

Article

Machine Learning-Enhanced Modeling of Heavy Metal Adsorption onto Coal Fly Ash-Derived Zeolite P

Benito A. Hernández-Guerrero ¹, Lorena Martínez ¹, Gabriel Peña-Rodríguez ²  and Fernando Trejo ^{1,*} 

¹ Instituto Politécnico Nacional, CICATA-Legaria, Legaria 694, Col. Irrigación, Ciudad de México 11500, Mexico; bhernandezg1407@alumno.ipn.mx (B.A.H.-G.); lore_3612@hotmail.com (L.M.)

² Grupo de Investigación en Instrumentación y Física de la Materia Condensada, Facultad de Ciencias Básicas, Universidad Francisco de Paula Santander, San José de Cúcuta A.A. 1055, Colombia; gabrielpr@ufps.edu.co

* Correspondence: ftrejoz@ipn.mx

Abstract

Zeolite P was synthesized by hydrothermal treatment of coal fly ash and applied to the individual removal of six heavy metals (Pb^{2+} , Ni^{2+} , Cu^{2+} , Cr^{3+} , Hg^{2+} , Cd^{2+}) from aqueous solutions. Characterization by SEM-EDS, FTIR, BET, XRD, zeta potential, and XPS revealed a BET surface area of $30 \text{ m}^2/\text{g}$, Si/Al ratio of 1.63, and pH_{pzc} of 3.2. Batch experiments at the natural solution pH of 3.9 in all cases ($C_0 = 10, 100, 200 \text{ mg/L}$; $t = 1\text{--}60 \text{ min}$) yielded an apparent selectivity sequence at $C_0 = 200 \text{ mg/L}$ of Hg^{2+} (10.47 mg/g) $>$ Pb^{2+} (9.12) $>$ Ni^{2+} (2.18) $>$ Cr^{3+} (2.05) $>$ Cu^{2+} (1.82) $>$ Cd^{2+} (1.26), where Hg^{2+} and Pb^{2+} reached near-equilibrium while the remaining metals were still approaching it at $t = 60 \text{ min}$. Weber–Morris and Boyd analyses confirmed three sequential diffusion stages with a concentration-dependent shift from film to intraparticle diffusion control through the narrow GIS channels ($3.1 \times 4.5 \text{ \AA}$). Ion exchange was identified as the dominant mechanism based on convergent kinetic, diffusion, XPS, and selectivity–electronegativity evidence ($r = +0.76$). A leakage-free machine learning framework combining physicochemical descriptors with experimental variables was tested under three cross-validation strategies of increasing stringency. Gradient Boosting achieved $R^2 = 0.979 \pm 0.043$ (repeated K-Fold) and $R^2 = 0.880$ on six completely held-out kinetic curves. An ablation study confirmed that physicochemical descriptors are essential (experimental-only models yielded negative R^2). SHAP feature importance rankings were consistent with established ion exchange selectivity theory. This work demonstrates that group-level validation, physics-informed descriptors, and systematic ablation testing are able to identify both the capabilities and the boundaries of small-dataset ML when testing for metal kinetics prediction.

Keywords: coal fly ash; zeolite P; heavy metal adsorption; machine learning; SHAP analysis; ablation test



Academic Editor: Laura Bulgariu

Received: 5 March 2026

Revised: 27 March 2026

Accepted: 31 March 2026

Published: 2 April 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

Coal fly ash (CFA) is one of the most abundant industrial solid wastes, with global production exceeding 750 million tons annually [1]. Improper disposal poses environmental and health risks, including groundwater contamination and airborne dust [2], yet CFA possesses high silica and alumina content, pozzolanic activity, and fine particle size that make it a valuable precursor for synthesizing zeolites and geopolymers [3]. CFA is classified as type C ($\text{SiO}_2 + \text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3 < 50 \text{ wt\%}$, from lignite or young bituminous coal) or type F ($>70 \text{ wt\%}$, from anthracite or older bituminous coal) [4]. Zeolites are three-dimensional

aluminosilicate frameworks bearing a net negative charge from $\text{Si}^{4+}/\text{Al}^{3+}$ substitution, compensated by exchangeable cations (Na^+ , K^+ , Ca^{2+} , Mg^{2+}), and typically offer a wide range of pore structures at lower cost than natural zeolites [5].

Among CFA-derived zeolites, zeolite P has attracted attention for heavy metal remediation due to its favorable pore geometry, ion exchange selectivity, and relatively simple hydrothermal synthesis conditions [6]. Synthesis is mainly governed by reaction time, NaOH concentration, and temperature, with higher Si/Al ratios reducing zeolite yield [7]. Several activation routes have been reported, including direct hydrothermal treatment [8,9], alkali fusion followed by hydrothermal processing [10,11], and molten salt methods [12]. The resulting materials have demonstrated effective removal of diverse contaminants: BTX vapors [13,14], dyes [15], tungsten [16], radioactive elements [17–19], and heavy metal cations, including Pb^{2+} , Cu^{2+} , Zn^{2+} , Ni^{2+} , Co^{2+} , Cd^{2+} , and Cr^{3+} [20–25].

Heavy metal contamination in water remains a critical environmental concern, with industrial discharges releasing persistent and bioaccumulative toxic metals, i.e., Pb, Ni, Cu, Cr, Hg, and Cd, into aquatic ecosystems [26,27]. Adsorption is a preferred remediation technique owing to its cost-effectiveness and operational simplicity [28], and CFA-derived zeolites offer the dual advantage of waste valorization and effective metal removal.

Adsorption kinetics are conventionally described by mechanistic models such as pseudo-first-order (PFO), pseudo-second-order (PSO), and Elovich equations [29]. While these provide interpretable rate parameters, they are fitted individually per metal–concentration curve and cannot capture nonlinear interactions when using multi-metal systems [30].

Machine learning approaches have demonstrated enhanced predictive capabilities for adsorption datasets [31], including Random Forest (RF), Gradient Boosting (GB), XG-Boost [32], SVR, KNN, CNN, MLP [33], ANN, and ANFIS [34]. Recent studies have applied ML to predict heavy metal adsorption on iron (oxyhydr)oxides using SHAP-based feature interpretation [35], to forecast adsorption efficiency on biochar using ensemble models with 353 literature-derived samples [36], and to develop interpretable Gradient Boosting frameworks predicting adsorption capacities of six heavy metals across 89 MOF composites [37]. However, these studies generally rely on random train/test splits and/or use one-hot encoding rather than physics-informed descriptors.

Most ML adsorption studies suffer from four methodological gaps: (1) data leakage from random splits that place correlated time points in both train and test sets; (2) single-strategy validation that cannot distinguish interpolation from extrapolation; (3) categorical metal encoding that discards physicochemical information needed for cross-metal generalization; and (4) feature importance reported without ablation, leaving rankings unvalidated against actual performance loss.

This work introduces a leakage-free ML framework for predicting adsorption kinetics of six heavy metals individually evaluated (Pb^{2+} , Ni^{2+} , Cu^{2+} , Cr^{3+} , Hg^{2+} , Cd^{2+}) on zeolite P synthesized from coal fly ash. The specific contributions are: (1) a group-level external validation protocol in which entire kinetic curves (Metal $\times$ C_0 combinations) are held out, ensuring that every time point of a test curve is genuinely unseen during model development; (2) three hierarchically strict cross-validation strategies—Repeated K-Fold (interpolation), Group K-Fold (extrapolation to unseen curves), and Leave-One-Metal-Out (extrapolation to unseen metals)—that progressively quantify the limits of the model; (3) a physics-informed feature space encoding metal identity through Shannon ionic radii [38], Pauling electronegativities [39], Marcus hydration enthalpies [40], and Pearson softness parameters [41], rather than relying on one-hot encoding; (4) an ablation study demonstrating that models trained on experimental variables alone (C_0 , t) yield negative R^2 , confirming that the physicochemical descriptors are essential to predictive performance;

(5) SHAP-based interpretability analysis [42] linking ML feature importance to established ion exchange selectivity theory for zeolites [43]; and (6) an honest assessment of model boundaries: the strictly negative R^2 under Leave-One-Metal-Out validation demonstrates that predicting the adsorption behavior of a completely novel metal from only six training metals remains beyond the capacity of this dataset, establishing a clear boundary condition for future work.

2. Materials and Methods

2.1. Materials

Coal fly ash (CFA) was obtained from a coal-fired power plant in Mexico. Particles were used as received. Sodium hydroxide pellets ($\geq 99\%$, Sigma-Aldrich, St. Louis, MO, USA) were used for hydrothermal synthesis. The following analytical-grade metal salts ($\geq 99\%$ purity, Sigma-Aldrich) were used to prepare synthetic contaminated water: lead nitrate ($\text{Pb}(\text{NO}_3)_2$), nickel chloride hexahydrate ($\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$), copper sulfate pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$), chromium chloride hexahydrate ($\text{CrCl}_3 \cdot 6\text{H}_2\text{O}$), mercury acetate ($\text{Hg}(\text{CH}_3\text{COO})_2$), and cadmium nitrate tetrahydrate ($\text{Cd}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$). All solutions were prepared using deionized water (resistivity $> 18 \text{ M}\Omega \cdot \text{cm}$). Hydrochloric acid and sodium hydroxide (0.1 M each) were used for pH adjustment during zeta potential measurements.

2.2. Zeolite P Synthesis from Coal Fly Ash

NaOH solutions (0.5 and 1.5 M) were mixed with CFA in a 30 mL Teflon-lined stainless steel autoclave at a liquid-to-solid ratio of 6:1 (v/w , mL/g). Hydrothermal treatment was carried out at two temperatures (100 and 150 °C) for 24 h under autogenous pressure. Four synthesis conditions were evaluated: (1) 0.5 M NaOH, 100 °C; (2) 1.5 M NaOH, 100 °C; (3) 0.5 M NaOH, 150 °C; and (4) 1.5 M NaOH, 150 °C. After reaction, the products were washed repeatedly with deionized water until the filtrate was colorless, dried at 110 °C for 4 h, and stored in a desiccator overnight. The optimal synthesis condition was selected based on maximum heavy metal removal from a screening experiment using a solution (Pb^{2+} , Hg^{2+} , Cu^{2+} , Ni^{2+} , Cd^{2+} , and Cr^{3+} , each at 10 mg/L in 30 mL) with two adsorbent doses (0.15 and 0.60 g). The material synthesized under the optimal condition was used for all subsequent batch adsorption experiments.

2.3. Characterization Techniques

Textural properties were determined by N_2 physisorption at 77 K (Quantachrome Nova Touch 4LX, Quantachrome Instruments, Boynton Beach, FL, USA). Samples were degassed at 250 °C for 3 h prior to analysis. Specific surface area was calculated by the Brunauer–Emmett–Teller (BET) method [44].

Zeta potential was measured as a function of pH (Micromeritics NanoPlus HD, Micromeritics Instrument Corporation, Norcross, GA, USA) over the range pH 2–11 using 0.1 g of zeolite P dispersed in 30 mL of deionized water, with 0.1 M HCl and 0.1 M NaOH as titrants. The pH at the point of zero charge (pH_{pzc}) was determined from the intersection of the zeta potential curve with zero millivolts. Crystalline phases were identified by X-ray diffraction (Bruker D8 Advance, Bruker Corporation, MA, USA, $\text{CuK}\alpha$ radiation, $\lambda = 1.5406 \text{ \AA}$, 40 kV, 40 mA). Phase identification was performed against the crystallographic database (International Centre for Diffraction Data). Le Bail refinement [45] was performed using the FullProf Suite [46].

Morphology and elemental mapping were obtained by scanning electron microscopy with energy-dispersive X-ray spectroscopy (JEOL JSM-6390, JEOL USA, Peabody, MA, USA). Bulk chemical composition was determined by X-ray fluorescence spectroscopy (Bruker S2 Ranger, Pd source, Bruker AXS Inc., Madison, WI, USA).

Surface chemical composition was analyzed by X-ray photoelectron spectroscopy (Thermo Scientific K-Alpha, Thermo Fisher Scientific, Waltham, MA, USA, monochromatic AlK α source, 1486.6 eV) under ultra-high vacuum ($<10^{-9}$ mbar) with a pass energy of 60 eV. The O 1s peak (531 eV) was used as the binding energy reference. Survey spectra were acquired for the CFA precursor, fresh zeolite P, and spent zeolite P (after heavy metal adsorption). Only survey spectra were collected; high-resolution scans with peak fitting for quantitative analysis of metal oxidation states and Na $^{+}$ /M $^{2+}$ exchange stoichiometry were not performed.

2.4. Batch Adsorption Experiments

Single-metal aqueous solutions containing Pb $^{2+}$, Hg $^{2+}$, Cu $^{2+}$, Ni $^{2+}$, Cd $^{2+}$, or Cr $^{3+}$ were prepared at initial concentrations (C_0) of 10, 100, and 200 mg/L from the respective metal salts dissolved in deionized water. Each metal was tested individually (single-metal systems). From an environmental perspective, C_0 = 10 mg/L represents moderately contaminated waters, such as dilute acid mine drainage, urban stormwater runoff, and post-treatment industrial effluents, where heavy metal concentrations typically range from 1 to 50 mg/L [47,48]. C_0 = 100 mg/L represents concentrated industrial discharges, including electroplating rinse water, battery manufacturing effluents, and metal finishing wastewater, where individual metal concentrations commonly reach 50–200 mg/L before treatment [48,49]. C_0 = 200 mg/L represents the upper bound of these concentrated streams and was included specifically to test the adsorbent under high-loading conditions where site saturation, pore congestion, and speciation effects become limiting.

Batch adsorption experiments were conducted in 50 mL beakers containing the optimally synthesized zeolite P at a dose of 2 g/L (0.06 g in 30 mL). The temperature was maintained at 25 ± 1 °C using a thermostatic water bath, and constant agitation was provided at 300 rpm using a stirring plate. All experiments were conducted at the natural pH of the synthetic contaminated water (pH 3.9), which was not adjusted to avoid introducing competing ions or altering metal speciation. While this pH is relevant to acid mine drainage, electroplating effluents, and tannery wastewaters, it is lower than typical surface water or municipal wastewater (pH 6–8). The adsorption capacities reported here therefore represent conservative estimates under electrostatically unfavorable conditions (pH near pH_{pzc} = 3.2), and higher capacities would be expected at neutral to mildly alkaline pH.

Because each metal was dissolved from a different salt, the counter-anion composition varied across metals: NO $_3^{-}$ for Pb $^{2+}$ and Cd $^{2+}$, Cl $^{-}$ for Ni $^{2+}$ and Cr $^{3+}$, SO $_4^{2-}$ for Cu $^{2+}$, and CH $_3$ COO $^{-}$ for Hg $^{2+}$. Since all experiments were conducted in single-metal systems, each solution contained only one counter-anion type at a concentration stoichiometrically determined by the metal salt. These anions are not expected to participate directly in cation exchange on the negatively charged zeolite framework, but their presence at elevated concentrations contributes to the ionic strength of the solution and may influence metal speciation through ion pair formation, particularly for SO $_4^{2-}$ with Cu $^{2+}$ [50].

Aliquots were filtered through 0.22 μ m PTFE syringe filters (polytetrafluoroethylene) prior to ICP-OES analysis. PTFE was selected for its established chemical inertness toward dissolved metal cations because the fluorocarbon backbone contains no ionizable or complexing functional groups. The brief contact time during syringe filtration (~2–5 s) precludes significant adsorptive losses.

The adsorption capacity at time t was calculated using Equation (1):

$$q_t = \frac{(C_0 - C_t) \times V}{m} \quad (1)$$

where C_0 and C_t are the initial and time-dependent metal concentrations (mg/L), V is the solution volume (L), and m is the adsorbent mass (g). Each Metal $\times C_0$ combination yields one kinetic curve comprising 12 q_t values, giving 6 metals $\times$ 3 concentrations $\times$ 12 time points = 216 data points organized into 18 kinetic curves.

2.5. Traditional Kinetic Models

Four kinetic models were fitted to the experimental q_t versus t data by nonlinear regression using the Levenberg–Marquardt algorithm implemented in SciPy v1.11 [51]. Goodness of fit was evaluated by the coefficient of determination (R^2), root mean square error (RMSE), and reduced chi-squared statistic (χ^2/ν).

2.5.1. Pseudo-First-Order (PFO) Model

$$q_t = q_e(1 - e^{-k_1 t}) \quad (2)$$

where q_e is the equilibrium adsorption capacity (mg/g) and k_1 is the PFO rate constant (min^{-1}). This model assumes physisorption with a rate proportional to the number of unoccupied sites [52].

2.5.2. Pseudo-Second-Order (PSO) Model

$$q_t = \frac{k_2 q_e^2 t}{1 + k_2 q_e t} \quad (3)$$

where k_2 is the PSO rate constant (g/mg·min). This model assumes chemisorption involving electron sharing or exchange between the adsorbate and adsorbent [53].

2.5.3. Elovich Model

$$q_t = \frac{1}{\beta} \ln(1 + \alpha \beta t) \quad (4)$$

where α is the initial adsorption rate (mg/g·min) and β is related to the extent of surface coverage (g/mg). This model describes chemisorption on energetically heterogeneous surfaces [54].

2.5.4. Avrami Model

$$q_t = q_e \left[1 - e^{-(k_a t)^n} \right] \quad (5)$$

where k_a is the Avrami kinetic constant and n is the Avrami exponent reflecting the adsorption mechanism complexity. Despite being originally formulated for phase-change kinetics, the Avrami model has been applied to adsorption processes where PFO and PSO cannot satisfactorily describe the uptake dynamics [55,56]. The best-fitting model for each Metal $\times C_0$ curve was selected based on the highest R^2 and lowest χ^2/ν .

2.6. Diffusion Models

To elucidate the rate-controlling transport mechanism, three complementary diffusion models were applied to the kinetic data at $C_0 = 100$ and 200 mg/L. Curves at $C_0 = 10$ mg/L reached equilibrium within 2–4 min, providing insufficient temporal resolution for diffusion analysis.

2.6.1. Weber–Morris Intraparticle Diffusion Model

$$q_t = k_{id} \cdot t^{1/2} + C_i \quad (6)$$

where k_{id} is the intraparticle diffusion rate constant ($\text{mg/g} \cdot \text{min}^{0.5}$) and C_i is the intercept proportional to the boundary layer thickness (mg/g). If intraparticle diffusion is the sole rate-controlling step, the plot of q_t versus $\sqrt{t}$ passes through the origin [57]. Multilinearity in this plot indicates that adsorption proceeds through two or more sequential rate-limiting stages [58]. Three linear segments were fitted, corresponding to: Phase I ($t = 1\text{--}4$ min, external film diffusion), Phase II ($t = 6\text{--}15$ min, transition/pore-mouth dehydration), and Phase III ($t = 20\text{--}60$ min, intraparticle diffusion).

2.6.2. Boyd Film Diffusion Model

For fractional attainment $F(t) = q_t/q_e$:

$$B_t = -0.4977 - \ln(1 - F) \quad (7)$$

for $F > 0.85$, and:

$$B_t = \left(\sqrt{\pi} - \sqrt{\pi - \frac{\pi^2 F}{3}} \right)^2 \quad (8)$$

for $F \leq 0.85$. A linear plot of B_t versus t passing through the origin indicates film diffusion control; a non-zero intercept indicates intraparticle diffusion control [59].

2.6.3. Effective Diffusion Coefficient

The Vermeulen approximation [60] was used to estimate the effective intraparticle diffusion coefficient (D_e):

$$-\ln(1 - F^2) = \frac{4\pi^2 D_e t}{d^2} \quad (9)$$

where d is the average particle diameter (estimated as $50 \mu\text{m}$ based on SEM analysis). D_e was obtained from the slope of $-\ln(1 - F^2)$ versus t . Values in the range $10^{-11}\text{--}10^{-13} \text{ m}^2/\text{s}$ are characteristic of microporous zeolitic ion exchange [61], while values of approximately $10^{-9} \text{ m}^2/\text{s}$ correspond to free-solution ionic diffusion.

2.7. Equilibrium Isotherm Models

Apparent equilibrium data were obtained by plotting the adsorption capacity at $t = 60$ min ($q_{t,60}$) against the residual equilibrium concentration ($C_e = C_0 - q_{t,60} \times \text{dose}$) at each of the three initial concentrations. Four isotherm models were fitted by nonlinear regression:

2.7.1. Langmuir Model

$$q_e = \frac{q_m K_L C_e}{1 + K_L C_e} \quad (10)$$

where q_m is the maximum monolayer capacity (mg/g) and K_L is the Langmuir constant (L/mg). The dimensionless separation factor $R_L = 1/(1 + K_L C_0)$ was calculated to assess favorability: $0 < R_L < 1$ indicates favorable adsorption [62].

2.7.2. Freundlich Model

$$q_e = K_F \cdot C_e^{1/n} \quad (11)$$

where K_F is the Freundlich constant ($\text{mg}/\text{g} \cdot (\text{L}/\text{mg})^{1/n}$) and n is the heterogeneity factor. Values of $1/n < 1$ indicate favorable adsorption [63].

2.7.3. Temkin Model

$$q_e = \frac{RT}{b_T} \ln(K_T C_e) \quad (12)$$

where K_T is the Temkin equilibrium binding constant (L/mg), b_T is the Temkin constant related to the heat of adsorption (kJ/mol), R is the gas constant ($8.314 \text{ J}/\text{mol} \cdot \text{K}$), and T is the absolute temperature (298.15 K) [64].

2.7.4. Dubinin–Radushkevich (D-R) Model

$$q_e = q_m \cdot \exp(-\beta \varepsilon^2) \quad (13)$$

where β is a constant related to the mean free energy of adsorption (mol^2/kJ^2) and $\varepsilon = RT \cdot \ln(1 + 1/C_e)$ is the Polanyi potential. The mean free energy $E = 1/\sqrt{2\beta}$ classifies the adsorption mechanism: $E < 8 \text{ kJ}/\text{mol}$ indicates physisorption, $8 < E < 16 \text{ kJ}/\text{mol}$ indicates ion exchange, and $E > 16 \text{ kJ}/\text{mol}$ indicates chemisorption [65].

It should be noted that all isotherms were fitted to only three data points per metal (one per initial concentration), yielding one degree of freedom for two-parameter models. Additionally, only Hg^{2+} and Pb^{2+} had reached near-equilibrium at $t = 60 \text{ min}$; for the remaining metals, $q_{t,60}$ represents an apparent kinetic value rather than a true thermodynamic equilibrium capacity. These isotherm fits are therefore reported as preliminary estimates. A dedicated equilibrium study with 6–8 concentrations and extended contact times (up to 24 h) is recommended for definitive isotherm characterization.

2.8. Physical Metal Ion Descriptors

The eight metal physicochemical descriptors (ionic radius, electronegativity, hydration enthalpy, Pearson softness, hydrated radius, charge, charge density, ionic potential) are taken directly from established thermochemical reference tables [38–41,66] and have well-documented roles in ion exchange selectivity theory for zeolites [67–69]. The five experimental features (C_0 , t , and their $\log/\sqrt{t}$ transforms) encode the functional forms of classical kinetic models [52,57].

The six interaction descriptors were designed to encode specific physicochemical hypotheses about the metal–adsorbent system: pore accessibility (whether the hydrated ion fits through the GIS 8-ring aperture [70]), ion exchange potential (electrostatic driving force scaled by exchange capacity [67]), electrostatic factor (Coulombic attraction under the experimental pH [71]), HSAB match (Pearson acid–base compatibility [41]), dehydration penalty (energy cost of hydration shell removal [40]), and surface capacity. We acknowledge, however, that the interaction descriptors are deterministic functions of the metal properties and fixed adsorbent constants and, therefore, carry no information beyond what is already present in the raw descriptors.

Two derived descriptors were computed: charge density (Z/r^2), which quantifies the electrostatic field intensity at the ion surface and controls interaction strength with framework oxygens [67], and ionic potential (Z/r), which correlates with the tendency to form inner-sphere complexes [71]. All descriptors are external constants sourced from the literature prior to any data collection; they do not depend on the adsorption measurements and are therefore mathematically identical before and after the train/test split, eliminating any possibility of information leakage through feature engineering. Table 1 shows a summary of these physicochemical descriptors.

Table 1. Physicochemical descriptors of heavy metal ions. r = ionic radius; χ = Pauling electronegativity; ΔH_{hyd} = hydration enthalpy; σ = softness; r_{hyd} = hydrated radius.

Metal Ion	r (Å)	χ	ΔH_{hyd} (kJ/mol)	σ	r_{hyd} (Å)	Z	Z/r^2	Z/r
Pb ²⁺	1.19	2.33	−1481	0.58	4.01	2	1.41	1.68
Hg ²⁺	1.02	2.00	−1824	0.96	4.30	2	1.92	1.96
Cd ²⁺	0.95	1.69	−1807	0.69	4.26	2	2.22	2.11
Cu ²⁺	0.73	1.90	−2100	0.38	4.19	2	3.75	2.74
Ni ²⁺	0.69	1.91	−2106	0.32	4.04	2	4.20	2.90
Cr ³⁺	0.615	1.66	−4402	0.08	4.61	3	7.93	4.88

2.9. Metal–Adsorbent Interaction Features

Six physics-based interaction descriptors were computed to encode specific hypotheses about metal–adsorbent interactions on zeolite P:

(1) Pore accessibility ($d_{\text{pore}}/r_{\text{hyd}}$): the ratio of the GIS 8-ring channel (3.1×4.5 Å) to the hydrated ionic radius. d_{pore} was calculated as $d_{\text{pore}} = \sqrt{(3.1 \times 4.5)} = 3.73$ Å, that is, the geometric mean of the GIS 8-ring channel aperture dimensions [72]. All six metals yield $d_{\text{pore}}/r_{\text{hyd}} < 1$ (range: 0.81–0.93), confirming that partial dehydration is required for channel entry, consistent with the three-phase Weber–Morris diffusion kinetics observed experimentally.

(2) Ion exchange potential ($\text{CEC} \times Z/r^2$): the product of the cation exchange capacity (3.2 meq/g) and charge density, representing a first-order approximation of the thermodynamic driving force for exchange [67,68].

(3) Electrostatic factor ($|\text{pH} - \text{pH}_{\text{pzc}}| \times Z$): the Coulombic attraction between the charged surface at the experimental pH (3.9) relative to pH_{pzc} (3.2) and the metal cation charge [71].

(4) HSAB match ($1 - |\sigma_{\text{metal}} - \sigma_{\text{zeolite}}|$): quantifies compatibility between the metal ion and adsorbent framework according to the Pearson hard–soft acid–base principle [41]. The zeolite framework hardness was estimated as $\sigma = 0.35$ based on its low Si/Al ratio (1.63) and aluminosilicate oxygen donor character.

(5) Dehydration penalty ($|\Delta H_{\text{hyd}}|/1000$): the energy cost of removing the hydration shell, normalized to a dimensionless scale (1.48–4.40).

(6) Surface capacity ($\text{BET} \times Z/100$): the available binding sites per unit mass normalized by ionic charge. This yields only two distinct values: 0.60 for divalent metals and 0.90 for Cr³⁺.

