

## Research Article

# Protected Geographical Indication Identification of a Chinese Green Tea (Anji-White) by Near-Infrared Spectroscopy and Chemometric Class Modeling Techniques

Lu Xu,<sup>1</sup> Peng-Tao Shi,<sup>1</sup> Xian-Shu Fu,<sup>1</sup> Hai-Feng Cui,<sup>1</sup> Zi-Hong Ye,<sup>1</sup> Chen-Bo Cai,<sup>2</sup> and Xiao-Ping Yu<sup>1</sup>

<sup>1</sup> Zhejiang Provincial Key Laboratory of Biometrology and Inspection & Quarantine, College of Life Sciences, China Jiliang University, Hangzhou 310018, China

<sup>2</sup> Department of Chemistry and Life Science, Chuxiong Normal University, Chuxiong 675000, China

Correspondence should be addressed to Zi-Hong Ye; [zhye@cjlu.edu.cn](mailto:zhye@cjlu.edu.cn) and Xiao-Ping Yu; [yxp@cjlu.edu.cn](mailto:yxp@cjlu.edu.cn)

Received 13 June 2012; Accepted 3 September 2012

Academic Editor: Tomasz Ruman

Copyright © 2013 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.

This paper reports a rapid identification method for a Chinese green tea with PGI, Anji-white tea, by class modeling techniques and NIR spectroscopy. 167 real and representative Anji-white tea samples were collected from 8 tea plantations in their original producing areas for model training. Another 81 non-Anji-white tea samples of similar appearance were collected from 7 important tea producing areas and used for validation of model specificity. Diffuse NIR spectra were measured with finely ground tea powders. OCPLS and SIMCA were used to describe the distribution of representative Anji-white tea objects and predict the authenticity of new objects. For data preprocessing, smoothing, derivatives, and SNV were applied to improve the raw spectra and classification performance. It is demonstrated that taking derivatives and SNV can improve classification accuracy and reduce the complexity of class models by removing spectral background and baseline. For the best models, the sensitivity and specificity were 0.886 and 0.951 for OCPLS, 0.886 and 0.938 for SIMCA with SNV spectra, respectively. Although it is difficult to perform an exhaustive analysis of all types of potential false objects, the proposed method can detect most of the important non-Anji-white teas in the Chinese market.

## 1. Introduction

As one of the most popular beverages in the world, tea is favored because of its pleasurable aroma, taste, and putative healthy effects [1–7]. Teas can be generally grouped into three principal types: unfermented, partially fermented, and fully fermented according to the degree of fermentation [8].

China has a long history of tea cultivation, processing, and consumption. Among various species, green tea accounts for the bulk of the total production and is favored by most Chinese consumers. In China, tea producing areas cover most of the central and southern provinces with vast differences in geographical and natural conditions, where the tea species,

cultivation techniques, and processing procedures are different. Therefore, almost all of the famous teas in China are named after their origins. Among the most famous teas, Anji-white tea with protected geographical indication (PGI) has a somewhat confusing name because it is a typical green tea. It is called a “white” tea because its leaves are very light in color due to its low chlorophyll and polyphenol contents [9]. Its processing procedure makes it a green tea by withering, pan firing, and shaping, followed by firing gently over charcoal. This specific variety of tea bush reported in ancient literatures was rediscovered growing wild in the 1980s at an altitude over 800 m. Now it is cultivated in the mountains of Anji County, along the Huangpu River

in a spectacular area, where there are heavy mist and vast forests of bright green bamboo. The flat and straight leaves produce a lasting fragrance unique in taste. It is recognized by traditional tea-tasting specialists that the high quality of Anji-white tea should be attributed to its species, growing environments, and processing procedure. Therefore, the PGI authentication of Anji-white tea is demanded to identify false products and ensure consumer interests. Numerous researches have contributed to the chemical compositions of teas influenced by various factors [10–15], including species, season, age of the leaves (plucking position), climate, and horticultural conditions (soil, water, minerals, fertilizers, etc.). Such investigations are important to understand the biological and healthy effects of teas but usually lack a comprehensive view of chemical compositions. Actually, because the chemical compositions of teas are very complex, it is very difficult to perform a thorough component analysis of teas and represent the quality/class of teas by the contents of a few chemical components. In traditional sensory analysis, the quality of teas is evaluated by professional tea tasters. Because training a qualified tea taster may take years and is very expensive [16], it is suitable to evaluate tea quality by some instrumental techniques.

Recently, spectroscopy coupled with chemometric methods has been widely applied in food analysis [17–22]. The principle of such techniques is that chemical compositions of complicated samples are represented by multivariate spectral signals; then relevant and useful information concerning food quality/parameters can be extracted by multivariate analysis methods. Near infrared (NIR) spectroscopy has been one of the most commonly used spectroscopic techniques in food quality evaluation and has some advantages over traditional chemical analysis methods, including lower sample preparation requirements, reduced analysis time and cost, the ability to simultaneous multicomponent analysis, and the potential use for online analysis [23].

