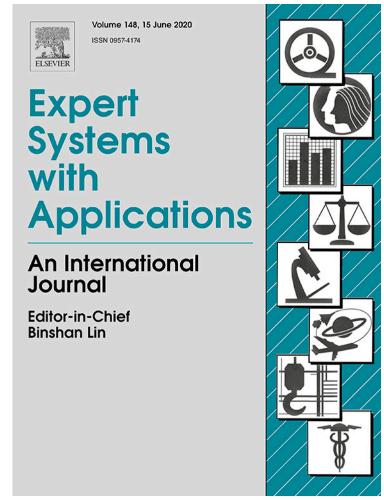
Journal Pre-proofs



Review

Data to Intelligence: The Role of Data-Driven Models in Wastewater Treatment

Majid Bahramian, Recep Kaan Dereli, Wanqing Zhao, Matteo Giberti, Eoin Casey

PII: S0957-4174(22)02472-1
DOI: https://doi.org/10.1016/j.eswa.2022.119453
Reference: ESWA 119453

To appear in: *Expert Systems with Applications*

Received Date: 4 September 2021
Revised Date: 22 October 2022
Accepted Date: 21 December 2022

Please cite this article as: Bahramian, M., Kaan Dereli, R., Zhao, W., Giberti, M., Casey, E., Data to Intelligence: The Role of Data-Driven Models in Wastewater Treatment, *Expert Systems with Applications* (2022), doi: https://doi.org/10.1016/j.eswa.2022.119453

Data to Intelligence: The Role of Data-Driven Models in Wastewater Treatment

*Majid Bahramian[1, *], Recep Kaan Dereli[1], Wanqing Zhao[2], Matteo Giberti[1], Eoin Casey[1]*

[1]*School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin, Ireland*

[2]*School of Computing Sciences, Smart Emerging Technologies, University of East Anglia, Norwich Research Park, Norwich, UK*

E-mail addresses:

Majid.bahramian@ucd.ie (Majid Bahramian)

Recep.dereli@ucd.ie (Recep Kaan Dereli)

W.Zhao2@uea.ac.uk (Wanqing Zhao)

matteo.giberti@ucdconnect.ie (Matteo Giberti)

eoin.casey@ucd.ie (Eoin Casey)

[*]*Corresponding author:* Majid Bahramian

E-mail addresses: majid.bahramian@ucd.ie, majidbahramian401@gmail.com

List of Abbreviations

AAD	Absolute Average Deviation
ABC	Artificial Bee Colony
ACO	Ant Colony Optimization
AD	Anaerobic Digestion
AI	Artificial Intelligence
AM	Attention Mechanism
ANFIS	Adaptive Neuro Fuzzy Inference system
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ASM2d	Activated Sludge Model No. 2 with denitrification
BOD	Biological Oxygen Demand
BSM1	Benchmark Simulation Model No.1
CI	Computational Intelligence
CNN	Convolutional Neural Network
COD	Chemical Oxygen Demand
CSTR	Complete Stirred Tank Reactor
DBN	Deep Belief Networks
DE	Differential Evolution
DNN	Deep Neural Network
DO	Dissolved Oxygen
DT	Decision Trees
EFAST	Extended Fourier Amplitude Sensitivity Test
EQI	Effluent Quality Index
FCM	Fuzzy C-means
FFA	Firefly Optimization Algorithm
FL	Fuzzy Logic
FO	Forward Osmosis
FV	Factorial Variance
GA	Genetic Algorithm
GHG	Greenhouse Gases
GP	Genetic Programming
GRU	Gated Recurrent Unit
GWP	Global Warming Potential
GWO	Grey Wolf Optimization
IA	Index of agreement
IMF	Intrinsic Mode Function
IQR	Interquartile Ranging
KPLS-RVM	Kernel Partial Least Squares with Relevance Vector Machine
LS-SVM	Least Square Support Vector Machine
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
MB	Methylene Blue
MBR	Membrane Bioreactor
MCR	Mass Content Ration
MG	Malachite Green
ML	Machine Learning
MLR	Multi-Layer Regression

MLVSS	Mixed Liquor Volatile Suspended Solids
MLANN	Multilayer Artificial Neural Network
MLP-PSO	Multilayer Perceptron hybridized with Particle Swarm Optimization
MOSSO	Multi Objective Shark Smell Optimization
MRE	Mass Removal Efficiency
MSE	Mean Square Error
M5T	M5 Tree
NMSE	Normalized Mean Square Error
NSE	Nash-Sutcliffe Error
OCI	Operational Cost Index
OCRNN	Over Complete Recurrent Neural Network
PBIAS	Percent bias
PCA	Principle Component Analysis
PMOC	Polar Mobile Organic Compounds
$P\text{-}PO_4$	Phosphate Phosphorus
PSO	Particle Swarm Optimization
QBPSA	Quantum-Behaved Particle Swarm Algorithms
R	Correlation coefficient
R^2	Coefficient of Determination
RBFNN	Radial Basis Function Neural Network
RF	Random Forest
RMSE	Root Mean Square Error
RO	Reverse Osmosis
RSM	Response Surface Methodology
R^2	Coefficient of Determination
SBR	Sequencing Batch Reactor
SC	Subtractive Clustering
SCADA	Supervisory Control and Data Acquisition
SI	Scattering Index
SOM	Self-Organizing Map
SVI	Sludge Volume Index
SVM	Support Vector Machine
SVR	Support Vector Regression
TDS	Total Dissolved Solids
THM	Trihalomethane
TLBO	Teaching Learning Based Optimization
TN	Total Nitrogen
TP	Total Phosphorus
TSK	Takagi–Sugeno–Kang
TSS	Total Suspended Solids
VFA	Volatile Fatty Acids
WWTP	Wastewater Treatment Plant

Abstract:

Increasing energy efficiency in wastewater treatment plants (WWTPs) is becoming more important. An emerging approach to addressing this issue is to exploit development in data science and modelling. Deployment of sensors to measure various parameters in WWTPs opens greater opportunities for exploiting the wealth of data. Artificial intelligence (AI) is emerging as a solution for automation and digitalization in the wastewater sector. This review aims to comprehensively investigate, summarize and analyze recent developments in AI methods applied to the modelling of WWTPs. The review shows that among the standalone models, Artificial Neural Networks (ANN) was the most popular model followed by, in descending order: Decision Trees (DT), Fuzzy Logic (FL), Genetic algorithm (GA) and Support Vector Machine (SVM). In the case of incomplete data, FL was the most frequently used method as it uses linguistic expert rules to find an approximation for the missing data. Regarding accuracy and precision, hybrid models demonstrated relatively better performance than the standalone ones. Among these models, the Machine Learning (ML)-metaheuristic, which integrates an AI model with a bioinspired optimization method, was the most preferred type as it was used in more than 45% of the hybrid models. Correlation coefficient (*R*), Correlation of Determination (R^2) and Root Mean Square Error (*RMSE*) were the frequently used metrics for model performance evaluation. Finally, the review shows that despite recent developments, industrial deployment is still lacking. The industrial application requires close interaction of interested parties, among which research institutes, private sector and public sector play an inevitable role. The future research should focus on mitigating the barriers for more in-depth collaboration of interested parties and finding new paths for more cooperative and harmonized activity of them.

Key words: *Artificial intelligence, Machine learning, Modeling, Optimization, Deep learning, Wastewater treatment.*

1. Introduction

Rapid urbanization, population growth, increasingly strict effluent quality requirements, and nearly-zero emission targets are several challenges facing the management of WWTPs. A total of 227 million cubic meters of municipal wastewater are generated annually which is expected to rise to 555 million cubic meters in 2030 (United Nations World Water Assessment Programme, 2017). Statistics has shown that only 42% of wastewater is properly managed and/or reused (United Nations World Water Assessment Programme, 2017). A variety of factors are responsible for the inappropriate management of wastewater, of which, insufficient plant capacity and inadequate operation are the most important (Zhang et al., 2020). One solution to improve WWTP performance is to implement advanced process control strategies, which are based on data measured from the unit processes. Frequently this data is underutilized for a variety of reasons (Bishoff et al., 2021; Fan et al., 2021; Newhart et al., 2019; Pang et al., 2019; United Nations World Water Assessment Programme, 2017). Either due to the size and complexity of the data generated by WWTPs or the lack of in-house capability around data science, it is challenging for the authorities in WWTPs to collect, store and manage useful data. Nevertheless, there has been a growing interest in the application of data science in modelling, operation and control of WWTPs, often straddling the fields of AI.

Due to the complexity of WWTPs, and the need for perpetual operation, large-scale, multi-source data sets is inevitably generated. To understand, analyze and extract practical information from the generated data, expert knowledge of process engineering and data science are required. As shown in Fig. 1, although it has seen a high number of the AI models in WWTPs, there have been few reviews, in particular critical AI investigations of applications in simulation and optimization of WWTPs (Abdallah et al., 2020; Bhagat et al., 2020; Corominas et al., 2018; Dürrenmatt and Gujer, 2012; Newhart et al., 2019; Yetilmezsoy et al., 2011; Zhao et al., 2020). In this regard, the purpose of this review is to critically evaluate and scrutinize the available literature on the application of data-driven techniques in WWTPs. Table 1 compares the present review with the available literatures in terms of the number of articles covered, application fields, period of study and scope of AI models. For instance, Bhagat et al. (2020) debated the pros and cons of various AI models in predicting the removal and concentration of heavy metals in wastewater. Abdallah et al. (2020)

discussed the application of AI in performance prediction of solid waste management. Zhao et al. (2020) provided a bibliometric analysis and systematic review of the application of AI in wastewater treatment. However, most of the reviews focused on the prediction models or task as a whole, and thus failed to investigate and categorize the underlying AI methods. In other words, none of them discussed the models in more details by scrutinizing, synthesizing, and classifying the forecasting methods employed. This is more noticeable in hybrid models, where the hybridization of different models with different concepts is categorized under the same title in majority of the reviews. The categorization is characterized by data incompleteness and uncertainty; temporal and spatial variability; domain features; and system boundaries, all of which mandates a new classification scheme.

The complexity of physical models used for simulation, optimization, and control of wastewater treatment, introduced the opportunity for parametric models for simulation and prediction of the modelled process. Additionally, the determination of a physical model of a process under study is very difficult, especially in the case of such complex processes as wastewater treatment, and the calibration of such a model requires the performance of quite laborious laboratory experiments to obtain appropriate data and requires time-consuming calculations related to the identification of model parameters. In this regard, the analyzed articles deal with parametric models. and their purpose is primarily to test the usefulness and compare in terms of accuracy of various machine learning methods, including in particular artificial intelligence methods.

In this regard, the present review aims to find the answer for the following questions? (1) What is the state-of-the-art in application of data-driven models in WWTPs? (2) what is the specific application of these models in process design, operation, and control? (3) what could be the perspectives on research direction? The present review not only provides a comprehensive and robust summary of the literature with detailed classification of the AI models, but also for WWTPs a critical assessment of the state-of-the-art in industrial application of data-driven prediction and classification techniques was presented. Moreover, a deeper look into the interaction between the research and industry was carried out to extract new highlights for future research on more practical deployment of data-driven models in wastewater sector. We also examined the details of metaheuristic algorithms used in the data-driven models. We aim to cover this subject by providing the purpose, application field, pertinent variables as well as the details of the frequently used metaheuristic algorithms.

[Table 1, here]

2. Methodology

2.1. Research questions

The main objective of this paper is to analyze, identify and present the state-of-the-art in data-driven models for WWTPs. The review will provide a direction for future research by identifying the gaps in the available literature. To do this, four specific research questions have been defined: (1) what are the various applications of data-driven models in WWTPs? (2) what are the data-driven models and algorithms used in wastewater applications? (3) how is the performance of various data-driven models compared to other methods? and (4) what are the strengths and limitations of data-driven techniques in field of wastewater treatment?

2.2. Search and selection criteria

This section describes how the available literature was searched and analyzed. The aim was to analyze a representative sample of publications quoting the use of forecasting models in WWTPs from 2000 up to 2021. Only peer-reviewed scientific publications and theses in Google Scholar and Scopus databases were considered. A preliminary inclusion and exclusion criteria were applied to limit and filter the number of the retrieved articles. The inclusion criteria employed a series of mandatory keywords and strings for the acceptance of the retrieved articles (Staples and Niazi, 2008). This incorporated ‘‘Wastewater Treatment Plant” or “Wastewater” with the following strings: ‘‘artificial intelligence”, ‘‘machine learning”, ‘‘K means”, ‘‘Fuzzy modelling”, “GA”, “Decision Trees”, “Support Vector Machine” and ‘‘artificial neural networks”. The exclusion strings were ‘‘sewer systems”, ‘‘waste and sludge management”, “medical and hospital wastes”, “dental wastewater”, “municipal solid wastes” and “construction waste and debris”.

Screening of articles using the search strings, based on the relevance of title, abstract and keywords, returned 303 publications. These publications were thoroughly inspected, in which 22 duplicate studies under different titles, but similar content was further excluded. The content of the remaining 281 studies were qualitatively assessed and extracted for data synthesis.

3. Overview of review results

3.1. Review statistics

Fig. 1 represents the distribution of AI models in the reviewed studies. The popularity of single (individual) models was higher than the hybrid models as single models were considered in 62% of the reviewed studies. Among single models, ANN was the most popular model, with feedforward neural network, radial basis function neural network, Kohonen self-organizing neural network, recurrent neural network, convolutional neural network and long / short term memory as the most popular ones. Moreover, the possibility of easily modifying the architecture to find the optimum architecture is one of the main advantages of neural networks (Hwangbo et al., 2021; Aish et al., 2015; Almomani, 2020; Ayodele et al., 2021). In case of hybrid models, ML-ML models were the most frequently used as they contributed to about 52% of the reviewed studies. Following ML-ML models, ML-metaheuristic, metaheuristic-metaheuristic, ML-ML-metaheuristic and ML-ML-ML models were detected in the hybrid models. Generally, hybridization is the combination of two distinct models, in parallel or series, to reach out a more precise model than the single models. The frequent pattern in the literature is the comparison of the modelling performance of the hybrid model with the underlying single models which was the case for 71% of the reviewed papers.

[Figure 1, here]

The annual trend of the AI models applied to WWTPs are represented in Fig. 2. The prevalence of ANN over other models is apparent throughout the reviewed studies. Following ANN, FL has been the most preferred model among AI models. It worth noting that the GA is recently receiving increased attention due to its efficiency in finding the approximate solution for optimization and search problems.

[Figure 2, here]

Fig. 3 shows the frequency of data-driven models with different applications within WWTPs. ANN was the most commonly used prediction model in studies focusing on the removal of conventional (138 case studies) and typical pollutants (164 case studies). It was also the preferred model in energy consumption prediction in treatment plants by contributing to more than 41% of the reviewed studies (117 case studies). Greenhouse gases emissions such as nitrous oxide, methane and carbon dioxide from WWTPs is one of the hot topics in laboratory and applied research in the field of environmental science. Fig. 3 also highlights the fact that the emission of greenhouse gases has not been frequently studied (7%) by using data-driven modeling techniques and the focus of the research was mainly on pollutant removal (52%, conventional and typical) and energy consumption (17%). The remaining portion of the reviewed studies focused on control strategies (12%) and economic aspects (12%) of wastewater treatment plants. The latter category mainly considered the costs associated with energy consumption (73%), mainly predicting the cost and energy required for the aeration of the biological reactor (88%) and a small number of case studies predicting the energy required for operation of pumps (12%). Some case studies considered the economic aspects of different processes of wastewater treatment, such as primary sedimentation (Struk-Sokołowska et al., 2020), electro-coagulation (Zakeri et al., 2021), reactive (Wang et al., 2021) and extractive distillation (Su et al., 2020) and activated sludge tank (Revollar et al., 2018). The ability in implicitly detecting the nonlinear relations, availability of multiple training algorithms, more fault tolerance and most importantly the ability to work with incomplete data and missing data, are the reported features of ANNs. The conventional pollutant metrics include biochemical oxygen demand (BOD), total suspended solids (TSS), fecal coliform, pH, oil and grease (EPA, 2018). Other pollutants including priority and toxic pollutants are categorized as typical (nonconventional) pollutants. The greenhouse gases emissions are reported in the form of $CO_{2\text{-eq}}$.

[Figure 3, here]

WWTPs use a variety of process to treat and reuse different types of wastewaters. Generally, three distinct processes of WWTPs can be recognized in the reviewed studies: Biological, Chemical and Physical processes. The classification of the reviewed studies based on the type of the process involved in the prediction, is given in Table 2. The breakdown of the data is as follows: biological processes (49%), chemical (38%) and physical processes (12%).

[Table 2, here]

ML and metaheuristic techniques have been used for better understanding of wastewater treatment processes. Both techniques can serve as practical tools for observation and analysis of data from the real time processes to provide operators with meaningful and predictive insights. A classification tree of different ML and metaheuristic techniques employed for modelling, optimization, and clustering of the wastewater treatment process is given in Fig. 4.

[Figure 4, here]

3.2. Single models

3.2.1. Artificial Neural Networks (ANN)

ANNs are effective in modelling processes with incomplete or uncertain data as well as capturing the nonlinearity. ANNs are designed to simulate the reactions of a biological nervous system when faced with real world tasks. Typically, neural networks consist of an input layer, hidden layers, and an output layer, each consisting of a number of nodes linked to every node in the subsequent layer by directed weighted edges (Fine, 2005). Different types of neural networks, have been extensively reviewed in the literature, among them feedforward neural networks trained by backpropagation algorithms are the most popular architecture (Golzar et al., 2020; Guo et al., 2015; Haghiri et al., 2018; Hamed et al., 2004; Han and Qiao, 2011; Hayder et al., 2014; Holubar et al., 2002; Kundu et al., 2013; Kusiak et al., 2013; Libotean et al., 2009; López et al., 2017; Lu et al., 2019; Ma et al., 2011; Mamandipoor et al., 2020; Mandal et al., 2015; Messikh et al., 2015; Mjalli et al., 2007). The most frequently tested back propagation algorithms were Levenberg-Marquardt (74%), Quasie-Newton (16%), one-step secant scaled conjugate gradient (5%), gradient descent with momentum and adaptive learning rate (4%), gradient descent with adaptive learning rate (1%), batch training with weight and bias learning rules (5). The most efficient training algorithm was selected based on the lowest Mean Square Error (MSE) and highest R^2 values.

ANNs have been widely employed to model various WWTP processes due to their robustness, easy calibration, and ability to recognize the complex relationships between variables in multivariate systems. However, one of the drawbacks for ANNs, is the high level of complexity of the network which was frequently reported on literature (Majhi and Panda, 2011; Otchere et al., 2021; Chiu et al., 2017; Fu et al., 2020). Moreover, ANNs are highly prone to overfitting. ANNs work with numerical data but translating information to numerical indicators isn't always straightforward and directly influences the model performance, which is highly dependent on the expert knowledge. In addition, it is often difficult to find the optimal structure and weights for ANNs. The learning of the network is performed over the training data, which could be much different than the testing and validation data. This would reduce the generalization ability of the model when applied to unseen data or operational conditions of the system.

Most ANN applications were conducted based on experimental data collected from full scale or lab scale treatment plants. The obtained data was often partitioned into three parts (training, validation, and testing) (77% of the reviewed studies) or two parts (training and testing) (23% of the reviewed studies). *Purelin, Sigmoid* and *Tansig* were the three frequently used transfer functions in the literature. While feedforward backpropagation was the most favorable network type, the three-layer network (input, hidden and output layer) was the most frequently mentioned typology in the reviewed studies. The performance of the model was evaluated by different metrics such as *R* (coefficient of correlation), *R^2 (Correlation of determination), RMSE (Root Mean Square Error), SSE (Sum of squared errors), Absolute Average Deviation (AAD%),* *Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Factorial Variance (FV), Index of Agreement (IA), Nash-Sutcliffe Error (NSE)* and *Percent Bias* (*PBIAS*). Among them *R*, R^2, *RMSE* and *MAPE* were the most used statistical indicators as they together account to more than 78% of the reviewed studies. Much detailed information regarding the application, performance, input and response variables, typology, architecture, and data partitioning of the ANN models can be found on the Table S1 (Supplementary Material).

3.2.2. Support vector machines (SVM)

SVM separates the hyperplane between data of two classes. The algorithm maximizes the margin between the optimum lines (support vector) generated from the data classification. SVM was employed in 22 modelling studies for different purposes. Fang et al. (2019) used a multi-kernel SVM based multi-feature fusion algorithm to predict the membrane permeation in a treatment plant.

The limitations of using single kernel function were avoided by using a multi-kernel function. In addition, the use of multi-kernel function improved the prediction performance of SVM. Multi-kernel functions have the ability to select an optimal kernel and parameters from a larger set of kernels, reducing bias due to kernel selection while allowing for a more automated machine learning procedure. A composite model integrating latent variables of kernel partial least squares with relevance vector machine (KPLS-RVM) was used by Liu et al. (2020) for effluent COD prediction in a full scale treatment plant and the results were compared with that of a least square support vector machine (LS-SVM). To cope with the high dimensionality and complex collinearity of nonlinear process, the dominant variables were selected by using kernel partial least square projection technique. Then relevance vector machine was used to construct a predictive function between the dominant variables and output variables. The *R* of the model was approximately 19.65% higher than the respective value for LS-SVM.

Foroughi et al. (2020) compared the efficiency between Multi-layer Regression (MLR) and LS-SVM in modeling Tetracycline removal from wastewater by a three-dimensional electrochemical system. The LS-

SVM model showed relatively smaller *MSE, RMSE, AAD* and Mean Absolute Error (MAE) values. SVM showed stronger correlation than MLR in modelling the second-order reaction rate constant of ozone for more than 130 organic compounds (Huang et al., 2020). SVM was used by Kazemi et al. (2021) to develop a soft sensor for Volatile Fatty Acid (VFA) control in anaerobic digestion. The results demonstrate the superiority of Support Vector Regression (SVR) over Adaptive Neuro Fuzzy Inference System (ANFIS) in providing more accurate residual errors, and correctly detecting VFA faults (the difference between the measurements and the respective statistical control limit). (Zaghloul et al., 2020). Liu et al. (2019) used LS-SVM for online prediction of effluent COD in an anaerobic WWTP. Mahata et al. (2020) reported the superiority of SVM over ANN in modelling the hydrogen production from dark fermentation with R^2 and *RMSE* values of 0.988 and 0.0103, respectively. Different soft computing techniques were used by Najafzadeh and Zeinolabedini (2019) to predict daily influent flowrate of a domestic WWTP. Performance analysis indicated that SVM (RMSE = 1435.4 and MAE = 1031.1) and FFBP-NN (RMSE = 1445.9 and MAE = 1036.7) techniques provided more accurate prediction of flow rates compared to ANFIS (RMSE = 1515.6 and MAE = 1075.4) and ANN (RMSE = 1501 and MAE = 1048.7). SVM was superior over Random Forest (RF) in predicting the N_2O emissions in a full scale WWTP (Vasilaki et al., 2020). Among the reviewed SVM studies, 72% used SVM for prediction and 28% were used for optimization purposes, respectively. The detection and distinction capabilities of SVM made it a useful optimization tool. Unlike ANN, SVM is not solved for the local optima and requires manual interpretations. A closer look at the datasets where SVM was used either for prediction or optimization reveals that SVM is relatively better in datasets where the number of dimensions (parameters) are bigger than the number of the samples. Despite its unique ability in feature detection, SVM is not effective in datasets with high range of noise and biases. This limits its application only to optimization problems and other data-driven models are more effective than SVM in prediction purposes.

