

Comparative Analysis of the Impact on Air Quality Due to the Operation of La Oroya Metallurgical Complex using the Grey Clustering Method

Alexi Delgado¹, Luis Vasquez², Luis Espinoza³, Manuel Mejía⁴
Erick Yauri⁵, Chiara Carbajal⁶, Enrique Lee Huamani⁷

Mining Engineering Section, Pontificia Universidad Católica del Perú, Lima-Perú^{1, 2, 3, 4, 5}
Administration Program, Universidad de Ciencias y Humanidades, Lima-Perú⁶
Image Processing Research Laboratory, Universidad de Ciencias y Humanidades, Lima-Perú⁷

Abstract—Air pollution is one of the biggest problems worldwide due to the increase of burning of fossil fuels by industries around the world. In the present work, the air quality study will be carried out with the grey clustering method, since the data obtained presents a certain level of uncertainty. In order to obtain a correct analysis of air quality, the comparison was made in two different years with the same monitoring stations. The air quality assessment was carried out in three monitoring stations located in three different districts of the province of La Oroya (La Oroya Antigua, minor town of Huari and Santa Rosa de Sacco), in which, they installed sampling equipment for the evaluation on the basis of 10 particulate matter (PM10) and sulfur dioxide (SO₂). In each point of study a positive result was obtained, where an improvement in air quality can be seen, this is due to the reduction of mining activity in the study area. These results show the improvement over the years. Finally, this method can also be used by any organization in the nation for water or air quality studies.

Keywords—Air quality assessment; Grey clustering method; particulate matter (PM10); sulfur dioxide (SO₂)

I. INTRODUCTION

Air pollution has become a serious problem for large cities due to transportation activities [1] or industrialization [2] as studies show [3], [4]. In this paper, we apply the grey clustering method to assess air quality, which is based on grey systems theory [5], [6]. The grey clustering method can be applied using incidence matrices or weight functions. In this work, air quality will be measured using the center-point triangular whitening weight functions (CTWF) method because it helps respondents to find good answers at the central point of the intervals called grey classes [7].

In this work, we carry out an evaluation of air quality in the city of La Oroya, Junín, Peru. This city has a massive problem related to air pollution, of great magnitude in some districts, due to various factors, mainly due to mining activity [8], [9]. Therefore, the main objective of this work is the evaluation of air quality at three points in the city of La Oroya, applying the grey clustering method. Based on this, a study focused on establishing a comparison and evaluation of air quality according to standard levels between the period of 2007 and 2012 was carried out. Because air quality assessment is a subject with a high level of uncertainty, we firmly believe

that applying this method would be of great help since the CTWF method considers uncertainty in its analysis and gives weight to the criteria [10]–[14].

For such purpose, the present investigation is organized as follows. In Section 2, a literature review about the CTWF method is given. Then, in Section 3, the methodology is explained step by step. After that, in Section 4, the results and discussions are presented. Finally, the conclusions are provided in Section 5.

II. LITERATURE REVIEW

Studies by various authors were analyzed in order to obtain useful knowledge for the elaboration and application of the grey clustering method for the evaluation of air quality.

First, it was essential to establish the most important parameters to be used in order to assess air quality. In order to do this we rely on The Association between Air Pollution and Population Health Risk for Respiratory Infection: A Case Study of Shenzhen China [2], in which the probable relationship between exposure to different types of air pollution with respiratory problems in certain periods of time is analyzed. This concludes that there is indeed a relationship between air pollution and respiratory health problems and that the main pollutants were PM₁₀, PM_{2.5} AND NO₂.

Similarly, we have considered the article “Short - term Effects of Ambient Gaseous Pollutants and Particulate Matter on Daily Mortality in Shanghai, China” [15], which analyzes the effect of pollutants and gases on daily mortality in China over time periods. In this study the direct relationship between pollutants PM₁₀, SO₂, NO₂ with non-accidental mortality and cardiopulmonary diseases in China is verified.

