EGU26-7477, updated on 13 Jun 2026
https://doi.org/10.5194/egusphere-egu26-7477
EGU General Assembly 2026

Informing Regional-Scale Air Quality Monitoring through Multimodal Data Integration and Probabilistic Spatiotemporal Modeling

Michelle Hummel and Byeongseong Choi
Department of Civil Engineering, University of Texas at Arlington, Arlington, Texas, United States of America

Low-cost sensors (LCS) are increasingly deployed to enhance the spatial resolution of air quality monitoring networks, particularly in communities that are sparsely covered by regulatory-grade instruments. Despite the potential of LCS to provide new insights into local-level pollutant concentrations, issues with accuracy, calibration drift, and maintenance requirements pose significant challenges for ensuring that collected data reliably support scientific analysis and decision-making. Data fusion techniques that integrate LCS data with other observations, such as regulatory-grade ground measurements and remotely sensed satellite data, show promise for improving the accuracy and resolution of air quality predictions. However, standardized approaches for data fusion are not yet available, limiting the use of multimodal data for decision-making.

Here, we present a community-scale air quality monitoring framework that integrates data from LCS networks, regulatory-grade monitoring stations, and satellite-derived observations through a stochastic advection–diffusion (SAD) modeling approach for fine particulate matter ($PM_{2.5}$). The proposed framework leverages multimodal spatiotemporal data to generate high-resolution $PM_{2.5}$ fields while explicitly accounting for uncertainty arising from sparse observations, sensor heterogeneity, and measurement error within a probabilistic state-space formulation. We apply the framework for a statewide case study in Texas, USA, using archived LCS observations from the PurpleAir network, regulatory-grade $PM_{2.5}$ measurements from the U.S. EPA Air Quality System (AQS), and satellite-derived $PM_{2.5}$ products from NASA's Tropospheric Emissions: Monitoring of Pollution (TEMPO) mission (early-release $PM_{2.5}$ products). We examine how different combinations of data sources contribute to predictive performance, providing insight into the relative value of low-cost, regulatory, and satellite observations within the integrated framework. The results demonstrate the value of combining LCS and satellite data within a physics-informed probabilistic framework to support community-scale air quality assessment, sensor network design, and adaptive environmental decision-making.