

Visible and Near-Infrared Spectroscopic Discriminant Analysis Applied to Identification of Soy Sauce Adulteration

Chunli Fu¹, Jiemei Chen¹, Lifang Fang², Tao Pan^{3*}

¹Department of Biological Engineering, Jinan University, Guangzhou, China

²Guangdong Langtao Science and Technology, Co., Ltd., Dongguan, China

³Department of Optoelectronic Engineering, Jinan University, Guangzhou, China

Email: *466945939@qq.com

How to cite this paper: Fu, C.L., Chen, J.M., Fang, L.F. and Pan, T. (2022) Visible and Near-Infrared Spectroscopic Discriminant Analysis Applied to Identification of Soy Sauce Adulteration. *American Journal of Analytical Chemistry*, 13, 51-62.

<https://doi.org/10.4236/ajac.2022.132005>

Received: January 18, 2022

Accepted: February 25, 2022

Published: February 28, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International

License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The identification of soy sauce adulteration can avoid fraud, and protect the rights and interests of producers and consumers. Based on two measurement models (1 mm, 10 mm), the visible and near-infrared (Vis-NIR) spectroscopy combined with standard normal variate-partial least squares-discriminant analysis (SNV-PLS-DA) was used to establish the discriminant analysis models for adulterated and brewed soy sauces. Chubang soy sauce was selected as an identification brand (negative, 70). The adulteration samples (positive, 72) were prepared by mixing Chubang soy sauce and blended soy sauce with different adulteration rates. Among them, the “blended soy sauce” sample was concocted of salt water (NaCl), monosodium glutamate (C₅H₁₀NNaO₅) and caramel color (C₆H₈O₃). The rigorous calibration-prediction-validation sample design was adopted. For the case of 1 mm, five waveband models (visible, short-NIR, long-NIR, whole NIR and whole scanning regions) were established respectively; in the case of 10 mm, three waveband models (visible, short-NIR and visible-short-NIR regions) for unsaturated absorption were also established respectively. In independent validation, the models of all wavebands in the cases of 1 mm and 10 mm have achieved good discrimination effects. For the case of 1 mm, the visible model achieved the optimal validation effect, the validation recognition-accuracy rate (RAR_v) was 99.6%; while in the case of 10 mm, both the visible and visible-short-NIR models achieved the optimal validation effect (RAR_v = 100%). The detection method does not require reagents and is fast and simple, which is easy to promote the application. The results can provide valuable reference for designing small dedicated spectrometers with different measurement modal and different spectral regions.

Keywords

Visible and Near-Infrared Spectroscopy, Soy Sauce, Adulteration Identification, Partial Least Squares-Discriminant Analysis, Standard Normal Variate

1. Introduction

High-quality brewed soy sauce is mainly made from soybeans, starch and wheat, which are fermented by microorganisms to form a flavorful liquid condiment. Soy sauce is one of the most widely used condiments in the world. Due to the huge consumer market, counterfeiting and adulteration by criminals often occur in order to obtain high profits. For example, use blended soy sauce that was made up of salt water, monosodium glutamate and caramel color, or mix blended soy sauce into high-quality brewed soy sauce to adulterate. This behavior may cause some components to be excessive or produce harmful components in soy sauce. The identification of soy sauce adulteration can avoid fraud, protect the rights and interests of producers and consumers, and contribute to food safety.

Normally, the identification methods for soy sauce adulteration mainly adopt the quantitative analysis of characteristic components [1] [2] [3]. It usually requires the quantitative analysis of a variety of minute quantity characteristic components, which involves a variety of high-end measurement methods that are complex and expensive. A quick and easy detection method that can be used in the field has important application value.

Near-infrared (NIR) spectroscopy mainly reflects the vibration absorption of the overtones and combination frequencies of the hydrogen-containing group X-H, thus, it has significant absorption for most agricultural products and foods. This method usually does not require reagents, and it can measure samples directly, with the advantage of being quick and easy. Combined with the visible light region, visible and near-infrared (Vis-NIR) spectroscopy has been applied in many fields, such as agriculture and food [4] [5] [6] [7], environment [8] [9], and biomedicine [10] [11] [12] [13].

The spectral qualitative discriminant analysis is based on the spectral similarity of the same type of samples and the spectral difference of different types of samples. For the identification of samples with small differences in component content, the qualitative discriminant analysis is more convenient than quantitative analysis. It is a hot research direction in recent years, and has been applied in many fields, such as melon genotype [14], edible oil types [15], transgenic sugarcane leaf [16] [17], milk powder adulteration identification [18] [19], wine identification [20] [21] [22], rice seed authenticity identification [23] as well as thalassemia screening [17] [24].

Vis-NIR spectroscopy had high application potential, but its application research in soy sauce analysis mainly focused on the quantitative analysis of quality indicators, e.g., total acid, total nitrogen and amino acid nitrogen [25]-[30]. Little re-

search of soy sauce qualitative discriminant analysis was reported, such as detection of adulteration of soy sauce by brine [31], and classification of soy sauce [30].