The performance of NIR spectroscopy analysis depends heavily on the proper use of chemometric methods. It has been pointed out that PGI authentication is a typical one-class problem [24], where a decision needs to be made on whether a new object should be accepted or rejected by a target class. In such cases, the commonly used classification methods discriminating two or more predefined classes are unsuitable for several reasons. Firstly, PGI identification requires identifying various unknown false objects, which is difficult to be exhaustively collected and analyzed. Moreover, a discrimination/classification model would be highly complex and have poor generalization performance if it includes many different classes of training objects. Therefore, a class model is required to describe the representative samples belonging to the target class and predict the identities of the unknown objects. A class model aims at describing the distribution of a target class and has reduced model complexity. However, the sensitivity and specificity of a class model should be sufficiently validated to ensure its usefulness. The sampling procedure should be representative and comprehensive to include most if not all of the significant variations likely to be encountered in future test materials [18].

TABLE 1: Analyzed tea samples.

Code	Producing area	Class <sup>b</sup>
A1 (22 <sup>a</sup> )	Changlong king tea plantation, Anji, Zhejiang	A
A2 (20)	Damingshan tea plantation, Anji, Zhejiang	A
A3 (21)	Xucunwan tea plantation, Anji, Zhejiang	A
A4 (21)	Xilongshankelai tea plantation, Anji, Zhejiang	A
A5 (20)	Huangpujiangyuan tea plantation, Anji, Zhejiang	A
A6 (20)	Yuqingshan tea plantation, Anji, Zhejiang	A
A7 (22)	Dipu tea plantation, Anji, Zhejiang	A
A8 (21)	Dalongshan tea plantation, Anji, Zhejiang	A
N1 (9)	Guangshan tea plantation, Xinyang, Henan	N
N2 (11)	Subu tea plantation, Luan, Anhui	N
N3 (10)	Gongchayuan tea plantation, Zhenghe, Fujian	N
N4 (14)	Tianmuhu tea plantation, Liyang, Jiangsu	N
N5 (10)	Baiyun tea plantation, Jingan, Jiangxi	N
N6 (11)	Jinggu tea plantation, Puer, Yunnan	N
N7 (16)	Xianrantang tea plantation, Fuding, Fujian	N

<sup>a</sup> Number of tea samples from the same tea producing areas.

<sup>b</sup> A: Anji-white teas; N: non-Anji-white teas.

With the above considerations, the objective of this paper is to develop a rapid and well-validated PGI authentication method for Anji-white tea by using class modeling techniques and NIR spectroscopy, with emphasis on representative sample collection and validation of class models.

## 2. Materials and Methods

**2.1. Tea Samples Analyzed.** 167 authentic Anji-white and 81 other main non-Anji-white tea samples were collected directly from the market branches of tea plantations in original producing areas with official certifications. All of the samples were made of green tea leaves picked before Qingming Festival 2011 (April 5, 2011). The detailed information concerning samples is shown in Table 1. All of the samples were stored in a cool, dark, and dry place with integral packaging before spectroscopic analysis.

**2.2. FTIR Spectroscopy.** Diffuse NIR spectra were collected using a Bruker-TENSOR37 FTIR spectrometer (Bruker Optics, Ettlingen, Germany) in the wavelength range from  $4000^{-1}$  to  $12000\text{ cm}^{-1}$ . Tea samples were finely grounded into particles using an agate pestle and mortar and filtered through a 40-mesh sieve. The powders were then packed into a NIR sample cup. The sample cup was filled fully and compacted naturally without further pressing. For each sample, 128 scans were performed with a resolution of  $8\text{ cm}^{-1}$  at room temperature using OPUS software. An increase in scanning time did not significantly improve the signal. The average of the 128 scans was used as a raw spectrum for further data analysis. The scanning interval was  $3.857\text{ cm}^{-1}$ ; therefore, each raw spectrum had 2074 individual data points.

