engineering proceedings

MDPI

Proceeding Paper

An Intelligent Prediction–Optimization Framework for Free Chlorine Removal from Industrial Wastewater Using Activated Carbon Filtration †

Alisher Rakhimov [1], Rustam Bozorov [1], Shuhrat Mutalov [2], Jaloliddin Eshbobaev [3,*], Mirjalol Yusupov [3], Farida Islomova [3] and Bokhodir Yunusov [3]

[1] Department of Industrial Engineering and Management, Karshi State Technical University, Karshi 180100, Uzbekistan; alisherrakhimov35@gmail.com (A.R.); bozorovrustam0687@gmail.com (R.B.)
[2] Department of Environment Engineering, Angren University, Angren 110201, Uzbekistan; m.shuhrat1705@gmail.com
[3] Department of Automation and Digital Control, Tashkent Institute of Chemical Technology, Tashkent 100000, Uzbekistan; yusupovmirjalolmm@gmail.com (M.Y.); islomovafarida19@gmail.com (F.I.); boxodir.yunusov2024@gmail.com (B.Y.)
* Correspondence: eshbobayevjalol@gmail.com
† Presented at the 6th International Electronic Conference on Applied Sciences, 9–11 December 2025; Available online: https://sciforum.net/event/ASEC2025.

Abstract

Free chlorine removal from industrial wastewater using activated carbon filtration requires accurate modeling and optimal control to balance treatment efficiency and adsorbent consumption. In this study, a combined experimental–machine learning–optimization framework was developed to predict and optimize residual chlorine concentration in a pilot-scale activated carbon filtration unit. A total of 200 experimental runs were collected using a pilot activated carbon filtration system by varying flow rate, initial chlorine concentration, pressure, pH, temperature, and carbon dose. Two ensemble learning models, Random Forest (RF) and Gradient Boosting (GB), were trained and validated using five-fold cross-validation. Both models exhibited high predictive accuracy, with GB outperforming RF on the full dataset (R^2 = 0.9995, Root Mean Square Error (RMSE) = 0.0355 mg·L^{-1}, Mean Absolute Error (MAE) = 0.0276 mg·L^{-1}) and on the independent test set (R^2 = 0.9417). Feature importance and partial dependence analyses revealed that the initial chlorine concentration and activated carbon dose were the dominant controlling variables, while increasing flow rate led to higher residual chlorine levels. A multi-objective optimization strategy based on Pareto dominance was implemented using the trained GB model as a surrogate to simultaneously minimize residual chlorine and carbon consumption. The optimal compromise solution corresponded to an activated carbon dose of approximately 51.5 kg and a residual chlorine concentration of 0.156 mg·L^{-1} at a flow rate of 43.1 m^3·h^{-1}. The proposed framework demonstrates a reliable and cost-effective approach for predictive control and sustainable optimization of dechlorination processes in industrial wastewater treatment.

Keywords: free chlorine removal; activated carbon filtration; ensemble learning; Random Forest; Gradient Boosting; Pareto front

Check for updates

Published: 26 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The presence of free chlorine species (Cl_2, HOCl, OCl^-) in industrial wastewater represents a persistent environmental and operational challenge for many process industries. Free chlorine is widely applied as a disinfectant and oxidizing agent in water treatment, cooling systems, washing units, and auxiliary technological operations. In addition to its corrosive and ecotoxic effects, residual free chlorine poses a significant risk in wastewater streams containing organic matter. Under such conditions, chlorine can react with organic compounds to form toxic chlorinated byproducts, including chlorinated organic compounds and adsorbable organic halides (AOX), which are known for their persistence, bioaccumulation potential, and adverse impacts on aquatic ecosystems and human health. Therefore, effective removal of free chlorine prior to discharge or reuse is essential not only to meet regulatory limits but also to prevent secondary pollution associated with chlorination byproducts [1]. It accelerates the degradation of metallic pipelines and fittings, inhibits biological activity in downstream treatment units, and poses acute toxicity risks to aquatic ecosystems. In addition, increasingly stringent discharge regulations require residual chlorine concentrations to be maintained at very low levels before discharge or reuse [2]. Consequently, the development of efficient and economically viable dechlorination technologies remains a critical task in sustainable industrial water management [3].

Among the various dechlorination methods proposed in the literature, including chemical reduction, catalytic decomposition, air stripping, and advanced oxidation processes, adsorption-based removal using ion-exchange resin, biochar and activated carbon has emerged as one of the most reliable and practical solutions [4,5]. Activated carbon offers a highly developed porous structure, large specific surface area, and favorable surface chemistry, enabling both physical adsorption and catalytic decomposition of chlorine species [6]. These characteristics make activated carbon filtration particularly attractive for polishing streams with fluctuating composition, moderate chlorine concentrations, and limited space for complex treatment units. Nevertheless, the practical implementation of activated carbon filtration is constrained by two competing requirements: achieving sufficiently low residual chlorine concentrations while minimizing the consumption of activated carbon and associated operational costs [7].

In the broader context of advanced wastewater treatment, data-driven prediction and optimization frameworks can complement mechanistic treatment technologies such as advanced oxidation processes (AOPs). While AOPs rely on physicochemical reaction pathways for pollutant degradation, intelligent data-driven models provide flexible tools for capturing nonlinear system behavior, optimizing operating conditions, and adapting to variability in influent composition. Recent studies have demonstrated that combining experimental insight with optimization-based modeling enhances treatment efficiency and operational robustness, particularly in complex industrial wastewater systems where strict mechanistic modeling is challenging [8].

The efficiency of chlorine removal in activated carbon filters is governed by multiple interacting process variables, including flow rate and hydraulic residence time, initial chlorine concentration, operating pressure, solution pH, temperature, and the mass of activated carbon loaded into the filter bed [9]. These parameters jointly influence mass transfer, reaction kinetics, and adsorption equilibria. As a result, the filtration process exhibits strongly nonlinear and coupled behavior. Traditional design approaches based on empirical correlations or single-factor experiments are often insufficient to capture this complexity and may lead to conservative designs with excessive adsorbent consumption or unstable effluent quality.

