

Research Article

Discriminating the Geographical Origins of Chinese White Lotus Seeds by Near-Infrared Spectroscopy and Chemometrics

Lu Xu,^{1,2} Chen-Bo Cai,³ Yuan-Bin She,² and Li-Juan Chen¹

¹College of Material and Chemical Engineering, Tongren University, Tongren, Guizhou 554300, China

²College of Chemical Engineering, Zhejiang University of Technology, Hangzhou 310014, China

³College of Chemistry and Life Science, Chuxiong Normal University, Chuxiong 675000, China

Correspondence should be addressed to Lu Xu; lxchemo@163.com and Yuan-Bin She; sheyb@zjut.edu.cn

Received 26 March 2015; Accepted 28 May 2015

Academic Editor: Maria Carmen Yebra-Biurrun

Copyright © 2015 Lu Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The traceability of a Chinese white lotus seed (WLS) with Protected Designation of Origin (PDO) was investigated using near-infrared (NIR) spectroscopy and chemometrics. Three chemometrics methods, discrimination analysis (DA), class modeling, and a newly proposed strategy, the fusion of DA and class modeling, were investigated to compare their capacity to trace the geographical origins of WLS. Least squares support vector machine (LS-SVM) was developed to distinguish the PDO WLS from non-PDO WLS of four main producing areas. A class modeling technique, one-class partial least squares (OCPLS), was developed only using the data of PDO WLS. By the fusion of LS-SVM and OCPLS, the best prediction sensitivity and specificity were 0.900 and 0.973, respectively. The results indicate that fusion of DA and class modeling can enhance the specificity for detection of non-PDO products. The conclusion is that DA and class modeling should be combined for tracing food geographical origins.

1. Introduction

Consumers require explicit and accurate information to make informed choices about their diet and the foods they buy. Choice of special foods may reflect lifestyle or religious concerns (e.g., vegetarianism, preference for organic products or products with Protected Denomination of Origin (PDO), and absence of pork for Jews and Muslims) or the health concerns (e.g., preference for some functional foods and absence of certain foods that may cause allergies) [1–3]. Because modern food industries provide a lot of processed foods and it is usually difficult to distinguish specific ingredients or the origins of foods by the naked eye, it is important for the producers and sellers to make honest and accurate description and labeling of their products. However, it is economically profitable for producers and sellers to make food adulterations and frauds, for example, replacing or diluting high-cost ingredients with cheaper ones and purposely mislabeling non-PDO products as a PDO product [4, 5]. Therefore, rapid discrimination methods are required

to distinguish the adulterations and frauds from the authentic foods.

Pattern recognition methods have been widely used for food classification and discrimination. The most frequently used pattern recognition methods in food analysis are classification or discrimination analysis (DA) and class modeling techniques (CMTs) [6–9]. Both DA and CMTs can learn from the labeled or known objects and predict the future objects. However, to tackle the problem of food adulterations and frauds, both DA and CMTs have encountered some difficulties [10–12]. DA models are trained from two or more known classes, so their applications should be limited to predictions of objects from the predefined classes; otherwise, if a new object comes from an untrained class, the prediction would be unreliable or wrong [13]. In contrast, CMTs are trained using only the data of one class (e.g., PDO food to be controlled) and can answer the question of whether a future object is from the target class or not. Because CMTs do not use any information concerning fraud or adulterated objects, they cannot ensure the specificity for detecting

various adulterations or frauds. Therefore, for tracing the geographical origins of PDO foods, it is more reasonable to combine DA and CMTs rather than relying on one of them as traditionally performed.

Lotus (*Nelumbo nucifera* Gaertn.) is a perennial aquatic plant in the monotypic family Nelumbonaceae. It is widely cultivated in China, India, Thailand, South Korea, and Japan, as well as the US [14, 15]. Recent researches have demonstrated that lotus seeds contain a variety of bioactive components (e.g., flavonoids, phospholipids, proteins, amino acids, vitamins, sugars, unsaturated fatty acids, and essential minerals) [16, 17] and its extract displays significant antipyretic, cooling, astringent, antioxidant, and cytoprotective effects and demulcent properties [18–20]. In China, lotus seeds are a valuable functional food and are used for soups, congee, and many other dishes. The bitter dried germ of lotus seeds is also used as a restorative herbal tea. Chemical compositions and quality of lotus seeds from different producing areas are quite different and have been investigated and compared extensively [21]. However, few researches have been devoted to the characterization and discrimination of the geographical origins of Chinese lotus seeds.

