Prediction of Marked Age of Mature Vinegar Based on Fourier Transform Near Infrared Spectroscopy

Zhengguang An[1], Huishan Lu[*1], Huanyu Jiang[2], Yibin Ying[2]

[1] College of Mechanical Engineering & Automatization, North University of China, 3 Xueyuan St., Taiyuan 030051, P. R. China
[2] College of Biosystems Engineering and Food Science, Zhejiang University, 268 Kaixuan St., Hangzhou 310029, P. R. China
464024678@qq.com, 13934597379@139.com, hyjiang@zju.edu.cn, ybying@zju.edu.cn

Abstract. To evaluate the applicability of near infrared (NIR) spectroscopy for discrimination between mature vinegar with different marked age (1 year, 3 years, and 5 years), transmission spectra were collected in the spectral range from 800 nm to 2500 nm in 1, 2, and 5 mm optical path-length rectangular quartz cuvettes with air as reference at room temperature. Discriminant models were developed based on discriminant analysis (DA) together with raw, first and second derivative spectra. The calibration result for raw spectra was better than that for first and second derivative spectra. The percentage of samples correctly classified for raw were 100% for 1 year old, 100% for 3 years old and 97.5% for 5 years old, respectively. In validation analysis, for 1 year, 3 years, and 5 years old sample groups, the percentage of samples correctly classified was 100%, 100%, and 100%, respectively. The results demonstrated that NIR spectroscopy could be used as a rapid and reliable method for classification of mature vinegar with different marked age.

1 Introduction

Fermented mature vinegar, which has a long history in China, is treated as the favorite condiments, health-care products and even medicines by Chinese people. The commercial mature vinegar consists of water, acetic acid, sugars, and other secondary constituents that contribute to the smell, taste and preserving qualities. The composition of mature vinegar is 80% (v/v) of water and a great variety of other compounds like organic acids, alcohols, minerals, polyphenols, amino acids, etc. accounting for the other 20%.

In recent years, near-infrared (NIR) spectroscopy has gained wide acceptance in different fields by virtue of its advantages over other analytical techniques, the most salient of which is its ability to record spectra for solid and liquid samples without any pre-treatment. This characteristic makes it especially attractive for straightforward, speedy characterization of natural and synthetic products. NIR spectroscopy has increasingly been adopted as an analytical tool in various fields,

such as the petrochemical, pharmaceutical, environmental, clinical and biomedical sectors. One of the most common applications of near-infrared spectroscopy combined with pattern recognition methods has been to discriminate between samples belonging to one of several distinct groups based on spectral properties. Some researches had reported the application of NIR spectroscopy for the aging of wine vinegar or Chinese rice wines [1], for discriminate between samples from different origins [2], for the prediction of the quality parameters of wine vinegar or wines [3-6]. However, although it has been proven that original NIR spectra of vinegar samples might be used to develop classification models with good abilities to discriminate between different origins, NIR data, especially from liquid samples, are affected by different types of perturbations such as optical interferences (light scatter), temperature differences and turbidity, making the use of pre-processing methods and inverse calibration essential. As a result of all the foregoing, NIR spectra contain not only chemical but also physical information which may be irrelevant and can mask the chemical information in the spectra (including information closely related to sample origin), and might deteriorate classification models developed from raw NIR spectra. Therefore, the application of suitable chemometrical methods to NIR spectra in order to minimize the contribution of physical effects and thus enhance the chemical information contained therein, could be seen as an important stage in model development and improvement. The main aim of this paper is to study the aging of mature vinegar during storage period.

