Review result of "Estimation of $PM_{2.5}$ Concentration in China Using Linear Hybrid Machine Learning Model." (AMT-2021-64) by Song et al.

Response to RC1:

referee's comments are given in blue,

our responses are given in red.

RC1: The submitted article develops a method to estimate PM2.5 values over China using a linear combination of three machine learning model. The innovative of this approach is the method to have an ensemble PM2.5 data from multiple machine learning model outputs. The research method is solid, and the results are convincing.

Response: We would like to thank the editor and referee for carefully reading the manuscript and providing detailed and constructive comments, which have helped a lot in improving the manuscript. We quote each comment below, followed by our response.

RC1: The background of the research does not cover all of the most recent machine learning produced PM2.5 products over China and provide convincing reason of why this approach is superior to the rest products. The big advantage of using AHI is the high temporal data (sub-hourly), however, the results section does not reflect this advantage.

Response: Due to the early start of this study, the latest research progress was not quoted when writing the research background. To make up for these deficiencies, we will add 18 references to the manuscript. These references are listed at the page 8-10 of this document.

The advantage that AHI can provide high temporal resolution data is also discussed, but for some reasons it was not included in the previous version of the manuscript. In the revised manuscript we have added this content. The results are shown in the figure below.

Figure 6 shows the scatterplot fitted with the inversion results of the mixed model from 9:00-17:00 Local Time. The model R^2 ranged from 0.556 to 0.88 at different times. Except for 17:00 when the model had the worst performance, the model R^2 exceeded 0.7 at other times, indicating that the model had a good performance. The optimal performance time is 13:00, R^2 is 0.88. According to the results, the hourly differences in model performance were significant.

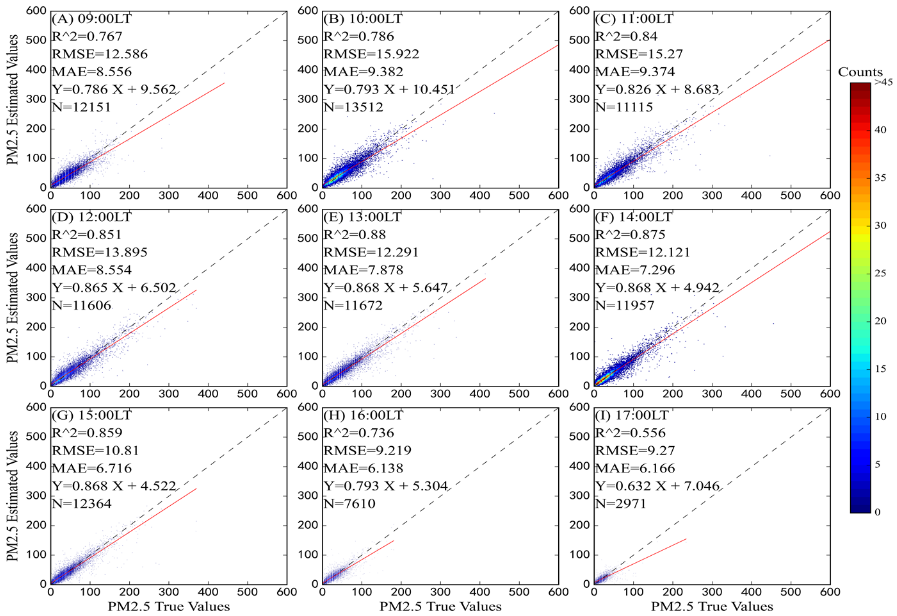

Figure 6 Hourly validation of model performance

The temporal distribution of $PM_{2.5}$ is shown in Figure 10, The $PM_{2.5}$ concentration began to rise from 9:00, and peaked at 55.65μg/m3 between 10:00 and 11:00 every day. After that, it maintained a high concentration until 15:00, and began to decrease. In the most polluted areas of China, the peak concentration of $PM_{2.5}$ can reach 85.05μg/m^3, while the peak in the less polluted areas is only about 40μg/m^3. On a national scale, daily $PM_{2.5}$ concentrations fluctuate little.

