Adsorption Performance of Torrefied Wood Chips for Volatile Organic Compounds and Ethylene Gas

Hyeon Cheol Kim, Si Young Ha, and Jae-Kyung Yang *

Machine learning models were developed to predict volatile organic compound and ethylene gas adsorption performance of freshness-preserving agents based on torrefied oak wood chips. Oak chips were torrefied at 350 °C for 20 min and processed into three particle sizes. A dataset of 39 experimental points was collected, comprising 8 input variables (particle size, torrefied wood content, commercial content, bulk density, compressed density, porosity, total content, and final bulk density) and 2 output variables (VOC and ethylene adsorption levels). Data augmentation techniques were applied to overcome dataset limitations. Three machine learning algorithms were implemented: Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Support Vector Regression (SVR). For ethylene adsorption, SVR achieved superior performance with R^2 = 0.934, RMSE = 5.06, and MAE = 1.997.). For VOC adsorption, RF demonstrated highest accuracy with R^2 = 0.962, RMSE = 1.11, and MAE = 0.845. Torrefied wood content was positively correlated with ethylene adsorption (r = 0.43). Porosity was negatively correlated (r = -0.76). Higher porosity gave reduced ethylene capture efficiency, consistent with a negative relationship between pore structure and adsorption. The effectiveness of machine learning was demonstrated in predicting gas adsorption performance. The work provides practical guidelines for designing torrefied wood-based freshness-preserving systems.

https://doi.org/10.15376/biores.21.2.4538-4561

Keywords: Torrefaction; Oak wood chips; Gas adsorption; Machine learning; Freshness preservation; Data augmentation

Contact information: Department of Environmental Materials Science/Institute of Agriculture and Life Science, Gyeongsang National University, Jinju, 52828, Republic of Korea;
** Corresponding author: jkyang@gnu.ac.kr*

INTRODUCTION

Global wood production reaches 4 billion m³ annually, with the global wood market projected to reach $174.3 billion in 2024 (FAO 2024). The forest sector contributed over $663 billion to global GDP in 2015, demonstrating the significant economic impact of high value-added utilization of wood resources (Li *et al.* 2022).

Forest biomass, as a renewable resource, has gained attention as a key material for achieving carbon neutrality and addressing climate change challenges (Berndes *et al.* 2016; Selivanov *et al.* 2023). While conventional wood utilization has primarily focused on primary processed products such as construction materials and pulp raw materials, recent research has actively explored transitions to high value-added products including biofuels, biochemicals, and functional materials (Amidon and Liu 2009; Woo and Turner 2019). Wood fuel production alone reached 1.97 billion m³ in 2022, indicating the rapid expansion of energy conversion from wood biomass (FAO 2024). This paradigm shift holds significant importance not only for improving the economic efficiency of the wood industry but also for developing environmentally friendly materials that can replace fossil-based alternatives (Brunet-Navarro *et al.* 2021).

Torrefaction is a thermal treatment technology that processes biomass in an oxygen-limited environment at relatively low temperatures (250 to 350 °C), significantly improving the physicochemical properties of wood as a pretreatment process (Chen *et al.* 2019, Lu *et al*. 2012). During this process, hemicellulose and some cellulose decompose to form a porous structure similar to biochar, while enhancing the energy density of the raw material (Chen and Kuo 2011; Tumuluru *et al.* 2021). Hardwood species including oak are evaluated as suitable raw materials for torrefaction treatment due to their solid wood texture and high density (Simonic *et al*. 2020). Torrefied wood exhibits increased hydrophobicity and improved grindability, making it highly applicable across various fields (Chen *et al.* 2019; Tumuluru *et al.* 2021). Particularly, the surface functional group changes generated during the torrefaction process (Xu *et al*. 2024), along with the porous structure, significantly enhance its potential for utilization as adsorption material (Doddapaneni and Kikas. 2023), attracting attention as an environmentally friendly alternative to fossil-based materials.

Although torrefied wood has been proposed as a potential adsorbent for ethylene gas and volatile organic compounds generated during the post-ripening process of agricultural products (Ha *et al.* 2025; Doddapaneni and Kikas 2023), its adsorption performance under various formulation conditions has not been systematically characterized. Ethylene, known as the "death or ripening hormone," is a plant hormone that accelerates fruit and vegetable maturation, serving as a major cause of quality deterioration during storage (Saltveit 1999). Indeed, excessive ethylene gas exposure is a major cause of postharvest losses, with fruits and vegetables accounting for 20 to 50% of global food losses in the supply chain (Kader 1985; Schudel *et al.* 2023). Internal ethylene concentrations in mature fruits can exceed 100 μL/L, rapidly accelerating quality deterioration of surrounding agricultural products (Saltveit 1999). Additionally, volatile organic compounds (VOCs) cause loss of flavor components in agricultural products and promote microbial growth, thereby adversely affecting freshness preservation (Martínez-Romero *et al.* 2007). Torrefied wood chips can be utilized as natural freshness-preserving agents that effectively adsorb these harmful gases and extend the storage life of agricultural products (Ha *et al.* 2025). Unlike conventional chemical preservatives, biochar-based ethylene absorbers are receiving significant attention as environmentally friendly and safe alternatives (Charoensuk *et al*. 2024).

The adsorption performance of freshness-preserving agents is determined by complex interactions among various structural parameters including particle size, torrefied material content, commercial base material content, bulk density, true density, and porosity. Traditional experimental approaches alone require significant time and cost for optimizing such multi-variable systems(Manatura *et al.* 2023). Machine learning methods offer a data-efficient complement by modeling complex variable interactions and enabling prediction across a broader parameter space than direct experimentation alone.

The introduction of machine learning technology is essential to address these challenges. Recent studies have demonstrated excellent performance of machine learning models in gas adsorption and biochar/torrefied material predictions, which are directly relevant to this research field (Zhang *et al.* 2023). Machine learning models have successfully predicted biochar porous structure properties, with validation R^2 reaching 0.94 for surface area and pore volume predictions (Li *et al*. 2023). The LightGBM model demonstrated superior prediction performance for biochar CO_2 adsorption capacity ($R^2 = 0.956$) (Zhao *et al*. 2025). The Gradient Tree Boosting model developed to predict the solid yield and higher heating value of torrefied biomass achieved a high accuracy of $R^2 = 0.93$ (Onsree *et al*. 2022). These machine learning algorithms can model nonlinear and complex relationships and achieve high prediction accuracy even from limited experimental data (Dalmau *et al*. 2025; Ghattas and Manzon 2023). In particular, data augmentation techniques can overcome limitations of small datasets and improve model generalization performance, making them essential tools for efficient freshness-preserving agent development (Liu *et al.* 2025).

Previous studies on gas adsorption and freshness-preserving agents mainly focused on single gas adsorption performance or relied on traditional statistical analysis methods, failing to reflect complex interactions in multi-variable systems (Pilkington *et al.* 2014). Moreover, most studies required substantial experimental data, limiting practical applications. With the advancement of machine learning technology, data augmentation techniques that can achieve high prediction performance even with small datasets have been developed, presenting new approaches to overcome these limitations (Jha *et al.* 2019). Particularly, there is a growing need for developing integrated models that predict simultaneous adsorption performance of both ethylene and volatile organic compounds, and research on comprehensive approaches that simultaneously consider structural characteristics and composition ratios of torrefied wood is required. Addressing these technical challenges to provide practical freshness-preserving agent design guidelines is recognized as a major objective in the current research field.

The primary objective of this study is to develop machine learning models that can accurately predict the volatile organic compound and ethylene gas adsorption performance of freshness-preserving agents based on torrefied oak chips. The specific research objectives are as follows. First, to systematically analyze physicochemical property changes of oak chips torrefied at 350 °C for 20 min and derive optimal conditions through particle size classification (< 1 mm, < 2 mm, < 4 mm). Second, to identify correlations between 8 input variables including particle size, torrefied material content, commercial base material content, bulk density, compressed density, and porosity, and 2 output variables including volatile organic compounds adsorption rate and ethylene adsorption rate based on 39 experimental data points. Third, to overcome limitations of small datasets by applying data augmentation techniques and comparatively evaluate prediction performance of Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Support Vector Regression (SVR) models. Fourth, to quantitatively assess model accuracy and generalization capability using coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE).

EXPERIMENTAL

Materials

The oak wood chips were provided by Poonglim Co., Ltd. (Republic of Korea) and were standardized to a uniform size of 3 cm × 3 cm. The chips were oven-dried at 105 °C for more than 24 h to completely remove moisture.

Torrefaction Process

Torrefaction was conducted using a laboratory-scale muffle furnace (TMF-3200, EYELA, Tokyo, Japan). The torrefaction conditions for oak wood chips were selected based on the results of the first-year study. Specifically, the chips were heated at 350 °C for 20 min in the electric muffle furnace. Prior to torrefaction, each batch of oak chips was wrapped in aluminum foil and glass fiber to minimize contact with oxygen. After the torrefaction process, the samples were cooled under a nitrogen atmosphere and stored in zip-lock bags until further processing.

