The Grape wine evaluation research based on TOPSIS and WRSR methods

Shihao Yang[1, a], Kaiyan Li[2, b] and Zekai Han[3, c]

[1]School of Science, North China University of Science and Technology, Tangshan 063000, China

[2]School of Economics, North China University of Science and Technology, Tangshan 063000, China

[3]School of Information Engineering, North China University of Science and Technology, Tangshan 063000, China

[a]18531548802@163.com, [b]562726475@qq.com, [c]872270165@qq.com

Abstract

The paper synthetically applies Gray Correlation Analysis, TOPSIS and WRSR into comprehensive evaluation of the quality of wine through the evaluation results of tasters. Firstly, we establish a comprehensive evaluation index system concerning four attributes and ten indexes of wine. Based on the evaluating statistics by tasters about twenty samples of the same kind of wine, we use them to analyze their gray correlations. The twenty groups of evaluation results showed that the otherness are less. Secondly, we build TOPSIS evaluation model and WRSR evaluation model, aiming at get every score of the quality of wine by calculating the evaluation results of twenty samples respectively. Finally, according to the percentage scoring system raised by Robert Parker, a famous wine taster in America, we get the average value of TOPSIS evaluation score and WRSR evaluation score based on different weights. In addition, we consider it as the final evaluation score of this kind of wine. The three models in our paper can be applied easily and feasibly. They overcame the subjectivity and one-sidedness of simple scoring method, providing new ideas and methods for wine quality evaluation and its grading.

Keywords

Grape wine; Evaluation; grey relational analysis; TOPSIS; non-integral Rank-sum ratio.

1. Introduction

With the development of living standards and taste of people, the demand for wine increases accordingly [6]. Therefore how to evaluate and authenticate wine scientifically and reasonably is well worth concerning and discussing. Traditional evaluation of wine generally hires some qualified tasters to mark every index respectively after tasting wine. Finally sum up all statistics to give evaluated score of wine [10]. However, owing to the difference of evaluation standard and style, the common mathematical method cannot be applied [5].

Wine quality evaluation and certification in academia has carried out a lot of research, some based on the data of wine-tasting score, some based on grape or its chemical composition. Wang Baixing and Feng Jishe based on five aspects of dry red wine, appearance, aroma, taste, typical, comprehensive quality score, applied fuzzy comprehensive evaluation and the principle of rank weighted average to estimate mouth feeling coordination of the red wine and analyze the results [2]. Yang Yi, Li Yanlai used t-test analysis, variance analysis, principal component analysis and grey relational analysis to analyze the problems in quality evaluation of grape wine including the credibility of wine judges' evaluation results, the relations between grape wine physiochemical indicators and grape wine quality, and wine quality grading etc. [3]. In this study of Yang Huafeng etc., red grape wine made from the same variety of wine grape but of different quality were detected by UV-visible

spectroscopy, and the correlations between wine quality and UV-visible absorption spectroscopy parameters were analyzed [7]. Chen Jin used F-test, rank sum test, and other significant analysis method, analysis on the grape-wine-tasting evaluation results for significance is made, used grape factor analysis, analysis of the confidence interval to establish wine quality classification model, and analyzed wine grapes and wine links between physical and chemical indicators by typical correlation analysis [6].

On the whole, there have been many researches made useful exploration and accumulated rich experience concerning the evaluation of wine and grading. But there are some deficiencies exist as follows: reuse statistical methods redundantly, neglecting other methods of application [11]; emphasize the evaluation of wine itself or chemical composition surface of grapes [7], the evaluation of tasters are much less; the choices of indexes mostly seem objective, neglecting the comprehensiveness and applicability of evaluation indexes of wine. Based on deficiencies mentioned above, our paper begins with the evaluation system which contains four kinds of attributes and ten indexes, building a comprehensive evaluation model about Grey Correlation, TOPSIS, and WRSR and evaluating empirically twenty groups' statistics of tasters.

