

Nondestructive Determination of Kiwifruit SSC using Visible/Near-Infrared Spectroscopy with Genetic Algorithm

Ma Shibang*

School of Mechanical and Electrical Engineering, Nanyang Normal University, Nanyang 473061, China

Received 12 December 2020; Accepted 22 February 2021

Abstract

Soluble solid content (SSC) is one of the most important internal quality attributes that affect the quality and price of fresh fruit. Spectral technology has been widely used to detect the attributes of agricultural products. However, the accuracy and stability of a prediction model are the main factors that affect its practical application. Visible (Vis)-near-infrared (NIR) spectroscopy combined with genetic algorithm (GA) was used to realize the rapid and nondestructive detection of kiwifruit SSC and improve the accuracy and stability of prediction model. The laboratory Vis-NIR spectroscopy system was built to collect the reflectance spectra of 120 kiwifruit samples at 400–2200 nm. Reflectance spectra were obtained after different pretreatments, namely, multiplicative scatter correction (MSC) and its combination with Savitzky-Golay smoothing (SG) method. The prediction models of multiple linear regression (MLR), partial least squares regression (PLSR), and least squares support vector machine (LS-SVM) were constructed for the prediction of kiwifruit SSC using the full-spectrum and effective wavelengths selected by GA. Results demonstrate that MSC+SG is the best pretreatment, the performances of the models established using the effective wavelengths selected by GA are better than those of the full-spectrum models, and the best performance is achieved by the GA-LS-SVM model, which is able to predict kiwifruit SSC with a correlation coefficient of prediction of 0.973 and a standard error of prediction of 0.362. The prediction accuracy is improved. This study demonstrates that the GA-LS-SVM model built using Vis-NIR spectroscopy with GA can rapidly and nondestructively determine kiwifruit SSC. The proposed algorithm provides evidence for the further development of a device for the nondestructive and rapid determination of kiwifruit SSC.

Keywords: nondestructive determination, genetic algorithm; kiwifruit, soluble solid content, visible/near-infrared spectroscopy

1. Introduction

Kiwifruit is one of the important fruits consumed all over the world for its healthy properties. Kiwifruit is widely consumed because it is rich in vitamins, sugar, minerals, and other nutrients. It is popular with consumers as the "fruit king" [1, 2]. Sweetness or soluble solid content (SSC, mostly sugars) is one of the most important internal quality indicators of fruit. SSC is also a key factor in determining fruit maturity and harvest time [3, 4, 5, 6, 7]. SSC affects the taste, price, and consumers' purchasing choices of fruit [8, 9, 10] and further affects the market competitiveness and economic value of fruits [11, 12]. The determination of fruit SSC has important reference value to evaluate the quality, harvest time, and grading sale of fruit. Traditional methods for the measurement fruit SSC require samples from internal tissues and are carried out using a digital refractometer. These methods are reliable, but destroy the sample and are time consuming. Sample destructive test can coarsely assess the quality of a whole batch of fruit. However, this method is not suitable for the rapid, nondestructive, and online real-time detection of fruit and does not meet the requirement of consumers and market. Spectrum-based detection methods are convenient, rapid, and nondestructive, have good measurement reproducibility, and have been widely used

during the past few decades in the detection of product quality in the agricultural field. Thus, many studies have detected fruit quality using spectral technology [13-18]. Liu et al. [19] used visible (Vis)-near-infrared (NIR) spectroscopy in the spectral range of 350–1800 nm to nondestructively measure the SSC and firmness of intact pear fruit by developing a partial least squares regression (PLSR) prediction model. The results showed that the prediction correlation coefficients (R) were 0.912 and 0.854 and the root mean standard errors of prediction (RMSEPs) were $0.662 \text{ } ^\circ\text{Bx}$ and 1.232 N for SSC and firmness, respectively. Thus, Vis-NIR spectrometry is an accurate, reliable, and nondestructive method for assessing the internal quality indices of intact pear fruit. However, the prediction model was built using the full spectrum and its prediction accuracy was low. Gao et al. [20] used NIR spectroscopy for the rapid nondestructive detection of the sugar content of red globe grape, and the results showed that the coefficient correlation of prediction set (R_p) was 0.980.

