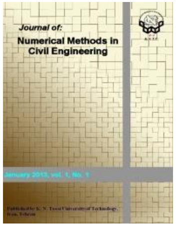
Numerical Methods in Civil Engineering, 10-2 (2025) 86-101



Numerical Methods in Civil Engineering

Journal Homepage: https://nmce.kntu.ac.ir/

The use of machine learning techniques in water and wastewater treatment processes: opportunities and challenges

Nader Biglarijoo*, Amin Shams, Vahid Hadipour*****

ARTICLE INFO

RESEARCH PAPER

Article history:
Received:
November 2025
Revised:
January 2026
Accepted:
February 2026

Keywords:
IoT-based monitoring
Machine learning
Soft sensors
Water treatment
Wastewater treatment

Abstract:

Artificial intelligence (AI) and machine learning (ML) are increasingly transforming water and wastewater treatment by enabling robust prediction, control, and optimization of complex physicochemical and biological processes that conventional mechanistic models struggle to capture. This review provides a comprehensive and methodologically structured synthesis of AI/ML applications across major treatment operations, including coagulation–flocculation, membrane filtration, adsorption, disinfection byproduct (DBP) formation, and wastewater treatment systems such as water quality monitoring, test design and laboratory scale tests. Using a systematic PRISMA-guided literature selection and scientometric analysis of publications from 1998 to 2025, fifty-five studies were critically evaluated to assess algorithmic trends, performance characteristics, and domain-specific applicability. Traditional ML algorithms (KNN, SVM, RF), deep learning architectures (CNN, RNN, LSTM, NARX), and metaheuristic optimization tools (GA, PSO, GEP) were examined alongside hybrid and ensemble models used to address nonlinear, multivariate water-quality relationships. Results show that AI/ML models consistently outperform empirical and mechanistic baselines in predicting coagulant dosage, membrane fouling, permeate flux, pollutant adsorption efficiency, and DBP formation, while emerging approaches such as soft sensors, and IoT-enabled monitoring are enabling real-time operational decision support. The review highlights future research opportunities in digital twins, physics-informed ML, transfer learning, explainable AI (XAI), graph neural networks (GNNs), and autonomous process control using reinforcement learning. By consolidating algorithmic mechanisms, and application domains, this work provides a rigorous and forward-looking perspective on the role of AI/ML in developing resilient and energy-efficient treatment systems.

1. Introduction

Ensuring access to clean and safe water remains one of the most pressing global challenges as industrialization, population growth, and climate change continue to intensify pressures on water resources [1]. Water and wastewater treatment plants (WTPs and WWTPs) are increasingly required to meet stricter regulatory standards while operating under growing uncertainties related to influent variability, emerging contaminants, and resource limitations [2,3].

Conventional treatment systems rely heavily on empirical rules, operator experience, and mechanistic models that often fall short when confronted with nonlinear relationships, complex physicochemical interactions, and rapidly changing operational conditions. As a result, treatment optimization and real-time decision-making remain major challenges for many facilities [4-6]. In recent years, artificial intelligence (AI) and machine learning (ML) have emerged as transformative technologies capable of addressing these limitations. By learning directly from historical and real-time process data, AI/ML models can uncover intricate patterns, model nonlinear responses, and provide fast, accurate predictions that support both process understanding and operational control [5-7]. The increasing availability of high-frequency sensor data, advances in

---

* *Corresponding author: Faculty of Civil Engineering, Semnan University, Semnan 35131-19111, Iran. Email: nader.biglary@semnan.ac.ir.*
** *Faculty of Civil Engineering, Semnan University, Semnan 35131-19111, Iran.*
*** *Faculty of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.*

https://doi.org/10.61882/NMCE.2511.1106

computing power, and the integration of AI with Internet of Things (IoT) technologies have further accelerated their adoption in the water sector [8,9].

A wide range of AI/ML techniques has now been applied across water and wastewater treatment processes. These include traditional algorithms such as k-nearest neighbors (KNN), support vector machines (SVM), random forests (RF), and principal component analysis (PCA); advanced neural network architectures such as multilayer perceptron (MLP), recurrent neural networks (RNN), long short-term memory (LSTM) networks, convolutional neural networks (CNN), and nonlinear autoregressive models with exogenous inputs (NARX); and optimization algorithms such as genetic algorithms (GA), particle swarm optimization (PSO), and gene expression programming (GEP). Meanwhile, hybrid and ensemble models that integrate the strengths of two or more algorithms have also shown significant potential, particularly in complex processes requiring both nonlinear modeling and parameter optimization [10-12].

These methodological developments have led to notable advancements across nearly all treatment processes. In coagulation–flocculation, AI models have demonstrated high accuracy in predicting coagulant dosage under variable influent conditions [13]. In membrane filtration, AI techniques have been used to predict flux decline, optimize backwash cycles, identify fouling mechanisms, and enhance contaminant removal. Adsorption studies have leveraged ANN, GRNN, and hybrid models to predict metal and organic pollutant removal efficiencies with extremely high correlations [14,15]. AI applications in disinfection byproduct (DBP) formation modeling have improved prediction accuracy for THMs, HAAs, bromate, and other DBPs by incorporating nonlinear relationships and dimensionality reduction techniques such as PCA. In wastewater treatment, deep learning, soft sensors, and IoT-integrated AI platforms have enabled real-time water quality monitoring, membrane bioreactor optimization, nutrient removal enhancement, and advanced process control [13].

Despite the remarkable progress, several challenges limit the widespread implementation of AI/ML in the water sector. These include data quality issues, the need for large and representative datasets, model interpretability concerns, the transferability of models across facilities, and the integration of AI tools into existing treatment infrastructure. Addressing these limitations is essential for transitioning AI from research-oriented applications to operational decision-making tools [14-16].

Although several reviews have examined AI/ML applications in water and wastewater treatment, most are limited to classical models (ANN, SVM, RF) and do not cover the rapidly emerging trends such as digital twins, advanced ensembles, and IoT-integrated real-time control [4,5,17]. This review distinguishes itself by (i) expanding the methodological scope to include hybrid evolutionary–deep learning models, (ii) synthesizing developments in soft sensors and IoT frameworks, and (iii) critically identifying challenges and future directions such as XAI, transfer learning, and GNN-based modeling—areas largely absent in previous reviews. Therefore, this review consolidates current knowledge on AI/ML applications in water and wastewater treatment, offering (i) an overview of commonly used algorithms and their mechanisms, (ii) a critical discussion of their strengths and limitations, (iii) an in-depth assessment of their applications in major treatment processes, and (iv) insights into future research directions.

2. Selection of Papers

The review covers 1998–2025, capturing the evolution of AI/ML from early ANN applications to modern deep learning and hybrid frameworks. The selection of articles was based on a scientometric approach, which enables a quantitative analysis of research trends, publication patterns, and thematic developments in the related field of research [17-20].

The data were collected from bibliographic databases, including Scopus, Web of Science, PubMed, IEEE, and Google Scholar. The key topics searched in this study included: artificial intelligence, wastewater treatment, water treatment, coagulation, flocculation, disinfection, membrane, machine learning, disinfection, biological treatment, adsorption, secondary treatment, tertiary treatment, advanced treatment, maintenance, data monitoring, costs, anomaly detection, water quality, energy, smart systems, and primary treatment. Studies were included if they: applied AI/ML or metaheuristic algorithms to water or wastewater treatment; were peer-reviewed journal articles; presented quantifiable outcomes; were published in English. Studies were excluded if they: lacked technical depth or quantitative results; were purely conceptual without implementation; did not relate to water/wastewater applications.

The literature was identified using a PRISMA-based work [17-20]. AI visualization and mapping tools such as Connected Papers, Litmaps, ResearchRabbit, and Semantic Scholar were utilized to expand the network of related works and track emerging research clusters.

Fig. 1 illustrates the most relevant journals to the subject of AI/ML in WTPs and WWTPs. As shown in Fig. 1, according to the search for references done in this paper, some journals have focused more on the use of AI or ML in treatment processes related to this work such as '*Science of the total environment', 'Process safety and environmental protection', 'Journal of membrane science', and 'Chemical Engineering Journal'.* Other journals such as '*Membranes*', '*Desalination and Water Treatment*', and '*International*

Journal of Environmental Research' were of secondary importance.

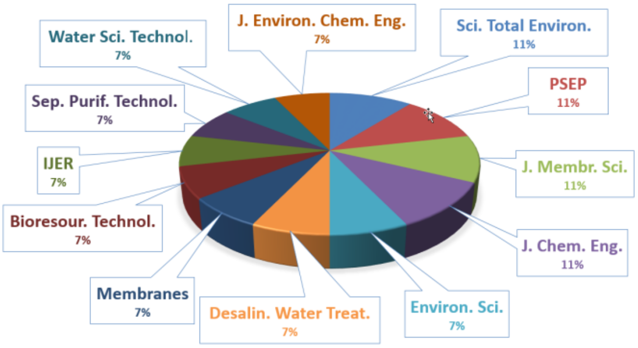


Fig. 1: The percentage of journals publishing in this field.