**2.3. Outlier Diagnosis and Data Preprocessing.** All of the data analysis was performed on MATLAB 7.0.1 (Mathworks,

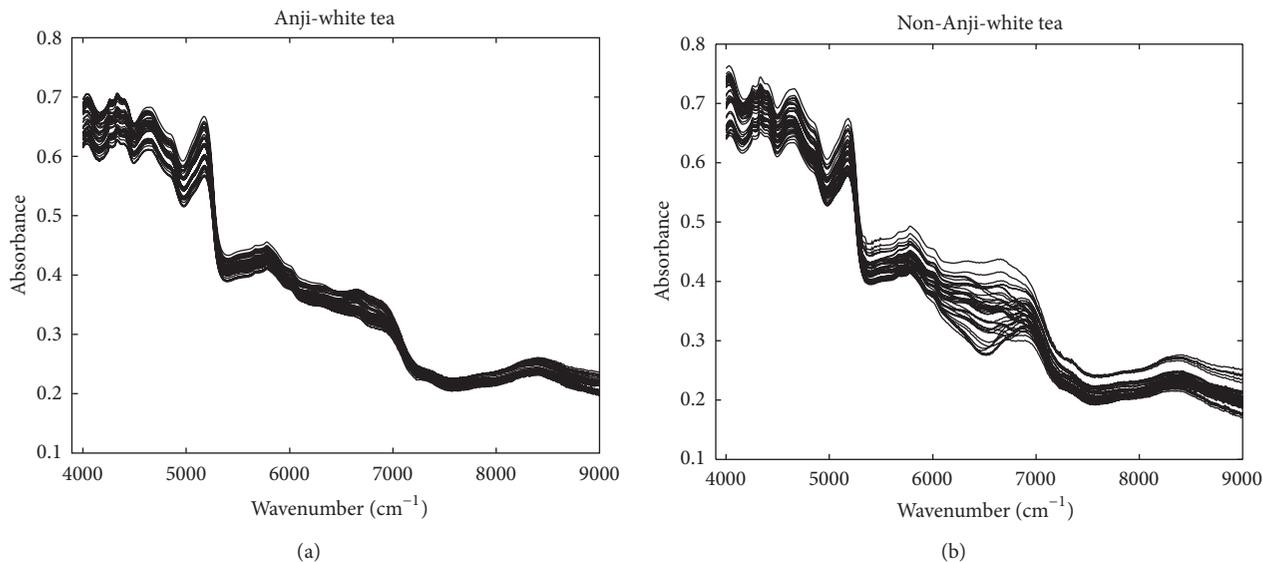


FIGURE 1: Some of the NIR spectra of real Anji-white (a) and non-Anji-white (b) tea samples.

Sherborn, MA, USA). In practical data analysis, outliers in the data would cause model bias or even breakdown of the models. Therefore, robust principal component analysis (rPCA) [25] was used to detect the outliers. rPCA can overcome the masking effects caused by the presence of multiple outliers. Considering the high-dimensional nature of the NIR spectral data (for the raw spectra,  $P = 2074$ ), an improved rPCA [26] was used, which was shown to be more numerically stable for high-dimensional data and have a moderate computational cost. According to the computed score distance (SD) and orthogonal distance (OD), an rPCA diagnosis plot classifies the samples into four groups: regular data (with small SD and small OD), good PCA-leverage points (with large SD and small OD), orthogonal outliers (with small SD and large OD), and bad PCA-leverage points (with large SD and large OD).

The data with outliers removed were then split into a representative training set and a test set by the Kennard and Stone (K-S) algorithm [27]. K-S algorithm selects a representative test set in such a way that the objects are scattered uniformly in the range of training objects. Because the distributions of tea samples from each producing area were not the same, the K-S method was performed separately for teas from different producing areas. For class models analysis, the training and test samples from each producing area selected by K-S algorithm were then put together to form a training and test set.

Smoothing, taking derivatives, and standard normal variate (SNV) [28] were used to improve the training and predicting performance of class models. Smoothing can suppress random noise in spectra and improve the signal-to-noise ratio (SNR). The S-G polynomial fitting algorithm [29] was used considering its popularity and simplicity. Taking derivatives can enhance spectral resolution and remove

baseline and background, so first-order and second-order derivatives were used. To prevent the degradation of SNR by differencing, derivatives were also computed by S-G algorithms. SNV was proved to be effective in reducing scattering effects and correcting the interference caused by variations of optical path. In this paper, SNV was used to reduce the spectral variations caused by the possible differences of powder bulk density.

**2.4. Class Modeling Techniques.** Recently, a new class modeling technique was proposed by us using one-class partial least squares (OCPLS) [30] regression. It was used for authentication of pure sesame oils by mid-infrared spectra and was demonstrated to have a comparable performance to soft independent modeling of class analogy (SIMCA) [24]. OCPLS develops a partial least squares regression model relating the features to a class response vector 1 with all the elements being ones. The use of 1 as a response vector means all the objects in the same class should be distributed as compact as possible. Unlike SIMCA which projects the raw variables onto a few principal components (PCs) explaining most of the data variances, OCPLS considers both the explained variances and compactness of a class by projecting the raw features onto the class average. The modeling error of the response variable is assumed to have a normal distribution and used as the distance measurement from an object to the class center. The class center is estimated as the mean of modeling error. Since OCPLS can be performed in the framework of multivariate calibration, estimation of its model complexity is more straightforward than for SIMCA; for example, a well-established  $F$ -test combined with Monte Carlo cross validation (MCCV) [31] was demonstrated to be effective in reducing the risk of overfitting [32].

