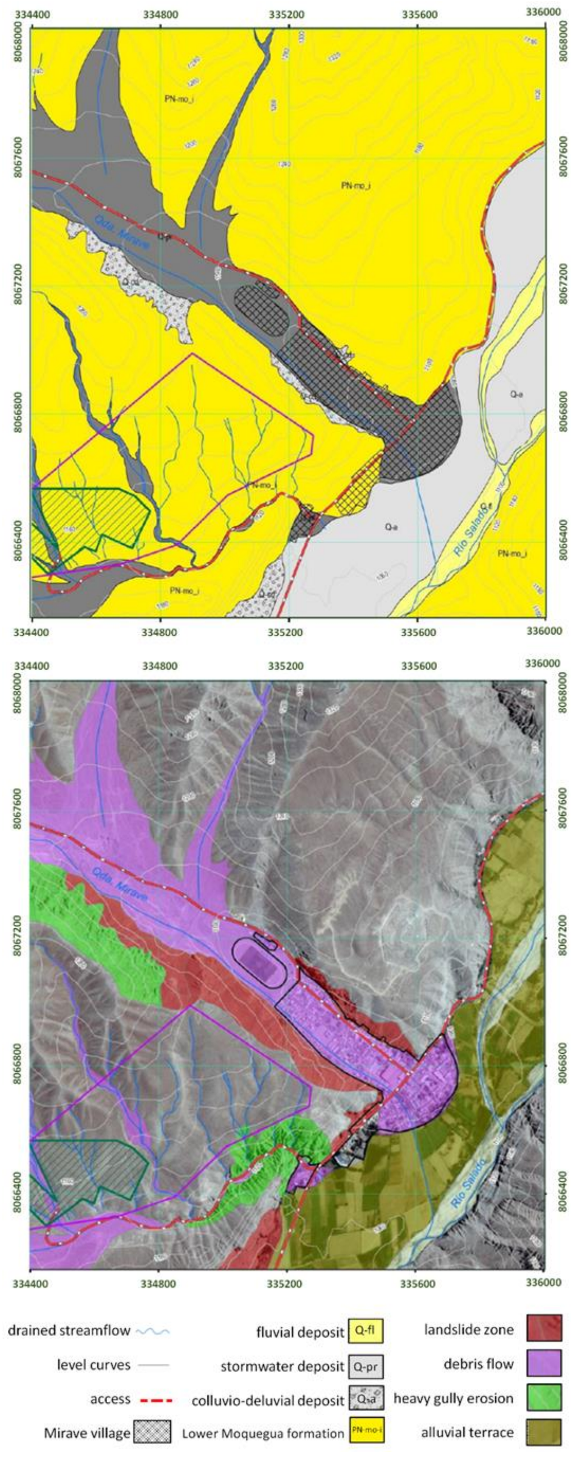
Revista Ambiente & Água
An Interdisciplinary Journal of Applied Science
UNIVERSIDADE DE TAUBATÉ
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AMBIENTAIS



ISSN = 1980-993X (Online)
http://www.ambi-agua.net

Library catalog entry by Liliane Castro CRB/8-6748

Revista Ambiente & Água - An Interdisciplinary Journal of Applied Science / Instituto de Pesquisas Ambientais em Bacias Hidrográficas. Taubaté. v. 17, n.2 (2006) - Taubaté: IPABHi, 2022.
Quadrimestral (2006 – 2013), Trimestral (2014 – 2016), Bimestral (2017), Publicação Contínua a partir de Janeiro de 2018.

Resumo em português e inglês.
ISSN 1980-993X

1. Ciências ambientais. 2. Recursos hídricos. I. Instituto de Pesquisas Ambientais em Bacias Hidrográficas.

CDD - 333.705
CDU - (03)556.18

TABLE OF CONTENTS

COVER:

This map shows both geological features and hazards as a result of mass movements in the Mirave Basin, Peru, which is classified as a hyperarid environment. This lithological map shows the recent alluvial and storm-water deposits resulting from debris flow. In this study, the evaluation of the source of debris flows in the Mirave Basin has identified that erosion processes on hillsides maximize geologic hazards and associated risks to human communities. Source: ORTEGA, J.C.B. *et al*. Effect of lithological and geotechnical characteristics on the generation of debris flows in the arid basin of Mirave, Peru. **Rev. Ambient. Água, Taubaté, vol. 17 n. 2, p. 1-18, 2022.** ***doi:10.4136/ambi-agua.2785***

ARTICLES

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Bioelectroremediation of hexadecane in electrical cells containing *Aspergillus niger* immobilized in alginate

ARTICLES doi:10.4136/ambi-agua.2805

Received: 05 Oct. 2021; Accepted: 16 Feb. 2022

Jayson Baldera Saavedra[1]; Lisbet Judith Noriega Pérez[1]; Claudio Eduardo Quiñones Cerna[2]; José Alfredo Cruz Monzón[3]; Fernando Javier Hurtado Butrón[4]; Mario Esparza Mantilla[5*]

[1]Departamento Académico de Ingeniería Ambiental. Universidad Nacional de Trujillo, Avenida Juan Pablo II, 13011, Trujillo, La Libertad, Peru. E-mail: ing.baldera@hotmail.com, lisbet.noriega.pe@gmail.com
[2]Facultad de Medicina Humana. Universidad Privada Antenor Orrego, Avenida América Sur 3145, 13008, Trujillo, La Libertad, Peru. E-mail: claudioqc24@gmail.com
[3]Departamento Académico de Química. Universidad Nacional de Trujillo, Avenida Juan Pablo II, 13011, Trujillo, La Libertad, Peru. E-mail: jcruzm@unitru.edu.pe
[4]Departamento Académico de Física. Universidad Nacional de Trujillo, Avenida Juan Pablo II, 13011, Trujillo, La Libertad, Peru. E-mail: fhurtado@unitru.edu.pe
[5]Laboratorio de Genética, Reproducción y Biología Molecular (GENERBIM). Escuela Profesional de Medicina Humana. Facultad de Medicina. Universidad Privada Antenor Orrego, Avenida América Sur 3145, 13008, Trujillo, La Libertad, Peru.
[*]Corresponding author. E-mail: mrodrigount@yahoo.com

ABSTRACT

Bioelectroremediation (BER) uses electrical current to stimulate catabolism of environmental pollutants, such as petroleum spills. However, applying current to soil may have adverse effects on the microorganisms involved in petroleum degradation. We identify and evaluate the BER capabilities of a strain of *Aspergillus niger* obtained from petroleum-contaminated soil. Spores of this strain immobilized in alginate spheres (2 g) were mixed with 100 g of hexadecane-contaminated sandy loam soil and exposed to 5, 10, or 15 mA direct current in a 200 cm^3 cell with copper electrodes. Soil hexadecane concentration was measured by gas chromatography. More than 94% of hexadecane was removed from the soil within 12 days for the currents tested, and the *A. niger* grew to 6 x10^6 CFU g^{-1} in 15 days at 10 mA current. The maximum hexadecane degradation was achieved using a 10-mA current for 20 days, but more than 99% of the hexadecane was removed by the fifth day. These results suggest that the use of spore-containing alginate beads promotes growth and petroleum biodegradation of *A. niger* exposed to electrical currents.

Keywords: *Aspergillus*, bioelectroremediation, hexadecane, petroleum, soil.

Bioeletrorremediação de hexadecano em células elétricas contendo *Aspergillus niger* imobilizado em alginato

RESUMO

A bioeletrorremediação (BER) usa corrente elétrica para estimular o catabolismo de poluentes ambientais, como derramamentos de petróleo. No entanto, a aplicação de corrente ao solo pode ter efeitos adversos sobre os microrganismos envolvidos na degradação do petróleo.

Identificamos e avaliamos as capacidades de BER de uma cepa de *Aspergillus niger* obtida de solo contaminado com petróleo. Os esporos desta cepa foram imobilizados em esferas de alginato (2 g) foram misturados com 100 g de solo franco-arenoso contaminado com hexadecano e expostos a 5, 10 ou 15 mA de corrente contínua em uma célula de 200 cm^3 com eletrodos de cobre. A concentração de hexadecano no solo foi medida por cromatografia gasosa. Mais de 94% do hexadecano foi removido do solo em 12 dias para as correntes testadas, e o A. niger cresceu para 6 x 10^6 UFC g^{-1} em 15 dias na corrente de 10 mA. A degradação máxima do hexadecano foi alcançada usando uma corrente de 10 mA por 20 dias, mas mais de 99% do hexadecano foi removido no quinto dia. Esses resultados sugerem que o uso de grânulos de alginato contendo esporos promove o crescimento e a biodegradação do petróleo de A. niger exposto a correntes elétricas.

Palavras-chave: *Aspergillus*, bioeletrorremediação, hexadecano, petróleo, solo.

1. INTRODUCTION

Given the long-lasting environmental effects of petroleum spills, developing procedures for rapid remediation of contaminated soils has attracted considerable attention. (Kumar *et al.*, 2021). These remediation procedures are divided into two broad classes - bioremediation and physicochemical technologies. Bioremediation may be the less expensive alternative, but it requires more time (Raj *et al.*, 2018). Thus, ways to improve the efficiency of bioremediation will likely make it more competitive and widely applied.

One promising advance in bioremediation is bioelectroremediation (BER), which involves stimulating biological degradation of contaminants with low-amperage direct current (Gidudu and Chirwa, 2019). It has been shown that the application of a current modifies the metabolic responses in microorganisms, causing acceleration of biodegradation (Sarankumar *et al.*, 2020). Furthermore, when used for petroleum-contaminated soil, BER increases bioavailability of petroleum-derived pollutants, thus making it less recalcitrant to microbial degradation (Fan *et al.*, 2017).

Among hydrocarbons, hexadecane (HXD), a common soil pollutant, is used frequently as a petroleum contamination proxy in bioremediation experiments (Dehghani *et al.*, 2013) because, like other hydrocarbons, it has low water solubility (0.9 µg L^{-1}) but is biologically labile (Okoye *et al.*, 2020). Using the HXD model to test traditional bioremediation has yielded some successful results using different microorganisms, such as *Pseudomonas aeruginosa* (Cruz *et al.*, 2021). Additionally, Yuan *et al.* (2013) achieved an HXD biodegradation of 53.7% using a bacterial consortium with a constant voltage gradient of 1.3 V cm^{-1} for 42 days in pristine soil; and Wang *et al.* (2016) used alternative bioremediation and electrokinetic technologies with soil bacteria to obtain 78.5% HXD degradation after 45 days of treatment. Using the fungus *Aspergillus brasiliensis*, Velasco-Alvarez *et al.* (2011) achieved a biodegradation of 96% with a current of 0.42 mA cm^{-2}.

While it is true BER has been successful in some laboratory-scale experiments, several parameters need to be explored further to create a coherent set of recommendations for successful BER use in the field. Therefore, evaluating ecological parameters, different microorganisms, pH, voltage gradient, electrochemical cell design, microorganism resistance to electric current, design and configuration of electrodes, costs, and energy efficiency are important factors for improving efficient elimination of pollutants and BER recommendations (Li, *et al.*, 2020; Annamalai and Sundaram, 2020).

To this end, we characterized a strain of *Aspergillus niger* isolated from petroleum-contaminated soil and evaluated its BER potential by exposing spores immobilized in alginate to different currents in HXD-containing soil.

2. MATERIAL AND METHODS

2.1. Isolation of *Aspergillus niger* strain from contaminated soil

Contaminated soil (1 kg) was collected from near Huamacucho, Peru (7°48'40.3" S 78°03'41.8" W, Figure 1A) where an oil spill occurred 3 days earlier. The soil (10 g) was combined with 100 ml of sterile saline solution (0.85% NaCl), homogenized, and allowed to settle at room temperature for 15 min. A 100 µL aliquot of the supernatant was spread on the surface of a petri dish containing potato dextrose agar (PDA, glucose 15 g L^{-1}, potato infusion 200 g L^{-1} and agar agar 15 g L^{-1}; pH 6) and incubated at 30ºC for 5 days. A colony with *Aspergillus*-like density and morphology with black conidia was selected and replicated once in tubes with PDA agar to obtain pure cultures (Al-Dossary *et al.*, 2019).

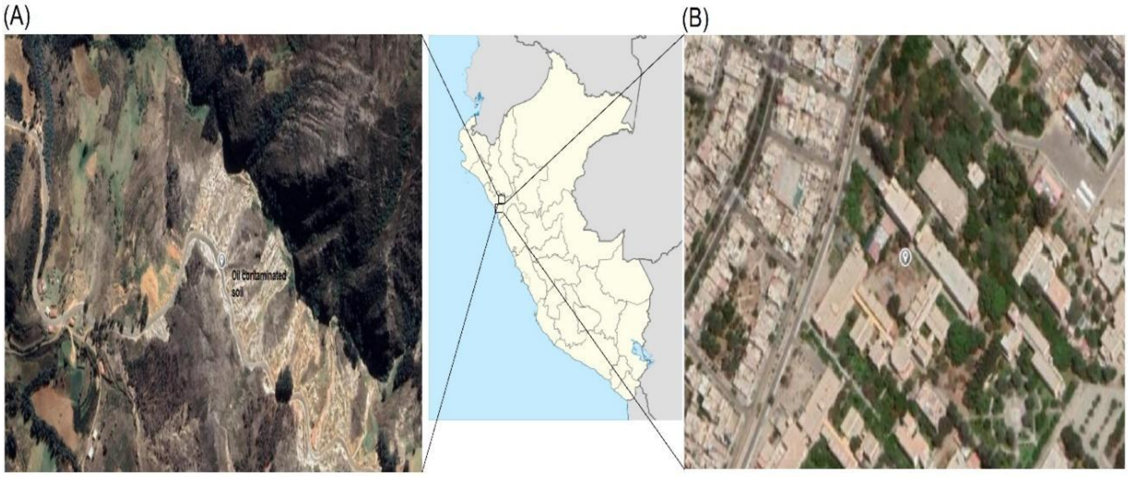


Figure 1. (A) Sampling site of oil-contaminated soil used in the isolation of *A. niger* near Huamachuco, Peru. (B) Location of the oil-free sampling site for the experimental trials, in Trujillo, Peru.
Source: Google Maps.

2.2. Morphological and molecular characterization of *A. niger*

A microculture of the fungus was used to observe microscopic morphological structures (conidiophore, vesicle and phialid) by the Riddell technique (Ordóñez-Valencia *et al.*, 2018). For genomic characterization, DNA was extracted from a pure 5-day old culture (Yeast DNA Extraction Kit, Thermo Scientific, Part No. 78870). As previously described, 18S rRNA sequences were amplified by PCR, using primers ITS-1 and ITS-4 and the amplified rDNA was characterized on a 1.5% agarose gel (Oduro-Mensah *et al.*, 2018). The purified PCR product was sequenced (Macrogen USA) and aligned to related known sequences using BLAST (National Centre for Biotechnology Information) (Buehler *et al.*, 2017). A phylogenetic tree was built from related rDNA sequences with MEGA-X (Kumar *et al.*, 2018).

2.3. Immobilization of fungal spores

A. niger spores were immobilized in calcium alginate spheres, according to the method described by Buehler *et al.* (2017). A spore suspension of the culture of *A. niger* containing $4x10^8$ CFU per mL in 1.5% Tween 80 (Sigma Aldrich, USA) was mixed 1:1 with sterile 3% sodium alginate solution (w v^{-1}). This mixture was added dropwise to sterile 0.2 M $CaCl_2$ using a 10 mL syringe. The resulting 3-5 mm spheres were rinsed with sterile water to discard excess calcium ions. The calcium alginate spheres contained $2x10^7$ UFC per gram and were preserved in a sterile solution of 0.1% NaCl at 4°C.

2.4. Addition of HXD to oil-free soil

Oil-free soil (4 kg) was obtained from the campus of National University of Trujillo, Peru (8°06′45.7″ S 79°02′19.1″ W) (Figure 1B), dried at room temperature, and homogenized with a mortar and sieved to a grain size of 1.15 µm (Table 1). The sieved soil was sterilized in an oven at 180°C for 1 h. The pH and conductivity of the soil was analyzed using a Consort 5010, while elemental composition was determined by atomic absorbance. A solution containing 180 mg of HXD (>98%, Merck, USA) dissolved in 1.9 mL n-heptane (≥ 99%, Merck, USA) was mixed thoroughly with 1 kg of soil. The soil was then spread on trays to let the n-heptane evaporate at room temperature for 24 hours (Velasco-Alvarez *et al.*, 2011).

2.5. Electrochemical chamber

The electrochemical chamber followed the specifications of Sumbarda-Ramos *et al.* (2010) with some modifications. The reaction chamber was a glass cylinder (external diameter of 5.0 cm, height of 9.8 cm). On either end of the reaction chamber was an electrolytic cell (diameter of 6.5 cm, height of 3.3 cm) containing a copper electrode with a contact area of 37.37 cm^2 (Figure 2). Microporous cellulose discs were used to facilitate ion exchange and current between the electrolytic cells and the reaction chamber. The chamber was sterilized with UV-A radiation before use.

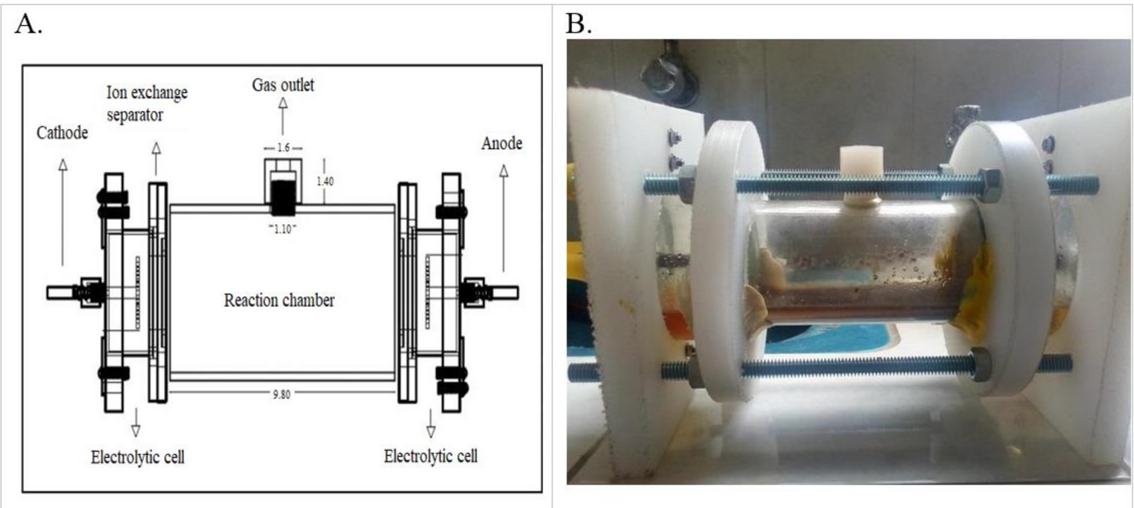


Figure 2. (A) Design and (B) photo of electrochemical chamber in use.

2.6. Bioelectrodegradation experiments

In each experimental run, 100 g of HXD-contaminated soil, 2 g of inoculated spheres, 43.5 mL of aqueous medium (250 mM $NaNO_3$, 22 mM KH_2PO_4, 7.5 mM $MgSO_4$•$7H_2O$, 40 mM KCl, 2.1 µM $FeSO_4$•$7H_2O$, 0.36 µM $CuSO_4$•$5H_2O$, 3.3 µM $ZnSO_4$•$7H_2O$, 0.17 µM $MnSO_4$•$7H_2O$, pH 5) was added to the reaction chamber (Volke-Sepúlveda *et al.*, 2006). Phosphate buffer (0.1 M, pH 7, 30 mL) was added to the electrolytic cells. The electrochemical chamber was hermetically secured with teflon and steel screws. Constant direct current (5, 10, or 15 mA) was applied to the electrodes for 6 or 12 days at room temperature (23±2ºC); current was monitored with a multimeter (Prasek, PR-301C). Controls omitted both current and inoculum but were otherwise treated the same as the experimental runs. The resulting soil samples (6 experimental and 2 control) were then analyzed for HXD content.

2.7. Determination of residual HXD

Soil samples were analyzed for HXD by EPA Method 8015C at an accredited laboratory (Corlan SAC, Lima, Perú) using a gas chromatograph (Perkin Elmer CLARUS 690 GC) with a capillary-type injector and a flame ionization detector with a sensitivity greater than 0.0131 C g^{-1}. Soil samples were tested in triplicate. The chromatographic data were analyzed by TotalChrom software (Lima *et al.*, 2020).

2.8. Statistical analysis

The results from different experimental conditions were compared using ANOVA and Tukey HSD tests by Minitab statistical software 18, with a level of significance $p < 0.05$ (Mena *et al.*, 2015).

2.9. Confirmation of optimized results

After completing the first set of experiments, a second series of confirmatory tests were completed using the same methods at a current of 10 mA and incubation times of 5, 10, 15, and 20 days with a similar control. In addition to HXD analysis of soil samples, 1 g of each soil sample was analyzed for *A. niger* colony-forming units (CFU) by serial dilution on Sabourand agar plates which were incubated at 30ºC for 5 days post-inoculation.

3. RESULTS AND DISCUSSION

3.1. Isolation of an *A. niger* strain

Soil contaminated with crude oil near Huamachuco, Peru was selected for isolation of an *A. niger* strain that has an active capacity to degrade hydrocarbons using standard purification techniques (Okoye *et al.*, 2020). This strain had macroscopic and microscopic morphology consistent with *A. niger*: a black globose aspergillar head, smooth brown conidiophores and globose spores or conidia. PCR amplification of the ITS-1 region of the 18S rRNA gene produced 612-617 bp oligonucleotides; sequences were compared to the Genbank database (BLAST). Phylogenetic analysis revealed that the isolated strain was 99% similar to *Aspergillus niger* (KY702576) (Figure 3). The 18S rRNA nucleotide sequence of the isolated strain was deposited in the GenBank database under the accession number MT180482 with designation QH27.

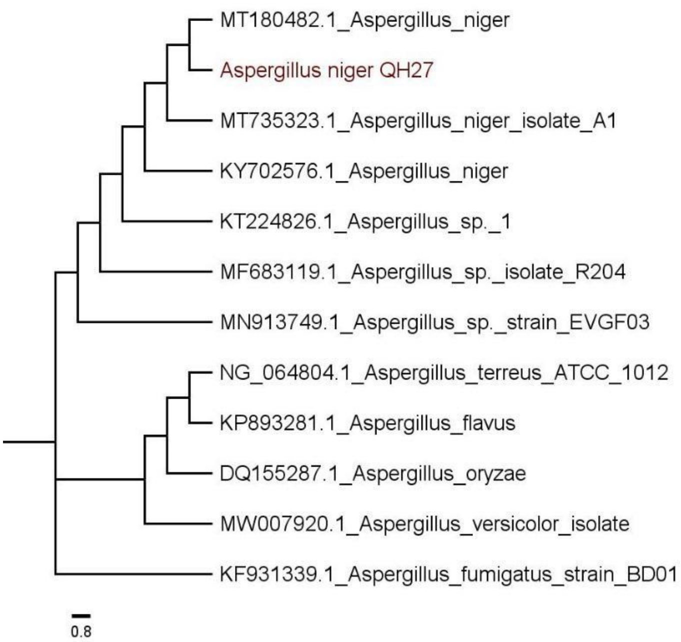


Figure 3. Phylogenetic analysis of 18S rRNA sequences from *Aspergillus niger* QH27.

3.2. Characterization of oil-free soil

In this investigation, a sandy loam soil with acceptable characteristics for BER was used (Table 1). Soil texture is acceptable for remediation experiments to evaluate organic compounds (Abdel-Moghny *et al.*, 2012). The soil sample had a total cation exchange capacity (CEC) of 11.59, which is appropriate for BER; but soils with CEC >15 experience physiochemical, hydrological, and mechanical changes when exposed to an electric field (Panjaitan and Andi, 2017).

Table 1. Physicochemical properties of the sandy loam soil used for BER studies.

Physical / chemical parameter	Value
pH (1:1)	8.0
C.E. (1:1) (dS/m)	0.8
CaCO3 (%)	8.4
M.O. (%)	1.1
P (ppm)	78
K (ppm)	427
Sand (%)	79
Silt (%)	10
Clay (%)	11
Textural class	Sandy loam
Ca^{+2} (cmol(+) kg^{-1})	7.42
Mg^{+2} (cmol(+) kg^{-1})	2.44
K^{+} (cmol(+) kg^{-1})	1.01
Na^{+} (cmol(+) kg^{-1})	0.72
Al^{+3} + H^{+} (cmol(+) kg^{-1})	0.00
Total Cation Exchange Capacity	11.59

3.3. Degradation of HXD using BER with *A. niger* QH27

BER results in several physical processes in soil: electrolysis, electromigration, electroosmosis and electrophoresis, which drag water, ions, compounds, colloids and/or microorganisms through the pores of the soil. Thus, soil components are more thoroughly mixed, causing the degradation of pollutants on the surface of the soil particles, and/or on the electrodes by electrochemical oxidation. More importantly, these processes increase the bioavailability of pollutants to microorganisms, and may also promote the transfer of pollutants and nutrients into cells (Ramadan *et al.*, 2018).

The degradation of HXD by BER using *A. niger* was monitored under 3 different currents (5, 10, and 15 mM) at two incubation times (6 and 12 days). A control that omitted both *A. niger* spheres and current was also analyzed at both timepoints. Each soil sample was analyzed in triplicate (Table 2). All initial experimental conditions resulted in a degradation greater than 75%, with a maximum degradation of 99.836 ± 0.002% when 10 mA was applied for 12 days (Table 2). The control showed negligible HXD loss.

BER results obtained here using fungal biomass differ from other investigations with bacteria, such as Hassan *et al.* (2019), who achieved a 20-30% degradation of diesel hydrocarbons with an electric current of approximately 220 mA, using *Acinetobacter calcoaceticis, Sphingobacterium multivorum* and *Sinorhizobium.* Likewise Vaishnavi *et al.* (2021) obtained a remediation of 84% of the diesel in soil using *Staphylococcus epidermidis* EVR4 with 2 V. Taken together, these data point to more efficient BER with fungus when compared to bacteria.

Table 2. HXD elimination by BER under different currents and endpoints. The % HXD elimination is the mean ($\bar{x}$) with standard deviation (σ) of 3 analyses of a soil sample.

Current (mA)	Time (days)	% HXD elimination ($\bar{x} \pm \sigma$)
5	6 12	89.972±0.002 98.612±0.003
10	6 12	99.3335±0.0004 99.836±0.002
15	6 12	79.964±0.007 94.833±0.001
0 (control)	6 12	0.001±0.001 0.001±0.002
10 (repeat)	5 10 15 20	99.278±0.0004 99.745±0.003 99.968±0.0003 99.994±0.0002
0 (control repeat)	20	0.0004±0.0002

One concern with BER is that the current may alter the microorganisms involved in bioremediation. For example, it has been shown that electrical current may inhibit cell differentiation, minimize biomass production and promote the formation of spores or CFU. In order to overcome this potential difficulty, *A. niger* spores were immobilized in alginate, which likely shortens the latency time during degradation by promoting cellular growth (Velasco-Alvarez *et al.*, 2017).

In order to test *A. niger* growth in the presence of electrical current and to confirm the BER results of the first experiments, a second series of experiments was performed (Table 2, bottom). These experiments used a constant current of 10 mA for 5 to 20 days and also monitored fungal growth. Fungal concentration steadily increased to 6×10^6 CFU g^{-1} after 15 days (Figure 4). Furthermore, HXD degradation was nearly complete by day 5, with the remaining <1% of HXD continuing to decrease (Table 2, bottom). These results confirm that the *A. niger* QH27 immobilized in alginate spheres both grows efficiently and uses HXD as a carbon source while exposed to electrical current (Habibul *et al.*, 2016).

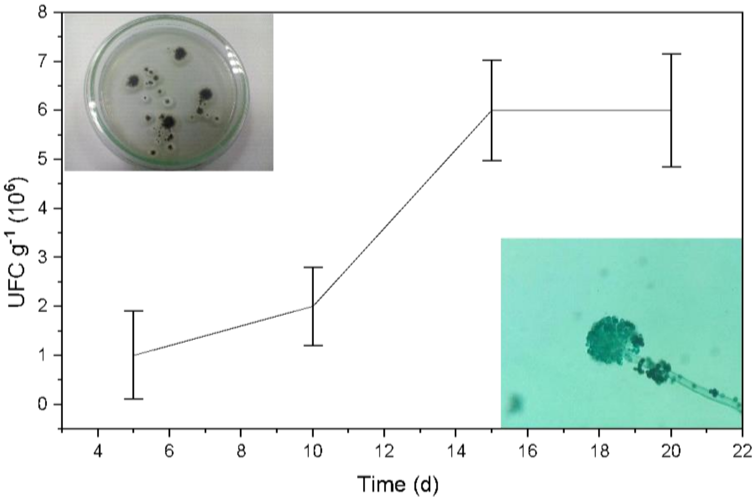


Figure 4. Growth curve of *A. niger* QH27 over 20 days exposed to 10 mA as measured by serial dilution (for example, Petri dish, top inset). *A. niger* was confirmed microscopically (bottom inset, 400x, Olympus CX31).

A one-way ANOVA ($F_{12,26}$=8.4x10^8, P=0) and Tukey's HSD test between runs and controls indicated statistically significant differences between all experimental runs tested, except for some control results. However, multiple repetitions of the entire experiment would likely increase error bars to roughly 10%. Therefore, nearly all treatments tested, except for 15 and 5 mA at 6 days, would almost completely remove HXD. Thus, an evaluation of energy consumption may be considered in choosing an appropriate current.

4. CONCLUSION

We identified a strain of *A. niger* that can rapidly grow and metabolize HXD as a carbon source while exposed to electrical current. Furthermore, we show that immobilizing spores in alginate spheres is one way to promote cellular growth and minimize latency. Under these conditions, *A. niger* QH27 can nearly eliminate HXD in loamy soils in as little as 5 days in an electrochemical chamber at 23ºC when exposed to 10 mA of direct current. This model opens the possibilities of testing additional strains and fungi for BER capacity exposed to a variety of pollutants, which may contribute to more efficient removal of contamination by hydrocarbons and their derivatives from soils.

5. REFERENCES

ABDEL-MOGHNY, T.; MOHAMED, R.; EL-SAYED, E.; MOHAMMED A.; SNOUSY, M. Effect of soil texture on remediation of hydrocarbons-contaminated soil at El-Minia District, Upper Egypt. **ISRN Chemical Engineering**, v. 2012, p. 406598, 2012. https://doi.org/10.5402/2012/406598

AL-DOSSARY, M.; ABOOD, S.; AL-SAAD, H. Biodegradation of crude oil using *Aspergillus* species. **Journal of Biology, Agriculture and Healthcare**, v. 9, n. 1 p. 60-64, 2019.

ANNAMALAI, S.; SUNDARAM, M. Electro-bioremediation: An Advanced Remediation technology for the treatment and management of contaminated soil. *In*: SAXENA, G.; BHARAGAVA, R. N. (eds.). **Bioremediation of Industrial Waste for Environmental Safety**. Singapore: Springer Nature, 2020. p. 183-214. https://doi.org/10.1007/978-981-13-3426-9_8

BUEHLER, A. J.; EVANOWSKI, R.; MARTIN, N.; BOOR, K.; WIEDMANN, M. Internal transcribed spacer (ITS) sequencing reveals considerable fungal diversity in dairy products. **Journal of Dairy Science**, v. 100, n. 11, p. 8814–8825, 2017. https://doi.org/10.3168/jds.2017-12635

CRUZ, J.; QUIÑONES, C.; SAAVEDRA, J.; URQUIZO, D.; ESPARZA, M. Biodegradation of phenol by *Pseudomonas aeruginosa* isolated from oil contaminated environments in Peru. **Bioscience Research**, v. 18, n. 2, p. 1294–1300, 2021.

DEHGHANI, M.; TAATIZADEH, S.; SAMAEI, M. Biodegradation of n-hexadecane in acinetobacter radioresistens liquid culture. **Health Scope**, v. 2, n. 3, p. 162–167, 2013. https://doi.org/10.17795/jhealthscope-14262

FAN, R.; GUO, S.; LI, T.; LI, F.; YANG, X.; WU, B. Contributions of electrokinetics and bioremediation in the treatment of different petroleum components. **Clean Soil Air Water Journal**, v. 43, n. 2, p. 251–259, 2017. https://doi.org/10.1002/clen.201300659

GIDUDU, B.; CHIRWA, E. M. N. Approaching electrodes configurations in bio-electrokinetic deoiling of petrochemical contaminated soil. **Chemical Engineering Transactions**, v. 76, p. 199–204, 2019. https://doi.org/10.3303/CET1976034

HABIBUL, N.; HU, Y.; SHENG, G. Microbial fuel cell driving electrokinetic remediation of toxic metal contaminated soils. **Journal of Hazardous Materials**, v. 318, p. 9–14, 2016. https://doi.org/10.1016/j.jhazmat.2016.06.041

HASSAN, I.; MOHAMEDELHASSAN, E.; YANFUL, E.; YUAN, Z. Mitigation of soil contaminated with diesel fuel using bioelectrokinetics. **Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering**, v. 54, n. 5, p. 416–426, 2019. https://doi.org/10.1080/10934529.2018.1558903

KUMAR, S.; RAI, A.; SINGH, R.; KUMAR, V.; KUMAR, D.; KUMAR, J., KUMAR A.; MALYAN, S. Bioelectroremediation technologies in remediation of environmental pollutants: challenges and future prospects. *In*: KRISHNAMOORTHY, S.; MANICKAM, P. (eds.). **Bioremediation for Environmental Sustainability**. Amsterdam: Elsevier, 2021. p. 147-165. https://doi.org/10.1016/b978-0-12-820318-7.00007-1

KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C.; TAMURA, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547–1549, 2018. https://doi.org/10.1093/molbev/msy096

LI, P.; WU, Z.; HU, H.; ZHANG, Y.; XIAO, T.; LU, X. *et al.* Efficient flexible perovskite solar cells using low-cost cu top and bottom electrodes. **ACS Applied Materials and Interfaces**, v. 12, n. 23, p. 26050–26059, 2020. https://doi.org/10.1021/acsami.0c06461

LIMA, S.; OLIVEIRA, A. F.; GOLIN, R.; LOPES, V. C.; CAIXETA, D. S.; LIMA, Z. M.; MORAIS, E. B. Isolation and characterization of hydrocarbon-degrading bacteria from gas station leaking-contaminated groundwater in the Southern Amazon, Brazil. **Brazilian Journal of Biology**, v. 80, n. 2, p. 354–361, 2020. https://doi.org/10.1590/1519-6984.208611

MENA, E., RUIZ, C.; VILLASEÑOR, J.; RODRIGO, M. A.; CAÑIZARES, P. Biological permeable reactive barriers coupled with electrokinetic soil flushing for the treatment of diesel-polluted clay soil. **Journal of Hazardous Materials**, v. 283, p. 131–139, 2015. https://doi.org/10.1016/j.jhazmat.2014.08.069

ODURO-MENSAH, D.; OCLOO, A.; LOWOR, S. T.; BONNEY, E. Y.; OKINE, L. K.; ADAMAFIO, N. A. Isolation and characterisation of theobromine-degrading filamentous fungi. **Microbiological Research**, v. 206, p. 16–24, 2018. https://doi.org/10.1016/j.micres.2017.09.006

OKOYE, A. U.; CHIKERE, C. B.; OKPOKWASILI, G. C. Isolation and characterization of hexadecane degrading bacteria from oil- polluted soil in Gio Community, Niger Delta, Nigeria. **Scientific African**, v. 9, p. 1-9, 2020. https://doi.org/10.1016/j.sciaf.2020.e00340

ORDÓÑEZ-VALENCIA, C.; FERRERA-CERRATO, R.; ALARCÓN, A.; HERNÁNDEZ-CUEVAS, L. V.; LARSEN, J. Early morphological development of sclerotia by Sclerotinia sclerotiorum in the presence of potassium bicarbonate. **Mexican Journal of Phytopathology**, v. 36, n. 3, p. 363–378, 2018. https://doi.org/10.18781/r.mex.fit.1803-4

PANJAITAN, N.; ANDI, A. Electrokinetic phenomena of cation exchange and its effect on the behaviour of expansive clays. **International Journal of GEOMATE**, v. 13, n. 38, p. 173–177, 2017. https://doi.org/10.21660/2017.38.74846

RAJ, K.; SARDAR, U. R.; BHARGAVI, E.; DEVI, I.; BHUNIA, B.; TIWARI, O. N. Advances in exopolysaccharides based bioremediation of heavy metals in soil and water: A critical review. **Carbohydrate Polymers**, v. 199, p. 353–364, 2018. https://doi.org/10.1016/j.carbpol.2018.07.037

RAMADAN, B. S.; SARI, G. L.; ROSMALINA, R. T.; EFFENDI, A. J.; HADRAH. An overview of electrokinetic soil flushing and its effect on bioremediation of hydrocarbon contaminated soil. **Journal of Environmental Management**, v. 218, p. 309–321, 2018. https://doi.org/10.1016/j.jenvman.2018.04.065

SARANKUMAR, R. K.; SELVI, A.; MURUGAN, K.; RAJASEKAR, A. Electrokinetic (EK) and Bio-electrokinetic (BEK) remediation of hexavalent chromium in contaminated soil using alkalophilic bio-anolyte. **Indian Geotechnical Journal**, v. 50, n. 3, p. 330–338, 2020. https://doi.org/10.1007/s40098-019-00366-6

SUMBARDA-RAMOS, E. G.; GUERRERO-GUTIERREZ, O. X.; MURILLO-RIVERA, B.; GONZÁLEZ, I.; OROPEZA-GUZMAN, M. T. Electrokinetic treatment for clayed and sandy soils. **Journal of Applied Electrochemistry**, v. 40, n. 6, p. 1255–1261, 2010. https://doi.org/10.1007/s10800-010-0097-7

VAISHNAVI, J.; DEVANESAN, S.; ALSALHI, M. S.; RAJASEKAR, A.; SELVI, A.; SRINIVASAN, P. *et al.* Biosurfactant mediated bioelectrokinetic remediation of diesel contaminated environment. **Chemosphere**, v. 264, p. 128377, 2021. https://doi.org/10.1016/j.chemosphere.2020.128377

VELASCO-ALVAREZ, N.; GONZÁLEZ, I.; DAMIAN-MATSUMURA, P.; GUTIÉRREZ-ROJAS, M. Enhanced hexadecane degradation and low biomass production by *Aspergillus niger* exposed to an electric current in a model system. **Bioresource Technology**, v. 102, n. 2, 1509–1515, 2011. https://doi.org/10.1016/j.biortech.2010.07.111

VELASCO-ALVAREZ, N.; GUTIÉRREZ-ROJAS, M.; GONZÁLEZ, I. Alterations in *Aspergillus brasiliensis* (*niger*) ATCC 9642 membranes associated to metabolism modifications during application of low-intensity electric current. **Bioelectrochemistry**, v. 118, p. 1–7, 2017. https://doi.org/10.1016/j.bioelechem.2017.06.008

VOLKE-SEPÚLVEDA, T.; GUTIÉRREZ-ROJAS, M.; FAVELA-TORRES, E. Biodegradation of high concentrations of hexadecane by *Aspergillus niger* in a solid-state system: Kinetic analysis. **Bioresource Technology**, v. 97, n. 14, p. 1583–1591, 2006. https://doi.org/10.1016/j.biortech.2005.07.026

WANG, S.; GUO, S.; LI, F., YANG, X.; TENG, F.; WANG, J. Effect of alternating bioremediation and electrokinetics on the remediation of n-hexadecane-contaminated soil. **Scientific Reports**, v. 6, p. 23833, 2016. https://doi.org/10.1038/srep23833

YUAN, Y.; GUO, S. H.; LI, F. M.; LI, T. T. Effect of an electric field on n-hexadecane microbial degradation in contaminated soil. **International Biodeterioration and Biodegradation**, v. 77, p. 78–84, 2013. https://doi.org/10.1016/j.ibiod.2012.10.012

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Indirect caffeine modeling in an urban river

ARTICLES doi:10.4136/ambi-agua.2810

Received: 04 Nov. 2021; Accepted: 16 Feb. 2022

Luis Otávio Miranda Peixoto[1*]; Luana Mayumi Takahasi Marques[2]; Alinne Mizukawa[1]; Julio Cesar Rodrigues de Azevedo[2]

[1]Departamento de Hidráulica e Saneamento. Universidade Federal do Paraná (UFPR), Avenida Coronel Francisco H. dos Santos, n° 100, Bloco H, CEP: 81530-000, Curitiba, PR, Brazil. E-mail: alimizu@gmail.com
[2]Departamento Acadêmico de Química e Biologia. Universidade Tecnológica do Paraná (UTFPR), Rua Deputado Heitor de Alencar Guimarães, n° 5000, Bloco EC, CEP: 81280-340, Curitiba, PR, Brazil.
E-mail: luanam@alunos.utfpr.edu.br, jcrazevedo@hotmail.com
[*]Corresponding author. E-mail: luisotaviopeixoto@gmail.com

ABSTRACT

Caffeine is used worldwide as a chemical tracer to identify anthropic pressures on urban water resources. Nevertheless, its quantification demands great financial investments. This research created a model that would indirectly determine a range of possible caffeine concentrations along an urban river, without the need for extensive laboratory work. The model is based on Canonical Correlation Analysis (CCA), which can correlate two sets of different-sized independent and dependent variables in order to generate a single empirical equation. This equation takes as input the concentrations of ammonia nitrogen and orthophosphate, as well as the total population and the population inhabiting irregular housing areas. From the model's results, it was possible to elaborate a spectrum of possible concentrations of caffeine along the Atuba River (Curitiba-Brazil). The tendency of water quality degradation of this river was also predicted. This model could become a useful preliminary analysis for water resource managers and researchers alike.

Keywords: caffeine, canonical correlation analysis, water quality modeling.

Modelagem indireta de cafeína em um rio urbano

RESUMO

A cafeína é utilizada mundialmente como um traçador químico para identificar pressões antrópicas em recursos hídricos urbanos. No entanto, a sua quantificação demanda grandes investimentos financeiros. Este estudo tem como objetivo criar um modelo que determinaria, indiretamente, uma banda de possíveis concentrações de cafeína ao longo de um rio urbano, sem a necessidade de esforço laboratorial. O modelo se baseia na análise de correlação canônica, que é capaz de correlacionar dois conjuntos de variáveis, dependentes e independentes, de diferentes dimensões e gerar uma equação que resumiria a relação. Esta equação utiliza como entrada as concentrações de nitrogênio ammoniacal e ortofosfato, e como saída a população total, e habitantes em zonas irregulares, assim como a concentração de cafeína. A partir do modelo, foi possível elaborar uma banda de possíveis concentrações de cafeína ao longo do rio. Este modelo possui a capacidade de ser utilizado como uma ferramenta preliminar para gestores e pesquisadores.

Palavras-chave: análise de correlação canônica, cafeína, qualidade das águas superficiais.

1. INTRODUCTION

The rise in urban population in the last decades has put increasing pressures on urban water bodies. These pressures include increased water demand, wastewater influx and pollution discharge onto these systems. The degradation of the water quality of urban rivers has become a challenge faced by society and academia when attempting to supply the population with water with enough quality and quantity (Woodhouse and Muller, 2017; Han *et al.*, 2018).

Caffeine is a stable compound under different environmental conditions. It is also very soluble and not volatile. It may be used as an indication of the discharge of wastewaters into an urban river. Due to these characteristics, several studies have indicated the viability of caffeine (1,3,7 – trimethylxanthine) as a chemical tracer for the identification of domestic wastewater discharge, compared with other microbiological or chemical indicators (Dafouz *et al.*, 2018; Mizukawa *et al.*, 2019). Also, there exists a possible correlation between high concentrations of caffeine and the presence of viral genome in natural waters has been observed (Gourmelon *et al.*, 2010, Sidhu *et al.*, 2013, Kumar *et al.*, 2019, Rimoldi *et al.*, 2020).

The determination of caffeine concentrations in natural waters can be costly and labor intensive. The machinery, the specialized personnel, the laboratory facilities and glassware, the chemical reagents and consumables are neither easy to acquire nor inexpensive (Colim *et al.*, 2019). Therefore, budgetary limitations constrain repeated caffeine analyses. To circumvent this problem, a model which could indirectly estimate caffeine concentrations in an urban river would be a useful tool for water resource managers and researchers, due to lower costs and faster assessment of the water quality situation. A statistical method that could be used in this context and for which can correlate different sets of multiple independent and dependent variables could be correlated and transformed into an equation is the Canonical Correlation Analysis (CCA) (Hotelling, 1936; Malacarne, 2014; Tiyasha *et al.*, 2020).

CCA has seen an increase in use in environmental sciences for its ability to identify and quantify possible associations among groups of different-sized sets of independent and dependent variables. Some examples of these properties: Gershunov *et al*. (2018) have used CCA to quantitatively assess the effect of the precipitation on water quality of coastal waters in the USA; Wei *et al.* (2018) have observed the relationship among 5 types of heavy metal pollution to 9 different traditional water quality parameters and concluded that for the case of the Dongtinghu Lake orthophosphate, E. Coli and the concentration of dissolved organic presented the highest correlation to the heavy metal pollution; Khalil *et al.* (2011) have implemented a CCA-based neural network to model and forecast water quality parameters of the Nile River, Egypt, using the rainfall in 50 different sub catchments of the river as input.

This study aims to present a model that can predict caffeine concentrations along an urban river, using the Atuba River as the study area. The parameters used as inputs for this model are the concentrations of ammonia nitrogen and orthophosphate and the total population (and the population in irregular housing) in this river's basin. The concentrations for the compounds were set as being dependent on the socio economic parameters. Such a model could improve the understanding of the behavior of contaminants along an urban river and its relationship to social characteristics of the population.

2. MATERIAL AND METHODS

2.1. STUDY AREA

The Atuba River Basin is located entirely within the Curitiba Metropolitan Area (CMA) in the State of Paraná (Southern Brazil). It has a total area of 129.94 km^2. Its main course (the

Atuba River) is 21.57 km long. About 562,700 people inhabited this basin at the time of the sampling. This population is spread through 4 different municipalities (Curitiba, Pinhais, Almirante Tamandare, and Colombo).

The Atuba River was chosen for this study for three reasons: 1) The river's basin is totally urbanized; 2) Though it occupies only 0.8% of the CMA's total area, it is responsible for 15% of the CMA's total gross income; 3) It is socially and physically diverse along its main course.

In early April/2019, 20 surface water samples were collected from the river. The locations of the sampling sites are shown in Figure 1**.** They were designated from P1 to P20, by their proximity to the river's spring. The zone of hydrological influence of each sampling site is also presented. These were used to establish the population of influence for each site.

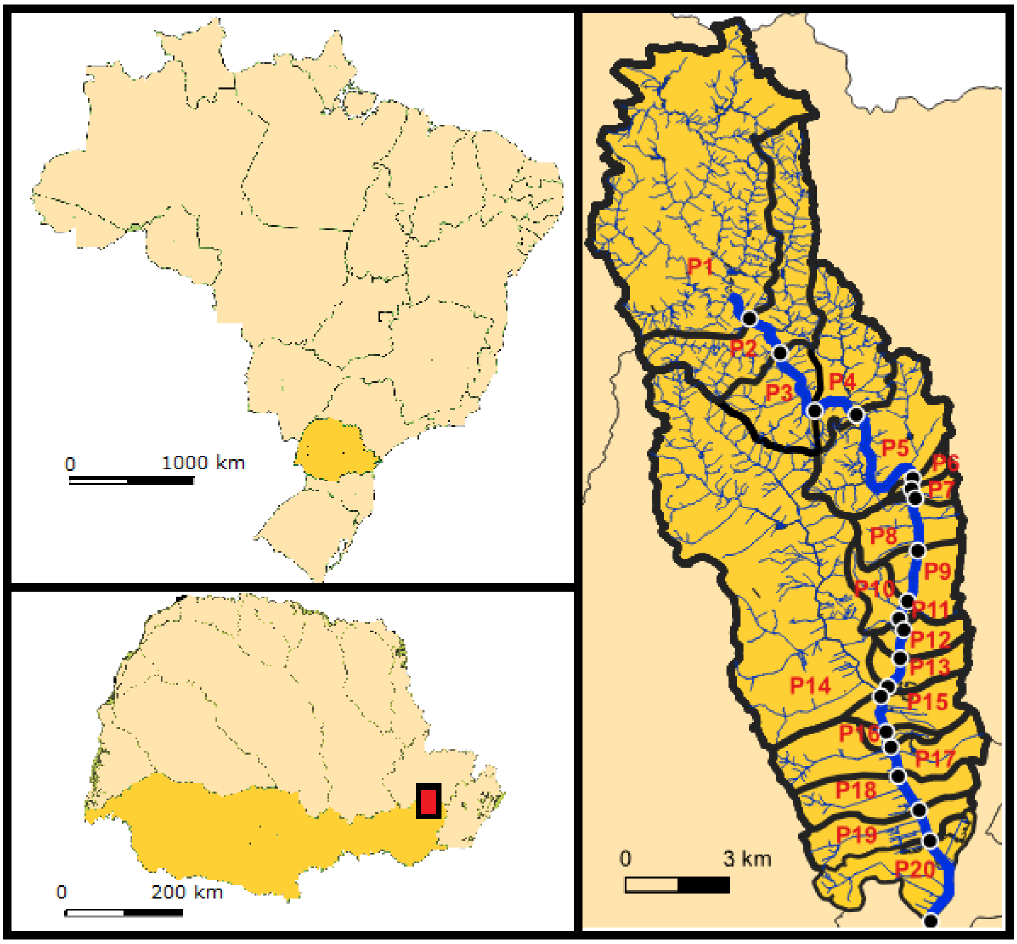


Figure 1. a) Location of the State of Paraná within Brazil; b) Location of the Iguassu River Basin (yellow) and the Atuba River Basin (red) within the State of Parana; c) Location of the sampling sites, and zones of hydrological influence, on the Atuba River Basin.

The socio economic profiles of the inhabitants are unevenly distributed throughout the length of the river. Table 1 presents the distribution of the population for each of the sampling sites, as well as the distribution of the residents that inhabit irregular housing within each zone of hydrological influence.

The socio economic diversities of the profiles within the zones of hydrological influence for the 20 sampling sites may be divided into three specific regions: a) zones for sites P1 to P10 are characterized by sparse regularized horizontal housing; b) zones for sites P10 to P16 are characterized by denser high-income vertical housing; c) zones for sites P17 to P20 are characterized by denser low-income irregular horizontal housing.

Table 1. Distribution of total population and inhabitants is irregular for the zones of hydrological influence of the 20 sampling sites along the Atuba River.

Site	POP	APOP	POPi	APOPi	Site	POP	APOP	POPi	APOPi
P1	58,526	58,526	0	0	**P11**	498	264,423	86	3,780
P2	36,966	95,492	841	841	**P12**	8,768	273,191	263	4,043
P3	21,329	116,821	560	1,401	**P13**	13,068	286,259	194	4,237
P4	35,676	152,497	312	1,713	**P14**	133,831	420,090	2,885	7,122
P5	45,075	197,572	549	2,262	**P15**	15,687	435,777	506	7,628
P6	3,405	200,977	91	2,353	**P16**	8,407	444,184	550	8,178
P7	6,460	207,437	162	2,515	**P17**	43,525	487,709	2,687	10,865
P8	17,912	225,349	716	3,231	**P18**	33,241	520,950	2,943	13,808
P9	19,390	244,739	270	3,501	**P19**	25,320	546,270	4,269	18,077
P10	19,186	263,925	193	3,694	**P20**	16,430	562,700	2,038	20,115

POP = Total inhabitants; APOP = Accumulated Population; POPi = Inhabitants in irregular housing; APOPi = Accumulated population in irregular housing.

Another point of interest for this basin is the existence of the wastewater treatment plant Atuba Sul (WWTP-Atuba Sul). This complex is located between the sampling sites P19 and P20. It is responsible for the treatment of most of the sewage collected in the northern area of the CMA. This facility's design is composed of Upflow Anaerobic Sludge Blanket (UASB) reactors, followed by dissolved air flotation systems. These processes are not designed to efficiently remove ammonia nitrogen.

2.2. SAMPLING

For the determination of the concentration of nutrients (ammonia nitrogen and orthophosphate), twenty samples of surface water of the Atuba River were collected in April 2019. These samples, collected in a 5 L Van Dorn bottle, were stored in twenty 500 mL polyethylene terephthalate (PET) bottles. These bottles had been previously decontaminated by a 5% v/v. HCl solution.

For the caffeine analyses, surface water samples from eight sites were collected (P1, P3, P7, P10, P13, P15, P19, and P20). The samples were stored in 1 L amber bottles until analyzed. These bottles had been decontaminated by a 5% v/v. Extran® detergent solution (Merck Milipore – Darmstadt, Germany).

After collection, all bottles were stored in thermally isolated boxes and immediately conducted to the laboratory.

2.3. CHEMICAL ANALYSES

The analyses for the determination of the concentration of both nutrients (ammonia nitrogen and orthophosphate) were performed in accordance with APHA *et al.* (2012). Orthophosphate concentrations were determined by the molybdate/ascorbic acid colorimetric method. Ammonia nitrogen concentrations were obtained through the nitroprusside/phenol spectrophotometric method.

For the determination of caffeine concentrations, the chromatographic analysis was performed in accordance with the method presented by Ide *et al.* (2013). The surface water samples were filtered through a membrane of cellulose acetate (0.45 µm). These filtered samples proceeded to flow through 6 mL octadecylsilane cartridges (Agilent - Santa Clara, USA). These cartridges had been conditioned by hexane, ethyl acetate and acidified ultra-pure water (pH close to 3) in a velocity that varied from 6 to 8 mL min^{-1}. After being vacuum dried, a solution of acetone and acetonitrile (v/v, 1:1) eluted the cartridges. The extracts were then put in flat-bottomed glass balloons and retro evaporated at 40°C. The contents left in the flat-bottomed glass balloons were dissolved, yet again, in 1 mL of acetonitrile and passed into a 2 mL vial.

The vials were sent for injection in a liquid chromatograph Agilent (Santa Clara, USA), Model 1260, with a photodiode array. This equipment featured an octadecylsilane column with a length of 250 mm, internal diameter of 4.6 mm, and a particle diameter of 5 µm. Then, 5 µL of sample were injected at a speed of 1.0 $mL.min^{-1}$ on isocratic mode with a composition of 1:1 of ultra-pure water and acetonitrile (HPLC grade purity, ≥ 99.9% – Sigma-Aldrich (St. Louis, USA)).

The monitored wavelength for caffeine was 273 nm and its retention time was 2.4 min. The calibration curves were established with concentrations ranging from 0.02 mg L-1 to 2.0 mg L-1 which resulted in a wide linear range with a regression coefficient (R^2) of 0.9971 with the equivalent equation: 1,742,329x + 67,869. The recovery test was performed with 6 different concentrations of caffeine and resulted in 49.3%±9.3%. Reproducibility and repeatability values of 0.8 and 4.0, respectively, expressed as coefficients of variation, were considered satisfactory (<15%). The repeatability was assessed with the injection of analytes in quintuplicate (three different concentrations) in the same day and reproducibility was assessed with the injection of analytes in quintuplicate (same concentrations as in repeatability) in three different days: 1st day, 7th day and 14th day. The limit of detection (LD) of 12.0 ng.L-1 and limit of quantification (LQ) of 40 ng.L-1, calculated using the standard deviation of at least three blank samples and IC is the inclination of the analytical curve (multiplied by 3 for LD and 10 to LQ).

2.4. MODEL GENERATION

The Pearson coefficient is a tool used to determine the linear correlation between two variables. As the Pearson coefficient is based solely on linear behaviors, non-linearly correlated variables might appear to be uncorrelated. Despite this limitation, the Pearson coefficient is widely used, due to its ease of implementation and its numerically tangible inputs and outputs. The Pearson coefficient is defined as (Equation 1):

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}} \tag{1}$$

The Pearson coefficient is the basis for the Canonical Correlation Analysis method (CCA). The CCA is a statistical method based upon the principles of the Principal Component Analysis. This method aims to explore the dependence between multiple dependent variables (DVs) and independent variables (IVs).

The CCA presents itself as a robust tool to evaluate the relevance of complex relationships between environmental variables (Khalil *et al.*, 2011; Di Felici *et al.*, 2012; Wei *et al.*, 2018; Mangadze *et al.*, 2019).

The method of the CCA was first presented by Hotelling (1936). This process consists of the synthetic creation of two canonical variable matrices (CV). These are composed of two vectors, called $\hat{A}$ and $\hat{B}$. The elements of these vectors, when multiplied by the sample values, would determine the optimal coefficients for the maximal Pearson correlation between the IVs and the DVs. That could be summarized as (Equation 2):

CV_{IV} has the maximum Pearson correlation to CV_{DV}:

$$CV_{IV} = A_1IV_1 + A_2IV_2 \dots A_nIV_n \quad and \quad CV_{DV} = B_1DV_1 + B_2DV_2 \dots B_iDV_i \tag{2}$$

CV_{IV} = Coefficients of the canonical variables matrix for the independent variables; CV_{DV} = Coefficients of the canonical variables matrix for the dependent variables; A_n = n^{th} coefficient of the Â vector; B_i = ith coefficient of the ^B vector; IV_n = nth independent variable; DV_i = ith independent variable.

Therefore, the matrices of coefficients $\hat{A}$ and $\hat{B}$ create a weighted sum of each variable in such a way that the Pearson correlation between the complex CV_{IV} and CV_{DV} is the best possible.

For the determination of these matrices, it is considered the covariance between every single variable, including itself, by means of a covariance matrix (**S**). According to this, the final set of canonical values depend upon the number of the variables (dependent or independent) that have the lowest number of instances (Malacarne, 2014)

After the normalization of the canonical weights to the experimental data, it is possible to approximate the equation $CV_{IV} \approx CV_{DV} + \tau$ (τ is a transposition constant). This approximation enables the creation of an equation capable of interpolating within the experimental data. For the scope of this study, this equation would be capable of determining the quantitative relationship between the socio economic data (total population and population residing in irregular housing) with the concentration of orthophosphate, ammonia nitrogen and caffeine. This quantitative relationship would be turned into an equation.

It is not within the scope of this paper to explain thoroughly the implementation of the CCA. Therefore, the authors point to the didactic works of (Hotelling, 1936; Malacarne, 2014; Rencher, 2002; Ferreira, 2018) for further information on the subject.

2.5. ERROR ANALYSES

To determine the statistical validity of the model, three different procedures were executed, the root mean square error (RMSE), the chi-squared (χ^2) test for association and the determination of the margin of error.

The RMSE is a widely used measure of error of a model. The lower its value, the closer the distribution of the modeled values are to the observed ones. It is given by the following equation (Equation 3).

$$RMSE = \sqrt{\frac{\sum(v_m - v_o)^2}{n}} \tag{3}$$

RMSE = Root Mean Square Error; v_m = value modeled; v_o = value observed; n = nº of samples

One of the ways to test the hypothesis that the correlation/independence between two variables is statistically significant is by the Chi-Squared test for Association. Its statistic is given by the following formula (Equation 4):

$$\chi^2 = \sum \frac{(v_m - v_o)^2}{v_o} \tag{4}$$

χ^2 = **Chi-squared statistic; M = Modeled value; O = observed value.**

The Chi-squared statistic would then be confronted to a curve of the probability density distribution of a Chi-distribution for the same degree of freedom, to discover its equivalent p-value. In the case of the test for association, a p-value lower than 0.05 means that the two variables entered are independent. A p-value higher than 0.05 would then mean that the two variables are associated.

The function that produces a chi-distribution for k degrees of freedom is given by (Equations 5 and 6, where Γ is the gamma distribution, also provided):

$$\begin{cases} for\ x\ \geq 0,\ f(x;k) = \dfrac{x^{k-1}e^{-\frac{x^2}{2}}}{2^{\frac{k}{2}-1}\Gamma\left(\frac{k}{2}\right)} \\ \quad for\ x < 0,\ \ f(x;k) = 0 \end{cases} \quad \rightarrow \quad \Gamma\left(\frac{k}{2}\right) = \int_0^{\infty} x^{\frac{k}{2}} e^{-x} dx \tag{5}$$

To determine the limits of the model, the margin of error was calculated:

$$M = 1.96 \frac{\sigma}{\sqrt{n}} \tag{6}$$

M = Margin of error; σ = standard deviation of the modeled values.

3. RESULTS AND DISCUSSION

3.1. Chemical Analyses

The observed concentrations for ammonia nitrogen, orthophosphate, and caffeine for the 20 sampling sites located along the Atuba River are presented on Table 2.

Table 2. Concentrations in the Atuba River for ammonia nitrogen, orthophosphate, and caffeine.

	Chemical Parameters				Chemical Parameters		
Site	**N_A (mg L^{-1})**	**PO_4^{3-} (mg L^{-1})**	**CAF (µg L^{-1})**	**Site**	**N_A (mg L^{-1})**	**PO_4^{3-} (mg L^{-1})**	**CAF (µg L^{-1})**
P1	0.111	0.381	0.639	**P11**	2.959	0.655	-
P2	0.219	0.422	-	**P12**	2.929	0.721	-
P3	0.236	0.452	0.642	**P13**	3.315	0.743	3.124
P4	1.764	0.542	-	**P14**	3.213	0.720	-
P5	2.070	0.589	-	**P15**	3.250	0.724	4.087
P6	2.257	0.590	-	**P16**	3.276	0.682	-
P7	2.162	0.607	1.007	**P17**	3.629	0.828	-
P8	2.569	0.654	-	**P18**	3.298	0.697	-
P9	2.492	0.655	-	**P19**	3.219	0.700	7.168
P10	2.327	0.660	1.186	**P20**	8.873	0.870	8.524

N_A = Ammonia nitrogen; PO_4^{3-} = Orthophosphate; CAF = Caffeine.

The concentrations for all parameters analyzed presented a rising behavior along the course of the river. The smallest concentration was observed on P1, which was closest to the river's spring, while the largest value was observed on the site P20. The site P20 was the one located the furthest from the spring, as well as being the only site which was influenced by the discharge from the WWTP Atuba Sul. The concentration rise found through the sites P19 and P20 may point to an inefficient treatment of the sewage processed on this WWTP regarding the ammonia nitrogen removal

Likewise, the spike in caffeine concentration starting at P13 may be explained by the demographic density found within these zones of hydrological influence. The largest concentration raise was 3.081 µg L^{-1}, between sites P15 and P19. This may be due to the higher density of irregular housing found in this region. As Katukiza *et al.* (2012) and Kelman (2015) proposed, these areas tend to be hotspots for hydric pollution for two reasons: i) lower governmental incentive to connect these residences to the public sewage and solid waste collection systems; ii) the irregular aspect of the domicile creates insecurity for their inhabitants, which usually are not willing to invest in individual wastewater treatment solutions. These factors point to the importance of the study of the social aspects of the residents of a river's basin to better understand the source and fate of water pollution in urban rivers.

Table 3 presents the concentrations determined for the Atuba River by other researchers that also encompassed the Atuba River in their study areas.

Table 3. Concentrations for caffeine (CAF), ammonia nitrogen (N_A) and orthophosphate (PO_4^{3-}) for other sampling campaigns performed on the Atuba River.

Sampling Date	Reference	Caffeine concentration ($\mu g\ L^{-1}$)			N_A ($mg\ L^{-1}$)	PO_4^{3-} ($mg\ L^{-1}$)
		Min	Max	Mean		
April/2014	**Mizukawa *et al.* (2019)**	0.65	5.375	2.89	9.77 (±15.6)	n.a.
June/2014		1.58	5.36	3.58	17.5 (±19.4)	n.a.
October/2014		0.53	4.39	3.06	10.6 (±6.7)	n.a.
March/2015		0.3	4.96	3.35	2.0 (±2.0)	n.a.
June/2015		1.46	3.44	2.79	18.4 (±9.9)	n.a.
February/2010	**Ide *et al.* (2017)**	8.97	6.9	7.93	n.a.	n.a.
May/2010		0.39	2.14	1.26	n.a.	n.a.
August/2010		5.14	6.73	5.93	n.a.	n.a.
November/2010		0.5	5.58	3.04	n.a.	n.a.
April/2012 Feb/2013*	**Osawa *et al.* (2015)**	n.a.	n.a.	n.a.	16.2 (±26.8)	0.48 (±0.88)
April/2011-Nov/2011*	**Kramer *et al.* (2015)**	n.a.	n.a.	n.a.	25.91 (±20.72)	1.84 (±1.71)

N_A = Ammonia Nitrogen; PO_4^{-3} = Orthophosphate; * = Four sampling campaigns were performed through this period.

Mizukawa *et al.* (2019) analyzed four sampling sites (named P1, P7, P9 and P20 in this study) and Ide *et al.* (2017) analyzed two (named P19 and P20). Though, the caffeine concentrations found by both groups were lower than the ones determined in this study. This result could be explained by the date of the sampling, which influences the amount of people residing on this basin, as well as the volume of wastewater processed by the WWTP Atuba Sul. According to the Brazilian Institute for Geography and Statistics (IBGE, 2019), the population for the area grew 6.4% from 2010 to 2014, 3.9% from 2014 to 2019 (a total of 10.4% from 2010 to 2019).

The concentrations for ammonia nitrogen determined by Mizukawa *et al.* (2019), Osawa *et al.* (2015) and Kramer *et al.* (2015) were higher than those observed in this study. This might be due to an improvement in ammonia nitrogen removal efficiency by the WWTP Atuba Sul. A reason that might explain a higher concentration of caffeine, yet a lower concentration of ammonia nitrogen, is the dilution of these contaminants by the rainy season, which could indicate that during a drier season the concentration of caffeine might be higher.

The concentration of orthophosphate observed in this study was on par with the ones observed by Osawa *et al.* (2015) and Kramer *et al.* (2015).

The concentrations of caffeine observed by other researchers in Brazil and other countries are presented in Table 4.

Table 4. Concentrations for caffeine observed by other researchers.

Country	Location	Caffeine concentration (μg L^{-1})	Reference
Brazil	Sinos River	3.73 (± 6.76)	Peteffi *et al.* (2019)
	Dourados River	0.14 (± 0.33)	Sposito *et al.* (2018)
	Brilhante River	0.02 (± 0.02)	
Taiwan	Taiwan Strait (Seawater)	0.002 (± 0.002)	Fang *et al.* (2019)
Ukraine	Dnieper River	19.2	Ho *et al.* (2020)
China	Shijing River	(Mass flow) 446.57 g d -1	Yuan *et al.* (2020)

The concentrations observed by the studies shown on Table 4 (Lopez-Doval *et al.*, 2016; Sposito *et al.*, 2018; Peteffi *et al.*, 2019; Fang *et al.*, 2019; Yuan *et al.*, 2020) corroborate the idea that caffeine contamination present on urban water bodies is not a problem confined only to the Atuba River Basin, or even to Brazil, but a worldwide problem. On a smaller scale, caffeine was also found on sea waters across the planet. Thus, the monitoring of caffeine presents itself as a viable way to understand the extent of the anthropic pressures put upon water resources.

Therefore, the existence of a tool (in this case, a model) able to determine indirectly the caffeine concentration on an urban river could improve the understanding of water contamination, across the world.

3.2. MODEL GENERATION

The model created aimed to determine the concentrations of caffeine in an indirect way. The model used as input the following social parameters: total population and population residing in irregular housing within the zones of hydrological influence of each sampling site. The chemical parameters used as inputs were the concentrations of ammonia nitrogen and orthophosphate, found on the same sites.

The first step used to generate the model for the caffeine concentration to determine which of the variables would be independent or dependent. It was decided that the chemical parameters were a portrait of the behavior of the population which resides in the basin, and not the opposite. Therefore, the total population and the population residing in irregular housing were set as the independent variables. Consequently, the chemical parameters were set as the dependent variables.

The second step was to determine the variance and covariance of all variables. The results for this procedure are presented on Table 5.

Table 5. Variance-Covariance matrix for the social and chemical parameters of the Atuba River.

S	IVs		DVs		
	P_T	P_i	CAF	N_A	PO_4
P_T	23188528890	780977474	339923	20833	15643
P_i	780977474	30550886	12846	7724	490.8
CAF	339923	12846	5.629	3.291	0.210
N_A	20833	7724	3.291	3.094	0.184
PO_4	15643	490.8	0.210	0.184	0.015

IVs = Independent Variables; DVs = Dependent Variables; P_T = Total Population; P_i = Population in irregular housing; CAF = Caffeine; N_A = Ammonia nitrogen; PO_4 = Orthophosphate.

