

Research Article

Electronic Tongue Combined with Chemometrics to Provenance Discrimination for a Green Tea (Anji-White Tea)

Si-Min Yan, Zi-Feng Hu, Cheng-Xin Wu, Lu Jin, Gong Chen, Xian-Yu Zeng, and Jia-Qi Zhu

Shanghai Institute of Quality Inspection and Technical Research, Shanghai 201114, China

Correspondence should be addressed to Jia-Qi Zhu; zhujq@sqi.org.cn

Received 14 January 2017; Accepted 28 February 2017; Published 9 May 2017

Academic Editor: Hai-Yan Fu

Copyright © 2017 Si-Min Yan 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 aims to provide a stable instrumental method for provenance discrimination of Anji-White tea by its distinctive taste. 180 authentic and 60 counterfeit white tea samples were collected for specific geographical origins detection; all of them were measured by electronic tongue coupled with 7 independent sensors. Therefore, chemometrics methods, principal component analysis (PCA), and partial least squares discriminant analysis (PLSDA) were performed in classification. The PCA distribution shows that, in provenance analysis, PCA is a simple and reliable tool for small sample sets, but for sets with large objects, PCA seems powerless in classification. Therefore, PLSDA was applied to develop a classification model. The prediction sensitivity and specificity of PLSDA, respectively, reached 0.917 and 0.950. This study demonstrates the potential of combining electronic tongue system and chemometrics as an effective tool for specific geographical origins detection in Anji-White tea.

1. Introduction

Green tea, made from the leaves of the *Camellia sinensis* plant, is one of the most popular beverages in the world. Moreover, it is an excellent source of antioxidants such as polyphenols, polysaccharides, and amino acids, so green tea has been widely consumed as a healthy drink for its preventing functions on obesity, cancer, liver, and cardiovascular diseases [1–3]. Tea plants are widely distributed in over 30 countries and play a significant role in their economies; the property and chemical components of green tea are influenced by many factors, such as climate, fertilization conditions, geographical origins, and processing procedures. Among these factors, the specific geographical origins are universally accepted as an important aspect [4–6]. Therefore, in China, most famous teas are named after their geographical origins, such as the Anji-White tea, the West Lake-Longjing tea, the Anxi-Tieguanyin tea, and the Wuyi-Rock tea.

Anji-White tea (AWT) is produced in Anji County (Zhejiang province, China) and has been awarded the protected geographical indication (PGI). It is called a “white” tea because its leaves are very light in color due to its low chlorophyll and polyphenol contents [7]. AWT has abundant amino acids and its fragrance can last for long time. However,

the actual yield of AWT is limited and can hardly afford the increasing market demand; some merchants fraudulently label “Anji-White tea” indication to non-Anji-White teas (NAWT) for illegal profits. Although these counterfeit products are inferior to the authentic AWT, their appearances are similar and can hardly be distinguished just by naked eyes. Until now, for provenance, the conventional discrimination way is sensory analysis; unknown tea sample was tasted by professional tea tasters; then its quality was described based on a series of sensory scores. However, the result of sensory analysis depends largely on the subjective judgment of the taster, and it is very expensive and time-consuming to train a professional tea taster. Therefore, an urgent demand exists for developing a nonhuman technique to discriminate the geographical origins of white teas.

In recent years, analytical methods based on instruments have been widely reported in food quality control such as wine, milk, and juice [8, 9]. In these researches, electronic tongue sensors possessed good stability and sensitivity. Moreover, a good correlation between human and electronic tongue judgment has been observed, which makes it a promising alternative to human sensory analysis of teas.

This paper was focused on developing an instrumental technique for discriminating the provenance of Anji-White

tea by electronic tongue system and chemometrics, to model the complicated relationship between the taste and the geographical origin; principal component analysis (PCA) and partial least squares discriminant analysis (PLSDA) models were used in classification [10–13].

2. Materials and Methods

2.1. Tea Samples. In this study, 180 authentic AWT samples were collected from local plantations in 30 white tea production sites of Anji County. And 60 NAWT samples were collected from 10 different production sites such as Miaoxi, Jingan, and Guangde. All of the samples were preserved in a cold storage covered with lightproof packaging before analysis. The detailed information about the samples is listed in Table 1.

