Predicting the whiteness index of cotton fabric with a least squares model

Wan Sieng Yeo · Woei Jye Lau

Abstract The textile bleaching process that involves hot hydrogen peroxide (H_2O_2) solution is commonly practised in cotton fabric manufacture. The purpose of the bleaching process is to remove color from the cotton, achieving a permanent white before proceeding to dyeing or shape matching. Normally, the visual ratings of whiteness on the cotton are measured based on whiteness index (WI). However, it is found that there is little research on chemical predictive modelling of the cotton fabric's WI compared to experimental study. Analytics using predictive modelling can forecast the outcomes, leading to better-informed cotton quality assurance and control decisions. Up to date, there is limited study applying least square support vector regression (LSSVR) model in the textile domain. Hence, the present study aims to develop a multi-output LSSVR (MLSSVR) model using bleaching process variables and results obtained from two different case studies to predict the WI of cotton. The predictive accuracy of the MLSSVR model was measured by root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). The obtained results were compared with other regression models including partial least square regression, predictive fuzzy model, locally weighted partial least square regression, and locally weighted kernel partial least square regression. Our findings indicate that the proposed MLSSVR model performed better than other models in predicting the WI as it showed significantly lower values of RMSE and MAE. Furthermore, it provided the highest R^2 values which are up to 0.9999.

Abbreviations

A	A matrix consisting Lagrange multipliers
b	A threshold value
b_k	Kernel parameter for locally weighted kernel partial least square
CIE	Commission on Illumination
CPU	Central processing unit (CPU)
H, P, Q, K, Ω, 0, S	A definite matrix
H_2O_2	Hydroxide peroxide

W. S. Yeo (✉)
Department of Chemical Engineering, Faculty of Engineering and Science, Curtin University Malaysia, CDT 250, 98009 Miri, Sarawak, Malaysia
e-mail: christineyeo@curtin.edu.my

W. J. Lau
School of Chemical and Energy Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

LSSVR	Least squares support vector regression
LOO	Leave-one-out
LW-KPLSR	Locally weighted Kernel partial least square regression
LW-PLSR	Locally weighted partial least square regression
LV	Latent variable
MAE	Mean absolute error
p	A positive hyperparameter of radial basis function kernel function
PE	Prediction error
PLSR	Partial least square regression
MLSSVR	Multi-output least square support vector regression
N_T	Total number of data sets
N_1	Number of training data sets
N_2	Number of testing data sets
R^2	R-squared or the coefficient of determination
RBF	Radial basis function
RMSE	Root mean square error
SVM	Support vector machine
T	Temperature of bleaching process
t	Time of bleaching process
V, v_i	A vector in multi-output least square support vector regression
WI	Whiteness index
W	Weighed value vector
x_i, x	Input vector
x_c and y_c	Chromaticity coordinates of the bleached cotton fabric samples
x_n and y_n	Chromaticity coordinates of the illuminant
y^i, y, Y	Output vector
Y_L	Lightness
Z	A mapping to some high or even unlimited/infinite dimensional Hilbert space or feature space via the nonlinear mapping function ϕ with n_h dimensions
γ, λ	Two positive real regularised parameters in the multi-output least square support vector regression
$\phi(x)$	A nonlinear mapping function
ξ	A vector containing slack variables
Ξ	A matrix consisting of slack variables with an order of $l \times m$
α	A vector consisting of Lagrange multipliers
ℓ	The Lagrangian function

Introduction

Cotton is widely used to make fabrics such as garments, bedding, curtains, and carpets (Wang et al. 2018). Hence, it is an important fiber in the textile industry. About 25 million tons of cotton are produced annually around the world and most of them ($\sim 50\%$) are used to make the clothes (Ahmad et al. 2021). Cotton fabric is popular because it has advantages including softness, biodegradability, comfort, hypoallergenic, breathability, and environmentally friendly (being a natural fiber) (Xie et al. 2013). However, similar to other natural fibers, cotton fiber contains natural pigments that cause it to appear in yellowish-brown in natural (Oliveira et al. 2018). It must be noted that environmental factors such as soil, dust, smoke, -dirt, and insects could also possibly affect the color of cotton fiber. In general, the yellowish-brown of cotton is visually associated with soiling or the lack of cleanliness and it is an attribute that must be removed.

Industrially, cotton fabric bleaching is a chemical oxidation process used to remove the yellowish-brown coloration from cotton by damaging the colorant (Oliveira et al. 2018). In other words, the bleaching process is responsible for eliminating the coloring materials from the cotton fiber to have a pure white appearance. A white fabric is highly desirable as it gives the impression of clean and pure (Jung and Sato 2013). Hence, the whiteness degree of cotton fabric is the main requirement of bleaching. Besides, the bleaching process that uses strong reducing or oxidizing agents is able to get rid of potentially hazardous contaminants such as bacteria, molds, and fungi (Gültekin 2016).

Hydrogen peroxide (H_2O_2) is one of the commonly used bleaching agents and is highly effective to oxidize the coloring matters (Oliveira et al. 2018). It is more preferable compared to chlorine bleach because it is gentler and less toxic (Bajpai 2007). Additionally, optical brighteners can be added to the bleaching process to increase whiteness levels (Oliveira et al. 2018). After the bleaching process, the whiteness index (WI) that indicates the whiteness

degree of the cotton is measured. Whiteness is defined in colorimetric terms as a color with the highest luminosity, no hue, and no saturation. The WI is calculated from the data computed by colorimetric instruments such as colorimeter and spectrophotometer. The higher the WI, the greater the whiteness degree of the measured cotton and vice versa (Topalovic et al. 2007). If the preferred white fabric has a high reflectance, then the ideal reflectance for textile materials should approach 100 (Ferdush et al. 2019).

The WI of the bleached cotton fabric is perpendicular to the time duration of the bleaching process and the amount of H_2O_2 (Haque and Islam 2015). Contrastingly, the bursting strength of the cotton fabric decreases at prolonged duration of the bleaching process and the increase in the H_2O_2 concentration. On the other hand, the higher temperature can improve the rate of bleaching and shorten the processing time (Abdul and Narendra 2013). Therefore, a colorimetric analysis is usually conducted to assess and investigate the bleaching procedure on the cotton samples. Artificial neural networks and adaptive neuro-inference systems have been used as prediction models in the textile domain. However, these models require large amount of data for model parameters optimisation and are quite time consuming. To address this issue, a fuzzy predictive model had been developed by Haque et al. (2018) using a fuzzy logic designer app in MATLAB to predict the WI of cotton using the bleaching process parameters that are nonlinear. Nevertheless, this fuzzy model is unable to predict the WI for the bleaching process parameters that are not within the ranges of the input data. It does not have the capability of machine learning models such as least square support vector regression (LSSVR).

Machine learning models including LSSVR models can learn information directly from data and understand their performance across a wide range of inputs (Wexler et al. 2019). LSSVR model exhibits good predictability to forecast the desired output variable, especially for nonlinear data. Hence, it received more attention and interest from researchers in many different areas over the years (Moosavi et al. 2021; Xu et al. 2013; Zhang and Wang 2021). But it was found that minimal research is conducted using the LSSVR model on color relevant studies including WI prediction.