However, Gao et al. only built a PLSR model, and the stability of the model was uncertain. Marques et al. [21] used a new handheld ultra-compact NIR spectrometer to determine the SSC of 'Tommy Atkins' mango by building a PLSR calibration model, and the results showed that the coefficient of determination (R^2) and RMSEP values were 0.92 and $0.55 \text{ } ^\circ\text{Bx}$, respectively. However, only a PLSR prediction model was used, and the stability of the model

*E-mail address: mshibang@126.com

ISSN: 1791-2377 © 2021 School of Science, IHU. All rights reserved.

doi:10.25103/jestr.141.11

was uncertain. Oliveira-Folador et al. [22] used NIR and mid-infrared spectroscopy to evaluate the quality traits of fresh passion fruit pulp. The predictive result of the PLSR model for SSC was $R^2 = 0.95$. However, the raw spectrum for building the prediction model was not pretreated, and external interference information was not eliminated; therefore, the stability of the prediction model was low. Huang et al. [23] used a new noncontact multi-channel spectroscopic system for the Vis–NIR range of 550–1650 nm to nondestructively detect internal defect in apples. The results showed that the classification model of partial least squares discriminant analysis was significantly better classification with an overall accuracy of 91.5%. However, only the qualitative analysis was carried out, and a prediction model was not built. Tian Xi et al. [24] used an online spectrum measurement system in the wavelength range of 650–1000 nm by combining algorithm-selected effective wavelengths and different spectral preprocessing methods to determine SSC of intact apple. The results showed that the combination of mean normalization and 11 points smoothing was the best spectral preprocessing method, and the performance of the best model was $R = 0.9043$ and RMSEP=0.4787.

These aforementioned studies show that Vis–NIR spectroscopy can predict the SSC of different fruits, such as pear, apple, grape, and mango. These studies built prediction models based on the full spectrum. The influences of different pretreatment methods and effective wavelength variables on the prediction precision of a prediction model have not been studied. Studies on the rapid and nondestructive prediction of kiwifruit SSC by Vis–NIR spectroscopy combined with genetic algorithm (GA) are also lacking. The accuracy and stability of a spectral prediction model have great influence on its application and popularization. Therefore, the accuracy and stability of prediction model is necessary to improve.

In the present study, we attempted to develop a method to predictive kiwifruit SSC by Vis–NIR spectroscopy combined with GA; studied the effects of different spectral preprocessing methods, full-spectrum variables, effective wavelength variables, and different prediction models on the prediction accuracy, and determined the best spectral pretreatment method and best predictive model for kiwifruit SSC. This study provides basis for the further development of practical detection equipment. The remainder of this paper is organized as follows. Section 2 describes the materials and methods. Section 3 makes a discussion on the results. Section 4 summarizes the conclusions.

2. Materials and Methods

2.1 Vis–NIR spectroscopy detection system

A Vis–NIR spectroscopy detection system in the wavelength range of 400–2200 nm was established and used to acquire the spectrum of kiwifruit samples in this study. The sketch of the detection system is shown in Fig. 1. The Vis–NIR spectroscopy detection system consisted of a Vis–NIR spectrometer, a light source unit equipped with optical fibers, a computer installed with data acquisition software, a sample holder, and a shield case. The spectroscopy detection system was enclosed in the shield case to minimize effect of ambient light. The spectral resolution of the Vis–NIR

detection system was 0.8 nm in the visible band and 10 nm in the NIR band.

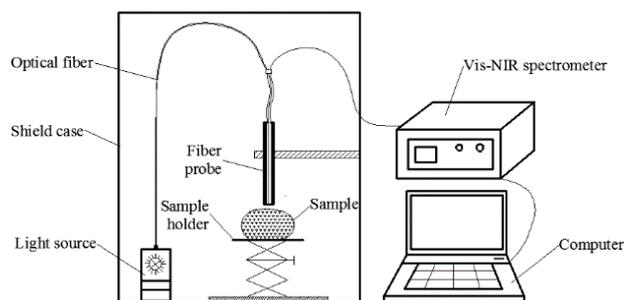


Fig. 1. Sketch of the detection system with visible and near-infrared spectroscopy.

2.2 Sample preparation

One hundred twenty intact kiwifruit samples were purchased from a local fruit market. The samples were transported to the laboratory and stored in refrigerator at 0°C . Before the experiment, the samples were randomly selected and placed in the laboratory under room temperature ($20 \pm 2^{\circ}\text{C}$) for more than 6 h to allow the kiwifruits' temperature to stabilize. Then, the spectral data and SSC values of each kiwifruit sample were collected using the Vis–NIR spectroscopy detection system and digital refractometer, respectively.

2.3 Vis–NIR reflectance spectral data collection

Fig. 2 provides a view of the test system. The Vis–NIR reflectance spectra of the samples were collected through the established Vis–NIR spectroscopy detection system. Kiwifruit was placed on the sample holder, and the distance between the fiber probe and kiwifruit surface was adjusted to 20 mm. Four different measurement positions were marked every 90° around each kiwifruit's equator. The mean of the four positions' spectra was used as the final spectral data of the tested sample for the establishment of the predictive model. This approach can decrease the influence of heterogeneous SSC distribution in the whole kiwifruit and improve the predictive accuracy of the model.