The partial least squares-discriminant analysis (PLS-DA) is an effective spectral discriminant analysis method. Based on sample categorical variables (positive 1 and negative 0), PLS quantitative analyses have been performed. Samples are classified according to their predicted categorical variables. Based on matrix theory and applied examples, some previous studies have shown that the discrimination effect of PLS-DA method is superior to that of PCA-LDA in most cases [32] [33] [34]. In the present study, PLS-DA was used to establish NIR discriminant analysis models. Standard Normal Variate (SNV) is a common spectral preprocessing method. It associates spectral changes with the component concentrations, increasing the difference between spectra, thereby improving the robustness and prediction ability of the models [35] [36] [37]. In the present study, SNV combined with PLS-DA, denoted as SNV-PLS-DA, was used to establish NIR discriminant analysis models.

In the present paper, the Vis-NIR spectroscopy combined with the SNV-PLS-DA method was used to establish the discriminant analysis models for adulterated and brewed soy sauces. Among them, the famous Chubang brand soy sauces were used as identification samples; the blended soy sauces made up of salt water, monosodium glutamate and caramel color, were mixed into the Chubang brand samples in different proportions and were used as adulteration samples.

In addition, from the spectral experiments of water-system samples, it was found that the use of a short transmission measurement optical path can avoid saturated absorption in the long-wavelength region, and the use of a long transmission measurement optical path can highlight the difference in absorption in the short-wavelength region. In view of this, the cuvettes of 1 mm and 10 mm were used to measure the transmission spectra of soy sauce samples. For the case of 1 mm, five waveband models (visible, short-NIR, long-NIR, whole NIR and whole scanning regions) were established and compared respectively; in the case of 10 mm, three waveband models (visible, short-NIR and visible-short-NIR regions) for unsaturated absorption were also established and compared respectively. It provided valuable reference for designing small dedicated spectrometers with different measurement modals and different spectral regions.

2. Materials and Methods

2.1. Experimental Materials, Instruments, and Measurement Methods

The Chubang soy sauce was collected from regular sale channels as identification brand (70 bottles, 1 sample for each bottle). The mother solution sample (100 ml) of “blended soy sauce” was concocted of salt water (NaCl, wt 15%), monosodium glutamate ($C_5H_{10}NNaO_5$, 10 g) and caramel color ($C_6H_8O_3$, 4 g). Among them, the concentrations of salt water and monosodium glutamate were based on salt

(18 g/100ml) and amino acid nitrogen (≥ 1.3 g/100ml) contents of Chubang soy sauce; the content of caramel color was based on the color of Chubang soy sauce. The adulteration samples (positive) were prepared by mixing the Chubang soy sauce and the blended soy sauce samples according to different adulteration rates (3% - 69% with tolerance 3%, 100%). Among them, the adulteration samples contained 24 adulterated rates, each rate had 3 samples, a total number of 72. Two types of samples, a total number of 142, were used for spectral measurement.

The XDS Rapid ContentTM Liquid Grating Spectrometer (FOSS, Denmark) and transmission accessories with 1 mm and 10 mm cuvettes were used for spectral measurement. Spectral scope ranged as 400 - 2498 nm with a 2 nm wavelength interval. Wavebands of 400 - 1100 nm and 1100 - 2498 nm were used for Si and PbS detection, respectively. Using the cuvettes of two optical path lengths of short and long (1 mm, 10 mm) respectively, each sample was measured five times. The experimental temperature and humidity were $25^{\circ}\text{C} \pm 1^{\circ}\text{C}$ and $45\% \pm 1\%$, respectively. The obtained spectral data set (negative 350, positive 360, a total of 710) of each measurement modal (1 mm, 10 mm) was used for modeling and validation respectively.

2.2. Calibration-Prediction-Validation Design and Evaluation Indicators

A rigorous calibration-prediction-validation “three-stage” sample experiment design was adopted. Samples of each category were randomly divided into calibration, prediction, and validation sets respectively. The calibration and prediction sets (all including two types of samples) were used for modeling and parameter optimization; and the independent validation samples (also including two types of samples) that not involved in modeling were used to validate the selected models, thereby obtain objective evaluation.

Among them, the negative samples (70 samples, 350 spectra) were randomly divided into calibration (24 samples, 120 spectra), prediction (24 samples, 120 spectra) and validation (22 samples, 110 spectra) sets. Among the positive samples, three samples were formulated for each adulteration rate, and they were divided into calibration, prediction, and validation sets, respectively. The all positive samples were divided to the calibration (24 samples, 120 spectra), prediction (24 samples, 120 spectra), and validation (24 samples, 120 spectra) sets. The calibration-prediction-validation division for the spectra of two types of samples is shown in **Table 1**.

