Response to Reviewers

15 September, 2021

Dear Editor:

I would like to thank you and the two anonymous reviewers for the attention paid to my submission entitled "Reproducibility of research during COVID-19: examining the case of population density and the basic reproductive rate from the perspective of spatial analysis" (GEAN-06-21-053).

The reviewers are generally positive about the contents of the paper, and there are no comments requiring additional analytical work. The most serious comment, from Reviewer 2, concerns the *tone* and *presentation* of the paper. You ask that I revise the paper to ensure that I am fair to a junior researcher (Reese Sy). This is a reasonable request. In this letter I respond to the thoughtful feedback I received and how I revised the paper in response.

I would like to assure the reviewers and you that my intention in using their work as an example was to constructively engage with Sy, White, and Nichols (SWN). In fact, even before submission to GA I had already taken steps to ensure that I was not acting like a complete jackass and a bully. I hope that my responses will persuade you of this.

Sincerely,

The author

Reviewer 1

The study aims to demonstrate the value of reproducible research, which involves making code and data publicly available. A strength of the manuscript is the expansion of relevant concerns with the SWN model, from the perspective of spatial analysis, and extensive and detailed justification for every model specification choice in the reanalysis. The inclusion of a RMarkdown file also guides the reader in understanding the models used in the reanalysis.

I have no major comments that can improve the study, as I believe the study is analytically sound and well-written. The only minor comment I have is the length of the background of the association of density and R0, which could be condensed further. Overall, the study is a good example of a manuscript that demonstrates the value of reproducible research, and would be an outstanding contribution to the scientific literature.

> Many thanks for your comments and positive assessment of the paper. I lightly edited the background section for conciseness, while trying to keep the paper as self-contained as possible.

Reviewer 2

The stated goal of this article is to promote reproducible research. This is done by reviewing another paper (SWN) that followed the standards advocated by this paper.

> Thank you for this comment. In fact, I reviewed at least 22 papers for this research. Of these, SWN was almost alone in being reproducible. Furthermore, of those that were reproducible, SWN was the only one that I thought I could engage with in an interesting way. The papers that I reviewed, along with notes about their reproducibility, are in the table below.

Paper	Reproducibility Notes
Ahmad et al. (2020)	Code not shared, data are provided but data pre-processing is not documented
Amadu et al. (2021)	Code not shared, data provenance is documented, data are not shared
Bhadra, Mukherjee, and Sarkar (2021)	Code and data not available; only data provenance is documented
Brandtner et al. (2021)	Code is available on demand, data are shared
Cruz et al. (2020)	Code is shared, sources of data are documented, data are not actually provided
Feng et al. (2020)	Code not shared, data provenance is documented, data are not shared
Feyman et al. (2020)	Code and data are provided
Fielding-Miller, Sundaram, and Brouwer (2020)	Code not shared, data provenance is documented, data are not actually provided
Hamidi, Ewing, and Sabouri (2020)	Code and data not available
Hamidi, Sabouri, and Ewing (2020)	Code and data not available
Inbaraj, George, and Chandrasingh (2021)	Code not shared, data are not shared
Lee et al. (2020)	Code not provided, data sources are documented but data are not shared (privacy & confidentiality)
Noury et al. (2021)	Code not shared, data are shared
Pequeno et al. (2020)	Raw data supplied in supplemental files, code not available
Roy and Ghosh (2020)	Code not shared, data are provided and data processing is documented
Skórka et al. (2020)	Code and data are shared but data processing is not documented,
Souris and Gonzalez (2020)	Code not provided, data sources are documented but data are not shared
Stephens, Chernyavskiy, and Bruns (2021)	Code and data are shared
Sy, White, and Nichols (2021)	Code and data are shared
Wang et al. (2021)	Code not shared, data are available
White and Hébert-Dufresne (2020)	Code and data are in repository
Wong and Li (2020)	Code not provided, data are shared but data pre-processing is not documented

My summary of the paper: 1) It is partly a new "COVID-19 paper" because it provides a new way to understand the population density hypothesis. 2) It is partly a "response paper" because it provides in-depth criticism of a recent paper. 3) It is partly an "advocacy paper" because it is clearly promoting open science. I will address these in order.

> Thank you for this insightful summary, which I think captures quite well the different missions of the paper.

1) The new models shown look sound and are an improvement over other models in this field. The density hypothesis is often discussed, so it is important to continuously improve our understanding of it.

Thank you for your positive assessment of the technical aspects of the paper.

