medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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-ND 4.0 International license .

Are we there yet? An adaptive SIR model for continuous estimation of COVID-19 infection rate and reproduction number in the United States

Mark B Shapiro, Fazle Karim, Guido Muscioni, Abel Saju Augustine
Anthem, Inc., USA

Abstract

The dynamics of the COVID-19 epidemic vary due to local population density and policy measures. When making decisions, policy makers consider an estimate of the effective reproduction number $\mathcal{R}_t$ which is the expected number of secondary infections by a single infected individual. We propose a simple method for estimating the time-varying infection rate and reproduction number $\mathcal{R}_t$ using a sliding window approach applied to a Susceptible-Infectious-Removed model. The infection rate is estimated using the reported cases for a seven-day window to obtain continuous estimation of $\mathcal{R}_t$. We demonstrate that the proposed adaptive SIR (*aSIR*) model can quickly adapt to an increase in the number of tests and associated increase in the reported cases of infections. Our results also suggest that intensive testing may be one of the effective methods of reducing $\mathcal{R}_t$. The *aSIR* model was applied to data at the state and county levels.

Introduction

We are in the middle of a global COVID-19 pandemic caused by the SARS-CoV-2 virus. As of September 2, 2020, over 6 million individuals in the United States have been reported positive for SARS-CoV-2. Modeling studies are key for understanding factors that drive the spread of the disease and for developing mitigation strategies. Early modeling efforts forecasted very large numbers of infected individuals which would overwhelm healthcare systems in many countries (Arenas et al., 2020; Ferguson et al., 2020; Mitjà et al., 2020). These forecasts served as a call to action for policy makers to introduce policy measures including social distancing, travel restrictions, and eventually lockdowns to avoid the predicted catastrophe (Adam, 2020; Anderson et al., 2020; Enserink & Kupferschmidt, 2020). The mitigating policy measures have been successful in changing the dynamics of the epidemic and "flattening the curve" so that fewer people needed to seek treatment at any given time and as such not overwhelm the healthcare system.

One of the most fundamental metrics that describes the epidemic's dynamics is the reproduction number $\mathcal{R}_t$ which is the expected number of secondary infections by a single infectious individual (Delamater et al., 2019). The idea that the course of an epidemic is determined by the rate of contact between susceptible and infectious individuals was proposed by William Hamer in 1906 (Hamer, 1906). Later, Kermack and McKendrick (Kermack & McKendrick, 1927) showed that epidemics stop not when there are no susceptible individuals

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

left, but rather when each infected individual can infect on average fewer than one more person. The reproduction number $\mathcal{R}_t$ depends on three factors: 1) the likelihood of infection per contact, 2) the period during which infectious individuals freely interact with those susceptible to contract the disease, and 3) rate of contact. The likelihood of infection per contact (factor 1) is determined by pathogen virulence and also by protective measures such as social distancing or wearing masks. Free interactions between infectious and susceptible individuals (factor 2) occur until the infectious individual is self-quarantined or hospitalized, either when the individual tests positive or symptoms become severe. Finally, the rate of contact (factor 3) is strongly affected by public health measures to mitigate risk (*WHO*, 2019), such as lockdowns during the COVID-19 epidemic. Thus, the reproduction number is determined by the biological properties of the pathogen and multiple aspects of social behavior. When $\mathcal{R}_t$>1 the number of cases is expected to grow exponentially. The epidemic is contained when $\mathcal{R}_t$ decreases and remains below 1. Real-time estimation of $\mathcal{R}_t$ is critical for determining the effect of implemented mitigation measures and planning for the future.

We propose a method for the continuous estimation of infection rate and reproduction number $\mathcal{R}_t$ that reflects the effects of mitigation measures as well as immunity acquired by those who recover from the disease. We estimate $\mathcal{R}_t$ using a Susceptible-Infectious-Removed (SIR) model (Kermack & McKendrick, 1927) that describes the dynamics of population compartments as follows: individuals start as Susceptible, are infected with the virus and become Infectious, and then move to the Removed compartment once they are quarantined or hospitalized, recover, or die. The SIR model is one of the simplest epidemiological models that still captures the main properties of an epidemic (Anderson, 1991; Anderson & May, 1991) and it has been widely used in epidemic modeling studies. In the majority of SIR modeling studies, the model parameters were constant. An SIR model with constant parameters, however, cannot be applied to the COVID-19 epidemic because various mitigating measures were introduced as the epidemic progressed. The effect of policy changes on COVID-19 dynamics has been modeled using a combination of an SIR model and Bayesian inference (Dehning et al., 2020; Karnakov et al., 2020). In these modeling studies, the infection spreading rate was assumed to be piece-wise linear between the three dates of policy changes. In another approach, continuous estimation of the reproduction number and the effect of mitigation measures were obtained based on estimates of the distribution of the serial interval between the symptom onset in the primary and secondary cases (Cori et al., 2013; Wallinga & Lipsitch, 2007; Wallinga & Teunis, 2004). The Bayesian inference methods as well as methods based on estimations of the serial interval include multiple parameters whose values are not estimated from the data. In contrast, we propose an adaptive SIR model (*aSIR*) in which only one parameter, the removal rate, is taken from the literature, while the second parameter, the infection rate, is continuously estimated from the data using a sliding window. A continuous estimate of the reproduction number $\mathcal{R}_t$ is then calculated using the infection rate estimate. The SIR model is described as a system of differential equations, and the key idea in our proposed method is that the initial conditions for each window are taken as values estimated for the previous window. The only additional hyperparameter is the length of