This section outlines how the pool of literature was narrowed from an initial set of 250 publications to a final selection of 55 studies, all of which address AI and ML applications in water and wastewater treatment processes. To illustrate this screening and refinement procedure, the PRISMA 2020 framework was employed. PRISMA 2020 is an internationally recognized standard designed to ensure transparency and rigor in systematic review reporting [17,18]. The following classification provides a structured overview of the search, screening, and eligibility stages, making clear which studies were retained or excluded and the rationale behind those decisions.

i) Identification

- Records identified through database searching: 150
- Records removed before screening (duplicates, irrelevant titles, automation tools): 40
- Records after removal: 110

ii) Screening

- Records screened by title/abstract: 40
- Records excluded (not relevant, outside scope, poor quality): 50
- Full-text articles assessed for eligibility: 60

iii) Eligibility

- Full-text articles excluded (reasons: insufficient data, non-AI/ML, not water-related, non-English): 16
- Studies included in qualitative synthesis: 55

iv) Included

- Final studies included in the systematic review: 55

It is notable that clear temporal evolution was observed across the 55 studies included in this review.

Early ANN-Centric Period: Research during this phase relied almost exclusively on multilayer perceptron and feedforward ANN models. Applications were primarily limited to coagulation modeling, turbidity prediction, and early membrane filtration studies. ANN-based coagulant dosing models dominated this era [13,17,18].

Classical ML Expansion: From 2010 onward, there was a clear shift toward more sophisticated models such as SVM, RF, ANFIS, and PCA-assisted architectures. Studies increasingly targeted membrane fouling prediction, adsorption modeling, and DBPs. Metaheuristic optimization (GA, PSO) also emerged to improve parameter calibration [3-5].

Deep Learning and Soft Sensors: Beginning in 2018, deep learning saw rapid adoption, particularly LSTM and CNN models for fouling image analysis, TMP prediction, and time-series forecasting in WWTPs. The rise of IoT-enabled monitoring platforms contributed to the transition toward soft sensing, real-time prediction, and multimodal data fusion [5,18,19].

Next-Generation AI: In the most recent period, the field diversified toward advanced ensemble learners (XGBoost), hybrid ML–mechanistic models, transfer learning approaches for cross-plant generalizability, and early exploration of digital twins. There is also emerging interest in explainable AI (XAI), as well as reinforcement learning and graph neural networks for distributed water systems [3,5,19,20].

3. Overview of AI/ML Techniques

3.1. Introduction to AI/ML methods

AI is an umbrella term encompassing all computational models that mimic human cognitive functions. However, machine learning can be classified as supervised (SVM, RF, DT, XGBoost), unsupervised (PCA, K-means, SOM), deep learning (CNN, RNN, LSTM, Autoencoders), metaheuristic optimization algorithms (GA, PSO, DE, GEP), and hybrid and integrated approaches (ANN-GA, PSO-SVR, CNN-Autoencoder, PCA-ANN, ML–Mechanistic hybrids) [15-20].

Different AI/ML methods are selected according to their potential. The rational for selecting some of the AI/ML methods is as follows: SVM (effective for nonlinear coagulation modelling); RF/GBM (robust for noisy water-quality datasets); LSTM (excels for influent fluctuation prediction); CNN (ideal for fouling image analysis); Metaheuristics (optimize dosage, backwash cycles, and MBR operating conditions) [3,19,20].

3.2. Common techniques

Although several types of algorithms and techniques have been applied in water treatment processes, a brief introduction of the most commonly used algorithms and methods is provided to facilitate the understanding of the subject.

3.2.1. K-Nearest Neighbor (KNN)

KNN is a classification and regression algorithm that predicts a data point's label based on its K nearest neighbors [17-20]. It handles nonlinear relationships and diverse data types well, but struggles with high-dimensional data, noise, and large datasets due to computational inefficiency. Fig. 2(a) illustrates KNN classification with K = 3 and K = 5 [21,22]. KNN is suitable for water treatment because many water quality parameters (e.g., turbidity, TOC, UV254) exhibit local similarity patterns that KNN can exploit without requiring a predefined model structure [21].

3.2.2. Fuzzy Neural Network (FNN)

FNNs combine fuzzy logic and artificial neural networks to effectively handle uncertainty and ambiguity [18,23]. They employ fuzzy inference, membership functions, and normalization alongside neural network weights, with a typical architecture of input, membership, inference, normalization, and output layers [20,24]. Compared to traditional neural networks, FNNs provide stronger problem-solving abilities and are widely applied in control systems, pattern recognition, and predictive analytics (Fig. 2(b)) [17,19]. FNNs are beneficial for treatment plants where raw-water quality fluctuates significantly, capturing uncertainty and imprecision [19].

3.2.3. Recurrent Neural Network (RNN)

RNNs are neural networks with feedback connections designed for sequential data processing. They leverage parameter sharing, memory retention, and Turing completeness to model time series dynamics [25]. Long Short-Term Memory (LSTM) networks extend RNNs with gating mechanisms to overcome conventional limitations. Fig. 2(c) shows the basic reused as inputs [17-19]. RNNs capture sequential patterns and are ideal for processes with strong temporal dependencies such as aeration control, membrane fouling, or COD removal [18].

3.2.4. Support Vector Machine (SVM)

SVMs are used for classification and regression by mapping data into higher-dimensional spaces with kernel functions to find an optimal hyperplane that maximizes the margin between classes [19,20,22]. They are effective in handling high-dimensional data and nonlinear relationships while avoiding local optima. However, SVMs can be computationally expensive on large datasets, sensitive to noise and outliers, and require careful kernel and hyper-parameter selection [17,18,26]. Fig. 2(d) illustrates the SVM decision boundary and margins. SVMs are effective for predicting water-treatment outcomes when relationships are nonlinear but datasets are not large enough for deep learning [26].

3.2.5. Random Forest (RF)

RF are ensemble models for classification and regression that combine multiple decision trees trained on random subsets of data [18,20,27]. Predictions are made by voting (classification) or averaging (regression). RFs handle high-dimensional data and nonlinear relationships well, with strong resistance to overfitting, but can be computationally intensive and sensitive to noise and outliers [17,19,28]. Fig. 2(e) depicts RF classification. RF is highly suitable for wastewater treatment due to its robustness to noise and ability to capture interactions among numerous operational variables [17].

3.2.6. Self-Organizing Map (SOM)

SOMs are clustering and dimensionality reduction algorithms that project high-dimensional data onto a lower-dimensional space using competitive learning while preserving data topology. They simplify visualization and classification and are widely applied in clustering, data visualization, and related tasks [17,18,20,29]. Fig. 2(f) illustrates the SOM method. SOM is valuable for clustering water quality or membrane fouling patterns, enabling operators to classify treatment conditions without labeled data [18].

3.2.7. Convolutional Neural Network (CNN)

CNNs are deep feedforward models built on convolutional operations, widely used in deep learning for their strong feature learning ability. They extract hierarchical features with convolutional layers, reduce dimensionality through pooling, and perform classification or regression via fully connected layers [17,18,30]. Fig. 2(g) shows a typical CNN structure. CNNs excel at learning spatial and spectral patterns from images and multichannel data such as membrane fouling images [18].

3.2.8. Principal Component Analysis (PCA)

PCA is a statistical method for dimensionality reduction that transforms high-dimensional data into a lower-dimensional space while retaining most of the original variance. It identifies orthogonal principal components ranked by the amount of variance they explain, thereby helping to reveal the underlying data structure [17,19,20,31]. Fig. 2(h) illustrates PCA with components PC_1 and PC_2. PCA reduces dimensionality in large water-quality datasets, improving model performance and identifying dominant pollution sources [20].

3.2.9. Genetic Algorithm (GA)

GAs are optimization methods inspired by natural evolution, using selection, crossover, and mutation to evolve better solutions. They excel at global search in multidimensional spaces, avoiding local optima, and are effective for complex nonlinear optimization problems due to their adaptability and robustness [17,18,32]. Fig. 2(i) illustrates the GA process. GA is effective for optimizing operating conditions (dosage, aeration rate, pH, HRT), especially when the search space is large and nonlinear [32].

3.2.10. Genetic Programming (GP)

GP is an evolutionary method based on Genetic Algorithms that evolves computer programs through operations such as selection, crossover, and mutation. It automatically generates candidate programs and refines them over generations, making it well-suited for complex nonlinear problems and tasks without explicit analytical solutions [17,18,33]. Fig. 2(j) illustrates the GP process. GP automatically generates mathematical expressions describing water-quality relationships, helping derive interpretable models for membrane flux or pollutant removal [17].

3.2.11. Particle Swarm Optimization (PSO)

PSO is an optimization technique inspired by the collective behavior of biological populations. It updates particle positions and velocities based on individual and group best solutions to search for optima. PSO effectively handles complex nonlinear problems and often avoids local optima, but its performance is sensitive to initial conditions and may require multiple runs [17,18,19,34]. Fig. 2(k) depicts the PSO process. PSO is widely used for calibrating ANN models and optimizing operational parameters in water treatment due to its fast convergence [19].

3.2.12. Deep Neural Network (DNN)

DNNs are artificial neural networks with multiple hidden layers that learn hierarchical data representations by combining lower-level features into higher-level abstractions. They are effective for modeling complex nonlinear relationships and extracting intricate patterns but require large datasets and high computational resources for training [17,19,35]. Fig. 2(l) shows a typical DNN structure. DNNs model intricate nonlinear relationships in large-scale water treatment datasets, such as full-plant SCADA sensor networks [19].

3.2.13. Model Predictive Control (MPC)

MPC is an advanced control method that predicts system behavior over a time horizon using a dynamic model and solves an optimization problem at each step to determine optimal control actions under constraints. Only the first action is implemented before recalculating, enabling effective handling of multivariable systems. Fig. 2(m) illustrates a typical MPC structure [17,18,36]. MPC is ideal for real-time water treatment control because it predicts future system states and optimizes control actions under constraints [18].