Through the **S** matrix, it was possible to determine the vectors with the canonical values for the independent (A) and dependent (B) variables. There are two instances of canonical values for each set of variables, these values are represented by the columns of their respective vectors (Equations 7 and 8).

$$A = \begin{bmatrix} -0{,}00000318 & 0.00001732 \\ -0.00009672 & -0.00047545 \end{bmatrix} \tag{7}$$

$$B = \begin{bmatrix} -0{,}379812 & -0.20066088 \\ 0.05470996 & -0.79124554 \\ -1.88825624 & 16.42867095 \end{bmatrix} \tag{8}$$

The next step was to establish which of the sets of canonical values are the most appropriate; for this purpose, the $R_{U,V}^{-1}$ matrix was used. The $R_{U,V}^{-1}$ matrix determines the value of the linear correlation among the coefficients in both matrix A and B (Equation 9).

$$R_{U,V}^{-1} = \begin{bmatrix} 1.01365326 & 0 \\ 0 & 1.26481073 \end{bmatrix} \quad \begin{cases} r_1 = 0.9865 \\ r_2 = 0.7063 \end{cases} \tag{9}$$

The first column of the matrices A and B presented a correlation of r = 0.9865, which is significantly higher than the second column (r = 0.7063). Therefore, the first set of canonical values was the chosen one. These values, along with the observed variables, were then formed into an equation (Equation 10):

$$-0.00000318P_T - 0.00009672P_i = -0.379812CAF + 0.05470996N_A - 1.88825624PO_4 + 0.2033 \tag{10}$$

To make the equation more easily understandable, the values were transformed to into integers (Equations 11 and 12):

$$\frac{2P_T}{61} + P_i = 3981CAF - \frac{1720N_A}{3} + 19786PO4 - 8409 \tag{11}$$

Solving for CAF:

$$CAF = \frac{P_T}{112270} + \frac{P_i}{3981} - 0.144\,N_A + 4.970\,PO_4 - 2.112 \tag{12}$$

Where: CAF = caffeine concentration (μg L^{-1}); P_T = Total Population; P_i = Population in irregular housing; N_A = concentration of ammonia nitrogen (mg L^{-1}); PO_4 = concentration of orthophosphate (mg L^{-1}).

The modeled values for the caffeine concentrations for the sites P1, P3, P7, P10, P13, P15, P19 and P20 (the ones in which caffeine was analyzed) were determined. The correlation coefficient between the ones modeled was r = 0.984. It presented a total RMSE of 1.18 μg L^{-1}. The culprit for this error statistic is the consistent overestimate of the model (excepting P13), when compared to the observations. The Chi-squared statistic for these two datasets (degrees of freedom = 7) was $\chi^2 = 0.9661$, which provided a p-value of 0.955. This led to the exclusion of the null hypothesis (that these two datasets were independent).

The correlation graph for the observed and modeled values is presented on Figure 2.

The values for the concentrations of caffeine for the twenty sampling sites were then calculated. These are presented in Table 6.

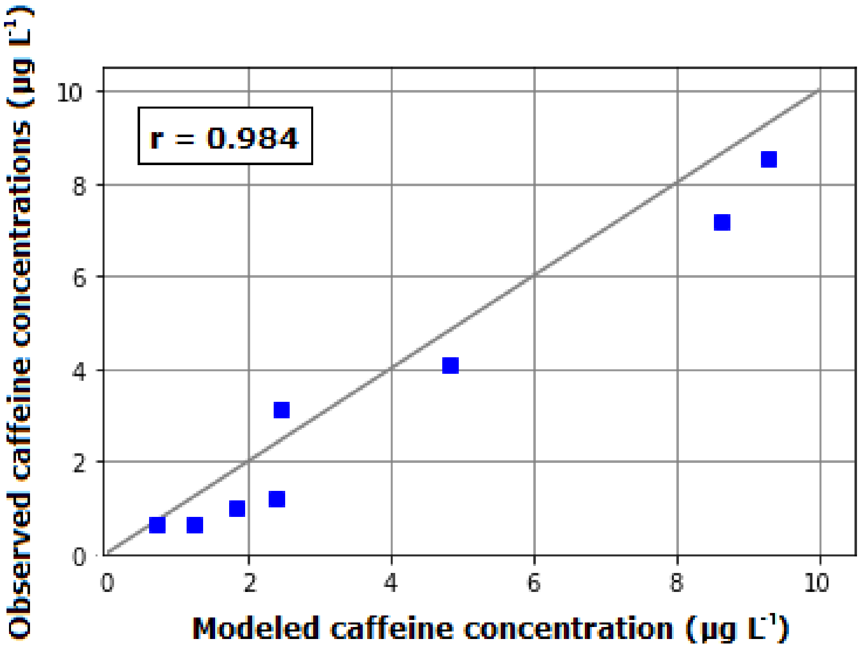


Figure 2. Correlation of the concentration of caffeine observed on the sampling sites P1, P3, P7, P10, P13, P15, P19 and P20 and the values predicted for the same sites using the CCA-generated model.

Table 6. Concentrations of caffeine (µg L^{-1}) found for the 20 sampling sites through the model generated by CCA compared with the values observed at 8 sampling sites across the Atuba River.

Site	Modeled caffeine concentration (mg L^{-1})	Observed Caffeine concentration (µg L^{-1})	Site	Modeled caffeine concentration (mg L^{-1})	Observed Caffeine concentration (µg L^{-1})
P1	0.728	0.639	**P11**	2.539	-
P2	1.069	-	**P12**	2.355	-
P3	1.244	0.642	**P13**	2.472	3.124
P4	1.407	-	**P14**	4.550	-
P5	1.754	-	**P15**	4.815	4.087
P6	1.834	-	**P16**	5.248	-
P7	1.832	1.007	**P17**	5.672	-
P8	2.000	-	**P18**	7.340	-
P9	2.225	-	**P19**	8.640	7.168
P10	2.401	1.186	**P20**	9.283	8.524

The distribution of the modeled values presented a standard deviation of 3.853, which led to a margin of error of 1.6887 µg L^{-1} (probability = 95%, z = 1.96). Figure 3, below, presents the comparison of the model (and its probable spectrum) to the studies that analyzed caffeine concentrations performed on the Atuba River.

The modeled spectrum covers all the values observed for this sampling campaign. This was expected, as these were the concentrations used to generate the model. The values for the other campaigns performed on the Atuba River (by Ide *et al.* (2017) and Mizukawa *et al.* (2019) are not, for the most part, encompassed by the model. This could be explained by the differences that occurred in the study area during this nine-year span. While the concentrations found on sites P19 and P20 were consistently lower than those found on this study, the ones found on site P7 were consistently higher. This could be explained by the different behavior of the population, as the latter two sampling sites are in a lower income, irregular housing area.

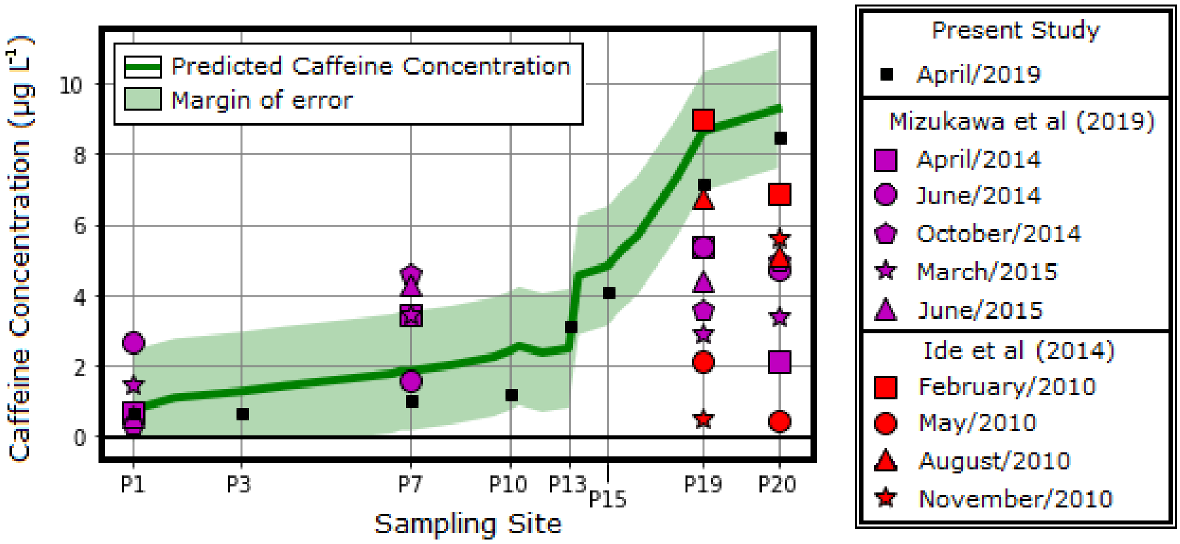


Figure 3. Distribution of three studies performed on the Atuba River compared with the modeled caffeine concentrations.

Another topic that has risen concern recently is the presence of viruses in water resources. This model might be a tool to better understand the behavior of these microorganisms in the aquatic environment. Gourmelon *et al.* (2010) and Kumar *et al.* (2019) have observed concomitantly the concentrations of caffeine and viral genomes in wastewater, rivers and lakes. Their results might indicate somewhat of a correlation between the concentrations of these two parameters. As their results pointed out, higher concentrations of caffeine were usually followed by a higher concentration of viral genome and a higher presence of the bacteria *Escherichia coli*. Their studies analyzed the concentrations of genomes from Norovirus, Hepatitis A, Aichivirus, Pepper Mild Mottle Virus and F-specific RNA bacteriophages. The results observed by Sidhu *et al.* (2013) have also pointed to some correlation between higher concentrations of caffeine and the presence of viral genome in water systems.

Due to the pandemic the world was under in 2021, the analysis of viral presence (especially SARS-CoV-2) in water resources have raised even more concerns. Rimoldi *et al.* (2020) have observed some possible correlation between the presence of this specific virus on waters with high concentrations of caffeine (above 0.44 $\mu g\ L^{-1}$). Interesting future studies could be done to better understand this correlation.

The model was able to interpolate a possible range of concentrations of caffeine that could be observed along the river. It took as input only socio economic data and the concentrations of nutrients, whose determination is cheaper and less labor intensive. The information provided by this model, when applied to other basins, might be useful to managers, researchers, and decision-makers to better understand the pressures a given urban river might be under.

4. CONCLUSIONS

The concentrations for ammonia nitrogen and orthophosphate for 20 sampling sites along the Atuba River were quantified. Also, the population of influence that inhabited each of the zones of hydrological influence was defined (from which the total population and population residing in irregular housing for each sampling site were established). For 8 of these sampling sites, the concentrations of caffeine were determined. Using these observed datasets, a model was created through the manipulation of the CCA method. The study generated a spectrum of probable ranges of caffeine concentrations for the course of the river, based on 4 inputs: the total population; the population inhabiting irregular housing; then population residing in the zone of hydrological influence; and the concentrations of ammonia nitrogen and orthophosphate at the sampling site. Even though the water resource system of the Atuba River is complex, being home for over half a million people and receiving the discharge from the WWTP Atuba Sul, the model had a satisfactory performance. More studies should be performed to understand the effects of generalization for water resource systems that are under different pressures, as well as different climatic, cultural, and infrastructural situations.

Due to the characteristics of caffeine as a chemical tracer for anthropic pressures on water systems, a model that could predict indirectly its concentration along a river might provide more information at a lower cost. The method for generating the model's equation could be applied to different urban rivers, becoming a useful tool for water-resource managers, researchers, and decision-makers to better understand the environmental impact the population inhabiting a given basin will have on its water resources. This model could be particularly beneficial for regions that do not have systematic sampling campaigns, by being able to determine zones which would demand higher attention from managers and decision-makers, with a small number of sporadic non-uniformly spread data samples.

5. ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) and the Brazilian National Council for Scientific and Technological Development (CNPq) by Bolsa Productividade (proc. 302736/2016-6) and Call MCTIC/CNPq N° 28/2018 (proc. 407157/2018-2). The authors thank the Graduate Program for Water Resources and Environmental Engineering of the Universidade Federal do Paraná (PPGERHA-UFPR) and the Multidisciplinary Nucleus for Research on Environmental Technologies of the Universidade Tecnológica Federal do Paraná (NIPTA-UTFPR).

6. REFERENCES

APHA; AWWA; WEF. **Standard Methods for the examination of water and wastewater**. 22nd ed. Washington, 2012. 1496 p.

COLIM, A. N.; NASCIMENTO, P. C.; WIETHAN, B. A.; ADOLFO, F. R.; DRESCH, L. R.; CARVALHO, L. M. *et al.* Reversed-Phase High-Performance Liquid Chromatography for the Determination of 15 Rare Earth Elements in Surface Water Sample Collected in a Mining Area from Lavras do Sul/RS, Brazil. **Cromatographia**, v. 82, n. 5, 2019. https://link.springer.com/article/10.1007/s10337-019-03709-w

DAFOUZ, R.; CÁCERES, N.; RODRÍGUEZ-GIL, J. L.; MASTROIANNI, N.; DE ALDA, M. L.; BARCELÓ, D. *et al.* Does the presence of caffeine in the marine environment represent an environmental risk? A regional and global study. **Science of the Total Environment**, v. 615, 632–642, 2018. https://doi.org/10.1016/j.scitotenv.2017.09.155

DI FELICI, V.; MARCINELLI, R.; PROULX, R.; CAMPIGLIA, E. A multivariate analysis for evaluating the environmental and economical aspects of agroecosystem sustainability in central Italy. **Journal of Environmental Management**. v. 98, 119-126, 2012. https://doi.org/10.1016/j.jenvman.2011.12.015

FANG, T.-H.; LIN, C.-W.; KAO, C.-H. Occurrence and distribution of pharmaceutical compounds in the Danshuei River Estuary and the Northern Taiwan Strait. **Marine Pollution Bulletin**, v. 146, p. 509-520, 2019. https://doi.org/10.1016/j.marpolbul.2019.06.069

FERREIRA, D. F. **Estatística Multivariada**. 3. ed. Lavras: Editora UFLA, 2018. p. 624.

GERSHUNOV, A.; BENMARHINA, T.; AGUILERA, R. Human health implications of extreme precipitation events and water quality in California, USA: a canonical correlation analysis. **Lancet Planetary Health**, v. 2, n. 1, p. S9, 2018. https://doi.org/10.1016/S2542-5196(18)30094-9

GOURMELON, M.; CAPRAIS, M. P.; MIESZKIN, S.; MARTI, S.; WÉRY, N.; JARDÉ, E. *et al.* Development of microbial and chemical MST tools to identify the origin of the faecal pollution in bathing and shellfish harvesting waters in France. **Water Research**, v. 44, 4812-4824, 2010. https://doi.org/10.1016/j.watres.2010.07.061

HAN, L.; ZHOU, W.; LI, Y.; QIAN, Y. Urbanization strategy and environmental changes: An insight with relationship between population change and fine particle pollution. **Science of the Total Environment**, v. 642, 789-799, 2018. https://doi.org/10.1016/j.scitotenv.2018.06.094

HO, K. T. *et al*. Contaminants, mutagenicity and toxicity in the surface waters of Kyiv, Ukraine. **Marine pollution bulletin**, v. 155, p. 111153, 2020. https://doi.org/10.1016/j.marpolbul.2020.111153

HOTELLING, H. Relations between two sets of variables. **Biometrika**, v. 28, n. 3/4, 321-327, 1936. https://doi.org/10.1007/978-1-4612-4380-9_14

IDE, A. H.; OSAWA, R. A.; MARCANTE, L. O.; PEREIRA, J. L. G. F. S. C.; AZEVEDO, J. C. R. Occurrence of pharmaceutical products, female sex hormones and caffeine in a subtropical region in Brazil. **CLEAN-Soil Air Water**, v. 45, n. 9, 2017.

IDE, A. H.; OSAWA, R.; MARCANTE, L. O.; PEREIRA, J. L. G. F. S. C.; AZEVEDO, J. C. R. Utilização da cafeína como indicador de contaminação por esgotos domésticos na bacia do Alto Iguaçu. **Brazilian Journal of Water Resources**, v. 18, n. 2, p. 201-211, 2013. https://doi.org/10.1002/clen.201700334

IBGE. **População nos censos demográficos, segundo os municípios das capitais 1872/2010**. 2019. Available on: https://censo2010.ibge.gov.br/sinopse/index.php?dados=6.

KATUKIZA, A. Y.; RONTELTAP, M.; NIWAGABA, C. B.; FOPPEN, J. W. A.; KANSIIME, F.; LENS, P. N. L. Sustainable sanitation technology options for urban slums. **Biotechnology Advances**, v. 30, n. 5, 964–978, 2012. https://doi.org/10.1016/j.biotechadv.2012.02.007

KELMAN, J. Water Supply to the Two Largest Brazilian Metropolitan Regions. **Aquatic Procedia,** v. 5, 13–21, 2015. **https://doi.org/10.1016/j.aqpro.2015.10.004**

KHALIL, B.; OUARDA, T. B. M. J.; SAINT-HILAIRE, A. Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis. **Journal of Hydrology**, v. 405, 277-287, 2011. https://doi.org/10.1016/j.jhydrol.2011.05.024

KRAMER, R. D.; MIZUKAWA, A.; IDE, A. H.; MARCANTE, L. O.; SANTOS, M. M.; AZEVEDO, J. C. R. Determinação de anti-inflamatórios na água e sedimento e suas relações com a qualidade da água na bacia do Alto Iguaçu, Curitiba-PR. **Brazilian Journal of Water Resources,** v. 20, n. 3, 657–667, 2015.

KUMAR, M.; RAM, B.; HONDA, R.; POOPIPATTANA, C.; CANH, P. D.; CHAMINDA, T. *et al.* Concurrence of antibiotic resistant bacteria (ARB), viruses, pharmaceuticals and personal care products (PPCPs) in ambient waters of Guwahati, India: Urban vulnerability and resilience perspective. **Science of the Total Environment**, v. 693, 133-147, 2019. https://doi.org/10.1016/j.scitotenv.2019.133640

LOPEZ-DOVAL, J. C.; MONTAGNER, C. C.; ALBUQUERQUE, A. F.; MOSCHINI-CARLOS, V.; UMBUZEIRO, G.; POMPEO, M. Nutrients, emerging pollutants and pesticides in a tropical urban reservoir: Spatial distributions and risk assessment. **Science of the Total Environmen***t*, v. 575, 1307-1324, 2016. https://doi.org/10.1016/j.scitotenv.2016.09.210

MALACARNE, R. L. Canonical Correlation Analysis. **The Mathematica Journal**, v. 16, 6-16, 2014.

MANGADZE, T.; TAYLOR, J. C.; FRONEMAN, W. P.; DALU, T. Water quality assessment in a small austral temperate river system (Bloukrans River system, South Africa): Application of multivariate analysis and diatom indices. **South African Journal of Botany**, v. 125, p. 353-359, 2019. https://doi.org/10.1016/j.sajb.2019.08.008

MIZUKAWA, A.; FILIPPE, T. C.; PEIXOTO, L. O. M.; SCIPIONI, B.; LEONARDI, I. R.; AZEVEDO, J. C. R. Caffeine as a chemical tracer for contamination of urban rivers. **Brazilian Journal of Water Resources**, v. 24, 2019. https://doi.org/10.1590/2318-0331.241920180184

OSAWA, R. A.; IDE, A. H.; SAMPAIO, N. M. F. M.; AZEVEDO, J. C. R. Determinação de fármacos anti-hipertensivos em águas superficiais na região metropolitana de Curitiba. **Brazilian Journal of Water Resources**, v. 20, n. 4, p. 1039-1050, 2015.

PETEFFI, G. P.; FLECK, J. D.; KAEL, I. M.; ROSA, D. C.; ANTUNES, M. V.; LINDEN, R. Ecotoxicological risk assessment due to the presence of bisphenol A and caffeine in surface waters in the Sinos River Basin - Rio Grande do Sul – Brazil. **Brazilian Journal of Biology**, v. 79, p. 712-721, 2019. https://doi.org/10.1590/1519-6984.189752

RENCHER, A. C. **Methods of Multivariate Analysis**. 2. ed. Danvers: John Wiley & Sons, 2002. p. 727.

RIMOLDI, S. G.; STEFANI, F.; GIGANTIELLO, A.; POLESELLO, S.; COMANDATORE, F.; MILETO, D. *et al.* Presence and infectivity of SARS-CoV-2 virus in wastewaters and rivers. **Science of the Total Environment**, v. 744, p. 140-148, 2020. https://doi.org/10.1016/j.scitotenv.2020.140911

SIDHU, J. P. S.; AHMED, W.; GERNJAK, W.; ARYAL, R.; MCCARTHY, D.; PALMER, A. *et al.* Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. **Science of the Total Environment**, v. 463-464, p. 488-496, 2013. https://doi.org/10.1016/j.scitotenv.2013.06.020

SPOSITO, J. C. V.; MONTAGNER, C. C.; CASADO, M.; NAVARRO-MARTIN, L.; SOLORZANO, J. C. J.; PINA, B. *et al.* Emerging contaminants in Brazilian rivers: Occurrence and effects on gene expression in zebrafish (*Danio rerio*) embryos. **Chemosphere**, v. 209, p. 696-704, 2018. https://doi.org/10.1016/j.chemosphere.2018.06.046

TIYASHA, T. M.; TUNG, Z. M.; YASEEN. A survey on river water quality modeling using artificial intelligence models: 2000–2020. **Journal of Hydrology**, v. 585, p. 124-186, 2020. https://doi.org/10.1016/j.jhydrol.2020.124670

WEI, H.; YU, H.; ZHANG, G.; PAN, H.; LV, C.; MENG, F. Revealing the correlations between heavy metals and water quality, with insight into the potential factors and variations through canonical correlation analysis in an upstream tributary. **Ecological Indicators,** v. 90, p. 485-493, 2018. https://doi.org/10.1016/j.ecolind.2018.03.037

WOODHOUSE, P.; MULLER, M. Water governance – an historical perspective on current debates. **World development**, v. 92, p. 225-241, 2017. https://doi.org/10.1016/j.worlddev.2016.11.014

YUAN, X.; HU, J.; LI, S.; YU, M. Occurrence, fate, and mass balance of selected pharmaceutical and personal care products (PPCPs) in an urbanized river. **Environmental Pollution**, v. 266, p. 115-124, 2020. https://doi.org/10.1016/j.envpol.2020.115340

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Accumulated litter, nutrient stock and decomposition in an Atlantic Forest fragment

ARTICLES doi:10.4136/ambi-agua.2787

Received: 31 Aug. 2021; Accepted: 07 Feb. 2022

Diêgo Gomes Júnior[1]; Marcos Vinicius Winckler Caldeira[2]; Dione Richer Momolli[3*]; William Macedo Delarmelina[4]; Henrique Machado Dias[2]; Ranieri Ribeiro Paula[2]; Kelly Nery Bighi[2]

[1]Gerência Local de Alegre. Instituto de Defesa Agropecuária e Florestal do Espírito Santo (IDAF-ES), Rua Olívio Correia Pedrosa, n° 566, CEP: 29500-000, Alegre, ES, Brazil. E-mail: diegogomesj@gmail.com
[2]Departamento de Ciências Florestais e da Madeira. Universidade Federal do Espírito Santo (UFES), Avenida Governador Lindemberg, n° 316, CEP: 29550-000, Jerônimo Monteiro, ES, Brazil.
E-mail: mvwcaldeira@gmail.com, henridias@yahoo.com.br, ranierirpaula@gmail.com, kellyneryb@yahoo.com.br
[3]Departamento de Engenharia Florestal. Universidade Federal de Santa Maria (UFSM), Avenida Roraima, n° 1000, CEP: 97105-900, Santa Maria, RS, Brazil.
[4]Técnico em Florestas. Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo (IFES), Campus Ibatiba, Avenida 7 de Novembro, n° 40, CEP: 29395-000, Ibatiba, ES, Brazil.
E-mail: william.delarmelina@ifes.edu.br
[*]Corresponding author. E-mail: dionemomolli@gmail.com

ABSTRACT

Litter dynamics is one of the fundamental processes for the growth and maintenance of native forest fragments, being considered the main pathway for nutrient cycling in forests. Therefore, studies on accumulated litter and nutrient content provide information for a better understanding of nutrient dynamics. The aim of the study was to evaluate leaf litter and nutrient stock in different seasons and the instantaneous rate of decomposition in an Atlantic Forest Fragment over two years. Litter sampling was carried out in 12 permanent plots with dimensions of 20 m x 50 m. Litter dry mass and nutrient concentration were determined. The average annual accumulation of litter was 5269 kg ha^{-1}. There was no statistically significant difference in the amount of litter between seasons. There was a statistically significant difference in the contents of N, P, K Ca and Mg between the different seasons. Nutrient stocks were 123.3, 92.8, 13.2, 11.8, 9.6, 3.0 kg ha^{-1} year^{-1} for Ca, N, Mg, S, K and P respectively. The total of nutrients in the accumulated litter was 253.7 kg ha^{-1} year^{-1}. The litter renewal time was 281 days. The times required for 50% and 95% litter decomposition were 196 and 850 days. The average litter stocks, nutrients and decomposition are in line with other studies, indicating that the Atlantic Forest fragment seasonal semi-deciduous presents indicators of nutrient cycling. These results show that the ecosystem is sustainable from the point of view of cycling and nutrient release.

Keywords: ecosystem functions, nutrient cycling, secondary succession.

Serapilheira acumulada, estoque de nutrientes e decomposição em um fragmento de Floresta Atlântica

RESUMO

A dinâmica da serapilheira é um dos processos fundamentais para o crescimento e manutenção dos fragmentos florestais nativos, sendo considerada a principal via de ciclagem de nutrientes nas florestas. Portanto, estudos sobre a serapilheira acumulada e conteúdo de nutrientes subsidiam informações para uma melhor compreensão da dinâmica dos nutrientes. O objetivo do estudo foi avaliar o estoque de serapilheira e nutrientes em diferentes estações do ano e a taxa instantânea de decomposição em um fragmento de Floresta Atlântica ao longo de dois anos. A amostragem da serapilheira foi realizada em 12 parcelas permanentes com dimensões de 20 m x 50 m. Determinou-se a massa seca da serapilheira e o teor dos nutrientes. O acúmulo médio anual de serapilheira foi de 5.269 kg ha^{-1}. Não houve diferença estatística na quantidade de serapilheira entre as estações do ano. Houve diferença estatística nos teores de N, P, K Ca e Mg entre as diferentes estações do ano. Os estoques de nutrientes foram de 123.3, 92.8, 13.2, 11.8, 9.6 e 3.0 kg ha^{-1} para Ca, N, Mg, S, K e P, respectivamente. O total de nutrientes na serapilheira acumulada foi de 253,7 kg ha^{-1} ano^{-1}. O tempo de renovação da serapilheira foi de 281 dias. O tempo para a decomposição de 50% e 95% da mesma foi de 196 e 850 dias, respectivamente. A produção média de serapilheira, o estoque de nutrientes e a decomposição estão de acordo com outros estudos realizados, indicando que o fragmento de Floresta Atlântica sazonal semidecidual apresenta indicadores de ciclagem de nutrientes. Esses resultados encontrados mostram que o ecossistema se encontra sustentável sob ponto de vista da ciclagem e liberação dos nutrientes.

Palavras-chave: ciclagem dos nutrientes, indicadores ecológicos, sucessão secundária.

1. INTRODUCTION

The Atlantic Forest biome is a biodiversity hotspot (Myers *et al*., 2000); however, it has been significantly impacted in recent decades by anthropogenic activities, with only 12.4% of the 131 million hectares of native forest remaining (SOS Mata Atlântica, 2020). The continued exploitation of natural resources is responsible for forest fragmentation and consequent loss of biodiversity (Fahrig, 2003). Both the fragmentation of forest remnants and the decrease in biodiversity result in a reduction of carbon stock and nutrient cycling capacity (Bello *et al*., 2015; Wolf *et al*., 2013).

The state of Espírito Santo is located in the Atlantic Forest biome; however, in 2019, only 10.5% of the state area had this natural forest formation (SOS Mata Atlântica, 2020). The forest cover in the south of the state is formed by small fragments of seasonal semideciduous forest, due to a series of anthropogenic disturbances (Godinho *et al*., 2014).

An understanding of the degree of disturbance of forest ecosystems, such as degree of fragmentation and capacity for biomass production, can be obtained through the assessment of nutrient cycling (Haag, 1985; Balieiro *et al*., 2004) and accumulated litter provides qualitative and quantitative support, being one of the main parameters when evaluating nutrient cycling (Scoriza *et al*., 2014). Nutrient cycling returns nutrients absorbed by plants to the soil in the form of plant tissues that decompose to make them available again (Odum, 1988). This process is of great importance for natural systems, especially in tropical areas, where the soils are highly weathered (Laliberte *et al*., 2013). Investigating nutrient cycling is an important strategy in forest regeneration programs (Caldeira *et al*, 2019).

The organic material constituted by leaves, branches, bark and plant reproductive material which is deposited in the soil is called “litter” (Kramer and Kozlowski, 1960; Fassebender, 1993). Assessing the seasonality of litter deposition and nutrient content for periods of less than two years may not provide an understanding of nutrient dynamics throughout the biogeochemical cycle. This information helps to choose the most suitable species for the formation and enrichment of the fragments (Caldeira *et al*., 2008).

The accumulation of litter on the soil and the regulation of decomposition is mainly determined by the successional stage of the ecosystem, the degree of disturbance and meteorological variables such as rainfall and air temperature (Scoriza *et al*., 2014; Machado *et al*., 2015). Other studies also indicate that the topographic gradient, the seasons of the year and the litter nutritional composition are important for the regulation of the decomposition rates (Santos *et al*., 2019; Machado *et al*., 2015). Moreover, nutritional composition strongly influences the biological activity and thus decomposition and release of nutrients to the soil (Simpson *et al.*, 2007; Krishna and Mohan, 2017).

Based on these reasons, quantifying litter stock and its nutrients is essential for forest restoration strategies. Considering that meteorological variables and nutritional quality influence litter decomposition, this study evaluated litter and nutrient stock in different seasons and the decomposition rate in an Atlantic Forest Fragment over two years.

2. MATERIAL AND METHODS

2.1. Location and characterization of the study area

The study area is located in a Private Natural Heritage Reserve (RPPN), Fazenda Boa Esperança (Brasil, 1998). The RPPN is located in the municipality of Cachoeiro de Itapemirim, southern Espírito Santo, at coordinates UTM/SIRGAS2000 268275.48 E and 7707754.70 S. The RPPN has a total area of 517 ha, comprising four forest fragments. The present study, located in the Itapemirim River Watershed, was carried out in a forest fragment with an area of 358.86 ha, belonging to the Burarama – Pacotuba – Cafundó ecological corridor. The average altitude of the 12 plots was 110 m, ranging from 91 m to 160 m. The average slope was 6.7%, ranging from 1% to 25%.

The vegetation in the present study area is classified as Submontane Seasonal Semideciduous Forest (IBGE, 2012). In a study on the structure of the arboreal component of the area, Archanjo *et al*. (2012) observed high richness of late secondary species and low density of early succession groups (Table 1), indicating that it is a well-preserved forest fragment with an advanced stage of succession. The characterization of the vegetation in the studied plots (Archanjo *et al*., 2012) and the topography (Delarmelina, 2015) are described in Table 1.

The climate of the region, according to the classification of Köppen, is of the Aw type (tropical with a dry season in winter) (Alvares *et al*., 2014), with an average temperature of the minimum of the coldest month of 11.8ºC, and the average of the highs of the hottest month of 34ºC (Pezzopane *et al*., 2012). According to rainfall characterization maps of Espírito Santo, the annual rainfall in the study area is between 1200 and 1300 mm (Incaper, 2017).

Precipitation data, for the study period and for the historical series (1987 - 2016), were obtained from the National Water Agency (ANA) station (02041002), located in the municipality of Castelo - ES, approximately 12 km away from the study area. The temperature data are from the automatic surface meteorological station of the National Institute of Meteorology (INMET) (Alegre-A617), located in the municipality of Alegre – ES, approximately 26 km from the studied area. The history of monthly averages for this variable was obtained with data from the same station, in the period 2006 to 2016, Figure 1.

Table 1. Characterization of the RPPN vegetation in Cachoeiro de Itapemirim, ES. Brazil.

Parameters	Values	Parameters	Values
Density (ha^{-1})*	1823	Pioneers (%) **	0.2
Number of species	258	Early Secondary Stage (%)	26.5
Number of families	54	Late Secondary Stage (%)	58.1
Diversity index (H')	4.13	Uncharacterized (%)	15.2
Equability (J)	0.74		

Dominant Families	Most common species
Fabaceae (44 sp)	*Astronium concinnum*
Myrtaceae (27 sp)	*Pseudopiptadenia contorta*
Euphorbiaceae (14 sp)	*Neoraputia alba*
Sapotaceae (13 sp)	*Astronium graveolens*
Rubiaceae (12 sp)	*Gallesia integrifólia*

*DBH ≥ 5 cm; ** Proportion of individuals by succession category.
Source: Delarmelina (2015), adapted by the author.

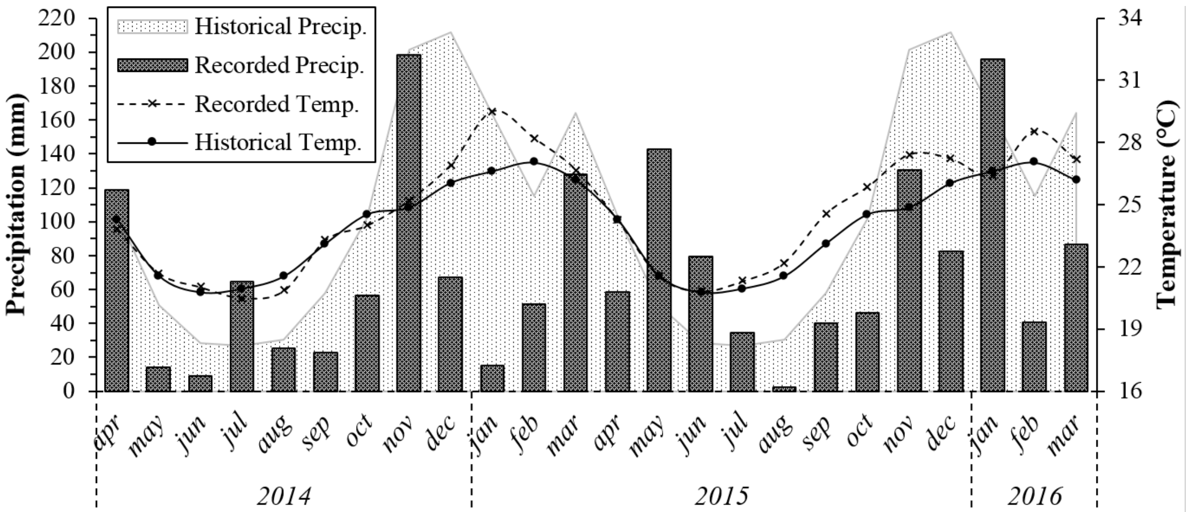


Figure 1. Climatic diagram referring to the study period and historical series of the RPPN in Cachoeiro de Itapemirim, ES, Brazil.

To determine the attributes of the soil the samples were chipped, dried in the shade and sieved (2 mm of mesh), obtaining the fine air-dried soil (TFSA). Later, they were used to determine physical and chemical attributes, according to Embrapa (2011). By means of TFSA, the granulometric composition, sand, silt and clay was obtained. The pH was determined in water (ratio 1:2.5); P, extracted with Mehlich-1 Extractor and read in a molecular absorption spectrophotometer; K and Na, extracted with Mehlich-1 Extractor and read in flame spectrophotometry; Ca^{2+} and Mg^{2+}, extracted with KCl 1 mol L^{-1} and determined by an atomic absorption spectrophotometer; Al^{3+} , extracted with KCl 1 mol L^{-1} and determined by titration; the H + Al (potential acidity), extracted with a 0.5 mol L^{-1} Ca acetate solution buffered at pH 7.0 and determined by titration.

Based on the values obtained from the chemical analysis, the sum of bases (SB) was determined, as well as the effective cation exchange capacity (t), the potential cation exchange capacity (T), the base saturation (V) and the saturation by aluminum (m), according to Embrapa (2011). As shown in Table 2, the highest clay contents were found in the deepest horizons and in the Ferrassol profile. The sand fraction predominated over clay in Planosol and Cambisol profiles. Soil pH ranged from 3.3 to 6.07 and higher acidity was found in Ferrasol type.

Table 2. Chemical and physical attributes of soil.

Horiz.	Depth	Sand	Silt	Clay	pH	P	K	Ca^{2+}	Mg^{2+}	Al^{3+}	SB	t	T	V	m
	cm	%			H_2O	mg dm^{-3}		cmol$_c$ dm^{-3}						%	
							Planosol								
A1	0-12	83.3	6.3	10.4	6.0	13.8	70.0	5.5	1.6	0.0	7.2	7.2	9.5	75.8	0.0
A2	12-41	88.0	5.1	6.9	5.8	8.5	27.0	1.2	0.7	0.0	2.0	2.0	3.3	60.4	0.0
E	41-58	87.9	6.1	6.0	6.1	14.5	14.0	0.8	0.4	0.0	1.2	1.2	2.0	60.6	0.0
Btg	58-85	70.7	6.5	22.8	5.1	0.5	3.0	3.4	3.6	1.1	7.0	8.1	10.3	68.1	13.5
Cg	85+	45.6	4.8	49.6	5.2	1.9	5.0	3.8	5.6	4.4	9.6	14.0	24.8	38.7	31.5
							Cambisol								
A1	0-7	53.8	17.1	29.1	5.9	7.3	162.0	6.6	2.8	0.0	9.8	9.8	11.8	83.1	0.0
A2	7-17	53.4	19.2	27.4	6.0	1.9	50.0	4.4	1.8	0.0	6.4	6.4	8.9	71.9	0.0
Bi	17-32	50.3	18.1	31.6	5.7	0.7	37.0	3.8	1.8	0.0	5.7	5.7	8.0	71.2	0.0
C	32-47	42.6	19.1	38.3	5.6	0.4	37.0	3.1	2.0	0.0	5.1	5.1	6.6	77.4	0.0
Cr	47+				-	-	-	-	-	-	-	-	-	-	-
							Ferrasol								
A1	0-6	32.7	7.4	59.9	4.4	3.1	53.0	2.0	0.6	0.8	2.7	3.5	9.3	29.0	22.9
AB	6-21	30.9	5.3	63.8	4.0	1.5	27.0	0.4	0.3	1.8	0.7	2.5	6.6	11.1	70.9
Bw1	21-47	24.4	5.1	70.5	3.3	0.5	11.0	0.2	0.1	1.7	0.3	2.0	4.8	6.6	84.2
Bw2	47-98	20.3	3.0	76.7	4.3	0.1	2.0	0.1	0.4	1.4	0.5	1.9	5.6	9.1	73.3
BC	98+	47.7	4.4	47.9	5.3	0.1	0.0	0.0	0.6	0.8	0.7	1.5	3.3	20.5	54.4

Horiz: Horizon; P: Phosphorus; K: Potassium; Ca^{2+}: Calcium; Mg^{2+}: Magnesium; Al^{3+}: Aluminum; SB: Sum of exchangeable bases; t: Effective cation exchange capability; T: Cation exchange capacity at pH 7.0; V: base saturation; m: Aluminum saturation.

The contents of K, Ca^{2+} and Mg^{2+}, as well as SB and t were higher in the superficial horizons of the Cambisol and Ferrasol profiles. In general, the soils in the region of the Planosol and Cambisol profiles were more fertile than the soil in the region of the Ferrasol profile, as verified by the higher values of P, SB and V.

2.2. Collection and processing of litter

The sampling of litter was carried out monthly from April 2014 to March 2016. For this, 12 permanent plots with dimensions of 20 m x 50 m were demarcated (Figure 1), totaling 12000 m². The plots were systematically distributed in the area, 350 meters away from each other. Monthly, 8 litter samples were collected per plot, totaling 96 monthly samples. The collection area of each sample was determined by a square with 0.25 m on the side and an area of 0.0625 m².

The litter traps were systematically installed in the plots, with a collector close to each vertex and one in the center of the plot, totaling five collectors per plot, 60 collectors in the entire studied area. The collectors were made in a square format, with PVC material and 2 mm nylon mesh, having 0.75 m on each side (area of 0.5625 m²) and 1 m above the ground.

After collection, the samples were stored in paper bags and dried in an air circulation oven at 65°C. When they reached constant weight, the material was weighed on an analytical balance (0.001 g) to obtain the dry mass. With the dry mass data, the total litter per unit area (kg ha^{-1}) was calculated. The samples were shredded in a Willey-type mill with 1 mm mesh (20 mesh) sieves. Subsequently, they were sent for chemical analysis of the plant material.

2.3. Chemical analysis

The chemical analyses of the macronutrients (N, P, K, Ca, Mg and S) contents of the litter were carried out at Laboratory for Agronomic, Environmental Analysis and Preparation of

Chemical Solutions. Nitrogen was extracted by sulfuric digestion and determined in a Kjeldahl distiller, while the other nutrients were extracted by nitro perchloric digestion, with phosphorus and sulfur being determined by optical spectrophotometry, and potassium, calcium and magnesium determined by atomic absorption spectrophotometry (Tedesco *et al*., 1995).

The amount of macronutrients (kg ha^{-1} $month^{-1}$) of the litter was obtained by multiplying the dry mass (kg ha^{-1} $month^{-1}$) by the nutrient content (g kg^{-1}), according to Equation 1 (Cuevas and Medina, 1986).

$$AN = [\text{Nutrient}] \text{ x DM} \quad (1)$$

Where:

AN = Amount of nutrients (kg ha^{-1} $monthly^{-1}$);

[Nutrient] = Nutrient content (g kg^{-1});

DM = Dry mass (kg ha^{-1} monthly $^{-1}$).

2.4. Litter decomposition

The indirect estimate of the decomposition constant (k) was obtained from the relationship between the annual litter input (L) and the annual mean of the accumulated litter (Xss), according to Equation 2, proposed by Olson (1963).

$$k = L/Xss \quad (2)$$

Where:

k = decomposition constant;

L = annual litter production (kg ha^{-1}); and

Xss = annual average of litter on the ground (kg ha^{-1}).

The average renewal time was estimated using the 1 / k ratio. The time required for decomposition of 50 and 95% of litter was through Equations 3 and 4, respectively (Shanks and Olson, 1961).

$$T50\% = 0{,}693/k \quad (3)$$

$$T95\% = 3/k \quad (4)$$

2.5. Statistics and data analysis

Data analysis was performed using the R software (R CORE TEAM, 2016). For the amount and concentration of nutrient in litter, the 12 plots represented the repetitions while the seasons were the treatments: summer (January, February and March), autumn (April, May and June), winter (July, August, and September) and spring (October, November, and December). After statistically significant difference was verified by applying the ANOVA Test (p-value ≤ 0.05).

The analysis of the influence of meteorological variables on litter was verified through Spearman's correlation between the amounts of litterfall dry mass and the meteorological elements (air temperature, precipitation and evapotranspiration) for the period of this study.

3. RESULTS AND DISCUSSION

3.1. Leaflitter storage

The average accumulated litter in the seasons over the two years was 5269 kg ha^{-1}, not statistically different by the t test ($p \leq 0.05$). Considering the averages of the first and second year, litter stocks were 5457 and 5079 kg ha^{-1} respectively. Regarding litterfal, the highest values were found in the winter seasons and the lowest values in the autumn season (Figure 2).

In the accumulated monthly litter, the variation occurs irregularly. In the first year, the months of August, October, November, December, January and February had greater accumulation, while in the second year the greatest accumulation occurred between May and October and in the month of February.

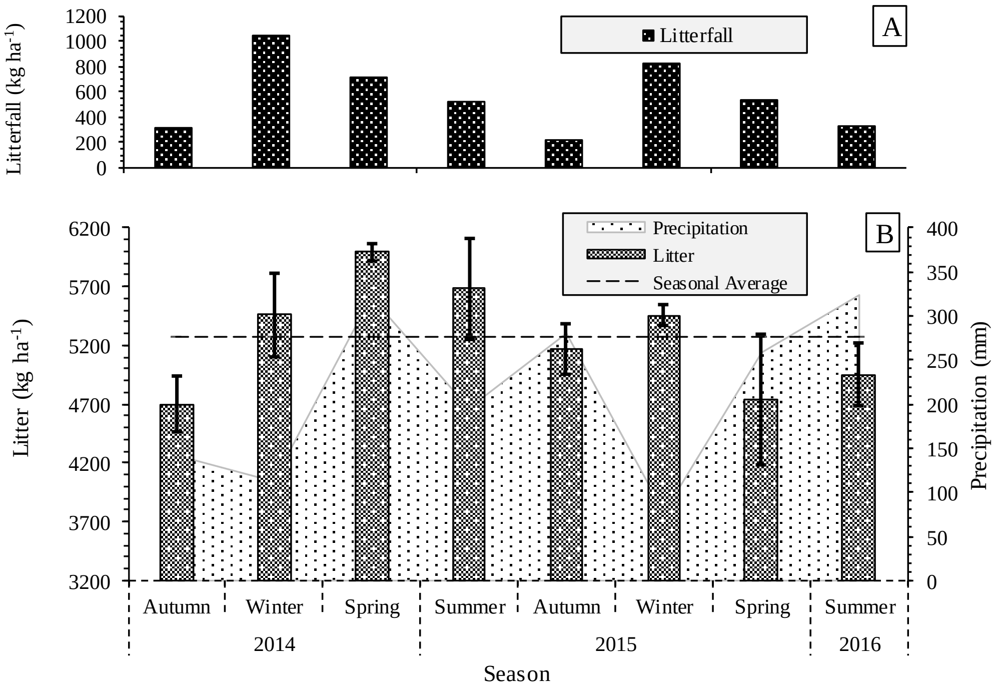


Figure 2. Seasonal Litterfall (A) and Litter (B) variation over two years.

The variation of litter accumulated on the forest floor, observed in the present study, occurs as a function of the difference between litter deposition and its decomposition, not showing a seasonal pattern. According to Delarmelina (2015), it is expected that there is no distinction of a seasonal pattern of litter accumulation due to greater temporal variability of litter deposited, indicating different dynamics of variation for both processes.

Assessing the litter stock on the forest floor in a Submontane Seasonal Semi-deciduous Forest, Godinho *et al*. (2014) found an accumulated 5500 kg ha^{-1}. This value is similar to the mean values in our study, 5269 kg ha^{-1} over two years. In an area undergoing restoration in a region of the Dense Ombrophilous Lowland Forest, Caldeira *et al*, (2019) evaluated litter accumulation under different spacing and number of species. The authors found values ranging from 3910 kg ha^{-1} for the treatment with greater spacing with fewer species to 5110 kg ha^{-1} for the treatment with smaller spacing and with greater species diversity. The average litter accumulation was 4440 kg ha^{-1}.

The amount of litter found in this study was slightly lower than that found in other studies considering the same forest formation. In a Seasonal Semideciduous Forest in the Cerrado biome, Martins *et al*. (2021) found 8300 kg ha^{-1} of accumulated litter, or 57% more. Vital *et al*. (2004), considering the same forest formation located in the state of São Paulo, found an average annual accumulation of 6227 kg ha^{-1}, that is, 18% more than in our study. Santos *et al*. (2019) evaluating the same type of forest formation in the state of Rio de Janeiro found an average stock of 7295 kg ha^{-1}, ranging from 5760 kg ha^{-1} in the rainy season to 8830 kg ha^{-1} in the dry season. One of the possible answers for the lower litter accumulation is justified by the fact that it is a secondary forest succession, being 26% Early Secondary Stage.

Comparing the litter stock in an area under restoration with a primary Atlantic forest, Correia *et al.* (2016) concluded that even after 23 years, the restored area did not reach the levels of accumulated litter in the primary forest. The study by Câmara (2020) shows that litter accumulation was greater in the fragment under advanced stage of regeneration 3529 kg ha^{-1} against only 2146 kg ha^{-1} in the fragment in the early stage. Another work that contributes to this justification is the research by Pinto *et al*. (2009), who, evaluating litter accumulated in Seasonal Semideciduous Forest, found 7008 kg ha^{-1} and 4648 kg ha^{-1} in formations in the maturity and initial stages, respectively.

Litter accumulation is influenced by biological activity, abiotic conditions and litter production rate. The greater the litter production, the greater the tendency of litter accumulation on the ground. The successional stage is another determining factor, thus, more preserved ecosystems present greater production (Pezzato and Wisniewski, 2006; Pinto *et al.* 2009).

There was no seasonality in litter accumulation for the first year, indicating that there was no effect of climatic elements on this variable in this period (Table 3). However, in the second year, the accumulation was significantly and negatively correlated with precipitation, not having a significant correlation with the other climatic elements.

Table 3. Spearman correlation between litterfall and meteorological variables over two years in the RPPN Fazenda Boa Esperança, Cachoeiro de Itapemirim, ES.

Year	Temp.[1]	Precipitation[2]								Et0[3]
		0	1	2	3	4	5	6	7	
2014-2015	0.25	-0.07	-0.17	-0.26	-0.22	-0.46	-0.33	-0.44	-0.57	0.47
2015-2016	-0.37	-0.67*	-0.19	-0.18	0.51	0.39	-0.38	-0.14	-0.09	-0.05

[1]Average air temperature; [2]Precipitation (0: referring to the month of collection; 1: referring to the first month prior to the collection; 2: referring to the second month prior to the collection; 3 referring to the third month prior to the collection; 4: referring to the fourth month previous to the collection to the collection; 5: referring to the accumulated precipitation of three months before the collection; 6: referring to the accumulated precipitation of four months before the collection; and 7: referring to the accumulated precipitation of five months before the collection); [3]Evapotranspiration.
* Significant by t test ($p \leq 0.05$).

3.2. Chemical composition of litter deposited on the soil

There was a statistically significant difference between the seasons for all nutrients, except for sulfur, Figure 3. The mean concentration of nutrients in litter accumulated throughout the seasons followed the order: Ca > N > Mg > S > K > P, representing 23.3, 17.7, 2.5, 2.2, 1.8 and 0.6 g kg^{-1} respectively.

The concentrate of the nutrients phosphorus, potassium, calcium and magnesium decreased in the seasons of the second year evaluated, showing a statistically significant difference ($p \leq 0.05$). Graphically, the seasonal means of sulfur decreased; however, they did not differ statistically ($p \leq 0.05$).

There was a variation in the order of the amount of nutrients observed in the two years of the present study. When compared to other studies, the first year had the same decreasing order observed by Delamerlina (2015): Ca > N > Mg > K > S > P, differing from the work by Godinho *et al*. (2014), where the descending order was Ca > N > K > Mg > S > P. Studying the nutrient concentration in litter accumulated from leaves and branches, in a Dense Ombrophylous Lowland Forest, Caldeira *et al,* (2019) found a similar order: Ca > N > Mg > K > S > P.

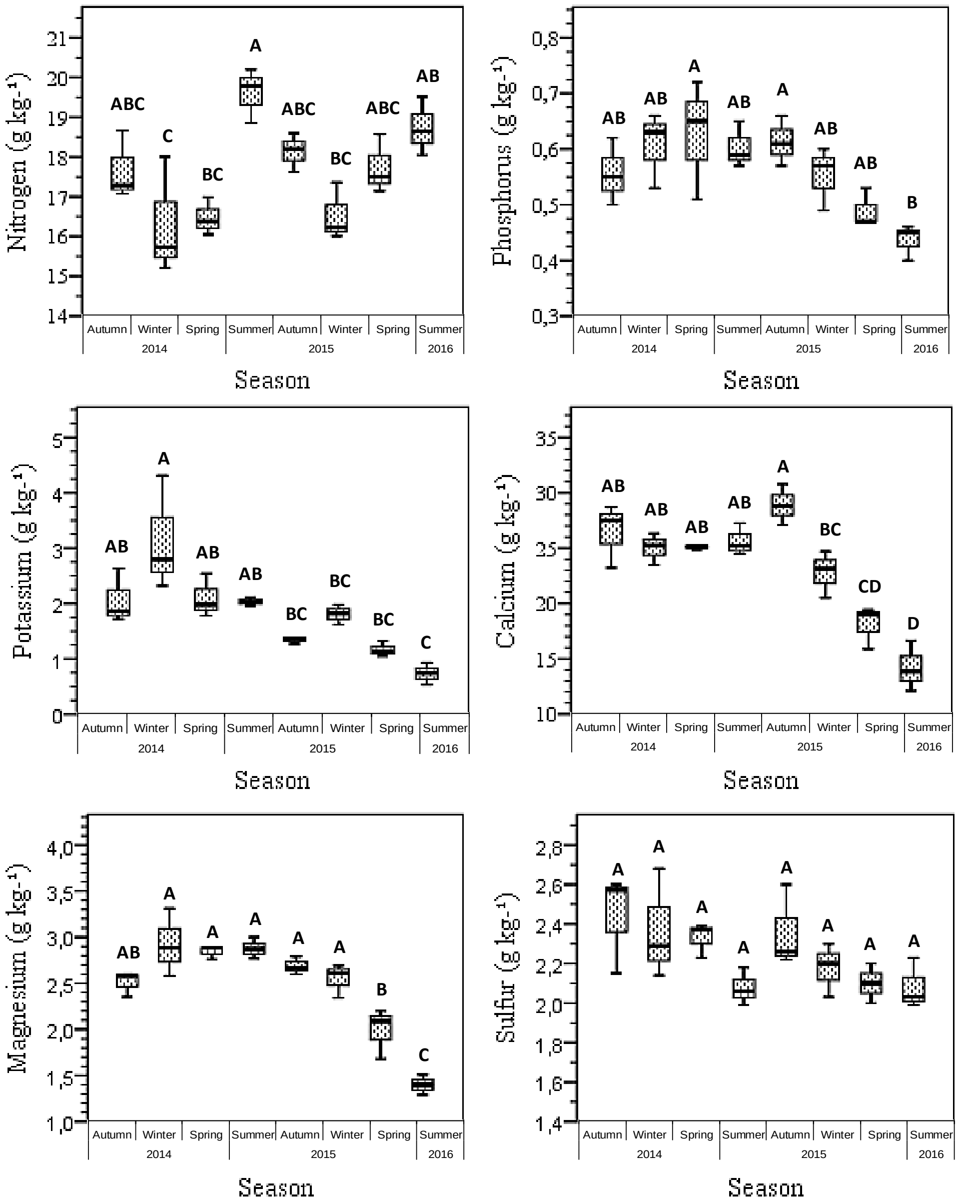


Figure 3. Seasonality of nutrient concentration in accumulated litter over the seasons in a fragment of Atlantic Forest.
Different letters between seasons in a given nutrient represent a statistically significant difference ($p < 0.05$). Bars: Confidence Interval (95%).

Evaluating nutrient cycling in a Seasonal Semideciduous Forest in the state of Rio de Janeiro, Santos *et al*. (2019) found the following order of concentration in both the dry and rainy seasons: Ca > N > Mg > P > K. Also considering the study by Santos *et al*. (2019), the average concentration of nutrients was close to the present research.

Assessing nutrient cycling along a secondary succession of semi-evergreen tropical forest in South-Eastern Mexico, Sánchez-Silva *et al.* (2018) found that leaf litter contents were lower in the dry season, the same behavior in the present study.

The average annual stock, considering the sum of all nutrients in the accumulated litter, was 253.7 kg ha^{-1} $year^{-1}$. Considering the means of each year separately, the nutrient stock was 278.3 and 229.2 kg ha^{-1} for the first and second year, respectively. The 2015 summer season was the one with the highest nutrient stock, 299.6 kg ha^{-1}, followed by the spring of 2014 with 296.5 kg ha^{-1} and the autumn season of 2015 with 279.5 kg ha^{-1}. The smallest amount of nutrients was found in the last season evaluated, summer 2016 with 186.2 kg ha^{-1}.

The nutrient calcium was the most representative, contributing 123 kg ha^{-1} year^{-1}, or 48.6% of the total accumulated, Figure 4. Following the order of importance, we have N, Mg, S, K and P, contributing with the amounts of 92.8, 13.2, 11.8, 9.6, 3.0 kg ha^{-1} year^{-1}, representing 36.6, 5.2, 4.7, 3.8 and 1.2% of the total.

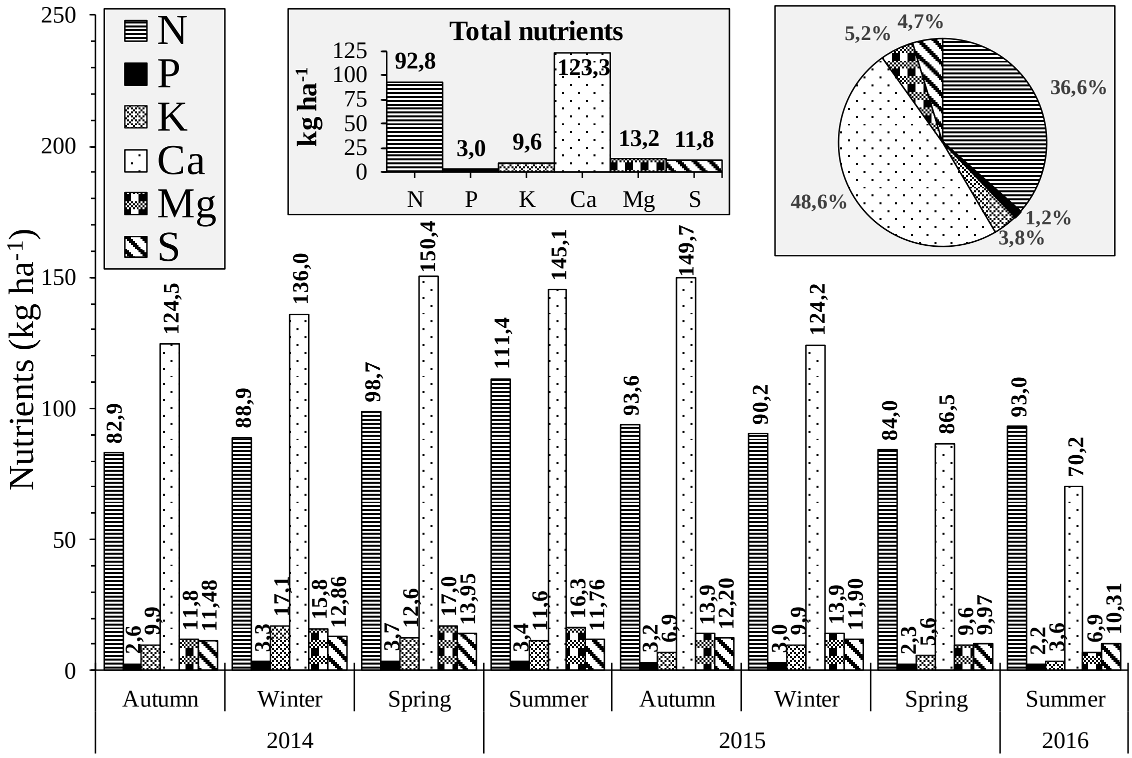


Figure 4. Annual average and seasonal amounts of nutrients in accumulated litter in a fragment of Forest Atlantic.

The largest amounts of Ca and N observed are associated with a large amount of these nutrients in the plant tissue, and Ca still has a slower release by the decomposition process as it is a structural component of plant tissue, with low mobility in this tissue (Godinho *et al*., 2014). Among all the elements analyzed, Ca has the lowest mobility within the plant, its concentration is low in the cytosol of cells, with a tendency to accumulate in the apoplast, vacuole and endoplasmic reticulum (Taiz and Zeiger, 2017).

Considering the sum of Ca + N, the participation of these macronutrients was 85.2%. These values are similar to the study by Godinho *et al.* (2014), in which they found a participation of 87.3%. Assessing the litter accumulated in a Montana Dense Ombrophilous Forest, Freitas *et al*. (2015) found a higher nutrient stock, being 293 kg ha$^{-1;}$ however, the participation of N + Ca was similar to the present study, 83, 7%.

Contrary to Ca, the low P content is due to the high mobility of the element from physiologically mature tissues to younger tissues (Malavolta *et al.*, 1997). Regarding K^+, the highest concentrations were verified in the season with the lowest rainfall recorded, winter 2014. This behavior was also observed by Vital *et al*. (2004), in which the decrease in precipitation provided considerable increases in K+. The explanation for these results is anchored in the fact that the nutrient is absorbed and transported in the ionic form K+, which makes it easily leached from the tissue surface during periods of rainfall. From a functional point of view, the element does not have a structural character but it is essential for osmotic regulation such as stomatal opening and closing (Taiz and Zeiger, 2017).

In an area under restoration process, Caldeira *et al.* (2017) evaluated the nutrient redistribution rate in two native tree species. For the authors, nutrients N, P and K were the most redistributed with means of 50.4, 70.9, 68.0% respectively. Considering the average redistribution only for the *Bixa arborea* species, the researchers found higher redistributions, 59.7, 84.9 and 72.6% for N, P and K, respectively.

In a Seasonal Semideciduous Forest, the average nutrient stock for the dry and wet season, Santos *et al.* (2019) found 102.5, 96.7, 37.3, 4.9 and 3.0 kg ha^{-1} for Ca, N, Mg, P and K, representing a total of 244.4 kg ha^{-1}. Considering the same nutrients as the work presented, the present study recorded a total of 241.9 kg ha^{-1}.

According to Caldeira *et al*. (2019), the return of nutrients to the soil through accumulated litter, is extremely important in improving and maintaining soil fertility. Also according to Delarmina (2015), because they are part of organic compounds, the nutrients from accumulated litter are less susceptible to leaching than those stored in the soil.

3.3. Litter decomposition

The mean decomposition constant for the two years was k = 1.29 (Table 4). The first year had k = 1.43, indicating a faster decomposition rate when compared to the average of the two years. The second year of monitoring resulted in a decomposition coefficient k = 1.13, indicating a decrease in decomposition when compared to the first year of evaluation and the average of the entire period studied.

The average renewal time 1/k was 281 days, decreasing to 255 days in the first period. The renewal time in the second period was longer, requiring 321 days. Estimates for the average decomposition of 50% and 95% of litter were 196 and 850 days, respectively. It is noted, however, that this time decreases substantially in the first year, requiring only 176 and 764 days, respectively. Approximately 969 days were required for 95% of the litter to decompose.

Table 4. Decomposition constant k, renewal time T50% and T95% of litter accumulated in the RPPN Fazenda Boa Esperança, Cachoeiro de Itapemirim, ES.

Year	k	1/k	T 50%	T 95%	1/k	T 50%	T 95%
		Year			Days		
1st year	1.43	0.70	0.48	2.09	255	176	764
2nd year	1.13	0.88	0.61	2.65	321	224	969
Average	1.29	0.77	0.54	2.33	281	196	850

As shown, the decomposition coefficients in the different periods are within the range considered acceptable. The values of K considered high, according to Olson (1963), are characteristic of tropical forests and these range from 1 to 4. For the present study, the average was 1.29. Evaluating the decomposition coefficients in 3 fragments under increasing stages of regeneration in Seasonal Semideciduous Forest, Câmara (2020) found much lower coefficients: 0.93, 0.99 and 1.05. According to the author, the ecosystem is not yet re-established, with deficiencies still persisting in the processes of transformation of plant material and nutrient cycling.

Evaluating the same type of forest formation, Vital *et al*. (2004); Cunha *et al.* (1997); Patricia and Morellato (1992) reported k = 1.71, 1.2, 1.6 respectively. Pinto *et al.* (2009) compared the decomposition rate in two seasonal semideciduous forest fragments and the value of k was 1.36 and 1.26 in the fragment in the initial stage and in the maturity stage, respectively. These results are similar to the present study. According to the authors, the estimated litter half-life was 186 and 200 days for the early and mature forest, respectively. For 95% of the litter to be decomposed, the estimated time was 807 and 869 days. respectively.

Although there has been variation in the decomposition constant over the years evaluated, it is known that litter in a natural environment is the result of interactions between the characteristics of the environment, the decomposing fauna, and the composition of the plant residue (Swift *et al.*, 1979; Correia and Andrade, 1999). The variation in the accumulated litter decomposition constant is closely linked to the alternation of weather conditions provided over the two years analyzed, especially the rainfall regime (Sanches *et al*., 2009).

According to Sánchez-Silva *et al*. (2018), the average time needed to decompose 50% of the litter in a secondary succession of semi-evergreen tropical forest was 6 months, or 180 days. These results are similar to our findings. The authors observed that the remaining leaf mass was greater in mature forest than in forests in less developed stages.

In tropical forest fragments whose structure presents greater development and with the presence of different strata, the smallest variation in light, temperature and humidity favors the diversification of decomposing organisms, favoring the activity of these organisms in the decomposition process (Pereira *et al*., 2013). In these environments, according to Sanches *et al*. (2009), the seasonality of precipitation directly regulates the activities of decomposing organisms, as the return of precipitation after a dry period increases the biodiversity of these organisms, contributing to the increase in the rate of litter decomposition.

According to Pires *et al*. (2006), accumulated litter regulates several processes in the dynamics of ecosystems, being responsible for the stability of the nutrient cycling system, releasing them to the soil through decomposition. In weathered tropical forest soils with low natural fertility, the maintenance of litter accumulated on the forest floor becomes very important in the nutrient cycling process (Godinho *et al*., 2014).

4. CONCLUSION

The average litter stocks and nutrients on the forest floor are in line with other studies, indicating that the Seasonal Semi-Deciduous Atlantic Forest Fragment presents indicators of sustainability in nutrient cycling.

The statistical analysis did not show a seasonal pattern for the amount of litter accumulated or the influence of meteorological variables, but there was a statistical difference between the concentrations of nutrients in the different seasons of the year.

The time required for the decomposition of 50% and 95% of litter was 196 and 850 days. These results show that the ecosystem is sustainable from the point of view of cycling and nutrient release.

5. ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

6. DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

7. REFERENCES

ALVARES, C. A. *et al*. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711-728, 2014. https://doi.org/10.1127/0941-2948/2013/0507

ARCHANJO, K. M. P. A.; SILVA, G. F.; CHICHORRO, J. F.; SOARES, C. P. B. Estrutura do componente arbóreo da Reserva Particular do Patrimônio Natural Cafundó, Cachoeiro de Itapemirim, Espírito Santo, Brasil. **Floresta**, v. 42, n. 1, p. 145-160, 2012. http://dx.doi.org/10.5380/rf.v42i1.26311

BALIEIRO, F. C.; FRANCO, A. A.; PEREIRA, M. G.; CAMPELLO, E. F. C.; DIAS, L. E. S. M. et al. Dinâmica de serapilheira e transferência de nitrogênio ao solo em plantios de *Pseudosamanea guachapele* and *Eucalyptus grandis*. **Pesquisa Agropecuária Brasileira**, v. 39, n. 6, p. 597-601, 2004. https://doi.org/10.1590/S0100-204X2004000600012

BELLO, C.; GALETTI, M.; PIZO, M. A.; MAGNAGO, L. F. S.; ROCHA, M. F.; LIMA, R. A. F. et al. Defaunation affects carbon storage in tropical forests. **Science Advances**, v. 1, n. 11, e1501105, 2015. https://doi.org/10.1126/sciadv.1501105

BRASIL. Portaria n° 62-N, de 19 de maio de 1998. **Diário Oficial da União - República Federativa do Brasil**, Brasília, 20 maio 1998. Seção 1, p. 105.

CALDEIRA, M. V. W.; VITORINO, M. D.; SCHAADT, S. S.; MORAES, E.; BALBINOT, R. Quantificação de serapilheira e de nutrientes em uma Floresta Ombrófila Densa. **Semina**: Ciências Agrárias, v. 29, n. 1, p. 53-68, 2008.

CALDEIRA, M. V. W. *et al*. Litter as an Ecological Indicator of Forest Restoration Processes in a Dense Ombrophylous Lowland Forest. **Floresta e Ambiente**, v. 26, n. spe1 e20180411, 2019. https://doi.org/10.1590/2179-8087.041118

CALDEIRA, M. V. W.; MOREIRA, F. L.; GONÇALVES, E. O.; CAMPANHARO, I. F.; CASTRO, K. C. Teor e redistribuição de nutrientes nas folhas de Bixa arborea Huber. e Joannesia princeps Vell. em área de restauração florestal. **Scientia Forestalis**, v. 45, n. 114, p. 305-317, 2017. https://dx.doi.org/10.18671/scifor.v45n114.07

CÂMARA, Y. B. **Aporte, acúmulo e decomposição de serapilheira em três fragmentos de mata atlântica com diferentes estágios de regeneração**. 2020. 75F. Dissertação (Mestrado) - Universidade Federal do Rio Grande do Norte, Natal, 2020.

