Active Learning-Based Guided Synthesis of Engineered Biochar for CO_2 Capture

Author(s):
Yuan, Xiangzhou; Suvarna, Manu; Lim, Juin Yau; Pérez-Ramírez, Javier; Wang, Xiaonan; Ok, Yong Sik

Publication date:
2024-04-16

Permanent link:
https://doi.org/10.3929/ethz-b-000668710

Originally published in:
Environmental Science & Technology 58(15), https://doi.org/10.1021/acs.est.3c10922

Environmental Science & Technology

pubs.acs.org/est

Article

Active Learning-Based Guided Synthesis of Engineered Biochar for CO_2 Capture

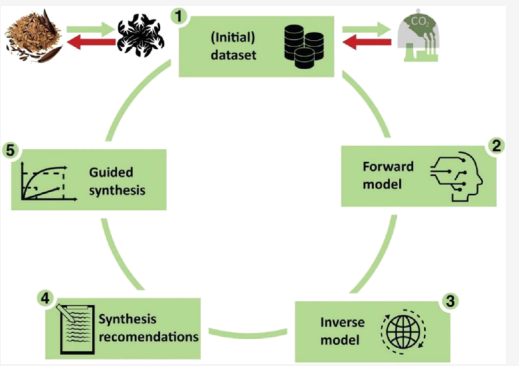
Xiangzhou Yuan,[⊥] Manu Suvarna,[⊥] Juin Yau Lim, Javier Pérez-Ramírez, Xiaonan Wang,* and Yong Sik Ok*

Cite This: *Environ. Sci. Technol.* 2024, 58, 6628−6636

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Biomass waste-derived engineered biochar for CO_2 capture presents a viable route for climate change mitigation and sustainable waste management. However, optimally synthesizing them for enhanced performance is time- and labor-intensive. To address these issues, we devise an active learning strategy to guide and expedite their synthesis with improved CO_2 adsorption capacities. Our framework learns from experimental data and recommends optimal synthesis parameters, aiming to maximize the narrow micropore volume of engineered biochar, which exhibits a linear correlation with its CO_2 adsorption capacity. We experimentally validate the active learning predictions, and these data are iteratively leveraged for subsequent model training and revalidation, thereby establishing a closed loop. Over three active learning cycles, we synthesized 16 property-specific engineered biochar samples such that the CO_2 uptake nearly doubled by the final round. We demonstrate a data-driven workflow to accelerate the development of high-performance engineered biochar with enhanced CO_2 uptake and broader applications as a functional material.



KEYWORDS: *inverse design, machine learning, particle swarm optimization, carbon neutrality, environmental sustainability, UN SDG 13*

1. INTRODUCTION

Carbon capture and storage (CCS) has attracted significant attention to mitigate climate change and limit global warming to 1.5 °C,[1,2] fueled by the steady increase in atmospheric concentration and annual growth rate of CO_2 emissions.[3] Several technologies, including solvent-based absorption, membrane and cryogenic separations, and adsorption over porous solids, have found prominence for postcombustion CO_2 capture, where the latter has emerged as a popular choice owing to its low cost and operational energy requirement, high selectivity, and ease of regeneration.[4−6] Different classes of adsorbents including zeolites, metal organic frameworks (MOFs), activated carbons, metal oxides, and silica have been extensively investigated over the past couple decades.[4,7] Among these, engineered biochar, a form of activated carbon, presents a promising functional material for postcombustion CO_2 capture due to its advantages including abundant carbon precursors (i.e., biomass and organic waste), tunable porosity, mild operating conditions, high feasibility for practical applications, low energy requirement for regeneration, and excellent carbon negative emission properties.[4,8,9] Moreover, upcycling biomass waste into engineered biochar serves as a sustainable waste-to-resource strategy, while the byproduct syngas can be combusted to generate thermal energy for synthesizing engineered biochar.[10,11]

The rational design of high-performing engineered biochar for CO_2 adsorption necessitates pore structure engineering (including the total surface area, pore volume, and pore size distribution), which is accomplished through carbonization and activation processes,[12−15] along with modification strategies such as doping with N, S, or Mg, among others.[16−19] It is well acknowledged that these steps are the controllable synthesis parameters; thus, a wide gamut of their possible parameter space is explored following the trial-and-error approach, which is time- and resource-intensive. This demands novel strategies that are capable of accelerating the experimental efforts to design property-specific engineered biochar materials.

Recent strides in machine learning (ML) have proven effective in guiding the synthesis of application-specific inorganic materials,[20−22] organic compounds,[23,24] adsorbents,[25,26] and membranes.[27] In the biochar community, several ML-based studies have been reported that predicted the yield and physical−chemical properties of biochar,[28−31] organic

Received: December 25, 2023
Revised: February 21, 2024
Accepted: February 22, 2024
Published: March 18, 2024

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

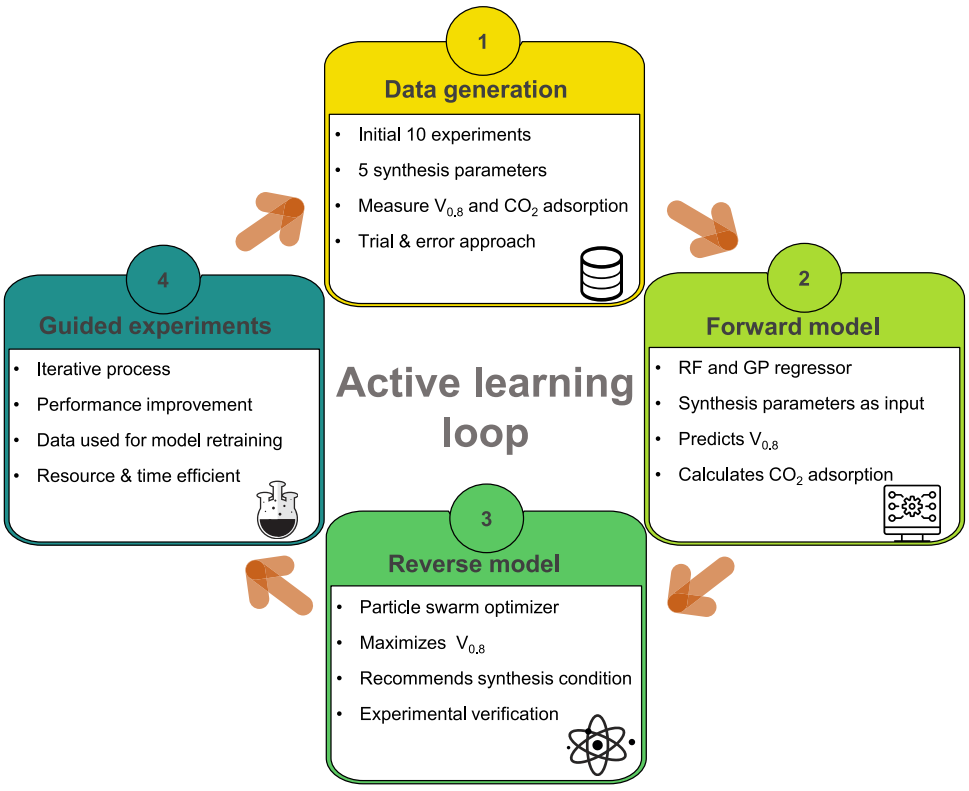


Figure 1. Active learning schematic for upcycling rice husk into engineered biochar for CO_2 capture toward environmental sustainability. Beginning with 10 initial experiments, a forward model is developed which maps $V_{0.8}$ as a function of five synthesis parameters and is correlated to CO_2 adsorption capacity. The inverse model then recommends experiments worth investigating to maximize the performance. Across three active learning cycles, the CO_2 adsorption of the synthesized engineered biochars almost doubles.

contaminant removal,[32−34] heavy metal immobilization,[35−38] and CO_2 adsorption on biochar[39] or leveraged concepts of natural language processing (NLP) to text-mine scientific articles and create a database of synthetic recipes.[40] Notably, although a handful of these studies provide experimental validation,[31,37,41] the majority remain purely theoretical. Thus, there is an evident dearth of studies that utilize ML concepts and tools to synthesize biochar with the desired properties and performance.

