Study of the Effectiveness of the Removal of Heavy Metals from the Irrigation Canal with Cerium Oxide Nanoparticles Using Neutrosophic Statistics

Juan Clímaco O. Núñez[1,*], Becquer F. Camayo-Lapa[1], Ever F. Ingaruca-Álvarez[1], Erick A. Huamán Alvarado[1], Humberto Dax B. Mancilla[1], Julio Cesar Álvarez Orellana[1], Katia Ninozca F. Ledesma[2]

[1] National University of Central Perú.

[2]César Vallejo University. Perú.

Emails: ospino303@gmail.com; bcamayo@uncp.edu.pe; eingaruca@uncp.edu.pe; erickhuaman@uncp.edu.pe; hbonilla@uncp.edu.pe; jalvarez@uncp.edu.pe; Kfloresl1@ucvvirtual.edu.pe

Abstract

For the treatment of contamination produced by the variant presence of heavy metals such as lead (Pb), copper (Cu), zinc (Zn) and arsenic (As) in the waters of the Irrigation Canal of the Left Bank of the Mantaro River (CIMIRM is Spanish), a purification procedure was carried out using different doses of cerium oxide nanoparticles (CeO_2) and evaluating their effectiveness in the elimination of these metals in the aforementioned mass of water. As a first step, the water from the CIMIRM canal was characterized using Modular Ultraviolet-Visible Spectrophotometry techniques with high NIR sensitivity and Inductively Coupled Plasma Mass Spectrometry (ICP-MS), to measure the concentrations of heavy metals. Additionally, an analysis of the CeO_2 nanoparticles was carried out using techniques to confirm their size and structure. The efficacy of the treatment was determined statistically using a four-stage four-factor factorial design, comparing the differences in the control groups and target groups. The classic statistical test used is the Wilcoxon rank sum test. One of the problems of the simulation of the study carried out in the laboratory is the lack of accuracy because the concentration of heavy metals in the Mantaro River varies during the year. This is why a single crisp value is not enough to study the effectiveness of treatments. One solution to this problem is to use Neutrosophic Statistics, where the data is replaced by Neutrosophic Numbers or intervals instead of crisp values.

Keywords: Nanoparticles; Factorial Design; Wilcoxon rank sum test; Neutrosophic Statistics; Neutrosophic Number.

1 Introduction

Currently, the water quality of irrigation canals for agricultural use is affected by many contaminants especially heavy metals that accumulate in crop fields and are then transmitted to agricultural products, generating multiple diseases. In the capital of the Junín region of Peru, there is a canal (CIMIRM) located on the left bank of the Mantaro River that supplies water for agricultural use to several municipalities and districts of the province of Jauja. The concentration of heavy metals present in the different areas of the irrigation streams of the CIMIRM canal is difficult to reduce during their journey due to their highly turbulent flow, which makes it necessary to have an advanced removal mechanism due to the high concentration of heavy metals. Therefore, agricultural irrigation with poor-quality water has been identified as an important source of contamination of soil and groundwater by heavy metals.

For now, experimental removal methods have not yet been investigated to solve this problem of metal concentration in irrigation canals. An alternative solution to these problems requires modern techniques, such as the emerging use of nanomaterials with efficient removal properties. Regarding these problems, very little has been studied on the interaction of nanomaterials with different concentrations of heavy metals in the waters of irrigation canals.

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

In response to this problem, some research has been developed to establish the relationship between heavy metal nanomaterials, concerning removal, adhesion, and efficiency. According to a study, it was shown that the removal of heavy metals under controlled conditions using distilled water can be reduced by up to 75%; but they were not carried out with real water samples. With this premise, the present investigation emerges where the objective has been established to evaluate the efficiency of the removal of heavy metals Pb, Cu, Zn, and As, at different concentrations, assuming different contamination scenarios of an irrigation canal, with the removal properties of CeO_2 nanoparticles. Nanomaterials are considered an efficient tool for the elimination of organic and inorganic contaminants such as heavy metals present in wastewater. The variation in the concentration of nanoparticles indicates the profile of removal efficiencies to obtain relevant data on the interaction of nanoparticle-heavy metals, which provides better knowledge of the emptiness present in this application of nanotechnology in water treatment.

