medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

Spatial-temporal relationship between population mobility and COVID-19 outbreaks in South Carolina: A time series forecasting analysis

Chengbo Zeng, MS [a,b,c] Jiajia Zhang, PhD [a,c,d] Zhenlong Li, PhD [a,c,e] Xiaowen Sun, MS [a,c,d]
Bankole Olatosi, PhD [a,c,f] Sharon Weissman, MD [c,g] Xiaoming Li, PhD [a,b,c]

a) South Carolina SmartState Center for Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA
b) Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA
c) University of South Carolina Big Data Health Science Center, Columbia, SC, USA
d) Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA
e) Geoinformation and Big Data Research Lab, Department of Geography, School of Art and Science, University of South Carolina, Columbia, SC, USA
f) Department of Health Services, Policy, and Management, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA
g) School of Medicine, University of South Carolina, Columbia, SC, USA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2021.01.02.21249119; this version posted January 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

Corresponding author:

Chengbo Zeng, MS, PhD Candidate

South Carolina SmartState Center for Healthcare Quality

University of South Carolina Big Data Health Science Center

Department of Health Promotion, Education, and Behavior

Arnold School of Public Health, University of South Carolina

Discovery I, 915 Greene Street, Columbia, SC 29208, United States

Phone: 803-477-5298

Email: czeng@email.sc.edu

Word count: 2,348

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Key points: How did population mobility impact and predict the coronavirus disease 2019 (COVID-19) outbreaks in South Carolina (SC)?

Findings: This study leveraged disease surveillance data, Twitter-based mobility data, and Poisson time series model to investigate the relationship between population mobility and COVID-19 daily incidences at both state- and county- levels in SC. Population mobility was positively associated with the COVID-19 outbreaks and could provide an acceptable prediction of COVID-19 daily new cases within two-week window.

Meaning: These findings suggested that population mobility affects the COVID-19 transmission and using Twitter-based mobility data could provide acceptable prediction for further COVID-19 spread as well as inform proactive measures for disease control.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Abstract

Importance: Population mobility is closely associated with coronavirus 2019 (COVID-19) transmission, and it could be used as a proximal indicator to predict future outbreaks, which could inform proactive non-pharmaceutical interventions for disease control. South Carolina (SC) is one of the states which reopened early and then suffered from a sharp increase of COVID-19.

Objective: To examine the spatial-temporal relationship between population mobility and COVID-19 outbreaks and use population mobility to predict daily new cases at both state- and county- levels in SC.

Design, setting, and participants: This longitudinal study used disease surveillance data and Twitter-based population mobility data from March 6 to November 11, 2020 to describe their changes, examine their correlation, and forecast daily COVID-19 new cases in two-week window in SC and its top five counties with the largest number of cumulative confirmed cases. Poisson count time series model was employed to carry out the research goals.

Main outcome and measure: The main outcome was daily new case which was calculated by subtracting the cumulative confirmed cases of previous day from the total cases. Population mobility was assessed using the number of users with travel distance larger than 0.5 mile which was calculated based on their geotagged twitters.

Results: Population mobility was positively associated with state-level daily COVID-19 incidence and those of the top five counties (i.e., Charleston, Greenville, Horry, Spartanburg, Richland). At the state-level, final model with time window within the last 7-day had the smallest prediction error, and the prediction accuracy was as high as 98.7%, 90.9%, and 81.6% for the next 3-, 7-, 14- days, respectively. Among Charleston, Greenville, Horry, Spartanburg, and Richland counties, the best predictive models were established based on their observations in

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 last 9-, 14-, 28-, 20-, and 9- days, respectively. The 14-day prediction accuracy ranged from 60.3% to 74.5%.

Conclusions and relevance: Population mobility was positively associated with COVID-19 incidences at both state- and county- levels in SC. Using Twitter-based mobility data could provide acceptable prediction for COVID-19 daily new cases. Population mobility measured via social media platform could inform proactive measures and resource relocations to curb disease outbreaks and their negative influences.

Keywords: COVID-19; Mobility; Incidence; South Carolina

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Introduction

Since the first confirmed case of Coronavirus Disease 2019 (COVID-19) in the United States (US) on January 21, 2020, the countrywide COVID-19 outbreaks have surged quickly. As of December 29, there were 19,566,140 cumulative confirmed cases and 338,769 COVID-19 related deaths in US.[1] South Carolina (SC), a state located in Southeastern US, had the first confirmed cases on March 6, 2020. From March to May, the trends of daily new cases were flat with an average of daily increased cases less than 500. However, with the early implementation of reopening policies, the daily new cases in SC have risen sharply since June. On July 14, the COVID-19 cases in SC surpassed 60,000, with more than 2,200 daily new cases, the second highest increase in one day in US.[2] Between August and October, the transmission rate slowed down with the further implementation of non-pharmaceutical interventions (NPIs), such as dine-in service restriction and face-covering requirement, but increased steadily after October. By December 29, there were 300,602 reported cases and 5,198 deaths in SC.[3] Given the rapid transmission of COVID-19 in SC, more research is needed to identify potential early predictors of increasing transmission rates which could then be used to inform proactive NPIs to suppress statewide disease transmission.

