

Journal of the
Operational Research
Society



Bayesian inference of COVID-19 pandemic in the U.S. using adjusted SIR and SEIR modeling

Journal:	<i>Journal of the Operational Research Society</i>
Manuscript ID	TJOR-2020-OP-0768
Manuscript Type:	Original Paper
Date Submitted by the Author:	02-Aug-2020
Complete List of Authors:	Yang, Qianhui; University of North Carolina at Chapel Hill,
Keywords:	Forecasting, Markov Processes, Mathematical Programming
Abstract:	In the United States, 3,882,167 confirmed cases and 141,677 deaths of people who tested positive for COVID-19 by July 22, 2020. Epidemiology modelings are important for the authority to understand the pandemic and making trajectory decisions in the future. By performing Bayesian inference using Markov-Chain Monte-Carlo (MCMC) algorithm on the adjusted compartmental models with multiple sets of changing points in the U.S, we found three changing points coincides with the national lockdown and gradual reopen. This shows the evidence of national intervention and setback, following with the prediction of the length and the pandemic peak of different future scenarios. Overall, we hereby confirm the effect the government intervention and reinforce the importance of slowing down the national reopen to decrease the total mortality cases and shorten the length of the pandemic.

SCHOLARONE™
Manuscripts

ORIGINAL PAPER

Bayesian inference of COVID-19 pandemic in the U.S. using adjusted SIR and SEIR modeling

Qianhui Yang^a

^aDepartment of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, U.S.A.

ARTICLE HISTORY

Compiled August 2, 2020

ABSTRACT

In the United States, 3,882,167 confirmed cases and 141,677 deaths of people who tested positive for COVID-19 by July 22, 2020. Epidemiology modelings are important for the authority to understand the pandemic and making trajectory decisions in the future. By performing Bayesian inference using Markov-Chain Monte-Carlo (MCMC) algorithm on the adjusted compartmental models with multiple sets of changing points in the U.S, we found three changing points coincides with the national lockdown and gradual reopen. This shows the evidence of national intervention and setback, following with the prediction of the length and the pandemic peak of different future scenarios. Overall, we hereby confirm the effect the government intervention and reinforce the importance of slowing down the national reopen to decrease the total mortality cases and shorten the length of the pandemic.

KEYWORDS

COVID-19, U.S, SIR, SEIR, Bayesian inference, pandemic forecast, social distancing, government intervention

1. Introduction

The new corona-virus COVID-19 started in Wuhan, China last November has caused global pandemic Velavan and Meyer (2020); Wu and McGoogan (2020) , the World Health Organization declaring the cononavirus a pandemic on 11 March 2020 Cucinotta and Vanelli (2020) . By July 23rd, 14971036 people has been infected with 618017 fatalities; According to Chinese data, the novel strain SARS-CoV-2 shows a basic reproduction number (R_0) from 2.0 to 3.5 Read, Bridgen, Cummings, Ho, and Jewell (2020); Zhao et al. (2020) , which is higher than SARS Liu, Gayle, Wilder-Smith, and Rocklöv (2020) . In addition, the COVID-19 infected patients could show no symptom for up to 14 days Lauer et al. (2020) , making the transmission hard to prevent.

In the U.S, the pandemic is still not fully controlled with 3882167 people infected and 141677 death by July 22. The U.S government has been taken multiple actions since the outbreak from March 22, including staying at home order, face-covering required in public, gathering banned, out-of-state travel restriction and

CONTACT Qianhui Yang. Author. Email: jessieyy@ad.unc.edu

closure of the school, restaurant ,and retail orders.

However, the pandemic situation in the U.S is now more complicated with an upward spike that has a larger base number but having overwhelming limited medical resources countrywide. Making strategical decisions are going to be significantly important for the future of the country at this time-point.

Epidemiology models are fundamental to help us understand the situation of the pandemic and more importantly, it could provide a trustful prediction for the decision-makers in order to plan the best for the country Enserink and Kupferschmidt (2020); Kermack, McKendrick, and Walker (1927). Compartmental models are widely used in the infectious disease, SIR models are one of the most commonly used models Boudrioua and Boudrioua (2020); Roddam (2001). SIR models partition the population into three groups: susceptible, infected and recovered. Later, more models have been developed, the SEIR model adds the incubation period and reporting delay into account Li and Muldowney (1995); Peng, Yang, Zhang, Zhuge, and Hong (2020). The discrete-time SIR model adds deceased patients into account Casella (2020). $SVC_{Sy}C_{Asy}IR$ adds the vaccination and temporal immunity into account Otoo, Opoku, Charles, and Kingsley (2020). SIDARTHE model divides the infection into eight stages with different severity of symptoms Giordano et al. (2020).

In this study, we set up multiple models to test the underline epidemiology parameters and changing points, applying the Bayesian Markov-Chain Monte-Carlo (MCMC) algorithm for parameter estimation.

First, we employ the simplified exponential SIR model to compare the main parameters with the other published models; Then, we consider the delay reporting time, weekend modification, incubation time, and most importantly, the changing point into our models combined with SIR and SEIR modeling; Lastly, based on our inference and assume reasonable government intervention, we predict the peak date of the second outbreak and the possible overall infected cases in the U.S.

We discovered the evidence of lock-down and reopen of the U.S. during the COVID-19 pandemic. In addition, according to our prediction, only under strong government intervention could prevent the virus from spreading even further. Slowing down the country reopen seems the highest priority of the country before the pandemic gets worse.

2. Methods

2.1. SIR and SEIR Modeling and Adjustments

Compartmental models have been widely used for infectious disease simulation Blackwood and Childs (2018); Hethcote (2000). Among them, Susceptible-Infectious-Recover (SIR) and Susceptible-Exposed-Infectious-Recover (SEIR) are two most powerful and concise compartmental models in COVID-19 studies world wide Dehning et al. (2020); NGUEMDJO, MENO, DONGFACK, and VENTELOU (2020); Radulescu and Cavanagh (2020).