Likewise, statistical information was necessary for the application of the grey clustering method, so the information provided by DIGESA in 2007 and 2012 [16] was used, which consists of the results obtained on the measurement of polluting agents for air, such as PM₁₀, SO₂, etc. It is worth mentioning that the data provided by DIGESA have a high degree of reliability, so any result obtained through them will also have a high degree of credibility.

Besides that the variables to be analyzed with the grey clustering method were selected. This selection was based on which variables were most relevant and have the greatest impact in terms of air pollution, in addition to their relation to the elements that originate from the mining industry [17]. The mentioned parameters are used in the grey clustering methodology to obtain a classification of the analyzed parameter.

Finally, the research developed in “Evaluating impact of air pollution on different diseases in Shenzhen, China” [18] was considered as a review for the development of the present paper. Such study assess the adverse health effects of air pollution. This study investigates the excess risk of 6 air pollutants (PM 10, PM 2.5, SO2, NO2 ...) for 21 disease groups. Daily air quality data and 1.6 million outpatient visit records from Shenzhen, China are used in the study. Where the results show that associations between air pollutants and diseases vary across different disease groups.

III. METHODOLOGY

The methodology presented in this work, to compare the degree of contamination, is the grey clustering.

To use this, we first define a set of “n” criteria (our classification of data to compare), a set of “s” grey classes (good, fair, bad, very bad) and a set of “m” objects (study points), according to sample values $[x]_{ij}$. The steps of the method are as follows:

Step 1: The center-points $\lambda_1, \lambda_2, \dots, \lambda_s$ of the grey classes are defined.

Step 2: Triangular functions are based on the relation to the number of “n” grey classes, which allows us to obtain the correspondence rule for triangular functions as represented in Fig. 1.

For a $f_j^k(x_{ij})$ calculated by (1).

$$f_j^k(x_{ij}) = \begin{cases} \frac{x-\lambda_2}{\lambda_3-\lambda_2}, & x \in [\lambda_2, \lambda_3] \\ \frac{\lambda_n-x}{\lambda_n-\lambda_3}, & x \in [\lambda_3, \lambda_n] \\ 0, & x \notin [\lambda_2, \lambda_n] \end{cases} \quad (1)$$

Step 3: The comprehensive clustering coefficient σ_i^k for object $i, i = 1, 2, \dots, m$, in the grey class $k, k = 1, 2, \dots, s$, is defined by (2).

$$\sigma_i^k = \sum_{j=1}^n f_j^k(x_{ij}) \cdot \eta_j \quad (2)$$

Where $f_j^k(x_{ij})$ is the CTWF, and η_j is the weight of criterion j .

Step 4: If $\max_{1 \leq k \leq s} \{\sigma_i^k\} = \sigma_i^{k^*}$, it is decided that object i belongs to grey class k^* . If there are several objects in grey class k^* , these objects could be ordered according to the values of $\sigma_i^{k^*}$.

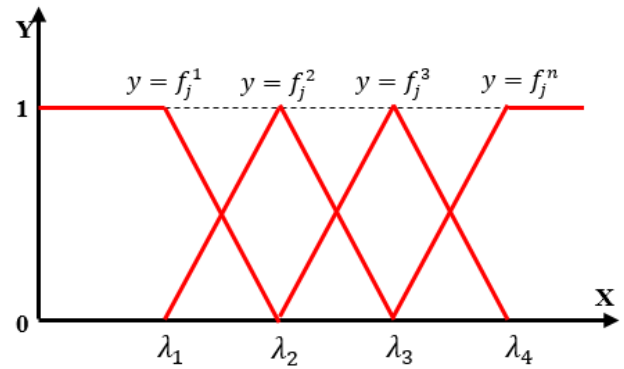


Fig. 1. Triangular Function.

IV. CASE STUDY

A. Monitoring Points

This study is based on the evaluation of air quality, which includes the La Oroya district as a scope, where three fixed stations were established in the following areas: Oroya Antigua, minor town of Huari and Santa Rosa de Sacco, installing equipment sampling for particle evaluation [19] as illustrated in Fig. 2.



Fig. 2. Location of Monitoring Points.

The monitoring points along with their locations and codes can be found in Table I.