3.2.14. Gene Expression Programming (GEP)

GEP is an evolutionary algorithm that evolves models using chromosomes encoded as linear genomes, which are later expressed as nonlinear expression trees. Combining concepts from genetic algorithms and genetic programming, GEP is effective at uncovering complex data relationships and is applied in symbolic regression, classification, and modeling [19,20]. Fig. 2(n) depicts a typical GEP structure. GEP produces interpretable mathematical models suitable for predicting pollutant removal or membrane fouling [19].

3.2.15. Extreme Gradient Boosting (XGBoost)

XGBoost is a powerful machine learning algorithm based on gradient boosting decision trees. It builds models in a stage-wise fashion, where each new tree corrects the errors of the previous ones. XGBoost is known for its speed, accuracy, and ability to handle large datasets, missing values, and overfitting through regularization. It is widely used for classification, regression, and ranking tasks in data science competitions and real-world applications [17,37]. A schematic description of a typical XGBoost structure is shown in Fig. 2(o). XGBoost efficiently handles irregular, incomplete water-quality datasets and provides high accuracy with built-in regularization [17].

3.2.16. Nonlinear AutoRegressive model with eXogenous inputs (NARX)

NARX is a type of recurrent neural network used for modeling and predicting time series data. It captures nonlinear relationships by using past values of the target variable (autoregression) and past values of external (exogenous) inputs to forecast future outputs. NARX is well-suited for dynamic systems and is commonly applied in system identification, control, and signal processing tasks [19,25]. A schematic description of a typical NARX structure is shown in Fig. 2(p). NARX is ideal for dynamic systems like coagulation or MBRs because it explicitly incorporates past values of both inputs and outputs.

3.2.17. Long Short-Term Memory (LSTM)

LSTM is a type of RNN designed to learn and remember long-term dependencies in sequential data. It uses memory cells and gating mechanisms (input, forget, and output gates) to control the flow of information and avoid issues like vanishing gradients. LSTM is widely used in tasks, such as time series forecasting, natural language processing, and speech recognition due to its ability to handle complex temporal patterns [20,38]. A schematic description of a typical LSTM structure is shown in Fig. 2(q). LSTMs are powerful for long-term dependency modeling, making them well-suited for predicting slow-changing fouling processes or diurnal wastewater patterns.

3.2.18. Decision Tree (DT)

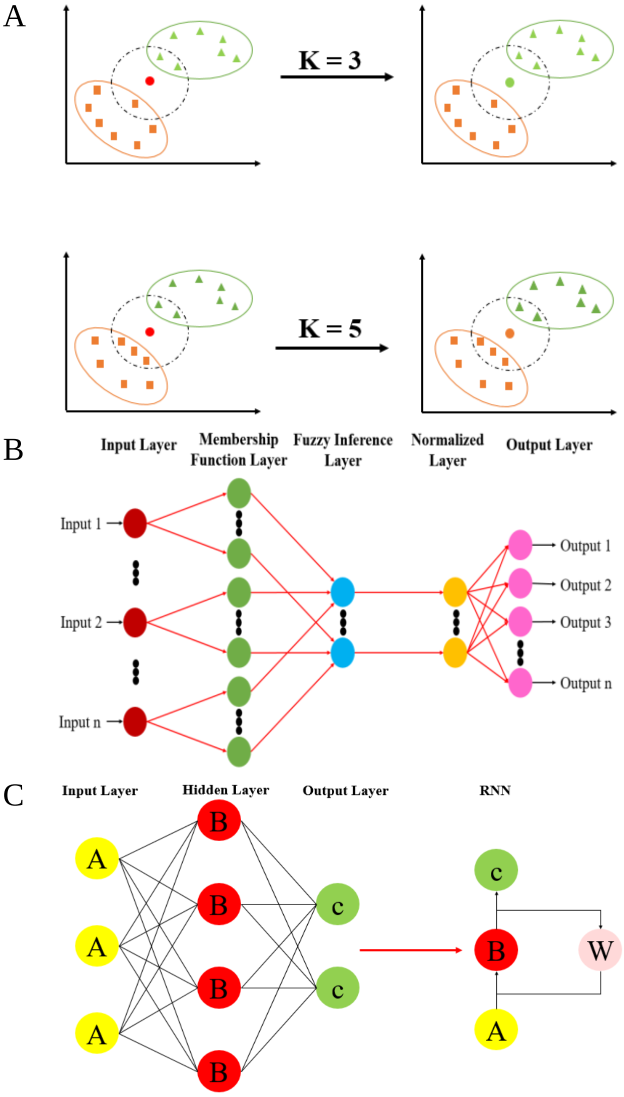
DT is a supervised learning algorithm used for classification and regression tasks. It splits the data into branches based on feature values, forming a tree-like structure where each internal node represents a decision based on an attribute, and each leaf node represents an output or class label. Decision Trees are easy to interpret, handle both numerical and categorical data, and are commonly used for their simplicity and effectiveness [17,19,20]. A schematic description of a typical DT structure is shown in Fig. 2(r). DTs provide interpretable rules for operational decision-making, helping operators understand key drivers of treatment performance [20].

3.2.19. Hybrid AI Models

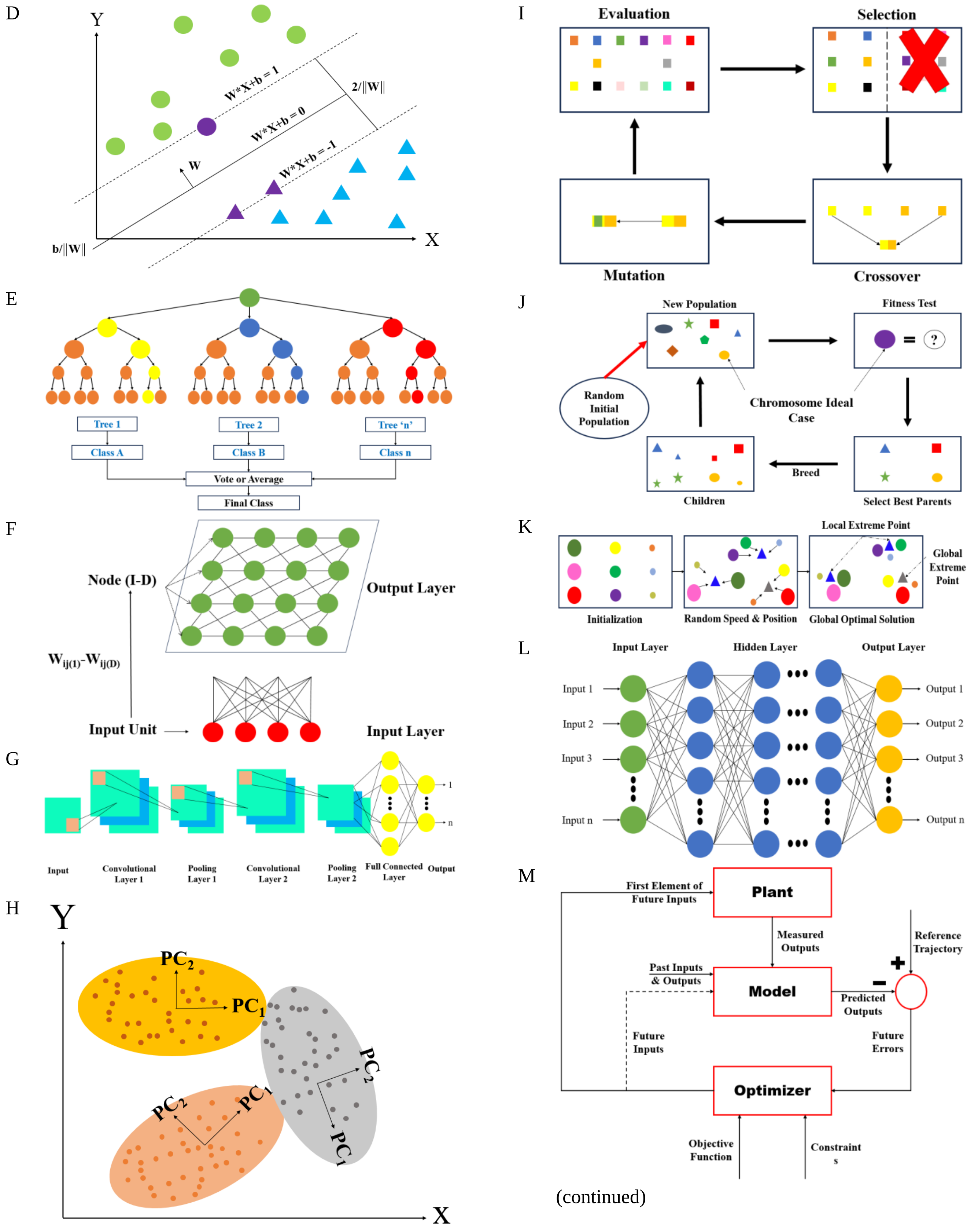
Finally, hybrid AI models can be used to overcome the main drawbacks of the individual models to enhance the performance of the integrated model. In fact, they exhibit better learning and prediction qualifications when facing more complicated nonlinear challenges. When RNN, PSO, SVM and GA are combined with other models, more helpful hybrid models like GA-ANN, GA-SVR, PSO-RNN, ANN-GANN, PSO-SVM, and PSO-ANN could be achieved [39]. These models proved to be effective in solving complicated and new environmental concerns. Hybrid models combine complementary strengths (e.g., ANN + PSO or SVM + GA), enabling superior prediction and optimization in highly nonlinear water treatment processes [17,20].