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

the sliding window. The proposed method retains the conceptual and computational simplicity of SIR-type models and can be easily extended through the introduction of additional compartments supported by data.

Data

The data on daily and cumulative confirmed cases between February 29, 2020 and September 2, 2020 were obtained from John Hopkins University (JHU) and the dates of interventions by states (e.g. state of emergency, stay-at-home order) were obtained from Wikipedia (Wikipedia, 2020). The JHU data were available at two levels of aggregation: county and state. JHU considers many sources for reporting these data; county level information is extracted from the website of the state's departments of health (Johns Hopkins University, 2020) and state level data are extracted directly from the Centers for Disease Control website.

Model

The SIR model is a system of ordinary differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{I}{N} S \\ \frac{dI}{dt} &= \beta \frac{I}{N} S - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \qquad (1)$$

Here, S is the number of susceptible individuals; I is the number of infectious individuals who freely interact with others and can transmit the infection; R is the number of individuals removed from the other two compartments because they are quarantined or hospitalized, recover and acquire immunity, or die. Several COVID-19 government data sources provide the daily number of newly confirmed cases as well as a cumulative number of confirmed cases. Careful consideration is required to determine if these numbers should be attributed to the I or R compartment. In the US, once an individual has been confirmed COVID-19 positive that person is expected to be either self-isolated or hospitalized. Therefore, we assigned the data on confirmed cases to the R compartment, and we fit the model on the cumulative number of confirmed cases.

The infection rate β is $\beta = p * c$, where p is the probability of infection during contact with an infectious individual, and c is the average number of contacts per day. We have no data that would allow us to estimate p and c separately, so we directly estimate β as is usually done when using SIR models.

The removal rate γ determines the rate with which the infected are removed from I to the R compartment. In the context of the COVID-19 epidemic, γ is determined by the time it takes for severe symptoms to appear so the person gets tested and is self-quarantined or hospitalized.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

Therefore, we will assume the duration of the infectious period as the average time it takes for the infected person to become isolated, not the overall time to recover. We assume that the person is infectious from the day they get infected before the symptoms appear (Arons et al., 2020; Z. Hu et al., 2020; Tindale et al., 2020). The average time to develop symptoms has been reported as 5 to 6 days (Li et al., 2020; Linton et al., 2020; *WHO*, 2020). We assume the infectious period before developing severe symptoms is 6 days, so $\gamma = 1/6$.

Time-variant Parameter Estimation

The *aSIR* model contains two parameters, β and γ, with $\gamma = 1/6$ taken from the literature and β estimated using the reported data for each region of interest. The time-variant $\beta(t)$ was estimated using a sliding window of τ =7 days and step of s =1 day, with the estimated values for S and I from the previous window used as the initial conditions for the next window. The reproduction number was calculated as $\mathcal{R}_t(t) = \beta(t)/\gamma$.

1. For the first window, we determined the date when the number of confirmed cases began to increase exponentially. This is important because for many states or counties, very few confirmed cases were initially reported for a number of days or even weeks, which suggests that either the epidemic had not started or the true number of infected people was not known. It is not reasonable to apply a SIR model for this initial period. We took the onset of the epidemic as the first of the four consecutive days in which the number of reported confirmed cases rose in at least three days. The initial conditions for system (1) for window 0 were $S_0(0) = N$, where N is the population in the region of interest, $I_0(0) = 1$, and $R_0(0) = 0$. Infection rate β_i and $S(t)$, $I(t)$ for $t \in [0, \tau-1]$ were estimated given the initial conditions and actual R.
2. Slide the window by s=1 point. For the new $i+1$ window, take the initial conditions as the estimated values from the previous window $S_{i+1}(0) = S_i(s)$, $I_{i+1}(0)= I_i(s)$, and actual $R_{i+1}(0)=R(s)$. Use actual values of $R(t)$, estimate infection rate β_{i+1} and $S_{i+1}(t)$, $I_{i+1}(t)$.
3. For each window, calculate $\mathcal{R}_{t.i} = \beta_i/\gamma$, assign the $\mathcal{R}_{t.i}$ to the last time point of the window. To get a smooth estimate of $\mathcal{R}_t$ we used a rolling average of 5 points.

