

Applying Grey Clustering and Shannon's Entropy to Assess Sediment Quality from a Watershed

Alexi Delgado¹, Betsy Vilchez², Fabian Chipana³, Gerson Trejo⁴
Renato Acari⁵, Rony Camarena⁶, Víctor Galicia⁷, Chiara Carbajal⁸

Mining Engineering Section, Pontificia Universidad Católica del Perú, Lima, Peru¹
Environmental Engineering Department, National University of Engineering, Lima, Peru^{2, 3, 4, 5, 6, 7}
Administration Program, Universidad de Ciencias y Humanidades, Lima, Peru⁸

Abstract—The evaluation of the quality of sediments is a complex issue in the Peruvian reality, mainly because there is no sampling protocol or norm for comparison, which leads to the assessment of sediments without a comprehensive analysis of their quality. In the present study, the quality of the sediments in the upper basin of the Huarmey river was evaluated in 30 monitoring points and 7 parameters, which are: arsenic, cadmium, copper, chromium, mercury, lead and zinc, which were compared according to the standards recommended by the Environmental Quality Guidelines for Sediments in freshwater bodies of Canada (Canadian Environmental Quality Guidelines - CEQG, 2002. Sediment Quality Guidelines for Protection of Aquatic Life - Fresh water according to Canadian Council of Ministers of the Environment (CCME)). The results of the evaluation, by grey clustering method and Shannon entropy, showed that 13 monitoring points resulted in good sediment quality, 1 monitoring point had moderate quality and 16 monitoring points presented poor quality; therefore, it can be concluded that the effluents and discharges of the mining activities that take place in the aforementioned location have a negative impact on environmental quality. Finally, the results obtained can be of great help for OEFA, the regional government, the municipalities and any other body that has oversight functions, since they will allow them to be more objective and precise decisions.

Keywords—Grey clustering; sediment quality; Shannon entropy

I. INTRODUCTION

The water resource of the Huarmey river basin in Ancash, Peru, represents the vital element for the supply of population, agricultural, livestock, mining, energy and ecological use [1], being important its optimal, rational and sustainable use. However, due to the continuous complaints from the surrounding population expressing their discomfort over an alleged environmental impact on the water and sediment of the upper basin of the Huarmey river, an environmental monitoring was carried out, where there is a record of the existence of mining activity in exploration and operation stage (Minera Huinac SAC) [2].

In the upper basin of the Huarmey river, districts of La Merced, Aija, Huacllan and Succha, province of Aija, department of Ancash, Peru. The sediment quality assessment was carried out, for which there were 30 monitoring points carried out by the Organization for Environmental Assessment and Enforcement, a specialized technical organization in charge of compliance with Peruvian environmental regulations

[2]. It is important to mention that, at present, Peru does not have a sediment sampling protocol and nor does it have regulations or quality standards to evaluate this component. Therefore, standards recommended by the Environmental Quality Guidelines for Sediments in freshwater bodies of Canada (Canadian Environmental Quality Guidelines - CEQG, 2002) will be used. Sediment Quality Guidelines for Protection of Aquatic Life - Fresh water according to the Canadian Council of Ministers of the Environment [3].

For the evaluation of the sediment quality, we will use the Grey Clustering method, as well as the Shannon entropy. Grey Clustering is a methodology that is based on a theory of fuzzy sets and can be applied by grey incidence of matrices or whitenization functions because unlike traditional statistical methods [4], this methodology considers the uncertainty of the fuzzy type which is present in the environment within their analysis. For the case study, the "Center-point Triangular Whitenization Functions - CTWF" will be used, since the CTWF is mainly applied to test if the observation objects belong to predetermined classes, known as grey classes [5] as evidenced in the studies of selection of innovative strategies [6] and in the evaluation of the quality of sediments by grey incidence [7]. On the other hand, the Shannon entropy method is also an artificial intelligence approach developed by Claude E. Shannon (Shannon and Weaver, 1994) that addresses the uncertainty due to the dispersion of the data [8], therefore this method was also used to determine the weights of the evaluation criteria within the CTWF method [9].

Therefore, the specific objective of the present work is to analyze and value the quality of sediments by Grey method Clustering and Shannon entropy in the upper basin of the Huarmey river monitored in May 2016 based on the standards recommended by the Guidelines Environmental Quality for Sediments in freshwater bodies of Canada (Canadian Environmental Quality Guidelines - CEQG, 2002. Sediment Quality Guidelines for Protection of Aquatic Life - Fresh water according to Canadian Council of Ministers of the Environment [3].

Thus, on the following the study is formed by Section II which summarizes the literature review; Section III, in which the CTWF method is explained in detail. After Section IV, will be the section where the case study is described, then the results and their discussion are presented in Section V. Finally, the conclusions are presented in Section VI.

II. LITERATURE REVIEW

Delgado et al., in 2017, developed a research that studied the water quality of the Santa River, in which different points were analyzed according to the parameters established by MINAM-Peru (DS No 015-2015). In this sense, 21 monitoring points of the Santa River basin were analyzed. It was concluded that the grey clustering method showed interesting results and that it could be applied to other studies on water quality or the environment in general. In this regard, the results showed that 47.6% of the monitoring points presented a good quality of water for human consumption, which could be purified by applying a disinfection; 33.3% of the monitoring points presented a moderate quality of water for human consumption, which could be purified with a conventional treatment; and 19.1% of the monitoring points presented a low quality of water for human consumption, which could be purified by applying a special treatment [10].