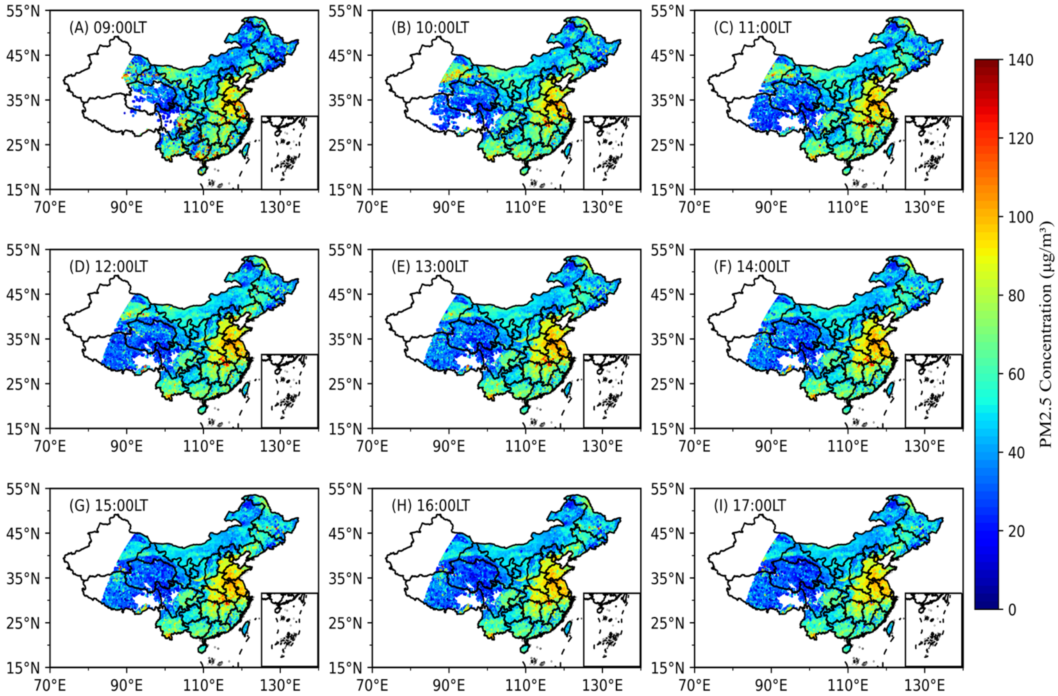

Figure 10 Hourly distribution of $PM_{2.5}$ in China in 2019

RC1: The most contribution of this study is the linear hybrid ML model. However, the paper does not explain details of this procedure. For example, why using linear combination, and how are the coefficients are determined? Instead of a simple regression, complexed error evaluations of individual

ML PM2.5 data may provide insights on a better way of combining these model outputs.

Response: Wolpert et al. (1992) pointed out that the combination of multiple models can improve the robustness and generalization ability of the model. In other words, machine learning models can be integrated in the same way as multi-mode ensemble forecasting. Thus, we could further improve the accuracy of the fitting by hybrid model.

The coefficient is determined by multiple linear regression model. Firstly, we use three sub-models to calculate the predicted value under the corresponding model. Then, multiple linear regressions are performed between the calculated predicted values and the label values in the original data. Finally, the output coefficients and intercepts of the multiple linear regression model are taken as the parameters of the **RGD-LHMLM**.

RC1: The parameter impotency is listed but no further explanation of parameter selection is mentioned.

Response: We mainly used feature importance to analyze the contribution of different parameters to the model. This can provide an explanation of the interpretability of the model. The selection of parameters is mainly based on the variable information provided in some references. Finally, these characteristics we screened are all physical quantities that have a certain influence on $PM_{2.5}$, such as AOD, boundary layer height, relative humidity, population density.

RC1: Bias analysis as functions of other influence factors is needed to better understand the uncertainties in PM2.5 product.

Response: We use formula (5) and formula (6) to calculate the value of the Bias and the generalization error of the Bias (GEB). It is generally believed that when we take the generalization error, the Bias must be expressed in the form of a square. The average GEB between estimated $PM_{2.5}$ based on the RGD-LHMLM and measured $PM_{2.5}$ are shown in Table 1.

The results show that the average GEB of the mixed model is smaller, and the deviation between the predicted data and the label data is lower.

$$\mathrm{Bias} = y_{\mathrm{label,i}} - y_{\mathrm{predict,i}} \quad (5)$$

$$\mathrm{GEB} = \frac{\sum_{i=1}^{N}(y_{\mathrm{label,i}} - y_{\mathrm{predict,i}})^2}{N} \quad (6)$$

Table 1 Comparison of model accuracy

Model	Fitting				Validation			
	R^2	RMSE	MAE	GEB	R^2	RMSE	MAE	GEB
RF	0.95	6.99	4.05	114.19	0.79	14.89	9.33	208.97
GBRT	0.96	6.87	4.52	110.00	0.81	14.09	9.18	198.65
DNN	0.97	5.03	3.49	59.16	0.80	14.45	9.06	221.86
RGD-LHMLM	0.98	4.39	3.00	44.97	0.84	12.92	8.01	166.95

