Supplementary Material

Spatiotemporal Weighted for Improving the Satellite-Based High-Resolution Ground PM2.5 Estimation Using the Light Gradient Boosting Machine

Xinyu Yu [123,†], Mengzhu Xi [123,†] ,Liyang Wu [123,*] and Hui Zheng [123]

[1] College of Geography and Environmental Science, Henan University, Kaifeng 475004, China
[2] Key Laboratory of Geospatial Technology for Middle and Lower Yellow River Regions, Ministry of Education, College of Environment and Planning, Henan University, Kaifeng 475004, China
[3] Henan Key Laboratory of Integrated Air Pollution Control and Ecological Security, Kaifeng 475004, China
* Correspondence: wuly@henu.edu.cn
† These authors contributed equally to this work.

Table S1. Summary of the data sources used in this study.

Variable	Description	Spatial Resolution	Temporal Resolution	Data Source
PM2.5	PM2.5	-	hourly	CNEMC (http://www.cnemc.cn)
AOD	Aerosol optical depth (550 nm)	1 km × 1 km	Daily	MODIS products (MCD19A2) (https://ladsweb.nascom.nasa.gov/)
BLH	Boundary layer height	0.25°×0.25°	Hourly	ERA5 reanalysis products (https://cds.climate.copernicus.eu/)
RH	Relative humidity			
TEM	2-m air temperature			
SP	Surface pressure			
WS	10-m wind speed			
NDVI	Normalized difference vegetation index	1 km × 1 km	Monthly	MODIS products (https://ladsweb.nascom.nasa.gov/)
LUCC	Land-use cover	500 m × 500 m	Annual	
DEM	Elevation	30 m × 30 m	-	SRTM (http://gdex.cr.usgs.gov/gdex/)

Table S2. Comparison of performances of the STW-LightGBM model and other AOD-based models applied in previous studies at the national scale in China.

Reference	Model	Period	Spital Resolution	Cross Validation		Regression Equation
				R^2	RMSE ($\mu g/m^3$)	
Li, *et al.* [1]	Geo-BPNN	2015	10 km	0.84	15.23	-
	Geo-GRNN			0.82	16.93	-
	Geo-DBN			0.88	13.03	-
He and Huang [2]	GTWR	2015	3 km	0.80	18.00	Y=0.81X+11.69
	Two-stage			0.72	21.01	Y=0.73X+16.67
	D-GWR			0.71	21.21	Y=0.79X+12.92
You, *et al.* [3]	GWR	2014	3 km	0.79	18.60	Y=0.83X+9.44
Li, *et al.* [4]	GRNN	2013–2014	3 km	0.816	20.93	Y=0.616X+22.90
Xue, *et al.* [5]	ML	2013–2016	3 km	0.53	30.40	Y=0.53X+25.3
	ML + GAM			0.61	27.80	Y=0.61X+21.2
Chen, *et al.* [6]	RF	2014-2016	10 km	0.83	18.08	Y=1.07X-4.64
Chen, *et al.* [7]	XGBoost	2014–2015	3 km	0.86	14.98	-
Zhang, *et al.* [8]	GBDT	2017	3 km	0.81	11.57	Y=0.78X+9.97
Wei, *et al.* [9]	RF	2016	1 km	0.81	17.91	Y=0.77X+12.56
	STRF			0.85	15.57	Y=0.82X+9.64
Wei, *et al.* [10]	STET	2013-2018	1 km	0.887	14.6	Y=0.86X+7.11
He, *et al.* [11]	ASTR	2013-2018	1 km	0.59	27.18	Y=0.61X+20.97
This study	STW-LightGBM	2015-2020	1 km	0.92	10.00	Y=0.9X+0.47