2.2. Electronic Tongue Analysis. 3 grams of each sample was added to 150 mL boiling distilled water and infused for 10 minutes. Then, the infusion was filtered into a glass beaker and cooled to 25°C in water bath for analysis. An ASTREE II electronic tongue system (Alpha M.O.S., Toulouse, France) coupled with a reference electrode and 7 independent liquid sensors (ZZ, BA, BB, CA, GA, HA, and JB) was performed in instrumental sensory analysis. In order to achieve the reliable data, the instrument was calibrated by standard NaCl (0.01 mol/L), HCl (0.01 mol/L) and Monosodium Glutamate (0.01 mol/L) solutions before the test. The signal acquisition interval was 1 s, the acquisition time was 180 s, and the stable signal exported at 180 s was saved for provenance analysis. The cross-sensitivity and selectivity of the sensor array are displayed in Table 2.

2.3. Principal Component Analysis. The whole data analysis was performed on MATLAB 7.14.0.739 (Mathworks, Sherborn, MA). As the 7 sensors exported the responses, respectively, multivariate statistical model is necessary in classification. Principal component analysis (PCA) is an unsupervised chemometrics tool used for pattern recognition and dimension reduction [14]. The data could be reconstructed by linearly uncorrelated variables (principal components) with considerable information loss. In this paper, 30 AWT samples from 5 provenances were discriminated using PCA model; the first three components were applied, so the objects were distributed in a three-dimensional PCA score figure. Similar, 30 NAWT samples were classified. Then, all of the samples in this study were analyzed by three-dimensional PCA model in the same way.

2.4. Partial Least Squares Discriminant Analysis. Partial Least Squares Discriminant Analysis (PLSDA) has been regarded as one of cornerstones in chemometrics because it can handle the large p (the number of variables in the data) problem very well [15]. In PLSDA, representative samples are necessary for model training, as the math model developed, unknown objects could be sorted. In this paper, both AWT and NAWT samples were divided into a training class and a predicting class. To ensure that samples in predicting class

TABLE 1: Detailed information of the samples.

Number	Producing area	Batch size ^a	Types ^b
A1	Ganxiqiao, Anji, Zhejiang	6	A
A2	Xucunwan, Anji, Zhejiang	6	A
A3	Xilong, Anji, Zhejiang	6	A
A4	Xizhuang, Anji, Zhejiang	6	A
A5	Huangjin, Anji, Zhejiang	6	A
A6	Gaoyu, Anji, Zhejiang	6	A
A7	Liangpeng, Anji, Zhejiang	6	A
A8	Dipu, Anji, Zhejiang	6	A
A9	Zhangjiashan, Anji, Zhejiang	6	A
A10	Dushantou, Anji, Zhejiang	6	A
A11	Xiaoyuan, Anji, Zhejiang	6	A
A12	Sangong, Anji, Zhejiang	6	A
A13	Lujia, Anji, Zhejiang	6	A
A14	Kuntong, Anji, Zhejiang	6	A
A15	Guanyinqiao, Anji, Zhejiang	6	A
A16	Gucheng, Anji, Zhejiang	6	A
A17	Qinglong, Anji, Zhejiang	6	A
A18	Meixi a, Anji, Zhejiang	6	A
A19	Meixi b, Anji, Zhejiang	6	A
A20	Jishan a, Anji, Zhejiang	6	A
A21	Jishan b, Anji, Zhejiang	6	A
A22	Lihuangdu, Anji, Zhejiang	6	A
A23	Shangsu, Anji, Zhejiang	6	A
A24	Waihuangdu, Anji, Zhejiang	6	A
A25	Nanbeizhuang a, Anji, Zhejiang	6	A
A26	Nanbeizhuang b, Anji, Zhejiang	6	A
A27	Ximu, Anji, Zhejiang	6	A
A28	Luxi, Anji, Zhejiang	6	A
A29	Qiankengqiao, Anji, Zhejiang	6	A
A30	Tiangai, Anji, Zhejiang	6	A
N1	Miaoxi, Huzhou, Zhejiang	6	N
N2	Changxing, Huzhou, Zhejiang	6	N
N3	Yuyao, Ningbo, Zhejiang	6	N
N4	Wencheng, Wenzhou, Zhejiang	6	N
N5	Jingan, Yichun, Jiangxi	6	N
N6	Linghuashan, Jiangxi	6	N
N7	Gannan, Jiangxi	6	N
N8	Ziyang, Wuyuan, Jiangxi	6	N
N9	Guangde, Anhui	6	N
N10	Qiyang, Yongzhou, Hunan	6	N

^aThe number of the tea samples from the same production area.