Delgado, in 2020, in Peru, mentions that evaluating the quality of surface waters is a complex issue that involves the comprehensive analysis of several parameters that are altered by natural or anthropogenic causes. In this sense, the Grey Clustering method, which is based on the Grey Systems theory [11], and the Shannon Entropy, based on the artificial intelligence approach, provide an alternative to evaluate water quality in a comprehensive manner considering the uncertainty within the analysis. In the mentioned study, the water quality in the upper basin of the Huallaga River was evaluated taking into account the results of the monitoring of twenty-one points carried out by the National Water Authority 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 [12].

Environmental Assessment and Control Agency (OEFA by its Spanish acronym), in Peru, 2016 water and sediment quality monitoring was carried out from May 20 to 30 at 30 monitoring points belonging to the upper basin of the Huarney River, which is formed by the Llactún and La Merced rivers with their respective tributaries, that when joined form the Santiago River which receives the contribution of the stream of the same name, in which there is a record of the existence of mining activity in exploration and operation stage (Minera Huinac S. A.C.). Finally, downstream it takes the name of the Aija River, which receives the contributions of the Mallqui and Allma rivers, where concentrations of total arsenic were exceeded in the 30 monitoring points, copper in 26 points, mercury in 10 points, lead in 18 points, zinc in 25 points of the reference values of the Canadian standard [2].

Chu and Tan, in 2014, in China, carried out the analysis of 39 samples of surface sediments, from the coastal ocean of Jiangsu to evaluate their quality. Making use of the Grey Clustering method for its evaluation and generating results classified into three categories (clean, light pollution and intense pollution). Of the thirty-nine samples, there are eleven clean samples, twenty light pollution samples, and eight heavy pollution samples. When analyzing the underlying reasons, pollutants dumped into the sea due to increased industrial and

agricultural activities that contributed to the pollution. Therefore, more emphasis should be placed on the management of the surface tidal flat sediment environment, especially on the treatment of the pollution source to improve the sediment quality for the sustainable development of the coastal zone [7].

Delgado, in 2018, in Peru, states that the assessment of pollution and the quality of the air is a serious problem for big cities, considering the increasing pollution of the air. In this sense, the evaluation of this problem using the grey clustering method, which is based on the theory of the grey system, has great advantages since it considers this uncertainty within the analysis. In such study, an evaluation of air quality was carried out in three monitoring points located in three different districts of the city of Lima, Peru, which are San Martín de Porres, Carabayllo and Puente Piedra. The results revealed that the three monitoring points presented good air quality in accordance with Peruvian law. Nevertheless, this could be because the districts in which the monitoring points are located are relatively new. Finally, the results of this study could help local and central authorities to make the best decision on the evaluation of air quality [5].

National Water Authority (ANA by its Spanish acronym), in November 2015, evaluated the quality of the surface water of the Huarney river basin, concluding that the effects on the bodies of water belonging to and / or tributaries to the Huarney basin were located in the upper area of said basin, being the Montecristo, Huinac, Hercules and Santiago streams, as well as the Llactún river, which presented concentration values of the content of metals such as aluminum, arsenic, cadmium, copper, iron, manganese, lead and zinc, which exceed the value established in the ECA- Water, due to the fact that mining companies are installed in this area at the head of the basin [13].

The application of the Grey Clustering and Shannon Entropy methodology in sediment quality is an innovative method; this is due to the scarce existing bibliography that applies similar methodologies and the non-existence of sediment quality standards in Peru. That is why the importance of the present work lies in the application of these powerful methodologies in a new field of study, the Peruvian context, in addition, it allows perceiving the quality in an environmental component of the upper basin of the Huarney River, Peru.

III. METHODOLOGY

In this section, we will proceed to describe the center-point triangular whitenization weight functions (CTWF) method, which can be described as follows: First assume that there is a set of m objects, a set of n criteria and a set of s grey classes, according to the sample value ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$) Then the CTWF method is appreciated in a flowchart in Fig. 1 and are developed with the following steps [12], [14]–[16].

A. Step 1: Determining the Center Points

The criteria ranges are divided into three grey classes, which are: $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ and these values are determined using the Canadian sediment standard.

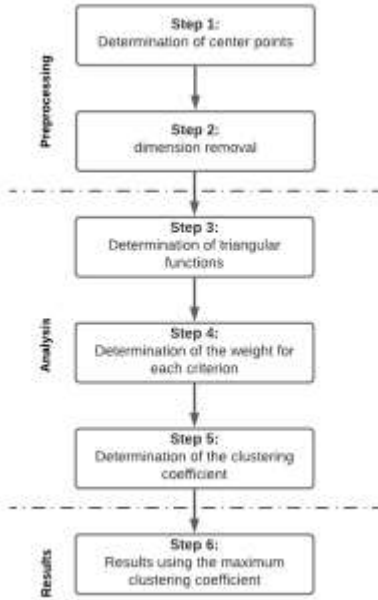


Fig. 1. Flowchart of the CTWF Method.