Near-infrared (NIR) spectroscopy has been successfully applied to quality control and monitoring of various food products [22]. Compared with traditional chemical analysis methods, NIR analysis has some advantages: (1) reduced sample treatment, analysis time, and cost; (2) the feasibility for nondestructive and online analysis; and (3) simultaneous analysis or characterization of multicomponents. Therefore, NIR spectroscopy is a convenient and economic tool for analyzing the large number of WLS samples from the farmer's markets and small retailers. This work aimed at developing a rapid and nondestructive method for distinguishing a Chinese PDO WLS using NIR spectroscopy. A new strategy was supposed by combining the traditionally used classification and class modeling methods. The performance of the new strategy on improving the detection specificity was investigated by comparing it with separate classification and class modeling methods.

2. Materials and Methods

2.1. Collection of Samples. A set of 109 representative samples of the PDO Jianning-WLS were collected from the original producing area (Jianning, Fujian) and 120 non-PDO WLS objects were collected from five main producing areas, namely, Jiangsu (26), Hunan (25), Hubei (27), Jiangxi (20), and Zhejiang (22). All the samples were harvested in 2013 and were kept with intact packaging in a cool and dark place before NIR analysis.

2.2. NIR Analysis. The NIR spectra were measured with the bare WLS kernels using the diffuse reflectance principle by a Bruker-TENSOR37 FTIR spectrometer (Bruker Optics, Ettlingen, Germany). A fiber probe was used to illuminate a kernel and collect the scattered light. The probe directly contacted with equatorial region of a kernel. Considering the internal composition of a kernel can be varied in different

parts, the spectrum was obtained as the average of three spectra scanned at different positions by rotating a kernel manually. Each spectrum was the average of 64 scans, and more scans did not reduce the noise significantly. The working range of the spectrometer was 4000–12000 cm^{-1} and the scanning interval was 1.929 cm^{-1} with a resolution of 4 cm^{-1} , so each raw spectrum had 4148 wavelengths. The working temperature was kept at 25°C and the humidity was kept at a steady level for the spectrometer. To reduce the influence of possible background shifts during measurement, the NIR analysis for all the objects was performed randomly and the internal gold background was measured as the reference for every hour.

2.3. Preprocessing, Outlier Diagnosis, and Data Splitting. All the data preprocessing and chemometrics computations were performed on Matlab 7.0.1 (MathWorks, Sherborn, MA). Different options were investigated to optimize data preprocessing. Smoothing can reduce random errors in the data and enhance the signal-to-noise ratio (SNR). The algorithm of least squares fitting by Savitzky and Golay (S-G) [23] was used for smoothing considering its effectiveness and simplicity. Taking second-order derivative (D2) spectra can enhance spectral resolution and remove linear baselines, so taking D2 spectra was also applied. Because direct differencing can degrade the SNR by inflating noise, the D2 spectra were also computed using the S-G algorithm. Standard normal variate (SNV) [24] was performed to reduce the influence of scattering effects and path variations caused by the rough surfaces of kernels.

For both DA and CMTs, outliers in the training set will cause significant errors of parameter estimations and even breakdown of the models, while outliers in the test set would make the prediction results unreliable for evaluation of model performance. Outliers can be caused by many factors and it is not trivial to detect high-dimensional outliers. Considering the high-dimensional nature of NIR spectra and the possible masking effects caused by the coexistence of multioutliers, the robust Stahel-Donoho estimate (SDE) of outlyingness [25] was adopted for outlier detection for the PDO and non-PDO WLS samples. This method projects each high-dimensional data point onto randomly generated unit vectors for many times (e.g., 500 or 1000). Because outliers tend to deviate from the bulk of the normal data, their random projections also tend to deviate those of the normal objects. The robust SDE outlyingness is defined based on the median absolute deviation (MAD) of projections.

After removal of outliers, the DUPLEX algorithm [26] was used to divide the measured spectral data into two representative sets: one for training and the other for validation. DUPLEX proceeds as follows: (1) select the two objects with the largest Euclidian distance and put them in the training set; (2) take the two farthest points among the remaining points as testing objects; (3) repeat steps (1) and (2) until one has had enough objects in the test set; and (4) put all the remaining objects in the training set. By alternatively selecting the two farthest points for the training set and test set, the test set and the training set will have a nearly equal distribution. Because

the distributions of objects from different producing areas were heterogeneous, DUPLEX was performed separately on the objects from each producing area.

2.4. Chemometric Models. Support vector machine (SVM) has been proved to be a robust and effective tool to tackle both classification and regression problems. It has had many successful applications in food analysis [27]. Least squares support vector machine (LS-SVM) [28] can be seen as a simplified and fast version of SVM algorithms. Unlike the traditional SVM algorithms, which try to solve a quadratic programming problem, LS-SVM obtains the solution by solving a set of linear equations. Therefore, LS-SVM is much faster than the traditional SVM algorithms. In this paper, a kernel transformation using the Gaussian function was performed to model nonlinear relationship considering the complexity of multiclass data.