Then the bias of the mixed model in different $PM_{2.5}$ concentration ranges was analyzed. As shown in the figure below: The average bias of the mixed model in different $PM_{2.5}$ concentration ranges was analyzed, and the result is shown in the figure 4. when the $PM_{2.5}$ concentration is less than 60 μg/m^3, the average bias of the model is less than 0. As the $PM_{2.5}$ concentration increases, the model deviation gradually increases. In other words, when the $PM_{2.5}$ concentration is small, the predicted value of the model will generally overestimate $PM_{2.5}$, and when the $PM_{2.5}$ further increases, it will underestimate the $PM_{2.5}$ concentration.

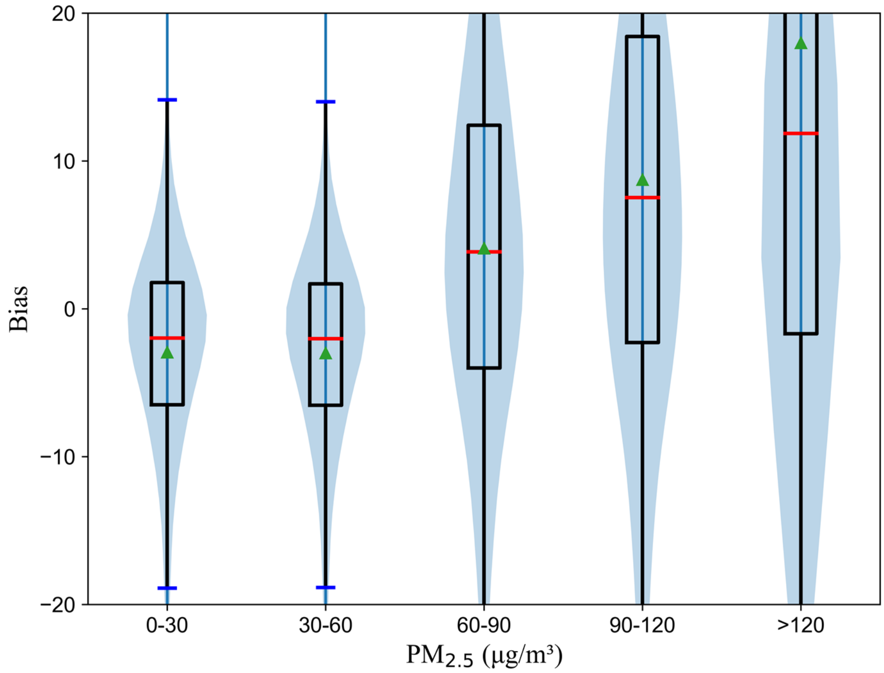

Figure 4 Bias between model predicted values and label values

We have compared other studies with our own and listed the results in Table 1:

Table 1

Model	R^2	RMSE	MAE	Reference
Stacking model	0.85	17.3	10.5	(Chen et al., 2019)
Two-stage random forests (YRD)	0.86	12.4	/	(Tang et al., 2019)
LME (BTH)	0.86	24.5	14.2	(Wang et al., 2017)
GTWR	0.78	20.10	/	(Xue et al., 2020b)
STLG	0.85	13.62	8.49	(Wei et al., 2021a)
RGD-LHMLM	0.84	12.92	8.01	This paper

References

(Yin et al., 2021;Xue et al., 2021;Wei et al., 2021b;Mao et al., 2021;Chen et al., 2021;Xue et al., 2020a;Wei et al., 2020;Wei et al., 2019a;Zeng et al., 2021;Guo et al., 2021;Qin et al., 2017;Li et al., 2016;Zheng et al., 2017;Gui et al., 2019;Shen et al., 2016;Li et al., 2021;Yang et al., 2016;Gui et al., 2020) (China, 2012;Yoshida et al., 2018)(Zhang et al., 2019)(Wei et al., 2019b)

Chen, B. J., You, S. X., Ye, Y., Fu, Y. Y., Ye, Z. R., Deng, J. S., Wang, K., and Hong, Y.: An interpretable self-adaptive deep neural network for estimating daily spatially-continuous PM2.5 concentrations across China,Sci Total Environ, 768,https://doi.org/10.1016/j.scitotenv.2020.144724, 2021.

