Journal of Science and Technology – Engineering and Technology for Sustainable Development
Volume 36, Issue 5, November 2026, 060–067

Robust Multi-Objective Optimization of Peanut Shell Biochar Production under Uncertainty Using Gaussian Process and Non-dominated Sorting Genetic Algorithm II

Phuc Vu Hong, Dat Nguyen Hoang, Do Quang Huy, Huy Nguyen Quang, Duy Nguyen Duc, Thieu Trinh Viet *

School of Mechanical Engineering, Hanoi University of Science and Technology, Ha Noi, Vietnam
**Corresponding author email: thieu.trinhviet@hust.edu.vn*

Abstract
As fossil fuels continue to dominate the global energy mix, the thermochemical conversion of biomass via pyrolysis offers a carbon-neutral pathway to produce high-value biochar. However, optimizing this highly nonlinear process remains challenging, especially when constrained by limited experimental data. This study investigates the pyrolysis of peanut shell-an abundant agricultural residue in ASEAN countries-aiming to simultaneously maximize biochar yield and fixed carbon content. To overcome the small dataset limitation (N = 9), a Gaussian Process Regression (GPR) model was developed. Incorporating Leave-One-Out Cross-Validation (LOOCV) and hyperparameter tuning, the GPR approach effectively mitigated overfitting while providing intrinsic uncertainty quantification-a distinct advantage over conventional machine learning methods. The surrogate model demonstrated high predictive accuracy (R^2 equal 0.9975 for yield and 0.9341 for fixed carbon). By coupling GPR with the NSGA-II algorithm, a robust multi-objective optimization framework was established. The generated Pareto front identified optimal operating conditions that secured high fixed carbon levels (e.g., above 75 wt%) while maintaining total predictive uncertainty below 2.5%. Ultimately, this data-efficient framework provides a highly reliable strategy for optimizing biomass pyrolysis and scaling sustainable bioenergy systems with minimal experimental overhead.

Keywords: Biomass pyrolysis; biochar yield; Gaussian process regression; NSGA-II; uncertainty quantification.

Impress

1. Introduction

As global energy demand increases, energy security and climate change remain major challenges of the 21st century. Despite strong commitments to achieving Net Zero emissions, fossil fuels continue to dominate primary energy consumption. According to the International Energy Agency (IEA), global coal consumption remains at 8.85 billion tons [1]. Consequently, biomass has attracted attention as a promising alternative due to its carbon-neutral potential and renewable nature.

To utilize this resource efficiently, pyrolysis is widely employed as a thermochemical method to convert biomass into valuable bioenergy products, particularly biochar. Biochar serves as an effective material for adsorption, carbon capture, and soil amendment. However, its physicochemical properties strongly depend on operating parameters, including temperature, heating rate, and residence time. The relationship between these variables and the final product characteristics is highly nonlinear, complicating the process optimization. In this study, peanut shell was selected as the feedstock due to its abundance as an underutilized agricultural residue in Vietnam and the broader ASEAN region. Annually, Vietnam alone generates approximately 94.7 million tons of agricultural residues, with open burning practices accounting for about 43% of global biomass burning [2, 3].

A major challenge in optimizing biomass pyrolysis is balancing prediction accuracy and experimental cost. Traditional optimization methods require extensive experimental runs, which are time-consuming and expensive. Consequently, pyrolysis studies frequently rely on small datasets. Previous research has primarily applied Design of Experiments (DoE) methods, such as Taguchi or Box–Behnken designs, typically involving fewer than 20 experimental runs. For instance, Vinh [4] examined the effect of pyrolysis temperature on biochar properties within a narrow range due to the extended duration required for each experiment. Gorshkov [5] tested seven different samples but restricted the temperature range to 400–600 °C. The limited availability of experimental data restricts the generalization and reliability of these findings.

To address the limitations of small datasets, machine learning (ML) algorithms have been proposed as alternatives to conventional approaches. Li [6] applied artificial neural networks (ANN) combined with adaptive neuro-fuzzy inference system (ANFIS) to predict fixed carbon (FC) content and biochar yield. Their model, trained on 226 samples, achieved high predictive accuracy, yielding R^2 values of 0.964 for biochar yield and 0.918 for FC. Liu [7] compared four

p-ISSN 3093-3242
e-ISSN 3093-3579
https://doi.org/10.51316/jst.194.etsd.2026.36.5.8
Received: Mar 23, 2026; Revised: Apr 24, 2026;
Accepted: May 9, 2026; Online: Jul 28, 2026

models – Support Vector Regression (SVR), Fuzzy Regression (FR), Backpropagation Artificial Neural Network (BP-ANN), and XGBoost–using 271 data points, identifying XGBoost as the most robust approach. They also demonstrated that data augmentation techniques improved model performance. Yapıcı [8] reported that GPR outperformed SVR in predicting biogas yield from LDPE waste pyrolysis. However, a significant gap remains conventional ML and ANN models generally require large datasets for reliable training. When applied to small datasets, these models are prone to overfitting and struggle to perform well in optimization tasks. Furthermore, prior studies rarely quantify prediction uncertainty, limiting the applicability of the results in practical scenarios.