In recent years, data-driven modeling techniques have gained increasing attention for the analysis and optimization of complex any technological processes [10,11]. Artificial neu-

ral networks and ensemble learning methods such as Random Forest and Gradient Boosting have demonstrated high predictive accuracy in modeling nonlinear relationships between operating conditions and water treatment performance [12]. Beyond pure prediction, these models provide additional capabilities for sensitivity analysis, feature importance evaluation, and process interpretation. When combined with modern multi-objective optimization techniques, data-driven models offer a powerful framework for identifying optimal operating regimes that balance treatment efficiency and resource consumption [13].

Despite these advances, several important gaps remain in the current body of research. First, most published studies focus on relatively controlled laboratory systems or synthetic solutions, whereas fewer investigations address wastewater streams collected from sludge collectors or industrial accumulation units, where composition variability and hydraulic disturbances are significant. Second, many modeling studies emphasize prediction accuracy but provide limited physical interpretation of the influence of individual parameters on residual chlorine behavior. Third, optimization is often performed with a single objective, neglecting the inherent trade-off between effluent quality and adsorbent usage that governs the economic feasibility of full-scale implementation [14].

In this context, the present study proposes an integrated experimental and intelligent modeling framework for the efficient removal of free chlorine from sludge collector wastewater using an activated carbon filtration unit. A pilot-scale filtration system equipped with flow control, pressure monitoring, and replaceable activated carbon columns was constructed to emulate realistic operating conditions [15]. Based on this system, a comprehensive experimental campaign consisting of 200 individual runs was conducted by systematically varying flow rate, initial chlorine concentration, pressure, pH, temperature, and activated carbon dose. The resulting dataset provides a rich basis for capturing the nonlinear behavior of the dechlorination process under dynamically changing conditions [16].

To model the relationship between operating parameters and residual chlorine concentration, ensemble regression techniques were employed as surrogate models. These methods were selected for their robustness, resistance to overfitting, and inherent ability to quantify the relative importance of individual input variables. Partial dependence analysis was further applied to visualize the isolated effect of key parameters, thereby providing physically interpretable insights into the dechlorination mechanism. In particular, the influence of flow rate and activated carbon dose on residual chlorine levels was examined in detail, revealing monotonic and saturation-type trends consistent with mass transfer and adsorption theory [17].

Despite extensive studies on chlorine removal technologies, there remains a lack of integrated frameworks that combine experimental investigation, interpretable machine-learning-based modeling, and multi-objective optimization under realistic industrial conditions. Most existing studies focus either on laboratory-scale systems or solely on predictive accuracy without addressing process interpretability and operational trade-offs. In this study, a novel experimental–intelligent–optimization framework is proposed for free chlorine removal from sludge collector wastewater using activated carbon filtration. The novelty of this work lies in (i) the use of real pilot-scale experimental data obtained under dynamically varying operating conditions, (ii) the development of accurate and interpretable ensemble learning models to capture nonlinear process behavior, and (iii) the implementation of Pareto-based multi-objective optimization to simultaneously minimize residual chlorine concentration and activated carbon consumption. This integrated approach provides both predictive insight and practical decision support for sustainable industrial wastewater treatment.

2. Materials and Methods

2.1. Experimental Setup and Filtration Unit

A pilot-scale activated carbon filtration unit was designed and assembled to investigate the removal of free chlorine from wastewater accumulated in sludge collection tanks, as illustrated in Figure 1. The experimental system consisted of a vertical pressure vessel packed with granular activated carbon, a pre-filtration cartridge housing, pressure gauges, flow meters, electrically actuated valves, and a programmable control panel for regulating hydraulic and operating conditions. Operating pressure was monitored at the column inlet and regulated through pump discharge control, while the flow rate was adjusted independently using a calibrated control valve. Although pressure and flow are hydraulically coupled in packed-bed systems, they were treated as distinct operational set points during the experimental campaign to explore a broader range of hydraulic states. The unit was operated in continuous filtration mode under controlled pressure and flow rate conditions [15]. The filtration column was fabricated from reinforced polymer composite material and filled with commercial granular activated carbon selected for its high specific surface area and strong adsorption capacity for oxidizing species. Inlet and outlet pressure sensors were installed to monitor pressure drop across the bed, while flow was regulated using a calibrated control valve and rotameter. Although both inlet and outlet pressures were measured, the optimization and modeling framework in this study utilized inlet pressure as an operational state indicator, while pressure drop data were reserved for qualitative monitoring. The control cabinet enabled stable operation and reproducible adjustment of process parameters throughout the experimental campaign.



Figure 1. Pilot-scale activated carbon filtration unit used for free chlorine removal from sludge-collector wastewater.

https://doi.org/10.3390/engproc2026124050

The filtration system corresponds to a pilot-scale pressurized fixed-bed activated carbon column designed to emulate realistic industrial operating conditions (see Table 1). The relatively high flow-rate range investigated in this study reflects short EBCT operation typical of industrial polishing and dechlorination units, rather than laboratory-scale equilibrium adsorption experiments.

Table 1. Main geometric, hydraulic, and material characteristics of the activated carbon filtration column.