In summary, a total of 710 spectra (negative 350, positive 360) were used for calibration (negative 120, positive 120, a total of 240), prediction (negative 120, positive 120, a total of 240) and validation set (negative 110, positive 120, a total of 230).

2.3. PLS-DA

Based on the calibration and prediction sets of Vis-NIR spectra, the framework

Table 1. Calibration-prediction-validation division for the spectra of two types of samples.

	Negative	Positive	Total
Calibration	120	120	240
Prediction	120	120	240
Validation	110	120	230
Total	350	360	710

of PLS-DA algorithm are as follows. (1) Each positive and negative sample was assigned the categorical variable (C) values 1 and 0, respectively. (2) The number of PLS latent variables (LV) was set as 1 to 10. Based on the spectra and categorical variables of calibration samples, the PLS coefficients for each LV were calculated. (3) Based on the spectrum of each sample (calibration, prediction) and the obtained PLS coefficients, the prediction value (\tilde{C}) of categorical variable for the sample was calculated for each LV. (4) When $\tilde{C} \geq 0.5$, the sample was judged to positive, and when $\tilde{C} < 0.5$, the sample was judged to negative. (5) The optimal LV was determined according to the optimal recognition-accuracy rate of the prediction set.

2.4. Model Evaluation Indicators

According to the actual category (positive and negative) of the samples, the nine recognition-accuracy rates (RARs) corresponding to the positive or negative or total and calibration or prediction or total were proposed and calculated as follows:

$$\text{RAR}_C^+ = \frac{\tilde{N}_C^+}{N_C^+}, \quad \text{RAR}_C^- = \frac{\tilde{N}_C^-}{N_C^-}, \quad \text{RAR}_C = \frac{\tilde{N}_C^+ + \tilde{N}_C^-}{N_C^+ + N_C^-} \quad (1)$$

$$\text{RAR}_P^+ = \frac{\tilde{N}_P^+}{N_P^+}, \quad \text{RAR}_P^- = \frac{\tilde{N}_P^-}{N_P^-}, \quad \text{RAR}_P = \frac{\tilde{N}_P^+ + \tilde{N}_P^-}{N_P^+ + N_P^-} \quad (2)$$

$$\text{RAR}^+ = \frac{\tilde{N}_C^+ + \tilde{N}_P^+}{N_C^+ + N_P^+}, \quad \text{RAR}^- = \frac{\tilde{N}_C^- + \tilde{N}_P^-}{N_C^- + N_P^-}, \quad \text{RAR}_{\text{Total}} = \frac{\tilde{N}_C^+ + \tilde{N}_C^- + \tilde{N}_P^+ + \tilde{N}_P^-}{N_C^+ + N_C^- + N_P^+ + N_P^-} \quad (3)$$

where N_C^+ , N_C^- , N_P^+ , and N_P^- were the numbers of actual positive and negative samples in the calibration and prediction sets, respectively; \tilde{N}_C^+ , \tilde{N}_C^- , \tilde{N}_P^+ , and \tilde{N}_P^- were the numbers of correctly recognized positive and negative samples in the calibration and prediction sets; From the negative, positive, calibration and prediction aspects, these indicators comprehensively evaluated the prediction effect of the discrimination model. The optimal LV was determined according to the maximum $\text{RAR}_{\text{Total}}$.

Next, the validation samples were identified following the previous steps. And referring to the actual category, the positive and negative validation RARs (RAR_V^+ and RAR_V^-) and the total validation RAR (RAR_V) were calculated as follows:

$$\text{RAR}_V^+ = \frac{\tilde{N}_V^+}{N_V^+}, \quad \text{RAR}_V^- = \frac{\tilde{N}_V^-}{N_V^-}, \quad \text{RAR}_V = \frac{\tilde{N}_V^+ + \tilde{N}_V^-}{N_V^+ + N_V^-} \quad (4)$$

where N_v^+ and N_v^- were the numbers of actual positive and negative samples in the validation set, respectively; and \tilde{N}_v^+ , \tilde{N}_v^- were the numbers of correctly recognized positive and negative samples in the validation set, respectively.

The computer algorithms for the above mentioned method were designed using MATLAB v7.6 software.

3. Results and Discussion

3.1. PLS-DA Modeling

In the case of 1 mm transmission measurement modal, the Vis-NIR spectra of all soy sauce samples in the whole scanning region (400 - 2498 nm) are shown in **Figure 1(a)**. In the case of 10 mm transmission measurement modal, the spectra of all samples were only displayed in the unsaturated spectral region (400 - 1388 nm) due to saturated absorption and noise in the long wavelength region (1400 - 2498 nm), as shown in **Figure 1(b)**. Among them, the green and red lines represent the negative and positive samples, respectively. In both cases, the spectra of negative sample group located at the edge of the spectra of positive sample group and were very similar to the edge spectra shape of the positive. Significant spectral differences between the two types of samples were not observed. Discriminant analysis based on chemometrics method was required.

