Food Chemistry: X 17 (2023) 100539

Contents lists available at ScienceDirect

Food Chemistry: X

journal homepage: www.sciencedirect.com/journal/food-chemistry-x

Identification of aged-rice adulteration based on near-infrared spectroscopy combined with partial least squares regression and characteristic wavelength variables

Zhanming Li [a], Jiahui Song [a], Yinxing Ma [a], Yue Yu [a,*], Xueming He [b], Yuanxin Guo [a], Jinxin Dou [c], Hao Dong [d,*]

[a] School of Grain Science and Technology, Jiangsu University of Science and Technology, Zhenjiang 212100, China
[b] College of Food Science and Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China
[c] Academy of National Food and Strategic Reserves Administration, Beijing 100037, China
[d] College of Light Industry and Food Sciences, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China

A R T I C L E I N F O

Keywords:
Storage
Late japonica rice
Acid value
Chemometrics
Competitive adaptive reweighting sampling

A B S T R A C T

The long-term storage of rice will inevitably be involved in the deterioration of edible quality, and aged rice poses a great threat to food safety and human health. The acid value can be employed as a sensitive index for the determination of rice quality and freshness. In this study, near-infrared spectra of three kinds of rice (Chinese Daohuaxiang, southern japonica rice, and late japonica rice) mixed with different proportions of aged rice were collected. The partial least squares regression (PLSR) model with different preprocessing was constructed to identify the aged rice adulteration. Meanwhile, a competitive adaptive reweighted sampling (CARS) algorithm was used to extract the optimization model of characteristic variables. The constructed CARS-PLSR model method could not only reduce greatly the number of characteristic variables required by the spectrum but also improve the identification accuracy of three kinds of aged-rice adulteration. As above, this study proposed a rapid, simple, and accurate detection method for aged-rice adulteration, providing new clues and alternatives for the quality control of commercial rice.

1. Introduction

The long-term storage of rice will inevitably be involved in the deterioration of edible quality, and aged rice poses a great threat to food safety and human health. The continuous decrease in the edible quality of aged rice, which cannot be accepted by consumers, eventually leads to the utilization of aged rice in the industry, resulting in huge economic losses (Saikrishna, Dutta, Subramanian, Moses, & Anandharamakrishnan, 2018; Śliwińska-Bartel, Burns, & Elliott, 2021). The mixing of fresh and aged rice in the market has attracted great attention. Several traders reprocess the low-price aged rice by adding pigments, paraffin, etc. to improve its appearance and sell it to the market for huge profits (Wang et al., 2022; Wu, Li, Bai, Yu, & Zhang, 2019). Although the system of food safety standards has been continuously improved, such incidents have not been well controlled. Meanwhile, the current rice-quality testing methods are dissatisfactory because of their cumbersome and time-consuming properties (Jain, Mishra, Mishra, & Thakare, 2020; Xu, Liu, & Zhang, 2021). Thus, a simple and flexible detection method must be developed to meet the requirement of the identification of rice adulteration.

Changes in free fatty acid content always occur before physical changes during the process of rice aging (Tran et al., 2005). As a parameter affected by storage conditions, the acid value can be employed as a sensitive index for the determination of rice quality and freshness (Zhai et al., 2022).

Near-infrared spectroscopy (NIRS) presents unique advantages, such as high efficiency, convenience, and environmental protection, and it has been widely used in the identification of adulteration and the origin of agricultural and food products (Zhou, Yu, Hu, & Li, 2020). NIRS has been used to predict the chemical composition of sausages, and the results showed that NIRS can be used to estimate accurately the fat content of traditional country-style sausages, moisture, and protein content (Eleni, Vasileios, & Paraskevi, 2021). The commonly used partial least squares regression (PLSR) model shows promise for data structure

* Corresponding authors.
E-mail addresses: yuyueoffice2012@126.com (Y. Yu), donghao@zhku.edu.cn (H. Dong).

https://doi.org/10.1016/j.fochx.2022.100539
Received 24 June 2022; Received in revised form 10 November 2022; Accepted 3 December 2022
Available online 8 December 2022

simplification by combining modular form and epistemic methods (Huerta, Leiva, Liu, Rodríguez, & Villegas, 2019). Screening the representative characteristic variables from the NIR spectra of the measured samples can effectively simplify the model and improve the transferability of the model (Ren, Ning, & Zhang, 2021; Su et al., 2021).

Importantly, different preprocessing methods for spectral data are greatly involved in the accuracy of analytical results (Rady, Fischer, Reeves, Logan, & James Watson, 2019). Spectral preprocessing is the application of various mathematical means to data processing of spectra to eliminate other sources of spectral variation, such as the baseline drift during experiments, thus minimizing the effect of spectra and their variation on the accuracy of NIRS analysis models (Kademi, Ulusoy, & Hecer, 2019; Pasquini, 2018). In addition, by treating each wavelength variable in the spectral data as an individual, the feature variables with large absolute values of regression coefficients in the PLSR model can be selected by using competitive adaptive reweighted sampling (CARS) and extracted to optimize the model, reduce redundant wavelength variables, and improve the discriminative accuracy (Li, Huang, Song, Zhang, & Min, 2019). Standard normal variables (SNVs), which are corrected PLSR and cross-validation, were utilized to treat the collected spectral data (400–2500 nm) of 66 meat samples, and the results showed that NIRS was helpful in the successful prediction of the fat and protein contents of beef (Maduro Dias et al., 2021).