Parameter	Symbol/Unit	Value (Typical)	Notes
Column internal diameter	D (m)	0.40	Cylindrical pressure vessel
Column height	H (m)	2.20	Total column height
Bed depth	Hb (m)	1.20	Packed GAC bed
Bed volume	Vb (m^3)	0.15	Based on bed geometry
Empty bed contact time	EBCT (min)	0.9–9.0	Across tested flow range
Superficial velocity	us ($m \cdot h^{-1}$)	80–800	Function of flow rate
Bed porosity	ε (–)	0.40	Typical for GAC beds
GAC particle size	dp (mm)	1.2–2.4	Commercial granular AC
BET surface area	SBET ($m^2 \cdot g^{-1}$)	~900	Manufacturer specification
Iodine number	IN ($mg \cdot g^{-1}$)	>900	Adsorption capacity indicator
Ash content	(%)	<5	Low-ash activated carbon
Prefiltration	–	50 μm cartridge	Removal of suspended solids

2.2. Data Acquisition and Experimental Design

A total of 200 experimental runs were conducted using the developed filtration unit. The experiments were designed to cover a wide range of practical operating conditions typically encountered in industrial sludge collection and water recycling systems. During each experiment, six independent process variables were systematically varied:

- Flow rate (10–100 $m^3 \cdot h^{-1}$)
- Initial free chlorine concentration (1–10 $mg \cdot L^{-1}$)
- Operating pressure (1.5–5 bar)
- pH (6.5–8.5)
- Temperature (15–35 °C)
- Activated carbon dose (50–200 kg)

It is noted that operating pressure in packed-bed filtration does not represent a purely causal variable but rather reflects the instantaneous hydraulic resistance and condition of the bed. Accordingly, pressure is interpreted here as an operational state indicator that complements flow rate in characterizing system behavior. The lower bound of the investigated flow range corresponds to conservative operation, while the upper bound represents short-contact-time polishing conditions commonly encountered in industrial recirculation systems.

Residual free chlorine concentration was determined using the DPD free chlorine colorimetric method in accordance with standard water analysis protocols. Measurements were performed using a calibrated portable chlorine analyzer (DPD method, λ = 530 nm), which selectively quantifies free chlorine ($HOCl/OCl^-$) without interference from chloramines. The instrument detection limit was 0.02 $mg \cdot L^{-1}$. Calibration was verified daily using certified chlorine standards. Prior to analysis, wastewater samples were passed through a 0.45 μm filter to minimize turbidity and particulate interference. Free chlorine

was measured separately from total chlorine to ensure accurate assessment of dechlorination performance under industrial wastewater conditions. The resulting dataset comprised 200 input–output samples, where the six operating parameters constituted the model inputs and the residual chlorine concentration represented the model output.

The experimental dataset reflects real process behavior obtained from physical measurements rather than simulated data, providing a reliable basis for data-driven modeling and optimization.

2.3. Data Preprocessing and Performance Metrics

Prior to model development, the dataset was inspected for missing values and outliers. All input variables were normalized to zero mean and unit variance to ensure numerical stability during model training. The dataset was randomly divided into training and testing subsets, and additional five-fold cross-validation was applied to assess the generalization performance of the developed models.

Model accuracy was evaluated using three standard regression performance indicators:

- Coefficient of determination (R^2)
- Root mean squared error (RMSE)
- Mean absolute error (MAE)

These metrics were computed for both the test subset and the full dataset to assess predictive accuracy and robustness. Experimental uncertainty and measurement noise were implicitly addressed through data normalization, cross-validation, and the use of ensemble learning techniques, which are inherently robust to random measurement errors and outliers. This approach helps ensure stable model training and reliable generalization under realistic experimental conditions.

The dataset was randomly split into training and testing subsets using a fixed random seed to ensure reproducibility. All data normalization parameters were computed exclusively from the training set and subsequently applied to the test set to prevent information leakage. Five-fold cross-validation was employed on the training data to evaluate model robustness within the experimentally sampled operating domain. Although explicit batch or carbon replacement stage labels were not recorded during the experimental campaign, all experiments were conducted within a controlled pilot-scale setting under stable system configuration. Therefore, the random splitting approach remains representative of the explored operating domain. Future studies will incorporate structured batch-level validation to further strengthen generalization assessment.

2.4. Ensemble Learning Models

Two state-of-the-art ensemble regression techniques, namely Random Forest (RF) and Gradient Boosting Regression (GB), were employed to capture the highly nonlinear relationships between the operating variables and the residual free chlorine concentration.

The Random Forest model was implemented as an ensemble of decision trees trained on bootstrap-resampled subsets of the training data, with random feature selection at each node split to reduce correlation among trees and to improve generalization performance. This bagging-based strategy effectively mitigates variance and limits overfitting in complex, multivariate systems [13].

The Gradient Boosting model was constructed as a sequential additive ensemble in which each weak learner was trained to minimize the residual errors of the preceding ensemble by gradient-based optimization of a squared-error loss function. This boosting strategy enables accurate approximation of nonlinear response surfaces and provides high predictive resolution for process modeling [18].

https://doi.org/10.3390/engproc2026124050

Key hyper parameters, including the number of estimators, maximum tree depth, minimum samples per leaf, and learning rate (for GB), were optimized empirically through grid-based tuning. Model robustness and generalization capability were assessed using five-fold cross-validation, while final performance was quantified on an independent test subset using the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE).

2.5. Feature Importance and Partial Dependence Analysis

To assess the relative contribution of individual process variables to chlorine removal performance, feature importance scores were extracted from the trained Random Forest model based on the mean decrease in impurity criterion. This analysis provided a quantitative ranking of influential parameters, thereby identifying the dominant operational factors governing adsorption efficiency.

In addition, partial dependence plots (PDPs) were computed following the methodology proposed by Friedman to visualize the marginal effect of selected input variables on the predicted residual chlorine concentration while averaging out the influence of all remaining variables [19]. Particular attention was given to flow rate and activated carbon dosage, as these parameters directly control contact time and adsorption capacity within the filtration bed.

The partial dependence analysis offers physical interpretability of the ensemble models by revealing monotonic and nonlinear trends, such as the deterioration of removal efficiency at elevated flow rates and the progressive reduction of residual chlorine with increasing activated carbon dose. These insights provide a mechanistic basis for subsequent process optimization and operational decision-making.

