ReGeo ISSN: 2177-3246

AV-PREDICT: A MULTI-MODEL SOFTWARE FOR PREDICTING TURBIDITY REDUCTION IN ALOE VERA-BASED JAR TEST WATER TREATMENT

AV-PREDICT: UM SOFTWARE MULTIMODELO PARA PREDIÇÃO DA REDUÇÃO DE TURBIDEZ EM ENSAIOS DE JAR TEST COM COAGULANTE À BASE DE ALOE VERA

AV-PREDICT: UN SOFTWARE MULTIMODELO PARA LA PREDICCIÓN DE LA REDUCCIÓN DE LA TURBIDEZ EN ENSAYOS DE JARRAS CON COAGULANTE A BASE DE ALOE VERA

https://doi.org/10.56238/revgeov17n7-091

Data de submissão: 01/06/2026 **Data de publicação:** 16/07/2026

Danieli Soares de Oliveira[1], **Clainer Bravin Donadel**[2]

ABSTRACT

Natural coagulants such as Aloe vera have attracted increasing attention as sustainable alternatives for turbidity removal in water treatment. This paper presents AV-Predict, an open-source, multi-model software designed to predict turbidity reduction in jar test experiments using Aloe vera-based coagulation. The software integrates a publicly available experimental dataset with six machine learning algorithms within a user-friendly computational environment, allowing users to estimate treatment performance, compare predictive models, and explore the influence of operational parameters on clarification efficiency. By providing reproducible workflows and rapid prediction of final turbidity and removal efficiency, AV-Predict reduces the need for extensive laboratory experimentation while supporting research, education, and decision-making. The public availability of both the source code and the experimental dataset further promote transparency, reproducibility, and the adoption of data-driven approaches for the optimization of sustainable water treatment technologies.

Keywords: *Aloe vera*. Natural Coagulant. Machine Learning. Predictive Modeling. Jar Test. Water Treatment.

RESUMO

Coagulantes naturais, como a Aloe vera, têm despertado crescente interesse como alternativas sustentáveis para a remoção de turbidez no tratamento de água. Este trabalho apresenta o AV-Predict, um software multimodelo de código aberto desenvolvido para predizer a redução de turbidez em ensaios de Jar Test utilizando coagulação à base de Aloe vera. O software integra um conjunto de dados experimentais de acesso público a seis algoritmos de aprendizado de máquina em um ambiente computacional de fácil utilização, permitindo estimar o desempenho do tratamento, comparar modelos preditivos e analisar a influência dos parâmetros operacionais na eficiência da clarificação. Ao fornecer fluxos de trabalho reprodutíveis e predições rápidas da turbidez final e da eficiência de remoção, o AV-

[1] Ph.D. in Environmental Engineering. Institution: Instituto Federal do Espírito Santo – campus Cariacica (IFES). E-mail: danieli@ifes.edu.br

[2] Ph.D. in Environmental Engineering. Institution: Instituto Federal do Espírito Santo – campus Vitória (IFES). E-mail: cdonadel@ifes.edu.br

Predict reduz a necessidade de extensos experimentos laboratoriais, apoiando atividades de pesquisa, ensino e tomada de decisão. A disponibilização pública do código-fonte e do conjunto de dados experimentais também promove a transparência, a reprodutibilidade e a adoção de abordagens orientadas por dados para a otimização de tecnologias sustentáveis de tratamento de água.

Palavras-chave: *Aloe vera*. Coagulante Natural. Aprendizado de Máquina. Modelagem Preditiva. Jar Test. Tratamento de Água.

RESUMEN
Los coagulantes naturales, como la Aloe vera, han despertado un creciente interés como alternativas sostenibles para la remoción de turbidez en el tratamiento de agua. Este trabajo presenta AV-Predict, un software multimodelo de código abierto desarrollado para predecir la reducción de la turbidez en ensayos de jarras utilizando coagulación a base de Aloe vera. El software integra un conjunto de datos experimentales de acceso público con seis algoritmos de aprendizaje automático en un entorno computacional de fácil uso, lo que permite estimar el desempeño del tratamiento, comparar modelos predictivos y analizar la influencia de los parámetros operativos en la eficiencia de la clarificación. Al proporcionar flujos de trabajo reproducibles y predicciones rápidas de la turbidez final y de la eficiencia de remoción, AV-Predict reduce la necesidad de realizar extensos ensayos de laboratorio y contribuye a la investigación, la enseñanza y la toma de decisiones. La disponibilidad pública tanto del código fuente como del conjunto de datos experimentales también fomenta la transparencia, la reproducibilidad y la adopción de enfoques basados en datos para la optimización de tecnologías sostenibles de tratamiento de agua.

