Problem Chosen	2023 ShuWei Cup Summary Sheet	Team Control Number
B		2023110620859

Catalytic reaction of cotton straw pyrolysis based on weighted kinetic modeling

With the increasing global demand for renewable energy, cotton straw is considered as an important bio-resource due to its abundant biomass components such as cellulose and lignin. In this paper, the effective utilization and sustainable development of cotton straw will be investigated based on the pyrolysis products under different catalytic pyrolysis reactions and pyrolysis gas product data.

For Problem 1, scatter plots of product mixing ratio and gas yield versus catalyst ratio for each catalyst combination were firstly drawn, and the preliminary trend was determined after comprehensively analyzing the errors in combination with related literature. After that, linear fitting and polynomial fitting were utilized to derive the relationship between the yield of pyrolysis products and the mixing ratio of the corresponding pyrolysis combination.

For problem 2, for the three pyrolysis combinations, we first visualized the data using line plots, and the data were depicted statistically by observing the variation of each data point in the image and the effect of the ratios. Secondly, stacked plots were used to delve into the effects of different pyrolysis combination ratios on the pyrolysis gas yield of each group. In addition, in order to emphasize and demonstrate the correlation between pyrolysis combination ratios and pyrolysis gas yields, Pearson's correlation coefficient was used for the calculation. Through the results of the study, we found that there is a certain statistical correlation between the data, which further confirms our hypothesis.

For problem 3, according to the P-P plot and K-S test, we obtained that the distributions of the nine groups of product yield data and gas yield data were in line with normal distribution. Therefore, the independent sample t-test was used to analyze the differences, and it was concluded that under the catalytic effect of the same proportion of desulfurization ash, there were significant differences in the yields of CE and LG pyrolysis products and pyrolysis gas components.

For problem 4, we used grouped histograms to compare the differences in pyrolysis gas fraction yields under the mixing ratios of FGD ash/cotton stalks, FGD ash/CE, and FGD ash/LG, to explore the mechanism of catalytic pyrolysis reaction. Then, we modeled the reaction kinetics of catalytic pyrolysis of CE and LG, respectively. Then, we examined the effect of the co-catalytic pyrolysis process on the reaction kinetic parameters, and performed weighted average calculations for the model mixtures in the co-catalytic pyrolysis process. Eventually, we found that the LG and CE catalytic pyrolysis basically overlapped. As can be seen from the kinetic parameters, the effect of co-catalytic pyrolysis is obvious.

For problem 5, First, we modeled the product yield with support vector machine and used SVR technique to establish the relationship function with the proportion of pyrolysis combinations and selected the best function to predict the product yield by adjusting the kernel function and model parameters. Ridge regression and polynomial regression were also used to predict the pyrolysis gas yield. Polynomial regression, on the other hand, solves the nonlinear relationship between product yield and pyrolysis reaction combination ratio by fitting a quadratic polynomial model, which can be predicted more accurately by combining the prediction results of the three methods.

key word： Polynomial regression ;Correlation analysis ;Reaction kinetics ;SVM ;Pyrolysis of cotton stalks

Content

1. Introduction 1
1.1 Background 1
1.2 Work 1
2.1 Data analysis 2
2.2 Analysis of question one 2
2.3 Analysis of question two 3
2.4 Analysis of question three 3
2.5 Analysis of question four 3
3. Symbol and Assumptions 4
3. 1 Symbol Description 4
3.2 Fundamental assumptions 4
4. Model 4
4.1 Model establishment and solution of problem 1 4
4.1.1 Data Preparation 5
4.1.2 Model establishment 7
4.1.3 Conclusion and analysis 13
4.2 Model establishment and solution of problem 2 14
4.2.1 Data Preparation 14
4.2.2 Conclusion and analysis 15
4.3 Model establishment and solution of problem 3 18
4.3.1 Normal distribution test 18
4.3.1.2 T-test 20
4.4 Model establishment and solution of problem 4 20
4.5 Model establishment and solution of problem 5 22
4.5.1 Support vector machine based prediction models 22
4.5.2 Predictive modeling based on ridge regression 24
4.5.2 Forecasting model based on polynomial regression 25
6. Strengths and Weakness 26
References 27

1. Introduction

1.1 Background

The excessive dependence of human beings on non-renewable fossil energy makes the global sustainable development strategy is facing a huge challenge, so it is urgent to find a renewable energy source to solve the world's energy problems. Biomass energy has the advantages of renewable, abundant reserves, low pollution, etc., is the only energy source that can realize "carbon sequestration", and in recent years, the use of biomass energy has received widespread attention. As an important biomass resource, cotton stalks can produce various forms of renewable energy through pyrolysis. However, conventional pyrolysis results in low yield and poor quality products, and catalytic pyrolysis can be directed to regulate the biomass pyrolysis products to improve their quality. Therefore, investigating the mechanism and effects of catalysts during the pyrolysis process, is of great significance for the efficient utilization and sustainable development of cotton stalk.

A chemical engineering laboratory used the model compound method to establish the pyrolysis combinations of desulfurization ash with cotton stalk and desulfurization ash with model compounds at different mixing ratios in order to investigate the catalytic mechanism and effects of desulfurization ash on the generation of pyrolysis products from the cotton stalk. Based on the fact that CE (Cellooligosaccharide) and LG (lignin) can represent the biomass properties of cotton straw itself, these two components were chosen as model compounds, which can control the experimental conditions more finely in order to study the targeted catalytic effect of desulfurized ash on different biomass components.

1.2 Work

Annex I shows pyrolysis product yields of three pyrolysis combinations and Annex II shows pyrolysis gas yields of three pyrolysis combinations.In this paper, mathematical modeling is used to solve the following problems.

- **Question 1:** Firstly, analyze the correlation between pyrolysis product yields and the corresponding pyrolysis combinations.Secondly, explain whether desulfurization ash as a catalyst plays a significant role in facilitating the pyrolysis of cotton stalks, cellulose, and lignin.
- **Question 2:** Analyze the effects of different pyrolysis combination ratios on the pyrolysis gas yields of each group and produce images for interpretation.
- **Question 3:** Explore whether there are significant differences in the yields of CE and LG pyrolysis products as well as pyrolysis gas fractions under the catalytic effect of the same proportion of desulfurization ash and explain the reasons.

- **Question 4:** Establish a catalytic reaction mechanism model of desulfurized ash for model compounds such as CE and LG and model the reaction kinetics for analysis.
- **Question 5:** Predict the yield or quantity of pyrolysis products using mathematical modeling or artificial intelligence learning methods.

2. Problem analysis

2.1 Data analysis

Annex I shows the relationship between the yield of pyrolysis products and the ratio of the corresponding pyrolysis combinations under the three pyrolysis combinations. Firstly, by observing the data in the table, it is found that the sum of the proportion of the four pyrolysis products under each pyrolysis combination ratio is 100% and there are no missing values, so it can be initially judged that there are no outliers or missing values in the data in Annex I. Next, the box plots of the yields of the four pyrolysis products under the ratios of each pyrolysis combination were plotted, and it was found that there were only a few outliers, but the effect of the outliers was ignored in consideration of the limited sample size. Next, the q-q plots were plotted and subjected to the Shapiro-Wilk test, and it was found that the data all conformed to a normal distribution. By performing descriptive statistics on the data, it was found that the statistical indicators of the data were within reasonable limits. Finally, the overall pattern of data variation was reacted by plotting a line graph. Annex II shows the relationship between pyrolysis gas yield and the ratio of the corresponding pyrolysis combinations under the three pyrolysis combinations, and the data were analyzed along the same lines as in Annex I.

2.2 Analysis of question one

Question one required a cross-sectional study of the data for each pyrolysis combination in Annex I. We analyzed the relationship between the yields of pyrolysis products (tar, water, coke residue, and syngas) and the ratios of the corresponding pyrolysis combinations, respectively, and explained whether desulfurization ash as a catalyst is important in promoting the pyrolysis of cotton straw, cellulose, and lignin. Firstly, we plotted box plots of the Annex I data and analyzed them with descriptive statistics and found no outliers or missing values. Secondly, we based on the scatter plots of the yields of the four pyrolysis products under each pyrolysis combination versus the ratio of pyrolysis combinations, and initially derived the variation pattern between the two. Next, the Pearson correlation coefficient was used to determine whether the data were correlated or not and a heat map was drawn to visualize the magnitude of the correlation. Finally, based on the principle of the least squares method, a linear fitting model, a quadratic polynomial fitting model and a cubic polynomial fitting model were established according to the specific characteristics of

each group of scatter plots, and the optimal model was selected for the specific solution, which finally led to the conclusion that desulfurization ash was used as a catalyst to promote the pyrolysis of cotton straw, cellulose and lignin.

2.3 Analysis of question two

Question 2 requires us to analyze the effect of different pyrolysis ratios on the pyrolysis gas yield of each group and make corresponding images for interpretation. Firstly, we preprocessed the data in Annex II and found that there were no outliers and missing values by plotting box plots, and then we visualized the effect of different pyrolysis combination ratios on the pyrolysis gas yield of each group by plotting scatter plots, stacked plots and thermograms. Based on the visualization results, we will analyze and explain the effect of desulfurization ash on each gas yield.

2.4 Analysis of question three

Question 3 required a variability analysis of the yields of CE and LG pyrolysis products as well as pyrolysis gas fractions under the catalytic action of the same proportion of desulfurized ash. First, the data were plotted by Q-Q plots and subjected to the Shapiro-Wilk test and found to be normally distributed. Secondly, the stacked diagrams were drawn to visualize the differences in the yields of CE and LG pyrolysis products and pyrolysis gas components under the catalytic effect of the same proportion of desulfurization ash, and the conclusion of significant difference was initially made. Finally, the differences were analyzed by T-test.

