Measurement 97 (2017) 156–164

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Year prediction of a mild aroma Chinese liquors based on fluorescence spectra and simulated annealing algorithm ☆

Zhu Weihua [a,*], Chen Guoqing [b], Zhu Zhuowei [b], Zhu Feng [a], Geng Ying [a], He Xiang [a], Tang Chunmei [a]

[a] *College of Science, Hohai University, Nanjing, Jiangsu, 210098, PR China*
[b] *College of Science, Jiangnan University, Wuxi, Jiangsu, 214122, PR China*

ARTICLE INFO

Article history:
Received 27 January 2016
Received in revised form 25 October 2016
Accepted 4 November 2016
Available online 5 November 2016

Keywords:
Fluorescence spectra of liquor
Correlation analysis
Simulated annealing algorithm
Partial least square method
Year forecast

ABSTRACT

Based on the three-dimensional fluorescence spectra of original degree Chinese liquors with different years, which belongs to a well-known series in China, this article establishes the year forecast model of Chinese liquors. Firstly, the analysis of correlation between three-dimensional fluorescence spectra and liquor's year shows that the year information is mainly distributed in the area of excitation wavelength of 200–230 nm and 250–320 nm and the emission wavelength of 400–500 nm. Secondly, by analyzing the multiple correlations of two-dimensional fluorescence spectra with the excitation wavelength of 220 nm, it is found that the multiple correlation of original spectra is stronger than that under wavelet decomposition. The detail signals have weak multiple correlation, while the approximate signal has strong multiple correlation. Finally, the spectral forecast model related to the year and the statistical distribution of year information are studied based on the nonlinear partial least square regression method. The results show that a year forecast model with high accuracy can be obtained by the simulated annealing algorithm, up to 0.96 year. The original fluorescence spectra contain abundant year information. The year information of wavelet decomposition spectra has layered distribution, and the year information is mainly concentrated in the third and the fourth layer. Furthermore, by the comparison of fluorescence spectral models in above three typical excitation wavelength interval, the year forecast model with high precision can be established in the excitation wavelength of 200–230 nm. The partial least square regression algorithm can effectively eliminates the multiple correlations among the fluorescence spectra. The research results provide not only a convenient optical method for the year calibration of the liquor, but also an important reference for the fluorescence spectra research of multi-component gradient system.

1. Introduction

There are a huge consumer market and diverse consumer groups in the field of liquor in China. However, the frequent occurrence of fake liquor is a serious threat to the public health and safety. As a fermented food of large production and sales, the safety of liquor is highly concerned by the society [1].

In many fraud cases, false information about the year of liquor has become the most serious issue. Because the different vintages have similar taste, some wineries and retailers mark the false year of liquor in order to earn large illegal benefits. Insiders point out: the lack of vintage calibration standard and industry standard makes it difficult to distinguish true or false vintage products in the market and leads to the particular year liquor becomes only a sale stunt [2].

As is known, two main components of liquor are water and ethanol. Besides, the liquor contains a small amount (the whole quality is less than 3%) but a large number (the species are more than 300) of organic compounds, such as alcohols, esters, acids, aldehydes, and phenols. Qiao point out that the liquor undergoes a series of slowly and continually physical or chemical changes in the storing time [3], which cause complex change among various components in liquor. Therefore, the vintage belongs to a typical multicomponent graded system. The complex composition of liquor makes it difficult in the analysis and detection of the liquor [4–6].

At present, the alcohol detection methods usually include infrared spectroscopy, dynamic light scattering technique, chromatographic analysis, atomic absorption spectral analysis, and so on.

☆ Project supported by the Foundation of National Natural Science (Grant No. 61378037).

* Corresponding author.
E-mail address: weihua_zhu@126.com (Z. Weihua).

http://dx.doi.org/10.1016/j.measurement.2016.11.016
0263-2241/

Sara et al. have determined the content change of hexanal caused by oxidation using combined methods of solid-phase microextraction, gas chromatography, and mass-spectrometry [7]. As a mature spectrum technology, molecular fluorescence spectroscopy has been widely used in component analysis and physical test, for its advantages of high sensitivity, good selectivity, low sample consumption, and convenience. For example: Yang et al. [8] have examined the authenticity of liquor based on three-dimensional fluorescence spectra; Liu et al. [9] have carried out the determination of resveratrol in red wine based on the three-dimensional spectra; Du et al. [10] have studied the organic decomposition using the three-dimensional fluorescence spectra. However, due to the fluorescence spectra with large line-width of molecular, spectra of various substances will produce different degrees of cross interference when large number of fluorescent substances exists in detection system. Concerning this issue, Zhu et al. [11–15] found that the liquor can produce strong fluorescence when excited by UV (ultraviolet) light. The vintages of different or same species specie with different years have various features in fluorescence spectra. Thus, by the correlation analysis between liquor fluorescence spectra and year, we can establish a new method to test the year of liquor.

Based on the three-dimensional fluorescence spectra of a certain Chinese liquors, this paper studies the correlation between the fluorescence spectra and the year of liquor, as well as the multiple correlations between fluorescence spectra with different wavelengths. Based on the nonlinear partial least squares method, we learn the distribution of information related to the year in the fluorescence spectra, and obtain the year prediction model of liquor in high accuracy.