2.4 Measurement of real SSC value

After spectra collecting, The SSC reference values of kiwifruits were measured immediately after spectra collection using the traditional method. First, the skin of the marked position on the kiwifruit sample was peeled and the flesh was cut out. Second, the juice was squeezed using a fruit squeezer (Model: HY-150; Hangzhou, China). Then, SSC was measured using a digital refractometer (Model: PR-101; Atgo Co., Ltd., Tokyo, Japan) with a refractive index accuracy of ± 0.1 and a range of 0%–45% with temperature correction. The mean of the SSC reference values of the four marked positions was taken as the final reference value of a tested sample. Therefore, 120 SSC reference values were obtained.

MLR is one of the statistical methods [27] that attempt to model the relationship between two or more interpretive variables and a response variable by fitting a linear equation into the observed data. The model for MLR is:

$$y_i = a_0 + a_1x_{i,1} + a_2x_{i,2} + \dots + a_jx_{i,j} + e_i \quad (1)$$

where y_i is the dependent variable, a_0 is a constant, $x_{i,j}$ is an independent variable, a_j is the vector of regression coefficient, and e_i is a random measured error.

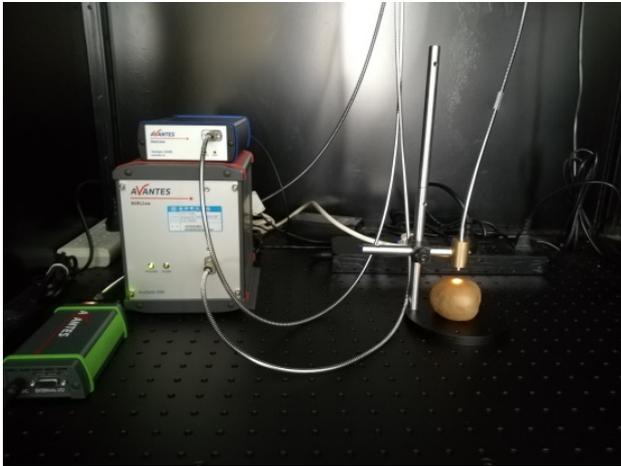


Fig. 2. The view of the test system.

PLSR is a widely used multivariate linear regression method in present spectroscopic analysis [28]. PLSR simultaneously considers the target chemical property matrix Y (the properties of interest) and the variable matrix X (the spectrum) and finds the fundamental relations between X and Y . In this study, PLSR analysis was performed to develop a regression model for the prediction of kiwifruit SSC. PLSR was also applied as a regression method to extract the latent variables (LVs). LVs were considered as new eigenvectors of the original spectra to reduce the dimensionality and compress the original spectral data. Certain selected LVs were employed as the inputs of PLSR. The number of LVs is a critical parameter in establishing a PLSR model. The optimal number of LVs is helpful to avoid overfitting or underfitting the model. Full cross-validation was used in the development of the PLSR model to determine the optimal number of LVs by root-mean-square error of cross-validation to prevent overfitting problems [29].

LS-SVM, which was proposed by Suykens and Vandewalle, is a simplified and improved model of support vector machine. LS-SVM is used in classification and regression problems. LS-SVM uses a set of linear equations for training and improves the convergence speed of the algorithm. Therefore, LS-SVM has been widely used in classification and regression [30]. The prediction belongs to the regression problem. For regression problems, the training set is supposed to be (x_i, y_i) , where $i = 1, 2, 3, \dots, n$; x_i is the input variable, and y_i is the output variable. The optimization problem and the constraint condition of LS-SVM can be described as follows:

$$\begin{cases} \min J(w, b, \xi) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^n \xi_i^2 \\ s.t. y_i [w^T \varphi(x_i) + b] = 1 - \xi_i, i = 1, \dots, n \end{cases} \quad (2)$$

where w is the weight vector; ξ_i is the slack variable for x_i ; γ is the regularization factor, which is used to adjust

the confidence interval of LS-SVM and the proportion of empirical risk; and b is a partial vector.

Proper kernel function and optimal LS-SVM parameters should be solved before the application of LS-SVM. Linear kernel, polynomial kernel, and radial basis function (RBF) are frequently obtained as the kernel function of LS-SVM. RBF could handle the nonlinear relationships between the spectra and target attributes, reduce the computational complexity of the training procedure, and perform well. Therefore, RBF was used as the kernel function of LS-SVM in this study. The regularization γ and RBF kernel function parameter σ^2 are two important parameters that determine the learning ability, prediction ability, and generalization ability of LS-SVM. In this study, K-fold cross-validation was applied to find the optimal values of γ and σ^2 [31].

2.5.2 Genetic algorithm (GA) method

GA is a numerical optimization method based on random search and is particularly suitable for dealing with complex nonlinear optimization problems and combination optimization problems. GA begins with a population initialized randomly over the search space of the optimization problem. GA generates improved approximate solutions or optimal solutions through an iterative procedure by simulating the Darwinian evolution principle of “survival of the fittest.” A detailed description of the algorithm can be referred to the literature [32]. GA has been successfully used as a feature selection technique, as it improves the predictive ability of prediction models and makes model predictions very easy. In this study, the GA-selected feature variables of the best-parameter spectra were used as input variables for the development of prediction models.