2.5 Analysis of question four

For problem 4, firstly, we need to carry out an in-depth theoretical analysis and experimental study on the experimental data demonstrating the differences between CE and LG and combining the chemical properties and structures in order to determine their specific mechanisms of action in the catalytic reaction of desulfurization ash. Next, we can establish the kinetic model of the desulfurization ash catalytic reaction of CE and LG based on the experimental data and theoretical analysis, including the determination of parameters such as the reaction rate equation, activation energy, and the number of reaction stages. Through the fitting of experimental data and mathematical modeling, we can describe the kinetic properties of the desulfurization ash catalytic reaction of CE and LG. Finally, we can validate and optimize the established model to ensure that it can accurately describe the mechanism and kinetic properties of the desulfurization ash catalytic reaction of CE and LG.

3. Symbol and Assumptions

3. 1 Symbol Description

Tab 1:Symbol Description

Synbol	meaning	unit (of measure)
wt	Mixing ratio	%
daf	Yield	mL/g
xi	Catalyst ratio	无
v	Average reaction rate	g/h
m0	Initial mass	g
mt	Instantaneous mass	g
m1	Final mass	g
X	Conversion rate	%

3.2 Fundamental assumptions

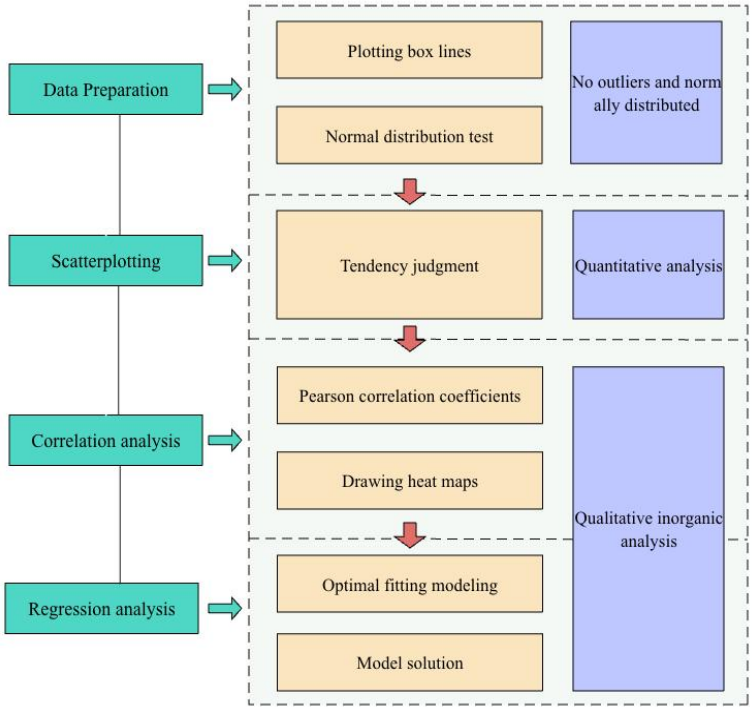
1. The pyrolysis reaction is carried out under constant temperature conditions and the effect of temperature on the yield of pyrolysis products is considered to be constant.
2. the catalytic effect of the studied desulfurized fly ash at different mixing ratios is quantifiable and can be described by a polynomial regression model.
3. the generation of pyrolysis products is predictable and independent of factors such as reaction time.
4. other external factors that may affect the pyrolysis process, such as atmosphere, reaction pressure, etc., are not considered.
5. the gas yield is influenced by the interaction between the studied desulfurized fly ash and biomass components, which can be described in terms of mixing ratios.
6. The variation of the gas yield is determined by the temperature and the pyrolysis reaction catalyzed by the desulfurized fly ash, and other factors do not have a significant effect for this model.

4. Model

4.1 Model establishment and solution of problem 1

After finding that the data have no outliers and conform to normal distribution through data preprocessing, we initially derive the variation pattern of the four pyrolysis products under each pyrolysis combination based on the scatter plots of the yield of the four pyrolysis products and the ratio of pyrolysis combinations. Next, the Pearson correlation coefficient was used to determine whether the data were

correlated. Finally, linear fitting models, quadratic polynomial fitting models and cubic polynomial fitting models were established based on the specific characteristics of each group of scatter plots, and the optimal model was selected for solving the problem, which finally led to the conclusion that desulfurization ash was used as a catalyst to promote the pyrolysis of cotton straw, cellulose and lignin.

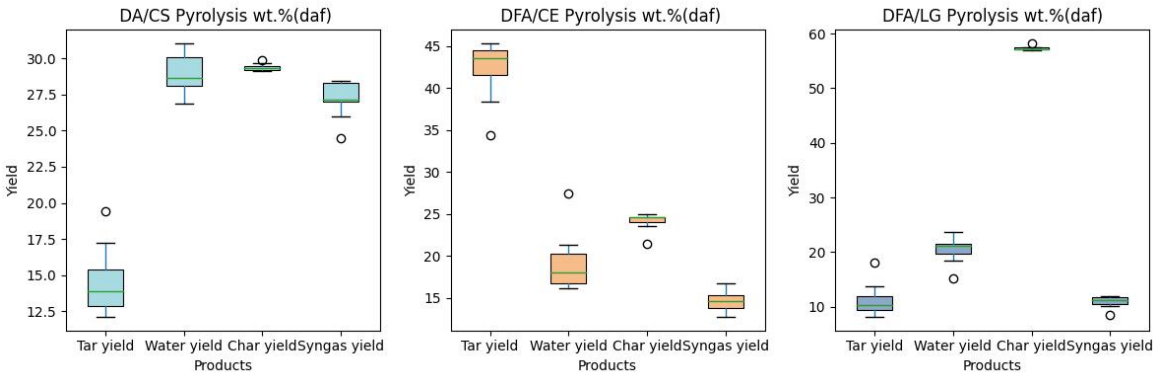


Flowchart of the first question

4.1.1 Data Preparation

4.1.1.1 Outlier checking

A box plot is a type of statistical chart commonly used to display information on the dispersion of a set of data, which facilitates an accurate and stable depiction of the discrete distribution of data and data processing. We made the following box-line plots using the yield of pyrolysis products from each pyrolysis combination at different ratios as sample data:

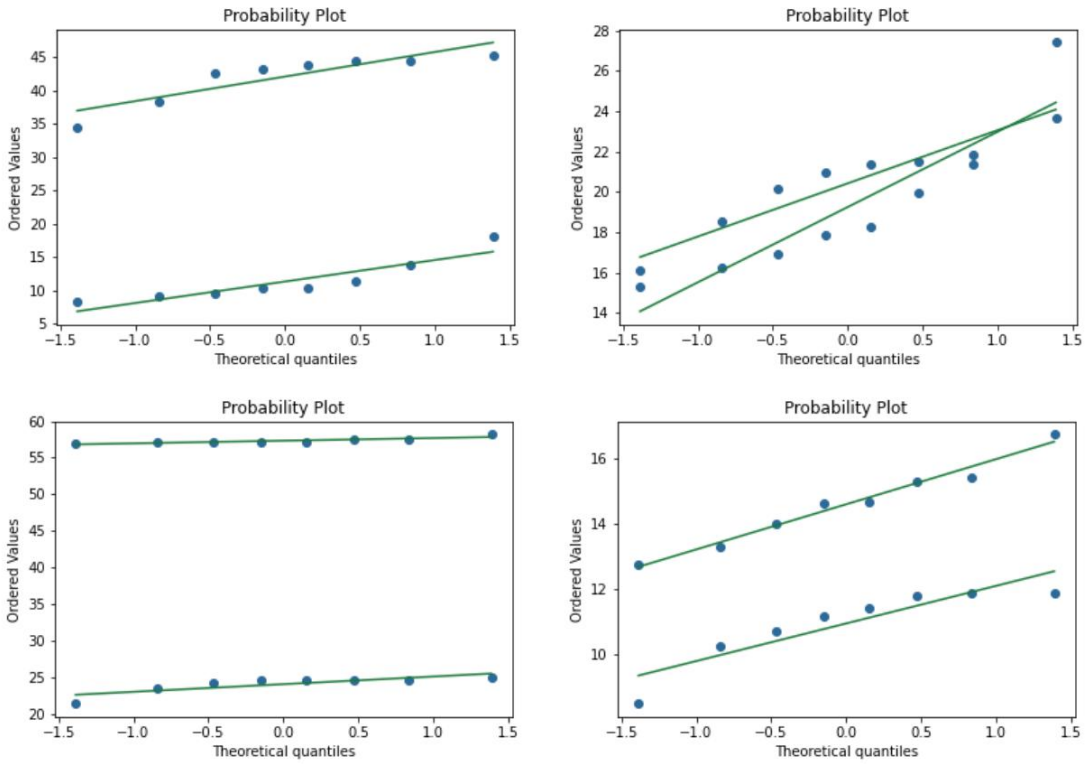


Box-line plots of pyrolysis product yield at different ratios

Off-mean values exceeding 4 times the standard deviation of the coefficient factor are judged as abnormal sizes, and the number of abnormal values is found to be small by observing the length of the box and line segments of each party, and the effect of the abnormal values can be ignored considering the small sample size.

4.1.1.2 Normal distribution test

Quantile Quantile Plot graphically compares the probability distributions of different quartiles of the observed and predicted values (distribution under the assumption of normality), so as to test whether they coincide with the law of normal distribution, the higher the overlap between the scatter and the straight line the more it obeys normal distribution, and the greater the difference between the scatters the more it does not obey the normal distribution. By plotting the scatter plot in Annex I, it is found that the data set is distributed around a straight line, indicating that the data conforms to a normal distribution.