In this proof-of-concept, we take the first steps to address this gap. Our previous studies uncovered two critical findings: (i) engineered biochar's narrow micropores under 0.8 nm ($V_{0.8}$) exhibit a linear correlation with CO_2 adsorption[42] and (ii) the textural properties including pore volume and surface area outweigh the functional groups for CO_2 adsorption at 25 °C and 1 bar.[39] Based on these findings, we hypothesize that tailoring $V_{0.8}$ in engineered biochar can achieve the desired CO_2 adsorption capacity.[9,39] To test this hypothesis, we used rice husk as a carbon precursor and created a data-driven framework that identifies optimal synthesis conditions to maximize the CO_2 adsorption capacity. We train a model on an initial set of experimental data, which learns the complex and nonlinear relationship between the synthesis parameters and $V_{0.8}$ of the engineered biochar and in turn recommends the synthesis conditions necessary to maximize $V_{0.8}$, ultimately enhancing the CO_2 adsorption. The data from these experiments are iteratively fed to the model for further training, prediction, and experimental recommendations, thereby establishing a closed active learning loop[43,44] (Figure 1). Across three active learning cycles, we synthesized 16 engineered biochar samples, progressively tuning their $V_{0.8}$ to achieve a nearly doubled CO_2 adsorption capacity by the final round.

2. METHODOLOGY

2.1. Experimental Protocols and Scope. All of the experiments encompassing synthesis, characterization, and CO_2 adsorption tests of engineered biochar samples are categorized into seed and guided experiments. Seed experiments rely on the experimental researcher's intuition and are trial-and-error in nature, while guided experiments follow suggestions from the active learning framework. Regardless of the experiment type, all of the engineered biochar samples are synthesized through rice husk carbonization and KOH activation. It is essential to note that various engineered biochar preparation methods exist, involving activating agents like NaOH and postactivation modifications such as N-, S-, O-, or Mg-doping.[17−19] However, our scope focuses on carbonization followed by KOH activation for consistency in the data. This choice allows our model to learn nonlinear correlations between synthetic parameters and engineered biochar's textural properties, enabling meaningful experiment recommendations. Furthermore, including a wide array of synthesis procedures would increase complexity, posing the challenge of the "dimensionality curse" with too many variables for a small data regime.[43,45]

2.1.1. Synthesis of Engineered Biochar. We chose carbonization followed by KOH activation, a verified and practical method for preparing high-performance CO_2 adsorption engineered biochar samples.[46−48] In this synthesis route, five parameters, including carbonization temperature (T_C) and residence time (R_C), activation temperature (T_A) and

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

residence time (R_A), and mass ratio of KOH and biochar (K/B), were considered essential, as these could be varied or adjusted conveniently during experimentation. Post activation, the engineered biochar samples were collected and washed using distilled water to remove unreacted chemicals and dried at 110 °C for 24 h.

2.1.2. Characterization and CO_2 Adsorption Tests. All of the synthesized engineered biochar samples were characterized to determine their textural properties and evaluated for their CO_2 adsorption capacities. The samples were first pre-degassed at 150 °C for 12 h to remove gases adsorbed in samples, followed by N_2 adsorption and desorption isotherms at −196 °C using the Micromeritics analyzer (ASAP 2020). Thereafter, the total pore volume (V_{Total}) was calculated using the Horvath−Kawazoe equation at p/p_0 = 0.99, and the cumulative pore volume limited by narrow micropores less than 0.8 nm ($V_{0.8}$) was obtained by using nonlocal density functional theory (NLDFT) with the standard slit model. CO_2 adsorption capacities were determined at 25 °C and 1 bar using the volumetric sorption analyzer (ASAP 2020, Micromeritics).

2.2. Active Learning Framework. Our active learning strategy combines the regression ability of ML with the optimization assets of a particle swarm optimizer (PSO) suitable for single objective functions.[49,50] Herein, the former is also referred to as a forward model as it predicts $V_{0.8}$ (property of interest) based on synthesis parameters (input to the model), whereas the latter is termed as an inverse model as it recommends optimal synthesis conditions (output of the optimizer) to maximize $V_{0.8}$ (objective function). Since CO_2 adsorption (performance) linearly correlated to $V_{0.8}$ (property), maximizing the latter invariably increases the former, which is the ultimate goal of this work.

2.2.1. Forward Model. As the first step, a random forest (RF) and Gaussian process (GP) regressor were trained on an initial set of 10 experiments using T_C, R_C, T_A, R_A, and K/B as input features to predict $V_{0.8}$, which was then used to calculate the CO_2 adsorption empirically using the linear relation[42]

$$CO_2 \text{ adsorption} = 8 \times V_{0.8} + 1.21 \quad (1)$$

This combined process of predicting $V_{0.8}$ and empirically calculating the CO_2 adsorption is referred to as the forward model. Subsequently, a feature importance analysis using the SHAP (Shapley Additive exPlanations) analysis was performed on the best-performing algorithm.[44,51]

This approach aims to determine the most important input features that affect the target properties and thereby enhance the interpretability of ML predictions,[49,52] while ensuring alignment with domain knowledge and experimental experience.[43,44,53] The latter is crucial for building user confidence in ML predictions, ultimately facilitating guided synthesis and process optimization.

2.2.2. Inverse Model. The ML algorithm acts as a surrogate to the PSO algorithm, which explores the synthesis parameters within specified boundary conditions to maximize $V_{0.8}$ and is referred to as the inverse model.[25,27,49] In each run of the inverse model, the PSO algorithm selects the next potential set of n conditions (generations) to be tested experimentally, where the number of generations are determined by the modeler.[50] Each batch serves as an iteration in the inverse design strategy, and each condition within a single iteration serves as an experimental data point worth validation.[23,24] The PSO-recommended experimental conditions are tested and simultaneously added to the initial data set, further training the forward model and subsequent PSO to sequentially maximize $V_{0.8}$. This iterative process continues until no further improvements in CO_2 adsorption capacities are observed or until the optimizer reaches a global maximum/minimum.[20,43]

The particulars of the various rounds of active learning cycles are summarized in Table 1. Details of the regressor algorithms, hyperparameter tuning, error metrics, SHAP, and PSO algorithm are described in the Supporting Information (Sections S1−S5). All of the modeling activities were performed in Python (version 3.6) using the open-source libraries: Scikit-learn for developing ML models, Shap for model interpretation, and PySwarm for optimization.

Table 1. Overview of the Active Learning Cycle with the Essential Details[a]

nomenclature	sample prepared	type of experiments	number of experiments
Seed	EB 1 − EB 10	initial	10
*R*1	EB 11 − EB 15	guided	5
*R*2	EB 16 − EB 21	guided	6
*R*3	EB 22 − EB 26	guided	5

[a]Note: EB-*x*, *x* represents the number of each prepared sample.

3. RESULTS AND DISCUSSION

3.1. Seed Experiments and Model Prediction. We initiate the active learning framework through the seed experiments, which comprised 10 experimental conditions from synthesis to the CO_2 adsorption performance of the engineered biochar. The N_2 adsorption and desorption isotherms for 10 engineered biochar samples prepared at −196 °C during the seed experiments are summarized in Table S1 and depicted in Figure S1. All isotherms belong to type I of the International Union of Pure and Applied Chemistry (IUPAC) classification, implying that all of the prepared engineered biochar samples were typical microporous carbon materials.[20,24] Moreover, the pore size distributions (PSDs) of the samples, which are measured following the N_2 isotherms, and the NLDFT with the standard slit models are presented in Figure 2a. For all of the samples in the seed experiments, the microporous structures were well-developed and dominant pore sizes were smaller than 2.0 nm, implying their effectiveness to adsorb CO_2 gas.[10,11] The CO_2 adsorption isotherms at 25 °C for the seed samples are depicted in Figure 2b. Because all of the synthesis conditions were performed under the trial-and-error approach, it resulted in significant variation in the CO_2 adsorption performance. The mean amount of CO_2 adsorbed during the seed cycle was observed to be 2.17 mmol g^{-1} with the least at 0.92 mmol g^{-1} and maximum at 2.88 mmol g^{-1}, respectively.