The recently proposed one-class partial least squares (OCPLS) [12, 29] was used as a class modeling technique. OCPLS can be performed as a special PLS regression relating the noncentered spectral features and a responses vector of **1** (with all elements being ones). Two types of distances can be derived from the developed OCPLS model, namely, score distance (SD) and absolute value of the centered model residual (ACR). SD is based on the Mahalanobis distance computed using the significant OCPLS components. ACR can measure the distance from an object to the fitted OCPLS model. In terms of the computed SD and ACR, four types of objects can be defined: regular or normal objects (with small SD and small ACR), good leverage objects (with large SD and small ACR), class outliers (with small SD and large ACR), and bad leverage objects (with large SD and large ACR). For detection of non-PDO objects, good leverage objects, class outliers, and bad leverage objects can be detected as different types of outliers [30] as each of them has one or/and two distance measures different from the regular objects. Monte Carlo cross validation (MCCV) [31] was used to select the number of primary OCPLS latent variables based on the model residuals.

To take advantage of the benefits of DA and CMTs, a strategy of models fusion was proposed. Suppose the PDO WLS as the target class; by separately using DA or CMTs, a new object will be predicted as PDO or non-PDO objects. The rule of models fusion is shown in Table 1, which can be summarized as that a new object can be accepted as a PDO object only when it is accepted by both DA and class models. When the predictions of DA and class modeling are inconsistent as for cases 2 and 3, the new object will be predicted as non-PDO. For case 2, a new object is accepted by the class modeling but rejected by the DA model. It is very likely that the object is a non-PDO object and is overlapped with the PDO objects if the DA model has a high accuracy for discrimination. In case 3, the new object is rejected by the class model but accepted by the DA model. The most possible reason for this result is that the object is from a non-PDO class other than all the non-PDO classes used for training the DA model if the class modeling has a good sensitivity. The

TABLE 1: Fusion of discrimination analysis (DA) and class modeling for discrimination of PDO/non-PDO WLS.

	Class modeling	DA	Combined model
Case 1	Acceptance ^a	Acceptance	Acceptance
Case 2	Acceptance	Rejection	Rejection
Case 3	Rejection	Acceptance	Rejection
Case 4	Rejection	Rejection	Rejection

^aA new object is predicted as PDO (acceptance) or non-PDO (rejection) object.

classification of PDO and non-PDO by models fusion is also shown in Figure 1.

2.5. Comparison of Models. Denote the PDO WLS as “positives” and all the non-PDO WLS as “negatives”; sensitivity can describe the model ability to correctly identify the positives and specificity can reflect the model capacity to correctly predict the negatives [32]. In this work, both sensitivity (Se) and specificity (Sp) were considered to compare different models and data preprocessing. The definitions of Se and Sp are

$$\begin{aligned} \text{Se} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{Sp} &= \frac{\text{TN}}{\text{TN} + \text{FP}}, \end{aligned} \quad (1)$$

where TP, FN, TN, and FP are the numbers of true positives, false negatives, true negatives, and false positives, respectively.

3. Results and Discussions

The spectral interval of 9000–12000 cm^{-1} was contaminated with significant background shifts, so this interval was not used for further data analysis. The raw NIR spectra (4000–9000 cm^{-1}) of the PDO and non-PDO WLS are demonstrated in Figure 2. Due to peak overlapping, it is difficult to perform accurate assignments of the peaks. As seen from Figure 2, the spectra of PDO and non-PDO objects have very similar absorbance patterns in the range of 4000–9000 cm^{-1} . However, the spectra of the non-PDO objects have much more variations, most of which may be caused by scattering effects and the roughness of kernel surfaces. Therefore, proper data preprocessing is necessary to remove unwanted spectral variations in the raw data.

Figure 3 shows the preprocessed spectra of PDO and non-PDO WLS objects. For S-G smoothing, the polynomial order was 2 and the window width was 15 considering the narrow scanning interval (1.929 cm^{-1}). For D2 spectra, with the polynomial order of 4, different window widths were tried and a window of 19 obtained a good signal-to-noise ratio. As seen from Figure 3, taking D2 spectra can enhance peak resolution and obtain some detailed information. By SNV transformation, a large part of the variations caused by scattering and surface roughness was removed for the non-PDO objects.

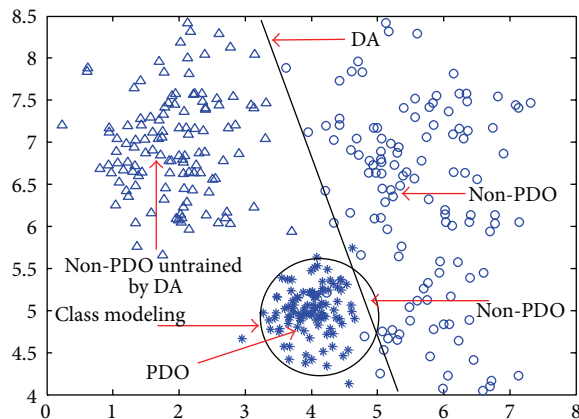


FIGURE 1: The discrimination of PDO and non-PDO samples by fusion of DA and class modeling.