Q-Q plots of pyrolysis product yield at different ratios

4.1.1.3 Descriptive statistical analysis

The descriptive statistics such as maximum value, minimum value, extreme deviation and standard deviation of pyrolysis product yield for each pyrolysis combination at different ratios are shown in the table below:

Tab 1:Descriptive statistical analysis of pyrolysis product yield from DFA/CS

DFA/CS	MAX	MIN	Range	AVG	STP
Tar yield (wt.%)	19.46	12.13	7.33	15.04	2.37
Water yield (wt.%)	31.02	26.84	4.18	28.62	2.33

Char yield (wt.%)	29.87	29.11	0.76	29.47	0.261
Syngas yield (wt.%)	28.45	24.49	3.96	27.05	1.51

Tab 2:Descriptive statistical analysis of pyrolysis product yield from DFA/CE

DFA/CE	MAX	MIN	Range	AVG	STP
Tar yield (wt.%)	45.28	34.42	10.86	42.71	4.5
Water yield (wt.%)	27.42	16.14	11.28	18.39	3.65
Char yield (wt.%)	24.91	21.43	3.48	24.6	1.21
Syngas yield (wt.%)	16.73	12.75	3.98	14.83	1.85

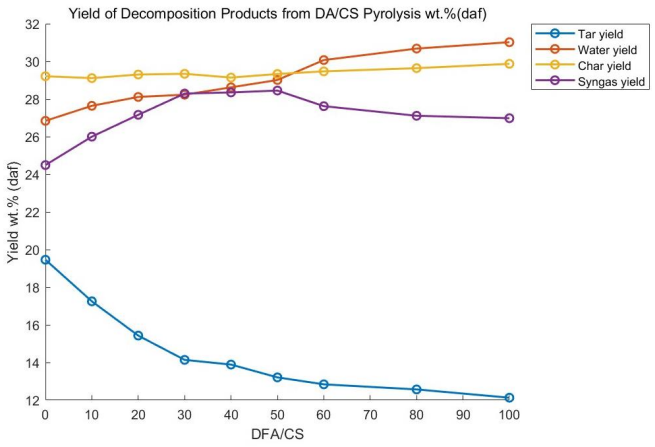
Tab 3:Descriptive statistical analysis of pyrolysis product yield from DFA/LG

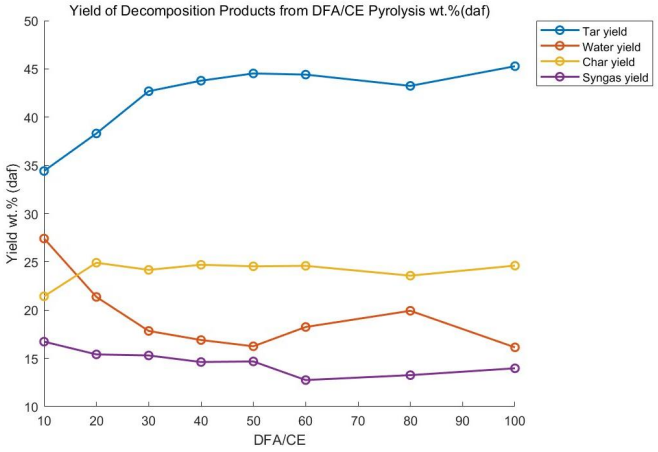
DFA/LG	MAX	MIN	Range	AVG	STP
Tar yield (wt.%)	18.06	8.19	9.87	10.07	3.23
Water yield (wt.%)	23.69	15.3	8.39	19.92	2.97
Char yield (wt.%)	58.17	56.98	1.19	57.23	0.379
Syngas yield (wt.%)	11.88	8.47	3.41	10.85	1.6

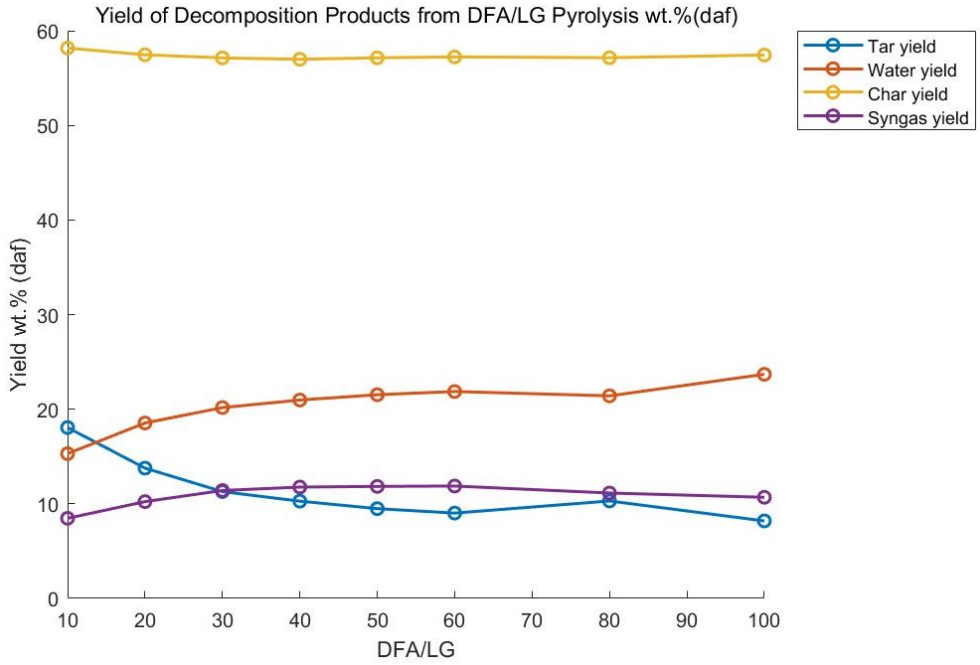
4.1.2 Model establishment

4.1.2.1 Initial judgment based on scatterplot

In order to more intuitively analyze the relationship between the yield of pyrolysis products and the ratios of corresponding pyrolysis combinations, we consider visualizing the data on the ground under various pyrolysis combinations and draw scatter plots as follows:







Scatter plot of pyrolysis product yield vs mixing ratio

4.1.2.2 Correlation analysis

In this paper, the correlation between the pyrolysis product yield and the corresponding pyrolysis combination ratios for each pyrolysis combination was derived by calculating the Pearson correlation coefficient and visualized in the form of a heat map as shown in Figure 5-8.

In Eqs.(5-10), x represents the mixing ratio and y_i represents the yield of each pyrolysis product.

$$p_{x,y} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} = \frac{E[(x-x_i)(y-y_i)]}{\sigma_x \sigma_y} \tag{5-1}$$

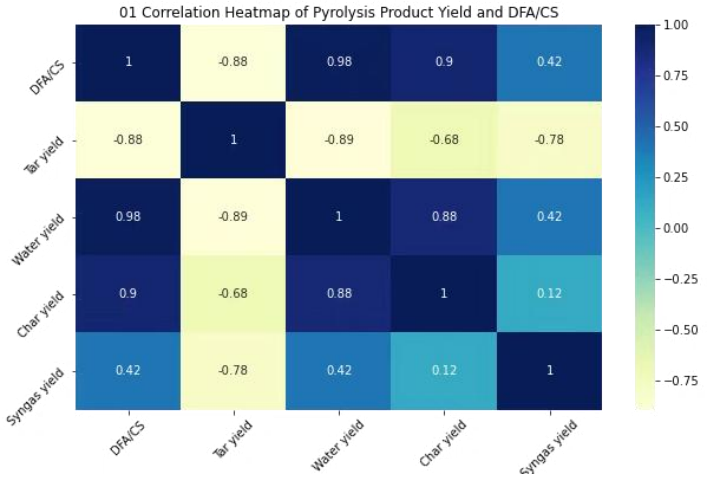


Figure 5:Correlation heatmap of pyrolysis product yield from DFA/CS

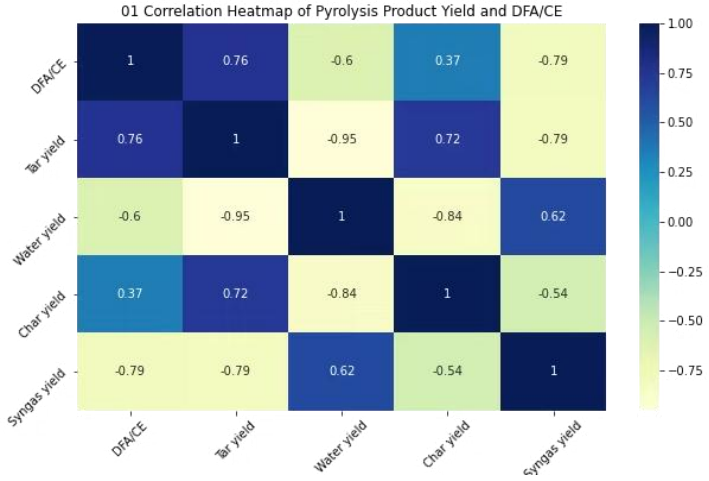


Figure 6:Correlation heatmap of pyrolysis product yield from DFA/CE

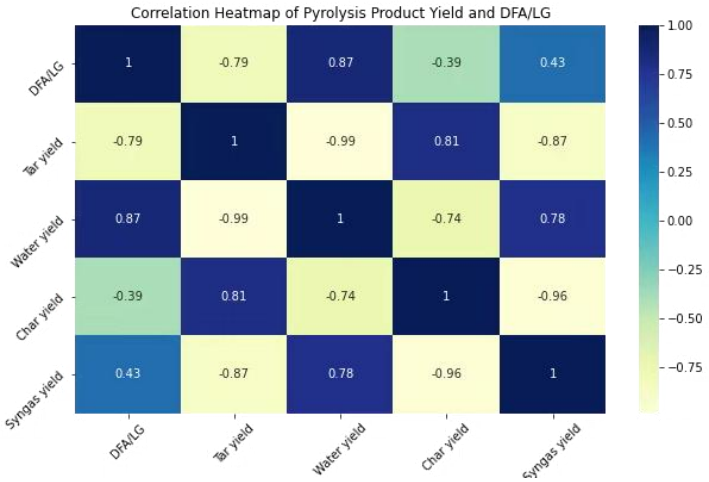


Figure 7:Correlation heatmap of pyrolysis product yield from DFA/LG

From the Pearson's correlation coefficient calculation results and the heat map, it can be seen that the Pearson's correlation coefficients of 8 out of 12 groups of data are greater than 0.6.Therefore, there is a strong correlation between the yield of pyrolysis products under different pyrolysis combinations and the ratios of the corresponding pyrolysis combinations.