Palabras clave: *Aloe vera*. Coagulante Natural. Aprendizaje Automático. Modelado Predictivo. Ensayo de Jarras. Tratamiento de Agua.

1 INTRODUCTION

Water treatment is essential to ensure safe water for human consumption, agriculture, and industrial applications (Kekes; Tzia; Kolliopoulos, 2023). Among the various treatment stages, coagulation plays a crucial role in destabilizing suspended particles and promoting their aggregation, enabling efficient turbidity removal. While conventional chemical coagulants such as aluminum and iron salts are widely used, increasing environmental and health concerns have driven the search for sustainable and biodegradable alternatives (Diver; Nhapi; Ruziwa, 2023).

Natural coagulants derived from plant-based materials have gained attention due to their low toxicity, biodegradability, and local availability. In particular, *Aloe vera* has demonstrated promising performance in turbidity removal, showing effective coagulation properties and potential applicability in sustainable water treatment systems (Benalia; Derbal; Khalfaoui; Bouchareb *et al.*, 2021; Febrianti; Permana; Ariani, 2024; Katubi; Amari; Harharah; Eldirderi *et al.*, 2021). Experimental studies have demonstrated that turbidity reduction efficiency using *Aloe vera* depends strongly on operational conditions such as coagulant dosage, initial turbidity, and mixing parameters. Recent datasets and experimental investigations have provided valuable information on *Aloe vera*-based coagulation performance under controlled jar test conditions, supporting further computational and predictive analyses (Oliveira; Nascimento; Donadel, 2025; Oliveira; Donadel, 2026).

Despite the availability of experimental data, predicting turbidity reduction efficiency under varying operational conditions remains a challenge. Laboratory experimentation is time-consuming and resource-intensive, and the lack of accessible computational tools limits the ability of researchers and practitioners to explore parameter sensitivity and optimize coagulation performance efficiently. Data-driven predictive modeling offers a promising approach to address this limitation, enabling rapid estimation of treatment performance based on existing experimental datasets.

To address this gap, this work presents AV-Predict, a data-driven, multi-model software designed to predict turbidity reduction efficiency in jar test experiments using *Aloe vera* as a natural coagulant. The software integrates experimental datasets and implements predictive modeling techniques to provide an accessible computational environment for performance estimation and analysis. By facilitating reproducibility, supporting parameter exploration, and enabling efficient prediction of turbidity removal performance, AV-Predict contributes to

advancing the practical application and optimization of natural coagulants in sustainable water treatment systems.

2 SOFTWARE STRUCTURE

The AV-Predict computational tool was developed in Python. The software integrates an experimental dataset and predictive modeling capabilities within an interactive computational environment. Table 1 summarizes the main software metadata, including version information, repository location, licensing, and technical requirements.

Table 1

AV-Predict software metadata and technical specifications

Code metadata description	Value
Current code version	v1.0.0
Permanent link to repository used for this code version	https://github.com/clainer-ifes/AV-Predict
Legal code license	MIT License
Code versioning system used	Git
Software code languages, tools and services used	Python 3.12
Compilation requirements, operating environments and dependencies	AV-Predict requires Python 3.12 and standard scientific libraries. The software is platform-independent and the required dataset is included in the code repository.
Developer documentation/manual	Available in the GitHub repository README file

Source: Prepared by the authors.

2.1 EXPERIMENTAL EVALUATION IN JAR TESTS - DATABASE

The software includes a publicly available experimental dataset derived from jar test evaluations, as described in (Oliveira; Donadel, 2026). The dataset is available in the file Dataset.xlsx and must be loaded during code execution. It contains the following parameters:

a) Coagulant concentration [mL/L];
b) Coagulation/flocculation time [min];
c) Mixing speed [rpm];
d) Global velocity gradient [s^{-1}];
e) Trial number [-];
f) Sedimentation time [min]; and
g) Final turbidity [NTU].