2.1.1. SIR and SEIR Modeling

In the SIR model, it is assumed three groups of individuals keep changing back and forth in the models, S(susceptible individuals), I (infected individuals) and R (recovered individuals) during time period of t , with total population of N and two parameters: spreading rate(λ) and recover rate (μ). SIR model can be described in three equations.

$$\frac{dS}{dt} = -\lambda \frac{SI}{N} \quad (1)$$

$$\frac{dI}{dt} = \lambda \frac{SI}{N} - \mu I \quad (2)$$

$$\frac{dR}{dt} = \mu I \quad (3)$$

In the SEIR model, between the stage of susceptible and infectious, a stage of exposed (E) is defined. A person could have 14 days of incubation period with no symptoms Lauer et al. (2020), during this period changing from being infected to become infectious with the rate of σ . Similarly, SEIR model can be described in four equations.

$$\frac{dS}{dt} = -\lambda \frac{SI}{N} \quad (4)$$

$$\frac{dE}{dt} = \lambda \frac{SI}{N} - \sigma E \quad (5)$$

$$\frac{dI}{dt} = \sigma E - \mu I \quad (6)$$

$$\frac{dR}{dt} = \mu I \quad (7)$$

2.1.2. Report Delay Adjustment, Change Points Modulation and weekend modulation

Our study applied the adjusted SIR and SEIR model considering the Report Delay D in the incoming infected cases Dehning et al. (2020) :

$$C_t = I_{t-D}^{new} \quad (8)$$

In addition, our study including a weekend modulation to alleviate difference of cases report between the weekend and the weekday Dehning et al. (2020):

$$f(t) = (1 - f_w) * (1 - |\sin(\frac{\pi}{7}t - \frac{1}{2}\Phi_w)|) \quad (9)$$

Due to the special situation in the U.S with country with fluctuation report cases and different status in states, in order to discover the best fit, multiple sets of changing points are tested in our study. We assume that the change of the spreading rate is linear and completed during the time period Δt Dehning et al. (2020):

$$C_t = I_{t-D}^{new}(1 - f(t)) \quad (10)$$

2.2. Bayesian Parameter Inference

Bayesian inference uses likelihood multiplies the prior distribution to determine the posterior distribution, prior is the distribution of the parameters before data, likelihood is the distribution of data condition on its parameter, posterior distribution is the distribution of the parameter after considering the data. In our study:

$$p(\Theta|\hat{C}) \propto p(\hat{C}|\Theta)p(\Theta) \quad (11)$$

where Θ represents the set of parameters and \hat{C} represents the real-world data.

2.2.1. Markov Chain Monte Carlo (MCMC) Method

Our study uses the python package PyMC3 with No-U-Turn(NUTS) algorithm, and multiple, independent Markov chains. Initialization of MCMC through automatic differentiation variational inference (ADVI). Totally 4 chains with each chains 1000 tuning and 1000 draw iterations. The target accept rate is set to 0.99, with the default accept rate in PyMC3 result in Pareto k diagnostic values higher than 0.7 when we add the weekend modification. The \hat{R} is verified to be below 1.05 in PyMC3.

2.2.2. The Likelihood

Our study applies the student's t-distribution as the likelihood function to make the MCMC more robust with regarding to the outliers.

$$p(\hat{C}|\Theta) \sim t(\hat{C}(\Theta), \sigma\sqrt{\hat{C}(\Theta)}) \quad (12)$$

2.2.3. Priors

Our study applies the simple SIR modeling at the onset phase in the U.S to obtain the initial spreading rate(λ), recover rate(μ) and reporting delay(D), and later using this set of initial parameter to further estimate the parameters after the changing points. According to Dehning et al. (2020), in order to emphasize the differences between the spreading rate and the recover rate, these two parameters are chosen to become a narrow log-normal

$$(\mu \sim \text{LogN}(\log(\frac{1}{8}), 0.2))$$

and a wide log-normal($\sigma \sim \text{LogN}(\log(0.5), 0.2)$). The reporting delay is also considered as a $\sim \text{LogN}(\log(8), 0.2)$. The initial infected $I_0 \sim \text{HalfCauchy}(100)$ is set to distributed as the Half-Cauchy to make sure the magnitude does not exceeded the parameter, similar for the scale factor $\sigma \sim \text{HalfCauchy}(10)$.

In our final model, several sets of changing points have been tried to test for the best model. Assuming that the recover rate keeps the same(μ), in the one changing point model, we assume that λ_1 decreases 50 percent to 0.2 after the first changing point. In the two changing point model, we assume that λ_2 goes back to λ_0 after the second changing point. In the three changing point model, we assume that λ_3 increase to 0.5 after the third changing point. In the four changing point model, we assume that λ_3 decrease to 0.2 and λ_4 then increase to 0.5.

In each changing points, the time period Δt is set to distributed as $LogN(log(3), 0.3)$. In weekly modulation as we mentioned in the eqnarray $f(t) = (1 - f_w) * (1 - |sin(\frac{\pi}{7}t - \frac{1}{2}\Phi_w)|)$, the amplitude $f_w \sim Beta(0.7, 0.17)$ and the phase $\Phi_w \sim vonMises(0, 0.01)$. All of our changing points are selected Monday of the week to make sure the model are used properly.

2.2.4. Model Assessment, Selection, Inference after Selection

Our study applies Bayesian leave-one-out(LOO) cross-validation to assess the best fit model among different changing points sets. The almost unbiased estimate for the model.

$$\hat{elpd}_{LOO} = \frac{1}{n} \sum_{i=1}^n \log p(y_i | x_i, C_{-1}, M_k) \quad (13)$$

Where M_k represents K-fold cross-validation. The fitting is also estimated by shape parameter k by fitting a generalized Pareto distribution to the tail of the weight distribution, the model with $k > \frac{1}{2}$ are considered bad fitting in our study.

