

SIMULTANEOUS ESTIMATION OF THE EFFECTIVE REPRODUCING NUMBER AND THE DETECTION RATE OF COVID-19

YORIYUKI YAMAGATA

ABSTRACT. A major difficulty to estimate R (the effective reproducing number) of COVID-19 is that most cases of COVID-19 infection are mild or asymptomatic, therefore true number of infection is difficult to determine. This paper estimates the daily change of R and the detection rate simultaneously using a Bayesian model. The analysis using synthesized data shows that our model correctly estimates R and detects a short-term shock of the detection rate. Then, we apply our model to data from several countries to evaluate the effectiveness of public healthcare measures. Our analysis focuses Japan, which employs a moderate measure to keep “social distance”. The result indicates a downward trend and now R becomes below 1. Although our analysis is preliminary, this may suggest that a moderate policy still can prevent epidemic of COVID-19.

1. INTRODUCTION

In the wake of the COVID-19 epidemic, the Japanese government gradually employed public health measures against COVID-19. In the first stage, the stronger quarantine measures at the border were implemented. Once patients who had no connection to Wuhan appeared inside the border, The government started to track these patients as far as possible and tried to find people contacted to these patients. Once these “track and quarantine clusters” tactics were overwhelmed by the number of patients, the government started to ask behavior changes to people, culminating “declaration of the emergency” on April 6. Public facilities like libraries were closed. Large shopping malls and entertainment businesses, such as movie theatres, were asked to be closed. Restaurants are asked to shorten their operating hours and stop providing alcoholic beverages at night. Working from home was encouraged, and the citizens were advised to avoid crowded areas and generally avoid going outside unnecessarily. However, the Japanese legal system does not have enough mechanisms to enforce these policies. Thus the effectiveness of these policies can be questioned. Many people are still commuting to their office because many companies lack the necessary ability to allow their employees to work from home. Many small restaurants and cafes are still running their business because of the lack of financial compensation.

To estimate the policy effect to COVID-19, we need estimate daily changes of the effective reproducing number R . A major difficulty to estimate R of COVID-19 is that most cases of COVID-19 infection are mild or asymptomatic, therefore the true number of infection is difficult to determine. This paper estimates the daily change of R and the detection rate simultaneously using a Bayesian model. The analysis

using synthesized data shows that our model correctly estimates R and detects a short-term shock of the detection rate. Our analysis focuses Japan, which employs a moderate measure to keep “social distance”. The result indicates a downward trend and now R becomes below 1. Although our analysis is preliminary, this may suggest that a moderate policy still can prevent epidemic of COVID-19.

Then, we apply our model to data from several countries to evaluate the effectiveness of public healthcare measures. The comparison between Denmark and Sweden reveals that lock-down is very effective in short term. However, In Sweden, which did not employ lock-down, R also reduced and now R of both countries are roughly same. This might suggest that lock-down is not effective long run, or Denmark is disadvantageous against COVID-19 comparing to Sweden.

Further, we applied our method to China, Italy and US and show that these countries are also about exiting from epidemic.

2. RELATED WORKS

Several works employ data-driven methods to predict and measure the public health measure of COVID-19.

Anastassopoulou et al. [1] apply the SIRD model to Chinese official statistics, estimating parameters using linear regression. The reporting rate is not estimated from data but assumed. By these models and parameters, they predict the COVID-19 epidemic in Hubei province.

Diego Caccavo [2] and independently Peter Turchin [6] apply modified SIRD models, in which parameters change overtime following specific function forms. Parameters govern these functions are estimated by minimizing the sum-of-square-error. However, using the sum-of-square method causes over-fitting and always favors a complex model, therefore it is not suitable to assess policy effectiveness. Further, fitting the SIRD model in the early stage of infection is difficult, as pointed out in stat-exchange¹. Using a Bayesian method, we avoid these problems to some degree, because a Bayesian method estimates parameter distribution instead of a point estimate. Thus, we can assess the degree of confidence of each parameter. Further, by well-established statistical methods, we can compare the explanatory power of different models.

Flaxman et al. [4] use a Bayesian model to estimate policy effectiveness. The methodology is different from us because they assume immediate effects from the policies implemented. Further, they use a discrete renewal process, a more advanced model than the SIRD model. They use parameters estimated from studies of clinical cases while we use a purely data-driven method.

3. METHOD AND MATERIALS

3.1. Model. We use the discrete-time SIR model but assume that the number of move between each category is stochastic and follows Poisson distribution.

