

# Fabric and Yarn Color Prediction Based on Random Forest and Weaving Parameters



Jinxing Liang<sup>1,2\*</sup>, Junyi Huang<sup>1</sup>, Wen Wu<sup>1</sup> and Guanghao Wu<sup>1</sup>

<sup>1</sup>School of Computer and Artificial Intelligence, Wuhan Textile University, Wuhan, Hubei 430200, China

<sup>2</sup>Key Laboratory of Intelligent Perception and Computing in Textile Industry, Wuhan, Hubei 430200, China

Submission: March 11, 2025; Published: March 19, 2025

\*Corresponding author: Jinxing Liang, School of Computer and Artificial Intelligence, Wuhan Textile University, China

## Abstract

Texture is a critical factor influencing the color measurement and analysis of textiles. In this study, we use six yarns of different colors to weave fabric samples with different textures by adjusting weaving parameters such as weft density, reed count, and pattern using an SGA598 fully automatic rapier loom. The yarn and fabric color data were measured using an X-Rite Color i7 spectrophotometer. Correlation analysis between fabric weaving parameters and their color data revealed a nonlinear relationship. Therefore, we construct the forward and backward color prediction model for yarn and fabric based on weaving parameters using the Random Forest (RF) algorithm. The model was validated using the k-fold cross-validation and compared with multi-layer perceptron (MLP) and multi-output linear regression (MOLR) models. Results demonstrated that the forward color prediction model (yarn to fabric) achieved a coefficient of determination ( $R^2$ ) of 0.9997, with average predicted color differences of  $\Delta E_{ab}$  and  $\Delta E_{00}$  being 0.342 and 0.241. The backward color prediction model (fabric to yarn) achieved an  $R^2$  of 0.9999, with average predicted color differences of  $\Delta E_{ab}$  and  $\Delta E_{00}$  being 0.166 and 0.115. The RF model's prediction accuracy significantly outperformed that of the MLP and MOLR models, providing support for color measurement and analysis in textiles.

**Keywords:** Color Prediction; Yarn; Fabric; Weaving Parameters; Random Forest Algorithm

## Introduction

Color is one of the key factors influencing the quality of textile products and consumer satisfaction. In the modern textile industry, accurate analysis and control of fabric color have become crucial for enhancing product competitiveness [1]. However, the presentation of fabric color is not only affected by the inherent color of the yarn but also significantly influenced by texture characteristics determined by weaving parameters, posing a major challenge for accurate color analysis [2]. For fabric, weaving parameters such as warp and weft density, as well as weaving methods, interact to form complex light absorption and scattering properties, directly affecting the measurement results of its color [3,4]. This is particularly evident in textiles with pronounced texture structures, often leading to color differences between products and expected designs, thereby increasing the cost and cycle of product development [5].

In recent years, researchers have explored various methods to address the impact of texture characteristics on color measurement.

For instance, physical optical models such as the Kubelka-Munk theory [6] and multiple scattering theory [7] have been used to explain the mechanisms of texture's influence on color from an optical perspective. However, these models are often based on simplified assumptions, making it difficult to accurately describe light scattering behaviors under complex texture structures. Some studies have attempted to improve prediction accuracy by introducing correction factors, but these approaches are typically limited to specific types of textiles and lack generalizability [8]. Additionally, efforts have been made to establish quantitative description systems for texture characteristics [9,10]. But it is worth noting, due to the high-dimensional nature of texture parameters and their complex interactions, traditional linear modeling methods struggle to capture the intrinsic relationships between texture and color [11].

In recent years, data-driven modeling approaches have gradually provided new insights for addressing this complex

issue [12]. Studies by Lorente et al. have demonstrated the advantages of machine learning algorithms in tackling intricate problems such as textile demand forecasting [13]. Medina et al. found that machine learning models outperform purely statistical prediction models in terms of accuracy [14]. Zhu et al. proposed a hybrid model combining Convolutional Neural Networks (CNN) and Random Forest (RF) for the classification of animal fiber microscopic images [15], showcasing the unique advantages of machine learning in handling high-dimensional, nonlinear features and offering new technical pathways for textile property analysis.

Although the above studies provide valuable references for understanding the relationship between fabric texture and color, current quantitative methods for describing the impact of texture characteristics on color measurement remain incomplete. There is a lack of systematic modeling approaches based on weaving parameters, making it difficult to translate theoretical research into practical applications for the textile industry. Therefore, this study aims to establish the forward and backward color prediction model for yarn and fabric based on the Random Forest algorithm, providing methodological support for textile enterprises to predict and analyze fabric color based on customer samples. This approach is expected to reduce product development cycles and costs while improving production efficiency.

### Methods

#### Correlation analysis between weaving parameters and color attributes

We first explored the relationships between fabric weaving parameters and CIELab color values (L, a, b) as well as chroma (C) and hue angle (h) using Pearson correlation analysis. The results are presented in Table 1. The correlation coefficient between weft density and lightness (L) was -0.872, with a significance level of  $p$ -value  $< 0.001$ , indicating a significant negative correlation. Specifically, as the weft density increased, the fabric lightness decreased significantly, primarily due to changes in texture characteristics and light absorption/scattering properties caused by higher weft density. In contrast, the correlation coefficients between weft density and color values (a and b), chroma (C), and hue angle (h) were 0.254, 0.007, 0.125, and -0.045, respectively, with  $p$ -values much greater than 0.05, demonstrating a significant nonlinear relationship. This suggests that the mechanism by which texture influences color is more complex.

