Mathematical Modeling and Algorithm Application
ISSN: 3006-0842 | Vol. 8, No. 3, 2026

Mechanism of n-Hexane Insolubles in Biomass-Coal Co-Pyrolysis and Optimization of Their Mixing Ratios

Xiaoyu Liu*

College of Materials and Chemical Engineering, Hezhou University, Hezhou 542899, China
* **Corresponding author**: Xiaoyu Liu (Email: lxy2397730704@163.com)

Abstract: This study thoroughly investigates the influence mechanism of n-hexane insolubles (INS) and their blending ratio on product yields during the co-pyrolysis of biomass and coal. First, Spearman's correlation coefficient analysis revealed intrinsic relationships between INS content and yields of tar, water, and char. Results indicate significant correlations between INS and tar/char yields, while its effect on water yield is negligible. Subsequently, to further elucidate yield variation patterns under complex operating conditions, variance analysis was employed to investigate the interactive effects of INS and blending ratios on product yields. Findings revealed that different INS-blend combinations significantly altered tar growth rates and coke residue decline trends. Building upon these findings, a nonlinear optimization model was constructed to maximize energy conversion efficiency and product utilization. The Sequential Quadratic Programming (SLSQP) algorithm was employed to identify optimal blending ratios and INS dosage under specified cost and technical constraints. The research outcomes provide scientific theoretical foundations and operational guidance for optimizing biomass energy utilization, enhancing energy efficiency, and achieving sustainable energy production.

Keywords: Spearman correlation analysis; Analysis of variance; SLSQP algorithm.

1. Introduction

With the rapid advancement of global industrialization and modernization, the conflict between growing energy demands and environmental protection has become increasingly prominent. Finding sustainable and environmentally friendly energy solutions has emerged as a global priority. The co-pyrolysis technology of biomass and coal has garnered significant attention due to its immense potential in integrating renewable energy with fossil fuels. Against this backdrop, accurately identifying the influence patterns of key components—such as insoluble n-hexane substances (INS)—on yield within reaction feedstocks, while addressing the issue of interaction interference among multiple variables, represents a critical challenge for enhancing energy utilization efficiency. Although previous studies have explored the fundamental distribution of pyrolysis yields, there remains room for improvement in deeply analyzing synergistic effects between components and seeking optimal ratios under complex constraints. The innovation of this section lies in not only using non-parametric Spearman's correlation coefficients for qualitative assessment of component influences but also quantifying interaction effects between blending ratios and specific components via ANOVA. Combined with the SLSQP algorithm, this approach achieves yield maximization optimization. The overall research methodology follows a logical sequence from correlation assessment and interaction effect analysis to multi-objective blending ratio optimization, aiming to enhance resource productivity in pyrolysis processes through systematic modeling[1-2].

Correlation Analysis of n-Hexane Insolubles on Pyrolysis Yields At the study's outset, quantitative values were transformed into ordered data using Spearman's model to calculate correlation coefficients between indicators and conduct significance tests. Analysis revealed a significant positive correlation between n-hexane insolubles and tar yield, along with a strong intrinsic link to coke residue yield[3]. However, water yield showed minimal direct influence from n-hexane insolubles content, exhibiting no statistical significance.

Interaction Effect Analysis Between n-Hexane Insolubles and Mixing Ratio Further investigation employed linear model equations to evaluate the combined effects of continuous and categorical variables, using analysis of variance (ANOVA) to determine statistical significance. Findings revealed that mixing ratio significantly influenced tar, water, and coke yield, while n-hexane insolubles primarily exerted significant effects on tar and coke yield. Interaction effect analysis diagrams clearly demonstrate that as INS content increases, the growth rates of tar yield under different blending ratios exhibit distinct differences. Specifically, high blending ratios combined with high INS content significantly enhance tar output while simultaneously accelerating the decline in coke residue yield[4-5].

Development of a Co-pyrolysis Ratio Strategy Based on Multi-Objective Optimization After clarifying the variable relationships, this study established an optimization model aimed at maximizing tar, water, and coke residue yields, setting the ranges of mixing ratios and INS content as constraints. The model accounts for technical constraints such as equipment capacity, temperature, and pressure, while also incorporating raw material costs as an economic constraint. The nonlinear constrained problem was solved iteratively using the SLSQP algorithm, ultimately determining the optimal blending ratio and INS concentration. This approach maximizes product utilization and energy conversion efficiency while satisfying industrial production limitations[6].