2) Since this will not be published in PLOS, the original authors will not have the opportunity to respond. It does not feel fair. The authors of the GA paper might argue, "this is not a response paper." SWN seems to be a strange choice for a larger example. The first author is a graduate student. No one should be immune from scientific criticism (the selection bias criticism is valid), but choosing a generally good graduate-led paper (if I correctly understand the GA authors) feels wrong. If the GA authors chose the papers led by well-established authors such as Paez or Wong, I would not feel so queasy.

Thank you for the thoughtful response to this aspect of my paper. I would like to begin my response by noting that PLoS ONE desk-rejected my paper, which deprived me of an opportunity to engage in scientific discourse with SWN (apparently PLoS ONE does not think that science should be discursive). Even worse, by forcing me to take my paper elsewhere, PLoS ONE deprived SWN of the opportunity to respond to my paper in their journal of choice. This is, in my view, a very weird aspect of PLoS ONE's editorial policy. Secondly, I did not choose SWN for the example to pick on them, but rather because they were, as I noted above, almost the only example of reproducible research that I found in my research. I chose SWN partly of necessity, and partly as a homage, not because I particularly wished to criticize them.

More generally, I fully agree with you and the editor that criticism should be professional, in particular when directed at the work of junior researchers. Graduate students should be nurtured, not crushed. I tried to be respectful of Sy, White, and Nichols, which is why I carefully checked myself to avoid any and all arguments that could be construed as ad hominem. I tried to engage with the science, not with them personally. I also made an effort to recognize that I, just like everyone else, have strengths and limitations as researcher; I don't pretend to have the last work on density and the spread of COVID, and recognize that others may be better equipped to analyze these data more competently than me.

Incidentally, your comment brought to mind memories of Gordon and Richardson's (1989) response to Newman and Kenworthy's (1989) paper on Gasoline Consumption and Cities. In their response, Gordon and Richardson concluded by suggesting that "[p]erhaps Newman and Kenworthy would be well advised to seek out another planet, preferably unpopulated, where they can build their compact cities from scratch with solar-powered transit." I read this many years ago as a graduate student, and to the date I still find it shockingly unprofessional. I must admit that I lack the viciousness and ferocity of Gordon and Richardson. Rather, my aim is to be a good scientist, a decent professional, but if at all possible an even better human being. For this reason I took steps, from the moment I completed the first draft of the paper, to ensure that I was operating within the parameters of my moral compass:

1. I submitted the paper to PLoS ONE so that, if accepted, it would be a companion to the original SWN paper. As I noted above, PLoS ONE issued a desk rejection within two days: apparently their editorial policy precludes "commentary"; this leads me to believe that a response from SWN would also not have been acceptable in PLoS ONE. Only after being rejected there, did I try a different journal for the paper.

2. Immediately after I pushed the first version of the paper to my repository, thereby making it public, I communicated with Reese Sy, Laura White, and Brooke Nichols. This was my e-mail to them on May 21, 2021:.

Dear Ms. Sy, Dr. White, and Dr. Nichols, I would like to express my appreciation for your recent work on population density and the spread of COVID-19. I came across your paper while doing a search for a review assignment, and it caught my attention given some of my earlier work on COVID-19. Your paper is a relatively rare example of reproducible research, and it allowed me to relatively quickly replicate your results. I submitted a paper to PLoS ONE that builds on your analyses, and I thought that it would be just courteous to share it with you. Please

*see: *repository*. The README page needs some tidying up (for some reason a figure is not displayed). The source document is an R Markdown document which, along with the pdf you can find in this folder: R0-Density-Reanalysis. I also packaged the data for ease of distribution, and the workflow for data preparation is documented in this folder: raw-data-processing). If you have any questions or comments about this, I would be happy to hear from you.*

As you can see from the salutation, I was well-aware that Ms. Sy is a graduate student. This was the response from Prof. White to my email:

Dear Dr. X, Thanks for sharing this. I enjoyed reading your paper and find your results really interesting. I am happy to see you addressing the selection bias issues and considering nonlinearities in the association between density and R0. As you noted, these were issues that we chose not to address, but I am happy to see this picked up on. I agree with your premise the making our research reproducible only leads to better science. Hopefully we all get a bit better at this going forward. Best regards, Laura

This was followed by a response from Prof. Nichols:

Dear X, I would like to echo Laura's words. I really enjoyed reading your work- and this is the type of thing we should do more often in science! Cheers, Brooke

Last but not least, this was the response from Ms. Sy:

Dear X, I agree with Laura and Brooke - thank you for sharing this with us. I completely agree with what you wrote about reproducible research as vital to examine the same data from a variety of different perspectives. This strengthens scientific discourse! Have a wonderful day. Best, Reese

Why did I ask SWN to comment on a first version of the paper? First, I strongly felt it was the courteous and decent thing to do. And secondly, I think that science is better and more fun when it is done in a spirit of cooperation instead of antagonism. Their responses do not give a sense that they felt the paper was unfair to any of them as a team, let alone to Ms. Sy as an individual. Quite the opposite. I was very happy (not to mention relieved!) when their responses were as supportive as they were.

