World Scientific Research Journal Volume 8 Issue 11, 2022
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

A Study on The Performance Control of Spray-fused Nonwoven Materials Based on Multiple Linear Regression

Zhi Liu[1, a], Youyou Liu[2, b], Xuan Tang[1, c], Liangliang Yang[1, d], Chenghao Gu[1, e]

[1]Nanjing Institute of Technology, Nanjing, China

[2]Zhengzhou Institute of Aviation Industry Management, Zhengzhou, China

[a]1428632509@qq.com, [b]1395173697@qq.com, [c]1784291289@qq.com

Abstract

Meltblown nonwoven materials have poor compressive resilience due to their inherently fine fibers, and their performance cannot be guaranteed. However, with the development of science and technology, scientists have developed the interlayer meltblown method, and the interlayer meltblown nonwoven material with "Z-shaped" structure has been released. If a relationship model can be established to analyze the relationship between process parameters and structural variables, structural variables and product performance, it is of great practical significance to establish the mechanism of product performance regulation. To address problem 1: In this paper, the raw data is first analyzed and processed to fill in the missing interpolation rates. The rate of change of product performance parameters was calculated for each structural variable before and after interpolation for each group of experiments. The data were then visualized and analyzed, and it was concluded that all five variables before and after interpolation had similar trends and obvious correlations; the compression resilience rate was not obviously correlated. Finally, the Spearman correlation coefficient model was established, and the intercalation rate had a significant effect on compression resilience and filtration resistance, and the intercalation rate was negatively correlated with compression resilience with a correlation coefficient of -0.358, and positively correlated with filtration resistance with a correlation coefficient of 0.375. In response to question two: first the data were pre-processed, followed by a multiple linear regression model for variance testing and then a heteroskedasticity test to determine that no heteroskedasticity existed. Next, a VIF test for multicollinearity was performed and it was concluded that the regression model did not have the effect of multicollinearity. The standard regression model was then solved for and predictions were made using this model, some of the results are shown below, please see Table 10 for detailed results. In response to question three: the relationship between the variables was first studied by calculating Spearman's correlation coefficients for the structural variable parameters. A multiple linear regression model was then developed based on the interaction term and the regressed model was subjected to the White's test for heteroskedasticity, which determined the presence of heteroskedasticity. A VIF test for multicollinearity was then performed, which yielded the presence of significant multicollinearity effects between the parameters. A stepwise regression operation was then performed to reflect the relationship between structural variables and product performance to calculate the best performance. Finally, based on the above operation, a single objective programming model was developed and solving this model resulted in the highest filtration efficiency of 86.743% when the receiving distance was 20 cm and the hot air speed was 1195 r/min, at which point the parameters of the structural variables were 2.4077 mm thickness, 95.740% porosity and 85.203% compression resilience. For question 4: Based on the regression equation established in question 3 combined with the regression equation

obtained in question 2, a bi-objective planning model was developed to minimize the filtration resistance and maximize the filtration efficiency. This was then converted to a single-objective planning model combined with sensitivity analysis to obtain the following results.

Keywords

Multiple linear regression model; Spearman's correlation coefficient; Sensitivity analysis; bi-objective programming model.

1. INTRODUCTION

1.1. Background of the Problem

As an important raw material in the process of mask making, although meltblown nonwoven materials are of great interest both at home and abroad for their good filtration performance, simple preparation process, low cost and light weight, the fibers are too fine resulting in poor compression resilience and cannot guarantee their performance.

In order to solve these problems, scientists developed the interlayer meltblown method, and the "Z-shaped" structure of the interlayer meltblown nonwoven material came into existence. The process is complicated by the number of preparation parameters, the interaction between parameters and the influence of the intercalation airflow, and the scientists' study of the process parameters that determine the structural variables and the constructive variables that determine the product properties.

In this context, it would be of great help to establish a mechanism for product performance regulation if the relationship between process parameters and structural variables and between structural variables and product performance could be modeled.

1.2. Presentation of the Problem

To address the following question, this paper provides the relevant data: question C data.xlsx.

Based on the above background, review of relevant literature, and contextualized mathematical modeling with an understanding of the context of expertise, solve the following problems.

(1) To study the changes in structural variables and product properties after intercalation and to analyze the effect of the intercalation rate on this change.

(2) To study the relationship between process parameters and structural variables and to predict the data of structural variables in Table 1.

(3) To study the relationship between structural variables and product performance, between structural variables and between product performance. On the basis of problem II, the process parameters are studied so that the filtration efficiency is maximized.

(4) According to the actual situation, taking into account the conditions and requirements of all aspects, study the process parameters, so that the filtration efficiency is as high as possible and the filtration resistance is as low as possible.

2. ANALYSIS OF THE PROBLEM

2.1. Analysis of Question One

Problem 1 requires a study of the changes in structural variables, product properties after intercalation and an analysis of the effect of the intercalation rate on this change. In this paper,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

the raw data can first be analyzed and processed to fill in the missing intercalation rates. The rate of change of each structural variable, product performance parameter before and after interpolation is calculated for each group of experiments, followed by descriptive analysis to analyze the correlation of interpolation on structural variables and product performance and point out the change pattern. The data were then visualized and analyzed by drawing a visualization for analysis. Finally, a relevant mathematical model was developed to analyze the effect of intercalation rate on these changes.

2.2. Analysis of Question Two

Problem 2 requires a study of the relationship between process parameters and structural variables to predict the structural variable data in Table 1. Firstly the data will be pre-processed and subsequently based on the proposal a multiple linear regression prediction model can be developed for variance testing followed by a heteroskedasticity test to determine if heteroskedasticity exists by. Secondly VIF test for multicollinearity is performed to test whether this regression model possesses the effect of multicollinearity. Finally then a standard regression model is solved and this model is used to make predictions to produce the required predictions for problem two.

2.3. Analysis of Question Three

Problem III requires the study of the relationship between structural variables and product performance, between structural variables, and between product performance. On the basis of problem two, the process parameters are studied so that the filtration efficiency is maximized. This problem can be solved by first calculating the Spearman's correlation coefficient for the structural variable parameters to study the relationship between the variables. It is also clear from the topic that there are more process parameters and there is also interaction between the parameters, so a multiple linear regression model can be developed based on the interaction term and the heteroscedasticity White's test can be performed using the regressed model to determine if there is heteroscedasticity. A VIF test for multicollinearity is then performed to test for the effect of multicollinearity between the parameters. Subsequently a stepwise regression operation is performed to reflect the relationship between structural variables and product performance to calculate the best performance. Finally, based on the above operations, a relevant single objective programming model can be developed to obtain the results of this question.