^bA: Anji-White tea; N: non-Anji-White tea.

were distributed uniformly around the training class, the Kennard and Stone (K-S) algorithm was used for division [16]. Afterwards, the AWT training class and the NAWT training class were taken as a total training class. Similarly, the AWT predicting class and the NAWT predicting class

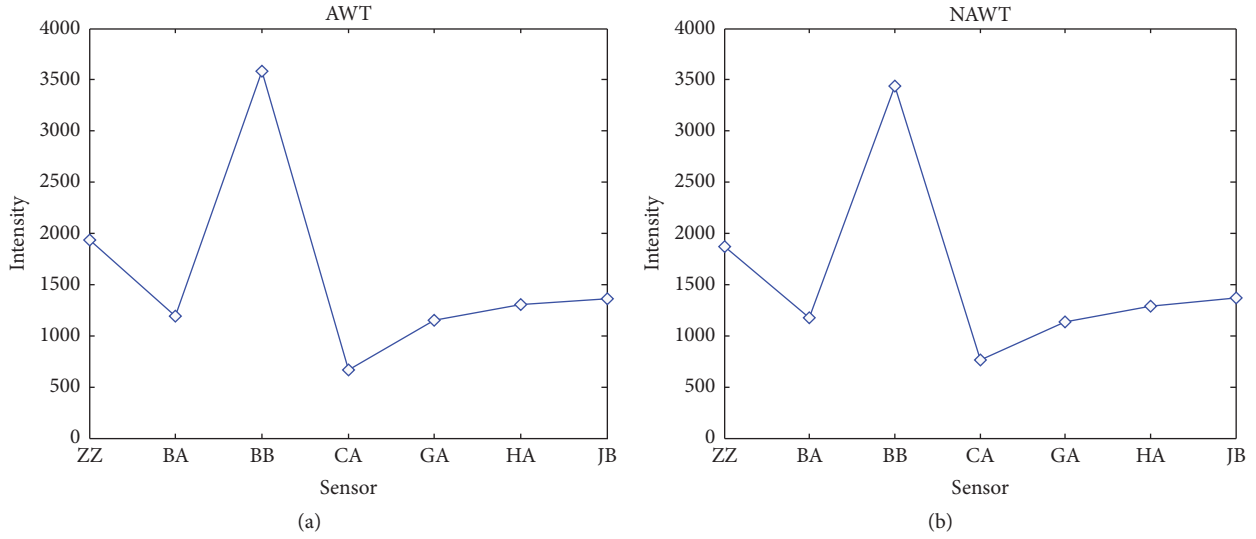


FIGURE 1: The average responses of 7 sensors to AWT and NAWT.

TABLE 2: Sensor sensitivity of the electronic tongue system.

Basic tastes	Substances	ZZ	BA	BB	CA	GA	HA	JB
Sour	Citric acid	7	6	7	7	7	6	6
Salty	KCl	7	4	4	5	4	4	4
Sweet	Glucose	7	4	7	7	4	4	4
Bitter	Caffeine	5	4	4	5	4	4	4
Savory	L-arginine	6	4	6	5	5	4	5

were also taken as a whole. Therefore, the total training class is arranged in $n \times p$ matrix X including p variables with n objects. $n \times 1$ response vector Y demonstrates the corresponding category of each object in matrix X . The value 1 and value -1 are set to denote the AWT and NAWT objects, and the value 0 is the cut-off value. In prediction, an unknown object will be classified into AWT class if its response value is above 0; otherwise, it will be classified into NAWT class.

For PLSDA, latent variables control the complexity of model. Inadequate latent variables will lead to an underfitted model; excessive latent variables will raise the risk of overfitting, so the number of latent variables is an important parameter in classification. Thus, Monte Carlo cross-validation (MCCV) was adopted to optimize the model [17]; the parameter with the minimal misclassification rate of MCCV (MRMCCV) was calculated by following formula:

$$\text{MRMCCV} = \left(\frac{\sum_{i=1}^T E_i}{\sum_{i=1}^T P_i} \right) \times 100\%, \quad (1)$$

where T is the times of splitting, E_i is the number of misclassified objects for the i th splitting, and P_i is the number of total prediction objects.

To evaluate the performance of PLSDA model, sensitivity and specificity were computed [18], the AWT was denoted

as “positive,” and the NAWT was denoted as “negative,” and then sensitivity and specificity were calculated as follows:

$$\text{Sens.} = \frac{TP}{TP + FN}, \quad (2)$$

$$\text{Spec.} = \frac{TN}{TN + FP},$$

where TP, FN, TN, and FP, respectively, represent the number of true positives, false negatives, true negatives, and false positives.

