| Problem Chosen | 2023 ShuWei Cup Summary Sheet | Team Control Number |
|---|---|---|
| B | | 2023111421466 |

# Catalytic Pyrolysis of Cotton Straw Based on Weighted Prediction Model


## Abstract

In recent years, cotton stalk is an important source of biomass energy because of the increasing demand for renewable energy worldwide. On this basis, the study on the high efficiency utilization and sustainable development of cotton straw resources was carried out combining with the pyrolysis products, gas and gas.

Preprocess the data given in the question. First of all, the data given by a rough analysis, scatter plot. By doing q-q graph and Shapiro-Wilk test on the given data set, we can judge whether the given data set accords with normal distribution. For this sequence, we use the 3-Sigma criterion to judge outliers, and relate the results with the actual situation. The results are normal distribution.

The first problem is to analyze the relationship between the yield of pyrolysis products and the mixture ratio of pyrolysis combinations. First of all, under the condition of different proportions, the graph of thermovalent products is drawn and discussed. At the same time, Pearson correlation analysis method was used to analyze the correlation of different ratios of product yield. The important role of desulfurized lime as catalyst was confirmed by comparing the reaction with and without desulfurized lime. Engage in analytical work.

The second problem is that each combination of the three cracking methods is plotted in Matlab. Based on the images, the differences of pyrolysis products with different ratios were discussed to explain the differences of pyrolysis gases. In question 3, the yields of CE and LG pyrolysis products are very different, as is the production of pyrolysis gases. Significant analysis was performed at the same mixing ratio.

In problem 4, we use the results of the first problem and the second, and use a graph to determine the mechanism. On this basis, the corresponding linear or nonlinear relationship between products is constructed, and the corresponding reaction mechanism model is established. On this basis, the corresponding dynamic model is established. Aiming at the fifth problem, this paper constructs a regression model by using the fourth problem and forecasts it. On this basis, grey prediction is combined with LSTM to improve the accuracy of the model. Finally, the optimal weight of each index is obtained by calculating the weight of each index. Finally, the prediction results are given. Finally, the prediction effect of the model is evaluated.

Keywords: data preprocessing, correlation analysis, significance analysis, forecasting mode

# Contents

# 1. Introduction

## 1.1 Background

Under the background of the increasing demand for new energy in the world, biomass as a mature renewable energy has been attached importance to. Cotton stalk is a kind of agricultural waste, which contains a lot of cellulose and lignin, is a very valuable biomass resources. Pyrolysis of cotton stalk is a renewable resource, but the quality and yield of pyrolysis products are affected by many factors such as pyrolysis temperature and catalyst. Therefore, it is of great theoretical and practical significance to study the mechanism and characteristics, catalytic mechanism and influence mechanism of cotton stalk pyrolysis products (see the Appendix for the definition of terms).

Using cotton straw as raw material and simulated compound as raw material, pyrolysis treatment was carried out and optimized. On this basis, the mechanism of desulphurization ash on the pyrolysis product of cotton stalk was studied by pyrolysis experiment. At the same time, the controllability and stability of the system and its contribution to the pyrolysis of cotton straw should also be taken into account in the selection of model compounds.

## 1.2 Restatement

(1) For each pyrolysis combination in Annex I, analyze the relationship between the yield of pyrolysis products (tar, water, coke residue, syngas) and the mixture ratio of the corresponding pyrolysis combination, and explain whether the desulfurized ash as a catalyst plays an important role in promoting pyrolysis of cotton straw, cellulose and lignin?

(2) According to Annex II, for each of the three pyrolysis combinations, the effect of the mixture ratio of the pyrolysis combination on the yield was discussed and the effect of each pyrolysis gas group was explained by making corresponding images.

(3) Are there significant differences in the yields of CE and LG pyrolysis products and pyrolysis gas components catalyzed by the same proportion of FGD ash?
(4) How to establish the catalytic reaction mechanism model of desulfurized ash of CE, LG and other model compounds, and establish the reaction kinetics model for analysis?
(5) Use mathematical models or artificial intelligence learning methods to produce predictions of the yield or quantity of pyrolysis products under limited data conditions.

# 2. Problem analysis

## 2.1 Data analysis

Preprocess the data given in the question. First, a rough analysis of the data is made, and then a scatter plot is drawn. By doing q-q graph and Shapiro-Wilk test on the given data set, we can judge whether the given data set accords with normal distribution. For this sequence, we use the 3-Sigma criterion to judge outliers, and relate the results with the actual situation. The results are normal distribution. Analysis of question one

The first problem is to analyze the relationship between the yield of pyrolysis products and the mixture ratio of pyrolysis combinations. First of all, under the condition of different proportions, the graph of thermovalent products is drawn and discussed. At the same time, Pearson correlation analysis method was used to analyze the correlation of different ratios of product yield. The important role of desulfurized lime as catalyst was confirmed by comparing the reaction with and without desulfurized lime. Do an analysis.

## 2.2 Analysis of question two and three

The second problem is that each combination of the three cracking methods is plotted in Matlab. Based on the images, the differences of pyrolysis products with different ratios were discussed to explain the differences of pyrolysis gases. In question

3, the yields of CE and LG pyrolysis products are very different, as is the production of pyrolysis gases. For the same mixing ratio, we do a significant analysis.

### 2.3 Analysis question four and five

In problem 4, we use the results of the first problem and the second and use a graph to determine the mechanism. On this basis, the corresponding linear or nonlinear relationship between products is constructed, and the corresponding reaction mechanism model is established. On this basis, the corresponding dynamic model is established. Aiming at the fifth problem, this paper constructs a regression model by using the fourth problem and forecasts it. On this basis, grey prediction is combined with LSTM to improve the accuracy of the model. Finally, the optimal weight of each index is obtained by calculating the weight of each index. Finally, the prediction results are given. Finally, the prediction effect of the model is evaluated.

## 3. Symbol and Assumptions

### 3.1 Symbol Description

The important notations used to facilitate our modeling and resolution are as follows:

| notation | clarification |
|---|---|
| $\Delta y_t$ | is the first order difference of the |
| $p$ | weights obtained from the entropy weighting method calculation |
| $f\left[x_0, x_1, ..., x_n\right]$ | lag order |
| $x_i$ | The $i_{th}$ missing term |
| n $n$ | Number of missing terms |
| $Cov(x, y)$ | Sample covariance |
| $S_x$ | Sample standard deviation |

| | |
|---|---|
| $r_{xy}$ | Sample Pearson's correlation coefficient |
| $y_t, y_{t-1}$ | is the value of the variable at moments $t$ and $t-1$ |
| $\xi$ | obeys a normal distribution |
| $\varepsilon_t$ | White noise series |
| $\beta_1, \beta_2, \cdots, \beta_4$ | Coefficients of the respective variables |
| $y_{预测}$ | Predictive data |
| $\mathrm{y}_{真实}$ | Real data |
| $\mathrm{forecast}(t)$ | Predicted value |
| $\mathrm{a}ctual(t)$ | True value |

# 4. Model

For better modeling, we make the following assumptions:

1.Data completeness: It is assumed that the information provided is complete and there are no missing values or error records.

2.Data representativeness: It is assumed that the established data can better reflect the pyrolysis process of cotton stalks under different conditions.

3.Consistency of catalytic effect: It is assumed that the desulfurized ash as the catalyst will have the same catalytic effect on each sample and will not be affected by the characteristics of the samples (e.g., origin, variety, etc.).

4.Stability of reaction conditions: In order to ensure the comparability of results, it is assumed that all pyrolysis tests are conducted under the same conditions.

5.Assumption of independence: It is assumed that the data points are independent of each other, i.e., the characteristics and results of one data point are not related to other data points.