Thus, in this study, an effective LSSVR model, namely multi-output LSSVR (MLSSVR) is developed using the bleaching process variables and results obtained from two different case studies to estimate the WI of bleached cotton. Then, the accuracy of LSSVR is evaluated by calculating the coefficient of determination (R^2), root mean square error (RMSE), and absolute mean error (MAE). Additionally, the obtained results are compared with partial least square regression (PLSR), predictive fuzzy model, locally weighted partial least square regression (LW-PLSR), and locally weighted kernel partial least square regression (LW-KPLSR) models.

Materials and methods

This section explains the bleaching process, post-treatment of cotton fabric, and WI measurement. It is followed by the MLSSVR model development, regression models parameters setting, and accuracy of the predictive performance measurement. Lastly, computer hardware and software configuration specifications are illustrated.

Bleaching process of cotton fabric and whiteness index

Case study 1

In this case study, the experimental data was taken from Haque et al. (2018). A single jersey cotton knitted fabric of 130 g/cm^2 was used as the fabric samples and a 12.5 g of fabric sample with a 1:10 material liquor ratio was treated in each bleaching time. The chemicals used in the bleaching process are H_2O_2 as bleaching agent (with its concentration set at 1.8 g/L, 2 g/L, and 2.2 g/L), 2 g/L of sodium hydroxide as caustic soda, and 1 g/L of kappazon H53 peroxide stabilizer. For each H_2O_2 concentration, the bleaching process was operated at six individual temperatures (T) (i.e., 78, 83, 88, 93, 98, 103 and 108 °C) and four different times (t) (i.e., 20, 30, 40, and 50 min).

Case study 2

Similar study was also conducted by Haque and Islam (2015) using a single jersey cotton fabric but with

lower weight (110 g/cm^2). H$_2$O$_2$ was also used as the bleaching agent with three different concentrations (1.8 g/L, 2 g/L, and 2.2 g/L). However, the bleaching solution was formulated differently. It contained 1 g/L of sodium hydroxide and 1 g/L of sodium silicate stabilizer (instead of kappazon H53 peroxide stabilizer). In terms of operation, the authors studied the effects of temperature (78, 88, 98, and 108 °C) and time (20, 30, 40 and 50 min) at four different intervals.

After bleaching process, both case studies subjected the bleached fabric samples to hot washing at 70 °C followed by cold washing at 27 °C. It was then squeezed by hand and dried at 70 °C for 30 min. Lastly, the WI for each bleached fabric sample was measured using a reflectance spectrophotometer (Data color 650). For case study 2, all the bleached fabric samples were further analysed using a bursting strength testing instrument (Autoburst, SDL Atlas) and by the ISO 13038–1 method. Figure 1 is the process flow involving bleaching operations, post-treatment of fabric samples, color, and bursting strength measurements.

Commission on Illumination (CIE) WI is one of the widely used color measurement methods for computing a WI to measure the degree of whiteness of bleached cotton fabric (Xu et al. 2015). This CIE WI generally refers to measurements made under D65 illumination, which is a standard representation of outdoor daylight. The CIE WI under CIE 1964 10° standard observer can be represented by Eq. 1 (Haque et al. 2018; Jafari and Amirshahi 2008).

$$WI = Y_L + 800(x_n - x_c) + 1700(y_n - y_c) \quad (1)$$

where Y_L is the lightness, x_c and y_c are chromaticity coordinates of the bleached cotton fabric samples, and x_n and y_n are chromaticity coordinates of the illuminant. However, the CIE WI has a constraint as shown in Eq. 2 (Jafari and Amirshahi 2008).

$$40 < WI < (5Y_L - 280) \quad (2)$$

Multi-output least square support vector regression model development

In this study, LSSVR model is developed from the bleaching process parameters to predict the WI of cotton fabrics. LSSVR model is a nonlinear prediction model that derives the support vector machine (SVM) theory (Liu and Yoo 2016). Different from the SVM, LSSVR gives a better solution for the reduction of the computational burden where a set of linear equations in a dual space is utilized. In this study, a MLSSVR model was adopted. Due to the multi-output setting in this MLSSVR, it becomes a more efficient training model. The examples of the results of the MLSSVR can be found in the work Xu et al. (2013) in which the
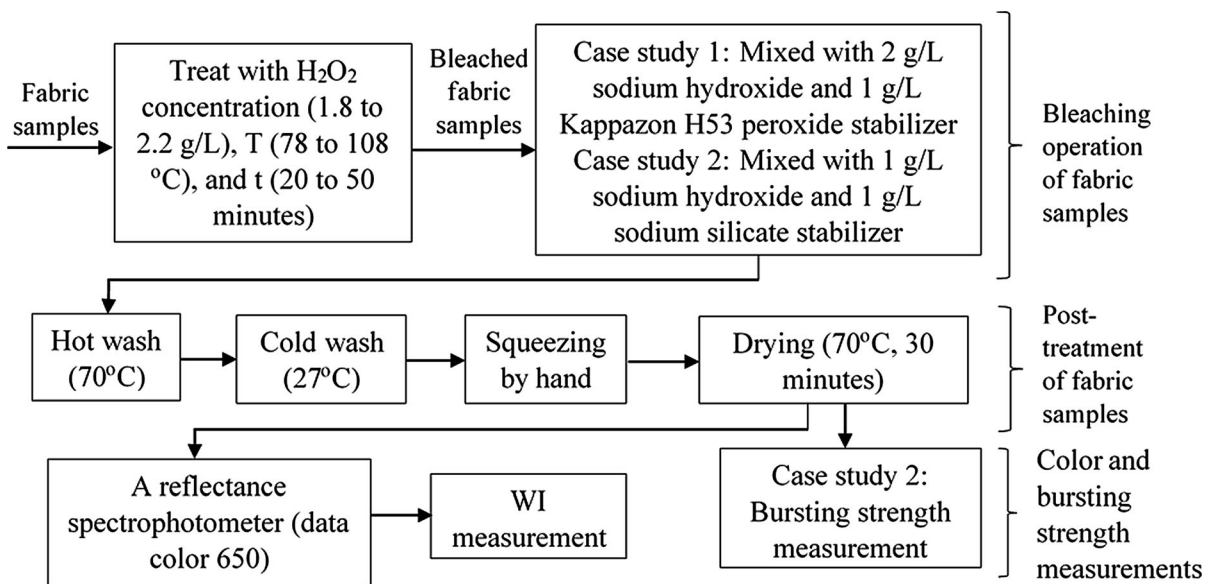

Fig. 1 Flowchart explaining the bleaching operations, post-treatment of fabric samples, color, and bursting strength measurements

MLSSVR were tested using synthetic, corn, polymer, and broomcorn data sets. However, this MLSSVR model has yet been used to examine cotton fabric data set.