Population mobility is a potential early indicator of COVID transmission as population mobility reflects the influences (both positive and negative) of NPIs, reopening actions, social distancing practices and public holidays.[4-6] For instance, at the early stage of COVID-19 epidemic, the SC Governor issued a series of NPIs, such as shelter-in-place and school and non-essential business closure, to reduce social interaction. These NPIs showed a positive effect in suppressing the statewide COVID-19 spread. Later in May, the reopening policies and public holidays diluted the implementation of NPIs leading to the increased social interactions and

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

statewide COVID-19 spread.[7,8] As it may be difficult to measure the real-time impact of reopening policies, public holidays and fidelity of NPIs implementation, population mobility could be a proximal indicator allowing for real-time COVID-19 transmission forecasting.

Social media platforms, such as Twitter, could collect geospatial information and closely monitor the change of population mobility.[9,10] Indeed, the tremendous volume of user-generated geoinformation from social media could promote the real-time or near real-time surveillance of population mobility and provide timely data of how population mobility respond to different phases of COVID-19 outbreak, policy reactions, and public holidays.[11-13] Several studies have leveraged mobility data from social media (e.g., Google, Facebook, Twitter) to investigate the relationship between population mobility and COVID-19 transmission.[6,8,14-16] These studies identified a consistently positive relationship between population mobility and COVID-19 incidence. However, few studies used population mobility as a predictor to forecast further outbreaks and to evaluate the prediction accuracy in addition to correlation analysis. One study by Wang and Yamamoto predicted COVID-19 daily new cases in Arizona using disease surveillance data, Google Community Mobility report, and partial differential equation. They found acceptable prediction for the next 3-day.[16] This study only classified Arizona into three regions (i.e., central, northern, southern of Arizona) and evaluated the prediction accuracy for the next 3-day which did not cover the duration of viral incubation (i.e., 14-day). More studies are needed to investigate the relationship between population mobility using social media data and COVID-19 transmission at both state- and county- levels and over longer timeframes.

Leveraging disease surveillance data and Twitter-based population mobility, the current study aimed to construct time series models of COVID-19 daily new cases, investigate the

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

relationship between them, and evaluate the prediction accuracy of daily new cases for the next two-week window at both state- and county- levels in SC.

Methods

COVID-19 incidence data

Cumulative confirmed cases of COVID-19 through November 11, 2020 at both state- and county-levels in SC were collected from The New York Times, which was deposited in Github (https://github.com/nytimes/covid-19-data). Within the study period (March 6, 2020 [date of 1st COVID diagnosis in SC] to November 11, 2020 [251st day]), daily new cases were calculated by subtracting the cumulative confirmed cases of previous day from the total cases for the entire state and its five counties with largest numbers of cumulative confirmed cases (i.e., Charleston, Greenville, Horry, Spartanburg, and Richland). The study protocol was approved by the Institutional Review Boards at the University of South Carolina.

Population mobility

Population mobility was assessed using the number of people (Twitter users) with moving distance larger than 0.5 mile per day in SC and the selected counties. The methodology of extracting daily population movement (origin-destination flows) from geotagged tweets has been described in detailed elsewhere.[17,18] Briefly, geotagged tweets during the study periods were collected and used for calculation. Only users who post at least twice a day or posted tweets on at least two consecutive days were included in the calculation. Daily travel distance was calculated for each user based on the derived origin-destination flows and used to generate a variable of how many people moved each day (with travel distance larger than 0.5 mile). This method of capturing population mobility through Twitter had been validated by previous study.[17]

Statistical analysis

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

First, daily new cases of COVID-19 and population mobility at both state- and county-levels were described using line charts in R version 3.6.3 (The R Foundation, "ggplot" package). Daily new cases and mobility were also described using fives quantiles (i.e., minimum, 25th percentile, 50th percentile, 75th percentile, and maximum) by each month.

