Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., author comment AC2
https://doi.org/10.5194/acp-2022-502-AC2, 2022

Reply on RC2

Yu Lin et al.

Author comment on "Decoupling impacts of weather conditions on interannual variations in concentrations of criteria air pollutants in South China – constraining analysis uncertainties by using multiple analysis tools" by Yu Lin et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-502-AC2, 2022

We greatly appreciate this reviewer for providing the constructive comments, which have helped us improve the paper quality. We have addressed all of the comments carefully, as detailed below.

Analyzing long-term trends by excluding the effects of meteorological factors is critical in the assessment of anthropogenic air pollution factors. In this paper, the authors have used three different methods to decouple meteorological effects and investigate the trends of different pollutants in South China. I find the comparison of these three methods valuable and novel even though the trends were only consistent in 30% of the conditions between these approaches. The manuscript is well-written and has a proper flow to it. The problem statement and introduction are well-written. The discussion of results is clear. However, I think the method section should be expanded and better explained. Here are some general comments for improvement:

RF and BRTs Modeling can be explained better. In particular, how the train-test splitting was applied is not explained thoroughly as it is important in model development. Was this random or sequential? For time series with long-term trends, this split should not be applied randomly, as might be customary in most of the random forest models in other fields, and should be applied sequentially. This is due to the fact that random split will bring extra information to the test validation (e.g. seasonal or weekly trends) that should not be available to the test and cause data leakage.

Response: The software Packages used in this study designed the train-set splitting randomly. This has been clarified in the revision. In the revision, we also added "The independent input variables included temporal variables (hour, day, weekday, week and month), observational concentrations and meteorological parameters (ws, wd, at, rh, dp, blh, tcc, ssr, sp and tp), the top three most influential variables in each modeling were listed in Table S3." Since temporal variables (hour, day, weekday, week and month) have been added in the machine learning, the random or sequential train-test splitting should not affect the performance of machine learning prediction. Moreover, the authors agree with the software developers for the random train-set splitting by considering emission changes through the study period.

The modeling work needs a feature importance analysis. This is very important since some of the features might not add anything to the model and can be simply eliminated from *the analysis. Also, it shows the most influential meteorological factor on the trends. Some additional information can be added to the discussion section about the reasons for observing some of the trends. For example, if authors hypothesize specific regulation(s) as the reason for a specific deweathered trend, that can be added in the discussion section in addition to the introduction.*

Response: Thanks for the advice. The results have been added in Supporting Information Table S3. However, no conclusive results can be obtained on the most influential meteorological factors regardless only top 1 and top 3 to be considered. Some of the results have been added in Results and discussion.

The error or confidence intervals should be added to trend figures (e.g. figures 4 and 5).

Response: The error bars of original annual averages have been added in the revision accordingly. Like all air quality modeling results, the predicted values have no error bars.

Line 37: change "..two-three year.." to "…two-three years"
Line 128: change "predicated" to "predicted"
Line 138: change "indicates" to "indicate"
Line 193: change "decreases" to "decrease".
Line 232: change "conducted to" to "conducted on"
Line 239: change "obtained between" to "obtained by"
Line 269: change "annul" to "annual"

Response: Thanks for the comments. We have revised the manuscript accordingly.

Line 100 and figure 1: "Hourly meteorological data … were obtained from the meteorological observational station at a nearby airport". The meteorology factors, especially wind direction, change rapidly spatially at nearshore sites similar to the ones used in this work. Please mention that the meteorological stations were the closest available to the air quality sites if that is the case. Otherwise, please try to use the closest possible station in your database. Also, this should be mentioned as a source of error in the analysis.

Response: In each city, the hourly averages of air pollutant concentrations at multiple sites in a city were used for machine learning. Thus, the input meteorological data should reflect synoptic weather conditions. Airports usually have a widely open space and the meteorological data at the nearest airports can be reasonably assumed as the synoptic weather conditions.

In the new version, it has been revised as "Like most of studies in the literature (Dai et al., 2021; Ma et al., 2021; Mallet, 2020; Vu et al., 2019; Wang et al., 2020), the meteorological data from the nearest airports were used for the two machine learning methods. The data reflected synoptic weather conditions and were particularly applicable for modelling the hourly averages of air pollutant concentrations at multiple sites in a city."

Figure 3: the range of predicted values is considerably smaller than the observed value. This is an inherited issue with RF and BRT models and should be explained in the text.

Response: In the revision, we added "Note that two machine learning methods always underpredicted the larger values of $PM_{2.5}$ concentrations which occurred less frequently. The same underprediction has also been reported in air quality modelling $PM_{2.5}$ concentrations, which could be due to missing mechanisms enhancing formation of $PM_{2.5}$ under poor dispersion conditions (Chang et al., 2020; Liu et al., 2021a; Shen et al., 2022; Zheng, et al., 2015). For these infrequent cases, the training for two machine learning methods may not be sufficient enough to yield good prediction."