Response to Reviewers

Dear Editor:

Thank you for your decision letter of May 5th, concerning our jointly authored submission *A spatio-temporal analysis of the environmental correlates of COVID-19 incidence in Spain*.

We would like to express our gratitude to the three anonymous reviewers for their thoughtful comments, which we addressed in this revision.

In what follows, we describe the actions taken in response to the feedback received. We trust that you, and the reviewers, will find that we have addressed all recommendations in a thorough and complete manner.

We look forward to your opinion regarding the revisions of this paper.

Sincerely,

The authors

Reviewer 1

Narrative

This timely paper examines the association between environmental variables and COVID-19 incidence in Spain. It is well-written and contains a robust statistical analysis. The SUR method is very robust and appropriate to address the authors' research questions; while it is very well articulated in the text. The results and discussion are well presented and the exhibited decrease in Rho (spatial autocorrelation) over time is particularly very insightful.

I believe that the paper is suitable for publication after some minor revisions.

Thank you for your kind comments and positive assessment of the paper.

Please see my specific suggestions and comments below:

Introduction and Purpose

Page 1, Lines 42-43: "the SARS-CoV2 virus has threatened to overrun health systems the world over" – sentence is awkward and should be rephrased to "around the world" or "global health systems".

Page 1, Line 47: "arrested the spread" should be "mitigated the spread of the virus".

Page 1, Line 48: "helped to keep a bad situation from becoming even worse" is awkward and should be rephrased. Do you have any citations regarding the efficacy of social distancing?

These items were edited as per your suggestion, and we have included references as requested.

Page 2, Lines 24-28: You need to add the degree symbol ° in front of C. Please do so throughout the paper.

Done.

Page 3, Lines 32-42: This paragraph does not really belong in the introduction. You are discussing the key findings and results here, which belong in the discussion and conclusion. Instead, I suggest adding some hypotheses and research questions.

The introduction was reworked and most of this material was moved to the Background. This helped to make the introduction shorter and more engaging.

Page 3, Line 47: The link should be a footnote, rather than in the main body of the text.

Changed to a footnote.

Context and Data Methods

Page 3, Line 51: "January 31th" should be "January 31st, 2020".

Corrected.

Page 3, Lines 9-10: Is the relatively slow start due to testing lags? You may want to add something to clarify. In the United States, many people who died or were infected with the virus in January and February were considered influenza cases and not COVID-19 cases.

This is true. We have edited the text to note the relationship with testing capabilities.

Page 4, Lines 13-14: Was Madrid the original hotspot of COVID-19 because of testing resources and population?

The original hotspot was actually Álava in Pais Vasco; after reorganizing the sections, we hope that it is clearer that our hypothesis is that population and GDP are related to the early flare of the outbreak in Álava and Madrid (the first and second wealthiest Autonomous Communities in Spain, in GDP per capita).

Section 2.2: This is all great information, but it almost reads like a literature review and could be condensed and discussed in the introduction. I think the paper can be rearranged to Introduction and Context; then Data and Methods.

The sections have been reorganized following this suggestion. The introduction is now a little bit shorter and more in-depth discussion of the state of knowledge is now in the Background.

Page 6, Lines 29-30: US should be abbreviated as U.S. Please fix throughout paper.

Done.

Table 1: Is the sunshine variable missing from the table?

It was. We have corrected this oversight.

Methods

Page 8, Line 15. "varaibles" is misspelled. Should be "variables".

Fixed.

Is the model actually a space-time model, rather than spatial? I don't think it is explicitly stated.

Indeed, the model is space-time, with a lagged dependent variable for space, and temporal correlations in the residuals. We have emphasized this in the methods.

The methods section is well written and model structure is clearly articulated. Although spatial SUR is not the only approach to answer your research questions, it is a very robust statistical approach. I am curious how the results from the SUR approach compares to Geographic and Temporally Weighted Regression and Space-Time Conditional Autoregressive Models (MCMC or INLA).

We now comment on other potential modelling approaches in the conclusions.

Analysis

Section 4.1 is well-written.

Figure 1 needs to be improved. The class breaks in the legends are inconsistent, making it difficult to compare incidence rates. The projection seems to be a bit odd. Mainland Spain should be much larger, while the Canary Islands should be in separate insets, or remove altogether. Finally, the black line for the Madrid label masks many provinces.

We changed the scale for the colors to be consistent for all maps. After your comment about the connectivity of the islands, we decided to exclude them, and therefore the maps now show only the coterminous provinces. After this, the line for Madrid is less obtrusive, and overall the figures are easier to read.

Suggest caption for Figure 2 should be "Autonomous Communities in Spain (sorted from north to south)"

Changed.

