Modeling the spatiotemporally dynamic transmissibility of Covid-19

Hao Yang[a]*, Xiaobai Yao[a*]

[a] Geography Department, University of Georgia
* hy96161@uga.edu; xyao@uga.edu

Keywords: Covid-19, effective reproductive number, spatiotemporal analysis

Introduction

The ongoing COVID-19 pandemic highlights the tremendous impact of infectious disease on people's lives and welfare and the need to develop rigorous theories and tools to understand and mitigate the disease's spreading. The rate of spread is a crucial indicator of the extent of the potential damage the disease may cause. The rates vary across space and in time. However, few studies have paid attention to the estimation and understanding of local transmissibility. From a geographical perspective, measuring local transmission rates and understanding contributing factors to the change of these rates over time are important research topics.

Aiming to measure the place- and time-specific rate of transmissibility, we choose to adopt the effective reproductive number (R^e) which is a metric to measure the time-specific transmissibility of an infectious disease while it is ongoing (Codeço et al., 2018; Farrington & Whitaker, 2003; Towers et al., 2014). The general concept of reproductive number in infectious disease research refers to the average number of secondary infections generated by an infectious case when an epidemic is ongoing. Due to geographical variations in susceptibility, social dynamics, mobility patterns, control policy, and the dynamic changes of them over time, the transmission rate of the epidemic varies by location and time. Whereas reproduction numbers of Covid-19 for many cities, regions, and countries have been extensively investigated and discussed (Bryant & Elofsson, 2020; Chen, 2020; Flaxman et al., 2020; Korolev, 2021; Arroyo-Marioli et al., 2021; Linka et al., 2020; Wilasang et al., 2021), little attention has been paid to the estimation of R^e at a fine spatial and temporal granularity. A bottleneck problem is the lack of practical estimation methods of effective reproductive numbers. A useful estimation method of R^e should be time- and place-specific.

Many studies have investigated the impact of human mobility, risk attitudes, and government policies on the spread of Covid-19 (Bryant & Elofsson, 2020; Caserotti et al., 2021; Flaxman et al., 2020b; Kraemer et al., 2020; Li et al., 2021; Zhou et al., 2020). Most of them examined the relationship with a single factor, such as human mobility and its impact on Covid-19 (Bryant & Elofsson, 2020), government policy and its impact on Covid-19 (Flaxman et al., 2020c), or at best two factors, such as the effect of human mobility and control measures on Covid-19 (Kraemer et al., 2020; Li et al., 2021). However, the interactions among these contributing factors and their impacts on the spread of the Covid-19 pandemic are more intricate. Thus, a comprehensive examination is in demand.

Therefore, the objective of this study is twofold. First, it presents a generally applicable R^e estimation modeling framework to obtain R^e values for any specific place and time during an epidemic, as long as regularly reported case data are available. Second, it examines the relationship between the R^e and a set of context factors including local human mobility and Covid intervention policy conditions.

Methods

The study develops a modeling framework and algorithm to estimate the space-time-specific effective reproduction numbers based on daily cases. A place-specific model is calibrated for each county in the United States to calculate time-specific (e.g. daily or weekly) R^e values based on daily cumulative circumstances in that county. The method is presented with a case study using daily county-level case data in the United States in 2020 and 2021. Then the spatiotemporal datasets of R^e values are coupled with data of context factors in corresponding spatial and temporal units.

After calibrating the R^e dataset, the study applies several machine learning and spatial analysis techniques to investigate the relationships between transmissibility rates and selected context factors such as human mobility changes. Among the modeling techniques applied in the study, the geographically weighted regression (GWR) is used to investigate the changing relationship in space, while exponential regression modeling is applied to establish county-specific models to examine the relationships from the temporal perspective.

