SINTEZA 2025
INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION TECHNOLOGY, COMPUTER SCIENCE, AND DATA SCIENCE

STUDENT SESSION

THE ROLE OF ARTIFICIAL INTELLIGENCE IN WASTEWATER TREATMENT

Milica Radojković*,
[0009-0001-8185-7211]
Milena Rikalović
[0000-0002-1809-5461]

Singidunum University,
Belgrade, Serbia

Abstract:

This paper analyses the role of artificial intelligence (AI) in wastewater treatment. It highlights how artificial neural networks (ANNs) can effectively process complex data, such as the correlation between biological and chemical parameters, which are often not linearly related. These models excel in handling large datasets, making them more effective than traditional methods for monitoring wastewater. Although AI is portrayed in media, it also serves as a valuable tool in (bio)technological processes, especially in complex systems that manage large datasets. The use of AI, particularly machine learning (ML), is growing in wastewater treatment plants (WWTP)s to optimise their operations. So far, AI implementation in this filed has led to more effective data analysis results than those achieved through conventional statistical approaches.

Keywords:

Artificial Intelligence, Wastewater Treatment, Data Science, Machine Learning.

INTRODUCTION

In modern times, the importance of water as a natural resource and polluted water treatment is one of the most important topics in contemporary science and (bio)technology, considering environmental protection. The special emphasis is on industrial wastewater and the challenges faced by treatment plants due to extensive urbanisation and industrialization since the middle of the last century. This aligns with sustainable development goals and rational resource utilization, which is highlighted by increased water stress on a global level. [1]

Effective wastewater treatment is essential not only for maintaining a safe water supply but also for preserving ecosystems and safeguarding human health. Nevertheless, conventional treatment methods often prove inadequate in tackling the complex challenges presented by modern–day pollutants. [2]

Correspondence:
Milica Radojković

e-mail:
milica.radojkovic.24@singimail.rs

DOI: 10.15308/Sinteza-2025-561-566

As a result, the reclamation and reuse of treated wastewater have become increasingly important strategies. This process is widely acknowledged as one of the most efficient solutions to the issue of water scarcity, playing a crucial role in the sustainable management of water resources by enabling reuse across various sectors. [3]

Therefore, it is imperative to improve and further develop wastewater treatment technologies by adopting modern methods. This need is driven by the escalating concern for water quality, which has become a significant global issue. These challenges are further intensified by current technological limitations in delivering comprehensive and reliable water quality. [3]

One of the modern approaches to this process is artificial intelligence (AI), which enables the analysis of large amounts of data, optimization of parameters, and prediction of results in real time. The concept of AI has been evolving since the 1950s, and in recent decades it has become increasingly popular thanks to advances in computer power and algorithms. [3]

AI is used in various areas, including planning, data mining, decision making, language processing, and even humanoid robots. Its application in wastewater treatment opens new possibilities for more efficient and sustainable management of this resource. [3]

The present study analyses and summarizes the present application of AI for wastewater treatment, as an important tool for this field in applied science. [3]

2. COMPONENTS OF ARTIFICIAL INTELLIGENCE

2.1. DATA

Data forms the core of artificial intelligence, which is capable of handling both quantitative and qualitative information. While AI can process a wide range of data types, such as text, images, and audio, the extent to which it can be applied to numerical or categorical data in wastewater treatment plants (WWTPs) is still under exploration. [4]

Quantitative data relate to precise numerical values and can be either continuous or discrete. In water treatment plants, most of the data is continuous, including flow rates, oxygen concentration, and energy use. [4]

Qualitative information includes details such as identification labels, the presence of faults, and the system's operational status (normal or abnormal). Typically, numerical data are presented in tabular format, where rows represent individual observations, and columns are referred to as attributes or variables that describe specific data characteristics. [4]

It can be stated that artificial intelligence (AI) operates as a model driven by information that is dependent on the quality of the information it processes. The greater the quantity of data available, the greater the accuracy of the AI design becomes – hence, extensive datasets are crucial for efficient training. When insufficient data is used, challenges such as overfitting the training data and poor performance on test data may arise. Issues such as the trained model's greater sensitivity to noise or high dimensionality often arise alongside these problems. [4]

In recent research, special emphasis has been placed on the data pre-processing and conditioning as a prerequisite for developing robust AI applications. This includes cleaning raw sensor inputs, normalising scales, and selecting relevant features that are crucial in dynamic systems like WWTPs. High-quality, structured, and representative data allow AI algorithms to effectively learn the underlying process behaviour and adjust predictions in real time, thereby improving operational control and resource optimisation. As such, the data are not only the fuel but also the framework for deploying successful AI models in complex treatment environments. [5]

