Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 740754, 7 pages
http://dx.doi.org/10.1155/2014/740754

Research Article

Multivariate Methods Based Soft Measurement for Wine Quality Evaluation

Shen Yin,[1] Lei Liu,[2] Xin Gao,[1] and Hamid Reza Karimi[3]

[1] *College of Engineering, Bohai University, Liaoning 121013, China*
[2] *College of Mathematics and Physics, Bohai University, Liaoning 121013, China*
[3] *Department of Engineering, Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway*

Correspondence should be addressed to Lei Liu; bhushxll@gmail.com

Received 22 April 2014; Accepted 13 May 2014; Published 3 June 2014

Academic Editor: Josep M. Rossell

Soft measurement is a new, developing, and promising industry technology and has been widely used in the industry nowadays. This technology plays a significant role especially in the case where some key variables are difficult to be measured by traditional measurement methods. In this paper, the quality of the wine is evaluated given the wine physicochemical indexes according to multivariate methods based soft measurement. The multivariate methods used in this paper include ordinary least squares regression (OLSR), principal component regression (PCR), partial least squares regression (PLSR), and modified partial least squares regression (MPLSR). By comparing the performance of the four methods, the MPLSR prediction model shows superior results than the others. In general, to determine the quality of the wine, experienced wine tasters are hired to taste the wine and make a decision. However, since the physicochemical indexes of wine can to some extent reflect the quality of wine, the multivariate statistical methods based soft measure can help the oenologist in wine evaluation.

1. Introduction

The evaluation of wine quality is highly significant due to their effects on wine classification and target marketing. The quality of wine reflects several major influences, including quality ratings, the reputation of the winery, and profitability [1]. Therefore, it is imperative to evaluate wine quality for both the food industry and wine science community [2, 3]. In view of the evaluation of wine quality, the traditional method is manual inspection or analysis of the chemical compounds. These methods cost huge financial inputs and time. The studies of wine quality evaluation are abundant internationally. The support vector machine (SVM) can build the wine quality classification model. The probabilistic neural network can evaluate the wine quality based on the mineral elements content of red wine. The visualization method of wine quality evaluation proposed can evaluate the wine quality according to the physicochemical property. However, in this paper, the multivariate methods based on soft measurement are the appropriate tools for chemical and physical measurements based on the wine physicochemical indexes.

Soft measurement technology is widely used [4–6]. The production process and the automatic control theory can be combined perfectly by soft measurement technique. In the industrial process, it seems common that some key control variables cannot be measured due to technique and economy factors [7]. However, based on soft measurement, a mathematical model can be built according to the relation of detectable process variable and in this way the undetectable process variable can be measured or estimated [8]. Due to these detectable variables and estimable variables, the productive process is controllable.

With the method of online analysis applied, the investment and the instrument maintenance continue to increase. Manual inspection is a time-consuming and complex work. The soft measurement is superior, with the development of computer technology. The primary methods to build soft measurement model include mechanism based modelling, knowledge based modelling, and data based modelling according to different prediction models. Using mechanism based technique to build the model, the known physical and chemical laws are a must and sometimes the mechanism in

production process needs to be understood deeply. Moreover, this application is limited, owing to the difficulty in modelling and period length. Although the knowledge based modelling is a simple, easy-to-understand, and convenient method, it is not suitable for the high precision and the knowledge rules extracted. The process data based modelling method can build statistical regression model based on multivariate statistical analysis theory, using a large number of production data. This method is superior in modelling, maintenance, and precision. In this paper, the model based modelling is utilized and the model is built by the multivariate methods based on soft measurement.

In this paper, in order to evaluate red wine quality, the multivariate methods based on soft measurement are used. This algorithm can help construct the fitted wine quality model and predict wine quality. We make a comparison among the models established by using OLSR, PCR, PLSR, and MPLSR, respectively, and then select out the best model in order to improve the model accuracy [9]. The methods are superior to manual measurement in the facet of wine quality evaluation. Due to the serious stability of OLSR, the models of PCR, PLSR, and MPLSR are built. These models can also help improve the production process. Furthermore, it can be useful in target marketing, it can help identify the most relevant factors, and can help classify wines such as premium brands (useful for setting prices) [2]. The physicochemical indexes that can impact the quality of wine are also proposed in this paper. There are 87 physicochemical indexes collected, in which a set containing 20 samples is used as calibration set and one containing 7 samples is used as verification set. We predict the wine quality when these physicochemical indexes are just known.

