Applying Grey Systems and Inverse Distance Weighted Method to Assess Water Quality from a River

Alexi Delgado[1], Anthonny Fernandez[2], Eduardo Lozano[3], Dennis Miguel[4], Félix León[5], Jhosep Arteta[6], Ch. Carbajal[7]
Mining Engineering Section, Pontificia Universidad Católica del Perú, Lima, Peru[1]
Environmental Engineering Faculty, Universidad Nacional de Ingeniería, Lima, Peru[2, 3, 4, 5, 6]
Administration Program, Universidad de Ciencias y Humanidades, Lima, Peru[7]

***Abstract*—The Cañete River basin, in Peru, has suffered an increase in pollution due to various causes, among the main ones being the lack of knowledge, culture of individuals and municipal authorities, economic activities, among others. We analyze this degree of pollution reflected in the upper part of the Cañete river basin through the Grey Clustering method, based on grey systems, which is presented as a good alternative to evaluate water quality in a comprehensive way, making use of Historical data from the monitoring program for the years 2014 and 2015 with nine monitoring points carried out by the National Water Authority (ANA), 6 parameters were defined to evaluate: Hydrogen potential (pH), Biochemical oxygen demand, Chemical demand of Oxygen, Total Suspended Solids, Total Manganese and Total Iron based on the PRATI index. For the spatial distribution, interpolation surfaces of the clustering coefficients were created, using the Spatial Analyst extension of the ArcGIS software, which provides tools to create, analyze and map data in raster format or surfaces. The interpolation method used is Inverse Distance Weighted (IDW). The results of the evaluation showed that in 2014 the monitoring points determined, through the Grey Clustering method, a level of contamination "Uncontaminated" at each point except for point P7 which gives us an "Acceptable" level according to the PRATI indices, while for the year 2015 points P1 and P2 indicate a level of contamination "Moderately contaminated", point 3 an "Acceptable" level, after points P4 to P9 they present a level of "Not contaminated". Finally, the Grey Clustering analysis method will determine the water quality in the 9 monitoring points of the upper-middle basin of the Cañete River in the years 2014 and 2015. Allowing to observe the reduction of water quality in points P1 and P2 for the period of the years 2014 and 2015 respectively, being crucial to achieve water resource management among local governments that can insert awareness and sustainable development policies.**

Keywords*—Grey systems; inverse distance weighted (IDW); water quality

I. Introduction

In the Cañete River basin, agriculture constitutes the main socioeconomic activity in the valley and is the activity with the highest water consumption. The intensive and unplanned use of the water resources of the Cañete River Basin puts at risk the modification of the characteristics of water availability and quality, so it is important to maintain an adequate quality of the water resource (ANA). For this, it is necessary to evaluate the current state of water resources such as rivers that are the receivers of effluents [1].

As rivers are the main receivers of effluents generated by the population, the quality parameters of river waters have deteriorated in recent years, basically due to the misuse of wastewater management by the population and the negligence of the authorities in solving water pollution through treatment and avoidance, affecting a large part of the agricultural activity of the valley in the lower part of the Cañete River basin and consequently the health of the population [2].

In the Cañete river basin, an increase in pollution is observed due to various causes, among the main ones are environmental ignorance, the culture of the people and the municipal authorities, being crucial to achieve the management of water resources among local governments where There are awareness policies and sustainable development can be inserted in the towns of its jurisdiction avoiding the contamination of said resource with the discharge of drains into rivers and the development of industrial, mining, agricultural economic activities and among other human activities [3].

Therefore, this work has the general objective of estimating the degree of contamination reflected in the river water quality data in nine monitoring points carried out by the ANA in the Alto Cañete basin in 2014 and 2015 using the Grey Clustering method, comparing it with an international scale the PRATI Index, for the study of water quality. To achieve the purpose of the research study, the following specific objectives were established:

1) Define the parameters to be evaluated, based on the PRATI index, to determine the degree of contamination of the rivers.

2) Carry out the Grey Clustering methodology for the monitoring points.

3) Rank the bodies of water by level of contamination.

There have been many recent investigations on the application of the grey clustering methodology to evaluate the environmental quality of different environmental components, but the one of interest in this study is on the quality of the water bodies that can, be affected by different sources such as economic activities anthropic, natural and among others, for this reason some investigations related to this environmental component are presented below.

Alexis Delgado et al., In 2017 evaluated the water quality of the Rímac River in its main tributaries; Río Blanco, Aruri, Río Rímac, Río Mayo and Río Santa in which the CTWF method was applied, which is based on Grey's systems theory. on the other hand, the data used were obtained from the ANA in which they were analyzed according to the water quality parameters such as O2, BOD, COD, SS, NH3 and NO3. Obtaining as results using the prati la clairvoyance index that the tributaries are classified as not contaminated, however, the Santa River presents greater vulnerability to contamination of the water quality [4].

Fu and Zou, in 2018, applied the Grey Clustering methodology to evaluate the water quality of 12 monitoring points in the Yellow River basin, using as real data those obtained through the monitoring of water quality from the Ministry of Environmental Protection China in May 2016, and as standard data, the surface water environmental quality standard. The results conclude that the general quality of the water in said basin is good, thus reaffirming the real situation of the basin. It is even concluded that the Grey Clustering method applied in the evaluation of water quality is reasonable, feasible and it is convenient to calculate [5].

Alexi Delgado et al. In 2019 they carried out the study called "Water quality in areas close to mining: Las Bambas, Peru" located the study place in the district of Tambobamba, province of Cotabambas, in the Apurímac region. The contamination levels of the water bodies due to the Las Bambas mine towards the area of influence that includes the Challhuahuacho and Ferrobamba rivers were evaluated. Obtaining as results that from the six monitoring points located, carried out between February 2017 and March 2019 by ANA, the rivers would have a low or high degree of contamination [6].