CORREIA, G. G. de S. *et al*. Estoque de serapilheira em floresta em restauração e em Floresta Atlântica de tabuleiro no sudeste brasileiro. **Revista Árvore**, v. 40, 2016. https://doi.org/10.1590/0100-67622016000100002

CORREIA, M. E. F.; ANDRADE, A. G. Formação de serrapilheira e ciclagem de nutrientes. *In*: SANTOS, G. A; CAMARGO, F. A. de. **Fundamentos da matéria orgânica do solo: ecossistemas tropicais e subtropicais.** Porto Alegre: Genesis, 1999. p. 209-214.

CUEVAS, E.; MEDINA, E. Nutrient dynamics within amazonian Forest ecosystems. **Oecologia**, v. 68, p. 466-472, 1986. https://doi.org/10.1007/BF01036756

CUNHA, G. C. **Aspectos da ciclagem de nutrientes em diferentes fases sucessionais de uma floresta estacional do Rio Grande do Sul.** 1997. 86f. Dissertação (Mestrado em Ciências/Ciências Florestais) – Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 1997.

DELARMELINA, W. M. **Fertilidade, estoque de carbono orgânico do solo e serapilheira em uma Floresta Estacional Semidecidual Submontana**. 2015. 123 f. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal do Espírito Santo, Jerônimo Monteiro, 2015.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Manual de métodos de análise de solos.** 2. ed. Rio de Janeiro: Embrapa Solos, 2011.

FAHRIG. L. Effects of habitat fragmentation on biodiversity. **Annual Review of Ecology, Evolution and Systematics**, v. 34, p. 487–515, 2003. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

FASSEBENDER, H. W. **Modelos edafológicos de sistemas agroforestales**. 2. ed. Turrialba: CATIE, 1993. 491 p,

FUNDAÇÃO SOS MATA ATLÂNTICA. **Atlas dos Remanescentes Florestais da Mata Atlântica - Período 2018-2019**. 2020. Available at: http://mapas.sosma.org.br. Access: 23 June 2021.

FREITAS, C. A. A. *et al*. Serapilheira acumulada em complexo rupestre de granito, Mimoso do Sul, ES. **Revista Árvore**, v. 39, n. 4, p. 671-681, 2015. https://doi.org/10.1590/0100-67622015000400009.

GODINHO, T. de O. *et al*. Quantificação de biomassa e nutrientes na serapilheira acumulada em trecho de Floresta Estacional Semidecidual Submontana, ES. **Cerne**, v. 20, n. 1, 2014. https://doi.org/10.1590/S0104-77602014000100002.

HAAG, H. P. **Ciclagem de nutrientes em florestas tropicais**. Fundação Cargill, 1985. 144p.

INCAPER. **Assistência Técnica e Extensão Rural**. 2017. Available at: http://hidrometeorologia.incaper.es.gov.br/?pagina=atlas_pluvio. Access: 19 Jan. 2017.

IBGE. **Manual técnico da vegetação brasileira.** 2 ed. Rio de Janeiro, 2012. 271 p.

KRAMER, J. P.; KOZLOWSKI, T. **Fisiologia das Árvores**. Lisboa: Fundação Calouste Gulbekian, 1960. 745 p

KRISHNA, M. P.; MOHAN, M. Litter decomposition in forest ecosystems: a review. **Energy, Ecology and Environment,** v. 2, p. 236–249, 2017. https://doi.org/10.1007/s40974-017-0064-9.

LALIBERTE, E.; GRACE, J. B.; HUSTON, M. A.; LAMBERS, H.; TESTE, F. P.; TURNER, B. L. et al. How does pedogenesis drive plant diversity? **Trends in Ecology & Evolution**, v. 28, p. 331–340, 2013. https://doi.org/10.1016/j.tree.2013.02.008

MACHADO, D. L. *et al*. Ciclagem de nutrientes em diferentes estágios sucessionais da Mata Atlântica na bacia do Rio Paraíba do Sul, RJ. **Bioscience Journal**, v. 31, n. 4, p. 1222-1237, 2015.

MALAVOLTA, E.; VITTI, G.; OLIVEIRA, S. A. **Avaliação do estado nutricional das plantas: princípios e aplicações**. 2. ed. Piracicaba: Potafos, 1997. 319 p.

MARTINS, T. O.; SILVA-NETO, C. M.; SIQUEIRA, K. N.; CARVALHO, H. C. S.; MORAES, D. C.; SILVA, P. H. F. et al. Accumulated litter and nutrient stock in biomass and soil in forest formations in the cerrado. **Scientia Forestalis**, v. 49, n. 129, e3427. 2021. https://doi.org/10.18671/scifor.v49n129.18

MYERS, N.; MITTERMEIER, R. A.; MITTERMEIER, C. G.; DA FONSECA, G. A. B.; KENT, J. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, p. 853–858, 2000.

ODUM, E. P. The strategy of ecosytem development. **Science**, v. 164, p. 262-270, 1988.

OLSON, J. S. Energy-storage and balance of producers and decomposers in ecological-systems. **Ecology**, v. 44, n. 2, p. 322-331, 1963.

PATRICIA, L.; MORELLATO, C. Nutrient cycling in two south-east brazilian forests I-Litterfall and litter standing crop. **Journal Tropical Ecology**, v. 8, p. 205-205, 1992. https://doi.org/10.1017/S0266467400006362

PEREIRA, G. H. A.; PEREIRA, M. G.; ANJOS, L. H. C.; AMORIM, T. A.; MENEZES, C. E. G. Decomposição da serapilheira, diversidade e funcionalidade de invertebrados do solo em um fragmento de Floresta Atlântica. **Bioscience Journal**, v. 29, n. 5, p. 1317-1327, 2013.

PEZZATO, A. W.; WISNIEWSKI, C. Produção de serapilheira em diferentes seres sucessionais da Floresta Estacional Semidecidual no oeste do Paraná. **Floresta**, v. 36, n. 1, p. 111-120, 2006. http://dx.doi.org/10.5380/rf.v36i1.5596

PEZZOPANE, J. E. M.; CASTRO, F. S.; PEZZOPANE, J. R. M.; CECÍLIO, R. A. **Agrometeorologia: aplicações para o Espírito Santo**. Alegre: UFES, 2012. 178p.

PINTO, S. I. C.; MARTINS, S. V.; BARROS, N. F.; DIAS, H. C. T. Ciclagem de nutrientes em dois trechos de Floresta Estacional Semidecidual na reserva florestal Mata do Paraíso em Viçosa, MG, Brasil. **Revista Árvore**, v. 33, n. 4, p. 653-663, 2009.

PIRES, L. A.; BRITEZ, R. M.; MARTEL, G. E.; PAGANO, S. N. Produção, acúmulo e decomposição da serapilheira em uma restinga da Ilha do Mel, Paranaguá, PR, Brasil. **Acta Botânica Brasilica**, v. 20, n. 1, p. 173-184, 2006. https://doi.org/10.1590/S0102-33062006000100016

R CORE TEAM. **R: A Language and Environment for Statistical Computing.** Vienna, 2016.

SÁNCHEZ-SILVA, S.; JONG, B. H. J.; ARYAL, D. R.; HUERTA-LWANGA, E.; MENDOZA-VEGA, J. Trends in leaf traits, litter dynamics and associated nutrient cycling along a secondary successional chronosequence of semi-evergreen tropical forest in South-Eastern Mexico. **Journal of Tropical Ecology**, v. 34, n. 6, p. 364–377, 2018. https://doi.org/10.1017/s0266467418000366

SANCHES, L.; VALENTINI, C. M. A.; BIUDES, M. S.; NOGUEIRA, J. S. Dinâmica sazonal da produção e decomposição de serapilheira em floresta tropical de transição. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 13, n. 2, p. 183–189, 2009.

SANTOS, G. L. *et al*. Ciclagem de nutrientes em diferentes condições topográficas em Floresta Estacional Semidecidual, Pinheiral-RJ. **Ciência Florestal**, Santa Maria-RS, v. 29, n. 4, p. 1737-1747, 2019. https://doi.org/10.5902/1980509826880.

SCORIZA, R. N.; PIÑA-RODRIGUES, F. C. M. Influência da precipitação e temperatura do ar na produção de serapilheira em trecho de Floresta Estacional em Sorocaba, SP. **Floresta**, v. 44, n. 4, p. 687-696, 2014. http://dx.doi.org/10.5380/rf.v44i4.34274

SHANKS, R.; OLSON, J. S. First year breakdown of leaf litter in Southern Appalachian Forests. **Forest Science**, v. 134, p. 194-195, 1961. https://doi.org/10.1126/science.134.3473.194

SIMPSON, A. J.; SIMPSON, M. J.; SMITH, E.; KELLEHER, B. P. Microbially derived inputs to soil organic matter: are current estimates too low? **Environmental Science & Technology,** v. 41, p. 8070–8076, 2007. https://doi.org/10.1021/es071217x

SWIFT, M. J.; HEAL, O. W.; ANDERSON, J. M. **Decomposition in Terrestrial Ecosystems**. Oxford: Blackwell, 1979. 372p.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**. 6. ed. Porto Alegre: Artmed, 2017. 888 p.

TEDESCO, M. J.; GIANELLO, C.; BISSANI, C. A.; BOHNEN, H.; VOLKWEIS, S. J. **Análise de solo, plantas e outros materiais**. 2. ed. Porto Alegre: Universidade Federal do Rio Grande do Sul, 1995. 174p.

VITAL, A. R. T.; GUERRINI, I. A.; FRANKEN, W. K.; FONSECA, R. C. B. Produção de serapilheira e ciclagem de nutrientes de uma floresta estacional semidecidual em zona ripária. **Revista Árvore**, v. 28, n. 6, p. 793-800, 2004. https://doi.org/10.1590/S0100-67622004000600004

WOLF, A.; DOUGHTY, C. E.; MALHI, Y. Lateral Diffusion of Nutrients by Mammalian Herbivores in Terrestrial Ecosystems. **PLoS ONE**, v. 8, n. 8, p. 1–10, 2013. https://doi.org/10.1371/journal.pone.0071352.

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Participatory assessment to define indicators for monitoring water-based payment of ecosystem services programs in Brazil

ARTICLES doi:10.4136/ambi-agua.2796

Received: 15 Sep. 2021; Accepted: 07 Feb. 2022

Ana Paula Dias Turetta*; Elaine Cristina Cardoso Fidalgo; Rachel Bardy Prado; Azeneth Eufrasino Schuler; Heitor Luiz da Costa Coutinho (in memorian)

Centro Nacional de Pesquisa em Solos. Empresa Brasileira de Pesquisa Agropecuária (Embrapa Solos), Rua Jardim Botânico, n° 1024, CEP: 22460-000, Rio de Janeiro, RJ, Brazil. E-mail: elaine.fidalgo@embrapa.br, rachel.prado@embrapa.br, azeneth.schuler@embrapa.br, heitor.coutinho@embrapa.br
*Corresponding author. E-mail: ana.turetta@embrapa.br

ABSTRACT

The cascade model is a commonly applied framework to evaluate ecosystem services provision, highlighting their benefits to society and assigning non-monetary and monetary values to these services. Adapting this model, we present a methodology to establish the most suitable indicators for monitoring Payment for Ecosystem Services (PES) programs related to water resources in Brazil. Through a participatory process, a set of ecosystem functions indicators were assigned for each ecosystem service (ES) considered in the study. The indicators were then classified following these criteria: clarity, viability, sensitivity, and relevance. The indicators were organized by their final score according to each evaluated criterion. The results demonstrated that "clarity" and "relevance" criteria were those most important for the experts to choose an indicator. In general, we could also observe a preference for analytical and well-established indicators in the literature for all ES evaluated. The indicators list presented can support the PES program monitoring in Brazil. Additionally, the methodology developed can be easily applied in other areas and provide the definition of the most suitable indicators to monitor water-based PES in different Brazilian contexts.

Keywords: cascade model, ecosystem functions, ecosystem service indicators, tropical areas.

Avaliação participativa para definição de indicadores para monitoramento de programas de pagamento de serviços ecossistêmicos hídricos no Brasil

RESUMO

O modelo em cascata é uma estrutura comumente utilizada para avaliar a provisão de serviços ecossistêmicos, destacando seus benefícios para a sociedade e atribuindo valores não monetários e monetários a esses serviços. Adaptando este modelo, apresentamos uma metodologia para estabelecer os indicadores mais adequados para monitorar programas de Pagamento por Serviços Ecossistêmicos (PSA) hídricos no Brasil. Por meio de um processo participativo, um conjunto de indicadores para as funções ecossistêmicas foi atribuído a cada serviço ecossistêmico (SE) considerado no estudo. Em seguida, os indicadores foram classificados segundo os seguintes critérios: clareza, viabilidade, sensibilidade e relevância. Os resultados demonstraram que os critérios clareza e relevância foram os que os especialistas consideraram como os mais importantes para escolher um indicador. No geral, podemos observar que para todos os SE avaliados, os indicadores de maior pontuação foram aqueles analíticos e bem estabelecidos pela literatura. A lista de indicadores pode apoiar o monitoramento de PSA na Floresta Atlântica. Além disso, a metodologia desenvolvida pode ser facilmente aplicada em outras áreas e fornecer suporte para a definição dos indicadores mais adequados para monitorar o PSA baseado em água em diferentes contextos nos trópicos.

Palavras-chave: áreas tropicais, funções do ecossistema, indicadores de serviços ecossistêmicos, modelo em cascata.

1. INTRODUCTION

Ecosystem functions (EF) can be defined as "the capacity of natural processes and components to provide goods and services that satisfy human needs, directly or indirectly" (De Groot, 1992; De Groot *et al.*, 2002).

This idea can be considered one of the key aspects regarding the ecosystem services (ES) concept, defined as the direct and indirect benefits people obtain from ecosystems, contributing to human well-being (MEA, 2005; TEEB, 2010).

More recently, the Common International Classification of Ecosystem Services (CICES) presented the vision of the final services (FS), which retain a connection to the underlying ecosystem functions, processes and structures that generate them and most directly affect the human well-being (Haines-Young and Potschin, 2018).

These approaches have in common the aim to integrate a broad range of ecosystem functions and services, considering their interdependencies and social demand to improve the decision-making measures (Primmer and Furman, 2012). An example is the cascade model (Haines-Young and Potschin, 2010) which is a commonly applied framework consisting of steps, starting at generating ES processes until their benefits and non-monetary/monetary values (Spangenberg *et al.*, 2014). It helps to operationalize the links among ecosystem properties (biophysical structure or process), ecosystem functions, ecosystem services and, contribute to the valuation procedure. Realistically, those links are not as simple and linear as they appear in this framework, but such an approach is useful to show the relations that are generated by the ESs, and consequently to better plan interventions (Pavan and Ometto, 2018).

De Groot *et al.* (2010a), working with the cascade framework, separated the benefits from their values. They argued that, if benefits are seen as gains in well-being generated by ecosystems, then it is clear that different groups may value these gains in different ways at different times, and indeed in different places (Fisher *et al.*, 2009). Despite this modification, the fundamental tenet of the ecosystem service paradigm remains: namely, that a service is only a service if a human beneficiary can be identified and that it is important to distinguish between the "final services" that contribute to people's well-being and the "intermediate ecosystem structures and functions" that give rise to them (Potschin and Haines-Young, 2011). Thus, we observe a clear focus on a more anthropocentric interpretation with a utilitarian background, which was already discussed by other authors. In Schoröter *et al.* (2014), it is possible to find a structured debate between opponents and proponents of the ecosystem services concept.

Additionally, Costanza *et al.* (2017) stated that the connections between ecosystem processes, functions, and benefits to humans are complex, nonlinear, and dynamic. These complex connections are poorly represented by a linear 'cascade', which assumes simple linkages and effects.

The key messages that seem to emerge from these debates is that, in relation to the cascade idea, whether or not it involves three, four or more steps, or how particular boxes are labelled, the fundamental task is to understand the mechanisms that link ecological systems to human well-being. The intention of the cascade idea is to highlight the essential elements that must be considered in any full analysis of an ecosystem service and the kind of relationships that exist among them (Potschin and Haines-Young, 2011).

For any ecosystem service, there are various attributes that could be measured, from the state of the underlying system, through the functioning of the system, to the services it provides, and the benefits gained by society (Spangenberg *et al.*, 2014). Key metrics that support and inform this process are the ecosystem service indicators (ESI).

ESI can be applied for different aspects of this 'flow', from the ecosystems that provide services, to the benefits that are obtained by people. They include measures of ecosystem processes and functions, their use (benefit) and impact (De Groot, 1992; Balmford *et al.*, 2008; Tallis *et al.*, 2008; De Groot *et al.*, 2010a; 2010b). The main challenge is to define the most suitable indicators, which meet the specific requirement for each ES. This means, understanding what needs to be known, and then choosing an appropriate combination among the plethora of potential indicators (Berghöfer and Schneide, 2015).

Successful natural resource management is dependent on effective knowledge exchange and utilization (Roux *et al.*, 2006; Fazey *et al.*, 2013). Knowledge exchange (KE) are processes that generate, share and/or use knowledge through various methods appropriate to the context, purpose, and participants involved. KE includes concepts such as sharing, generation, coproduction, co management, and brokerage of knowledge (Fazey *et al.*, 2013). There is no single optimum approach for integrating local and scientific knowledge and encouraging a shift in science from the development of knowledge integration products to the development of problem-focused, knowledge integration processes. These processes need to be systematic, reflexive and cyclic so that multiple views and multiple methods are considered in relation to an environmental management problem (Raymond *et al.*, 2010).

Fleischman and Briske (2016) highlight professional ecological knowledge (PEK), which differs from local ecological knowledge (LEK) because it is not grounded in direct experience of natural resources to support human livelihoods, and that it differs from scientific knowledge because it is not directly derived from systematic inquiry. PEK is a unique knowledge source. It includes best management practices, procedural manuals, and technical guides that often come to be thought of as verified scientific knowledge by personnel who use them. The knowledge used by professional resource managers, particularly those in public agencies, is important because these managers play an important role in decision making about public and private land use around the world.

The KE process can be facilitated by a participatory process. The experience from some projects has shown the potential of stakeholder engagement in natural resources management processes (sometimes referred to as ''diversity analysis'', e.g., Pain, 2004). This approach has been seen as a way of generating information on the ''relevant actors'' to understand their behavior, interests, agendas, and influence on decision-making processes (Brugha and Varvasovsky, 2000). Thus, stakeholder engagement can be useful for the definition of indicators resulting in more realistic, meaningful and achievable options than those set by top-down methods (Better Evaluation, 2014). Moreover, the indicators established in a participatory way can reduce costs and ensure continued monitoring of ecosystem functions and services.

In participatory processes, it is recommended to use a stakeholder analysis method to identify the most relevant stakeholders for each case that will be included in the further processes. The participation mechanism can occur at any stage of the evaluation process: its design, data collection, analysis, reporting or managing the study (OECD, 2011a; Guijt, 2014).

In the past two decades, the number of PES schemes has significantly increased; currently there are around 550 PES programs worldwide and approximately half of these are being implemented in Latina America (Salzman *et al.*, 2018). In Brazil, the most well-known water-based PES is the Water Producer Program of the National Water Agency (ANA), ongoing since 2005. This is a national program to stimulate the implementation of water-based PES projects in the strategic basins for restoration and water supply (ANA, 2012). The official website of the Water Producer Program informs that there are 29 projects underway (ANA, 2021).

While monitoring of PES projects is essential to identify PES effectiveness and their environmental and socioeconomic consequences, the lack of adequate monitoring has been identified as a major bottleneck of these programs worldwide (Pagiola and Platais, 2007; Engel *et al.*, 2008; Pagiola *et al.*, 2012). Lima *et al.* (2021) presents an overview of monitoring water-based PES in Brazil, pointing out its main characteristics such as analyzed parameters, frequency and also identifying gaps and proposing future perspectives.

Considering these aspects, we present the results of a participatory process for ESI selection for water-based Payment for Ecosystem Services (PES) in Brazil. We propose a set of indicators able to be used to monitor the results of interventions by these PES water-based projects. Additionally, the methodology developed can be easily applied in other areas and provide the definition of the most suitable indicators to monitor water-based PES in different contexts in Brazil.

2. MATERIAL AND METHODS

This study was developed in three steps: pre-selection of EF, ES and ESI; an expert participative workshop; ESI definition and ranking.

Step 1. Pre-selection of EF, ES and ESI

The pre-selection of EF and ES was based on Costanza *et al.* (1997) and MEA (2005). We identified and worked with the ES directly associated with water resources –water supply, water regulation, erosion control, soil quality and habitat regulation (Figure 1).

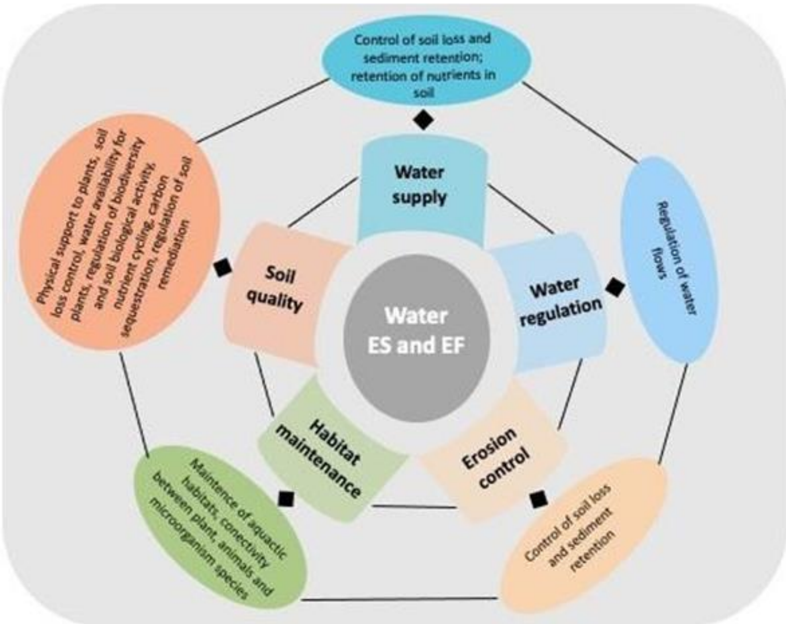


Figure 1. Ecosystem services (ES) and functions (EF) directly associated with water resources considered in this study, pre-selected from Costanza *et al.* (1997) and MEA (2005).

We then analyzed the indicators that are being used by the water-based PES programs in Brazil (Pocidônio and Turetta, 2012). From this evaluation, a preliminary indicator list was generated considering the suitability and effectiveness of each indicator, following the criteria proposed by OECD (2011b) in order to support the start of the participatory process.

Next, all of this information was organized following the cascade conceptual model

(Haines-Young and Potschin, 2010; Martin-Lopez *et al*., 2014) associated with two components: Structure and Function. We considered the structure as the ability of the biophysical environment to provide a particular ES. The function was considered the ecosystem mechanism by which the services are generated –for example, one of the forest cover functions is the potential of slowing the surface water flow, that is linked to sediment retention and soil loss control, a core capacity for the water supply/water quality ecosystem services (Figure 2).

Thus, structure and function are underpinning elements that determine the capacity of the ecosystem to deliver particular services (Haines-Young and Potschin, 2018).

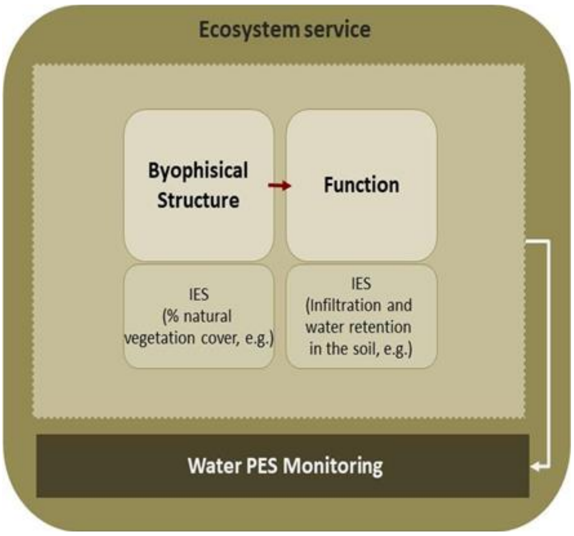


Figure 2. Conceptual model adopted for the selection of ESI.

Step 2. The expert participative workshop

A stakeholder analysis aims to evaluate and understand stakeholders from the perspective of an organization, or to determine their relevance to a project or policy (Brugha and Varvasovsky, 2000). Considering a variety of methods to identify stakeholders, as described by Reed *et al.* (2009), we selected the "project expert's consultation". We selected this method to add analytical horsepower, input external change force and stimulate exchange of knowledge in the scientific community.

Thus, a group of forty-two experts related to PES and with diverse backgrounds and expertise joined a workshop held at Embrapa Solos, in Rio de Janeiro, Brazil, on June 25th and 26th, 2013 (Turetta *et al.*, 2013).

The workshop was based on different dynamics, working with the whole group for general presentations, discussions, and decisions, and in separate working groups, according to their expertise: Group 1: Water regulation and erosion control; Group 2: Water supply; Group 3: Soil quality; Group 4: Habitat maintenance.

The first activity for the workshop participants was to analyze the preliminary list of EF and correlate with the indicators defined in Step 1. The experts were invited to review the list and to include and/or exclude EF and indicators following the recommendations in order to reduce the subjectivity and better standardize the process of ESI selection:

- Always taking into account the conceptual model defined in Step 1 (Figure 2);

- Selecting ESI for local scale application, as the monitoring of water-based PES is mostly "*in situ*";

- Prioritizing indicators with thresholds already referenced in the literature and/or in

Brazilian legislation;

- Indicators should be appropriate for the baseline (initial condition) and for the monitoring of water-based PES intervention impacts.

Step 3. ESI definition and ranking

Once the groups had finished the review and definition of EF and indicators for each ES established in Step 1, we invited the experts to assign the scores “1” (low), “2” (medium) or “3” (high) for each ESI considering the criteria: clarity, viability, sensitivity, relevance (Table 1):

Table 1. Criteria to ESI evaluation.

Criteria	Rationale
Clarity	Efficiency of understanding and communicating (association of the indicator response with the phenomenon) and simplicity of use of the indicator by the decision maker
Viability	Analysis costs and easiness of obtaining and analyzing the indicator
Sensitivity	Ability to detect the impacts and changes in the ecosystem
Relevance	Applicability of the indicator to demonstrate the ecosystems’ function

No weighing factor was applied during the evaluation process, i.e., all criteria influenced the evaluation equally.

As a result, a list containing the ecosystem functions for each ES and the indicators for monitoring interventions, by criteria, was generated.

We ran the “Mode” analysis to access the most frequent score of each indicator per criterion on Excel© software to obtain the final indicators matrix.

3. RESULTS AND DISCUSSION

3.1. Overall results

Selection of effective indicators is best achieved by developing conceptual models of the ecosystem and using these to pinpoint indicators that provide the required information (Queensland Government, 2020). Based on that, the results presented here followed the analysis considering the ESI per EF presented in the conceptual model applied in this study (Figure 2).

An overall analysis that emerged from the ESIs selection process is that the ecosystem services that are connected to a high number of functions are those harder to evaluate, as there are many aspects to be considered, including quantitative and qualitative characteristics.

The highest amount of EFs identified by the experts was in the "Soil Quality" ES, while the "Water supply" ES had the lowest EFs associated. (Table 2). However, “Water Supply” presented the highest number of correlated ESI, which reflects the number of parameters and indexes already established for the assessment of water quality status.

“Water regulation” and “Erosion control” were the services with fewer ESI associated with each EF and it was possible to distinguish that the “Water regulation” concentrated most of the ESIs regarding flow aspects, while “Erosion control” presented indicators focused on the soil loss parameters (Table 2).

In some cases, the same ESIs were suggested for more than one ES (Table 2). "Turbidity", for example, was recorded as "Water supply" and "Erosion control" services. It is not a surprise as this parameter is associated with the presence of suspended solids in the water, an important feature for both services.

Table 2. ESI organized by ES, EF and criteria. Dark gray represents mode 3 (it means, the highest mode value for the indicator performance per criterion); light gray represents score 2, and white, score 1, respectively.

Water supply					
Ecosystem Function	*Indicators*	*Clarity*	*Viability*	*Sensibility*	*Relevance*
Control of soil loss and sediment retention	Turbidity	3	3	3	3
	Total solids	3	3	3	3
	Suspended solids	3	3	2	2
Retention of nutrients in soil	Dissolved Oxygen (DO)	3	3	3	3
	pH	3	3	2	2
	Biochemical Oxygen Demand (BOD)	2	2	2	2
	Thermotolerant coliforms	3	2	2	2
	Total coliforms	3	2	3	3
	Water temperature	3	3	1	1
	Total nitrogen (TN)	2	3	3	3
	Nitrate	2	3	3	3
	Nitrite	2	1	2	2
	Ammonia Nitrogen	2	1	3	3
	Total Phosphorus (TP)	3	3	3	3
	Dissolved Inorganic Phosphorus (DIP)	2	2	2	2
	Total Dissolved Phosphorus (TDP)	1	1	2	2
	Nitrogen / phosphorus ratio (N/P)	2	2	2	2
	Heavy metals	2	2	2	2
	Pesticides	3	1	1	3

Continue…

Continued…

	Dissolved Organic Carbon (DOC)
	Total Organic Carbon (TOC)
	Dissolved Inorganic Carbon (DIC)
	Cations (sodium, calcium, potassium, magnesium)
	Anions (carbonates, bicarbonates, chlorides, sulfates, nitrates)
	Total Hardness
	Chlorophyll A
	Oils and greases
	Salinity
	Alkalinity
Retention of nutrients in soil	Presence of aquatic macrophytes
	Hormones
	Antibiotics
	Surface-active agents
	Chemical Oxygen Demand (COD)
	Metals (micronutrients)
	E.coli
	Virus
	Salmonella
	Electric Conductivity (EC)
	Phytoplankton algae

Continue…

Continued…					
Retention of nutrients in soil	Color				
	Benthic organisms				
	Standardized Toxicological Bioassays (Acute and Chronic)				
	Index of Biological Integrity (IBI)				
	Diversity Index				
	Species richness				
	Equitability (aquatic vertebrates, aquatic invertebrates - insects, zooplankton				
	Dissolved Oxygen (DO)				
	Cyanobacteria				
Water Regulation					
Ecosystem Function	*Indicators*	*Clarity*	*Viability*	*Sensibility*	*Relevance*
Maintenance of groundwater recharge	Groundwater level				
Maintenance of springs	Flow rate				
	Groundwater level				
Maintenance of reference flow	Groundwater level				
	Base flow coefficient (Qbase/precipitation)				
	Reference flow (Q7,0 or Q95)				
Attenuation of extreme events (floods)	Peak flow				
	Frequency of extreme events				
Retention of soil water	Runoff coefficient				
	Soil physical-water properties				
Continue…					

Continued…

Climate regulation	Heat attenuation				
	Evaporative fraction (Etr/Eto)				
Erosion Control					
Ecosystem Function	*Indicators*	*Clarity*	*Viability*	*Sensibility*	*Relevance*
Reduction of surface erosion	Estimated soil loss per area				
	Occurrence of soil erosion (number of erosion points per area)				
	Turbidity				
	Amount of sediment retained in barriers (physical/vegetative barriers, dams, terraces)				
	Sedimentation rate in the reservoir				
	Sediment flow in the canal				
Reduction of channel erosion	Canyon fault				
	Sediment flow in the canal				
	Geomorphic channel facility				
	Clogging				
Reduction of the sediment supply in the water body	Amount of sediment retained in barriers (physical/vegetative barriers, dams, terraces)				
	Sediment flow in channels				
	Turbidity				
	Sedimentation rate in the reservoir				
Soil Quality					
Ecosystem Function	*Indicators*	*Clarity*	*Viability*	*Sensibility*	*Relevance*
Provide physical support to plants	Resistance to penetration				
	Water Infiltration				

Continue…

Continued…

	Aggregates Stability
	Structure Degree
Control of soil loss	Degree of intensity of the erosive process
	Thickness A Horizon
	Diversity of species of the production system
	% of exposed soil
	Erosion per area
Water availability for plants	Soil organic matter (SOM)
	Aggregation
	Ratio of dead coverage
	Water available
	Thickness of the A + B horizon
Soil biodiversity regulation and biological activity	Presence of earthworms and spiders
	CO efflux
	Enzymatic activity
	Decomposition rate
	Macro and mesofauna diversity
	Diversity of nematodes
Nutrient cycling (Soil Fertility)	Soil organic matter (SOM)
	Diversity of species of the production system
	Cation Exchange Capacity (CEC)

Continue…

Continued…

	Biological fixation of N				
	Light organic matter				
	Decomposition rate				
	Amount of residual biomass				
	Nutrient Content				
	Export index				
	pH				
C sequestration	Carbon stock				
	Soil organic matter (SOM)				
Regulation of soil remediation potential	Soil organic matter (SOM)				
	Potential risk of contamination				
	Soil enzymatic activity				
	Waste from agrochemicals				
	Heavy metals (concern about the use of alternative inputs)				
	Aggregation				
Habitat Protection					
Ecosystem Function	*Indicators*	*Clarity*	*Viability*	*Sensibility*	*Relevance*
Conservation status of terrestrial habitats	Species threatened of extinction				
	Occurrence of invasive alien species				

Continue…

Continued…

	Habitat area
	Successional stage
	Diversity of species of native fauna and flora (richness and abundance of individuals by species)
	Species richness (number of species per area)
	Presence of key species (flora)
	Genetic diversity (genetic bench in situ)
	Litter (quality and quantity)
Conservation status of aquatic habitats	Composition of native fish communities and/or aquatic insects
	Abundance and Wealth of fish and/or aquatic insects
	Diversity of fish and/or aquatic insects
	Turbidity
	Presence of riparian forest
	Water temperature
	Biochemical Oxygen Demand (BOD)
	Presence of invasive species
	Presence of bioindicator species (fauna and flora)
	Endemism
	Ecological flow
Gene flow of plant, animal and microorganism species	Forest fragmentation index
	Proximity to protected areas

Continue…

Continued…

	Permeability of the matrix of uses
	Structural and functional connectivity (dependent species)
	Landscape metrics
	Presence of pollinating species
	Presence of dispersing species
	Presence of barriers in rivers
	Presence of future barriers in river projects
	Recruitment of species, presence of regenerating stratum
Biological control (for production)	Food production
	Demand for the use of agrochemicals
	Occurrence of natural enemies
Pollination (for food production)	Food production
	Quantity and quality of litter
	Occurrence of functional species (decomposers, pollinators, N-fixers, etc.)
	Nutrient cycling
Soil quality	Quantity and quality of litter
	Occurrence of functional species (decomposers, pollinators, N-fixers, etc.)
	Nutrient cycling
	Soil quality indicators (chemical, physical, biological and microbiological)
Gene Bank in situ	Diversity of species of fauna and flora

UNESCO (2020) highlights that it is urgent to improve water quality monitoring in order to effectively deal with the complexity of tracking a large number of parameters, including new, emerging pollutants. However, for effective action, water quality must be understood in the framework of hydrological processes based on the water quality and hydrological monitoring especially because most of the current monitoring processes are mainly based on ineffective traditional approaches and are jeopardized by a lack of scientific knowledge and technical skills. Thus, we believe that the ESI proposed to "water supply" and "water regulation" services in our study can help to fill this gap, as it is possible to use them in integrative water services monitoring.

The EFs associated with the "Soil Quality" service highlighted the diversity of functions that the soil provides for the ecosystem. Soil organic matter (SOM) and parameters related to soil aggregation were one of the ESI suggested to monitor different EF. This shows their relevance for soil quality and health, especially in tropical areas (Table 2).

Changes in land use or land management practices can influence soil properties such as organic matter content, aggregates, and density (USEPA, 2006). It indicates the potential of these properties to perform as indicators to monitoring PES interventions.

The "Habitat protection" presented a similar performance of "Soil Quality" ES, with a high diversity of EFs and ESIs (Table 2). Hatziiordanou *et al.* (2019) highlighted that a systematic approach to assess the habitat ecosystem service has not yet emerged. The same authors stated that, to evaluate this service it is important to observe the anthropogenic impact on biodiversity. The combination of these criteria could provide information about conservation measures. In our study, the EF and indicators selected by the experts to evaluate this service follow this rationality and present a set of solutions to contribute for the habitat service monitoring (Table 2).

3.2. ESI performance per criteria

One of the trickiest concerns in indicator research is to reach a final list with an ideal number of choices. If the quantity of indicators is simply too high, it defeats the aim altogether (Nathan and Reddy, 2010). Thus, the selection criteria adopted in this study aimed to find out the most suitable indicator package to monitor interventions in water PES programs.

3.2.1. Water supply

"Retention of nutrients in soil" was the EF with the highest number of ESI (Table 2). Most of these indicators are connected to Brazilian water legislation, such as the CONAMA 357/2005. This regulation provides the classification of water bodies and environmental guidelines, as well as the conditions and standards for the discharge of effluents and other measures. In this resolution some parameters are defined as standards to declare the water quality status. Some examples are: chemical oxygen demand, turbidity, pH, among others. These indicators comprise the main indexes of water quality for human supply.

Also, regarding the "Retention of nutrients in soil - relation with eutrophication and water contamination" function, dissolved oxygen, electric conductivity, presence of aquatic macrophytes, and total phosphorus were the indicators that scored higher. These parameters got the score "3" for all criteria (Table 2); index of biological integrity, diversity index, salmonella and virus were the indicators that scored the lowest.

This result follows the same rationale of Griffiths *et al.* (2018), which shows that biological indicators are better considered as a step in developing a practical monitoring scheme instead of a final indicator, as there are operational issues to be solved such as ease of application, robustness, sensitivity, laboratory accuracy, throughput, economic value, and descriptiveness.

The "Viability" criterion was the one in which the highest number of indicators (17) received the lowest score (Table 2). The reasons might include specificities about sampling, analysis and costs.

3.2.2. Water regulation

This ES had six associated EFs – “Maintenance of groundwater recharge”; “Maintenance of springs”; “Maintenance of reference flow”; “Attenuation of extreme events (floods)”; “Retention of soil water”; “Climate regulation”. All the EFs received few indicators for their monitoring (Table 2).

The “Flow rate” indicator got the best scores for all the criteria, although it was related to only one EF, “Maintenance of springs”. The indicator “Groundwater level” got a high score for two EFs: “Maintenance of groundwater recharge”, “Maintenance of reference flow” (Table 2).

The “Runoff coefficient” indicator was cited twice: for EF “maintenance of reference discharge” and “soil water retention”. The advantage of this indicator is that it integrates different aspects of hydrological processes. Runoff coefficient is the ratio between the total volume of water flowed by surface runoff and the total volume of precipitation. The Runoff Coefficient can be applied for one rainfall event or a period with several events. This kind of ratio integrates the entire watershed rainfall response, showing how much of the volume of water from precipitation the watershed can retain within the soil and how much will be released through runoff. Thus, it comprises the result of a set of flow processes conditioned by: (i) soil properties, such as water infiltration capacity in the soil, soil hydraulic conductivity, soil porosity; (ii) characteristics of terrain (slope, presence of microrelief structures, vegetation cover); geological characteristic of the basin; (iii) aspects of the rainfall (volume, duration and intensity of the events) (Merz and Günter, 2009).

The “water table level” was pointed out as an indicator of three functions: “aquifer recharge maintenance”, “the maintenance of river discharge” and “springs’ conservation”. This indicator requires the installation of sensors in wells for its measurement, and its metric integrates a set of processes that reflects the response of the watershed to rainfall regime.

The indicators with the lowest score (four indicators) were set in the “Sensitivity” criteria (Table 2). The hypothesis is that these indicators were considered broad, without an established parameter for measurement.

Erosion control

The indicator “Amount of sediment retained in barriers (physical/vegetative barriers, dams, terraces)” got the highest score for the four criteria, associated with different EFs (“Reduction of surface erosion” and “Reduction of the sediment supply in the water body”). However, it is important to be clear that the performance of this indicator can be influenced by the implementation of these barriers. Points with a high percentage of terraces (greater than 50%) were subjected to a slight degree of erosion since, in such cases, land is usually not cultivated or cultivation is carried out along the contour lines (Kosmas *et al.*, 2014). In this case, it is recommended to consider the other indicators that were also high scored, such as “Turbidity” and “Estimation of soil loss per area” (Table 2).

3.2.3. Soil quality

Soil quality, compared with the other ES, presented a more extensive indicator option for monitoring. This service has eight associated EFs and a set of associated indicators. The indicators that received the highest score – “3” for all criteria – were: Aggregates Stability; Resistance to penetration; Aggregation; Presence of earthworms and spiders and CEC. It is interesting to observe that all of these ESI are related to soil structure, directly or indirectly (Table 2).

Aggregation is considered one of the most suitable indicators for soil quality and crop production (Arshad and Martin, 2002). There is evidence about the close linkage between soil organic carbon (SOC) and aggregation (Martins *et al.*, 2009; Plaza-Bonilla *et al.*, 2013). Aggregate stability is a relevant indicator of soil susceptibility to runoff and erosion, especially in tropical areas where intense rainfall is frequent (Barthes and Roose, 2002). Additionally, Coq *et al.* (2007) demonstrated the influence of earthworms in large macroaggregates. Similar results are reported by Marichal *et al.* (2014) finding that the total macroinvertebrate density was significantly correlated with macro porosity and that these characteristics relate to the SOC content.

Soil organic matter (SOM), although a minor component in most soils, is primarily responsible for structure, function, and sustainability of the ecosystems (Turetta *et al.*, 2019). For this reason, the "organic matter content" was suggested as an indicator for the EFs "Water availability for plants"; "Nutrient cycling (Soil Fertility)"; "C sequestration" and "Regulation of soil remediation potential" highlighting the influence of this parameter in many soil processes.

3.2.4. Habitat protection

Most of the indicators that scored higher – "abundance and wealth of fish" and/or aquatic insects"; "Turbidity"; "Presence of riparian forest"; "biochemical oxygen demand" – are linked to the EF "Conservation status of aquatic habitats". This function is closely related to "Water supply" service and apparently, they followed the same rationale to suggest the most suitable indicators, such as "Turbidity" and "BOD of water" (Table 2).

It is interesting to observe a few suggested indicators, such as "Presence of invasive species" and "Presence of bioindicator species (fauna and flora)" for monitoring this service. The use of indicator species to monitor or assess environmental conditions is a firmly established tradition connected to environmental studies since the 1970s (Noss, 1990). In this case, the recommendation is to replace these indicators for others that can encompass multiple levels of biological organization.

Additionally, many parameters related to landscape metrics were suggested as indicators for this ES. Several studies indicate that such metrics are quite appropriate to describe the state of biodiversity (Walz, 2011). Still, the same author enhances that the results of such studies are strongly dependent on the scale of investigation and the underlying database. Furthermore, we highlight that the scale and database are aspects that should be carefully considered for all indicators.

"Pollination" and "Soil Quality" are classified as ES (Costanza *et al.* 1997; MEA, 2005). However, these services were suggested as indicators for monitoring this "Habitat protection" ES. It suggested a common conceptual confusion and distortion regarding ecosystem services concept, as already identified by Schoröter *et al.* (2014).

For "Habitat protection", four indicators received the lowest score: one in the "clarity" criterion; one in "relevance" and two in "viability". However, when we selected the ESIs with the highest score, the "sensitivity" criterion showed the lowest number of indicators compared to the other criteria (Table 2). An overall assumption is that for the experts, despite a higher number of options for monitoring this service, only few indicators are indeed sensitive to detect changes in this servisse.

3.3. Recommendations for the use of IES

For the best use of our findings, it is important to observe some aspects of the PES programs to be monitored, such as scale of the evaluation, thresholds, and others. These aspects must be assessed considering the database available, methods applied to identify the indicators and their application in public policies (Table 3).

Table 3. Aspects to be considered for the use of suggested IESs.

Attribute	Database	Method	Public Policy
Scale	✓		
Time differentiation	✓		
Quality and documentation	✓	✓	
Technological, scientifically grounded		✓	
Compatibility with international standards	✓	✓	✓
Implementation for modelling	✓		
Reference or threshold	✓	✓	✓

Source: Modified from Sieber (2019).

The database available appears a relevant aspect to be considered, whether for indicator surveys and defining references or for building scenarios (Table 3). Also, this is a core aspect for the baseline definition, which is usually a big problem for monitoring issues. The method is also an important aspect closely related to the analysis aspect and recognized standards.

Additionally, another suggestion to increase the practicality and adoption of our indicator matrix (Table 2) is to favor the indicators that received the highest score (Mode 3), since they were better evaluated by the experts and tend to be the most analytical and well established in the literature.

4. CONCLUSIONS

We presented a well-defined framework for transparent and operative selection of ecosystem services indicators, which can be used to monitor interventions on water-based PES.

In a broad field such as the ecosystem services assessment, the cascade model has proven to be useful to define the limits of the study, clarifying the flow of EF/intermediate services and how they can be impacted by PES interventions.

Four practical criteria (clarity, viability, sensitivity, and relevance) provided guidance to the experts to identify and select the best indicators for each function. The “viability” and “sensitivity” criteria received the lowest scores by the experts. In general, we could also observe in all services a preference for analytical and well-established indicators in the literature.

The methodological approach considering a pre-evaluation of intermediate services and indicators and its later discussion and validation by the experts proved to be an innovative approach for the ES indicators arena, promoting knowledge exchange in the scientific community. By doing so, we also expect to promote a more horizontal flow of information and decisions which can be helpful and easy to reproduce in other situations, improving the operationalization and governance of ecosystem services and connecting theory to practice.

The main recommendation of this study is to recognize there is no “ideal” indicator. Thus, it is recommended to consider a set of indicators for evaluation to cover possible gaps that a single indicator could present.

5. REFERENCES

ANA (Brasil). **Manual Operativo do Programa produtor de Água**. Brasília, 2012 84p.

ANA (Brasil). **Projetos**. 2021. Available at https://www.gov.br/ana/pt-br/acesso-a-informacao/acoes-e-programas/programa-produtor-de-agua/projetos Access: January 10th, 2022.

ARSHAD, M. A.; MARTIN, S. Identifying critical limits for soil quality indicators in agro-ecosystems. **Agriculture, Ecosystems and Environment**, v. 88, p. 153–160, 2002. https://doi.org/10.1016/S0167-8809(01)00252-3

BALMFORD, A.; RODRIGUES, A. S. J.; WALPOLE, M.; TEN BRINK, P.; KETTUNEN, M.; BRAAT, L. *et al.* **The economics of ecosystems and biodiversity**: scoping the science. ENV/070307/2007/486089/ETU/B2. Final report. Cambridge: European Commission, 2008.

BARTHES, B.; ROOSE, E. Aggregate stability as an indicator of soil susceptibility to runoff and erosion; validation at several levels. **Catena**, v. 47, n. 2, p. 133-149, 2002. https://doi.org/10.1016/S0341-8162(01)00180-1

BERGHÖFER, A.; SCHNEIDER, A. **Indicators for Managing Ecosystem Services – Options & Examples**. Leipzig; Eschborn: Helmholtz Zentrum für Umweltforschung (UFZ); Deutsche Gesellschaft für Internationale Zusammenarbeit, 2015. 49p.

BETTER EVALUATION. **Stakeholder mapping and analysis**. 2014. Available on www.betterevaluation.org/en/evaluation-options/mapping_stakeholders. Access March 1st, 2020.

BRUGHA, R.; VARVASOVSKY, Z. Stakeholder analysis: a review. **Health Policy and Planning,** v. 15, p. 239-246, 2000. https://doi.org/10.1093/heapol/15.3.239

COQ, S.; BARTHES, B. G.; OLIVER, R.; RABARY, B.; BLANCHART, E. Earthworm activity affects soil aggregation and organic matter dynamics according to the quality and localization of crop residues—An experimental study (Madagascar). **Soil Biology and Biochemistry**, v. 39, n. 8, p. 2119-2128, 2007. https://doi.org/10.1016/j.soilbio.2007.03.019

COSTANZA, R.; D'ARGE, R.; DE GROOT, R.; FARBERK, S.; GRASSO, M.; HANNON, B. *et al.* 1997. The value of the world's ecosystems services and natural capital. **Nature**, v. 387, p. 253-260, 1997. https://doi.org/10.1038/387253a0

COSTANZA, R.; DE GROOT, R.; BRAAT, L.; KUBISZEWSKI, I.; FIORAMONTI, L.; SUTTON, P. *et al.* Twenty years of ecosystem services: How far have we come and how far do we still need to go? **Ecosystem Services**, v. 28, p. 1-16, 2017. http://dx.doi.org/10.1016/j.ecoser.2017.09.008

DE GROOT, R. S. **Functions of Nature Evaluation of Nature in Environmental Planning, Management and Decision Making**. Groningen: Wolters-Noordhoff, 1992. Available on https://edepot.wur.nl/211708. Access 26th February 2020.

DE GROOT, R. S.; WILSON, M.; BOUMANS, R. A Typology for the Classification Description and Valuation of Ecosystem Functions, Goods and Services. **Ecological Economics**, v. 41, n. 3, 2002. https://doi.org/10.1016/S0921-8009(02)00089-7

DE GROOT, R. S.; ALKEMADE, R.; BRAAT, L.; HEIN, L.; WILLEMEN, L. Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. Ecological Complexity, v. 7, p. 260–272, 2010a. https://doi.org/10.1016/j.ecocom.2009.10.006

DE GROOT, R. S.; FISHER, B.; CHRISTIE, M.; ARONSON, J.; BRAAT, L.R.; HAINES-YOUNG, J. *et al.* Integrating the ecological and economic dimensions in biodiversity and ecosystem service valuation. *In:* KUMAR, P. (ed.) **The Economics of Ecosystems and Biodiversity**: Ecological and Economic Foundations. London: Earthscan, 2010b. https://doi.org/10.4324/9781849775489

ENGEL, S. S.; PAGIOLA, S.; WUNDER, S. Designing payments for environmental services in theory and practice: An overview of the issues. **Ecological Economics**, v. 65, n. 4, p. 663-674, 2008. https://doi.org/10.1016/j.ecolecon.2008.03.011

FAZEY, I.; EVELY, A. C.; REED, M. S.; STRINGER, L. C.; KRUIJSEN, J.; WHITE, P. C. L. *et al.* Knowledge exchange: a review and research agenda for environmental management. **Environmental Conservation,** v. 40, p. 19-36, 2013. http://dx.doi.org/10.1017/S037689291200029X

FISHER, B.; TURNER, R. K.; MORLING, P. Defining and classifying ecosystem services for decision making. **Ecological Economics**, v. 68, p. 643-653, 2009. https://doi.org/10.1016/j.ecolecon.2008.09.014

FLEISCHMAN, F.; BRISKE, D. D. Professional ecological knowledge: an unrecognized knowledge domain within natural resource management. **Ecology and Society**, v. 21, n. 1, p. 32, 2016. http://dx.doi.org/10.5751/ES-08274-210132

GRIFFITHS, B. F.; FABER, J. H.; BLOEM, J. Applying Soil Health Indicators to Encourage Sustainable Soil Use: The Transition from Scientific Study to Practical Application. **Sustainability**, v. 10, n. 9, p. 3021, 2018 https://doi.org/10.3390/su10093021

GUIJT, I. **Participatory Approaches**. Methodological Briefs – Impact Evaluation No. 5. Florence: UNICEF Office of Research, 2014. https://doi.org/10.13140/RG.2.1.4948.1768

HAINES-YOUNG, R.; POTSCHIN, M. The links between biodiversity, ecosystem services and human well-being. In: RAFFAELLI, D.; FRID, C. F. (Eds.). **Ecosystem Ecology**: a new synthesis. BES Ecological Reviews Series, CUP. Cambridge, 2010.

HAINES-YOUNG, R.; POTSCHIN, M. Revision of the Common International Classification for Ecosystem Services (CICES V5.1): A Policy Brief. **One Ecosystem**, v. 3, n. e27108, 2018. https://dx.doi.org/10.3897/oneeco.3.e27108

HATZIIORDANOU, L.; FITOKA, E.; HADJICHARALAMPOUS, N.; VOTSI, N.E.; PALASKAS, D.; MALAK, D. A. Indicators for mapping and assessment of ecosystem condition and of the ecosystem service habitat maintenance in support of the EU Biodiversity Strategy to 2020. **One Ecosystem**, v. 4, n. e32704, 2019. https://dx.doi.org10.3897/oneeco.4.e32704

KOSMAS, C.; KAIRIS, O. R.; KARAVITIS, C. H.; RITSEMA, C.; SALVATI, L. *et al.* Evaluation and selection of indicators for land degradation and desertification monitoring: types of degradation, causes, and implications for management. **Environmental Management**, v. 54, n. 5, p. 971-982, 2014.

LIMA, A. P. M.; PRADO, R. B.; LATAWIEC, A. E. Payment for water-ecosystem services monitoring in Brazil. **Revista Ambiente & Água**, v. 14, n. 4, 2021. https://doi.org/10.4136/ambi-agua.2684

MARICHAL, R.; GRIMALDI, M.; FEIJOO, A.; OSZWALD, J.; PRAXEDES, C.; COBO, D. H. R. *et al.* Soil macroinvertebrate communities and ecosystem services in deforested landscapes of Amazonia. **Applied Soil Ecology**, v. 83, p. 177-185, 2014. https://doi.org/10.1016/j.apsoil.2014.05.006

MARTÍN-LÓPEZ, B.; GÓMEZ-BAGGETHUN, E.; GARCÍA-LLORENTE, M.; MONTES, C. Trade-offs across value-domains in ecosystem services assessment. **Ecological Indicators**, v. 37, p. 220–228, 2014. http://dx.doi.org/10.1016/j.ecolind.2013.03.003

MARTINS, M. R.; CORÁ, J. R.; JORGE, R. F.; MARCELO, A. V. Crop type influences soil aggregation and organic matter under no-tillage. **Soil and Tillage Research**, v. 104, p. 22-29, 2009. https://doi.org/10.1016/j.still.2008.11.003

MERZ, R.; GÜNTER, B. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. **Water Resources Research**, v. 45, n. W01405, 2009. https://doi.org/10.1029/2008WR007163

MILLENIUM ECOSYTEM ASSESSMENT. **Millennium Ecosystem Assessment**. Washington, DC.: World Resources Institute, 2005.

NATHAN, H. S. K.; REDDY, B. S. **Selection Criteria for Sustainable Development Indicators**. Mumbai: Indira Gandhi Institute of Development Research, 2010. Available on https://core.ac.uk/download/pdf/6256597.pdf Access on December 7th 2020.

NOSS, R. Indicators for Monitoring Biodiversity: A Hierarchical Approach. **Conservation Biology**, v. 4, n. 4, 1990. https://doi.org/10.1111/j.1523-1739.1990.tb00309.x

OECD. **Future Global Shocks – Improving Risk Governance**. Paris, 2011a.

OECD. **Compendium of OECD well-being indicators**. Paris, 2011b.

PAIN, R. Social geography: participatory research. **Progress in Human Geography**, v. 28, n. 5, 2004. https://doi.org/10.1191/0309132504ph511pr

PAGIOLA, S.; PLATAIS, G. **Payments for Environmental Services**: From Theory to Practice. Washington: World Bank, 2007.

PAGIOLA, S.;VON GLEHN, H. C.; TAFFARELLO, D. **Experiências de pagamentos por serviços ambientais no Brasil**. São Paulo: Secretaria do Meio Ambiente, 2012. 336 p.

PAVAN, A. L. R.; OMETTO, A. R. Ecosystem Services in Life Cycle Assessment: A novel conceptual framework for soil. Science of the Total Environment, v. 1, n. 643, p. 1337-1347, 2018. https://doi.org/10.1016/j.scitotenv.2018.06.191

PLAZA-BONILLA, D.; CANTERO-MARTINEZ, C.; VIÑAS, P.; ÁLVARO-FUENTES, J. Soil aggregation and organic carbon protection in a no-tillage chronosequence under Mediterranean conditions. **Geoderma**, v. 193-194, p. 76-82, 2013. https://doi.org/10.1016/j.geoderma.2012.10.022

POCIDÔNIO, E. A. L.; TURETTA, A. P. D. **Programas de Pagamento por Serviços Ambientais no Brasil**. Dados eletrônicos. Embrapa Solos, 2012. 25 p. http://www.cnps.embrapa.br/publicacoes/ .

POTSCHIN, M. P.; HAINES-YOUNG, R. H. Ecosystem services: Exploring a geographical perspective. **Progress in Physical Geography**, v. 35, p. 575, 2011. https://doi.org.10.1177/0309133311423172

PRIMMER, E.; FURMAN, E. Operationalizing ecosystem service approaches for governance: Do measuring, mapping and valuing integrate sector-specific knowledge systems? **Ecosystem Services**, v. 1, p. 85–92, 2012. https://doi.org/10.1016/j.ecoser.2012.07.008

QUEENSLAND GOVERNMENT. **Ecosystem health indicators**. 2020. Available on https://environment.des.qld.gov.au/management/water/health-indicators#physico_chemical_indicators Access 14 Dec. 2020.

RAYMOND, C. M.; FAZEY, I.; REED, M. S.; STRINGER, L. C.; ROBINSON, G. M.; EVELY, A. C. Integrating local and scientific knowledge for environmental management. **Journal of Environmental Management**, v. 91, p. 1766-1777, 2010. http://dx.doi.org/10.1016/j.jenvman.2010.03.023

REED, M. S.; GRAVES, A.; DANDY, N.; POSTHUMUS, H.; HUBACEK, K.; MORRIS, J. *et al.* 2009. Who's in and why? A typology of stakeholder analysis methods for natural resource management. **Journal of Environmental Management**, v. 90, p. 1933-1949, 2009. https://doi:10.1016/j.jenvman.2009.01.001

ROUX, D. J.; ROGERS, K. H.; BIGGS, H. C.; ASHTON, P. J.; SERGEANT, A. 2006. Bridging the science–management divide: moving from unidirectional knowledge transfer to knowledge interfacing and sharing. **Ecology and Society**, v. 11, n. 1, p. 4, 2006.

SALZMAN, J.; BENNETT, G.; CARROLL, N.; GOLDSTEIN, A.; JENKINS, M. The Global Status and Trends of Payments for Ecosystem Services. **Nature Sustainability**, v. 1, n. 3, p. 136-144, 2018. https://doi.org/10.1038/s41893-018-0033-0

SCHRÖTER, M.; VAN DER ZANDEN, E. H.; VAN OUDENHOVEN, A. P. E.; REMME, R. P.; SERNA-CHAVEZ, H. M.; DE GROOT, R. S. *et al.* Ecosystem Services as a Contested Concept: A Synthesis of Critique and Counter-Arguments. **Conservation Letters**, v. 7, n. 6, p. 514–52, 2014. https://doi.org/10.1111/conl.12091

SIEBER, S. **Tailoring SI to farm environment**. Oral communication. Berlin: Humbold University, 2019.

SPANGENBERG, J. H.; VON HAAREN, C.; SETTELE, J. The ecosystem service cascade: Further developing the metaphor. Integrating societal processes to accommodate social processes and planning, and the case of bioenergy. Ecological Economics, v. 104, p. 22-32, 2014. https://doi.org/10.1016/j.ecolecon.2014.04.025

TALLIS, H.; KAREIVA, P.; MARVIER. M.; CHANG, A. An ecosystem services framework to support both practical conservation and economic development. **Proceedings of the National Academy of Science of the United States of America**, v. 105, n. 28, p. 9457–9464, 2008. https://doi.org/10.1073/pnas.0705797105

TEEB. **Ecological and Economic Foundations.** Abingdon: Routledge, 2010. 410 p.

TURETTA, A. P. D.; PRADO, R. B.; COUTINHO, H. L. DA C.; FIDALGO, E. C. C.; SCHULER, A. E. *et al.* **Memória da Oficina Ranqueamento de Indicadores de Serviços Ambientais**. Rio de Janeiro: Embrapa Solos, 2013. 67 p.

TURETTA, A. P. D.; NOVOTNY, E. H.; REBELLO, C. M. Soil organic matter quality as indicator of ecosystem services. *In*: WORLD CONGRESS OF SOIL SCIENCE, 21., 2018, Rio de Janeiro. **Soil science: beyond food and fuel: proceedings[...]** Viçosa, MG: SBCS, 2019. v. 2, p. 474.

UNESCO. **Monitoring water quality and wastewater**. 2020. Available on https://en.unesco.org/waterquality-IIWQ/activities-projects/monitoring-water-quality-wastewater Access on 14 December 2020.

USEPA. **The development and use of soil quality indicators for assessing the role of soil in environmental interactions**. Science Report SC030265. Washington, 2006. 248p. Available www.environment-agency.gov.uk. Access 14 December 2020.

WALZ, U. Landscape Structure, Landscape Metrics and Biodiversity. **Living Reviews in Landscape Research**, v. 5, 2011.

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Effect of lithological and geotechnical characteristics on the generation of debris flows in the arid basin of Mirave, Peru

ARTICLES doi:10.4136/ambi-agua.2785

Received: 28 Aug. 2021; Accepted: 09 Mar. 2022

Julio César Bizarreta Ortega[1]; Marko Antonio Lopez Bendezu[2]; Alexandre Almeida Del Savio[2*]; Fausto Alfredo Canales[3]

[1]Instituto Latino-Americano de Tecnologia, Infraestrutura e Território. Universidade Federal da Integração Latino-Americana, Avenida Silvio Américo Sasdelli, n° 1842, CEP: 85866-000, Foz do Iguaçu, PR, Brazil. E-mail: jcbizarreta@gmail.com
[2]Instituto de Investigación Científica. Department of Civil Engineering. Universidad de Lima, Avenida Javier Prado Este, n° 4600, 15023, Santiago de Surco, Lima, Peru. E-mail: mlopezb@ulima.edu.pe
[3]Department of Civil and Environmental. Universidad de la Costa, Calle 58#55-66, 080002, Barranquilla, Atlántico, Colombia. E-mail: fcanales@cuc.edu.co
[*]Corresponding author. E-mail: delsavio@gmail.com

ABSTRACT

Studies concerning the debris flows in mountain areas are relevant because of their potential negative effects on the human communities and infrastructure in their areas of influence. To advance the understanding of the theoretical basis, this study qualitatively analyzes the effect of lithological characteristics and soil type on the generation of debris flows in the arid basin of Mirave, in southern Peru, as a consequence of extensive rainfall. Two debris flow events are evaluated, which occurred in the studied area in March 2015 and February 2019. The method used to achieve the objective combines the use of satellite images, field data collection, and lab tests results to estimate the relative importance of the abovementioned characteristics in the generation of debris flows. The results suggest that the poor presence of clay and the predominance of sandy-loamy structured soils in the Mirave Basin make them unstable when erosion occurs. In addition, the features of broken down materials generated from residual and colluvial soils in the primary area of study are one of the main causes of debris flows in the region.

Keywords: arid basin, debris flows, geotechnical characteristics, lithology.

Efeito das características litológicas e geotécnicas na geração de fluxos de detritos na bacia árida de Mirave, Peru

RESUMO

A relevância dos estudos sobre os fluxos de detritos em áreas montanhosas se deve aos seus efeitos negativos potenciais sobre as comunidades humanas e a infraestrutura em suas áreas de influência. Visando contribuir para o avanço no entendimento do embasamento teórico, o presente estudo analisa qualitativamente o efeito das características litológicas e do tipo de solo na geração de fluxos de entulho na bacia árida de Mirave, no sul do Peru, em decorrência de chuvas intensas. São avaliados dois eventos de fluxo de detritos, que ocorreram na área estudada em março de 2015 e fevereiro de 2019. O método utilizado para atingir o objetivo combina o uso de imagens de satélite, coleta de dados de campo e resultados de testes de laboratório para estimar a importância relativa do características supracitadas na geração de fluxos de detritos. Os resultados sugerem que a fraca presença de argila e a predominância de solos estruturados arenoso-argiloso na bacia do Mirave os tornam instáveis quando ocorre a erosão. Além disso, as características dos materiais degradados gerados a partir de solos residuais e coluviais na área primária de estudo são uma das principais causas dos fluxos de detritos na região.

Palavras-chave: bacia árida, características geotécnicas, fluxos de detritos, litologia.

1. INTRODUCTION

Based on their magnitude and frequency, debris landslides have been studied around the world in relation to the potential negative impacts on the human communities in their area of influence. Debris flows and rockfalls are some of the main geomorphological processes in the Andes, particularly in the arid Andean zones (Moreiras, 2006). When such flows get activated in basins, they follow the path of natural drainage; those that run along natural channels (Chen *et al.*, 2009) are referred to as channeled debris flow. Usually, an initiating or generating zone, a flow transportation or conducting zone, and a deposition zone are present in a debris flow event (Imaizumi *et al.*, 2016). The initiation zones of debris flow due to extensive precipitations are mainly affected by availability, type, and classification of soil, lithology, current strength, the slope, soil usage and cover, and the density of previous events (Esper Angillieri *et al.*, 2020; Meyer *et al.*, 2014). Lithology studies and soils are fundamental for understanding the types of mechanisms linked to an event, as well as to evaluate the zones that are prone to the initiation of debris flow. This is because the types of rocks and soil play a key role in the formation and buildup of surface unconsolidated sediments, which induce debris flows (Di *et al.*, 2019). In general, the literature states that the lithology and the types of soils are critical in the occurrence of this phenomenon (Cannon and Reneau, 2000; Esper Angillieri, 2020; Ding *et al.*, 2020). The debris flow is mainly generated in steep slope zones ranging from 20° to 45°, initiated by surface breakdowns of slopes in the heads or hillside slopes of gullies, or sometimes in the bed of steep drainage channels, during events of extreme runoff (Hungr, 2005).

This research work is focused on the arid basin located in the town of Mirave, in the Tacna region of the Sechura–Atacama Desert. This desert rests on the western side of the Andes Range and the Pacific Ocean. This area has a lot of arid basins along which the dry rivers run; these are local rivers exhibiting the debris flow, and these rivers may remain dry for long periods of time, normally measured in years. When such basins become active, they are generally manifested as an event of debris flow. The mechanisms and the location of areas prone to the generation of debris flow in these regions are barely studied, as well as the effect of lithological aspects and soil characteristics.

Debris flows are expected in many Peruvian locations due to steep slopes, high mountains, extremely arid western foothills with rocks and soils susceptible to removal by rainwater. However, debris flows have only been reported in Mirave for 2015 and 2019, and according to the elderly residents, in 1927. The paper describes the lithological and geotechnical characteristics of the Mirave Basin and complements the work by Del Savio *et al.* (2019) by discussing the effect of these characteristics, thus improving the understanding of the debris flow generation in this region.

By using the finite differences method provided by the commercial software FLO-2D, Del Savio *et al.* (2019) simulated the debris flow that occurred in 2019 in the same region. This software applies a numerical integration of flow motion and conservation equations to simulate the debris movement of a 2D scenario (Chang *et al.*, 2020). The numerical simulation does not entirely reproduce the physical phenomenon but uses the liquid hydrographs of flow for different periods of time to determine the affected area and the zones of debris deposition. Del Savio *et al.* (2019) provide the parameters to numerically simulate a model in FLO-2D, including the debris flow rates, hydrographs, and volume of the debris flow in the four micro-ravines of the Mirave Basin; and the rheological parameters of the debris flow. However, one of its limitations is that it does not consider the characteristics of surface lithology of the micro basin. In this context, this research aims to qualitatively evaluate the effect of lithological characteristics and soil types in the generation of debris flow in the Mirave Basin, which is part of the Sechura–Atacama Desert, in southern Peru.

By focusing on the Mirave Basin area, this research contributes to the literature related to debris flows in hyperarid environments. In South America, similar case studies on debris flow in the Andean region include the paper by Melendez *et al.* (2021), who applied a TETIS hydrological simulation model on the Lurin River Basin on the arid coast of Peru. The information for the hydrological model calibration was available; however, geotechnical and lithological data was not included in the modeling, missing some accuracy related to the desertic conditions of the region. Lara *et al.* (2018) carried out a geomorphological study to determine the characteristics of the arid basin of the La Ciénaga River in Argentina. They found that the basin is in an active tectonic environment and that all the sub-basins analyzed share an elongated shape, which allows a rapid concentration of water that intensifies the power of flash floods, very similar to the Mirave Sub-basins. Besides these examples, the review paper by Moreiras *et al.* (2021) compiled research related to this type of event in the semiarid Central Andean region, highlighting the importance of improving the knowledge on the subject and implementing preventive measures to protect the Andean communities.