This paper is organized as follows. Section 2 reviews the multivariate statistical methods, including OLSR, PCR, PLSR, and MPLSR. In Section 3, the data sets are introduced and analyzed. The wine prediction models are established based on the multivariate statistical methods and then a comparison among these models is provided. Finally, the conclusions are presented in the last section.

2. Modeling Algorithms

In this section, the multivariate statistical analysis is used to solve the wine problems. OLSR, PCR, PLSR, and MPLSR are introduced briefly. Owing to the multicollinearity or the less number of samples than variables, OLSR lost its effectiveness. PCR is utilized to solve the problem of collineation. The information plays an important role in the data set based on the cumulative percent variance (CPV). Unlike PCR, PLSR especially focuses on the internal height linearly dependent variables. MPLSR has the advantage of both PCR and PLSR and is better than them.

2.1. Ordinary Least Squares Regression (OLSR). Least squares method is a kind of mathematical optimization technique. The matching data is found by minimizing the error sum of squares. The unknown data is obtained by the least squares method simply. The difference between quadratic sum of actual data and quadratic sum of calculated data is very small so that the least squares method can also be used for curve fitting. In general, the OLSR algorithmn is used to solve problems only involving one dimensional response variables. However, OLSR can also solve the problems of more than one dimensional response variables. The algorithm is briefly introduced as follows.

Step 1. Collect n samples of the measurable variables and response variables and then normalize them to zero mean and unit variance, denoted as $U = [u_1 \cdots u_m] \in R^{n\times m}$, $Y = [y_1 \cdots y_p] \in R^{n\times p}$. The following steps will be executed only when U^TU is a full rank matrix.

Step 2. Minimize the sum of errors e_i:

$$\begin{aligned}\beta_i &= \arg\min e_i^T e_i \\ &= \arg\min \left(y_i - U\beta_i\right)^T \left(y_i - U\beta_i\right) \quad i = 1 \cdots p.\end{aligned} \tag{1}$$

Because of the invertibility of U^TU, the β_i can also be calculated by

$$\beta_i = \left(U^TU\right)^{-1}U^T y_i. \tag{2}$$

Step 3. Use the e_i and β_i to form $E \in R^{n\times p}$ and $M \in R^{m\times p}$, respectively, and the final model can be indicated as

$$Y = UM + E. \tag{3}$$

2.2. Principal Component Regression (PCR). Principal components analysis invented by Pearson in 1901 [10] can analyze the data and establish the mathematical model. The method plays an important role in the data of principal components (i.e., characteristic vectors) and their weights (i.e., the eigenvalues [11]) through the study of the characteristics of covariance to decompose matrix [12].

Since 1980s, PCA has been successfully applied in numerous areas including data compression, image processing, feature extraction, pattern recognition, and process monitoring [13, 14]. Since PCR is superior in dimensionality reduction, it may solve the variable multicollinearity in practice and OLSR's stability problem [15]. The collinearity problem is effectually solved. The irrelevant aggregative indicators, original indicators variant, take the place of the original ones based on PCR. The aggregative indicators are used to show the original indicators in this statistical approach. PCR has been widely used in many fields [16–18].

Step 1. Gather measurable variables and the responses variables and normalize them to zero mean and unit variance, presented as $U = [u_1 \cdots u_m] \in R^{n\times m}$ and $Y = [y_1 \cdots y_p] \in R^{n\times p}$.

Step 2. Accomplish singular value decomposition (SVD) on the covariance matrix of independent variable set:

$$\begin{aligned}&\frac{1}{n-1}U^TU = P\Lambda P^T, \\ &\Lambda = \operatorname{diag}\left(\lambda_1 \cdots \lambda_m\right) \quad \lambda_1 > \cdots > \lambda_m > 0\end{aligned} \tag{4}$$

in which P is the so-called loading matrix of covariance matrix.

