Integrating GIS and Machine Learning for Urban Air Quality Assessment

Mehri Davtalab

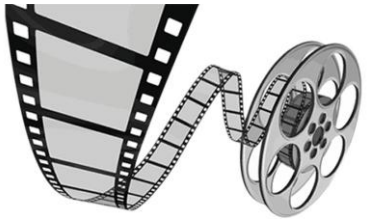
Researcher/ Postdoc fellows at FTMC, Vilnius, Lithuania

Warsaw, 10/17/2025


CMENTARZ POLSKI


366
Ś. P.
KAZIMIERZ
TWARZOWSKI
*1935 +1942
R.I.P.
365
Ś. P.
DANUTA
TARASIUK
*1941


KHOSROW SINAI'S
THE LOST
REQUIEM
BELIEVE THAT FROM ASHES OF DEATH THE SEED OF
LIFE IS BORN



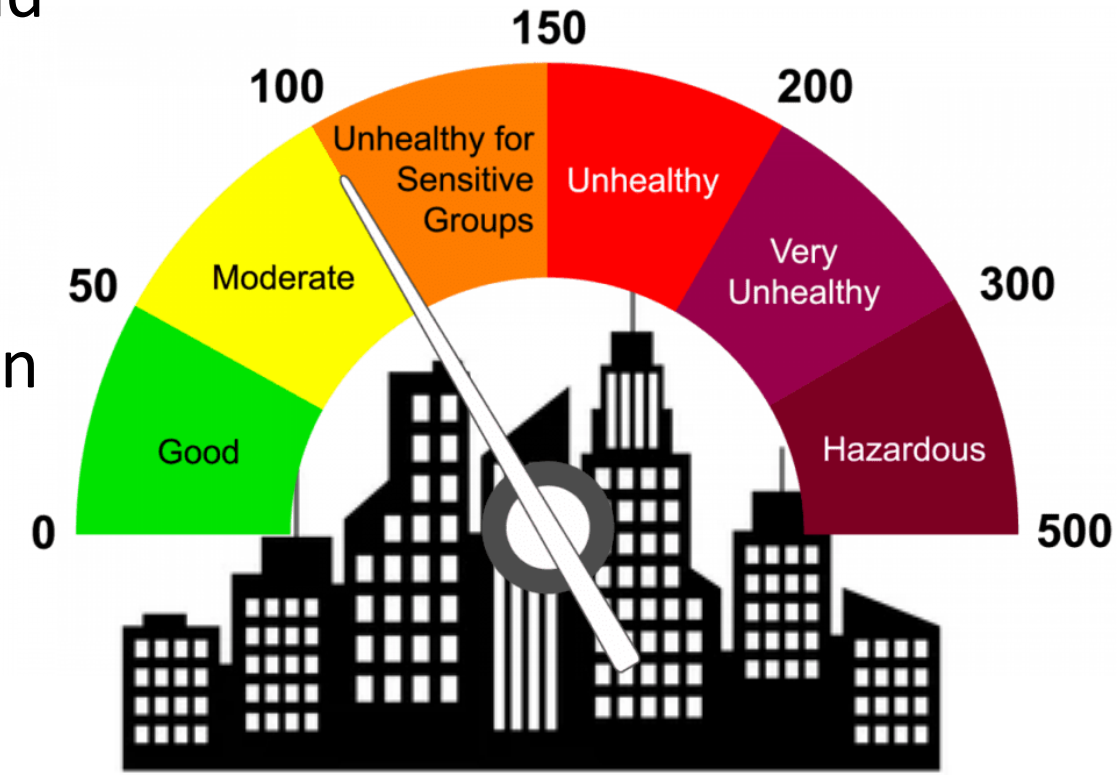
Introduction

- Urban areas occupy only 3% of the Earth's land but accommodate over 50% of the global population.
- The dense population intensify resource consumption, energy use, and waste production.
- Urban air quality assessment is essential to ensure sustainable urban development and protect human health.