3.2.20. Emerging ML Paradigms for WTPs and WWTPs

Modern ensemble methods—particularly LightGBM, XGBoost, CatBoost, and stacked generalization—offer superior performance for multi-variable, non-linear water quality datasets due to their ability to reduce bias and variance simultaneously [3,5,13]. In addition, Reinforcement learning (RL) is emerging as a promising framework for dynamic optimization, enabling self-learning control strategies that adapt to fluctuating influent conditions. Meanwhile, in water networks, sensor nodes form interconnected graphs. Graph Neural Networks (GNNs) can naturally model these relational dependencies, enabling improved leak detection, anomaly detection, pipe failure prognosis, and distributed water quality prediction [17-20]. Physics-informed machine learning integrates mechanistic process knowledge with data-driven models, embedding conservation laws, mass-balance equations, and kinetic expressions directly into the learning framework [25,27,29,34]. A digital twin is a dynamic, virtual replica of a treatment unit or entire plant that continuously synchronizes with real-time operational data. Digital twins allow the operators to test chemical dosing strategies, aeration profiles, membrane backwashing schedules, or energy-efficiency scenarios without risking plant performance. Recent studies have begun integrating ML-based soft sensors, reinforcement learning, and mechanistic simulators within digital twins to support proactive decision-making and predictive maintenance [17,19,25]. The adoption of Internet of Things (IoT) platforms in water treatment plants has led to high-frequency, multi-sensor data streams that augment traditional laboratory sampling. IoT-enabled soft sensors use ML models to estimate hard-to-measure variables (e.g., COD, ammonia, DBP precursors) based on online surrogates like UV254, ORP, DO, or turbidity. These systems improve real-time monitoring accuracy and support closed-loop process control strategies [3,4,19].



(continued)


D
Y
X
W*X+b = 1
W*X+b = 0
W*X+b = -1
2/||W||
b/||W||
W
E
Tree 1
Tree 2
Tree 'n'
Class A
Class B
Class n
Vote or Average
Final Class
F
Node (I-D)
Output Layer
W_ij(1)-W_ij(D)
Input Unit
Input Layer
G
Input
Convolutional Layer 1
Pooling Layer 1
Convolutional Layer 2
Pooling Layer 2
Full Connected Layer
Output
1
n
H
Y
X
PC2
PC1
PC2
PC1
PC2
PC1
I
Evaluation
Selection
Mutation
Crossover
J
New Population
Fitness Test
Random Initial Population
Chromosome Ideal Case
?
Breed
Children
Select Best Parents
K
Local Extreme Point
Global Extreme Point
Initialization
Random Speed & Position
Global Optimal Solution
L
Input Layer
Hidden Layer
Output Layer
Input 1
Input 2
Input 3
Input n
Output 1
Output 2
Output 3
Output n
M
Plant
Model
Optimizer
First Element of Future Inputs
Measured Outputs
Past Inputs & Outputs
Predicted Outputs
Reference Trajectory
+
−
Future Inputs
Future Errors
Objective Function
Constraints

(continued)

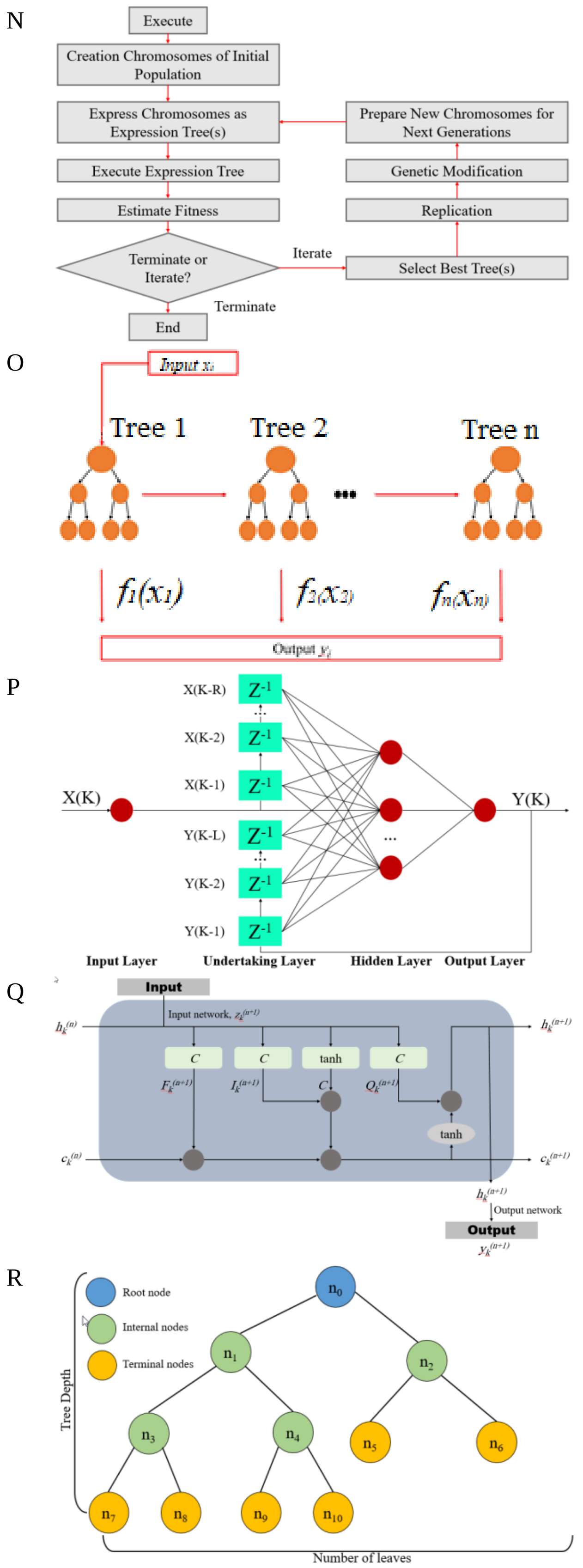


Fig. 2: A schematic description of AI/ML models and techniques. (A) KNN; (B) FNN; (C) RNN; (D) SVM; (E) RF; (F) SOM; (G) CNN; (H) PCA; (I) GA; (J) GP; (K) PSO; (L) DNN; (M) MPC; (N) GEP; (O) XGBoost; (P) NARX; (Q) LSTM; (R) DT [3,13,17,19,20].

3.2. Benefits and drawbacks of AI/ML

Among all the aforementioned algorithms and methods, the significant ones have certain advantages and disadvantages. Table 1 shows the main benefits and drawbacks of each method in brief to highlight the applicability of them in real-world problems. The importance of understanding these pros and cons is that it helps to choose the most efficient method for different purposes and the nature of available data.

Table 1: Benefits and drawbacks of some AI/ML methods [17-20].

Algorithm	Pros	Cons
KNN	Simple, no training; Good for nonlinear patterns	Slow with large datasets; sensitive to noise; not suitable for high dimensions
SVM / SVR	High-dimensional and nonlinear data; avoids Local minima	Computationally expensive; requires careful kernel selection
DT	Easy to interpret; Handles categorical & numerical data	Prone to overfitting; sensitive to small data changes
RF	Robust, reduces overfitting; high accuracy	Less interpretable; slower for large ensembles
XGBoost	High predictive accuracy; Handles missing data; strong regularization	Computationally intensive; many hyperparameters
ANN / MLP	Good for complex nonlinear relationships	Requires large data; prone to overfitting; “black-box” nature
RNN / LSTM / NARX	Excellent for time-series prediction; captures temporal dependencies	Long training time; requires large, sequential datasets
CNN	Powerful feature extraction (images, spectra); good for pattern recognition	High computational demand; needs large labeled datasets
Fuzzy Logic / ANFIS	Interpretable; handles uncertainty and imprecision well	Not ideal for high-dimensional data; rule explosion possible
PCA	Effective dimensionality reduction; reduces noise	Loses variance/information; linear method only
GA / GP / GEP	Strong for global search & optimization; avoids local minima	Slow convergence; computationally expensive
PSO	Fast, simple, good for nonlinear optimization	Sensitive to parameter tuning; may get stuck in local optima
Hybrid Models	Combine strengths; improved accuracy and robustness	Higher complexity; greater computational cost

3.3. Common activation functions

In AI/ML models, activation functions play a significant role in identifying non-linearity into the network that makes learning sophisticated patterns and connections in the dataset possible. Without the presence of them, neural networks can only learn linear functions regardless of the number of layers. This would restrict their power that makes them act similarly to a basic linear regression model [17,19,20,40]. In Table 2, the most common activation functions which are used in AI/ML models are presented.

3.4. The flowchart of AI/ML model development

The flowchart for developing an AI/ML model outlines the step-by-step process involved in building intelligent systems. It typically includes the following stages: problem identification, data collection, data processing, model selection, training and test data separation, model evaluation, and model tuning. Fig. 3 illustrates schematically different stages and processes in developing an AI/ML model.

Table 2: The most common activation functions used in models [3,13,18,20].