The above correlation analysis reveals that while some weaving parameters exhibit strong linear relationships with lightness, fabric production processes generally involve nonlinear interactions among multiple parameters, resulting in an overall nonlinear influence on color data. Additionally, complex optical properties arising from fabric structure variations, such as multiple scattering, surface reflection, and angle dependency, further enhance the nonlinear relationship between weaving

parameters and color data, making it difficult to describe using simple physical models. Therefore, this study employs the Random Forest (RF) algorithm to construct a color prediction model for yarn and fabric based on weaving parameters. By building multiple decision trees and integrating their predictions, the model effectively captures nonlinear relationships and interaction effects among variables, providing support for quantitative color analysis in fabric production.

#### Random Forest algorithm

The Random Forest algorithm is an ensemble learning method based on decision trees, proposed by Breiman [16]. This algorithm improves model accuracy and robustness by constructing multiple decision trees and aggregating their predictions. The basis inner the Random Forest is decision trees, which recursively partition the feature space to classify or predict samples. This approach efficiently handles high-dimensional data, is less prone to overfitting, and exhibits strong resistance to outliers and noise, making it suitable for complex nonlinear regression problems.

For regression tasks, the final prediction of the Random Forest is the average of the predictions from all decision trees. For a single decision tree  $h_t(x)$ , it establishes a mapping between input features and target variables by recursively partitioning the feature space. The construction of decision trees follows the principle of minimizing the objective function. Given an input feature vector  $x$ , the prediction output of the Random Forest regression model is expressed as Equation (1):

$$y' = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (1)$$

where,  $T$  is the number of decision trees in the forest, and  $h_t(x)$  is the prediction vector of the  $t$ -th decision tree for input  $x$ ,  $y'$  is the predicted result.

#### K-fold cross-validation

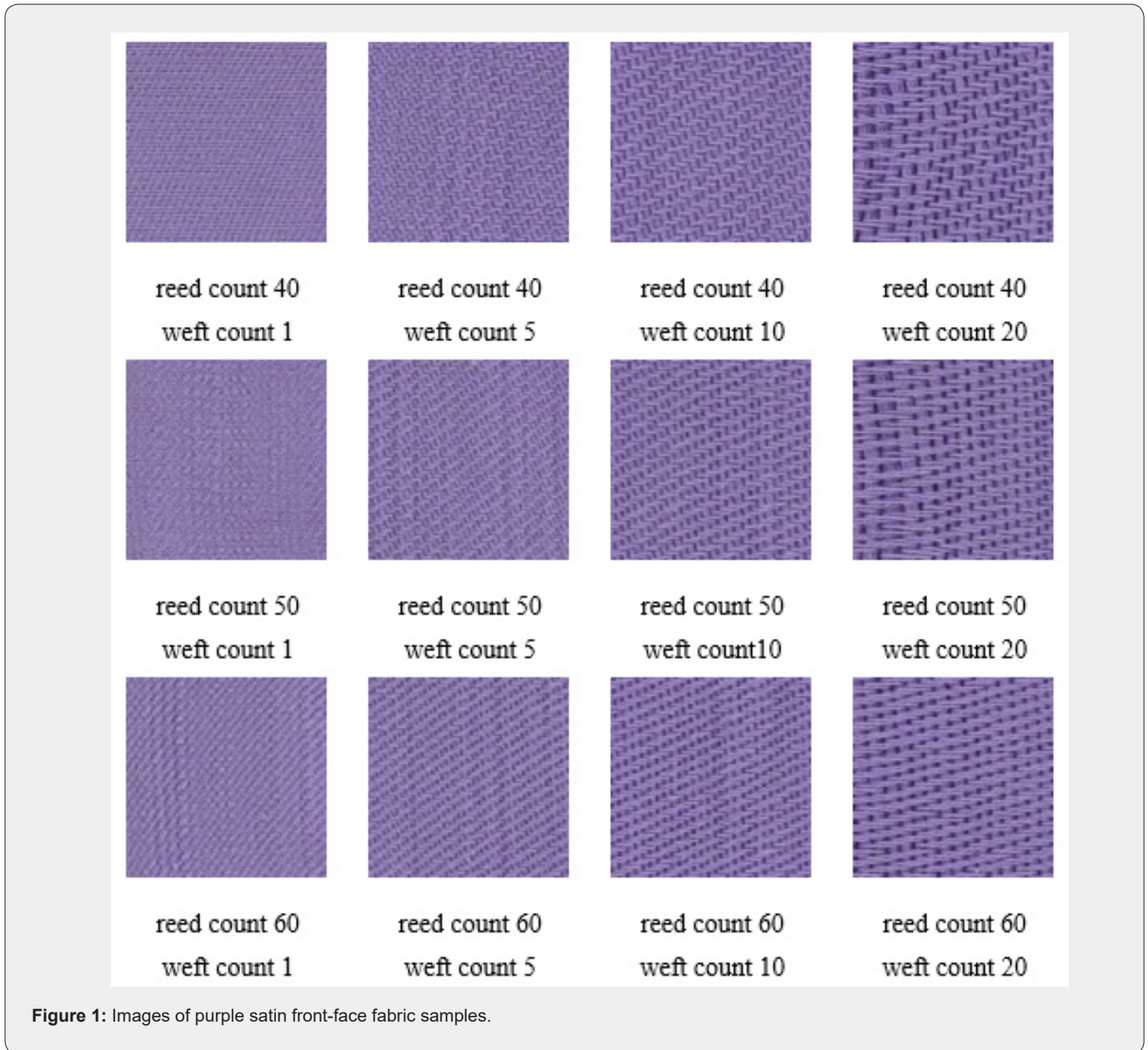
The generalization ability and stability of the color prediction model are critical aspects. This study employs  $k$ -fold cross-validation to train and evaluate the model, providing reliable validation under limited samples. The core idea is to assess model performance through data partitioning and validation iterations. Specifically, the dataset  $D$  is divided into  $k$  disjoint subsets  $D_1, D_2, \dots, D_k$ , each of size  $n/k$ , where  $n$  is the total number of samples. In the  $j$ -th iteration ( $j \in \{1, 2, \dots, k\}$ ),  $D_j$  is used as the validation set, while the remaining  $k-1$  subsets are used as the training set. Each fold's training process is conducted independently, and the results from all  $k$  iterations are aggregated to evaluate the model's performance. In this study, a sample weight adjustment mechanism was introduced during training to enhance the model's predictive accuracy, as detailed in the model construction section.