B. Step 2: Dimension Removal

At this point it is assumed that there are objects for evaluation and n criteria or grey classes, which form the following matrix $Z = \{Z_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$. In this sense it is normalized by each criterion C_j ($j = 1, 2, \dots, n$). The normalized value P_{ij} , which is calculated by (1).

$$P_{ij} = \frac{Z_{ij}}{\sum_{j=1}^n Z_{ij}} \quad (1)$$

C. Step 3: Determination of Triangular Functions

The grey classes are expanded in the directions of each parameter used and for this purpose the Canadian standard for sediments will be used as a reference, which provides values to measure the quality. In this research the Canadian standard provides us with three quality levels for each parameter analyzed, so we will have three functions for each parameter analyzed. The new sequence of center points in λ_1, λ_2 y λ_3 . For class $k = 1, 2, 3, j = 1, 2, \dots, n$ for an observed value x_{ij} . The computation of the Central Point Triangular Whitenization Functions (CTWF) is shown by (2) – (4). A visual representation is shown in Fig. 2.

$$f_j^1(x_{ij}) = \begin{cases} 1 & x \in [0, \lambda_j^1] \\ \frac{\lambda_j^2 - x}{\lambda_j^2 - \lambda_j^1} & x \in < \lambda_j^1, \lambda_j^2 > \\ 0 & x \in [\lambda_j^2, +\infty] \end{cases} \quad (2)$$

$$f_j^i(x_{ij}) = \begin{cases} \frac{x - \lambda_j^{i-1}}{\lambda_j^i - \lambda_j^{i-1}} & x \in < \lambda_j^{i-1}, \lambda_j^i > \\ \frac{\lambda_j^{i+1} - x}{\lambda_j^{i+1} - \lambda_j^i} & x \in < \lambda_j^i, \lambda_j^{i+1} > \\ 0 & x \in [0, \lambda_j^{i-1}] \cup [\lambda_j^{i+1}, +\infty] \end{cases} \quad (3)$$

$$f_j^n(x_{ij}) = \begin{cases} \frac{x - \lambda_j^{n-1}}{\lambda_j^n - \lambda_j^{n-1}} & x \in < \lambda_j^{n-1}, \lambda_j^n > \\ 1 & x \in [\lambda_j^n, +\infty > \\ 0 & x \in [0, \lambda_j^{n-1}] \end{cases} \quad (4)$$

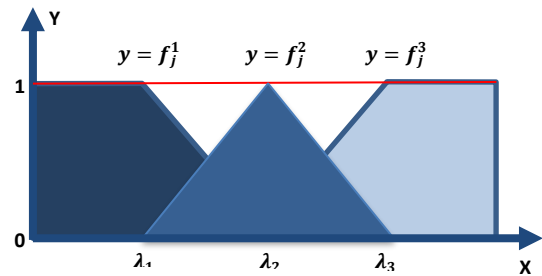


Fig. 2. CTWF Representation.

D. Step 4: Determining Weight for each Criterion

In this step, use is made of Shannon's Entropy weight method, which developed the measure "H", which satisfies the following properties [9], [12], [14], [15], [17]:

- H is a positive continuous function.
- If all p_i are equivalent $p_i = 1/n$, in that sense, H should be a growing monotonous function of n .
- For all $n \geq 2, H(P_1, P_2, \dots, P_n) = h(p_1 + p_2, p_3, \dots, P_n) + (p_1 + p_2)H(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2})$

Shannon shows that only functions satisfying this condition are computed by (5).

$$H_{Shannon} = -\sum_{i=1}^n p_i \log(p_i) \quad (5)$$

Where: $0 \leq p_i \leq 1; \sum_{i=1}^n p_i = 1$

Regarding the entropy weight methodology, it can be demonstrated according to the following definition [9], [12], [15]. As shown above, m objects are shown for evaluation and n evaluation criteria, which form the following matrix $x = \{x_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$. After that, the following steps will continue.

1) The matrix $x = \{x_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$ is normalized by each criterion C_j . The normalization evaluates P_{ij} and are calculated by (6).

$$f_j^1(x_{ij}) = P_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (6)$$

2) The entropy of each criterion is calculated by (7).

$$H = -k \sum_{i=1}^n p_{ij} \ln(p_{ij}) \quad (7)$$

Where, k is constant, $k = (\ln(m))^{-1}$

3) The degree of divergence of intrinsic information of each criterion C_j is calculated by (8).

$$div_j = 1 - H_j \quad (8)$$

4) The weight of the entropy of each criterion C_j is calculated by (9).

$$w_j = \frac{div_j}{\sum_{j=1}^n div_j} \quad (9)$$

Where, w_j it's equal to the n_j

E. Step 5: Determining the Clustering Coefficient

The clustering coefficient σ_i^k per object $i, i = 1, 2, \dots, m$, with respect to the grey classes $k, k = 1, 2, \dots, s$, is calculated by (10).

$$\sigma_i^k = \sum_{j=1}^n f_j^k(x_{ij}) n_j \quad (10)$$

Where $f_j^k(x_{ij})n_j$ is the CTWF of the k^{th} grey class of the j^{th} criterion, and n_j is the weight of criterion j , to establish these the Shannon Entropy method will be used.

F. Step 6: Results using the Maximum Clustering Coefficient

Finally, the value of $\max_{1 \leq k \leq s} \{\sigma_i^k\} = \sigma_i^k$ has to be calculated, based on that the object belongs to each grey class is opted. When there are several objects in some grey class, these objects can be ordered according to the magnitudes of their clustering coefficients.