¹<https://stats.stackexchange.com/questions/446712/fitting-sir-model-with-2019-ncov-data-doesnt-converge>

$$\begin{aligned}
(1) \quad & NI(t) \sim \text{Poisson}\left(\frac{\beta IS}{P}\right) \\
(2) \quad & NR(t) \sim \text{Poisson}(\gamma I) \\
(3) \quad & I(t+1) = I(t) + NI(t) - NR(t) - ND(t) \\
(4) \quad & S(t+1) = S(t) - NI(t) \\
(5) \quad & R(t+1) = R(t) + NR(t)
\end{aligned}$$

The effective reproduction rate can be written

$$(6) \quad R = \frac{\beta}{\gamma} \cdot \frac{S}{P}$$

When R becomes < 1 then the infection starts to decline.

We cannot expect that these values are directly observable, because many (or most) cases are mild or asymptomatic. Therefore, we introduce the detection rate q and let the number of cumulative observed cases C_{obs} and recovered R_{obs} as

$$\begin{aligned}
(7) \quad & NI_{\text{obs}}(t) \sim \text{Poisson}(qNI(t)) \\
(8) \quad & NR_{\text{obs}}(t) \sim \text{Poisson}(\gamma I_{\text{obs}}(t)) \\
(9) \quad & I_{\text{obs}}(t+1) = I_{\text{obs}}(t) + NI_{\text{obs}}(t) - NR_{\text{obs}}(t) \\
(10) \quad & R_{\text{obs}}(t+1) = R_{\text{obs}}(t) + NR_{\text{obs}}(t)
\end{aligned}$$

We assume that β and q change day to day bases while other parameters are fixed. To get a reasonable estimate, we assume prior distributions somewhat arbitrary chosen.

$$\begin{aligned}
(11) \quad & S(0) = I(0) \sim \text{Student_t}(3, 0, 1) \\
(12) \quad & \beta(0) \sim \text{Student_t}(3, 0, 1) \\
(13) \quad & \beta(t+1) \sim \text{Student_t}(3, \beta(t), \sigma_b) \\
(14) \quad & \sigma_b \sim \text{Exponential}(1)
\end{aligned}$$

while the prior of $q(t)$ is the uniform distribution over $[0, 1]$. To make our model robust, we choose Student-t as the prior for β . We perform a sensitivity analysis of the prior for σ_b to ensure that the choice of priors does not strongly affect the result.

3.2. Data. The number of confirmed cases of each country up to May 11 were drawn from data repository ² by Johns Hopkins University Center for Systems Science and Engineering. The dataset also contains the number of recovery, but as pointed out in README the number is underestimated. For example, the number of recovery in Norway is only 32 in May 11, which is impossible with 8132 confirmed cases. Therefore, we estimate $\gamma = 0.04$ per day by Chinese data and assume that γ is constant across all countries.

²<https://github.com/CSSEGISandData/COVID-19>

3.3. Experiment. First, we performed model validation using synthesized data of the scenarios in which β and q are constant, β is constant but q is piecewise constant, β reduces linearly and q is constant with white noise and β, q are constant where β is high enough to make almost all people obtain immunity. The result is presented in Section 4.1.

Then the real-world data were fed to Stan [3] for parameter estimation by our Bayesian model. We simplified our model to ease modeling in Stan. Because latent discrete variables cannot be used in Stan, we used real numbers for $NI(t)$ and $NR(t)$. We used normal approximation $\mathcal{N}(\lambda, \sqrt{\lambda})$ for the Poisson distribution used for $NI(t)$. For $NR(t)$, we replaced stochastic laws to deterministic laws $NR(t) = \gamma I$ to avoid a numerical issue.

Parameter estimation used 10,000 iterations with 5,000 iterations for warm-up and 5,000 iterations for sampling. Four (default number of Stan) independent computations were performed simultaneously and used to estimate \hat{R} . If $\hat{R} < 1.1$ then we regard that the estimation is converged.

To make sure that our results are meaningful, we compared the performance of our model with a model (we refer it by “the constant model” **const**), which assumes constant β and q and a model (we refer it by “the constant detection rate model” **const-q**). **const** estimates constant β and q simultaneously but **const-q** takes q as a part of given data. Parameters were estimated for **const** and **const-q** in the same way and LOO, a standard measure of model performance, was compared. Because the exact computation of LOO-CV is computationally expensive, approximation PSIS-LOO-CV [7] and WAIC [8] were compared. Further, we check the reliability of these estimation by checking the Pareto-k for the importance weight distribution. The computation of PSIS-LOO-CV and WAIC is performed by ArviZ [5].

Models and the computation history used for this experiment are public at GitHub ³.