Two of the interaction descriptors, i.e., dehydration penalty ($|\Delta H_{\text{hyd}}|/1000$) and surface capacity ($\text{BET} \times Z/100$), involve numerical rescaling by arbitrary constants (10^3 and 10^2 , respectively). These scaling factors were introduced to normalize the numerical range for gradient-based and distance-based algorithms (SVR, Ridge, Elastic Net) but are mathematically irrelevant for tree-based models, which are invariant to monotonic feature transformations [73]. The dehydration penalty is redundant with the raw hydration enthalpy already present in the feature set, and the surface capacity reduces to a coarse proxy for ionic charge since the BET surface area is constant across all experiments. The ablation study confirms that neither descriptor contributes measurable predictive value beyond the raw metal descriptors ($\Delta R^2 < 0.001$). These features are retained in the reported results, but readers should interpret them as rescaled versions of existing descriptors rather than independent physicochemical variables.

Three time-derived features ($\log(1 + t)$, $\sqrt{t}$, $\log(1 + C_0)$) were appended alongside the raw time and C_0 , yielding a total of 19 input features. The logarithmic and square-root transforms are the functional forms of established kinetic models: $\sqrt{t}$ appears in the Weber–Morris intraparticle diffusion model (Equation (6)) and $\log(t)$ linearizes pseudo-first-order

kinetics. Every feature is a deterministic function of either (a) external literature constants or (b) known experimental conditions (t , C_0). No feature is fitted on the data, eliminating any possibility of information leakage through feature engineering.

2.10. Machine Learning Models

Five regression algorithms were evaluated: Random Forest (RF), Gradient Boosting (GB), Ridge Regression, Elastic Net, and Support Vector Regression (SVR). Tree-based models (RF, GB) were used without feature scaling, as decision trees are invariant to monotonic transformations. Linear and kernel-based models (Ridge, Elastic Net, SVR) included a StandardScaler inside a scikit-learn Pipeline so that scaling parameters were fitted exclusively on training folds, preventing information leakage [74]. Hyperparameters were pre-specified based on literature defaults: RF ($n\_estimators = 200$, $max\_depth = 10$, $min\_samples\_split = 5$, $min\_samples\_leaf = 3$, $max\_features = \sqrt{p}$) and GB ($n\_estimators = 200$, $max\_depth = 5$, $learning\_rate = 0.05$, $subsample = 0.8$, $min\_samples\_split = 5$, $min\_samples\_leaf = 3$). Hyperparameters were not tuned on the development data, eliminating the need for nested cross-validation. All random seeds were fixed (seed = 42) to ensure full reproducibility.

2.11. Validation Strategy

2.11.1. Group-Level External Split

The 18 kinetic curves were divided into a development set (12 curves, 144 samples, 66.7%) and an external test set (6 curves, 72 samples, 33.3%) prior to any model development. The split was stratified at the metal level: exactly one curve per metal was randomly held out (seed = 42), ensuring that every metal was represented in both sets while guaranteeing that no time point from an external curve was ever used during training. This is a methodological improvement over row-level splits with stratification on metal, which can place different time points of the same kinetic curve in both training and test sets, allowing trivial interpolation-based prediction [75,76]. The external set was locked and used only once for the final performance evaluation.

2.11.2. Internal Cross-Validation

Three CV strategies of increasing stringency were applied exclusively to the development set: (i) Repeated 5-Fold CV (5 folds $\times$ 10 repeats = 50 iterations) estimates interpolation performance within the training distribution; (ii) Group K-Fold (groups = Metal $\times$ C_0 kinetic curves, $k = \min(6, n\_groups)$) estimates curve-level extrapolation by ensuring that entire kinetic curves are held out in each fold; and (iii) Leave-One-Metal-Out (LOMO) estimates the models ability to predict a completely novel metal by training on five metals and predicting the sixth. Metrics reported include R^2 , MAE, and RMSE, with 95% parametric confidence intervals for Repeated K-Fold scores.

2.11.3. Statistical Comparison

Pairwise Wilcoxon signed-rank tests with Bonferroni correction for 10 pairwise comparisons were used to assess whether differences between models were statistically significant at $\alpha = 0.05$ [77].

2.12. Interpretability: SHAP Analysis

SHapley Additive exPlanations (SHAP) were computed for the tree-based models using TreeExplainer [42]. SHAP decomposes each individual prediction into additive contributions from each feature based on cooperative game theory. Mean absolute SHAP values provide a global importance ranking, while SHAP summary (beeswarm) plots reveal the direction and magnitude of each feature's influence on the predicted q_t . SHAP

importance was cross-validated against Mean Decrease in Impurity (MDI) and permutation importance to ensure robustness [78]. SHAP is employed as a model-interpretability tool to evaluate whether the learned feature importance is consistent with—not proof of—established ion exchange selectivity theory; the correlative nature of SHAP analysis is explicitly acknowledged.

2.13. Ablation Study

To quantify the contribution of each feature category to predictive performance, an ablation study was conducted by training the three best-performing models (RF, GB, Ridge) on four progressively inclusive feature sets: (i) Experimental-Only (5 features: C_0 , t , $\log(1 + t)$, $\sqrt{t}$, $\log(1 + C_0)$); (ii) Descriptors-Only (14 features: 8 metal descriptors + 6 interaction terms, no time or concentration); (iii) Metal Descriptors + Experimental (13 features: 8 metal descriptors + 5 experimental, no interaction terms); and (iv) Full feature set (all 19 features). Each configuration was evaluated under all three CV strategies (Repeated K-Fold, Group K-Fold, LOMO) and on the external test set. Statistical significance of performance differences between the Full and Experimental-Only sets was assessed by Wilcoxon signed-rank tests on the 50 paired K-Fold R^2 scores.

2.14. Selectivity–Descriptor Correlation Analysis

To quantify the relationship between metal physicochemical properties and adsorption capacity independently of the ML model, Pearson and Spearman correlation coefficients were computed between each metal descriptor (electronegativity, ionic radius, hydration enthalpy, softness, hydrated radius) and the apparent adsorption capacity at $C_0 = 200$ mg/L and $t = 60$ min across the six metals ($n = 6$). This analysis provides a direct, model-free assessment of which physicochemical properties govern uptake capacity and serves as an independent benchmark against which SHAP feature importance rankings can be evaluated.

2.15. Reproducibility

All adsorption experiments were performed as single measurements without independent replicates. Computational reproducibility is ensured by a fixed random seed (42) for all stochastic operations, explicit pre-specified hyperparameters (not tuned on this dataset), and a split table documenting development versus external curve assignments. All 19 features are deterministic transforms of raw experimental variables (C_0 , t , $\log(1 + t)$, $\sqrt{t}$, $\log(1 + C_0)$) or look-up values from published thermochemical tables [38–41,66]. No feature is fitted on data (no target encoding, no PCA, no statistical aggregations). This means the feature engineering step is identical regardless of when it is applied, before or after the train/test split, and introduces no information leakage.

3. Results

3.1. Characterization of Coal Fly Ash and Synthesized Zeolite P

An experiment was conducted to determine the optimal synthesis conditions of zeolite from fly ash (optimal conditions for hydrothermal synthesis: 1.5 M NaOH at 150 °C). The target variable was the percentage of total metal ions removed from the solution. The study analyzed the effect of adsorbent amount on the removal of metal ions. According to Table A1, zeolites produced with a higher molar concentration of NaOH and material amount showed a significant improvement in metal removal. On the other hand, zeolites produced with a low molar concentration of NaOH and a low material amount did not remove metal ions significantly.

The optimal zeolite amount was 0.6 g. The fact that zeolite P readily forms as an equilibrium phase at high temperatures demonstrates the fast growth rate of crystallization, which is not possible at lower temperatures due to slow nucleation [24].

3.1.1. XRD Analysis

X-ray diffraction was used to determine the crystalline phases present in the fly ash sample (see Figure A1a). The sample is mainly composed of mullite [79] and silicon dioxide [80]. The concentration of NaOH plays a crucial role in zeolite synthesis as it dissolves mullite and quartz to form an aluminosilicate gel [81]. Under optimal synthesis conditions, zeolite P was the primary product, as shown in Figure A1b. The as-synthesized material will be referred to as zeolite P [82]. Figure A1b shows the experimental (red) and calculated (black) diffraction patterns using the Le Bail method. The difference between the calculated and experimental diffractograms is shown at the bottom of the figure (blue line). The diffractogram also shows the peak positions for zeolite P, mullite, and silicon dioxide. The crystalline fraction obtained from the diffractogram indicates a conversion of 71.1% of fly ash to zeolite P (area under the curve). The mullite content in the final product is 27.1%, with the SiO₂ content below 2%. The fly ash used in this study contains 58% mullite and 42% silicon dioxide. Characteristic diffraction peaks at $2\theta = 12.4^\circ$, 17.6° , 21.6° , 28.1° , and 33.4° corresponded to (110), (200), (211), (310), and (222) reflections of the zeolite phase [3,83].

The synthesis conditions of the zeolite resulted in partial dissolution of the mullite, while almost all the silicon dioxide was converted. In previous reports, the yield of zeolite obtained through hydrothermal synthesis has varied from 20% to 60%. However, these variations are due to factors such as the type and composition of the fly ash, the alkali used and its concentration (KOH, NaOH, LiOH, etc.), the temperature and time of the reaction, and the liquid-to-solid ratio [83–86]. Table A2 summarizes the mineralogical phases identified based on the crystallographic cards, where the crystal system, space groups, and lattice parameters are observed.

3.1.2. SEM-EDS and X-Ray Fluorescence Spectroscopy Analyses

Figure A2 shows the SEM images of fly ash and zeolite P at 2500 $\times$ and 7000 $\times$ magnifications. Fly ashes consist of spheres with a diameter ranging from 2 to 10 μm [87]. After hydrothermal synthesis, the spheres underwent a change in morphology, resulting in the formation of cauliflower-like structures.

Figure A3 shows the elemental mapping for zeolite P, which exhibits a uniform distribution of elements. The brighter and larger surfaces of the particles are filled by the most abundant elements, such as oxygen, aluminum, silicon, and sodium. The size of the small spheres or ceno-spheres varies depending on the carbon and iron content, with diameters ranging from 20 to 200 μm and a shell wall thickness ranging from 2 to 10 μm [88]. The carbon present in the sample is attributed to the residues of unburnt carbon. The elements found in highest abundance were aluminum, silicon, iron, calcium, and sodium. Reports indicate that iron-rich fly ashes are primarily composed of iron oxide and amorphous aluminosilicate, with calcium being the fourth most abundant element found in the spheres [89].

X-ray fluorescence spectroscopy was used to determine the chemical composition of fly ash and zeolite P. The fly ash is classified as type F, with concentrations of SiO₂, Al₂O₃, and Fe₂O₃ higher than 70%. The ideal Si/Al atomic ratio for converting fly ash to zeolite ranges from 1:1 to 2:1 because higher ratios may decrease the yield of the zeolite phase during hydrothermal synthesis [7,86]. The Si/Al atomic ratio decreased during the conversion process. Table A3 summarizes the composition of both fly ash and zeolite P.

The analysis indicated a Si/Al ratio of 1.63, within the characteristic range for zeolite P (1.0–2.5) [70]. The concentrations of Fe_2O_3 , Al_2O_3 , and CaO remained almost unchanged, while the concentration of SiO_2 decreased due to the reaction. It has been observed that the Na concentration increased significantly because NaOH was used in the synthesis process to dissolve the mullite and silicon oxide in the fly ash.

3.1.3. FTIR Analysis

Figure A4 shows the FTIR spectra of fly ash and zeolite P in the region associated with the asymmetric stretching vibrations of Si–O–Si and Si–O–Al linkages ($950\text{--}1255\text{ cm}^{-1}$). The fly ash exhibits a broad band centered at 1057 cm^{-1} , characteristic of its predominantly amorphous aluminosilicate structure. After hydrothermal treatment, this band shifts to 986 cm^{-1} for zeolite P, indicating increased incorporation of Al into the tetrahedral framework [90]. This shift toward lower wavenumbers reflects the formation of a more ordered aluminosilicate structure, in agreement with the appearance of well-defined diffraction peaks assigned to zeolite P in the XRD analysis [91].

Additional bands observed in the $500\text{--}420\text{ cm}^{-1}$ region correspond to bending vibrations of Si–O–Al linkages. The band at 434 cm^{-1} in fly ash is replaced by a band at 427 cm^{-1} in zeolite P, which is associated with vibrations of the secondary building units of the zeolite framework [92]. This fact further supports the crystallization of zeolite P inferred from XRD. Moreover, the asymmetric and symmetric stretching vibrations of Si–O–Si and Si–O–Al bridges shift from 794 cm^{-1} in fly ash to 740 cm^{-1} in zeolite P, providing additional spectroscopic evidence of framework reorganization during the formation of zeolite, consistent with the phase evolution observed in the XRD patterns [13].

3.1.4. BET Surface Area and Porosity

The specific surface area (SSA), average pore volume (APV), and average pore diameter (APD) of zeolite P are $30\text{ m}^2/\text{g}$, $0.07\text{ cm}^3/\text{g}$, and 2.9 nm , respectively. In contrast, the corresponding properties of fly ash are $1\text{ m}^2/\text{g}$, $0.01\text{ cm}^3/\text{g}$, and 3.6 nm . Zeolite P has a significantly higher specific surface area and pore volume than fly ash, although its average pore diameter is slightly smaller. Despite the increasing SSA in zeolite P, this value is lower than values reported for pure synthetic zeolite P ($50\text{--}150\text{ m}^2/\text{g}$) [93], attributable to residual non-zeolitic phases from fly ash and bigger crystallites reducing the access to porosity.

3.1.5. Zeta Potential

The point of zero charge ($\text{pH}_{\text{pzc}} = 3.2$) indicates a slightly negative surface charge under the experimental pH conditions ($\text{pH} = 3.9$), favoring electrostatic attraction of cationic metal species [94].

3.1.6. Physicochemical Properties of Synthesized Zeolite P

Some of the key physicochemical properties used in the ML approaches are summarized in Table 2. The BET surface area was determined as $30\text{ m}^2/\text{g}$ using nitrogen adsorption at 77 K . The point of zero charge (pH_{pzc}) was measured as 3.2 by the zeta potential. The pore diameter of 3.73 Å corresponds to the geometric mean that is characteristic of the GIS framework topology. The cation exchange capacity (CEC) of 3.2 meq/g , and Si/Al ratio of 1.63 (experimental) are consistent with aluminum-rich zeolite P variants exhibiting ion exchange potential [3].

3.2. Heavy Metal Adsorption Kinetics

The kinetic profiles for heavy metal adsorption onto zeolite P are shown in Figure 1. All metals developed rapid initial uptake within the first 10 min, gradually approaching equilibrium. At $C_0 = 10\text{ mg/L}$, equilibrium was attained within 5–10 min for most metals,

reflecting the high driving force at low coverage and availability of sites. At higher concentrations (100 and 200 mg/L), equilibrium required more time, as progressive site occupation reduced the adsorption rate.

Table 2. Physicochemical properties of zeolite P used in this study.

Property	Value	Unit
BET specific surface area	30.0	m ² /g
pH _{pzc}	3.2	—
Pore diameter	3.73	Å
CEC	3.2	meq/g
Si/Al ratio	1.63	—

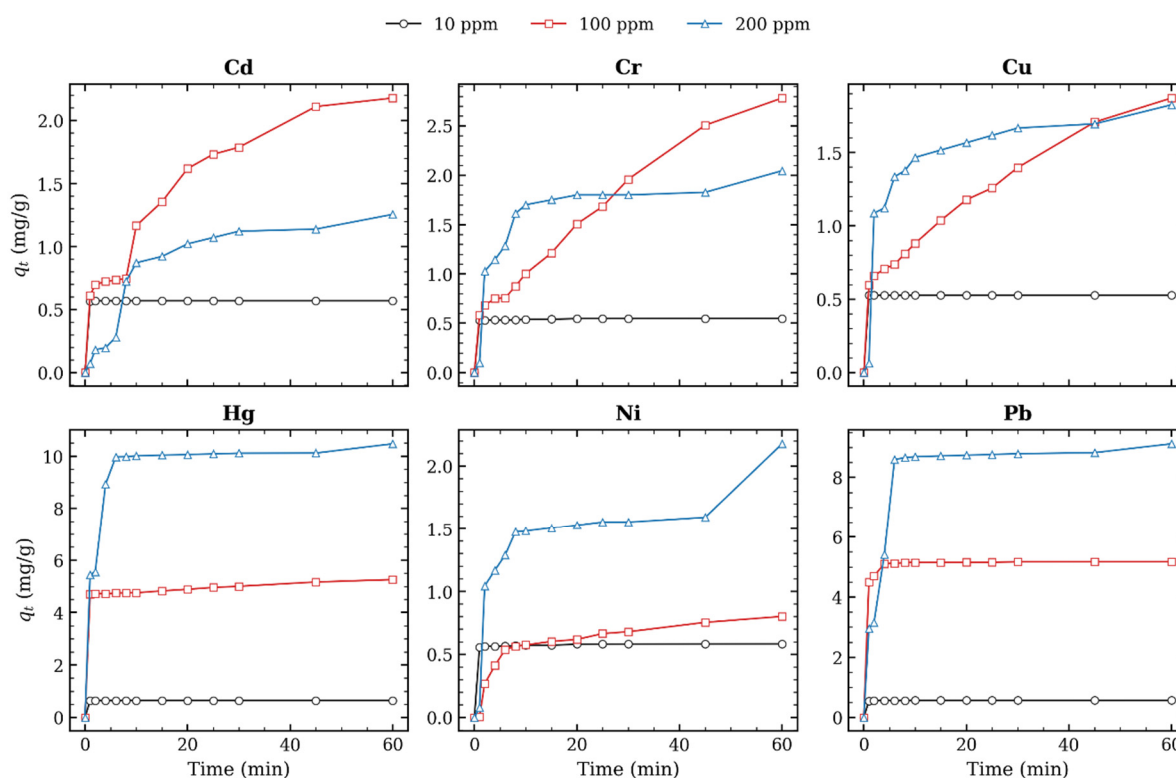


Figure 1. Relation between contact time and adsorption capacity at different initial concentrations, where q_t is the adsorption capacity at time t .

The following behavior of q_t is observed at the three different concentrations:

At $C_0 = 10$ mg/L: Hg^{2+} (0.63) > Pb^{2+} (0.56) $\approx$ Ni^{2+} (0.56) $\approx$ Cd^{2+} (0.55) $\approx$ Cu^{2+} (0.54) $\approx$ Cr^{3+} (0.53 mg/g).

At $C_0 = 100$ mg/L: Hg^{2+} (5.26) > Pb^{2+} (5.17) >> Cr^{3+} (2.78) > Cd^{2+} (2.18) > Cu^{2+} (1.87) > Ni^{2+} (0.80 mg/g).

At $C_0 = 200$ mg/L: Hg^{2+} (10.47) > Pb^{2+} (9.12) >> Ni^{2+} (2.18) > Cr^{3+} (2.05) > Cu^{2+} (1.82) > Cd^{2+} (1.26 mg/g).

Thus, distinct kinetic behaviors were observed depending on metal type and concentration as follows:

Mercury (Hg^{2+}) and lead (Pb^{2+}) exhibited the highest adsorption capacities among the studied metals. At $C_0 = 200$ mg/L, Hg reached $q_{t,\text{max}} = 10.47$ mg/g and Pb achieved $q_{t,\text{max}} = 9.12$ mg/g after 60 min. Both metals showed rapid initial uptake during the first 10–15 min, followed by a gradual approach to equilibrium. This behavior is consistent with the high affinity of zeolite P for these divalent cations with favorable ionic radii and

charge density for ion exchange. Pb^{2+} and Hg^{2+} at all concentrations and all six metals at $C_0 = 10 \text{ mg/L}$ follow a classical pattern, consistent with Mechnou et al. [95] and PSO kinetics reported for zeolite systems [96].

Chromium (Cr^{3+}) and cadmium (Cd^{2+}) displayed intermediate adsorption capacities with $q_{t,\text{max}}$ values of 2.78 and 2.18 mg/g at $C_0 = 100 \text{ mg/L}$, respectively. Notably, Cr showed prolonged uptake throughout the contact period, suggesting both ion exchange and surface complexation mechanisms. Since $\text{CrCl}_3 \cdot 6\text{H}_2\text{O}$ was used as the precursor, Cr exists exclusively as Cr(III) in the trivalent oxidation state at the experimental pH of 3.9. The dominant species is $[\text{Cr}(\text{H}_2\text{O})_6]^{3+}$ with minor CrOH^{2+} . The anomalous kinetic behavior is attributed to the extreme kinetic inertness of $[\text{Cr}(\text{H}_2\text{O})_6]^{3+}$, the highest hydration enthalpy (-4402 kJ/mol), steric constraints from the largest hydrated radius ($4.61 \text{ \AA}$), and concentration-dependent oligomerization to polynuclear hydroxo-bridged species [97–99]. Cr(III) undergoes concentration-dependent oligomerization to form polynuclear hydroxo-bridged species [99] that exceed the channel dimensions and must dissociate before exchange can proceed. In contrast, Pb^{2+} and Hg^{2+} , which have faster dehydration kinetics and more favorable pore accessibility, follow the expected thermodynamic trend at 60 min.

However, Cd^{2+} and Cr^{3+} at $C_0 = 100\text{--}200 \text{ mg/L}$ exhibit a three-phase departure: rapid initial uptake (Phase I), stagnation (Phase II), and secondary re-acceleration (Phase III). This behavior reflects the fundamental structural difference between mesoporous activated carbons (BET $500\text{--}2400 \text{ m}^2/\text{g}$, pore diameters $2\text{--}50 \text{ nm}$, surface physisorption) and microporous zeolite P (BET $30 \text{ m}^2/\text{g}$, GIS eight-ring channels, ion exchange). The channel dimensions are comparable to or smaller than the hydrated radii of Cd^{2+} ($4.26 \text{ \AA}$) and Cr^{3+} ($4.61 \text{ \AA}$), creating a dehydration bottleneck ($\Delta H_{\text{hyd}} = -1807$ and -4402 kJ/mol [40]) absent in mesoporous systems. Phase I corresponds to fast external-site exchange, Phase II to dehydration-gated stagnation at the pore mouth, and Phase III to re-acceleration once the first ions transit the bottleneck and allow queued ions to follow. The apparent inversion in adsorption capacity at $C_0 = 200 \text{ mg/L}$ relative to $C_0 = 100 \text{ mg/L}$ observed for Cd^{2+} ($q_{t,60} = 1.26$ vs. 2.18 mg/g) and Cr^{3+} ($q_{t,60} = 2.05$ vs. 2.78 mg/g) reflects a kinetic rather than thermodynamic limitation. Neither concentration fully reached equilibrium at $t = 60 \text{ min}$, as evidenced by the positive slopes of both curves at the final time point. The dramatically suppressed initial uptake at $C_0 = 200 \text{ mg/L}$ ($q_{t,1} = 0.071 \text{ mg/g}$ for Cd vs. 0.612 at $C_0 = 100 \text{ mg/L}$) indicates severe transport limitations in the narrow GIS channels, where high ionic concentrations at the pore mouth create mutual steric hindrance that retards intraparticle diffusion [72,100]. Additionally, enhanced metal hydrolysis at $C_0 = 200 \text{ mg/L}$ depresses the local pH near the adsorbent surface, increasing H^+ competition for exchange sites; this effect is particularly severe at the working pH of 3.9, which is already close to the pH_{pzc} of the zeolite (3.2) [101].

Copper (Cu^{2+}) and nickel (Ni^{2+}) showed the lowest adsorption capacities ($q_{t,\text{max}} < 2.2 \text{ mg/g}$) due to weaker interactions with zeolite P. At low initial concentration, all metals rapidly reached equilibrium in a few minutes with q_t values around $0.5\text{--}0.6 \text{ mg/g}$, suggesting that available adsorption sites were quickly saturated under these conditions.

3.3. Traditional Kinetic Model Performance

Four kinetic models (PFO, PSO, Elovich, and Avrami) were fitted to each of the 18 experimental kinetic curves (six metals $\times$ three initial concentrations) by nonlinear regression. The best-fitting model for each curve was selected based on the highest R^2 and lowest reduced χ^2/ν . The PSO model provided the best fit for 10 of 18 curves (56%), the Avrami model for seven curves (39%), and PFO for one curve (6%). The Elovich model was not the best fit for any curve. Table 3 summarizes the kinetic model for the best-fitting

model for each Metal $\times$ C_0 combination. The complete set of fitted parameters for all four models is provided in the Appendix A section (Tables A4 and A5).

Table 3. Best-fitting conventional kinetic model for each Metal $\times$ C_0 combination.

Metal Ion	C_0 (mg/L)	q_e , Exp	q_e , Calc	Best Model	R^2	RMSE	χ^2/ν
Pb ²⁺	10	0.56	0.56	PSO	0.9144	0.003	0.0002
Pb ²⁺	100	5.17	5.21	PSO	0.9425	0.051	0.0062
Pb ²⁺	200	9.12	8.91	Avrami	0.9438	0.524	0.9036
Ni ²⁺	10	0.56	0.56	PSO	0.9402	0.002	0.0000
Ni ²⁺	100	0.80	0.82	PSO	0.9348	0.054	0.1672
Ni ²⁺	200	2.18	1.86	PSO	0.8161	0.202	0.5825
Cu ²⁺	10	0.54	0.54	PSO	0.9143	0.003	0.0002
Cu ²⁺	100	1.87	50.0	Avrami	0.9576	0.083	0.1346
Cu ²⁺	200	1.82	1.83	PSO	0.8748	0.157	0.5056
Cr ³⁺	10	0.53	0.53	PSO	0.9400	0.002	0.0000
Cr ³⁺	100	2.78	50.0	Avrami	0.9692	0.124	0.3639
Cr ³⁺	200	2.05	1.85	PFO	0.9188	0.146	0.4026
Hg ²⁺	10	0.63	0.63	PSO	0.9405	0.002	0.0000
Hg ²⁺	100	5.26	11.24	Avrami	0.7612	0.087	0.0188
Hg ²⁺	200	10.47	10.22	Avrami	0.9305	0.449	0.3742
Cd ²⁺	10	0.55	0.55	PSO	0.9262	0.002	0.0001
Cd ²⁺	100	2.18	10.60	Avrami	0.9447	0.132	0.2417
Cd ²⁺	200	1.26	1.14	Avrami	0.9559	0.087	0.2292

The PSO model provided the best fit predominantly at $C_0 = 10$ mg/L (all six metals) and for metals with fast, monotonic uptake at higher concentrations (Pb²⁺ at $C_0 = 100$ mg/L, Ni²⁺ at $C_0 = 100$ – 200 mg/L, Cu²⁺ at $C_0 = 200$ mg/L, Cd²⁺ at $C_0 = 10$ mg/L). The Avrami model was best specifically for metals exhibiting the three-phase kinetic pattern (Cd²⁺ and Cr³⁺ at $C_0 = 100$ mg/L), sigmoidal uptake (Hg²⁺ at $C_0 = 100$ – 200 mg/L), and the slow progressive regime (Pb²⁺ at $C_0 = 200$ mg/L, Cd²⁺ at $C_0 = 200$ mg/L). PFO was best for only one curve (Cr³⁺ at $C_0 = 200$ mg/L), where the near-linear uptake under far-from-equilibrium conditions is well described by the exponential approach. The Elovich model was not the best fit for any curve, consistent with its assumption of energetically heterogeneous surfaces, which is less applicable to well-defined exchange sites of zeolite P [53,55].