2. Construction of evaluation index system for Wine

There is no specific and absolute quantification standard to distinguish between good and bad wine, currently authoritative wine scoring system is put forward by Robert Parker, the famous American wine critic. Parke uses the scoring system of 50 to 100 points [13]. According to its rating criteria, each wine can get 50 points on the basis of points. Another 50 points are composed of four elements, the appearance, aroma, flavor, overall quality or potential. The specific score indicators are as follows:

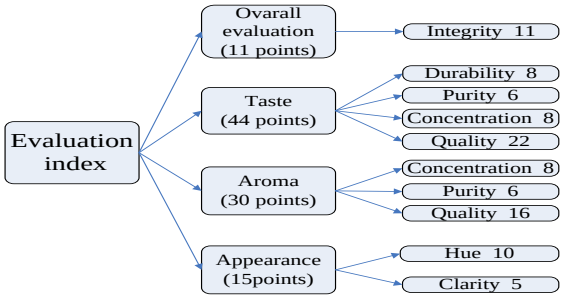

Fig. 1 Diagram of index system

According to the above indicators, each wine will get a score, different scores represent different quality. In the scoring system, the quality of the wine represented by the different points in the system, which is shown Table 1:

Table 1 Wine quality grading table

Score	Grade	Explanation
96-100	Extraordinary	top wine
90-95	Outstanding	wine with advanced taste characteristics
80-89	Above average	pure taste, making fine wines
70-79	Average	slightly flawed, but the taste is still no serious problem of wine
60-69	Below average	have significant defects or have unwelcome odors
50-59	Unacceptable	both unbalanced and very dull, it is not recommended for drinking

3. Evaluation Model Introduction

3.1 grey relational analysis [1]

As a changeable and developmental system, correlation analysis actually is the quantitative analysis in the process of dynamic development trend. The basic idea is judging whether sequence curve geometric shapes are connected tightly according to their level of similarity, The closer the curve, the

closer the development trend, the greater the degree connection. And vice versa. There are calculation steps:

(1)Standardized processing of 0-1 index value;

As there are different orders of magnitude between the evaluation indicators, they can't be compared directly. We must standard the original index value by 0-1. The formula is as follows:

$$b_{ij} = \frac{a_{ij} - a_j^{\min}}{a_j^{\max} - a_j^{\min}} \tag{1}$$

In this formula, $a_j^{\max}$ and $a_j^{\min}$ indicate the maximum and minimum values of the j-th index in all samples.

(2) Confirming the reference sequence;

(3) Confirming the compared sequence;

(4) Solving the correlation coefficient;

For a reference sequence χ_0, the comparing sequence is χ_i, we can express the differences between every comparing curves and reference curves at every point in this relationship:

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k |x_0(k) - x_i(k)| + \zeta \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \zeta \max\limits_i \max\limits_k |x_0(k) - x_i(k)|} \tag{2}$$

In this equation, $\xi_i(k)$ is a relative difference of the kth index comparison curve χ_i with the reference curve χ_0, and the relative value of the said χ_i to χ_0 in the correlation coefficient of K index. ζ is the resolution factor, $\zeta \in [0,1]$, and it is introduced to reduce the influence of the extreme value on the calculation. In actual use, should be based on the degree of correlation between sequence resolution coefficient selections, generally the most appropriate $\zeta \le 0.5$.

(5) Solving the grey weighted correlation coefficient of the sample.