Step 3. Determine the number of principle components with the appropriate criteria [19] and calculate the score matrix T:

$$T = \mathrm{UP}_{\mathrm{pc}} \in R^{n\times l}, \tag{5}$$

where

$$\begin{aligned} P_{\mathrm{pc}} \subset P &= \left[P_{\mathrm{pc}} \;\; P_{\mathrm{res}}\right] \in R^{m\times m};\\ P_{\mathrm{pc}} &= \left[p_1 \cdots p_l\right] \in R^{m\times l};\\ P_{\mathrm{res}} &= \left[p_{l+1} \cdots p_m\right] \in R^{m\times(m-l)}; \end{aligned} \tag{6}$$

l is the number of the principal components.

Step 4. Perform OLS regression between the score matrix T and the dependent matrix Y and obtain the final model:

$$\begin{aligned} \beta_i &= \left(T^T T\right)^{-1} T^T y_i, \quad i = 1 \cdots p,\\ B &= \left[\beta_1 \cdots \beta_p\right] \in R^{m\times p},\\ Y &= TB + E. \end{aligned} \tag{7}$$

Namely, $Y = UM + E$, $M = P_{\mathrm{pc}}B$.

2.3. Partial Least Squares Regression (PLSR). Partial least squares method can be applied in the case where the number of explanatory variables is very high. PLS generalizes and combines features from principal component analysis and multiple regression. Least squares regression and principal component regression method extract different factor scores. The purpose of PCR is to extract the relevant information to ensconce matrix U and predict the value of variable Y. Then we can use these independent variables to improve the quality of prediction model. When the correlation of some useful variables is low, the reliability of the final prediction model of PCR will go down. This technique has certain defects and is too hard in solving this problem. Nevertheless, PLS can decompose the variables U and Y and extract the components from U, Y at the same time. Latent variables (LVs) of U and Y have a strong relation and ensure the PLSR algorithmn based model can make prediction from measurable variables [20]. We decide a few factors to participate in the model. PLSR is widely used in many areas, such as fault detection, wine analysis, and chemistry [9, 21, 22]. We consider the standard PLSR method as follows.

Step 1. Collect the measurable variables U and the response variables Y and normalize the data sets of U and Y, expressed as $U = [u_1 \cdots u_m] \in R^{n\times m}$ and $Y = [y_1 \cdots y_p] \in R^{n\times p}$, in which u_i and y_i are normalized to zero mean and unit variance.

Step 2. Calculate the following equations γ times iteratively:

$$\begin{aligned} \left(p_k, q_k\right) &= \arg \max_{\|p_k\|=1, \|q_k\|=1} p_k^T U^T Y q_k, \quad k = 1 \cdots \gamma\\ a_k &= U_k p_k \qquad b_k = Y_k q_k\\ U_k &= a_k c_k^T + U_{k+1}, \qquad Y_k = a_k r_k^T + Y_{k+1}\\ c_k &= \frac{U_k^T a_k}{\|a_k\|^2} \qquad r_k = \frac{Y_k^T a_k}{\|a_k\|^2}\\ U_{k+1} &= U_k - a_k c_k^T, \qquad Y_{k+1} = Y_k - a_k r_k^T, \end{aligned} \tag{8}$$

where the p_k, q_k and a_k, b_k are the loading vector and score vector of U and Y, respectively, and the γ is the number of latent variables (LVs) which is usually determined by the cross validation criteria [22].

Step 3. Store p_k, a_k, c_k, r_k into P, A, C, R, and the result of standard PLSR algorithm can be expressed as

$$\begin{aligned} A &= UP, \quad U = AC^T + E,\\ Y &= AR^T + F. \end{aligned} \tag{9}$$

Namely, $Y = UM + F$, $M = PR^T$.

2.4. Modified Partial Least Squares Regression (MPLSR). As a matter of fact, the problem of multicollinearity and the less number of samples than variables are two common phenomena. When OLSR is applied, the problem of multicollinearity leads to a serious stability problem. PCR and PLSR can solve the problem of highly collinear. PCR algorithm solves the collinearity problems efficiently by introducing principal components (PCs). PLSR considers both outer relations (U and Y block individually) and inner relation (linking both blocks), but PCR only considers the outer relation of U block. PCR and PLSR run the risk of losing useful information in selected PCs or LVs by these dimension reduction methods. In order to solve these problems, Yin et al. proposed the MPLSR [23]. MPLSR has been validated on the industrial benchmark of Tennessee Eastman process for fault detection and a good result has been obtained.