These variables serve as input features and target outputs for predictive model training and validation.

2.2 INPUT DATA FOR THE COMPUTATIONAL CODE

During execution, the user is prompted to define the predictive model, specify model configuration parameters, and provide input conditions for turbidity prediction.

2.2.1 Predictive model selection

The software supports multiple predictive models, including Linear Regression, K-Nearest Neighbors Regressor, Support Vector Regressor (RBF kernel), Random Forest Regressor, Gradient Boosting Regressor, and Artificial Neural Network (Multilayer Perceptron Regressor). Linear Regression is used as the default model unless another option is selected.

2.2.2 Hyperparameter configuration

Hyperparameters define the structure and learning behavior of predictive models and must be specified prior to training. Hyperparameter requirements depend on the selected model. For example, Linear Regression does not require hyperparameter tuning, whereas Artificial Neural Networks require specification of parameters such as the number of hidden layers and neurons per layer.

2.2.3 Input parameters for prediction

The software requires the following operating conditions:

i. Initial turbidity [NTU]: Reference turbidity measured at the beginning of the experiment, prior to sedimentation;
ii. Coagulant concentration [mL/L]: Dosage of the *Aloe vera* coagulant applied in the jar test;
iii. Coagulation/flocculation time [min]: Total duration of the coagulation and flocculation stages;
iv. Mixing speed [rpm]: Rotational speed of the mixing device during agitation; accepts integer values only;
v. Global velocity gradient [s^{-1}]: Measure of shear intensity generated during mixing; accepts integer values only;

vi. Sedimentation time [min]: Duration of the settling phase before turbidity sampling; accepts integer values only.

These parameters define the operating conditions under which turbidity removal performance is predicted.

2.2.4 Data Flow in the computational code

The execution workflow begins with loading the experimental dataset included with the software. The user then selects a predictive model, configures hyperparameters if applicable, and provides input conditions. The software processes the input data and applies the selected model to estimate final turbidity and turbidity removal efficiency. Model training and prediction are performed using the integrated experimental dataset, ensuring reproducible and consistent results. The software generates numerical outputs and graphical visualizations illustrating predicted turbidity removal performance. All computed variables remain accessible within the execution environment, allowing further analysis and interpretation.

3 IMPACTS

AV-Predict provides an accessible and reproducible software solution for predicting turbidity reduction efficiency using *Aloe vera*–based coagulation. By integrating a publicly available experimental dataset with multiple predictive modeling approaches, the software enables rapid estimation of treatment performance under varying operating conditions without the need for extensive laboratory experimentation. The tool supports researchers, engineers, and students by facilitating model comparison, parameter sensitivity analysis, and performance prediction within a user-friendly computational environment. AV-Predict contributes to advancing sustainable water treatment research by enabling efficient evaluation and optimization of natural coagulant applications. The software promotes reproducibility, accelerates experimental interpretation, and supports decision-making in laboratory-scale clarification studies, ultimately facilitating the broader adoption of environmentally sustainable treatment strategies.

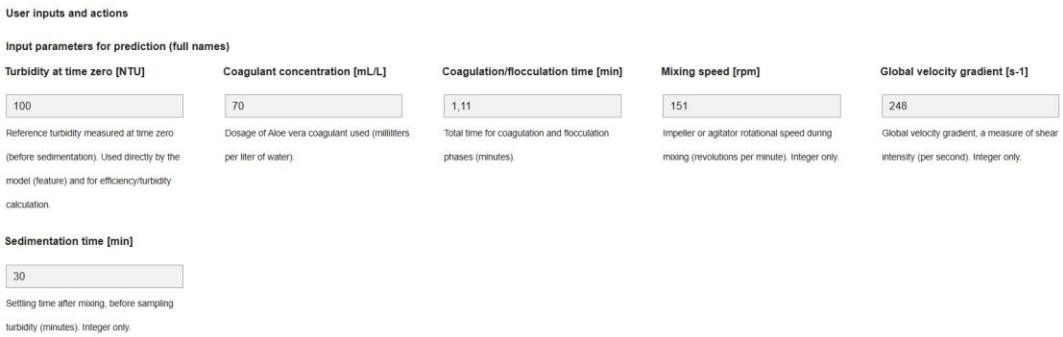
4 RESULTS AND FUTURE WORK

Figure 1 presents the graphical user interface of AV-Predict used to define the input parameters for prediction. The operational parameters provided to the model include turbidity