Results

We fit the model for each state and county in the United States. The model performance was evaluated by calculating the quality of fit as the root mean squared error between the actual and fitted P data for all windows concatenated (*wRMSE*). The fit was excellent with *wRMSE* < 6 across all states. We also calculated a 1-day forecast, 3-day forecast, and 7-day forecast of R after each window (Fig. 1A). The mean absolute prediction error (MAPE) for the forecasts is given in Table 1. The 1-day forecast error did not exceed 2.6% across all states while the 7-day forecast error was large and reached 16.2% for the state of New York. In particular, the 7-day

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

forecast strongly overestimated the number of cases when the reproduction number was high and changing fast (Fig. 1).

The estimated time course of $\mathcal{R}_t$ for the state of New York and for Nassau county, one of the most affected in the beginning of COVID-19 epidemic, are shown in Figure 1. The estimated daily number of infectious individuals rapidly increased and then gradually declined after the lockdown was introduced on March 22, 2020 (Fig. 1A). The estimated reproduction number also declined after the lockdown began (Fig. 1B). The time course of $\mathcal{R}_t$ shows weekly seasonality which likely reflects the effect of social interactions and possibly the effect of fluctuations in case reporting on weekdays vs. weekends. For New York state and Nassau county, $\mathcal{R}_t$ exhibited an initial increase which may reflect the fact that the epidemic in the New York region was continuously seeded by travelers arriving to JFK airport until a ban on international travel was introduced on March 12, 2020. It may also reflect the fact that not all severe cases were initially recognized and reported as COVID-19. In Florida, $\mathcal{R}_t$ decreased close to 1 by mid-April but then began increasing at the end of May (Fig. 1B). In June 2020, Florida authorities introduced more stringent measures to control the epidemic which is reflected in decreasing $\mathcal{R}_t$ in the second half of July 2020. The opening of multiple states since June 2020 has been accompanied by $\mathcal{R}_t$ rising above 1 (not shown here), and close monitoring of reproductive number is needed to contain another wave of the epidemic.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

