JPIChE 54 (2) 2026: 14-22

PAKISTAN INSTITUTE OF CHEMICAL ENGINEERS Since 1969

Journal of the Pakistan Institute of Chemical Engineers

Journal homepage: www.piche.org.pk/journals

DOI: https://doi.org/10.54693/piche.05422

Check for updates

An AI-Enabled Digital Twin Framework for Real-Time Prediction and Optimization of Industrial Wastewater Treatment Processes

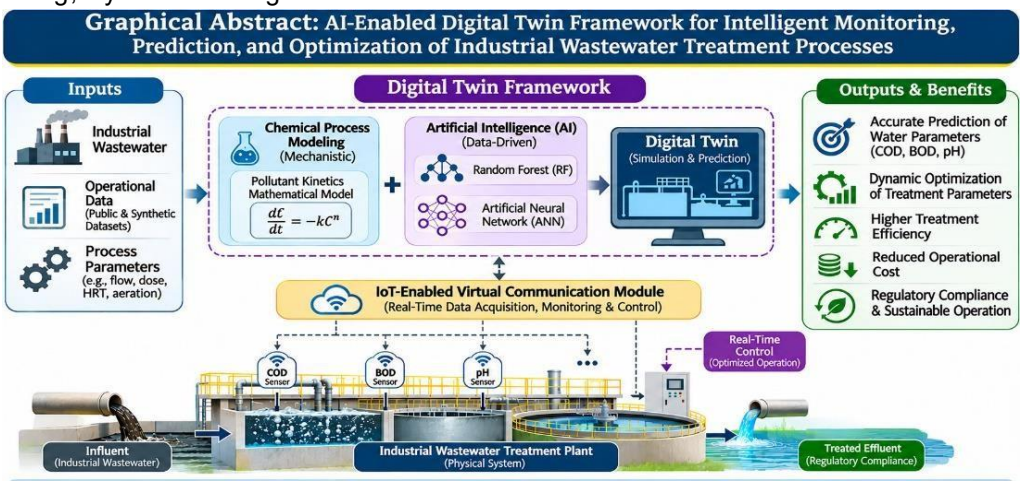
A. Rehman[1], A. Israr[2]*, A.S. Ghouri[3]

Submitted: 27/04/2026, Accepted: 10/10/2026, Published: 19/10/2026

Abstract

Continuous monitoring and optimization of industrial wastewater treatment processes are essential to meet regulatory requirements and ensure operational efficiency. Traditional methodologies for wastewater treatment are primarily based on physical experimentation and mechanistic modeling. This study aims at developing a digital twin model to optimize the treatment processes in industrial wastewater treatment through a computational method. The suggested framework incorporates a combination of chemical process modeling and artificial intelligence techniques to develop a simulation model to predict the water parameters such as Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), and pH. A mathematical model for pollutant kinetics coupled with Random Forest (RF) and Artificial Neural Network (ANN) methods will be utilized. Moreover, a virtual communication module is included to simulate the IoT-enabled real-time data acquisition and control process. The system is tested using both public and synthetically generated datasets with different operating scenarios. Comparative study reveals that the hybrid AI-mechanistic model achieves superior predictive accuracy and flexibility compared to traditional methods. Moreover, the system is capable of optimizing the treatment parameters dynamically, which helps in managing the processes in an intelligent manner. The findings demonstrate that the integration of AI into chemical engineering can lead to the development of a viable solution for intelligent and cost-effective wastewater treatment systems.

Keywords: Digital Twin; Wastewater Treatment; Machine Learning; Reaction Kinetics; Process Optimization; IoT Monitoring; Hybrid Modeling.



Graphical Abstract

[1] Department of Information Technology, Balochistan University of Information Technology, Engineering and Management Sciences, Pakistan.
[2] Department of Electrical Engineering, Balochistan University of Information Technology, Engineering and Management Sciences, Pakistan.
[3] Department of Chemical Engineering, Balochistan University of Information Technology, Engineering and Management Sciences, Pakistan.
*__Corresponding Author:__ Adil Israr (adil.israr@buitms.edu.pk)

1. Introduction:

There has been significant generation of effluents from industries due to rapid industrial development, causing various environmental problems and health hazards. Various industries include textile industry, chemical industry, and pharmaceutical industry, which produce effluents containing different types of organic and inorganic compounds, including toxic substances such as heavy metals and hazardous organic residues. There is a need for treating such effluents as per the regulatory standards in order to minimize environmental damage [1], [2]. In traditional effluent treatment plants, there are different physical, chemical, and biological treatments carried out based on several parameters.

