Environmental and Microbial Technology
https://www.sciltp.com/journals/emt

Article

AI-Driven Prediction of Uranium Adsorption on Acid-Modified Biochar: Integrating Large Language Models with Interpretable Machine Learning

Jingyang Sun [1], Sut Ian Chan [2] and Shuting Zhuang [3,*]

[1] School of Ecology & Environment, Renmin University of China, Beijing 100872, China
[2] School of Information, Renmin University of China, Beijing 100872, China
[3] School of Chemistry & Life Resources, Renmin University of China, Beijing 100872, China
* Correspondence: zst@ruc.edu.cn

How To Cite: Sun, J.Y.; Chan, S.I.; Zhuang, S.T. AI-Driven Prediction of Uranium Adsorption on Acid-Modified Biochar: Integrating Large Language Models with Interpretable Machine Learning. *Environmental and Microbial Technology* **2026**, *1*(1), 5.

Received: 7 December 2025
Revised: 21 December 2025
Accepted: 30 December 2025
Published: 6 January 2026

Abstract: Efficient uranium recovery from radioactive wastewater is a pressing challenge in sustainable nuclear energy development. Here, we present an innovative AI-driven framework that integrates large language models with machine learning to optimize uranium adsorption by acid-modified biochar. By automatically extracting and structuring literature data, we compiled a high-quality dataset comprising 589 experimental data points. Among four tree-based ensemble models, CatBoost delivered the best performance (test R^2 = 0.98, RMSE = 24.78). Feature importance indicates that adsorption conditions (71.51%) as the most influential factors, significantly outweighing biochar preparation conditions (15.27%) and physicochemical properties (13.22%). SHapley Additive exPlanations (SHAP) analysis further provided insights into how key features influence uranium adsorption, identifying important variables and their impact patterns. Finally, we developed a user-friendly graphical user interface that enables rapid, intelligent prediction of uranium adsorption capacity, supporting data-driven experimental design. This study provides a referenceable, AI-powered solution for radioactive wastewater treatment and offers a transferable framework for the intelligent remediation of other environmental pollutants.

Keywords: uranium adsorption; acid-modified biochar; large language models; machine learning; intelligent prediction

1. Introduction

The rapid expansion of nuclear energy necessitates effective management of its inevitable byproducts, particularly uranium (U), which poses significant environmental risks due to its long-lived radioactivity and persistent toxicity [1–3]. Efficient removal of uranium from contaminated wastewater is crucial to mitigate its impact on ecosystems and human health, and to ensure the sustainable development of nuclear energy. Among existing technologies, adsorption is a mainstream approach for uranium-contaminated water remediation due to its operational simplicity, low cost, and high efficiency [4–6]. However, the selection of suitable adsorbents—capable of high-efficiency uranium removal at low cost—remains a critical challenge [7–9].

Biochar, a carbon-rich material produced by pyrolyzing biomass under oxygen-limited or anaerobic conditions, has gained considerable attention as a potential adsorbent due to its high surface area, porous structure, and environmental sustainability [10–12]. Its adsorption performance is enhanced by multiple mechanisms, such as surface complexation, chemical reduction, and ion exchange, all of which facilitate the uptake of uranium ions [13,14]. In particular, acid modification has proven effective in improving biochar's adsorption capacity by introducing oxygen-containing functional groups like carboxyl, hydroxyl, and carbonyl, which enhance surface

hydrophilicity and expand the porous structure [15–17]. Despite these advances, achieving optimal uranium adsorption performance remains a challenge, as it requires a deep understanding of the complex interactions between material properties, preparation conditions, and adsorption factors.

Traditional optimization methods, relying on repeated experimental trials, are labor-intensive and time-consuming. Furthermore, these methods struggle to account for the complex, nonlinear interactions between the various factors influencing adsorption performance [18–20]. In recent years, artificial intelligence (AI) has emerged as a powerful tool to address these challenges [21,22]. Specifically, large language models (LLMs) offer a promising solution for automating the extraction of material properties from scientific literature, thereby accelerating data collection and facilitating the creation of high-quality datasets [23]. For instance, Polak and Morgan [24] proposed the ChatExtract method that can fully automate highly accurate data extraction using an advanced conversational LLM, with precision and recall both close to 90% from the best conversational LLMs like GPT-4. In parallel, machine learning (ML) algorithms excel at identifying hidden patterns in complex data, making them well-suited for predicting the behavior of multi-variable systems, such as uranium adsorption [25,26]. The application of AI in pollutant adsorption modeling has been successfully demonstrated for heavy metals, pharmaceuticals, and per- and polyfluoroalkyl substances (PFAS), showcasing its potential for optimizing environmental remediation processes [27–29]. For example, Fabregat-Palau et al. [29] developed a novel ML tool, PFASorptionML. By leveraging PFAS-specific and soil-dependent input parameters, this tool has wide applicability and can predict the K_d of 47 types of PFAS.

However, significant limitations persist in current research. First, the application of LLMs for automated material data extraction remains nascent. While promising, these tools still suffer from limited semantic parsing of numerical data and high computational costs for training, necessitating a more generalizable and efficient framework [30]. Furthermore, existing studies exhibit a "binary disconnect" that limits the understanding of adsorption mechanisms: one stream emphasizes preparation parameters but neglects how they shape physicochemical properties and their link to adsorption performance [31]; the other prioritizes physicochemical properties while treating them as static, failing to integrate dynamic preparation logic and clarify their role in adsorption mechanisms [32–34]. Most importantly, the integration of LLMs and ML techniques—key for constructing high-quality datasets and reliable predictive models—remains largely unexplored, leaving a critical gap in bridging data-driven discovery with mechanistic understanding.

To address these challenges, we propose an integrated AI-driven workflow that combines LLM-based data acquisition with interpretable ML modeling for the optimization of uranium adsorption by acid-modified biochar. This framework introduces several key innovations: (1) a novel, automated method for extracting high-precision material data from the literature, utilizing LLMs and advanced prompt engineering techniques; (2) a comparison of four mainstream ensemble ML models (Random Forest, XGBoost, CatBoost, and LightGBM) for predicting uranium adsorption capacity (Q_e), with hyperparameter optimization to improve prediction accuracy; (3) a data-driven analysis of the key features affecting adsorption performance, including SHapley Additive exPlanations (SHAP) analysis and partial dependence plots (PDPs), to provide insights into the underlying adsorption mechanisms; and (4) the development of a user-friendly graphical user interface (GUI) that enables researchers to rapidly assess the uranium adsorption performance of various acid-modified biochar materials.

This study represents a novel approach by combining the power of LLMs and ML techniques to optimize the uranium adsorption capacity of acid-modified biochar. It breaks through traditional research bottlenecks by automating data extraction, enhancing model interpretability, and seamlessly integrating data-driven insights with predictive capabilities in a unified workflow. Beyond uranium removal, the framework we present offers a referenceable solution for the development of other environmentally functional materials, demonstrating the broad potential of AI-driven strategies in addressing global pollution challenges.

2. Methodologies

2.1. Data Collection and Preprocessing

A targeted literature search was first conducted in the Web of Science database using the keyword combination "Biochar" + "Uranium adsorption" for the period 2015–2025, which yielded 261 potentially relevant publications. Subsequently, a two-stage screening process was implemented to identify eligible studies. In the first stage, an automated screening was performed using DeepSeek-R1 to exclude irrelevant publications according to predefined criteria (Text S1). In the second stage, expert validation was conducted, wherein publications retained from the initial screening were manually reviewed by domain experts to confirm relevance and assess data quality.

Following the screening process, a structured data extraction framework integrating multiple LLMs with rule-based validation and expert supervision was developed. Specifically, three LLMs were employed to extract

structured information from the selected articles, namely: DeepSeek-R1, GPT-4o and Doubao (Beta). To ensure consistency across models, a standardized prompt template was designed (Text S2). A two-tier verification mechanism was then applied to enhance data reliability: (1) Cross-model consistency checking, where each data point required agreement from at least two of the three LLMs to be considered preliminarily consistent; conflicting cases were flagged for further review. A quantitative scoring system (0–100 points) was also established to evaluate consistency and reliability (Table S1). (2) Expert supervision was used to resolve discrepancies among model outputs and to manually verify critical parameters. For data presented graphically (e.g., adsorption isotherms and kinetics), numerical values were extracted using WebPlotDigitizer to minimize human reading error and improve quantitative accuracy. All extracted parameters were converted into standardized units to facilitate dataset integration.

Categorical variables were encoded accordingly: acid type was ordinally encoded based on descending acid strength as follows: HCl (6), HNO_3 (5), PA (4), H_3PO_4 (3), CH_3COOH (2), and HA (1). The sequence of pyrolysis and modification (Seq_P_M) was represented as a binary variable, with "0" indicating modification before pyrolysis and "1" indicating pyrolysis before modification. To handle missing data, four imputation methods (multiple imputation by chained equations (MICE) [36], mean imputation, median imputation, and mode imputation) were systematically evaluated. The optimal method was selected based on minimal disturbance to the original data distribution for missing value filling. Additionally, Spearman correlation analysis was conducted to assess linear relationships between numerical variables. To mitigate multicollinearity in subsequent machine learning modeling, at least one of any pair of strongly correlated variables ($|r| > 0.9$) was removed, thereby enhancing model accuracy and interpretability. All data preprocessing, analytical procedures, ML model development, and interpretation were conducted in PyCharm (version 2025.1). A schematic summary of the integrated methodology is provided in Figure 1.

