Capítulo 8

Analysis of environmental pollution for Mexico City and New York City during 2006-2018

Como citar:
Vázquez Álvarez, L. *et al.* (2023). Analysis of environmental pollution for Mexico City and New York City during 2006-2018. En V. Meriño *et al*. (Eds.) (2023). Gestión del Conocimiento. Perspectiva Multidisciplinaria, libro 58. (pp. 141-160). Fondo Editorial de la Universidad Nacional Experimental Sur del Lago, Jesús María Semprum. https://doi.org/10.59899/Ges-cono-58-C8

Laura N. Vázquez Álvarez

https://orcid.org/0009-0003-0973-0501

Bachelor's degree in Economic Sciences by the Autonomous University of Ciudad Juárez. Corporate Investigator at Diligent Corporation. This work is based on the thesis dissertation "Análisis de la contaminación ambiental para la Ciudad de México y Nueva York durante 2006-2018" (UACJ, 2020). The authors thank Dr. Patricia Barraza, Ramsés Jiménez, and the audience of *The Colloquium of Economic Research in the University of Cd. Juarez* in November 2020.

Correo: lauranoemi_vazquez@hotmail.com

David Vázquez Guzmán

https://orcid.org/0000-0001-8254-9766

Doctorado en Economía. University of Stirling, Maestría en Economía. University of California – Riverside. Maestría en Administración. ITESM. Ingeniería en Sistemas Computacionales. Instituto Tecnológico de Chihuahua II. Profesor Investigador en Economía en el departamento de Ciencias Sociales en la Universidad Autónoma de Ciudad Juárez.

Correo: david.vazquez@uacj.mx

Ana Luz Ramos Soto
https://orcid.org/0000-0001-8167-2631
Máster en alta dirección por la Organización Continental de Excelencia Educativa. Diplomado Elementos Básicos de Gestión Estratégica para el Desarrollo Local y Regional. Doctorado en Ciencias de Planificación de Empresas y Desarrollo Regional. Maestría en Ciencias en Planificación de Empresas y Desarrollo Regional. Licenciatura en Economía. Profesora Investigadora de Tiempo Completo en la Facultad de Contaduría y Administración en la Universidad Autónoma de Benito Juárez de Oaxaca.
Correo: analuz_606@yahoo.com.mx

Ricardo Alarcón Alcántara
Máster en Administración. Doctor en Ciencias de la Administración. Profesor Investigador de Tiempo Completo. Integrante del Cuerpo Académico Desarrollo y Negocios, Coordinador General de Tutorías y Asesorías de la Facultad de Contaduría y Administración de la Universidad Autónoma "Benito Juárez" de Oaxaca.
Correo: riicardo.alarcon3@gmail.com

Abstract

This study provides a comprehensive report of the main pollutants that affect the cities of New York and Mexico during the period of 2006-2008. It entails the use of principal component analysis and the Air Quality Index to predict the main component of air pollution affecting each city. For Mexico City, the two components explained almost 70% of the data were, Carbon Monoxide and Nitrogen Dioxide. As for New York City, the two components that explained almost 70% of the data were Sulfur Dioxide and Carbon Monoxide. In conclusion, pollution in Mexico City can be largely explained by the presence of vehicles and its geographical location, as for New York City, it was due to its high population density and vehicle presence.

Keywords: environmental economics, air quality index, air pollution, principal component analysis

Análisis de la contaminación ambiental para la Ciudad de México y la Ciudad de Nueva York durante 2006-2018

Resumen

Este estudio proporciona un informe completo de los principales contaminantes que afectan a las ciudades de Nueva York y México durante el periodo 2006-2008. Implica el uso del análisis de componentes principales y el Índice de Calidad del Aire para predecir el componente principal de la contaminación del aire que afecta a cada ciudad. Para la Ciudad de México, los dos componentes que explicaron casi el 70% de los datos fueron el monóxido de carbono y el dióxido de nitrógeno. En cuanto a la ciudad de Nueva York, los dos componentes que explicaron casi el 70% de los datos fueron el dióxido de azufre y el monóxido de carbono. En conclusión, la contaminación en la Ciudad de México se puede explicar en gran medida por la presencia de vehículos y su ubicación geográfica, en tanto que en la Ciudad de Nueva York se debió a su alta densidad poblacional y presencia vehicular.

Palabras clave: economía ambiental, índice de calidad del aire, contaminación del aire, análisis de componentes principales

Introduction

In recent years, economists have been focused on environmental problems and the importance and implications of these issues. In addition, over time, the interaction between the level of development of a country and environmental degradation has been understood (Todaro & Smith, 2015). One of the issues that is usually on the minds of citizens is global warming and the quality of the environment. In terms of global warming, "The Fourth Assessment Report of the Intergovernmental Panel on Climate Change", dispelled many uncertainties about global warming, and established that the warming of the climate system is now unequivocal and a fact that marks reality (Dale, Fant, Strzepek, Lickley, & Solomon, 2017).