Even with technological advances in wastewater treatment, there are challenges associated with the monitoring and control of these processes. Conventional methods are mostly based on measurements conducted in laboratories and rely mainly on static models, which may not always be adequate in representing the dynamic behavior of real-life systems [3], [4]. This could result in poor process operation, unnecessary expenses, and variations in output quality. In recent times, data-driven techniques have been found useful in predicting outcomes for complex environmental systems. Machine learning (ML) has shown considerable promise in this regard [5], [6]. Nonetheless, data-driven models could suffer from a lack of interpretability.

The idea of a digital twin has evolved into a potential tool that could help reconcile knowledge of physical processes with data analysis techniques. A digital twin implies creating a virtual representation of the physical process that is subject to constant observation, simulation, and optimization. Its use has been discussed in recent research focused on applying digital twins in industrial settings and smart manufacturing plants [7], [8], [9], [10]. Digital twin approaches have become increasingly popular in the realm of water and wastewater systems, where they could assist in improving process control and management [10], [11], [12]. However, combining digital twins with artificial intelligence and modeling of chemical processes is still in the development stages [14], [15], [16].

While modern studies heavily deploy machine learning for standalone water quality forecasting [13] or utilize IoT infrastructure strictly as a passive data logging pipeline [15], existing frameworks almost exclusively treat wastewater systems under isolated paradigms. Current hybrid digital twins typically use AI to estimate physical variables or rely on parallel, decoupled executions where the physics and data-driven layers do not continuously govern each other.

The specific methodological contribution of this work lies in its enclosed closed-loop coupling mechanism. Instead of using machine learning as a parallel black-box simulator, the proposed architecture embeds the first-order reaction kinetics directly into the objective framework, forcing the data-driven layer to model the physical residual variance $Y_{DT} = \alpha Y_{chem} + (1 - \alpha) Y_{ML}$. This specific architecture preserves process boundary constraints under extreme hydraulic fluctuations while requiring significantly lower training datasets than fully unconstrained deep learning pipelines.

For this research, digital twin architecture for real-time optimization of wastewater treatments is designed using artificial intelligence techniques. The model utilizes the principles of first-principles chemical models by applying the concept of pollutant degradation kinetics [3]. Machine learning techniques are used in conjunction with chemical modeling to improve the effectiveness and efficiency of prediction under changing conditions.

The key contributions of the study include the following:

(i) Development of a hybrid modeling approach that utilizes chemical kinetics and machine learning to predict wastewater treatment;

(ii) Design of a digital twin system for real-time monitoring and optimization; and

(iii) Computational validation of the approach using literature- and simulation-based datasets.

This research offers a scalable and affordable solution for intelligent wastewater management and emphasizes the importance of merging artificial intelligence and chemical engineering techniques.

2. Literature Review:

2.1 Digital Twin Applications in Process Industries:

The development of digital twin technology can be considered an essential step toward monitoring and optimization in industrial settings. This innovation is related to the creation of a virtual model that allows for continuous updates based on real-time or artificial data input. The first theoretical and systematic work on digital twins was developed for the industry [7], [9]. Further

research made it possible to extend the scope of application of digital twin technology to operations and system optimization [8], [10], [17].

The use of digital twins in environmental and water systems has just entered an early development stage. Some recent research papers have discussed their applications in wastewater treatment plants for monitoring and forecasting processes [11], [12], [18], [19]. Nevertheless, most developed models and systems are mostly descriptive or partially coupled and lack full integration with chemical reactions and adaptive artificial intelligence systems [20].

2.2 Machine Learning in Wastewater Treatment Systems:

The application of machine learning methods for predicting water quality characteristics like Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), turbidity, and pollutant concentration has been prevalent. Machine learning methods like ANN, RF, and deep learning are more accurate at predicting than traditional regression models [5], [6], [21]. Recent developments have focused on optimizing the process by exploring the use of deep learning and hybrid methods [5], [13], [22], [23]. Although these have been achieved, one disadvantage of present-day machine learning methods is their poor integration with chemical mechanism-based modeling. This has largely been due to treating the process as a black-box problem [24], [25], [26].

2.3 IoT and Communication in Environmental Monitoring:

The introduction of IoT technology has led to the possibility of monitoring in real-time and distributed collection of data in environmental systems. IoT platforms have made it possible to offer real-time sensing, communication, and cloud-based analysis for water quality monitoring systems [14], [15], [27]. Some recent studies have used IoT technology to monitor water quality and the environment [16], [28]. Although, in most cases, IoT-based solutions pay attention to collecting data and not using these collected data for predicting chemical reactions or any other processes using artificial intelligence.