**Experiment**

**Sample preparation and color measurement**

This study utilized an SGA598 fully-automatic rapier loom and six yarns of different colors (red, yellow, blue, green, purple, and gray) to fabricate fabric samples with varying texture characteristics by controlling fabric texture and yarn density. The weaving textures included plain weave, twill face, twill back,

satin face, and satin back, totaling five variations. The longitudinal density of the yarn was controlled by adjusting the reed count, set to 40, 50, and 60, respectively. The transverse density was regulated by the number of weft yarns, set to 1, 5, 10, and 20. Consequently, a total of  $6 \times 5 \times 3 \times 4 = 360$  fabric samples were produced, with each color of yarn twisted from 32-count yarn. As an example, the satin front-face fabric sample for purple yarn is shown in Figure 1.



After sample preparation, the color attributes of both yarns and fabrics under CIE1964 and 10 degree standard observer color matching functions was measured using an X-Rite Color i7 bench-top spectrophotometer in large aperture mode. Among these, the color measurement of the yarn is accomplished through the method of yarn winding. To mitigate measurement noise, an

average of 10 measurements was taken for each yarn and fabric sample.

**Data feature representation**

For modeling, one-hot encoding was employed to convert weaving parameters into numerical features recognizable by the

model. As categorical variables, texture variables were expanded into five binary features, each corresponding to a specific texture, with values of 0 or 1. Weaving parameters such as reed count and weft count were normalized using the Z-score standardization method to eliminate the influence of different units and scales. The Z-score normalization is expressed as Equation (2):

$$x^* = \frac{x - \mu}{\sigma} \quad (2)$$

where  $x$  is the original feature value,  $\mu$  is the mean of the feature,  $\sigma$  is the standard deviation, and  $x^*$  is the normalized feature value.

Additionally, the input and output of the model include color attributes for both yarn and fabric, specifically comprising five components of lightness (L), red-green chromaticity (a), yellow-blue chromaticity (b), color saturation (C), and hue angle (h). By concatenating the waving parameters and yarn color attributes of each fabric as the model input and using the fabric's color attributes as the output, the forward color prediction model (yarn to fabric) was constructed, as shown in Equation (1). Conversely, by using the fabric's weaving parameters and fabric color attributes as input and the yarn's color attributes as output, the backward color prediction model (fabric to yarn) was established.

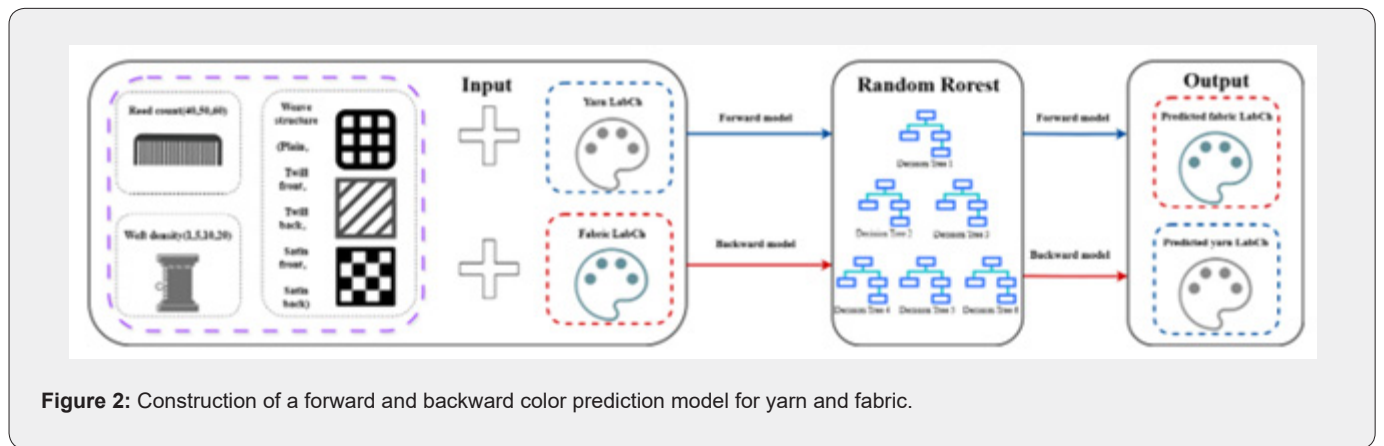


Figure 2: Construction of a forward and backward color prediction model for yarn and fabric.