2. Model Establishment and Solution

2.1. Correlation Analysis of n-Hexane Insolubles on Pyrolysis Yield

2.1.1. Establishment of Spearman Model

Considering that the Spearman correlation coefficient is used to describe the correlation between two ordinal datasets, we first rank the scores of the indicators, convert the quantitative value sequences in this paper into ordinal data sequences, then calculate the Spearman correlation coefficients between the scores of each indicator, and test the significance of the correlation coefficients[7-8]. The specific calculation formula for the Spearman correlation coefficient between any two datasets is as follows:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{1}$$

where d_i represents the rank difference between the scores of the two indicators. The variation range of r_s is [-1, 1]. A negative value indicates a negative correlation, and a positive value indicates a positive correlation. The closer the absolute value is to 1, the stronger the correlation between the two evaluation indicators. In this paper, an absolute value of r_s greater than 0.6 is considered a high correlation, less than 0.4 a low correlation, and between 0.4 and 0.6 a moderate correlation[9-10].

2.1.2. Solution of Spearman Model

Correspondence of various indicators:

The calculation results of Spearman coefficients between the scores of various indicators are shown in Figure 1 below:

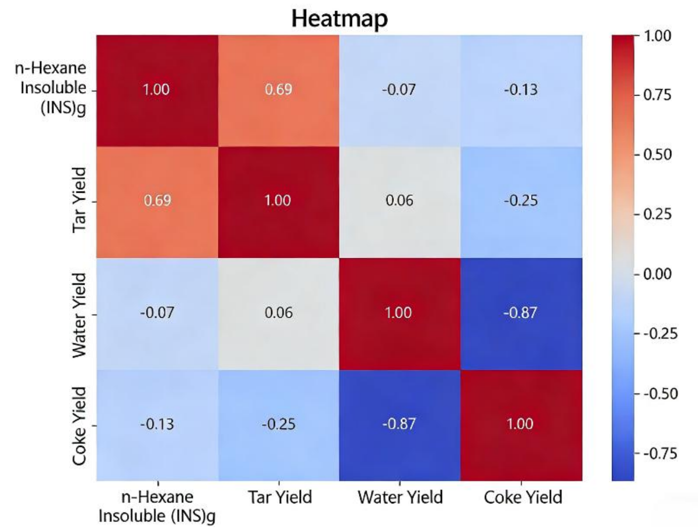

Figure 1 Heat Map of Spearman Coefficients

Description of the box plot of yields (tar yield, water yield, char yield): The box plot in Figure 2 below shows the distribution of three different yields. The box plot displays the median, quartiles, and outliers of the data.

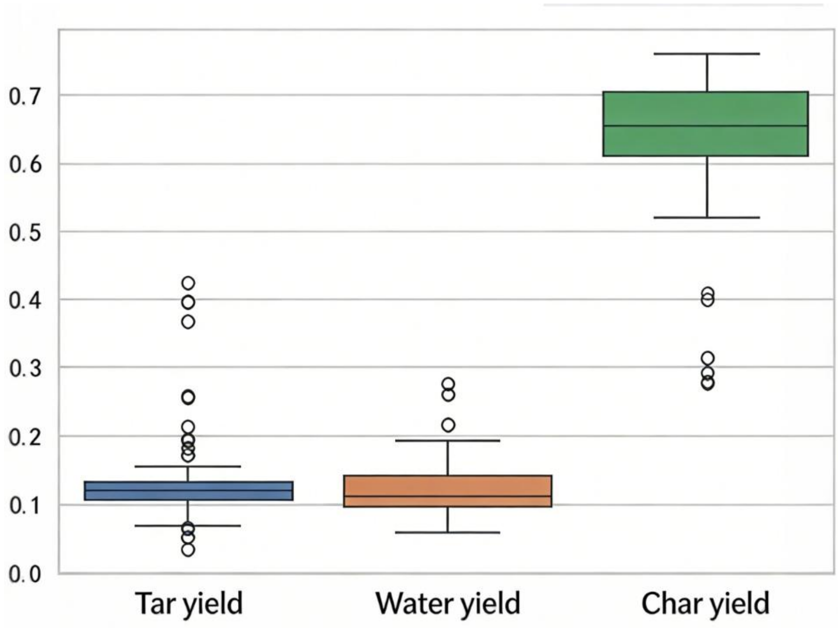

Figure 2 Box Plot of Yields

Tar yield: The median of tar yield is relatively low, and the data distribution is relatively concentrated.