2.4 Chemical Process Modeling in Wastewater Systems:

Models of chemical engineering continue to be essential for explaining the mechanisms of pollutants' breakdown, adsorption, and kinetics in wastewater management systems [29]. The use of classical models, including activated sludge methods and reaction engineering theory, is based on solid scientific reasoning [2], [3], [4]. Although frequently used in design and modeling, they cannot adequately represent nonlinear and dynamic industrial scenarios if applied separately[30]. Adsorption-based techniques have also been widely studied for removal of heavy metals in wastewater systems [31].

Table 1: Comparative Review of Existing Studies and Identified Research Gaps

No.	Study	Approach	Domain	Strength	Limitation	Key Insight for This Study
1	[7]	Digital Twin Concept	Manufacturing	Strong Digital Twin (DT) foundation	No environmental application	Establishes DT architecture concept
2	[9]	Systematic Review	DT systems	Clear classification	No implementation model	Defines DT research structure
3	[10]	Operations modeling	Industry systems	Optimization framework	No AI integration	Highlights need for AI coupling
4	[8]	Smart manufacturing DT	Industrial systems	Architecture design	Limited process chemistry	Shows DT adaptability gap
5	[5]	ML water modeling	Wastewater	High prediction accuracy	No mechanistic link	Demonstrates AI-only limitation
6	[6]	ANN models	COD/BOD prediction	Strong predictive performance	Black-box nature	Motivates hybrid modeling
7	[13]	Deep learning	Wastewater optimization	High accuracy	No interpretability	Need for physics-based integration
8	[21]	AI prediction	Water quality	Real-time forecasting	No process coupling	Lack of chemical integration

9	[11]	Digital twin monitoring	Wastewater plants	Monitoring capability	Weak AI integration	DT-AI disconnect
10	[12]	DT platform	Water systems	System architecture	No chemical kinetics	Missing mechanistic modeling

The existing works on digital twins, machine learning techniques, and wastewater treatment modeling were reviewed and compared in Table 1. It was observed from the literature review that although there are many investigations on digital twin concepts in manufacturing and industrial systems, their use in environmental engineering is rare [7], [8]. Machine learning algorithms are quite effective in estimating water quality; however, all these studies employ a data-driven approach without considering the underlying chemistry behind the system [5], [6], [13]. In a similar fashion, IoT-enabled monitoring solutions offer real-time data acquisition capability but fail to integrate any predictive/optimization layer into their operation. Overall, the identified studies show that there is currently a rather fragmented approach to research within this topic, whereby digital twins, AI algorithms, and chemical engineering approaches are developed separately from each other [32].

2.5 Research Gap Identification:

While considerable progress has been made in digital twin technologies, machine learning techniques, and IoT-based monitoring systems, one crucial area that remains inadequately researched is the integration of these techniques for wastewater treatment plants. Current literature primarily discusses the application of digital twins to industrial processes or the use of machine learning models for predictions without considering the underlying chemical kinetics [9], [10]. In addition, most AI models applied in wastewater treatment studies consider only the black-box nature of the process. Furthermore, IoT-driven monitoring schemes are separate from intelligent prediction and optimization schemes [14], [15]. Hence, there exists a definite requirement for a comprehensive digital twin system that combines chemical reaction models, machine learning techniques, and communication technologies to create intelligent wastewater treatment processes.