By contrast, considerable effort should be exerted to support the identification of rice adulteration with aged rice, ensuring food safety and human health. Herein, the acid value content of rice was presented as an index to determine the degree of rice aging. NIRS technology was utilized to collect the spectral data of aged-rice adulteration in three kinds of rice, including Chinese Daohuaxiang, southern japonica rice, and late japonica rice. The PLSR model combined with CARS algorithm was used in this study to simplify the model and improve the accuracy.

2. Materials and methods

2.1. Materials

The rice samples used in this experiment, including 20 kg Chinese Daohuaxiang, southern japonica rice, and late japonica rice, were all fresh and provided by the National Grain Reserve (Wuxi, Jiangsu Province, China). The hulled and polished rice (4 kg) was used for aging at the temperature of 37 ℃ and humidity of 85 %. The fresh polished rice was stored at −40 ℃ for the subsequent experiments.

2.2. Rice aging

The acid value of rice was determined accordingly (China National Standard, 2006; Aricetti & Tubino, 2012). The samples were processed by a rice milling machine to produce milled rice, and the milled rice was crushed by a pulverizer and passed through an 80-mesh sieve. A total of 10 g rice flour was obtained from the sieve, placed in a 250 mL cone bottle with a plug grinding mouth, and accurately added with 50 mL absolute ethanol with a pipetting gun. After shaking for 10 min and allowing the mixture to stand for 1–2 min, 25.0 mL filtrate was transferred, and 50 mL distilled water was added. Phenolphthalein–ethanol was used as an indicator and titrated with 0.01 mol/L potassium hydroxide in 95 % ethanol standard titration solution. Finally, the acid value was calculated (1) and presented using the average of parallel experiments.

$$\text{Acid value (mg KOH/100g dry value)} = (V1 - V0) \times C \times 56.1 \times (50/25) \times [100/(100 - w) \times m] \times 100 \quad (1)$$

In the formula:

V1: Volume of potassium hydroxide–95 % ethanol solution consumed by titrating the sample (mL);

V0: Volume of potassium hydroxide–95 % ethanol solution consumed for the titration blank (mL);

C: Accurate concentration of potassium hydroxide–95 % ethanol solution (mol/L);

50: Absolute ethanol for sample extraction (mL).

100: Mass converted to 100 g (dry) sample (g).

m: Mass of sample (g);

w: Percentage of moisture sample, mass of moisture in every 100 g sample (g).

2.3. Samples preparation

The acid value of aged rice was measured every-seven days, and aging ended when the acid value reached 25 mg KOH/100 g dry value. The aged rice was mixed with the same type of fresh rice at several ratios (0 %, 5 %, 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, and 100 %). After crushing the mixed rice samples into powder, samples of rice flour with different adulterated ratios were obtained using an 80-mesh sieve. Twenty samples were set for each mixing ratio, and 260 test samples mixed with aged rice were prepared for each rice sample. One hundred and ninety-five samples were used for calibration set and 65 samples were proposed as the prediction set (Firmani, De Luca, Bucci, Marini, & Biancolillo, 2019; Liu, Li, Peng, Ma, & Yan, 2021; Mishra & Nikzad-Langerodi, 2021).

2.4. NIRS data acquisition

All spectra were collected at room temperature (25 ± 2 °C) using a multiplex spectrometer (FX2000, Ideaoptics, Shanghai, China). The NIR diffuse reflectance technology was selected to test the samples. Morpho software was used to collect the sample spectral data. The acquisition parameters were as follows: spectral resolution of 8 nm, spectral scanning wavelength in the range of 900–1700 nm, average spectral scanning times of 32, and integration time of 20 ms. The rice samples were added into the cuvette, and each rice sample was scanned thrice from three different angles (0°, 120°, and 240°). Then, the average value of the three measurements was used as the input data of the model to improve the accuracy of the spectral data. Fig. S1 shows the schematic of the NIRS data-collecting systems for rice samples with a self-made optical cuvette.

2.5. Data preprocessing

When using the model to analyze the predicted samples, the same data preprocessing method in the model correction stage must be selected. Combined with the indicators used, five preprocessing methods, including min–max normalization (MMN), multiple scattering correction (MSC), smoothing (Five points), SNV, first derivative, and second derivative, were selected to preprocess the collected spectra separately or superimposed. In order to avoid the increase of noise after derivative processing, data is smoothed (Five points) before the treatment of first and second derivatives. Matlab program (R2021a) was used to analyze the collected NIR spectra data set of three kinds of rice with different ratios of aged rice.

2.6. CARS for key variable selection

To improve the accuracy of the model, by taking each wavelength variable in the NIRS as an independent individual, we used CARS to select the characteristic variables with large absolute values of the regression coefficient in the PLS model. CARS method is a feature variable selection method combining Monte Carlo sampling and PLS model regression coefficient. In the CARS algorithm, the points with higher absolute value weight of regression coefficient in the PLS model are retained as a new subset through adaptive weighted sampling each time, and the points with lower weight are removed. Then, the PLS model is established based on the new subset. After multiple calculations, the