SNV is an effective method of spectral preprocessing. It can reduce the spectral error caused by scattering and improve the prediction accuracy and stability of the model. In the present paper, SNV was used to correct the spectra of soy sauce samples. The SNV spectra of the two measurement modals are shown in **Figure 2**. Further, the average SNV spectra of the negative and positive samples of the two models were also calculated, see **Figure 3**. In the case of 1 mm, a small difference was observed between the average spectra of two types of samples; while in the case of 10 mm, a significant difference was observed between the average spectra, but the fluctuation of positive spectral group was very significant, so that the difference between the two spectral groups was still insignificant.

Firstly, the direct PLS-DA models were established by using the raw spectra.

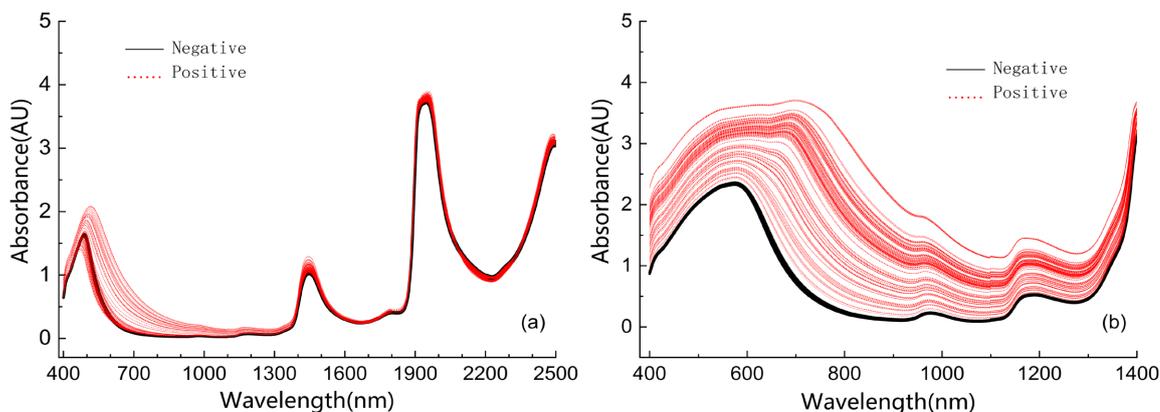


Figure 1. Vis-NIR spectra of two types of soy sauce samples: (a) 1 mm; (b) 10 mm.

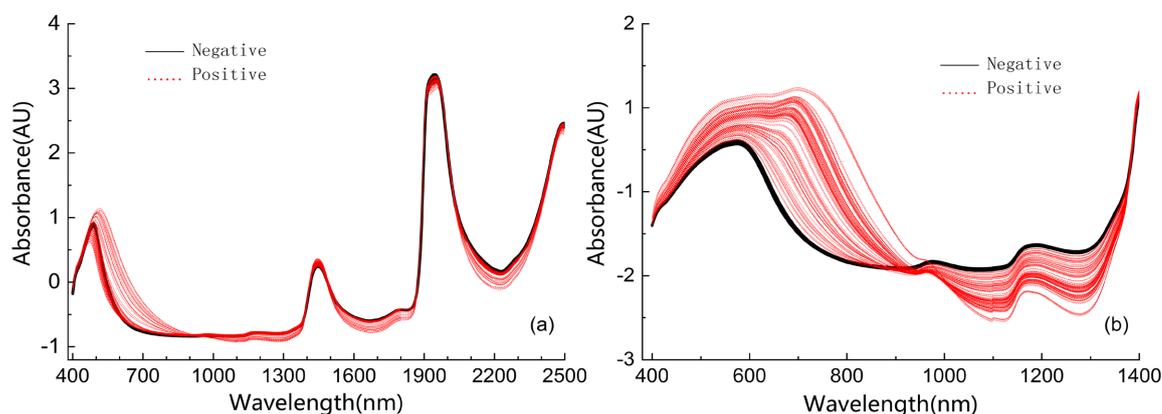


Figure 2. SNV spectra of two types of soy sauce samples: (a) 1 mm; (b) 10 mm.

