medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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-NC 4.0 International license .

Mitigating COVID-19 outbreak via high testing capacity and strong transmission-intervention in the United States

Shi Chen,[1] Qin Li,[1*] Song Gao,[2*] Yuhao Kang[2], Xun Shi[3]

[1]Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA
[2]GeoDS Lab, Department of Geography, University of Wisconsin-Madison, Madison, WI 53706, USA
[3]Department of Geography, Dartmouth College, Hanover, NH 03755, USA

*To whom correspondence should be addressed; E-mail: qinli@math.wisc.edu; song.gao@wisc.edu.

Most models of the COVID-19 pandemic in the United States do not consider geographic variation, and their relevance to public policies is not straightforward. We developed a mathematical model that characterizes infections by state and incorporates inflows and outflows of interstate travelers. Modeling reveals that curbing interstate travel when the disease is already widespread will make little difference. Meanwhile, increased testing capacity (facilitating early identification of infected people and quick isolation) and strict social-distancing and self-quarantine rules are effective in abating the outbreak. The modeling has also produced state-specific information. For example, for New York and Michigan, isolation of persons exposed to the virus needs to be imposed within 2 days to prevent a broad outbreak, whereas for other states this period can be 3.6 days. This model could be used to determine resources needed before safely lifting state policies on social distancing.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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-NC 4.0 International license .

Introduction

The Coronavirus disease (COVID-19) is an ongoing pandemic that poses a global threat. As of March 26, 2020, more than 520,000 cases of COVID-19 have been reported in over 200 countries and territories, resulting in approximately 23,500 deaths (*9–12, 24, 26, 28, 35, 40*). In the United States, the first known positive case was identified in Washington state on January 20, 2020 (*18*). By March 26, the epidemic had been rapidly spreading across many communities and present in all 50 states, plus the District of Columbia; the number of confirmed cases in the United States rose to 78,786 with 1,137 deaths.

To combat the spread of COVID-19, the government has taken actions in various dimensions, including banning or discouraging domestic and international travels, announcing stay-at-home orders to curb non-essential interactions for reducing transmission rate, and urging commercial laboratories to increase test capacity. To curb traveling, on January 31, the United States government announced travel restrictions on travelers from China; on February 29, it announced travel ban against Iran and advised travel with caution to Europe (*1*) ; on March 11, it announced travel restrictions on most of European countries. To reduce human-interactions, on March 13, a national emergency was declared; as of March 28, 39 states had issued either statewide or regionally stay-at-home or shelter-in-place order, requiring residents to stay indoors except for essential activities. To increase test capacities, on February 4, the United States Food and Drug Administration (FDA) approved the United States Centers for Disease Control and Prevention (CDC)'s test, which was later to be proved inconclusive (*2*); on February 29, the FDA relaxed its rules for some laboratories, allowing them to start testing before the agency granting its approvals; on March 27, FDA issued an Emergency Use Authorization to a medical device maker, the Abbott Labs, for the use of a coronavirus test that delivers quick testing results (*3*).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

So far, since there is no treatment or vaccine for SARS-COV-2 available, these actions have been taken largely based on classic epidemic controls. Works on evaluating similar measures in other countries, especially China, started to emerge (*14, 24*). For example, the effect of travel restriction on delaying the virus spread in China has been reported (*9, 36*). However, it is still unclear what control and intervention measures would have actual effect, and to what extent, on abating the spread of COVID-19 in the United States. As the United States has very different political, administrative, social, pubic health and medical systems, as well as culture from China, this remains to be a critical question to address, especially considering that some measures and policies come with extremely high economic and societal costs.

There have been numerous modeling works projecting or predicting the trend of the COVID-19 pandemic regionally or globally (*16, 21*). Most of the works apply a global model to the entire study area, either a region, a country, or the entire globe. Rarely the variation of different parts within one area and the interactions among those parts are taken into consideration. However, a country like the United States features diversity in all aspects. On the one hand, the overall situation of the entire country is a result emerging from local situations and their interactions, and thus, ignoring the local interactions can hardly lead to a high-quality overall model; on the other hand, as all interventions and policies finally have to be adapted to the local situation, a localized modeling will be much more relevant to the real-world practices. Spatially and network-related epidemic models can describe the geographical spread of viral dynamics (*24, 27, 30, 37*). Recent studies have shown the importance of incorporating timely human mobility patterns derived from mobile phone big data and global flight networks into the epidemiology modeling process and in public health studies (*7–9, 15, 20, 22, 24, 29, 33, 41, 42*). Without accurate models that incorporate human mobility patterns and spatial interactions (*8, 42*), it is rather challenging to quantify the sensitivity of parameters, and using the linkage to real practices to make sensible policy suggestions.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