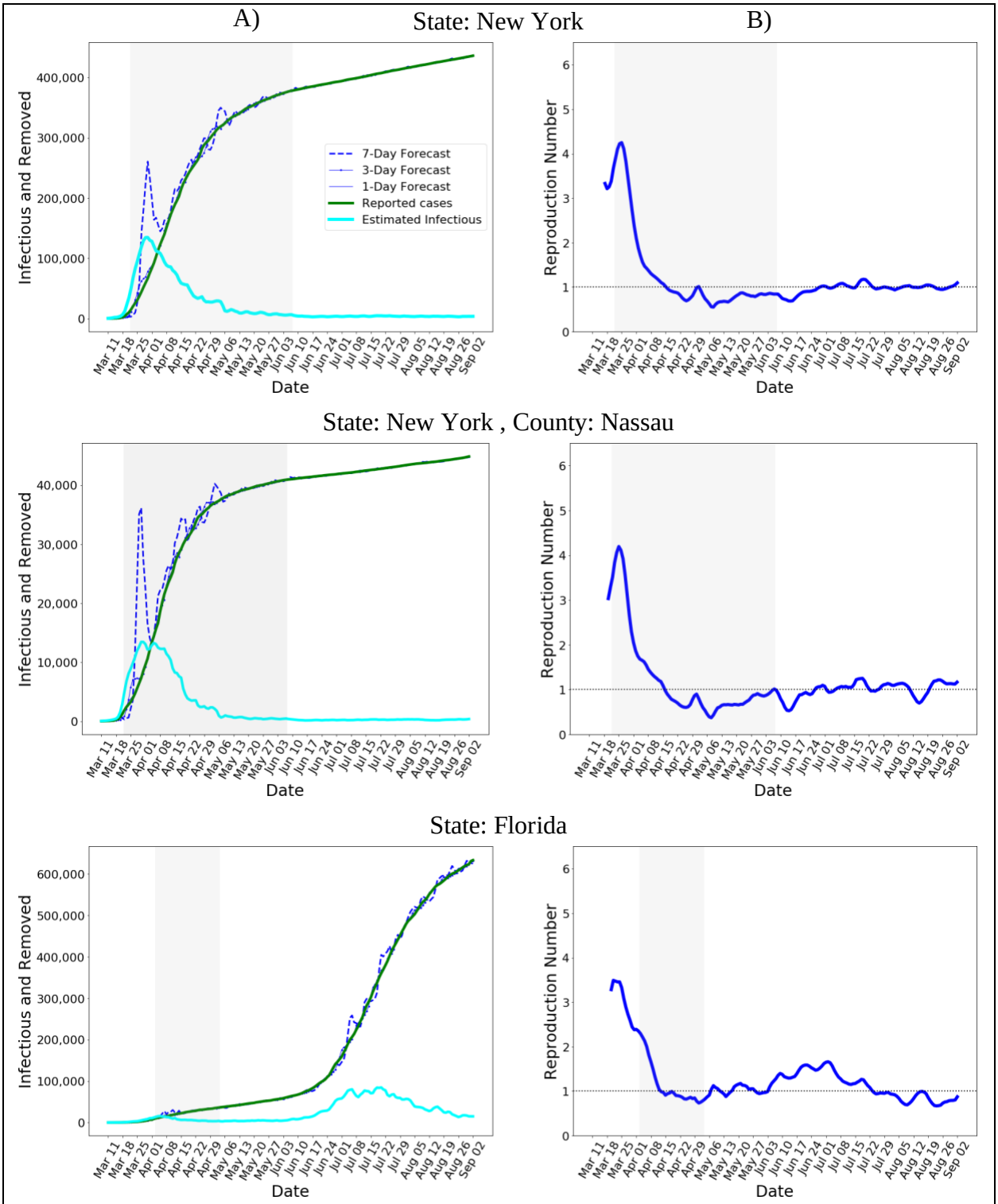

Figure 1. **A)** Estimated Infectious and forecasted Removed. **B)** Estimated reproduction number $\mathcal{R}_t$. The shaded region indicates the dates of the lockdown. While the 1-day and 3-day forecasts are accurate, the 7-day forecast exhibits large errors when $\mathcal{R}_t$>1 and is changing fast.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

We also compared *aSIR* with the model by Cori and colleagues (Cori et al., 2013) implemented as R package *EpiEstim* and a model implemented by Kevin Systrom and Thomas Vladeck available from a popular and influential COVID-19 data tracking website *rt.live* (Fig. 2). In *EpiEstim*, we assumed equal probability of infection within the infectious period of 6 days, the $\mathcal{R}_t$ estimate was smoothed with a 7-point rolling average window, same as in *aSIR*. While all three models show similar estimates when $\mathcal{R}_t$ is close to 1, their estimates differ considerably in in the beginning of the epidemic. In particular, the *rt.live* model returned lower max $\mathcal{R}_t$ than the other two models, and estimated that $\mathcal{R}_t$ already decreased to 1 by the time the lockdown was announced in NY state on March 22, 2020 (Fig. 2, shaded region). The *EpiEstim* and the *aSIR* models estimated similar peak values of $\mathcal{R}_t$, and both models estimated that $\mathcal{R}_t$ dereased close to 1 in the first week of April, 2020. Although both models show a rapidly decreasing $\mathcal{R}_t$ in March, the *aSIR* model shows a lagged change. We are not aware of ground truth data however to determine which model produces a more accurate estimate.

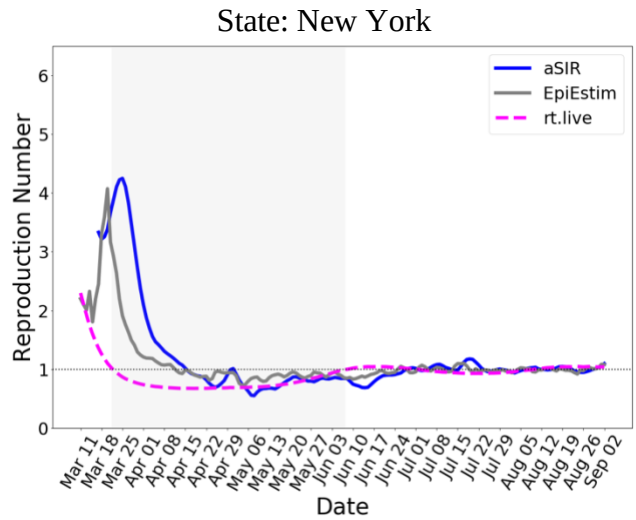

Figure 2. Comparison of models that generate continuous $\mathcal{R}_t$ estimates. The three $\mathcal{R}_t$ estimates differ widely in the beginning of the epidemic. In particular, the $\mathcal{R}_t$ estimated by *rt.live* model decreased to 1 by the lockdown onset on March 22 (shaded region).