Estimate place-time-specific effective reproduction numbers (R^e)

We develop a modeling framework to compute $R^e_{t,m}$ (effective R value at time t in place m) from daily reported cases. In our experiment, the county level is chosen because this is the finest spatial level where daily Covid-19 case data have been made available. The modeling framework first establishes the mathematical relationship between the time-series values of $R^e_{t,m}$ and the growth rates $\lambda_{t,m}$. Then based on the reported cumulative cases, it constructs the time series of active cases $I_{t,m}$ which can be used to calculate time-series growth rates $\lambda_{t,m}$. Finally, the time series of $R^e_{t,m}$ for each county can be derived from $\lambda_{t,m}$. The outline of the $R^e_{t,m}$ estimation model is illustrated in Figure 1.

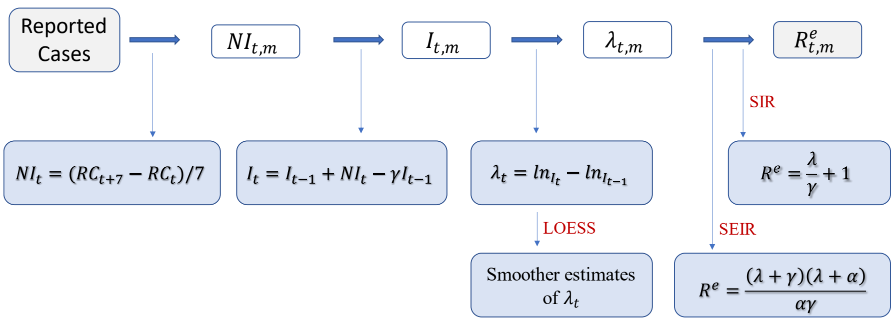

Figure 1. Logical steps in the space-time $R^e_{t,m}$ estimation algorithm

Spatiotemporal analysis

The pubic health literature has identified many contributing factors for the propagation of Covid-19, including human mobility patterns, government interventions, public risk perception, vaccination coverage, and many others. The effects of these context factors on the spread of the disease are nonstationary (Cazelles et al., 2005), meaning that the disease-context_factor associations vary across space and time. Therefore, findings drawn from a single place and time are not generalizable to other conditions. Analytical methods for modeling spatial-temporal nonstationarity in geographic relationships is essential. From the spatial perspective, we use the GWR model which takes non-stationary variables into consideration and models the local relationships between these context factors and the local R^e. From the temporal perspective, an exponential regression model is applied to explore the impact of context variables on the changes of R^e.

GWR model

GWR is an extension of ordinary least squares regression (OLS) by allowing the relationships between the independent and dependent variables to vary over space (Brunsdon et al., 1996). The model can be expressed mathematically by Equation (1) regression.

$$y_i = \beta_{i0} + \sum_k \beta_{ik} x_{ik} + \varepsilon_i \quad (1)$$

where y_i is the observation of the dependent variable in ith location, x_{ik} is the observation of the kth independent variable in ith location, ε_i is the error term, β_{ik} is the coefficient to be estimated for the respective independent x_i at location k. What makes GWR special is the localized relationship. Instead of generating global coefficients for respective variables, a GWR model allows the coefficients to be specific to a location i (e.g. county i). Therefore, this study applied GWR to examine the spatially varying relationship between R_0^e and socioeconomic and demographic characteristics of local areas at any timestamp t.

Exponential regression method

The exponential regression model and a few other machine learning techniques have been applied to understand how human behavioral and policy factors in a specific place affect the change of R_t^e over time. The exponential relationship is expressed in Equation (2). Here R_t^e is scaled from the baseline R_0^e in County m and evolves with changes in human mobility patterns, intervention policies, risk perception, vaccine rate, and mutations of the virus.

$$R_{m,t}^e = R_{m,0}^e e^{\sum_{i=1}^{n} \alpha_{m,i} * x_{m,i}} \qquad (2)$$

where $R_{m,t}^e$ is the effective reproduction number at time t in place (e.g. county) m. $x_{m,i}$ is the indicator variable for context factor i ($i \in 1, \dots, n$) and $\alpha_{m,i}$ is the coefficient for the variable. Note that each local place, a county in the case study, is modeled locally. Thus, the final result is a set of local exponential regression models, one for each county. The same strategy is used for all the other four techniques.