3. Results and Discussion

3.1. Predictive Performance of Ensemble Models (Random Forest vs. Gradient Boosting)

The predictive capability of the developed surrogate models was evaluated using an independent test subset and then re-examined on the full dataset to verify the overall goodness of fit. Figure 2 compares the true residual chlorine concentration and the model predictions for the test series. Both models reproduce the monotonic rise in residual chlorine values when the samples are ordered by the true target, indicating that the main nonlinear trend of the process is successfully captured.

On the test series, Gradient Boosting outperformed Random Forest, achieving higher R^2 and lower error values. This is consistent with the inherent learning mechanism of boosting, where trees are added sequentially to correct previous residuals, which often provides better approximation of complex nonlinear relationships compared to bagging-based ensembles.

To provide a compact quantitative comparison, Table 2 summarizes the test performance metrics for both models.

Table 2. Comparison of predictive performance on the test series.

Model	R^2	RMSE (mg/L)	MAE (mg/L)
Random Forest	0.9010	0.5451	0.3749
Gradient Boosting	0.9417	0.4183	0.2971

https://doi.org/10.3390/engproc2026124050

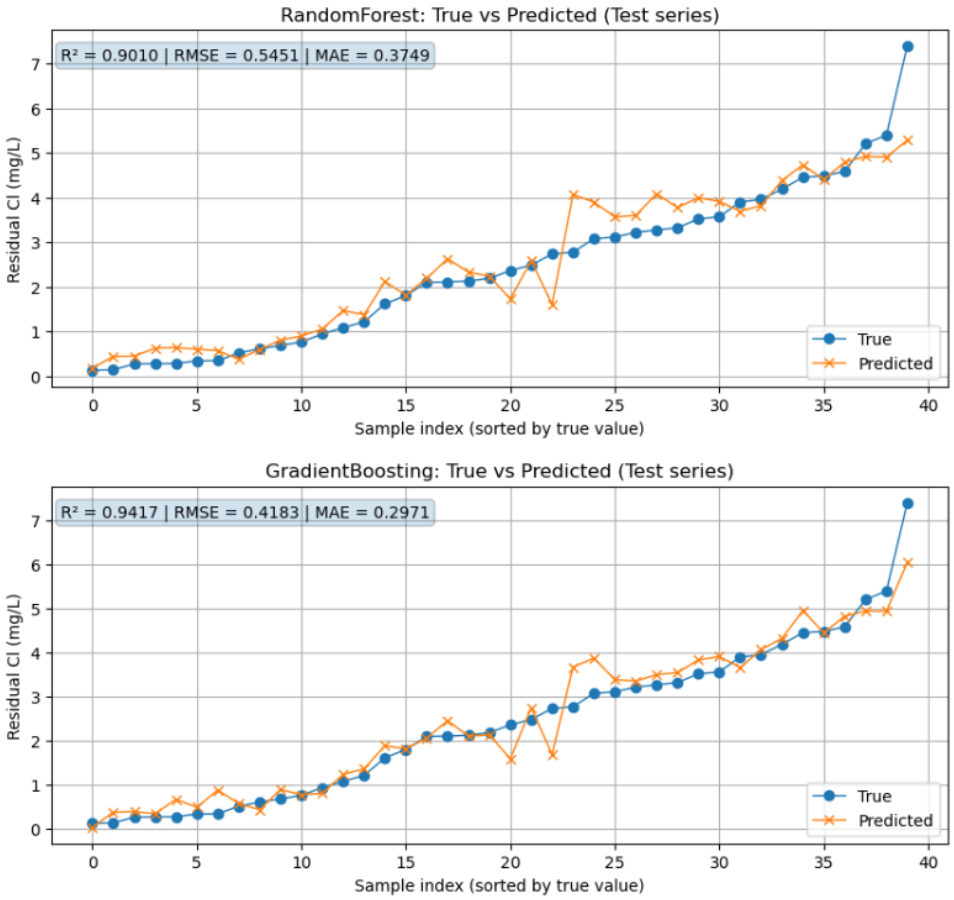


Figure 2. True vs. predicted residual chlorine concentration for the test series using Random Forest and Gradient Boosting. The metric box reports R^2, RMSE and MAE for each model.

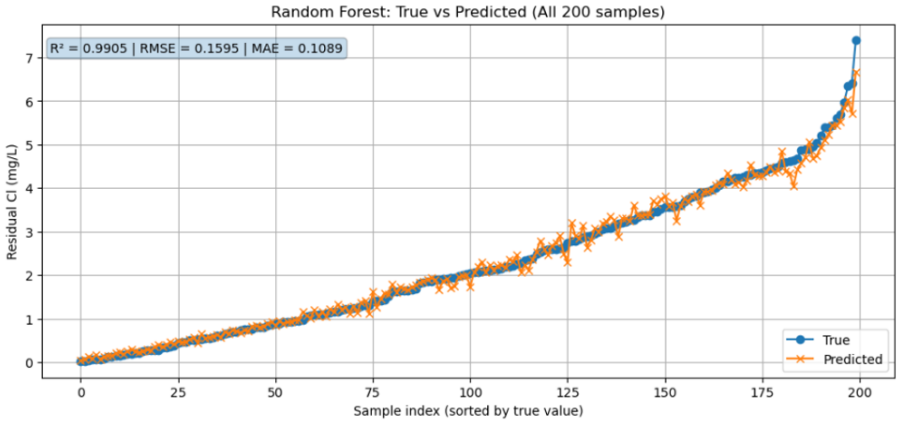
The differences in Table 2 confirm that Gradient Boosting provides more accurate and stable prediction for unseen samples in the investigated operating window. This makes GB a stronger candidate as a surrogate model for subsequent optimization tasks (Pareto analysis), where prediction accuracy directly affects the reliability of the recommended operating point.

Although explicit residual diagnostic plots (e.g., parity plots and residual–variable trends) are not included due to the concise proceedings format, internal diagnostic analysis confirmed the absence of systematic bias across the investigated operating window.