Second, Poisson count time series model was used to model the impact of population mobility on the daily new cases of COVID-19 at state-level. Time series models were built at the various time windows. At the first round selection, a total of 17 time windows (by a 7-day increment) were considered including 1 to 7 days, 1 to 14 days,…, and 1 to 119 days. The daily new cases from the 1st to 234th days were used as the training dataset and those from the next 3-day (235th ~ 237th) were used as testing dataset for the purpose of model evaluation. With the smallest prediction error (Formula 1) and good interpretation, model with the best time window was selected. Once the best time window in the first round selection was determined, the second and third round selections would be conducted to narrow down the time window and obtain the final model with the smallest prediction error. The final model was used to predict the COVID-19 daily new cases for the next 3-, 7-, and 14- days (238th ~ 251st days). Cumulative difference (Formula 2) between observed and predicted cases and mean absolute percentage accuracy (Formula 3) by each timeframe were reported.[16]

$$\sqrt{\sum_{d=1}^{3}\left(\frac{x_o - x_p}{x_o}\right)^2} \qquad \text{(Formula 1)}$$

$$\sum_{d=1}^{n}\left|x_0 - x_p\right| \qquad \text{(Formula 2)}$$

$$1 - \frac{\sum_{d=1}^{n}\left|x_o - x_p\right|}{\sum_{d=1}^{n} x_o} \qquad \text{(Formula 3)}$$

Notes: *d*: day; *n*: next 3-, 7-, or 14- days; *o*: observed value; *p*: predicted value; *x*: daily new cases.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Finally, similar analytic procedure was performed to construct the final model at the county-level for each of the top five counties (i.e., Charleston, Greenville, Horry, Spartanburg, and Richland) in SC. Poisson count time series model was conduct using the R package ("tscount").

Results

Descriptive statistics

Table 1 shows the descriptive statistics of COVID-19 new cases at both state- and county-levels. By October 31, there were 176,612 cumulative COVID-19 confirmed cases in SC. The cumulative confirmed cases in Charleston, Greenville, Horry, Spartanburg, and Richland were 17,384, 18,021, 12,591, 9,290, and 17,531, respectively. At the state-level, the daily new cases from March to the end of May were less than 500. From June to the middle of July, the daily new cases elevated, with 2,217 new COVID-19 patients on July 14. After that, the transmission rate decreased, with most of the daily new cases less than 1,500. After October, the daily new cases steadily increased.

At the county- level, the top five counties showed a similar trend of COVID-19 outbreaks and accounted for more than 40.0% of the total cases in SC. The daily new cases increased earlier in Greenville than the other four counties.

Regarding the population mobility, the trends at both state- and county- levels were similar. The numbers of people in SC (Twitter users in our data) with a moving distance of more than 0.5 mile decreased from 1,400 to 550 between March 6 and April 9, 2020. Although there were slight increases from the middle of April to that of June, the numbers were consistently around 1,000 after this timeframe. At the county-level, each of the five counties had less than 200 people with moving distance larger than 0.5 mile after the middle of March.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Model selection of time series analyses

Following the model selection procedure, Poisson count time series model of COVID-19 incidence at the state-level was constructed using daily new cases and population mobility. Population mobility was positively associated with state-level COVID-19 daily new cases (β=0.818, 95%*CI*: 0.761~0.876), and model using the past 7- day (1~7 days) as time window had the smallest prediction error (Table 2). The prediction error of new cases in the next 3- day (235th ~ 237th) was 0.218.

At the county-level, similar modelling procedure was employed. Population mobility was consistently and positively associated with COVID-19 new cases across the top five counties. The best time windows for Charleston, Greenville, Horry, Spartanburg, and Richland were 9-, 14-, 28-, 20-, and 9- days, respectively. Table 2 presents the detailed results of final model, correlation analysis, and 3-day prediction error at both state- and county- levels.

COVID-19 daily new cases forecasting

Table 2 also presents the results of forecasting and prediction accuracy. Using final models with the selected time windows, COVID-19 daily new cases were forecasted for the next 14-day at both state- and county- levels. At the state- level, the 3-day cumulative difference and prediction accuracy were 42 and 98.7%, respectively. As compared to the 3-day predication accuracy, the 7- and 14- day accuracy reduced to 90.9% and 81.6%. At the county- level, among the top five counties, the 3-day prediction accuracy ranged from 69.0% to 99.3%. The prediction accuracy deceased in Charleston, Greenville, and Spartanburg with the increase of time span. In contrast, the prediction accuracy in Horry and Richland increased in 7-day prediction but reduced in 14-day prediction. The 14- day prediction accuracy among Horry and Richland were close to their values in 3-day prediction.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Discussion

This study leveraged disease surveillance data and Twitter-based population mobility to test the relationship between mobility and COVID-19 daily new cases and forecast the future transmission rates during the next 14 days at both state- and county- levels in SC. Results revealed that population mobility was significantly and positively associated with COVID-19 daily new cases. Using the selected models to forecast COVID-19 transmission, we found that although the prediction accuracy at state- level and most of the selected counties decreased with the increase of time span, the prediction accuracy was acceptable. To the best of our knowledge, this is the first study that combined correlation analysis and forecasting together and investigated the impacts of population mobility on COVID-19 spread at both state- and county- levels.