Several observations merit comment. At $C_0 = 10$ mg/L, all metals show very high PSO k_2 values (28–46 g/mg·min) and initial rates $h = k_2 q_e^2$ of 8–18 mg/g·min, reflecting the rapid saturation of the limited number of metal ions on readily accessible exchange sites. The PSO-calculated q_e values at this concentration (0.53–0.63 mg/g) are in excellent agreement with experimental values, confirming that true equilibrium was reached. At $C_0 = 100$ – 200 mg/L, the PSO k_2 values drop by 1–3 orders of magnitude (0.009–0.276 g/mg·min), consistent with the transition from surface-exchange-controlled kinetics to intraparticle-diffusion-controlled kinetics as loading increases [96]. For several Avrami fits at $C_0 = 100$ mg/L (Cu, Cr, Cd), the optimizer converged to boundary values ($q_e \approx 50$ mg/g, $k_a \approx 0$) that lack physical meaning but nonetheless captured the overall curve shape; these parameters should be interpreted with caution. The Avrami exponent n provides mechanistic information: $n \approx 1.0$ – 1.5 for Pb and Cd at $C_0 = 200$ mg/L suggests a combination of nucleation-like and diffusion-controlled uptake, while $n < 0.5$ for Hg at $C_0 = 100$ suggests a purely diffusion-limited process [55].

The dominance of PSO kinetics is consistent with a chemisorption/exchange mechanism involving electron sharing between the metal cation and the zeolite framework oxygens [53]. For microporous zeolites, PSO kinetics have been specifically associated

with ion exchange-controlled uptake rather than physisorption [96]. The Avrami model's superior performance for several high-concentration curves indicates that simple PSO cannot capture the multi-stage transport dynamics that emerge when intraparticle diffusion through the narrow GIS channels becomes rate-limiting [55].

3.4. Diffusion Mechanisms Analysis

3.4.1. Weber–Morris Intraparticle Diffusion

Plots of q_t versus $\sqrt{t}$ for all metals at $C_0 = 100$ and 200 mg/L exhibit multilinearity, confirming at least three sequential rate-controlling steps [57,58]: Phase I ($t = 1\text{--}4$ min, external film diffusion), Phase II ($t = 6\text{--}15$ min, pore-mouth dehydration and entry into the GIS 8-ring channels), and Phase III ($t = 20\text{--}60$ min, intraparticle diffusion). For Cd^{2+} and Cr^{3+} , k_{id2} exceeded k_{id1} (Cd at $C_0 = 100$ mg/L: $k_{id2} = 0.480$ vs. $k_{id1} = 0.105$ mg/g·min $^{-0.5}$), quantitatively confirming the re-acceleration phenomenon attributable to a dehydration “gate-opening” effect at the pore mouth. For Pb^{2+} and Hg^{2+} , k_{id} decreased monotonically from Phase I to III, consistent with conventional diffusion kinetics. Non-zero C-intercepts across all phases indicate that intraparticle diffusion is not the sole rate-controlling step [58]. Figure 2 shows the Weber–Morris plots for all metals at $C_0 = 100$ and 200 mg/L. In addition, Table 4 shows the different values of k_{id} for $C_0 = 100$ mg/L.

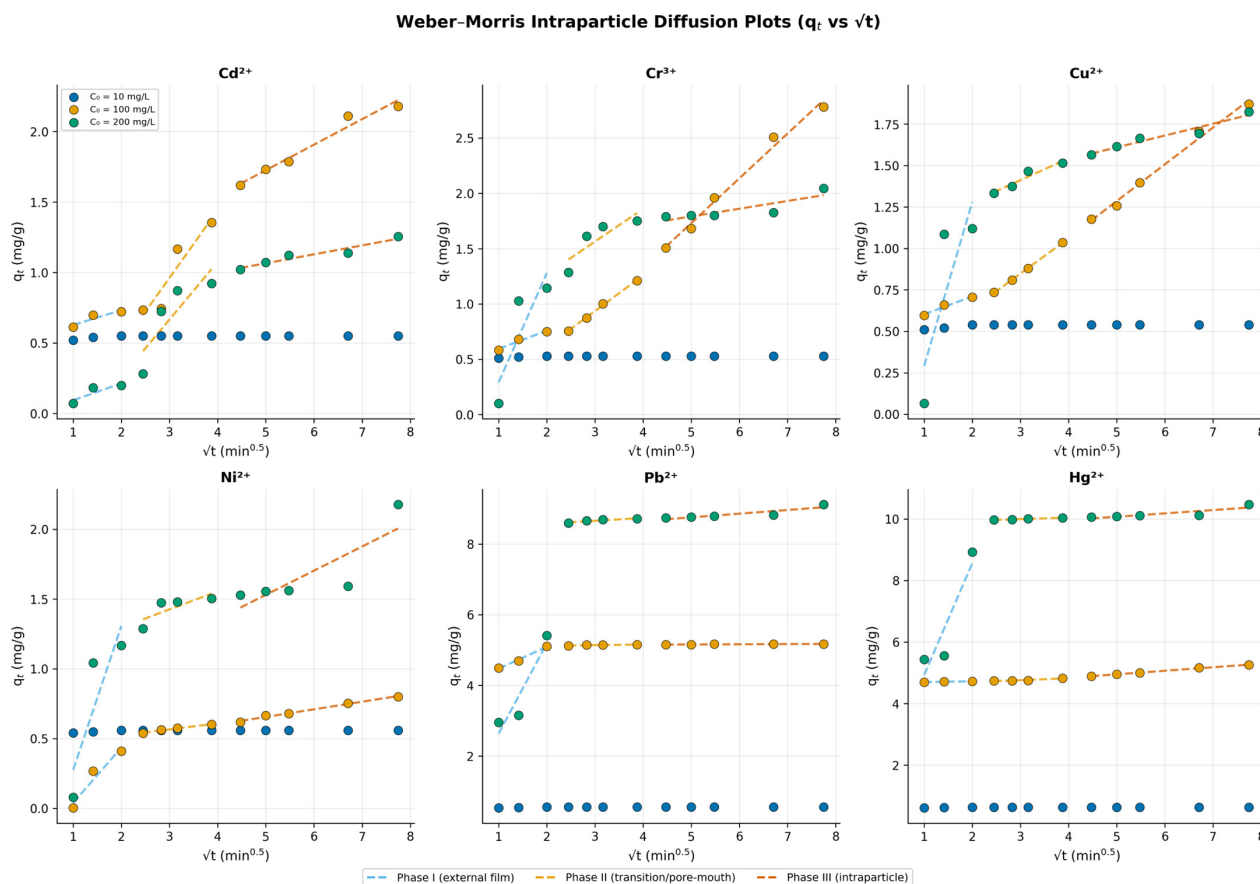


Figure 2. Weber–Morris q_t vs. $\sqrt{t}$ plots for all metals at $C_0 = 100$ and 200 mg/L (2×3).

Table 4. Weber–Morris intraparticle diffusion parameters at $C_0 = 100$ mg/L. k_{id} in mg/g·min $^{-0.5}$; C in mg/g.

Metal Ion	k_{id1}	C_1	k_{id2}	C_2	k_{id3}	C_3
Pb^{2+}	0.619	3.855	0.018	5.086	0.005	5.130
Ni^{2+}	0.396	−0.356	0.044	0.436	0.055	0.383

Table 4. Cont.

Metal Ion	k_{id1}	C_1	k_{id2}	C_2	k_{id3}	C_3
Cu^{2+}	0.108	0.495	0.212	0.213	0.222	0.177
Cr^{3+}	0.164	0.431	0.322	−0.029	0.405	−0.293
Hg^{2+}	0.026	4.673	0.061	4.581	0.115	4.381
Cd^{2+}	0.105	0.522	0.480	−0.477	0.181	0.820

3.4.2. Boyd Film Diffusion

Boyd plots (B_t vs. t) discriminate between film and intraparticle diffusion control [59]. At $C_0 = 100$ mg/L, near-zero intercepts for Cd^{2+} (−0.116) and Cr^{3+} (−0.236) indicate significant film diffusion contribution, whereas at $C_0 = 200$ mg/L, positive intercepts (Cu: 0.401; Cr: 0.380) confirm that intraparticle diffusion becomes rate-limiting at higher ionic loading. This concentration-dependent transition is consistent with the pore-congestion mechanism in the narrow GIS channels [61,100]. Figure 3 shows the Boyd film diffusion at $C_0 = 100$ and 200 mg/L. Table 5 summarizes the Boyd film diffusion analysis for the highest concentrations.

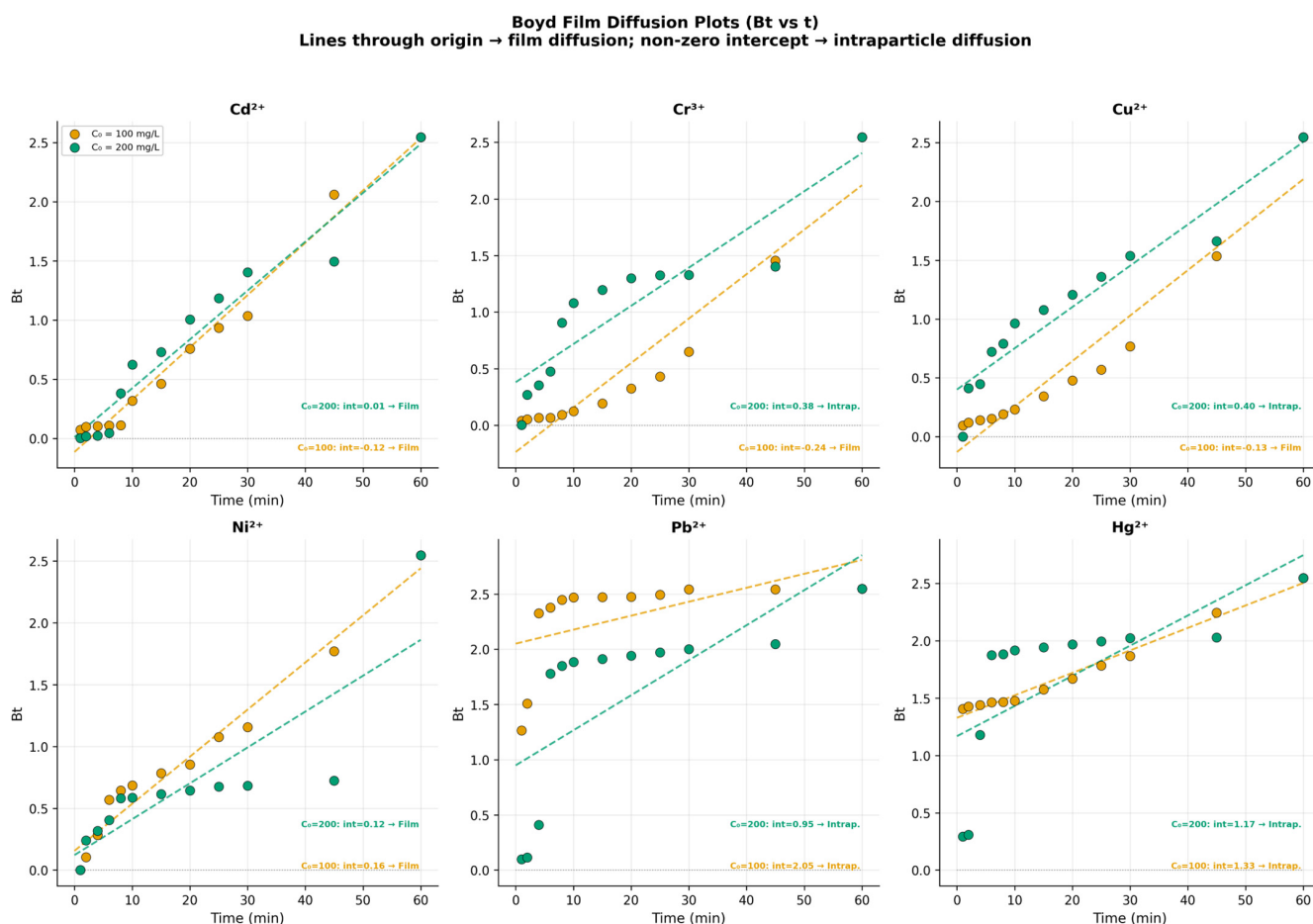
Figure 3. Boyd B_t vs. t plots for all metals at $C_0 = 100$ and 200 mg/L.

Table 5. Boyd film diffusion analysis.

Metal Ion	C_0	q_e , Est	B Slope	Intercept	R^2	Rate-Limiting Step
Pb^{2+}	100	5.43	0.013	2.051	0.295	Intraparticle
Pb^{2+}	200	9.58	0.032	0.949	0.493	Intraparticle
Ni^{2+}	100	0.84	0.038	0.156	0.965	Film

Table 5. Cont.

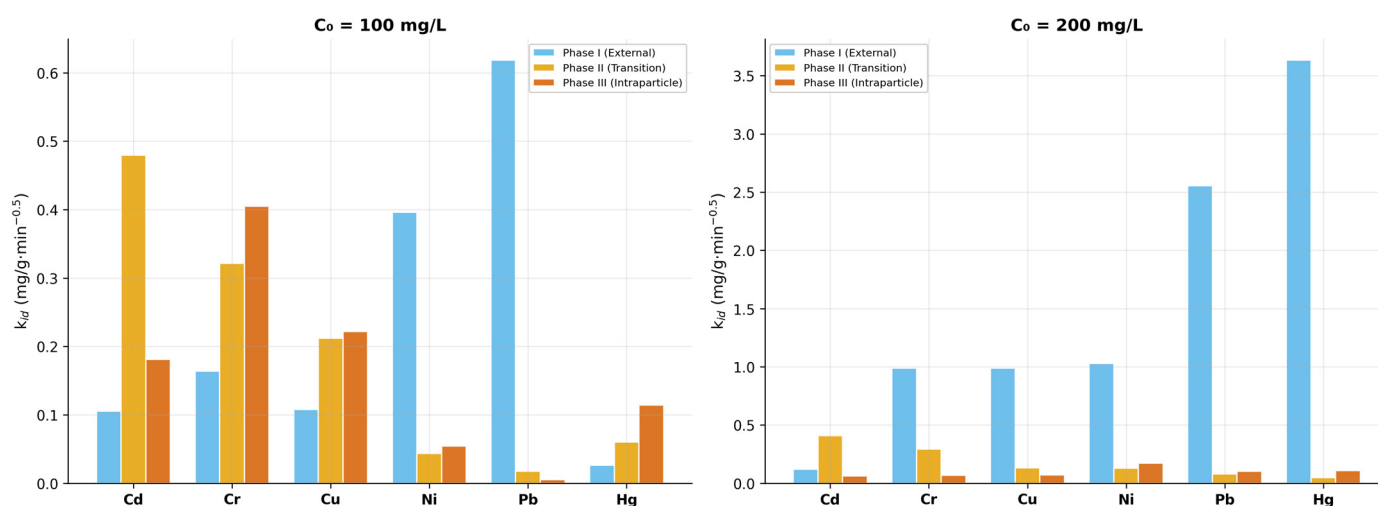
Metal Ion	C ₀	q _e , Est	B Slope	Intercept	R ²	Rate-Limiting Step
Ni ²⁺	200	2.29	0.029	0.122	0.720	Film
Cu ²⁺	100	1.96	0.039	−0.130	0.934	Film
Cu ²⁺	200	1.92	0.035	0.401	0.916	Intraparticle
Cr ³⁺	100	2.92	0.039	−0.236	0.909	Film
Cr ³⁺	200	2.15	0.034	0.380	0.830	Intraparticle
Hg ²⁺	100	5.52	0.020	1.329	0.985	Intraparticle
Hg ²⁺	200	11.00	0.026	1.168	0.477	Intraparticle
Cd ²⁺	100	2.29	0.044	−0.116	0.982	Film
Cd ²⁺	200	1.32	0.041	0.010	0.949	Film

3.4.3. Effective Diffusion Coefficients

Vermeulen-derived effective diffusivities (D_e) ranged from 10^{-10} to 10^{-9} cm²/s for all metals (Table 6), characteristic of microporous zeolitic diffusion [43,61] and two to three orders of magnitude below free-solution ionic diffusion ($\sim 10^{-6}$ cm²/s), confirming that the zeolite pore structure imposes substantial transport resistance. Figure 4 depicts the Weber–Morris diffusion rate constants.

Table 6. Effective intraparticle diffusion coefficients (Vermeulen approximation, $d = 50$ μm).

Metal Ion	C ₀ (mg/L)	D _e (cm ² /s)	Magnitude
Pb ²⁺	100	1.28×10^{-10}	$10^{-9.9}$
Pb ²⁺	200	3.11×10^{-10}	$10^{-9.5}$
Ni ²⁺	100	3.72×10^{-10}	$10^{-9.4}$
Ni ²⁺	200	2.84×10^{-10}	$10^{-9.5}$
Cu ²⁺	100	3.76×10^{-10}	$10^{-9.4}$
Cu ²⁺	200	3.43×10^{-10}	$10^{-9.5}$
Cr ³⁺	100	3.83×10^{-10}	$10^{-9.4}$
Cr ³⁺	200	3.30×10^{-10}	$10^{-9.5}$
Hg ²⁺	100	1.97×10^{-10}	$10^{-9.7}$
Hg ²⁺	200	2.59×10^{-10}	$10^{-9.6}$
Cd ²⁺	100	4.31×10^{-10}	$10^{-9.4}$
Cd ²⁺	200	4.03×10^{-10}	$10^{-9.4}$

Weber–Morris Diffusion Rate Constants (k_{id}) by Phase and MetalFigure 4. Comparison of k_{id} across metals and concentrations at $C_0 = 100$ and 200 mg/L.

Collectively, these results demonstrate that adsorption proceeds through three sequential diffusion stages, whose rate-limiting step shifts from external film diffusion at $C_0 = 100$ mg/L to intraparticle diffusion at $C_0 = 200$ mg/L, and that the anomalous re-acceleration for Cd^{2+} and Cr^{3+} corresponds to the dehydration-gated entry into the GIS micropore system.

3.5. Equilibrium Isotherm Analysis

Four isotherm models (Langmuir, Freundlich, Temkin, and Dubinin–Radushkevich) were fitted to the apparent equilibrium data (q_t at $t = 60$ min vs. C_e) at three initial concentrations per metal. It should be noted that all isotherms were fitted to only three data points per metal (one degree of freedom for two-parameter models), and only Hg^{2+} and Pb^{2+} had reached near-equilibrium at $t = 60$ min; for the remaining metals, $q_{t,60}$ represents an apparent kinetic value rather than a true thermodynamic equilibrium capacity. These results are therefore reported as preliminary estimates.

For Pb^{2+} and Hg^{2+} , Langmuir and Freundlich provided excellent fits ($R^2 > 0.99$), with Langmuir separation factors ($0 < R_L < 1$) confirming favorable adsorption at all three concentrations [62]. The Freundlich $1/n$ values for these metals (0.84 and 0.98) indicate favorable uptake. For the remaining metals, isotherm fits were moderate to poor ($R^2 = 0.38\text{--}0.97$), reflecting the fact that equilibrium had not been reached. The plots for Langmuir, Freundlich, Temkin, and D-R are shown in Figure 5, whereas the Langmuir separation factor (R_L) is shown in Figure 6.

Adsorption Isotherms (q_e at $t=60$ min vs C_e ; dose=2 g/L, pH=3.9, $T=25^\circ\text{C}$)

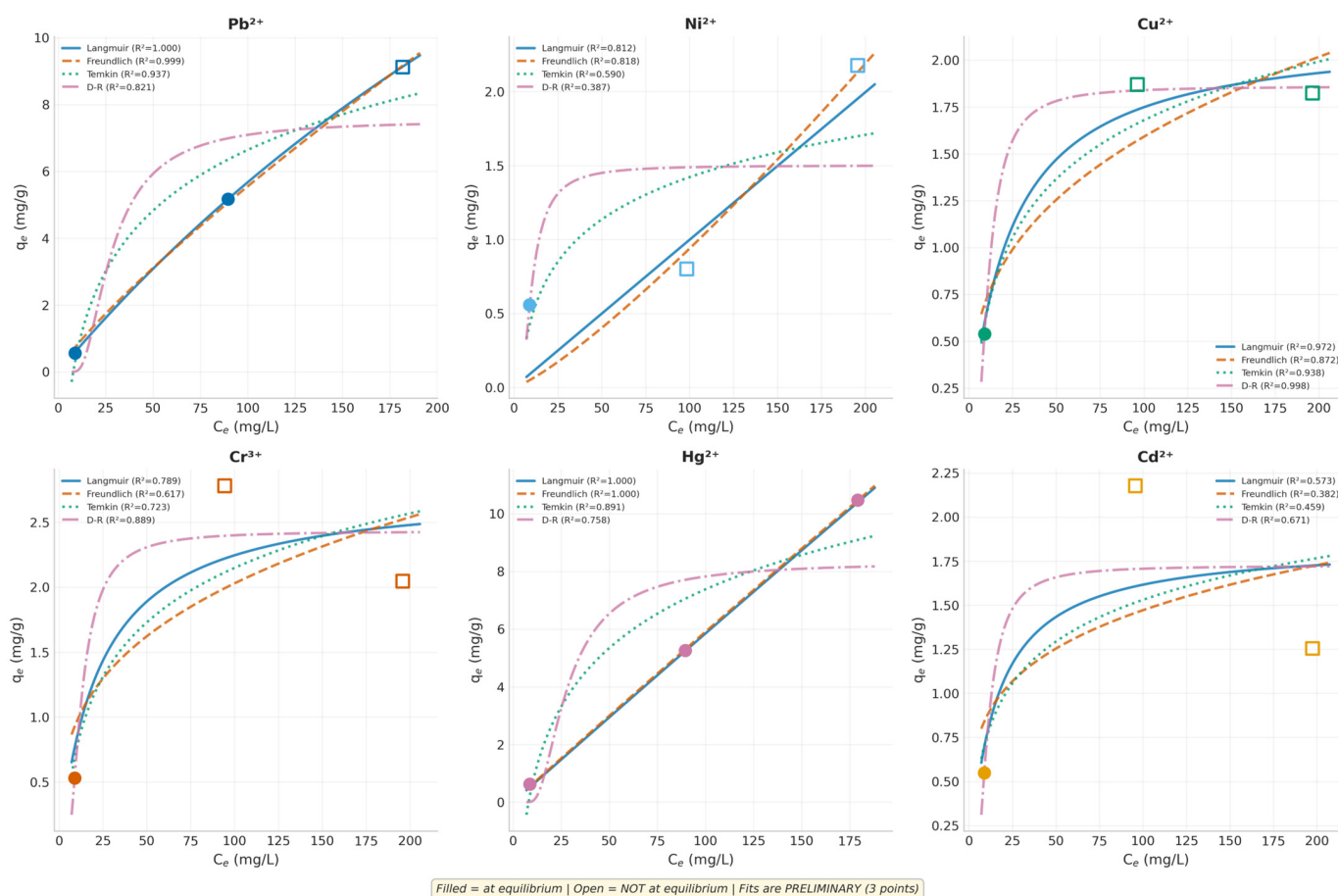


Figure 5. Isotherm plots (q_e vs. C_e) for all 6 metals with Langmuir, Freundlich, Temkin, and D-R model fits.

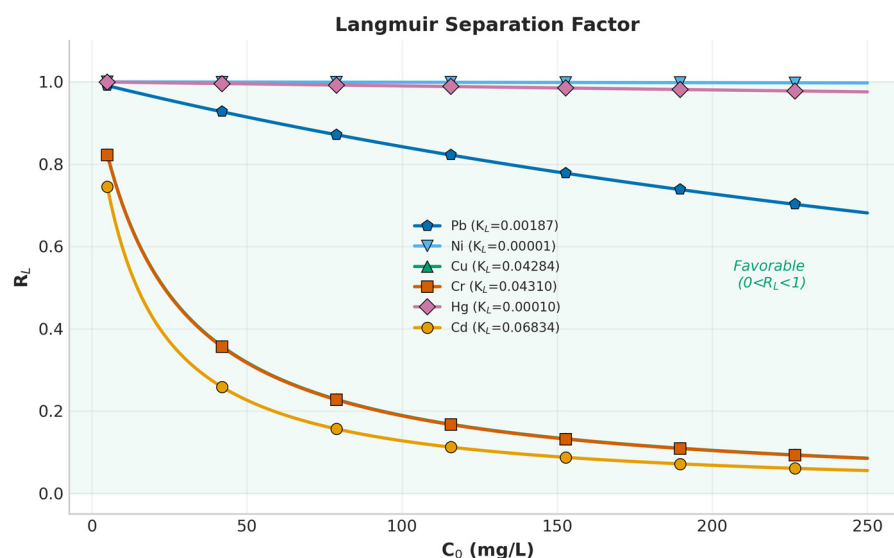


Figure 6. Langmuir separation factor R_L vs. C_0 .

The D-R mean free energy values fell below 8 kJ/mol for all metals, which formally classifies the mechanism as physisorption. However, this classification should be interpreted with extreme caution, given the three-point fitting limitation and the non-equilibrium status of four metals. The ion exchange mechanism established by the kinetic and diffusion analyses (Section 3.11) is far more robustly supported than the D-R energy classification, and the discrepancy is attributed to the inadequacy of the D-R model under these data constraints rather than to a genuine mechanistic contradiction.

3.6. Machine Learning Model Performance

The ML framework was applied to a dataset comprising 216 observations (six metals $\times$ three concentrations $\times$ 12 time points) organized into 18 kinetic curves, with 19 input features derived from metal physicochemical descriptors (eight), metal–adsorbent interaction terms (six), and experimental variables (five). Prior to any model development, the dataset was split at the group level into a development set (12 curves, 144 samples, 66.7%) and an external test set (six curves, 72 samples, 33.3%), with exactly one curve per metal held out (seed = 42). The external set was locked and used only once for the final performance evaluation.

3.6.1. Internal Cross-Validation (Development Set)

Five regression algorithms were evaluated under Repeated Five-Fold CV (5 folds $\times$ 10 repeats = 50 iterations) on the development set (Table 7). Gradient Boosting achieved the best performance ($R^2 = 0.979 \pm 0.043$, 95% CI [0.967, 0.991], MAE = 0.164 mg/g, RMSE = 0.325 mg/g), followed by Random Forest ($R^2 = 0.890 \pm 0.059$). Ridge, Elastic Net, and SVR achieved lower R^2 values (0.453–0.714), indicating that the nonlinear interactions between metal descriptors and experimental conditions are poorly captured by linear or kernel-based models under the current hyperparameter settings. Pairwise Wilcoxon signed-rank tests with Bonferroni correction confirmed that Gradient Boosting significantly outperformed all other models ($p < 0.001$ for all comparisons).

