

Article

Explainable Machine Learning for Methylene Blue Removal Under Irradiation: Assessing Nominal Cu-Loading in AC@NiO

Nesrin Bulut ¹, Metin Zontul ² , Seda Karateke ^{3,*} , Orhan Baytar ¹ , Ceren Orak ^{4,5}  and Sabit Horoz ^{6,7} 

¹ Department of Chemical Engineering, Siirt University, Siirt 56100, Türkiye; buluttnesrin@gmail.com (N.B.); orhanbaytar@siirt.edu.tr (O.B.)

² Department of Computer Engineering, Sivas University of Science and Technology, Sivas 58000, Türkiye; metinzontul@sivas.edu.tr

³ Department of Software Engineering, Istanbul Atlas University, Istanbul 34408, Türkiye

⁴ Department of Chemical Engineering, Sivas University of Science and Technology, Sivas 58100, Türkiye; cerenorak@sivas.edu.tr

⁵ Sivas Energy Storage Research and Application Center (SEDAM), Sivas University of Science and Technology, Sivas 58100, Türkiye

⁶ Department of Engineering Fundamental Sciences, Sivas University of Science and Technology, Sivas 58100, Türkiye; sabit.horoz@sivas.edu.tr

⁷ Nanophotonics Application and Research Center-CÜNAM, Sivas Cumhuriyet University, Sivas 58140, Türkiye

* Correspondence: seda.karateke@atlas.edu.tr

Abstract

The time-dependent removal performance under irradiation of activated-carbon-supported nickel oxide (AC@NiO) samples containing nominal Cu-loadings of 0%, 3%, 5%, 7.5%, and 10% was examined using experimental concentration data and an explainable machine-learning (ML) framework. Reaction time and nominal Cu-loading were used as predictors. Preliminary screening with C_t/C_0 as the common target showed that Random Forest (RF) outperformed K-Nearest Neighbors, Multi-Layer Perceptron, and Support Vector Regression. The target-specific regularized RF models were fitted for C_t/C_0 and $\ln(C_0/C_t)$; normalized apparent removal efficiency was derived deterministically as $\eta/100 = 1 - C_t/C_0$. Across 10 controlled random seeds, the mean leave-one-out cross-validation R^2 values were 0.9691 ± 0.0013 for C_t/C_0 and the derived $\eta/100$, and 0.9152 ± 0.0090 for $\ln(C_0/C_t)$. A fixed chronological evaluation, trained at $t \leq 100$ min and evaluated at $100 < t \leq 120$ min, yielded mean R^2 values of 0.9668 ± 0.0073 and 0.8890 ± 0.0163 , respectively. Because the composition-specific RF predictions were constant across this boundary interval, these scores are interpreted as later-time boundary diagnostics rather than evidence of temporal extrapolation. Point-estimate SHAP, impurity-based, and permutation-based analyses assigned greater predictive importance to reaction time, whereas moving-block bootstrap results showed that this ranking was sensitive to temporal resampling. Among the five tested compositions, the nominal 5% Cu-containing sample exhibited the most favorable apparent removal profile. This finding is specific to the investigated conditions and does not establish a universal optimum or a causal physicochemical mechanism.

Keywords: methylene blue; AC@NiO; nominal Cu-loading; explainable machine learning; SHAP; permutation importance; chronological evaluation



Academic Editors: Haidi Xu, Xuehua Yu and Yi Jiao

Received: 18 July 2026

Revised: 28 August 2026

Accepted: 1 September 2026

Published: 7 September 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The increasing discharge of persistent organic contaminants into aquatic environments has intensified the need for effective wastewater treatment technologies [1]. Conventional

biological and physicochemical processes may not completely remove chemically stable pollutants, allowing residual contaminants to remain in treated water and potentially affect aquatic ecosystems and human health [2]. Among advanced oxidation processes (AOPs), heterogeneous photocatalysis has attracted considerable attention because semiconductor materials can generate reactive oxygen species (ROS) capable of promoting the degradation and, under suitable conditions, mineralization of organic pollutants [3,4].

Transition-metal oxides, including nickel oxide (NiO), have been investigated as photocatalytic and adsorptive materials because of their chemical stability, availability, and electronic properties [5,6]. Nevertheless, the photocatalytic performance of unmodified NiO can be limited by its optical response and the recombination of photogenerated electron–hole pairs, which reduce the fraction of charge carriers available for surface reactions [7,8]. Hybridization with activated carbon (AC) can increase accessible surface area and improve pollutant adsorption, while elemental doping has been explored as a means of modifying the structural, optical, and charge-transfer properties of NiO-based materials [9–11]. These considerations motivate the investigation of Cu-containing AC@NiO catalysts as combined adsorption–photocatalysis systems.

Assessing such multicomponent samples requires analysis of the entire time-dependent removal profile under irradiation rather than comparison at a single reaction time. Reaction time and catalyst composition can affect the observed response simultaneously, and their relationships with pollutant removal may be nonlinear or non-monotonic. Moreover, apparent removal efficiency, the residual concentration ratio C_t/C_0 , and the logarithmically transformed concentration response $\ln(C_0/C_t)$ are not independent outputs. Apparent removal efficiency and C_t/C_0 are deterministically related, whereas $\ln(C_0/C_t)$ is a nonlinear transformation of the same residual concentration measurement. Consequently, their predictive performance must be evaluated on the appropriate output scales without interpreting agreement among the derived responses as independent physical validation.

Response surface methodology and related polynomial regression approaches are widely used to analyze catalytic and treatment experiments [12,13]. These methods can be effective when the selected polynomial structure adequately represents the response surface. However, increasingly complex higher-order and interaction terms may be required for nonlinear systems, and model reliability can deteriorate when the number of estimated terms becomes large relative to the available experimental observations [14]. This limitation is particularly relevant for small catalyst datasets containing both temporal and compositional variables.

ML provides an alternative framework for modeling multivariable and nonlinear environmental processes [15,16]. Random Forest (RF) regression combines randomized decision trees and can represent nonlinear associations and interactions without requiring a prespecified polynomial response surface [17,18]. K-Nearest Neighbors (KNN), Multi-Layer Perceptron (MLP), and Support Vector Regression (SVR) offer complementary nonparametric, neural-network, and kernel-based modeling strategies and are therefore suitable comparative baselines [19–22]. Nevertheless, predictive accuracy alone is insufficient to establish that an ML model will generalize to observations collected under a different validation structure. In time-dependent datasets, randomized cross-validation may place measurements from neighboring reaction times in both the training and validation subsets, potentially producing an optimistic assessment of later-time predictive performance.

The interpretability of ML models is also important in catalyst research because accurate predictions do not, by themselves, explain how individual input variables influence model outputs [23,24]. Explainable Artificial Intelligence (XAI) methods, including SHapley Additive exPlanations (SHAP), can characterize the magnitude and direction of feature contributions at both local and global levels [25–27]. For tree-based regression, SHAP can

be complemented by the normalized mean decrease in impurity (MDI), which summarizes the relative impurity reduction assigned to each predictor. These measures describe associations represented by the fitted model; they should not be interpreted as percentages of explained variance, causal effects, or direct evidence of physicochemical mechanisms.

Model interpretation must also be accompanied by validation procedures that reflect the intended prediction task. General interpretability principles and recent ML-assisted photocatalyst studies emphasize the importance of evaluating predictive models transparently and within clearly defined applicability domains [28,29]. For kinetic data, an out-of-time validation design, in which a model is fitted using earlier reaction times and evaluated on a later-time interval, provides a direct assessment of temporal generalization and reduces the risk of mixing earlier and later observations. Such an evaluation does not prove that the model has recovered an underlying kinetic law, but it tests whether the fitted relationship remains predictive within the specified later-time horizon.

Against this background, the present study develops a staged, chronologically evaluated, and explainable ML framework for analyzing the time-dependent removal performance under irradiation of AC@NiO samples containing nominal Cu-loadings of 0%, 3%, 5%, 7.5%, and 10%. Reaction time and nominal Cu-loading are used as predictors. Separate RF models are fitted for the residual concentration ratio $r = C_t/C_0$ and the logarithmically transformed response $\ell = \ln(C_0/C_t)$, whereas normalized apparent removal efficiency is derived as $e = \eta/100 = 1 - r$. The main contributions of the study are as follows:

- RF, KNN, MLP, and SVR are compared using only $r = C_t/C_0$ as the common screening target, avoiding the treatment of deterministically related responses as independent model-selection outcomes.
- The selected RF model is evaluated separately for the two fitted responses, r and ℓ , across 10 controlled random seeds using five-fold CV and LOOCV. Predictions for e are derived from those for r and are not treated as independent predictive evidence.
- Later-time predictive performance is examined through an out-of-time diagnostic in which observations at $t \leq 100$ min are used for training and observations at $t > 100$ min are reserved for evaluation.
- SHAP, normalized MDI, marginal and grouped-conditional permutation importance, PDP, ALE, a discrete response map, and a moving-block bootstrap are used as complementary analyses of predictive associations and their stability.

Together, these stages separate model comparison, robustness assessment, chronological diagnosis, and post hoc interpretation while preserving the algebraic relationships among the response variables. The framework is used to identify the most favorable composition among the investigated nominal Cu-loadings without interpreting it as a globally optimal formulation or evidence of a causal physicochemical mechanism.

The remainder of the manuscript is organized as follows. Section 2 reviews previous studies on Cu-modified NiO-based materials, carbon-supported catalyst systems, nonlinear ML modeling, chronological evaluation, and explainable AI in photocatalytic applications. Section 3 describes the catalyst-preparation procedure, removal experiments under irradiation, response-variable definitions, candidate regression algorithms, cross-validation strategies, chronological diagnostic design, performance metrics, and the explainability and stability analyses. The complete computational sequence and its conceptual representation are also provided in Algorithm 1 and Figure 1, respectively. Section 4 presents the time-dependent removal profiles, candidate-model comparison, target-specific RF robustness results, out-of-time diagnostic performance, complementary feature-attribution analyses, response-pattern diagnostics, stability results, and linear and monotonic correlation patterns. Finally, Section 5 summarizes the main findings and experimental and characterization requirements for broader generalization.

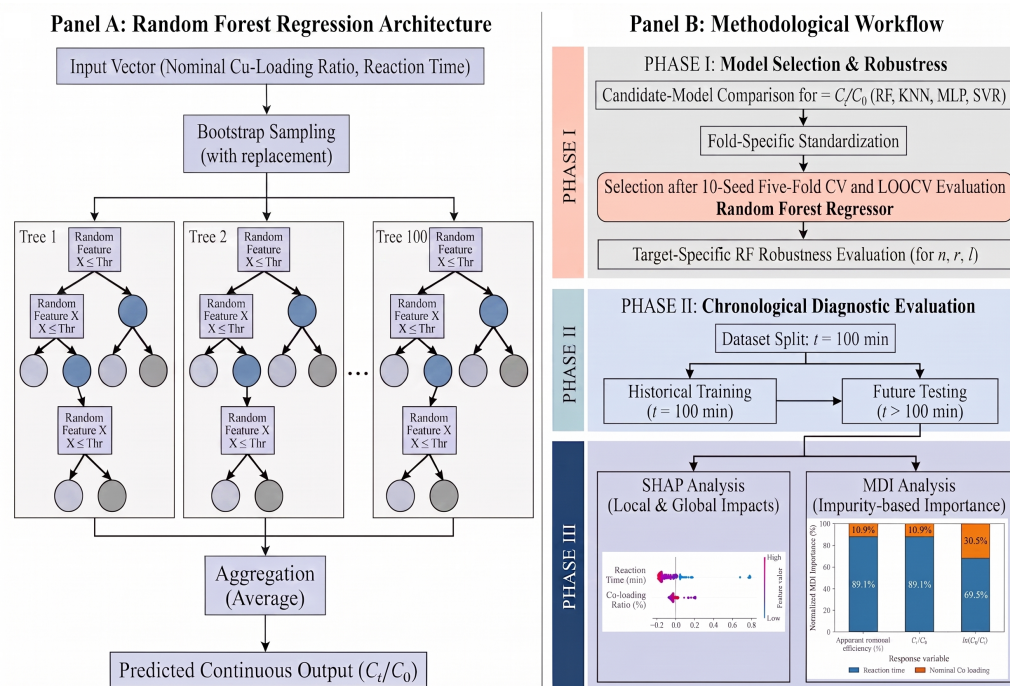


Figure 1. Schematic representation of the RF architecture and the three-phase modeling workflow.

Algorithm 1 Chronologically Validated ML and RF-XAI Framework for Removal-under-Irradiation Modeling

Input: Experimental dataset $\mathcal{D} = \{(\mathbf{x}_i, \eta_i, r_i, \ell_i)\}_{i=1}^N$, where $\mathbf{x}_i = [d_i, t_i]^T$, $r_i = C_{t,i}/C_0$, $\eta_i = 100(1 - r_i)$, and $\ell_i = \ln(C_0/C_{t,i})$; normalized efficiency $e_i = \eta_i/100$; primary screening target r ; target-specific CV set $\mathcal{Y}_{CV} = \{r, \ell\}$; fitted target set $\mathcal{Y} = \{r, \ell\}$; candidate-model set $\mathcal{A} = \{\text{RF}, \text{KNN}, \text{MLP}, \text{SVR}\}$; seed set $\mathcal{S} = \{1, \dots, 10\}$; chronological cutoff $t_c = 100$ min

Output: Candidate-model comparison statistics, target-specific RF robustness summaries, chronological diagnostic predictions, SHAP contributions, and normalized MDI scores

Phase I-A: Candidate-Model Comparison for $r = C_t/C_0$

- 1: **for all** $a \in \mathcal{A}$ **do**
- 2: Initialize $\mathcal{M}_{a,r}^{5F} \leftarrow \emptyset$ and $\mathcal{M}_{a,r}^{LOO} \leftarrow \emptyset$
- 3: **end for**
- 4: **for all** $s \in \mathcal{S}$ **do**
- 5: Generate and retain a shuffled five-fold partition Π_s using `random_state = s`
- 6: **for all** $a \in \mathcal{A}$ **do**
- 7: Initialize model a using seed s when the estimator contains stochastic components
- 8: For KNN, MLP, and SVR, apply fold-specific standardization within the cross-validation pipeline
- 9: Generate pooled five-fold predictions

$$\hat{\mathbf{r}}_{5F}^{(a,s)} \leftarrow \text{CROSSVALPREDICT}(a, \mathbf{X}, \mathbf{r}, \Pi_s)$$

- 10: Compute

$$\mathbf{m}_{a,s,r}^{5F} \leftarrow \text{METRICS}(\mathbf{r}, \hat{\mathbf{r}}_{5F}^{(a,s)})$$

- 11: Append $\mathbf{m}_{a,s,r}^{5F}$ to $\mathcal{M}_{a,r}^{5F}$

Algorithm 1 *Cont.*

12: Generate pooled LOOCV predictions

$$\hat{\mathbf{r}}_{\text{LOO}}^{(a,s)} \leftarrow \text{LOOCVPREDICT}(a, \mathbf{X}, \mathbf{r})$$

13: Compute

$$\mathbf{m}_{a,s,r}^{\text{LOO}} \leftarrow \text{METRICS}(\mathbf{r}, \hat{\mathbf{r}}_{\text{LOO}}^{(a,s)})$$

14: Append $\mathbf{m}_{a,s,r}^{\text{LOO}}$ to $\mathcal{M}_{a,r}^{\text{LOO}}$

15: **end for**

16: **end for**

17: **for all** $a \in \mathcal{A}$ **do**

18: Summarize R^2 , MAE, and MSE as the mean and standard deviation across the controlled runs

19: **end for**

20: Rank the candidate models using their cross-validated performance for $r = C_t/C_0$

21: Retain RF for the target-specific, chronological, and explainability analyses

Phase I-B: Target-Specific RF Robustness

22: **for all** $y \in \mathcal{Y}_{\text{CV}}$ **do**

23: Initialize $\mathcal{M}_{\text{RF},y}^{5\text{F}} \leftarrow \emptyset$ and $\mathcal{M}_{\text{RF},y}^{\text{LOO}} \leftarrow \emptyset$

24: **for all** $s \in \mathcal{S}$ **do**

25: Initialize the final RF configuration using seed s

26: Generate pooled five-fold predictions using Π_s

$$\hat{\mathbf{y}}_{5\text{F}}^{(s)} \leftarrow \text{CROSSVALPREDICT}(\text{RF}, \mathbf{X}, \mathbf{y}, \Pi_s)$$

27: Compute

$$\mathbf{m}_{s,y}^{5\text{F}} \leftarrow \text{METRICS}(\mathbf{y}, \hat{\mathbf{y}}_{5\text{F}}^{(s)})$$

28: Append $\mathbf{m}_{s,y}^{5\text{F}}$ to $\mathcal{M}_{\text{RF},y}^{5\text{F}}$

29: Generate pooled LOOCV predictions

$$\hat{\mathbf{y}}_{\text{LOO}}^{(s)} \leftarrow \text{LOOCVPREDICT}(\text{RF}, \mathbf{X}, \mathbf{y})$$

30: Compute

$$\mathbf{m}_{s,y}^{\text{LOO}} \leftarrow \text{METRICS}(\mathbf{y}, \hat{\mathbf{y}}_{\text{LOO}}^{(s)})$$

31: Append $\mathbf{m}_{s,y}^{\text{LOO}}$ to $\mathcal{M}_{\text{RF},y}^{\text{LOO}}$

32: **end for**

33: Summarize the target-specific R^2 , MAE, and MSE values across the 10 controlled runs

34: **end for**

35: For each seed and validation protocol, derive $\hat{\mathbf{e}} = \mathbf{1} - \hat{\mathbf{r}}$ and compute the corresponding metrics against $\mathbf{e} = \mathbf{1} - \mathbf{r}$