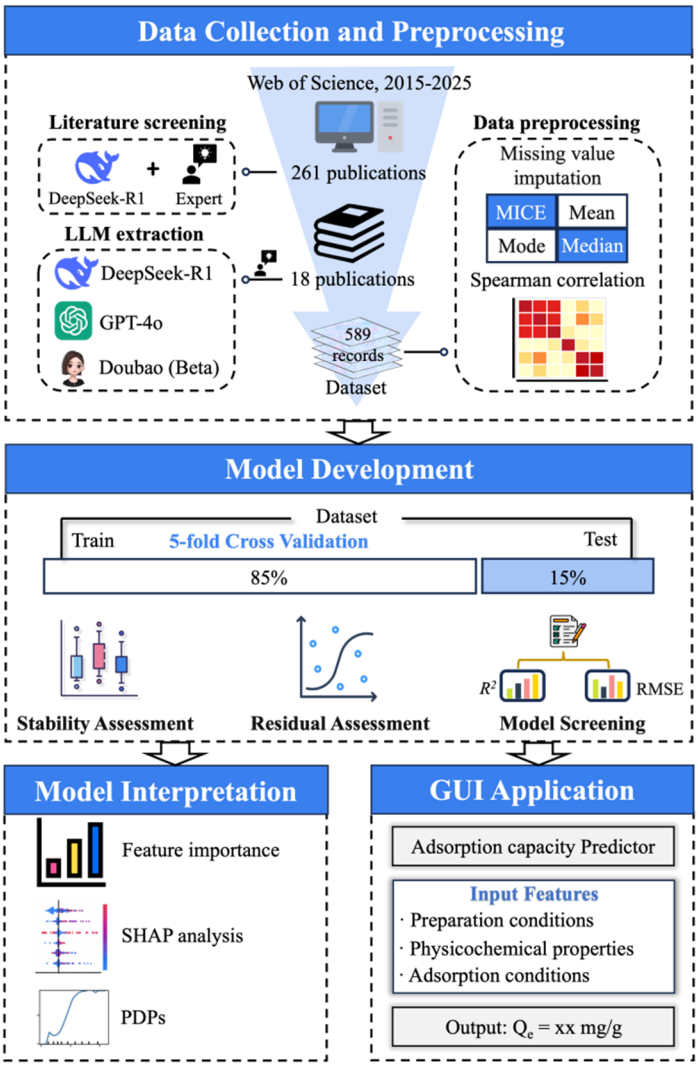

Figure 1. A schematic overview of the study workflow for predicting the adsorption capacity of acid-modified biochar for uranium using AI technology.

2.2. Machine Learning Modeling

To predict the uranium adsorption capacity of acid-modified biochar, four ensemble tree-based machine learning models (Random Forest, XGBoost, CatBoost, and LightGBM) were selected for performance evaluation. Ensemble tree-based approaches were adopted as the modeling framework due to their established advantages in capturing complex nonlinear relationships and enhancing prediction robustness [35,36]. Specifically, compared to conventional single-model or non-ensemble techniques, tree-based ensemble methods combine predictions from multiple decision trees through bagging (e.g., Random Forest) or boosting (e.g., XGBoost, CatBoost, LightGBM) strategies [37]. This integration effectively reduces both bias and variance, improves generalization performance on unseen data, and has consistently achieved superior predictive accuracy across a wide range of applications [33]. This modeling choice is further supported by previous studies in which ensemble tree-based models have proven effective in predicting adsorption-related properties in biochar systems, aligning closely with the objectives of the present work [25,32,34].

The dataset was partitioned into training and test sets using three different split ratios (70:30, 80:20, and 85:15) to identify the optimal partitioning strategy. Based on preliminary evaluations, the 85:15 ratio was found to yield the best performance (Table S2). Therefore, this ratio was adopted for the final model training. Hyperparameter optimization for all models was performed using the Optuna framework with the Tree-structured Parzen Estimator Sampler (TPESampler), which enhances prediction accuracy and model stability [38,39]. The hyperparameters and their corresponding tuning ranges are provided in Table S3. To mitigate overfitting during model training, a 5-fold cross-validation procedure was applied across all ensemble models.

2.3. Model Performance Evaluation

Predictive performance was rigorously assessed using the coefficient of determination (R^2) and root mean square error (RMSE), calculated according to Equations (1) and (2), respectively:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - y_i^{pred})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_i^{pred})^2}{n}} \quad (2)$$

Here, y_i denotes the actual value, y_i^{pred} represents the predicted value, and $\overline{y}$ is the mean of the actual values. Higher R^2 and lower RMSE values indicate better model performance [35]. Therefore, priority should be given to those with a higher R^2 and a lower RMSE based on the model's performance on the test set.

To evaluate model stability, a repeated 5-fold cross-validation approach was employed. The entire dataset was randomly split 50 times, and for each split, a 5-fold cross-validation was conducted. The R^2 and RMSE values for both training and test folds were recorded, generating a distribution of performance metrics for each model. This process allowed for the assessment of model robustness against data sampling variability. The model exhibiting the smallest variance in performance metrics across the 50 iterations was considered the most stable.

In addition, we performed a random-seed sensitivity analysis to ensure that the reported performance is not dependent on a single stochastic initialization. Specifically, we re-trained the optimized CatBoost model under multiple random seeds while keeping the train/test split strategy unchanged, and summarized the distribution of R^2 and RMSE across seeds (Table S5).

Finally, residual analysis was conducted to diagnose model behavior and identify potential biases. Residuals (defined as the differences between experimentally observed values and model-predicted values) were computed for test set predictions. The distribution of residuals was visually examined using scatter plots of residuals versus predicted values. A well-performing model is expected to display residuals randomly dispersed around zero, without observable patterns or systematic deviations.

2.4. Model Behavior Interpretation

To elucidate the underlying mechanisms governing uranium adsorption, the relationships between input features (preparation conditions, physicochemical properties and adsorption conditions) and the target variable (Q_e) were systematically analyzed. This analysis aimed to quantify the influence of individual factors on adsorption performance. Initially, feature importance analysis was performed to evaluate the relative contribution of each input variable to Q_e. Furthermore, SHAP analysis was employed to quantify the specific impact of individual features on Q_e predictions, thereby identifying key variables governing adsorption capacity and elucidating the direction of their effects [31]. To provide a comprehensive interpretation of feature contributions, SHAP

dependence plots and PDPs were utilized, illustrating how individual input variables influence model outputs across their value ranges. The integration of these interpretability techniques systematically deciphered the mechanisms by which preparation conditions, biochar characteristics, and adsorption parameters collectively affected uranium adsorption.

3. Results and Discussion

3.1. Automated Literature Screening and Data Extraction

Conventional data acquisition in scientific research is often constrained by its time-consuming and labor-intensive nature. The recent emergence of LLM-based intelligent extraction techniques offers a promising alternative to mitigate these limitations [40]. Through structured human-computer interaction, LLMs can automatically process batches of unstructured documents to extract targeted information [41]. However, existing studies indicate that LLMs may still encounter challenges in accurately interpreting unstructured text, and their extraction efficiency can be suboptimal. These limitations can be partially alleviated through prompt engineering and model fine-tuning [42,43]. Moreover, multi-LLM cross-validation has been proposed as a means to enhance extraction accuracy and reliability, though expert supervision remains indispensable throughout the process.

To address these considerations, we developed a collaborative LLM-expert framework for efficient dataset construction. For literature screening, DeepSeek-R1 was employed to filter publications based on four predefined criteria (e.g., biochar focus and acid modification), as mentioned in Text S1. Subsequent expert validation confirmed near-perfect accuracy, with only one misclassification, demonstrating the model's strong screening capability. Given this high performance, multi-LLM cross-validation was deemed unnecessary at this stage to conserve computational resources.

For data extraction, three LLMs (DeepSeek-R1, GPT-4o, and Doubao (Beta)), were deployed using a pre-validated structured prompt template. Results indicated that multi-LLM cross-validation significantly improved extraction accuracy compared to single-LLM approaches (Figure 2a). Although this method increased processing time (Figure 2b) and computational cost, it remained substantially more efficient than manual extraction. The average extraction scores of biochar preparation conditions and biochar physicochemical properties under multi-LLM cross-validation are summarized in Figure 2c,d, respectively. Notably, the method performed well on textual data but was less effective for graphical content (e.g., pH_{pzc}), reflecting a known limitation of current LLMs in accurately interpreting visual information. Consequently, these feature values and adsorption parameters were all extracted using WebPlotDigitizer to prevent data loss and ensure accuracy.