Preparation of a Freshness Preservative Using Torrefaction Oak Chips

The torrefied oak chips were ground into powder using a commercial grinder operating at 34,000 rpm. The resulting torrefied material was sieved using standard mesh screens to obtain three particle size fractions: <1 mm (20 mesh pass), <2 mm (10 mesh pass), and <4 mm (5 mesh pass). All sieved samples were stored in zip-lock bags until further use in subsequent processes. For the preparation of freshness-preserving agents containing torrefied powder, a commercial freshness-preserving base material was provided by SunGel Co., Ltd. (Republic of Korea). The agents were formulated by varying the particle size and input amount of the torrefied material, as well as the amount of the freshness-preserving base. Detailed formulations are summarized in Table 1.

Evaluation of Bulk Porosity of Uncompressed Torrefied Wood Powder Bed

The torrefied oak wood was placed in a 25 mL glass cylinder, and its weight was measured. The real density was calculated by dividing the weight by the volume. The sample was then dried in a dry oven at 105 °C for more than 12 h, followed by compression using a 500 g weight. The weight after compression was recorded to determine the relative density of the unpacked bed, following standard procedures for biomass characterization (ASTM D2395-17). The relative porosity of the unpacked shelf-life extender containing torrefied oak was calculated according to Eq. 1 (Table 1). Note that this calculation does not indicate a true density due to the fact that the compressed bed may still contain air, both between the particles and in smaller particles within particles.

$$\boldsymbol{Bed\ porosity}\ (\%) = \left(\mathbf{1} - \frac{\boldsymbol{Unpacked\ Density\ (mg/cm^3)}}{\boldsymbol{Packed\ Density\ (mg/cm^3)}}\right) \times \mathbf{100} \tag{1}$$

Evaluation of Gas Adsorption Capacity of the Freshness-Preserving

Apples, known for releasing substantial amounts of volatile organic compounds and ethylene gas (Fidler and North 1969), were selected as the test crop. The apples (*Malus pumila* Mill.) used in the experiment were purchased from a local market. Only apples harvested at the mature stage, exhibiting uniform color, and free from defects or mechanical damage were selected. The selected apples were sealed in airtight plastic containers together with the freshness-preserving agent. A tube was connected from the sealed container to a gas analyzer (AMO-1000, ASN, Korea) to measure the release of volatile organic compounds and ethylene gas. The volatile organic compounds and ethylene gas adsorption performance were calculated based on the control group without the freshness-preserving agent (Table 1). The gas concentrations released from both the control and treated groups (Table 1, Sample No. 21) reached a plateau at 7 days of storage, indicating adsorption equilibrium (Fig. 1). Stabilization of gas concentrations was further confirmed by extending the observation period to 28 days, which showed no statistically significant differences from the 7-day values. Consequently, the data obtained at this 7-day equilibrium point were established as the baseline to normalize and calculate the relative gas adsorption percentages for all experimental groups.

Table 1. Volatile Organic Compounds, Ethylene Gas Adsorption Raw Data Table According to Composition and Freshness-preserving Characteristics

	Input Data Variables							Output Data Variables	
No.	PS (mm)	TWC (g)	CFC (g)	TC (g)	BD (Mg/m^3)	CD (Mg/m^3)	PO (%)	VOA (%)	EGA (%)
1	0	0	4	4	1.15	4.34	73.46	30	47.1
4	1	1	3	4	0.13	0.4	68.99	30	55.5
7	1	2	2	4	0.14	0.31	54.11	20.5	77
10	1	3	1	4	0.16	0.29	46.72	30.5	68.1
13	1	4	0	4	0.27	0.51	46.03	25.5	62.8
16	2	1	3	4	0.3	0.91	67.37	20	62.8
19	2	2	2	4	0.16	0.52	69.43	10.5	83.2
22	2	3	1	4	0.15	0.52	70.66	11	79.6
25	2	4	0	4	0.14	0.61	77.94	25.5	64.7
28	4	1	3	4	0.13	0.58	78.22	30	47.1
31	4	2	2	4	0.1	0.65	84.02	25.5	70.7
34	4	3	1	4	0.13	0.85	85.17	25.5	60.7
37	4	4	0	4	0.17	2.11	91.9	25.5	58
2	0	0	7	7	0.12	0.49	75.57	25	52.9
5	1	2	5	7	0.17	0.43	61.78	25.5	60.2
8	1	3.5	3.5	7	0.16	0.33	51.41	20.5	78
11	1	5	2	7	0.13	0.26	49.17	25.5	73.3
14	1	7	0	7	0.49	0.92	46.97	20.5	71.7
17	2	2	5	7	0.28	0.91	69.39	20	68.6
20	2	3.5	3.5	7	0.16	0.49	67.55	10.5	84.3
23	2	5	2	7	0.15	0.5	70.39	15	80.1
26	2	7	0	7	0.09	0.44	79.38	10.5	73.5
29	4	2	5	7	0.1	0.54	80.57	30	57.6
32	4	3.5	3.5	7	0.12	0.97	87.18	20.5	70.7
35	4	5	2	7	0.08	0.6	85.77	25.5	62.8
38	4	7	0	7	0.17	1.9	91.04	25.5	71.2
3	0	0	14	14	0.09	0.38	75.76	31	63.4
6	1	4	10	14	0.16	0.39	59.88	25.5	62.8
9	1	7	7	14	0.11	0.25	54.63	15.5	79.1
12	1	10	4	14	1.15	1.95	41.05	25.5	75.4
15	1	14	0	14	0.47	0.91	48.63	20.5	73.8
18	2	4	10	14	0.13	0.41	67.63	15.5	73.3
21	2	7	7	14	0.15	0.5	69.68	10.5	84.8
24	2	10	4	14	0.13	0.45	71.91	13.5	82.7
27	2	14	0	14	0.09	0.41	77.6	10.5	71.1
30	4	4	10	14	0.14	0.7	80.61	30	68.6
33	4	7	7	14	0.13	0.9	85.75	20	78
36	4	10	4	14	0.18	1.45	87.3	25.5	62.8
39	4	14	0	14	0.11	1.15	90.26	20.5	72.5

PS: Particle size of torrefied oak wood, TWC: torrefied wood content, CFC: commercial freshness-preserving content, TC: Total content of torrefied wood and commercial freshness-preserving BD: bulk density, CD: Compressed density, PO: porosity, VOA: volatile organic compounds absorption, EGA: ethylene gas absorption

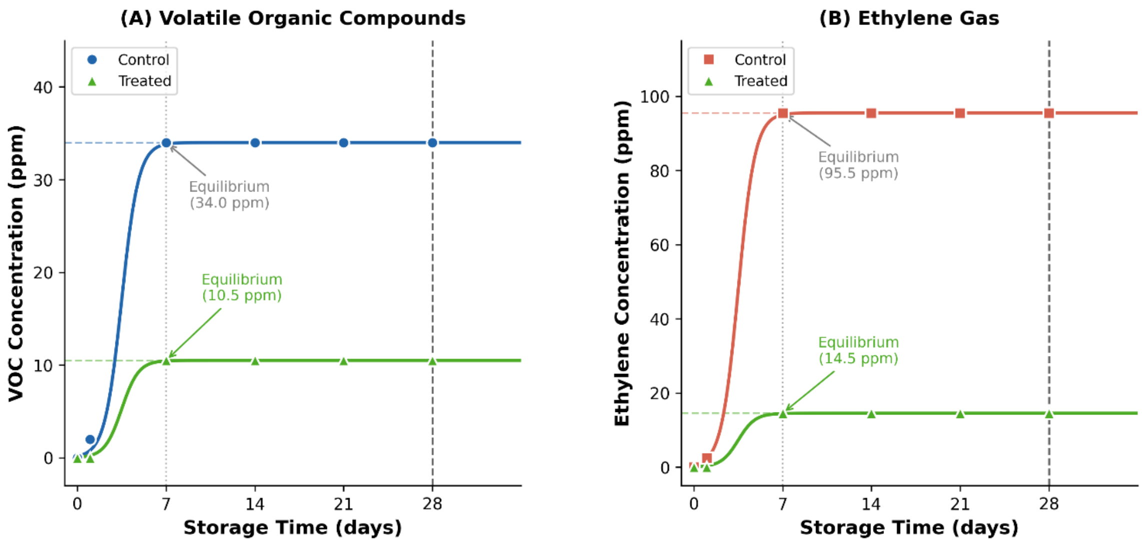

Fig. 1. Kinetic release and adsorption of (A) volatile organic compounds and (B) ethylene gas from apples with and without freshness-preserving agent treatment during 28 days of storage

Prior to the adsorption experiments, the amounts of volatile organic compounds (VOC) and ethylene gas released from apple samples were quantified to establish baseline conditions. As illustrated in Fig. 1, the concentrations of VOC and ethylene gas within the sealed containers gradually increased during the storage period, reaching equilibrium at approximately 7 days. At this equilibrium point, the VOC concentration of the control group (without freshness-preserving agent) was approximately 34.0 ppm, and the ethylene concentration was approximately 95.5 ppm (Fig. 1). Based on these values, the amount of gas absorbed by the freshness-preserving agent was expressed as a percentage calculated from the gas concentrations detected at Day 7 across all experimental groups (Table 1).