To overcome these constraints, this study presents a methodology designed specifically for small datasets (comprising nine experimental samples) using peanut shell feedstock. GPR was utilized to construct the surrogate model and facilitate multi-objective optimization via the NSGA-II algorithm. Leave-One-Out Cross-Validation (LOOCV) and grid search were implemented to minimize overfitting. Unlike standard ML methods, GPR is well-suited for limited data and provides predictive uncertainty estimation. Following the generation of the response surface, NSGA-II was applied to identify the optimal operating region-represented by the Pareto front-that simultaneously evaluates biochar yield, FC content, and their respective uncertainties.

2. Methodology

2.1. Fuel Preparation

Peanut shells sourced from Vietnam were utilized as the feedstock in this study. The raw material was washed with distilled water to remove impurities and dried at 105 °C for 2 hours. Subsequently, the dried biomass was sieved to isolate particles ranging from 200 to 400 µm in size. Prior to each experiment, the samples were re-dried under identical conditions to eliminate residual moisture. All drying processes were performed in a laboratory drying cabinet (Thermo-plus GMP500) with a temperature control accuracy of ±0.5%.

2.2. Experiment Conditions

The pyrolysis experiments were conducted in an electric furnace equipped with a SUS304 stainless steel tubular reactor (60 mm inner diameter), as illustrated in Fig. 1. A *K*-type thermocouple with an accuracy of 1.5% was employed to monitor and control the furnace temperature and heating rate. The pyrolysis process was performed under an N2 atmosphere. The operating temperature was varied from 400 to 600 °C in increments of 100 °C, and the heating rate was set between 5 and 15 °C/min in increments of 5 °C/min. The residence time was maintained at 75 minutes for all experimental runs. Upon completion of the pyrolysis phase, the resulting biochar was cooled to room temperature under continuous N2 flow and subsequently stored in sealed desiccators.

The biochar yield was calculated using the following equation:

$$Biochar\ yield\ (\%) = \frac{m_{biochar}}{m_{initial}} \times 100 \quad (1)$$

where $m_{biochar}$ is the mass of the biochar product after pyrolysis and $m_{initial}$ is the initial mass of the dried sample.

The biochar yield was determined using an analytical scale (model Sartorius Quintix 224-1S) with a precision of ±0.1 mg. The FC content of the biochar was determined according to the ASTM D5142 standard using a standard laboratory furnace (model Nabertherm L 9/11/B180 (L-090K1CN)) with a temperature accuracy of ±1.5 °C.

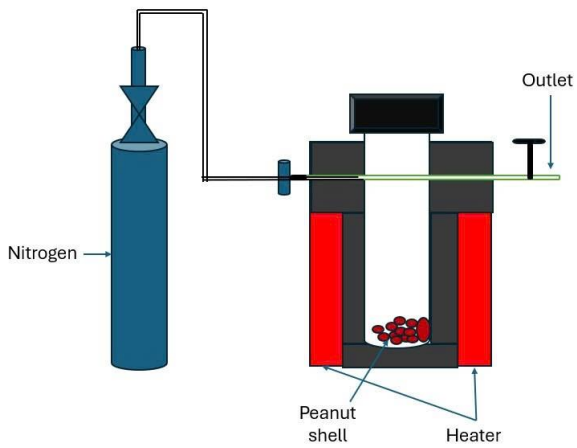

Fig. 1. Schematic of the pyrolysis apparatus

2.3. Gaussian Process

Gaussian Process (GP) is a probabilistic regression model in which a latent function *f(x)* is assumed to follow a Gaussian Process prior, characterized by a mean function and a covariance function [9]. The GP model is mathematically formulated as:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad (2)$$

where *m(x)* is the mean function and *k(x, x')* is the covariance function (kernel). For simplicity and without loss of generality, the mean function is conventionally assumed to be zero.

Given a training dataset comprising *N* samples $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$, the observed output is expressed as:

$$y_i = f(x_i) + \varepsilon_i \quad (3)$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma_n^2)$ represents independent and identically distributed Gaussian noise with variance σ_n^2.

The covariance matrix K of the training inputs is constructed as:

$$K = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \cdots & k(x_N, x_N) \end{bmatrix} \quad (4)$$

Incorporating the observation noise, the covariance matrix of the targets becomes:

$$\mathbf{K} + \sigma_n^2 \mathbf{I} \quad (5)$$

For a new test point x_*, the posterior predictive distribution of y_* follows a Gaussian distribution with the predictive mean and variance given by:

$$\mu_* = \mathbf{k}_*^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}$$
$$\sigma_*^2 = k(x_*, x_*) - \mathbf{k}_*^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_* \quad (6)$$

where $\mathbf{k}_* = [k(x_1, x_*), \dots, k(x_N, x_*)]^\top$. In these equations, $\mathbf{K}$ denotes the kernel matrix of the training data, σ_n^2 is the noise variance, and $\mathbf{I}$ represents the identity matrix.