2.4. Analysis of Question Four

The fourth question requires the study of the process parameters that make the filtration efficiency as high as possible and the filtration resistance as low as possible, taking into account the conditions and requirements of the actual situation. Analyzing this question, we can find out the optimal combination of receiving distance and hot air speed, and to make the filtration efficiency high while minimizing the filtration resistance. By combining the regression equation between structural variables and product performance established in the third question with the regression equation between process parameters and structural variables obtained in the second question, we can obtain a dual-objective planning model that minimizes the filtration resistance and maximizes the filtration efficiency. By searching for the extreme value points of the dual objective planning, the optimal combination of receiving distance and hot air velocity is sought. To obtain the results of this model, it can be first transformed into a single-objective planning model combined with sensitivity analysis to find the optimal process parameters to meet the requirements of the topic.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

3. BASIC ASSUMPTIONS

1. assuming that the data given in the question are true and reliable.
2. assumes that the standardized reference value is the mean of the data.
3. assumes that no other relevant variables have an effect on product performance.

4. DESCRIPTION OF SYMBOLS

symbolic	Symbol Description	Symbolic unit
$y_i(j)$	The group number is the jth parameter of the i uninterpolated group	
$y_i'(j)$	The group number is the jth parameter of the i interpolation group	
$z_i(j)$	The rate of change of the jth parameter with group number i	
VIF	expansion factor	
x_1	reception distance	cm
x_2	Hot air speed	r/min
x_3	thickness	mm
x_4	porosity	
x_5	Compression resilience	
y_1	Filtration resistance	Pa
y_2	Filtration efficiency	
y_3	permeability	mm/s
r_s	Correlation coefficient	

5. DATA PROCESSING AND CORRELATION ANALYSIS

5.1. Data pre-processing

Firstly, this paper observes and analyzes the original data and finds that the uninterpolated group lacks interpolation rate. In order to fully and concisely reflect the effective information of the data and the subsequent model solution, the original data data1 is now treated as follows.

Step1: Fill in the missing interpolation rate for the uninterpolated group in the original data table with "0%", taking into account practical considerations.

Step2: Calculate the rate of change of parameters for each structural variable, product performance before and after interpolation for each group of experiments.

$$z_i(j)=\frac{y_i'(j)-y_i(j)}{y_i(j)}\times 100\% \quad i\in[1,25],j\in[1,6] \tag{1}$$

In the above equation $y_i'(j)$, $j\in(1,6)$ represents the thickness, porosity, compression resilience, filtration resistance, filtration efficiency and air permeability of the mask after intercalation, respectively, and $z_i(j)$, $j\in(1,6)$ represents the rate of change of mask thickness,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

porosity, compression resilience, filtration resistance, filtration efficiency and air permeability after intercalation, respectively.

Some of the data after processing are as follows (see appendix for complete data$data1'$).

Table 1. Data after data1 processing

group number	Thickness mm			...	Air permeability mm/s			interpolation rate
	1#	2#	rate of change		1#	2#	rate of change	
1	1.715	2.810	63.848	...	777.100	1019.670	31.215	36.440
2	1.830	2.910	59.016	...	795.570	968.630	21.753	24.740
3	1.890	3.425	81.217	...	564.930	643.400	13.890	31.450
4	2.095	3.400	62.291	...	474.500	603.170	27.117	19.370
5	2.235	3.845	72.036	...	347.230	405.830	16.876	31.190
6	1.390	2.015	44.964	...	685.970	660.300	-3.742	16.530
7	1.705	2.750	61.290	...	477.570	642.900	34.619	19.730
8	1.550	3.165	104.194	...	326.330	408.600	25.211	21.720
9	1.915	2.810	46.736	...	268.830	329.730	22.654	18.620
...	...	...	...	...	...	...	...	...
22	1.000	2.020	102.000	...	211.400	234.130	10.752	19.350
23	0.995	2.060	107.035	...	182.630	226.030	23.764	24.020
24	1.110	2.085	87.838	...	154.200	191.100	23.930	35.880
25	1.365	2.425	77.656	...	137.170	156.330	13.968	32.950

5.2. Descriptive Statistics

To further dissect the data, descriptive statistics are presented in Table 1, yielding the following table.

Table 2. Descriptive statistics of the treated data

	Variable	Mean	Std. Dev.	Min	Max
Thickness mm	1#	1.519	0.383	0.870	2.235
	2#	2.607	0.549	1.740	3.845
	rate of change	75.672	27.604	11.024	137.023
Porosity (%)	1#	92.349	2.091	87.230	95.030
	2#	95.861	0.918	94.000	97.300
	rate of change	3.839	1.787	0.945	7.761
Compression resilience (%)	1#	79.442	9.326	43.910	90.360
	2#	86.623	4.433	75.970	94.590
	rate of change	11.367	21.726	-11.704	102.141
Filter resistance Pa	1#	29.784	16.462	8.130	65.770
	2#	24.157	14.784	7.200	62.033
	rate of change	-19.796	9.544	-38.722	-0.702
Filtration efficiency (%)	1#	34.992	25.131	1.733	80.033
	2#	49.387	18.874	19.967	86.967
	rate of change	197.974	368.953	1.123	1423.370
Air permeability mm/s	1#	347.604	189.575	137.170	795.570
	2#	422.134	236.817	156.330	1019.670
	rate of change	20.643	11.519	-3.742	44.354
	interpolation rate	23.036	11.569	2.500	50.870

From the above table, it can be seen that the average rate of change of thickness and filtration efficiency before and after intercalation is large, reaching 75.672% and 197.974% respectively; the average rate of change of porosity is the smallest, at 3.839%; and among the six parameters, only the average rate of change of filtration resistance before and after intercalation is negative, so the filtration resistance is negatively correlated with intercalation or not.