However, K-Fold CV permits different time points from the same curve to appear in training and test folds (interpolation). When the validation strategy was tightened to Group K-Fold, all models exhibited dramatically negative R^2 values (i.e., GB: -391.5 , RF: -7343.2), indicating that predicting an entirely unseen kinetic curve from within the 12-curve development set is extremely challenging given the limited number of groups. This

is not surprising because with only 12 development curves spanning diverse metals and concentrations, holding out two curves at a time ($\approx 17\%$) can create train–test distributional mismatches that produce arbitrarily poor predictions [102]. The LOMO results further confirm this: for most metals, R^2 was strongly negative, with only Pb achieving positive R^2 under RF and GB. This means that five metals do not contain enough diversity to predict the sixth metal, which is a finding that underscores the limitations of small datasets.

Table 7. Repeated K-Fold CV results ($5 \times 10 = 50$ folds) on the development set.

Model	R^2 Mean	R^2 Std	R^2 95% CI	MAE	RMSE	Gap *
Gradient Boosting	0.979	0.043	[0.967, 0.991]	0.164	0.325	0.021
SVR	0.956	0.059	[0.939, 0.974]	0.296	0.541	0.028
Random Forest	0.890	0.059	[0.873, 0.907]	0.631	0.913	0.081
Ridge	0.656	0.139	[0.616, 0.696]	1.312	1.615	0.084
Elastic Net	0.505	0.105	[0.475, 0.535]	1.476	1.984	0.073

Note: * Gap = Train R^2 – CV R^2 ; values < 0.05 indicate low overfitting risk.

The validation strategy comparison is shown in Figure 7. Figure 7a displays RF and GB R^2 across the three internal CV strategies, clearly illustrating the decline from K-Fold to GroupKFold to LOMO. The grouped bar chart compares R^2 scores of Random Forest and Gradient Boosting across three cross-validation strategies: standard K-Fold (interpolation, where random time points can be located in both train and test), GroupKFold (extrapolation, where entire Metal $\times C_0$ kinetic curves form indivisible groups), and Leave-One-Metal-Out (LOMO, the strictest test, where all data for one metal is held out). The red dashed line marks $R^2 = 0.80$ as a reference threshold. Figure 7b contrasts internal (CV) and external R^2 for all five models. The grouped bar chart comparing cross-validated R^2 (blue bars) against external test set R^2 (green bars) for all models having some predictive capability for unseen curves under the current data distribution, but limited extrapolation ability to unseen metals, shows similar bar heights, while large gaps between blue and green bars indicate overfitting to the development data. Figure 7c shows learning curves for RF and GB. The line plots for Random Forest and Gradient Boosting show training R^2 (filled circles, solid lines) and test R^2 (squares, dashed lines) as functions of training set size (20%, 40%, 60%, 80%, 100% of the development set). The shaded bands represent ± 1 standard deviation across folds. The learning curves for Random Forest and Gradient Boosting show that both training and CV test R^2 converge and plateau by approximately 80% of the development set (~ 115 samples), confirming that the dataset size is sufficient to capture the learnable signal.

3.6.2. External Validation (Held-Out Kinetic Curves)

Six kinetic curves (one per metal) were held out as the external test set comprising 72 samples (33%): Cd at $C_0 = 10$ mg/L, Cr at $C_0 = 100$ mg/L, Cu at $C_0 = 10$ mg/L, Hg at $C_0 = 100$ mg/L, Ni at $C_0 = 10$ mg/L, and Pb at $C_0 = 100$ mg/L. The remaining 12 curves (144 samples, 67%) formed the development set. Complete assignments are listed in Table 8. The split is fully reproducible from `numpy.random.RandomState` (seed = 42) applied to the alphabetically sorted metal list. The split procedure in detail was as follows: For each metal, three kinetic curves ($C_0 = 10, 100, 200$ mg/L) were randomly shuffled and the first was assigned to the external set. This guarantees (i) every metal appears in both sets, (ii) entire kinetic curves are held out (no time point leakage), and (iii) the split is reproducible from the random seed. The external set includes three curves at $C_0 = 10$ mg/L (Cd, Cu, Ni) and three curves at $C_0 = 100$ mg/L (Cr, Hg, Pb).

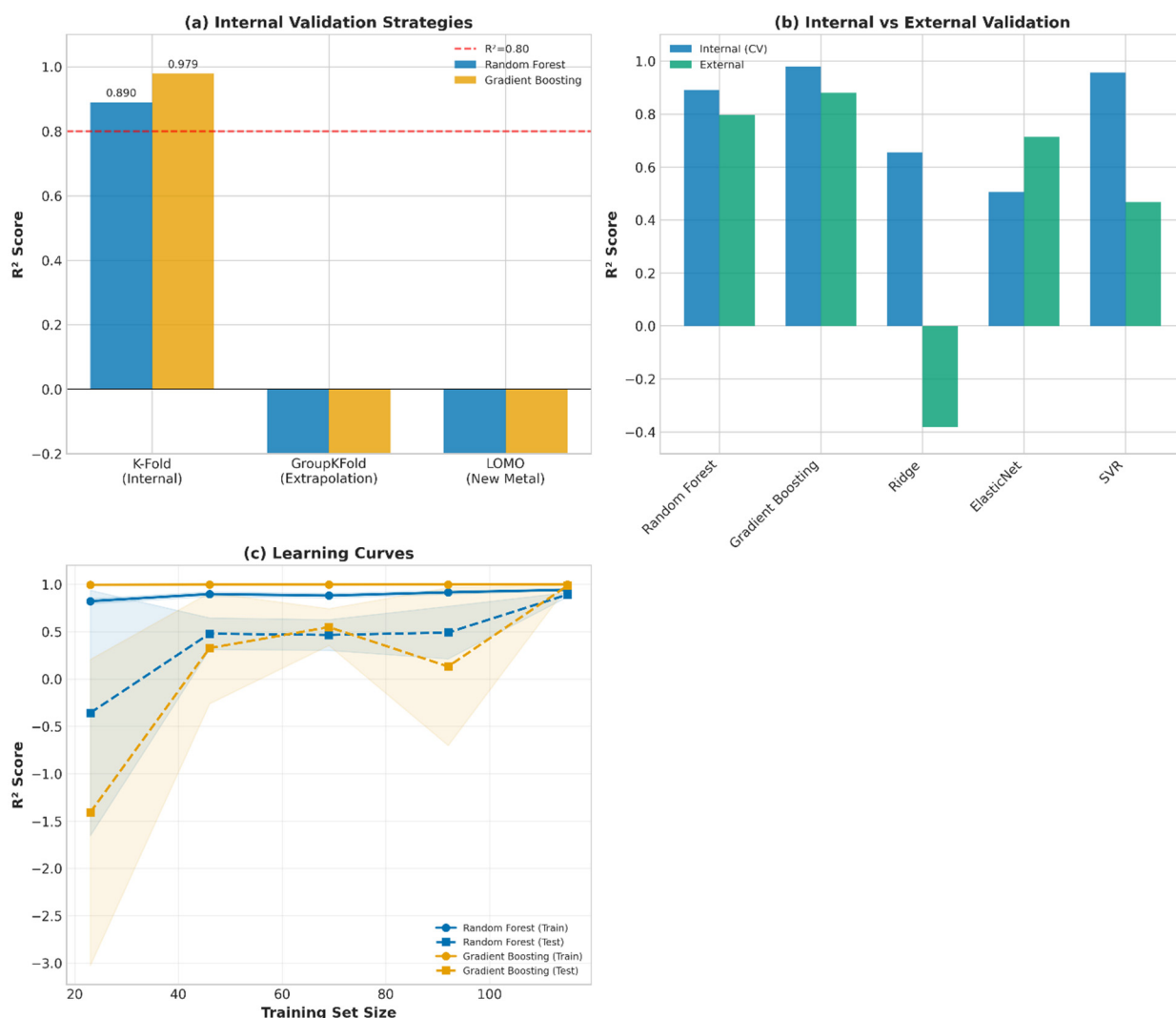


Figure 7. Validation of five ML models for predicting q_t (mg/g) on zeolite P. Dataset: A total of 216 observations split into development (12 curves, 144 samples) and external test (6 curves, 72 samples) sets; no time point from any external curve was seen during training. **(a)** R^2 across three CV strategies: Repeated K-Fold (interpolation), GroupKFold (unseen curves), LOMO (unseen metal). Dashed red = $R^2 = 0.80$. **(b)** Internal CV (blue) vs. external (green) R^2 . **(c)** Learning curves (circles/solid = train; squares/dashed = CV test; shaded = ± 1 SD).

Table 8. Complete listing of kinetic curve assignments for the group-level data split. One curve per metal (6 of 18 total, 33%) was randomly assigned to the external test set using a stratified random procedure (random seed = 42), ensuring every metal was represented in both sets. Each curve comprises 12 time points (1, 2, 4, 6, 8, 10, 15, 20, 25, 30, 45, 60 min). External test set (72 samples); development set (144 samples).

Metal Ion	C_0 , mg/L	Assigned Set	n	q_t Range, mg/g	Kinetic Behavior
Cd^{2+}	10	External	12	0.520–0.550	Fast equilibrium (<2 min)
Cr^{3+}	100	External	12	0.584–2.782	Slow, progressive uptake
Cu^{2+}	10	External	12	0.510–0.540	Fast equilibrium (<4 min)
Hg^{2+}	100	External	12	4.697–5.257	Near-equilibrium plateau
Ni^{2+}	10	External	12	0.540–0.560	Fast equilibrium (<4 min)
Pb^{2+}	100	External	12	4.492–5.169	Fast initial, slow plateau
Cd^{2+}	100	Development	12	0.612–2.178	Slow, progressive uptake
Cd^{2+}	200	Development	12	0.071–1.255	Slow, progressive uptake

Table 8. Cont.

Metal Ion	C ₀ , mg/L	Assigned Set	n	q _t Range, mg/g	Kinetic Behavior
Cr ³⁺	10	Development	12	0.510–0.530	Fast equilibrium (<2 min)
Cr ³⁺	200	Development	12	0.101–2.046	Slow, progressive uptake
Cu ²⁺	100	Development	12	0.596–1.870	Slow, progressive uptake
Cu ²⁺	200	Development	12	0.065–1.824	Slow, progressive uptake
Hg ²⁺	10	Development	12	0.610–0.630	Fast equilibrium (<2 min)
Hg ²⁺	200	Development	12	5.440–10.472	Fast initial, slow rise
Ni ²⁺	100	Development	12	0.004–0.801	Slow, progressive uptake
Ni ²⁺	200	Development	12	0.079–2.177	Slow, progressive uptake
Pb ²⁺	10	Development	12	0.530–0.560	Fast equilibrium (<2 min)
Pb ²⁺	200	Development	12	2.948–9.120	Two-stage uptake

Table 9 presents the external validation results. Gradient Boosting achieved the highest external $R^2 = 0.880$ (MAE = 0.477 mg/g, RMSE = 0.704 mg/g) followed by Random Forest with $R^2 = 0.797$. Critically, Ridge regression collapsed to $R^2 = -0.383$ on the external set despite a CV R^2 of 0.656, confirming severe overfitting for linear models on this nonlinear problem. SVR also degraded substantially (external $R^2 = 0.468$ vs CV $R^2 = 0.956$), suggesting that its RBF kernel memorized the development-set distribution. The gap for GB ($\Delta R^2 = \text{CV} - \text{External} = 0.099$) falls within the “Good” threshold (<0.10), indicating satisfactory predictive capability from training to truly unseen curves.

Table 9. External validation results on the held-out test set (6 complete kinetic curves, 72 samples).

Model	R ²	MAE	RMSE	Gap	Status
GB	0.880	0.477	0.704	0.099	Good
RF	0.797	0.658	0.915	0.093	Good
Elastic Net	0.714	0.988	1.087	−0.209	Excellent *
SVR	0.468	1.268	1.482	0.489	Overfitting
Ridge	−0.383	1.896	2.389	1.039	Overfitting

Note: * Elastic Net negative gap (external > CV) reflects its low-variance structure; CV underestimates it because KFold samples are small.

Figure 8 displays parity plots for the best model on both the development and external sets and for all models on the external set. In the development-set parity plot, GB predictions cluster tightly along the 1:1 line across all metals and concentrations. On the external set, the scatter increases particularly for high-q_t samples, but the overall trend remains realistic, with $R^2 = 0.880$. Notably, Pb at C₀ = 100 mg/L (one of the held-out curves) is predicted well at intermediate times but shows some under-prediction near equilibrium, likely because the model was not trained on any Pb curve at C₀ = 100 mg/L. These systematic deviations are consistent with the tree-based model’s tendency to regress toward the training mean at the distribution extremes [74].

The per-metal external breakdown reveals that the overall $R^2 = 0.880$ is driven by the range of q_t across metals: metals with higher q_t (Pb, Hg at C₀ = 100 mg/L) contribute more to the total variance. When computed per-metal (12 points each), R^2 values are negative for Cd, Cu, and Ni, whose held-out curves are at C₀ = 10 mg/L (very low q_t range, 0.51–0.56 mg/g), where prediction errors of even 0.1 mg/g produce negative R^2 . This is a dataset-structure artefact: for near-constant kinetic curves at low C₀, R^2 is ill-defined. The MAE for these metals is actually very small (<0.5 mg/g), confirming that the model absolute error is low even where R^2 is negative.

Figure 9 presents a comprehensive residual analysis of the GB model on the external test set. The residual histogram (Figure 9a) is approximately centered at zero

($\mu = -0.04$ mg/g) with a standard deviation of 0.70 mg/g, indicating no systematic bias. The residuals-against-predicted plot (Figure 9b) shows no strong heteroscedasticity (no funnel pattern), although residual variance is slightly larger at higher predicted q_t values, which is a common pattern in adsorption data where high-concentration curves span a wider q_t range. The Q-Q plot (Figure 9c) reveals an approximately normal residual distribution, with slight heavy tails at the extremes, which is characteristic of small-sample deviations. Systematic deviations at the tails (S-shape) suggest a non-Gaussian error distribution is taking place. The boxplot by metal (Figure 9d) shows that residual medians are close to zero for all metals exhibiting the widest interquartile ranges for Pb and Hg, corresponding to their higher q_t magnitudes on the held-out curves. The relative errors are comparable across all six metals, indicating that the model does not systematically favor or penalize any particular metal.

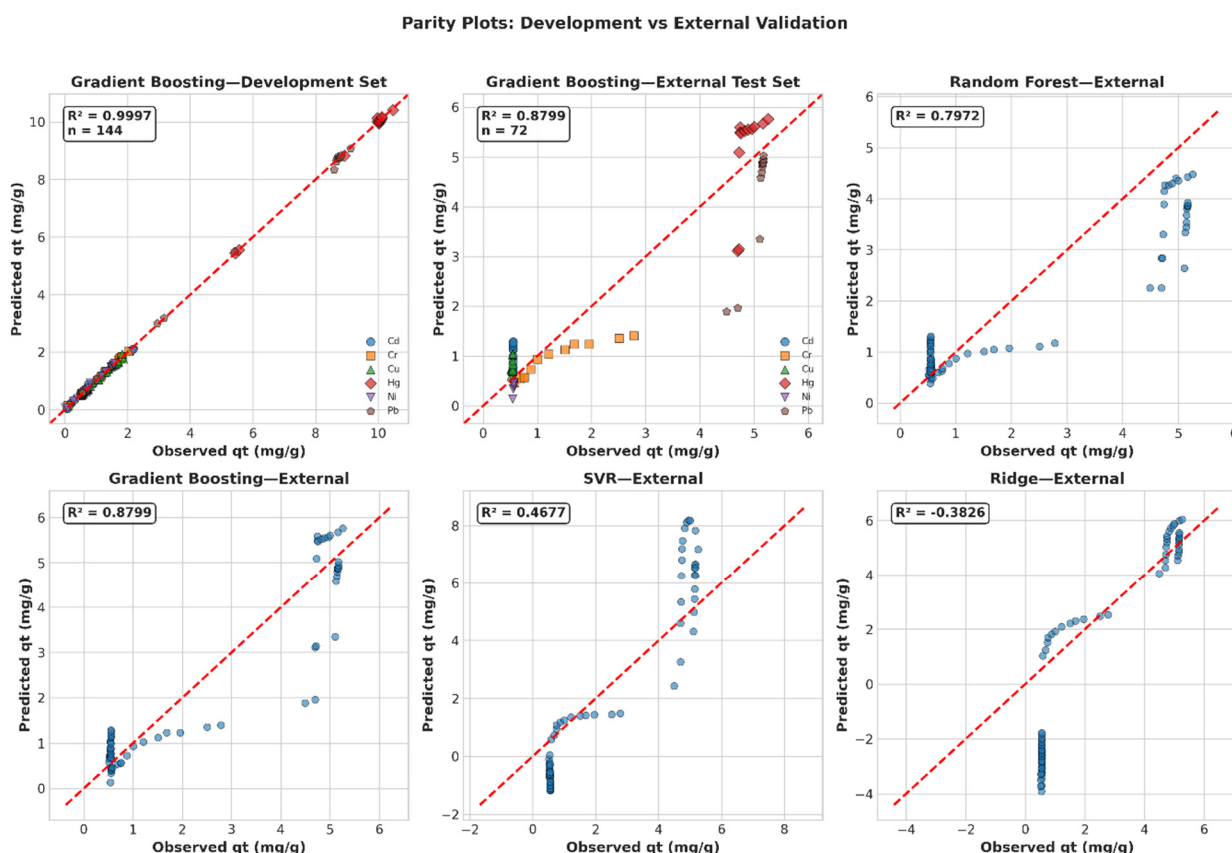


Figure 8. Parity plots (predicted vs. observed q_t). Red diagonal = 1:1 line; shaded band = ± 0.5 mg/g. Metal symbols consistent across all figures: Pb = brown pentagons; Hg = red diamonds; Cd = blue circles; Cu = green triangles; Ni = light purple triangles; Cr = orange squares. (**Top**) Gradient Boosting on development set (**left**, $n = 144$) and external set (**center**, $n = 72$). (**Bottom**) External set for GB, SVR, and Ridge. Insets report R^2 .

3.7. Kinetic Curve Predictions

Figure 10 depicts experimental kinetic data with GB predictions for all 18 Metal $\times$ C_0 curves, clearly distinguishing development-set curves (filled markers, solid lines) from external-set curves (open markers, dashed lines).

For the development-set curves, the model captures both the rapid initial uptake and the plateau region with high fidelity. For the external-set curves, the model correctly draws the qualitative kinetic shape (rapid initial rise followed by plateau) but may over- or under-predict the equilibrium loading by 0.3–1.2 mg/g. The most challenging external curves at $C_0 = 100$ mg/L are Pb (slightly under-predicted near equilibrium) and Hg (slight

over-prediction at early times). These discrepancies identify the Metal $\times$ C_0 regimes where additional training data would improve the model. Notably, the model reproduces the stagnation-to-re-acceleration transition for Cd^{2+} at $C_0 = 100$ mg/L (q_t jumping from 0.74 to 1.17 mg/g between $t = 8$ and 10 min) without being explicitly programmed with any kinetic model equations. This indicates that the tree-based splitting structure implicitly encodes the metal-specific transport transitions through conditional interactions between time, hydrated radius, and hydration enthalpy descriptors.

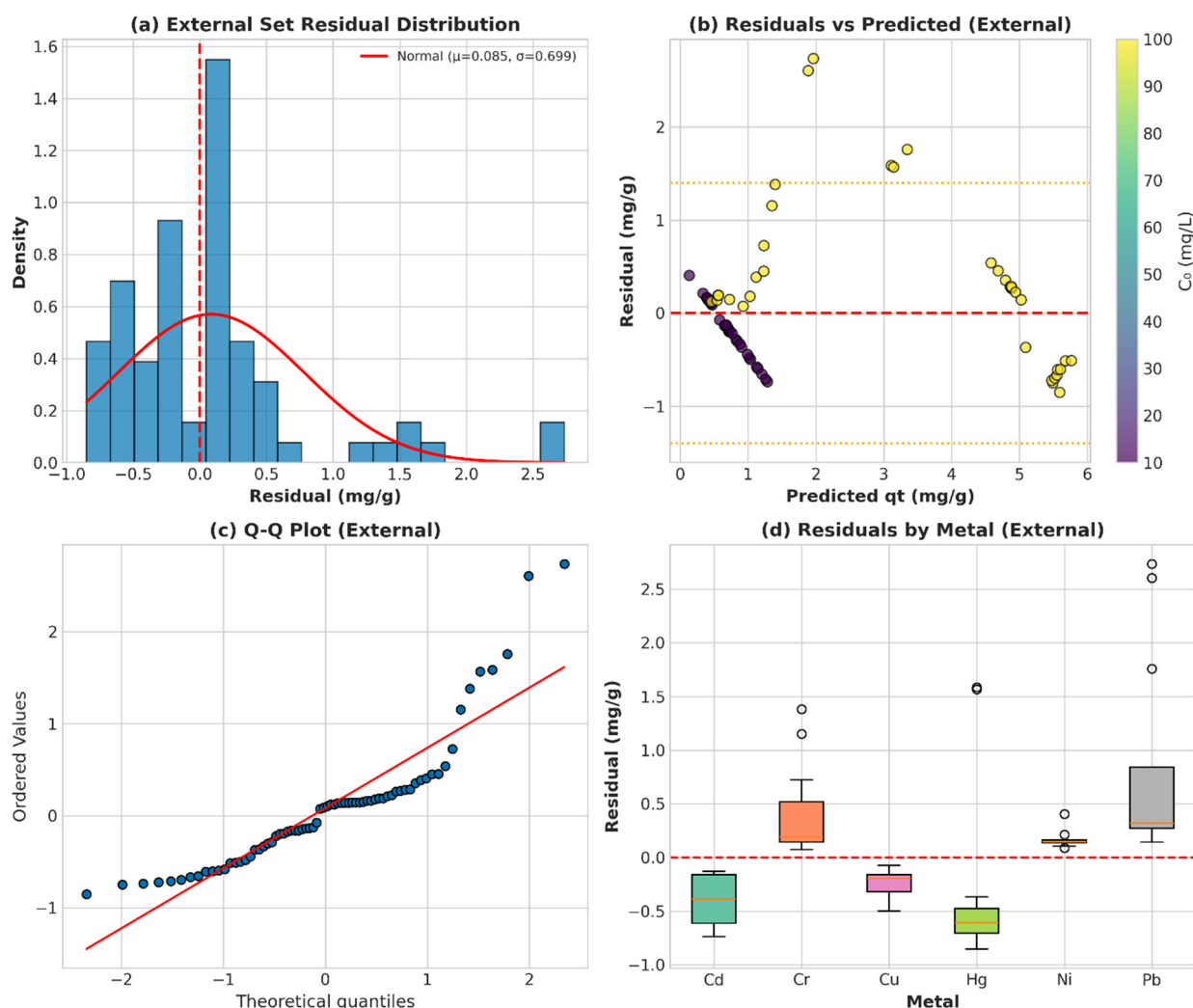


Figure 9. Residual analysis (observed–predicted q_t , mg/g) for Gradient Boosting on the external set ($n = 72$). (a) Histogram with Gaussian fit (red; μ , σ in legend). (b) Residuals vs. predicted; orange band = $\pm 2\sigma$. Metal colors consistent with all figures. (c) Normal Q-Q plot. (d) Box plots by metal (box = IQR; line = median; whiskers = $1.5 \times \text{IQR}$; dashed red = zero).

As a summary, traditional kinetic models (PFO, PSO, Elovich, Avrami) are fitted independently to each of the 18 curves, producing 18 separate parameter sets (36–54 total parameters) with no transferability: the PSO k_2 for Pb^{2+} at $C_0 = 100$ mg/L cannot predict Ni^{2+} at $C_0 = 200$ mg/L, nor Pb^{2+} at an untested concentration.

The ML model is a single unified predictor trained once on 12 curves that achieves three capabilities per-curve models cannot: (i) cross-curve prediction of six held-out kinetic curves (external $R^2 = 0.880$), including metals at concentrations absent from training; (ii) metal-descriptor encoding that, in principle, enables prediction for new metals from their physicochemical properties alone, though LOMO analysis demonstrates this currently fails with six training metals; and (iii) global feature importance via SHAP, which explains

why one metal is faster than another through descriptor rankings, whereas per-curve constants (k_2 , α , n) only describe the rate for each individual condition.

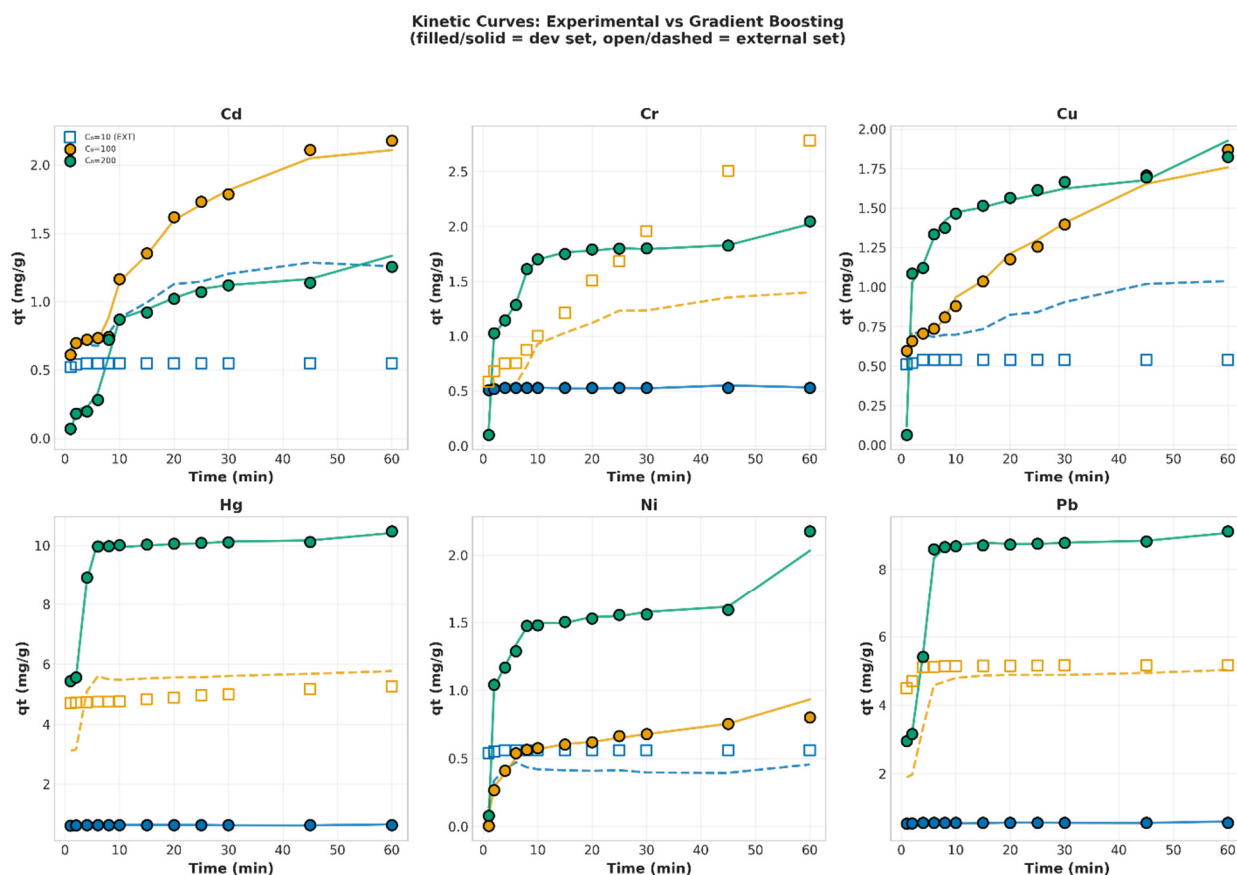


Figure 10. Experimental kinetics and Gradient Boosting predictions for six metals at $C_0 = 10$ mg/L (blue), 100 mg/L (orange), and 200 mg/L (green). pH 3.9, dose 2 g/L, and 25 °C. Filled circles/solid lines = development set; open markers/dashed lines = external test set (never seen during training). External curves: Cd at 10 mg/L, Cr at 100 mg/L, Cu at 10 mg/L, Hg at 100 mg/L, Ni at 10 mg/L, and Pb at 100 mg/L.

