Long-term trends in $PM_{2.5}$ pollution in China: Combining MODIS and VIIRS satellites via machine learning

Jing Wei & Zhanqing Li

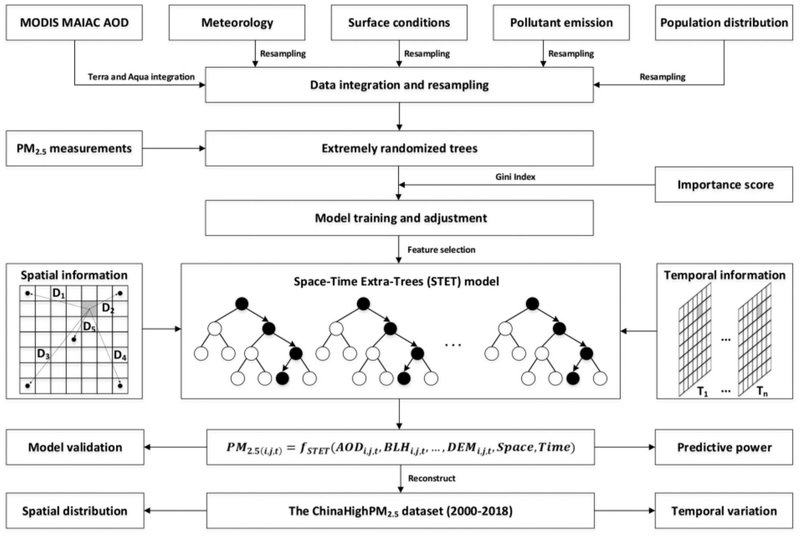
Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA

Introduction

- Air pollution has become one of the major environmental risks affecting human health and can raise the risk of death. $PM_{2.5}$ has become the fifth leading lethal risk around the world.
- Currently, $PM_{2.5}$ estimates and predictions were less accurate with overall low accuracies at coarse spatial resolutions due to the limitations of the models themselves and data sources.
- More accurate methods especially machine learning that improve the prediction accuracy of $PM_{2.5}$ concentrations are thus needed, and this would make up for the gap in studies on $PM_{2.5}$ variations across China.
- We produce for the first time a high-resolution and high-quality $PM_{2.5}$ dataset for China, reconstructing the period since 2000, and perform the investigation of long-term spatiotemporal $PM_{2.5}$ variations.

Materials and methods

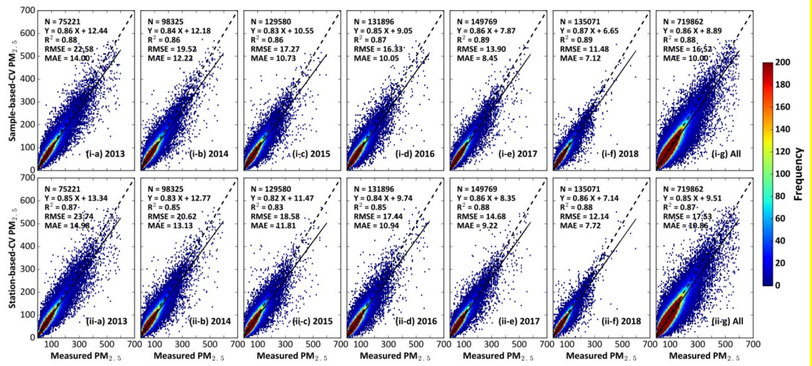
- **Method**: We proposed a space-time extremely randomized trees (STET) model by involving the spatiotemporal information into the machine learning model to construct the robust AOD-$PM_{2.5}$ conversion relationship.



- **AOD data:** MODIS MAIAC 1 km (MCD19A2) AOD product and VIIRS Deep Blue 6 km (AERDB) AOD product
- **Meteorological data**: Boundary layer heights , temperature, relative humidity, precipitation, evaporation, surface pressure, wind speed, and wind direction, were extracted from the ERA5 reanalysis.
- **Auxiliary data**: CAMS Emission Inventory including aerosol precursors (i.e., NH_3, NO_x, and SO_2), MERRA2 aerosol simulations, NDVI, DEM, and LandScan population distribution, et al.

