medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Bayesian nowcasting with Laplacian-P-splines

Bryan Sumalinab*[1,2], Oswaldo Gressani[1], Niel Hens[1,3], Christel Faes[1]

[1]Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat), Data Science Institute (DSI), Hasselt University, Hasselt, Belgium

[2]Department of Mathematics and Statistics, College of Science and Mathematics, Mindanao State University - Iligan Institute of Technology, Iligan City, Philippines

[3]Centre for Health Economic Research and Modelling Infectious Diseases (CHERMID), Vaccine & Infectious Disease Institute, Antwerp University, Antwerp, Belgium

Abstract

During an epidemic, the daily number of reported (infected/death) cases is often lower than the actual number of cases due to underreporting. Nowcasting aims to estimate the cases that have not yet been reported and combine it with the already reported cases to obtain an estimate of the daily cases. In this paper, we present a fast and flexible Bayesian approach to nowcasting combining P-splines and Laplace approximations. The main benefit of Laplacian-P-splines (LPS) is the flexibility and faster computation time compared to Markov chain Monte Carlo (MCMC) algorithms that are often used for Bayesian inference. In addition, it is natural to quantify the prediction uncertainty with LPS in the Bayesian framework, and hence prediction intervals are easily obtained. Model performance is assessed through simulations, and the method is applied to the Belgian COVID-19 mortality cases for the year 2021. Simulation results show that our model has good predictive performance except when the nowcast date is near the peak date, where it has lower prediction interval coverage.

Keywords: Nowcasting, Laplacian-P-splines, Epidemic, COVID-19, Reporting delay.

*Corresponding author. *E-mail address*: bryan.sumalinab@uhasselt.be

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

1 Introduction

Nowcasting is a term used for estimating the occurred-but-not-yet-reported-events (Donker et al., 2011; Van de Kassteele et al., 2019). In epidemiology, real-time updates of new symptomatic/infected individuals are helpful to assess the present situation and provide recommendations for rapid planning and for implementing essential measures to contain an epidemic outbreak. The exact number of new daily cases is frequently subject to reporting delays, resulting in underreporting of the real number of infected individuals for that day. Failing to account for the reporting delays will lead to possibly biased predictions that might have an effect on policy making (Gutierrez et al., 2020). The main goal of nowcasting is to estimate the actual number of new cases by combining the (predicted) not-yet-reported-cases with the already reported cases.

Several early references that establish the statistical framework for this type of problem can be found in the paper of Lawless (1994) which demonstrates how nowcasting can be used not only in disease surveillance but also in other contexts such as warranty and insurance claims. More recently, Höhle and an der Heiden (2014) applied nowcasting to the outbreak of Shiga toxin-producing Escherichia coli in Germany and also to the SARS-CoV-2 outbreak (Glöckner et al., 2020; Günther et al., 2021). Their approach is formulated within a hierarchical Bayesian framework that consists of estimating the epidemic curve by using a quadratic spline based on a truncated power basis function and the time-varying reporting delay distribution is approximated by a discrete time survival model. Van de Kassteele et al. (2019) pointed out that a potential drawback of such a method is the long computation time required by the Markov chain Monte Carlo (MCMC) algorithm and therefore proposed an alternative fast and flexible modelling strategy based on bivariate P-splines (Penalized B-splines). P-splines (Eilers and Marx, 1996) provide a flexible smoothing tool used to describe trends in the data. It introduces a penalty parameter that controls the roughness of the fit and counterbalances the flexibility of a rich B-splines basis (Eilers and Marx, 1996; Eilers et al., 2015). Other attractive features of the P-splines smoother are the relatively simple structure of the penalty matrix that is effortlessly

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

computed and a natural extension to the Bayesian framework (Lang and Brezger, 2004). Based on the approach of Van de Kassteele et al. (2019), the number of cases are structured in a two-dimensional table (with calendar time as the first dimension and delay time as the second dimension), yielding the data matrix used as an input in the model. The reporting intensity is assumed to be a smooth surface and is modelled using two-dimensional P-splines.

In this paper, we extend the method of Van de Kassteele et al. (2019) by proposing a new nowcasting methodology based on Laplacian-P-splines (LPS) in a fully Bayesian framework. A key advantage of working with the Bayesian approach is the ease to obtain the predictive distribution and quantify the uncertainty associated with the predictions. In addition, the posterior distribution of the penalty parameter can be explored, and hence its uncertainty can be accounted for. The Laplace approximation uses a second-order Taylor expansion to approximate the posterior distribution of the regression parameters by a Gaussian density. It is a sampling-free method with the major advantage of faster computational time as opposed to MCMC approaches that are commonly used in Bayesian inference. Therefore, given the flexibility of previously mentioned P-splines smoothers and the computational benefit of Laplace approximations, it can be a helpful tool in the daily monitoring of new cases during an epidemic period. Laplacian-P-splines already proved to be useful in survival models (Gressani and Lambert, 2018; Gressani et al., 2022), generalized additive models (Gressani and Lambert, 2021) and also in epidemic models for estimating the effective reproduction number (Gressani et al., 2021). We build on the work of Gressani and Lambert (2021) to extend the Laplacian-P-splines methodology to nowcasting, thereby providing a fast and flexible (fully) Bayesian alternative to Van de Kassteele et al. (2019). To evaluate the (predictive) performance of our method, a simulation study is implemented and several performance measures are reported such as the mean absolute percentage error (MAPE), symmetric mean absolute percentage error (SMAPE) and prediction interval coverage. Finally, we apply our method to the COVID-19 mortality data in Belgium for the year 2021. The data and R codes used to implement the method are available on GitHub through this link https://github.com/bryansumalinab/Laplacian-P-spline-nowcasting.git.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2 Methodology

2.1 Bayesian model formulation

In this section, we build upon the work of Van de Kassteele et al. (2019) to formulate a fully Bayesian model based on P-splines. Let $y_{t,d}$ denote the number of cases that occurred at time $t = 1, 2, ..., T$ (corresponding to the calendar day) and reported with a delay of $d = 0, 1, 2..., D$ days. The information on cases can be summarized in matrix form:

$$Y = \begin{bmatrix} y_{1,0} & y_{1,1} & y_{1,2} & \cdots & y_{1,D} \\ y_{2,0} & y_{2,1} & y_{2,2} & \cdots & y_{2,D} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y_{T-(D-1),0} & y_{T-(D-1),1} & y_{T-(D-1),2} & \cdots & \boldsymbol{y_{T-(D-1),D}} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y_{T-1,0} & y_{T-1,1} & \boldsymbol{y_{T-1,2}} & \cdots & \boldsymbol{y_{T-1,D}} \\ y_{T,0} & \boldsymbol{y_{T,1}} & \boldsymbol{y_{T,2}} & \cdots & \boldsymbol{y_{T,D}} \end{bmatrix},$$

with cases that have not yet been reported (at time T) highlighted in bold. The not-yet-reported cases correspond to (t, d) combinations satisfying $t > T - d$. The main objective is to predict the total number of cases, $y_t = \sum_{d=0}^{D} y_{t,d}$, for $t = T - (D - 1), \ldots, T$ for which the nowcasted and already reported cases can be combined.

Let $\mathcal{D} := \boldsymbol{y} = (y_1, y_2 \ldots, y_n)'$ denote the vector of the observed number of cases by stacking the columns of matrix Y for the reported cases, where each entry corresponds to each (t, d) combination of reported cases $y_{t,d}$. The model assumes that the number of cases are Poisson distributed, i.e., $y_{t,d} \sim \text{Poisson}(\mu_{t,d})$ with mean $\mu_{t,d} > 0$. Following Van de Kassteele et al. (2019), the (log) mean number of cases is modeled with two dimensional B-splines:

$$\log(\mu_{t,d}) = \beta_0 + \sum_{j=1}^{K_T} \sum_{k=1}^{K_D} \theta_{j,k} b_j(t) b_k(d) + \sum_{l=1}^{p} \beta_l z_l(t, d),$$

where β_0 is the intercept; $b_j(\cdot)$ and $b_k(\cdot)$ are univariate B-splines basis functions specified in the time and delay dimensions, respectively; and $z_l(t, d)$ represents additional covariates with

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

regression coefficients β_l. In matrix notation:

$$\log(\boldsymbol{\mu}) = B\boldsymbol{\theta} + Z\boldsymbol{\beta}, \tag{1}$$

where the matrices B and Z correspond to the basis functions and covariates, respectively, and vectors $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are the associated parameters to be estimated (Details in Appendix A1).

In the philosophy of P-splines (Eilers and Marx, 1996), we use a rich (cubic) B-splines basis and counterbalance the associated flexibility by imposing a discrete roughness penalty on contiguous B-spline coefficients. For the two dimensional P-splines, the penalty can be obtained based on row-wise (direction of calendar time) and column-wise (direction of reporting delay) differences for matrix $\Theta = (\theta_{j,k})$ with $j = 1, \ldots, K_T$ and $k = 1, \ldots, K_D$ (Appendix A1) (see Durbán et al. (2002) and Fahrmeir et al. (2013) (pp. 507-508)). Let D_t^m and D_d^m denote the mth order row-wise and column-wise difference matrix with dimensions $(K_T - m) \times K_T$ and $(K_D - m) \times K_D$, respectively. In this paper, we use a second order ($m = 2$) difference penalty. For ease of notation, let $D_t = D_t^m$ and $D_d = D_d^m$. The difference matrix for vector $\boldsymbol{\theta}$ can be obtained by expanding the difference matrix into $I_{K_D} \otimes D_t$ and $D_d \otimes I_{K_T}$, where $\otimes$ denotes the Kronecker product. Using this notation, the row-wise and column-wise difference penalty can be written, respectively, as:

$$\|(I_{K_D} \otimes D_t)\boldsymbol{\theta}\|^2 = \boldsymbol{\theta}'(I_{K_D} \otimes D_t)'(I_{K_D} \otimes D_t)\boldsymbol{\theta} = \boldsymbol{\theta}'(I_{K_D} \otimes D_t'D_t)\boldsymbol{\theta} \quad \text{and}$$
$$\|(D_d \otimes I_{K_T})\boldsymbol{\theta}\|^2 = \boldsymbol{\theta}'(D_d \otimes I_{K_T})'(D_d \otimes I_{K_T})\boldsymbol{\theta} = \boldsymbol{\theta}'(D_d'D_d \otimes I_{K_T})\boldsymbol{\theta}.$$

