Comment on acp-2022-767
Anonymous Referee #2

Referee comment on "Technical note: Improving the European air quality forecast of Copernicus Atmosphere Monitoring Service using machine learning techniques" by Jean-Maxime Bertrand et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-767-RC2, 2022

The authors propose a treatment to an ensemble of CTMs to improve air quality forecasting. This has been done repeatedly in the past using various approaches. The novelty here is the systematic use of ML methods which seems to produce promising results, that is, outscoring the current ensemble method. As I am not a ML expert, I cannot judge the technical implementation of the various algorithms tested. Overall though, I would say the analysis presented makes sense and I trust the authors that the treatments they propose ‘are doing the right thing for the right reason’.

My advice to the editor is to accept the manuscript for publication, pending some clarifications that I invite the authors to consider:

- Quality of figure needs improving
- Specify in plain words what is meant by raw ensemble – is that the unbiased ensemble mean? Possibly I have overlooked, but I cannot locate a definition in the text
- Please comment on what would make your ML methodology better/preferred to other ensemble-improving methods (as for example: https://acp.copernicus.org/articles/14/11791/2014/, https://acp.copernicus.org/articles/13/7153/2013/)
- I believe the authors could make stronger conclusions had they tested their methodology on high pollution episodes, which are notoriously more difficult to predict.
- On the same line of the comment above, would the use of the proposed ML method improve on the predicting of exceedances for regulated pollutant? Please consider adding a comment on these.
- What do you think are the implication of your proposed methodology on gridded output?
- Please avoid the use of acronyms in the conclusions