Review_YWU30306_Kirsten de Nooijer

This review was prepared as part of graduate program course work at Wageningen University, and has been produced under supervision of Prof Wouter Peters. The review has been posted because of its good quality, and likely usefulness to the authors and editor. This review was not solicited by the journal.

Review of "*Global Impact of COVID-19 Restrictions on the Surface Concentrations of Nitrogen Dioxide and Ozone*" by Keller et al. (2020)

The paper by Keller et al. (2020) aims to create a 'business as usual' model output to compare with the changes in nitrogen dioxide (NO_2) and ozone (O_3) concentrations observed during the 2020 lockdown situation. To create this 'business as usual' output (how emissions in the first half of 2020 would have been without lockdown) the authors adapt the NASA GEOS-CF model to include seasonal variability of air pollutants. The result is a biased-corrected model (BCM) that also includes meteorological and compositional information. Publicly available data on NO_2 and O_3 is used from 5,756 observation sites, from which most are in Europe, North America and China. A machine learning algorithm is used to predict the time-varying bias at each observation site. Reductions in NO_2 range from 60% in more severely affected cities as Wuhan, to little difference in less affected cities as Rio de Janeiro. They estimate a reduction of NO_x ($NO+NO_2$) of 2.9 TgN during the first half of 2020, equivalent to 5.1% of the annual anthropogenic total. Following changes in O_3 concentrations is more difficult due to competing influences of non-linear atmospheric chemistry. The analysis does indicate a flattening of the O_3 diurnal cycle with O_3 increasing during the night and decreasing during the day. They do expect that the importance of photochemical production will increase in the Norther Hemisphere, resulting in an overall decrease in surface O_3, if NO_x emissions continue to decrease as a result of COVID-19 restrictions.

The main reason for writing this paper is to report new knowledge on a very recent (and still ongoing) event that is of global importance. Keller et al. correctly identify the knowledge-gap for a need of quantification of the reduction in global emissions, fitting the scope of the journal. Their research contributes to the reporting and is based on suitable methods for quantification of reductions. Therefore, this overall well written research should be published. However, not in this current state. My major comment is on the absence of 3 important things that need to be added before publication: (1) statistics, (2) calculation steps and (3) definition of 'lockdown'. The 3 major comments are followed by a list of minor comments.

<u>Major comments:</u>

Major comment (1): Statistics need to be included. The aim of the paper is to quantify, uncertainties should be quantified as well. The reported numbers are easily disregarded without the proper statistics (e.g. p-values, t-tests or z-tests) and uncertainty ranges. This problem is present in figures 4 and 6 (but applies for all given emission changes in the manuscript). Here the difference between the BCM prediction and observations is shown, but without noting whether this difference is significant (or perhaps falls within the uncertainty range of the BCM prediction).

Example 1, line 145: "*For Wuhan, we find a reduction in NO_2 of **60%** relative to the expected BCM value for February and March 2020, and similar decreases are found over Milan (**60%**) and New York (**45%**) starting in mid-March and lasting through April (Fig. 4; Tables A1-A3).*"

How certain are these numbers? Is it between 62% and 58%? Or between 70% and 50%? I urge the authors to please quantify the uncertainty of these numbers by providing uncertainty ranges or mentioning of significance. This could be implemented similar to Le Quéré et al. (2020), here reductions in emissions are provided by stating the range (representing ±1σ) instead of a single number.

Example 2, line 228: "*Compared to the BCM model, there has been an increase in the concentration of night time O_3 (midnight-5.00 local time, Fig. 8a) by* ***1 part per billion by volume*** *(ppbv = nmol mol-1) compared to the BCM, whereas O_x shows a decrease of* ***1 ppbv*** *(Fig. 8b).*"

Is this reported 1 ppbv difference significant? I highly suggest you to report whether the modelled change is significantly different from the observations. The recent paper by Liu et al., (2020), also referenced by Keller et al., does report significance and thereby makes a more compelling case. Liu et al. derived uncertainty from 10000 Monte Carlo simulations from monthly statistics to estimate a 68% confidence interval. This procedure could be followed here as well. Another suggestion is to provide a paragraph on uncertainty estimation for the machine learning algorithm in the method section, similar to Petetin et al. (2020). Perhaps here the method of Hengl et al. (2017) could be useful. They describe a procedure for machine-learning uncertainty estimation with the use of the program R and the package 'xgboost'.