This article aims to verify the effectiveness of contaminated water treatment for the removal of heavy metals. One of the problems that exists is the annual variation in the concentration of heavy metals in the Mantaro River. That is why if we want to obtain precision in the results, we must have imprecise data that takes into account the maximum and minimum concentration values of each heavy metal in this river.

To achieve this objective, a statistical study is carried out with data in the form of neutrosophic numbers, where two samples studied within a laboratory are compared [1-3]. There is a set of samples, some of them are taken as a control sample and the rest are subjected to various treatments. These results are compared statistically to determine which is the best treatment. To gain precision, the crisp values are replaced by neutrosophic numbers that allow for simulating the annual fluctuation of concentrations of heavy metals in the river. This is part of the so-called Neutrosophic Statistics, which consists of the generalization of classical statistical methods for data or parameters in the form of intervals, or when the size of the population or sample is not exactly known [4-15]. Specifically in this article, we apply a four-factor factorial design applied to data in the form of neutrosophic numbers of heavy metal concentrations for different treatments, to select the most effective of them [16]. The Wilcoxon rank sum test was selected on these data to determine if there is a significant improvement when the treatments are compared [17, 18]. This is a non-parametric test that does not require any distribution of the data.

2 Materials and Methods

This section is dedicated to exposing the basic elements of Neutrosophic Statistics.

Definition 1: ([13]) Let X be a universe of discourse. A *Neutrosophic Set* (NS) is characterized by three membership functions, $u_A(x), r_A(x), v_A(x): X \rightarrow\ ]^-0, 1^+[$, which satisfy the condition $^-0 \le \inf u_A(x) + \inf r_A(x) + \inf v_A(x) \le \sup u_A(x) + \sup r_A(x) + \sup v_A(x) \le 3^+$ for all x∈X. $u_A(x), r_A(x), v_A(x)$ are the membership functions of truthfulness, indeterminacy, and falseness of x in A, respectively, and their images are standard or non-standard subsets of $]^-0, 1^+[$.

Definition 2: ([13]) Let X be a universe of discourse. A *Single-Valued Neutrosophic Set* (SVNS) A on X is a set of the form:

$$A = \{\langle x, u_A(x), r_A(x), v_A(x)\rangle : x \in X\} \quad (1)$$

Where $u_A, r_A, v_A: X \rightarrow [0,1]$, satisfy the condition $0 \le u_A(x) + r_A(x) + v_A(x) \le 3$ for all x∈X. $u_A(x), r_A(x)$, and $v_A(x)$ denote the membership functions of truthfulness, indeterminate, and falseness of x in A, respectively. For convenience, a *Single-Valued Neutrosophic Number* (SVNN) will be expressed as $A = (a, b, c)$, where $a, b, c \in [0,1]$ and satisfy $0 \le a + b + c \le 3$.

Neutrosophic Statistics extends classical statistics, such that we deal with set values rather than crisp values, [13].

Neutrosophic Descriptive Statistics is comprised of all techniques to summarize and describe the neutrosophic numerical data characteristics.

Neutrosophic Inferential Statistics consists of methods that allow the generalization from a neutrosophic sampling to a population from which the sample was selected.

Neutrosophic Data is the data that contains some indeterminacy. Similarly to classical statistics, it can be classified as:

- *Discrete neutrosophic data*, if the values are isolated points.
- *Continuous neutrosophic data*, if the values form one or more intervals.

Another classification is the following:

- *Quantitative (numerical) neutrosophic data*; for example, a number in the interval (we do not know

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

exactly), 47, 52, 67, or 69 (we do not know exactly);

- *Qualitative (categorical) neutrosophic data*; for example: blue or red (we don't know exactly), white, black or green or yellow (not knowing exactly).

The *univariate neutrosophic data* is a neutrosophic data that consists of observations on a neutrosophic single attribute.

Multivariable neutrosophic data is neutrosophic data that consists of observations on two or more attributes.

A *Neutrosophical Statistical Number* N has the form $N = d + I$, [1-3], where d is called *the determinate part* and I is called *the indeterminate part*.

A *Neutrosophic Frequency Distribution* is a table displaying the categories, frequencies, and relative frequencies with some indeterminacy. Most often, indeterminacy occurs due to imprecise, incomplete, or unknown data related to frequency. As a consequence, relative frequency becomes imprecise, incomplete, or unknown too.