## 4.1 Data pre-processing

The Kolmogorov-Smirnov test is a nonparametric test used to test that a set of data satisfies a certain probability distribution, especially for conforming to a normal distribution. Its basic idea is to determine whether these samples satisfy this theoretical distribution by comparing the statistical distribution function of the samples with the statistical distribution function and deriving the maximum difference between the two. When the maximum difference is less than a certain threshold, we regard it as meeting this theoretical distribution. In the case where the difference between the two is small, it is presumed that the theoretical distribution of this sample is known to itself. Usually, the distribution under the original assumption is one-dimensional continuous (e.g. normal, uniform, exponential, etc.) and discrete (e.g. Poisson, etc.). That is: a set X follows a kind of one-dimensional continuous distribution F::

$$Z = \sqrt{n} max_i (|F_n(x_{i-1}) - F(x_i)|, |F_n(x_i) - F(x_i)|)$$

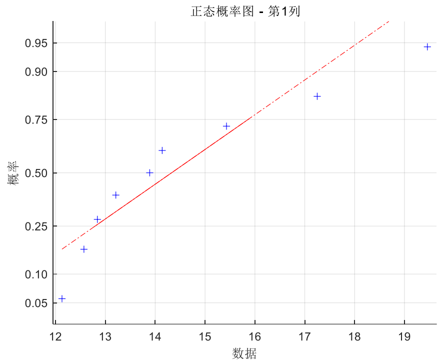
The H-True----- distribution converges to the Kolmogorov-Smirnov distribution. That is, when drawing samples from a one-dimensional continuous distribution F：

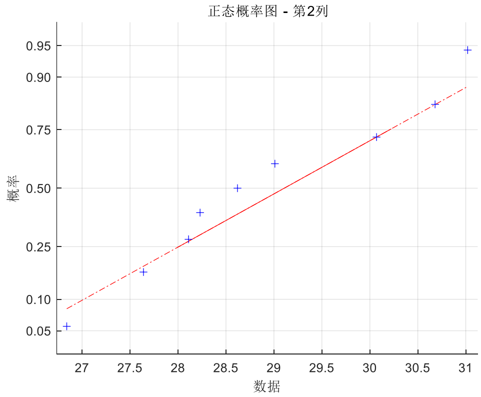
$$Z \to K = \sup\ |B(F(x))|$$

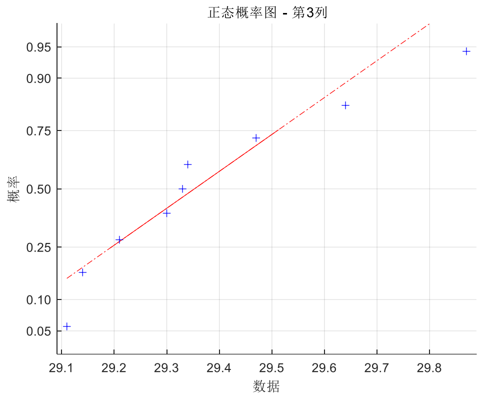
Note:The distribution of the random variable K does not depend on F when F is continuously distributed.

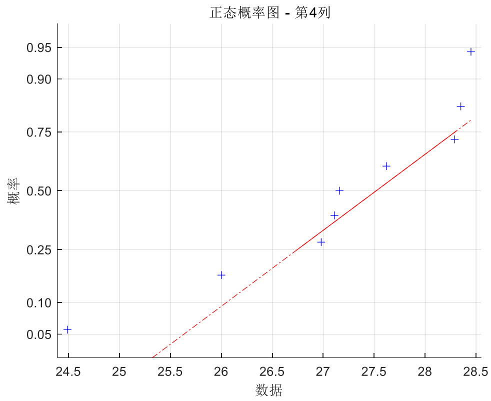
The result of the KS test is usually a p-value, and if the p-value is less than the significance level (typically 0.05), the original hypothesis is rejected, i.e., the two samples are considered to be from different distributions.

We need to determine the distribution of the shipments by plotting the Q-Q graph using SPSS and by performing the Kolmogorov-Smirnov test, the results of the Kolmogorov-Smirnov test analysis are shown in the figure below:




After analyzing the information, it was determined that most of the information conformed to the normal distribution. So we introduce the principle. The criterion for determining outliers, where a given set of information conforms to the normative distribution, 99.7% of the information falls within the range of the mean three standard deviations, see Fig. 3. If there is a value that is outside of the three standard deviations, then that value is considered to be an outlier. The formula for determining outliers is as follows.

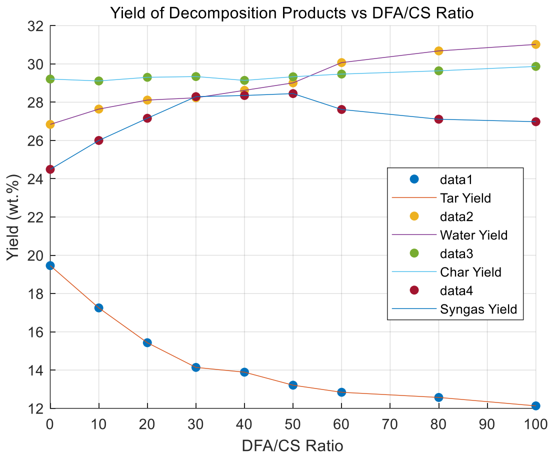
$$P(|x-\mu| > 3\sigma) \leq 0.003$$

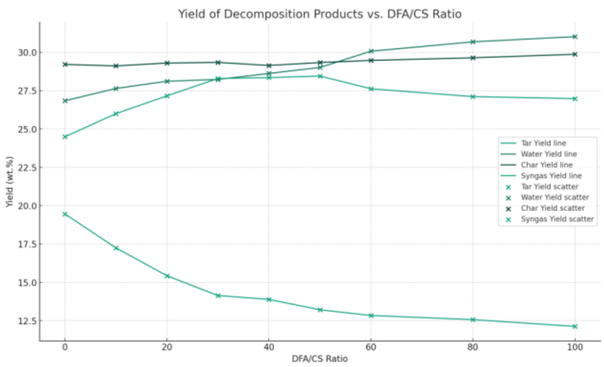
Where an item of information in the collected data set is the mean of all the information under that index and the standard deviation of all the information under that index.Finally, the given data set is made free of anomalies by manual judgment and by considering realistic conditions. Subsequent modeling can be done directly

## 4.1 Problem 1 Modeling and Solving

Only the results of the data in the group of DFA/CS are shown, and the results related to DFA/CE and DFA/LG are put in the supporting materials for presentation.

For the data set DFA/CS, a scatterplot was plotted using python as follows




The results show that there is a significant negative correlation between tar yield and coal fiber content in coal, with a correlation coefficient around -0.88. As the DFA/CS ratio increases, the tar yield decreases gradually; there is a strong positive correlation between the water yield and the DFA/CS ratio (the correlation coefficient is around 0.98). As the DFA/CS ratio increased, the yield increased;

The results showed that there was a significant positive correlation between DFA/CS ratio and coke slag yield with a correlation coefficient around 0.90. As the DFA/CS ratio increases, the coke slag yield increases. At present, the correlation between syngas yield and DFA/CS is poor (correlation only 0.42) except for the DFA/CS ratio.

When the DFA/CS ratio increased, the tar yield decreased significantly, but the water yield and coke slag yield increased, while the syngas yield did not change significantly. This is due to the reaction mechanism of each component during pyrolysis. When the DFA/CS ratio increased, some components were transformed into water or tar slag, but the transformation of tar was inhibited.

It can be seen from the scatter plot that most of the indices are correlated with each other. Here, the correlation analysis model is constructed to evaluate the correlation between the evaluation indexes by introducing the person correlation. Sample co-correlation:

$$Cov(x,y)=\frac{\sum_{i=1}^{n}(X_I-X)(Y_i-Y)}{n-1}$$

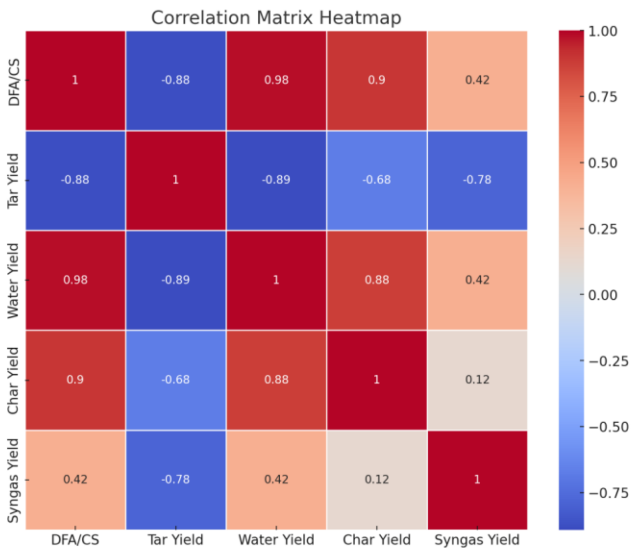
Sample standard deviation:

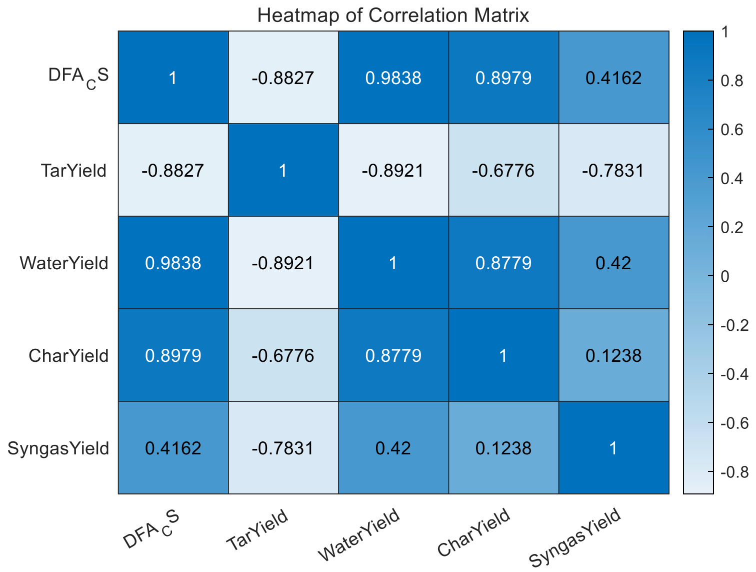
$$S_x=\sqrt{\frac{\sum_{i=1}^{n}(X_i-X)^2}{n-1}}$$

Sample Pearson correlation coefficient:

$$r_{xy}=\frac{Cov(X,Y)}{S_X S_Y}$$

The correlation coefficient method was used to analyze the correlation between the products in the DFA/CS portfolio. To enhance the visualization of the calculation results, the matrix was plotted using python in this paper:



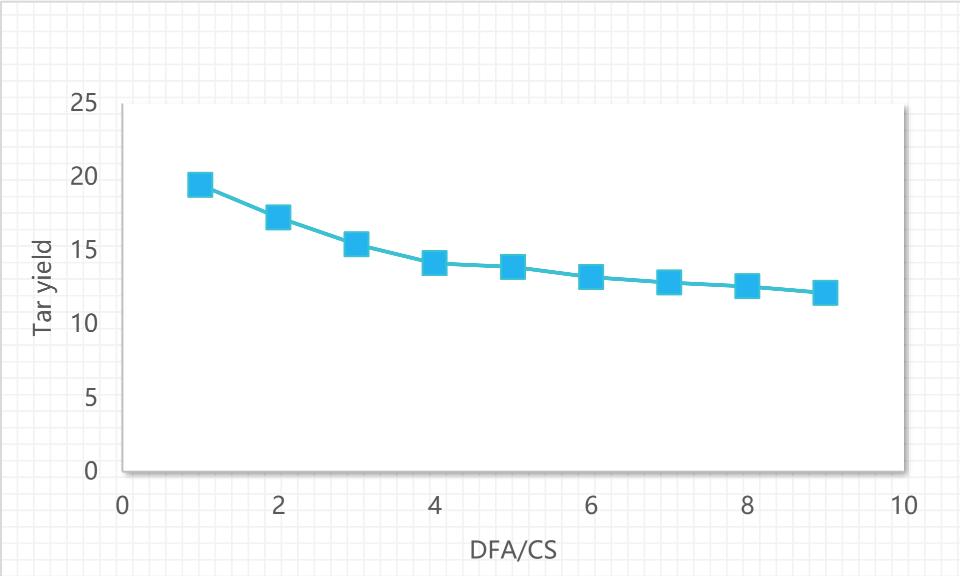



The thermodynamic curves show a good correlation between the pyrolysis product yield and the DFA/CS ratio. As shown in Fig:

The results show that the tar yield decreases significantly when the DFA/CS ratio increases. The results show that the water production rate increases as the DFA/CS ratio increases and the yield increases. In addition, as the DFA/CS ratio increases, the coke slag yield also shows a good correlation, and its correlation reaches 0.9. The correlation between syngas yield and DFA/CS is not strong, and the correlation coefficient is only 0.0 1, which suggests that other factors have a great influence on the syngas yield.

In addition, the tar yield showed a significant negative correlation with both water yield and coke residue yield. It indicates that in the pyrolysis process, part of the tar is converted into water and coke residue.

Considering that there is a good correlation between the indicators, we chose DFA/CS and ar field as the research object and carried out regression analysis on them. The measurement results were used as a basis to judge the reaction mechanism. Firstly, a graph was plotted between these two variables as follows




As can be seen from the line graph, there is a negative correlation or quadratic polynomial link between them. We used matlab to write the program instead of generic. Both linear and quadratic polynomial fitting methods were used, and the fit of the two methods was compared, and the one with the better fit was chosen as the final correlation.

Table 1: Mechanism determination pseudo-code

| **% Initialize DFA/CS ratio and tar yield data** |
|---|
| Define DFA_CS as [0, 10, 20, 30, 40, 50, 60, 80, 100] |
| Define TarYield as [19.46, 17.25, 15.43, 14.14, 13.89, 13.21, 12.84, 12.57, 12.13] |
| % Linear fitting of tar yields |
| A first order polynomial was used to fit DFA_CS and TarYield to obtain the fit coefficients fit_lin |
| Calculate the tar yield value of the fitted line lin_y |
| % Plotting data points and fitting lines |
| Creating a new graph window |
| Plotting Scatter Plot for DFA_CS and TarYield |
| Keeping the graphs so that more can be plotted on the same plot |
| Plot linear fit lines for DFA_CS and lin_y |
| % Display the linear fit equation |
| Convert the fit coefficients fit_lin to a symbolic polynomial lin_poly |
| Print out the linear fit equation |

| |
|---|
| % Fit a polynomial to the tar yield |
| Choose an appropriate polynomial order for the fit, here choose 2nd order. |
| Use the 2nd order polynomial to fit DFA_CS and TarYield, and get the fit coefficients fit_poly |
| Calculate the tar yield value of the polynomial fit curve poly_y |
| % Plot data points and polynomial fit curve |
| Creating a new graph window |
| Plotting Scatter Plots for DFA_CS and TarYield |
| Keeping graphs so you can plot more on the same graph |
| Plotting polynomial fit curves for DFA_CS and poly_y |

Finally, the following results were produced in the following two areas

R-squared for Linear Fit: 0.77911

R-squared for Polynomial Fit: 0.96936

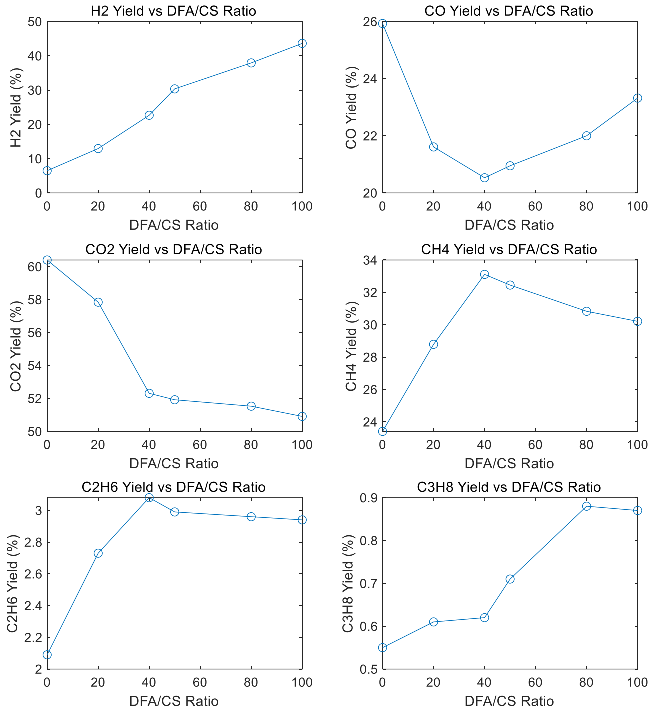
After comparison, we chose to fit the polynomial results.

As for the results of other models, we will put them in the appendix.

Decreasing trend of tar yield: When the DFA/CS ratio increases (i.e., the more desulfurization ash is added), then the tar yield shows a significant decreasing trend.

Modeling and Solving Problem 2

The curves of various gases in the DFA/CS system at different ratios were plotted by matlab and the results are given below:




The CO yield decreased with the increase of DFA/CS ratio in the initial stage and then gradually leveled off; this indicates that the CO production in the flue gas was suppressed to a certain extent when the DFA/CS ratio was small, but its effect gradually decreased with the increase of catalyst dosage.