The kernel function governs the correlation structure of the data. In this study, the Radial Basis Function (RBF) kernel was employed, expressed as:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{d=1}^{D} \frac{(x_d - x'_d)^2}{l_d^2}\right) \quad (7)$$

where σ_f^2 represents the signal variance and l_d denotes the characteristic length scale of the d-th input variable.

The optimal hyperparameters of the GP model are determined by maximizing the log marginal likelihood (LML) of the training data:

$$\log p(\mathbf{y} \mid \mathbf{X}) = -\frac{1}{2} \mathbf{y}^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} - \frac{1}{2} \log|\mathbf{K} + \sigma_n^2 \mathbf{I}| - \frac{N}{2} \log(2\pi) \quad (8)$$

Upon completion of the training phase, the GP model simultaneously provides the predictive mean and variance, offering a comprehensive probabilistic description of the output distribution.

2.4. Non-Dominated Sorting Genetic Algorithm

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) is a prominent evolutionary algorithm employed for multi-objective optimization [10]. Its primary objective is to identify the Pareto-optimal set within the design space. A solution is mathematically defined as Pareto optimal when no single objective can be improved without a simultaneous degradation of at least one other objective.

A general multi-objective optimization problem is formulated as:

$$\min_{\mathbf{x} \in \Omega} \mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_M(\mathbf{x})] \quad (9)$$

where $\mathbf{x}$ represents the design variable vector, Ω denotes the feasible domain, and $f_i(\mathbf{x})$ are the individual objective functions.

The core mechanism of NSGA-II relies on non-dominated sorting, partitioning the population into distinct Pareto fronts based on dominance relationships. Solutions occupying the first Pareto front are strictly non-dominated and are assigned the highest priority. To preserve solution diversity within each front, NSGA-II incorporates a crowding distance metric, which evaluates the relative density of solutions within the objective space. This mechanism mitigates premature convergence and ensures a uniformly distributed Pareto front.

Throughout the evolutionary process, NSGA-II implements an elitism strategy by merging parent and offspring populations. The superior individuals are subsequently selected based on their non-dominated rank and crowding distance to propagate to the next generation. This approach guarantees the retention of high-quality Pareto solutions throughout the optimization process. As NSGA-II operates independently of predefined weighting factors and simultaneously identifies multiple optimal trade-off solutions, it is highly effective for evaluating trade-offs among conflicting objectives.

2.5. Evaluation Metrics

To evaluate the model's predictive performance, five statistical metrics were employed: the coefficient of determination (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Let y_i and $\hat{y}_i$ denote the experimental and predicted values of the i-th sample, $\bar{y}$ the mean of the experimental values, and n the total number of samples. The metrics are mathematically defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (10)$$

R^2 measures the proportion of variance in the experimental data explained by the model.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (11)$$

MSE evaluates the average squared difference, inherently penalizing large deviations [11].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (12)$$

RMSE represents the standard deviation of prediction errors, sharing the same unit as the output variable to facilitate interpretability [12].

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (13)$$

MAE computes the average magnitude of absolute errors, providing a robust evaluation that is less sensitive to outliers.

Inpress

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| \quad (14)$$

MAPE expresses the error as a scale-independent percentage [13], facilitating comparison across datasets; however, it exhibits numerical instability when actual values (y_i) approach zero. To ensure full reproducibility, the specific hyperparameters used for the NSGA-II algorithm were set as follows: a population size of 100, 200 generations, a crossover rate of 0.9, and a mutation rate of 0.5.

3. Results and Discussion

3.1. Effect of Temperature and Heating Rate on Biochar Yield and Fixed Carbon Content

The pyrolysis of peanut shells was evaluated between 400 and 600 °C at heating rates of 5, 10, and 15 °C/min. In this study, the residence time was maintained at a constant 75 minutes across all experiments to isolate thermal effects. The results indicate that temperature is the dominant factor, inversely affecting biochar yield while increasing the FC content. Furthermore, the heating rate exhibits a non-monotonic relationship, with 10 °C/min optimally promoting secondary charring.

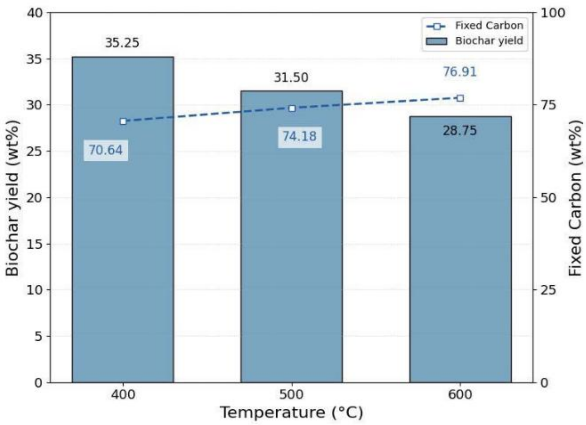

Fig. 2. Effect of pyrolysis temperature on biochar yield and fixed carbon content at 5 °C/min

