93. Multisource Data Integration for Spatial Prediction in Air Quality Monitoring

Carmine Laudato⋆, Monica Sebillo, Pietro Battistoni, and Daniele Sofia,3

1 LabGIS, University of Salerno, Fisciano, Italy 2 Sense Square S.r.l., Corso Garibaldi 33, 84123 Salerno (SA), Italy 3 Department of Computer Engineering, Modeling, Electronics and Systems, University of Calabria, Via P. Bucci, cubo 39/c, Rende (CS), 87036, Italy

Continuous air-quality monitoring and spatial reconstruction provide a basis for future pollution early warning services. Integrating regulatory measurements and low-cost observations with meteorological variables and atmospheric-model outputs can support these tasks, while accounting for data quality and uncertainty. The research conducted at the LabGIS laboratory, University of Salerno, is oriented towards intelligent Digital Twins for complex environmental and urban systems. A key challenge is to make heterogeneous, multi-scale information flows interoperable and coherent while preserving data quality, traceability, and uncertainty throughout the data life cycle. Within this framework, the paper describes a multisource geospatial pipeline for daily spatial reconstruction of particulate matter with aerodynamic diameters up to 2.5 and 10 µm (PM2.5 and PM10) in the municipality of Milan. The pipeline integrates regional environmental protection agency (ARPA) measurements and Sensy sensor data with meteorological variables and Copernicus atmospheric-model outputs, accounting for different spatial and temporal support and preserving source provenance. Integrated observations and derived features support reconstruction at unobserved locations through machine learning models, including Random Forest. Evaluation combines spatial exclusion and temporal separation, baseline comparisons, and feature ablations. An independent regulatory replication in Campania and an internal consistency assessment of the University of Salerno campus sensor network extend the evaluation. Future work will address forecasting and strengthen the integration of physical and data-driven approaches, supporting hybrid, interpretable solutions through geovisual analytics and reproducible assessment protocols.