3. In addition, when submitting the paper to PLoS ONE (and after that to Science of the Total Environment, Landscape and Urban Planning, and Geographical Analysis), I always recommended that at least one of the authors of SWN be invited to review the paper.

It seems clear to me that PLoS ONE is not interested in responses to research, only in research. But does PLoS ONE's editorial policy really mean that SWN cannot respond? If fairness to SWN is important, the Editor of GA could easily invite them to respond to my paper in GA! Since PLoS ONE does not seem to care about this conversation, why shouldn't we take it to continue it elsewhere?

3) A well-documented paper containing data and code is easier to replicate and improve than a paper without these items. This seems obvious. The paper possibly indicates to journal editors that response papers are more likely to be submitted under open science. It possibly forms a cautionary tale that authors need to be extra diligent if they make their work easily replicable. However, using a case study method, as the GA paper does, seems to hurt the original paper/authors more than promote open science.

I agree with you that this seems obvious. The reality is that, as I noted above, the examples of replicable research are still few and far between, and in the whole literature of density and COVID, SWN was in the minority. I think my message is only partially directed to editors: for example, PLoS ONE already has a data and code policy - and yet, it is so seldom followed, let alone enforced, as to be next to useless. My message, instead, is mainly directed to researchers: we need to do much better than the very lowly expectations of journals. I already explained that I did not choose SWN to pick on them, but rather because they were one of the precious few papers that were reproducible. My communications with them do not give the impression that

they were hurt by this paper, but rather felt it was an encouraging product of open science. And, if in doubt, GA could always invite them to respond.

I want to think the GA paper comes from a place that respects the SWN paper, but the implementation of the GA paper does not give that impression. However, I do like the improved COVID-19 density model.

I am sorry that you were left with the impression that I was not respectful of SWN. I have no reason to think that they themselves felt disrespected. In my revisions I have been extra-careful to ensure that I engage with the material in a respectful and professional way.

Other suggestions

Schweinsberg et al. (2021) supports the view that different researchers reach different conclusions (https://doi.org/10.1016/j.obhdp.2021.02.003).

Thank you for suggesting this paper. I have included it in the discussion and references.

Pages 4,5. I think the authors have swapped the meanings of "large" and "small" scales.

Thank you for pointing this out. The text has been revised to avoid confusion.

The authors did not mention the time period studied.

SWN obtained publicly available daily COVID-19 cases among United States counties from the New York Times and for each county assumed that the exponential growth period began one week prior to the second daily increase in cases. We assumed that the period of exponential growth lasted approximately 18 days, as previous research have shown the COVID-19 exponential period to be around 20–24 days in New York City.

I assume that the data for the GA paper comes from the Johns Hopkins database. This data is sometimes revised, so this particular data set is an example of citations being insufficient to achieve exact replicability.

The data were obtained from the New York Times. Since the New York Times uses GitHub for their data repository, data versioning is available. See: https://github.com/nytimes/covid-19-data.

Please define the basic reproductive number for non-epidemiologist readers earlier in the paper.

Done.

The first part of the discussion section contains additional results. The discussion seems to start on page 18, line 41.

I reorganized the discussion following this comment.

These opinions on the SWN paper and its authors seem unnecessary. Page 7, line 30. "reasonable modeling choice" Page 7, line 40. "sensible controls" Page 8, lines 34,35. "whether a graduate student or a seasoned scientist" Page 18, lines 46-47. "required different technical skills than those available to SWN"

I appreciate your attention to these statements, which have been revised to make them more a matter of factly and less opinionated. Some statements were deleted altogether since they were unnecessary.

Page 8, line 19. It seems that the authors call themselves "expert"s. This seems unnecessary.

This was more of a general statement rather than an exlusive reference to me, and has been changed to "independent researcher" to avoid implying expertise.

Page 9, lines 49-52. There are many reasons to delete this footnote.