3.2.3. Genetic Algorithm

Inspired by Darwin's theory of natural evolution, GA searches for the most suitable individuals to carry out the reproduction. The idea can be applied to find the optimum solution among a set of available solutions.

GA was used in 28 studies. It was widely used for process optimization (70% of the reviewed studies). These processes include heat-activated oxidation (Rahmani et al., 2020), reactive and extractive distillation (Su et al., 2020), pervaporation (Suggala and Bhattacharya, 2003), hybrid distillation (Tashvigh and Nasernejad, 2017; C. Wang et al., 2021), flocculation (Dawood and Li, 2013), adsorption (Onu et al., 2021; Picos-benítez et al., 2020), RO (Al-Obaidi et al., 2017; Al-Obaidi et al., 2018), membrane fouling (Lee et al., 2009), wastewater leakage (Attwa and Zamzam, 2020), Benchmark Simulation Model No.1 (BSM1) and Activated Sludge Models (ASMs) (Zhu et al., 2015), and microbial fuel cells (Sedaqatvand et al., 2013). The remaining portion (30%) used GA for prediction purposes. Al-Obaidi et al. (2017) used GA for to find the best operating parameters for the optimum rejection of chlorophenol in a single spiral RO membrane module. Operating feed flowrate, operating pressure and temperature were considered as decision variables. The specifications of 1e-4 to1e-3 m^3/s, 4 to 24.77 atm and 15 to 40 °C were considered as the minimum and maximum constraints for decision variables of operating feed flowrate, applied pressure and temperature, respectively. Moreover, as suggested by the membrane manufacturer, an allowable pressure drop of 1.38 atm throughout the membrane length was also inserted to the constraints. The maximum generation of 500, population size of 50, crossover probability of 0.6 and mutation probability of 0.1, were considered as the GA parameters for finding the optimal solutions. Niu et al., (2020b) used GA to optimize the performance of Deep Believe Network (DBN) in predicting the effluent quality of a full-scale papermaking WWTP. As shown on Eq. 1 and 2, this study used the reciprocal of MSE of effluent COD and suspended solids as the matching function.

$$f(x)_1 = {}^{1}\!/_{mse\,(\hat{A}-A)} = {}^{1}\!/_{\sum_{i=1}^{n}(\hat{a}_i - a_i)^2} \tag{1}$$

$$f(x)_1 = {}^{1}\!/_{mse\,(\hat{B}-B)} = {}^{1}\!/_{\sum_{i=1}^{n}(\hat{b}_i - b_i)^2} \tag{2}$$

In these equations, n is the number of data points, $\hat{A} = \{\widehat{a_1}, \widehat{a_2}, ..., \widehat{a_n}\}$ is the predicted values for the concentration of COD in the effluent of the treatment plant and $A = \{a_1, a_2, ...,a_n\}$ is the measured values for the COD concentration in the effluent. Similarly, $\hat{B} = \{\widehat{b_1}, \widehat{b_2}, ..., \widehat{b_n}\}$ is the predicted values for the concentration of suspended solids in the effluent of the treatment plant and $A = \{b_1, b_2, ...,b_n\}$ is the measured values of the suspended solids in the effluent of the treatment plant. The six variables of influent flowrate, pH, temperature, dissolved oxygen (DO), influent COD, and influent suspended solids, with respective ranges of 1681-2983.2 1e+4 m^3/d, 2.21-3.82, 6.4-7.88 ℃, 329-946 mg/l, 17.8-32.7 mg/l, 1.88-4.12 mg/l, were considered as decision variables. The maximum generation of 100, population size of 20, random crossover and mutation probability were other specifications of the GA for searching and locating the optimal solutions. GA is acknowledged for its superior optimization performance, specifically its ability to globally optimize several parameters (Rahmani et al., 2020; Sellami et al., 2020; Su et al., 2020; Suggala and Bhattacharya, 2003). GA uses the evolutionary characteristics of the population to search for meaningful information and determine the search direction based on natural selection. Wang et al. (2020) mentioned its adaptability in using natural evolution, mutation, and crossover to find the optimum solution.

Despite its efficiency in optimization problems, GA has also several limitations. The most notable one is that GA is a computationally expensive and a large number of generations need to be performed before reaching an optimum solution. Moreover, there are many considerations and parameters involved in the implementation of the algorithm. For example, the formulation of an objective function, encoding of decision variables, selection of population size, mutation and crossover rates, and the generation of the new population should be carried out carefully.

Table 3 gives an overview of the application of GA in the reviewed studies. GA was widely used for process optimization in the reviewed studies (50%). The optimized processes and parameters include adsorption, photocatalytic and reverse osmosis processes, design and optimization of bioreactor, microbial fuel cells, and oxidation ditches, treatment of micropollutants and pharmaceutical pollutants, balancing demand and response strategy, minimizing capital and operational costs, optimizing aeration and pumps energy consumption, influent and effluent quality and quantity, and minimizing wastewater leakage within the treatment plant. Following parameter optimization, GA was also applied to improve the prediction accuracy in different hybrid models by contributing to more than 30% of the reviewed models. In this sense, the forecasted parameters include Hydrogen (H_2) production, pH, Sludge Volume index (SVI), coagulant dosage, effluent COD – Nitrate and Phosphate, effluent quality index (EQI) and kinetics of SBR and Complete Mixed Tank Reactor (CSTR). Moreover, GA was also applied to optimize the structure of the main prediction model (10.71% of the reviewed studies). These models include deep belief networks (DBN), Multilayer Artificial Neural Networks (MLANN) and Principal component analysis (PCA). Finally, in two out of 28 reviewed papers, GP, a further development of GA, was applied for prediction of multiple variables, i.e., the membrane fouling rate and permeate flux decline in microfiltration process (Shokrkar et al., 2012; Tashvigh and Nasernejad, 2017). Shokrkar et al., (2012) applied GP to find a mathematical function of independent variables for operating the microfiltration of oily wastewater (temperature, transmembrane pressure, oil concentration, crossflow velocity). The ultimate goal of the genetic programming was to find a mathematical function to model the membrane permeate flux as a function of these independent variables. By assigning the matching function as the bias between the evolved output and the target output, the lowest matching function value was considered the preferred one. The following parameters were considered during development stage of the GP model: terminal set: X1 = filtration time (min), X2 = oil concentration (ppm), X3 = cross-flow velocity (m/s), X4 = temprature (◦C), X5 = trans-membrane pressure (bar); function set: {+,-,×, exp, sqrt}; maximum number of iterations: 90; population size: 1000; genetic operators: crossover and mutation; mutation probability: 2%; initial random tree depth: 4; maximum random tree depth: 7; Elitism parameter: Keepbest; Initialization method: Ramped half–half; selection method: Lexictour. The study showed that the results gained from GP model had an average error of less than 5% with the experimental values.

[Table 3, here]

3.2.4. Decision Trees

Decision trees split the variables into different nodes and select the nodes which show the most homogeneous sub-nodes.

One of the main advantages of decision trees is that the consideration of all possible outcomes of a decision is enforced. The consequences of each decision branches (leaves) are provided in a comprehensive analysis. Moreover, the algorithm of the technique is easy to understand, and it has the possibility to smoothly be integrated to other prediction and optimization tools. However, along with its simple structure, decision trees suffer from low accuracy and stability issues. 12% of the reviewed papers focused on the decision trees. Among them, 82% was used for regression and 18% was used for classification and feature detection, respectively.

Zhou et al. (2019) employed Random Forest (RF) for influent flow rate forecasting in WWTP. By means of trees, Rahimian and Behnam (2020) classified the number of active pumps in a full scale treatment plant. Similarly, Byliński et al. (2019) predicted the odor intensity and its properties in a biological WWTP with DTs. Soares et al. (2020) used RF to predict basic dye biosorption process with orange waste. Similarly, Teychene et al. (2020) used RF to analyze the rejection mechanism of polar mobile organic compounds (PMOC) by RO and nanofiltration techniques. The RF model emphasized that the NF rejection mechanism is governed by electrostatic interaction and sieving effects. In contrast, PMOC rejection of the RO membrane strongly depends on the topological polar surface area of the PMOC.

Song et al. (2020) applied RF technique to three biological nitrogen removal processes to identify the N_2O production mechanism in activated sludge tanks in different times of the year. They reported that the biomass-normalized dissolved inorganic carbon concentration and specific ammonia oxidation activity were the most influential parameters affecting N_2O emissions from the aerated zones of activated sludge bioreactors. For the anoxic tanks, dissolved-organic-carbon-to-NO_2^-/NO_3^- ratio was found to be the most important factor.

Deepnarain et al. (2019) employed a novel DT for modelling the sludge bulking in a full-scale WWTP operated as a 3-stage Phoredox process. While a RF model was employed to find the correlation between influent wastewater characteristics and SVI, a classification tree model was used to determine the dominant environmental factors affecting the proliferation of filamentous microorganisms.

(Jiang et al., 2019) reported that RF outperformed Linear Regression and Support Vector Regression (SVR) in modelling the performance of an activated carbon system enhanced by an ultrasonic assisted impregnation unit for hydrothermal carbonization and pyrolysis of wheat, corn and sorghum straws. Jung and Kim (2016) used DT algorithm to determine the parameters affecting the energy consumption in a WWTP.

Çelik et al. (2013) used a decision tree based on Gini algorithm to estimate pH, BOD, COD, and TSS of a full-scale domestic WWTP. The accuracy of the model in predicting pH, BOD, COD and TSS was 64.62%, 81.25%, 57.81% and 65.62%, respectively. Granata et al. (2017) compared the performance of SVR and Regression Trees (RT) in predicting the effluent BOD_5, COD, TSS and Total Dissolved Solids (TDS) in a full-scale treatment plant. With regards to R^2 and *RMSE*, the results showed the superior performance of SVM in predicting effluent TSS, TDS, and COD, nevertheless, both models showed similar performance in predicting BOD. Zhu et al. (2019) applied ANN and RF to model the adsorption of heavy metals to biochar. The results indicated that the RF model showed better accuracy and predictive performance for adsorption efficiency (R^2=0.973) than ANN model (R^2=0.948). similarly, RF outperformed gradient boosting trees and artificial neural network for tetracycline and sulfamethoxazole adsorption on carbon-based materials (Zhu et al., 2021b).

Hocaoglu (2017) used DT to assess the wastewater reuse alternatives of Mediterranean hotels. To acquire the highest water saving, mixed domestic wastewater reuse may be the best suitable alternative for hotels which have more than 250 rooms, large irrigated landscaped area of bigger than 100 m^2 per room and operated seasonally. However, for hotels which are operated throughout the whole year, having limited irrigated landscaped area and under design or construction, grey water reuse may be the best suitable alternative for the highest water saving.

3.2.5. Fuzzy logic

Fuzzy logic theory provides qualitative reasoning into the quantitavie estimation, thus enabling generalized rules as an option for decision making.

The robustness of the FL is one of the advantages of this model. This means that the model is not susceptible to the changes within the system. Moreover, the reasoning process is often simple. FL was used in 30 case studies in the reviewed papers. Mamdani and Sugeno were the most frequently used fuzzy architectures. A large number of studies used incomplete datasets when using FL throughout the modelling (Chang et al., 2012; Chung and Kim, 2014; Fiter et al., 2005; Fonseca et al., 2018; Gupta et al., 2017; Karimi et al., 2011; Mahjouri et al., 2017a; Marsili-Libelli and Giunti, 2002; Mazhar et al., 2019; Murnleitner et al., 2002; Nadiri et al., 2018; Patel et al., 2020; Revollar et al., 2018; Santín et al., 2018, 2015b, 2015a; Tejaswini et al., 2020b).

Due to ability to convert the linguistic information to mathematical rules, FL was widely used for modelling various processes within the WWTPs.(Bae et al., 2006; Belchior et al., 2012; Bello et al., 2014; Bertanza et al., 2020; Boiocchi et al., 2016; Borges et al., 2016; Carrasco et al., 2002; Chang et al., 2012; Fiter et al., 2005; Fonseca et al., 2018; Marsili-Libelli and Giunti, 2002; Murnleitner et al., 2002; Patel et al., 2020; Revollar et al., 2018; Santín et al., 2018, 2015a). This broad spectrum was not limited to the application area of FL, but also for the input variables of the model. From DO lag-time (Bae et al., 2006), ammonia concentration (Bertanza et al., 2020), $R_{NatAMM} = \frac{|NO^-_{3_{in}} - NO^-_{3_{out}}|}{|NH^+_{4_{in}} - NH^+_{4_{out}}|}$ (Boiocchi et al., 2016), turbidity (Fonseca et al., 2018; Borges et al., 2016), and rainfall data (Borges et al., 2016), to pH and the filling level of tanks/reactors (Murnleitner et al., 2002) were considered as input variables for FL modelling. Though the output of fuzzy modelling had a broad variation, it can be categorised into three distinct groups, i.e. quality, cost and carbon emissions.. For instance, Revollar et al. (2018) used DO concentration as an indicator for energy consumption, nitrate concentration at the end of the anoxic zone as an emission indicator and ammonium concentration as the effluent quality indicator. Santín et al. (2018) used N_2O as the sole emission from the plant, ignoring other important greenhouse gases (GHGs), in particular methane and CO_2 which are abundantly present in treatment plants (Metcaff and Eddy, 2002). Tejaswini et al. (2020b) used Operational Cost Index (OCI) to evaluate the energy perfromance, and EQI to assess the water quality in the effluent of the plant.

In general, 40% of the reviewed studies used FL for predicting the energy consumption, 27% for effluent quality forecasting and the remaining portion (33%) for selection of the most proper treatment technology. For instance, Karimi et al. (2011) and Mahjouri et al. (2017a) developed a decision making criteria based on FL for selection of the appropriate treatment technology for iron and steel industry. Gupta et al., (2017) employed fuzzy intelligence for evaluation of growth parameters and metabolites extraction of microalgae. Chung and Kim, (2014) worked on a fuzzy multi-criteria logic to prioritize locations of treated wastewater use considering climate change scenarios. Biogas production, overload and coagulant dosing control were among other applications of the FL modelling within WWTPs.

In general, there is strong evidence for the superiority of FL in modelling processes with incomplete data. Moreover, fuzzy-logic controllers provide the possibility of incorporating expert knowledge into the design for the processes to be controlled. Moreover, FL can deal with distorted or imprecise data, and it allows to model uncertain knowledge and information in this regard.

3.3. Hybrid models

Hybrid models use different combinations of data-driven models. The purpose is to fuse the advantages of various modeling strategies by a skillful combination. We detected a large variation of hybrid models with different combinations and application fields. The hybrid methods in the reviewed studies can be divided into following five categories:

3.3.1. ML-ML models

This class combines two Machine Learning (ML) models with each other, and it accounts to almost 38% of the reviewed hybrid models. Standalone models such as ANN, SVM, GA, FL and DT are combined in various forms to increase the performance and efficiency of the prediction. For instance, to forecast the wastewater influent/effluent BOD, COD and TSS, Lotfi et al. (2020) combined autoregressive integrated moving average (ARIMA) model with adaptive neuro fuzzy inference system (ANFIS). Prior to the modelling task, moving average smoothing was used to remove the noise in the time series. The smoothing determined that the data with values three times lower or higher than the standard deviation should be considered outlier and must be replaced with the mean of reliable data before and after it. This pre-processing technique significantly increases the reliability of the modelling. After combining ARIMA with ANFIS, the coefficient R^2 increased by 18%, the efficiency of the NSE increased by 20%, the MAE, MAPE, and RMSE errors decreased by 78%, 61%, and 65%, and the SI decreased to 88%. Asadi et al. (2020) used a multi-layered perceptron network (MLP) and an (ANFIS) with grid partition (GP), subtractive clustering (SC) and fuzzy C-means (FCM) clustering to predict the biogas production (including methane, carbon dioxide and hydrogen sulfide) from the anaerobic digester of a full-scale WWTP. Clustering techniques were used to determine the independent variables in the input data, making it possible to reduce the number of model input variables. This significantly reduced the model's computational time while showing similar or higher performance. While the values of RMSE, R and IA for models with clustering were 0.58, 0.92 and 0.94, the respective values for models without clustering were 1.2, 0.66 and 0.28. Similarly, Sulthana et al. (2014) used a fuzzy principal component analysis to model the reduction in COD and BOD of treated wastewater. PCA was employed to maximize the variance in the dataset matrix and reduce dimensionality. The partitioning of the PCA scores was carried out with a Takagi–Sugeno–Kang (TSK) fuzzy model based on a Fuzzy C-means (FCM) clustering. The FCM clustering algorithm was applied to the PCA score matrix, wherein each score is assigned to a centroid of cluster with different membership values. The proposed and hybrid fuzzy principal component regression model had the capacity of describing the PCA score clusters fuzzy systems, as well as the non-linear modeling intelligence to capture the relation between fuzzy partitioned PCA score clusters and target output, in this case reduction in BOD and COD.

A neuro-fuzzy control system was employed by Bernardelli et al. (2020) to predict the main process variables and providing the right amount of aeration to achieve an efficient and economical operation within a municipal WWTP. A novel combination of convolutional neural network (CNN) and SVM was used by Fan and Zhang (2020) to predict the 5-day BOD of a full-scale activated sludge plant. The wavelet denoising ANFIS showed satisfactory performance in predicting wastewater effluent discharge quality (Fu et al., 2020). Huang et al. (2010) used a neural fuzzy model for on-line estimation of nutrient dynamics in an anoxic/oxic process. Mahshidnia and Jafarian (2016) used an adaptive neural fuzzy intelligent system to forecast the removal of Malachite Green (MG) from domestic wastewater. A neuro fuzzy network was also used by Fernandez et al. (2009) to improve wastewater flow-rate forecasting.

Mahjouri et al., (2017) used a novel combination of fuzzy Delphi and fuzzy analytic hierarchy process to identify the key indicators in selecting the optimal wastewater treatment technology. A hybridization of unsupervised self-organizing map (SOM) with MLP has been reported by Rustum and Adeloye (2012) to be more efficient and accurate than the individual models in predicting the effluent BOD concentration of primary clarifiers.

3.3.2. ML-ML-ML models

Another classification uses the sequential combination of three singular models to implement the prediction. Heo et al. (2020) used a novel sequential combination of convolutional neural network (CNN), Gated Recurrent Unit (GRU), and deep neural network (DNN) in a specialized multimodal ensemble learning-based algorithm generated by different intrinsic mode functions (IMFs) to forecast the wastewater influent loading rate. The architecture of the proposed model is given on Fig. 5.

[Figure 5, here]

The model was tested by predicting the COD, SS, TN and TP loads on long-term (daily), and short-term (hourly) with multi-steps forecast horizons. The proposed hybrid model outperformed other five Recurrent Neural Network-based reference models in capturing the spatial and temporal patterns of fluctuating influent loads. Compared to other ANN models, the proposed ensemble model showed better performance in a slightly shorter computation time. With regards to MAE, MAPE (%) and mean absolute scaled error (MASE), the model increased the accuracy by 30%, 25% and 42%, respectively.

3.3.3. ML-Metaheuristic models

One of the widely used type of hybrid models which accounted to more than 51% of the reviewed hybrid models, was the combination of a single ML model with a Metaheuristic algorithm. Generally integrated to the output of the ML model, metaheuristic algorithm was following three different purposes: increasing the prediction performance; optimization of the structure of the ML models; and finding the optimum values of different parameters (parameter optimization). The details of each of these models with their application are given in Tables 6 and 7. Here, we present some of the applications within the WWTPs.

To predict the biogas production (specifically methane) rate Akbaş et al. (2015) employed a hybrid MLP and PSO techniques. Similarly, Pan and Cao (2020) employed Artificial bee colony (ABC) to improve the prediction performance of the ANN modelling the effluent BOD of a biological WWTP. To precisely predict the permeate flux decline during microfiltration of oily wastewater, Badrnezhad and Mirza (2014) used a combination of ANN and GA. Balasubramani et al. (2020) used Particle Swarm Optimization (PSO) to increase the prediction efficiency of ANN in forecasting the micropollutants removal from domestic wastewater. Ehteram et al. (2020) used Multilayer Perception hybridized with Particle Swarm Optimization

(MLP-PSO) to model the TDS and permeate flow rate of the RO process. The results were compared against support vector machine (SVM) and M5T models. MLP-PSO outperformed other models in TDS and permeate flux prediction of RO process. Zhang et al. (2019) used a hybrid model by combining ANN with GA to forecast the overall performance of a biological WWTP.

Baki and Aras (2018) compared the performance of two optimization models namely teaching learning-based algorithm (TLBO) and ABC algorithm in training the structure of an ANN model for predicting the influent BOD of a WWTP. Boztoprak et al. (2016) integrated GA and ABC into a Cellular Neural Network for predicting the SVI at a full-scale activated sludge plant.

Dehghani et al. (2019) combined Grey Wolf Optimization (GWO) with ANFIS to recursively predict the short- and long-term influent flowrate of a biological sewage treatment plant. Farzin et al. (2020) compared the performance of different stand-alone AI models with a novel hybrid combination of LS-SVM and firefly optimization algorithm (FFA) for electrochemical removal of drugs from domestic wastewater. A combination of differential evolution (DE) algorithm and ANN was used for pH, TSS, COD and BOD modelling of aerated lagoons in a full scale WWTP (Godini et al., 2020). Huang et al. (2011) integrated GA to a neuro fuzzy system to study the coagulation process of a WWTP in a paper mill. A novel fuzzy wavelet neural network based on the GA was used to estimate the effluent quality and biogas production rates in a full-scale anaerobic wastewater treatment process (Huang et al., 2019). PSO was employed by Huang et al. (2009) to optimize the parameters of LS-SVR in predicting the effluent quality of a WWTP. Chemical oxygen demand, biochemical oxygen demand, total nitrogen, ammonium nitrogen and TSS, were generated by BSM1. The parameters of LS-SVR were optimized by PSO in an attempt to obtain a more accurate model. The results indicated that the proposed model is capable of being adapted to different weather conditions. A Parallel Particle Swarm Optimization-Long Short Term Memory Model was integrated to SVR to predict the BOD, TN, TP and COD of the influent to a WWTP (Yan et al, 2019). Nassef et al. (2019) integrated PSO into an ANFIS model to enhance the modelling of lipid generation from microalgae during the microwave pretreatment.