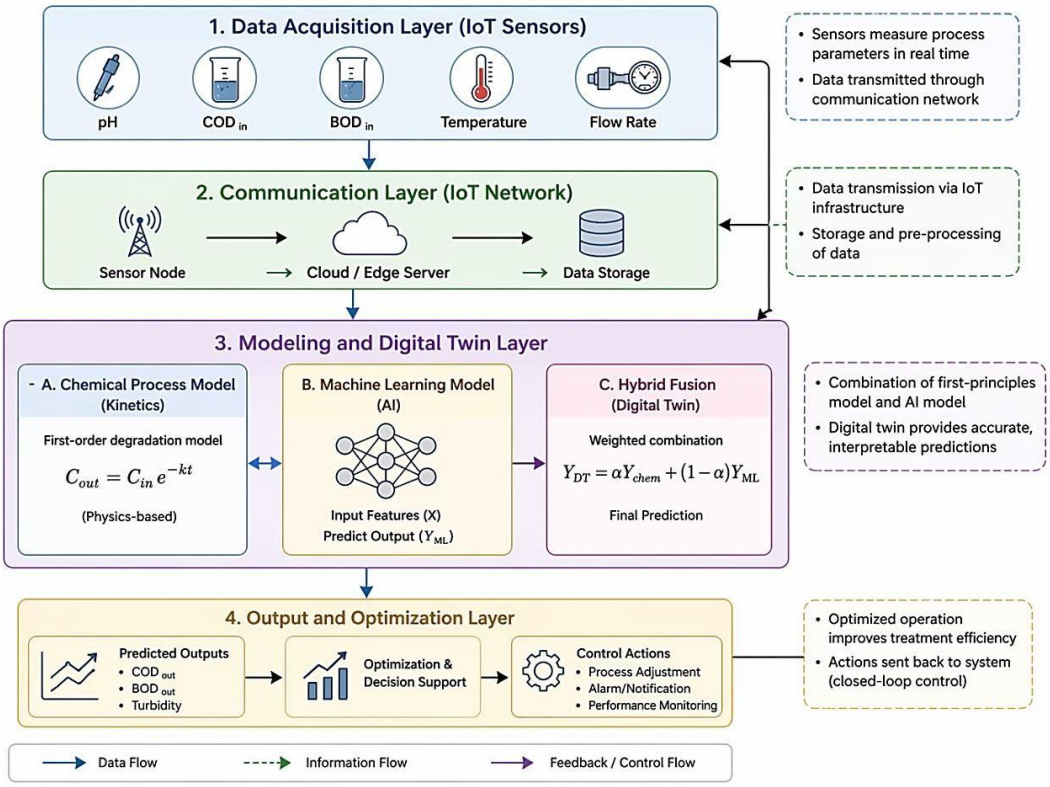


Figure 1: Proposed Digital Twin Framework

3. Proposed Framework and Methodology:

3.1 Overview of the Proposed Framework:

The digital twin architecture suggested with AI capabilities is shown in Figure 1. The process starts with the data acquisition layer where the following parameters – pH, COD, BOD, temperature, and flow rate – are collected. The data are sent to the communication layer, which describes the data exchange and data storage modeled based on IoT technologies. The modeling layer is the core of the system and consists of two models – chemical kinetics and machine learning. Moreover, it employs a hybrid approach to the fusion of the data and generates accurate predictions. Finally, the output layer provides the water quality parameter predictions.

The optimization capability of the proposed digital twin is structurally designed to operate as a predictive Decision Support System (DSS). Rather than using an unconstrained algorithmic solver, the framework uses heuristic parameter sweeping over the controllable process variables (such as hydraulic retention time and input flow rates) to evaluate 'what-if' operational scenarios and identify the most stable processing windows.

3.2 Chemical Process Modeling Layer:

The wastewater treatment model is depicted using a simple first-order kinetic expression, a common practice when studying the fate and behavior of pollutants in chemical and environmental engineering processes[2], [3].

$$C_{out} = C_{it}\, e^{-kt}$$

The selection of a pseudo-first-order degradation model serves as an established baseline for tracking bulk carbonaceous organic matter removal (represented by COD and BOD) under continuous, stable operating regimes where substrate concentrations are not fully limiting. This assumption remains valid under these conditions:
(i) the hydraulic flow regime approximates a plug-flow or well-mixed tank configuration at steady state,
(ii) microbial or chemical catalyst concentrations remain relatively constant relative to the changing organic load, and
(iii) macro-toxic shocks do not completely inhibit biological activity.
While complex multi-substrate kinetics exist, this first-order representation provides a computationally lean physical anchor necessary for real-time digital twin scaling without imposing prohibitive mathematical overhead.
Where:

- C_{in} = influent pollutant concentration
- C_{out}= effluent pollutant concentration
- k= reaction rate constant
- t= retention time

This model provides the physics-based baseline for the digital twin, ensuring chemical interpretability of the system. Adsorption kinetics and regeneration mechanisms have also been investigated in wastewater treatment processes [33].

3.3 Machine Learning Prediction Layer:

To capture nonlinear behavior not represented in classical chemical models, a supervised machine learning module is integrated into the framework.

Let the input feature vector be defined as:

$$X = [pH, COD_{in}, BOD_{in}, T, Flow\ Rate]$$

The output vector is:

$$Y = [COD_{out}, BOD_{out}, Turbidity]$$

The predictive model is defined as:

$$Y = f(X)$$

Where $f(X)$is approximated using machine learning algorithms such as RF and ANN [6], [13].

