ORIGINAL ARTICLE

Nondestructive determination of soluble solids content and pH in red bayberry (*Myrica rubra*) based on color space

Jie Feng[1,2] · Lingling Jiang[3] · Jialei Zhang[1,2] · Hong Zheng[3] · Yanfang Sun[1,2] ·
Shaoning Chen[1,2] · Meilan Yu[1,2] · Wei Hu[1,2] · Defa Shi[4] · Xiaohong Sun[5] ·
Hongfei Lu[1,2]

Abstract Color has strong relationship with food quality. In this paper, partial least square regression (PLSR) and least square-support vector machine (LS-SVM) models combined with six different color spaces (NRGB, CIE-LAB, CMY, HSI, I1I2I3, and YCbCr) were developed and compared to predict pH value and soluble solids content (SSC) in red bayberry. The results showed that PLSR and LS-SVM models coupled with color space could predict pH value in red bayberry (r = 0.93–0.96, RMSE = 0.09–0.12, MAE = 0.07–0.09, and MRE = 0.04–0.06). In addition, the minimum errors (RMSE = 0.09, MAE = 0.07, and MRE = 0.04) and maximum correlation coefficient value (r = 0.96) were found with the PLSR based on CMY, I1I2I3, and YCbCr color spaces. For predicting SSC, PLSR models based on CIELAB color space (r = 0.90, RMSE = 0.91, MAE = 0.69 and MRE = 0.12)

and HSI color space (r = 0.89, RMSE = 0.95, MAE = 0.73 and MRE = 0.13) were recommended. The results indicated that color space combined with chemometric is suitable to non-destructively detect pH value and SSC of red bayberry.

Keywords Red bayberry · pH value · Soluble solids content · PLSR · LS-SVM

Jie Feng, Lingling Jiang and Jialei Zhang have contributed equally to this work.

✉ Xiaohong Sun
  xhsun2000@163.com

✉ Hongfei Lu
  luhongfei0164@163.com

1 College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou 310018, China

2 Zhejiang Province Key Laboratory of Plant Secondary Metabolism and Regulation, Hangzhou 310018, China

3 School of Pharmaceutical Sciences, Wenzhou Medical University, Wenzhou 325035, China

4 School of Civil Engineering and Architecture, Zhejiang University of Science and Technology, Hangzhou 310023, China

5 Yuanpei College, Shaoxing University, Shaoxing 312000, China

Introduction

Bayberry (*Myrica rubra* Siebold and Zuccarini), which belongs to the genus *Myrica* in the family Myricaceae, is cultivated in Southeast China since more than 2000 years. It is an important economic Asian fruit crop that produces luscious fruit with appealing flavor (Perkins et al. 2017). The fresh fruit can be processed into a variety of forms, including jam, juice, wine, preserved and canned in syrup (Li et al. 2018). It is highly regarded by consumers, which is abundant with high nutritional components such as anthocyanins, carbohydrates, organic acids, flavonoids, and vitamins (Cheng et al. 2016). Soluble solids content (SSC) and pH, two of the internal quality indices, play an important role in determining fruit maturity and harvest time (Huang et al. 2018). During red bayberry ripening, SSC increased and the acidity decreased (Zhang et al. 2005). The SSC reflects sugar content such as glucose, fructose, and sucrose which is related to fruit taste (Jordan et al. 2000). The pH value of red bayberry, to some extent, is due to organic acids including citric acid, malic acid, oxalic acid and tartaric acid. It is not only related to fruit taste, but also affects the stability of anthocyanins which provides colors ranging from salmon-pink through red and from violet to nearly black of red bayberry. In addition, the

attractive colors and the sugar/acid ratios are likely to be the primary attributes contributing to consumer preference. Therefore, the measurements of SSC and pH of red bayberry are important in fruit quality issue.

In recent years, several methods have been developed for detecting the fruit SSC for apples and other fruits, such as infrared spectroscopy technology (Guo et al. 2015; Pu et al. 2016; Guo et al. 2016; Cen et al. 2006; Paz et al. 2008; Fan et al. 2009; Moghimi et al. 2010; Ma et al. 2018), hyperspectral scattering technology (Peng and Lu 2008; Mendoza et al. 2011), and electronic nose sensors (Xu et al. 2018). However, there is no information available on using color space to detect the SSC and pH of red bayberry. Color as an important component of food quality relevant to market acceptance, can be used to quantify the distribution of ingredients for quality evaluation using computer vision. Researches showed that color has been successfully tested for detecting the quality of various food products, such as bruises detection on red bayberry (Lu et al. 2011), color measurements and pattern recognition of paprika (Palacios-Morillo et al. 2016), color quantification of processed ginger (Zhou et al. 2016), evaluation of acrylamide contents in biscuit (Lu and Zheng 2012), prediction of anthocyanins, ascorbic acid, total phenols, flavonoids and antioxidant activity in red bayberry juice (Zheng et al. 2011), maturity evaluation of date (Zhang et al. 2014), and prediction of color and firmness in banana (Xie et al. 2018). In addition, the image processing software, the lowering cost of digital camera and computational technique have made color image processing more versatile and less expensive. With a digital camera, color images can be obtained and saved in three color sensors per pixel in which each sensor captures the intensity of the light in the red (R), green (G) or blue (B) spectrum. Other color spaces such as CMY (Cyan, Magenta and Yellow), HSI (Hue, Saturation and Intensity), I1I2I3 (I1, I2 and I3 in Ohta color space), YCbCr (luminance, blue and red) and CIELAB (L*a*b*) can be transformed by RGB color space. Generally speaking, different color spaces have different characteristics and we should select a suitable color space for a specific visual task. Thus, it is important to choose an appropriate color space for achieving a best result.

