

Proposal on the Influence of Emissions and Vegetation Coverage on Air Quality in Major Chinese Cities

Project Members

ZENG, Zhi Qi	1155215115
JIANG, Wen Ming	1155215099
QIN, Lang	1155215108

April 12, 2024

Motivation

Air quality is a globally recognized priority concern. Air pollution not only affects the economy but can have harmful effects on people's lives. For example, air pollution can cause acid rain and haze, which can lead to water and soil hazards. Air pollution can lead to the depletion of the ozone layer, which can lead to increased exposure to ultraviolet radiation and an increased risk of skin cancer [2]. Reduced air quality affects people's health, which in turn affects healthcare costs and the productivity of the labour force.

In the decade from 2004 to 2014, China's air pollution control policies have achieved decent results. Therefore, the study of factors affecting air quality is conducive to better formulation of government policies on optimizing air quality, which is conducive to economic growth, environmental protection as well as human health and safety.

Problem Statement

In China, many cities have banned fireworks and firecrackers in consideration of safety issues and air pollution. In order to solve the problem of air pollution, the government vigorously promotes new energy vehicles. The air quality problem seems to be taken more and more seriously. Much of the air quality problem is down to factors such as energy consumption, exhaust emissions and the degree of greening.

Studies have shown that green areas in urban parks have a positive impact on air pollution control, while industrial and vehicular emissions have a negative effect on air pollution [1].

Therefore, we hope to find out the factors that most affect air quality by analyzing the data from the perspective of the influencing factors of air quality problems, so as to provide some help for urban planning as well as policy making.

Data Source

In order to analyze the relevant influencing factors of the AQI, we have selected the AQI data of three cities as samples, namely Beijing, Chongqing and Shanghai [3].

We obtained data on industrial, vehicle and domestic emissions in various Chinese cities from

2004 to 2014 from the official website of the National Bureau of Statistics of China [4].

In addition, we obtained data on the greenness and vegetation coverage of the three cities [5].

Data Processing Approaches

This project aims to examine the listed factors that would influence the AQI of a certain city in China and find out in what extent do these factors change the AQI.

We purpose to reprocess the raw dataset at first. At the very beginning we would use data visualization to show the relation between AQI and its factor. When the effects of these factors are determined, we begin to prepare these statistics by detecting and handling outliers and further nondimensionalization.

After that we would adopt supervised learning to build models, trying to find out the weighing of these variables to influence AQI. Meanwhile, we would use several algorithms to discover the most accurate model by comparing their deviation.

After the model-building process, data visualization would be used again to show the level of correlation between these variables and AQI.

Implementation / Experimentation

In visualize the raw data, we intend to use time series plot to see how each dependent and independent variables varies over the period and potentially relates to the others. Then, to nondimensionalize these independent variables, we consider using normalization to map these statistics into a range from 0 to 1.

During the model-building process, we consider adopting libraries in Python, for instance, scikit-learn, pandas, numpy etc. and implement algorithms including multiple linear regression analysis, PLS regression and random forest to generate the weighing of different variables and calculate the MSE of these outputs to select the most accurate model.

At last, a correlation matrix would be used to identifying which factors might influence AQI more significantly.

Expected Conclusions

Through utilizing those methods based on database of 2004-2014, we expect to derive a function shows the relative weighting of each factor (industrial, vehicle and domestic emissions, vegetarian coverage). Then we could anticipate identifying key drivers of air

pollution and quantifying their respective impacts on air quality levels. Studying how these factors individually and collectively influence air quality can provide a scientific basis for policymaking, fostering environmental improvement and sustainable development.

We hope this research will not only give us normal people a deeper insight, but also empower policymakers, urban planners, and environmental agencies to formulate evidence-based interventions and prioritize tailored strategies aimed at mitigating air pollution and enhancing the overall environmental sustainability and public health in Chinese cities.

References

- [1] Hongshan Ai, Xi Zhang, Zhengqing Zhou (2023). The impact of greenspace on air pollution: Empirical evidence from China.
<https://doi.org/10.1016/j.ecolind.2023.109881>
- [2] Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and health impacts of air pollution: a review. *Frontiers in public health*, 8, 505570. <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2020.00014/full>
- [3] World Air Quality Historical Database city: Beijing, Chongqing and Shanghai.
<https://aqicn.org/station/beijing/> <https://aqicn.org/historical/#city:chongqing>
<https://aqicn.org/historical/#city:shanghai>
- [4] Waste gas emissions in the East, Central and Western regions (2004 - 2014). National Bureau of Statistics of China. <https://www.stats.gov.cn/>
- [5] Area of green space in Chinese gardens (2004 - 2014). CEIC data.
<https://www.ceicdata.com/zh-hans/china/area-of-garden-and-green-prefecture-level-city/area-of-garden--green-beijing>