The authors Alexi Delgado et al., 2020 carried out the study called "Evaluation of the quality of surface water in the upper basin of the Huallaga river, in Peru, using grey systems and Shannon entropy" where the quality of the water in the basin was evaluated Upper Huallaga river taking into account the results of the monitoring of twenty-one points carried out by the National Water Authority (ANA) analyzing nine parameters of the PRATI Index. The results showed that all the monitoring points of the Huallaga River were classified as uncontaminated, which means that the discharges, generated by economic activities, are carried out through treatment plants that meet the quality parameters [7].

II. Data

A. Water Quality Parameters of the Cañete River Basin

According to the ANA in its report IT 148-2019-ANA-CAÑETE determines 3 studies to identify polluting sources in the years 2010, 2014, 2017 regarding the Cañete river basin that could be affecting the quality of its waters, in this report It is described that finding the location of these sources is of utmost importance to establish measures for their recovery and preservation of the water resource, the data that it gives us is that there are 40 sources of contamination distributed throughout the basin, which is found in greater numbers In the upper part of the Cañete river basin, the ANA in the report mentions that the river is affected by the discharges of wastewater of domestic origin, as the main sources of contamination due to the lack of maintenance of the equipment of water treatment or in the absence of these. These reports can also be corroborated with the study carried out by the University of Cañete in 2018 that determines the influence of river pollution with the development of the Province of Cañete.

For the evaluation of surface water quality through Grey Clustering, use is made of monitoring reports from a reliable source. The Cañete Fortaleza Water Administrative Authority, monitors the quality of the natural bodies of surface water in the Cañete river basin, establishing itself to carry out the monitoring network of 15 monitoring points in 2014 and 20 monitoring points of monitoring in 2015 throughout the basin, as a consequence of the results of the identification of polluting sources determined in the Technical Report made by the administrative authority of the basin (Nº139-2014-ANA-AAA.CF- ALA.MOC-AT/LEAP). In the same way, taking as a reference the results of the water quality monitoring for the years 2014 and 2015 (Nº060-2014-ANA-DGCRH/GOCRH y Nº086-2015-ANA-DGCRH/GOCRH), 9 monitoring points were determined located in the upper basin of the Cañete River, due to the presence of new activities that could generate alterations to the water quality after 2014.

The influence is established by non-experimental methodology, using the survey technique. The surveys are carried out with 100 inhabitants living on the banks of the Cañete River (the districts of Zuniga, Pacaran and Lunahuana are included in the surveys). The study concludes with a perception of moderate and high contamination of the Cañete riverbed due to the discharge of wastewater from the growing tourist activity within the districts, as well as the discharge of domestic water from the growing city, having clear what are the discharges [8].

Having clear the type of discharge that predominates, the basic parameters were chosen in a monitoring study such as pH, BOD, COD, SST and two additional parameters were also added with which are iron and manganese for study purposes as shown in Table I.

TABLE I. PARAMETERS PRIORITIZED FOR THE EVALUATION OF WATER QUALITY

Parameters	Units	Notation
pH	pH	C1
BOD	ppm	C2
DQO	ppm	C3
STS	mg/L	C4
Mn	ppm	C5
Fe	ppm	C6

III. Methodology

The development of this study is established through three approaches. At the beginning, we will proceed with the description of the study area and the different water bodies, to be considered, in said area to focus the investigation. Then, to evaluate the water bodies according to their quality, we will proceed with the Grey Clustering method. And finely, the spatial distribution of the Water quality will be analyzed using ArcGis.

The analysis of surface water quality was carried out in the upper part of the Cañete River basin (Fig. 1), which is in the central zone of Peru and has a surface area of 1,756.05 km2.

A. Case Study

In 2014 and 2017, the National Water Authority has carried out identification of polluting sources in the Cañete River basin, allowing to know in detail the activities that would be affecting the quality of the water bodies in the Cañete River basin The results indicate that there is an increase in the number of discharges of domestic wastewater discharged to the receiving body from 12 to 19. There is also a couple that are the main reference for the programming of management actions in the Cañete Basin. In order to establish measures for its recovery and conservation of surface waters [9] which is represented in Fig. 2.

For the evaluation of the surface water quality of the upper Cañete River basin, information was collected from 9 monitoring points obtained from the water quality monitoring carried out from December 15 to 19, 2014 [10] and from October 19 to 23, 2015 (ANA, 2015) by the Cañete Fortaleza Water Administrative Authority and the Mala Omas Cañete Local Water Authorities which is represented in Fig. 3.

B. Data Processing: Grey Clustering

This section describes the method established in Grey Clustering, based on grey systems. Being originally developed by Deng [11] in his Grey System Theory. Where the central point triangular standardization weight function (CTWF) will be used to test if the observation objects belong to predetermined classes, so that they can be treated accordingly to their characteristics [12].

1) Explanation of the Grey Clustering method:

a) Step 1: Data Sizing: The dimensioning of the water quality monitoring data of each environmental indicator and of the selected Water Quality Index parameters is carried out.

b) Step 2: Grey Clustering Classification: The Grey classification is established from the Environmental Quality Index chosen by the CTWF method, as can be seen in Fig. 4.

Fig. 1. Cañete River Basin, Peru.

