

EGU2020-12943

<https://doi.org/10.5194/egusphere-egu2020-12943>

EGU General Assembly 2020

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Correction for the Measurements of Particulate Matter Sensors through Machine Learning

Zhen Cheng and Qiaofeng Guo

School of Environmental Science and Engineering, Shanghai Jiao Tong University, Shanghai, China(chengz88@sjtu.edu.cn)

Instruments based on light scattering used to measure total suspended particulate (TSP) concentrations have the advantages of fast response, small size and low cost as compared to the gravimetric reference method. However, the relationship between scattering intensity and TSP mass concentration varies nonlinearly with both environmental conditions and particle properties, making it difficult to make corrections. This study applied four machine learning models (support vector machine, random forest, gradient boosting regression trees and an artificial neural network) to correct scattering measurements for TSP mass concentrations. A total of 1141 hourly records of collocated gravimetric and light scattering measurements taken at 17 urban sites in Shanghai, China were used for model training and validation. All four machine learning models improved the linear regressions between scattering and gravimetric mass by increasing slopes from 0.4 to 0.9-1.1 and coefficients of determination from 0.1 to 0.8-0.9. Partial dependence plots indicate that TSP concentrations determined by light scattering instruments increased continuously in the PM_{2.5} concentration range of ~0-80 $\mu\text{g}/\text{m}^3$; however, they leveled off above PM₁₀ and TSP concentrations of ~60 and 200 $\mu\text{g}/\text{m}^3$, respectively. The TSP mass concentrations determined by scattering showed an exponential growth after relative humidity exceeded 70%, in agreement with previous studies on hygroscopic growth of fine particles. This study demonstrates that machine learning models can effectively improve the correlation between light scattering measurements and TSP mass concentrations with filter-based methods. Interpretation analysis further provides scientific insights into the major factors (e.g., hygroscopic growth) that cause scattering measurements to deviate from TSP mass concentrations besides other factors like fluctuation of mass density and refractive index.

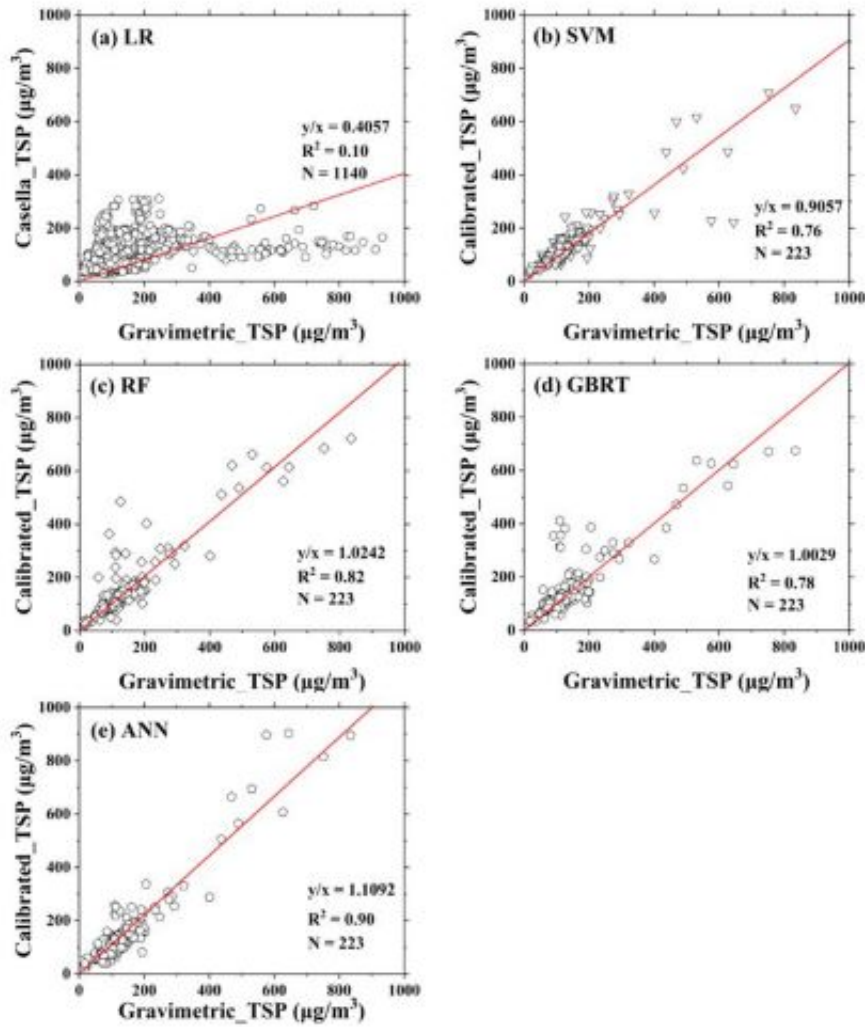


Figure 1. Comparison of TSP concentrations determined by light scattering and machine learning model outputs with those by gravimetric analyses. **(a)** LR: Linear Regression; **(b)** SVM: Support Vector Machine; **(c)** RF: Random Forest; **(d)** GBRT: Gradient Boosting Regression Tree; **(e)** ANN: Artificial Neural Network. y/x represents the slope, R^2 is the coefficient of determination, N means the volume of the dataset.