4.1.2.3 Fitting Model Establishment

(1) Establishment of the fitting function

By observing the shape characteristics of the 12 scatter plots, the scatter plots of pyrolysis product yield versus mixing ratio were initially categorized into two types: linear and presence of peaks. Therefore, it can be preliminarily determined that three fitting functions are used: linear function

$$y_i = ax + b \tag{4-1}$$

quadratic function

$$y_i = ax^2 + bx + c \tag{4-2}$$

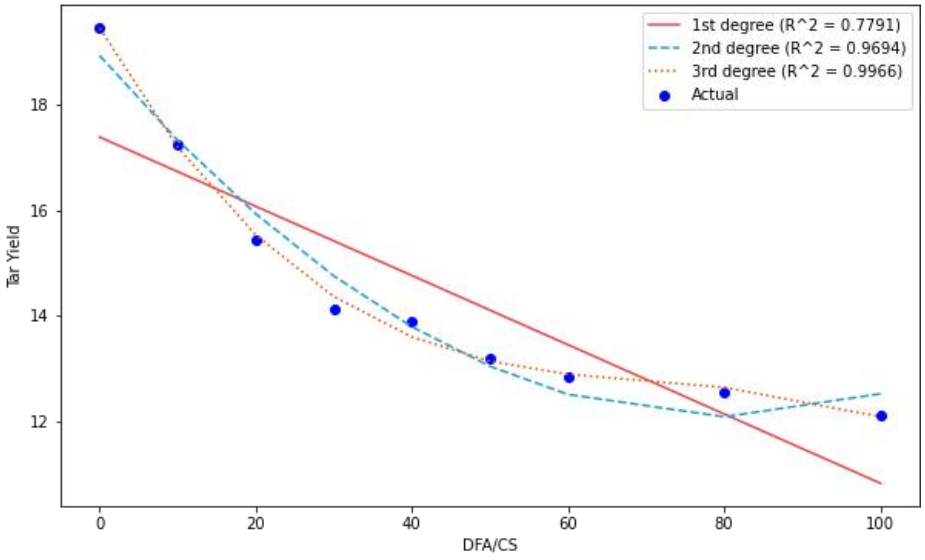
cubic polynomial function

$$y_i = ax^3 + bx^2 + cx + d \tag{4-3}$$

(2) Establishment of the fitting model

① Tar yield and mixing ratio fitting model

We fit the scatter plots of tar yield versus mixing ratio for four pyrolysis combinations based on the principle of least squares with a linear function y=ax+b a quadratic function $y = ax^2 + bx + c$ and a cubic polynomial function $y = ax^3 + bx^2 + cx + d$ and analyze them comparatively, as in the case of DFA/CS, as follows:



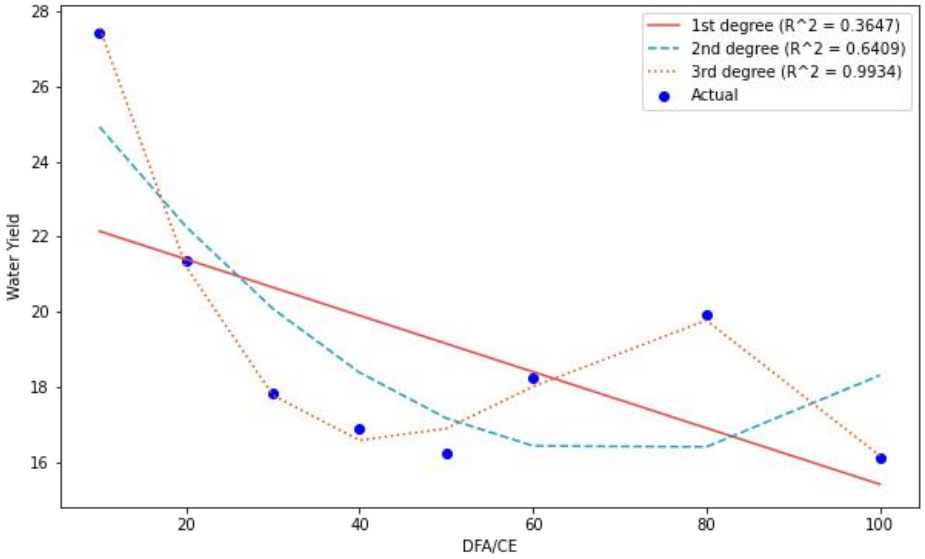
Comparison of linear, quadratic and cubic fits

From the comparison graph, it can be visualized that for DFA/CS, the

relationship between tar yield and mixing ratio is obviously better fitted by a cubic polynomial with R^2 = 0.9966, which is significantly larger than the linear and quadratic polynomial fits, which suggests that the cubic polynomial model can better describe the relationship between tar yield and mixing ratio under DFA/CS. Eventually, all four combinations were subjected to the above process to determine the optimal fitting model with specific equations.

② Water yield and mixing ratio fitting model

We fit the scatter plots of tar yield versus mixing ratio for four pyrolysis combinations based on the principle of least squares with a linear function y=ax+b, a quadratic function $y = ax^2 + bx + c$, and a cubic polynomial function $y = ax^3 + bx^2 + cx + d$ and analyze them comparatively, as in the case of DFA/CE, as follows:

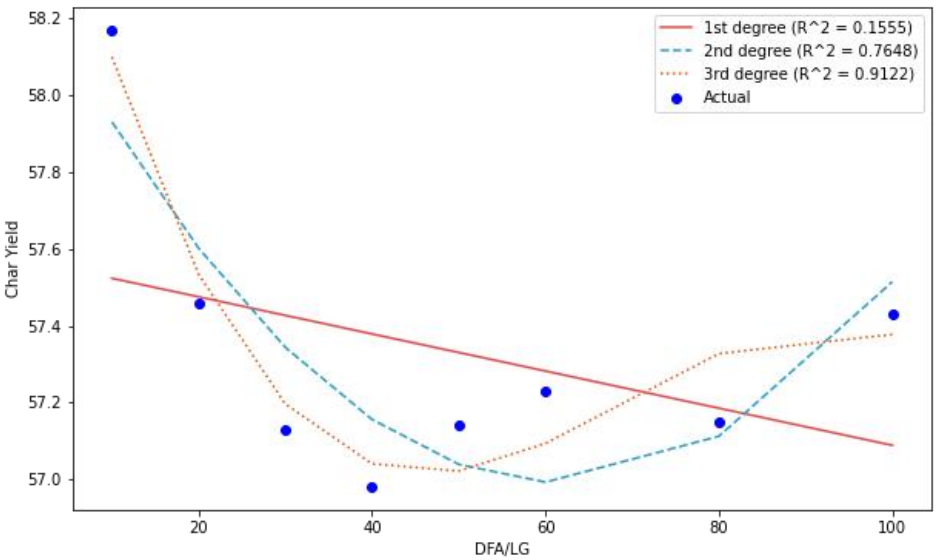


Comparison of linear, quadratic and cubic fits

From the comparison graph, it can be visualized that for DFA/CE, the relationship between aquatic yield and mixing ratio is clearly better fitted by a cubic polynomial, with R^2 = 0.9934, which is significantly larger than the linear and quadratic polynomial fits, which suggests that the cubic polynomial model can better describe the relationship between water yield and mixing ratio under DFA/CE. Eventually, all four combinations were subjected to the above process to determine the optimal fitting model with specific equations.

③ Char yield and mixing ratio fitting model

We fit the scatter plots of char yield versus mixing ratio for four pyrolysis combinations based on the principle of least squares with a linear function y=ax+b, a quadratic function $y = ax^2 + bx + c$, and a cubic polynomial function $y = ax^3 + bx^2 + cx + d$ and analyze them comparatively, as in the case of DFA/LG, as follows:

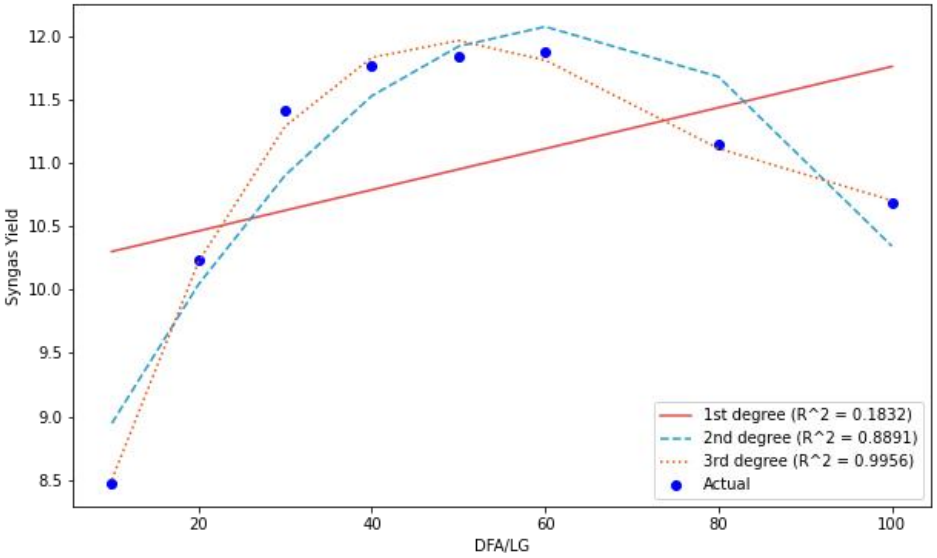


Comparison of linear, quadratic and cubic fits

From the comparison graph, it can be visualized that for DFA/LG, the relationship between aquatic yield and mixing ratio is clearly better fitted by a cubic polynomial, with R^2 = 0.9122, which is significantly larger than the linear and quadratic polynomial fits, which suggests that the cubic polynomial model can better describe the relationship between char yield and mixing ratio under DFA/LG. Eventually, all four combinations were subjected to the above process to determine the optimal fitting model with specific equations.