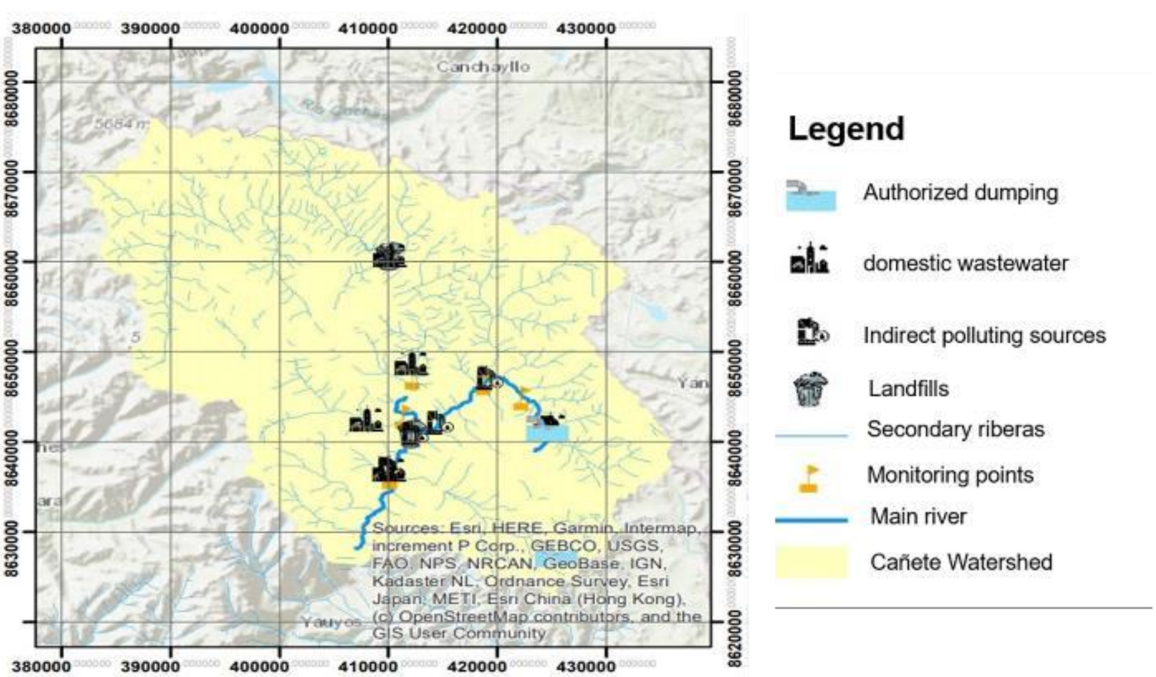

Fig. 2. Pollutant Sources of the upper Cañete River Basin.

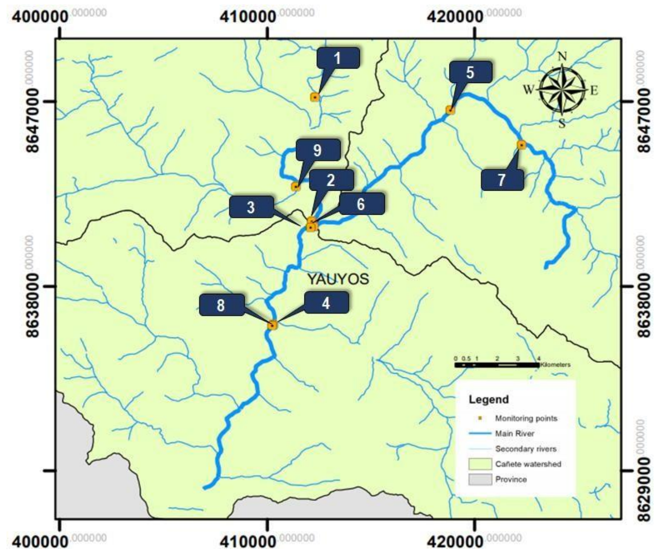

Fig. 3. Control Points for the Quality of Surface Waters in the Upper Cañete River Basin.

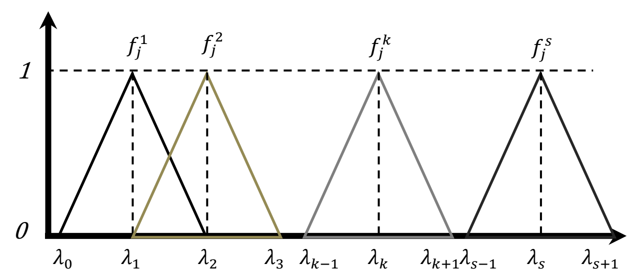

Fig. 4. Central Point Triangular Standardization Weight Function in Relation to the Selected Environmental Quality Index.

If there is a greater range of environmental quality, more functions will be increased, following the same procedure. Once the Grey classification has been established using the CTWF method, these are calculated from the k-nth class grey, k = 1, 2, 3, 4, 5…n of the j-nth parameter, j = 1, 2 …n, for a value of water quality monitoring Xij according to the following functions [13]:

$$f_j^1(x_{ij}) = \begin{cases} 0\,; x \notin [0, \lambda_j^2] \\ 1\,; x \in [0, \lambda_j^1] \\ \frac{\lambda_j^2 - x}{\lambda_j^2 - \lambda_j^1}\,; x \in [\lambda_j^1, \lambda_j^2] \end{cases} \qquad (1)$$

$$f_j^m(x_{ij}) = \begin{cases} 0\,; x \notin [\lambda_j^{m-1}, \lambda_j^{m+1}] \\ \frac{x - \lambda_j^{m-1}}{\lambda_j^m - \lambda_j^{m-1}}\,; x \in [\lambda_j^{m-1}, \lambda_j^m] \\ \frac{\lambda_j^{m+1} - x}{\lambda_j^{m+1} - \lambda_j^{m-1}}\,; x \in [\lambda_j^m, \lambda_j^{m+1}] \end{cases} \qquad (2)$$

$$f_j^k(x_{ij}) = \begin{cases} 0\,; x \notin [\lambda_j^k, \propto] \\ \frac{x - \lambda_j^k}{\lambda_j^{k+1} - \lambda_j^k}\,; x \in [\lambda_j^k, \lambda_j^{k+1}] \\ 1\,; x \in [\lambda_j^{k+1}, \propto] \end{cases} \qquad (3)$$

c) Step 3: Weight of each parameter of the selected Environmental Quality Index: We proceed with the calculation of the weight of each parameter of the environmental quality index in an objective way, known as arithmetic weight, which is calculated according to Equation 4.