Neutrosophic Survey Results are survey results that contain some indeterminacy.

A *Neutrosophic Population* is a population not well determined at the level of membership (i.e., not sure if some individuals belong or do not belong to the population).

A *simple random neutrosophic sample* of size n from a classical or neutrosophic population is a sample of n individuals such that at least one of them has some indeterminacy.

A *stratified random neutrosophic sampling* is the pollster groups of the (classical or neutrosophic) population by strata according to a classification. Then the pollster takes a random sample (of appropriate size according to a criterion) from each group. If there is some indeterminacy, we deal with neutrosophic sampling.

Additionally, we describe some concepts of interval calculus, which should be useful in this paper.

Given $N_1 = a_1 + b_1 I$ and $N_2 = a_2 + b_2 I$ two neutrosophic numbers, some operations between them are defined as follows, [1-3]:

$N_1 + N_2 = a_1 + a_2 + (b_1 + b_2)I$ (Addition),

$N_1 - N_2 = a_1 - a_2 + (b_1 - b_2)I$ (Difference),

$N_1 \times N_2 = a_1 a_2 + (a_1 b_2 + b_1 a_2 + b_1 b_2)I$ (Product),

$\frac{N_1}{N_2} = \frac{a_1 + b_1 I}{a_2 + b_2 I} = \frac{a_1}{a_2} + \frac{a_2 b_1 - a_1 b_2}{a_2(a_2 + b_2)} I$ (Division).

Additionally, given $I_1 = [a_1, b_1]$ and $I_2 = [a_2, b_2]$ we have the following operations between them ([1-3]):

1. $I_1 \leq I_2$ if and only if $a_1 \leq a_2$ and $b_1 \leq b_2$.
2. $I_1 + I_2 = [a_1 + a_2, b_1 + b_2]$ (Addition);
3. $I_1 - I_2 = [a_1 - b_2, b_1 - a_2]$ (Subtraction),
4. $I_1 \cdot I_2 = [\min\{a_1 \cdot b_1, a_1 \cdot b_2, a_2 \cdot b_1, a_2 \cdot b_2\}, \max\{a_1 \cdot b_1, a_1 \cdot b_2, a_2 \cdot b_1, a_2 \cdot b_2\}]$ (Product),
5. $\frac{I_1}{I_2} = \left[\frac{a_1}{b_1}, \frac{a_2}{b_2}\right]$, always that $0 \notin I_2$ (Division),
6. $\sqrt{I} = \left[\sqrt{a}, \sqrt{b}\right]$, always that $a \geq 0$ (Square root),
7. $I^n = \underbrace{I \cdot I \cdot \cdots \cdot I \cdot I}_{n \text{ times}}$.

3 Results

Spherical nanoparticles of Cerium Oxide purchased from the company MKNano in Canada with an approximate diameter of 25 nm were used. The hydrodynamic diameter was verified with the help of the Dynamic Light Scattering (DLS) equipment Nicomp Z3000, USA. For this analysis, nanoparticle solutions with concentrations between 10 and 40 mg/L were prepared, which were sonicated with the Branson ultrasonic homogenizer, SFX 550, USA, to achieve a 20% maximum intensity at 10 minutes to generate specific properties of environmental applications.

After characterization and knowing the concentrations of the heavy metals Pb, Cu, Zn, and As in the waters of the CIMIRM canal, synthetic water solutions with different concentrations of the four heavy metals were prepared in the laboratory, see Table 1.

Table 1: Concentration of synthetic water for the CIMIRM canal.

Solution	Species	Concentration (mg/L)
	Pb	0.02
1	Cu	0.02
	Zn	0.02

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

	As	0.02
2	Pb	1
	Cu	1
	Zn	1
	As	1
3	Pb	2
	Cu	2
	Zn	2
	As	2
4	Pb	8
	Cu	8
	Zn	8
	As	8

These four concentrations are studied separately, but the idea is to predict the behavior of the concentrations in the entire range of values including those that do not appear explicitly. Thus, each concentration is denoted by $\mu_{Pb} = 0.02 + 7.98I$, $\mu_{Cu} = 0.02 + 7.98I$, $\mu_{Zn} = 0.02 + 7.98I$, and $\mu_{As} = 0.02 + 7.98I$, where $I = [0, 1]$, is the concentration of lead, copper, zinc, and arsenic in the water, respectively. The results of the chemical processing of water with the four concentration values (0.02, 1, 2, and 8 mg/L) can be interpolated when it comes to some intermediate value, in this way the study is carried out with values in the form of neutrosophic numbers or intervals.