The correlation coefficient only expresses the degree of association between every indicator statistics. Considering the difference of weights between all indicators, we aim to compare the absolute degree of association between overall curve χ_i and reference curve χ_0 .We can use the Gray weight relation to present, the formula is:

$$r_{0-i} = \frac{1}{n}\sum_{k=1}^{n} w_k \xi_i(k) \tag{3}$$

3.2 TOPSIS [5]

TOPSIS method, which is a Technique for Order Preference by Similarity to an Ideal Solution, is one of the common and effective methods for analyzing Multi-objective decision.For assessment and sequence by measuring index value vector of every object by evaluation and the relative distance of the ideal solution and negative ideal solution of evaluations.At the same time,calculating comprehensive index of each evaluation object. The steps are as follows:

(1)Confirming the standardized decision matrix;

(2)Disposing the raw data in type 0-1 standardize;

(3)Structuring weighted norm matrix c_{ij} ;

(4)Confirming the ideal solution c^+ and the negative ideal solution c^-;

(5)Calculating the distance values between each target values and ideal solutions, d_i^+ and d_i^+;

(6)Comprehensive evaluation index determined.

Calculating the relative proximity between evaluated objects and ideal solutions, as the comprehensive evaluation index.

$$c_i = \frac{d_i^-}{d_i^- + d_i^+} \tag{4}$$

Especially, if c_i value is larger, meaning that the evaluated object is closer with the positive ideal solution. It is easy to draw a conclusion that the evaluated object is better.

3.3 WRSR method [4]

WRSR method (non-integral Rank-sum ratio), it compiles the rank of the index value, which is similar to the linear interpolation method, which exists a quantitative linear relationship with the original index value, thereby it overcoming the shortcomings of quantitative information of the original index value easily lost when RSR method compiles rank. The calculation steps are as follows:

(1) To list raw data table and compile rank;

For high-priority index, the method of compile rank is

$$R = 1 + (n-1)\frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{5}$$

In the formula, R is the rank, n is the number of samples, X is the original index value, $X_{\min}$ $X_{\max}$ are respectively the minimum, maximum value of the original index.

(2) To calculate weighted RSR $wRSR_i$ [9];

(3) To determine the distribution of WRSR [5] (calculate probability unit);

(4) To calculate the linear regression equation;

(5) Grading and sorted.

After getting the results of TOPSIS and RSR method, to calculate their average according to the weight of w_T and w_w, namely comprehensive evaluation value of wine.

4. The empirical evaluation of wines

This study used data of China Undergraduate Mathematical Contest in Modeling title A in 2012 [12]. Please 10 qualified wine-tasting evaluate the 20 samples of a wine and score each index after tasting them. Seeing appendix for collecting data of samples and evaluating this wine by using the grey relational analysis method, TOPSOS and WRSR method.

4.1 Correlation analysis

We can use the formula to standardize the original data, the results are showed in appendix. In order to study the relevance of all kinds of samples, we choose the first sample as a reference sequence.

$X_0 = \{0.867, 0.588, 0.667, 0.370, 0.609, 0.529, 0.308, 0.313, 0.375, 0.300\}$, other samples are the compared sequences.

In this formula, we take ζ is 0.5, $\min_i \min_k |x_0(k) - x_i(k)| = 0$, $\max_i \max_k |x_0(k) - x_i(k) = 0.867|$. Gray weight relation of samples are showed in the Table 2:

Table 2 Table of Gray weight relation value of samples

r_{0-i}	r_{0-2}	r_{0-3}	r_{0-4}	r_{0-5}	r_{0-6}	r_{0-7}	r_{0-8}	r_{0-9}	r_{0-10}
Gray weight relation	0.071	0.074	0.075	0.080	0.080	0.064	0.061	0.080	0.077
r_{0-i}	r_{0-11}	r_{0-12}	r_{0-13}	r_{0-14}	r_{0-15}	r_{0-16}	r_{0-17}	r_{0-18}	r_{0-19}
Gray weight relation	0.064	0.067	0.061	0.072	0.055	0.078	0.077	0.045	0.055

Judging from the gray weight relation value in Table 2, the correlation broadens between samples, which means the difference seems small of each other. The evaluation difference of the wine by tasters is smaller, the liability is bigger than before. Judging from the Fig. 2, we can also see the close relationship between all samples.