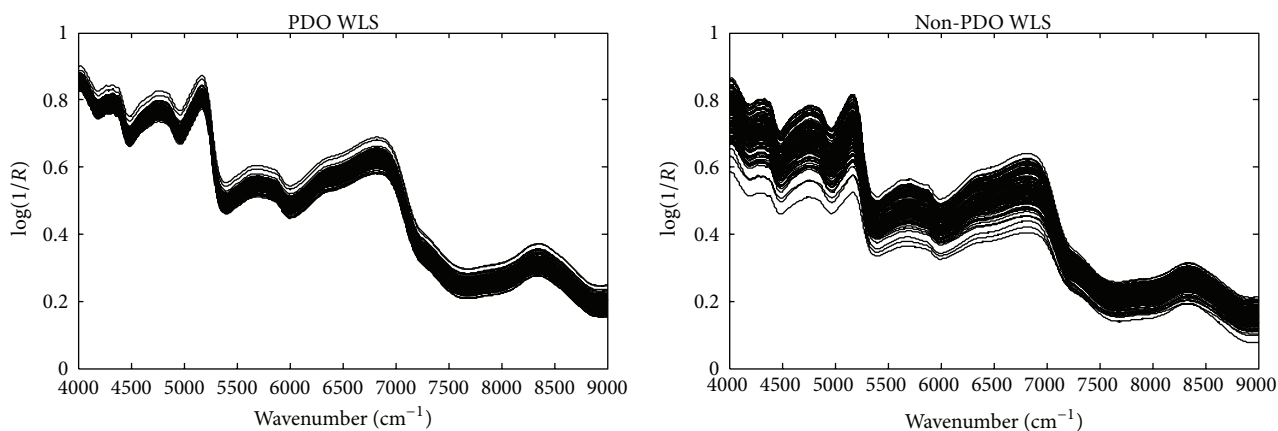


FIGURE 2: The raw NIR spectra of PDO and non-PDO WLS objects.

Outlier detection and data splitting were performed on the raw data (4000–9000 cm^{-1}). Because the distributions of WLS from different producing areas were different, the SDE outlier diagnosis was performed separately on each group (objects from the same producing area). For SDE, the number of projections in this work was 1000. According to the $3\text{-}\sigma$ rule, a SDE value over 3 will indicate an outlier. As a result, 5 and 6 outliers were detected and excluded from the Jianning-WLS and Jiangsu-WLS. The outliers can be attributed to the scattering effects caused by rough surfaces of kernels. DUPLEX method was then used to divide each group into two sets: one for training and the other for prediction. The data splitting is shown in Table 2.

With different data preprocessing, LS-SVM models were developed using the data of 74 PDO and 77 non-PDO objects, and OCPLS models were developed using only the 74 PDO objects. To compare the performance of models, the same test set containing 30 PDO and 37 non-PDO objects was used for prediction. MCCV was used to estimate the modeling errors of OCPLS and the number of significant components. The original training set was randomly divided for 100 times and, considering the size of training set was not very large, each time 20% of the training objects were left out for prediction

TABLE 2: Splitting of the data for training and prediction.

Producing areas	Total objects ^a	For training	For prediction
Fujian (PDO)	104	74	30
Hubei	31	21	10
Jiangsu	24	16	8
Hunan	33	22	11
Zhejiang	26	18	8

^aOutliers have been removed.

in MCCV. An economic OCPLS model was selected if including more components cannot significantly reduce the predicted residual sum of squares (PRESS) value. A 10-fold cross validation was used to simultaneously optimize the two model parameters of LS-SVM, kernel width parameter (σ) and the regularization parameter (γ), by minimizing the classification error rates of MCCV.

With different preprocessing methods, the optimized model parameters and prediction results of LS-SVM, OCPLS, and models fusion were summarized in Table 3. Spectral smoothing cannot significantly improve the classification performance, while D2 and SNV can enhance the model