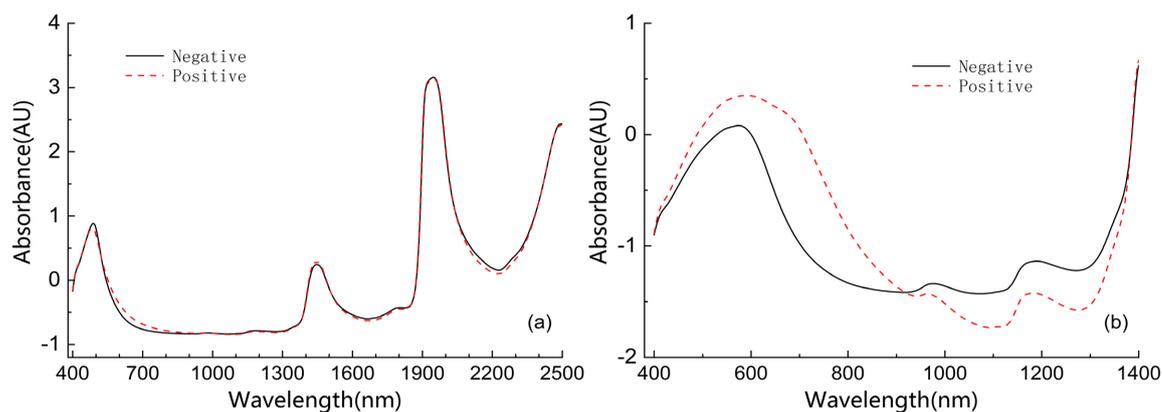


Figure 3. Average SNV spectra of two types of samples: (a) 1 mm; (b) 10 mm.

In the case of 1 mm, the direct PLS-DA models were established based on five spectral regions (visible, short-NIR, long-NIR, whole NIR and whole scanning regions), respectively. The results of the corresponding nine indicators are summarized in **Table 2**. Among them, the model of 400 - 780 nm obtained the optimal RAR_{Total} (99.4%). In the case of 10 mm, the direct PLS-DA models were established based on three spectral regions (visible, short-NIR and visible-short-NIR regions), respectively. The results of the corresponding nine indicators are summarized in **Table 3**. Among them, the models of 400 - 780 nm and 400 - 1388 nm obtained the optimal RAR_{Total} (100%).

Secondly, the SNV-PLS-DA model was established by using the SNV spectra. In the case of 1 mm, the SNV-PLS-DA models were established based on five spectral regions (visible, short-NIR, long-NIR, whole NIR and whole scanning regions), respectively. The results of the corresponding nine indicators are summarized in **Table 4**. Among them, the model of 400 - 780 nm obtained the optimal RAR_{Total} (100%). In the case of 10 mm, the SNV-PLS-DA models were established based on three spectral regions (visible, short-NIR and visible-short-NIR regions), respectively. The results of the corresponding nine indicators are summarized in **Table 5**. Among them, the models of 400 - 780 nm and 400 - 1388 nm obtained the optimal RAR_{Total} (100%).

Table 2. Modelling effects of the direct PLS-DA models for the case of 1 mm.

Waveband (nm)	LV	RAR_C^-	RAR_C^+	RAR_p^-	RAR_p^+	RAR_C	RAR_p	RAR^-	RAR^+	RAR_{Total}
400 - 780	9	100%	100%	100%	97.5%	100%	98.8%	100%	98.8%	99.4%
780 - 1400	8	99.2%	97.5%	100%	96.7%	98.3%	98.3%	99.6%	97.1%	98.3%
1400 - 2498	8	100%	93.3%	99.2%	87.5%	96.7%	93.3%	99.6%	90.4%	95.0%
780 - 2498	10	100%	95.8%	99.2%	88.3%	97.9%	93.8%	99.6%	92.1%	95.8%
400 - 2498	8	100%	100%	100%	95.8%	100%	97.9%	100%	97.9%	99.0%

Table 3. Modelling effects of the direct PLS-DA models for the case of 10 mm.

Waveband (nm)	LV	RAR_C^-	RAR_C^+	RAR_p^-	RAR_p^+	RAR_C	RAR_p	RAR^-	RAR^+	RAR_{Total}
400 - 780	4	100%	100%	100%	100%	100%	100%	100%	100%	100%
780 - 1388	10	100%	95.0%	100%	95.8%	97.5%	97.9%	100%	95.4%	97.7%
400 - 1388	5	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 4. Modelling effects of the SNV-PLS-DA models for the case of 1 mm.

Waveband (nm)	LV	RAR_C^-	RAR_C^+	RAR_p^-	RAR_p^+	RAR_C	RAR_p	RAR^-	RAR^+	RAR_{Total}
400 - 780	9	100%	100%	100%	100%	100%	100%	100%	100%	100%
780 - 1400	6	100%	100%	100%	98.3%	100%	99.2%	100%	99.2%	99.6%
1400 - 2498	9	100%	97.5%	99.2%	90.0%	98.8%	94.6%	99.6%	93.8%	96.7%
780 - 2498	10	100%	100%	99.2%	95.8%	100%	97.5%	99.6%	97.9%	98.8%
400 - 2498	10	100%	100%	100%	95.0%	100%	97.5%	100%	97.5%	98.8%

Table 5. Modelling effects of the SNV-PLS-DA models for the case of 10 mm.