Accordingly, the core of the study is twofold. First, to localize the modeling, we developed a mathematical compartmental model system that simultaneously characterizes the spatiotemporal dynamics of infections in 51 areas (50 states and the District of Columbia). Each state or district has its own model, and all models simultaneously take into account inflows and outflows of interstate travelers. Second, to improve the practical relevance, we chose to use three parameters that can directly correspond to possible practical means to discover, combat, and control the spread of the disease, and quantify their impact on the final output of the model. The three parameters include: 1) the transmission rate b, which corresponds to the local social-distancing enforcement, e.g., the stay-home order; 2) the detection and reporting rate r, which corresponds to the testing capacity; and 3) the travel ratio α_t, which corresponds to the ratio of interstate travel volume compared to that of 2019 during the same period.

The modeling is a dynamic projection process. We employed daily and state-specific historical data to incrementally calibrate the model, and then used the calibrated model to predict future scenarios under different control and intervention measures. During this process, we ran data assimilation methods to identify parameter values that optimally fit the current situation (see more details in the supplementary material). To project into the future, we set different values for the parameters to create different control and intervention scenarios, and then ran the simulation to see their impact on the model results. The final output of the model is the total number of confirmed cases in a state on a particular day.

The current strategy in the United States is to isolate people who have the symptoms of COVID-19. An ideal scenario is to have an 100% reporting rate, i.e., every infected case gets confirmed and thus isolated quickly. Another ideal setting is to have everyone who was in contact with the infected gets identified and isolated quickly as well. Our model incorporated these considerations and examined such direct isolation of the exposed compartment in detail. We particularly investigated the impact of quickness of such actions through mathematical

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

modeling and scenario analysis.

A notable result from our modeling is that the impact of interstate travel restriction on the model output is modest. This can be explained by that when the disease has already widespread in all states, the relatively small number of cases in the travelers will cause little difference to the local situation, compared with the effects of local social-distancing and isolation rules and the increase of testing capacity.

Methods

The mathematical model that simulates the spatiotemporal dynamics of state-level infections in the United States is a modified compartmental model in epidemiology by taking into account the variation of the 51 administrative units and their interactions (*6, 17, 19*). It consists of 51 ordinary differential equation (ODE) systems, with each one characterizing the evolution of susceptible (S), exposed (E), reported (I), unreported (A) and resolved (R) cases per state (see more details in the supplementary material and also Fig. S1). The 51 ODE systems are then coupled through the state-to-state travel network flows (see Fig. S2) that were extracted from the aggregated SafeGraph mobility data and weighted by α_t (*25, 31*). Unlike most other models, we also incorporate the potential asymptomatic transmission. This makes the derivation of the basic reproduction number R_0 different. Besides, each ODE system also includes two unknown parameters: the transmission rate (b) and the report rate for each state (r). The unknown parameters are inferred based on the number of confirmed cases in each state for the period of March 1-March 20, 2020. The source of infection case data is (*10*).

The COVID-19 transmission dynamics (the ODE system) was simulated using the Forward Euler method, with each day discretized into 24 smaller time periods to ensure the numerical stability (see Fig. S3). The parameter fitting was conducted under the Bayesian formulation that combines the effect of the underlying dynamics governed by the ODE system, serving as

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

the prior knowledge, and the collected data, appearing in the likelihood function, to generate the posterior distribution that characterized the behavior of the state variables, including S, E, I, A, R, as well as the two unknown parameters, b and r. For this classical data assimilation problem, we employed the Ensemble Kalman Filter method that was derived from the Kalman filter and tailored to deal with problems with high-dimensional state variables (*13, 32*). The method proves to be effective when the measuring operator is linear and the underlying dynamics is Gaussian-like. It has been applied to a vast of problems that do not strictly satisfy the Gaussianity requirement. To apply this method, we generated 2000 samples according to the prior distribution, and evolve the samples through the dynamics of the ODE system. The samples were then rectified at the end of each day, using the announced number of confirmed cases, for tuning the two unknown parameters b and r.

At the beginning of the simulation, March 1st, only a few states had non-zero confirmed cases. The true numbers of exposed people and unreported cases on that day, however, are unknown. These two numbers are also the state variables that need to be inferred to using the collected data. On March 1, we put a non-informative prior with range $[0, 500]$ and $[0, 200]$ over the exposed latent population and unreported infectious population in each state, respectively. Fig. S4-S13 show the data assimilation results for different states including the number of people in different compartmental groups and their temporal changes with 95% credible intervals. The average reporting rate r over all states is 0.2266 at the end of March 20 through the data assimilation method.