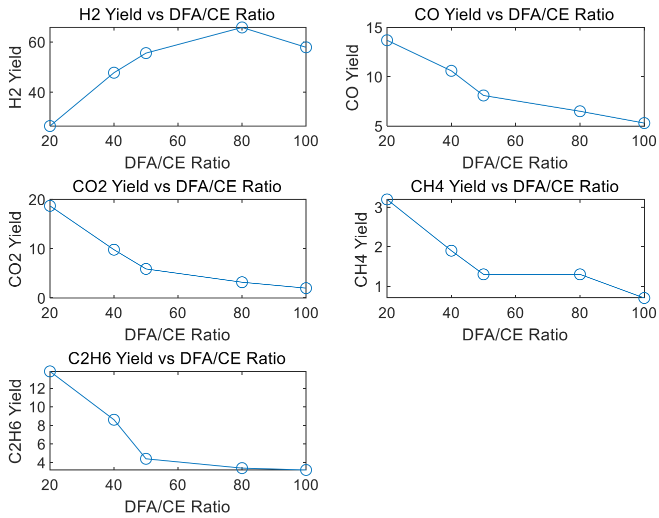
CO2 production:There is a general trend of decreasing CO2 production. This result suggests that the desulfurization ash plays a contributing role to some extent rather than complete oxidation, which reduces the CO2 production.

Methane Yield:The methane yield appeared to be maximum at moderate DFA/CS ratio conditions and then slowly decreased. It indicates that the desulfurization ash has a certain promotion effect on methane production, but its effect gradually decreases with the increase of catalyst ratio. The yields of ethane (C2H6), propane (C3H8), propylene(C3H6),ethylene(C2H4),andbutane(C4H10)yields(C2H6,C3H8,C3H6,C2H 4,C4H10) showed a tendency to ebb and flow, or increase, or decrease, or stabilize.

This difference may be due to the pathway of desulfurization ash on various hydrocarbons. For example, desulfurization ash may accelerate the cleavage of carbon chains to produce shorter hydrocarbons, or have an effect on the equilibrium of unsaturated saturated hydrocarbons.

The results show that the addition of desulfurized ash can effectively control the generation rate of pyrolysis products and change the chemical reaction pathways and kinetics of the pyrolysis process to some extent. The results show that desulfurization ash has a certain effect on the rate, direction and product distribution of the pyrolysis process.

Analyze it with the following line graph




Hydrogen Yield:

In the 100/100 range, the hydrogen yield increases with increasing DFA/CE ratio, but the decrease is not significant. This can indicate that at higher concentrations, the catalyst produces better promotion of hydrogen, but at some point, some equilibrium or other constraint is reached that makes it less efficient.

Carbon Monoxide Production:

As the DFA/CE ratio increases, the CO yield keeps decreasing. This suggests that the addition of desulfurization ash has some inhibitory effect on CO generation or drives the conversion of CO to CO2.

CO2 yield:

The results showed that the CO2 yield decreased significantly with the addition of desulfurization ash, which was mainly due to the partial oxidation of the carbon source by the desulfurization ash, resulting in the generation of more CO and H2 than CO2.

Methane production:

As the DFA/CE ratio increased, the CH4 yield decreased significantly, indicating that the desulfurization ash promoted the degradation of CH4 or affected the conversion path of CH4.

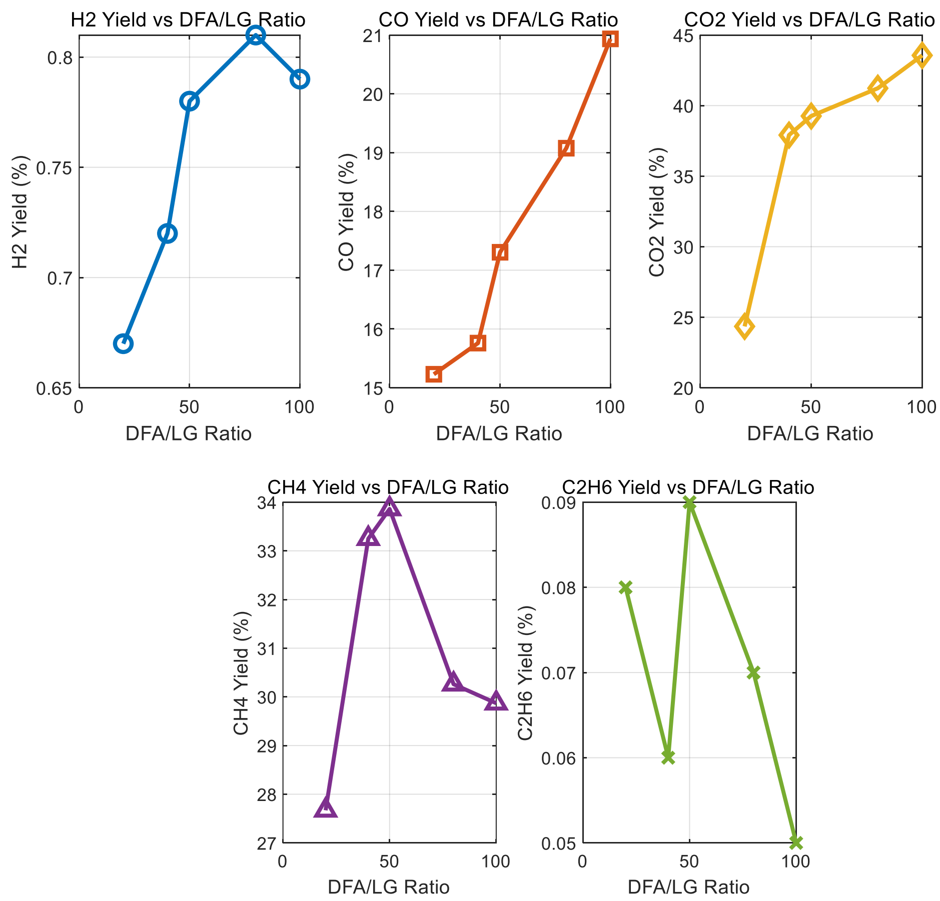
Ethane yield (C2H6):

Similar to methane, the ethane yield decreases significantly when the DFA/CE ratio increases, indicating that the catalyst has a strong influence on the heavy hydrocarbon yield or the heavy hydrocarbon yield.

As a whole, when the DFA/CE ratio increases, the hydrogen yield increases while the other gases yield decreases, indicating that the desulfurization ash plays a role in promoting hydrogen production and changing the reaction equilibrium. It was shown that the desulfurization ash has an important effect on the generation of carbonaceous gases (H2, CO, etc.) such as CH4, CH4.

DFA/LG Group

For the given data, plot a line graph as follows



Hydrogen Yield:

The H2 yield generally increased as the DFA/LG ratio increased and slightly decreased up to 100/100. It was found that as the ratio of desulfurized ash to lignin increased, the hydrogen production yield showed a trend of increasing and then decreasing, most likely due to reaching the peak of catalytic efficacy.

Carbon monoxide yield:

CO yield increased with the increase of DFA/LG ratio. This indicates that the increase in SO2 content in the flue gas contributes to the CO production.

CO2 yield:

CO2 yield increased with increase in DFA/LG ratio. This indicates that the desulfurization ash contributes to the full oxidation of lignin components.

Methane yield:

The CH4 yield showed a maximum value when the DFA/LG ratio was 50/100 and then decreased gradually. This indicates that methane is formed most efficiently at the appropriate concentration of desulfurized ash, and the subsequent desulfurized ash residue will play a catalytic role.

Ethane yield (C2H6):

The yield of C2H6 showed an overall decreasing trend, but the change was small; this suggests that the desulfurization ash has little effect on ethane formation, or it may be that the desulfurization ash has an inhibitory effect on ethane formation.

The results of the study showed that the composition of the generated gas product changed significantly after the addition of desulfurization ash, especially the effect on the production of hydrogen, CO, and methane. This may be due to the effect of desulfurization ash on pyrolysis pathways, such as promoting pyrolysis and oxidation, which in turn changes the distribution of products.