Step 1. Gather all the n samples of measurable variables and response variables and stack them into U and Y, respectively. Normalize them to zero mean and unit variance, denoted as $U = [u_1 \cdots u_m] \in R^{n\times m}$ and $Y = [y_1 \cdots y_p] \in R^{n\times p}$.

Step 2. Calculate the regression coefficient matrix M:

$$\begin{aligned} \frac{1}{n} Y^T U &\approx M^T \frac{U^T U}{n},\\ M &= \begin{cases} \left(U^T U\right)^{-1} U^T Y, & \text{if } \operatorname{rank}\left(U^T U\right) = m,\\ \left(U^T U\right)^{\dagger} U^T Y, & \text{if } \operatorname{rank}\left(U^T U\right) < m \end{cases} \end{aligned} \tag{10}$$

in which the $(U^T U)^{\dagger}$ is the pseudoinverse of $U^T U$.

Table 1: The wine quality of one sample given by human experts.

Human expert	1	2	3	4	5	6	7	8	9	10	Mean points
Presentation (20 points)	9	10	10	10	10	9	10	10	10	7	9.5
Fragrance (30 points)	25	20	21	21	23	18	23	25	24	23	22.3
Mouth-feel (40 points)	31	26	29	25	28	32	28	32	27	26	28.4
Overall feeling (10 points)	9	8	8	9	9	8	9	9	8	9	8.6
Sum points (100 points)	74	64	68	65	70	67	70	76	69	65	68.8

Table 2: The wine quality of one sample given by human experts.

Group	1	2
The standard deviation	7.3426	3.978

Step 3. Obtain the final model of MPLSR:

$$Y = UM + E_y, \tag{11}$$

where E_y is the residue part of Y which is uncorrelated with U.

3. Results and Discussion

In order to classify and identify red wine, [24] proposed the technology based on the spectrum and pattern recognition. Nevertheless, this paper is devoted to comparing these methods and predicting the wine quality for the purpose of wine classification and target marketing.

In this section, we use three multivariate statistical methods based on soft measurement, including principal component regression (PCR), partial least squares regression (PLSR), and modified partial least squares regression (MPLSR) to find which is the best method according to the relationship between the red wine physicochemical indexes and the wine quality. We ignore ordinary least squares regression (OLSR) which is not for this situation. The fitting efficiency and prediction ability of these methods provided are compared by relevant figures or indexes and the best model can be searched.

3.1. Data Preprocessing. The data set (mathematical modeling official site: http://www.mcm.edu.cn) in this paper includes the red wine qualities, physicochemical indexes, and aroma substances perfectly, of which the last two are collectively known as physicochemical indexes in the following.

The scores, given by professional tasters in two groups, have obvious differences. Solving the scores is a primary task. We analyze two sets of scores from four aspects: presentation (20 points), fragrance (30 points), mouth-feel (40 points), and overall feeling (10 points) and the total is 100. For example, the important information in one sample can be found in Table 1.

In the data set, 27 samples of wine in the first group are equal to those in the second group. There are 10 tasters on different levels in each group. The missing data in the first group is replaced by average value (AVG). The first group's standard deviation (SD) is 7.3426 and the other one's is 3.978. The average value of samples in both groups and standard deviation of two groups are calculated in Table 2, respectively.

$$AV = \frac{1}{N}\sum_{i=1}^{N} x_i$$
$$SD = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(x_i - AV)^2}. \tag{12}$$

Obviously, the standard deviation of the second group calculated is smaller than the one of the first group. The second group provides a more reliable result than the first one. So we choose the second one as the actual wine quality to compare with fitted wine quality. To construct the predicting models, the set of 20 random samples collected is used as calibration set. The remainder samples are used as the verification set to verify the prediction ability of the models.

3.2. Modeling and Comparison. In order to guarantee the influence between indexes and quality of wine, some assumptions are made as follows.

(1) In this paper, the vinification processes are identical, and the environment of vinification is the same.

(2) The scores given by the expert tasters approach the real quality scores.

(3) The grapes used to vinify a kind of wine are of one species.