Let $\lambda_t > 0$ and $\lambda_d > 0$ denote the row-wise and column-wise penalty parameter, respectively. The penalty for the two dimensional B-splines is then given by:

$$\begin{aligned}&\lambda_t\boldsymbol{\theta}'(I_{K_D} \otimes D_t'D_t)\boldsymbol{\theta} + \lambda_d\boldsymbol{\theta}'(D_d'D_d \otimes I_{K_T})\boldsymbol{\theta}\\ &= \boldsymbol{\theta}'\Big(\lambda_t(I_{K_D} \otimes D_t'D_t) + \lambda_d(D_d'D_d \otimes I_{K_T})\Big)\boldsymbol{\theta}.\end{aligned}$$

Let us define the penalty matrices $P_t = D_t'D_t + \delta I_{K_T}$ and $P_d = D_d'D_d + \delta I_{K_D}$, where δ is a small

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

number (say $\delta = 10^{-6}$), to ensure that the penalty matrices are full rank and thus invertible. Following Lang and Brezger (2004), the penalty can be translated in the Bayesian framework by specifying a Gaussian prior for each column and row vector of Θ. We write Θ in terms of its columns and rows:

$$\Theta = (\boldsymbol{\theta}_1^c, \ldots, \boldsymbol{\theta}_{K_D}^c) \qquad \text{(columns of } \Theta) \qquad \text{and}$$
$$\Theta' = (\boldsymbol{\theta}_1^{r'}, \ldots, \boldsymbol{\theta}_{K_T}^{r'}) \qquad \text{(rows of } \Theta).$$

The priors are then given by $(\boldsymbol{\theta}_k^c|\lambda_t) \sim \mathcal{N}_{dim(\boldsymbol{\theta}_k^c)}(\mathbf{0}, (\lambda_t P_t)^{-1})$ and $(\boldsymbol{\theta}_j^r|\lambda_d) \sim \mathcal{N}_{dim(\boldsymbol{\theta}_j^r)}(\mathbf{0}, (\lambda_d P_d)^{-1})$ for $k = 1, \ldots, K_D$ and $j = 1, \ldots, K_T$, respectively. It follows that the joint prior for $\boldsymbol{\theta}$ can be written as:

$$\begin{aligned}
p(\boldsymbol{\theta}|\lambda_t, \lambda_d) &= \left(\prod_{k=1}^{K_D} p(\boldsymbol{\theta}_k^c|\lambda_t)\right)\left(\prod_{j=1}^{K_T} p(\boldsymbol{\theta}_j^r|\lambda_d)\right) \\
&\propto \exp\left(-\frac{1}{2}\lambda_t \sum_{k=1}^{K_D} \boldsymbol{\theta}_k^{c'} P_t \boldsymbol{\theta}_k^c\right) \exp\left(-\frac{1}{2}\lambda_d \sum_{j=1}^{K_T} \boldsymbol{\theta}_j^r P_d \boldsymbol{\theta}_j^{r'}\right) \\
&= \exp\left(-\frac{1}{2}\lambda_t \boldsymbol{\theta}'(I_{K_D} \otimes P_t)\boldsymbol{\theta}\right) \exp\left(-\frac{1}{2}\lambda_d \boldsymbol{\theta}'(P_d \otimes I_{K_T})\boldsymbol{\theta}\right) \\
&= \exp\left(-\frac{1}{2}\boldsymbol{\theta}'\left(\lambda_t(I_{K_D} \otimes P_t) + \lambda_d(P_d \otimes I_{K_T})\right)\boldsymbol{\theta}\right).
\end{aligned}$$

Gaussian priors are assumed for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, namely $\boldsymbol{\beta} \sim \mathcal{N}_{dim(\boldsymbol{\beta})}(\mathbf{0}, V_{\boldsymbol{\beta}}^{-1})$ with $V_{\boldsymbol{\beta}} = \zeta I_{p+1}$ (small ζ, e.g. $\zeta = 10^{-5}$) and $(\boldsymbol{\theta}|\boldsymbol{\lambda}) \sim \mathcal{N}_{dim(\boldsymbol{\theta})}(\mathbf{0}, \mathcal{P}^{-1}(\boldsymbol{\lambda}))$, where $\boldsymbol{\lambda} = (\lambda_t, \lambda_d)'$ is the penalty vector and $\mathcal{P}(\boldsymbol{\lambda}) = \lambda_t(I_{K_D} \otimes P_t) + \lambda_d(P_d \otimes I_{K_T})$ the global penalty matrix. Denote by $X = (B, Z)$ the global design matrix, $\boldsymbol{\xi} = (\boldsymbol{\beta}', \boldsymbol{\theta}')'$ the latent parameter vector and $Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}} = \begin{bmatrix} V_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathcal{P}(\boldsymbol{\lambda}) \end{bmatrix}$ the precision matrix for $\boldsymbol{\xi}$. The full Bayesian model is then summarized as follows:

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

$$
\begin{aligned}
&(y_i|\boldsymbol{\xi}) \sim \text{Poisson}(\mu_i) \;\text{ with }\; \log(\boldsymbol{\mu}) = X\boldsymbol{\xi},\\
&(\boldsymbol{\xi}|\boldsymbol{\lambda}) \sim \mathcal{N}_{dim(\boldsymbol{\xi})}(\mathbf{0}, (Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}})^{-1}),\\
&(\lambda_t|\delta_t) \sim \mathcal{G}\left(\frac{\nu}{2}, \frac{\nu\delta_t}{2}\right),\\
&(\lambda_d|\delta_d) \sim \mathcal{G}\left(\frac{\nu}{2}, \frac{\nu\delta_d}{2}\right),\\
&\delta_t \sim \mathcal{G}(a_\delta, b_\delta),\\
&\delta_d \sim \mathcal{G}(a_\delta, b_\delta),
\end{aligned}
$$

where $\mathcal{G}(a, b)$ denotes a Gamma distribution with mean a/b and variance a/b^2. This robust prior specification on the penalty parameters follows from Jullion and Lambert (2007). They have shown that when the hyperparameters a_δ, b_δ are chosen to be equal and small enough (say 10^{-4}), then the resulting fit is robust to the value of ν (e.g. $\nu = 3$ in this paper).

2.2 Laplace approximation to the conditional posterior of $\boldsymbol{\xi}$

To obtain a Laplace approximation for the set of regression parameters $\boldsymbol{\xi} = (\boldsymbol{\beta}', \boldsymbol{\theta}')'$, we use the posterior of $\boldsymbol{\xi}$ conditional on the vector of penalty parameters $\boldsymbol{\lambda}$. Then, the gradient and Hessian of the log-posterior are computed and used in a Newton-Raphson algorithm to obtain the Gaussian approximation to the conditional posterior distribution of $\boldsymbol{\xi}$.

Note that for a Poisson distributed y_i with mean μ_i, we have $f(y_i) \propto \exp\left(y_i\gamma_i - s(\gamma_i)\right)$, where $\gamma_i = \log(\mu_i)$ and $s(\gamma_i) = \exp(\gamma_i) = \mu_i$. Using Bayes' rule, the posterior of $\boldsymbol{\xi}$ conditional on the penalty vector $\boldsymbol{\lambda}$ is:

$$
\begin{aligned}
p(\boldsymbol{\xi}|\boldsymbol{\lambda}, \mathcal{D}) &\propto \mathcal{L}(\boldsymbol{\xi}; \mathcal{D}) p(\boldsymbol{\xi}|\boldsymbol{\lambda})\\
&\propto \exp\left(\sum_{i=1}^{n} (y_i\gamma_i - s(\gamma_i))\right) \exp\left(-\frac{1}{2}(\boldsymbol{\xi}' Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}} \boldsymbol{\xi})\right)\\
&= \exp\left(\sum_{i=1}^{n} (y_i\gamma_i - s(\gamma_i)) - \frac{1}{2}(\boldsymbol{\xi}' Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}} \boldsymbol{\xi})\right),
\end{aligned}
$$

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

where $\mathcal{L}(\boldsymbol{\xi};\mathcal{D})$ denotes the likelihood function. Hence, the log-conditional posterior of $\boldsymbol{\xi}$ can be obtained as:

$$\log p(\boldsymbol{\xi}|\boldsymbol{\lambda},\mathcal{D}) \doteq \sum_{i=1}^{n}\Big(y_i\gamma_i - s(\gamma_i)\Big) - \frac{1}{2}(\boldsymbol{\xi}^{'}Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}\boldsymbol{\xi}), \tag{2}$$

where $\doteq$ denotes equality up to an additive constant.

It is known for Poisson generalized linear models that the gradient (score vector) and Hessian of the first term on the right hand side of (2) are given, respectively, by (see Agresti (2013), p. 136):

$$\nabla_{\boldsymbol{\xi}} \log \mathcal{L}(\boldsymbol{\xi};\mathcal{D}) = X^{'}(\boldsymbol{y}-\boldsymbol{\mu}) \quad \text{and} \quad \nabla^2_{\boldsymbol{\xi}} \log \mathcal{L}(\boldsymbol{\xi};\mathcal{D}) = -X^{'}WX,$$

where $W = \text{diag}(w_1,\ldots,w_n)$ is a diagonal matrix with $w_i = (Var(y_i)[g'(\mu_i)]^2)^{-1} = \mu_i$ and $g(\cdot)$ is the log-link function such that $g(\mu_i) = \log(\mu_i)$. Hence, the gradient and Hessian for the log-conditional posterior of $\boldsymbol{\xi}$ are:

$$\nabla_{\boldsymbol{\xi}} \log p(\boldsymbol{\xi}|\boldsymbol{\lambda},\mathcal{D}) = X^{'}(\boldsymbol{y}-\boldsymbol{\mu}) - Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}\boldsymbol{\xi} \quad \text{and} \quad \nabla^2_{\boldsymbol{\xi}} \log p(\boldsymbol{\xi}|\boldsymbol{\lambda},\mathcal{D}) = -(X^{'}WX + Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}).$$