Figure 2 illustrates the overall architecture of the forward and backward color prediction model for yarn and fabric based on the Random Forest algorithm. Taking the forward model (from yarn to fabric) as an example, its input features  $X$  are defined as Equation (3):

$$X = \begin{bmatrix} k_1, w_1, t_1, L_{yarn,1}, a_{yarn,1}, b_{yarn,1}, C_{yarn,1}, h_{yarn,1} \\ \vdots \\ k_i, w_i, t_i, L_{yarn,i}, a_{yarn,i}, b_{yarn,i}, C_{yarn,i}, h_{yarn,i} \\ \vdots \\ k_n, w_n, t_n, L_{yarn,n}, a_{yarn,n}, b_{yarn,n}, C_{yarn,n}, h_{yarn,n} \end{bmatrix} \quad (3)$$

where  $k_i$  is the reed count,  $w_i$  is the weft count,  $L_{yarn,i}$ ,  $a_{yarn,i}$ ,  $b_{yarn,i}$ ,  $C_{yarn,i}$  and  $h_{yarn,i}$  are the color feature vectors of the yarn, and  $t_i$  is the texture category vector constructed using one-hot encoding. The subscript  $i$  denotes the  $i$ -th sample, with a range from 1 to  $n$ . The target variable  $y_i$  is defined as Equation (4):

$$y_i = [L_{fabric}, a_{fabric}, b_{fabric}, C_{fabric}, h_{fabric}] \quad (4)$$

where  $y_i$  contains the five color components of the fabric. Using the Random Forest model, for each decision tree  $T_k$ , given a sample dataset  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ , the goal of each tree is to split based on the features  $x_i$  to predict the target variable  $y_i$ .

At each level of the tree, a feature  $x_i$  and a split value  $q$  are selected to minimize the mean squared error (MSE) of the node, as shown in Equation (5):

$$split(x_i, q) = \arg \min_t \sum_{i=1}^n (y_i - f(x_i))^2 \quad (5)$$

where  $f(x_i)$  is the predicted value at the current node. The final prediction of the Random Forest regression is the average of the predictions from all decision trees, as shown in Equation (1).

During the model training process, grid search is employed to select the optimal hyper-parameter combination for hyper-parameter tuning. The goal of hyper-parameter tuning is to minimize the mean squared error in cross-validation. Given a hyper-parameter combination  $h$ , the model is optimized by minimizing the mean squared error in cross-validation, as shown in Equation (6):

$$h^* = \arg \min_h \frac{1}{k} \sum_{i=1}^k MSE(h, D_{train,i}, D_{val,i}) \quad (6)$$

where  $k$  is the number of folds in cross-validation (set to 5 and 10 in this experiment),  $MSE(h, D_{train,i}, D_{val,i})$  is the mean squared error of the hyper-parameter combination  $h$  on the  $i$ -th fold, and



$h^*$  is the optimized hyper-parameter combination. Similarly, by using fabric color and weaving parameters as the input and yarn color as the output, a backward color prediction model (from fabric to yarn) can be constructed.

**Evaluation metrics**

The study employs multiple evaluation metrics, including the coefficient of determination ( $R^2$ ) and color difference measures such as CIELab ( $\Delta E_{ab}$ ) and CIEDE2000 ( $\Delta E_{00}$ ), to assess model performance. The coefficient of determination ( $R^2$ ) quantifies the extent to which the model explains the variability of the dependent variable. A value closer to 1 indicates a better fit, as shown in Equation (7):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

where  $n$  is the number of samples,  $y_i$  is the true value of the  $i$ -th sample, and  $y'_i$  is the predicted value of the  $i$ -th sample,  $\bar{y}$  is the mean of the true values. A value of  $R^2$  closer to 1 indicates a better fit of the model to the data. Additionally, color differences in CIELab ( $\Delta E_{ab}$ ) and CIEDE2000 ( $\Delta E_{00}$ ) were used to assess the

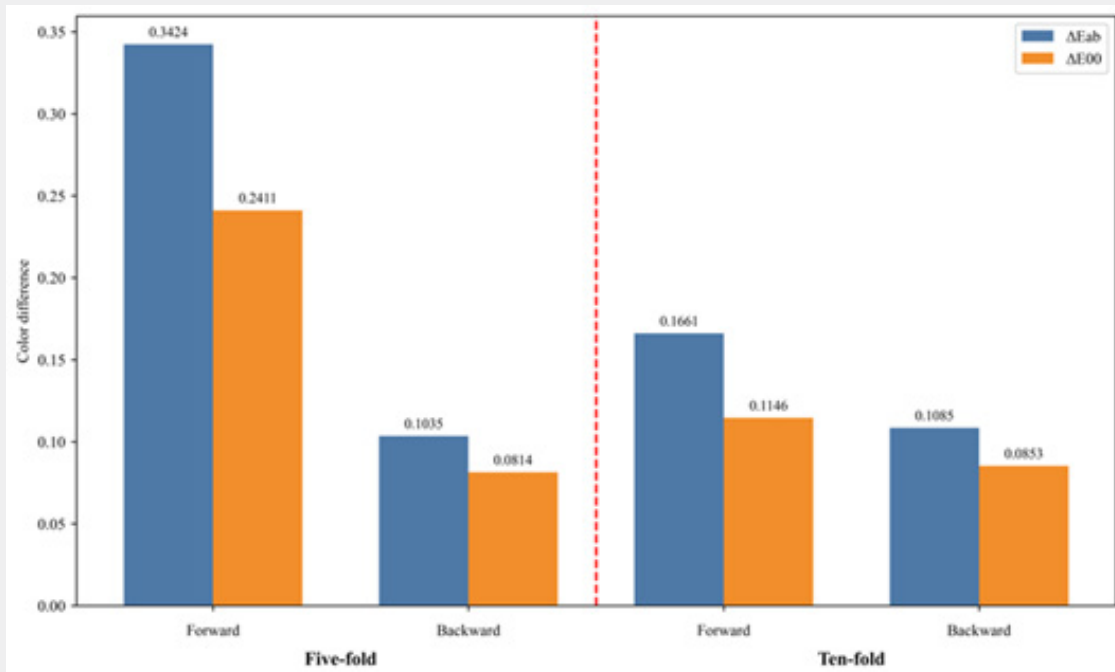
differences between true and predicted color values. The CIELab color difference is calculated as Equation (8). The CIEDE2000 color difference calculation follows the method described in reference [17].

$$\Delta E_{ab} = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2} \quad (8)$$

where  $(L_1, a_1, b_1)$  and  $(L_2, a_2, b_2)$  are the true and predicted color values, respectively.