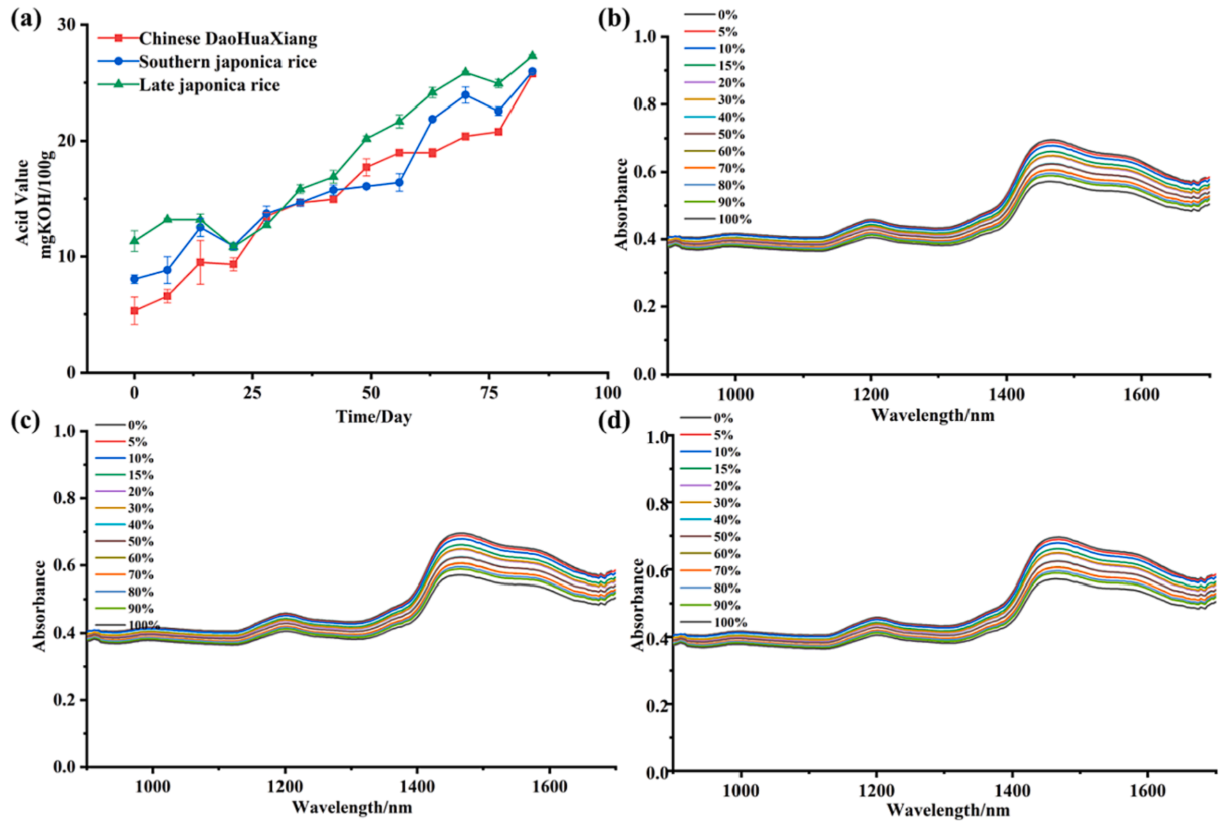

Fig. 1. Variation trend of acid values of three rice varieties(a). NIR spectra of (b) Chinese Daohuaxiang, (c) southern japonica rice, and (d) late japonica rice with different ratios of aged rice.

wavelength in the subset with the minimum root mean square error of cross-validation (RMSECV) of PLS model interactive verification is selected as the characteristic wavelength (Zheng et al., 2012).

2.7. PLSR model establishment

As a commonly used quantitative analysis method, PLSR realizes regression modeling and data structure simplification by combining model-based and epistemic methods, which is an organic combination of principal component analysis and multivariate linear analysis. It uses the property matrix Y and spectral matrix X to find a linear regression model in the new space projection, to find the multi-dimensional vector in the X matrix that can explain the Y matrix accurately, and obtain the basic relationship of the above two matrix data. PLSR can address the multicollinearity of variables in chemometrics and is applicable to the problem that the number of variables is greater than the sample size (the property matrix Y is the mixing proportion).

2.8. Model evaluation index

The correlation coefficient for cross-validation (R^2c), root mean - standard error of the calibration set (RMSEC), the correlation coefficient for prediction (R^2p), and root mean square error for prediction set (RMSEP) were used to evaluate the model. R^2c and R^2p represented the correlation between the predicted and real values. RMSEC was used to evaluate the accuracy of the prediction results of the calibrated set samples. RMSEP was used to measure the accuracy of the prediction results of the prediction set samples (Khan & Byun, 2020).

3. Results and discussion

3.1. Acid value

Fig. 1 shows the changes in the acid value of three kinds of rice under aging conditions. The slow change in acid value in the early stage of aging might be involved in the fat oxidation, which was observed in the induction period, and the increase rate was significantly accelerated in the stage of rapid oxidation and the increased speed is faster under the hydrolysis. During the whole aging process, the acid value increased, and at the 13th week of aging, the acid value of all three kinds of rice exceeded the target value (25 mg KOH/100 g dry value). The aged rice was prepared and used for subsequent experiments.

3.2. Spectral data analysis

Spectra data of three kinds of rice with different aged rice ratios were collected. As shown in Fig. 1, several absorption peaks were observed in the range of 900–1700 nm in the NIR spectra. The positions of absorption peaks were the same with several deviations, but the intensities were different. Such a result was observed because in the aging process, rice lipids were oxidized and hydrolyzed in the storage process. Oils were hydrolyzed by lipase to produce glycerol and fatty acids, increasing the content of free fatty acids (Ferreira, de Oliveira, Basso, Mendes, & Hirata, 2019). The organic contents of C—H, O—H, N—H, S—H, and other hydrogen-containing groups were different, resulting in various absorption peaks (Liu, Sun, Tan, & Liu, 2020). In summary, the spectral similarity of the three types of rice samples with different ratios of aged rice was high, and distinguishing them directly through spectral data analysis was difficult. Chemometric models were used to conduct in-depth data mining and analysis of the NIR spectral data of three rice adulteration samples with different ratios of aged rice.