Data Augmentation and Preprocessing

Traditionally, machine learning models operate more effectively when provided with a broad range of input data to prevent prediction errors (Shorten and Khoshgoftaar 2019). In this study, the experimental dataset (Table 1) consisted of 39 data entries, categorized into eight input variables particle size, torrefied material content, raw material content, final bulk density, specific density, and porosity and two output variables, namely volatile organic compounds adsorption rate and ethylene adsorption rate. The size of this dataset was insufficient to develop a predictive model with strong generalization capability (Mikołajczyk and Grochowski 2018). Therefore, data augmentation was applied to enhance the reliability of the metadata and to simulate real-world conditions (Feng *et al.* 2019) (Eq. 2).

$$\tilde{\boldsymbol{X}}_{\boldsymbol{ij}}{}^{(r)} = \boldsymbol{X}_{\boldsymbol{ij}}\left(\boldsymbol{1}+\boldsymbol{\epsilon}_{\boldsymbol{ij}}^{(r)}\right),\ \boldsymbol{\epsilon}_{\boldsymbol{ij}}^{(r)} \sim \boldsymbol{\mu}(-\boldsymbol{\lambda},\boldsymbol{\lambda}) \tag{2}$$

$$\tilde{\boldsymbol{y}}_{\boldsymbol{it}}{}^{(r)} = \boldsymbol{X}_{\boldsymbol{it}}\left(\boldsymbol{1}+\boldsymbol{\delta}_{\boldsymbol{it}}^{(r)}\right),\ \boldsymbol{\delta}_{\boldsymbol{it}}^{(r)} \sim \boldsymbol{\mu}(-\boldsymbol{\lambda},\boldsymbol{\lambda}) \tag{3}$$

As shown in Eq. (2), for each observation i, feature j, and target dimension t, r denotes the augmentation iteration index, and λ corresponds to the noise level ($0<\lambda<1$). The terms ϵ and δ are mutually independent and are independently sampled across observations and features. The final training matrix is expressed in Eq. 4.

$$\boldsymbol{X}^* = \begin{bmatrix} \tilde{\boldsymbol{X}}^{(1)} \\ \dots \\ \tilde{\boldsymbol{X}}^{(r)} \end{bmatrix}, \boldsymbol{y}^* = \begin{bmatrix} \tilde{\boldsymbol{y}}^{(1)} \\ \dots \\ \tilde{\boldsymbol{y}}^{(r)} \end{bmatrix} \tag{4}$$

Predictive Modeling of Gas Adsorption Performance for Freshness-preserving Agent

Machine learning algorithms were employed to construct predictive models for volatile organic compounds and ethylene gas adsorption performance of the freshness-preserving agent. Three regression algorithms were selected based on their ability to capture non-linear relationships between input features and adsorption performance: Multiple Linear Regression *(MLR),* Random Forest (RF), Extreme Gradient Boosting (XGB), and Support Vector Regression (SVR).

Multiple Linear Regression (MLR) baseline model

Multiple Linear Regression (MLR) was employed as a baseline reference model to evaluate whether the relationship between the eight input variables and the two adsorption output variables (VOC and ethylene adsorption rates) could be adequately described by a linear additive framework. MLR assumes that each output variable is a weighted linear combination of the input features, making it one of the most interpretable and computationally efficient modeling approaches in statistical analysis. The MLR model assumes a linear relationship of the form (5),

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \varepsilon \tag{5}$$

where $\hat{y}$ is the predicted adsorption rate, β_0 is the intercept, β_1 through β_m are regression coefficients for each of the m = 8 input variables x_1 through x_m, and ε is the residual error term. Model parameters were estimated by ordinary least squares (OLS) using scikit-learn LinearRegression (Pedregosa *et al*. 2011). No regularization was applied in order to represent the standard MLR formulation as a baseline.

Random forest model

Random Forest (RF) is an ensemble learning algorithm that integrates multiple decision trees to enhance prediction accuracy and reduce the risk of overfitting (Breiman 2001). Each tree is constructed from a random subset of training data and features, and the final prediction is obtained by averaging the outputs of all trees. This approach is particularly effective in modeling non-linear relationships and handling high-dimensional feature spaces (Friedman 2001; Meinshausen and Ridgeway 2006). In the present study, RF was employed to capture the complex interactions among structural parameters of the freshness-preserving agent and their influence on adsorption performance. The hyperparameter types and search ranges for RF are summarized in Table 2.

Table 2. Hyperparameter Search Ranges and Optimal Values of Random Forest (RF) Model for Predicting Ethylene and Volatile Organic Compounds Adsorption Performance

Hyperparameters	Range	Number of Cases	
		Ethylene	Volatile organic compounds
Number of estimators	10 to 1000	651	137
Maximum depth	1 to 30	27	23
Minimum samples split	1 to 10	2	3
Minimum samples leaf	1 to 5	1	1
Maximum features	sqrt, 0.5, 0.7, 1.0	1.0	1.0

Extreme gradient boosting model

Extreme Gradient Boosting (XGB) is an advanced boosting algorithm designed to optimize both computational efficiency and prediction performance (Chen and Guestrin 2016). Unlike bagging-based approaches, XGB sequentially constructs trees where each subsequent tree corrects the errors of the previous ones (Friedman 2001). By applying regularization techniques and optimized gradient descent, XGB minimizes overfitting while maintaining high prediction accuracy (Chen and Guestrin 2016). In this research, XGB was selected to provide robust predictions of volatile organic compound and ethylene gas adsorption performance, owing to its proven capability in modeling highly non-linear and imbalanced datasets (Nielsen 2016; Torlay *et al.* 2017). The hyperparameter types and search ranges for XGB are summarized in Table 3.

Table 3. Hyperparameter Search Ranges and Optimal Values of Extreme Gradient Boosting (XGB) Model for Predicting Ethylene and Volatile Organic Compounds Adsorption Performance

Hyperparameters	Range	Number of Cases	
		Ethylene	Volatile organic compounds
number of estimators	10 to 1000	632	708
learning rate	0.02 to 0.2	0.1713	0.2579
maximum depth	1 to 20	2	13
subsample ratio	0.3 to 1.0	0.9236	0.7299
column sample ratio	0.4 to 0.9	0.6347	0.8021

Support vector regression model

Support Vector Regression (SVR) is a kernel-based learning algorithm that maps input features into higher-dimensional space to establish a regression function within a specified margin of tolerance (Drucker *et al*. 1996). SVR seeks to minimize prediction error while maintaining model generalization by employing kernel functions such as radial basis function (RBF). This makes SVR particularly suitable for datasets with limited sample sizes or non-linear patterns. In the current study, SVR was implemented to evaluate its ability to generalize the adsorption behavior of the freshness-preserving agent across varying experimental conditions. The hyperparameter types and search ranges for RF are summarized in Table 4.

Table 4. Hyperparameter Search Ranges and Optimal Values of Support Vector Regression (SVR) Model for Predicting Ethylene and Volatile Organic Compounds Adsorption Performance

Hyperparameters	Range	Number of Cases	
		Ethylene	Volatile organic compounds
kernel	rbf, poly, sigmoid	rbf	rbf
penalty parameter C	0.1 ~ 100	6.4511	68.4263
epsilon	0.001 ~ 1	0.2744	0.0353
polynomial degree	2 to 6	4	5
Gamma	'scale', 'auto', 0.001, 0.01, 0.1, 1	1	1
coef0	0.0 to 0.5	0.0368	0.6099

Hyperparameter tuning

To ensure optimal predictive performance, hyperparameter tuning was conducted for all three regression algorithms (RF, XGB, and SVR). The RandomizedSearchCV method was employed with 120 randomized iterations combined with five-fold cross-validation (Bergstra and Bengio 2012). This approach enabled systematic exploration of parameter spaces while maintaining computational efficiency (Pedregosa *et al.* 2011). Through this process, the most appropriate parameter combinations for each algorithm were identified, thereby reducing the risk of overfitting and enhancing model generalization to unseen data (Kohavi 1995).

Model performance evaluation

The predictive accuracy and generalization capability of the machine learning models were evaluated using three statistical indicators: the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) (Willmott and Matsuura 2005). R^2 represents the proportion of variance in the observed data that can be explained by the model, with values closer to 1.0 indicating superior predictive performance (Eq. 6) (Kvålseth 1985). RMSE is defined as the square root of the mean of the squared differences between the predicted and observed values, providing a measure of the model's prediction error in the same units as the target variable (Eq. 7) (Chai and Draxler 2014). MAE calculates the mean of the absolute differences between the predicted and observed values, reflecting the average magnitude of prediction errors regardless of direction (Eq. 8) (Willmott and Matsuura 2005). A well-trained model should yield R^2 values close to 1.0, while RMSE and MAE should be as low as possible (Chicco *et al.* 2021). These metrics were computed for both the training and validation datasets to ensure that the models achieved high accuracy without overfitting.

$$R^2 = 1 - \frac{\sum(t-y)^2}{\sum(t-\bar{t})^2} \tag{6}$$

$$RMSE = \sqrt{\frac{\sum(y-\hat{y})^2}{n}} \tag{7}$$

$$MAE = \frac{\sum|y-\hat{y}|}{n} \tag{8}$$

RESULTS AND DISCUSSION

Characterization of Torrefied Oak Wood Chips

The physicochemical and adsorption-related properties of torrefied oak wood chips were characterized to provide a comprehensive understanding of the material prior to freshness-preserving agent preparation. Torrefaction at 350 °C for 20 min under limited-oxygen conditions induced substantial structural transformations, including partial decomposition of hemicellulose and amorphous cellulose, resulting in a biochar-like porous architecture with enhanced surface hydrophobicity (Chen and Kuo 2011; Niu *et al*. 2019). These thermally induced changes are known to increase the specific surface area and generate microporous and mesoporous structures conducive to gas adsorption (Doddapaneni and Kikas. 2023).