2.2.5. Data and Programming

Data are retrieved and then filtered from the online repository of the coronavirus visual dashboard managed by John Hopkins University (<https://coronavirus.jhu.edu/>); All the programming are performed in Jupyter Notebook and table are formed in SAS; The Code can be accessed on Github(<https://github.com/Jessie1024>).

3. Results

3.1. Initial inference by simple SIR modeling during the beginning of the outbreak

First, we start with the simple SIR modeling from the beginning of the outbreak. We perform Bayesian inference by Markov-Chain Monte-Carlo(MCMC) sampling on the spreading rate (λ), recover rate(μ), Delay period (D), initial infected number (I_0) and scale factor (σ) (**FIG 1**).

The initial model estimates the median of spreading rate $\lambda = 0.46$, recovery rate $\mu = 0.12$, number of infections at begin $I_0 = 46$, delay time $D = 8.5$, scale factor $\sigma = 8$, reproductive number $R_0 = 3.79$.

These parameters match the published material, where the reproductive number of COVID-19 in the U.S. was reported between 2.6 and 4.0 Gunzler and Sehgal (2020); Shao, Cheng, and Chen (2020); Zhou et al. (2020), and delay time 8.5 days makes sense as for the sum of incubation and testing period.

From both the log plot and linear plot, the exponential model fits, which indicates an exponential growth at the beginning of the outbreak in the U.S.(**FIG 1. A-C**).

3.2. SIR and SEIR Modeling with multiple changing points and modifications

Due to the complexity of the situation in the *U.S.* with multiple outbreaks in different states, our study applies models with zero to four changing points with the weekend modification and delay, then compare their model fitness (**Table 1**).

The results show that under the 85% acceptance rate, the two changing points model and the three changing point model fits the best situation in the *U.S.* while the simple SIR model contained no changing point shows the lowest fitness level. The simple SIR model results in $\hat{elpd}_{LOO} = -1122.43$, with $SE = 12.37$, $p_{loo} = 9.86$ and $pareto_k = 97.3\%$. The best fit four models have the changing point on March 23 and May 4, March 23 and May 18, March 23 and May 11, Mar 23 May 4 and June 8, with the $\hat{elpd}_{LOO} = -1050.05, -1059.14, -1043.83, -1035.58$. The higher the \hat{elpd}_{LOO} , the lower the SE, the higher the p_{loo} and the greater the $pareto_k$, the better the model fit the data Vehtari, Gelman, and Gabry (2017). These changing points match the *U.S.* shutting down and reopening, which is around March 23 and reopen at the beginning of May.

To precisely estimate the parameter and especially fix the low $pareto_k$ to up to 100%, we change the acceptance rate to 99% for the seven of the models (**Table 2**), this results in an overall $pareto_k$ increases, SE decreases and p_{loo} increases. The SIR models with the changing point March 23, May 4 and June 8, changing point March 23, April 27 and June 15, changing point March 23 and May 4, changing point March 23 and May 18 and changing point March 16 and May 4 has $\hat{elpd}_{LOO} = -1114.16, -1115.67, -1131.72, -1120, 1161.06$. The SEIR models with the changing point March 23 and May 4 and changing point March 23, May 4 and June 8 have $\hat{elpd}_{LOO} = -1094.81, -1124.53$.

In the SIR model at changing point Mar 23 and May 4, the inferred median of spreading rate went from 0.48 [0.4, 0.58] to 0.26 [0.19, 0.35] to 0.30 [0.23, 0.40], while the recover rate is 0.26, the inferred changing point in this model is March 13 [Mar 8, Mar 15] and May 22 [May 17, May 24] (**Supplementary 1**);

In the SIR model at set changing point at Mar 23 and May 11, the inferred changing point in this model is March 15 [March 10, Mar 20] and May 29 [May 24, May 31], the inferred median of spreading rate went from 0.48 [0.41, 0.58] to 0.24 [0.19, 0.33] then comes back to 0.28 [0.23, 0.37], the median of the recovered rate is 0.24 [0.19, 0.33] (**Supplementary 2**);

In the SIR model at set changing point at Mar 23, May 4 and June 8, the inferred changing point in this model is March 17 [March 14, Mar 22], May 3 [May 1, May 5], June 6 [June 4, June 8] inferred median of spreading rate went from 0.44 [0.38, 0.54] to 0.19 [0.15, 0.25] then comes back to 0.24 [0.2, 0.3], the median of the recovered rate is 0.20 [0.16, 0.25] (**Supplementary 3**);

In the SIR model at set changing point at Mar 23 Apr 27 Jun 15, the inferred

changing point in this model is March 17 [March 13, Mar 21] and Apr 26 [Apr 24, Apr 28] and Jun 10 [Jun 8, Jun 12] the inferred median of spreading rate went from 0.41 [0.36, 0.48] to 0.15 [0.12, 0.20] then comes back to 0.20 [0.16, 0.25], the median of the recovered rate is 0.14 [0.12, 0.21](**Supplementary 4**);

In the SEIR model at set changing point at Mar 23 and May 11, the inferred changing point in this model is March 15[March 10, Mar 20] and May 10 [May 7, May 13], the inferred median of spreading rate went from 2.02 [1.22, 3.44] to 0.62[0.37, 0.92] then comes back to 0.58 [0.30, 0.89], the median of the recovered rate is 0.4[0.27, 0.58](**Supplementary 5**);

In the SEIR model at set changing point at Mar 23, May 4 and June 8, the inferred changing point in this model is March 22 [March 15, Mar 24] ,May 3 [May 1, May 5] and June 5[June 3, June 7], the inferred median of spreading rate went from 1.23[0.78, 1.68] to 0.36[0.26, 0.53] then comes back to 0.49 [0.35, 0.70], the median of the recovered rate is 0.37[0.27, 0.54](**Supplementary 6**);

Overall, we conclude from our best fit models that the first changing point of the U.S pandemic is around Mar 20 and the second changing point is between the end of May to the beginning of the June, the spreading rate went from significantly greater than recover rate to lower than recover rate and eventually increases back to slightly higher than the recover rate(**Table 4**);