ISSN: 2472-3703
DOI: 10.6911/WSRJ.202211_8(11).0027

5.3. Data Visualization and Analysis

In order to more intuitively and conveniently analyze the changes in structural variables and product performance after interpolation, this paper considers visualizing the experimental data before and after each group of interpolation and draws a line statistical graph as follows.

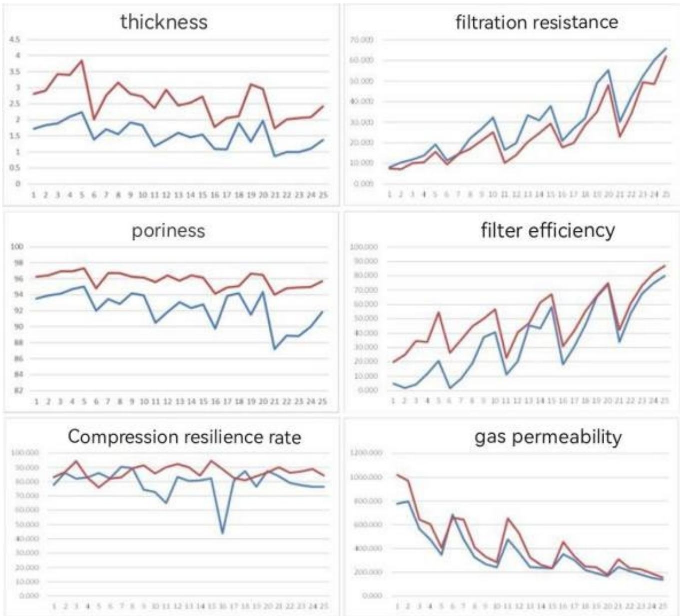

Figure 1. Structural variables, product performance line graph

According to the above figure, except for the compressive resilience, all the other five variables before and after interpolation have similar trends, and the parameter differences are almost constant, so the parameter change after interpolation should be a fixed value, and the slight fluctuations may be due to the process parameters or the interpolation rate; the change of the compressive resilience is obviously not fixed, so the correlation between this parameter and whether or not to interpolate is not obvious.

5.4. Spearman Correlation Coefficient Modeling and Solving

Spearman's rank correlation [2] is a method for studying the correlation between two variables based on rank information. It is based on the difference of each pair of ranks of two columns of paired ranks, Spielman rank correlation does not require data conditions as strict as the product-difference correlation coefficient, as long as the observations of the two variables are paired rank rating information, or rank information obtained from the transformation of observations of continuous variables, regardless of the overall distribution pattern of the two variables, the size of the sample size, can be used Spielman rank correlation to be studied at , with the following formula.

Step1: Calculation of correlation coefficient

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

In the above equation n is the number of samples and d_i is the grade difference between $y_i(j)$与$y_i'(j)$.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

Step2: Hypothesis testing

The following assumptions are made for the correlation coefficients

$$H_0: r_s = 0 \qquad H_1: r_s \neq 0 \tag{3}$$

Construct the statistic $r_s\sqrt{n-1} \sim N(0,1)$, calculate the test value $r_s\sqrt{n-1}$, and find the corresponding P-value compared to 0.05. The correlation coefficients and P-values were calculated as follows.

Table 3. Correlation coefficients

Correlation coefficient	thickness	porosity	Compression resilience	Filtration resistance	Filtration efficiency	permeability
interpolation rate	0.152	-0.148	-0.358	0.375	-0.03	-0.135
p-value	0.467	0.481	0.078	0.065	0.887	0.519

From the above table it can be seen that there is no parameter with p-value less than 0.05, so relaxed to 0.1, at this time there are two parameters with p-values meeting the requirements, 0.078 and 0.065, so the intercalation rate has a significant effect on the compression resilience and filtration resistance, and the intercalation rate is negatively correlated with the compression resilience with a correlation coefficient of -0.358, and positively correlated with the filtration resistance with a correlation coefficient of 0.375 .

6. MULTIPLE LINEAR REGRESSION MODELING

6.1. Data pre-processing

In order to fully and concisely represent the valid information of the data, we first pre-processed the annex data. For illustration purposes, we first set the receiving distance and hot air velocity to x_1, x_2 ; the thickness, porosity and compressive resilience to x_3, x_4, x_5 ; and the filtration resistance, filtration efficiency and permeability to y_1, y_2, y_3 .

6.2. Multiple Linear Regression

Step1: Build a multiple linear regression model

The model for the multiple linear regression analysis [3] was.

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \tag{4}$$

where β_0, β_1, β_2, $\cdots$, β_m is the biased regression coefficient, uncorrelated with $x_1, x_2, x_3, \cdots, x_m$, and ε is the random error term.

To investigate the relationship between the structural variables thickness, porosity, compressional resilience and the process parameters acceptance distance and hot air velocity, while our existing data acceptance distance and hot air velocity are quantitative indicators, so we do not need to create virtual indicators. The following table provides a general description of each indicator.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

Table 4. Description of variables

variable relationship	Variable Type	Variable Name	instructions
independent variable	Quantitative indicators	reception distance	Distance from the point of solution injection to the position of receiving the solution injected during the experiment. cm
		Hot air speed	Speed of the solution sprayed in air. r/min
implicit variable		thickness	Thickness of meltblown nonwoven material formed. mm
		porosity	The ratio of voids to the volume of material in a meltblown nonwoven material.
		Compression resilience	Resilience to compression. Greater elasticity indicates greater resistance to compression.

Descriptive statistics on these data yield

Table 5. Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Receiving distance cm	75	30	7.118685	20	40
Hot air speed r/min	75	1000	142.3737	800	1200
Thickness mm	75	2.607073	0.4722402	1.697782	3.527112
porosity	75	95.87915	0.820229	93.73874	97.28438
Compression resilience	75	86.60788	1.216714	83.51687	88.66131

We can see that there are 75 observations in each data set, and the standard deviations of thickness, porosity, and compressional resilience are small, and the difference between the maximum and minimum values is small, indicating that they do not vary significantly with receiving distance, and hot air velocity.