Population mobility could reflect the impacts of NPIs, reopening policies, and public holidays and estimate the social movement during the current COVID-19 pandemic. It is closely related to the COVID-19 outbreaks, which is in accordance with that of prior research.[6,8,14-16] This study added values to previous studies by examining the impacts of population mobility on COVID-19 incidence and evaluate its prediction efficacy at both state- and county- levels in SC during the two-week window. Although this indicator could only reflect the mobility among people who used Twitter, the results still revealed the positive correlation between mobility and COVID-19 spread.

Additionally, using Twitter-based mobility data to predict COVID-19 daily new cases could provide acceptable accuracy, which could also justify the validity and prediction efficacy of this indicator. The high prediction accuracy at the state-level was consistent with Wang's finding in Arizona.[16] However, such a high prediction accuracy could not be found at the county-level. One possible explanation for these findings is that we did not capture the influences of

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

contextual factors (i.e., population density) and the roles of mitigating factors (e.g., wearing face mask, practicing social distancing).[16,19,20] Additionally, the Twitter-based mobility did not differentiate the social movement at different locations, such as movement around parks, workplace, and retail places, which showed different impacts on the COVID-19 incidence.[6] Furthermore, in this study, we only captured population mobility at state- and county- levels while population mobility at zip code level might provide more accurate prediction. Finally, compared with mobility data from other platforms (e.g., Facebook, Google, Safegraph, Apple), our Twitter-based mobility only estimated how many people with moving distance larger than a specific value. Nevertheless, the findings generating from this study confirmed the spatial-temporal relationship between Twitter-based mobility and COVID-19 outbreaks in SC as well as the prediction efficacy of population mobility.

Use of population mobility data has potential implications in future research and practices to curb COVID-19 outbreaks. From a research perspective, studies on mobility and COVID-19 could be studied at state-, county-, and/or zip code levels. In addition, mobility around different locations could provide detailed information regarding COVID-19 transmission, identify the most relevant mobility associated with daily new cases, and inform tailored interventions on social distancing by different locations to control disease outbreaks. Furthermore, the geospatial difference in the prediction accuracy of population mobility on daily new cases by county suggested that contextual factors, such as demographic characteristics and implementation fidelity of NPIs at county-level, should be accounted for in future research. Finally, as the incubation and transmission of COVID-19 are closely associated with time-varying factors, such as temperature and weather, their impacts are also need to be considered in forecasting studies.[21] Regarding the practice of disease control and prevention, leveraging social media platform to

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

monitor daily population mobility could be used to improve the predictions of further COVID-19 spread, inform proactive NPIs, and arrange healthcare resources to reduce disease morbidity and mortality.[22,23]

Conclusion

Population mobility was positively associated with COVID-19 transmission at both state- and county- levels in SC. Using Twitter-based mobility data could provide acceptable prediction for COVID-19 daily new cases. The application of social media platforms to monitor population mobility and predict COVID-19 spread could inform proactive measures to curb disease outbreaks and buffer their negative influences.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Acknowledgement

This study was supported by the National Institute of Health (NIH) Research Grant R01AI127203-01A by National Institute of Allergy and Infectious Diseases and National Science Foundation (NSF) Grant No. 2028791.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

Reference

1. Centers for Disease Control and Prevention. Coronavirus Disease 2019 (COVID-19): Cases & Deaths in the US. 2020. Retrived from: https://www.cdc.gov/coronavirus/2019-ncov/index.html
2. Bohatch E. SC sees more than 2,200 coronavirus cases, second highest increase in one day. *The State*. 2020. Retrived from: https://www.thestate.com/news/coronavirus/article244215817.html
3. South Carolina Department of Health and Environmental Control. SC Demographic Data (COVID-19). 2020. Retrived from: https://www.scdhec.gov/infectious-diseases/viruses/coronavirus-disease-2019-covid-19/sc-demographic-data-covid-19
4. Drake TM, Docherty AB, Weiser TG, Yule S, Sheikh A, Harrison EM. The effects of physical distancing on population mobility during the COVID-19 pandemic in the UK. *Lancet Digit Health.* 2020;2(8):e385-e387. doi: 10.1016/S2589-7500(20)30134-5
5. Wang S, Liu Y, Hu T. Examining the change of human mobility adherent to social restriction policies and its effect on COVID-19 cases in Australia. *Int J Environ Res Public Hea*lth. 2020;17(21). doi: 10.3390/ijerph17217930
6. Sulyok M, Walker M. Community movement and COVID-19: A global study using Google's Community Mobility Reports. *Epidemiol Infect.* 2020;148:e284. doi: 10.1017/S095026882002757
7. Gibbs H, Liu Y, Pearson CAB, et al. Changing travel patterns in China during the early stages of the COVID-19 pandemic. *Nat Commun*. 2020;11(1):5012. doi: 10.1038/s41467-020-18783-0