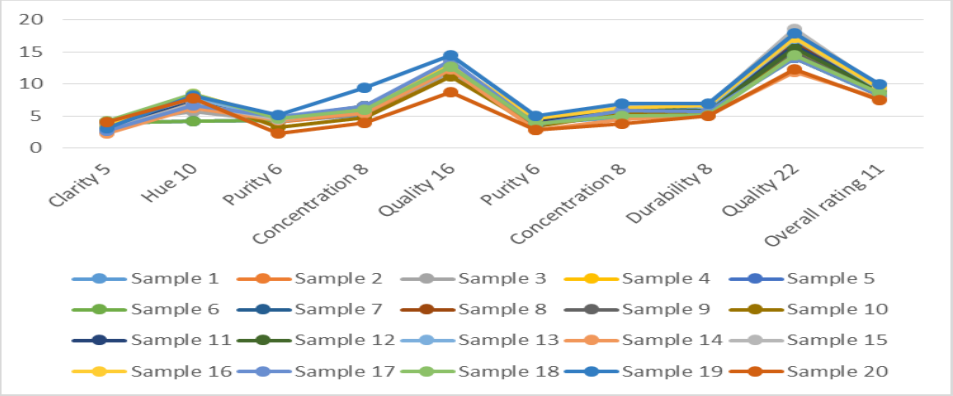

Fig. 2 the line chart of indicators evaluation of samples

As it can be seen from Fig. 2, each line is more closely, meaning that the higher the similarity of the first sample with other samples between 19, also shows there are minor differences between the 20 samples.

4.2 Evaluation of TOPSIS method

(1)Confirming the standardized decision matrix;

Disposing the raw data in normalized and same trending by using Office 2013. The standardized decision matrix we have got are as attached. What we can see from the decision matrix is that every data points are between 0 and 1, some of these include many 0 or 1. So explaining that there are some assessment value are the same in the raw data.

(2)Structuring weighted norm matrix;

In the evaluation of wine, there are ten indexes have determined in advance , including clarity, pure degrees , concentration and overall impression etc. and so the index weighting matrix is as follows:

$$w_j = \begin{bmatrix} 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.06 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.08 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.16 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.06 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.08 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.08 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.22 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.11 \end{bmatrix}^T$$

By $c_{ij} = b_{ij} \bullet w_j$,we can get the weighted norm matrixes c_{ij} , and the results are as attached.

(3)Confirming the positive ideal solution and the negative ideal solution [5];

Confirm the ideal solutions are as follows from the weighted norm matrixes:

Positive ideal solutions:

$$c^+ = (0.050,0.100,0.060,0.080,0.160,0.060,0.080,0.080,0.248,0.110)$$

Negative ideal solutions: $c^- = (0,0,0,0,0,0,0,0,0,0)$

(4) Calculating the distance between the target value and the ideal solution, and calculating the comprehensive evaluation index according to the formula (4). The calculation results are shown in Table 3.

Table 3 TOPSIS distance table

Samples	c^+	c^-	c_i	Sequence
1	0.2209	0.1657	0.4285	
2	0.5698	0.6923	0.5485	
3	0.1819	0.1941	0.5162	
4	0.1656	0.2147	0.5645	
5	0.2230	0.1803	0.4470	
6	0.2230	0.1803	0.4470	
7	0.1110	0.2686	0.7076	
8	0.1055	0.2731	0.7212	
9	0.1853	0.1952	0.5130	
10	0.2170	0.1702	0.4362	0.5739
11	0.1372	0.2429	0.6390	
12	0.1382	0.2454	0.6398	
13	0.0949	0.2820	0.7481	
14	0.3084	0.1238	0.2865	
15	0.0728	0.3284	0.8185	
16	0.0797	0.2986	0.7893	
17	0.3347	0.4340	0.5646	
18	0.1963	0.2074	0.5137	
19	0.0415	0.3489	0.8938	
20	0.3453	0.0957	0.2171	

As it can be seen from Table 3, the maximum value of the ideal distance c^+ is 0.5698, the minimum value is 0.0415, and the difference between the two is 0.5283, while the difference between the maximum value and the minimum value of the negative ideal distance c^- is 0.5966. At the same time, from the comprehensive index c_i also known that the difference between the 20 samples is not big, but it can be sorted.