Based on the data generated during the seed experiments, two types of ML algorithms, namely, RF and GP regressors, were devised to predict $V_{0.8}$ using synthesis parameters as model inputs. The RF regressor exhibited superior performance with R^2 and RMSE values of 0.88 and 0.024 cm^3 g^{-1}, respectively, compared to GP with corresponding R^2 and RMSE values of 0.74 and 0.035 cm^3 g^{-1}, respectively (Table S2). Given the small size of the seed data set, we expect that RF, an ensemble algorithm must have fit the small-sized data set more effectively by making use of bootstrap sampling and statistical averaging.[28,50] Finally, given the better prediction

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

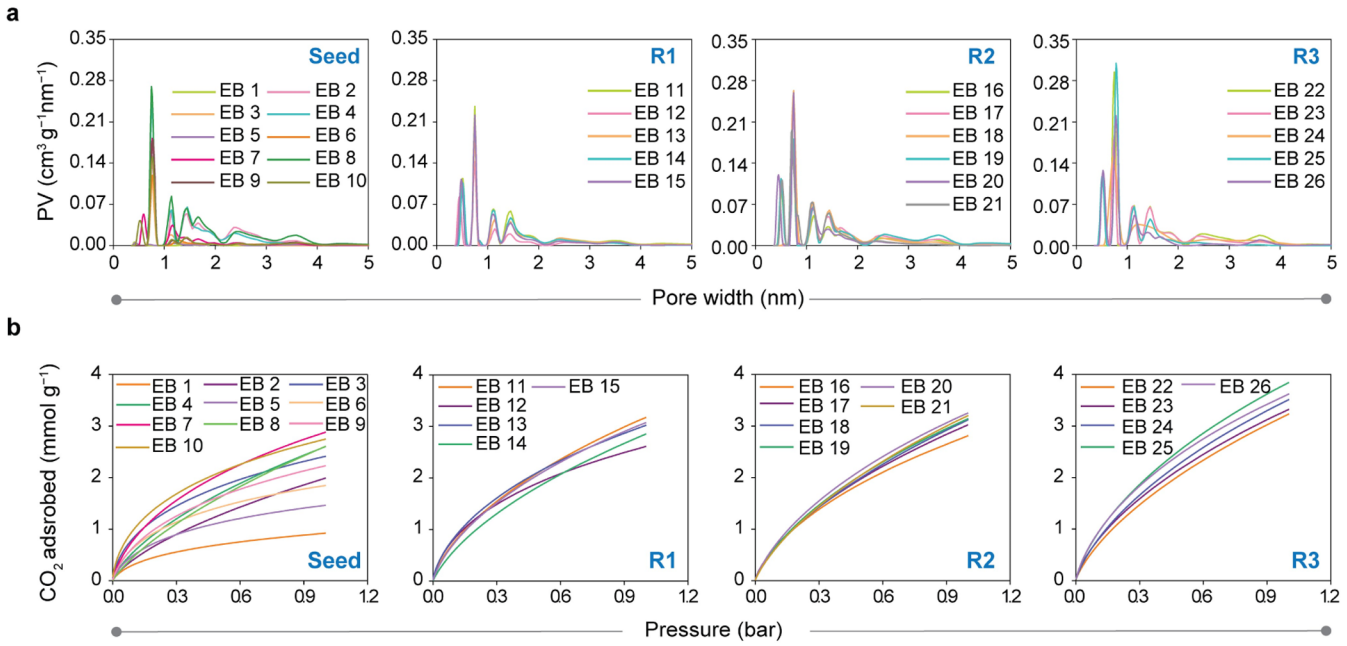


Figure 2. Pore size distributions (PSDs) of (a) ten samples for seed cycle, five samples for *R*1 cycle, six samples for *R*2 cycle, and five samples for *R*3 cycle. EB-*x*, *x* represents the number of each sample. (b) Experimental CO_2 uptake at 25 °C and 1 bar using all data from seed to *R*3 experiments. EB-*x*, *x* represents the number of each prepared sample.

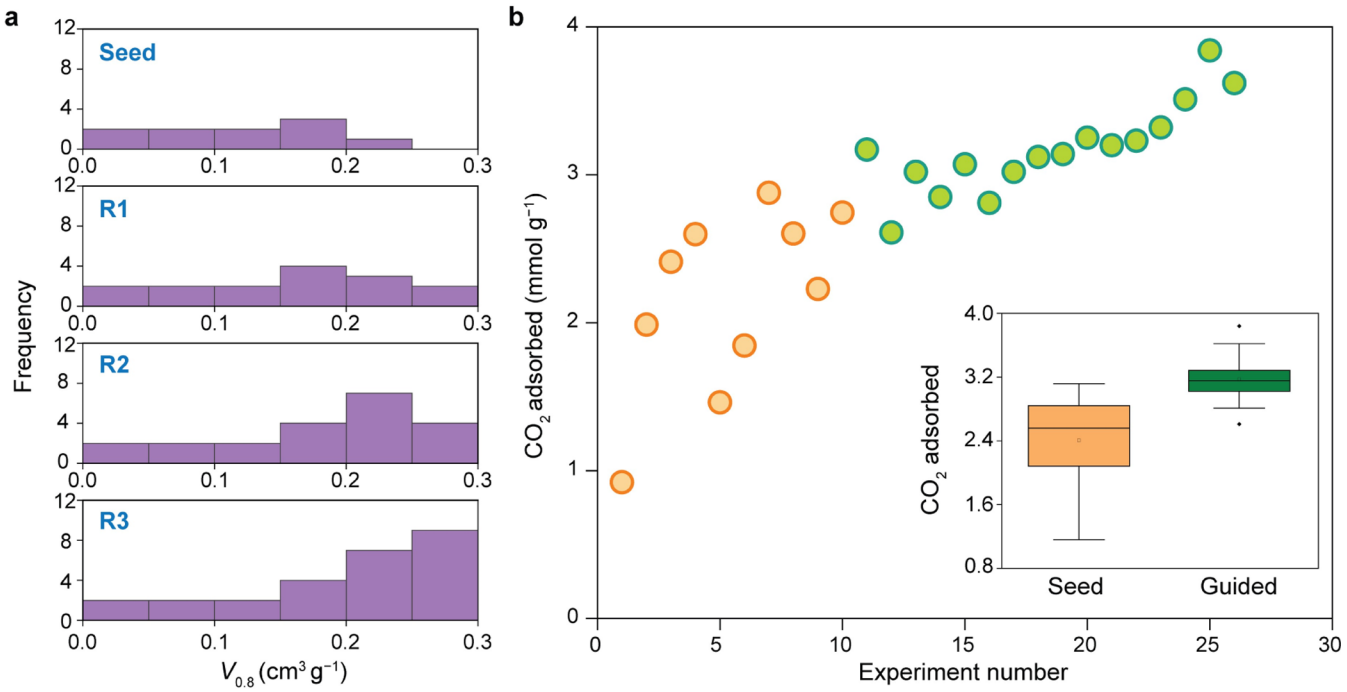


Figure 3. Evolution of (a) $V_{0.8}$ in the form of histogram over the active learning cycle. A steady increase in the size of $V_{0.8}$ and the corresponding frequency is observed after each iteration. (b) CO_2 adsorption capacity at 25 °C and 1 bar as a scatter plot, over the various experiments performed during active learning. A linear increase in the CO_2 adsorption capacities is observed with progressive experiments. The subplot in panel (b) shows a mean CO_2 uptake of 2.17 mmol g^{-1} for seed experiments, whereas those for guided experiments reach 3.24 mmol g^{-1}.

performance of the RF over GP, the former was hereon chosen as the choice of algorithm for SHAP analysis and as a surrogate for the subsequent optimization in the active learning framework.