2. METHODS

2.1. Location and Characterization of the Study Area

Figure 1 shows the location of the Tacna region in the south of Peru and the red lines outline the Mirave Basin, which belongs to the Ilabaya District. The town of Mirave lies at an average altitude of 1100 masl, at the micro basin mouth of the Mirave ravine (Medina Allca and Luque Poma, 2016), at the UTM coordinate: 8066400 North; 334400 East; datum WGS 84, and Zone 19S. The inhabitants of the town (approximately 800 people) did not want to be relocated despite knowing that they inhabit a high-risk geological area due to mass movement (Ilabaya, 2016).

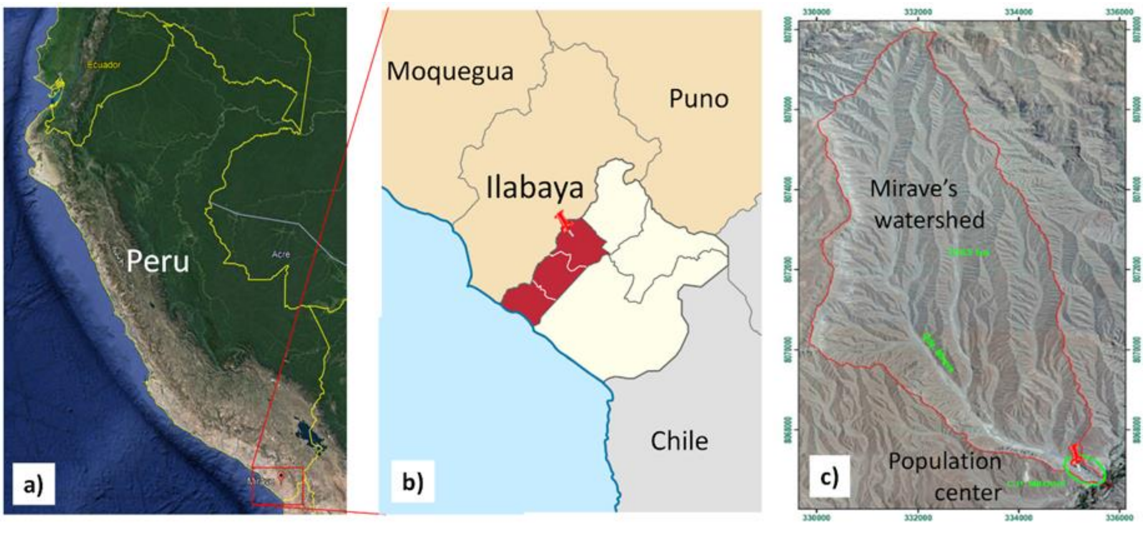


Figure 1. a) Location of the Tacna region in Peru; b) Map of Tacna where the District of Ilabaya is situated; c) Basin and town of Mirave.

The sub-basin is characterized by the erosion of sediments known as the Moquegua formation, corresponding to the period starting during the Eocene (~50 mya) up to the Pliocene (~4 mya) (Decou *et al.*, 2013), and a small part of the Toquepala group from the Early Cretaceous, called the Paralanque formation (Bellido Bravo, 1979).

Guerrero *et al.* (2013) classified the climates of the Sechura–Atacama Desert as hyperarid, arid, and semiarid climates. The debris flows occurring in these regions generally run on natural channels that remain dry most of the year. The nearest meteorological station is the Ilabaya station, situated 5 km from the high sub-basin of Mirave, which lies at an altitude of 1375 masl. According to the Servicio Nacional de Meteorología e Hidrología del Peru (SENAMHI, 2021) the rainfalls only occur in the summer months (December, January, February, and March). Taking some recent years as reference, the climate classification applied by Guerrero *et al.* (2013) would characterize the year 2016 as a hyper arid climate year (i.e., with a precipitation $\leq$ 5 mm/year), the years 2017 and 2018 as arid climate years (5 $<$ precipitation $\leq$ 50 mm/year), and the years 2015 and 2019 as semi arid climate years (50 $<$ annual precipitation $<$ 250 mm/year). The latter ones are particularly interesting for our study because they coincide with the occurrence of debris flows. The sudden rains in arid areas such as Mirave may range from moderate to heavy, which may trigger debris flows and floods, according to the findings of Stolle *et al.* (2015) and Houston (2006) for arid zones. Considering the effect of other extreme climate events, the 2015 event in Mirave was not associated with the El Niño phenomenon, whereas the 2019 event coincided with the coastal El Niño phenomenon.

The occurrence of debris flows in the Mirave area is associated with exceptional rainfall, and they are classified into old and recent debris flows. The origin of the deposits corresponds to ancient debris and mudflows from the sedimentary outcrops of the Lower Moquegua Formation. Mostly during January, atmospheric patterns related to the Southern Hemisphere's summer season modulate the spatial distribution of rainfall, air temperatures, humidity, and other variables on the South American continent (SENAMHI, 2021). In this season, the Intertropical Convergence Zone moves to the south near the equator, the presence of the South Atlantic Convergence Zone and an anticyclonic circulation pattern of high-altitude wind flows (12 km) known as the High from Bolivia, is activated. These systems are the main drivers of convective rainfall in large parts of Peru and facilitate the development of storms (Medina Allca and Luque Poma, 2016).

The Ilabaya station recorded an 8-mm precipitation on March 26, 2015, a date when the debris flow occurred in Mirave. On previous days, from March 22, 2015, total daily rainfalls of 1, 8, 2, and 4 mm were recorded. According to the Instituto Nacional de Defensa Civil (INDECI, 2019) and the SENAMHI (2021), this event destroyed 60 houses and the street pavement, but no deaths were reported. Notably, retention dams were built before the event along the stream's course. These works reduced the effects of the debris flows, as described by Del Savio *et al.* (2019). Figure 2 shows the location of the dams built in the studied micro basin.

The event of February 8, 2019 had a greater magnitude than that of 2015 and was triggered by a rainfall of approximately 35 mm in the upper Mirave Basin, as recorded at the Ilabaya station. This debris flow occurred two days after a minor rainfall (SENAMHI, 2021). According to the data from government agencies (INDECI, 2019), this event resulted in two deaths, affecting 905 people and destroying 76 houses, among other damages.

From the field study conducted after the events, the occurrence of these disasters can be explained mainly because the town center is located in the alluvial range of the sub-basin, i.e., the natural drainage area of the debris that flows into the river is occupied by houses and other structures. The main identified factors that favored the debris flow in Mirave are (i) slopes with current erosion processes (gullies) that feed material into the stream channel; (ii) slopes that range from 25° to 70° that facilitate erosion of soil cover and rock; iii) unconsolidated conglomerate rock with a sandy-loamy matrix that is easily eroded by rainfall, accelerating the formation of loose material and contributing to the sediment load toward the gully; and iv) the absence of vegetation cover, which allows for a rapid acceleration of slope erosion.

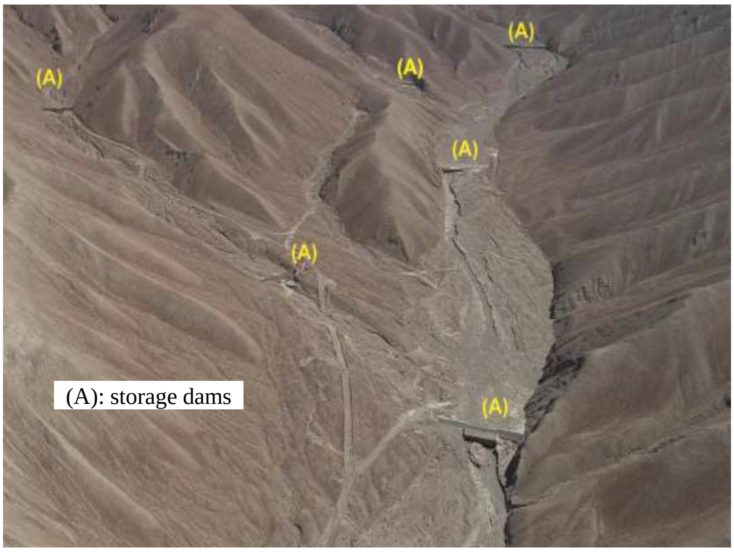


Figure 2. Location of sediment storage dams.

The retaining dams or energy dissipating walls were filled by the materials transported by the debris flows. While passing through the urban area, they first moved through the stadium (Figure 3a) having 10-m high outer walls, which were completely buried, as shown in Figure 3b. The mass of the flow destroyed part of the walls and reached the stands and sport fields. Throughout its entire course, the flow destroyed the houses in its path.

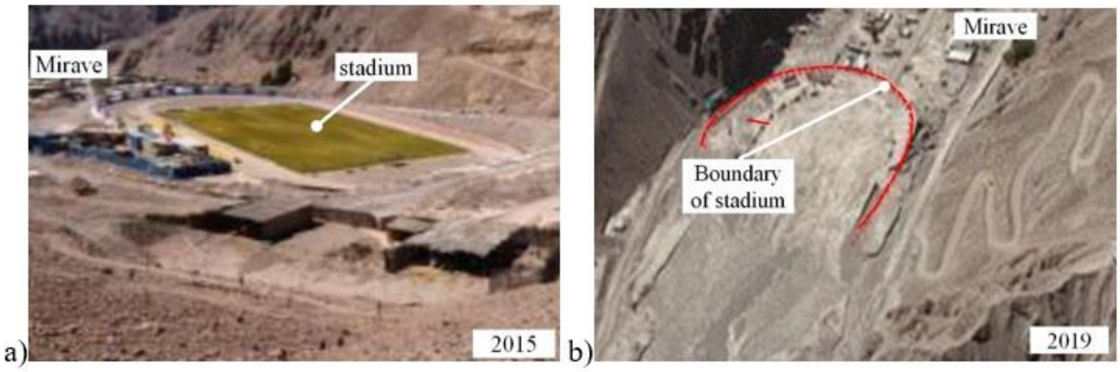


Figure 3. A debris flow occurred on a) March 26th, 2015, and b) February 8th, 2019.

2.2. Lithology and Soils in Mirave's Sub-Basin

The geological analysis of the study area was developed based on the geological map of the Moquegua quadrangle (Martinez and Zuloaga, 2000), the interpretation of aerial photographs (Del Savio *et al.*, 2019), Google Earth satellite images, and their own in-situ observations.

Observations were made throughout the Mirave drainage basin, mainly in the upper basin and sites near the watershed, to qualitatively determine the relative importance of lithology on the onset of debris flows due to rainfall. These observation rounds included twelve soil sample collection points and covered the type and quality of rock masses, soils resulting from weathering and transport, vegetation, slope, and the presence of erosional features. Google Earth satellite images were used in inaccessible or unexplored areas. Based on the aforementioned data, the lithological map was prepared, as shown in Figure 4, which in addition to the lithological units, shows the recent alluvial and storm-water deposits resulting from debris flow, located from the drainage areas to the alluvial fan that reaches the Salado River.

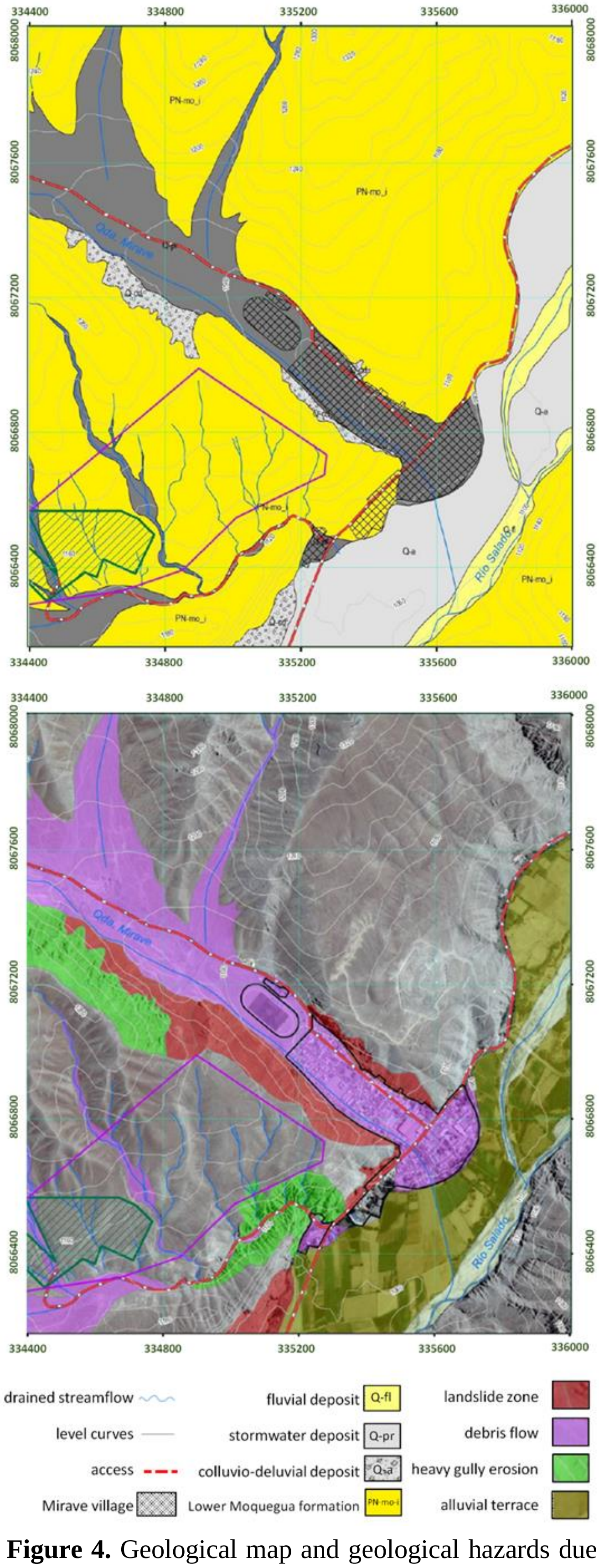
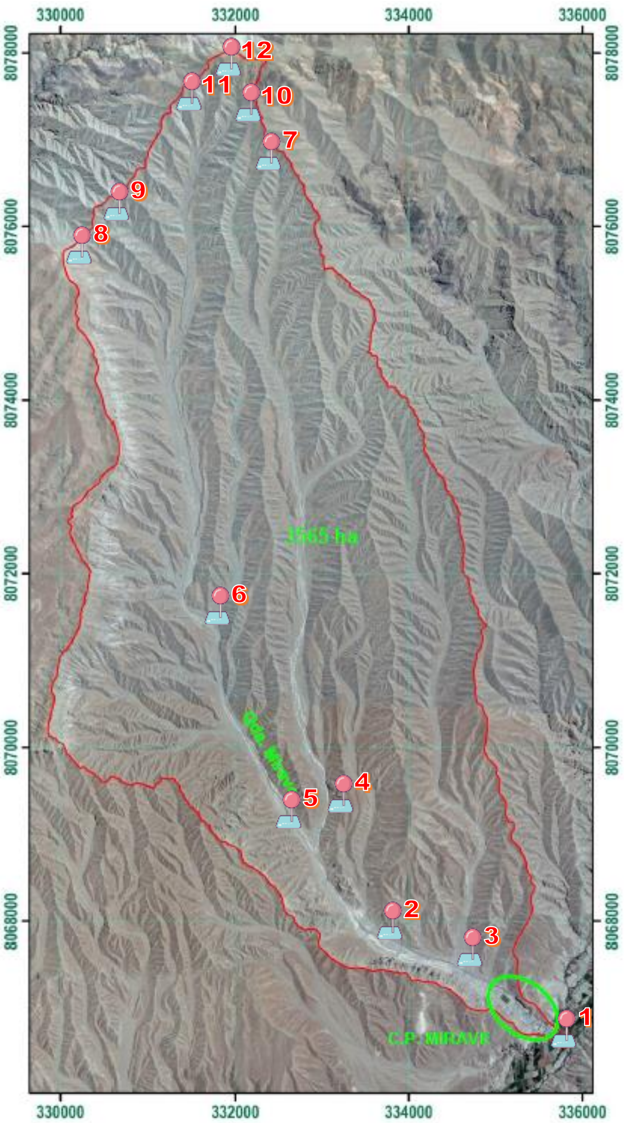


Figure 4. Geological map and geological hazards due to mass movements of Mirave.

The twelve soil sample collection points are shown in Figure 5, along with the location of the town of Mirave in the alluvial fan formed at the mouth of the basin. The geotechnical properties of the soils were evaluated by conducting field and laboratory tests. Altered samples were collected at a depth of 30 cm for moisture tests (ASTM D2216-19, 2020a), Atterberg limits (ASTM D4318-17e1, 2020d), specific gravity tests (ASTM D5550-14, 2020e), grain-size tests (ASTM D6913M-17, 2020f), soil classification (ASTM D2487-17e1, 2020b), and direct shear under drained conditions (ASTM D3080/D3080M-11, 2020c). The direct shear test was performed using particles smaller than 4.8 mm in diameter and with the average density of the soil.



Sampling points:

1	335707.89 m E	8066642.09 m S
2	333850.35 m E	8067677.79 m S
3	334639.54 m E	8067520.66 m S
4	332845.92 m E	8069951.51 m S
5	332273.36 m E	8069779.99 m S
6	331775.86 m E	8071585.78 m S
7	332414.79 m E	8076782.27 m S
8	330062.45 m E	8075712.58 m S
9	330333.20 m E	8075921.82 m S
10	332185.99 m E	8077410.54 m S
11	331586.27 m E	8077599.94 m S
12	332024.01 m E	8077853.18 m S

Figure 5. Delimitation map of the Mirave Micro-basin and sampling points.

3. RESULTS AND DISCUSSION

3.1. Lithological Characteristics

In the area corresponding to the Mirave Sub-basin, four areas with different lithology were observed, corresponding to Paleogene–Neogene sedimentary rocks and Quaternary deposits. These include conglomerates of the upper Moquegua formation, conglomerates of the lower Moquegua formation, andesites of the Paralanque formation, and recent alluvial deposits (sands and gravels). Matrix-supported conglomerates (from the lower Moquegua formation) have polymictic clasts with subrounded to subangular characteristics and poorly sorted features, with a maximum particle size of 5 cm. These materials are present in the lower basin forming escarpments, especially at the outflow of the drainage. In the middle and upper part, these materials form steeply sloping reliefs with a noticeable presence of colluvial and residual materials. A large number of gullies are concentrated in the primary area, which indicates the erodible characteristics of the materials. Figure 6a shows this type of material located at the headwaters. These materials contain fine whitish particles very similar to volcanic tuffs in some parts of the matrix.

The second lithological area contains clast-supported conglomerates (from the upper Moquegua formation), and it is located in escarpments on the left bank and upper Mirave Basin. These materials are located in the watershed on the right bank of the basin, covering an area of approximately 6 km. They are characterized by their lithic aspect with the presence of poorly sorted subrounded polymictic clasts with an average size of 30 cm; however, they can be as large as 80 cm. In Figure 6b, part of a steep side can be observed with some of these materials. The formation of loose clasts on the slopes is a typical feature. The gullies in these areas present an abrupt initiation followed by a smooth one, with accumulation of larger materials in the natural channels.

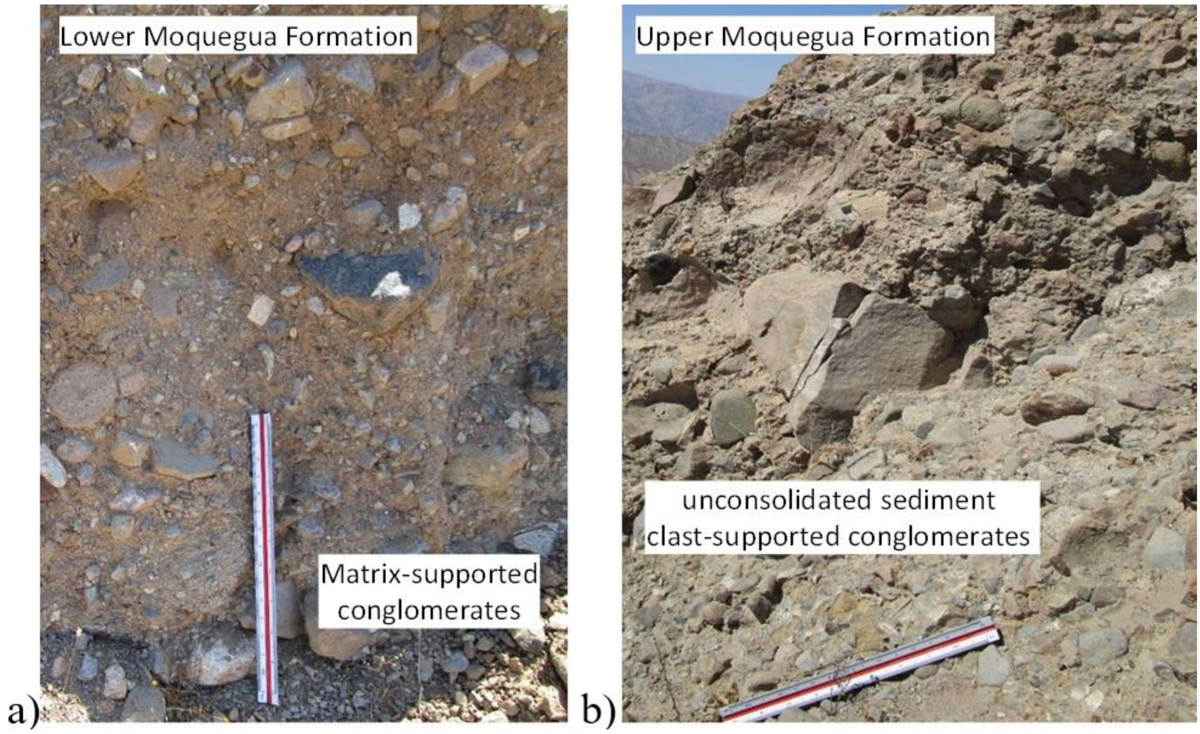


Figure 6. a) Matrix-supported conglomerates and b) clast-supported conglomerates.

The third lithological area contains andesites, which emerge in the left head of the Mirave basin, where the highest point of the basin is located at 2404 masl. The hillsides of these rock formations have an average slope of 30°. Fractured rock massifs and thin layers of soil indicate the action of weathering (Figure 7a). The soils of the hillsides exhibit angular and subangular rock fragments of different sizes (blocks, boulders, cobbles, and gravels). The action of water in this area has resulted in the formation of some small streams that transport the materials contained in the riverbed to the lowest parts. In the 2019 catastrophic event, angular blocks were observed in the Mirave dejection cone, which indicate that they originated in this region. The last lithological area is the alluvial deposits, formed by different debris-flow sediments. Figure 7b shows these deposits that are located in the upper zones. These are characterized by being unconsolidated, and their particles are frequently removed when debris flows are activated.

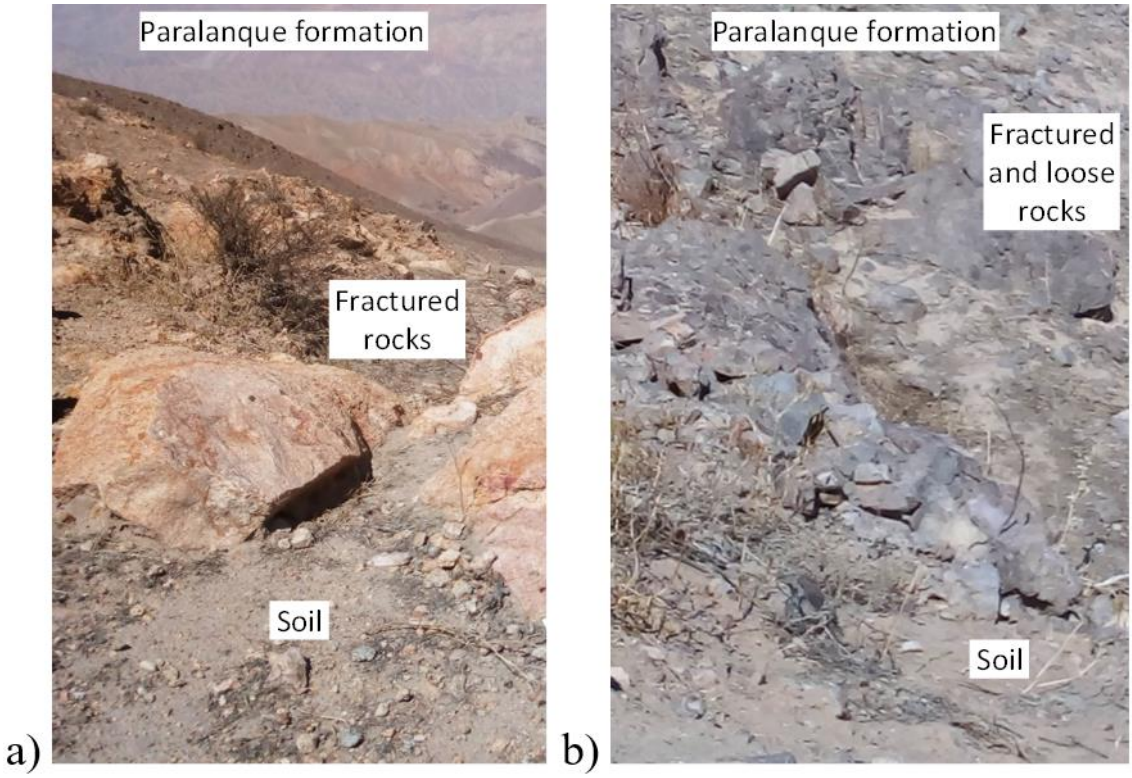


Figure 7. Rock undergoing physical a) alteration and b) weathering.

3.2. Deposit Classification

The samples collected from the field were analyzed by conducting gradation test, Atterberg limit test, specific gravity test, natural moisture content determination, and direct shear test for only one sample. These analyses were performed at the Soils and Rocks Laboratory of the Civil Engineering Department of Universidad de Lima. Undisturbed samples of twelve 30–40-cm deep pits were located at the head of the Mirave Sub-basin, as shown in Figure 5.

The particle-size distribution (or granulometric curve) of the samples obtained through the ASTM D6913/D6913M-17 standard is shown in Figure 8, which presents soil gradation between 75 mm (3 in) and 0.075 mm. Table 1 shows the results of the analyses conducted for the soil samples and presents a detailed summary of the sieving particle analysis. Of the 12 samples tested, 9 are granular soils or silty gravels, which do not present cohesion or plasticity due to the null presence of clay. All the plasticity tests carried out with the ASTM D4318-17e1 standard did not show plasticity, so it can be concluded that the fines present in the sample are all non-plastic silt, which is a type of granular soil, although they have higher fine content such as of 15 or 20%. Sample 1, located in the town of Mirave, non-plastic silt and with 98% fines, it is understood that the majority is silt and there is no clay due to its lack of plasticity. Finally, from the field inspection of the basin and the debris flow event of February 2019 (Del Savio *et al.*, 2019), the classification and characterization of the Mirave Basin is: Morphologically is a debris flows in riverbed; geological structure is an erosion of residual soils and rock weathering; mechanism of occurrence is landslides and accumulation of material in the riverbed; transportation mechanism is highly non-stationary and pulsating; effect on the river bed is erosive and granulometry is coarse granulometry (pebbles, gravel, sands and silts).

According to the granulometric curve of soils, sand and gravel particles can be classified as well or poorly graded. To this end, two parameters were determined: 1) the uniformity coefficient (C_u), which evaluates how evenly the soil granulometry is distributed; and 2) the gradation coefficient (C_c), which evaluates the progression of the variation in the size of the soil particles (Wibisono *et al.*, 2018). According to ASTM D6913M-17, well-graded sands and gravels must satisfy $C_u \geq 6$ and $C_u \geq 4$, respectively, and $1 \leq C_c \leq 3$; if they do not satisfy any of these criteria, they will be poorly graded. From the lab results of granulometry, C_u was found to vary from 8.9 to 887.7 and C_c ranged from 0.6 to 8, which lead to the conclusion that the evaluated soils are mixed and can be classified between well and poorly graded. Based on the soil plasticity tests (ASTM D4318-17e1, 2020d), the soils were determined to be non-plastic (NP) because of the absence of fine soils (silt or clays) greater than 12% in their composition, which is common in soils without cohesion. The results from the specific gravity of solids (ASTM D5550-14, 2020e) were in the range of 2.65–2.67, which are typical values for sands. The direct shear test (ASTM D3080/D3080M-11, 2020c) was conducted using particles smaller than 4.8 mm in diameter and with an average density of 1.68 g/cm^3, resulting in an average friction angle of 40.4° and cohesion equal to zero.

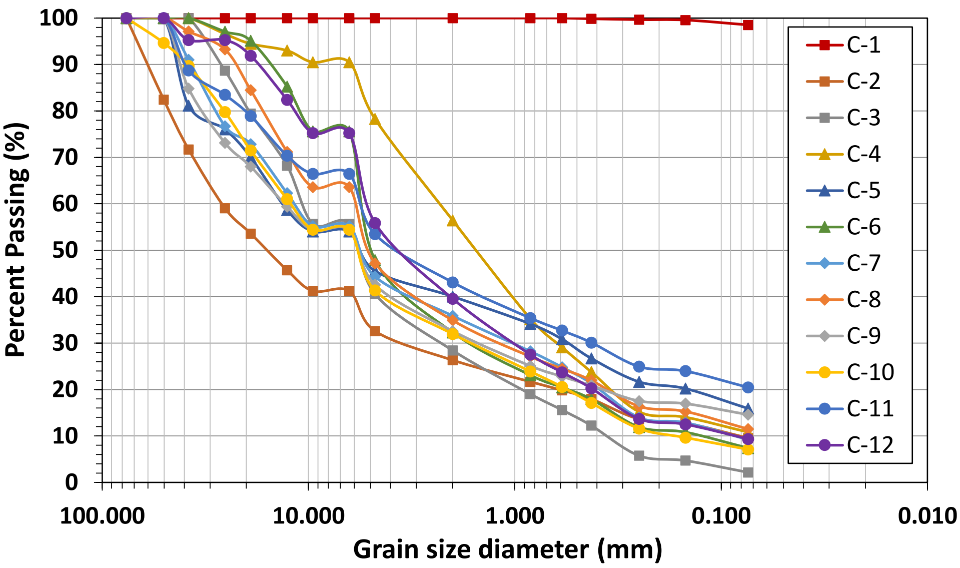


Figure 8. Granulometric curves of samples collected.

The soil erosion factor (K) can be used to measure the susceptibility of soil particles to detachment and transport by surface runoff. For this study, we used the Sharpley and Williams, (1990), *Revised Universal Soil Loss Equation* (RUSLE) concept to determine the soil erosion of the samples brought to the laboratory. The factors affecting erosion are the texture, structure, permeability, and organic content of the soil. Among these factors, texture is the most influential. Several authors use different methods to indirectly determine soil erosion (Renard *et al.*, 1997; Sharpley and Williams, 1990; Wischmeier and Smith, 1978; Hooke, 2003; Romero-Díaz *et al.*, 2007).

For the fraction of sand, silt and clay, in this work used on the Erosion Productivity Impact Calculator - EPIC model (Sharpley and Williams, 1990). The Equation 1 is as follows:

$$K = \left\{0.2 + 0.3exp\left[-0.0256SAN\left(1 - \frac{SIL}{100}\right)\right]\right\} x \left(\frac{SIL}{CLA+SIL}\right)^{0.3} x \left[1 - \frac{0.25C'}{C'+exp(3.72-2.95C')}\right] x \left[1 - \frac{0.7SN_1}{SN_1+exp(-5.51+22.9SN_1)}\right] \quad (1)$$

Where *SAN* is the sand fraction in %; *SIL* is the silt fraction in %; *CLA* is the clay fraction in %; *C'* is the soil total organic carbon content in %; and SN_1=1-*SAN*/100.

Then, to find the modified K value considering the gravel fraction as a key parameter, the proposal by Chang-Xing (2009) and used in works such as Wang *et al.*, 2013 and Hu *et al.* 2019, to modify soil erodibility for different ranges of gravel content. The modification coefficient M (relative soil erosion) was determined by the following Equation 2:

$$M = \{0.0781 exp(-0.0249G)\ 0.294 - 0.0123G\ 1 - 0.0829G\ G > 20\%\ 10\% < G \leq 20\%\ G \leq 10\% \quad (2)$$

Where G is the gravel content. The modified soil erodibility (K_r) can be calculated by the following Equation 3:

$$K_r = K \cdot M \quad (3)$$

As explained above, in the 12 samples silt predominates over clay; therefore, assuming that the clay content is 10% of the fines, the values of K and Kr presented in Table 1 were calculated. It can be seen that the Kr range increases notably with values of 0.48 and 1.12, as minimum and maximum respectively. This indicates the important effect of gravel content in the calculation of soil erosion. Considering that the silt content is greater than the clay content, Kr values greater than 0.4 will be obtained, which indicates that they are the most erodible of all soils. This means that they separate easily and tend to form high runoff rates. Finally, the annual soil losses in the study area show that the erosion factor is high.

From the soil analysis, it was found that the soil has a high content of non-plastic silt (they do not have enough colloids) and results in an area susceptible to erosion. The unconsolidated conglomerate of the formation presents an unstable sandy-silty matrix favoring an increase in the production of loose materials. Therefore, it was quantified that the erosion processes of all the slopes adjoining the channel of the stream generate loose material that, in the event of extraordinary rains, is dragged into the channel. Finally, the study will continue to take into consideration the information from the rain gauge recently installed in Mirave, in order to have a better understanding of the triggering factors of flows and slope erosion.

3.3. Discussion

Fine soils or particles smaller than 0.005 mm with an IP of less than 4 do not have sufficient colloids to support erosion (Grabowski *et al.*, 2011; Grissinger, 1966; Sherard *et al.*, 1976; Smerdon and Beasley, 1959; Lyle and Smerdon, 1965). Briaud *et al.* (2017) reported that SP materials (poorly graded sand), SM (silty sand), and ML (low plasticity silt) exhibit obvious erosional features. Therefore, the presence of NP silts in the matrix of the soils under analysis in this study may be related to the areas susceptible to erosion. The latter does not mean that plastic materials do not undergo erosion, because, as mentioned by Partheniades (2006), a plastic clay can be easily eroded under moving water depending on the water chemistry.

Areas with erosional evidence linked to debris flows in the studied region show that materials with variable percentages of sand and silt ranging from 40% to 55% and without plasticity are susceptible to erosional processes. Surface hydraulic erosion processes act with greater intensity on materials without plasticity and with particles between 0.01 and 1-mm diameter (Yang and Liu, 2016); i.e., silts and sands. The matrix-supported conglomerates, which are predominant in the Mirave Sub-basin, exhibit an unstable sandy-silt matrix when undergoing erosional processes. The instability of the matrix disintegrates the unconsolidated conglomerate of the Moquegua formation, favoring an increase in the production of loose materials. The loose materials of the hillside, the loose materials in the gullies, and the disaggregated materials from previous landslides are usually the result of the debris flows (Brayshaw and Hassan, 2009; Glade, 2005; Takahashi, 2009). These processes occur at the head of the Mirave Micro Basin with a higher intensity, which is similar to what has been recorded in other studies (Brayshaw and Hassan, 2009; Jakob, 2005).

Three profiles with different lithology are analyzed in Figure 9 to discuss how lithological features affect the generation of different types of debris flows, all located in the headwaters of the Mirave Micro Basin.

Table 1. Report of tests analysis for samples.

ID	USCS	w (%)	G_s	%G	%S	%F	LL	PL	D_{10} (mm)	D_{30} (mm)	D_{60} (mm)	C_U	C_C	K	K_r
C-1	ML	5.7	-	0.0	1.5	98.5	NP	NP	0.002	-	-	-	-	0.483	0.483
C-2	GP-GM	1.5	2.65	67.4	23.0	9.5	NP	NP	0.082	3.327	26.184	320.5	5.2	0.363	0.530
C-3	GW	2.1	-	59.4	38.4	2.2	NP	NP	0.355	2.240	10.520	29.7	1.3	0.305	0.542
C-4	SP-SM	1.5	-	21.7	67.4	10.8	NP	NP	0.054	0.622	2.304	43.0	3.1	0.247	1.124
C-5	GM	1.4	-	54.5	29.6	15.9	NP	NP	0.015	0.552	13.309	888.7	1.5	0.346	0.695
C-6	GP-GM	1.1	-	52.1	40.5	7.4	NP	NP	0.126	1.613	5.399	42.8	3.8	0.304	0.649
C-7	GW-GM	1.8	2.64	55.7	35.1	9.2	NP	NP	0.086	1.028	11.577	134.1	1.1	0.321	0.627
C-8	GP-GM	3.2	-	52.8	35.7	11.5	NP	NP	0.043	1.162	5.961	139.7	5.3	0.322	0.675
C-9	GM	1.2	-	57.5	28.0	14.6	NP	NP	0.019	1.483	13.069	682.5	8.8	0.350	0.653
C-10	GW-GM	1.5	2.66	58.7	34.2	7.1	NP	NP	0.164	1.620	12.143	74.2	1.3	0.322	0.583
C-11	GM	2	-	46.5	33.0	20.5	NP	NP	0.008	0.421	5.504	671.1	3.9	0.340	0.832
C-12	SW-SM	1.1	-	44.1	46.5	9.3	NP	NP	0.086	1.005	5.062	59.2	2.3	0.291	0.758

USCS: Unified Soil Classification System; w: water content: Gs: specific gravity of soils; LL, PL: Liquid and Plastic Limit; %G, %S, %F: Gravel, Sand and Fine content; Dx: x% of the soil particles are finer than this size; Cu; uniformity coefficient; Cc: coefficient of curvature; K: soil erodibility factor; K_r: modified soil erodibility

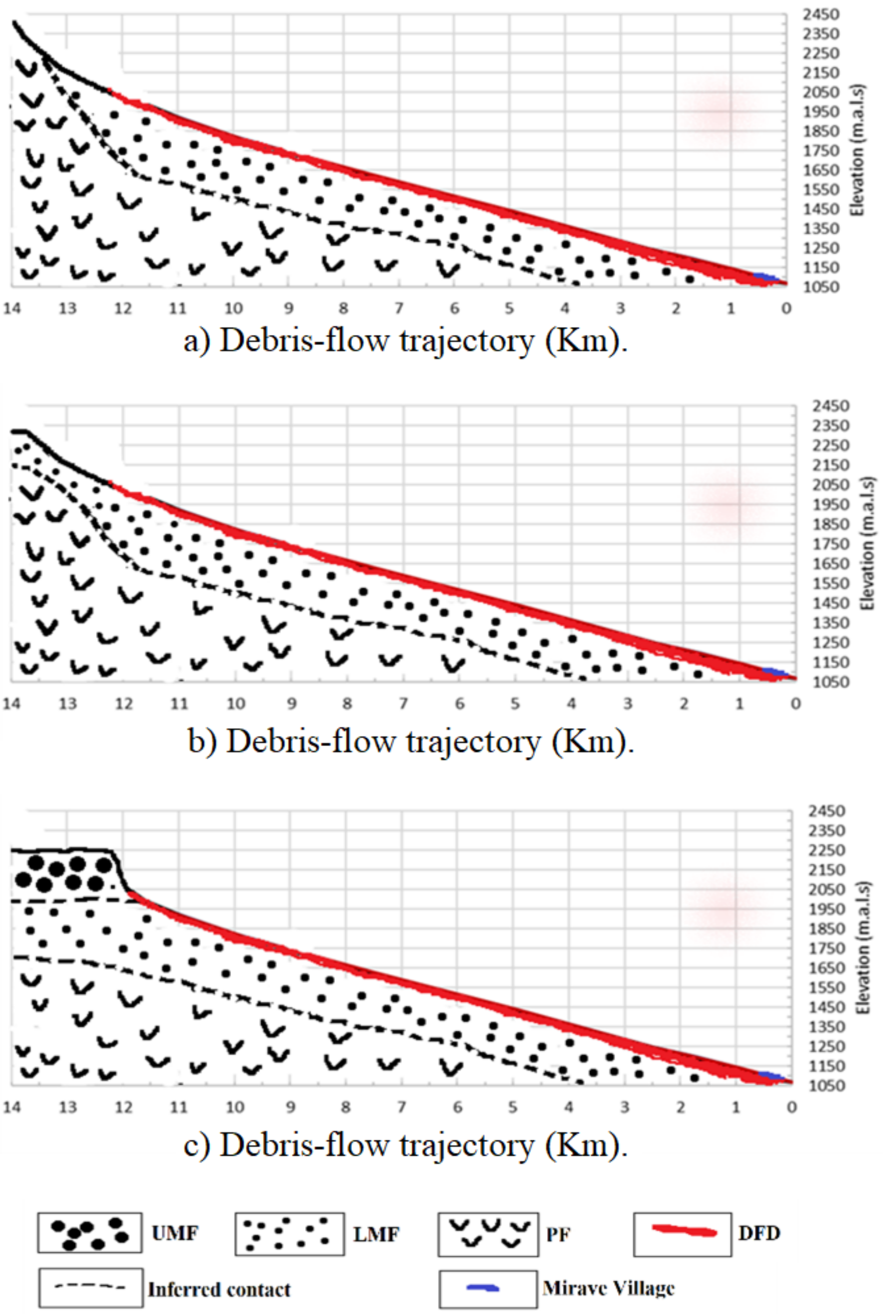


Figure 9. a) Andesite head; b) supported array conglomerate head; and c) head of supported clast conglomerates.

The first one occurs in areas of rocky outcrops of volcanic origin (Figure 9a), where the debris flow is generated by the action of water concentrated in furrows, which, because of the high slopes of these formations, is strong enough to transport angular debris contained in its course to the lower parts. In the field study conducted after the 2019 event, it was found that the blocks have angular shapes, a material that comes from the Mirave dejection. The transport of angular rock type blocks required a greater hydraulic energy in the form of water courses for their movement, which indicates that they may be related to large events, such as the one that took place in 2019. Downstream flows down the rills at high speed may contribute to the generation of debris flows downstream, specifically in the gully channel in which loose sediments are accumulated.

The matrix-supported conglomerates of the lower Moquegua formation exhibit irregular distributions and contain polymictic clasts. Two types of these materials can be identified with different erosive behavior in terms of water (Figure 9b). Water erosion in these materials has formed numerous gullies, with different depths and extensions and without the presence of a water table. These gullies are distributed throughout most of the basin. The interior of the gullies shows intense erosive activity and accumulation of materials transported by gravity and water. The advance of the gullies in the upper Mirave Basin affects hillsides with steep slopes (approximately 30°), causing potential landslides at the contact of the gully and the hillside, which can lead to the onset of debris flows. The surface soils of approximately 20-cm diameter cover the slopes that are unaffected by the gullies and present vegetation only in the summer months (January, February, and March).

The lithological unit comprises nonuniform consolidated and clast-supported conglomerates of the upper Moquegua formation (Figure 9c). Physical weathering acts slowly to disaggregate the conglomerate clasts, and environmental agents transport these loose materials down the hillside. They can be accumulated on the hillsides and in the channels of pre-existing gullies of the lower Moquegua formation, which exhibit a more moderate relief. High speed streams fall as jets or small waterfalls onto the lower areas and are connected to different creeks. With that energy, they enter the areas of colluvial materials and in some cases (especially when the force of the water is strong, such as in big storms), can liquefy these materials and initiate the flow of debris.

4. CONCLUSIONS

Given the arid climate of the Mirave region, the formation of clay surface soils is not typical. The low percentage of this type of minerals indicates low cohesion; thus, the matrix-supported conglomerates and their corresponding residual and colluvial soils are in advanced erosion processes. The erosion of the matrix of these conglomerates causes the collapse of these materials, which are the most abundant in the debris flow generating zones of the Mirave Sub-basin. At the headwaters, a change in sediment production has been found, primarily generated from residual and colluvial soils, enhancing the debris flow generation mechanism.

The main geotechnical features in these arid mountain basins are particle size, plasticity, and erodibility. The absence of plasticity and the abundance of potentially erodible particles because of their size (sand and silt) strengthen the erosive features of the region, i.e., the abundance of gullies and rills. However, to quantify erosion, future studies should conduct rock tests, in addition to considering the use of methodologies for calculating the erosion factor according to taxonomic classes, with a greater number of samples that allow for a stricter statistical analysis.

In this study, the evaluation of the source of debris flows in the Mirave Basin has identified that erosion processes on hillsides maximize geologic hazards and risks to human communities established in the alluvial fan of the basin.

5. ACKNOWLEDGMENTS

This research was supported by the National Innovation Program for Competitiveness and Productivity (Innóvate Perú) of the Ministry of Production of Peru and the Scientific Research Institute (IDIC) of the Universidad de Lima with Project Number 687-INNOVATEPERU-RPR-2017.

6. REFERENCES

ASTM INTERNATIONAL. **ASTM D2216-19**. Standard Test Methods for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass. West Conshohocken, 2020a.

ASTM INTERNATIONAL. **ASTM D2487-17e1**. Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System). West Conshohocken, 2020b.

ASTM INTERNATIONAL. **ASTM D3080/D3080M-11**. Standard Test Method for Direct Shear Test of Soils Under Consolidated Drained Conditions. West Conshohocken, 2020c.

ASTM INTERNATIONAL. **ASTM D4318-17e1.** Standard Test Method for Specific Gravity of Soil Solids by Gas Pycnometer. West Conshohocken, 2020d.

ASTM INTERNATIONAL. **ASTM D5550-14.** Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils. West Conshohocken, 2020e.

ASTM INTERNATIONAL. **ASTM D6913M-17**. Standard Test Methods for Particle-Size Distribution (Gradation) of Soils Using Sieve Analysis. West Conshohocken, 2020f.

BELLIDO BRAVO, E. **Geology of the Moquegua Quadrangle. Page: 35-u**. Lima: Instituto Geológico Minero y Metalúrgico, 1979.

BRAYSHAW, D.; HASSAN, M. A. Debris flow initiation and sediment recharge in gullies. **Geomorphology**, v. 109, n. 3-4, p. 122-131, 2009. https://doi.org/10.1016/j.geomorph.2009.02.021

BRIAUD, J.-L.; GOVINDASAMY, A. V.; SHAFII, I. Erosion charts for selected geomaterials. **Journal of Geotechnical and Geoenvironmental Engineering**, v. 143, n. 10, p. 04017072, 2017. https://ascelibrary.org/doi/full/10.1061/(ASCE)GT.1943-5606.0001771

CANNON, S. H.; RENEAU, S. L. Conditions for generation of fire-related debris flows, Capulin Canyon, New Mexico. **Earth Surface Processes and Landforms**, v. 25, n. 10, p. 1103-1121, 2000. https://doi.org/10.1002/1096-9837(200009)25:10%3C1103::AID-ESP120%3E3.0.CO;2-H

CHANG, M.; LIU, Y.; ZHOU, C.; CHE, H. Hazard assessment of a catastrophic mine waste debris flow of Hou Gully, Shimian, China. **Engineering Geology**, v. 275, p. 105733, 2020. https://doi.org/10.1016/j.enggeo.2020.105733

CHANG-XING, S. H. I. Effects of gravel content on soil erodibility and the methods of estimating soil erodibility factor K. **Chinese Journal of Soil Science,** v. 40, n. 6, p. 1398–1401, 2009.

CHEN, J.-C.; LIN, C.-W.; WANG, L.-C. Geomorphic characteristics of hillslope and channelized debris flows: A case study in the Shitou area of central Taiwan. **Journal of Mountain Science**, v. 6, n. 3, p. 266-273, 2009. https://doi.org/10.1007/s11629-009-0250-0

DECOU, A.; VON EYNATTEN, H.; DUNKL, I.; FREI, D.; WÖRNER, G. Late Eocene to Early Miocene Andean uplift inferred from detrital zircon fission track and U–Pb dating of Cenozoic forearc sediments (15-18°S). **Journal of South American Earth Sciences**, v. 45, p. 6-23, 2013. https://doi.org/10.1016/j.jsames.2013.02.003

DEL SAVIO, A. A.; QUISCA ASTOCAHUANA, S. I.; CASTILLO NAVARRO, L. F. Numerical simulation of debris flows of the catastrophic event of February 2019 in Mirave–Peru. **Revista Ambiente & Água**, v. 14, n. 6, p. 1-20, 2019. https://doi.org/10.4136/ambi-agua.2437

DI, B.; ZHANG, H.; LIU, Y. *et al.* Assessing susceptibility of debris flow in Southwest China using gradient boosting machine. **Scientific Reports**, v. 9, n. 1, p. 12532, 2019. https://doi.org/10.1038/s41598-019-48986-5

DING, M.; HUANG, T.; ZHENG, H.; YANG, G. Respective influence of vertical mountain differentiation on debris flow occurrence in the Upper Min River, China. **Scientific Reports**, v. 10, n. 1, p. 11689, 2020. https://doi.org/10.1038/s41598-019-48986-5

ESPER ANGILLIERI, M. Y. Debris flow susceptibility mapping using frequency ratio and seed cells, in a portion of a mountain international route, Dry Central Andes of Argentina. **CATENA**, v. 189, p. 104504, 2020. https://doi.org/10.1016/j.catena.2020.104504

ESPER ANGILLIERI, M. Y.; PERUCCA, L.; VARGAS, N. Spatial and temporal analysis of debris flow occurrence in three adjacent basins of the western margin of Grande River: Quebrada de Humahuaca, Jujuy, Argentina. **Geografiska Annaler: Series A, Physical Geography**, v. 102, n. 2, p. 83-103, 2020. https://doi.org/10.1080/04353676.2020.1744075

GLADE, T. Linking debris-flow hazard assessments with geomorphology. **Geomorphology**, v. 66, n. 1-4, p. 189-213, 2005. https://doi.org/10.1016/j.geomorph.2004.09.023

GRABOWSKI, R. C.; DROPPO, I. G.; WHARTON, G. Erodibility of cohesive sediment: The importance of sediment properties. **Earth-Science Reviews**, v. 105, n. 3-, p. 101-120, 2011. https://doi.org/10.1016/j.earscirev.2011.01.008

GRISSINGER, E. H. Resistance of selected clay systems to erosion by water. **Water Resources Research**, v. 2, n. 1, p. 131-138, 1966. https://doi.org/10.1029/WR002i001p00131

GUERRERO, P. C.; ROSAS, M.; ARROYO, M. T. K.; WIENS, J. J. Evolutionary lag times and recent origin of the biota of an ancient desert (Atacama-Sechura). **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 28, p. 11469-11474, 2013. https://doi.org/10.1073/pnas.1308721110

HOOKE, J. Coarse sediment connectivity in river channel systems: a conceptual framework and methodology. **Geomorphology**, v. 56, n. 1-2, p. 79-94, 2003. https://doi.org/10.1016/S0169-555X(03)00047-3

HOUSTON, J. Variability of precipitation in the Atacama Desert: its causes and hydrological impact. **International Journal of Climatology**, v. 26, n. 15, p. 2181-2198, 2006. https://doi.org/10.1002/joc.1359

HU, S.; LI, L.; CHEN, L.; CHENG, L.; YUAN, L.; HUANG, X. *et al.* Estimation of Soil Erosion in the Chaohu Lake Basin through Modified Soil Erodibility Combined with Gravel Content in the RUSLE Model. **Water**, v. 11, n. 9, p. 1806, 2019. https://doi.org/10.3390/w11091806

HUNGR, O. Classification and terminology. **Debris-flow Hazards and Related Phenomena**. p. 9-23, 2005. Berlin: Springer Praxis Books.

ILABAYA. District Municipality. **Risk assessment report on debris flow risk in the town of Mirave, district of Ilabaya, province of Jorge Basadre, Tacna Department.** Lima: Centro Nacional d Estimación, Prevención y Reducción del Riesgo de Desastres, 2016.

IMAIZUMI, F.; TSUCHIYA, S.; OHSAKA, O. Behavior of boulders within a debris flow initiation zone. **International Journal of Erosion Control Engineering**, v. 9, n. 3, p. 91-100, 2016. https://doi.org/10.13101/ijece.9.91

INDECI. **Complementary Report Number 408 - 09/02/2019/COEN - INDECI/02:20 Hours (Report N° 01) Huaico en el distrito de Ilabaya–Tacna**. 2019. Available: https://www.indeci.gob.pe/emergencias/tacna-jorge-basadre-ilabaya-mirave-huaico-reporte-complementario-01 Access: 8 Aug. 2020.

JAKOB, M. A size classification for debris flows. **Engineering Geology**, v. 79, n. 3-4, p. 151-161, 2005. https://doi.org/10.1016/j.enggeo.2005.01.006

LARA, G.; PERUCCA, L.; ROTHIS, M. Morphometric, geomorphologic and flood hazard analysis of an arid mountain river basin, central pre-Andes of Argentina. Southwestern South America. **Geografia Fisica e Dinamica Quaternaria, GFDQ**, v. 41. p. 83-97, 2018

LYLE, W. M.; SMERDON, E. T. Relation of compaction and other soil properties to erosion resistance of soils. **Transactions of the ASAE**, v. 8, n. 3, p. 419-422, 1965. https://doi.org/10.13031/2013.40536

MARTINEZ, W.; ZULOAGA, A. **Explanatory Report of the geology of the Moquegua Quadrangle. (35-u)**. 2000. Available: https://repositorio.ingemmet.gob.pe/handle/20.500.12544/2044. Access: 5 June 2021.

MEDINA ALLCA, L.; LUQUE POMA, G. **Evaluation of geological hazards in Mirave town center and Alto Mirave sector.** Lima: INGEMMET, 2016.

MELÉNDEZ, D.; RAMOS-FERNÁNDEZ, L.; VELÁSQUEZ, T.; ALTAMIRANO, L. Simulation with a conceptual distributed hydrological model on a daily scale in a semi-arid basin of the Lurin River, Peru. **Idesia,** v. 39, n. 1, 2021.

MEYER, N. K.; SCHWANGHART, W.; KORUP, O.; NADIM, F. Roads at risk–traffic detours from debris flows in southern Norway. **Natural Hazards and Earth System Sciences Discussions**, v. 2, p. 6623-6651, 2014. https://doi.org/10.5194/nhess-15-985-2015

MOREIRAS, S. M. Frequency of debris flows and rockfall along the Mendoza river valley (Central Andes), Argentina: Associated risk and future scenario. **Quaternary International**, v. 158, n. 1, p. 110-121, 2006. https://doi.org/10.1016/j.quaint.2006.05.028

MOREIRAS, S. M.; SEPÚLVEDA, S. A.; CORREAS-GONZÁLEZ, M.; LAURO, C.; VERGARA, I.; JEANNERET, P. *et al.* Debris Flows Occurrence in the Semiarid Central Andes under Climate Change Scenario. **Geosciences,** v. 11, 43, 2021. https://doi.org/10.3390/geosciences11020043

PARTHENIADES, E. **Engineering properties and hydraulic behavior of cohesive sediments**. Boca Raton: CRC Press, 2006.

RENARD, K. G.; FOSTER, G. R.; WEESIES, G. A.; MCCOOL, D. K.; YODER, D. C. **Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)**. Washington, D.C.: U.S. Dept. of Agriculture, 1997.

ROMERO-DÍAZ, A.; ALONSO-SARRIÁ, F.; MARTÍNEZ-LLORIS, M. Erosion rates obtained from check-dam sedimentation (SE Spain). A multi-method comparison. **CATENA**, v. 71, n. 1, p. 172-178, 2007. https://doi.org/10.1016/j.catena.2006.05.011

SENAMHI. **National hydrometeorological data**. Available: https://www.senamhi.gob.pe/servicios/?p=estaciones. Access: 7 July 2021.

SHARPLEY, A. N.; WILLIAMS, J. R. (Orgs.). **EPIC, Erosion/Productivity Impact Calculator. 1, Model documentation**. Springfield: Agricultural Research Service, 1990.

SHERARD, J. L.; DUNNIGAN, L. P.; DECKER, R. S. Identification and nature of dispersive soils. **Journal of the Geotechnical Engineering Division**, v. 102, n. 4, p. 287-301, 1976. https://doi.org/10.1061/AJGEB6.0000256

SMERDON, E. T.; BEASLEY, R. P. **The tractive force theory applied to stability of open channels in cohesive soils**. Columbia: University of Missouri, College of Agriculture, 1959. p. 1-36.

STOLLE, A.; LANGER, M.; BLÖTHE, J. H.; KORUP, O. On predicting debris flows in arid mountain belts. **Global and Planetary Change**, v. 126, p. 1-13, 2015. https://doi.org/10.1016/j.gloplacha.2014.12.005

TAKAHASHI, T. A review of Japanese debris flow research. **International Journal of Erosion Control Engineering**, v. 2, n. 1, p. 1-14, 2009. https://doi.org/10.13101/ijece.2.1

WANG, B.; ZHENG, F.; RÖMKENS, M. Comparison of soil erodibility factors in USLE, RUSLE2, EPIC and Dg models based on a Chinese soil erodibility database. **Acta Agriculturae Scandinavica, Section B — Soil & Plant Science**, v. 63, n. 1, p. 69-79, 2013. https://doi.org/10.1080/09064710.2012.718358

WIBISONO, G.; NUGROHO, S. A.; UMAM, K. The influence of sand's gradation and clay content of direct shear test on clayey sand. **IOP Conference Series: Materials Science and Engineering**, v. 316, p. 012038, 2018.

WISCHMEIER, W. H.; SMITH, D. D. **Predicting rainfall erosion losses: a guide to conservation planning**. Washington, D.C.: U.S. Dept. of Agriculture, 1978.

YANG, J.; LIU, X. Shear wave velocity and stiffness of sand: the role of non-plastic fines. **Géotechnique**, v. 66, n. 6, p. 500-514, 2016. https://doi.org/10.1680/jgeot.15.P.205

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Methodological approaches for imputing missing data into monthly flows series

ARTICLES doi:10.4136/ambi-agua.2795

Received: 15 Sep. 2021; Accepted: 07 Mar. 2022

Michel Trarbach Bleidorn[1]; Wanderson de Paula Pinto[2]; Isamara Maria Schmidt[1*]; Antonio Sergio Ferreira Mendonça[1]; José Antonio Tosta dos Reis[1]

[1]Departamento de Engenharia Ambiental. Universidade Federal do Espírito Santo (UFES), Avenida Fernando Ferrari, n° 514, CEP: 29075-910, Vitória, ES, Brazil. E-mail: michelbleidorn@gmail.com, anserfm@terra.com.br, jatreis@gmail.com
[2]Núcleo Integrado de Pesquisa em Engenharia Ambiental. Faculdade da Região Serrana (FARESE), Rua Jequitibá, n° 121, CEP: 29645-000, Santa Maria de Jetibá, ES, Brazil. E-mail: wandersondpp@gmail.com
[*]Corresponding author. E-mail: isamaraschmidt@gmail.com

ABSTRACT

Missing data is one of the main difficulties in working with fluviometric records. Database gaps may result from fluviometric stations components problems, monitoring interruptions and lack of observers. Incomplete series analysis generates uncertain results, negatively impacting water resources management. Thus, proper missing data consideration is very important to ensure better information quality. This work aims to analyze, comparatively, missing data imputation methodologies in monthly river-flow time series, considering, as a case study, the Doce River, located in Southeast Brazil. Missing data were simulated in 5%, 10%, 15%, 25% and 40% proportions following a random distribution pattern, ignoring the missing data generation mechanisms. Ten missing data imputation methodologies were used: arithmetic mean, median, simple and multiple linear regression, regional weighting, spline and Stineman interpolation, Kalman smoothing, multiple imputation and maximum likelihood. Their performances were compared through bias, root mean square error, absolute mean percentage error, determination coefficient and concordance index. Results indicate that for 5% missing data, any methodology for imputing can be considered, recommending caution for arithmetic mean method application. However, as the missing data proportion increases, it is recommended to use multiple imputation and maximum likelihood methodologies when there are support stations for imputation, and the Stineman interpolation and Kalman Smoothing methods when only the studied series is available.

Keywords: Doce river, imputation, missing data.

Abordagens metodológicas para imputação de dados faltantes de vazões médias mensais

RESUMO

A falta de dados é uma das principais dificuldades no trabalho com registros fluviométricos. As lacunas no banco de dados podem resultar de problemas nos componentes das estações fluviométricas, interrupções no monitoramento e falha dos observadores. A análise de séries incompletas gera resultados incertos, impactando negativamente a gestão dos recursos hídricos. Assim, a consideração adequada dos dados faltantes é muito importante para garantir a qualidade de informação. Este trabalho teve como objetivo analisar, comparativamente, metodologias de imputação de dados faltantes em séries temporais de vazões fluviais mensais, considerando, em um estudo de caso, o Rio Doce, localizado no Sudeste do Brasil. Os dados faltantes foram simulados nas proporções de 5%, 10%, 15%, 25% e 40% seguindo um padrão de distribuição aleatória e ignorando os mecanismos de geração de falhas. Foram utilizadas dez metodologias de imputação de dados faltantes: média aritmética, mediana, regressão linear simples e múltipla, ponderação regional, interpolação spline e Stineman, suavização de Kalman, imputação múltipla e máxima verossimilhança. Seus desempenhos foram comparados por meio dos indicadores viés, raiz do erro quadrático médio, erro absoluto médio percentual, coeficiente de determinação e índice de concordância. Os resultados indicam que para 5% de dados faltantes, qualquer metodologia de imputação pode ser considerada, recomendando cautela na aplicação da média aritmética. No entanto, à medida que a proporção de dados faltantes aumenta, recomenda-se o uso das metodologias imputação múltipla e máxima verossimilhança quando houver estações de suporte para imputação, e os métodos de interpolação Stineman e suavização de Kalman quando apenas as séries estudadas estiverem disponíveis.

Palavras-chave: dados faltantes, imputação, Rio Doce.

1. INTRODUCTION

Knowledge about the water regime in a river basin is fundamental in hydrological studies, being an indispensable factor for adequate water-resource management (Moreira, 2006). In this sense, Brazilian Law 9,433 (Brasil, 1997), which instituted the National Water Resources Policy, indicates, among the management instruments, the National Water Resources Information System. In this context, the hydrometeorological monitoring network maintained by the National Water Agency (ANA) and the availability of the databases generated in the Hydrological Information System (HIDROWEB) become relevant.

The instrument National Water Resources Information System, however, is still incipient in relation to the others, mainly due to the limited number of monitoring stations and the incompleteness of the data generated (Fioreze and Oliveira, 2010). The existence of gaps in the historical series is due to technical or maintenance problems, non-ideal climatic conditions, instrumental failures or device errors during data collection, human error during data entry, calibration processes and/or data damage due to malfunction of machines storage, construction and organization of hydrometric databases (Gao, 2017; Johnston, 1999; Peña-Angulo *et al.*, 2019; Tencaliec, 2017; Tucci, 1997).

According to McKnight *et al.* (2007), “in general, the term missing data means that some type of information about the phenomenon in which we are interested is missing” and, therefore, the sample is called incomplete. The analysis of incomplete flow-time series produces negative impacts, especially on stochastic decompositions of series, compromising information such as trend, stationarity, cycle and seasonality (Box and Cox, 1964).

As discussed by Roth *et al.* (1999) and Pigott (2001), the existence of missing values in time series generally decreases the capacity and precision of statistical analysis approaches and contributes to biased estimates of the relationship between variables, which may cause inaccurate assumptions in data set exploration which can negatively impact water resources management, for example, in determination of maximum permissible uptake and ecological flows, extreme flows estimation, flows forecasting, hydraulic systems designs, among others. Due to this fact, the reconstruction of incomplete series and the treatment of missing data must be seen as a priority in the data preparation procedure (Hamzah *et al.*, 2020).

Because missing data imputation is a useful tool in water-resource management studies (Barnetche and Kobiyama, 2006), several authors have worked on the application of techniques for imputing missing data in hydrological studies resulting in a variety of methods ranging from simple imputation by mean or median to widely used statistical methods such as Regional Weighting (Ely *et al.*, 2021); interpolations (linear, quadratic and cubic) (Gyau-Boakye and Schultz, 1994, Hamzah *et al.*, 2020); methods based on linear regressions (single and multiple) (Kamwaga *et al.*, 2018; Khalifeloo *et al.*, 2015); Self Organizing Map (SOM) and Soil and Water Assessment Tool (SWAT) (Kim *et al.*, 2015); to more advanced and robust methods, such as different Artificial Neural Network approaches (Canchala-Nastar *et al.*, 2019; Elshorbagy *et al.*, 2000; Nkiaka *et al.*, 2016; Starrett *et al.*, 2010; Vega-Garcia *et al.*, 2019); machine learning methods (Heras and Matovelle, 2021; Rado *et al.*, 2019); satellite radar altimetry and multiple imputation (Ekeu-Wei *et al.*, 2018); combination of regression and autoregressive integrated moving average (ARIMA) models called dynamic regression (Tencaliec *et al.*, 2015); Singular Spectrum Analysis (SSA) and Multichannel Singular Spectrum Analysis (MSSA) (Semiromi and Koch, 2019); among many others. The many methods that can be used for hydrological missing data imputation resulted in literature reviews as can be seen in Ben Aissia *et al.* (2017) and Hamzah *et al.* (2020).

Ventura *et al.* (2016) carried out a study to compare statistical methods for filling gaps and to verify which method presents better results for meteorological data series. Three weather stations located in Porto Alegre, Rio de Janeiro and Manaus cities, in Brazil, were chosen. Failures were simulated in real data series and the performances of four methods were compared: simple average, moving average, simple linear regression and multiple linear regression. To verify the obtained results, the mean absolute error and the correlation coefficient were used. The results showed excellent performance of the multiple linear regression method for the variables temperature, humidity and dew point, while the simple average had the best result for the variable atmospheric pressure. None of the four methods presented good results for the variable solar radiation.

Nunes *et al.* (2009) carried out a study with the objective of publicizing the Multiple Imputation (MI) method. The authors selected a 470 surgical patient death outcome data set and adjusted logistic models to it. Two incomplete data sets were generated, one presenting 5% and the other 20% of missing data for the variable albumin. Models were adjusted to the complete series, to the series presenting missing data and the series filled by using MI. The estimates obtained by the analysis of the series presenting missing data and with the filled series were different, mainly for those presenting 20% of missing data. The utilized MI was efficient, because the results achieved with the series filled by imputations were close to those obtained with the complete series. The results obtained considering series filled by using MI were superior to those obtained for series with missing data.

Junger and Leon (2015) presented an imputation method via Maximum Likelihood (ML) that is suitable for multivariate time series using the EM algorithm (Expectation and Maximization) under the assumption of normal distribution. The authors used a database related with tem PM_{10} monitoring stations located in São Paulo city, Brazil. Different approaches were considered to filter the temporal component. A simulation study was carried out to compare the proposed and some frequently used methods of quality and performance. The simulations showed that when the amount of missing data was less than 5%, the complete data analysis generated satisfactory results, regardless of the mechanism that generated the missing data. Imputation quality began to degenerate when missing data proportions exceeded 10%. The proposed imputation method presented good accuracy and precision in different configurations with respect to the missing observation patterns. Most imputations obtained valid results, even under the non-random losses mechanism.

Sattari *et al.* (2017) evaluated different methods of imputing missing data in monthly rainfall time series collected at six stations in southern Iran. Imputation methodologies analyzed include arithmetic mean, inverse distance interpolation, linear regression, multiple imputations, multiple linear regression analysis, non-linear iterative partial least squares algorithm, NR method, single best estimator, UK traditional method and M5 decision model tree. Results showed that arithmetic averaging method, multiple linear regression method and nonlinear iterative partial least squares algorithm perform best. Multiple regression methods provided a successful missing precipitation data estimation. Multiple imputation methods produced the most accurate results for precipitation data from five dependent stations. Finally, the decision-tree algorithm is explicit, and therefore it is used when insights into decision making are needed.

Chen *et al.* (2019) verified the impact of using different methods for imputing missing data in rainfall series on the forecasting hydrological and non-point (H/NPS) pollution performance using the Soil and Water Assessment Tool (SWAT) model. Multiple imputation (MI) and maximum likelihood methods using expectation-maximization bootstrap algorithm (EMB) were considered. Different imputed data sets effects were investigated through a case study in the Daning River Basin, Three Gorges Reservoir Region, China. Results indicate that rainfall data imputation and H/NPS model performance obtained by EMB algorithm are superior to MI performance. Authors highlight the important implications for choosing appropriate imputation methods in H/NPS models to solve data scarcity problems for watershed studies.

Hamzah *et al.* (2022) evaluated the performance of multiple imputations by chained equations (MICE) approach to predicting recurrence in streamflow datasets. To evaluate and verify MICE approach effectiveness in treating missing streamflow data, complete historical daily streamflow series from 2012 to 2014 were used. Later, MICE methods coupled with multiple linear regression (MLR) were applied to restore streamflow rates in Malaysia's Langat River Basin from 1978 to 2016. The best estimation methods are validated with tests such as adjusted R-squared (Adj R^2), residual standard error (RSE) and mean absolute percentage error (MAPE). Findings revealed that the classification and regression tree (CART) method combined with MLR outperformed the other approaches tested, with highest Adj R^2 value and lowest RSE and MAPE values observed regardless of missing conditions.