3. Results and Discussions

Average responses of the sensors to AWT and NAWT are showed in Figure 1. As seen from Figure 1, the responses of BA, GA, HA, and JB to AWT and NAWT are very approximate, the AWT objects have stronger intensities for ZZ and BB sensors, and NAWT objects get higher CA signal than AWT objects.

The PCA score results are demonstrated in Figure 2. Three principal components of the data were extracted in analysis. The variance information of three PCA models is listed in Table 3. Seen from Table 3, three PCA models (models A, B, and C) perform well in statistics because all of their cumulative variances exceed 98%. As plotted in Figure 2, the AWT/NAWT samples from 5 different provenances are clearly discriminated in model A/B. But in model C, 180 AWT and 60 NAWT objects overlap with each other; in other words, PCA can hardly satisfy the demand of provenance discrimination.

PLSDA is an effective technique in classification. In this paper, the K-S algorithm was used to obtain the training class and the predicting class; finally, 120 AWT and 40 NAWT samples were picked for training; other 60 AWT and 20 NAWT samples were used for prediction. MCCV was then used to estimate the number of latent variables (from 1 to 7); the training class was randomly split into secondary training

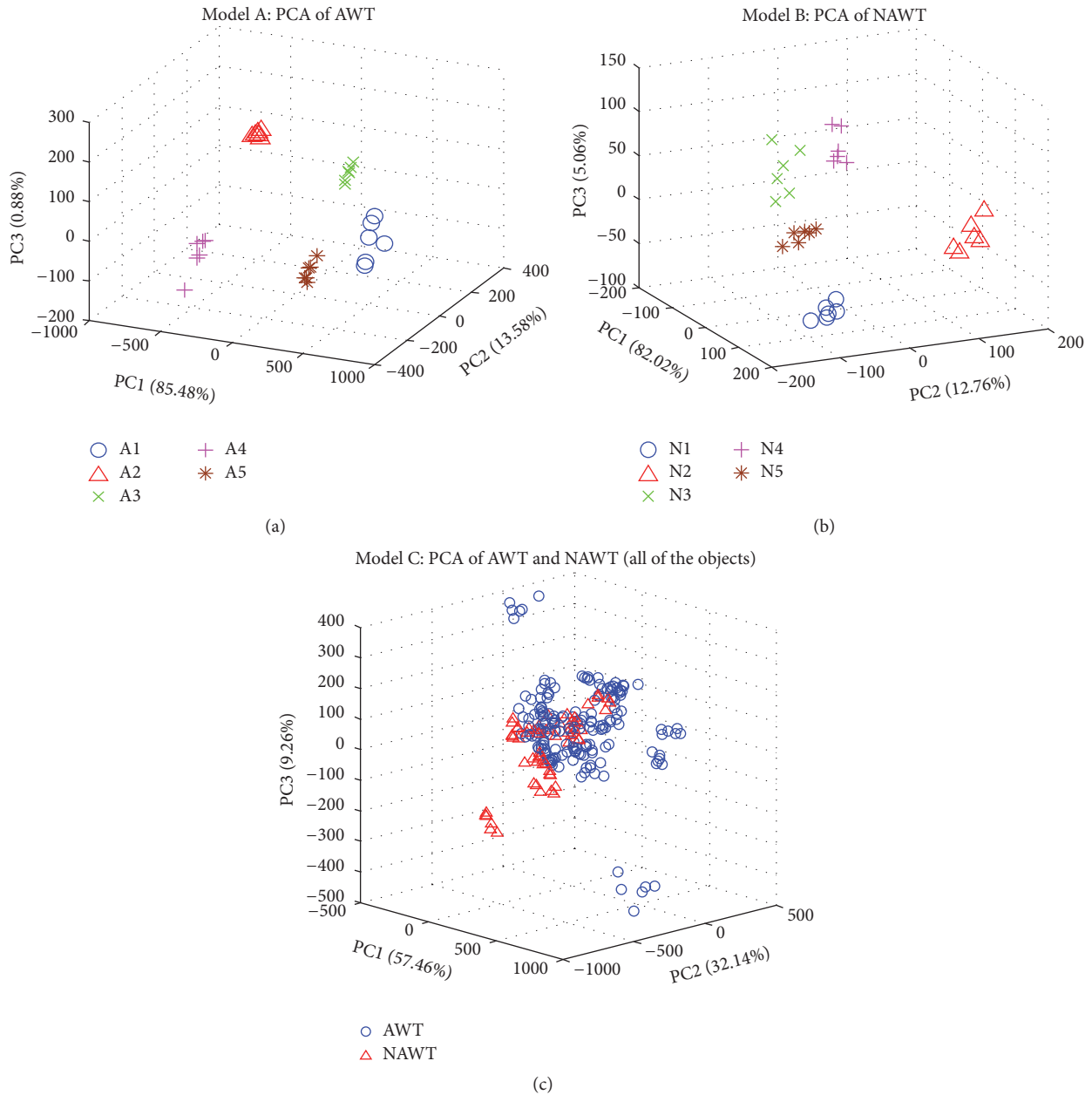


FIGURE 2: PCA scores of the samples in three-dimensional space formed by three principal components. The sample information is listed on the right; colors and marks are used for denotation.

TABLE 3: The variance of each component captured by PCA.

Principal component	Variance (%)		
	Model A	Model B	Model C
PC1	85.48	82.02	57.46
PC2	13.58	12.76	32.14
PC3	0.88	5.06	9.26
Cumulation	99.94	99.84	98.86

class (50%) and secondary predicting class (50%) for 100 times; by calculating and simulating, the minimal MRMCCV value was obtained with a four-component model. Therefore,