In this paper, different color spaces coupled with partial least square regression (PLSR) models and least square-support vector machine (LS-SVM) models were used for detecting the SSC and pH values of red bayberry. The aim of this research is (1) to evaluate the effectiveness of models on inspecting the SSC and pH values originated from red bayberry image analysis, (2) to compare the performance of PLSR and LS-SVM models combined with different color spaces, (3) to lay a foundation for evaluation of other nutrient contents (phenols, flavonoids,

anthocyanins) and choosing optimal red bayberry to fruits processing (bayberry juice, dried bayberry or canned bayberry).

Materials and methods

Fruit materials and image acquisition

Red bayberry (*M. rubra* Sieb. & Zucc. cv. 'Biqi') was hand-harvested from an orchard in Cixi (Ningbo, China), and transported by refrigeration at 5 °C for 3 h to the laboratory. After transporting to our laboratory, a total of 50 fruits were selected according to its uniform size, while the physically damaged or diseased ones were removed. Then, a Canon EOS 50D camera with a Canon EF-S 18–55 mm f/3.5–5.6 IS lens was used to acquire red bayberry images, and all image acquisitions were carried out at least in triplicate under the same condition. After acquisition of images, SSC measurement was made with a portable refractometer (Shanghai Tianlei Instrument Co. Ltd., Shanghai, China) with accuracy of 0.002 Brix, and pH value of bayberry samples was measured by a pH meter (Mettler-Toledo Delta 320) with accuracy of 0.01 pH unit. Both of these measurements were performed immediately after image acquisition.
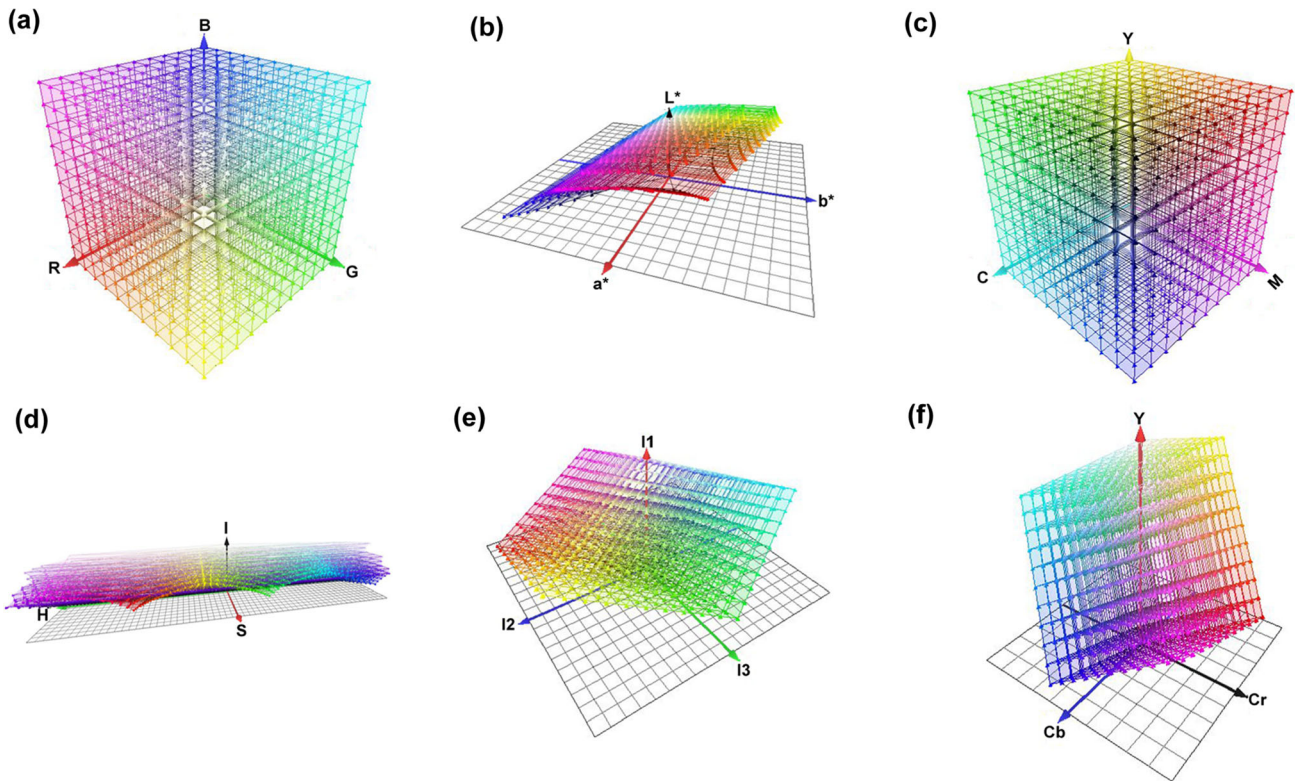
Color spaces

A color space is a method by which we can specify, create and visualize color (Ford and Roberts 1998). Generally, there are three types of color spaces, namely hardware-orientated space, human-orientated space, and instrumental space (Wu and Sun 2013). However, there is no any color space which is better than others and suitable to all kinds of images yet. To study the effect of different color spaces on the prediction performance of SSC and pH value of red bayberry, six different color spaces were evaluated, i.e., NRGB (normalized RGB), CIELAB, CMY, HSI, I1I2I3, and YCbCr. 3D demonstration of these color space images is illustrated in Fig. 1 (Wu and Sun 2013).

RGB corresponds to the three primary colors: red, green and blue, respectively. To reduce the dependence on lighting, the RGB color components are normalized by the following equation (García-Mateos et al. 2015):

$$\begin{cases} n_r = R/(R + G + B) \\ n_g = G/(R + G + B) \\ n_b = B/(R + G + B) \end{cases} \quad (1)$$

where n_r, n_g and n_b are the normalized values between 0 and 1, and the sum of these components is 1.