Results

Following the proposed Covid-19 effective reproduction number estimation framework, the study creates a spatiotemporal dataset of time-series R^e in the United States. In this case study, the spatial granularity is county level and the temporal scale is daily, while the estimation framework is generally applicable to other spatial and temporal scales. Figure 2 shows a snapshot of the produced spatiotemporal dataset of Covid-19 R^e in U.S. counties. This will help the local government monitor county-wide infection status and trends and understand what sort of messages or policies are most effective.

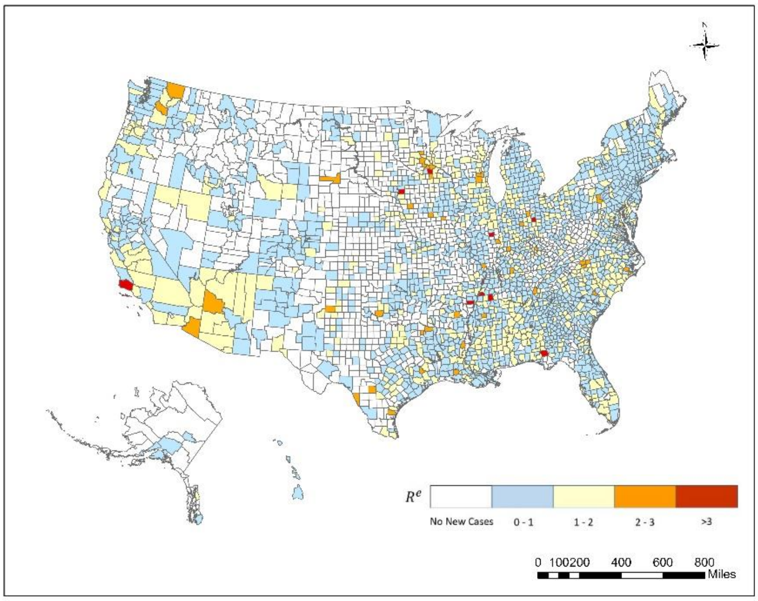

Figure 2. Estimated R^e of Covid-19 in the United States across counties on May 1st, 2020

An exponential regression models are constructed for each county separately. As an example, Table 1 shows the results for Los Angeles County. The results for other countries have a similar overall pattern, although specific values of coefficients differ. The study finds that R^e exhibits a positive association with the changes in human mobility targeting location categories of retail and recreation, grocery and pharmacy, and transit stations. It also finds a negative association with changes in human mobility of visiting parks. In addition, the vaccination rate and the proportion of removed cases in a place are found to be negatively correlated with R^e. Furthermore, facial covering regulations and gathering restrictions have a positive effect on reducing transmissibility.

County	R_0	Retail	Park	Transit Stations	Facial Covering	Vaccine Rate	Removed Cases	Delta variant	R^2
Los Angeles County	4.5	0.008	-0.003	0.005	-0.263	-0.034	-0.037	1.708	0.853

Table 1. The results of the exponential regression model for Los Angeles County

The GWR model is applied to analyze the relationship between local geographic characteristics and the R^e at t_0 which is the time when the pandemic was first announced in the United States. Figure 3 shows the spatially varying relationships between R^e and the context variables. Generally speaking, the results suggest that at the beginning of the pandemic in the United States, R^e exhibits a positive relationship with the proportion of male population, the percentage of votes for Republicans, and the air quality index of PM 2.5, while showing a negative association with the share of the elderly population (65 years and older). However, the significance and specific association of each variable vary across space.