2.5.3 Model evaluation standard

The stability, reliability, and dynamic adaptability of the prediction models were used as the evaluation criteria of model performance. Four statistical indices, namely, the correlation coefficient of calibration set (R_c), the standard error of calibration (SEC), the correlation coefficient of prediction set (R_p), and the standard error of prediction (SEP), were calculated as shown in Equations (3)–(6) to evaluate the performance of the established models.

$$R_c = \frac{\sqrt{\sum_{i=1}^{n_c} (\hat{y}_i - y_i)^2}}{\sqrt{\sum_{i=1}^{n_c} (y_i - y_m)^2}} \quad (3)$$

$$R_p = \frac{\sqrt{\sum_{i=1}^{n_p} (\hat{y}_i - y_i)^2}}{\sqrt{\sum_{i=1}^{n_p} (y_i - y_m)^2}} \quad (4)$$

$$SEC = \sqrt{\frac{1}{n_c - 1} \sum_{i=1}^{n_c} (\hat{y}_i - y_m)^2} \quad (5)$$

$$SEP = \sqrt{\frac{1}{n_p - 1} \sum_{i=1}^{n_p} (\hat{y}_i - y_m)^2} \quad (6)$$

where, \hat{y}_i is the i th predicted value, y_i is the i th measured value, y_m is the mean of the calibration or prediction set, n_c is the number of samples in the calibration set, and n_p is the number of samples in the prediction set. Generally, models with higher R_c and R_p or with lower SEC and standard error of prediction (SEP) are more satisfactory than models with lower R_c and R_p or with higher SEC and SEP.

In this study, data analysis was performed in the Matlab R2010 software (The Mathworks, Inc., Natick, MA, USA).

3. Results and discussion

3.1 Spectral analysis of sample and SSC values

Mahalanobis distance outlier detection approach was used before establishing the models to detect the samples' spectral data and SSC values to identify the potential abnormal samples and eliminate the influence of unavoidable outliers in the training sample on the predictive ability and accuracy of the predictive models [33]. Six outlier samples were detected and removed from the 120 samples. The original spectra of the remaining 114 kiwifruit samples are shown in Fig. 3. These spectra had some crossovers and overlapping; however, the trends of the spectra were quite consistent. An obvious absorption peak was found around 1190 nm, which was assigned to the second overtone of the C–H band. The original spectra have evident scattering, noise, and baseline shift. Therefore, preprocessing methods (MSC and SG) were applied to the original spectra. The preprocessed spectra shown in Fig. 4 have no baseline shift and noise.

All samples were randomly divided into a calibration set and a prediction set at a ratio of 3:1. A calibration set with 86 samples and a prediction set with 28 samples were obtained. The statistical characteristics of the SSC of kiwifruits in the calibration and prediction sets are shown in Table 1. The maximum SSC in the calibration set was larger than that in the prediction set, and the minimum SSC in the prediction set was smaller than that in the calibration set. The standard deviation (SD) of the calibration and prediction sets was close; thus, the sample division was rational.

3.2 Prediction models established using full spectrum

After the samples were divided into the calibration and prediction sets, the full spectrum was pretreated by MSC or MSC+SG and then used as the input variable to build the prediction models. The MLR, PLSR and LS-SVM models for SSC based on the full spectrum were built. The results of

prediction are shown in Table 2. As shown in Table 2, the performances of the three models preprocessed with MSC+SG were better than those preprocessed with MSC only. Therefore, MSC+SG was selected as the preprocessing method for building the prediction models. The performance of the LS-SVM model was the best among the three models with $R_c = 0.898$, $SEC=0.567$, $R_p = 0.797$, and $SEP=0.815$, and the PLSR model was better than the MLR model. However, the prediction accuracies of the three models are not high. The main reason may be that the full spectrum contains a large amount of useless redundant information, which reduces the prediction ability and accuracy of the model. Thus, useless wavelength information should be removed and effective wavelength variables should be selected from the full spectrum to improve the prediction ability and accuracy of the model.

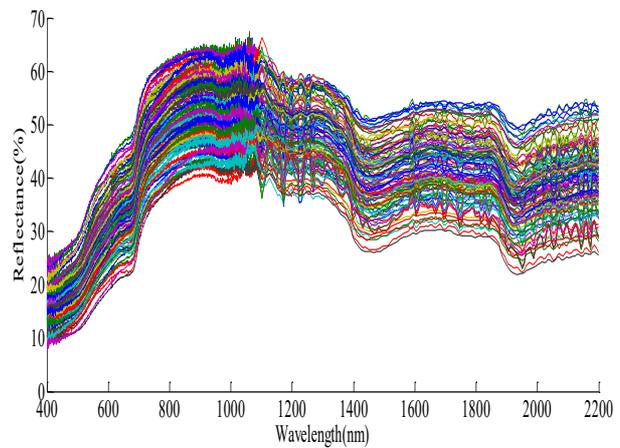


Fig. 3. Original spectra of all kiwifruit samples.

