An NMR-Based Evaluation Model for Chinese Liquor Using KPCA and Cloud Model

Linlin Cheng[1,*], Jia Zheng[2], Xingzhong Xiong[3], Mingju Chen[1,3], Xingyue Zhang[1,3]
[1]*School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China*
[2]*Technology Research Center, Wuliangye Yibin Co., Ltd., Yibin, China*
[3]*Intelligent Perception and Control Key Laboratory of Sichuan Province, Sichuan University of Science & Engineering, Yibin, China*
**Corresponding Author*

Abstract: With the growing market value of Chinese liquor (Baijiu), smart technologies are increasingly being adopted in this field. Intelligent and accurate classification of Baijiu remains at an early stage in the market, representing a practical research direction that can translate time cost into economic benefits. This study employs nuclear magnetic resonance (NMR) spectroscopy combined with machine learning techniques to classify Baijiu and establish an objective evaluation model for its sensory characteristics. First, principal component analysis (PCA) and kernel principal component analysis (KPCA) were applied to reduce the dimensionality of NMR data. Point cloud models reflecting different sensory features of Baijiu were constructed in both PCA and KPCA spaces. It was observed that the point cloud model based on KPCA exhibited better aggregation density, which more effectively represents the classification outcomes of various Baijiu types. Subsequently, the ranges of point cloud regions and characteristic descriptors of Baijiu were analyzed. Using the well-classified KPCA-NMR point cloud model, regional correlations between sensory descriptors and cloud models were established. Finally, regression models-including linear, quadratic, exponential, and polynomial regression-were developed by correlating KPCA-transformed NMR data with sensory evaluation scores. Experimental results demonstrated that the KPCA-based point cloud models of different Baijiu categories are more separable in space, facilitating the distinction of their specific features. Furthermore, the integration of cloud model regions with objective evaluation terms yielded a polynomial regression model that achieved the highest correlation coefficient. This model outperformed the other three by 256.97%, 43.07%, and 238.84% in goodness-of-fit, respectively. In tests involving unknown grades of Baijiu, the proposed model achieved a classification accuracy of 95.31%, which was the closest to the results obtained from sensory evaluation.

Keywords: Chinese Liquor; Kernel Principal Component Analysis; Nuclear Magnetic Resonance Spectroscopy; Cloud Model

1. Introduction

Baijiu is a mixture composed of various chemical components, primarily consisting of H2O, C2H5OH, and 1% to 2% trace components [1-6]. These trace components not only influence the color, taste, and aroma of Baijiu, but they are also a critical factor in determining the quality of the Baijiu [7-10]. The sensor techniques are used to build Baijiu spectrum data through detecting some of the trace components of Baijiu, including gas chromatography [11,12], liquid chromatography [13-15], GC-MS [16-18], infrared spectrometry [19-21], electronic nose [22,23], and NMR [24,25]. In particular, NMR technology collects sample atlas information using MRI machine, which is recognized for its strength in representation and reproducibility.

NMR utilizes electromagnetic waves to excite nuclei with non-zero spin in magnetic fields. This method has proven to be beneficial for food science materials [26-31]. Unlike other traditional analysis methods, NMR does not only enable error-free detection of protons components in wine, but also simplifies and speeds up the operational steps in a non-destructive manner [32-34]. NMR is main method to examination of the main components

Copyright @ STEMM Institute Press

of the red wine in many western countries, Marko Viskić [35] et al used the correctly implemented NMR technique to perform the tasks of differentiating red wines of different vintages, regions and varieties. Gougeon [36] et al used H-1 NMR metabolomics technique on Bordeaux wines to differentiate Bordeaux wines from French red wines, with final classification rates above 71%. Sulfur dioxide, a preservative additive commonly used in the red wine industry, has an effect on metabolites in wine, such as acetaldehyde and diacetyl, etc. Cassino [37] et al used H-1 NMR spectrum to highlight the structural changes of metabolites in the presence of SO2, and the results obtained corresponded to those determined by the OIV standard protocol. On the domestic side in China, Luan [38] et al. used NMR techniques to examine the chemical components of four different brands of beers, to obtain the NMR spectral profiles of the beers, and to investigate the differences between the main components of different brands of beer in combination with statistical methods. The difference in the alcoholic content of Baijiu brings about differences in taste and color. Liu [39] et al used NMR hydrogen spectrum to determine the alcoholic content of Baijiu, which was simple to operate and provided a quick method for enterprises and testing departments. In summary, NMR spectrum technology has the advantages of obvious features, good reproducibility, high sampling accuracy and strong recognition ability [40-42], and the recognition technology based on NMR spectrum has been applied to many fields and has achieved certain results. NMR spectrum technology has an irreplaceable and important role in different research projects for various wines, and the research results often provide a solid foundation for subsequent researchers' experiments.

We synthesize the commonalities existing in different Baijiu and consider that the alcohols, acids, esters and other trace components in liquors have H and C elements, which have different peak characteristics in NMR spectrum, and NMR spectrum can accurately and sensitively reflect the changes of trace components in Baijiu, so that the quality levels of Baijiu with different characteristics can be distinguished. The application of NMR spectrum technology to the characterization of Baijiu is still in the exploratory stage.

In view of this, we proposed Baijiu information point cloud models and Baijiu characteristic cloud models using NMR spectrum based on the above uncovered characteristics and the potential intelligent needs of the market, and finally constructed a Baijiu sensory judgment model. The judgment model fits the current low-cost and high-efficiency intelligent Baijiu classification technology that is urgently needed in the Chinese Baijiu market. On the Baijiu manufacturing side, the model can satisfy any Baijiu manufacturer to conduct internal pre-evaluation of a bottle of Baijiu for the purpose of quality testing to determine the price of that bottle of Baijiu. On the Baijiu market side, the model can help buyers intelligently determine whether the Baijiu they buy is counterfeit or not [43], and avoid being duped. Our contribution can be summarized as follows:

(1) We used NMR to obtain NMR spectrums of different grades of Baijiu, compared and analyzed the PCA and KPCA to achieve the performance of NMR spectrums data de-correlation and dimensionality reduction, and built point cloud probability models of different grades of Baijiu in the KPCA space.

(2) The identification words association was established, and the established point cloud models of Baijiu can realize the accurate identification and classification of Baijiu, which can lay a good theoretical foundation for the classification and identification of Baijiu through NMR spectrum technology.

(3) We used the KPCA dimensionality reduction NMR spectrum data combined with the sensory evaluation scores to do regression analysis and establish regression models for sensory evaluation of Baijiu to achieve objective evaluation of the sensory characteristics of Baijiu.

2. Materials and Methods

2.1 Testing Instruments and Testing Samples

The experimental instrument was a Bruker NMR spectrometer AVANCE III HD 600 MHz, and the parameters of the NMR spectrometer were shown in Table 1. 1600 samples of strong aromatic Baijiu were selected from southern Sichuan. After the sensory evaluation of the Baijiu samples, the samples can be classified into four grades: special, excellent, first, and second, and the Baijiu samples were numbered.

The AVANCE III HD 600 MHz is equipped with a new digital radio frequency synthesizer,

 Copyright @ STEMM Institute Press

which further enhances receiver performance, and a redesigned preamplifier for increased sensitivity, accuracy, and improved performance. The steps for the NMR test are as follows.

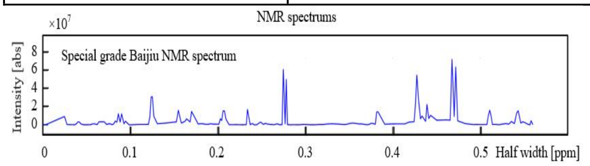
(1) Calibration of the NMR parameters for linearity and sensitivity to the NMR installation requirements.

(2) Deuterium oxide, dipotassium phosphate and phosphoric acid were used to configure the buffer solution with pH 2.0. Buffer solution of 100 μL was added to 900 μL of the base wine sample, mixed well, and the pH of the buffer solution was adjusted to about 4.0 using HCl and NaOH. An aliquot of 500 μL of the mixture was accurately taken and transferred to a 5 mm NMR test tube, to be measured.