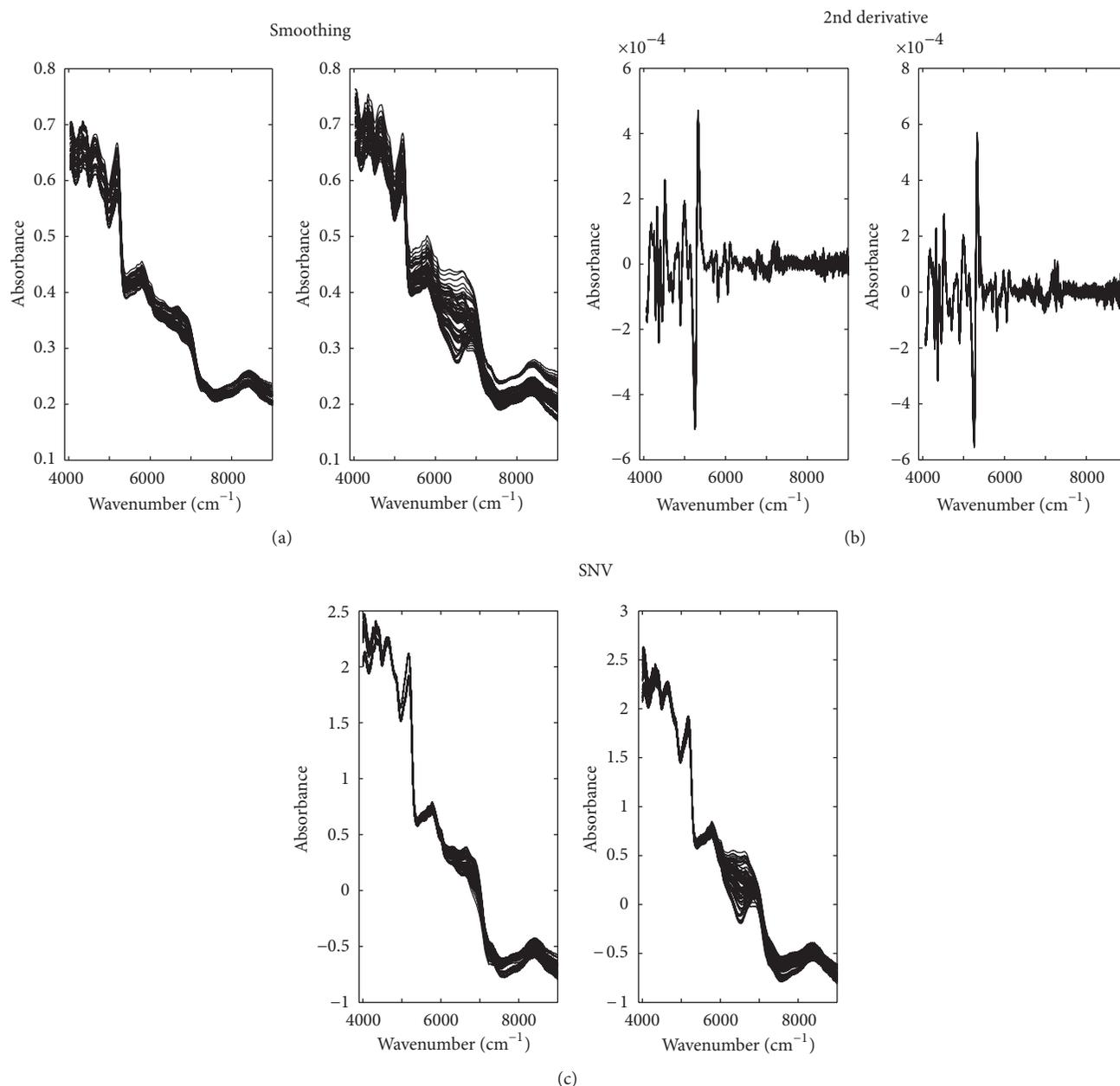


FIGURE 2: Some of the preprocessed spectra of Anji-white (left) and non-Anji white (right) tea samples by smoothing, taking derivatives, and SNV.

SIMCA describes the class structure of the training objects by the PCs space spanned by a few significant PCs. The magnitude of residual error can be used as a distance measurement from an object to the class center. To reject or accept a new object, its residuals can be tested with an  $F$ -test procedure. It was realized that the residual error could be underestimated when it is computed directly from PCA of the training samples. This would lead to a large number of objects that are wrongly rejected (a large  $\alpha$ -error); therefore, residuals predicted by leave-one-out cross-validation (LOOCV) rather than the training residuals were used [33]. This procedure was shown to be effective in reducing the number of false outliers.