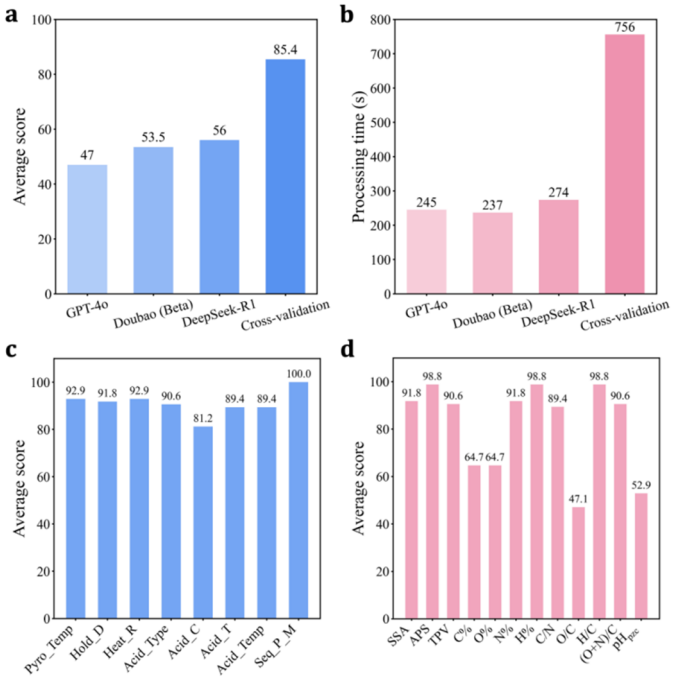

Figure 2. Comparison of extraction performance under the LLM-expert collaboration framework. (**a**) Average scores of different literature extraction strategies; (**b**) Processing time of different literature extraction strategies; (**c**) Average scores of biochar preparation conditions extraction; (**d**) Average scores of biochar physicochemical properties extraction.

The multi-stage quality control procedure we implemented specifically targets the known failure modes of LLMs, such as unit confusion, context drift across tables/figures, and omission of conditional qualifiers, in order to mitigate potential errors introduced by LLM-assisted extraction. Under the multi-LLM cross-validation strategy, the same paper was independently parsed by different LLMs, and the extracted fields were compared to identify disagreements and suspicious values. In addition, rule-based validation was applied to all extracted records, including standardized unit normalization, range/logic checks (e.g., non-negative constraints and physically plausible bounds for adsorption capacity and operating conditions), and consistency checks across related variables. Manual verification was conducted on records of inconsistent or rule-based check markers across LLMs, and spot checks were carried out on randomly sampled subsets of papers. These are crucial steps to confirm extraction accuracy.

This integrated closed-loop workflow, which encompasses automated literature screening, intelligent data extraction, and expert validation, overcame the accuracy limitations of single-LLM approaches for complex scientific data and established a reusable framework. Within this structure, LLMs handled standardized data processing, while experts provided critical validation and knowledge structuring. This collaborative effort yielded a high-quality dataset on biochar uranium adsorption, comprising 589 rigorously validated records, which provides a robust foundation for subsequent model development and mechanistic analysis.

3.2. Dataset Characteristics and Preprocessing

The constructed dataset incorporates three key components: biochar preparation conditions, physicochemical properties, and adsorption conditions. As summarized in Table 1, the initial dataset comprises 24 input features and 1 target variable. Pyrolysis temperature is one of the key parameters for regulating the uranium adsorption performance of biochar. In the dataset, the pyrolysis temperatures are mainly distributed between 500 and 700°C. Within this temperature range, the organic matter in biomass can be fully decomposed, forming abundant microporous and mesoporous structures.

Table 1. Descriptive statistics of parameters in the original dataset.

Variable Name	Type	Mean	Std	Min	Max	Data Missing
Pyrolysis temperature (°C)	Numerical	543.5	185.02	180	850	0%
Hold duration (h)	Numerical	2.04	1.41	0.5	8	0%
Heating rate (°C/min)	Numerical	4.42	4.54	0	10	0%
Acid_Type	Categorical	4.15	1.19	1	6	0%
Acid concentration (mol/L)	Numerical	3.83	3.50	0.16	14.6	4%
Acid treatment time (h)	Numerical	13.82	9.74	0.5	24	4%
Acid treatment temperature (°C)	Numerical	46.15	25.66	25	90	0%
Sequence of pyrolysis and modification	Categorical	0.65	0.49	0	1	0%
SSA (m^2/g)	Numerical	573.81	483.77	17.05	1521.38	4%
APS (nm)	Numerical	5.22	2.38	2.2	9.6	54%
TPV (cm^3/g)	Numerical	0.56	0.45	0.089	1.71	12%
C (%)	Numerical	70.49	22.76	1.21	91.74	23%
O (%)	Numerical	16.17	7.63	2.45	41.3	23%
N (%)	Numerical	3.56	1.79	2.4	7.5	69%
H (%)	Numerical	2.1	0.85	1.5	2.7	92%
C/N	Numerical	18.95	7.84	10.3	28.4	69%
O/C	Numerical	0.34	0.44	0.08	2.02	27%
H/C	Numerical	0.67	0.05	0.63	0.7	92%
(O + N)/C	Numerical	0.31	0.11	0.12	0.51	69%
pH_{pzc}	Numerical	2.79	1.43	1	7	35%
pH	Numerical	4.86	1.53	1	12	0%
T (K)	Numerical	303.54	13.32	278	338	0%
C_0 (mg/L)	Numerical	67.65	68.82	0.5	500	0%
SLR (g/L)	Numerical	0.63	1.92	0.02	20	0%
Q_e (mg/g)	Numerical	146.18	175.99	0.8	1030	0%

Furthermore, a high specific surface area (SSA) combined with sufficient surface micropores provides high-density adsorption sites, which are essential for effective uranium uptake. Following acid modification, the SSA of biochar was significantly enhanced, with an average value of 553.81 m^2/g and a maximum of 1521.38 m^2/g. The oxygen-to-carbon (O/C) ratio, which directly reflects the abundance of oxygen-containing functional groups, influences the surface charge distribution and adsorption selectivity of biochar. The O/C ratios in the dataset are centered around 0.34, indicating a richer presence of surface functional groups compared to unmodified biochar, as supported by a previous study [34]. Notably, the uranium adsorption capacity of acid-modified biochar

demonstrated considerable potential, with an average Q_e of 146.18 mg/g and a maximum value reaching 1030 mg/g. These results underscore the efficacy of acid-modified biochar as a promising adsorbent for uranium removal.

To address missing values in the dataset, we systematically compared four imputation strategies (Figure 3). For variables with <50% missingness, MICE provided the best overall balance between information retention and distributional fidelity, producing imputed values that were most consistent with the observed data. By contrast, when missingness exceeded 50%, none of the tested methods reproduced the original distribution satisfactorily (Figure S1), and these variables were therefore removed to avoid introducing substantial imputation-driven bias and to preserve dataset reliability. On the one hand, H and N-related descriptors are not consistently reported across studies, which limits cross-study comparability. On the other hand, compared with pH-dependent surface charge and the availability of reactive binding sites governed by oxygen-containing functional groups and structural properties, derived ratios such as H/C and C/N mainly act as rough representatives for aromaticness/condensation or N content and cannot directly quantify uranium binding capacity, which is relatively consistent with the current understanding in this field [46,47,58]. Therefore, we excluded these variables and prioritized better-covered, mechanistically interpretable descriptors to improve dataset consistency and model robustness. The distributions of the retained variables without missing values are provided in Figure S2.