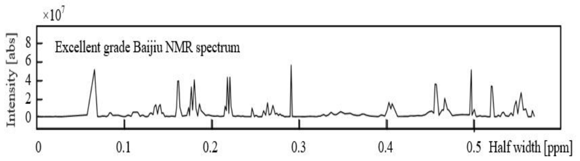
(3) The test was performed with an autosampler, and the topspin software was used to perform Fourier transform, baseline correction, phase adjustment, and integration interval setting on the NMR spectrum, and then the NMR spectrum of Baijiu was derived, as shown in Figure 1. Those with 93 points or more were identified as special grade base wines, 88 to 92.9 points as excellent grade base wines, 80 to 87.9 points as first grade base wines, and 70 to 79.9 points as second grade base wines. In response to the relevant requirements, the specific manufacturers information cannot be disclosed, so the base wines information collected from the relevant NMR spectrum were displayed by number, partly as shown in Table 2.

Table 1. Experimental Instruments

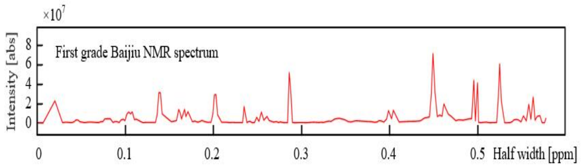
Instruments Names	Specifications/Models	Manufacturers
NMR spectrometer	AVANCE III HD 600 MHz	Bruker Corporation
Pipette gun	10-100 μL/100-1000 μL	Dragon Laboratory Instruments
Pipette gun	100 μL	Dragon Laboratory Instruments
Pipette gun	1000 μL	Dragon Laboratory Instruments
NMR tube	5 mm	NORELL Corporation



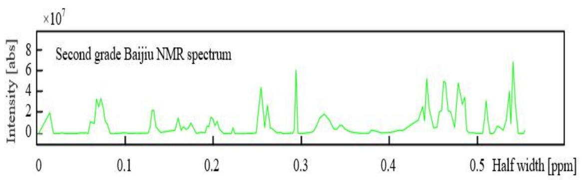
(a) Special Grade Baijiu NMR Spectrum



(b) Excellent Grade Baijiu NR Spectrum



(c) First Grade Baijiu NMR Spetrum



(d) Second Grade Baijiu NMR Spectrum

Figure 1. NMR Spectrums of Different Grades Baijiu

2.2 Sensory Identification of Baijiu Grades

The sommeliers observe the color of Baijiu through their eyes, smell its aroma through their sense of smell, and judge its taste through their sense of taste, with all three elements of information combined in order to provide a quantitative evaluation of its style. This process is known as sensory identification, and in order to be effective, it requires a clean and odorless environment, as well as fresh air, suitable light, and a constant temperature of around 20 to 25 degrees Celsius with a humidity of around 60%.

The sensory evaluation standards for Baijiu were based on the Chinese National Standard GB/T10345-2007 "Methods of Analysis for Baijiu"; the tasting panel was composed of professional sommeliers. The Baijiu samples were scored according to a total score of 100 points by the dark evaluation method, with color, aroma, taste, and style each accounting for 10, 25, 50, and 15 points, respectively, and then the average score of each judge was obtained.

As can be seen from Table 2, first-grade Baijiu is characterized by a poor aroma, sour and muddy flavor, slightly astringent taste, and a lack of balance and mellowness. Second-grade Baijiu is characterized by a strong spice, mixed flavors, astringency, and an imbalance. Excellent-grade Baijiu is characterized by a cellar fragrance, a slightly dated and mixed taste, more mellowness, and better balance. Special-grade Baijiu is characterized by a mellow aroma, a semi-aged lingering taste, very

Copyright @ STEMM Institute Press

mellow and sweet flavors, and a distinct character.

2.3 Spectrum Cloud Models and Kernel Principal Component Analysis

2.3.1 Spectrum cloud models

The mathematical method that combines the theoretical approach of fuzzy phenomena with random phenomena in statistics so that qualitative and quantitative descriptions form a one-to-one mapping relationship is referred to as a cloud model for describing non-deterministic relationships, thereby linking natural language and data language.

The cloud model is built mainly based on the three-characteristic information of expectation (E_x), entropy (E_n) and hyper entropy (H_e). In the cloud model, the location of the center of gravity in a point cloud region is the E_x of that point cloud region, which best represents the coordinates of the corresponding qualitative concept of that quality Baijiu on the number field, that is, reflects the central location of the Baijiu quality information. The range of the area where the point cloud is located is the range of acceptable values for the qualitative concept, which in this experiment is the range of the number field that can be described by the sensory evaluation language, indicating the fuzziness of the qualitative concept. E_n is used in the cloud model to represent the range of number fields. The credibility of this evaluation language is assessed using the H_e , which measures the cohesiveness and degree of cohesion of a point cloud cluster. This index helps indicate the affiliation of the point cloud cluster.

In the cloud model, the forward cloud generator converts qualitative concepts into quantitative values, and the inverse cloud generator converts quantitative values into qualitative concepts, and numerical characteristics (E_x , E_n and H_e) reflecting material information can be obtained, and the cloud generators works in the following specific steps.

Table 2. Results of Baijiu Tasting based on NMR Spectrum

Samples Number	Storage Time [years]	Fermentation Time [days]	Sensory Identification	Evaluation	Scores
51	2	90	First	Mixed sauce, Unbalanced	84.4
52	3	90	First	Sauce wine, Severely sour	87.8
53	1	7	First	Muddy odor, Additives	87.8
54	1	90	First	Mud flavor, Grain aroma	87.9
55	2	90	First	Sauce aroma, Not mellow	80.5
56	6	45	First	Heavy mud flavor, Not mellow	86.8
57	3	90	First	Mud flavor, Strong aroma, Bland	80.2
58	2	90	First	Sauce wine	81.9
59	4	90	First	Poor aroma	87.6
60	4	120	Second	Strong flavor of spices	78.5
61	5	90	Second	Strong aftertaste	70.5
62	0.5	90	Second	Astringent	79.1
63	3	90	Second	Adding spices, Not balanced	75.5
64	2	90	Second	Mixed flavors, Strong aftertaste	79.7
65	2	120	Second	Rancid acid	79.4
66	5	180	Second	Heavy mud flavor, Ammonia smell, Not balanced	76.3
67	7	180	Second	Mixed flavors, Astringent aftertaste	79.8
68	2	90	Second	Astringent	79.2
69	2	90	Second	Astringent	70.4
70	3	120	Excellent	Slightly dated	92.8
71	3	210	Excellent	Slightly dated	92.9
72	2	120	Excellent	Sauce wine, Relatively mellow	92.4

Firstly, calculate the averages of the input data samples for each group.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, i = 1, 2, ..., n \tag{1}$$

Where, n is the number of experimental repetitions and x_i is the i -th data sample.

The absolute center distance of the first-order samples is calculated from the average of the

 Copyright @ STEMM Institute Press

sample array.

$$d = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|, i = 1,2,...,n \quad (2)$$

Then, the sample variance of this data set is calculated as follows.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, i = 1,2,...,n \quad (3)$$

Finally, the average value of the substance sample information can be obtained as E_x , E_n and H_e are the ambiguity and uncertainty of the substance information, respectively, which are given by the following equations.

$$E_x = \frac{1}{n}\sum_{i=1}^{n}x_i, i = 1,2,...,n \quad (4)$$

$$E_n = \sqrt{\frac{\pi}{2}}\bullet\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|, i = 1,2,...,n \quad (5)$$

$$H_e = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 - E_n^{\ 2}}, i = 1,2,...,n \quad (6)$$

According to the formulas of the cloud model generator, the m-dimensional E_x obtained from the material data can be fed into the cloud generator with the m-dimensional E_n and H_e , and then n random numbers with normal distribution characteristics can be randomly generated from E_n and $H_e^{\ 2}$. This step is repeated m times to make $i = 1,2,\ldots,m$, a one-dimensional normal random number x with E_{xi} and E_{ni} can be generated. After that, the determinacy of each sampling point is calculated according to the qualitative concept.

$$\mu = \exp\sum_{i=1}^{m}\left[-\frac{(x_i - E_{xi})^2}{2E_{nn_i}^{\ 2}}\right] \quad (7)$$

$\{x_1, x_2, ..., x_m\}$ together with the deterministic u to form a point cloud. By repeating the above steps n times, m-dimensional point clouds can be obtained and the final model of Baijiu is generated.