Since the twenty samples are taken from the same kind of wine, by getting c_i values of the twenty samples, removing the maximum and minimum value and solving the average value, finally the comprehensive assessment of this kind of wine is 0.5739.

4.3 Evaluation of WRSR method

According to the steps of the WRSR method[8], the calculation results are shown in Table 4:

Table 4 calculation results of WRSR method [9]

Samples	*wRSR*	f	$\sum f$	$\tilde{R}$	$\bar{R}$	$(\bar{R}/n)\times 100\%$	*probit*	WRSR	Final score
20	7.0736	1	1	1	1	5.0	3.3551	7.7115	
13	7.7994	1	2	2	2	10.0	3.7184	8.0664	
9	8.7281	1	3	3	3	15.0	3.9636	8.3059	
1	8.7403	1	4	4	4	20.0	4.1584	8.4962	
5、6	8.7967	2	6	5,6	5.5	27.5	4.4022	8.7343	
3	9.0925	1	7	7	7	35.0	4.6147	8.9418	
8	9.1901	1	8	8	8	40.0	4.7467	9.0708	9.4472
16	9.2359	1	9	9	9	45.0	4.8743	9.1954	
17	9.2669	1	10	10	10	50.0	5.0000	9.3182	
2	9.4173	1	11	11	11	55.0	5.1257	9.4410	
4	9.4555	1	12	12	12	60.0	5.2533	9.5656	
10	9.8062	1	13	13	13	65.0	5.3853	9.6945	
11	9.9114	1	14	14	14	70.0	5.5244	9.8304	

6	10.1783	1	15	15	15	75.0	5.6745	9.9770
7	10.2682	1	16	16	16	80.0	5.8416	10.1402
12	10.3124	1	17	17	17	85.0	6.0364	10.3305
15	10.5412	1	18	18	18	90.0	6.2816	10.5700
14	10.7653	1	19	19	19	95.0	6.6449	10.9248
18	11.2579	1	20	20	20	98.8	7.2571	11.5228
20	7.0736	1	1	1	1	5.0	3.3551	7.7115

As can be seen from table 4, the weighted RSR of samples 5 and 6 are the same, so the two samples are compared with the other 18 samples as a whole. At the same time, we can see that the difference of RSR between each samples is small from value of the weighted RSR.

Cumulative frequency corresponding to the $probit_i$ as independent variables, with $wRSR_i$ as dependent variable, using statistical software to do regression analysis. The results are shown in Table 5 and Table 6:

Table 5 Variance analysis of regression equation

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17.621	1	17.621	316.449	.000[a]
	Residual	.947	17	.056		
	Total	18.567	18			

a. Predictors: (Constant), probit.

b. Dependent Variable: wRSR.

From table 5 to see, F=316.449, Sig=0.000, linear regression height of wRSR and probit is significant, effect of linear regression is better.

Table 6 Regression equation coefficient and T-test

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.434	.288		15.401	.000
	probit	.977	.055	.974	17.789	.000[a]

a. Dependent Variable: wRSR.

From table 5 to see, the T test of regression coefficient probit was 17.789, Sig=0.000, and the results were consistent with the F test, and the effect was good. And it can get the regression equation:

$$wRSR = 4.434 + 0.977 \times probit \tag{6}$$

Its effect, the F test of the equation and the T test of the regression coefficient is better. The WRSR estimation values are shown in Table 5. The overall evaluation of this wine is shown in the last column of Table 5, its processing method is same as the TOPSIS method.

According to the results of TOPSIS method and the WRSR method, determining the weight of 0.4 and 0.6 respectively, calculating the average of the two, the comprehensive evaluation value is 7.9639, so the wine can be thought of a good grade.