Using a Newton-Raphson algorithm, the Laplace approximation for the conditional posterior of $\boldsymbol{\xi}$ is, by abuse of notation, $\widetilde{p}_G(\boldsymbol{\xi}|\boldsymbol{\lambda},\mathcal{D}) = \mathcal{N}(\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}},\widehat{\Sigma}_{\boldsymbol{\lambda}})$ (Gressani and Lambert, 2021), where $\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}$ and $\widehat{\Sigma}_{\boldsymbol{\lambda}}$ correspond to the mode and variance-covariance matrix, respectively, given by:

$$\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}} = (X^{'}\widetilde{W}X + Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}})^{-1}\widetilde{\boldsymbol{w}} \quad \text{and}$$
$$\widehat{\Sigma}_{\boldsymbol{\lambda}} = (X^{'}\widetilde{W}X + Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}})^{-1},$$

where $\widetilde{\boldsymbol{w}}$ is the vector at convergence resulting from the sequence $\widetilde{\boldsymbol{w}}^{(0)}$,$\widetilde{\boldsymbol{w}}^{(1)}$, $\widetilde{\boldsymbol{w}}^{(2)}$,..., with $\widetilde{\boldsymbol{w}}^{(0)} = X^{'}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\xi}^{(0)})) + X^{'}W(\boldsymbol{\xi}^{(0)})X\boldsymbol{\xi}^{(0)}$, the initial condition $\boldsymbol{\xi}^{(0)} = (0,0,\ldots,0,0)'_{dim(\boldsymbol{\xi})}$, $\boldsymbol{\mu}(\boldsymbol{\xi}^{(0)}) = \exp(X\boldsymbol{\xi}^{(0)})$, and $W(\boldsymbol{\xi}^{(0)}) = \text{diag}(X\boldsymbol{\xi}^{(0)})$.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2.3 Hyperparameter optimization

In this section, we derive the (approximate) posterior distribution of the hyperparameters belonging to the penalization part, i.e., λ and δ. We first derive the joint posterior of λ and δ and then proceed with an integration to isolate the marginal posterior distribution of the penalty parameter.

Let $\boldsymbol{\eta} = (\lambda_t, \lambda_d, \delta_t, \delta_d)'$ denotes the vector of hyperparameters. Using Bayes' theorem, the marginal posterior of $\boldsymbol{\eta}$ is:

$$\begin{aligned} p(\boldsymbol{\eta}|\mathcal{D}) &= \frac{p(\boldsymbol{\xi}, \boldsymbol{\eta}|\mathcal{D})}{p(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathcal{D})} \\ &= \frac{p(\mathcal{D}|\boldsymbol{\xi}, \boldsymbol{\eta})p(\boldsymbol{\xi}, \boldsymbol{\eta})}{p(\mathcal{D})p(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathcal{D})} \\ &\propto \frac{\mathcal{L}(\boldsymbol{\xi}; \mathcal{D})p(\boldsymbol{\xi}|\boldsymbol{\eta})p(\boldsymbol{\eta})}{p(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathcal{D})}. \end{aligned}$$

Following Rue et al. (2009), the above posterior can be approximated by replacing $p(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathcal{D})$ with $\widetilde{p}_G(\boldsymbol{\xi}|\boldsymbol{\lambda}, \mathcal{D})$ obtained in Section 2.2 and by evaluating the latent vector at $\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}$. Note that $\gamma_i = \mathbf{x}_i'\boldsymbol{\xi}$, where $\mathbf{x}_i'$ corresponds to the ith row of the design matrix X. Also, the determinant $|Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}|$ in $p(\boldsymbol{\xi}|\boldsymbol{\lambda})$ is $|Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}| \propto |\mathcal{P}(\boldsymbol{\lambda})|$. Hence, the approximated marginal posterior of $\boldsymbol{\eta}$ is:

$$\begin{aligned} \widetilde{p}(\boldsymbol{\eta}|\mathcal{D}) &\propto \frac{\mathcal{L}(\boldsymbol{\xi}; \mathcal{D})p(\boldsymbol{\xi}|\boldsymbol{\eta})p(\boldsymbol{\eta})}{\widetilde{p}_G(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathcal{D})}\Big|_{\boldsymbol{\xi}=\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}} \\ &\propto \exp\left(\sum_{i=1}^{n}(y_i\mathbf{x}_i'\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}} - s(\mathbf{x}_i'\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}))\right) \\ &\times |\mathcal{P}(\boldsymbol{\lambda})|^{\frac{1}{2}}\exp\left(-\frac{1}{2}\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}'Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}\right) \\ &\times (\lambda_t\lambda_d)^{(\frac{\nu}{2}-1)}\,\delta_t^{(\frac{\nu}{2}+a_\delta-1)}\exp(-\delta_t(b_\delta + \frac{\nu}{2}\lambda_t)) \\ &\times \delta_d^{(\frac{\nu}{2}+a_\delta-1)}\exp(-\delta_d(b_\delta + \frac{\nu}{2}\lambda_d)) \\ &\times |X'\widetilde{W}X + Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}|^{-\frac{1}{2}}. \end{aligned}$$

To obtain the marginal posterior of $\boldsymbol{\lambda}$, we need to integrate out the hyperparameters δ_t and δ_d

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

from $\widetilde{p}(\boldsymbol{\eta}|\mathcal{D})$ as follows:

$$\int_0^\infty \delta_t^{(\frac{\nu}{2}+a_\delta-1)} \exp(-\delta_t(b_\delta + \frac{\nu}{2}\lambda_t))\, d\delta_t \propto (b_\delta + \frac{\nu}{2}\lambda_t)^{-(\frac{v}{2}+a_\delta)} \qquad \text{and}$$
$$\int_0^\infty \delta_d^{(\frac{\nu}{2}+a_\delta-1)} \exp(-\delta_d(b_\delta + \frac{\nu}{2}\lambda_d))\, d\delta_d \propto (b_\delta + \frac{\nu}{2}\lambda_d)^{-(\frac{v}{2}+a_\delta)}.$$

Hence,

$$\begin{aligned}\widetilde{p}(\boldsymbol{\lambda}|\mathcal{D}) &= \int_0^\infty \int_0^\infty \widetilde{p}(\boldsymbol{\eta}|\mathcal{D}) d\delta_t\, d\delta_d \\ &\propto |X'\widetilde{W}X + Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}}|^{-\frac{1}{2}} \exp\left(\sum_{i=1}^n (y_i \mathbf{x}_i' \widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}} - s(\mathbf{x}_i' \widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}})) - \frac{1}{2}\widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}' Q_{\boldsymbol{\xi}}^{\boldsymbol{\lambda}} \widehat{\boldsymbol{\xi}}_{\boldsymbol{\lambda}}\right) \\ &\times |\mathcal{P}(\boldsymbol{\lambda})|^{\frac{1}{2}} (\lambda_t \lambda_d)^{(\frac{\nu}{2}-1)} \left((b_\delta + \frac{\nu}{2}\lambda_t)(b_\delta + \frac{\nu}{2}\lambda_d)\right)^{-(\frac{v}{2}+a_\delta)}.\end{aligned}$$

The posterior mode (obtained via Newton-Raphson) is used as a point estimate for the penalty vector. To ensure numerical stability, we log-transformed the penalty vector $\boldsymbol{v} = (v_t, v_d)' = (\log(\lambda_t), \log(\lambda_d))'$. Using the method of transformations, we obtain the posterior of $\boldsymbol{v}$ given by:

$$\begin{aligned}\widetilde{p}(\boldsymbol{v}; y) &\propto |X'\widetilde{W}X + Q_{\boldsymbol{\xi}}^{\boldsymbol{v}}|^{-\frac{1}{2}} \exp\left(\sum_{i=1}^n (y_i \mathbf{x}_i' \widehat{\boldsymbol{\xi}}_{\boldsymbol{v}} - s(\mathbf{x}_i' \widehat{\boldsymbol{\xi}}_{\boldsymbol{v}})) - \frac{1}{2}\widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}' Q_{\boldsymbol{\xi}}^{\boldsymbol{v}} \widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}\right) \\ &\times |\mathcal{P}(\boldsymbol{v})|^{\frac{1}{2}} \left(b_\delta + \frac{\nu}{2}\exp(\lambda_t)(b_\delta + \frac{\nu}{2}\exp(\lambda_d)\right)^{-(\frac{v}{2}+a_\delta)} \\ &\times (\exp(v_t)\exp(v_d))^{(\frac{\nu}{2})}.\end{aligned}$$

The exponent of the last term above reduces to $\frac{v}{2}$ because of multiplying $(\exp(v_d)\exp(v_t))^{(\frac{\nu}{2}-1)}$ with the Jacobian of the transformation $J = \exp(v_d)\exp(v_t)$. Moreover, the log-posterior of $\boldsymbol{v}$ is given by:

$$\begin{aligned}\log \widetilde{p}(\boldsymbol{v}; y) \doteq &-\frac{1}{2}\log|X'\widetilde{W}X + Q_{\boldsymbol{\xi}}^{\boldsymbol{v}}| + \sum_{i=1}^n y_i \boldsymbol{x}_i' \widehat{\boldsymbol{\xi}}_{\boldsymbol{v}} - \sum_{i=1}^n s(\boldsymbol{x}_i' \widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}) - \frac{1}{2}\widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}' Q_{\boldsymbol{\xi}}^{\boldsymbol{v}} \widehat{\boldsymbol{\xi}}_{\boldsymbol{v}} \\ &+ \frac{1}{2}\log|P(\boldsymbol{v})| - \left(\frac{\nu}{2} + a_\delta\right)\left(\log(b_\delta + \frac{\nu}{2}\exp(v_d)) + \log(b_\delta + \frac{\nu}{2}\exp(v_t))\right) \\ &+ \frac{\nu}{2}(v_d + v_t).\end{aligned}$$

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2.4 Nowcasting with prediction interval

To obtain the mean nowcast estimate with the prediction interval, note that $\log(\mu_i) = \boldsymbol{x}_i^{'}\boldsymbol{\xi}$ or equivalently $\log(\mu_{t,d}) = \boldsymbol{x}_{t,d}^{'}\boldsymbol{\xi}$, where $\widetilde{p}_G(\boldsymbol{\xi}|\boldsymbol{v},\mathcal{D}) = \mathcal{N}(\widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}})$. Thus, $\widetilde{p}(\log(\mu_{t,d})|\boldsymbol{v},\mathcal{D}) = \mathcal{N}(\boldsymbol{x}_{t,d}^{'}\widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}, \boldsymbol{x}_{t,d}^{'}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}}\boldsymbol{x}_{t,d})$. The mean estimate for the not-yet-reported cases is calculated as $\widehat{\mu}_{t,d} = \exp(\boldsymbol{x}_{t,d}^{'}\widehat{\boldsymbol{\xi}}_{\boldsymbol{v}})$ for all (t,d) combinations with $t > T-d$. Then, the estimate for the total number of cases for each t is obtained by summing the already reported cases and the mean estimate for the not-yet-reported cases, i.e. $\widehat{\mu}_t = \sum_{\{d:t\leq T-d\}} y_{t,d} + \sum_{\{d:t>T-d\}} \widehat{\mu}_{t,d}$.