2.3.2. Kernel Principal Component Analysis

PCA is primarily a series of linear transformations of the original dataset that preserves data with high variance and eliminates data with low variance, thus reducing the amount of data and achieving the objective of dimensionality reduction. Nevertheless, when non-linear correlations between data exist, the PCA de-correlation capability is lacking. KPCA projects the data from the original space to a higher-dimensional feature space through a nonlinear mapping, and then extracts the principal components in the high-dimensional feature space with the PCA method, allowing the linear partitioning of nonlinear relation Baijiu data to a high-dimensional space.

The set of spectrum data samples is X ($m \times n$ dimensions), and now X is mapped into the high-dimensional space $\phi(x_i)$.

$$C = \frac{1}{n-1}\sum_{i=1}^{n}\phi(x_i)\phi(x_i)^T = \frac{1}{n}[\phi(x_1),...,\phi(x_n)]\begin{bmatrix}\phi(x_1)^T \\ \cdots \\ \phi(x_n)^T\end{bmatrix} \quad (8)$$

Let $X^T = [\phi(x_1),...,\phi(x_n)]$, then we have: $C = \frac{1}{n-1}X^T X$.

Introduce kernel function:

$$K = XX^T = \begin{bmatrix}\phi(x_1)^T \\ \cdots \\ \phi(x_n)^T\end{bmatrix}[\phi(x_1),...,\phi(x_n)] = \begin{bmatrix}k(x_1,x_1) & \cdots & k(x_1,x_n) \\ \vdots & \ddots & \vdots \\ k(x_n,x_1) & \cdots & k(x_n,x_n)\end{bmatrix} \quad (9)$$

Let the eigenvector of K be u , the eigenvalue is λ, to find the eigenvalues, the eigenvectors have:

$$(XX^T)u = \lambda u \quad (10)$$

Multiplying both the left and right sides of the above equation by an X^T has:

$$X^T(XX^T)u = \lambda X^T u \quad (11)$$

So:

$$(X^T X)(X^T u) = \lambda(X^T u) \quad (12)$$

Also, because $(n-1)\bullet C = X^T X$, so the eigenvalues of the matrices K and C are identical and are both λ . The eigenvector of C is $X^T u$, for which a unitization is obtained as:

$$\upsilon = \frac{1}{\|X^T u\|}X^T u = \frac{1}{\sqrt{u^T XX^T u}}X^T u = \frac{1}{\sqrt{u^T Ku}}X^T u = \frac{1}{\sqrt{u^T \lambda u}}X^T u = \frac{1}{\sqrt{\lambda}}X^T u \quad (13)$$

Then find the projection of x on υ :

$$\upsilon^T\phi(x_j) = (\frac{1}{\sqrt{\lambda}}\mathbf{X}^T\mathbf{u})^T\phi(x_j) = \frac{1}{\sqrt{\lambda}}\mathbf{u}^T\mathbf{X}\phi(x_j) = \frac{1}{\sqrt{\lambda}}\mathbf{u}^T\begin{bmatrix}\phi(x_1)^T \\ \vdots \\ \phi(x_n)^T\end{bmatrix}\phi(x_j) = \frac{1}{\sqrt{\lambda}}\mathbf{u}^T\begin{bmatrix}k(x_1,x_j) \\ \vdots \\ k(x_n,x_j)\end{bmatrix} \quad (14)$$

In this paper, we used the Gaussian kernel of the radial basis kernel function with the expression: $k(x,y) = e^{(-\gamma\|x-y\|^2)}$, where the γ parameter was adjustable.

2.4 Ethical Statement

Institutional Review Board (IRB) or equivalent approval was not sought for this study for the following reasons: the study was conducted on white wine, not human subjects, and was of a product evaluation nature; the participating sommeliers performed their regular duties as professional evaluators and did not involve additional risk; the data collected were limited to a professional evaluation of the quality of the white wine, and did not contain sensitive personal information; the sommelier's participation was entirely voluntary and within the scope of their normal duties; and IRB review is not normally required for this type of product evaluation study under general research ethics guidelines. The participation of the sommelier was entirely voluntary and within the scope of his or her normal duties; and IRB review is not normally required for product evaluation studies of this type according to universal research ethics guidelines. Nonetheless, the research process was conducted in strict compliance with relevant industry standards and ethical codes to ensure the professionalism and impartiality of the study.

2.5 Statement of Informed Consent

This Chinese Liquor evaluation study did not utilize a formal written or verbal informed consent process. Participating sommeliers were performing their regular duties as professional evaluators and were not subjects of the study. Their participation was entirely voluntary and they were fully aware of the purpose and process of the evaluation. The study only collected professional evaluations of the quality of white wines and did not involve personal information or additional risks. Considering the nature of the study and the minimal risks, formal informed consent was not considered necessary. Nevertheless, we ensured that all participants were fully informed about the purpose of the study and how the results of their evaluations would be used, and that their participation in itself demonstrated approval and consent to the study.

3. Results and Discussion

3.1 NMR Spectrum Cloud Models and Sensory Evaluation of Baijiu

3.1.1 Point cloud of Baijiu information based on PCA and KPCA

To reduce external interference and unnecessary information, the corresponding values of the data points to be processed are first subjected to first-order and second-order differentiation using forward difference. Subsequently, the data is processed using the Savitzky-Golay polynomial convolution smoothing method. Additionally, the multiple scatter correction method is employed to eliminate the effects caused by sample scattering, such as baseline shifts or offsets. Finally, redundant information is removed through the application of wavelet transform. PCA and KPCA are then employed to reduce the dimensionality of the Baijiu NMR spectrums data, respectively, and cloud models are constructed to illustrate the rank of Baijiu in order to study their connection.

The principle of maximum variance is used in principal component analysis to complete the calculation of principal components, and the number of principal components is finally determined by the actual need and the size of the cumulative contribution rate. The specific implementation process is as follows: firstly, the eigenvalue λ and its corresponding eigenvector ω of the data set to be processed are deduced by using the calculation method of principal component analysis, then the corresponding variance contribution rate of each component is calculated and reordered according to the size of the value from the largest to the smallest, and finally the cumulative variance contribution rate is solved. Remember that the value of the variance contribution rate of the i -th principal component is a_i, and the corresponding cumulative contribution rate is β_i, and the specific solution is shown as follows:

$$a_i = \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j}, i = 1, 2, ..., p, j = 1, 2, ..., p \tag{15}$$

$$\beta_i = \frac{\sum_{m=1}^{i} \lambda_m}{\sum_{j=1}^{p} \lambda_j}, i = 1, 2, ..., p, j = 1, 2, ..., p, m = 1, 2, ..., i \tag{16}$$

PCA and KPCA were used to process the Baijiu grade information data, and the cumulative variance contribution of the first two dimensions of the data both reached 0.9965, PC1 was 0.9921 and PC2 was 0.0044. This phenomenon indicates that after the data are processed by PCA and KPCA, the first two coefficients (PC1, PC2 or KPC1, KPC2) can represent the main characteristics of the NMR spectrum information, so that these data can be used instead of the original data for classification and identification processing, and cloud models can be established to express the characteristic information of different grades of Baijiu by the first two characteristic coefficients. The two-dimensional data were input into the inverse cloud generator to obtain numerical features reflecting the information of Baijiu, namely E_x , E_n and H_e .Table 3 and Table 4 show the grade information characteristic values of

 Copyright @ STEMM Institute Press

different Baijiu grades based on PCA and KPCA, respectively.

Table 3. Characteristic Values of Different Baijiu Grades Based on PCA

Grades	E_x		E_n		H_e	
	PC1	PC2	PC1	PC2	PC1	PC2
Special	-1.35×10^8	-1.84×10^7	1.83×10^6	2.87×10^6	1.91×10^6	2.72×10^6
Excellent	-1.63×10^7	-6.15×10^7	2.25×10^8	8.13×10^7	2.35×10^8	9.35×10^7
First	8.64×10^7	1.02×10^8	3.52×10^8	3.22×10^8	3.10×10^8	3.57×10^8
Second	6.49×10^7	-2.25×10^7	3.10×10^8	2.44×10^8	3.24×10^8	2.36×10^8

From Table 3, we can see that for the first principal component PC1, the E_x , E_n and H_e values of first grade and second grade Baijiu are not very different, indicating that the first grade and second grade Baijiu are of similar quality. E_x for special grade Baijiu are very different from those of other grades Baijiu and are more easily distinguishable from other grades Baijiu in terms of grade. As can be seen from Table 4, the eigenvalues corresponding to KPC1 and KPC2 of the special grade Baijiu are the smallest, and there is a gap between them and the eigenvalues of other grades of Baijiu. For the entropy values, there is not much difference between the entropy values of special grade and second grade Baijiu, and it can be concluded that they have a similar degree of ambiguity. The similar H_e values of first-grade and second-grade Baijiu indicate some ambiguity in the grade information for these two qualities. This suggests a high probability that the qualities of these two grades of Baijiu are not easily distinguishable. By inputting the eigenvalues of each of the four grades of Baijiu into the two-dimensional cloud generator, point cloud models of the Baijiu NMR spectrum can be obtained, as shown in Figure 2.