④ Syngas yield and mixing ratio fitting model

We fit the scatter plots of syngas yield versus mixing ratio for four pyrolysis combinations based on the principle of least squares with a linear function y=ax+b, a quadratic function $y = ax^2 + bx + c$, and a cubic polynomial function $y = ax^3 + bx^2 + cx + d$ and analyze them comparatively, as in the case of DFA/CE, as follows:



Comparison of linear, quadratic and cubic fits

From the comparison graph, it can be visualized that for DFA/LG, the relationship between aquatic yield and mixing ratio is clearly better fitted by a cubic polynomial, with R^2 = 0.9956, which is significantly larger than the linear and quadratic polynomial fits, which suggests that the cubic polynomial model can better describe the relationship between syngas yield and mixing ratio under DFA/LG. Eventually, all four combinations were subjected to the above process to determine the optimal fitting model with specific equations.

4.1.3 Conclusion and analysis

4.1.3.1 Relationship between pyrolysis product yield and mixing ratio

(1) DFA/CS pyrolysis product yield analysis

✧ The tar yield shows **a decreasing trend** with increasing mixing ratio, indicating that desulfurization ash may have a positive contribution to reducing tar production.

✧ The water yield, char and syngas yield, on the other hand, showed **an overall increasing trend** with the increase of mixing ratio, indicating that desulfurization ash may have a positive contributing effect in promoting the generation of the three.

(2) Analysis of DFA/CE pyrolysis product yield

✧ The tar yield **first rises and then slightly decreases** with the increase of mixing ratio, indicating that the desulfurization ash may have a certain promotion effect on reducing the generation of tar, but the effect is weakened at higher ratios.

✧ Water yield and char **have a little change** with the increase of the ratio, indicating that desulfurization ash may have less effect on the generation of water and char.

✧ The syngas yield **decreased slightly**, suggesting that desulfurization ash may have contributed to the reduction of syngas water.

(3) Analysis of DFA/LG pyrolysis product yield

✧ The tar yield shows **an overall decreasing trend** with the increase of mixing ratio, indicating that the desulfurization ash may have a positive contribution to the reduction of tar production.

✧ The water yield and syngas yield **increased slightly**, indicating that the desulfurization ash may have a certain role in promoting the generation of water and syngas.

✧ The overall char yield is **relatively stable**, indicating that desulfurization ash has a small effect as a catalyst.

4.1.3.2 Effectiveness of desulfurization ash as a catalyst

(1) **Pyrolysis of cotton stalks:** desulfurization ash plays a positive role in the pyrolysis of cotton stalks, especially in reducing tar yield.

(2) **Pyrolysis of cellulose:** desulfurization ash also contributes to the pyrolysis of

cellulose, especially at mixed ratios, but this effect is reduced at higher ratios.

(3) **Pyrolysis of lignin:** desulfurization ash shows more obvious promotion effect in the pyrolysis of lignin, mainly in the reduction of tar yield.

In summary, desulfurization ash as a catalyst in the pyrolysis process of cottonwood straw, cellulose and lignin all play a certain role in promoting. This effect showed some differences among different substances, especially in reducing the tar yield more significantly. However, this catalytic effect may change as the proportion of desulfurized ash increases, with the exact effect depending on the type of the original oh. Therefore, in practical applications, the selection of an appropriate proportion of desulfurized ash is crucial for improving pyrolysis efficiency and product quality.

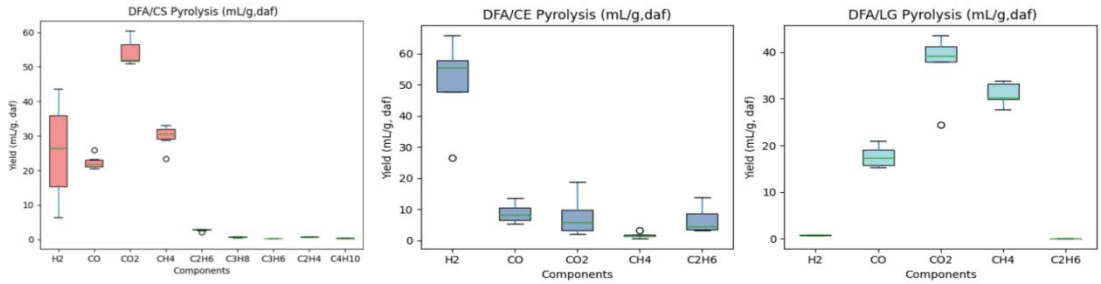
4.2 Model establishment and solution of problem 2

After preprocessing the data in Annex II and finding no outliers or missing values, python was used to draw scatter plots of pyrolysis gas yield versus mixing ratio under each pyrolysis combination, to visualize the effect of different pyrolysis combination ratios on the pyrolysis gas yield in each group, and analyze and interpret the results based on the visualization.

4.2.1 Data Preparation

4.2.1.1 Outlier checking

A box plot is a type of statistical chart commonly used to display information on the dispersion of a set of data, which facilitates an accurate and stable depiction of the discrete distribution of data and data processing. We made the following box-line plots using the yield of pyrolysis gas from each pyrolysis combination at different ratios as sample data:



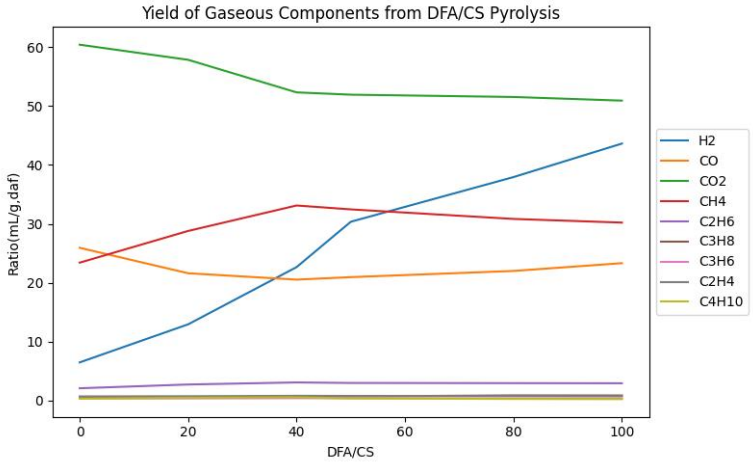
Box-line plots of pyrolysis gas yield at different ratios

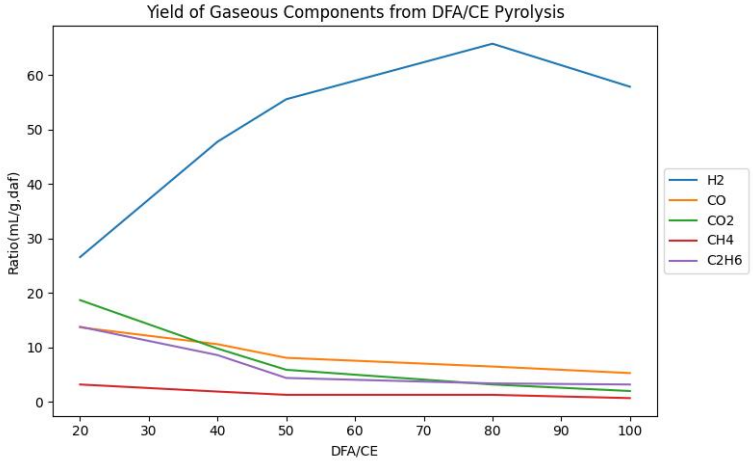
Off-mean values exceeding 4 times the standard deviation of the coefficient factor are judged as abnormal sizes, and the number of abnormal values is found to be small by observing the length of the box and line segments of each party, and the effect of the abnormal values can be ignored considering the small sample size.

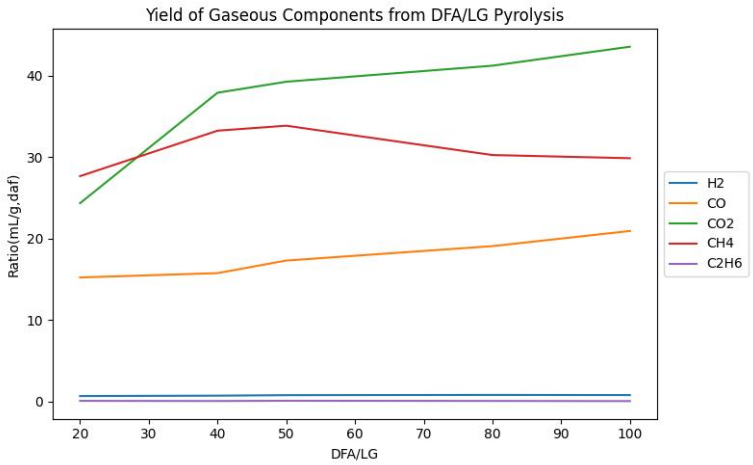
4.2.1.1 Data visualization

A folding graph of pyrolysis gas yield versus mixing ratio for each pyrolysis

combination was plotted using python as follows:







4.2.2 Conclusion and analysis

4.2.2.1 Effect of mixing ratio on pyrolysis gas yield in DFA/CS groups

(1) H_2 yield:

The H_2 yield showed an overall increasing trend with increasing DFA/LG ratio, but slightly decreased at 100/100. This suggests that the hydrogen yield increases and

then decreases as the ratio of desulfurized ash (DFA) to lignin (LG) increases, probably because the peak of catalytic efficiency is reached.