When talking about environmental aspects, there is a tendency to categorize according to the area that is polluted, for example, air, water, land, etc. Air quality is one of the main concerns within metropolitan areas and this is why the study will focus on the analysis of air pollution.

Air quality directly affects a person's quality of life and like the weather, it has the potential to change suddenly. This is why there are agencies responsible for disseminating information on air quality. A key tool is the instruments that help in the measurement of environmental conditions. In the case of air pollution, the index that stands out is the Air Quality Index (AQI). The index is used to provide information on local air quality and warn the population about the possible effects that could occur (AirNow, 2017).

In previous studies, there have been authors who relate environmental variables with some economic determinants in order to analyze the relationship they present in greater depth. From studies that implement economic growth and environmental pollution through the use of the Environmental Kuznets Curve (Adu & Denkyirah, 2018); to those who create their own indices to calculate air quality (Bishoi, Prakash & Jain, 2009). That is, environmental analysis has been joining the topics of interest to economists.

This is why the study intends to determine which are the main pollutants in each city, and also to study the dynamic evolution of the Air Quality Index instrument. It is intended to find the causes of pollution by using the Principal Component Analysis (PCA) for each city and thus obtain a better understanding of the behavior of the tool. In addition, the use of other econometric tools will be explored to more fully understand the behavior of the AQI.

Theoretical framework

Review of study approaches related to the environment-economic growth

When comparing the relationship between economic growth and the environment, different ways of approaching the subject arise. Lim, in 1997 determined that economic growth generates pressure on the environment, however, he found how there is also a relationship where economic growth does not affect negatively. In the particular case of South Korea, the study presented by Lim shows how in 1980 there was a change in the rigidity of environmental regulation and policies aimed at caring for the environment became stricter (Lim, 1997).

Meanwhile, there are others who have studied the inverted U theory with more countries. It is argued that industrialization and agricultural modernization may reduce the quality of the environment in the early stages of development, however, this trend will reverse as an economy continues its development process.

Selden and Song compare this relationship to four of the most prominent air pollutants: particulate matter, sulfur dioxide, nitrogen oxides, and carbon monoxide. Per capita emissions of the four pollutants are found to exhibit inverted U relationships with GDP per capita. While this suggests that emissions will decline in the long term, continued rapid growth in global emissions is predicted for decades to come (Selden & Song, 1994).

Principal component analysis as a predictor of the level of air pollution

Econometrics has tools that facilitate the interpretation of extensive databases, one of which is the Principal Component Analysis. Academics from Sultan Zainal Abidin University succeeded in demonstrating the importance of historical data in sampling plan strategies to achieve the desired research objectives, as well as to highlight the possibility of determining the optimal number of sampling parameters, which will reduce costs. and the sampling time (Azid et al., 2014).

In the article titled "Prediction of the Level of Air Pollution Using Principal Component Analysis and Artificial Neural Network Techniques: a Case Study in Malaysia", the researchers focus on recognizing the patterns of air quality in Malaysia. They obtained data from the Department of the Environment (DOE) and managed to collect eight air quality parameters at ten monitoring stations over 7 years (2005-2011) (Azid et al., 2014).
Using the data mentioned above, principal component analysis could be used to identify sources of contamination at the study sites. In conclusion, the PCA managed to identify that CH4, NmHC, THC, O3 and PM10 are the most significant parameters for this case study.

Contextual framework

Mexico City and New York City

The population increase that has occurred in recent years has been related to urbanization and the introduction of globalization. Today, 55% of the world's population lives in urban areas, a proportion that is expected to increase to 68% by 2050. Projections show that urbanization, the gradual change in the residence of the human population from rural areas to Combined with the overall increase in the world's population, it could add another 2.5 billion people to urban areas by 2050, with about 90% of this increase in Asia and Africa, according to a new United Nations dataset. Nations, 2018).

Urbanization acquires greater relevance when the climatological differences between rural and urban areas are analyzed, since the impacts of this phenomenon on the climate worldwide can be confirmed. To make sense of the choice of these two cities, the United Nations mentions that the most urbanized regions include North America (with 82% of its population living in urban areas in 2018), Latin America and the Caribbean (81%), Europe (74%) and Oceania (68%) (United Nations, 2018).

There are studies that show the consequences of the increase in population and urbanization in the imbalance of the environment. The expansion process of cities and their population density are considered important (Cartaxo, Valois, Miranda, & Costa, 2018). In the case of New York, the total population is a little over eight million inhabitants (United States Census Bureau, 2010) and in Mexico City there are almost nine million inhabitants (INEGI, 2015). Therefore, the analysis made with these two cities will be representative of the interpretation that arises.