Fig. 2 to Fig. 7 present the resulting biochar yields and FC contents. The observed parameter dependencies align with established kinetic degradation models for lignocellulosic biomass [14-16]. Temperature inversely affected biochar yield while proportionally enhancing FC content across all heating regimes. For example, at a 5 °C/min heating rate, elevating the temperature from 400 to 600 °C reduced the biochar yield from 35.25% to 28.75% (an approximate 6.5% decrease). This mass loss primarily stems from the thermal degradation of the lignocellulosic matrix-specifically cellulose and hemicellulose. Between 300 and 450 °C, the cleavage of glycosidic bonds and the decomposition of unstable oxygenated functional groups (such as carboxyl and carbonyl) drive the intense devolatilization of CO, CO_2, and water vapor. This mechanism corroborates findings by Gorshkov [3], who reported a comparable approximately 6.2% mass reduction for nutshell biomass within a similar thermal range. Beyond 500 °C, the reaction dynamics shift toward carbonization. The more thermally resilient lignin framework undergoes gradual decomposition while devolatilization slows. This environment favors secondary condensation and aromatization reactions. Consequently, the FC content peaked at 78.67% under the most severe operational conditions (600 °C, 15 °C/min), coupled with a corresponding drop in volatile matter from approximately 23% to 14% at 5 °C/min. This progressive increase in FC reflects improved structural stability and enhanced carbon retention within the biochar matrix.

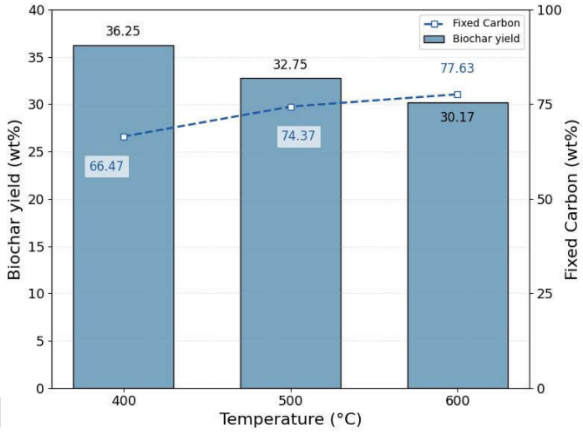

Fig. 3. Effect of pyrolysis temperature on biochar yield and fixed carbon content at 10 °C/min

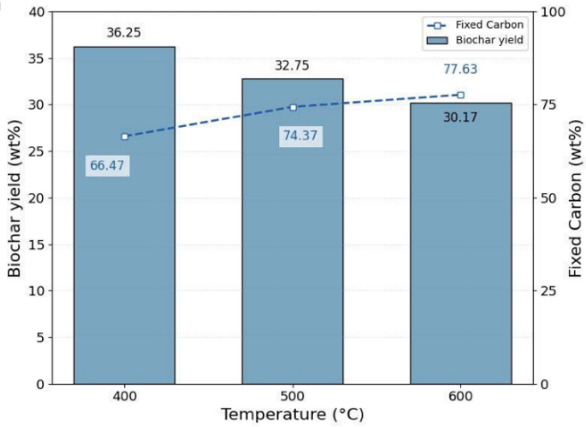

Fig. 4. Effect of pyrolysis temperature on biochar yield and fixed carbon content at 15 °C/min

While slower heating rates are theoretically assumed to maximize char yield by minimizing internal thermal gradients, the experimental data revealed a non-monotonic relationship. The highest biochar yields were consistently recorded at a moderate heating rate of 10 °C/min. At 500 °C, for instance, the 10 °C/min rate

yielded 32.75% biochar, outperforming both 5 °C/min (31.50%) and 15 °C/min (29.25%). This local maximum indicates the promotion of secondary charring. As noted by Zhao *et al.* [17], an optimal heating rate provides volatile compounds with sufficient residence time within the pore network to undergo secondary cracking and repolymerization. This process amplifies the final solid mass without causing structural collapse from thermal shock. Conversely, increasing the heating rate to 15 °C/min suppressed the char yield by 2–4%. Rapid heating generates steep temperature gradients between the particle's surface and core, inducing internal thermal stress and fragmenting the pore structure. Under these conditions, volatiles are expelled too rapidly to participate in secondary recombination pathways, leading to a net reduction in yield.