For ANN-based modeling:

$$Y = \sigma(W_2 \cdot \sigma(W_1 X + b_1) + b_2)$$

The ANN architecture is configured as a multi-layer perceptron comprising an input layer, two hidden layers with 16 and 8 neurons, respectively, utilizing the ReLU activation function, and a linear output layer. These structural hyperparameters and the Adam optimizer learning rate (0.001) were established via a standard grid-search optimization procedure to minimize validation error.

Where:

- $W_1, W_2$are weight matrices

- $b_1, b_2$are bias terms
- $\sigma$is the activation function

This layer improves prediction accuracy by learning complex nonlinear relationships in wastewater behavior [21], [34].

3.4 Digital Twin Integration Mechanism:

The digital twin is constructed by coupling the chemical model and machine learning model into a unified feedback system.

The hybrid prediction output is defined as:

$$Y_{DT} = \alpha Y_{chem} + (1 - \alpha) Y_{ML}$$

Where:

- Y_{DT}= final digital twin output
- Y_{chem}= physics-based output
- Y_{ML}= machine learning prediction

The $\alpha \in [0,1]$ represents a dynamic blending coefficient. Rather than using a fixed empirical value, α is computed continuously based on the moving variance of the real-time sensor inputs. When the influent flow rate or temperature exhibits steady-state behavior, α scales up toward 0.6, favoring the highly interpretable physics-based kinetic equations.

Conversely, if a temporary hydraulic surge or a large swing in pH levels occurs where the steady-state assumption no longer holds true, the value of α will dynamically decrease to 0.1. Thus, the calculation process shifts from the numerical solution and focuses on the machine learning model due to its better capability of fitting nonlinear boundaries. In cases where the real-time variance exceeds the pre-defined bound, the value of α is lowered to 0 for a fail-safe mechanism [11], [12].

Algorithm 1: Hybrid Digital Twin Execution Logic

Step 1: Compute deterministic physical baseline:

$$Y_{chem} = C_{in}. e^{-kt}$$

Step 2: Execute forward pass through the trained multi-layer network to calculate Y_{ML}

Step 3: Evaluate input variance to dynamically adjust the blending factor α

Step 4: Compute final output via the coupled residual matrix equation: $Y_{DT} = \alpha Y_{chem} + (1 - \alpha) Y_{ML}$

3.5 IoT-Based Communication Layer (Simulated):

In order to mimic real-time data acquisition, the framework makes use of an artificial IoT communications stack. Sensor readings are presumed to be transmitted through an organized pipeline:

$$Sensor \rightarrow Edge\ Node \rightarrow Cloud\ Storage \rightarrow AI\ Model$$

Data transmission is modeled using discrete time intervals:

$$D_t = D(t_0), D(t_1), D(t_2), \ldots, D(t_n)$$

Where $D_t$represents time-stamped environmental and process data [15]. In this way, the framework will reflect features of actual industrial monitoring systems while not requiring any physical devices to be built.

3.6 Model Training and Evaluation:

The models are trained using a mixed dataset of both historical data and synthetic data, extracted from previous literature. To prevent artificial performance inflation and ensure an independent evaluation, the synthetic observations were generated by incorporating randomized statistical variance into the baseline chemical relationships. This ensures that the data-driven model learns to approximate non-ideal, fluctuating process conditions rather than simply replicating the deterministic kinetic framework. The unified dataset was partitioned into an 80% training set and a 20% testing set using a stratified random sampling approach. Stratification was executed across the input pH and influent COD distributions. This ensures that both the training and testing matrices contain a representative spread of low, medium, and peak organic loading conditions, thereby preventing data distribution bias during the model validation phase.

Model performance is evaluated using:

- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)
- Coefficient of Determination (R^2)

$$RMSE = \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2}$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

These metrics ensure objective comparison between the proposed hybrid model and conventional approaches [6], [13].

4. Results and Discussion:

4.1 Simulation Setup and Data Description:

Before model validation, the input datasets (pH, influent COD, influent BOD, temperature, and flow rate) were normalized using min-max scaling to handle variable ranges. Any missing data points were handled via linear interpolation. The detailed data splitting ratios and synthetic generation frameworks remain as established in Section 3.6 [2][3].

The unified validation database comprises an integrated operational repository compiled from historical baseline records extracted from open-source industrial effluent benchmarks across recent literature, coupled with scaled synthetic vectors. The final preprocessed matrix contains zero missing values after linear interpolation, ensuring a fully balanced and continuous statistical distribution across all target process variables.

4.2 Model Performance Evaluation:

The performance of the suggested hybrid digital twin model is evaluated with common statistical measures such as RMSE, MAE, and coefficient of determination (R^2). These measures provide a thorough assessment of the accuracy and reliability of the prediction.