3.3. Forecast of the ending of the pandemic and the total infected and death number in the U.S.

Applying the models above, we predict the future of the U.S pandemic in the two months (**FIG 3,4**). The SEIR models with three changing points predict the daily new cases will increase to 250 K at the middle of the August and the cumulative cases are going to reach 5 million at the beginning of August; Another SEIR model with two changing points predicts the daily new cases will increase to 150K at the middle of the August and the cumulative cases are going to reach 7 million at the middle of the August(**FIG 3**);

Three of the SIR models with two changing points and three changing points predict the daily new cases increases to 250K at the middle of the August and the cumulative cases are going to reach 7 million at the beginning of the August (**FIG 4**);

However, since U.S government already start the second round of intervention, here we use the simple exponential SIR model assuming the spreading rate decreases 10%, 20%, 15%, 25% at the length of 3 days and 30 days (**FIG 5**)(**Table 3**). In the case of strong intervention which we assume the spreading rate decrease between 15 and 25 percent, the peak will arrive before the middle of August with around 100K cases per day; The prediction of the total infected cases till the end of the pandemic will be less than ten million if the actions are taken quickly and the decrease rate is over than 20%.

However, if the actions are taken slow or the decrease rate is lower than 20%,

it is highly possible that the growth of this pandemic is going to last more than a month and the total infected number could be extremely high in this scenario. If the decrease rate is only 10%, the total cases will be over 40 million; If the decrease rate is 15%, the total cases will be over 20 million.

4. Discussion

Our study employs the Bayesian algorithm with the adjusted SIR and SEIR model to infer the fundamental epidemiology parameter in the U.S. COVID-19 pandemic from Mar 8 to July 8, predicting the peak and the ending of this round of outbreak.

At the beginning of the outbreak from March 8th to March 23rd, the median of the reproductive number R_0 is 3.79, which means about 4 people will contract COVID-19 from one person with the disease. This is extremely high in the U.S compared to other infectious disease such as MERs and H1N1 Cowling et al. (2010) . The delay time is 8.5, technically, this including the incubation period and the time between the testing and result. The scale factor of 8 means that for each reported infected case, there are eight cases that show no symptoms but can still be infectious. This matches the WHO report with a factor of 10 Team et al. (2020) (**FIG 1**).

The U.S. government starts the lock-down from the middle of March, SIR and SEIR models start to play its role after the lock-down. At first, multiple adjusted models with the different number changing points are tested and the two CP(changing points) and three CP models show to fit the best of the U.S. situation with larger $epld_{100}$, lower SE of $epld_{100}$ and higher p_{100} and $pareto_k$ (**Table 1**).

We further increase the acceptance rate from the default 85%to99% since the $pareto_k$ rate seems low in the default modeling. We fit multiple sets of changing points(only Monday, because of the weekend modification) to see the best fit models in the U.S. situation, which results in many good fit models with high $epld_{100}$, p_{100} , $pareto_k$ and low SE of $epld_{100}$. It is hard to pick the best models at this time point because they all have their own advantages and explanation in the real world. For example, the *SIRMar23Apr27June15modeling* infers that the spreading rate decrease from 0.41 to 0.15 then back 0.20 at the changing point at Mar 17, Apr 26 and June 10. This matches the country lock-down and reopens in stages. The *SEIRMar23May4June8modeling* infers that the spreading rate decrease from 1.23 to 0.36 then comes back to 0.49 on March 22, May 3 and June 5, which also matches the U.S. situation that we have mentioned above. All of these models show the spreading rate went from over the recovery rate to below it at the lock-down point and then go above the recovery rate after the country reopens. This explains that the pandemic is closely related to the government interventions with the re-open increase the spreading rate and worse pandemic situation at the U.S (**Table 3**) .

The last part of our study is the prediction with our presented parameters try to answer the following question: When will this pandemic end and how many people will end up infected in this pandemic? To answer these questions, we assume that the country will take stronger intervention and slow down the reopen stages. The result indicates that the peak of the pandemic extremely relate to the government strategy, with very little decrease of spreading rate affect thousands of infected cases

and monthly shorter of the pandemic period. Here our model predicts that in the case of strong intervention, the peak will arrive in the middle of August with less than 10 million infected in the end. However, if the situation stays uncontrolled, the pandemic could possibly last to next year with a total infected number of over 40 million in the end. If the death rate stays the same with 3.6%, the final death cases would be 1.44 million.

5. Conclusion

We utilize the adjusted SIR and SEIR models to help us understand the current U.S. pandemic situation. The changing points of the model coincide with the time of government intervention and actions during the pandemic. Two or three changing points since the March 8th outbreak of COVID-19, first decrease from $R_0 > 1$ to $R_0 < 1$ at around March 15, second increase from $R_0 < 1$ to $R_0 > 1$ at the May. These two changing points coincide with the country’s lock-down and reopen date in the U.S.

Our study also predicts that without strong intervention, the total death could up to 1.44 million and the pandemic will not end until February next year. Slowing down the country reopen and keeping a social distance is vital for the country.

A future improvement of the study including develops the model in states and cities, adding the vaccination in the model ,and making more prediction models base on more complicated scenarios.

6. Acknowledgement

The author would like to thank her mentor Dr. Kosorok, for his encourage and support.

7. Disclosure statement

The author declares no conflicts of interest in preparing this article.

8. Funding

The author receive no specific funding for this work.