Waveband (nm)	LV	RAR_C^-	RAR_C^+	RAR_p^-	RAR_p^+	RAR_C	RAR_p	RAR^-	RAR^+	RAR_{Total}
400 - 780	6	100%	100%	100%	100%	100%	100%	100%	100%	100%
780 - 1388	10	100%	100%	100%	96.7%	100%	98.3%	100%	98.3%	99.2%
400 - 1388	7	100%	100%	100%	100%	100%	100%	100%	100%	100%

In summary, the discriminant effect of SNV spectra was slightly better than that of the raw spectra.

3.2. Independent Validation

The 230 validation samples that not participated in the modeling process were used to validate the optimal SNV-PLS-DA models. The validation effects (RAR_v^- , RAR_v^+ , RAR_v) for the cases of 1 mm and 10 mm are summarized in **Table 6**.

The results showed that the SNV-PLS-DA models achieved good validation discriminant effects for both the 1 mm and 10 mm cases. In the case of 1 mm,

Table 6. Validation effects of the SNV-PLS-DA models for the cases of 1 mm and 10 mm.

	Wavebands (nm)	LV	RAR_{ν}^{-}	RAR_{ν}^{+}	RAR_{ν}
1 mm	400 - 780	9	100%	99.2%	99.6%
	780 - 1400	6	100%	97.9%	98.9%
	1400 - 2498	8	100%	96.2%	98.1%
	780 - 2498	8	100%	97.9%	98.9%
	400 - 2498	8	100%	98.3%	99.1%
10 mm	400 - 780	6	100%	100%	100%
	780 - 1388	8	100%	98.8%	99.4%
	400 - 1388	7	100%	100%	100%

the visible model achieved the optimal validation effect ($RAR_{\nu} = 99.6\%$); while in the case of 10 mm, both the visible and visible-short-NIR models achieved the optimal validation effect ($RAR_{\nu} = 100\%$). These experimental results can provide valuable reference for designing small dedicated spectrometer with different measurement modals and different spectral regions.

4. Conclusions

In this paper, based on two measurement models (1 mm, 10 mm), the Vis-NIR spectroscopy combined with the SNV-PLS-DA method was used to establish the discriminant analysis models for adulterated and brewed soy sauces. In the case of 1 mm, five waveband models were established and compared respectively; in the case of 10 mm, three waveband models for unsaturated absorption were also established and compared respectively.

The validation results showed that the models of all wavebands in the cases of 1 mm and 10 mm have achieved good discrimination effects. In the case of 1 mm, the visible model achieved the optimal validation effect ($RAR_{\nu} = 99.6\%$); while in the case of 10 mm, both the visible and visible-short-NIR models achieved the optimal validation effect ($RAR_{\nu} = 100\%$). These experimental results can provide valuable reference for designing small dedicated spectrometers with different measurement modals and different spectral regions. The detection method does not require reagents and is fast and simple, which is easy to promote the application.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No.61078040), and the Science and Technology Project of Guangdong Province of China (No.2014A020213016, No.2014A020212445).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Kim, D.B., Kim, S., Lee, S. and Yoo, M. (2017) Analytical Method Validation and Monitoring of Levulinic Acid in Soy Sauce and Their Application to Food. *Analytical Science and Technology*, **30**, 295-302. <https://doi.org/10.5806/AST.2017.30.5.295>
- [2] Park, S.K., Kim, C.T., Lee, J.W., Jhee, O.H., Om, A.S., Kang, J.S. and Moon, T.W. (2007) Analysis of Ethyl Carbamate in Korean Soy Sauce Using High-Performance Liquid Chromatography with Fluorescence Detection or Tandem Mass Spectrometry and Gas Chromatography with Mass Spectrometry. *Food Control*, **18**, 975-982.
- [3] Zhu, Y.H., Yang, Y., Zhou, Z.X., Li, G.R., Jiang, M., Zhang, C. and Chen, S.Q. (2010) Direct Determination of Free Tryptophan Contents in Soy Sauces and Its Application as an Index of Soy Sauce Adulteration. *Food Chemistry*, **1**, 159-162.
- [4] Tumsavas, Z. (2017) Application of Visible and Near Infrared Reflectance Spectroscopy to Predict Total Nitrogen in Soil. *Journal of Environmental Biology*, **38**, 1101-1106. [https://doi.org/10.22438/jeb/38/5\(SI\)/GM-29](https://doi.org/10.22438/jeb/38/5(SI)/GM-29)
- [5] Chen, H.Z., Pan, T., Chen, J.M. and Lu, Q.P. (2011) Waveband Selection for NIR Spectroscopy Analysis of Soil Organic Matter Based on SG Smoothing and MWPLS Methods. *Chemometrics and Intelligent Laboratory Systems*, **107**, 139-146.
- [6] Pan, T., Han, Y., Chen, J.M., Yao, L.J. and Xie, J. (2016) Optimal Partner Wavelength Combination Method with Application to Near-Infrared Spectroscopic Analysis. *Chemometrics and Intelligent Laboratory Systems*, **156**, 217-223. <https://doi.org/10.1016/j.chemolab.2016.05.022>
- [7] Pudelko, A. and Chodak, M. (2020) Estimation of Total Nitrogen and Organic Carbon Contents in Mine Soils with NIR Reflectance Spectroscopy and Various Chemometric Methods. *Geoderma*, **368**, Article ID: 114306.
- [8] Sousa, A.C., Lucio, M.M.L.M., Bezerra, O.F., Marcone, G.P.S., Pereira, A.F.C., Dantas, E.O., Fragoso, W.D., Araujo, M.C.U. and Galvao, R.K.H. (2007) A Method for Determination of COD in a Domestic Wastewater Treatment Plant by Using Near-Infrared Reflectance Spectrometry of Seston. *Analytica Chimica Acta*, **588**, 231-236.
- [9] Pan, T., Chen, Z.H., Chen, J.M. and Liu, Z.Y. (2012) Near-Infrared Spectroscopy with Waveband Selection Stability for the Determination of COD in Sugar Refinery Wastewater. *Analytical Methods*, **4**, 1046-1052. <https://doi.org/10.1039/c2ay05856a>
- [10] Pan, T., Wu, Z.T. and Chen, H.Z. (2012) Waveband Optimization for Near-Infrared Spectroscopic Analysis of Total Nitrogen in Soil. *Chinese Journal of Analytical Chemistry*, **40**, 920-924. <https://doi.org/10.3724/SP.J.1096.2012.10987>
- [11] Pan, T., Liu, J.M., Chen, J.M., Zhang, G.P. and Zhao, Y. (2013) Rapid Determination of Preliminary Thalassaemia Screening Indicators Based on Near-Infrared Spectroscopy with Wavelength Selection Stability. *Analytical Methods*, **5**, 4355-4362. <https://doi.org/10.1039/c3ay40732b>
- [12] Chen, J.M., Yin, Z.W., Tang, Y. and Pan, T. (2017) Vis-NIR Spectroscopy with Moving-Window PLS Method Applied to Rapid Analysis of Whole Blood Viscosity. *Analytical and Bioanalytical Chemistry*, **409**, 2737-2745. <https://doi.org/10.1007/s00216-017-0218-9>
- [13] Chen, J.M., Peng, L.J., Han, Y., Zhang, J. and Pan, T. (2018) A Rapid Quantification Method for the Screening Indicator for β -Thalassemia with Near-Infrared Spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **193**, 499-506.
- [14] Seregély, Z., Deák, T. and Bisztray, G.D. (2004) Distinguishing Melon Genotypes Using NIR Spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, **72**, 195-203.