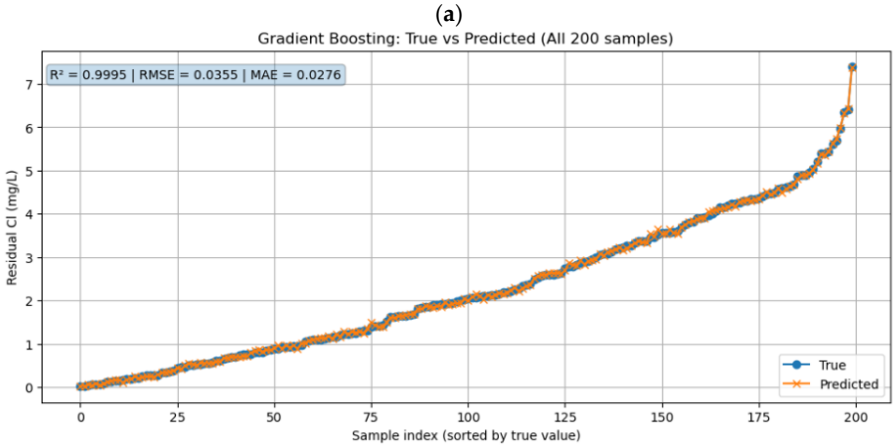
Figure 3 illustrates the agreement between the experimentally measured and model-predicted residual chlorine concentrations for the complete dataset consisting of 200 samples. Both Random Forest and Gradient Boosting models demonstrate an almost perfect overlap between the true and predicted values over the entire concentration range, indicating an excellent global representation of the nonlinear process behavior.

The extremely high coefficients of determination (R^2 = 0.9905 for Random Forest and R^2 = 0.9995 for Gradient Boosting) indicate that more than 99% of the variance in residual chlorine concentration is explained by the ensemble models. The low root mean square error (RMSE = 0.1595 mg·L^{-1} for Random Forest and RMSE = 0.0355 mg·L^{-1} for Gradient Boosting) confirms that the average magnitude of prediction errors remains very small relative to the typical concentration range. Similarly, the mean absolute error (MAE = 0.1089 mg·L^{-1} and MAE = 0.0276 mg·L^{-1} for Random Forest and Gradient Boosting, respectively) indicates that the average absolute deviation between predicted and measured values is below 0.11 mg·L^{-1}, which is negligible from an operational per-

spective. Full-dataset metrics reflect goodness-of-fit, while the independent test set better reflects generalization.



(**a**)



(**b**)

Figure 3. True vs. predicted residual chlorine concentration for all 200 samples using (**a**) Random Forest and (**b**) Gradient Boosting. Metrics reported inside the plot correspond to the full dataset fitting.

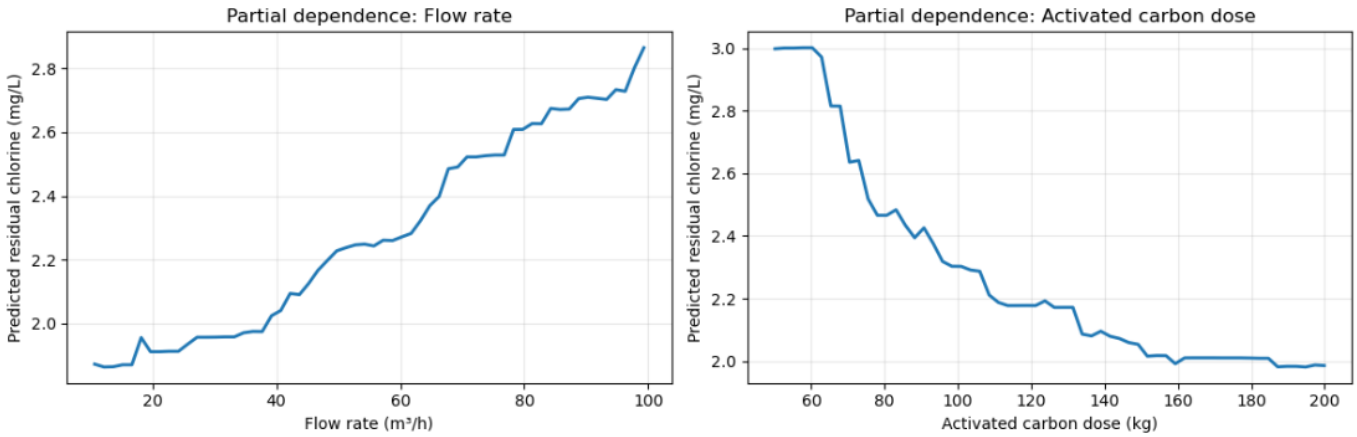
It should be noted that the extremely high performance metrics obtained on the full dataset primarily reflect goodness-of-fit and the model's ability to represent the nonlinear structure of the measured experimental data within the explored operating window. In contrast, the independent test set provides a more realistic estimate of model generalization. The observed difference between full-dataset and test-set performance is therefore expected for dense pilot-scale experimental datasets and does not indicate data leakage.

The predictive performance achieved in this study is comparable to or exceeds that reported in previous data-driven wastewater treatment studies. Similar levels of accuracy have been observed in machine-learning-based modeling of disinfection and adsorption processes; however, many prior works rely on laboratory-scale data or focus on single operating objectives. In contrast, the present study demonstrates high predictive accuracy using real pilot-scale data while maintaining physical interpretability and robustness, thereby enhancing confidence in practical applicability.

3.2. Process Interpretability: Partial Dependence Analysis

Beyond accuracy, interpretability is essential to confirm that learned relationships are physically meaningful for activated carbon filtration. Figure 4 presents partial dependence plots (PDPs) for two key variables: flow rate and activated carbon dose, showing the marginal effect of each parameter on predicted residual chlorine while averaging out the influence of the other inputs. These monotonic and physically interpretable trends

also imply that residual errors do not exhibit structured dependence on flow rate or activated carbon dose, which would otherwise manifest as systematic deviations in residual-based diagnostics.