Water yield: The median of water yield is the highest, and the data distribution is wide, indicating large fluctuations.

Char yield: The distribution of char yield is between that of tar and water, with few outliers.

2.1.3. Results

In summary, n-hexane insolubles (INS) have a significant impact on tar yield and char yield, but no significant impact on water yield.

2.2. Interaction Effect Investigation Between n-Hexane Insolubles and Mixing Ratios

2.2.1. Model Establishment

First, we use the read_excel function of Pandas to load data from Excel files. The loaded data is stored in a Pandas DataFrame, which facilitates subsequent data operations and analysis. Visualization is performed using the processed dataset.

We take the visualization results of tar yield as an example for illustration, as shown in Figure 3 below:

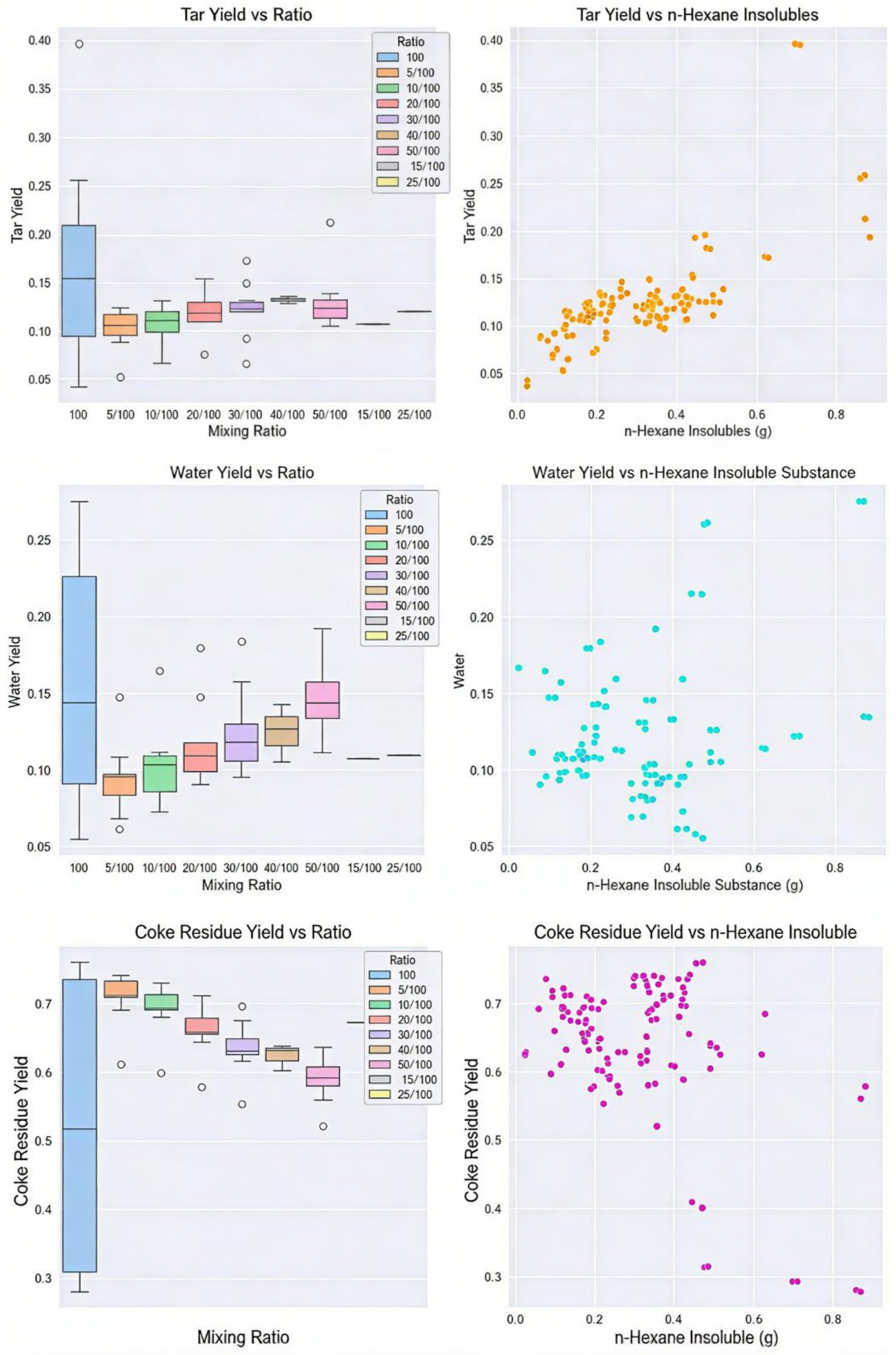

Figure 3 Visualization Results

Box plot (top left figure): Tar yield shows different distribution characteristics under different mixing ratios. The box plot for the mixing ratio of 100 (red box) shows a relatively high median and large data dispersion, including some low outliers. Other mixing ratios such as 0, 5/100, etc., show moderate tar yields, but with different dispersion and outlier conditions.

Scatter plot (top right figure): Tar yield shows certain fluctuations with the increase of n-hexane insoluble content, without an obvious linear relationship. The data is relatively concentrated when the n-hexane insoluble content is low, and the variation range of tar yield increases with the increase of content.

2.2.2. Model Solution

We set linear model formulas containing continuous variables and categorical variables for three different yields (tar yield, water yield, and char yield) respectively. We use the ols function of the Statsmodels library to construct linear models and the anova_lm function to perform analysis of variance (ANOVA) to evaluate the statistical significance of variables in the models. The specific results are shown in Table 1 below:

Table 1 Statistical Significance of Analysis of Variance

ANOVA Results for Tar Yield	ANOVA Results for Water Yield	ANOVA Results for Char Yield