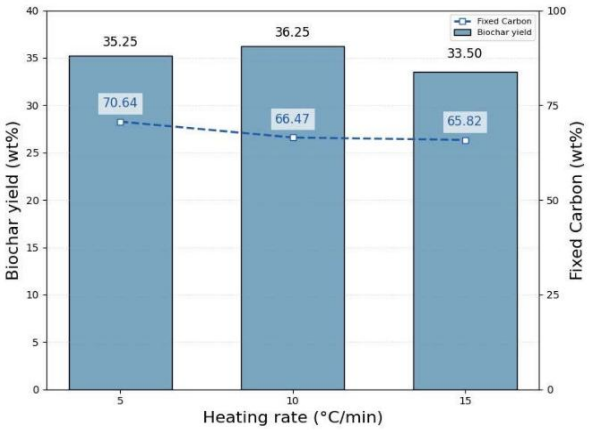

Fig. 5. Effect of pyrolysis heating rate on biochar yield and fixed carbon content at 400 °C

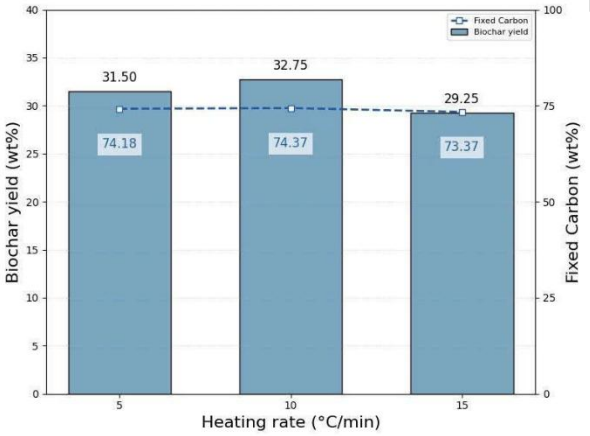

Fig. 6. Effect of pyrolysis heating rate on biochar yield and fixed carbon content at 500 °C

Selecting optimal pyrolysis parameters ultimately requires balancing biochar quantity (yield) and quality (FC content). Experimental evidence suggests that operating between 500 and 600 °C at a heating rate of 10 °C/min provides the most favorable compromise. This specific regime effectively exploits secondary charring mechanisms to mitigate yield loss while ensuring a high degree of carbonization.

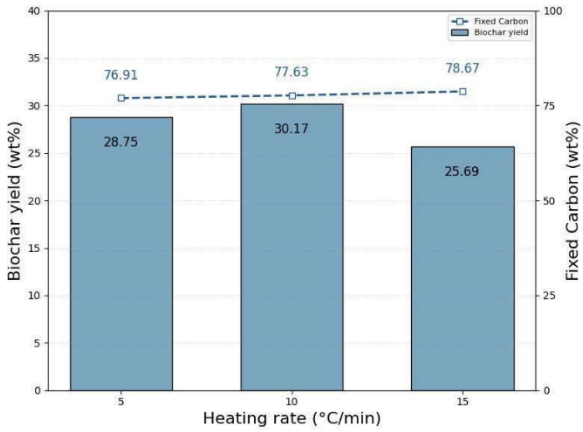

Fig. 7. Effect of pyrolysis heating rate on biochar yield and fixed carbon content at 600 °C

3.2. Gaussian Process Model Performance

To address the limited experimental dataset (N=9), GPR models were developed to predict BC and FC content. To prevent overfitting, an RBF kernel was selected, as its continuous and smooth properties are well-suited for modeling nonlinear thermochemical phenomena like biomass pyrolysis. This kernel effectively isolates signal from noise, capturing the intrinsic process variations while accounting for experimental measurement errors. Hyperparameters were optimized via grid search. Due to the small sample size, model generalization was validated using LOOCV, iteratively utilizing eight samples for training and one for testing. Unlike standard machine learning approaches such as ANN or SVR that tend to overfit small datasets, this GPR framework provides inherent uncertainty quantification, demonstrating higher reliability over deterministic models.

Table 1. LOOCV performance of the optimized GPR model for biochar yield and fixed carbon content

Metrics	Average BC Yield Training set	BC Yield Testing set	Average FC Training set	FC Testing set
R2	1.0000	0.9975	0.9905	0.9341
RMSE	0.0008	0.1597	0.4101	1.1223
MAE	0.0007	0.1410	0.3088	0.9358
MAPE (%)	0.0021	0.4718	0.4337	1.1665