IV. CASE STUDY

A. Description of the Study Area

The study area is in the upper basin of the Huarmey River, district of La Merced, Aija, Huacllan and Succha, province of Aija, department of Ancash, Peru as shown in Fig. 3. This is due to the results of water and sediment quality information, carried out from May 20 to 30, 2016. For which there were 30 monitoring points conducted by the Organization for Environmental Evaluation and Enforcement [2].



Fig. 3. Location of the Department of Ancash in Peru.

B. Description of Study Objects

For the evaluation of sediment quality in the upper basin of the Huarmey River La Merced, Aija, Huacllan and Succha districts, province of Aija, department of Ancash, information was collected from 30 monitoring points obtained from Report No. 266-2016 OEFA/DE- SDCA, carried out May 20 to 30, 2016 [2], and are shown in Table I and Fig. 4.

TABLE I. MONITORING POINTS IN THE UPPER HUARMEY RIVER BASIN

Point	Code	Coordinates UTM - WGS 84 18S		
		East (m)	North (m)	Elevation (m.a.s.l.)
1	SED-HUA1	206381	8925199	3901
2	SED-QHuin1	206427	8925324	3906
3	SED-RLlac1	206461	8925143	3893
4	SED-RLlac2	206506	8924857	3856
5	SED-HUA2	206585	8924591	3841
6	SED-RLlac3	206711	8924450	3818
7	SED-HUA3	206948	8923944	3813
8	SED-RLlac4	207197	8924040	3781
9	SED-HUA4	207347	8923936	3789
10	SED-RLlac5	207254	8923762	3757
11	SED-HUA5	207220	8923686	3769
12	SED-RLlac6	207363	8923601	3748
13	SED-RLlac10	209764	8921211	3411
14	SED-RLlac7	208328	8922234	3618
15	SED-RLlac8	208479	8922051	3584
16	SED-HUA6	208442	8922203	3620
17	SED-HUA8	209528	8922005	3367
18	SED-HUA10	211354	8921662	3446
19	SED-RLMer3	212160	8920813	3158
20	SED-HUA9	212696	8922670	3292
21	SED-RLMer1	212970	8923072	3305
22	SED-RLMer2	212762	8922439	3269
23	SED-RAija1	209384	8916087	2774
24	SED-RSant2	209581	8916347	2832
25	SED-Rall2	214163	8915170	3290
26	SED-RLlac11	211204	8918533	3040
27	SED-RLMer4	211293	8918643	3049
28	SED-QSant5	211348	8918603	3066
29	SED-RSant1	211155	8918424	3066
30	SED-RMall1	215369	8914339	3417

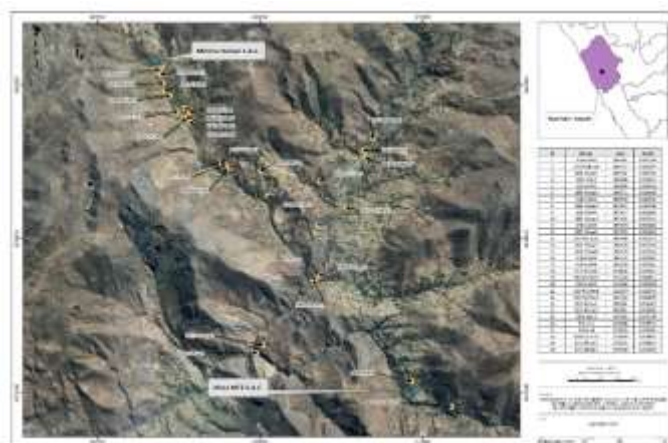


Fig. 4. Sediment Quality Monitoring Points in the upper Huarney River Basin.

C. Description of Evaluation Criteria

The evaluation criteria for the present study are determined by the Canadian sediment quality parameters, which are presented in Table II:

TABLE II. SEDIMENT QUALITY ASSESSMENT CRITERIA

Criteria	Units	Notation
Total Arsenic	mg/Kg	C ₁
Total Cadmium	mg/Kg	C ₂
Total Copper	mg/Kg	C ₃
Total chromium	mg/Kg	C ₄
Total Mercury	mg/Kg	C ₅
Total Plomium	mg/Kg	C ₆
Total Zinc	mg/Kg	C ₇

D. Definition of Grey Classes

The Grey Classes for the evaluation are 3 and are based on the sediment quality levels of the international standard - Canadian Sediment Quality Guidelines (CSQG) according to the Canadian Council of Ministers of the Environment [3] where it contemplates the limits to the ISQG (Interim Sediment Quality Guideline) values: the concentration below which no adverse biological effects are expected and PEL (Probable Effect Level): concentration above which adverse biological effects are frequently found. Proposing a third category considered as Moderate, which is presented in Table III.

E. Calculations using the CTWF Method

Step 1: Center points

Based on the international standard CSQG, the central values of the parameters to be analyzed were obtained and denominated as equivalent ISQG: Good and PEL: Poor. These values are shown in Table IV.

Step 2: Dimension removal

The non-dimension values for each parameter according to Table III on the average concentration of each parameter are shown in Table V.