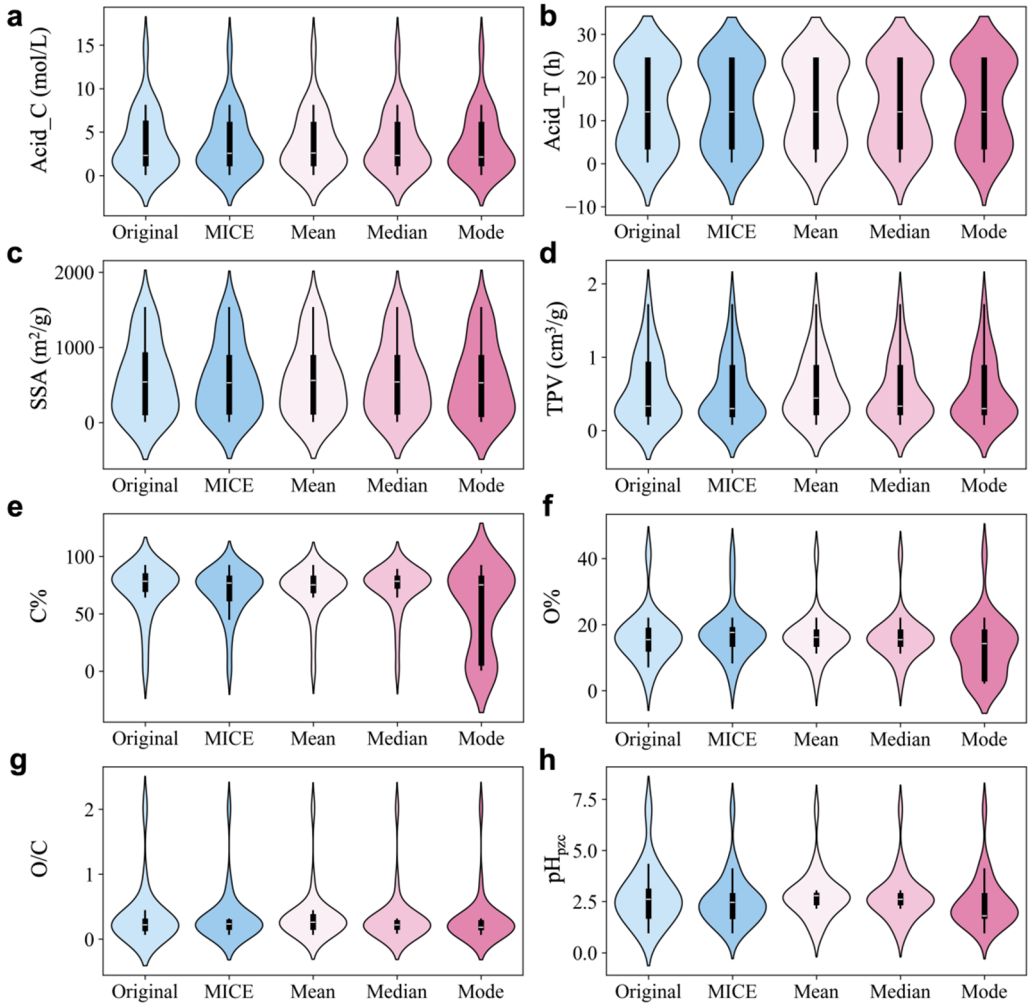

Figure 3. Performance comparison of different missing-value handling strategies (MICE: multiple imputation by chained equations; Mean: mean imputation; Median: median imputation; Mode: mode imputation). (**a**) Acid concentration (Acid_C); (**b**) Acid treatment time (Acid_T); (**c**) SSA; (**d**) TPV; (**e**) C%; (**f**) O%; (**g**) O/C; (**h**) pH_{pzc}.

Notably, feature scaling was intentionally omitted in the preprocessing pipeline, as tree-based ensemble algorithms are inherently robust to features with varying scales. To assess inter-feature relationships, Spearman correlation analysis was conducted. Given the relative insensitivity of ensemble tree models to moderate multicollinearity, a conservative threshold of $|r| > 0.9$ was adopted for feature selection. As shown in Figure 4, a strong correlation was observed between SSA and TPV ($r = 0.91$). Based on its documented greater influence on active site availability and adsorption performance [35], SSA was retained while TPV was removed. Similarly,

elemental composition metrics (C% and O%) were excluded while O/C ratio was retained, as it more directly reflects the relative abundance of oxygen-containing functional groups that govern biochar adsorption mechanisms. This feature refinement process yielded a final set of 16 variables for subsequent model development.

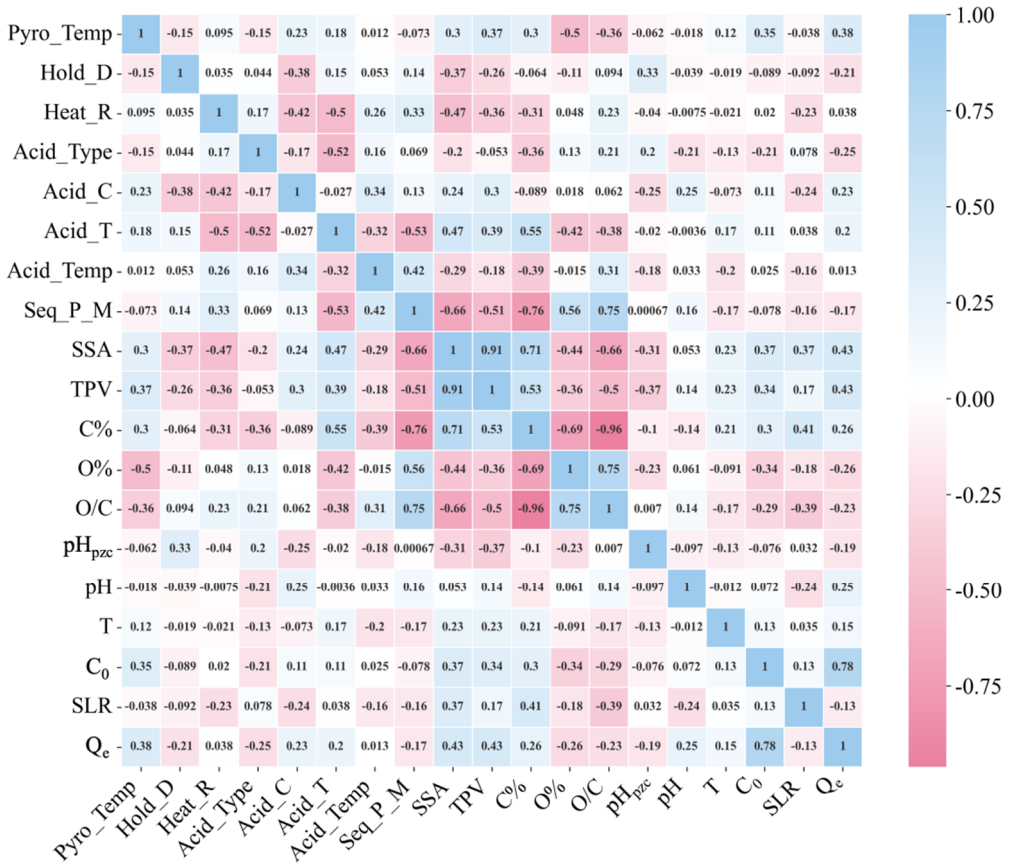

Figure 4. Spearman correlation coefficient matrix between input features and target variable.

3.3. Predictive Modeling of Uranium Adsorption

3.3.1. Model Performance and Evaluation

The predictive performance of four machine learning models for uranium adsorption capacity was systematically evaluated. As shown in Figure 5, the close alignment between predicted and observed values, together with the comparable distributions of training and test predictions, indicates that all models learned stable relationships from the dataset rather than memorizing the training samples. Consistently high test performance (R^2 > 0.95 for all models) further suggests that the models generalize well to test data under the same data-generation conditions. In particular, CatBoost achieved the best accuracy (R^2 = 0.98, RMSE = 24.78), meaning it explains most of the variance in adsorption capacity while maintaining a relatively small absolute prediction error in the same physical units as Q_e. Importantly, the absence of an obvious train-test performance gap in Figure 5 provides a qualitative but direct indication that overfitting is limited, supporting the reliability of these metrics within the data-supported domain. The superior performance of CatBoost is consistent with its ordered boosting strategy and its ability to reduce prediction shift, which can improve robustness compared with conventional gradient-boosting approaches [44]. Although XGBoost, Random Forest, and LightGBM also yielded high prediction accuracy, initial evaluations indicated that CatBoost possessed greater predictive accuracy and robustness.