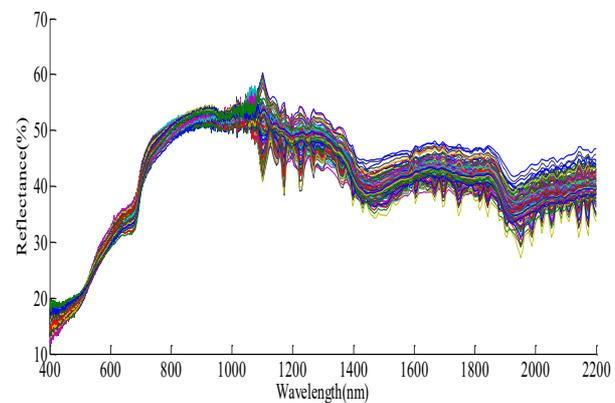


Fig. 4. Kiwifruit samples' spectra pretreated by MSC and SG methods.

Table 1. Statistical results of the SSCs of the calibration and prediction sets

Set	Number of samples	Maximum	Minimum	Mean	SD
Calibration set	86	13.6	8.2	10.4	1.14
Prediction set	28	13.4	7.8	10.3	1.33

Table 2. Prediction results of kiwifruit SSC based on different models and preprocessing methods

Pretreatment method	MLR				PLSR				LS-SVM			
	R_c	SEC	R_p	SEP	R_c	SEC	R_p	SEP	R_c	SEC	R_p	SEP
MSC	0.756	0.916	0.636	0.886	0.775	0.726	0.771	0.481	0.854	0.611	0.763	0.874
MSC+SG	0.786	0.710	0.734	0.909	0.805	0.680	0.762	0.487	0.898	0.567	0.797	0.815

3.3 Prediction models established using effective variables

The GA method was used to select the effective wavelength variables from the full spectrum, and the effective variables were used as input variables to establish the prediction models of MLR, PLSR, and LS-SVM to reduce useless information and improve the prediction precision. In the GA method, the population size of chromosome was set to 30; the probabilities of mutation and cross-over were set to 1% and 50%, respectively; and the number of run times was set to 100. The initial population of GA was generated randomly every time, and GA was randomly run 5 times to select the effective wavelength variables. Fig. 5 illustrates one of the five variable selection processes, where the selected parameter-combination spectra variables were above the dotted line. On the whole, the GA-selected variables were similar in each of the five selection processes; centered at the wavelengths of 760, 1360, 1440, 1640, and

2080 nm; and greatly reduced information redundancy and the number of variables required for developing the prediction model. Fig. 6 shows the distribution of the GA-selected effective variables at 400–2200 nm. The optimized variables were used to develop the prediction models of SSC (referred to as GA-MLR, GA-PLSR, and GA-LS-SVM). Table 3 shows the prediction results of the three different prediction models built using the effective wavelength variables selected by running GA five times. The results show that the accuracies of the models built using the effective wavelength variables are better than those of the models built with the full spectrum (Table 2). The prediction results of GA-LS-SVM were relatively satisfactory as indicated by $R_c = 0.987$, $SEC = 0.365$, $R_p = 0.973$, and $SEP = 0.362$ (Fig. 7). These results also showed that GA was able to extract effective wavelength variables and therefore greatly improved the models' prediction accuracy.

