Machine learning methods in spatio-temporal prediction of atmospheric pollution.

Ewa Skubalska-Rafajłowicz,

Faculty of Information and Communication Technology,

Wroclaw University of Science and Technology, Poland.

Abstract:

Spatially dense maps of air pollution at a specific moment in time can be generated using physical models, such as ADMS-Urban. Unfortunately, due to high computational cost, spatially dense maps are computed for selected, distant moments in time and then verified with low temporal resolution (e.g., monthly or yearly averages). Meanwhile, in densely populated urban areas, information on pollution levels at short, hourly intervals is desirable. In this lecture, we will demonstrate the feasibility of using machine learning models to generate air pollution maps from sparse historical data and current measurements from a limited number of monitoring stations. Such spatial data are highly correlated. Particularly strong correlations typically occur between closely spaced grid points. This allows for a reduction in the output dimension of models using principal component analysis (PCA). We will present neural network models trained based on measurements from a small number of monitoring stations and coefficient values in a reduced principal component space. We will present various methods for evaluating space-time models as single- and multi-output models, accounting for time dependencies at individual grid points or only for spatial dependencies with respect to selected time points. The models will be evaluated using real-world data from a central district in Wrocław, Poland, considering various pollutants such as PM2.5 and NOx, presented on spatially dense pollution maps with ca. 2 × 10^5 points located at a 10 x 10 m grid.