From the prepared solutions of nanoparticles of 10, 50, 250, and 1250 mg/L, 1 mL of each was taken with the help of a micropipette and poured into the test tubes containing the solutions of the four metals under investigation. A factorial research design was used. A *factorial design* of each level of one factor occurs associated with each level of each of the other factors [16]. For example, suppose there are three factors with levels $I = 3, J = 2$ and $K = 4$. Then, they have$3 \times 2 \times 4 = 24$ treatment combinations. If there are three observations per combination then 72 experimental units will be needed. When the factorial design is feasible to carry out, good economic advantages are obtained and it allows greater accuracy in joint evaluations when compared to evaluations done only one at a time.

In this problem, 16 experiments were carried out ($4 \times 4 = 16$ which means 4 treatments for each of the 4 heavy metals).

In this problem there are 4 solutions of different concentrations of microparticles applied to 4 samples with different concentrations of heavy metals, this is a total of 16 samples to study. The results were indicated in Figures 1-4:

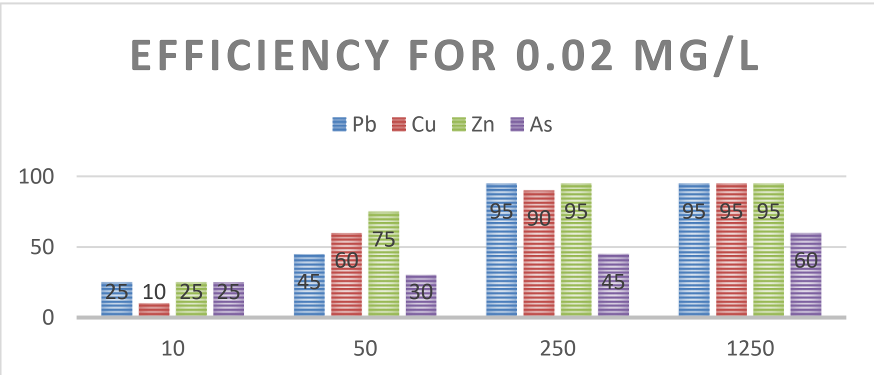

Figure 1: Removal efficiency of 0.02 mg of Pb, Cu, Zn, and As per liter (in percent) for concentrations of 10, 50, 250, and 1250 of CeO_2 microparticles.

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

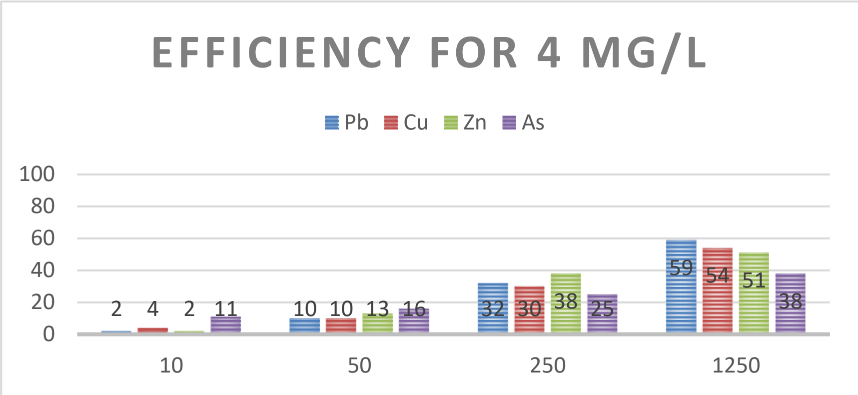

Figure 2: Removal efficiency of 1 mg of Pb, Cu, Zn, and As per liter (in percent) for concentrations of 10, 50, 250, and 1250 of CeO_2 microparticles.