5. Conclusion

This study is based on the evaluation results of 10 indicators by wine critic members, the evaluations for wine are through grey relational analysis, TOPSIS sort and united no integer rank level RSR. Through access to information, and we can proposed four ten categories wine evaluation system which systemic summary the evaluation factor of wine roundly and according to the current authority of the wine scoring system. Based on this, we can determine stability and reliability of the evaluation by using gray correlation analysis and confirming whether there is a big difference of evaluation results by the different wine-tasting number' views for the same wine in geometrically level. From

the empirical results, although relative distance of some curves of samples are larger, but on the whole, the correlations of each curves are about 0.068, so it is a suitable results.

What is the advantage of the model prominently is using TOPSIS sort and non-integer rank level RSR synthetically, what are the important methods of evaluation and sort of samples. The main idea of TOPSIS sort is through measuring index evaluation value vector of each evaluated objects and relative distance between ideal solution and negative ideal solution of evaluations [5]. WRSR method is based on RSR method to compile rank for index value in similar to the linear interpolation mode. Ultimately determine the order of the object being evaluated [4]. The research get the evaluation result of wine by averaging the sort results of two methods and next to combine and average these with the different weights again. The two results do not the same magnitude, but the substance of score proportion which represents is the same. From this perspective, the results of the evaluation have more binding, more scientific and reliability.

References

[1] Du Dong, Pang Qinghua and Wu Yan: *Modern Comprehensive Evaluation Method and Case Selection* (Tsinghua University Press, China 2008), p.112-116.
[2] Wang Baixing, Feng Jishe: Application of Fuzzy Synthetic Evaluation in the Sense Estimation of Dry Red Wine, Food and Nutrition in China, Vol.17 (2011) No.8, p.33-37.
[3] Yang Yi, Li Yanlai: Quality Evaluation of Grape Wine Based on Comprehensive Statistics Analysis Model, Liquor-Making Science & Technology, Tol.252 (2015) No.6, p.118-123.
[4] Zhang Lixin, Cui Lijie: AN Evaluation Study on Agricultural Sustainability of Shandong Province-Based on Non-integer Rank WRSR, East China Economic Management, Vol.29 (2015) No.7, p.14-19.
[5] Wang He, Song Peige and An Lin: Comprehensive Evaluation of the Situation of Maternal HealthCare with TOPSIS and RSR, Chinese Journal of Health Statistics, Vol.32 (2015) No.2, p.240-242.
[6] Cheng Jin: Mathematical Model Based on Statistics Grape Wine Evaluation, Journal of Yancheng Institute of Technology (Natural Science Edition), Vol.27 (2014) No.1, p.18-21.
[7] Yang Huafeng, Yang Yong, Dong Qin, et al. Application of UV-visible Spectroscopy in Grape Wine Quality Evaluation, Liquor-Making Science & Technology, Tol.221 (2012) No.11, p.58-64.
[8] Chi xiong, Shan Limin, Yao Minli, et al. Application of RSR in Comprehensive Evaluation of Food Hygiene Supervision Work Quality, Chinese Journal of Public Health Management, Vol.22 (2006) No.2, p.114-115.
[9] Miao Huibo, Chang Wenhu and Peng Yingchun: Comprehensive Evaluation of Maternal and Child Health Status by Using Weighted RSR, Maternal and Child Health Care of China, Vol.22 (2007) No.7, p.849-851.
[10] Gao Xin, Yang Rumei, Guan Xianfei, et al. Wine Professional Sommelier's Credibility Model and Evaluation Method, Journal of Yunann Agricultural University, Vol.29 (2014) No.2, p.235-240.
[11] Li Yunlai, Li Jiming and Jiang Zhongjun: Application of Statistical Analysis in the Evaluation of Grape Wine Quality, Liquor-Making Science & Technology, Tol.178 (2009) No.4, p.79-82.
[12] Original data on http://www.mcm.edu.cn/problem/2012/2012.html.
[13] http://www.wine-world.com/.