Model performance

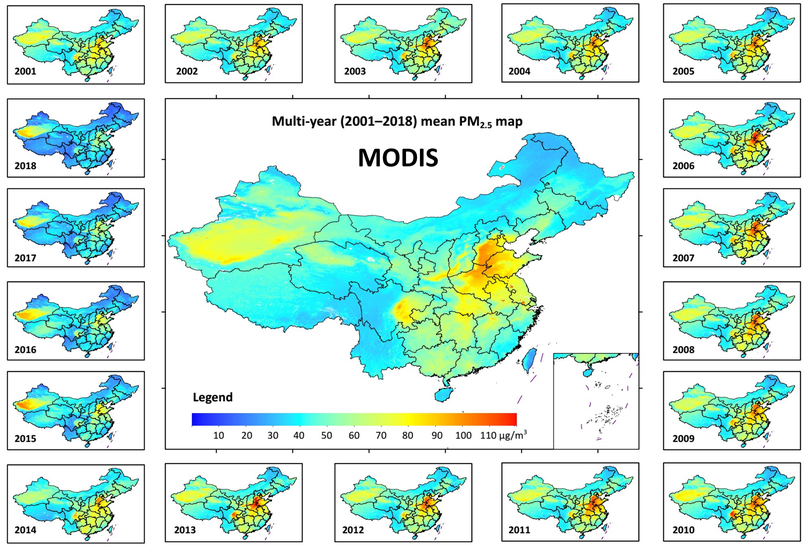
- Two independent approaches including out-of-sample and out-of-station ten-fold cross-validation methods are performed to evaluate the overall accuracy and spatial prediction ability of daily $PM_{2.5}$ estimates and predictions.

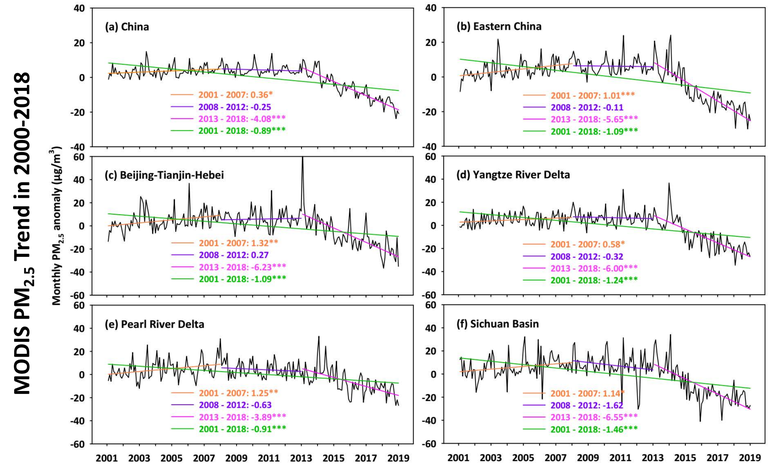


Overall accuracy: CV-R^2 = 0.88, and RMSE = 16.5 μg/m^3
Prediction accuracy: CV-R^2 = 0.87, and RMSE = 17.5 μg/m^3

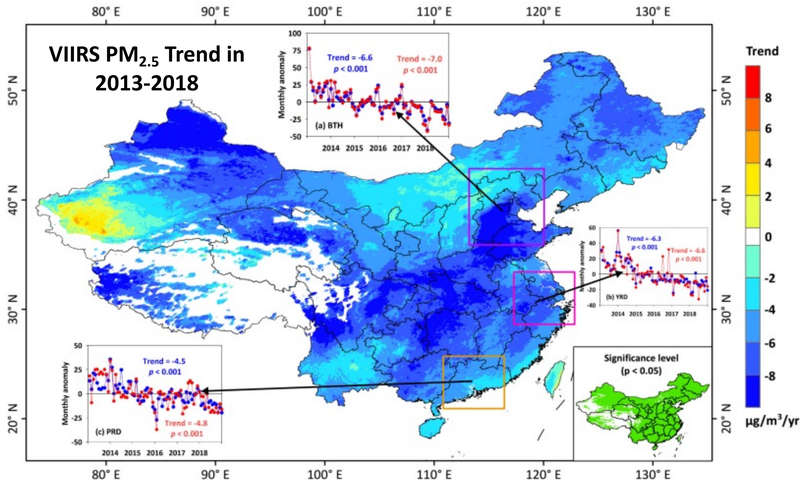
Spatiotemporal variations

- We created and extended long-term $PM_{2.5}$ records dating back to 2000, filling the gap in observations before 2013.
- More than 88% of the country suffers from a $PM_{2.5}$ exposure risk (annual $PM_{2.5}$ > 35 μg/m^3), especially the North China Plain.
- $PM_{2.5}$ pollution has changed greatly and diversely across China during the last two decades.





- In China and all typical regions, $PM_{2.5}$ pollution showed significantly increasing trends since 2000 to around 2007, and remained high until 2013, after which it declined sharply, benefiting from large emission reductions.



- VIIRS can also well capture the temporal trends in $PM_{2.5}$ pollution and could be adopted to extend the EOS long-term $PM_{2.5}$ data records using ML to the next few decades in the post MODIS/MISR era.

Reference

- Wei, J., Li, Z.*, et al. Reconstructing 1-km-resolution high-quality $PM_{2.5}$ data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. *Remote Sensing of Environment*, 2021, 252, 112136.
- Wei, J.*, Li, Z.*, et al. Extending the EOS long-term $PM_{2.5}$ data records since 2013 in China: application to the VIIRS Deep Blue aerosol products. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60, 4100412.