4. RESULTS

4.1. Model validation. Fig. 1 and Fig. 2 show the estimated β and q for synthesized data using $\beta = 0.07$ and $q = 0.2$. Fig 1 also shows the estimated β by **const-q** model. the true $q = 0.2$ is given to **const-q** model as data.

The results show that both models can estimate β correctly, while completely fail to estimate q . Estimated q is biased to 1 and noisy. Therefore, our method cannot estimate the absolute level of the detection rate. However, the estimate of β is still accurate so our method can be used to estimate β . In fact, Fig. 3 and Fig. 5 shows that our model is robust against changing q .

Fig. 3 and Fig. 4 show the estimated β and q for synthesized data using constant $\beta = 0.07$ but changing q which goes 1 to 0.8 in Apr. 1.

Fig 3 shows β estimated by **const-q** model drops at April 1. Although the estimate by our model also drops around the same date, the drop is less significant. This indicates that allowing q vary makes an estimated β robust against a sudden change of q .

Fig. 5 and Fig 6 show the estimated β and q for synthesized data using linearly decreasing β and noisy q . Fig. 5 shows that our method is smooth while the estimate of **const-q** is noisy. Again, this shows that our model is more robust against changing q . Further, Fig 6 indicates that while the estimated q is biased toward

³<https://github.com/yoriyuki/BayesianCOVID19/tree/used-for-paper-v2/notebook>

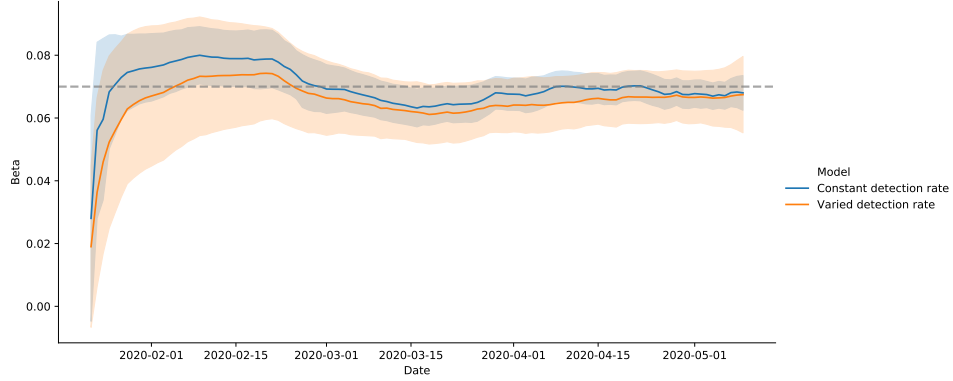


FIGURE 1. Estimated β for data generated with constant β and q . The blue line indicates the estimate of **const-q** model while the orange line indicates the estimate of our model. The dotted line indicates the true β . **const-q** model is given the correct q . The solid lines show means and shades with same colors show standard deviation.

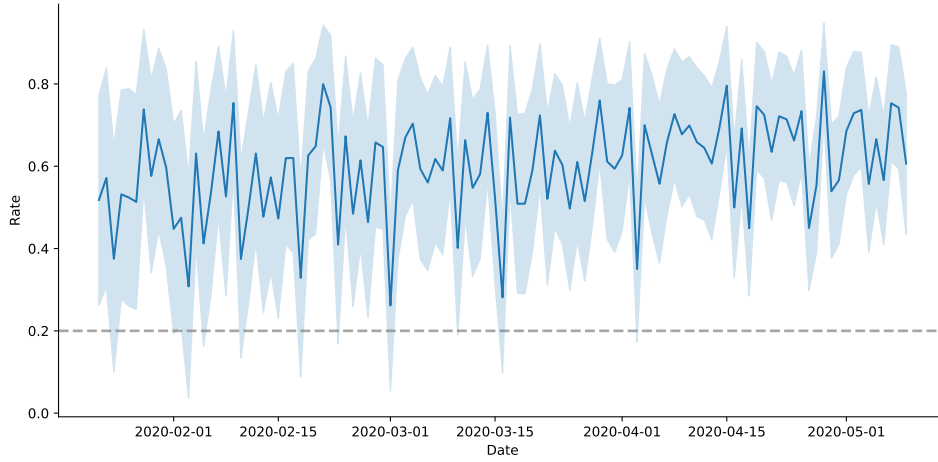


FIGURE 2. Estimated q for data generated with constant β and q . The solid lines show means and shades with same colors show standard deviation. The dotted line indicates the true q .

1, the short-term change coincides the ground truth. Therefore, the estimated q is still useful to find a short-term shock to the detection rate.