**Results and Analysis**

Tables 2 and 3 present the experimental results for the forward model (predicting fabric color from yarn color) and the backward model (predicting yarn color from fabric color) based on weaving parameters, respectively. The results include statistical metrics such as the average (Avg.), standard deviation (Std.), median (Med.), maximum (Max.), 90<sup>th</sup> percentile (90%), and 95<sup>th</sup> percentile (95%) of the color differences  $\Delta E_{ab}$  and  $\Delta E_{00}$ , as well as the coefficient of determination ( $R^2$ ). Visual comparisons of the color differences and  $R^2$  values are shown in Figures 3 and 4, respectively. Due to differences in the input and output of the forward and backward models, their performance also varies to some extent.



**Figure 3:** Comparisons of the forward and backward color prediction results in terms of color differences.

From the results in Table 2, we can see that under 5-fold cross-validation, the forward model achieves average color differences  $\Delta E_{ab}$  of 0.3424 and  $\Delta E_{00}$  of 0.2411. And under 10-fold cross-validation, the average color differences are 0.1661 and 0.1146

for  $\Delta E_{ab}$  and  $\Delta E_{00}$ , respectively, both significantly lower than the commonly used perceptible color difference threshold in textile production ( $\Delta E_{00}=1.0$ ). This indicates a high consistency between the model's predictions and the actual measurements.

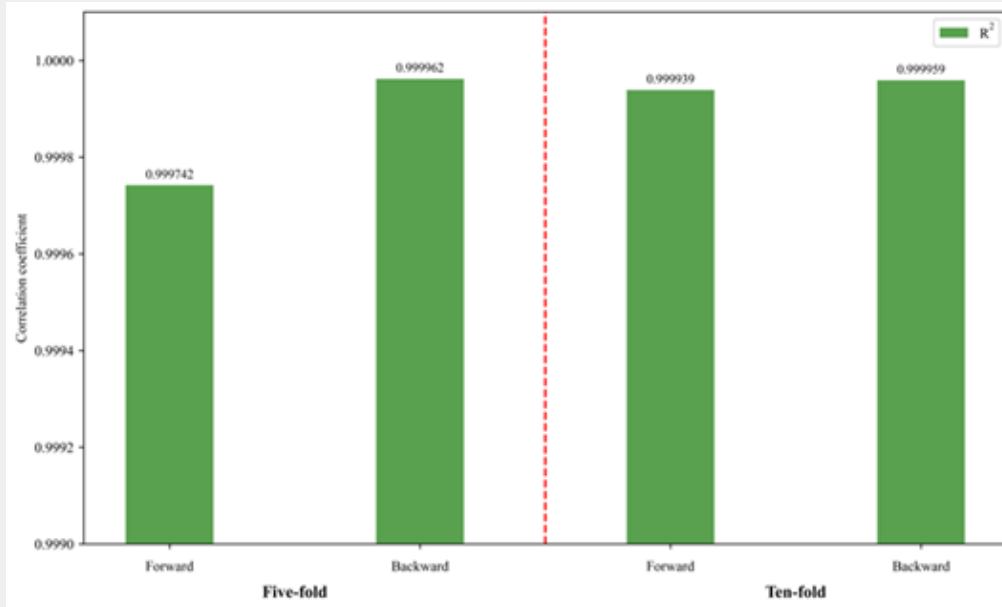


Figure 4: Comparisons of the forward and backward color prediction results in terms of R2.

Table 1: Correlation coefficients between weaving parameters and the color attributes of fabric samples.

Weaving parameters	L	p-value	a	p-value	b	p-value	C	p-value	h	p-value
Weft	-0.872	0	0.254	0.189	0.007	0.121	0.125	0.189	-0.045	0.37
Reed	-0.029	0.33	-0.1	0.249	0.032	0.427	-0.133	0.329	0.025	0.463
Texture	0.162	0.413	0.221	0.109	0.258	0.306	0.232	0.151	0.101	0.095

Table 2: Forward color prediction results for fabric using yarn color and weaving parameters.

Results	Forward model			
	5-fold		10-fold	
	$\Delta E_{ab}$	$\Delta E_{00}$	$\Delta E_{ab}$	$\Delta E_{00}$
Avg.	0.3424	0.2411	0.1661	0.1146
Std.	0.2346	0.1774	0.115	0.0826
Med.	0.2891	0.2013	0.1405	0.0939
Max.	1.5549	1.0974	0.8882	0.5684
90%	0.2149	0.1772	0.233	0.2194
95%	0.6895	0.6254	0.3717	0.2727
R <sup>2</sup>	0.9997		0.9999	

Table 3: Backward color prediction results for yarn using fabric color and weaving parameters.

Results	Backward model			
	5-fold		10-fold	
	$\Delta E_{ab}$	$\Delta E_{00}$	$\Delta E_{ab}$	$\Delta E_{00}$
Avg.	0.1035	0.0814	0.1085	0.0853
Std.	0.0917	0.0772	0.094	0.0793
Med.	0.0742	0.0556	0.0835	0.0614
Max.	0.5769	0.5044	0.5956	0.5253
90%	0.2149	0.1772	0.2285	0.1908
95%	0.2727	0.2369	0.2918	0.2522
R <sup>2</sup>	0.99996		0.99996	

Under 5-fold cross-validation, the standard deviation of  $\Delta E_{ab}$  for the forward model is 0.2346, with a median of 0.2891 and a 95% percentile of 0.6895. Under 10-fold cross-validation, the standard deviation of  $\Delta E_{ab}$  decreases to 0.1150, with a median of 0.1405 and a 95% percentile of 0.3717. All these values are below the perceptible color difference threshold, and the median color difference is close to the average with a small standard deviation, indicating a relatively uniform distribution of prediction errors. This illustrates the vast majority of predictions exhibit very small color differences, with no significant skewness.