## 4.2 Problem 3 Modeling and Solving

The matlab t-test for two independent samples is used in this paper, and some of the data in Appendix I is given here as an example.

| circumstance | p-value |
| --- | --- |

| | |
|---|---|
| Tar yield p-value | 0.000000 |
| Water yield p-value | 0.476739 |
| Char yield p-value | 0.000000 |
| Syngas yield p-value | 0.000031 |

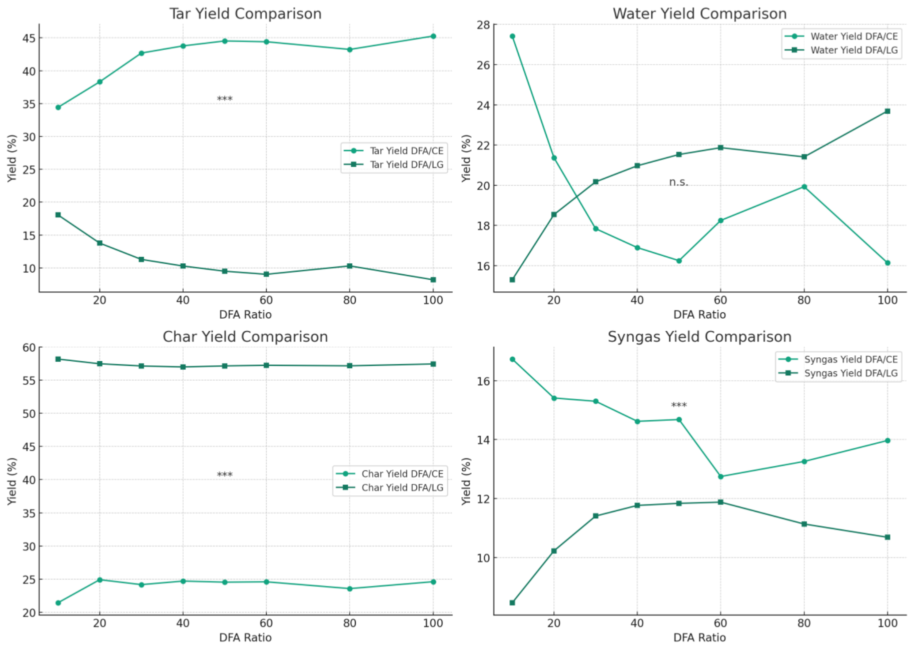
P-value for tar yield: very small (almost 0), much less than 0.05, indicating that the DFA/CE to DFA/LG ratio is statistically significant. That is, in both cases, we have enough evidence to show that the change in tar yield is not accidental.

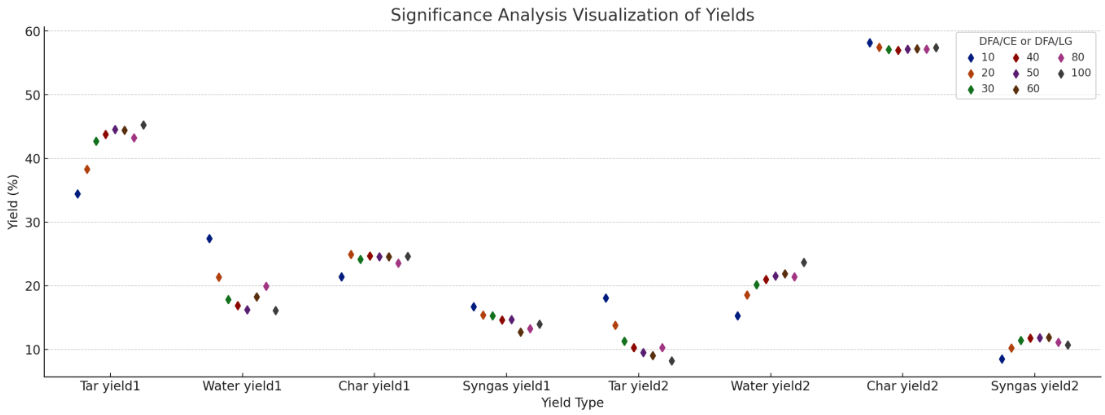
The P-value for water yield is 0.476739, which is greater than 0.05, indicating that there is no statistical significance in water yield for the two ratios. This suggests that the difference in water yield in both cases is not by chance but is influenced by environmental factors.

The P-value for coke yield is very small (almost 0) and much less than 0.05, indicating that there is a statistically significant difference in coke yield between the two proportions. The results of the study show that there is a significant difference in the effect of desulfurization ash as a catalyst on coke yield in the two biomass pyrolysis reactions.

Syngas yield P=0.000031, which is significantly less than 0.05, indicates that there is a statistically significant difference in syngas yield between the two ratios. This suggests that different pyrolysis conditions have a significant effect on the syngas yield.

In conclusion, the yields of the pyrolysis products, except for the moisture yield, differed significantly under DFA/CE and DFA/LG conditions. This indicates that, to a certain extent, the desulfurization ash has a certain promotion effect on the thermal cracking of cellulose and lignin, and its effect is related to the type and proportion of pyrolysis products. This apparent difference may be due to the fact that the desulfurization ash changes the pyrolysis reaction pathway, which in turn affects the final composition of the generated gas. In order to show this difference more intuitively, we have created the following visualization.




Based on the results shown above, the following describes the yield of each pyrolysis product as a function of the DFA/CE or DFA/LG ratio:

Tar yields: in DFA/CE (Thayerd1), the tar yields fluctuate with increasing ratio, but there is an overall increasing trend. In the DFA/LG group (Thayer d2), there was a general trend of decreasing tar yield as the ratio increased. The experimental results show that at higher desulfurization rates, there is a large difference in tar yield between the two conditions.

Wateryield: In the DFA/CE (Wateryield1) treatment, the yield varied over time with a general decreasing trend. In the DFA/LG combination (Wateryield2), the yield increased as the ratio increased. This indicates that under DFA/LG condition, the desulfurization ash contributes more to the water yield.

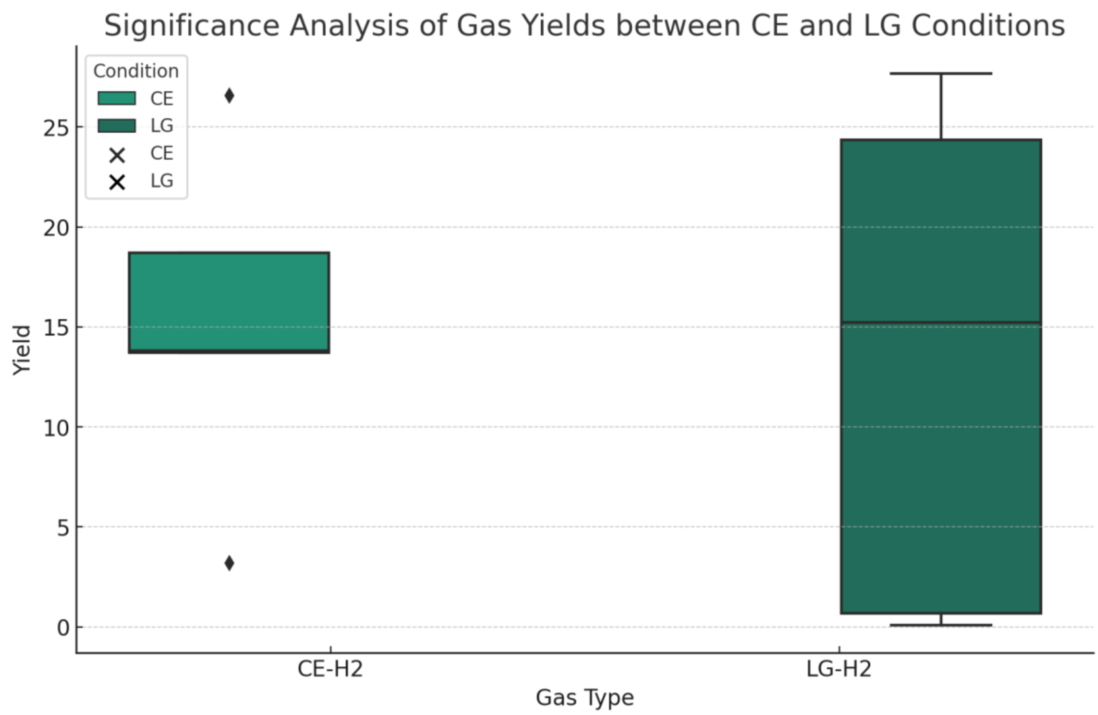
Coke production: Under the DFA/CE (Charyield1) condition, the coke production remains basically unchanged, with no significant upward or downward trend. Under the DFA/LG combination (Charyield2) working condition, the coke yield has been maintained at a high level, showing some fluctuations. The results show that the coke yield varies greatly between the two conditions, especially at the DFA/LG ratio.