The next step involved SHAP analysis of the RF regressor to unveil feature importance. Though SHAP interpretation does not imply causality, they aid in determining correlations between the input features and target variables and thus help deduce how the ML algorithm arrives at its prediction.[51,52] The synthesis parameters influence based on the seed experiments ranked as $T_C > T_A > R_C > K/B > R_A$. These findings align with the existing literature, wherein it is agreed that operating temperatures are more important than residence time for both carbonization and activation.[54−56] Moreover, during the KOH activation, chemical reactions occur at high

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

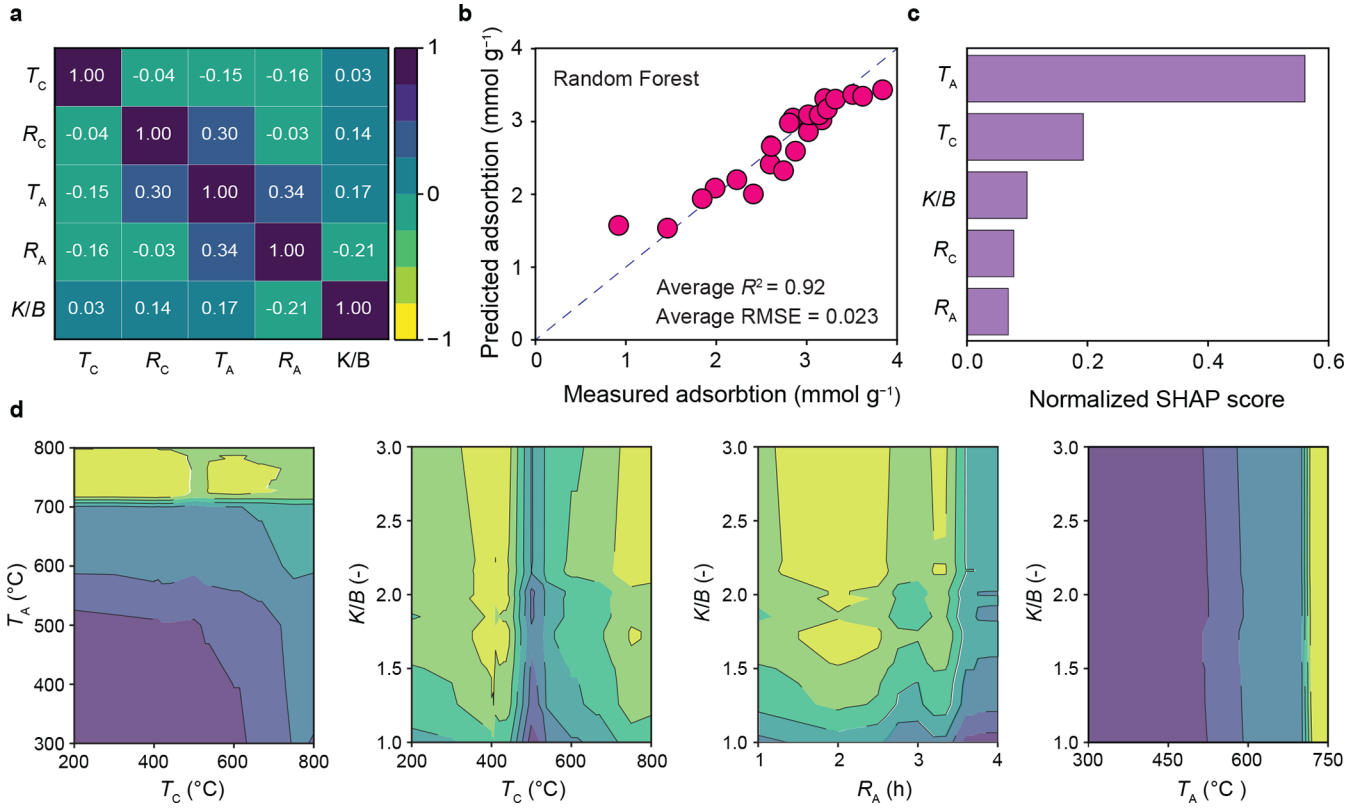


Figure 4. (a) Collinearity heatmap between all of the synthesis parameters investigated in the study, where no strong correlation is observed. (b) Average prediction performance of the RF algorithm over the entire active learning process. The blue dashed lines represent the line of equality ($x = y$). (c) Feature importance ranking represented as normalized SHAP values after the final round of active learning. (d) Two-way interaction plot between all of the synthetic parameters investigated, where the yellow regions depict maximal performance in terms of CO_2 adsorption. The following abbreviations are used: T_A for activation temperature, R_A for activation duration, T_c for carbonization temperature, R_C for carbonization duration, and K/B for mass ratio of KOH to biochar.

operating temperatures to efficiently develop microporous structures, highlighting the critical role of K/B over R_A.

3.2. Active Learning. After training the RF regressor on the seed experiments (10 experimental conditions), the optimization process was initiated in batch mode for PSO using local penalization. We defined maximization of $V_{0.8}$ as the objective function, as it was dependent on the synthesis parameters (and thus tunable) and linearly correlated to the CO_2 adsorption capacity at 25 °C and 1 bar. Thus, it served as an ideal link between the synthesis parameters and the CO_2 adsorption tests.

3.2.1. Maximizing $V_{0.8}$ and CO_2 Adsorption. Based on the seed experimental data, the RF-based surrogate model was developed, and its predictions were thereafter leveraged by the PSO to suggest five additional experimental points (i.e., half the size of the seed experiments). This formed our first round of active learning ($R1$). The PSO sampling of five experiments was not arbitrary but rather based on the modeler's notion of sampling 50% of the seed data set to capture sufficient experimental design space ($10 \times 0.5 = 5$). We synthesized engineered biochar samples based on the synthesis recommendations by the PSO ($R1$, summarized in Table S2) and characterized their textural properties, followed by the evaluation of the CO_2 adsorption capacities. Figure S1 presents the similar N_2 adsorption and desorption isotherms of the five engineered biochar samples synthesized during the $R1$ cycle, suggesting that they were also microporous carbon materials. As shown in Figure 2a, most of the pore sizes are less than 2 nm, which could further improve the CO_2 adsorption capacities.[42] The CO_2 adsorption isotherms at 25 °C of all five prepared samples are depicted in Figure 2b, exhibiting improved CO_2 adsorption capacities compared with seed experiments. Specifically, the average $V_{0.8}$ and CO_2 adsorption capacity during the latter experiments were 0.12 cm^3 g^{-1} and 2.17 mmol g^{-1}, whereas those from guided experiments during $R1$ were 0.23 cm^3 g^{-1} and 2.94 mmol g^{-1}, respectively. In summary, the CO_2 adsorption capacities increased by 1.35-fold, given the approximately 2-fold increase in $V_{0.8}$ (Figure 3).

During the second round ($R2$), the RF-based surrogate model was trained on 15 data points (seed and $R1$ experiments) and fed to the PSO, which in turn suggested the next 6 experimental conditions for synthesizing CO_2 adsorbents from rice husk. Here, again, the PSO sampling of 6 experiments was chosen by the modeler under the pretext of sampling 40% of the current data set ($15 \times 0.4 = 6$). On synthesizing the engineered biochar samples based on the suggested experiments, the PSDs of these 6 samples as shown in Figure 2a displayed a similar trend as the earlier synthesized samples but with improvements in $V_{0.8}$. As a result, the average CO_2 adsorption in $R2$ increased to 3.09 mmol g^{-1}, which was approximately 1.4 times more than those in seed experimental data. The maximum CO_2 adsorption in this cycle was found to be 3.25 mmol g^{-1}, approximately 13% more than the best CO_2 adsorption observed during the seed experiments.