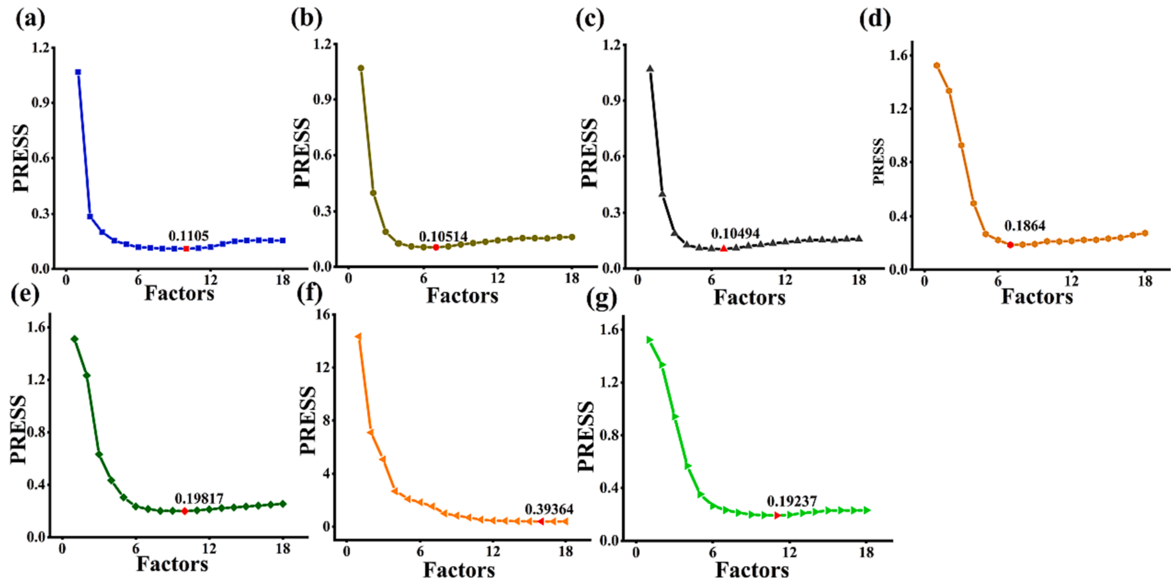

Fig. 2. Changes in the PRESS value of the PLSR model under different pre-treatments of Chinese Daohuaxiang: (a) MMN, (b) MSC, (c) SNV, (d) smoothing, (e) first derivative, (f) second derivative, and (g) raw data (RAW).

Table 1
Classification results of the PLSR model under different preprocessing of Chinese Daohuaxiang (CDHX), southern japonica rice (SJR), and late japonica rice (LJR).

Preprocessing methods	Rice types	Optimal factor number	Training set		Prediction set	
			R^2c	RMSEC	R^2p	RMSEP
RAW	CDHX	11	0.996987	0.025159	0.618506	0.252805
	SJR	15	0.997089	0.024732	0.597654	0.256009
	LJR	15	0.997632	0.022322	0.597209	0.263094
MMN	CDHX	10	0.998258	0.01916	0.626319	0.251943
	SJR	11	0.997058	0.024862	0.595638	0.255856
	LJR	13	0.997641	0.018521	0.618621	0.256704
MSC	CDHX	7	0.998014	0.020452	0.617383	0.25247
	SJR	7	0.996271	0.0282	0.580267	0.258538
	LJR	10	0.997114	0.024628	0.619879	0.255979
SNV	CDHX	7	0.998018	0.020431	0.619904	0.25276
	SJR	15	0.998438	0.018148	0.603548	0.254585
	LJR	10	0.996863	0.025666	0.622471	0.25276
Smoothing	CDHX	7	0.996242	0.028072	0.627392	0.250512
	SJR	13	0.996567	0.02684	0.597479	0.257261
	LJR	12	0.996833	0.025789	0.605525	0.260937
First derivative	CDHX	10	0.996989	0.02515	0.606555	0.256777
	SJR	10	0.994789	0.032997	0.585074	0.257232
	LJR	16	0.996451	0.027286	0.578202	0.26591
Second derivative	CDHX	16	0.998857	0.015529	0.58213	0.25814
	SJR	14	0.995548	0.030526	0.588297	0.258525
	LJR	15	0.997242	0.024078	0.559978	0.27119

3.3. Rice adulteration identification based on the PLSR model

The collected NIR spectra of the three kinds of rice adulteration samples were divided into training and prediction sets, and the PLSR model was constructed by using the NIR spectra of the training set. Fig. 2, S2, and S3 showed the distribution of the predicted residual sum of squares (PRESS) in the PLSR model under different preprocessing of the training sets with different proportions of Chinese Daohuaxiang, southern japonica rice, and late japonica rice. With the increase in the number of factors, PRESS decreased continuously. When PRESS reached the minimum value, and the number of factors continuously increased, the value of the PRESS became stable, indicating that the factor number at the minimum PRESS was optimal (Sarkodie & Strezov, 2018). In the process of establishing the PLSR model, the number of factors selected by the model presented a great influence on the accuracy of the model. The fewer the number of factors, the lesser the influence of the model, which may lead to the decrease of the model accuracy. The higher the number of factors, the more comprehensive the point calculated by the model and the closer to the real situation. However, an extremely large number of factors may not only increase the computational complexity of the model but also increase the variables with low or no correlation, leading to the overfitting of the analysis results (Yan et al., 2019).