Abu Romman *et al.* (2021) compared ten imputation methods that were used to impute rainfall depth data in an arid region of the Mediterranean. Series mean, linear interpolation, linear trend, arithmetic mean, normal ratio, inverse distance weighting, linear regression with GPCC data, linear regression with satellite data, stepwise multiple linear regression and multiple imputation were used for these imputations. The results showed that for intervals between 5 and 20% of failures, the stepwise multiple linear regression method produced best results with a root mean square error (RMSE) and mean absolute error (MAE) less than 7 and 2 mm, respectively. This was followed by the Monte Carlo Markov chain expectation-maximization-based multiple imputation method, which had an RSME and MAE of 1.01 and 0.08 mm, respectively, when the series had 20% failures. On the other hand, satellite data use for imputation was adequate for failures between 10 and 15%.

Other studies can be highlighted, especially related to medicine and health (Camargos *et al.*, 2011; Carreras *et al.*, 2021; Khan *et al.*, 2021; Nunes, 2007; Payrovnaziri *et al.*, 2021), air pollution (Choi *et al.*, 2021; Ghazali *et al.*, 2021; Pinto, 2013), engineering, mainly civil and traffic (Abdelgawad *et al.*, 2015; Jiang *et al.*, 2021), meteorology (Afrifa-Yamoah *et al.*, 2020; Bier and Ferraz, 2017; Costa *et al.*, 2021; Ferrari and Ozaki, 2014; García-Peña *et al.*, 2014), agriculture (Jiao *et al.*, 2016; Nishina *et al.*, 2017; Swenson, 2014), energy (Barbosa *et al.*, 2018; Pelisson, 2021) and education (Vinha and Laros, 2018).

Currently, it is unclear in the literature which method is the most appropriate to deal with missing value imputation in river-flow time series. Considering this scenario, this research evaluates ten imputation techniques, based on single and multiple imputation methods: Arithmetic Mean (AM), Median (M), Simple Linear Regression (SLR), Multiple Linear

Regression (MLR), Regional Weighting (RW), Spline Interpolation (SPLINE), Stineman interpolation (STINE), Kalman Smoothing with (KALMAN), Multiple Imputation (MI) and Maximum Likelihood (ML). It is important to emphasize that there are still few studies that use MI and ML imputation methodologies dealing with missing value imputation in river-flow time series, evidencing the need to develop works analyzing their performance, and these results can help sustainable water resource management. In this context, the present research objective is to comparatively analyze methodologies for imputing missing data by an application to Doce River, Brazil, monthly average fluviometric flow-time series.

2. MATERIAL AND METHODS

2.1. Study Area

The Doce River watershed, Figure 1, is located in Southeast Brazil, occupying portions of Minas Gerais and Espírito Santo states between the parallels 17°45' and 21°15' South latitude and the meridians 39°55' and 43º45' West longitude. The Doce River presents 853 km total extension and 83,465 km^2 drainage area (Coelho, 2007). Of this area, 86% belongs to Minas Gerais state and the remaining 14% to Espírito Santo state, being, therefore, a federal dominion watershed.

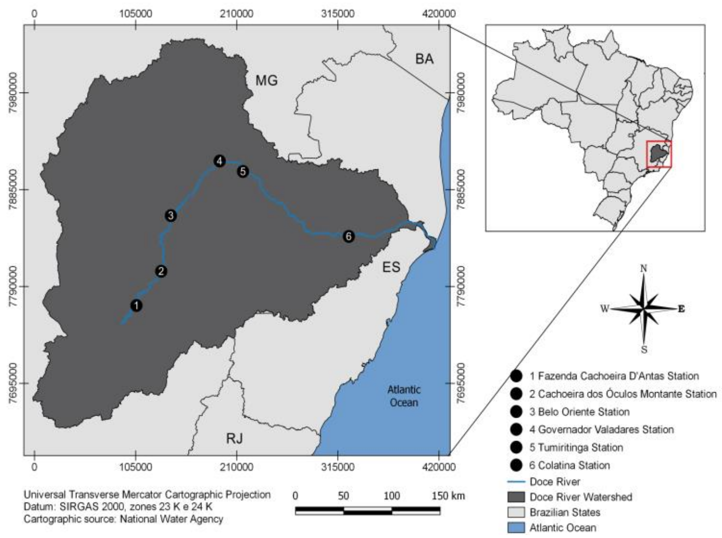


Figure 1. Doce River Watershed.

2.2. Data

Six ANA hydrometeorological network Doce River fluviometric stations were considered in this work. The monthly flow data was obtained from the HIDROWEB Hydrological Information System (Agência Nacional de Águas, 2018). Station-selection criteria were related to the existence of consistent historical average monthly flow-time series for at least 20 years, according to literature recommendations (Longhi and Formiga, 2011; Sarmento, 2007).

Historical series stationarity analysis, that verifies mean and variance identity of two distinct hydrological series subperiods, utilized Fisher's F and t-Student tests, in order to evaluate time series hydrological regime behavior changes, which can be caused by several factors, such as reservoir construction upstream of the fluviometric station, water withdrawal for irrigation agricultural activities use and even local climate regime changes over time (Sousa *et al.*, 2009). For the analysis, the software SisCAH 1.0 - Computational System for Hydrological Analysis - was used (Sousa *et al.*, 2009).

The selected stations are shown in Figure 1. The list of stations considered is presented in Table 1, with respective ANA identification codes, identification nomenclature adopted in this study, geographic coordinates and drainage areas. Colatina station is located in Espírito Santo state, and the others in Minas Gerais state. After analyzing the available data, the study base period 1987 to 2014 was determined.

Table 1. Selected fluviometric stations information.

Code	ID	Station	X (m)	Y (m)	Zone	Drainage Area (km^2)
56425000	E1	Fazenda Cachoeira D'Antas	743325	7766305	23 K	10,100
56539000	E2	Cachoeira dos Óculos - Montante	764419	7762315	23 K	15,900
56719998	E3	Belo Oriente	775691	7760371	23 K	24,200
56850000	E4	Governador Valadares	189099	7753279	24 K	40,500
56920000	E5	Tumiritinga	221838	7752450	24 K	55,100
56994500	E6	Colatina	329007	7749346	24 K	76,400

2.3. Missing data patterns

According to Silva (2012), the presence of missing data in a database can be characterized by observed failure behavior patterns which is of paramount importance to describe missing value locations in the series. First, Collins *et al.* (1991) described and divided the pattern of missingness into two groups: *general* (random) and *special* patterns. Special patterns including univariate missing data, unit nonresponse, and monotone missing data. “Missing general” or “random pattern” is where missing data occur in any of the variables in any position. In a “special pattern” case, if there is only one variable with missing data while the other variables are completely recorded, the pattern is called “univariate missing data”. Additionally, when the multivariate pattern is detected, means that the missing values arise in more than one variable; if there are missing values on a block of variables for the same set of cases, and the remaining of the variables are all complete, the missing data pattern is called “unit nonresponse”; and, the pattern is said to be “monotone” whenever observations are ordered and item k is missing, and all $k + 1,..., n$ cases are also missing (Collins *et al.*, 1991; Silva, 2012). For the present study, the general pattern will be considered, because it assumes that the missing data follow a random distribution.

2.4. Missing data mechanisms

Little and Rubin (2019) classified three possible ways that data may go missing: Missing Completely at Random (MCAR), Missing at Random (MAR) and Missing Not at Random (MNAR). As discussed by Hamzah *et al.* (2020) and Gao *et al.*, (2018), MCAR describes data where the gaps are distinct from any of the variables in the dataset. In any event, the missing values probably correlated to other observed values, yet not to missing values; in that case, the missingness is assumed to be MAR. Missing data which are dependent on the observed value is known as MNAR. Missing data can be presented in the form of a probabilistic process that describes the association among the measured variables and the probability of missing value. For more details about Missing Data Mechanisms, see Little and Rubin (2019); based on these authors, missing value in streamflow study is determined as MCAR because the missingness episode in streamflow data of an area is not influenced by data in that area or any area. This mechanism of missing data is classified as ignorable, that is, there isn't need to model the mechanism as part of the estimation process (Allison, 2001).

2.5. Missing data imputation methods

The missing data imputation methodologies can be classified in three general classes: Single Imputation (SI), Multiple Imputation (MI) and Estimation. The basic principle of SI methods is to impute one value to each database missing data and, then, analyze it as if there weren't missing data (McKnight *et al.*, 2007). The MI is characterized by being a method that consists of imputing a certain missing data by a data set and, after analyzing it, determining the best value to be taken to impute it. The Estimation method consists in estimating parameters that govern the missing values distribution from the observed data. For the present study, the Maximum Likelihood (ML) methodology will be employed. In addition to this general classification and more widely used in the literature to deal with missing data, imputation methodologies can also be classified as univariate or multivariate. The first occurs when the series itself provides information that can be used by imputation methodology. The second one is when it is necessary to use support stations to impute interest series. In this study, AM, M, SPLINE, STINE and KALMAN were classified as univariate imputation methods. In turn, SLR, MLR, RW, MI and ML are multivariate methods.

2.5.1. Single Imputation Methods

Arithmetic Mean (AM)

This is the most commonly used single imputation technique where the missing values are replaced with the mean value of the time series. The mean of a series of values $x_1, x_2, ..., x_n$ is given by Equation 1.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

Median (M)

Median imputation is another simple method often appropriate for highly skewed data (Junger and Leon, 2015). This method calculates the median of the variable based on all cases that have data for any variable and replaces the series missing values with the median of the variable (Kabir *et al.*, 2019). The median of a series of values $x_1, x_2, ..., x_n$ is can be obtained by Equation 2 after sorting the dataset in ascending order:

$$m(x) = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & if\ n\ an\ even\ number \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & if\ n\ an\ odd\ number \end{cases} \qquad (2)$$

Simple (SLR) and Multiple Linear Regression (MLR)

Imputation with linear regression uses information from complete series to fill the missing values of the interest series. It can be of two types: (i) imputation by simple linear regression, when the purpose is to predict the dependent series behavior as a function of only one independent series (Equation 3) or (ii) imputation by multiple linear regression, when the dependent series behavior is a function of more than one independent series (Equation 4).

$$X = \beta_0 + \beta Y + \varepsilon \qquad (3)$$

In Equation 3, X represents the linear regression equation dependent series; β_0 a linear coefficient vector; β the angular coefficient; Y the independent station and ε the model residuals. As in SLR case, only one independent series is used to adjust the regression, Pearson correlation coefficient was used as a criterion for choosing this series and, after the test was applied, the station time series that presented highest correlation coefficient with E6, which is the interest station, was then chosen to perform imputation.

$$X = \beta_0 + \sum_{i=1}^{n} \beta_i Y_i + \varepsilon \qquad (4)$$

In Equation 4, X represents the linear regression equation dependent series; β_0 a linear coefficient vector; β_i the angular coefficients; Y_i the independent station and ε the model residuals. For MLR, stations E1, E2, E3, E4 and E5 series were used as independent series to fit the regression with E6.

Regional Weighting (RW)

Imputation by regional weighting method establishes linear regressions between the station that has series with missing data, *X*, and each neighboring station series, Y_1, Y_1, ..., Y_n, incorporating distance information (Mello *et al.*, 2017; Pruski *et al.*, 2004). From each linear regression performed, a correlation coefficient is obtained for the data to be estimated. Equation 5 denotes the regional weighting method.

$$X = \frac{1}{n}\sum_{i=1}^{n}\frac{N_X}{N_i}D_i \tag{5}$$

In Equation 5, N_X and Ni represents the monthly average flows data for the station with missing data to be imputed and the order "i" neighboring station monthly average flow, respectively ($m^3.s^{-1}$); D_i denotes the values observed in the order "i" neighboring stations during the month of occurrence in the station with the data to be imputed ($m^3.s^{-1}$); and n is the number of neighboring stations considered.

Spline Interpolation (SPLINE)

According to Wijesekara and Liyanage (2020), for $n + 1$ pair of observations $\{(t_i, x_i): i = 0,1, \dots n\}$, the shape of spline is modeled by interpolating between all the pairs of observations (t_{i-1}, x_{i-1}) and (t_i, x_i) with polynomials described in Equation 6.

$$x = q_i(t), i = 1,2, \dots, n \tag{6}$$

Stineman interpolation (STINE)

This is an advanced interpolation method where interpolation occurs based on (i) whether values of the specified points ordinates change monotonically and (ii) the slopes of the line segments joining specified points change monotonically (Turicchi *et al.*, 2020).

Kalman Smoothing (KALMAN)

The Kalman filter calculates the mean and variance of the unobserved state, given the observations. This filter is a recursive algorithm; the current best estimate is updated whenever a new observation is obtained. Kalman Smoothing takes the form of a backwards recursion and it can be used to compute a smoothed estimator of the disturbance vector (Wijesekara and Liyanage, 2020).

In the present work, R package *imputeTS* was used for Spline Interpolation, Stineman Interpolation and Kalman smoothing imputations.

2.5.2. Multiple Imputation

In an attempt to develop a method that could reflect uncertainty over missing data imputations, Rubin (1987) created the MI method, in which each missing value is replaced by a set of plausible values representing this uncertainty about the value to be imputed. MI consists of the following three steps: (i) m complete databases Y_{obs}, Y_{mis} are obtained through appropriate imputation techniques; (ii) separately, m data banks are analyzed by a traditional statistical method, as if they were indeed complete data groups; (iii) the m results found in step (ii) are combined in a simple and suitable way for obtaining the so-called repeated imputation inference.

In this research, the imputation model used (i) is adjusted by the Bayesian Paradigm: from the result of the posterior distribution, a set of random extractions is made for the missing data from the observed data, thus obtaining the complete database, that is, multiple imputations are made using the linear regression method ($Y = \alpha + \beta X$), Y ~ N (Xβ; I σ^2), where the response variable Y will be the variable to be imputed for which the parameters are estimated from its own posterior distribution. The predicted values for Y_{obs} and Y_{mis} are calculated and, for each predicted Y_{mis}, the observed unit with the closest predicted value is sought using it as the value to be imputed. The variability between imputations is generated through the steps used to estimate β and σ and which are repeated m times; in the next step (ii) Q of each imputed data set is estimated; finally, in step (iii), the m results obtained can be combined using the rules proposed by Rubin (1987). The idea is that from each analysis the estimates for the parameter of interest Q are obtained, that is, Q_i for i = 1,2,3 ..., m. According to Schafer (1999), Q can be any scalar measure to be estimated, such as mean, correlation, regression coefficient or odds ratio. Then, the combined estimate will be the average of the individual estimates $\bar{Q} = \frac{1}{m}\sum_{i=1}^{m}\hat{Q}_i$. For the combined variance, the variance is first calculated within the imputations $\bar{U} = \frac{1}{m}\sum_{i=1}^{m}U_i$ and the variance between imputations $B = \frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{Q}_i - \bar{Q}\right)^2$. Then, the total variance, which is the combined variance, will be $T = \bar{U} + \left(1 + \frac{1}{m}\right)B$.

The literature standardizes m = 5 imputations, a value that, as indicated by Nunes (2007) emerged from the researcher's experiences. MI is implemented in the *mice* R software library, developed by Buuren and Groothuis-Oudshoorn (2010).

2.5.3. Maximum Likelihood

The ML method's basic principle is to choose as an estimation of parameters those values that, if true, would maximize the probability of observing what was actually observed (Allison, 2001). The parameters estimation is performed by maximizing the likelihood function, which, in most cases, can't be performed analytically. It is necessary to employ numerical methods, such as the EM algorithm. This maximization method is very popular when the data considered in the estimation aren't complete (Allison, 2001).

The EM algorithm is an iterative procedure consisting of two steps repetition: Estimation (E) and maximization (M). The process begins with mean vector and covariance matrix estimation by using only the complete data. According to Junger and Leon (2015), being $\boldsymbol{x}_t, (t = 1, \dots, m)$, the t^{th} random vector $\boldsymbol{X}$ realization, with multivariate normal distribution and m components not observed. The vector $\boldsymbol{x}_t$ can be arranged in such a way that the m missing components are placed in the first positions, that is, $\boldsymbol{x}_t = \left(x_{t1}, \dots, x_{tm}, x_{t(m+1)}, \dots, x_{tp}\right)^T$, and represented as $\boldsymbol{x}_t = (\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2})^T$. Consider B windows with different covariance regimes over time. The estimation of the mean vector at the instant t and window *b, (b = 1, ..., B)*, can be partitioned following the same configuration as that corresponding to $\boldsymbol{x}_t$ components, according to equations 7 and 8.

$$\tilde{\mu}_t = \begin{bmatrix}\tilde{\mu}_{t1}\\ \tilde{\mu}_{t2}\end{bmatrix} \tag{7}$$

$$\textstyle\sum_b = \begin{bmatrix}\widetilde{\boldsymbol{\Sigma}}_{b11} & \widetilde{\boldsymbol{\Sigma}}_{b12}\\ \widetilde{\boldsymbol{\Sigma}}_{b21} & \widetilde{\boldsymbol{\Sigma}}_{b22}\end{bmatrix} \tag{8}$$

The imputation algorithm consists of (i) replace the missing values for estimates values; (ii) estimate the normal model parameters μ and $\boldsymbol{\Sigma}$ and each univariate time series μ_t level; (iii) re-estimate the missing values considering just the updated parameters and each time series level. This process is repeated until the estimated values stop to vary (Junger and Leon, 2015). Junger and Leon (2015) report that initial estimates $\tilde{\mu}_0$ and $\widetilde{\boldsymbol{\Sigma}}_0$ are, respectively, the mean vector and the sample covariance matrix, considering only the observed data. In iteration $(k+1)$ of the step E for the EM algorithm, the missing values are imputed as the conditional mean to the observed values and the parameters estimated in the previous interaction appropriated using equation (9), with the contributions to covariance estimated by Equations 10 and 11.

$$\tilde{\boldsymbol{x}}_{t1}^{(k+1)} = E\left[\boldsymbol{X}_{t1} \middle| \boldsymbol{x}_{t2}, \tilde{\mu}_t^{(k)}, \tilde{\boldsymbol{\Sigma}}_b^{(k)}\right] = \tilde{\mu}_{t1}^{(k)} + \tilde{\boldsymbol{\Sigma}}_{b12}^{(k)} \tilde{\boldsymbol{\Sigma}}_{b22}^{(k)-1}\left(\boldsymbol{x}_{t2} - \tilde{\mu}_{t2}^{(k)}\right) \tag{9}$$

$$\widetilde{\boldsymbol{x}_{t1}\boldsymbol{x}_{t1}^T}^{(k+1)} = E\left[\boldsymbol{X}_{t1}\boldsymbol{X}_{t1}^T \middle| \boldsymbol{x}_{t2}, \tilde{\mu}_t^{(k)}, \tilde{\boldsymbol{\Sigma}}_b^{(k)}\right] = \tilde{\boldsymbol{\Sigma}}_{b11}^{(k)} - \tilde{\boldsymbol{\Sigma}}_{b12}^{(k)} \tilde{\boldsymbol{\Sigma}}_{b22}^{(k)-1} \tilde{\boldsymbol{\Sigma}}_{b21}^{(k)} + \tilde{\boldsymbol{x}}_{t1}\tilde{\boldsymbol{x}}_{t1}^T \tag{10}$$

$$\widetilde{\boldsymbol{x}_{t1}\boldsymbol{x}_{t2}^T}^{(k+1)} = E\left[\boldsymbol{X}_{t1}\boldsymbol{X}_{t2}^T \middle| \boldsymbol{x}_{t2}, \tilde{\mu}_t^{(k)}, \tilde{\boldsymbol{\Sigma}}_b^{(k)}\right] = \tilde{\boldsymbol{x}}_{t1}\tilde{\boldsymbol{x}}_{t2}^T \tag{11}$$

According to Junger and Leon (2015), in step M the μ_b and $\boldsymbol{\Sigma}_b$ reviewed maximum likelihood estimates are computed, considering implicit the interaction index $(k+1)$, $\tilde{\mu}_b = \sum_{t=1}^{n_b} \frac{\tilde{x}_{bt}}{n_b}$ and $\tilde{\boldsymbol{\Sigma}}_b = \sum_{t=1}^{n_b} \frac{\widetilde{x_{bt}x_{bt}^T}}{n_b} - \tilde{\mu}_b\tilde{\mu}_b^T$. The $\tilde{\mu}_b$ estimate is used only for $\tilde{\boldsymbol{\Sigma}}_b$ calculation.

Junger and Leon (2015) emphasize the need to use additional models to estimate the contribution of the time component for each univariate series, that is, μ_t value. The temporal trajectory of the series considered in this study was modeled with use of cubic non-parametric splines because, according to the same authors, this trajectory was the one that presented the best performance in relation to the regression models and *ARIMA* (Autoregressive Integrated Moving Average) models for air pollution time series missing data imputation.

Therefore, consider that μ_t can be estimated by a smooth function g_j with $j = 1, ..., p$. The curve g_j, in turn, is estimated in such a way that the functional $S(g_j) = \sum_{k=1}^{K}[X_t - g(v_k)]^2 + \lambda\int_a^b (g'')^2\,dx$ be minimized. The points $v_1, v_2, ..., v_k$, ordered in the interval [a, b], are the nodes and λ is the curve smoothing parameter (Junger and Leon, 2015). This results in a natural cubic spline (Green and Silverman, 1993). Each variable X_J has its level given by $\mu_{jt} = g(X_{jt})$.

In this paper, the EM algorithm proposed by Junger (2008) was used, which makes use of the *mtsdi* platform in the R software, which was implemented by the author.

2.6. Performance Indexes

For the performance evaluation of different missing data, imputation methods in effecting the estimates were employed as can be seen below (Equations 12, 13, 14, 15 and 16).

BIAS

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \tilde{x}_i) \tag{12}$$

Root Mean Square Error (RMSE)

$$\frac{1}{N}\sqrt{\sum_{i=1}^{N}(x_i - \tilde{x}_i)^2} \tag{13}$$

Mean Absolute Percentage Error (MAPE)

$$\frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - \tilde{x}_i}{x_i}\right| \times 100 \tag{14}$$

Coefficient of Determination (R^2)

$$\frac{1}{N}\sum_{i-1}^{N}\frac{\sum_{i=1}^{N}[(x_i - \bar{x})(\tilde{x}_i - \bar{\tilde{x}})]}{\hat{\sigma}(x_i)\hat{\sigma}(\tilde{x})} \tag{15}$$

Concordance Index (d_2)

$$1-\left[\frac{\sum_{i=1}^{N}(x_i-\tilde{x}_i)^2}{\sum_{i=1}^{N}(|x_i-\bar{x}|+|\tilde{x}_i-\bar{x}|)^2}\right] \tag{16}$$

The BIAS measure quantifies underestimation and overestimation estimates with respect to the average observations (Bier and Ferraz, 2017; Junger, 2008); RMSE and MAPE are accuracy estimates measures. In equations 12-16, N represents the number of missing data in the modeled data set, x_i the observed data, $\tilde{x}_i$ the imputed data, $i = 1, \dots, m$, $\bar{x}$ the observed values mean and $\bar{\tilde{x}}$ the imputed data mean (Pinto, 2013).

2.7. Application

From the Doce River monthly average flow-time series, an algorithm was created to simulate five incomplete data banks, with 5%, 10%, 15%, 25% and 40% missing data. The routines for simulating the missing data percentages, imputations and analyzes were implemented by using the R software (R Development Core Team, 2018). The website for that software access is *http://www.r-project.org*.

3. RESULTS AND DICUSSION

For a better understanding of the studied average monthly flow variable, each station's descriptive statistics are presented in Table 2. It's observed that, during the study period, the monthly average flow rates ranged from 162.48 $m^3.s^{-1}$ in E1 to 797.77 $m^3.s^{-1}$ in E6. These values represent the lowest and highest average monthly flows, considering all stations. A high standard deviation and coefficient of variation values can be observed, an aspect that allows us to conclude that the means aren't representative, a fact that, according to Bayer *et al.* (2012), can be associated with the large data intra-annual variability, indicating a seasonal component. It is also noted that the observed extreme monthly average flow values, maximum and minimum, were 3,528.32 and 41.58 $m^3.s^{-1}$, at Stations E6 and E1, respectively. The first was registered on January 12, 2013, while the second was observed on October 1, 2014.

The fluviometric stations' monthly average flow distributions present positive asymmetry values ranging from 1.79 to 2.04. For kurtosis, the variation was from 3.24 to 4.63. The mean values are greater than the respective median and mode values, and the data distribution values' degree of concentration is classified as leptokurtic, according to Fonseca and Martins (2011).

Figure 2 shows the six stations' average monthly flow-time evolution. As can be observed, the time series present intra-annual variability pattern, with floods periods followed by drought periods, which characterizes seasonality property, already indicated by the descriptive statistics and confirmed in Pinto *et al.* (2015) studies for E6 monthly mean flows and by Bleidorn *et al.* (2018) studies, for minimum monthly flows for all stations considered in this study. It's relevant to note that both studies suggest a seasonal component with a 12-month period, with a flood period from November to May, and drought period from June to October.

Table 2. Average monthly flow variable descriptive measures.

Descriptive measures	E1	E2	E3	E4	E5	E6
Mean ($m^3.s^{-1}$)	162.48	214.60	325.18	524.07	646.85	797.77
Median ($m^3.s^{-1}$)	132.74	171.45	256.43	395.25	499.78	554.74
Standard Deviation ($m^3.s^{-1}$)	92.34	129.15	204.40	345.43	418.09	604.33
Coefficient of Variance (%)	56.83	60.18	62.85	65.91	64.63	75.75
Minimum Value ($m^3.s^{-1}$)	41.58	52.49	81.63	112.20	131.03	194.20
Maximum Value ($m^3.s^{-1}$)	616.22	827.27	1,219.04	2,051.50	2,391.05	3,528.32
Kurtosis	4.03	3.58	3.80	3.50	3.24	4.63
Asymmetry	1.79	1.74	1.77	1.78	1.77	2.04

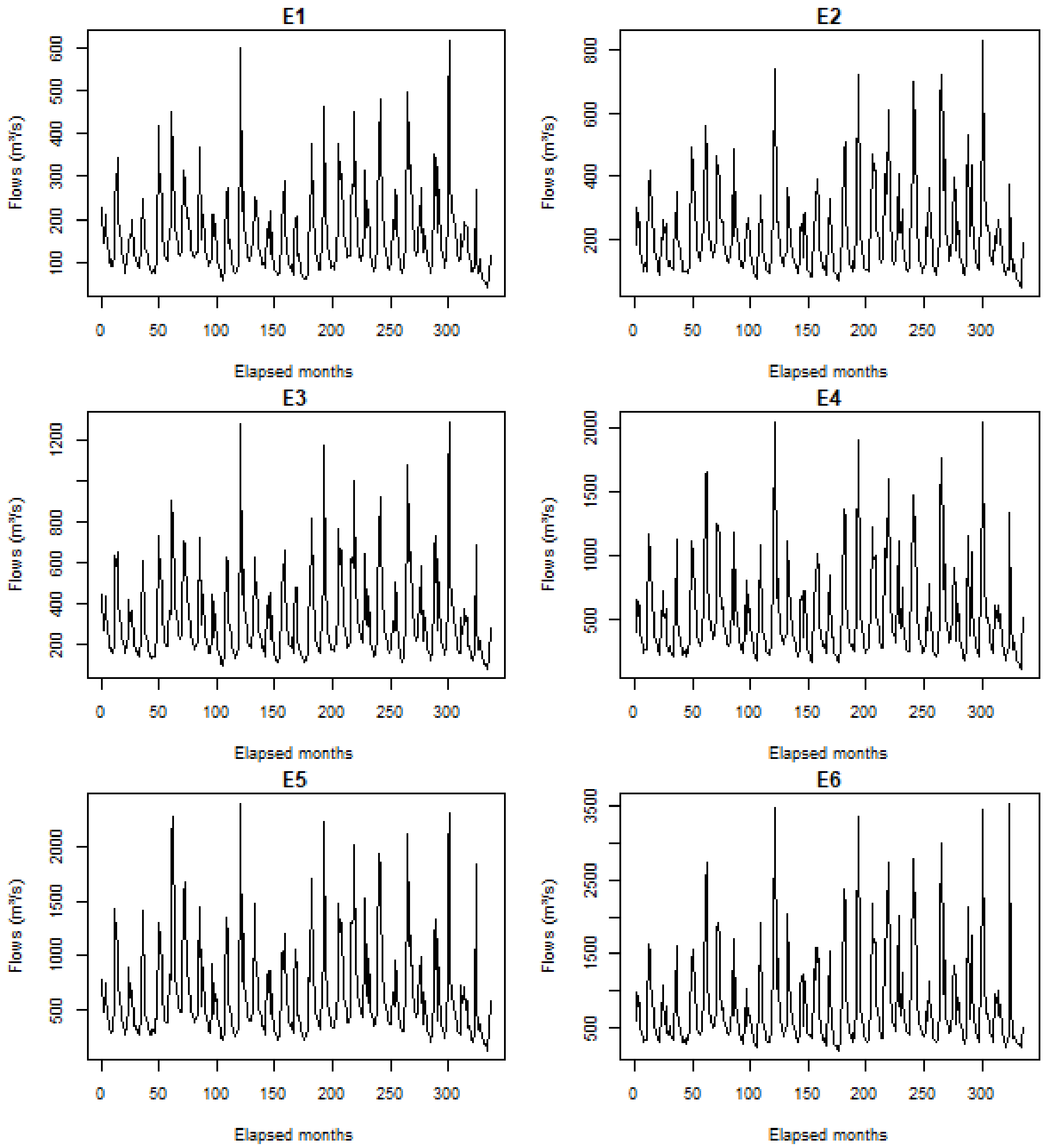


Figure 2. Fluviometric stations average monthly flow time series.

According to the Fischer *F* and *t*-Student tests, the time series under study are homogeneous, that is, they do not present variances and mean changes over time, considering five-year analysis periods. This fact is important to validate studied stations' data consistency. The imputation methods considered in this work require that the data follow a Normal distribution; however, according to Shapiro and Wilk (1965) and Bera and Jarque (1981) tests, the data normality null hypothesis is rejected. According to Junger (2008) and Sabino *et al.* (2014), environmental data, commonly, do not follow Normal distribution. Thus, statistical tests are performed to check the application of a transformation that could stabilize the variance of the observations. In the case under study, the smoothing parameter λ was estimated, as suggested by Box and Cox (1964) to define the type of transformation. According to Reisen *et al.* (2008), often the transformation not only stabilizes the variance but also improves the data distribution approximation to the Normal distribution. Thus, all imputations were performed using the natural logarithm transformation in order to improve the approximation with the

Normal distribution and to stabilize the variance. The imputed data was back transformed for subsequent analysis.

Linear Pearson correlation coefficients for the average monthly flows *r* between stations under study indicate a highly homogeneous distribution, because among the pairs of stations the lowest value was 0.9240, as shown in Table 3. This condition was expected, because the stations E1 and E6 present the highest distance between stations identified in the study area, which suggests that better performance can be achieved by using the multivariate imputation methods (SLR, MLR, RW, MI and ML).

Table 3. Average monthly flow data: Pearson correlations between stations.

	E1	E2	E3	E4	E5	E6
E1	1					
E2	0.9875	1				
E3	0.9797	0.9869	1			
E4	0.9604	0.9698	0.9897	1		
E5	0.9419	0.9524	0.9762	0.9928	1	
E6	0.9240	0.9368	0.9616	0.9782	0.9794	1

In order to validate the missing data imputation methods considered in this study, the E6 database underwent 1000 replications for each failure ratio. Table 4 shows the performance-indicator means for the missing data imputation methods. In general, there is a decreasing imputation quality gradient presented by the performance indexes (PI) because of missing data increase. The SI (AM, M, SLR, MLR and RW) methods showed considerably low BIAS values and high RMSE and MAPE values in relation to SI (SPLINE, STINE e KALMAN) and mainly to multivariate attribution (MI and ML) methods. SI (AM, SLR, MLR and RW) methods showed a considerable increase in BIAS when missing data proportion increased, making imputed series underestimated in relation to observed one. M showed a similar behavior, however, with lower intensity. The increase in bias values was observed to a lesser extent for SI methods (SPLINE, STINE and KALMAN) and mainly for MI and ML, suggesting a small imputed data variance loss in relation to observed ones. In addition, the methods that lost most quality for R^2 and d_2 indicators were SI (AM, M, SLR, MLR and RW) followed by SI (SPLINE, STINE and KALMAN) and with little loss for MI and ML.This confirms MI and ML methods good performance, being that the last corroborates with the efficiency proven by Junger (2008) and reforced by Burger *et al.* (2018).

Both AM and M methodologies are central tendency measures. However, the second is a better alternative for variables that do not follow a Normal distribution, that is, the median better represents the central tendency of a distribution that presents large deviations from the Normal distribution (McKnight *et al.*, 2007; Pinto, 2013). Therefore, for variables that do not follow a Normal distribution, such as the river average flow rate, under study, the imputation methodology by AM is not efficient, especially for faults above 5% even under natural logarithm transformation. This conclusion was also found in Pinto (2013), where the detrimental effect of this method on the imputation of the variable PM_{10} (Inhalable Particulate Matter with aerodynamic diameter less than 10 µm) is reduced only in samples with a small percentage of missing data (5%). The efficiency of the imputation methodology by M decreases slightly for faults above 10%. Hence, this method is not indicated for these situations.

Table 4. Performance indicators for missing data imputation methodologies.

Faults	PI	AM	M	SPLINE	STINE	KALMAN	SLR	MLR	RW	M1	ML
5%	BIAS	0.0169	0.0098	0.0001	0.0004	0.0001	0.0187	0.0214	0.0169	0.0001	<0.0001
	RMSE	0.0084	0.0077	0.0045	0.0039	0.0039	0.0049	0.0057	0.0043	0.0020	<0.0001
	MAPE	0.3870	0.3725	0.1939	0.1649	0.1684	0.2877	0.3305	0.2598	0.0999	0.0737
	R	0.9668	0.9719	0.9905	0.9927	0.9927	0.9889	0.9850	0.9913	0.9982	0.9990
	d_2	0.9829	0.9856	0.9952	0.9963	0.9963	0.9941	0.9920	0.9954	0.9991	0.9995
10%	BIAS	0.0667	0.0221	0.0001	0.0007	0.0001	0.0703	0.0762	0.0671	0.0002	<0.0001
	RMSE	0.0154	0.0111	0.0066	0.0057	0.0057	0.0124	0.0138	0.0117	0.0028	<0.0001
	MAPE	1.0273	0.7502	0.4083	0.3397	0.3473	1.0846	1.1748	1.0350	0.2010	0.1489
	R	0.8974	0.9425	0.9800	0.9851	0.9850	0.9404	0.9273	0.9469	0.9963	0.9979
	d_2	0.9452	0.9698	0.9898	0.9924	0.9923	0.9653	0.9577	0.9693	0.9982	0.9989
15%	BIAS	0.1504	0.0413	0.0002	0.0014	0.0003	0.1540	0.1623	0.1498	0.0003	<0.0001
	RMSE	0.0246	0.0142	0.0085	0.0072	0.0072	0.0220	0.0236	0.0211	0.0035	<0.0001
	MAPE	2.2230	1.1391	0.6391	0.5228	0.5347	2.3717	2.4979	2.3092	0.3054	0.2247
	R	0.7811	0.9080	0.9675	0.9764	0.9763	0.8469	0.8268	0.8575	0.9944	0.9968
	d_2	0.8754	0.9505	0.9834	0.9879	0.9878	0.9022	0.8892	0.9087	0.9972	0.9984
25%	BIAS	0.4099	0.0929	0.0004	0.0020	<0.0001	0.4152	0.4284	0.4101	0.0005	<0.0001
	RMSE	0.0475	0.0196	0.0120	0.0099	0.0100	0.0456	0.0477	0.0448	0.0047	<0.0001
	MAPE	6.1715	1.9624	1.1655	0.9337	0.9602	6.4132	6.6123	6.3317	0.5232	0.3852
	R	0.5245	0.8285	0.9357	0.9548	0.9542	0.6210	0.5986	0.6290	0.9900	0.9944
	d_2	0.6966	0.9033	0.9668	0.9764	0.9759	0.7204	0.7050	0.7254	0.9949	0.9971
40%	BIAS	1.0423	0.2412	0.0001	0.0032	0.0003	1.0485	1.0652	1.0423	0.0008	0.0002
	RMSE	0.0921	0.0297	0.0174	0.0140	0.0141	0.0909	0.0931	0.0901	0.0070	<0.0001
	MAPE	15.8970	3.8042	2.1929	1.6855	1.7512	16.1982	16.4453	16.098	0.9293	0.6654
	R	0.2729	0.6580	0.8655	0.9087	0.9062	0.3686	0.3539	0.3723	0.9781	0.9885
	d_2	0.5526	0.7902	0.9289	0.9511	0.9489	0.5617	0.5557	0.5635	0.9888	0.9941

In the hydrology area, the idea of using a single imputation method to recover the missing value (mainly AM and M) is widely used by several authors, as can be seen Ben Aissia *et al.* (2017); Gao *et al.* (2018); Kabir *et al.* (2019); Norliyana *et al.* (2017) and Rahman *et al.* (2017) works. However, despite the method's simplicity the researchers agree that reconstructing the missing value using the same "number" does not reflect the variation that would likely occur if the variables were observed. The real numbers possibly differ from those imputed. Thus, the variation of those same variables is underestimated.

SLR methodology presents better results in relation to MLR according to all performance indicators and regardless of the number of failures. MLR involves more independent variables to predict an adjustment model and, theoretically, its complexity results in a higher efficiency than the corresponding SLR. However, for the SLR methodology, an algorithm was developed in which the independent station was taken as a function of the correlation coefficient *r* higher value and, thus, alternate stations were candidates according to the replication progress. For the imputation methodology via MLR, based on the correlation coefficient *r* (which assumed values greater than 0.924 between the station pairs), all the other stations (E1, E2, E3, E4 and E5) were considered as adjusted model independent variables, for the dependent variable E6. This justifies the better performance of SLR over MLR in this study. The imputation methodology via RW, which is widely used in hydrological variables series imputation studies, showed superior efficiency than the corresponding to other SI (AM, M, SLR and MLR) methods only for the 5% failure rate and presented less efficiency than those corresponding to SI (SPLINE, STINE and KALMAN) in addition to MI and ML methods, independently of the missing data percentage.

Oliveira *et al.* (2010) concluded that the MLR method showed, according to the performance indexes BIAS, MAE and *r*, the best results in annual rainfall faults imputation for six stations located in the state of Goiás, Brazil, followed by vector regional combined with multiple potential regression (MPR), RW, regional vector combined with multiple linear regression, regional weighting based on linear regression, regional vector combined with regional weighting and, lastly, regional vector method. This result corroborates those found by (i) Ventura *et al.* (2016), that, considering BIAS, MAE and *r*, verified the superiority of the MLR method for temperature, relative humidity and dew point meteorological variables, considering Manaus, Rio de Janeiro and Porto Alegre Brazilian cities stations series; (ii) those found by Mello *et al.* (2017) that, according to the MAE, for precipitation series corresponding to eight rainfall stations located in Santa Catarina state, Brazil, the MLR method was the one that presented the best performance, followed by RW, regional weighting based on linear regressions and, finally, SLR; and (iii) those found by Bier and Ferraz (2017) who concluded that, for average compensated temperature for meteorological stations located in Rio Grande do Sul state, Brazil, the MLR and the RW methods were the most adequate for missing data estimates, whereas for precipitation there was no method that could be considered best. According to the same authors, this was due to the fact that neighboring stations' precipitation data was less correlated if compared to average temperature compensated data, generating estimates less related to the observed series and presenting larger estimative errors.

Low BIAS, RMSE and MAPE values and high R^2 and d_2 values confirm univariate single imputation methods (SPLINE, STINE and KALMAN) good performance, however, inferior to MI and ML multiple multivariate imputation methods. Additionally, it is possible to verify a slight superiority of imputation quality via ML methodology in relation to MI, regardless of missing data proportion. Figures 3 and 5 show the simulated results for 5 and 40% missing data proportions and Figures 4 and 6 show a visual comparison between observed and imputed data. This comparison shows good performance by all methodologies for the 5% missing data proportion. Therefore, for this smaller data failure percentage, any imputation methodology could be assumed without causing significant changes in the time series characteristics. Thus, caution is suggested in the use of AM, even for small failure proportions. However, for the 40% proportion, the analysis shows SI methods (SPLINE, STINE and KALMAN) good performance and superiority of MI and ML methods, while for other SI methodologies (AM, M, SLR, MLR and RW) imputed data variability underestimation is notable. Therefore, it can be concluded that the use of such methodologies is not recommended to carry out imputation with high percentages of missing data, as the time series characteristics can be extremely altered.

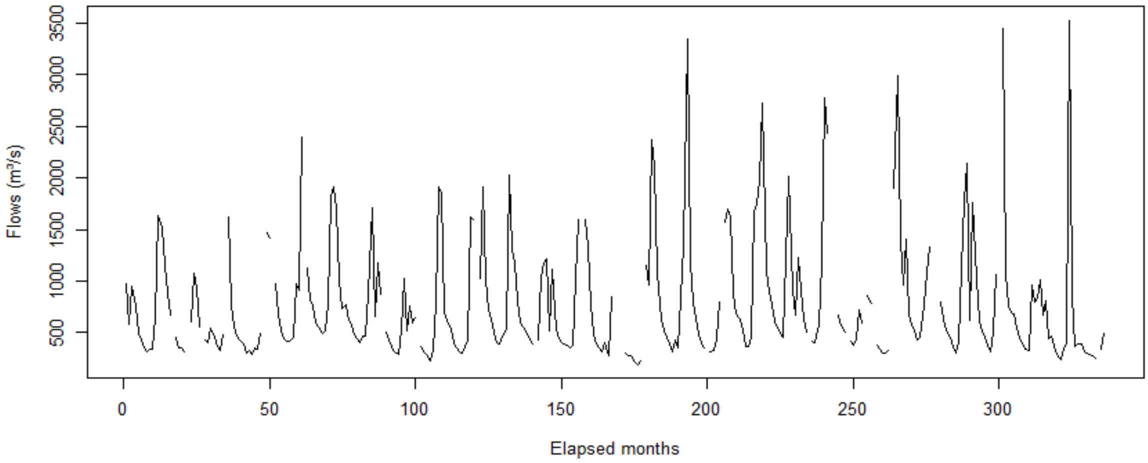


Figure 3. Average monthly flow time series for the E6 station, considering 5% missing data.

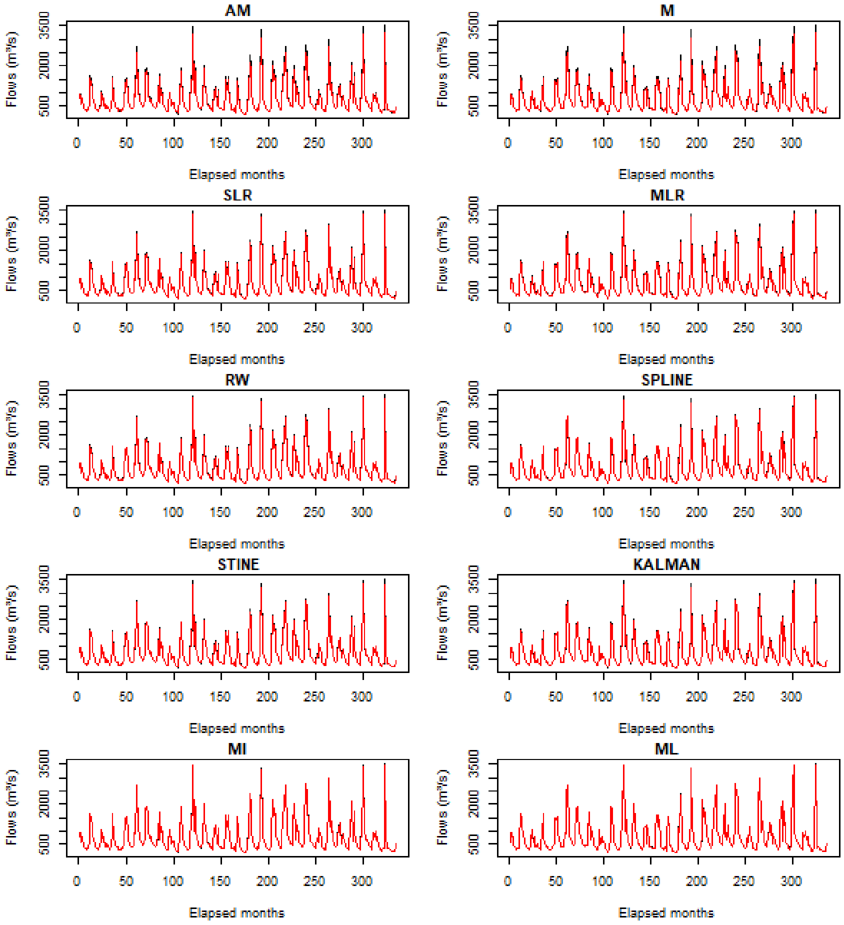


Figure 4. Observed (black line) and imputed (red line) average monthly flows time series for the E6 station, for each imputation methodology, considering 5% missing data.

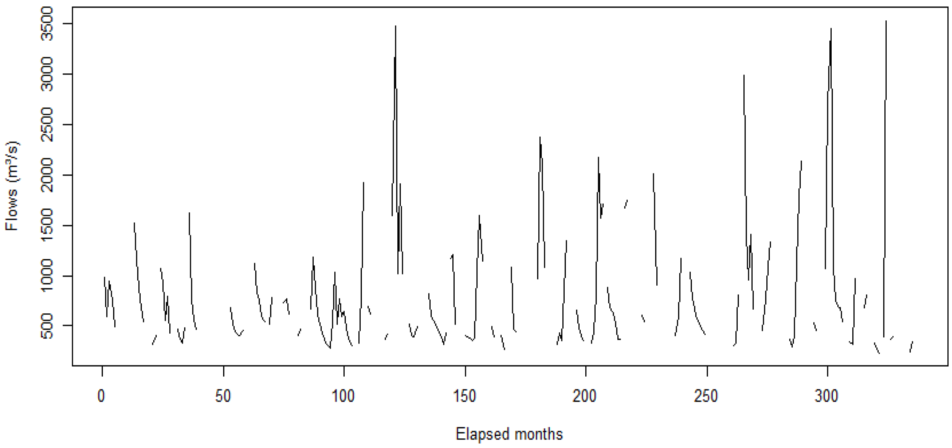


Figure 5. Average monthly flow time series for the E6 station, considering 40% missing data.

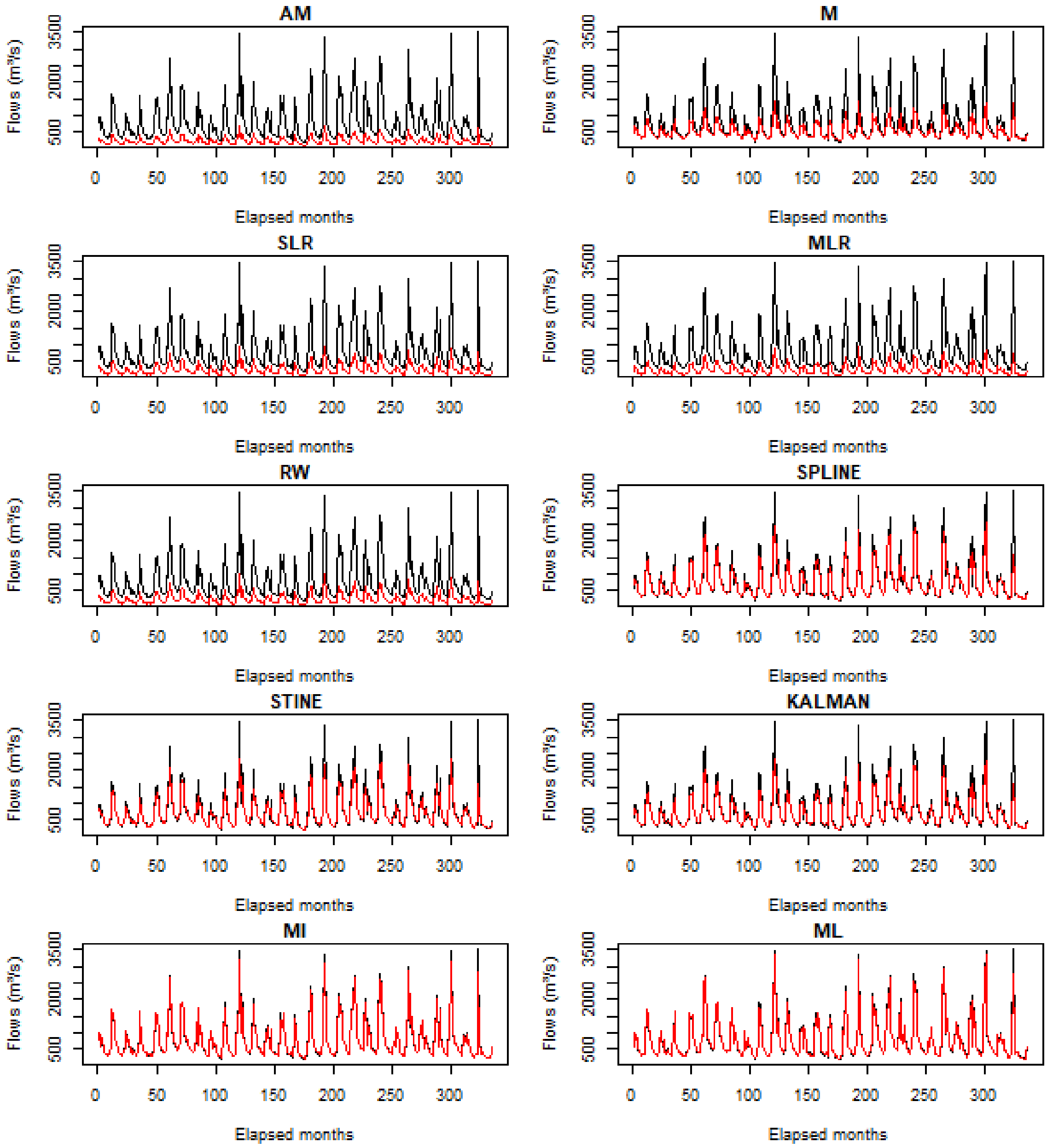


Figure 6. Observed (black line) and imputed (red line) average monthly flows time series for the E6 station, for each imputation methodology, considering 40% missing data.

Table 5 shows relative difference results between the quality indicators used in imputation methods evaluation in this study for Doce River flow series in a scenario where missing data percentage increased from 5% to 40%. As can be seen, all methodologies tested in this study showed a loss of quality with the missing data proportion increase and, based on that, the table was organized in order to order the methodologies compared in this study from the one with greatest to the least loss of quality according to indicators used. From this, we can conclude that, in cases where there are support stations for imputation, the best methodologies for imputing missing data in flow data are the multivariate ones MI and ML, while, in cases where there are no support stations for imputation, the best options of imputation methodologies to be used are the univariate ones Kalman Smoothing and Stineman interpolation.

Table 5. Relative difference in the imputation method quality when failures increased from 5% to 40%.

Imputation Methodology	BIAS	RMSE	MAPE	R^2	d_2
AM	-98.38%	-90.88%	-97.57%	-71.77%	-43.78%
MLR	-97.99%	-93.88%	-97.99%	-64.07%	-43.98%
SLR	-98.22%	-94.61%	-98.22%	-62.73%	-43.50%
RW	-98.38%	-95.23%	-98.39%	-62.44%	-43.39%
M	-95.94%	-74.07%	-90.21%	-32.30%	-19.83%
SPLINE	0.00%	-74.14%	-91.16%	-12.62%	-6.66%
KALMAN	-66.67%	-72.34%	-90.38%	-8.71%	-4.76%
STINE	-87.50%	-72.14%	-90.22%	-8.46%	-4.54%
MI	-87.50%	-71.43%	-89.25%	-2.01%	-1.03%
ML	-50.00%	0.00%	-88.92%	-1.05%	-0.54%

Based on these results, it can be said that, multivariate MI and ML methodologies application in river flow rates missing data imputation, confirmed their good performance, already proven for imputations considering other variables (Junger, 2008; Pinto, 2013; Nunes *et al.*, 2010; Vinha and Laros, 2018), as occurred in cases of application univariate methodologies KALMAN and STINE whose good performance had also been confirmed to handle missing data in exchange rate data (Burger *et al.*, 2018). Therefore, it is suggested that both methodologies should be taken as quality references for fluviometric variables missing data imputation, especially the ML. Lastly, Table 6 shows the main advantages and disadvantages for each methodology evaluated in the present study.

In this way, fluviometric monitoring and historical series data consistency as well as the adequate missing data treatment are fundamental for water-resource studies and projects, turning possible adequate water-resource use planning, river basin management, flow forecasting, industrial and agricultural public supply, navigation, basic sanitation, water concession and granting, academic studies and water use conflicts resolution.

4. CONCLUSIONS

Reliable flow series are essential for studies related to water resources. However, incomplete series can result even from the operation of large monitoring networks with data quality control. This fact indicates the importance of techniques that allow imputation of missing data. However, these techniques have been little used in hydrological studies. In this study, it was verified that, for the Doce River flow series, the multiple imputation and maximum likelihood methodologies, considered as references in the imputation of missing data for series involving several environmental and social variables, performed better than those most commonly utilized. It was also concluded that the improvement in the results of the imputations in relation to the others increases as the flow series present greater proportions of the missing data.

Table 6. Advantages and disadvantages for each methodology evaluated in the present study.

Methodology	Advantages	Disadvantages
AM	Easy to implement	Decreases data variability
M	Easy to implement	Decreases data variability
SLR	Requires additional variables with an acceptable correlation	In situations under and overestimation conditions, the variability of the data decreases
MLR	Requires additional variables with an acceptable correlation	In situations under and overestimation conditions, the variability of the data decreases
RW	Requires additional variables with an acceptable correlation	Requires additional variables that have acceptable correlations. Tends to underestimate data variability
SPLINE	It's a nonlinear approach. Provides a "smooth" interpolant. Doesn't usually get "wiggly" like higher-order polynomial interpolation can	Has a limited ability to predict oscillations from univariate data. Requires a bit more work than linear interpolation to implement
STINE	Produces a imputation known to be robust against sporadic outliers and performs better than spline interpolations, where abrupt changes are observed; Solves the non-monotonic problem of linear and spline interpolation	Is not as smooth method as the linear and spline methods
KALMAN	It avoids the influence of possible structural breaks during the missing values estimation	Assumes a large knowledge of probabilistic theory, specifically Gaussian conditional properties of random variables, which may limit its study and application scope
MI	Incorporates the uncertainty in each database generated, reducing the standard error of the final imputed values	Requires statistical knowledge and computational sophistication
ML	Because is based on available data, extracts information from its behavior	Requires statistical knowledge and computational sophistication

Simulations showed that, for 5% missing data, all the imputation methods present good performances. For this small missing data proportion, statistical efficiency is not compromised. Even so, for this missing data amount, the arithmetic mean methodology, which presents the worse results, should be avoided. The imputation method quality begins to become different for proportions above 10%. The arithmetic mean, simple and multiple linear regression and regional weighting imputation methodologies may be considered limited, even though the last three methods are multivariate and under the condition that the stations showed high homogeneity among them. Therefore, multivariate methods that consider the single imputation principle were not efficient. On the contrary, univariate methodologies Spline interpolation, Stine interpolation and Kalman Smoothing being single methods showed good results in imputation at high proportions of missing data. Also, multiple imputation and maximum likelihood multivariate methods were shown to be more robust and accurate for missing data treatment of the variable under study, being recommended for failure imputation procedures. For this, it is fundamental that the variables' data present homogeneity among them. Therefore, such methods are recommended for the proper treatment of missing data, which allows us to guarantee quality in imputed time series, thus being able to subsidize the planning and control of water resources.

The method of imputation by multiple linear regression can be improved, since the adjustment of $n_i < 5$ neighboring stations were based on the correlation coefficient and, according to the replication progress of replication, it is possible that one or more stations do not present a qualified correlation coefficient. As for multiple imputation, it is concluded that its efficiency can be increased as an ideal *m* imputation value is established for the variable under study, considering different proportions of missing data. Therefore, it is recommended that future studies be developed because the procedure efficiency is related to the number of *m* databases generated and, in general, the higher the percentage of missing values, the greater is *m*. It is also recommended that studies be developed regarding the imputation of missing data for river flows in daily and monthly series.

5. REFERENCES

ABDELGAWAD, H.; ABDULAZIM, T.; ABDULHAI, B.; HADAYEGHI, A.; HARRETT, W. Data Imputation and Nested Seasonality Time Series Modelling for Permanent Data Collection Stations: Methodology and Application to Ontario. **Canadian Journal of Civil Engineering**, v. 42, n. 5, p. 287–302, 2015. https://doi.org/10.1139/cjce-2014-0087

ABU ROMMAN, Z.; AL-BAKRI, J.; AL KUISI, M. Comparison of methods for filling in gaps in monthly rainfall series in arid regions. **International Journal of Climatology**, v. 41, n. 15, p. 6674–6689, 2021. https://doi.org/10.1002/joc.7219.

AFRIFA-YAMOAH, E.; MUELLER, U. A.; TAYLOR, S. M.; FISHER. A. J. Missing data imputation of high-resolution temporal climate time series data. **Meteorological Applications**, v. 27, n. 1, p. e1873, 2020. https://doi.org/10.1002/met.1873.

AGÊNCIA NACIONAL DE ÁGUAS (Brasil). **HIDROWEB: Sistema de Informações Hidrológicas**. Brasília, 2018. Available at: http://hidroweb.ana.gov.br/default.asp. Access in: Apr. 27, 2018.

ALLISON, P. D. **Missing data**. Thousand Oaks: Sage publications, 2001. (Quantitative Applications in the Social Sciences, 136).

BARBOSA, M. T. M.; LIMA, A.; KUEHNE, B. T.; BATISTA, B. G.; FILHO, D.M.L.; PEIXOTO, M.L. M. Imputação de dados faltantes no monitoramento de consumo energético residencial em Smart Grids. *In*: WORKSHOP DE COMPUTAÇÃO URBANA (COURB), 2., 2018, Campos do Jordão. **Anais**[...]. Porto Alegre: Sociedade Brasileira de Computação, 2018.

BARNETCHE, D.; KOBIYAMA, M. Aplicação do hycymodel no preenchimento de falhas de monitoramento de vazões. **Geosul**, v. 21, n. 42, p.185-194, 2006.

BAYER, D. M.; CASTRO, N. D. R.; BAYER, F. M. Modelagem e previsão de vazões médias mensais do rio Potiribu utilizando modelos de séries temporais. **RBRH**, v. 17, n. 2, p. 229-239, 2012. https://doi.org/10.21168/rbrh.v17n2.p229-239.

BEN AISSIA, M. A.; CHEBANA, F.; OUARDA, T. B. M. J. Multivariate missing data in hydrology – Review and applications. **Advances in Water Resources**, v. 110, p. 299–309, 2017. https://doi.org/10.1016/j.advwatres.2017.10.002.

BERA, A. K.; JARQUE, C. M. Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. **Economics Letters**, v. 7, n. 4, p. 313-318, 1981. https://doi.org/10.1016/0165-1765(81)90035-5.

BIER, A. A.; FERRAZ, S. E. T. Comparação de metodologias de preenchimento de falhas em dados meteorológicos para estações no sul do Brasil. **Revista Brasileira de Meteorologia**, v. 32, n. 2, p. 215-226, 2017. https://doi.org/10.1590/0102-77863220008.

BLEIDORN, M. T.; SCHMIDT, I. M.; KNAAK, J.; LIMA, G. B.; PINTO, W. P.; BRAUM, E. S. Análise comparativa de vazões ecológicas e disponibilidade hídrica mensal e anual para o rio Doce. *In*: CONGRESSO FLORESTAL LATINO-AMERICANO, 7., 12-15 Jun. 2018, Vitória. **Anais[…]** Vitória: Programa de Pós-graduação em Ciências Florestais; Universidade Federal do Espírito Santo, 2018. Available at: https://www.even3.com.br/anais/conflat/95094-analise-comparativa-de-vazoes-ecologicas-e-disponibilidade-hidrica-mensal-e-anual-para-o-rio-doce/. Access in Oct. 10, 2018.

BOX, G. E.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society: Series B (Methodological**, v. 26, n. 2, p. 211-243, 1964. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x.

BRASIL. Presidência da República. Lei n. 9.433, de 9 de janeiro de 1997. Institui a Política Nacional de Recursos Hídricos, cria o Sistema Nacional de Gerenciamento de Recursos Hídricos, regulamenta o inciso XIX do art. 21 da Constituição Federal, e altera o art. 1º da Lei nº 8.001, de 13 de março de 1990, que modificou a Lei nº 7.990, de 28 de dezembro de 1989. **Diário Oficial [da] União**: seção 1, Brasília, DF, 09 jan. 1997.

BURGER, S.; SILVERMAN, S.; van VUUREN, G. Deriving Correlation Matrices for Missing Financial Time-Series Data. **International Journal of Economics and Finance**, v. 10, n. 10, p. 105-120, 2018. https://doi.org/10.5539/ijef.v10n10p105.

BUUREN, S. V.; GROOTHUIS-OUDSHOORN, K. mice: Multivariate imputation by chained equations in R. **Journal of statistical software**, v. 45, n. 3, p. 1-68, 2010. https://doi.org/10.18637/jss.v045.i03.

CAMARGOS, V. P.; CÉSAR, C. C.; CAIAFFA, W. T.; XAVIER, C. C.; PROIETTI, F. A. Imputação múltipla e análise de casos completos em modelos de regressão logística: uma avaliação prática do impacto das perdas em covariáveis. **Cadernos de Saúde Pública**, v. 27, n. 12, p. 2299-2313, 2011. https://doi.org/10.1590/S0102-311X2011001200003.

CANCHALA-NASTAR, T.; CARVAJAL-ESCOBAR, Y.; ALFONSO-MORALES, W.; CERÓN, W. L.; CAICEDO, E. Estimation of missing data of monthly rainfall in southwestern Colombia using artificial neural networks. **Data in Brief**, v. 26, p. 104517, 2019. https://doi.org/10.1016/j.dib.2019.104517.

CARRERAS, G.; MICCINESI, G.; WILCOCK, A.; PRESTON, N.; NIEBOER, D.; DELIENS, L. *et al.* Missing Not at Random in End of Life Care Studies: Multiple Imputation and Sensitivity Analysis on Data from the ACTION Study. **BMC Medical Research Methodology**, v. 21, n. 13, 2021. https://doi.org/10.1186/s12874-020-01180-y.

CHEN, L.; XU, J.; WANG, G.; SHEN, Z. Comparison of the Multiple Imputation Approaches for Imputing Rainfall Data Series and Their Applications to Watershed Models. **Journal of Hydrology**, v. 572, p. 449–460, 2019. https://doi.org/10.1016/j.jhydrol.2019.03.025.

CHOI, C.; JUNG, H.; CHO, J. An Ensemble Method for Missing Data of Environmental Sensor Considering Univariate and Multivariate Characteristics. **Sensors**, v. 21, n. 22, p. 7595, 2021. https://doi.org/10.3390/s21227595.

COELHO, A. L. N. **Alterações Hidrogeomorfológicas no Médio-Baixo Rio Doce/ES.** 2007. Tese (Doutorado em Geografia) - Instituto de Geociências, Departamento de Geografia, Universidade Federal Fluminense, Niterói, 2007.

COLLINS, A. J.; FOX, J.; LONG, J. S. Modern Methods of Data Analysis. **The Statistician**, v. 40, n. 4, p. 458, 1991. https://doi.org/10.2307/2348744.

COSTA, R. L.; GOMES, H. B.; PINTO, D. D. C.; ROCHA JÚNIOR, R. L.; SILVA, F. D. S.; GOMES, H. B. *et al.* Gap Filling and Quality Control Applied to Meteorological Variables Measured in the Northeast Region of Brazil. **Atmosphere**, v. 12, n. 10, p. 1278, 2021. https://doi.org/10.3390/atmos12101278.

EKEU-WEI, I. T.; BLACKBURN, G. A.; PEDRUCO, P. Infilling missing data in hydrology: Solutions using satellite radar altimetry and multiple imputation for data-sparse regions. **Water**, v. 10, n. 10, p. 1-22, 2018. https://doi.org/10.3390/w10101483.

ELSHORBAGY, A. A.; PANU, U. S.; SIMONOVIC, S. P. Group-based estimation of missing hydrological data: II. Application to streamflows. **Hydrological Sciences Journal**, v. 45, n. 6, p. 867-880, 2000. https://doi.org/10.1080/02626660009492389.

ELY, D. F.; LIMBERGER, L.; MANGILI, F. B.; GAMERO, P.; SCHMENGLER, M. Análise de métodos para o preenchimento de falhas aplicados em séries de dados pluviométricos do estado do Paraná (Brasil). **Raega - O Espaço Geográfico em Análise**, v. 51, p. 122–142, 2021.

FERRARI, G. T.; OZAKI, V. Missing data imputation of climate datasets: Implications to modeling extreme drought events. **Revista Brasileira de Meteorologia**, v. 29, n. 1, p. 21-28, 2014. https://doi.org/10.1590/S0102-77862014000100003.

FIOREZE, A. P.; OLIVEIRA, L. F. C. Uses of hydric resources at the Santa Bárbara River hydrographical basin, Goiás state, Brazil. **Pesquisa Agropecuária Tropical**, v. 40, n. 1, p. 28-35, 2010. https://doi.org/10.5216/pat.v40i1.3869.

FONSECA, J. S. D.; MARTINS, G. D. A. **Curso de estatística**. São Paulo: Atlas, 2011.

GAO, Y. **Dealing with missing data in hydrology - Data analysis of discharge and groundwater timeseries in Northeast Germany**. 2017. Doctoral dissertation (Doctorate degree in Natural Sciences) - Department of Earth Sciences, Freie Universität Berlin, Berlin, 2017.

GAO, Y.; MERZ, C.; LISCHEID, G.; SCHNEIDER, M. A review on missing hydrological data processing. **Environmental Earth Sciences**, v. 77, n. 2, p. 1-12, 2018. https://doi.org/10.1007/s12665-018-7228-6.

GARCÍA-PEÑA, M.; ARCINIEGAS-ALARCÓN, S.; BARBIN, D. Imputação de dados climáticos utilizando a decomposição por valores singulares: uma comparação empírica. **Revista Brasileira de Meteorologia**, v. 29, n. 4, p. 527-536, 2014. https://doi.org/10.1590/0102-778620130005.

GHAZALI, S. M.; SHAADAN, N.; IDRUS, Z. A Comparative Study of Several EOF Based Imputation Methods for Long Gap Missing Values in a Single-Site Temporal Time Dependent (SSTTD) Air Quality (PM10) Dataset. **Pertanika Journal of Science & Technology**, v. 29, n. 4, 2021. https://doi.org/10.47836/pjst.29.4.21.

GREEN, P.; SILVERMAN, B. **Nonparametric Regression and Generalized Linear Models**. New York: Chapman and Hall/CRC, 1993.

GYAU-BOAKYE, P.; SCHULTZ, G. A. Filling gaps in runoff time series in West Africa. **Hydrological Sciences Journal**, v. 39, n. 6, p. 621-636, 1994. https://doi.org/10.1080/02626669409492784.

HAMZAH, F. B.; MOHAMAD HAMZAH, F.; MOHD RAZALI, S. F.; EL-SHAFIE, A. Multiple Imputations by Chained Equations for Recovering Missing Daily Streamflow Observations: A Case Study of Langat River Basin in Malaysia. **Hydrological Sciences Journal**, v. 67, n. 1, p. 137-149, 2022. https://doi.org/10.1080/02626667.2021.2001471.

HAMZAH, F. B.; MOHD HAMZAH, F.; MOHD RAZALI, S. F.; JAAFAR, O.; ABDUL JAMIL, N. Imputation methods for recovering streamflow observation: A methodological review. **Cogent Environmental Science**, v. 6, n. 1, p. 1-21, 2020. https://doi.org/10.1080/23311843.2020.1745133.

HERAS, D.; MATOVELLE, C. Machine-learning methods for hydrological imputation data: analysis of the goodness of fit of the model in hydrographic systems of the Pacific-Ecuador. **Revista Ambiente & Água**, v. 16, 2021. https://doi.org/10.4136/ambi-agua.2708.

JIANG, L.; ZHAO, T.; FENG, C.; ZHANG, W. Improvement of Random Forest by Multiple Imputation Applied to Tower Crane Accident Prediction with Missing Data. **Engineering, Construction and Architectural Management**, 2021 https://doi.org/10.1108/ECAM-07-2021-0606.