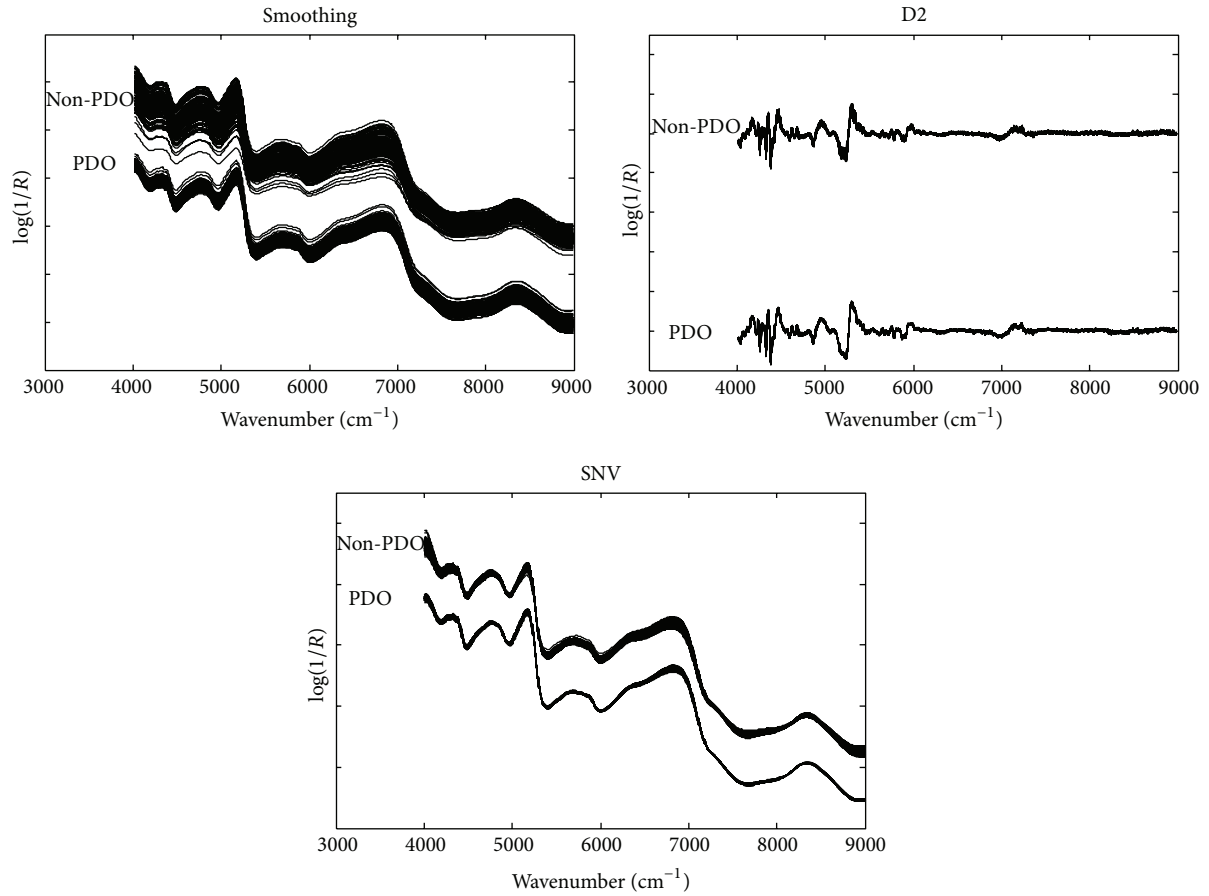


FIGURE 3: The smoothed, D2, and SNV spectra of PDO and non-PDO WLS objects. An extra shift of $\log(1/R)$ was added to distinguish PDO and non-PDO objects.

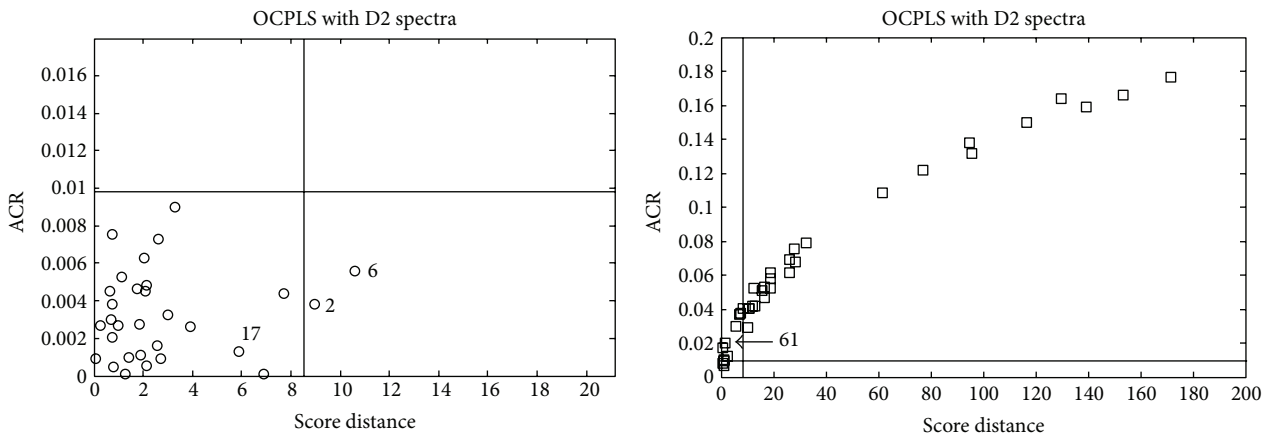


FIGURE 4: Prediction results of D2-OCPLS for PDO (objects 1–30, \circ) and non-PDO (objects 31–67, \square) objects.