The results of the PLSR classification model were established by determining the optimal number of factors under different preprocessing methods for Chinese Daohuaxiang, southern japonica rice, and late japonica rice (Table 1). The prediction results of the PLSR model using the first- and second-order derivative smoothing were slightly poorer than those of the original spectrum. The reason may the increased noise and reduced accuracy of the model due to derivative conversion. These preprocessing steps were not considered for the subsequent spectral optimization. The prediction results of the PLSR model established by normalization method, MSC, SNV, and smoothing method were better than those of the PLSR model established by the original spectrum. In the rice adulteration analysis of Chinese Daohuaxiang, the model

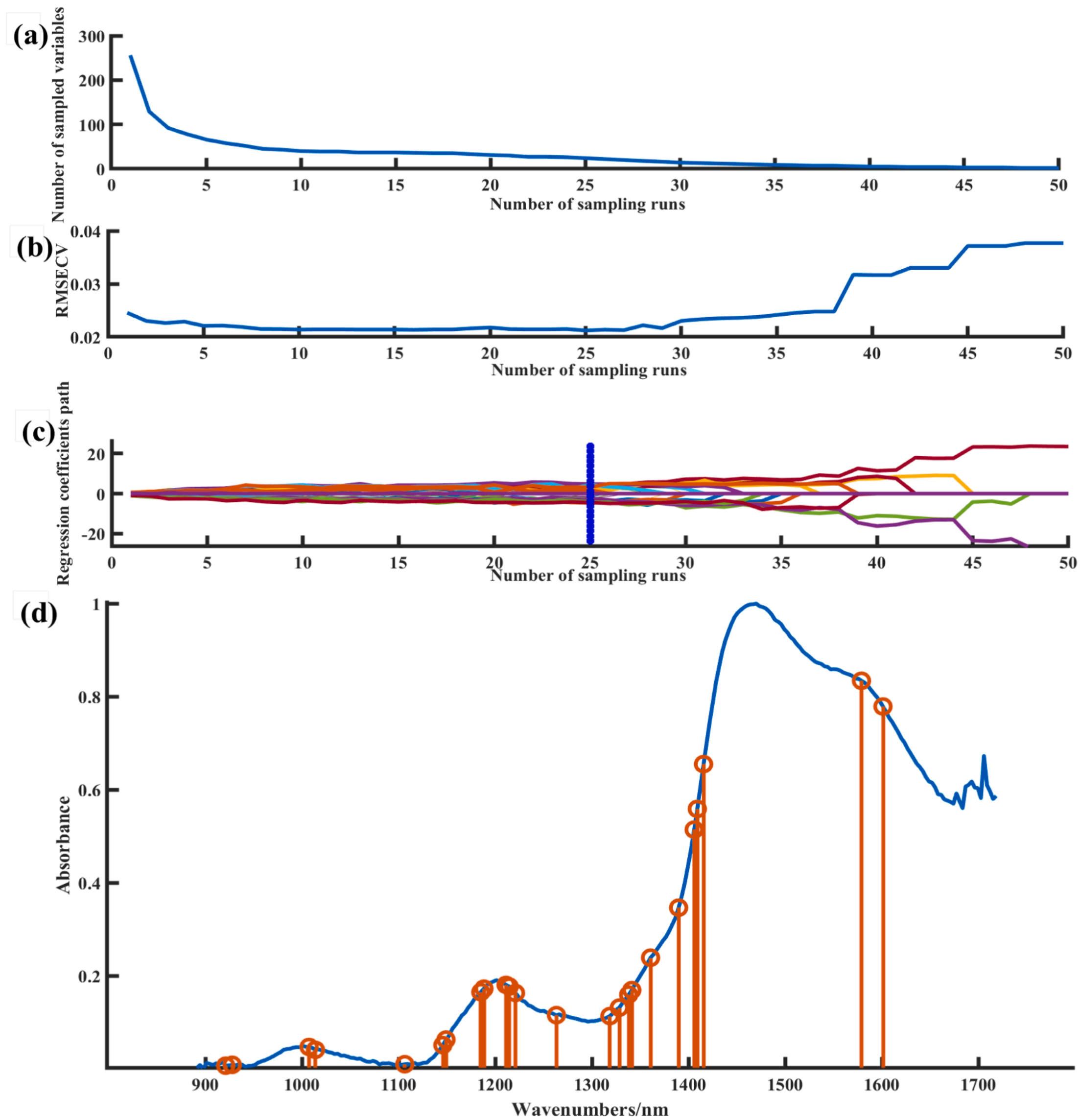

Fig. 3. Changes in the number of selected variables (a); changes in RMSECV (b); regression coefficients of each variable during the calculations of CARS algorithm (c); and distribution of selected characteristic variables by CARS (d).

accuracy obtained by smoothing treatment was the highest, with R^2p and RMSEP reaching 0.6263 and 0.251943, respectively. In the rice adulteration DA of southern japonica rice, the model accuracy obtained by SNV preprocessing was the highest, with R^2p and RMSEP of 0.6035 and 0.254585, respectively. Similarly, in the analysis of late japonica rice, the model accuracy obtained by SNV preprocessing was the highest, with R^2p and RMSEP amounting to 0.6225 and 0.25276, respectively. However, on the whole, the R^2p of rice adulteration analysis of three kinds of rice under different preprocessing was still dissatisfactory. Thus, the discriminant model must be optimized further and the characteristic variables in NIRS be extracted to improve the accuracy of the model.