References

Blackwood, J., & Childs, L. (2018, Dec.). An introduction to compartmental modeling for the budding infectious disease modeler. *Letters in Biomathematics*, 5(1), 195-221. Retrieved from <https://lettersinbiomath.journals.publicknowledgeproject.org/index.php/lib/article/view/81>

Boudrioua, M. S., & Boudrioua, A. (2020). Predicting the covid-19 epidemic in algeria using the sir model. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/early/2020/06/06/2020.04.25.20079467>

- Casella, F. (2020). Can the covid-19 epidemic be controlled on the basis of daily test reports? *IEEE Control Systems Letters*, 1-1.
- Cowling, B. J., Lau, M. S. Y., Ho, L.-M., Chuang, S.-K., Tsang, T., Liu, S.-H., ... Lau, E. H. Y. (2010, nov). The effective reproduction number of pandemic influenza: prospective estimation. *Epidemiology (Cambridge, Mass.)*, 21(6), 842-846. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/20805752>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3084966/>
- Cucinotta, D., & Vanelli, M. (2020, Mar.). Who declares covid-19 a pandemic. *Acta Bio Medica Atenei Parmensis*, 91(1), 157-160. Retrieved from <https://www.mattioli1885journals.com/index.php/actabiomedica/article/view/9397>
- Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., & Priesemann, V. (2020). Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. *Science*. Retrieved from <https://science.sciencemag.org/content/early/2020/05/14/science.abb9789>
- Enserink, M., & Kupferschmidt, K. (2020). With covid-19, modeling takes on life and death importance. *Science*, 367(6485), 1414-1415. Retrieved from <https://science.sciencemag.org/content/367/6485/1414.2>
- Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., & Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 26(6), 855-860. Retrieved from <https://doi.org/10.1038/s41591-020-0883-7>
- Gunzler, D., & Sehgal, A. R. (2020). Time-varying covid-19 reproduction number in the united states. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/early/2020/04/15/2020.04.10.20060863>
- Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42(4), 599-653. Retrieved from <https://doi.org/10.1137/S0036144500371907>
- Kermack, W. O., McKendrick, A. G., & Walker, G. T. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772), 700-721. Retrieved from <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1927.0118>
- Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., ... Lessler, J. (2020, may). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of internal medicine*, 172(9), 577-582. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/32150748>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7081172/>
- Li, M. Y., & Muldowney, J. S. (1995). Global stability for the seir model in epidemiology. *Mathematical Biosciences*, 125(2), 155 - 164. Retrieved from <http://www.sciencedirect.com/science/article/pii/0025556495927565>
- Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklöv, J. (2020, 02). The reproductive number of covid-19 is higher compared to sars coronavirus. *Journal of Travel Medicine*, 27(2). Retrieved from <https://doi.org/10.1093/jtm/taaa021> (taaa021)
- NGUEMDJO, U. K., MENO, F., DONGFACK, A., & VENTELOU, B. (2020). Simulating the progression of the covid-19 disease in cameroon using sir models. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/early/2020/05/21/2020.05.18.20105551>
- Otoo, D., Opoku, P., Charles, S., & Kingsley, A. P. (2020). Deterministic epidemic model for (svcsycasyir) pneumonia dynamics, with vaccination and temporal immunity. *Infectious Disease Modelling*, 5, 42 - 60. Retrieved from <http://www.sciencedirect.com/science/article/pii/S2468042719300399>
- Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). *Epidemic analysis of covid-19 in china by dynamical modeling*.
- Radulescu, A., & Cavanagh, K. (2020). *Management strategies in a seir model of covid 19 community spread*.
- Read, J. M., Bridgen, J. R., Cummings, D. A., Ho, A., & Jewell, C. P.

(2020). Novel coronavirus 2019-ncov: early estimation of epidemiological parameters and epidemic predictions. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/early/2020/01/28/2020.01.23.20018549>

Roddam, A. W. (2001, 02). Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation: O diekmann and jap heesterbeek, 2000, chichester: John wiley pp. 303, £39.95. isbn 0-471-49241-8. *International Journal of Epidemiology*, 30(1), 186-186. Retrieved from <https://doi.org/10.1093/ije/30.1.186>

Shao, N., Cheng, J., & Chen, W. (2020). The reproductive number r_0 of covid-19 based on estimate of a statistical time delay dynamical system. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/early/2020/02/20/2020.02.17.20023747.1>

Team, C. D. C. C.-. R., Jorden, M. A., Rudman, S. L., Villarino, E., Hoferka, S., Patel, M. T., ... Starita, L. M. (2020, jun). Evidence for Limited Early Spread of COVID-19 Within the United States, January-February 2020. *MMWR. Morbidity and mortality weekly report*, 69(22), 680-684. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/32497028> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7315848/>

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27, 1413-1432.

Velavan, T. P., & Meyer, C. G. (2020). The covid-19 epidemic. *Tropical Medicine & International Health*, 25(3), 278-280. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/tmi.13383>

Wu, Z., & McGoogan, J. M. (2020, 04). Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: Summary of a report of 72FIX ME!!!!314 cases from the chinese center for disease control and prevention. *JAMA*, 323(13), 1239-1242. Retrieved from <https://doi.org/10.1001/jama.2020.2648>

Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... Wang, M. H. (2020, mar). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*, 92, 214-217. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/32007643> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7110798/>

Zhou, X., Ma, X., Hong, N., Su, L., Ma, Y., He, J., ... Long, Y. (2020). Forecasting the worldwide spread of covid-19 based on logistic model and seir model. *medRxiv*. Retrieved from <https://www.medrxiv.org/content/early/2020/04/08/2020.03.26.20044289>

Appendices

Supplement 1. The Distribution of prior and posterior parameters for SIR model with changing point March 23 and May 4. **Supplement 2.** The Distribution of prior and posterior parameters for SIR model with changing point March 23 and May 11. **Supplement 3.** The Distribution of prior and posterior parameters for SIR model with changing point March 23, May 4 and June 8. **Supplement 4.** The Distribution of prior and posterior parameters for SIR model with changing point March 23, April 27 and June 15 . **Supplement 5.** The Distribution of prior and posterior parameters for SEIR model with changing point March 23 and May 11. **Supplement 6.** The Distribution of prior and posterior parameters for SIR model with changing point March 23, May 4 and June 8.

