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

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 NO_2 reductions in Europe estimated from satellites, surface stations and air quality models”??) as well as in the first paragraph of the introduction. 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.

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.

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).

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.

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)

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?).

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

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

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).

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?

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

L284: Provide citation for "policy measures across Europe"

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

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?

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

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".

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!

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

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.

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

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…”

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.

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

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

L326: Spell out TNO

L399: This sentence doesn’t make sense

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