2 Materials and method

2.1 Samples

In this work, a total of two brands of mature vinegars were obtained in local market named Ninghuafu and Donghu, with different marked age (1 year, 3 years, and 5 years) of each variety. All of these mature vinegars were commonly used in Chinese people's daily life. Before the experiment, the mature vinegar samples were stored in the laboratory at a constant temperature of 25 ± 1 ^{0}C for more than 48h to have an equalization room temperature. The samples were all original vinegar liquid without dilution. A total of 150 samples (25 samples for each age of each variety) were prepared for further treatments. Twenty samples were selected randomly from each age of each variety and a total of 120 mature vinegar samples were used in the calibration set, whereas, 30 samples (5 for each age of each variety) were selected randomly as the validation set from the remaining samples. And the details were listed in Table 1.

Tab. 1. Distribution of mature vinegar samples in the calibration and validation set

Variety	Age (year)	Number of samples	
		Calibration set	Validation set
Donghu	1	20	5
	3	20	5
	5	20	5
Ninghuafu	1	20	5
	3	20	5
	5	20	5

2.2 Spectral Measurements

Samples taken from freshly opened bottles of mature vinegar were scanned in transmission mode (800-2500 nm) using a scanning spectrometer, Nexus FT-NIR (Thermo Nicolet Corp., Madison, WI), with an interferometer, an InGaAs detector, and a broad-band light source (quartz tungsten halogen, 50 W). NIR spectral data were collected using OMNIC software (Thermo Nicolet Corp.) and stored in absorbance format. Samples were scanned in demountable liquid cells of different optical path lengths (1, 2, and 5 mm; Pike Technologies, Madison, WI) with air as the reference at room temperature. All samples were shaken before scanning. The mirror velocity was 0.9494 cm s-1, and the resolution was 2 cm-1. The spectrum of each sample was the average of 32 successive scans. The spectral regions with an absorbance value equal to or higher than 1.5 were not used in spectral analysis due to the zero transmissivity and the fact that they are considered saturated.

2.3 Chemometrics and Data Analysis

Chemometrics analysis was performed using the commercial software package, TQ Analyst v6.2.1 (Thermo Nicolet Corporation, Madison, WI, USA). Discriminant analysis (DA) was used to classify mature vinegar samples with different vinegar age. It was used to determine the classes of known samples which were most similar to an unknown sample by computing the unknown's distance from each class center in Mahalanobis distance.

Principal component analysis (PCA) was performed before DA models were developed. PCA was performed in order to reduce the number of variables showing co-linearity. Thus, the samples were in a new reduction k-dimensional space (k<n). From the k factor score, Mahalanobis distance which indicated how different a sample spectrum was from the average spectrum of the sample set was calculated. It was defined by the following equation:

$$MD_i = \left(t_i - \bar{t}\right)S_k - 1\left(t_i - \bar{t}\right)T$$

Where MD_i is the Mahalanobis distance, t_i is the score vector of ith sample, $\bar{t}$ is the mean score vector of the sample set, S_k is the scores covariance matrix.
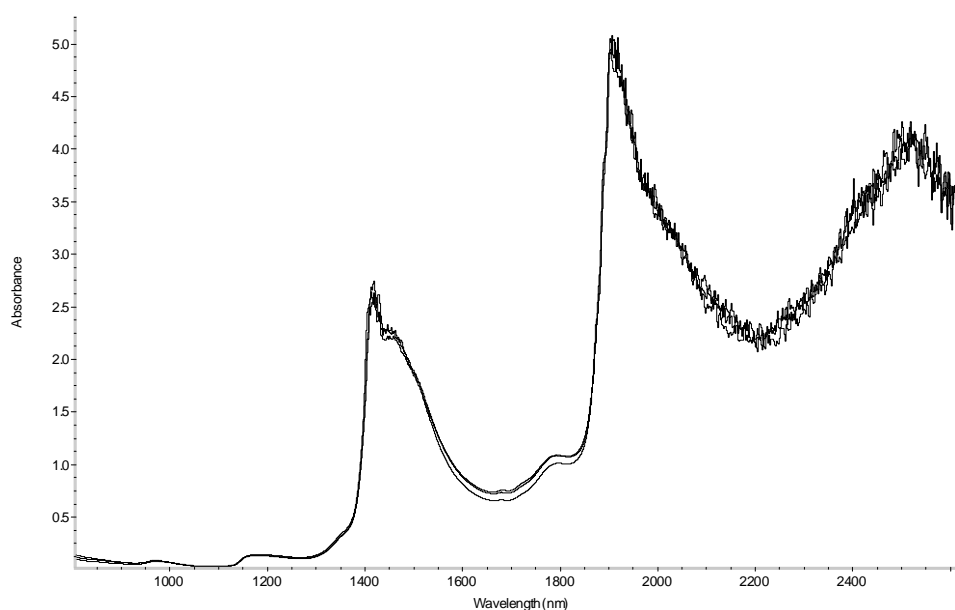
The DA models were then validated by using it to predict the age of samples in the validation set.