The prediction interval for the nowcast values is obtained by sampling from the posterior predictive distribution of the log-mean number of cases by following the five steps below:

1. For each (t,d) combinations with $t > T-d$ (corresponding to the not-yet-reported cases), generate 1000 random samples $(\widehat{\boldsymbol{y}}_{t,d})$ from a Gaussian distribution with mean $\boldsymbol{x}_{t,d}^{'}\widehat{\boldsymbol{\xi}}_{\boldsymbol{v}}$ and variance $\boldsymbol{x}_{t,d}^{'}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}}\boldsymbol{x}_{t,d}$.

2. Exponentiate the sampled values from the previous step to obtain the average reporting intensities $\widehat{\boldsymbol{\mu}}_{t,d} = \exp(\widehat{\boldsymbol{y}}_{t,d})$.

3. Compute the average prediction for the not-yet-reported cases for each t. That is, compute $\widehat{\boldsymbol{\mu}}_t = \sum_d \widehat{\boldsymbol{\mu}}_{t,d}$ for $t > T-d$.

4. For each t, sample $\widehat{\boldsymbol{y}}_t$ containing 1000 values from Poisson$(\widehat{\boldsymbol{\mu}}_t)$.

5. Finally, compute the quantiles of the sampled values $\widehat{\boldsymbol{y}}_t$ that correspond to the desired prediction interval.

2.5 Delay distribution

To obtain the smooth estimate of the delay distribution, only the first term on the right hand side of equation (1) is used (excluding the week effects), as explained in the paper of Van de Kassteele et al. (2019). Specifically, the procedure to compute the delay distribution is as follows:

1. Compute the contribution of the smoothing term in equation (1) to the reporting intensity for all (t,d) combinations: $\boldsymbol{\mu}^{\text{smooth}} = \exp(B\boldsymbol{\theta})$.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2. Arrange $\boldsymbol{\mu}^{\text{smooth}}$ into a $T \times (D+1)$ matrix with entries $\mu_{t,d}^{\text{smooth}}$.

3. For each $t = 1, \ldots, T$, compute the reporting delay distribution given by: $\hat{f}_t(d) = \frac{\mu_{t,d}^{\text{smooth}}}{\sum_{d=0}^{D} \mu_{t,d}^{\text{smooth}}}$.

3 Simulations

A simulation study is implemented in order to evaluate the predictive performance of the proposed method. The procedure to perform the simulations is as follows:

1. Consider a function $f(t)$ that represents the mean epidemic curve of all cases such that $\mu(t) = \exp(f(t))$ for $t = 1, ..., T$.

2. For each t, generate a random sample y_t from a Poisson distribution with mean $\mu(t)$.

3. To account for possible delays $d = 0, 1, 2, \ldots, D$, generate samples from a multinomial distribution with probabilities $(p_0, p_1, p_2, \cdots, p_D$ such that $\sum_{d=0}^{D} p_d = 1)$, i.e.,

$$(y_{t,0}, y_{t,1}, y_{t,2}, \ldots, y_{t,D}) \sim Multinomial(y_t, p_0, p_1, p_2, \cdots, p_D).$$

   This sample represents the reported number of cases for each (t, d) combination.

4. Repeat steps 1 to 3 for 1000 times to generate 1000 possible realizations and compute the desired accuracy measures.

We consider two functions $f(t)$ inspired from the paper of Noufaily et al. (2016) given by:

$$f_1(t) = \theta_1 + \theta_2 \sin\left(\frac{2\pi t}{150}\right),$$
$$f_2(t) = \theta_1 + \theta_2 \sin\left(\frac{2\pi t}{150}\right) + \theta_3 \sqrt{t},$$

for $t = 1, \ldots, 365$. That is, we assume a one year (365 days) time window in the simulation. In terms of delay probabilities, we consider a maximum delay of $D = 7$ days with probabilities $(p_0, p_1, p_2, \cdots, p_7) = (0.0, 0.1, 0.4, 0.2, 0.1, 0.1, 0.05, 0.05)$. This is illustrated in Figure 1. For example, at time $T - 6$, only the cases that have a delay of 7 days (with probability $p_7 = 0.05$) are not yet reported, or equivalently, 95% of the cases have already been reported. For time

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

$T-5$, only 90% of the cases have been reported, as cases with delays of 6 and 7 days are yet to be reported. On the nowcast day (time T), no case is reported, that is, the delay is 100%.

Two scenarios are considered for the first function $f_1(t)$: (i) the first one is having a small number of cases with values of $\theta_1 = 2$ and $\theta_2 = 1$; (ii) the second scenario has relatively larger cases with $\theta_1 = 3$ and $\theta_2 = 1$. These functions are denoted by $f_{11}(t)$ and $f_{12}(t)$, respectively. The second function $f_2(t)$ has values $\theta_1 = 0$, $\theta_2 = 0.4$ and $\theta_3 = 0.2$. The first function has a symmetric curve with three peaks as shown in Figures 2a and 2b. On the other hand, the second function is not periodic as opposed to the first function (see Figure 2c). The plot for (one realization of) simulated cases based on these functions are shown in Figure 3 with (blue dashed) vertical lines corresponding to the different nowcast dates.

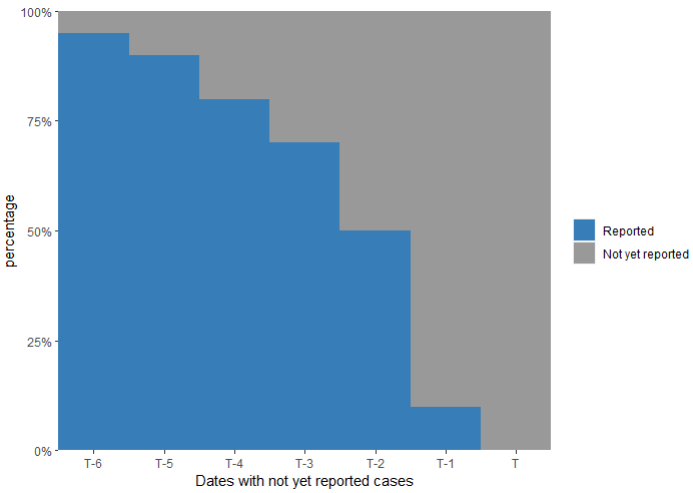
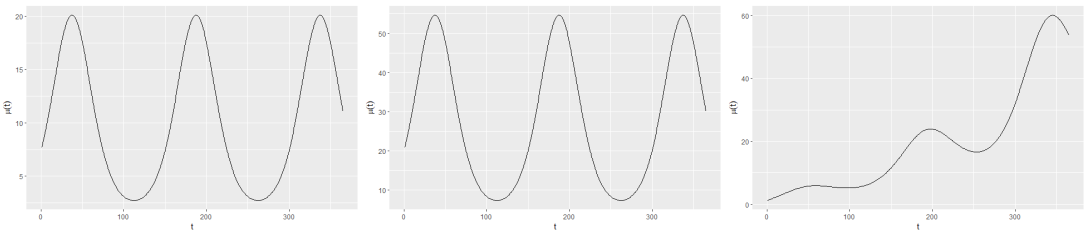


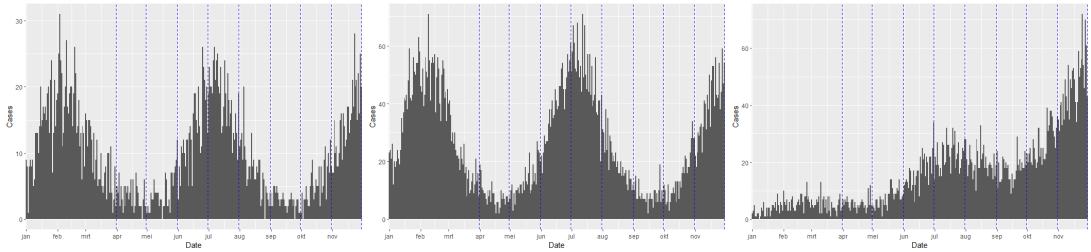
Figure 1: Illustration of the delay probabilities considered in the simulation.



(a) Plot for $\mu(t) = \exp(f_{11}(t))$ (b) Plot for $\mu(t) = \exp(f_{12}(t))$ (c) Plot for $\mu(t) = \exp(f_2(t))$

Figure 2: Plot of the functions considered in the simulation.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .



(a) $\mu(t) = \exp(f_{11}(t))$ (b) $\mu(t) = \exp(f_{12}(t))$ (c) $\mu(t) = \exp(f_2(t))$

Figure 3: Plot of the simulated cases for different functions. Blue dashed vertical lines - different nowcast dates.

The mean absolute percentage error (MAPE) and symmetric mean absolute percentage error (SMAPE) are chosen to measure the predictive accuracy of our methodology. Moreover, for each simulation, we consider a 95% prediction interval and determine the (true) unreported cases that falls within the interval to obtain the prediction interval coverage. The formulas to compute these measures are given by:

$$MAPE_t = \frac{1}{S}\sum_{s=1}^{S}\left|\frac{y_{st} - \hat{y}_{st}}{y_{st}}\right| \times 100\%,$$

$$SMAPE_t = \frac{1}{S}\sum_{s=1}^{S}\frac{|y_{st} - \hat{y}_{st}|}{(|y_{st}|+|\hat{y}_{st}|)/2} \times 100\%,$$

where $S = 1000$ is the number of generated realizations, y_{st} is the true value (generated from step 2 of the simulation in Section 3 at iteration s) and $\hat{y}_{st}$ is the corresponding predicted value (mean nowcast estimate $\widehat{\mu}_t$ in Section 2.4). The MAPE and SMAPE metrics are scale-independent and measure the error of prediction from the actual value in terms of percentage. One disadvantage of MAPE is that it could have undefined values when the actual values are zero, which is evident in the formula where the true value is in the denominator. In our setting, this happened especially with very small delay probabilities where there is a high probability of zero unreported cases, and so we exclude this case in computing the MAPE. Another way to address this issue is to use the SMAPE indicator that has values with lower and upper bounds of 0% and 200%, respectively.