TABLE I. MONITORING POINTS

code	Monitoring point	Location	District	Altitude	Coordinates	
					East	North
P1	INSTITUCIÓN EDUCATIVA NACIONAL N° 31149	Av. Brasil N° 222 minor town of Huari	La Oroya	3682 m	409394	8712744
P2	Vivienda	St. Dos de Mayo	Oroya Antigua	3728 m	401953	8726184
P3	Municipalidad Distrital de Santa Rosa de Sacco	Street Mariano Melgar N° 208	Santa Rosa de Sacco	3789 m	397482	8723112

B. Evaluation Parameters

For the evaluation of air quality, the Peruvian law [20] establishes four levels of air quality for each parameter. The parameter values are shown in Table II.

Likewise the data of each parameter collected in 2007 is shown in Table III.

Besides, the data of each parameter collected in 2012 is shown in Table IV.

Then, the dimensioned data for each monitoring point during 2007 is displayed in Table V.

The same data regarding the year of 2012 is given in Table VI.

Then the steps of the methodology according to Section 3 are applied.

TABLE II. AIR QUALITY LEVELS

Levels		Parameters	
		C1 (PM10)	C2 (SO2)
λ1	Good	"0 -75"	"0 -10"
λ2	Moderate	"76 - 150"	"11 - 20"
λ3	Bad	"151 - 250"	"21 - 500"
λ4	Care threshold	> 250	> 500

TABLE III. DATA OF EACH PARAMETER FOR EACH MONITORING POINT IN 2007

2007	P1	P2	P3
C1	42.44	44.92	70.74
C2	287.92	1237.27	64.41

TABLE IV. DATA OF EACH PARAMETER FOR EACH MONITORING POINT IN 2012

2012	P1	P2	P3
C1	38.7	28.4	48.5
C2	11	33	7.66

TABLE V. DIMENSIONED DATA 2007

2007	P1	P2	P3
C1	0.266	0.282	0.444
C2	1.467	6.305	0.328

TABLE VI. DIMENSIONED DATA 2012

2012	P1	P2	P3
C1	0.243	0.178	0.304
C2	0.056	0.168	0.039

Step 1: From Table II, of the air levels, the mean and dimensioned points were determined for each value of λ, this is shown in Table VII.

TABLE VII. DIMENSIONED PARAMETERS

	λ1	λ2	λ3	λ4
C1	0.235	0.706	1.255	1.804
C2	0.025	0.076	1.325	2.573

Step 2: From Table VII, the triangular functions [21] for each parameter C1 and C2 will be obtained.

As an example, these functions will be calculated for a single monitoring point as shown in (3) – (10); in which (3) – (6) are a demonstration regarding the year of 2007, while (7) – (10) displays the data of the year 2012.

$$f_j^1(x) = \begin{cases} 1, & x \in [0,0.235] \\ \frac{0.706-x}{0.706-0.235}, & x \in < 0.235,0.706 > \\ 0, & x \in [0.706, \infty > \end{cases} \quad (3)$$

$$f_j^2(x) = \begin{cases} \frac{x-0.235}{0.706-0.235}, & x \in [0.235, 0.706], \\ \frac{1.255-x}{1.255-0.706}, & x \in < 0.706, 1.255 > \\ 0, & x \notin [0.235, 1.255] \end{cases} \quad (4)$$

$$f_j^3(x) = \begin{cases} \frac{x-0.706}{1.255-0.706}, & x \in [0.706, 1.255], \\ \frac{1.804-x}{1.804-1.255}, & x \in < 1.255, 1.804 > \\ 0, & x \notin [0.706, 1.804] \end{cases} \quad (5)$$

$$f_j^4(x) = \begin{cases} \frac{x-1.255}{1.804-1.255}, & x \in [1.255, 1.804], \\ 1, & x \in < 1.804, \infty > \\ 0, & x \in [0, 1.255 > \end{cases} \quad (6)$$

$$f_j^1(x) = \begin{cases} 1, & x \in [0,0.025] \\ \frac{0.076-x}{0.076-0.025}, & x \in < 0.025,0.076 > \\ 0, & x \in [0.076, \infty > \end{cases} \quad (7)$$