Phase II: Chronological Out-of-Time Diagnostic Evaluation

36: Sort $\mathcal{D}$ in ascending order of reaction time t_i

37: Define the historical training subset

$$\mathcal{D}_{\text{train}} \leftarrow \{(\mathbf{x}_i, \eta_i, r_i, \ell_i) \in \mathcal{D} : t_i \leq t_c\}$$

38: Define the later-time test subset

$$\mathcal{D}_{\text{test}} \leftarrow \{(\mathbf{x}_i, \eta_i, r_i, \ell_i) \in \mathcal{D} : t_i > t_c\}$$

39: **for all** $y \in \mathcal{Y}$ **do**

Algorithm 1 *Cont.*

```

40:   Fit the final RF model  $\mathcal{F}_y^{\text{OOT}}$  using only  $\mathcal{D}_{\text{train}}$ 
41:   Generate later-time predictions
      
$$\hat{\mathbf{y}}_{\text{OOT}} \leftarrow \text{PREDICT}(\mathcal{F}_y^{\text{OOT}}, \mathbf{X}_{\text{test}})$$

42:   Compute
      
$$\mathbf{m}_y^{\text{OOT}} \leftarrow \text{METRICS}(\mathbf{y}_{\text{test}}, \hat{\mathbf{y}}_{\text{OOT}})$$

43: end for
44: Derive later-time percentage-scale efficiency predictions as  $\hat{\eta}_{\text{OOT}} = 100(1 - \hat{\mathbf{r}}_{\text{OOT}})$  and
   compute the corresponding metrics
Phase III: RF Explainability Assessment
45: for all  $y \in \mathcal{Y}$  do
46:   Refit the final RF specification  $\mathcal{F}_y^*$  using the complete dataset for post hoc interpretation
47:   Initialize the TreeExplainer
      
$$\mathcal{E}_y \leftarrow \text{TREEEXPLAINER}(\mathcal{F}_y^*)$$

48:   Compute the SHAP contribution matrix
      
$$\Phi^{(y)} \leftarrow \text{SHAPVALUES}(\mathcal{E}_y, \mathbf{X})$$

49:   for all  $j \in \{d, t\}$  do
50:     Calculate mean absolute SHAP importance
      
$$I_{j,y}^{\text{SHAP}} \leftarrow \frac{1}{N} \sum_{i=1}^N |\phi_j^{(y)}(\mathbf{x}_i)|$$

51:   end for
52:   Extract normalized MDI importance
      
$$\mathbf{I}_y^{\text{MDI}} \leftarrow \text{FEATUREIMPORTANCE}(\mathcal{F}_y^*)$$

53: end for
54: Obtain efficiency-scale SHAP contributions from the ratio model as  $\Phi^{(\eta)} = -100\Phi^{(r)}$ ;
   retain the same normalized MDI proportions as for  $r$ 
55: Return Candidate-model comparison summaries, target-specific RF robustness summaries,
   chronological diagnostic metrics, SHAP contributions, and normalized MDI importance scores

```

2. Related Studies

Cu modification of NiO and the incorporation of carbonaceous supports have both been investigated as strategies for improving pollutant removal. Hameeda et al. [30] synthesized Cu-doped NiO nanoparticles by co-precipitation and reported changes in optical properties, morphology, and visible-light photocatalytic dye degradation relative to undoped NiO. The observed improvement was attributed to Cu-induced modifications in the electronic and structural properties of the material. In a complementary approach, Hashem et al. [31] prepared NiO–CuO/activated-carbon nanocomposites for the removal of Pb(II) and Cd(II) ions from water. Their results demonstrated the feasibility of combining metal oxides with activated carbon to integrate surface adsorption with the functionality of the supported oxide phases. These studies provide materials-based motivation for examining the time-dependent removal performance of Cu-containing AC@NiO catalysts.

A related body of research has compared conventional polynomial response-surface models with nonlinear ML methods for environmental treatment processes. Kang et al. [32]

evaluated polynomial regression and several ML algorithms for predicting decolorization efficiency in a UV/H₂O₂ textile-wastewater treatment system. Their results showed that the evaluated ML models could provide higher predictive accuracy than the selected polynomial model, while SHAP analysis helped describe the model-based contributions of the process variables. This comparison illustrates the potential value of ML when a fixed polynomial response surface is insufficient to represent nonlinear relationships. It does not, however, imply that ML automatically provides mechanistic or causal explanations.

Machine-learning methods have also been applied directly to photocatalytic degradation datasets. Anandhi and Iyapparaja [33] compared regression algorithms for predicting the photocatalytic degradation of pharmaceutical compounds and dyes using compiled experimental data. Their findings indicated that ensemble models can represent heterogeneous relationships among catalyst, pollutant, and operating variables more effectively than simple linear baselines. Similarly, Schossler et al. [34] developed an interpretable ML framework for predicting the TiO₂-mediated photocatalytic degradation of air contaminants. Their workflow combined data preprocessing and augmentation with SHAP-based feature attribution to examine how the fitted model associated experimental variables with the predicted response. Together, these studies demonstrate the growing use of nonlinear predictive models and post-hoc interpretation methods in photocatalysis research.

A closely related study by Baytar et al. [35] investigated the time-dependent photocatalytic performance of Mn-doped Zn₂SnO₄. In that study, conventional statistical analyses were supplemented with an optimized Multi-Layer Perceptron model and SHAP analysis. A leave-one-dose-out cross-validation strategy was used to evaluate predictive performance when an entire dopant level was excluded from model fitting. That work demonstrated the usefulness of combining cross-validation and feature-attribution methods when analyzing temporal and compositional variables in a doped photocatalyst system. The present study builds on this general modeling framework but applies RF regression and adds a chronological out-of-time evaluation to assess later-time predictive performance.

Validation design is particularly important when observations are correlated or sequentially ordered. Kaufman et al. [36] systematically described data-leakage mechanisms through which information unavailable at prediction time can enter model training or evaluation and produce overly optimistic performance estimates. For time-dependent degradation data, randomized partitioning can place observations from neighboring reaction times in both the training and validation subsets. A chronological split reduces this type of temporal mixing, although performance at a single later-time interval does not establish that a model has recovered the underlying kinetic mechanism or will extrapolate beyond the investigated experimental domain.

The RF models used in the present study are based on the ensemble-tree framework described by Breiman, whereas the SHAP analysis follows the game-theoretic feature-attribution framework introduced by Lundberg and Lee [37]. In the present application, SHAP is complemented by normalized mean decrease in impurity (MDI), marginal and grouped-conditional permutation importance, response-pattern diagnostics, and a moving-block bootstrap stability analysis. These quantities are interpreted as properties of the fitted models rather than as causal effects, percentages of explained variance, or direct evidence of physicochemical mechanisms.

Taken together, the existing literature provides both a materials basis for investigating Cu-containing activated-carbon-supported NiO and a methodological basis for applying nonlinear and interpretable ML models to photocatalytic data. The present study extends these lines of research by combining target-specific RF evaluation across multiple random seeds, a chronological later-time diagnostic, and complementary explainability and stability analyses. This framework is used to distinguish model-based temporal and compositional

associations and to identify the most favorable sample composition only among the nominal Cu-loadings and experimental conditions examined in this study.

3. Materials and Methods

All chemicals were of analytical grade and were used as received without further purification. Nickel(II) nitrate hexahydrate ($\text{Ni}(\text{NO}_3)_2 \cdot 6 \text{H}_2\text{O}$, 98%, Sigma–Aldrich, St. Louis, MO, USA.) and copper(II) nitrate trihydrate ($\text{Cu}(\text{NO}_3)_2 \cdot 3 \text{H}_2\text{O}$, 99%, Sigma–Aldrich, St. Louis, MO, USA.) were used as the nickel and copper precursors, respectively. These precursors were used to prepare the AC@NiO catalysts with nominal Cu-loading levels of 0%, 3%, 5%, 7.5%, and 10%.

Activated carbon (AC) was produced from locally collected okra (*Abelmoschus esculentus*) stalks. The stalks were thoroughly washed with deionized water to remove adhering surface impurities and then dried at 80 °C for 24 h. After drying, the stalks were cut into small pieces and impregnated with potassium hydroxide (KOH) at a biomass-to-KOH mass ratio of 1:1. The impregnated material was carbonized in a tube furnace at 500 °C for 2 h under a nitrogen atmosphere using a heating rate of 5 °C min^{−1}.

Following carbonization, the resulting AC was washed successively with dilute hydrochloric acid (HCl) and deionized water until the washing solution reached approximately neutral pH. The washed material was subsequently dried at 80 °C for 12 h, ground, and sieved to obtain particles smaller than 150 µm (100 mesh). The resulting material was used as the carbonaceous support for the preparation of the Cu-containing AC@NiO catalysts.

Sodium hydroxide (NaOH, 99%, Merck, Darmstadt, Germany) was used as the precipitating agent during catalyst synthesis and was prepared as a 1 mol L^{−1} aqueous solution. Deionized water with a resistivity of 18.2 MΩ·cm and absolute ethanol (99.5%, Merck) were used for solution preparation and washing. Methylene blue (MB; dye content 95%, Sigma–Aldrich) was selected as the model organic pollutant for evaluating the removal performance of the synthesized catalysts.

3.1. Catalyst Synthesis

The Cu-containing AC@NiO catalysts were prepared by a co-precipitation method followed by thermal treatment. For each synthesis, 0.50 g of nickel(II) nitrate hexahydrate ($\text{Ni}(\text{NO}_3)_2 \cdot 6 \text{H}_2\text{O}$) was dissolved in 50 mL of deionized water and magnetically stirred at room temperature for 30 min. For the Cu-containing samples, copper(II) nitrate trihydrate ($\text{Cu}(\text{NO}_3)_2 \cdot 3 \text{H}_2\text{O}$) was added in an amount calculated to obtain nominal Cu fractions of 3%, 5%, 7.5%, and 10%. The nominal Cu fraction was defined on a molar basis as

$$x_{\text{Cu}}(\%) = 100 \left(\frac{n_{\text{Cu}}}{n_{\text{Cu}} + n_{\text{Ni}}} \right), \quad (1)$$

where n_{Cu} and n_{Ni} denote the molar amounts of Cu and Ni introduced through their respective nitrate precursors. A nominal 0% Cu precursor formulation was prepared under the same conditions to produce the 0% reference catalyst.

Subsequently, 0.25 g of the okra-stalk-derived AC was added to each precursor solution. The suspensions were ultrasonicated for 30 min to facilitate the dispersion of the AC and contact between the carbon surface and dissolved metal precursors. The AC mass was kept constant at 0.25 g for all five nominal Cu compositions to maintain a consistent carbon-support loading throughout the catalyst series.

The nominal Cu fractions of 0%, 3%, 5%, 7.5%, and 10% were selected to evaluate the removal response over a graded compositional range extending from nominal 0% Cu reference sample to the highest Cu-containing formulation investigated. These values represent nominal precursor compositions and do not, by themselves, demonstrate substitutional Cu

incorporation into the NiO lattice. Confirmation of lattice substitution, secondary-phase formation, or Cu-induced structural and electronic changes would require appropriate physicochemical characterization.

A 1 mol L^{−1} NaOH solution was added dropwise to each suspension under continuous stirring until the pH reached 10. The resulting suspensions were then maintained at 60 °C under continuous stirring for 2 h to promote precipitation and aging of the metal-containing phases on the AC support. The solids were recovered by centrifugation at 4000 rpm for 10 min and washed repeatedly with deionized water and absolute ethanol to remove residual soluble species. Washing was continued until the supernatant reached approximately neutral pH. The recovered solids were subsequently dried at 80 °C for 12 h.

The dried precursors were thermally treated in a muffle furnace at 450 °C for 3 h in air using a heating rate of 5 °C min^{−1}. After thermal treatment, the samples were allowed to cool to room temperature inside the furnace, ground using an agate mortar, and stored in a desiccator until the removal experiments were performed. The same preparation procedure was applied separately to the nominal 0%, 3%, 5%, 7.5%, and 10% Cu compositions.

3.2. Physicochemical Characterization

The optical properties of the nominal 0% and 5% Cu-containing AC@NiO samples were examined by UV–Vis diffuse-reflectance spectroscopy (UV–Vis diffuse-reflectance spectroscopy (DRS) was performed using a UV-2600 spectrophotometer equipped with an integrating sphere (Shimadzu Corporation, Kyoto, Japan)) over the wavelength range of 200–800 nm, using BaSO₄ as the reference standard. The measured diffuse reflectance, R , was transformed using the Kubelka–Munk function, $F(R) = \frac{(1-R)^2}{2R}$.

Apparent optical band-gap energies were estimated from plots of $[F(R)hv]^2$ versus hv , corresponding to a direct-allowed Tauc representation. The linear regions near the absorption edges were extrapolated to $[F(R)hv]^2 = 0$.

3.3. Removal Experiments Under Irradiation

The time-dependent removal performance of the prepared catalysts was evaluated in a batch photoreactor using methylene blue (MB) as the model pollutant. For each experiment, 0.10 g of catalyst was dispersed in 100 mL of an aqueous MB solution with an initial concentration of 20 mg L^{−1}, corresponding to a catalyst loading of 1.0 g L^{−1}. The solution pH and temperature were maintained at 7 and 25 °C, respectively.

Before irradiation, each catalyst suspension was magnetically stirred in the dark for 30 min as a pre-irradiation conditioning step. The concentration measured at the end of this period and immediately before illumination was defined operationally as C_0 . This definition does not demonstrate that adsorption equilibrium was attained or isolate the photocatalytic contribution to the subsequent concentration decrease. The irradiation source was then switched on, and this point was designated as $t = 0$ for the irradiation experiment.

During irradiation, 3 mL aliquots were collected at $t = 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110$, and 120 min. Each aliquot was immediately passed through a 0.45 µm membrane syringe filter to separate suspended catalyst particles from the liquid phase. The absorbance of the filtered solution was measured at the characteristic MB absorption wavelength of 664 nm using a UV–Vis spectrophotometer. The residual MB concentration at each sampling time, C_t , was determined from the corresponding absorbance using an external calibration curve.

The residual concentration ratio was calculated as

$$r_t = \frac{C_t}{C_0}, \quad (2)$$

and the apparent removal efficiency during irradiation was calculated as

$$\eta_t(\%) = 100 \left(1 - \frac{C_t}{C_0} \right). \quad (3)$$

The logarithmically transformed concentration response was calculated as

$$y_{\ln,t} = \ln \left(\frac{C_0}{C_t} \right). \quad (4)$$

The quantities η_t , C_t/C_0 , and $\ln(C_0/C_t)$ were calculated from the same post-dark-normalized concentration measurements and therefore represent mathematically related forms of the same response rather than independent experimental outcomes. These quantities describe the apparent concentration decrease during irradiation relative to the operational post-dark baseline, C_0 , and do not separately quantify adsorption, photolysis, or photocatalytic transformation. The resulting time-dependent observations were subsequently used as the target variables in the ML analyses.

As a descriptive kinetic benchmark, the linearized apparent pseudo-first-order representation was fitted separately for each nominal Cu-loading using ordinary least squares over the complete 0–120 min interval: $\ln(C_0/C_t) = b + k_{\text{app}}t$.

The intercept b was estimated rather than constrained to zero to permit assessment of departures from the idealized relationship. The fitted slope, intercept, and coefficient of determination were reported in Table S1. These regressions were fitted to the complete composition-specific profiles and therefore provide in-sample kinetic descriptors rather than cross-validated prediction metrics.

3.4. Machine Learning and Explainable AI Methodology

3.4.1. Data Configuration and Algorithm Selection

The experimental dataset contained 65 observations corresponding to five nominal Cu-loading levels (0%, 3%, 5%, 7.5%, and 10%) and 13 reaction times (0–120 min). Nominal Cu-loading and reaction time were used as the two predictor variables. For the i th observation, the input vector was defined as

$$\mathbf{x}_i = \begin{bmatrix} d_i \\ t_i \end{bmatrix}, \quad (5)$$

where d_i is the nominal Cu-loading (%) and t_i is the reaction time (min).

Three separate target representations were considered: degradation efficiency, η_i (%); residual concentration ratio, $r_i = C_{t,i}/C_0$; and the logarithmically transformed concentration response, $\ell_i = \ln(C_0/C_{t,i})$. These quantities were calculated from the same measured MB concentrations according to Equations (2)–(4). Therefore, they do not constitute statistically independent experimental outcomes. Apparent removal efficiency is an affine transformation of C_t/C_0 , whereas $\ln(C_0/C_t)$ is a nonlinear monotonic transformation of the same concentration ratio.

A separate regression model was fitted for each target representation. For the target-specific multi-seed analysis, apparent removal efficiency was represented in normalized fractional form as $\eta/100$. Consequently, the corresponding MAE and MSE values were dimensionless and directly comparable with those obtained for C_t/C_0 . For graphical presentation, out-of-time evaluation, and SHAP interpretation, the normalized predictions and attribution values were transformed back to the percentage scale where explicitly indicated. Identical performance for $\eta/100$ and C_t/C_0 was treated as an expected consequence of their deterministic affine relationship rather than as independent evidence of model validity.

No linear functional form was imposed on the relationships among reaction time, nominal Cu-loading, and the response variables. Four nonlinear regression algorithms were therefore evaluated using the same predictor variables, target definitions, and cross-validation partitions.

Random Forest (RF) regression was included as a tree-based ensemble method. RF constructs regression trees using bootstrap samples and randomized feature selection and averages the individual tree predictions. This structure can represent nonlinear associations and interactions while reducing prediction variance relative to a single regression tree. K-Nearest Neighbors (KNN) regression was included as a local nonparametric method that estimates an output from nearby observations in predictor space [38]. A Multi-Layer Perceptron (MLP) was evaluated as a neural-network-based nonlinear function approximator [39]. Support Vector Regression (SVR) with a radial basis function kernel was included as a regularized kernel-based regression method [40,41].