TABLE III. QUALITY INTERVALS FOR THE INTERNATIONAL CSQG STANDARD

Parameter (mg/Kg)	Quality Index Condition		
	ISQG	Moderate	Pel
Total Arsenic	< 5.9	5.9 - 17	17 <
Total Cadmium	< 0.6	0.6 - 3.5	3.5 <
Total Copper	< 35.7	35.7 - 197	197 <
Total chromium	< 37.3	37.3 - 90	90 <
Total Mercury	< 0.17	0.17 - 0.486	0.486 <
Total Plomium	< 35	35 - 91.3	91.3 <
Total Zinc	< 123	123 - 315	315 <

TABLE IV. CENTRAL VALUES OF THE PARAMETERS FOR THE INTERNATIONAL STANDARD CSQG

Parameter (mg/Kg)	Quality Index Condition		
	Good	Moderate	Bad
Total Arsenic	5.9	11.45	17
Total Cadmium	0.6	2.05	3.5
Total Copper	35.7	116.35	197
Total chromium	37.3	63.65	90
Total Mercury	0.17	0.328	0.486
Total Plomium	35	63.15	91.3
Total Zinc	123	219	315

TABLE V. NON-DIMENSION VALUES FOR THE CSQG INTERNATIONAL STANDARD

Parameter (mg/Kg)	Quality Index Condition		
	Good	Moderate	Bad
Total Arsenic	0.515	1.000	1.485
Total Cadmium	0.293	1.000	1.707
Total Copper	0.307	1.000	1.693
Total chromium	0.586	1.000	1.414
Total Mercury	0.518	1.000	1.482
Total Plomium	0.554	1.000	1.446
Total Zinc	0.562	1.000	1.438

Similarly, based on the monitoring results obtained from Report No. 266-2016 OEFA/DE- SDCA, dimensionless values were obtained for the 30 selected monitoring points, as the example we show in Table VI for 5 monitoring points.

TABLE VI. DATA WITHOUT MONITORING DIMENSION IN THE CASE STUDY

Param eters	C1 - Total Arsenic	C2 - Total Cadmi um	C3 - Total copper	C4 - Total Chro mium	C5 - Total Merc ury	C6 - Total Lead	C7 - Total Zinc
1	1.43	0.24	0.39	0.28	0.30	0.31	0.66
2	49.96	4.79	8.38	0.18	4.33	8.90	8.04
3	33.45	7.51	4.63	0.31	3.08	9.15	12.17
4	44.10	3.71	4.37	0.28	2.53	52.89	7.46
5	1.48	0.55	0.46	0.46	0.15	0.56	1.31

Step 3: Determining triangular functions and their functions

Replacing the values obtained from Table VI in the equations of the Whitenization functions, as an example for total arsenic functions equations are shown in (11) – (13). Its graphic representation is displayed in Fig. 5.

$$f_1^1(x) = \begin{cases} 1, x \leq 0.515 \\ \frac{1.00-x}{0.485}, 0.515 \leq x \leq 1.00 \\ 0, x \geq 1.00 \end{cases} \quad (11)$$

$$f_1^2(x) = \begin{cases} \frac{x-0.515}{0.485}, 0.515 \leq x \leq 1.00 \\ \frac{1.485-x}{0.485}, 1.00 \leq x \leq 1.485 \\ 0, x \leq 0.515 \cup 1.485 \leq x \end{cases} \quad (12)$$

$$f_1^3(x) = \begin{cases} \frac{1.485-x}{0.485}, 1.00 \leq x \leq 1.485 \\ 1, x \geq 1.485 \\ 0, x \leq 1.00 \end{cases} \quad (13)$$

Step 4: Determining weight for each criterion

The clustering weight (n_j) of each parameter was determined with Shanon's entropy. The following procedure was followed for this purpose.

- 1) Standardized parameter values from the Canadian standard, which are presented in Table VII.
- 2) The entropy (H_j) of each criterion (C_j) was calculated through (7). The results are shown in Table VIII.

3) Finally, the entropy weights w_j were found according by using (9), and equated to the clustering weights n_i of each parameter. The values are presented in Table IX.

Step 5: Determining the clustering coefficient

The values of the clustering coefficients were calculated using (10). The results of the first 2 monitoring points are shown in Table X.

Step 6: Results using the maximum clustering coefficient

Finally, we calculate the value of $\max_{1 \leq k \leq s} \{\sigma_i^k\} = \sigma_i^k$ for each grey class according to each monitoring point by adding a comparison by quality scale and with it we get Table XI.

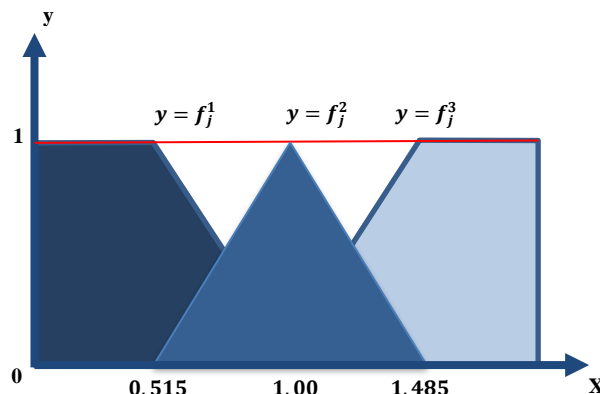


Fig. 5. Whitenization Functions for Total Arsenic.

