Reply Letter to the Comments

Manuscript ID: Preprint egusphere-2023-2425

Manuscript Title: Impact of Weather Patterns and Meteorological Factors on $PM_{2.5}$ and O_3 during the Covid-19 Lockdown in China

Authors: Fuzhen Shen, Michaela I. Hegglin, Yue Yuan

Dear Editor, dear Reviewers,

We would like to thank you and the reviewers for the thoughtful comments regarding our manuscript. These comments have helped us greatly improve the interpretation of our findings, and are important for our future work as well. We have carefully revised the manuscript accordingly. Our point-to-point responses to the reviewers' comments, which are repeated in italics, are given below.

To Reviewer #2

General comment:

This manuscript classified different types of synoptic-scale weather patterns during the first two months from 2015 to 2020 over China. Based on the ML (Machine Learning) models, the authors provided a quantitative assessment of meteorological factors in driving the predictions of PM2.5 and O3 under the specific weather system. The authors provided useful information about the anomalies of PM2.5 and O3 during the study period. This study is well within the scope of ACP. However, the discussions of relative results from the ML analyses were not well be demonstrated. While there is a need for minor revision, particularly in the discussion sections. I suggest that this paper could be published in the journal of ACP in case of the comments is addressed by the authors.

Authors' Reply #0: We thank the reviewer for his/her overall positive assessment of our study. We have made modifications to our manuscript according to the reviewer's helpful suggestions below, which helped to improve the manuscript further.

Highlight the new findings of this study. The authors should demonstrate the creative results, especially to differentiate those in previous studies. I think, studies on meteorological effects on driving the predictions of PM2.5 and O3 have been widely obtained. The authors should introduce more studies about them, and their comparisons with each other should be summarized and discussed in the "Discussion" part.

Authors' Reply #1: We thank the reviewer for this comment. Indeed, many studies have estimated $PM_{2.5}$ and O_3 by using different prediction models[1], but they are limited to explain the final predictions[2-5], especially to provide details of specific input features. In our study, the SHAP module coupled to the GBM

model was run to quantify the local importance of the specific input variables during the haze event in 2020 (***Revised Manuscript version Line****:350-352*).

References:

[1] Wu Y, Lin S, Shi K, Ye Z, Fang Y. Seasonal prediction of daily PM2. 5 concentrations with interpretable machine learning: a case study of Beijing, China. Environmental Science and Pollution Research. 2022 Jun;29(30):45821-36.

[2] Xiao Q, Chang HH, Geng G, Liu Y. An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. Environmental science & technology. 2018 Oct 24;52(22):13260-9.

[3] Zhang M, Wu D, Xue R. Hourly prediction of PM 2.5 concentration in Beijing based on Bi-LSTM neural network. Multimedia Tools and Applications. 2021 Jul;80:24455-68.

[4] Jin H, Chen X, Zhong R, Liu M. Influence and prediction of PM2. 5 through multiple environmental variables in China. Science of The Total Environment. 2022 Nov 25;849:157910.

[5] Weng, X., Forster, G. L., and Nowack, P.: A machine learning approach to quantify meteorological drivers of ozone pollution in China from 2015 to 2019, Atmos. Chem. Phys., 22, 8385–8402, https://doi.org/10.5194/acp-22-8385-2022, 2022.

Analytical method appeared adequate; however some key procedural and QA/QC details are missing. Please provide more details in the manuscript, including the time resolutions of field and reanalysis data, and the uncertainties of ML analysis.

Authors' Reply #2: We thank the reviewer for his/her thoughtful point. More detailed descriptions of the data and ML method have been added to provide the missing information (***Revised Manuscript version Line****:113,405-407*).

Gradient Boosting Machine (GBM) was selected for a quantitative assessment of meteorological factors in driving the predictions of PM5 and O3. Does the authors try to compare it with other models, such as random forest, etc.

Authors' Reply #3: We thank the reviewer for his/her comment. In fact, we had first used a flexible neural tree which is capable to perform automatic feature selection and function approximation[1]. However, the prediction results were not satisfactory and we decided to adopt the GBM model instead. The GBM model has some advantages for the prediction task, like better accuracy, higher efficiency, and capability of handling large-scale data[2]. As a result, GBM was implemented.

References:

[1] Yuehui Chen, Bo Yang, Jiwen Dong, Ajith Abraham, Time-series forecasting using flexible neural tree model, Information Sciences, Volume 174, Issues 3–4, 11 August 2005, Pages 219-235, ISSN 0020-0255, http://dx.doi.org/10.1016/j.ins.2004.10.005

[2] Shen F, Hegglin MI, Luo Y, Yuan Y, Wang B, Flemming J, Wang J, Zhang Y, Chen M, Yang Q, Ge X. Disentangling drivers of air pollutant and health risk changes during the COVID-19 lockdown in China. npj climate and atmospheric science. 2022 Jun 30;5(1):54.

Since the authors focus on the COVID-19 lockdown period, and how about the influence of emission reductions on the anomalies of PM2.5 and O3?

Authors' Reply #4: We thank the reviewer for his/her thoughtful question. In general, anthropogenic reduction has dominated control of the decline in the primary air pollutants but the influence is complex on the secondary air pollutants, including $PM_{2.5}$ and O_3. For example, in NCP, our study here demonstrate that large $PM_{2.5}$ anomalies in SWP-I (-27.6 μg/m^3) and SWP-III (-27.9 μg/m^3) in 2020 were dominated by the anthropogenic reduction. However, the anomalies of $PM_{2.5}$ and O_3 in SWP-II are subject to regional variations due to meteorology, which is what we are studying here specially in the haze event(***Revised Manuscript version Line****:299-306*).

Overall, this paper was well organized, but I still find some explanations for lack of evidence. Please try to improve it.

Specific comments:

Lines 16 to 17: Please clarify the sentence.

Authors' Reply #5: The reviewer might be confused by why North East China (NEC) was abruptly introduced. The reason is that North East China (NEC) and NCP were both significantly under the impact of the double-centre high-pressure system. We had stressed this unnecessary point. To avoid this confusion, NEC was deleted from this sentence(***Revised Manuscript version Line****:18)*.

Line 75: What does the "100%" mean?

Authors' Reply #6: We thank the reviewer for pointing out this inaccurate statement. This sentence should be "achieving the classification 100% automatically" (***Revised Manuscript version Line****:76)*.

Lines 187 to 189: Please clarify the sentence.

Authors' Reply #7: We thank the reviewer for his/her detailed reading of our previous article. We have now modified this ambiguous statement, which highlights the contribution of the meteorological effect(***Revised Manuscript version Line****:202-203)*.

Lines 244 to 245: This sentence is unclear.

Authors' Reply #8: We thank the reviewer for pointing out this confusing statement. We now have changed this sentence to "NCP and NEC (SC) have higher(lower) than expected concentrations for $PM_{2.5}$ (O_3)" (***Revised Manuscript version Line****:261-262)*.