Measuring influential factors for air pollution in New-York-Newark-Jersey City by grey relation analysis

Baoxin Chen, Qin Lu, Tao Zhang*
School of Humanities and Social Sciences, Macao Polytechnic Institute, Macao, China
*Corresponding author: taozhang@ipm.edu.mo

Abstract. This study investigates variations of air quality index (AQI) for air pollutants in the metropolitan area (New-York-Newark-Jersey City; NY-NJ-PA) during 2010-2019 to identify the key influential factors for air pollution using grey relation analysis method. The results show that the annual averages of daily AQI values for multiple pollutants in NY-NJ-PA generally decreased except O3. The AQI values of O3 fluctuate greatly. The main factors affecting AQI values of CO, PM2.5, O3, SO2, and PM10 are the emissions of CO2, SO2, and NOx, whereas the main factors affecting AQI values of NO are the emissions of SO2, population, and NOx.

1. Introduction
While economic growth has enabled the development of human society and the improvement of people’s living standard, it also induces the problems of air pollution and energy security. As U.S. government officially returns back to the Paris climate agreement in 2021[1], it is not only a commitment to global environmental protection, but a promise to the health of American citizens through policy formulation. Therefore, it is important to monitor and investigate the variations of air pollutants. This study is going to examine the variations of air quality index (AQI) for air pollutants in the metropolitan area (New-York-Newark-Jersey City) to identify the key influential factors for air pollution by grey relation analysis method.

The main indicators of air pollutants for usual monitoring include particulate matter (PM2.5 / PM10), CO, SO2, NO2, and O3. The exposure of air pollutants is mainly directly related to the combustion of petroleum and other liquid fuels, natural gas, coal, industrial emissions, and vehicle emissions. Among them, O3 is not emitted directly, it is formed by the optical reaction of atmospheric antioxidants and hydrocarbons.

It is widely known that air pollutants can damage the lungs, airways and cardiovascular, causing cardiovascular and respiratory diseases. Moreover, they can induce heart attacks, strokes, heart failure as well as respiratory failure. It also increases the risk of morbidity and death from diseases such as diabetes and tumors. It is reported that about 500,000 lung cancer deaths and 1.6 million COPD deaths can be attributed to air pollution, which may also account for 19% of all cardiovascular deaths and 21% of all stroke deaths [2]. It is shown that air pollution is closely related to a large number of diseases, such as bone diseases, skin and immune diseases, cognitive function and neurological diseases, conjunctivitis, dry eye, blepharitis, inflammatory bowel disease, intravascular coagulation. It must be noted that CO is an odorless, colorless and toxic gas. Long-term exposure to low concentrations of CO
can cause chronic poisoning reactions, and high concentrations of CO can be fatal. PM2.5 is related to psychomotor developmental delay [3, 4].

Although air pollutants may damage people’s health greatly, it could be monitored and controlled. The EPA has established the American air quality standards by dividing the AQI into six levels for public health. “0 to 50” indicates that the air quality level is "good", the air quality is satisfactory, and air pollution is almost without risk; "51 to 100" represents the air quality level at "medium", the air quality is acceptable, but some people may be at risk especially those extremely sensitive to air pollution; "101 to 150" means that the air quality level is "unhealthy for sensitive people", members of sensitive groups may experience undesirable health effects, but ordinary people are less likely to be affected; "151 to 200" means that the air quality level is "bad", some of ordinary people may experience undesirable health effects, and members of sensitive groups may experience more serious effects; AQI of "201 to 300" represents the air quality level at "very unhealthy", and a health warning to everyone increases: "301 or above" represents the air quality level "dangerous", and everyone is highly likely to be affected [5].

NY-NJ-PA is a cluster of counties with a high degree of economic integration and high population density, reaching the scale of a metropolitan. To provide a basis for policy formulation, according to U.S. Census data, air quality directly affects the health of 19,216,182 people in this area. For the sake of clarifying the variations of air pollution in this area, this study compares the time series air-pollution indicators in the past 10 years, and calculates the grey correlation of the data series based on the geometric relationship of the data, in order to find the influential relationship between the main factors.

2. Grey relation analysis

Grey relation analysis (GRA) is a method of multi-attribute decision-making analysis based on grey relation theory. GRA compares the array of related factors and only requires a small sample size. It can quantify the dynamic processes and determine the degree of correlation between factors. It has been widely applied in the socio-economic system [6]. The calculation steps of GRA include selecting analysis data, data transformation, and calculating correlation coefficient and correlation degree [7].

2.1. Data source and data processing

2.1.1. Data source. This study selected the six major influential factors potentially affecting the air quality indicators in New-York-Newark-Jersey City from 2010 to 2019, including Population, Real GDP (Thousands of chained 2012 dollars), Personal income (thousands of dollars), CO2 emissions (million metric tons), SO2 emissions (million metric tons), and NOx emissions (million metric tons). The data come from Independent Statistics and Analysis of the U.S. Energy Information Administration (EIA) [7] and Bureau of Economic Analysis of the U.S. Department of Commerce (BEA) [8].

