Interactive comment on "Estimating lockdown induced European NO_2 changes" by Jérôme Barré et al.

Anonymous Referee #1

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.

Received and published: 6 November 2020

Barré et al. (2020) describes estimates of the magnitude of NO2 reduction that can be attributed to COVID-19 lockdown measures in Europe in 2020. The paper is highly relevant given the continuation of COVID-19 and associated restrictions, interesting and well written. The argument for needing meteorological normalisation is, I think, sound, but the introduction would benefit from some more discussion why this is so, focussing on the role of different meteorological variables on NO2 concentration. The methodology employed to achieve the meteorological normalisation is sound for surface estimates. I think it would make much more sense to use surface mixing ratio estimates from TROPOMI rather than column values, given the fact that (a) the paper focuses on urban air quality where exposure by large populations is at the surface level and (b) surface temperature and wind are key predictor variables. The CAMS modelling section is well presented, but so far lacks punch. It would substantially benefit the paper to discuss the emission reductions further as this would demonstrate not just that the NO2 did reduce, but why. Discussing the sectors responsible for emission reduction and therefore NO2 concentration reduction, and whether this is consistent across Europe, would provide the chemistry focus required for Atmospheric Chemistry and Physics.
Overall, I would recommend publication once these comments are addressed. Expanded comments on the above discussion are made below, followed by minor technical comments.

Unfortunately, the TROPOMI instrument as a nadir viewing geometry does not provide surface mixing ratio retrievals but only tropospheric columns retrievals. Due to the nature of the measurements (reflected radiation from earth surface in UV-Visible part of the spectrum in the case of TROPOMI) is it very well established in the scientific community that such type of remote sensing measurements cannot provide retrieved surface concentration solely but only vertically integrated retrieved content with provided weighting function called the averaging kernel.

Regarding the emission suggestion, we state in the introduction: "These lockdowns drastically reduced traffic and also activity levels in most industries (Guevara et al., 2020; Le Quéré et al., 2020). These sectors represent a large share of NO_x emissions (51% according to EEA 2020a)." The references mentioned in the text already discuss extensively the sectors for emission reductions and also show the large share of NO_x emission in the impacted sectors. This is recalled in the modelling section where the same references are mentioned and also adding Colette et al. (2020) that refers on how this was implemented in the modelling procedure. We do not think then adding extensive discussion from already published results in the literature will be an added value to the article, but we now add more description on the emission reduction estimates, recalling the information given by Guevara et al., 2020 and Colette et al., 2020 in the modelling section and adding a table showing the emission inventory reduction factors used per country and per sector.

Expanded Comments:

1. The introduction needs to provide the reader with some more context on why studying NO2 during the lockdown is important. I think a paragraph could be included to this end, outlining the unique nature of this real world emissions-reduction experiment and its potential to help us understand potential broadscale impacts of future pollution reduction measures.

We have complemented the introduction accordingly.

2. The introduction highlights that considering meteorology is important for NO2 prediction – I agree, but the introduction would benefit from a little more context on why this is the case. Just a couple of sentences are necessary on, e.g., boundary layer heights and NO:NO2 temperature dependent ratios that make this point clear.

We have detailed the introduction accordingly.

3. Figure 2: Temperature is a really important factor for the NOx partitioning and warm temperature anomalies are highlighted in the text for early 2020. Therefore, I think Fig. 2 would benefit from showing the distribution of temperature as well as PBLH and wind speed.

We have added the temperature distributions in the figure.

4. Line 180-182: Please consider this sentence: "This illustrates that such "beforeduring" type of satellite comparisons is misleading and unfit for assessing the effects of COVID-19 lockdown because it is very sensitive to seasonal variations of weather regimes and emissions." What you have shown, to this point in the paper, is that different baselines provide different results for the 'lockdown NO2 change' and that the weather was different. You have not strictly proven, yet, the link between the two. At this point in the paper, you either need to prove the causal link with data, provide references for the statement 'because it is very sensitive. . ." or mute the sentence to something like "This illustrates that such 'before-during' satellite comparisons clearly provide very different results as to the effect of lockdown on European NO2. This led us to investigate weather considering meteorology may provide a more consistent picture".

We have clarified the sentence.

5. Why did you use TROPOMI column NO2 rather than surface mixing ratios? The mixing ratio can be determined from the NO2 tropospheric vertical profile. This would seem to me more relevant for urban air quality than tropospheric column values, and would provide consistency with surface observations. In addition, this would make your surface meteorological predictor values much more relevant, at the moment (effectively) surface temperature and wind are being used to predict the whole column.

Please see the response above. In addition, the surface mixing ratio cannot be straightforwardly determined from TROPOMI measurement. We are not sure what the reviewer means by "The mixing ratio can be determined from the NO2 tropospheric vertical profile."? Then which NO2 tropospheric vertical profile? From measurements (in situ I presume), from models? Column based quantities are retrieved from the TROPOMI measurements not profiles.

6. Why were NO2 modelled concentrations in the predictor variables? I think to demonstrate the importance of meteorological normalisation, you should show that the GBM gives good prediction independent of NO2 concentration estimates.