Air quality index

In order to understand the dynamic evolution of air quality, it is necessary to implement variables that measure these effects. In the case of environmental pollution in Mexico City, the person in charge of granting the air rating is the Mexican Official Standard and the index is calculated for five of the criteria pollutants: sulfur dioxide, carbon monoxide, nitrogen dioxide, ozone and suspended particles. These particles are used to calculate the Air Quality Index, formerly called IMECA (Government of Mexico City, 2019). In 2006, there was a change in the IMECA regulations, the modifications were inspired by the changes made in 1999 to the Pollution Standard Index (PSI). Currently, the PSI contains information relevant to public health reviews with updated national standards (García, 2011).

In the case of New York, in order to analyze the evolution of environmental pollution, the AQI will be used, it is an index that notifies air quality on a daily basis. In this case, the United States Environmental Protection Agency (EPA) is in charge of the diffusion and uses five main air pollutants regulated by the Clean Air Act: ozone at ground level, particulate pollution, carbon monoxide, sulfur dioxide and nitrogen dioxide (AirNow, 2016).

In table 1 the index is classified into six sections according to the environmental conditions it has. 0 being the best air quality condition and 500 being unsafe and dangerous conditions.

Table 1
(AirNow, 2016)

Valores del índice de la calidad del aire (AQI)	Niveles preocupantes para la salud
Cuando el AQI se encuentra en esta gama:	*...la condición de la calidad del aire es:*
0 a 50	Buena
51 a 100	Moderada
101 a 150	Insalubre para grupos sensibles
151 a 200	Insalubre
201 a 300	Muy insalubre
301 a 500	Peligrosa

Vehicle pool

As a control variable, the vehicle pool of the cities is introduced with the intention of analyzing what effect this has on the AQI particles. The relevance of the automotive fleet is due to the fact that vehicles are the greatest compromisers of air quality in the United States, and produce approximately one third of all air pollution in the United States. (National Geographic Staff, 2019).

A consequence of the accelerated increase in population is the observation of the number of motorized vehicles in cities. The role of the automotive and transportation sector has played an important role in the increase in emissions worldwide. In the countries of the European Community, transport contributes 75% of all carbon monoxide (CO), 40% of hydrocarbons (HC) and 48% of nitrogen oxides (NOx). In the United States, the main source of CO is transportation, despite reductions in emissions that have occurred since 1970 due to increasingly stringent emission control standards and improvements in energy efficiency (Cartaxo et al., 2018).

Among the different pollutants taken into account for the measurement of environmental pollution, carbon monoxide is considered the vehicular pollutant with the greatest influence on the loss of atmospheric air quality (Cartaxo et al., 2018). This is why the comparison of the number of vehicles in cities is essential when analyzing the air quality index.

Methodological framework

Principal components analysis

This analysis is a common multivariate statistical technique, applied in many sciences, which mathematically characterizes the spatio-temporal variability of large data sets. One of its main goals is to identify, through data reduction, the recurrent and independent modes of variations (or signals) within a very large data set, thus summarizing the essential information of that data set so that they can be analyzed. draw meaningful and descriptive conclusions. This is achieved by classifying the initially correlated data into a hierarchy of statistically independent modes of variation that successfully explain less and less of the total variance (Eder, Bash, Foley & Pleim, 2013).

Analysis is also considered a method to identify the structure within a set of variables. Many analyzes involve large numbers of variables that are difficult to interpret. The use of PCA or factor analysis helps to find interrelationships between variables (generally called elements) to find a smaller number of unifying variables called factors (Mooi, Sarstedt, & Mooi-Reci, 2018).

Another function that PCA has is to solve the problem that arises with data that present a high level of multicollinearity (Maddala, 1992). However, the main function is to condense the information and interpret the data by omitting variables that are less significant from the data set with minimal loss of original data (Azid et al., 2014).

Operationalization of the variables

Among the variables that will be handled, those that make up the AQI are included. The index, first introduced by the EPA in 1998, ranked ambient air quality according to concentrations of major air pollutants such as PM10, PM2, ozone, SO2, NO2, and CO. Subsequently, a similar index-based approach to expressing health risk was developed in France, Great Britain, Germany, and Mexico (Kowalska et al., 2009).

However, in 2006, Mexico launched environmental regulation for the federal district NADF-009-aire-2006, which establishes the requirements to prepare the metropolitan air quality index (Secretaría del Medio Ambiente, 2006). This new norm is aligned with the measurement standards of the index in the United States, which turns out to be optimal for the comparison of the two cities to be studied.