Table 4. Characteristic Values of Different Baijiu Grades Based on KPCA

Grades	E_x		E_n		H_e	
	PC1	PC2	PC1	PC2	PC1	PC2
Special	-1.59×10^{11}	-4.69×10^{11}	5.60×10^{10}	8.81×10^{10}	6.23×10^{10}	8.59×10^{10}
Excellent	-1.24×10^{11}	-2.02×10^{11}	1.72×10^{11}	5.64×10^{11}	1.79×10^{11}	5.59×10^{11}
First	3.33×10^{11}	3.28×10^{11}	8.57×10^{11}	5.74×10^{11}	1.10×10^{12}	6.85×10^{11}
Second	-5.02×10^{10}	3.44×10^{11}	1.32×10^{11}	6.87×10^{11}	1.28×10^{11}	6.83×10^{11}

In Figure 2, the vertical axis is the affiliation of the point cloud, which indicates the certainty of the Baijiu grade. The blue color represents the special grade Baijiu point cloud, the green color represents the excellent grade Baijiu point cloud, the black color represents the second grade Baijiu point cloud, and the red color represents the first grade Baijiu point cloud. The position of each grade of Baijiu is determined by the values of E_x corresponding to the PCA or KPCA coefficients in Table 3 and Table 4, and the range of its axes is determined by the values of E_n , while the "bell-shaped" affiliation of each point cloud is determined by H_e . The cloud model of each grade in Figure 2(a) is characterized by a "bell shape", and only the cloud model of the second grade Baijiu has a poor "bell shape" and insufficient affiliation, reflecting the lack of stability of its metric tendency and quality; the affiliation of the corresponding special grade Baijiu is particularly good and its ambiguity is minimal. Among the four different grades of Baijiu, the cloud models of excellent grade Baijiu and second grade Baijiu are closer together, indicating some similarity in their grade's information. The position results of the special grade Baijiu were farther away from the other grades Baijiu, thus making it easier to distinguish them from the other grades Baijiu in terms of grade. In Figure 2(b), the cloud models are inductively inferred from the E_x , E_n and H_e corresponding to KPC1 and KPC2 obtained by dimensionality reduction of the original data. Similar to the results based on PCA, the point cloud space locations of different grades of Baijiu are different, but the cloud models of different grades of Baijiu NMR spectrums are located farther away from each other in the KPCA domain, and the NMR spectrum of the same grade Baijiu is better aggregated and occupies less cloud space. It can be seen that KPCA better reflects the characteristics of different grades of Baijiu, and better classification results will be obtained in the KPCA domain.

3.1.2 Construction of Baijiu characteristics cloud models based on PCA and KPCA

Since clouds are composed of point clouds in point cloud clusters, and each point cloud is a mapping from qualitative to quantitative space, each one contributes to the determination of qualitative concepts. By studying their relationship, the qualitative outputs collected and processed by the instrument can be transformed into a human language of information about the senses. The corresponding linguistic information for the four grades of special grade, excellent grade, first grade, and second grade, is mainly: mellow aroma and semi-aged flavor; rich cellar fragrance, slightly dated, slightly mixed, and slightly mellow; poor aroma, sour and slightly astringent; strong aroma, astringent, and mixed.

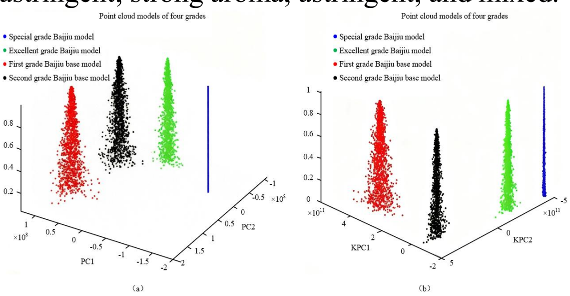


Figure 2. Cloud Models of Different Grades Baijiu (a) PCA, (b) KPCA

According to the relationship between the location of the central region of the point cloud and the linguistic information of each level, the word that was selected most frequently had a frequency of $\sqrt{0.01\times n_1}$ in the one-dimensional normal distribution. If the position of the w2ord in one-dimension ranges from $[Ex-kEn, Ex+kEn]$, the probability that the point cloud X falls in the interval $(-\infty, Ex+kEn)$ is:

$$P(X < Ex+kEn) = \frac{1}{2} + \frac{\sqrt{0.01\times n_1}}{2} \approx \alpha \quad (17)$$

Where, $X < Ex + k_1En$ is the range of values of the term in the cloud model coordinate system. We can look up the coordinate value k corresponding to the probability an according to the standard normal distribution table. The standard normal distribution formula is:

$$\delta \sim N(\mu, \sigma^2) \rightarrow \eta = \frac{\delta-\mu}{\sigma} \sim N(0,1) \quad (18)$$

Where, μ is the expected value and σ is the variance. Based on the point clouds of Baijiu grades information built by PCA and KPCA, we calculate the ranges of cloud models corresponding to each word of each grade of Baijiu, so as to complete the relationship between point clouds and words.

$$\frac{(x-E_{x_1})^2}{(kE_{n_1})^2} + \frac{(y-E_{x_2})^2}{(kE_{n_2})^2} \le 1 \quad (19)$$

Where, E_{x1} and E_{x2} are the expectations of two data with major categorical characteristics, and E_{n1} , E_{n2} is the entropy corresponding to E_{x1}, E_{x2}.

According to the above correlations formulas between the sensory words of Baijiu and the point cloud ranges of base wine grades information, we calculated the divisions of words in the cloud models built based on PCA and based on KPCA, respectively, and derived the correlation results between the cloud point regions and the evaluation words, and conducted the following collations and analysis.

The sensory words for each quality of Baijiu identified according to the sensory tasting appraisal panel were combined with the point cloud regions association formulas to divide the corresponding regions based on the grades information of Baijiu built by PCA and KPCA, and the results were shown in Figure 3.

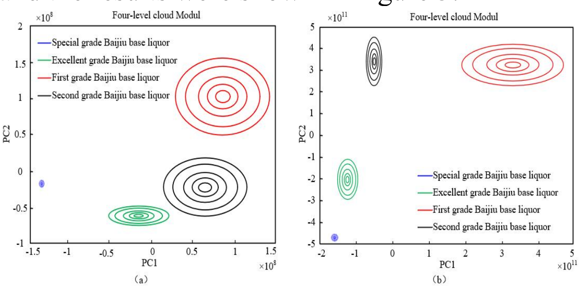


Figure 3. Cloud Models' Regions Maps of Different Grades of Baijiu (a) PCA, (b) KPCA

In Fig. 3, from right to left, blue represents the point cloud region of special grade Baijiu, green represents the point cloud region of excellent grade Baijiu, black represents the point cloud region of second grade Baijiu, and red represents the point cloud region of first grade Baijiu. Each grade of Baijiu is divided by an oval line, and each area corresponds to the corresponding judging word. The point cloud distribution results of the PCA and KPCA domains are somewhat similar, with both classification results showing that the special grade Baijiu are farther away from the other quality Baijiu, the relative position of the first grade Baijiu remain unchanged, and the relative positions of the other two grades of Baijiu have changed significantly. Compared with the PCA domain cloud models, the cloud models of Baijiu grades information built based on KPCA have the feature of greater relative distance between various types of Baijiu. It can be seen that KPCA maps the spectrum data to a high-dimensional space and then downscales it, which is more effective in achieving the classification of different grades of Baijiu.

The point cloud ranges were determined by taking the values within 90% of the probability of normal distribution. The formulas for the point cloud range of each quality of Baijiu in Table 5 and Table 6 correspond to the elliptical regions of each grade of Baijiu from the inside to the outside in Figure 3.