For forecasting, we performed scenario studies of two types. First, we ran the mathematical model by applying the initial data obtained as of March 20 into the future for the next 40 days, but with different configurations of (b, r, α_t). The simulation results out of this setting were then compared with those from the setting that the three parameters remained unchanged for each state. To quantify and visualize the difference, we compared the increase of the percentage of

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

the non-affected population when the measures of stay-at-home, increasing test rate, and travel bans were enacted.

The second scenario was about a more ideal situation: every confirmed case would get isolated immediately, as well as those who had been exposed to those confirmed cases, no matter if those who had been exposed had started to show symptoms or not. We built a new mathematical model that incorporated such isolations to study the effect of them. A new quarantined compartment (Q) was introduced into the model. Through the simulation, we examined the correlation between the average action-taking time (i.e., temporal lag in putting a person into quarantine denoted by D_q) and the increase of non-infected population. In both scenario studies, the simulation was run with the Forward Euler ODE solver, during which each day was divided into 24 intervals to achieve a numerical stability.

Findings

The interstate travel in the United States has been rapidly dropping recently (*39*), which is going to have a tremendous impact on the country's economy and society. However, this study finds that the mitigation effect that this change can bring about by itself, regarding abating the COVID-19 spread, is very limited as the outbreak has already widespread in all states. On the other hand, our modeling indicates that lowering the transmission rate and the unreported-case rate would have a great effect on abating the infection. Mapped to the public policy, this means that curbing the interstates travel alone is not going to reduce the spread of COVID-19, while the policies such as ordering stay-at-home and social distancing that reduces human interaction and increasing the testing capacities of all clinics around the nation could significantly decrease the total number of infections. While we did not conduct a rigorous economic analysis in this study, it seems reasonable to assume that producing more testing kits and conducting more tests will be significantly less expensive than completely shutting off business and curbing travels.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

This potentially means that the economy would not have been hit as hard if we could have increased the detection capacity. These findings are in line with the observation in Germany and Korea (*4, 34*), both of which had a large percentage of the population tested, while opted out lockdowns of cities.

Fig. 1 shows the effect on spatiotemporal dynamics of infectious population across states by setting the coefficients at different configurations. We set $r = 1-\alpha_r(1-r_0)$ and $b = \alpha_b b_0$, where r_0 and b_0 are the report and transmission rate as of March 20, 2020 using data assimilation fitting result. By decreasing α_r from 1 to 0, we increase the report rate from the original r_0 to 1, and by decreasing α_b we decrease the transmission rate. Most states, except a few such as NY, MI, and CA, see drastic improvement when the transmission rate is decreased and the testing(reporting) rate is increased, but the reduction of interstate traffic alone is not as effective. Our modelling reveals that once the epidemic in an area has reached a certain stage, the difference that can be caused to the local situation by the relatively small number of imported cases due to the interstate travel is insignificant. According to our modeling, all states in the United States have reached that stage. Therefore, as long as those travelers follow the social-distancing rules and the local government provides sufficient testing capacity, there is no apparent urge to curb interstate travel. This is in line with the finding in (*20, 36*), in which the authors projected the pick up of the spreading in other parts of China outside of Wuhan with about 3 days delay, and in the world outside China within a 2-3 weeks of delay, assuming no further screening is in place. Different from China where the city of Wuhan is clearly the epicenter of the COVID-19 outbreak and the travel ban quickly gets the rest of China under control, most of the states in the United States have already had signs of community spread by March 20, 2020 (*5*), and banning other states will hardly make much difference to the local situation. In addition, Fig. 2 shows the corresponding prediction time series of infectious population in top 15 states under two scenarios (see also Fig. S14): (A) the reported rate and the transmission rate are unchanged

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

as of March 20, 2020, with $\alpha_r = \alpha_b = 1$, in which most states will continue their exponential growth before reaching their peak; (B) with $\alpha_r = \alpha_b = 0.1$, that is, when the transmission rate b is much smaller and the reported rate r is much higher (closer to 1), we can "flatten the curve" on the virus.