Fig. 7 and Fig. 8 show the estimate of β and q for data generated with constant β and q . Unlike Fig. 1 and Fig. 2, β and time horizon are chosen so that most of population are infected in the end. The result shows that the estimate of β becomes unreliable when the infection starts saturated. Therefore, our model is only useful for the case of a low infection rate. The estimated β by **const-q** swings widely, therefore is omitted.

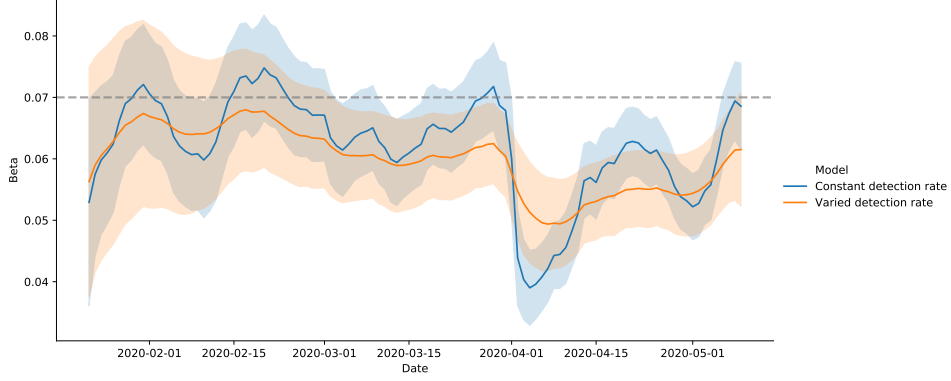


FIGURE 3. Estimated β for data generated with constant β and step-like q which starts with 1 and reduces its value to 0.8. The blue line indicates the estimate of **const-q** model while the orange line indicates the estimate of our model. The solid lines show means and shades with same colors show standard deviation. The dotted line indicates the true β . **const-q** model is given $q = 1$.

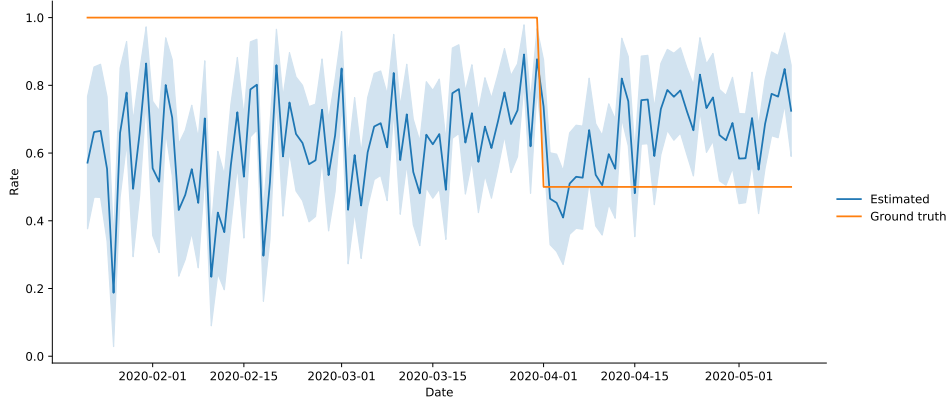


FIGURE 4. Estimated q for data generated with constant β and step-like q . The solid lines show means and shades with same colors show standard deviation. The orange line indicates the true q .

In summary, the estimate of β , which is a key parameter for estimate R , by our model is robust against sudden changes and noise in q , even though the estimate of q itself is unreliable. When the infection begins to saturate among population, the estimate of β as well as q becomes unreliable.

4.2. Analysis of Japan.

Next, we present our analysis of Japanese data.

First, we applied three models, **const**, **const-q** and our model (**varied-q**) to Japanese data. For all parameters of all models, $\hat{R} < 1.1$ holds so convergence is excellent. Then two information criteria, PSIS-LOO-CV and WAIC are applied to **const**, **const-q** and **varied-q**. The results are shown in Table 1.