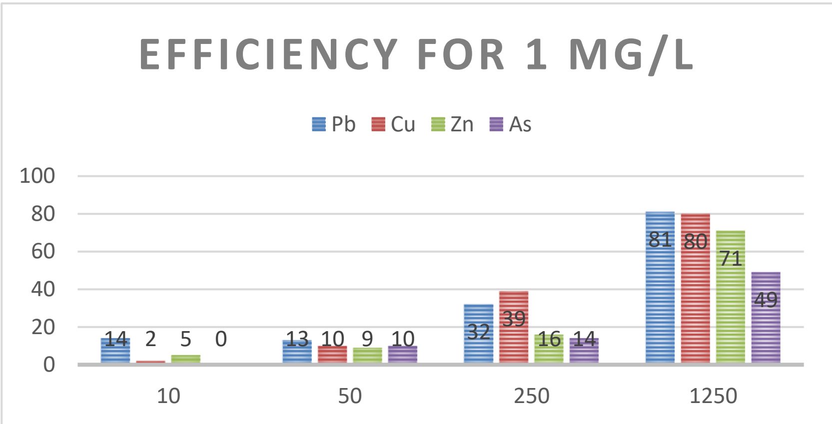

Figure 3: Removal efficiency of 4 mg of Pb, Cu, Zn, and As per liter (in percent) for concentrations of 10, 50, 250, and 1250 of CeO_2 microparticles.

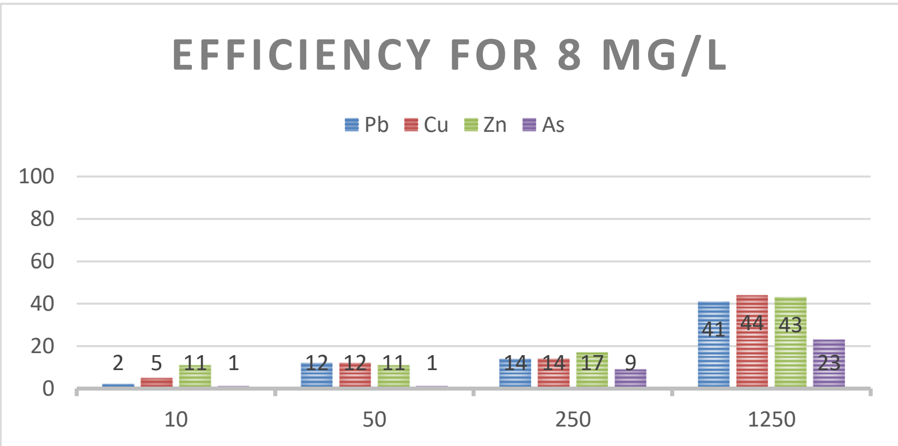

Figure 4: Removal efficiency of 8 mg of Pb, Cu, Zn, and As per liter (in percent) for concentrations of 10, 50, 250, and 1250 of CeO_2 microparticles.

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

The removal efficiency of each of the 4 heavy metals for each of the four treatments is summarized in the neutrosophic graphs that appear in Figures 5-8.

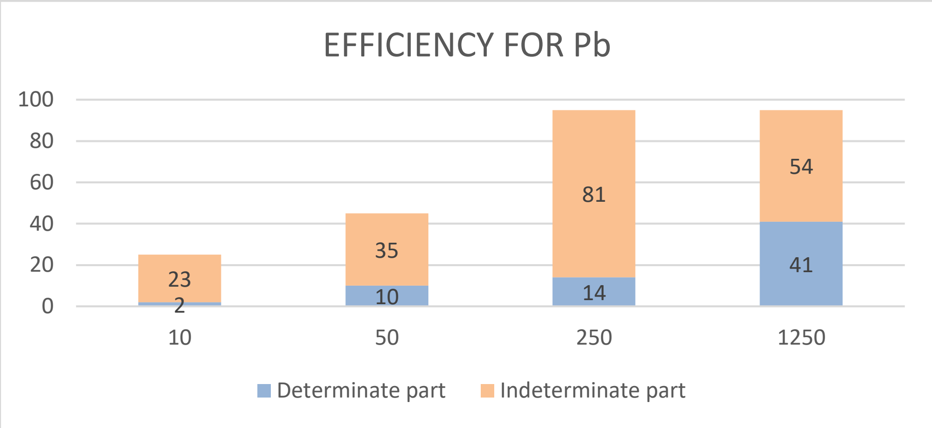

Figure 5: Efficiency shown in a neutrosophic graph for each of the lead treatments. The determinate part appears in blue and the indeterminate part appears in red.