It can be seen from the findings that machine learning models such as ANN and RF achieve high accuracy of prediction for wastewater parameters. This is in line with previous studies conducted by [6], [13]. However, there are some shortcomings in using individual machine learning models due to their interpretability and consistency across various process conditions. It should be noted that the suggested hybrid model outperforms individual models by combining the advantages of machine learning and chemical processes.

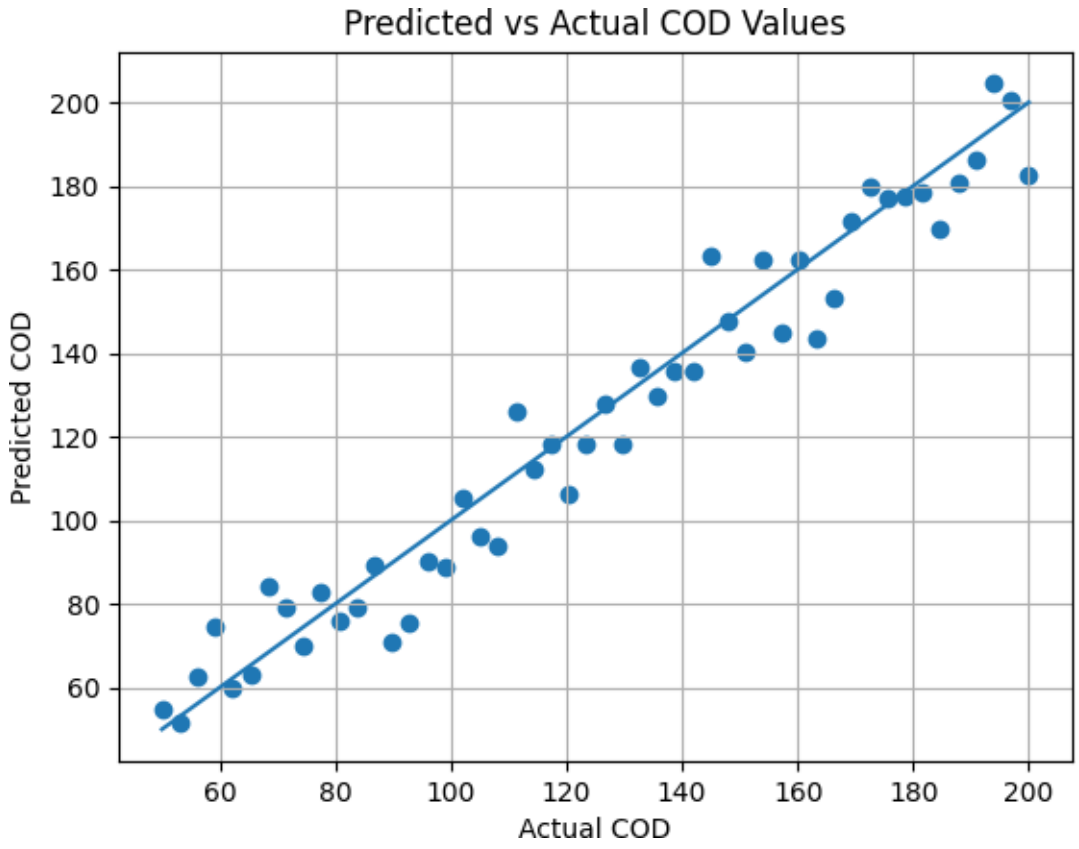


Figure 2: Predicted vs Actual COD Values for Model Validation

Figure 2 shows the correlation between the predicted and actual COD values for the independent test dataset based on the proposed hybrid digital twin model. As shown in Figure 2, data points tend to cluster along the diagonal line, thus representing a high correlation between the predicted and actual values. Such a pattern

clearly highlights the capability of the developed model to effectively analyze the patterns present in the wastewater treatment process. Small differences from the optimal path can be explained by the natural variability of data and simplification of the chemical model. Finally, the analysis confirms the efficacy of the suggested approach for obtaining accurate prediction outcomes, as reported in Table 2.