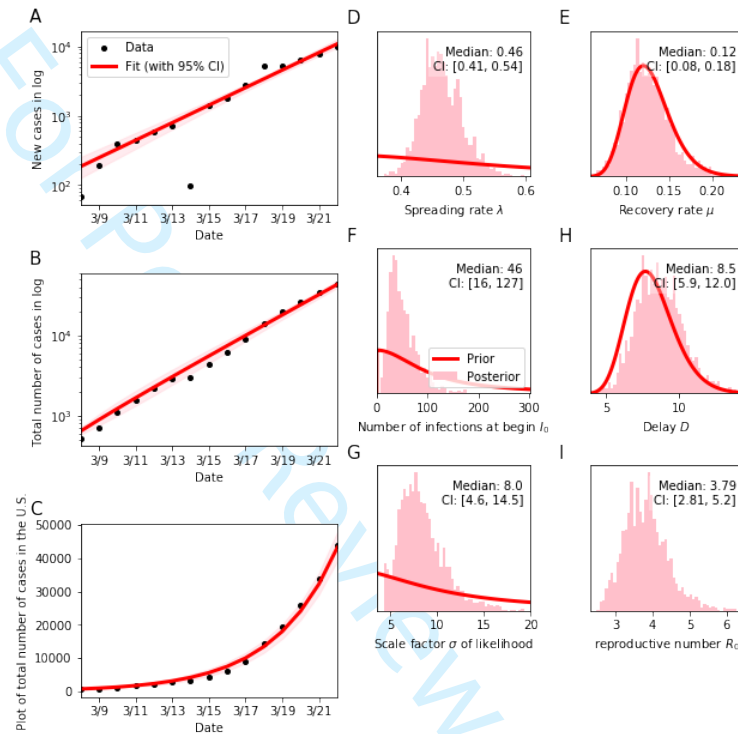


Figure 1. Initial inference by simple SIR modeling from March 8th to March 23rd. **A.** The number of new cases in the U.S. in log scale. **B.** The total number of cases in the U.S. in log scale. **C.** The total number of cases in linear scale. **D.** The distribution of prior and posterior of the estimated spreading rate λ . **E.** The distribution of prior and posterior for the estimated recovery rate μ . **F.** The distribution of prior and posterior of the estimated number of infection at beginning I_0 . **H.** The distribution of prior and posterior of the estimated number of delay (D). **G.** The distribution of prior and posterior of the estimated scale factor of likelihood **I.** The distribution of the reproductive rate R_0 .

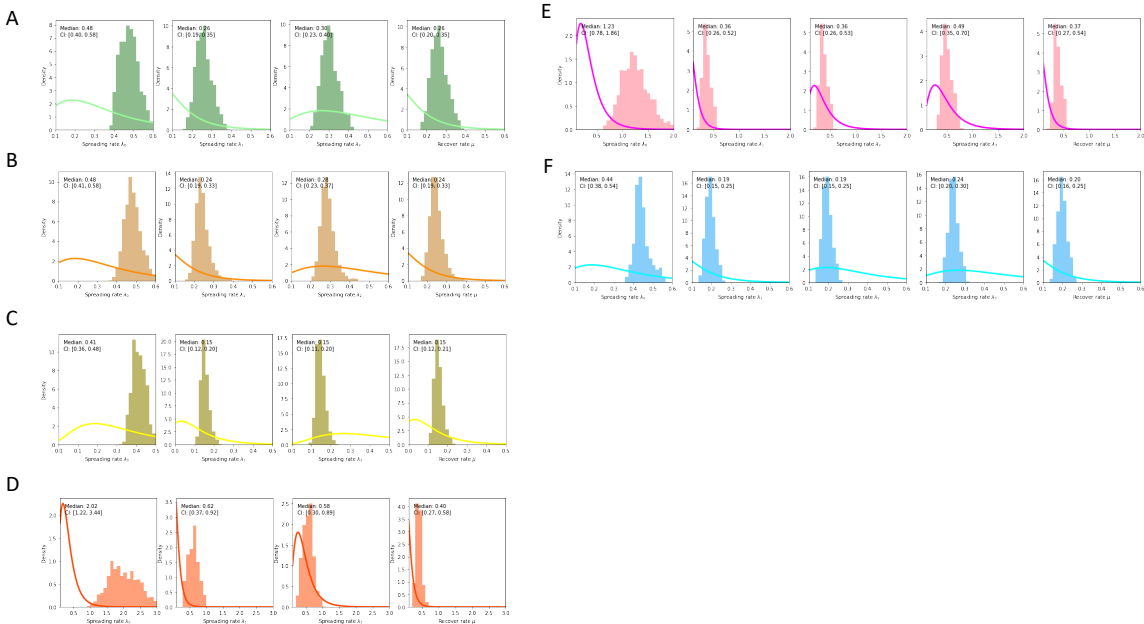


Figure 2. The distribution of the recover rate and spreading rate in the following models **A.** SIR model changing point at Mar 23 and May 4; **B.** SIR model changing point at Mar 23 and May 4; **C.** SIR model changing point at Mar 23, Apr 27 and June 15; **D.** SEIR model changing point at Mar 23, May 11; **E.** SEIR changing point at Mar 23, May 4 and June 8.