3.4. CARS-PLSR model

CARS algorithm was used to simplify the model to eliminate redundant variables in NIRS and improve the stability and accuracy of the PLSR model. Fig. 3a showed the process of CARS optimization of characteristic variables under the normalization preprocessing of Chinese Daohuaxiang. With the increase in sample collection times, the

Table 2
Feature variables of Chinese Daohuaxiang, southern japonica rice, and late japonica rice spectrum extracted by CARS algorithm under different preprocessing.

Preprocessing methods	Operation times			Variables number		
	Chinese Daohuaxiang	Southern japonica rice	Late japonica rice	Chinese Daohuaxiang	Southern japonica rice	Late japonica rice
RAW	27	27	25	20	20	24
MMN	25	32	23	24	12	29
MSC	27	31	25	20	13	24
SNV	28	28	25	18	18	24
Smoothing	28	27	30	18	20	14

Table 3
Different preprocessing PLSR models of Chinese Daohuaxiang after the screening of feature variables.

Preprocessing methods	Optimal factor number	Training set		Prediction set	
		R^2c	RMSEC	R^2p	RMSEP
RAW	7	0.997139	0.024521	0.969242	0.078805
MMN	7	0.998373	0.01916	0.997996	0.020557
MSC	4	0.998186	0.019548	0.995823	0.029529
SNV	4	0.998155	0.019713	0.997698	0.021989
Smoothing	6	0.996713	0.028072	0.997329	0.023523

number of redundant variables in the NIRS of samples decreased continuously, the regression coefficient path changed continuously, and the RMSECV decreased continuously and reached the minimum. When the sampling operation time was 25, the RMSECV reached the minimum, increased rapidly, and then became stable (Fig. 3b). Thus, the redundant variables in the spectrum were effectively removed, and the number of characteristic variables in the spectrum decreased from the initial number of 256 to 24.

Similarly, the feature extraction of spectral data under different preprocessing of three types of rice samples was carried out using the CARS algorithm (Table 2). The CARS-PLSR model was constructed through optimization of the original model by extracting the characteristic variables, and the optimal factor number of the CARS-PLSR model was determined by PRESS. Under the optimal factor number, the results of the PLSR model were determined (Tables 3, S1, and S2). In addition, the MMN-CARS-PLSR model showed a higher accuracy than the other models in the adulteration DA of Chinese Daohuaxiang, with the R^2p of 0.998 and RMSEP of 0.020557. The SNV-CARS-PLSR model presented a higher accuracy than the other models in the adulteration analysis of southern japonica rice, with the R^2p of 0.994 and RMSEP of 0.036187. The SNV-CARS-PLSR model also presented a higher accuracy for late japonica rice, with the R^2p of 0.996 and RMSEP of 0.031132.

In comparison with previous studies, the CARS-PLSR model in this research presented a comparable classification accuracy with potential applications (Kamruzzaman, Kalita, Ahmed, ElMasry, & Makino, 2022). The PLS-DA approach provided satisfactory results in the discrimination of adulterated rice samples with a total classification rate of 82.4 % (Le Nguyen Doan, Nguyen, Marini, & Biancolillo, 2021). The optimum PLSR model based on the NIRS data shows promise in determining rice adulteration (Lapcharoensuk et al., 2019). In addition, the utilization of calibration transfer on the spectral data for the PLSR model provides a strategy for the achievement of a stable model with a satisfied standard deviation of RMSEP, allowing the prediction of Thai jasmine rice adulteration without the information on blended rice (Wongsaipun, Theanjumpol, & Kittiwachana, 2021). Moreover, when combined with CARS to achieve a reduced set of variables, the presented PLS model is a feasible approach for the identification of notoginseng powder adulteration (Chen, Tan, Lin, & Li, 2019).

The results showed that the discriminant accuracy of the established model for extracting characteristic wavelength variables was significantly improved by CARS. Especially, the analysis based on the SNV-PLSR-CARS model indicated excellent application potential for the evaluation of aged-rice adulteration. This finding was consistent with the report indicating that CARS algorithm is used to screen feature variables to improve the accuracy of models (Matuszyk et al., 2020; Pei & Huang, 2016; Song, Du, Li, Tang, & Huang, 2020). In comparison with other algorithms, the application of the CARS-PLS model for the determination of chlorpyrifos residue in rice has achieved optimum results (Jiang, Mehedi Hassan, Jiao, Li, & Chen, 2021). It was also reported that the CARS model combined with SVM and SPA presented the highest prediction precisions when used for the analysis of spectrum data (Ahmad et al., 2021; Su, Bakalis, & Sun, 2019). As above, CARS algorithm was achieved to extract the optimization model of characteristic variables effectively and reduce redundant wavelength variables for the improvement of prediction accuracy.