a classification model with 4 latent variables was developed and the predicting class was used to evaluate the performance of this model; Figure 3 shows the training and predicting results of the PLSDA model. For training, the sensitivity reached 0.933 (112/120) and the specificity reached 0.950 (38/40). In prediction, the sensitivity of PLSDA model is 0.917 (55/60), and the specificity is 0.950 (19/20); only 5 AWT and 1 NAWT objects are misclassified in prediction.

4. Conclusion

The feasibility of combining electronic tongue and chemometrics for provenance detecting of Anji-White tea was

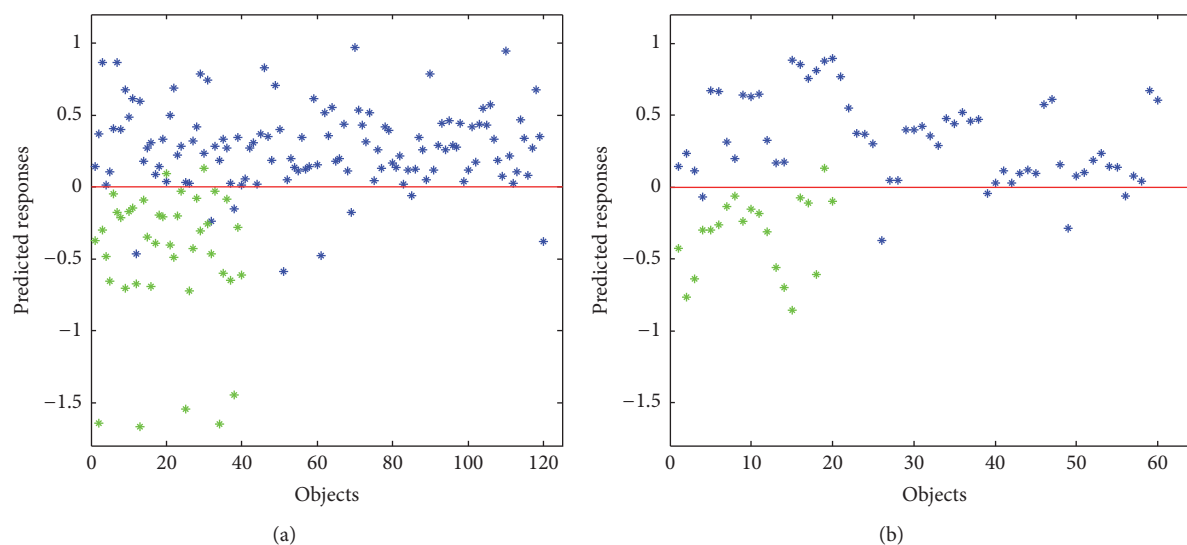


FIGURE 3: Training (a) and predicting (b) results of PLSDA model based on the data of electronic tongue. The AWT objects were marked by blue asterisks and the NAWT objects were denoted by green asterisks.

investigated. Electronic tongue is a rapid and stable instrument, and it takes little sample loss in analysis, so it is suitable for food quality control, especially the taste analysis of tea products. PCA and PLSDA models were performed in classification. The PCA results demonstrate that objects from small sample set could be clearly discriminated by their geographical origins in a three-dimensional PCA score plot. However, PCA is an unsupervised tool; the provenance or other sample features are not involved in separating; consequently, the AWT and NAWT objects overlap seriously as all of the objects are analyzed in a PCA model. PLSDA has a better performance in provenance discrimination than PCA; both the sensitivity and the specificity are satisfying in prediction. But PLSDA is more complicated in modeling, it requires plentiful and representative samples for training, and it takes a long time for model optimization. In our future work, more classification models like artificial neural network and support vector machine will be investigated for provenance discrimination.