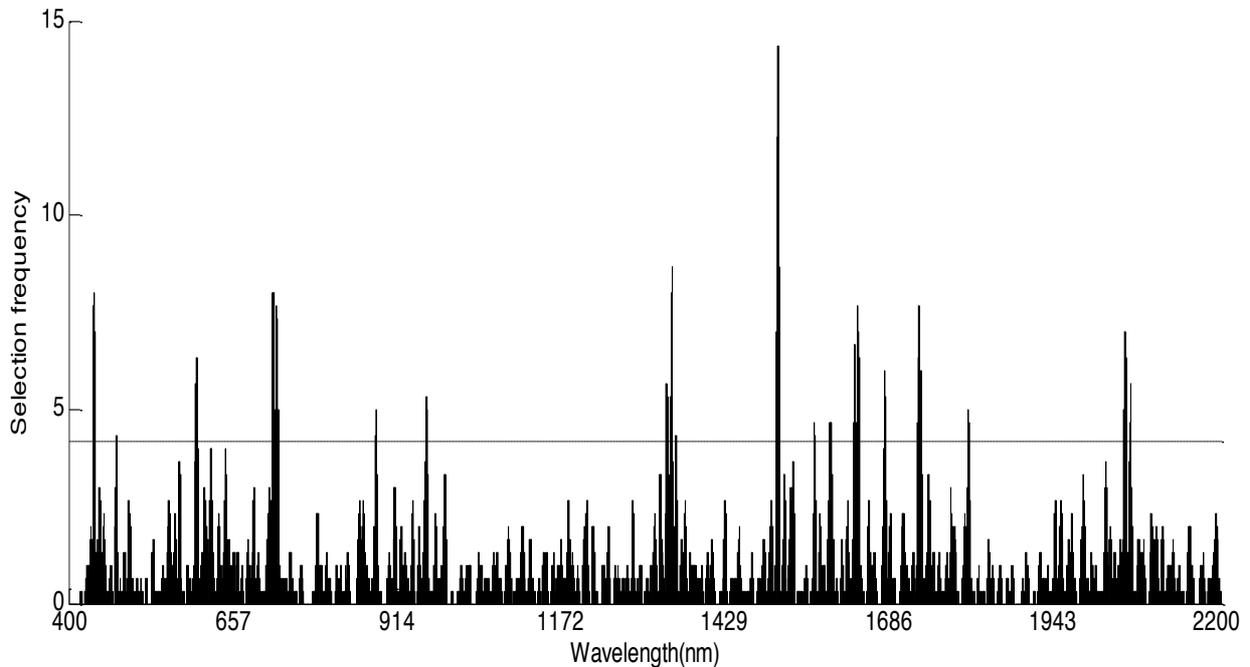


Fig. 5. Frequencies of variables selected by GA.

Table 3. Prediction results of GA-MLR, GA-PLSR, and GA-LS-SVM

GA Times	GA-MLR				GA-PLSR				GA-LS-SVM			
	R_c	SEC	R_p	SEP	R_c	SEC	R_p	SEP	R_c	SEC	R_p	SEP
1	0.926	0.576	0.905	0.515	0.946	0.453	0.928	0.492	0.962	0.374	0.950	0.373
2	0.941	0.505	0.906	0.550	0.930	0.569	0.920	0.568	0.945	0.406	0.937	0.478
3	0.905	0.585	0.896	0.544	0.969	0.364	0.919	0.463	0.961	0.391	0.930	0.450
4	0.928	0.528	0.911	0.513	0.927	0.596	0.916	0.495	0.951	0.455	0.937	0.414
5	0.931	0.517	0.930	0.510	0.950	0.469	0.903	0.519	0.987	0.365	0.973	0.362

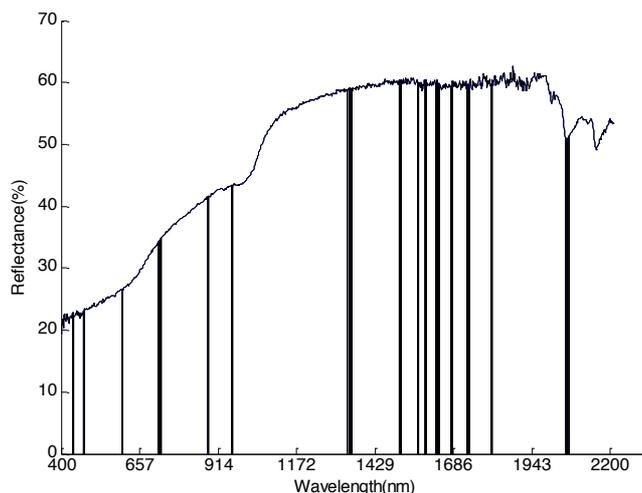


Fig. 6. Distribution of the GA-selected effective variables in the spectral region of 400–2200 nm

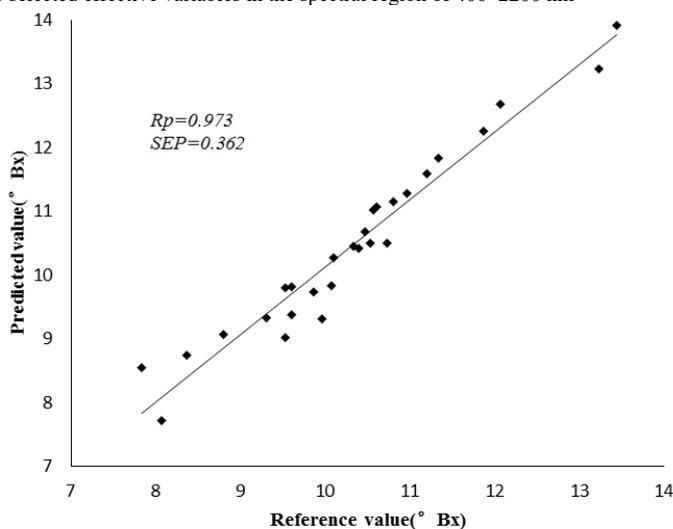


Fig. 7. Prediction and reference of SSC value with GA–LS-SVM model

4. Conclusion

Different preprocessing methods, GA-selected effective wavelength variables, and prediction models were comprehensively analyzed to study the performance of Vis–NIR spectroscopy combined with GA in the determination of kiwifruit SSC and to improve the accuracy and stability of prediction. The notable conclusions of this study are as follows:

- (1) Vis–NIR spectroscopy combined with GA can rapidly and nondestructively predict the SSC of kiwifruit.
- (2) The MSC+SG preprocessing method was better than MSC alone in developing prediction models.
- (3) The prediction accuracies of the MLR, PLSR, and LS-SVM models built with the effective wavelength variables selected by GA were superior those built with full-spectrum variables. The best prediction model for kiwifruit SSC was the GA–LS-SVM model, which showed improved prediction accuracy with $R_p = 0.973$ and $SEP=0.362$ °Bx.

This study indicates that Vis–NIR spectroscopy combined with GA may rapidly and accurately predict kiwifruit SSC without destroying kiwifruit. This method is very likely to be applied to the real-time detection of fruit quality and attributes in the future and provides a basis for the further development of practical nondestructive and rapid detection equipment for fruit. However, only one method of GA was used in this study. Thus, more algorithms should be studied in the future to further improve the accuracy and robustness of the prediction model.

Acknowledgement

This work was supported by the Science and Technology Research Project of Henan Province (Project No.192102110200).

This is an Open Access article distributed under the terms of the Creative Commons Attribution License.



References

- Lu Z. M., Wang X. L., Cao M. M., Li Y. Y., Su J. L., Gao H., "Effect of 24-epibrassinolide on sugar metabolism and delaying postharvest senescence of kiwifruit during ambient storage". *Scientia Horticulturae*, 253, 2019, pp. 1-7.
- Hu H. L., Zhou H. SH., Li P. X., "Lacquer wax coating improves the sensory and quality attributes of kiwifruit during ambient storage". *Scientia Horticulturae*, 244, 2019, pp. 31-41.
- Umer M. J., Lei G., Haileslassie G., Luqman B. S., Pingli Y., Zhao SH.J., Lu X. Q., He N., Zhu H. J., Liu W. G., "Expression pattern of sugars and organic acids regulatory genes during watermelon fruit development". *Scientia Horticulturae*, 265(109102), 2020, pp. 1-13.
- Li X. B., Li CH. N., Sun J., Aaron J., "Dynamic changes of enzymes involved in sugar and organic acid level modification during blueberry fruit maturation". *Food Chemistry*, 309(125617), 2020, pp. 1-14.
- Burdon J., Lallu N., Pidakala P., Barnett A., "Soluble solids accumulation and postharvest performance of 'Hayward' kiwifruit". *Postharvest Biology and Technology*, 80, 2013, pp. 1-8.
- Shimeles T., Han R. C., Do S. P., Yeon M. L., Jong H. C., Min W. B., Kwon H., Sung M. P., Cheon S. J., "Ripening quality of kiwifruit cultivars is affected by harvest time". *Scientia Horticulturae*, 261(108936), 2020, pp. 1-9.
- Rahat U., Saranjam K., Muhammad B., Farwa N., Muhammad S., "Non-invasive assessment of mango ripening using fluorescence spectroscopy". *Optik*, 127, 2016, pp. 5186-5189.
- Rajkumar P., Wang N., Elmasry G., Raghavan G. S. V., Garipey Y., "Studies on banana fruit quality and maturity stages using hyperspectral imaging". *Journal of Food Engineering*, 108, 2012, pp. 194-200.
- Carlos H. C., Gayle M. C., Paul M., "Consumer acceptance of 'Brooks' and 'Bing' cherries is mainly dependent on fruit SSC and visual skin color". *Postharvest Biology and Technology*, 28, 2003, pp. 159-167.
- You J. L., Seok H. E., "Kiwifruit cultivar 'Halla gold' functional component changes during preharvest fruit maturation and postharvest storage". *Scientia Horticulturae*, 234, 2018, pp. 134-139.
- Jiang B., He J. R., Yang SH. Q., Fu H. F., Li T., Song H. B., He D. J., "Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues". *Artificial Intelligence in Agriculture*, 1, 2019, pp. 1-8.
- Seong B. O., Sowbiya M., Yong-Bum K., Mi H. S., Jin G. K., "Characteristic of fruit development for optimal harvest date and postharvest storability in 'Skinny Green' baby kiwifruit". *Scientia Horticulturae*, 222, 2017, pp. 57-61.
- Sirinnapa S., Jinda S., Sumio K., "Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy". *Postharvest Biology and Technology*, 31, 2004, pp. 137-145.