Activation Function & Description	Formula
Sigmoid Function	$f(x) = \dfrac{1}{1 + e^{-x}}$
ReLU Function	$f(x) = \max(0.x)$
ELU Function	$f(x) = \begin{cases} \alpha(e^{x} - 1).x < 0 \\ x.x \geq 0 \end{cases}$
Ramp Function	$\begin{cases} 0.x < T_1 \\ \dfrac{x - T_1}{T_2 - T_1}.T_1 \leq x \leq T_2 \\ 1.x \geq T_2 \end{cases}$
Tanh Function	$f(x) = \tanh(x)$
Leaky ReLU Function	$f(x) = \text{x if x} > 0. \text{else } \alpha x$
Heaviside	$f(x) = \begin{cases} 0.x < T \\ 1.x \geq T \end{cases}$

3.5. Popular frameworks and tools

Frameworks and tools in AI/ML are software libraries and platforms that introduce ready-made functions, modules, and infrastructures to help researchers build, train, test, evaluate and deploy models into real-world cases. They save time by presenting pre-built components, supporting complicated operations such as training neural networks, data preprocessing, and model testing, offering scalability and cloud integration, and deploying models simply into real applications. Table 3 presents the popular frameworks and tools [27,40].

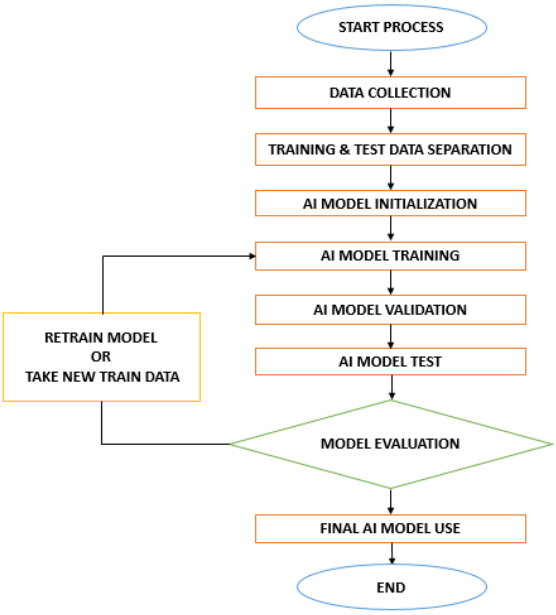


Fig. 3: Flowchart outlining the different stages in developing an AI/ML model [18].

3.5. Popular frameworks and tools

Frameworks and tools in AI/ML are software libraries and platforms that introduce ready-made functions, modules, and infrastructures to help researchers build, train, test, evaluate and deploy models into real-world cases. They save time by presenting pre-built components, supporting complicated operations such as training neural networks, data preprocessing, and model testing, offering scalability and cloud integration, and deploying models simply into real applications. Table 3 presents the popular frameworks and tools [27,40].

Table 3: Popular frameworks and tools [17-20].

Tool/Framework	Description
TensorFlow	- Google open-source library - Creating and teaching deep learning models. - Supports low-level customization and high-level APIs.
Scikit-learn	- An efficient Python library for traditional ML. - Great for beginners and small-scale projects.
PyTorch	- Developed by Meta (Facebook) - A dynamic framework for deep learning with a Pythonic style. - Popular in research and production.
Keras	- High-level API for creating neural networks. - Easy to use and good for quick prototyping.
MATLAB	- A proprietary platform used in academia and engineering. - Includes toolboxes for ML and deep learning.
Jupyter Notebooks	- Not a framework but a key tool - An interactive environment.
RapidMiner/ KNIME	- Visual (drag-and-drop) platforms for ML, good for non-programmers. - Useful in business applications.

3.6. A summary of highlighted papers

In this section, a summary of objectives, datasets, techniques, and outcomes for representative studies is presented. As providing data for all 55 references might be unnecessary, some highlighted ones are presented in Table 4.

4. The Use of AI/ML in Water Treatment

AI techniques have been utilized to create models and enhance various processes within water treatment plants (WTPs), including coagulation and flocculation, membrane filtration, adsorption, the formation and control of disinfection byproducts (DBPs), and the management of water quality. In the following, these applications are briefly discussed.

4.1 Coagulation and flocculation

Coagulation–flocculation (C/F) is a key water-treatment process used to remove natural organic matter and colloids, with alum dosed based on jar tests and online water-quality data. Sudden changes in raw water—such as from rainfall or drought—make dosage selection difficult, motivating the use of AI for more reliable predictions. AI models typically estimate coagulant dosage using 3–12 inputs, including turbidity, pH, temperature, alkalinity, TOC/DOC, color, UV254, conductivity, and other routine parameters [19,30].

Table 4: A summary of techniques, objectives, and outcomes.

Ref.	**Treatment Process**	**Objective**	**Dataset / Scale**	**Input Variables**	**AI / ML Technique(s)**	**Key Outcomes**
[1]	Coagulation–Flocculation	Predict optimal coagulant dosage	Full-scale WTP	Turbidity, pH, alkalinity, DOC	ANN	High prediction accuracy ($R^2 > 0.95$) and improved dosing stability
[4]	Adsorption	Predict dye removal efficiency	Batch adsorption experiments	Initial concentration, pH, dosage, time	RF	RF achieved lowest RMSE among tested models
[5]	Adsorption	Model heavy metal removal	Lab-scale	pH, adsorbent dose, contact time	ANN–GA	Reduced prediction error compared with standalone ANN
[7]	Membrane Filtration (MBR)	TMP forecasting	Full-scale MBR	TMP history, MLSS, flow rate	LSTM	Captured long-term fouling trends
[8]	Membrane Fouling	Fouling type classification	Experimental	SEM images	CNN	>90% classification accuracy
[10]	Disinfection Byproducts	Model HAA formation	Pilot-scale	TOC, bromide, contact time	XGBoost	Higher accuracy than ANN and SVM
[11]	Wastewater Treatment	Predict effluent COD	Full-scale WWTP	Flow, influent COD, DO	RF	Robust prediction under influent variability
[13]	WWTP Monitoring	Soft sensor development	Full-scale	Online sensors (ORP, DO, pH)	ANN	Enabled real-time COD estimation
[14]	Predictive Maintenance	Estimate membrane replacement time	Full-scale MBR	TMP, energy use, flow	AI-based degradation model	Accurate remaining useful life prediction
[15]	Process Optimization	Optimize aeration energy	Simulation-based	DO, airflow, load	PSO–ANN	Reduced energy demand while maintaining effluent quality

Among these models, MLP with backpropagation (BPN) is the most common, though Extreme Learning Machines (ELM) offer much faster training (<1 s compared to 60–120 s for BPN). Frequently used activation functions include sigmoid, tansig, ReLU, and radial-basis functions (RBF). Hybrid and alternative approaches—such as SVR-KNN-RBF, ANFIS-GRID/SUB, GRNN, CNNs, and variable-structure networks—have also been applied, often outperforming traditional methods [19,25,41].

Studies show strong predictive performance across different regions and plant sizes. For example, RBF-based models in Alberta, ANFIS-GRID in Algeria, ANFIS-SUB in Iran, and CNN models for synthetic and natural waters all achieved high R^2 values (>0.85–0.96). Time-lagged and recurrent networks such as NARX have also delivered high accuracy ($R^2 \approx 0.91$–0.95). Comparisons frequently show GRNN, SA-MLP, and optimized MLP models outperforming RBF-ANN, fuzzy regression, or GA-MLP models. ANN-based systems have also been used for predicting TOC removal and optimizing C/F costs, with one German WTP estimating ~15% cost savings using ANN-assisted coagulation control [17,20,24]. Despite the promising performance of ANN, SVM, and hybrid models in predicting coagulation efficiency, none of the reviewed studies evaluated model transferability across different water sources, indicating a gap in external validation [17].

4.2 Membrane processes

Research on membrane technologies focuses on predicting fouling behavior, permeate flux, flux decline, and contaminant or metal removal. For microfiltration, Fuzzy

Logic (FL) and GA-optimized models accurately predicted flux decline, and FL models achieved R^2 values of 0.96–0.99. Semi-empirical and ANN models showed comparable performance for flux decline and membrane rejection [18,20].

In ultrafiltration, ANN-based pulsation increased permeation flux by 21%, and MLP-BPN models reliably predicted transmembrane pressure in pilot systems with sequential backwashes. Hybrid regression–ANN models also effectively estimated flux under varying conditions such as protein content, membrane type, and filtration mode. Fouling prediction using Darcy’s law with ANN achieved very low MSE, and studies confirmed strong effects of proteins and colloids on reversible and irreversible fouling. PCA applied to FEEM data further improved fouling recognition [17,18].

For nanofiltration and reverse osmosis, PCA-assisted monitoring helped detect fouling early, while ANN, SVM, and RF models predicted flux and salt rejection with $R^2 > 0.94$. CNN models used high-resolution fouling images to predict NF/RO flux decline and fouling thickness. PCA–GA hybrid modeling also optimized backwash timing, yielding up to 9.8% energy savings. ANN models accurately predicted contaminant removal by nanofiltration in municipal waters [3,20].

Machine learning has also been applied to polymer inclusion membranes (PIMs) for metal removal. MLP-ANN outperformed MLR in predicting heavy-metal removal, and ANN exceeded ANFIS in modeling Cr (VI) transport, though some studies reported ANFIS to be superior depending on data quality and structure. ANN models successfully estimated arsenic removal ($R^2 > 0.93$), and ANN, RNN, and GRNN techniques have been used to study cadmium removal [17,19]. Although deep learning techniques such as CNN and LSTM have shown strong accuracy in fouling prediction, few studies incorporated real-time sensor noise or operational disturbances, revealing a gap in robustness testing [20].