The concentrations of the criteria pollutants O3, NO2, SO2 and CO will be expressed in parts per million (ppm), while the concentrations of PM10 and PM2.5 will be expressed in micrograms per cubic meter (µg/m3). Both types of particles are measured and reported at local conditions of pressure and temperature (Secretaría del Medio Ambiente, 2006).

Databases

For Mexico City, the Government of Mexico City publishes on its website the Air Quality Index for each hour every day. It includes the value of ozone, sulfur dioxide, nitrogen dioxide, carbon monoxide and PM10. The agency in charge of the measurement is SIMAT (Mexico City Atmospheric Monitoring System) (SIMAT, 2019).

In the case of New York, the United States Environmental Protection Agency publishes the files with the available data on its page to download. The files are updated twice a year: once in June to capture the full data for the previous year, and once in December to capture the data for the summer (ozone season). EPA and its federal, tribal, state and local partners have developed a website, AIRNow, to provide the public with easy access to national air quality information. At www.airnow.gov, you'll find daily AQI forecasts and real-time AQI conditions for more than 300 cities in the United States, with links to more detailed state and local air quality websites (Agency & Division, 2014).

The data was collected and the respective conversions were made so that the data were homologated. Image 1 shows how the data was presented in the Excel data processor, where an AQI value is assigned per day and pollutant.

Figure 1
AQI Database (Own elaboration)

date	COMPOSEDCDMX	CDMXO3	CDMXSO2	CDMXNO2	CDMXCO	CDMXPM10	COMPOSEDNYC	NYO3	NYSO2	NYCO	NYNO2	NYPM10	NYCNO2tm1
1/1/2006	66.00	37.71	13.63	24.08	32.58	75.33	75.00	2.00	39.00	6.10	25.00	7.00	
1/2/2006	67.00	34.96	11.54	22.04	24.67	59.42	72.00	9.00	52.00	6.78	29.00	14.00	25.00
1/3/2006	71.00	34.58	12.67	26.46	20.38	58.29	29.00	23.00	13.00	4.79	29.00	17.00	29.00
1/4/2006	74.00	30.38	26.04	30.21	24.33	55.33	51.00	17.00	21.00	4.85	37.00	16.00	29.00
1/5/2006	76.00	27.50	35.54	20.54	19.54	59.67	61.00	2.00	37.00	6.21	33.00	10.00	37.00
1/6/2006	77.00	29.17	14.88	14.88	9.96	57.67	36.00	16.00	24.00	4.70	24.00	4.00	33.00
1/7/2006	76.00	22.75	38.63	23.13	17.71	51.04	50.00	19.00	50.00	3.83	36.00	6.00	24.00
1/8/2006	75.00	34.54	56.92	29.17	23.21	57.54	63.00	5.00	41.00	4.89	32.00	19.00	36.00
1/9/2006	75.00	34.79	27.13	36.42	30.83	62.29	67.00	6.00	67.00	5.37	42.00	14.00	32.00
1/10/2006	76.00	34.29	23.63	29.54	22.46	71.63	56.00	10.00	41.00	4.39	36.00	6.00	42.00
1/11/2006	77.00	34.38	37.83	32.17	20.75	65.58	104.00	9.00	39.00	9.09	43.00	12.00	36.00
1/12/2006	94.00	41.67	11.96	37.58	25.96	70.08	55.00	7.00	55.00	5.72	44.00	9.00	43.00
1/13/2006	94.00	32.67	11.33	23.17	20.54	76.13	84.00	21.00	46.00	7.39	41.00	13.00	44.00
1/14/2006	95.00	35.17	8.25	20.46	11.33	65.00	44.00	19.00	9.00	4.70	37.00	8.00	41.00

Before starting with the PCA, it was necessary to standardize the data. The raw air quality variables were standardized by transforming the Z scale to a mean of 0.0 and a variance of 1.0 applying the equation:

$$Z_{ij} = (X_{ij} - \mu)/\sigma$$

where Zij is the value j of the standard score of the measured variable i, Xij is the observation number j of variable i; μ is the mean value of the variable; and σ is the standard deviation. The Z scale transformation method was used to ensure that the different air quality variables had the same weight in the statistical analysis process. Furthermore, these transformations will homogenize the variance (Azid et al., 2014).

Information processing and analysis

The Kaiser-Meyer-Olkin (KMO) test was carried out to measure the adequacy of the sampling. These matrices measure the sampling adequacy for each variable along the diagonal and the negatives of the partial correlation/covariances on the diagonals. The diagonal elements must be greater than 0.5 as a minimum if the sample is adequate for the variables (Azid et al., 2014).