JIAO, S.; TIEZZI, F.; HUANG, Y.; GRAY, K.A.; MALTECCA, C. The Use of Multiple Imputation for the Accurate Measurements of Individual Feed Intake by Electronic Feeders. **Journal of Animal Science**, v. 94, n. 2, p. 824-832, 2016. https://doi.org/10.2527/jas.2015-9667.

JOHNSTON, C. A. **Development and evaluation of infilling methods for missing hydrologic and chemical watershed monitoring data**. 1999. Thesis (Master of Science degree in Environmental Engineering) - Faculty of the Virginia Polytechnic Institute and State University, Virginia, 1999.

JUNGER, W. L. **Análise, imputação de dados e interfaces computacionais em estudos de séries temporais epidemiológicas**. 2008. Tese (Doutorado em Saúde Coletiva) – Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2008.

JUNGER, W. L.; LEON, A. P. Imputation of missing data in time series for air pollutants. **Atmospheric Environment**, v. 102, p. 96-104, 2015. https://doi.org/10.1016/j.atmosenv.2014.11.049.

KABIR, G.; TESFAMARIAM, S.; HEMSING, J.; SADIQ, R. Handling incomplete and missing data in water network database using imputation methods. **Sustainable and Resilient Infrastructure**, v. 5, n. 6, p. 365-377, 2019. https://doi.org/10.1080/23789689.2019.1600960.

KAMWAGA, S.; MULUNGU, D. M.; VALIMBA, P. Assessment of empirical and regression methods for infilling missing streamflow data in Little Ruaha catchment Tanzania. **Physics and Chemistry of the Earth, Parts A/B/C**, v. 106, p. 17-28, 2018. https://doi.org/10.1016/j.pce.2018.05.008.

KHALIFELOO, M. H.; MOHAMMAD, M.; HEYDARI, M. Multiple imputation for hydrological missing data by using a regression method (Klang river basin). **International Journal of Research Engineering and Technology**, v. 4, n. 6, p. 519-524, 2015. https://doi.org/10.15623/ijret.2015.0406090.

KHAN, N. A.; TORRALBA, K. D.; ASLAM, F. Missing data in randomised controlled trials of rheumatoid arthritis drug therapy are substantial and handled inappropriately. **RMD open**, v. 7, n. 2, p. e001708, 2021. https://doi.org/10.1136/rmdopen-2021-001708.

KIM, M.; BAEK, S.; LIGARAY, M.; PYO, J.; PARK, M.; CHO, K. H. Comparative studies of different imputation methods for recovering streamflow observation. **Water**, v. 7, n. 12, p. 6847-6860, 2015. https://doi.org/10.3390/w7126663.

LITTLE, R. J. A.; RUBIN, D. B. **Statistical analysis with missing data**. Hoboken: John Wiley & Sons, 2019. https://doi.org/10.1002/9781119013563.

LONGHI, E. H.; FORMIGA, K. T. M. Metodologias para determinar vazão ecológica em rios. **Revista Brasileira de Ciências Ambientais (Online)**, v. 20, p. 33-48, 2011.

McKNIGHT, P. E.; McKNIGHT, K. M; SIDANI, S.; FIGUEREDO, A. J. **Missing data**: a gentle introduction. New York: Guilford Press, 2007.

MELLO, Y. R.; KOHLS, W.; OLIVEIRA, T. M. N. Uso de diferentes métodos para o preenchimento de falhas em estações pluviométricas. **Boletim de Geografia**, v. 35, n. 1, p. 112-121, 2017. https://doi.org/10.4025/bolgeogr.v35i1.30893.

MOREIRA, M. C. **Gestão de recursos hídricos: sistema integrado para otimização da outorga de uso da água**. 2006. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Federal de Viçosa, Viçosa, 2006.

NISHINA, K.; ITO, A.; HANASAKI, N.; HAYASHI, S. Reconstruction of spatially detailed global map of NH_4^+ and NO_3^- application in synthetic nitrogen fertilizer. **Earth System Science Data**, v. 9, n. 1, p. 149-162, 2017. https://doi.org/10.5194/essd-9-149-2017.

NKIAKA, E.; NAWAZ, N. R.; LOVETT, J. C. Using self-organizing maps to infill missing data in hydro-meteorological time series from the Logone catchment, Lake Chad basin. **Environmental Monitoring and Assessment**, v. 188, n. 7, p. 1-12, 2016. https://doi.org/10.1007/s10661-016-5385-1.

NORLIYANA, W.; ISMAIL, W.; ZAWIAH, W.; ZIN, W.; IBRAHIM, W. Estimation of rainfall and stream flow missing data for Terengganu, Malaysia by using interpolation technique methods. **Malaysian Journal of Fundamental & Applied Sciences**, v. 13, n. 3, p. 213–217, 2017. https://doi.org/10.11113/mjfas.v13n3.578.

NUNES, L. N. **Métodos de imputação de dados aplicados na área de saúde**. 2007. Tese (Doutorado em Epidemiologia) - Faculdade de medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

NUNES, L. N.; KLÜCK, M. M.; FACHEL, J. M. G. Uso da imputação múltipla de dados faltantes: uma simulação utilizando dados epidemiológicos. **Cadernos de Saúde Pública**, v. 25, n. 2, p. 268-278, 2009. https://doi.org/10.1590/S0102-311X2009000200005.

NUNES, L. N.; KLÜCK, M. M.; FACHEL, J. M. G. Comparação de métodos de imputação única e múltipla usando como exemplo um modelo de risco para mortalidade cirúrgica. **Revista Brasileira de Epidemiologia**, v. 13, n. 4, p. 596-606, 2010.

OLIVEIRA, L. F.; FIOREZE, A. P.; MEDEIROS, A. M.; SILVA, M. A. Comparação de metodologias de preenchimento de falhas de séries históricas de precipitação pluvial anual. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 14, n. 11, p. 1186-1192, 2010. https://doi.org/10.1590/S1415-43662010001100008.

PAYROVNAZIRI, S. N.; XING, A.; SALMAN, S.; LIU, X.; BIAN, J.; HE, Z. Assessing the impact of imputation on the interpretations of prediction models: A case study on mortality prediction for patients with acute myocardial infarction. **AMIA Joint Summits on Translational Science**, v. 2021, p. 465-474, 2021.

PELISSON, A. A. **Aprendizado de máquina para previsão de geração de energia fotovoltaica em dados de estações solarimétricas**. 2021. Dissertação (mestrado) - Universidade Federal de Santa Catarina. Joinville, 2021.

PEÑA-ANGULO, D.; NADAL-ROMERO, E.; GONZÁLEZ-HIDALGO, J. C.; ALBALADEJO, J.; ANDREU, V.; BAGARELLO, V. *et al.* Spatial variability of the relationships of runoff and sediment yield with weather types throughout the Mediterranean basin. **Journal of Hydrology**, v. 571, p. 390–405, 2019. https://doi.org/10.1016/j.jhydrol.2019.01.059.

PIGOTT, T. D. A review of methods for missing data. **Educational Research and Evaluation**, v. 7, n. 4, p. 353–383, 2001. https://doi.org/10.1076/edre.7.4.353.8937.

PINTO, W. P. **O uso da metodologia de dados faltantes em séries temporais com aplicações a dados de concentração (PM_{10}) observados na Região da Grande Vitória**. 2013. Dissertação (Mestrado em Engenharia Ambiental) - Universidade Federal do Espírito Santo, Vitória, 2013.

PINTO, W. P.; LIMA, G. B.; ZANETTI, J. B. Análise comparativa de modelos de séries temporais para modelagem e previsão de regimes de vazões médias mensais do Rio Doce, Colatina-Espírito Santo. **Ciência e Natura**, v. 37, n. 3, p. 1-11, 2015. http://dx.doi.org/10.5902/2179460X17143.

PRUSKI, F. F.; PEREIRA, S. B.; NOVAES, L. F. D.; SILVA, D. D. D.; RAMOS, M. M. Precipitação média anual e vazão específica média de longa duração, na Bacia do São Francisco. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 8, n. 2-3, p. 247-253, 2004. https://doi.org/10.1590/S1415-43662004000200013.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing. Vienna, 2018.

RADO, O.; FANAH, M.A.; TAKTEK, E. Performance analysis of missing values imputation methods using machine learning techniques. *In*: ARAI, K.; BHATIA, R.; KAPOOR, S. (eds.). **Intelligent Computing-Proceedings of the Computing Conference**. Springer Cham, 2019. p. 738-750. https://doi.org/10.1007/978-3-030-22871-2_51.

RAHMAN, N. A.; DENI, S. M.; RAMLI, N. M. Generalized linear model for estimation of missing daily rainfall data. *In*: AIP CONFERENCE, 2017. **Proceedings[…]** AIP Publishing LLC, 2017. https://doi.org/10.1063/1.4981003.

REISEN, V. A.; MOLINARES, F. A. F.; TEIXEIRA, E. C. Modelagem de séries temporais sazonais na presença de outliers: Estudo de caso da vazão máxima mensal do rio Jucu, ES, Brasil. **RBRH,** v. 13, n. 2, p. 45-53, 2008. http://dx.doi.org/10.21168/rbrh.v13n2.p45-53.

ROTH, P. L.; SWITZER, F. S.; SWITZER, D. M. Missing data in multiple item scales: A Monte Carlo analysis of missing data techniques. **Organizational Research Methods**, v. 2, n. 3, p. 211–232, 1999. https://doi.org/10.1177/109442819923001.

RUBIN, D. B. Procedures with nonignorable nonresponse. *In*: RUBIN, D. B. **Multiple Imputation for Nonresponse in Surveys**. New York: John Wiley & Sons, 1987. p. 202-240.

SABINO, C. V. S.; LAGE, L. V.; ALMEIDA, K. C. B. Uso de métodos estatísticos robustos na análise ambiental. **Engenharia Sanitária e Ambiental**, v. 19, n. SPE, p. 87-94, 2014. https://doi.org/10.1590/S1413-41522014019010000588.

SARMENTO, R. **Termo de referência para a elaboração de estudos sobre a vazão ecológica na bacia do rio São Francisco**. Referência: Edital, (05) do ano de 2006, PROJETO 704BRA2041 da Organização das Nações Unidas para a Educação, a Ciência e a Cultura – UNESCO. 2007. Available at: http://cbhsaofrancisco.org.br/2017/?wpfb_dl=1584. Access in Apr. 20, 2018.

SATTARI, M.; REZAZADEH-JOUDI, A.; KUSIAK, A. Assessment of different methods for estimation of missing data in precipitation studies. **Hydrology Research**, v. 48, n. 4, p. 1032-1044, 2017. https://doi.org/10.2166/nh.2016.364.

SCHAFER, J. L. Multiple imputation: a primer. **Statistical methods in medical research**, v. 8, n. 1, p. 3-15, 1999. https://doi.org/10.1177%2F096228029900800102.

SEMIROMI, M. T.; KOCH, M. Reconstruction of groundwater levels to impute missing values using singular and multichannel spectrum analysis: application to the Ardabil Plain, Iran. **Hydrological Sciences Journal**, v. 64, n. 14, p. 1711-1726, 2019. https://doi.org/10.1080/02626667.2019.1669793.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, v. 52, n. 3/4, p. 591-611, 1965. https://doi.org/10.2307/2333709.

SILVA, M. J. C. **Imputação múltipla: comparação e eficiência em experimentos multi ambientais**. 2012. Dissertação (Mestrado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2012.

SOUSA, H. T.; PRUSKI, F. F.; BOF, L. H. N.; CECON, P. R.; SOUZA, J. R. C. **SisCAH 1.0: Sistema computacional para análises hidrológicas**. Brasília: ANA; Viçosa, MG: UFV, 2009.

STARRETT, S. K.; STARRETT, S. K.; HEIER, T.; SU, Y.; TUAN, D.; BANDURRAGA, M. Filling in missing peak flow data using artificial neural networks. **ARPN Journal of Engineering and Applied Sciences**, , v. 5, n. 1, p. 49-55, 2010.

SWENSON, N. G. Phylogenetic imputation of plant functional trait databases. **Ecography**, v. 37, n. 2, p. 105-110, 2014. https://doi.org/10.1111/j.1600-0587.2013.00528.x.

TENCALIEC, P. **Developments in statistics applied to hydrometeorology: Imputation of streamflow data and semiparametric precipitation modeling**. 2017. Thesis (Doctorate degree in Applied Mathematics) - University Grenoble Alpes, Grenoble, 2017.

TENCALIEC, P.; FAVRE, A. C.; PRIEUR, C.; MATHEVET, T. Reconstruction of missing daily streamflow data using dynamic regression models. **Water Resources Research**, v. 51, n. 12, p. 9447-9463, 2015. https://doi.org/10.1002/2015WR017399.

TUCCI, C. E. **Hidrologia: Ciência e Aplicação**. Porto Alegre: Editora da UFRGS/Edusp/ABRH, 1997.

TURICCHI, J.; O'DRISCOLL, R.; FINLAYSON, G.; DUARTE, C.; PALMEIRA, A. L.; LARSEN, S. C. *et al.* Data imputation and body weight variability calculation using linear and nonlinear methods in data collected from digital smart scales: simulation and validation study. **JMIR mHealth and uHealth**, v. 8, n. 9, p. e17977, 2020. https://doi.org/10.2196/17977.

VEGA-GARCIA, C.; DECUYPER, M.; ALCÁZAR, J. Applying cascade-correlation neural networks to in-fill gaps in Mediterranean daily flow data series. **Water**, v. 11, n. 8, p. 1691, 2019. https://doi.org/10.3390/w11081691.

VENTURA, T. M.; SANTANA, L. L. R.; MARTINS, C. A.; FIGUEIREDO, J. M. Análise da aplicabilidade de métodos estatísticos para preenchimento de falhas em dados meteorológicos (Analysis methods of application for statistical data in meteorology). **Revista Brasileira de Climatologia**, v. 19, n. 12, p. 168-177, 2016. http://dx.doi.org/10.5380/abclima.v19i0.44989.

VINHA, L. G. A; LAROS, J. A. Dados ausentes em avaliações educacionais: comparação de métodos de tratamento. **Estudos em Avaliação Educacional**, v. 29. n. 70, p. 156-187, 2018. https://doi.org/10.18222/eae.v0ix.3916.

WIJESEKARA, W.M.L.K.N.; LIYANAGE, L. Comparison of imputation methods for missing values in air pollution data: Case study on Sydney air quality index. *In*: ARAI, K.; BHATIA, R.; KAPOOR, S. (eds.). **Advances in Information and Communication**: Proceedings of the 2020 Future of Information and Communication Conference (FICC). Springer Cham, 2020. p. 257-269. https://doi.org/10.1007/978-3-030-39442-4_20.

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Removal of turbidity and color in domestic wastewater using aqueous seed extract of *Cassia fistula*

ARTICLES doi:10.4136/ambi-agua.2803

Received: 04 Oct. 2021; Accepted: 09 Mar. 2022

Arnulfo Antonio Tarón Dunoyer*; Luis Enrique Guzmán Carrillo; Rafael Emilio González Cuello

Faculty of Engineering. Food Engineering Program. University of Cartagena, Street 30, n° 48-152, 130014, Zaragocilla, Cartagena, Colombia. E-mail: lguzmanc1@unicartagena.edu.co, rgonzalezc1@unicartagena.edu.co
*Corresponding author. E-mail: atarond@unicartagena.edu.co

ABSTRACT

Many substances of plant origin are extracted for use in the primary treatment of domestic wastewater. In most cases, they are used as coagulating and flocculating agents and are derived from seeds, leaves, bark or sap, roots, and fruits of trees and plants. In this research, the use of *Cassia fistula* seed was evaluated for the removal of turbidity and color in domestic wastewater from a pumping station in the city of Cartagena (Colombia). The optimal dose of *C. fistula* seed powder was determined by jar test using an E&Q F6-300 digital flocculator. Physicochemical parameters such as turbidity and color were determined, following the recommendations of APHA (Standard Methods for Water and Wastewater), expressing the results in UNT (Total Nephelometric Units) for turbidity, and UPt-Co (Platinum-Cobalt Units) for the color. The results obtained show that with a dose of 160 mgL^{-1} of the coagulant extracted from the *C. fistula* seed, a value of 34.14 NTU is reached for removal of 62.18% with respect to the initial turbidity value. The color decreases reaching a minimum value of 88.59 UPC for removal of 64%, at a dose of 160 mgL^{-1} of natural coagulant. The seed *C. fistula* exhibited good coagulating properties at low doses and can be an important alternative for the removal of color and turbidity in wastewater.

Keywords: color, removal, turbidity, water treatment.

Remoção de turbidez e cor em efluentes domésticos usando extrato aquoso de sementes de *Cassia fistula*

RESUMO

Muitas substâncias de origem vegetal são extraídas para serem utilizadas no tratamento primário de águas residuais domésticas. Na maioria dos casos, elas são usadas como agentes coagulantes e floculantes e são derivadas de sementes, folhas, cascas ou seiva, raízes e frutos de árvores e plantas. Nesta pesquisa, avaliou-se o uso da semente de Cassia fístula para a remoção de turbidez e cor em águas residuais domésticas de uma estação de bombeamento na cidade de Cartagena (Colômbia). A dose ideal de pó de semente de C. fístula foi determinada por teste de jarro usando um floculador digital E&Q F6-300. Parâmetros físico-químicos como turbidez e cor foram determinados, seguindo as recomendações da APHA (Métodos Padrão para Água e Esgoto), expressando os resultados em UNT (Unidade Nefelométrica Total) para turbidez e UPt-Co (Unidade de Platina-Cobalto) para a cor. Os resultados obtidos mostram que com uma dose de 160 mgL^{-1} do coagulante extraído da semente de C. fístula, atinge-se um valor de 34,14 NTU para remoção de 62,18% em relação ao valor de turbidez inicial. A cor diminui atingindo um valor mínimo de 88,59 UPC para remoção de 64%, na dose de 160 mgL^{-1} de coagulante natural. A semente de C. fistula exibiu boas propriedades de coagulação em baixas doses e pode ser uma alternativa importante para a remoção de cor e turbidez em águas residuárias.

Palavras-chave: cor, remoção, tratamento de água, turbidez.

1. INTRODUCTION

The quality of both drinking and domestic wastewater is determined in terms of physical, chemical, and biological parameters and its monitoring is extremely important since it is known that the consumption of contaminated water accelerates the appearance of many health problems (Mirzabeygi *et al.*, 2017; Soleimani *et al.*, 2018). High concentrations of organic compounds and suspended particles in the water increase turbidity, serving as a medium for transmitting pathogenic organisms; therefore, the removal of turbidity is an important process in water treatment (Hameed *et al.*, 2018). The removal of turbidity and color in water treatment is basically achieved simply and inexpensively through processes such as coagulation and flocculation. The residual waters produced by domestic activities are usually contaminated by particles, many of which are suspended solids. Removal of these contaminants generally requires the use of coagulants.

Coagulation is one of the oldest processes, widely used in the treatment of drinking and wastewater. Coagulation is a process that removes impurities, especially suspended particles and colloids in water, by destabilizing and agglomerating the particles into larger aggregates. This allows the aggregates to settle quickly and can subsequently be easily separated from the water (Jiang, 2015).

In recent years, the paradigm in the treatment of drinking water and industrial wastewater has changed the culture of water operators, by adopting and implementing sustainable development in the operation. One of the realistic practices is to replace the chemicals used in the treatment process with "green" chemicals that cause less environmental impact in terms of production, consumption, and secondary waste management. In this context, natural coagulant seems to fit this picture and may be an option over conventional inorganic coagulants.

It is reported that natural coagulants are generally easily obtainable from plant raw material; produce less biodegradable sludge (potentially reducing the costs associated with its disposal) and are less affected by the pH of the water (Mohd-Salleh *et al.*, 2019; Saleem and Bachmann, 2019) Inorganic coagulants such as alum are widely used to remove turbidity from the water (Hussain *et al.*, 2019; Ramavandi, 2014). Currently, numerous substances extracted from plants with coagulant activity (natural coagulants) have been discovered. Natural coagulants have proven their clotting efficacy as reported in a substantial number of research articles. However, the widespread acceptance and application of natural coagulants in the water industry remain low.

A review is necessary to promote the use of natural coagulants, highlighting the current development and efforts to improve the capacity of natural coagulants (Ang and Mohammad, 2020; Tarón *et al.*, 2017). Several explanations have been suggested in the literature regarding the coagulant activity of the extracts of natural coagulants (Guzmán *et al.*, 2013).

Ndabigengesere *et al.* (1995) proposed that proteins with a molecular weight of 13 kDa present in plant extracts are active components for coagulation. The chemical composition of the extracts obtained from natural coagulants does not clearly explain the coagulation activity

(Bui *et al.*, 2016; Guzmán *et al.*, 2015; Šciban *et al.*, 2009). Some authors have suggested that the active components of coagulation in plant extracts are not proteins, but some type of organic polyelectrolyte (Sanghi *et al.*, 2002; Okuda *et al.*, 2001).

There are several natural plant extracts such as *Moringa oleifera*, *Jatropha curcas*, *Cyamopsis tetragonoloba*, *Strychnos potatorum*, *Hibiscus sabdariffa*, and *Clidemia angustifolia* that have been used in coagulation and water purification for many years. These coagulants, for the most part, are derived from seeds, leaves, bark or sap, roots and fruits of trees and plants, or can be extracted from microorganisms, animal or plant tissues (Yuan and Manh, 2015). The cañafistula, carao, or cañadonga is a natural tree of Central America and the coastal areas of the Antilles, belonging to the family Fabaceae genus Cassia. In Colombia, México, and probably other countries, it is also known as a *Cassia fistula* Golden-Shower (Tarón *et al.*, 2017).

Guzmán *et al.* (2015), reported in their study that *Cassia fistula* seed powder has excellent properties to be used in the primary treatment of drinking water. In this research, the coagulant power of the *Cassia fistula* seed and its effect on the removal of color and turbidity in domestic wastewater was evaluated.

2. MATERIAL AND METHODS

2.1. Domestic wastewater sample

Domestic wastewater from the domestic wastewater pumping station located in the "La Cuchilla" sector (10.405672, 75.521783) of the city of Cartagena de Indias - Colombia was used. The sample was collected between 9:00 and 10:00 a.m., since it is assumed that a maximum discharge peak passes through this pumping station at that time and contains high levels of turbidity and color.

2.2. Initial physicochemical characterization of wastewater

The domestic wastewater was characterized physicochemically, following the methodology proposed by APHA *et al.* (2012). Turbidity was determined by the nephelometric method (method 2130B); using a Turbiquant 300 IR turbidimeter which measures turbidity in nephelometric turbidity units (NTU) using a formalin polymer as a standard solution.

Table 1 shows the values of the physicochemical parameters of the residual effluent.

Table 1. Initial physicochemical characterization of domestic wastewater.

Parameter	Value	Unit
Total alkalinity	218.0±1.00	mg $CaCO_3$ L^{-1}
BOD	128.1±0.81	mgL^{-1}
COD	219.4±0.76	mgL^{-1}
Total hardness	490.0±0.57	mg $CaCO_3$ L^{-1}
Conductivity	1210.2±0.8	μScm^{-1}
Turbidness	90.28±1.00	NTU
Colour	246.1±0.60	UPC
Total solids	701.0±0.22	mgL^{-1}

Values represent the mean of three determinations (N = 3). Source: the authors and adapted from Tarón et al., (2017).

2.3. Obtaining the coagulant

The natural coagulant was obtained according to the scheme proposed by Yin (2010). The seeds of *C. fistula* were collected manually, the seeds damaged by insects were discarded. Later they were subjected to sun exposure, for approximately 8 days. Immediately upon drying, these were mechanically ground, using a Pulvex Model 95 helical mill; the powder obtained was passed through a 40 mesh sieve. The coagulant solution was prepared by dissolving 25 g of seed powder in 100 mL of distilled water.

2.4. Jar Test

The standard jar test described by Satterfield (2005) was used in this investigation. This consisted of preparing seven solutions with 500 mL of residual water obtained from the pumping station, using one as a control; the remaining six were dosed with *C. fistula* seed extracts at concentrations of 120, 140, 150, 160, 180, and 200 mg / L respectively. The residual water and that mixed with the coagulating agent were initially stirred at 100 rpm for 1 minute, followed by slow stirring at a speed of 40 rpm for 30 minutes; subsequently, the samples were allowed to stand for 60 minutes (sedimentation time). For this, an E&Q F6-300 digital flocculator was used. After the sedimentation time, samples of the supernatant liquid are taken for analysis.

The percentage turbidity removal was determined using Equation 1.

$$percent\ turbidity\ removal = \frac{T_o - T_f}{T_o} * 100 \quad (1)$$

Where; T_o is the initial turbidity value and T_f is the final turbidity value, for each dose of coagulant.

The color was measured by visual comparison of the sample using a Lovibond PFX 195 colorimeter applying the 2120B method and the results of the color evaluation were expressed in platinum-cobalt units (UPC).

The percentage of color removal was determined using Equation 2.

$$percent\ color\ removal = \frac{C_o - C_f}{C_o} * 100 \quad (2)$$

Where C_o is the initial value of the color and C_f is the final value, for each dose of coagulant.

2.5. Statistical analysis

The percentage of removal of turbidity and color was used as response variables. These data were analyzed using the analysis of variance (one-way ANOVA) in order to determine statistically significant differences ($p < 0.05$) between the samples. SPSS software (Version 17.0 for Windows) and Tukey's multiple comparison test were used. All tests were done in triplicate.

3. RESULTS AND DISCUSSION

Table 2 shows the values of the physicochemical parameters of domestic wastewater before and after the coagulation process, using a dose of 160 mgL^{-1}. The results show that the coagulant presents excellent coagulation conditions, which shows a decrease in the levels of the analyzed physicochemical parameters, finding significant differences at $p < 0.05$. When comparing the results with some coagulants used, there is correspondence between them, such as the case of Aluminum Sulfate (Al_2SO_4), which presents similar removal percentages, but using a dose of 20 mgL^{-1}. Prakash *et al.* (2014), reported similar results, but using Moringa oleifera seeds as coagulant.

Table 2. Physicochemical characterization of the domestic residual effluent after the coagulation process with *C. fistula* seed at a dose of 160 mgL^{-1}.

Parameter	**Residual effluent (Ci)**	**Aqueous extract of Seed of *C. fistula***	**$(Al_2SO_4)^*$**	**Unid**
Total alkalinity	218.0±1.00[a]	261.1±08[b]	290.0	mg $CaCO_3$ L^{-1}
Biochemical Oxygen Demand (BOD_5)	128.1±0.81[a]	52.38±20[b]	31.92	mgL^{-1}
Chemical Oxygen Demand (COD)	219.4±0.76[a]	211.1±10[a]	245.0	mgL^{-1}
Total hardness	490.0±0.57[a]	401.3±00[b]	331.5	mg $CaCO_3$ L^{-1}
Conductivity	1210.2±0.8[a]	1346.1±1[b]	1330	μScm^{-1}
Turbidness	90.28±1.00[a]	34.14±17[b]	9.500	NTU
Colour	246.1±0.60[a]	88.59±23[b]	42.00	UPC
Total solids	701.0±0.22[a]	230.2±05[b]	149.4	mgL^{-1}

* taken from Tarón *et al.* (2017), (Al_2SO_4) * at a dose of 20 mgL^{-1}. Different letters in each row show statistically significant differences at $p < 0.05$. C_i: Initial conditions.
Source: the authors.

The BOD_5 value decreases considerably with respect to the initial value, finding statistically significant differences at $p < 0.05$, obtaining a maximum value in removal of 57.02%, using a coagulant dose of 160 mgL^{-1}, these results correspond to those reported for Aluminum Sulfate (Al_2SO_4), which achieved a decrease in the biochemical oxygen demand of 75.08% (Tarón *et al.*, 2017).

Regarding the chemical oxygen demand (COD), no statistically significant differences were found at $p > 0.05$ at a dose of 160 mgL^{-1} of coagulant; it may be feasible, possibly due to the action of various compounds of an organic nature and not biodegradable by conventional methods and refractory to biological oxidation.

Alkalinity, unlike the other physicochemical parameters, increased its concentration level; this behavior can be explained, supported by the possible restoration of the system by the dissolution of the flocs formed (Guzmán *et al.*, 2015).

The domestic wastewater sample initially contained a fairly high level of turbidity (90.28 ± 1.00), characteristic for this type of effluent; the doses of coagulants used in the coagulation process decreased the levels of turbidity in the effluent, at different doses. In Table 3, the results of the turbidity levels in the domestic wastewater can be seen, taking into account the concentrations used in the aqueous extract of *C. fistula* seed.

According to Equation 1, raised in Section 2.4 of this manuscript, removal percentages of 47.09, 54.13, 50.60 and 54.00% are achieved, when using doses of 120, 140, 180 and 200 mgL^{-1} of coagulant, respectively, highlighting that the maximum removal occurs when using a dose of 160 mgL^{-1} of aqueous extract of *C. fistula* seed, in which a removal of 62.18% is achieved, these results are in accordance with those reported in the literature, using other natural coagulants, such as those of *M. oleifera* and chemical coagulants such as Aluminum Sulfate, which present a coagulation capacity greater than that of *C. fistula* seed.

Table 3. Behavior of turbidity and color in domestic waste effluent

Parameters *	**Coagulant dose (mgL^{-1})**				
	120	140	160	180	200
Turbidity	47.76±1[a]	41.41±82[b]	34.14±17[c]	44.59±40[d]	41.52±15[d]
Colour	169.34±06[a]	117.14±21[b]	88.59±23[c]	99.42±28[d]	98.19±17[d]

***** *different letters in each row show statistically significant differences at p <0.05***.**
Source: the authors.

An important aspect to highlight is the fact that after reaching the peak of maximum removal, this percentage (%) decreases, when increasing the coagulant dose from 160 mgL^{-1} to 180 and 200 mgL^{-1} (Figure 1). What would be expected would be that the removal would have a variation directly with the coagulant concentration; the fact that the turbidity level experiences an increase may be caused by an increase in the concentration of total solids, due to the restoration of the system by the dissolution of the flocs formed (Guzmán *et al.*, 2015).

The analysis of variance (ANOVA) shows that there is no statistically significant difference at p> 0,05, between the effluent turbidity levels, when a dose of 180 mgL^{-1} is used with that of 200 mgL^{-1}. According to the results, the aqueous extract of *C. fistula* seed has good coagulant properties that allow it to remove turbidity in waters with high content of solids and organic matter, although these percentages are lower than those presented by sulfate. aluminum, which is one of the most widely used coagulants in the primary treatment of raw and wastewater.

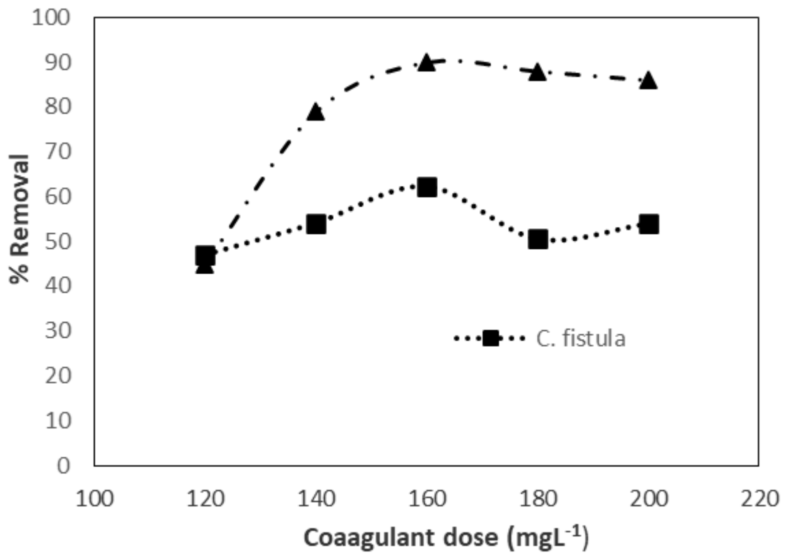


Figure 1. Turbidity removal percentage vs. dose of coagulant applied.
Tarón et al., 2017; doses of Aluminum sulfate [10, 15, 20, 25 and 30 mgL^{-1}].

In Figure 2, the behavior of the color in the effluent is shown, for a dose of 120 mgL^{-1} of coagulant, a removal of 31.19% is achieved, which supposes, being a low percentage of removal, is a promising result, taking into account that, in primary wastewater treatments, higher doses of these are used. These results correspond to those found by Barreto *et al.* (2018), but using avocado seeds (*Persea gratissima* Goerin).

The most efficient dose, in the removal of color, is obtained by using a coagulant concentration of 160 mg L^{-1}, achieving 64% of the total color removal, finding significant statistical differences at p> 0.05 when compared with the percentages obtained with other doses of coagulant. For doses of 180 and 200 mg L^{-1}, no statistical differences were found at p> 0.05.

It must be taken into account that, when using high doses of coagulant, above 160 mgL^{-1}, a re-stabilization of the dispersion can occur, which results in an increase in the turbidity and color values (Guzmán *et al.*, 2015).

The coagulant effect and the ability to remove color in domestic wastewater from the aqueous extract of the *C. fistula* seed is lower than that of aluminum sulfate (Al_2SO_4); however, this can be an advantage, when used in the treatment primary water and thus reduce the amounts used of inorganic polyelectrolytes (chemical coagulants).

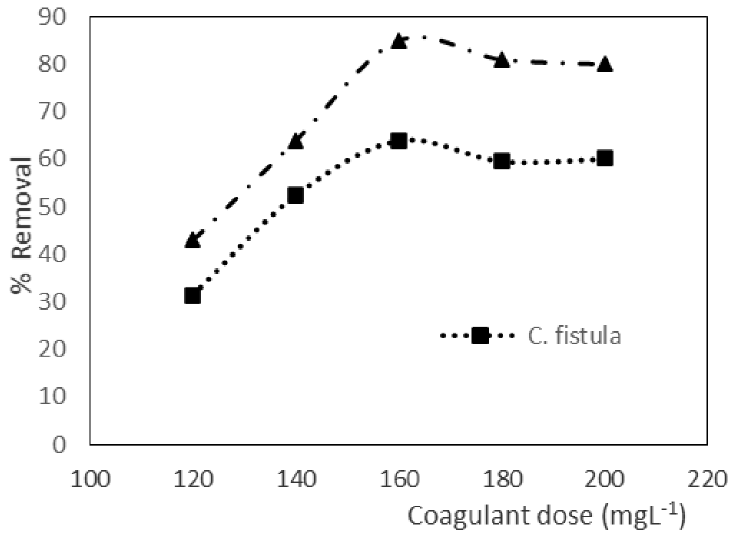


Figure 2. Percentage of color removal vs. dose of coagulant applied.
Tarón et al. (2017); doses of Aluminum sulfate [10, 15, 20, 25 and 30 mgL^{-1}].

4. CONCLUSIONS

The aqueous extract of the *C. fistula* seed powder has interesting qualities to be used in practice as an alternative natural coagulant in the primary treatment of domestic wastewater using dosages close to 160 mg L^{-1}. In this investigation, it was found that the coagulating power and the ability to remove color, turbidity, total solids, and BOD_5 from the aqueous extract of the seed of *C. fistula* is lower than that of aluminum sulfate, which indicates that further investigation is necessary regarding its feasibility for use as a secondary coagulant.

5. REFERENCES

ANG, W. L.; MOHAMMAD, A. W. State of the art and sustainability of natural coagulants in water and wastewater treatment. **Journal of Cleaner Production,** v. 262, n. 121267, p. 1-18, 2020. https://doi.org/10.1016/j.jclepro.2020.121267

APHA; AWWA; WEF. **Standard Methods for the examination of water and wastewater**. 22nd ed. Washington, 2012. 1496 p.

BARRETO, S.; VARGAS, D.; RUIZ, L.; GÓMEZ, S. Evaluación de coagulantes naturales en la clarificación de aguas. **Revista de Investigación Agraria y Ambiental,** v. 11, n.1, p. 105-116, 2018.

BUI, H. M.; PERNG, Y. S.; DUONG, G. The use of artificial neural networks for modeling coagulation of reactive dye wastewater using *Cassia fistula* Linn (CF) gum. **Journal of Environmental Science and Management,** v. 19, n. 1, p. 1-8, 2016. https://doi.org/10.47125/jesam/2016_1/01

GUZMÁN, L.; TARÓN, A.; NUÑEZ, A. Polvo de la semilla *Cassia fistula* como coagulante natural en el tratamiento de agua cruda. **Biotecnología en el Sector Agropecuario y Agroindustrial,** v. 13, n. 2, p. 123-129, 2015. https://doi.org/10.18684/BSAA(13)123-129

GUZMÁN, L.; VILLABONA, A.; TEJADA, C.; GARCÍA, R. Reducción de la turbidez del agua usando coagulantes naturales: una revisión. **Revista UDCA Actualidad & Divulgación Científica,** v. 16, n. 1, p. 253-262, 2013.

HAMEED, Y. T.; IDRIS, A.; HUSSAIN, S. A.; ABDULLAH, N.; MAN, H. C.; SUJA, F. A tannin based agent for coagulation and flocculation of municipal wastewater as a pretreatment for biofilm process. **Journal of Cleaner Prodution,** v. 182, p. 198-205, 2018. https://doi.org/10.1016/j.jclepro.2018.02.044

HUSSAIN, S.; SATTAR GHOURI, A.; AHMAD, A. Pine cone extract as natural coagulant for purification of turbid water. **Heliyon**, v. 5, n. 3, p. 1-11, 2019. https://doi.org/10.1016/j.heliyon.2019.e01420

JIANG, J. Q. The role of coagulation in water treatment. **Current Opinion in Chemical Engineering,** v. 8, p. 36-44, 2015. https://doi.org/10.1016/j.coche.2015.01.008

MIRZABEYGI, M.; YOUSEFI, N.; ABBASNIA, A.; YOUZI, H.; ALIKHANI, M.; MAHVI, A. H. Evaluation of groundwater quality and assessment of scaling potential and corrosiveness of water supply networks. **Iran. Journal of Water Supply: Research and Technology - Aqua**, v. 66, n. 6, p. 416–425, 2017. https://doi.org/10.2166/aqua.2017.128

MOHD-SALLEH, S. N. A.; MOHD-ZIN, N. S.; OTHMAN, N. A review of wastewater treatment using natural material and its potential as aid and composite coagulant. **Sains Malaysiana,** v. 48, n. 1, p. 155-164, 2019. http://dx.doi.org/10.17576/jsm-2019-4801-18

NDABIGENGESERE, A.; NARASIAH, K. S.; TALBOT, B.G. Active agents and mechanism of coagulation of turbid waters using *Moringa oleifera*. **Water Researh,** v. 29, n. 2, p. 703-710, 1995. https://doi.org/10.1016/0043-1354(94)00161-Y

OKUDA, T.; BAES, A. U.; NISHIJIMA, W.; OKADA, M. Coagulation mechanism of salt solution-extracted active component in *Moringa oleifera* seeds. **Water Research,** v. 35, n. 3, p. 830-834, 2001. https://doi.org/10.1016/S0043-1354(00)00296-7

PRAKASH, N. B.; SOCKAN, V.; JAYAKARAN, P. Waste Water Treatment by Coagulation and Flocculation. **International Journal of Engineering, Science and Innovative Technology (IJESIT),** v. 3, n. 2, p. 479-484, 2014.

RAMAVANDI, B. Treatment of water turbidity and bacteria by using a coagulant extracted from *Plantago ovata*. **Water Resources and Industry,** v. 6, p. 36-50, 2014. https://doi.org/10.1016/j.wri.2014.07.001

SALEEM, M.; BACHMANN, R. T. A contemporary review on plant-based coagulants for applications in water treatment. **Journal of Industrial and Engineering Chemistry,** v. 75, p. 281-297, 2019. https://doi.org/10.1016/j.jiec.2018.12.029

SANGHI, R.; BHATTTACHARYA, B.; SINGH, V. *Cassia angustifolia* seed gum as an effective natural coagulant for decolourisation of dye solutions. **Green Chemistry,** v. 4, 2002. https://doi.org/10.1039/B200067A

SATTERFIELD, Z. Jar testing. **Tech Brief**, v. 5, n. 1, p. 1-4, 2005.

ŠCIBAN, M.; KLAŠNJA, M.; ANTOV, M.; ŠKRBIŠ, B. Removal of water turbidity by natural coagulants obtained from chestnut. **Bioresource Technology**, v. 100, n. 24, p. 6639-6643, 2009. https://doi.org/10.1016/j.biortech.2009.06.047

SOLEIMANI, H.; ABBASNIA, A.; YOUSEFI, M.; MOHAMMADI, A. A.; KHORASGANI, F. C. Data on assessment of groundwater quality for drinking and irrigation in rural area Sarpol-e Zahab city, Kermanshah province, Iran. **Data in Brief,** v. 17, p. 148–156, 2018. https://doi.org/10.1016/j.dib.2017.12.061

TARON, A.; GUZMAN, L.; BARROS, I. Evaluación de la *Cassia fistula* como coagulante natural en el tratamiento primario de aguas residuales. **Revista Orinoquía,** v. 21, n. 1, p. 73-78, 2017. https://doi.org/10.22579/20112629.396

YIN, C. Emerging usage of plant-based coagulants for water and wastewater treatment. **Process Biochemistry,** v. 45, n. 9, p. 1437-144, 2010. https://doi.org/10.1016/j.procbio.2010.05.030

YUAN, S. P.; MANH, H. B. The feasibility of *Cassia fistula* gum with polyaluminum chloride for the decolorization of reactive dyeing wastewater. **Journal of Serbian Chemical Society,** v. 80, n. 1, p. 115-125, 2015. https://doi.org/10.2298/JSC140102041P

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Advantages, disadvantages and methods of applying mathematical models to evaluate water quality in reservoirs: a systematic review

ARTICLES doi:10.4136/ambi-agua.2804

Received: 05 Oct. 2021; Accepted: 07 Mar. 2022

Fabio Leandro da Silva[1*]; Ângela Terumi Fushita[2]; Marcela Bianchessi da Cunha-Santino[1]; Irineu Bianchini Júnior[1]

[1]Departamento de Hidrobiologia. Universidade Federal de São Carlos (UFSCar), Rodovia Washington Luiz, KM 235, CEP: 13565-905, São Carlos, SP, Brazil. E-mail: cunha_santino@ufscar.br, irineu@ufscar.br
[2]Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas. Universidade Federal do ABC (UFABC), Avenida dos Estados, n° 5001, CEP: 09210-580, Santo André, SP, Brazil. E-mail: angela.fushita@ufabc.edu.br
[*]Corresponding author. E-mail: fabioleandro@alumni.usp.br

ABSTRACT

Human activities are affecting reservoir water quality; consequently, methods are necessary to verify those impacts. Mathematical modeling improves the understanding of the anthropic impact on water quality, changes in limnological data, and helps formulate management strategies. However, it is necessary to consider the (dis)advantages as well as the methods used for water-quality assessment in reservoirs. This study conducted a systematic review in four databases: (i) PubMed/Medline; (ii) Scopus; (iii) Web of Science; and (iv) Wiley Online Library. We combined Boolean operators and words aiming to identify papers linked to the scope. Rayyan software allowed the initial screening of the found papers. Peer-reviewed papers and the use of mathematical models to assess reservoir water quality were the inclusion criteria. Exclusion criteria included articles in languages other than English, grey literature, and inaccessible articles. Our research found 169 articles, of which 39 were selected and only 13 were included in the review. Mathematical modeling has many benefits related to real-world problems, but the main disadvantages are process simplification, specific rules of the model, and lack of information or data monitoring. Kinetic equations, regression models, Monte Carlo analysis, finite segment models, modeling tools, zero-order rate equations, partial differential algebraic equations, and predictive analysis are the methods observed in mathematical modeling. This review provides information for unfamiliar managers who intend to use mathematical models to assess the water quality of reservoirs.

Keywords: limnologic tool, model inventory, water management.

Vantagens, desvantagens e métodos dos modelos matemáticos aplicados para avaliação da qualidade da água em reservatórios: uma revisão sistemática

RESUMO

As atividades humanas afetam a qualidade da água dos reservatórios, portanto, métodos são necessários para verificar esses impactos. A modelagem matemática auxilia no entendimento do impacto antrópico na qualidade da água, variações nos dados limnológicos e auxilia na formulação de estratégias de manejo. Entretanto, é preciso considerar as (des)vantagens, bem como os métodos utilizados para avaliação da qualidade da água em reservatórios. Este estudo conduziu uma revisão sistemática em quatro bases de dados: (i) PubMed/Medline; (ii) Scopus; (iii) Web of Science e (iv) Wiley Online Library. Foram combinamos operadores booleanos e palavras para encontrar os artigos ligados ao escopo. O *software* Rayyan permitiu a triagem inicial dos artigos encontrados. Artigos *peer-reviewed* e o uso de modelos matemáticos para avaliar a qualidade da água em reservatórios foram critérios de inclusão. Os critérios de exclusão foram artigos não publicados em inglês, literatura cinzenta e artigos inacessíveis. Foram encontrados 169 artigos, dos quais 39 foram selecionados e apenas 13 foram incluídos na revisão. A modelagem matemática tem muitos benefícios relacionados à problemas do mundo real, mas as principais desvantagens são a simplificação do processo, regras específicas e a falta de informações ou monitoramento de dados. Equações cinéticas, modelos de regressão, análise de Monte Carlo, modelos de segmento finito, ferramentas de modelagem, equações de taxa de ordem zero, equações algébricas diferenciais parciais e análise preditiva são os métodos observados. Esta revisão fornece informações para gestores não familiarizados que almejam o uso de modelos matemáticos para avaliar a qualidade da água de reservatórios.

Palavras-chave: ferramenta limnológica, inventário de modelo, manejo da água.

1. INTRODUCTION

Considering water-quality loss in reservoirs, approaches are commonly adopted for assessment and management action proposals (e.g. Dippong *et al.*, 2017; 2018; Bianchini Jr. and Cunha-Santino, 2018). Mathematical modeling greatly improves the understanding of these systems. This tool aims to construct a model that can verify changes in water quality, considering the initial conditions, the dependence of simulations/phenomena, and the final impact (Ziemińska-Stolarska and Skrzypski, 2012).

Previous research recognized the use of mathematical models for the assessment of reservoir water quality (Ward and Linch, 1996; Xu *et al.*, 2017, Crespo *et al.*, 2018, Bianchini Jr. *et al.*, 2019; Chen *et al.*, 2019; Absalon *et al.*, 2020). As demonstrated, this tool seems to be a robust way to verify the impact of human activities on the reservoir, mainly water quality loss, and the main drivers of change in the limnological variables.

Taking into account the results of mathematical modeling, actions can be formulated aiming to reduce the human interference on water quality, due to the identification of priority areas for restoration and action plans formulation (Anderson *et al.*, 1998; Westphal *et al.*, 2004; Zhou *et al.*, 2016). In addition, factors such as potential economic benefits, storage capacity, the impact of climate change, and the dynamics of biogeochemical variables in the reservoir can be verified (Nover *et al.* 2019; Siniscalchi *et al.*, 2020).

Many regions are facing problems because of water scarcity and the negative impacts on agriculture and public water supply (Rosa *et al.*, 2020). Reservoirs play an important role for society, since they provide benefits as supporting services due to the mineralization process (Chen *et al.*, 2019) and water provision for multiple purposes. However, due to alterations in hydrological conditions and changes in metabolism, human activities (urbanization, agriculture, etc.) have adversely affected the reservoirs (Shi *et al.*, 2021).

Mathematical modeling emerges as an alternative to reservoir management. This tool helps to understand the processes taking place in the reservoir, it provides the basis for the strategy, and strengthens water quality restoration (Thibodeaux and Aguilar, 2005; Siniscalchi *et al.*, 2020). However, there is a need to verify the benefits, disadvantages, and methodologies primarily used in the assessment of water quality in reservoirs. Such information helps to understand the tool and may provide information for unfamiliar managers and scientists.

In this study, we conducted a systematic review of the literature, considered primary research, and examined the (dis)advantages of the mathematical models applied to assess the reservoir's water quality. Furthermore, we verified the adopted methodologies in the studies.

2. MATERIAL AND METHODS

The present research was conducted following the guidelines from Collaboration for Environmental Evidence (2018), considering seven steps: (i) planning the systematic review; (ii) conducting the search; (iii) eligibility screening; (iv) data coding and extraction; (v) critical appraisal of study validity; (vi) data synthesis and (vii) interpreting findings. The RepOrting standards for Systematic Evidence Synthesis (ROSES) were also followed. Eventual derivations from the guidelines were detailed below. The primary research question was "What are the benefits and the disadvantages of the mathematical model(s)/approach(es) for the water quality assessment of freshwater reservoirs?". Furthermore, a secondary question was proposed: "Which research methodologies have primarily been used for water quality assessment and management of those freshwater reservoirs?". We based the main research question on the PICO strategy (Population, Intervention, Comparator, Outcomes), Table 1 describes each element of the strategy and the research expression.

Table 1. PICO strategy for the research question.

Description	Abbreviation	Components
Population	P	Freshwater reservoir
Intervention	I	Mathematical model/approach
Comparator	C	Not applicable
Outcomes	O	Water quality

A literature search was conducted between February and April of 2021, in four databases without data limits: (i) PubMed/Medline; (ii) Scopus; (iii) Web of Science; and (iv) Wiley Online Library. Using search terms and Boolean operators, the following expression was employed: ("mathematic* model*" OR "mathematic* approach*" OR "mathematical modeling" OR "modeling approach" OR "mathematical modelling" OR "modelling approach" OR "math model") AND ("freshwater reservoir*" OR "water reservoir*" OR "man-made lake*" OR "man-made reservoir*") AND ("water quality" OR "drinking-water quality" OR "water physicochemical parameters" OR "water biological parameters" OR "limnological parameters" OR "limnological variables" OR "water variables").

Zotero was the chosen bibliographic management software. The studies were screened based on inclusion criteria by two independent researchers (FLS and ATF) using the software Rayyan, verifying the title and abstract. Rayyan (http://rayyan.qcri.org) is a free web app developed by Ouzzani *et al.* (2016), destined to conduct an initial screening of titles and abstracts by a semi-automatic process. The independent reviewers were unaware of each other's decision, in case of disagreement between the two independent researchers (FLS and ATF), a third independent reviewer (MBCS) would solve the disagreements. The full textual analysis of the screened articles was operated by a single reviewer (FLS); however, two reviewers validated the information considering the full text. In the cases of exclusion, the reviewers left a comment specifying the reason for the no inclusion.

Inclusion criteria involved: (i) peer-reviewed articles published in English and (ii) articles that employed mathematical models/approaches to assess the water quality/parameters of reservoirs. On other hand, exclusion criteria were: (i) duplicated articles; (ii) articles in other languages than English; (iii) articles that did not meet the research objective (i.e. reviews, articles that not employed mathematical modeling to verify water quality in freshwater reservoirs); (iv) grey literature and (v) articles not available in full form or with restricting access. The validity level was attributed according to Martins and Carmo Júnior (2018), the study adequacy was verified (full or partial coverage) to the research question, a situation double-checked.

Regarding the critical appraisals of the included articles, we have adopted modified elements pointed out by Haddaway *et al.* (2014), similar to Cresswell *et al.* (2018). Two reviewers (IBJ and MBCS) verified the quality analysis of the papers, the following elements were considered: replication ('possible' or 'no possible'); level of methodological details ('high', 'moderate' or 'low'); results discussion ('consistent' or 'no consistent'); potential bias ('none evident' or 'evident'). Concerning replication (possible or not), it is important to note that, for the evaluation of this criterion, only the degree of data complexity (i.e., quantity and accessibility) was considered for the application (or benchmarking) of the model. Thus, the integrity of the model is not being evaluated, as the articles were previously peer-reviewed. We minimize the bias of this study opting for no temporal delimitation, conducting the process with independent reviewers, and including positive and negative results. The strategy adopted for data extraction and coding was narrative synthesis due to the lack of standardization of the studies. In the end, the findings were interpreted.

3. RESULTS

As shown in the flow chart (Figure 1), 169 articles were recovered from the selected publication databases. Initially, 39 articles were excluded based on repeated results. Conducting an initial review of the title and abstract, we excluded 99 articles. Furthermore, After the criteria adoption and the full reading, 13 articles were included in the review. These studies enabled us to answer the proposed questions.

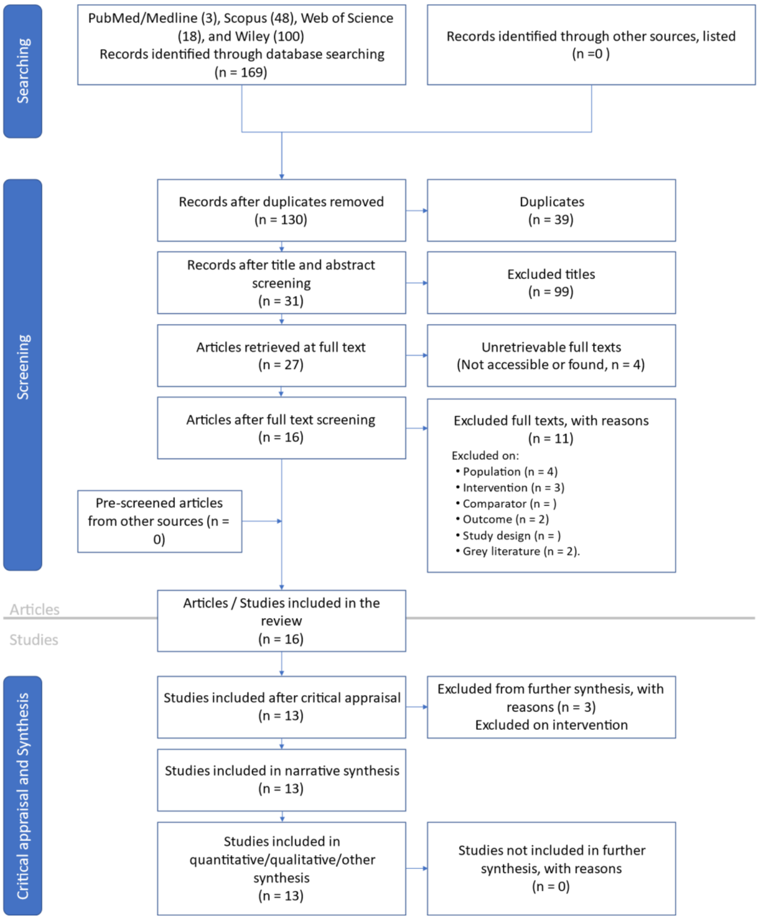


Figure 1. ROSE Flow diagram of the study selection.
Source: Haddaway *et al.* (2018)

Regarding the finds, the articles were published in 11 journals: (i) Water Air Soil Pollution (n=1); (ii) Water Research (n = 1); (iii) Journal of American Water Resources Association (n = 1); (iv) Journal of Water Science (n = 1); (v) Chemosphere (n = 2); (vi) Environmental Monitoring and Assessment (n = 2); (vii) Journal of Hydrology (n = 1); (viii) Science of Total Environment (n = 1); (ix) Computer Aided Chemical Engineering (n = 1); (x) Lake and Reservoir Management (n = 1) and (xi) Sensors (n = 1).

Considering the temporal distribution, the finds cover a period from 1998 to 2020 (Table 2). A single study was published in 1998, 2004, 2005, 2013 and 2016. Regarding the number of publications in the remaining years (2017, 2018, 2019 and 2020), two papers were found in each year. Relatively, the studies are concentrated in the last five years (from 2016 to 2020), because of the number of records (ca. 69%). The mathematical modeling studies were carried out in or the authors were from the United States of America - USA (n = 4), China (n = 3), Brazil (n = 2), Argentina (n = 2), France (n = 1), and Poland (n = 1). Most of the studies focused on temperate zones, followed by the subtropical ones. The critical appraisal can be verified in Table 2.

Table 2. Authors of studies, study area or study conditions, the country of the study, adequacy to the proposed questions and critical appraisal.

Authors (year)	Study area or study conditions	Country	Adequacy	Replication
Anderson *et al.* (1998)	Eastside Reservoir	USA	Total	Not Possible
Westphal *et al.* (2004)	Wachusett Reservoir	USA	Total	Possible
Thibodeaux and Aguilar (2005)	Hypothetical	USA	Partial	Possible
Cunha-Santino *et al.* (2013)	Laboratory conditions	Brazil	Partial	Possible
Zhou *et al.* (2016)	Yuqiao reservoir	China	Partial	Not Possible
Harris and Graham (2017)	Cheney Reservoir	USA	Partial	Possible
Xu *et al.* (2017)	Qingcaosha Reservoir	China	Total	Possible
Crespo *et al.* (2018)	January Lake	France	Partial	Not Possible
Siniscalchi *et al.* (2018)	Chasicó Lake/Paso de las Piedras	Argentina	Partial	Not Possible
Bianchini Jr. *et al.* (2019)	Piraju Reservoir	Brazil	Partial	Possible
Chen *et al.* (2019)	SY Reservoir (referred name)	China	Total	Possible
Absalon *et al.* (2020)	Paprocany Reservoir	Poland	Partial	Not Possible
Siniscalchi *et al.* (2020)	Paso de las Piedras Reservoir	Argentina	Partial	Not Possible

	Level of methodological details	Results discussion	Potential bias
Anderson *et al.* (1998)	Moderate	Consistent	Evident
Westphal *et al.* (2004)	High	Consistent	Evident
Thibodeaux and Aguilar (2005)	Moderate	No Consistent	Not Evident
Cunha-Santino *et al.* (2013)	Moderate	Consistent	Not Evident
Zhou *et al.* (2016)	High	Consistent	Not Evident
Harris and Graham (2017)	Moderate	Consistent	Evident
Xu *et al.* (2017)	High	Consistent	Not Evident
Crespo *et al.* (2018)	Moderate	Consistent	Evident
Siniscalchi *et al.* (2018)	Low	Not Consistent	Not Evident
Bianchini Jr. *et al.* (2019)	Moderate	Consistent	Not Evident
Chen *et al.* (2019)	High	Consistent	Not Evident
Absalon *et al.* (2020)	High	Consistent	Not Evident
Siniscalchi *et al.* (2020)	High	Consistent	Not Evident

3.1. Benefits and the disadvantages of the mathematical model(s)/approach(es) for the water quality assessment of freshwater reservoirs

Mathematical modeling allows the determination of pathogen concentrations (e.g., *Cryptosporidium*, *Giardia* spp, rotavirus, and poliovirus) in the epilimnetic and hypolimnetic region of the water reservoir, and the number of contaminated people because of recreation activities, considering the annual data values and the intensity of recreational activities

(Anderson *et al.*, 1998). This situation was proven in the East End Reservoir (USA) of Anderson *et al.* (1998), besides the benefits (comparison of results with available monitoring data, peak events, eventual risks for population), the lack of mechanistic information, the requirement of sampling data, the assumption simplification, limited information related to impacts of major water exposure, the assumptions involving the decomposition of fecal material and the dispersion of pathogens constitute limitations.

The mathematical model generated information about the diffusion or advection of total organic carbon (TOC) in the Wachusett Reservoir, but Westphal *et al.* (2004) pointed out that model calibration requires a large amount of data and the deviation caused by observational data to improve model performance. In addition, the authors point out that considering some relationships between light and nutrients, model refinement is possible, and it is necessary to obtain water samples to evaluate the effects of proposed plans and actions.

Through mathematical modeling, Thibodeaux and Aguilar (2005) quantified the dissolved organic carbon (DOC) in the bed and water column based on transport, bed sediment, and microbial production. It is worth mentioning that the use of algebraic expressions and differential equations enabled the authors to obtain the necessary empirical evidence, and the generated algorithm predicted the DOC in the hypothetical reservoir.

Mathematical modeling allows to verify the temporal mass changes of course particulate organic matter detritus (CPOM) in the case of new reservoirs and the oxygen consumption, a situation that can reflect on water quality and on the release of greenhouse gases in the short and long time (Cunha-Santino *et al.*, 2013). The authors strengthened the impact, including the risk of eutrophication, oxidation of debris, and increased biochemical oxygen demand (BOD).

Zhou *et al.* (2016) based on the Dyna-CLUE (Land Use Conversion and Its Impact in a Small Area) model and Grey Relational Analysis (GRA), verified the land use changes of the Yuqiao Reservoir (China). The use of GRA allows the development of solutions to real-world problems because of the consideration of changes in development scenarios. In view of the correlation with land use/land cover and land use change, it is possible to verify the restoration of the area near the reservoir, the impact of resettlement and the activities of agribusiness, and generate useful information about the loss of phosphorus (P) in the soil. The authors reinforce that actions based on the model enhance the reduction of P concentrations in the reservoir and how the management actions can reduce the P concentration (36-45%).

In Cheney Reservoir (USA), a comparison of 12 linear and nonlinear regression modeling techniques, aiming the prediction of cyanobacterial abundance and metabolites (microcystin and geosmin) using water quality data from 2004 to 2015, was conducted by Harris and Graham (2017). The authors reinforce that the models had a poor prediction to verify the maxima concentrations, a situation attributed to drought and rainfall events; the models predict cyanobacterial using seasonally variation and do not differ the formation conditions and years. It is highlighted that Cubist modeling has a unique structure, and its creation is based on a tree-based modeling method, a set of rules and terminal node data.

Xu *et al.* (2017) stand out that mathematical modeling has advantages in scenario simulation and identification of mitigation actions aimed at ensuring safe drinking water and managing emerging pollutants, such as atrazine and bisphenol in Qingcaosha Reservoir (China). The researchers verified that the employed model enhances the determination and exchange of atrazine in the reservoir, however, important variations can be ignored due to the long interval of the data input or rainfall events, some simulation bias in the winter, generation of uncertain because of the low frequency of observed data and information about the practical operation.

Crespo *et al.* (2018) verified the pollutant's distribution and optimal strategies for refilling water in January Lake (France) using mathematical modeling. It was possible to determine the localities with adequate water quality for recreation and water intake, considering the constant volume of the lake, the wind influence, water currents, and the water quality evolution.

Using mechanistic models and optimal control problem models, Siniscalchi *et al.* (2018) evaluated the salt concentration considering the humid climate scenario and the variables that affect the phytoplankton in the reservoir, aiming at a recovery strategy. Due to the numerical results, the authors believe that modeling is a useful tool for planning and verifying the impact on water bodies.

Bianchini Jr. *et al.* (2019) showed that the mathematical model used to verify the material mass balance in the Piraju Reservoir (Brazil) is a feasible method to assist in the water quality monitoring of watersheds. In addition, the authors emphasize that since freshwater ecosystem services can be inferred, the model can be used in water quality plans and scenarios involving aquatic bodies.

The carbon dynamics, climatic conditions and nutrient scenarios, algal blooms and systemic carbon conversion of SY Reservoir (China) were simulated by Chen *et al.* (2019). The benefits involved the slight tendency to underestimate the level of the water because of a tributary negation, and a negative bias of surface carbon dioxide in the function of an overestimation in SY reservoir.

Absalon *et al.* (2020) used a three-dimensional hydrodynamic and ecosystem model, in Paprocrany Reservoir (Poland), considering the following data from 1995 to 2014: temperature, dissolved oxygen, pH, total suspended solids, DOC, dissolved inorganic carbon, dissolved (in)organic nitrogen, particulate organic nitrogen, ammonia nitrogen, dissolved organic phosphorus, particulate organic phosphorus, phosphate organic phosphorus, silica, bacteria, phytoplankton, and zooplankton. The researchers have increased the demand for available data, and monitoring of watershed areas is necessary for assessment and scenario formulation. The analysis is helpful to verify the impact on water quality and changes in limnological variables in the reservoir area, analysis of mitigation options, and maintenance/improvement of water quality related to climate change (Absalon *et al.*, 2020).

Artificial wetlands models associated to a mechanistic reservoir model demonstrated potential as a remediation tool, the mathematical modeling study carried out in Paso de las Piedras Reservoir (Argentina) enhanced the comparison of different scenarios and opportunities to the obtainment of improvements strategies formulation to reduce nutrients loading and favor the water quality recovery in eutrophic systems (Siniscalchi *et al.*, 2020). Due to the information about gradients, differential algebra, and variable optimization, the authors considered this method an advantage, which is conducive to the planning and design of water supply systems.

3.2. Research methodologies for water quality assessment and management of freshwater reservoirs

In the case of Eastside Reservoir (USA), it was observed that the Monte Carlo technique was transformed into a finite-segment model (hybrid Monte Carlo finite-segment method). Multiple simulations occurred based on parameters (pathogen content, feces mass, inaction rate, infection rate, feces shed, reservoir segmentation, depth, volume variation assumptions, water body filling in the climatic seasons), which allowed verifying the dispersive transport, and the pathogen inputs, based on available data (Anderson *et al.*, 1998).

Westphal *et al.* (2004) adopted a quasi-mechanistic approach to verify the advection, settling, and diffusion of TOC in Wachusett Reservoir; a situation that resulted in a simple mass balance model with equations which considered the longitudinal reservoir division, the bathymetric information, two-dimensional structures, flow, volume, time step, diffusion component, and mixing component.

A two-step DOC release model inferred in a previous study, in the tea bag equation and continuous microorganisms' production was employed by Thibodeaux and Aguilar (2005), considering a hypothetical reservoir. These authors assumed that the readily quantification of DOC, the microbial production and transport of DOC (steady-state model and mass balance), boundary condition with to Fick's first law, the inter-phase transport, the water column DOC accumulation model, and a zero-order rate equation that was based on DOC kinetics. The following variables were considered: soil porosity, the concentration of carbon (mg/L), DOC concentration in the pore-water, the distance from the interface into the bed, DOC production rate, the DOC diffusivity, mass-transfer coefficient. In addition, there is a model for quantifying DOC in water bodies and DOC productivity based on bed depth (Thibodeaux and Aguilar, 2005).

Cunha-Santino *et al.* (2013) used a set of equations to verify the mineralization of CPOM under experimental conditions. These equations are based on the parameterization of the time evolution of elements (particulate organic carbon - POC, total inorganic carbon – TIC, TOC) and nonlinear regression calculations using the Levenberg-Marquardt algorithm. It took into account: POC mass loss, formation and mineralization of DOC, the formation of gases and inorganic substances, and the kinetics of the dissolved oxygen consumption during the mineralization process.

It is possible to verify the use of GRA to check problems that are based on a situation without information (black), with complete information (white), and with intermediary information (grey) (Zhou *et al.*, 2016). Data are necessary to calibrate the model and compare the simulated scenarios, multiple linear regression models are adopted to verify the model validation. Zhou *et al.* (2016) based the simulated scenarios of land-use change and adopted driving forces (e.g., P sources, agricultural practices) on the grey model and multiple linear regression analysis, using differential equations, and the grey relational grade to estimate the lack of data, as well as the Conversion of Land Use and Its Effects at Small Regional Extent model.

Harris and Graham (2017) use a predictive model combined with the use of training functions and random data selection. Each model has five repetitions and 10-fold cross-validation. The authors compared the 12 models according to the root mean square error (RMSE), linear and nonlinear regression models were included, and the comparison occurred with the observed response (variables concentrations). The authors verified the temporal variation patterns and the abundance of cyanobacteria and the metabolites, using cubist models (support vector machines, random forest, bagged trees), due to the predictive roles of water variables (chlorophyll a, nutrients, temperature, time of the year, iron, oxygen).

Xu *et al.* (2017) used a 3D-emerging model in the Qingcaosha Reservoir; this item is a numerical hydrodynamic model that counts with an orthogonal curvilinear coordinate which simulates transport, flow, and water temperature. The foundation of modeling is based on a grid and sounding method. Such models considered meteorological data, processes as degradation, and rainfall, diffusion, and reaction equations allowed the modulation of atrazine and bisphenol.

The description of water quality evolution (generic pollutant) in a reservoir, using numerical simulations and regarding the optimization problem based on an algorithm and the refilling process was possible in Jaunay Lake (France). In the mathematical modeling, Crespo *et al.* (2018) represented the lake surface geometry, denoted the pollutant concentration, the mensuration of the pollutant volume per area, the time for the model process, as well the influence of four effects (diffusion of pollutant, the wind, the water currents and spill/removal of pollutant). Based on the listed factors and the space-time, the formulated equation considered: the diffusion coefficient of pollutant, the wind velocity, the water velocity, the lake geometry, and the river velocity.

Siniscalchi *et al.* (2018) evaluated the salt reservoirs in Argentina (Chasicó Lake and Paso de las Piedras Reservoir) using eco-hydrological models, including the mass balance of variables, evaporation and kinetic equations in the algebraic equation system. The hydrological model shows the dynamic water mass balance based on an equation that takes into account the total mass of the lake (kg), water density, volume, precipitation, flow velocity of tributary rivers, flow velocity of groundwater, and evaporation.