The steady and linear improvements in CO_2 adsorption observed over two rounds of active learning prompted a third

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

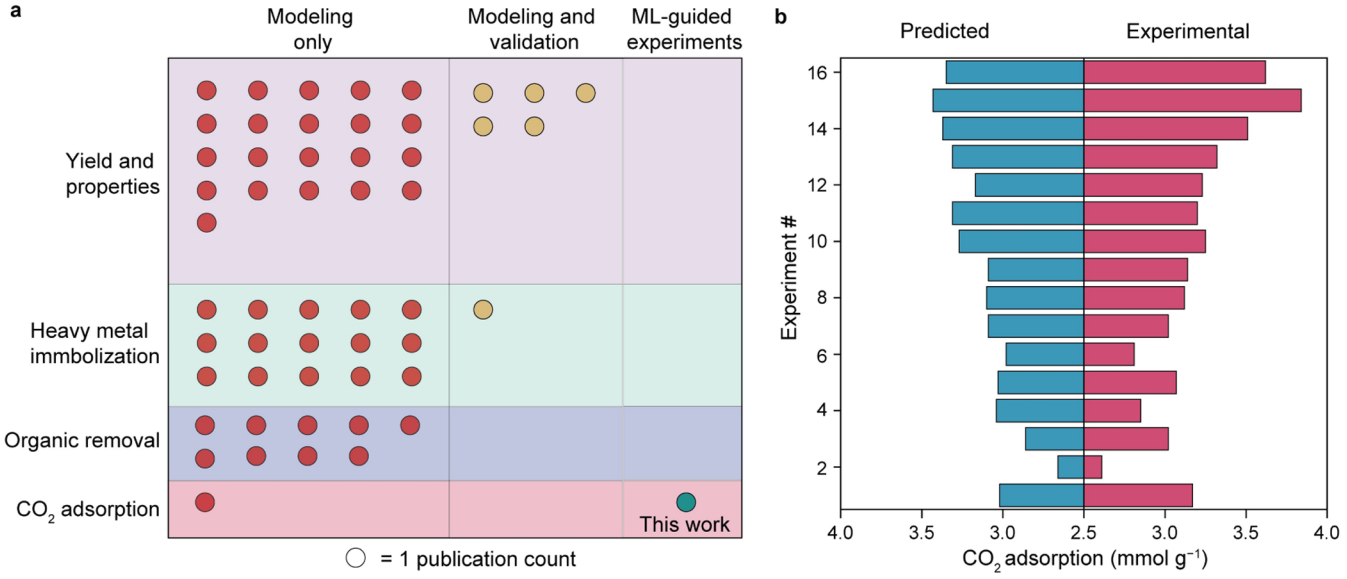


Figure 5. (a) Landscape of over 50 publications published between 2019 and 2023, leveraging machine learning for prediction of various biochar-specific properties and applications (y-axis). Approximately 90% of the published works are purely theoretical. (b) A series of 16 engineered biochar samples were synthesized during the active learning cycles by tuning their $V_{0.8}$ based on the model-recommended synthesis parameters, which ultimately affected the CO_2 adsorption. The predicted and experimentally measured CO_2 adsorption of these samples closely matched with each other.

round. With a total of 21 experimental data points until *R*2, we probed the PSO to generate five additional experimental data points for synthesizing high-performance CO_2 adsorbents from rice husk. The PSO sampling of five experiments in this round was chosen under the pretext of sampling 25% of the current data set ($21 \times 0.25 = 5$). By this round, it appeared that the PSO was hitting a global maximum with respect to parameters including T_C and K/B (further discussion in the next section). Nonetheless, the next set of five suggested experimental values, i.e., *R*3 were used to synthesize engineered biochar samples, and their performance was evaluated. The CO_2 adsorption isotherms at 25 °C of all five prepared samples during *R*3 are shown in Figure 2b and exhibited the best CO_2 adsorption capacities in this study. As observed in Figures 2b and 3b, three of the engineered biochar samples synthesized under this iteration had CO_2 adsorption capacities above 3.5 mmol g^{-1}, with the highest value of 3.84 mmol g^{-1}. From an overall comparative perspective, this implied that by the third round the CO_2 adsorption capacities had increased by 1.7−1.9 times compared with the seed experiments.

3.2.2. Improvements in Model Prediction during Active Learning. Beyond assessing the engineered biochar performance, we analyzed the evolution of the forward model's prediction ability. No correlations were observed among the synthesis conditions (inputs to the model), as shown in Figure 4a. Simultaneously, the RF-based forward model was trained using PSO-recommended experimental data after each active learning round, and its error metrics were observed as shown in Figure S2. During the initial *R*1 cycle, as the size of the data set increased from 10 to 15, the model performance also increased from $R^2 = 0.88$ to 0.91. Further on, retraining the RF algorithm with experimental data of *R*2 improved the predictions to $R^2 = 0.94$, finally saturating at the same values on subsequent training with *R*3 data. Throughout the active learning cycles, the RMSE values ranged between 0.021 and 0.024 mmol g^{-1}, with a mean $R^2 = 0.92$ as depicted in Figure 4b, indicating model robustness and generalization. The improvements in model prediction were expected with increased data quantity and quality, where the latter was fostered by clean, controlled experiment environments. Further details on the determination of optimal parameter values for T_C, R_C, T_A, R_A, and K/B are described in the Supporting Information (Section S6).

3.2.3. Rationalizing the Model Recommendations. The synthesis parameters including T_C, T_A, and K/B are vital for achieving desired microporosity in engineered biochar for effective CO_2 adsorption.[12,57] Interestingly, these parameters were identified as the top three features affecting performance by SHAP analysis (Figure 4c). It is worth noting that the model ranked these parameters with no prior mechanistic information on the reaction chemistry or kinetics but purely based on data-driven correlations, building user confidence. We rationalize the model recommendations by visually analyzing the two-way interaction among the synthesis parameters investigated and link them with literature claims to present plausible reasoning.

Integrated carbonization and activation, particularly with KOH, effectively enhance microporosity and pore structure in biochar due to the adequate valorization of the precursor.[14,57] Chemical activation with KOH interacts with carbon matrices, liberating gases to form micropores.[12,13,42] Our study identifies optimal conditions for high CO_2 adsorption with $T_A > 700$ °C and T_C at ca. 400 °C (Figure 4d). The activation of precursors carbonized at $T_C \leq 400$ °C results in high reactivity with the KOH solvent, yielding well-defined pore structures.[56] However, increasing T_C above 500 °C may hinder this process due to the presence of volatiles on the precursor surface formed during carbonization.[13,14] In the case of T_C and K/B (Figure 4d), high-performance regions were identified for T_C values between 400 and 405 and ratio values between 1.7 and 2.0. While from a modeling perspective these values suggest the PSO had converged at respective global and local minima, mechanistically, it indicted that the amount of KOH at T_C = 400−405 was suitable to react with most, if not all, of the

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

precursor, forming well-developed pores.[14,56] Similarly, for K/B and R_A (Figure 4d), the optimal performance was observed at R_A between 1.5 and 3.5 h, with K/B = 1.7−2, providing the necessary residence time for the activation process. A shorter duration (<1.5 h) would lead to incomplete reaction between KOH and the precursor, while the converse (>3.5 h) risks forming surface moieties that hamper porosity.[12] During chemical activation, T_A values ranging from 650 to 800 °C are typical for biomass-derived engineered biochar.[9,14] This also seemed the case in K/B and T_A (Figure 4d), where, irrespective of the K/B values, T_A > 700 °C favored high performance, with recommended values ranging between 724 and 800 °C and 1.5 and 2.0 for T_A and K/B, respectively, during $R3$ experiments. In summary, rationalizing the feature importance analysis revealed that the synthetic recommendations made by the model aligned closely with the established heuristic, reinforcing its validity and thus building user confidence.