Mohammadi et al. (2020) compared the performance of Levenberg Marquardt and particle swarm algorithms in training the MLP developed for modelling the alkylphenols removal in a moving bed biofilm reactor. The results showed that when the number of neurons in the hidden layer was increased from 8 to 10, the network trained with PSO algorithm showed slightly better performance (R=0.9997, MSE=2.526e-5, MAE=0.0041) than the traditional Levenberg-Marquardt algorithm (R=0.9989, MSE=2.582e-5, MAE=0.0043). The number. The number of optimal neurons was 5:9:3 for Levenberg Marquardt and 5:11:3 for PSO. Moreover, the *Tan-sigmoid* in the hidden layer and *Purelin* in the output layer were selected as the best transfer functions for the multilayer perception ANN.

Qi et al. (2020) compared the performance of Response Surface methodology (RSM), ANN-particle swarm and ANN-GA in predicting the optimum adsorption conditions of methylene blue (MB) from simulated wastewater by the mesoporous rGO/Fe/Co nanohybrids. Among the three models studied, the ANN-PSO model has the highest R^2 value and the lowest *MSE* value.

Seifi et al. (2020) studied the performance of three data-driven models including MLP, RBFNN and SVM hybridized by multi-objective shark smell optimization algorithm in predicting different prediction horizons of immediate, short-term, and long-term of influent time series of a wastewater treatment and reuse plant. The results indicated that the multilayer perception neural network model with multi objective shark smell optimization (MOSSO) algorithm produced better results than the other hybrid and standalone models in all prediction horizons (immediate, short term and long term).

Caraman et al. (2020) used Fuzzy-GA technique to develop a control strategy for determination of optimal setpoint of DO and nitrate. Tracking the optimum setpoint of dissolved oxygen in the fifth aeration zone and nitrate in the second anoxic zone of a biological nitrogen removal process was carried out by Han et al. (2018). They employed adaptive kernel function models enhanced by an improved multi-objective particle swarm optimization algorithm to predict the dissolved oxygen and nitrate concentrations inside the biological nitrogen removal process. A neuro-genetic approach was used for soft sensor control of a WWTP (Fernandez de Canete et al., 2021). Similarly, Chang and Li (2021) applied the binary particle swarm optimization to improve the efficiency of an Over-Complete Deep Recurrent Neural Network (ODRNN) to capture the nonlinear, non-Gaussian characteristics of the sensor data inside a WWTP. The simulation results with BSM1 illustrated the efficiency of the ODRNN over Deep Recurrent Neural Networks. Yue (2020) used Wale Optimization Algorithm to enhance the prediction capabilities of Kriging surrogate model in modelling the ozonation process of a pharmaceutical WWTP.

3.3.4. Metaheuristic-Metaheuristic models

Piotrowski et al. (2019) employed a hybrid combination of Artificial Bee Colony and direct search algorithm (DSA) for optimizing biological processes and operational cost of a sequencing batch rector. A two-layer optimization phase was integrated to the Activated Sludge Model No. 2 with denitrification (ASM2d) model of biological processes and the aeration system. A variety of different deterministic and non-deterministic optimization algorithms were applied and the sequence of ABC algorithm and the Direct Search Algorithm (DSA) was found to be the most optimum in improving the overall capacity, increasing effluent quality, and reducing the cost. Chen et al., (2016) used a sequence of artificial bee colony and quantum-behaved particle swarm algorithms (QBPSA) to develop a water quality monitoring model for effluent quality control in a WWTP. Two different hybridizations were considered. One hybrid strategy was to use sequential combination, and the other was to use parallel adaptive cooperative evolving of ABC and QBPSA. The performance comparison of the artificial bee colony, quantum-behaved particle swarm, their sequential combinations, and parallel adaptive dual populations revealed that the parallel dual population method is more efficient than the original algorithms.

3.3.5. ML-ML-metaheuristic models

This classification is similar to CI-CI classification with an optimization method integrated to the sequence of two CI models. This combination accounted to 8.5% of the reviewed studies. Zhang et al., (2021) used a combination of adaptive multi-objective particle swarm optimization and fuzzy neural network for tracking the obtained optimal DO set points for reducing energy consumption in a full scale treatment plant.

A RSM followed by ANN hybridized with GA was employed by Mohan et al. (2015) to predict the removal of Cr(V) in an adsorption process using cupric oxide nanoparticles. A novel sequential fusion CNN coupled by long short-term memory (LSTM) and attention mechanism (AM) was developed by Li et al. (2021) to monitor the water quality in a full-scale wastewater treatment system treating paper industry effluent.

The objective of the models, the purpose of PSO in that model and the publishing year is shown in Table S2 (Supplementary Material). Among the reviewed models, main portion of the models were using PSO to improve the overall prediction efficiency (55% of the reviewed models). 35.5% of the models applied PSO for parameter optimization and 9.5% of the models used PSO for improving the structure of the main prediction model.

Ant Colony Optimization (ACO) is one of the nature-inspired techniques for solving optimization problems. ACO algorithm works by finding optimal paths according to the random behavior of ants in searching for food. Upon finding a source of food (solution), it leaves a "marker" (pheromones) that show the path has food. ACO captures global minima in much lower number of iterations and with a lower value when compared with PSO. Moreover, there is lesser number of transients in the graph for square of error for ACO. The summary of the ACO application in the WWTPs is given on Table 4. In WWTPs, it is either used to improve the prediction performance (60%) or tuning a specific parameter to its optimum values (40%).

[Table 4, here]

4. Discussion
4.1. Accuracy

The modelling of WWTP is essential for process management and cost evaluation. In the reviewed models, the effectiveness of modelling was evaluated by different statistical indicators. R, R^2, MSE, $RMSE$, $MAPE$ and MAE were the most frequently used (173 models out of 281, in aggregate). Table 5 summarized the metrics used in the literature to evaluate the performance of the prediction models. In general, hybrid models showed slightly better performance than the standalone ones. The reported range of the acceptable R on the literature was between 0.62 and 0.99. The reported values of R and R^2 for all the standalone AI models were in the same range except for the ANN where a range between 0.6289 and 0.998 for R, and a range between 0.616 and 0997 for R^2 was detected. *RMSE* as one of the widely used model indicators showed a wide variability in the reviewed studies. The reported values for *RMSE* ranged between 0.0004 and 3486 indicating the uncertainties in the reported values for this indicator. In case of *MAPE*, the reported values ranged between 0.6% and 36.28%. The indicators used for accuracy in the models were different which is making it difficult to compare them effectively. Predictive accuracy should be based on the highest measure of difference between the observed values and predicted values, not on the average residuals. This could be neglected by lack of necessary filtration actions on the raw data and use of improper training techniques which influence the accuracy in different scales.

In general, this study highlighted that the hybrid methods showed better performance compared to standalone methods (Table 5). Similarly, numerous authors underpinned the efficiency of hybrid models (Asadi et al., 2020; Bernardelli et al., 2020; Dehghani et al., 2019; Farzin et al., 2020; Fernandez de Canete et al., 2021; Honggui Han et al., 2018a; Heo et al., 2021; Huang et al., 2019; Najah Ahmed et al., 2019; Park et al., 2020; Qi et al., 2020a; Seifi et al., 2020; Shamshirband et al., 2019; Shi and Xu, 2018; Zhang et al., 2019). In the case of stand-alone models, ANN seems better than the other models in terms of accuracy and wider adoption by the scientific community. The reason for this could be the ability of ANN in handling data with high unpredictability and non-constant variance (Ayodele et al., 2021). The main drawback of ANN is that it requires a large amount of data which is necessary for accurate prediction. Moreover, the training phase is sensitive to the type of the algorithm, bias and learning rates (Cheng et al., 2020). In case of incomplete data, FL, in particular the models with the optimized membership function, reflected better performance (Ansari et al., 2020). For instance, in case of control strategy, FL was preferred over nonlinear model predictive control methods. (Revollar et al., 2018; Bae et al., 2006; Bello et al., 2014; Bertanza et al., 2020; Santín et al., 2015; Tejaswini et al., 2020b). The availability of the data was the factor affecting the performance of both hybrid and stand-alone models (Alsadaie et al., 2016; Honggui Han et al., 2018b; Heddam et al., 2012; Hsu, 2009; Huang et al., 2016, 2011, 2010, 2009a).

[Table 5, here]

The literature indicates that the main source of data was local authorities (64% of the reviewed studies). The remaining portion was either gathered from laboratory experiments (12%) or full-scale treatment plants (24%). The availability of the input data is a critical factor in these models, which was rigorously criticized by previous review papers (Corominas et al., 2018). The researchers mentioned that they were left choiceless with no option other than working on the available data which influences the type, structure, efficiency and accuracy of the data-driven models. Additionally, majority of treatment plants are prone to high fluctuations, either in the influent stream or in the processes, which are mainly not captured by the prediction models. For instance, usually the magnitude of the influent loading rate is the crucial factor for operational management of the treatment plants. For instance, DO is adjusted based on the loading rate which is highly correlated with the energy consumption of the plant. So, the maximum and minimum loading is necessary for operational purposes. Most of the proposed prediction models are not capable of predicting the local maximum and minimum of different time series and the overall evaluation of the model performance is based on several statistical indicators, which automatically ignore the highest and lowest magnitude of the time series. To be implemented in the real treatment plant, the model must be capable of providing the peaks as well as the local minimums of the operational parameters in different time steps, a point which was abundantly neglected in the literature.

Simulation time is also a very important factor in modelling studies. Despite its importance in prediction, simulation time rarely (7%) mentioned in the reviewed studies. Simulation times was reported to range from 0.01 seconds to 20 minutes. The simulation time can be significant when a large dataset is used for modelling the processes within the plant. A model trained with large datasets requires longer evaluation process and if a very vibrant parameter such as DO, pH or ORP (Oxidation Reduction Potential) is neglected during the training process, the validity of the modelling procedure could be questioned. Small treatment plants are subjected to different dynamics with high fluctuations, spanning from minutes (pH and temperature) to years or decades (plant maintenance and construction). While working on dormant variables such as clarifier volume and sludge concentration, the simulation time would not be too long, however in case of volatile variables such as pH the simulation time will have significant effect on the validity of the model. For this purpose, some scholars suggested to develop a repository or inventory of the data to comprehensively represent the abundant dynamics within the treatment plant and adjust the simulation time based on the compulsory criteria provided by these inventories (Corominas et al., 2018; Newhart et al., 2019; Ye et al., 2020).

Another limitation that we identified, is related to the input data and how to prepare it for the modelling endeavor. High quality data is vital for efficient model development and validation. Most of studies tried to assess the model with inadequate number of data provided by state and local authorities, environmental protection agencies, international organizations, or the literature. The representativeness of the data is also questionable. Most of the values in the gathered data is not the exact measurements of the intended variables, but the signal received by the sensors located in different places. Sensors receive the magnitude of the signal generated in or adjacent to the place of the sensor, not necessarily in the entire boundary. It is possible that sensors do not necessarily represent the exact values of the parameters. For the sensor located in the middle of the aeration tank, the TSS signal is representative of the TSS concentration in the middle of the tank, not for the entire tank. This leads to unprecise and unrepresentative measurements inside the process. For better understanding of the input data, the source of data and the methods of measurements need to be clearly mentioned. To alleviate such errors, simultaneous use of several sensors at different locations within the aeration tank would be useful. Moreover, some sensors are in the regular need for calibration, maintenance and replacement to ensure the accurate measurements. Some scholars also suggested data curing and statistical significance of the gathered data before any necessary action on the modelling endeavor (Bijlsma et al., 2021; Hashimoto et al., 2021; Qiu et al., 2018).

4.2. Data quality and data pre-processing

4.2.1. Data quality

Quality of data is essential for machine learning and the occurrence of noisy data in any dataset can dramatically influence the accuracy of any ML technique and lead to decreased classified dataset. Before applying machine learning models or data mining algorithms to any process, the quality of data in terms of accuracy, completeness, consistency, timeliness, believability, and interpretability must be checked. Several studies tried to understand the effect of noisy datasets on machine classifiers. Two distinct group of approaches can be categorized for this purpose. First group is the approaches that try to learn directly from noisy labels and focus on noise-robust algorithms, e.g., Beigman & Klebanov (2009); Guan et al. (2019); Huang et al. (2017); Manwani & Sastry (2013); Van Horn et al. (2015). The second group is mainly focusing on the label-cleansing methods which tries to remove, replace, or correct mislabeled data (Brodley and Friedl, 1999). The procedure is to combine the noisy data with a small portion of clear data using semi-supervised techniques. Moreover, some case studies reported the capacity of DNN in working with noisy datasets (Rolnick et. al., 2018). Reis et. al. (2019) suggested using probability distribution functions, rather than deterministic quantities for the noisy data. The performance of Probabilistic Random Forest (PRF) and RF were compared when a set of noise were injected to the dataset. The PRF outperformed RF in all cases, with a slight increase in running time. The classification accuracy increased by 10% and 30% for noisy labels and noisy features, respectively.

4.2.2. Data preprocessing

Generally, the raw data generated in wastewater treatment needs to be preprocessed for effective prediction, which is not the case for the main portion of the literature (more than 95%). Where possible, the authors stated the unknown source of data or inability to amend the raw data due to ownership issues (Revollar et, al., 2018; Mazhar et. al., 2019; Sanitin et. al., 2018; Yel and Yalpir, 2011). However, data preprocessing prior to model development was considered in several deep learning studies and various methods were introduced in literature for data curing and cleaning.

[Figure 6, Here]

As shown on Figure 6, generally the data preprocessing task can be divided into five distinct classes. Data cleaning tries to identify the errors in the data and replace them with correct ones. Outlier removal, statistical imputation, and low variance exclusion are subclasses of data cleaning techniques. If the dataset has a Gaussian distribution, the standard deviation can be used as cut-off value for outlier detection. The interquartile ranging (IQR) methods can be used for a non-Gaussian distribution sample of data. The IQR is calculated as the difference between the 75^{th} and the 25^{th} percentiles (the third and first quartile) of the data and defines the box in a box and whisker plot. The IQR can be used to identify outliers by defining limits on the sample values that are a factor k of the IQR below the 25^{th} percentile or above the 75^{th} percentile. The common value for the factor k is 1.5. A factor k of 3 or more can be used to identify values that are extreme outliers or far outs when described in the context of box and whisker plots.

Feature selection finds the input variables that are highly relevant to the modelling task. Subclasses of feature selection methods are intrinsic, wrapper and filter methods. Intrinsic methods perform algorithms that automatically detect feature selection during training. Wrapper methods search subsets of features that perform according to a predictive model; and Filter methods select subsets of features based on their relationship with the target.

Data transforms implement new probability distribution for the input variables. The purpose is to transform the raw variables into Gaussian variables. Popular techniques on the literature are normalization and standardization. While normalization techniques transform data into a value between 0 and 1, standardization techniques rescale data in a form to have a mean of 0 and standard deviation of 1. Min-Max, Z score and decimal scaling methods can be used for normalization and Power, Box-Cox and Yeo-Johnson Transform can be mentioned as popular methods for data standardization.

Feature engineering is the process of producing new variables that are not in the learning set and represent the main features of the dataset. The aim is to extracts new variables from the raw data to detect the most influential variables on the learning process. Polynomial power transforms are frequently used to expose the relationship between the input variables and the target.

4.3. Dimensionality reduction compacts raw data to derive smaller projection. The procedure significantly reduces the number of the input variables in the training data. All feature selection methods can be categorized as dimensionality reduction. Furthermore, techniques from linear algebra like matrix factorization where the eigen decomposition and singular value decomposition of dataset are extracted, can be used to reduce the number of inputs into the model.**Simulation Platform**

Several AI software packages, and platforms have been utilized to model WWTPs with AI techniques. Table 6 summarizes these platforms in terms of their application frequency and availability. The table intends to guide researchers in the selection of the most suitable tool for their study. Approximately 69% of the reviewed models used MATLAB. Its Neural Network toolbox was frequently used, particularly in the training and testing of neural networks and MLP algorithms. One of the advantages of MATLAB over other platforms such as Python, is that it is a self-contained package. Moreover, Simulink-MATLAB system with its large library of functions, is a unique product. MATLAB also is a user-friendly platform, and its ANN and Fuzzy toolboxes are designed graphically and can be used with minimum knowledge of data science. This enables the faster understanding and expeditious implementation of models. Arguably this can also have downsides. Researchers with poor knowledge of data science using an "easy" software platform may not be able to use it in a critical way. Therefore, deeper analysis of methods used, and generated results is crucial for the validity of conclusions.

The other commonly used software was Design Expert (@*Stat-Ease Inc.*) contributing to about 11% of the reviewed models. The ability to provide comparative tests, screening, characterization, optimization, robust parameter design as well as mixture designs enabled the platform to outnumber other well-known platforms such as Python and Statistica (Adeogun et al., 2021; Bhatti et al., 2011; Aghilesh et al., 2021). Statistica, *R* programming and Python were the next frequently used platforms in the literature, contributing to about 5%, 4% and 4% of the reviewed studies, respectively. Statistica has different libraries for machine learning and data analysis. The *R* programming language is widely used for statistical analysis and its recent versions includes libraries for data analysis and machine learning. In the reviewed studies, Python language was rarely used to simulate AI models (2.4%). C++ has multiple AI libraries including OpenNN, OpenCV, BOOST, gflags, glog, and Tensor Flow. Despite having a variety of the AI libraries, C++ was seldom used in the literature (less than 1%).

[Table 6, here]

4.4. Transition from knowledge into practice

Almost 8% of the reviewed studies claim the extraction of commercial products as a result of the forecasting research. The remaining 92% represented the research in the form of an academic practice with further potential of commercial applications. In either case, a mutual collaboration with private sector is vital for wider application and usability of the developed models as mature products. AI companies can play a vital role on this matter by providing graphical and user-friendly platforms to enhance the usability of the techniques. S-CAN$^{@}$ is using automatic cleaning for sensors to guarantee the quality of the real time data. While such companies are focusing on the data quality in biological WWTPs, some big water companies are focusing on development of the digital twin for more efficient simulation of the wastewater treatment processes. Digital twins are continuously updated with historical as well as current data from different sources such as Supervisory Control and Data Acquisition (SCADA) systems, sensors, meters, and other measured sources to create a real time representation immediately after information becomes available. This feature enables utilities to better understand the past and current performance of their WWTP while helping them predict future performance and simulate the impact of potential changes in the virtual world before taking actions. With a robust algorithm which rely on data from more than 100 plants in operation Veolia@ has developed the Hubgrade digital twin module for holistic assessment of WWTP performance.

Some recent patents on data quality in WWTPs also provide practical application. Notably, the idea of using dual sensors in WWTPs seems to be a reasonable solution to increase data quality (Choi, 2014). In this method a control unit for controlling operation of the pumps based on a sensed value of the sensor is located inside the WWTP. The sensor comprises a first unit sensor and a second unit sensor which are working independently from each other. The control unit controls the operation of the pumps based on the highest sensed value of the first unit sensor or the second unit sensor. Therefore, in case of an overflow phenomenon, caused by malfunction of a sensor, can be prevented as the water level in the treatment tank is reliably sensed by the sensors which are independently installed.

In situ monitoring system is another emerging idea to ensure data quality in WWTPs. In this method a data processor, embedded with a library of real time historical WWTP data and a cloud storage unit, is connected to the sensors located on different places within the WWTP (Duden et al., 2013). The sensors are operative to transmit the data representing the characteristics of the measured parameter into the data processor. The data processor compares the received signal with the similar historical values in the library and if any significant statistical anomaly is not present, the processor transmits the signal to cloud storage unit.

Other practical applications of data-driven techniques in WWTPs are the soft sensors. These sensors employ different machine learning algorithms to make a sensible approximation of the water quality indicators. An example is the COD soft sensors (Miao, 2008). The COD soft sensor is designed against the problem of difficult COD on-line measurement by applying the rapid Extended Fourier Amplitude Sensitivity Test (EFAST) method for pruning redundant neurons, simplifying the neural network structure and carrying out the soft measurement of COD according to the non-linear characteristic of the wastewater treatment process.

Another example of successful methods in practice is FL which not only used for diagnosis and control of treatment (Pires et al., 2005) but also for monitoring the greenhouse gases emissions (Santín et al., 2018). To introduce such reliable and advanced techniques, the research community should be capable of proposing new alternatives and practitioners should be able to guarantee the data quality as much as possible. There also is a need for deeper interaction of researchers with different background. To bring knowledge into practice, close interaction of researchers with multidisciplinary backgrounds would be necessary for successful application of the developed techniques in industry.

5. Conclusion

A comprehensive review of the available literature on the application of AI models to prognosticate different dynamics inside the WWTPs was presented. The forecasting models along with the optimization techniques assist operators and utilities to assess the impact of current and future scenarios on WWTPs. The accuracy of these models lays on several parameters such as data quality and its cleaning method, and the selection of appropriate model type and structure. The complexity and nature of the research problem along with the objectives of the study are the critical determinants of method selection. Moreover, the availability and quality of the necessary data is another important factor influencing the method choice. The review of more than 280 AI models in WWTPs, revealed the use of 188 single models and 115 hybrid models. Among single models, ANN was used in 77 case studies, followed by DT, FL, GA and SVM. In case of hybrid models, the combination of ML and metaheuristic algorithms outnumbered other combinations. The other popular hybrid models, in descending order, were ML-ML, ML-ML-metaheuristic, metaheuristic-metaheuristic and ML-ML-ML.

In case of accuracy measurement, different statistical indicators were used to measure. R^2, *RMSE* and *R* and were the most frequently used indicators. Hybrid models performed relatively better than standalone models (single models). In case of incomplete data, FL was the favorite method among standalone models.

Some models neither studied nor considered in training and validation stages. For instance, reinforcement learning has the potential to forecast many features of WWTP data. There are several obstacles for wide application of reinforcement learning, of which the notable one could be the need for large training data to produce reliable results. The further application of AI, in particular data-driven models, is heavily relying on the quality of the data generated in WWTPs. Though the generated data is quantitively significant, the quality of the generated data is not satisfactory for extraction of meaningful information. The current data-rich, information-poor status on the literature is pushing the attentions toward the generation, storage and maintenance of more representative and high-quality data in WWTPs.