Salt reservoirs (Chasicó Lake and Paso de las Piedras Reservoir) in Argentina were evaluated with ecohydrological models, Siniscalchi *et al.* (2018) included the mass balance of variables, evaporation, and kinetic equations in algebraic equations systems. Evaporation was calculated according to the radiation effects, the net radiation, wind, and vapor saturation deficit. The salt concentration was obtained by algebraic equations that considered water column height, rainfall, and flow rate. Regarding water quality, for this purpose, the authors included the mass balance of the reservoir' trophic chain (nutrients, phytoplankton, zooplankton, zoo planktivorous and piscivorous fishes) and forcing functions (solar radiation, temperature, contributor river's inflows, nutrients concentrations, ecological state). Also, the authors employed a differential-algebraic equation (DAE) to the optimal control of the problems.

In Piraju Reservoir (Brazil), the description of the mass balance of substances used a sampling campaign from 2003 to 2007 and two equations that englobe an assimilation factor (alpha), the first-order reaction rate constant, hydraulic flushing, the volume of the reservoir, the daily load of substance, the upstream flow rate, and the time needed to reach the equilibrium of the concentration (Bianchini Jr. *et al.*, 2019). The authors assumed that the reservoir is in a steady state, represented by a continuous stirred tank reactor. In addition, the researchers emphasized that the reservoir is a hybrid system with a step input; the change in the alpha coefficient indicates the retention or release of the substance being evaluated.

Tools such as Delft3D are used as modeling ones. In the case of SY Reservoir, modules (WAQ, FLOW) and the hydrodynamic model were employed to verify the dispersive/advective fluxes, furthermore, orthogonal curvilinear grid and the k-epsilon model were employed, using data from meteorological data, as air temperature, solar radiation, cloud cover, etc. (Chen *et al.*, 2019). The ecological Delft3D-WAQ model used by the researchers takes into account key processes (inorganic carbon balance, nutrient cycling, phytoplankton dynamics) and ecological aspects (species competition, limiting nutrients, light, temperature), the modulation was based on separate variables and occurred the simulation in the hydrodynamic/ecological models, the calibration occurred considering a trial-and-error method and the performance analysis was based on root mean square error, Nash-Sutcliffe efficiency, and root mean square difference.

The Aquatic Ecosystem 3D Model (AEM3D) was employed to verify the variability of water quality and flow, the effects of the increase related to water supply, the effects of a barrier that limits the water supply, and the climate change impacts. In the modeling simulation, Absalon *et al.* (2020) included the water flow (direction and velocity), retention time, transport of a virtual marker, temperature changes, and the variations in the concentration of nutrients and plankton. The authors used the Reynolds averaged Navier-Stokes and Reynolds average kinematic boundary to simulate the heat exchange, phytoplankton and zooplankton dynamics in the reservoir.

In the case of Paso de las Piedras Reservoir, the mathematical model was based on the dynamic of the limnological variables mass balance (zooplankton groups, phytoplankton groups, nitrogen species, phosphorus species, organic carbon, and dissolved oxygen), and considered the sediment resuspension, a partial differential equation system, and the inflows and outflows of water in the system (Siniscalchi *et al.*, 2020). This model was based on partial differential-algebraic equations that are transformed in an algebraic equation system; the integrated model aims to adopt a vector parameterization approach to an optimal control problem, verified by an optimization algorithm. In this sense, the authors adopted a mass balance, forcing functions and kinetic equations for wetlands construction, with the intent to restore water quality. It was considered that macrophytes, orthophosphate, organic phosphorus, nitrate, ammonium, organic nitrogen, POC, DOC, dissolved oxygen, the input from tributaries, the factor relative with the tributary diverted fraction into wetlands, generation/consumption (rate equations - respiration, natural death and organism growth) in wetlands, and output flow rate (Siniscalchi *et al.*, 2020).

3.3. Finds synthesis

Figure 2 shows a synthesis of the benefits, disadvantages, methods and the data necessary for the modeling models applied to the analysis of water quality in reservoirs. Considering all the included papers in this review, it is clear that many methods support the analysis, and a variety of water quality parameters and scenarios can be evaluated. However, data from monitoring campaigns are essential for the model application, as well as validation. The need to simplify ecological processes or assumptions, the frequency of data monitoring, the lack of data, and the limited information on the object of analysis are the most common designated shortcomings. In Table 3, the main information of the selected articles is pointed out.

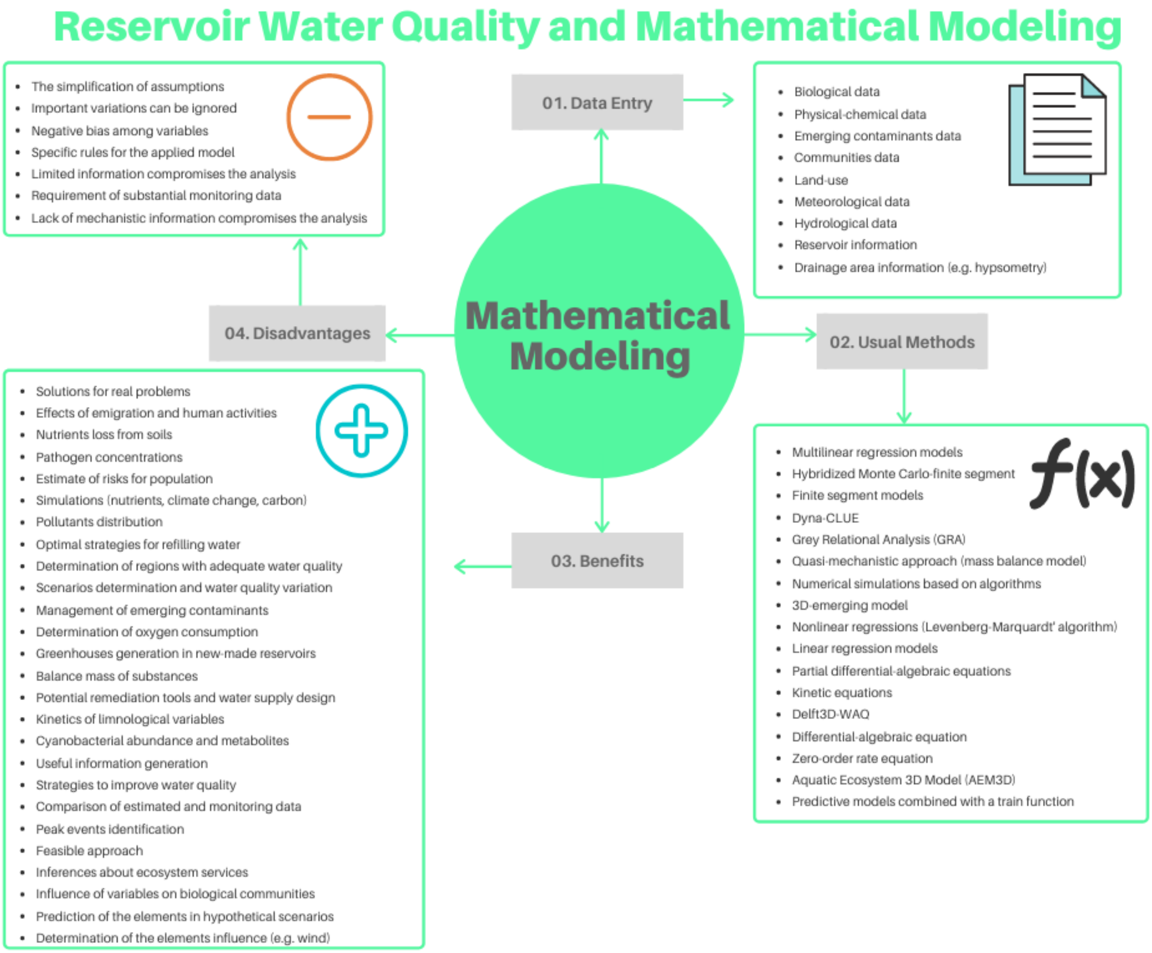


Figure 2. Mathematical modeling in reservoirs: data entry, usual methods, benefits, and disadvantages. Based on: Absalon *et al.* (2020); Anderson *et al.* (1998), Bianchini Jr. *et al.* (2019), Chen *et al.* (2019), Crespo *et al.* (2018), Cunha-Santino *et al.* (2013), Harris and Graham (2017), Siniscalchi *et al.* (2018; 2020), Thibodeaux and Aguilar (2005), Westphal *et al.* (2004), Xu *et al.* (2017), Zhou *et al.* (2016).

4. DISCUSSION

Considering the importance of aquatic ecosystems and the water quality loss the mathematical modeling plays an important role due to strategy formulation (e.g., limitation of human activities around the reservoir), the pollution influence on water, the limnological variables patterns, and the guideline attendance (Kerachian and Karamouz, 2007; Dippong *et al.*, 2018). Mathematical modeling subsidizes the improvement of water quality and the proper management of reservoirs, because of the contribution related to the new made systems and the optimization of the existing ones (Bianhini Jr. and Cunha-Santino, 2018).

Table 3. Synthesis of the selected studies.

Authors	Methods	Advantages	Disadvantages/Limitations
Anderson *et al.* (1998)	Use of a Hybridized Monte Carlo finite segment model to predict the pathogen concentration in the Eastside Reservoir, considering the spatial and temporal variability.	Assessment of the recreational activities impacts on water quality; simulation of pathogen concentration; prediction of annual mean values, considering acceptable levels and treatment practices; comparison of the results with available sampling data; peak events identification.	Requirement of sampling data to identify peak events; lack of mechanistic information; simplification of assumptions.
Westphal *et al.* (2004)	Quasi-mechanistic approach that resulted in a mass balance model with equations to simulate TOC.	Use of historical input series for calibration; simulation of mechanistic elements (e.g. diffusion); the predictive strength can be improved using updated data; credibility to be extended into long planning and operational strategies; characterization of thermal/seasonal structure.	Calibration demands large amounts of data to improve the accuracy; simplification of elements; underprediction/overprediction due to TOC sources; eventual multivariate dependencies in the variables.
Thibodeaux and Aguilar (2005)	Quantification of the DOC in the bed and water column based on a mathematical model, using differential equations, algebraic expressions and transport kinetics in laboratory conditions.	It was possible to obtain an algorithm that predicts the DOC in a hypothetical reservoir. The steps led to the understanding of DOC release, considering basic mechanisms and resulting in the generalization of results.	Assumptions related to temperature are necessary. Also, uncertain intervals need to be considered. Sufficient data are necessary for the modeling. Differential equations are necessary to describe microbial processes.
Cunha-Santino *et al.* (2013)	The CPOM mineralization was described by equations and non-linear regression with an iterative algorithm (Levenberg-Marquardt). The oxygen consumption was described using a first-order kinetics model.	The mathematical model of decomposition kinetics allowed us to verify the half-life of CPOM and the effects on water quality from short to long-term, including oxygen consumption and implications to eutrophication.	Not clear in the text.

Continue...

Continued...			
Zhou *et al.* (2016)	Use of Dyna-CLUE model, grey relational analysis (GRA), and grey model (GM) to simulate P levels	Prediction of P concentration changes associated with land use; identification of the ecological and environmental effects of changes in land use. Less data requirement.	There is a need to adapt the modeling approach for other watersheds.
Harris and Graham (2017)	Use of training functions and random data to verify the temporal variation patterns and the abundance of cyanobacteria. The predictive model included (non)linear regression and five repetitions.	Prediction of cyanobacterial, geosmin, and microcystin abundance. It was possible to verify important predictor variables. Also, the model demanded fewer explanatory variables. Improvements are possible adjusting the models.	The cubist model was more robust in the cases of larger cyanobacterial abundances or geosmin concentrations. The maxima concentration was not predicted due to seasonal changes (environmental variation was not captured). The models do not distinguish inter-annual/intra-annual differences. Long-term data are necessary.
Xu *et al.* (2017)	A 3D-emerging model (Delft3D-FLOW/Delft3D-WAQ) was used, the numerical hydrodynamic simulates some parameters (transport, flow, water temperature) and counts with an orthogonal curvilinear coordinate	The modeling was capable of describing the patterns of water temperature, salinity and emerging contaminants over time and space. It was possible to verify the transport and biodegradation of the contaminants, insight into risk of contaminants and base for decision-making. Observational data was employed to adjust the model.	The complications of the water quality model demand different criteria of goodness-of-fit. May high input data and long intervals ignore important variations. Low observed that it can generate uncertainty in prediction.
Crespo *et al.* (2018)	A mathematical model was performed considering numerical simulations and the problem optimization, based on an algorithm and refiling process.	Optimal strategies were obtained for refilling water, and optimal locations were identified ensuring water quality. The generic pollutant distribution associated with refilling location was obtained, as well as prospective refilling.	The model assumes that water volume is constant, the pollutant remains at the water surface, and its distribution is influenced by wind and water currents.
Siniscalchi *et al.* (2018)	Differential algebraic equations (Kinect, mass balance, evaporation) represented the ecohydrological models. A dataset (10 years) was used for calibration.	Optimal control problems were possible; the modeling addressed salinity and flooding issues. Also, eutrophication problems were considered, as well as restoration profiles, biomanipulation processes, and useful elements for decision-making were generated.	The entire model has 110 algebraic equations, as well as 42 differential equations.
Continue...			

Continued...

Bianchini Jr. *et al.* (2019)	Description of limnological mass balance from, using two equations.	The model allowed us to verify the retention capacity of the reservoir, showing the numbers of retentions and inferences about the physical processes, using an alpha parameter. It is a feasible approach in water monitoring.	The model had some premises: the reservoir is a completely mixed system; the system can be represented by a zero-dimensional model and is in a steady state.
Chen *et al.* (2019)	A three-dimensional ecological model was employed to simulate the carbon dynamics, the climate conditions, nutrients scenarios, algal blooms, and the systematic carbon transformations.	Scenarios were evaluated, considering climate and nutrients, including the systematic carbon transformations. The model allowed us to verify changes in the trophic state associated with CO2 and water volume. Simulations (algal blooms, carbon dynamics) were possible. The model can be coupled with other models (e.g. watersheds).	A negative bias of surface CO_2 was verified; but the model was validated. Simulated CO_2 concentrations were obtained using derived data from semiempirical equations, so uncertain and potential sources of error can be verified. The atmospheric contribution was not considered. The parameters in the Wanninkhof equation have 20% of uncertainty.
Absalon *et al.* (2020)	The Aquatic Ecosystem 3D Model (AEM3D) was used to assess variability in water quality and flow, considering many parameters (water flow, retention time, transport, temperature, plankton, and nutrients concentration).	Impacts on water quality, flow and limnological variables over time were verified based on the model, including scenarios formulation related to pollutants, climate change and algal bloom. Actions formulation was possible. Also, the main source of problems was identified.	The hydrodynamics and thermodynamics of the reservoir were represented by > 100 equations that represent the system processes. No sufficient available data for 2016 compromised the parameterization of water quality and inflows.
Siniscalchi *et al.* (2020)	It involves an integration of mechanistic models and partial differential-algebraic equations. The analysis considered water variables, mass balance of biogeochemical variables (including taxonomic groups), inflows and outflows.	The approach provided temporal profiles for biogeochemical variables, contributing to planning and restoration measures. Experimental data calibrated the model. Furthermore, optimal control and problem design were considered. The study showed as advantage the information on gradients and the variables optimization.	Not clear in the text.

Previous research showed how mathematical modeling is useful for water quality assessment. Ward and Linch (1996) focused on the use of mathematical modeling for recreational benefits with biological variables and environmental changes; the results reinforce the possibility of water quality improvements allowed by models. Guzman *et al.* (2017) evaluated the influence of landscape cover change, sediment loads, water infiltration and a long-data period, favoring the elaboration of rehabilitation projects. Using mathematical modeling, Cid *et al.* (2011) identified the most polluted areas and the main pollution sources of a reservoir in Argentina, using limnological data and tridimensional models.

In this study we verified such possibilities; many situations linked to water quality and the reservoir drainage area can be explored using mathematical modeling. However, mathematical modeling has many assumptions, and factors such as strategy and system complexity need to be considered (Nover *et al.* 2019). Also, it is necessary that the attainment of optimal operation rules, the adjustment of algorithms, use of hybrid methods, the need of input data (observational data) for the verification of hydro-environmental processes, robustness, acceptable performance, prediction, accuracy, deal with the limitation of models, and the determination of parameters (Bezsonnyi *et al.*, 2017; Karami *et al.*, 2019, Latif *et al.*, 2022).

Reservoirs have several impacts associated with the damming that impact directly the water quality and ecological processes (Winton *et al.*, 2019). A model can be applied to similar systems, sometimes it is not necessary to use specific programs (Bezsonnyi *et al.*, 2017). According to Zieminska-Stolarska and Skrzypski (2012), the most common issues evaluated for mathematical modeling (physical, analytical, or numerical) are related to water quality compromising, basically the models can be divided into: (i) one-dimensional models - common to verify changes in parameters; (ii) two-dimensional models - usual in reservoirs due to longitudinal and depth profile; and (iii) three-dimensional models (3D) - considers the spatial distribution.

In the 1970s and 1980s, the need to assess water quality in reservoirs inspired mathematical modeling. Scientists tried to integrate ecological processes, water quality, and hydrodynamics based on assumptions and dimensional representations (e.g. 1D, 2D, 3D, or more) (Orlob, 1992). Nowadays, we can see that the found results provide evidence about advances in mathematical modeling and yet water quality, ecological processes, the reservoir hydrodynamic, and dimensional representations are elements present in the analysis, the benefits enhance decision-making, but limitations and the resources for the modeling process need to be considered.

Indeed, the model’s selection by users demands verification of the complexity, the temporal scales, needful data, personal knowledge, the project, and calibration (Yuan *et al.*, 2020). The model can have many assumptions; it is important to emphasize the need to consider factors that do not exist in the model, especially political and system complexity (Nover *et al.* 2019). Criteria can be adopted for mathematical models’ selections, evaluation can be auxiliaries in the selection of the most appropriate model (e.g. Chinyama *et al.*, 2014).

Hundreds of mathematical models can be found for water quality assessment; the guidance of developed countries supports the standardization of model use, but this is not the case in emerging countries (Wang *et al.*, 2013). Mathematical modeling is a feasible tool that saves resources and time for water quality assessment in reservoirs (Heidarzadeh *et al.*, 2021). In view of the achievements and contributions to management, this study points out the tools used for reservoir water quality monitoring and decision-making.

Considering future research involving mathematical modeling and water quality in reservoirs, there is a need to solve the main disadvantages (e.g. bias, specific rules, simplifications, etc.) and maintain the assessment of the system complexity, including the main processes. The advantages can be considered in the selection of the appropriate model for reservoir analysis; however, data availability and mechanistic information compromise the analysis.

It is worth mentioning that the limitation of this study is related to the standards of attendance and the small number of articles included in the review. In addition, the sample supports the obtained composition related to the subject. In future studies, we recommend a review to verify the advantages and disadvantages of the mathematical modeling methods that are mainly used to assess the water quality of reservoirs.

5. FINAL CONSIDERATIONS

This review provides information on the role of mathematical modeling in the assessment of reservoir water quality. The advantages, disadvantages, primarily used methodologies and data required for calibration and analysis of mathematical modeling for water quality assessment in reservoirs are pointed out. The study found that the benefits of mathematical modeling are diverse, and most models allow verification of limnological variable models and the formulation of management actions. Some disadvantages, mainly limited data and simplifications, can compromise mathematical modeling. Regression, mechanical methods and algebraic equations are the most commonly used methods.

Benefits from mathematical modeling are the analysis of real problems involving population growth, climate changes and human activities' development; the nutrients and pollutant's dynamics and contributions from the reservoir area; optimal strategies; determination of areas with adequate quality water; the mass balance of substances; the kinetics of limnological variables; inference about ecosystem services as the supporting ones; the influence of physical-chemical variables on biological communities; and prediction about hypothetical scenarios (e.g., new made reservoirs).

The main found disadvantages are the necessity of simply process and real world situations; important variations can be ignored due to monitoring interval and the modeling conditions; the analyzed variables can demonstrate bias because the causality of variables in some cases; models have specific rules, given the structure and the analytical components; the lack of information or limitations about some themes can compromise the mathematical modeling (e.g., contaminated people due to water contact, lack of mechanistic information related to the reservoir as flow and volume); and the inexistence of a large base of monitoring data can compromise the analysis and the determination of patterns.

Data input requires information from the monitoring program; it is necessary that data be input from communities (phytoplankton, zooplankton, cyanobacterial abundance), limnological data (nutrients, heavy metals, solids, temperature, etc.), biologic data (chlorophyll-a, pathogens), landscape cover in the reservoir area, meteorological data (temperature, wind, precipitation, air humidity), reservoir characteristics (volume, area, retention time, river flow, etc.), new compounds (such as pesticides).

The research methodologies primarily used for water quality assessment of reservoirs include: kinetic equations, multilinear/linear regression models, hybridized Monte-Carlo analysis, finite segment models, GRA, mass balance models, modeling tools (Delft3D WAQ, 3D-emerging, AEM3D, Dyna-CLUE), zero order rate equation, partial differential-algebraic equations, and predictive analysis allied to train function. Further research should evaluate the pros and cons of the research methods that have been identified.

6. ACKNOWLEDGMENT

Financial support was provided by the National Council of Technological and Scientific Development (Process: 158927/2018-4).

7. REFERENCES

ABSALON, D.; MATYSIK, M.; WOZNICA, A.; LOZOWSKI, B.; JAROSZ, W.; ULANCZYK, R.; BABCZYNSKA, A.; PASIERBINSKI, A. Multi-Faceted Environmental Analysis to Improve the Quality of Anthropogenic Water Reservoirs (Paprocany Reservoir Case Study). **Sensors**, v. 20, n. 9, p. 1-30, 2020. https://doi.org/10.3390/s20092626

ANDERSON, M. A.; STEWART, M. H.; YATES, M. V.; GERBA, C. P. Modeling the impact of body-contact recreation on pathogen concentrations in a source drinking water reservoir. **Water Research**, v. 32, p. 3293–3306, 1998. https://doi.org/10.1016/S0043-1354(98)00128-6

BEZSONNYI, V.; TRETYAKOV, O.; KHALMURADOV, B.; PONOMARENKO, R. Examining the dynamics and modeling of oxygen regime of Chervonooskil water reservoir. **Eastern-European Journal of Enterprise Technologies**, v. 5, n. 10, p. 32–38, 2017. https://doi.org/10.15587/1729-4061.2017.109477

BIANCHINI JR, I.; FUSHITA, Â. T.; CUNHA-SANTINO, M. B. Evaluating the retention capacity of a new subtropical run-of-river reservoir. **Environmental Monitoring Assessment**, v. 191, n. 3, p. 1-15, 2019. https://doi.org/10.1007/s10661-019-7295-5

BIANCHINI JR, I.; CUNHA-SANTINO, M. B. Reservoir management: an opinion to how the scientific community can contribute. **Acta Limnologica Brasiliensia**, v. 30, p. e301, 2018. https://doi.org/10.1590/s2179-975x13217

CHEN, Z.; HUANG, P.; ZHANG, Z. Interaction between carbon dioxide emissions and eutrophication in a drinking water reservoir: A three-dimensional ecological modeling approach. **Science of the Total Environment**, v. 663, p. 369–379, 2019. https://doi.org/10.1016/j.scitotenv.2019.01.336

CHINYAMA, A.; OCHIENG, G. M.; NHAPI, I.; OTIENO, F. A. O. A simple framework for selection of water quality models. **Reviews in Environmental Science and Bio/Technology**, v. 13, p. 109–119, 2014. https://doi.org/10.1007/s11157-013-9321-3

CID, F. D.; ANTÓN, R. I.; PARDO, R.; VEGA, M.; CAVIEDES-VIDAL, E. Modeling spatial and temporal variations in the water quality of an artificial water reservoir in the semiarid Midwest of Argentina. **Analytica Chimica Acta**, v. 705, p. 243–252, 2011. https://doi.org/10.1016/j.aca.2011.06.013

COLLABORATION FOR ENVIRONMENTAL EVIDENCE. **Guidelines and Standards for Evidence synthesis in Environmental Management 2018**. Available at: www.environmentalevidence.org/information-for-authors. Access: 10 Jul. 2021.

CRESPO, M.; ORSONI, J.; BORTOLI, J.; RAPAPORT, A.; ROUSSEAU, A.; JAUZEIN, V. Optimal discharge locations to refill hydric reservoirs with reused water-application to the jaunay lake case study. **Revue des sciences de l'eau/Journal of Water Science**, v. 31, n. 4, p. 377–385, 2018. https://doi.org/10.7202/1055595ar

CRESSWELL, C. J.; CUNNINGHAM, H. M.; WILCOX, A.; RANDALL, N. P. What specific plant traits support ecosystem services such as pollination, bio-control and water quality protection in temperate climates? A systematic map. **Environmental Evidence**, v. 7, n. 1, p. 1-13, 2018. https://doi.org/10.1186/s13750-018-0120-8

CUNHA-SANTINO, M. B.; BITAR, A. L.; BIANCHINI JR., I. Chemical constraints on new man-made lakes. **Environmental Monitoring and Assessment**, v. 185, n. 12, p. 10177–10190, 2013. https://doi.org/10.1007/s10661-013-3322-0

DIPPONG, T.; MIHALI, C.; GOGA, F.; CICAL, E. Seasonal evolution and depth variability of heavy metal concentrations in the water of Firiza-Strîmtori Lake, NW of Romania. **Studia Ubb Chemia**, v. 62, n. 1, p. 213-228, 2017. https://doi.org/10.24193/subbchem.2017.1.19

DIPPONG, T.; MIHALI, C.; NĂSUI, D.; BERINDE, Z.; BUTEAN, C. Assessment of Water Physicochemical Parameters in the Strîmtori-Firiza Reservoir in Northwest Romania. **Water environment research: a research publication of the Water Environment Federation**, v. 90, n. 3, p. 220–233, 2018. https://doi.org/10.2175/106143017X15054988926578

GUZMAN, C.D.; ZIMALE, F. A.; TEBEBU, T. Y.; BAYABIL, H. K.; TILAHUN, S. A.; YITAFERU, B.; RIENTJES, T. H. M.; STEENHUIS, T. S. Modeling discharge and sediment concentrations after landscape interventions in a humid monsoon climate: The Anjeni watershed in the highlands of Ethiopia. **Hydrological Processes**, v. 31, n. 6, p. 1239-1257, 2017. https://doi.org/10.1002/hyp.11092

HADDAWAY, N. R.; MACURA, B.; WHALEY, P.; PULLIN, A. S. ROSES flow diagram for systematic reviews. Version 1.0. **Figshare**, Online resource, 2018. https://doi.org/10.6084/m9.figshare.5897389

HADDAWAY, N. R.; STYLES, D.; PULLIN, A. S. Evidence on the environmental impacts of farmland abandonment in high altitude/mountain regions: a systematic map. **Environmental Evidence**, v. 3, n. 1, p. 1-19, 2014. https://doi.org/10.1186/2047-2382-3-17

HARRIS, T. D.; GRAHAM, J. L. Predicting cyanobacterial abundance, microcystin, and geosmin in a eutrophic drinking-water reservoir using a 14-year dataset. **Lake and reservoir management**, v. 33, n. 1, p. 32-48, 2017. https://doi.org/10.1080/10402381.2016.1263694

HEIDARZADEH, N.; MAHDAVI, H.; YAGHOUTI, M. Reservoir water-quality simulation using simplified mathematical models (case study: Seymareh Reservoir). **Marine and Freshwater Research**, v. 76, 2021. https://doi.org/10.1071/MF20334.

KARAMI, H.; FARZIN, S.; JAHANGIRI, A.; EHTERAM, M.; KISI, O.; EL-SHAFIE, A. Multi-Reservoir System Optimization Based on Hybrid Gravitational Algorithm to Minimize Water-Supply Deficiencies. **Water Resources Management**, v. 33, n. 8, p. 2741-2760, 2019. https://doi.org/10.1007/s11269-019-02238-3

KERACHIAN, R.; KARAMOUZ, M. A stochastic conflict resolution model for water quality management in reservoir–river systems. **Advances in Water Resources**, v. 30, n. 4, p. 866-882, 2007. https://doi.org/10.1016/j.advwatres.2006.07.005

LATIF, S. D.; BIRIMA, A. H.; AHMED, A. N.; HATEM, D. M.; AL-ANSARI, N.; FAI, C. M.; EL-SHAFIE, A. Development of prediction model for phosphate in reservoir water system based machine learning algorithms. **Ain Shams Engineering Journal**, v. 13, n. 1, p. 101523, 2022. https://doi.org/10.1016/j.asej.2021.06.009

MARTINS, T. S.; CARMO JUNIOR, G. N. R. Avaliação de Impacto Ambiental: Uma Revisão Sistemática sob a Ótica Metodológica. **E&S Engineering and Science**, v. 7, p. 29–41, 2018. https://doi.org/10.18607/ES201876616

NOVER, D. M.; DOGAN, M. S.; RAGATZ, R.; BOOTH, L.; MEDELLÍN-AZUARA, J.; LUND, J. R.; VIERS, J. H. Does More Storage Give California More Water? **JAWRA Journal of the American Water Resources Association**, v. 55, n. 3, p. 759–771, 2019. https://doi.org/10.1111/1752-1688.12745

ORLOB, G. T. Water-Quality Modeling for Decision Making. **Journal of Water Resources Planning and Management**, v. 118, n. 3, p. 295-307, 1992. https://doi.org/10.1061/(ASCE)0733-9496(1992)118:3(295)

OUZZANI, M.; HAMMADY, H.; FEDOROWICZ, Z.; ELMAGARMID, A. 2016. Rayyan—a web and mobile app for systematic reviews. **Systematic Reviews**, v. 5, n. 1, p. 1-10, 2016. https://doi.org/10.1186/s13643-016-0384-4

ROSA, L.; CHIARELLI, D. D.; RULLI, M. C.; DELL'ANGELO, J.; D'ODORICO, P. Global agricultural economic water scarcity. **Science Advances**, v. 6, n. 18, p. eaaz6031, 2020. https://doi.org/10.1126/sciadv.aaz6031

SINISCALCHI, A. G.; DI MAGGIO, J.; ESTRADA, V.; SOLEDAD DIAZ, M. Integrated mathematical models for drinking water reservoirs and constructed wetlands as a tool for restoration planning. **Journal of Hydrology**, v. 586, p. 124867, 2020. https://doi.org/10.1016/j.jhydrol.2020.124867

SINISCALCHI, A. G.; FRITZ, L.; PRIETO, G. C.; ESTRADA, V.; HOFFMEYERM M.; LARA, R. J.; DIAZ, M. S. Modelling and advanced dynamic optimisation strategies for hydrological and water quality management in continental water bodies. **Computer Aided Chemical Engineering**, v. 43, p. 271-277, 2018. https://doi.org/10.1016/B978-0-444-64235-6.50048-6

SHI, Y.; ZHANG, L.; LI, Y.; ZHOU, L.; ZHOU, Y.; ZHANG, Y.; HUANG, C.; LI, H.; ZHU, G. Influence of land use and rainfall on the optical properties of dissolved organic matter in a key drinking water reservoir in China. **Science of the Total Environment**, v. 699, p. 134301, 2020. https://doi.org/10.1016/j.scitotenv.2019.134301

THIBODEAUX, L. J.; AGUILAR, L. Kinetics of peat soil dissolved organic carbon release to surface water. Part 2. A chemodynamic process model. **Chemosphere**, v. 60, n. 9, p. 1190-1196, 2005. https://doi.org/10.1016/j.chemosphere.2005.02.047

WARD, F. A.; LYNCH, T. P. Integrated river basin optimization: modeling economic and hydrologic interdependence 1. **Journal of the American Water Resources Association**, v. 32, n. 6, p. 1127-1138, 1996. https://doi.org/10.1111/j.1752-1688.1996.tb03483.x

WANG, Q.; LI, S.; JIA, P.; QI, C.; DING, F. A Review of Surface Water Quality Models. **The Scientific World Journal**, v. 2013, p. 1-7, 2013. https://doi.org/10.1155/2013/231768

WESTPHAL, K. S.; CHAPRA, S.C.; SUNG, W. Modeling TOC and UV-254 absorbance for reservoir planning and operation. **Journal of the American Water Resources Association**, v. 40, n. 3, p. 795-809, 2004. https://doi.org/10.1111/j.1752-1688.2004.tb04459.x

WINTON, R. S.; CALAMITA, E.; WEHRLI, B. Reviews and syntheses: Dams, water quality and tropical reservoir stratification. **Biogeosciences**, v. 16, n. 8, p. 1657-1671, 2019. https://doi.org/10.5194/bg-16-1657-2019

XU, C.; ZHANG, J.; BI, X.; XU, Z.; HE, Y.; GIN, K. Y. H. Developing an integrated 3D-hydrodynamic and emerging contaminant model for assessing water quality in a Yangtze Estuary Reservoir. **Chemosphere**, v. 188, p. 218-230, 2017. https://doi.org/10.1016/j.chemosphere.2017.08.121

YUAN, L.; SINSHAW, T.; FORSHAY, K, J. Review of Watershed-Scale Water Quality and Nonpoint Source Pollution Models. **Geosciences**, v. 10, n. 1, p. 25, 2020. https://doi.org/10.3390/geosciences10010025

ZIEMIŃSKA-STOLARSKA, A.; SKRZYPSKI, J. Review of Mathematical Models of Water Quality. **Ecological Chemistry and Engineering**, v. 19, n. 2, p. 197-211, 2012. https://doi.org/10.2478/v10216-011-0015-x

ZHOU, B.; XU, Y.; VOGT, R. D.; LU, X.; LI, X.; DENG, X.; YUE, A.; ZHU, L. Effects of Land Use Change on Phosphorus Levels in Surface Waters-a Case Study of a Watershed Strongly Influenced by Agriculture. **Water Air and Soil Pollution**, v. 227, n. 5, p. 1-14, 2016. https://doi.org/10.1007/s11270-016-2855-6

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Evaluation of biological wastewater treatment in stabilization lagoons from Punta Carnero, Salinas – Ecuador

ARTICLES doi:10.4136/ambi-agua.2822

Received: 17 Dec. 2021; Accepted: 07 Mar. 2022

Juan José Humanante Cabrera[1*]; Lucrecia Cristina Moreno Alcivar[1,2]; Carlos Alberto Deza Navarrete[1]; Ana Mercedes Grijalva Endara[1,3]; Juan Humanante Moreno[4]; Joan Alberto Suárez Tomalá[2]

[1]Graduate School. National University of Tumbes, University Avenue, s/n, 24001, Pampa Grande, Tumbes, Peru. E-mail: lmoreno@upse.edu.ec, cdezan@untumbes.edu.pe, ana.grijalvae@ug.edu.ec
[2]Department Engineering Sciences Faculty. Santa Elena Peninsula State University, Santa Elena Principal Avenue, 240350, La Libertad, Santa Elena, Ecuador. E-mail: lmoreno@upse.edu.ec, joan.suareztomala@upse.edu.ec
[3]Department Chemical Sciences Faculty. Guayaquil University, Avenue 10 NO Guayaquil, 090514, Guayaquil, Guayas, Ecuador. E-mail: ana.grijalvae@ug.edu.ec
[4]Postgraduate master's degree. Guayaquil University, Avenue 10 NO Guayaquil, 090514, Guayaquil, Guayas, Ecuador. E-mail: juan.humanante@eppetroecuador.ec
[*]Corresponding author. E-mail: jhumanante@hotmail.com

ABSTRACT

This research evaluated the wastewater treatment system of the Punta Carnero sector, in relation to pollutant efficiency load removal, final effluent quality and impact on the ecosystem, and finally to determine if the final discharge can be reused for agricultural irrigation. The research was based on the affluent and effluent characterization of the system, carried out in three phases: i) Taking of simple samples, analyzed in an accredited water laboratory and analysis of the contaminant loads efficiency; ii) Review of results compared to the Table of "Discharge limits to a freshwater receiving body"; iii) Examination of results based on the "Water Quality Criteria for Agricultural Irrigation" Table of the Ecuadorian regulation TULSMA (2015). BOD (62.42%), COD (62.41%) and FC (53.58%) removal efficiencies did not comply with current Ecuadorian regulations. The quality of the effluent with respect to the parameters evaluated for discharges to a freshwater receiving body denoted a non-optimal quality of final discharge, negatively impacting the ecosystem. Finally, the evaluation determined parameters that exceed the water quality criteria for agricultural irrigation allowed: Oils-Fats (5.65 mg/l), FC (62,900 NMP/100ml), Hg (0.00141 mg/l), OD (8.86 mg/l). After evaluating the wastewater treatment system, it was determined that the pollutant load removal efficiency and effluent quality is not optimal for discharge into a receiving water body, so it's not suitable for reuse in agricultural irrigation.

Keywords: affluent, effluent, water quality.

Avaliação do tratamento biológico de águas residuais em lagoas de estabilização de Punta Carnero, Santa Elena – Equador

RESUMO

O objetivo da pesquisa foi avaliar o sistema de tratamento de águas residuais do setor de

Punta Carnero, em relação à sua eficiência na remoção de cargas poluentes, a qualidade do efluente final e seu impacto no ecossistema e, finalmente, determinar se o descarte final pode ser reutilizado para irrigação agrícola. A pesquisa consistiu na caracterização da influência e do efluente do sistema, desenvolvido em três fases: i) Retirada de amostras simples, testadas em um laboratório de água credenciado e análise da eficiência das cargas poluentes; ii) Revisão dos resultados comparados com a Tabela de "Limites de descarga para um corpo receptor de água doce", iii) Exame dos resultados de acordo com a Tabela de "Critérios de qualidade da água para irrigação agrícola" da regulamentação equatoriana TULSMA (2015). A eficiência de remoção da BDO (62,42%), COD (62,41%) e CF (53,58%) não estão em conformidade com os regulamentos atuais do Equador. A qualidade do efluente com respeito aos parâmetros avaliados para descargas em um corpo receptor de água doce, denotam uma qualidade não ideal de descarga final, afetando negativamente o ecossistema; finalmente, a avaliação determinou parâmetros que excedem os critérios de qualidade da água para irrigação agrícola: Óleos-Gorduras (5,65 mg/l), CF (62900 NMP/100ml), Hg (0,00141 mg/l), OD (8,86 mg/l). Após avaliação do sistema de tratamento de águas residuais, foi determinado que sua eficiência na remoção de cargas poluentes e sua qualidade de efluentes não é ótima para ser descarregada em um corpo de água receptor, portanto, não é adequada para reutilização na irrigação agrícola.

Palavras-chave: afluente, efluente, qualidade da água.

1. INTRODUCTION

In Latin America, the application of biological treatments to treat wastewater of domestic origin is more readily available (Vargas *et al.*, 2020). This treatment is known as "conventional" because it's common and widely used, which usually involves low costs (Roy and Saha, 2021). The treatments are simple, efficient and cost-effective, and use microorganisms to degrade a large part of biodegradable waste from wastewater effluents (Al-Qodah *et al.*, 2020; Jung and Pauly, 2011). The main objectives of these treatments are the elimination of pathogenic microorganisms, suspended solids and the reduction of organic matter to an acceptable level (Grigorieva *et al.*, 2013; Leite *et al.*, 2005).

Stabilization lagoons are passive systems built to treat wastewater by biological processes (Florentino *et al.*, 2019); they are used for the upgrading of liquid effluents from domestic, agricultural and industrial sources (Jimoh *et al.*, 2019; Nuñez and Fragoso, 2020). Due to the low implementation cost, ease of operation, minimal energy consumption and high efficiency in the reduction of pathogenic organisms, this is the type of wastewater treatment most used in underdeveloped countries (Araújo *et al.*, 2016; Li *et al.*, 2018; Romero and Castillo, 2018). There are generally three types of lagoon systems: anaerobic, facultative and maturation or aerobic, each with different design and treatment characteristics (Dos Santos and Van Haandel, 2021; Matsumoto and Sánchez, 2016). Anaerobic lagoons operate in the absence of oxygen and have depths of 3 to 5 meters (Perú, 2007; Cortés *et al.*, 2017); facultative lagoons decompose the Biochemical Oxygen Demand (BOD) or organic matter (Sánchez and Matsumoto, 2013), by aerobic, anaerobic and facultative bacteria, with depths ranging from 1.5 to 2.5 meters (Sánchez *et al.*, 2011; Treviño and Cortés, 2016). Maturation lagoons are used at the end of treatment, and their depths vary from 0.5 to 1.5 meters (Cárdenas *et al.*, 2005; Tilley, 2011). The efficiency of stabilization lagoons depends mainly on factors such as depth, hydraulic retention time, temperature, bacteria and algae (Bezerra *et al.*, 2020).

González and Chiroles (2011) argue that for every liter of wastewater, at least eight liters of freshwater are polluted, so it is important to consider that untreated or inadequately treated effluents are the main source of pollution to natural water bodies (Cedeño, 2020), changing their chemistry and seriously altering the ecosystems that depend on them (Lahera, 2010).

Wastewater treatment plants are built to reduce the impact of polluted effluents (Morera *et al.*, 2017); however, many facilities leave aside the external effects that may occur in the operation and maintenance phase in each of these production units (Hernández *et al.*, 2017), so that the implementation of environmental regulations and standards are presented in order to define water quality criteria for discharge and ensure the non-contamination of water resources (Marçal and Silva, 2017). In Ecuador, problems have arisen due to the absence of sufficient treatment and physical infrastructure (Montero *et al.*, 2020), and it's estimated that only 10% of wastewater is discharged into bodies of water with some type of effluent treatment (Sato *et al.*, 2013; Velasco *et al.*, 2019). Ecuadoran public institutions charged with regulating wastewater discharge include the Environmental Management Departments of Provincial Prefectures and Mayor's Offices to the highest environmental authority in the country, which is the Ministry of the Environment (Peña *et al.*, 2018). The reuse of wastewater for agricultural irrigation has advantages associated with the improvement of the fertility of agricultural soils by providing organic matter (Silva *et al.*, 2018), but the effluent discharged must be evaluated from the agronomic and bacteriological point of view (Cisneros and Saucedo, 2016). Therefore, TULSMA (2015) states that the use of wastewater for irrigation is prohibited, except for that which has been previously treated and complies with the water quality levels for agricultural irrigation.

The scientific relevance of this research is characterized by exposing the high degree of contamination of the treated wastewater and potential risk to human contamination and environmental degradation through the analysis of Biochemical Oxygen Demand, Chemical Oxygen Demand, Total Nitrogen and Fecal Coliforms that can be harmful to the habitants, flora, and fauna of the study area. The wastewater treatment system using stabilization lagoons is located on Punta Carnero Road in Salinas Canton, Province of Santa Elena – Ecuador. It currently has a system of seven lagoons: three anaerobic, two facultative and two aerobic or maturation, which discharges its previously treated effluent into the Achayan River, which finally drains to Punta Carnero Beach towards the Pacific Ocean (Humanante, 2016). This system is part of the eight lagoon systems in the Santa Elena Province under the responsibility of the company AGUAPEN-EP, which has been providing services to the province since 1998 (Suárez and Panchana, 2021). This research evaluated the affluent and effluent of the wastewater treatment efficiency, the possible effects on the ecosystem produced by the quality of the final effluent, and finally determined if the final discharge is optimal for agricultural irrigation. The investigation was based on water quality analyses performed in an accredited water laboratory, through simple samples of wastewater taken at the inlet and outlet of the lagoon system on April 15, April 28 and July 6, 2021. The results of water quality tests were compared with the Unified Text of Secondary Environmental Legislation of Ecuador in relation to the General Standards Table for effluent discharge into freshwater bodies, which regulates the following parameters: Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Nitrogen (N) and Phosphorus (P). Finally, the results were also compared with water quality criteria for agricultural or irrigation water table, which regulates the following parameters: Oils-Fats, Aluminum, Arsenic, Beryllium, Boron, Cadmium, Zinc, Cobalt, Copper, Fecal Coliforms, Chromium, Fluoride, Iron, Parasite Eggs, Lithium, Floating Material, Mercury, Manganese, Molybdenum, Nickel, Nitrites, Dissolved Oxygen, pH, Lead, Selenium, Sulfates and Vanadium.

2. MATERIAL AND METHODS

To carry out the investigation, the characteristics of the raw wastewater must first be known (Table 1). The research was developed in the wastewater treatment system of the Punta Carnero Sector in Salinas Canton, Santa Elena Province (Ecuador), located at UTM coordinates 509066 E, 9750151 N. Wastewater arriving to the system comes from the cantons of La

Libertad and Salinas, which in turn is collected and transported through sanitary sewer networks to the wastewater treatment system. At present, the Stabilization lagoon system is composed of three anaerobic lagoons, two facultative and two aeration or maturation lagoons, which are inadequate for the current population. Biological treatment processes are carried out in the lagoons: anaerobic treatment in anaerobic lagoons and aerobic treatment in facultative and maturation lagoons, which are connected to each other by means of 400 mm diameter PVC pipes through distribution chambers; finally, the treated water passes through a chlorination labyrinth and its final effluent is discharged into a natural waterway (Humanante, 2016). The research was carried out in three phases.

Table 1. Characteristics of the raw wastewater.

Characteristics of wastewater	Description
Physical	Turbidity Color Odor Total solids Temperature
Chemical	Chemical Oxygen Demand (COD) Total Organic Carbon (TOC) Nitrogen Phosphorus Chlorides Sulfates Alkalinity pH Heavy Metals Trace Elements Priority Pollutants Chemical Oxygen Demand (COD) Total Organic Carbon (TOC) Biochemical Oxygen Demand (BOD) Oxygen required for nitrification
Biological	Microbial population

Source: Wijaya and Soedjono (2018).

2.1. Pollutant load removal efficiency

The first phase involves determining the value of the pollutant load removal efficiency of the wastewater treatment system. First, two simple samples were taken in the lagoon system (Figure 1), simple samples representing the composition of the water for that specific time and location (Romero, 2004). The first simple sample was taken in the affluent of the system, water distribution chamber with outlet to anaerobic lagoons at UTM Coordinates 509203 E 9750365 N; the second simple sample was taken in the effluent of the water outlet system of the Parshall Flume with discharge to the Achayan River at UTM Coordinates 509042 E 9749661 N. Next, the samples taken in glass and plastic bottles of one liter capacity were refrigerated at a temperature of 4.3ºC, and one milliliter of sulfuric acid was added to the glass bottle. The samples were then transported to an accredited water testing laboratory, where they were analyzed and finally the test results were received approximately fifteen days after the samples were taken and left. Third, once the test results were obtained, the results were processed using the equation (Romero, 2004) to determine the pollutant load removal efficiency of the wastewater treatment system (Equation 1).

$$E(\%) = \frac{(S_0 - S)}{S_0} \times 100 \quad (1)$$

The equation shows: E (%) is the pollutant load removal efficiency in percent; S_0 is the affluent pollutant load in milligrams per liter (mg/l) and S is the effluent pollutant load in milligrams per liter (mg/l).

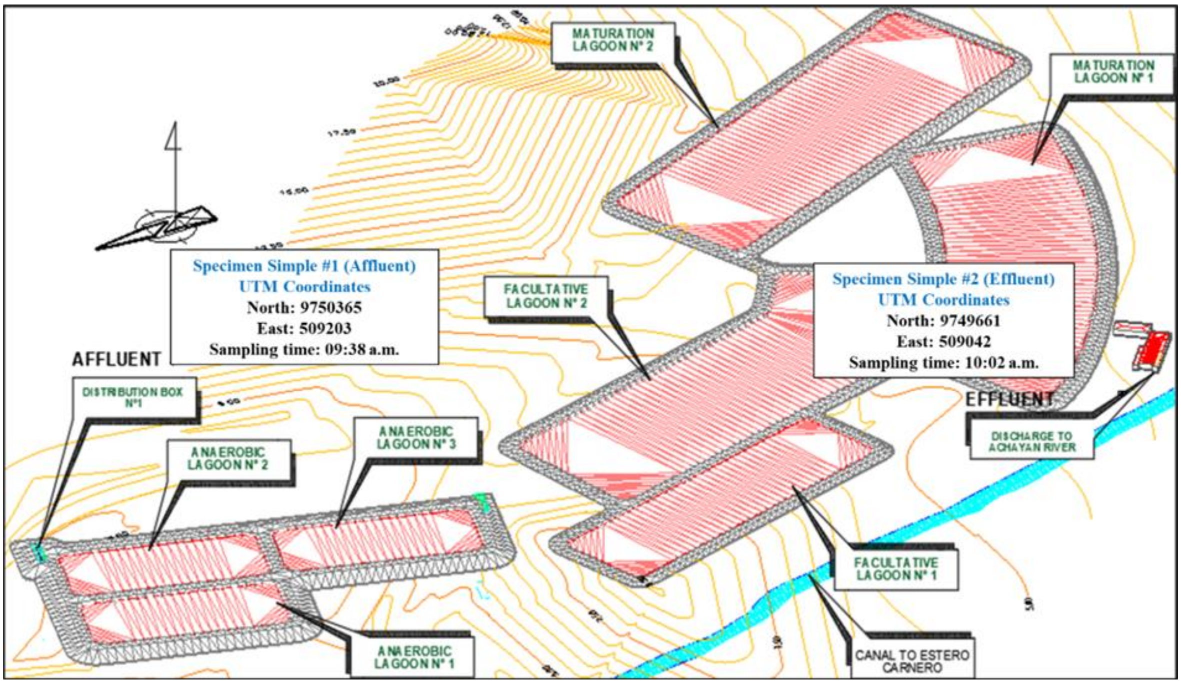


Figure 1. Wastewater treatment system of the Punta Carnero sector, Salinas - Ecuador (Humanante, 2016), simple sampling location.

2.2. Effluent water quality for discharge to a freshwater body

To determine the quality of effluent from the wastewater treatment system, in this second phase, simple samples were taken from the final effluent (discharge to the Achayan River) on three dates:

S1: Simple sample taken directly from the discharge, packaged in two plastic bottles and one glass bottle, both with a capacity of one liter (April 15, 2021).

S2: Simple sample taken directly from the discharge, packaged in two plastic bottles and one glass bottle, both with a capacity of one liter (April 28, 2021).

S3: Simple sample taken directly from the discharge, packaged in two plastic bottles and one glass bottle, both with a capacity of one liter (July 6, 2021).

After the results of the tests of the simple samples taken at the accredited laboratory, “Grupo Quimico Marcos” were obtained; these data were processed in the MINITAB statistical software to determine the mean and standard deviation of the system final effluent. Once these indicators were obtained and represented graphically S1, S2 and S3, the quality of the water discharged into the Achayan River was determined. Once the quality of the discharged water was known, the general standard for effluent discharge to freshwater bodies was presented in relation to the parameters evaluated in the laboratory: Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Total Nitrogen (N) and Total Phosphorus (P), and the possible effects caused by poor wastewater treatment and its impact on the environment (Table 2).

Table 2. General standard for effluent discharge to freshwater bodies and possible effects of exceeding the maximum permissible limits in the discharge of treated wastewater to the Achayan River (freshwater body).

Parameters	Expressed as	Unity	M.A.L.*	Possible effects
Biochemical Oxygen Demand	BDO	mgO_2/l	100	Damage to flora and fauna (Raffo and Ruiz, 2014).
Chemical Oxygen Demand	COD	mgO_2/l	200	Absence of oxygen in aquatic organisms (Mayta and Mayta, 2017).
Total Nitrogen	N	mg/l	10	Affects plant and animal life (Espinosa *et al.*, 2013).
Total Phosphorus	P	mg/l	50	Algae growth and absence of oxygen (Reyes *et al.*, 2017).

*Maximum Allowable Limit of discharge to a freshwater body.
Source: Section 5.2.4. from TULSMA (2015).

2.3. Effluent water quality for agricultural irrigation water reuse

The third phase of the research was to determine if the water that had been treated and then discharged to the Achayan River (freshwater body) can be destined for water reuse for agricultural irrigation. For this purpose the guidelines established in the Unified Text of Secondary Environmental Legislation of Ecuador 2015 must be followed, where TULSMA (2015) states that water for agricultural use is understood as that used for crop irrigation and other related or complementary activities established by the competent bodies.

In this phase, the results obtained from the accredited water laboratory are compared with the water quality criteria for agricultural irrigation (Table 3); this table also mentions the environmental impacts caused by poor quality water reused for agricultural irrigation (Perú, 2006).

Table 3. Water quality criteria for agricultural irrigation

Parameters	Expressed as	Unity	Quality Criterion	Environmental Impacts*
Aluminum	Al	mg/l	5.0	Plant growth impacts
Arsenic	As	mg/l	0.1	Harvest yield impacts
Beryllium	Be	mg/l	0.1	Plant growth impacts
Boron	B	mg/l	0.75	Harvest yield impacts
Cadmium	Cd	mg/l	0.05	Plant disorders
Chromium	Cr	mg/l	0.1	Unacceptable harvests
Cobalt	Co	mg/l	0.01	Plant disorders
Copper	Cu	mg/l	0.2	Plant growth impacts
Dissolved Oxygen	OD	mg/l	3	Decreased oxygen
Fecal Coliforms	CF	NMP/100ml	1000	Diseases to the population
Floating Matter	-	-	Absence	Harvest yield impacts
Fluorine	F	mg/l	1.0	Plant growth impacts
Iron	Fe	mg/l	5.0	Plant growth impacts
Lead	Pb	mg/l	5.0	Plant growth impacts
Lithium	Li	mg/l	2.5	Harvest yield impacts
Manganese	Mn	mg/l	0.2	Plant disorders
Mercury	Hg	mg/l	0.001	Food chain problems

Continue...

Continued...

Molybdenum	Mo	mg/l	0.01	Plant growth impacts
Nickel	Ni	mg/l	0.2	Plant disorders
Nitrites	NO_2	mg/l	0.5	Eutrophication
Oils and Fats	Visible	mg/l	Absence	Plant growth impacts
Parasite Eggs	-	-	Absence	Diseases to the population
pH	pH	-	6-9	Plant growth impacts
Selenium	Se	mg/l	0.02	Harvest yield impacts
Sulfates	SO_4^2	mg/l	250	Plant leaf burns
Vanadium	V	mg/l	0.1	Plant growth impacts
Zinc	Zn	mg/l	2	Plant growth impacts

***Environmental Impacts (Perú, 2006).**
Source: Section 5.1.3. from TULSMA (2015).

3. RESULTS AND DISCUSSION

The research shows the following results based on the evaluation carried out in three phases:

3.1. Pollutant load removal efficiency results

The pollutant load removal efficiency (Table 4) was determined based on the affluent, which is the incoming pollutant load (S_0), and the effluent, which is the outgoing pollutant load (S) of the wastewater treatment system evaluated. A total of 28 parameters were evaluated, each of which was analyzed in the laboratory and the results obtained from the tests were processed in Excel software (Table 3), A graph of individual values is also shown, classifying the parameters that denoted an efficiency greater than 50%, less than 50%, no efficiency (0%) and negative values that indicate that the system is greater than that entering it (Figure 2).

Table 4. Pollutant load removal efficiency.

Parameters	Unity	Affluent S_0	Effluent S	Removal Efficiency (%)
Aluminum	mg/l	1.084	0.002	99.82
Arsenic	mg/l	0.00215	0.00215	0.00
Beryllium	mg/l	0.001	0.002	-100.00
Boron	mg/l	0.391	0.373	4.60
Cadmium	mg/l	0.001	0.001	0.00
Chromium	mg/l	0.003	0.001	66.67
Cobalt	mg/l	0.00041	0.00041	0.00
Copper	mg/l	0.035	0.00116	96.69
BDO	mgO_2/l	379.2	142.5	62.42
Dissolved Oxygen	mg/l	5.15	8.86	-72.04
COD	mgO_2/l	632.63	237.8	62.41
Fecal Coliforms	NMP/100ml	135500	62900.0	53.58
Fluorine	mg/l	1.01	0.57	43.56
Iron	mg/l	0.984	0.034	96.54
Lead	mg/l	0.005	0.003	40.00
Lithium	mg/l	0.01	0.009	10.00
Manganese	mg/l	0.088	0.066	25.00
Mercury	mg/l	0.00141	0.00141	0.00

Continue...

Continued...				
Molybdenum	mg/l	0.001	0.001	0.00
Nickel	mg/l	0.004	0.00123	69.25
Nitrites	mg/l	0.115	0.115	0.00
Oils and Fats	mg/l	9.76	5.65	42.11
Selenium	mg/l	0.00306	0.00306	0.00
Sulfates	mg/l	165	94	43.03
Total Phosphorus	mg/l	6.3	5.9	6.35
Total Nitrogen	mg/l	53	56	-5.66
Vanadium	mg/l	0.003	0.002	33.33
Zinc	mg/l	0.156	0.023	85.26

Ecuadorian regulations only regulate three parameters for pollutant load removal efficiency. The removal of BOD and COD was 62.42% and 62.41%, respectively, which indicates that the efficiency of these parameters doesn't reach the values established in current Ecuadorian regulations (INEN, 2012), where it's established that there must be a 70-85% removal rate in order for the treatment in the lagoon system to be considered good. On the other hand, in the microbiological parameter of Fecal Coliforms, the removal efficiency was 53.58%, which doesn't meet the requirements of the Environmental Secondary Legislation Text of 2015, where it provides for a removal of 99.9% of this pollutant, which relates that the treatment is not adequate.

The removal efficiency of COD and BOD of about 60% is relatively insufficient especially if the initial COD values in the raw water is high. This usually suggests the use of a pretreatment step before the biological step (Al-Qodah *et al*, 2019).

Heavy metal removal efficiencies were obtained in the treatment system (Figure 2), in which: Al (99.82%), Cu (96.69%), Fe (96.54%), Zn (85.26%), Ni (69.25%), Cr (66.67%), BOD (62.42%), COD (62.41%) and CF (53.58%), had a pollutant load removal greater than 50%; F (43.56%), SO_4 (43.03%), Oils and Fats (42.11%), Pb (40.00%), V (33.33%), Mn (25.00%), Li (10.00%) and B (4.60%), its pollutant load removal was below 50%; As, Cd, Co, Hg, Mo, NO_2 and Se, no removal of pollutant load; finally, N (-5.66%), OD (-72.04) and Be (-100.00), presented an inefficient removal in their treatment since the final load increased in relation to their initial load.

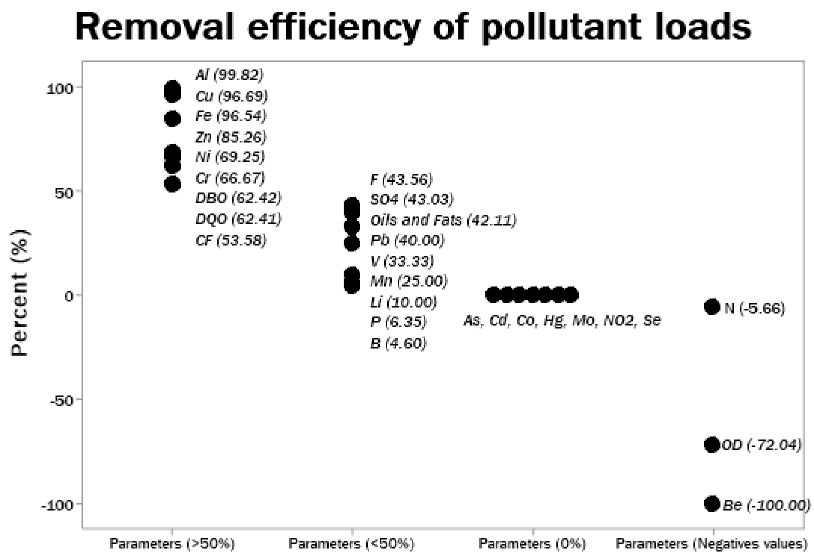


Figure 2. Removal Efficiency of pollutant loads.

3.2. Final effluent discharge quality results

It was determined that the effluent quality is not optimal for discharge with respect to BOD, COD and Total Nitrogen, while Total Phosphorus complies with present regulations.

The effluent quality of BOD (Figure 3), shows that it doesn't comply with the maximum permissible limit of 100 milligrams of oxygen per liter, which means that exceeding the established limit would cause negative effects to the flora and fauna of the sector where the water is discharged. On the other hand, the effluent quality of COD (Figure 4), like BOD, doesn't comply with the 200 milligrams of oxygen per liter established in the mentioned regulation, which denotes possible negative effects such as the absence of oxygen in aquatic microorganisms in the sector; Similarly, Total Nitrogen levels exceed the maximum permissible limit of 50 milligrams per liter (Figure 5), which indicates possible effects such as the vulnerability of plants and animals to high levels of N. However, the effluent quality of Total Phosphorus (Figure 6) complies with the maximum permissible limit of 10 milligrams per liter, which indicates that there are no problems in its discharge.

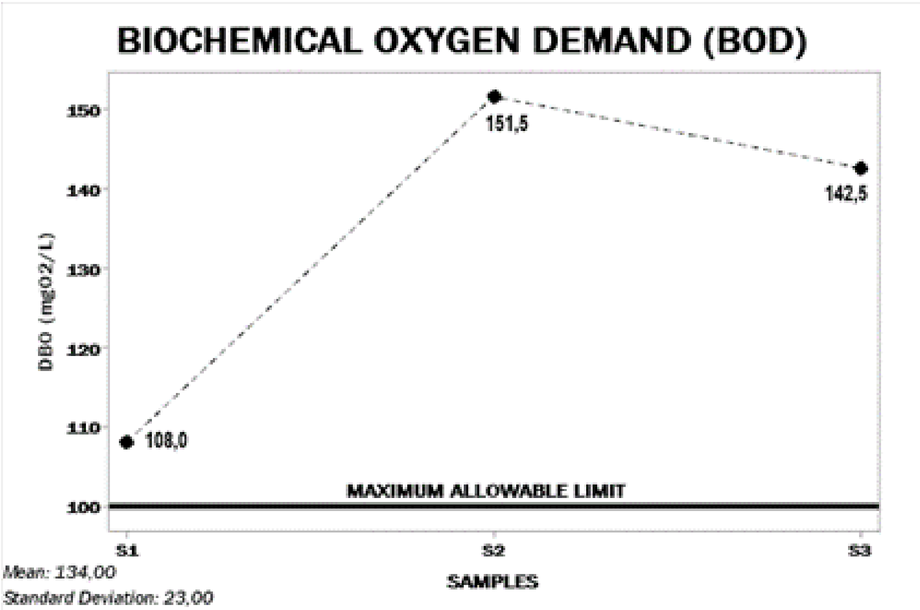


Figure 3. BOD Effluent Quality.

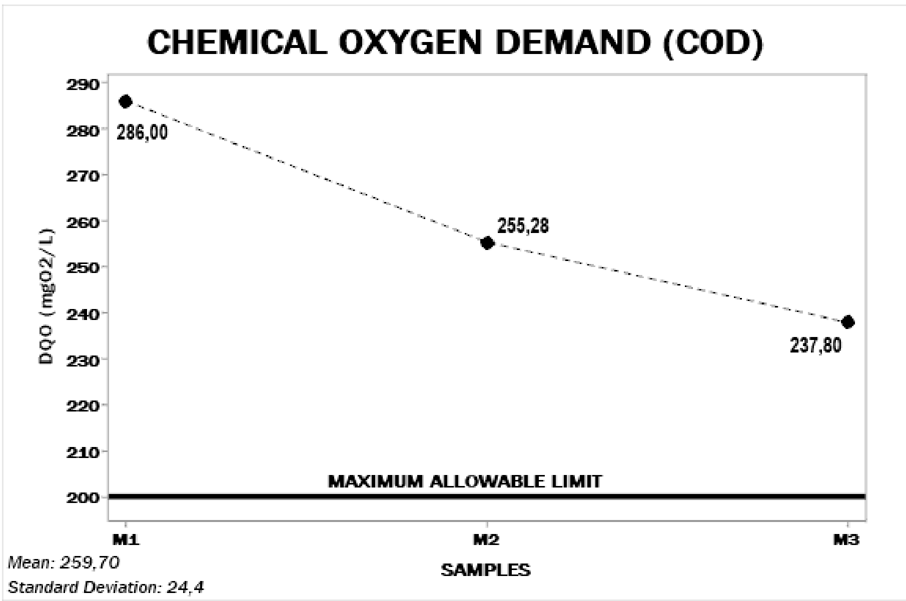


Figure 4. COD Effluent Quality.

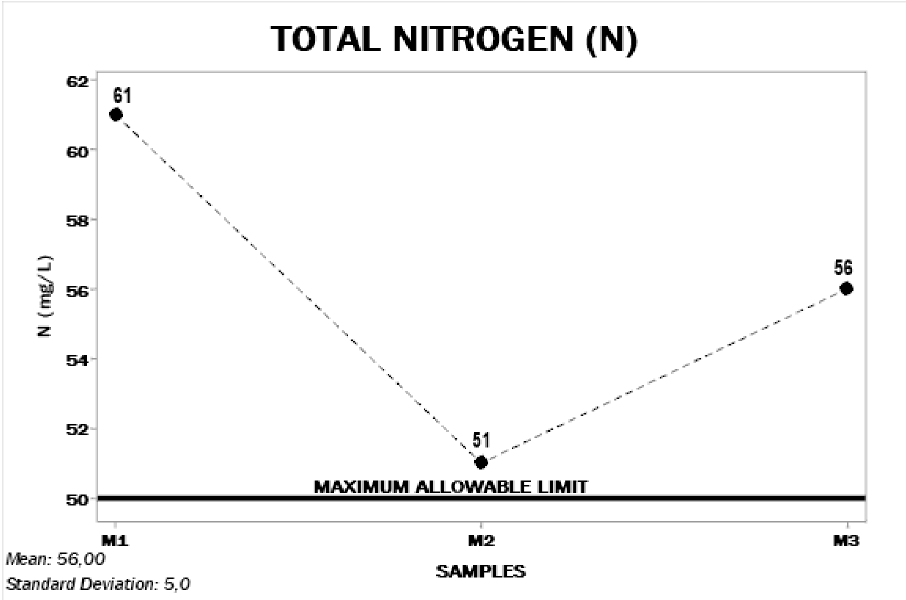


Figure 5. N Effluent Quality.

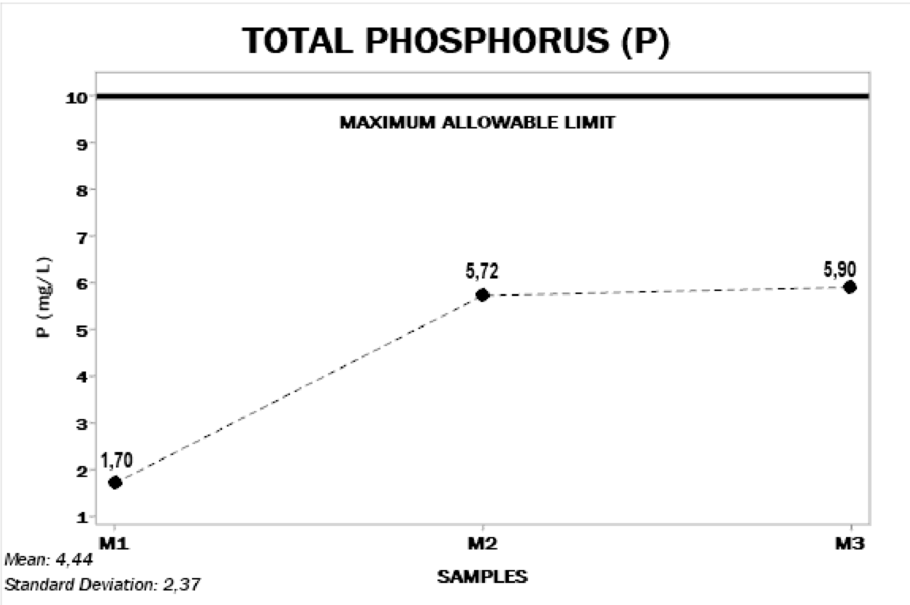


Figure 6. P Effluent Quality.

3.3. Results comparison obtained in the laboratory

Of the 27 parameters evaluated with the agricultural irrigation water quality criteria of the Unified Text of Secondary Environmental Legislation, the following comply with the requirements of the regulations Al (0.002 < 5.0), As (0.00215 < 0.1), Be (0.002 < 0.1), B (0.373 < 0.75), Cd (0.001 < 0.05), Cr (0.001 < 0.1), Co (0.00041 < 0.01), Cu (0.00116 < 0.2), Floating Matter (Absence), F (0.57 < 1), Fe (0.034 < 5), Pb (0.003 < 5), Li (0.009 < 2.5), Mn (0.066 < 0.2), , Mo (0.001 < 0.01), Ni (0.00123 < 0.2), NO_2 (0.115 < 0.5), Parasite eggs (Absence), pH (7.88 < 6-9), Se (0.00306 < 0.02), Sulfates (94 < 250), V (0.002 < 0.1) and Zn (0.023 < 2), while OD (8.86 > 3), Fecal Coliforms (62900 > 1000), Hg (0.00141 > 0.001), and Oils and Fats (5.65 > Absence) exceed agricultural irrgiation water quality criteria.