The purpose of all assumptions above is that only the wine physicochemical indexes can affect the quality of wine. We emphasize the relations between the physicochemical indexes and the quality of red wine while ignoring others.

OLSR is not applicable to the situation of multicollinearity in the calibration set. It is not utilized when the number of samples is smaller than the number of variables. In this paper, OLSR is not employed to establish the model.

The other multivariate statistic methods based on soft measurement, including PCR, PLSR, and MPLSR, are utilized to establish models. We propose the root mean squared error of calibration (RMSEC) and root mean error of prediction (RMSEP); RMSEC evaluates the fitting efficiency and RMSEP

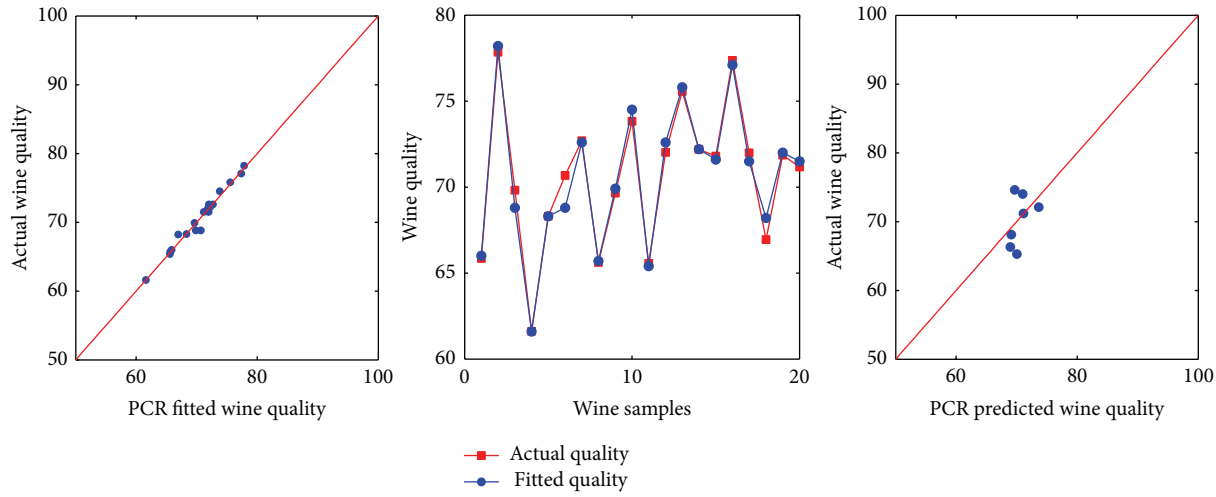

Figure 1: PCR fitted quality versus actual quality and predicted quality versus actual quality.

demonstrates the prediction ability of different methods [9]. RMSEC and RMSEP can be described as follows:

$$\text{RMSEC} = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_{\text{cal},i}\right)^2}{N}}$$
$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{M}\left(y_i - \hat{y}_{\text{pred},i}\right)^2}{M}}, \quad (13)$$

where y_i is the measured value for the ith sample; $\hat{y}_{\text{cal},i}$ and $\hat{y}_{\text{pred},i}$ are the method fitted and model predicted response values for the ith sample, respectively; and N and M are the number of calibration samples and verification samples, respectively.

The CPV[2] (cumulative percent variance) [19] is firstly employed to select the number of PCs with 98% as CPV and 17 PCs are obtained for PCR. With the help of cross validation [22, 25], 4 LVs are chosen for PLSR. There is no need for MPLSR to select the PCs or LVs that are the main difficulties for PCR and PLSR in maintaining the integrity of information from the origin data set.

According to Figures 1, 2, and 3 and Table 3, these different multivariate statistical methods will significantly impact the effect of the predicting models. Due to the problems of multicollinearity and the less number of samples than variables in the calibration set, the OLSR is totally not fitted to establish the prediction model. In order to reach the minimal values of RMSEC, MPLSR obtains a satisfying result for the fitting efficiency. However, PCR, PLSR, and MPLSR have a similar prediction ability in this paper.