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

8. Xiong C, Hu S, Yang M, Luo W, Zhang L. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc Natl Acad Sci USA.* 2020;117(44):27087-27089. doi: 10.1073/pnas.2010836117

9. Hawelka B, Sitko I, Beinat E, Sobolevsky S, Kazakopoulos P, Ratti C. Geo-located Twitter as proxy for global mobility patterns. *Cartogr Geogr Inf Sc.* 2014;41(3):260-271. doi: 10.1080/15230406.2014.890072

10. Martin Y, Li Z, Ge Y. Towards real-time population estimates: introducing Twitter daily estimates of residents and non-residents at the county level. *arXiv preprint* 2020. arXiv:2011.13482

11. Yang C, Sha D, Liu Q, et al. Taking the pulse of COVID-19: A spatiotemporal perspective. *International Journal of Digital Earth.* 2020;13(10):1186-1211. doi: 10.1080/17538947.2020.1809723

12. Li Z, Li X, Porter D, et al. Monitoring the spatial spread of COVID-19 and effectiveness of the control measures through human movement using big social media data: A study protocol. *JMIR Res Protoc.* 2020. doi: 10.2196/24432

13. Huang X, Li Z, Jiang Y, et al. The characteristics of multi-source mobility datasets and how they reveal the luxury nature of social distancing in the US during the COVID-19 pandemic. *medRxiv*. 2020. doi: 10.1101/2020.07.31.20143016

14. Bisanzio D, Kraemer MUG, Bogoch, II, Brewer T, Brownstein JS, Reithinger R. Use of Twitter social media activity as a proxy for human mobility to predict the spatiotemporal spread of COVID-19 at global scale. *Geospat Health.* 2020;15(1). doi: 10.4081/gh.2020.882.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

15. Gatalo O, Tseng K, Hamilton A, Lin G, Klein E. Associations between phone mobility data and COVID-19 cases. *Lancet Infect Dis.* 2020. doi: 10.1016/S1473-3099(20)30725-8

16. Wang HY, Yamamoto N. Using a partial differential equation with Google Mobility data to predict COVID-19 in Arizona. *Math Biosci Eng.* 2020;17(5):4891-4904. doi: 10.3934/mbe.2020266

17. Huang X, Li Z, Jiang Y, Li X, Porter D. Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLoS One.* 2020;15(11):e0241957. doi: 10.1371/journal.pone.0241957

18. Li Z, Huang X, Ye X, Li X. ODT flow explorer: Extract, query, and visualize human mobility. *ArXiv [Preprint] 2020 Nov 26:arXiv:201112958v1* 2020.

19. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *Lancet Infect Dis.* 2020;20(11):1247-1254. doi: 10.1016/S1473-3099(20)30553-3

20. Gatalo O, Tseng K, Hamilton A, Lin G, Klein E, Program CDCM-H. Associations between phone mobility data and COVID-19 cases. *Lancet Infect Dis.* 2020. doi: 10.1016/S1473-3099(20)30725-8

21. Carlson CJ, Gomez ACR, Bansal S, Ryan SJ. Misconceptions about weather and seasonality must not misguide COVID-19 response. *Nat Commun.* 2020;11(1):4312. doi: 10.1038/s41467-020-18150-z

22. Kurian SJ, Bhatti AUR, Ting HH, Storlie C, Shah N, Bydon M. Utilization of mobility data in the fight against COVID-19. *Mayo Clin Proc Innov Qual Outcomes.* 2020. doi: 10.1016/j.mayocpiqo.2020.10.003

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

23. Buckee CO, Balsari S, Chan J, et al. Aggregated mobility data could help fight COVID-19. *Science.* 2020;368(6487):145-146. doi: 10.1126/science.abb8021

Table 1. Descriptive statistics of population mobility and COVID-19 new cases at both state- and county- levels