Next, we investigated the effect of an abrupt increase in testing on the $\mathcal{R}_t$ estimate (Fig. 3). We assumed a step-wise 50% increase in testing that persisted after April 12 (Fig. 3, left panel). Both *aSIR* and *EpiEstim* models exhibited a spike in $\mathcal{R}_t$. It can also be argued that if testing increases then infectious individuals may be detected and quarantined sooner, resulting in a shorter infectious period and larger removal rate γ, in turn lowering $\mathcal{R}_t$. We did not model a possible increase in γ. Instead, we assumed that the underlying dynamics of the epidemic did not

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

change, and within 2 weeks both models returned to the $\mathcal{R}_t$ time course estimated without the testing increase.

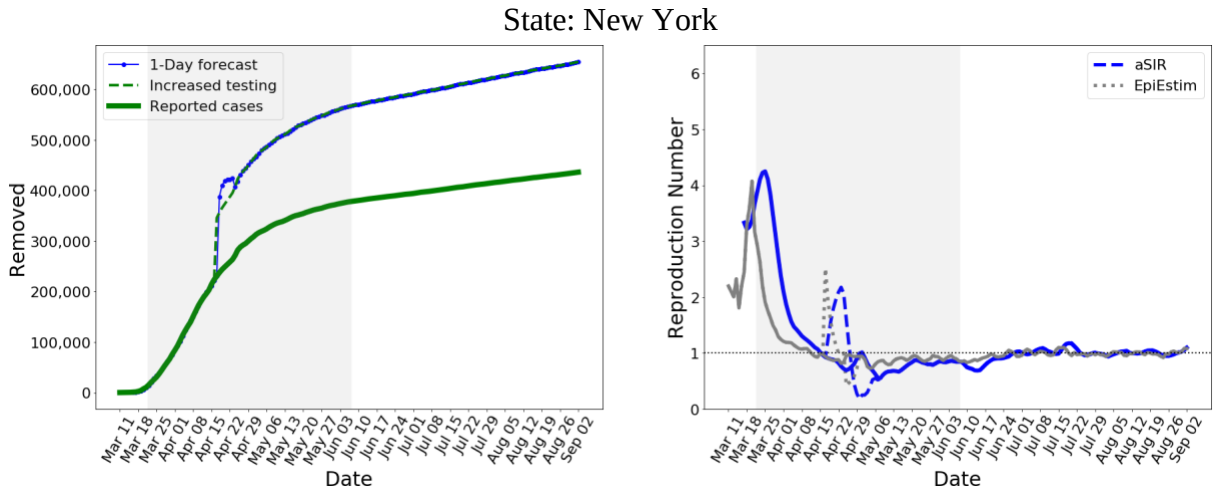

Figure 3. The effect of a step-wise 50% increase in testing (left panel, dashed line). The 1-day forecast by the *aSIR* model adapts within a week. For $\mathcal{R}_t$ estimate, both *EpiEstim* and our *aSIR* models produced a spike followed by a decrease (right panel, dashed lines) before returning to the unperturbed $\mathcal{R}_t$ time course (solid lines).

Discussion

We have developed a simple approach to adaptively estimate time-varying parameters of the SIR model using reported data on the number of confirmed cases. This approach adds to the already large literature on COVID-19 modeling in two ways. First, we estimate the parameters of the SIR model using a sliding window of a limited duration, 7 days, to account for fast changes in transmissibility and contact patterns in response to changes in social behavior and government mitigation measures. The window duration is a hyperparameter that can be changed as needed, the trade-off being the accuracy of the parameter estimates versus rapid reaction to changes in the underlying epidemic. Because the proposed model is so simple, a number of scenarios can be explored as needed.

Second, we attribute the data on reported cases to the Removed compartment rather than Infectious. This modeling decision is based on the realities of the COVID-19 epidemic in the US where confirmed positive individuals are supposed to self-isolate or are hospitalized. Although these individuals remain infectious and can infect other family members or caretakers even when self-isolated or hospitalized, they are not freely interacting with the susceptible population as would be required to attribute them to the *I* compartment. It has been proposed to add a new X compartment in the SIR model to model symptomatic quarantined infectious individuals (Maier & Brockmann, 2020). We have no data to independently estimate this additional parameter of quarantine rate, however. For the same reason, we did not use the Susceptible-Exposed-Infected-Removed (SEIR) model because we are not aware of reliable data about the duration of the

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

exposed period during which an infected person is not yet infectious. Moreover, it has been found that the SIR model performed better than an SEIR model in representing the information contained in the confirmed-case data on COVID-19 (Roda et al., 2020).