4. Conclusion

To ensure food safety, the aged-rice adulteration analysis of three kinds of rice was performed by NIRS combined with the PLSR model. The results showed that the PLSR model established by the normalization method, MSC, SNC, and smoothing preprocessing presented better prediction accuracy than the PLSR model established by the original spectrum. However, the accuracy remained dissatisfactory. To eliminate the redundant spectral variables and improve the stability and accuracy of the analysis model, we employed the CARS algorithm to simplify the model. The MMN-CARS-PLSR model showed a higher accuracy in the aged-rice adulteration analysis of Chinese Daohuaxiang than the other models, with R^2p of more than 0.99 and RMSEP of 0.020557. Compared with the other models, the SNV-CARS-PLSR model presented a higher accuracy in the purity analysis of southern japonica rice, whereas it showed a higher accuracy in aged-rice adulteration analysis of late japonica, with R^2p of 0.9955 and RMSEP of 0.031132. In conclusion, the constructed CARS-PLSR model not only greatly reduced the number of characteristic variables required by the spectrum but also improved the identification accuracy of aged-rice adulteration for Chinese Daohuaxiang, southern japonica rice, and late japonica rice. The proposed simple and accurate detection method provides a new alternative for the quality control of commercial rice.

Declaration of Competing Interest

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.

Data availability

Data will be made available on request.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (Grant No. 31902204).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2022.100539.

References

Ahmad, H., Sun, J., Nirere, A., Shaheen, N., Zhou, X., & Yao, K. (2021). Classification of tea varieties based on fluorescence hyperspectral image technology and ABC-SVM algorithm. *Journal of Food Processing and Preservation, 45*(3), e15241.

Aricetti, J. A., & Tubino, M. (2012). A visual titration method for the determination of the acid number of oils and fats: A green alternative. *Journal of the American Oil Chemists' Society, 89*(11), 2113–2115.

Chen, H., Tan, C., Lin, Z., & Li, H. (2019). Quantifying several adulterants of notoginseng powder by near-infrared spectroscopy and multivariate calibration. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 211*, 280–286.

China National Standard, GB/T 20569-2006. Guidelines for evaluation of paddy storage character. 2006.

Eleni, K., Vasileios, P., & Paraskevi, M. (2021). Feasibility of application of near infrared reflectance (NIR) spectroscopy for the prediction of the chemical composition of traditional sausages. *Applied Sciences, 11*(23), 11282.

Ferreira, M. M., de Oliveira, G. F., Basso, R. C., Mendes, A. A., & Hirata, D. B. (2019). Optimization of free fatty acid production by enzymatic hydrolysis of vegetable oils using a non-commercial lipase from *Geotrichum* candidum. *Bioprocess and Biosystems Engineering, 42*(10), 1647–1659.

Firmani, P., De Luca, S., Bucci, R., Marini, F., & Biancolillo, A. (2019). Near infrared (NIR) spectroscopy-based classification for the authentication of Darjeeling black tea. *Food Control, 100*, 292–299.

Huerta, M., Leiva, V., Liu, S., Rodríguez, M., & Villegas, D. (2019). On a partial least squares regression model for asymmetric data with a chemical application in mining. *Chemometrics and Intelligent Laboratory Systems, 190*, 55–68.

Jain, S., Mishra, P. K., Mishra, J., & Thakare, V. V. (2020). Design and analysis of H-Shape patch sensor for rice quality detection. *Materials Today: Proceedings, 29*, 581–586.

Jiang, L., Mehedi Hassan, M., Jiao, T., Li, H., & Chen, Q. (2021). Rapid detection of chlorpyrifos residue in rice using surface-enhanced Raman scattering coupled with chemometric algorithm. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 261*, Article 119996.

Kademi, H. I., Ulusoy, B. H., & Hecer, C. (2019). Applications of miniaturized and portable near infrared spectroscopy (NIRS) for inspection and control of meat and meat products. *Food Reviews International, 35*(3), 201–220.

Kamruzzaman, M., Kalita, D., Ahmed, M. T., ElMasry, G., & Makino, Y. (2022). Effect of variable selection algorithms on model performance for predicting moisture content in biological materials using spectral data. *Analytica Chimica Acta, 1202*, Article 339390.

Khan, P. W., & Byun, Y. C. (2020). Genetic algorithm based optimized feature engineering and hybrid machine learning for effective energy consumption prediction. *IEEE Access, 8*, 196274–196286.

Lapcharoensuk, R., Chalachai, S., Sinjaru, S., Singsriand, P., Hongwiangjan, J., & Yaemphochai, N. (2019). Quantitative detection of pepper powder adulterated with rice powder using Fourier-transform near infrared spectroscopy. *IOP Conference Series: Earth and Environmental Science* (Vol. 301, p. 012068): IOP Publishing.

Le Nguyen Doan, D., Nguyen, Q. C., Marini, F., & Biancolillo, A. (2021). Authentication of rice (*Oryza sativa* L.) using near infrared spectroscopy combined with different chemometric classification strategies. *Applied Sciences, 11*(1), 362.

Li, Q., Huang, Y., Song, X., Zhang, J., & Min, S. (2019). Moving window smoothing on the ensemble of competitive adaptive reweighted sampling algorithm. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 214*, 129–138.

Liu, J., Sun, S., Tan, Z., & Liu, Y. (2020). Nondestructive detection of sunset yellow in cream based on near-infrared spectroscopy and interval random forest. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 242*, Article 118718.

Liu, Y., Li, Y., Peng, Y., Ma, S., & Yan, S. (2021). A feasibility quantitative analysis of free fatty acids in polished rice by fourier transform near-infrared spectroscopy and chemometrics. *Journal of Food Science, 86*(8), 3434–3446.

Maduro Dias, C. S. A. M., Nunes, H. P., Melo, T. M. M. V., Rosa, H. J. D., Silva, C. C. G., & Borba, A. E. S. (2021). Application of near infrared reflectance (NIR) spectroscopy to predict the moisture, protein, and fat content of beef for gourmet hamburger preparation. *Livestock Science, 254*, Article 104772.