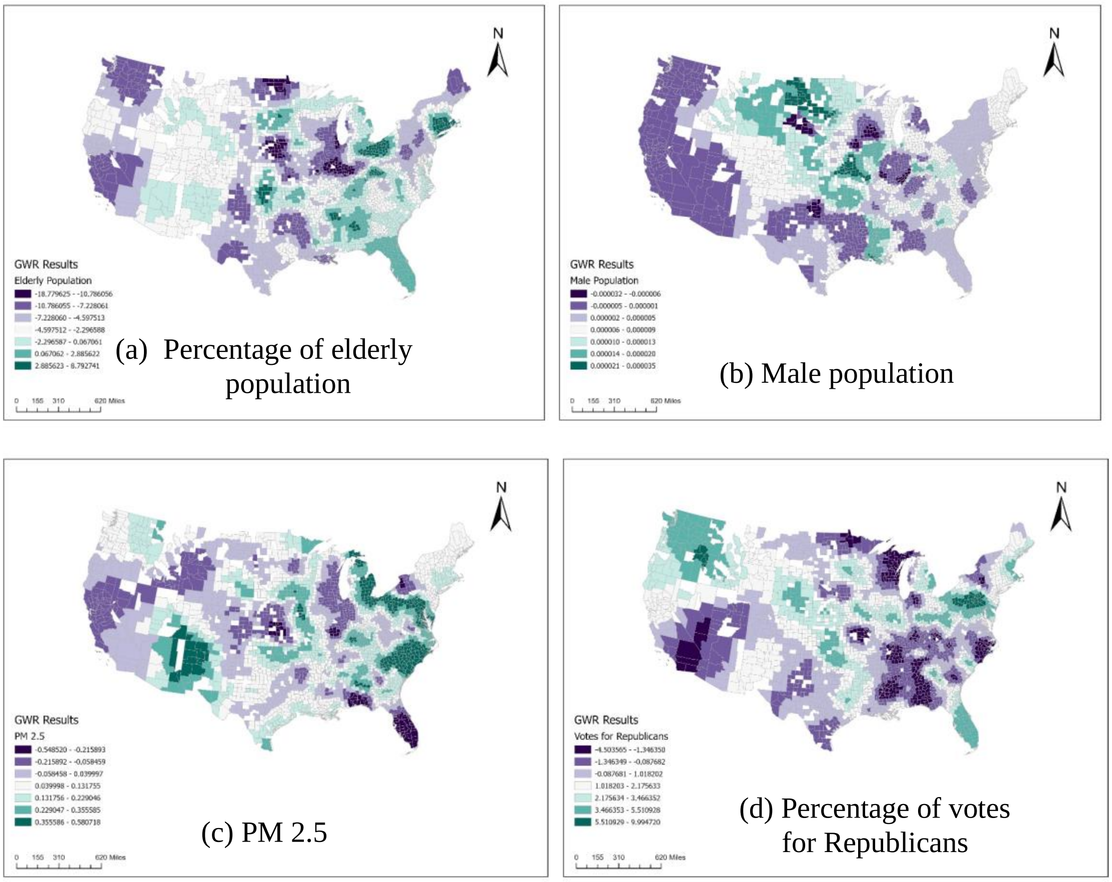

Figure 3. GWR results: local coefficients of selected independent variables

Discussion and Conclusion

The study presents an R^e estimation framework to obtain effective productive number (R^e) of Covid-19 and produces a spatiotemporal dataset of R^e at the county level in the United States. Moreover, the study investigates the relationship between the spatiotemporal dynamics of R^e with local human behavior factors. It was found that R^e is positively associated with human mobility (except mobility in parks) and the existence of Delta variant, and it is negatively associated with government intervention policies, mobility in parks, and depletion of the susceptibles. The level of impact by each factor varies geographically and differs by Covid-19 variants.

The scientific contributions are multifold. First, the R^e estimation framework is generally applicable to any part of the world at any spatial scale of which regularly reported time-series case data are available. This will allow researchers and policymakers to examine spatially and temporal changing R^e patterns so as to identify high-risk areas in space and time. Second, by modeling the relationship between R_e and local human behavior factors, people can use the model to better understand the effectiveness of various intervention policies.

References

Brunsdon, C., Fotheringham, S., & Charlton, M. (1996). Geographically weighted regression modelling spatial non-stationarity.