We assume that x_3, x_4, x_5 denotes thickness, porosity, and compressive resilience, respectively, and that the factors affecting them are the acceptance distance and hot air speed in the preparation of meltblown nonwoven materials for the intercalation layer x_1, x_2. And let there be a linear relationship between the dependent variables and their respective variables, then the linear overall regression model between them can be expressed as

$$\begin{cases} x_3 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_1 \\ x_4 = \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \varepsilon_2 \\ x_5 = \beta_0'' + \beta_1'' x_1 + \beta_2'' x_2 + \varepsilon_2 \end{cases} \tag{5}$$

where ε is the random error term $\varepsilon \sim N(0, \sigma^2)$.

Step2: Analysis of variance

An ANOVA was first conducted between the thickness and the two independent variables, setting the original hypothesis H_0: $\beta_1 = \beta_2 = 0$ as follows.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

Table 6. Analysis of variance

Source	SS	df	MS	F(2, 72)=1436.82
Model	16.0994	2	8.0497	Prob > F=0.000
Residual	0.4034	72	0.0056	R-squared = 0.9756
Total	16.5028	74	0.2230	Adj R-squared=0.9749

We obtained the joint significance test F(2, 72)=1436.82, p<0.001, so the original hypothesis is rejected, i.e., that β_1、β_2 is not all zero and SSR=16.0994, SSE=0.4034, SST=16.5028, according to the following equation.

$$R^2 = 1 - \frac{SSE}{SST} \tag{6}$$

$$SSE = \sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{7}$$

$$R^2_{adjusted} = 1 - \frac{SSE/（n-k-1）}{SST/（n-1）} \quad （k\text{为自变量的个数}） \tag{8}$$

We get $R^2 = 0.9756$. $R^2_{\text{adjusted}} = 0.9749$, the better the fit is when we introduce more independent variables, we tend to use the adjusted goodness of fit, if the reduction of the SSE by the introduced independent variables is particularly small, then the integrated goodness of fit will be small instead, and our goodness of fit is close to 1 before and after integration, which indicates that the regression line we designed fits the observations better, and our regression is a predictive regression, theR^2 The higher precision of indicates that we have better prediction afterwards.

Similarly, we do an ANOVA on porosity and compressive resilience and get the following table

Table 7. Analysis of variance

	Source	SS	df	MS	F(2, 72)=148.97
porosity	Model	40.0956	2	20.0480	Prob > F=0.000
	Residual	9.6845	72	0.1346	R-squared = 0.8054
	Total	49.7854	74	0.6728	Adj R-squared=0.8000
Compression resilience	Source	SS	df	MS	F(2,72)=12.96
	Model	28.9986	2	14.4993	Prob > F=0.000
	Residual	80.5505	72	1.1188	R-squared = 0.8647
	Total	109.5491	74	1.4804	Adj R-squared=0.8443

We obtained before and after integration of R^2 0.8504, 0.8000 and 0.8647, 0.8443, respectively. Again, since the goodness of fit after integration is close to 1, we can state that the regression line we designed fits the observations well and the adjustment before and after R^2 difference is not significant and the effect of multicollinearity can be disregarded.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

However, while the joint significant tests are all P<0.001, we reject the original hypothesis at the 99% confidence level. That is, β_1、 β_2 is not all zero.

Step3: Heteroskedasticity test

We first used Stata software to plot a scatter plot between the residuals and the fitted values, but intuitively we had difficulty observing the presence of heteroskedasticity.

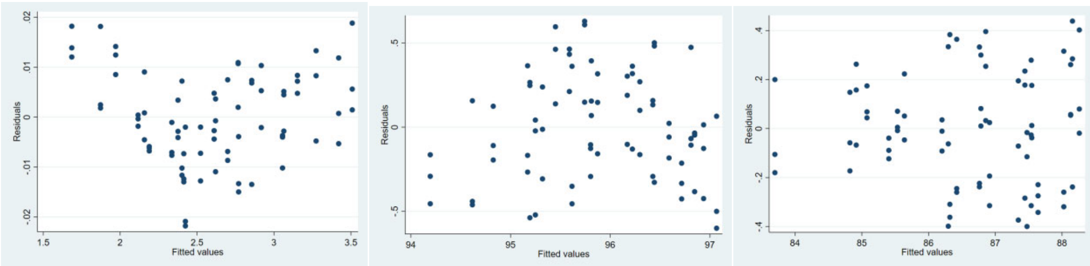

Figure 2. Scatter plot of residuals and fitted values for the standard regression model

To determine more precisely whether heteroskedasticity exists, we performed White tests on the standard regression model to obtain $\chi^2(5)=2.49$, p=0.7778>0.05; $\chi^2(5)=2.49$, p=0.063>0.05; and $\chi^2(5)=2.31$, p=0.8051>0.05, respectively. none of the original hypotheses could be rejected at the 95% confidence level, so it can be assumed that there is no heteroskedasticity.

Step4: VIF test for multicollinearity

To test for the presence of multicollinearity we use the variance inflation factor VIF.

$$VIF_i = \frac{1}{1-R^2{}_{1-k|i}} \tag{9}$$

$R^2{}_{1-k|i}$ is the goodness-of-fit obtained by regressing the ith independent variable as the dependent variable on the remaining k-1 independent variables. Define the regression model $VIF=\max\{VIF_1,VIF_2,VIF_3,VIF_4,VIF_5\}$, if $VIF>10$, then it can be considered that there is multicollinearity. The results of the test are as follows.

Table 8. VIF test results

Variable	VIF	1/VIF
reception distance	1	1
Hot air speed	1	1
Mean VIF	1	

It follows from Mean VIF=1<10 that the regression model does not have the effect of multicollinearity. Therefore we can directly perform the standardized regression.

Step5: Standardized regression model solving

To explore the effect of each variable on the dependent variable, we used standard regression.

Brief description of the principle: all variables x_1, x_2 were put into the model, after which an attempt was made to remove one independent variable from the model to see if there was a

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

significant change in the variance of the dependent variable explained by the whole model, while the one with the least explanatory power was removed; and the process was repeated until there were no independent variables that could be removed. The results of the standardized regression were then obtained as follows.