Syngas yield: Under DFA/CE (Syngasyield1), the syngas yield showed a tendency of decreasing and then increasing with the increase of the ratio. Under DFA/LG (Syngasyield2), the syngas yield decreases as the ratio increases. The syngas yield increased with the increase of the added amount, indicating the addition of desulfurized ash to the syngas.

Overall, the results of the visualization experiments showed that there was a significant difference in the yield of pyrolysis products under different reaction conditions. The results also showed that the effect of desulfurization ash on the two pyrolysis pathways was different under DFA/CE and DFA/CE conditions. Under DFA/LG conditions, the coke yields were generally high, which was strongly related to its thermal cracking pathway with cellulose. On this basis, the variation of syngas yield can well reflect the effect of desulfurization ash on gas products.

The analysis of variance of the data in Annexure II was carried out using the functions of matlab and the following results were obtained

| **circumstance** | **p-value** |
|---|---|
| P-value for H2 | 0.82285 |
| P-value for CO p-value | 0.87813 |
| P-value for CO2 | 0.81235 |
| P-value for CH4 | 0.88345 |
| P-value for C2H6 | 0.71669 |




In terms of visualization, this study presents the gas yields under different reaction conditions (CE vs. LG) using box plots and scatter plots, respectively. The box plots show the range of distribution of gas yields, intermediate and outliers in each case. The scatter plot presents each particular data point so that we can see how these data are distributed.From the graph, we can see the distribution of gas yields in various scenarios. For example, when the median line of the box plot (the center line of the box plot) in one case is compared to the other, it indicates that there is a difference in gas yield between the two cases. However, because the values of P in the previous studies were very large, we were unable to demonstrate that the difference was statistically significant.

Reaction mechanism of pyrolysis products

Tar Yield:

$$y = 0.001x^2 - 0.17x + 18.92$$

R-squared for Polynomial Fit: 0.96936

Water Yield:

$y = 0.001x^2 - 0.0521x + 26.9329$

R-squared for Polynomial Fit: 0.97242

Char Yield:

$y = 0.0001x^2 - 0.0007x + 29.1948$

R-squared for Polynomial Fit: 0.89768

Syngas Yield:

$y = 0.001x^2 + 0.1196x + 24.9484$

S-squared for Polynomial Fit: 0.83259

## 4.3 Modeling and Solving Problem 4

As can be seen in Figs. 8,9,10, the output of the gas is strongly nonlinear. For specific nonlinear problems, it is difficult to analyze them directly. For this reason, polynomials of the first, second, third and fourth orders, and first-order exponentials and quadratic exponentials are chosen as fitting methods. Each combination is fitted several times and the one with the best fit is selected as the output, based on which the matlab program is written.

| **Set DFA_CS as [0, 20, 40, 50, 80, 100]** |
|---|
| Set gas data H2, CO, CO2, CH4, C2H6, C3H8, C3H6, C2H4, C4H10 with respective values |
| Initialize empty lists bestFits, bestGOFs, bestModels, bestEquations |
| Set DFA_CS_range from the minimum positive value of DFA_CS to its maximum |
| Define list of models: 'poly1', 'poly2', 'poly3', 'poly4', 'exp1', 'exp2' |

Define list of gases: 'H2', 'CO', 'CO2', 'CH4', 'C2H6', 'C3H8', 'C3H6', 'C2H4', 'C4H10'

Create a new figure for subplots

For each gas in gases list:

Retrieve the yData for the gas

Set bestAdjR2 to negative infinity

For each model in models list:

If the model requires positive x values and the first DFA_CS is zero, skip this model

Otherwise:

Set xData as DFA_CS

Set yData_fit as yData for the current gas

If the model works only with positive x values:

Filter xData and yData_fit to exclude the zero in DFA_CS

Fit the model to xData and yData_fit

If the model's adjusted R-square is better than bestAdjR2:

Update bestAdjR2 with the new value

Add the fit result to bestFits list

Add the goodness-of-fit to bestGOFs list

Add the model to bestModels list

Generate the fit equation and add to bestEquations list

Plot the best fit for the current gas on a subplot:

Plot xData vs yData as points

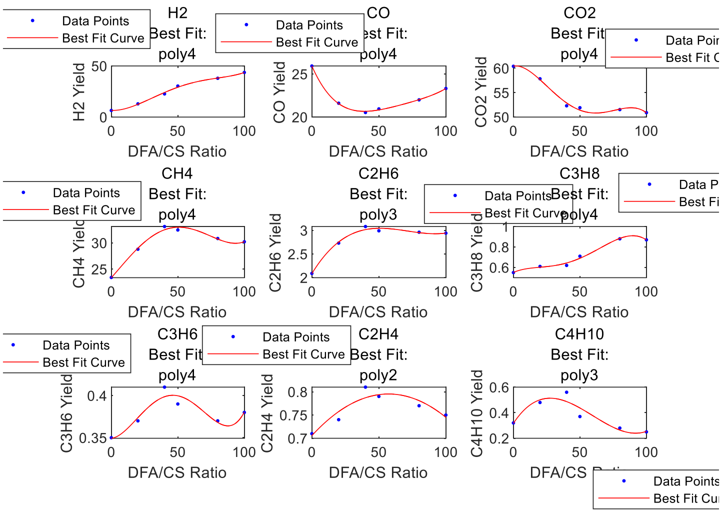
Plot DFA_CS_range vs the best fit as a line

Set title, labels, and legend

Print the best fit equation and its R-square value for the current gas

Adjust layout of the figure to prevent overlap

Finally, the result is obtained as shown below




Some of the results are shown below and the full results are shown in Appendix III. Using the reaction mechanism we can see that for the two gases

H2 Best Fit Equation：y="-0.037643x + 0.056423 "+" 0.784114" "x" ^"-1"

H2 Best R-square: 0.903654

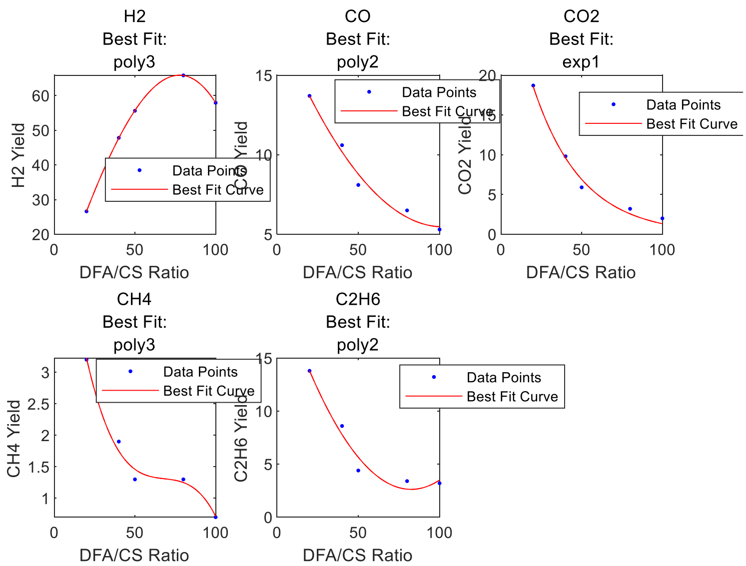
The fitted equation contains quadratic, primary and x-1 terms and the R-squared value is 0.982587 which shows that the fit is good and the model explains the variation in the data well.

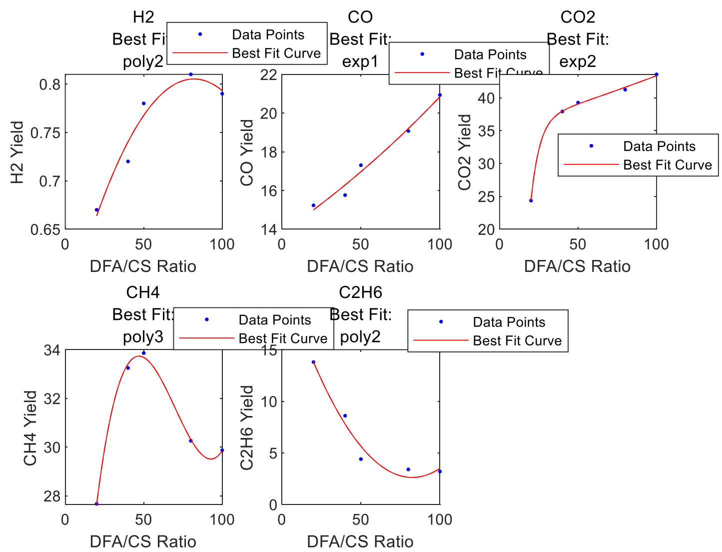
CH4 Best Fit Equation:

CH4 Best R-square: 0.990376

The fitted equation contains quadratic, primary and x-1 terms.The R-squared value is 0.971356 which shows high goodness of fit.

