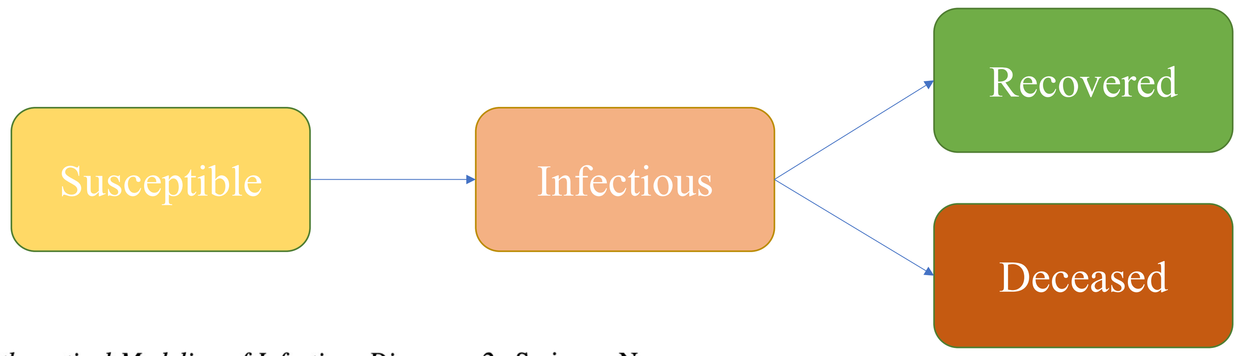
Epidemic Log Gaussian Cox Process

Francois Buet-Golfouse, *University College London*, ucahfbu@ucl.ac.uk
Hans Roggeman, *Independent Researcher*, hraoyama@gmail.com

A Semi-Parametric Approach

- Our approach is based on parametric mathematical epidemiological models, known as compartmental models, representing the spread of an epidemic.
- In particular, our focus, here, is on a Susceptible-Infectious-Recovered-Deceased approach on Covid-19 data from Germany. Our model is applicable to more complicated setups.



Michael Li. *An Introduction to Mathematical Modeling of Infectious Diseases: 2*. Springer, New York, NY, 1st ed. 2018 edition edition, February 2018. ISBN 978-3-319-72121-7.

- This framework can be translated in terms of *Poisson process* by looking at new infections i_t (defined as the the number of *susceptible* people at time $t-1$ who have become *infectious* at time t. Similarly, r_t and d_t are the new recovered and deceased respectively.
- In short, we introduce a Gaussian Process prior on the *log-intensities* $(\lambda, \gamma, \delta)$, motivated by epidemiology:

$$\begin{aligned} i_t &\sim \text{Poisson}(\lambda_t) \\ r_t &\sim \text{Poisson}(\gamma_t) \\ d_t &\sim \text{Poisson}(\delta_t) \end{aligned} \quad + \quad \begin{aligned} \log\lambda_t &= f_t^\lambda \sim \text{GP}(t, \log I_{t-1} + \log S_{t-1}) \\ \log\gamma_t &= f_t^\gamma \sim \text{GP}(t, \log I_{t-1}) \\ \log\delta_t &= f_t^\delta \sim \text{GP}(t, \log I_{t-1})^1. \end{aligned}$$

James Hensman and Theodore Kypraios. Variational Bayesian non-parametric inference for infectious disease models. In *Machine Learning for Healthcare Technologies*, pp. 181–202. October 2016. URL https://digital-library.theiet.org/content/books/10.1049/pbhe002e_ch9. Publisher: IET Digital Library.

A Semi-Parametric Approach

- As is usually done via MCMC, Laplace methods or VI, one tries to compute the *posterior* distribution of the GP to make predictions at future steps.
- We thus propose a new, simpler, approach, based on a *local* Laplace expansion on the log-likelihood.
- First, for $\xi = \lambda, \gamma, \delta$, the true posterior is given by

$$p(\mathbf{f}^\xi|\mathbf{y}^\xi) \propto \exp\left(\mathbf{y}^{\xi^T}\mathbf{f}^\xi - \sum_{i=1}^T e^{f_i^\xi} - \frac{1}{2}\left(\mathbf{f}^\xi - \mathbf{m}^\xi\right)^T \mathbf{K}^{\xi^{-1}}\left(\mathbf{f}^\xi - \mathbf{m}^\xi\right)\right)$$