All models were implemented using the Python environment and package versions reported in the publicly available source-code repository, with the regression algorithms implemented through the `scikit-learn` library [42]. The preliminary candidate-model comparison was conducted using fixed baseline configurations. KNN regression used $k = 5$ with distance-based weighting and the Euclidean distance metric. The MLP regressor contained hidden layers of 64 and 32 units, used the ReLU activation function and Adam optimizer, and was trained with `learning_rate_init = 0.001` for a maximum of 500 iterations. SVR employed a radial basis function kernel with $C = 1.0$, $\epsilon = 0.1$, and $\gamma = \text{scale}$. The baseline RF model used 100 regression trees and bootstrap sampling.

For KNN, MLP, and SVR, predictor standardization was implemented within a `scikit-learn` Pipeline. The scaler was fitted exclusively to each training fold and subsequently applied to the corresponding validation fold. RF was fitted using the original predictor scales because tree-based partitioning does not require feature standardization. No MinMax or other scaling transformation was applied to the target variables.

The preliminary comparison reported in Table 1 used the naturally dimensionless residual concentration ratio $r = C_t/C_0$ as the common screening target and was intended only to compare the baseline candidate algorithms. Following this screening stage, RF was retained and evaluated using the final target-specific analysis protocol. To mitigate the risk of memorizing the limited sample space ($N = 65$) and to improve out-of-sample generalization, the final RF implementation was strictly regularized. It consisted of 100 regression trees and used bootstrap sampling, the squared-error splitting criterion, a constrained maximum tree depth (`max_depth = 5`), `min_samples_split = 2`, `min_samples_leaf = 1`, and `max_features = 1.0`. Its random state was set to the corresponding controlled seed in each run.

An additional model-complexity sensitivity analysis was performed for the primary C_t/C_0 response. The reference RF (`max_depth = None`, `min_samples_leaf = 1`) was compared with maximum depths of 3 and 5, minimum leaf sizes of 2 and 3, and the combined setting of maximum depth 5 and minimum leaf size 2. All configurations retained 100 trees, with the remaining RF settings held fixed. Seeds 0–9 were used with matched validation partitions across configurations within each seed. Five-fold CV and LOOCV metrics were calculated from the pooled out-of-fold predictions for each run. The chronological diagnostic used 55 observations at $t \leq 100$ min for training and 10 observations at 110 and 120 min for evaluation. Results were summarized as the mean and sample standard deviation across seeds. For five-fold CV, this variability includes both partitioning and RF fitting; for LOOCV and the fixed chronological split, it reflects stochastic RF fitting. It does not represent experimental uncertainty. In this sensitivity analysis, normalized apparent

removal efficiency was derived as $1 - C_t/C_0$ and was not evaluated as an independent target. No automatic model selection was performed in this additional analysis.

Table 1. Comparative predictive performance of the candidate machine learning models for C_t/C_0 under five-fold CV and LOOCV. Results are reported as the mean $\pm$ standard deviation across 10 controlled random-seed evaluations.

CV Method	Model	R^2 (Mean $\pm$ Std)	MAE (Mean $\pm$ Std)	MSE (Mean $\pm$ Std)
Five-Fold CV	KNN ($k = 5$)	0.5939 ± 0.0522	0.1132 ± 0.0083	0.0296 ± 0.0038
Five-Fold CV	MLP (64, 32)	0.4568 ± 0.0670	0.1470 ± 0.0089	0.0398 ± 0.0048
Five-Fold CV	Random Forest	0.9354 ± 0.0197	0.0505 ± 0.0038	0.0051 ± 0.0016
Five-Fold CV	SVR (RBF kernel)	0.7663 ± 0.0301	0.0988 ± 0.0055	0.0167 ± 0.0022
LOOCV	KNN ($k = 5$)	0.6910 ± 0.0000	0.0906 ± 0.0000	0.0221 ± 0.0000
LOOCV	MLP (64, 32)	0.4716 ± 0.0477	0.1463 ± 0.0059	0.0387 ± 0.0035
LOOCV	Random Forest	0.9670 ± 0.0025	0.0382 ± 0.0016	0.0026 ± 0.0002
LOOCV	SVR (RBF kernel)	0.7924 ± 0.0000	0.0919 ± 0.0000	0.0147 ± 0.0000

The target-specific RF models address the residual concentration ratio $r = C_t/C_0$ and the logarithmically transformed concentration response $\ell = \ln(C_0/C_t)$. For the primary-response results in Table 2, normalized apparent removal efficiency and its predictions were derived as $e = \eta/100 = 1 - r$ and $\hat{e} = 1 - \hat{r}$, respectively, rather than fitted as an independent response. Percentage-scale efficiency is obtained by multiplying the normalized quantity by 100. Tables 1 and 2 correspond to the preliminary model-screening stage and the subsequent target-specific evaluation, respectively; their RF results are not replicate estimates from an identical procedure.

Table 2. Target-specific performance metrics of the regularized Random Forest (RF) model under five-fold CV and leave-one-out cross-validation (LOOCV). Results are reported as the mean $\pm$ standard deviation across 10 controlled random-seed evaluations.

CV Method	Output Variable	R^2 (Mean $\pm$ SD)	MAE (Mean $\pm$ SD)	MSE (Mean $\pm$ SD)
Five-fold CV	C_t/C_0 and $\eta/100$ ^a	0.9546 ± 0.0137	0.0420 ± 0.0045	0.0030 ± 0.0009
Five-Fold CV	$\ln(C_0/C_t)$	0.8542 ± 0.0248	0.3860 ± 0.0265	0.2662 ± 0.0453
LOOCV	C_t/C_0 and $\eta/100$ ^a	0.9691 ± 0.0013	0.0341 ± 0.0008	0.0020 ± 0.0001
LOOCV	$\ln(C_0/C_t)$	0.9152 ± 0.0090	0.3107 ± 0.0132	0.1549 ± 0.0164

^a The residual concentration ratio (C_t/C_0) and normalized apparent removal efficiency ($\eta/100$) share identical mathematical performance metrics (R^2 , MAE, MSE) due to their deterministic affine relationship ($\hat{e} = 1 - \hat{r}$).

3.4.2. Cross-Validation and Multi-Seed Robustness

Within-domain predictive performance was evaluated using shuffled five-fold cross-validation and leave-one-out cross-validation (LOOCV) [43]. The two fitted responses were $r = C_t/C_0$ and $\ell = \ln(C_0/C_t)$. Normalized apparent removal efficiency was evaluated through the deterministic relations $e = 1 - r$ and $\hat{e} = 1 - \hat{r}$, rather than through a third independently fitted RF model.

For five-fold cross-validation, the 65 observations were randomly shuffled and partitioned into five subsets. In each iteration, four subsets were used for model fitting and the remaining subset was used for validation. This procedure was repeated until every observation had received one out-of-fold prediction. The complete set of 65 out-of-fold predictions was then used to calculate R^2 , MAE, and MSE for that run.

To assess sensitivity to data partitioning and stochastic model initialization, the five-fold procedure was repeated using 10 prespecified random seeds [44]. For each seed, the same shuffled folds were used for RF, KNN, MLP, and SVR, thereby enabling a paired

comparison under identical training and validation subsets. The seed controlled the fold construction and, for RF and MLP, the stochastic components of model fitting. Thus, RF and MLP could vary because of both partitioning and initialization, whereas KNN and SVR could vary only across the shuffled partitions. For KNN, MLP, and SVR, predictor standardization was performed within a scikit-learn Pipeline: the StandardScaler was fitted exclusively on each training fold and then applied to the corresponding validation fold. This fold-specific preprocessing prevented information leakage from the validation data. RF was fitted to the original predictor scales because tree-based models do not require feature standardization.

LOOCV was used as a complementary small-sample validation procedure. For each of the $N = 65$ observations, one observation was used for validation and the remaining $N - 1 = 64$ observations were used for model fitting. The process was repeated until each observation had been held out once, producing one out-of-sample prediction for every observation. Because the LOOCV partitions are fixed, changing the random seed affects only models containing stochastic fitting components. Accordingly, RF and MLP may exhibit seed-dependent variation under LOOCV, whereas deterministic KNN and SVR implementations are expected to return identical LOOCV predictions when their hyperparameters and preprocessing procedures remain unchanged.

For each model, target, and validation strategy, performance metrics were first calculated from the complete set of out-of-fold predictions within each seed. The reported mean and standard deviation were then calculated across the ten seed-level metric values. Thus, the reported standard deviations characterize sensitivity to the controlled partitioning and/or model-initialization conditions; they do not represent experimental uncertainty or confidence intervals for the true generalization error.

Because LOOCV training sets overlap extensively and the corresponding prediction errors are not statistically independent, LOOCV variability should not be interpreted as an uncertainty-free estimate of generalization performance [45,46]. Similarly, shuffled five-fold cross-validation can distribute observations from adjacent reaction times between the training and validation subsets. The ability of the selected model to predict observations collected after a defined temporal cutoff was therefore assessed separately using the out-of-time validation procedure described in Section 3.4.3.

3.4.3. Out-of-Time (Chronological) Validation Strategy

A fixed chronological split was used to assess later-time predictions for compositions already represented in the training data. The training subset comprised 55 observations at $t \leq 100$ min, and the evaluation subset comprised 10 observations at 110 and 120 min. The later-time observations were excluded from fitting each chronological RF model. The original single-run assessment was supplemented by the multi-seed model-complexity comparison described in Section 3.4.1, using the same temporal cutoff.

The RF model family had been selected using the complete dataset during preliminary screening. In addition, the same chronological subset was examined across the sensitivity configurations. These evaluations are therefore post-selection, later-time diagnostics, not an untouched, independent final test of the complete model-selection procedure. The reference RF configuration was retained; the sensitivity analysis did not select a replacement model for the downstream analyses.

The investigated numerical ranges were $0 \leq t \leq 120$ min and nominal Cu-loading between 0% and 10%, under the experimental conditions used in this study. Only the five Cu-loading levels of 0%, 3%, 5%, 7.5%, and 10% were measured. Predictions at intermediate loadings are model-based interpolation and have not been independently validated. Predictions outside the investigated ranges or under different operating conditions are

unsupported. Being inside these ranges does not itself establish predictive reliability for new catalyst batches.

The holdout times of 110 and 120 min are outside the training-time range ending at 100 min. However, an RF does not extend a continuous temporal trend beyond this boundary. Its terminal-leaf behavior was checked by comparing predictions at 100, 110, and 120 min for each composition, configuration, and seed. As a simple chronological reference, each composition's measured C_t/C_0 at 100 min was also carried forward to both later times. These predictions were formed without fitting a model or using the later-time outcomes and were then scored against the measured later-time values.

3.4.4. XAI Implementation

Two complementary feature-attribution methods were applied to the selected RF models: normalized mean decrease in impurity (MDI) and SHAP. MDI summarizes the accumulated reduction in the RF regression criterion attributed to each predictor across the ensemble, whereas SHAP decomposes individual model predictions into additive feature contributions [47]. Both methods were applied separately to the RF model fitted for each target representation.

For a split node v , the weighted decrease in impurity was defined as

$$\Delta\mathcal{I}(v) = p(v)\mathcal{I}(v) - p(v_L)\mathcal{I}(v_L) - p(v_R)\mathcal{I}(v_R), \quad (6)$$

where $\mathcal{I}(v)$ is the regression impurity at node v , $p(v)$ is the proportion of observations reaching that node, and v_L and v_R are its left and right child nodes, respectively. The normalized MDI assigned to predictor j was calculated as

$$I_j^{\text{MDI}} = \frac{\sum_{m=1}^M \sum_{\substack{v \in \mathcal{T}_m \\ s(v)=j}} \Delta\mathcal{I}(v)}{\sum_{k=1}^p \sum_{m=1}^M \sum_{\substack{v \in \mathcal{T}_m \\ s(v)=k}} \Delta\mathcal{I}(v)}, \quad (7)$$

where M is the number of trees, $\mathcal{T}_m$ is the set of split nodes in tree m , $s(v)$ identifies the predictor used at node v , and p is the total number of predictors. The MDI values satisfy

$$\sum_{j=1}^p I_j^{\text{MDI}} = 1. \quad (8)$$

For presentation as percentages, each normalized value was multiplied by 100. These percentages represent relative shares of the accumulated impurity reduction assigned by the fitted RF model. They are not percentages of explained variance and should not be interpreted as physical, causal, or mechanistic contributions [47].

MDI can favor predictors that provide a larger number of potential split points. In the present dataset, reaction time had 13 distinct values, whereas nominal Cu-loading had five. Consequently, the MDI comparison may partly reflect differences in predictor cardinality. The MDI results were therefore interpreted together with SHAP rather than as a stand-alone ranking of physicochemical importance.

To complement MDI with an importance measure that is not calculated from the number or magnitude of tree splits, cross-validated marginal permutation importance was calculated for the primary response, C_t/C_0 . For each of 10 controlled Random Forest seeds, shuffled five-fold cross-validation was performed. Within each validation fold, the values of one predictor were randomly permuted while the other predictor was retained

unchanged, and the resulting deterioration in predictive performance was evaluated over 50 independent permutations.

For predictor j , permutation importance was quantified using two complementary performance measures:

$$\Delta R_j^2 = R_{\text{base}}^2 - R_{\text{perm},j}^2 \quad (9)$$

and

$$\Delta \text{MAE}_j = \text{MAE}_{\text{perm},j} - \text{MAE}_{\text{base}}, \quad (10)$$

where the subscript “base” denotes performance on the unpermuted validation fold and “perm” denotes performance after permutation of predictor j . Positive values therefore indicate a deterioration in validation performance after disruption of the corresponding predictor. The fold- and permutation-level values were first averaged within each RF seed, after which the mean and sample standard deviation were calculated across the 10 seeds. The resulting standard deviations quantify model-seed variability rather than experimental uncertainty. Because this procedure used shuffled cross-validation, it represents within-domain variable importance and not out-of-time generalization.

A grouped-conditional permutation procedure was additionally applied using the same cross-fitted predictions. Nominal Cu-loading was permuted within each fixed reaction-time stratum, whereas reaction time was permuted within each fixed Cu-loading stratum. Fifty permutations were performed for each predictor, permutation scheme, and RF seed. The resulting decreases in OOF R^2 and increases in OOF MAE were calculated relative to the corresponding unpermuted cross-fitted predictions. This grouped procedure was interpreted as a design-based approximation to conditional importance rather than as a causal importance measure or conditional-randomization test.

The stability of the RF importance estimates was examined using 1000 synchronized moving-block bootstrap repetitions. In each repetition, blocks containing three adjacent reaction-time points were sampled with replacement, and the same sampled time sequence was applied to all five nominal Cu-loading groups. Thus, the Cu-loading balance was retained while partially preserving the temporal structure. A new RF model was fitted in every repetition, and the MDI, marginal PFI, and grouped-conditional PFI values were recorded. Each PFI estimate was based on 50 permutations and evaluated over the unchanged 5×13 reference grid. The distributions were summarized using their medians and 2.5th–97.5th percentiles. These percentile ranges characterize model and resampling stability rather than experimental uncertainty.

SHAP was used to obtain signed, observation-specific feature contributions and a complementary global importance measure. For an observation $\mathbf{x}_i$, the explained RF prediction was represented by the additive decomposition

$$\hat{f}(\mathbf{x}_i) = \phi_0 + \sum_{j=1}^p \phi_j(\mathbf{x}_i), \quad (11)$$

where ϕ_0 is the expected model output and $\phi_j(\mathbf{x}_i)$ is the SHAP contribution of predictor j to the prediction for observation i . A positive SHAP value indicates that the predictor increased the model output relative to ϕ_0 , whereas a negative value indicates that it decreased the model output.

For the complete feature set $F = \{1, \dots, p\}$, the SHAP value for predictor j was defined as

$$\phi_j(\mathbf{x}_i) = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [v_{\mathbf{x}_i}(S \cup \{j\}) - v_{\mathbf{x}_i}(S)], \quad (12)$$

where S is a feature subset that excludes predictor j , and $v_{\mathbf{x}_i}(S)$ is the expected model output when the features in S are fixed at their observed values for $\mathbf{x}_i$. The combinatorial weight averages the marginal contribution of predictor j over all possible feature orderings [37].

Because the selected models were tree ensembles, SHAP values were calculated using TreeExplainer, which provides a model-specific procedure for explaining tree-based predictions [48]. Global SHAP importance for predictor j was calculated as the mean absolute SHAP value across the N observations:

$$I_j^{\text{SHAP}} = \frac{1}{N} \sum_{i=1}^N |\phi_j(\mathbf{x}_i)|. \quad (13)$$

Mean absolute SHAP values retain the units of the corresponding model output. Therefore, their numerical magnitudes were compared only between reaction time and nominal Cu-loading within the same target model and not directly across apparent removal efficiency, C_t/C_0 , and $\ln(C_0/C_t)$. Signed local SHAP values were used to examine the direction and distribution of the model-based associations for individual observations.

Because apparent removal efficiency is an affine transformation of C_t/C_0 , the corresponding RF predictions and SHAP contributions are expected to exhibit the same structure with opposite directions and a 100-fold difference in magnitude when identical model settings are used. This relationship is an algebraic consequence of the target definitions and was not interpreted as independent evidence that the model learned a mass-balance mechanism.

First-order accumulated local effects (ALE) were calculated for the primary response, C_t/C_0 , to provide a complementary representation of the local associations learned by the RF model. Because nominal Cu-loading and reaction time were observed on discrete experimental grids, continuous interpolation and artificial intermediate predictor values were not used. The ALE calculations were therefore restricted to the experimentally observed predictor levels.