3.3. Active Learning Implications for Guiding Biochar Synthesis. Recent years have seen an increase in the number of ML-based studies on various applications. However, our literature analysis reveals that opportunities to translate the collective insights from such studies to design property-specific biochars have remained minimal (Figure 5a, Supporting Information Table S1). To address this limitation, we developed a data-driven framework, which uses small data (typical in experimental research groups) to accelerate the development of engineered biochar materials, demonstrated through 16 samples synthesized over three active learning cycles with desired textural properties and enhanced CO_2 adsorption capacities. The model's prediction and the measured CO_2 adsorption were in close agreement with each other (Figure 5b), highlighting the efficacy of our approach.

The generality of our work arises from its data-driven nature and is not limited to particular feedstocks or specific synthetic parameters. Data-driven models have found prominence in recent years as they do not rely on the process mechanism but rather learn correlations between a set of inputs and target functions, making them adaptable to various applications. Thus, we are confident that our framework can be used or adapted by researchers to synthesize property-specific biochar for their respective applications by choosing the feedstocks and synthesis parameters of choice. By a similar extension, this strategy could also be used for the development of materials such as adsorbents, catalysts, and membranes by exploiting their structure−property performance relations provided sufficient (ca. 20−30) experimental data points are available to initiate the active learning loop. In the current study, our prime focus was directed toward developing the guided synthesis protocol from conception to implementation. Understanding the mechanisms and kinetics of CO_2 adsorption (i.e., functional groups, hydrophilicity, and surface charge density) for these engineered biochar samples will be the point of future research endeavors.

■ ASSOCIATED CONTENT

Supporting Information

The curated data set and ML model developed in this study are open sourced at GitHub (https://github.com/ssuvarnamanu/active-learning-for-biochar-design). The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.3c10922.

Details of the ML algorithms are explained in Section S1, and hyper-parameters, error metrics, and SHAP analysis are described in Sections S2−S4. Formulation of the PSO and evolution of the model predictions and optimization are provided in Sections S5 and S6, respectively. N_2 adsorption isotherms are depicted in Figure S1. Model performance in terms of R^2 is shown Figure S2. Details of the experimental data set are presented in Table S1. Algorithm-specific tuned hyper-parameter are listed in Table S2. Literature overview of ML studies pertaining to biochar applications is described in Table S3 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Xiaonan Wang − *Department of Chemical Engineering, Tsinghua University, Beijing 100084, China*; Email: wangxiaonan@mail.tsinghua.edu.cn

Yong Sik Ok − *Korea Biochar Research Center, APRU Sustainable Waste Management Program & Division of Environmental Science and Ecological Engineering, Korea University, Seoul 02841, Republic of Korea*; orcid.org/0000-0003-3401-0912; Email: yongsikok@korea.ac.kr

Authors

Xiangzhou Yuan − *Ministry of Education of Key Laboratory of Energy Thermal Conversion and Control, School of Energy and Environment, Southeast University, Nanjing 210096, China; Korea Biochar Research Center, APRU Sustainable Waste Management Program & Division of Environmental Science and Ecological Engineering, Korea University, Seoul 02841, Republic of Korea*; orcid.org/0000-0002-6480-3983

Manu Suvarna − *Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland*

Juin Yau Lim − *Korea Biochar Research Center, APRU Sustainable Waste Management Program & Division of Environmental Science and Ecological Engineering, Korea University, Seoul 02841, Republic of Korea*

Javier Pérez-Ramírez − *Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland*; orcid.org/0000-0002-5805-7355

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.3c10922

Author Contributions

[⊥]X.Y. and M.S. contributed equally to this work. X.Y.: conceptualization, experiments and analyses, data collection, writing (review and editing), and visualization; M.S.: conceptualization, modeling, writing (review and editing), and visualization; J.Y.L.: writing (review and editing); J.P.-R.: writing (review and editing); X.W.: conceptualization, writing (review and editing), and supervision; and Y.S.O.: conceptualization, writing (review and editing), and supervision.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

government (MSIT) (No. 2021R1A2C2011734). This research was also supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF—2021R1A6A1A10045235). This research was partly supported by the OJEong Resilience Institute, Korea University, Korea, the Start-up Research Fund of Southeast University (RF1028623274), China, and the National Science and Technology Major Project of China (No. 2022ZD0117501), China.