We further investigate the effect of increased testing capacity and report rate. As shown in Figure 3a, most states see drastic improvement when the report rate increases. All states, by April 29, see monotonically exponential reduction of infections. The impact is strong in states such as MA, AZ, FL, and OR, but relatively weak in states such as NY, MI and IL. In Figure 3b, we study the effect of α_r and α_b on the basic reproduction rate R_e in NY (see other states in Fig. S15). It can be seen that merely raising the report rate cannot fully make $R_e < 1$. To mitigate the spread of COVID-19 in these states, a proactive approach needs to be taken, and quick detection and isolation of the exposed population need to be in place instead of being delayed until the onset of the symptoms. This measure can prevent the exposed population from potentially infecting other susceptible people. In Figure 3c, we plot the increase of infections in terms of D_q (i.e., the temporal lag in putting a person into quarantine) for the states that are sensitive to change of D_q, including NY, NJ, IL, GA, MI, CO, WI, LA, TX, PA, MA, and TN. The longer one waits to inform and isolate the exposed population, the more infection one observes. For example, there is a sharp transition for NY and MI. If the average detection and isolation time is more than 2 days, the number of infections will significantly increase.

The results again showed the importance of sufficient testing and strong transmission-intervention measures such as self-quarantine and stay-at-home policy (*38*). These policies can help quickly identify the source of infection and isolate them before they infect the remaining population. This measure presumably comes with a lower economical cost.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

Discussion

Modeling and analyzing the spread of COVID-19, and assessing the effect of various policies could be instrumental to national and international agencies for health response planning (*9, 14, 21, 28, 36, 38*). We show that the effect of interstate travel reduction is at most modest in the United States when the outbreak has already widespread in all states. On the other hand, we need to impose strong transmission-reduction intervention and increased testing capacity and report rate to contain the spread of virus. The result is in agreement with previous findings (*9, 12, 14, 23, 36, 40*), suggesting that the effect of travel ban at a later stage of the outbreak is rather modest. This is also in line with the fact that the outbreaks still occurred in Europe even upon the strong travel ban on the earlier epicenter of Wuhan and its surrounding cities in China. We also show that the transmission-reduction intervention such as policies on the social-distancing and shelter-in-place rules, and the increase of testing rate, which facilitates immediate isolation upon exposure, will significantly reduce the total infected population. Such effect is mostly visible for the states of NY, NJ, MI, and IL. Particularly, our modeling results showed that for states such as NY and MI, to achieve an optimal infection reduction, a more proactive approach needs to be taken to quickly identify the exposed population and isolate them within two days of exposure in order to ensure the infection reduction. The result is in agreement with previous findings (*24, 28*).

We do need to emphasize that the model itself does not distinguish different ways of traveling across states. Indeed, if the interstate travel is conducted mostly through transiting through busy airports and train stations, and the social-distancing policy is not strictly imposed, then the high population density at these places will bring up the transmission rate b locally in space and time, leading to a higher infection rate. This is a severe consequence, but it should not be counted as the direct result of relaxing travel restrictions.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

Moving forward, we estimate that the decline in travel has a modest effect on the mitigation of the pandemic. We need a stronger transmission-reduction intervention and increased detection and report rate in place to prevent the further spread of the virus. The results could potentially be used to design a optimal containment scheme for mitigating and controlling the spread of COVID-19 in the United States.

References

1. Vox news, available at https://www.vox.com/policy-and-politics/2020/2/29/21159273/coronavirus-death-trump-health-officials-travel-ban-iran.

2. New Yorks Times Report, available at https://www.nytimes.com/2020/02/12/health/coronavirus-test-kits-cdc.html.

3. USA Today Report, available at https://www.usatoday.com/story/news/health/2020/03/28/coronavirus-fda-authorizes-abbott-labs-fast-portable-covid-test/2932766001/.

4. Why Germany's Coronavirus Death Rate Is Far Lower Than In Other Countries, available at https://www.npr.org/2020/03/25/820595489/why-germanys-coronavirus-death-rate-is-far-lower-than-in-other-countries.

5. USA Today Report, available at https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html.

6. F. BRAUER, *Compartmental models in epidemiology*, in Mathematical epidemiology, Springer, 2008, pp. 19–79.

7. D. BROCKMANN AND D. HELBING, *The hidden geometry of complex, network-driven contagion phenomena*, Science, 342 (2013), pp. 1337–1342.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

8. C. O. Buckee, S. Balsari, J. Chan, M. Crosas, F. Dominici, U. Gasser, Y. H. Grad, B. Grenfell, M. E. Halloran, M. U. Kraemer, et al., *Aggregated mobility data could help fight COVID-19*, Science (New York, NY), (2020).

9. M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, et al., *The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak*, Science, (2020).

10. E. Dong, H. Du, and L. Gardner, *An interactive web-based dashboard to track COVID-19 in real time*, The Lancet Infectious Diseases.