The two approaches are complementary, not competing. Per-curve models provide interpretable rate constants for each specific condition; the ML model provides cross-condition predictive capability and descriptor-level mechanistic insight. Table 10 shows a summary of both approaches.

Table 10. Summary of traditional kinetic models and the ML model.

Criterion	Per-Curve Models	Unified ML Model
Scope	1 curve (1 metal, 1 C_0)	All 18 curves simultaneously
Total parameters	36–54 (18×2 –3)	Fixed hyperparameters, trained once
Predict unseen C_0	No	Yes (external $R^2 = 0.880$)
Predict unseen metal	No	Limited (LOMO $R^2 < 0$)
Mechanistic insight	Rate constants per curve	Global descriptor importance (SHAP)

3.8. Leave-One-Metal-Out Validation

Leave-One-Metal-Out (LOMO) represents the strictest validation test: one metal is entirely excluded from training, and the model must predict its adsorption behavior based solely on the five remaining metals. This evaluates whether the physicochemical descriptor space provides sufficient coverage for cross-metal extrapolation. Under LOMO, all models

yielded strictly negative R^2 for every feature set configuration, indicating that predicting the adsorption behavior of a completely novel metal remains beyond the capacity of this six-metal dataset. The LOMO failure is particularly severe for Cd^{2+} ($R^2 = -85.21$ with the full feature set), reflecting its anomalous three-phase kinetic behavior and borderline pore accessibility that place it in a unique region of the descriptor space not adequately represented by the five remaining metals. Pb^{2+} yielded the only near-positive LOMO R^2 (0.93), likely because its high electronegativity ($\chi = 2.33$) and large ionic radius (1.19 Å) fall within a region partially interpolated by Hg^{2+} ($\chi = 2.00$, $r = 1.02$ Å). The LOMO limitation is due to the dataset size rather than attributable to a specific feature engineering choice. The LOMO analysis establishes a boundary condition: the model has predictive capability for unseen curves under the current data distribution but has limited extrapolation ability to unseen metals. Transparent failure reporting is essential for responsible ML practice and prevents readers from overestimating model applicability [75,76].

3.9. Ablation Study: Feature Set Contribution

To determine whether the physicochemical descriptors are essential or merely decorative, an ablation study was conducted by training three models (RF, GB, Ridge) on four progressively inclusive feature sets under multiple evaluation strategies (Table 11).

Table 11. Ablation study summary: R^2 for three models across four feature sets under Repeated K-Fold CV and external validation.

Feature Set	Strategy	RF R^2	GB R^2	Ridge R^2	Behavior
Experimental-Only	K-Fold	−0.057	−0.242	0.147	
Experimental-Only	External	−0.314	−0.310	0.063	
Descriptors-Only	K-Fold	0.165	0.169	0.176	
Descriptors-Only	External	0.891	0.907	0.878	
Metal Desc + Exp	K-Fold	0.899	0.979	0.641	
Metal Desc + Exp	External	0.748	0.596	−0.203	
Full (19 features)	K-Fold	0.890	0.979	0.656	Best
Full (19 features)	External	0.797	0.880	−0.383	Best

When trained on only the five experimental features (C_0 , t , $\log(1 + t)$, $\sqrt{t}$, $\log(1 + C_0)$), all models yielded negative or near-zero R^2 under Repeated K-Fold (GB: $R^2 = -0.242$; RF: $R^2 = -0.057$; Ridge: $R^2 = 0.147$) and on the external test set (GB: $R^2 = -0.310$; RF: $R^2 = -0.314$). Without metal-specific descriptors, the model cannot distinguish between metals and predicts a single average kinetic curve for all six, which fits none. This is the strongest evidence that the physicochemical descriptors are essential, not auxiliary, to predictive performance. A model that knows the concentration and time but not the identity of the metal has no basis for predicting that Hg^{2+} will adsorb ten times more than Cd^{2+} at $C_0 = 200$ mg/L.

The Metal Descriptors + Experimental feature set (13 features, excluding all six interaction terms) achieved $R^2 = 0.979$ under Repeated K-Fold for Gradient Boosting, identical to the Full feature set (19 features, $R^2 = 0.979$, $\Delta R^2 < 0.001$). This confirms that tree-based models learn relevant metal–adsorbent interactions implicitly through their conditional splitting structure, without requiring pre-computed interaction terms. The six interaction descriptors (pore accessibility, ion exchange potential, electrostatic factor, HSAB match, dehydration penalty, surface capacity) are retained in the manuscript for mechanistic interpretability, as they enable direct mapping to ion exchange theory [67,68], but are acknowledged as redundant for prediction.

A noteworthy result is that the Descriptors-Only configuration (14 features, no time or concentration) achieved external $R^2 = 0.907$ for Gradient Boosting, substantially out-

performing the Experimental-Only configuration ($R^2 = -0.310$). This means that metal physicochemical properties alone carry more predictive information than the experimental variables alone. The descriptors encode the metal-specific exchange behavior that distinguishes one curve from another, while the experimental variables (C_0 , t) encode only the shared temporal shape common to all curves.

Wilcoxon signed-rank tests on the 50 paired K-Fold R^2 scores (Table 12) confirmed that the Full feature set significantly outperformed the Experimental-Only set for all three models ($p = 1.78 \times 10^{-15}$ for RF, GB, and Ridge). The effect sizes ($\Delta R^2 = +0.509$ to $+1.221$) are large by any standard, confirming that the performance difference is not attributable to random CV fold variation.

Table 12. Wilcoxon signed-rank tests comparing Full feature set vs. Experimental-Only under Repeated 5-Fold CV (50 paired R^2 scores). * indicates $p < 0.001$.

Model	Full R^2 (50 Folds)	Exp-Only R^2	ΔR^2	Wilcoxon p
Random Forest	0.890 ± 0.059	-0.057 ± 0.298	$+0.947$	$1.78 \times 10^{-15} *$
Gradient Boosting	0.979 ± 0.043	-0.242 ± 0.496	$+1.221$	$1.78 \times 10^{-15} *$
Ridge	0.656 ± 0.139	0.147 ± 0.186	$+0.509$	$1.78 \times 10^{-15} *$

Figure 11 shows the ablation study evaluated for four feature sets under three validation strategies of increasing stringency, revealing how each feature group contributes to performance and where the model fails.

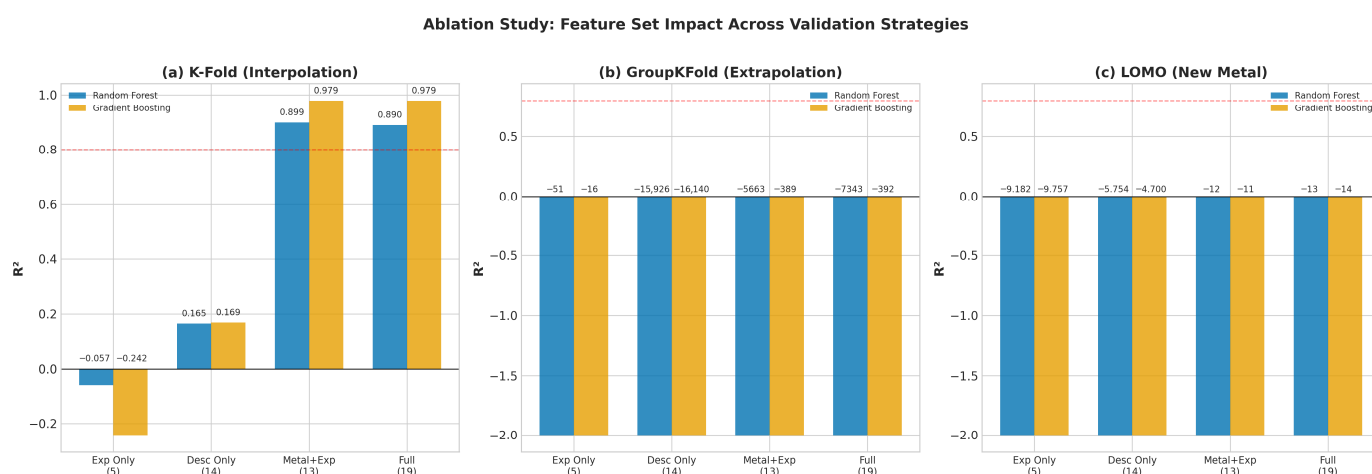


Figure 11. Ablation study: (a) R^2 comparison across four feature sets for three models; (b,c) external parity plot for Full vs. Experimental-Only.

When the model is tested on randomly held-out samples from the same kinetic curves it was trained on, the progression is clear (Figure 11a). Experimental-Only (C_0 , t , 5 features) yields negative R^2 for both RF (-0.057) and GB (-0.242), meaning the model predicts worse than simply guessing the mean q_t . This is the most important result of the ablation: without knowing which metal is present, the model cannot function, i.e., the same C_0 and t produce $q_t = 0.56$ for Pb at $C_0 = 10$ mg/L but $q_t = 9.12$ for Pb at $C_0 = 200$ mg/L. Adding Descriptors-Only (14 features, no C_0 or t) recovers modest performance ($R^2 \approx 0.17$), confirming that metal identity carries signal, but the model cannot trace kinetic curves without time. Metal + Experimental (13 features) jumps to $R^2 = 0.979$ for GB and 0.899 for RF, essentially matching the Full model (19 features, $R^2 = 0.979$ for GB, 0.890 for RF). The negligible difference between 13 and 19 features ($\Delta R^2 < 0.001$ for GB) proves that the six interaction terms are redundant for tree-based models, which reconstruct multiplicative interactions internally through sequential splits.

If entire kinetic curves are held out, all feature sets collapse to catastrophically negative R^2 values (ranging from -16 to $-16,140$), as shown in Figure 11b. This is not a failure of the feature engineering, as it reveals that the model memorizes curve-specific patterns during interpolation (Figure 11a) but cannot extrapolate to unseen Metal $\times$ C_0 combinations using GroupKFold alone. The extreme negative values arise because GroupKFold with only 12 development curves creates folds where entire metals may be absent from training, approximating the LOMO condition. This result is consistent with the small dataset size (12 curves) and the discrete nature of the metal identity descriptors.

When an entire metal is held out, all feature sets yield R^2 between -4.7 and -9.8 , confirming that six metals are insufficient for cross-metal generalization regardless of which features are provided (Figure 11c). Notably, the Full model ($R^2 \approx -13$ to -14) performs no better than Experimental-Only ($R^2 \approx -9$ to -10), demonstrating that even physics-informed descriptors cannot compensate for the absence of a metal from the training distribution when only six metals are available. This is a boundary condition, not a flaw: it establishes that expanding to 15–20 metals is required before cross-metal prediction becomes feasible.

3.10. SHAP-Based Model Interpretation

SHAP (SHapley Additive exPlanations) was employed as a model interpretability tool to evaluate whether the Gradient Boosting model-learned feature importance is consistent with, not proof of, established ion exchange selectivity theory [42]. The correlative nature of SHAP analysis is explicitly acknowledged: SHAP values quantify each feature's contribution to the model predictions, not to the physical system causality.

3.10.1. Global Feature Importance

The global SHAP importance ranking (Figure 12), measured by mean $|\text{SHAP}|$ across all 144 development set predictions, identified initial concentration (C_0) as the dominant predictor overall. This is expected from mass action kinetics. Among the metal-level descriptors, ionic radius emerged as the most important, followed by ion exchange potential, charge density, and electronegativity.

The surface capacity feature (which encodes BET surface area) contributed minimally, consistent with the low BET surface area ($30 \text{ m}^2/\text{g}$) and the conclusion from the traditional analysis that surface physisorption is not the primary mechanism. This value reflects the restricted access of the N_2 probe molecule (kinetic diameter $3.64 \text{ \AA}$) to the internal micropore system of the GIS framework, whose eight-membered ring apertures ($3.1 \times 4.5 \text{ \AA}$) are partially obstructed by extraframework cations at low Si/Al ratios [72]. However, heavy metal removal by low-silica zeolites proceeds predominantly through ion exchange within the crystalline lattice rather than surface physisorption [100,103]. The high framework charge density (Si/Al = 1.63) ensures abundant exchangeable sites that are not captured by BET measurements. The concentration-dependent kinetic profiles observed in this work are consistent with this mechanism: at $C_0 = 10 \text{ mg/L}$, equilibrium was reached within 2–4 min (exchange at readily accessible external and near-surface sites), whereas at $C_0 = 100\text{--}200 \text{ mg/L}$, progressive uptake over 60 min indicates rate limitation by intraparticle diffusion through the narrow GIS channels [61,104]. The SHAP analysis independently supports this interpretation, as hydrated radius and dehydration penalty (descriptors encoding the difficulty of entering narrow pores) ranked among the most important predictors, while the surface-capacity feature contributed minimally. Low BET surface areas ($20\text{--}60 \text{ m}^2/\text{g}$) are consistently reported for CFA-derived GIS-type zeolites that nonetheless achieve effective heavy metal removal [3,105], confirming that this is a structural characteristic of the framework type rather than a synthesis limitation.

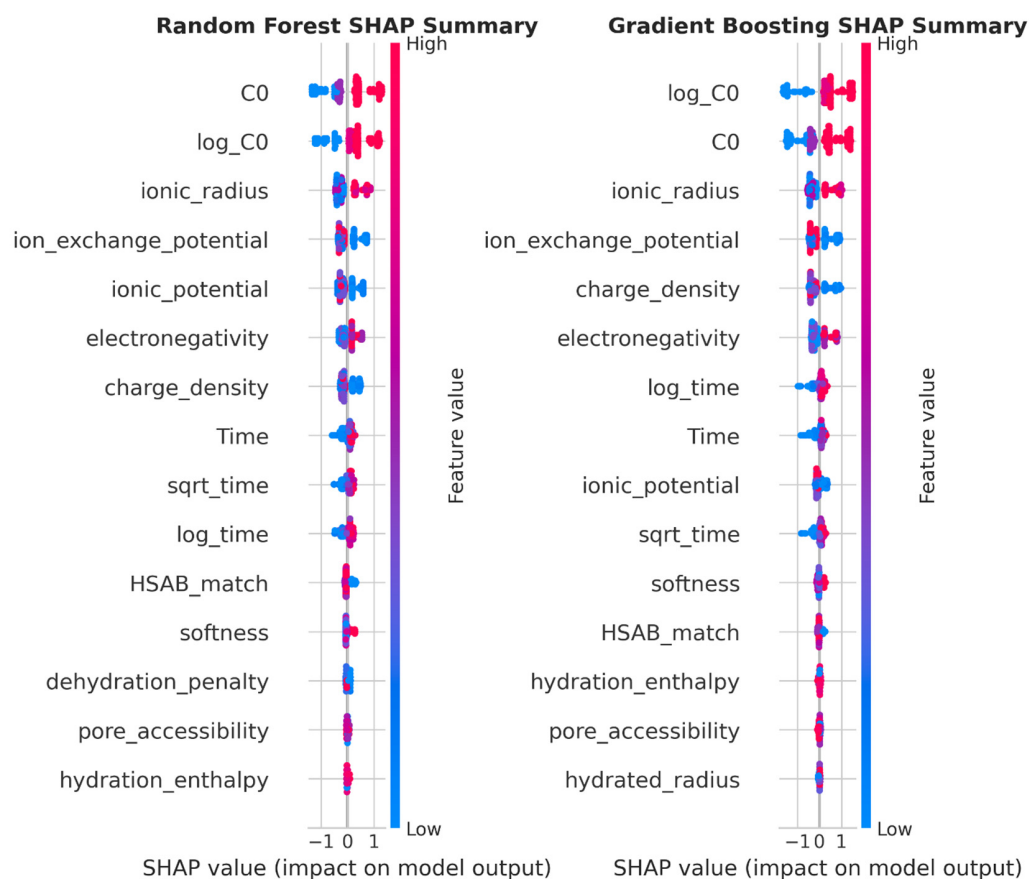


Figure 12. SHAP summary (beeswarm) plots: (left) Random Forest; (right) Gradient Boosting. Red color indicates high feature values, while blue color indicates low ones.

Figure 13 presents three complementary feature importance analyses as follows: a) Mean Decrease Impurity or MDI (Figure 13a,b) permutation importance (Figure 13b), and c) SHAP importance (Figure 13c). Figure 13a shows a horizontal bar chart of the top 12 features ranked by Gini impurity-based importance, with side-by-side bars for Random Forest (blue) and Gradient Boosting (orange). MDI is the measure of how much each feature reduces the weighted impurity across all tree splits. Note that MDI tends to favor high-cardinality continuous features and can be biased when features have different scales or cardinalities. Figure 13b plots a horizontal bar chart of the top 12 features ranked by permutation importance (green bars with error bars showing ± 1 standard deviation from 20 randomization repeats). Permutation importance measures the decrease in model performance when a feature's values are randomly shuffled, making it model-agnostic and unbiased toward feature scale. The error bars reflect the variability in the importance estimate. Finally, Figure 13c shows a horizontal bar chart of the top 12 features ranked by mean absolute SHAP value, with side-by-side bars for RF and GB. SHAP provides theoretically grounding based on cooperative game theory. Comparing all three subplots (MDI, permutation, SHAP) reveals whether the feature ranking is consistent across importance metrics, which strengthens interpretability claims.

All three methods converge on the same top-tier features: initial concentration C_0 and its log-transform are the most important, followed by ionic radius, ion exchange potential, ionic potential, and electronegativity. Time-related features (time, $\sqrt{t}$, $\log(1 + t)$) rank lower, which is physically consistent: once the metal identity and C_0 are specified, the shape of the kinetic curve is mostly determined, with time acting as a within-curve interpolator.

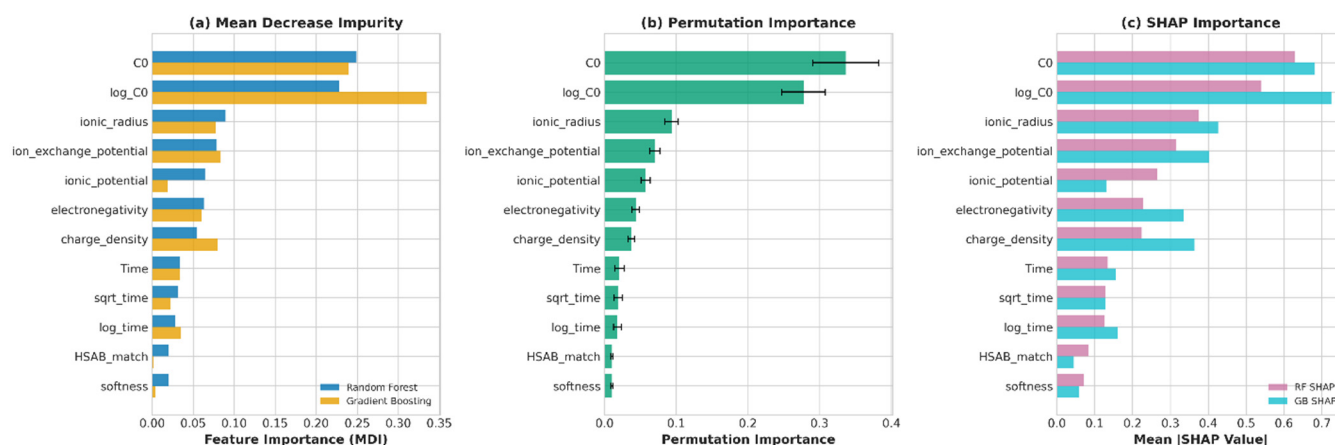


Figure 13. Feature importance comparison: (a) Mean Decrease Impurity for RF and GB; (b) permutation importance; (c) mean |SHAP| values.

3.10.2. Convergence with Eisenman–Sherry Selectivity Theory

The Gradient Boosting model was never informed of ion exchange theory, yet its SHAP-derived feature rankings independently reproduce four predictions of the Eisenman–Sherry selectivity theory [67,68]:

(i) Ionic radius emerges as the dominant metal-level predictor, consistent with the size-dependent balance between Coulombic attraction and the dehydration penalty that governs selectivity in low-silica zeolites. The SHAP beeswarm plot shows that large ionic radius values (Pb^{2+} : 1.19 Å, Hg^{2+} : 1.02 Å) receive positive SHAP contributions (moving q_t upward), while small radii (Cr^{3+} : 0.615 Å, Ni^{2+} : 0.69 Å) receive negative contributions. This matches the classical prediction that larger cations with lower charge density are preferentially exchanged in zeolites with the GIS framework topology [68,69].

(ii) Hydration enthalpy shows negative SHAP contributions at extreme values. Cr^{3+} , with $|\Delta H_{\text{hyd}}| = 4402 \text{ kJ/mol}$ (more than double any divalent metal), consistently receives the most negative SHAP values for this feature, meaning the model has learned that very high dehydration costs suppress uptake. This matches the dehydration-gating mechanism in narrow-pore zeolites [43,61], independently confirmed by the Weber–Morris three-phase diffusion analysis presented in Section 3.4.1.

(iii) Softness contributes positively for metals with values near the estimated framework hardness ($\sigma \approx 0.35$). Metals with intermediate-to-high softness (Hg^{2+} : $\sigma = 0.96$; Pb^{2+} : $\sigma = 0.58$; Cd^{2+} : $\sigma = 0.69$) receive positive SHAP contributions, while the hardest metal (Cr^{3+} : $\sigma = 0.08$) receives negative contributions. This is consistent with HSAB theory [41], which predicts that the partially soft framework oxygen sites at Si–O–Al bridges favor softer cations.

(iv) The SHAP interaction between time and ionic radius reveals that the model has learned the concentration-dependent kinetic regimes: for metals with large ionic radii (Pb^{2+} , Hg^{2+}), the SHAP contribution of time diminishes after $t \approx 4 \text{ min}$ (fast equilibrium), whereas for metals with high hydration enthalpy (Cr^{3+} , Cd^{2+}), time continues to contribute positively through $t = 60 \text{ min}$ (slow, progressive uptake not yet at equilibrium). This conditional behavior maps precisely onto the three-phase diffusion kinetics confirmed by the Weber–Morris analysis and the metal-specific behavioral groups described in Section 3.4.1.

3.10.3. Counterfactual Validation via the Ablation Study

SHAP analysis alone cannot answer counterfactual questions (“what happens if we remove these features?”). The ablation study (Section 3.9) addresses this directly: removing the physicochemical descriptors that SHAP ranks as important causes all models to yield

negative R^2 ($p = 1.78 \times 10^{-15}$), confirming that these features are necessary for prediction, not merely correlated with the target. Conversely, removing the interaction descriptors that SHAP ranks as unimportant produces no performance change ($\Delta R^2 < 0.001$), validating the SHAP low-importance assignments. This bidirectional consistency between SHAP importance and ablation performance provides stronger evidence than either method alone.

3.10.4. Limitations of SHAP Interpretation

Several limitations must be acknowledged. SHAP cannot distinguish between correlated descriptors: if ionic radius and ionic potential (Z/r) are highly correlated, SHAP may distribute importance between them in ways that do not reflect their independent mechanistic contributions. Readers should therefore interpret SHAP results at the level of descriptor groups (size-related, charge-related, hydration-related) rather than individual features. SHAP cannot establish causal directionality: the observation that ionic radius has high SHAP importance does not prove that ionic radius causes selectivity, only that the model uses it as a strong predictor. SHAP is therefore presented as one element of a triangulation strategy, alongside the Weber–Morris diffusion analysis, the Boyd film diffusion model, the ablation study, and convergence with classical selectivity theory, rather than as a standalone mechanistic proof.

As a summary, Figure 14 embeds physical interpretation with the ML results. Figure 14a plots each metal in the softness–ionic radius descriptor space, with marker size proportional to the maximum equilibrium loading (q_e at $t = 60$ min, $C_0 = 200$ mg/L). The highest loadings correspond to Hg^{2+} and Pb^{2+} , both large and soft ions, while the lowest corresponds to Cd^{2+} and Cu^{2+} . Figure 14b presents the full selectivity sequence: $Hg^{2+} > Pb^{2+} > Ni^{2+} > Cr^{3+} > Cu^{2+} > Cd^{2+}$ at $C_0 = 200$ mg/L. This order is broadly consistent with the Eisenman–Sherry selectivity theory for cation exchangers, where a combination of ionic radius, charge, and dehydration energy rules preference [41,67]. Thus, visualization connects the ML predictions to HSAB theory: softer, larger ions (Pb^{2+} , Hg^{2+}) tend to cluster differently from harder, smaller ions (Cr^{3+} , Ni^{2+}), and the bubble sizes reveal whether the selectivity sequence correlates with position in the properties space.

Figure 14c displays the CV R^2 , external R^2 , and gap side by side for each model, which provides a visual summary of model selection. Gradient Boosting has both the highest external R^2 and a moderate gap, which constitutes the best trade-off between accuracy and predictive capability. A summary of the results is shown as a subplot panel.

3.11. Ion Exchange Mechanism Identification

Above the zero-point charge, the negatively charged surface of zeolites attracts M^{2+} ions by electrostatic forces, resulting in a rapid rate of metal removal during the initial contact time between zeolite and metal ions [106]. However, some changes were observed by spectroscopic techniques. For example, the FTIR spectrum of the spent zeolite P (zeolite P that is characterized after the adsorption tests) in Figure A4 shows lower transmittance in the bands that can be attributed to the presence of metal bonds. The bands located between 740 and 790 cm^{-1} correspond to asymmetric and symmetric stretching vibrations of the Si–O–Si and Si–O–Al bonds, as well as the band between 500 and 420 cm^{-1} . Signals between 670 and 660 cm^{-1} are assigned to symmetric stretching vibration of Al–O and Si–O groups, whereas those ranging from 670 to 570 cm^{-1} are allocated to bending vibrations of Al–O, Si–O, and Al–O–Si groups. All these vibrations are assigned to the T–O bending mode [107].

