2023 2nd International Conference on Health Big Data and Intelligent Healthcare (ICHIH) | 979-8-3503-0932-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICHIH60370.2023.10396000

Optimization of Performance Models for Interlayer Meltblown Nonwovens with Random Forest

PengJie Wang
Zhuhai College of Science and Technology
Alibaba Cloud College of Big Data Application
Zhuhai, China
1297276320@stu.zcst.edu.cn

Ya Li
Information Technology Application Innovation Platform Department YGSOFT INC.
Zhuhai, China
liya2@ygsoft.com

YuHeng Liang
Zhuhai College of Science and Technology
Alibaba Cloud College of Big Data Application
Zhuhai, China
Liangyuheng@stu.zcst.edu.cn

Xiaolin Zhu
Zhuhai College of Science and Technology
Alibaba Cloud College of Big Data Application
Zhuhai, China
* Corresponding author: fdsfsdxno@foxmail.com

TongFei Li
Zhuhai College of Science and Technology
Alibaba Cloud College of Big Data Application
Zhuhai, China
litongfei@zcst.edu.cn

Wei Lv
Zhuhai College of Science and Technology
Alibaba Cloud College of Big Data Application
Zhuhai, China
luwei@zcst.edu.cn

***Abstract*—In the context of rapid societal and technological advancements, there has been a pronounced elevation in public health awareness. Among various protective measures, meltblown nonwoven materials have emerged as an indispensable component in protective masks, offering a crucial respiratory barrier for users. This study endeavors to provide an in-depth examination of this pivotal material. By systematically analyzing extensive datasets of meltblown fabric manufacturing processes and taking into account key parameters such as receiving distance, air velocity, and material thickness, this paper, grounded on historical production parameters, harnesses advanced machine learning techniques to meticulously design and optimize the performance control model for interlayer melt-blown nonwoven materials. The central aim of this research is to further enhance the efficacy of this material in health protection applications.**

Keywords-component; Interlayer meltblown, nonwoven materials, PCA Dimensionality Reduction, Multivariate Nonlinear Regression, Multi-objective Linear Programming, Random Forest, Neural Networks.

I. INTRODUCTION

Interpolated meltblown nonwoven materials[1] are versatile and widely used in a variety of applications, and their properties have a direct impact on the quality and performance of the final product. In order to meet the demands of different applications, researchers and manufacturers have been seeking an effective way to control and optimize the properties of these materials. In this context, this study aims to design and optimize a performance control model for intercalated meltblown nonwoven materials to meet the demands of different applications and improve their competitiveness.

Interpolated meltblown nonwoven materials have a wide range of applications including, but not limited to, medical supplies, filtration materials, automotive interiors, construction materials and packaging industries. Their wide range of applications is due to their many advantageous properties such as lightweight, abrasion resistance, filtration performance and customizability. However, in order to achieve optimal performance in these various applications, various key performance parameters such as mechanical properties, filtration efficiency, air permeability and chemical stability must be precisely controlled.

Currently, the performance control of intercalated meltblown nonwoven materials relies heavily on empirical and experimental methods, which limits the optimization and customization of their performance. With advances in materials science and computational methods, there is an opportunity to develop model-based approaches for predicting and optimizing material properties, thereby accelerating the development and marketing of new products. This study aims to fill the gap in this research area by providing an innovative approach for performance control of intercalated meltblown nonwoven materials.

Given the critical position of intercalated meltblown nonwoven materials in modern industry and the urgent need for performance control, this study has received much attention. By developing an effective performance control model, we can better meet the needs of different applications, improve material customizability, and reduce production costs. This will help drive technological advances in the field of nonwoven materials and provide innovative solutions for sustainable development.