First, the KMO test for Mexico City is shown, followed by the sample KMO test for New York City. The tests were performed in STATA with the use of previously standardized data. In both cases, the KMO statistic is greater than 0.5, therefore, the data is viable for PCA.

.factortest z1cdmxo3 z1cdmxso2 z1cdmxno2 z1cdmxco z1cdmxpm10

Determinant of the correlation matrix

Det = 0.232

Bartlett test of sphericity

Chi-square = 6936.058
Degrees of freedom = 10
p-value = 0.000
H0: variables are not intercorrelated

Kaiser-Meyer-Olkin Measure of Sampling Adequacy

KMO = 0.504

. factortest z1nyo3 z1nyso2 z1nyco z1nyno2 z1nypm10

Determinant of the correlation matrix

Det = 0.398

Bartlett test of sphericity

Chi-square = 1025.146
Degrees of freedom = 10
p-value = 0.000
H0: variables are not intercorrelated

Kaiser-Meyer-Olkin Measure of Sampling Adequacy

KMO = 0.593

Principal Components Analysis

Once the use of the data was confirmed, the PCA proceeded. When executing the function in the first instance, the data is returned with the number of components corresponding to the number of variables with which it was started, however, it is necessary to select an optimal number of components.

One way to decide the number of factors (or components) is to extract all factors with an eigenvalue greater than 1. The reason for this is that each factor with an eigenvalue greater than 1 represents more variance than a single variable (taking into account that before starting the analysis, the variables were standardized so that the variance was exactly 1) (Mooi et al., 2018). This method is also known as the Guttman rule (Larsen & Warne, 2010) or the Kaiser Criterion (Mooi et al., 2018).

The results for the case of Mexico City are shown below, followed by the case of New York. The results have already gone through a filtering process and only the components that obtained an eigenvalue greater than 1 are shown. In addition, only the values with a correlation value greater than 0.3 are shown, since it is usually interpreted as a median correlation and any value lower than that is considered low or negligible (Vinuesa, 2016). Illustration 2 corresponds to the results obtained for CDMX and illustration 3 for NYC.

Illustration 2

CDMX PCA (Own elaboration)

```
12 . pca z1cdmxo3 z1cdmxso2 z1cdmxno2 z1cdmxco z1cdmxpm10, comp(2)  mineigen (1) blanks(0.3)

   Principal components/correlation                  Number of obs    =       4748
                                                     Number of comp.  =          2
                                                     Trace            =          5
       Rotation: (unrotated = principal)             Rho              =     0.6690
```

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	2.1687	.992576	0.4337	0.4337
Comp2	1.17612	.341003	0.2352	0.6690
Comp3	.835116	.181544	0.1670	0.8360
Comp4	.653572	.487074	0.1307	0.9667
Comp5	.166498	.	0.0333	1.0000

Principal components (eigenvectors) (blanks are abs(loading)<.3)

Variable	Comp1	Comp2	Unexplained
z1cdmxo3	0.4030		.5481
z1cdmxso2		0.6106	.4128
z1cdmxno2	0.6264		.1467
z1cdmxco	0.5516	-0.3429	.2018
z1cdmxpm10		0.6502	.3459

Illustration 3

NYC PCA(Own elaboration)

```
8 . pca z1nyo3 z1nyso2 z1nyco z1nyno2 z1nypm10, comp(2) blanks(0.3)

  Principal components/correlation                  Number of obs    =      1117
                                                    Number of comp.  =         2
                                                    Trace            =         5
      Rotation: (unrotated = principal)             Rho              =    0.6858

      --------------------------------------------------------------------------
         Component |   Eigenvalue   Difference         Proportion   Cumulative
      -------------+------------------------------------------------------------
             Comp1 |      1.92205      .415057             0.3844       0.3844
             Comp2 |      1.50699      .901552             0.3014       0.6858
             Comp3 |      .605439     .0455015             0.1211       0.8069
             Comp4 |      .559938      .154353             0.1120       0.9189
             Comp5 |      .405584            .             0.0811       1.0000
      --------------------------------------------------------------------------

  Principal components (eigenvectors)  (blanks are abs(loading)<.3)

      ------------------------------------------------
          Variable |    Comp1     Comp2 | Unexplained
      -------------+--------------------+-------------
            z1nyo3 |             0.6639 |      .2651
           z1nyso2 |   0.5809           |      .3174
            z1nyco |   0.5620           |      .3909
           z1nyno2 |   0.5478    0.3088 |      .2796
          z1nypm10 |             0.6633 |       .318
      ------------------------------------------------
```

Once the analysis has been carried out, a varimax that aims to minimize component complexity by making large loads more significant and the least significant small loads within each component (Azid et al., 2014). That is, this procedure aims to maximize the dispersion of loadings within the factors, which means that some variables will have high loadings, while the loadings of the remaining variables will be considerably smaller (Mooi et al., 2018).