XPS survey spectra (Figure 15), normalized prior to comparison, confirmed the elemental composition of the fly ash precursor, fresh zeolite P, and spent zeolite P (Hg , $C_0 = 200$ mg/L). The fly ash spectrum showed a sharp peak at ~530 eV (attributed to Fe or Ca oxides) with a poorly defined Na signal, whereas the fresh zeolite P spectrum exhibited

a clear Na 1s signal (~ 1071 eV) and sharper Si/Al peaks, confirming successful hydrothermal conversion into the crystalline GIS framework with Na^+ as the charge-compensating cation. After adsorption, the Na 1s signal was qualitatively reduced in intensity, and signals attributable to adsorbed heavy metals appeared (expected binding energies: Pb 4f ~ 138 eV, Hg 4f ~ 101 eV, Cd 3d ~ 405 eV, Cu 2p ~ 933 eV, Ni 2p ~ 855 eV, Cr 2p ~ 577 eV), consistent with the partial replacement of extraframework Na^+ by metal cations. The Si 2p and Al 2p framework peaks (~ 75 – 100 eV) were preserved in the spent material, confirming that the tetrahedral Si–O–Al framework remained structurally intact during adsorption, a hallmark of ion exchange rather than framework dissolution or surface precipitation. The O 1s region (~ 530 eV), dominant in all three spectra, may exhibit slight shifts reflecting changes in the oxygen coordination environment as Na–O–framework bonds are replaced by M^{2+} –O–framework bonds. The Si–OH bond is also found at 533.85–533.87 eV [108], while Al_2O_3 , present in the raw material, is observed in the Al 2p peak [109].

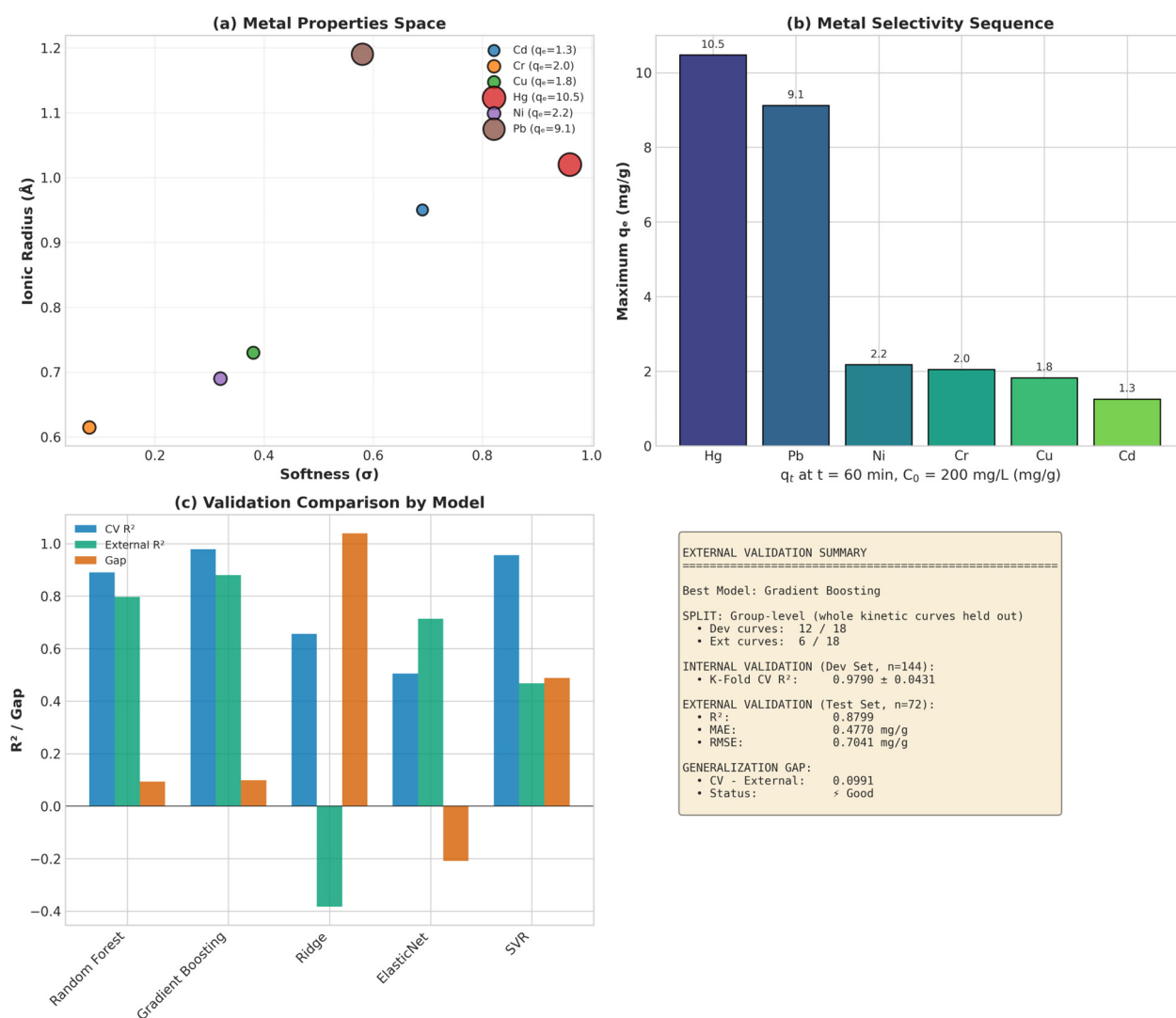


Figure 14. Physical interpretation and summary. (a) Softness (σ) vs. ionic radius (Å); bubble size for q_t at $C_0 = 200$ mg/L; dotted line = zeolite P hardness (≈ 0.35). (b) Selectivity at $C_0 = 200$ mg/L, $t = 60$ min; values for Ni, Cr, Cu, and Cd are apparent (equilibrium not reached). (c) Gap: CV R^2 (blue) – external R^2 (green); orange = gap (excellent) < 0.05 ; (good) $= 0.05$ – 0.10 ; (poor) ≥ 0.10 .

It should be noted that the XPS data are limited to survey spectra; high-resolution scans with peak fitting for quantitative determination of Na^+ depletion stoichiometry, metal oxidation states, and bonding environments were not performed.

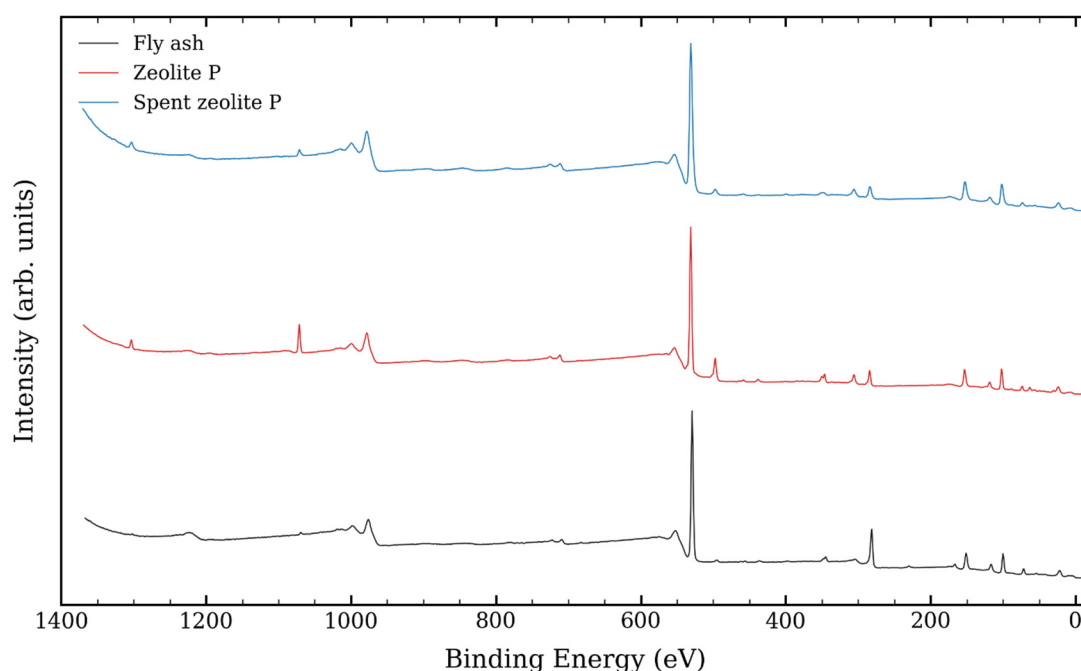


Figure 15. Survey XPS spectra of fly ash, zeolite P, and spent zeolite.

The dominant removal mechanism was identified as ion exchange based on the convergence of independent lines of evidence. XPS survey spectra showed a reduction in the Na 1s signal in spent zeolite P relative to the fresh material, accompanied by the appearance of heavy metal signals, consistent with the partial replacement of extraframework Na^+ by metal cations from solution. The preservation of Si 2p and Al 2p framework peaks confirms that the crystalline structure remained intact [110]. This spectroscopic evidence converges with: (i) PSO kinetics indicative of chemisorption/exchange [53]; (ii) multilinear Weber–Morris plots confirming multi-stage intraparticle diffusion through microporous channels; (iii) Boyd analysis showing concentration-dependent transition from film to intraparticle diffusion control; (iv) effective diffusion coefficients ($D_e \approx 10^{-10}$ – 10^{-9} cm^2/s) characteristic of intracrystalline zeolitic exchange [61]; (v) SHAP-derived feature importance dominated by ion exchange-relevant descriptors (ionic radius, hydration enthalpy, charge density) rather than surface area-dependent features; and (vi) a selectivity sequence at $C_0 = 200$ mg/L ($\text{Hg}^{2+} > \text{Pb}^{2+} > \text{Ni}^{2+} > \text{Cr}^{3+} > \text{Cu}^{2+} > \text{Cd}^{2+}$) that correlates with electronegativity ($r = +0.76$), consistent with the Eisenman–Sherry ion exchange selectivity theory for low-silica zeolites [67,68]. The low BET surface area (30 m^2/g) and minimal SHAP contribution of the surface capacity feature argue against surface physisorption as the primary pathway, while the acidic experimental pH (3.9) and absence of new crystalline phases in the spent material exclude surface precipitation as a major mechanism. High-resolution XPS with peak deconvolution is recommended as a priority for future work to confirm the $\text{Na}^+/\text{M}^{2+}$ exchange stoichiometry and metal binding environments.

The behavior of Cd^{2+} and Cr^{3+} merits a profound analysis. The three-phase kinetic profile observed for Cd^{2+} and Cr^{3+} is characteristic of multi-stage transport through microporous adsorbents and is well described by the Weber–Morris intraparticle diffusion model [57]. Multilinearity in the q_t -versus- $\sqrt{t}$ plot confirms that adsorption proceeds through at least two sequential rate-controlling steps [58]. Phase I corresponds to external film diffusion and exchange at readily accessible surface sites; Phase II reflects a dehydration bottleneck at the narrow GIS eight-ring channel apertures, where metal ions with hydrated radii exceeding the channel dimension (Cd^{2+} : 4.26 Å; Cr^{3+} : 4.61 Å) must undergo energy-intensive partial dehydration ($\Delta H_{\text{hyd}} = -1807$ and -4402 kJ/mol, respectively)

before entering the pore system [40,61]; Phase III represents intraparticle diffusion into the interior framework, triggered once the first dehydrated ions successfully transit the pore neck and vacate the bottleneck positions for subsequent ions [43]. The extended stagnation of Cr^{3+} relative to Cd^{2+} is consistent with its kinetic inertness. The water exchange rate of $[\text{Cr}(\text{H}_2\text{O})_6]^{3+}$ ($\sim 2.4 \times 10^{-6} \text{ s}^{-1}$) is approximately six orders of magnitude slower than for divalent transition metals [98]. This pattern is absent for Pb^{2+} and Hg^{2+} , whose lower hydration enthalpies and more favorable dehydration kinetics permit rapid, near-monotonic uptake without a pore-mouth bottleneck.

The Weber–Morris analysis confirms this quantitatively: Cd^{2+} and Cr^{3+} show multilinear q_t -vs.- $\sqrt{t}$ plots with $k_{id2} > k_{id1}$ (Cd at $C_0 = 100 \text{ mg/L}$: $k_{id1} = 0.105$, $k_{id2} = 0.480 \text{ mg/g}\cdot\text{min}^{-0.5}$), whereas Pb^{2+} shows monotonically decreasing k_{id} values consistent with the classical pattern. Similar multi-stage behavior has been reported for microporous zeolites by Sprynskyy et al. [96] and is theoretically predicted when channel dimensions approach ion sizes [43].

3.12. Practical Implications and Limitations

The external R^2 of 0.8799 obtained here under a leakage-free group-level split is notably lower than $R^2 > 0.95$. This gap does not indicate poorer modeling; rather, it reflects the removal of interpolation leakage. When individual time points from the same curve appear in both train and test, the model benefits from autocorrelation within the curve, producing R^2 values that overstate the predictive capability. Our results demonstrate that once this artefact is removed, the honest performance of a tree-ensemble model on a six-metal, 18-curve dataset is ~ 0.88 , which is still useful for screening and semi-quantitative prediction but not sufficiently precise for engineering design without additional data. We strongly recommend that future ML adsorption studies adopt group-level (curve-level) external splits and report Group K-Fold and LOMO results, alongside K-Fold, so that readers may distinguish interpolation from extrapolation performance [111–114].

However, the following limitations of this work persist, and they must be considered by readers:

Dataset size and metal coverage. Six metals and three concentrations yield only 216 observations (18 unique conditions). While sufficient for within-distribution prediction (external $R^2 = 0.880$), the dataset is at the lower boundary for tree-based models. LOMO validation yields strictly negative R^2 for all models and feature sets, demonstrating that six metals are insufficient for cross-metal extrapolation. An estimated 15–20 metals would be needed to adequately cover the ~ 7 effective dimensions of the descriptor space.

Single pH. All experiments were conducted at pH 3.9, the natural pH of the metal salt solutions, without pH adjustment. This pH is near the pH_{pzc} of zeolite P (3.2), and the results may not be extrapolated to neutral or alkaline conditions where metal speciation, surface charge, and competing H^+ exchange differ substantially. A systematic pH-dependence study is needed.

Single temperature. All experiments were conducted at $25 \pm 1^\circ\text{C}$. Thermodynamic parameters (ΔG° , ΔH° , ΔS°) could not be determined from single-temperature data and require isotherm measurements at a minimum of three temperatures via van't Hoff analysis.

Single-metal systems and counter-anion diversity. All experiments used single-metal solutions; competitive adsorption in multicomponent systems was not evaluated. The six metal salts introduced different counter-anions (NO_3^- for Pb^{2+} and Cd^{2+} , Cl^- for Ni^{2+} and Cr^{3+} , SO_4^{2-} for Cu^{2+} , CH_3COO^- for Hg^{2+}), which may influence free cation concentration through ion pair formation, most significantly CuSO_4 ($\log K = 2.3$) and Hg –acetate complexation ($\log K = 3.5$). Future studies should adopt a uniform counter anion (all nitrate salts) to eliminate this confounding.

Non-equilibrium status and three-point isotherms. The 60 min contact time was adequate for Hg^{2+} and Pb^{2+} but insufficient for Ni^{2+} , Cu^{2+} , Cr^{3+} , and Cd^{2+} to reach equilibrium, particularly at $C_0 = 200$ mg/L. The kinetic anomaly ($q_{t,200} < q_{t,100}$) for Cr^{3+} and Cd^{2+} may resolve with extended contact times if the 200 mg/L curves eventually surpass the 100 mg/L values, as predicted thermodynamically. All isotherm models were fitted to only three data points per metal (one degree of freedom for two-parameter models), limiting the reliability of isotherm parameter estimates. A dedicated equilibrium study (6–8 concentrations, ≥ 24 h contact time) is needed.

XPS limited to survey spectra. Only survey-level XPS was performed; high-resolution scans with peak fitting were not conducted. Quantitative determination of Na^+ depletion stoichiometry, metal oxidation states, and bonding environments was therefore not possible. The ion exchange mechanism was established through convergent evidence lines rather than through XPS alone. High-resolution XPS with peak deconvolution is a priority for future work.

SHAP is correlative, not causal. SHAP values quantify feature contributions to model predictions, not to physical system causality. SHAP cannot distinguish between correlated descriptors and cannot establish causal directionality. It is presented as a consistency check within a triangulation framework, not as standalone mechanistic proof. Definitive mechanistic confirmation would require EXAFS, high-resolution XPS, or DFT calculations.

Interaction descriptors are redundant. The six interaction features (pore accessibility, ion exchange potential, electrostatic factor, HSAB match, dehydration penalty, surface capacity) are deterministic functions of the metal descriptors and fixed adsorbent constants. The ablation study confirms they contribute no incremental predictive information ($\Delta R^2 < 0.001$). They are retained for interpretability but acknowledged as redundant for prediction.

D-R energy classification contradicts kinetic evidence. All D-R mean free energy values ($E < 8$ kJ/mol) formally classify the mechanism as physisorption, contradicting the ion exchange mechanism established by other evidence lines. This discrepancy is attributed to the three-point fitting limitation and non-equilibrium status rather than to a genuine mechanistic contradiction.

All adsorption experiments were conducted as single measurements without replicates. Although the internal consistency of the kinetic curves (smooth monotonic profiles, convergence at $C_0 = 10$ mg/L, high model fit quality, and successful ML external validation) supports data reliability, formal error bars based on replicate measurements could not be computed. Future studies should include triplicate experiments to enable statistical uncertainty quantification.

4. Conclusions

This study demonstrated the successful synthesis of zeolite P from coal fly ash and its application for heavy metal removal from aqueous solutions, supported by a machine learning framework for adsorption modeling.

Zeolite P was synthesized via hydrothermal treatment ($\text{BET} = 30$ m²/g, $\text{Si}/\text{Al} = 1.63$, $\text{pH}_{\text{pzc}} = 3.2$). Batch adsorption experiments were carried out at different initial concentrations ($C_0 = 10, 100, 200$ mg/L) of different metal ions (Pb^{2+} , Cd^{2+} , Hg^{2+} , Ni^{2+} , Cu^{2+} , Cr^{3+}) into aqueous solutions.

The experimental selectivity sequence ($C_0 = 200$ mg/L) $\text{Hg}^{2+} > \text{Pb}^{2+} > \text{Ni}^{2+} > \text{Cr}^{3+} > \text{Cu}^{2+} > \text{Cd}^{2+}$ has been explained through three converging physicochemical factors: electronegativity (governs cation–framework bonding strength, Pearson $r = +0.76$ with q_t), hydration enthalpy (governs the kinetic barrier to dehydration and pore entry), and pore accessibility (governs steric fit within the GIS channels). The superiority of Hg^{2+}

over Pb^{2+} despite lower electronegativity is attributed to HSAB soft–soft covalent bonding with partially soft framework oxygen sites. The collapse of the selectivity sequence at $C_0 = 10 \text{ mg/L}$ (all metals: $q_t \approx 0.53\text{--}0.63 \text{ mg/g}$) confirms that metal-specific differentiation requires sufficient loading to engage the intraparticle exchange pathway.

The physical adsorption of metal ions into zeolite P is well described by ion exchange, as observed by XPS and supported by PSO kinetics indicative of chemisorption/exchange; (i) multilinear Weber–Morris plots; (ii) Boyd analysis showing concentration-dependent transition from film to intraparticle diffusion control; (iii) effective diffusion coefficients; and (iv) SHAP-derived feature importance dominated by ion exchange-relevant descriptors (ionic radius, hydration enthalpy, charge density) rather than surface area-dependent features.

Gradient Boosting, trained on physics descriptors (19 features), achieved a Repeated K-Fold CV $R^2 = 0.979$ and an external-set $R^2 = 0.8799$ under a leakage-free group-level split. The gap of 0.099 confirms that the model has some predictive capability for unseen curves under the current data distribution but has limited extrapolation ability to unseen metals.

Group K-Fold and Leave-One-Metal-Out cross-validation produced strongly negative R^2 for all models, revealing that extrapolation to truly novel metals are not feasible with only 18 kinetic curves. This result highlights the data-efficiency challenge in ML adsorption and motivates multi-adsorbent, multi-laboratory data pooling.

SHAP analysis identified initial concentration, ionic radius, ion exchange potential, and electronegativity as the dominant predictors, linking ML importance to classical ion exchange selectivity theory and providing mechanistic transparency.

The ablation study demonstrated that physicochemical descriptors are essential for prediction: models trained on experimental variables alone (C_0 , t) yielded negative R^2 (GB: -0.242 ; Wilcoxon $p = 1.78 \times 10^{-15}$), while adding metal descriptors recovered $R^2 = 0.979$. The six interaction terms were redundant for tree-based models ($\Delta R^2 < 0.001$), which reconstruct multiplicative interactions internally. Electronegativity, hydration enthalpy, and softness as top features were consistent across three independent importance methods and aligned with Eisenman–Sherry ion exchange theory. Under Leave-One-Metal-Out validation, all models and feature sets yielded $R^2 < 0$, establishing that cross-metal extrapolation requires expansion beyond the current six-metal dataset.

Future work should include an expanded dataset with additional metals, adsorbents, and environmental conditions. A systematic pH study spanning pH 3–9 is recommended to: (i) determine the optimal pH for each metal, (ii) incorporate pH as a continuous ML feature, and (iii) validate whether the current model electrostatic factor descriptor captures the dominant pH effect or whether additional pH-dependent descriptors (metal speciation fractions, surface protonation degree) are required. In addition, temperature variations and competing ions need to be evaluated. Nested cross-validation for hyperparameter optimization and ensemble stacking may also enhance predictive accuracy without introducing data leakage.

Author Contributions: B.A.H.-G.: Software and writing—original draft. L.M.: Experimentation and writing—original draft. G.P.-R.: Formal analysis and writing—original draft. F.T.: Formal analysis, writing—review and editing, and funding acquisition. All authors have read and agreed to the published version of this manuscript.

Funding: This research was funded by the Secretaría de Investigación y Posgrado from Instituto Politécnico Nacional (Project SIP 20250057).

Data Availability Statement: The complete dataset and Python v3.14.3 code are available from the corresponding author upon reasonable request.

Acknowledgments: BAHG thanks the Secretaría de Investigación y Posgrado from Instituto Politécnico Nacional for financial support through the “Proyectos de Desarrollo Tecnológico o Innovación para Alumnos del IPN”. BAHG and LM also thank SECIHTI for a scholarship grant.

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

Abbreviations

ML	Machine learning
SEM	Scanning electron microscopy
EDS	Energy-dispersive X-ray spectroscopy
FTIR	Fourier transformed infrared spectroscopy
BET	Brunauer–Emmett–Teller technique
XRD	X-ray diffraction
XPS	X-ray photoelectron spectroscopy
HSAB	Hard and soft acids and bases theory
MAE	Mean absolute error
RMSE	Root mean squared error
LOMO	Leave-One-Metal-Out
SHAP	SHapley Additive exPlanation
r	Ionic radius in Å
ΔH_{hyd}	Hydration enthalpy in kJ/mol
χ	Pauling electronegativity
σ	Softness
r_{hyd}	Hydrated radius in Å
RF	Random Forest
GB	Gradient Boosting
SVR	Support Vector Regression
CFA	Coal fly ash
ICP-OES	Inductively coupled plasma optical emission spectroscopy

Appendix A

Table A1. Removal of metal ions depending on the type and amount of zeolite.

Zeolite Synthesis Conditions	Adsorbent Amount, g	Metal Ion Removal, %
0.5 M, 100 °C	0.60	63
0.5 M, 100 °C	0.15	8
0.5 M, 150 °C	0.60	55
0.5 M, 150 °C	0.15	2
1.5 M, 100 °C	0.60	56
1.5 M, 100 °C	0.15	9
1.5 M, 150 °C	0.60	79
1.5 M, 150 °C	0.15	39

Note: M is the concentration of NaOH in mol/L.

Table A2. Crystalline parameters obtained from XRD patterns for zeolite P.

Mineral	ICCD PDF	Crystal System	Space Group	Lattice Parameters, Å
Zeolite P	01-080-0699	Tetragonal	I 41/a m d	a = b = 9.999, c = 10.070
Mullite	01-074-4144	Orthorhombic	P b a m	a = 7.569, b = 7.682, c = 2.884
SiO ₂	01-087-2096	Hexagonal	P 32 2 1	a = b = 4.913, c = 5.405

Table A3. Chemical composition obtained through X-ray fluorescence spectroscopy in fly ash and zeolite P (contents are given in wt%).

Composition	Fly Ash	Zeolite P
SiO ₂	48.31	42.60
Al ₂ O ₃	19.29	22.20
Na ₂ O	9.82	13.50
Fe ₂ O ₃	9.20	8.62
CaO	4.18	5.06
MgO	1.89	2.02
K ₂ O	1.87	1.88
TiO ₂	1.85	2.29
SO ₃	1.65	0.82
P ₂ O ₅	1.16	0.24
Cl	0.52	0.57
ZnO	0.14	0.12
ZrO ₂	0.12	0.08
Si/Al ratio	2.12	1.63

Table A4. Kinetic parameters of pseudo-first-order (PFO) and pseudo-second-order (PSO) models. $q_{e,exp}$ = experimental capacity at $t = 60$ min; $q_{e,calc}$ = model-calculated equilibrium capacity; k_1 = PFO rate constant; k_2 = PSO rate constant; $h = k_2 q_e^2$ = initial adsorption rate; R^2 = coefficient of determination; RMSE = root mean square error. Nonlinear regression by Levenberg–Marquardt algorithm (SciPy v1.11). Experimental conditions: pH 3.9, dose 2 g/L, $T = 25$ °C, 300 rpm.

