Input to MLs

Stockholm air quality forecast system

Local meteorological forecast via API SMHI & Met Norway

Meteorological diagnostic wind model (Airviro)

Local emission data base and non-exhaust road dust model (NORTRIP)

Land-use and topographic data

Urban Gaussian dispersion model (Airviro)

Street Canyon model (Airviro, OSPM)

CAMS Regional AQ Forecasts Via API

Deterministic forecast of PM$_{10}$, NO$_x$, O$_3$

Meteorological variables:
wind speed, wind direction, relative humidity, cloud cover, precipitation, temperature, pressure, boundary layer height (+ sine and cosine of cyclic variables)

Calendar variables:
hour of day, day of week, Julian day, year (+ sine and cosine of cyclic variables)

Measurements+ :

NO$_x$, PM$_{10}$, O$_3$

Autocorrelation:
lagged of NO$_x$, PM$_{10}$, O$_3$

Feature Engineering:
mean/max/.../Q3/min of concentration from the rolling historical period

4 separate training data sets:
1 urban background and 3 street canyon sites

Random forest   XGBoost   LSTM

GAM Ensemble model

4 test data sets 1-day forecast   **4 test data sets 2-day forecast**   **4 test data sets 3-day forecast**

Random forest 1-, 2-, 3-day   XGBoost 1-, 2-, 3-day   LSTM 1-, 2-, 3-day   GAM 1-, 2-, 3-day

Evaluation of 1-day, 2-day, 3-day hourly forecasts at 4 sites
PM$_{10}$, NO$_x$ and O$_3$
Comparison between deterministic forecasts & 4 ML forecasts
R^2, MAPE, nRMSE, nMAE, MSE, Pearson, MQI
Method of feature importance ranking
MDI, Permutation, SHAP, Gradient-based method