$$n_j^k = \frac{1/\lambda_j^k}{\sum_{j=1}^{m} 1/\lambda_j^k} \qquad (4)$$

d) Step 4: Determination of the environmental quality classification of the evaluated point: Once the values evaluated in the Whitenization functions and the weights of each parameter have been determined, the Clustering coefficient is calculated for each value obtained in the different grey classifications by means of Equation 5 with the one with the highest value being σ_i^k the value that defines the environmental quality classification of the evaluated point according to Equation 6.

$$\sigma_i^k = \sum_{j=1}^{n} f_j^k(x_{ij}).n_j \qquad (5)$$

$$\max_{1 \le k \le s}\{\sigma_i^k\} = \sigma_i^{k^*} \qquad (6)$$

2) Application of the Grey Clustering method:

a) Step 1: PRATI water quality standard and its determination of central points: Table II shows the classification of the water quality level according to the PRATI Index. Being assigned as follows:

λ_1: Not Contaminated.

λ_2: Acceptable.

λ_3: Moderately Acceptable.

λ_4: Contaminated.

λ_5: Highly Contaminated.

Then we proceed to determine the midpoints of each level of water quality established by the PATRI index. These points corresponding to each level (Table III) will be the central points to consider for the aforementioned symbology: $\lambda 1$, $\lambda 2$, $\lambda 3$, . . , $\lambda 5$.

TABLE II. PRATI INDEX STANDARD DATA FOR WATER QUALITY ASSESSMENT

Parameters	Quality Status Index				
	λ_1	λ_2	λ_3	λ_4	λ_5
pH	6.5-8.0	8.0-8.4	8.4-9.0	9.0-10.1	>10.1
BOD (ppm)	0.0-1.5	1.5-3.0	3.0-6.0	6.0-12.0	>12.0
COQ (ppm)	0-10	10-20	20-40	40-80	>80
SST (mg/L)	0-20	20-40	40-100	100-278	>278
NH3 (ppm)	0-0.1	0.1-0.3	0.3-0.9	0.9-2.7	>2.7
NO3 (ppm)	0-4	4-12	12-36	36-108	>108
Cl (ppm)	0-50	50-150	150-300	300-620	>620
Mn (ppm)	0-0.05	0.05-0.17	0.17-0.50	0.50-1.00	>1.00
Fe (ppm)	0-0.1	0.1-0.3	0.3-0.9	0.9-2.7	>2.7

TABLE III. CORE VALUES OF THE PRATI INDEX

Parameters	Nomenclature	λ_1	λ_2	λ_3	λ_4	λ_5
pH	C1	7.25	8.20	8.70	9.55	10.40
BOD	C2	0.75	2.25	4.50	9.00	13.50
COD	C3	5.00	15.00	30.00	60.00	90.00
TSS	C4	10.00	30.00	70.00	189.00	308.00
Mn	C8	0.03	0.11	0.34	0.75	1.16
Fe	C9	0.05	0.20	0.60	1.80	3.00

Table IV and Table V present the results of the monitors in 9 points of the Cañete River in the upper middle basin of the same name, carried out by the ANA in the years 2014 (Nº060-2014-ANA- DGCRH/GOCRH) and 2015 (Nº086-2015-ANA-DGCRH/GOCRH).

b) Step 2: Sizing the PRATI Index and the water quality monitoring data: Table VI shows the oversized values of the PRATI Index, as well as the oversized values (Table VII) of the monitoring carried out by ANA in 2014 and 2015.

c) Step 3: Clustering weights of the parameters: Based on the values in Table IV, the values of the Clustering weights of each parameter of the Water Quality Index - PRATI were determined according to Equation 4. The values are shown in Table VIII.

d) Step 4: Obtaining the Whitening functions and their evaluation: To obtain the Whitenization functions, with the five Grey classifications, the values of Table III were substituted in Equation 1, Equation 2 y Equation 3. Next, the data in Table VI was evaluated in the Whitenization functions formed, yielding the following values shown in Table IX and so on for the remaining points except for P3 and P6.

e) Step 5: Results using the max. Clusterization Coefficients: To conclude, we proceed to calculate the highest value of $\{\sigma k\}$, based on Equation 6. With this, it was possible to determine to which Grey Class each monitoring point belongs and what water quality each of the nine monitoring points, the results are observed in Table X and Table XI, for the years 2014 and 2015 respectively.