The idea of MLSSVR comes from the multi-output case done by An et al. (2009). It is letting $Y = [y_{i,j}] \in \Re^{l \times m}$ where $y_{i,j}$ is the (i,j)-th of an output, $\Re$ is the set of real numbers, and $l \times m$ is the order of a matrix. With a given total number of data sets, N_T, i.e., $\{(x_i, y^i)\}_{i=1}^l$ where $x_i \in \Re^d$ and $y^i \in \Re^m$ are the input vector and output vector, respectively. And the multi-output regression has an objective to estimate an output vector $y \in \Re^m$ from a given input vector $x \in \Re^d$ where this regression problem can be built as learning a mapping from $\Re^d$ to $\Re^m$. Multi-output regression solves the problem by searching the weighed value vector, $W = (w_1, w_2, ..., w_m) \in \Re^{n_h \times m}$ and a threshold value, $b = (b_1, b_2, ..., b_m)^T \in \Re^m$ that minimises the following objective function with constraints (Eqs. 3 and 4). According to the structural risk minimization principle, the actual risk bound by the empirical risk is minimized by the objective function (Eq. 3) with its constraint (Eq. 4). The real signification on these equations is to minimize the frequency of error in the training data set used to develop the MLSSVR model. Hence, the probability of error on the testing data set is expected to be small. For more details about the structural risk minimization principle, one is referred to the work of Vapnik (1992).

$$\begin{array}{c} \min \\ W \in \Re^{m \times n_h}, b \in \Re^m \end{array} \Im(W, \Xi)$$
$$= \frac{1}{2} trace(W^T W) + \gamma \frac{1}{2} trace(\Xi^T \Xi), \qquad (3)$$

$s.t.\ Y = Z^T W + repmat(b^T, l, 1) + \Xi, \qquad (4)$

where γ is a positive real regularised parameter, $\xi = (\xi_1, \xi_2, ..., \xi_l)^T \in \Re^l$ is a vector containing slack variables, $Z = (\phi(x_1), \phi(x_2), ..., \phi(x_l)) \in \Re^{n_h \times l}$, $\phi: \Re^d \to \Re^{n_h}$ is a mapping to some high or even unlimited/ infinite dimensional Hilbert space or feature space via the nonlinear mapping function ϕ with n_h dimensions, and $\Xi = (\xi_1, \xi_2, ..., \xi_m) \in \Re_+^{l \times m}$ is a $l \times m$ matrix consisting of slack variables with $\Re_+$ the subset of positive ones. The slack variables are used to improve not only the generalization model performance but also to yield more compact and lower complexity model (Tang et al. 2015). These slack variables are non-negative and can tolerate misclassification in the training data that is used to develop the MLSSVR model.

It can be said that the solution to the regression problem shown in Eqs. 3 and 4 disconnects between the different output variables and only need to use Cholesky decomposition, conjugate gradient, or single value decomposition, etc. to compute a single inverse matrix once that is shared by all the vectors $w_i(\forall_i \in N_m)$. Unlike the single-output case, its solution to the regression problem needs to be solved multiple times. Hence, the multi-output regression is much more efficient than the single-output regression.

According to Xu et al. (2013), to formulate the intuition of Hierarchical Bayes, all $w_i \in \Re^{n_h} (i \in N_m)$ is assumed to be written as $w_i = w_0 + v_i$, where the vectors $v_i \in \Re^{n_h} (i \in N_m)$ are small when the different outputs are same to each other, otherwise the mean vector $w_0 \in \Re^{n_h}$ are small. It can be said that w_0 takes the information of the commonality and $v_i(i \in N_m)$ brings the information of the specialty. $w_0 \in \Re^{n_h}$, $V = (v_1, v_2, ..., v_m) \in \Re^{n_h \times m}$, and $b = (b_1, b_2, ..., b_m)^T \in \Re^m$ are solved spontaneously to minimise the following objective function with constraints (Eqs. 5 and 6):

$$\begin{array}{c} \min \\ w_0 \in \Re^{n_h}, V \in \Re^{n_h}, b \in \Re^m \end{array} \Im(w_0, V, \Xi)$$
$$= \frac{1}{2} w_0^T w_0 + \frac{1}{2} \frac{\lambda}{m} trace(V^T V) + \gamma \frac{1}{2} trace(\Xi^T \Xi), \qquad (5)$$

$s.t.\ Y = Z^T W + repmat(b^T, l, 1) + \Xi, \qquad (6)$

where $\Xi = (\xi_1, \xi_2, ..., \xi_m) \in \Re^{l \times m}$, $W = (w_0 + v_1, w_0 + v_2, ..., w_0 + v_m) \in \Re^{n_h \times m}$, $\lambda, \gamma \in \Re_+$ are two positive real regularised parameters, and $Z = (\phi(x_1), \phi(x_2), ..., \phi(x_l)) \in \Re^{n_h \times l}$.

The Lagrangian function for the problem shown in Eqs. 5 and 6 is defined as (Eq. 7):

$$\ell(w_0, V, b, \Xi, A) = \Im(w_0, V, \Xi) - trace(A^T(Z^T W + repmat(b^T, l, 1) + \Xi - Y)), \qquad (7)$$

where $A = (\alpha_1, \alpha_2, ..., \alpha_m) \in \Re^{l \times m}$ is a matrix containing of Lagrange multipliers. The Karush-Kuhn-Tucker conditions for optimality result in the linear equations as follows (Eq. 8):

$$\begin{cases} \dfrac{\partial \ell}{\partial w_0} = 0 \Rightarrow w_0 = \sum_{i=1}^{m} Z\alpha_i, \\[2mm] \dfrac{\partial \ell}{\partial V} = 0 \Rightarrow V = \dfrac{m}{\lambda} ZA, \\[2mm] \dfrac{\partial \ell}{\partial b} = 0 \Rightarrow A^T 1_l = 0_l, \\[2mm] \dfrac{\partial \ell}{\partial \Xi} = 0 \Rightarrow A = \gamma \Xi, \\[2mm] \dfrac{\partial \ell}{\partial A} = 0 \Rightarrow Z^T W + repmat(b^T, l, 1) + \Xi - Y = 0_{l \times m} \end{cases}$$

$$(8)$$

From Eq. 8, the mean vector, w_0 is a linear combination of $v_1, v_2, ..., v_m$. As mentioned earlier since $\forall_i \in N_m$, so w_i is assumed to be $w_i = w_0 + v_i$ in which w_i is also a linear combination of $v_1, v_2, ..., v_m$. Hence, the following objective function (Eqs. 9 and 10) can obtain an equivalent optimisation problem with constraints including only the V and b.

$$\min_{V \in \Re^{n_h}, b \in \Re^m} \Im(V, \Xi) = \frac{1}{2} \frac{\lambda^2}{m^2} V 1_m 1_m^T V^T$$
$$+ \frac{1}{2} \frac{\lambda}{m} trace(V^T V)$$
$$+ \gamma \frac{1}{2} trace(\Xi^T \Xi), \quad (9)$$

$$s.t.\ Y = Z^T V + repmat(\frac{\lambda}{m} Z^T V 1_m, 1, m)$$
$$+ repmat(b^T, l, 1) + \Xi \quad (10)$$

From Eq. 9, MLSSVR figures out a trade-off between small size vectors for every output,$trace(V^T V)$ and nearness of all vectors to the mean vector, $V 1_m 1_m^T V^T$. Like the standard LSSVR, W and Ξ are discharged to get the linear system as expressed in Eq. 11.

$$\begin{bmatrix} 0_{ml \times m} & P^T \\ P & H \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0_m \\ y \end{bmatrix}_m \quad (11)$$

where $P = blockdiag \overbrace{(1_l, 1_l, ..., 1_l)}^{m} \in \Re^{ml \times m}$, $H = \Omega + \gamma^{-1} I_{ml} + (\frac{m}{\lambda}) Q \in \Re^{ml \times ml}$, $\Omega = repmat(K, m, m)$ $\in \Re^{ml \times ml}$, $Q = blockdiag \overbrace{(K, K, ...K)}^{m} \in \Re^{ml \times ml}$, and $K = Z^T Z \in \Re^{l \times l}$ are definite matrices while $\alpha = (\alpha_1^T, \alpha_2^T, ..., \alpha_m^T)^T \in \Re^{ml}$ and $y = (y_1^T, y_2^T, ..., y_m^T)^T \in \Re^{ml}$ are vectors. Hence, the linear system shown in the Eq. 11 has $(l + 1) \times m$ equations.