3.3. Contribution Ratio of Wine Physicochemical Indexes. In this paper, 87 wine physicochemical indexes are collected in the calibration set and these indexes can provide adequate information. The large wine physicochemical indexes to be measured are a time consuming and complex work. We all know some wine physicochemical indexes do not contribute to the wine quality. In order to avoid doing lengthy and complex calculations, it is necessary to analyze the contribution of every wine physicochemical index and select the main wine physicochemical indexes.

In order to analyze the contribution of every grape physicochemical index to the wine quality, a contribution ratio (CR) is introduced and it can be formulated as follows:

$$y - \beta_0 = BX, \qquad \text{CR}_i = \frac{\beta_i x_i}{y - \beta_0}, \quad i = 1 \cdots m,$$
$$B = [\beta_1, \beta_2 \cdots \beta_m] \in R^{1\times m}, \quad (14)$$
$$X = [x_1, x_2 \cdots x_m]^T \in R^{1\times m},$$

where B is the coefficient matrix of regression equation and β_0 is a constant. The benefit of this definition for CR is that the sum of all CR_i ($i = 1 \cdots m$) equals 1.

The figure, contribution of wine physicochemical indexes, is shown in Figure 4. As can be seen from the figure, the great influence on the wine quality is cis-resveratrol, color, acetic acid, 2-methylpropyl ester, butanedioic acid, diethyl ester, and so forth. On the contrary, some physicochemical indexes can be ignored, such as transresveratrol, TR, limonene, 1-butanol,3-methyl, styrene, and so forth. Since some variables can be controlled in the production process, this information can be used to improve the wine quality. The detailed information is shown in [9].

4. Conclusion

In this paper, multivariate methods based on soft measurement, ordinary least squares regression (OLSR), principal component regression (PCR), partial least squares regression (PLSR), and modified partial least squares regression (MPLSR) are reviewed. With these methods and the real data obtained in practice, wine quality prediction models have been constructed. One purpose of this work is to choose the best method among OLSR, PLS, PCR, and MPLSR. The

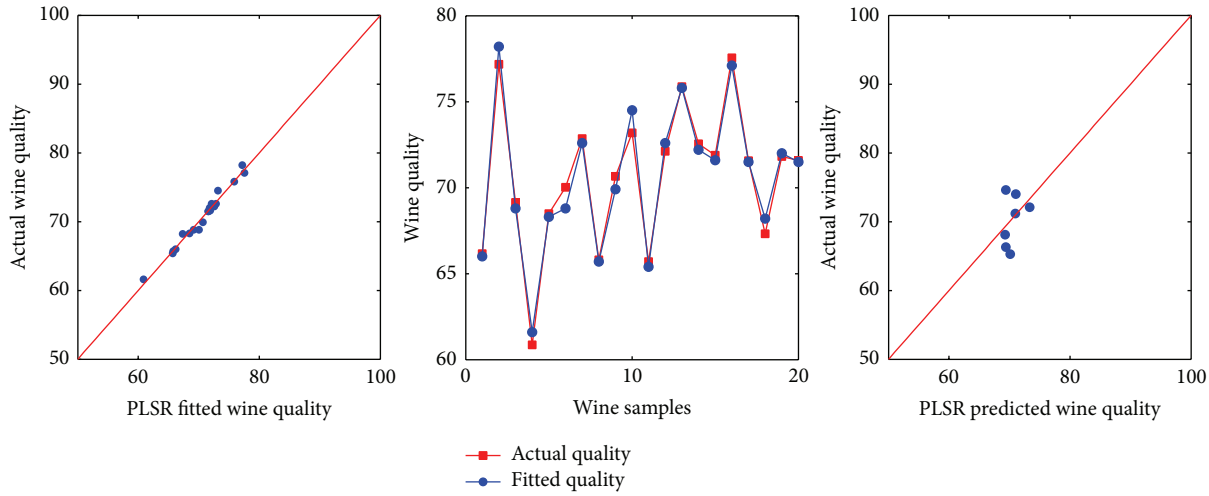

Figure 2: PLSR fitted quality versus actual quality and predicted quality versus actual quality.