The textural parameters of the torrefied oak wood chips, including bulk density (BD) and compressed density, were determined experimentally and are reported in Table 1. The bulk density of torrefied wood fractions ranged from 0.08 to 1.15 Mg/m³, and the compressed density ranged from 0.25 to 4.34 Mg/m³, reflecting the influence of particle size and formulation composition on the packing structure. The calculated porosity values ranged from 41.05% to 91.90% (Table 1), indicating a highly porous material framework. These textural parameters are directly relevant to gas adsorption performance, as porosity governs the availability of internal void spaces for molecular capture (Rouquerol *et al*. 2013).

Surface morphology of the torrefied oak wood chips was qualitatively consistent with the characteristics reported for biomass torrefied at similar temperatures, including roughened surface textures and the development of inter-particle void spaces due to thermal degradation of cell wall components (Simonic *et al*. 2020; Xu *et al*. 2024). Surface functional group changes associated with torrefaction—including the reduction of hydroxyl (-OH) groups and the formation of carbonyl (C=O) and aromatic C=C structures—are well-documented and contribute to the enhanced hydrophobicity and altered adsorption affinity of the material (Xu *et al*. 2024).

Correlation Analysis of Structural Parameters and Gas Adsorption Performance

The Pearson correlation matrix (Fig. 2) reveals distinct and contrasting relationships between the structural parameters of the freshness-preserving agent and the two target gas adsorption performances. Interestingly, ethylene adsorption showed a negative correlation with porosity ($r = -0.25$), which appears counterintuitive but can be explained by the specific pore size requirements for ethylene molecules (2.85 Å kinetic diameter) (Keller and Staudt 2005). Excessive porosity, primarily representing inter-particle gaps in this study, may result in inefficient gas-solid interaction, thereby hindering the effective capture of ethylene molecules.

Torrefied oak wood content showed the strongest positive correlation with ethylene adsorption ($r = 0.43$), indicating that increasing the proportion of torrefied wood directly enhances ethylene capture capacity, attributable to the hydrophobic surface characteristics and residual microporous structure developed during torrefaction. Compressed density exhibited a notable negative correlation with ethylene adsorption ($r = -0.36$), suggesting that denser material configurations are less conducive to ethylene uptake, while bulk density showed a weak negative correlation ($r = -0.16$), consistent with the observation that lower bulk density alone does not favor ethylene adsorption when pore sizes exceed the optimal range for molecular capture (Yang 2003).

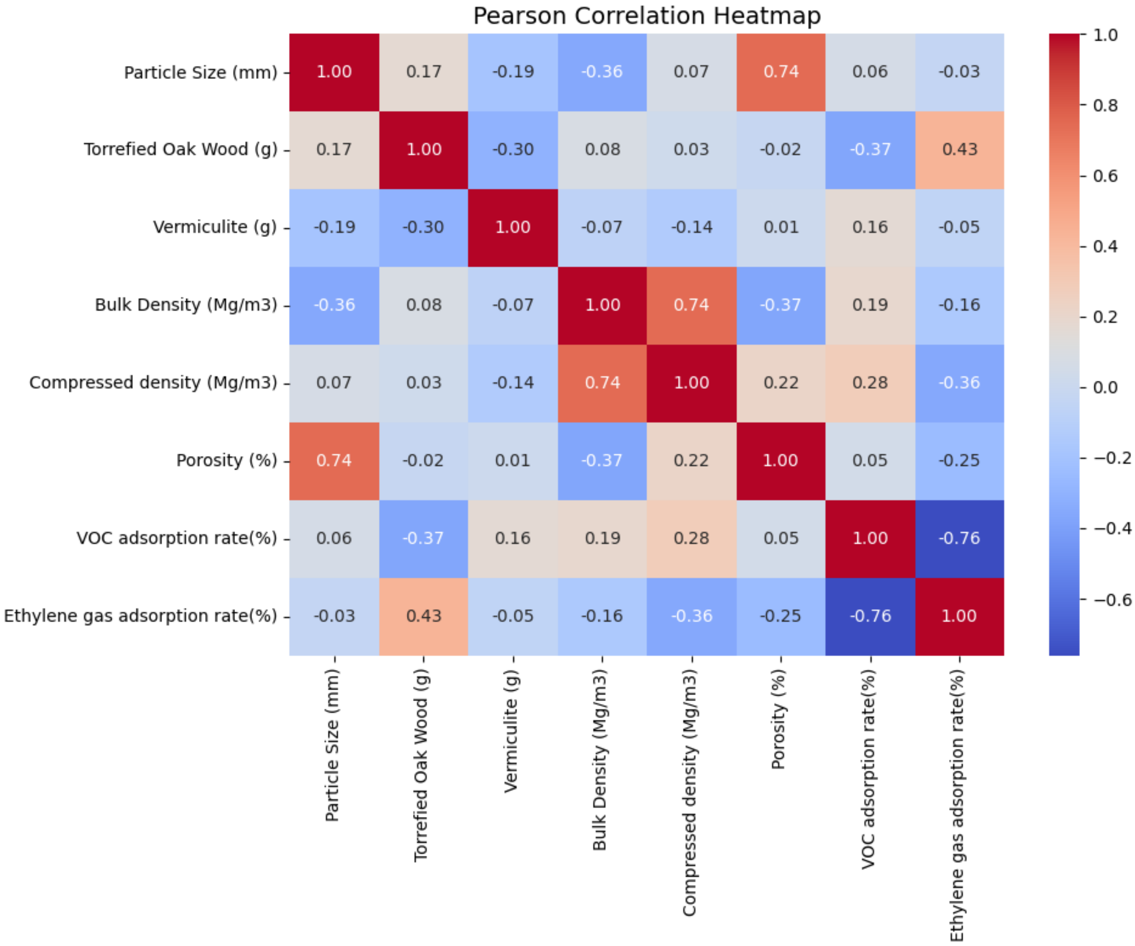

Fig. 2. Pearson correlation analysis between ethylene and volatile organic compounds adsorption and the prepared freshness-preserving agent

Volatile organic compounds adsorption showed a negative correlation with torrefied oak wood content (r = -0.37), suggesting that an increase in torrefied wood content led to diminished effectiveness in the adsorption of larger, more complex VOCs (Bansal and Goyal 2005). This result may be attributable to the surface chemical characteristics of the torrefied material, where hydrophobic surfaces developed during torrefaction show limited affinity for the polar functional groups common in VOC molecules (Chen *et al.* 2014). A weak positive correlation with vermiculite content (r = 0.16) suggests that the commercial base material contributed more to VOC capture, possibly through surface adsorption mechanisms rather than pore-filling (Yang 2003). Bulk density showed a modest positive correlation with VOC adsorption (r = 0.19), thereby providing substantiation for the hypothesis that denser material configurations may facilitate greater surface contact points per unit volume for VOC capture (Rouquerol *et al.* 2013). The strong negative intercorrelation between VOC and ethylene adsorption rates (r = -0.76) is particularly significant, demonstrating a fundamental trade-off between the two adsorption targets governed by contrasting structural requirements. This finding justifies the gas-specific model development strategy adopted in the present study, wherein SVR and RF were independently optimized for ethylene and VOC prediction, respectively, underscoring the necessity of machine learning approaches capable of capturing these complex, gas-specific nonlinear interactions across the multidimensional input feature space.

Machine Learning Model Hyperparameter Optimization

Hyperparameter tuning was conducted for all three regression algorithms (RF, XGB, and SVR) using the RandomizedSearchCV method with 120 randomized iterations combined with five-fold cross-validation. This approach enabled systematic exploration of parameter spaces while maintaining computational efficiency.

Random Forest Hyperparameter Optimization

The optimal hyperparameters for RF showed distinct patterns between gas types (Table 2). For ethylene prediction, RF required a larger ensemble (n_estimators = 651) with deeper trees (max_depth = 27), indicating the need for complex model architecture to capture ethylene's intricate adsorption behavior. In contrast, volatile organic compounds prediction achieved optimal performance with a smaller ensemble (n_estimators = 137) and shallower trees (max_depth = 23), suggesting that volatile organic compounds adsorption patterns are more straightforward to model. Both tasks benefited from minimal overfitting control (min_samples_split = 2-3, min_samples_leaf = 1) and access to all features (max_features = 1.0).

XGBoost Hyperparameter Optimization

XGBoost optimization revealed significant differences in optimal configurations between gas types (Table 3). Ethylene prediction required a moderate ensemble size (n_estimators = 632) with shallow trees (max_depth = 2) and conservative learning rate (learning_rate = 0.1713), emphasizing the need for gradual learning to avoid overfitting on the complex ethylene adsorption patterns. Volatile organic compounds prediction benefited from a larger ensemble (n_estimators = 708) with deeper trees (max_depth = 13) and higher learning rate (learning_rate = 0.2579), indicating that volatile organic compounds patterns can accommodate more aggressive learning strategies. The subsample and column sample ratios differed substantially (ethylene: subsample = 0.9236, colsample = 0.6347; volatile organic compounds: subsample = 0.7299, colsample = 0.8021), suggesting different regularization requirements for each gas type.