Matuszyk, E., Sierka, E., Rodewald, M., Bae, H., Meyer, T., Kus, E., … Baranska, M. (2020). Differential response of liver sinusoidal endothelial cells and hepatocytes to oleic and palmitic acid revealed by Raman and CARS imaging. *Biochimica et Biophysica Acta (BBA) – Molecular Basis of Disease, 1866*(6), Article 165763.

Mishra, P., & Nikzad-Langerodi, R. (2021). A brief note on application of domain-invariant PLS for adapting near-infrared spectroscopy calibrations between different physical forms of samples. *Talanta, 232*, Article 122461.

Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives–A review. *Analytica Chimica Acta, 1026*, 8–36.

Pei, M., & Huang, X. (2016). Determination of trace phenolic acids in fruit juice samples using multiple monolithic fiber solid-phase microextraction coupled with high-performance liquid chromatography. *Analytical Methods, 8*(18), 3831–3838.

Rady, A., Fischer, J., Reeves, S., Logan, B., & James Watson, N. (2019). The effect of light intensity, sensor height, and spectral pre-processing methods when using NIR spectroscopy to identify different allergen-containing powdered foods. *Sensors, 20*(1), 230.

Ren, G., Ning, J., & Zhang, Z. (2021). Multi-variable selection strategy based on near-infrared spectra for the rapid description of dianhong black tea quality. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 245*, Article 118918.

Saikrishna, A., Dutta, S., Subramanian, V., Moses, J., & Anandharamakrishnan, C. (2018). Ageing of rice: A review. *Journal of Cereal Science, 81*, 161–170.

Sarkodie, S. A., & Strezov, V. (2018). Assessment of contribution of Australia's energy production to CO_2 emissions and environmental degradation using statistical dynamic approach. *Science of the Total Environment, 639*, 888–899.

Śliwińska-Bartel, M., Burns, D. T., & Elliott, C. (2021). Rice fraud a global problem: A review of analytical tools to detect species, country of origin and adulterations. *Trends in Food Science & Technology, 116*, 36–46.

Song, X., Du, G., Li, Q., Tang, G., & Huang, Y. (2020). Rapid spectral analysis of agro-products using an optimal strategy: Dynamic backward interval PLS–competitive adaptive reweighted sampling. *Analytical and Bioanalytical Chemistry, 412*(12), 2795–2804.

Su, W. H., Bakalis, S., & Sun, D. W. (2019). Chemometrics in tandem with near infrared (NIR) hyperspectral imaging and Fourier transform mid infrared (FT-MIR) microspectroscopy for variety identification and cooking loss determination of sweet potato. *Biosystems Engineering, 180*, 70–86.

Su, W. H., Yang, C., Dong, Y., Johnson, R., Page, R., Szinyei, T., … Steffenson, B. J. (2021). Hyperspectral imaging and improved feature variable selection for automated determination of deoxynivalenol in various genetic lines of barley kernels for resistance screening. *Food Chemistry, 343*, Article 128507.

Tran, T. U., Suzuki, K., Okadome, H., Ikezaki, H., Homma, S., & Ohtsubo, K. I. (2005). Detection of changes in taste of japonica and indica brown and milled rice (*Oryza sativa* L.) during storage using physicochemical analyses and a taste sensing system. *Journal of Agricultural and Food Chemistry, 53*(4), 1108–1118.

Wang, H., Wang, Y., Wang, R., Liu, X., Zhang, Y., Zhang, H., & Chi, C. (2022). Impact of long-term storage on multi-scale structures and physicochemical properties of starch isolated from rice grains. *Food Hydrocolloids, 124*, Article 107255.

Wongsaipun, S., Theanjumpol, P., & Kittiwachana, S. (2021). Development of a universal calibration model for quantification of adulteration in Thai jasmine rice using near-infrared spectroscopy. *Food Analytical Methods, 14*(5), 997–1010.

Wu, P., Li, C., Bai, Y., Yu, S., & Zhang, X. (2019). A starch molecular basis for aging-induced changes in pasting and textural properties of waxy rice. *Food Chemistry, 284*, 270–278.

Xu, J., Liu, K., & Zhang, C. (2021). Electronic nose for volatile organic compounds analysis in rice aging. *Trends in Food Science & Technology, 109*, 83–93.

Yan, H., Song, X., Tian, K., Gao, J., Li, Q., Xiong, Y., & Min, S. (2019). A modification of the bootstrapping soft shrinkage approach for spectral variable selection in the issue of over-fitting, model accuracy and variable selection credibility. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 210*, 362–371.

Zhai, Y., Pan, L., Luo, X., Zhang, Y., Wang, R., & Chen, Z. (2022). Effect of electron beam irradiation on storage, moisture and eating properties of high-moisture rice during storage. *Journal of Cereal Science, 103*, Article 103407.

Zheng, K., Li, Q., Wang, J., Geng, J., Peng, C., Tao, S., … Du, Y. (2012). Stability competitive adaptive reweighted sampling (SCARS) and its applications to multivariate calibration of NIR spectra. *Chemometrics & Intelligent Laboratory Systems, 112*, 48–54.

Zhou, D., Yu, Y., Hu, R., & Li, Z. (2020). Discrimination of *Tetrastigma hemsleyanum* according to geographical origin by near-infrared spectroscopy combined with a deep learning approach. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 238*, Article 118380.