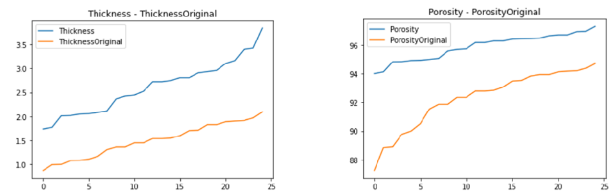
To solve the relationship between structural variables and product properties and itself for the interpolated meltblown nonwoven materials, principal component analysis in data dimensionality reduction[2] is used and transformed into digital visualization. Secondly, a nonlinear regression model[3] was established, with filtration efficiency as the dependent variable and the rest of the indicators as the independent variables, to derive a nonlinear regression equation, which gives a functional expression for the filtration efficiency of the product. Then, based on the least squares method, the model is solved by Spss to get the objective function[4] of filtration resistance. This question uses multi-objective optimization, which has two objective functions: maximization of filtration efficiency and minimization of filtration resistance. The multi-objective function is converted into one objective function by using positive/negative indicators, with the constraints of receiving distance not more than 100cm, hot air speed not more than 2000 r/min, thickness not more than 3mm, and compression resilience not less than 85%, and then a linear programming model is established and solved by the Random Forest algorithm[5] to obtain the solution. After setting the parameter range, the process parameters are: receiving speed 22cm, hot air speed 2000r/min.

II. METHOD ANALYSIS

A. Material structure analysis based on interpolated melt-blown nonwoven materials

In this study, we focus on key issues in the production and preparation of intercalated meltblown nonwoven materials.

For the influence of material structure variables: By analyzing the line graph before and after intercalation in Figure 1, we found that the impact on thickness and porosity parameters is greater after intercalation, resulting in a significant increase in their amplitude. As for the compression resilience, the impact after intercalation is greater in the early stage, smaller in the later stage, and shows a gradually increasing trend.

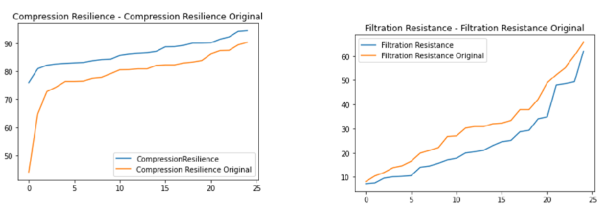


(a) Thickness change before and after intercalation

(b) Changes in porosity before and after insertion

Fig.1 Insert thickness and porosity Before and after change

For changes in product performance: Analyzing the line graphs before and after insertion in Fig. 2, it can be concluded that the effect on filtration resistance and filtration efficiency after insertion is relatively small, resulting in smaller filtration resistance and higher filtration efficiency. In terms of air permeability, the effect of the layers is also small, leading to a slight increase in air permeability.

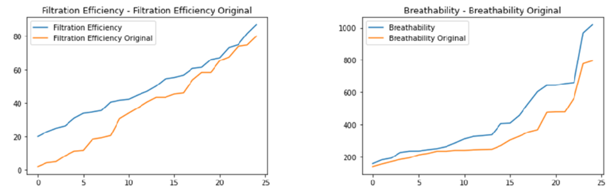


(a) plug the layer before and after the compression resilience rate of change

(b) Change in filtration resistance before and after insertion of layers

Fig.2 Insertion Resilience and Filtration Resistance

Correlation analysis of insertion rate Fig. 3: Correlation analysis reveals the relationship between different parameters. There is a significant positive correlation between porosity and thickness, a significant positive correlation between filtration resistance and filtration efficiency, a significant negative correlation between filtration resistance and air permeability, and a significant negative correlation between filtration efficiency and air permeability. These results contribute to a better understanding of the mechanisms by which insertion affects product performance.



(a) Change in filtration efficiency before and after interpolation

(b) Change in air permeability before and after interlayering

Fig.3 Insert filtration efficiency and permeability

B. Data distribution of structural variables over process parameters

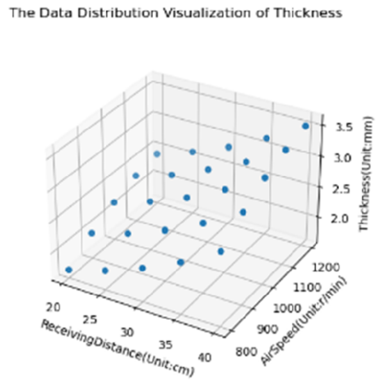


Fig.4 Thickness data distribution

Quantifying the effect of factor interactions through the thickness analysis data plots in Fig. 4, a linear regression model can be considered using the following equations (1)(2), with thickness (d) and porosity (ø) as the regression variables, and receiving distance (l) and hot air velocity (v) as the response variables, respectively [6]. Here, the distribution of thickness and porosity over the receiving distance and hot air velocity can be approximated as a plane.