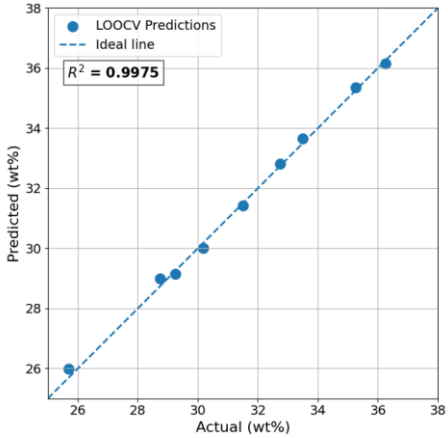

Fig. 8. LOOCV validation results of the GPR model for biochar yield prediction

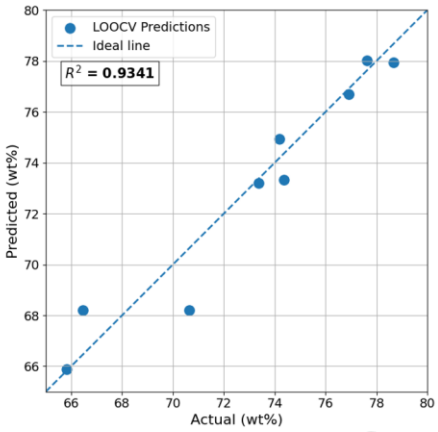

Fig. 9. LOOCV validation results of the GPR model for fixed carbon prediction

Table 1 and Fig. 8 to Fig. 9 summarize the LOOCV performance. The BC model demonstrated high precision, achieving a testing R^2 of 0.9975 and a MAPE of 0.47%. The perfect training fit (R^2 = 1.0000) paired with minimal testing error confirms the model successfully avoided severe overfitting. The FC model yielded a slightly lower, yet highly reliable testing R^2 of 0.9341 and a MAPE of 1.17%. The moderate variance between the FC training (R^2= 0.9905) and testing metrics indicates expected prediction error on unseen data, remaining well within acceptable thresholds for thermochemical modeling. A critical advantage of the GP approach is its ability to quantify predictive uncertainty. In this study, the total uncertainty was formulated by integrating the intrinsic GP predictive variance (representing spatial data distribution) with the external error variance derived from LOOCV (representing limited data and measurement noise):

$$\sigma_{total}^2 = \sigma_{GP}^2 + \sigma_{LOO}^2 \quad (15)$$

where σ_{GP} is the GP standard deviation and σ_{LOO} is the LOOCV error standard deviation. This composite framework ensures a mathematically robust surrogate model for subsequent multi-objective optimization.

