

MORE ACCURATE PREDICTION OF PM 2.5 FROM MODIS SATELLITE USING TOP-OF-ATMOSPHERE REFLECTANCE

Mridula Mavuri[1], Subhajit Chakrabarty[1], Devesh Sarda[1], Udaysinh Rathod[1], Tirthankar Chakraborty[2]
Louisiana State University Shreveport
Pacific Northwest National Laboratory`

Corresponding Author: Mridula Mavuri
Email: mridulam71@lsus.edu
doi: 10.34107/UDUK9890419

ABSTRACT

Air pollution, particularly the concentration of fine particulate matter (PM 2.5), remains a pressing global health concern due to its strong links to cardiovascular and respiratory diseases. Traditional satellite-based PM 2.5 estimation approaches rely heavily on Aerosol Optical Depth (AOD), an intermediate optical parameter that introduces uncertainty due to retrieval errors, cloud interference, and atmospheric correction assumptions. Recent research has shown promise in directly using Top-of-Atmosphere Reflectance (TOAR) from satellite sensors as a more stable and globally available alternative. Building on this emerging direction, this study presents a Transformer-based deep learning framework that directly forecasts daily PM 2.5 concentrations from TOAR and meteorological variables, entirely bypassing AOD. The objective is to develop a globally adaptable and temporally robust model capable of forecasting PM 2.5 levels over a seven-day horizon, with transfer learning as a key contribution enabling cross-continental generalization.. The model was trained on daily data from 1,012 U.S. Environmental Protection Agency (EPA) ground stations for the year 2023, using seven MODIS TOAR spectral bands and six ERA5-Land meteorological variables. It was subsequently fine-tuned using data from 902 European Environmental Agency (EEA) stations. A custom Transformer architecture was designed to process 30-day sequences of multivariate inputs and produce 7-day PM 2.5 forecasts. On the U.S. test set, the model achieved a mean absolute error (MAE) of 1.338 μg/m³ and an R^2 score of 0.9454, consistently maintaining R^2 values above 0.94 on Day 1 and a stable score of 0.9287 on Day 7. When adapted to European data via transfer learning, the model achieved an overall R^2 of 0.9222 and an MAE of just 0.205 μg/m³, demonstrating strong generalization across continents with minimal retraining. These results significantly outperform several baseline models tested in prior studies, including LSTM, Deep Belief Networks, and XGBoost, in both temporal resilience and geographic transferability. Our findings show that Transformer architecture is well-suited to capturing both short-term fluctuations and long-term trends in PM 2.5 driven by meteorology and atmospheric transport. The use of transfer learning notably reduces the dependence on region-specific data, making this approach highly viable for deployment in regions with limited ground-monitoring infrastructure. By eliminating the dependency on AOD, relying solely on globally available TOAR and meteorological inputs, and demonstrating strong cross-regional performance, this study introduces a scalable and adaptable solution for real-time PM 2.5 forecasting. The proposed framework further supports near-real-time operational deployment, holding promise for global air quality surveillance, early warning systems, public health decision-making, and environmental policy, especially in data-scarce or underserved regions where predictive tools are most urgently needed.

Keywords: Particulate Matter, Geospatial, Satellite, Top-of-Atmosphere Reflectance, Air Pollution

INTRODUCTION

Air pollution remains one of the leading environmental threats to public health and global sustainability. Among the many airborne pollutants, fine particulate matter (PM 2.5) defined as particulate matter with an aerodynamic diameter of less than 2.5 micrometers has been extensively linked to severe health outcomes, including cardiovascular and respiratory diseases, reduced lung function, adverse birth