(2) CO yield:

CO yield showed a continuous upward trend with increasing DFA/LG ratio. This may indicate that the increase in desulfurized ash favors the production of carbon monoxide, which may be related to the promotion of oxidation reactions.

(3) CO_2 yield:

CO_2 yield showed a significant increasing trend with the increase of DFA/LG ratio. This may imply that the desulfurization ash promoted the complete oxidation of the lignin fraction.

(4) CH_4 yield:

The CH4 yield peaked at a DFA/LG ratio of 50/100 and then decreased. This may indicate that methane was most efficiently produced at moderate DFA concentrations, while further methane conversion may have been promoted subsequently due to the catalytic effect of DFA.

(5) C_2H_6 yield:

The C_2H_6 yield showed an overall decreasing trend, although the change was not significant. This may indicate that ethane formation is less affected by desulfurization ash, or that its production pathway is inhibited at high desulfurization ash concentrations.

Overall, these data suggest that the addition of desulfurization ash changes the composition of the gas products, particularly for increases in hydrogen and carbon monoxide and changes in the yields of methane and carbon dioxide. This may be related to the effect of desulfurization ash on pyrolytic reaction pathways, such as facilitating cracking or oxidizing processes, which alters the distribution of gas products.

4.2.2.2 Effect of mixing ratio on pyrolysis gas yield in DFA/CE groups

(1) H2 yield:

H_2 yields generally tend to increase with increasing DFA/CE ratio, but decrease slightly at 100/100. This may indicate that the catalyst is more efficient at promoting hydrogen production at higher concentrations, but at some point there is a slight decrease in efficiency, perhaps due to reaching equilibrium or other limiting conditions of the reaction.

(2) CO yield

CO yield continued to decrease with increasing DFA/CE ratio. This suggests that the addition of desulfurized ash may have inhibited the formation of carbon monoxide or facilitated the conversion of carbon monoxide to carbon dioxide.

(3) CO_2 yield:

CO2 yield decreased significantly with increasing DFA/CE ratio, which may be due to the presence of FGD ash promoting partial oxidation of carbon, resulting in the

production of more CO and H_2 than CO_2.

(4) CH_4 yield:

CH4 yield decreased significantly with increasing DFA/CE ratio, which may imply that the desulfurization ash promotes the decomposition of methane or changes the chemical pathway of methane production.

(5) C_2H_6 yield:

Similar to methane, ethane yield also decreased significantly with increasing DFA/CE ratio, suggesting that the catalyst may have influenced the stability of heavier hydrocarbons or their generation pathways.

Overall, the trend of increasing hydrogen yield and decreasing other gas yields with increasing DFA/CE ratios may indicate that FGD ash acts as a catalyst that facilitates the generation of hydrogen and influences the chemical reaction equilibrium during pyrolysis. The catalytic effect of the desulfurization ash may promote the production of gases with lower carbon content (e.g., hydrogen and carbon monoxide) while inhibiting the formation of heavy carbon gases (e.g., methane and ethane).

4.2.2.3 Effect of mixing ratio on pyrolysis gas yield in DFA/LG groups

(1) H2 yield:

The H_2 yield showed an overall increasing trend with increasing DFA/LG ratio, but slightly decreased at 100/100. This suggests that as the ratio of desulfurized ash (DFA) to lignin (LG) increases, the hydrogen production increases and then decreases, probably because the peak of catalytic efficiency is reached.

(2) CO yield:

CO yield showed a continuous upward trend with increasing DFA/LG ratio. This may indicate that the increase in desulfurized ash favors the production of carbon monoxide, which may be related to the promotion of the oxidation reaction.

(3) CO_2 yield:

CO_2 yield showed a significant upward trend with the increase in DFA/LG ratio. This may imply that the desulfurization ash promoted the complete oxidation of the lignin fraction.

(4) CH_4 yield:

The CH_4 yield peaked at a DFA/LG ratio of 50/100 and then decreased. This may indicate that methane was most efficiently produced at moderate concentrations of desulfurization ash, and that further conversion of methane may have been promoted subsequently due to the catalytic effect of desulfurization ash.

(5) C_2H_6 yield

The C_2H_6 yield showed an overall decreasing trend, albeit with little change. This may indicate that ethane formation is less affected by desulfurization ash, or that its production pathway is inhibited at high desulfurization ash concentrations.

Overall, these data suggest that the addition of FGD ash changes the composition of the gas products, particularly for increases in hydrogen and carbon monoxide and

changes in the yields of methane and carbon dioxide. This may be related to the effect of FGD ash on pyrolytic reaction pathways, such as facilitating cracking or oxidizing processes, thus changing the distribution of gas products.

4.2.2.4 Conclusion

By analyzing the gas component yields of the three pyrolysis combinations, we can see that FGD fly ash has different effects on different components. In the DFA/CS combination, the FGD fly ash significantly affects the gas component yields, especially the hydrogen and carbon monoxide yields. In the DFA/CE and DFA/LG combinations, the effect of FGD fly ash on gas yields was also significant, especially in increasing hydrogen yields. These findings illustrate the different effects of desulfurized fly ash as a catalyst in different pyrolysis combinations, which are potentially valuable for improving the yield of specific gas components.

4.3 Model establishment and solution of problem 3

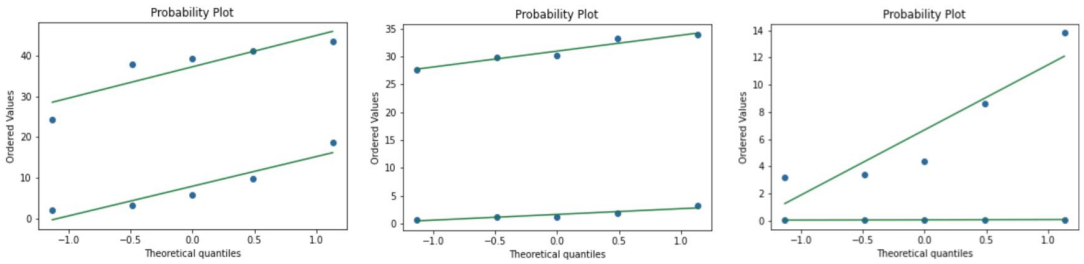
In order to analyze the variability of the yield of CE and LG pyrolysis products and the yield of pyrolysis gas components under the catalytic action of the same proportion of desulfurization ash, firstly, the normal distribution test was carried out by plotting Q-Q diagrams and performing the Shapiro-Wilk test. Second, the variability of the yield of CE and LG pyrolysis products and the yield of pyrolysis gas components before the catalytic effect of the same proportion of desulfurization ash was visualized by plotting stacked diagrams. Finally, the variability was analyzed by t-test.

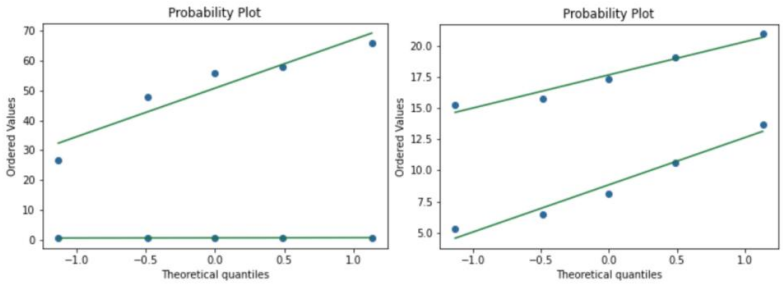
4.3.1 Normal distribution test

From the data preprocessing of the first question, it can be seen that the pyrolysis product data all conform to normal distribution, so only the pyrolysis gas rate data were tested for normal distribution.

4.3.1.1 Test for normal distribution based on Q-Q plots

Q-Q plots of CE and LG pyrolysis gas yields using python are shown below:





Q-Q plots of CE and LG pyrolysis gas yields

The Q-Q graph shows that all the data are concentrated around a straight line, indicating that the data conforms to a normal distribution.

4.3.1.2 S-W test

The empirical distribution function F_n for n independent and identically distributed ordered observations X is defined as:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[-\infty,X]}(xi)$$

$I_{[-\infty,X]}(X_i)$ is an exponential function, which is equal to 1 if $x_i \leq x$ and it is equal to 0 otherwise.The Kolmogorov-Smirnov statistic for a given cumulative distribution function F(x) is:

$$D_n = \sup|F_n(x) - F(x)|$$

A test that compares a frequency distribution f(x) with a theoretical distribution g(x) or two distributions of observations by the KS test. The original hypothesis H_0: The total coefficients of the original adverse reactions follow a normal distribution.

D is the maximum value of the difference between F_0x and F_nx. At the 95% confidence level, when the actual observation D > D(n,α) (as indicated by the significance value of the output being less than 0.05), then the hypothesis is rejected. H_0, otherwise the H_0 hypothesis is accepted.