Including the business-as-usual NO_2 surface concentrations from the CAMS regional forecasts will help the GB model to simulate more accurately NO_2 columns where the meteorological information solely is not enough as the training set is very small. Using the business-as-usual CAMS NO_2 forecasts or not in the predictors do not interfere with the validity of the meteorological normalisation method. We used business as usual predictors only and whether they are meteorological information, chemical forecast information or else do not change the methodology to predict NO_2 columns to compare with the factual or real NO_2 columns. We now clarify that the NO_2 forecast information used is business as usual information.

7. Table 3: Please clarify whether the average changes in this table are means or medians, ideally consistent across all comparisons.

We have clarified the caption.

8. Line 283-284: "Using the last three years is long enough to capture weather variability at each site, but not too long with regards to long-term reduction of NO2 happening as a result of policy measures across Europe" – perhaps the authors could clarify this rather vague statement by indicating expected (or citing known) NO2 trends across Europe as a result of policy measures. In the previous paragraph, the authors note that their method underestimates NO2 in the pre-lockdown period by 8 %, could this be partly due to decreasing NO2 trends driven by policy or emissions?

Since GBM models are trained based on past data, strong trends between the training period (2017-2019) and the period of interest (2020) may indeed affect the reliability of the ML-based predictions. In order to limit this issue, we introduced the date index feature. If no substantial trend exists since 2017, the GBM models will not use this feature and will benefit from the entire (2017-2019) training dataset for learning the influence of the meteorology on NO2 surface mixing ratios. Conversely, if a strong trend is present, the GBM models will give more importance to this date index feature during their training, which in practise will allow predicting NO2 mixing ratios (in 2020) in the range of values observed during the last part of the training dataset. This is discussed in more detail in Petetin et al. (2020). We clarified the statement and also added a reference for trend in the text.

9. I think the strength of the paper would be improved by discussing in more detail what led to the NO2 decrease – the modelling section seems to offer this opportunity. Was it reduction in industry, aviation, road transport, all of the above equally or something else that was primarily responsible for the NO2 change, and was this consistent across Europe?

Please see the response to the main comment above. The modelling section has been complemented accordingly.

10. It would be interesting to discuss how this kind of weather-normalisation 'business as usual' prediction could be implemented for air quality forecasting, in addition to event/emission change analysis.

To reinforce the usefulness of the method on observation-based emission change analysis, we added the following sentence in the conclusion: "Beyond the quantification of the impact of COVID-19-related restrictions on pollutant concentrations, the weather-normalization methodology used in this study is of general interest for assessing the impact of any type of emission changes (e.g. regulation) on air quality (Grange et al., 2018, 2019)."

It is important to note that, state of the art air quality forecasting is using models that do include the meteorological variability intrinsically in their predictions as such models represent and predict explicitly the evolution of the pollutants in the atmosphere. So, no need of weather normalisation here. CAMS operationally run such forecast models daily at 10km resolution over Europe. In addition, data assimilation is performed within those models to include the information from observations in an optimal way. With the current methodology presented here BAU parameters coming from air quality and numerical weather prediction models are needed. So, using such weather normalized technique to forecast air quality will be redundant. One of the advantages and currently looked at application of machine learning techniques on air quality forecasting focuses more on designing downscaling methods to predict phenomena beyond the model's resolution, but this is rather out of the scope of the paper.

Technical comments:

1. Be consistent with subscripting of "x" in NOx

Fixed

2. Line 72: please revise the first sentence of the paragraph, 'very changing' is poor grammar – perhaps 'highly variable' would be better

We have corrected the sentence

3. I'm not sure Table 1 is necessary, it is so small and the information is clearly stated in the text anyway.

We have removed table 1

4. Line 123: strange font difference in 0.1x0.1o

The fonts have been changed

5. Figure 3 (and subsequent similar figures): Given that you consider urban areas down to 0.5 million inhabitants, I recommend adding some more circles to your population circle-size legend (maybe 0.5 m, 1 m, 2 m, 4 m, 8 m)

We have adjusted the plots

6. Figure 3: (and subsequent similar figures) Subscripts please on the NO2 in the colour bar label

We have adjusted the plots

7. Figure 3 (and subsequent similar figures): please just clarify, the%change is relative to each baseline scenario? – I suggest including this clarification in the figure caption.

We have clarified the captions.

8. Table 2/line 245: are the outliers included in the statistics presented in table 2? If they're included, might they explain the significant RMSE?

Yes, a strong RMSE in the performance results will most likely generate outlier in the NO2 changes estimates. We clarify the text accordingly.

9. Lines 263-265: Perhaps my personal choice, but I would write "X % reduction" not the double negative "-X% reduction". This would also be consistent with the way it is written in the paragraph starting Line 347.

Corrected

10. Line 307: should be ". . .measurements do not directly translate to. . ."

Corrected

11. Line 323: model rather than models

Corrected

12. Line 338: I'm curious if there is a metric which could help determine the stringency of lockdown measures in different countries? At the moment, knowledge of the scale of lockdowns and COVID-19 consequences are fresh in our minds, but people may not have a feel for that reading this in the future. I think some discussion of what constitutes a 'more stringent' vs 'less stringent' lockdown is warranted.

There is such a metric developed by Oxford university. It accounts for various parameters to define a stringency index. We now mention this index with the reference in the text.