$$f_j^2(x) = \begin{cases} \frac{x-0.025}{0.076-0.025}, & x \in [0.025, 0.076], \\ \frac{1.325-x}{1.325-0.076}, & x \in < 0.076, 1.325 > \\ 0, & x \notin [0.025, 1.325] \end{cases} \quad (8)$$

$$f_j^3(x) = \begin{cases} \frac{x-0.076}{1.325-0.076}, & x \in [0.076, 1.325], \\ \frac{2.573-x}{2.573-1.325}, & x \in < 1.325, 2.573 > \\ 0, & x \notin [0.076, 2.573] \end{cases} \quad (9)$$

$$f_j^4(x) = \begin{cases} \frac{x-1.325}{2.573-1.325}, & x \in [1.325, 2.573], \\ 1, & x \in < 2.573, \infty > \\ 0, & x \in [0, 1.325 > \end{cases} \quad (10)$$

The values in Table VII were replaced in the triangular functions to calculate the CTWF values in each criterion for both years of evaluation. Such results can be found in Table VIII and Table IX.

Step 3: To find the clustering coefficient, the CTWF values in Tables VIII and Table IX must be multiplied by a criterion weight. This weight is a function of the inverses of the dimensioned values in Table VII, such weights are given in Table X.

The values of the coefficient obtained are presented in Table XI and Table XII, in which the maximum value is highlighted in yellow in order to emphasize it.

Step 4: Finally, the classification of each sampling point is made according to the maximum clustering coefficient values for each year, obtaining the following:

TABLE VIII. CTWF VALUES FOR EACH MONITORING POINT FOR 2007

P1	C1	C2
f1	0.9348	0.0000
f2	0.0653	0.0000
f3	0.0000	0.8861
f4	0.0000	0.1139
P2	C1	C2
f1	0.9008	0.0000
f2	0.0992	0.0000
f3	0.0000	0.0000
f4	0.0000	1.0000
P3	C1	C2
f1	0.5565	0.0000
f2	0.4435	0.7985
f3	0.0000	0.2015
f4	0.0000	0.0000

TABLE IX. CTWF VALUES FOR EACH MONITORING POINT FOR 2012

P1	C1	C2
f1	0.9836	0.4010
f2	0.0164	0.5990
f3	0.0000	0.0000
f4	0.0000	0.0000
P2	C1	C2
f1	1.0000	0.0000
f2	0.0000	0.9267
f3	0.0000	0.0733
f4	0.0000	0.0000
P3	C1	C2
f1	0.8540	0.7346
f2	0.1460	0.2654
f3	0.0000	0.0000
f4	0.0000	0.0000

TABLE X. CRITERION WEIGHT FOR EACH PARAMETER

	λ_1	λ_2	λ_3	λ_4
C1	0.098	0.098	0.514	0.588
C2	0.902	0.902	0.486	0.412

TABLE XI. VALUES OF THE CLUSTERING COEFFICIENTS FOR 2007

2007	λ_1	λ_2	λ_3	λ_4
P1	0.091	0.006	0.431	0.047
P2	0.088	0.010	0.000	0.412
P3	0.054	0.764	0.098	0.000

TABLE XII. VALUES OF THE CLUSTERING COEFFICIENTS FOR 2012

2012	λ_1	λ_2	λ_3	λ_4
P1	0.458	0.542	0.000	0.000
P2	0.000	0.836	0.036	0.000
P3	0.663	0.239	0.000	0.000

Regarding the year of 2007:

- For P1, $\max_{1 \leq k \leq s} \{\sigma_i^k\} = 0.431$, where $k=3$. Therefore, P1 belongs to bad grey class.
- For P2, $\max_{1 \leq k \leq s} \{\sigma_i^k\} = 0.412$, where $k=4$. Therefore, P2 belongs to care threshold grey class.
- For P3, $\max_{1 \leq k \leq s} \{\sigma_i^k\} = 0.764$, where $k=2$. Therefore, P3 belongs to moderate grey class.