Furthermore, from 5-fold to 10-fold cross-validation, both the average and maximum color differences show a significant decreasing trend. The average  $\Delta E_{ab}$  decreases from 0.3424 to 0.1661, and the average  $\Delta E_{00}$  decreases from 0.2411 to 0.1146. The maximum  $\Delta E_{ab}$  decreases from 1.5549 to 0.8882, and the maximum  $\Delta E_{00}$  decreases from 1.0974 to 0.5684. The standard deviations also decrease accordingly, indicating that increasing the number of cross-validation folds allows the model to learn data features more thoroughly, thereby effectively improving prediction accuracy. Under 10-fold cross-validation, the model achieves a coefficient of determination ( $R^2$ ) of 0.9999, further demonstrating the superiority of the Random Forest model in capturing the nonlinear relationships among yarn color, weaving parameters, and fabric color.

Analysis of results in Table 3 reveals that under 10-fold cross-validation, the backward model achieves average color differences of 0.1085 and 0.0853 for  $\Delta E_{ab}$  and  $\Delta E_{00}$ , respectively, significantly lower than the corresponding metrics of the forward model. This indicates higher accuracy in predicting yarn color from fabric color. The standard deviations of the backward model with  $\Delta E_{ab}=0.0940$  and  $\Delta E_{00}=0.0793$  are small and similar to the forward model, indicating high stability in prediction results. The median values are lower than the average, suggesting that a few larger color differences may have slightly increased the average. However, the overall prediction error distribution remains concentrated at low levels. The maximum color differences ( $\Delta E_{ab}=0.5956$ ,  $\Delta E_{00}=0.5253$ ) and 95% percentiles ( $\Delta E_{ab}=0.2918$ ,  $\Delta E_{00}=0.2522$ ) indicate that even in extreme cases, the model's prediction errors remain within an acceptable range.

Additionally, the backward model achieves a coefficient of determination ( $R^2$ ) of 0.99996, indicating an extremely high level of fit and the ability to accurately capture the relationship between fabric color and yarn color. Unlike the forward model, the backward model shows very slight differences between 5-fold and 10-fold cross-validation ( $\Delta E_{ab}=0.1035$  vs. 0.1085,  $\Delta E_{00}=0.0814$  vs. 0.0853), suggesting greater robustness and lower sensitivity to the number of cross-validation folds.

Further comparison of the above results reveals that the average color difference of the backward model is significantly lower than that of the forward model. This discrepancy may stem

from the nature of the prediction tasks. In the forward model (predicting fabric color from yarn color), the input consists of yarns of the same color but with different weaving parameters, aiming to predict the corresponding fabric colors. As we all know, fabric color is influenced by multiple factors, including yarn color, texture structure, weft density, and reed count, representing a mapping from a single yarn color feature space to a multi-texture fabric color feature space.

In contrast, the backward model (predicting yarn color from fabric color) takes fabrics of different colors and weaving parameters as input to predict the color of the same yarn, representing a mapping from a multi-texture fabric color feature space to a single yarn color feature space. When calculating the loss function, the forward model aggregates the prediction errors for each fabric color, while the backward model only calculates the prediction errors for six yarn colors. Consequently, the large sample size of the forward model, combined with potential measurement errors, increases the complexity of the forward model, resulting in lower accuracy compared to the backward model, which has a simpler loss function calculation.

Additionally, from the perspective of the Random Forest algorithm's principles, its ensemble of decision trees excels at handling data points with discrete distributions in the feature space. In the backward model, fabric samples with different weaving parameter combinations may form clearer cluster structures in the feature space, facilitating decision tree splitting and learning. In the forward model, fabric colors generated from the same yarn under different weaving parameters may exhibit greater dispersion and non-linearity, increasing the difficulty of model learning.

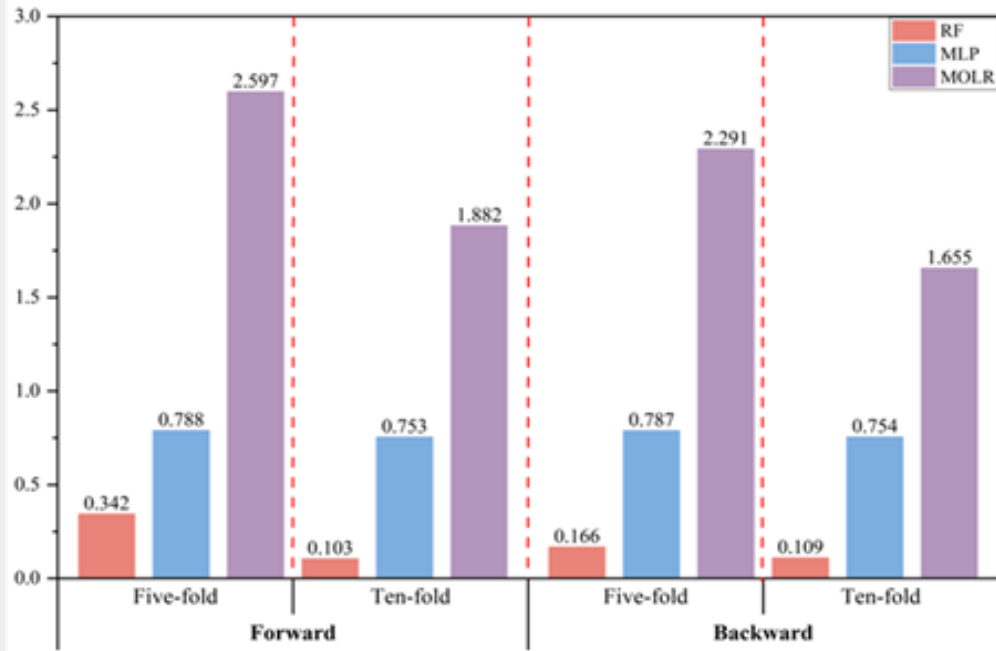
In summary, both the forward and backward prediction models achieve average color differences under 10-fold cross-validation that are well below the human-perceptible threshold ( $\Delta E_{00}=1.0$ ), demonstrating the practical value of the Random Forest-based color prediction. In practice, the forward model (yarn to fabric) is particularly suitable for workflows such as textile design and weaving parameter optimization. On the other hand, the backward model (fabric to yarn), with its higher stability and accuracy, is more applicable to tasks such as fabric color quality control and textile color reverse analysis.