Support Vector Regression Hyperparameter Optimization

SVR optimization demonstrated contrasting parameter requirements between gas predictions (Table 4). Both applications utilized RBF kernels with gamma = 1, confirming the effectiveness of radial basis function mapping for capturing non-linear gas adsorption relationships. However, the penalty parameter C showed dramatic differences: ethylene required moderate regularization (C = 6.4511) while volatile organic compounds demanded strong regularization (C = 68.4263), indicating that volatile organic compounds prediction requires more flexible decision boundaries. The epsilon values also varied significantly (ethylene: ε = 0.2744; volatile organic compounds: ε = 0.0353), with volatile organic compounds requiring tighter tolerance zones for accurate predictions.

Model Performance Evaluation and Comparison

Ethylene gas adsorption prediction

The comparative analysis of machine learning models for ethylene adsorption prediction (Table 5, Fig. 3) revealed distinct performance characteristics. The MLR model yielded the lowest predictive performance, with a validation R^2 of 0.303 and high prediction errors (RMSE = 53.457, MAE = 5.512), confirming that the relationships governing ethylene adsorption cannot be adequately described by a linear additive framework. SVR demonstrated superior performance with the highest R^2 value of 0.934

on validation data and the lowest prediction errors (RMSE = 5.06, MAE = 1.997). The tight alignment between predicted and experimental values in Fig. 3 indicates excellent generalization capability. This performance can be attributed to SVR's kernel-based approach, which effectively captures the non-linear relationships between structural parameters and ethylene adsorption through radial basis function mapping. The XGB model also achieved reliable performance (R^2 = 0.926) with moderate prediction errors (RMSE = 5.65, MAE = 1.56). While slightly less accurate than SVR, XGB maintained consistent performance across training and validation datasets, indicating good model stability without severe overfitting. In contrast, the Random Forest (RF) model displayed wider scatter and lower accuracy (R^2 = 0.879), indicating its limited capacity to capture the nonlinear adsorption mechanisms of ethylene. The substantial difference between training (R^2 = 0.914) and validation performance (R^2 = 0.879) suggests that RF struggled with generalization despite requiring the most complex hyperparameter configuration.

Table 5. Comparison of Training and Validation Performance of MLR, RF, SVR, and XGB Models in Ethylene Gas Adsorption Prediction

Model	R^2_train	RMSE_train	MAE_train	R^2_Test	RMSE_val	MAE_val
Ethylene Gas Adsorption Predict_MLR	0.301	67.195	7.106	0.303	53.457	5.512
Ethylene Gas Adsorption Predict_RF	0.914	8.219	2.339	0.879	9.268	2.431
Ethylene Gas Adsorption Predict_SVR	0.951	4.700	1.828	0.933	5.060	1.997
Ethylene Gas Adsorption Predict_XGB	0.945	5.259	1.706	0.926	5.648	1.562

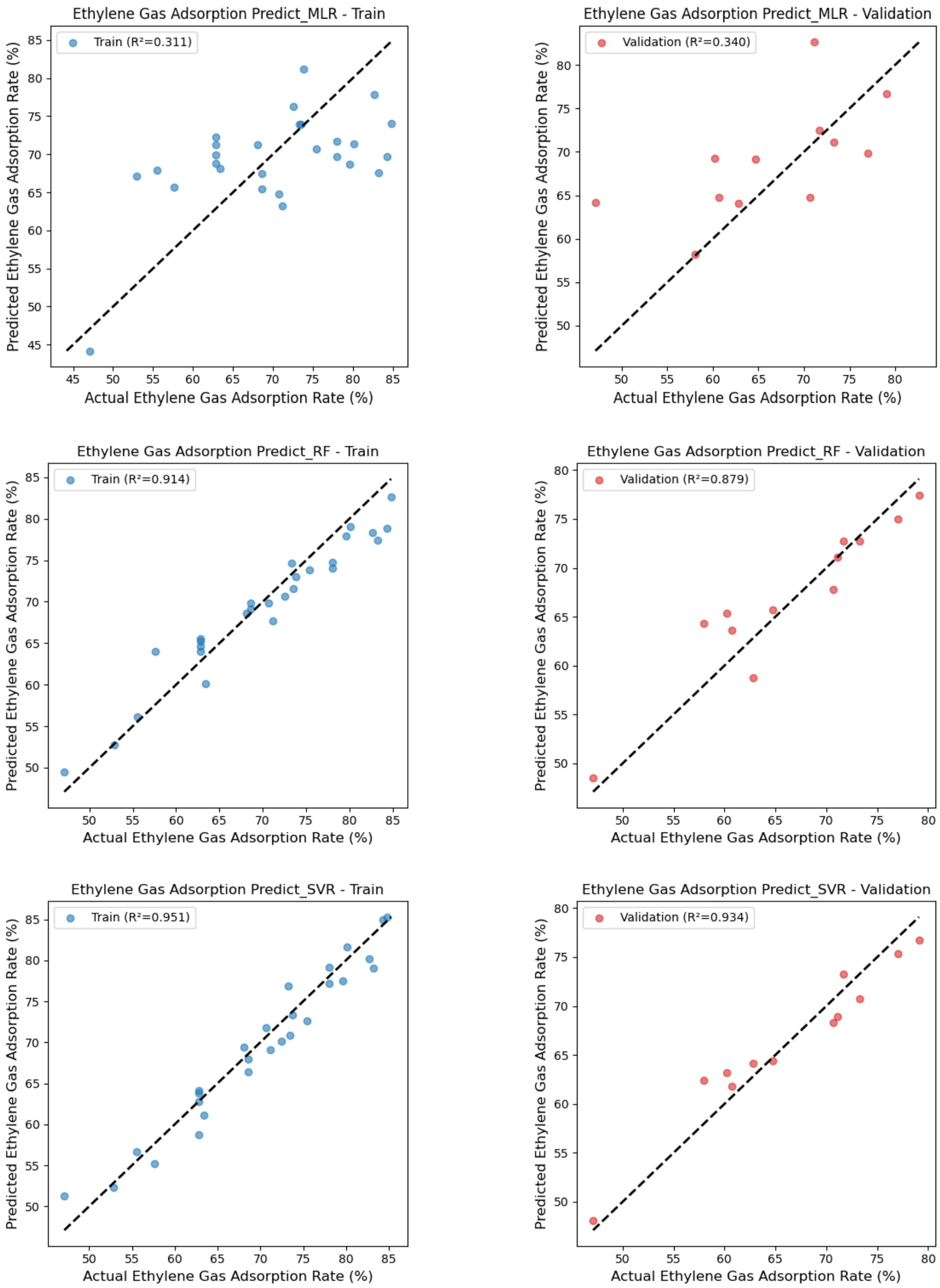

Fig. 3. Comparison of predicted and experimental values for ethylene gas adsorption using RF, SVR, and XGB models

Volatile Organic Compounds Adsorption Prediction

In contrast to ethylene, the adsorption of volatile organic compounds was best predicted by the RF model (Table 6, Fig. 4). The MLR model yielded the lowest predictive performance, with a validation R^2 of 0.412 and high prediction errors (RMSE = 67.195, MAE = 7.106), confirming that VOC adsorption behavior cannot be adequately described by a linear additive framework. RF achieved the highest R^2 value (0.962) with very low prediction error (RMSE = 1.11, MAE = 0.845), outperforming both SVR and XGB. The minimal difference between training (R^2 = 0.957) and validation performance (R^2 = 0.962) demonstrates excellent generalization without overfitting. The ensemble learning structure of RF effectively handles interactions among multiple structural variables, such as porosity, bulk density, and particle size. The SVR model initially showed high training accuracy ($R^2 = 0.990$) but suffered from reduced performance in validation ($R^2 = 0.931$), indicating mild overfitting.

Despite the high penalty parameter (C = 68.4263) intended to prevent overfitting, the model learned training-specific patterns that did not generalize well to new volatile organic compounds adsorption scenarios. XGB delivered intermediate accuracy ($R^2 = 0.911$) between SVR and RF, maintaining reasonable prediction capability without severe overfitting issues. The balanced hyperparameter configuration (deeper trees with moderate learning rate) successfully captured volatile organic compounds adsorption complexity while maintaining model stability.