$$d = a_1 \times l + a_2 \times V + W_1 \quad (1)$$

$$\phi = b_1 \times b_2 \times V + W_2 \quad (2)$$

where w_1 and w_2 are constant terms and a_1, a_2, b_1, b_2 are the coefficients of the linear terms of the independent variables.

C. Linear regression correlation formula

1) Beta(3)(standardized regression coefficient calculation):

$$\beta_j = \frac{b_j}{\sigma_x \sqrt{VIF_j}} \tag{3}$$

2) t = regression coefficient/standard error of regression coefficient; t = constant term/standard error of constant term.

3) VIF (4)(Variance Inflation Factor):

$$VIF_j = \frac{1}{1 - R_j^2} \tag{4}$$

4) Adjust the R-squared formula(5):

$$R^2 = \left\{1 - \left[\frac{(1-R^2)(n-1)}{(n-k-1)}\right]\right\} \tag{5}$$

5) F Final value(6):

$$F = \frac{\Sigma(\hat{y}-\bar{y})^2/p}{\Sigma(y-\hat{y})^2/(n-p-1)} \tag{6}$$

From the analysis of the results of the F-test [7] in Table Ⅰ, it can be seen that the significance P-value is 0.000****, which is significant at the level, rejecting the original hypothesis that the regression coefficient is 0, and the model basically meets the requirements For the covariate performance of the covariates, the VIF [8] is less than 10, and the model does not have the problem of multicollinear variables, and the model is well-established.

TABLE I.Table of results of linear regression analysis where the dependent variable is thickness

Linear regression analysis results n = 75									
	Standardized coefficient		Standardized factor	1	p	R^2	VIF	Adjust R^2	F
	B	Standard error	Beta						
constant	0.86	0.072	-	11.982	0.000***	-	0.976	0.975	F=1436822 P=0.000***
Receiving Distance (cm)	0.054	0.001	0.817	44.336	0.000***	1			
Thermal air speed (r/min)	0.002	0	0.555	30.133	0.000***	1			
Dependent variable: thickness mm									

III. MODEL BUILDING AND SOLVING

Since it is necessary to solve for the optimal solution when the product's filtration efficiency is the highest by choosing the values of the receiving distance and the hot air speed, a nonlinear regression model is considered. Taking the filtration efficiency as the dependent variable and the rest of the indicators as independent variables, the nonlinear regression equation was established to obtain the optimal regression equation for the filtration efficiency of the product by solving the equation using the principal component analysis method in the data dimensionality reduction. The relationship between structural variables and product performance, as well as the relationship between the two variables themselves is analyzed using data dimensionality reduction and data visualization to form a chart, so that the relationship is described more intuitively. The nonlinear regression model[9] is solved to obtain the highest filtration efficiency of the product and the corresponding factor values.

A. Principal Component Analysis (PCA) model for data dimensionality reduction

PCA, an algorithm for data dimensionality reduction. Principal Component Analysis, a linear combination of multiple indicators with some correlation, with the goal of dimensionality reduction by explaining as much information as possible in the original data in the least number of dimensions, after dimensionality reduction of the variables are linearly independent of each other, and the finalized new variables are linear combinations of the original variables, and the further back the principal components in the variance is also a small proportion of the variance, and the weaker the ability to synthesize the original information.