In the simulation, every end of the month from March to November are set as nowcast dates. For each nowcast date, the prediction measures are computed for the dates having unreported

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

cases, that is, for $t = T - (D-1), \ldots, T$. As we have a maximum delay of 7 days, there will be seven dates (including the nowcast date) that involves prediction of unreported cases. Tables 1-3 summarize the MAPE results for the different functions being considered. The error increases as it gets closer to the nowcast day, that is, from 6 days prior to the nowcast day to the nowcast day. This is expected, as there is less data available at each delay time as we approach the nowcast day. Indeed, on day T, about 95% of the cases that occurred on day $T-6$ have been reported, while on day T no case is reported yet. The errors are relatively higher in March and April, where we have the least available data. For functions $f_{12}(t)$ and $f_2(t)$, the MAPE on the nowcast day ranges between 20% - 35% for May to November. This is reasonable considering that we do not have data available on the nowcast day because all cases have a delay of at least one day. Moreover, the MAPE results using LPS are similar to the ones obtained using Van de Kaasteele et al.'s method (values in parentheses in the Table). There is no clear pattern that indicates one approach has higher or lower error than the other.

Results for the prediction interval coverage can be found in Tables 4-6. For function $f_{11}(t)$, almost all of the values are higher than the 95% nominal prediction interval. For function $f_{12}(t)$, the prediction interval coverage is close to 95% for most cases, although the coverage of the nowcast day for June and November are somewhat too low (60.04% and 73.95%, respectively). This could be resulting from reaching the peak (curve change) and the fact that no data is available for the nowcast day, that is, there is no zero delay. Moreover, the function $f_2(t)$ has high coverage except on the nowcast day for October and November, having higher simulated cases. There is moderate to strong undercoverage in the prediction interval with Van de Kassteele et al.'s method for 4 to 6 days before the nowcast day, while LPS is more conservative and has coverages that are closer to the nominal value. On the other hand, from 2 days before the nowcast day to the nowcast day, Van de Kaasteele et al.'s method has higher coverage than LPS, reflecting much wider prediction intervals. The results for SMAPE of the three functions are also displayed in Appendix A2 with the percentage error of functions $f_{12}(t)$ and $f_2(t)$ ranging between 20% - 40% on the nowcast day. Similar to MAPE results, LPS and Van de Kaasteele et al.'s method have comparable SMAPE values with no systematic difference between the two approaches.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1: Mean absolute percentage error (MAPE) for the number of cases at day $T-t\,(t=0,\ldots,6)$ using function $f_{11}(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	9.9 (4.8)	15.2 (9.8)	23.1 (19.8)	35.9 (23.6)	52.1 (37.5)	63.7 (50.0)	80.4 (53.3)	40.1 (28.4)
Apr	10.5 (5.5)	18.5 (10.7)	28.0 (19.2)	38.0 (29.2)	52.8 (35.0)	64.8 (47.4)	75.3 (54.8)	41.1 (28.8)
May	8.8 (4.8)	13.6 (9.9)	17.7 (16.0)	24.6 (12.0)	36.0 (18.0)	41.0 (21.5)	42.1 (22.8)	26.3 (15.0)
June	4.7 (5.0)	7.1 (9.7)	9.9 (16.0)	13.5 (12.8)	18.3 (16.6)	22.3 (22.4)	22.7 (22.7)	14.1 (15.0)
July	5.1 (5.0)	8.1 (10.0)	11.2 (15.2)	15.4 (13.7)	22.5 (23.9)	25.5 (28.8)	26.7 (32.6)	16.4 (18.4)
Aug	9.7 (5.0)	15.9 (10.3)	22.3 (20.1)	29.5 (23.6)	40.4 (36.8)	56.4 (47.0)	65.7 (50.5)	34.3 (27.6)
Sept	9.6 (5.3)	17.3 (10.0)	27.0 (20.4)	33.2 (28.8)	45.2 (34.7)	53.4 (50.3)	62.4 (51.1)	35.4 (28.7)
Oct	8.0 (4.8)	12.2 (9.2)	16.0 (20.4)	19.8 (18.3)	24.8 (24.1)	32.3 (33.5)	30.5 (34.9)	20.5 (20.8)
Nov	4.8 (5.0)	6.8 (10.2)	9.3 (15.1)	12.0 (12.3)	15.8 (16.8)	19.1 (19.8)	20.8 (22.4)	12.6 (14.5)

Table 2: Mean absolute percentage error (MAPE) for the number of cases at day $T-t\,(t=0,\ldots,6)$ using function $f_{12}(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	5.3 (4.9)	8.1 (9.6)	12.6 (16.7)	18.9 (19.6)	28.1 (31.4)	36.2 (45.4)	35.2 (46.6)	20.6 (24.9)
Apr	7.7 (5.5)	11.7 (10.1)	16.2 (20.3)	24.4 (19.1)	33.9 (26.0)	40.6 (39.3)	43.3 (43.7)	25.4 (23.4)
May	5.1 (5.0)	7.4 (10.4)	9.9 (13.7)	13.2 (11.3)	18.8 (32.3)	20.8 (29.5)	24.6 (32.8)	14.2 (19.3)
June	2.7 (4.9)	4.3 (9.9)	6.2 (14.2)	8.1 (10.2)	10.7 (26.5)	15.8 (25.5)	26.0 (26.7)	10.6 (16.9)
July	3.1 (5.0)	4.6 (10.2)	6.5 (12.0)	8.5 (19.5)	11.8 (47.0)	15.6 (46.7)	20.7 (51.4)	10.1 (27.4)
Aug	5.9 (4.8)	8.7 (10.2)	12.0 (17.4)	15.7 (15.4)	23.8 (21.7)	28.9 (32.5)	31.5 (37.5)	18.1 (19.9)
Sept	7.5 (4.9)	11.1 (9.5)	14.7 (19.8)	19.0 (19.6)	28.0 (25.2)	34.0 (37.8)	34.4 (40.6)	21.2 (22.5)
Oct	4.6 (5.0)	6.4 (9.9)	9.0 (17.0)	10.7 (14.1)	13.6 (14.5)	19.8 (22.5)	24.0 (24.5)	12.6 (15.4)
Nov	2.7 (4.9)	4.0 (10.0)	5.4 (13.7)	6.6 (10.9)	8.7 (31.2)	16.3 (29.5)	22.6 (30.5)	9.5 (18.7)

Table 3: Mean absolute percentage error (MAPE) for the number of cases at day $T-t\,(t=0,\ldots,6)$ using function $f_2(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	9.9 (5.0)	15.9 (10.2)	23.3 (20.1)	32.1 (23.0)	46.5 (38.6)	57.6 (48.3)	68.7 (52.5)	36.3 (28.2)
Apr	8.6 (4.6)	14.5 (9.9)	19.2 (20.1)	25.4 (19.3)	35.1 (28.7)	44.8 (44.0)	55.8 (43.8)	29.1 (24.4)
May	6.2 (5.0)	9.7 (9.8)	13.9 (13.8)	18.3 (11.8)	23.8 (26.6)	27.9 (24.8)	29.0 (25.1)	18.4 (16.7)
June	4.3 (5.1)	6.8 (10.0)	9.3 (15.8)	12.6 (12.1)	17.5 (15.7)	21.6 (20.2)	23.1 (20.7)	13.6 (14.2)
July	3.9 (4.9)	6.1 (9.7)	8.7 (15.1)	10.7 (11.1)	15.5 (20.6)	18.4 (20.5)	21.4 (21.4)	12.1 (14.7)
Aug	4.7 (5.2)	7.3 (10.1)	9.6 (14.9)	12.0 (12.6)	17.1 (18.4)	21.6 (21.2)	22.8 (24.3)	13.6 (15.3)
Sept	4.5 (5.2)	6.7 (10.2)	9.3 (16.6)	11.7 (13.1)	15.7 (16.3)	21.5 (21.0)	22.8 (22.9)	13.2 (15.0)
Oct	3.7 (5.0)	5.1 (10.1)	7.1 (15.7)	9.0 (10.5)	11.0 (13.5)	18.5 (16.0)	24.5 (16.3)	11.3 (12.4)
Nov	2.5 (5.0)	3.7 (9.9)	5.3 (13.9)	6.7 (10.6)	8.6 (28.8)	17.8 (24.3)	26.4 (24.4)	10.1 (16.7)

Table 4: Prediction interval coverage for the number of cases at day $T-t\,(t=0,\ldots,6)$ using function $f_{11}(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	100.0 (77.7)	99.9 (90.7)	99.3 (92.1)	98.2 (98.0)	95.4 (99.2)	95.1 (96.7)	94.9 (98.1)	97.5 (93.2)
Apr	100.0 (86.7)	99.9 (96.2)	99.5 (91.2)	99.6 (97.8)	99.3 (98.2)	97.5 (97.1)	95.9 (97.0)	98.8 (94.9)
May	100.0 (41.5)	99.5 (44.1)	99.2 (60.8)	98.4 (95.9)	95.5 (98.0)	95.3 (98.1)	96.1 (99.0)	97.7 (76.8)
June	99.4 (40.8)	99.5 (50.7)	97.4 (60.2)	95.7 (95.0)	93.6 (99.0)	93.6 (98.3)	94.9 (98.8)	96.3 (77.5)
July	99.9 (46.9)	99.5 (56.6)	97.7 (76.2)	96.3 (97.1)	95.7 (98.1)	97.0 (98.3)	98.2 (99.4)	97.8 (81.8)
Aug	100.0 (78.9)	100.0 (92.0)	99.5 (93.2)	99.5 (98.2)	99.3 (98.8)	97.4 (98.2)	97.9 (97.3)	99.1 (93.8)
Sept	100.0 (85.3)	100.0 (96.2)	99.8 (91.3)	99.6 (96.8)	99.7 (96.8)	99.0 (93.1)	98.4 (95.5)	99.5 (93.6)
Oct	99.9 (72.1)	99.9 (85.1)	99.2 (82.4)	99.2 (94.7)	97.9 (98.1)	96.7 (96.7)	98.7 (98.7)	98.8 (89.7)
Nov	99.6 (40.6)	98.9 (41.0)	96.6 (63.9)	96.1 (96.1)	96.0 (98.5)	96.2 (99.2)	95.7 (99.1)	97.0 (76.9)