CIELAB is an international standard for color measurement developed by the Commission International

Fig. 1 3D demonstration of six different color space images. **a** RGB, **b** CIELAB, **c** CMY, **d** HSI, **e** I1I2I3, **f** YCbCr

Eclairage (CIE) in 1976. It consists of a lightness component ($L*$ value, from 0 to 100), along with two chromatic components ($a*$ value and $b*$ value, from − 120 to 120), in which $a*$ extends from green to red and $b*$ from blue to yellow as illustrated in Fig. 1b. The CIELAB values were obtained using the formulas as follows (Wu and Sun 2013):

$$\begin{cases} X = 0.607 \times R + 0.174 \times G + 0.201 \times B \\ Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \\ Z = 0.000 \times R + 0.066 \times G + 1.117 \times B \end{cases} \quad (2)$$

$$\begin{cases} L* = 116 \times \left(\dfrac{Y}{Y_0}\right)^{\frac{1}{3}} - 16 \\ a* = 500 \times \left[\left(\dfrac{X}{X_0}\right)^{\frac{1}{3}} - \left(\dfrac{Y}{Y_0}\right)^{\frac{1}{3}}\right] \\ b* = 200 \times \left[\left(\dfrac{Y}{Y_0}\right)^{\frac{1}{3}} - \left(\dfrac{Z}{Z_0}\right)^{\frac{1}{3}}\right] \end{cases} \quad (3)$$

where $X/X_0 > 0.01$, $Y/Y_0 > 0.01$ and $Z/Z_0 > 0.01$. (X_0, Y_0, Z_0) are X, Y, Z values for the standard white.

The CMY color model stands for Cyan, Magenta and Yellow which are the complements of Red, Green and Blue respectively as shown in Fig. 1c. The values can be achieved by the following equation (Dahat and Chavan 2016):

$$\begin{cases} C = 255 - R \\ M = 255 - G \\ Y = 255 - B \end{cases} \quad (4)$$

HSI color space (see Fig. 1d) is intuitive, which are motivated by the human vision system in a sense. It can be separated into three components, i.e. hue, saturation and intensity. They were calculated as Eq. (5) (Deb et al. 2009):

$$\begin{cases} H = cos^{(-1)}\left\{ \dfrac{\frac{1}{2}[(R-G)+(R-B)]}{\left[(R-G)^2 + (R-B)(G-B)\right]^{1/2}} \right\} \\ S = 1 - 3[min(R,G,B)]/(R+G+B) \\ I = (R+G+B)/3 \end{cases} \quad (5)$$

where H represents the dominant color of an area, S measures the colorfulness of an area in proportion to its brightness, and I is related to the color luminance.

As shown in Fig. 1e, I1I2I3 color space can be achieved by a linear transformation of RGB (García-Mateos et al. 2015):

$$\begin{cases} I1 = \dfrac{1}{3}(R + G + B) \\ \quad I2 = \dfrac{1}{2}(R - B) \\ I3 = \dfrac{1}{4}(2G - R - B) \end{cases} \quad (6)$$

And for YCbCr color space, it can be seen in Fig. 1f. that Cb represents the difference between the blue channel and a reference value and Cr is between red and a reference value. The values can be achieved by Eq. (7) (García-Mateos et al. 2015):

$$\begin{cases} Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \\ \quad Cb = 0.564(B - Y) + 128 \\ \quad Cr = 0.713(R - Y) + 128 \end{cases} \quad (7)$$

Partial least square regression (PLSR)

PLSR is a linear algorithm for modeling the relationship between two data sets (Wold et al., 2001). In this study, the color space values and the quality indicators (SSC and pH value) were used to form the explanatory matrix ($\mathbf{X}$) and dependent matrix ($\mathbf{Y}$), respectively. The development of PLSR prediction models involves two basic steps: training and test phases. In the training phase, 70% of the data were randomly selected to generate the model. In this phase, tenfold cross validation (10-CV) was used to choose the optimum number of PLSR components with the smallest prediction error, which avoids overfitting of the model. After training models, an independent data set (30% of the total data) was utilized to test the prediction performance. The PLSR was performed using the software MATLAB (R2010a, the MathWorks Inc., USA). Conventional analysis of linear regression was carried out using OriginLab (OriginPro, version 7.5).

Least square-support vector machine (LS-SVM)

The least-squares support vector machine (LS-SVM) is a variant of the standard SVM that can address linear and nonlinear multivariable calibration and solve multivariable calibration problems relatively quickly (Yu et al. 2016). And proper kernel function and optimal kernel parameters are the crucial elements for LS-SVM. In this study, radial basis function (RBF) kernel was used as the kernel function due to its effectiveness and speed in training process. A grid-search technique and leave one out (LOO) cross-validation was applied to find out the optimal parameter values including regularization parameter gam (γ) and RBF kernel function parameter sig2 (σ^2). All LS-SVM algorithms were implemented with MATLAB (R2010a, the MathWorks Inc., USA) and a LS-SVM toolbox for

MATLAB (LS-SVM v1.7, Suykens, Leuven, Belgium) under Windows XP.