The images shown below indicate the values of the components greater than 0.3 in the components that resulted with an eigenvalue greater than 1. In addition, it is seen how the varimax rotation method is applied.

Illustration 4

CDMX PCA with rotation (Own elaboration)

```
14 . rotate, varimax blanks(0.3)

  Principal components/correlation                    Number of obs    =      4748
                                                      Number of comp.  =         2
                                                      Trace            =         5
      Rotation: orthogonal varimax (Kaiser off)       Rho              =    0.6690
```

Component	Variance	Difference	Proportion	Cumulative
Comp1	2.01868	.692546	0.4037	0.4037
Comp2	1.32613	.	0.2652	0.6690

Rotated components (blanks are abs(loading)<.3)

Variable	Comp1	Comp2	Unexplained
z1cdmxo3	0.4845		.5481
z1cdmxso2		0.6644	.4128
z1cdmxno2	0.5947		.1467
z1cdmxco	0.6415		.2018
z1cdmxpm10		0.7036	.3459

Illustration 5

NYC PCA with rotation (Own elaboration)

```
9 . rotate, varimax blanks(0.3)

  Principal components/correlation                    Number of obs    =      1117
                                                      Number of comp.  =         2
                                                      Trace            =         5
      Rotation: orthogonal varimax (Kaiser off)       Rho              =    0.6858
```

Component	Variance	Difference	Proportion	Cumulative
Comp1	1.92031	.411588	0.3841	0.3841
Comp2	1.50873	.	0.3017	0.6858

Rotated components (blanks are abs(loading)<.3)

Variable	Comp1	Comp2	Unexplained
z1nyo3		0.6749	.2651
z1nyso2	0.5700		.3174
z1nyco	0.5585		.3909
z1nyno2	0.5666		.2796
z1nypm10		0.6555	.318

Discussion

Mexico City

Once the analysis was carried out, a total of 2 main components were considered significant. Cumulatively, these two components explain 66.9% of the contaminants taken into account for the AQI measurement. Of the first component, which explains 40.37% of the data, the pollutants that turned out to be significant were Carbon Monoxide and Nitrogen Dioxide. In the second significant component, 26.53% of the data was explained and the variables that turned out to be more significant were Sulfur Dioxide and Suspended Particles (Pm10).

In the case of the main component, there is a significant combination of Carbon Monoxide and Nitrogen Dioxide. In the case of CO, it is mainly a product of the incomplete combustion of carbon-containing fuels, such as gasoline, natural gas, oil, coal, and wood. The largest anthropogenic source of CO is vehicle emissions (National Research Council, 2002). According to the Ministry of the Environment (SEDEMA), the main producers of Nitrogen Oxides are vehicles that use combustion engines, producing nitrogen oxides and in Mexico City they are the main source of emission of this pollutant (SEDEMA, 2020b).

As a justification for the first component, in the Metropolitan Area of Mexico City, the vehicle fleet more than doubled since in a decade, going from 3.7 million motorized units in 2005 to 9.5 million in 2015 (INEGI, 2018). Of the 9.5 million vehicles in circulation registered in the Valley of Mexico in 2015, 57% are registered in Mexico City (Amador, 2017).

In addition to the exaggerated presence of cars, places that have high concentrations of CO tend to have topographical or meteorological characteristics that exacerbate pollution; for example, strong temperature inversions or the existence of nearby hills that inhibit wind flow can limit the dispersion of pollutants (National Research Council, 2002). In the case of Mexico City, it is characterized by being surrounded by mountains and having a winter with clear skies, ideal places for investments and, therefore, a greater presence of CO (Jimenéz, 2001).

For the second main component, Sulfur Dioxide and suspended particles can be present together since the oxidation of Sulfur Dioxide transforms the gas into suspended particles. In the case of Mexico City, sulfur oxides are an important precursor of secondary aerosols (SEDEMA, 2020a). In addition, the presence of the Popocatépetl volcano, forest fires, and the high number of cars also contribute to the presence of suspended particles (SEDEMA, 2020c).

New York City

In the case of NYC, there were also 2 components that turned out to be significant. Between these two components, 68.58% of the total explanation of the data was accumulated. In the case of the first component, 38.41% of the information was captured and the Sulfur Dioxide, Carbon Monoxide and Nitrogen Dioxide variables stood out with more significance. Regarding the second component, 30.17% of the information was explained and the variables that stood out were Ozone and Suspended Particles (Pm10).