TABLE IV. VALUES OF THE MONITORING PARAMETERS OF THE NINE POINTS OF THE UPPER MIDDLE BASIN OF THE CAÑETE RIVER CARRIED OUT IN 2014

Results of water quality monitoring for the year 2014						
Points	***C1***	***C2***	***C3***	***C4***	***C5***	***C6***
P. 1	0.959	0.167	0.125	0.012	0.006	0.021
P. 2	0.964	0.167	0.125	0.012	0.003	0.013
P. 3	0.959	0.167	0.125	0.012	0.015	0.035
P. 4	0.963	0.167	0.125	0.012	0.003	0.016
P. 5	0.882	0.167	0.125	0.063	0.148	0.168
P. 6	0.884	0.167	0.125	0.012	0.028	0.036
P. 7	0.947	0.333	0.25	0.064	0.232	0.286
P. 8	0.85	0.167	0.125	0.063	0.005	0.022
P. 9	0.827	0.333	0.25	0.071	0.02	0.203

TABLE V. VALUES OF THE MONITORING PARAMETERS OF THE NINE POINTS OF THE UPPER MIDDLE BASIN OF THE CAÑETE RIVER CARRIED OUT IN 2015

Results of water quality monitoring for the year 2015						
Points	***C1***	***C2***	***C3***	***C4***	***C5***	***C6***
P. 1	0.956	0.633	0.303	0.018	15.142	0.42
P. 2	0.965	0.5	0.403	0.008	16.998	0.004
P. 3	0.946	0.617	0.403	0.008	26.004	0.006
P. 4	0.961	0.583	0.505	0.008	16.674	0.004
P. 5	0.974	0.5	0.505	0.008	16.255	0.086
P. 6	0.956	0.5	0.403	0.008	37.322	0.005
P. 7	0.968	0.6	0.303	0.008	14.314	0.196
P. 8	0.959	0.667	0.303	0.008	7.755	0.039
P. 9	0.823	0.667	0.303	0.008	5.619	0.019

TABLE VI. OVERSIZED VALUES OF THE PRATI INDEX

Parameters	Nomenclature	λ_1	λ_2	λ_3	λ_4	λ_5
pH	C1	0.822	0.930	0.986	1.083	1.179
BOD	C2	0.125	0.375	0.750	1.500	2.250
COD	C3	0.125	0.375	0.750	1.500	2.250
TSS	C4	0.082	0.247	0.577	1.557	2.537
Mn	C8	0.053	0.231	0.704	1.576	2.437
Fe	C9	0.044	0.177	0.531	1.593	2.655

TABLE VII. DIMENSIONING OF THE WATER QUALITY PARAMETERS OF THE MONITORING CARRIED OUT BY THE ANA FOR THE YEARS 2014 AND 2015

Sizing of the water quality parameters for the year 2014						
Points	***C1***	***C2***	***C3***	***C4***	***C5***	***C6***
P. 1	0.959	0.167	0.125	0.012	0.006	0.021
P. 2	0.964	0.167	0.125	0.012	0.003	0.013
P. 3	0.959	0.167	0.125	0.012	0.015	0.035
P. 4	0.963	0.167	0.125	0.012	0.003	0.016
P. 5	0.882	0.167	0.125	0.063	0.148	0.168
P. 6	0.884	0.167	0.125	0.012	0.028	0.036
P. 7	0.947	0.333	0.250	0.064	0.232	0.286
P. 8	0.850	0.167	0.125	0.063	0.005	0.022
P. 9	0.827	0.333	0.250	0.071	0.020	0.203
Sizing of the water quality parameters for the year 2015						
Points	***C1***	***C2***	***C3***	***C4***	***C5***	***C6***
P. 1	0.956	0.633	0.303	0.018	15.142	0.42
P. 2	0.965	0.500	0.403	0.008	16.998	0.004
P. 3	0.946	0.617	0.403	0.008	26.004	0.006
P. 4	0.961	0.583	0.505	0.008	16.674	0.004
P. 5	0.974	0.500	0.505	0.008	16.255	0.086
P. 6	0.956	0.500	0.403	0.008	37.322	0.005
P. 7	0.968	0.600	0.303	0.008	14.314	0.196
P. 8	0.959	0.667	0.303	0.008	7.755	0.039
P. 9	0.823	0.667	0.303	0.008	5.619	0.019

TABLE VIII. CLUSTERING WEIGHTS FOR EACH PARAMETER

Clustering weights of each parameter						
Points	*C1*	*C2*	*C3*	*C4*	5	*C6*
λ_1	0.018	0.118	0.118	0.179	235	0.333
λ_2	0.053	0.130	0.130	0.198	212	0.276
λ_3	0.116	0.153	0.153	0.199	162	0.216
λ_4	0.222	0.160	0.160	0.154	153	0.151
λ_5	0.290	0.152	0.152	0.135	141	0.129

TABLE IX. FUNCTIONS VALUES FOR EACH PARAMETER

Whitening function evaluated at P3 (2014)							
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	σ_i^k
f_1	0.000	0.832	1.000	1.000	1.000	1.000	**0.962**
f_2	0.482	0.168	0.000	0.000	0.000	0.000	0.047
f_3	0.518	0.000	0.000	0.000	0.000	0.000	0.060
f_4	0.000	0.000	0.000	0.000	0.000	0.000	0.000
f_5	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Whitening function evaluated at P6 (2014)							
f_1	0.424	0.832	1.000	1.000	1.000	1.000	**0.970**
f_2	0.576	0.168	0.000	0.000	0.000	0.000	0.052
f_3	0.000	0.000	0.000	0.000	0.000	0.000	0.000
f_4	0.000	0.000	0.000	0.000	0.000	0.000	0.000
f_5	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Whitening function evaluated at P3 (2015)							
f_1	0.000	0.000	0.000	1.000	0.000	1.000	**0.512**
f_2	0.720	0.356	0.927	0.000	0.000	0.000	0.205
f_3	0.280	0.644	0.073	0.000	0.000	0.000	0.143
f_4	0.000	0.000	0.000	0.000	0.000	0.000	0.000
f_5	0.000	0.000	0.000	0.000	1.000	0.000	0.141
Whitening function evaluated at P6 (2015)							
f_1	0.000	0.000	0.000	1.000	0.000	1.000	**0.512**
f_2	0.540	0.667	0.927	0.000	0.000	0.000	0.236
f_3	0.460	0.333	0.073	0.000	0.000	0.000	0.116
f_4	0.000	0.000	0.000	0.000	0.000	0.000	0.000
f_5	0.000	0.000	0.000	0.000	1.000	0.000	0.141

3) Spatial distribution of the condition of water quality: To perform the spatial distribution, it was established to create interpolation surfaces of the clustering coefficients, where the Spatial Analyst extension of the Arcgis software is used, which provides tools to create, analyze and map data in raster format or surfaces. Water quality interpolations will be carried out to estimate the pollution values in the selected section of the Cañete River Basin. The interpolation method used is the Inverse Distance Weighted or "IDW" (Inverse Distance Weighted). Where the IDW assumes that each water quality monitoring point has a local influence that decreases with distance. This is observed in more detail in Fig. 5 of the years 2014 and 2015, therefore, these images were the result of the spatial distribution analysis thrown by ArcGIS.