It is a fact that the paper was desk rejected by the journals for the reasons stated. The fact that one of them was PLoS ONE helps to explain why publishing this paper elsewhere is not unfair to SWN. I did delete the editorial comment to the effect that "[t]his does not inspire much confidence in the commitment of journals to reproducibility in research."

Citation typos: page 6, line 44; page 6, line 51; page 7, line 9; page 7, line 20; page 8, line 51

Errors corrected.

Typos Page 9, line 20. “RWN” should be “SWN” Page 11, line 21. “fourth” should be “fifth”

Typos corrected.

Mapping Figure 1: A narrower border on counties would better highlight the fill color. Figure 2: Rainbow colors are difficult to interpret. A divergent color band would be better. Figure 3: A single color could display counties with data. This will link the map to the binary variable in the model.

Changes in the cartography include narrower borders for the boundaries and the use of different color palettes. I used palettes from the ‘scico‘ package. It provides perceptually uniform and colourblind safe palettes.

I wish to thank you again for the care paid to reviewing this submission. I hope that you will find my responses to your concerns to be fair and sufficient.

References

Ahmad, Khansa, Sebhat Erqou, Nishant Shah, Umair Nazir, Alan R. Morrison, Gaurav Choudhary, and Wen-Chih Wu. 2020. “Association of Poor Housing Conditions with COVID-19 Incidence and Mortality Across US Counties.” Journal Article. *PLOS ONE* 15 (11): e0241327. https://doi.org/10.1371/journal.pone.0241327.

Amadu, Iddrisu, Bright Opoku Ahinkorah, Abdul-Rahaman Afitiri, Abdul-Aziz Seidu, Edward Kwabena Ameyaw, John Elvis Hagan, Eric Duku, and Simon Appah Aram. 2021. “Assessing Sub-Regional-Specific Strengths of Healthcare Systems Associated with COVID-19 Prevalence, Deaths and Recoveries in Africa.” Journal Article. *PLOS ONE* 16 (3): e0247274. https://doi.org/10.1371/journal.pone.0247274.

Bhadra, Arunava, Arindam Mukherjee, and Kabita Sarkar. 2021. “Impact of Population Density on Covid-19 Infected and Mortality Rate in India.” Journal Article. *Modeling Earth Systems and Environment* 7 (1): 623–29. https://doi.org/10.1007/s40808-020-00984-7.

Brandtner, Christof, Luís M. A. Bettencourt, Marc G. Berman, and Andrew J. Stier. 2021. “Creatures of the State? Metropolitan Counties Compensated for State Inaction in Initial u.s. Response to COVID-19 Pandemic.” Journal Article. *PLOS ONE* 16 (2): e0246249. https://doi.org/10.1371/journal.pone.0246249.

Cruz, Christian Joy Pattawi, Rachel Ganly, Zilin Li, and Stuart Gietel-Basten. 2020. “Exploring the Young Demographic Profile of COVID-19 Cases in Hong Kong: Evidence from Migration and Travel History Data.” Journal Article. *PLOS ONE* 15 (6): e0235306. https://doi.org/10.1371/journal.pone.0235306.

Feng, Yongjiu, Qingmei Li, Xiaohua Tong, Rong Wang, Shuting Zhai, Chen Gao, Zhenkun Lei, et al. 2020. “Spatiotemporal Spread Pattern of the COVID-19 Cases in China.” Journal Article. *PLOS ONE* 15 (12): e0244351. https://doi.org/10.1371/journal.pone.0244351.

Feyman, Yevgeniy, Jacob Bor, Julia Raifman, and Kevin N. Griffith. 2020. “Effectiveness of COVID-19 Shelter-in-Place Orders Varied by State.” Journal Article. *PLOS ONE* 15 (12): e0245008. https://doi.org/10.1371/journal.pone.0245008.

Fielding-Miller, Rebecca K., Maria E. Sundaram, and Kimberly Brouwer. 2020. “Social Determinants of COVID-19 Mortality at the County Level.” Journal Article. *PLOS ONE* 15 (10): e0240151. https://doi.org/10.1371/journal.pone.0240151.

Gordon, Peter, and Harry W Richardson. 1989. “Gasoline Consumption and Cities: A Reply.” *Journal of the American Planning Association* 55 (3): 342–46.

Hamidi, Shima, Reid Ewing, and Sadegh Sabouri. 2020. “Longitudinal Analyses of the Relationship Between Development Density and the COVID-19 Morbidity and Mortality Rates: Early Evidence from 1,165 Metropolitan Counties in the United States.” Journal Article. *Health & Place* 64: 102378. https://doi.org/10.1016/j.healthplace.2020.102378.