TABLE VII. NORMALIZED VALUES FOR EACH PARAMETER

	C1	C2	C3	C4	C5	C6	C7
	Total arsenic	total cadmium	total copper	total chromium	total mercury	total lead	total zinc
Good	0.172	0.098	0.102	0.195	0.173	0.185	0.187
Moderate	0.333	0.333	0.333	0.333	0.333	0.333	0.333
Bad	0.495	0.569	0.564	0.471	0.494	0.482	0.479

TABLE VIII. DEGREE OF DIVERGENCE

Group	C1	C2	C3	C4	C5	C6	C7
H_j	0.93	0.83	0.84	0.95	0.93	0.94	0.94

TABLE IX. CLUSTERING WEIGHT OF EACH CRITERION

Group	C1	C2	C3	C4	C5	C6	C7
w_j	0.114	0.257	0.246	0.082	0.112	0.096	0.092

TABLE X. VALUES OF TRIANGULAR WHITENING FUNCTIONS (CTWF) FOR THE FIRST 2 MONITORING POINTS

P1	Criteria	C1	C2	C3	C4	C5	C6	C7	Results
	$f_1^1(x)$	0.000	1.000	0.882	1.000	1.000	1.000	0.771	0.114
	$f_1^2(x)$	0.108	0.000	0.118	0.000	0.000	0.000	0.229	0.021
	$f_1^3(x)$	0.892	0.000	0.000	0.000	0.000	0.000	0.000	0.050
P2	Criteria	C1	C2	C3	C4	C5	C6	C7	Results
	$f_2^1(x)$	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.016
	$f_2^2(x)$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$f_2^3(x)$	1.000	1.000	1.000	0.000	1.000	1.000	1.000	0.488

V. RESULTS AND DISCUSSION

TABLE XI. COMPARISON OF SEDIMENT QUALITY ACCORDING TO THEIR MAX VALUE

Results	Maxik	Quality	Quality scale
P11	0.133	Good	Better sediment quality
P17	0.123	Good	
P22	0.123	Good	
P21	0.121	Good	
P19	0.120	Good	
P7	0.119	Good	
P20	0.118	Good	
P27	0.114	Good	
P1	0.114	Good	
P16	0.113	Good	
P9	0.111	Good	
P18	0.099	Good	
P5	0.088	Good	Good sediment quality
P30	0.080	Moderate	
P2	0.488	Bad	Lower sediment quality
P3	0.488	Bad	
P4	0.488	Bad	
P15	0.472	Bad	
P13	0.454	Bad	
P28	0.454	Bad	
P23	0.450	Bad	
p8	0.432	Bad	
P12	0.432	Bad	
P26	0.432	Bad	
P6	0.417	Bad	
P29	0.336	Bad	
P10	0.303	Bad	
P14	0.290	Bad	
P24	0.269	Bad	
P25	0.131	Bad	Poor sediment quality

A. About the Case Study

It is observed in Table XII that 13 (43%) monitoring points resulted with the sediment quality good, while 1 monitoring point (3%) resulted with moderate quality and 16 (54%) monitoring points with poor quality. In addition, the comparison of quality level can be performed according to the maximum clustering coefficient ($\text{Max } \sigma_i^k$). In addition, we analyze according to each quality category into good, moderate and bad, respectively. For a better understanding further details regarding their location and differentiation are displayed in colors in Fig. 6.

- **Good category:** This means that no adverse biological effects are expected in these points and it is also observed that monitoring point P11 has the best water quality and point P5 the lowest quality within this category, this may happen because the points are located in tributaries to the main river and also the monitoring point P11 is more distant from the bridle path compared to the monitoring point P5 [2]. Likewise, the other points show good quality, possibly because they were sampled in tributaries of the main river.
- **Moderate category:** Means that at this point adverse biological effects are expected being monitoring point P30 the only point in this category, this may be because it is the farthest downstream point from the MTZ S.A.C mine compared to monitoring point P25 which is 100 m downstream of the mine [2].
- **Bad category:** It means that biological effects are frequently found in these points and also the monitoring point P25 presents the best sediment quality while point P2 presents the lowest sediment quality within this category, this could happen because the points are located within the main river where the mines discharge their effluents and also points P2, P3 and P4 are located downstream and closer to the Huinac mine, which is why they may have the lowest sediment quality compared to point P25, which has the best quality in the category since it is located 100 m downstream of the MTZ S.A.C. mine [2], but compared to points P2, P3 and P4, this mine has better control of its effluents.



Fig. 6. Results based on the Maximum Coefficient on the Case Study Map.

In relation to other studies, Chu & Tan [7] applied the Grey Clustering method to evaluate the surface sediment quality of the Jiangsu coastal ocean, and showed that more emphasis should be placed on the management of the surface tidal flat sediment environment, especially on the treatment of the pollution source to improve the sediment quality for the sustainable development of the coastal zone. In the water quality assessment conducted by Delgado et al., [12] in the Huallaga River basin, it was shown that of the 21 monitoring points analyzed all of these were found - according to the Pratti index - to be uncontaminated. A point to highlight from the previous study is that Shannon Entropy was used to determine the value of the weights, a feature that was not used in the study by Liping et al study [18] where the weights were determined using the arithmetic mean. Therefore, the present study is characterized by analyzing a topic that is little addressed from the perspectives of the Grey Clustering method, which is the analysis of sediments, but it also integrates a component that will provide greater objectivity and precision, Shannon Entropy, to determine the value of the weights.

B. Proposals for Poor Quality Points

According to the results and discussion on the case study, it is possible to evidence the contamination of sediment quality by heavy metals and that can be attributable to the effluents of the mining companies (mainly Huinac), being the monitoring points downstream of these, therefore, it is proposed that this

incorporates constant monitoring of sediment quality and thus also take into account the natural state of the streams and rivers that may be mineralized areas, and if responsibility is warranted by this to proceed to better control their effluents by treatment plants.