For the interpolation of the water quality data, the 8 monitoring points were located to generate the area of influence of the 50-meter section around the Cañete River channel so that it can be visible on an adequate scale. The area of influence established by the monitoring from the ANA, starts from downstream from the town of Llapay and the Laraos river (RCañe4) to downstream from the Viti district (RCañe1) and upstream from the Huancachi town center (Ralis3).

The value determined by the Grey Clustering for the Z value field was used and the size of the grid or cell was defined as 1 meter. The area of influence is used to delimit the interpolated values. Finally, the IDW water quality surfaces were generated for each of the areas of influence. Where the maximum clustering coefficients were classified to classify the stretch of river in 5 classes established in the PRATI Index:

- Not contaminated.
- Acceptable.
- Moderately polluted.
- Contaminated.
- Highly polluted.

TABLE X. EVALUATION OF WHITENING FUNCTIONS IN THE NINE POINTS FOR THE YEAR 2014

Parameter	**Maximum Clusterization Coefficient**	**Water Quality Level**
P1	0.9623	Not contaminated
P2	0.9623	Not contaminated
P3	0.9623	Not contaminated
P4	0.9623	Not contaminated
P5	0.5404	Not contaminated
P6	0.9700	Not contaminated
P7	0.5927	Acceptable
P8	0.9756	Not contaminated
P9	0.5092	Not contaminated

TABLE XI. EVALUATION OF WHITENING FUNCTIONS IN THE NINE POINTS FOR THE YEAR 2015

Parameter	**Maximum Clusterization Coefficient**	**Water Quality Level**
P1	0.308	Moderately polluted
P2	0.799	Moderately polluted
P3	0.512	Not contaminated
P4	0.512	Not contaminated
P5	0.407	Not contaminated
P6	0.512	Not contaminated
P7	0.423	Acceptable
P8	0.546	Not contaminated
P9	0.564	Not contaminated

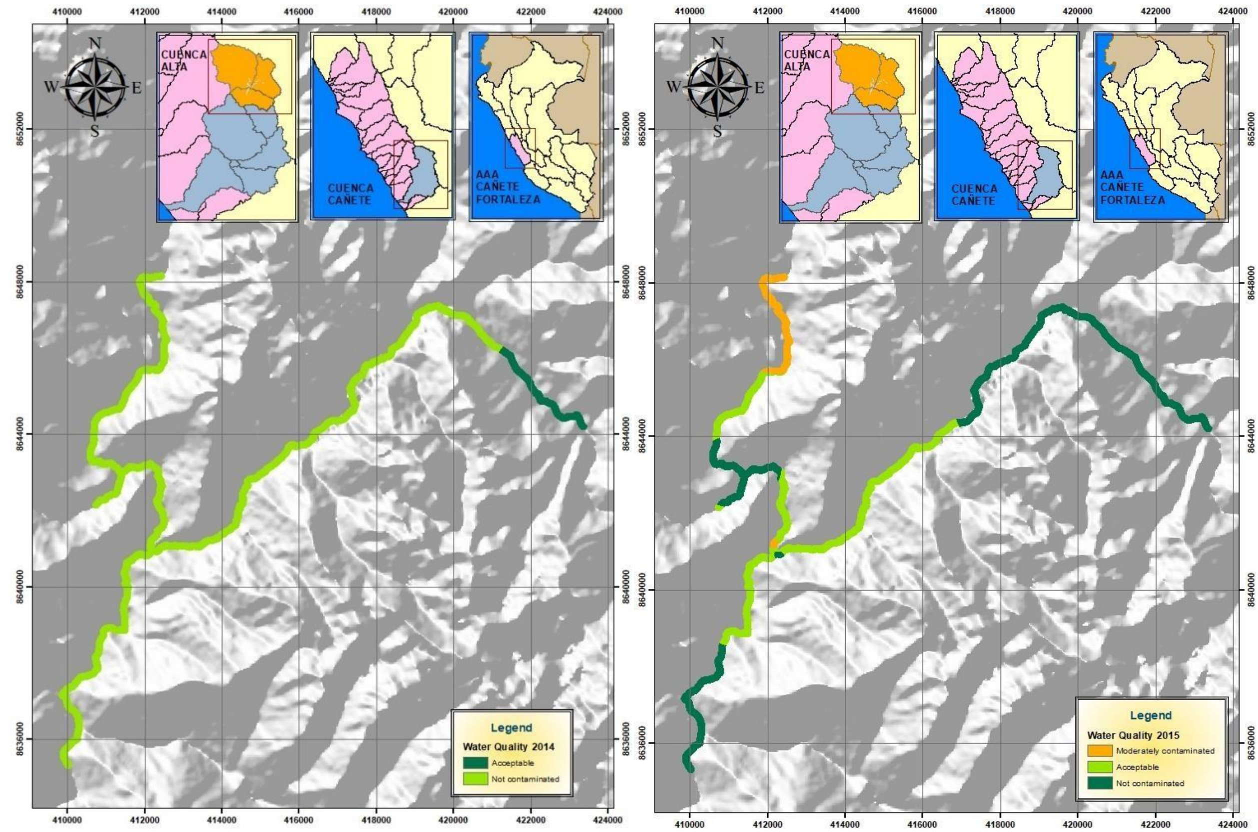

Fig. 5. Distribution of Water Quality in 2014 (Left) and 2015 (Right).