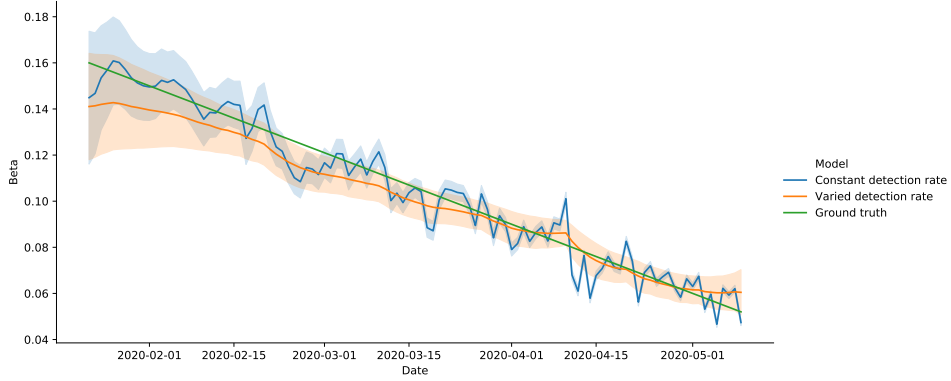


FIGURE 5. Estimated β for data generated with linearly decreasing β and noisy q . The blue line indicates the estimate of **const-q** model while the orange line indicates the estimate of our model. The solid lines show means and shades with same colors show standard deviation. The green line indicates the true β . **const-q** model is given $q = 0.5$.

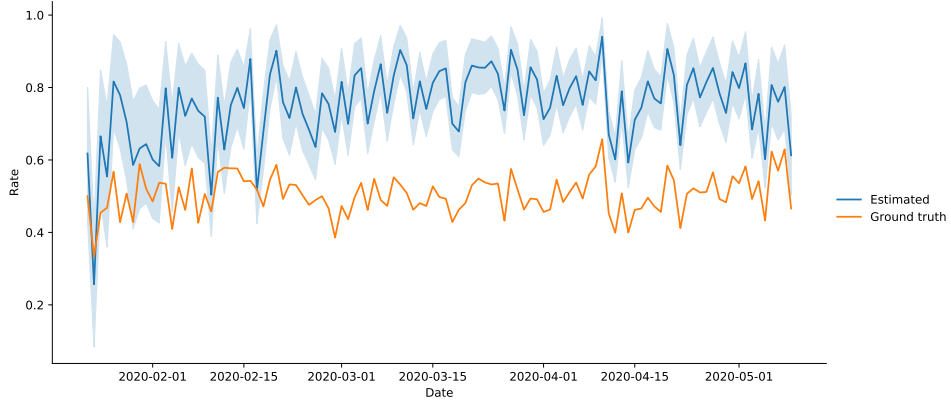


FIGURE 6. Estimated q for data generated with constant β and noisy q . The solid lines show means and shades with same colors show standard deviation. The orange line indicates the true q .

Table 1 shows that our model, **varied-q**, is the best model. However, we need to be careful because the difference between **const-q** is not large. Further, there is a data point in which the Pareto-k for the importance weight distribution is larger than 0.7. Therefore, the model is not robust and influenced too much from small numbers of data points.

Keeping this in mind, we tentatively choose **varied-q** and we analyze the result of **varied-q**.

Fig. 9 shows an estimated R for each day in Japan by **varied-q**. The first upward trend may not be very reliable because there is a few data for infection. The downward trend from the mid. February until the mid. March could be

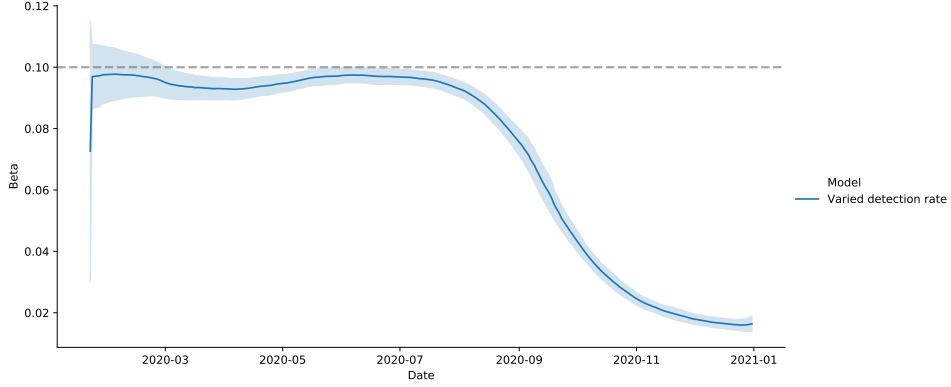


FIGURE 7. Estimated β for data generated with constant β and q which make most of population is infected. The blue line indicates the estimate of our model. The solid lines show means and shades with same colors show standard deviation. The dotted line indicates the true β .