1) Modeling steps

Based on research questions, indicators and data are selected. Following this, the indicator data is standardized, a process which is automatically executed by the Factor function in the SPSS software. Subsequently, the correlation between the indicators is determined. The number of principal components, denoted as m, is then ascertained. After determining the expression for the principal component, these components are named. Finally, the comprehensive principal component value is calculated, upon which further evaluation and research are conducted.

a) Features are normalized [10](Feature Normalization):

For each feature x in the dataset its normalized representation(7) is:

$$x_i' = \frac{x_i - \mu_i}{\sigma_i} \tag{7}$$

Where μ_i is the mean of the feature and σ_i is its standard deviation.

b) Calculate the dimensionality reduction matrix(8)

First calculate the covariance matrix of the sample features. Matrixized calculation is used.

$$C_{ov} = \frac{1}{m}\sum_{i=1}^{n}\left(x^{(i)}\right)\cdot\left(x^{(i)}\right)^{T}$$

$$X = \begin{bmatrix} \left(X^{(1)}\right)^{T} \\ \vdots \\ x^{(m)})^{T} \end{bmatrix} = \begin{bmatrix} X_1^{(1)} & X_2^{(1)} & \dots & X_n^{(1)} \\ X_1^{(2)} & X_2^{(2)} & \dots & X_n^{(2)} \\ X_1^{(m)} & X_2^{(m)} & \dots & X_n^{(m)} \end{bmatrix} \quad (8)$$

$$C_{ov} = \frac{1}{m}X^{T} * X\ (x{:}\ m \times n)$$

$$C_{ov} = n \times m \times m \times n = n \times n$$

Next, the eigenvalues and eigenvectors of the covariance matrix are calculated. The algorithm of singular value decomposition is used to calculate(9).

$$[U, S, V] = S_{vd}(C_{ov})$$
$$U{:}\ n \times n\ \text{(reduced dimensional matrix)} \quad (9)$$
$$U = \left[U^{(1)}, u^{(2)}, \dots, u^{(k)}, \dots, u^{(n)}\right]$$

(Samples can be reduced to k dimensions)

2) Objective function

Since the ultimate goal is to make the filtration efficiency (f) as high as possible while the filtration resistance (q) is as small as possible, in order to facilitate the calculations, a positive/negative indicator is used to handle the transformation into a total indicator(10).

$$\text{obj: max}f\&\&\text{min}q{-}{>}\ \text{max}0.5^{*}f + 0.5^{*}q \quad (10)$$

That is, maximum 0.5*f-0.5*(0.218×q) = 0.207×l+0.241×d+0.222 × ø + 0.121 × v - 0.148 × λ + 0.118 × r, obtaining the relationship between filtration resistance and filtration efficiency shown in Fig. 5.

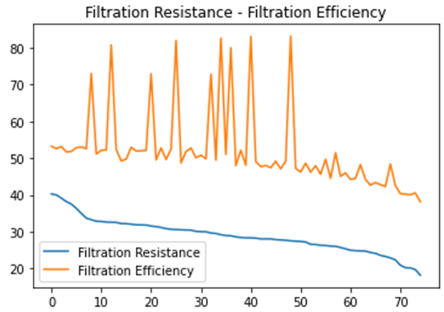


Fig.5 Relationship between filtration resistance and filtration efficiency

3) Constraints

Based on the existing data, we aim to achieve the highest possible filtration efficiency and minimize filter resistance. To do this, we determine the range of values for various influencing factors, leading to constraints such as: the receiving distance being between 0 and 100 cm, the hot air speed ranging from 0 to 2000 r/min, the thickness being from 0 to 3 mm, and the compression resilience lying between 0 and 85%. Given these constraints, a multi-objective linear programming model is established(11).

$$\text{obj: max}f\&\&\text{min}q{-}{>}\ \text{max}0.5^{*}f + 0.5^{*}q \quad (11)$$

B. Random Forest Algorithm Fusion

1) Explanation of related terms:

The described concept is the conditional probability distribution of the class given specific characteristic conditions. This is employed in the Random Forest Algorithm, which combines multiple decision trees. For each instance, the dataset is randomly selected with replacement, and a subset of the features is also randomly chosen for the output.

2) Algorithm steps.

Assuming the training set T has a size of N and contains M features, and given a random forest size of K, we traverse the random forest K times. For each pass, using a put-back sampling method on training set T, we sample N times to generate a new sub-training set D. We then randomly select m features, ensuring m is less than M. Using the newly formed training set D and these m features, we learn a complete decision tree.