### 3. Results and Discussion

Some of the spectra of the authentic Anji-white tea and Non-Anji-white tea samples are demonstrated in Figure 1. The spectral range of  $9000\text{--}12000\text{ cm}^{-1}$  carries poor chemical information and has a very low SNR, so this wavelength range was excluded from further data analysis. Seen from Figure 1, all the teas have very similar absorbance bands in the range of  $4000\text{--}9000\text{ cm}^{-1}$ . The wide bands in  $8000\text{--}9000\text{ cm}^{-1}$  can be attributed to the second overtone of  $\text{C-H}$  stretching. Peaks in  $6000\text{--}7000\text{ cm}^{-1}$  involve the contributions of  $\text{O-H}$  stretching vibrations and stretching vibration of  $\text{N-H}$  ( $\sim 6800\text{ cm}^{-1}$ )

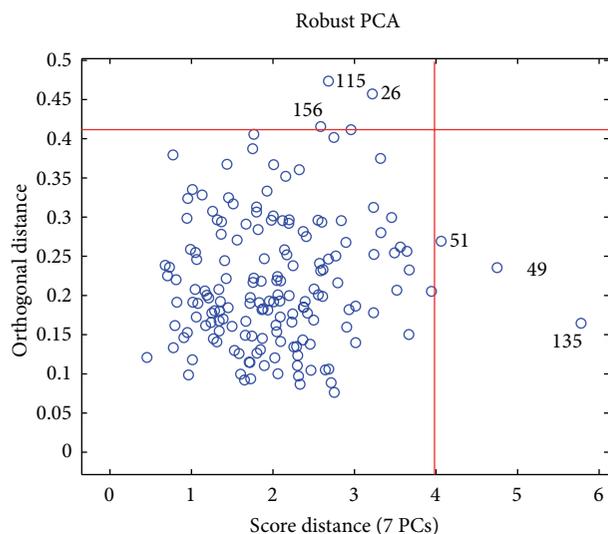


FIGURE 3: Outlier diagnosis plot obtained by robust PCA with 7 principal components of 167 Anji-white tea samples.

in amino acids. Other obvious bands include  $5600\text{ cm}^{-1}$  (fundamental stretching of  $\text{-C-H}$ ),  $5200\text{ cm}^{-1}$  (combination of  $\text{O-H}$  and  $\text{C-O}$  stretching),  $4700\text{ cm}^{-1}$  (combination of  $\text{O-H}$  bending and  $\text{C-O}$  stretching), and  $4300\text{ cm}^{-1}$  (combination of  $\text{C-H}$  stretching and  $\text{-CH}_2$  deformation). The raw spectra are highly overlapped and characterized by a poor peak resolution, so accurate assignments of specific peaks are very difficult.

The low level of details in the raw spectra can be attributed to the contributions of multicomponents and the shifts and distortions resulted from their interactions. Though different tea varieties had very similar absorbance patterns, the relative intensities of different bands were different. Therefore, class modeling techniques are useful to extract the subtle information from spectral data for characterizing real Anji-white teas.

To sharpen the classification performance of class models, smoothing, first- and second-order S-G derivatives, and SNV were used to preprocess the raw spectra. Some of the preprocessed spectra of Anji-white and non-Anji-white teas are plotted in Figure 2. Seen from Figure 2, although smoothed spectra can slightly improve the SNR of the raw spectra, they have the risk of losing some useful high-frequency information in the raw data. The second-order derivative spectra can remove most of the baselines and enhance some detailed information. SNV spectra can reduce some spectral variations while enhancing others. The actual effects of data preprocessing should be evaluated by classification performance.

Outlier detection was performed by rPCA of the raw spectra. The number of PCs was determined by robust pooled predicted residual sum of squares (PRESS) values. Following the rule of thumb, the first seven PCs were selected as to account for 95.32% (>95%) of the total data variance. The rPCA diagnosis plot of the 167 authentic Anji-white tea samples is shown in Figure 3. OD is a measure of the distance from the sample to the model space spanned by selected PCs,

and SD describes the sample dispersion in the class projected onto the model space. Therefore, both orthogonal outliers and bad PCA-leverage points should be excluded from the training set. Because the real Anji-white tea samples came from different producing areas, there might be considerable difference in the contents of different chemical components. Therefore, good PCA-leverage points (objects 49, 51 and 135) were retained to represent the spectral variations among the real Anji-white tea samples from different producing areas. In Figure 3, three orthogonal outliers (objects 26, 115, and 156) were removed. The K-S algorithm was then used to split the remaining 164 Anji-white samples into a training set with 120 samples and a test set of 44 samples. Therefore, the test set had 44 positive (Anji) objects and 81 negative (non-Anji) objects.