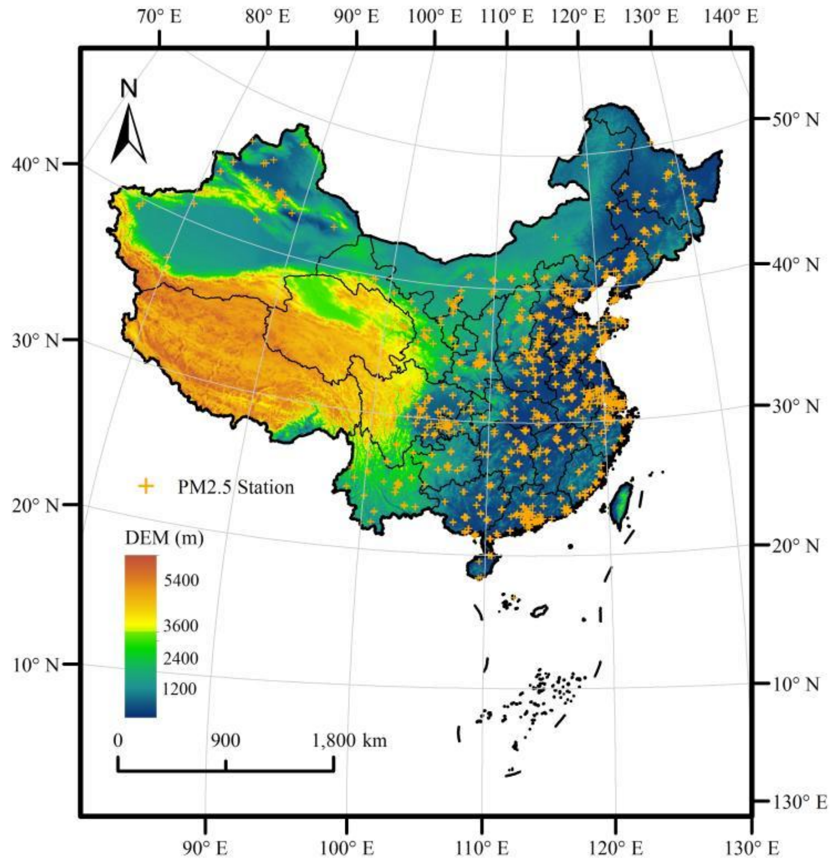

Figure S1. Spatial distribution of PM2.5 monitoring stations included in this study. The background is digital elevation, and the yellow cross is PM2.5 station.

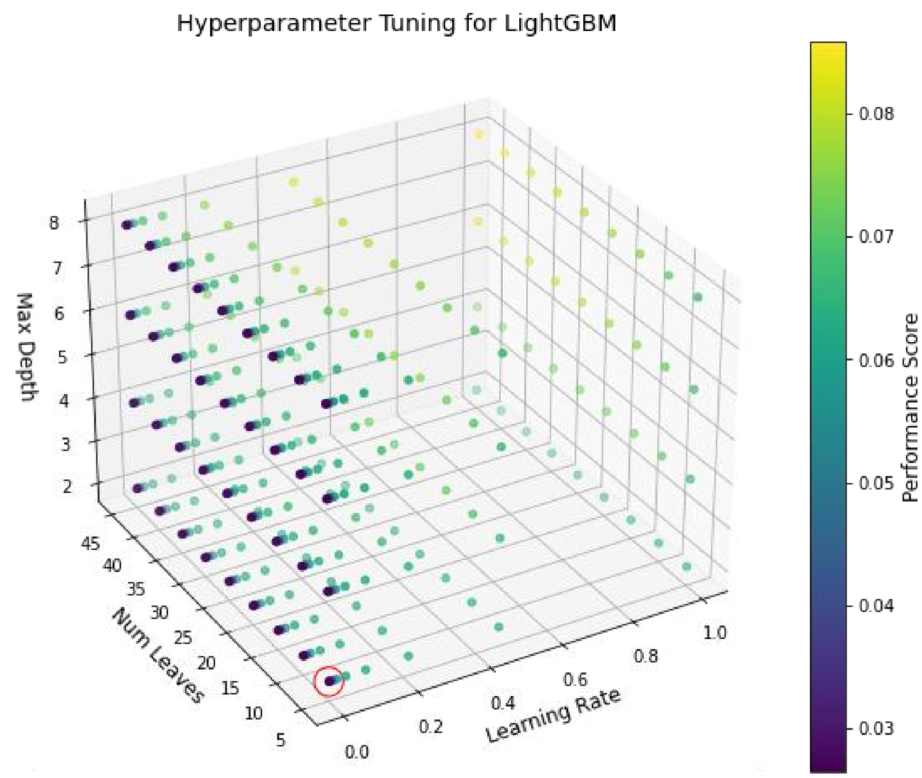

Figure S2. The visualization result of STW-LightGBM parameter adjustment.