Metal Ion	C_0 (mg/L)	q_e, Exp (mg/g)	Pseudo-First-Order (PFO)				Pseudo-Second-Order (PSO)					
			$q_e, Calc$ (mg/g)	k_1 (min ⁻¹)	R^2	RMSE	$q_e, Calc$ (mg/g)	k_2 (g/mg·min)	h (mg/g·min)	R^2	RMSE	Best Model
Pb ²⁺	10	0.560	0.559	2.913	0.7160	0.0050	0.563	28.4690	9.009	0.9140	0.0030	PSO
Pb ²⁺	100	5.169	5.124	1.957	0.7700	0.1010	5.211	1.1590	31.463	0.9430	0.0510	PSO
Pb ²⁺	200	9.120	8.969	0.299	0.9380	0.5480	9.855	0.0430	4.202	0.8800	0.7670	Avrami
Ni ²⁺	10	0.560	0.559	3.340	0.8140	0.0030	0.562	46.0220	14.520	0.9400	0.0020	PSO
Ni ²⁺	100	0.801	0.702	0.195	0.9200	0.0600	0.815	0.2760	0.184	0.9350	0.0540	PSO
Ni ²⁺	200	2.177	1.649	0.288	0.7920	0.2150	1.865	0.1890	0.657	0.8160	0.2020	PSO
Cu ²⁺	10	0.540	0.539	2.874	0.7180	0.0050	0.543	28.4080	8.362	0.9140	0.0030	PSO
Cu ²⁺	100	1.870	1.643	0.084	0.6880	0.2250	1.901	0.0580	0.209	0.7800	0.1890	Avrami
Cu ²⁺	200	1.824	1.635	0.289	0.8630	0.1650	1.832	0.2030	0.682	0.8750	0.1570	PSO
Cr ³⁺	10	0.530	0.529	3.283	0.8160	0.0030	0.532	45.9330	12.986	0.9400	0.0020	PSO
Cr ³⁺	100	2.782	2.934	0.040	0.8950	0.2290	3.928	0.0090	0.139	0.9090	0.2130	Avrami
Cr ³⁺	200	2.046	1.848	0.242	0.9190	0.1460	2.079	0.1510	0.653	0.9140	0.1500	PFO
Hg ²⁺	10	0.630	0.629	3.462	0.8110	0.0030	0.632	46.1960	18.434	0.9410	0.0020	PSO
Hg ²⁺	100	5.257	4.891	3.158	0.0950	0.1700	4.948	2.5100	61.453	0.3110	0.1490	Avrami
Hg ²⁺	200	10.472	10.132	0.537	0.9150	0.4960	10.766	0.0850	9.799	0.8940	0.5540	Avrami
Cd ²⁺	10	0.550	0.549	2.902	0.9260	0.0020	0.553	32.0190	9.779	0.9260	0.0020	PSO
Cd ²⁺	100	2.178	2.114	0.075	0.8720	0.2010	2.597	0.0310	0.208	0.8950	0.1820	Avrami
Cd ²⁺	200	1.255	1.230	0.085	0.9430	0.0990	1.563	0.0510	0.125	0.9280	0.1110	Avrami

Table A5. Kinetic parameters of the Elovich and Avrami models. α = initial adsorption rate; β = surface coverage parameter; k_a = Avrami constant; n = Avrami exponent reflecting mechanism complexity. Negative Elovich R^2 values indicate the model is worse than the mean. For curves where the Avrami optimizer converged to boundary values ($q_e \approx 50$ mg/g, $k_a \approx 0$), the fit captured the overall shape, but parameters lack physical meaning.

Metal Ion	C_0 (mg/L)	q_e, Exp (mg/g)	Elovich				Avrami					
			α (mg/g·min)	β (g/mg)	R^2	RMSE	$q_e, Calc$ (mg/g)	k_a (min ⁻¹)	n	R^2	RMSE	Best Model
Pb ²⁺	10	0.560	1.0×10^6	35.660	−7.0000	0.0270	0.834	10	0.020	0.5880	0.0060	PSO
Pb ²⁺	100	5.169	1.0×10^6	3.456	−0.0180	0.2130	6.877	10	0.061	0.6730	0.1210	PSO
Pb ²⁺	200	9.120	16.875	0.630	0.7400	1.1270	8.910	0.298	1.167	0.9440	0.5240	Avrami

Table A5. Cont.

Metal Ion	C ₀ (mg/L)	q _e , Exp (mg/g)	Elovich				Avrami					
			α (mg/g·min)	β (g/mg)	R ²	RMSE	q _e , Calc (mg/g)	k _a (min ^{−1})	n	R ²	RMSE	Best Model
Ni ²⁺	10	0.560	1.0 × 10 ⁶	35.560	−23.500	0.0300	0.854	10	0.012	0.5670	0.0040	PSO
Ni ²⁺	100	0.801	0.341	5.773	0.8970	0.0680	0.706	0.191	0.962	0.9200	0.0600	PSO
Ni ²⁺	200	2.177	1.596	2.860	0.7830	0.2200	1.657	0.283	0.955	0.7920	0.2150	PSO
Cu ²⁺	10	0.540	1.0 × 10 ⁶	37.060	−6.2800	0.0260	0.802	10	0.021	0.5880	0.0060	PSO
Cu ²⁺	100	1.870	0.481	2.570	0.8790	0.1400	50.000	~0	0.360	0.9580	0.0830	Avrami
Cu ²⁺	200	1.824	1.833	3.017	0.8010	0.1990	1.651	0.279	0.917	0.8640	0.1640	PSO
Cr ³⁺	10	0.530	1.0 × 10 ⁶	37.690	−20.460	0.0280	0.807	10	0.013	0.5670	0.0040	PSO
Cr ³⁺	100	2.782	0.181	0.884	0.9260	0.1930	50.000	~0	0.529	0.9690	0.1240	Avrami
Cr ³⁺	200	2.046	1.584	2.549	0.8360	0.2070	1.851	0.241	0.981	0.9190	0.1460	PFO
Hg ²⁺	10	0.630	1.0 × 10 ⁶	31.400	−31.420	0.0340	0.965	10	0.011	0.5670	0.0040	PSO
Hg ²⁺	100	5.257	1.0 × 10 ⁶	3.594	−0.2340	0.1990	11.241	~0	0.037	0.7610	0.0870	Avrami
Hg ²⁺	200	10.472	246.04	0.834	0.7090	0.9190	10.221	0.547	0.806	0.9310	0.4490	Avrami
Cd ²⁺	10	0.550	1.0 × 10 ⁶	36.300	−9.4600	0.0280	0.827	10	0.017	0.5220	0.0060	PSO
Cd ²⁺	100	2.178	0.348	1.617	0.9170	0.1610	10.600	~0	0.430	0.9450	0.1320	Avrami
Cd ²⁺	200	1.255	0.162	2.324	0.9020	0.1300	1.142	0.104	1.494	0.9560	0.0870	Avrami

Appendix B

This section involves a description of equations and solving methods for the ML frameworks.

Random Forest (RF) [73]

Consider an ensemble of $B = 200$ decision trees, each trained on a bootstrap sample. The prediction is the average:

$$\hat{y}(x) = (1/B) \sum_{b=1}^B T_b(x)$$

where $T_b(x)$ is the prediction of tree b . Each tree is grown by recursively partitioning the feature space to minimize within-node MSE. At each candidate split, a random subset of $m = \sqrt{p}$ features (where $p = 19$, so $m \approx 4$) is evaluated, reducing inter-tree correlation. Splitting stops when $\text{min\_samples\_leaf} = 3$ or $\text{max\_depth} = 10$ is reached. The out-of-bag (OOB) error from the ~37% of samples not used in each bootstrap provides a built-in validation estimate.

The Random Forest hyperparameters were $n\_estimators (B) = 200$; $\text{max\_depth} = 10$; $\text{min\_samples\_split} = 5$; $\text{min\_samples\_leaf} = 3$; $\text{max\_features} (m) = \sqrt{p} \approx 4$; and $\text{feature scaling} = \text{none}$ (tree-based, scale-invariant).

Gradient Boosting (GB) [115]

Gradient Boosting builds $M = 200$ trees sequentially, each correcting the residual errors of the previous ensemble. Starting from the mean prediction:

$$F_0(x) = \text{mean}(y)$$

each subsequent tree $h_m(x)$ is fitted to the negative gradient (pseudo-residuals) of the loss function:

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x), \quad m = 1, 2, \dots, M$$

$$r_{im} = -\partial L(y_i, F_{m-1}(x_i)) / \partial F_{m-1}(x_i)$$

For squared error loss $L = \frac{1}{2}(y - F)^2$, the pseudo-residuals simplify to:

$$r_{im} = y_i - F_{m-1}(x_i)$$

Each tree $h_m(x)$ is a shallow tree ($\text{max\_depth} = 5$) trained on a random subsample of 80% of the data (stochastic gradient boosting). The learning rate $\eta = 0.05$ shrinks

each tree's contribution, reducing overfitting. The Gradient Boosting hyperparameters were `n_estimators` (M) = 200; `max_depth` = 5; `learning_rate` (η) = 0.05; `subsample` = 0.8 (stochastic); `min_samples_split` = 5; `min_samples_leaf` = 3; and `feature_scaling` = none (tree-based, scale-invariant).

Ridge Regression [116]

Ridge Regression minimizes the L2-penalized sum of squared errors:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \{ \sum_{i=1}^n (y_i - x_i^T \beta)^2 + \alpha \|\beta\|_2^2 \}$$

This admits a closed-form solution:

$$\hat{\beta} = (X^T X + \alpha I)^{-1} X^T y$$

where $\alpha = 10$ controls regularization strength. The L2 penalty shrinks all coefficients toward zero but never to exactly zero, retaining all features. Features are standardized before fitting:

$$\tilde{x}_j = (x_j - \mu_j) / \sigma_j$$

where μ_j and σ_j are computed on the training set only (not the test set) to prevent data leakage.

Elastic Net [117]

Elastic Net combines L1 (Lasso) and L2 penalties:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \{ (1/2n) \sum_i (y_i - x_i^T \beta)^2 + \alpha \rho \|\beta\|_1 + \alpha(1-\rho)/2 \|\beta\|_2^2 \}$$

where $\alpha = 0.5$ is the total penalty strength and $\rho = l_1\_ratio = 0.5$ balances L1 and L2 contributions equally. It is solved iteratively by coordinate descent. The L1 component performs automatic feature selection by shrinking some coefficients to exactly zero, while the L2 component handles correlated features. Features standardized as for Ridge.

Support Vector Regression (SVR) [118]

SVR finds a function that deviates from y_i by at most ε while being as flat as possible in a high-dimensional kernel space:

$$\min_{w, b, \xi, \xi^*} \quad \frac{1}{2} \|w\|^2 + C \cdot \sum_{i=1}^n (\xi_i + \xi_i^*)$$

subject to:

$$y_i - (w^T \varphi(x_i) + b) \leq \varepsilon + \xi_i$$

$$(w^T \varphi(x_i) + b) - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0 \quad \forall i$$

where $C = 10$ controls the trade-off between flatness and tolerance for deviations beyond $\varepsilon = 0.1$ and $\varphi(x)$ maps inputs into a high-dimensional space via the radial basis function (RBF) kernel:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

with $\gamma = 1/(p \cdot \operatorname{Var}(X))$, denoted as $\gamma = \text{'scale'}$ in scikit-learn. The dual problem is solved by sequential minimal optimization (SMO). Features standardized as for Ridge.

Cross-Validation Strategies

Three strategies of increasing stringency were applied exclusively to the development set (144 samples):

Strategy 1: Repeated K-Fold (interpolation test).

Here, 5 folds $\times$ 10 repeats = 50 iterations. In each iteration, ~80% of the samples are used for training and ~20% for validation. Samples from the same kinetic curve may appear in both sets, so this tests whether the model interpolates within the training distribution. The 50 R^2 scores yield a parametric 95% CI.

Strategy 2: Group K-Fold (curve extrapolation test).

Groups = Metal $\times$ C_0 kinetic curves (12 groups in the development set). Entire curves are held out per fold. This tests whether the model can predict unseen Metal $\times$ C_0 combinations when the same metal appears at other concentrations in the training set.

Strategy 3: Leave-One-Metal-Out (metal extrapolation test).

Train on five metals (all their curves), predict the sixth metal (all its curves). This is the strictest test: it evaluates whether the model can extrapolate to a completely novel metal from its physicochemical descriptors alone. Result: All models yielded $R^2 < 0$, demonstrating that six metals are insufficient for cross-metal extrapolation.

Statistical model comparison.

The Wilcoxon signed-rank test is applied to the 50 paired K-Fold R^2 scores:

$$H_0: \text{median}(R^2_{\text{model 1}} - R^2_{\text{model 2}}) = 0$$

$$T = \sum R_i^+ \quad (\text{sum of positive ranks of signed differences})$$

Bonferroni correction for $C(5,2) = 10$ pairwise comparisons: $\alpha_{\text{adj}} = 0.05/10 = 0.005$.

External validation: Final models trained once on the full development set (144 samples) and applied once to the locked external set (72 samples). This is the first and only time the external data is used. Metrics reported: R^2 , MAE, RMSE, and per-metal residual analysis (normality, heteroscedasticity, bias).

SHAP (SHapley Additive exPlanations)

SHAP [42] decomposes each individual prediction $f(x)$ into additive feature contributions:

$$f(x) = \varphi_0 + \sum_{j=1}^P \varphi_j(x)$$

where $\varphi_0 = E[f(X)]$ is the baseline (mean prediction across all training samples) and φ_j is the Shapley value for feature j , computed as the weighted average of its marginal contribution across all possible feature coalitions S :

$$\varphi_j = \sum_{S \subseteq \{1, \dots, P\} \setminus \{j\}} \frac{|S|! (p - |S| - 1)!}{p!} \cdot [f(S \cup \{j\}) - f(S)]$$

For tree-based models, TreeExplainer [119] computes exact Shapley values in polynomial time $O(TLD^2)$, where T = number of trees, L = maximum leaves per tree, and D = maximum depth. This avoids the exponential cost of brute-force Shapley computation.

Global feature importance is defined as the mean absolute SHAP value across all $n = 144$ development samples:

$$I_j = (1/n) \sum_{i=1}^n |\varphi_j(x_i)|$$

The SHAP beeswarm plot shows the distribution of φ_j values across all samples, with colors encoding the feature value (red = high, blue = low). This reveals both the magnitude and direction of each feature's effect.

Validation: SHAP rankings were cross-validated against two independent importance methods: (1) Mean Decrease in Impurity (MDI, built into RF/GB) and (2) permutation importance (20 repeats). The top five features were consistent across all three methods. SHAP is presented as a consistency check with established Eisenman–Sherry ion exchange selectivity theory, not as causal proof.

Ablation Study

The ablation study quantifies the contribution of each feature group by training models on progressively restricted feature sets and measuring the resulting performance change. Four feature sets are compared, as shown in Table A6.

Table A6. Ablation feature sets.

Feature Set	N	Features Included	Purpose
A: Experimental-Only	5	C_0 , t , $\log(1 + t)$, $\sqrt{t}$, $\log(1 + C_0)$	Predict without metal identity?
B: Descriptors-Only	14	Metal descriptors (8) + interaction (6)	Descriptors generalize across curves?
C: Metal + Experimental	13	Metal descriptors (8) + experimental (5)	Full minus 6 interaction terms
D: Full (all features)	19	All 19 features	Maximum information

For each feature set F and model M , the ablation R^2 is:

$$R^2_{\text{ablation}}(F, M) = 1 - \frac{\sum (y_i - M_F(x_i))^2}{\sum (y_i - \bar{y})^2}$$

where M_F denotes model M trained using only the features in set F . The contribution of a feature group G is quantified as the drop in R^2 when that group is removed:

$$\Delta R^2_G = R^2_{\text{Full}} - R^2_{\text{Full} \setminus G}$$

Statistical significance of each ablation comparison is assessed by the Wilcoxon signed-rank test on 50 paired K-Fold R^2 scores. Key results: (1) Experimental-Only yields negative R^2 for all models (GB: -0.242), confirming that physicochemical descriptors are essential; (2) Metal + Exp achieves $R^2 = 0.979$, identical to Full ($\Delta R^2 < 0.001$), indicating that the six interaction terms are redundant for tree-based models that can learn interaction effects internally; and (3) Full vs. Experimental-Only: Wilcoxon $p = 1.78 \times 10^{-15}$ for all three models tested.

Appendix C

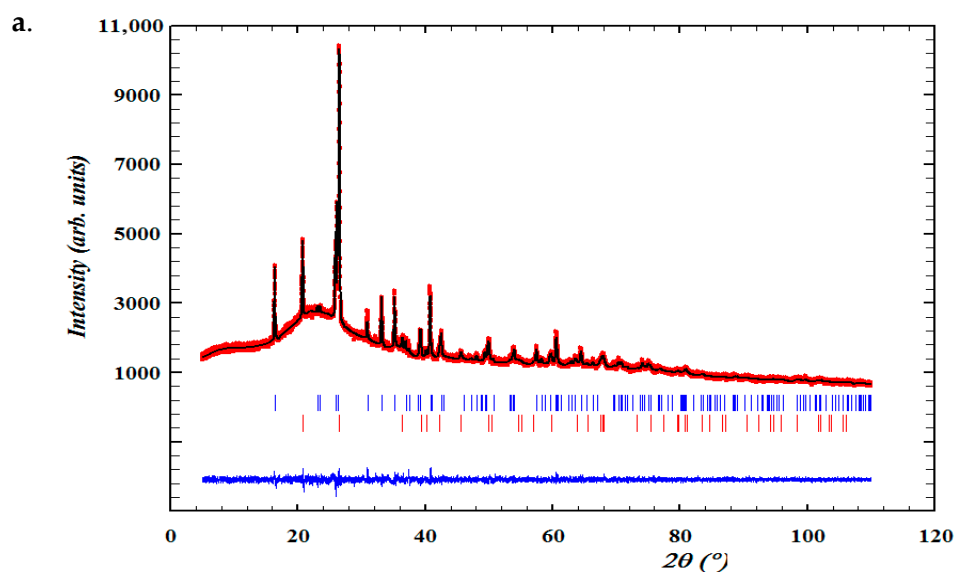


Figure A1. Cont.

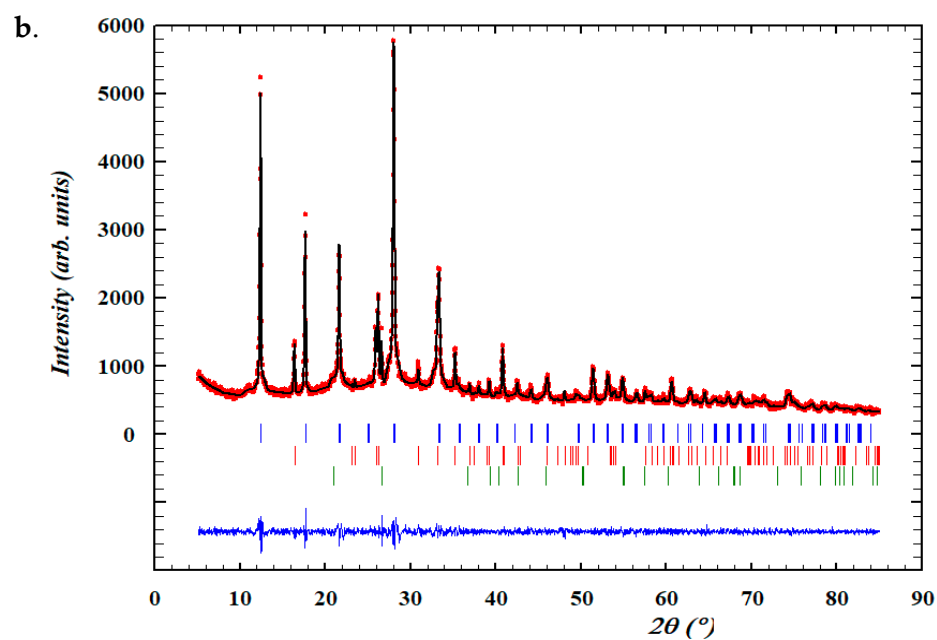


Figure A1. X-ray diffraction patterns for fly ash composed by 58% of mullite (blue line) and 42% of silicon dioxide (red line) (a), and zeolite P composed by 71.1 % of zeolite (blue line), 27.1% of mullite (red line) and 1.8% of silicon dioxide (green line) (b) Mullite is defined as $\text{Al}_{4.52}\text{Si}_{1.48}\text{O}_{9.74}$; silicon dioxide is defined as SiO_2 ; zeolite P is defined as $\text{Na}_{3.552}(\text{Al}_{3.632})(\text{H}_2\text{O})_{10.656}$.

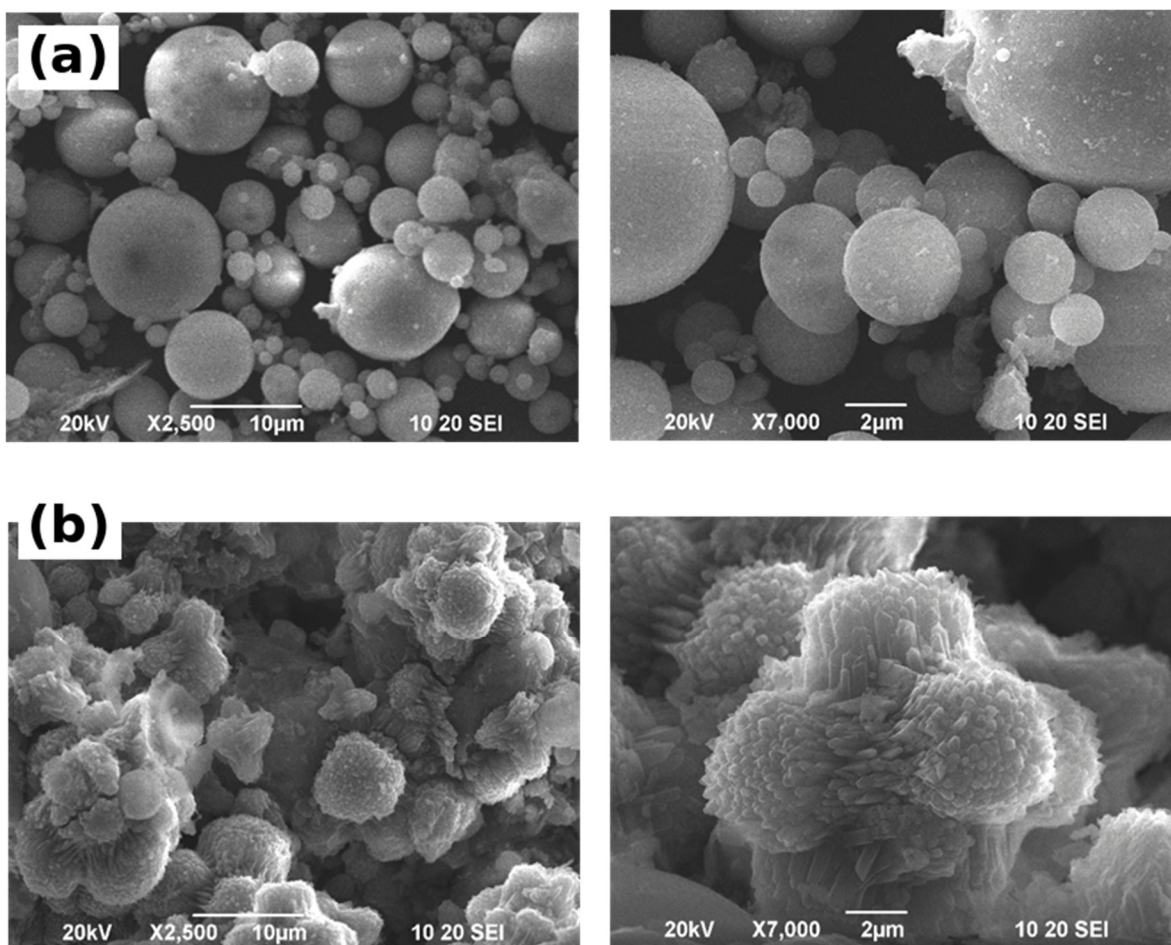


Figure A2. Images obtained by SEM at 2500X and 7000X for: (a) fly ash and (b) zeolite P.

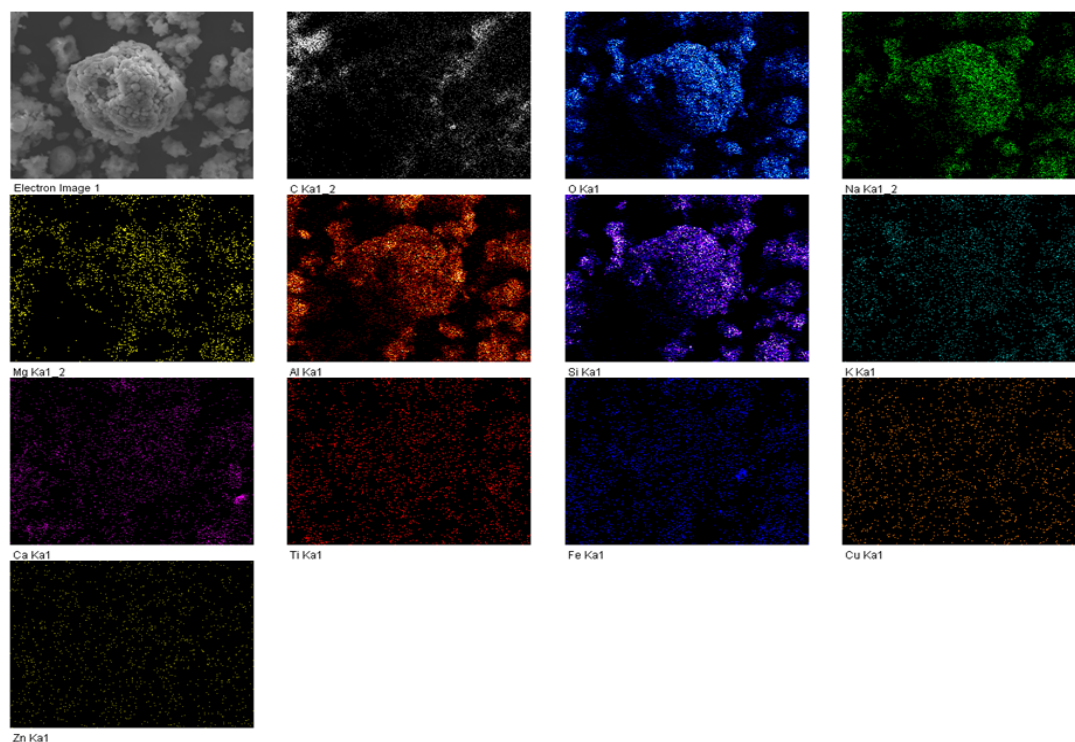


Figure A3. Elemental mapping through energy-dispersive X-ray spectroscopy for zeolite P.