When talking about Sulfur Dioxide, its presence can be attributed to various sources such as the oil and gas industry, pipeline operations, marine operations or metal smelting (British Columbia Health Link, 2020). In the case of New York City, high SO2 levels are associated with boilers using No. 4 or No. 6 oil. Burning these high-sulfur fuels is common in neighborhoods with many large residential buildings (Kheirbek et al., 2013).

As for Carbon Monoxide and Nitrogen Dioxide, NO2 is mostly attributed to the presence of automobiles; maritime and air transport; and pollution generated by city buildings (Kheirbek et al., 2013). Now, since the largest anthropogenic source of CO in the United States is vehicle emissions (National Research Council, 2002), it is not surprising that this variable was also significant in the first principal component.

In the second instance, Ozone and Suspended Particles are mentioned as significant variables for the second main component. Ozone is caused by the combination of other pollutants (Volatile Organic Compounds, Nitrogen Oxides) and light (World Health Organization, 2004). In the case of New York, much of the pollution comes from motor vehicles, as they are the main anthropogenic sources of ground-level ozone precursors (Kheirbek et al., 2013).

Concluding remarks

The study of the environment has been a topic that has been incorporated into academic studies due to the growing interest in the conservation of the planet and sustainability. Along with the awakening to environmental concerns, studies devoted to the economic impacts of water, soil, and air pollution have emerged more frequently. Within the present study, it was dedicated to observing air pollution through the use of the Air Quality Index and Econometrics in order to be able to compare two cities in uneven economic situations, and to be able to examine what the differences are.

Air quality monitoring programs have generated a huge, multidimensional and complex data set, which require environmental techniques for data analysis and interpretation of the underlying information. In this study, we apply PCA to identify the pollution sources according to the city studied and analyze the different causes.

It was concluded that pollution in Mexico City is due in large part to a growing presence of vehicles and their geographical location. On the other hand, in the case of New York, its contamination is distributed in a more uniform way. Although there is a strong presence of cars in the city, its high population density and large buildings are the main source of environmental pollution.

References

Adu, D. T. & Denkyirah, E. K. (2018). Economic growth and environmental pollution in West Africa: Testing the Environmental Kuznets Curve hypothesis. Kasetsart Journal of Social Sciences, 8–15. https://doi.org/10.1016/j.kjss.2017.12.008

Agency, U. S. E. P. & Division, I. (2014). Air Quality Index (AQI). Encyclopedia of Quality of Life and Well-Being Research, (February), 120–120. https://doi.org/10.1007/978-94-007-0753-5_100115

AirNow. (2016). Air Quality Index (AQI) Basics. Retrieved May 10, 2019, from https://airnow.gov/index.cfm?action=aqibasics.aqi

AirNow. (2017). Air Quality Index - A Guide to Air Quality and Your Health. Retrieved September 7, 2019, from https://airnow.gov/index.cfm?action=aqi_brochure.index

Amador, O. (2017). La Ciudad de México duplicó el parque vehicular en 10 años. El Economista. Retrieved from https://www.eleconomista.com.mx/politica/La-Ciudad-de-Mexico-duplico-el-parque-vehicular-en-10-anos-20170801-0114.html

Ambiente, M., General, A., Unidas, N., Ambiente, M., Ambiente, O. N. U. M., Ambiente, M., … Econ, C. (2018). Comunicado de Prensa Núm. 260 / 18 de junio DE 2018 "Estadísticas a propósito del Día Mundial del Medio Ambiente (5 de junio)." 5–7.

Azid, A., Juahir, H., Toriman, M. E., Kamarudin, M. K. A., Saudi, A. S. M., Hasnam, C. N. C., … Yamin, M. (2014). Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: A case study in Malaysia. Water, Air, and Soil Pollution, 225(8). https://doi.org/10.1007/s11270-014-2063-1

Bishoi, B., Prakash, A. & Jain, V. K. (2009). A comparative study of air quality index based on factor analysis and US-EPA methods for an urban environment. Aerosol and Air Quality Research, 9(1), 1–17.

Cartaxo, E., Valois, I., Miranda, V. & Costa, M. (2018). Issuances of automotive vehicles and the impacts on air quality in the largest city in the Brazilian Amazon. Sustainability (Switzerland), 10(11). https://doi.org/10.3390/su10114091

Dale, A., Fant, C., Strzepek, K., Lickley, M. & Solomon, S. (2017). Climate model uncertainty in impact assessments for agriculture: A multi-ensemble case study on maize in sub-Saharan Africa. Earth's Future, 5(3), 337–353. https://doi.org/10.1002/2017EF000539

Eder, B., Bash, J., Foley, K. & Pleim, J. (2013). Incorporatin Principal Component Analysis into Air Quality Model Evaluation. Atmospheric Environment. https://doi.org/10.1016/j.atmosenv.2013.10.015