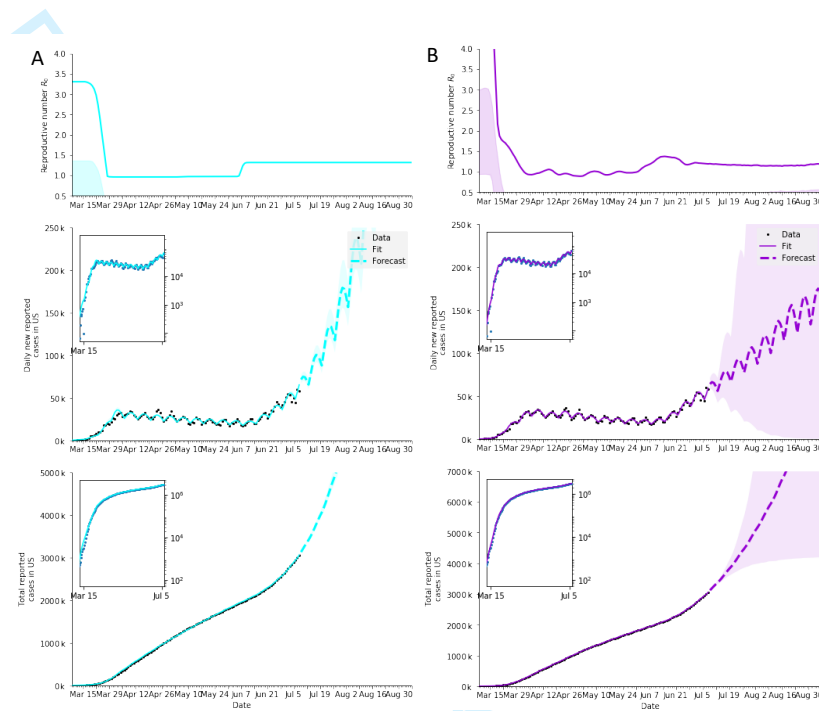


Figure 3. SEIR model fit with multiple changing points **A.** SEIR model with changing points at March 23, May 4 and June 8 (**Top**) reproductive rate changing (**Middle**) daily new cases in the U.S. fit and reported data (**Bottom**) Total new cases in the U.S. fit and reported data **B.** SEIR model with changing points at March 23, May 11 (**Top**) reproductive rate changing (**Middle**) daily new cases in the U.S. fit and reported data (**Bottom**) Total new cases in the U.S. fit and reported data.

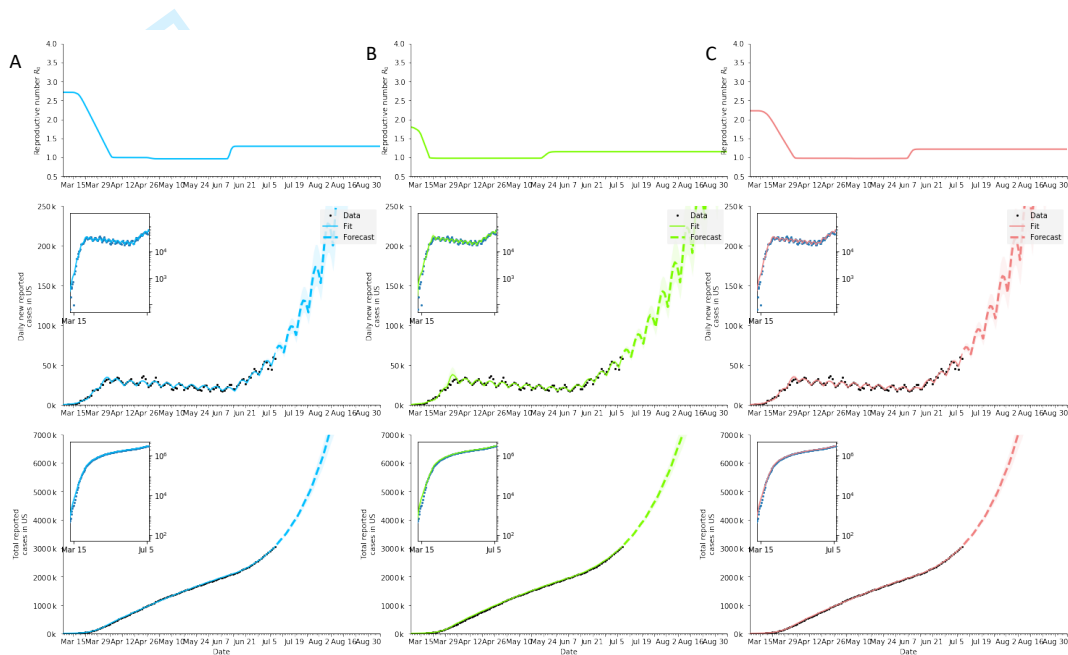


Figure 4. SIR model fit with multiple changing points **A.** SIR model with changing points at March 23, April 27 and June 15 (**Top**) reproductive rate changing (**Middle**) daily new cases in the U.S. fit and reported data (**Bottom**)Total new cases in the U.S. fit and reported data **B.** SIR model with changing points at March 23, May 4 and June 8 (**Top**)reproductive rate changing (**Middle**) daily new cases in the U.S. fit and reported data (**Bottom**)Total new cases in the U.S. fit and reported data. **C.**SIR model with changing points at March 23, May 4 and June 8 (**Top**)reproductive rate changing (**Middle**) daily new cases in the U.S. fit and reported data (**Bottom**)Total new cases in the U.S. fit and reported data.