Hamidi, Shima, Sadegh Sabouri, and Reid Ewing. 2020. "Does Density Aggravate the COVID-19 Pandemic?" Journal Article. *Journal of the American Planning Association* 86 (4): 495–509. https://doi.org/10.1080/01944363.2020.1777891.

Inbaraj, Leeberk Raja, Carolin Elizabeth George, and Sindhulina Chandrasingh. 2021. "Seroprevalence of COVID-19 Infection in a Rural District of South India: A Population-Based Seroepidemiological Study." Journal Article. *PLOS ONE* 16 (3): e0249247. https://doi.org/10.1371/journal.pone.0249247.

Lee, Minha, Jun Zhao, Qianqian Sun, Yixuan Pan, Weiyi Zhou, Chenfeng Xiong, and Lei Zhang. 2020. "Human Mobility Trends During the Early Stage of the COVID-19 Pandemic in the United States." Journal Article. *PLOS ONE* 15 (11): e0241468. https://doi.org/10.1371/journal.pone.0241468.

Newman, Peter WG, and Jeffrey R Kenworthy. 1989. "Gasoline Consumption and Cities: A Comparison of US Cities with a Global Survey." *Journal of the American Planning Association* 55 (1): 24–37.

Noury, Abdul, Abel François, Olivier Gergaud, and Alexandre Garel. 2021. "How Does COVID-19 Affect Electoral Participation? Evidence from the French Municipal Elections." Journal Article. *PLOS ONE* 16 (2): e0247026. https://doi.org/10.1371/journal.pone.0247026.

Pequeno, Pedro, Bruna Mendel, Clarissa Rosa, Mariane Bosholn, Jorge Luiz Souza, Fabricio Baccaro, Reinaldo Barbosa, and William Magnusson. 2020. "Air Transportation, Population Density and Temperature Predict the Spread of COVID-19 in Brazil." Journal Article. *PeerJ* 8: e9322. https://doi.org/10.7717/peerj.9322.

Roy, Satyaki, and Preetam Ghosh. 2020. "Factors Affecting COVID-19 Infected and Death Rates Inform Lockdown-Related Policymaking." Journal Article. *PLOS ONE* 15 (10): e0241165. https://doi.org/10.1371/journal.pone.0241165.

Skórka, Piotr, Beata Grzywacz, Dawid Moroń, and Magdalena Lenda. 2020. "The Macroecology of the COVID-19 Pandemic in the Anthropocene." Journal Article. *PLOS ONE* 15 (7): e0236856. https://doi.org/10.1371/journal.pone.0236856.

Souris, Marc, and Jean-Paul Gonzalez. 2020. "COVID-19: Spatial Analysis of Hospital Case-Fatality Rate in France." Journal Article. *PLOS ONE* 15 (12): e0243606. https://doi.org/10.1371/journal.pone.0243606.

Stephens, Kenton E., Pavel Chernyavskiy, and Danielle R. Bruns. 2021. "Impact of Altitude on COVID-19 Infection and Death in the United States: A Modeling and Observational Study." Journal Article. *PLOS ONE* 16 (1): e0245055. https://doi.org/10.1371/journal.pone.0245055.

Sy, Karla Therese L., Laura F. White, and Brooke E. Nichols. 2021. "Population Density and Basic Reproductive Number of COVID-19 Across United States Counties." Journal Article. *PLOS ONE* 16 (4): e0249271. https://doi.org/10.1371/journal.pone.0249271.

Wang, Fangxiong, Ziqian Tan, Zaihui Yu, Siqi Yao, and Changfeng Guo. 2021. "Transmission and Control Pressure Analysis of the COVID-19 Epidemic Situation Using Multisource Spatio-Temporal Big Data." Journal Article. *PLOS ONE* 16 (3): e0249145. https://doi.org/10.1371/journal.pone.0249145.

White, Easton R., and Laurent Hébert-Dufresne. 2020. "State-Level Variation of Initial COVID-19 Dynamics in the United States." Journal Article. *PLOS ONE* 15 (10): e0240648. https://doi.org/10.1371/journal.pone.0240648.

Wong, David W. S., and Yun Li. 2020. "Spreading of COVID-19: Density Matters." Journal Article. *PLOS ONE* 15 (12): e0242398. https://doi.org/10.1371/journal.pone.0242398.