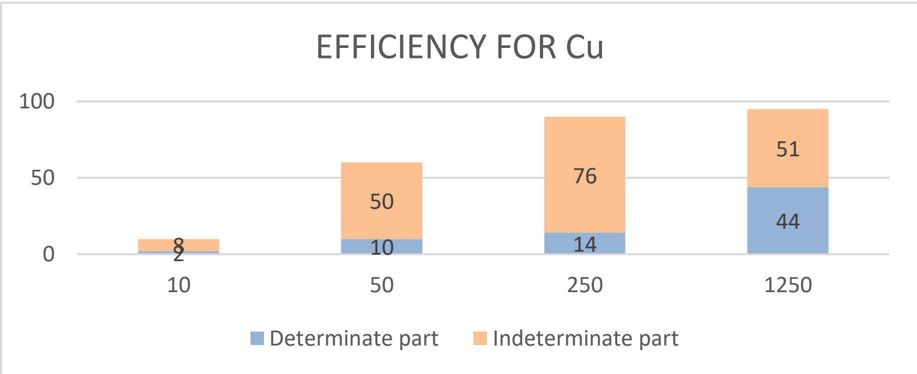

Figure 6: Efficiency shown in a neutrosophic graph for each of the copper treatments. The determinate part appears in blue and the indeterminate part appears in red.

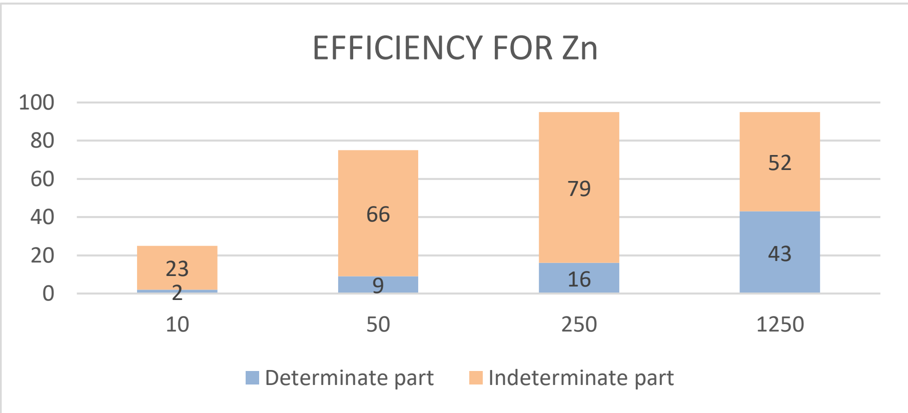

Figure 7: Efficiency shown in a neutrosophic graph for each of the zinc treatments. The determinate part appears in blue and the indeterminate part appears in red.

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

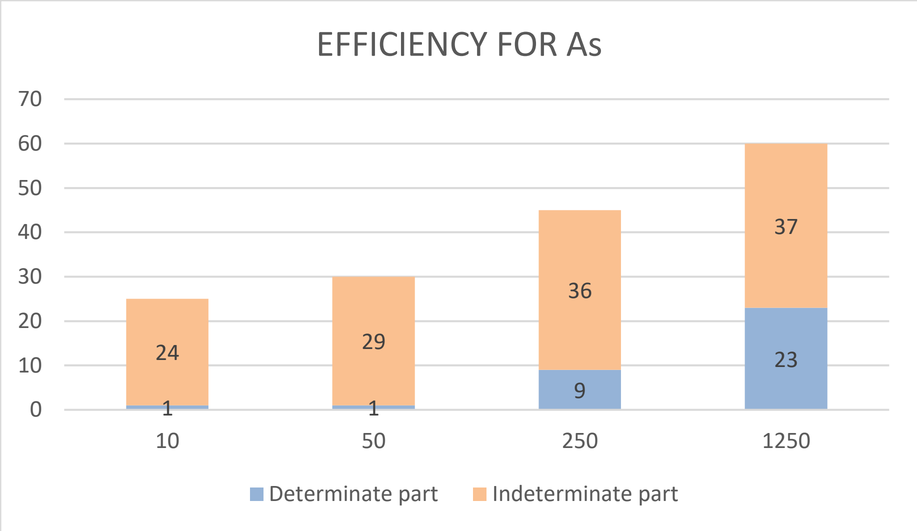

Figure 8: Efficiency shown in a neutrosophic graph for each of the arsenic treatments. The determinate part appears in blue and the indeterminate part appears in red.