Table 9. Regression table

	Thickness mm	porosity	Compression resilience
Receiving distance cm	0.054***	0.084***	-0.069***
	(44.336)	(14.042)	(-3.984)
Hot air speed r/min	0.002***	0.003***	-0.003***
	(30.133)	(10.039)	(-3.170)
_cons	-0.860***	90.349***	91.410***
	(-11.982)	(256.770)	(90.101)
N	75	75	75

*** p<0.01 ** p<0.05 * p<0.1

Taking thickness as an example, each unit increase in receiving distance can increase the thickness by 0.054 units, controlling for other variables constant, and similarly, each unit increase in hot air velocity can increase the thickness by 0.002 units, controlling for other variables constant, from which three regression equations can be derived as

$$\begin{cases} x_3 = 0.054x_1 + 0.002x_2 - 0.86 + \varepsilon \\ x_4 = 0.084x_1 + 0.003x_2 + 90.349 + \varepsilon \\ x_5 = -0.069x_1 - 0.003x_2 + 91.41 + \varepsilon \end{cases} \tag{10}$$

From the three equations above, it can be seen that the receiving distance and hot air velocity are positively correlated with thickness and porosity and negatively correlated with compressional resilience, with the receiving distance contributing a greater amount of evaluation than the hot air velocity.

Step6: Prediction of the model

Based on the above three regression equations, we can predict the thickness, porosity and compressive resilience for receiving distances of 23cm, 28cm, 33cm and 38cm and hot air speeds of 850, 950, 1150 and 1250r/min, respectively. The predicted results are shown in the following table.

Table 10. Projected results

Receiving distance (cm)	Hot air speed(r/min)	Thickness mm	Porosity (%)	Compression resilience (%)
38	850	2.892	96.091	86.238
33	950	2.822	95.971	86.283
28	1150	2.952	96.151	86.028
23	1250	2.882	96.031	86.073
38	1250	3.692	97.291	85.038
33	1150	3.222	96.571	85.683
28	950	2.552	95.551	86.628
23	850	2.082	94.831	87.273

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

7. APPLICATION OF MULTIPLE LINEAR REGRESSION MODELS WITH CROSS-CORRELATION

7.1. Correlation Test

To further explore the relationship between the variables, Spearman's correlation coefficient was calculated using spss for the structural variable parameters and the results were calculated as follows.

Table 11. Table of Spearman's correlation coefficients

	Thickness mm	Porosity (%)	Compression resilience (%)
Thickness mm	1.000	0.921**	-0.473**
Porosity (%)	0.921**	1.000	-0.355**
Compression resilience (%)	-0.473**	-0.355**	1.000

**. Significant correlation at the 0.01 level (two-tailed).

	Filter resistance Pa	Filtration efficiency (%)	Air permeability mm/s
Filter resistance Pa	1.000	0.789**	-0.648**
Filtration efficiency (%)	0.789**	1.000	-0.778**
Air permeability mm/s	-0.648**	-0.778**	1.000

**. Significant correlation at the 0.01 level (two-tailed).

From the above table, we can see that thickness is positively and strongly correlated with porosity, negatively correlated with compression resilience, and porosity is negatively correlated with compression resilience; filtration resistance is positively correlated with filtration efficiency, negatively correlated with air permeability, and filtration efficiency is negatively correlated with air permeability.

7.2. Multiple Linear Regression Models Considering Interaction Terms

Interaction terms are interactions between multiple eigenvalues and are not equivalent to correlations. Interaction terms can be judged as significant or not by p-value. Based on the study in 7.1, the following interaction terms are considered in building the multiple linear regression model: $x_3 * x_4$, $x_3 * x_5$, $x_4 * x_5$ and $x_3 * x_4 * x_5$, and are denoted as x_6、 x_7、 x_8 respectively. Taking $y1$、 $y2$、 $y3$ as the dependent variable and , , x_3 x_4 x_5 and x_6、 x_7、 x_8 as independent variables for multiple linear regression, the regression results are presented in the following table.

Table 12. Table of regression coefficient analysis

		x3	x4	x5	x6	x7	x8	_cons
y1	Coef.	317.786	-14.003	-12.119	-1.416	-2.180	0.192	844.514
	Std. Err.	167.374	78.272	82.433	1.708	0.735	0.871	7417.952
y2	Coef.	966.234	169.521	137.238	-13.319	3.398	-1.581	-14911.480
	Std. Err.	243.363	124.183	130.727	2.918	1.985	1.410	11503.820
y3	Coef.	-8649.942	239.680	424.712	72.892	20.210	-4.858	-19244.530
	Std. Err.	3206.497	1415.828	1487.194	34.914	22.863	16.061	130954.500

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

The three models were calculated to be 0.816, 0.842, and 0.837 respectively by R^2. A White's test for heteroskedasticity was performed on the regressed models, and it was concluded that heteroskedasticity existed in all three models, so the originally constructed statistic failed and the hypothesis test for the model was temporarily unavailable, and the standard errors are now adjusted, i.e., the statistic and hypothesis test are constructed again using the OLS + robust standard error approach with the following p-values.

Table 13. Table of p-values for hypothesis testing

models		t	P>t	lower bound of the confidence interval	Upper confidence interval
	x3	1.900	0.062	-16.204	651.775
	x4	-0.180	0.859	-170.191	142.186
y1	x5	-0.150	0.884	-176.611	152.373
	x8	0.220	0.826	-1.545	1.929
	_cons	0.110	0.910	-13957.780	15646.810
	x5	1.050	0.298	-123.623	398.099
	x6	-4.570	0.000	-19.141	-7.497
y2	x7	1.710	0.091	-0.563	7.359
	x8	-1.120	0.266	-4.396	1.234
	_cons	-1.300	0.199	-37867.010	8044.041
	x3	-2.700	0.009	-15048.410	-2251.477
	x4	0.170	0.866	-2585.560	3064.920
	x5	0.290	0.776	-2542.939	3392.362
y3	x6	2.090	0.041	3.222	142.563
	x7	0.880	0.380	-25.412	65.833
	x8	-0.300	0.763	-36.907	27.191
	_cons	-0.150	0.884	-280560.200	242071.100

In summary although the model's R^2 is large and the model as a whole passes the F-test, it can be seen from the table above that there are more t-tests for the regression coefficients that are not significant, so consider whether this is due to the effect of multicollinearity, the following is calculated for the model with the effect of multicollinearity VIF as shown in the table below.