OCPLS and SIMCA models were developed to describe the distribution of real Anji-white tea samples. For SIMCA, the improved decision region [33] was adopted to reduce the risk of having a large number of objects wrongly rejected. Cross-validation was performed to evaluate the number of significant PCs; the criterion of 95% total explained variances was also considered. For OCPLS, Monte Carlo cross-validation (MCCV) with 10% objects left out was used to determine the number of PLS components and the sampling time was 100. The PRESS values by MCCV were subject to an  $F$ -test [32]. As suggested, a significant level of 0.25 was adopted to select the least number of latent variables with a PRESS value not significantly larger than the minimum value according to the  $F$ -test. Sensitivity and specificity were used to evaluate the performance of different models and preprocessing options. The training and prediction results of test samples by SIMCA and OCPLS were shown in Table 2. Seen from Table 2, preprocessing generally improved the classification performance in terms of sensitivity and specificity. However, the models based on spectra smoothing seem to have inferior performance, which might be attributed to the possible loss of detailed frequency information. Second derivative and SNV significantly improved the class models by reducing the baseline and backgrounds. The model complexity of SIMCA and OCPLS based on such preprocessing was reduced. For the best models, the sensitivity and specificity were 0.886 and 0.951 for OCPLS and 0.886 and 0.938 for SIMCA with SNV spectra, respectively. For both OCPLS and SIMCA, the best class models were obtained by SNV preprocessing and the prediction results were demonstrated in Figures 4 and 5. The comparison of different preprocessing methods demonstrated that the spectral variations caused by scattering effects and baseline shifts played a more important role than an inferior SNR.

## 4. Conclusions

Rapid identification methods of a PGI green tea, Anji-white tea, were developed by using NIR spectroscopy and chemometric class modeling techniques. With SNV preprocessing, OCPLS (sensitivity 0.886 and specificity 0.951) and SIMCA (sensitivity 0.886 and specificity 0.938) achieved best classification performance in terms of prediction sensitivity and specificity. The analysis results indicate removal of spectral

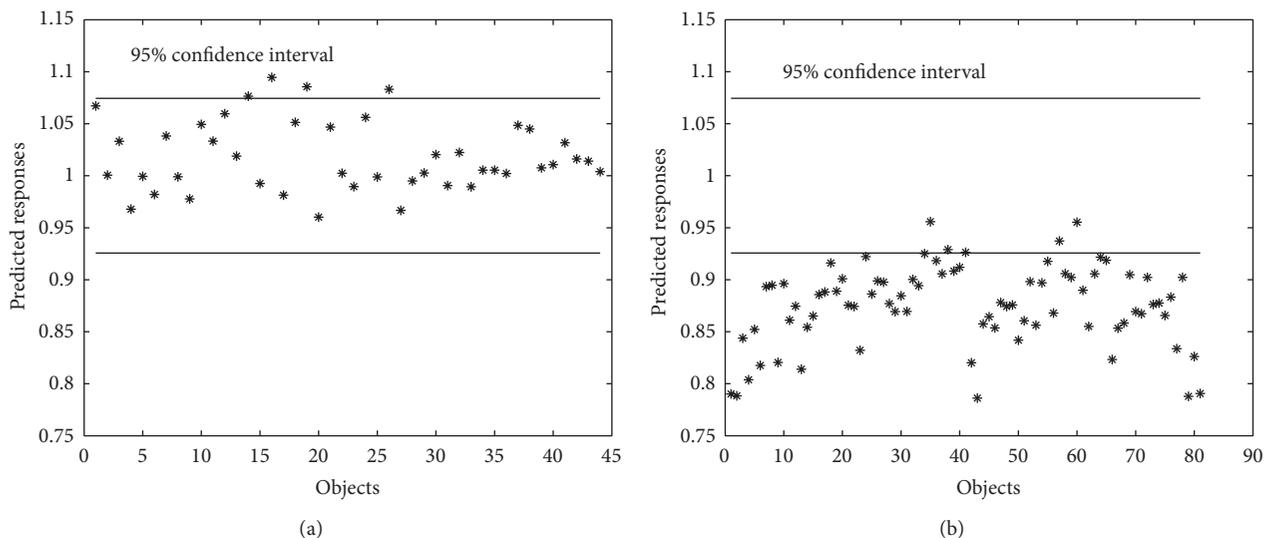


FIGURE 4: Results obtained by SNV-OCPLS with 8 latent variables for (a) 44 positive test objects, and (b) 81 negative test objects.