at time zero, coagulant concentration, coagulation/flocculation time, mixing speed, global velocity gradient, and sedimentation time. The example values shown in the interface are 100 NTU, 70 mL/L, 1.11 min, 151 rpm, 248 s^{-1}, and 30 min, respectively.

Figure 1

User interface of AV-Predict showing the prediction input module

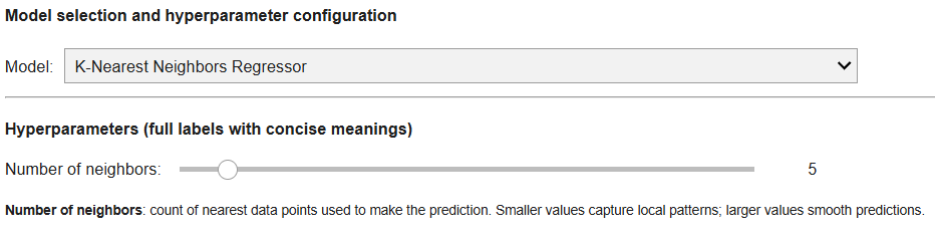


Source: Prepared by the authors.

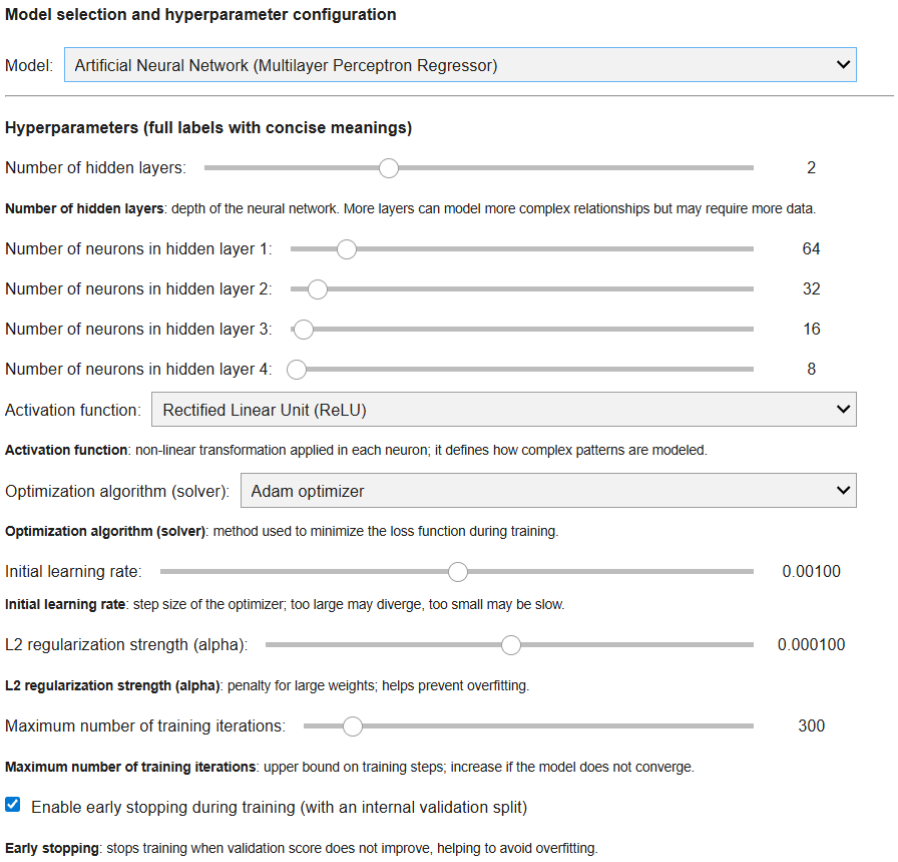
Figure 2 presents two example applications of AV-Predict, illustrating model selection and the configuration of hyperparameters, when applicable. The first example (Figure 2a) illustrates the use of the K-Nearest Neighbors Regressor, while the second example (Figure 2b) demonstrates the configuration of the Artificial Neural Network (Multilayer Perceptron Regressor). These examples highlight the flexibility of the tool and the range of modeling approaches available for estimating turbidity reduction efficiency using *Aloe vera*–based coagulation.