		Minimum	25th percentile	50th percentile	75th percentile	Maximum
Population mobility						
State-level						
	March	658	809	1,010	1,109	1,438
	April	554	617	670	697	786
	May	630	754	812	848	940
	June	756	848	871	910	993
	July	818	870	896	937	1,039
	August	767	828	863	884	1,035
	September	784	831	875	907	1,021
	October	789	843	898	965	1,085
County-level						
Charleston						
	March	81	104	126	142	195
	April	62	75	83	92	98
	May	73	93	104	121	140
	June	96	109	116	126	154
	July	95	109	117	121	133
	August	88	99	109	120	134
	September	94	110	116	126	150
	October	95	113	122	132	142
Greenville						
	March	103	115	139	156	177
	April	82	93	106	114	134
	May	100	113	119	127	132
	June	104	117	124	133	162
	July	107	124	135	140	153
	August	111	129	140	146	168
	September	114	128	140	144	158
	October	104	133	138	149	169
Horry						

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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.

March	77	84	87	116	158
April	53	64	71	80	97
May	76	87	100	128	151
June	103	113	125	133	162
July	100	116	123	140	171
August	89	112	118	137	160
September	79	96	107	117	143
October	71	99	105	116	151
Spartanburg					
March	40	67	82	89	106
April	34	43	47	50	61
May	47	51	56	62	72
June	50	62	65	72	78
July	51	67	76	85	101
August	50	65	70	77	94
September	55	62	65	70	74
October	52	59	67	79	92
Richland					
March	58	76	82	93	120
April	53	68	73	78	84
May	61	69	77	84	115
June	65	77	86	93	105
July	59	76	82	95	105
August	72	79	89	95	109
September	72	82	89	97	125
October	72	84	92	100	119
COVID-19 new cases					
State-level					
March	0	3	18	74	158
April	62	131	154	204	275
May	82	129	164	228	467
June	236	476	757	1,115	1,755
July	972	1,520	1,726	1,855	2,374

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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.

August	456	722	937	1,214	1,583
September	301	624	863	1,190	2,665
October	381	789	912	1,057	1,706
County-level					
Charleston					
March	0	0	1	8	32
April	0	3	5	12	48
May	0	1	6	8	23
June	11	34	69	200	373
July	85	164	221	303	418
August	25	53	95	105	218
September	0	35	46	65	425
October	13	34	50	61	89
Greenville					
March	0	1	5	11	18
April	0	9	19	28	54
May	7	14	21	33	150
June	47	71	115	147	245
July	49	129	167	196	276
August	14	40	53	95	184
September	6	41	75	113	289
October	27	87	107	140	197
Horry					
March	0	1	2	3	5
April	0	2	5	9	18
May	0	4	5	10	26
June	17	47	99	133	221
July	63	103	145	189	358
August	16	30	41	56	115
September	4	20	30	46	70
October	15	48	73	90	139
Spartanburg					
March	0	0	0	2	7

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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.

	April	1	4	6	11	32
	May	1	4	7	14	61
	June	5	18	34	44	72
	July	18	48	63	84	125
	August	11	25	44	62	92
	September	2	18	50	99	215
	October	0	46	78	96	147
Richland						
	March	1	3	6	14	37
	April	3	15	25	32	56
	May	5	15	19	26	33
	June	12	44	67	81	155
	July	57	108	138	165	234
	August	39	79	93	124	408
	September	34	77	96	142	766
	October	24	51	67	78	130

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 2. The impacts of population mobility on COVID-19 outbreaks in SC

	State- level	County- level				
		Charleston	**Greenville**	**Horry**	**Spartanburg**	**Richland**
Model training						
Time windows	1-7	1-9	1-14	1-28	1-20	1-9
Coefficient of population mobility (95%*CI*)	0.818 (0.761,0.876)	0.486 (0.338,0.634)	0.278 (0.165,0.390)	0.395 (0.275,0.515)	0.220 (0.118,0.422)	0.167 (0.067,0.246)
Model evaluation (3-day prediction error)	0.218	1.752	0.217	2.778	0.363	0.435
Forecasting						
Prediction						
238th day	1,097	64	128	71	86	64
239th day	1,031	68	135	43	102	78
240th day	1,029	69	130	53	58	77
241st day	1,091	67	142	51	80	77
242nd day	1,034	74	160	34	79	72
243rd day	1,073	73	130	67	65	78
244th day	1,049	79	149	41	69	77
245th day	1,096	82	138	44	97	80
246th day	1,085	85	149	48	87	86
247th day	1,096	88	140	39	80	88
248th day	1,105	91	146	47	77	88
249th day	1,104	94	150	38	71	89
250th day	1,113	98	149	35	62	91
251st day	1,114	101	147	48	78	92
Observation						
238th day	1,100	78	133	47	89	92
239th day	1,003	54	155	42	86	122
240th day	1,018	71	127	39	101	76
Cumulative difference	**42**	**30**	**28**	**40**	**66**	**81**
3-day accuracy (%)	**98.7**	**85.1**	**99.3**	**69.0**	**76.0**	**72.2**
241st day	1,411	96	186	49	124	123
242nd day	894	49	138	36	48	62
243rd day	1,035	67	92	67	61	72
244th day	918	59	164	43	45	76

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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.