We examined the CE and LG pyrolysis gas yield data with the following results:

Component	Passed S-W Test	p-value
H_2	Yes	7.06E-05
CO	Yes	0.001362
$C0_2$	Yes	0.000188
CH_4	Yes	9.62E-09
C_2H_6	Yes	0.011604

From the results in the table, it can be seen that the pyrolysis gas yield data follows a normal distribution under both CE and LG.

4.3.1.2 T-test

Since the pyrolysis product yield and pyrolysis gas yield data of CE and LG passed the normal distribution test, we performed an independent samples t-test using python, and obtained the following results.

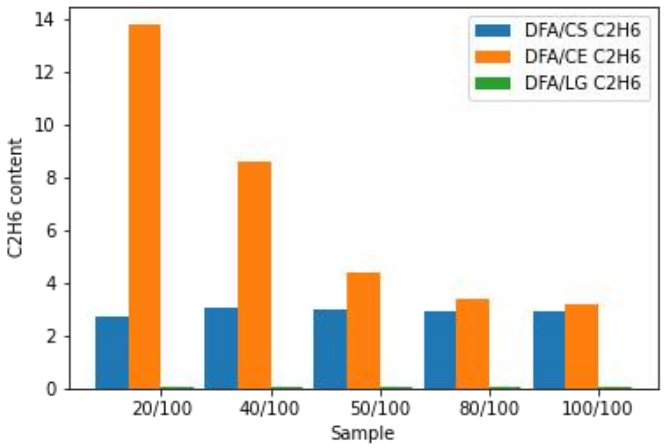
Component	t-statistic	p-value
Tar yield	17.60781	6E-11
Water yield	-0.73117	0.476739
Char yield	-78.39	6.57E-20
Syngas yield	6.03453	3.07E-05

Component	t-statistic	p-value
H2	7.480264	7.06E-05
CO	-4.79655	0.001362
CO2	-6.4989	0.000188
CH4	-24.0205	9.62E-09

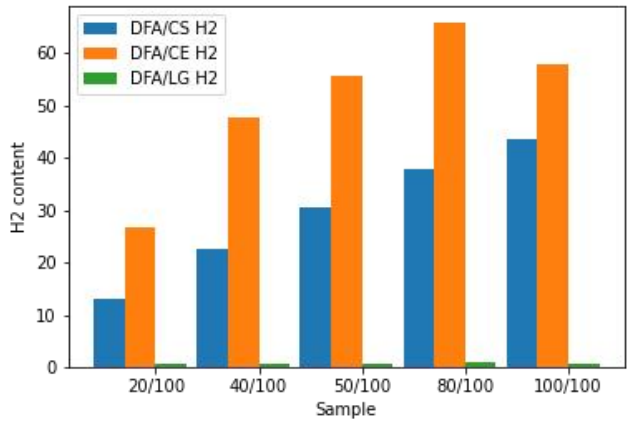
As shown in the above table, the p-values are less than 0.05, and there is a significant difference between the yields of CE and LG pyrolysis products as well as pyrolysis gas fractions under the catalytic effect of the same proportion of FGD ash.

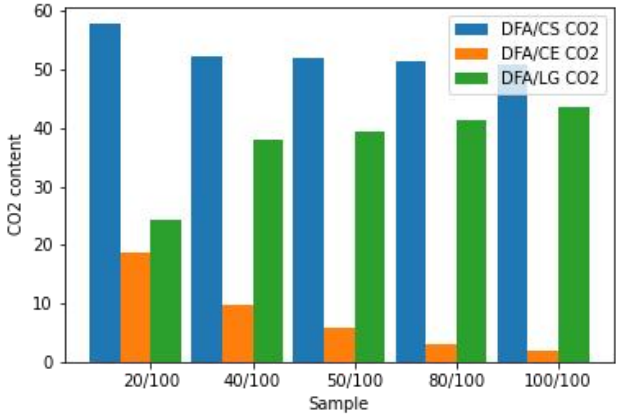
4.4 Model establishment and solution of problem 4

According to the modeling solving results of problem 3, the desulfurization ash has a significant effect on the gas components produced by the pyrolysis of cotton rod, CE and LG. To demonstrate these differences more intuitively and explore the mechanism of the catalytic pyrolysis reaction, we drew grouped histograms of pyrolysis gas component yield under desulfurization ash / cotton bar, desulfurization ash / CE, and desulfurization ash / LG mixing ratios.



As can be observed from Figure (a), the resulting H2 is mainly from the decomposition of CE. Desulphurization ash significantly promoted the chemical reaction of hydrogen produced by CE pyrolysis. In general, breakage of ether bonds and carboxyl functional groups is the main source of CO and CO2 during pyrolysis of lignocellulosic biomass.





It can be seen from Figure (b), (c): in the cotton rod pyrolysis process, CE and LG are important sources of CO and CO2; with the increase of desulfurization ash content, its catalysis inhibits the reaction of CE decomposition to form CO and CO2, and promotes the chemical reaction of LG pyrolysis to form CO and CO2.

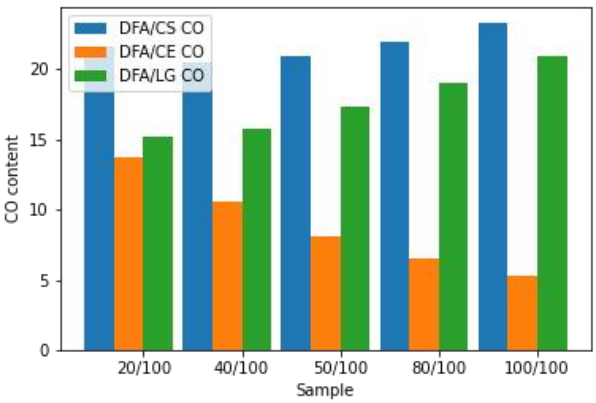
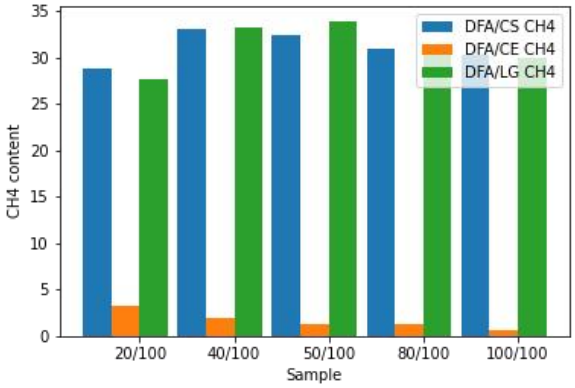


Figure (d) shows that the CH4 in desulphization ash / cotton rod pyrolysis gas mainly comes from the fracture of LG-rich methoxy side chain, and the CH4

produced by the thermal decomposition of CE is very little, and the desulfurization ash has obvious inhibitory effect on the process.



As can be seen from Figure Figure (e), the thermal decomposition of LG has a limited ability to produce C2H6. With the increase of desulfurization ash content, the production of C2H6 generated by CE pyrolysis decreased significantly, which indicates that the promotion of CE pyrolysis inhibited the process of ring opening and the production of C2H6 radical in CE structural units. Desulphurization ash had almost no effect on rod pyrolysis to generate C2H6, which indicated that CE and LG components during rod pyrolysis inhibited the free radical reaction of C2H6.

In conclusion, the chemical reaction paths of the model compounds of CE and LG can be inferred as follows:

(1) CE pyrolysis:

CE-> H2 (large) + CO + CO2 + CH4 (small) + C2H6 (large) + aliphatic hydrocarbons, aromatic hydrocarbons and heteroatom compounds + C (coke) + H2O + other gases

(2) LG pyrolysis:

LG-> H2 (small) + CO + CO2 + CH4 (large) + C2H6 (small) + aliphatic hydrocarbons, aromatic hydrocarbons and heteroatomic compounds + C (coke) + H2O + other gases

4.5 Model establishment and solution of problem 5

4.5.1 Support vector machine based prediction models

We utilized the epsilon-SVR technique within the SVR learning method to predict changes in pyrolysis reaction product yields. Our aim was to establish a functional relationship between various ratios of DFA (Desulfurization Ash) and CS (Cotton Stalks) and the resulting product yields. In our study, we adopted the support vector back as our modeling tool, employing a multi-component model for the pyrolysis reaction, encompassing 'Tar', 'Water yield', 'Water yield', 'Char yield', and 'Syngas yield' for individual prediction. Specifically, 'Char yield' and 'Syngas yield', both being integral components, were separately predicted to ascertain the most optimal model parameters.

First, we take the ratio of DFA/CS as a variable i and denote the index value of

pyrolysis reaction product yield for this ratio by x. The task of SVR is to find a function x = f(t), i.e., a function of the relationship between the time t and the yield of the pyrolysis reaction product x. We expect that the SVR algorithm will be able to predict the optimal model parameters for each of the four components. We expect the SVR algorithm to make the error between the predicted product yield x and the true indicator value x within an acceptable threshold.

Our functional expression is:

$$\chi = f(t) = \omega \cdot \phi(t) + y$$

In this formulation, ϕ (t) is the kernel function of the eigenvectors, w is the weight vector, and y is the bias term. We use SVR to fit the relationship between the yield of pyrolysis reaction products and time, and find the optimal function to predict the yield of pyrolysis reaction products by adjusting the kernel function and model parameters.

We opted for three kernel functions: linear, rbf (radial basis function), and poly (polynomial kernel function) to establish the functional relationship. Among these, the RBF kernel, known for its efficiency in handling nonlinear relationships, proved suitable for pyrolysis yield prediction.

To determine the best model, we conducted an exhaustive search through Grid Search, exploring various parameter combinations within a defined parameter grid. This involved evaluating 108 different parameter combinations through 5-fold cross-validation, resulting in a total of 540 fits. Eventually, the optimal hyperparameters were obtained through this process.