Let $z_{j,k-1}$ and $z_{j,k}$ denote two consecutive observed levels of predictor j . The mean local prediction difference over the corresponding interval was calculated as

$$\hat{\delta}_{j,k} = \frac{1}{n_{j,k}} \sum_{i: x_{ij}=z_{j,k}} \left[\hat{f}(z_{j,k}, \mathbf{x}_{i,-j}) - \hat{f}(z_{j,k-1}, \mathbf{x}_{i,-j}) \right], \quad (14)$$

where $\mathbf{x}_{i,-j}$ denotes the remaining predictor and $n_{j,k}$ is the number of observations at the upper endpoint of interval k . The estimated local effects were accumulated over consecutive intervals and subsequently centered to have a weighted mean of zero over the empirical distribution of the corresponding predictor. Consequently, positive and negative ALE values indicate increases and decreases, respectively, in the model prediction relative to its centered mean rather than absolute changes in the measured response.

For nominal Cu-loading, the ALE values were calculated only at the five experimentally investigated levels of 0%, 3%, 5%, 7.5%, and 10%. For reaction time, ALE values were calculated at the 13 experimentally observed time points between 0 and 120 min. The analysis was repeated for 10 controlled RF seeds. Mean ALE values and sample standard deviations were then calculated across the 10 seed-specific estimates. These standard deviations quantify model-seed variability and should not be interpreted as bootstrap or experimental uncertainty. Finally, neither MDI nor SHAP establishes causality. The resulting attributions describe how the fitted RF models used reaction time and nominal Cu-loading to generate predictions within the observed dataset. They do not directly demonstrate Cu-induced lattice substitution, defect formation, charge-transfer behavior, or other physicochemical mechanisms.

In addition to SHAP and MDI, one-way partial dependence was calculated for the primary response, C_t/C_0 , to examine the marginal response represented by the fitted RF model [28]. For predictor x_j at a specified value z , the partial-dependence estimate was calculated as

$$\widehat{\text{PD}}_j(z) = \frac{1}{N} \sum_{i=1}^N \hat{f}(z, \mathbf{x}_{i,-j}), \quad (15)$$

where $\hat{f}$ denotes the fitted RF regression model and $\mathbf{x}_{i,-j}$ contains the observed values of the remaining predictor for observation i .

For each controlled random seed, an RF model with the prespecified hyperparameters was fitted to the complete dataset, and the resulting partial-dependence estimates were calculated separately. The reported curves represent the mean across 10 controlled RF seeds, and the error bars or shaded intervals denote ± 1 standard deviation across these fitted models.

For nominal Cu-loading, partial dependence was evaluated only at the experimentally investigated levels of 0%, 3%, 5%, 7.5%, and 10%. Reaction-time dependence was evaluated at the experimentally observed time points.

A discrete two-dimensional RF response map was constructed for C_t/C_0 using the five nominal Cu-loading levels and 13 reaction times included in the experimental dataset. Each cell was calculated as the mean prediction across 10 controlled RF seeds. Only the observed 5×13 grid was used, without interpolation, smoothing, extrapolation, or artificial intermediate Cu-loading values.

3.4.5. Performance Evaluation Metrics

Predictive performance was evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and mean squared error (MSE). For both five-fold cross-validation and leave-one-out cross-validation (LOOCV), the out-of-fold predictions were pooled within each controlled run before the performance metrics were calculated. Consequently, each cross-validation run produced one R^2 , one MAE, and one MSE value from the complete set of 65 out-of-fold predictions rather than separate metric values that were subsequently averaged across folds.

For the chronological evaluation, the metrics were calculated exclusively from the predictions for the $N_{\text{test}} = 10$ observations satisfying $t > 100$ min. At the prediction stage, each evaluation observation was excluded from the data used to fit its corresponding fold-specific or chronological-training model.

Let $\mathcal{E}$ denote an evaluation set containing $n_{\mathcal{E}}$ observations. The coefficient of determination was calculated as

$$R^2 = 1 - \frac{\sum_{i \in \mathcal{E}} (y_i - \hat{y}_i)^2}{\sum_{i \in \mathcal{E}} (y_i - \bar{y}_{\mathcal{E}})^2}, \quad (16)$$

where

$$\bar{y}_{\mathcal{E}} = \frac{1}{n_{\mathcal{E}}} \sum_{i \in \mathcal{E}} y_i \quad (17)$$

is the mean observed response in the corresponding evaluation set. MAE and MSE were calculated as

$$\text{MAE} = \frac{1}{n_{\mathcal{E}}} \sum_{i \in \mathcal{E}} |y_i - \hat{y}_i|, \quad (18)$$

and

$$\text{MSE} = \frac{1}{n_{\mathcal{E}}} \sum_{i \in \mathcal{E}} (y_i - \hat{y}_i)^2, \quad (19)$$

respectively. Here, y_i and $\hat{y}_i$ denote the observed and predicted responses for observation i .

The R^2 score quantifies prediction error relative to the variability of the observed response around its evaluation-set mean. A value close to 1 indicates strong agreement, whereas a negative value indicates that the prediction errors are larger than those obtained by using the evaluation-set mean as a constant baseline [49]. MAE represents the average absolute prediction error on the original response scale, whereas MSE gives greater weight to larger residuals because the errors are squared [50,51].

Because the response representations have different numerical scales, their MAE and MSE values were interpreted separately. For the target-specific multi-seed analysis, apparent removal efficiency was evaluated in normalized fractional form as $\eta/100$; therefore, its MAE and MSE were dimensionless. For the chronological evaluation, apparent-removal-efficiency predictions were transformed back to the percentage scale, with MAE reported in percentage points and MSE in squared percentage points. The error metrics for C_t/C_0 and $\ln(C_0/C_t)$ were dimensionless.

Apparent removal efficiency and the residual concentration ratio are related by the affine transformation $\eta = 100(1 - r)$, where $r = C_t/C_0$. Therefore, when their predictions satisfy the corresponding relationship $\hat{\eta} = 100(1 - \hat{r})$, their performance metrics obey

$$R_{\eta}^2 = R_r^2, \quad \text{MAE}_{\eta} = 100 \text{MAE}_r, \quad \text{MSE}_{\eta} = 10^4 \text{MSE}_r. \quad (20)$$

Accordingly, equal R^2 values and proportionally scaled error metrics for these two target representations are expected mathematical consequences of their definitions and do not constitute independent performance results.

For the multi-seed analyses, the complete validation procedure was performed separately for each of the ten prespecified seeds. Each seed produced one value for each performance metric from its pooled out-of-fold predictions. The reported values were summarized as the mean and standard deviation of these ten seed-level metrics. The standard deviation therefore describes sensitivity to the controlled data-partitioning and/or model-initialization settings and should not be interpreted as experimental uncertainty or as a confidence interval for population-level predictive performance.

3.4.6. Algorithmic Workflow of the Predictive Framework

Algorithm 1 formalizes the computational workflow, while Figure 1 provides a complementary visual representation of the RF architecture and its integration into the analytical framework. Algorithm 1 specifies the sequence of model comparison, target-specific robustness evaluation, chronological diagnosis, and post hoc interpretation, whereas Figure 1 summarizes the relationships among these stages.

As shown in Figure 1A, the RF model receives nominal Cu-loading and reaction time as its predictor vector. Bootstrap samples drawn with replacement are used to construct an ensemble of 100 regression trees. Each tree recursively partitions the predictor space using feature-threshold decisions and produces a continuous response estimate. The individual tree estimates are then averaged to obtain the final RF prediction. This ensemble structure enables the representation of nonlinear temporal patterns and composition-dependent associations while reducing dependence on any single regression tree.

Figure 1B connects this estimator architecture to the three main phases formalized in Algorithm 1. Phase I comprises two analytically distinct stages. In Phase I-A, RF, KNN, MLP, and SVR are compared using only the predefined primary screening target $r = C_t/C_0$. Restricting candidate-model comparison to this single target prevents normalized apparent removal efficiency from being treated as an independent model-selection outcome, because $e = \eta/100 = 1 - r$. The candidate models are evaluated across 10 controlled seeds using shuffled five-fold CV and LOOCV. For KNN, MLP, and SVR, predictor standardization is performed within the cross-validation pipeline, with the scaler fitted exclusively to

each training fold and subsequently applied to the corresponding validation fold. RF is evaluated on the original predictor scales because tree-based partitioning does not require feature standardization.

After RF is retained, Phase I-B evaluates its target-specific robustness separately for the two fitted responses: the residual concentration ratio $r = C_t/C_0$ and the logarithmically transformed response $\ell = \ln(C_0/C_t)$. The final RF configuration is evaluated across the same 10 controlled seeds using five-fold CV and LOOCV, and the resulting R^2 , MAE, and MSE values are summarized as the mean and standard deviation across the controlled runs. Normalized apparent removal efficiency and its predictions are derived as $e = 1 - r$ and $\hat{e} = 1 - \hat{r}$; their metrics therefore represent the same predictive outcome on an inversely related scale rather than an independent model fit.

Phase II applies the chronological out-of-time diagnostic evaluation to the selected RF configuration. The model is fitted separately for the two fitted responses, C_t/C_0 and $\ln(C_0/C_t)$, using the historical observations collected at $t \leq 100$ min and is evaluated on the later observations at $t > 100$ min. Apparent-removal-efficiency predictions are derived from the C_t/C_0 predictions. This fixed temporal separation avoids randomized mixing of historical and later-time measurements and examines whether the fitted RF relationship remains predictive within the investigated $100 < t \leq 120$ min interval. Because the analysis uses a single temporal cutoff and catalyst compositions already represented in the training data, it is interpreted as a later-time diagnostic rather than as evidence of unrestricted extrapolation or recovery of an explicit physicochemical kinetic law.

In Phase III, the final RF specification is refitted separately for r and ℓ using the complete dataset for post hoc interpretation; the efficiency-scale attributions are obtained from the affine transformation of r . SHAP quantifies the direction and magnitude of observation-level feature contributions, whereas normalized MDI provides a complementary global summary of the impurity reduction assigned to reaction time and nominal Cu-loading. These analyses describe associations represented by the fitted RF models and are analytically distinct from the out-of-sample performance evaluations conducted in Phases I and II. Accordingly, neither the SHAP nor MDI results are interpreted as direct evidence of causal, structural, electronic, or mechanistic effects.

The Python 3.13 scripts implementing the candidate-model comparison, target-specific multi-seed evaluation, chronological diagnostic analysis, and SHAP and MDI procedures are publicly available at <https://github.com/metinzontul/ML-Photocatalysis-Kinetics> (accessed on 31 August 2026).

4. Results and Discussion

4.1. Time-Dependent Removal Performance Under Irradiation

Figure 2 presents the time-dependent apparent removal profiles of the AC@NiO samples containing nominal Cu-loadings of 0%, 3%, 5%, 7.5%, and 10%. The same post-dark-normalized concentration data are shown as (a) apparent removal efficiency during irradiation, (b) residual concentration ratio C_t/C_0 , and (c) logarithmically transformed concentration response $\ln(C_0/C_t)$.

All five nominal Cu-containing compositions exhibited rapid initial apparent removal during the first 10–20 min of irradiation, followed by a more gradual concentration decrease at longer times. At 10 min, C_t/C_0 ranged from approximately 0.23 to 0.46, depending on the nominal Cu-loading. Because independent dark-only and catalyst-free photolysis controls were unavailable, this initial decrease cannot be attributed exclusively to photocatalysis and may include contributions from continued adsorption, direct photolysis, and photocatalytic transformation.

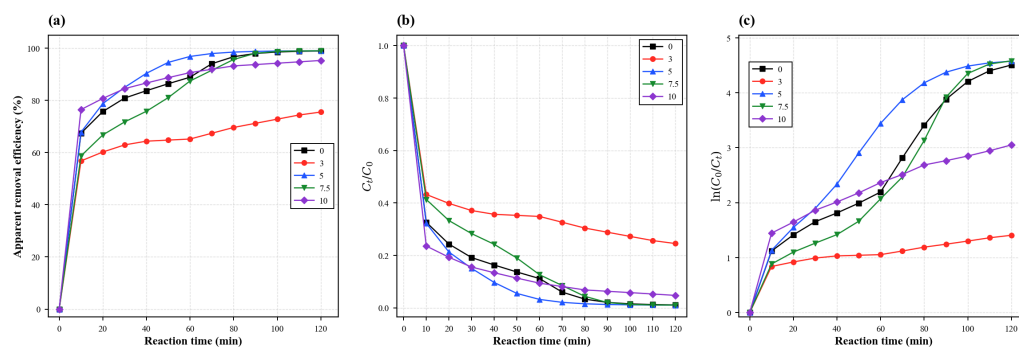


Figure 2. Time-dependent responses of nominal Cu-containing AC@NiO samples under irradiation: (a) apparent removal efficiency, (b) C_t/C_0 , and (c) $\ln(C_0/C_t)$.

The nominal 5% Cu-containing sample exhibited the lowest residual concentration over most of the investigated interval, with C_t/C_0 decreasing below approximately 0.05 by 60 min and reaching approximately 0.01 at 120 min. The nominal 10% sample showed a rapid initial response but retained a higher final residual concentration than the nominal 0%, 5%, and 7.5% samples. The nominal 7.5% sample exhibited an intermediate trajectory and approached the low final residual concentrations of the nominal 0% and 5% samples at longer irradiation times. In contrast, the nominal 3% sample retained the highest C_t/C_0 over most of the interval and reached approximately 0.20–0.22 at 120 min, corresponding to an apparent removal efficiency of approximately 78–80% relative to the post-dark baseline.

The transformation $\ln(C_0/C_t)$ was used to examine the concentration profiles on a logarithmic scale. For a process adequately represented by an ideal apparent pseudo-first-order relationship,

$$\ln\left(\frac{C_0}{C_t}\right) = k_{\text{app}}t, \quad (21)$$

where k_{app} denotes the apparent rate constant. Under this idealized representation, a plot of $\ln(C_0/C_t)$ against irradiation time would be approximately linear, pass through the origin, and have a slope of k_{app} [52].

For descriptive comparison among the five compositions, ordinary least-squares regressions were performed over the complete 0–120 min interval using $\ln(C_0/C_t) = b + k_{\text{app}}t$. The intercept b was estimated rather than constrained to zero so that departures from the idealized relationship in Equation (21) could be assessed. The composition-specific fitting results are reported in Supplementary Table S1.

The fitted k_{app} values ranged from 0.00754 min^{-1} for the nominal 3% sample to 0.03895 min^{-1} for the nominal 7.5% sample. The nominal 5% sample yielded $k_{\text{app}} = 0.03756 \text{ min}^{-1}$ and $R^2 = 0.9310$. Across the five compositions, the R^2 values ranged from 0.6830 to 0.9686, demonstrating that the linearized pseudo-first-order representation did not describe all profiles equally well over the complete irradiation interval. The nonzero fitted intercepts are also consistent with departures from the idealized zero-intercept relationship.

These regressions are presented as descriptive kinetic benchmarks. Their in-sample R^2 values are not directly comparable with the cross-validated or chronological R^2 values reported for the ML models. Furthermore, because the measured concentration decrease may include contributions from continued adsorption, direct photolysis, and photocatalytic transformation, the fitted k_{app} values should be interpreted as apparent removal-rate parameters rather than intrinsic photocatalytic rate constants. Neither the fitted slopes nor deviations from linearity are used to infer a rate-limiting step, surface-activation process, intermediate-species competition, or other mechanistic transition.

Overall, the observed profiles indicate a non-monotonic association between nominal Cu-loading and apparent removal within the tested compositional range. Non-monotonic loading effects have also been reported in other modified photocatalyst systems [53]; however, this comparison does not establish an equivalent mechanism in the present system. The nominal 5% composition is identified as the most favorable only among the five discrete loadings investigated and should not be interpreted as a local, global, or universal optimum. Because the profiles were not supported by independent experimental replicates, differences among the compositions are interpreted descriptively rather than as statistically tested differences. These trajectories provide the experimental basis for the subsequent target-specific ML analyses.

4.2. Physicochemical Characterization of the Nominal 0% and 5% Cu-Containing AC@NiO Samples

4.2.1. XRD Analysis

The XRD patterns of the undoped and nominal 5% Cu-containing AC@NiO samples are presented in Figure 3. The principal diffraction reflections observed at approximately $2\theta = 37^\circ$, 43° , 63° , and 75° were indexed to the (111), (200), (220), and (311) planes, respectively, of cubic NiO with the $Fm\bar{3}m$ space group. These reflections are consistent with the formation of a crystalline cubic NiO phase in both investigated samples.

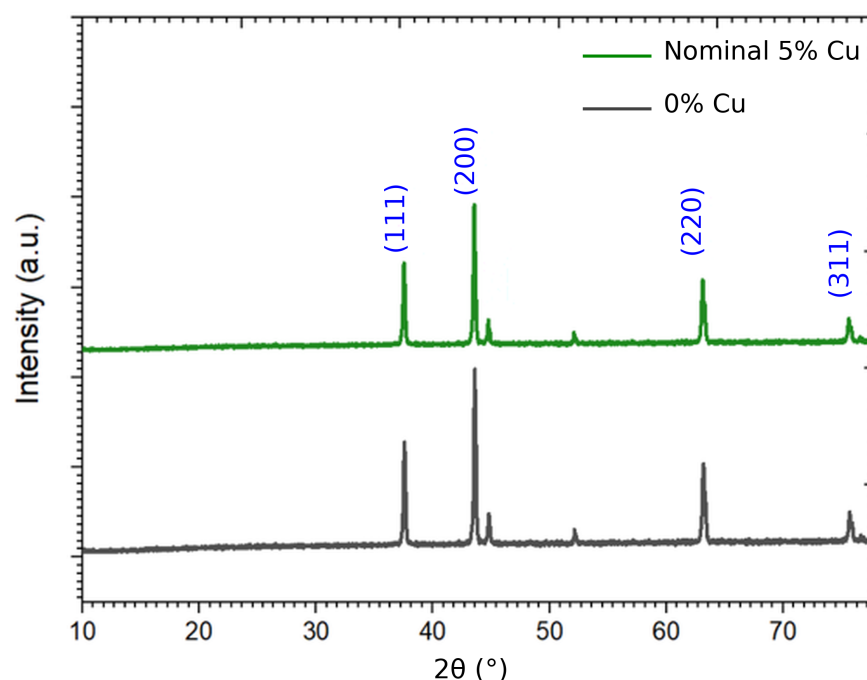


Figure 3. XRD patterns of the undoped (0% Cu) and nominal 5% Cu-containing AC@NiO samples.