C(Mixing Ratio): Sum of Squares (sum_sq): 0.103658 Degrees of Freedom (df): 9 F-value: 4.558299, indicating that the mixing ratio has a significant impact on tar yield. p-value (PR(>F)): 0.000034, very small, indicating that different levels of mixing ratio have a statistically significant impact on tar yield.	C(Mixing Ratio): Sum of Squares (sum_sq): 0.030958 Degrees of Freedom (df): 9 F-value: 2.283420, relatively low, but still statistically significant. p-value (PR(>F)): 0.020964, indicating that the mixing ratio has statistical significance on water yield at the 5% significance level.	C(Mixing Ratio): Sum of Squares (sum_sq): 0.286610 Degrees of Freedom (df): 9 F-value: 4.289856, indicating that the mixing ratio has a significant impact on char yield. p-value (PR(>F)): 0.000072, very small, indicating that different levels of mixing ratio have a significant impact on char yield.
INS_g: Sum of Squares (sum_sq): 0.041865 Degrees of Freedom (df): 1 F-value: 16.568858, indicating that n-hexane insoluble content has a significant impact on tar yield. p-value (PR(>F)): 0.000083, very small, further confirming the statistical significance of n-hexane insoluble content.	INS_g: Sum of Squares (sum_sq): 0.003403 Degrees of Freedom (df): 1 F-value: 2.258894, relatively low, close to insignificant. p-value (PR(>F)): 0.135391, relatively large, indicating that the impact of n-hexane insoluble content on water yield is not statistically significant.	INS_g: Sum of Squares (sum_sq): 0.113100 Degrees of Freedom (df): 1 F-value: 15.235472, indicating that n-hexane insoluble content has a significant impact on char yield. p-value (PR(>F)): 0.000155, very small, confirming its statistical significance.
Residual: Sum of Squares (sum_sq): 0.313313 Degrees of Freedom (df): 124	Residual: Sum of Squares (sum_sq): 0.186797 Degrees of Freedom (df): 124	Residual: Sum of Squares (sum_sq): 0.920508 Degrees of Freedom (df): 124

These ANOVA results reveal the degree and significance of the impact of mixing ratio and n-hexane insoluble content on different types of yields. For tar yield and char yield, both mixing ratio and n-hexane insoluble content have a significant impact on the yields, while for water yield, although the mixing ratio has a significant impact, the impact of n-hexane insoluble content is not significant. These insights help to further adjust the mixing ratio in the production process and control the amount of n-hexane insolubles to optimize yields.