Cluster analysis

Cluster analysis is one of the most useful statistical tools used in chemometrics for classifying a given population into groups (clusters), based on similarity or closeness measures. Hierarchical clustering, as a common used clustering algorithm, has different agglomerative cluster methods: Ward's minimum variance method, weighted pair-group method using centroids (WPGMC), the single linkage method, the complete linkage method, the average linkage method (UPGMA) and a maximum likelihood estimate algorithm (Berge et al. 2003). It calculates the distances between all samples using a defined metric such as Euclidean, Manhattan, Camberra and Pearson distances (Ragno et al. 2007). In this study, cluster analysis was displayed in order to find out similarities among the prediction models. UPGMA and the squared Euclidean distance were applied as the hierarchic agglomerative cluster algorithm and distance elaboration, respectively. The cluster analysis was conducted using PAST software (Version 2.17c) (Hammer et al. 2001).

Results and discussion

Prediction of SSC and PH value using PLSR models based on color space

In this study, the mean value and standard deviation of pH and SSC in red bayberry were 3.36 ± 0.20 and 12.57 Brix ± 1.42, respectively. The average values of six different color spaces (NRGB, CIELAB, CMY, HSI, I1I2I3, and YCbCr) were evaluated as the inputs for the prediction models. As seen in Tables 1 and 2, the number of components in the PLSR to predict pH and SSC is 3 for all of the color spaces. To test the performance of models, several parameters between experimental and predicted values were calculated (Tables 1 and 2): correlation coefficient (r), root mean squared error (RMSE), mean absolute error (MAE), mean relative error (MRE). For predicting SSC (in Table 1), the PLSR models based on CIELAB color space could get the highest r value (r = 0.90) and lowest errors (RMSE = 0.91 Brix, MAE = 0.69 Brix and MRE = 0.12). From Table 2, it can be observed that the r value for PLSR models based on NRGB, CIELAB, CMY, HSI, I1I2I3 and YCbCr color spaces to detect pH value is higher than 0.95, and the error is lower than 0.10. Figure 2 shows the correlation between the experimental and the model-predicted values. It indicated that the predictive capacity for detecting pH value was better than SSC.

Table 1 The results of partial least square regression (PLSR) and least square-support vector machine (LS-SVM) models of soluble solids content (SSC) using different color spaces

Models	Parameters	Color space					
		NRGB	CIELAB	CMY	HSI	I1I2I3	YCbCr
PLSR	Factors	3	3	3	3	3	3
	r^a	0.90	0.90	0.89	0.89	0.89	0.89
	RMSE[b]	0.91	0.91	0.97	0.95	0.97	0.97
	MAE[c]	0.77	0.69	0.81	0.73	0.81	0.81
	MRE[d]	0.14	0.12	0.14	0.13	0.14	0.14
LS-SVM	γ^e	70.77	27.21	2.6×10^4	1.9×10^5	2.1×10^3	529.27
	σ^{2f}	163.19	43.21	1.0×10^5	3.7×10^3	368.47	456.06
	r	0.91	0.73	0.82	0.91	0.90	0.88
	RMSE	1.06	1.36	1.22	1.10	1.07	1.14
	MAE	0.85	0.99	1.01	0.86	0.80	0.91
	MRE	0.15	0.18	0.18	0.15	0.14	0.16

[a]Correlation coefficient; [b]root mean squared error; [c]mean absolute error; [d]mean relative error; [e]regularization parameter gam; [f]parameter sig2

Table 2 The results of partial least square regression (PLSR) and least square-support vector machine (LS-SVM) models of pH using different color spaces

Models	Parameters	Color space					
		NRGB	CIELAB	CMY	HSI	I1I2I3	YCbCr
PLSR	Factors	3	3	3	3	3	3
	r^a	0.95	0.95	0.96	0.95	0.96	0.96
	RMSE[b]	0.10	0.10	0.09	0.10	0.09	0.09
	MAE[c]	0.08	0.08	0.07	0.08	0.07	0.07
	MRE[d]	0.05	0.06	0.04	0.05	0.04	0.04
LS-SVM	γ^e	70.77	27.21	2.6×10^4	1.9×10^5	2.1×10^3	529.27
	σ^{2f}	163.19	43.21	1.0×10^5	3.7×10^3	368.47	456.06
	r	0.94	0.93	0.95	0.96	0.96	0.96
	RMSE	0.11	0.11	0.10	0.12	0.11	0.10
	MAE	0.09	0.09	0.08	0.09	0.09	0.08
	MRE	0.06	0.06	0.06	0.06	0.06	0.06

[a]Correlation coefficient; [b]root mean squared error; [c]mean absolute error; [d]mean relative error; [e]regularization parameter gam; [f]parameter sig2

Prediction of SSC and pH value using LS-SVM models based on color space

LS-SVM as a variant of SVM was applied to building SSC and pH value prediction models in this study. In the training phase, 70% of the data were randomly selected to develop the models. To obtain a good prediction performance, the optimal parameters (γ and σ^2) were optimized by grid-search and LOO cross-validation. As seen in Tables 1 and 2, the selected optimal value of γ and σ^2 for pH and SSC prediction models are 70.77 and 163.19 for NRGB, 27.21 and 43.21 for CIELAB, 2.6×10^4 and 1.0×10^5 for CMY, 1.9×10^5 and 3.7×10^3 for HSI, 2.1×10^3 and 368.47 for I1I2I3, and 529.27 and 456.06 for YCbCr, respectively.