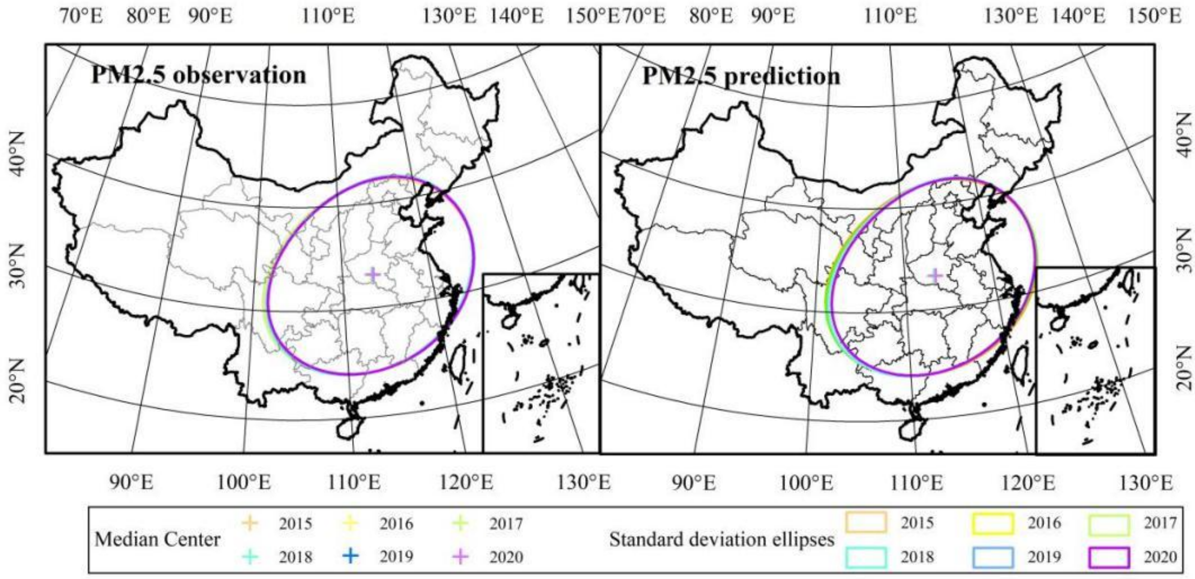

Figure S3. China's PM2.5 observed and predicted standard error ellipse from 2015 to 2020.