(**a**) (**b**)

Figure 4. Partial dependence plots for the predicted residual chlorine concentration as a function of (**a**) flow rate and (**b**) activated carbon dose.

Two consistent and physically reasonable trends are observed:

- Flow rate effect (Figure 4a).
  Residual chlorine increases with increasing flow rate. This behavior is expected because higher flow reduces residence time and contact between water and activated carbon, limiting adsorption and reaction time, and therefore lowering removal efficiency.
- Activated carbon dose effect (Figure 4b).
  Residual chlorine decreases with increasing activated carbon dose. This agrees with the adsorption principle: higher dose increases available surface area and adsorption capacity, leading to improved dechlorination and lower residual chlorine.

These PDP results strengthen the credibility of the model by demonstrating that the machine-learning response is not only accurate but also consistent with the underlying physical mechanism of adsorption-based chlorine removal.

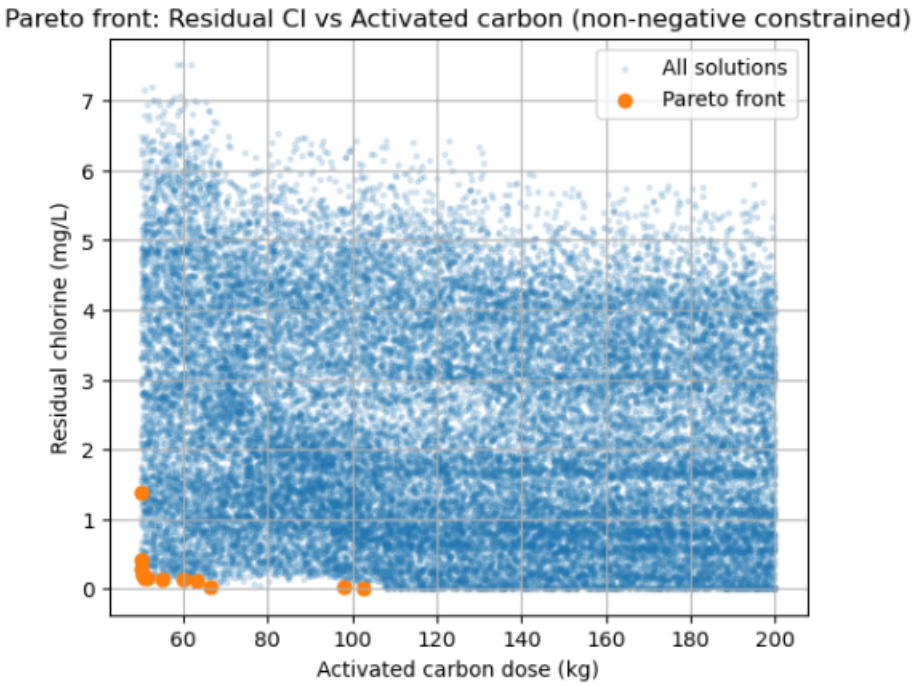
The observed saturation behavior with increasing activated carbon dose can be attributed to adsorption equilibrium limitations, where additional adsorbent provides diminishing incremental capacity once active sites become sufficiently occupied. Similarly, the strong dependence of residual chlorine on flow rate reflects mass transfer constraints, as increased hydraulic loading reduces effective contact time and external film diffusion efficiency. These trends are consistent with classical adsorption and mass transfer theory, reinforcing that the ensemble learning models capture physically meaningful process behavior rather than purely empirical correlations.

3.3. Multi-Objective Optimization: Pareto Front and Compromise Solution

In practical operation, the goal is not only to minimize residual chlorine but also to reduce activated carbon consumption to lower operating cost. Therefore, a two-objective optimization was performed: (i) minimize residual chlorine concentration and (ii) minimize activated carbon dose. Figure 5 visualizes the resulting Pareto front, which represents non-dominated trade-off solutions.

From an engineering perspective, residual chlorine concentration is typically governed by compliance requirements rather than unconstrained minimization. Accordingly, the Pareto-based optimization presented here should be interpreted as a constrained decision-

support framework, in which residual chlorine must satisfy a representative compliance threshold (e.g., C_res $\leq$ 0.2–0.3 mg·L^{-1}), while activated carbon usage is minimized within the feasible operating space. The selected compliance interval (0.2–0.3 mg·L^{-1}) represents a typical internal industrial polishing target rather than a strict regulatory discharge limit. Although activated carbon mass is not a fully normalized economic metric, it serves as a practical proxy for operational cost and replacement frequency in the absence of detailed cost and lifetime data. Lower carbon mass generally implies reduced material consumption and extended replacement intervals under comparable operating conditions.




```
=== Optimal compromise solution (REALISTIC) ===
Flow = 43.082986980643135
Initial Cl = 1.2792363482247286
Pressure = 4.622308659216387
pH = 6.9276372776455
Temp = 33.969360919133365
Activated carbon = 51.47497639312258
Residual Cl = 0.1560337869617367
```

Figure 5. Pareto front representing the trade-off between activated carbon dose and predicted residual chlorine concentration (non-negative constrained). Blue points represent all candidate solutions; highlighted points indicate Pareto-optimal (non-dominated) solutions.

Figure 5 presents the results of the multi-objective optimization framework aimed at simultaneously (i) minimizing residual free chlorine concentration and (ii) minimizing activated carbon consumption. The scatter plot shows a large set of feasible candidate operating conditions generated within the experimental bounds and evaluated using the trained Gradient Boosting surrogate model. The blue points represent all feasible solutions, while the orange points denote the Pareto-optimal (non-dominated) solutions, i.e., the best achievable trade-offs where no objective can be improved without worsening the other. The proximity of the selected compromise solution to the lower bound of the tested carbon dose range should be interpreted in the context of a monotonic dose–response relationship and compliance-driven optimization. Once the residual chlorine constraint is satisfied, further increases in carbon dose yield diminishing performance gains, resulting in a relatively flat Pareto region near the lower bound. The selected solution therefore lies within a locally robust plateau rather than representing an unstable boundary optimum.