Cumulative difference	**670**	**110**	**147**	**45**	**175**	**144**
7-day accuracy (%)	**90.9**	**76.7**	**85.2**	**85.9**	**68.3**	**76.8**
245th day	769	57	63	51	22	87
246th day	1,233	63	159	54	152	73
247th day	1,870	101	299	94	165	124
248th day	946	77	121	47	69	65
249th day	703	49	107	36	55	43
250th day	1,347	63	200	101	59	83
251st day	1,257	93	177	86	60	148
Cumulative difference	**2,858**	**272**	**541**	**217**	**452**	**329**
14-day accuracy (%)	**81.6**	**72.1**	**74.5**	**72.6**	**60.3**	**73.6**

Note: CI: Confidence interval.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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 .

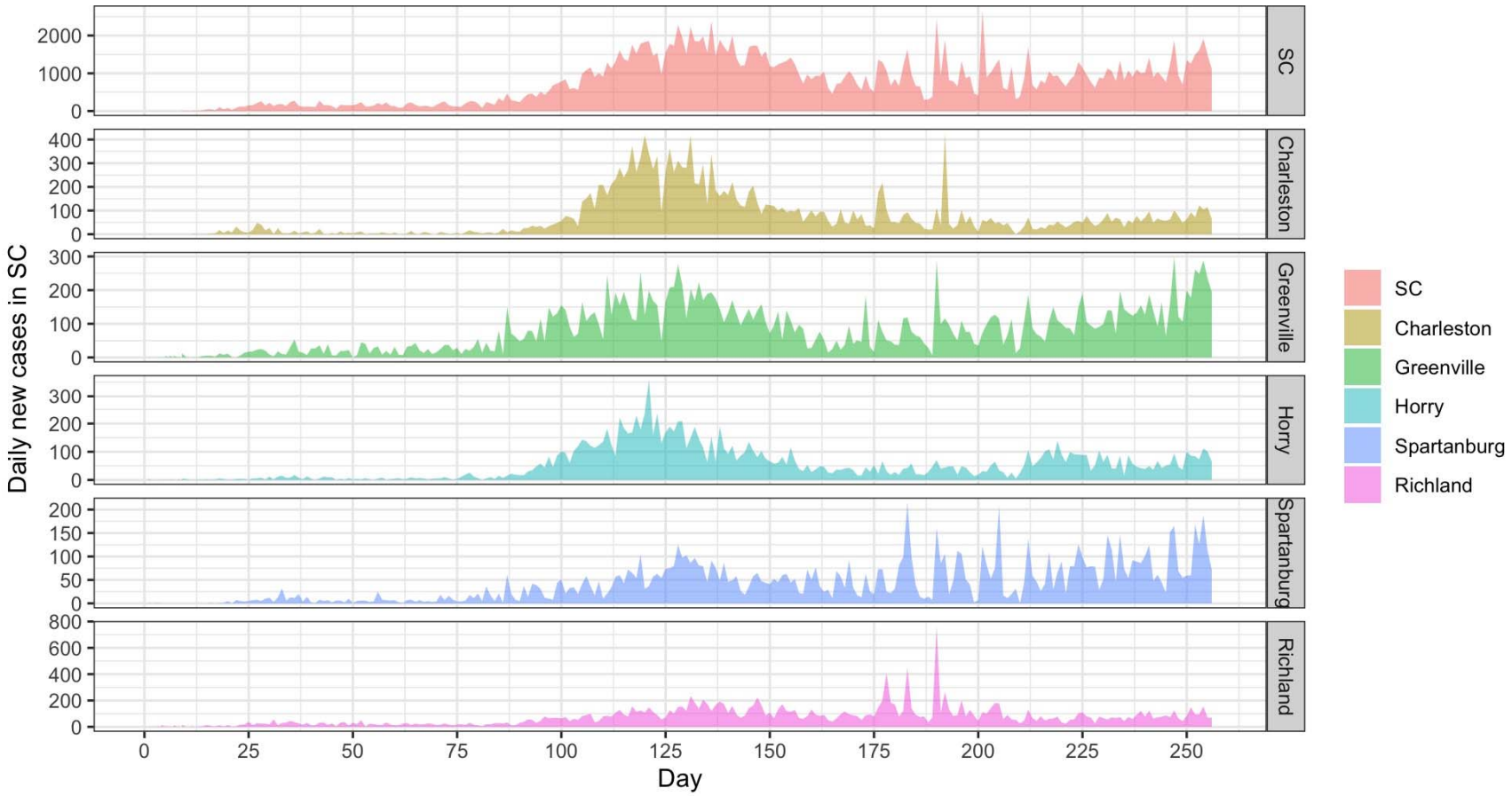

Figure 1. Daily COVID-19 new cases at both state- and county-levels in SC

Note: SC: South Carolina.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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.