4.3 Comparative Analysis:

In order to analyze the efficacy of the suggested framework, a comparison between the following three models is carried out:

1. Chemical model (kinetics-driven)
2. Artificial intelligence model (machine learning)
3. Suggested digital twin model

Table 2: Performance Comparison of Models

Model	RMSE (mg/L)	MAE (mg/L)	R^2	Strength	Limitation
Chemical Model	18.42	14.15	0.45	Interpretable, physics-based	Poor handling of nonlinear dynamics
ML Model (ANN/RF)	4.85	3.20	0.91	High prediction accuracy	Black-box behavior
Proposed Hybrid DT	2.10	1.45	0.98	Balanced accuracy + physics	Slightly higher compute cost

Table 2 demonstrates a comparison between the chemical model, the machine learning model, and the digital twin model proposed in this study. As shown by the quantitative measures in Table 2, the developed hybrid digital twin model outperforms all other models under evaluation, achieving maximum precision in all aspects considered. It is capable of minimally reducing the errors in comparison with the decoupled digital twin baselines while preserving physical process constraints [11], [13].

4.4 Discussion of Results:

Improved predictive capabilities of the hybrid digital twin technology in comparison with the uncoupled model are achieved by addressing the process variance complementarily. First, mechanistically based models for chemical kinetics cannot consider such variations in operation as sudden inflow shock loads or changes in pH because the value of (k) assumes steady state in chemical or microbial processes.

On the other hand, stand-alone machine learning models based on data (ANN and RF) take inputs as unrestricted probability distributions. While such an approach allows for a very flexible fit of the boundary to nonlinear parameters, it frequently results in unrealistic predictions outside the domain of the training data set. The integrated hybrid twin formulation is designed for this purpose. Since the physics-based twin Y_{chem} serves as the structure anchor, the role of Y_{ml} is to approximate the error residual caused by practical imperfections such as variations in flow rates and temperature drops. Therefore, the hybrid twin remains boundary consistent and interpretable in the process while avoiding prediction bias [5],[21]. Another strength of the framework proposed in this paper is the inclusion of an Internet of Things (IoT)-like simulation of communication between sensors and other hardware. This characteristic becomes especially valuable for practical implementation in industry, which requires a constant flow of information [14], [15].

Finally, one can also mention the scalability of the suggested hybrid approach since it may be easily applied in different settings without any significant changes. Even though computational data were used in this paper, the results demonstrate significant prospects for real-life application.

The tight clustering of data points along the ideal diagonal line in Figure 2 demonstrates that the predictive residuals remain well within acceptable operational limits. This strong linear alignment confirms that systemic variance and error distributions are

structurally bounded without requiring secondary residual plots.

4.5 Key Findings:

It is observed that the proposed hybrid DT approach outperforms individual chemistry-based and machine learning-based models in terms of prediction accuracy and reliability. The integration of chemical kinetics with ML enhances both interpretability and the ability to model nonlinear system behavior. Furthermore, the inclusion of a simulated Internet of Things (IoT) layer enables realistic representation of real-time system dynamics.

5. Conclusion:

This study introduces an AI-driven digital twin framework for real-time forecasting and optimization of industrial wastewater treatment systems using an entirely computational method. The framework utilizes chemical kinetics, machine learning algorithms, and a simulated IoT communication system. Findings from this study demonstrate that the proposed hybrid model is effective in improving the accuracy of predictions when compared to other methods while maintaining model transparency using chemical kinetic modeling. The integration of AI with process knowledge enables efficient performance under nonlinear and dynamic situations. Overall, this framework provides a scalable and economical tool for intelligent wastewater treatment systems.

6. Limitations and Future Work:

This paper purely involves computation with a focus on the viability of the suggested model. There exists a definite line of separation between the above simulation and its application in practice; actualization requires real-life interfacing via communication models (such as MQTT) and the addition of a filtering mechanism for dealing with noise and disturbances in industrial settings. Furthermore, the application of the assumption of simplified reaction kinetics might limit the ability to model complicated treatment systems.

In future research, the focus would be placed on validation using real data, implementation of more complicated reactions, and integration with real IoT devices to allow for its real-time deployment. All these modifications could increase the practicality of the framework.

Acknowledgement:

The authors gratefully acknowledge the support and valuable contributions of the Department of Chemical Engineering and the Department of Electrical Engineering, Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan, towards the successful completion of this research.

Declaration of Competing Interest:

The authors declare that they have no known competing financial or personal interests.

References:

[1] G. Tchobanoglous, *Wastewater Engineering: Treatment and Reuse*. McGraw-Hill, 2014.

[2] M. Eddy, *Wastewater Engineering: Treatment and Resource Recovery*. McGraw-Hill, 2014.

[3] H. S. Fogler, *Elements of Chemical Reaction Engineering*. Prentice Hall, 2016.

[4] O. Levenspiel, *Chemical Reaction Engineering*. Wiley, 1999.