Chen, J. P., Yin, J. H., Zang, L., Zhang, T. X., and Zhao, M. D.: Stacking machine learning model for estimating hourly PM2.5 in China based on Himawari 8 aerosol optical depth data,Sci Total Environ, 697,https://doi.org/10.1016/j.scitotenv.2019.134021, 2019.

China: Ambient air quality standards. GB 3095-2012., China Environmental Science Press, Beijing, 2012.

Gui, K., Che, H. Z., Wang, Y. Q., Wang, H., Zhang, L., Zhao, H. J., Zheng, Y., Sun, T. Z., and Zhang, X. Y.: Satellite-derived PM2.5 concentration trends over Eastern China from 1998 to 2016: Relationships to emissions and meteorological parameters,Environ Pollut, 247, 1125-1133,https://doi.org/10.1016/j.envpol.2019.01.056, 2019.

Gui, K., Che, H. Z., Zeng, Z. L., Wang, Y. Q., Zhai, S. X., Wang, Z. M., Luo, M., Zhang, L., Liao, T. T., Zhao, H. J., Li, L., Zheng, Y., and Zhang, X. Y.: Construction of a virtual PM2.5 observation network in China based on high-density surface meteorological observations using the Extreme Gradient Boosting model,Environ Int, 141,https://doi.org/10.1016/j.envint.2020.105801, 2020.

Guo, B., Zhang, D. M., Pei, L., Su, Y., Wang, X. X., Bian, Y., Zhang, D. H., Yao, W. Q., Zhou, Z. X., and Guo, L. Y.: Estimating PM2.5 concentrations via random forest method using satellite, auxiliary, and ground-level station dataset at multiple temporal scales across China in 2017,Sci Total Environ, 778,https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.146288, 2021.

Li, H. M., Yang, Y., Wang, H. L., Li, B. J., Wang, P. Y., Li, J. D., and Liao, H.: Constructing a spatiotemporally coherent long-term PM2.5 concentration dataset over China during 1980-2019 using a machine learning approach,Sci Total Environ, 765,https://doi.org/10.1016/j.scitotenv.2020.144263, 2021.

Li, Z. Q., Zhang, Y., Shao, J., Li, B. S., Hong, J., Liu, D., Li, D. H., Wei, P., Li, W., Li, L., Zhang, F. X., Guo, J., Deng, Q., Wang, B. X., Cui, C. L., Zhang, W. C., Wang, Z. Z., Lv, Y., Xu, H., Chen, X. F., Li, L., and Qie, L. L.: Remote sensing of atmospheric particulate mass of dry PM2.5 near the ground: Method validation using ground-based measurements,Remote Sens Environ, 173, 59-68,https://doi.org/10.1016/j.rse.2015.11.019, 2016.

Mao, F. Y., Hong, J., Min, Q. L., Gong, W., Zang, L., and Yin, J. H.: Estimating hourly full-coverage PM2.5 over China based on TOA reflectance data from the Fengyun-4A satellite,Environ Pollut, 270,https://doi.org/10.1016/j.envpol.2020.116119, 2021.

Qin, K., Wang, L. Y., Wu, L. X., Xu, J., Rao, L. L., Letu, H., Shi, T. W., and Wang, R. F.: A campaign for investigating aerosol optical properties during winter hazes over Shijiazhuang, China,Atmos Res, 198, 113-122,https://doi.org/10.1016/j.atmosres.2017.08.018, 2017.

Shen, Z. X., Cao, J. J., Zhang, L. M., Zhang, Q., Huang, R. J., Liu, S. X., Zhao, Z. Z., Zhu, C. S., Lei, Y. L., Xu, H. M., and Zheng, C. L.: Retrieving historical ambient PM2.5 concentrations using existing visibility measurements in Xi'an, Northwest China,Atmos Environ, 126, 15-20,https://doi.org/10.1016/j.atmosenv.2015.11.040, 2016.