Declaration of Competing Interest

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

Acknowledgement

This publication has been financially supported by Science Foundation Ireland under the SFI Strategic Partnership Programme Grant No. SFI/15/SPP/E3125. The opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Science Foundation Ireland.

References

Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., Mahfood, B., 2020. Artificial intelligence applications in solid waste management: A systematic research review. Waste Manag. 109, 231–246. https://doi.org/10.1016/j.wasman.2020.04.057

Adeogun, A.I., Bhagawati, P.B., Shivayogimath, C.B., 2021. Pollutants removals and energy consumption in electrochemical cell for pulping processes wastewater treatment: Artificial neural network, response surface methodology and kinetic studies. J. Environ. Manage. 281, 111897. https://doi.org/10.1016/j.jenvman.2020.111897

Aish, A.M., Zaqoot, H.A., Abdeljawad, S.M., 2015. Artificial neural network approach for predicting reverse osmosis desalination plants performance in the Gaza Strip. Desalination 367, 240–247. https://doi.org/10.1016/j.desal.2015.04.008

Akbaş, H., Bilgen, B., Turhan, A., M., An integrated prediction and optimization model of biogas production system at a wastewater treatment facility, Bioresource Technology, 196, 2015, 566-576, https://doi.org/10.1016/j.biortech.2015.08.017.

Al-Obaidi, M.A., Li, J.P., Alsadaie, S., Kara-Zaïtri, C., Mujtaba, I.M., 2018. Modelling and optimisation of a multistage Reverse Osmosis processes with permeate reprocessing and recycling for the removal of N-nitrosodimethylamine from wastewater using Species Conserving Genetic Algorithms. Chem. Eng. J. 350, 824–834. https://doi.org/10.1016/j.cej.2018.06.022

Al-Obaidi, M.A., Li, J.P., Kara-Zaïtri, C., Mujtaba, I.M., 2017. Optimisation of reverse osmosis based wastewater treatment system for the removal of chlorophenol using genetic algorithms. Chem. Eng. J. 316, 91–100. https://doi.org/10.1016/j.cej.2016.12.096

Ali Akbar Miran Beigi, 2020. A Reliable and Rapid Method for Simultaneous Determination of Furfural and Hydroxymethyl Furfural in Oil Refinery Wastewaters by Ant Colony/Partial Least-Squares Analysis. J. Anal. Chem. 75, 1486–1496. https://doi.org/10.1134/S1061934820110039

Almomani, F., 2020. Prediction the performance of multistage moving bed biological process using artificial neural network (ANN). Sci. Total Environ. 744, 140854. https://doi.org/10.1016/j.scitotenv.2020.140854

Alsadaie, S., Zanil, M.F. bin, Hussain, A., Mujtaba, I.M., 2016. Development of Hybrid Fuzzy-GMC Control System for MSF Desalination Process, Computer Aided Chemical Engineering. Elsevier

Masson SAS. https://doi.org/10.1016/B978-0-444-63428-3.50156-9

Alves Goulart, D., Dutra Pereira, R., 2020. Autonomous pH control by reinforcement learning for electroplating industry wastewater. Comput. Chem. Eng. 140. https://doi.org/10.1016/j.compchemeng.2020.106909

Alwan, G., 2012. Adaptive Genetic PH Control of a Wastewater Treatment Unit via LAB View. J. Chem. Process Eng. Res. 5, 22–32.

Ansari, M., Othman, F., El-sha, A., 2020. Science of the Total Environment Optimized fuzzy inference system to enhance prediction accuracy for in fl uent characteristics of a sewage treatment plant 722. https://doi.org/10.1016/j.scitotenv.2020.137878

Antwi, P., Zhang, D., Xiao, L., Kabutey, F.T., Quashie, F.K., Luo, W., Meng, J., Li, J., 2019a. Modeling the performance of Single-stage Nitrogen removal using Anammox and Partial nitritation (SNAP) process with backpropagation neural network and response surface methodology. Sci. Total Environ. 690, 108–120. https://doi.org/10.1016/j.scitotenv.2019.06.530

Antwi, P., Zhang, D., Xiao, L., Kabutey, F.T., Quashie, F.K., Luo, W., Meng, J., Li, J., 2019b. Modeling the performance of Single-stage Nitrogen removal using Anammox and Partial nitritation (SNAP) process with backpropagation neural network and response surface methodology. Sci. Total Environ. 690, 108–120. https://doi.org/10.1016/j.scitotenv.2019.06.530

Asadi, A., Verma, A., Yang, K., Mejabi, B., 2017. Wastewater treatment aeration process optimization: A data mining approach. J. Environ. Manage. 203, 630–639. https://doi.org/10.1016/j.jenvman.2016.07.047

Asadi, M., Guo, H., McPhedran, K., 2020. Biogas production estimation using data-driven approaches for cold region municipal wastewater anaerobic digestion. J. Environ. Manage. 253, 109708. https://doi.org/10.1016/j.jenvman.2019.109708

Asfaram, A., Ghaedi, M., Hajati, S., Goudarzi, A., 2016. Synthesis of magnetic γ-Fe2O3-based nanomaterial for ultrasonic assisted dyes adsorption: Modeling and optimization. Ultrason. Sonochem. 32, 418–431. https://doi.org/10.1016/j.ultsonch.2016.04.011

Attwa, M., Zamzam, S., 2020. An integrated approach of GIS and geoelectrical techniques for wastewater leakage investigations: Active constraint balancing and genetic algorithms application. J. Appl. Geophys. 175, 103992. https://doi.org/10.1016/j.jappgeo.2020.103992

Ayodele, B.V., Alsaffar, M.A., Mustapa, S.I., Cheng, C.K., Witoon, T., 2021. Modeling the effect of process parameters on the photocatalytic degradation of organic pollutants using artificial neural networks. Process Saf. Environ. Prot. 145, 120–132. https://doi.org/10.1016/j.psep.2020.07.053

Babić, S., Mutavdžić, D., Ašperger, D., Horvat, A.J.M., Kaštelan-Macan, M., 2007. Determination of veterinary pharmaceuticals in production wastewater by HPTLC-videodensitometry. Chromatographia 65, 105–110. https://doi.org/10.1365/s10337-006-0109-2

Badrnezhad, R., Mirza, B., 2014. Modeling and optimization of cross-flow ultrafiltration using hybrid neural network-genetic algorithm approach. J. Ind. Eng. Chem. 20, 528–543. https://doi.org/10.1016/j.jiec.2013.05.012

Bae, H., Seo, H.Y., Kim, S., Kim, Y., 2006. Knowledge-based control and case-based diagnosis based upon empirical knowledge and fuzzy logic for the SBR plant. Water Sci. Technol. 53, 217–224. https://doi.org/10.2166/wst.2006.024

Bagheri, M., Mirbagheri, S.A., Ehteshami, M., Bagheri, Z., 2015. Modeling of a sequencing batch reactor treating municipal wastewater using multi-layer perceptron and radial basis function artificial neural networks. Process Saf. Environ. Prot. 93, 111–123. https://doi.org/10.1016/j.psep.2014.04.006

Bai, J., Liu, H., Yin, B., Ma, H., 2015. Modeling of enhanced VFAs production from waste activated sludge by modified ADM1 with improved particle swarm optimization for parameters estimation. Biochem. Eng. J. 103, 22–31. https://doi.org/10.1016/j.bej.2015.06.015

Baki, O.T., Aras, E., 2018. Estimation of BOD in wastewater treatment plant by using different ANN algorithms. Membr. Water Treat. 9, 455–462. https://doi.org/10.12989/mwt.2018.9.6.455

Bakr, M.H., Nasr, M., Ashmawy, M., Tawfik, A., 2021. Environmental Technology & Innovation Predictive performance of auto-aerated immobilized biomass reactor treating anaerobic effluent of cardboard wastewater enriched with bronopol (2-bromo-2-nitropropan-1 , 3-diol) via artificial neural network. Environ. Technol. Innov. 21, 101327. https://doi.org/10.1016/j.eti.2020.101327

Balasubramani, K., Sivarajasekar, N., Muthusaravanan, S., Ram, K., Naushad, M., Ahamad, T., Sharma, G., 2020. Efficient removal of antidepressant Flupentixol using graphene oxide/cellulose nanogel composite: Particle swarm algorithm based artificial neural network modelling and optimization. J. Mol. Liq. 319, 114371. https://doi.org/10.1016/j.molliq.2020.114371

Beigman, E., Klebanov, B.B. (2009). Learning with Annotation Noise. Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP, Singapore, 280–287.

Belchior, C.A.C., Araújo, R.A.M., Landeck, J.A.C., 2012. Dissolved oxygen control of the activated sludge wastewater treatment process using stable adaptive fuzzy control. Comput. Chem. Eng. 37, 152–162. https://doi.org/10.1016/j.compchemeng.2011.09.011

Bello, O., Hamam, Y., Djouani, K., 2014. Fuzzy dynamic modelling and predictive control of a coagulation chemical dosing unit for water treatment plants. J. Electr. Syst. Inf. Technol. 1, 129–143. https://doi.org/10.1016/j.jesit.2014.08.001

Beltramo, T., Klocke, M., Hitzmann, B., 2019. Prediction of the biogas production using GA and ACO input features selection method for ANN model. Inf. Process. Agric. 6, 349–356. https://doi.org/10.1016/j.inpa.2019.01.002

Beltramo, T., Ranzan, C., Hinrichs, J., Hitzmann, B., 2016. Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm. Biosyst. Eng. 143, 68–78. https://doi.org/10.1016/j.biosystemseng.2016.01.006

Béraud, B., Steyer, J.P., Lemoine, C., Latrille, E., Manic, G., Printemps-Vacquier, C., 2007. Towards a global multi objective optimization of wastewater treatment plant based on modeling and genetic algorithms. Water Sci. Technol. 56, 109–116. https://doi.org/10.2166/wst.2007.592

Bernardelli, A., Marsili-Libelli, S., Manzini, A., Stancari, S., Tardini, G., Montanari, D., Anceschi, G., Gelli, P., Venier, S., 2020. Real-time model predictive control of a wastewater treatment plant based on machine learning. Water Sci. Technol. 81, 2391–2400. https://doi.org/10.2166/wst.2020.298

Bernier, J., Rocher, V., Dochain, D., n.d. of of Management of of Management of of treatment improving Management of of treatment improving treatment basin for. IFAC-PapersOnLine 52, 358–363. https://doi.org/10.1016/j.ifacol.2019.06.088

Bertanza, G., Menoni, L., Baroni, P., 2020. Energy saving for air supply in a real WWTP: Application of a fuzzy logic controller. Water Sci. Technol. 81, 1552–1557. https://doi.org/10.2166/wst.2020.084

Besharati Fard, M., Mirbagheri, S.A., Pendashteh, A., Alavi, J., 2020. Estimation of effluent parameters of slaughterhouse wastewater treatment with artificial neural network and B-spline quasi interpolation. Int. J. Environ. Res. 14, 527–539. https://doi.org/10.1007/s41742-020-00274-1

Bhagat, S.K., Tung, T.M., Yaseen, Z.M., 2020. Development of artificial intelligence for modeling wastewater heavy metal removal: State of the art, application assessment and possible future research. J. Clean. Prod. 250, 119473. https://doi.org/10.1016/j.jclepro.2019.119473

Bhatti, M.S., Kapoor, D., Kalia, R.K., Reddy, A.S., Thukral, A.K., 2011. RSM and ANN modeling for electrocoagulation of copper from simulated wastewater: Multi objective optimization using genetic algorithm approach. Desalination 274, 74–80. https://doi.org/10.1016/j.desal.2011.01.083

Bijlsma, L., Picó, Y., Andreu, V., Celma, A., Estévez-Danta, A., González-Mariño, I., Hernández, F., López de Alda, M., López-García, E., Marcé, R.M., Miró, M., Montes, R., Pérez de San Román-Landa, U., Pitarch, E., Pocurull, E., Postigo, C., Prieto, A., Rico, A., Rodil, R., Valcárcel, Y., Ventura, M., Quintana, J.B., 2021. The embodiment of wastewater data for the estimation of illicit drug consumption in Spain. Sci. Total Environ. 772, 144794. https://doi.org/10.1016/j.scitotenv.2020.144794

Bishoff, D., AlSayed, A., Hosen, S., Menon, P., ElDyasti, A., 2021. Effect of COD on methanotrophic growth and the anaerobic digestibility of its sludge as a further step for integration in WWTPS. J. Environ. Manage. 290, 112543. https://doi.org/10.1016/j.jenvman.2021.112543

Boiocchi, R., Gernaey, K. V., Sin, G., 2016. Control of wastewater N2O emissions by balancing the microbial communities using a fuzzy-logic approach. IFAC-PapersOnLine 49, 1157–1162. https://doi.org/10.1016/j.ifacol.2016.07.359

Borges, R.M., Mattedi, A., Munaro, C.J., Gonçalves, R.F., 2016. A modular diagnosis system based on fuzzy logic for UASB reactors treating sewage 309–317. https://doi.org/10.2166/wst.2016.156

Boztoprak, H., Özbay, Y., Güçlü, D., Küçükhemek, M., 2016. Prediction of sludge volume index bulking using image analysis and neural network at a full-scale activated sludge plant. Desalin. Water Treat. 57, 17195–17205. https://doi.org/10.1080/19443994.2015.1085909

Brand, N., Ostfeld, A., 2011. Optimal Design of Regional Wastewater Pipelines and Treatment Plant Systems. Water Environ. Res. 83, 53–64. https://doi.org/10.2175/106143010x12780288628219

Brodley, C.E., Freidl, M.A., 1999. Identifying Mislabeled Training Data. Artificial Intel. Research. 131 - 167.

Byliński, H., Sobecki, A., Gebicki, J., 2019. The use of artificial neural networks and decision trees to predict the degree of odor nuisance of post-digestion sludge in the sewage treatment plant process. Sustain. https://doi.org/10.3390/su11164407

Cai, W., Lesnik, K.L., Wade, M.J., Heidrich, E.S., Wang, Y., Liu, H., 2019. Incorporating microbial community data with machine learning techniques to predict feed substrates in microbial fuel cells. Biosens. Bioelectron. 133, 64–71. https://doi.org/10.1016/j.bios.2019.03.021

Caraman, S., Luca, L., Vasiliev, I., Barbu, M., 2020. Optimal-setpoint-based control strategy of a wastewater treatment process. Processes 8, 1–16. https://doi.org/10.3390/PR8101203

Carrasco, E.F., Rodríguez, J., Pual, A., Roca, E., Lema, J.M., 2002. Rule-based diagnosis and supervision of a pilot-scale wastewater treatment plant using fuzzy logic techniques. Expert Syst. Appl. 22, 11–20. https://doi.org/10.1016/S0957-4174(01)00045-8

Çelik, U., Yurtay, N., Sertkaya, C., 2013. Wastewater Effluent Prediction Based on Decision Tree. J. Selçuk Un 1, 138–148.

Chang, N. Bin, Chen, W.C., Shieh, W.K., 2001. Optimal control of wastewater treatment plants via integrated neural network and genetic algorithms. Civ. Eng. Environ. Syst. 18, 1–17. https://doi.org/10.1080/02630250108970290

Chang, P., Li, Z.Y., 2021. Over-complete deep recurrent neutral network based on wastewater treatment process soft sensor application. Appl. Soft Comput. 105, 107227. https://doi.org/10.1016/j.asoc.2021.107227

Chang, W., Chang, J., Yin, H., 2012. on Fuzzy Decision Making 2395–2400.

Chen, G., Wang, J., Li, R., 2016. Parameter identification for a water quality model using two hybrid swarm intelligence algorithms. Soft Comput. 20, 2829–2839. https://doi.org/10.1007/s00500-015-1684-7

Chen, H., Chen, A., Xu, L., Xie, H., Qiao, H., Lin, Q., Cai, K., 2020. A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. Agric. Water Manag. 240, 106303. https://doi.org/10.1016/j.agwat.2020.106303

Chen, H.W., Yu, R.F., Ning, S.K., Huang, H.C., 2010. Forecasting effluent quality of an industry wastewater treatment plant by evolutionary grey dynamic model. Resour. Conserv. Recycl. 54, 235–241. https://doi.org/10.1016/j.resconrec.2009.08.005

Chen, Y., Xu, J., Yu, H., Zhen, Z., Li, D., 2016. Three-Dimensional Short-Term Prediction Model of Dissolved Oxygen Content Based on PSO-BPANN Algorithm Coupled with Kriging Interpolation. Math. Probl. Eng. 2016. https://doi.org/10.1155/2016/6564202

Chen, Y., Yu, G., Long, Y., Teng, J., You, X., Liao, B., Lin, H., 2019. Bioresource Technology Application of radial basis function arti fi cial neural network to quantify interfacial energies related to membrane fouling in a membrane bioreactor. Bioresour. Technol. 293, 122103. https://doi.org/10.1016/j.biortech.2019.122103

Cheng, T., Harrou, F., Kadri, F., Sun, Y., Leiknes, T., 2020. Forecasting of Wastewater Treatment Plant Key Features Using Deep Learning-Based Models: A Case Study. IEEE Access 8, 184475–184485. https://doi.org/10.1109/access.2020.3030820

Chiu, Y.C., Chiang, C.W., Lee, T.Y., 2017. Prediction of biochemical oxygen demand at the upstream catchment of a reservoir using adaptive neuro fuzzy inference system. Water Sci. Technol. 76, 1739–1753. https://doi.org/10.2166/wst.2017.359

Chung, E.S., Kim, Y., 2014. Development of fuzzy multi-criteria approach to prioritize locations of treated wastewater use considering climate change scenarios. J. Environ. Manage. 146, 505–516. https://doi.org/10.1016/j.jenvman.2014.08.013

Corominas, L., Garrido-Baserba, M., Villez, K., Olsson, G., Cortés, U., Poch, M., 2018. Transforming data into knowledge for improved wastewater treatment operation: A critical review of techniques. Environ. Model. Softw. 106, 89–103. https://doi.org/10.1016/j.envsoft.2017.11.023

da Silva Ribeiro, T., Grossi, C.D., Merma, A.G., dos Santos, B.F., Torem, M.L., 2019. Removal of boron from mining wastewaters by electrocoagulation method: Modelling experimental data using artificial neural networks. Miner. Eng. 131, 8–13. https://doi.org/10.1016/j.mineng.2018.10.016

Dawood, A.S., Li, Y., 2013. Modeling and optimization of new flocculant dosage and pH for flocculation: Removal of pollutants from wastewater. Water (Switzerland) 5, 342–355. https://doi.org/10.3390/w5020342

De Miranda Ramos Soares, A.P., De Oliveira Carvalho, F., De Farias Silva, C.E., Da Silva Gonçalves, A.H., De Souza Abud, A.K., 2020. Random Forest as a promising application to predict basic-dye biosorption process using orange waste. J. Environ. Chem. Eng. 8, 103952. https://doi.org/10.1016/j.jece.2020.103952

de Moraes, N.F.S., Santana, R.M.R., Gomes, R.K.M., Santos Júnior, S.G., de Lucena, A.L.A., Zaidan, L.E.M.C., Napoleão, D.C., 2021. Performance verification of different advanced oxidation processes in the degradation of the dye acid violet 17: reaction kinetics, toxicity and degradation prediction by artificial neural networks. Chem. Pap. 75, 539–552. https://doi.org/10.1007/s11696-020-01325-9

Deepnarain, N., Nasr, M., Kumari, S., Stenström, T.A., Reddy, P., Pillay, K., Bux, F., 2019. Decision tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. Process Saf. Environ. Prot. 126, 25–34. https://doi.org/10.1016/j.psep.2019.02.023

Dehghani, M., Seifi, A., Riahi-Madvar, H., 2019. Novel forecasting models for immediate-short-term to long-term influent flow prediction by combining ANFIS and grey wolf optimization. J. Hydrol. 576, 698–725. https://doi.org/10.1016/j.jhydrol.2019.06.065

Do, K., Schmitt, F., 2020. Artificial Intelligence Model for Forecasting of Membrane Fouling in Wastewater Treatment by Membrane Technology. Model. Membr. Membr. Process. 301–325. https://doi.org/10.1002/9781119536260.ch9

Du, X., Ma, Y., Wei, X., Jegatheesan, V., 2020. Optimal parameter estimation in activated sludge process based wastewater treatment practice. Water (Switzerland) 12. https://doi.org/10.3390/W12092604

Dürrenmatt, D.J.Ô., Gujer, W., 2012. Data-driven modeling approaches to support wastewater treatment plant operation. Environ. Model. Softw. 30, 47–56. https://doi.org/10.1016/j.envsoft.2011.11.007

Ehteram, M., Salih, S.Q., Yaseen, Z.M., 2020. Efficiency evaluation of reverse osmosis desalination plant using hybridized multilayer perceptron with particle swarm optimization 15278–15291.