3.2 Construction of Sensory Evaluation Models for Baijiu Based on KPCA

On the basis of KPCA dimensionality reduction of NMR data, it was then combined with the sensory evaluation results and their statistical analysis to establish scientific objective and digital evaluation models of Baijiu quality.

The composite index g for the characteristics of Baijiu constructed from NMR spectrums data, g is the sum of the products of the KPCA F_j and the contribution ratio R_j.

$$\backslash g = \sum_{j=1}^{m} R_j F_j \tag{20}$$

With the composite index g as the independent variable and the scores values u generated by sensory score as the dependent variable, the mathematical models $u = f(g)$ of sensory score value of Baijiu were established. The models were used to determine the quality of Baijiu quantitatively.

For the elements g and the corresponding scores u in Table 7, the following kinds of regression analysis were done, and the final results were shown in Table 8, and their corresponding schematic diagrams were shown in Figure 4.

Table 5. Cloud Point Ranges of Different Chinese Spirit Grades Based on PCA

Base wine grades	Ranges	Probability	Evaluation Languages
Special	$\left(\frac{x+1.35\times10^{8}}{4.58\times10^{4}}\right)^2+\left(\frac{y+1.84\times10^{7}}{7.18\times10^{4}}\right)^2<1$	20%	Mellow aroma, Semi aged flavor, lingering taste, Mellow Sweet, characteristic
	$\left(\frac{x+1.35\times10^{8}}{9.52\times10^{4}}\right)^2+\left(\frac{y+1.84\times10^{7}}{1.49\times10^{5}}\right)^2<1$	40%	
	$\left(\frac{x+1.35\times10^{8}}{1.54\times10^{5}}\right)^2+\left(\frac{y+1.84\times10^{7}}{2.41\times10^{5}}\right)^2<1$	60%	
	$\left(\frac{x+1.35\times10^{8}}{2.36\times10^{5}}\right)^2+\left(\frac{y+1.84\times10^{7}}{3.70\times10^{5}}\right)^2<1$	80%	
	$\left(\frac{x+1.35\times10^{8}}{3.02\times10^{5}}\right)^2+\left(\frac{y+1.84\times10^{7}}{4.74\times10^{5}}\right)^2<1$	90%	
Excellent	$\left(\frac{x+1.63\times10^{7}}{5.63\times10^{6}}\right)^2+\left(\frac{y+6.15\times10^{7}}{2.03\times10^{6}}\right)^2<1$	20%	Cellar fragrance, slightly dated, mixed, mellow in the mouth, relatively balanced
	$\left(\frac{x+1.63\times10^{7}}{1.17\times10^{7}}\right)^2+\left(\frac{y+6.15\times10^{7}}{4.23\times10^{6}}\right)^2<1$	40%	
	$\left(\frac{x+1.63\times10^{7}}{1.89\times10^{7}}\right)^2+\left(\frac{y+6.15\times10^{7}}{6.83\times10^{6}}\right)^2<1$	60%	
	$\left(\frac{x+1.63\times10^{7}}{2.90\times10^{7}}\right)^2+\left(\frac{y+6.15\times10^{7}}{1.05\times10^{7}}\right)^2<1$	80%	
	$\left(\frac{x+1.63\times10^{7}}{5.63\times10^{6}}\right)^2+\left(\frac{y+6.15\times10^{7}}{2.03\times10^{6}}\right)^2<1$	90%	Poor aroma, sour and muddy flavor, slightly

First	$\left(\frac{x-8.64\times10^{7}}{8.8\times10^{6}}\right)^2+\left(\frac{y-1.02\times10^{8}}{8.05\times10^{6}}\right)^2<1$	20%	astringent, not balanced
	$\left(\frac{x-8.64\times10^{7}}{1.83\times10^{7}}\right)^2+\left(\frac{y-1.02\times10^{8}}{1.67\times10^{7}}\right)^2<1$	40%	
	$\left(\frac{x-8.64\times10^{7}}{2.96\times10^{7}}\right)^2+\left(\frac{y-1.02\times10^{8}}{2.70\times10^{7}}\right)^2<1$	60%	
	$\left(\frac{x-8.64\times10^{7}}{4.54\times10^{7}}\right)^2+\left(\frac{y-1.02\times10^{8}}{4.15\times10^{7}}\right)^2<1$	80%	
	$\left(\frac{x-8.64\times10^{7}}{1.37\times10^{8}}\right)^2+\left(\frac{y-1.02\times10^{8}}{1.26\times10^{8}}\right)^2<1$	90%	
Second	$\left(\frac{x-6.46\times10^{7}}{7.75\times10^{6}}\right)^2+\left(\frac{y+2.25\times10^{7}}{6.10\times10^{6}}\right)^2<1$	20%	Strong flavor of spices, mixed flavors, astringent, not balanced
	$\left(\frac{x-6.46\times10^{7}}{1.61\times10^{7}}\right)^2+\left(\frac{y+2.25\times10^{7}}{1.27\times10^{7}}\right)^2<1$	40%	
	$\left(\frac{x-6.46\times10^{7}}{2.60\times10^{7}}\right)^2+\left(\frac{y+2.25\times10^{7}}{2.05\times10^{7}}\right)^2<1$	60%	
	$\left(\frac{x-6.46\times10^{7}}{4.00\times10^{7}}\right)^2+\left(\frac{y+2.25\times10^{7}}{3.15\times10^{7}}\right)^2<1$	80%	
	$\left(\frac{x-6.46\times10^{7}}{1.21\times10^{8}}\right)^2+\left(\frac{y+2.25\times10^{7}}{9.52\times10^{7}}\right)^2<1$	90%	

Table 6. Cloud Point Ranges of Different Chinese Spirit Grades Based on KPCA

Base wine grades	Ranges	Probability	Evaluation Languages
Special	$\left(\frac{x+1.59\times10^{11}}{1.40\times10^{9}}\right)^2+\left(\frac{y+4.69\times10^{11}}{2.20\times10^{9}}\right)^2<1$	20%	Mellow aroma, semi-aged flavor, lingering taste, Mellow Sweet, characteristic
	$\left(\frac{x+1.59\times10^{11}}{2.91\times10^{9}}\right)^2+\left(\frac{y+4.69\times10^{11}}{4.58\times10^{9}}\right)^2<1$	40%	
	$\left(\frac{x+1.59\times10^{11}}{4.70\times10^{9}}\right)^2+\left(\frac{y+4.69\times10^{11}}{7.40\times10^{9}}\right)^2<1$	60%	
	$\left(\frac{x+1.59\times10^{11}}{7.22\times10^{9}}\right)^2+\left(\frac{y+4.69\times10^{11}}{1.14\times10^{10}}\right)^2<1$	80%	
	$\left(\frac{x+1.59\times10^{11}}{9.24\times10^{9}}\right)^2+\left(\frac{y+4.69\times10^{11}}{1.45\times10^{10}}\right)^2<1$	90%	
Excellent	$\left(\frac{x+1.24\times10^{11}}{4.3\times10^{9}}\right)^2+\left(\frac{y+2.02\times10^{11}}{1.41\times10^{10}}\right)^2<1$	20%	Cellar fragrance, slightly dated, mixed, mellow in the mouth, relatively balanced
	$\left(\frac{x+1.24\times10^{11}}{8.94\times10^{9}}\right)^2+\left(\frac{y+2.02\times10^{11}}{2.93\times10^{10}}\right)^2<1$	40%	
	$\left(\frac{x+1.24\times10^{11}}{1.44\times10^{10}}\right)^2+\left(\frac{y+2.02\times10^{11}}{4.74\times10^{10}}\right)^2<1$	60%	