The two patterns show broadly similar positions of the major diffraction reflections. No new reflections unique to the nominal 5% Cu-containing sample can be unambiguously assigned to crystalline CuO or Cu₂O. However, the absence of clearly resolved secondary-phase reflections does not demonstrate substitutional incorporation of Cu into the NiO lattice and does not exclude the presence of amorphous, highly dispersed, or low-concentration Cu-containing species below the detection limit of XRD. The minor unassigned features observed in the patterns were not used for phase attribution or mechanistic interpretation.

4.2.2. Scanning Electron Microscopy (SEM) Analysis

The surface morphologies of the 0% Cu and nominal 5% Cu-containing AC@NiO samples were examined by SEM, as presented in Figure 4. Both samples exhibit heterogeneous surfaces composed of irregularly shaped and agglomerated particles, together with rough surface textures and visible interparticle voids. These observations indicate an open and heterogeneous micrometre-scale morphology for both AC@NiO samples.

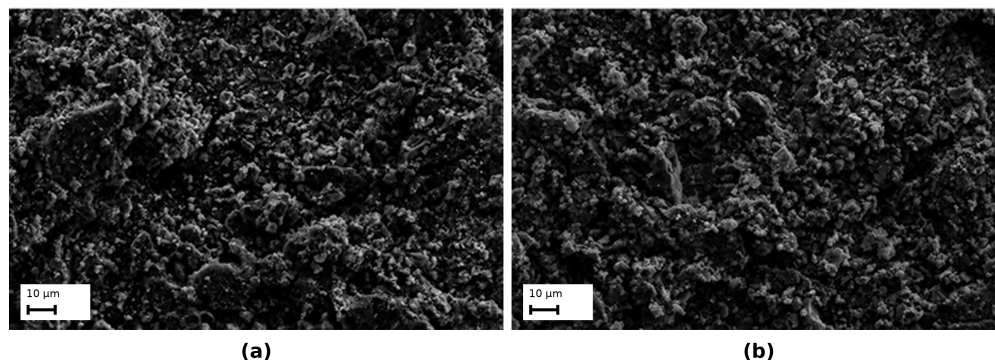


Figure 4. SEM micrographs of the (a) nominal 0% Cu reference and (b) nominal 5% Cu-containing AC@NiO samples recorded at a magnification of 1000 $\times$. The scale bars represent 10 μm .

The overall morphology of the nominal 5% Cu-containing sample is broadly comparable to that of the 0% Cu sample. Although local variations in particle packing and apparent surface roughness are visible, no pronounced modification of the micrometre-scale morphology is observed at the investigated magnification. Thus, the nominal Cu addition did not substantially alter the overall irregular and agglomerated morphology of the AC@NiO material. The elemental compositions of the samples were subsequently examined by EDX analysis.

4.2.3. EDX Analysis

The local elemental compositions of the nominal 0% Cu reference and nominal 5% Cu-containing AC@NiO samples were examined by EDX analysis, and the corresponding spectra are presented in Figure 5. For the nominal 0% Cu reference sample, EDX detected C, O, and Ni at 87.48, 6.24, and 6.28 at.%, respectively. The dominant C signal is consistent with the activated-carbon support, whereas the joint detection of Ni and O is consistent with the Ni-containing oxide component identified as crystalline NiO by XRD.

For the nominal 5% Cu-containing sample, C, O, Ni, and Cu were detected at 87.78, 5.88, 5.13, and 1.21 at.%, respectively. The Cu signal confirms the presence of Cu within the analyzed region but does not establish substitutional incorporation into the NiO lattice, its oxidation state, a spatially uniform distribution, or the bulk Cu concentration. Because EDX provides local and semiquantitative compositional information, the measured Cu atomic percentage was not equated directly with the nominal precursor composition. No additional prominent elemental signals were detected within the detection limits of the displayed spectra.

4.2.4. UV–Vis DRS and Tauc Analysis

The UV–Vis diffuse-reflectance profiles, expressed using the Kubelka–Munk function $F(R)$, and the corresponding direct-transition Tauc representations are presented in Figure 6. The two samples exhibited closely similar spectral profiles, with a pronounced response in the UV region at approximately 260–280 nm, followed by a gradual decrease toward the visible region. The nominal 5% Cu-containing sample showed slightly higher $F(R)$ values, particularly between approximately 350 and 450 nm. This difference indicates a modest alteration of the optical response following nominal Cu addition; however, it was interpreted descriptively.

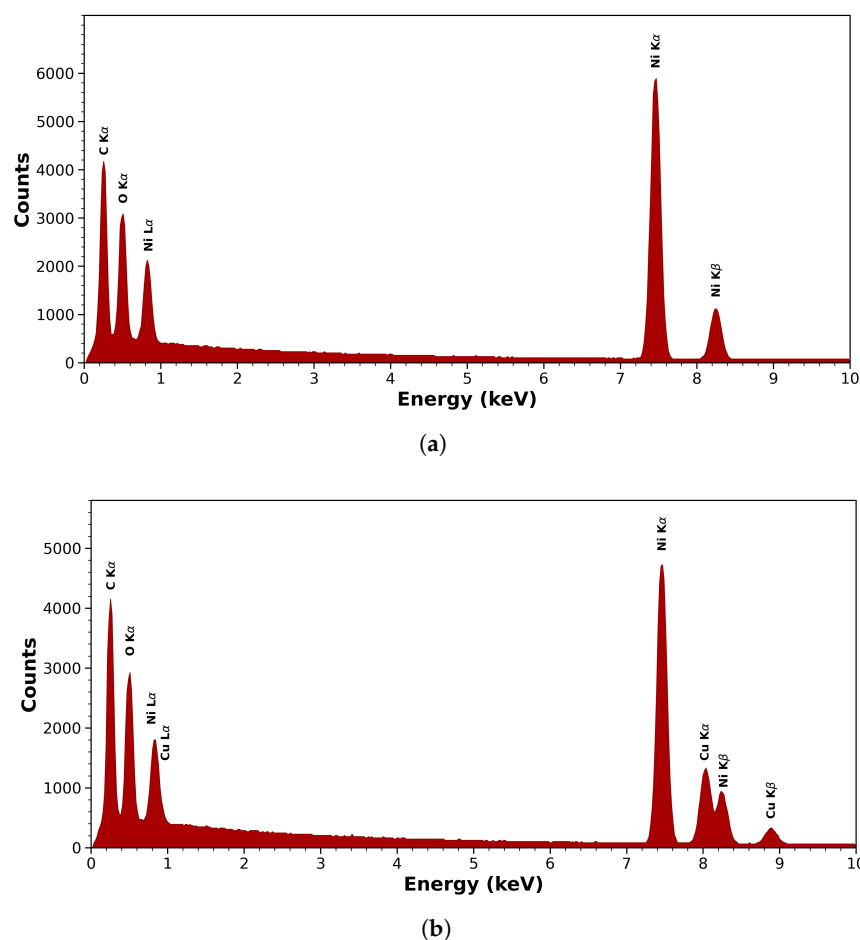


Figure 5. Spectra of the AC@NiO samples: (a) nominal 0% Cu reference sample and (b) nominal 5% Cu-containing sample.

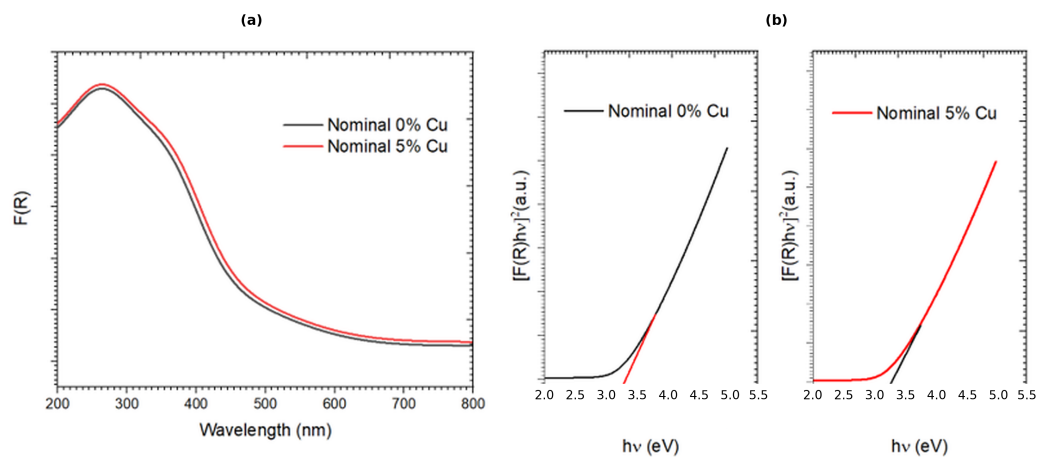


Figure 6. UV-Vis diffuse-reflectance results for the nominal 0% Cu reference and nominal 5% Cu-containing AC@NiO samples: (a) Kubelka–Munk function, $F(R)$, as a function of wavelength; and (b) direct-transition Tauc representations, $[F(R)hv]^2$, as a function of photon energy. Black and red denote the nominal 0% and 5% Cu-containing samples, respectively. The straight-line segments show the selected linear regions and their extrapolation to $[F(R)hv]^2 = 0$.

Under the direct-transition assumption adopted in the Tauc representations, extrapolation of the selected linear regions to $[F(R)hv]^2 = 0$ yielded estimated optical band-gap values of approximately 3.12 eV for the nominal 0% Cu reference and 3.08 eV for the nominal 5% Cu-containing sample. The resulting difference of approximately 0.04 eV is

small. In the absence of independently repeated spectra and fitting-uncertainty estimates, it should be interpreted as a nominal shift under the adopted Tauc representation rather than as conclusive evidence of physically meaningful band-gap narrowing. Moreover, the DRS and Tauc results alone do not establish Cu incorporation into the NiO lattice, the formation of Cu-induced defect states, Ni vacancies or Ni^{3+} species, interfacial charge transfer, or suppression of electron-hole recombination.

4.3. Comparative Model Selection and Multi-Seed Robustness

Before conducting the detailed target-specific analysis, the predictive performance of four nonlinear regression algorithms—K-Nearest Neighbors (KNN), Multi-Layer Perceptron (MLP), Random Forest (RF), and Support Vector Regression (SVR)—was evaluated using the residual concentration ratio C_t/C_0 as the primary dimensionless response. The same predictor variables, namely reaction time and nominal Cu-loading, were used for all models.

Model performance was assessed using shuffled five-fold cross-validation and leave-one-out cross-validation (LOOCV). To examine sensitivity to shuffled data partitions and stochastic model initialization, the evaluation was repeated across 10 controlled random seeds. Identical five-fold partitions were used for all candidate algorithms within each seed, thereby supporting a paired comparison under the same training and validation subsets. The resulting performance metrics are presented in Table 1.

As shown in Table 1, RF provided the highest mean predictive performance among the four candidate algorithms under both validation strategies. Under LOOCV, RF achieved an R^2 of 0.9670 ± 0.0025 , compared with 0.7924 ± 0.0000 for SVR, 0.6910 ± 0.0000 for KNN, and 0.4716 ± 0.0477 for MLP. RF also produced the lowest LOOCV error values, with an MAE of 0.0382 ± 0.0016 and an MSE of 0.0026 ± 0.0002 .

The relatively small standard deviations obtained for RF indicate limited sensitivity to the examined random initializations. The zero standard deviations reported for KNN and SVR under LOOCV arise because these models are deterministic when their hyperparameters and preprocessing procedures are fixed and because the LOOCV partitions are identical across runs. These values should therefore not be interpreted as experimental uncertainty or as evidence that KNN and SVR are intrinsically more stable than RF.

Based on its higher cross-validated R^2 values and lower prediction errors, RF was retained for the subsequent target-specific, chronological, and explainability analyses. The performance values in Table 1 correspond to the preliminary model-screening stage; the results obtained using the final target-specific RF configuration are reported separately in Table 2.

4.4. Target-Specific Performance and Multi-Seed Robustness of the RF Model

The preliminary model-screening results presented in Table 1 identified RF as the best-performing candidate among KNN, MLP, RF, and SVR. Following this screening stage, the final RF hyperparameters were fixed before conducting the target-specific evaluation. The final RF configuration was evaluated separately for the two fitted responses, the residual concentration ratio C_t/C_0 and the logarithmically transformed concentration response $\ln(C_0/C_t)$. Normalized apparent removal efficiency and its predictions were derived from C_t/C_0 , not fitted as a third response. The analysis was repeated across 10 controlled random seeds using shuffled five-fold cross-validation and leave-one-out cross-validation (LOOCV).

Normalized efficiency was used in the numerical evaluation because

$$\frac{\eta}{100} = 1 - \frac{C_t}{C_0}. \quad (22)$$

Accordingly, the R^2 , MAE, and MSE values reported for $\eta/100$ are the same predictive outcome as those obtained for C_t/C_0 on an inversely related dimensionless scale. For graphical presentation, the derived normalized-efficiency predictions were transformed back to the percentage scale. The target-specific performance results are presented in Table 2.

As shown in Table 2, the final regularized RF model yielded mean R^2 values of 0.9546 ± 0.0137 under five-fold cross-validation and 0.9691 ± 0.0013 under LOOCV for C_t/C_0 and $\eta/100$. Because these responses are deterministically related, their identical performance metrics represent the same predictive outcome on inversely related scales and should not be interpreted as independent evidence of model performance.

For $\ln(C_0/C_t)$, the corresponding mean R^2 values were 0.8542 ± 0.0248 under five-fold cross-validation and 0.9152 ± 0.0090 under LOOCV. The lower predictive agreement obtained for this response indicates that the RF model reproduced the logarithmically transformed concentration response less accurately than the affinely related response representations. Here, $\ln(C_0/C_t)$ denotes a transformed concentration response rather than an independently estimated apparent rate constant.

The higher mean scores obtained under LOOCV do not imply that LOOCV was necessarily more stringent than five-fold cross-validation. Each LOOCV model was fitted using $N - 1 = 64$ observations, and the resulting training sets overlapped substantially. Because the LOOCV partitions were fixed across the 10 runs, the reported standard deviations primarily quantify stochastic variation in RF fitting rather than variation in fold assignment. They should not be interpreted as experimental uncertainty or confidence intervals for generalization performance.

Figure 7 provides a visual assessment of the LOOCV predictions. Panels (a) and (b) represent the same predictive outcome on affinely related response scales and do not provide independent evidence of model performance. The separately fitted logarithmic-response model in panel (c) has a lower R^2 in the displayed run.

4.5. Robustness Against Temporal Data Leakage: Out-of-Time Validation

To evaluate model performance at later reaction times while minimizing temporal mixing between the training and test subsets, an out-of-time chronological validation strategy was implemented. Random shuffling was disabled because of the sequential structure of the kinetic dataset.

The dataset contained 65 observations and was partitioned using the temporal cut-off $t_c = 100$ min. The RF model was fitted using the 55 historical observations satisfying $t \leq 100$ min and evaluated on the remaining 10 later-time observations satisfying $t > 100$ min. For each chronological RF fit, the later-time observations were excluded from model fitting. The resulting performance metrics are reported in Table 3. For apparent removal efficiency, MAE is expressed in percentage points and MSE in squared percentage points; the remaining error metrics are dimensionless.

Table 3. Out-of-time chronological validation results for the regularized RF model trained on historical observations ($t \leq 100$ min, $N_{\text{train}} = 55$) and evaluated on unseen later-time observations ($t > 100$ min, $N_{\text{test}} = 10$). Results are reported as the mean $\pm$ standard deviation across 10 controlled random-seed evaluations.

Validation Strategy	Output Variable	R^2 Score	MAE	MSE
Chronological ($t \leq 100$ min $\rightarrow t > 100$ min)	C_t/C_0	0.9668 ± 0.0073	0.0152 ± 0.0020	0.0003 ± 0.0001
Chronological ($t \leq 100$ min $\rightarrow t > 100$ min)	Apparent removal efficiency (%)	0.9668 ± 0.0073	1.5191 ± 0.2016	2.8900 ± 0.6400
Chronological ($t \leq 100$ min $\rightarrow t > 100$ min)	$\ln(C_0/C_t)$	0.8890 ± 0.0163	0.3700 ± 0.0244	0.1739 ± 0.0256

The regularized RF model achieved a mean chronological-evaluation R^2 of 0.9668 ± 0.0073 for both C_t/C_0 and apparent removal efficiency. The equality of these values and

the scaling of their error metrics follow from the deterministic relationship defined in Equation (3). Therefore, these two results should not be interpreted as independent performance outcomes.