Then, the solution of Eq. 11 can be written in term of $\alpha^* = (\alpha_1^{*T}, \alpha_2^{*T}, ..., \alpha_m^{*T})^T$ and b^*. Hence, the respective decision function for the multiple output is expressed as Eq. 12.

$$f(x) = \phi(x)^T W^* + b^{*T} = \phi(x)^T repmat(w_0^*, 1, m)$$
$$+ \phi(x)^T V^* + b^{*T}$$
$$= \phi(x)^T repmat(\sum_{il}^{m} Z\alpha_{il}^*, 1, m) + \frac{m}{\lambda} \phi(x)^T Z A^* + b^{*T}$$
$$= repmat(\sum_{il=1}^{m} \sum_{j=1}^{l} \alpha_{il,j}^* K(x, x_j), 1, m)$$
$$+ \frac{m}{\lambda} \sum_{j=1}^{l} \alpha^{j*} K(x, x_j) + b^{*T}$$

$$(12)$$

Same as the conventional LSSVR, the linear system of MLSSVR as displayed in Eq. 11 is not positive define, hence solving Eq. 11 instantly is hard. But it can be reconstructed into the following linear system (Eq. 13):

$$\begin{bmatrix} S & 0_{ml \times ml} \\ 0_{m \times m} & H \end{bmatrix} \begin{bmatrix} b \\ H^{-1} Pb + \alpha \end{bmatrix} = \begin{bmatrix} P^T H^{-1} y \\ y \end{bmatrix},$$

$$(13)$$

with $S = P^T H^{-1} P \in \Re^{m \times m}$. Notice that it is easy to display S that is a positive definite matrix. Then, this new linear system as shown in Eq. 13 can be solved using the following three steps:

Step 1 Solve η, and v from $H\eta = P$ and $Hv = y$,

Step 2 Compute $S = P^T \eta$,

Step 3 Obtain the solution:$b = S^{-1} \eta^T y$, $\alpha = v - \eta b$.

Thus, in MLSSVR, the solution of the training procedure can be obtained by solving two sets of linear equations with the same positive definite coefficient matrix $H \in \Re^{ml \times ml}$ and the inverse matrix of $S \in \Re_+^{m \times m}$ can be computed easily (Xu et al. 2013). Figure 2 illustrates the training procedure of MLSSVR involving the mathematical equations.

In this study, the radial basis function (RBF) kernel function adopted from Keerthi and Lin (2003) as shown in Eq. 14 is used in the MLSSVR.

$$k(x, z) = \exp(-p \|x - z\|^2), p > 0 \quad (14)$$

where p is the positive hyperparameter of RBF kernel function. Moreover, all the tuning parameters in MLSSVP including γ, λ, and p are tuned and optimised
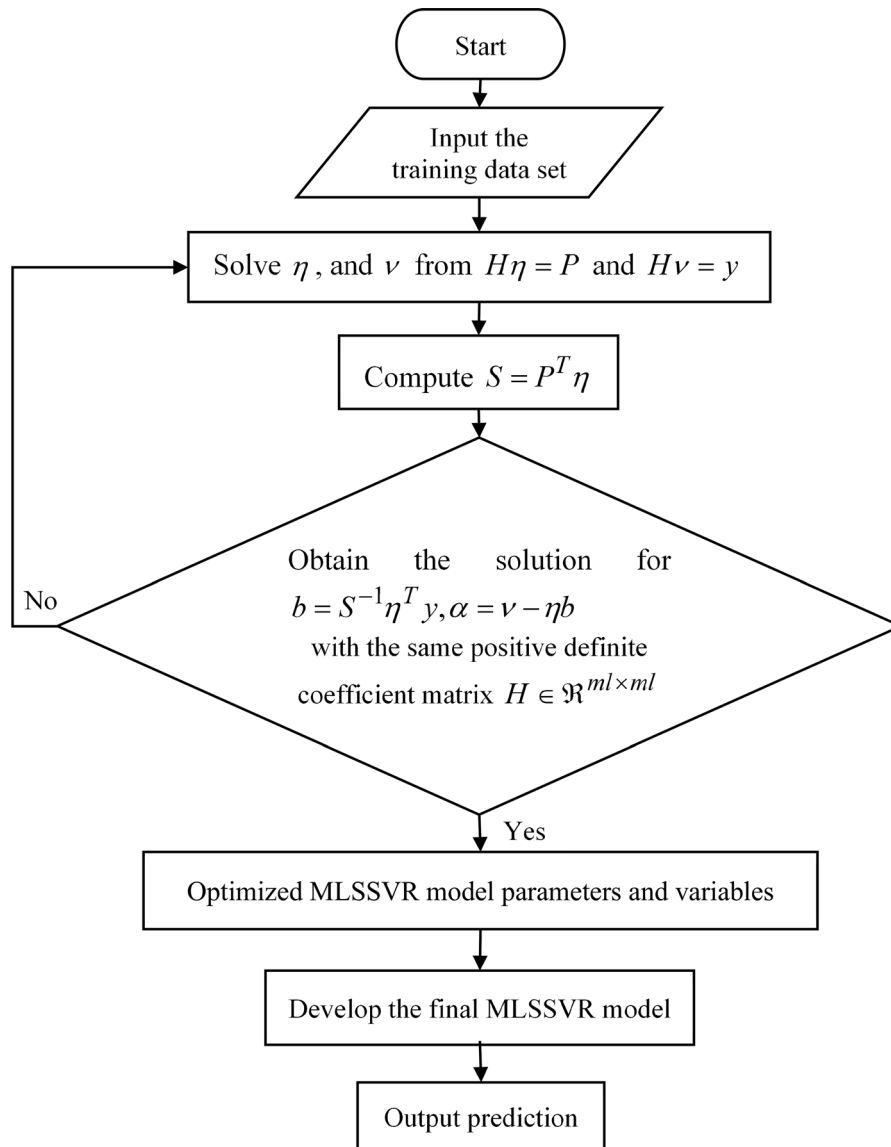
Fig. 2 Flowchart explaining the training procedure of MLSSVR involving the mathematical equations

using leave-one-out (LOO) procedure to obtain the average relative error, δ as shown in Eq. 15 (Xu et al. 2013).

$$\delta = \frac{1}{l} \sum_{i=1}^{l} \frac{\left| Y_i - \hat{Y}_i \right|}{Y_i} \qquad (15)$$

whereby Y_i shows the actual output and $\hat{Y}_i$ shows the predicted output.

Regression models parameters setting

For case study 1, a total of 56 bleaching data sets which consists of H_2O_2 concentration, T, t, and the WI of the cotton fabric were adopted from Haque et al. (2018). In this study, H_2O_2 concentration, T and t are served as the input variables for the regression models including MLSSVR, PLSR, LW-PLSR, and LW-KPLSR models while the WI of the bleached cotton fabric is denoted as the output variable. Meanwhile, for case study 2, 48 bleaching data sets including H_2O_2

concentration, T, t, bursting strength (psi) and the WI of the cotton fabric were taken from Haque and Islam (2015). Different from case study 1, an additional input variable—bursting strength of bleached cotton fabric is included in the regression models. For both case studies, their datasets were exported into MATLAB software and divided into two sets, i.e., 75% for training data and 25% for testing data. Hence, they were split into 40 data sets for case study 1 and 36 data sets for case study 2 (see Table 1) to serve as training data utilized to develop the regression models and fuzzy method. Then, 16 and 12 data sets for case study 1 and 2, respectively were used as testing data for validation purposes. Moreover, training data were also employed to evaluate the performance of the MLSSVR, PLSR, LW-PLSR, and LW-KPLSR models as well as the fuzzy method. Then, RMSE, MAE, and R^2 for all regression models and fuzzy method were determined and compared. Figure 3 shows a flow chart explaining the framework of the regression models and fuzzy method for the bleaching process.