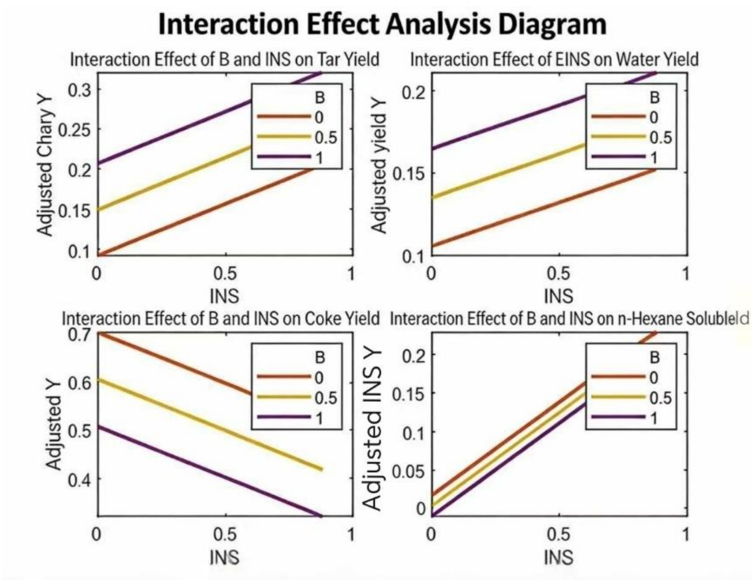

Figure 4 Interaction Effect Diagram

Figure 4 above shows the interaction effects of different mixing ratios (B) and n-hexane insoluble (INS) contents on various pyrolysis product yields (tar yield, water yield, char yield, n-hexane soluble yield). The slope and distribution of the lines in each figure reveal the interaction between mixing ratio and n-hexane insoluble content in predicting various yields. The detailed description of each figure is as follows:

Tar yield (top left figure)

This figure shows the change of tar yield when the n-hexane insoluble content increases from 0 to 1 under different mixing ratios (0, 0.5, 1). Under all mixing ratios, tar yield shows an upward trend with the increase of n-hexane insoluble content. The upward trend is the most significant when the mixing ratio is 1, showing a high growth rate of tar yield, while the growth rate is the lowest when the ratio is 0.

Water yield (top right figure)

Similarly, water yield increases with the increase of n-hexane insoluble content, and different mixing ratios affect the growth rate of water yield. Under the mixing ratios of 0.5 and 1, water yield increases significantly with the increase of n-hexane insoluble content, and the slope is the largest when the ratio is 1.

Char yield (bottom left figure)

Char yield decreases with the increase of n-hexane insoluble content, especially at high mixing ratios. The decreasing trend is the most significant when the mixing ratio is 1, indicating that char yield will decrease significantly at high mixing ratios and high n-hexane insoluble content.

n-hexane soluble yield (bottom right figure)

n-hexane soluble yield increases with the increase of n-hexane insoluble content, especially when the mixing ratios are 0 and 0.5. Although it also shows an increasing trend when the mixing ratio is 1, the growth rate is relatively gentle.

2.2.3. Results

This set of charts clearly reveals the impact of mixing ratio and n-hexane insoluble content on different pyrolysis product yields, as well as the significant interaction between the two. For different yield indicators, the combination of mixing ratio and n-hexane insoluble content has different impact trends, which has important guiding significance for adjusting the mixing ratio in the production process and the content of chemical additives.

2.3. Multi-Objective Optimization-Based Co-Pyrolysis Mixing Strategy Development

2.3.1. Establishment of Optimization Model

The goal is to maximize the tar yield, water yield, and char yield in the pyrolysis process by optimizing the mixing ratio B and n-hexane insoluble content INS. The detailed description of the optimization model is as follows:

Objective function:

In this optimization model, we try to maximize three different yields (tar yield, water yield, char yield) at the same time, and each yield is predicted by the corresponding linear model. The objective function is defined as a linear combination of these yields, where the weight of each yield is negative, indicating that we hope to maximize them.

$$\begin{aligned} f(B, INS) = & -\mathrm{predict}_{\mathrm{tarModel}}(B, INS, \mathrm{defaultValues}) \\ & -\mathrm{predict}_{\mathrm{waterModel}}(B, INS, \mathrm{defaultValues}) \\ & -\mathrm{predict}_{\mathrm{charModel}}(B, INS, \mathrm{defaultValues}) \end{aligned} \tag{2}$$