Using the same model applied to DFA/CE group, DFA/LG group. The visualization results were obtained as follows



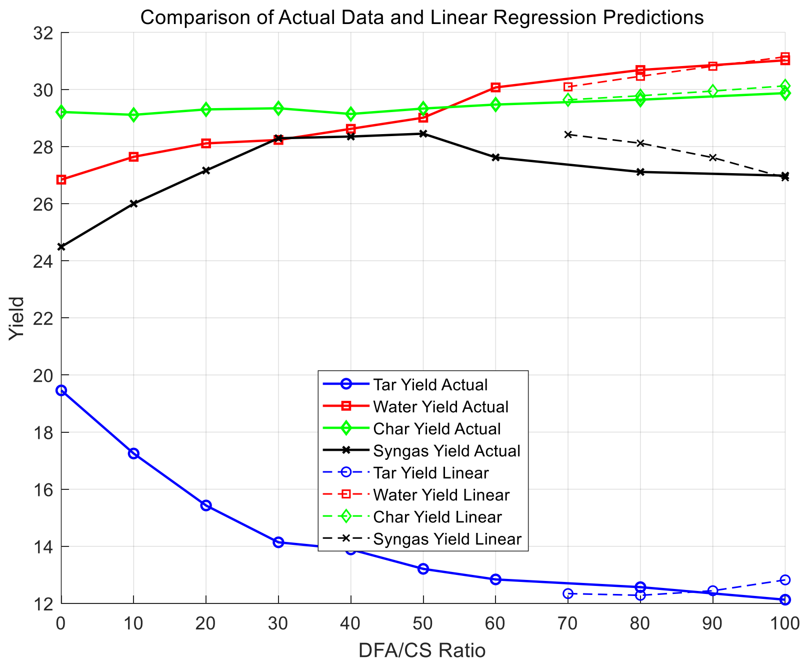

The specific functional relationship expressions are put in Appendix IV to show.

## 4.4 Modeling and Solving Problem 5

In this project, it is proposed to carry out the simulation studies of pyrolysis product yield and cracking gas under 70/100 and 90/100 conditions, respectively, under two conditions.

First, using the regression model established in the fourth problem, Tar yield, Water yield, Char yield, Syngas yield under 70/100 and 90/100 conditions, respectively, its results are as follows

| DFA/CS | 70/100 | 90/100 |
|---|---|---|
| Tar yield | 12.41 | 12.53 |
| Water yield | 30.0899 | 30.8119 |
| Char yield | 29.6358 | 29.9418 |
| Syngas yield | 28.4204 | 27.6124 |




On this basis, the forecast results are weighted and averaged by introducing two other forecasting models to improve the accuracy of the forecasts. Here, we use 70/100 and 90/100 DFA/CS models for forecasting respectively.

| DFA/CS | Tar yield | Water yield | Char yield | Syngas yield |
|---|---|---|---|---|
| 0 | 19.46 | 26.84 | 29.21 | 24.49 |
| 10 | 17.25 | 27.64 | 29.11 | 26 |
| 20 | 15.43 | 28.11 | 29.3 | 27.16 |
| 30 | 14.14 | 28.23 | 29.34 | 28.29 |
| 40 | 13.89 | 28.62 | 29.14 | 28.35 |
| 50 | 13.21 | 29.01 | 29.33 | 28.45 |
| 60 | 12.84 | 30.07 | 29.47 | 27.62 |
| 70 | | | | |

| | | | | |
|---|---|---|---|---|
| 80 | 12.57 | 30.68 | 29.64 | 27.11 |
| 90 | | | | |
| 100 | 12.13 | 31.02 | 29.87 | 26.98 |

Here we will build the GM(1,1) model to predict the data of the constituency.

# 5. Conclusion

The model demonstrated significant potential in the study of catalytic reactions of cotton straw pyrolysis, especially in analyzing and predicting pyrolysis products and pyrolysis gas products under different catalytic conditions. Since the model combines data preprocessing, correlation analysis, significance analysis, and multiple prediction methods, it can be applied not only to the pyrolysis process of cotton straw, but also has the potential to be generalized to other types of biomass pyrolysis studies. For example, by appropriately adjusting the model parameters, it can be applied to the pyrolysis analysis of other agricultural wastes such as wood chips and rice husk.

The efficiency and accuracy demonstrated by the model in processing and analyzing large datasets make it a valuable tool in the field of renewable energy, especially in biomass energy conversion studies. Through detailed analysis of pyrolysis products, the model can help researchers and engineers to optimize the biomass pyrolysis process and improve the efficiency of energy conversion, thus supporting the development of sustainable energy systems.

Considering the growing global demand for renewable energy and biomass energy, the model is of great practical significance for improving the effective utilization of biomass resources. It can not only help enhance energy output, but also contribute to reducing environmental pollution and promoting sustainable development.

With its comprehensiveness, flexibility and advanced analytical capabilities, the model is expected to find a wide range of applications in the field of biomass energy conversion and sustainable energy.

# References


[1] 张一林，郁芸君，陈珠明.人工智能、中小企业融资与银行数字化转型[J].中国工业经济，2021（12）：69-87.DOI10.19581/j.cnki.ciejournal.2021.12.003.

[2] 杨寅,刘勤,吕晓雷.中国企业智能财务应用现状及发展趋势分析——基于2021年调查问卷数据的例证[J].会计之友,2022(20):111-117.

[3] 刘薇,宁晓红.护士安宁疗护知信行及培训需求调查问卷的编制及信效度检验[J].实用老年医学,2022,36(11):1170-1173+1184.

[4] 史枚翎.移动APP对图书馆社会化阅读推广服务的影响因素研究——基于调查问卷的实证分析[J].图书馆研究与工作,2022(06):16-20.

[5] 贺宇,赵建航,孔朝阳.中国产品在东盟市场的印象研究——基于印尼、泰国、老挝三国调查问卷的分析[J].科学决策,2022(10):125-137.

[6] 吕承超,崔悦.乡村振兴发展:指标评价体系、地区差距与空间极化[J].农业经济问题,2021(05):20-32.DOI:10.13246/j.cnki.iae.2021.05.004.

[7] 裴征，朱晓伟，龚超.基于层次分析法和熵值法组合的DRG指标评价体系权重赋值研究[J].中国医院管理，2020，40（11）：69-72+83.

[8] 崔明明,聂常虹.基于指标评价体系的我国粮食安全演变研究[J].中国科学院院刊,2019,34(08):910-919.DOI:10.16418/j.issn.1000-3045.2019.08.009.

[9] 张宝友,杨玉香,孟丽君.物流服务质量评价模型与方法研究综述[J].中国流通经济,2021,35(02):49-60.DOI:10.14089/j.cnki.cn11-3664/f.2021.02.005.

[10] 谢幼如,常亚洁.绩效导向的教育信息化评价模型的构建[J].中国电化教育,2015(01):56-61+92.

[11] 刘晋,程彦斌,齐东川等.基于支持向量机的化工工艺安全评价模型构建及优化 研究[J].中国安全生产科学技术,2022,18(12):154-161.