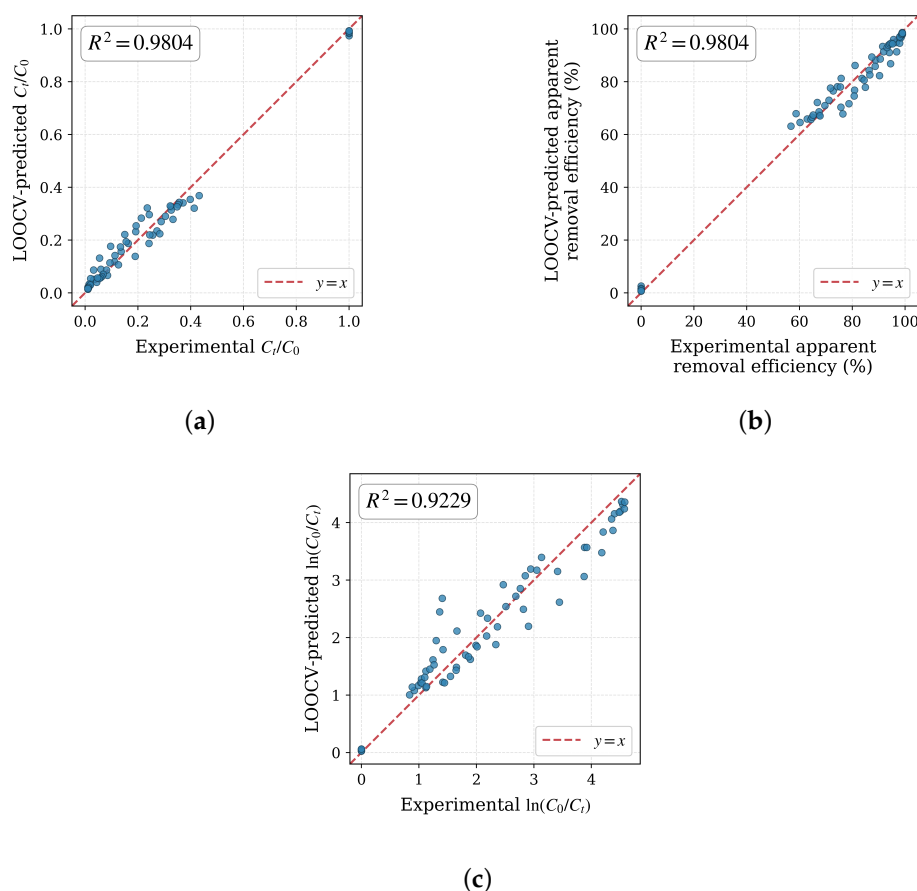


Figure 7. Experimental versus LOOCV-predicted responses for the reference RF configuration (seed 0): (a) C_t/C_0 , (b) apparent removal efficiency (%), and (c) $\ln(C_0/C_t)$. Panel (b) was derived from panel (a) using $100(1 - r)$ for both observations and predictions, where $r = C_t/C_0$. Red dashed lines indicate ideal agreement ($y = x$). Panel-specific R^2 values refer to this single run; Table 2 reports the mean and standard deviation across 10 controlled runs (seeds 0–9).

For the logarithmically transformed response $\ln(C_0/C_t)$, the regularized RF model achieved a mean chronological-evaluation R^2 of 0.8890 ± 0.0163 across the 10 controlled random seeds. Although this value was lower than those obtained for C_t/C_0 and apparent removal efficiency, it represents predictive agreement only within the fixed later-time evaluation set and does not establish temporal extrapolation.

The chronological results describe predictions for the same five compositions over a fixed later-time interval. Their interpretation is limited by the small evaluation subset and the post-selection design described in Section 3.4.3.

Table 4 reports the additional RF model-complexity comparison. None of the examined constraints increased mean R^2 under any of the three evaluation protocols. The depth-5 model retained relatively high predictive agreement, whereas the depth-3 restriction markedly reduced performance, particularly in the chronological diagnostic. Increasing the minimum leaf size, alone or together with the depth-5 constraint, also reduced mean R^2 . The unrestricted configuration was retained as the reference model. This comparison does not establish that it is free from small-sample overfitting or that it generalizes to independent data.

Table 4. RF complexity sensitivity for C_t/C_0 (mean $R^2 \pm$ sample SD across ten seeds).

Maximum Tree Depth	Minimum Leaf Size	Five-Fold CV	LOOCV	Chronological Diagnostic
None	1	0.9611 ± 0.0158	0.9785 ± 0.0014	0.9902 ± 0.0039
3	1	0.8759 ± 0.0137	0.8809 ± 0.0025	0.2878 ± 0.0295
5	1	0.9546 ± 0.0137	0.9691 ± 0.0013	0.9668 ± 0.0073
None	2	0.9160 ± 0.0480	0.9596 ± 0.0033	0.9568 ± 0.0074
None	3	0.8159 ± 0.0533	0.8999 ± 0.0094	0.8428 ± 0.0167
5	2	0.9146 ± 0.0477	0.9569 ± 0.0033	0.9467 ± 0.0079

None denotes unrestricted depth. All configurations used 100 trees, seeds 0–9, and matched validation partitions within each seed. Chronological training/evaluation used 55/10 observations. SD describes computational variability, not experimental uncertainty. The configuration with a constrained maximum depth ($\text{max_depth} = 5$) was selected as the reference model to reduce the risk of overfitting in the limited sample space.

For every tested configuration and seed, the chronological RF gave identical predictions at 100, 110, and 120 min for a fixed composition. The holdout scores therefore evaluate terminal-leaf boundary predictions, not an extrapolated temporal trend. Carrying each composition's measured 100-min value forward yielded $R^2 = 0.9865$, $\text{MAE} = 0.00719$, and $\text{MSE} = 1.17 \times 10^{-4}$ on the same 10 observations. This simple reference had a slightly lower R^2 but a lower MAE than the unrestricted RF (mean $\text{MAE} = 0.00784$). Thus, a high chronological R^2 alone does not demonstrate that the RF learned the subsequent time evolution. The reused holdout and row-wise cross-validation do not provide independent experimental validation.

4.6. Explainable AI: SHAP Analysis, Feature Importance, and Partial Dependence

To interpret the predictions of the RF model, complementary feature-attribution and response-pattern approaches were employed, including SHAP, impurity-based and permutation importance, PDP, and ALE. SHAP was used to quantify the magnitude and direction of individual feature contributions, whereas the normalized mean decrease in impurity (MDI) provided a global summary of the relative importance assigned to each predictor by the fitted RF model.

The mean absolute SHAP values reported in Table 5 indicate that reaction time made the largest contribution to predictions of apparent removal efficiency and C_t/C_0 . For apparent removal efficiency, the mean absolute SHAP value was 15.096 percentage points for reaction time and 5.617 percentage points for the nominal Cu-loading. The corresponding values for C_t/C_0 were 0.1510 and 0.0562, respectively.

The approximately 100-fold scaling between the SHAP values for apparent removal efficiency and C_t/C_0 follows from the deterministic relationship defined in Equation (3). This proportionality therefore represents an algebraic consistency check and should not be interpreted as independent evidence that the RF model learned a physical mass-balance law or avoided overfitting.

For the logarithmically transformed response $\ln(C_0/C_t)$, the mean absolute SHAP values were 0.9029 for reaction time and 0.4553 for the nominal Cu-loading. Thus, reaction time remained the dominant predictor, while the nominal Cu-loading showed a larger relative contribution for this target than for the two linearly related responses. Because SHAP values are expressed on the scale of the corresponding model output, their absolute magnitudes should not be compared directly across different target variables.

The SHAP summary plots in Figure 8 show the distribution and direction of the observation-level feature contributions. Higher reaction times generally increased the predicted apparent removal efficiency and $\ln(C_0/C_t)$ while decreasing the predicted residual concentration ratio C_t/C_0 , consistent with the temporal progression of pollutant removal.

Relative to reaction time, the Cu-loading ratio exhibited smaller and less monotonic SHAP contributions, with both the magnitude and direction of its effects varying across observations. This pattern indicates a nonuniform model-based association between the Cu-loading ratio and the predicted responses.

Table 5. Mean absolute SHAP values for reaction time and Cu-doping ratio. SHAP values are expressed on the scale of the corresponding model output and are therefore not directly comparable across target variables.

Target Output	Reaction Time (Mean SHAP)	Nominal Cu-Loading (Mean SHAP)
Efficiency (%)	15.096	5.617
C_t/C_0	0.1510	0.0562
$\ln(C_0/C_t)$	0.9029	0.4553

The normalized impurity-based importance values shown in Figure 9 provide a complementary global assessment. Reaction time was assigned 89.15% of the total normalized impurity-based importance for both apparent removal efficiency and C_t/C_0 , whereas the nominal Cu-loading was assigned 10.85%. For $\ln(C_0/C_t)$, the relative importance assigned to reaction time decreased to 69.51%, while that assigned to the nominal Cu-loading increased to 30.49%.

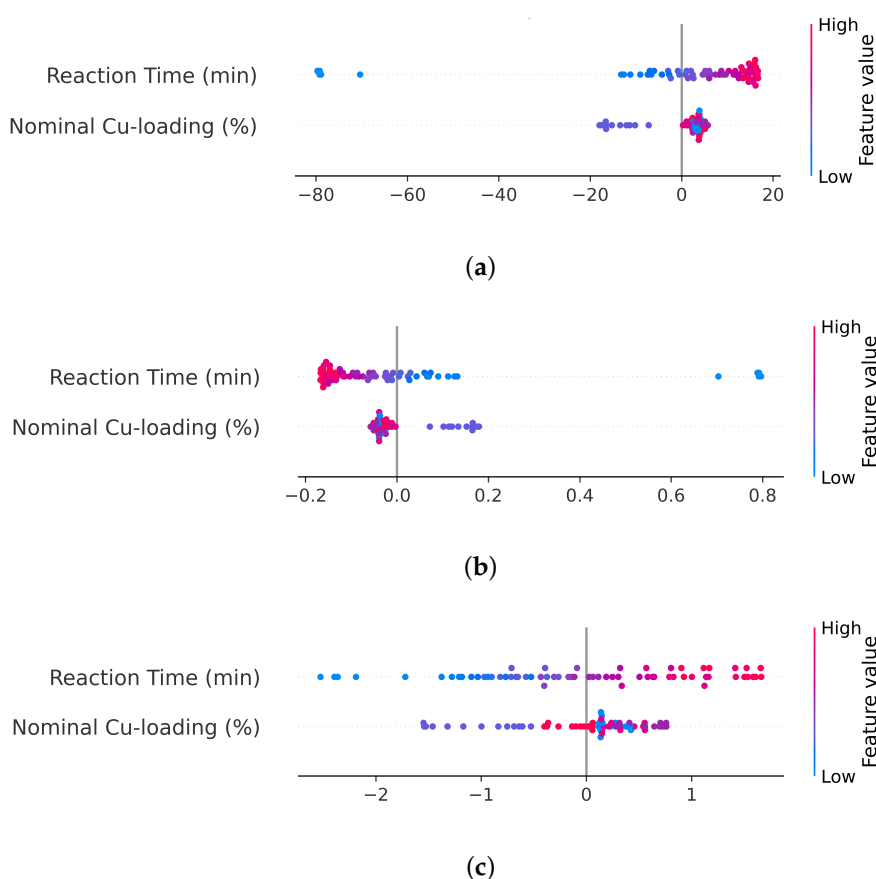


Figure 8. SHAP summary plots showing the magnitude and direction of the contributions of reaction time and nominal Cu-loading to predictions of (a) apparent removal efficiency, (b) C_t/C_0 , and (c) $\ln(C_0/C_t)$. The color scale represents the corresponding feature values, ranging from low (blue) to high (red).

These percentages represent shares of the total impurity reduction assigned by the fitted RF model; they should not be interpreted as percentages of explained variance. Nevertheless, the SHAP and MDI results consistently indicate that reaction time was the

dominant predictor for all three responses and that the relative predictive contribution of the nominal Cu-loading was greater for $\ln(C_0/C_t)$. These associations are model-based and do not, by themselves, establish causal effects, intrinsic reaction-rate constants, or Cu-induced morphological and electronic mechanisms.

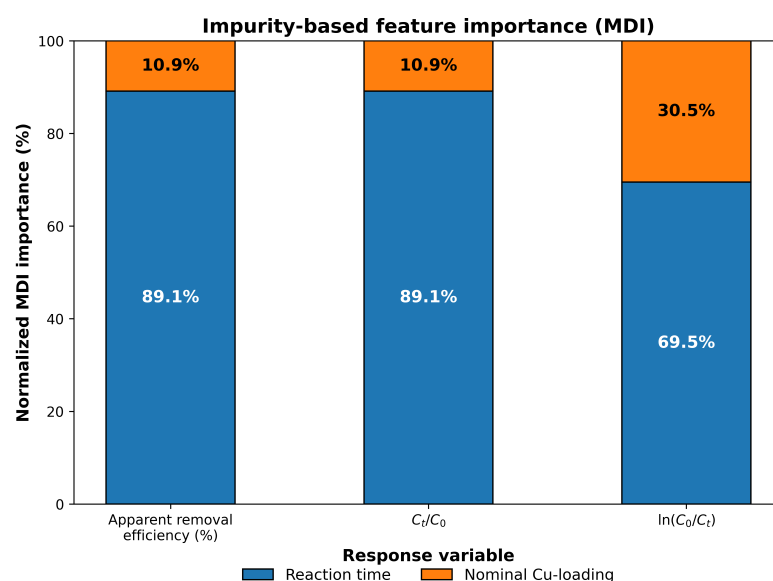


Figure 9. Normalized mean decrease in impurity (MDI) feature importance assigned by the RF model to reaction time and nominal Cu-loading for predicting apparent removal efficiency, C_t/C_0 , and $\ln(C_0/C_t)$.

Because reaction time contained 13 distinct values whereas nominal Cu-loading contained only five, the MDI ranking was further examined using cross-validated marginal permutation importance. As shown in Figure 10, permutation of reaction time produced a mean decrease in validation R^2 of 1.496 ± 0.122 , compared with 0.321 ± 0.086 for nominal Cu-loading. Similarly, permutation of reaction time increased validation MAE by 0.1548 ± 0.0072 , whereas permutation of nominal Cu-loading increased MAE by 0.0453 ± 0.0046 . Thus, both permutation-based performance measures indicate that the RF predictions were more sensitive to disruption of reaction time than to disruption of nominal Cu loading.

A decrease in R^2 greater than unity is mathematically possible because the permuted predictions can yield negative validation R^2 values. These results support the model-based dominance of reaction time without relying directly on the split frequencies used by MDI. Nevertheless, marginal permutation importance quantifies predictive dependence within the fitted model and should not be interpreted as a causal, mechanistic, or physicochemical contribution of either predictor.

Bars represent the mean importance across 10 controlled RF seeds, and error bars indicate ± 1 standard deviation across seeds. For each seed, importance was calculated using shuffled five-fold cross-validation with 50 permutations per validation fold. The error bars represent model-seed variability rather than experimental uncertainty.

The potential effect of predictor cardinality on the MDI ranking was examined further using cross-fitted marginal and grouped-conditional permutation importance. As shown in Figure 11, marginal permutation of reaction time produced a mean decrease in OOF R^2 of 1.593 ± 0.090 , compared with 0.177 ± 0.018 for nominal Cu-loading. The corresponding increases in OOF MAE were 0.1548 ± 0.0072 and 0.0453 ± 0.0046 , respectively.

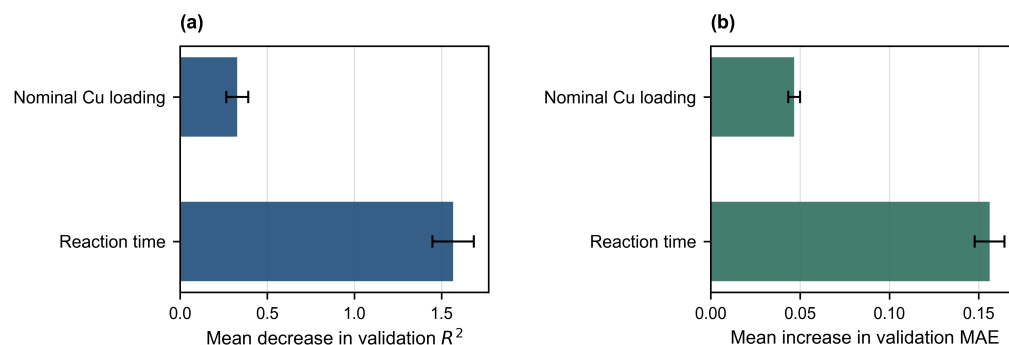


Figure 10. Cross-validated marginal permutation importance for the prediction of C_t/C_0 , quantified as (a) the mean decrease in validation R^2 and (b) the mean increase in validation MAE.

Under grouped-conditional permutation, disruption of reaction time within fixed Cu-loading strata produced an OOF R^2 decrease of 1.248 ± 0.139 and an OOF MAE increase of 0.1208 ± 0.0097 . Permutation of nominal Cu-loading within fixed reaction-time strata produced corresponding values of 0.061 ± 0.016 and 0.0167 ± 0.0041 . Thus, reaction time remained the more influential predictor under both permutation schemes.

The similarity between the marginal and grouped-conditional estimates is consistent with the balanced factorial structure of the dataset, in which each of the five nominal Cu-loading levels was evaluated at each of the 13 reaction times. A decrease in R^2 greater than unity is mathematically possible because the permuted OOF predictions can yield negative R^2 values. These importance values describe the sensitivity of RF predictions to disruption of the input variables and do not establish causal or physicochemical effects.