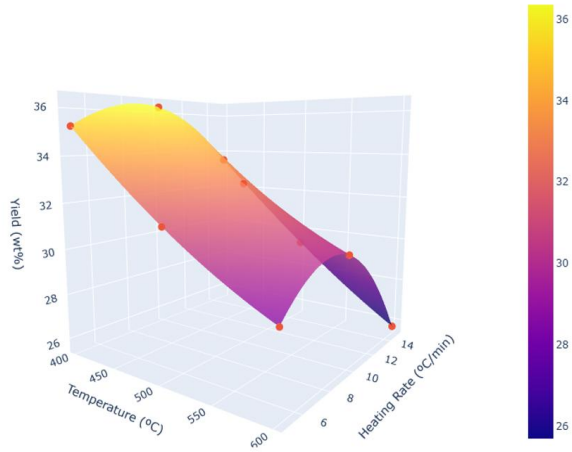

Fig. 10. Biochar yield response surface from GPR model

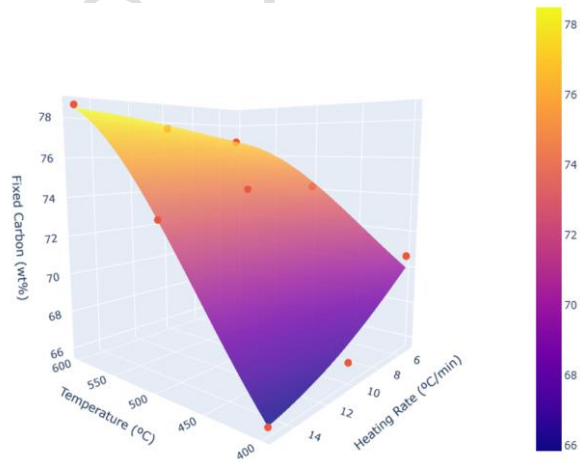

Fig. 11. Fixed carbon response surface from GPR model

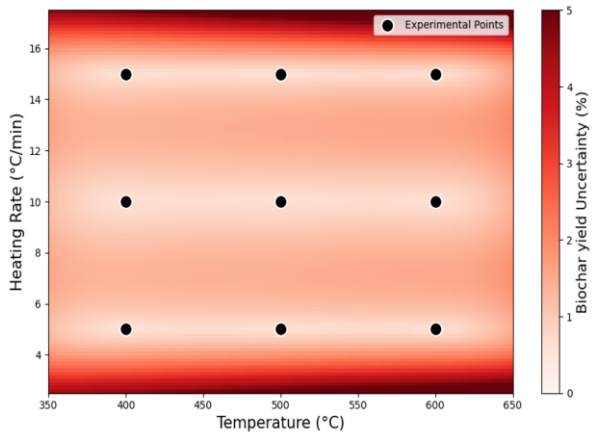

Fig. 12. Biochar yield uncertainty from GPR model

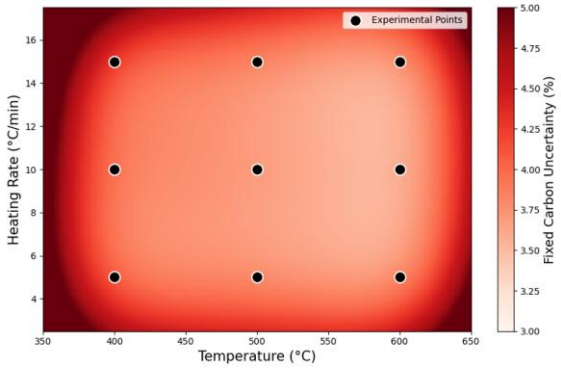

Fig. 13. Fixed carbon uncertainty from GPR model

3.3. Multi-Objective Optimization via NSGA-II

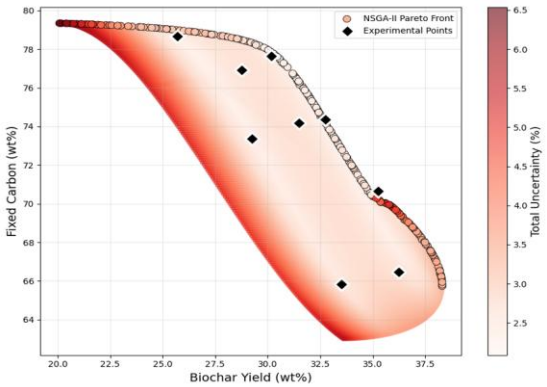

Fig. 14. Pareto front optimization using NSGA-II without uncertainty consideration

Fig. 14 presents the Pareto front generated by the deterministic NSGA-II algorithm, highlighting the inherent trade-off between biochar yield and FC content. The continuous and smooth frontier indicates stable algorithmic convergence and a comprehensive exploration of the objective space. Specifically, maximizing FC to 79–80 wt% restricts the yield to approximately 20–25 wt%, whereas increasing the yield to 35–38 wt% inevitably reduces FC to 66–70 wt%. This inverse relationship confirms the conflicting nature of the two objectives, validating the necessity of a multi-objective optimization approach. Furthermore, the uniform distribution of solutions along the Pareto front demonstrates the efficacy of the crowding-distance mechanism in preserving population diversity, thereby providing a comprehensive set of trade-off options rather than a singular local optimum.

The integration of predictive uncertainty into the NSGA-II framework visibly modifies the Pareto topology (Fig. 15). Instead of converging strictly on the extreme deterministic boundaries, the optimal solutions shift slightly inward. This topological deviation occurs because the algorithm penalizes operating regions associated with higher model uncertainty, effectively balancing absolute objective performance with predictive reliability. In Fig. 15, the star markers denote the strategically selected optimal points that satisfy two strict criteria: an FC content exceeding 75 wt% and a total predictive uncertainty below 2.5%. By adhering to these thresholds, the identified parameters offer highly robust operating conditions. This uncertainty-aware approach ensures high-quality carbonization while mitigating the risks associated with data limitations, thereby facilitating reliable engineering decision-making for scale-up applications.

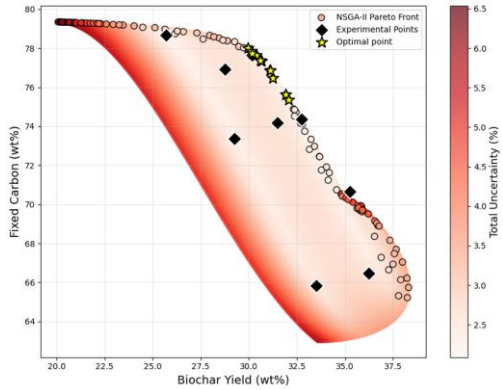

Fig. 15. Optimal point using NSGA-II with uncertainty consideration

4. Conclusion

This study proposed a data-efficient framework for modeling and optimizing peanut shell pyrolysis using a limited experimental dataset (N = 9). By employing Gaussian Process Regression (GPR) coupled with grid search and LOOCV, the methodology effectively mitigated overfitting. The resulting surrogate models demonstrated high predictive accuracy, achieving R^2 values of 0.9975 for biochar yield and 0.9341 for FC. A key advantage of this GPR approach, distinguishing it from conventional machine learning methods, is its ability to quantify intrinsic predictive uncertainties. Integrating these probabilistic models with the NSGA-II algorithm generated an uncertainty-aware Pareto front. This advanced optimization strategy identified stable operating conditions–specifically, operating between 500 and 600 °C at a heating rate of 10 °C/min–that secure high FC content (> 75 wt%) while strictly constraining the total prediction uncertainty to below 2.5%. Future research should extend this framework to other biomass feedstocks and evaluate the scalability of the GPR-NSGA-II model for continuous, industrial-scale pyrolysis reactors.

Acknowledgments

This research is funded by Hanoi University of Science and Technology (HUST) under project number T2025-PC-035.

References

[1] International Energy Agency, Coal 2023: Analysis and Forecast to 2026, International Energy Agency, Paris, France, Dec. 2023.