Fan, W., Guo, T., Gao, S., Lu, Y., Meng, Y., Huo, M., 2021. Evolution of dissolved organic matter during artificial groundwater recharge with effluent from underutilized WWTP and the resulting facilitated transport effect. Environ. Res. 193, 110527. https://doi.org/10.1016/j.envres.2020.110527

Fan, W., Zhang, Z., 2020. A CNN-SVR Hybrid Prediction Model for Wastewater Index Measurement. Proc. - 2020 2nd Int. Conf. Adv. Comput. Technol. Inf. Sci. Commun. CTISC 2020 90–94. https://doi.org/10.1109/CTISC49998.2020.00022

Fang, A., Cui, L., Zhang, Z., 2019. Multi-kernel SVM based multi-feature fusion algorithm for analyzing wastewater membrane permeation in environmental pollution. Desalin. Water Treat. 163, 415–420. https://doi.org/10.5004/dwt.2019.24216

Fang, F., Qiao, L.L., Cao, J.S., Li, Y., Xie, W.M., Sheng, G.P., Yu, H.Q., 2016. Quantitative evaluation of A2O and reversed A2O processes for biological municipal wastewater treatment using a projection pursuit method. Sep. Purif. Technol. 166, 164–170. https://doi.org/10.1016/j.seppur.2016.04.036

Farzin, S., Chianeh, F.N., Anaraki, M.V., Mahmoudian, F., 2020. Introducing a framework for modeling of drug electrochemical removal from wastewater based on data mining algorithms, scatter interpolation method, and multi criteria decision analysis (DID). J. Clean. Prod. 266, 122075. https://doi.org/10.1016/j.jclepro.2020.122075

Fawzy, M., Nasr, M., Adel, S., Nagy, H., Helmi, S., 2016. Environmental approach and artificial intelligence for Ni(II) and Cd(II) biosorption from aqueous solution using Typha domingensis biomass. Ecol. Eng. 95, 743–752. https://doi.org/10.1016/j.ecoleng.2016.07.007

Fernandez de Canete, J., del Saz-Orozco, P., Gómez-de-Gabriel, J., Baratti, R., Ruano, A., Rivas-Blanco, I., 2021. Control and soft sensing strategies for a wastewater treatment plant using a neuro-genetic approach. Comput. Chem. Eng. 144, 107146. https://doi.org/10.1016/j.compchemeng.2020.107146

Fernandez, F.J., Seco, A., Ferrer, J., Rodrigo, M.A., 2009. Use of neurofuzzy networks to improve wastewater flow-rate forecasting. Environ. Model. Softw. 24, 686–693. https://doi.org/10.1016/j.envsoft.2008.10.010

Fine, T.L., 2005. Fundamentals of Artificial Neural Networks [Book Reviews]. IEEE Trans. Inf. Theory 42, 1322. https://doi.org/10.1109/tit.1996.508868

Fiter, M., Güell, D., Comas, J., Colprim, J., Poch, M., Rodríguez-Roda, I., 2005. Energy saving in a wastewater treatment process: An application of Fuzzy Logic Control. Environ. Technol. 26, 1263–1270. https://doi.org/10.1080/09593332608618596

Fonseca, R.R., Franco, I.C., Thompson, J.P., da Silva, F.V., 2018. Turbidity control on dissolved air flotation process using fuzzy logic. Water Sci. Technol. 78, 2586–2596. https://doi.org/10.2166/wst.2019.015

Foroughi, M., Rahmani, A.R., Asgari, G., Nematollahi, D., Yetilmezsoy, K., Samarghandi, M.R., 2020. Optimization and Modeling of Tetracycline Removal from Wastewater by Three-Dimensional Electrochemical System: Application of Response Surface Methodology and Least Squares Support Vector Machine. Environ. Model. Assess. 25, 327–341. https://doi.org/10.1007/s10666-019-09675-9

Fu, Z., Cheng, J., Yang, M., Batista, J., Jiang, Y., 2020. Wastewater discharge quality prediction using stratified sampling and wavelet de-noising ANFIS model. Comput. Electr. Eng. 85. https://doi.org/10.1016/j.compeleceng.2020.106701

Ge, L., Ji, E., 2020. An improved artificial bee colony algorithm and its application in machine learning. J. Phys. Conf. Ser. 1650, 83–88. https://doi.org/10.1088/1742-6596/1650/3/032053

Ghaedi, M., Ansari, A., Bahari, F., Ghaedi, A.M., Vafaei, A., 2015. Spectrochimica Acta Part A : Molecular and Biomolecular Spectroscopy A hybrid artificial neural network and particle swarm optimization for prediction of removal of hazardous dye brilliant green from aqueous solution using zinc sulfide nanoparticle loaded. Spectrochim. ACTA PART A Mol. Biomol. Spectrosc. 137, 1004–1015. https://doi.org/10.1016/j.saa.2014.08.011

Ghaedi, M., Ghaedi, A.M., Ansari, A., Mohammadi, F., Vafaei, A., 2014. Spectrochimica Acta Part A : Molecular and Biomolecular Spectroscopy Artificial neural network and particle swarm optimization for removal of methyl orange by gold nanoparticles loaded on activated carbon and Tamarisk. Spectrochim. ACTA PART A Mol. Biomol. Spectrosc. 132, 639–654. https://doi.org/10.1016/j.saa.2014.04.175

Ghandehari, S., Montazer-Rahmati, M.M., Asghari, M., 2011. A comparison between semi-theoretical and empirical modeling of cross-flow microfiltration using ANN. Desalination 277, 348–355. https://doi.org/10.1016/j.desal.2011.04.057

Ghazali, A., Shirani, M., Semnani, A., Zare-Shahabadi, V., Nekoeinia, M., 2018. Optimization of crystal violet adsorption onto Date palm leaves as a potent biosorbent from aqueous solutions using response surface methodology and ant colony. J. Environ. Chem. Eng. 6, 3942–3950. https://doi.org/10.1016/j.jece.2018.05.043

Gheytaspour, M., Bigdarvish, O.H., 2018. Forecasting Oxygen Demand in Treatment Plant Using Artificial Neural Networks. Int. J. Adv. Eng. Res. Sci. 5, 50–57. https://doi.org/10.22161/ijaers.5.3.8

Ghorbani Mooselu, M., Nikoo, M.R., Latifi, M., Sadegh, M., Al-Wardy, M., Al-Rawas, G.A., 2020. A multi-objective optimal allocation of treated wastewater in urban areas using leader-follower game. J. Clean. Prod. 267. https://doi.org/10.1016/j.jclepro.2020.122189

Godini, K., Azarian, G., Kimiaei, A., Dragoi, E.N., Curteanu, S., 2021. Modeling of a real industrial wastewater treatment plant based on aerated lagoon using a neuro-evolutive technique. Process Saf. Environ. Prot. 148, 114–124. https://doi.org/10.1016/j.psep.2020.09.057

Godini, K., Azarian, G., Kimiaei, A., Niculina, E., Curteanu, S., 2020. l P re of. Process Saf. Environ. Prot. https://doi.org/10.1016/j.psep.2020.09.057

Golzar, F., Nilsson, D., Martin, V., 2020. Forecasting wastewater temperature based on artificial neural network (ANN) technique and Monte Carlo sensitivity analysis. Sustain. 12. https://doi.org/10.3390/SU12166386

Granata, F., Papirio, S., Esposito, G., Gargano, R., de Marinis, G., 2017. Machine learning algorithms for the forecasting of wastewater quality indicators. Water (Switzerland) 9, 1–12. https://doi.org/10.3390/w9020105

Güller, S., Silahtaroğlu, G., Akpolat, O., 2019. Communications in Statistics : Case Studies , Data Analysis and Applications Analysis waste water characteristics via data mining : A Muğla province case and external validation Analysis waste water characteristics via data mining : A ˘ province case and . Commun. Stat. Case Stud. Data Anal. Appl. 0, 1–10. https://doi.org/10.1080/23737484.2019.1604192

Guan, N., Liu, T., Zhang, Y., Tao, D., Davis, L.S., 2019. Truncated Cauchy Non-Negative Matrix Factorization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41, 1, 246-259, doi: 10.1109/TPAMI.2017.2777841.

Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y.M., Park, J. pyo, Kim, J.H., Cho, K.H., 2015. Prediction of effluent concentration in a wastewater treatment plant using machine learning models. J. Environ. Sci. (China) 32, 90–101. https://doi.org/10.1016/j.jes.2015.01.007

Gupta, S.K., Ansari, F.A., Nasr, M., Rawat, I., Nayunigari, M.K., Bux, F., 2017. Cultivation of Chlorella sorokiniana and Scenedesmus obliquus in wastewater: Fuzzy intelligence for evaluation of growth parameters and metabolites extraction. J. Clean. Prod. 147, 419–430. https://doi.org/10.1016/j.jclepro.2017.01.144

Haghiri, S., Daghighi, A., Moharramzadeh, S., 2018. Optimum coagulant forecasting by modeling jar test experiments using ANNs. Drink. Water Eng. Sci. 11, 1–8. https://doi.org/10.5194/dwes-11-1-2018

Hamed, M.M., Khalafallah, M.G., Hassanien, E.A., 2004. Prediction of wastewater treatment plant performance using artificial neural networks. Environ. Model. Softw. 19, 919–928. https://doi.org/10.1016/j.envsoft.2003.10.005

Han, Honggui, Liu, Z., Ge, L., Qiao, J., 2018a. Prediction of sludge bulking using the knowledge-leverage-based fuzzy neural network. Water Sci. Technol. 77, 617–627. https://doi.org/10.2166/wst.2017.574

Han, Hong-gui, Zhang, L., Liu, H., Qiao, J., 2018. Multiobjective design of fuzzy neural network controller for wastewater treatment process. Appl. Soft Comput. J. 67, 467–478. https://doi.org/10.1016/j.asoc.2018.03.020

Han, Honggui, Zhang, S., Qiao, J., Wang, X., 2018b. An intelligent detecting system for permeability prediction of MBR. Water Sci. Technol. 77, 467–478. https://doi.org/10.2166/wst.2017.562

Han, H.G., Qiao, J.F., 2011. Adaptive dissolved oxygen control based on dynamic structure neural network. Appl. Soft Comput. J. 11, 3812–3820. https://doi.org/10.1016/j.asoc.2011.02.014

Hashimoto, S., Matsukami, H., Ieda, T., Suzuki, G., 2021. Comprehensive screening of polybromochlorodibenzo-p-dioxins, dibenzofurans as mixed halogenated compounds in wastewater samples from industrial facilities by GC×GC/ToFMS and post-data processing. Chemosphere 276, 130085. https://doi.org/10.1016/j.chemosphere.2021.130085

Hayder, G., Ramli, M.Z., Malek, M.A., Khamis, A., Hilmin, N.M., 2014. Prediction model development for petroleum refinery wastewater treatment. J. Water Process Eng. 4, 1–5. https://doi.org/10.1016/j.jwpe.2014.08.006

Heddam, S., Bermad, A., Dechemi, N., 2012. ANFIS-based modelling for coagulant dosage in drinking water treatment plant: A case study. Environ. Monit. Assess. 184, 1953–1971. https://doi.org/10.1007/s10661-011-2091-x

Heo, S., Nam, K., Loy-Benitez, J., Yoo, C., 2020. Data-Driven Hybrid Model for Forecasting Wastewater Influent Loads Based on Multimodal and Ensemble Deep Learning. IEEE Trans. Ind. Informatics 3203, 1–10. https://doi.org/10.1109/TII.2020.3039272

Heo, S., Nam, K., Tariq, S., Lim, J.Y., Park, J., Yoo, C., 2021. A hybrid machine learning–based multi-objective supervisory control strategy of a full-scale wastewater treatment for cost-effective and sustainable operation under varying influent conditions. J. Clean. Prod. 291, 125853. https://doi.org/10.1016/j.jclepro.2021.125853

Hocaoglu, S.M., 2017. Evaluations of on-site wastewater reuse alternatives for hotels through water balance. Resour. Conserv. Recycl. 122, 43–50. https://doi.org/10.1016/j.resconrec.2017.01.022

Holubar, P., Zani, L., Hager, M., Fr, W., Radak, Z., Braun, R., 2002. Advanced controlling of anaerobic digestion by means of hierarchical neural networks 36, 2582–2588.

Hsu, L.C., 2009. Forecasting the output of integrated circuit industry using genetic algorithm based multivariable grey optimization models. Expert Syst. Appl. 36, 7898–7903. https://doi.org/10.1016/j.eswa.2008.11.004

Huang, M., Han, W., Wan, J., Ma, Y., Chen, X., 2016. Multi-objective optimisation for design and operation of anaerobic digestion using GA-ANN and NSGA-II. J. Chem. Technol. Biotechnol. 91, 226–233. https://doi.org/10.1002/jctb.4568

Huang, M., Ma, Y., Wan, J., Chen, X., 2015. A sensor-software based on a genetic algorithm-based neural fuzzy system for modeling and simulating a wastewater treatment process. Appl. Soft Comput. J. 27, 1–10. https://doi.org/10.1016/j.asoc.2014.10.034

Huang, M., Wan, J., Ma, Y., Zhang, H., Wang, Y., Wei, C., Liu, H., Yoo, C., 2011. A GA-based neural fuzzy system for modeling a paper mill wastewater treatment process. Ind. Eng. Chem. Res. 50, 13500–13507. https://doi.org/10.1021/ie201296p

Huang, M.Z., Wan, J.Q., Ma, Y.W., Li, W.J., Sun, X.F., Wan, Y., 2010. A fast predicting neural fuzzy model for on-line estimation of nutrient dynamics in an anoxic/oxic process. Bioresour. Technol. 101, 1642–1651. https://doi.org/10.1016/j.biortech.2009.08.111

Huang, Y., Li, T., Zheng, S., Fan, L., Su, L., Zhao, Y., Xie, H. Bin, Li, C., 2020. QSAR modeling for the ozonation of diverse organic compounds in water. Sci. Total Environ. 715, 136816. https://doi.org/10.1016/j.scitotenv.2020.136816

Huang, Z., Luo, J., Li, X., Zhou, Y., 2009a. Prediction of effluent parameters of wastewater treatment plant based on improved least square support vector machine with PSO. 2009 1st Int. Conf. Inf. Sci.

Eng. ICISE 2009 4058–4061. https://doi.org/10.1109/ICISE.2009.846

Huang, Z., Luo, J., Li, X., Zhou, Y., 2009b. Prediction of effluent parameters of wastewater treatment plant based on improved least square support vector machine with PSO 4058–4061.

Huang, Z., Wu, R., Yi, X., Liu, H., Cai, J., Niu, G., Huang, M., Ying, G., Baños, R., 2019. A Novel Model with GA Evolving FWNN for Effluent Quality and Biogas Production Forecast in a Full-Scale Anaerobic Wastewater Treatment Process. Complexity 2019. https://doi.org/10.1155/2019/2468189

Huang, S.; Zhang, H.; Pižurica, A. 2017. A Robust Sparse Representation Model for Hyperspectral Image Classification. Sensors, 17, 2087. https://doi.org/10.3390/s17092087.

Hwangbo, S., Al, R., Chen, X., Sin, G., 2021. Integrated Model for Understanding N2O Emissions from Wastewater Treatment Plants: A Deep Learning Approach. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.0c05231

Iqbal, J., Guria, C., 2009. Optimization of an operating domestic wastewater treatment plant using elitist non-dominated sorting genetic algorithm. Chem. Eng. Res. Des. 87, 1481–1496. https://doi.org/10.1016/j.cherd.2009.04.010

Jiang, W., Xing, X., Li, S., Zhang, X., Wang, W., 2019. Synthesis, characterization and machine learning based performance prediction of straw activated carbon. J. Clean. Prod. 212, 1210–1223. https://doi.org/10.1016/j.jclepro.2018.12.093

Jun, L.Y., Karri, R.R., Mubarak, N.M., Yon, L.S., Bing, C.H., Khalid, M., Jagadish, P., Abdullah, E.C., 2020a. Modelling of methylene blue adsorption using peroxidase immobilized functionalized Buckypaper/polyvinyl alcohol membrane via ant colony optimization. Environ. Pollut. 259, 113940. https://doi.org/10.1016/j.envpol.2020.113940

Jun, L.Y., Karri, R.R., Yon, L.S., Mubarak, N.M., Bing, C.H., Mohammad, K., Jagadish, P., Abdullah, E.C., 2020b. Modeling and optimization by particle swarm embedded neural network for adsorption of methylene blue by jicama peroxidase immobilized on buckypaper/polyvinyl alcohol membrane. Environ. Res. 183, 109158. https://doi.org/10.1016/j.envres.2020.109158

Jung, Y.-J., Kim, Y.-J., 2016. Relationship between Energy Consumption and Operational Variables at Wastewater Treatment Plant. J. Korean Soc. Water Environ. 32, 253–260. https://doi.org/10.15681/kswe.2016.32.3.253

K, A., Mungray, A., Agarwal, S., Ali, J., Chandra Garg, M., 2021. Performance optimisation of forward-osmosis membrane system using machine learning for the treatment of textile industry wastewater. J. Clean. Prod. 289, 125690. https://doi.org/10.1016/j.jclepro.2020.125690

Karimi, A. R.; Mehrdadi. N.; Hashemian, S. J.; Nabi Bidhendi, G. R.; Tavakkoli Moghaddam, R., 2011. Selection of Wastewater Treatment Process Based on Analytical Hierarchy Process. Int. J. Civ. Eng. Technol. 5, 27–33.

Karri, R.R., Sahu, J.N., Meikap, B.C., 2020. Improving efficacy of Cr (VI) adsorption process on sustainable adsorbent derived from waste biomass (sugarcane bagasse) with help of ant colony optimization. Ind. Crops Prod. 143, 111927. https://doi.org/10.1016/j.indcrop.2019.111927

Kazemi, P., Bengoa, C., Steyer, J.P., Giralt, J., 2021. Data-driven techniques for fault detection in anaerobic digestion process. Process Saf. Environ. Prot. 146, 905–915. https://doi.org/10.1016/j.psep.2020.12.016

Khan, S.U., Khan, H., Anwar, S., Khan, S., Boldrin Zanoni, M. V., Hussain, S., 2020. Computational and statistical modeling for parameters optimization of electrochemical decontamination of synozol red dye wastewater. Chemosphere 253, 126673. https://doi.org/10.1016/j.chemosphere.2020.126673

Khawaga, R.I., Abdel Jabbar, N., Al-Asheh, S., Abouleish, M., 2019. Model identification and control of chlorine residual for disinfection of wastewater. J. Water Process Eng. 32, 100936. https://doi.org/10.1016/j.jwpe.2019.100936

Kundu, P., Debsarkar, A., Mukherjee, S., 2013. Artificial Neural Network Modeling for Biological Removal of Organic Carbon and Nitrogen from Slaughterhouse Wastewater in a Sequencing Batch Reactor. Adv. Artif. Neural Syst. 2013, 1–15. https://doi.org/10.1155/2013/268064

Kusiak, A., Verma, A., Wei, X., 2013. A data-mining approach to predict influent quality 2197–2210. https://doi.org/10.1007/s10661-012-2701-2

Lariche, M.J., Soltani, Shahrzad, Davoudi Nezhad, H., Moradi, H., Soltani, Sheyda, Farsayad, F., Moradi Kazerouni, H., 2020. Developing supervised models for estimating methylene blue removal by silver nanoparticles. Energy Sources, Part A Recover. Util. Environ. Eff. 42, 1247–1254. https://doi.org/10.1080/15567036.2019.1602233

Lee, T.M., Oh, H., Choung, Y.K., Oh, S., Jeon, M., Kim, J.H., Nam, S.H., Lee, S., 2009. Prediction of membrane fouling in the pilot-scale microfiltration system using genetic programming. Desalination 247, 285–294. https://doi.org/10.1016/j.desal.2008.12.031

Li, X., Yi, X., Liu, Z., Liu, H., Chen, T., Niu, G., Yan, B., Chen, C., Huang, M., Ying, G., 2021. Application of novel hybrid deep leaning model for cleaner production in a paper industrial wastewater treatment system. J. Clean. Prod. 294, 126343. https://doi.org/10.1016/j.jclepro.2021.126343

Li, Z., Yan, X., 2019. Ensemble model of wastewater treatment plant based on rich diversity of principal component determining by genetic algorithm for status monitoring. Control Eng. Pract. 88, 38–51. https://doi.org/10.1016/j.conengprac.2019.04.008

Libotean, D., Giralt, J., Giralt, F., Rallo, R., Wolfe, T., Cohen, Y., 2009. Neural network approach for modeling the performance of reverse osmosis membrane desalting. J. Memb. Sci. 326, 408–419. https://doi.org/10.1016/j.memsci.2008.10.028

Liu, H., Yang, C., Huang, M., Yoo, C.K., 2020. Soft sensor modeling of industrial process data using kernel latent variables-based relevance vector machine. Appl. Soft Comput. J. 90, 106149. https://doi.org/10.1016/j.asoc.2020.106149

Liu, P., 2020. Jou rna lP. Econ. Lett. 109008. https://doi.org/10.1016/j.asoc.2021.107227

Liu, Z.J., Wan, J.Q., Ma, Y.W., Wang, Y., 2019. Online prediction of effluent COD in the anaerobic wastewater treatment system based on PCA-LSSVM algorithm. Environ. Sci. Pollut. Res. 26, 12828–12841. https://doi.org/10.1007/s11356-019-04671-8

López, M.E., Rene, E.R., Boger, Z., Veiga, M.C., Kennes, C., 2017. Modelling the removal of volatile pollutants under transient conditions in a two-stage bioreactor using artificial neural networks. J. Hazard. Mater. 324, 100–109. https://doi.org/10.1016/j.jhazmat.2016.03.018

Lotfi, K., Bonakdari, H., Ebtehaj, I., Delatolla, R., Zinatizadeh, A.A., Gharabaghi, B., 2020. A novel stochastic wastewater quality modeling based on fuzzy techniques. J. Environ. Heal. Sci. Eng. 18, 1099–1120. https://doi.org/10.1007/s40201-020-00530-8

Lu, H., Li, H., Liu, T., Fan, Y., Yuan, Y., Xie, M., Qian, X., 2019. Simulating heavy metal concentrations in an aquatic environment using artificial intelligence models and physicochemical indexes. Sci.