The most effective treatment for the removal of each of the heavy metals is when the highest concentration of CeO_2 is used, which is 1250 mg/L. However, it is noted that in the case of arsenic, the efficiency reaches a maximum of 60%, although for high concentrations of heavy metals, the result is only 20%.

Let us statistically analyze the purification efficiency when there is a CeO_2 concentration equal to 1250 mg/L. Note that the results in Figures 5-8 give values with an indeterminate part greater than the determine part. We will proceed to find this indeterminacy in a more refined way. For this, we will follow the following procedure:

1. Four samples of concentration values were formed, one for each heavy metal. Each sample consists of 30 elements.

The sample corresponding to each heavy metal consists of a concentration amount in mg/L of the metal, obtained randomly according to the formula $0.02 + 7.98rand$ where *rand* is a random number between 0 and 1. The sampling is defined such that there are 3 random values in the range from 0 to 0.1, 3 in the range of 0.1 to 0.2, and so on. This allows for further refinement of the results.

In summary, we have the sample sets:

- A sample of size 30 with concentrations distributed as indicated above, with Pb concentrations. Another sample is obtained with the same procedure as this one.
- A sample 1 and a sample 2 with Cu concentrations following the same procedure.
- The same as the previous points with Zn.
- The above procedure is repeated with As.

2. The heavy metal removal procedure based on CeO_2 with 1250 mg/L is applied to each element of each sample.

The elements in sample 2 constitute control groups that undergo a traditional metal removal procedure.

3. The efficiency of each element of the four samples is calculated by applying the proposed method for sample 1 and the traditional one for sample 2.

4. Ten efficiency intervals are formed, each bounded by the minimum and maximum efficiency obtained from each triple.

5. Now there are 40 elements in the form of intervals for the proposed method and their corresponding 40 intervals with the control method.

6. The Wilcoxon rank sum test is applied between both samples [17, 18].

Let us explain this procedure in more detail. Taking the lead as an example, two samples with different concentrations are formed, let us call them Sample 1 and Sample 2. Sample 1 contains concentrations with three random values between 0.02 and 0.02+7.98(0.1), three values between 0.02+7.98(0.1) and 0.02+7.98(0.2), and so on, until reaching 30 elements in the sample.

The same is done with Sample 2.

The proposed method is applied to all the elements of Sample 1 and the efficiency is calculated. A traditional method is applied to the elements of Sample 2.

For Sample 1, an interval is formed $x_0 = [min(e_1, e_2, e_3), max(e_1, e_2, e_3)]$ where e_1, e_2, e_3 are the efficiency values for the three concentrations generated between 0.02 and 0.02+7.98(0.1). By repeating this procedure we

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

will have 10 data in the form of an interval.

For Sample 2 the mathematical procedure is the same, but chemically the traditional method of removing heavy metals is applied.

A unique Sample 1 is then created with all values from sample 1 and a unique Sample 2 with all values from sample 2.

Now the single Sample 1 contains 40 data in interval form and the same situation happens with the single Sample 2.

The Wilcoxon rank sum test is applied between both tests adapted to the data in the form of intervals [17, 18]. The final result obtained was a p-value in the form of an interval equal to $[0.0157, 0.0397] < 0.05$, therefore the null hypothesis is rejected and the proposed method is considered to be better than the original.

4. Conclusion

The removal of heavy metals from the Irrigation Canal of the Left Bank of the Mantaro River in Peru, especially lead, copper, zinc, and arsenic, is of vital importance to preserve the lives of the people who inhabit this place. These waters are used in agricultural irrigation. This article contains a statistical study of the efficiency results of applying a cerium oxide-based removal method. These results were studied in a laboratory simulating the conditions of this part of the river. However, it has been proven that there are changes in the concentrations of these metals during a year. To simulate this difference in concentrations we use Neutrosophic Statistics, where the data are expressed in the form of intervals or equivalently Neutrosophic Numbers. These data were compared with the efficiency results after applying a traditional method to control samples. It was concluded that the proposed method is more efficient and that is statistically significant. Specifically, the Wilcoxon rank sum test adapted to the data in interval form was applied.