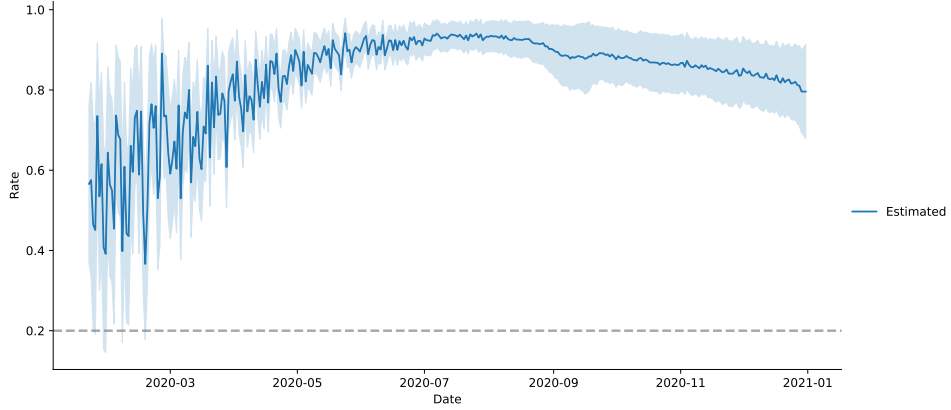


FIGURE 8. Estimated q for data generated with constant β and q which make most of population is infected. The solid lines show means and shades with same colors show standard deviation. The dotted line indicates the true q .

explained by public awareness and tracking effort of infection. After the mid. March, the tracking effort might be overwhelmed, thus created an upward trend until the beginning of April, when “the state of emergency” was declared to major urban areas. Since then, there was a downward trend, and now R is below 1. Thus, currently the infection is shrinking.

4.3. Sensitivity analysis. Before going analysis of other countries, we analyze sensitivity of the result to prior distributions using Japanese data as an example.

Fig. 9 shows an estimated R for each day in Japan by **varied- q** . The blue line shows the estimate of R based on the prior $\sigma_b \sim \text{Exponential}(1)$, which are used

Model	PSIS-LOO	DSE(PSIS-LOO)	WAIC	DSE(WAIC)
varied-q	792.209	0	720.258	0
const-q	809.442	6.44946	730.481	3.76378
const	5056.46	809.672	5768.74	1042.7

TABLE 1. Comparison of PSIS-LOO and WAIC for each model. **varied-q** is a model with varying β and q . **const-q** is a model with varying β and fixed q . **const** is a model with fixed β and q . DSE is the standard error of the distance from the best model.

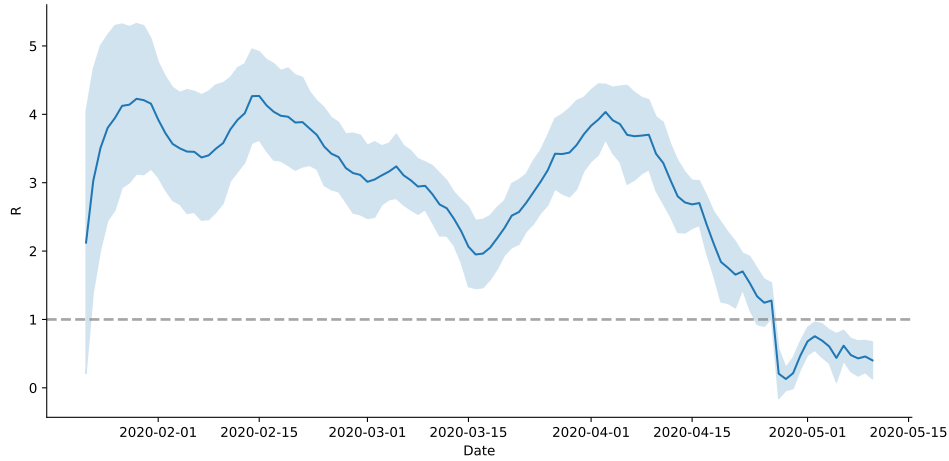


FIGURE 9. Daily change of R in Japan. The solid lines show means and shades with same colors show standard deviation.

throughout in this paper. The orange line shows the estimate of R based on the prior $\sigma_b \sim \text{Exponential}(0.1)$. There is no significant difference of estimates between two priors; therefore we can safely conclude that the effect of prior is small.

4.4. Comparison of Denmark and Sweden. Next, we compare Denmark and Sweden. These countries are economically and socially similar countries, but employed very different policies against COVID-19.

We applied the same procedure as Japan to data from two countries. All parameters are converged and information criteria favor **varied-q** model. For Sweden, we only applied data from Mar 1. because the confirmed cases only appear in February 27. The estimated R clearly shows that lock-down introduced March 13. was effective to put down R . However, in Sweden, which did not employ lock-down, R reduced gradually and now R is almost same between Denmark and Sweden. This might suggest that lock-down is not very effective in long run but might also due to unfavorable conditions to Denmark, for example, higher population density.