Regarding the year of 2012:

- For P1, $\max_{1 \leq k \leq s} \{\sigma_i^k\} = 0.542$, where $k=2$. Therefore, P1 belongs to moderate grey class
- For P2, $\max_{1 \leq k \leq s} \{\sigma_i^k\} = 0.836$, where $k=2$. Therefore, P2 belongs to moderate grey class
- For P3, $\max_{1 \leq k \leq s} \{\sigma_i^k\} = 0.663$, where $k=1$. Therefore, P3 belongs to good grey class

V. RESULT AND DISCUSSION

A. About the Case Study

In this section, the results obtained regarding the air quality of the city of La Oroya and its relationship with the activity of the city's metallurgical complex will be presented.

Regarding the results obtained in 2007, it was observed from the three monitoring points that none of them registered good air quality, in accordance with the air quality incidences of the Peruvian state [22].

Likewise, it is possible to establish that the three control points have different air qualities despite being in the same province. Thus, it is established that control point P3 presents better air quality than points P2 and P1, and that control point P1 presents better air quality than point P2, obtaining the following relation:

$$P3 > P1 > P2$$

On the other hand, regarding to the results achieved in 2012, it was observed that the P3 monitoring point recorded good air quality and, as for the other two points, air quality was improved, compared to 2007 [16].

It is also possible to establish that the three control points have different air qualities, which provides that the P3 has a better air quality than points P2 and P1. Following the order registered in 2007.

In a comparison between the results obtained in 2007 and 2012, we can see an improvement in air quality, which can be explained by the closure of the metallurgical complex La Oroya between June 2009 and August 2012, the month in which the zinc plant resumes its operations [23].

Finally, if a comparison is made between the distances of the control points and the metallurgical complex, it is observed that the tracking point P1 is 15 km away, P2 is 0.5 km and P3 5 km away; however, point P1 records a worse air quality than point P3, this is due to the direction of the air currents that are recorded in the area.

B. About the Methodology

In relation to the method, grey clustering, was applied in the analysis of 10 particulate matter (PM10) and sulfur dioxide (SO₂), with which it was possible to obtain clusterization equations and assign weights to each of the pollutants, which allows us to establish the quality of air from the city of La Oroya with respect to each pollutant analyzed.

It should be noted that in this work the weights were assigned to the specified pollutants starting from the use of the grey clustering method; however, it is necessary to establish an in-depth evaluation of the weight corresponding more precisely to each pollutant.

VI. CONCLUSIONS

The air quality in 2012 compared to 2007 shows an improvement in all the evaluated control points, which may be due to the fact that in this period the mining company entered into a total temporary freeze due to social and environmental conflicts. Likewise, it was also observed that the control points closer to the metallurgical complex did not necessarily record lower air quality relative to the more distant monitoring points, this may be due to the trajectory of air currents in the area.

It also has been demonstrated that along with the grey clustering and CTWF method an air quality analysis can be made based on specific points or on different points in the time series throughout the year, since it is based on the theory of the grey system, which analyzes the uncertainty and therefore its use in the evaluation of air quality is appropriate due to the high uncertainty that it possesses.

The method used in this work can be considered as an appropriate option for the evaluation of nitrogen dioxide in the air, since no continuous record of this contaminant has been found by DIGESA, in the period that this research was carried out. On the other hand, it is suggested to analyze the possible relationship between respiratory infections and the various air pollutants in La Oroya. Finally, it is recommended to analyze the air quality in the cities surrounding the city of La Oroya and its possible linkage with the metallurgical complex in future studies for a better understanding of the topic and its evolution in this context.