Table 6. Comparison of Training and Validation Performance of MLR, RF, SVR, and XGB Models in Volatile Organic Compounds Adsorption Prediction

Model	R^2_train	RMSE_train	MAE_train	R^2_Test	RMSE_val	MAE_val
VOC Adsorption Predict_MLR	0.301	67.195	7.106	0.303	53.457	5.512
VOC Adsorption Predict_RF	0.956	2.078	1.228	0.961	1.114	0.845
VOC Adsorption Predict_SVR	0.989	0.499	0.440	0.930	2.019	0.778
VOC Adsorption Predict_XGB	0.983	0.788	0.638	0.911	2.586	1.165

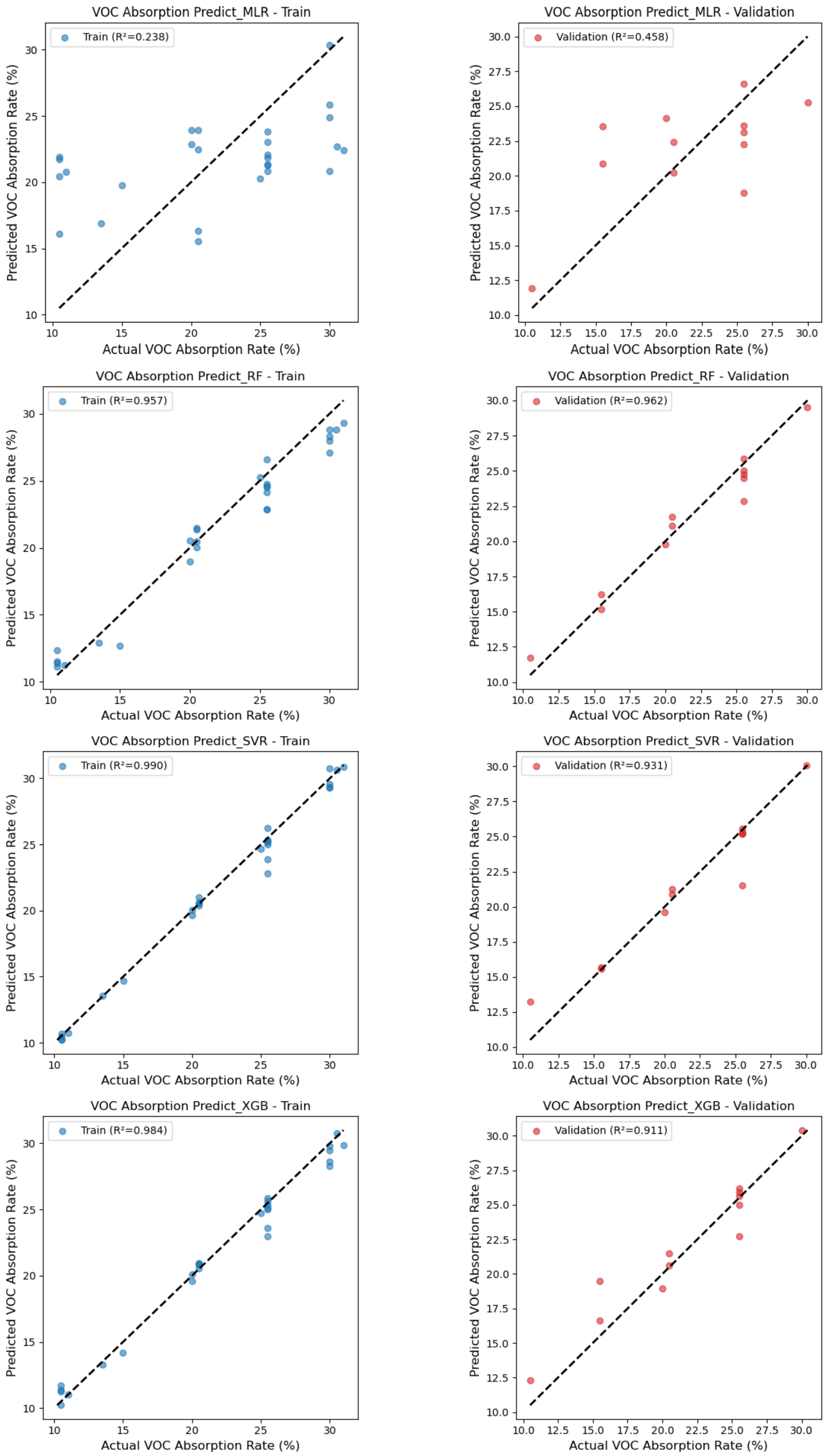

Fig. 4. Comparison of predicted and experimental values for volatile organic compounds adsorption using RF, SVR, and XGB models

Optimal Model Selection and Performance Analysis

SVR is recommended as the optimal choice for ethylene adsorption prediction due to its superior generalization capability ($R^2 = 0.934$) and robust performance under limited data conditions (Vapnik 2013; Smola and Schölkopf 2004). The kernel-based approach effectively models the penetration-dependent adsorption behavior of small ethylene molecules, where pore structure optimization is critical (Cortes and Vapnik 1995). The moderate hyperparameter requirements ($C = 6.4511$, $\varepsilon = 0.2744$) indicate that ethylene adsorption follows well-defined physical principles that can be captured without excessive model complexity (Cherkassky and Ma 2004). RF is preferred for volatile organic compounds adsorption prediction due to its exceptional accuracy ($R^2 = 0.962$) and ability to capture multi-variable interactions without overfitting (Breiman 2001). The ensemble approach successfully models the complex surface-interaction mechanisms that characterize volatile organic compounds adsorption on torrefied materials (Liaw and Wiener 2002). The simpler hyperparameter configuration required for optimal volatile organic compounds prediction (n_estimators = 137) demonstrates the efficiency of RF in handling multi-dimensional feature interactions (Hastie *et al.* 2009).

Comparison with Literature and Model Validation

Although conventional multiple linear regression (MLR) can describe individual pairwise relationships, it is unable to adequately capture the nonlinear cross-interactions among the eight structural input variables that simultaneously influence two distinct adsorption outputs. In the present study, MLR was evaluated as a baseline, yielding low predictive accuracy ($R^2 = 0.303$ for ethylene; $R^2 = 0.412$ for VOC), which confirms that the multi-variable relationships governing adsorption performance in this system cannot be adequately described by linear models alone (Tables 5, 6). Machine learning approaches, which can model complex nonlinear interactions across the full variable space, are therefore justified for this application.

The SVR model achieved the highest predictive accuracy for ethylene adsorption ($R^2 = 0.934$), outperforming all other evaluated algorithms in this study. This result compares favorably with Li *et al.* (2023), who reported $R^2 = 0.95$ using a light gradient boosting machine (LGBM) for CO_2 adsorption prediction of biochar based on activation temperature and gas flow rate as key input features. The comparable R^2 values between the two studies are particularly noteworthy given that the present study predicted ethylene adsorption directly under practical postharvest storage conditions, whereas Li *et al.* (2023) targeted indirect structural properties under idealized laboratory settings. For VOC adsorption prediction, RF achieved the highest R^2 of 0.962 among all models evaluated, surpassing recent deep learning and ensemble-based approaches reported in the literature. Specifically, Zhang *et al.* (2025) achieved $R^2 = 0.97$ using gradient boosting machines with total pore volume and micropore volume as primary features for biochar VOC adsorption prediction. Jeong *et al.* (2025) reported $R^2 = 0.94$ using a multimodal-based prediction model with VOC type and BET surface area as key input features for activated carbon VOC adsorption. Zhou *et al.* (2025) achieved $R^2 = 0.89$ using a physics-embedded deep neural network (PEDNN) with mesopore volume and total pore volume as key features. The RF model in this study achieved an $R^2 = 0.962$, demonstrating high accuracy comparable to the performance of reported machine learning models. This was achieved despite the experimental data conditions derived from actual reservoir preservation tests being far more restrictive than those of large-scale pure gas injection systems (Table 7).

Table 7. Comparison of Machine Learning Model Performance in Gas Adsorption Literature

Adsorbent Material	Target Compound	ML Algorithm	R^2	Key Features	Study
Torrefied oak wood	Ethylene	SVR	0.934	Particle size, porosity, density	Present Study
Torrefied oak wood	VOC	RF	0.962	Particle size, porosity, density	Present Study
Biochar	CO_2	LGBM	0.95	Activation temperature, Gas flow rate	Li *et al*., 2023
Biochar	VOC	Gradient boosting machines	0.97	Total pore volume, Micropore volume	Zhang *et al.* 2025
Activated carbons	VOC	Multimodal-based prediction model	0.94	VOC type BETdata	Jeong *et al.* 2025
Activated carbons	VOC	PEDNN	0.89	mesopore volume total pore volume	Zhou *et al.* 2025

LGBM: Light Gradient boosting machines
BET: Brunauer-Emmett-Teller
PEDNN: physical experience-driven neural network

Comparison with Other Biomass Adsorbents

To place the adsorption performance of the torrefied oak wood chips into a broader context, the results of the present study were compared with previously reported studies employing thermochemically treated biomass materials as gas adsorbents (Table 8). Previous studies have demonstrated that biochar and heat-treated biomass materials exhibit strong adsorption capacity toward ethylene and various volatile organic compounds due to their developed porous structures and increased surface functionality (Dutta *et al.* 2017; Doddapaneni *et al.* 2018). For example, Charoensuk *et al*. (2024) reported ethylene removal efficiencies of approximately 60 to 85% using agricultural biochar materials produced by pyrolysis. Similarly, Rajabi *et al.* (2021) demonstrated that biochar materials with well-developed microporous structures can effectively adsorb a wide range of VOCs. The adsorption efficiencies observed in the present study (ethylene adsorption up to 84.8%) fall within the range reported for thermally treated biomass adsorbents in the literature. These findings confirm that torrefied oak wood chips exhibit adsorption performance comparable to other biomass-derived adsorbents, while offering the additional advantage of being produced from forest biomass resources using relatively mild thermal treatment conditions.