4.5. Multi-national comparison. Fig. 13 shows daily estimate of R of China, Italy, Japan and US. The same method as Japan was applied to the rest of countries and adequacy of **varied-q** model was verified. For Italy, we only apply data from

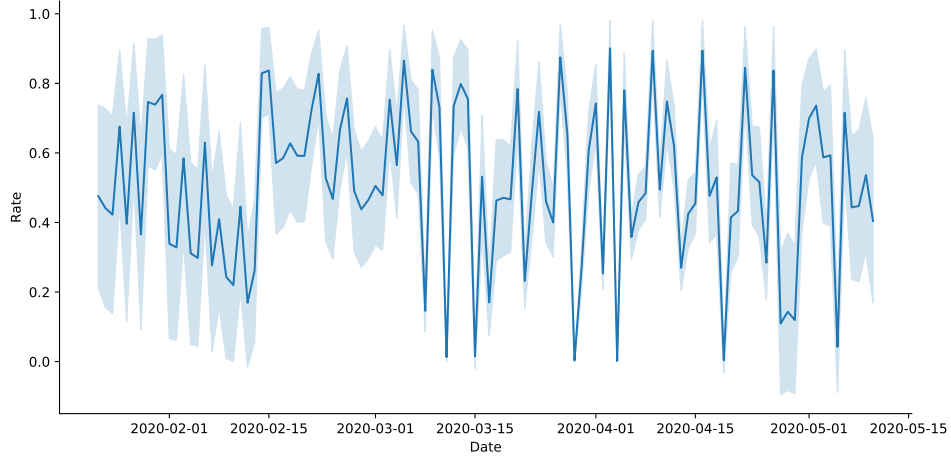


FIGURE 10. Daily change of the detection rate in Japan. The solid lines show means and shades with same colors show standard deviation.

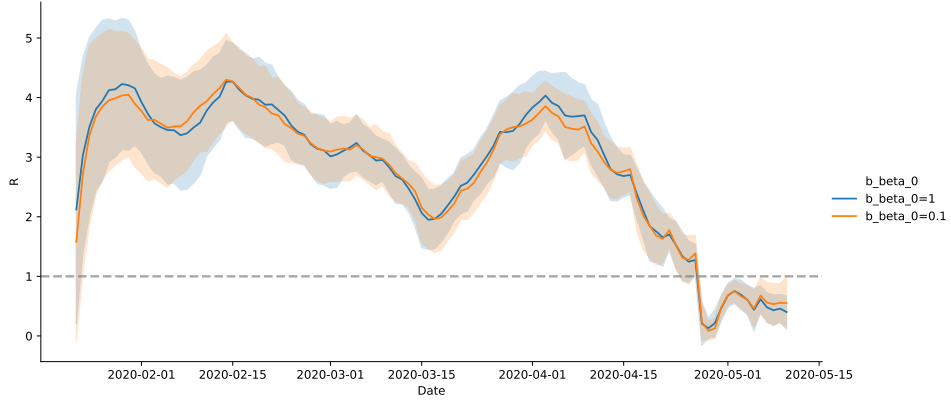


FIGURE 11. Daily estimate of R in Japan using different priors. The solid lines show means and shades with same colors show standard deviation.

March 1, 2020 because otherwise the parameter estimation did not converge. The same method was applied to Korea but the parameter estimation does not converge.

The results show that China, Italy, Japan, and the US are about to exit epidemic.

5. FUTURE WORK

The method used in this paper has several limitations.

First, as experiment used simulation data revealed, our method cannot determine the true level of detection rate, nor its long-term trends. To find the true detection rate, we need different kind of data, such as excess mortality.

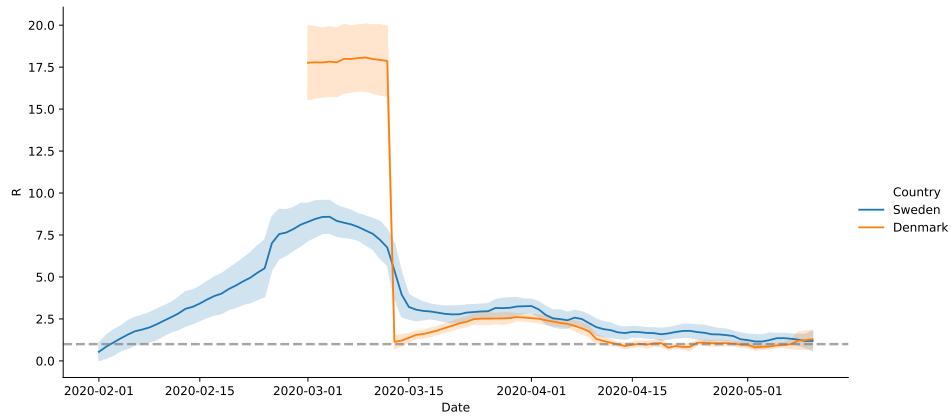


FIGURE 12. Daily estimate of R in Denmark and Sweden. The solid lines show means and shades with same colors show standard deviation.