IV. Result and Discussion

A. About the Case Study

It is shown that for the year 2014, in Table X, that 8 monitoring points resulted in an uncontaminated water quality except for monitoring point 7, which has an Acceptable water quality, however, a comparison can be made of the quality level according to the maximum Clusterization coefficient.

Likewise, for the year 2015, in Table XI, 6 monitoring points resulted in an "uncontaminated" water quality, while monitoring points 1 and 2 have a "Moderately contaminated" water quality except for the point of monitoring 3 that has an "Acceptable" water quality and a comparison of the quality level was made according to the maximum Clusterization coefficient.

According to Fig. 6, for the years 2014 and 2015, the maximum clustering coefficients decrease progressively and the majority is of "Unpolluted" quality, this because the study area is located in the upper basin of the Cañete River, where the affectation of the water resource is not high compared to the middle or lower part of a basin.

B. Application of the ArcGIS Tool for Digitizing the Results of the Grey Clustering Method

The applied methodology uses a mixture of two tools the ArcGIS and the Grey method that are combined to show us the quality of the water on a geographical plane, indicating a thematic graph and of clear order, with respect to the Grey method, it offers an alternative to evaluate the quality of the water, comprehensively considering the uncertainty within the analysis, as Delgado et al. tell us, 2020, due to taking as a section the upper part of the Chillón River Basin, a study of physicochemical parameters would not be sufficient because the river flow is not stationary and the concentrations over time are not therefore that this method takes that uncertainty in its methodology, therefore the method is well adapted for the evaluation of quality in the upper part or the stretch of the Cañete river, continuing with the use of the GIS tool (ArcGIS) defines us that it is system that allows to collect, organize, manage, analyze, share and distribute geographic information, that is why this tool projects the results to us throughout the section that was studied on a map where it could be analyzed in a more interactive way since it reflects the data in something visible [14].

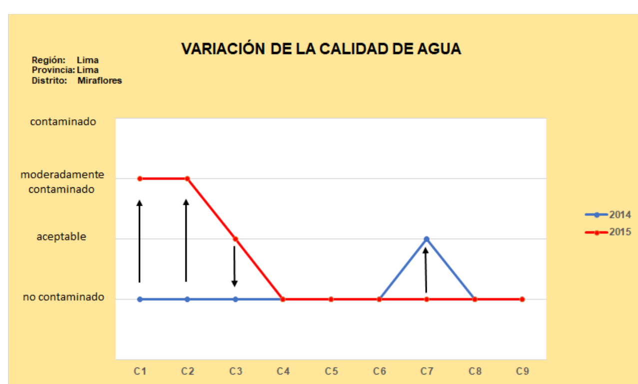

Fig. 6. Variation in Water Quality in 2014 and 2015.

C. Analysis of the Results in the Year 2014-2015

It is observed that in 2014 the monitoring points determined, through the Grey Clustering method, a null contamination at each point except for point P7 which gives us an acceptable level according to the PRATI indices, this is explained because according to the Report IT 060-2014-ANA-CAÑETE the values of the points taken do not present high concentrations with respect to the ECA-water [7]. Another point to consider is that there is no contamination by mining companies because the 2014 report does not show that the metals present concentrations that exceed the ECA, it is corroborated that a correct characterization of the parameters was taken because they were not took the parameter of thermo-tolerant coliforms because in the report of the aforementioned year it indicates in its results and discussion the coliforms parameter does not represent a risk because these concentrations of each point did not exceed the ECA and also for a point not taken from This report exceeds the ECA but studies indicate that it will not generate health hazards because they are not conservative to their activity since it is influenced by their temperature, biological activity and the physicochemical composition of water bodies.

For the year 2015, it presents a variation to the previous year because points P1 and P2 indicate a moderately contaminated contamination, point P3 shows us an acceptable level, after point P4 to P9 it presents us with an uncontaminated level. According to IT 086-2015-ANA CAÑETE, the Mn parameter presents higher concentrations compared to the previous year that even exceeds the ECA. This report indicates that the variation is due to the geochemical nature of the basin, however, the number of discharges has not changed and the concentrations in 2014 were below the ECA. This variation may be influenced by the presence of 3 authorized discharges to the Cañete River basin that appear on IT 086-2015-ANA CAÑETE, these are from the companies.

LNG SRL, whose discharges are from a mixed chamber, sewage from the manhole and industrial wastewater, respectively, this variation from point P1 to P4 can be contrasted with IT 148-2019- ANA CAÑETE since it indicates that point one o P1 is located below a district called "Vitis" which is at the height of a fish farm, while from point P2 to P4 there is a decrease in pollution due to factors such as a mixture of sections between the Cañete and Alis rivers.

D. Joint Analysis

The graph presented in the results gives us an indication of a variation in points P1 to P4 and at the other extreme points P6 to P8, having a variation in the year 2014-2015, according to report IT 148- 2019-ANA CAÑETE these points are found in the vast majority below districts such as Vitis, Laraos and a rise and fall is generated because the other points are present or in the mixing zone of rivers or the mixing zone of the sections [15].