Figure 2

Examples of predictive model selection and hyperparameter configuration in AV-Predict: (a) K-Nearest Neighbors Regressor; (b) Artificial Neural Network (Multilayer Perceptron Regressor)



(a)



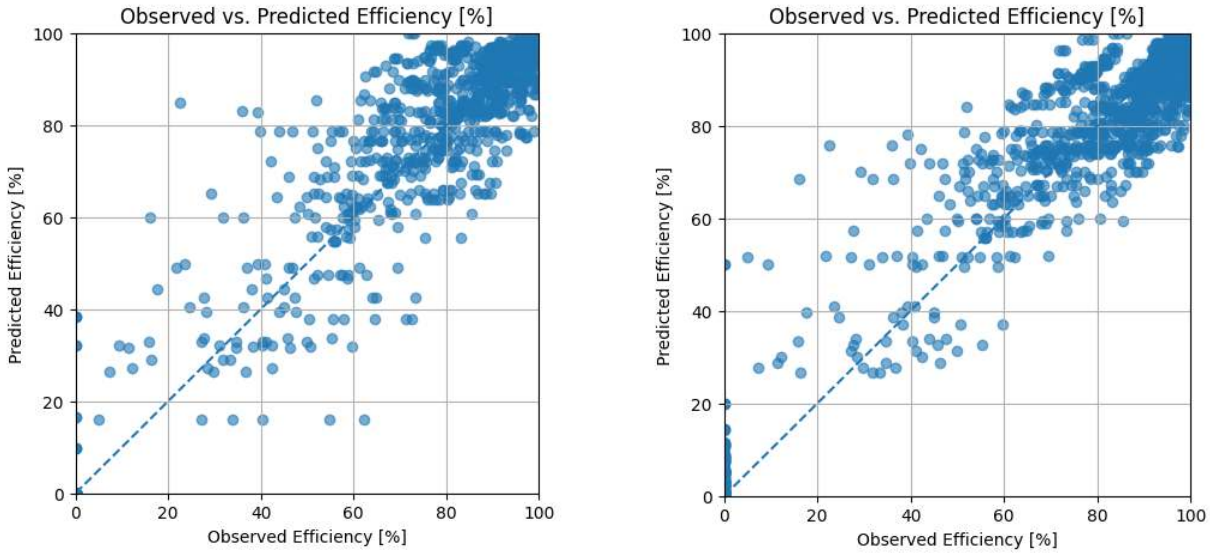
(b)

Source: Prepared by the authors.

Figure 3 presents the training results of the selected predictive models. Figure 3a shows the performance obtained with the K-Nearest Neighbors Regressor, whereas Figure 3b displays the training outcomes of the Artificial Neural Network (Multilayer Perceptron Regressor). Together, these results illustrate differences in model behavior and support the assessment of predictive performance under varying operational conditions.

Figure 3

Examples of training results obtained in AV-Predict: (a) K-Nearest Neighbors Regressor; (b) Artificial Neural Network (Multilayer Perceptron Regressor)



(a) (b)

Source: Prepared by the authors.

Table 2 summarizes the predictive performance of the K-Nearest Neighbors Regressor and the Artificial Neural Network (Multilayer Perceptron Regressor). The reported cross-validation metrics indicate consistent generalization capability for both models, with Group 5-fold CV R^2 values of 0.878 ± 0.168 and 0.867 ± 0.202, respectively. The corresponding RMSE values for final turbidity prediction were 12.184 ± 2.433 NTU for the K-Nearest Neighbors Regressor and 12.145 ± 2.086 NTU for the Artificial Neural Network.

For the simulation scenario, the input values adopted were those presented in the example of Figure 1. Under these operating conditions, the K-Nearest Neighbors Regressor predicted a turbidity removal efficiency of 91.22%, resulting in a final turbidity of 8.779 NTU, whereas the Artificial Neural Network predicted a removal efficiency of 92.65% and a final turbidity of 7.348 NTU. These results demonstrate comparable predictive behavior between the models, with slight differences in estimated removal performance.