2.1.2. Data transformation. The purpose of this study is to eliminate dimensional differences and transform data into comparable sequences. The process of data standardization has to be done firstly, based on obtaining the average and standard deviation of each sequence. And the original data are subtracted by the average. And then the subtracted data are divided by the standard deviation. Finally, a new standardized data series with a normalized distribution is obtained, with a mean value at 0 and a variance at 1.

2.2. Calculating the correlation coefficient and degree of correlation

2.2.1. Calculation formula of correlation coefficient $\zeta_i(k)$.

$$\zeta_i(k) = \frac{A \min + \rho A \max}{A \Delta i(k) + \rho A \max}$$  (1)
\( \Delta 0t(k) \) represents the absolute difference between two comparison sequences at time \( k \), \( \Delta \text{min} \) represents the minimum value of the absolute difference between the two comparison sequences at each time, \( \Delta \text{max} \) represents the maximum value of the absolute difference between the two comparison sequences at each time. When compared sequences intersect, \( \Delta \text{min} = 0 \). \( \rho \) is an identity coefficient, used to reduce the distortion caused by the extremely large values of the maximum absolute difference and adjust the differences of the results, its value is between \((0, 1)\), in this study \( \rho = 0.1 \).

2.2.2. Relevance calculation and sorting. The correlation of the two sequences is calculated by the average of the correlation coefficients of the two comparison sequences at each moment:

\[
R_w = \frac{1}{N} \sum_{k=1}^{N} \zeta(k)
\]  

(2)

\( N \) represents the number of sequences that need to be compared, and the correlation degree of each factor to the same parent sequence is arranged in descending order to obtain the main factor affecting the air quality index.

3. Result of analyzing air pollution degree and influential factors in NY-NJ-PA

3.1. Annual AQI trends of air pollutants in NY-NJ-PA
In NY-NJ-PA area, the AQI of air pollutants shows a significant downward trend, but the pollutant O3 fluctuates greatly. Ozone is a secondary pollutant, and its concentration is mainly affected by the CO and NOx in the ambient air. As for the concentration of precursors and the influence of photochemical reactions, industrial production will cause the concentration of these precursors to be increased.

![Figure 1. New-York-Newark-Jersey City air pollutant AQI trends](image-url)
3.2. Results of grey correlation factors (2010-2019)

According to Table 1, the main factors affecting air quality indices of CO, PM2.5, O3, SO2, and PM10 are CO2, SO2, and NOx emissions. And the main factors affecting NO API are SO2, population, and NOx emissions.

Table 1. Analysis of the grey correlation factors in the New-York-Newark-Jersey City area.

| Factor                        | CO       | PM2.5   | NO       | O3       | SO2       | PM10     |
|-------------------------------|----------|---------|----------|----------|-----------|----------|
| Population (Thousands of people) | 0.240    | 0.271   | 0.377    | 0.256    | 0.257     | 0.287    |
| Real GDP (Thousands of chained 2012 dollars) | 0.215    | 0.217   | 0.323    | 0.315    | 0.292     | 0.313    |
| Personal income (thousands of dollars) | 0.211    | 0.209   | 0.314    | 0.317    | 0.288     | 0.310    |
| CO2 emissions                 | 0.580    | 0.486   | 0.348    | 0.562    | 0.565     | 0.582    |
| SO2 emissions                 | 0.583    | 0.617   | 0.405    | 0.510    | 0.559     | 0.579    |
| NOx emissions                 | 0.533    | 0.577   | 0.375    | 0.489    | 0.439     | 0.508    |

4. Conclusion

According to the results in this study, the air quality in NY-NJ-PA is at the level of “good” in recent years. It shows that human activities, industrial production, and population will directly affect the air quality. Therefore, it is necessary to control the source of industrial emissions and reduce pollutants emitted by human activities vigorously.

Acknowledgements

This work was financially supported by the Research Project Fund of Macao Polytechnic Institute (Approval No.: RP/ESCHS-04/2020).

References

[1] https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/20/paris-climate-agreement/
[2] Schraufnagel D E, Balmes J R, Cowl C T, et al. Air pollution and noncommunicable diseases: A review by the Forum of International Respiratory Societies’ Environmental Committee, Part 2: Air pollution and organ systems[J]. Chest, 2019, 155(2): 417-426.
[3] Bae S, Kwon H. Current state of research on the risk of morbidity and mortality associated with air pollution in Korea[J]. Yonsei medical journal, 2019, 60(3): 243.
[4] Kim J B, Prunicki M, Haddad F, et al. Cumulative lifetime burden of cardiovascular disease from early exposure to air pollution[J]. Journal of the American Heart Association, 2020, 9(6): e014944.
[5] https://www.airnow.gov/qaqi/qaqi-basics/
[6] Cui L, Chan H K, Zhou Y, et al. Exploring critical factors of green business failure based on Grey Decision Making Trial and Evaluation Laboratory (DEMATEL)[J]. Journal of Business
[7] Lin S J, Lu I J, Lewis C. Grey relation performance correlations among economics, energy use and carbon dioxide emission in Taiwan[J]. Energy Policy, 2007, 35(3): 1948-1955.
[8] https://www.eia.gov/
[9] https://www.bea.gov/