	$\left(\frac{x+1.24\times10^{11}}{2.22\times10^{10}}\right)^2+\left(\frac{y+2.02\times10^{11}}{7.28\times10^{10}}\right)^2<1$	80%	
	$\left(\frac{x+1.24\times10^{11}}{2.84\times10^{10}}\right)^2+\left(\frac{y+2.02\times10^{11}}{9.31\times10^{10}}\right)^2<1$	90%	
First	$\left(\frac{x-3.33\times10^{11}}{2.14\times10^{10}}\right)^2+\left(\frac{y-3.28\times10^{11}}{1.44\times10^{10}}\right)^2<1$	20%	Poor aroma, sour and muddy flavor, slightly astringent, not balanced
	$\left(\frac{x-3.33\times10^{11}}{4.46\times10^{10}}\right)^2+\left(\frac{y-3.28\times10^{11}}{2.99\times10^{10}}\right)^2<1$	40%	
	$\left(\frac{x-3.33\times10^{11}}{7.20\times10^{10}}\right)^2+\left(\frac{y-3.28\times10^{11}}{4.82\times10^{10}}\right)^2<1$	60%	
	$\left(\frac{x-3.33\times10^{11}}{1.10\times10^{11}}\right)^2+\left(\frac{y-3.28\times10^{11}}{7.40\times10^{10}}\right)^2<1$	80%	
	$\left(\frac{x-3.33\times10^{11}}{1.41\times10^{11}}\right)^2+\left(\frac{y-3.28\times10^{11}}{9.47\times10^{10}}\right)^2<1$	90%	
Second	$\left(\frac{x+5.02\times10^{10}}{3.30\times10^{9}}\right)^2+\left(\frac{y-3.44\times10^{11}}{1.72\times10^{10}}\right)^2<1$	20%	Strong flavor of spices, mixed flavors, astringent, not balanced
	$\left(\frac{x+5.02\times10^{10}}{6.86\times10^{9}}\right)^2+\left(\frac{y-3.44\times10^{11}}{3.57\times10^{10}}\right)^2<1$	40%	
	$\left(\frac{x+5.02\times10^{10}}{1.11\times10^{10}}\right)^2+\left(\frac{y-3.44\times10^{11}}{5.77\times10^{10}}\right)^2<1$	60%	
	$\left(\frac{x+5.02\times10^{10}}{1.70\times10^{10}}\right)^2+\left(\frac{y-3.44\times10^{11}}{8.86\times10^{10}}\right)^2<1$	80%	
	$\left(\frac{x+5.02\times10^{10}}{2.18\times10^{10}}\right)^2+\left(\frac{y-3.44\times10^{11}}{1.13\times10^{11}}\right)^2<1$	90%	

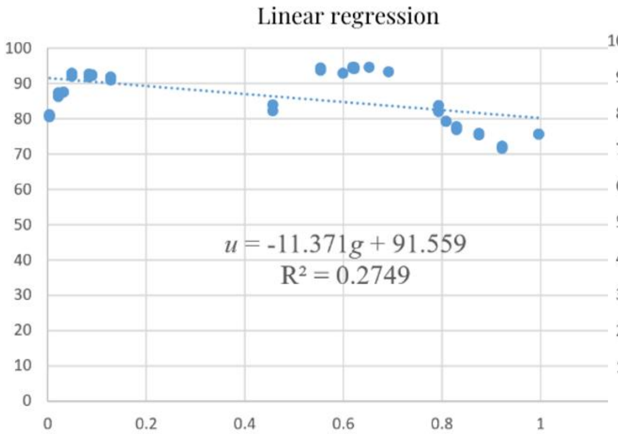
Table 7. Evaluation Indicators and Corresponding Sensory Scores

Samples numbers	123	124	127	128	134	138	139	143	70
g	0.72221	0.55408	0.92281	0.79163	0.75342	0.72221	0.55408	0.92281	0.12870
Sensory scores	93.2	94.5	94.0	94.6	94.2	93.7	94.3	94.6	92.8
Samples numbers	71	72	73	74	75	76	77	78	51
g	0.08419	0.09063	0.00461	0.59929	0.12870	0.08419	0.09063	0.00461	0.02278
Sensory scores	92.9	92.4	92.2	90.7	92.6	92.3	91.9	92.3	84.4
Samples numbers	52	53	54	55	56	57	58	59	60
g	0.99709	0.83023	0.79428	0.04335	0.02278	0.99709	0.83023	0.79428	0.04986
Sensory scores	87.8	87.8	87.9	80.5	87.8	80.2	87.9	87.6	78.5
Samples numbers	61	62	63	64	65	66	67	68	69
g	0.45711	0.71916	0.80966	0.87565	0.04986	0.45711	0.71916	0.80966	0.87565
Sensory scores	70.5	79.1	71.5	79.7	79.4	76.8	77.8	79.2	70.2

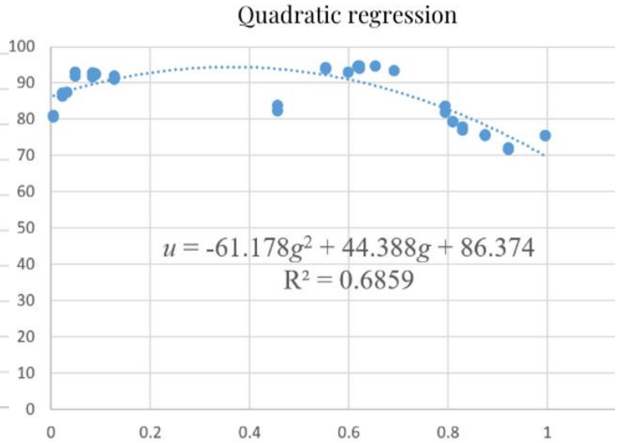
Table 8. Regression Analysis

regression analysis	Formulas	Correlation coefficients R^2
Linear regression	$u=-11.371g+91.559$	0.2749
Quadratic regression	$u=-61.178g^2+44.388g+86.374$	0.6859
Polynomial regression	$u=-4160.3g^6+16364g^5-23695g^4+15540g^3-4509.9g^2+459.01g+78.296$	0.9813

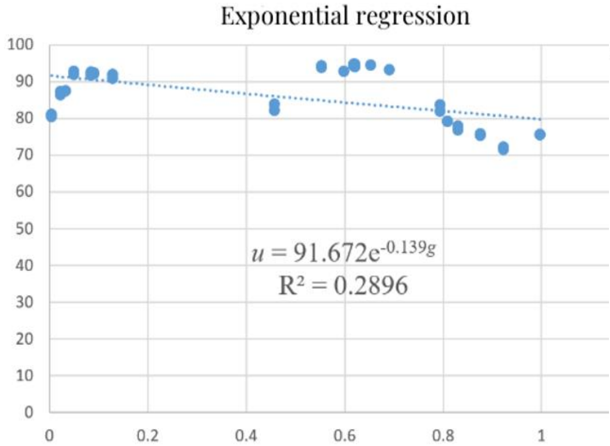
Exponential regression	$u = 91.672e^{-0.139g}$	0.2896



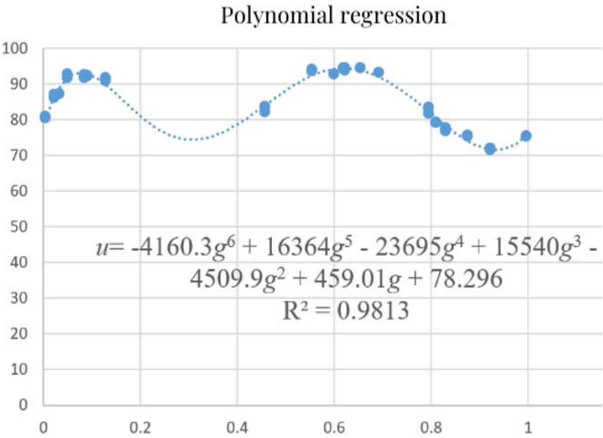
(a) Linear Regression



(b) Quadratic Regression



(c) Index Regression



(d) Polynomial Regression

Figure 4. Schematic Diagrams of Regression Analysis

From the correlation coefficients R^2 of several regression analyses in Table 8, it can be seen that the polynomial regression equation has the largest correlation coefficient, indicating that the polynomial regression equation can more accurately describe the relationship between g and u .Therefore, the equation for the KPCA domain-based Baijiu score of NMR spectrum is:

$$u = -4160.3g^6 + 16364g^5 - 23695g^4 + 15540g^3 - 4509.9g^2 + 459.01g + 78.296 \quad (21)$$

The polynomial regression equation was used to score the 64 collected Baijiu samples, rated them and compared them with the taster's sensory evaluations of the rating results, which were shown in Table 9. The correctness of the estimation equation using polynomial regression can reach 95.31%, which indicates that the estimation equation is more reliable.