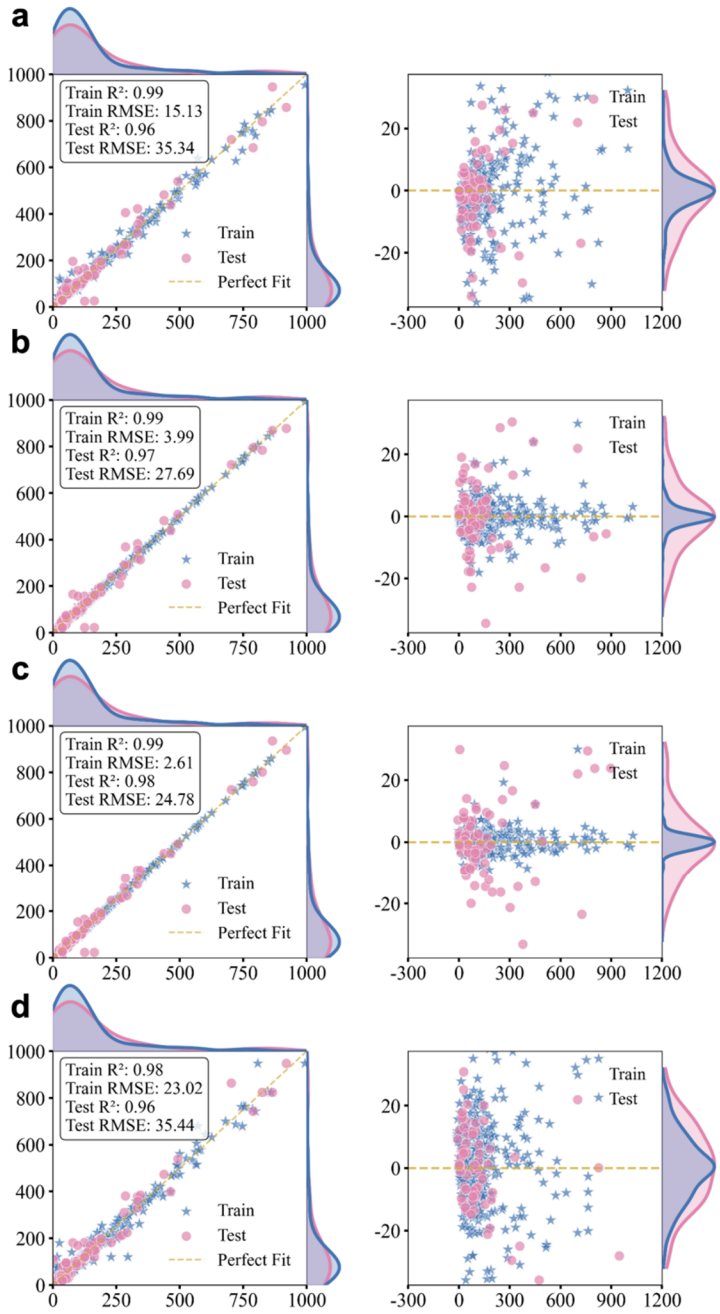

Figure 5. The model performance and residual analysis of (**a**) RF, (**b**) XGBoost, (**c**) CatBoost and (**d**) LightGBM. The density plots on the upper and right sides of the model performance plots respectively represent the distribution of the predicted values and actual values.

To further assess model stability, a repeated 5-fold cross-validation was conducted over 50 independent data splits (Figure 6). CatBoost again demonstrated the most consistent and reliable performance, showing the lowest variability across evaluation metrics, which confirms its strong generalization capacity under varying data sampling conditions. Random-seed sensitivity analysis indicated that the model performance remains stable across different seeds (Table S5), supporting the robustness of our conclusions. Residual analysis provided additional insights into model behavior (Figure 5). The residuals of the CatBoost model were distributed randomly around zero with negligible systematic bias, whereas the other models displayed more pronounced deviation patterns. This observation further supports that CatBoost produces well-calibrated predictions, which is an essential characteristic for credible adsorption performance estimation in real-world scenarios.

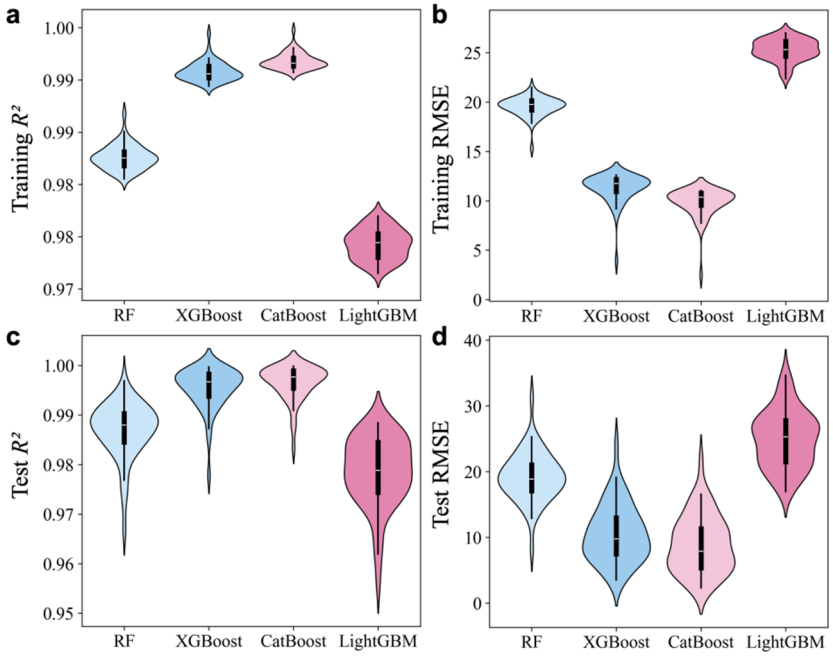

Figure 6. Stability evaluation results of different machine learning models (**a**: training R^2, **b**: training RMSE, **c**: test R^2, **d**: test RMSE).

Based on its exceptional accuracy, robustness, and minimal prediction bias, CatBoost was selected as the final model. Overall, these model evaluation and optimization efforts significantly improved prediction reliability, enabling more efficient assessment of uranium adsorption by biochar and supporting more cost-effective applications in adsorption-related research.

3.3.2. Mechanism Interpretation

This study integrated feature importance and SHAP analysis to systematically identify the key factors regulating uranium adsorption by acid-modified biochar. As shown in Figure 7a, feature importance analysis quantified the relative contribution of each input feature to CatBoost model's predictions, revealing that initial uranium concentration (C_0) was the most dominant factor, accounting for 54.7% of the total importance of all input features. This proportion far exceeds other parameters in the same category, confirming its leading role even among adsorption conditions. Specifically, C_0 directly enhances total adsorption capacity by increasing the contact probability between UO_2^{2+} and active adsorption sites on the biochar surface. Higher C_0 creates a stronger mass transfer driving force for UO_2^{2+}, promoting the occupation of available adsorption sites [45].

Quantitative analysis showed that adsorption conditions accounted for the largest proportion of contributions to uranium adsorption capacity (71.51%), followed by biochar preparation conditions (15.27%) and physicochemical properties (13.22%). The dominant role of adsorption conditions is attributed to their direct regulation of two critical processes: (1) the accessibility of active adsorption sites for UO_2^{2+}, and (2) the thermodynamic and kinetic behaviors of uranium. Both processes are influenced by key factors such as solid-to-liquid ratio (SLR) and solution pH [46,47].

Furthermore, SHAP analysis was used to quantify the importance and influence direction of each feature by evaluating its marginal contribution to model predictions (Figure 7b,c). In the plots, each point corresponds to a feature value of a sample, with red for high values and blue for low values. The size of the SHAP values reflects their impact strength: larger values represent stronger positive contributions to adsorption capacity, while smaller values indicate stronger negative effects. For example, higher values of C_0 and SSA were associated with larger SHAP values, corresponding to greater adsorption, whereas lower values of the pyrolysis and modification sequence led to smaller SHAP values and reduced adsorption. The y-axis arranges the features in order of their relative importance, from high to low, which is roughly consistent with the built-in feature importance ranking of the CatBoost model. However, some features like Acid_C and pH showed SHAP values that did not vary consistently with their own feature values, indicating non-linear relationships that require further investigation.

SHAP dependence plots and PDPs were used to explore the relationships between input features and Q_e, revealing how individual features influence the target variable. Figure 7d–i showed the SHAP dependence plots for the six most important features identified by SHAP method, and the SHAP dependence plots for the remaining

features were shown in Figure S3. For these key features, the trends observed in both the SHAP dependence and PDPs plots (Figure S4) are highly consistent. This consistency highlighted the internal coherence of the model's feature interpretation and indicated robust relationships between these features and Q_e.

