Review of Wang et al 2020

Wang et al. present an analysis of particulate matter over China. They use a Light Gradient Boosting Machine regression technique to combine satellite observations and model simulations to estimate surface $PM_{2.5}$ and PM_{10}. While the broad topics addressed in this manuscript (i.e. pollutant estimation, data-model fusion, machine learning, etc.) are important areas of research, I cannot recommend this paper for publication due to serious methodological issues and the fact that very similar and more comprehensive work has already been published.

Similarity to previous work

Van Donkelaar et al. (2019) and Hammer et al. (2020) use a far simpler statistical method to achieve similar performance at across larger temporal, spatial, and chemical scales. As a specific example, Hammer et al. (2020) use a relatively simple linear adjustment of satellite observations to achieve similar performance at similar resolution globally for 20 years.

Methodological Issues

Section 3.1

This data processing is inappropriate for the input data as described in the supplement. In particular, downscaling wind fields through bicubic interpolation does not preserve mass or energy, nor does it accurately reproduce any higher order variability in the wind fields. More advanced statistical approaches (e.g. Kirchmeier et al., 2014) are necessary to produce physically consistent information.

Section 3.3

The 10x cross validation scheme described in this section does not account for the very large amount of spatial and temporal correlation present in environmental data. This will substantially and inappropriately bias the estimated skill of the machine learning model (e.g. Hastie et al. 2001, Roberts et al., 2017). The authors should perform a more rigorous evaluation of the model skill consistent with the standard in the field (e.g. Barnes et al., 2020). For example, the authors could cross validate through block methods by removing longer time periods of data or removing entire spatial regions of data beyond the autocorrelation length scale.

Section 4.4.

The feature importance results in this section call in question the validity of the modelling approach here. The top ranked variables include those that were unphysically downscaled: wind speeds (U, V, and Ustar) and turbulent evaporation, as well as total column ozone. As the vast majority of the ozone within a total column is present in the stratosphere, any surface predictive capacity associated with that variable is dubious.

Additionally, the high importance ranking of NO_2 potentially indicates that the machine learning model is predicting based simply off the proximity to combustion sources. This explains why the model performance is so poor over more remote regions in Western China and limits the applicability of the method developed.

Figures

All figures with maps in this paper violate the ACP maps policy:
https://www.atmospheric-chemistry-and-physics.net/submission.html#mapsaerials
"In order to depoliticize scientific articles, authors should avoid the drawing of borders or use of contested topographical names."
The inset of the South China Sea does not aid in the scientific interpretation of the results presented in this manuscript in any way, and only confuses the figures.

Specific Comments:
L14 "Most of the existing approaches for the estimation of $PM_{2.5}$ and PM_{10} employed the remote sensing Aerosol Optical Depth (AOD) products as the main variate."

I don't believe this is the case. Most approaches to estimate $PM_{2.5}$ and PM_{10} come from in situ observations of aerosol mass and size distributions.

L29 "conducive to the researches". I'm not sure what this sentence means.

L38 Satellite observations are much more expensive than ground-based monitoring.

L45 This statement needs appropriate citation.

Section 2.2
This section is far too lacking in details to interpret or reproduce the work presented in this paper. The authors should explicitly state the variables used in the main body of the manuscript.

L120 This link does not work for me.

Section 3.2
How were the hyperparameters selected? Was there any optimization or search algorithm applied here?

Figure 5. The colors being split into only 5 bins makes assessing performance difficult. Consider using a continuous colorbar.

L329 The authors should explicitly show evidence for this incorrect estimation through sampling.

References
Barnes, Elizabeth A., Benjamin Toms, James W. Hurrell, Imme Ebert-Uphoff, Chuck Anderson, and David Anderson. "Indicator Patterns of Forced Change Learned by an Artificial Neural Network." Journal of Advances in Modeling Earth Systems 12, no. 9 (2020): e2020MS002195. https://doi.org/10.1029/2020MS002195.

Barnes, Elizabeth A., James W. Hurrell, Imme Ebert-Uphoff, Chuck Anderson, and David Anderson. “Viewing Forced Climate Patterns through an AI Lens.” Geophysical Research Letters 2019. https://doi.org/10.1029/2019GL084944.

Hammer, Melanie S., Aaron van Donkelaar, Chi Li, Alexei Lyapustin, Andrew M. Sayer, N. Christina Hsu, Robert C. Levy, et al. “Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998–2018).” Environmental Science & Technology 54, no. 13 (2020): 7879–90. https://doi.org/10.1021/acs.est.0c01764.
Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.

Kirchmeier, Megan C., David J. Lorenz, and Daniel J. Vimont. “Statistical Downscaling of Daily Wind Speed Variations.” Journal of Applied Meteorology and Climatology 53, no. 3 (2014): 660–75. https://doi.org/10.1175/JAMC-D-13-0230.1.

Roberts, David R., Volker Bahn, Simone Ciuti, Mark S. Boyce, Jane Elith, Gurutzeta Guillera-Arroita, Severin Hauenstein, et al. “Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure.” Ecography 40, no. 8 (2017): 913–29. https://doi.org/10.1111/ecog.02881.

Van Donkelaar, A, Randall V. Martin, Chi Li, and Richard T. Burnett. “Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors.” Environmental Science & Technology 2019. https://doi.org/10.1021/acs.est.8b06392.