- [15] Yang, H., Irudayaraj, J. and Paradkar, M.M. (2005) Discriminant Analysis of Edible Oils and Fats by FTIR, FT-NIR and FT-Raman Spectroscopy. *Food Chemistry*, **93**, 25-32.
- [16] Guo, H.S., Chen, J.M., Pan, T., Wang, J.H. and Cao, G. (2014) Vis-NIR Wavelength Selection for Non-Destructive Discriminant Analysis of Breed Screening of Transgenic Sugarcane. *Analytical Methods*, **6**, 8810-8816.
<https://doi.org/10.1039/C4AY01833H>
- [17] Yao, L.J., Xu, W.Q., Pan, T. and Chen, J.M. (2018) Moving-Window Bis-Correlation Coefficients Method for Visible and Near-Infrared Spectral Discriminant Analysis with Applications. *Journal of Innovative Optical Health Sciences*, **11**, Article ID: 1850005.
<https://doi.org/10.1142/S1793545818500050>
- [18] Capuano, E., Boerrigter-Eenling, R., Koot, A. and Ruth, S.M. (2015) Targeted and Untargeted Detection of Skim Milk Powder Adulteration by Near-Infrared Spectroscopy. *Food Analytical Methods*, **8**, 2125-2134.
<https://doi.org/10.1007/s12161-015-0100-3>
- [19] Xu, L.L., Li, W.Q., Zhu, H. and Wang, J.J. (2016) Detection of Adulteration of Milk Powder by Near Infrared Spectroscopy. *Journal of Food Safety and Food Quality- Archiv für Lebensmittelhygiene*, **7**, 3133-3137.
- [20] Santos, C.A.T.D., Pascoa, R.N.M.J., Sarraguca, M.C., Porto, P.A.L.S., Cerdeira, A.L., Gonzalez-Saiz, J.M., Pizarro, C. and Lopes, J.A. (2017) Merging Vibrational Spectroscopic Data for Wine Classification According to the Geographic Origin. *Food Research International*, **102**, 504-510.
- [21] Yu, J., Zhan, J.C. and Huang, W.D. (2017) Identification of Wine According to Grape Variety Using Near-Infrared Spectroscopy Based on Radial Basis Function Neural Networks and Least-Squares Support Vector Machines. *Food Analytical Methods*, **10**, 3306-3311. <https://doi.org/10.1007/s12161-017-0887-1>
- [22] Liao, S.X., Chen, J.M. and Pan, T. (2020) Visible and Near-Infrared Spectroscopic Discriminant Analysis Applied to Brand Identification of Wine. *American Journal of Analytical Chemistry*, **11**, 104-113.
- [23] Chen, J.M., Li, M.L., Pan, T., Pang, L.W., Yao, L.J. and Zhang, J. (2019) Rapid and Non-Destructive Analysis for the Identification of Multi-Grain Rice Seeds with Near-Infrared Spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **219**, 179-185.
- [24] Liu, K.Z., Tsang, K.S., Li, C.K., Shaw, R.A. and Mantsch, H.H. (2003) Infrared Spectroscopic Identification of β -Thalassemia. *Clinical Chemistry*, **49**, 1125-1132.
<https://doi.org/10.1373/49.7.1125>
- [25] Liu, T., Bao, C.F. and Ren, Y.L. (2011) Determination of Quality Properties of Soy Sauce by Support Vector Regression Coupled with SW-NIR Spectroscopy. *Chemical Research in Chinese Universities*, **27**, 385-391.
- [26] Qin, Q.Y., Zhao, J.W., Chen, Q.S., Hao, L. and Sun, Z.B. (2012) Rapid Measurement of Antioxidant Activity in Dark Soy Sauce by NIR Spectroscopy Combined with Spectral Intervals Selection and Nonlinear Regression Tools. *Analytical Methods*, **4**, 940-946.
<https://doi.org/10.1039/c2ay05766b>
- [27] Zhao, J.W., Qin, Q.Y., Chen, Q.S. and Hao, L. (2013) Simultaneous Determination of Amino Acid Nitrogen and Total Acid in Soy Sauce Using Near Infrared Spectroscopy Combined with Characteristic Variables Selection. *Food Science and Technology International*, **19**, 305-314. <https://doi.org/10.1177/1082013212452475>
- [28] Qin, O.Y., Chen, Q.S., Zhao, J.W. and Lin, H. (2013) Determination of Amino Acid Nitrogen in Soy Sauce Using Near Infrared Spectroscopy Combined with Charac-