Table 8. Comparison of Adsorption Performance of Thermally Treated Biomass Adsorbents Reported in Previous Studies

Study	Adsorbent Material	Target Gas	Adsorption Performance	Key Conditions
Charoensuk *et al.* (2024)	Agricultural biochar	Ethylene	60 to 85% ethylene removal	pyrolysis 400 to 600 °C
Rajabi et al. (2021)	Biochar	VOCs	65 to 90% VOC adsorption	high surface area biochar
Doddapaneni *et al.* (2018)	Torrefied biomass	VOCs	enhanced adsorption after torrefaction	250 to 300 °C
Dutta *et al.* (2017)	Biochar	Various VOCs	high adsorption capacity due to micropores	review study
Present study	Torrefied oak wood chips	Ethylene, VOCs	47 to 84% ethylene adsorption	torrefaction 350 °C

CONCLUSIONS

1. This study demonstrated the feasibility and effectiveness of using machine learning approaches to predict gas adsorption performance of torrefied oak-based freshness-preserving agents.
2. Key findings include: (1) Gas-specific optimization strategies are essential, as evidenced by the strong negative correlation (r = -0.76) between ethylene and volatile organic compounds adsorption performance; (2) Support Vector Regression (SVR) provides optimal prediction accuracy for ethylene adsorption (R^2 = 0.934), while Random Forest (RF) excels in volatile organic compounds adsorption prediction (R^2 = 0.962); (3) Hyperparameter optimization reveals that ethylene prediction requires more conservative model configurations while volatile organic compounds prediction benefits from more complex ensemble approaches; (4) The structural parameter analysis indicates that ethylene adsorption was negatively correlated with porosity (r = −0.25) and positively correlated with torrefied wood content (r = 0.43), while volatile organic compounds showed contrasting trends, suggesting that the two adsorption targets were governed by distinct physicochemical interactions that warrant further mechanistic investigation using pore characterization techniques such as nitrogen adsorption analysis.
3. The integration of torrefied biomass materials with machine learning prediction models offers a sustainable and efficient approach to freshness preservation technology, contributing to food waste reduction and improved storage efficiency in agricultural and food industry applications.

ACKNOWLEDGMENTS

This study was carried out with the support of ´R&D Program for Forest Science Technology (Project No. “RS-2023-KF002602613821165300003”)´ provided by Korea Forest Service (Korea Forestry Promotion Institute).

Data Availability

All datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Use of Generative AI

The authors acknowledge the use of AI assistance during manuscript preparation. This includes English translation, formatting of the Experimental section, formatting of references, and language refinement. All AI-generated content was thoroughly reviewed and verified by the authors to ensure accuracy and adherence to scientific standards.

REFERENCE CITED

Amidon, T. E., and Liu, S. (2009). "Water-based woody biorefinery," *Biotechnology Advances* 27(5), 542-550. https://doi.org/10.1016/j.biotechadv.2009.04.012

ASTM D2395-17 (2017). "Standard test methods for density and specific gravity (relative density) of wood and wood-based materials," American Society for Testing and Materials, West Conshohocken, PA.

Bansal, R. C., and Goyal, M. (2005). *Activated Carbon Adsorption*, CRC Press, Boca Raton, FL.

Bergstra, J., and Bengio, Y. (2012). "Random search for hyper-parameter optimization," *Journal of Machine Learning Research* 13(2), 281-305.

Berndes, G., Abt, B., Asikainen, A., Cowie, A., Dale, V., Egnell, G., Lindner, M., Marelli, L., Paré, D., Pingoud, K., and Yeh, S. (2016). "Forest biomass, carbon neutrality and climate change mitigation," *From Science to Policy* 3(7), 1-27. https://doi.org/10.36333/fs03

Breiman, L. (2001). "Random forests," *Machine Learning* 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

Brunet-Navarro, P., Jochheim, H., Cardellini, G., Richter, K., and Muys, B. (2021). "Climate mitigation by energy and material substitution of wood products has an expiry date," *Journal Cleaner Production* 303, article 127026. https://doi.org/10.1016/j.jclepro.2021.127026

Chai, T., and Draxler, R. R. (2014). "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geoscientific Model Development* 7(3), 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014

Charoensuk, P., Chaiwong, S., Suwunwong, T., Halley, P. J., and Suwantong, O. (2024). "Preparation and utilization of biochar from agricultural wastes as ethylene absorber for 'gros michel' banana ripening," *Industrial Crops and Products* 222, article 119860. https://doi.org/10.1016/j.indcrop.2024.119860

Chen, Y., Liu, B., Yang, H., Yang, Q., and Chen, H. (2014). "Evolution of functional groups and pore structure during cotton and corn stalks torrefaction and its correlation with hydrophobicity," *Fuel* 137, 41-49. https://doi.org/10.1016/j.fuel.2014.07.036

Chen, T., and Guestrin, C. (2016). "XGBoost: A scalable tree boosting system," in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM*, San Francisco, CA, pp. 785-794. https://doi.org/10.1145/2939672.2939785

Chen, W. H., Lu, K. M., Liu, S. H., Tsai, C. K., Lee, W. J., and Lin, T. C. (2019). "Biomass torrefaction: properties, applications, challenges, and economy," *Renewable and Sustainable Energy Reviews* 115, article 109395. https://doi.org/10.1016/j.rser.2019.109395

Chen, W. H., and Kuo, P. C. (2011). "Torrefaction and co-torrefaction characterization of hemicellulose, cellulose and lignin as well as torrefaction of some basic constituents in biomass," *Energy* 36(2), 803-811. https://doi.org/10.1016/j.energy.2010.12.036

Cherkassky, V., and Ma, Y. (2004). "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks* 17(1), 113-126. https://doi.org/10.1016/S0893-6080(03)00169-2

Chicco, D., Warrens, M. J., and Jurman, G. (2021). "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science* 7, article e623. https://doi.org/10.7717/peerj-cs.623

Cortes, C., and Vapnik, V. (1995). "Support-vector networks," *Machine Learning* 20(3), 273-297. https://doi.org/10.1007/BF00994018

Dalmau, D., Sigman, M. S., and Alegre-Requena, J. (2025). "Machine learning workflows beyond linear models in low-data regimes," *Chemical Science* 16(19), 8555-8560. https://doi.org/10.1039/D5SC00996K

Doddapaneni, T. R. K. C., Jain, R., Praveenkumar, R., Rintala, J., Romar, H., and Konttinen, J. (2018). "Adsorption of furfural from torrefaction condensate using torrefied biomass," *Chemical Engineering Journal* 334, 558-568. https://doi.org/10.1016/j.cej.2017.10.053

Doddapaneni, T. R. K. C., and Kikas, T. (2023). "Advanced applications of torrefied biomass: A perspective view," *Energies* 16(4). https://doi.org/10.3390/en16041635

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., and Vapnik, V. (1996). "Support vector regression machines," *Advances in Neural Information Processing* Systems 9, 155-16.

Dutta, T., Kwon, E., Bhattacharya, S. S., Jeon, B. H., Deep, A., Uchimiya, M., and Kim, K. H. (2017). "Polycyclic aromatic hydrocarbons and volatile organic compounds in biochar and biochar-amended soil: A review," *Global Change Biology Bioenergy* 9(6), 990-1004. https://doi.org/10.1111/gcbb.12363

FAO. (2024). "The State of the World's Forests 2024," Food and Agriculture Organization of the United Nations, Rome.

Feng, S., Zhou, H., and Dong, H. (2019). "Using deep neural network with small dataset to predict material defects," *Materials & Design* 162, 300-310. https://doi.org/ 10.1016/j.matdes.2018.11.060

Fidler, J. C., and North, C. J. (1969). "Production of volatile organic compounds by apples," *Journal of the science of food and agriculture* 20(9), 521-526. https://doi.org/10.1002/jsfa.2740200903

Friedman, J. H. (2001). "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics* 29(5), 1189-1232. https://doi.org/10.1214/aos/1013203451

Ghattas, B., and Manzon, D. (2023). "Machine learning alternatives to response surface models," *Mathematics* 11(15), article 3406. https://doi.org/10.3390/math11153406

Ha, S. Y., Kim, H. C., Lim, W. S., and Yang, J. (2025). "Heated wood-based ethylene scavenger for active packaging to prevent browning of *Lentinula edodes*," *BioResources* 20(3), 6286-6298. https://doi.org/10.15376/biores.20.3.6286-6298

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd Ed.), Springer, New York.