2. Experimental instrument and samples

2.1. Experimental instrument

The British Edinburgh FLS920P steady-transient fluorescence spectrometer is adopted as the experimental apparatus, S/N is 9647:1(Water peak Raman).The parameters are set as follows: the excitation wavelength ranges from 200 to 600 nm with step-length of 5 nm, and the emission wavelength ranges from 200 to 600 nm with step-length of 1 nm.

2.2. Experimental samples

The experimental samples are 25 kinds of original degree liquors with the years from 1 to 40, which are offered by a Chinese brand winery.

3. Fluorescence spectra and correlation analysis of liquor

Fig. 1 shows the three-dimensional fluorescence spectra of 4 typical Chinese liquors (5 years, 10 years, 20 years and 40 years) in contour. It can be seen that the contours of three-dimensional fluorescence spectra of Chinese liquors with different years are generally similar, except for local distribution pattern and fluorescence intensity. It can be explained that the density of certain components in liquor and its emitted fluorescence intensity change with the year. Therefore, there is obvious correlation between the fluorescence intensity and year of liquor.

3.1. Correlation between the three-dimensional fluorescence spectra and the year of liquor

The correlation function of fluorescence spectra and year of liquor can be written as follows

$$R_{Y/ij} = abs = \left\{ \frac{\sum_{n=1}^{N}(x_{n/ij} - \bar{x}_{ij})(y_n - \bar{y})}{\sqrt{\sum_{n=1}^{N}(x_{n/ij} - \bar{x}_{ij})^2} \cdot \sqrt{\sum_{n=1}^{N}(y_n - \bar{y})^2}} \right\} \tag{1}$$

where i and j are the excitation wavelength and the emission wavelength respectively. $x_{n/ij}$ is the fluorescence intensity. y_n is the year of liquor. $R_{Y/ij}$ is the correlation coefficient of $x_{n/ij}$ and y_n. N is the sample size. According to (1), the absolute value cloud of $R_{Y/ij}$ are calculated with 25 liquor samples (distributed in 1–40 years) as shown in Fig. 2. It can be found that the correlation degree between fluorescence spectra and the year of liquor is low in the whole excitation-emission coordination. The maximum correlation coefficient is less than 0.6. With the emission wavelength less than 400 nm, the correlation coefficient is less than 0.2 in most regions. With the emission wavelength of 400–500 nm, and the excitation wavelength of 200–230 nm and 250–320 nm, the correlation degree is relatively higher with the values among 0.3–0.5. Thus, the correlation between fluorescence spectra in a single wavelength and year of liquor is relatively weak. The information of year distributes in a large range of fluorescence spectra. It is difficult to predict the year of the liquor by the model with single spectral intensity, however, multiple regression forecasting model with high precision can be built.

Fig. 3 shows the correlation between fluorescence spectra and year with the excitation wavelengths of 220 nm, 300 nm and 580 nm and the emission wavelength of 400–600 nm. The correlation is small with the excitation wavelength of 580 nm, but it is larger and smoother with the excitation wavelength of 300 nm. Correlation curve exhibits a large fluctuation with the excitation wavelength of 220 nm, and its value is greater than that with the excitation wavelength of 300 nm in many areas. Therefore, the follow-up multiple regression modeling is based on these two-dimensional fluorescence spectra with the excitation wavelength of 220 nm. The modeling results with the excitation wavelengths of 300 nm and 580 nm are also presented for influence comparison by different excitation wavelengths in next discussion.

3.2. Multiple correlation of two-dimensional fluorescence spectra

Multivariate regression modeling is usually based on the least square theory, which requires no correlation among variables avoiding random in coefficient and failure of model. Fluorescence spectra are optical transition of molecules excited by UV. Due to the broadening of the spectral and spectral overlap of different levels, the fluorescence intensity at different wavelengths is often correlated, which is called multiple correlation. In order to eliminate or reduce the multiple correlation between variables for better modeling, people often do derivation or wavelet decomposition on data for pretreatment. This chapter focuses on the multiple correlation of two-dimensional fluorescence spectra data of 25 kinds of liquor samples when the excitation wavelength is 220 nm.

3.2.1. Multiple correlation of original fluorescence spectra

The multiple correlation function of fluorescence spectra and year of liquor can be described by

$$R_{ij} = abs \left\{ \frac{\sum_{n=1}^{N}(x_{n/i} - \bar{x}_i)(x_{n/j} - \bar{x}_j)}{\sqrt{\sum_{n=1}^{N}(x_{n/i} - \bar{x}_i)^2} \cdot \sqrt{\sum_{n=1}^{N}(x_{n/j} - \bar{x}_j)^2}} \right\} \tag{2}$$

where $x_{n/i}$ and $x_{n/j}$ are the fluorescence intensity with emission wavelength of i and j respectively. R_{ij} is the correlation coefficient of $x_{n/i}$ and $x_{n/j}$. N is the sample size. Fig. 4 shows the calculated multiple correlation value of fluorescence spectra with the excitation wavelength of 220 nm and emission wavelength of 300–600 nm.