The reported numbers of positive cases represent a fraction of infected individuals because of the limited availability of testing in March and April of 2020 with the result that only those who developed severe symptoms were tested. Up to 80% of the infected individuals may be asymptomatic or develop mild symptoms (Wu & McGoogan, 2020) and were not tested, so for that period our model applies only to the small sub-population who developed severe symptoms. This sub-population, however, is of particular interest because it represents those most at risk, and the reproduction number estimated from this limited data can be used to guide policy decisions aimed at protecting the most vulnerable population (Lourenco et al., 2020). At the same time, as the numbers of the tested individuals increase, the short sliding window approach makes our model adaptable to an ever-larger proportion of the population (Fig. 3).

Across all US states, the maximal $\mathcal{R}_t$ values of were estimated for New York (4.4) and Michigan (4.5) (Table 1) which is close to the mean value of 4.34 estimated for Italy (Wangping et al., 2020) but higher than that obtained by a stochastic transmission model (Abbott et al., 2020; Kucharski et al., 2020). The wide range of maximal values of $\mathcal{R}_t$ from 2.0 to 4.5 (Table 1) likely reflects the differences in contact rates due to population density (H. Hu et al., 2013; Sy et al., 2020). Increased social distancing is required to contain the spread of the epidemic (Prem et al., 2020), with more stringent mitigation measures, including lockdown, considered necessary to decrease the contact rate in high-density states and counties. Another measure to lower $\mathcal{R}_t$ is to increase the removal rate γ by intensive testing and quarantine of individuals tested positive. This targeted intervention would strongly decrease the interaction between the infectious and susceptible individuals and keep $\mathcal{R}_t$ <1 until a vaccine becomes available. Our model will allow researchers as well as policy makers to monitor the reproduction number in different geographical regions of the US, better understand the effect of government policies on the dynamics of the epidemic, and develop further mitigation strategies as we continue to battle COVID-19 (Friedman et al., 2020; Inglesby, 2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

Table 1. Reproduction numbers $\mathcal{R}_t$ and forecast accuracy for 50 states.

State	$\mathcal{R}_t$ **Max**	**MAPE 1-Day Forecast**	**MAPE 3-Day Forecast**	**MAPE 7-Day Forecast**
Alabama	2.9	1.5%	4.2%	10.0%
Alaska	2.8	1.6%	3.7%	11.3%
Arizona	3.3	1.3%	2.9%	10.1%
Arkansas	2.8	1.5%	4.0%	12.6%
California	2.5	1.7%	2.9%	6.3%
Colorado	2.6	1.1%	2.9%	7.3%
Connecticut	4.1	2.0%	3.1%	9.3%
Delaware	2.4	1.7%	2.9%	7.9%
District of Columbia	2.1	0.8%	1.8%	4.4%
Florida	3.6	2.0%	4.4%	9.3%
Georgia	3.0	1.8%	3.7%	7.4%
Hawaii	2.7	2.0%	3.6%	9.7%
Idaho	3.4	2.4%	4.8%	13.6%
Illinois	4.0	1.3%	2.5%	8.5%
Indiana	3.8	1.4%	4.0%	10.4%
Iowa	2.8	1.8%	3.6%	8.0%
Kansas	3.0	1.6%	3.5%	8.6%
Kentucky	3.0	2.6%	4.9%	11.2%
Louisiana	3.7	1.8%	4.0%	12.1%
Maine	2.0	1.2%	2.8%	6.7%
Maryland	3.3	1.2%	2.8%	6.2%
Massachusetts	3.4	1.3%	3.6%	9.7%
Michigan	4.5	1.6%	3.5%	12.8%
Minnesota	2.7	1.3%	2.9%	8.0%
Mississippi	2.9	1.2%	3.0%	9.3%
Missouri	3.6	1.8%	3.3%	11.4%
Montana	3.3	1.6%	3.7%	11.9%
Nebraska	2.5	2.0%	4.1%	9.7%
Nevada	2.9	2.3%	3.8%	10.0%
New Hampshire	2.3	1.7%	3.3%	8.5%
New Jersey	4.1	1.5%	2.3%	7.8%
New Mexico	2.3	2.2%	3.4%	7.3%
New York	4.4	1.5%	4.2%	16.2%
North Carolina	3.2	1.3%	2.4%	7.2%

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

North Dakota	2.4	1.8%	4.8%	12.6%
Ohio	3.3	1.2%	3.4%	9.8%
Oklahoma	3.1	1.4%	3.5%	10.2%
Oregon	2.5	1.3%	2.8%	6.6%
Pennsylvania	3.2	1.7%	2.9%	6.3%
Rhode Island	2.4	1.5%	3.1%	6.8%
South Carolina	3.5	2.1%	4.3%	10.6%
South Dakota	2.1	1.3%	3.2%	8.7%
Tennessee	3.5	2.2%	4.8%	12.5%
Texas	3.6	2.0%	3.9%	9.3%
Utah	3.2	1.4%	3.1%	8.3%
Vermont	2.9	0.8%	2.4%	7.7%
Virginia	2.5	1.1%	2.1%	5.1%
Washington	3.0	2.0%	4.8%	8.6%
West Virginia	3.5	1.6%	3.9%	14.0%
Wisconsin	3.6	1.5%	3.2%	10.0%
Wyoming	2.9	1.9%	4.7%	14.2%