[2] D. Beňová, K. Mareš, P. Hutla, T. Ivanova, J. Banout, and M. Kolaříková, Energy potential of agri residual biomass in Southeast Asia with the focus on Vietnam, Agronomy, vol. 11, no. 1, Jan. 2021, Art. no. 169. https://doi.org/10.3390/agronomy11010169

[3] N. H. Son, B. T. P. Loan, and N. D. Minh, The current status of agricultural wastes and residuals management and recycling in Vietnam, FFTC Agricultural Policy Platform, 2021.

[4] N. T. Vinh, L. N. Han, H. P. N. Thao, and H. T. T. Hien, Trial production of biochar pellets from de-oiled cashew nut shells, TNU Journal of Science and Technology, vol. 230, no. 2, pp. 38–46, 2025. https://doi.org/10.34238/tnu-jst.11673

[5] A. Gorshkov, N. Berezikov, A. Kaltaev, S. Yankovsky, K. Slyusarsky, R. Tabakaev, and K. Larionov, Analysis of the physicochemical characteristics of biochar obtained by slow pyrolysis of nut shells in a nitrogen atmosphere, Energies, vol. 14, no. 23, Dec. 2021, Art. no. 8075. https://doi.org/10.3390/en14238075

[6] Y. Li, R. Gupta, and S. You, Machine learning assisted prediction of biochar yield and composition via pyrolysis of biomass, Bioresource Technology, vol. 359, Sep. 2022, Art. no. 127511. https://doi.org/10.1016/j.biortech.2022.127511

[7] B. Liu, G. Zhang, S. Hong, and Z. Zhang, ANN assisted optimization of phosgene free hexamethylene diisocyanate production for economic and environmental benefits, Scientific Reports, vol. 15, Sep. 2025, Art. no. 33973. https://doi.org/10.1038/s41598-025-11730-3

[8] E. Yapıcı, H. Akgün, K. Özkan, Z. Günkaya, A. Özkan, and M. Banar, Prediction of gas product yield from packaging waste pyrolysis: Support vector and Gaussian process regression models, International Journal of Environmental Science and Technology, vol. 20, pp. 461–476, Feb. 2022. https://doi.org/10.1007/s13762-022-04013-1

[9] J. Wei, Y. Sun, D. Zhu, Q. Liu, X. Wang, and Z. Liu, PAL-GP: A parallel active learning framework based on Gaussian processes for efficient fatigue reliability assessment of structures, Structures, vol. 86, Apr. 2026, Art. no. 111294. https://doi.org/10.1016/j.istruc.2026.111294

[10] W. Bi, Y. Wu, Y. Xu, and A. Zhang, Optimized deployment–monitoring synchronized array method for passive sonobuoys based on an improved NSGA-II algorithm, Applied Soft Computing, vol. 195, Jun. 2026, Art. no. 115065. https://doi.org/10.1016/j.asoc.2026.115065

[11] K. Phuakpunk, B. Chalermsinsuwan, and S. Assabumrungrat, Comparative artificial neural network models for predicting kinetic parameters of biomass pyrolysis from the biomass characteristics, Results in Engineering, vol. 27, Sep. 2025, Art. no. 106701. https://doi.org/10.1016/j.rineng.2025.106701

[12] J. Xing, K. Luo, H. Wang, Z. Gao, and J. Fan, A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches, Energy, vol. 188, Dec. 2019, Art. no. 116077. https://doi.org/10.1016/j.energy.2019.116077

[13] P. J. García-Nieto, E. García-Gonzalo, J. P. Paredes-Sánchez, and L. A. Menéndez-García, Interpretable machine learning models for forecasting elemental composition of solid biomass fuels from proximate analyses in energy processes, Journal of the Energy Institute, vol. 123, Dec. 2025, Art. no. 102322. https://doi.org/10.1016/j.joei.2025.102322

[14] P. Talwar, M. A. Agudelo, and S. Nanda, Pyrolysis process, reactors, products, and applications: A review, Energies, vol. 18, no. 11, Jun. 2025, Art. no. 2979. https://doi.org/10.3390/en18112979

[15] S. N. Mohan, N. N. Nair, P. V. S. Nair, K. V. Babu, and K. M. George, Biochar as an exceptional bioresource for energy, agronomy, carbon sequestration, activated carbon and specialty materials, Waste Biomass Valorization, vol. 7, no. 2, pp. 201–235, Apr. 2016. https://doi.org/10.1007/s12649-015-9459-z

[16] A. K. Nanda, K. K. Berruti, A. Dalai, J. A. Kozinski, and S. C. Cameron, Advanced techniques in the production of biochar from lignocellulosic biomass and environmental applications, Biofuel Research Journal, vol. 3, no. 1, pp. 363–372, Mar. 2016. https://doi.org/10.1016/j.clema.2022.100137

[17] B. Zhao, D. O'Connor, J. Zhang, T. Peng, Z. Shen, D. C.W. Tsang, and D. Hou, Effect of pyrolysis temperature, heating rate, and residence time on rapeseed stem derived biochar, Journal of Cleaner Production, vol. 174, pp. 977–988, Feb. 2018. https://doi.org/10.1016/j.jclepro.2017.11.013

In press