### Model Comparison

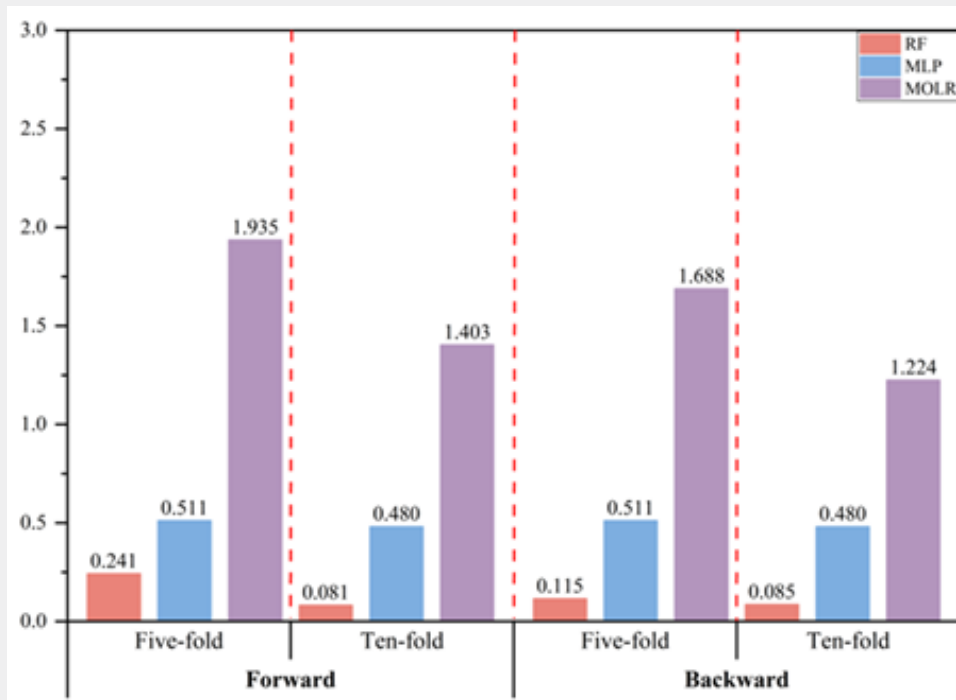
This study further compared the Random Forest-based (RF) color prediction model with a multi-layer perceptron (MLP) model and a multi-output linear regression (MOLR) model, using the same input and output data. The MLP model, with its deep network structure, can automatically learn and fit potential feature relationships. Tables 4 and 5 present the performance comparison results of the three models for the forward and reverse prediction tasks, respectively. The MOLR model is a natural extension of classical linear regression to multi-dimensional output spaces,

capable of simultaneously predicting multiple related continuous target variables while considering correlations among output variables. It is suitable for multi-dimensional prediction tasks

with inherent dependencies. Visual comparisons of the color differences and coefficients of determination ( $R^2$ ) are shown in Figures 5-7.



**Figure 5:** Comparison of the color difference  $\Delta E_{ab}$  of three models for forward and backward color prediction.



**Figure 6:** Comparison of the color difference  $\Delta E_{00}$  of three models for forward and backward color prediction.



**Table 4:** Comparison of color difference and  $R^2$  in forward color prediction results of three models.

Forward model						
Results	Five-fold			Ten-fold		
	$\Delta E_{ab}$	$\Delta E_{00}$	$R^2$	$\Delta E_{ab}$	$\Delta E_{00}$	$R^2$
RF	0.3424	0.2411	0.99974	0.1661	0.1146	0.99994
MLP	0.7885	0.5107	0.99905	0.7874	0.5108	0.99906
MOLR	2.5967	1.9352	0.99158	2.2912	1.6875	0.99367

**Table 5:** Comparison of color difference and  $R^2$  in backward color prediction results of three models.

Backward model						
Results	Five-fold			Ten-fold		
	$\Delta E_{ab}$	$\Delta E_{00}$	$R^2$	$\Delta E_{ab}$	$\Delta E_{00}$	$R^2$
RF	0.1035	0.0814	0.99996	0.1085	0.0853	0.99996
MLP	0.4805	0.753	0.9991	0.4803	0.7544	0.99909
MOLR	1.8815	1.4027	0.99158	1.655	1.2241	0.99582

From the results in Tables 4 and 5, as well as Figures 5-7, it is evident that the RF model achieves the highest color prediction accuracy for both the forward (yarn to fabric) and backward (fabric to yarn) prediction tasks, followed by the MLP model. The MOLR model performs significantly worse than both the RF and MLP models, demonstrating the superiority of the RF model in addressing textile color prediction problems.