specificity of OCPLS for identifying non-PDO objects, indicating that the scattering effects and baseline shifts contribute a large part of the unwanted variations in the raw data. For OCPLS, both D2 and SNV preprocessing can obtain good sensitivity for prediction of PDO objects. However, the specificity of OCPLS was not satisfying. By LS-SVM, relatively high sensitivity and specificity were obtained by

all the models. The best classification results (sensitivity of 0.900 and specificity of 0.973) obtained by models fusion were the same as those of LS-SVM with D2 spectra. The predictions of OCPLS and LS-SVM with D2 spectra were shown in Figures 4 and 5. For predictions of PDO objects (positives), OCPLS had 2 positives wrongly rejected (objects 2 and 6) and LS-SVM had 3 false negatives (objects 2, 6, and 17),

TABLE 3: Model parameters and prediction results of LS-SVM and OCPLS.

Data pretreatment	OCPLS			$\log(\gamma), \log(\sigma)$	LS-SVM		Models fusion	
	LVs ^a	Se ^b	Sp ^c		Se ^a	Sp	Se	Sp
Raw data	4	0.800 (24/30)	0.541 (20/37)	18.04, 14.64	0.867 (26/30)	0.919 (34/37)	0.800 (24/30)	0.919 (34/37)
Smoothing	4	0.833 (25/30)	0.541 (20/37)	18.26, 14.88	0.867 (26/30)	0.919 (34/37)	0.833 (25/30)	0.946 (35/37)
D2	3	0.933 (28/30)	0.703 (26/37)	11.12, 9.41	0.900 (27/30)	0.973 (36/37)	0.900 (27/30)	0.973 (36/37)
SNV	1	0.900 (27/30)	0.595 (22/37)	12.28, 11.90	0.900 (27/30)	0.919 (34/37)	0.900 (27/30)	0.946 (35/37)

^aNumber of OCPLS components.

^bSe: sensitivity; the numbers in the brackets indicate TP/(TP + FN).

^cSp: specificity; the numbers in the brackets indicate TN/(TN + FP).

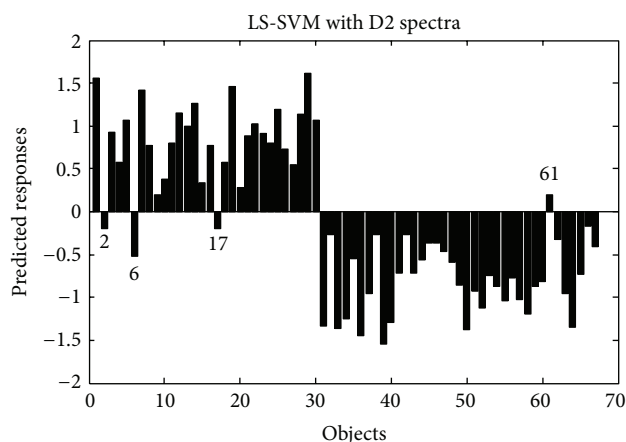


FIGURE 5: Prediction results of LS-SVM with D2 spectra. Objects 1–30 are PDO WLS and 31–67 are non-PDO objects.

so the fusion of models also had 3 false negatives according to Table 1 and its sensitivity (0.900) was slightly lower than that of OCPLS (0.933) and the same as that of LS-SVM. For predictions of non-PDO objects (negatives), OCPLS had 11 negatives wrongly accepted and LS-SVM had 1 false positive (object 61), so the fusion of models had 1 false positive and its model specificity was 0.973. Although the models fusion seems to obtain no improvements over LS-SVM for this data, one cannot conclude that fusion of DA and class modeling is meaningless. Theoretically, as seen from Table 1, fusion of DA and class modeling will have a sensitivity value not higher than that of the class modeling because it always rejects no less objects than the class modeling. However, it can always obtain a specificity value not lower than DA model for the same reason. As pointed out in Section 2.4, the above results demonstrate that if the class modeling can obtain a good sensitivity and the DA model can have a good classification accuracy, the fusion of DA and class modeling can improve the specificity of class modeling without losing much sensitivity. The above conclusion is easy to understand because fusion of models takes advantage of the information

concerning both positive (PDO) objects and negative (non-PDO) objects. For case 3 as in Table 1, it is more reasonable to use models fusion rather than DA.

4. Conclusions

The fusion of DA and class modeling was proposed to trace the geographical origins of a Chinese WLS by NIR spectroscopy. With proper data preprocessing like SNV and D2 to remove the variations caused by scattering effects and the roughness of kernel surface, NIR spectroscopy was demonstrated to be a suitable tool for rapid analysis of WLS kernels. The comparison of different modeling strategies demonstrates that in desirable conditions fusion of DA and class modeling can significantly enhance model specificity without loss of much sensitivity. Our future research will be focused on untargeted analysis of food adulterations by the proposed models fusion strategy.

