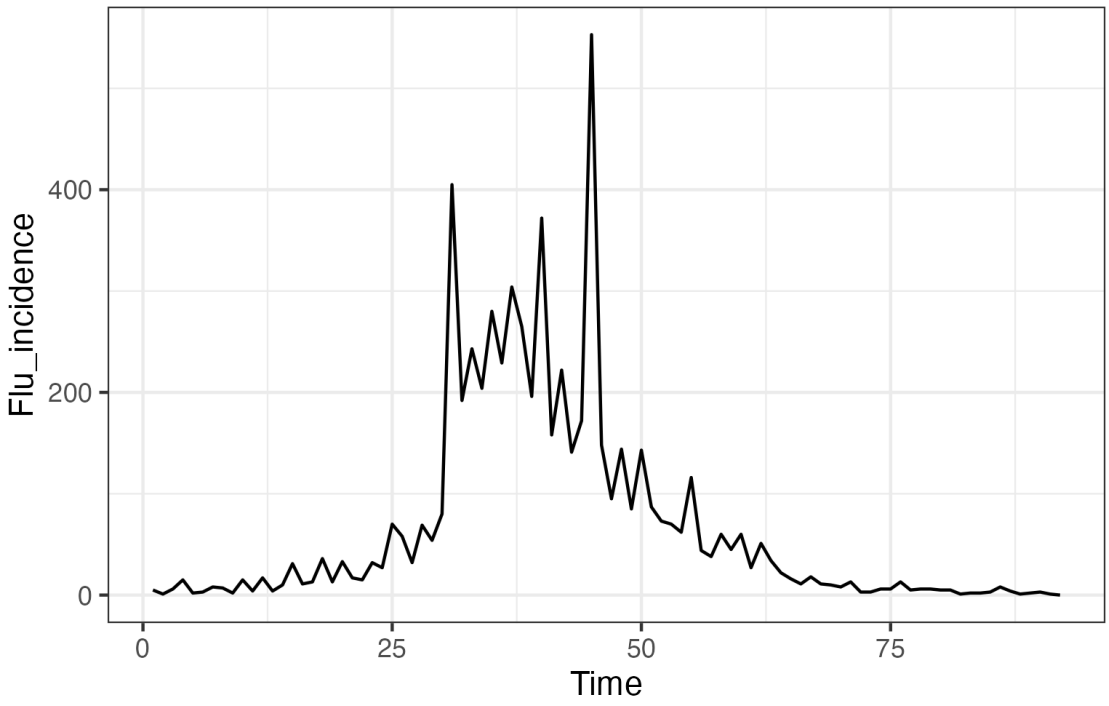

Figure 7: Pandemic influenza incidence counts in Baltimore, Maryland in 1918.

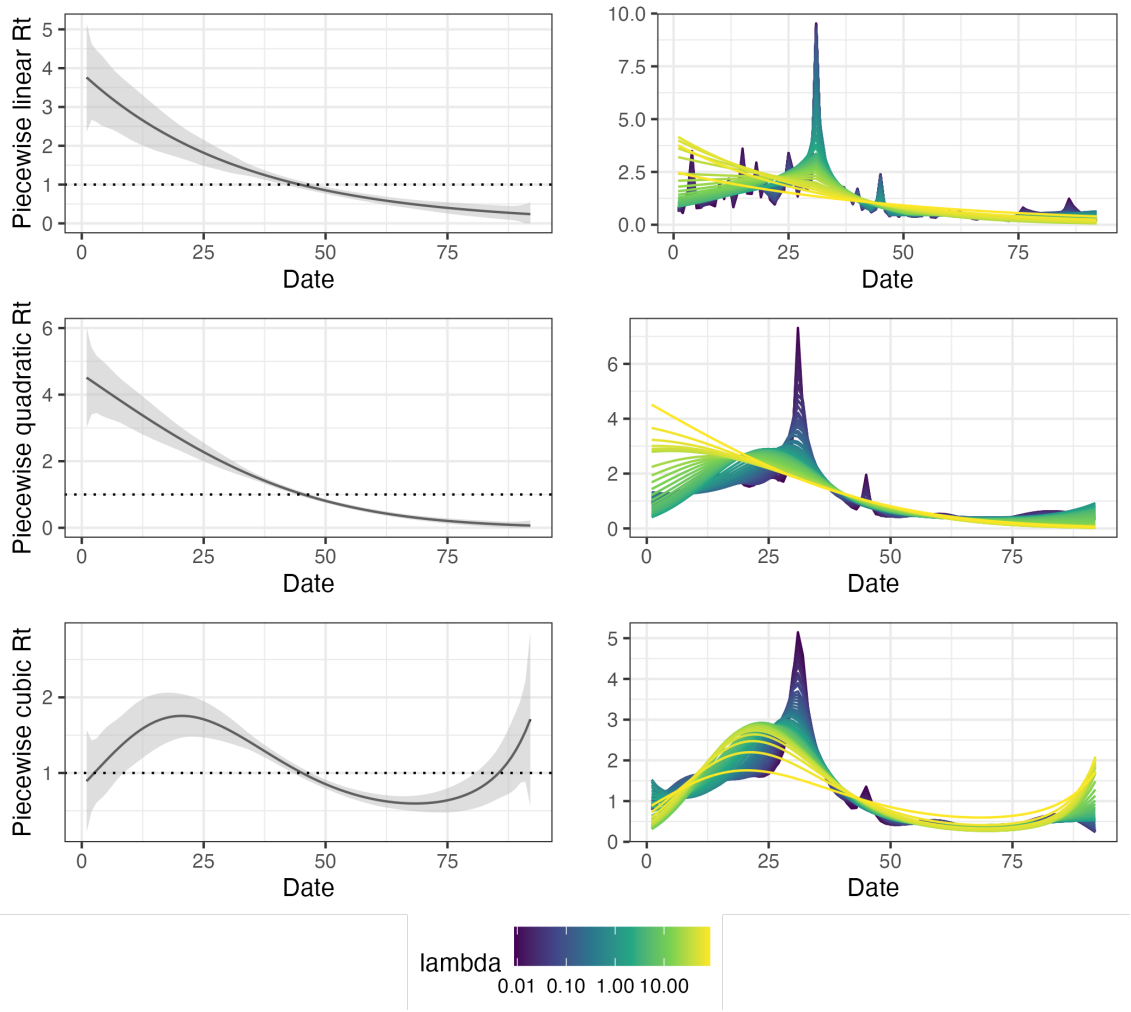

Figure 8: Estimated effective reproduction numbers for pandemic influenza incidence counts in Baltimore, Maryland in 1918. The left panels display the CV-tuned estimates with 95% confidence intervals. The right panels demonstrate estimates corresponding to 50 tuning parameters. The top, medium and bottom panels illustrate the estimated reproduction numbers ($\mathcal{R}_t$) using the Poisson trend filtering (in Equation (4)) with degrees $k = 1, 2, 3$ respectively.