After training the LS-SVM models, prediction performance was tested using an independent data set. Examining the agreement between the experimental data and the prediction values are presented in Fig. 3, and the r, RMSE, MAE and MRE are given in Tables 1 and 2. According to Table 1, the LS-SVM models based on NRGB, CIELAB, CMY, HSI, I1I2I3 and YCbCr could predict SSC in red bayberry with the RMSE of 1.06, 1.36, 1.22, 1.10, 1.07, and 1.14 Brix, and r values of 0.91, 0.73, 0.82, 0.91, 0.90, and 0.88, respectively. For the prediction of pH, the r value is more than 0.9, and the range of RMSE, MAE and MRE are 0.10–0.12, 0.08–0.09, and 0.06, respectively.
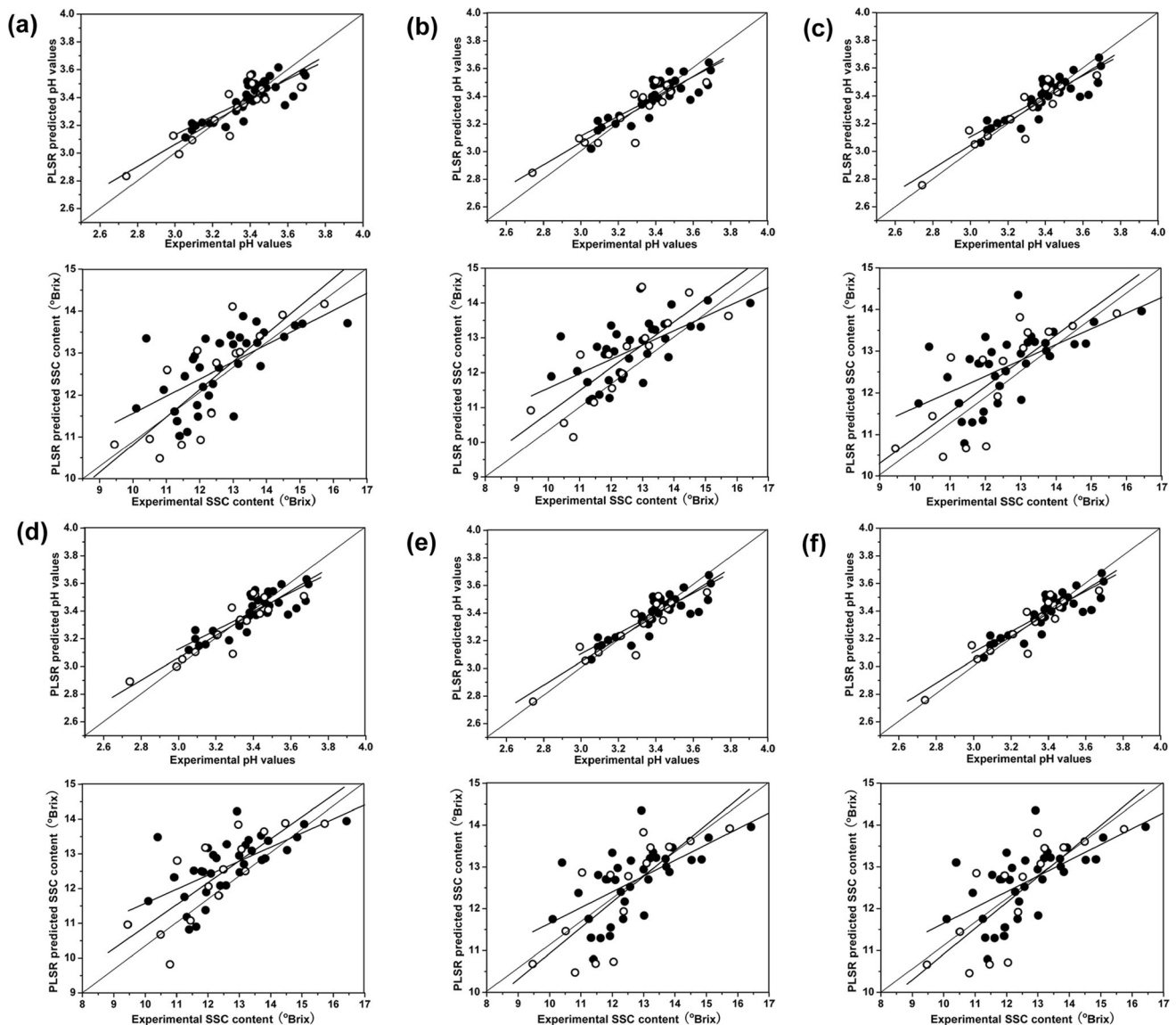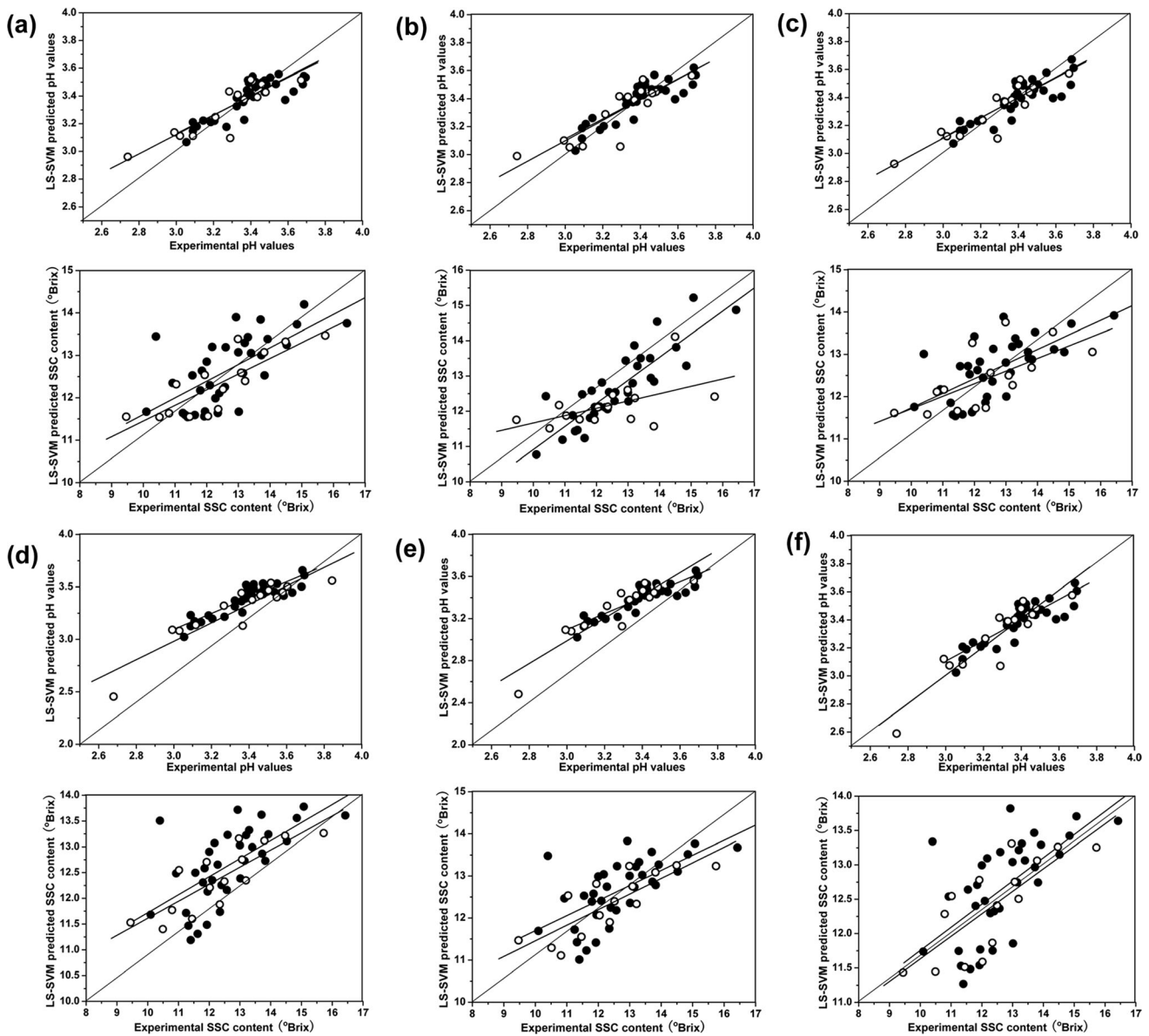
Fig. 2 Correlation of experimental values (○) and predicted values (●) for pH and SSC in red bayberry from PLSR models based on **a** RGB, **b** CIELAB, **c** CMY, **d** HSI, **e** I1I2I3, **f** YCbCr