Specifically, under 10-fold cross-validation for the forward color prediction model, the RF model achieves a  $\Delta E_{00}$  value of 0.1146, compared to 0.5108 for the MLP model and 1.6875 for the MOLR model. The prediction error of the RF model is approximately one-fifth that of the MLP model and less than one-tenth that of the MOLR. Similarly, for the backward color prediction model under 10-fold cross-validation, the RF model achieves a  $\Delta E_{00}$  value of 0.0853, compared to 0.4803 for the MLP model and 1.2241 for the MOLR model. Again, the RF model's prediction error is about one-fifth that of the MLP model and less than one-tenth that of the linear regression model. Moreover, the coefficients of determination ( $R^2$ ) for the Random Forest-based forward and backward models are slightly better than those of the MLP model and significantly better than those of the MOLR, further confirming the superior performance of the RF model.

Additionally, the average color differences of the Random Forest-based forward and backward models are significantly below the human-perceptible threshold ( $\Delta E_{00}=1.0$ ). While the MLP-based models exhibit higher average color differences than the RF models, they still remain close to the perceptible threshold. In contrast, the MOLR model show color differences significantly above the perceptible threshold ( $\Delta E_{00}=1.0$ ), indicating that their predictions may lead to noticeable color deviations in practical applications.

The superiority of the RF model can be attributed to its ensemble learning skills, which consists of multiple decision

trees. Each tree is trained independently and focuses on different aspects of the data, enabling a multi-perspective learning strategy that is particularly effective in capturing the complex nonlinear relationships between color and parameters. Furthermore, the RF voting mechanism and random feature selection provide inherent robustness against noise and outliers, making it more reliable in handling potential measurement errors.

Although MLP models can theoretically approximate any complex function, they are prone to local optima and over-fitting under limited training data, hindering their ability to fully model complex boundaries in high-dimensional feature spaces. The MOLR model assumes a linear relationship between features and target variables, which fundamentally mismatches the highly nonlinear of between parameters and color. In conclusion, the RF model demonstrates significant advantages in fabric and yarn color prediction, outperforming both MLP and MOLR in terms of prediction accuracy, stability, and practicality.

### Conclusion

Based on the nonlinear relationship between weaving parameters and fabric color, this study constructed forward (yarn to fabric) and backward (fabric to yarn) color prediction models via weaving parameters and using the RF algorithm. The study utilized six yarns of different colors to weave 360 samples under varying texture, reed count, and weft density. The proposed RF-based color prediction model was tested, and its performance was evaluated using metrics such as color difference and the coefficient of determination ( $R^2$ ). The experimental results were thoroughly discussed and analyzed, and comparisons were made with MLP and MOLR models. The results demonstrate that the RF model exhibits significant advantages in both forward and backward color prediction tasks, outperforming the MLP and MOLR models. The findings provide support for fabric color analysis and color prediction. Future work will focus on expanding the diversity of

experimental samples, conducting research on spectral modeling of yarn and fabric based on weaving parameters, and exploring advanced methods such as deep learning to further enhance the research.

### References

1. Cárdenas LM, Shamey R, Hinks D (2009) Key variables in the control of color in the textile supply chain. *International Journal of Clothing Science and Technology* 21(5): 256-269.
2. Humeau-Heurtier A (2019) Texture feature extraction methods: A survey. *IEEE access* 7: 8975-9000.
3. Alsmadi MK (2020) Content-based image retrieval using color, shape and texture descriptors and features. *Arabian Journal for Science and Engineering* 45(4): 3317-3330.
4. Pham MT, Lefèvre S, Merciol F (2018) Attribute profiles on derived textural features for highly textured optical image classification. *IEEE Geoscience and Remote Sensing Letters* 15(7): 1125-1129.
5. Malm V, Strååt M, Walkenström P (2014) Effects of surface structure and substrate color on color differences in textile coatings containing effect pigments. *Textile research journal* 84(2): 125-139.
6. He DC, Wang L (1991) Texture features based on texture spectrum. *Pattern recognition* 24(5): 391-399.
7. De Lucia M, Buonopane M (2004) Color prediction in textile application. *Optical Metrology in Production Engineering, France*.
8. Ben Salem Y, Nasri S (2010) Automatic recognition of woven fabrics based on texture and using SVM. *Signal, image and video processing* 4: 429-434.
9. Zhu Y, Duan J, Li Y (2022) Image classification method of cashmere and wool based on the multi-feature selection and random forest method. *Textile Research Journal* 92(7-8): 1012-1025.
10. Brodić D, Amelio A, Milivojević Z N (2017) Clustering documents in evolving languages by image texture analysis. *Applied Intelligence* 46(4): 916-933.
11. Yildirim P, Birant D, Alpyildiz T (2018) Data mining and machine learning in textile industry. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8(1): e1228.
12. Hameed I M, Abdulhussain S H, Mahmmod B M (2021) Content-based image retrieval: A review of recent trends. *Cogent Engineering* 8(1): 1927469.
13. Lorente-Leyva LL, Alemany MME, Peluffo-Ordóñez DH (2021) Demand forecasting for textile products using statistical analysis and machine learning algorithms. *Asian Conference on Intelligent Information and Database Systems Thailand*.
14. Medina H, Peña M, Siguenza-Guzman L (2021) Demand forecasting for textile products using machine learning methods. *International Conference on applied technologies Ecuador*: 301-315pp.
15. Zhu Y, Duan J, Wu T (2021) Animal fiber imagery classification using a combination of random forest and deep learning methods. *Journal of Engineered Fibers and Fabrics* 16: 15589250211009333.
16. Breiman L (2001) Random forests. *Machine learning* 45: 5-32.
17. Luo MR, Cui G, Rigg B (2001) The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application* 26(5): 340-350.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: 10.19080/CTFTTE.2025.09.555764

### Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
( Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission  
<https://juniperpublishers.com/online-submission.php>