Major comment (2): It's unclear how numbers in the result section are constructed from the represented data, no calculation steps are mentioned in the method section. Most importantly, how is the reduction in global NOx emissions of 2.9 TgN calculated?

Line 253 states the following: *"This results in anthropogenic emission adjustment factors of* ***0.3 to 1.4*** *(Fig. A7)."*

Because of the lack of clarification on calculation steps or argumentation, it is unclear how the adjustment factors of 0.3 and 1.4 are determined. Is perhaps the approach of Mendoza & Russel (2001) used to derive adjustment factors for NO_x emissions? Please refer to the used methodology or provide the calculation steps. The in the manuscript referred figure A7 does not provide the calculations either (even though this seems to be suggested). Figure A7 only shows the monthly average perturbations applied to the 2018 anthropogenic base emissions, ranging from 0.5 to 1.5. As a consequence, the resulting quantification of reduction in emissions loses credibility.

Lines 262-266: *"Based on bottom-up emissions estimates for 2015 from the Emission Database for Global Atmospheric Research (EDGAR v5.0_AP, Crippa et al., 2018, 2020) and using a constant concentration/emissions ratio of 0.8 based on the best fit line obtained from the model sensitivity simulation (dashed purple line in Fig. 9a),* ***we calculate*** *that the total reduction in anthropogenic NO_x emissions due to COVID-19 containment measures during the first six months of 2020 amounted to* ***2.9 TgN*** *(**Fig. 9b** and **Table 2**)."*

It is clear a calculation is performed, but not how. How is the quite important 2.9 TgN reduction in anthropogenic NO_x emission due to COVID-19 containment constructed? The 2.9 TgN is not in the referred Table 2 nor in Figure 9b. I urge the authors to provide the taken calculation steps resulting in the (quite important) 2.9 TgN reduction in anthropogenic NO_x emission. This will improve the credibility of that given number.

Major comment (3): The manuscript mentions 'lockdown' situations but does not provide a definition of 'lockdown'. The restrictions vary per country (Ravindran & Shah, 2020) and the definition will have consequences on changes in NO_2 emissions. Some countries only enforced restrictions based on time, while keeping most forms of transport, schools and business open. Others have been reported to only had restrictions for part of the country. Please provide a definition of 'lockdown'.

Lines 156-162: *For Taipei and Rio de Janeiro, the observations and the BCM show little difference (Fig. 4), consistent with* ***the less stringent*** *quarantine measures in these places. Other cities with only short-term NO_2 reductions of less than 25% include Atlanta (USA), Budapest (Hungary), and Melbourne (Australia), again correlating with the* ***comparatively relaxed*** *containment measures in these places (Fig. A1-A3). In contrast, Tokyo (Japan) and Stockholm (Sweden), which also implemented* ***a less aggressive COVID-19 response****, exhibit NO_2 reductions comparable to those of cities with official lockdowns (>20%), suggesting that economic and human activities were similarly subdued in those cities."*

This suggests that degrees of reduction in NO_2 emissions are linked to severity in measures taken by local governments (e.g. lines 156-162), however, the severity of measures per country are not characterised. I suggest providing an overview of 'lockdowns' via a table including severity of measures and start and end dates. As an example, take a look at Ravidran & Shah (2020), where countries were classified on severity by introducing colour codes.

Line 142: *The start and end dates for these are from https://en.wikipedia.org/wiki/COVID19_pandemic_lockdowns or based on local knowledge."*

Because of Wikipedia's quickly changing contents, stating the start and end dates in a table will be an improvement on the derived results and will be more concrete than the stated 'local knowledge'.