Comparison of the results

Figure 4 shows the dendrogram representing the relationships between the models clustered by hierarchical cluster analysis on the basis of r, RMSE, MAE, MRE. Prediction models divided into the same cluster, meaning that the prediction ability of these models is approximately equal. According to Table 1 and Fig. 4a, the results showed that the optimal models for predicting SSC were PLSR models based on CIELAB (r = 0.90, RMSE = 0.91 Brix, MAE = 0.69 Brix, and MRE = 0.12) and HSI color spaces

(r = 0.89, RMSE = 0.95 Brix, MAE = 0.73 Brix, and MRE = 0.13). The results were better than Shao and He (2007) using Vis/NIR spectroscopy techniques for nondestructive measurement of SSC and pH of bayberry juice. The values of correlation coefficient of SSC obtained in this research was slightly better than those obtained by Peirs et al. (2001) with apple of 0.84 and those obtained by Lu (2001) with sweet cherries of 0.89, and was similar with the results obtained by Shao et al. (2007) with tomatoes of 0.90. Although the results in our study were worse than those obtained by Moghimi et al. (2010) with kiwifruit of

Fig. 3 Correlation of experimental values (○) and predicted values (●) for pH and SSC in red bayberry from LS-SVM models based on **a** RGB, **b** CIELAB, **c** CMY, **d** HSI, **e** I1I2I3, **f** YCbCr

0.93, they were still better than many other fruits above. So, the PLSR models based on CIELAB and HSI color space were good methods to predict SSC. For predicting pH (in Table 2 and Fig. 4b), the minimum errors (RMSE = 0.09, MAE = 0.07, and MRE = 0.04) and maximum r value (r = 0.96) were found with PLSR model based on CMY, I1I2I3, and YCbCr color spaces. However, it is worthy to notice that both PLSR and LS-SVM models based on all of color space could get a high r value (r = 0.93–0.96) and low errors (RMSE = 0.09–0.12, MAE = 0.07–0.09, and MRE = 0.04–0.06). In addition, the performance of these models to predict pH in red bayberry fruit was better than those obtained by Gómez

et al. (2006) using Satsuma mandarin fruit with r = 0.8 and RMSEP = 0.18; and by Moghimi et al. (2010) in kiwifruit with r = 0.943 and RMSEP = 0.076. Therefore, color space combined with chemometrics could be an appropriate method to predict SSC and pH value in fruit.

Conclusion

This paper proposes the application of PLSR and LS-SVM to develop a non-destructive technique to predict soluble solids content and pH value in red bayberry based on six different color spaces (NRGB, CIELAB, CMY, HSI,

