Review of Barré et al., Estimating lockdown induced European NO2 changes, submitted to ACP, 2020

Anonymous Referee #3

We would like to thank the reviewer for their comments that helped to improve the paper's quality. Please read our answers in italic fonts below.

The authors quantify the reduction in NO2 levels over Europe that resulted from the decline in emitting activity during the Spring 2020 lockdowns, themselves resulting from government responses to the COVID19 pandemic. They do this by using satellite NO2 column data, surface measurements and model simulations, while also demonstrating the importance of accounting for year-to-year variability in weather conditions that would otherwise influence the NO2 signal on top of any emission changes. They conclude with a brief synthesis and comparison of the different methods, showing the estimated NO2 reductions for large European urban areas.
Overall, this is an interesting application and demonstration of state-of-the-art measurement, analysis and modelling tools to a timely topic. My questions, comments and concerns about the science are minor and outlined below. However, my main issue is more around presentation and structure. To me, the manuscript currently reads like a series of disconnected stories that are only weakly united at the end, with a rather thin discussion and summary. I expand on this comment and make some suggestions below, but I think addressing it would be a sizeable task (hence suggesting "major revisions"). I would urge the authors to consider this point since I think it would ultimately leave them with a much more readable (and citable!) piece of research.

Major comments – Structure, presentation and focus

A key selling point of this research is the multiple approaches that the authors have applied, yet this is not really front and centre to the reader, except in the Abstract. I would suggest reflecting this contribution in the title (e.g., "Lockdown-induced NO2 reductions in Europe estimated from satellites, surface stations and air quality models"??) as well as in the first paragraph of the introduction.

We have changed the title.

Currently, the introduction is rather focussed on reporting individual lockdown studies (which can probably be synthesised more) and discussing actual and potential misapplications of TropOMI data. There is not much information or discussion on what can be gleaned from surface observations and models, let alone why an approach with all three might be novel and more robust.

The information about surface observations and models were already there but it was rather disorganised. We have now re-arranged and re-written the introduction. Synthetised some paragraphs (i.e., satellite) and adding more information on others (e.g., surface observations). It is now more in accordance with the rest of the paper.

I would suggest that the authors then consider the presentation of the methods and results. One way would be to describe the measurement and model details and analysis approaches in one section, followed by a results section that begins with the current Figure 8 (which is the main take home message). Subsequent sections could then explore the differences between the approaches (e.g., combining some of the other maps?) as well as highlighting what are more well-known or secondary aspects, such as the need to consider meteorological normalisation. A final discussion section could consider the uncertainties in each approach in more detail.

We understand the reviewers' point of view, but we cannot operate such changes at this stage of the paper's submission. This will completely break the flow of explanations and will require re-writing almost entirely the paper on top of the already major changes requested in this review. We however have reorganised the three section on satellite observation, surface observations and air quality models with methods and results subsections. We now split the former "Summary and discussion" section into two: one section about the comparisons of the three approaches (former figure 8 now figure 9 discussion) and one general and final conclusion.

Even if the above suggestion is not followed, the interpretation and discussion around the current Figure 8 certainly needs more attention and discussion. The submitted manuscript is rather scant on detail in comparing the outcome of the different approaches, how independent they are (e.g., are the model or surface measurements used in the satellite retrieval method or validation?), or how they may be used to provide some validation of each other or increase the overall confidence (e.g., as per IPCC type language like "very likely" etc. when there are several lines of evidence).

We have now extended the discussions for each of the individual results and the comparisons around the final figure (now figure 9). An entire section is now devoted to the comparisons of the different methods and provide extended explanation of why they could differ or not. The section expands over two pages.
Note that the NO2 forecasts used in the training set and predictors is business as usual information as no assimilation performed therefore making it independent form the model estimates provided in section 4. We clarified the text accordingly.

Finally, related to the presentation, I would encourage the authors to revisit the readability/flow and grammar of the manuscript. For the former, I often found that paragraphs did not nicely follow on from one another, reading instead like disparate bullet points. Additionally, many longer paragraphs could be broken down into more readable chunks. Regarding grammar, to my mind there are several examples of curious word choice and word order. I accept that this may just be my preference coming through, but I would encourage the authors to proofread any resubmission.

The flow and grammar have been revised.

Specific comments

Introduction: Somewhere, I would find space to acknowledge earlier work on weather normalization of AQ observations. One suggestion (but not limited to this!) is David Carslaw, whose blogs on the impact of lockdowns on NO2 refer to his published work (e.g., see: https://ee.ricardo.com/news/blog-update-on-covid-19-and-changes-in-air-pollution)

The references were added in the introduction with few additional sentences. Note that the (published) work of David Carslaw is acknowledged through the references to Grange et al. (2018, 2019) (Carslaw being the second author of these studies).

L108: Here or elsewhere (methods or results?) it would be good to be explicit about the otherwise implicit assumptions about BAU – i.e., that you're assuming emitting activity would be similar to previous years (for the weather normalized techniques), or as per the projected 2020 emissions data (for the simulations…although are these indeed be the same as previous years?).

This point was answered in the other referee comments. We now refer to the EEA 2020a report in the satellite and surface observation sections to acknowledge the potential contribution of the NOx emission trend in the BAU predictions.