4. CONCLUSIONS

The removal efficiency of the wastewater treatment system doesn't meet the efficiencies established in existing regulations, so it is determined that the biological treatment is not optimal. It is proposed as an improvement to achieve these efficiencies that there should be more frequent control of the stabilization lagoons, as well as more frequent removal of the silt that is present in these lagoons, which doesn't allow the water entering the system to be treated in an efficient manner.

According to the results of the study, the quality of the water discharged into the receiving body (Achayan River) is poor, which could have serious consequences in the future, since high concentrations of BOD, COD and N directly affect plants and animals in the sector, and indirectly affect people who bathe in the sea.

The treated wastewater is not suitable for reuse for agricultural irrigation because it doesn't fully comply with the criteria established in Table 2 of the Ecuadorian Environmental Regulations.

5. REFERENCES

AL-QODAH, Z.; AL-QUDAH, Y.; ASSIREY, E. Combined biological wastewater treatment with electrocoagulation as a post-polishing process: A review. **Separation Science and Technology**, v. 55, n. 13, p. 2334-2352, 2020. https://doi.org/10.1080/01496395.2019.1626891

AL-QODAH, Z.; AL-QUDAH, Y.; OMAR, W. On the performance of electrocoagulation-assisted biological treatment processes: a review on the state of the art. **Environmental Science and Pollution Research,** v. 26, n. 28, p. 28689-28713, 2019. https://doi.org/10.1007/s11356-019-06053-6

ARAÚJO, G.; LIMA, I.; ARAÚJO, A.; SILVA, M. Avaliação experimental e modelagem matemática de filtros anaeróbios como alternativa de baixo custo para remoção de algas de efluentes de lagoas de estabilização. **Engenharia Sanitária e Ambiental**, v. 21, n. 4, p. 687-696, 2016. https://doi.org/10.1590/S1413-41522016134641

BEZERRA, E.; DE SOUZA, I.; SAAVEDRA, N. Variability in phytoplankton community structure and influence on stabilization pond functioning. **Revista Ambiente & Água**, v. 15, n. 2, p. 1-13, 2020. https://doi.org/10.4136/ambi-agua.2507

CÁRDENAS, C.; JAEGER, C.; VILLASMIL, H.; PERRUOLO, T.; YABROUDI, S.; LÓPEZ, F.; HERRERA, L.; CASTEJÓN, O. Evaluation of the units that conform the wastewater treatment plant south Maracaibo. **Journal of the Engineering Technical Faculty of the Zulia's University**, v. 28, n. 2, p. 97-109, 2005.

CEDEÑO, H. Analysis of the water quality parameters of the dead river effluent for possible reuse of Manta. **Knowledge Pole: Scientific-Professional Journal**, v. 5, n. 2, p. 579-604, 2020. http://dx.doi.org/10.23857/pc.v5i2.1299

CISNEROS, O.; SAUCEDO, H. Wastewater reuse in agriculture. **Water Technology Institute Mexican**, 2016. 170 p.

CORTÉS, F.; TREVIÑO, A.; ESPINOZA, A.; SÁENZ, A.; ALCORTA, M.; GONZÁLEZ, J. *et al.* Optimization in the design of a wastewater treatment system integrated by three stabilization lagoon. **Water Technology and Sciences**, v. 8, n. 4, p. 139-155, 2017. https://doi.org/10.24850/j-tyca-2017-04-09

DOS SANTOS, S. L.; VAN HAANDEL, A. Transformation of Waste Stabilization Ponds: Reengineering of an Obsolete Sewage Treatment System. **Water 2021**, v. 13, n. 9, p. 1193, 2021. https://doi.org/10.3390/w13091193

ESPINOSA, M.; LEÓN, Y.; RODRIGUEZ, X. Problem of the determination of nitrogen species (total nitrogen and ammonia) in wastewater. **CENIC Journal. Chemical Sciences**, v. 44, p. 1-12, 2013.

FLORENTINO, A.; COSTA, M.; NASCIMENTO, J.; FARES ABDALA, E.; MOTA, C.; DOS SANTOS, A. Identification of microalgae from waste stabilization ponds and evaluation of electroflotation by alternate current for simultaneous biomass separation and cell disruption. **Engenharia Sanitaria e Ambiental**, v. 24, n. 1, p. 177-186, 2019. https://doi.org/10.1590/S1413-41522019193972

GONZÁLEZ, M.; CHIROLES, S. Safe use and microbiological risks of wastewater for agriculture. **Cuban Journal of Public Health**, v. 37, n. 1, p. 61-73, 2011.

GRIGORIEVA, E. V.; BONDARENKO, N. V.; KHAILOV, E. N.; KOROBEINIKO, A. Analysis of optimal control problems for the process of wastewater biological treatment. **Math Journal: Theory and Applications**, v. 20, n. 2, p. 103-118, 2013.

HERNÁNDEZ, A.; QUIMIS, A.; MOLINA, G.; MORENO, L. The treatment of wastewater in Portoviejo Canton and its potential environmental impact. **UNESUM-Sciences: Multidisciplinary Scientific Journal**, v. 1, n. 2, p. 47-58, 2017. https://doi.org/10.47230/unesum-ciencias.v1.n2.2017.17

HUMANANTE, J. **Reduction of coliforms in effluent stabilization lagoon of Salinas - Libertad applying bacteria**. 2016. Thesis (Master's Degree in Sanitary Engineering) - Guayaquil University, Guayaquil, 2016.

INEN. **Normas para estudio y diseño de sistemas de Agua potable y disposición de aguas residuals PPara poblaciones mayores a 1000 habitantes**. Quito, 2012. p. 420.

JIMOH, T. A.; KESHINRO, M. O.; COWAN, K. A. Microalgal–Bacterial Flocs and Extracellular Polymeric Substances: Two Essential and Valuable Products of Integrated Algal Pond Systems. **Water, Air, & Soil Pollution**, v. 230, n. 4, p. 95, 2019. https://doi.org/10.1007/s11270-019-4148-3

JUNG, H.; PAULY, D. 4.19 - Water in the Pulp and Paper Industry. *In:* WILDERER, P. (ed.). **Treatise on Water Science**. Oxford: Elsevier, 2011. p. 667-683.

LAHERA, V. Sustainable Infrastructure: Wastewater Treatment Plants. **Journal of Territorial Studies**, v. 12, n. 2, p. 58-69, 2010.

LEITE, V.; ATHAYDE, G.; SOUSA, J.; LOPES, W.; PRASAD, S.; SILVA, S. Tratamento de águas residuárias em lagoas de estabilização para aplicação na fertirrigação. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 9, n. 1, p. 71-75, 2005. https://doi.org/10.1590/1807-1929/agriambi.v9nsupp71-75

LI, M.; ZHANG, H.; LEMCKERT, C.; ROIKO, A.; STRATTON, H. On the hydrodynamics and treatment efficiency of waste stabilisation ponds: From a literature review to a strategic evaluation framework. **Journal of Cleaner Production**, v. 183, p. 495-514, 2018. https://doi.org/10.1016/j.jclepro.2018.01.199

MARÇAL, D.; SILVA, C. Avaliação do impacto do efluente da estação de tratamento de esgoto ETE-Pirajá sobre o Rio Parnaíba, Teresina (PI). **Engenharia Sanitária e Ambiental**, v. 22, n. 4, p. 761-772, 2017. https://doi.org/10.1590/S1413-41522017148242

MATSUMOTO, T.; SÁNCHEZ, I. Performance of the Sewage Treatment Plant of São João de Iracema (Brazil). **Ingeniería**, v. 21, n. 2, p. 176-186, 2016. https://doi.org/10.14483/udistrital.jour.reving.2016.2.a04

MAYTA, R.; MAYTA, J. Removal of Chromium and Chemical Oxygen Demand of Tannery Wastewater by Electrocoagulation. **Journal of the Peruvian Chemical Society**, v. 83, n. 3, p. 331-340, 2017.

MONTERO, F.; MOLINA, C.; PILLCO, B.; SARDUY, L.; DIÉGUEZ, K. Evaluation of the Environmental Impact of the Wastewater Treatment Plant Construction. Case Pindo Chico River, Puyo, Pastaza, Ecuador. **Science, Environment and Climate**, v. 3, n. 1, p. 23-39, 2020. https://doi.org/10.22206/cac.2020.v3i1.pp23-39

MORERA, S.; COROMINAS, L.; RIGOLA, M.; POCH, M., COMAS, J. Using a detailed inventory of a large wastewater treatment plant to estimate the relative importance of construction to the overall environmental impacts. **Water Research**, v. 122, p. 614-623, 2017. https://doi.org/10.1016/j.watres.2017.05.069

NUÑEZ, J.; FRAGOSO, P. Use of aquatic macroinvertebrates as an evaluation system for stabilizing El Salguero lagoon (Colombia). **Technological Information**, v. 31, n. 3, p. 277-284, 2020. http://dx.doi.org/10.4067/S0718-07642020000300277

PEÑA, S.; MAYORGA, J.; MONTOYA, R. Proposal for the treatment of wastewater from the city of Yaguachi (Ecuador). **Science and Engineering**, v. 39, n. 2, p. 161-167, 2018.

PERÚ. Comisión Nacional del Agua. **Manual de Agua Potable, Alcantarillado y Saneamiento**. Lima, 2007. 234p.

PERÚ. Ministerio de Salud. Dirección General de Salud Ambiental e Inocuidad Alimentaria. **Riego de Vegetales y Bebederos de Animales**. Lima, 2006. 134p.

RAFFO, E.; RUIZ, E. Characterisation of waste water and biochemical oxygen demand. **Industrial Data**, v. 17, n. 1, p. 71-80, 2014.

REYES, M.; ZÁRATE, A.; CARRILLO, S.; DURÁN, C. Removal of Phosphorus in a Laboratory Scale System of Artificial Wetlands. **Central Chemistry**, v. 2, n. 1, p. 25-32, 2017. https://doi.org/10.29166/quimica.v2i1.546

ROMERO, J. Tratamiento de aguas residuals. *In:* Romero Rojas, J. A. **Tratamiento de aguas residuales; teoria y principios de diseño**. Bogotá: Escuela Colombiana de Ingenieria, 2004. p. 1248.

ROMERO, T.; CASTILLO, Y. Updating of status of Mayabeque city stabilization lagoons. **Hydraulic and Environmental Engineering**, v. 39, n. 2, p. 72-85, 2018.

ROY, M.; SAHA, R. 6 - Dyes and their removal technologies from wastewater: A critical review. *In*: BHATTACHARYYA, S.; MONDAL, N. K.; PLATOS, J.; SNÁŠEL, V.; KRÖMER, P. (eds). **Intelligent Environmental Data Monitoring for Pollution Management.** Elsevier, 2021. p. 127-160.

SÁNCHEZ, I.; MATSUMOTO, T. Bathymetric survey and performance of waste stabilization ponds system. **Journal of Agricultural Science**, v. 30, n. 1, p. 65-78, 2013.

SÁNCHEZ, R.; ROSA, E.; MORENO, M. Analysis of the reliability of operation of optional lagoons primaries in Villa Clara-Cuba. **Chemical Technology**, v. 31, n. 1, p. 23-38, 2011.

SATO, T.; QADIR, M.; YAMAMOTO, S.; ENDO, T.; ZAHOOR, A. Global, regional, and country level need for data on wastewater generation, treatment, and use. **Agricultural Water Management**, v. 130, p. 1-13, 2013. https://doi.org/10.1016/j.agwat.2013.08.007

SILVA, J.; TORRES, P.; MADERA, C. Domestic wastewater reuse in agriculture. A review. **Colombian Agronomy**, v. 26, n. 2, p. 347-359, 2018.

SUÁREZ, J.; PANCHANA, R. **Statistical evaluation of physical, chemical and bacteriological analysis parameters of the effluents of the wastewater treatment system located on the side of Punta Carnero road in Salinas Canton.** 2021. Thesis (Civil Engineering) – Santa Elena Peninsula State University, La Libertad, 2021.

TILLEY, E. *et al.* **Compendium of sanitation systems and technologies**. 2nd ed. Dübendorf: EAWAG, 2011. p. 161.

TREVIÑO, A.; CORTÉS, F. Reduced design method for stabilization lagoon. **Mexican Journal of Agricultural Sciences**, v. 7, n. 4, p. 729-742, 2016.

TULSMA. **Libro VI de la calidad ambiental**. Quito: Environment Ministry of Ecuador, 2015. p. 184.

VARGAS, A.; CALDERÓN, J.; VELÁSQUEZ, D.; CASTRO, M., NÚÑEZ, D. Biological system analysis for domestic wastewater treatment in Colombia. **Ingeniare: Chilean Journal of Engineering**, v. 28, p. 315-322, 2020. https://dx.doi.org/10.4067/S0718-33052020000200315

VELASCO, G.; MONCAYO, J.; CHUQUER, D. Diagnosis of wastewater treatment system of Manta, **Infoanalytic Journal**, v. 7, n. 1, p. 27-39, 2019. https://doi.org/10.26807/ia.v7i1.93

WIJAYA, I.; SOEDJONO, E. Physicochemical characteristic of municipal wastewater in tropical area: Case study of Surabaya City, Indonesia. **IOP Conf. Series: Earth and Environmental Science,** v. 135, p. 1-6, 2018. https://doi.org/10.1088/1755-1315/1/012018

Ambiente & Água - An Interdisciplinary Journal of Applied Science
ISSN 1980-993X – doi:10.4136/1980-993X
www.ambi-agua.net
E-mail: ambi.agua@gmail.com

Assessment of water and sediment quality variation due to organic and conventionally irrigated pre-germinated rice-field cultivation

ARTICLES doi:10.4136/ambi-agua.2825

Received: 23 Dec. 2021; Accepted: 11 Mar. 2022

Enio Sebastião Vicente de Jesus[1]; Letícia Schönhofen Longoni[2]; Luciano Kayser Vargas[2]; Anelise Beneduzi[2*]

[1]Programa de Pós-graduação em Avaliação de Impactos Ambientais. Universidade La Salle (Unilasalle), Avenida Victor Barreto, n° 2288, CEP: 92010-000, Canoas, RS, Brazil. E-mail: eniovicentedejesus@gmail.com
[2]Departamento de Diagnóstico e Pesquisa Agropecuária. Secretaria da Agricultura, Pecuária e Desenvolvimento Rural (SEAPDR), Rua Gonçalves Dias, n° 570, CEP: 90130-060, Porto Alegre, RS, Brazil.
E-mail: leticia-longoni@agricultura.rs.gov.br, lucianokayser@gmail.com
[*]Corresponding author. E-mail: anebeneduzi@gmail.com

ABSTRACT

This study evaluated the water and sediment quality of samples collected from different points in organic (O1, O2, O3, and O4) and conventional (C1, C2, and C3) pre-germinated rice fields located in Viamão/RS. Quality indicators such as phosphorus, dissolved oxygen, manganese, iron, turbidity, and BOD_5 reduced water quality beyond Classes 3 and 4, as defined by CONAMA Resolution 357/2005. Based on the aluminum levels, the water samples collected from all the points were categorized as Class 4; furthermore, the IQA classified the quality of water samples from Points O4 and C1 as “bad.” The COD/BOD ratio was high, demonstrating that the biodegradable fraction was considerably low. The conductivity of water at points O4, C1, C2, and C3, exceeded 100 µS/cm, as defined by CETESB, indicating impacted environments. The levels of zinc in C1 and nickel in O2, C2, and C3 in the sediment exceeded the quality reference values established by the FEPAM Ordinance 85/2014. In general, the lowest water quality was observed in the samples collected from Points O2 and O4, and the lowest sediment quality was observed in the samples collected from all points in the conventional rice fields and from Point O2 in the organic rice field.

Keywords: CETESB, coliforms, CONAMA no 357/2005, FEPAM Ordinance no 85/2014, irrigated rice, physicochemical analysis.

Variação da qualidade da água e do solo de cultivo convencional e orgânico de arroz irrigado pré-germinado

RESUMO

Neste estudo, avaliou-se a qualidade da água e do sedimento de amostras coletadas em diferentes pontos de arrozais orgânicos (O1, O2, O3 e O4) e convencionais (C1, C2 e C3) pré-germinados em Viamão/RS. Indicadores de qualidade como fósforo, oxigênio dissolvido, manganês, ferro, turbidez e DBO_5 reduziram a qualidade da água além das classes 3 e 4, conforme definido pela Resolução CONAMA 357/2005. Com base nos níveis de alumínio, as amostras de água coletadas em todos os pontos foram categorizadas como classe 4; além disso, o IQA classificou a qualidade das amostras de água dos pontos O4 e C1 como "ruim". A relação

DQO/DBO foi alta, demonstrando que a fração biodegradável era consideravelmente baixa. A condutividade da água nos pontos O4, C1, C2 e C3, ultrapassou 100 µS/cm, conforme definido pela CETESB, indicativo de ambientes impactados. Os teores de zinco em C1 e níquel em O2, C2 e C3 no sedimento superaram os valores de referência de qualidade estabelecidos pela Portaria FEPAM 85/2014. De maneira geral, a mais baixa qualidade da água foi observada nas amostras coletadas nos pontos O2 e O4, e a mais baixa qualidade do sedimento foi observada nas amostras coletadas em todos os pontos nos arrozais convencionais e no ponto O2 nos arrozais orgânicos.

Palavras-chave: análises físico-químicas, arroz irrigado, CETESB, coliformes, CONAMA no 357/2005, Portaria FEPAM no 85/2014.

1. INTRODUCTION

The State of Rio Grande do Sul (RS) is the largest rice producer in Brazil, with a harvested area of 964,537 ha and a total production of 7,241,458 tons in 2018/2019 (IRGA, 2019). Although the cultivation of pre-germinated rice is uncommon, it is the cultivation technique adopted in the metropolitan region of Porto Alegre (RS). As a result, the watercourses in this region frequently receive effluents contaminated with agricultural inputs and pesticides, suspended solids, and organic residues carried by the drainage water of rice fields. Consequently, the supply of quality water to the population is impaired, leading to the suspension of the public water supply system.

Water management is essential for crop performance in flood-irrigated rice, as water, in addition to weed control, interferes with nutrient availability, and increases incidences of certain pests and diseases (Gomes *et al.*, 2008). The pre-germinated rice cultivation system uses large volumes of water that aid in the formation of mud during the initial soil preparation, thus providing favorable conditions for seeding. The lowering of the water layer level carried out 3–5 days after sowing releases effluents with a high polluting load into the watercourses. Thus, this cultivation system considerably affects the water quality of the Gravataí River Basin because of the release of effluents with high turbidity and high concentrations of suspended solids. In 2016, the release of drained water from rice fields, which was highly turbid, caused an interruption in the water collection, treatment, and distribution system for public supply, in the municipality of Gravataí.

The sustainable practice of rice farming in Brazil undergoes occasional conceptual changes, mainly to address environmental concerns, especially in terms of the quality of effluent water from farming (Mattos and Martins, 2009). Various efforts have been made to address this issue, such as the intensified search for technological alternatives for organic rice production systems, including but not limited to the discontinued use of pesticides prevalent in traditional farming. When not used according to technical recommendations, these pesticides can contaminate the environment and adversely affect the aquatic and soil organisms within production systems and in their surroundings. Similarly, fertilizers, especially nitrogen and phosphorus, can cause eutrophication of both surface and groundwater, leading to oxygen depletion and severe consequences on aquatic ecosystems (Mattos and Martins, 2009). According to the IRGA (2019), organic rice plantations/cultivation systems in RS occupied 6,000 ha in 2017/2018, with 4,600 ha in settlements of the “Sem Terra” Movement. The average productivity was observed to be 5,000 kg/ha, and the production cost was half that of traditional cultivation (IRGA, 2019).

The pre-germinated rice cultivation system is mainly characterized by the initial drainage of the crop, carried out a few days after sowing, to ensure the proper establishment of the crop. This type of water management, still used by many rice growers, causes detrimental effects on the environment, such as the loss of nutrients (nitrogen, phosphorus, potassium, calcium, and sodium) and total solids; furthermore, it aids the transport of pesticides adsorbed on soil particles into water sources. In addition, they cause an increase in water turbidity, re-infestation of the area by weeds, and imply additional use of water (Scivittaro *et al.*, 2010).

Therefore, in this study, we aimed to evaluate the quality of water and sediment in both conventional and organic cultivation systems of irrigated pre-germinated rice crops located in the municipality of Viamão (RS) through physicochemical and microbiological analyses.

2. MATERIAL AND METHODS

2.1. Collection of water and sediment samples

The water samples were collected on November 30, 2020 (beginning of cultivation period) and March 12, 2020 (end of cultivation period), in organic and conventional pre-germinated rice fields in Viamão (RS). The water samples were stored in sterilized glass bottles (500 mL), and their preservation was inherent to each analysis. Additionally, sediment samples (1 kg) were collected on November 11, 2019, and placed in sanitized plastic bags. Seven points were sampled for collection (Figure 1), four in the organic rice field (O1, O2, O3, and O4) and three in the conventional field (C1, C2, and C3). O1 was in the Águas Claras Dam (-30.070590 and -50.874573), which supplies water for irrigating the organic rice fields. The O2 and O3 were in the drainage channels of the organic rice fields (-30.063640 and -50.886690; -30.04350 and -50.884500). The O4 was also located in a drainage channel (-30.030999 and -50.867704), close to the border of the conventional rice field. Drain water from the cultivated areas of organic rice fields converged at O4. C1 and C2 were located in the drainage channels of the conventional rice fields (-30.021599 and -50.880455; -30.012875 and -50.895273). C3 was located in the Rio Gravataí adduction channel (-30.004383 and -50.923700), and is the main source of water for irrigating the conventional rice fields.

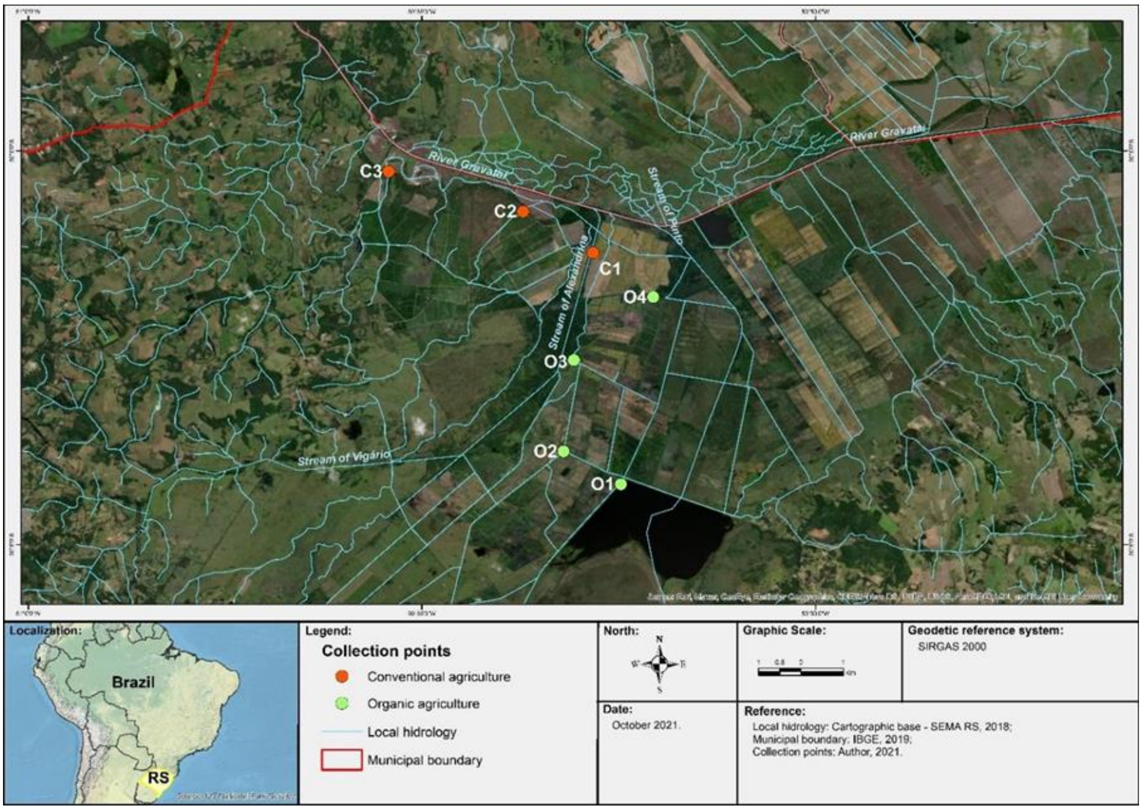


Figure 1. Location of the sampling points in organic and conventional rice cultivation areas in Viamão (RS).

2.2. Physicochemical and microbiological analysis

Electrical conductivity (EC), temperature, turbidity, and pH were measured in situ using HI 9829 (Hanna Instruments) and EXO1 (YSI) multiparameter probes. Physicochemical analyses of the water samples were carried out at the Environmental Analysis Laboratory/FEPAM: total dissolved solids (mg/L), total chloride (mg Cl/L), dissolved oxygen (DO) (mg O_2/ml), chemical oxygen demand (COD) (mg O_2/L), biochemical oxygen demand (BOD_5) (mg O_2/L), total phosphorus (mg P/L), ammonia nitrogen (mg N/L), aluminum (mg Al/L), cadmium (mg Cd/L), lead (mg Pb/L), copper (mg Cu/L), chromium (mg Cr/L), iron (mg Fe/L), manganese (mg Mn/L), nickel (mg Ni/L), and zinc (mg Zn/L) (APHA *et al.*, 2005). The samples obtained were classified according to CONAMA Resolution no 357 (CONAMA, 2005). The identification of the metals present in the sediment was carried out at the Soil Analysis Laboratory/UFRGS, using inductively coupled plasma optical emission spectrometry (ICP-OES, Perkin Elmer), which allowed the categorization of the samples according to CONAMA Resolution no. 420 (CONAMA, 2009) and FEPAM Ordinance no. 85 (FEPAM, 2014). Additional elements analyzed in the sediments were phosphorus (P%), potassium (K%), calcium (Ca%), magnesium (Mg%), sulfur (S%), sodium (mg Na/kg), arsenic (mg As/kg), selenium (mg Se/kg), barium (mg Ba/kg), vanadium (mg V/kg), and cobalt (mg Co/kg) (APHA *et al.*, 2005).

For the quantification of thermotolerant coliforms (TC), the most probable number (MPN) technique was used. The dilutions were inoculated on a chromogenic substrate (Colilert, Idexx) in five sets of five tubes and then incubated for 24 h at 35°C. The substrates that changed their color to yellow and emitted fluorescence were considered positive for thermotolerant coliforms.

For the integrated analysis of physicochemical and microbiological parameters, principal component analysis (PCA) was performed using Past 3.14 software.

2.3. Water Quality Index

The water quality index (WQI) was calculated according to CETESB (2016), based on which the water quality of a source was graded on a scale from very poor to very good (Table 1).

Table 1. WQI classification (CETESB, 2016).

Statement	WQI range
Very good	79 < WQI ≤ 100
Good	51 < WQI ≤ 79
Fair	36 < WQI ≤ 51
Poor	19 < WQI ≤ 36
Very poor	WQI ≤ 19

3. RESULTS

3.1. Microbiological and physicochemical analysis of water at sampled points

Based on the DO values, water from O2 (3.57 mg O_2/L) and C1 (2.57 mg O_2/L) were considered to belong to Class 4 (Table 2), and O1 and C2 were categorized as Class 3 and Class 2, respectively (Table 2). At O3 and O4, the DO values were much lower (0.48 and 0.96 mg O_2/L) than the detection limits established (CONAMA, 2005), making it impossible to classify them above Class 4. Based on turbidity levels, water from O4 and C1 were considered to belong to Class 2; furthermore, as C3 had the highest level of turbidity (151.4 NTU), it could not be categorized above Classes 3 and 4. On the basis of the levels of Fe, O1, O3, O4, C1, and C2 were categorized as Class 3 and O2 (10.9 mg Fe/L) and C3 (5.18 mg Fe/L) above the classes. Mn levels from O4 and C1 indicated that these samples belong to Class 3; furthermore, at O2 with a level of 1.25 mg Mn/L, the class was higher than 4. According to the Zn levels from C1, the water samples were classified as Class 3. Based on BOD_5, water from O2 and O3 were considered to belong to Class 3 (CONAMA, 2005). COD ranged from 47 mg O_2/L (for O3 and C2) to 80 mg O_2/L (for O1), and COD/BOD ratio was high (>3.5, Table 1). According to the total P levels, the samples were categorized above Classes 3 and 4. The pH of C3 was 6.22, which is within the range stipulated by legislation (6.0–9.0, CONAMA, 2005); however, at all the other points, the pH values were between 4.55 and 5.75. The highest EC levels were observed at C2 (117 µS cm^{-1}) for the samples collected on March 12, 2020 (Table 2). Based on total chloride, total dissolved solids, ammonia nitrogen, Cd, Cu, and Ni levels, the water from all points were categorized as Class 1. Water samples collected from O2 on March 12, 2020, had the lowest water quality primarily because their pH values and Fe, Mn, and P levels were above the stipulated class limit, and based on the DO level, the water was classified as Class 4 (Table 2).

The samples collected on November 30, 2020, from O2, O3, C1, and C2 were classified as Class 4 water, based on the DO levels, with values ranging between 2.14 and 3.88 mg O_2/L (Table 3). The DO from O4 (0.52 mg O_2/L) was substantially lower than the values established, and its classification was not possible (CONAMA, 2005). Based on turbidity values, water from C3 was considered to belong to Class 2. However, the turbidity values of water from O2, O3, O4, C1, and C2 (ranging from 263 to 1800 NTU) also exceeded the permissible limits stipulated by legislation (CONAMA, 2005). Furthermore, based on the Fe concentration, the water from O1 and C3 were categorized as Class 3; however, as samples from all other points exhibited high Fe concentration (from 8.38 to 55.7 mg Fe/L), they could not be placed in a class greater than 4. Based on the Mn levels, water from O2, O4, C1, C2, and C3 were considered Class 3. Moreover, based on the Zn concentrations, the samples from O4 were categorized as Class 3. According to the total Cr the samples from O4 were categorized as Class 4 (0.051 mg Cr/L). On the basis of the BOD_5 levels, O4 were classified as Class 2; furthermore, it was observed that the BOD_5 were higher in the water from O3 (14 mg O_2/L) than any of the classes. The COD ranged from 273 mg O_2/L in O4 to 33 mg O_2/L in C3. The COD/BOD ratio was high (Table 3) and was >3.5. High concentrations of Al were found at all samples, ranging from 0.247 in O1 to 281 mg Al/L in O4, making it impossible to categorize them into any of the classes. Total phosphorus showed similar results. The pH of the water from all the points in the conventional rice fields was within the range stipulated by legislation (6.0– 9.0) (CONAMA, 2005); however, the pH of points in the organic rice field varied between 4.94 and 5.81. The EC at O4, C1, C2, and C3, was greater than 100 µS/cm; this result is indicative of impacted environments (CETESB, 2016). For total dissolved solids, the value of 684 mg/L in C1 exceeded the framework limit of CONAMA Resolution no 357 (CONAMA, 2005). On the basis of total chloride, Cd, Pb, Cu, Ni, and ammonia nitrogen levels, the water from all points were classified as Class 1. Based on the parameters presented in Table 3, water samples collected from O3 and O4 on November 30, 2020, had the lowest water quality. Water from O3 was considered low quality primarily based on their pH, BOD_5, and turbidity values. The Al, Fe, and P levels were above the stipulated class limits, and according to the DO levels, the samples were categorized as Class 4. The pH, OD, and turbidity values, and Al, Fe, and P levels of water from O4 were above the stipulated class limit; furthermore, based on the total Cr levels, the samples were categorized as Class 4 (Table 3).

The highest values for TC were observed at O4 (240 MPN/100 mL) and C1 (3300 MPN/100 mL), on November 30, 2020, classifying these water samples as Class 2 and Class 3, respectively (Table 3), which is water suitable for the irrigation of cereals as well as other purposes, according to CONAMA Resolution no 357 (CONAMA, 2005).

Table 2. Classification of water samples collected in organic and conventionally irrigated rice fields based on the physicochemical parameters and proportion of thermotolerant coliforms of the samples as class 1 (green), class 2 (blue), class 3 (orange), class 4 (red) and with quality lower than class 3 or 4 (yellow) (Brasil, 2005). Parameters that are not highlighted were not included in the resolution. Sampling date: March 12, 2020. *NR: not in the resolution.

Parameters	Organic rice fields				Conventionally rice fields			Class 3 CONAMA no 357/2005
	O1	O2	O3	O4	C1	C2	C3	
Electric conductivity (µS/cm)	6	34.0	29.0	26,0	46,0	117,0	88.0	NR
Dissolved oxygen (mg/L)	4.36	3.57	0.48	0.96	2.57	5.22	6.79	4.0
pH	5.00	4.99	4.55	4.57	4.70	5.75	6.22	6.0-9.0
Water temperature (°C)	32.2	29.5	26.0	26.9	24.9	26.3	29.5	NR
Turbidity (NTU)	38.7	31.5	10.5	46.9	70.6	31.5	151.4	100
Cadmium (mg/L)	<0.006	<0.006	<0.006	<0.006	<0.006	<0.006	<0.006	0.01
Copper (mg/L)	<0.004	0.004	<0.004	<0.004	<0.004	<0.004	<0.004	0.013
Chrome (mg/L)	0.907	10.9	1.09	2.43	3.00	2.13	5.18	5.0
Manganese (mg/L)	0.031	1.25	0.059	0.129	0.106	0.056	0.100	0.5
Nickel (mg/L)	<0.011	<0.011	<0.011	<0.011	<0.011	<0.011	<0.011	0.025
Zinc (mg/L)	0.064	<0.020	<0.005	0.008	0.306	0.015	0.014	5.0
Total chloride (mg/L)	8.7	20.6	10.9	10.1	11.4	22.2	16.6	250
BOD_5 (mg/L)	2	9	6	3	3	1	1	10.0
COD (mg/L)	80.0	62.0	47	78	65.0	47	61.0	NR
COD/BOD	40	6.9	7.8	26	21.7	47	61	NR
Total phosphorus (mg/L)	0.135	0.217	0.080	0.138	0.176	0.210	0.298	0.030
Ammoniacal nitrogen (mg/L)	<0.064	<0.064	<0.064	<0.064	0.088	<0.064	<0.064	13.3
Total dissolved solids (mg/L)	70	63	71	95	123	112	191	500
Thermotolerant coliforms (MPN/100mL)	4	13	50	4	6	17	8	4000

Table 3. Classification of water samples collected in organic and conventionally irrigated rice fields based on the physicochemical parameters of the samples as Class 1 (green), Class 2 (blue), Class 3 (orange), Class 4 (red) and with quality lower than Class 3 or 4 (yellow) (CONAMA, 2005). Parameters that are not highlighted were not included in the resolution. Sampling date: November 30, 2020. *NR: not in the resolution.

Parameters	Organic rice fields				Conventionally rice fields			Class 3 CONAMA no 357/2005
	O1	O2	O3	O4	C1	C2	C3	
Electric conductivity (µS/cm)	34.50	58.60	57.70	102,00	154,10	209.30	130.70	NR
Dissolved oxygen (mg O_2/L)	7.90	3.35	2.14	0.52	3.46	3.88	6.58	4.0
pH	4.94	5.77	5.62	5.81	6.42	6.63	6.82	6.0-9.0
Water temperature (°C)	25.7	25.3	25.5	25.4	25.9	25.3	26.6	NR
Turbidity (NTU)	13.80	263	345	1,800	1,600	760	81	100
Aluminum (mg Al/L)	0.247	21.1	27.4	281	196	84.2	4.81	0.2
Cadmium (mg Cd/L)	<0.002	<0.002	<0.002	<0.002	<0.002	<0.002	<0.002	0.01
Lead (mg Pb/L)	<0.004	<0.004	<0.004	<0.004	<0.004	<0.004	<0.004	0.033
Copper (mg Cu/L)	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.004	0.0013
Chrome (mg Cr/L)	<0.004	<0.007	0.008	0.051	0.037	0.017	<0.004	0.05
Iron (mg Fe/L)	0.413	8.38	9.53	55.7	47.3	21.5	4.57	5.0
Manganese (mg Mn/L)	0.027	0.162	0.060	0.447	0.477	0.301	0.151	0.5
Nickel (mg Ni/L)	<0.004	<0.004	<0.004	0.011	0.010	0.004	<0.004	0.025
Zinc (mg Zn/L)	0.003	0.018	0.062	0.467	0.125	0.047	0.018	5.0
Total chloride (mg Cl/L)	9.4	10.8	11.7	20.3	30.2	37.9	14.6	250
BOD_5 (mg O_2/L)	2.0	2	14	4	3	2	1	10
COD (mg O_2/L)	49	54	64	273	206	123	33	NR
COD/BOD	24.5	27	4.6	68.3	68.7	61.5	33	NR
Total phosphorus (mg P/L)	0.146	0.52	0.547	2.70	2.70	1.61	1.10	0.030
Ammoniacal nitrogen (mg N/L)	<0.064	<0.064	<0.064	<0.064	1.67	1.20	0.488	13.3
Total dissolved solids (mg/L)	22	96	67	488	684	173	30	500
Thermotolerant coliforms (MPN/100mL)	1	17	11	240	3300	21	7.8	4000

3.2. Water Quality Index

The water samples collected from O3, O4, and C1 on March 12, 2020, were classified as "fair," and those collected from O1, O2, C2, and C3 were classified as "good” (Table 4). However, the samples collected on November 30, 2020, had relatively lower WQI values. The water from O4 and C1 were classified as "poor,” those from O3 and C2 were classified as "fair," and those from O1, O2, and C3 were classified as "good.”

Table 4. WQI of the water samples collected from different organic and conventionally irrigated rice fields.

Sampling date	Organic rice fields				Conventionally rice fields		
	O1	**O2**	**O3**	**O4**	**C1**	**C2**	**C3**
March 12, 2020	66.85	57.18	40.68	43.81	50.87	72.80	69.99
WQI	Good	Good	Fair	Fair	Fair	Good	Good
November 30, 2020	77.40	51.30	40.50	28.38	34.18	39.46	72.33
WQI	Good	Good	Fair	Poor	Poor	Fair	Good

3.3. Principal component analysis of physicochemical and microbiological parameters of the water samples

During the collection on March 12, 2020 (Figure 2), Component 1 separated points O2, C2, and C3 from points C1, O1, O3, and O4. Furthermore, O2 and O3 were separated from the other points by Component 2, with O3 from organic rice fields being the least impacted. The air temperature was separated from the other attributes by Component 1, and Component 2 separated iron and chloride from the other variables. Turbidity was related to total phosphorus, conductivity, DO, and pH. During the collection on November 30, 2020 (Figure 3), Component 1 separated points O4, C1, and C2; Component 2 grouped points O1, O2, and O3 and separated these three points from C3, indicating that these points from the organic rice field are less impacted. DO was separated by Component 1, and Component 2 separated and grouped nickel, COD, aluminum, chromium, iron, turbidity, and total dissolved solids. Conductivity was more associated with pH and ammonia nitrogen; furthermore, phosphorus and manganese were closely associated.

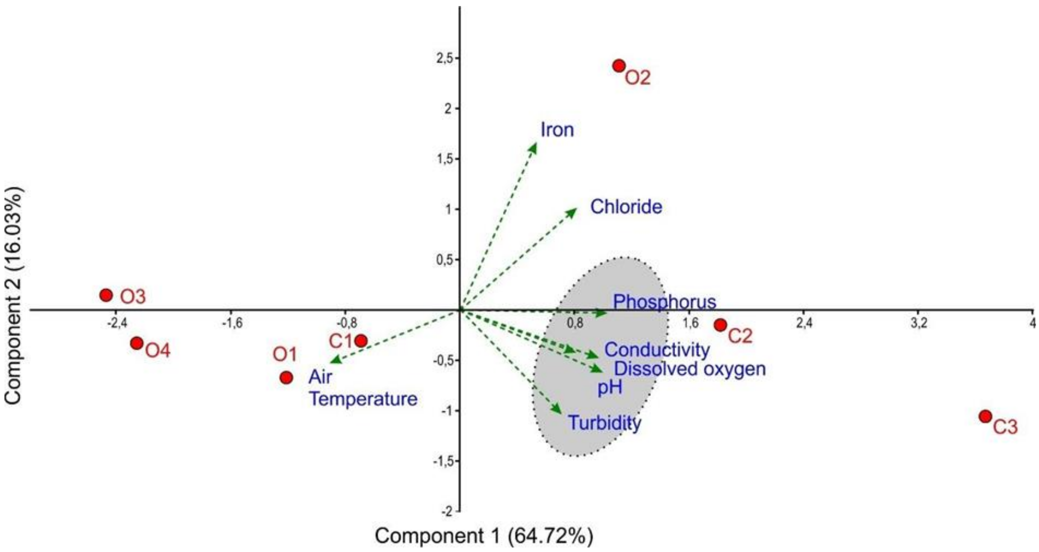


Figure 2. PCA of the physicochemical and microbiological parameters analyzed at different points in the organic and conventionally irrigated rice fields. Component 1, with 64.72% affinity, separates from Component 2 with 16.03% affinity. Sampling date: March 12, 2020.

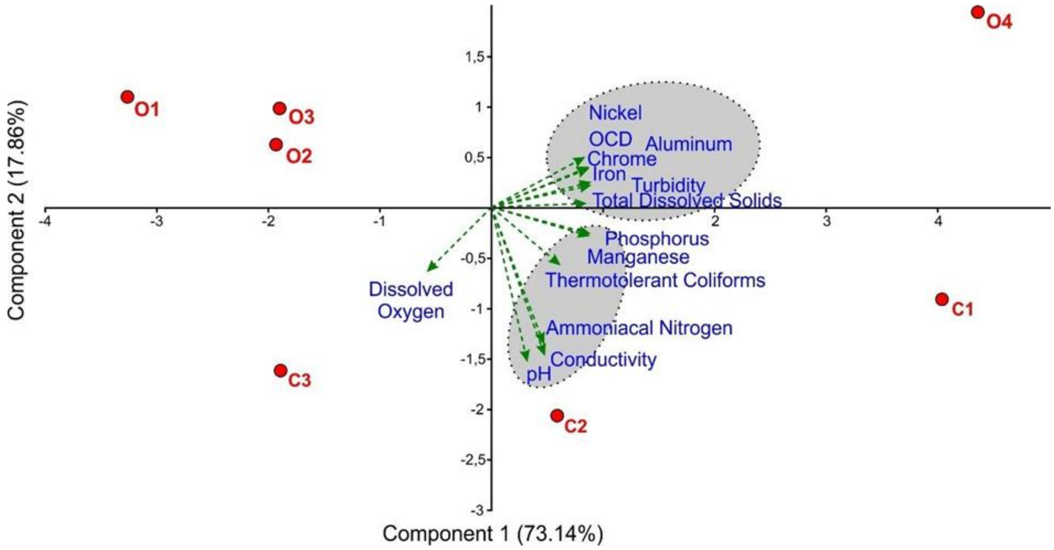


Figure 3. PCA of the physicochemical and microbiological parameters analyzed at different points in the organic and conventionally irrigated rice fields. Component 1, with 73.14% affinity, separates from Component 2 with 17.86% affinity. Sampling date: November 30, 2020.

3.4. Physicochemical analysis of the sediment at the sampled points

According to Table 5, sediment from C1 of the conventional tillage presented the highest levels of P, K, Cu, Zn, Na, Al, Co, and Ba. The samples collected from O2 (organic rice field) had the highest levels for Ca, Mg, S, Fe, Mn, and Va. However, C2 had the highest levels of Ni and P, and O3 had higher Cr levels. Based on the quality reference values (QRV, FEPAM, 2014) the QRV of Zn was 30 mg Zn/kg in C1; this was above the stipulated (29 mg Zn/kg). O2, C2, and C3, had Ni levels of 8, 12, and 8 mg Ni/kg, respectively, all of which were above the established (7 mg Ni/kg). The sediment from conventional tillage, along with that from the O2 from organic tillage, had the lowest quality among the sampled points (Table 5).

Table 5. Concentrations of chemical elements at the sediments sampled on November 11, 2019 in organic and conventionally irrigated rice fields compared with the levels allowed in CONAMA Resolution no 420 (CONAMA, 2009) and FEPAM Ordinance no 85 (FEPAM, 2014).

Parameters	**Organic rice fields**			**Conventionally rice fields**			**CONAMA Resolution no 420/2009**	**FEPAM Ordinance no 85/2014**
	O2	**O3**	**O4**	**C1**	**C2**	**C3**		
Phosphorus (P%)	0.03	0.01	0.01	0.04	0.02	0.02	NR	NR
Potassium (K%)	0.09	0.03	0.03	0.10	0.07	0.08	NR	NR
Calcium (Ca%)	0.24	0.05	0.05	0.21	0.08	0.10	NR	NR
Magnesium (Mg%)	0.18	0.03	0.02	0.09	0.06	0.06	NR	NR
Sulfur (S%)	0.06	0.03	0	0.06	0.03	0.02	NR	NR
Copper (mg/kg)	3	0	1	4	2	2	200	11
Zinc (mg/kg)	22	9	11	30	22	20	450	29
Iron (Fe%)	1.7	0.29	0.28	1.1	0.84	0.73	NR	NR

Continue...

Continued...

Manganese (mg Mn/kg)	442	53	27	179	52	115	NR	NR
Sodium (mg Na/kg)	131	76	93	166	122	145	NR	NR
Cadmium (mg Cd/kg)	<0.2	<0.2	<0.2	<0.2	<0.2	<0.2	3	0.42
Chromium (mg Cr/kg)	11	17	7	11	15	13	150	21
Nickel (mg Ni/kg)	8	7	4	6	12	8	70	7.0
Lead (mg Pb/kg)	7	2	2	8	9	7	180	16
Aluminum (Al%)	1.9	0.59	0.99	3.2	2.2	2.1	NR	NR
Arsenic (mg As/kg)	<0.2	<0.2	<0.2	<0.2	<0.2	<0.2	35	NR
Selenium (mg Se/kg)	<4	<4	<4	<4.	<4	<4	NR	NR
Cobalt (mg Co/kg)	6	2	2	7	6	5	35	7.0
Barium (mg Ba/kg)	75	23	30	95	51	65	300	NR
Vanadium (mg V/kg)	36	7	8	35	30	27	NR	76

*NR: not in the resolution.

As shown in Figure 4, PCA Component 1 separated C1, C3, and O2 from C2, O3, and O4. O2, O3, and O4 were separated from the other points by Component 2. Furthermore, Component 2 separated P, K, N, and organic carbon from pH, Mg, Mn, Fe, and Ca.

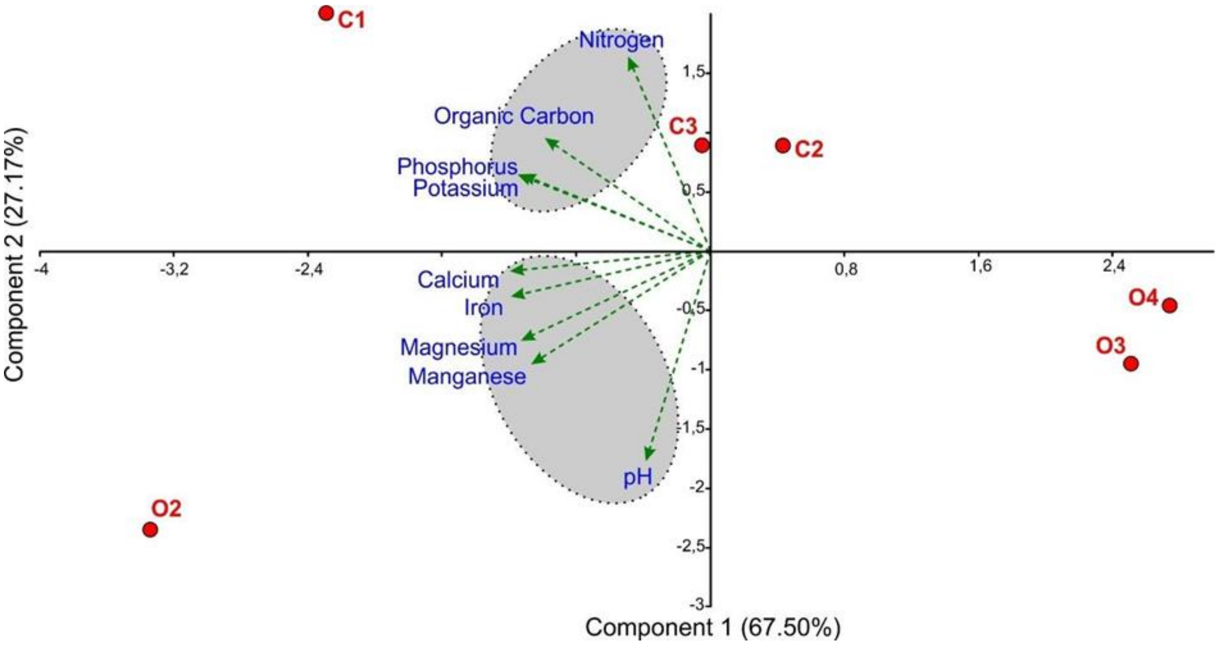


Figure 4. PCA of the physicochemical and microbiological elements analyzed at different points in the organic and conventionally irrigated rice fields. Component 1, with 67.50% affinity, separates from Component 2 with 27.17% affinity. Sampling date: November 11, 2019.

4. DISCUSSION

In this study, the quality of water and sediment was evaluated in two pre-germinated irrigated rice fields (organic and conventional). Both the cultivation areas were located in the Gravataí River Basin (Viamão/RS), in the stretch identified as Middle Gravataí classified as Class 3 (Bourscheid, 2012) based on CONAMA Resolution no 357 (CONAMA, 2005). The water sampled from Points O1 and C3 is used for irrigation and comes from the Águas Claras Dam and the Gravataí River, respectively. The water from Points O3, O4, C1, and C2, is used for drainage and is reused for irrigating organic and conventional rice fields via pumping. According to Mondstock (2015), rice crops can have a great impact on water quality, especially when taking into consideration the runoff from cultivated land. Furthermore, during periods of intense rainfall, when the soil is unable to absorb the entire volume of accumulated water, drainage channels help in restoring flow.

CONAMA Resolution no 357 (CONAMA, 2005) classifies Class 3 fresh waters as those that can be used for irrigation of arboreal, cereal and forage crops as well as other purposes. The results of the physicochemical analysis of the water samples in this study indicated high concentrations of P, turbidity, DO, BOD_5, Mn, and Fe beyond the permissible limits established for Classes 3 and 4 (CONAMA, 2005). At all points sampled, including O1 (Águas Claras dam) and C3 (adduction channel connected to the Gravataí River), P levels exceeded the permitted limits of 0.05 mg P/L indicating high eutrophication. The P from agricultural areas is relevant as it is an indicator of water quality; other indicators such as suspended solids and turbidity are associated with P transport (Scivitaro *et al.*, 2010). In a study conducted by Scivitaro *et al.* (2010) in Capão do Leão (RS), the P levels found in the waters during the drainage period of pre-germinated rice cultivation ranged from 9 to 27 mg P/L, exceeding the maximum value defined for Class 3 (CONAMA, 2005). The previous flooding of the soil in this type of cultivation promotes physical, chemical, and biological changes in relation to the original condition. Among these changes, the increase in the P availability in solution is significant; this increase, in addition to the partial dissolution of the phosphate fertilizer applied during pre-sowing, explains the high P levels in drainage water (Scivitaro *et al.*, 2010). Furthermore, P and turbidity levels (>100 NTU) were high for Class 3 (CONAMA, 2005), with the highest turbidity levels observed at O4 (1800 NTU). This high value coincides with the period of crop cultivation (November 30, 2020), soil preparation, sowing, water depth reduction, and drainage. Drainage causes the loss of total solids, which causes an increase in turbidity in the water, as well as the loss of nutrients and other materials, such as pesticides that are adsorbed on suspended soil particles and can be transported to water sources (Mattos *et al.*, 2012).

DO is also one of the primary parameters affecting the environment, with natural water bodies having high levels of DO, essential for the maintenance of aquatic life (Britto *et al.*, 2016; Silva and Pereira Filho, 2010). According to CONAMA Resolution no 357 (CONAMA, 2005), the limit for DO in Class 4 is >2 mg O_2/L. Water samples collected from O1 (7.90 mg O_2/L, Águas Claras Dam) and C3 (6.79 mg O_2/L, the Gravataí River adduction channel) had the highest DO levels. Based on these values, the water samples were categorized as Class 1, that is, water that can be used: a) for human consumption, after simplified treatment; b) the protection of aquatic communities; c) primary contact recreation, such as swimming, water skiing, and diving; d) the irrigation of vegetables that are consumed raw, fruits that grow close to the ground, and the ones that are eaten raw without removing the skin; and e) the protection of aquatic communities in indigenous lands (CONAMA, 2005). The lowest DO value was observed in the water samples collected from O4, < 2.0 mg O_2/L, which is the minimum value for classification in Class 4, that is, waters utilized for navigation and landscape harmony (CONAMA, 2005). Britto *et al.* (2016) attributed DO values <1.0 mg O_2/L to the presence of agricultural production residues in the water body. In the study conducted in the Itajaí hydrographic basin (SC), to assess the quality of water used in irrigated rice farming, relatively low concentrations of DO were observed in the drainages, with an average of 3.7 mg O_2/L, while in the abstractions, the average was 6.6 mg O_2/L. This variation reflects the negative influence of rice farming on water quality (Silva and Pereira Filho, 2010).

According to CETESB (2016), organic discharges result in the greatest increase in BOD in a water body, which completely depletes oxygen in the water, consequently leading to the disappearance of aquatic life. The greater the BOD, the greater is the degree of water pollution (CETESB, 2016). In this study, high levels of BOD were found at all sampling points in the organic rice fields, with the highest value in O3 (14 mg O_2/L) at the time of crop cultivation, surpassing the limit of Class 3 (<10 mg O_2/L) (CONAMA, 2005). Silva and Pereira Filho (2010) analyzed the quality of irrigation water used in rice farming and showed that the average levels of BOD were 3.2 mg O_2/L and 4.8 mg O_2/L in abstraction and drainage, respectively. In this study, BOD values were higher in practically all samples collected from drainages, thus indicating a greater amount of organic matter in returning waters to the Gravataí River after passing through rice fields. Associating BOD with ammoniacal nitrogen and DO reinforces this idea, as BOD was higher in returning water, indicating increased ammonia concentration and decreased DO.

COD is an extremely useful parameter when used in conjunction with BOD to observe the biodegradability of dumps (Britto *et al.*, 2016). However, no reference for COD exists in the CONAMA Resolution no 357 (CONAMA, 2005). Molozzi *et al.* (2006) reported that during the rice maturation stage, COD values were higher in drainage water than that in irrigation water. It was only after the harvest that irrigation water had a higher COD than that of drainage water. In this study, the COD values ranged from 47 mg O_2/L (in O3 and C2), during the rice maturation and harvest phases, to 273 mg O_2/L (O4) during the soil preparation and rice sowing phases. In general, the COD was higher in the conventional rice field than that in the organic rice field; this value can be attributed to the application of chemical fertilizers and pesticides (Molozzi *et al.*, 2006). In this study, the COD/BOD ratio was high, demonstrating that the biodegradable fraction was very low, which suggests the need for physicochemical treatment of these waters (Von Sperling, 1996).

High levels of the metals Fe, Mn, Zn, and Ni were found in samples collected for the present study. The highest Fe concentrations were found during the cultivation period of the crop and at O4 (55.7 mg Fe/L). Fe concentrations >5 mg Fe L^{-1}, which is the maximum value for this parameter for Class 3 (CONAMA, 2005), were found at some points, mainly at the end of cultivation and harvest. Fe is toxic to vegetables at concentrations >5 mg Fe L^{-1}, when it makes P and Mo unavailable, causing nutritional deficiency (Almeida, 2010). Moreover, the red yellow alic argis soil of this region contains iron, hematite, and goethite with a high propensity for erosion (EMBRAPA, 2018), which may also contribute to these high levels. Santos and Hernandez (2013) also found high Fe concentrations at all sampled points of the agricultural basin of the Ipê Stream in Ilha Solteira (SP). The Mn level in O2 was 1.25 Mn mg/L, which is higher than the maximum value allowed for inclusion in the classes (0.50 mg Mn/L, CONAMA, 2005). According to Quinatto *et al.* (2019), Mn rarely reaches concentrations of 1.0 mg Mn/L in natural surface waters and is usually present in amounts of <0.2 mg Mn/L. The concentration of Zn in the sediment collected from C1 (30 mg Zn/kg) exceeded the QRV (29.0 mg Zn/kg, FEPAM, 2014); this increase can be attributed to the application of fertilizer in the conventional field. In the study by Lavnitcki *et al.* (2020) in the Ponte Grande River Basin (Lages/SC), Zn concentration in the sediments ranged from 57.86 to 210.90 mg Zn/kg between the sampled points, being found in springs, junctions and mainly in urbanized areas due to the release of domestic effluents and rainwater. Sanches Filho *et al.* (2015) conducted a study on the São Lourenço River (RS) and found high Zn concentrations (between 27.9 and 83.6 mg Zn/kg) for most of the sampling points; however, the high concentrations were attributed to the geology of the region. In this study, the highest Ni concentration in the sediment was 12.0 mg Ni/kg at C2, which is a drainage channel in the conventional rice field. O2, O3, C2 and C3 also presented high concentrations of Ni above the QRV of 7.0 mg Ni/kg (FEPAM 2014). Sanches *et al.* (2014) reported that the Ni concentration at sediment sampling Point 3 (20 mg Ni/kg) was high enough to present a possible risk to aquatic life in the São Lourenço Stream.

The water quality index (WQI) was classified as "poor" and indicated a decrease in the water quality of samples collected from O4 (organic farming) and C1 (conventional farming) on November 30, 2020. The samples collected from other points were either "fair" or "good" water. Lopes *et al.* (2008), based on the isolated analysis of the variables that comprise the WQI, reported that the isolated value of this index is not sufficient for an accurate analysis of water quality. According to the authors, the fluctuations of the WQI variables compensate each other, keeping the index relatively stable at a level; however, this relative "stability" masks important fluctuations in the environment and must be monitored and analyzed with greater care. An environment can fall into the "optimal" range of the WQI even if some substances are at concentrations that are toxic to the biota and WQI does not include important potential contaminants, such as pesticides (Cunha *et al.*, 2013).

Although the electrical conductivity (EC) is not determined in CONAMA Resolution no 357 (CONAMA, 2005), CETESB (2016) considers values >100 µS cm^{-1} indicative of impacted environments. The EC of samples collected from C2 on March 12, 2020, and those collected from O4, C1, C2, and C3 on November 30, 2020, exceeded 100 µS/cm, indicative of impacted environments. In a study by Scivitaro *et al.* (2010), the EC indices of the post-sowing drainage water of pre-germinated rice remained >150 µS cm^{-1}, indicating the changes in water composition, especially in the concentration of minerals, owing to the dissolution of salts of the applied fertilizers. In conventional farming, EC values are higher than those in organic rice farming because of the application of fertilizers. EC is one of the most important criteria regarding the quality of irrigation water, as high values imply a risk of salinization of the soil and corrosive characteristics of the water. Although EC is a good indication of changes in the mineral composition of water, it does not indicate the relative amounts of the various components (Scivitaro *et al.*, 2010).

5. CONCLUSION

Based on different parameters, the lowest water quality was observed in Points O2 and O4 (organic rice field). However, for the sediment, the lowest quality was observed in the conventional rice fields along with Point O2 in the organic rice field. Although high levels of metals and other parameters provide essential and beneficial micronutrients to crops, they introduce toxic and potentially carcinogenic heavy metals to the production chain, which through cumulative soil contamination, can reach the food chain and cause various health problems in humans (Nava *et al.*, 2011). The techniques used in organic agriculture, when carried out correctly, improve soil fertility, increase the capacity to retain water and nutrients, reduce erosion and leaching, consequently, the loss of quality of this resource (Kamiyama *et al.*, 2011). This is possible by applying the key concepts of organic agriculture and soil conservation, such as the maintenance of soil cover, crop rotation, increase in organic matter content, and favorable biological activities occurring in the soil. In addition, the soil in an organic production system has a greater capacity to retain potential contaminants, thus reducing the chances of their percolation (Morgera *et al.*, 2012).

6. REFERENCES

ALMEIDA, O. A. de. **Qualidade da água de irrigação.** Cruz das Almas: Embrapa Mandioca e Fruticultura, 2010. 234p.

APHA; AWWA; WEF. **Standard Methods for the Examination of Water and Wastewater**. 21. ed. Washington, 2005.

BOURSCHEID – ENGENHARIA E MEIO AMBIENTE S. A. **Processo de Planejamento da Bacia do Rio Gravataí, Plano de Bacia. Plano de Recursos Hídricos da Bacia Hidrográfica do Rio Gravataí – Relatório Final**. Porto Alegre: Secretaria do Meio Ambiente do Estado do Rio Grande do Sul, 2012. 19p.

BRITTO, F. B.; SILVA, T. M. M; VASCO, A. N.; AGUIAR NETTO, A. O.; CARVALHO, C. M. Impactos da produção do arroz inundado na qualidade da água do Rio Betume, Sergipe. **Revista Agropecuária Técnica**, v. 37, n. 1, p. 44-54, 2016. https://doi.org/10.25066/agrotec.v37i1.25254

CETESB. **Apêndice E- Significado Ambiental e Sanitário das Variáveis de Qualidade das Águas e dos Sedimentos e Metodologias Analíticas e de Amostragem.** 2016. Available in: https://cetesb.sp.gov.br/aguas-interiores/wp-content/uploads/sites/12/2017/11/Ap%C3%AAndice-E-Significado-Ambiental-e-Sanit%C3%A1rio-das-Vari%C3%A1veis-de-Qualidade-2016.pdf. Access: Nov. 2021.

CONAMA (Brasil). Resolução nº 357 de 17 de março de 2005. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. **Diário Oficial [da] União**: seção 1, Brasília, DF, n. 053, p. 58-63, 18 mar. 2005.

CONAMA (Brasil). Resolução nº 420 de 28 de dezembro de 2009. Dispõe sobre critérios e valores orientadores de qualidade do solo quanto à presença de substâncias químicas e estabelece diretrizes para o gerenciamento ambiental de áreas contaminadas por essas substâncias em decorrência de atividades antrópicas. **Diário Oficial [da] União**: seção 1, Brasília, DF, n. 249, p. 81-84, 30 dez. 2009.

CUNHA, R. W.; GARCIA JR., M. D. N.; ALBERTONI, E. F.; PALMA-SILVA, C. Qualidade de água de uma lagoa rasa em meio rural no sul do Brasil. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 17, n. 7, p. 770-779, 2013. https://doi.org/10.1590/S1415-43662013000700012

EMBRAPA. **Sistema Brasileiro de Classificação do Solo**. 5. ed. Brasília: Embrapa, 2018. 356p.

FEPAM. **Portaria FEPAM nº 85/2014**. 2014. Available in: http://www.fepam.rs.gov.br/legislacao/arq/Portaria085-2014.pdf. Access: Nov. 2021.

GOMES, A. S.; SCIVITTARO, W. B.; PETRINI, J. A.; FERREIRA, L. H. G. **A água**: Distribuição, regulamentação e uso na agricultura, com ênfase ao arroz irrigado. Pelotas: Embrapa Clima Temperado, 2008. 46 p.

IRGA. **Boletim de Resultados da Lavoura - Safra 2018/19 – Arroz Irrigado e Soja em Rotação**. 2019. Available in: http://www.irga.rs.gov.br. Access: Nov. 2021.

KAMIYAMA, A.; MARIA, I.C.; SOUZA, D.C.C.; SILVEIRA, A.P.D. Percepção ambiental dos produtores e qualidade do solo em propriedades orgânicas e convencionais. **Bragantia**, v.70, n. 1, p.176-184, 2011. https://doi.org/10.1590/S0006-87052011000100024

LAVNITCKI, L.; BECEGATO, V. A.; VILELA, P. B.; BAUM, C. A.; DUMINELLI, E. C.; TONIAZZO, F. *et al.* Concentração de Metais Pesados na Bacia do Rio Ponte Grande no Município de Lages-SC. *In*: ZUFFO, A. M. (org.). **Engenharia sanitária e ambiental**: tecnologias para a sustentabilidade. Ponta Grossa: Atena Editora, 2020. p. 53-70.

LOPES, F. B.; TEIXEIRA, A. S.; ANDRADE, E. M.; AQUINO, D. N.; ARAÚJO, L. F. P. Mapa da qualidade das águas do rio Acaraú, pelo emprego do IQA e Geoprocessamento. **Revista Ciência Agronômica**, v. 39, n. 3, p. 392-402, 2008.

MATTOS, M. L. T.; MARTINS, J. F. S. **Cultivo de arroz irrigado orgânico no Rio Grande do Sul**. Pelotas: Embrapa Clima Temperado, 2009. 160p.

MATTOS, M. L. T; SCIVITTARO, W. B.; PETRINI, J. A.; DOS SANTOS, I. M. B. **Qualidade da água de drenagem em cultivo de arroz pré-germinado.** Pelotas: Embrapa Clima Temperado, 2012. 6p.

MOLOZZI, J.; PINHEIRO, A.; SILVA, M. R. da. Qualidade da água em diferentes estádios de desenvolvimento do arroz irrigado. **Pesquisa Agropecuária Brasileira**, v.41, n. 9, p.1393-1398, 2006. https://doi.org/10.1590/S0100-204X2006000900007

MONDSTOCK, C. M. **Água de irrigação e drenagem do perímetro de irrigação do Arroio Duro de 2005 a 2012**. Cachoeirinha: IRGA/Divisão de Pesquisa, 2015. 96 p.

MORGERA, E.; CARO, C.B.; DURÁN, G.M. Organic agriculture and the Law. **FAO Legislative Studies**, v. 107, p. 6-10, 2012.

NAVA, I. A.; GONÇALVES JUNIOR, A. C.; NACKE, H.; GUERINI, V. L.; SCWANTES, D. Disponibilidade dos metais pesados tóxicos cádmio, chumbo e cromo no solo e tecido foliar da soja adubada com diferentes fontes de NPK+Zn. **Ciência e Agrotecnologia**, v. 35, n. 5, p. 884-892, 2011. https://doi.org/10.1590/S1413-70542011000500004

QUINATTO, J.; ZAMBELLI, N. L. D. N.; SOUZA, D. H.; RAFAELI NETO, S. L.; CARDOSO, J. T.; SKORONSKI, E. Avaliação da qualidade da água de um rio urbano utilizando o conceito de carga de poluentes: o caso do Rio Carahá (Lages, Brasil). **Revista Ambiente & Água**, v. 14, n. 1, 2019. https://doi.org/10.4136/ambi-agua.2252

SANCHES FILHO, P. J.; ZIEBELL, R.; BETEMPS, G. R. Determinação de metais pesados no sedimento do Arroio São Lourenço-São Lourenço do Sul (RS). **Revista Ibero-Americana de Ciências Ambientais**, v. 6, n. 1, p. 213-221, 2015. https://doi.org/10.6008/SPC2179-6858.2015.001.0017

SANCHES, S. M.; HIPÓLITO, S. G.; PASSOS, C. Vigilância e controle da qualidade da água do Córrego Pirapitinga no município de Ituiutaba–MG. **Revista Agrogeoambiental**, v. 6, n. 3, p. 71-82, 2014. https://doi.org/10.18406/2316-1817v6n32014677

SANTOS, G. O.; HERNANDEZ, F. B. T. Uso do solo e monitoramento dos recursos hídricos no córrego do Ipê, Ilha Solteira, SP. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 17, n.1, p. 60-68, 2013.

SCIVITTARO, W. B.; SILVA, M. T.; BRAUN, J. A. F.; MATTOS, M. L. T. **Monitoramento da Qualidade da Água de Drenagem Inicial de Lavoura de Arroz Pré-germinado**. Pelotas: Embrapa Clima Temperado, 2010. 8 p.

SILVA, D. D. C.; PEREIRA FILHO, J. Qualidade química da água em função de seu uso na rizicultura irrigada na região do baixo estuário do rio Itajaí. **Revista de Estudos Ambientais**, v.12, n. 2, p. 26-37, 2010.

VON SPERLING, M. **Introdução à qualidade das águas e ao tratamento de esgotos.** 2. ed. Belo Horizonte: Editora UFMG, 1996. 243p.