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 5: Prediction interval coverage for the number of cases at day $T-t\,(t=0,\ldots,6)$ using function $f_{12}(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	99.9 (50.4)	99.4 (66.3)	96.3 (86.1)	93.9 (98.8)	89.9 (99.1)	90.9 (99.0)	95.1 (99.5)	95.0 (85.6)
Apr	99.9 (67.5)	99.7 (82.3)	97.5 (86.4)	96.9 (98.0)	95.7 (99.2)	94.4 (99.2)	95.9 (98.8)	97.1 (90.2)
May	99.7 (9.1)	98.4 (11.1)	96.8 (75.4)	94.8 (95.7)	92.3 (99.3)	93.9 (98.9)	91.1 (99.8)	95.3 (69.9)
June	99.2 (9.6)	97.5 (12.8)	95.6 (73.7)	92.3 (94.9)	92.4 (98.8)	86.8 (98.2)	63.2 (99.3)	89.6 (69.6)
July	99.3 (13.4)	98.5 (29.5)	96.1 (83.4)	95.4 (98.0)	94.8 (98.7)	94.0 (99.6)	89.1 (98.6)	95.3 (74.5)
Aug	99.7 (55.5)	99.6 (69.7)	98.5 (86.4)	97.9 (98.8)	96.2 (99.6)	96.8 (99.0)	97.5 (98.8)	98.0 (86.8)
Sept	100.0 (68.4)	99.7 (83.8)	99.0 (82.2)	99.0 (96.3)	97.0 (98.9)	96.2 (98.6)	98.1 (98.6)	98.4 (89.5)
Oct	99.7 (38.7)	99.0 (43.4)	97.1 (64.2)	97.5 (97.6)	97.6 (99.5)	91.9 (99.0)	91.5 (99.5)	96.3 (77.4)
Nov	99.5 (8.2)	97.1 (10.5)	96.2 (76.9)	96.6 (97.3)	96.4 (99.7)	84.6 (99.4)	73.6 (99.4)	92.0 (70.2)

Table 6: Prediction interval coverage for the number of cases at day $T-t\,(t=0,\ldots,6)$ using function $f_2(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	100.0 (75.1)	100.0 (89.1)	99.1 (91.1)	98.0 (97.2)	94.4 (99.0)	93.7 (97.5)	92.7 (98.6)	96.8 (92.5)
Apr	99.9 (76.2)	99.7 (88.1)	99.0 (89.8)	99.0 (96.6)	97.5 (99.3)	95.2 (97.4)	95.1 (97.8)	97.9 (92.2)
May	99.6 (8.0)	98.9 (9.5)	97.5 (57.3)	96.3 (87.3)	94.9 (98.1)	94.8 (95.8)	96.2 (97.9)	96.9 (64.8)
June	99.8 (34.2)	98.3 (40.2)	96.7 (53.1)	94.8 (93.1)	93.4 (98.8)	92.1 (99.1)	93.1 (98.8)	95.4 (73.9)
July	99.6 (31.3)	99.1 (39.2)	96.7 (63.3)	97.1 (95.6)	95.5 (98.1)	95.7 (98.4)	93.1 (99.3)	96.7 (75.0)
Aug	99.4 (38.7)	98.9 (46.5)	97.3 (71.9)	96.6 (95.6)	96.9 (98.4)	94.8 (98.2)	96.4 (98.1)	97.2 (78.2)
Sept	99.7 (39.3)	99.6 (45.3)	96.4 (60.3)	95.9 (93.4)	96.6 (98.4)	93.0 (97.2)	95.4 (98.2)	96.7 (76.0)
Oct	99.2 (21.5)	98.6 (20.9)	96.0 (42.2)	95.7 (90.8)	96.9 (99.2)	86.9 (96.7)	81.4 (98.2)	93.5 (67.1)
Nov	99.2 (5.7)	97.8 (7.5)	96.5 (60.4)	94.4 (93.1)	96.0 (98.4)	79.6 (97.8)	59.4 (98.3)	89.0 (65.9)

4 Real Data application

We apply our method to COVID-19 mortality data in Belgium, for the year 2021. The raw data is available on the website of the Sciensano research institute (see sciensano.be/covid19data). The data contains the cumulative number of mortality cases, reported up to the day of the file. The file is updated every day, and in this way, the number of cases and reporting delays are obtained. The data is structured in matrix format with the date of death as rows and number of days of reporting delay as columns. Figure 4 shows the total number of cases with (blue dashed) vertical lines corresponding to different nowcast dates used for illustration. As no cases are reported immediately on the day of death in the data, all cases have a delay of at least one day. The proportion of cases reported for each delay is shown in Table 7. Majority of the cases (66%) were reported with a delay of 2 days. Followed by a delay of 1 day and 3 days in 14% and 12% of the cases, respectively. The rest of the delays account for 8% of the cases. The reporting delay is truncated to a maximum of seven days.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

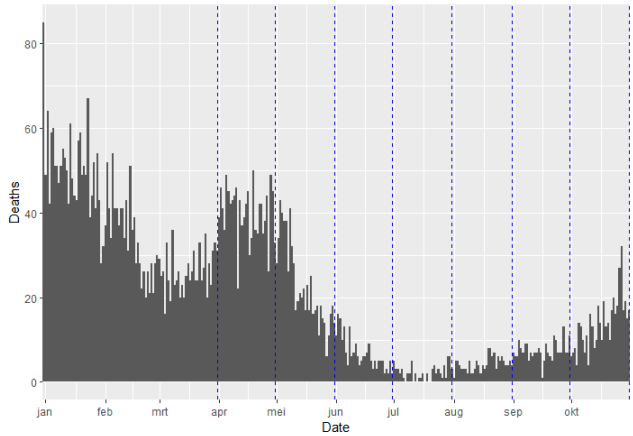


Figure 4: Covid-19 death cases in Belgium for the year 2021. Blue dashed vertical lines - different nowcast dates.

Table 7: Proportion of cases reported for each delay for the mortality data.

delay (days)	0	1	2	3	4	5	6	7
proportion	0.000	0.143	0.664	0.118	0.047	0.017	0.007	0.004

In this section, we present the model results at different nowcast dates. In the analysis, we choose $K_T = 40$ and $K_D = 10$ as the number of B-splines for the time and delay dimension, respectively. Moreover, a cubic B-splines basis with a second order penalty is used. In addition, we also consider the day of the week effect as additional covariate in the model.

Figure 5 presents the nowcasting results using the mortality data at different nowcast dates (namely, at every end of the month). The blue color represents the reported cases for the past 14 days on the nowcast date, the gray color are cases that have not yet been reported, and the orange line corresponds to the nowcast prediction with the 95% prediction interval. It can be seen that the nowcast predictions are fairly close to the observed cases (gray). In addition, all of the observed cases fall within the prediction interval. The nowcasts at the day of nowcasting (i.e. last date presented in the plots) for March and April are slightly underestimated. Based on the simulation results, possible reasons for the underestimation are: (1) the number of cases in this period is higher as compared to other periods; (2) the amount of information used in this period is still limited (from January till March or April); and (3) at the nowcast date, no information is available as there is a 100% delay of cases.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

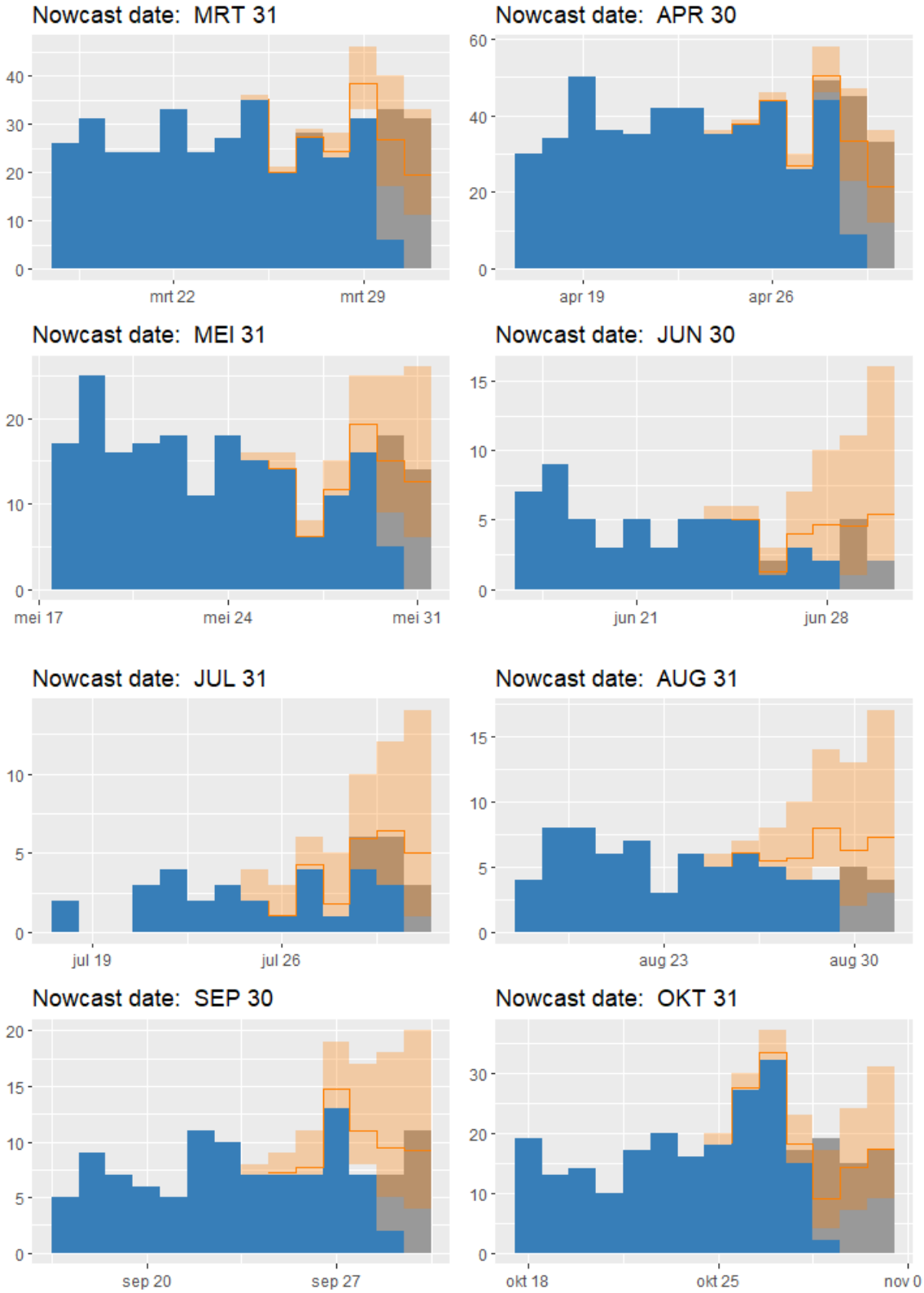


Figure 5: Nowcast plot for mortality data with different nowcast dates. Blue - reported cases ; Gray - Not-yet-reported cases; Orange - nowcast with prediction interval.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The estimated delay density for different nowcast dates is shown in Figure 6. The steps to obtain the time-varying delay distribution is outlined in Section 2.5. Figure 6 shows that the delay distribution is fairly constant through time, with density being highest for a delay of 2 days. This confirms the observed reporting intensity in our data that most cases are reported with a two-day delay. Furthermore, the density is highest between April and June, which relates to a higher number of reported cases (see Figure 4).