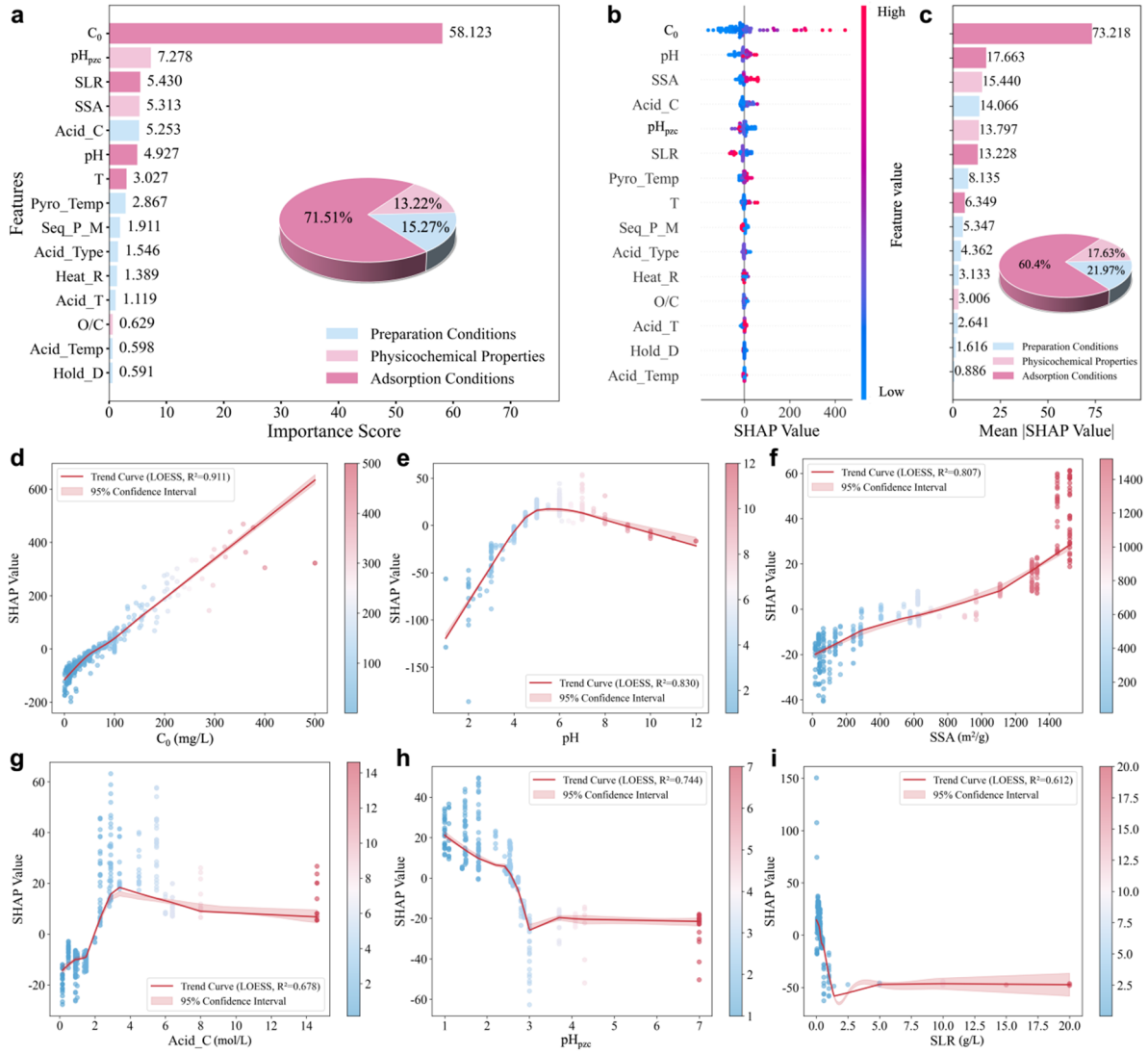

Figure 7. Global feature importance and SHAP-based interpretation of input variables in the CatBoost model. (**a**) feature importance ranking and grouped contribution; (**b**) SHAP beeswarm plot; (**c**) Mean |SHAP| values and grouped contribution; (**d**–**i**) SHAP dependence plots for the six most important features.

Among adsorption conditions, Q_e showed a clear positive correlation with the C_0 and a clear negative correlation with SLR, while exhibiting a non-linear relationship with pH. Figure 7d demonstrated that as C_0 increases, Q_e first rises rapidly, then slows down and gradually plateaus, following a three-stage pattern: rapid increase, deceleration and saturation [48,61]. It is widely accepted that under the same initial adsorption conditions, increasing C_0 significantly boosts Q_e [49,50,62]. Additionally, Figure 7i showed that as SLR increases, Q_e gradually decreases. This occurs because SLR reflects the amount of biochar added per unit volume of solution, while Q_e measures the adsorption capacity per unit mass of biochar. When SLR rises, the relative amount of uranium contacting each unit mass of biochar declines, leading to lower Q_e [51]. However, the uranium removal rate actually increases in this case. Generally, adding more adsorbent improves adsorption performance, but the cost of practical water treatment must be considered. The non-linear SHAP patterns observed for pH reflect the coupled effects of surface charge regulation and uranium speciation. Specifically, pH significantly influences uranium adsorption by altering the surface charge of the adsorbent and the speciation of the adsorbate through electrostatic interactions [52]. Figure 7e shows that: when pH < 5, adsorption capacity increases with rising pH, likely due to electrostatic attraction between positively charged UO_2^{2+} species and oxygen-containing functional groups on the biochar surface [53]; when pH is in the range of 5–7, uranium removal efficiency is optimal; further

increasing pH reduces removal efficiency, as U(VI) transforms into negatively charged ions (e.g., $(UO_2)_3(OH)_7^-$ and $(UO_2)(OH)_3^-$) that generate electrostatic repulsion with the biochar surface, inhibiting adsorption [54].

The physicochemical properties of biochar also play a key role in uranium adsorption. A higher SSA and a lower pH_{pzc} facilitate uranium adsorption on the biochar surface (Figure 7f,h). A larger SSA provides more active sites for adsorption, enhancing interfacial interactions [55]. For example, Li et al. [56] synthesized phosphorus-doped biochar from pomelo peel, achieving a maximum SSA of 1521 m^2/g and a uranium adsorption capacity of 603 mg/g. Additionally, pH_{pzc} helps determine the optimal environmental pH for uranium adsorption by biochar: when pH is below pH_{pzc}, the biochar surface carries a positive charge; when pH is above pH_{pzc}, it carries a negative charge. Typically, the pH_{pzc} of biochar is generally below 7 [57]. Therefore, within the optimal pH range (approximately 5–7), a lower pH_{pzc} makes the biochar surface more likely to carry a negative charge, enhancing electrostatic interactions with uranium and improving adsorption performance [58].

Appropriately increasing the acid concentration during the modification process helps improve the adsorption performance of biochar (Figure 7g). Additionally, Figure S3h suggested that phosphoric acid is a suitable modifier—phosphoric acid modification not only increases the specific surface area and porosity of biochar (providing more adsorption sites), but also introduces abundant oxygen-containing functional groups (e.g., -COOH, -OH). These groups can coordinate and complex with UO_2^{2+}, enhancing adsorption [59]. Meanwhile, phosphoric acid modification reduces the pH_{pzc} of biochar, making its surface more negatively charged across a wider pH range and strengthening electrostatic adsorption of uranium. This study also suggests that the order of pyrolysis and modification is associated with adsorption performance: biochar produced by modification before pyrolysis tends to show higher uranium adsorption in the compiled dataset (Figure S3i). A plausible explanation is that pre-activation can optimize surface properties prior to carbonization through targeted functional-group introduction and pore development [60]. The sample count of each sequence is reported in Table S4, and the balance of the data further supports this claim.

It should be emphasized that SHAP does not establish causal relationships; rather, it explains how the trained model attributes prediction changes to variations in input features based on patterns learned from the dataset. Therefore, the observed SHAP trends are interpreted as data-driven associations that are evaluated for their physical and chemical plausibility by comparison with established adsorption mechanisms reported in the literature.

3.3.3. GUI for Predictive Application

To bridge the gap between predictive modeling and practical implementation, we developed a user-friendly GUI application using Python, which integrates the optimized CatBoost model (Figure 8). This interactive platform allows users to input 15 key feature parameters to predict the uranium adsorption capacity of acid-modified biochar. The interface is organized into three dedicated input modules, logically structured to reflect the experimental workflow. Users can enter values within suggested ranges (which are basically consistent with the span of the dataset in this study) and obtain instant predictions by clicking the “Predict” button. Although the GUI does not enforce strict bounds to permit limited extrapolation, which can be useful for exploratory experimental design, users are cautioned against extensive extrapolation as it may notably reduce prediction reliability.

The resulting adsorption capacity is shown along with a confidence interval reported using quantile regression. Quantile regression provides a more robust way to estimate the uncertainty compared to the method based on the predictive variance of the CatBoost model. This approach gives a better indication of the range within which the true adsorption capacity is likely to lie, especially when dealing with data points that are far from the training data domain. This design strikes a balance between flexibility and a safeguard against unreliable extrapolation. Moreover, it eliminates the need for programming knowledge, making the tool accessible to a wide range of researchers and practitioners.

By providing an efficient and user-friendly pathway to leverage machine learning predictions, this GUI serves as a practical bridge between data-driven modeling and real-world uranium adsorption studies. It significantly accelerates experimental design and optimization, contributing to more intelligent and resource-efficient research in radioactive wastewater treatment.

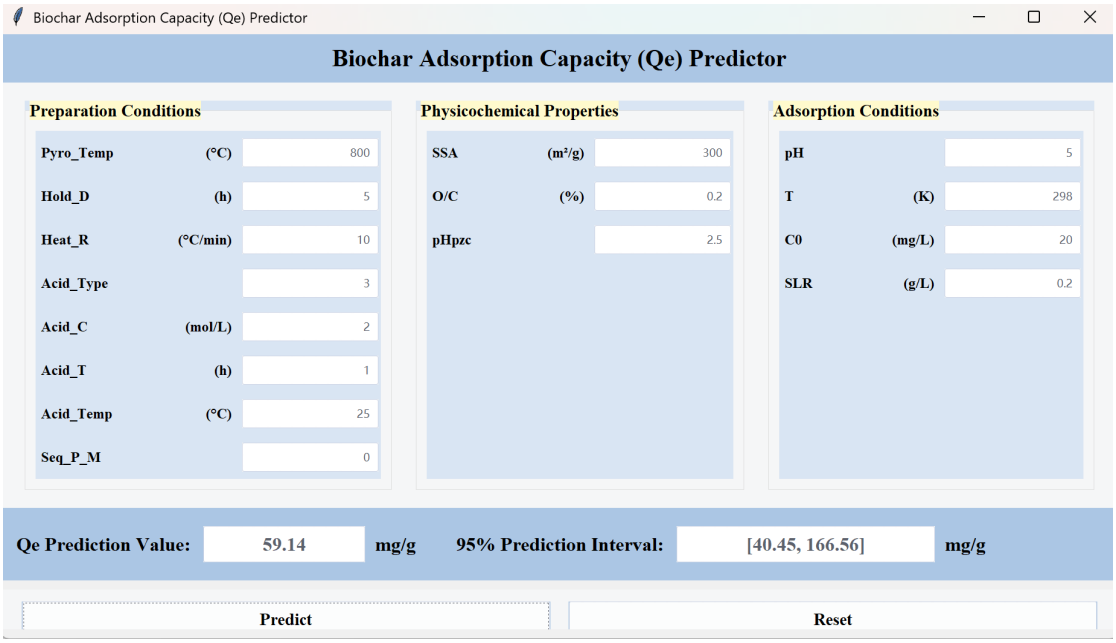

Figure 8. A user-friendly GUI application for the rapid prediction of uranium adsorption capacity on acid-modified biochar using the optimized CatBoost model. https://github.com/stephanieee528/Biochar_Adsorption_Predictor (accessed on 17 December 2025).