Table 14. VIF test results

Variable	VIF	1/VIF
x8	78876.77	0.000013
x5	75567.59	0.000013
x3	40593.57	0.000025
x4	30061.29	0.000033
x6	21502.46	0.000047
x7	18444.63	0.000054

Obviously, the VIF in the table is much larger than 10, indicating the existence of serious multicollinearity among the six parameters. For the multicollinearity, the stepwise regression method is chosen to solve it in this paper.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

7.3. Backward Stepwise Regression Model

In contrast to forward stepwise regression, all variables were first put into the model, after which an attempt was made to remove one of the independent variables from the model to see if there was a significant change in the variance of the dependent variable explained by the whole model, after which the one with the least explanatory power was removed; this process was iterated until no independent variable qualified for removal. In this paper, the stepwise regression operation was implemented using stata software, α taking 0.1, and to avoid heteroskedasticity, robust standard errors were used directly, resulting in the following regression model.

Table 15. Stepwise regression coefficient analysis table

		Coef.	Std. Err.	t	P>t	confidence limit (math.)	confidence limit
	x3	153.540	56.402	2.720	0.008	41.050	266.031
	x4	-5.334	1.288	-4.140	0.000	-7.902	-2.765
y1	x8	0.058	0.020	2.940	0.004	0.019	0.098
	x7	-1.859	0.664	-2.800	0.007	-3.182	-0.535
	_cons	76.384	94.109	0.810	0.420	-111.311	264.078
	x3	1277.678	290.924	4.390	0.000	697.448	1857.908
	x4	38.854	8.177	4.750	0.000	22.546	55.162
y2	x8	-0.057	0.014	-4.100	0.000	-0.084	-0.029
	x6	-13.523	3.035	-4.460	0.000	-19.577	-7.469
	_cons	-3148.425	673.113	-4.680	0.000	-4490.906	-1805.945
	x3	-6467.441	3525.740	-1.830	0.071	-13499.310	564.425
	x8	-1.864	0.971	-1.920	0.059	-3.801	0.073
y3	x5	191.003	108.493	1.760	0.083	-25.380	407.386
	x6	68.153	36.765	1.850	0.068	-5.173	141.479
	_cons	-827.366	1564.609	-0.530	0.599	-3947.881	2293.148

The p-value of each regression coefficient is less than 0.1 and passes the hypothesis test, then the multiple linear regression model can be expressed as

$$\begin{aligned} y_1 &= 76.384 + 153.54x_3 - 5.334x_4 + 0.058x_4x_5 - 1.859x_3x_5 \\ y_2 &= -3148.425 + 1277.678x_3 + 38.854x_4 - 0.057x_4x_5 - 13.523x_3x_4 \\ y_3 &= -827.366 - 6467.441x_3 - 1.864x_4x_5 + 191.003x_5 + 68.153x_3x_4 \end{aligned} \tag{11}$$

The goodness of fit of the three models are 85.3%, 83.7% and 84.9%, which are close to 1. This indicates that the above models can reflect the relationship between structural variables and product performance better and can be used to calculate the best performance.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

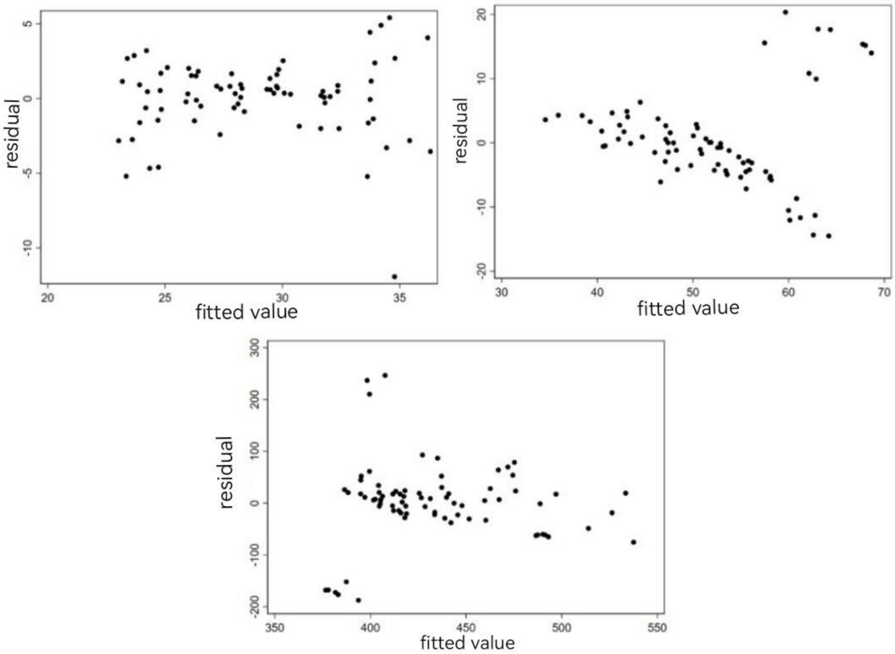

Figure 3. Residual plots

From the residual plots, it can be seen that the maximum values of the residuals are around 5, 19 and 220, respectively, and the standardized residuals are concentrated in the distribution near the straight line, which can be considered that the standardized residuals satisfy the normal distribution and the errors pass, and the established regression equations are reasonable and have small errors.

7.4. Optimization Modeling of Filtration Efficiency

Using the regression equations derived above for process parameters versus structural variables and structural variables versus product performance, a regression equation for process parameters versus product performance can be derived, which can be used to develop a single-objective planning model to obtain the highest filtration efficiency.

1) Decision variables: independent variables of the filter efficiency regression equation x_1, x_2

2) Objective function.

$$\begin{cases} x_3 = 0.054x_1 + 0.002x_2 - 0.86 \\ x_4 = 0.084x_1 + 0.003x_2 + 90.349 \\ x_5 = -0.069x_1 - 0.003x_2 + 91.41 \\ y_2 = -3148.425 + 1277.678x_3 + 38.854x_4 \\ \qquad -0.057x_4x_5 - 13.523x_3x_4 \end{cases} \tag{12}$$

$$\begin{aligned} \max\ y_3 = & -156.820759 - 0.06101041x_1^2 - 0.00008114x_2^2 \\ & -0.00443643x_1x_2 + 7.14601954x_1 + 0.26304686x_2 \end{aligned} \tag{13}$$

3) Constraint: the range of values of the independent variable does not exceed the maximum range in data3

①Receiving distance constraint.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

$$20 \leqslant x_1 \leqslant 40 \tag{14}$$

② Hot air speed constraint.

$$800 \leqslant x_2 \leqslant 1200 \tag{15}$$

In summary, the filtration efficiency is modeled as follows.

$$\begin{aligned} \max\ y_3 = & -156.820759 - 0.06101041x_1^2 - 0.00008114x_2^2 \\ & - 0.00443643x_1x_2 + 7.14601954x_1 + 0.26304686x_2 \end{aligned} \tag{16}$$

$$s.t. \quad \begin{cases} 20 \leqslant x_1 \leqslant 40 \\ 800 \leqslant x_2 \leqslant 1200 \end{cases} \tag{17}$$

The above model was solved using Matlab software and the results obtained are shown in the following table.