Highlights

- (i) NIR was used for tracing geographical origins of WLS.
- (ii) Fusion of CMTs and DA was suggested.
- (iii) The proposed strategy can improve model specificity.

Conflict of Interests

All authors declare no conflict of interests.

Authors' Contribution

Lu Xu and Chen-Bo Cai equally contributed to this work.

Acknowledgments

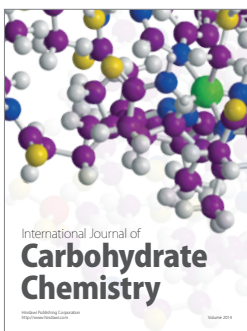
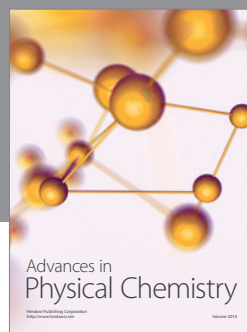
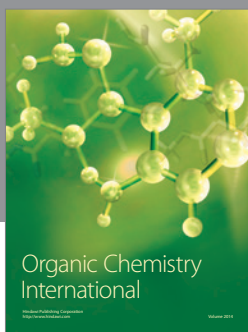
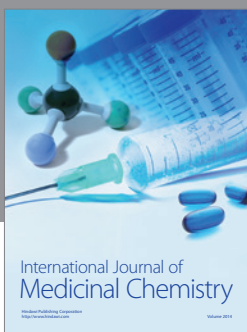
Yuan-Bin She is grateful to the financial support from the General and Youth Projects of National Natural Science of China (Grants nos. 21476270, 21276006, and 21205145).

Lu Xu is supported by the Open Research Program (no. GCTKF2014007) of State Key Laboratory Breeding Base of Green Chemistry Synthesis Technology (Zhejiang University of Technology) and the Research Fund for the Doctoral Program of Tongren University (no. trxyDH1501).