N_T, N_1, N_2, and latent variable (LV) represent the total numbers of data sets, numbers of training data sets, numbers of testing data sets, and number of LV, respectively. In this study, LV is set as 1 while the kernel parameter (b_k) for LW-KPLSR with log kernel function is set as 1. The value of phi in the LW-PLSR and LW-KPLSR is fixed at 0.1 (Yeo et al. 2017). Besides, some parameters for MLSSVR model which are, γ, λ, and p were tuned using LOO technique to get the optimal results. The summarised parameters setting for MLSSVR, PLSR, LW-PLSR, and LW-KPLSR models are displayed in Table 1.

Accuracy of the predictive performance measurement

In this study, the performance of the prediction models is evaluated using RMSE, MAE, R^2 and prediction error (PE). Both RMSE and MAE are goodness-of-fit indicators that describe differences in observed and predicted values (Harmel et al. 2010). RMSE as shown in Eq. 16 is the square root of the total of the squared differences between the actual and expected output. Thus, a lower RMSE implies better accuracy and predictive performance (Hocaoğlu et al. 2008).

$$RMSE = \sqrt{\frac{\sum_i \left(Y_i - \hat{Y}_i\right)^2}{n}} \tag{16}$$

whereby n shows the number of samples.

MAE calculates the average absolute difference between the actual and predicted output value. The formula to calculate MAE is displayed in Eq. 17:

$$MAE = \frac{1}{n} \sum_1^n \left| Y_i - \hat{Y}_i \right| \tag{17}$$

As can be seen from Eq. 18, R^2 is obtained by comparing the total of the squared errors to the total of the squared deviations about its mean. R^2 is used to measure the goodness of fit between real and predicted variables with value ranging between 0 and 1 (Jaeger et al. 2017).

$$R^2 = 1 - \frac{\sum_i \left(Y_i - \hat{Y}_i\right)^2}{\sum_i \left(Y_i - \overline{Y}\right)^2} \tag{18}$$

whereby $\overline{Y}$ represents the mean value of the actual output.

Additionally, PE (%) can be calculated using Eq. 19 (Guang et al. 1995).

$$PE = \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right| \times 100 \tag{19}$$

Computer hardware and software configuration specifications

In this study, all simulation works were performed on the same computer and software system to ensure the consistency of the results from all regression models. The hardware and software configuration

Table 1 Values used for the regression models in two different case studies

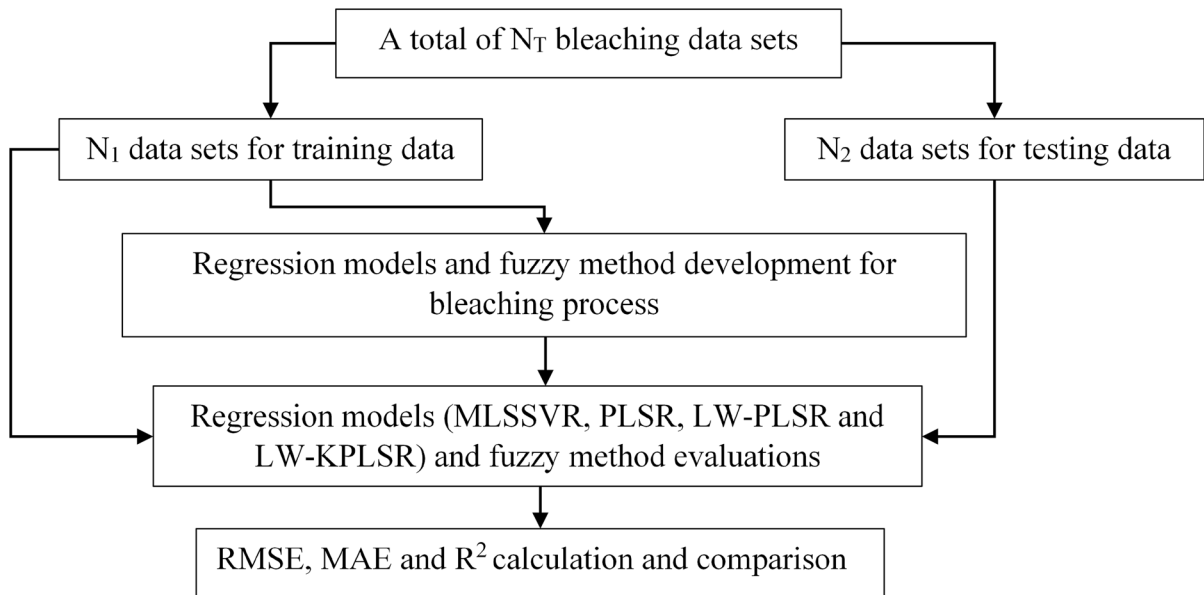| Parameters	N_T	N_1	N_2	LV	phi	b_k	γ	λ	p
Values (case study 1)	56	40	16	1	0.1	1	15	10	3
Values (case study 2)	48	36	12	1	0.1	1	15	10	3

Fig. 3 Framework of regression models and fuzzy method for the bleaching process

specifications of the used Asus ZenBook UX305 laptop are Window 10 (64 bit), MATLAB (version R2021a), 2.20 GHz Intel core M3-6Y30, 4.0 GB random-access memory and 128 GB solid state drive storage.

Results and discussion

As mentioned earlier, for case study 1, an LSSVR model that is called MLSSVR was developed using several key parameters of bleaching process, and these parameters including H_2O_2 concentrations, T, t, and WI of cotton fabric samples are nonlinear. Case study 2 is very similar to case study 1 except one additional parameter, i.e., bursting strength of cotton fabric samples was included to build MLSSVR model. The higher value of WI indicates the greater degree of whiteness of the cotton fabric. Whiter cotton fabric is desired before it is dyed, printed or other wet-treatments (Ferreira et al. 2019; Kabir et al. 2014). Many studies reported that WI of cotton fabric can be increased by increasing H_2O_2 concentrations and T (Abdul and Narendra 2013; Ferdush et al. 2019). Nevertheless, high concentrations of H_2O_2 could break up the unsaturated bonds (e.g., C=C) and decreases the bursting strength of cotton fabric (Ferdush et al. 2019; Haque and Islam 2015; Tang

et al. 2016). Hence, optimum bleaching process parameters are required to determine the effectiveness of the bleaching process to produce a targeted cotton fabric whiteness.

The predictive modelling techniques including the MLSSVR model can be used to estimate the outcome or the results of the bleached cotton fabric such as its WI using the bleaching process parameters. Initially, a fuzzy predictive model was constructed by Haque et al. (2018) for a bleaching process using a MATLAB app that is called a fuzzy logic designer. However, this method is unable to predict beyond the range of the input data. Hence, in this study, an MLSVVR was developed using the bleaching process parameters reported in two different case studies to overcome the limitations of this fuzzy method. Moreover, other regression models including PLSR, LW-PLSR, and LW-KPLSR models were also built using the same process parameters from case study 1 and 2.