V. Conclusion

The Grey Clustering analysis method allowed determining the water quality in the 9 monitoring points of the upper-middle basin of the Cañete River in the years 2014 and 2015. Allowing to observe the reduction of the water quality in the points P1 and P2 for the period of the years 2014 and 2015 respectively. The first point is in the upper area of the Basin and the second point is at the confluence of the Alis and Cañete rivers. Thus, the loss of quality of the water resource is observed. The deterioration of the water quality in these points may cause, in the future, if the necessary mechanisms for the conservation of the water resource are not used, effects on the health of the surrounding populations and damage to their agriculture and livestock.

In the Grey Clustering methodology, in the stage of the evaluation of the Whitening Functions and use of the maximum Clusterization coefficients, it is observed that the points close to the point P2 have a water quality level cataloged as No Contaminated, however in the evaluations mentioned at the beginning of the paragraph it was observed that these points, P3 and P6, have a high probability of qualifying as Moderately Contaminated as they have Classification Coefficients close to these values. With this, it can be concluded that the points that are at the confluence of the Ríos Alis and Cañete are or are likely to vary their water quality, this due to the influence of the actions of the population found close to this area of the Cañete River.

The implementation of the IDW methodology complements what was developed in the Grey Clustering methodology, because it positions the quality level in a section by distributing it on a geographical map, generating a greater perception of the places that It has been altered and establishing a relationship between the present activities or geographical characteristics of the place, in which a greater analysis would be sought in the section in which it has been altered to verify the degree of impact and see a greater relationship against to the present activities.

References

[1] W. Zhuang, X. Zhao, F. Zhu, J. Liu, and H. Y. Xu, "Application of water quality evaluation model based on gray correlation analysis and artificial neural network algorithm," in Proceedings of 2017 9th International Conference On Modelling, Identification and Control, ICMIC 2017, Mar. 2018, vol. 2018-March, pp. 993–997, doi: 10.1109/ICMIC.2017.8321601.

[2] S. A. Kumar, H. Kumar, V. Dutt, and H. Soni, "Self-Health Analysis with Two Step Histogram based Procedure using Machine Learning," in 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Feb. 2021, pp. 794–799, doi: 10.1109/ICICV50876.2021.9388427.

[3] V. E. Cabrera, P. E. Hildebrand, and J. W. Jones, "Modelling the effect of household composition on the welfare of limited-resource farmers in Coastal Cañete, Peru," Agric. Syst., vol. 86, no. 2, pp. 207–222, Nov. 2005, doi: 10.1016/J.AGSY.2004.08.009.

[4] A. Delgado and I. Romero, "Social impact assessment on a hydrocarbon proyect using triangular whitenization weight functions," CACIDI 2016 - Congreso Aergentino de Ciencias de la Informatica y Desarrollos de Investigacion. 7785998, doi: 10.1109/CACIDI.2016.7785998.

[5] X. Q. Fu and Z. H. Zou, "Water Quality Evaluation of the Yellow River Basin Based on Gray Clustering Method," IOP Conf. Ser. Earth Environ. Sci., vol. 128, no. 1, 2018, doi: 10.1088/1755-1315/128/1/012139.

[6] A. Delgado and I. Romero, "Applying the Grey Systems Theory to Assess Social Impact from an Energy Project," in 2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Aug. 2018, pp. 1–4, doi: 10.1109/INTERCON.2018.8526372.

[7] V. Bax, W. Francesconi, and A. Delgado, "Land-use conflicts between biodiversity conservation and extractive industries in the Peruvian

Andes," J. Environ. Manage., vol. 232, 2019, doi: 10.1016/j.jenvman.2018.12.016.

[8] V. Duong, A. Ahmed, and O. Farook, "A model template green environment initiative for recycling plastic bottles with progressive entrepreneurship partnership," PICMET 2018 - Portl. Int. Conf. Manag. Eng. Technol. Manag. Technol. Entrep. Engine Econ. Growth, Proc., pp. 1–5, 2018, doi: 10.23919/PICMET.2018.8481786.

[9] A. Delgado and I. Romero, "Environmental conflict analysis on a hydrocarbon exploration project using the Shannon entropy," in Proceedings of the 2017 Electronic Congress, E-CON UNI 2017, Jun. 2017, vol. 2018-January, pp. 1–4, doi: 10.1109/ECON.2017.8247309.

[10] A. N. del A. D. de G. de C. de los R. Hídricos, "Resultado del monitoreo de la calidad de agua de la cuenca del río Cañete (realizado del 15 al 19 de Diciembre de 2014): Informe técnico," Aut. Nac. del Agua, 2014, Accessed: Oct. 27, 2021. [Online]. Available: https://repositorio.ana.gob.pe/handle/20.500.12543/2574.

[11] Sifeng Liu and Jeffrey Yi Lin Forrest, Grey Systems: Theory and Applications . 2014.

[12] S. Liu and Y. Lin, Grey Systems Theory and Applications, vol. 68. Springer Berlin Heidelberg, 2011.

[13] A. Delgado, A. Aguirre, E. Palomino, and G. Salazar,, "Applying triangular whitenization weight functions to assess water quality of main affluents of Rimac river," Proceedings of the 2017 Electronic Congress, E-CON UNI 2017, 2018-January, pp. 1-4, doi: 10.1109/ECON.2017.8247308.

[14] C. M. Viana, P. Abrantes, and J. Rocha, "Introductory Chapter: Geographic Information Systems and Science," Geogr. Inf. Syst. Sci., no. November, 2019, doi: 10.5772/intechopen.86121.

[15] S. C. Wu, K. Y. Ke, H. T. Lin, and Y. C. Tan, "Optimization of Groundwater Quality Monitoring Network Using Risk Assessment and Geostatistic Approach," Water Resour. Manag., vol. 31, no. 1, pp. 515–530, 2017, doi: 10.1007/s11269-016-1545-x.