11. J. M. Drake, S. K. Chew, and S. Ma, *Societal learning in epidemics: intervention effectiveness during the 2003 SARS outbreak in singapore*, PloS One, 1 (2006).

12. Z. Du, L. Wang, S. Cauchemez, X. Xu, X. Wang, B. Cowling, and L. Meyers, *Risk for transportation of 2019 novel coronavirus disease from Wuhan to other cities in China*, Emerging infectious diseases, 26 (2020).

13. G. Evensen, *The ensemble kalman filter for combined state and parameter estimation*, IEEE Control Systems Magazine, 29 (2009), pp. 83–104.

14. L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, D. G. Bonsall, and C. Fraser, *Quantifying dynamics of SARS-CoV-2 transmission suggests that epidemic control and avoidance is feasible through instantaneous digital contact tracing*, medRxiv, (2020).

15. G. Grasselli, A. Pesenti, and M. Cecconi, *Critical care utilization for the COVID-19 outbreak in lombardy, italy: early experience and forecast during an emergency response*, JAMA, (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

16. J. HELLEWELL, S. ABBOTT, A. GIMMA, N. I. BOSSE, C. I. JARVIS, T. W. RUSSELL, J. D. MUNDAY, A. J. KUCHARSKI, W. J. EDMUNDS, F. SUN, ET AL., *Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts*, The Lancet Global Health, (2020).

17. H. W. HETHCOTE, *The mathematics of infectious diseases*, SIAM review, 42 (2000), pp. 599–653.

18. M. L. HOLSHUE, C. DEBOLT, S. LINDQUIST, K. H. LOFY, J. WIESMAN, H. BRUCE, C. SPITTERS, K. ERICSON, S. WILKERSON, A. TURAL, ET AL., *First case of 2019 novel coronavirus in the United States*, New England Journal of Medicine, (2020).

19. W. O. KERMACK AND A. G. MCKENDRICK, *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 (1927), pp. 700–721.

20. M. U. KRAEMER, C.-H. YANG, B. GUTIERREZ, C.-H. WU, B. KLEIN, D. M. PIGOTT, L. DU PLESSIS, N. R. FARIA, R. LI, W. P. HANAGE, ET AL., *The effect of human mobility and control measures on the COVID-19 epidemic in China*, Science, (2020).

21. A. J. KUCHARSKI, T. W. RUSSELL, C. DIAMOND, Y. LIU, J. EDMUNDS, S. FUNK, R. M. EGGO, F. SUN, M. JIT, J. D. MUNDAY, ET AL., *Early dynamics of transmission and control of COVID-19: a mathematical modelling study*, The Lancet Infectious Diseases, (2020).

22. S. LAI, I. I. BOGOCH, N. RUKTANONCHAI, A. G. WATTS, Y. LI, J. YU, X. LV, W. YANG, Y. HONGJIE, K. KHAN, ET AL., *Assessing spread risk of wuhan novel coronavirus within and beyond China, January-April 2020: a travel network-based modelling study*, (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

23. S. LAI, N. W. RUKTANONCHAI, L. ZHOU, O. PROSPER, W. LUO, J. R. FLOYD, A. WESOLOWSKI, C. ZHANG, X. DU, H. YU, ET AL., *Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak: an observational and modelling study*, medRxiv, (2020).

24. R. LI, S. PEI, B. CHEN, Y. SONG, T. ZHANG, W. YANG, AND J. SHAMAN, *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)*, Science, (2020).

25. Y. LIANG, S. GAO, Y. CAI, N. Z. FOUTZ, AND L. WU, *Calibrating the dynamic huff model for business analysis using location big data*, Transactions in GIS, (2020).

26. M. LIPSITCH, T. COHEN, B. COOPER, J. M. ROBINS, S. MA, L. JAMES, G. GOPALAKRISHNA, S. K. CHEW, C. C. TAN, M. H. SAMORE, ET AL., *Transmission dynamics and control of severe acute respiratory syndrome*, Science, 300 (2003), pp. 1966–1970.

27. A. L. LLOYD AND R. M. MAY, *Spatial heterogeneity in epidemic models*, Journal of theoretical biology, 179 (1996), pp. 1–11.

28. B. F. MAIER AND D. BROCKMANN, *Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in Mainland China*, arXiv preprint arXiv:2002.07572, (2020).

29. S. MELONI, N. PERRA, A. ARENAS, S. GÓMEZ, Y. MORENO, AND A. VESPIGNANI, *Modeling human mobility responses to the large-scale spreading of infectious diseases*, Scientific Reports, 1 (2011), p. 62.