Conflicts of Interest

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

References

- [1] A. G. Dulloo, C. Duret, D. Rohrer et al., "Efficacy of a green tea extract rich in catechin polyphenols and caffeine in increasing 24-h energy expenditure and fat oxidation in humans," *The American Journal of Clinical Nutrition*, vol. 70, no. 6, pp. 1040–1045, 1999.
- [2] R. J. Ruch, S.-J. Cheng, and J. E. Klaunig, "Prevention of cytotoxicity and inhibition of intercellular communication by antioxidant catechins isolated from Chinese green tea," *Carcinogenesis*, vol. 10, no. 6, pp. 1003–1008, 1989.
- [3] K. Imai, K. Suga, and K. Nakachi, "Cancer-preventive effects of drinking green tea among a Japanese population," *Preventive Medicine*, vol. 26, no. 6, pp. 769–775, 1997.
- [4] T. S. Pilgrim, R. J. Watling, and K. Grice, "Application of trace element and stable isotope signatures to determine the provenance of tea (*Camelliasinensis*) samples," *Food Chemistry*, vol. 118, no. 4, pp. 921–926, 2010.
- [5] G. Ma, Y. Zhang, J. Zhang et al., "Determining the geographical origin of Chinese green tea by linear discriminant analysis of trace metals and rare earth elements: taking *Dongting Biluochun* as an example," *Food Control*, vol. 59, pp. 714–720, 2016.
- [6] J. Brzezicha-Cirocka, M. Grembecka, and P. Szefer, "Monitoring of essential and heavy metals in green tea from different geographical origins," *Environmental Monitoring and Assessment*, vol. 188, no. 3, article 183, 2016.
- [7] 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.
- [8] E. W. Nery and L. T. Kubota, "Integrated, paper-based potentiometric electronic tongue for the analysis of beer and wine," *Analytica Chimica Acta*, vol. 918, pp. 60–68, 2016.
- [9] O. Benjamin and D. Gamrasni, "Electronic tongue as an objective evaluation method for taste profile of pomegranate juice in comparison with sensory panel and chemical analysis," *Food Analytical Methods*, vol. 9, no. 6, pp. 1726–1735, 2016.
- [10] J. Saavedra, A. Córdova, L. Gálvez, C. Quezada, and R. Navarro, "Principal component analysis as an exploration tool for kinetic modeling of food quality: a case study of a dried apple cluster snack," *Journal of Food Engineering*, vol. 119, no. 2, pp. 229–235, 2013.
- [11] G. Dijksterhuis, M. Flipsen, and P. Punter, "Principal component analysis of TI-curves: Three methods compared," *Food Quality and Preference*, vol. 5, no. 1-2, pp. 121–127, 1994.
- [12] G. Hough, A. N. Califano, N. C. Bertola et al., "Partial least squares correlations between sensory and instrumental measurements of flavor and texture for reggianito grating cheese," *Food Quality and Preference*, vol. 7, no. 1, pp. 47–53, 1996.

- [13] S. M. Yan, J. P. Liu, and L. Xu, "Rapid discrimination of the geographical origins of an Oolong tea (Anxi-Tieguanyin) by near-infrared spectroscopy and partial least squares discriminant analysis," *Journal of Analytical Methods in Chemistry*, vol. 2014, Article ID 704971, 6 pages, 2014.
- [14] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1–3, pp. 37–52, 1987.
- [15] M. Barker and W. Rayens, "Partial least squares for discrimination," *Journal of Chemometrics*, vol. 17, no. 3, pp. 166–173, 2003.
- [16] R. W. Kennard and L. A. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, no. 1, pp. 137–148, 1969.
- [17] Q. S. Xu and Y. Z. Liang, "Monte Carlo cross validation," *Chemometrics and Intelligent Laboratory Systems*, vol. 56, no. 1, pp. 1–11, 2001.
- [18] 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
<https://www.hindawi.com>