All results from these regression models and fuzzy method for case study 1 are summarized in Table 2 for comparison purpose and the results for fuzzy method were adopted from Haque et al. (2018). Moreover, the results from the regression models and fuzzy method for case study 2 are presented in Table 3 where the results for fuzzy method were generated from the fuzzy logic designer app in the MATLAB. In Tables 2 and 3, $RMSE_1$, MAE_1, and R^2_1 are the RMSE, MAE,

Table 2 Comparison of the results obtained from different models for case study 1

Results	MLSSVR	Fuzzy method	PE (%)	PLSR	PE (%)	LW-PLSR	PE (%)	LW-KPLSR	PE (%)
Kernel function	RBF	–	–	–	–	–	–	Log kernel	–
$RMSE_1$	**0.1606**	0.7373	359	2.1014	1209	0.3335	108	0.4755	196
MAE_1	**0.1126**	0.6133	445	1.5981	1320	0.2560	127	0.4088	263
R^2_1	**0.9985**	0.9673	3	0.6469	35	0.9934	1	0.9863	1
$RMSE_2$	**0.3339**	0.5358	60	1.2194	265	0.8122	143	0.6714	101
MAE_2	**0.2388**	0.4781	234	1.0427	337	0.7101	197	0.5861	145
R^2_2	**0.9829**	0.9549	3	0.8357	15	0.9103	7	0.9334	5

Table 3 Comparison of the results obtained from different models for case study 2

Results	MLSSVR	Fuzzy method	PE (%)	PLSR	PE (%)	LW-PLSR	PE (%)	LW-KPLSR	PE (%)
Kernel function	RBF	–	–	–	–	–	–	Log kernel	–
$RMSE_1$	**0.0408**	3.3523	8107	2.1117	5070	0.2304	464	0.4025	885
MAE_1	**0.0274**	1.5914	5706	1.5815	5670	0.2024	638	0.3311	1108
R^2_1	**0.9999**	0.4206	58	0.6589	34	0.9970	0.30	0.9906	0.93
$RMSE_2$	**0.2972**	6.4790	2080	1.8877	535	0.8185	175	0.9051	205
MAE_2	**0.2302**	6.1358	2566	1.6081	599	0.7822	240	0.8721	279
R^2_2	**0.9810**	− 0.1150	112	0.6828	30	0.8825	10	0.8337	15

and R^2 for training data, respectively whereas $RMSE_2$, MAE_2, and R^2_2 are the RMSE, MAE, and R^2 for testing data, respectively. From both tables, among these regression models and fuzzy method, MLSSVR with RBF kernel function provided the best predictive performance. However, the results of MLSSVR for case study 2 (Table 3), especially the results for training data ($RMSE_1$, MAE_1, and R^2_1) are much better compared to the case study 1 (Table 2).

Based on Tables 2 and 3, PLSR did not work well for both case studies as its RMSE and MAE values for case studies 1 and 2 are significantly higher than MLSSVR. Furthermore, its R^2 values for both case studies are lower than 0.84. PLSR which is a linear model gave the worse results as it is unable to cope with the nonlinear process data from bleaching process (Wang et al. 2021). Additionally, compared to fuzzy method, MLSSVR showed 60–445% (for case study 1) and 2080–8107% (for case study 2) lower RMSE and MAE values, and 3–112% higher R^2 values. The fuzzy method involves membership function, fuzzy logic operators, and if-then rules. There are three conceptual components such as a rule

case that include a selection of fuzzy rules, a database which explains the membership functions used in the fuzzy rules, and a reasoning mechanism that shows the inference way upon the rules to derive an output (Kovac et al. 2013). From Table 2, for testing data set, fuzzy method performed better than PLSR, LW-PLSR, and LW-KPLSR in which the $RMSE_2$ and MAE_2 for fuzzy method are lower and its R^2_2 values are higher. However, in Table 3, fuzzy method performed poorer than PLSR, LW-PLSR, and LW-KPLSR.

In addition, the overall results show that the fuzzy method worked badly compared to MLSSVR. This may be due to the helps of LOO in the MLSSVR to determine the optimal tuning parameters and the RBF kernel function that helps to map the original data into a high dimensional space for better prediction of the nonlinear data. For both LW-PLSR and LW-KPLSR, it is found that MLSSVR demonstrated 101–263% for case study 1 and 175–1108% for case study 2 lower RMSE and MAE values and 1–7% for case study 1 and 0.30–15% for case study 2 higher R^2 values. On the other hand, LW-PLSR and LW-KPLSR gave better

results than fuzzy method for the training data set of case study 1 where their $RMSE_1$ and MAE_1 are lower and R^2_1 are higher. Moreover, for case study 2, LW-PLSR and LW-KPLSR also performed better than fuzzy method and PLSR. This may be due to the presence of locally weighted algorithm in both LW-PLSR and LW-KPLSR which improves their predictive performance for the training data. These locally weighted based models (LW-PLSR and LW-KPLSR) are using a PLSR model that is constructed by weighting data with Euclidian distances similarity measurement (Kaneko and Funatsu 2016).

From Figs. 4 and 5, it is obvious that the predicted outputs for training data and testing data of case stud1 1 and 2 from MLSSVR are closer to the actual data compared to other regression models and fuzzy method. From Fig. 4b for testing data, except for MLSSVR, other regression models and fuzzy method performed poorer compared to training data in Fig. 4a in which their predicted outputs are far from the actual output. Similarly, besides MLSSVR in Fig. 5a, b for training and testing data, the predicted outputs from the rest of the regression model and fuzzy model are far apart from the actual data. Moreover, Fig. 6a, b illustrate the correlation between the actual and predicted values of output from MLSSVR for testing data of case study 1 and 2, respectively. From these figures, notice that all data points are close to the line which indicate that the predicted outputs from MLSSVR are near to the actual data for testing data

of both case studies. In general, all the results show that MLSSVR copes much better than the rest of the methods. Hence, it can conclude that MLSSVR is an effective method to predict the WI using the bleaching process parameters.