Jeong, H., Choi, J., Moon, J., Yoon, Y., and Cho, K. H. (2025). "Multimodal deep learning-based prediction of activated carbon adsorption capacities for volatile organic compound removal," *Journal of Cleaner Production* 519, article 145999.

https://doi.org/10.1016/j.jclepro.2025.145999
Jha, K., Doshi, A., Patel, P., and Shah, M. (2019). "A comprehensive review on automation in agriculture using artificial intelligence," *Artificial Intelligence in Agriculture* 2, 1-12. https://doi.org/10.1016/j.aiia.2019.05.004
Kader, A. A. (1985). "Ethylene-induced senescence and physiological disorders in harvested horticultural crops," *HortScience* 20(1), 54-57.
Keller, J. U., and Staudt, R. (2005). *Gas Adsorption Equilibria: Experimental Methods and Adsorptive Isotherms*, Springer, New York.
Kohavi, R. (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the 14th International Joint Conference on Artificial Intelligence* 41(2), 1137-1143.
Kvålseth, T. O. (1985). "Cautionary note about R^2," *American Statistician* 39(4), 279-285. https://doi.org/ 10.1080/00031305.1985.10479448
Li, H., Ai, Z., Yang, L., Zhang, W., Yang, Z., Peng, H., and Leng, L. (2023). "Machine learning assisted predicting and engineering specific surface area and total pore volume of biochar," *Bioresource Technology* 369, article 128417. https://doi.org/10.1016/j.biortech.2022.128417
Li, Y., Mei, B., Linhares-Juvenal, T., Formenton Cardoso, N., and Tshering, C. (2022). "Forest sector contribution to national economies 2015 – The direct, indirect and induced effects on value added, employment and labour income," *Forestry Working Paper* (33). Rome, FAO. https://doi.org/10.4060/cc2387en
Liaw, A., and Wiener, M. (2002). "Classification and regression by randomForest," *R News* 2(3), 18-22.
Liu, D., Kababji, S. E., Mitsakakis, N., Pilgram, L., Walters, T., Clemons, M., Pond, G., El-Hussuna, A., and Emam, K. E. (2025). "Synthetic data generation for augmenting small samples," *arXiv Preprint* arXiv:2501.18741. https://doi.org/10.48550/arXiv.2501.18741
Lu, K., Lee, W., Chen, W., Liu, S., and Lin, T. (2012). "Torrefaction and low temperature carbonization of oil palm fiber and eucalyptus in nitrogen and air atmospheres," *Bioresource Technology* 123, 98-105. https://doi.org/10.1016/j.biortech.2012.07.096
Manatura, K., Chalermsinsuwan, B., Kaewtrakulchai, N., Kwon, E. E., and Chen, W. H. (2023). "Machine learning and statistical analysis for biomass torrefaction: A review," *Bioresource Technology* 369, article 128504. https://doi.org/10.1016/j.biortech.2022.128504
Martínez-Romero, D., Bailén, G., Serrano, M., Guillén, F., Valverde, J. M., Zapata, P., Castillo, S., and Valero, D. (2007). "Tools to maintain postharvest fruit and vegetable quality through the inhibition of ethylene action: A review," *Critical Reviews in Food Science and Nutrition* 47(6), 543-560. https://doi.org/10.1080/10408390600846390
Meinshausen, N., and Ridgeway, G. (2006). "Quantile regression forests," *Journal of Machine Learning Research* 7(6), 983-999.
Mikołajczyk, A., and Grochowski, M. (2018). "Data augmentation for improving deep learning in image classification problem," *2018 International Interdisciplinary PhD Workshop (IIPhDW), IEEE*, pp. 117-122. https://doi.org/10.1109/IIPHDW.2018.8388338
Nielsen, D. (2016). *Tree Boosting with XGBoost: Why Does XGBoost Win 'Every' Machine Learning Competition?* Master's Thesis, Norwegian University of Science and Technology, Trondheim, Norway.
Niu, Y., Lv, Y., Lei, Y., Liu, S., Liang, Y., Wang, D., and Hui, S. E. (2019). "Biomass torrefaction: properties, applications, challenges, and economy," *Renewable and*

Sustainable Energy Reviews 115, article 109395. https://doi.org/10.1016/j.rser.2019.109395

Onsree, T., Tippayawong, N., Phithakkitnukoon, S., and Lauterbach, J. (2022). "Interpretable machine-learning model with a collaborative game approach to predict yields and higher heating value of torrefied biomass," *Energy* 249, article 123676. https://doi.org/10.1016/j.energy.2022.123676

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., *et al*. (2011). "Scikit-learn: Machine learning in Python," *The Journal of Machine Learning Research* 12, 2825-2830

Pilkington, J. L., Preston, C., and Gomes, R. L. (2014). "Comparison of response surface methodology (RSM) and artificial neural networks (ANN) towards efficient extraction of artemisinin from *Artemisia annua*," *Industrial Crops and Products* 58, 15-24. https://doi.org/10.1016/j.indcrop.2014.03.016

Rajabi, H., Mosleh, M. H., Prakoso, T., Ghaemi, N., Mandal, P., Lea-Langton, A., and Sedighi, M. (2021). "Competitive adsorption of multicomponent volatile organic compounds on biochar," *Chemosphere* 283, article 131288. https://doi.org/10.1016/j.chemosphere.2021.131288

Rouquerol, J., Rouquerol, F., Llewellyn, P., Maurin, G. and Sing, K. S. (2013). *Adsorption by Powders and Porous Solids: Principles, Methodology and Applications* (2nd Ed.), Academic Press, Oxford, UK.

Saltveit, M. E. (1999). "Effect of ethylene on quality of fresh fruits and vegetables," *Postharvest Biology and Technology* 15(3), 279-292. https://doi.org/10.1016/S0925-5214(98)00091-X

Schudel, S., Shoji, K., Shrivastava, C., Onwude, D., and Defraeye, T. (2023). "Solution roadmap to reduce food loss along your postharvest supply chain from farm to retail," *Food Packaging and Shelf Life* 36, article 101057. https://doi.org/10.1016/j.fpsl.2023.101057

Selivanov, E., Cudlin, P., and Horáček, P. (2023). "Carbon neutrality of forest biomass for bioenergy: a scoping review," *iForest - Biogeosciences and Forestry* 16(2), 70-77. https://doi.org/10.3832/ifor4160-015

Shorten, C., and Khoshgoftaar, T. M. (2019). "A survey on image data augmentation for deep learning," *Journal of Big Data* 6(1), article 60. https://doi.org/10.1186/s40537-019-0197-0

Simonic, M., Goricanec, D., and Urbancl, D. (2020). "Impact of torrefaction on biomass properties depending on temperature and operation time," *Science of the Total Environment* 740, article 140086. https://doi.org/10.1016/j.scitotenv.2020.140086

Smola, A. J., and Schölkopf, B. (2004). "A tutorial on support vector regression," *Statistics and Computing* 14(3), 199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Torlay, L., Perrone-Bertolotti, M., Thomas, E., and Baciu, M. (2017). "Machine learning–XGBoost analysis of language networks to classify patients with epilepsy," *Brain Informatics* 4(3), 159-169. https://doi.org/10.1007/s40708-017-0065-7

Tumuluru, J. S., Kremer, T., Wright, C. T., and Boardman, R. D. (2021). "Biomass torrefaction process, product properties, reactor types, and moving bed reactor design concepts," *Frontiers in Energy Research* 9, article 728140. https://doi.org/10.3389/fenrg.2021.728140

Vapnik, V. (2013). *The Nature of Statistical Learning Theory*, Springer Science & Business Media, New York.

Willmott, C. J., and Matsuura, K. (2005). "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research* 30(1), 79-82. https://doi.org/10.3354/cr030079

Woo, H., and Turner, P. (2019). "A review of recent research on carbon neutrality in forest bioenergy feedstocks," *International Journal of Environmental Sciences & Natural Resources* 19(2), article 556014. https://doi.org/10.19080/IJESNR.2019.19.556014

Xu, J., Zhu, L., Cai, W., Ding, Z., Chen, D., Zhang, W., Xing, C., Wang, K., and Ma, Z. (2024). "Upgrading the quality of biomass by advanced oxidative torrefaction pretreatment: Rebuilding the oxidative torrefaction mechanism based on hemicellulose, cellulose, and lignin," *Chemical Engineering Journal* 494, article 153044. https://doi.org/10.1016/j.cej.2024.153044

Yang, R. T. (2003). *Adsorbents: Fundamentals and Applications*, John Wiley & Sons, Hoboken, NJ.

Zhang, W., Chen, Q., Geng, J., Zhao, Q., Tian, J., Li, Y., Liu, S., Yang, Z., Qu, W., Yang, J., Li, L., Leng, L., and Li, H. (2025). "Predicting adsorption of volatile organic compounds onto biochars with machine learning and potential applications," *ACS ES&T Engineering* 5(12), 3326-3336. https://doi.org/10.1021/acsestengg.5c00547

Zhang, W., Chen, R., Li, J., Huang, T., Wu, B., Ma, J., Wen, Q., Tan, J., and Huang, W. (2023). "Synthesis optimization and adsorption modeling of biochar for pollutant removal via machine learning," *Biochar* 5(1), article 25. https://doi.org/10.1007/s42773-023-00225-x

Zhao, C., Yue, W., Xia, Q., Yang, H., Chen, A., and Liu, X. (2025). "Prediction of CO_2 adsorption performance of biochar based on machine learning," *Biomass Bioenergy* 201, article 108129. https://doi.org/10.1016/j.biombioe.2025.108129

Zhou, H., Xie, Q., Gu, S., Li, X., Zhou, Y., Liang, D., and Liu, J. (2025). "A feature analysis-based physical experience-driven neural network for predicting VOCs adsorption performance of activated carbon," *Carbon* 241, article 120387. https://doi.org/10.1016/j.carbon.2025.120387

Article submitted: September 7, 2025; Peer review completed: February 7, 2026; Revised version received: March 4, 2026; Further revised version received and accepted: March 18, 2026; Published: April 3, 2026.
DOI: 10.15376/biores.21.2.4538-4561