I have the same suggestions for Figure 3 as I did for Figure 1. The resolution is poor, and the mainland should be much larger and the center of the map frames.

Done.

Figure 4: Are the correlations statistically significant? Please provide p-values.

The purpose of the daily correlation analysis in Figure 4 is to see how much day to day variability there is in the statistical associations with incidence. The p-values of these correlation coefficients are somewhat misleading because of the smaller samples used in their calculation. The information from this Figure helped us decide which variables to constrain in the SUR systems. At that point, the significance of the correlation is given by the p-values of the coefficients of the model. Instead of reporting the p-values for the daily correlations, we added a sentence to the effect that overall for the period under examination, the pairwise correlations between the independent variables and incidence are significant at $p < 0.05$, with the exception of the three sunshine variables. More details concerning the correlation tests can be found in the R markdown document in the chunk `correlation-tests`.

What is the justification of using Rook instead of Queen? Queen would strengthen the examination of spatial interaction and dependence. I think you are losing important information with a Rook weight matrix.

For this revision, we used the rook criterion, and the results do not change in any major ways.

Is it justifiable to assume that the islands are adjacent? I think that introduces a degree of uncertainty, especially after the lockdown order.

This is a good point. We repeated the analysis, and removing the three islands in the sample (Baleares, Las Palmas, and Tenerife) does not change the results in any major way. The results appear to be relatively robust to these changes in W. Conceptually, we agree that it makes more sense to exclude the islands given the lockdown implications. New footnotes comment on this, and the alternative analyses can be consulted in the source R markdown document.

Figure 5: Please increase the font size of the pooled-R2 labels.

Done.

Results and Discussion

I suggest that this section should just be "Discussion", while "Analysis" should be renamed to "Results".

Done.

Page 20, Line 32: "pandemia" is misspelled. Should be "pandemic".

Typo fixed.

Conclusion

Page 24, Lines 43-44: "Thirdly, all environmental data are based on a single station in a province". This is a bit concerning since a province is essentially similar to a state. I suppose the justification that you selected a single station based on population distribution is fair, but an interpolation could have improved the accuracy and precision of the environmental variables.

Many thanks for this comment. We were also curious about using a summary of several stations for the environmental variables. To test this, we interpolated the environmental variables using a k-points algorithm. When using these variables, instead of seeing improvements on accuracy and precision, the goodness-of-fit of the models declines, the intervals of confidence become wider, and consequently several coefficients lose significance. It seems to us that the interpolated variables introduce undesirable noise: on the one hand, they smooth the environmental variables over the province, but on the other hand they also include stations in places that are not necessarily informative (such as airports, up in Sierra Nevada, etc.) Having selected the stations manually seems to have kept this information "crisp". We recognize the possibility that a different interpolation algorithm might improve this, and recommend this as a topic for future research. A proper comparison of interpolation algorithms would involve experimenting with several methods to derive several environmental variables for a period of 51 days. For the purpose of this paper, we have recreated the data package to include both datasets (single station and average of stations). In the paper we discuss only the results of the first.

I also think that you should not finish your paper with limitations. I suggest moving the limitations to the discussion, then finish with a strong conclusion regarding the main findings and public health impacts.

Thank you for this suggestion, which we have adopted.

Reviewer 2

Comments to the Author This is a well written paper exploring environmental impacts on incidence of CV19 with spatial models, controlling for a number of socioeconomic variables. There is value in this paper, but a number of issues need to be considered:

Thank you for your positive comments and feedback

It would be helpful for the authors choose a name for the virus (e.g. SARS-COV2, COVID19) and use it consistently or explain if there are differences in when/how these names are use. Additionally, make sure presentation is consistent: e.g. sars-cov2 vs. sars-cov-2, COVID/Covid, etc.

SARS-CoV2 is the virus, and COVID-19 the disease it causes. We have clarified this, and revised the text to make consistent use of these terms.

pg 5, ln 13 - percentage*

Typo was corrected.

pg 6, ln 46 - I'm not sure I fully understand the weighting ... given the results indicate this approach doesn't produce a good fit, perhaps it can left out

We feel that this is, although a minor point, of interest for modellers and would prefer to retain it. We have revised the discussion of the weights and hopefully they are more clear now.

pg 8 ln 15 "variable" misspelled, ln 16 "restrictions" misspelled.

Typos were corrected

use of mean temp/humidity - does daily variation matter? For example do some regions have lows near 2 and highs near 20 c, where others have lows near 5 and highs near 6?