REFERENCES

(1) Mac Dowell, N.; Fennell, P. S.; Shah, N.; Maitland, G. C. The Role of CO_2 Capture and Utilization in Mitigating Climate Change. *Nat. Clim. Change* **2017**, *7* (4), 243−249.
(2) Download Report: Global Warming of 1.5°C an IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. https://www.ipcc.ch/sr15/download (accessed December 04, 2022).
(3) US Department of Commerce. Global Monitoring Laboratory - Carbon Cycle Greenhouse Gases. https://gml.noaa.gov/ccgg/trends/mlo.html (accessed December 04, 2022).
(4) Dissanayake, P. D.; You, S.; Igalavithana, A. D.; Xia, Y.; Bhatnagar, A.; Gupta, S.; Kua, H. W.; Kim, S.; Kwon, J.-H.; Tsang, D. C. W.; Ok, Y. S. Biochar-Based Adsorbents for Carbon Dioxide Capture: A Critical Review. *Renewable Sustainable Energy Rev.* **2020**, *119*, No. 109582.
(5) Singh, G.; Lakhi, K. S.; Sil, S.; Bhosale, S. V.; Kim, I.; Albahily, K.; Vinu, A. Biomass Derived Porous Carbon for CO_2 Capture. *Carbon* **2019**, *148*, 164−186.
(6) Guo, S.; Li, Y.; Wang, Y.; Wang, L.; Sun, Y.; Liu, L. Recent Advances in Biochar-Based Adsorbents for CO_2 Capture. *Carbon Capture Sci. Technol.* **2022**, *4*, No. 100059.
(7) Ahmed, R.; Liu, G.; Yousaf, B.; Abbas, Q.; Ullah, H.; Ali, M. U. Recent Advances in Carbon-Based Renewable Adsorbent for Selective Carbon Dioxide Capture and Separation-A Review. *J. Cleaner Prod.* **2020**, *242*, No. 118409.
(8) Li, S.; Yuan, X.; Deng, S.; Zhao, L.; Lee, K. B. A Review on Biomass-Derived CO_2 Adsorption Capture: Adsorbent, Adsorber, Adsorption, and Advice. *Renewable Sustainable Energy Rev.* **2021**, *152*, No. 111708.
(9) Yuan, X.; Wang, J.; Deng, S.; Suvarna, M.; Wang, X.; Zhang, W.; Hamilton, S. T.; Alahmed, A.; Jamal, A.; Park, A.-H. A.; Bi, X.; Ok, Y. S. Recent Advancements in Sustainable Upcycling of Solid Waste into Porous Carbons for Carbon Dioxide Capture. *Renewable Sustainable Energy Rev.* **2022**, *162*, No. 112413.
(10) Yuan, X.; Kumar, N. M.; Brigljević, B.; Li, S.; Deng, S.; Byun, M.; Lee, B.; Lin, C. S. K.; Tsang, D. C. W.; Lee, K. B.; Chopra, S. S.; Lim, H.; Ok, Y. S. Sustainability-Inspired Upcycling of Waste Polyethylene Terephthalate Plastic into Porous Carbon for CO_2 Capture. *Green Chem.* **2022**, *24* (4), 1494−1504.
(11) Yuan, X.; Wang, J.; Deng, S.; Dissanayake, P. D.; Wang, S.; You, S.; Yip, A. C. K.; Li, S.; Jeong, Y.; Tsang, D. C. W.; Ok, Y. S. Sustainable Food Waste Management: Synthesizing Engineered Biochar for CO_2 Capture. *ACS Sustainable Chem. Eng.* **2022**, *10* (39), 13026−13036.
(12) Heidarinejad, Z.; Dehghani, M. H.; Heidari, M.; Javedan, G.; Ali, I.; Sillanpää, M. Methods for Preparation and Activation of Activated Carbon: A Review. *Environ. Chem. Lett.* **2020**, *18* (2), 393−415.
(13) Kamran, U.; Park, S.-J. Chemically Modified Carbonaceous Adsorbents for Enhanced CO_2 Capture: A Review. *J. Cleaner Prod.* **2021**, *290*, No. 125776.
(14) Chen, Y.; Zhang, X.; Chen, W.; Yang, H.; Chen, H. The Structure Evolution of Biochar from Biomass Pyrolysis and Its Correlation with Gas Pollutant Adsorption Performance. *Bioresour. Technol.* **2017**, *246*, 101−109.
(15) Qiu, T.; Li, C.; Guang, M.; Zhang, Y. Porous Carbon Material Production from Microwave-Assisted Pyrolysis of Peanut Shell. *Carbon Res.* **2023**, *2* (1), 45.
(16) Guo, L.; Yang, J.; Hu, G.; Hu, X.; Wang, L.; Dong, Y.; DaCosta, H.; Fan, M. Role of Hydrogen Peroxide Preoxidizing on CO_2 Adsorption of Nitrogen-Doped Carbons Produced from Coconut Shell. *ACS Sustainable Chem. Eng.* **2016**, *4* (5), 2806−2813.
(17) Nazir, G.; Rehman, A.; Park, S.-J. Role of Heteroatoms (Nitrogen and Sulfur)-Dual Doped Corn-Starch Based Porous Carbons for Selective CO_2 Adsorption and Separation. *J. CO2 Util.* **2021**, *51*, No. 101641.
(18) Ma, X.; Li, L.; Zeng, Z.; Chen, R.; Wang, C.; Zhou, K.; Li, H. Experimental and Theoretical Demonstration of the Relative Effects of O-Doping and N-Doping in Porous Carbons for CO_2 Capture. *Appl. Surf. Sci.* **2019**, *481*, 1139−1147.
(19) Liu, W.-J.; Jiang, H.; Tian, K.; Ding, Y.-W.; Yu, H.-Q. Mesoporous Carbon Stabilized MgO Nanoparticles Synthesized by Pyrolysis of MgCl2 Preloaded Waste Biomass for Highly Efficient CO_2 Capture. *Environ. Sci. Technol.* **2013**, *47* (16), 9397−9403.
(20) Mekki-Berrada, F.; Ren, Z.; Huang, T.; Wong, W. K.; Zheng, F.; Xie, J.; Tian, I. P. S.; Jayavelu, S.; Mahfoud, Z.; Bash, D.; Hippalgaonkar, K.; Khan, S.; Buonassisi, T.; Li, Q.; Wang, X. Two-Step Machine Learning Enables Optimized Nanoparticle Synthesis. *npj Comput. Mater.* **2021**, *7* (1), 1−10.
(21) Tang, B.; Lu, Y.; Zhou, J.; Chouhan, T.; Wang, H.; Golani, P.; Xu, M.; Xu, Q.; Guan, C.; Liu, Z. Machine Learning-Guided Synthesis of Advanced Inorganic Materials. *Mater. Today* **2020**, *41*, 72−80.
(22) Zhuo, Q.; Liang, Y.; Hu, Y.; Shi, M.; Zhao, C.; Zhang, S. Applications of Biochar in Medical and Related Environmental Fields: Current Status and Future Perspectives. *Carbon Res.* **2023**, *2* (1), 32.
(23) Shields, B. J.; Stevens, J.; Li, J.; Parasram, M.; Damani, F.; Alvarado, J. I. M.; Janey, J. M.; Adams, R. P.; Doyle, A. G. Bayesian Reaction Optimization as a Tool for Chemical Synthesis. *Nature* **2021**, *590* (7844), 89−96.
(24) Angello, N. H.; Rathore, V.; Beker, W.; Wołos, A.; Jira, E. R.; Roszak, R.; Wu, T. C.; Schroeder, C. M.; Aspuru-Guzik, A.; Grzybowski, B. A.; Burke, M. D. Closed-Loop Optimization of General Reaction Conditions for Heteroaryl Suzuki-Miyaura Coupling. *Science* **2022**, *378* (6618), 399−405.
(25) Burns, T. D.; Pai, K. N.; Subraveti, S. G.; Collins, S. P.; Krykunov, M.; Rajendran, A.; Woo, T. K. Prediction of MOF Performance in Vacuum Swing Adsorption Systems for Postcombustion CO_2 Capture Based on Integrated Molecular Simulations, Process Optimizations, and Machine Learning Models. *Environ. Sci. Technol.* **2020**, *54* (7), 4536−4544.
(26) Zhang, K.; Zhong, S.; Zhang, H. Predicting Aqueous Adsorption of Organic Compounds onto Biochars, Carbon Nanotubes, Granular Activated Carbons, and Resins with Machine Learning. *Environ. Sci. Technol.* **2020**, *54* (11), 7008−7018.
(27) Gao, H.; Zhong, S.; Zhang, W.; Igou, T.; Berger, E.; Reid, E.; Zhao, Y.; Lambeth, D.; Gan, L.; Afolabi, M. A.; Tong, Z.; Lan, G.; Chen, Y. Revolutionizing Membrane Design Using Machine Learning-Bayesian Optimization. *Environ. Sci. Technol.* **2022**, *56* (4), 2572−2581.
(28) Zhu, X.; Li, Y.; Wang, X. Machine Learning Prediction of Biochar Yield and Carbon Contents in Biochar Based on Biomass Characteristics and Pyrolysis Conditions. *Bioresour. Technol.* **2019**, *288*, No. 121527.
(29) Li, J.; Pan, L.; Suvarna, M.; Tong, Y. W.; Wang, X. Fuel Properties of Hydrochar and Pyrochar: Prediction and Exploration with Machine Learning. *Appl. Energy* **2020**, *269*, No. 115166.
(30) Pathy, A.; Meher, S.; B, P. Predicting Algal Biochar Yield Using EXtreme Gradient Boosting (XGB) Algorithm of Machine Learning Methods. *Algal Res.* **2020**, *50*, No. 102006.
(31) Li, H.; Ai, Z.; Yang, L.; Zhang, W.; Yang, Z.; Peng, H.; Leng, L. Machine Learning Assisted Predicting and Engineering Specific