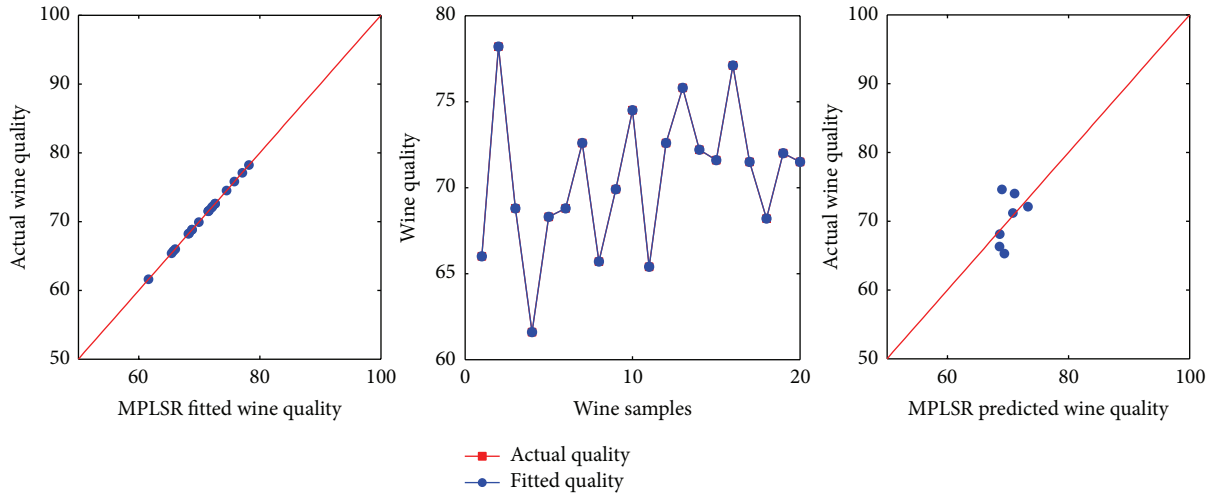

Figure 3: MPLSR fitted quality versus actual quality and predicted quality versus actual quality.

model built by MPLSR performs the best among the four models with the superior fitting efficiency represented by RMSEC. The models built by PCR, PLSR, and MPLSR have a similar prediction ability represented by RESEP. Several wine physicochemical indexes have significant contributions to the final wine quality while others are insignificant. The efficiency of the MPLSR model is validated on a larger data set, while this method is based on the objective tests, aiding the speed and quality of the oenologist performance. The outcome of the work is useful for the wine industry. From the perspective of industries, they can save manpower and financial resources in this method.

With the science and technology transformed into productivity, the advanced technology is widely used in industry. At the same time, soft measurement is developing. Soft measurement has achieved comprehensive results in the process control research. This technology breaks the traditional pattern of single input single output (SISO). The computer technology is conducive to soft measurement, improving the availability of soft-sensing technique and reducing the difficulty of application. We believe that soft measurement will have a wide application prospect, with the integration of computer technology and advanced technology.

Table 3: RMSEC and RMSEP values.

	PCR	PLSR	MPLSR
RMSEC	0.62	0.60	0
RMSEP	3.07	3.20	3.01

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

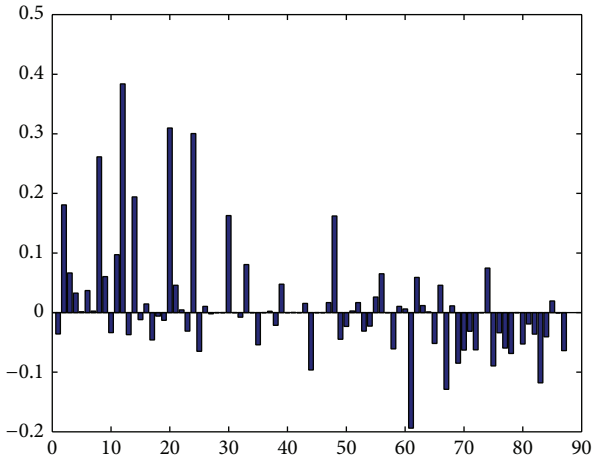

Figure 4: Contribution ratio of wine physicochemical indexes in MPLSR.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61304102) and the Natural Science Foundation of Liaoning Province, China (no. 2013020002).

References

[1] G. Schamel and K. Anderson, "Wine quality and varietal, regional and winery reputations: hedonic prices for Australia and New Zealand," *Economic Record*, vol. 79, no. 246, pp. 357–369, 2003.