REFERENCES

- [1] V. S. Limaye et al., Development of ahmedabad's air information and response (Air) plan to protect public health, *Int. J. Environ. Res. Public Health*, vol. 15, no. 7, Jul. 2018.
- [2] X. Xia, A. Zhang, S. Liang, Q. Qi, L. Jiang, and Y. Ye, The association between air pollution and population health risk for respiratory infection: A case study of Shenzhen, China, *Int. J. Environ. Res. Public Health*, vol. 14, no. 9, Sep. 2017.
- [3] A. Delgado and A. Aguirre, Air Quality level Assessment through the Grey Clustering Analysis on Lima, Peru. .
- [4] A. Delgado, P. Montellanos, and J. Llave, Air quality level assessment in Lima city using the grey clustering method, in *IEEE ICA-ACCA 2018 - IEEE International Conference on Automation/23rd Congress of the Chilean Association of Automatic Control: Towards an Industry 4.0 - Proceedings*, 2019.
- [5] S. Liu and Y. Lin, *Grey Systems Theory and Applications*, vol. 53. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.
- [6] S. Liu and Y. Lin, *Introduction to grey systems theory*, vol. 68. 2010.
- [7] A. Delgado, A. Espinoza, P. Quispe, P. Valverde, and C. Carbajal, Water quality in areas surrounding mining: Las Bambas, Peru, *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, 2019.
- [8] El Comercio Perú, La Oroya: todo lo que tienes que saber sobre el conflicto, *El Comercio Perú*, Junín, 12-Aug-2015.
- [9] A. Delgado, R. Cuba, H. Jamanca, A. Sampen, and C. Carbajal, Social impact assessment in la oroya, peru applying the grey clustering method, *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 1, 2019.
- [10] A. Delgado, H. Reyes, I. Romero, and C. C. Mancilla, Social impact assessment using the grey clustering method: A case study on a mining project, 2019, pp. 1–5.
- [11] A. Delgado, E. Luna, M. Hernández, K. Montero, and C. Carbajal, Assessment of the air quality in four cities with near mining activity in mexico, using the grey clustering method, *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, 2019.
- [12] A. Delgado, C. Carbajal, H. Reyes, and I. Romero, Social Impact Assessment on a Mining Project in Peru Using the Grey Clustering Method and the Entropy-Weight Method, *Commun. Comput. Inf. Sci.*, vol. 1096 CCIS, pp. 116–128, 2019.
- [13] A. Delgado, N. Rojas, J. Oblitas, B. Andrés, A. Huerta, and C. Carbajal, Water quality assessment using the grey clustering analysis on a river of Taxco, Mexico, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 4717–4723, 2020.
- [14] A. Delgado, J. Culqui, G. Tasayco, A. Millán, E. Tirado, and C. Carbajal, Quality assessment of surface water associated with a copper mine in peru using grey systems, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 6660–6668, 2020.
- [15] G. Chen et al., Short-term effects of ambient gaseous pollutants and particulate matter on daily mortality in Shanghai, China, *J. Occup. Health*, vol. 50, no. 1, pp. 41–47, Jan. 2008.
- [16] Ministerio de Salud del Perú, La Oroya, pp. 1–37, 2007.
- [17] Y. Zhang, J. Ni, J. Liu, and L. Jian, Grey evaluation empirical study based on center-point triangular whitenization weight function of Jiangsu Province industrial technology innovation strategy alliance, *Grey Syst. Theory Appl.*, vol. 4, no. 1, pp. 124–136, Jan. 2014.
- [18] L. L. Chen, J. Xu, Q. Zhang, Q. H. Wang, Y. Q. Xue, and C. R. Ren, Evaluating impact of air pollution on different diseases in Shenzhen, China, *IBM J. Res. Dev.*, vol. 61, no. 6, pp. 21–29, Nov. 2017.
- [19] I. Sociales, *Anuario de Estadísticas Ambientales*, 2012.
- [20] D. S. N. 063-2007 PCM, Decreto Supremo, D. Of. El Peru, pp. 1–12, 2008.
- [21] L. Zhou and S. Xu, Application of Grey Clustering Method in Eutrophication Evaluation of Wetland, 2007.
- [22] Ministerio del Ambiente, Resolución Ministerial N°181-2016-MINAM. Indic de calidad del aire, pp. 1–6, 2016.
- [23] Ministerio de Salud, Informe N° 003939-2012/DEPA/DIGESA, Vigil. Sanit. la Calid. del Aire por el reinicio las Act. del Complejo Met. Doe Run Perú, 2012.