Table 16. Solution results

Receiving distance (cm)	Hot air speed(r/min)	Thickness mm
20	1195	2.4077
Porosity (%)	Compression resilience (%)	Filtration efficiency (%)
95.740	85.203	86.743

From the data in the above table, it can be seen that the highest filtration efficiency of 86.743% was achieved when the receiving distance was 20 cm and the hot air speed was 1195 r/min, at which time the structural variables parameters were 2.4077 mm thickness, 95.740% porosity and 85.203% compression resilience.

8. APPLICATION OF THE BI-OBJECTIVE PLANNING MODEL

8.1. Model Building and Solving

The problem is to find the optimal combination of receiving distance and hot air velocity, and to make high filtration efficiency with minimum filtration resistance. By combining the regression equation between structural variables and product performance established in the third question with the regression equation between process parameters and structural variables obtained in the second question, we can obtain a dual-objective planning model between minimizing filtration resistance and maximizing filtration efficiency. The extreme value points of the dual-objective planning [4] are searched for the optimal combination of receiving distance and hot air velocity.

8.2. Development of a bi-objective Planning Model

8.2.1 Planning models for filtration efficiency

A regression fit of the structural variables thickness, porosity, compressive resilience and filtration efficiency yields the regression equation for maximizing filtration efficiency.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

$$\max\ f_1 = -3148.425 + 1277.678x_3 + 38.854x_4 - 0.057x_4x_5 - 13.523x_3x_4 \tag{18}$$

By looking at the coefficients of the decision variables it is easy to see that porosity has the greatest effect on filtration efficiency, while the interaction term between porosity and compressive resilience has the least effect on filtration efficiency.

In order to take into account the requirements of all conditions, we make the receiving distance less than 100 cm, the hot air speed less than 2000 r/min, the thickness less than 3 mm, and the compressional resilience not less than 85%, i.e., the constraints are.

$$s.t.\ \begin{cases} 0 < x_1 \leqslant 100 \\ 0 < x_2 \leqslant 2000 \\ 0 < x_3 \leqslant 3 \\ 0 < x_4 \leqslant 100 \\ 85 < x_5 < 100 \\ 0.054x_1 + 0.002x_2 - x_3 = 0.86 \\ 0.084x_1 + 0.003x_2 - x_4 = -90.349 \\ 0.069x_1 + 0.003x_2 + x_5 = 91.41 \end{cases} \tag{19}$$

The above constraints give both the relationships between product thickness, porosity, compressive resilience and receiving distance, and hot air velocity, and the process parameters can be combined with the filtration efficiency by searching for the extreme value point of this function to find the process parameters at maximum filtration performance through the regression equation fitted in the second question between the process parameters and the structural variables.

8.2.2 Planning model for filtration resistance

The regression equation fitted by the third question between the filtration efficiency and the structural variables shows easily that thickness has the greatest effect on the filtration resistance, followed by porosity and, to a lesser extent, the interaction term between porosity and compressive resilience. The regression equation allows the objective function between the minimized filtration resistance and the structural scalar to be listed as.

$$\min\ f_2 = 76.384 + 153.54x_3 - 5.334x_4 + 0.058x_4x_5 - 1.859x_3x_5 \tag{20}$$

and the above constraints between the decision variables are satisfied.

8.2.3 Bi-objective planning model

In order to maximize the filtration efficiency while also weighing the effect of filtration resistance, a dual-objective planning model of filtration efficiency-filtration resistance is further developed in this paper. Based on the above notational definition, the model is as follows.

Decision variables: receiving distance x_1 , hot air speed x_2 , thickness x_3 , porosity x_4 , compressional resilience x_5 .

Objective function.

$$\begin{gathered} \min\ f_1 = 3148.425 - 1277.678x_3 - 38.854x_4 + 0.057x_4x_5 + 13.523x_3x_4 \\ \min\ f_2 = 76.384 + 153.54x_3 - 5.334x_4 + 0.058x_4x_5 - 1.859x_3x_5 \end{gathered} \tag{21}$$

Constraints: The five independent variables take values in a range that does not exceed the conditions and requirements needed for actual production.

Receiving distance constraint.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

$$0 < x_1 \leqslant 100 \tag{22}$$

② Hot air speed constraint.

$$0 < x_2 \leqslant 2000 \tag{23}$$

Thickness constraint.

$$0 < 0.054x_1 + 0.002x_2 - 0.86 \leqslant 3 \tag{24}$$

④ Porosity constraint.

$$0 < 0.084x_1 + 0.003x_2 + 90.349 < 100 \tag{25}$$

⑤ Compressional resilience constraint.

$$85 \leqslant 91.41 - 0.069x_1 - 0.003x_2 \leqslant 100 \tag{26}$$

In summary, the following bi-objective planning model was developed.

$$\begin{aligned} \min\ f_1 &= 3148.425 - 1277.678x_3 - 38.854x_4 + 0.057x_4x_5 + 13.523x_3 \\ \min\ f_2 &= 76.384 + 153.54x_3 - 5.334x_4 + 0.058x_4x_5 - 1.859x_3x_5 \end{aligned} \tag{27}$$

$$s.t. \begin{cases} 0 < x_1 \leqslant 100 \\ 0 < x_2 \leqslant 2000 \\ 0 < x_3 \leqslant 3 \\ 0 < x_4 \leqslant 100 \\ 85 < x_5 < 100 \\ 0.054x_1 + 0.002x_2 - x_3 = 0.86 \\ 0.084x_1 + 0.003x_2 - x_4 = -90.349 \\ 0.069x_1 + 0.003x_2 + x_5 = 91.41 \end{cases} \tag{28}$$