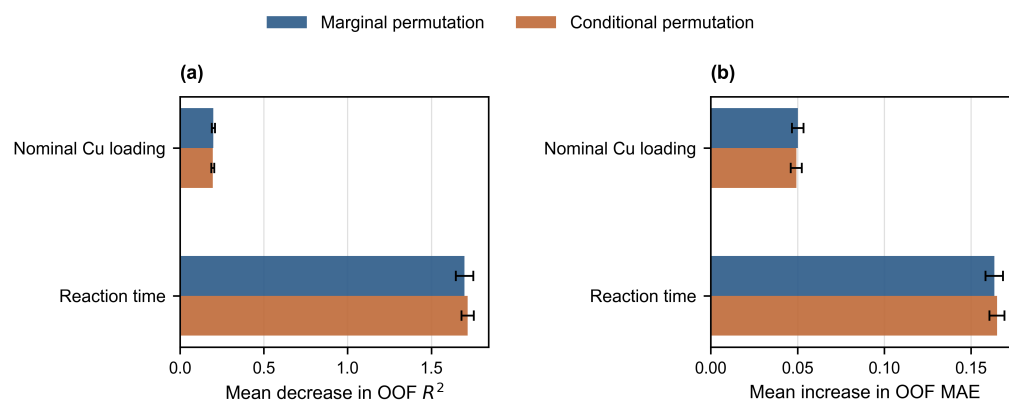


Figure 11. Cross-fitted marginal and grouped-conditional permutation importance for predicting C_t/C_0 , quantified as (a) the mean decrease in OOF R^2 and (b) the mean increase in OOF MAE.

Under grouped-conditional permutation, nominal Cu-loading was shuffled within fixed reaction-time strata, whereas reaction time was shuffled within fixed Cu-loading strata. Bars represent the mean across 10 controlled RF seeds, and error bars indicate ± 1 standard deviation across seeds. Fifty permutations were performed for each predictor, permutation scheme, and seed. The error bars represent model-seed variability rather than experimental uncertainty.

The stability of the feature-importance estimates under synchronized time-block re-sampling is presented in Figure 12.

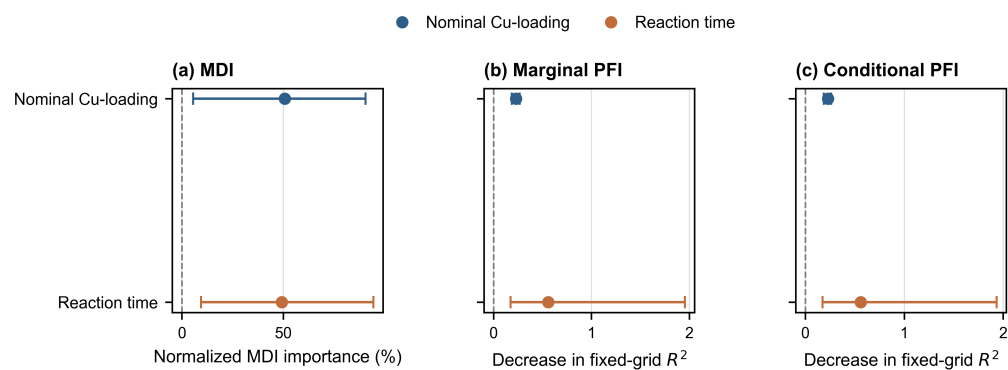


Figure 12. Model-based block-bootstrap stability of the RF feature-importance estimates for C_t/C_0 .

The blue and orange symbols represent nominal Cu-loading and reaction time, respectively. The points show the bootstrap medians, and the horizontal bars show the corresponding 2.5th–97.5th percentile ranges. The dashed vertical line denotes zero importance. Panel (a) presents normalized MDI importance, whereas panels (b) and (c) present the decrease in fixed-grid R^2 for marginal and grouped-conditional PFI, respectively.

The MDI estimates exhibited substantial resampling variability, with median importance values of 49.22% for nominal Cu-loading and 50.78% for reaction time and wide, strongly overlapping percentile ranges. Marginal PFI produced median R^2 decreases of 0.228 for nominal Cu-loading and 0.564 for reaction time. The corresponding grouped-conditional estimates were 0.229 and 0.563. Thus, reaction time had the greater median PFI, but the overlapping percentile ranges indicate that the importance ranking was not uniformly stable under time-block resampling.

These percentile ranges describe model and resampling stability rather than experimental uncertainty because independently synthesized catalyst batches and repeated measurements were not available.

To complement the SHAP, MDI, and permutation-importance results, one-way partial dependence was examined for the primary response, C_t/C_0 . The resulting partial-dependence estimates for nominal Cu-loading and reaction time are presented in Figure 13.

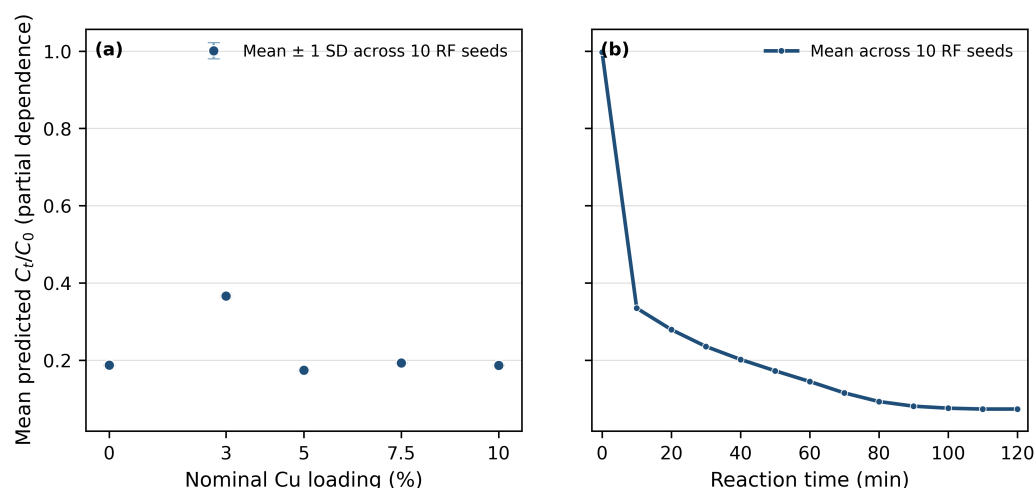


Figure 13. One-way Random Forest partial-dependence estimates for C_t/C_0 as functions of (a) nominal Cu-loading, evaluated only at the experimentally investigated loading levels, and (b) reaction time.

Symbols and lines in Figure 13 represent the mean partial-dependence estimates across 10 controlled RF seeds, whereas the error bars and shaded band denote ± 1 stan-

dard deviation across seeds. These intervals quantify model-seed variability rather than experimental uncertainty.

As shown in Figure 13a, the dependence of the predicted C_t/C_0 on nominal Cu-loading is non-monotonic within the investigated compositional range. The nominal 3% Cu-loading composition exhibits the highest mean partial-dependence value, whereas the nominal 5% composition exhibits the lowest mean model-predicted C_t/C_0 among the five investigated levels. The corresponding estimates for the 0%, 7.5%, and 10% compositions are comparatively similar. The limited variation across RF seeds suggests that the fitted estimates are relatively insensitive to RF initialization, but it does not establish statistically significant differences between the catalyst compositions.

The Cu-loading panel contains only the experimentally investigated nominal levels of 0%, 3%, 5%, 7.5%, and 10%. Because the immediate neighborhood of 5%, particularly 4% and 6%, was not experimentally sampled, the observed model-based minimum cannot be interpreted as evidence that 5% is a local or global optimum. It identifies 5% only as the investigated level with the lowest mean predicted C_t/C_0 .

Figure 13b shows a pronounced decrease in the mean predicted C_t/C_0 during the initial reaction period, followed by a progressively smaller decrease at longer reaction times. This pattern is consistent with the experimentally observed time-dependent removal profiles presented in Figure 2. Nevertheless, the partial-dependence curve represents the temporal response learned by the RF model and should not be interpreted as evidence of a specific reaction mechanism or kinetic law.

To complement the partial-dependence analysis, discrete first-order accumulated local effects (ALE) were examined for the primary response, C_t/C_0 . The ALE estimates for nominal Cu-loading and reaction time are presented in Figure 14. The nominal Cu-loading analysis was restricted to the five experimentally investigated levels and did not involve interpolation at unobserved intermediate Cu-loading values.

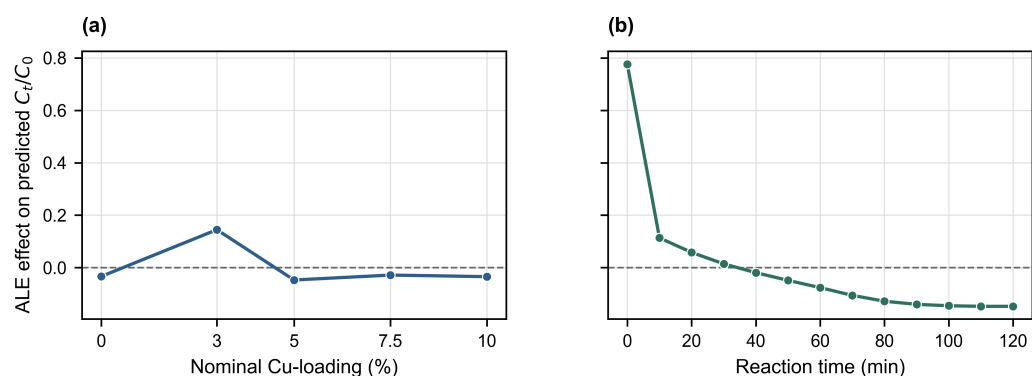


Figure 14. Accumulated local effects (ALE) of (a) nominal Cu-loading and (b) reaction time on the RF-predicted C_t/C_0 . Markers indicate the experimentally represented feature levels. The horizontal dashed line denotes the zero-effect reference; positive and negative ALE values indicate model-predicted effects above and below the centered average prediction, respectively.

In Figure 14, the lines and symbols represent the mean ALE values across 10 controlled RF seeds, and the error bars and shaded bands indicate ± 1 standard deviation across seeds. Because the ALE values were centered over the empirical predictor distribution, positive and negative values indicate increases and decreases, respectively, in the predicted C_t/C_0 relative to the centered mean. The displayed variability reflects model-seed variability rather than experimental or bootstrap uncertainty.

As shown in Figure 14a, the mean ALE effects at nominal Cu-loading levels of 0%, 3%, 5%, 7.5%, and 10% were -0.0339 ± 0.0011 , 0.1445 ± 0.0023 , -0.0474 ± 0.0014 , -0.0286 ± 0.0009 , and -0.0347 ± 0.0015 , respectively. The nominal 5% Cu-loading level

therefore exhibited the most negative mean ALE effect among the investigated levels, whereas the nominal 3% level exhibited the most positive effect. However, these model-based results do not establish 5% as a local or global optimum because intermediate Cu-loading levels, particularly 4% and 6%, were not experimentally investigated.

The reaction-time ALE curve in Figure 14b decreased sharply between 0 and 10 min and then declined more gradually, changing from 0.7759 ± 0.0036 at 0 min to approximately -0.1478 ± 0.0012 at 120 min. This association is consistent with the time-dependent removal profiles in Figure 2, but it does not demonstrate a specific kinetic law or photocatalytic mechanism.

Because each nominal Cu-loading level was evaluated at every reaction time, the discrete ALE estimates are closely related to the centered partial-dependence estimates and are interpreted as a complementary local-effect representation rather than as independent validation of the PDP findings. The joint model-predicted response to reaction time and nominal Cu-loading is presented in Figure 15.

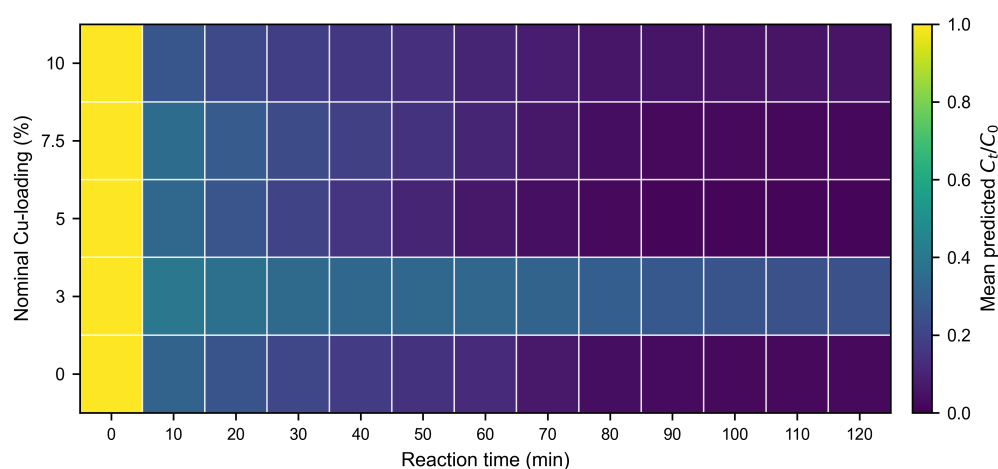


Figure 15. Discrete two-dimensional RF response map of the mean predicted C_t/C_0 across reaction time and nominal Cu-loading.

Each cell represents one of the 65 experimentally investigated Cu-loading–reaction-time combinations. The color scale ranges from 0 to 1. Darker purple colors indicate lower predicted C_t/C_0 and therefore greater predicted removal, whereas lighter green and yellow colors indicate higher residual concentration ratios and lower predicted removal.

The dominant color transition occurs along the reaction-time axis, reflecting the decrease in predicted C_t/C_0 with increasing reaction time. The nominal 3% Cu-loading level retains comparatively high predicted residual concentrations over most nonzero reaction times. The nominal 10% Cu-loading level exhibits the lowest mean predicted C_t/C_0 at 10–30 min, whereas the nominal 5% level exhibits the lowest values from approximately 40 to 120 min. Thus, the fitted compositional ranking varies with reaction time.

The map is restricted to the five investigated Cu-loading levels and does not establish 5% as a local or global optimum. It represents fitted RF predictions rather than experimental uncertainty, statistical significance, or evidence of a physicochemical mechanism.

4.7. Linear and Monotonic Correlation Analysis

Pearson product–moment and Spearman rank correlation matrices were calculated to characterize the pairwise linear and monotonic associations among the nominal Cu-loading, reaction time, degradation efficiency, C_t/C_0 , and $\ln(C_0/C_t)$. The resulting matrices are presented in Figure 16.

Apparent removal efficiency and C_t/C_0 exhibited a correlation of -1.00 in both matrices. This result follows directly from their deterministic affine relationship, $\eta = 100(1 -$

C_t/C_0), and should therefore be interpreted as an algebraic consequence of the response definitions rather than as an independent statistical validation of mass conservation.

The Spearman correlations of apparent removal efficiency and C_t/C_0 with $\ln(C_0/C_t)$ were 1.00 and -1.00 , respectively. These perfect rank correlations are expected because all three responses are monotonic transformations of the same residual concentration measurement. Their corresponding Pearson correlations were 0.81 and -0.81 , reflecting the nonlinear logarithmic transformation between the response scales. These differences do not independently establish a particular kinetic mechanism or an asymptotic decay law.

Reaction time showed a positive association with degradation efficiency, with Pearson and Spearman coefficients of 0.68 and 0.76, respectively. The corresponding associations with C_t/C_0 were -0.68 and -0.76 . For $\ln(C_0/C_t)$, the Pearson and Spearman coefficients were 0.77 and 0.76, respectively. These results indicate a strong temporal ordering of the degradation responses. The difference between the Pearson and Spearman coefficients for apparent removal efficiency further suggests that its relationship with reaction time is monotonic but not strictly linear.

The nominal Cu-loading exhibited weak marginal correlations with apparent removal efficiency (0.08 for Pearson and 0.13 for Spearman), C_t/C_0 (-0.08 and -0.13), and $\ln(C_0/C_t)$ (0.04 and 0.13). These values indicate a weak overall bivariate association when observations from all reaction times are pooled. However, weak marginal correlations do not exclude nonlinear, conditional, or time-dependent associations within the fitted RF model. Accordingly, the correlation analysis provides a descriptive complement to the SHAP and MDI results presented in Section 4.6, rather than independent evidence of causal or mechanistic effects.

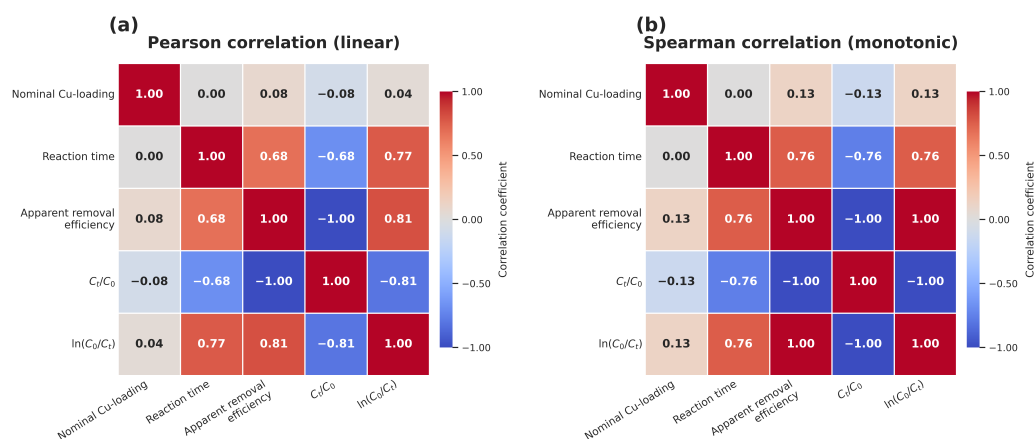


Figure 16. Correlation matrices for nominal Cu-loading, reaction time, and the three response representations: (a) Pearson linear correlation coefficients and (b) Spearman rank correlation coefficients. Red and blue cells indicate positive and negative correlations, respectively, with color intensity representing the magnitude of the correlation coefficient.