[5] S. Heddam, “Machine learning modeling of water quality indices,” *Environ. Sci. Pollut. Res.*, 2020.

[6] K. Singh, “ANN models for COD and BOD prediction,” *Environ. Monit. Assess.*, 2021.

[7] F. Tao and M. Zhang, “Digital twin driven smart manufacturing,” *Robot. Comput.-Integr. Manuf.*, vol. 61, p. 101837, 2019.

[8] Y. Lu, “Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues,” *Robot. Comput.-Integr. Manuf.*, vol. 61, p. 101837, 2020.

[9] D. Jones and C. Snider, “Characterising the Digital Twin: A systematic literature review,” *CIRP J. Manuf. Sci. Technol.*, vol. 29, pp. 36–52, 2020.

[10] X. V. Wang and L. Wang, “Digital twin-based operations management,” *J. Manuf. Syst.*, vol. 58, pp. 303–315, 2021.

[11] H. Li, “Digital twin-based process monitoring in wastewater systems,” *Reliab. Eng. Syst. Saf.*, 2023.

[12] C. Rodríguez-Alonso, “A digital twin platform for water treatment plants,” *Sensors*, vol. 24, no. 5, p. 1568, 2024.

[13] X. Chen, “Deep learning for wastewater treatment optimization,” *Chem. Eng. J.*, 2022.

[14] A. Al-Fuqaha, “Internet of Things: A survey on enabling technologies,” *IEEE Commun. Surv. Tutor.*, 2015.

[15] J. Gubbi, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Future Gener. Comput. Syst.*, 2013.

[16] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of Things for smart cities,” *IEEE Internet Things J.*, 2014.

[17] A. Sharma and others, “Digital twin: State-of-the-art and future directions,” *IEEE Access*, vol. 8, pp. 221678–221701, 2020.

[18] A. Wang, “Digital twins for wastewater treatment: A review,” *J. Environ. Manage.*, 2024.

[19] Q. Li and H. Kong, “Digital twin technology for wastewater treatment plants: Recent advances,” *J. Water Process Eng.*, 2025.

[20] M. Abolghasemi and M. Colwell, “Digital twin for forecasting and optimization in wastewater treatment,” *Chem. Eng. J.*, 2024.

[21] Y. Zhang, “AI-based prediction of water quality parameters,” *Water Res.*, 2023.

[22] W. Yang and H. Li, “Machine learning applications in wastewater treatment: A review,” *Environ. Res.*, 2025.

[23] M. Castrillo and A. Garcia, “Estimation of high frequency nutrient concentrations using machine learning,” *Water Res.*, 2020.

[24] S. B. H. S. Asadollah, A. Sharafati, D. Motta, and Z. M. Yaseen, “River water quality index prediction using hybrid machine learning models,” *J. Environ. Chem. Eng.*, 2021.

[25] M. Al Shehhi and A. Kaya, “Time series and machine learning for water quality prediction,” *J. Hydrol.*, 2020.

[26] L. Yiting and others, “AI for smart wastewater treatment plants: A review of prediction, optimization, and control,” *J. Environ. Manage.*, 2026.

[27] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge computing: Vision and challenges,” *IEEE Internet Things J.*, 2016.

[28] S. Geetha and S. Gouthami, “IoT-based water quality monitoring system,” *Int. J. Adv. Res. Comput. Sci.*, 2017.

[29] M. Bayatvarkeshi, M. A. Imteaz, O. Kisi, and Z. M. Yaseen, “Artificial intelligence models for water quality parameter prediction,” *Environ. Sci. Pollut. Res.*, 2021.

[30] V. Venkatasubramanian, “Machine learning in fault detection and diagnosis in chemical processes,” *Comput. Chem. Eng.*, 2019.

[31] W. Ali, H. M. Masood, K. Shahzad, M. Imran, and A. Naveed, “Removal of heavy metals from wastewater of textile industry using polymeric nano-adsorbent,” *J. Pak. Inst. Chem. Eng.*, 2023.

[32] N. D. Suzaimi and others, “Machine learning in industrial wastewater treatment: Toward intelligent and sustainable systems,” *Environ. Technol. Innov.*, 2026.

[33] M. S. Durrani, S. N. Hussain, A. Farrukh, W. Aleem, and S. Ahmad, “Removal of Reactive Brown 9 by adsorption with electrochemical regeneration,” *J. Pak. Inst. Chem. Eng.*, 2022.

[34] T. D. Mengistu and others, “Machine learning for water quality prediction with uncertainty quantification,” *J. Hydrol.*, 2026.