C. About the Methodology

The Grey Clustering method is a useful methodology to analyze the environmental system with respect to environmental quality, since among the advantages of Fuzzy Sets it is considered that it can be applied when the internal mechanisms of the system are unknown or the concept to be measured is imprecise, in this case of the study area. It is a mathematical theory of uncertainty to model situations where traditional instruments do not lead to optimal results due to the existence of uncertainty problems [19]–[21].

Based on the study conducted, the Grey Clustering Analysis Methodology (GCA) was chosen as the method for Environmental Quality Assessment over the Delphi methods [22] and the AHP method [23], where the advantages and disadvantages involving these mentioned methodologies are shown in Table XII.

The Shannon entropy method allows determining the clustering weights (n_i) for each criterion in an objective way, without the need to consult experts, which makes this method a more efficient and integrated method for sediment quality assessment.

TABLE XII. COMPARATIVE TABLE OF ADVANTAGES AND DISADVANTAGES OF ENVIRONMENTAL QUALITY ASSESSMENT METHODS IN THE FIELD OF RESEARCH

Environmental Quality Assessment Method	Advantages	Disadvantages
Grey Clustering Methodology	<ul style="list-style-type: none"> -Defines a new and reduced set of indicators, linear combinations of the initial ones, which allows comparison at all levels evaluated. -High degree of flexibility, due to the membership functions, which allows it to include uncertainty in its analysis. -It presents an objective weighting system. -It allows an integral analysis between the different thematic areas, by means of fuzzy inference rules. -Process a large set of data and reduce its dimensionality with a minimum loss of information [24]. 	The methodology has not been disseminated and its applications to environmental fields are recent.
Delphi Methodology	<ul style="list-style-type: none"> -Subjective experience and critical input. -Complex, large, multidisciplinary problems with considerable uncertainties. -Possibility of unexpected breakthroughs. -Particularly long-time frames. -Achieving consensus in areas of uncertainty or in situations lacking causality. -Focus on issues where multiple stakeholder groups are potentially involved. -Delphi studies can be relatively simple to design and flexible in the way those designs are combined [25]. 	<ul style="list-style-type: none"> -Causal models cannot be built or validated. -Opinions from a large group are required. -Deficiencies of the researcher or panel members may arise. -The researcher imposes his preconceived ideas on respondents.
Analytic Hierarchy Process (AHP) Methodology	<ul style="list-style-type: none"> -Allows a complex problem to be broken down and analyzed in parts. -Allows quantitative and qualitative criteria to be measured using a common scale. -It facilitates the understanding of the problem by the decision-maker or by those involved in the analysis stage. -It is easy to use and allows its solution to be complemented with mathematical optimization methods [26]. 	<ul style="list-style-type: none"> -Additional analysis is required to establish preconditions. Since minimum points of agreement among stakeholders regarding objectives, criteria, weights, etc. are required. -The analysis includes a certain degree of subjectivity. Because data from different sources are used, the analysis includes some degree of subjectivity.

VI. CONCLUSION

In the present quality study of the sediments of the surface water bodies of the upper basin of the Huarmey river, the 30 monitoring points of the basin could be classified, from which it was determined that 13 (43%) of the monitoring points presented good sediment quality, 1 (3%) moderate and 16 (54%) of poor quality. In the points that presented a good quality of sediments, it is attributed to the fact that they were located in tributaries of the main river, for the monitoring point that presented a moderate quality it can be attributed to being located at a significant distance from the MTZ SAC mine. Those that presented a poor quality of sediments, which corresponds to most of the points sampled, can be attributed to the fact that they are located within the main river, close to the activities of the mining industries (mainly Huinac) and the discharges of their effluents.

The study used the Grey Clustering method and Shannon's Entropy, as for the Grey Clustering methodology it turns out to be one of the most effective since it considers the uncertainty within the analysis; in addition, the analysis was enhanced with the Entropy methodology of Shannon, which allows developing the analysis process objectively, without the need for expert judgment. As a result of these, it allowed to generate classifications of the quality of sediments, which are pertinent in the application in the Peruvian context due to the lack of regulations regarding sediments.

This evaluation information obtained is relevant because it allows generating timely decision-making, in relation to the current context, by public entities and the central government with powers in the upper basin of the Huarmey River, Peru. And finally, the study serves as the basis for future research

regarding the quality of sediments, in addition, to be complemented with subsequent studies of characterization of water, soil and sediments by natural formations in the Huarmey river basin.