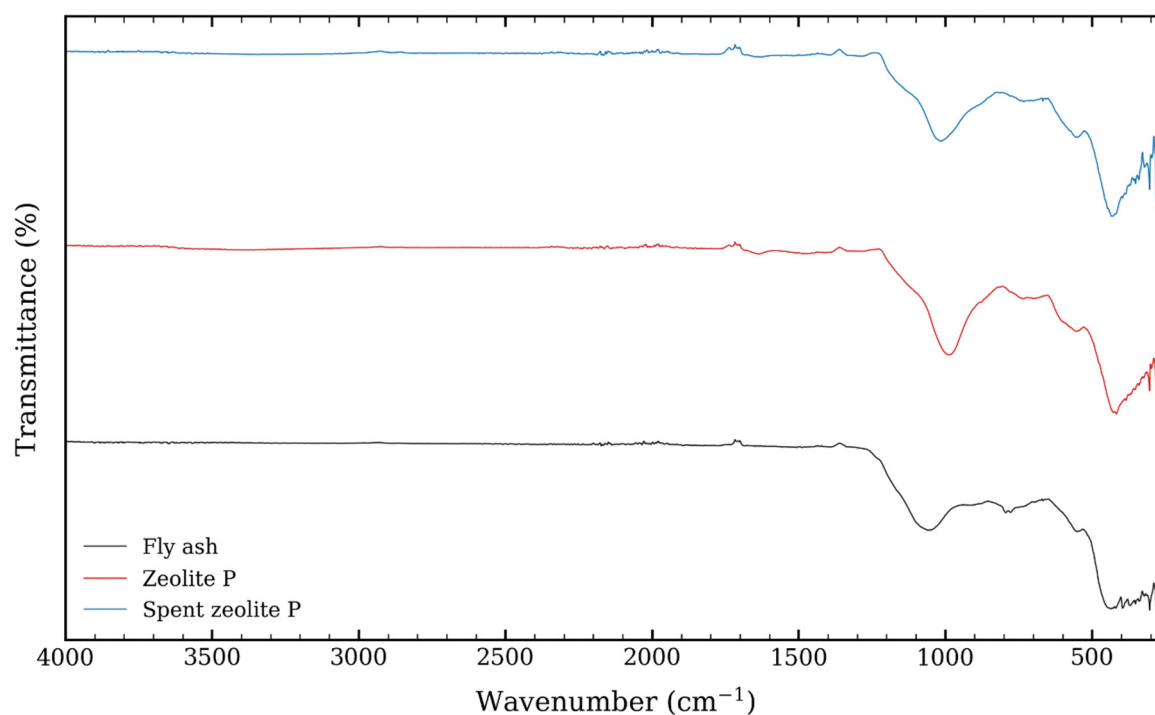


Figure A4. FTIR spectra for fly ash, zeolite P and spent zeolite P.

References

1. Yao, Z.T.; Ji, X.S.; Sarker, P.K.; Tang, J.H.; Ge, L.Q.; Xia, M.S.; Xi, Y.Q. A comprehensive review on the applications of coal fly ash. *Earth-Sci. Rev.* **2015**, *141*, 105–121. [[CrossRef](#)]
2. Ahmaruzzaman, M. A review on the utilization of fly ash. *Prog. Energy Combust. Sci.* **2010**, *36*, 327–363. [[CrossRef](#)]
3. Querol, X.; Moreno, N.; Umaña, J.C.; Alastuey, A.; Hernández, E.; López-Soler, A.; Plana, F. Synthesis of zeolites from coal fly ash: An overview. *Int. J. Coal Geol.* **2002**, *50*, 413–423. [[CrossRef](#)]

4. Joseph, I.V.; Tosheva, L.; Doyle, A.M. Simultaneous removal of Cd(II), Co(II), Cu(II), Pb(II), and Zn(II) ions from aqueous solutions via adsorption on FAU-type zeolites prepared from coal fly ash. *J. Environ. Chem. Eng.* **2020**, *8*, 103895. [\[CrossRef\]](#)
5. Belviso, C. State-of-art applications of fly ash from coal and biomass: A focus on zeolite synthesis processes and issues. *Prog. Energy Combust. Sci.* **2018**, *65*, 109–135. [\[CrossRef\]](#)
6. Bukhari, S.S.; Behin, J.; Kazemian, H.; Rohani, S. Conversion of coal fly ash to zeolite utilizing microwave and ultrasound energies: A review. *Fuel* **2015**, *140*, 250–266. [\[CrossRef\]](#)
7. Längauer, D.; Čablík, V.; Hredzák, S.; Zubrik, A.; Matik, M.; Danková, Z. Preparation of synthetic zeolites from coal fly ash by hydrothermal synthesis. *Materials* **2021**, *14*, 1267. [\[CrossRef\]](#)
8. Cundy, C.S.; Cox, P.A. The hydrothermal synthesis of zeolites: Precursors, intermediates and reaction mechanism. *Microporous Mesoporous Mat.* **2005**, *82*, 1–78. [\[CrossRef\]](#)
9. Derkowski, A.; Franus, W.; Beran, E.; Czimerová, A. Properties and potential applications of zeolitic materials produced from fly ash using simple method of synthesis. *Powder Technol.* **2006**, *166*, 47–54. [\[CrossRef\]](#)
10. Ojha, K.; Pradhan, N.C.; Samanta, A.N. Zeolite from fly ash: Synthesis and characterization. *Bull. Mat. Sci.* **2004**, *27*, 555–564. [\[CrossRef\]](#)
11. Kazemian, H.; Nghdali, Z.; Ghaffari Kashani, T.; Farhadi, F. Conversion of high silicon fly ash to Na-P1 zeolite: Alkaline fusion followed by hydrothermal crystallization. *Adv. Powder Technol.* **2010**, *21*, 279–283. [\[CrossRef\]](#)
12. Park, M.; Choi, C.L.; Lim, W.T.; Kim, M.C.; Choi, J.; Heo, N.H. Molten-salt method for the synthesis of zeolitic materials: I. Zeolite formation in alkaline molten-salt system. *Microporous Mesoporous Mat.* **2000**, *37*, 81–89. [\[CrossRef\]](#)
13. Bandura, L.; Panek, R.; Rotko, M.; Franus, W. Synthetic zeolites from fly ash from fly ash for an effective trapping of BTX in gas stream. *Microporous Mesoporous Mat.* **2016**, *223*, 1–9. [\[CrossRef\]](#)
14. Ramesh, K.; Reddy, D.D. Zeolites and their potential uses in agriculture. In *Advances in Agronomy*; Sparks, D.L., Ed.; Elsevier: Cambridge, MA, USA, 2011. [\[CrossRef\]](#)
15. Sivalingam, S.; Sen, S. Efficient removal of textile dye using nanosized fly ash derived zeolite-x: Kinetics and process optimization study. *J. Taiwan Inst. Chem. Eng.* **2019**, *96*, 305–314. [\[CrossRef\]](#)
16. Ogata, F.; Iwata, Y.; Kawasaki, N. Properties of novel adsorbent produced by hydrothermal treatment of waste fly ash in alkaline solution and its capability for adsorption of tungsten from aqueous solution. *J. Environ. Chem. Eng.* **2015**, *3*, 333–338. [\[CrossRef\]](#)
17. Mimura, H.; Yokota, K.; Akiba, K.; Onodera, Y. Alkali hydrothermal synthesis of zeolites from coal fly ash and their uptake properties of cesium ion. *J. Nucl. Sci. Technol.* **2001**, *38*, 766–772. [\[CrossRef\]](#)
18. Dyer, A.; Keir, D. Nuclear waste treatment by zeolites. *Zeolites* **1984**, *4*, 215–217. [\[CrossRef\]](#)
19. Hafez, M.B.; Nazmy, A.F.; Salem, F.; Eldesoki, M. Fixation mechanism between zeolite and some radioactive elements. *J. Radioanal. Chem.* **1978**, *47*, 115–119. [\[CrossRef\]](#)
20. Visa, M. Synthesis and characterization of new zeolite materials obtained from fly ash for heavy metals removal in advanced wastewater treatment. *Powder Technol.* **2016**, *294*, 338–347. [\[CrossRef\]](#)
21. Bosso, S.T.; Enzweiler, J. Evaluation of heavy metal removal from aqueous solution onto scolecite. *Water Res.* **2002**, *36*, 4795–4800. [\[CrossRef\]](#)
22. Li, G.; Li, M.; Zhang, X.; Cao, P.; Jiang, H.; Luo, J.; Jiang, T. Hydrothermal synthesis of zeolites-calcium silicate hydrate composite from coal fly ash with co-activation of Ca(OH)₂-NaOH for aqueous heavy metals removal. *Int. J. Min. Sci. Technol.* **2022**, *32*, 563–573. [\[CrossRef\]](#)
23. Koshy, N.; Singh, D.N. Fly ash zeolites for water treatment applications. *J. Environ. Chem. Eng.* **2016**, *5*, 1460–1472. [\[CrossRef\]](#)
24. Lee, M.G.; Yi, G.; Ahn, B.J.; Roddick, F. Conversion of coal fly ash into zeolite and heavy metal removal characteristics of the products. *Korean J. Chem. Eng.* **2000**, *17*, 325–331. [\[CrossRef\]](#)
25. Kobayashi, Y.; Ogata, F.; Nakamura, T.; Kawasaki, N. Synthesis of novel zeolites produced from fly ash by hydrothermal treatment in alkaline solution and its evaluation as an adsorbent for heavy metal removal. *J. Environ. Chem. Eng.* **2020**, *8*, 103687. [\[CrossRef\]](#)
26. Vardhan, K.H.; Kumar, P.S.; Panda, R.C. A review on heavy metal pollution, toxicity and remedial measures. *J. Mol. Liq.* **2019**, *290*, 111197. [\[CrossRef\]](#)
27. Ali, H.; Khan, E.; Ilahi, I. Environmental chemistry and ecotoxicology of hazardous heavy metals. *J. Chem.* **2019**, *2019*, 6730305. [\[CrossRef\]](#)
28. Burakov, A.E.; Galunin, E.V.; Burakova, I.V.; Kucherova, A.E.; Agarwal, S.; Tkachev, A.G.; Gupta, V.K. Adsorption of heavy metals on conventional and nanostructured materials for wastewater treatment purposes: A review. *Ecotox. Environ. Saf.* **2018**, *148*, 702–712. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Largitte, L.; Pasquier, R. A review of the kinetics adsorption models and their application to the adsorption of lead by an activated carbon. *Chem. Eng. Res. Des.* **2016**, *109*, 495–504. [\[CrossRef\]](#)
30. Anastopoulos, I.; Kyzas, G.Z. Are the thermodynamic parameters correctly estimated in liquid-phase adsorption phenomena? *J. Mol. Liq.* **2016**, *218*, 174–185. [\[CrossRef\]](#)

31. Zhu, X.; Wang, X.; Ok, Y.S. The application of machine learning methods for prediction of metal sorption onto biochars. *J. Hazard. Mater.* **2022**, *378*, 120727. [[CrossRef](#)]
32. Long, X.; Huangfu, X.; Huang, R.; Liang, Y.; Wu, S.; Wang, J. The application of machine learning methods for prediction of heavy metal by activated carbons, biochars, and carbon nanotubes. *Chemosphere* **2024**, *354*, 141584. [[CrossRef](#)]
33. Al-Zyoud, M.; Mostafa, S.A.; Khersan, I.; Gowrishankar, J.; Sahu, P.K.; Singla, S.; Sabirov, S.; Khudayberganov, I.; Sherzod, S. Effects of physical properties on the heavy metal adsorption of biochar via a robust approach. *Curr. Res. Biotechnol.* **2026**, *11*, 100367. [[CrossRef](#)]
34. Kumari, B.; Seedat, N.; Moothi, K.; Roopchand, R. Artificial intelligence-assisted modelling of heavy metal adsorption using cellulose-based and bio-wasted adsorbents: A focus on ANN and ANFIS architectures. *Results Eng.* **2025**, *28*, 107147. [[CrossRef](#)]
35. Liu, J.; Zhao, J.; Du, J.; Peng, S.; Tan, S.; Zhang, W.; Yan, X.; Wang, H.; Lin, Z. Machine learning predicts heavy metal adsorption on iron (oxyhydr)oxides. *Sci. Tot. Environ.* **2024**, *950*, 175370. [[CrossRef](#)]
36. Yaseen, Z.M.; Alhalimi, F.L. Heavy metal adsorption efficiency prediction using biochar properties: A comparative analysis for ensemble machine learning models. *Sci. Rep.* **2025**, *15*, 13434. [[CrossRef](#)]
37. Wang, J.; Long, X.; Huang, R.; Liang, Y.; Wu, S.; Huanfu, X.; Liu, H. Simulation, prediction and optimization of heavy metal adsorption by metal-organic frameworks with machine learning. *Environ. Res.* **2025**, *285*, 122612. [[CrossRef](#)] [[PubMed](#)]
38. Shannon, R.D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. Sect. A* **1976**, *32*, 751–767. [[CrossRef](#)]
39. Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, USA, 1960.
40. Marcus, Y. Thermodynamics of solvation of ions. Part 5—Gibbs free energy of hydration at 298.15 K. *J. Chem. Soc. Faraday Trans.* **1991**, *87*, 2995–2999. [[CrossRef](#)]
41. Pearson, R.G. Hard and soft acids and bases. *J. Am. Chem. Soc.* **1963**, *85*, 3533–3539. [[CrossRef](#)]
42. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *arXiv* **2017**, arXiv:1705.07874. [[CrossRef](#)]
43. Barrer, R.M. *Hydrothermal Chemistry of Zeolites*; Academic Press: London, UK, 1982.
44. Brunauer, S.; Emmett, P.H.; Teller, E. Adsorption of gases in multimolecular layers. *J. Am. Chem. Soc.* **1938**, *60*, 309–319. [[CrossRef](#)]
45. Le Bail, A.; Duroy, H.; Fourquet, J.L. Ab-initio structure determination of LiSbWO₆ by X-ray powder diffraction. *Mater. Res. Bull.* **1988**, *23*, 447–452. [[CrossRef](#)]
46. Rodríguez-Carvajal, J. Recent advances in magnetic structure determination by neutron powder diffraction. *Phys. B Cond. Matter* **1993**, *192*, 55–69. [[CrossRef](#)]
47. Akcil, A.; Koldas, S. Acid mine drainage (AMD): Causes, treatment and case studies. *J. Clean. Prod.* **2006**, *14*, 1139–1145. [[CrossRef](#)]
48. Fu, F.; Wang, Q. Removal of heavy metal ions from wastewaters: A review. *J. Environ. Manag.* **2011**, *92*, 407–418. [[CrossRef](#)] [[PubMed](#)]
49. Barakat, M.A. New trends in removing heavy metals from industrial wastewater. *Arab. J. Chem.* **2011**, *4*, 361–377. [[CrossRef](#)]
50. Smith, R.M.; Martell, A.E. *Critical Stability Constants: Vol. 4, Inorganic Complexes*; Plenum Press: New York, NY, USA, 1976.
51. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)] [[PubMed](#)]
52. Lagergren, S. Zur theorie der sogenannten Adsorption gelöster stoffe. *Kungliga Svenska Vetenskapsakademiens. Handlingar* **1898**, *24*, 1–39.
53. Ho, Y.S.; McKay, G. Pseudo-second order model for sorption processes. *Process Biochem.* **1999**, *34*, 451–465. [[CrossRef](#)]
54. Roginsky, S.Z.; Zeldovich, Y.B. An equation for the kinetics of activated adsorption. *Acta Physicochimica U.R.S.S.* **1934**, *1*, 554–559. [[CrossRef](#)]
55. Lopes, E.C.; dos Anjos, F.S.; Vieira, E.F.; Cestari, A.R. An alternative Avrami equation to evaluate kinetic parameters of the interaction of Hg(II) with thin chitosan membranes. *J. Colloid Interface Sci.* **2003**, *263*, 542–547. [[CrossRef](#)]
56. Avrami, M. Kinetics of phase change. I: General theory. *J. Chem. Phys.* **1939**, *7*, 1103–1112. [[CrossRef](#)]
57. Weber, W.J.; Morris, J.C. Kinetics of adsorption on carbon from solution. *J. Sanit. Eng. Div.* **1963**, *89*, 31–60. [[CrossRef](#)]
58. Wu, F.-C.; Tseng, R.-L.; Juang, R.-S. Initial behavior of intraparticle diffusion model used in the description of adsorption kinetics. *Chem. Eng. J.* **2009**, *153*, 1–8. [[CrossRef](#)]
59. Boyd, G.E.; Adamson, A.W.; Myers, L.S. The exchange adsorption of ions from aqueous solutions by organic zeolites. II. Kinetics. *J. Am. Chem. Soc.* **1947**, *69*, 2836–2848. [[CrossRef](#)] [[PubMed](#)]
60. Vermeulen, T. Theory for irreversible and constant-pattern solid diffusion. *Ind. Eng. Chem.* **1953**, *45*, 1664–1670. [[CrossRef](#)]
61. Ruthven, D.M. *Principles of Adsorption and Adsorption Processes*; John Wiley & Sons: Hoboken, NJ, USA, 1984.
62. Hall, K.R.; Eagleton, L.C.; Acrivos, A.; Vermeulen, T. Pore- and solid-diffusion kinetics in fixed-bed adsorption under constant-pattern conditions. *Ind. Eng. Chem. Fundamen.* **1966**, *5*, 212–223. [[CrossRef](#)]
63. Freundlich, H. Über die Adsorption in Lösungen. *Z. Phys. Chem.* **1906**, *57*, 385–470. [[CrossRef](#)]
64. Temkin, M.I.; Pyzhev, V. Kinetics of ammonia synthesis on promoted iron catalysts. *Acta Physicochimica U.R.S.S.* **1940**, *12*, 327–356.

65. Helfferich, F. *Ion Exchange*; McGraw-Hill: Columbus, OH, USA, 1962.
66. Nightingale, E.R. Phenomenological theory of ion solvation. Effective radii of hydrated ions. *J. Phys. Chem.* **1959**, *63*, 1381–1387. [\[CrossRef\]](#)
67. Eisenman, G. Cation selective glass electrodes and their mode of operation. *Biophys. J.* **1962**, *2*, 259–323. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Sherry, H.S. The ion-exchange properties of zeolites. I. Univalent ion exchange in synthetic faujasite. *J. Phys. Chem.* **1966**, *70*, 1158–1168. [\[CrossRef\]](#)
69. Dyer, A. Ion-exchange properties of zeolites and related materials. In *Introduction to Zeolite Science and Practice*, 3rd ed.; Čejka, J., van Bekkum, H., Corma, A., Schüth, F., Eds.; Elsevier: Amsterdam, The Netherlands, 2007; pp. 525–553.
70. Breck, D.W. *Zeolite Molecular Sieves: Structure, Chemistry, and Use*; Wiley: New York, NY, USA, 1974.
71. Spósito, G. *The Surface Chemistry of Soils*; Oxford University Press: Oxford, UK, 1984.
72. Baerlocher, C.; McCusker, L.B.; Olson, D.H. *Atlas of Zeolite Framework Types*, 6th ed.; Elsevier: Amsterdam, The Netherlands, 2007.
73. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
74. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.
75. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Aroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [\[CrossRef\]](#)
76. Ploton, P.; Mortier, F.; Réjou-Méchain, M.; Barbier, N.; Picard, N.; Rossi, V.; Dormann, C.; Cornu, G.; Viennois, G.; Bayol, N.; et al. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* **2020**, *11*, 4540. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Demšar, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30. Available online: <https://jmlr.org/papers/volume7/demsar06a/demsar06a.pdf> (accessed on 25 March 2026).
78. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed.; Lulu.com: Durham, NC, USA, 2022.
79. International Centre for Diffraction Data (ICDD). *Powder Diffraction File, Card No. 01-074-4144*; ICDD: Newtown Square, PA, USA, 2024.
80. International Centre for Diffraction Data (ICDD). *Powder Diffraction File, Card No. 01-087-2096*; ICDD: Newtown Square, PA, USA, 2024.
81. Phair, J.W.; Van Deventer, J.S.J. Characterization of fly-ash-based geopolymeric binders activated with sodium aluminate. *Ind. Eng. Chem. Res.* **2002**, *41*, 4242–4251. [\[CrossRef\]](#)
82. International Centre for Diffraction Data (ICDD). *Powder Diffraction File, Card No. 01-080-0699*; ICDD: Newtown Square, PA, USA, 2024.
83. Franus, W.; Wdowin, M.; Franus, M. Synthesis and characterization of zeolites prepared from industrial fly ash. *Environ. Monit. Assess.* **2014**, *186*, 5721–5729. [\[CrossRef\]](#)
84. Hollman, G.G.; Steenbruggen, G.; Janssen-Jurkovičová, M. A two-step process for the synthesis of zeolites from coal fly ash. *Fuel* **1999**, *78*, 1225–1230. [\[CrossRef\]](#)
85. Fernández-Pereira, C.; Luna-Galiano, Y.; Rodríguez-Piñero, M.A.; Vale, J.; Querol, X. Utilisation of zeolitised coal fly ash as immobilising agent of a metallurgical waste. *J. Chem. Technol. Biotechnol.* **2002**, *77*, 305–310. [\[CrossRef\]](#)
86. Wdowin, M.; Franus, M.; Panek, R.; Bandura, L.; Franus, W. The conversion technology of fly ash into zeolites. *Clean Technol. Environ. Policy* **2014**, *16*, 1217–1223. [\[CrossRef\]](#)
87. Wang, Z.; Fan, Y.; Li, Y.; Qu, F.; Wu, D.; Kong, H. Synthesis of zeolite/hydrous lanthanum oxide composite from coal fly ash for efficient phosphate removal from lake water. *Microporous Mesoporous Mat.* **2016**, *222*, 226–234. [\[CrossRef\]](#)
88. Kolay, P.K.; Singh, D.N.; Murti, M.V.R. Synthesis of zeolites from lagoon ash. *Fuel* **2001**, *80*, 739–745. [\[CrossRef\]](#)
89. Kutchko, B.G.; Kim, A.G. Fly ash characterization by SEM-EDS. *Fuel* **2006**, *85*, 2537–2544. [\[CrossRef\]](#)
90. Deng, L.; Xu, Q.; Wu, H. Synthesis of zeolite-like material by hydrothermal and fusion methods using municipal solid waste fly ash. *Procedia Environ. Sci.* **2016**, *31*, 662–667. [\[CrossRef\]](#)
91. Elaiopoulos, K.; Perraki, T.; Grigoropoulou, E. Mineralogical study and porosimetry measurements of zeolites from Scaloma area, Thrace, Greece. *Microporous Mesoporous Mat.* **2008**, *112*, 441–449. [\[CrossRef\]](#)
92. Schüth, F.; Schmidt, W. Microporous and mesoporous materials. *Adv. Mater.* **2002**, *14*, 629–638. [\[CrossRef\]](#)
93. Ates, A.; Hardacre, C. The effect of various treatment conditions on natural zeolites: Ion exchange, acidic, thermal and steam treatments. *J. Colloid Interface Sci.* **2012**, *372*, 130–140. [\[CrossRef\]](#)
94. Kosmulski, M. Isoelectric points and points of zero charge of metal (hydr)oxides: 50 years after Parks' review. *Adv. Colloid Interface Sci.* **2016**, *238*, 1–61. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Mechnou, I.; Oukkass, S.; El Kartouti, A.; Hlaibi, M.; Saleh, N. High adsorption capacity of diclofenac and paracetamol using OMWW based W-carboxylate doped activated carbon. *Sep. Purif. Technol.* **2025**, *378*, 134719. [\[CrossRef\]](#)

96. Sprynskyy, M.; Buszewski, B.; Terzyk, A.P.; Namieśnik, J. Study of the selection mechanism of heavy metal (Pb^{2+} , Cu^{2+} , Ni^{2+} , and Cd^{2+}) adsorption on clinoptilolite. *J. Colloid Interface Sci.* **2006**, *304*, 21–28. [CrossRef]
97. Rai, D.; Sass, B.M.; Moore, D.A. Chromium(III) hydrolysis constants and solubility of chromium(III) hydroxide. *Inorg. Chem.* **1987**, *26*, 345–349. [CrossRef]
98. Helm, L.; Merbach, A.E. Inorganic and bioinorganic solvent exchange mechanisms. *Chem. Rev.* **2005**, *105*, 1923–1960. [CrossRef]
99. Stunzi, H.; Marty, W. Early stages of the hydrolysis of chromium(III) in aqueous solution. 1. Characterization of a tetrameric species. *Inorg. Chem.* **1983**, *22*, 2145–2150. [CrossRef]
100. Inglezakis, V.J. The concept of “capacity” in zeolite ion-exchange systems. *J. Colloid Interface Sci.* **2005**, *281*, 68–79. [CrossRef] [PubMed]
101. Álvarez-Ayuso, E.; García-Sánchez, A.; Querol, X. Purification of metal electroplating waste waters using zeolites. *Water Res.* **2003**, *37*, 4855–4862. [CrossRef] [PubMed]
102. Berrar, D. Cross-validation. In *Encyclopedia of Bioinformatics and Computational Biology*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 542–545. [CrossRef]
103. Velarde, L.; Nabavi, M.S.; Escalera, E.; Antti, M.-L.; Akhtar, F. Adsorption of heavy metals on natural zeolites: A review. *Chemosphere* **2023**, *328*, 138508. [CrossRef] [PubMed]
104. Nevenka, R.; Senila, M.; Cadar, O.; Senila, L. Simultaneous removal of heavy metals (Cu, Cd, Cr, Ni, Zn and Pb) from aqueous solutions using thermally treated Romanian zeolitic volcanic tuff. *Molecules* **2022**, *27*, 3938. [CrossRef]
105. Rios, A.; Williams, C.; Roberts, C. Removal of heavy metals from acid mine drainage (AMD) using coal fly ash, natural clinker and synthetic zeolites. *J. Hazard. Mater.* **2008**, *156*, 23–35. [CrossRef]
106. Cheng, T.; Chen, C.; Tang, R.; Han, C.H.; Tian, Y. Competitive adsorption of Cu, Ni, Pb, and Cd from aqueous solution onto fly ash-based Linde F(K) zeolite. *Iran. J. Chem. Chem. Eng.* **2018**, *37*, 61–72. [CrossRef]
107. Liu, Y.; Yan, C.; Zhao, J.; Zhang, Z.; Wang, H.; Zhou, S.; Wu, L. Synthesis of zeolite P1 from fly ash under solvent-free conditions for ammonium removal from water. *J. Clean. Prod.* **2018**, *202*, 11–22. [CrossRef]
108. Medykowska, M.; Wiśniewska, M.; Szewczuk-Karpisz, K.; Panek, R. Interaction mechanism of heavy metal ions with the nanostructured zeolites surface—Adsorption, electrokinetic and XPS studies. *J. Mol. Liq.* **2022**, *357*, 119144. [CrossRef]
109. Mordekovitz, Y.; Shoval, Y.; Froumin, N.; Hayun, S. Effect of structure and composition of non-stoichiometric magnesium aluminate spinel on water adsorption. *Materials* **2020**, *13*, 3195. [CrossRef]
110. Inglezakis, V.J.; Loizidou, M.D.; Grigoropoulou, H.P. Ion exchange of Pb^{2+} , Cu^{2+} , Fe^{3+} , and Cr^{3+} on natural clinoptilolite: Selectivity determination and influence of acidity on metal uptake. *J. Colloid Interface Sci.* **2003**, *261*, 49–54. [CrossRef] [PubMed]
111. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]
112. Cerqueira, V.; Torgo, L.; Mozetič, I. Evaluating time series forecasting models: An empirical study on performance estimation methods. *Mach. Learn.* **2020**, *109*, 1997–2028. [CrossRef]
113. Cawley, C.; Talbot, N. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Mach. Learn. Res.* **2010**, *11*, 2079–2107. Available online: <http://jmlr.csail.mit.edu/papers/v11/cawley10a.html> (accessed on 23 March 2026).
114. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef]
115. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
116. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **2000**, *42*, 80–86. [CrossRef]
117. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [CrossRef]
118. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: Berlin/Heidelberg, Germany, 1995.
119. 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]

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.