Fig. 4 Dendrogram representing the relationships between the prediction models clustered by hierarchical cluster analysis on the basis of r, RMSE, MAE, and MRE, **a** SSC, **b** pH value
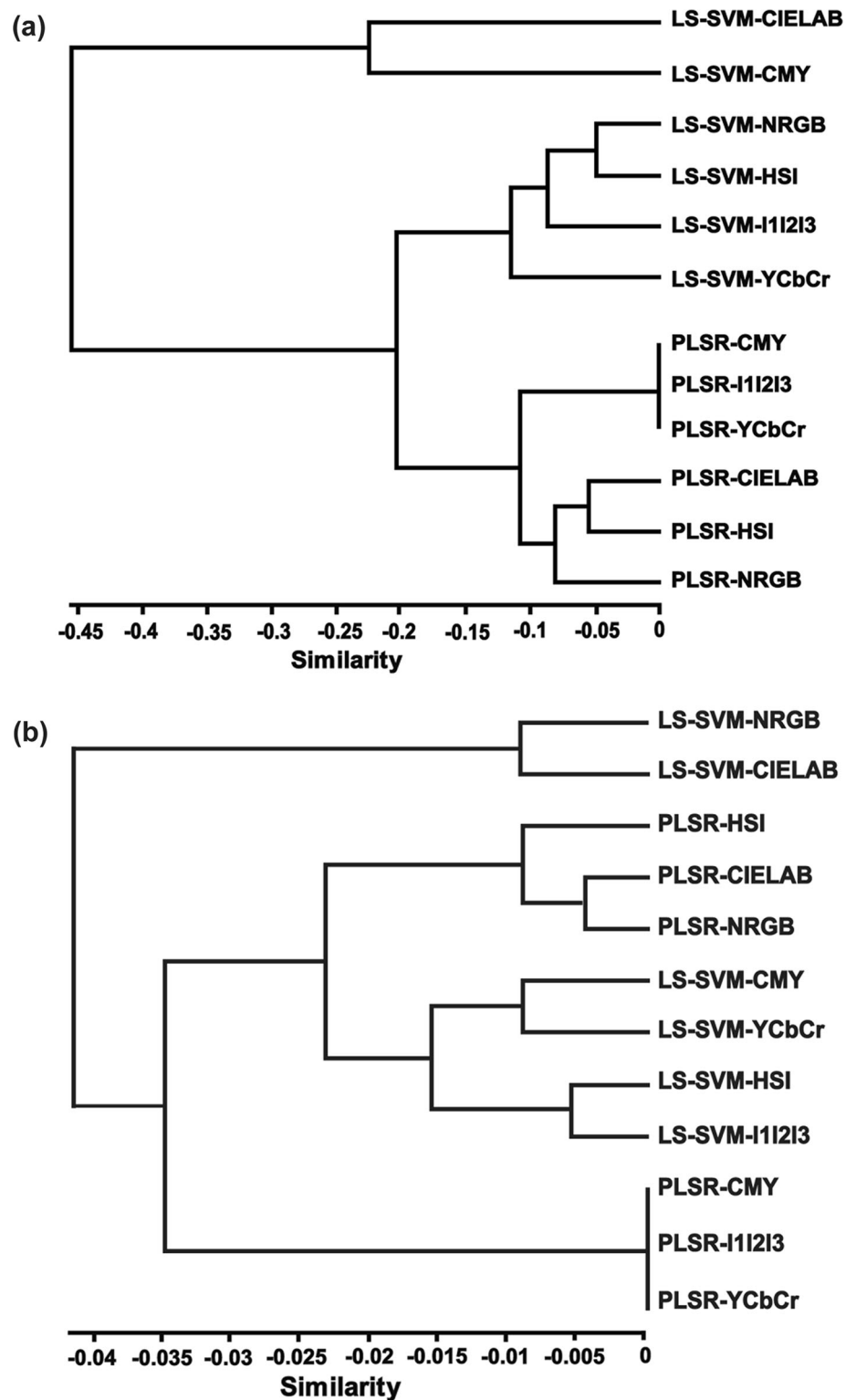
I1I2I3, and YCbCr). The results showed that PLSR and LS-SVM models combined with color space (r = 0.93–0.96, RMSE = 0.09–0.12, MAE = 0.07–0.09, and MRE = 0.04–0.06) as a potential tool can be used to predict pH value in red bayberry, and the PLSR models based on

CIELAB and HSI color space (r = 0.89–0.90, RMSE = 0.91–0.95, MAE = 0.73–0.77, and MRE = 0.13–0.14) is adequate for the prediction of SSC in this study. It indicates the possibility of developing a potentially non-destructive and cost-effective technique using color space and

chemometrics for facilitating quality detection of red bayberry. And we consider it available for detecting other fruits as well in further study.

References

Berge ACB, Atwill ER, Sischo WM (2003) Assessing antibiotic resistance in fecal *Escherichia coli* in young calves using cluster analysis techniques. Prev Vet Med 61(2):91–102

Cen H, He Y, Huang M (2006) Measurement of soluble solids contents and pH in orange juice using chemometrics and vis-NIRS. J Agric Food Chem 54(20):7437–7443

Cheng H, Chen J, Chen S, Xia Q, Liu D, Ye X (2016) Sensory evaluation, physicochemical properties and aroma-active profiles in a diverse collection of Chinese bayberry (*Myrica rubra*) cultivars. Food Chem 212(1):374–385

Dahat AV, Chavan PV (2016) Secret sharing based visual cryptography scheme using CMY color space. Procedia Comput Sci 78:563–570

Deb K, Kang SJ, Jo KH (2009) Statistical characteristic in HSI color model and position histogram based vehicle license plate detection. Intell Serv Robot 2:173–186

Fan G, Zha JW, Du R, Gao L (2009) Determination of soluble solids and firmness of apples by Vis/NIR transmittance. J Food Eng 93(4):416–420

Ford A, Roberts A (1998) Color space conversions (technical report). Westminster University, London

García-Mateos G, Hernández-Hernández JL, Escarabajal-Henarejos D, Jaén-Terrones S, Molina-Martínez JM (2015) Study and comparison of color models for automatic image analysis in irrigation management applications. Agric Water Manag 151:158–166

Gómez AH, He Y, Pereira AG (2006) Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. J Food Eng 77(2):313–319