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636

Surface Area and Total Pore Volume of Biochar. *Bioresour. Technol.* **2023**, *369*, No. 128417.
(32) Zhu, X.; He, M.; Sun, Y.; Xu, Z.; Wan, Z.; Hou, D.; Alessi, D. S.; Tsang, D. C. W. Insights into the Adsorption of Pharmaceuticals and Personal Care Products (PPCPs) on Biochar and Activated Carbon with the Aid of Machine Learning. *J. Hazard. Mater.* **2022**, *423*, No. 127060.
(33) Nguyen, X. C.; Nguyen, T. T. H.; Hang, N. T. T.; Thai, V. N.; Doan, T. O.; Duong, T. T.; Duong, T. N.; Hwang, Y.; Lam, V. S.; Ly, Q. V.Insight into the Adsorption of Nutrients from Water by Pyrogenic Carbonaceous Adsorbents Using a Bootstrap Method and Machine Learning *ACS EST Water* 2022 DOI: 10.1021/acsestwater.2c00301.
(34) Nguyen, X. C.; Ly, Q. V.; Nguyen, T. T. H.; Ngo, H. T. T.; Hu, Y.; Zhang, Z. Potential Application of Machine Learning for Exploring Adsorption Mechanisms of Pharmaceuticals onto Biochars. *Chemosphere* **2022**, *287*, No. 132203.
(35) Zhu, X.; Wang, X.; Ok, Y. S. The Application of Machine Learning Methods for Prediction of Metal Sorption onto Biochars. *J. Hazard. Mater.* **2019**, *378*, No. 120727.
(36) Palansooriya, K. N.; Li, J.; Dissanayake, P. D.; Suvarna, M.; Li, L.; Yuan, X.; Sarkar, B.; Tsang, D. C. W.; Rinklebe, J.; Wang, X.; Ok, Y. S. Prediction of Soil Heavy Metal Immobilization by Biochar Using Machine Learning. *Environ. Sci. Technol.* **2022**, *56* (7), 4187−4198.
(37) Zhang, W.; Ashraf, W. M.; Senadheera, S. S.; Alessi, D. S.; Tack, F. M. G.; Ok, Y. S. Machine Learning Based Prediction and Experimental Validation of Arsenite and Arsenate Sorption on Biochars. *Sci. Total Environ.* **2023**, *904*, No. 166678.
(38) Ullah, H.; Khan, S.; Chen, B.; Shahab, A.; Riaz, L.; Lun, L.; Wu, N. Machine Learning Approach to Predict Adsorption Capacity of Fe-Modified Biochar for Selenium. *Carbon Res.* **2023**, *2* (1), 29.
(39) Yuan, X.; Suvarna, M.; Low, S.; Dissanayake, P. D.; Lee, K. B.; Li, J.; Wang, X.; Ok, Y. S. Applied Machine Learning for Prediction of CO_2 Adsorption on Biomass Waste-Derived Porous Carbons. *Environ. Sci. Technol.* **2021**, *55* (17), 11925−11936.
(40) Paula, A. J.; Ferreira, O. P.; Souza Filho, A. G.; Filho, F. N.; Andrade, C. E.; Faria, A. F. Machine Learning and Natural Language Processing Enable a Data-Oriented Experimental Design Approach for Producing Biochar and Hydrochar from Biomass. *Chem. Mater.* **2022**, *34* (3), 979−990.
(41) Wang, R.; Zhang, S.; Chen, H.; He, Z.; Cao, G.; Wang, K.; Li, F.; Ren, N.; Xing, D.; Ho, S.-H. Enhancing Biochar-Based Nonradical Persulfate Activation Using Data-Driven Techniques. *Environ. Sci. Technol.* **2023**, *57* (9), 4050−4059.
(42) Yuan, X.; Lee, J. G.; Yun, H.; Deng, S.; Kim, Y. J.; Lee, J. E.; Kwak, S. K.; Lee, K. B. Solving Two Environmental Issues Simultaneously: Waste Polyethylene Terephthalate Plastic Bottle-Derived Microporous Carbons for Capturing CO_2. *Chem. Eng. J.* **2020**, *397*, No. 125350.
(43) Li, J.; Suvarna, M.; Li, L.; Pan, L.; Pérez-Ramírez, J.; Ok, Y. S.; Wang, X. A Review of Computational Modeling Techniques for Wet Waste Valorization: Research Trends and Future Perspectives. *J. Cleaner Prod.* **2022**, *367*, No. 133025.
(44) Zhong, S.; Zhang, K.; Bagheri, M.; Burken, J. G.; Gu, A.; Li, B.; Ma, X.; Marrone, B. L.; Ren, Z. J.; Schrier, J.; Shi, W.; Tan, H.; Wang, T.; Wang, X.; Wong, B. M.; Xiao, X.; Yu, X.; Zhu, J.-J.; Zhang, H. Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environ. Sci. Technol.* **2021**, *55* (19), 12741−12754.
(45) Kim, M.; Ha, M. Y.; Jung, W.-B.; Yoon, J.; Shin, E.; Kim, I.; Lee, W. B.; Kim, Y.; Jung, H. Searching for an Optimal Multi-Metallic Alloy Catalyst by Active Learning Combined with Experiments. *Adv. Mater.* **2022**, *34* (19), No. 2108900.
(46) Qu, J.; Wang, Y.; Tian, X.; Jiang, Z.; Deng, F.; Tao, Y.; Jiang, Q.; Wang, L.; Zhang, Y. KOH-Activated Porous Biochar with High Specific Surface Area for Adsorptive Removal of Chromium (VI) and Naphthalene from Water: Affecting Factors, Mechanisms and Reusability Exploration. *J. Hazard. Mater.* **2021**, *401*, No. 123292.
(47) Ludwinowicz, J.; Jaroniec, M. Effect of Activating Agents on the Development of Microporosity in Polymeric-Based Carbon for CO_2 Adsorption. *Carbon* **2015**, *94*, 673−679.
(48) Li, C.; Li, Y.; Shao, Y.; Zhang, L.; Zhang, S.; Wang, S.; Li, B.; Cui, Z.; Tang, Y.; Hu, X. Activation of Biomass with Volatilized KOH. *Green Chem.* **2023**, *25* (7), 2825−2839.
(49) Li, J.; Suvarna, M.; Pan, L.; Zhao, Y.; Wang, X. A Hybrid Data-Driven and Mechanistic Modelling Approach for Hydrothermal Gasification. *Appl. Energy* **2021**, *304*, No. 117674.
(50) Suvarna, M.; Jahirul, M. I.; Aaron-Yeap, W. H.; Augustine, C. V.; Umesh, A.; Rasul, M. G.; Günay, M. E.; Yildirim, R.; Janaun, J. Predicting Biodiesel Properties and Its Optimal Fatty Acid Profile via Explainable Machine Learning. *Renewable Energy* **2022**, *189*, 245−258, DOI: 10.1016/j.renene.2022.02.124.
(51) Suvarna, M.; Preikschas, P.; Pérez-Ramírez, J. Identifying Descriptors for Promoted Rhodium-Based Catalysts for Higher Alcohol Synthesis via Machine Learning. *ACS Catal.* **2022**, *12* (24), 15373−15385.
(52) Suvarna, M.; Araújo, T. P.; Pérez-Ramírez, J. A Generalized Machine Learning Framework to Predict the Space-Time Yield of Methanol from Thermocatalytic CO_2 Hydrogenation. *Appl. Catal.,B* **2022**, *315*, No. 121530.
(53) Liu, X.; Lu, D.; Zhang, A.; Liu, Q.; Jiang, G. Data-Driven Machine Learning in Environmental Pollution: Gains and Problems. *Environ. Sci. Technol.* **2022**, *56* (4), 2124−2133.
(54) Chang, Y.-M.; Tsai, W.-T.; Li, M.-H. Characterization of Activated Carbon Prepared from Chlorella-Based Algal Residue. *Bioresour. Technol.* **2015**, *184*, 344−348.
(55) Jang, E.; Choi, S. W.; Lee, K. B. Effect of Carbonization Temperature on the Physical Properties and CO_2 Adsorption Behavior of Petroleum Coke-Derived Porous Carbon. *Fuel* **2019**, *248*, 85−92.
(56) Zornitta, R. L.; Barcelos, K. M.; Nogueira, F. G. E.; Ruotolo, L. A. M. Understanding the Mechanism of Carbonization and KOH Activation of Polyaniline Leading to Enhanced Electrosorption Performance. *Carbon* **2020**, *156*, 346−358.
(57) Singh, G.; Lee, J.; Karakoti, A.; Bahadur, R.; Yi, J.; Zhao, D.; AlBahily, K.; Vinu, A. Emerging Trends in Porous Materials for CO_2 Capture and Conversion. *Chem. Soc. Rev.* **2020**, *49* (13), 4360−4404.

https://doi.org/10.1021/acs.est.3c10922
Environ. Sci. Technol. 2024, 58, 6628−6636