30. D. MOLLISON, *Spatial contact models for ecological and epidemic spread*, Journal of the Royal Statistical Society: Series B (Methodological), 39 (1977), pp. 283–313.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

31. T. PRESTBY, J. APP, Y. KANG, AND S. GAO, *Understanding neighborhood isolation through spatial interaction network analysis using location big data*, Environment and Planning A: Economy and Space, p. 0308518X19891911.

32. S. REICH AND C. COTTER, *Probabilistic forecasting and Bayesian data assimilation*, Cambridge University Press, 2015.

33. D. B. RICHARDSON, N. D. VOLKOW, M.-P. KWAN, R. M. KAPLAN, M. F. GOODCHILD, AND R. T. CROYLE, *Spatial turn in health research*, Science, 339 (2013), pp. 1390–1392.

34. E. SHIM, A. TARIQ, W. CHOI, Y. LEE, AND G. CHOWELL, *Transmission potential of COVID-19 in south korea*, medRxiv, (2020).

35. H. TIAN, Y. LIU, Y. LI, M. U. KRAEMER, B. CHEN, C.-H. WU, J. CAI, B. LI, B. XU, Q. YANG, ET AL., *Early evaluation of transmission control measures in response to the 2019 novel coronavirus outbreak in China*, medRxiv, (2020).

36. H. TIAN, Y. LIU, Y. LI, C.-H. WU, B. CHEN, M. U. KRAEMER, B. LI, J. CAI, B. XU, Q. YANG, ET AL., *An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china*, Science, (2020).

37. H. C. TUCKWELL, L. TOUBIANA, AND J.-F. VIBERT, *Spatial epidemic network models with viral dynamics*, Physical Review E, 57 (1998), p. 2163.

38. J. WANG, K. TANG, K. FENG, AND W. LV, *When is the COVID-19 pandemic over? evidence from the stay-at-home policy execution in 106 Chinese cities*, Available at SSRN: https://ssrn.com/abstract=3561491, 106 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

39. M. S. WARREN AND S. W. SKILLMAN, *Mobility changes in response to COVID-19*, Descartes Labs, (2020).

40. J. T. WU, K. LEUNG, AND G. M. LEUNG, *Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study*, The Lancet, 395 (2020), pp. 689–697.

41. C. ZHOU, F. SU, T. PEI, A. ZHANG, Y. DU, B. LUO, Z. CAO, J. WANG, W. YUAN, Y. ZHU, ET AL., *COVID-19: Challenges to GIS with big data*, Geography and Sustainability, (2020).

42. X. ZHU, A. ZHANG, S. XU, P. JIA, X. TAN, J. TIAN, T. WEI, Z. QUAN, AND J. YU, *Spatially explicit modeling of 2019-nCoV epidemic trend based on mobile phone data in Mainland China*, medRxiv, (2020).

Acknowledgments

We would like to thank the SafeGraph Inc. for providing the anonymous and aggregated human mobility and place visit data. We would also like to thank all individuals and organizations for collecting and updating the COVID-19 observation data and reports. **Funding:** S.G. and Q.L. acknowledge the funding support provided by the National Science Foundation (Award No. BCS-2027375). Q.L. and S.C. acknowledge the Data Science Initiative of UW-Madison. X.S. acknowledges the Scholarly Innovation and Advancement Awards of Dartmouth College. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. **Author contributions:** Research design and conceptualization: Q.L., S.C., S.G.; Data collection and processing: S.C., S.G., Y.H.K.; Mathematical model implementation: Q.L., S.C.;

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

Result analysis: Q.L., S.G., X.S.; Visualization: S.C., S.G., Y.H.K.; Project administration: Q.L. S.G., X.S.; Writing: all authors.