Introduction

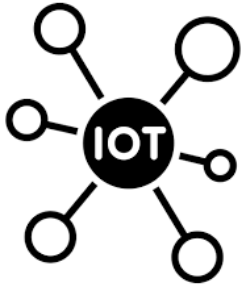
- Urbanization reshapes urban structure, altering land use, population density, and transportation networks.

- These structural changes influence urban air quality and amplify environmental pressures.



What is GIS?

- 80% of all data has a spatial component
- A system for **capturing, storing, analyzing and visualizing** spatial (location-based) data



Integrating GIS and ML

- Understanding how urban structure influences air pollution can be enhanced using Geographic Information System (GIS) for spatial analysis and Machine Learning (ML) for uncovering complex patterns

1 **Effects of spatial urban form on $PM_{2.5}$ concentration**

2 **Multi-pollutant air quality assessment around urban schools using machine learning**

1 Effects of spatial urban form on $PM_{2.5}$ concentration

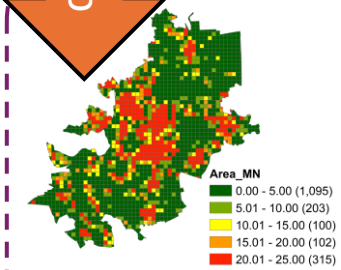
- Green space alone doesn't guarantee cleaner air. What really makes a difference is green space **configuration**.

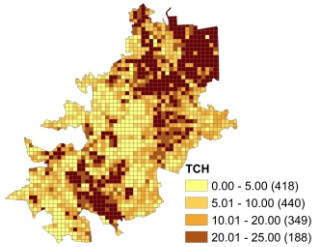
- Built-up land aggregation, shape, and proportion significantly influence $PM_{2.5}$ concentrations.

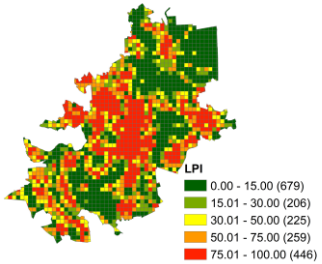
1 Effects of spatial urban form on $PM_{2.5}$ concentration