Future developments will incorporate additional modeling approaches, expanded datasets, and enhanced visualization features to further improve usability and predictive accuracy.

Table 2

Predictive performance metrics and simulation results for the selected models using the input parameters illustrated in Figure 1

Parameter	K-Nearest Neighbors Regressor	Artificial Neural Network (Multilayer Perceptron Regressor)
Group 5-fold CV R^2 (final turbidity)	0.878 ± 0.168	0.867 ± 0.202
Group 5-fold CV RMSE (final turbidity)	12.184 ± 2.433 NTU	12.145 ± 2.086 NTU
Predicted Turbidity Removal Efficiency [%]	91.22%	92.65%
Predicted final Turbidity [NTU]	8.779	7.348

Source: Prepared by the authors.

5 CONCLUSIONS

This paper presented AV-Predict, an open-source, multi-model software developed to predict turbidity reduction in jar test water treatment using *Aloe vera* as a natural coagulant. The software integrates a publicly available experimental dataset with multiple machine learning algorithms within a flexible and reproducible computational framework, allowing users to compare predictive models and estimate treatment performance under different operating conditions. By reducing the need for extensive laboratory experimentation and supporting rapid parameter exploration, AV-Predict represents a practical tool for research, education, and decision support in sustainable water treatment. Furthermore, the public availability of both the source code and the experimental dataset promotes transparency, reproducibility, and the wider adoption of data-driven approaches for natural coagulant research. Future developments will focus on expanding the experimental database, incorporating additional predictive models and automated optimization strategies, and enhancing data visualization capabilities, thereby broadening the applicability of the software to sustainable water treatment studies.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING

The APC was supported by the National Council for Scientific and Technological Development (CNPq) through a Research Productivity Fellowship. The experimental apparatus was funded by the Foundation for Research Support of Espírito Santo (FAPES), IFES and CNPq.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Clainer Bravin Donadel: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Danieli Soares de Oliveira**: Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

ACKNOWLEDGMENTS

Authors thank Federal Institute of Espírito Santo (IFES), CNPq and FAPES. During the preparation of this manuscript, the authors used AI tools for the purposes of language editing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

REFERENCES

Benalia, A.; Derbal, K.; Khalfaoui, A.; Bouchareb, R. *et al.* Use of Aloe vera as an Organic Coagulant for Improving Drinking Water Quality. **Water**, v.13, n. 15, p. 2024, DOI: 10.3390/w13152024.

Diver, D.; Nhapi, I.; Ruziwa, W. R. The potential and constraints of replacing conventional chemical coagulants with natural plant extracts in water and wastewater treatment. **Environmental Advances**, 13, p. 100421, 2023/10/01/ 2023.

Febrianti, N.; Permana, R. N. W.; Ariani, I. K. Utilization of aloe vera as a biocoagulant for turbidity, total dissolved solid (TDS), and iron (Fe) removal in well water. **IOP Conference Series: Earth and Environmental Science**, 1312, n. 1, p. 012007, 2024/02/01 2024.

Katubi, K. M.; Amari, A.; Harharah, H. N.; Eldirderi, M. M. *et al.* Aloe vera as Promising Material for Water Treatment: A Review. **Processes**, v.9, n. 5, p. 782, DOI: 10.3390/pr9050782.

Kekes, T.; Tzia, C.; Kolliopoulos, G. Drinking and Natural Mineral Water: Treatment and Quality–Safety Assurance. **Water**, v.15, n. 13, p. 2325, DOI: 10.3390/w15132325.

Oliveira, D. S. d.; Donadel, C. B. Experimental dataset on the application of aloe vera gel as a natural coagulant for water treatment. **Data in Brief**, 64, p. 112366, 2026/02/01/ 2026.

Oliveira, D. S.; Nascimento, R. S.; Donadel, C. B. Aloe Vera in Water Treatment: Toward a Greener Future for Environmental Engineering. **Sustainability**, v.17, n. 9, p. 4163, DOI: 10.3390/su17094163.