Lines 21-22: *Reductions in NO_2* ***correlate*** *with timing and intensity of COVID-19 restrictions, ranging from 60% in severely affected cities (e.g., Wuhan, Milan) to little change (e.g., Rio de Janeiro, Taipei)."*

Also, the manuscript mentions correlations in timing and intensity of COVID-19 restrictions and reductions in NO_2 (e.g. lines 21-22). A quantification of this correlation is however missing. Are these findings only based on eying the figures? Was a correlation test performed? I recommend adding quantification of the correlations.

<u>Minor comments:</u>

Table 1: The links for AEROS (Japan) and EPA Victoria (Australia, Melbourne) do not work.

119: Provide an argumentation on why all observations below or above 2 standard deviations from the mean are removed, contrary to Ma et al. (2020) where observations below or above 3 standard deviations were removed.

Figures 2 and 3: The presentation of the machine learning statistics could be simplified in form of a table. I fail to see how the representation of the machine learning statistics in a graph are useful to the reader (including the location#, since no information is supplied to deduct which location# is which location). I suggest replacing figures 2 and 3 by a table providing statistical performance, similar to Table 4 of Ivatt & Evans (2019).

Figures 4 and 6: Reductions in % are difficult to read in the figures, one must go back to the text for the actual numbers. Consider including the numbers in the figures, so they stand stronger by themselves. Both figures could be shortened on the x-as as well, starting at 2019. The (incomplete) data from 2018 does not contribute to the results. I would even consider replacing both figures 4 and 6 entirely by new figures that better meet the objective of quantifying the difference in reductions of NO_2 and O_3 concentrations (including notification of significance or uncertainty ranges, see major comment 1).

191: Consider replacing the vague terms ‘some countries’ and ‘most countries’. These results are stronger when presented in numbers, for example: '42 out of 46 countries...'

208: Reconsider the phrasing of this result. Belgium, Italy, Luxembourg and Switzerland do not all four show pronounced peaks in early April, based on Figure 7.

221-225: Consider including chemical equations of the mentioned processes to improve readability of this paragraph.

252-258: Move this text to methods, it seems out of place here in the result section.

305-309: Move this text to methods as well, it seems out of place her in the conclusion section.

Figure 10: Consider moving this figure to the result section instead of below the conclusion.

References:

Hengl, T., Leenaars, J. G., Shepherd, K. D., Walsh, M. G., Heuvelink, G. B., Mamo, T., ... & Wheeler, I. (2017). Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutrient Cycling in Agroecosystems, 109(1), 77-102.

Ivatt, P. D. and M. J. Evans. Improving the prediction of an atmospheric chemistry transport model using gradient boosted regression trees, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-753, in review, 2019.

Le Quéré, C., Jackson, R.B., Jones, M.W. et al. Temporary reduction in daily global CO2 emissions during the COVID-19 forced confinement. Nat. Clim. Chang., https://doi.org/10.1038/s41558-020-0797-x, 2020

Liu, Z., Deng, Z., Ciais, P., Lei, R., Davis, S. J., Feng, S., ... & Zhu, B. (2020). COVID-19 causes record decline in global CO2 emissions.

Ma, J., Ding, Y., Cheng, J. C., Jiang, F., Tan, Y., Gan, V. J., & Wan, Z. (2020). Identification of high impact factors of air quality on a national scale using big data and machine learning techniques. Journal of Cleaner Production, 244, 118955

Mendoza-Dominguez, A., & Russell, A. G. (2001). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. Atmospheric Environment, 35(16), 2879-2894.

Petetin, H., Bowdalo, D., Soret, A., Guevara, M., Jorba, O., Serradell, K., and C. Pérez García-Pando. Meteorology-normalized impact of COVID-19 lockdown upon NO2 pollution in Spain, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-446, in review, 2020.

Ravindran, S., & Shah, M. (2020). Unintended consequences of lockdowns: Covid-19 and the shadow pandemic (No. w27562). National Bureau of Economic Research.