*Indeed, which is why there are experiments with the moving averages of the environmental variables: 1) to smooth the daily variation in accordance to the incubation period of the disease. The narrowest window is 8 days, and the two 11-day windows weight the daily variations in different ways: uniformly for the lagw variables, and giving greater weight to days closer to the median incubation period of 4-5 days in the lag*11*w variables.*

pg 10 ln 50 - sentence grammar

could W be defined using commuter flow to gain a better sense of interaction? Are there past examples in the literature that define W for provinces across Spain? Perhaps there's more interaction between Madrid and Barcelona, etc.

In principle, it is possible, but unfortunately commuter flow data are not readily available. The Ministry of Transportation is working on flow data for Spain based on mobile data, to be released at a future (and yet undefined) date. In our view, once in lockdown, hierarchical flows are perhaps less relevant than cross-boundary flows, and even those should become relatively less important as the lockdown persists: this is, we believe, what the autocorrelation coefficient is telling us. See this paragraph at the end of the discussion:

> *Finally, we defined neighborhoods based on adjacency. It would be interesting to compare other connectivity criteria, for instance based on domestic transportation instrastructure and services. We flag this as a matter for future research.*

pg 15 ln 22 - 'transit' systems

Typo corrected.

given multiple equations would it be appropriate to account for multiple testing and adjust accordingly?

Not with SUR, because the multiple equations are estimated jointly, and the calculations account for the size of the sample.

pg 18 ln 32-33, do the authors mean 'due to *violations* of the restrictions of movement'?

> *This sentence was revised to conform to the slight changes in the results after the models were reestimated without the islands.*

pg 19, not sure the anecdote is appropriate - suggest this be deleted

> *Deleted.*

pg 21, ln 52 - can this recommendation be backed up with further evidence about lockdown effectiveness?

> *We think that this is a reference to pg 20, ln 52. This paragraph was too wordy, and we have edited it for clarity. Instead of effectiveness of the lockdown, we tie this discussion to risk compensation, which seems to a clearer case.*

pg 22 ln 36 - I'm not sure I follow this statement ("Without a lockdown. . . "). Also, please remove "!". I've tried to work through this logic, but can't quite seem to figure out how you have come to this conclusion. Wouldn't cases generated elsewhere also be affected by temperatures in those locations? Would this not make the process more complex/less clear?

> *You are right that the process becomes more complicated, and explaining it is also not simple. We were not as clear as we had hoped, but made an effort with this new paragraph:*
>
> > *What do these effects mean? Under a situation of lockdown, inter-regional contagion is reduced, as expected, and the total effects of the variables tend towards their direct effects. In those first few days covered by our analysis the total effect of all variables is greater due to the spatial contagion effect. This also makes analysis and intervention more complex: the contagion effect essentially acts like a multiplier, whereby developments in each province spill over to their neighbors. Once the contagion effect has been tamed, each province can be "treated" independently from its neighbors.*

please double check spelling/grammar throughout.

> *We have proof-read the paper and hopefully fixed all remaining spelling/grammar mistakes.*

Reviewer 3

Comments to the Author

The paper is well written and very timely. I truly enjoyed reading it and I think it should be granted publication only with some very minor revisions. Although I understand and I use econometric and statistical models, I am not an expert on the technical part and therefore I will not comment on that specifically. My comments are more "general".

Many thanks for your positive assessment and suggestions for revisions.

There is a clear need of looking at environmental factors such as those you are considering. However, there has been a huge debate on pollution as a possible environmental factor affecting the spread of the disease. I would suggest that you clarify somewhere that you look only at weather-related factor and not pollution and explain why (bad data I assume?).

Indeed, working with meterological stations was already challenging, and we are not aware of pollution data at the scale of the analysis. We now discuss the role of pollution as an environmental factor in need of further attention.

Working on the topic myself I often wonder how the results would change if the disease had hit somewhere else first? Clearly that would have affected the spread... Could this be a problem in your case?

This is an interesting question and a problem insofar as we lack a counterfactual, This is where the use of simulation techniques are useful, and the work of Muller et al. with MATSim is relevant in this respect; see:

> *Muller, S. A., Balmer, M., Neumann, A., & Nagel, K. (2020). Mobility traces and spreading of COVID-19. medRxiv, 2020.2003.2027.20045302. doi:10.1101/2020.03.27.20045302*

Very minor:

I would write NUTS2, NUTS3... I have almost never seen NUTII and NUTIII...

We made this change.

Page 6: when you say you use a rather large geographical unit of analysis. I get that, to be quick, you have to use whatever data are available in the short run. However, I do not get what you mean by 'the analysis must therefore be considered "ecological" '. Either you explain it better or you simply remove the sentence acknowledging your results are preliminary, but could be easily applied to a finer scale once, and if, better data become available.

We agree, and removed the sentence to avoid overcomplicating the discussion. Earlier on in the paper we already had explained that this is a population health approach.

Here and there I would remove some "however".

Done.