- Dropping indices, we show that this can be well approximated by the following, where $\boldsymbol{H} = \text{diag}(\boldsymbol{y})$ and $\hat{\boldsymbol{f}} = \log \boldsymbol{y}$:

$$p(\mathbf{f}|\mathbf{y}) \propto \exp\left(-\frac{1}{2}\left(\mathbf{f}-\hat{\mathbf{f}}\right)^T \mathbf{H}\left(\mathbf{f}-\hat{\mathbf{f}}\right) - \frac{1}{2}(\mathbf{f}-\mathbf{m})^T\mathbf{K}^{-1}(\mathbf{f}-\mathbf{m})\right)$$

- Finally, the posterior is approximated by a Gaussian distribution $\phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$, with parameters given by:

$$\begin{aligned} \boldsymbol{\mu} &= \boldsymbol{\Sigma}\left(\mathbf{K}^{-1}\mathbf{m} + \mathbf{H}\hat{\mathbf{f}}\right) \\ \boldsymbol{\Sigma} &= \left(\mathbf{K}^{-1} + \mathbf{H}\right)^{-1}. \end{aligned}$$

- Note that, importantly, and contrary to other methods, the local Laplace posterior is available entirely in closed-form and does not require any iterations.

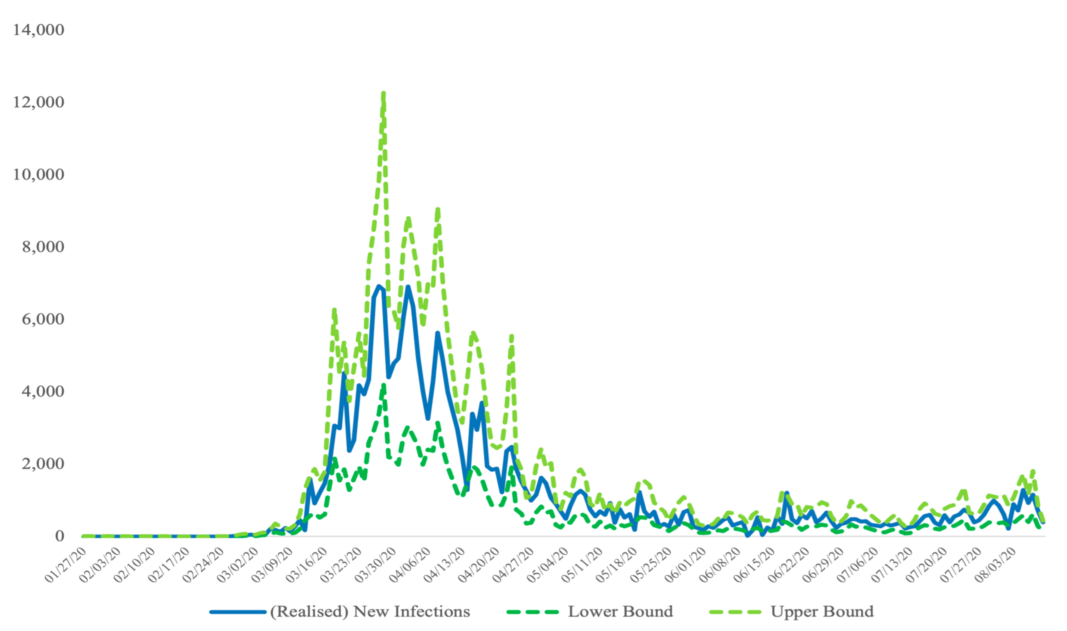
Ming Teng, Farouk Nathoo, and Timothy D. Johnson. Bayesian computation for Log-Gaussian Cox processes: a comparative analysis of methods. *Journal of Statistical Computation and Simulation*, 87(11), 2017. URL https://www.tandfonline.com/doi/abs/10.1080/00949655.2017.1326117.