2.2. ALGORITHMS

AI algorithms are commonly classified in various ways, including: supervised, semi-supervised, unsupervised, and reinforcement learning algorithm (depending on the learning process); linear and non-linear algorithms (based on the nature of the function they model); and approaches such as machine learning (ML) and deep learning (DL). [4]

In the context of WWTP algorithms, there are hundreds of AI models available, and the most suitable one for a specific issue is not always known beforehand. [6]

The most effective method for selecting an AI algorithm is through random sampling and the iterative trial-and-error method (TaE), which involves comparing multiple algorithms in parallel. [4]

Additionally, the selection of the algorithm is influenced by the nature of the data, the type of issue being tackled, and the intended results. [4]

Generally speaking, AI algorithms are diverse, with a wide range of models to choose from. It is not feasible to identify the optimal algorithm for a particular issue beforehand. [4] Moreover, the choice of algorithm depends on the nature of the data, the problem being addressed, and the desired outputs. [6]

Additionally, considerations such as computational efficiency, scalability, and interpretability of the model also play an important role in algorithm selection. For example, while complex models like deep neural networks may offer high accuracy, they often require greater computational resources and may be more difficult to interpret. Therefore, balancing model performance with practical implementation constraints is essential for achieving reliable and sustainable outcomes in real-world applications. [7]

Recent studies underscore that although deep learning architectures-particularly convolutional and recurrent neural networks-demonstrate considerable strength in pattern recognition and time-series forecasting, their black-box nature presents notable limitations in terms of explain ability, which can be crucial in safety-critical settings such as WWTP (Tomar et al., 2019). [2]In comparison, more straightforward models like decision trees or support vector machines may yield slightly lower accuracy, yet they offer enhanced transparency and interpretability (Bukhari et al., 2021). [8]Therefore, the balance between model performance and interpretability should be assessed on a case-by-case basis, particularly when AI systems are employed to monitor chemical and biological parameters within WWTPs. [8]

3. EXAMPLES OF ARTIFICIAL INTELLIGENCE IN WASTEWATER TREATMENT

Artificial intelligence surged in popularity during the 1980s, and its applications have since expanded across various industries. However, utilizing AI to address urgent challenges in wastewater treatment remains complex, due to the need for specialized skills (e.g, computing systems), knowledge (e.g., data science and statistical methods), and resources (e.g., sufficient datasets). [4] Moreover, selecting an inappropriate problem can lead to the failure of AI implementation. Previous studies have utilised a variety of approaches, such as predicting one effluent quality indicator from another, incorporating data from different stages of the process, or combining influent quality metrics with operational factors, such as return sludge flow rate, sludge volume index, food-to-microorganism ratio, sludge retention time, and energy or chemical consumption. [4]

For example, Zhao et al. (2016) [6]created an artificial neural network (ANN) model aimed at forecasting the effluent levels of total phosphorus (TP), biological oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), and ammonium nitrogen at a wastewater treatment facility located in China. The model was built using input data that included raw wastewater quality indicators-such as influent levels of total phosphorus (TP), biological oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), ammonium nitrogen (NH_4 –N), and pH together with data on energy and chemical consumption such as electricity, coagulants, and flocculants. [9]

Russell & Norving (2016) [6] employed artificial intelligence, specifically a soft sensor built on neural network principles, designed to estimate effluent parameters (including COD, TN, TSS, O_2, NO_3, NH_4, and alkalinity) and represent the non-linear dynamics of the wastewater treatment process. [4] This tool was also utilised for the optimal control of aeration, pumping, and disposal costs, while ensuring compliance with effluent regulations. [4] The strategy employed to tackle the challenge involved the development of an AI-based neural soft sensor, a neural identification model, and an all-encompassing control strategy forWWTPs. [4] The input variables (i.e., secondary variables) included wastewater concentrations measured through a physical online sensor system. The most suitable input variables for the AI algorithm were determined using the Principal Component Analysis (PCA) technique. [4] The computational detector featured a neural network structure consisting of two hidden layers with neuron counts of 100, 55, 25, and 3. The dataset consisted of 4,200 data points (sampling time Ts=15 minutes). The authors employed the neural identification model developed to assess the dynamic behaviour of WWTPs with respect to COD, TN, and TSS. [4] A DNN (Deep Neural Network) model was used to map the temporal evolution and management parameters (influent flow rate and NH4–N). One more, Principal component anlysis (PCA) for variable selection. The developed framework consisted of a standalone neural network employing a feedforward configuration of 50 – 35 – 15 – 1 neuron was trained using 1,500 data entries recorded at 15-minute intervals (Ts=15min). The outcomes demonstrated that the neurogenetic controller achieved and R^2 value in the range of 0.88 to 0.91, alongside a mean absolute percentage deviation between 2.99% and 4.52%. Compared to the conventional decoupled PID control, the neurogenetic method reduced average costs by 12.89%. [4]

