Examining Variable Significance in Air Quality Across Cities of Different Urban Form in the USA

Pierre-Antoine Lequeu

Artificial Intelligence in Urban Planning and Design 2

Research Idea

As the air quality is becoming a main concern for people, many studies focus on highlighting the major causes of bad air quality. However, the majority focuses on metropolitan cities and draw conclusion for this type of city only.

Spatio-temporal prediction and factor identification of urban air quality using support vector machine

Chih-Chun Liu, Tzu-Chi Lin, Kuang-Yu Yuan, Pei-Te Chiueh, 2022

Air pollution assessment in Seoul, South Korea, using an updated daily air quality index

Hyunjung Lee, Jinhyo Lee, Seokryul Oh, Sookuk Park, Helmut Mayer, 2022

Research question

A limited number of studies have explored the variations in the significance of air quality variables based on city size. Nonetheless, it is essential to acknowledge that daily travel patterns and sources of pollution are significantly different depending on the city urban form (population, area…)

- Therefore, how does the variable significance in air quality changes in different urban form ?
- On what urban form should urban planner focus when developing a city ?

Previous Work

Urban air quality in a mid-size city — $PM_{2.5}$ composition, sources and identification of impact areas: From local to long range contributions

Stefania Squizzato, Marta Cazzaro, Elena Innocente, Flavia Visin, Philip K. Hopke, Giancarlo Rampazzo, 2017

« *Urban air quality represents a major public health burden and is a long-standing concern to European citizens. Combustion processes and traffic-related emissions represent the main primary particulate matter (PM) sources in urban areas* »

Urban and air pollution: a multi-city study of long-term effects of urban landscape patterns on air quality trends

Lu Liang and Peng Gong, 2020

« *This empirical study identifies the key urban form determinants of decadal-long fine particulate matter ($PM_{2.5}$) trends in all 626 Chinese cities at the county level and above* »

Dataset

- **DEAP: Deciphering Environmental Air Pollution**

https://www.kaggle.com/datasets/mayukh18/deap-deciphering-environmental-air-pollution/data

Large-scale spatio-temporal dataset giving daily pollutant concentration, traffic, pollution from power plants and meteorological factors in 50+ cities in the USA over a two-year period

Include cities like Los-Angeles (3.8M), Chicago (2.7M), Houston (2.2M) but also little rock, AR (200k) or boise, ID (230k)

Deciphering Environmental Air Pollution with Large Scale City Data

Mayukh Bhattacharyya, Sayan Nag and Udita Ghosh

https://www.ijcai.org/proceedings/2022/0698.pdf

Methodology

1. Multiple models have proven effective on air quality work : SVM, GBM, LSTM. First, try these models on the dataset to determine the most effective on Air Quality Prediction
2. Apply it to different subdataset with cities of the same urban form
3. Compare the results for different datasets

Expected Results

Partial Dependence Plot and Feature Importance for the different subdatasets.

Example of expected result :

Traffic might have a very different importance depending on city density

Small cities might be more affected by powerplants

Thank you for your attention