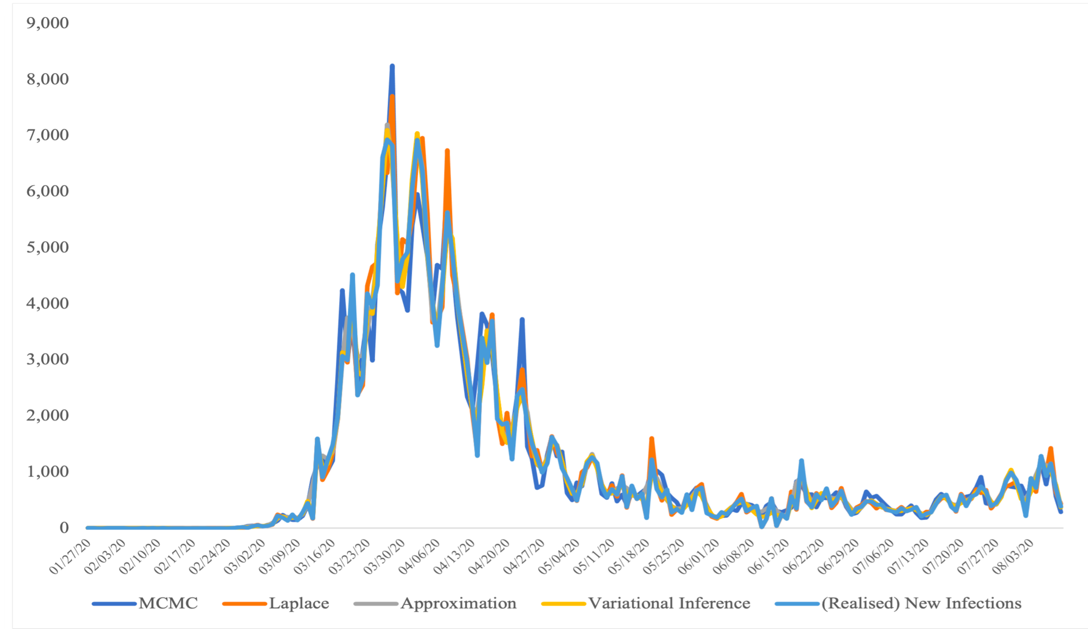
Total Variance and Uncertainty

- To consider performance of the overall model and estimation techniques, we point out the it is important to consider the total variance, given by

$$\mathbb{V}[\mathbf{y}_t^\xi] = \mathbb{E}[\mathbf{y}_t^\xi] + \mathbb{E}[\mathbf{y}_t^\xi]^2 \cdot \left(e^{\sigma^2_{\xi,t}} - 1\right)$$

- In particular, one should note that the second term (stemming from uncertainty around the log-intensity) is usually the main contributor to overall variance.
- All classical methods, as well as our local Laplace approximation, were run with a Matern-32 kernel (the same lengthscale and variance settings were used):





- The model performs rather well and manages to reproduce some of the seasonality and spikes observed in the data. This is one of the advantages of a semi-parametric approach thanks to Gaussian Processes.

Results

- We give a quantitative overview of our results contrasting methods and find that the performance of the *local Laplace* method is very close to the Laplace method, and has a performance which is slightly worse but equivalent to VI, while being available in closed-form.

NEW INFECTIONS

Method	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)
Markov Chain Monte Carlo	201.71	353.20
Laplace Approximation	86.28	183.50
Local Laplace	127.85	217.73
Variational Inference	102.21	175.18

NEW RECOVERIES

Method	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)
Markov Chain Monte Carlo	248.70	522.68
Laplace Approximation	183.51	452.69
Local Laplace	183.51	452.69
Variational Inference	183.53	359.86

NEW DEATHS

Method	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)
Markov Chain Monte Carlo	12.47	29.72
Laplace Approximation	10.70	28.00
Local Laplace	10.70	28.00
Variational Inference	11.17	27.98

- To summarise, we have proposed a new multivariate semi-parametric framework (which we also extended to a multi-output GP) to model the spread of epidemics, and put forward a novel, closed-form (hence quick, transparent and reproducible), posterior approximation, yielding good results.
- Future research includes considering this approach on graph Poisson processes.

Additional References

Virginia Aglietti, Theodoros Damoulas, and Edwin Bonilla. Efficient Inference in Multi-task Cox Process Models. *PMLR*, 89, March 2019. URL http://arxiv.org/abs/1805.09781. arXiv: 1805.09781.

James Hensman and Zoubin Ghahramani. Scalable Variational Gaussian Process Classification. *JMLR: W&CP*, 38:10, 2015.

Zhaozhi Qian, Ahmed M Alaa, and Mihaela van der Schaar. When and How to Lift the Lockdown? Global COVID-19 Scenario Analysis and Policy Assessment using Compartmental Gaussian Processes. *Advances in Neural Information Processing Systems*, pp. 12, 2020.

Jarno Vanhatalo, Ville Pietiläinen, and Aki Vehtari. Approximate inference for disease mapping with sparse Gaussian processes. *Statistics in Medicine*, 29(15):1580–1607, 2010. ISSN 1097-0258. doi: https://doi.org/10.1002/sim.3895. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.3895. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.3895.