4. Conclusions

This study presents an integrated AI-driven framework that combines large language models for automated data extraction with interpretable machine learning to predict uranium adsorption performance on acid-modified biochar, addressing a key challenge in radioactive wastewater management. Among the tested models, CatBoost achieved the best overall performance and was selected as the primary predictor. While adsorption conditions dominate the prediction, feature importance and SHAP analyses consistently identify pH_{pzc}, SSA, and acid treatment-related variables, underscoring the coupled effects of surface charge regulation and reactive site availability. These patterns provide practical guidance for the rational design of adsorbents and for prioritizing experimental data collection toward high-impact variables.

Several limitations should be noted. Using only tree-based ensemble models is difficult to integrate physical constraints and is limited in handling high-dimensional data. Future research may consider neural networks and hybrid physics-informed ML approaches, which will be explored after the database is expanded. In addition, while the workflow is pollutant-agnostic in principle, its application to new pollutants requires building dedicated datasets and retraining due to uranium-specific features and learned relationships.

Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/260106155037017 1/EMT-25120049-Supplementary-Materials.pdf. Text S1: Predefined criteria for literature screening. Text S2: System Prompt. Table S1: Quantitative scoring rule for data extraction using LLMs. Table S2: Four tree-based ensemble ML model prediction results under different division ratios. Table S3: The hyperparameters and their ranges of the model used in this study. Table S4: Sample counts for the two preparation sequences used to define Seq_P_M in the compiled dataset. Table S5: Seed sensitivity analysis of the CatBoost model performance. Figure S1: Performance comparison of MICE, Mean, Median and Mode in feature missing value imputation with missing ratios exceeding 50%. Figure S2: The distribution of the remaining features. The histogram represents the distribution of category features. Figure S3: The SHAP dependence plots of the remaining features. Figure S4: Model interpretation based on partial dependence plots of uranium adsorption capacity on input features.

Author Contributions

J.S.: Writing—review & editing, Writing—original draft, Methodology, Software, Formal analysis, Data curation, Visualization. S.I.C.: Software, Visualization. S.Z.: Writing—review & editing, Supervision, Funding acquisition, Conceptualization. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (Grant No. 52500105).

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

GUI application code is available at [https://github.com/stephanieee528/Biochar_Adsorption_Predictor (accessed on 17 December 2025)]. Data supporting the findings of this study will be available from the corresponding author upon reasonable request.

Acknowledgments

The authors gratefully acknowledge the financial support by the National Natural Science Foundation of China (Grant No. 52500105). This research was supported by Public Computing Cloud, Renmin University of China.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

References

1. Wang, J.; Zhuang, S. Extraction and adsorption of U(VI) from aqueous solution using affinity ligand-based technologies: an overview. *Rev. Environ. Sci. Biotechnol.* **2019**, *18*, 437–452.
2. Li, Z.-L.; Li, S.-F.; Zhang, Z.-M.; et al. Extracellular electron transfer-dependent bioremediation of uranium-contaminated groundwater: Advancements and challenges. *Water Res.* **2025**, *272*, 122957.
3. Zhang, D.; Fang, L.; Liu, L.; et al. Uranium extraction from seawater by novel materials: A review. *Sep. Purif. Technol.* **2023**, *320*, 124204.
4. Boussouga, Y.-A.; Joseph, J.; Stryhanyuk, H.; et al. Adsorption of uranium(VI) complexes with polymer-based spherical activated carbon. *Water Res.* **2024**, *249*, 120825.
5. Chen, D.; Ren, Z.; Shi, M.; et al. Covalently constructed porous polyamidoxime nanofibers for enhanced uranium capture. *Water Res.* **2025**, 124944.
6. Liu, Y.-L.; Cao, P.; Zhang, Q.; et al. Solar-activated ZnS@MXene heterostructure for integrated radioactive wastewater treatment and energy harvesting. *Water Res.* **2025**, *286*, 124254.
7. Bone, S.E.; Cliff, J.; Weaver, K.; et al. Complexation by Organic Matter Controls Uranium Mobility in Anoxic Sediments. *Environ. Sci. Technol.* **2020**, *54*, 1493–1502.
8. Sun, J.; Yi, X.; Yuan, C.; et al. Adsorptive removal of radioactive technetium by nanomaterials. *Rev. Environ. Sci. Biotechnol.* **2025**, *24*, 115–144.
9. Xie, Y.; Chen, C.; Ren, X.; et al. Emerging natural and tailored materials for uranium-contaminated water treatment and environmental remediation. *Prog. Mater. Sci.* **2019**, *103*, 180–234.
10. Guilhen, S.N.; Mašek, O.; Ortiz, N.; et al. Pyrolytic temperature evaluation of macauba biochar for uranium adsorption from aqueous solutions. *Biomass Bioenergy* **2019**, *122*, 381–390.
11. Ravindiran, G.; Rajamanickam, S.; Janardhan, G.; et al. Production and modifications of biochar to engineered materials and its application for environmental sustainability: A review. *Biochar* **2024**, *6*, 62.
12. Xiong, X.; Liu, J.; Xiao, T.; et al. Remediation of uranium-contaminated water and soil by biochar-based materials: A review. *Biochar* **2025**, *7*, 41.
13. Huang, F.; Dong, F.; Chen, L.; et al. Biochar-mediated remediation of uranium-contaminated soils: evidence, mechanisms, and perspectives. *Biochar* **2024**, *6*, 16.
14. Jun, B.-M.; Jung, J.-Y.; Oh, M.; et al. Uranium removal from radioactive wastewater using biochar-based adsorbents: A review on synthesis, performance, and mechanism. *J. Water Process Eng.* **2025**, *75*, 107956.