Guo W, Shang L, Zhu X, Nelson SO (2015) Nondestructive detection of soluble solids content of apples from dielectric spectra with ANN and chemometric methods. Food Bioprocess Technol 8(5):1126–1138

Guo Y, Ni Y, Kokot S (2016) Evaluation of chemical components and properties of the jujube fruit using near infrared spectroscopy and chemometrics. Spectrochim Acta A 153:79–86

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontol Electron 4:1–9

Huang Y, Lu R, Chen K (2018) Assessment of tomato soluble solids content and pH by spatially-resolved and conventional Vis/NIR spectroscopy. J Food Eng 236:19–28

Jordan RB, Walton EF, Klages KU, Seelye RJ (2000) Postharvest fruit density as an indicator of dry matter and ripened soluble solids of kiwifruit. Postharvest Biol Technol 20(2):163–173

Li Y, Zhang L, Chen F, Lai S, Yang H (2018) Effects of vacuum impregnation with calcium ascorbate and disodium stannous citrate on Chinese red bayberry. Food Bioprocess Technol 11(7):1300–1316

Lu R (2001) Predicting firmness and sugar content of sweet cherries using near-infrared diffuse reflectance spectroscopy. Trans ASABE 44(5):1265–1271

Lu H, Zheng H (2012) Fractal colour: a new approach for evaluation of acrylamide contents in biscuits. Food Chem 134(4):2521–2525

Lu H, Zheng H, Hu Y, Lou H, Kong X (2011) Bruise detection on red bayberry (*Myrica rubra* Sieb. & Zucc.) using fractal analysis and support vector machine. J Food Eng 104(1):149–153

Ma T, Li X, Inagaki T, Yang H, Tsuchikawa S (2018) Noncontact evaluation of soluble solids content in apples by near-infrared hyperspectral imaging. J Food Eng 224:53–61

Mendoza F, Lu R, Ariana D, Cen H, Bailey B (2011) Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. Postharvest Biol Technol 62(2):149–160

Moghimi A, Aghkhani MH, Sazgarnia A, Sarmad M (2010) Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. Biosyst Eng 106(3):295–302

Palacios-Morillo A, Jurado JM, Alcázar A, Pablos F (2016) Differentiation of Spanish paprika from Protected Designation of Origin based on color measurements and pattern recognition. Food Control 62:243–249

Paz P, Sánchez MT, Pérez-Marín D, Guerrero JE, Garrido-Varo A (2008) Nondestructive determination of total soluble solid content and firmness in plums using near-infrared reflectance spectroscopy. J Agric Food Chem 56(8):2565–2570

Peirs A, Lammertyn J, Ooms K, Nicolaï BM (2001) Prediction of the optimal picking date of different apple cultivars by means of Vis/NIR-spectroscopy. Postharvest Biol Technol 21(2):189–199

Peng Y, Lu R (2008) Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content. Postharvest Biol Technol 48(1):52–62

Perkins ML, Yuan Y, Joyce DC (2017) Ultrasonic fog application of organic acids delays postharvest decay in red bayberry. Postharvest Biol Technol 133:41–47

Pu H, Liu D, Wang L, Sun DW (2016) Soluble solids content and pH prediction and maturity discrimination of lychee fruits using visible and near infrared hyperspectral imaging. Food Anal Method 9(1):235–244

Ragno G, Luca MD, Ioele G (2007) An application of cluster analysis and multivariate classification methods to spring water monitoring data. Microchem J 87(2):119–127

Shao Y, He Y (2007) Nondestructive measurement of the internal quality of bayberry juice using Vis/NIR spectroscopy. J Food Eng 79(3):1015–1019

Shao Y, He Y, Gómez AH, Pereir AG, Qiu Z, Zhang Y (2007) Visible/near infrared spectrometric technique for nondestructive assessment of tomato 'Heatwave' (*Lycopersicum esculentum*) quality characteristics. J Food Eng 81(4):672–678

Wold S, Trygg J, Berglund A, Antti H (2001) Some recent developments in PLS modeling. Chemom Intell Lab Syst 58(2):131–150

Wu D, Sun DW (2013) Colour measurements by computer vision for food quality control—a review. Trends Food Sci Technol 29(1):5–20

Xie C, Chu B, He Y (2018) Prediction of banana color and firmness using a novel wavelengths selection method of hyperspectral imaging. Food Chem 245:132–140

Xu S, Sun X, Lu H, Yang H, Ruan Q, Huang H, Chen M (2018) Detecting and monitoring the flavor of tomato (*Solanum lycopersicum*) under the impact of postharvest handlings by physicochemical parameters and electronic nose. Sensors 18(6):1847

Yu H, Chen Y, Hassan SG, Li D (2016) Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO. Comput Electron Agric 122:94–102

Zhang WS, Chen KS, Zhang B, Sun CD, Cai C, Zhou CH, Xu WP, Zhang WQ, Ferguson IB (2005) Postharvest responses of Chinese bayberry fruit. Postharvest Biol Technol 37(3):241–251

Zhang D, Lee DJ, Tippetts BJ, Lillywhite KD (2014) Date maturity and quality evaluation using color distribution analysis and back projection. J Food Eng 131:161–169

Zheng H, Jiang L, Lou H, Hu Y, Kong X, Lu H (2011) Application of artificial neural network (ANN) and partial least-squares regression (PLSR) to predict the changes of anthocyanins, ascorbic acid, total phenols, flavonoids, and antioxidant activity during storage of red bayberry juice based on fractal analysis and red, green, and blue (RGB) intensity values. J Agric Food Chem 59(2):592–600

Zhou SJ, Meng J, Huang ZP, Jiang SZ, Tu TQ (2016) A method for discrimination of processed ginger based on image color feature and a support vector machine model. Anal Methods 8:2201–2206