4.3 Adsorption

AI, particularly ANN-based models, has been widely used to predict pollutant removal in adsorption processes. MLP-ANN has accurately estimated removal efficiencies for metals such as As (III), Hg^{2+}, Ni^{2+}, and others using advanced adsorbents, often achieving very high R^2 values (0.98–0.999). Similar models have been applied to nanocomposites like MOF–LDH materials and ordered nanostructure adsorbents, consistently showing strong predictive ability.

Comparative studies using FF-BPN, CCF-BPN, GRNN, and Gradient Boosting found GRNN to be the most effective for predicting Pb^{2+}, Cu^{2+}, and Ni^{2+} removal using biochar. MLP-ANN combined with the LM algorithm has also been used to model inorganic and organic pollutant levels, while PC-ANN outperformed MLR and PCR in predicting As species. ANN models supported by SOM segmentation successfully predicted $KMnO_4$ concentrations [17,18,24].

Hybrid techniques such as ANN–PSO further improved predictions for methylene blue extraction compared to RSM or standalone ANN models. Overall, ANN-based approaches consistently provide the best performance for modeling adsorption of metals and organics—including Pb^{2+}, Cu^{2+}, Ni^{2+}, Cd^{2+}, Hg^{2+}, Cr (VI), and As (III)—across a wide range of adsorbents [19,20]. Metaheuristic-enhanced models improved adsorption predictions; however, no benchmark datasets exist for comparing models across adsorbents or operating conditions, representing a major research gap [20].

4.4 DBPs formation and control

AI is widely used to model and control disinfection byproduct (DBP) formation, helping optimize disinfectant dosing and reduce health risks from long-term DBP exposure. Traditional empirical and mechanistic DBP models often struggle due to site-specific variability and uncertainty, making AI a valuable alternative. PCA has been applied to reduce data dimensionality and identify major factors influencing DBPs—such as THMs, HAAs, nitrosamines—while lowering computational complexity. Key input variables typically include DOC, pH, temperature, and disinfectant dose [3,42].

ANN models have successfully predicted bromate and numerous DBPs, achieving R^2 values of 0.81–0.98. Hybrid PCA–ANN approaches showed strong performance for THMs, HAAs, and TOX, while autoencoders combined with PCA effectively reduced dimensionality for THM/HAA prediction. Other integrated models, such as linear regression–ANN, ANN–SVM, and ANN–gene expression programming, have also demonstrated high accuracy, often exceeding correlation coefficients of 0.92 [17,43,44]. Studies consistently used small or site-specific datasets for DBP prediction, and none investigated cross-season or multi-plant generalization, highlighting a gap in long-term validation. Fig. 4 shows a graphical summary of the use of AI/ML in WTPs to create a better understanding of their effectiveness.

5. The Use of AI/ML in Wastewater Treatment

AI models and IoT framework and conventional approaches are effective for designing smart WWT systems and sewage reuse [3]. An AI model is a helpful and powerful tool for prediction, optimization, and modeling of the WWT process and has been mostly used in different aspects of WWT like removal of dyes, heavy metals, nutrients, organics, solids, microbial contamination, drugs and pesticides from water

[17]. In terms of research, AI models are commonly used in laboratory-scale research and process design. When it comes to practical applications, process design includes process parameter optimization and process performance prediction. In optimization and prediction, a large deal of data is needed to establish and train AI models that could be gained by monitoring water quality [18].

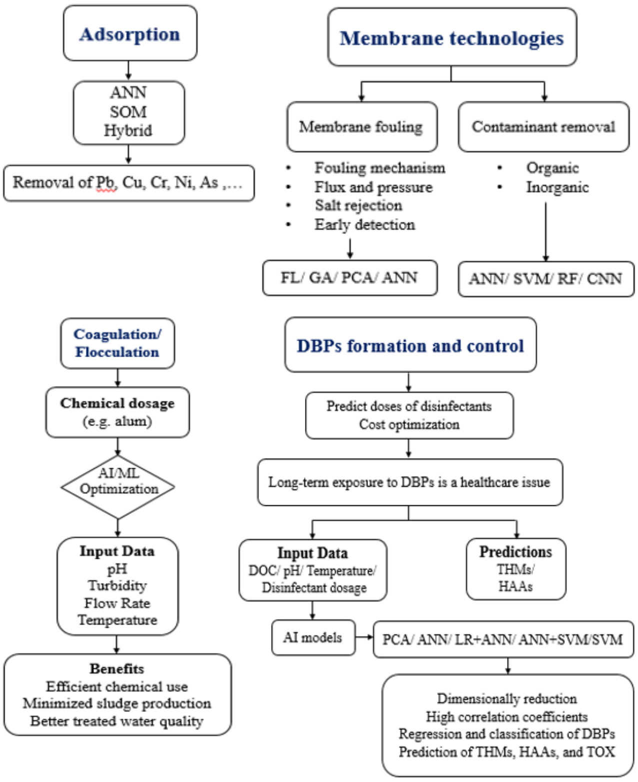


Fig. 4: The summary of the use of AI/ML in water treatment processes.

5.1. *Water Quality Monitoring for Data Acquisition*

AI-supported water quality monitoring in WWT relies heavily on sensor-generated data, with studies typically using 60–80% of data for training and the rest for testing. Sensors enable continuous measurement of influent parameters—such as BOD, COD, DO, pH, flow rate, temperature, and pollutant levels—and effluent parameters like BOD, DO, pH, and residual contaminants to assess WWTP performance. As data availability grows, real-time, data-driven modeling has become increasingly important [20,45].

Advanced approaches integrate spectroscopy with AI: for example, combining CNNs with laser-induced Raman and fluorescence spectroscopy to detect micropollutants in real time ($R^2 = 0.74$), outperforming traditional WWTP monitoring methods. Blending fluorescence and Raman signals with AI has proven effective for image-based recognition in microbiological water analysis, micropollutant detection, and operational adjustments [18,46].

Real-time remote monitoring systems have also been developed using PCA, ANN, and multivariate statistical process control, providing operators with key insights for efficient WWTP operation. IoT paired with AI offers high-accuracy water quality prediction and supports decision-making for public water management. Soft sensors—data-driven virtual sensors—have been shown to reliably monitor biological WWTP processes and significantly reduce maintenance needs, even in long-term on-site deployments [47,48].

5.2. *Laboratory-Scale Research*

Laboratory-scale research supports the development of new WWT technologies because it is inexpensive, controllable, and repeatable. At this scale, AI is mainly used to design and optimize membrane systems and bioreactors, especially membrane bioreactors (MBRs), which combine biological treatment with membrane filtration. MBRs provide high biomass levels, low sludge production, short HRT, small footprints, and high-quality effluent, while anaerobic MBRs (AnMBRs) offer even greater energy efficiency [3,17,50].

Membrane fouling remains the major challenge, reducing permeability and raising operating costs. AI models—applied to forward osmosis, RO, NF, UF, and MF—help predict or manage fouling. RBF-ANN models have been used to study interfacial energy on membrane surfaces, while ANN and hybrid ANN–RSM models accurately predict flux and improve process optimization [18,51].

AI also enhances MBR operational optimization, using inputs such as temperature, pH, DO, salinity, HRT, and pollutant loads. Multi-algorithm frameworks (ANN, SVR, ANFIS) have achieved high accuracy in predicting aerobic granular sludge reactor’s performance. BPNN and wavelet NN models effectively predict COD and TN removal in various MBR configurations, highlighting pH as a key operating factor [20,52].

5.3. *Process Design*

WWT involves complex process design and operating conditions, and AI/ML models help simplify these complexities by capturing nonlinear relationships between inputs (e.g., time, temperature, pH, influent quality) and outputs (e.g., removal efficiency, adsorption efficiency, effluent quality). Common AI applications include predicting effluent quality, assessing WWT performance, and optimizing energy use and operating parameters. For process design, AI is mainly used for optimizing operating conditions and predicting system performance [3,53].

Hybrid optimization models—such as ANN-GA, ANN-PSO, and RSM—have been used to improve biomass production, enhance decontamination efficiency, and optimize electro-oxidation processes. Studies show significant gains, including improved algal biomass productivity, high COD removal, and more accurate prediction of optimal conditions. SOM models have also been applied to improve nutrient removal by analyzing long-term aeration cycles [18,20,54].

AI combined with spectroscopy (FTIR, SEM, Raman) enables rapid identification of functional groups and detection of micropollutants, with a CNNs providing strong spectral classification performance. Advanced deep learning models, such as CLSTMA, GRU, CNN-1D, and LSTM, have shown high accuracy for real-time WWT monitoring, biomass prediction, and N_2O emission modeling. Enhanced hybrid models like FFNN-LSSVM outperform conventional techniques, offering faster and more reliable predictions of key effluent parameters such as BOD and ammonia [3,19,53,55].

5.4. Emerging trends in treatment processes

Transfer learning enables models trained on one WWTP (with abundant data) to be adapted to another plant with limited data, addressing one of the major barriers in real-world AI deployment [45,46]. Moreover, anomaly detection is a critical application because early identification of shock loads, contamination peaks, or equipment malfunctions can prevent regulatory violations and environmental damage. Methods in this regard include Autoencoders, Isolation Forests, One-Class SVM, and GNN-based anomaly detectors [13,20]. Another important use is predictive maintenance. Predictive maintenance using RUL estimation—based on vibration sensors, pressure sensors, energy consumption patterns, and membrane TMP trends—can markedly reduce downtime and operational cost [17,29].