The effectiveness of the treated effluent is influenced by both the characteristics of the incoming wastewater and the operational conditions within the treatment plant. [4] Across various studies, input variables used in the models differed significantly, as multiple factors influence WWTP performance. [6]

A significant portion of research in this domain (around 52%) has drawn upon indicators of influent volume and composition as input variables. This trend highlights a common preference among researchers for applying artificial neural networks (ANNs) to capture the intricate and non-linear interactions between influent characteristics and effluent quality. [4] For instance, in a study by Bekkari & Zeddouri, [10] influent parameters such as pH, temperature, TSS, TKN (Total Kjeldahl Nitrogen, a measure of the amount of nitrogen in an environmental sample), BOD, and COD where used to predict the COD level in treated water from a WWTP in Algeria. [10]

Wastewater treatment monitoring data often involves complex and non-linear chemical relationships. Due to this complexity, ANNs, which are inherently suited for non-linear modelling, can accurately predict pollutant removal in WWTPs. [7]

ANNs are widely utilised in water-related research due to their capacity to learn complex non-linear, multi - input\output relationships from historical data during the training process. They are particularly valuable when there is typically based on physical and chemical principles. [7]

Mechanistic models often depend on material and energy balances and empirical formulae, however, these can be inaccurate and require numerous assumptions to remain manageable. As a result, such models carry significant limitations. [7]

4. ADVANCEMENTS AND THE INCORPORATION OF AI INTO WASTEWATER TREATMENT PRACTICES

Biological WWTPs are highly complex and adaptive systems, shaped by the interplay of microbial activity in the incoming wastewater and the multitude of processes involved. As a result, accurately modelling these systems and evaluating their efficiency presents significant difficulties, which has led to a surge in research interest. Recently, AI methodologies, along with associated tools, have become increasingly valuable for forecasting effluent quality or assessing the efficiency of contaminant removal in both *large-scale* and experimental wastewater treatment systems. [4]

Regulatory requirements and discharge limits for effluent from WWTPs are becoming increasingly stringent. The challenging physical environment and demanding conditions within WWTPs make the transparent model-based approach (a method relying on detailed process understanding) impractical for sensor development. [4]

Artificial intelligence (AI) is increasingly being employed to support the creation of virtual measurement tools, enhance automated control mechanisms, and improve optimisation processes in wastewater processing systems. For example, BOD5 (Biological Oxygen Demand over five days) represents a key indicator in evaluating the condition of water for specific uses, yet the standard test is both time consuming and costly. Al–based soft sensors are being developed to address these challenges. [4] For instance, Osman and Li (2020) [11] developed a BOD5 sensor using six input variables, achieving a high R^2 value of 0.99. Furthermore, soft sensors have been created to assist in the control of WWTPs determined in relation to environmental weather patterns, given that such factors can alter both the rate of incoming flow and the concentration of specific constituents. [4]

To improve predictive performance, the study employed genetic algorithms to optimise the quantity of nodes within every concealed layer of the neural network framework. [11] This evolutionary approach enabled the identification of the most suitable network architecture for BOD_5 prediction, thereby enhancing the model's accuracy and generalisation capability. [11]

Feed-forward neural networks operate through a three-layer structure: an input layer containing a vector of pre-processed parameters, an intermediary layer incorporating non-linear transformations, followed by the concluding layer that compares predicted values to actual results. These networks are often trained using the Levenberg-Marquardt algorithm, which enhances convergence and accuracy. Moreover, integrating lagged input data through auto-regressive models helps improve predictive performance in dynamic conditions. [12]

Despite their effectiveness, ANN-based models can be computationally intensive, especially when processing large datasets, making them less feasible for smaller treatment facilities with limited hardware resources. Support vector machines, on the other hand, provide a more efficient alternative, delivering accurate predictions while maintaining lower computational demands. [12]

Adaptive systems based on neural networks and fuzzy logic offer a combined approach, merging the learning ability of neural networks with the flexibility of fuzzy logic to manage uncertainty and variability in biological treatment processes. These models are particularly useful when dealing with sensitive biomass systems, such as aerobic granular sludge reactors, where influent conditions may fluctuate frequently. [12]

Recent studies have also explored hybrid frameworks that combine ANFIS with support vector regression (SVR) in a two-stage modelling process. This approach allows individual prediction of key output parameters and provides greater flexibility i adjusting for errors, leading to improved overall model robustness and accuracy. [12]

5. CONCLUSIONS AND PERSPECTIVES

Based on the current application of AI in wastewater treatment plants, several key aspects can be highlighted in context prediction in AI applications, AI algorithms, soft and online sensors, algorithm comparison, and AI for image analysis. [10]