Bryant, P., & Elofsson, A. (2020). Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries. MedRxiv. https://doi.org/10.1101/2020.04.13.20063644

Caserotti, M., Girardi, P., Rubaltelli, E., Tasso, A., Lotto, L., & Gavaruzzi, T. (2021). Associations of COVID-19 risk perception with vaccine hesitancy over time for Italian residents. Social Science and Medicine, 272. https://doi.org/10.1016/j.socscimed.2021.113688

Chen, J. (2020). Pathogenicity and transmissibility of 2019-nCoV—A quick overview and comparison with other emerging viruses. Microbes and Infection, 22(2), 69–71. https://doi.org/10.1016/j.micinf.2020.01.004

Codeço, C. T., Villela, D. A. M., & Coelho, F. C. (2018). Estimating the effective reproduction number of dengue considering temperature-dependent generation intervals. Epidemics, 25, 101–111. https://doi.org/10.1016/j.epidem.2018.05.011

Farrington, C. P., & Whitaker, H. J. (2003). Estimation of effective reproduction numbers for infectious diseases using serological survey data. In Biostatistics. https://academic.oup.com/biostatistics/article/4/4/621/246942

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., Monod, M., Perez-Guzman, P. N., Schmit, N., Cilloni, L., Ainslie, K. E. C., Baguelin, M., Boonyasiri, A., Boyd, O., Cattarino, L., … Bhatt, S. (2020a). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257–261. https://doi.org/10.1038/s41586-020-2405-7

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., Monod, M., Perez-Guzman, P. N., Schmit, N., Cilloni, L., Ainslie, K. E. C., Baguelin, M., Boonyasiri, A., Boyd, O., Cattarino, L., … Bhatt, S. (2020b). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257–261. https://doi.org/10.1038/s41586-020-2405-7

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., Monod, M., Perez-Guzman, P. N., Schmit, N., Cilloni, L., Ainslie, K. E. C., Baguelin, M., Boonyasiri, A., Boyd, O., Cattarino, L., … Bhatt, S. (2020c). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257–261. https://doi.org/10.1038/s41586-020-2405-7

Gao, S., Rao, J., Kang, Y., Liang, Y., & Kruse, J. (n.d.). Mapping county-level mobility pattern changes in the United States in response to COVID-19. https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html

Korolev, I. (2021). Identification and estimation of the SEIRD epidemic model for COVID-19. Journal of Econometrics, 220(1), 63–85. https://doi.org/10.1016/j.jeconom.2020.07.038

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Covid-19 Data, O., Group, W., du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. v. (n.d.). The effect of human mobility and control measures on the COVID-19 epidemic in China. https://www.science.org

Li, Y., Li, M., Rice, M., Zhang, H., Sha, D., Li, M., Su, Y., & Yang, C. (2021). The impact of policy measures on human mobility, COVID-19 cases, and mortality in the US: A spatiotemporal perspective. International Journal of Environmental Research and Public Health, 18(3), 1–25. https://doi.org/10.3390/ijerph18030996

Martorell-Marugán, J., Villatoro-García, J. A., García-Moreno, A., López-Domínguez, R., Requena, F., Merelo, J. J., Lacasaña, M., de Dios Luna, J., Díaz-Mochón, J. J., Lorente, J. A., & Carmona-Sáez, P. (2021). DatAC: A visual analytics platform to explore climate and air quality indicators associated with the COVID-19 pandemic in Spain. Science of the Total Environment, 750. https://doi.org/10.1016/j.scitotenv.2020.141424

Towers, S., Patterson-Lomba, O., & Castillo-Chavez, C. (2014). Temporal Variations in the Effective Reproduction Number of the 2014 West Africa Ebola Outbreak. PLoS Currents. https://doi.org/10.1371/currents.outbreaks.9e4c4294ec8ce1adad283172b16bc908

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L., Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Chen, H., & Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet, 395(10229), 1054–1062. https://doi.org/10.1016/S0140-6736(20)30566-3