Constraints:

In this optimization problem, the value ranges of variables B and INS are limited to [0, 1], which reflects the operational or mixing ratio constraints in actual situations. For example, B and INS may represent the percentage content of a certain mixture, which must be between 0% and 100%.

$$0 \le B \le 1, 0 \le INS \le 1 \tag{3}$$

Multi-objective optimization:

Using the fgoalattain function, we set the goal attainment level, that is, the minimum tar yield, water yield, and char yield we hope to achieve. These goals are balanced by the weight vector, reflecting the importance or priority of different yields.

$$\mathrm{goal} = [0.3, 0.2, 0.5] \tag{4}$$

$$\mathrm{weight} = [-1, -1, -1] \tag{5}$$

Economic constraints:

Cost is an important factor restricting production. The optimization model should consider the cost of raw materials, especially when the cost of different input materials varies greatly.

$$\mathrm{COST}(B, INS) = c_B \cdot B + c_{INS} \cdot INS \le \mathrm{Budget} \tag{6}$$

where c_B and c_{INS} are the unit costs of mixing ratio and n-hexane insolubles respectively, and Budget is the budget upper limit.

There may be equipment or technical limitations in actual operation, such as the maximum bearing capacity of the reactor in terms of temperature and pressure.

$$T_{min} \le f(B, INS) \le T_{max} \tag{7}$$

$$P_{min} \le g(B, INS) \le P_{max} \tag{8}$$

Here, T_{max} and T_{min} represent the maximum and minimum allowable temperatures respectively, P_{max} and P_{min} are the pressure limits, and functions f and g describe the relationship between temperature, pressure and input variables.

To sum up, the following optimization model is established:

$$\begin{aligned} & \max f(B, INS) = -\mathrm{predict}_{\mathrm{tarModel}}(B, INS, \mathrm{defaultValues}) \\ & -\mathrm{predict}_{\mathrm{waterModel}}(B, INS, \mathrm{defaultValues}) - \mathrm{predict}_{\mathrm{charModel}}(B, INS, \mathrm{defaultValues}) \\ & \mathrm{s.t.}\ T_{min} \le f(B, INS) \le T_{max} \\ & \quad P_{min} \le g(B, INS) \le P_{max} \\ & \quad \mathrm{COST}(B, INS) = c_B \cdot B + c_{INS} \cdot INS \le \mathrm{Budget} \\ & \quad \mathrm{goal} = [0.3, 0.2, 0.5] \\ & \quad \mathrm{weight} = [-1, -1, -1] \\ & \quad 0 \le B \le 1 \\ & \quad 0 \le INS \le 1 \end{aligned} \tag{9}$$

To search for the optimal mixing ratio B and n-hexane insoluble content INS, we use the Sequential Least Squares Programming (SLSQP) algorithm. When using fgoalattain for optimization solution, we use the SQP (Sequential Quadratic Programming) algorithm. This algorithm is a commonly used method in nonlinear optimization and is suitable for handling such constrained optimization problems.

In each iteration, the algorithm solves a quadratic programming subproblem to approximate the original nonlinear problem:

$$\nabla^2 L(x_k) d + \nabla f(x_k) = 0, \nabla h_j(x_k)^T d + h_j(x_k) = 0, j = 1, \dots, l \tag{10}$$

where d is the search direction from the current point x_k to the new point.

Line search and parameter update: After finding the direction d, a suitable step size α is selected through line search to update the solution $x \leftarrow x + \alpha d$. Line search ensures that the objective function decreases in the direction that satisfies the constraints.

Convergence judgment: Check whether the update of the solution is small enough, whether the maximum number of iterations is reached, or other convergence criteria, such as whether the norm of the gradient is small enough.