Supplementary materials

Materials and Methods

Figs. S1 to S15

References *(S1-S13)*

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

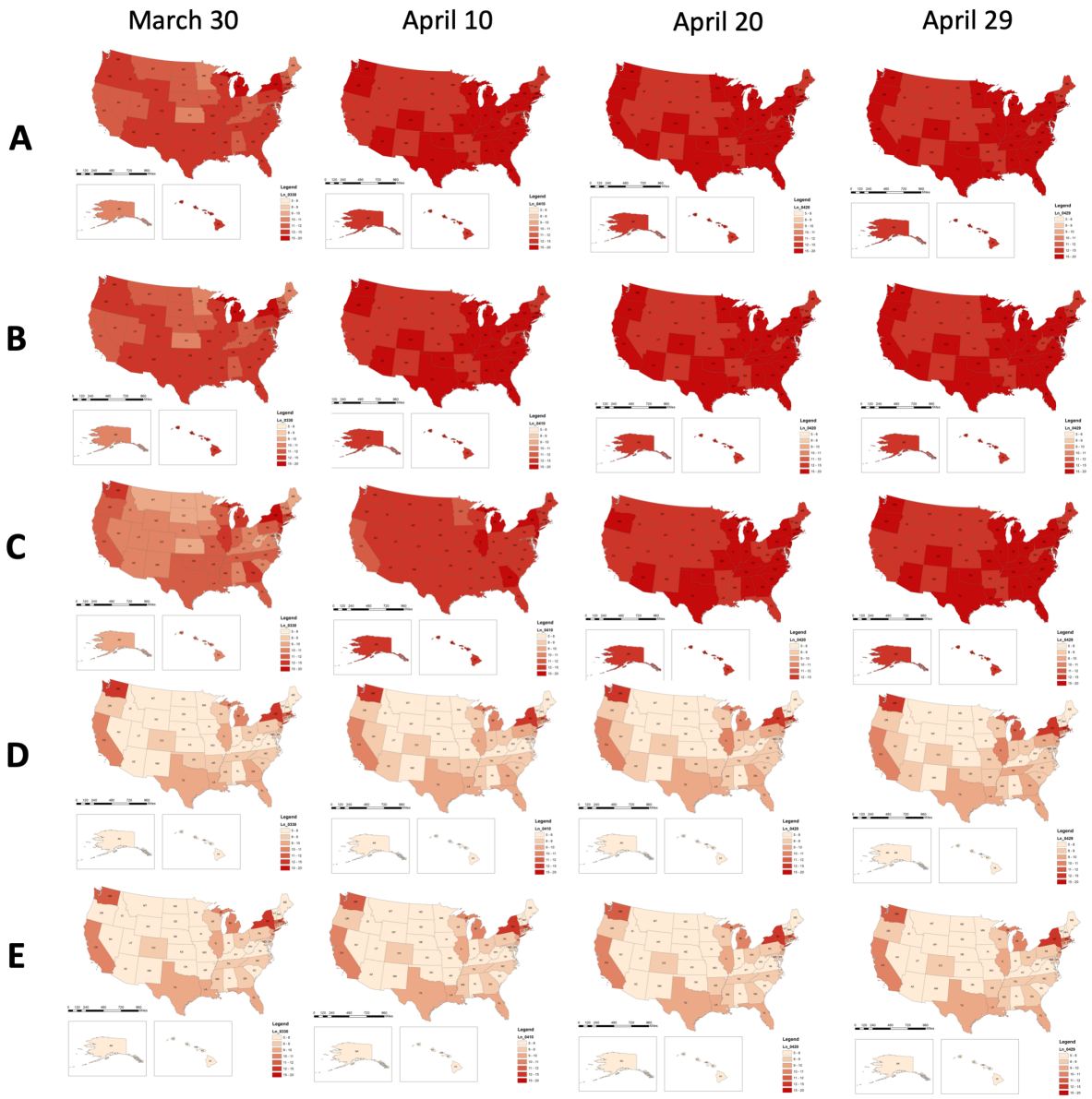

Figure 1: The spatiotemporal distribution of predicted infected population (in natural logarithm scale) across all states under different simulation scenarios: (A) $\alpha_r = 1$ and $\alpha_b = 1$, i.e., all parameters took the values of the initial configuration, obtained through data assimilation method using the numbers of confirmed cases during March 1 – March 20, 2020; (B) the travel flow was reduced to $\alpha_t = 0.05$, while other parameters values remained unchanged; (C) $\alpha_r = 0.1$ and $\alpha_b = 1$; (D) $\alpha_r = 1$ and $\alpha_b = 0.1$; (E) $\alpha_r = 0.1$, $\alpha_b = 0.1$. In the simulations, the transmission rate was set to be $b = \alpha_b b_0$ and the reporting rate was set to be $r = 1 - \alpha_r(1 - r_0)$. Where r_0 and b_0 were the reporting rate and the transmission rate on March 20, 2020, which are inferred from the data assimilation step.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