Conclusions

In the current study, LSSVR model, namely MLSSVR was developed using the experimental results of bleaching process from two different case studies to predict the WI of the cotton fabric. The input variables for case study 1 are H_2O_2 concentrations, T, and t while case study 2 contains 1 additional input variable, i.e., bursting strength. It is important to determine the optimal bleaching process parameters in order to achieve the highest WI of the cotton fabrics. Hence, the predictive modelling including MLSSVR plays an essential role to meet the targeted quality of cotton fabrics in the textile manufacturing processes. For both case studies, the developed MLSSVR could outperform other methods including fuzzy method, PLSR, LW-PLSR, and LW-KPLSR as it improved RMSE and MAE values by 60–8107%. Besides, its R^2 values in both case studies are also very high, reaching up to 0.9999. These results denote that MLSSVR model is a potential predictive model for the bleaching process in the textile domain. In future studies, integrating a locally weighted algorithm in the
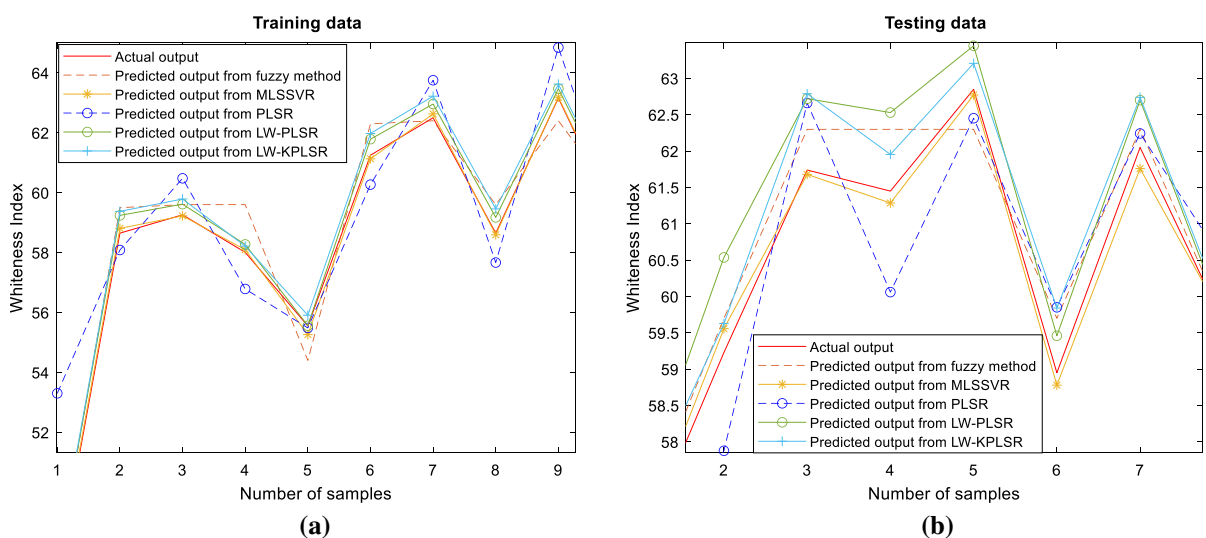

Fig. 4 Comparison of the predicted output values from regression models **a** for training data and **b** for testing data of case study 1
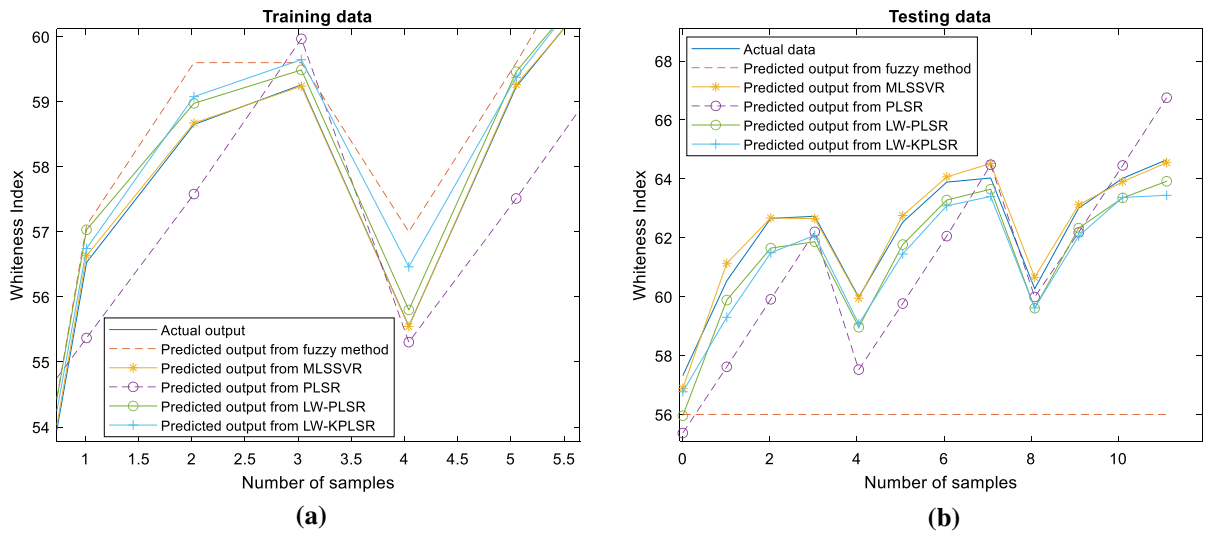
Fig. 5 Comparison of the predicted output values from regression models **a** for training data and **b** for testing data of case study 2
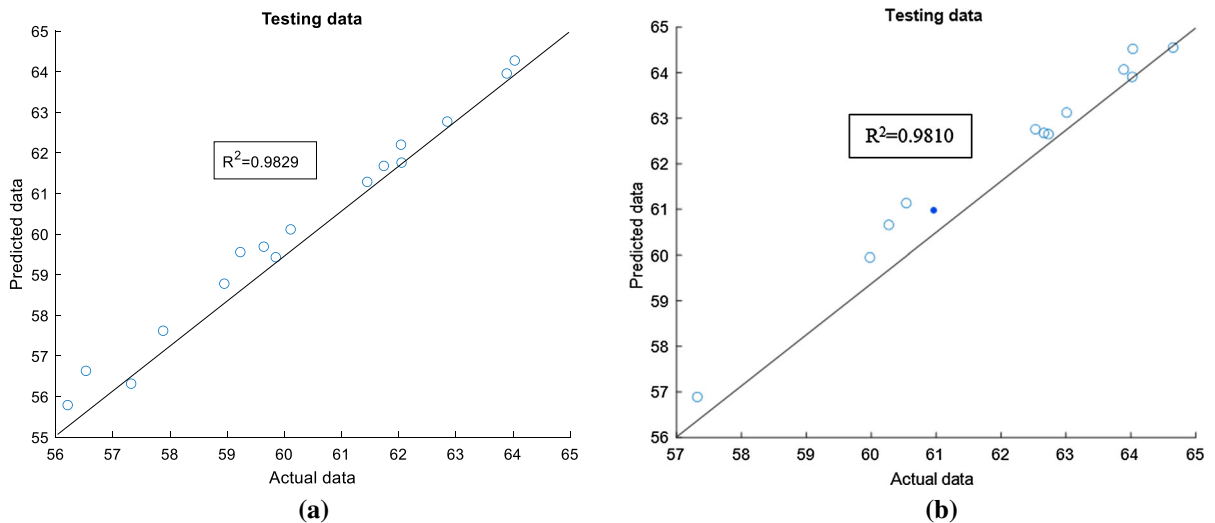

Fig. 6 Correlation between the actual and predicted values of output from MLSSVR for testing data of **a** Case study 1 and **b** Case study 2

MLSSVR model could be expected to enhance its predictive outcomes. Furthermore, more attention should be paid to the new aspects of the textile fiber bleaching process and the cost of the process optimization.

Author contributions The authors discussed the results and contributed to the final manuscript.

Declarations

Conflict of interest The authors have no conflict of interest to declare. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval This research program does not involve testing to be done on humans or animals. It also does not involve

any potentially dangerous equipment and hazardous substance of any kind. Therefore, no ethical issue will be expected in this research project.

Human or animal rights No animal studies or human participants involvement in the study, hence this research project is compliance with ethical standards.