García, B. (2011). Los Índices de Calidad del Aire: Alcances y Limitaciones. Conciencia Tecnológica, 74–76. Retrieved from http://documat.unirioja.es/descarga/articulo/3832437.pdf

Gobierno de la Ciudad de México. (2019). Índice de Calidad del Aire. Retrieved May 10, 2019, from http://www.aire.cdmx.gob.mx/default.php?opc=%27ZaBhnmI=&dc=%27Zw==

INEGI. (2015). Número de habitantes. Retrieved from http://cuentame.inegi.org.mx/monografias/informacion/df/poblacion/

INEGI. (2018). Estadística de Vehículos de Motor Registrados en Circulación. Dirección General Adjunta de Registros Administrativos Económicos.

Jimenéz, B. E. (2001). La Contaminación Ambiental en México. Editorial Limusa.

Kheirbek, I., Johnson, S., Ito, K., Anan, K., Matte, T., Kass, D., … Ross, Z. (2013). Air Survey. New York.

Kowalska, M., Ośródka, L., Klejnowski, K., Zejda, J. E., Krajny, E. & Wojtylak, M. (2009). Air quality index and its significance in environmental health risk communication. Archives of Environmental Protection, 35(1), 13–21.

Larsen, R. & Warne, R. T. (2010). Estimating confidence intervals for eigenvalues. Behavior Research Methods, 42(3), 871–876. https://doi.org/10.3758/BRM.42.3.871

Maddala, G. . (1992). Introduction to Econometrics (Second). New Jersey: Prentice Hall.

Mooi, E., Sarstedt, M. & Mooi-Reci, I. (2018). Market Research. The Process, Data, and Methods Using Stata. In [Book]. https://doi.org/10.1007/978-981-10-5218-7

National Geographic Staff. (2019). The environmental impacts of cars, explained. Retrieved from National Geographic website: https://www.nationalgeographic.com/environment/green-guide/buying-guides/car/environmental-impact/%0D

National Research Council. (2002). Challenge of Managing Carbon Monoxide Pollution in Fairbanks, Alaska. https://doi.org/10.17226/10378

Secretaría del Medio Ambiente (2006). Norma Ambiental Para El Distrito Federal Nadf-009-Aire-2006, Que Establece Los Requisitos Para Elaborar El Índice Metropolitano De La Calidad Del Aire. Gaceta Oficial Del Distrito Federal, 154–165. Retrieved from http://www.aire.cdmx.gob.mx/descargas/monitoreo/normatividad/NADF-009-AIRE-2006.pdf

SEDEMA. (2020a). Dióxido de azufre (SO2). Retrieved May 16, 2020, from Gobierno de la Ciudad de México website: www.aire.cdmx.gob.mx/default.php?opc=%27Y6BhnmKkYg==%27%0D

SEDEMA. (2020b). Óxidos de nitrógeno (NOX). Retrieved May 16, 2020, from Gobierno de la Ciudad de México website: http://www.aire.cdmx.gob.mx/default.php?opc='Y6BhnmKkZA=='

SEDEMA. (2020c). Partículas suspendidas. Retrieved May 16, 2020, from Gobierno de la Ciudad de México website: www.aire.cdmx.gob.mx/default.php?opc='Y6BhnmKkZg=='%0D

Selden, T. M. & Song, D. (1994). Environmental quality and development: Is there a kuznets curve for air pollution emissions? Journal of Environmental Economics and Management, Vol. 27, pp. 147–162. https://doi.org/10.1006/jeem.1994.1031

SIMAT. (2019). Datos derivados del monitoreo atmosférico. Retrieved from http://www.aire.cdmx.gob.mx/default.php?opc='aKBhnmI='&opcion=aw==

Todaro, M. P. & Smith, S. C. (2015). Economic Development (12th ed.). Retrieved from http://www.downloadslide.com

UN Environment. (2019). About UN Environment.

United States Census Bureau. (2010). Profile of General Population and Housing Characteristics: 2010 2010 Demographic Profile Data. Retrieved from 2010 Demographic Profile website: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=CF

Vázquez-Guzmán, D. (2008). Measurement of Income Inequality in Mexico: Methodology, Assessment and Empirical Relationship with Poverty and Human Development. PhD thesis. Supervisors: Dipak Ghosh and David N. F. Bell. Examiners: Prasanta K. Pattanaik and Sheila DowUniversity of Stirling, UK.

Vinuesa, P. (2016). Correlación: teoría y práctica. Retrieved from https://www.ccg.unam.mx/~vinuesa/R4biosciences/docs/Tema8_correlacion.html%0D

World Health Organization. (2004). Scientific Facts on Air Pollution Ozone. GreenFacts Digest, 3–6.