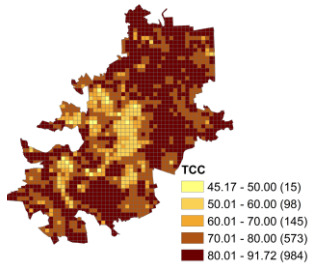
Built-up

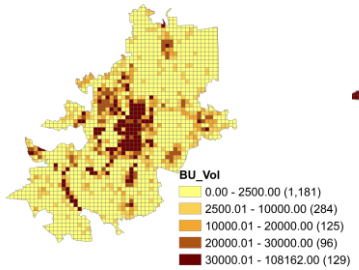
Greenery

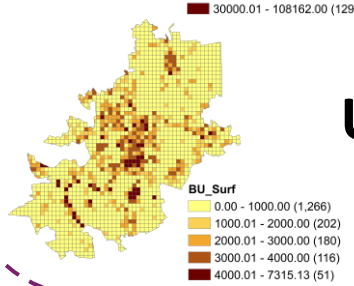




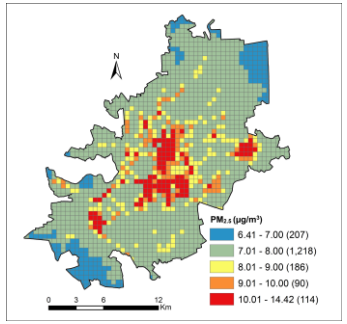


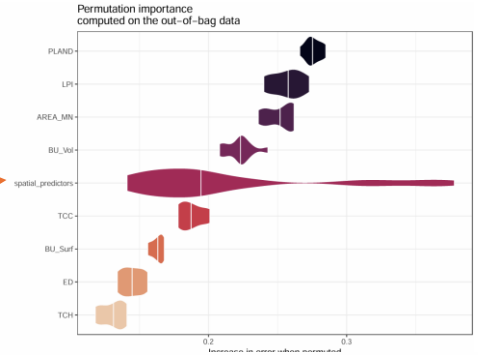




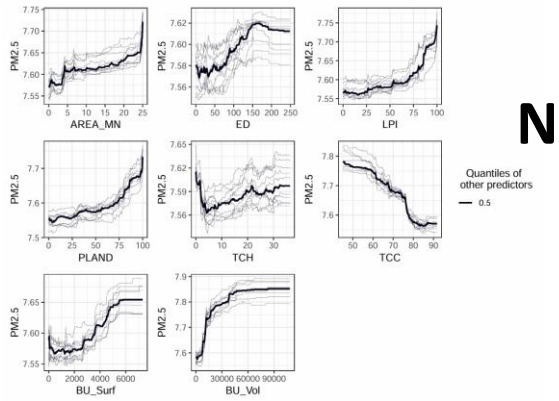


Urban form

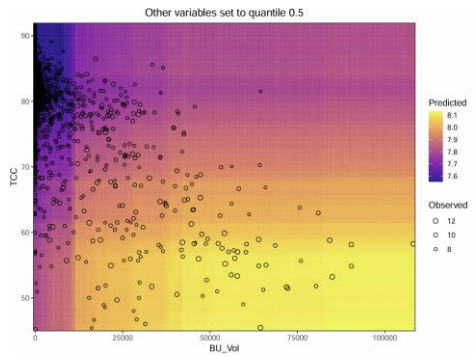




Ranking of factors



Non-linear relationship



Interaction effects

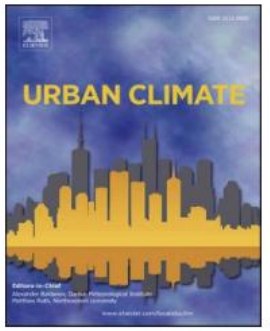
2

Urban Climate 62 (2025) 102567

Contents lists available at ScienceDirect

Urban Climate

journal homepage: www.elsevier.com/locate/uclim



Check for updates

Multi-pollutant air quality assessment around urban schools using machine learning

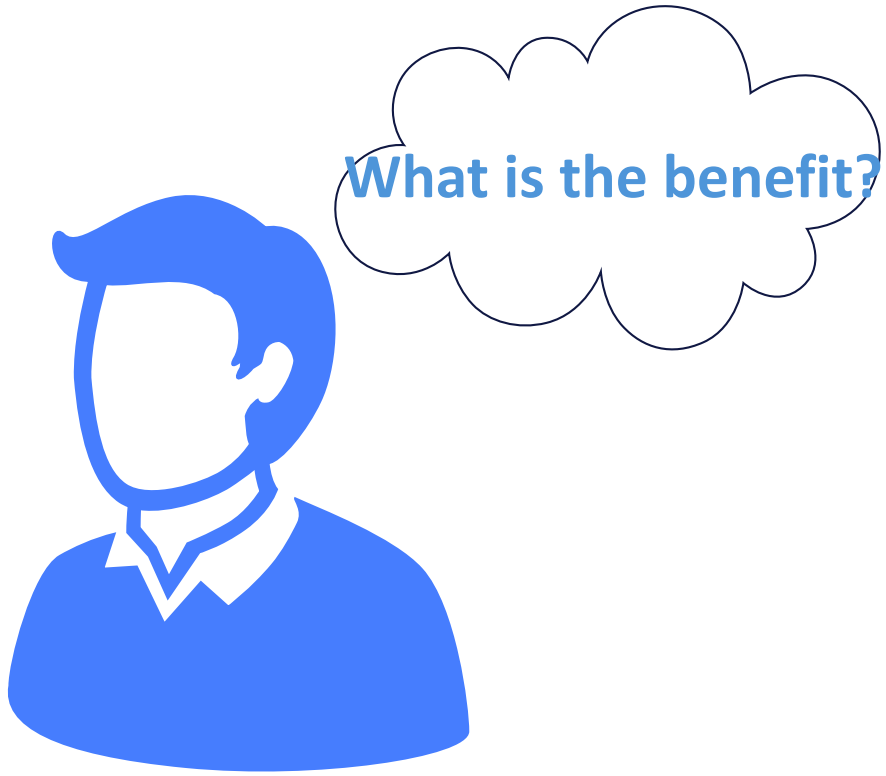
Mehri Davtalab [a,*], Lina Davulienė [a], Ieva Uogintė [a], Simonas Kecorius [a,b], Mario Lovrić [c,d], Steigvilė Byčenkienė [a]