References

[1] Chen, J., Ye, J., & Du, S. (2017). Scale effect and anisotropy analyzed for neutrosophic numbers of rock joint roughness coefficient based on neutrosophic statistics. Symmetry, 9, 208.

[2] Vasantha, W.B., Kandasamy, I. and Smarandache, F. (2018) Algebraic structure of neutrosophic duplets in neutrosophic rings. Infinite Study.

[3] Abodala, N. (2021). Foundations of Neutrosophic Number Theory. Neutrosophic Sets and Systems, 39, 120-132.

[4] V.Herrera, Alex. , G.Flores, Lenín. , L.G.Delgado, José. Treatment Alternatives to Gingival Hyperpigmentation Using Neutrosophic Correlation Coefficients. International Journal of Neutrosophic Science, vol. 18, no. 4, 2022, pp. 108-115.

[5] Patro, S. K., and Smarandache, F. (2016). The neutrosophic statistical distribution, more problems, more solutions. Infinite Study.

[6] Smarandache, F. (2022). Neutrosophic Statistics is an extension of Interval Statistics, while Plithogenic Statistics is the most general form of statistics (second version). International Journal of Neutrosophic Science, 19(1), 148-165.

[7] Aslam, M. (2020). Introducing Grubbs's test for detecting outliers under neutrosophic statistics–An application to medical data. Journal of King Saud University-Science, 32, 2696-2700.

[8] Khan, Z., Gulistan, M., Hashim, R., Yaqoob, N., and Chammam, W. (2020). Design of S-control chart for neutrosophic data: An application to manufacturing industry. Journal of Intelligent & Fuzzy Systems, 38, 4743-4751.

[9] Uktamova, D. (2024). Evaluating the Onshore Wind Farms based on Multi-Criteria Decision-Making Model under Neutrosophic Set. Journal of Journal of Sustainable Development and Green Technology, 4(2), 22-33. DOI: https://doi.org/10.54216/JSDGT.040203

[10] Aslam, M. (2021). Analyzing wind power data using analysis of means under neutrosophic statistics. Soft Computing, 25(10), 7087-7093.

[11] Duan, W. Q., Khan, Z., Gulistan, M., and Khurshid, A. (2021). Neutrosophic exponential distribution: modeling and applications for complex data analysis. Complexity, 2021, 1-8.

[12] H., A. Lucero, V. D., V. (2023). A Neutrosophic Analysis of Ecuador's Traffic Injury Legislation. Journal of International Journal of Neutrosophic Science, 20(1), 192-198. DOI: https://doi.org/10.54216/IJNS.200116

[13] Smarandache, F. (2022). Neutrosophic Statistics is an extension of Interval Statistics, while Plithogenic Statistics is the most general form of statistics (second version) (Vol. 2). Infinite Study.

[14] Woodall, W. H., Driscoll, A. R., and Montgomery, D. C. (2022). A review and perspective on neutrosophic statistical process monitoring methods. IEEE Access, 10, 100456-100462.

[15] Josefina, L. E., V. E., R. G., J. (2023). Study of the Effectiveness of Hypopressive Abdominal Training as

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024

Physiotherapeutic Treatment in Lumbosciatica in Adults Using Neutrosophic Statistics. International Journal of Neutrosophic Science, 21(4), 43-53
[16] Antony, J. (2023). Design of experiments for engineers and scientists. Elsevier.
[17] Lin, T., Chen, T., Liu, J., and Tu, X. M. (2021). Extending the Mann-Whitney-Wilcoxon rank sum test to survey data for comparing mean ranks. Statistics in Medicine, 40(7), 1705-1717.
[18] Kishore, K., and Jaswal, V. (2022). Statistics Corner: Wilcoxon-Mann-Whitney Test. Journal of Postgraduate Medicine, Education and Research, 56(4), 199-201.

DOI: https://doi.org/10.54216/IJNS.250120
Received: January 01, 2024 Revised: April 01, 2024 Accepted: June 28, 2024