Total Environ. 694, 133591. https://doi.org/10.1016/j.scitotenv.2019.133591

Ma, Y., Huang, M., Wan, J., Wang, Y., Sun, X., Zhang, H., 2011. Prediction model of DnBP degradation based on BP neural network in AAO system. Bioresour. Technol. 102, 4410–4415. https://doi.org/10.1016/j.biortech.2011.01.004

Mahata, C., Ray, S., Das, D., 2020a. Optimization of dark fermentative hydrogen production from organic wastes using acidogenic mixed consortia. Energy Convers. Manag. 219, 113047. https://doi.org/10.1016/j.enconman.2020.113047

Mahata, C., Ray, S., Das, D., 2020b. Optimization of dark fermentative hydrogen production from organic wastes using acidogenic mixed consortia. Energy Convers. Manag. 219, 113047. https://doi.org/10.1016/j.enconman.2020.113047

Mahjouri, M., Ishak, M.B., Torabian, A., Abd Manaf, L., Halimoon, N., Ghoddusi, J., 2017a. Optimal selection of Iron and Steel wastewater treatment technology using integrated multi-criteria decision-making techniques and fuzzy logic. Process Saf. Environ. Prot. 107, 54–68. https://doi.org/10.1016/j.psep.2017.01.016

Mahjouri, M., Ishak, M.B., Torabian, A., Manaf, L.A., Halimoon, N., 2017b. The application of a hybrid model for identifying and ranking indicators for assessing the sustainability of wastewater treatment systems. Sustain. Prod. Consum. 10, 21–37. https://doi.org/10.1016/j.spc.2016.09.006

Mahmoodi-Babolan, N., Heydari, A., Nematollahzadeh, A., 2019. Removal of methylene blue via bioinspired catecholamine/starch superadsorbent and the efficiency prediction by response surface methodology and artificial neural network-particle swarm optimization. Bioresour. Technol. 294, 122084. https://doi.org/10.1016/j.biortech.2019.122084

Mahmoodi, N.M., Keshavarzi, S., Ghezelbash, M., 2017. Synthesis of nanoparticle and modelling of its photocatalytic dye degradation ability from colored wastewater. J. Environ. Chem. Eng. 5, 3684–3689. https://doi.org/10.1016/j.jece.2017.07.010

Mahshidnia, M., Jafarian, A., 2016. Forecasting Wastewater Treatment Results with an ANFIS Intelligent System. Eng. Technol. Appl. Sci. Res. 6, 1175–1181. https://doi.org/10.48084/etasr.745

Majhi, B., Panda, G., 2011. Robust identification of nonlinear complex systems using low complexity ANN and particle swarm optimization technique. Expert Syst. Appl. 38, 321–333. https://doi.org/10.1016/j.eswa.2010.06.070

Mamandipoor, B., Majd, M., Sheikhalishahi, S., Modena, C., Osmani, V., 2020. Monitoring and detecting faults in wastewater treatment plants using deep learning. Environ. Monit. Assess. 192. https://doi.org/10.1007/s10661-020-8064-1

Mandal, S., Mahapatra, S.S., Sahu, M.K., Patel, R.K., 2015. Artificial neural network modelling of As(III) removal from water by novel hybrid material. Process Saf. Environ. Prot. 93, 249–264. https://doi.org/10.1016/j.psep.2014.02.016

Manwani, N., Sastry, P. S., 2013. Noise Tolerance Under Risk Minimization, IEEE Transactions Cybernetics, 43, 1146-1151

Maria, R., Alves, D.B., Augusto, C., 2009. Optimal Control Law Development in a Sequential Batch Reactor T hrough Mixed Integer Particle Swarm Dynamic Optimization. Elsevier Inc. https://doi.org/10.1016/S1570-7946(09)70681-9

Marsili-Libelli, S., Giunti, L., 2002. Fuzzy predictive control for nitrogen removal in biological wastewater treatment. Water Sci. Technol. 45, 37–44. https://doi.org/10.2166/wst.2002.0546

Mazhar, S., Ditta, A., Bulgariu, L., Ahmad, I., Ahmed, M., Nadiri, A.A., 2019. Sequential treatment of paper and pulp industrial wastewater: Prediction of water quality parameters by Mamdani Fuzzy Logic model and phytotoxicity assessment. Chemosphere 227, 256–268. https://doi.org/10.1016/j.chemosphere.2019.04.022

Mei, Y., Yang, J., Lu, Y., Hao, F., Xu, D., Pan, H., Wang, J., 2019. BP–ANN model coupled with particle swarm optimization for the efficient prediction of 2-chlorophenol removal in an electro-oxidation system. Int. J. Environ. Res. Public Health 16. https://doi.org/10.3390/ijerph16142454

Messikh, N., Chiha, M., Ahmedchekkat, F., Al Bsoul, A., 2015. Application of radial basis function neural network for removal of copper using an emulsion liquid membrane process assisted by ultrasound. Desalin. Water Treat. 56, 399–408. https://doi.org/10.1080/19443994.2014.936513

Mjalli, F.S., Al-Asheh, S., Alfadala, H.E., 2007. Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. J. Environ. Manage. 83, 329–338. https://doi.org/10.1016/j.jenvman.2006.03.004

Mohadesi, B.A.A.R.M., 2019. Modeling and prediction of water quality parameters using a hybrid particle swarm optimization – neural fuzzy approach. Int. J. Environ. Sci. Technol. 16, 4823–4832. https://doi.org/10.1007/s13762-018-1896-3

Mohammad, A.T., Al-Obaidi, M.A., Hameed, E.M., Basheer, B.N., Mujtaba, I.M., 2020. Modelling the chlorophenol removal from wastewater via reverse osmosis process using a multilayer artificial neural network with genetic algorithm. J. Water Process Eng. 33, 100993. https://doi.org/10.1016/j.jwpe.2019.100993

Mohammadi, F., Bina, B., Karimi, H., Rahimi, S., Yavari, Z., 2020. Modeling and sensitivity analysis of the alkylphenols removal via moving bed biofilm reactor using artificial neural networks: Comparison of levenberg marquardt and particle swarm optimization training algorithms. Biochem. Eng. J. 161, 107685. https://doi.org/10.1016/j.bej.2020.107685

Mohan, S., Singh, Y., Verma, D.K., Hasan, S.H., 2015. Synthesis of CuO nanoparticles through green route using Citrus limon juice and its application as nanosorbent for Cr(VI) remediation: Process optimization with RSM and ANN-GA based model, Process Safety and Environmental Protection. Institution of Chemical Engineers. https://doi.org/10.1016/j.psep.2015.05.005

Mojiri, A., Ohashi, A., Ozaki, N., Aoi, Y., Kindaichi, T., 2020. Integrated anammox-biochar in synthetic wastewater treatment: Performance and optimization by artificial neural network. J. Clean. Prod. 243, 118638. https://doi.org/10.1016/j.jclepro.2019.118638

Mokhtari Nesfchi, M., Ebrahimian Pirbazari, A., Khalil Saraei, F.E., Rojaee, F., Mahdavi, F., Fa'al Rastegar, S.A., 2021. Fabrication of plasmonic nanoparticles / cobalt doped TiO2 nanosheets for degradation of tetracycline and modeling the process by artificial intelligence techniques. Mater. Sci. Semicond. Process. 122, 105465. https://doi.org/10.1016/j.mssp.2020.105465

Moral, H., Aksoy, A., Gokcay, C.F., 2008. Modeling of the activated sludge process by using artificial neural networks with automated architecture screening. Comput. Chem. Eng. 32, 2471–2478. https://doi.org/10.1016/j.compchemeng.2008.01.008

Murnleitner, E., Becker, T.M., Delgado, A., 2002. State detection and control of overloads in the anaerobic wastewater treatment using fuzzy logic. Water Res. 36, 201–211. https://doi.org/10.1016/S0043-1354(01)00186-5

Nadiri, A.A., Shokri, S., Tsai, F.T.C., Asghari Moghaddam, A., 2018. Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model. J.

Clean. Prod. 180, 539–549. https://doi.org/10.1016/j.jclepro.2018.01.139

Nair, V. V., Dhar, H., Kumar, S., Thalla, A.K., Mukherjee, S., Wong, J.W.C., 2016. Artificial neural network based modeling to evaluate methane yield from biogas in a laboratory-scale anaerobic bioreactor. Bioresour. Technol. 217, 90–99. https://doi.org/10.1016/j.biortech.2016.03.046

Najafzadeh, M., Zeinolabedini, M., 2019. Prognostication of waste water treatment plant performance using efficient soft computing models: An environmental evaluation. Meas. J. Int. Meas. Confed. 138, 690–701. https://doi.org/10.1016/j.measurement.2019.02.014

Najah Ahmed, A., Binti Othman, F., Abdulmohsin Afan, H., Khaleel Ibrahim, R., Ming Fai, C., Shabbir Hossain, M., Ehteram, M., Elshafie, A., 2019. Machine learning methods for better water quality prediction. J. Hydrol. 578. https://doi.org/10.1016/j.jhydrol.2019.124084

Nasr, M.S., Moustafa, M.A.E., Seif, H.A.E., El Kobrosy, G., 2012. Application of artificial neural network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. Alexandria Eng. J. 51, 37–43. https://doi.org/10.1016/j.aej.2012.07.005

Nassef, A.M., Rezk, H., Abdelkareem, M.A., Alaswad, A., Olabi, A., 2019. Application of fuzzy modelling and Particle Swarm Optimization to enhance lipid extraction from microalgae. Sustain. Energy Technol. Assessments 35, 73–79. https://doi.org/10.1016/j.seta.2019.06.003

Nawaz, A., Arora, A.S., Yun, C.M., Cho, H., You, S., Lee, M., 2019. Data Authorization and Forecasting by a Proactive Soft Sensing Tool-Anammox Based Process. Ind. Eng. Chem. Res. 58, 9552–9563. https://doi.org/10.1021/acs.iecr.9b00722

Newhart, K.B., Holloway, R.W., Hering, A.S., Cath, T.Y., 2019. Data-driven performance analyses of wastewater treatment plants: A review. Water Res. 157, 498–513. https://doi.org/10.1016/j.watres.2019.03.030

Niu, G., Yi, X., Chen, C., Li, X., Han, D., Yan, B., 2020a. A novel ef fl uent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. J. Clean. Prod. 265, 121787. https://doi.org/10.1016/j.jclepro.2020.121787

Niu, G., Yi, X., Chen, C., Li, X., Han, D., Yan, B., Huang, M., Ying, G., 2020b. A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. J. Clean. Prod. 265, 121787. https://doi.org/10.1016/j.jclepro.2020.121787

Onu, C.E., Nwabanne, J.T., Ohale, P.E., Asadu, C.O., 2021. Comparative analysis of RSM, ANN and ANFIS and the mechanistic modeling in eriochrome black-T dye adsorption using modified clay. South African J. Chem. Eng. 36, 24–42. https://doi.org/10.1016/j.sajce.2020.12.003

Otchere, D.A., Arbi Ganat, T.O., Gholami, R., Ridha, S., 2021. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. J. Pet. Sci. Eng. 108182. https://doi.org/10.1016/j.petrol.2020.108182

Pang, J., Yang, S., He, L., Chen, Y., Ren, N., 2019. Intelligent control/operational strategies in WWTPs through an integrated Q-learning algorithm with ASM2d-guided reward. Water (Switzerland) 11. https://doi.org/10.3390/w11050927

Park, S., Kim, K., Shin, C., Min, J.H., Na, E.H., Park, L.J., 2020. Variable update strategy to improve water quality forecast accuracy in multivariate data assimilation using the ensemble Kalman filter. Water Res. 176, 115711. https://doi.org/10.1016/j.watres.2020.115711

Parthasarathy, P., Narayanan, S.K., 2014. Effect of Hydrothermal Carbonization Reaction Parameters on.

Environ. Prog. Sustain. Energy 33, 676–680. https://doi.org/10.1002/ep

Patel, N., Ruparelia, J., Barve, J., 2020. Prediction of total suspended solids present in effluent of primary clarifier of industrial common effluent treatment plant: Mechanistic and fuzzy approach. J. Water Process Eng. 34, 101146. https://doi.org/10.1016/j.jwpe.2020.101146

Pelalak, R., Heidari, Z., Alizadeh, R., Ghareshabani, E., Nasseh, N., Marjani, A., Albadarin, A.B., Shirazian, S., 2021. Efficient oxidation/mineralization of pharmaceutical pollutants using a novel Iron (III) oxyhydroxide nanostructure prepared via plasma technology: Experimental, modeling and DFT studies. J. Hazard. Mater. 411, 125074. https://doi.org/10.1016/j.jhazmat.2021.125074

Picos-benítez, A.R., Martínez-vargas, B.L., Duron-torres, S.M., Brillas, E., Peralta-hernández, J.M., 2020. The use of artificial intelligence models in the prediction of optimum operational conditions for the treatment of dye wastewaters with similar structural characteristics 143, 36–44. https://doi.org/10.1016/j.psep.2020.06.020

Picos-Benítez, A.R., Martínez-Vargas, B.L., Duron-Torres, S.M., Brillas, E., Peralta-Hernández, J.M., 2020. The use of artificial intelligence models in the prediction of optimum operational conditions for the treatment of dye wastewaters with similar structural characteristics. Process Saf. Environ. Prot. 143, 36–44. https://doi.org/10.1016/j.psep.2020.06.020

Piotrowski, R., Paul, A., Lewandowski, M., 2019. Improving SBR performance alongside with cost reduction through optimizing biological processes and dissolved oxygen concentration trajectory. Appl. Sci. 9. https://doi.org/10.3390/app9112268

Pires, O.C., Palma, C., Moita, I., Costa, J.C., Alves, M.M., Ferreira, E.C., 2005. A Fuzzy-Logic Based Expert System for Diagnosis and Control of an Integrated Wastewater Treatment. Enpromer 1–8.

Podder, M.S., Majumder, C.B., 2016. The use of artificial neural network for modelling of phycoremediation of toxic elements As(III) and As(V) from wastewater using Botryococcus braunii. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 155, 130–145. https://doi.org/10.1016/j.saa.2015.11.011

Publishing, I.W.A., Science, W., 2012. A data-driven model for maximization of methane production in a wastewater treatment plant Andrew Kusiak and Xiupeng Wei 1116–1122. https://doi.org/10.2166/wst.2012.953

Qi, J., Hou, Y., Hu, J., Ruan, W., Xiang, Y., Wei, X., 2020a. Decontamination of methylene Blue from simulated wastewater by the mesoporous rGO/Fe/Co nanohybrids: Artificial intelligence modeling and optimization. Mater. Today Commun. 24, 100709. https://doi.org/10.1016/j.mtcomm.2019.100709

Qi, J., Hou, Y., Hu, J., Ruan, W., Xiang, Y., Wei, X., 2020b. Decontamination of methylene Blue from simulated wastewater by the mesoporous rGO / Fe / Co nanohybrids : Arti fi cial intelligence modeling and optimization. Mater. Today Commun. 24, 100709. https://doi.org/10.1016/j.mtcomm.2019.100709

Qiao, Z., Wang, Z., Zhang, C., Yuan, S., Zhu, Y., Wang, J., 2012. PVAm–PIP/PS composite membrane with high performance for CO_2/N_2 separation. AIChE J. 59, 215–228. https://doi.org/10.1002/aic

Qiu, Y., Li, J., Huang, X., Shi, H., 2018. A feasible data-driven mining system to optimize wastewater treatment process design and operation. Water (Switzerland) 10. https://doi.org/10.3390/w10101342

Rafati, M., Pazouki, M., Ghadamian, H., Jalilzadeh, A., 2021. Determine the most effective process control parameters on activated sludge based on particle swarm optimisation algorithm (Case Study : South wastewater treatment plant of Tehran). Int. J. Environ. Anal. Chem. 00, 1–22.

https://doi.org/10.1080/03067319.2020.1836169

Rahimian, P., Behnam, S., 2020. A novel data-driven and feature-based forecasting framework for wastewater optimization of network pressure management system. Int. J. Ind. Eng. Prod. Res. 31, 423–433. https://doi.org/10.22068/ijiepr.31.3.423

Rahmani, A., Salari, M., Tari, K., Shabanloo, A., Shabanloo, N., Bajalan, S., 2020. Enhanced degradation of furfural by heat-activated persulfate/nZVI-rGO oxidation system: Degradation pathway and improving the biodegradability of oil refinery wastewater. J. Environ. Chem. Eng. 8, 104468. https://doi.org/10.1016/j.jece.2020.104468

Ranade, N. V., Nagarajan, S., Sarvothaman, V., Ranade, V. V., 2021. ANN based modelling of hydrodynamic cavitation processes: Biomass pre-treatment and wastewater treatment. Ultrason. Sonochem. 72, 105428. https://doi.org/10.1016/j.ultsonch.2020.105428

Rashidi, M.M., Ali, M., Freidoonimehr, N., Nazari, F., 2013. Parametric analysis and optimization of entropy generation in unsteady MHD fl ow over a stretching rotating disk using arti fi cial neural network and particle swarm optimization algorithm. Energy 55, 497–510. https://doi.org/10.1016/j.energy.2013.01.036

Rastegaripour, F., Saboni, M.S., Shojaei, S., Tavassoli, A., 2019. Simultaneous management of water and wastewater using ant and artificial neural network (ANN) algorithms. Int. J. Environ. Sci. Technol. 16, 5835–5856. https://doi.org/10.1007/s13762-018-1943-0

Razvarz, S., Jafari, R., Yu, W., Golmankhaneh, A.K., 2017. PSO and NN modeling for photocatalytic removal of pollution in wastewater. 2017 14th Int. Conf. Electr. Eng. Comput. Sci. Autom. Control. CCE 2017 23. https://doi.org/10.1109/ICEEE.2017.8108825

Reis, I., Baron, D., Shahaf, S., 2018. Probabilistic Random Forest: A Machine Learning Algorithm for Noisy Data Sets. The American Astronomical Society. 157, 16. https://doi.org/10.3847/1538-3881/aaf101

Revollar, S., Álvarez, H., Lamanna, R., Vega, P., Goldar, A., 2018. Economic Model Predictive Control of a Wastewater Treatment Plant using Neural and Fuzzy models, Computer Aided Chemical Engineering. Elsevier Masson SAS. https://doi.org/10.1016/B978-0-444-64235-6.50215-1

Rolnick, D., Veit, A., Belongie, S., Shavit, N., 2018. Deep Learning is Robust to Massive Label Noise. Cornell University, Computer Science - Machine Learning.

Ruano, M. V., Ribes, J., Ferrer, J., Sin, G., 2011. Application of the Morris method for screening the influential parameters of fuzzy controllers applied to wastewater treatment plants. Water Sci. Technol. 63, 2199–2206. https://doi.org/10.2166/wst.2011.442

Rustum, R., Adeloye, A., 2012. Improved Modelling of Wastewater Treatment Primary Clarifier Using Hybrid Anns. Int. J. Comput. Sci. Artif. Intell. 2, 14–22. https://doi.org/10.5963/ijcsai0204002

Rustum, R., Adeloye, A.J., Scholz, M., 2008. Applying Kohonen Self-Organizing Map as a Software Sensor to Predict Biochemical Oxygen Demand. Water Environ. Res. 80, 32–40. https://doi.org/10.2175/106143007x184500

Sabour, M.R., Amiri, A., 2017. Comparative study of ANN and RSM for simultaneous optimization of multiple targets in Fenton treatment of landfill leachate. Waste Manag. https://doi.org/10.1016/j.wasman.2017.03.048

Santín, I., Barbu, M., Pedret, C., Vilanova, R., 2018. Fuzzy logic for plant-wide control of biological wastewater treatment process including greenhouse gas emissions. ISA Trans. 77, 146–166.

https://doi.org/10.1016/j.isatra.2018.04.006

Santín, I., Pedret, C., Vilanova, R., 2015a. Applying variable dissolved oxygen set point in a two level hierarchical control structure to a wastewater treatment process. J. Process Control 28, 40–55. https://doi.org/10.1016/j.jprocont.2015.02.005

Santín, I., Pedret, C., Vilanova, R., 2015b. Fuzzy Control and Model Predictive Control Configurations for Effluent Violations Removal in Wastewater Treatment Plants. Ind. Eng. Chem. Res. 54, 2763–2775. https://doi.org/10.1021/ie504079q

Sattar, A.A., Elhakeem, M., Rezaie-Balf, M., Gharabaghi, B., Bonakdari, H., 2019. Artificial intelligence models for prediction of the aeration efficiency of the stepped weir. Flow Meas. Instrum. 65, 78–89. https://doi.org/10.1016/j.flowmeasinst.2018.11.017

Sedaqatvand, R., Nasr Esfahany, M., Behzad, T., Mohseni, M., Mardanpour, M.M., 2013. Parameter estimation and characterization of a single-chamber microbial fuel cell for dairy wastewater treatment. Bioresour. Technol. 146, 247–253. https://doi.org/10.1016/j.biortech.2013.07.054

Seifi, A., Ehteram, M., Soroush, F., 2020. Uncertainties of instantaneous influent flow predictions by intelligence models hybridized with multi-objective shark smell optimization algorithm. J. Hydrol. 587, 124977. https://doi.org/10.1016/j.jhydrol.2020.124977

Sellami, A., Ali, M. Ben, Kairouani, L., 2020. Pareto-optimization of msf-ot/tvc desalination plant using surface response methodology and genetic algorithm. Desalin. Water Treat. 204, 10–21. https://doi.org/10.5004/dwt.2020.26229

Shamshirband, S., Jafari Nodoushan, E., Adolf, J.E., Abdul Manaf, A., Mosavi, A., Chau, K. wing, 2019. Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters. Eng. Appl. Comput. Fluid Mech. 13, 91–101. https://doi.org/10.1080/19942060.2018.1553742

Sharafati, A., Asadollah, S.B.H.S., Hosseinzadeh, M., 2020. The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty. Process Saf. Environ. Prot. 140, 68–78. https://doi.org/10.1016/j.psep.2020.04.045

Shen, W., Huang, F., Zhang, X., Zhu, Y., Chen, X., Akbarjon, N., 2018. On-line chemical oxygen demand estimation models for. Water Sci. Technol. 78, 310–319. https://doi.org/10.2166/wst.2018.299

Shi, S., Xu, G., 2018. Novel performance prediction model of a biofilm system treating domestic wastewater based on stacked denoising auto-encoders deep learning network. Chem. Eng. J. 347, 280–290. https://doi.org/10.1016/j.cej.2018.04.087

Shi, Y., Zhao, X., Zhang, Y., Ren, N., 2009. Back propagation neural network (BPNN) prediction model and control strategies of methanogen phase reactor treating traditional Chinese medicine wastewater (TCMW) 144, 70–74. https://doi.org/10.1016/j.jbiotec.2009.08.014

Shin, Y., Kim, Z., Yu, J., Kim, G., Hwang, S., 2019. Development of NOx reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA). J. Clean. Prod. 232, 1418–1429. https://doi.org/10.1016/j.jclepro.2019.05.276

Shokrkar, H., Salahi, A., Kasiri, N., Mohammadi, T., 2012. Prediction of permeation flux decline during MF of oily wastewater using genetic programming. Chem. Eng. Res. Des. 90, 846–853. https://doi.org/10.1016/j.cherd.2011.10.002

Singh, K.P., Gupta, S., 2012. Artificial intelligence based modeling for predicting the disinfection by-products in water. Chemom. Intell. Lab. Syst. 114, 122–131. https://doi.org/10.1016/j.chemolab.2012.03.014

Sivanantham, V., Narayana, P.L., Hyeong, K.J., Pareddy, P., Sangeetha, V., Kyoung – Seok, M., In, K.H., Sung, H.K., Reddy, N.S., 2021. Modeling and optimization of chlorophenol rejection for spiral wound reverse osmosis membrane modules. Chemosphere 268, 129345. https://doi.org/10.1016/j.chemosphere.2020.129345

Smarra, F., Jain, A., de Rubeis, T., Ambrosini, D., D'Innocenzo, A., Mangharam, R., 2018. Data-driven model predictive control using random forests for building energy optimization and climate control. Appl. Energy 226, 1252–1272. https://doi.org/10.1016/j.apenergy.2018.02.126

Song, M.J., Choi, S., Bae, W. Bin, Lee, J., Han, H., Kim, D.D., Kwon, M., Myung, J., Kim, Y.M., Yoon, S., 2020. Identification of primary effecters of N2O emissions from full-scale biological nitrogen removal systems using random forest approach. Water Res. 184, 116144. https://doi.org/10.1016/j.watres.2020.116144

Srinu Naik, S., Pydi Setty, Y., 2014. Optimization of parameters using response surface methodology and genetic algorithm for biological denitrification of wastewater. Int. J. Environ. Sci. Technol. 11, 823–830. https://doi.org/10.1007/s13762-013-0266-4

Struk-Sokołowska2020.pdf, n.d.