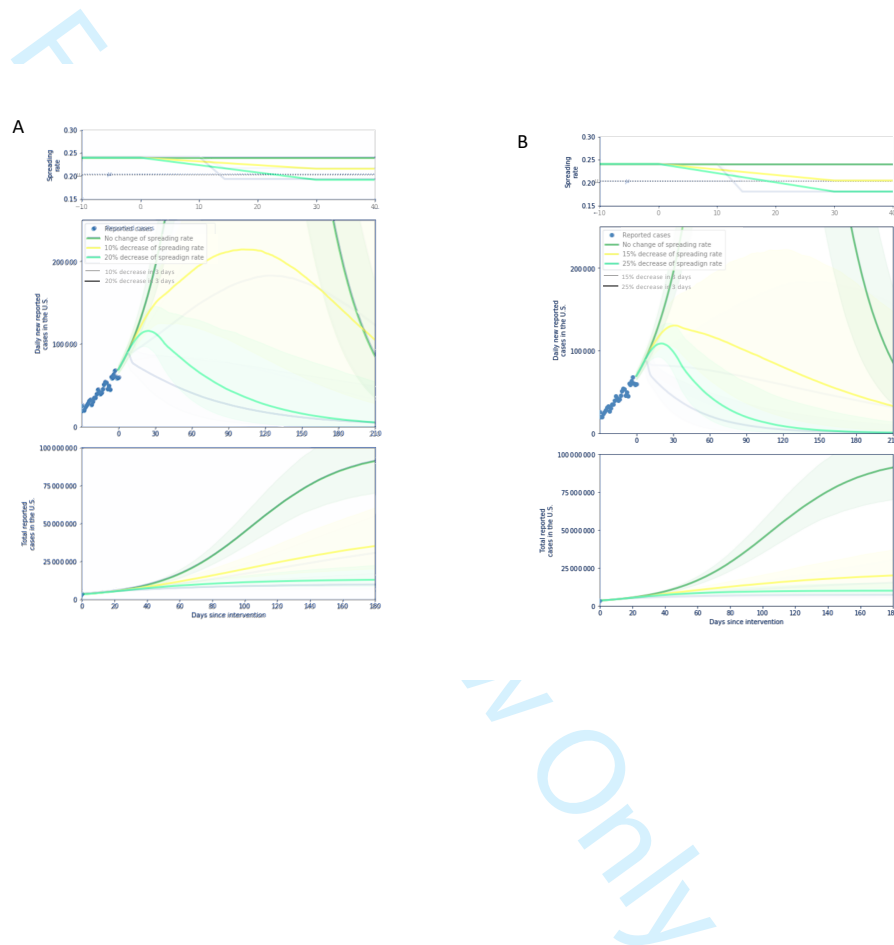


Figure 5. COVID-19 Pandemic prediction in the following 6 months since July 13 **A.** Pandemic prediction if spreading rate decrease 10% or 20% in the length of 3 days and 30 days **(Top)** Change of Spreading rate λ **(Middle)** daily new cases in the U.S **(Bottom)** Total infected cases in the U.S **B.** prediction if spreading rate decrease 15% or 25% in the length of 3 days and 30 days **(Top)** Changing of Spreading rate λ **(Middle)** daily new cases in the U.S **(Bottom)** Total infected cases in the U.S.

Model assessment with default 85% acceptance rate

Model	elpd_loo	SE	p_loo	pareto_k
SEIR_Mar_23_May_11_modified	- 1048.9	10.32	12.63	100%
SEIR_Mar_23_May_4_June_8_modified	- 1069.84	9.84	15.39	98.30%
SIR_Simple	- 1122.43	12.37	9.86	100%
SIR_Mar_23	- 1123.01	12.31	10.19	97.30%
SIR_Mar_23_Apr_27	- 1080.16	9.78	17.65	98.20%
SIR_Mar_23_Apr_27_June_15	- 1078.72	9.06	14.15	92.10%
SIR_Mar_23_Apr_20_May_11_June_15	- 1091.71	8.92	17.48	95.60%
SIR_Mar_23_May_4_June_8_modified	- 1035.58	10.1	12.45	93%
SIR_Mar_23_May_11_modified	- 1043.83	9.99	15.74	91.30%
SIR_Mar_23_May_18_modified	- 1059.14	10.42	21.31	94.80%
SIR_Mar_23_May_4_modified	- 1050.05	10.78	23.62	95.60%
SIR_Mar_23_May_4_June_8	- 1096.84	9.58	15.05	95.70%

Model assessment with 99% acceptance rate

Model	epld_loo	SE	p_loo	Pareto_k
SIR_Mar_23_May_4_June_8	- 1114.16	10.05	14.86	100%
SIR_Mar_23_Apr_27_June_15	- 1115.67	10.04	14.63	100.00%
SEIR_Mar_23_May_11	- 1094.81	11.09	27.01	95.90%
SIR_Mar_23_May_4	- 1131.72	11.28	24.48	95.10%
SIR_Mar_23_May_18	- 1120	10.39	19.79	92.60%
SEIR_Mar_23_May_4_June_8	- 1124.53	10.14	16.14	98.40%
SIR_Mar_16_May_4	- 1161.06	10.22	13.76	97.60%

Model with different changing point	R0	R1	R2	R3	R0 CI		R1 CI		R2 CI		R3 CI	
					Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
SIR_Mar_23_May_4_June_8	2.21	0.98	0.97	1.22	1.00	2.38	1.00	1.56	0.94	1.00	0.13	1.25
SIR_Mar_23_Apr_27_June_15	2.75	1.00	0.97	1.31	0.95	3.00	0.95	1.67	0.92	0.95	0.08	1.33
SEIR_Mar_23_May_11	5.06	1.56	1.43	0.00	1.59	4.52	1.59	3.41	1.11	1.53	0.00	0.00
SIR_Mar_23_May_4	1.81	0.98	1.14	0.00	1.00	2.00	1.00	1.75	1.15	1.14	0.00	0.00
SIR_Mar_23_May_18	1.97	0.98	1.15	0.00	1.00	2.16	1.00	1.74	1.21	1.12	0.00	0.00
SEIR_Mar_23_May_4_June_8	3.32	0.96	0.97	1.32	0.96	2.11	0.96	1.41	0.70	0.98	0.33	0.95

Predicted infected cases in 7 months in the U.S

model	Peak_time	Peak_No	month1	month2	month3	month4	month5	month6	month7
10% decrease in 3 days	125	182043	6242677	10008594	14726237	20035753	25424012	30329975	34549614
20% decrease in 3 days	8	87670	5610042	7102090	8091800	8717273	9121404	9372987	9535499
15% decrease in 3 days	10	85053	5893719	8263449	10432273	12324646	13912439	15296902	16352247
25% decrease in 3 days	5	82771	5375156	6300915	6739795	6947088	7048634	7091029	7106595
15% decrease in 30 days	30	130198	6658892	10382616	13650081	16408713	18444768	20032300	21266595
25% decrease in 30 days	20	108518	6380108	8397770	9334070	9752222	9930236	9996095	10027156
10% decrease in 30 days	105	214481	6814521	11902020	17971751	24472402	30389125	35303001	39199275
20% decrease in 30 days	30	113535	6513254	9245983	11008954	12065592	12700771	13088010	13333239
no change	110	855951