8.3. Solving the bi-objective Planning Model

To solve the optimal value of the dual-objective programming model, we transform it into a single-objective programming model, that is, the above two objective functions are linearly weighted, because in the actual processing production, we prefer to produce the filter efficiency of the mask, so we can set the weights of f_1 and f_2 to 0.6 and 0.4 respectively. Also note that the units of the two objective functions of the function are different, so we need to first carry out the objective function Standardization to remove the influence of the magnitude, and then weighting, we can use the average value of the filtration resistance and filtration efficiency in data3 as the reference value for standardization, that is, the two objective functions are divided by 52.03 and 28.90 respectively. then the weighted combined objective function is

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202211_8(11).0027

$$\min\ f = \frac{0.6 f_1}{52.03} + \frac{0.4 f_2}{28.90} \tag{29}$$

Then we transform into a single objective planning problem, which is solved using the Matlab toolbox to obtain

Table 17. Results of dual-objective planning

Receiving distance (cm)	Hot air speed(r/min)	Filter resistance Pa	Filtration efficiency (%)
55.8	2000	32.93	94.09

When the receiving distance is 55.8 cm and the hot air speed is 2000 r/min, the maximum filtration efficiency of 94.09% can be ensured, while the filtration resistance is ensured to be as small as possible at 32.93 Pa. Since in the actual production process, it is preferred to produce a larger filtration efficiency, the planning results of our solution are more favored to consider a larger filtration efficiency.

8.4. Sensitivity Analysis

Sensitivity analysis [5] is an uncertainty analysis technique that examines the extent to which a certain change in the relevant factors affects a key indicator or set of key indicators from the perspective of quantitative analysis. In essence, it is a method of explaining the pattern of the size of the key indicators affected by changes in these factors by changing the values of the relevant variables one by one.

We explore the effect of the weights on the results by changing the weights of f_1 and f_2 (since the sum of the weights is one, it is sufficient to change only one of the weights) Here we change the weights of f_2 and $w_2 \in (0.1, 0.5)$. The following figure shows the sensitivity analysis plot made by Matlab based on changing the weights.

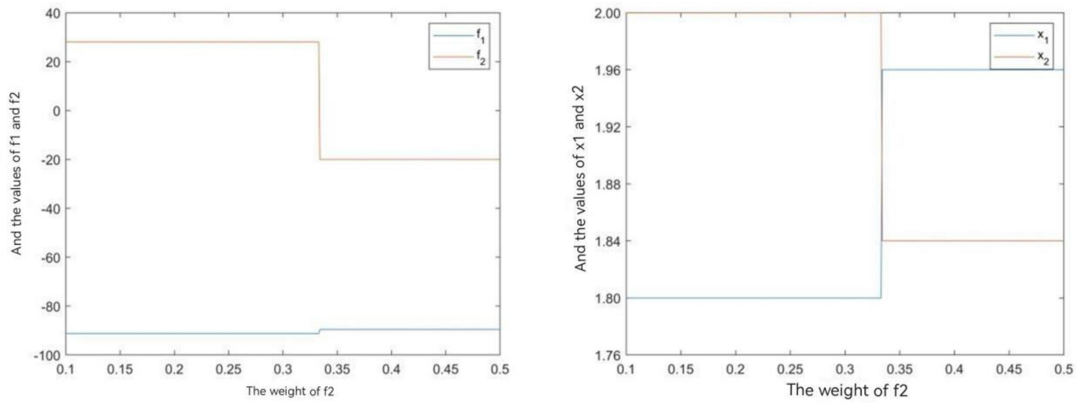

Figure 4. Sensitivity analysis graph by changing the weights

From the figure, it can be seen that the turning point of the weight of f_2 is about 0.333, when the weight of f_2 is less than this turning point, the values of x_1 and x_2 are taken as 1.84 and 2 respectively (after normalization), and when it is greater than the turning point, the values of x_1 and x_2 change more, the main reason is that when the weight of f_2 is less than 0.333, then the weight for the filtration efficiency becomes bigger, then the manufacturer will pay more attention to the production of mask's filtration efficiency, even though the filtration resistance of the production mask has become larger.

9. MODEL EVALUATION, IMPROVEMENT AND REPLICATION

9.1. Evaluation of the Model

Multiple linear regression model It is a simple and effective mathematical model that has been widely used in various fields. It has the advantage of being fast to model, does not require very complex calculations, and still runs fast with large amounts of data. The understanding and interpretation of each variable can be given based on the coefficients. And he also has the disadvantage of not fitting the nonlinear data well. So it needs to determine whether the variables are linear to each other first.

9.2. Extension of the Model

In addition to applying re this question, a multiple linear regression model can be used to predict customer value based on multiple factors, and once the model is built, different business strategies can be applied to customers of different values.

REFERENCES

[1] Han Ling,Hu Mengyuan,Ma Yingbo,Hao Donglian. Current status of research on medical nonwoven mask materials and their new technologies[J]. Journal of Xi'an Engineering University, 2020, 34(02): 20-25.DOI:10.13338/j.issn.1674-649x.2020.02.003.

[2] Jia Ke, Yang Zhe, Wei Chao, Zheng Liming, Li Yanbin, Bi Tianshu. Longitudinal protection of new energy transmission lines based on Spearman's rank correlation coefficient[J]. Power System Automation, 2020, 44(15):103-111.

[3] Li P, Mao Q, Wang X, Wang XD, Wang XY, Song T, Guo R. Establishment and application of operational cost prediction model based on multiple linear regression analysis[J]. Petroleum Planning and Design,2018,29(03):33-37+52.

[4] Yiming Feng,Lie Xiao,Dinghao Zheng. Research on credit strategy of micro and small enterprises based on dual-objective planning[J]. Mall Modernization, 2020 (20): 111-113.DOI:10.14013/j.cnki.scxdh.2020.20.042.

[5] Wu, Shu-Chen, Qi, Zong-Feng, Li, Jian-Xun. Intelligent global sensitivity analysis based on deep learning[J]. Journal of Shanghai Jiaotong University, 2022, 56 (07): 840-849.DOI:10.16183/j.cnki.jsjtu.2021.191.