References

- [1] M. Woolfe and S. Primrose, "Food forensics: using DNA technology to combat misdescription and fraud," *Trends in Biotechnology*, vol. 22, no. 5, pp. 222–226, 2004.
- [2] K. Sharma and M. Paradakar, "The melamine adulteration scandal," *Food Security*, vol. 2, no. 1, pp. 97–107, 2010.
- [3] J. C. Moore, J. Spink, and M. Lipp, "Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010," *Journal of Food Science*, vol. 77, no. 4, pp. R118–R126, 2012.
- [4] C. Fernandez, C. Astier, E. Rock, J.-B. Coulon, and J.-L. Berdagué, "Characterization of milk by analysis of its terpene fractions," *International Journal of Food Science and Technology*, vol. 38, no. 4, pp. 445–451, 2003.
- [5] R. Karoui, G. Mazerolles, and É. Dufour, "Spectroscopic techniques coupled with chemometric tools for structure and texture determinations in dairy products," *International Dairy Journal*, vol. 13, no. 8, pp. 607–620, 2003.
- [6] D. Cozzolino, H. E. Smyth, and M. Gishen, "Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins," *Journal of Agricultural and Food Chemistry*, vol. 51, no. 26, pp. 7703–7708, 2003.
- [7] D. Toher, G. Downey, and T. B. Murphy, "A comparison of model-based and regression classification techniques applied to near infrared spectroscopic data in food authentication studies," *Chemometrics and Intelligent Laboratory Systems*, vol. 89, no. 2, pp. 102–115, 2007.
- [8] M. Casale, C. Casolino, G. Ferrari, and M. Forina, "Near infrared spectroscopy and class modelling techniques for the geographical authentication of Ligurian extra virgin olive oil," *Journal of Near Infrared Spectroscopy*, vol. 16, no. 1, pp. 39–47, 2008.
- [9] P. Oliveri, V. di Egidio, T. Woodcock, and G. Downey, "Application of class-modelling techniques to near infrared data for food authentication purposes," *Food Chemistry*, vol. 125, no. 4, pp. 1450–1456, 2011.
- [10] M. Forina, P. Oliveri, S. Lanteri, and M. Casale, "Class-modeling techniques, classic and new, for old and new problems," *Chemometrics and Intelligent Laboratory Systems*, vol. 93, no. 2, pp. 132–148, 2008.
- [11] R. Karoui, G. Downey, and C. Blecker, "Mid-infrared spectroscopy coupled with chemometrics: a tool for the analysis of intact food systems and the exploration of their molecular structure-quality relationships-A review," *Chemical Reviews*, vol. 110, no. 10, pp. 6144–6168, 2010.
- [12] L. Xu, P.-T. Shi, Z.-H. Ye, S.-M. Yan, and X.-P. Yu, "Rapid analysis of adulterations in Chinese lotus root powder (LRP) by near-infrared (NIR) spectroscopy coupled with chemometric class modeling techniques," *Food Chemistry*, vol. 141, no. 3, pp. 2434–2439, 2013.
- [13] L. Xu, Z.-H. Ye, S.-M. Yan et al., "Combining local wavelength information and ensemble learning to enhance the specificity of class modeling techniques: identification of food geographical origins and adulteration," *Analytica Chimica Acta*, vol. 754, pp. 31–38, 2012.
- [14] J. Sung, J.-S. Sung, and H.-S. Shin, "Cytoprotective effects of lotus (*Nelumbo nucifera* Gaertner) seed extracts on oxidative damaged mouse embryonic fibroblast cell," *Food Science and Biotechnology*, vol. 20, no. 6, pp. 1533–1537, 2011.
- [15] J. Man, J. Cai, C. Cai, B. Xu, H. Huai, and C. Wei, "Comparison of physicochemical properties of starches from seed and rhizome of lotus," *Carbohydrate Polymers*, vol. 88, no. 2, pp. 676–683, 2012.
- [16] K. R. Sridhar and R. Bhat, "Nutraceutical potential of lotus seeds," *International Journal of Agricultural Technology*, vol. 3, pp. 143–155, 2007.
- [17] P. K. Mukherjee, D. Mukherjee, A. K. Maji, S. Rai, and M. Heinrich, "The sacred lotus (*Nelumbo nucifera*)—phytochemical and therapeutic profile," *Journal of Pharmacy and Pharmacology*, vol. 61, no. 4, pp. 407–422, 2009.
- [18] D.-H. Sohn, Y.-C. Kim, S.-H. Oh, E.-J. Park, X. Li, and B.-H. Lee, "Hepatoprotective and free radical scavenging effects of *Nelumbo nucifera*," *Phytomedicine*, vol. 10, no. 2-3, pp. 165–169, 2003.
- [19] C. P. Liu, W. J. Tsai, Y. L. Lin, J. F. Liao, C. F. Chen, and Y. C. Kuo, "The extracts from *Nelumbo nucifera* suppress cell cycle progression, cytokine genes expression, and cell proliferation in human peripheral blood mononuclear cells," *Life Sciences*, vol. 75, no. 6, pp. 699–716, 2004.
- [20] Z.-Q. Ling, B.-J. Xie, and E.-L. Yang, "Isolation, characterization, and determination of antioxidative activity of oligomeric procyanidins from the seedpod of *Nelumbo nucifera* Gaertn," *Journal of Agricultural and Food Chemistry*, vol. 53, no. 7, pp. 2441–2445, 2005.
- [21] Y.-B. Zheng, J.-Z. Wu, and W.-X. Lin, "Quality evaluation system of Lotus seeds by clustering analysis," *Science and Technology of Food Industry*, vol. 27, pp. 66–69, 2006.
- [22] M. M. Said, S. Gibbons, A. C. Moffat, and M. Zloh, "Rapid detection of sildenafil analogue in *Eurycoma longifolia* products using a new two-tier procedure of the near infrared (NIR) spectra database," *Food Chemistry*, vol. 158, pp. 296–301, 2014.
- [23] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.
- [24] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Applied Spectroscopy*, vol. 43, no. 5, pp. 772–777, 1989.
- [25] S. van Aelst, E. Vandervieren, and G. Willems, "A Stahel-Donoho estimator based on huberized outlyingness," *Computational Statistics & Data Analysis*, vol. 56, no. 3, pp. 531–542, 2012.
- [26] R. D. Snee, "Validation of regression models: methods and examples," *Technometrics*, vol. 19, no. 4, pp. 415–428, 1977.
- [27] V. G. Uarrota, R. Moresco, B. Coelho et al., "Metabolomics combined with chemometric tools (PCA, HCA, PLS-DA and SVM) for screening cassava (*Manihot esculenta* Crantz) roots during postharvest physiological deterioration," *Food Chemistry*, vol. 161, pp. 67–78, 2014.
- [28] L. Salguero-Chaparro and F. Peña-Rodríguez, "On-line versus off-line NIRS analysis of intact olives," *LWT—Food Science and Technology*, vol. 56, no. 2, pp. 363–369, 2014.
- [29] L. Xu, S.-M. Yan, C.-B. Cai, and X.-P. Yu, "One-class partial least squares (OCPLS) classifier," *Chemometrics and Intelligent Laboratory Systems*, vol. 126, pp. 1–5, 2013.

- [30] M. Hubert, P. J. Rousseeuw, and S. Verboven, "A fast method for robust principal components with applications to chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 60, no. 1-2, pp. 101-111, 2002.
- [31] Q.-S. Xu and Y.-Z. Liang, "Monte Carlo cross validation," *Chemometrics and Intelligent Laboratory Systems*, vol. 56, no. 1, pp. 1-11, 2001.
- [32] M. Forina, C. Armanino, R. Leardi, and G. Drava, "A class-modelling technique based on potential functions," *Journal of Chemometrics*, vol. 5, no. 5, pp. 435-453, 1991.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

