

Atmos. Chem. Phys. Discuss., referee comment RC1
<https://doi.org/10.5194/acp-2022-232-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2022-232

Anonymous Referee #1

Referee comment on "Statistical and machine learning methods for evaluating trends in air quality under changing meteorological conditions" by Minghao Qiu et al., Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2022-232-RC1>, 2022

General Comments:

The idea of evaluating the performance of statistical and machine learning methods used to correct meteorological variability in emission trends using model results is an interesting one, since there have been many recent papers that have been written that use these methods on real-world data sets, without any real metric of their effectiveness in recovering the true trends in emissions.

Overall, this is a well-written paper with a set of carefully designed experiments to assess the performance of different statistical methods to determine meteorology corrected emission trends. The writing is high-quality, includes proper citations, and the figures are clear and easy to understand. I recommend publishing with a few minor corrections to improve the readability of the manuscript and comprehensibility for researchers without a background in statistics.

There were several places where more detail is warranted. Particularly in the description of the application of the different models, there was not a lot of detail and it was difficult to determine how these methods were applied to the data sets. This is important in assessing the conclusions of the paper, that an RF model is preferable to the other statistical methods, as the specific implementation of each method could have a significant impact on its performance. This is particularly true for the machine learning methods.

Specific Comments:

I found the discussion of causal methods in lines 65-76 to be slightly confusing in this paper, since the paper was not focusing on assessing causal links, but rather on testing

counterfactuals—this link should have been more directly made clear, especially for the typical reader of this journal who doesn't have a background in statistics/causal inference.

It would have been useful to have some type of overview cartoon for the different experiments and their relationship to the terms in equation 1— Table 1 was useful, but I had to read the paper through twice before the different simulations were clear to me, and it would have been helpful to have some visual aid for this.

Lines 146. What fraction of the meteorology-concentration relationship is due to changes in natural emissions?

Is there a clear separation between the training and test data for the Random Forest? As I understand this was in part the point of the double-machine learning method, but this should be more clearly spelled out. It would obviously be problematic if both the training and test data are used to evaluate the performance of the RF method to recover the emission-driven trends, as this would give an artificially good performance for this method.

Lines 170-176. Can you spell out what you mean by the "uncorrected" method here? Does that mean the term $f_i(X_{it})$ is neglected in equation 1?

Section 3.4. How are the observations corrected? Are these using the meteorological correction models as determined from the GEOS-Chem model?