Su, Y., Yang, A., Jin, S., Shen, W., Cui, P., Ren, J., 2020. Investigation on ternary system tetrahydrofuran/ethanol/water with three azeotropes separation via the combination of reactive and extractive distillation. J. Clean. Prod. 273, 123145. https://doi.org/10.1016/j.jclepro.2020.123145

Suggala, S. V., Bhattacharya, P.K., 2003. Real coded genetic algorithm for optimization of pervaporation process parameters for removal of volatile organics from water. Ind. Eng. Chem. Res. 42, 3118–3128. https://doi.org/10.1021/ie020183t

Sulthana, A., Latha, K.C., Imran, M., Rathan, R., Sridhar, R., Balasubramanian, S., 2014. Non-linear modeling using fuzzy principal component regression for Vidyaranyapuram sewage treatment plant, Mysore - India. Water Sci. Technol. 70, 1040–1047. https://doi.org/10.2166/wst.2014.333

Sun, S., Bao, Z., Li, R., Sun, D., Geng, H., Huang, X., Lin, J., Zhang, P., Ma, R., Fang, L., Zhang, X., Zhao, X., 2017. Reduction and prediction of N2O emission from an Anoxic/Oxic wastewater treatment plant upon DO control and model simulation. Bioresour. Technol. 244, 800–809. https://doi.org/10.1016/j.biortech.2017.08.054

Szeląg2020.pdf, n.d.

Tanzifi, M., Tavakkoli Yaraki, M., Beiramzadeh, Z., Heidarpoor Saremi, L., Najafifard, M., Moradi, H., Mansouri, M., Karami, M., Bazgir, H., 2020. Carboxymethyl cellulose improved adsorption capacity of polypyrrole/CMC composite nanoparticles for removal of reactive dyes: Experimental optimization and DFT calculation. Chemosphere 255. https://doi.org/10.1016/j.chemosphere.2020.127052

Tashvigh, A.A., Nasernejad, B., 2017. Soft computing method for modeling and optimization of air and water gap membrane distillation — A genetic programming approach. Desalin. Water Treat. 76, 30–39. https://doi.org/10.5004/dwt.2017.20696

Taşkın, A., Zeybek, Z., Satar, B., Taşkın, S., 2016. For Domestic Wastewater Treatment, Finding Optimum Conditions by Particle Swarm Optimization and Experimental Design. International Journal of Environmental & Agriculture Research,2,1.

Tayyebi, S., Alishiri, M., 2014. The control of MSF desalination plants based on inverse model control by neural network. Desalination 333, 92–100. https://doi.org/10.1016/j.desal.2013.11.022

Tejaswini, E.S.S., Panjwani, S., Seshagiri Rao, A., 2020a. Design of hierarchical control strategies for biological wastewater treatment plants to reduce operational costs. Chem. Eng. Res. Des. 161, 197–205. https://doi.org/10.1016/j.cherd.2020.07.003

Tejaswini, E.S.S., Panjwani, S., Uday Bhaskar Babu, G., Seshagiri Rao, A., 2020b. Model based control of a full-scale biological wastewater treatment plant. IFAC-PapersOnLine 53, 208–213. https://doi.org/10.1016/j.ifacol.2020.06.036

Teychene, B., Chi, F., Chokki, J., Darracq, G., Baron, J., Joyeux, M., Gallard, H., 2020. Investigation of polar mobile organic compounds (PMOC) removal by reverse osmosis and nano fi ltration : rejection mechanism modelling using decision tree 975–983. https://doi.org/10.2166/ws.2020.020

Torregrossa, D., Hansen, J., Hernández-Sancho, F., Cornelissen, A., Schutz, G., Leopold, U., 2017. A data-driven methodology to support pump performance analysis and energy efficiency optimization in Waste Water Treatment Plants. Appl. Energy 208, 1430–1440. https://doi.org/10.1016/j.apenergy.2017.09.012

United Nations World Water Assessment Programme, 2017. Wastewater: The untapped resource, The United Nations World Water Development Report.

Van Horn, G., Branson, S., Farrell, R., Haber, S., Barry, J., Ipeirotis, P., Perona, P., Belongie, S.. 2015. Building a bird recognition app and large scale dataset with citizen scientists: The fine print in fine-grained dataset collection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 595-604, doi: 10.1109/CVPR.2015.7298658.

and Pattern Recognition, (CVPR), 2015.

Vasilaki, V., Danishvar, S., Mousavi, A., Katsou, E., 2020. Data-driven versus conventional N2O EF quantification methods in wastewater; how can we quantify reliable annual EFs? Comput. Chem. Eng. 141, 106997. https://doi.org/10.1016/j.compchemeng.2020.106997

Verma, A., Wei, X., Kusiak, A., 2013. Predicting the total suspended solids in wastewater: A data-mining approach. Eng. Appl. Artif. Intell. 26, 1366–1372. https://doi.org/10.1016/j.engappai.2012.08.015

Wang, C., Zhuang, Y., Liu, L., Zhang, L., Du, J., 2021. Design and comparison of energy-saving double column and triple column reactive-extractive hybrid distillation processes for ternary multi-azeotrope dehydration. Sep. Purif. Technol. 259, 118211. https://doi.org/10.1016/j.seppur.2020.118211

Wang, J., Wan, K., Gao, X., Cheng, X., Shen, Y., Wen, Z., Tariq, U., Piran, M.J., 2020. Energy and Materials-Saving Management via Deep Learning for Wastewater Treatment Plants. IEEE Access 8, 191694–191705. https://doi.org/10.1109/access.2020.3032531

Wang, Y., Li, J., Wang, S., Yang, J., Qi, C., Guo, H., Liu, X., Zhang, H., 2020. Operational optimization of wastewater reuse integrated energy system. Energy 200, 117309. https://doi.org/10.1016/j.energy.2020.117309

Wang, Y., Li, L., Wang, K., 2021. An online operating performance evaluation approach using probabilistic fuzzy theory for chemical processes with uncertainties 144. https://doi.org/10.1016/j.compchemeng.2020.107156

Water, I., Data, Q., 2019. Application of a Parallel Particle Swarm Optimization-Long Short Term Memory Model to 1–16.

Wei, X., Kusiak, A., 2012. Optimization of biogas production process in a wastewater treatment plant. 62nd IIE Annu. Conf. Expo 2012 1432–1440.

Wu, J., Cheng, H., Liu, Y., Huang, D., Yuan, L., Yao, L., 2020. Learning soft sensors using time difference–based multi-kernel relevance vector machine with applications for quality-relevant monitoring in wastewater treatment. Environ. Sci. Pollut. Res. 27, 28986–28999. https://doi.org/10.1007/s11356-020-09192-3

Xie, W., Zhang, R., Li, W., Ni, B., Fang, F., Sheng, G., Yu, H., 2011. Simulation and optimization of a full-scale Carrousel oxidation ditch plant for municipal wastewater treatment 56, 9–16. https://doi.org/10.1016/j.bej.2011.04.010

Xin, C., Shi, X., Wang, D., Yang, C., Li, Q., Liu, H., 2020. Multi-grained cascade forest for effluent quality prediction of papermaking wastewater treatment processes. Water Sci. Technol. 81, 1090–1098. https://doi.org/10.2166/wst.2020.206

Yang, S.S., Yu, X.L., Ding, M.Q., He, L., Cao, G.L., Zhao, L., Tao, Y., Pang, J.W., Bai, S.W., Ding, J., Ren, N.Q., 2021. Simulating a combined lysis-cryptic and biological nitrogen removal system treating domestic wastewater at low C/N ratios using artificial neural network. Water Res. 189, 116576. https://doi.org/10.1016/j.watres.2020.116576

Yang, T., Zhang, L., Wang, A., Gao, H., 2013. Fuzzy modeling approach to predictions of chemical oxygen demand in activated sludge processes. Inf. Sci. (Ny). 235, 55–64. https://doi.org/10.1016/j.ins.2012.07.021

Yasmin, N.S.A., Wahab, N.A., Anuar, A.N., 2020. Improved support vector machine using optimization techniques for an aerobic granular sludge. Bull. Electr. Eng. Informatics 9, 1835–1843. https://doi.org/10.11591/eei.v9i5.2264

Yazdankish, E., Foroughi, M., Azqhandi, M.H.A., 2020. Capture of I131 from medical-based wastewater using the highly effective and recyclable adsorbent of g-C3N4 assembled with Mg-Co-Al-layered double hydroxide. J. Hazard. Mater. 389, 122151. https://doi.org/10.1016/j.jhazmat.2020.122151

Ye, Z., Yang, J., Zhong, N., Tu, X., Jia, J., Wang, J., 2020. Tackling environmental challenges in pollution controls using artificial intelligence: A review. Sci. Total Environ. 699, 134279. https://doi.org/10.1016/j.scitotenv.2019.134279

Yel, E., Onen, V., Tezel, G., Yilmazturk, D., 2020. Performance prediction modeling of andesite processing wastewater physicochemical treatment via artificial neural network. Arab. J. Geosci. 13. https://doi.org/10.1007/s12517-020-05940-4

Yel, E., Yalpir, S., 2011. Prediction of primary treatment effluent parameters by Fuzzy Inference System (FIS) approach. Procedia Comput. Sci. 3, 659–665. https://doi.org/10.1016/j.procs.2010.12.110

Yetilmezsoy, K., Ozkaya, B., Cakmakci, M., 2011. Artificial intelligence-based prediction models for environmental engineering. Neural Netw. World 21, 193–218. https://doi.org/10.14311/NNW.2011.21.012

Youse, M., Ebrahim, M., Soltani, J., Roozbahani, A., 2018. Multi-objective particle swarm optimization model for conjunctive use of treated wastewater and groundwater 208, 224–231. https://doi.org/10.1016/j.agwat.2018.06.025

Yu, Z., Yousaf, K., Ahmad, M., Yousaf, M., Gao, Q., Chen, K., 2020. Bioresource Technology E ffi cient pyrolysis of ginkgo biloba leaf residue and pharmaceutical sludge (mixture) with high production of clean energy : Process optimization by particle swarm optimization and gradient boosting decision tree algorithm. Bioresour. Technol. 304, 123020.

https://doi.org/10.1016/j.biortech.2020.123020

Zaghloul, M.S., Hamza, R.A., Iorhemen, O.T., Tay, J.H., 2020. Comparison of adaptive neuro-fuzzy inference systems (ANFIS) and support vector regression (SVR) for data-driven modelling of aerobic granular sludge reactors. J. Environ. Chem. Eng. 8, 103742. https://doi.org/10.1016/j.jece.2020.103742

Zakeri, H.R., Yousefi, M., Mohammadi, A.A., Baziar, M., Mojiri, S.A., Salehnia, S., Hosseinzadeh, A., 2021. Chemical coagulation-electro fenton as a superior combination process for treatment of dairy wastewater: performance and modelling. Int. J. Environ. Sci. Technol. https://doi.org/10.1007/s13762-021-03149-w

Zeinolabedini, M., Najafzadeh, M., 2019. Comparative study of different wavelet-based neural network models to predict sewage sludge quantity in wastewater treatment plant. Environ. Monit. Assess. 191. https://doi.org/10.1007/s10661-019-7196-7

Zhang, L., Chao, B., Zhang, X., 2020a. Modeling and optimization of microbial lipid fermentation from cellulosic ethanol wastewater by Rhodotorula glutinis based on the support vector machine. Bioresour. Technol. 301, 122781. https://doi.org/10.1016/j.biortech.2020.122781

Zhang, L., Chao, B., Zhang, X., 2020b. Modeling and optimization of microbial lipid fermentation from cellulosic ethanol wastewater by Rhodotorula glutinis based on the support vector machine. Bioresour. Technol. 301, 122781. https://doi.org/10.1016/j.biortech.2020.122781

Zhang, L., Han, H., Qiao, J., 2021. Dynamic Multiobjective Optimal Control with Knowledge-Decision for Wastewater Treatment Process 72–78. https://doi.org/10.1109/iccss52145.2020.9336933

Zhang, X., Yan, S., Chen, J., Tyagi, R.D., Li, J., 2020. Physical, chemical, and biological impact (hazard) of hospital wastewater on environment: presence of pharmaceuticals, pathogens, and antibiotic-resistance genes. Curr. Dev. Biotechnol. Bioeng. 79–102. https://doi.org/10.1016/b978-0-12-819722-6.00003-1

Zhang, Y., Gao, X., Smith, K., Inial, G., Liu, S., Conil, L.B., Pan, B., 2019. Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. Water Res. 164, 114888. https://doi.org/10.1016/j.watres.2019.114888

Zhang, Y., Pan, B., 2014. Modeling batch and column phosphate removal by hydrated ferric oxide-based nanocomposite using response surface methodology and artificial neural network. Chem. Eng. J. 249, 111–120. https://doi.org/10.1016/j.cej.2014.03.073

Zhang, Z., He, X., Kusiak, A., 2015. Data-driven minimization of pump operating and maintenance cost. Eng. Appl. Artif. Intell. 40, 37–46. https://doi.org/10.1016/j.engappai.2015.01.003

Zhang, Z., Kusiak, A., Zeng, Y., Wei, X., 2016. Modeling and optimization of a wastewater pumping system with data-mining methods. Appl. Energy 164, 303–311. https://doi.org/10.1016/j.apenergy.2015.11.061

Zhang, Z., Zeng, Y., Kusiak, A., 2012. Minimizing pump energy in a wastewater processing plant. Energy 47, 505–514. https://doi.org/10.1016/j.energy.2012.08.048

Zhao, L., Dai, T., Qiao, Z., Sun, P., Hao, J., Yang, Y., 2020. Application of artificial intelligence to wastewater treatment: A bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse. Process Saf. Environ. Prot. 133, 169–182. https://doi.org/10.1016/j.psep.2019.11.014

Zhao, L.J., Chai, T.Y., Yuan, D.C., 2012. Selective ensemble extreme learning machine modeling of effluent quality in wastewater treatment plants. Int. J. Autom. Comput. 9, 627–633. https://doi.org/10.1007/s11633-012-0688-3

Zhou, P., Li, Z., Snowling, S., Baetz, B.W., Na, D., Boyd, G., 2019. A random forest model for inflow prediction at wastewater treatment plants. Stoch. Environ. Res. Risk Assess. 33, 1781–1792. https://doi.org/10.1007/s00477-019-01732-9

Zhu, A., Guo, J., Ni, B.J., Wang, S., Yang, Q., Peng, Y., 2015. A novel protocol for model calibration in biological wastewater treatment. Sci. Rep. 5, 29–32. https://doi.org/10.1038/srep08493

Zhu, X., Wan, Z., Tsang, D.C.W., He, M., Hou, D., Su, Z., Shang, J., 2021a. Machine learning for the selection of carbon-based materials for tetracycline and sulfamethoxazole adsorption. Chem. Eng. J. 406, 126782. https://doi.org/10.1016/j.cej.2020.126782

Zhu, X., Wan, Z., Tsang, D.C.W., He, M., Hou, D., Su, Z., Shang, J., 2021b. Machine learning for the selection of carbon-based materials for tetracycline and sulfamethoxazole adsorption. Chem. Eng. J. 406, 126782. https://doi.org/10.1016/j.cej.2020.126782

Zhu, X., Wang, X., Ok, Y.S., 2019. The application of machine learning methods for prediction of metal sorption onto biochars. J. Hazard. Mater. 378, 120727. https://doi.org/10.1016/j.jhazmat.2019.06.004

FIGURE CAPTIONS

Figure 1. Distribution of publications by AI model type.

Figure 2. Frequency and trend of AI techniques applied to wastewater treatment during 2000-2020.

Figure 3. Heatmap of the number of models with artificial neural network, support vector machine, genetic algorithm, decision tree, and Fuzzy logic with different applications in wastewater treatment.

Figure 4. The classification tree of the ML and metaheuristic techniques applied to wastewater treatment process. WNN: Wavelet Neural Network, BFA: Butterfly Algorithm; GWO: Grey Wolf Optimizer, FFA: Firefly Algorithm, WOA: Whale Optimization Algorithm, SSO: Salp Swarm Optimization, MFL: Mamdani Fuzzy Logic, TSFL: Takagi-Sugeno Fuzzy Logic, LFL: Larsen Fuzzy Logic, SCFL: Supervised Committee Fuzzy Logic, DSA: Direct Search Algorithm, WDN: Wavelet de noise.

Figure 5. Schematic representation of the hybrid combination of CNN, GRU and DNN (ME-DeepL) architecture (Heo et al., 2020).

Figure 6. Overview of data preparation techniques.

LIST OF TABLES

Revised List of Figures

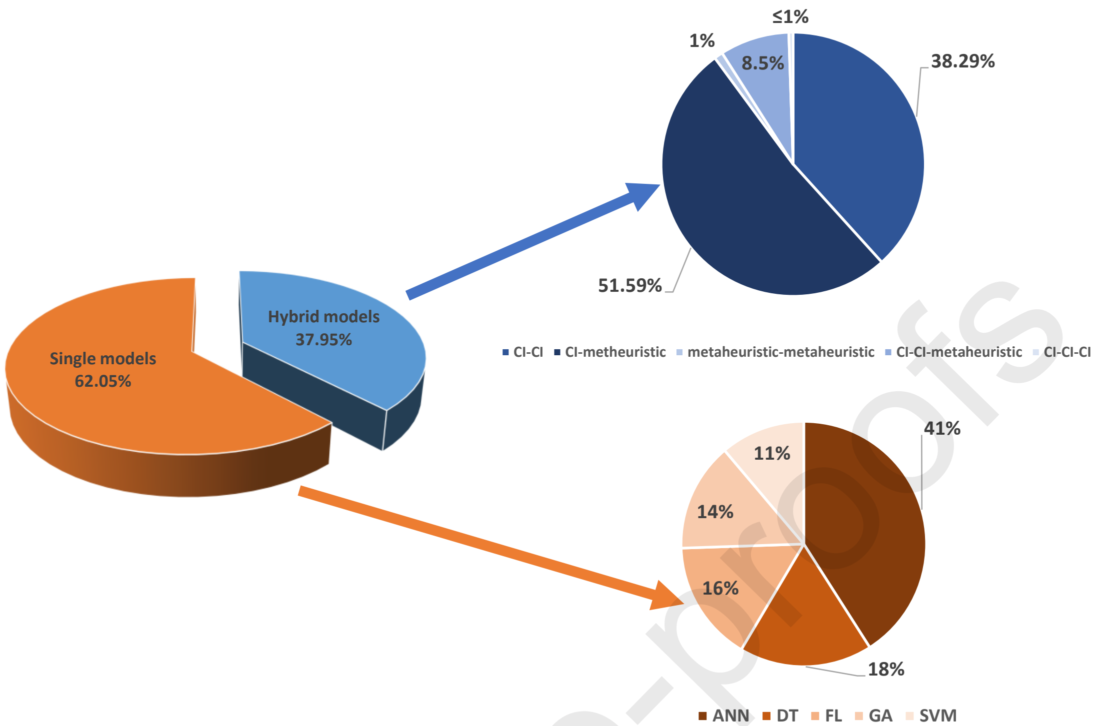

≤1%
1%
8.5%
38.29%
51.59%
CI-CI
CI-metheuristic
metaheuristic-metaheuristic
CI-CI-metaheuristic
CI-CI-CI
Hybrid models
37.95%
Single models
62.05%
41%
11%
14%
16%
18%
ANN
DT
FL
GA
SVM

Figure 1

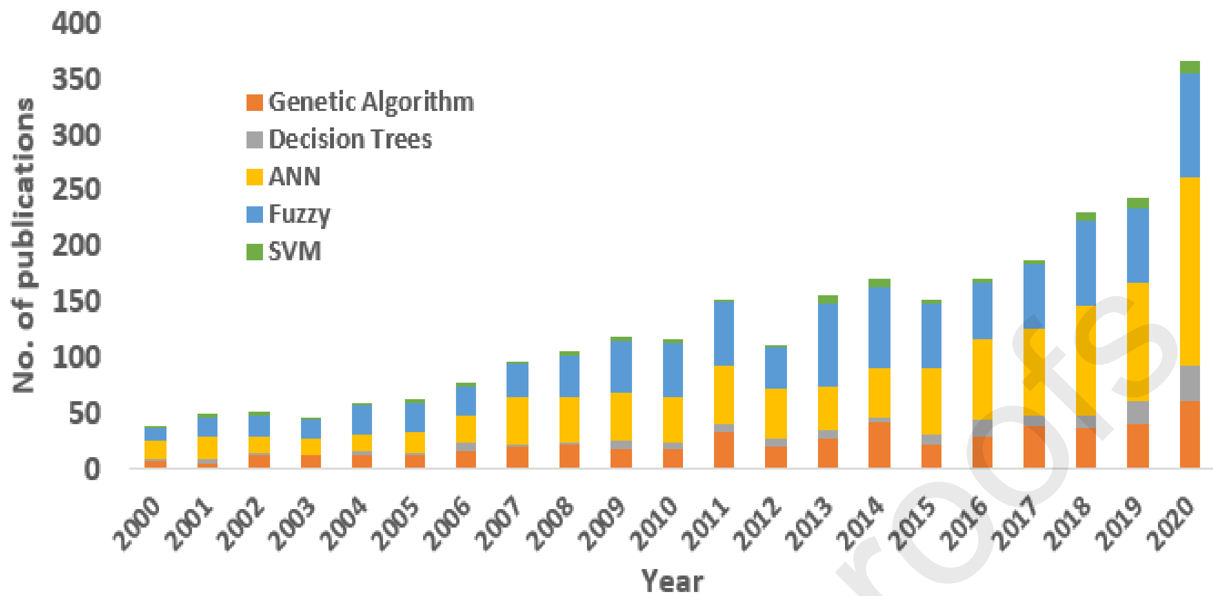

400
350
300
250
200
150
100
50
0
No. of publications
Genetic Algorithm
Decision Trees
ANN
Fuzzy
SVM
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Year

Figure 2

	ANN	SVM	GA	DT	Fuzzy
Greenhuse gas emissions	2	2	1	2	2
Concentration	3			1	
Operational management and control	3				5
Conventional pollutant removal	22	7	3	5	8
Typical pollutant removal	20	4	7	2	
Energy consumption	7	1	2	1	6
Other costs	3	1	3		1
	0	5	10	15	25