[a] Center for Physical Sciences and Technology (FTMC), LT-10257 Vilnius, Lithuania
[b] Institute of Epidemiology, Helmholtz Zentrum München, 85764 Neuherberg, Germany
[c] The Lisbon Council, 1040 Brussels, Belgium
[d] Institute for Anthropological Research, 10000 Zagreb, Croatia

Scan me

2 The aim

To develop a model that can simultaneously estimate the spatial distribution of multiple pollutants.



To provide policymakers with a tool to identify areas with elevated pollution in cities with limited monitoring stations.

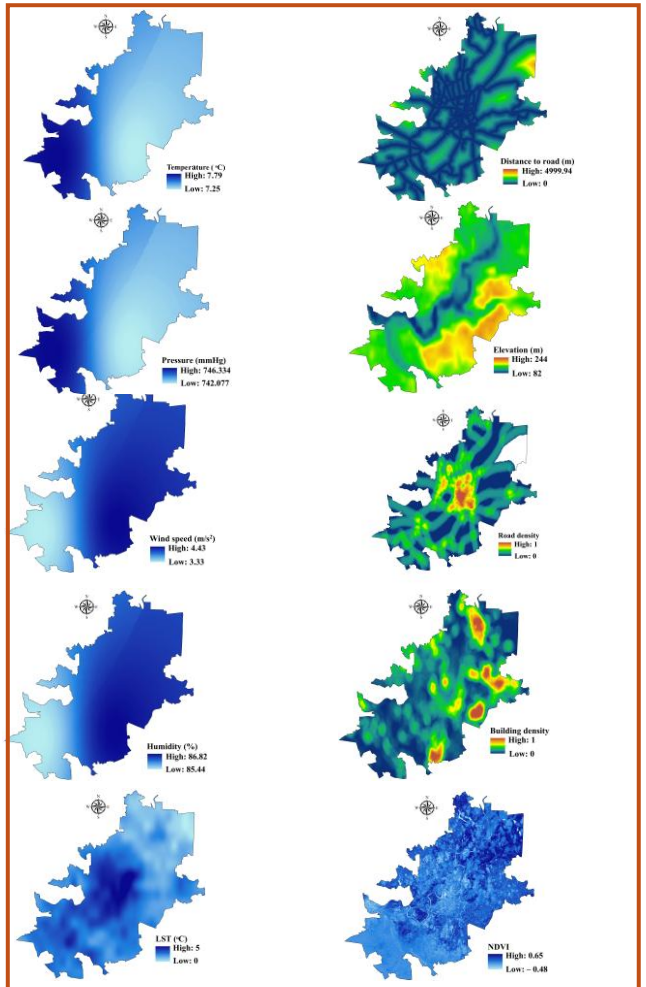
2 The problem

- **Limited Ground Stations:** Many cities have sparse monitoring stations, so air quality cannot be assessed across the entire city.

- **Sensitive Locations:** The lack of coverage is particularly concerning near sensitive locations, such as schools

- **Satellite Solutions:** Satellite data can cover the whole city and help solve scale issues, but most studies focus on one pollutant or build separate models for multi-pollutants.