Specific process:

The exact process of a random forest is shown in Fig. 6.

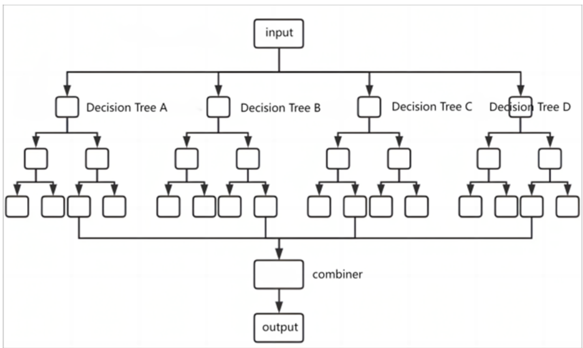


Fig.6 Random Forest

IV. EXPERIMENTAL RESULT AND ANALYSE

A. Analysis of PCA experiment results[11]

Considering the applicability of principal component analysis, the test is shown in Table II:

TABLE II.KMO test and Bartlett's test table

KMO test and Bartlett's test		
KMO value		0.652
Barteltt's test of sphericity	Approximate chi-square	555.479
	df	21.000
	p	0.000***

B. Analysis of experimental results of multiple nonlinear regression models

Decision variable:

Since the combination of process parameters is to be chosen so that the filtration efficiency is as high as possible while the filtration resistance is as low as possible, the receiving distance, hot air velocity, thickness, porosity, compression resilience, and permeability are used as decision variables(12), i.e., the decision variables are:

$$l, v, d, ø, \lambda, r \tag{12}$$

after the polynomial regression equation, the model is solved using nonlinear least squares (NLS) and with the use of principal component analysis method regression model in data dimensionality reduction. The model equation obtained is(13):

$$f=0.207\times l+0.241\times d+0.222\times ø+0.121\times v-0.148\times\lambda-0.218\times q+0.118\times r \tag{13}$$

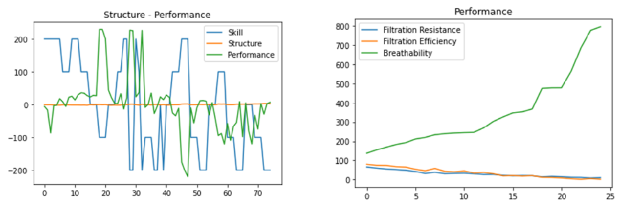
The data related to the regression equations are shown in Table III below:

TABLE III.Factor weight analysis table

name	Variance explained rate	Cumulative variance explained	weights
Principal Component 1	4.02619	57.517%	100%

The calculation of the weights of the principal component analysis shows that the weight of principal component 1 is 100%.

C. *Analysis of the experimental results of the multi-objective linear programming modeling is shown in Fig. 7.*



(a) Relationship between product performance and structural variables

(b) Relationship between product performance parameters

Fig.7 Structure and parameters of product performance

Solving for the constraints yields (14):

i.e. max $0.5*f-0.5*(0.218\times q)$
$=0.207\times l+0.241\times d+0.222\times ø+0.121\times v-0.148\times\lambda+0.118\times r$

$$\text{s.t.}\begin{cases} 0<l<100 \\ 0<v<2000 \\ 0<d<3 \\ 0<\lambda<85 \end{cases} \tag{14}$$

D. *Experimental Analysis of Random Forest Algorithm Fusion*

The optimal screenshot solved according to the random forest algorithm is shown in Figure8.

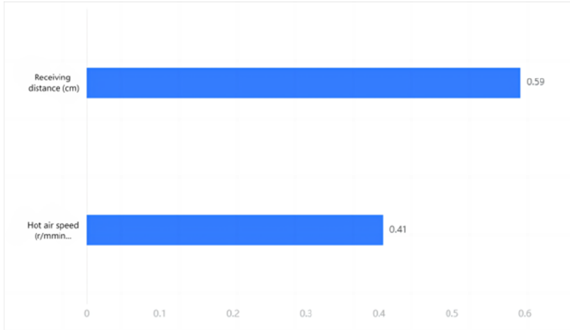


Fig.8 Receiving distance, thermal velocity characterization importance map

Multi-objective linear[12] programming is solved using the random forest algorithm. Get process parameters: receiving distance of 22cm, hot air speed of 2000r/min, can make the filtration efficiency as high as possible at the same time While striving to filter resistance as small as possible.