6. Future Research Direction

The present review provides a comprehensive and systematically structured assessment of 55 peer-reviewed studies using a PRISMA-based methodology, includes a detailed analysis of algorithmic and temporal trends, and integrates modern AI innovations not covered in earlier works. However, in some areas, not enough work is done; therefore, in the following, some important suggestions are presented:

6.1 Hybrid Mechanistic–AI Models

Most AI models function as “black boxes,” relying solely on data-driven learning. Future research should integrate mechanistic/physicochemical models with ML to improve model reliability, interpretability, and transferability across facilities.

6.2 Digital Twins for Water and Wastewater Systems

Developing AI-driven digital twins for coagulation, membrane filtration, and full-scale WWTPs can enable real-time optimization, failure prediction, and energy management under extreme climatic or influent variations.

6.3 Transfer Learning

Existing models often perform well only in the system on which they were trained. Transfer learning, domain adaptation, and meta-learning approaches are needed to generalize AI models between plants with different influent characteristics and operational practices.

6.4 Improved Data Acquisition and Sensor Technologies

Advanced spectroscopy–AI fusion, low-cost sensors, and robust soft sensors can generate more reliable datasets for AI training. The development of autonomous sensor fault detection and self-calibration algorithms will further enhance long-term monitoring.

6.5 AI for Emerging Contaminants

Future work should extend AI applications to complex micropollutants (PFAS, pharmaceuticals, endocrine disruptors, microplastics) where traditional models show limitations.

6.6 Real-Time Control and Automation

The development of AI-based supervisory control systems, reinforcement learning controllers, and predictive maintenance models could transition WWTPs into fully autonomous, self-optimizing facilities.

6.7 The use of XAI and GNN

Future research should place greater emphasis on explainable artificial intelligence (XAI) and graph neural networks (GNNs) to address current limitations in transparency and system representation within AI-based models for WTPs and WWTPs. XAI methods can improve model interpretability by identifying the relative influence of key operational and water-quality variables, thereby enhancing trust and practical applicability. In parallel, GNNs provide a promising framework for capturing the interconnected and networked structure of treatment processes and sensor systems, enabling more realistic modeling of spatial and process-level dependencies. The incorporation of these approaches may support the development of more reliable and transferable AI tools for treatment system analysis and operation.

7. Challenges and Limitations

While promising, the application of AI/ML in water and wastewater treatment faces several practical challenges. In the following, some of the most significant challenges and limitations are described.

7.1 Data Quality, Noise, and Missing Values

Treatment plants often generate incomplete, noisy, or inconsistent datasets due to sensor drift, manual sampling errors, and environmental disturbances. High-quality, representative datasets are essential for reliable AI/ML performance [3,17].

7.2 Lack of Standardization and Benchmark Datasets

There is no unified benchmark dataset for coagulation, membrane fouling, adsorption, or DBP formation. This makes it difficult to compare models or reproduce results across studies [18].

7.3 Model Interpretability and Operator Trust

Many AI models operate as black boxes, limiting transparency and acceptance by plant operators. Explainable AI tools are needed to interpret predictions and support regulatory compliance.

7.4 Limited Scalability and Real-World Deployment

Models performing well at laboratory scale often struggle when applied to full-scale, heterogeneous systems due to variable influent conditions, operational disturbances, and infrastructure constraints [3,17,20].

8. Conclusion

This review demonstrates that artificial intelligence and machine learning have become indispensable tools for advancing water and wastewater treatment processes. Across treatment stages—coagulation–flocculation, membrane filtration, adsorption, disinfection byproduct (DBP) control, and full wastewater treatment—AI/ML models consistently outperform traditional mechanistic and empirical approaches in prediction accuracy, adaptability to nonlinearity, and robustness under fluctuating water quality conditions. Neural-network-based models (MLP, CNN, RNN, LSTM, GRNN), hybrid optimization frameworks (ANN–GA, ANN–PSO, PCA–ANN), and data-driven soft sensing have enabled major improvements in contaminant removal efficiency, membrane flux prediction, fouling mitigation, coagulant dosage optimization, and DBP forecasting.

In wastewater treatment, the integration of IoT-enabled sensing, deep learning, and soft sensor architectures has made real-time monitoring and process automation increasingly feasible. Laboratory-scale research shows that AI is highly effective for optimizing membrane bioreactors (MBRs), anaerobic MBRs (AnMBRs), and biological treatment systems by improving the prediction of flux decline, sludge characteristics, nutrient removal, and operating conditions. Overall, AI/ML approaches allow operators to reduce chemical and energy consumption, enhance process stability, and support data-informed operational decision making.

Despite their proven potential, many AI applications remain confined to laboratory studies or offline simulations. Moving toward deployment in real-world treatment plants requires addressing critical challenges associated with data quality, model generalizability, and integration with existing control infrastructures. Nevertheless, with growing access to high-resolution data and advances in hybrid mechanistic–machine-learning modeling, AI is positioned to become a core technology shaping the next generation of intelligent, resilient, and energy-efficient water treatment systems.

References

[1] Biglarijoo, N., Shams, A., & Hadipour, V. (2025). Modeling Fenton Treatment for Landfill Leachate with the Focus on Initial COD and Sludge Generation. *Numerical Methods in Civil Engineering 9(3)*, 39-46.

[2] Biglarijoo, N., & Shams, A. (2024). Multivariate modeling of the Fenton process for enhanced COD removal and low sludge generation in landfill leachate treatment. *Journal of Applied Research in Water and Wastewater 11(2)*, 110-115.

[3] Altowayti, W.A.H., Shahir, S., Othman, N., Eisa, T.A.E., Yafooz, W.M.S., Al-Dhaqm, A., Yong Soon, C., Yahya, I.B., Rahim, N.A.N.B.C., Abaker, M., & Ali, A. (2022). The role of conventional methods and artificial intelligence in the wastewater treatment: a comprehensive review. *Processes 10*, 1832.

[4] Carreres-Prieto, D., Garcia, J.T., Carrillo, J.M., Vigueras-& Rodriguez, A. (2023). Towards highly economical and accurate wastewater sensors by reduced parts of the LED-visible spectrum. *Science of the total environment 871*, 162082.

[5] Emaminejad, S.A., Morgan, V.L., Kumar, K., Kavathekar, A., Ragush, C., Shuai,W., Jia, Z., Huffaker, R., Wells, G., & Cusick, R.D. (2022). Statistical and microbial analysis of bio-electrochemical sensors used for carbon monitoring at water resource recovery facilities. *Environmental science: water research and technology 8*, 2052–2064.

[6] Biglarijoo, N., Mirbagheri, S.A., Bagheri, M., & Ehteshami, M. (2017). Assessment of effective parameters in landfill leachate treatment and optimization of the process using neural network, genetic algorithm and response surface methodology. *Process safety and environmental protection 106*, 89-103.

[7] Chan, M. K., Shams, A., Wang, C. C., Lee, P. Y., Jahani, Y., & Mirbagheri, S. A. (2023). *A*rtificial neural network model for membrane desalination: A predictive and optimization study. *Computation 11(3)*, 68.

[8] Mirbagheri, S. A., Biglarijoo, N., & Nejad, M. K. (2014). Pilot plant studies for the removal of heavy metals from industrial wastewater using adsorbents. *Turkish Journal of Engineering and Environmental Sciences 38*(2), 159-166.

[9] Mirbagheri, S.A., Biglarijoo, N., Ahmadi, S., Razmara, P., & Yazdan Doost. A. (2016). Removing Fe, Zn and Mn from steel making plant wastewater using RO and NF membranes. *Iranian Journal of Health Sciences 4(4)*, 41-55

[10] Biglarijoo, N., Mirbagheri, S.A., Ehteshami, M., & Ghaznavi, S.M. (2016). Optimization of Fenton process using response surface methodology and analytic hierarchy process for landfill leachate treatment. *Process Safety and Environmental Protection 104 (Part A)*, 150-160.

[11] Mirbagheri, S.A., Ahmadi, S., & Biglarijoo, N. (2016). Denitrification of nitrate-contaminated groundwater in an anoxic rotating biological contactor: a case study. *Desalination and Water Treatment 57 (10)*, 4694-4700.

[12] Iorhemen, O.T., Hamza, R.A., & Tay, J.H. (2016). Membrane bioreactor (MBR) technology for wastewater treatment and reclamation: Membrane fouling. *Membranes 6*, 33.

[13] Valikhan Anaraki, M., Farzin, S., Ahmadianfar, I., & Shams, A. (2024). Development of a novel discharge routing method based on the large discharge dataset, Muskingum model, optimization methods, and multi-criteria decision making. *Journal of Soft Computing in Civil Engineering 8(4)*, 54–93.

[14] Mirbagheri, S.A., Biglarijoo, N., & Ahmadi, S. (2013). Efficinecy of two types of activated carbon columns to treat industrial wastewater: A case study. *International Journal of Geology 7(3)*, 88-93.

[15] Cai, Y., Li, X., Zaidi, A.A., Shi, Y., Zhang, K., Sun, P., & Lu, Z. (2019). Processing efficiency, simulation and enzyme activities analysis of an air-lift multilevel circulation membrane bioreactor (AMCMBR) on marine domestic sewage treatment. Periodica Polytechnica *Chemical Engineering 63*, 448–458.

[16] Chen, Y., Yu, G., Long, Y., Teng, J., You, X., Liao, B.Q., & Lin, H. (2019). Application of radial basis function artificial neural network to quantify interfacial energies related to membrane fouling in a membrane bioreactor. *Bioresource Technology* 293, 122103.