3   Results and discussion

3.1 Spectral Analysis

Figure 1 shows the average spectra of 1, 3 and 5 years old sample groups in the whole sample set without any preprocessing. No obvious spectral differences could be observed between the three spectra, and the spectra are highly overlapped except in the regions 950-980 nm, 1140-1160 nm, 1650-1700 nm, 1780-1800 nm, 1880-1920 nm, and 2200-2250. All spectra have intense absorption bands at 1450 nm, related to the first O-H overtone, and at 1900-1950 nm, related to the combination of stretch and deformation of the O-H group in water. The small absorption band at 1690 nm might be related to the -CH3 stretch first overtone or C-H groups in aromatic compounds, at 1782 nm related to the C-H stretch first overtone, at 2266 nm likely with C-H combination bands of methanol, and at 2302 nm with combination band of C-H stretch and

deformation of C-H from the -CH2 group (Yu et al., 2007; Cozzolino et al., 2003)[1,6].

Fig.1 Average spectra of 1, 3 and 5 years sample groups in the whole mature vinegar sample

3.2 Chemometrics analysis

3.2.1 Calibration Analysis

The calibration results of DA model developed on raw spectra are much better than that on first and second derivative spectra. There are only 1 sample is misclassified in the DA model developed on raw spectra. While there are 74 and 80 samples incorrectly classified in the DA model developed on first and second derivative spectra. The reason for the performance of raw spectra better than that of second derivative spectra is that DA with raw spectra describes 100% of the variability, while DA with first and second derivative explains 54.4 and 37.1%, respectively (The data, 100%, 54.4 and 37.1%, are obtained from the calibration results of the three models.). Thus, raw spectra were used for calibration and validation analysis. Table 2 shows the calibration results of DA models developed on raw, first and second derivative spectra. The results were in agreement with the literature for Chinese rice wine variety analysis (Hanyan Yu et al. 2003) [6].
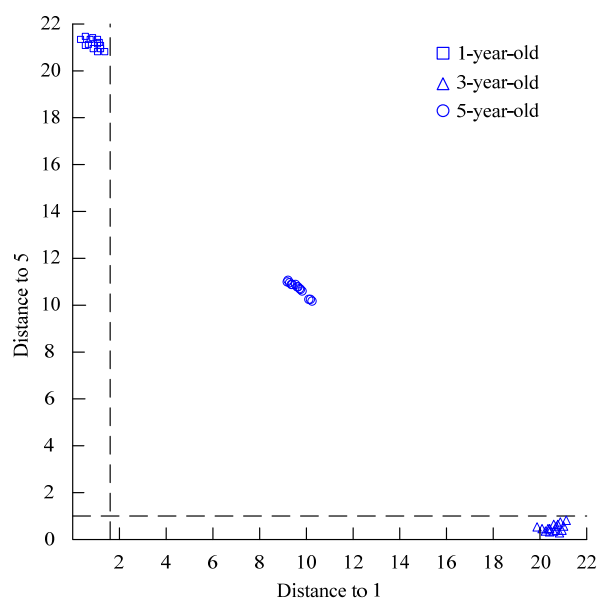
Tab. 2. Calibration results of DA models developed on raw, 1st and 2nd derivative spectra