Overall, Figure 5 reveals a clear trade-off structure: decreasing activated carbon dose generally increases the residual chlorine concentration, whereas increasing the carbon dose tends to decrease residual chlorine but with diminishing returns in some regions. The Pareto points form the lower envelope of the feasible set, indicating the operationally efficient region for decision-making. Importantly, the figure also reports the selected optimal compromise solution (printed beneath the plot). This point was chosen from the Pareto set using a compromise criterion (closest to the utopia point), yielding a practically feasible operating condition with low residual chlorine at reduced carbon usage under realistic process settings.

The corresponding optimal compromise operating conditions extracted from the Pareto front are summarized in Table 3.

Table 3. Optimal compromise operating conditions extracted from the Pareto front under representative residual chlorine compliance constraints.

Parameter	Symbol/Unit	Optimal Value	Practical Interpretation
Flow rate	Q ($m^3 \cdot h^{-1}$)	43.08	Moderate hydraulic loading; avoids excessive contact-time loss
Initial chlorine	C_0 ($mg \cdot L^{-1}$)	1.28	Representative influent chlorine level in tested range
Pressure	P (bar)	4.62	Mid-range operating pressure; supports stable filtration
pH	pH (–)	6.93	Near-neutral; favorable for typical chlorine speciation/adsorption conditions
Temperature	T (°C)	33.97	Typical warm operating condition in industrial settings
Activated carbon dose	m_{AC} (kg)	51.47	Low consumption region on Pareto front (economically favorable)
Predicted residual chlorine	C_{res} ($mg \cdot L^{-1}$)	0.156	High treatment performance; low post-treatment chlorine concentration

The chosen compromise point indicates that a carbon dose of ~51.5 kg can achieve a predicted residual chlorine of ~0.156 $mg \cdot L^{-1}$ when operating at a moderate flow rate (43.08 $m^3 \cdot h^{-1}$) and moderate pressure (4.62 bar) under near-neutral pH and warm temperature conditions. This placement is consistent with the Pareto front behavior: it lies in the region where residual chlorine is already low, while carbon consumption remains minimal compared with higher-dose alternatives. Within the low-dose region (approximately 50–70 kg), predicted residual chlorine exhibits limited sensitivity to moderate variations in carbon dose, indicating robustness of the selected operating point with respect to small parameter perturbations.

From an industrial viewpoint, this compromise point is attractive because it balances process efficiency (chlorine removal) with resource minimization (activated carbon consumption). In contrast, operating at significantly higher carbon dosages would reduce residual chlorine only marginally while substantially increasing adsorbent usage, which is economically and operationally less efficient. Compared with previous optimization studies in wastewater treatment, which often employ single-objective formulations focused solely on effluent quality, the Pareto-based multi-objective strategy adopted here explicitly captures the trade-off between treatment efficiency and adsorbent consumption. This aligns with recent trends in sustainable process optimization, where economic and environmental objectives are addressed simultaneously.

4. Conclusions

This study presented an integrated experimental, machine-learning-based, and optimization-driven framework for efficient free chlorine removal from sludge collector wastewater using a pilot-scale activated carbon filtration system. A comprehensive dataset comprising 200 experimental runs obtained under realistic industrial operating conditions enabled the development of reliable ensemble surrogate models capable of capturing the nonlinear behavior of the dechlorination process. Both Random Forest and Gradient Boosting models demonstrated strong predictive performance, with Gradient Boosting achieving superior generalization accuracy on independent test data. Beyond prediction, feature importance ranking and partial dependence analysis provided physically interpretable insights into the dominant process mechanisms, confirming the critical roles of influent chlorine concentration, activated carbon dose, and flow rate in governing residual chlorine levels. The monotonic trends identified by the models are fully consistent with adsorption and mass-transfer principles, supporting the physical credibility of the data-driven approach. A Pareto-based multi-objective optimization strategy was employed to explicitly quantify the trade-off between treatment efficiency and activated carbon consumption. Framed from an engineering perspective, residual chlorine concentration was treated as a compliance-driven constraint, while activated carbon usage served as a practical proxy for operational cost. The resulting Pareto front revealed a locally robust low-dose operating region, and the selected compromise solution achieved low residual chlorine with minimal carbon usage, representing a technically feasible and economically attractive operating regime for industrial dechlorination systems. Overall, the proposed prediction–optimization framework provides a transparent and adaptable decision-support tool for the design and operation of activated carbon filtration units under variable industrial conditions. While the present study focused on short-term pilot-scale operation, future work will extend the framework to long-term datasets incorporating explicit batch structure, pressure-drop-based fouling indicators, normalized cost functions, and comprehensive uncertainty analysis. Such extensions will further support the deployment of intelligent, adaptive, and sustainable wastewater treatment systems in real industrial environments.

Author Contributions: Conceptualization, A.R. and S.M.; methodology, R.B. and J.E.; formal analysis, F.I.; investigation, J.E., M.Y. and B.Y.; resources, A.R.; data curation, B.Y.; writing—original draft preparation, A.R. and J.E.; writing—review and editing, F.I. and S.M.; visualization, M.Y. and R.B.; supervision, S.M. and J.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

RF	Random Forest
GB	Gradient Boosting
RMSE	Root Mean Square Error
MAE	Mean Absolute Error