[17] Malviya, A., & Jaspal, D. (2021). Artificial intelligence as an upcoming technology in wastewater treatment: a comprehensive review. *Environmental Technology Reviews 10*(1), 177-187.

[18] Safeer, S., Pandey, R. P., Rehman, B., Safdar, T., Ahmad, I., Hasan, S. W., & Ullah, A. (2022). A review of artificial intelligence in water purification and wastewater treatment: Recent advancements. *Journal of Water Process Engineering 49*, 102974.

[19] Chowdhury, S., & Karanfil, T. (2024). Applications of artificial intelligence (AI) in drinking water treatment processes: Possibilities. *Chemosphere 356*, 141958

[20] Wang, Y., Cheng, Y., Liu, H., Guo, Q., Dai, C., Zhao, M., & Liu, D. (2023). A review on applications of artificial intelligence in wastewater treatment. *Sustainability 15*(18), 13557.

[21] Maier, H. R., Morgan, N., & Chow, C. W. (2004). Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. *Environmental Modelling & Software 19*(5), 485-494.

[22] Jayaweera, C. D., & Aziz, N. (2021). An efficient neural network model for aiding the coagulation process of water treatment plants. *Environment, Development and Sustainability 24*(1), 1069-1085.

[23] Zhang, F. (2025). Integration of CFD and machine learning for application in water treatment process modeling: Membrane ozonation process evaluation. *Chemometrics and Intelligent Laboratory Systems 257*, 105302.

[24] Mobarak-Qamsari, E., Kasra-Kermanshahi, R., Nosrati, M., & Amani, T. (2012). Enzymatic Pre-Hydrolysis of high fat Content Dairy Wastewater as a Pretreatment for Anaerobic Digestion. *International Journal of Environmental Research 6*(2), 475–480.

[25] Narges, S., Ghorban, A., Hassan, K., & Mohammad, K. (2021). Prediction of the optimal dosage of coagulants in water treatment plants through developing models based on artificial neural network fuzzy inference system (ANFIS). *Journal of Environmental Health Science and Engineering 19*(2), 1543-1553.

[26] Yamamura, H., Putri, E. U., Kawakami, T., Suzuki, A., Ariesyady, H. D., & Ishii, T. (2020). Dosage optimization of polyaluminum chloride by the application of convolutional neural network to the floc images captured in jar tests. *Separation and Purification Technology 237*, 116467.

[27] Kim, C. M., & Parnichkun, M. (2017). Prediction of settled water turbidity and optimal coagulant dosage in drinking water treatment plant using a hybrid model of k-means clustering and adaptive neuro-fuzzy inference system. *Applied Water Science 7(7)*, 3885-3902.

[28] Gomes, L. S., Souza, F. A. A., Pontes, R. S. T., Fernandes Neto, T. R., & Araújo, R. A. M. (2015). Coagulant dosage determination in a water treatment plant using dynamic neural network models. *International Journal of Computational Intelligence and Applications 14*(03), 1550013.

[29] Lamrini, B., & Le Lann, M. V. (2014). A decision support tool for technical processes optimization in drinking water treatment. *Desalination and Water Treatment 52* (22-24), 4079-4088.

[30] Zangooei, H., Delnavaz, M., & Asadollahfardi, G. (2016). Prediction of coagulation and flocculation processes using ANN models and fuzzy regression. *Water Science and Technology 74*(6), 1296-1311.

[31] Kennedy, M. J., Gandomi, A. H., & Miller, C. M. (2015). Coagulation modeling using artificial neural networks to predict both turbidity and DOM-PARAFAC component removal. *Journal of Environmental Chemical Engineering 3*(4), 2829-2838.

[32] Dharman, S., Chandramouli, V., & Lingireddy, S. (2012). Predicting total organic carbon removal efficiency and coagulation dosage using artificial neural networks. *Environmental Engineering Science 29*(8), 743-750.

[33] Mälzer, H. J., & Strugholtz, S. (2008). Artificial neural networks for cost optimization of coagulation, sedimentation and filtration in drinking water treatment. *Water Science and Technology: Water Supply*, *8*(4), 383-388.

[34] Madaeni, S. S., & Kurdian, A. R. (2011). Fuzzy modeling and hybrid genetic algorithm optimization of virus removal from water using microfiltration membrane. *Chemical Engineering Research and Design 89*(4), 456-470.

[35] Altunkaynak, A., & Chellam, S. (2010). Prediction of specific permeate flux during crossflow microfiltration of polydispersed colloidal suspensions by fuzzy logic models. *Desalination 253*(1-3), 188-194.

[36] Varank, G., Erkan, H., Yazıcı, S., Demir, A., & Engin, G. (2014). Electrocoagulation of tannery wastewater using monopolar electrodes: Process optimization by response surface methodology. *International Journal of Environmental Research 8*(1), 165–180.

[37] Delgrange, N., Cabassud, C., Cabassud, M., Durand-Bourlier, L., & Laine, J. M. (1998). Neural networks for prediction of ultrafiltration transmembrane pressure–application to drinking water production. *Journal of membrane science 150*(1), 111-123.

[38] Baxter, C. W., Zhang, Q., Stanley, S. J., Shariff, R., Tupas, R. R., & Stark, H. L. (2001). Drinking water quality and treatment: the use of artificial neural networks. *Canadian Journal of civil engineering 28*(S1), 26-35.

[39] Park, S., Baek, S. S., Pyo, J., Pachepsky, Y., Park, J., & Cho, K. H. (2019). Deep neural networks for modeling fouling growth and flux decline during NF/RO membrane filtration. *Journal of Membrane Science 587*, 117164.

[40] Yazdandoost, Z., & Zonoozi, M. H. (2025). Treatment of oil-water emulsion through coagulation-flocculation and sedimentation using waterworks sludge as a green agent. *Journal of Environmental Chemical Engineering 13*(3), 116548.

[41] Gonzales, R. R., Nakagawa, K., Hasegawa, S., Kumagai, K., Matsuoka, A., Okamoto, Y., & Matsuyama, H. (2025). Treatment of high-strength wastewater with coupled flocculation and membrane filtration prior to ammoniacal nitrogen enrichment by osmotically assisted reverse osmosis. *Separation and Purification Technology 352*, 128159.

[42] Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern recognition letters 31*(14), 2225-2236.

[43] Kulkarni, P., & Chellam, S. (2010). Disinfection by-product formation following chlorination of drinking water: artificial neural network models and changes in speciation with treatment. *Science of the Total Environment 408*(19), 4202-4210.

[44] Milot, J., Rodriguez, M. J., & Serodes, J. B. (2000). Modeling the susceptibility of drinking water utilities to form high concentrations of trihalomethanes. *Journal of environmental management 60*(2), 155-171.

[45] Singh, N.K., Yadav, M., Singh, V., Padhiyar, H., Kumar, V., Bhatia, S.K., & Show, P.L. (2023). Artificial intelligence and machine learning based monitoring and design of biological wastewater treatment systems. *Bioresource Technology 369*, 128486.

[46] Emaminejad, S.A., Morgan, V.L., Kumar, K., Kavathekar, A., Ragush, C., Shuai,W., Jia, Z., Huffaker, R., Wells, G., & Cusick, R.D. (2022). Statistical and microbial analysis of bio-electrochemical sensors used for carbon monitoring at water resource recovery facilities. *Environmental science: water research and technology 8*, 2052–2064.

[47] Yang, Y.; Kim, K.R., Kou, R., Li, Y., Fu, J., Zhao, L., & Liu, H. (2021). Prediction of effluent quality in a wastewater treatment plant by dynamic neural network modeling. *Process Safety & Environmental Protection 158*, 515–524.

[48] Mustafa, H.M., Mustapha, A., Hayder, G., & Salisu, A. (2021). Applications of IoT and artificial intelligence in water quality monitoring and prediction: a review. In proceedings of the 6th international conference on inventive computation technologies (ICICT), Coimbatore, India, 968–975.

[50] Kamali, M., Appels, L., Yu, X., Aminabhavi, T.M., & Dewil, R. (2021). Artificial intelligence as a sustainable tool in wastewater treatment using membrane bioreactors. *Chemical Engineering Journal 417*, 128070.

[51] Bagheri, M., Al-jaberi, K., Wunsch, D.C., & Burken, J.G. (2019). A deeper look at plant uptake of environmental contaminants using intelligent approaches. *Science of the Total Environment* 652, 561-569.

[52] Iorhemen, O.T., Hamza, R.A., & Tay, J.H. (2016). Membrane bioreactor (MBR) technology for wastewater treatment and reclamation: Membrane fouling. *Membranes 6*, 33.

[53] Bein, E., Zucker, I., Drewes, J.E., & Huebner, U. (2021). Ozone membrane contactors for water and wastewater treatment: A critical review on materials selection, mass transfer and process design. *Chemical Engineering Journal 413*, 127393.

[54] Woo, T.Y., Nam, K.J., Heo, S.K., Lim, J.Y., Kim, S.Y., & Yoo, C.Y. (2022). Predictive maintenance system for membrane replacement time detection using AI-based functional profile monitoring: Application to a full-scale MBR plant**.** *Journal of Membrane Science 649*, 120400

[55] Mehrizi, R. A., Mirbagheri, S. A., & Shams, A. (2024). Development of a generalized mathematical model for two-stage reverse osmosis desalination systems. *Computers and Chemical Engineering, 182*, 108562.