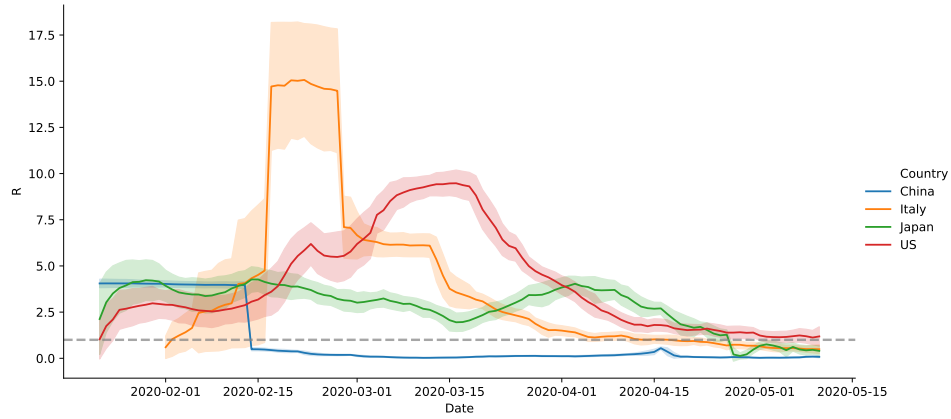


FIGURE 13. Daily estimate of R in several countries. The solid lines show means and shades with same colors show standard deviation.

Second, our model uses a naive SIR model and does not consider incubation period and reporting delay. We can reconstruct the date of exposure by back projection, so it would be interesting apply our method to such data.

ACKNOWLEDGEMENT

The author thanks to Kentaro Matsuura and the Tokyo.R Slack group for many suggestions and advice on Bayesian modeling. The author also thanks to Peter Turchin, from whose work the author's work started.

REFERENCES

- [1] Cleo Anastassopoulou, Lucia Russo, Athanasios Tsakris, and Constantinos Siettos. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *Plos One*, 15(3):e0230405, 2020.
- [2] Diego Caccavo. Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model. *medRxiv*, page 2020.03.19.20039388, 2020.
- [3] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76(1), 2017.
- [4] Seth Flaxman, Swapnil Mishra, Axel Gandy, Juliette T Unwin, Helen Coupland, Thomas A Mellan, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, Pablo N P Guzman, Nora Schmit, Lucia Cilloni, Kylie E C Ainslie, Marc Baguelin, Isobel Blake, Adhiratha Boonyasiri, Olivia Boyd, Lorenzo Cattarino, Constanze Ciavarella, Laura Cooper, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Bimandra Djaafara, Ilaria Dorigatti, Sabine Van Elsland, Rich Fitzjohn, Han Fu, Katy Gaythorpe, Lily Geidelberg, Nicholas Grassly, Will Green, Timothy Hallett, Arran Hamlet, Wes Hinsley, Ben Jeffrey, David Jorgensen, Edward Knock, Daniel Laydon, Gemma Nedjati-Gilani, Pierre Nouvellet, Kris Parag, Igor Siveroni, Hayley Thompson, Robert Verity, Erik Volz, Patrick Gt Walker, Caroline Walters, Haowei Wang, Yuanrong Wang, Oliver Watson, Xiaoyue Xi, Peter Winskill, Charles Whittaker, Azra Ghani, Christl A Donnelly, Steven Riley, Lucy C Okell, Michaela A C Vollmer, Neil M Ferguson, and Samir Bhatt. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. *Imperial College London*, (March):1–35, 2020.
- [5] Ravin Kumar, Colin Carroll, Ari Hartikainen, and Osvaldo A. Martin. ArviZ a unified library for exploratory analysis of Bayesian models in Python. *The Journal of Open Source Software*, 2019.
- [6] Peter Turchin. Analyzing covid-19 data with sird models, 2020.
- [7] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, 2017.
- [8] Sumio Watanabe. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11(Dec):3571–3594, 2010.

NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY (AIST), 1-8-31 MIDORIGAOKA, IKEDA, JAPAN

Email address: yoriyuki.yamagata@aist.go.jp