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

References

Abbott, S., Hellewell, J., Thompson, R. N., Sherratt, K., Gibbs, H. P., Bosse, N. I., Munday, J. D., Meakin, S., Doughty, E. L., Chun, J. Y., Chan, Y.-W. D., Finger, F., Campbell, P., Endo, A., Pearson, C. A. B., Gimma, A., Russell, T., CMMID COVID modelling group, Flasche, S., … Funk, S. (2020). Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Research*, *5*, 112. https://doi.org/10.12688/wellcomeopenres.16006.1

Adam, D. (2020). Special report: The simulations driving the world's response to COVID-19. *Nature*, *580*(7803), 316–318. https://doi.org/10.1038/d41586-020-01003-6

Anderson, R. M. (1991). Discussion: The Kermack-McKendrick epidemic threshold theorem. *Bulletin of Mathematical Biology*, *53*(1), 1. https://doi.org/10.1007/BF02464422

Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, *395*(10228), 931–934. https://doi.org/10.1016/S0140-6736(20)30567-5

Anderson, R. M., & May, R. M. (1991). *Infectious diseases of humans: Dynamics and control*. Oxford University Press.

Arenas, A., Cota, W., Gomez-Gardenes, J., Gómez, S., Granell, C., Matamalas, J. T., Soriano-Panos, D., & Steinegger, B. (2020). A mathematical model for the spatiotemporal epidemic spreading of COVID19. *MedRxiv*, 2020.03.21.20040022. https://doi.org/10.1101/2020.03.21.20040022

Arons, M. M., Hatfield, K. M., Reddy, S. C., Kimball, A., James, A., Jacobs, J. R., Taylor, J., Spicer, K., Bardossy, A. C., Oakley, L. P., Tanwar, S., Dyal, J. W., Harney, J., Chisty, Z., Bell, J. M., Methner, M., Paul, P., Carlson, C. M., McLaughlin, H. P., … Jernigan, J. A. (2020). Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *New England Journal of Medicine*, *382*(22), 2081–2090. https://doi.org/10.1056/NEJMoa2008457

Cori, A., Ferguson, N. M., Fraser, C., & Cauchemez, S. (2013). A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology*, *178*(9), 1505–1512. https://doi.org/10.1093/aje/kwt133

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*. https://doi.org/10.1126/science.abb9789

Delamater, P. L., Street, E. J., Leslie, T. F., Yang, Y. T., & Jacobsen, K. H. (2019). Complexity of the Basic Reproduction Number (R0). *Emerging Infectious Diseases*, *25*(1), 1–4. https://doi.org/10.3201/eid2501.171901

Enserink, M., & Kupferschmidt, K. (2020). With COVID-19, modeling takes on life and death importance. *Science*, *367*(6485), 1414–1415. https://doi.org/10.1126/science.367.6485.1414-b

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba Perez, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L., Van Elsland, S., … Ghani, A. (2020). *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand*. Imperial College London. https://doi.org/10.25561/77482

Friedman, E., Friedman, J., Johnson, S., & Landsberg, A. (2020). Transitioning Out of the Coronavirus Lockdown: A Framework for Evaluating Zone-Based Social Distancing. *Frontiers in Public Health*, *8*. https://doi.org/10.3389/fpubh.2020.00266

Hamer, W. (1906). The Milroy Lectures ON EPIDEMIC DISEASE IN ENGLAND—THE EVIDENCE OF VARIABILITY AND OF PERSISTENCY OF TYPE. *The Lancet*, *167*(4305), 569–574. https://doi.org/10.1016/S0140-6736(01)80187-2

Hu, H., Nigmatulina, K., & Eckhoff, P. (2013). The scaling of contact rates with population density for the infectious disease models. *Mathematical Biosciences*, *244*(2), 125–134. https://doi.org/10.1016/j.mbs.2013.04.013

Hu, Z., Song, C., Xu, C., Jin, G., Chen, Y., Xu, X., Ma, H., Chen, W., Lin, Y., Zheng, Y., Wang, J., Hu, Z., Yi, Y., & Shen, H. (2020). Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Science China. Life Sciences*, 1–6. https://doi.org/10.1007/s11427-020-1661-4

