

# Inference of seasonal forcings on reproduction numbers during epidemics

Lu Lu, Ziyao Tian, Fengyang Wang, Yi Xiao Zhang

## Background

Exponential growth of the number of symptomatic cases is usually observed in the initial stages of infectious disease epidemics. Typically, the rate of this growth is decomposed into two quantities:

- The serial interval: time between onset of symptoms of two adjacent individuals in a transmission chain
- The reproductive number,  $R$ : the expected number of new infections a single symptomatic individual will cause

The estimation of the latter quantity is important for policy, because infectious disease outbreaks will switch from an exponential growth regime to an exponential decay regime when the reproductive number falls below 1. The goal of mitigation or suppression policies is therefore to effect change on the reproductive number  $R$ .

Many infectious diseases are believed to be seasonal, including most respiratory viruses, like the SARS-CoV-2 virus implicated in the 2019–2020 coronavirus pandemic. Wang et al. (2020) have analysed how the weather of locations within China correlates with local point estimates of the reproductive number  $R$ . However, because their analysis is frequentist in nature, their estimates of the seasonal forcing do not come with any estimate of uncertainty.

## Bayesian Framework for Estimation of Time-Varying $R$

Thompson et al. (2019) outline a Bayesian framework for estimation of  $R(t)$  using symptomatic case counts, decomposed into local and imported cases. The EpiEstim R package implements ideas from this paper. A simplified summary of their approach, ignoring the imported cases:

- Estimation of (discrete) distribution of serial interval: Our project does not extend this portion, which estimates the quantities  $w_s$  that denote the probability density of interval length  $s$  under a posterior distribution for the serial interval.
- Estimation of time-varying  $R$ : Data is collected containing the number of symptomatic cases observed at time  $t$ :  $I(t)$ . The observed number of cases at each timestep is assumed to be drawn from a Poisson distribution:

$$I(t)|I(1), \dots, I(t-1), w, R(t) \sim \text{Poisson} \left( R(t) \sum_{s=1}^t I(t-s)w_s \right)$$

Then Bayesian inference can be used to compute the posterior distribution of  $R(t)$  conditional on the data and the serial interval distribution. (In the Thompson et al. paper, a gamma-distributed prior for  $R(t)$  [the conjugate prior for a Poisson likelihood] and a convenient serial interval distribution was chosen so that this posterior is proportional to an analytic expression.)

## Extension of Thompson et al. for $R$ subject to forcings

Our proposal is to extend the method of Thompson et al. to multiple locations, each with its own incidence time series  $I_1(t), \dots, I_n(t)$  such that

$$R_i(t) = \beta^T x_i(t) + \varepsilon_i(t)$$

where  $\beta$  is a parameter describing the locally linear effect of each seasonal forcing,  $x_i(t)$  is a vector of observed seasonal factors at time  $t$  and location  $i$ , and  $\varepsilon_i(t)$  denotes the component of  $R_i(t)$  not linearly predicted by seasonal conditions at location  $i$ . and the serial interval distribution  $(w_s)_{s \geq 1}$  are assumed to be characteristics of the disease independent of location. The observed number of cases at each timestep is

$$I_i(t) | I_i(1), \dots, I_i(t-1), w, x_i(t), \varepsilon_i(t) \sim \text{Poisson} \left( (\beta^T x_i(t) + \varepsilon_i(t)) \sum_{s=1}^t I(t-s) w_s \right)$$

Here the parameter of interest is  $\beta$  and the  $\varepsilon_i(t)$  are nuisance variables. We propose to use Markov chain Monte Carlo to compute a joint distribution over all the parameters and marginalize over the  $\varepsilon_i(t)$ . Because of concerns over confounding factors that may affect  $R$  and are correlated with temperature and humidity, we plan to add several other features to  $x(t)$  to help control for these factors. We also need to restrict the analysis to the pre-lockdown regime in each province, because the lockdown is likely to change the magnitude of seasonal forcings.

- Time (days)
- GDP per capita
- Population density

## Dataset

- The Italian government dataset at <https://github.com/pcm-dpc/COVID-19> contains the disease incidence for each province. This data contains date of confirmation instead of date of onset of symptoms, and so some procedure may be necessary to correct for this. Abbott et al. (2020) outlines a technique to correct for reporting delays using an exponential distribution of reporting delay fit to line list data; we may use the same distribution.
- We also will need to find temperature and humidity data. We have located several possible sources and are evaluating them.

## Directions for Future Research

- Looking at the climate and infection data in American and Canadian cities as the situation develops further. In particular, compare this with the results in Italy, and observe if the linear assumption between seasonal factors and  $R$  holds true across different climate regions. If not, try fitting while assuming a nonlinear relationship between  $R$  and seasonal factors.
- Look into statistical simulation models which consider intercity travel and geographic proximity.

## Citations

- (PRE-PRINT) Abbott, S., Hellewell, J., Munday, J. D., Chun, J. Y., Thompson, R. N., Bosse, N. I., Chan, Y. D., Russell, T. W., Jarvis, C. I., ... & Funk, S. (2020). Temporal variation in transmission during the COVID-19 outbreak. CMMID. Available at CMMID: <https://cmmid.github.io/topics/covid19/current-patterns-transmission/global-time-varying-transmission.html>

- Thompson, R. N., Stockwin, J. E., van Gaalen, R. D., Polonsky, J. A., Kamvar, Z. N., Demarsh, P. A., ... & Lessler, J. (2019). Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics*, 29, 100356.
- Wang, J., Tang, K., Feng, K., & Lv, W. (2020). High Temperature and High Humidity Reduce the Transmission of COVID-19. Available at SSRN: <https://ssrn.com/abstract=3551767> or <http://dx.doi.org/10.2139/ssrn.3551767>