Prediction in AI applications: Almost all AI applications in WWTPs focus on prediction. These predictions include outputs such as effluent quality parameters (e.g., COD, BOD_5, NH_4–N), energy consumption, influent data, aeration time, sludge bulking, and sludge settleability. [10]

A variety of artificial intelligence approaches, including foundational, combined, and advanced models, along with both supervised and unsupervised techniques, have been evaluated using real-world data. While artificial neural networks (ANN) are the most prevalent algorithm, deep learning (DL) techniques such as DNNs, LSTMs, CNNs, and hybrid approaches are steadily gaining more attention. [4]

Soft sensors are becoming essential in supporting the operation of WWTPs. Their integration into intelligent control systems adds more value than merely using model–based predictive sensors. [10]

Comparisons between algorithms applied to different datasets or studies remain relative. Although R^2 values and error metrics offer preliminary insight, synchronised data with consistent output units is necessary for producing increased trustworthiness of the findings. [10]

Real-time physical monitoring devices, which produce substantial amounts of information, are likewise utilised in developing AI frameworks based on data patterns. [4] These sensors are particularly useful for supporting WWTP operations, including the detection of faults or anomalies. [4]

Leveraging artificial intelligence for interpreting visual inputs, particularly in scenarios involving identifying aggregated microbial clusters and surface-adhered biological layers or predicting effluent quality, still requires further investigation. [4]

Recommendations for future development of AI in water waste treatment include: data solutions should be developed with a thorough understanding of the processes; *large-scale* data sharing and collaborative research efforts should be encouraged; feature selection techniques it is advisable to integrate dimensionality reduction and feature selection methods-such as principal component analysis, multiple regression techniques, and sequential elimination strategies should be incorporated into AI deployment approaches for WWTPs.; [10] and the integration of online sensors and the data–driven models capable of real–time learning and adaptation represents a promising future direction. [10]

Further exploration into AI–based image analysis for water quality assessment, biofilm detection, and sludge characterisation could significantly enhance the operational efficiency and maintenance of wastewater treatment systems. [3]

The rapid progress in machine learning methodologies opens up diverse possibilities for their use in the wastewater treatment industry. This review specifically explored the role of artificial neural networks in forecasting the effectiveness of WWTPs, focusing on parameters like effluent quality and pollutant removal rates. [9]

A systematic review methodology enabled a focused selection and analysis of studies on artificial neural networks in wastewater treatment, leading to a deeper understanding of model structures and parameter tuning. However, the review did not assess real-world applications of these models. One of the main challenges in applying AI to WWTPs remains the lack of reliable and high-quality data, which is crucial for model accuracy and effectiveness. However, it is crucial to underline that one of the key obstacles to applying these models is the challenge of obtaining reliable and accurate data. [9]

REFERENCES

[1] A. Capodaglio, "“The role of wastewater treatment in achieving sustainable development goals," *Environmental Science & Ecotechnology*, vol. 6, pp. 100-107, 2021.

[2] A. Burkov, The Hundred-Page Machine Learning Book, 2019.

[3] D. Z., Wastewater treatment plant monitoring via a deep learning approach, 2018.

[4] T. T. H. N. Q. B. T. X.-T. B. H. H. N. i. D. D. N. X. C. Nguyen, "Artificial intelligence for wastewater treatment," in *Current Developments in Biotechnology and Bioengineering: Advances in Biological Wastewater Treatment Systems,* 2022, 2022, p. 587–608.

[5] P. C. S. G. S. K. V. R. Tomar, AI Enabling Tehnologies: A Survey, 2019.

[6] S.N.P.Rusell, Artificial Intelligence: A modern approach, 2016.

[7] S. Z. Z, Modeling and Optimalisation of Paper-making Wastewater Treatment Based on Reinforcemen Learning, 2017.

[8] R. S. Bukhari, "Machine learning: Real-world uses and future challenges," *Tech Horizons Journal*, vol. 2, pp. 1-12, 2021.

[9] C. C. S. C. O. Marina Salim Dantas, "Artificial neural networks for performance prediction of full-scale wastewater treatment plants: a systematic review," *Water Science & Technology*, vol. 88, no. 6, p. 1447–1470, 2023.

[10] G. W. K., Deep Learning-based Management For Wastewater Treatment Plants Under Blockmain Environmen, 2020.

[11] M. S. Zaghloul, "Application of machine learning techniques to model a full-scale wastewater treatment plant with biological nutrient removal," *Journal of Environmental Chemical Engineering*, vol. 10, no. 3, 2022.

[12] A. Voravich.Ganthavee, "Artificial intelligence and machine learning for the optimization of pharmaceutical wastewater treatment systems," *Environmental Chemistry Letters*, vol. 22, pp. 2293-2318, 2024.