2 Our model

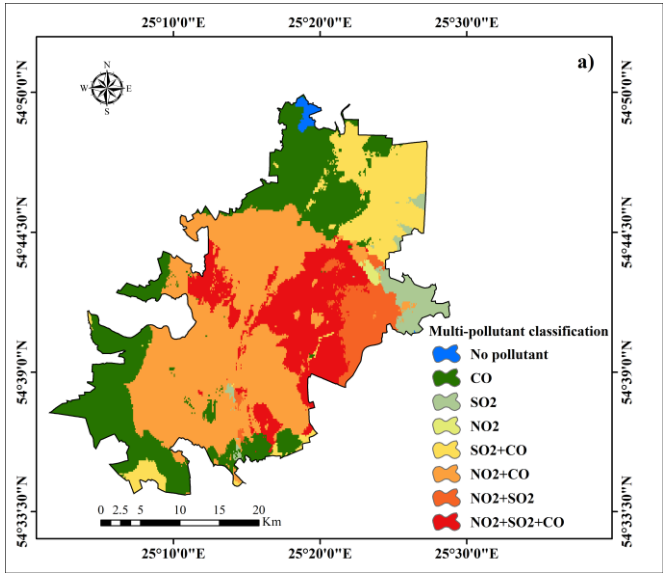


Urban form & meteorological data

CO

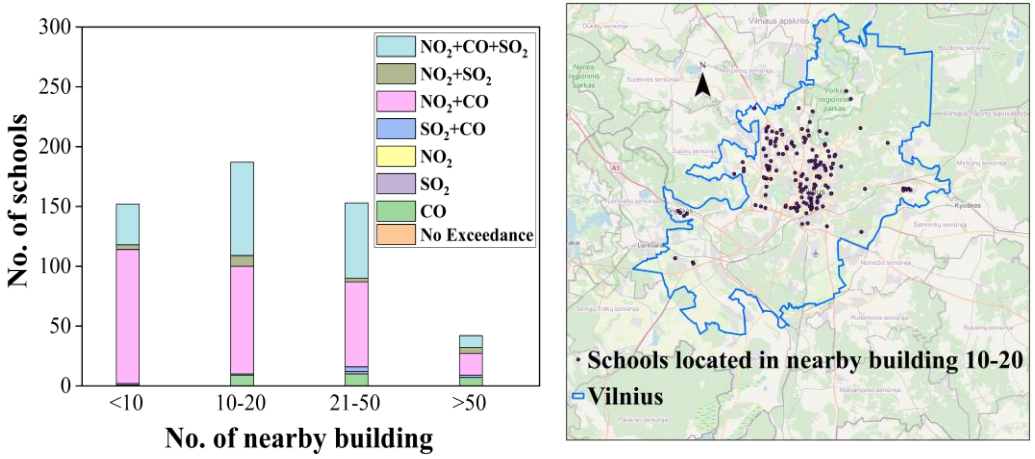
SO2

NO2

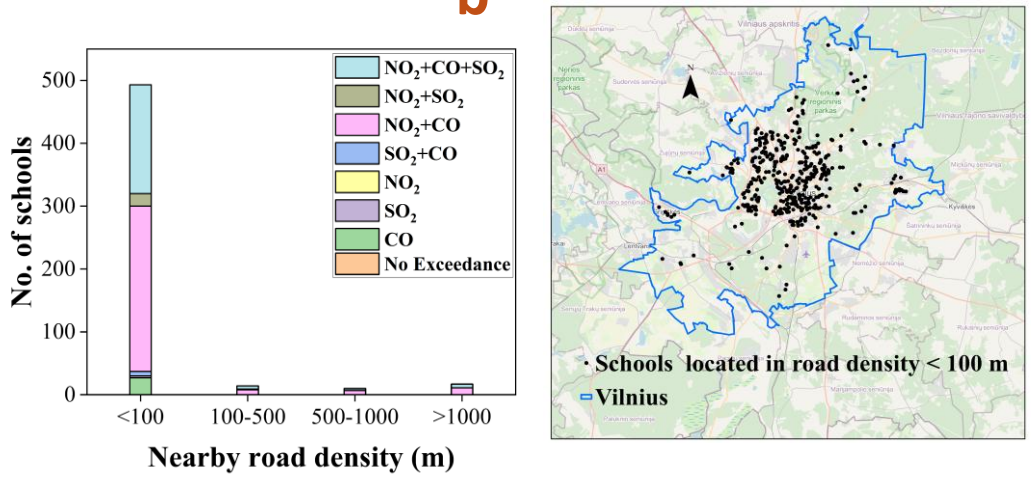


2 Spatial relationship between air quality around urban schools and urban form

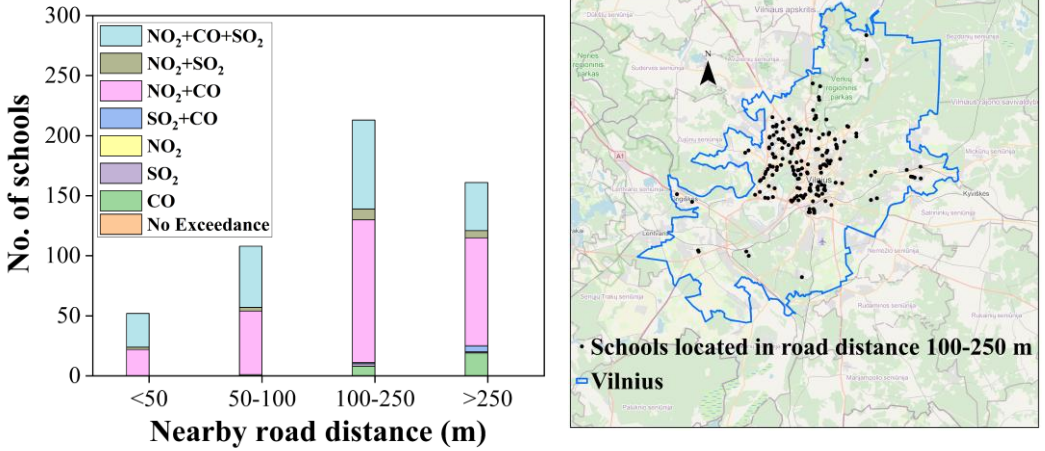
a



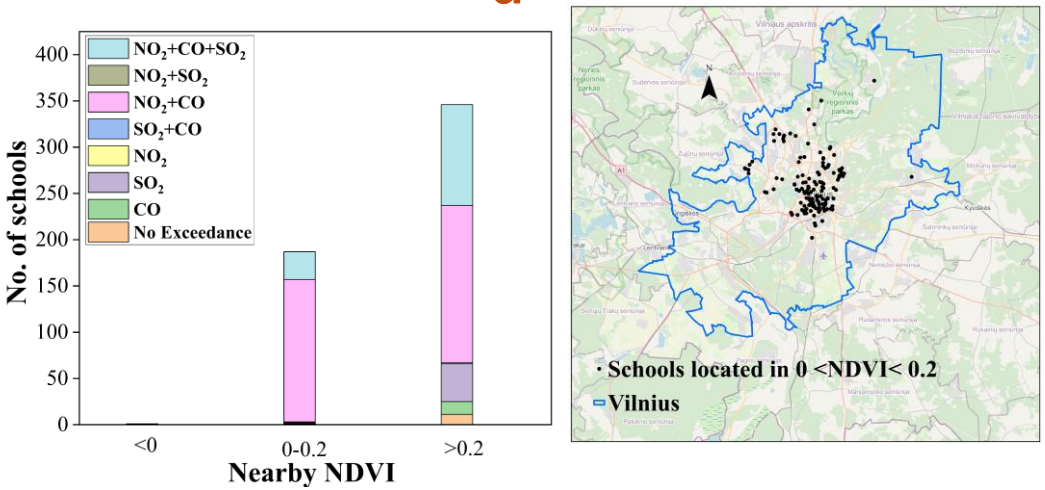
b



C



d



To summarize:

1. GIS provides spatial data integration and visualization, while ML enhances prediction accuracy and factor ranking. **Spatial Random Forest** models non-linear relationships, interaction effects, ranks influential predictors, and accounts for spatial dependences.

2. This study demonstrates the effectiveness of machine learning–based approaches for assessing air quality around urban schools in Vilnius with **limited ground monitoring stations**. By integrating satellite-derived pollutant data with urban form and meteorological variables, the Random Forest model shows superior performance in classifying **multi-pollutant** distributions.

Acknowledgements

Dr. Steigvilė Byčenkienė

LMT
Lietuvos mokslo taryba

EDIAQI
Evidence Driven Indoor Air Quality Improvement

Thank you!