Figure 3

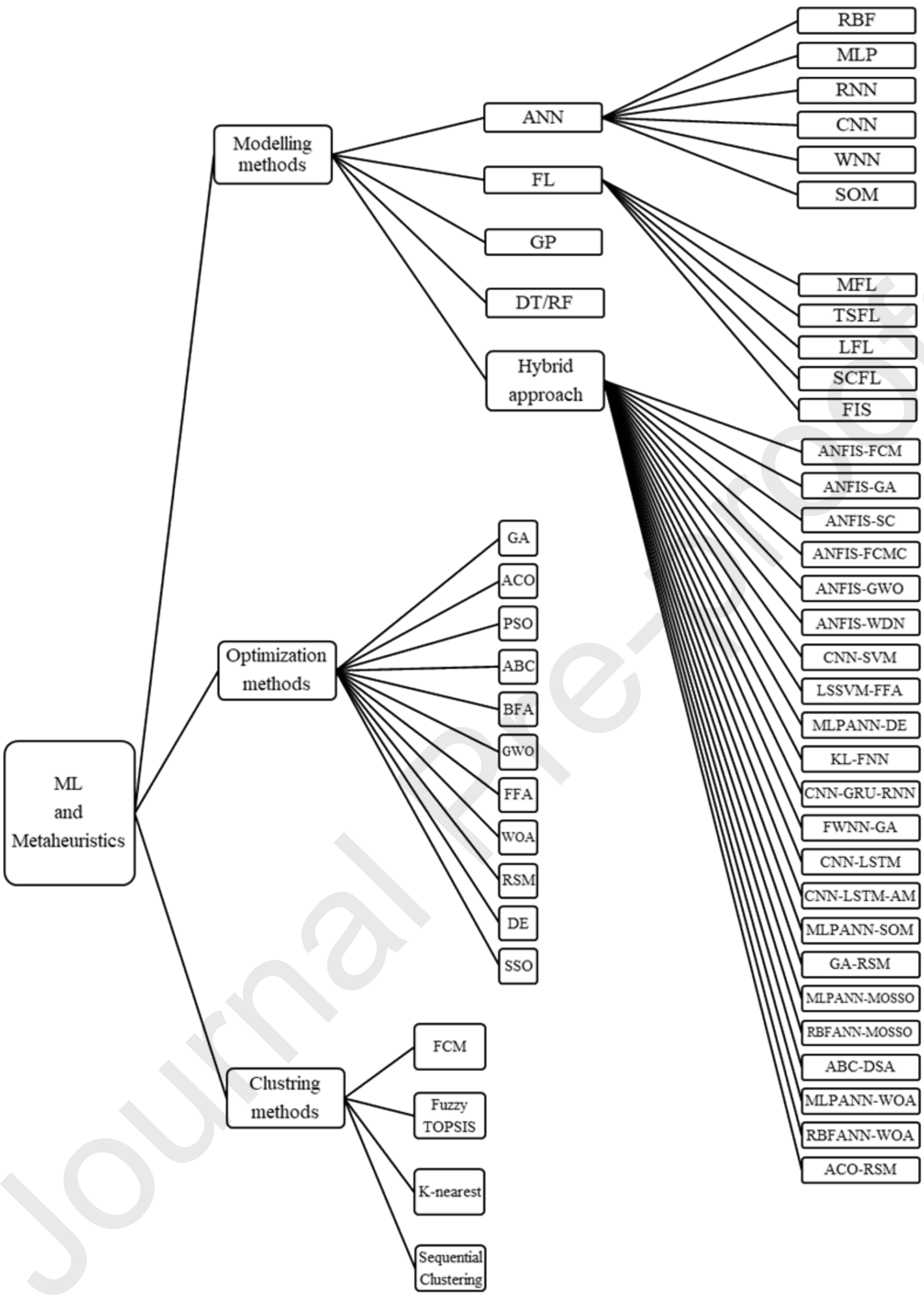

ML
and
Metaheuristics
Modelling methods
Optimization methods
Clustring methods
ANN
FL
GP
DT/RF
Hybrid approach
RBF
MLP
RNN
CNN
WNN
SOM
MFL
TSFL
LFL
SCFL
FIS
ANFIS-FCM
ANFIS-GA
ANFIS-SC
ANFIS-FCMC
ANFIS-GWO
ANFIS-WDN
CNN-SVM
LSSVM-FFA
MLPANN-DE
KL-FNN
CNN-GRU-RNN
FWNN-GA
CNN-LSTM
CNN-LSTM-AM
MLPANN-SOM
GA-RSM
MLPANN-MOSSO
RBFANN-MOSSO
ABC-DSA
MLPANN-WOA
RBFANN-WOA
ACO-RSM
GA
ACO
PSO
ABC
BFA
GWO
FFA
WOA
RSM
DE
SSO
FCM
Fuzzy TOPSIS
K-nearest
Sequential Clustering

Figure 4

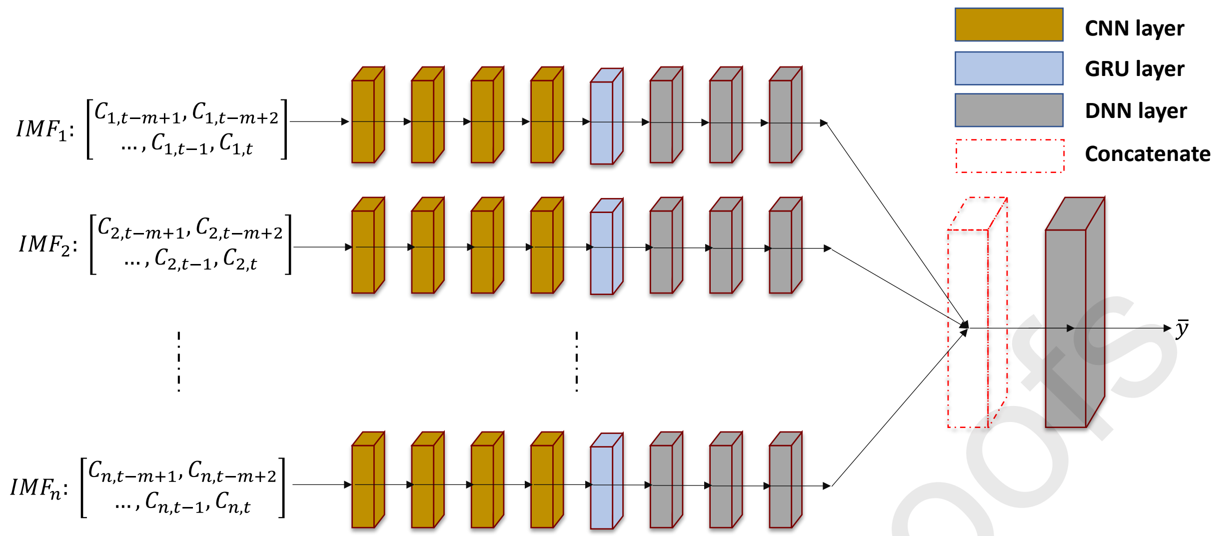

CNN layer
GRU layer
DNN layer
Concatenate
IMF_1: $\begin{bmatrix} C_{1,t-m+1}, C_{1,t-m+2} \\ \dots, C_{1,t-1}, C_{1,t} \end{bmatrix}$
IMF_2: $\begin{bmatrix} C_{2,t-m+1}, C_{2,t-m+2} \\ \dots, C_{2,t-1}, C_{2,t} \end{bmatrix}$
IMF_n: $\begin{bmatrix} C_{n,t-m+1}, C_{n,t-m+2} \\ \dots, C_{n,t-1}, C_{n,t} \end{bmatrix}$
$\bar{y}$

Figure 6

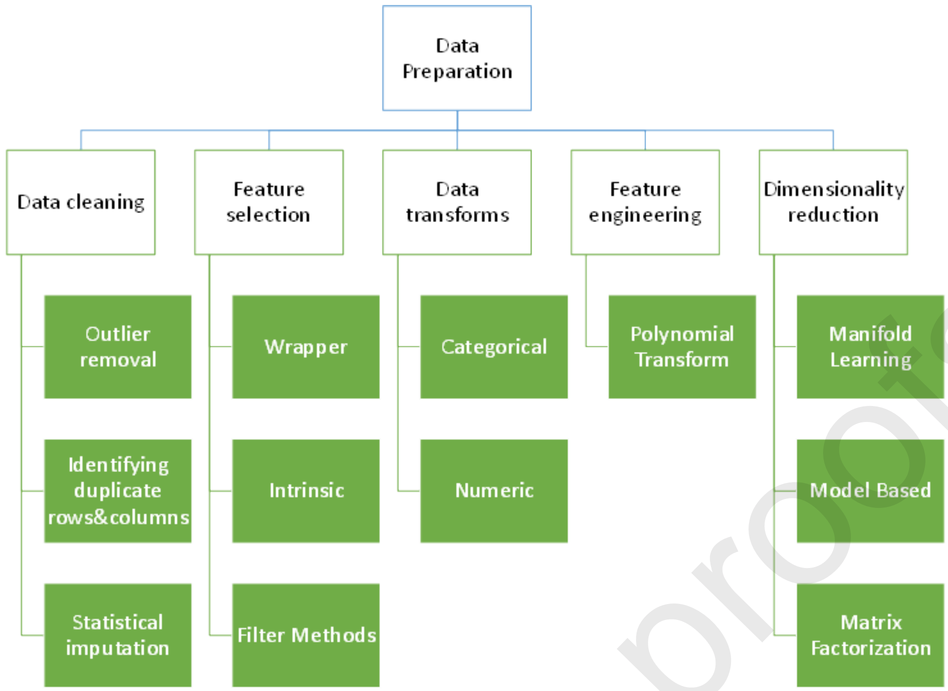


Figure 7

Revised List of Tables

Table 1. Previous review work focusing on the application of AI models in water and wastewater treatment.

Reference	Publication year	Application field	Review intention	Types of AI models	No. of reviewed studies	Period of review
(Zhao et al., 2020)	2020	Wastewater treatment	bibliometric analysis, systematic review	ANN, FL, GA, ES, MT, BN, PSO, SVM	124	1995 - 2019
(Abdallah et al., 2020)	2020	Solid waste management	systematic review	ANN, SVM, LR, DT, GA	81	2000 - 2020
(Bhagat et al., 2020)	2020	Heavy metal removal from wastewater	State of the art, application assessment,	ANN, RSM, MLR, FL, ANFIS, SVM, GP, SVR	231	2000 - 2019
(Newhart et al., 2019)	2019	Wastewater treatment plant	future research	ANN, FL	153	1994-2018
(Corominas et al., 2018)	2018	Wastewater treatment plant	Systematic review of data-driven process control methods	ANN, FL, ANFIS, SVM, FCM, DT	187	1990 - 2015
(Yetilmezsoy et al., 2011)	2011	Environmental systems	critical review of data-driven wastewater operation	ANN, FL, ANFIS, Quasi-newton, MLR	111	2000 - 2010
Present review	2021	Wastewater treatment	Narrative review of	ANN, FL, SVM, DT, GA, PSO,	281	2000 - 2020

			data-driven techniques	ACO, ANFIS, Hybrid		

Table 2. Application of AI methods in treatment performance of WWTPs.

Application	ANN	SVM	GA	DT	Fuzzy
Biological process[1]	(Antwi et al., 2019a; Ranade et al., 2021; Almomani, 2020; Besharati Fard et al.,	(Cai et al., 2019; Liu et al., 2019; Mahata et al., 2020; Szeląg et	(Niu et al., 2020; Du et al., 2020; Lee et al., 2009; Suggala and	(Byliński et al., 2019; Song et al.,	(Ansari et al., 2020; Bae et

	2020; Do and Schmitt, 2020; Hayder et al., 2014; Kundu et al., 2013; López et al., 2017; Ma et al., 2011; Moral et al., 2008; Mojiri et al., 2020; Nawaz et al., 2019; Shi et al., 2009; Zhang and Pan, 2014)	al., 2020; Wu et al., 2020; Yasmin et al., 2020; Zaghloul et al., 2020; Zhang et al., 2020)	Bhattacharya, 2003)	2020; Sharafati et al., 2020; Deepnarain et al., 2019)	al., 2006; Borges et al., 2016; Gupta et al., 2017; Marsili-Libelli and Giunti, 2002; Mazhar et al., 2019; Revollar et al., 2018; Tejaswini et al., 2020b)
Physical processes[2]	(Aish et al., 2015; Ghandehari et al., 2011; Libotean et al., 2009; Sattar et al., 2019)	(Fang et al., 2019; Najafzadeh and Zeinolabedini, 2019)	-	(Verma et al., 2013)	(Patel et al., 2020; Yel and Yalpir, 2011)
Chemical process[3]	(Asfaram et al., 2016; Chen et al., 2019; Mokhtari Nesfchi et al., 2021; Mungray et al., 2021; Onu et al., 2021; Pelalak et al., 2021; Shen et al., 2018; Yang et al., 2021; Yel et al., 2020; da Silva Ribeiro et al., 2019; Fawzy et al., 2016; Haghiri et al., 2018;Messikh et al., 2015)	(Farzin et al., 2020; Foroughi et al., 2020; Huang et al., 2020; Lariche et al., 2020)	(Yazdankish et al., 2020; Tanzifi et al., 2020; Rahmani et al., 2020; Picos-Benítez et al., 2020; Onu et al., 2021; Khan et al., 2020; Tashvigh and Nasernejad, 2017; Dawood and Li, 2013)	(Zhu et al., 2021; De Miranda Ramos Soares et al., 2020)	(Bello et al., 2014; Khawaga et al., 2019)

[1] The biological process includes: Conventional Activated Sludge (CAS), biological nitrogen and phosphate removal, Single-stage Nitrogen removal using Anammox and Partial nitrification (SNAP), enhanced nutrient removal biological process, up-flow anaerobic sludge blanket, membrane bioreactors (MBR), multi-stage biological reactors, sequencing batch reactor (SBR), biotrickling filter, biofilter, anaerobic filters, microbial fuel cells, anaerobic digestion, aerobic and anaerobic granular Sludge, aerobic-anaerobic-sequential biological treatment, dark fermentation, microalgae growth and kinetics, sludge treatment and biomass pretreatment.

[2] The physical process includes reverse osmosis (RO), crossflow microfiltration, membrane filtration, primary and secondary clarifiers, aeration and the blowers.

[3] The chemical process includes dye adsorption, flocculation and coagulation, interfacial energies adjacent to membrane surface, forward osmosis (FO), electrochemical and photo electrochemical degradation, sludge ultrasonication, ozonation and membrane distillation.

Table 3. Purpose of the GA in the reviewed studies.

Forecasted variable	Purpose of GA	Year	Reference

	Parameter and model structure optimization	Variable prediction		
N-nitrosamine removal	✓	-	2018	(Al-Obaidi et al., 2018)
Membrane fouling	-	✓	2009	(Hwang et al., 2009)
Hydrogen production	✓	-	2020	(Mahata et al., 2020b)
Removal of chlorophenol	✓	-	2017	(Al-Obaidi et al., 2017)
pH control	✓	-	2012	(Alwan, 2012)
Wastewater leakage	✓	-	2020	(Attwa and Zamzam, 2020)
Pharmaceuticals in wastewater	✓	-	2007	(Babić et al., 2007)
EQI, energy consumption for pumping and aeration	✓	-	2007	(Béraud et al., 2007)
Capital and operational costs	✓	-	2011	(Brand and Ostfeld, 2011)
Operational costs	✓	-	2001	(Chang et al., 2001)
Effluent COD	-	✓	2010	(Chen et al., 2010)
COD, NO_3^- and PO_4^{3-}	-	✓	2015	(Huang et al., 2015)
Influent flowrate, operational costs, discharge BOD	✓	-	2009	(Iqbal and Guria, 2009)
Online fault detection	✓	-	2019	(Li and Yan, 2019)
Photocatalytic degradation of dyes	✓	-	2017	(Mahmoodi et al., 2017)

Chlorophenol removal using RO	✓	-	2020	(Mohammad et al., 2020)
airflow rate, biofilm carrier, carbon source temperature and pH in Gas-Liquid-Solid bioreactor	✓	-	2014	(Srinu Naik and Pydi Setty, 2014)
Effluent quality index	✓	-	2020	(Niu et al., 2020b)
Design parameters of single chamber microbial fuel cell	✓	-	2013	(Sedaqatvand et al., 2013)
Permeation flux in microfiltration	-	✓	2012	(Shokrkar et al., 2012)
Overall operational cost	✓	-	2020	(J. Wang et al., 2020)
Design of a carrousel oxidation ditch	✓	✓	2011	(Xie et al., 2011)
COD removal in microbial lipid fermentation	✓	-	2020	(L. Zhang et al., 2020b)
Effluent quality index	✓	-	2012	(Zhao et al., 2012)
Kinetics of SBR and CSTR	✓	-	2015	(Zhu et al., 2015)
Total number of models	**23**	**5**		
%	**82**	**18**		

Table 4. The purpose of ACO in the reviewed hybrid models.

Forecasted	**Purpose of ACO**	**Year**	**Reference**

variable	Parameter optimization	Improve prediction performance		
Biogas production	✓	-	2019	(Beltramo et al., 2019)
DO concentration		✓	2016	(Y. Chen et al., 2016)
Crystal violet adsorption	✓	✓	2018	(Ghazali et al., 2018)
MB adsorption	✓	✓	2020	(Jun et al., 2020a)
Cr (VI) adsorption	✓		2020	(Karri et al., 2020)
Influent volume		✓	2012	(Qiao et al., 2012)
Furaldehydes 2-furaldehyde (F) and 5-hydroxymethyl-2-furaldehyde (HMF)	-	✓	2020	(Ali Akbar Miran Beigi, 2020)
Effluent flowrate	-	✓	2019	(Rastegaripour et al., 2019)
Total number of models	**4**	**6**		
%	**40**	**60**		

Table 5. Performance indicators of artificial intelligence models with the reported acceptance range in the reviewed studies.

Term	**Equation**	**No. of studies**	**Reported Acceptable Range**				
			ANN	**DT**	**SVM**	**Fuzzy**	**Hybrid**
Correlation coefficient (R)	$R = \frac{\sum XY}{n\sigma_x\sigma_y}, X = x - \bar{x}, Y = y$	56	[0.4789, 0.998]	[0.78, 0.98]	[0.78, 0.97]	[0.7, 0.95]	[0.92, 0.99]
Coefficient of determination (R^2)	$R^2 = \left(\frac{\sum XY}{n\sigma_x\sigma_y}\right)^2$	72	[0.616, 0.997]	[0.781, 0.906]	[0.672, 0.992]	[0.71, 0.992]	[0.89, 0.99]
Mean Square Error (MSE)	$MSE = \frac{1}{n}\sum(O_v - P_v)^2$	14	[0.010, 12.64]	[735.68, 4587.25]	[0.012, 0.0042]*	-	[0.000108, 4.0375]
Normalized MSE (NMSE)	$NMSE = \frac{1}{nS_{d_0}}\sum(O_v - P_v)^2$	8	[0.65, 0.89]	[0.57, 0.79]	[0.62, 0.75]	[0.77, 0.91]	[0.77, 0.99]
Root mean square error (*RMSE*)	$RMSE = \sqrt{\frac{1}{n}\sum(O_v - P_v)^2}$	62	[0.76, 4.91]	[103, 3486]	[0.047, 45.32]	[0.70, 45]	[0.0004, 1.76]
Men absolute percentage error (*MAPE*)	$MAPE = \frac{1}{n}\sum\frac{\lvert O_v - P_v\rvert}{O_v} \times 1$	26	-	[12.91%, 18.43%]	[0.6%, 1.8%]	[10%, 13%]	[0.381%, 1.28%]

Mean absolute error (*MAE*)	$MAE = \frac{1}{n}\sum\lvert O_v - P_v\rvert$	30	[0.0585, 6.6217]	[0.57, 40.94]	[0.037, 1031.1]	[0.02, 5.14]	[0.02, 3.87]
Nash-Sutcliffe Error (NSE)	$NSE = 1 - \frac{\sum(O_v - P_v)^2}{\sum(O_v - \overline{O_v})^2}$	3	[0.75, 0.98]	-	-	-	[0.65, 0.97]
Index of agreement (*IA*)	$IA = 1 - \left[\frac{\sum(P_v - O_v)}{\sum(\lvert P_v - \overline{O_v}\rvert + \lvert O_v}\right.$	1	[0.993, 0.998]	-	-	-	-
Absolute Average Deviation	$AAD\% = \frac{1}{n}\sum\left(\frac{P_v - O_v}{P_v}\right) \times 1$	1	[0.1838%, 0.5631%]	-	-	-	-
PBIAS	$PBIAS = \frac{\sum(O_v - P_v)}{\sum O_v}$	1	-	-	-	[0.77, 0.98]	-

Table 6. Summary of simulation platforms used in the reviewed studies.

Simulation platform	**Open source**	**Cost**	**No. of studies**	**%**
MATLAB	-	Paid	81	64.8
Design Expert system	✓	Free	14	11.2

Statistica	✓	Paid	7	5.6
R (programming language)	✓	Free	5	4
MINITAB	✓	Free	5	4
Python	✓	Paid	3	2.4
GPSX	✓	Paid	2	1.6
SPSS	-	Paid	2	1.6
NNMODEL	✓	Free	1	0.8
Weka	✓	Free	1	0.8
C ++	✓	Free	1	0.8
KNIME	✓	Free	1	0.8
NeuroShell2	✓	Free	1	0.8
DESASS	✓	Free	1	0.8

CRediT author statement

Majid Bahramian: Conceptualization, Methodology, Software **Recep Kaan Dereli**: Writing- Original draft preparation. **Wanqing Zhao**: Investigation. Paraphrasing **Matteo Giberti***:* Software, Validation **Eoin Casey**: Supervision, Writing- Reviewing and Editing,

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

ORCID information

Majid Bahramian https://orcid.org/0000-0002-7571-5567

Recep Kaan Dereli https://orcid.org/0000-0001-7853-270X

Wanqing Zhao https://orcid.org/0000-0003-0611-0529

Matteo Giberti https://orcid.org/0000-0002-7075-6626

Eoin Casey https://orcid.org/0000-0002-6471-7356

Highlights

- ANN outnumbered other standalone AI models (single models) applied to WWTPs.
- Hybrid models were relatively more accurate than the standalone models.
- Most of hybrid models were classified as CI-metaheuristic models.
- FL was the most suitable model for the incomplete data sets.
- Despite recent developments, industrial deployment is lacking.