Not only that, we also randomly selected 32 kinds of white wines on the market and scored them using regression equations and compared them with the sommelier's sensory evaluation of the scoring results, and the results are shown in Table 10, which has an accuracy rate of 90.62 %.

Table 9. The Results of Evaluation

Grades Rating Results	Match	Mismatch
Number of base wines	61	3
Percentage/%	95.31	4.69

Table 10. The Results of Evaluation

Grades Rating Results	Match	Mismatch
Number of base wines	29	3
Percentage/%	90.62	9.37

3.3 Discussion

The above results show that the established KPCA point cloud models for Baijiu have a good classification effect, and it is combined with objective judgment words to obtain the corresponding distribution cloud models for Baijiu types, and finally combined with the sensory evaluation results to form a numerical evaluation model with a high classification accuracy. In the domestic aspect of China in recent years, research on intelligent classification of Baijiu is in its infancy, and with reference to related research projects, two are worth discussing in our opinion.

Qian et al. [44] in 2021 established fingerprint profiles of flavor components in strong spiced Baijiu using GC-MS technique based on seven commercially available types of Baijiu, and combined similarity analysis and discriminant analysis (DA) to distinguish and identify these Baijiu. It was finally concluded that the Baijiu samples could be distinguished using DA with an accuracy of 100%. In this research paper, Qian et al. focus on roughly the same idea as this paper, both aiming to be able to form a fast and accurate classification of Baijiu, and Qian et al. achieve a final classification accuracy of 100%, which on the surface appears to be superior to ours in terms of accuracy. However, the DA used by Qian et al. is to summarize the regularity of things on a larger scale based on the fixed laws that have been mastered. There are obvious drawbacks to this analysis method, that is, the data laws of small samples do not fully reflect all the characteristics for two reasons: firstly, it is difficult to map small samples to larger samples with accurate and good play due to their own limitations; second, the application of DA is not similar to methods such as PCA, and DA does not ignore the properties that have very little impact and increase the importance of properties that carry strong information. The above drawbacks reflect the limitations of the work done by Qian et al. They only applied the classification law to the seven most common types of wine on the market for testing, lacking more comparisons and thus lacking practical applicability. In addition, the DA statistical method they apply can cause confusion and confusion in the nature of the data when faced with large data samples, so that the final decision model is prone to erroneous decisions when there are too many Baijiu samples.

With the rise of deep learning in various fields in recent years, a number of research scholars have also introduced deep learning into big data analytics. In the Baijiu intelligent technology industry, Liu et al. [45] landmark proposed a deep learning-based brand classification prediction model for Baijiu based on TensorFlow and Kera's frameworks. They collected the characteristic information of the Baijiu to be tested by means of an electronic tongue (array sensor) and combined it with known categories of Baijiu samples to be tested to establish a test sample dataset, and trained and tested the deep learning-based Baijiu brand classification prediction model through the training and test sets. The final classification accuracy of their proposed prediction model is 99%. The incorporation of deep learning neural networks into the classification of Baijiu is, in our opinion, an innovation that uses the power of neural networks to train models to "remember" the relevant Baijiu properties of the final generated Baijiu samples to achieve classification of unknown Baijiu. However, this method of thinking also has certain drawbacks. firstly, using deep learning to iterate and train on Baijiu feature information will eventually form a training model that can be used for testing, which only covers the feature information of the provided Baijiu types, and when faced with unknown Baijiu types, it is difficult to reason by analogy as in the case of objective evaluation of words. And if new feature information of Baijiu types needs to be incorporated, a training session from scratch is required, which is very time-consuming. Secondly, when deep learning is required for training, ordinary computers are unable to complete, requiring multi-GPU high computing rate servers, and a simple 4 GPU high server configuration has reached a value of nearly 100,000 yuan, which is generally unaffordable for research scholars; Thirdly, in the Baijiu market, the general buyer does not pay attention to the information provided by the composition characteristics of Baijiu commodities (such as the proportion of ingredients, etc.), but is concerned about the taste information provided. Therefore, while using deep learning to train on Baijiu characteristics for developing a classification prediction model is scientific, it lacks the human element of traditional sensory evaluation.

Compared with the above classical Baijiu classification studies, the prediction accuracy of our prediction model is slightly lower than them, but it can be seen to possess a very high generalization evaluation performance and high fault tolerance for unknown Baijiu categories. In contrast to the deep learning approach, we conduct field record examinations of Baijiu brewers and invite professional sommeliers to conduct professional evaluation of Baijiu taste in order to come up with identification words that fit the intelligent classification and the general public. When a new Baijiu category or even a foreign liquor category needs to be intelligently classified, it is easy and time-saving to simply incorporate the new identification terms. For the evaluation model derived from the DA analysis method, the accuracy rate can only reach 100% because the study only targets 7 known Baijiu categories (much lower than the 64 unknown Baijiu categories in this paper). Moreover, because of the obvious drawbacks of the DA analysis method, it can be boldly speculated that the misclassification rate will increase when the number of samples to be classified rises sharply and the information on the types of liquor increases.

4. Conclusion

In this paper, we used PCA and KPCA to reduce the dimensionality of Baijiu NMR spectrum and established point cloud models of different Baijiu grades in the domains of PCA and KPCA. The results showed that although the point cloud distribution results of the two methods were somewhat similar, the point cloud models of different Baijiu grades based on KPCA were more closely aggregated and the distance between different point cloud models were farther, so they can better reflect the characteristics of different Baijiu grades. After that, the division of the Baijiu grade information cloud models' regions built based on KPCA was completed on the basis of KPCA point cloud model, and the cloud regions ranges of Baijiu grade information and identification words were analyzed to realize the regional association between subjective evaluation words of Baijiu and the cloud models. Finally, the KPCA coefficients of the NMR spectral data were regressed on the sensory evaluation scores, and four types of classification evaluation models for Baijiu, namely linear regression, quadratic regression, exponential regression and polynomial regression, were established. From the results, it was found that the classification fit of the polynomial regression model was 0.9813, which was 256.97%, 43.07%, and 238.84% higher than the other three regression models, and the accuracy of the polynomial regression model in classifying unknown Baijiu grades was 95.31%.

Acknowledgments

This research was funded by the Wuliangye Group Industry-university-research cooperation project between Wuliangye Group and Sichuan University of Science & Engineering, grant number CXY2020ZR006.

References

[1] Gao L, Zhou J, He G. Effect of microbial interaction on flavor quality in Chinese baijiu fermentation. Frontiers in Nutrition. 2022; 9:960712.
[2] Hong J, Zhao D, Sun B. Research progress on the profile of trace components in baijiu. Food Reviews International. 2023; 39(3):1666-93.
[3] N. Shi ZYQ, L. Lei. Study on the Determination of Various Trace Components in Liquor by Gas Chromatography. Analysis and Testing. 2020.
[4] Sandhaus D. Drunk in China: Baijiu and the world's oldest drinking culture: U of Nebraska Press; 2019.
[5] Tu W, Cao X, Cheng J, Li L, Zhang T, Wu Q, et al. Chinese Baijiu: The perfect works of microorganisms. Frontiers in Microbiology. 2022; 13:919044.
[6] Wang B, Wu Q, Xu Y, Sun B. Multiple sugars promote microbial interactions in Chinese baijiu fermentation. LWT. 2021; 138:110631.
[7] Fan W, Zhao X, Du G, Chen J, Li J, Zheng J, et al. Metaproteomic analysis of enzymatic composition in Baobaoqu fermentation starter for Wuliangye baijiu. International Journal of Food Science & Technology. 2021; 56(8):4170-81.
[8] Song Z, Du H, Zhang M, Nie Y, Xu Y. Schizosaccharomyces pombe can reduce acetic acid produced by Baijiu spontaneous fermentation microbiota. Microorganisms. 2019; 7(12):606.
[9] Zhao Z, Sugimachi M, Yoshizaki Y, Yin X, Han XL, Okutsu K, et al. Impact of solid-state saccharification on the flavor of rice-flavor baijiu. Journal of Food Science. 2021; 86(11):4958-68.
[10] Li J, Yang P, Zhang J, Shui Z, Huo D, Hou C, et al. A metal ion-regulated colorimetric sensor array: discriminating Chinese Baijiu from other beverages. Analytical methods. 2019; 11(41):5274-81.
[11] Chen S, Lu J, Qian M, He H, Li A, Zhang J, et al. Untargeted headspace-gas chromatography-ion mobility spectrometry in combination with chemometrics for detecting the age of Chinese liquor (Baijiu).