Reference:
EEA: European Union emission inventory report 1990-2018 under the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP), EEA Report No 05/2020, 2020a.

Table 1: There is really no point in this Table, whose information could just be included in the text.

We have removed the table

L143: How was the PBL height calculated and/or where did it come from?

The PBL height comes from the IFS calculations. We clarified the text accordingly and provide the reference to the IFS documentation for calculation details.

L168: What are the criteria for "urban areas"? I am curious because it seems that the definition must include some of the surrounding metro areas (e.g., Southend, Essex, UK "proper" has a population < 200k), yet some major areas are excluded (e.g., the South Hampshire metro area in the UK has a population >1M).

We have used the free data base coming from (https://simplemaps.com/data/world-cities). Following the website description and Q&A:

*" **What counts as a city/town?***

Any populated place in the world as determined by U.S. government agencies. Neighborhoods within listed cities are not included.

Where does your data come from?

Cities for all non-U.S. countries comes from the National Geospatial-Intelligence Agency. US cities data comes from the U.S. Census Bureau and the U.S. Geological Survey. The basic database and population data comes from Natural Earth Data. Population density data comes from the Center for International Earth Science Information Network at Columbia University in partnership with NASA's Socioeconomic Data and Applications Center. We've made a concerted effort to source our data from public domain and permissively-licensed sources that will not restrict the rights of our customers."

So, it is possible that cities boundary definitions are not reflecting the overall population in the urban area. We choose to use the term urban area as for many cases a pixel of 10 km x 10 km is not representative of a given city but also of its surroundings. We now provide the reference site in the text for the database that has been used and clarified the sentence.

Section 2.3: A figure showing the performance of the ML technique would be helpful. E.g., time series for a particular location, showing its performance for the training and test data sets?

We have added the suggested additional figure providing times series for 2019 and 2020 for Madrid. We also have amended the text to include the figure description and discussion.

L264: I'm not sure what "perform better" means here.

The sentence has been clarified.

L284: Provide citation for "policy measures across Europe"

The reference is now included in the text.

L300: I'm confused by this sentence – is it related to comparing the surface observations against the satellite data? Please clarify.

Yes, it is. We have clarified the text.

Section 4: However this section gets worked into a revised manuscript, more information is needed, even if it just points to other studies. I would encourage separate sections on the modeling set up and the emissions, as well as how the activity data (etc) were used. Also, how does the model output compare to TropOMI?

From the other reviewer comments, we have clarified the emission scaling procedure with the necessary references. We also provided the country dependent reductions factors in the annex section.

The study focuses on the assessment of the relative reductions seen by various methods but not an evaluation and validation of the CAMS regional models using TROPOMI. Such activity of validation using surface data and satellite data is routinely performed quarterly on the operational production and delivered publicly. We have added clarifications and the link to the validation reports in the text.
A fully consistent comparison between model and TROPOMI estimates would require a satellite observation operator. Unfortunately, such operator is not officially available yet for the regional models. This point is discussed in the summary and discussion section of the first submitted version of the paper. This is now in the last paragraph of section 5.

L321: Do the 11 models need to be named? Perhaps just point to the citation?

The model names have been removed.

Section 5: As noted above, this section needs more discussion on Figure 8 and the difference between the results. Some additional specific comments follow:

L384: Explain/justify why it is "crucial…for air quality policy".

The text has been clarified.

L389: What is meant by "relevance" here? I would argue is more of convenience, since the plot will be missing out a large majority of Europe's total population!

We have removed the word relevance. It is however in the most densely populated areas that the pollutions changes are expected to be seen at a 10 km scale. Business as usual pollution levels have the tendency to be higher in large urban areas than small urban areas.

L412: Explain/describe “background footprint” and clarify the “representativeness” issues for more general readers.

We have clarified the text accordingly.

Figure 8: This is a great figure, but it is rather busy with the lines which prevents any clear message emerging from a glance. Hard to know what to suggest (put hourly station and model data in an appendix figure, so it’s comparing like with like?), but I would at least encourage the authors to make the zero line more obvious.

We have updated the figure with the zero line wider.

Figure 8: A separate issue to the above, the spread (IQR etc) needs a clearer definition. Is it a spatial and temporal spread?

In the case of figure 8 it is the temporal spread. We have clarified the caption.

Technical corrections (a full proofread is recommended)

L45: “…part of the nitrogen oxides…” – nitrogen oxides include a lot more than NO and NO2 (e.g., N2O, N2O5 etc). To me this is also an example of curious wording. Suggest “Nitrogen dioxide (NO2; together with NO, a constituent of NOx) is a very well-established…”

Fixed.

L72: “The storm Ciara...” -> “Storm Ciara…” (and in other cases too). I’m no expert but seems like the preferred orthography is to capitalize the “S” in Storm when referring to a named one.

Fixed

L140: “A number of named extratropical cyclones (Storms Ciara, …)”

Fixed

L269: This is an example long paragraph that could be broken into shorter ones.

The paragraph has been broken down into three.

L326: Spell out TNO

Done

L399: This sentence doesn’t make sense

We have now clarified the sentence

L429: This is not a stand-alone sentence (belongs as a clause of previous sentence).

Fixed