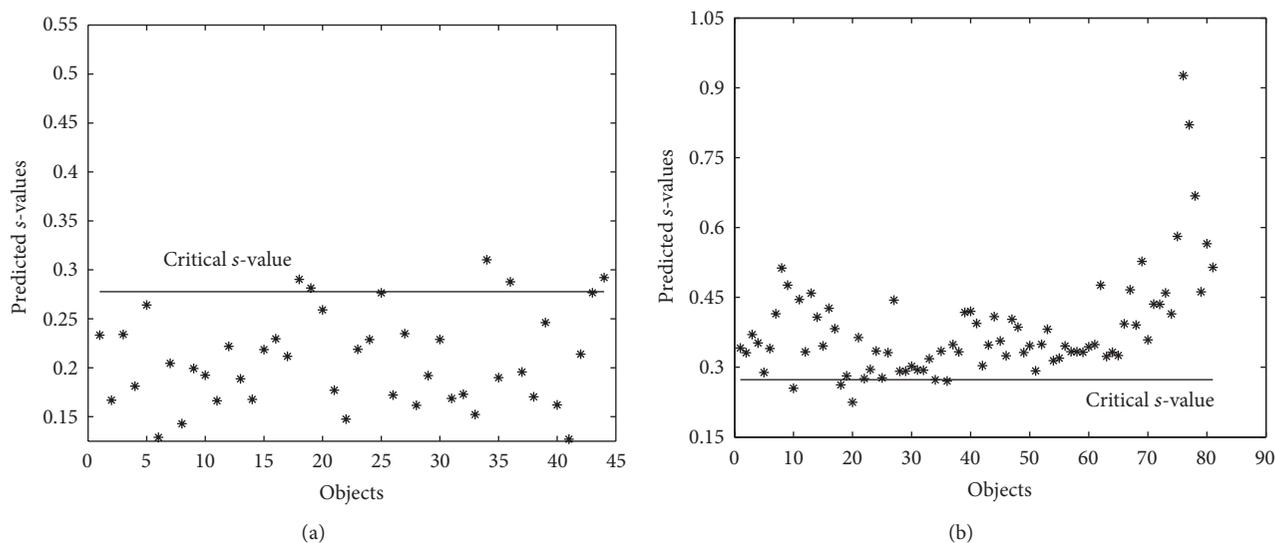


FIGURE 5: Results obtained by SNV-SIMCA with 7 principal components for (a) 44 positive test objects, and (b) 81 negative test objects.

TABLE 2: Predicting results obtained by OCPLS and SIMCA.

Model	Data treatment	TP <sup>b</sup>	FN	TN	FP	Sensitivity	Specificity
OCPLS	Raw data (11 <sup>a</sup> )	34	10	68	13	0.773	0.840
	Smoothing (10)	36	8	65	16	0.818	0.803
	2nd derivative (9)	37	7	73	8	0.841	0.901
	SNV (8)	39	5	77	4	0.886	0.951
SIMCA	Raw data (10)	35	9	70	11	0.796	0.864
	Smoothing (8)	33	11	68	13	0.750	0.840
	2nd derivative (9)	36	8	74	7	0.818	0.914
	SNV (7)	39	5	76	5	0.886	0.938

<sup>a</sup> Numbers of OCPLS or SIMCA latent variables.

<sup>b</sup> TP: number of true positives; FN: number of false negatives; TN: number of true negatives; FP: number of false positives. "Positive" and "negative" represent Anji- and non-Anji-white tea, respectively.

background, and baseline plays a more important role than a higher SNR. Taking derivatives and SNV transformation can not only improve classification accuracy but can also reduce the complexity of OCPLS and SIMCA models. Although it is hard to perform an exhaustive collection and analysis of all types of white teas, this study provides a reliable and effective tool to identify Anji-white tea against most of the important non-Anji-white teas in the Chinese market.

## Authors' Contribution

Lu Xu and Peng-Tao Shi equally contributed to this work.

## Acknowledgments

The authors appreciate the National Public Welfare Industry Projects of China (nos. 201210010 and 201210092) and Hangzhou Programs for Agricultural Science and Technology Development (no. 20101032B28) for their financial aids. C.-B. Cai is also grateful to the financial aid of the Applied and Basic Research Project of Yunnan Provincial Science and Technology Department (no. 2010CD087).