 Copyright @ STEMM Institute Press

Foods. 2021; 10(11):2888.
[12]Song X, Jing S, Zhu L, Ma C, Song T, Wu J, et al. Untargeted and targeted metabolomics strategy for the classification of strong aroma-type baijiu (liquor) according to geographical origin using comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry. Food Chemistry. 2020; 314:126098.
[13]Wang Z, He M, Chen B, Hu B. Triazine covalent organic polymer coated stir bar sorptive extraction coupled with high performance liquid chromatography for the analysis of trace phthalate esters in mineral water and liquor samples. Journal of Chromatography A. 2021; 1660:462665.
[14]Wu C, Xing X, Liu G, Su D, Li A, Gui S, et al. Effects of Nongxiangxing baijiu (Chinese liquor) on mild alcoholic liver injury revealed by non-target metabolomics using ultra-performance liquid chromatography quadrupole-time-of-flight mass spectrometry. Journal of bioscience and bioengineering. 2022; 134(1):62-9.
[15]Yan Y, Lu J, Nie Y, Li C, Chen S, Xu Y. Characterization of volatile thiols in Chinese liquor (Baijiu) by ultraperformance liquid chromatography–mass spectrometry and ultraperformance liquid chromatography–quadrupole-time-of-flight mass spectrometry. Frontiers in Nutrition. 2022; 9:1022600.
[16]Hu S, Wang L. Age discrimination of Chinese Baijiu based on midinfrared spectroscopy and chemometrics. Journal of Food Quality. 2021; 2021:1-14.
[17]Zhang X, Wang C, Wang L, Chen S, Xu Y. Optimization and validation of a head space solid-phase microextraction-arrow gas chromatography-mass spectrometry method using central composite design for determination of aroma compounds in Chinese liquor (Baijiu). Journal of Chromatography A. 2020; 1610:460584.
[18]Zhou R, Chen X, Xia Y, Chen M, Zhang Y, Li Q, et al. Research on the application of liquid-liquid extraction-gas chromatography-mass spectrometry (LLE-GC-MS) and headspace-gas chromatography-ion mobility spectrometry (HS-GC-IMS) in distinguishing the Baiyunbian aged liquors. International Journal of Food Engineering. 2021; 17(2):83-96.
[19]Jun Z, Yang Y, Yao Y, Zi-wen L, Jian W, Chang-jun H. Application of mid-infrared spectroscopy in the analysis of key indexes of strong flavour Chinese spirit’s base liquor. Spectroscopy and Spectral Analysis. 2022; 42(3):764-8.
[20]Yadav PK, Sharma RM. Classification of illicit liquors based on their geographic origin using Attenuated total reflectance (ATR)–Fourier transform infrared (FT-IR) spectroscopy and chemometrics. Forensic science international. 2019; 295: e1-e5.
[21]Chen D, Tan Z, Huang Z, Lv Y, Li Q. Detection of lethal fake liquors using digitally labelled gas-phase Fourier transform infrared spectroscopy. Spectroscopy Letters. 2019; 52(3-4):204-10.
[22]Hou H-R, Meng Q-H, Qi P-F, Jing T. A hand-held electronic nose system for rapid identification of Chinese liquors. IEEE Transactions on Instrumentation and Measurement. 2021; 70:1-11.
[23]Wu X, Zhu J, Wu B, Zhao C, Sun J, Dai C. Discrimination of Chinese liquors based on electronic nose and fuzzy discriminant principal component analysis. Foods. 2019; 8(1):38.
[24]Li Y, Fan S, Li A, Liu G, Lu W, Yang B, et al. Vintage analysis of Chinese Baijiu by GC and 1H NMR combined with multivariable analysis. Food Chemistry. 2021; 360:129937.
[25]Zhao T, Chen S, Li H, Xu Y. Determination of linoleic acid oxylipins in Chinese Baijiu using ultra-performance liquid chromatography with quadruple-time-of-flight mass spectrometry (UPLC-QTOF-MS) and nuclear magnetic resonance (NMR). Analytical Letters. 2019; 52(14):2165-79.
[26]Consonni R, Cagliani L. Quality assessment of traditional food by NMR analysis. Food Control. 2022; 142:109226.
[27]Consonni R, Cagliani LR. The potentiality of NMR-based metabolomics in food science and food authentication assessment. Magnetic Resonance in Chemistry. 2019; 57(9):558-78.
[28]Radziej S, Scherb-Forster J, Schlicht C, Eisenreich W. NMR-based identification of thickeners in membrane-filtered food premixtures. European Food Research and Technology. 2022; 248(7):1715-20.

[29]Sobolev AP, Ingallina C, Spano M, Di Matteo G, Mannina L. NMR-based approaches in the study of foods. Molecules. 2022; 27(22):7906.
[30]Hatzakis E. Nuclear magnetic resonance (NMR) spectroscopy in food science: A comprehensive review. Comprehensive reviews in food science and food safety. 2019; 18(1):189-220.
[31]Sobolev AP, Thomas F, Donarski J, Ingallina C, Circi S, Marincola FC, et al. Use of NMR applications to tackle future food fraud issues. Trends in Food Science & Technology. 2019; 91:347-53.
[32]Fels L, Ruf F, Bunzel M. Quantification of isomaltulose in food products by using heteronuclear single quantum coherence NMR-experiments. Frontiers in Nutrition. 2022; 9:928102.
[33]He W, Bertram HC. NMR-based metabolomics to decipher the molecular mechanisms in the action of gut-modulating foods. Foods. 2022; 11(17):2707.
[34]Shumilina E, Johansen T, Glasner C, Robert J, Rustad T, Dikiy A. Application of NMR spectroscopy and conventional analytical methods for the assessment of wastewater from food industry. Waste and biomass valorization. 2020; 11:1349-57.
[35]Viskić M, Bandić LM, Korenika A-MJ, Jeromel A. NMR in the service of wine differentiation. Foods. 2021; 10(1):120.
[36]Gougeon L, Da Costa G, Guyon F, Richard T. 1H NMR metabolomics applied to Bordeaux red wines. Food Chemistry. 2019; 301:125257.
[37]Cassino C, Tsolakis C, Gulino F, Vaudano E, Osella D. The effects of sulphur dioxide on wine metabolites: new insights from 1H NMR spectroscopy based in-situ screening, detection, identification and quantification. Lwt. 2021; 145:111296.
[38]Luan X, Shen G, Zheng Y, Feng J, Dong J. Compositional Quantitation and Brand Identification of Beer via NMR Approach Combined with Multivariate Statistical Analysis. Liquor-Making Science & Technology. 2016; (05):48-53. doi: 10.13746/j.njkj.2016043.
[39]Liu E-m, Li X-y, Zong X-y, Li L, Wang S-f. Detection of alcohol content in Chinese Baijiu based on 1H nuclear magnetic resonance spectrum. 2021.
[40]Masetti O, Sorbo A, Nisini L. NMR tracing of food geographical origin: The impact of seasonality, cultivar and production year on data analysis. Separations. 2021; 8(12):230.
[41]Riley IM, Nivelle MA, Ooms N, Delcour JA. The use of time domain 1H NMR to study proton dynamics in starch-rich foods: A review. Comprehensive Reviews in Food Science and Food Safety. 2022; 21(6):4738-75.
[42]Schmitt C, Schneider T, Rumask L, Fischer M, Hackl T. Food profiling: determination of the geographical origin of walnuts by 1H NMR spectroscopy using the polar extract. Journal of agricultural and food chemistry. 2020; 68(52):15526-34.
[43]Peng H, Li L, Wu J, Sheng X, Rao J, Wang X, et al. Classification method of Strong-flavor Baijiu base liquor based on electronic tongue technique. China Brewing. 2022; 41(03):158-62.
[44]Qian Y, Hu X, Sun Y, Li J, Huang Z, Li D, et al. Classification of strong-flavor Baijiu based on fingerprint and stoichiometry. China Brewing. 2021; 40(06):152-6.
[45]Xin L, Qiang H, Yong-shuai Z, Xian-guo T. Research on the application methods of classifying and recognizing Baijiu wine based on deep learning. Food and Machinery. 2021; 37(4):68-71, 9.

 Copyright @ STEMM Institute Press