REFERENCES

- [1] INRENA and Administración Técnica del Distrito de Riego Casma - Huarmey, "Evaluación de los recursos hídricos en las cuencas de los ríos Casma, Culebras y Huarmey: estudio hidrológico en la cuenca del río Huarmey (Informe final)," Aut. Nac. del Agua, Dec. 2007.
- [2] OEFA, "Informe del monitoreo ambiental de calidad de agua y sedimento, realizado del 20 al 30 de mayo de 2016, en la Cuenca alta del río Huarmey, distritos La Merced, Aija, Huacllan y Succha, provincia de Aija, departamento de Ancash," Lima, Informe Técnico W 2.66-201 6-OEFA/DE-SDCA, Dec. 2016.
- [3] CCME, "Canadian Sediment Quality Guidelines for the Protection of Aquatic Life Freshwater," 2002.
- [4] S. Liu and Y. Lin, "Grey Incidence and Evaluations," in Grey Systems: Theory and Applications, S. Liu and Y. Lin, Eds. Berlin, Heidelberg: Springer, 2011, pp. 51–105.
- [5] A. Delgado, P. Montellanos, and J. Llave, "Air quality assessment in Lima city using the grey clustering method," 2018 IEEE Int. Conf. Autom. Congr. Chil. Assoc. Autom. Control(ICA-ACCA), 2018.
- [6] 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, doi: 10.1108/gS-11-2013-0027.
- [7] K. J. Chu and M. Tan, "Assessment of Sediment Quality in Jiangsu Coastal Ocean Based on Grey Clustering Method," Applied Mechanics and Materials, 2014.
- [8] 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, 2018-January, pp. 1-4, doi: 10.1109/ECON.2017.8247309.
- [9] X.-S. He, Q.-W. Fan, and X.-S. Yang, "Chapter 6 - Probability theory for analyzing nature-inspired algorithms," in Nature-Inspired Computation

- and Swarm Intelligence, X.-S. Yang, Ed. Academic Press, 2020, pp. 77–88.
- [10] A. Delgado, A. Aguirre, E. Palomino, and G. Salazar, “Applying triangular whitenization weight functions to assess water quality of main affluents of Rimac river,” in Proceedings of the 2017 Electronic Congress, E-CON UNI 2017, 2018-January, pp. 1-4, doi: 10.1109/ECON.2017.8247308.
- [11] A. Delgado and I. Romero, “Social impact assessment on a hydrocarbon proyect using triangular whitenization weight functions,” in CACIDI 2016 - Congreso Argentino de Ciencias de la Informatica y Desarrollos de Investigación, 7785998, doi: 10.1109/CACIDI.2016.7785998.
- [12] A. Delgado, J. Vidal, J. Castro, J. Felix, and J. Saenz, “Assessment of Surface Water Quality on the Upper Watershed of Huallaga River, in Peru, using Grey Systems and Shannon Entropy,” Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 11, 2020, doi: 10.14569/IJACSA.2020.0111156.
- [13] ANA, “Resultados del monitoreo participativo de la calidad de agua superficial de la cuenca del río y mar de Huarmey (realizado del 28 de octubre al 04 de noviembre de 2015): Informe técnico,” Aut. Nac. del Agua, Dec. 2015.
- [14] V. Bax, W. Francesconi, and A. Delgado, “Land-use conflicts between biodiversity conservation and extractive industries in the Peruvian Andes,” Journal of Environmental Management, vol. 232, pp. 1028–1036, Feb. 2019, doi: 10.1016/j.jenvman.2018.12.016.
- [15] S. Liu, Grey Systems: Theory and Applications, vol. 68. Berlin, Heidelberg: Springer, 2011.
- [16] S. Liu, Y. Yang, and J. Forrest, Grey Clustering Evaluation Models. 2017.
- [17] A. Shemshadi, H. Shirazi, M. Toreihi, and M. J. Tarokh, “A fuzzy VIKOR method for supplier selection based on entropy measure for objective weighting,” Expert Syst. Appl., vol. 38, no. 10, pp. 12160–12167, 2011.
- [18] W. Liping, L. Kunrong, and Z. Weibo, “Application of Grey Clustering method for water quality evaluation in fenchuan River Yan’an Baota Area,” in 2011 International Symposium on Water Resource and Environmental Protection, 2011.
- [19] R. E. Bellman and L. A. Zadeh, “Decision-Making in a Fuzzy Environment,” Manage. Sci., vol. 17, no. 4, p. B-141, Dec. 1970, doi: 10.1287/mnsc.17.4.B141.
- [20] L. A. Zadeh, “Fuzzy sets,” Inf. Control, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.
- [21] A. Delgado, F. Gil, J. Chullunquía, T. Valdivia, and C. Carbajal, “Water Quality Analysis in Mantaro River, Peru, Before and After the Tailing’s Accident Using the Grey Clustering Method,” Int. J. Adv. Sci. Eng. Inf. Technol., vol. 11, no. 3, pp. 917–922, 2021, doi: 10.18517/IJASEIT.11.3.11928.
- [22] J. R. Avella, “Delphi Panels: Research Design, Procedures, Advantages, and Challenges,” Int. J. Dr. Stud., vol. 11, pp. 305–321, Sep. 2016.
- [23] E. Vergara Maldonado, “Pautas para la selección de las técnicas AHP, PROMETHEE y ábaco de Régnier modificado,” 2010.
- [24] A. Delgado, J. A. G. Achata, J. A. B. Valdivia, J. C. J. S. Montes, and C. Carbajal, “Grey Clustering Method for Water Quality Assessment to Determine the Impact of Mining Company, Peru,” Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 4, pp. 557–564, 2021, doi: 10.14569/IJACSA.2021.0120471.
- [25] A. Delgado and H. Flor, “Selection of the best air purifier system to urban houses using AHP,” in 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies, CHILECON 2017 - Proceedings, 2017-January, pp. 1-4, doi: 10.1109/CHILECON.2017.8229622.
- [26] H. Li, Y. Fu, W. Tang, and W. Yang, “The application of grey clustering analysis on social impact assessment of natural forest protection project,” in Proceedings 2010 IEEE International Conference on Information Theory and Information Security, ICITIS 2010, 2010, pp. 776–780, doi: 10.1109/ICITIS.2010.5689689.