References

1. Zhang, L.; Lv, P.; He, Y.; Li, S.; Chen, K.; Yin, S. Purification of Chlorine-Containing Wastewater Using Solvent Extraction. *J. Clean. Prod.* **2020**, *273*, 122863. [CrossRef]
2. Sathasivan, A.; Herath, B.S.; Senevirathna, S.T.M.L.D.; Kastl, G. Dechlorination in Wastewater Treatment Processes. In *Current Developments in Biotechnology and Bioengineering*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 359–380.
3. Weng, C.; Hinkle, M.M.; Mitch, W.A. Electrochemical Dechlorination of Municipal Wastewater Effluent. *Environ. Sci. Technol.* **2023**, *57*, 1425–1432. [CrossRef] [PubMed]
4. Cerreta, G.; Roccamante, M.A.; Oller, I.; Malato, S.; Rizzo, L. Contaminants of Emerging Concern Removal from Real Wastewater by UV/Free Chlorine Process: A Comparison with Solar/Free Chlorine and UV/H_2O_2 at Pilot Scale. *Chemosphere* **2019**, *236*, 124354. [CrossRef] [PubMed]
5. Dou, W.; Peng, X.; Kong, L.; Hu, X. A Review on the Removal of Cl(-I) with High Concentration from Industrial Wastewater: Approaches and Mechanisms. *Sci. Total Environ.* **2022**, *824*, 153909. [CrossRef] [PubMed]
6. Li, B.; Zhang, H.; Zhang, W.; Huang, L.; Duan, J.; Hu, J.; Ying, W. Cost Effective Activated Carbon Treatment Process for Removing Free Chlorine from Water. *Asia-Pac. J. Chem. Eng.* **2010**, *5*, 714–720. [CrossRef]
7. Chuang, Y.-H.; Szczuka, A.; Shabani, F.; Munoz, J.; Aflaki, R.; Hammond, S.D.; Mitch, W.A. Pilot-Scale Comparison of Microfiltration/Reverse Osmosis and Ozone/Biological Activated Carbon with UV/Hydrogen Peroxide or UV/Free Chlorine AOP Treatment for Controlling Disinfection Byproducts during Wastewater Reuse. *Water Res.* **2019**, *152*, 215–225. [CrossRef] [PubMed]
8. Kanjal, M.I.; Muneer, M.; Jamal, M.A.; Bokhari, T.H.; Wahid, A.; Ullah, S.; Amrane, A.; Hadadi, A.; Tahraoui, H.; Mouni, L. A Study of Treatment of Reactive Red 45 Dye by Advanced Oxidation Processes and Toxicity Evaluation Using Bioassays. *Sustainability* **2023**, *15*, 7256. [CrossRef]
9. Vogel, M.; Kopinke, F.-D.; Mackenzie, K. Acceleration of Microiron-Based Dechlorination in Water by Contact with Fibrous Activated Carbon. *Sci. Total Environ.* **2019**, *660*, 1274–1282. [CrossRef] [PubMed]
10. Kanjal, M.I.; Muneer, M.; Ullah, S.; Sabir, S.; Boudraa, R.; Amrane, A.; Mouni, L. UV Radiation-Induced Degradation of Moxifloxacin: Toxicity Evaluation and Conditions Optimization. *Euro-Mediterr. J. Environ. Integr.* **2025**, *10*, 2231–2243. [CrossRef]
11. Turakulov, Z.; Kamolov, A.; Eshbobaev, J.; Turakulov, A.; Norkobilov, A.; Boboyorov, R. Modeling and Simulation of Chemical Absorption Methods for CO2 Separation from Cement Plant Flue Gases. *Eng. Proc.* **2023**, *56*, 142. [CrossRef]
12. Eshbobaev, J.; Norkobilov, A.; Usmanov, K.; Turakulov, Z.; Kamolov, A.; Rejabov, S.; Farkhadova, S.; Eshbobaev, J.; Norkobilov, A.; Usmanov, K.; et al. An Optimized ANFIS Model for Predicting Water Hardness and TDS in Ion-Exchange Wastewater Treatment Systems. *Eng. Proc.* **2025**, *117*, 18. [CrossRef]
13. Zhou, P.; Li, Z.; Snowling, S.; Baetz, B.W.; Na, D.; Boyd, G. A Random Forest Model for Inflow Prediction at Wastewater Treatment Plants. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 1781–1792. [CrossRef]
14. Wang, K.-J.; Wang, P.-S.; Nguyen, H.-P. A Data-Driven Optimization Model for Coagulant Dosage Decision in Industrial Wastewater Treatment. *Comput. Chem. Eng.* **2021**, *152*, 107383. [CrossRef]
15. Rakhimov, A.; Bozorov, R.; Tuychiev, A.; Mutalov, S.; Eshbobaev, J.; Jabborov, A. Development of an ANFIS-Based Intelligent Control System for Free Chlorine Removal from Industrial Wastewater Using Ion-Exchange Resin. *Eng. Proc.* **2026**, *117*, 28. [CrossRef]
16. Han, Z.; Moh, E.S.X.; Santos, A.L.S.; Barcellos, I.C.; Peng, Y.; Huang, W.; Ye, J. Dechlorination of Wastewater from Shell-Based Glucosamine Processing by Mangrove Wetland-Derived Fungi. *Front. Microbiol.* **2023**, *14*, 1271286. [CrossRef] [PubMed]
17. Khawaga, R.I.; Abdel Jabbar, N.; Al-Asheh, S.; Abouleish, M. Model Identification and Control of Chlorine Residual for Disinfection of Wastewater. *J. Water Process Eng.* **2019**, *32*, 100936. [CrossRef]
18. Wang, X.; Wang, J.; Wang, B.; Wei, W.; Zhang, F.; Huang, M.; Liu, H. Enhancing Effluent Biochemical Oxygen Demand Prediction in Wastewater Treatment Plants: An Ensemble Model of Gradient Boosting and LSTM. *J. Environ. Chem. Eng.* **2025**, *13*, 117897. [CrossRef]
19. Roy, T.; Das, P.; Jagirdar, R.; Shhabat, M.; Abdullah, M.S.; Kashem, A.; Rahman, R. Prediction of Mechanical Properties of Eco-Friendly Concrete Using Machine Learning Algorithms and Partial Dependence Plot Analysis. *Smart Constr. Sustain. Cities* **2025**, *3*, 2. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.