4.8. Experimental Limitations and Future Work

The available experimental records did not permit retrospective reporting or independent evaluation of the light-source spectrum and irradiance at the reactor, source-to-reactor distance, detailed reactor geometry, stirring rate, dissolved-oxygen conditions, cumulative sampling-volume correction, or the absolute concentration decrease during the 30-min dark-conditioning period. Independent dark-only and catalyst-free photolysis controls were also unavailable. Consequently, the post- $t = 0$ concentration decrease is reported as apparent removal under irradiation and cannot be quantitatively separated into contributions from continued adsorption, direct photolysis, and photocatalytic transformation.

These limitations restrict the interpretation of the findings to the investigated experimental conditions.

Future studies should include independently synthesized replicate catalyst batches, repeated removal experiments, and denser compositional sampling around the nominal 5% Cu-loading level, particularly at 4% and 6%. The irradiation and reactor conditions should be fully documented, and quantitative dark-adsorption and catalyst-free photolysis controls should be included. Structural, optical, and surface characterization should also be performed for each composition. Integration of these experiments with external and group-based model validation would enable a more rigorous assessment of the reproducibility, physicochemical basis, and generalizability of the observed composition-dependent behavior.

5. Conclusions

This study developed a chronologically evaluated and explainable ML framework for examining the associations of reaction time and nominal Cu-loading with the time-dependent removal performance of AC@NiO samples under irradiation. Following comparative model screening using C_t/C_0 as the primary response, RF was selected for the subsequent analyses. The final regularized RF model showed limited sensitivity to the examined random initializations, achieving mean LOOCV R^2 values of 0.9691 ± 0.0013 for C_t/C_0 and the deterministically derived $\eta/100$, and 0.9152 ± 0.0090 for $\ln(C_0/C_t)$. The identical results for C_t/C_0 and $\eta/100$ arise from their deterministic affine relationship and do not constitute independent predictive evidence.

The fixed chronological evaluation yielded mean R^2 values of 0.9668 ± 0.0073 for C_t/C_0 and the corresponding apparent removal efficiency, and 0.8890 ± 0.0163 for $\ln(C_0/C_t)$ within the held-out $100 < t \leq 120$ min interval. However, the composition-specific RF predictions were constant across the 100–120 min boundary interval, and comparison with a carry-forward reference showed that a high R^2 alone does not establish learned temporal evolution. These results are therefore interpreted as later-time boundary diagnostics within the present dataset and validation design, rather than as evidence of external generalizability or temporal extrapolation.

SHAP, MDI, and permutation-based analyses assigned greater point-estimate importance to reaction time. However, the moving-block bootstrap analysis showed that this importance ranking was sensitive to temporal resampling and was not uniformly stable across bootstrap repetitions. The feature-attribution results therefore represent predictive associations rather than causal effects or evidence of a physicochemical mechanism.

Under the investigated conditions of an initial MB concentration of 20 mg L^{-1} , a catalyst dosage of 1.0 g L^{-1} , pH 7, 25°C , and irradiation times of 0–120 min, the nominal 5% Cu-containing sample exhibited the most favorable apparent removal profile among the five tested compositions. This observation neither demonstrates a particular structural configuration of Cu nor establishes 5% as a local, global, or universal optimum.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/catal16090810/s1>: Table S1: Descriptive ordinary least-squares fits of the linearized apparent pseudo-first-order representation.

Author Contributions: Conceptualization, N.B., O.B., M.Z. and S.H.; methodology, M.Z., S.K. and O.B.; software, M.Z. and S.K.; validation, M.Z. and S.K.; formal analysis, M.Z., S.K. and S.H.; investigation, M.Z. and S.H.; data curation, N.B., M.Z. and S.K.; writing—original draft preparation, M.Z., S.K. and S.H.; writing—review and editing, N.B., M.Z., S.K., O.B., C.O. and S.H.; visualization, M.Z. and S.K.; supervision, M.Z.; project administration, M.Z. and S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Siirt University Scientific Research Projects Coordination Unit, grant number 2019-SIÜFEB-023.

Data Availability Statement: The Python 3.13 source code used in this study is publicly available at <https://github.com/metinzontul/ML-Photocatalysis-Kinetics> (accessed on 31 August 2026). The experimental dataset is available from the corresponding author upon reasonable request.

Acknowledgments: This study was conducted within the scope of the Siirt University Research Foundation Project (Project No. 2019-SIÜFEB-023). The authors thank Siirt University for providing research facilities and support during this study. The authors used Gemini Pro 3.1 as a supportive tool for refining the manuscript's language, enhancing its structural flow, and ensuring clarity throughout the writing process. Additionally, Gemini Nano Banana 2 was employed to assist in the generation of Figure 1. Following the use of these AI tools, the authors thoroughly reviewed, validated, and finalized the content to ensure scientific accuracy. The authors take full responsibility for the information and arguments presented in this work.

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

Abbreviations

The following abbreviations are used in this manuscript:

AC	Activated carbon
AOP	Advanced oxidation process
CV	Cross-validation
KNN	K-nearest neighbors
LOOCV	Leave-one-out cross-validation
MAE	Mean absolute error
MB	Methylene blue
MDI	Mean decrease in impurity
ML	Machine learning
MLP	Multi-layer perceptron
MSE	Mean squared error
OOT	Out-of-time
RBF	Radial basis function
RF	Random forest
ROS	Reactive oxygen species
SHAP	SHapley Additive exPlanations
SVR	Support vector regression
UV-Vis	Ultraviolet-visible
XAI	Explainable artificial intelligence

References

- Orak, C.; Ersöz, G. Heterogeneous photo-Fenton-like degradation of oxytetracycline-containing wastewater. *J. Adv. Res. Nat. Appl. Sci.* **2024**, *10*, 182–189. [\[CrossRef\]](#)
- Gupta, A.D.; Singh, H.; Varjani, S.; Awasthi, M.K.; Giri, B.S.; Pandey, A. A critical review on biochar-based catalysts for the abatement of toxic pollutants from water via advanced oxidation processes (AOPs). *Sci. Total Environ.* **2022**, *849*, 157831. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xiong, J.; Zeng, H.-Y.; Chen, C.-R.; Li, S.; Xu, S.; An, D.-S. Ag₃VO₄-anchored ZnTi hydrotalcite microspheres with a rosette-like structure for tetracycline degradation. *Appl. Clay Sci.* **2021**, *208*, 106118. [\[CrossRef\]](#)
- Gonçalves, R.A.; Toledo, R.P.; Joshi, N.; Berengue, O.M. Green synthesis and applications of ZnO and TiO₂ nanostructures. *Molecules* **2021**, *26*, 2236. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ur Rahman, Z.; Shah, U.; Alam, A.; Shah, Z.; Shaheen, K.; Khan, S.B.; Khan, S.A. Photocatalytic degradation of cefixime using a CuO–NiO nanocomposite photocatalyst. *Inorg. Chem. Commun.* **2023**, *148*, 110312. [\[CrossRef\]](#)
- Alharbi, F.K.; Ben Aissa, M.A.; Ismail, M.; Albadri, A.E.A.E.; Modwi, A.; Al-Ayed, A.S.; Alharbi, A.A. Ag–NiO-functionalised mesoporous g-C₃N₄ nanosheets for the eradication of lead ions from wastewater. *Int. J. Environ. Anal. Chem.* **2024**, *105*, 6849–6865. [\[CrossRef\]](#)

7. Alharbi, F.K.; Ben Aissa, M.A.; Modwi, A.; Alharbi, A.A.; Alfadhl, F.; Saleh, S.M. Impact of Ag–NiO nanoparticles loaded on g-C₃N₄ nanosheets for cleaning contaminated water. *Inorg. Chem. Commun.* **2024**, *166*, 112608. [\[CrossRef\]](#)
8. Roguai, S. Tunable structural, optical, and photocatalytic properties of sulfur-doped NiO thin films: Influence of doping concentration and scavenger-assisted mechanism analysis. *J. Sol-Gel Sci. Technol.* **2026**, *117*, 46. [\[CrossRef\]](#)
9. Jain, A.; Balasubramanian, R.; Srinivasan, M.P. Hydrothermal conversion of biomass waste to activated carbon with high porosity: A review. *Chem. Eng. J.* **2016**, *283*, 789–805. [\[CrossRef\]](#)
10. Algarni, T.S.; Al-Mohaimeed, A.M.; Al-Odayni, A.B.; Abduh, N.A.Y. Activated carbon/ZnFe₂O₄ nanocomposite adsorbent for the efficient removal of crystal violet cationic dye from aqueous solutions. *Nanomaterials* **2022**, *12*, 3224. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Bulut, N.; Baytar, O.; Şahin, Ö.; Horoz, S. Synthesis of Co-doped NiO/AC photocatalysts and their use in photocatalytic degradation. *J. Aust. Ceram. Soc.* **2021**, *57*, 419–425. [\[CrossRef\]](#)
12. Orak, C.; Yüksel, A. Box–Behnken design for hydrogen evolution from sugar-industry wastewater using solar-driven hybrid catalysts. *ACS Omega* **2022**, *7*, 42489–42498. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Shariati, M.; Babaei, A.; Azizi, A. Synthesis of the tetranary magnetic nanocomposite as a good photocatalyst for the degradation of basic fuchsin dye in aqueous media under visible light: Characterization, response surface methodology, and kinetic study. *J. Mater. Res.* **2023**, *38*, 2666–2678. [\[CrossRef\]](#)
14. Ghattas, B.; Manzoni, D. Machine learning alternatives to response surface models. *Mathematics* **2023**, *11*, 3406. [\[CrossRef\]](#)
15. Navidpour, A.H.; Hosseinzadeh, A.; Huang, Z.; Li, D.; Zhou, J.L. Application of machine learning algorithms in predicting the photocatalytic degradation of perfluorooctanoic acid. *Catal. Rev. Sci. Eng.* **2022**, *66*, 687–712. [\[CrossRef\]](#)
16. Haq, Z.U.; Ullah, H.; Khan, M.N.A.; Naqvi, S.R.; Ahsan, M. Hydrogen production optimization from a sewage-sludge supercritical gasification process using machine learning methods integrated with a genetic algorithm. *Chem. Eng. Res. Des.* **2022**, *184*, 614–626. [\[CrossRef\]](#)
17. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
18. Bakır, R.; Orak, C.; Yüksel, A. Optimizing hydrogen evolution prediction: A unified approach using random forests, LightGBM, and a bagging regressor ensemble model. *Int. J. Hydrogen Energy* **2024**, *67*, 101–110. [\[CrossRef\]](#)
19. Tasneem, S.; Ageeli, A.A.; Alamier, W.M.; Hasan, N.; Safaei, M.R. Organic catalysts for hydrogen production from noodle wastewater: Machine-learning- and deep-learning-based analysis. *Int. J. Hydrogen Energy* **2024**, *52*, 599–616. [\[CrossRef\]](#)
20. Ren, T.; Wang, L.; Chang, C.; Li, X. Machine-learning-assisted multiphysics coupling performance optimization in a photocatalytic hydrogen-production system. *Energy Convers. Manag.* **2020**, *216*, 112935. [\[CrossRef\]](#)
21. Jyothirmai, M.V.; Dantuluri, R.; Sinha, P.; Abraham, B.M.; Singh, J.K. Machine-learning-driven high-throughput screening of transition-metal-atom- intercalated g-C₃N₄/MX₂ (M = Mo, W; X = S, Se, Te) heterostructures for the hydrogen evolution reaction. *ACS Appl. Mater. Interfaces* **2024**, *16*, 12437–12445. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Yang, W.; Li, H. Machine learning in wastewater treatment: A comprehensive bibliometric review. *ACS ES&T Water* **2025**, *5*, 511–524. [\[CrossRef\]](#)
23. Saadetejad, D.; Oral, B.; Can, E.; Yıldırım, R. Machine learning analysis of gas-phase photocatalytic CO₂ reduction for hydrogen production. *Int. J. Hydrogen Energy* **2022**, *47*, 19655–19668. [\[CrossRef\]](#)
24. Wayo, D.D.K.; Goliatt, L.; Ganji, M.D. DFT and hybrid classical–quantum machine learning integration for photocatalyst discovery and hydrogen production. *Rev. Chem. Eng.* **2025**, *41*, 741–774. [\[CrossRef\]](#)
25. Kalasampath, K.; Spoorthi, K.N.; Sajeev, S.; Kuppa, S.S.; Ajay, K.; Maruthamuthu, A. A literature review on applications of explainable artificial intelligence (XAI). *IEEE Access* **2025**, *13*, 41111–41140. [\[CrossRef\]](#)
26. Vassilev-Galindo, V.; Llorca, J. Explainable artificial intelligence for materials discovery: Application to catalysts for the HER and ORR. *Chem. Sci.* **2026**, *17*, 1058–1072. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Salih, A.M.; Raisi-Estabragh, Z.; Galazzo, I.B.; Radeva, P.; Petersen, S.E.; Lekadir, K.; Menegaz, G. A perspective on explainable artificial intelligence methods: SHAP and LIME. *Adv. Intell. Syst.* **2025**, *7*, 2400304. [\[CrossRef\]](#)
28. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*; Leanpub: Victoria, BC, Canada, 2020.
29. Armaković, S.J.; Armaković, S.; Bilić, A.; Savanović, M.M. ZnO-based photocatalysts: Synergistic effects of material modifications and machine learning optimization. *Catalysts* **2025**, *15*, 793. [\[CrossRef\]](#)
30. Hameeda, B.; Mushtaq, A.; Saeed, M.; Munir, A.; Jabeen, U.; Waseem, A. Development of Cu-doped NiO nanoscale material as an efficient photocatalyst for visible-light dye degradation. *Toxin Rev.* **2021**, *40*, 1396–1406. [\[CrossRef\]](#)
31. Hashem, B.Y.; Alswat, A.A.; Ali, S.L.; Al-Odaini, N.A.; Alshorifi, F.T. Facile synthesis of NiO–CuO/activated-carbon nanocomposites for use in the removal of lead and cadmium ions from water. *ACS Omega* **2022**, *7*, 47183–47191. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Kang, J.-K.; Lee, Y.-J.; Son, C.-Y.; Park, S.-J.; Lee, C.-G. Alternative assessment of machine learning to polynomial regression in response surface methodology for predicting decolorization efficiency in textile wastewater treatment. *Chemosphere* **2025**, *370*, 143996. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Anandhi, G.; Iyapparaja, M. Photocatalytic degradation of drugs and dyes using a machine learning approach. *RSC Adv.* **2024**, *14*, 9003–9019. [\[CrossRef\]](#) [\[PubMed\]](#)

34. Schossler, R.T.; Ojo, S.; Jiang, Z.; Hu, J.; Yu, X. A novel interpretable machine learning model approach for the prediction of TiO₂ photocatalytic degradation of air contaminants. *Sci. Rep.* **2024**, *14*, 13070. [[CrossRef](#)] [[PubMed](#)]
35. Baytar, O.; Zontul, M.; Orak, C.; Karateke, S.; Aydın, H.; Horoz, S. Interpretable machine learning approach for photocatalytic degradation in Mn-doped semiconductors using multilayer perceptron and SHAP analysis. *Catalysts* **2026**, *16*, 530. [[CrossRef](#)]
36. Kaufman, S.; Rosset, S.; Perlich, C.; Stitelman, O. Leakage in data mining: Formulation, detection, and avoidance. *ACM Trans. Knowl. Discov. Data* **2012**, *6*, 1–21.
37. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; pp. 4768–4777.
38. Stone, C.J. Consistent nonparametric regression. *Ann. Stat.* **1977**, *5*, 595–620. [[CrossRef](#)]
39. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [[CrossRef](#)]
40. Drucker, H.; Burges, C.J.C.; Kaufman, L.; Smola, A.J.; Vapnik, V. Support vector regression machines. In *Advances in Neural Information Processing Systems 9*; Mozer, M.C., Jordan, M.I., Petsche, T., Eds.; MIT Press: Cambridge, MA, USA, 1997; pp. 155–161.
41. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
42. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
43. Stone, M. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B Methodol.* **1974**, *36*, 111–133. [[CrossRef](#)]
44. Bouthillier, X.; Delaunay, P.; Bronzi, M.; Trofimov, A.; Nichyporuk, B.; Szeto, J.; Mohammadi Sepahvand, N.; Raff, E.; Madan, K.; Voleti, V.; et al. Accounting for variance in machine learning benchmarks. *Proc. Mach. Learn. Syst.* **2021**, *3*, 747–769.
45. Bengio, Y.; Grandvalet, Y. No unbiased estimator of the variance of *k*-fold cross-validation. *J. Mach. Learn. Res.* **2004**, *5*, 1089–1105.
46. Varoquaux, G. Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage* **2018**, *180*, 68–77. [[CrossRef](#)] [[PubMed](#)]
47. Louppe, G.; Wehenkel, L.; Sutura, A.; Geurts, P. Understanding variable importances in forests of randomized trees. In *Advances in Neural Information Processing Systems 26*; Burges, C.J.C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2013; pp. 431–439.
48. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [[CrossRef](#)] [[PubMed](#)]
49. Kvålseth, T.O. Cautionary note about R^2 . *Am. Stat.* **1985**, *39*, 279–285. [[CrossRef](#)]
50. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [[CrossRef](#)]
51. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [[CrossRef](#)]
52. Tran, H.D.; Nguyen, D.Q.; Do, P.T.; Tran, U.N.P. Kinetics of photocatalytic degradation of organic compounds: A mini-review and new approach. *RSC Adv.* **2023**, *13*, 16915–16925. [[CrossRef](#)] [[PubMed](#)]
53. Liu, Y.; Luo, Y.; Ye, J.; Zhao, L.; Chen, Y.; Dai, Y.; Xie, Y.; Wang, Y.; Zhang, Y. Photocatalytic hydrogen production performance of anatase and brookite TiO₂ based on oxygen vacancies and noble metal Pt modification. *J. Photochem. Photobiol. A Chem.* **2026**, *479*, 117242. [[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.