

**Note to readers with disabilities:** *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to [508 standards](#) due to the complexity of the information being presented. If you need assistance accessing journal content, please contact [ehp508@niehs.nih.gov](mailto:ehp508@niehs.nih.gov). Our staff will work with you to assess and meet your accessibility needs within 3 working days.

### **Supplemental Material**

#### **Deep Ensemble Machine Learning Framework for the Estimation of PM<sub>2.5</sub> Concentrations**

Wenhua Yu, Shanshan Li, Tingting Ye, Rongbin Xu, Jiangning Song, and Yuming Guo

#### **Table of Contents**

**Table S1.** The model hyper-parameters and computing cost information for study models.

**Table S2.** The variables information and sources.

**Table S3.** The seasonal distribution of daily average PM<sub>2.5</sub> ( $\mu\text{g}/\text{m}^3$ ) from 2015 to 2019 in Italy.

**Table S4.** The distribution comparison between observed PM<sub>2.5</sub> concentration and estimated PM<sub>2.5</sub> by RF, XGBoost, SL, and DEML in Italy from 2015 to 2019.

**Figure S1.** R<sup>2</sup> for each monitoring station based on predicted and observed PM<sub>2.5</sub> from 2015 to 2019 in Italy.

**Figure S2.** Italian regional distribution in Cluster-based spatial Cross-validation.

**Figure S3.** The dissimilarity index (DI) of the spatial locations in Italy with the ground stations in a specific day.

**Figure S4.** PM<sub>2.5</sub> imputation performance of the RF, XGBoost, and DEML by PM<sub>10</sub> from 2015 to 2019 in Italy.

**Figure S5.** The evaluation of the importance of the explanatory variables in PM<sub>2.5</sub> estimation by RF and XGboost from 2015 to 2019 in Italy.