

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2022-550-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2022-550

Anonymous Referee #2

Referee comment on "Climate-driven deterioration of future ozone pollution in Asia predicted by machine learning with multi-source data" by Huimin Li et al., Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2022-550-RC2>, 2022

The authors demonstrate a framework of using machine learning (ML) to project long-term (2020–2100) surface ozone levels over Asia. The machine learning algorithm (random forest, RF) is trained with ozone data from 2014 to 2018, along with data of meteorology, emissions and other auxiliary data. The trained RF is then used to make ozone projections based on meteorological fields from the four climate scenarios (i.e., SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5) of CMIP6.

This study adopts a data assimilation approach that combines simulations from chemical transport model and observations to better represent real-world ozone levels. This manuscript is within the scope of ACP and has a good scientific quality. I suggest that this manuscript is accepted after the authors address my comments below.

Specific comments:

- In section 2.3, a more detailed description of data assimilation approach should be added to the main text for readers to follow. As a minimum, the authors should include some citations for this section.
- In section 4, the authors mention that one of the limitations to this study is in only using observations across China for the data assimilation. I recommend that the authors also highlight this limitation in section 2.3. For instance, in lines 191 to 193, uncertainties of GEOS-chem simulation are only minimized in China.
- The sentence from lines 196 to 198 appears to suggest that all of the ozone concentrations from the study domain have been assimilated. I don't think this is the case for regions outside of China. I would suggest the authors to be more specific. For example, how have the ozone concentrations from outside of China been processed? Are these directly from the simulation of GEOS-chem?
- In section 2.2, the observational network of CNEMC has an inconsistent number of observational sites through 2014–2019 as the number of sites has grown. Does the

inconsistency of sites affect data assimilation? How the authors handle this potential issue?

- In section 2.4, could the authors give the ranges of the hyperparameters used in the tuning during cross validation and the final selected hyperparameters? Besides, Is the whole set of training data (i.e., all data from 2014-2018 over Asia) randomly split into 10-folds for the cross validation? If in this case, why does the caption of Fig. 2 indicate that the 10-fold cross-validation results are from the year 2019? I suggest that the authors clarify this and give more information regarding the cross-validation process. Moreover, I am concerned that spatial autocorrelation may exist in the cross-validation because of the random split of the training data. For instance, a grid kept for training while the adjacent grid that shares high similarity with this grid is used for validation. This may violate the assumption of data independence. See Ploton et al. (2020) (<https://doi.org/10.1038/s41467-020-18321-y>) that is relevant to the spatial autocorrelation issue.
- Same in section 2.4, variables such as month of the year (MOY) and geographical locations of model grids may not have actual physical meaning. I'm not sure why variables such of these are necessary. Could the authors provide some explanations?
- It seems that the authors construct a single RF emulator to model ozone over the entirety of Asia. One of the advantages of using a single emulator is in the large size of the training data. However, a single emulator is not able to provide information about feature importance for any specific regions. For instance, humidity in southern China is more important, while temperature and solar radiation may be the key features in northern China (e.g., Weng et al., 2022) (<https://doi.org/10.5194/acp-22-8385-2022>). The importance scores in Fig. 4 can only reflect the overall importance of the features from the whole study domain, and the interpretation of these scores should be treated with caution. For instance, if the study domain covers more regions with humidity as the key feature for suppressing ozone production, it is likely that humidity is weighted to be more important than other features. I suggest the authors to address and discuss this limitation.

Minor and technical comments:

- Line 77: Citation of Gong et al. (2019) should be replaced by Gong and Liao (2019). This should be consistent with the citation in Line 76.
- Line 206: Mis-spelling of author name. It should be "Rodriguez".
- In the supplementary, I'm not sure whether Fig. S11 and Fig. S12 follow the same caption as Fig. S8. Are these still percentage differences (%) between 2020–2029 and 2091–2100?

References:

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., and Pélissier, R.: Spatial validation reveals poor predictive performance of large-scale ecological mapping models, *Nat. Commun.*, 11, 1–11, <https://doi.org/10.1038/s41467-020-18321-y>, 2020.

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.