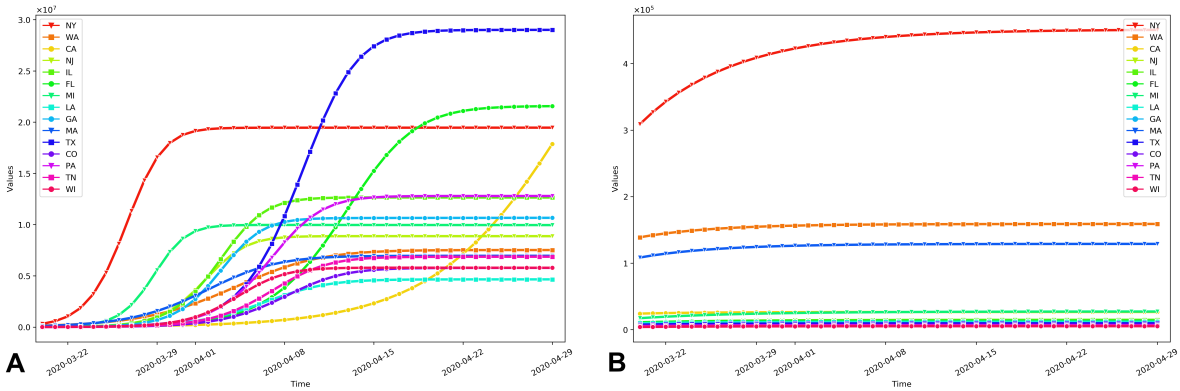

Figure 2: The prediction time series of the total infected population in the 15 most affected states under two scenarios: (A) $\alpha_r = \alpha_b = 1$, i.e., both the reported rate and the transmission rate remained unchanged; (B) $\alpha_r = \alpha_b = 0.1$, i.e., the transmission rate b was smaller and the reported rate r was larger (closer to 1) as $r = 1 - \alpha_r(1 - r_0)$.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.03.20052720; this version posted April 7, 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-NC 4.0 International license .

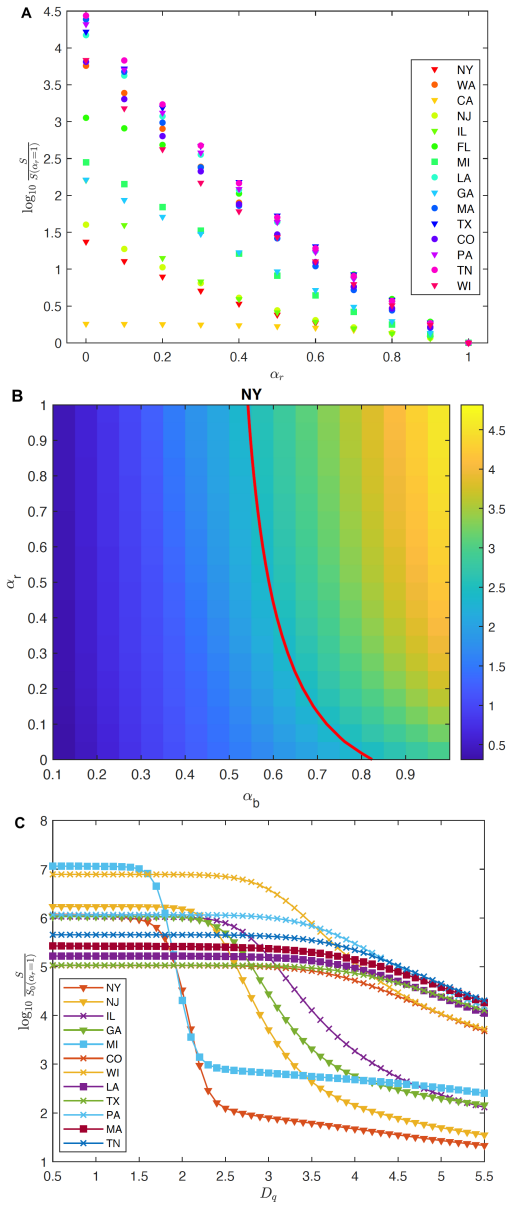

Figure 3: (A) Susceptible population (S) on April 29, 2020 as a function of α_r. $S(\alpha_r = 1)$ is the susceptible population on April 29 computed with the report rate set as the original report rate inferred from the data assimilation step. In all states, S increases as α_r decreases, meaning that more people stay unaffected when a higher report is enacted. (B) R_e, the basic reproduction number, on April 29 for different α_b and α_r in NY. The red line is the level set $R_e = 1$. It can be seen that increasing the reported rate helps diminish the reproductive number, but cannot reduce R_e under 1 if the original transmission rate b_0 is applied; (C) Susceptible population on April 29 for different D_q. $S(\alpha_r = 1)$ is the same as in (A). S significantly depends on the period from expose to quarantine.