After the training is completed, we will evaluate the built SVR model using the test set data to calculate its prediction performance of product yield under different DFA/CS ratios, and we used two metrics：

(1) Mean Squared Error：

Mean square error is a common metric for assessing the prediction results of regression models. For k-fold cross-validation, the MSE is calculated as follows:：

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

where n is the number of samples, is the true value, and is the model predicted value.

(2) R-squared score：

The R^2 score is an indicator of the goodness of fit of the regression model. For k-fold cross-validation, the R^2 score is calculated as follows:：

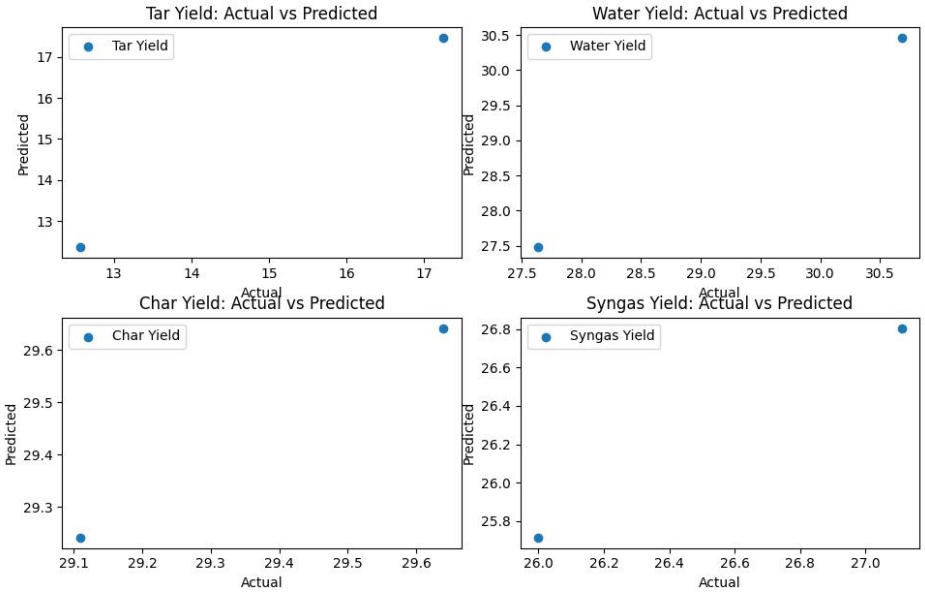
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

Where is the mean of the true values, and the other symbols have the same meaning as above.

	Mean Squared Error	R^2 Score
Tar yield	0.04496	0.99178
Water yield	1.46651	0.98666

Char yield	0.00867	0.87649
Syngas yield	0.08695	0.71771
Average	0.04475	0.89234

The SVM models achieved satisfactory performance in predicting the yields of the four products, boasting an average R^2 score of 0.8923. Additionally, the average mean squared error (MSE) stood at 0.04475, signifying minimal prediction errors, as depicted in the figure.



From the scatterplot, we can visualize the actual values versus the predicted values, and the model fit is outstanding.

4.5.2 Predictive modeling based on ridge regression

Although support vector machine (SVM) is a powerful machine learning algorithm, it has high computational complexity, complicated parameter tuning process, and limited ability to handle limited data when dealing with multivariate regression problems and nonlinear problems. Therefore, we chose to use ridge regression and polynomial regression, which are relatively easy to implement and can effectively deal with overfitting and nonlinear problems, to predict pyrolysis gas production.

Ridge regression is a modified least squares method that controls the complexity of the model as well as prevents overfitting phenomena by introducing an L2 regularization term. Specifically, ridge regression works by adding an L2 regularization term to the traditional linear regression loss function, namely:

$$L(w) = \sum(y_i - w_1x_1 - w_2x_2 - \cdots - w_nx_n - b_n)^2 + \lambda\sum w_i^2$$

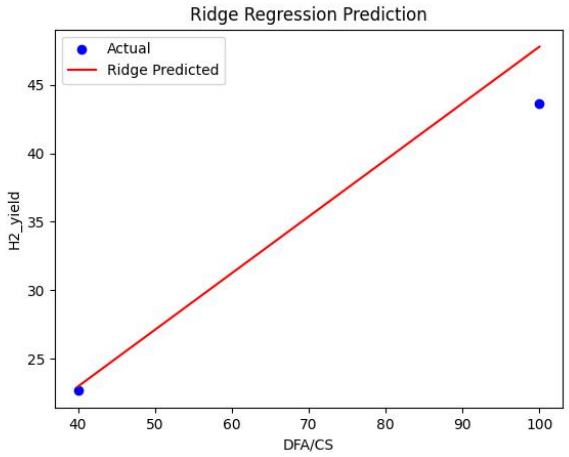
For pyrolysis gas yield prediction, we utilized the composition ratio of DFA and CS as input features. The data was split, with 80% used for training the model and the remaining 20% for testing its generalization ability. After training, the ridge

regression model predicted the test set results, which were then compared against the true values. Similar to the earlier evaluation method, we assessed the model's performance using mean square error and the coefficient of determination.

The results are shown in Figure below. The mean square error is 8.79, which is close to 0, proving that the model has some prediction accuracy. The coefficient of determination (R^2 Score) is 0.920, which is close to 1, and the model fits well.

	Mean Squared Error	R^2 Score
Ridge Regression	8.790381	0.92004

From the scatterplot , we can visualize how the actual values compare with the predicted values, and the model fit is more excellent.



4.5.3 Forecasting model based on polynomial regression

Through preliminary analysis and visualization of the data, we found that the relationship between product yield and pyrolysis reaction combination ratios is not a simple linear relationship and appears to exhibit a quadratic or curvilinear relationship. To improve the accuracy of the predictive model, we used a 2nd polynomial regression model and tried to find a more complex polynomial equation to fit the data, thus improving the predictive power.

First, the feature matrix is constructed: the higher powers of the features are computed based on the number of quadratic polynomials chosen. The feature matrix chosen for this model is the pyrolysis reaction combination ratios.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 20 \\ \vdots \\ 100 \end{bmatrix}$$

The processed feature matrix is:

$$X_{poly} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^n \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 20 & 20^2 \\ \vdots & \vdots & \vdots \\ 1 & 100 & 100^2 \end{bmatrix}$$

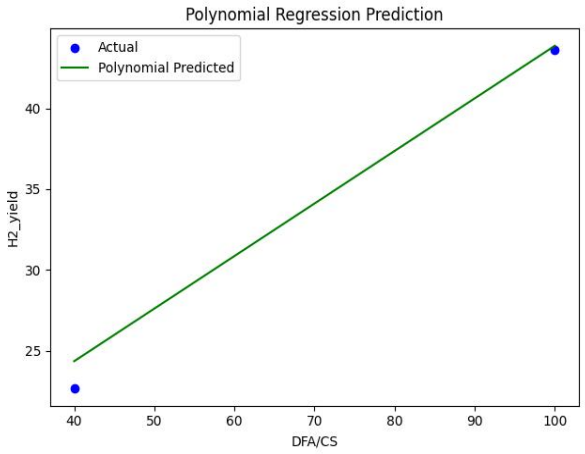
We apply the least squares method to train a linear regression model using the processed feature matrix X_poly and the dependent variable Y. The model will try to find the best weight vector W to minimize the prediction error. The model will try to find the optimal weight vector W in order to minimize the prediction error and finally get the prediction function:

$$Y = -0.0014936351706039148 * X^2 + 0.5345413385826874 * X^2 + 0.0$$

The mean square error (MSE) as well as the coefficient of determination (R^2) evaluation models are shown in below Table :

	Mean Squared Error	R^2 Score
Polynomial Regression	1.466519	0.98666

From the scatterplot , we can visualize the comparison between the actual values and the predicted values, and the predicted values are uniformly distributed around a straight line along with the actual values, the predictive ability of the model is strong, and the model fitting is excellent.



6.Strengths and Weakness

6.1 Strengths

Based on limited data, we consider as many factors as possible to perfect our computational model.

1. Polynomial regression: It captures and depicts the non-linear relationship between dependent and independent variables, can control the complexity of the model by adjusting the order of the polynomials, and can better accommodate curvilinear data. Also, it is more explanatory than other nonlinear models and can be used in

conjunction with linear regression methods through feature transformation.
2. Support Vector Machine: It has strong generalization ability, can handle high dimensional data and multiple types of feature data, and is suitable for small sample datasets. It uses kernel tricks and can handle nonlinear problems while having good stability and robustness and avoiding overfitting.

6.2 Weakness

1. Assumption limitation: the establishment of mathematical models relies on certain assumptions, and there may be differences between the assumptions and the actual situation, resulting in limitations on the accuracy of the model.
2. Data requirements: mathematical models require a large amount of experimental data for parameter estimation and validation, and if the data are insufficient or inaccurate, the reliability of the model will be affected.

References

[1] 杜启,唐初阳,安新元等.基于模型化合物的脱硫灰催化提质生物质热解气研究[J].可再生能源,2023,41(06):
[2] 宋春财,胡浩权.秸秆及其主要组分的催化热解及动力学研究[J].煤炭转化,2003(03):91-97.
[3] 杜海清. 木质类生物质催化热解动力学研究[D].黑龙江大学,2009.
[4] 陈明强,齐雪宜,王君等.棉秆催化热解特性及动力学建模研究[J].燃料化学学报,2011,39(08):
[5] 谭洪.生物质热裂解机理试验研究[D].浙江大学,2005.
[6] 宋成芳.生物质催化热解炭化的试验研究与机理分析[D].浙江工业大学,2013.