Table 8 shows the computational time required for doing the nowcasting analysis using mortality data. We used the *microbenchmark()* function from the *microbenchmark* package in R. The analysis was implemented on a device with an Intel(R) Core(TM) i5-1135G7, CPU running at a base frequency of 2.40GHz, and having 4 cores. Twenty function calls were evaluated for each nowcast date. The average real elapsed time is around 12 to 25 seconds, except for the nowcast date on October, for which we used almost all of the 2021 data with an average of 50 seconds.

Table 8: Computation time (in seconds) for 20 evaluations using mortality data with the LPS methodology.

(seconds)	min	mean	median	max
March	14.99	**15.46**	15.35	16.45
April	20.07	**20.43**	20.41	20.91
May	19.10	**19.44**	19.43	19.73
June	12.34	**12.53**	12.54	12.85
July	24.59	**24.81**	24.80	25.12
August	24.08	**24.65**	24.46	26.19
September	16.26	**16.52**	16.50	16.78
October	49.69	**50.16**	50.04	51.41

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

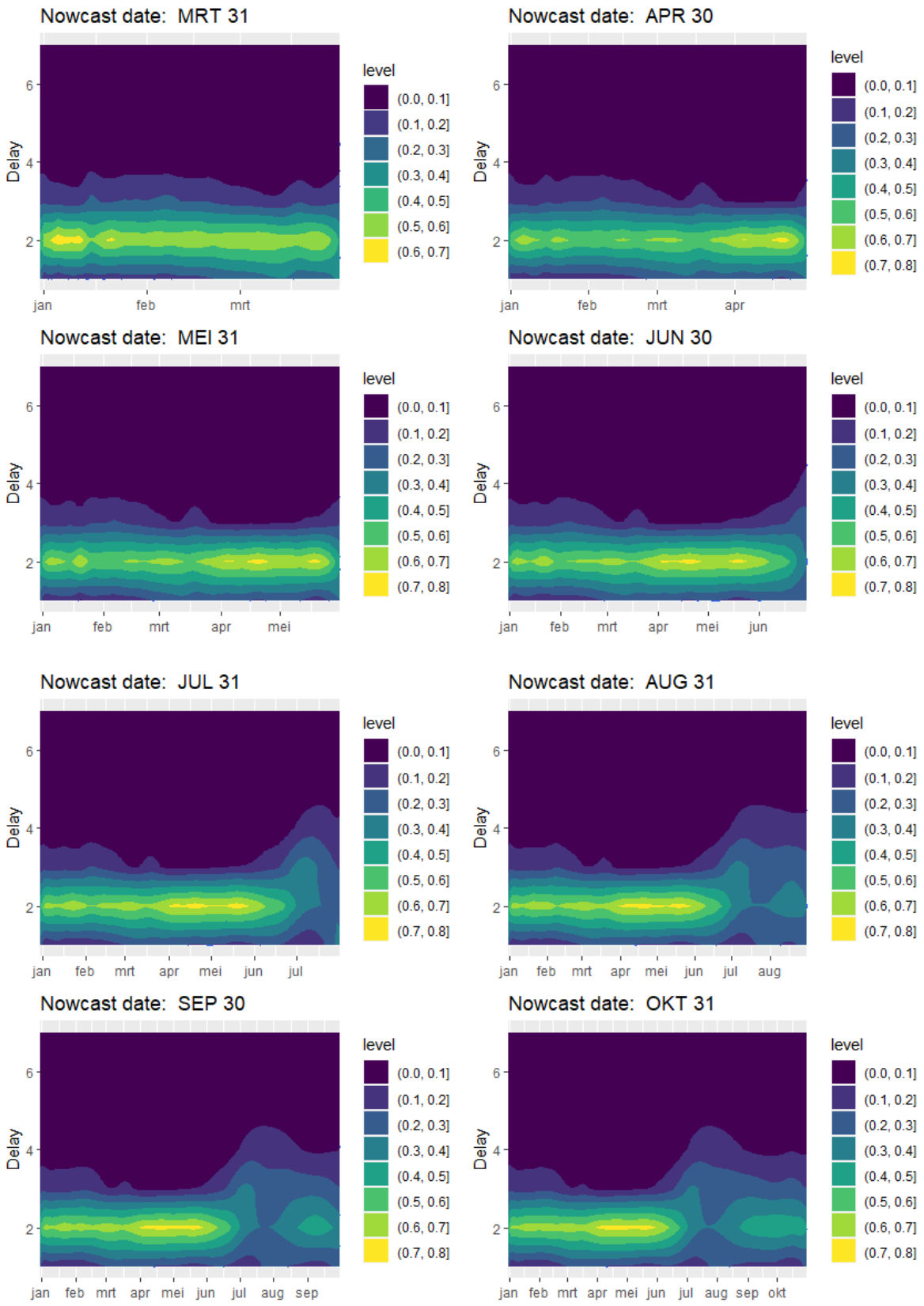


Figure 6: Estimated delay density for mortality data with different nowcast dates.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

5 Conclusion

This paper presents a novel Bayesian approach of the nowcasting model proposed by Van de Kassteele et al. (2019). One advantage of using the Bayesian approach is that the prediction intervals can be naturally obtained which is the main goal of nowcasting. Our proposed methodology is based on a combination of Laplace approximations and P-splines, making it both fast and flexible. We perform a simulation study to evaluate the (predictive) performance of our method. Simulation results show that nowcasts and prediction interval coverage are fairly close to the true value and nominal (95%) prediction interval, respectively, except when the nowcast date falls on the day of the peak where it has low coverage. This could be due to the higher values and a shift in the epidemic trend. Moreover, we also apply our method to the Belgium COVID-19 mortality data (2021). The nowcasts are close to the actual observed values, and all of the 95% prediction intervals for the different nowcast dates being considered contain the observed values.

This study has some limitations. First, it assumes that the counts are Poisson distributed. One possible consequence of this is that the variability in the data tends to be underestimated because of overdispersion, which can cause undercoverage in the prediction intervals. This can be addressed by using the negative Binomial distribution. Another issue is that the priors for the B-spline coefficients are based on the Kronecker product of the penalty matrices. Lang and Brezger (2004) pointed out that this could overfit the data, and they suggest using spatial smoothness priors. Moreover, as mentioned by Van de Kassteele et al. (2019) our model specification makes it difficult to model abrupt changes in the reporting process.

Competing Interest Statement

The authors have declared no competing interest.

Funding Statement

This project is funded by the European Union Research and Innovation Action under the H2020 work programme EpiPose (grant number 101003688).

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

References

Agresti, A. (2013). *Categorical Data Analysis*. John Wiley & Sons.

Donker, T., van Boven, M., van Ballegooijen, W. M., van't Klooster, T. M., Wielders, C. C., and Wallinga, J. (2011). Nowcasting pandemic influenza A/H1N1 2009 hospitalizations in the Netherlands. *European Journal of Epidemiology*, 26(3):195–201.

Durbán, M., Currie, I., and Eilers, P. H. C. (2002). Using P-splines to smooth two-dimensional Poisson data. In *Proceedings of 17th International Workshop on Statistical Modelling, Chania, Crete*, pages 207–214.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, 11(2):89–121.

Eilers, P. H. C., Marx, B. D., and Durbán, M. (2015). Twenty years of P-splines. *SORT: Statistics and Operations Research Transactions*, 39(2):149–186.

Fahrmeir, L., Kneib, T., Lang, S., and Marx, B. (2013). *Regression: Models, Methods and Applications*. Springer Science & Business Media.

Glöckner, S., Krause, G., and Höhle, M. (2020). Now-casting the COVID-19 epidemic: The use case of Japan, March 2020. *medRxiv*.

Gressani, O., Faes, C., and Hens, N. (2022). Laplacian-P-splines for Bayesian inference in the mixture cure model. *Statistics in Medicine*, 41(14):2602–2626.

Gressani, O. and Lambert, P. (2018). Fast Bayesian inference using Laplace approximations in a flexible promotion time cure model based on P-splines. *Computational Statistics & Data Analysis*, 124:151–167.

Gressani, O. and Lambert, P. (2021). Laplace approximations for fast Bayesian inference in generalized additive models based on P-splines. *Computational Statistics & Data Analysis*, 154:107088.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Gressani, O., Wallinga, J., Althaus, C., Hens, N., and Faes, C. (2021). EpiLPS: a fast and flexible Bayesian tool for near real-time estimation of the time-varying reproduction number. *medRxiv*.

Günther, F., Bender, A., Katz, K., Küchenhoff, H., and Höhle, M. (2021). Nowcasting the COVID-19 pandemic in Bavaria. *Biometrical Journal*, 63(3):490–502.

Gutierrez, E., Rubli, A., and Tavares, T. (2020). Delays in death reports and their implications for tracking the evolution of COVID-19. *Available at SSRN 3645304*.