## References

- [1] C. Dufresne and E. Farnworth, "Tea, Kombucha, and health: a review," *Food Research International*, vol. 33, no. 6, pp. 409–421, 2000.
- [2] N. Khan and H. Mukhtar, "Tea polyphenols for health promotion," *Life Sciences*, vol. 81, no. 7, pp. 519–533, 2007.
- [3] S. T. Saito, G. Gosmann, C. Pungartnik, and M. Brendel, "Green tea extract—patents and diversity of uses," *Recent Patents on Food, Nutrition & Agriculture*, vol. 1, no. 3, pp. 203–215, 2009.
- [4] A. B. Sharangi, "Medicinal and therapeutic potentialities of tea (*Camellia sinensis* L.)—a review," *Food Research International*, vol. 42, no. 5–6, pp. 529–535, 2009.
- [5] V. Sharma and L. J. M. Rao, "A thought on the biological activities of black tea," *Critical Reviews in Food Science and Nutrition*, vol. 49, no. 5, pp. 379–404, 2009.
- [6] Y. Wang and C. T. Ho, "Polyphenols chemistry of tea and coffee: a century of progress," *Journal of Agricultural and Food Chemistry*, vol. 57, no. 18, pp. 8109–8114, 2009.
- [7] N. T. Zaveri, "Green tea and its polyphenolic catechins: medicinal uses in cancer and noncancer applications," *Life Sciences*, vol. 78, no. 18, pp. 2073–2080, 2006.
- [8] J. K. Lin, C. L. Lin, Y. C. Liang, S. Y. Lin-Shiau, and I. M. Juan, "Survey of catechins, gallic acid, and methylxanthines in green, oolong, pu-erh, and black teas," *Journal of Agricultural and Food Chemistry*, vol. 46, no. 9, pp. 3635–3642, 1998.
- [9] K. Wei, L. Y. Wang, J. Zhou et al., "Comparison of catechins and purine alkaloids in albino and normal green tea cultivars (*Camellia sinensis* L.) by HPLC," *Food Chemistry*, vol. 130, no. 3, pp. 720–724, 2012.
- [10] Y. L. Lin, I. M. Juan, Y. L. Chen, Y. C. Liang, and J. K. Lin, "Composition of polyphenols in fresh tea leaves and associations of their oxygen-radical-absorbing capacity with antiproliferative actions in fibroblast cells," *Journal of Agricultural and Food Chemistry*, vol. 44, no. 6, pp. 1387–1394, 1996.
- [11] C. M. Peters, R. J. Green, E. M. Janle, and M. G. Ferruzzi, "Formulation with ascorbic acid and sucrose modulates catechin bioavailability from green tea," *Food Research International*, vol. 43, no. 1, pp. 95–102, 2010.
- [12] D. Komes, D. Horžić, A. Belščak, K. K. Ganić, and I. Vulić, "Green tea preparation and its influence on the content of bioactive compounds," *Food Research International*, vol. 43, no. 1, pp. 167–176, 2010.
- [13] L. Bazinet, M. Araya-Farias, A. Doyen, D. Trudel, and B. Têtu, "Effect of process unit operations and long-term storage on catechin contents in EGCG-enriched tea drink," *Food Research International*, vol. 43, no. 6, pp. 1692–1701, 2010.
- [14] T. Karak and R. M. Bhagat, "Trace elements in tea leaves, made tea and tea infusion: a review," *Food Research International*, vol. 43, no. 9, pp. 2234–2252, 2010.
- [15] A. Ghosh, P. Tamuly, N. Bhattacharyya, B. Tudu, N. Gogoi, and R. Bandyopadhyay, "Estimation of theaflavin content in black tea using electronic tongue," *Journal of Food Engineering*, vol. 110, no. 1, pp. 71–79, 2012.
- [16] G. Le Gall, I. J. Colquhoun, and M. Defernez, "Metabolite profiling using  $^1\text{H}$  NMR spectroscopy for quality assessment of green tea, *Camellia sinensis* (L.)," *Journal of Agricultural and Food Chemistry*, vol. 52, no. 4, pp. 692–700, 2004.
- [17] G. Downey, "Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics," *Trends in Analytical Chemistry*, vol. 17, no. 7, pp. 418–424, 1998.
- [18] 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.
- [19] A. Rohman and Y. B. Che Man, "Determination of extra virgin olive oil in quaternary mixture using FTIR spectroscopy and multivariate calibration," *Spectroscopy*, vol. 26, no. 3, pp. 203–211, 2011.
- [20] C. J. Du and D. W. Sun, "Learning techniques used in computer vision for food quality evaluation: a review," *Journal of Food Engineering*, vol. 72, no. 1, pp. 39–55, 2006.
- [21] Q. X. Zhang, Q. B. Li, and G. J. Zhang, "Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection," *Spectroscopy*, vol. 27, no. 2, pp. 93–105, 2012.
- [22] H. Huang, H. Yu, H. Xu, and Y. Ying, "Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: a review," *Journal of Food Engineering*, vol. 87, no. 3, pp. 303–313, 2008.
- [23] R. Karoui and J. De Baerdemaeker, "A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products," *Food Chemistry*, vol. 102, no. 3, pp. 621–640, 2007.
- [24] 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.
- [25] I. Stanimirova, B. Walczak, D. L. Massart, and V. Simeonov, "A comparison between two robust PCA algorithms," *Chemometrics and Intelligent Laboratory Systems*, vol. 71, no. 1, pp. 83–95, 2004.
- [26] 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.
- [27] R. W. Kennard and L. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, no. 1, pp. 137–148, 1969.

- [28] 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.
- [29] 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.
- [30] L. Xu, C. B. Cai, and D. D. Deng, "Multivariate quality control solved by one-class partial least squares regression: identification of adulterated peanut oils by mid-infrared spectroscopy," *Journal of Chemometrics*, vol. 25, no. 10, pp. 568–574, 2011.
- [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] D. M. Haaland and E. V. Thomas, "Partial least-squares methods for spectral analyses—1. Relation to other quantitative calibration methods and the extraction of qualitative information," *Analytical Chemistry*, vol. 60, no. 11, pp. 1193–1202, 1988.
- [33] R. De Maesschalck, A. Candolfi, D. L. Massart, and S. Heuerding, "Decision criteria for soft independent modelling of class analogy applied to near infrared data," *Chemometrics and Intelligent Laboratory Systems*, vol. 47, no. 1, pp. 65–77, 1999.



**Hindawi**

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