- Liu Y. D., Sun X. D., Zhang H. L., Ouyang A. G., "Nondestructive measurement of internal quality of Nanfeng mandarin fruit by charge coupled device near infrared spectroscopy". *Computers and Electronics in Agriculture*, 71S, 2010, pp. S10-S14.
- Bertone E., Venturello A., Leardi R., Geobaldo F., "Prediction of the optimum harvest time of 'Scarlet' apples using DR-UV-Vis and NIR spectroscopy". *Postharvest Biology and Technology*, 69, 2012, pp. 15-23.
- Gabrieli A. O., Sylvie B., Catherine M. C. R., Adaucto B. P., Fernanda C., "Comparison of NIRS approach for prediction of internal quality traits in three fruit species". *Food Chemistry*, 143, 2014, pp. 223-230.
- Parichat T., Kumpon W., Nadthawat M., Sakunna W., Chanida K., Viboon C., Danai B., "Non-destructive identification and estimation of granulation in 'Sai Num Pung' tangerine fruit using near infrared spectroscopy and chemometrics Sila Kittiwachana". *Postharvest Biology and Technology*, 153, 2019, pp.13-20.
- Cortés V., Blasco J., Aleixos N., Cubero S., Talens P., "Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review". *Trends in Food Science & Technology*, 85, 2019, pp. 138-148.
- Liu Y. D., Chen X. M., Ouyang A. G., "Nondestructive determination of pear internal quality indices by visible and near-infrared spectrometry". *LWT-Food Science and Technology*, 41, 2008, pp. 1720-1725.
- Gao SH., Wang Q. H., Li Q. X., Shi H., "Non-destructive Detection of Vitamin C, Sugar Content and Total Acidity of Red Globe Grape Based on Near-Infrared Spectroscopy". *Chinese Journal of Analytical Chemistry*, 47(6), 2019, pp. 941-949.
- Emanuel J. N. M., Sérgio T. F., Maria F. P., Celio P., "Rapid and non-destructive determination of quality parameters in the 'Tommy Atkins' mango using a novel handheld near infrared spectrometer". *Food Chemistry*, 197, 2016, pp. 1207-1214.
- Gabrieli O., Milene O. B., Eriel F. A., Catherine M. C. R., Sylvie Bm, Fernanda C., "Quality traits prediction of the passion fruit pulp using NIR and MIR spectroscopy". *LWT-Food Science and Technology*, 95, 2018, pp. 172-178.
- Huang Y. P., Lu R.F., Chen K. J., "Detection of internal defect of apples by a multichannel Vis/NIR spectroscopic system". *Postharvest Biology and Technology*, 161(111065), 2020, pp.1-8.
- Tian X., Fan SH. X., Li J. B., Xia Y., Huang W. Q., "Comparison and optimization of models for SSC on-line determination of intact apple using efficient spectrum optimization and variable selection algorithm". *Infrared Physics and Technology*, 102(102979), 2019, pp. 1-11.
- Zhao CH. J., "Kiwifruit flesh firmness determination by a NIR sensitive device and image multivariate data analyses". *Sensors and Actuators A : Physical*, 296, 2019, pp. 265-271.
- Tom F., Cecilia R., Ana G., José E. G., "On the geometry of SNV and MSC". *Chemometrics and Intelligent Laboratory Systems*, 96, 2009, pp. 22-26.
- Chen H. ZH., Pan T., Chen J. M., Lu Q. P., "Waveband selection for NIR spectroscopy analysis of soil organic matter based on SG smoothing and MWPLS methods". *Chemometrics and Intelligent Laboratory Systems*, 107, 2011, pp. 139-146.
- Hamid G., Golasa M., Hakimeh F., "Spectroscopic studies on Solvatochromism of mixed-chelate copper (II) complexes using MLR technique". *Spectrochimica Acta Part A*, 85, 2012, pp. 25-30.
- Li J. B., Zhang H. L., Zhan B. SH., Wang ZH. L., Jiang Y. L., "Determination of SSC in pears by establishing the multi-cultivar models based on visible-NIR spectroscopy". *Infrared Physics and Technology*, 102(103066), 2019, pp. 1-10.
- Xu J. C., Ren Q. W., Shen ZH. ZH., "Prediction of the strength of concrete radiation shielding based on LS-SVM". *Annals of Nuclear Energy*, 85, 2015, pp. 296-300.
- Leardi R., Amparo L. G., "Genetic algorithms applied to feature selection in PLS regression: how and when to use them". *Chemometrics and Intelligent Laboratory Systems*, 41, 1998, pp. 195-207.
- Christophe L., Olivier K., Yves D., "Christophe L. Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance". *Journal of Experimental Social Psychology*, 74, 2018, pp.150-156.