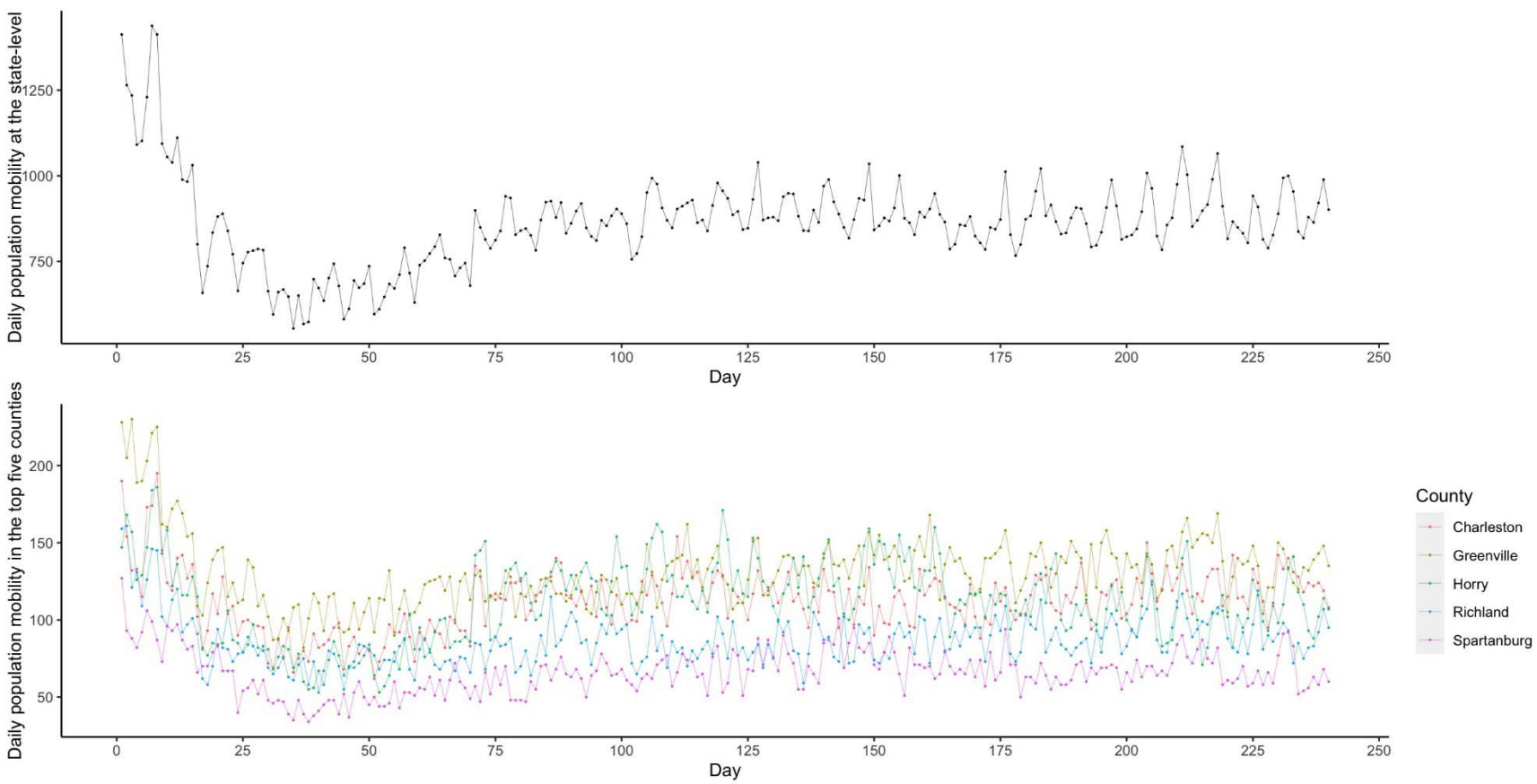

Figure 2. Daily population mobility at both state- and county-levels in SC

Note: SC: South Carolina.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.02.21249119; this version posted January 4, 2021. 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.