Data type	Age (year)	Number of samples	Number of samples misclassified	% of samples correctly classified
Raw spectra	1	40	0	100
	3	40	0	100
	5	40	1	97.5
1st derivative spectra	1	40	24	40
	3	40	23	42.5
	5	40	27	32.5
2nd derivative spectra		40	23	42.5
		40	31	22.5
		40	26	35

3.2.2 Validation Analysis

Thirty samples (5 for each age of each variety) were selected randomly as the validation set. There were no samples misclassified. Therefore, NIR spectroscopy can be used to classify mature vinegar samples with different vinegar age. The ability of the NIR spectroscopy to discriminate or identify mature vinegar is based on the vibrational responses of chemical bonds to NIR radiation. It is that the major components of mature vinegar with different vinegar age are significantly different. Most of the changes were related to the primary structural components of organic molecules (C–H, N–H and O–H bonds), which provided the necessary information for classification of mature vinegar with different vinegar age by NIR spectroscopy. Meanwhile, the spectral difference between the sample groups with different mature vinegar age can be detected from Figure 1, correspondingly.

The pairwise distance plots for the 1-, 3- and 5-year-old sample groups were showed in Figure 2, respectively. They show graphically the Mahalanobis distance between each standard and the two classes that are selected for the X- and Y-axis of the plot. The calibration standards for the class that is selected for the X- axis of the pairwise distance plot are similar to each other; the data points for those standards will be clustered in the upper left corner of the plot. Similarly, the data points for the standards selected for Y-axis will be clustered in the lower right corner of the plot. The greater the distance between the two clusters in the plot, the greater is the difference between the corresponding class and the other class. From Figure 2, it can be seen that the 3 sample groups can be clearly classified.

Fig.2 Pairwise distance plot for 1-, 3- and 5-year-old sample groups in the whole sample set

4   Conclusions

The applicability of FT-NIR spectroscopic technique for predicting age of mature vinegar is presented in this work. The percentages of samples correctly classified in calibration and validation analysis were 99% and 100%, respectively. The results indicated that NIR spectroscopy together with DA is a powerful tool for discrimination between mature vinegar samples with different vinegar age. Further studies are needed to extend to other varieties of mature vinegar.

Acknowledgement

References

1. Cozzolino, D., H. E. Smyth, and M. Gishen. 2003. Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. *J. Agric. Food Chem.* 51(26): 7703-7708.
2. Liu, L., D. Cozzolino, W. U. Cynkar, R. G. Dambergs, L. Janik, B. K. O'Neill, C. B. Colby, and M. Gishen. 2007. Preliminary study on the application of visible-near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. *Food Chem.* 106(2):781-786.
3. Pontes, M. J. C., S. R. B. Santos, M. C. U. Araujo, L. F. Almeida, R. A. C. Lima, E. N. Gaiao, and U. T. C. P. Souto. 2006. Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Res. Intl.* 39(2): 182-189.
4. Reid, L. M., A. Woodcock, C. P. O'Donnell, J. D. Kelly, and G. Downey. 2005. Differentiation of apple juice samples on the basis of heat-treatment and variety using chemometric analysis of MIR and NIR data. *Food Res. Intl.* 38(10): 1109-1115.
5. Saiz-Abajo, M. J., J. M. Gonzalez-Saiz, and C. J. Pizarro. 2004. Classification of wine and alcohol vinegar samples based on near-infrared spectroscopy: Feasibility study on the detection of adulterated vinegar samples. *J. Agric. Food Chem.* 52(25): 7711-7719.
6. Yu, H. Y., Y. Zhou, X. P. Fu, L. J. Xie, and Y. B. Ying. 2007. Discrimination between Chinese rice wines of different geographical origins by NIRS and AAS. *European Food Res. Tech.* 225(3-4): 313-320.