V. CONCLUSION

A. *Advantages of the model*

In our study, data was consolidated and tested for the normality of each parameter's distribution. By leveraging Pearson's correlation coefficient, combined with experimental findings, we enhanced the reliability of the results. Our approach utilized linear regression to fit data sets like "receiving distance - hot air velocity - thickness" and "receiving distance - hot air velocity - porosity". For more complex distributions such as "receiving distance - hot air velocity - compression resilience", we employed a BP neural network. This adaptability to different data distributions underpins the method's robust generalization and accuracy. Dimensionality reduction tackled challenges such as the "curse of dimensionality", reducing computational complexity, and enabling more intuitive data understanding. Overall, the chosen linear regression effectively captures the relationship between process parameters and filtration efficiency. Additionally, the model was further optimized using the random forest algorithm.

B. *Disadvantages of the model*

While nonlinear regression offers an improved fit, it encompasses not just the relationship between process parameters and filtration efficiency but also several other features. This suggests the derived curve may not depict a direct correlation between the process parameters and filtration efficiency.

ACKNOWLEDGMENTS

1) Zhuhai College of Science and Technology Doctoral Enhancement Program.

2) Support for Research Platforms and Projects in Guangdong Universities (New Generation Information Technology Key Area Specialization) 2023ZDZX1049.

REFERENCES

[1] Peng M, Jia H, Jiang L, et al. Study on structure and property of PP/TPU melt-blown nonwovens[J]. The Journal of The Textile Institute, 2019, 110(3): 468-475.

[2] Schreiber J B. Issues and recommendations for exploratory factor analysis and principal component analysis[J]. Research in Social and Administrative Pharmacy, 2021, 17(5): 1004-1011.

[3] Bulturbayevich M B, Baxromovna B L. Application of nonlinear regression models[C]//Conference Zone. 2022: 299-303.

[4] Kara Y, Molnár K. A review of processing strategies to generate melt-blown nano/microfiber mats for high-efficiency filtration applications[J]. Journal of Industrial Textiles, 2022, 51(1_suppl): 137S-180S.

[5] Speiser J L, Miller M E, Tooze J, et al. A comparison of random forest variable selection methods for classification prediction modeling[J]. Expert systems with applications, 2019, 134: 93-101.

[6] Khaire U M, Dhanalakshmi R. Stability of feature selection algorithm: A review[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(4): 1060-1073.

[7] Likitha B, Nakka J, Verma J, et al. Prediction of breast cancer analysis using machine learning algorithms and xgboost technique[C]//Data Science and Computational Intelligence: Sixteenth International Conference on Information Processing, ICInPro 2021, Bengaluru, India, October 22–24, 2021, Proceedings 16. Springer International Publishing, 2021: 298-313.

[8] Hou R, Zhou D, Nie R, et al. VIF-Net: An unsupervised framework for infrared and visible image fusion[J]. IEEE Transactions on Computational Imaging, 2020, 6: 640-651.

[9] Zhang L, Shi Z, Cheng M M, et al. Nonlinear regression via deep negative correlation learning[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 43(3): 982-998.

[10] Tan L, Li C, Xia J, et al. Application of Self-Organizing Feature Map Neural Network Based on K-means Clustering in Network Intrusion Detection[J]. Computers, Materials & Continua, 2019, 61(1).

[11] Huang Y, Shen L, Liu H. Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China[J]. Journal of Cleaner Production, 2019, 209: 415-423.

[12] Maiti I, Mandal T, Pramanik S. Neutrosophic goal programming strategy for multi-level multi-objective linear programming problem[J]. Journal of ambient intelligence and humanized computing, 2020, 11(8): 3175-3186.