Tang, D., Liu, D. R., Tang, Y. L., Seyler, B. C., Deng, X. F., and Zhan, Y.: Comparison of GOCI and Himawari-8 aerosol optical depth for deriving full-coverage hourly PM2.5 across the Yangtze River Delta,Atmos Environ, 217,https://doi.org/10.1016/j.atmosenv.2019.116973, 2019.
Wang, W., Mao, F. Y., Du, L., Pan, Z. X., Gong, W., and Fang, S. H.: Deriving Hourly PM2.5 Concentrations from Himawari-8 AODs over Beijing-Tianjin-Hebei in China,Remote Sens-Basel, 9,https://doi.org/10.3390/rs9080858, 2017.
Wei, J., Huang, W., Li, Z. Q., Xue, W. H., Peng, Y. R., Sun, L., and Cribb, M.: Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach,Remote Sens Environ, 231,https://doi.org/10.1016/j.rse.2019.111221, 2019a.
Wei, J., Li, Z., Sun, L., Peng, Y., Zhang, Z., Li, Z., Su, T., Feng, L., Cai, Z., and Wu, H.: Evaluation and uncertainty estimate of next-generation geostationary meteorological Himawari-8/AHI aerosol products,Sci Total Environ, 692, 879-891,https://doi.org/10.1016/j.scitotenv.2019.07.326, 2019b.
Wei, J., Li, Z. Q., Cribb, M., Huang, W., Xue, W. H., Sun, L., Guo, J. P., Peng, Y. R., Li, J., Lyapustin, A., Liu, L., Wu, H., and Song, Y. M.: Improved 1 km resolution PM2.5 estimates across China using enhanced space-time extremely randomized trees,Atmos Chem Phys, 20, 3273-3289,https://doi.org/10.5194/acp-20-3273-2020, 2020.
Wei, J., Li, Z., Pinker, R. T., Wang, J., Sun, L., Xue, W., Li, R., and Cribb, M.: Himawari-8-derived diurnal variations in ground-level PM2.5 pollution across China using the fast space-time Light Gradient Boosting Machine (LightGBM),Atmos. Chem. Phys., 21, 7863-7880,https://doi.org/10.5194/acp-21-7863-2021, 2021a.
Wei, J., Li, Z. Q., Lyapustin, A., Sun, L., Peng, Y. R., Xue, W. H., Su, T. N., and Cribb, M.: Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications,Remote Sens Environ, 252,https://doi.org/10.1016/j.rse.2020.112136, 2021b.
Xue, W. H., Zhang, J., Zhong, C., Ji, D. Y., and Huang, W.: Satellite-derived spatiotemporal PM2.5 concentrations and variations from 2006 to 2017 in China,Sci Total Environ, 712,https://doi.org/10.1016/j.scitotenv.2019.134577, 2020a.
Xue, W. H., Zhang, J., Zhong, C., Li, X. Y., and Wei, J.: Spatiotemporal PM2.5 variations and its response to the industrial structure from 2000 to 2018 in the Beijing-Tianjin-Hebei region,J Clean Prod, 279,https://doi.org/10.1016/j.jclepro.2020.123742, 2021.
Xue, Y., Li, Y., Guang, J., Tugui, A., She, L., Qin, K., Fan, C., Che, Y. H., Xie, Y. Q., Wen, Y. N., and Wang, Z. X.: Hourly PM2.5 Estimation over Central and Eastern China Based on Himawari-8 Data,Remote Sens-Basel, 12,https://doi.org/10.3390/rs12050855, 2020b.
Yang, Y., Liao, H., and Lou, S.: Increase in winter haze over eastern China in recent decades: Roles of variations in meteorological parameters and anthropogenic emissions,J Geophys Res-Atmos, 121, 13050-13065,https://doi.org/10.1002/2016jd025136, 2016.
Yin, J. H., Mao, F. Y., Zang, L., Chen, J. P., Lu, X., and Hong, J.: Retrieving PM2.5 with high spatio-temporal coverage by TOA reflectance of Himawari-8,Atmospheric Pollution Research, 12, 14-20,https://doi.org/10.1016/j.apr.2021.02.007, 2021.
Yoshida, M., Kikuchi, M., Nagao, T. M., Murakami, H., Nomaki, T., and Higurashi, A.: Common Retrieval of Aerosol Properties for Imaging Satellite Sensors,Journal of the Meteorological Society of Japan. Ser. II, 96B, 193-209,https://doi.org/10.2151/jmsj.2018-039, 2018.
Zeng, Z. L., Gui, K., Wang, Z. M., Luo, M., Geng, H., Ge, E. J., An, J. C., Song, X. Y., Ning, G. C., Zhai, S. X., and Liu, H. Z.: Estimating hourly surface PM2.5 concentrations across China from high-density meteorological observations by machine learning,Atmos Res,

254,https://doi.org/10.1016/j.atmosres.2021.105516, 2021.

Zheng, C. W., Zhao, C. F., Zhu, Y. N., Wang, Y., Shi, X. Q., Wu, X. L., Chen, T. M., Wu, F., and Qiu, Y. M.: Analysis of influential factors for the relationship between PM2.5 and AOD in Beijing,Atmos Chem Phys, 17, 13473-13489,https://doi.org/10.5194/acp-17-13473-2017, 2017.