# Appendix

Annex I Other model results

Tar Yield:
Linear Fit: y = -0.07*x + 17.38
R-squared for Linear Fit: 0.77911
Polynomial Fit: y = 0.00*x^2 + -0.17*x + 18.92
R-squared for Polynomial Fit: 0.96936

Water Yield:
Linear Fit: y = 0.04*x + 27.07
R-squared for Linear Fit: 0.96777
Polynomial Fit: y = -0.00*x^2 + 0.05*x + 26.93
R-squared for Polynomial Fit: 0.97242

Char Yield:
Linear Fit: y = 0.01*x + 29.09
R-squared for Linear Fit: 0.8063
Polynomial Fit: y = 0.00*x^2 + -0.00*x + 29.19
R-squared for Polynomial Fit: 0.89768

Syngas Yield:
Linear Fit: y = 0.02*x + 26.46
R-squared for Linear Fit: 0.17319
Polynomial Fit: y = -0.00*x^2 + 0.12*x + 24.95
R-squared for Polynomial Fit: 0.83259

2、Complete results for reaction mechanism
Reaction mechanism of DFA/CE group of pyrolysis products

Tar Yield:
Linear Fit: y = 0.09*x + 37.51
R-squared for Linear Fit: 0.5754
Polynomial Fit: y = -0.00*x^2 + 0.36*x + 32.22
R-squared for Polynomial Fit: 0.86617

Water Yield:
Linear Fit: y = -0.07*x + 22.90
R-squared for Linear Fit: 0.36466
Polynomial Fit: y = 0.00*x^2 + -0.34*x + 28.04
R-squared for Polynomial Fit: 0.64088

Char Yield:
Linear Fit: y = 0.01*x + 23.39
R-squared for Linear Fit: 0.13471

Polynomial Fit: y = -0.00*x^2 + 0.09*x + 21.94
R-squared for Polynomial Fit: 0.37426

Syngas Yield:
Linear Fit: y = -0.03*x + 16.20
R-squared for Linear Fit: 0.62448
Polynomial Fit: y = 0.00*x^2 + -0.11*x + 17.79
R-squared for Polynomial Fit: 0.85698
Reaction mechanism of DFA/LG group of pyrolysis products
Tar Yield:
Linear Fit: y = -0.08*x + 15.38
R-squared for Linear Fit: 0.62996
Polynomial Fit: y = 0.00*x^2 + -0.29*x + 19.50
R-squared for Polynomial Fit: 0.87416

Water Yield:
Linear Fit: y = 0.07*x + 16.91
R-squared for Linear Fit: 0.75001
Polynomial Fit: y = -0.00*x^2 + 0.19*x + 14.55
R-squared for Polynomial Fit: 0.87757

Char Yield:
Linear Fit: y = -0.00*x + 57.57
R-squared for Linear Fit: 0.15554
Polynomial Fit: y = 0.00*x^2 + -0.04*x + 58.33
R-squared for Polynomial Fit: 0.76475

Syngas Yield:
Linear Fit: y = 0.02*x + 10.14
R-squared for Linear Fit: 0.18319
Polynomial Fit: y = -0.00*x^2 + 0.14*x + 7.61
R-squared for Polynomial Fit: 0.88912

3. Complete results of gas fitting
H2 Best Fit Equation: 3.119481*x^3 + -2.413486*x^2 + -7.636714*x + 17.793309 + 28.383004*x^-1
H2 Best R-square: 0.982587
CO Best Fit Equation: 0.531694*x^3 + -1.177807*x^2 + 1.276710*x + 0.884053 + 20.788860*x^-1
CO Best R-square: 0.990761
CO2 Best Fit Equation: -2.224212*x^3 + 0.552478*x^2 + 6.327436*x + -4.391436 + 51.603495*x^-1
CO2 Best R-square: 0.958848
CH4 Best Fit Equation: 1.175847*x^3 + 1.058320*x^2 + -5.691765*x + 0.744921 +

32.955346*x^-1
CH4 Best R-square: 0.941432
C2H6 Best Fit Equation: 0.168309*x^2 + -0.314102*x + 0.039340 + 3.041718*x^-1
C2H6 Best R-square: 0.971356
C3H8 Best Fit Equation: -0.060375*x^3 + -0.037709*x^2 + 0.126709*x + 0.194898 + 0.680787*x^-1
C3H8 Best R-square: 0.946513
C3H6 Best Fit Equation: 0.021364*x^3 + 0.009134*x^2 + -0.059469*x + -0.006724 + 0.400144*x^-1
C3H6 Best R-square: 0.416607
C2H4 Best Fit Equation: -0.037694*x + 0.017963 + 0.793079*x^-1
C2H4 Best R-square: 0.742319
C4H10 Best Fit Equation: 0.105841*x^2 + -0.097641*x + -0.207335 + 0.446484*x^-1
C4H10 Best R-square: 0.638931
H2 Best Fit Equation: -1.637120*x^2 + -11.999472*x + 16.583320 + 60.575978*x^-1
H2 Best R-square: 0.999919
CO Best Fit Equation: 1.243776*x + -3.438598 + 7.844980*x^-1
CO Best R-square: 0.961805
CO2 Best Fit Equation: exp model equation here...
CO2 Best R-square: 0.985883
CH4 Best Fit Equation: -0.512466*x^2 + 0.495875*x + -0.254250 + 1.357300*x^-1
CH4 Best R-square: 0.942849
C2H6 Best Fit Equation: 2.921224*x + -4.481887 + 4.343020*x^-1
C2H6 Best R-square: 0.930370
H2 Best Fit Equation: -0.037643*x + 0.056423 + 0.784114*x^-1
H2 Best R-square: 0.903654
CO Best Fit Equation: exp model equation here...
CO Best R-square: 0.972261
CO2 Best Fit Equation: exp model equation here...
CO2 Best R-square: 0.995972
CH4 Best Fit Equation: 2.854572*x^2 + -3.186335*x + -3.229576 + 33.118868*x^-1
CH4 Best R-square: 0.990376
C2H6 Best Fit Equation: 2.921224*x + -4.481887 + 4.343020*x^-1
C2H6 Best R-square: 0.930370

32.955346*x^-1
CH4 Best R-square: 0.941432
C2H6 Best Fit Equation: 0.168309*x^2 + -0.314102*x + 0.039340 + 3.041718*x^-1
C2H6 Best R-square: 0.971356
C3H8 Best Fit Equation: -0.060375*x^3 + -0.037709*x^2 + 0.126709*x + 0.194898 + 0.680787*x^-1
C3H8 Best R-square: 0.946513
C3H6 Best Fit Equation: 0.021364*x^3 + 0.009134*x^2 + -0.059469*x + -0.006724 +

0.400144*x^-1
C3H6 Best R-square: 0.416607
C2H4 Best Fit Equation: -0.037694*x + 0.017963 + 0.793079*x^-1
C2H4 Best R-square: 0.742319
C4H10 Best Fit Equation: 0.105841*x^2 + -0.097641*x + -0.207335 + 0.446484*x^-1
C4H10 Best R-square: 0.638931
H2 Best Fit Equation: -1.637120*x^2 + -11.999472*x + 16.583320 + 60.575978*x^-1
H2 Best R-square: 0.999919
CO Best Fit Equation: 1.243776*x + -3.438598 + 7.844980*x^-1
CO Best R-square: 0.961805
CO2 Best Fit Equation: exp model equation here...
CO2 Best R-square: 0.985883
CH4 Best Fit Equation: -0.512466*x^2 + 0.495875*x + -0.254250 + 1.357300*x^-1
CH4 Best R-square: 0.942849
C2H6 Best Fit Equation: 2.921224*x + -4.481887 + 4.343020*x^-1
C2H6 Best R-square: 0.930370
H2 Best Fit Equation: -0.037643*x + 0.056423 + 0.784114*x^-1
H2 Best R-square: 0.903654
CO Best Fit Equation: exp model equation here...
CO Best R-square: 0.972261
CO2 Best Fit Equation: exp model equation here...
CO2 Best R-square: 0.995972
CH4 Best Fit Equation: 2.854572*x^2 + -3.186335*x + -3.229576 + 33.118868*x^-1
CH4 Best R-square: 0.990376
C2H6 Best Fit Equation: 2.921224*x + -4.481887 + 4.343020*x^-1
C2H6 Best R-square: 0.930370

4. Complete results of other combinations of question four
H2 Best Fit Equation: -1.637120*x^2 + -11.999472*x + 16.583320 + 60.575978*x^-1
H2 Best R-square: 0.999919
CO Best Fit Equation: 1.243776*x + -3.438598 + 7.844980*x^-1
CO Best R-square: 0.961805
CO2 Best Fit Equation: exp model equation here...
CO2 Best R-square: 0.985883
CH4 Best Fit Equation: -0.512466*x^2 + 0.495875*x + -0.254250 + 1.357300*x^-1
CH4 Best R-square: 0.942849
C2H6 Best Fit Equation: 2.921224*x + -4.481887 + 4.343020*x^-1
C2H6 Best R-square: 0.930370
H2 Best Fit Equation: -0.037643*x + 0.056423 + 0.784114*x^-1
H2 Best R-square: 0.903654
CO Best Fit Equation: exp model equation here...
CO Best R-square: 0.972261

CO2 Best Fit Equation: exp model equation here...
CO2 Best R-square: 0.995972
CH4 Best Fit Equation: 2.854572*x^2 + -3.186335*x + -3.229576 + 33.118868*x^-1
CH4 Best R-square: 0.990376
C2H6 Best Fit Equation: 2.921224*x + -4.481887 + 4.343020*x^-1
C2H6 Best R-square: 0.930370