Höhle, M. and an der Heiden, M. (2014). Bayesian nowcasting during the STEC O104: H4 outbreak in Germany, 2011. *Biometrics*, 70(4):993–1002.

Jullion, A. and Lambert, P. (2007). Robust specification of the roughness penalty prior distribution in spatially adaptive Bayesian P-splines models. *Computational Statistics & Data Analysis*, 51(5):2542–2558.

Lang, S. and Brezger, A. (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics*, 13(1):183–212.

Lawless, J. (1994). Adjustments for reporting delays and the prediction of occurred but not reported events. *Canadian Journal of Statistics*, 22(1):15–31.

Noufaily, A., Farrington, P., Garthwaite, P., Enki, D. G., Andrews, N., and Charlett, A. (2016). Detection of infectious disease outbreaks from laboratory data with reporting delays. *Journal of the American Statistical Association*, 111(514):488–499.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.

Van de Kassteele, J., Eilers, P. H. C., and Wallinga, J. (2019). Nowcasting the number of new symptomatic cases during infectious disease outbreaks using constrained P-spline smoothing. *Epidemiology*, 30(5):737.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

6 Appendix

A1.

In equation (1), we have a two-dimensional B-spline model:

$$\log(\boldsymbol{\mu}) = B\boldsymbol{\theta} + Z\boldsymbol{\beta},$$

where

- $B = B_D \otimes B_T$ is the Kronecker product of B-splines matrices B_T and B_D with dimension $T \times K_T$ and $D \times K_D$, given by:

$$B_T = \begin{bmatrix} b_1(1) & b_2(1) & \dots & b_{K_T}(1) \\ b_1(2) & b_2(2) & \dots & b_{K_T}(2) \\ \vdots & \vdots & \dots & \vdots \\ b_1(T) & b_2(T) & \dots & b_{K_T}(T) \end{bmatrix} \text{and } B_D = \begin{bmatrix} \tilde{b}_1(1) & \tilde{b}_2(1) & \dots & \tilde{b}_{K_D}(1) \\ \tilde{b}_1(2) & \tilde{b}_2(2) & \dots & \tilde{b}_{K_D}(2) \\ \vdots & \vdots & \dots & \vdots \\ \tilde{b}_1(D) & \tilde{b}_2(D) & \dots & \tilde{b}_{K_D}(D) \end{bmatrix}.$$

  As a result, B has dimension $TD \times K_T K_D$.

- Z is the design matrix for additional covariates with dimension $TD \times (p+1)$ and corresponding parameters $\boldsymbol{\beta}$. This is given by:

$$Z = \begin{bmatrix} 1 & z_1(1,1) & \dots & z_p(1,1) \\ \vdots & \vdots & \dots & \vdots \\ 1 & z_1(K_T,1) & \dots & z_p(K_T,1) \\ \vdots & \vdots & \dots & \vdots \\ 1 & z_1(1,K_D) & \dots & z_p(1,K_D) \\ \vdots & \vdots & \dots & \vdots \\ 1 & z_1(K_T,K_D) & \dots & z_p(K_T,K_D) \end{bmatrix}.$$

- $\boldsymbol{\theta} = \text{vec}(\Theta)$ has dimension $K_T K_D \times 1$ where $\text{vec}(\Theta)$ denotes the vectorization of a matrix Θ. It is obtain by vectorizing the matrix:

medRxiv preprint doi: https://doi.org/10.1101/2022.08.26.22279249; this version posted August 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

$$\Theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \dots & \theta_{1,K_D} \\ \theta_{2,1} & \theta_{2,2} & \dots & \theta_{2,K_D} \\ \vdots & \vdots & \dots & \vdots \\ \theta_{K_T,1} & \theta_{K_T,2} & \dots & \theta_{K_T,K_D} \end{bmatrix}$$

of B-splines coefficients $\theta_{j,k}$ by stacking the columns of Θ, that is,

$$\boldsymbol{\theta} = (\theta_{1,1}, \theta_{2,1}, \cdots, \theta_{K_T,1}, \theta_{1,2}, \theta_{2,2}, \cdots, \theta_{K_T,2}, \cdots, \theta_{1,K_D}, \theta_{2,K_D}, \cdots, \theta_{K_T,K_D})'.$$

A2. Result for symmetric mean absolute percentage error (SMAPE)

Table 9: Symmetric mean absolute percentage error (SMAPE) for the number of cases at day $T - t\,(t = 0, \dots, 6)$ using function $f_{11}(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	11.8 (5.6)	16.6 (12.3)	22.9 (25.3)	30.9 (25.8)	38.4 (33.6)	46.8 (47.8)	51.4 (49.1)	31.3 (28.5)
Apr	24.4 (7.9)	32.6 (15.5)	38.0 (28.0)	42.5 (42.3)	51.5 (45.9)	58.4 (64.0)	59.2 (63.7)	43.8 (38.2)
May	9.4 (5.1)	14.2 (10.7)	17.4 (17.9)	21.9 (13.0)	28.9 (16.5)	34.2 (20.9)	35.5 (21.9)	23.1 (15.1)
June	4.7 (5.3)	7.0 (10.6)	9.9 (18.1)	12.8 (13.8)	16.9 (15.3)	22.5 (21.8)	24.4 (21.9)	14.0 (15.2)
July	5.2 (5.3)	8.0 (10.9)	11.0 (17.1)	14.2 (13.9)	19.6 (20.6)	23.9 (25.6)	24.8 (28.0)	15.2 (17.3)
Aug	12.0 (6.1)	19.3 (13.1)	24.8 (26.5)	29.2 (27.8)	35.0 (35.0)	46.1 (47.0)	46.6 (50.5)	30.4 (29.4)
Sept	22.8 (7.1)	31.0 (14.1)	38.6 (29.9)	41.3 (39.4)	46.1 (46.9)	53.6 (67.2)	54.1 (63.5)	41.1 (38.3)
Oct	8.4 (5.4)	12.4 (10.5)	15.7 (24.7)	18.6 (20.6)	22.2 (24.6)	31.2 (35.7)	30.3 (35.4)	19.8 (22.4)
Nov	4.8 (5.3)	6.7 (11.0)	9.1 (16.9)	11.6 (13.2)	14.8 (15.5)	19.5 (19.4)	21.9 (21.4)	12.6 (14.7)

Table 10: Symmetric mean absolute percentage error (SMAPE) for the number of cases at day $T - t\,(t = 0, \dots, 6)$ using function $f_{12}(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	5.4 (5.2)	8.1 (10.5)	12.5 (19.0)	17.1 (18.2)	23.4 (24.0)	31.0 (38.3)	31.9 (40.1)	18.5 (22.2)
Apr	7.8 (6.2)	11.4 (11.6)	15.9 (24.5)	21.5 (22.2)	27.4 (26.3)	34.4 (41.4)	35.5 (44.8)	22.0 (25.3)
May	5.1 (5.2)	7.4 (11.0)	9.8 (14.9)	12.7 (10.9)	17.1 (26.2)	21.8 (24.8)	27.7 (27.0)	14.5 (17.1)
June	2.8 (5.1)	4.3 (10.6)	6.2 (15.5)	8.0 (10.1)	10.3 (22.3)	17.4 (22.3)	30.8 (23.1)	11.4 (15.6)
July	3.1 (5.2)	4.5 (10.9)	6.4 (13.0)	8.3 (17.3)	11.3 (33.9)	16.5 (34.3)	23.5 (36.5)	10.5 (21.6)
Aug	6.0 (5.2)	8.6 (11.2)	12.1 (20.0)	15.1 (16.4)	21.1 (20.7)	27.2 (34.9)	28.7 (38.0)	17.0 (20.9)
Sept	7.8 (5.4)	11.1 (11.0)	14.6 (23.9)	17.8 (22.4)	24.2 (25.9)	31.7 (42.6)	32.6 (43.7)	20.0 (25.0)
Oct	4.6 (5.3)	6.4 (10.7)	9.1 (19.2)	10.7 (15.5)	13.4 (14.8)	21.9 (24.9)	27.8 (26.6)	13.4 (16.7)
Nov	2.7 (5.0)	3.9 (10.6)	5.5 (15.0)	6.6 (10.6)	8.8 (25.5)	18.1 (24.7)	26.2 (25.3)	10.2 (16.7)

Table 11: Symmetric mean absolute percentage error (SMAPE) for the number of cases at day $T - t\,(t = 0, \dots, 6)$ using function $f_2(t)$, as well as the average over these days. Values in parentheses are results using Van de Kassteele et al.'s method.

	t							
Nowcast Dates	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**Average**
March	10.8 (5.7)	17.1 (12.5)	22.1 (25.2)	26.4 (23.8)	34.4 (32.0)	43.8 (40.8)	44.7 (40.6)	28.5 (25.8)
Apr	9.9 (5.3)	15.1 (11.8)	18.6 (25.5)	22.3 (21.0)	28.6 (26.0)	37.7 (37.2)	39.5 (38.8)	24.5 (23.6)
May	6.3 (5.2)	9.6 (10.4)	13.8 (15.1)	17.0 (11.6)	21.0 (22.0)	27.2 (22.0)	28.3 (22.1)	17.6 (15.5)
June	4.4 (5.4)	6.8 (10.8)	9.3 (17.8)	12.0 (13.1)	15.9 (14.5)	22.2 (20.3)	25.2 (20.6)	13.7 (14.6)
July	3.9 (5.1)	6.0 (10.4)	8.7 (16.7)	10.4 (11.3)	14.3 (18.0)	19.2 (19.1)	23.4 (19.6)	12.3 (14.3)
Aug	4.7 (5.5)	7.2 (11.0)	9.6 (16.8)	11.8 (13.3)	16.0 (16.7)	22.1 (21.4)	24.2 (24.3)	13.7 (15.6)
Sept	4.5 (5.5)	6.7 (11.1)	9.4 (18.7)	11.5 (14.3)	15.0 (15.8)	22.8 (22.7)	25.5 (24.6)	13.6 (16.1)
Oct	3.7 (5.2)	5.1 (10.8)	7.2 (17.4)	9.0 (11.2)	11.0 (12.6)	20.9 (16.8)	28.9 (16.8)	12.2 (13.0)
Nov	2.5 (5.1)	3.7 (10.5)	5.4 (15.2)	6.8 (10.3)	8.8 (24.2)	20.1 (21.1)	31.2 (21.1)	11.2 (15.4)