- teristic Variables Selection and Extreme Learning Machine. *Food and Bioprocess Technology*, **6**, 2486-2493. <https://doi.org/10.1007/s11947-012-0936-0>
- [29] Xu, J., Huang, F.R., Li, Y.P., Chen, Z. and Wang, Y. (2015) Rapid Detection of Total Nitrogen Content in Soy Sauce Using NIR Spectroscopy. *Czech Journal of Food Sciences*, **33**, 518-522. <https://doi.org/10.17221/229/2015-CJFS>
- [30] Hu, L.Q., Yin, C.L., Ma, S. and Liu, Z.M. (2019) Vis-NIR Spectroscopy Combined with Wavelengths Selection by PSO Optimization Algorithm for Simultaneous Determination of Four Quality Parameters and Classification of Soy Sauce. *Food Analytical Methods*, **12**, 633-643. <https://doi.org/10.1007/s12161-018-01407-1>
- [31] Natcha, T. and Panmanas, S. (2014) Detection of Adulteration of Soy Sauce by Brine Using Near Infrared Spectroscopy, 2nd International CAFÉ. *Agriculture and Agricultural Science Procedia*, **2**, 212-217.
- [32] Errikson, L., Johansson, E., Kettaneh, W.N., Trygg, J., Wikstrom, C. and Wold, S. (2005) Part I Basic Principles and Applications. In: Antti, H. and Holmes, E., Eds., *Multi- and Megavariate Data Analysis*, Umetrics Academy, Umeå, 263-336.
- [33] Chiang, L.H., Russell, E.L. and Braatz, R.D. (2000) Fault Diagnosis in Chemical Processes Using Fisher Discriminant Analysis, Discriminant Partial Least Squares, and Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, **50**, 243-252.
- [34] Barker, M. and Rayens, W. (2003) Partial Least Squares for Discrimination. *Journal of Chemometrics: A Journal of the Chemometrics Society*, **17**, 166-173. <https://doi.org/10.1002/cem.785>
- [35] Barnes, R.J., Dhanoa, M.S. and Lister, S.J. (1989) Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, **43**, 772-777. <https://doi.org/10.1366/0003702894202201>
- [36] Dhanoa, M.S., Lister, S.J., Sanderson, R. and Barnes, R.J., (1994) The Link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) Transformations of NIR Spectra. *Journal of Near Infrared Spectroscopy*, **2**, 43-47. <https://doi.org/10.1255/jnirs.30>
- [37] Han, L.L., Mao, P.S., Wang, G.X. and Wang, Y.H. (2008) Study on Vigour Test of Oat Seeds with Near Infrared Reflectance Spectroscopy. *Journal of Infrared and Millimeter Waves*, **27**, 86-90.