15. Mei, Y.; Zhuang, S.; Wang, J. Adsorption of heavy metals by biochar in aqueous solution: A review. *Sci. Total Environ.* **2025**, *968*, 178898.
16. Paschalidou, P.; Pashalidis, I.; Manariotis, I.D.; et al. Hyper sorption capacity of raw and oxidized biochars from various feedstocks for U(VI). *J. Environ. Chem. Eng.* **2020**, *8*, 103932.
17. Ahmed, W.; Mehmood, S.; Qaswar, M.; et al. Oxidized biochar obtained from rice straw as adsorbent to remove uranium(VI) from aqueous solutions. *J. Environ. Chem. Eng.* **2021**, *9*, 105104.
18. Sharma, M.; Anshika; Yadav, L.; et al. Breaking new ground: Innovative adsorbents for uranium and thorium ions removal and environmental cleanup. *Coord. Chem. Rev.* **2024**, *517*, 216008.
19. Li, M.; Liu, H.; Chen, T.; et al. Synthesis of magnetic biochar composites for enhanced uranium(VI) adsorption. *Sci. Total Environ.* **2019**, *651*, 1020–1028.
20. Liu, F.; Wang, S.; Zhao, C.; et al. Constructing coconut shell biochar/MXenes composites through self-assembly strategy to enhance U(VI) and Cs(I) immobilization capability. *Biochar* **2023**, *5*, 31.
21. Zheng, Z.; Rampal, N.; Inizan, T.J.; et al. Large language models for reticular chemistry. *Nat. Rev. Mater.* **2025**, *10*, 369–381.
22. Li, Y.; Gupta, R.; You, S. Machine learning assisted prediction of biochar yield and composition via pyrolysis of biomass. *Bioresour. Technol.* **2022**, *359*, 127511.
23. Shi, Y.; Rampal, N.; Zhao, C.; et al. Comparison of LLMs in extracting synthesis conditions and generating Q&A datasets for metal–organic frameworks. *Digit. Discov.* **2025**, *4*, 2676–2683.
24. Polak, M.P.; Morgan, D. Extracting accurate materials data from research papers with conversational language models and prompt engineering. *Nat. Commun.* **2024**, *15*, 1569.
25. Palansooriya, K.N.; Li, J.; Dissanayake, P.D.; et al. Prediction of Soil Heavy Metal Immobilization by Biochar Using Machine Learning. *Environ. Sci. Technol.* **2022**, *56*, 4187–4198.
26. Zhu, X.; Wan, Z.; Tsang, D.C.W.; et al. Machine learning for the selection of carbon-based materials for tetracycline and sulfamethoxazole adsorption. *Chem. Eng. J.* **2021**, *406*, 126782.
27. Zhao, S.; Guo, J.; Tang, Y.; et al. Applications of machine learning in heavy metal adsorption modeling: A review. *Sep. Purif. Technol.* **2025**, *377*, 134168.
28. Li, J.; Pan, L.; Huang, Y.; et al. Biochar design for antibiotics adsorption via a hybrid machine-learning-based optimization framework. *Sep. Purif. Technol.* **2024**, *348*, 127666.
29. Fabregat-Palau, J.; Ershadi, A.; Finkel, M.; et al. Modeling PFAS sorption in soils using machine learning. *Environ. Sci. Technol.* **2025**, *59*, 7678–7687.
30. Jiang, X.; Wang, W.; Tian, S.; et al. Applications of natural language processing and large language models in materials discovery. *npj Comput. Mater.* **2025**, *11*, 79.
31. Fu, W.; Feng, M.; Guo, C.; et al. Machine learning-driven prediction of phosphorus removal performance of metal-modified biochar and optimization of preparation processes considering water quality management objectives. *Bioresour. Technol.* **2024**, *403*, 130861.
32. Liu, C.; Balasubramanian, P.; An, J.; et al. Machine learning prediction of ammonia nitrogen adsorption on biochar with model evaluation and optimization. *NPJ Clean Water* **2025**, *8*, 13.
33. Lyu, H.; Xu, Z.; Zhong, J.; et al. Machine learning-driven prediction of phosphorus adsorption capacity of biochar: Insights for adsorbent design and process optimization. *J. Environ. Manag.* **2024**, *369*, 122405.
34. Da, T.-X.; Ren, H.-K.; He, W.-K.; et al. Prediction of uranium adsorption capacity on biochar by machine learning methods. *J. Environ. Chem. Eng.* **2022**, *10*, 108449.
35. Zhao, F.; Tang, L.; Song, W.; et al. Predicting and refining acid modifications of biochar based on machine learning and bibliometric analysis: Specific surface area, average pore size, and total pore volume. *Sci. Total Environ.* **2024**, *948*, 174584.
36. Hossain, M.M.; Sikder, R.; Hua, G.; et al. From Model Development to Mitigation: Machine Learning for Predicting and Minimizing Iodinated Trihalomethanes in Water Treatment. *Environ. Sci. Technol.* **2025**, *59*, 11638–11652.
37. Mienye, I.D.; Jere, N. A Survey of Decision Trees: Concepts, Algorithms, and Applications. *IEEE Access* **2024**, *12*, 86716–86727.
38. Yang, Q.; Bao, R.; Wang, Z.; et al. Unlocking prediction and optimal design of CO_2 methanation catalysts via active learning-enhanced interpretable ensemble learning. *Chem. Eng. J.* **2025**, *509*, 161154.
39. Yang, W.; Wu, Y.; Jia, W.; et al. Source term inversion of nuclear accident with random release durations based on machine learning. *J. Hazard. Mater.* **2025**, *488*, 137448.
40. Schilling-Wilhelmi, M.; Ríos-García, M.; Shabih, S.; et al. From text to insight: large language models for chemical data extraction. *Chem. Soc. Rev.* **2025**, *54*, 1125–1150.
41. Bran, A.M.; Cox, S.; Schilter, O.; et al. Augmenting large language models with chemistry tools. *Nat. Mach. Intell.* **2024**, *6*, 525–535.

42. Wu, J.; Shi, R.; Zhou, X.; et al. Empowering Chemistry Experts with Large Language Models for Literature Interpretation in Single-Atom Catalysis Toward Advanced Oxidation. *Angew. Chem., Int. Ed.* **2025**, e202520525.
43. Farris, B.R.; Leonard, K.C. Accelerating Catalysis Understanding via Large Language Model Data Extraction and Shallow Machine Learning Techniques. *JACS Au* **2025**, *5*, 5578–5589.
44. Jaffari, Z.H.; Abbas, A.; Lam, S.-M.; et al. Machine learning approaches to predict the photocatalytic performance of bismuth ferrite-based materials in the removal of malachite green. *J. Hazard. Mater.* **2023**, *442*, 130031.
45. Sun, Y.; Yuan, N.; Ge, Y.; et al. Adsorption behavior and mechanism of U(VI) onto phytic acid-modified biochar/MoS_2 heterojunction materials. *Sep. Purif. Technol.* **2022**, *294*, 121158.
46. Hu, R.; Xiao, J.; Wang, T.; et al. Engineering of phosphate-functionalized biochars with highly developed surface area and porosity for efficient and selective extraction of uranium. *Chem. Eng. J.* **2020**, *379*, 122388.
47. Jin, J.; Li, S.; Peng, X.; et al. HNO_3 modified biochars for uranium(VI) removal from aqueous solution. *Bioresour. Technol.* **2018**, *256*, 247–253.
48. Mei, Y.; Zhuang, S.; Wang, J. Biochar: a potential and green adsorbent for antibiotics removal from aqueous solution. *Rev. Environ. Sci. Biotechnol.* **2024**, *23*, 1065–1103.
49. Yu, P.; Li, Y.; Cai, Z.; et al. Simultaneous removal of Cd and ciprofloxacin hydrochloride by ZVI/biochar composite in water: Compound effects and removal mechanism. *Sep. Purif. Technol.* **2023**, *327*, 124821.
50. Zhao, Z.; Li, P.; Zhang, M.; et al. Unlocking the potential of Chinese herbal medicine residue-derived biochar as an efficient adsorbent for high-performance tetracycline removal. *Environ. Res.* **2024**, *252*, 118425.
51. Morshedy, A.S.; Taha, M.H.; El-Aty, D.M.A.; et al. Solid waste sub-driven acidic mesoporous activated carbon structures for efficient uranium capture through the treatment of industrial phosphoric acid. *Environ. Technol. Innov.* **2021**, *21*, 101363.
52. Yi, J.; Huo, Z.; Tan, X.; et al. Plasma-facilitated modification of pumpkin vine-based biochar and its application for efficient elimination of uranyl from aqueous solution. *Plasma Sci. Technol.* **2019**, *21*, 095502.
53. Yu, J.; Zhang, X.; Wang, H.; et al. Upgradation of water hyacinth for decontamination of uranium-containing radioactive wastewater with double environmental benefit. *Colloids Surf., A* **2025**, *705*, 135709.
54. Yu, S.; Wu, X.; Ye, J.; et al. Dual Effect of Acetic Acid Efficiently Enhances Sludge-Based Biochar to Recover Uranium From Aqueous Solution. *Front. Chem.* **2022**, *10*, 835959.
55. Jin, Q.; Cui, J. Fungi-enabled hierarchical porous magnetic carbon derived from biomass for efficient remediation of As(III)-contaminated water and soil: performance and mechanism. *Environ. Sci. Nano* **2023**, *10*, 1297–1312.
56. Li, H.; Zhang, X.; Luo, C.; et al. Pomelo peel derived phosphorus-doped biochar for efficient disposal of uranium-containing nuclear wastewater: Experimental and theoretical perspectives. *Sep. Purif. Technol.* **2024**, *333*, 125947.
57. Dou, S.; Ke, X.-X.; Shao, Z.-D.; et al. Fish scale-based biochar with defined pore size and ultrahigh specific surface area for highly efficient adsorption of ciprofloxacin. *Chemosphere* **2022**, *287*, 131962.
58. Zhou, Y.; Xiao, J.; Hu, R.; et al. Engineered phosphorous-functionalized biochar with enhanced porosity using phytic acid-assisted ball milling for efficient and selective uptake of aquatic uranium. *J. Mol. Liq.* **2020**, *303*, 112659.
59. Zeng, X.-Y.; Wang, Y.; Li, R.-X.; et al. Impacts of temperatures and phosphoric-acid modification to the physicochemical properties of biochar for excellent sulfadiazine adsorption. *Biochar* **2022**, *4*, 14.
60. Sun, Y.; Yue, Q.; Gao, B.; et al. Comparative study on characterization and adsorption properties of activated carbons with H_3PO_4 and $H_4P_2O_7$ activation employing Cyperus alternifolius as precursor. *Chem. Eng. J.* **2012**, 181–182, 790–797.
61. Zare, N.; Wu, T.; Zhang, D.; et al. Efficient removal of Congo red using adsorbents prepared via MOF-on-MOF strategy. *Chem. Eng. J.* **2026**, *527*, 171592.
62. Sharker, T.; Xiao, X.; Muff, J. Hybrid capacitive deionization using MgAl-LDHs-coated graphite felt electrodes for phosphate removal. *Chem. Eng. J. Adv.* **2026**, *25*, 100985.