2.3.2. Solution of Optimization Model

The optimization results are shown in Figure 5 below:

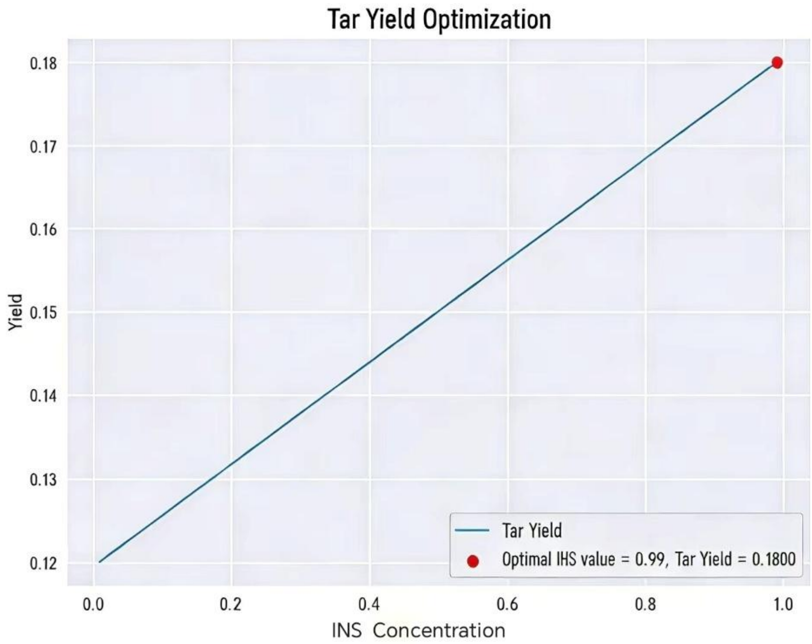

Figure 5 Optimization Results

3. Conclusions

Through modeling and analysis of the biomass-coal co-pyrolysis process, this study identified n-hexane-insoluble substances as the core factor driving variations in tar and coke yields. It systematically elucidated the interactive mechanism between these substances and the mixing ratio, ultimately providing an optimal raw material blending scheme that offers theoretical support for practical industrial operations. However, the analysis of variance employed in this study imposes stringent requirements on data normality and homogeneity of variance. Furthermore, the model may encounter multicollinearity issues when handling highly correlated variables, potentially compromising its stability. Future research should focus on incorporating more complex dynamic reaction parameters and exploring the integration of more robust machine learning algorithms to enhance predictive accuracy and optimization capabilities within diverse industrial settings.

References

[1] Kristanto J, Azis M M, Purwono S. Enhancing machine learning prediction of biomass pyrolysis yields through taxonomy-based preprocessing and rigorous model tuning[J]. Biomass and Bioenergy, 2026, 205108524-108524.

[2] Yani A, Wijayanti W, Sasongko N M, et al. Synergistic co-pyrolysis of polyethylene terephthalate and hibiscus rosa-sinensis: Catalytic mechanisms for optimizing pyrolysis yields[J]. South African Journal of Chemical Engineering, 2025, 53281-302.

[3] Nagawkar R B, Subramaniam S, Brown C R, et al. Computational models for the prediction of yields in the autothermal pyrolysis of biomass[J]. Chemical Engineering Science, 2025, 311121599-121599.

[4] Mariyam S, Alherbawi M, McKay G, et al. A predictive model for biomass waste pyrolysis yield: Exploring the correlation of proximate analysis and product composition[J]. Energy Conversion and Management: X, 2025, 25100831-100831.

[5] Ma H, Zhang Y, Qiu L, et al. Effect of Reaction Conditions on Energy Yield of Pyrolysis Gas from Apple Tree Branches[J]. ACS omega, 2024, 9(26): 28028-28036.

[6] Elhenawy Y, Fouad K, Bassyouni M, et al. Yield and energy outputs analysis of sawdust biomass pyrolysis[J]. Energy Conversion and Management: X, 2024, 22100583-.

[7] Ikhlass. B, S. A, F. T. Modelling and simulation of fast pyrolysis of pomace from three-phase olive mill targeting optimal yields of pyrolysis products[J]. Biofuels, 2024, 15(3): 349-361.

[8] Arumugam B D, Ganesan C M. Characterization of cassia fistula seed and shell co-pyrolysis yield optimization by response surface methodology and compositional analysis[J]. Process Safety and Environmental Protection, 2024, 184355-376.

[9] Rodrigues L, Budai A, Elsgaard L, et al. The importance of biochar quality and pyrolysis yield for soil carbon sequestration in practice[J]. European Journal of Soil Science, 2023, 74(4).

[10] Jun P Y, Hu T L, Hao L, et al. Effect of coal tar pitch modification on the structure and char yield of pyrolysis epoxy resin carbons[J]. Diamond & Related Materials, 2023, 137.