References

Abdul S, Narendra G (2013) Accelerated bleaching of cotton material with hydrogen peroxide. J Text Sci Eng 3:1000140. https://doi.org/10.4172/2165-8064.1000140

Ahmad S, Huifang W, Akhtar S, Imran S, Yousaf H, Wang C, Akhtar MS (2021) Impact assessment of better management practices of cotton: a sociological study of southern Punjab, Pakistan. Pak J Agric Sci 58:291–300. https://doi.org/10.21162/PAKJAS/21.227

An X, Xu S, Zhang L-D, Su S-G (2009) Multiple dependent variables LS-SVM regression algorithm and its application in NIR spectral quantitative analysis. Spectrosc Spectr Anal 29:127–130. https://doi.org/10.3964/j.issn.1000-0593(2009)01-0127-04

Bajpai D (2007) Laundry detergents: an overview. J Oleo Sci 56:327–340. https://doi.org/10.5650/jos.56.327

Ferdush J, Nahar K, Akter T, Ferdoush MJ, Jahan N, Iqbal SF (2019) Effect of hydrogen peroxide concentration on 100% cotton knit fabric bleaching. ESJ 15:254–263. https://doi.org/10.19044/esj.2019.v15n33p254

Ferreira ILS, Medeiros I, Steffens F, Oliveira FR (2019) Cotton fabric bleached with seawater: mechanical and coloristic properties. Mater Res 22:e20190085. https://doi.org/10.1590/1980-5373-MR-2019-0085

Guang W, Baraldo M, Furlanut M (1995) Calculating percentage prediction error: a user's note. Pharmacol Res 32:241–248. https://doi.org/10.1016/S1043-6618(05)80029-5

Gültekin BC (2016) Bleaching of SeaCell® active fabrics with hydrogen peroxide. Fibers Polym 17:1175–1180. https://doi.org/10.1007/s12221-016-6181-9

Haque A, Islam MA (2015) Optimization of bleaching parameters by whiteness index and bursting strength of knitted cotton fabric. Int J Sci Technol Res 4:40–43

Haque ANMA, Smriti SA, Hussain M, Farzana N, Siddiqa F, Islam MA (2018) Prediction of whiteness index of cotton using bleaching process variables by fuzzy inference system. Fash Text 5:1–13. https://doi.org/10.1186/s40691-017-0118-9

Harmel RD, Smith PK, Migliaccio KW (2010) Modifying goodness-of-fit indicators to incorporate both measurement and model uncertainty in model calibration and validation. Trans ASABE 53:55–63. https://doi.org/10.13031/2013.29502

Hocaoğlu FO, Gerek ÖN, Kurban M (2008) Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks. Solar energ 82:714–726. https://doi.org/10.1016/j.solener.2008.02.003

Jaeger BC, Edwards LJ, Das K, Sen PK (2017) An R 2 statistic for fixed effects in the generalized linear mixed model. J Appl Stat 44:1086–1105. https://doi.org/10.1080/02664763.2016.1193725

Jafari R, Amirshahi S (2008) Variation in the decisions of observers regarding the ordering of white samples. Color Technol 124:127–131. https://doi.org/10.1111/j.1478-4408.2008.00132.x

Jung H, Sato T (2013) Comparison between the color properties of whiteness index and yellowness index on the CIELAB. Text Coloration Finish 25:241–246. https://doi.org/10.5764/TCF.2013.25.4.241

Kabir SF, Iqbal MI, Sikdar PP, Rahman MM, Akhter S (2014) Optimization of parameters of cotton fabric whiteness. Eur Sci J 10:200–210

Kaneko H, Funatsu K (2016) Ensemble locally weighted partial least squares as a just-in-time modeling method. AIChE J 62:717–725. https://doi.org/10.1002/aic.15090

Keerthi SS, Lin C-J (2003) Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Comput 15:1667–1689. https://doi.org/10.1162/089976603321891855

Kovac P, Rodic D, Pucovsky V, Savkovic B, Gostimirovic M (2013) Application of fuzzy logic and regression analysis for modeling surface roughness in face milliing. J Intell Manuf 24:755–762. https://doi.org/10.1007/s10845-012-0623-z

Liu H, Yoo C (2016) A robust localized soft sensor for particulate matter modeling in Seoul metro systems. J Hazard Mater 305:209–218. https://doi.org/10.1016/j.jhazmat.2015.11.051

Moosavi SR, Vaferi B, Wood DA (2021) Auto-characterization of naturally fractured reservoirs drilled by horizontal well using multi-output least squares support vector regression. Arab J Geosci 14:1–12. https://doi.org/10.1007/s12517-021-06559-9

Oliveira BPd, Moriyama LT, Bagnato VS (2018) Colorimetric analysis of cotton textile bleaching through H2O2 activated by UV light. J Braz Chem Soc 29:1360–1365. https://doi.org/10.21577/0103-5053.20170235

Tang P, Ji B, Sun G (2016) Whiteness improvement of citric acid crosslinked cotton fabrics: H2O2 bleaching under alkaline condition. Carbohydr polym 147:139–145. https://doi.org/10.1016/j.carbpol.2016.04.007

Tang F, Tiňo P, Gutiérrez PA, Chen H (2015) The benefits of modeling slack variables in svms. Neural Comput 27:954–981. https://doi.org/10.1162/NECO_a_00714

Topalovic T, Nierstrasz VA, Bautista L, Jocic D, Navarro A, Warmoeskerken MM (2007) Analysis of the effects of catalytic bleaching on cotton. Cellulose 14:385–400. https://doi.org/10.1007/s10570-007-9120-5

Vapnik V (1992) Principles of risk minimization for learning theory. In: NIPS'91: Proceedings of the 4th international conference on neural information processing systems, pp 831–838. https://doi.org/10.5555/2986916.2987018

Wang X, Hu P, Zhen L, Peng D (2021) Drsl: deep relational similarity learning for cross-modal retrieval. Inf Sci 546:298–311. https://doi.org/10.1016/j.ins.2020.08.009

Wang D, Zhong L, Zhang C, Zhang F, Zhang G (2018) A novel reactive phosphorous flame retardant for cotton fabrics with durable flame retardancy and high whiteness due to

self-buffering. Cellulose 25:5479–5497. https://doi.org/10.1007/s10570-018-1964-3

Wexler J, Pushkarna M, Bolukbasi T, Wattenberg M, Viégas F, Wilson J (2019) The What-If Tool: interactive probing of machine learning models. IEEE Trans Vis Comput Gr 26:56–65. https://doi.org/10.1109/TVCG.2019.2934619

Xie K, Gao A, Zhang Y (2013) Flame retardant finishing of cotton fabric based on synergistic compounds containing boron and nitrogen. Carbohydr Polym 98:706–710. https://doi.org/10.1016/j.carbpol.2013.06.014

Xu C, Hinks D, Sun C, Wei Q (2015) Establishment of an activated peroxide system for low-temperature cotton bleaching using N-[4-(triethylammoniomethyl) benzoyl] butyrolactam chloride. Carbohydr Polym 119:71–77. https://doi.org/10.1016/j.carbpol.2014.11.054

Xu S, An X, Qiao X, Zhu L, Li L (2013) Multi-output least-squares support vector regression machines. Pattern Recogn Lett 34:1078–1084. https://doi.org/10.1016/j.patrec.2013.01.015

Yeo WS, Saptoro A, Kumar P (2017) Development of adaptive soft sensor using locally weighted Kernel partial least square model. Chem Prod Process Model 12:1–13. https://doi.org/10.1515/cppm-2017-0022

Zhang J, Wang Y (2021) Evaluating the bond strength of FRP-to-concrete composite joints using metaheuristic-optimized least-squares support vector regression. Neural Comput Appl 33:3621–3635. https://doi.org/10.1007/s00521-020-05191-0