Inglesby, T. V. (2020). Public Health Measures and the Reproduction Number of SARS-CoV-2. *JAMA*, *323*(21), 2186–2187. https://doi.org/10.1001/jama.2020.7878

Johns Hopkins University. (2020, July 19). *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University*. Retrieved from Coronavirus Resource Center: https://github.com/CSSEGISandData/COVID-19

Karnakov, P., Arampatzis, G., Kičić, I., Wermelinger, F., Wälchli, D., Papadimitriou, C., & Koumoutsakos, P. (2020). Data driven inference of the reproduction number (R0) for COVID-19 before and after interventions for 51 European countries. *MedRxiv*, 2020.05.21.20109314. https://doi.org/10.1101/2020.05.21.20109314

Kermack, W. O., & McKendrick, A. G. (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. JSTOR.

Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R. M., Sun, F., Jit, M., Munday, J. D., Davies, N., Gimma, A., van Zandvoort, K., Gibbs, H., Hellewell, J., Jarvis, C. I., Clifford, S., Quilty, B. J., Bosse, N. I., … Flasche, S. (2020). Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *The Lancet Infectious Diseases*, *20*(5), 553–558. https://doi.org/10.1016/S1473-3099(20)30144-4

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S. M., Lau, E. H. Y., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., … Feng, Z. (2020). Early Transmission Dynamics in Wuhan, China, of Novel

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

Coronavirus–Infected Pneumonia. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa2001316

Linton, N. M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A. R., Jung, S., Yuan, B., Kinoshita, R., & Nishiura, H. (2020). Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. *Journal of Clinical Medicine*, *9*(2). https://doi.org/10.3390/jcm9020538

Lourenco, J., Paton, R., Ghafari, M., Kraemer, M., Thompson, C., Simmonds, P., Klenerman, P., & Gupta, S. (2020). Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. *MedRxiv*, 2020.03.24.20042291. https://doi.org/10.1101/2020.03.24.20042291

Maier, B. F., & Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*, *368*(6492), 742–746. https://doi.org/10.1126/science.abb4557

Mitjà, O., Arenas, À., Rodó, X., Tobias, A., Brew, J., & Benlloch, J. M. (2020). Experts' request to the Spanish Government: Move Spain towards complete lockdown. *The Lancet*, *395*(10231), 1193–1194. https://doi.org/10.1016/S0140-6736(20)30753-4

Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza. (2019). World Health Organization; World Health Organization. http://www.who.int/influenza/publications/public_health_measures/publication/en/

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., Flasche, S., Clifford, S., Pearson, C. A. B., Munday, J. D., Abbott, S., Gibbs, H., Rosello, A., Quilty, B. J., Jombart, T., Sun, F., Diamond, C., Gimma, A., Zandvoort, K. van, … Klepac, P. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *The Lancet Public Health*, *5*(5), e261–e270. https://doi.org/10.1016/S2468-2667(20)30073-6

Q&A on coronaviruses (COVID-19). (2020, June 28). World Health Organization. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses

Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling*, *5*, 271–281. https://doi.org/10.1016/j.idm.2020.03.001

Sy, K. T. L., White, L. F., & Nichols, B. E. (2020). Population density and basic reproductive number of COVID-19 across United States counties. *MedRxiv*, 2020.06.12.20130021. https://doi.org/10.1101/2020.06.12.20130021

Tindale, L. C., Stockdale, J. E., Coombe, M., Garlock, E. S., Lau, W. Y. V., Saraswat, M., Zhang, L., Chen, D., Wallinga, J., & Colijn, C. (2020). Evidence for transmission of COVID-19 prior to symptom onset. *ELife*, *9*. https://doi.org/10.7554/eLife.57149

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20193896; this version posted September 15, 2020. 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-ND 4.0 International license .

Wallinga, J., & Lipsitch, M. (2007). How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings*. *Biological Sciences*, *274*(1609), 599–604. https://doi.org/10.1098/rspb.2006.3754

Wallinga, J., & Teunis, P. (2004). Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology*, *160*(6), 509–516. https://doi.org/10.1093/aje/kwh255

Wangping, J., Ke, H., Yang, S., Wenzhe, C., Shengshu, W., Shanshan, Y., Jianwei, W., Fuyin, K., Penggang, T., Jing, L., Miao, L., & Yao, H. (2020). Extended SIR Prediction of the Epidemics Trend of COVID-19 in Italy and Compared With Hunan, China. *Frontiers in Medicine*, *7*. https://doi.org/10.3389/fmed.2020.00169

Wikipedia. (2020). *U.S. state and local government responses to the COVID-19 pandemic*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/U.S._state_and_local_government_responses_to_the_COVID-19_pandemic

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*, *323*(13), 1239–1242. https://doi.org/10.1001/jama.2020.2648