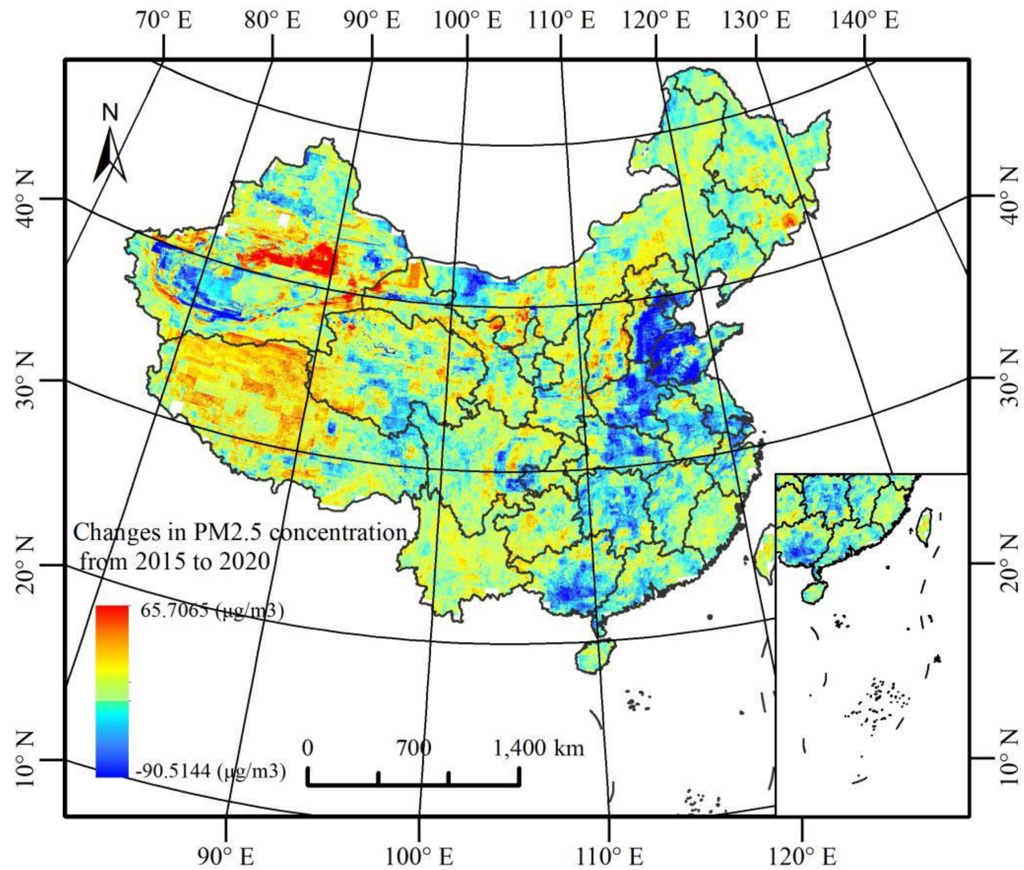

Figure S4. Changes in PM2.5 concentration in China from 2015 to 2020.

References

1. Li, T.; Shen, H.; Yuan, Q.; Zhang, X.; Zhang, L. Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. *Geophys. Res. Lett.* **2017**, 44, 11,985-911,993, doi:https://doi.org/10.1002/2017GL075710.
2. He, Q.; Huang, B. Satellite-based mapping of daily high-resolution ground PM2.5 in China via space-time regression modeling. *Remote Sens. Environ.* **2018**, 206, 72-83, doi:10.1016/j.rse.2017.12.018.
3. You, W.; Zang, Z.; Zhang, L.; Li, Y.; Pan, X.; Wang, W. National-Scale Estimates of Ground-Level PM2.5 Concentration in China Using Geographically Weighted Regression Based on 3 km Resolution MODIS AOD. *Remote Sens.* **2016**, 8, doi:10.3390/rs8030184.
4. Li, T.; Shen, H.; Zeng, C.; Yuan, Q.; Zhang, L. Point-surface fusion of station measurements and satellite observations for mapping PM2.5 distribution in China: Methods and assessment. *Atmos. Environ.* **2017**, 152, 477-489, doi:https://doi.org/10.1016/j.atmosenv.2017.01.004.
5. Xue, T.; Zheng, Y.; Tong, D.; Zheng, B.; Li, X.; Zhu, T.; Zhang, Q. Spatiotemporal continuous estimates of PM2.5 concentrations in China, 2000–2016: A machine learning method with inputs from satellites, chemical transport model, and ground observations. *Environ. Int.* **2019**, 123, 345-357, doi:https://doi.org/10.1016/j.envint.2018.11.075.
6. Chen, G.; Li, S.; Knibbs, L.D.; Hamm, N.A.S.; Cao, W.; Li, T.; Guo, J.; Ren, H.; Abramson, M.J.; Guo, Y. A machine learning method to estimate PM2.5 concentrations across China with remote sensing, meteorological and land use information. *Sci. Total Environ.* **2018**, 636, 52-60, doi:10.1016/j.scitotenv.2018.04.251.
7. Chen, Z.; Zhang, T.; Zhang, R.; Zhu, Z.-M.; Yang, J.; Chen, P.-Y.; Ou, C.-Q.; Guo, Y. Extreme gradient boosting model to estimate PM2.5 concentrations with missing-filled satellite data in China. *Atmos. Environ.* **2019**, 202, 180-189, doi:https://doi.org/10.1016/j.atmosenv.2019.01.027.
8. Zhang, T.; He, W.; Zheng, H.; Cui, Y.; Song, H.; Fu, S. Satellite-based ground PM2.5 estimation using a gradient boosting decision tree. Chemosphere **2021**, 268, 128801,

doi:https://doi.org/10.1016/j.chemosphere.2020.128801.

9. Wei, J.; Huang, W.; Li, Z.; Xue, W.; Peng, Y.; Sun, L.; Cribb, M. Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach. *Remote Sens. Environ.* **2019**, 231, doi:10.1016/j.rse.2019.111221.
10. Wei, J.; Li, Z.; Lyapustin, A.; Sun, L.; Peng, Y.; Xue, W.; Su, T.; 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.* **2021**, 252, doi:10.1016/j.rse.2020.112136.
11. He, Q.; Gao, K.; Zhang, L.; Song, Y.; Zhang, M. Satellite-derived 1-km estimates and long-term trends of PM2.5 concentrations in China from 2000 to 2018. *Environ. Int.* **2021**, 156, 106726, doi:https://doi.org/10.1016/j.envint.2021.106726.