[2] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, "Modeling wine preferences by data mining from physicochemical properties," *Decision Support Systems*, vol. 47, no. 4, pp. 547–553, 2009.

[3] S. Buratti, S. Benedetti, M. Scampicchio, and E. C. Pangerod, "Characterization and classification of Italian Barbera wines by using an electronic nose and an amperometric electronic tongue," *Analytica Chimica Acta*, vol. 525, no. 1, pp. 133–139, 2004.

[4] A. van den Bos, "Application of statistical parameter estimation methods to physical measurements," *Journal of Physics E: Scientific Instruments*, vol. 10, no. 8, article 753, 1977.

[5] J. Liang and J. Qian, "Multivariate statistical process monitoring and control: recent developments and applications to chemical industry," *Chinese Journal of Chemical Engineering*, vol. 11, no. 2, pp. 191–203, 2003.

[6] J.-M. Riviere, M. Bayart, J.-M. Thiriet, A. Boras, and M. Robert, "Intelligent instruments: some modelling approaches," *Measurement and Control*, vol. 29, no. 6, pp. 179–186, 1996.

[7] J. Yu and C. Zhou, "Soft-sensing techniques in process control," *Control Theory and Applications*, vol. 13, no. 2, pp. 137–144, 1996.

[8] J. Yu, "Soft-sensing technology and its application," *Process Automation Instrumentation*, vol. 29, no. 1, pp. 1–7, 2008.

[9] S. Yin, X. Zhu, and H. R. Karimi, "Quality evaluation based on multivariate statistical methods," *Mathematical Problems in Engineering*, vol. 2013, Article ID 639652, 10 pages, 2013.

[10] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine*, vol. 2, no. 6, pp. 559–572, 1901.

[11] P. J. A. Shaw, *Multivariate Statistics for the Environmental Sciences*, Hodder Arnold, London, UK, 2003.

[12] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[13] J. E. Jackson, *A User's Guide to Principal Components*, vol. 587, John Wiley & Sons, New York, NY, USA, 2005.

[14] I. Jolliffe, "Principal component analysis," in *Encyclopedia of Statistics in Behavioral Science*, John Wiley & Sons, New York, NY, USA, 2005.

[15] S. Weisberg, *Applied Linear Regression*, vol. 528, John Wiley & Sons, New York, NY, USA, 2005.

[16] C. Xie, D. Chen, R. Zheng, and Z. Xie, "Principal component analysis on aroma constituents of seven higharoma pattern oolong teas," in *Journal of South China Agricultural University*, vol. 20, pp. 113–117, 1998.

[17] J. Yang, H. Tong, and L. Jia, "Principal composition analysis of sensory and physiochemical quality of fermented bean curd," in *Transactions of the Chinese Society of Agricultural Engineering*, vol. 2, 2002.

[18] Y. Bao, Z. Hu, Y. Bai, and R. Guo, "Application of principal component analysis and cluster analysis to evaluating ecological safety of land use," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 22, no. 8, pp. 87–90, 2006.

[19] S. Valle, W. Li, and S. J. Qin, "Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods," *Industrial & Engineering Chemistry Research*, vol. 38, no. 11, pp. 4389–4401, 1999.

[20] P. Geladi and B. R. Kowalski, "Partial least-squares regression: a tutorial," *Analytica Chimica Acta*, vol. 185, pp. 1–17, 1986.

[21] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.

[22] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 58, no. 2, pp. 109–130, 2001.

[23] S. Yin, S. X. Ding, P. Zhang, A. Hagahni, and A. Naik, "Study on modifications of pls approach for process monitoring," *Threshold*, vol. 2, pp. 12389–12394, 2011.

[24] G.-F. Wu, Y.-H. Jiang, Y.-Y. Wang, and Y. He, "Discrimination of varieties of dry red wines based on independent component analysis and BP neural network," *Spectroscopy and Spectral Analysis*, vol. 29, no. 5, pp. 1268–1271, 2009.

[25] X. Su, Z. Li, Y. Feng, and L. Wu, "New global exponential stability criteria for interval-delayed neural networks," *Journal of Systems and Control Engineering*, vol. 225, no. 1, pp. 125–136, 2011.