

## Use of Visible and Near Infrared Spectroscopy and Least Squares-Support Vector Machine To Determine Soluble Solids Content and pH of Cola Beverage

FEI LIU AND YONG HE\*

College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China

Visible and near infrared (VIS/NIR) transmission spectroscopy and chemometric methods were utilized for the fast determination of soluble solids content (SSC) and pH of cola beverage. A total of 180 samples were used for the calibration set, whereas 60 samples were used for the validation set. Some preprocessing methods were applied before developing the calibration models. Several PLS factors, extracted by partial least squares (PLS) analysis, were used as the inputs of least squares-support vector machine (LS-SVM) model according to their accumulative reliabilities. The correlation coefficient ( $r$ ), root mean square error of prediction (rmsEP), bias, and RPD were 0.959, 1.136,  $-0.185$ , and 3.5 for SSC, whereas 0.973, 0.053, 0.017, and 4.1 for pH, respectively. An excellent prediction precision was achieved by LS-SVM compared with PLS. The results indicated that VIS/NIR spectroscopy combined with LS-SVM could be applied as a rapid and alternative way for the fast determination of SSC and pH of cola beverage.

**KEYWORDS:** VIS/NIR spectroscopy; cola beverage; soluble solids content (SSC) and pH; partial least squares (PLS); least squares-support vector machine (LS-SVM)

### INTRODUCTION

Cola beverage is a kind of soft drink and is very popular all over the world. Cola beverage has a dark brown color and sweet taste. It is composed of carbonated water, caramel, phosphoric acid, flavor, and nutritive sweeteners such as high fructose corn syrup, medium invert syrup, and liquid sucrose, whereas white granulated sugar might be used for certain cola beverages sourced in China. To extend the shelf life or improve the product taste, antioxidants, preservatives, and sweeteners are often added to food singly or in combination as food additives (1). The carbonation, low pH, and to a certain extent, the flavor are important to preserve in the cola beverage. While the addition of a specific preservative, such as sodium benzoate, is commonly used in cola beverages containing non-nutritive sweeteners. The soluble solids content and pH value are two important internal quality indices of cola beverages which have certain effects on the taste and flavor of drinking cola. With the improvement of peoples' living standard, people pay more attention to the health effects of drinking cola which contains lots of sugar that can cause weight problems. Some researchers studied some chemical components in cola soft drinks using different methods. Campuzano et al. (2) determined the glucose and fructose in cola using a bienzyme biosensor. Boyce (1) measured the aspartame, acesulfame, and benzoic acid levels using mixed micellar electrokinetic chromatography. However, these methods were time consuming, laborious, and costly. There were no reports

in the literature about the measurement of SSC and pH of cola using visible and near infrared (VIS/NIR) spectroscopy. Hence, it was necessary to develop a rapid and alternative method for the determination of SSC and pH of cola beverages using VIS/NIR spectroscopy.

VIS/NIR spectroscopy has attracted considerable attention and has been widely used as a rapid, low cost, and high precision analytical method in many fields, such as agriculture, pharmaceuticals, food, textiles, cosmetics, and polymer production industries (3). In the food industry, some researchers had applied VIS/NIR spectroscopy for the determination of particular chemical constituents which had strong influences in beverages, such as ethanol, sugars, organic acids, and phenolic compounds (4–7). Several researchers had applied VIS/NIR spectroscopy to determine soluble solids contents of fruit products in orange juice (8) and bayberry juice (9). Urbano-Cuadrado et al. (4) used near infrared reflectance spectroscopy and multivariate analysis to determine 15 parameters in different types of wine, such as ethanol, reducing sugars, total acidity, and pH. Iñón et al. (10) used middle and near infrared spectroscopy for the determination of the quality properties of beers. Yu et al. (11) used visible and near infrared spectroscopy to access the Brix and pH values of rice wines. Luis et al. (12) developed a rapid analysis of sugars in fruit juices by FT-NIR spectroscopy. Rambla et al. (13) applied PLS-NIR for the determination of total sugar, glucose, fructose, and sucrose in aqueous solutions of fruit juices. The calibration methods in the literatures were commonly used PLS analysis. PLS analysis only dealt with linear problems, which meant that PLS only took the linear

\* To whom correspondence should be addressed. Phone: +86-571-86971143. Fax: +86-571-86971143. E-mail: yhe@zju.edu.cn

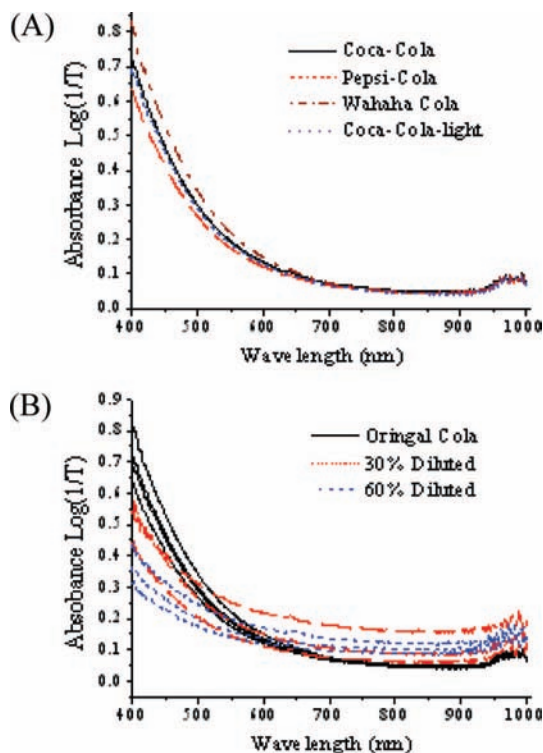


Figure 1. Absorbance spectra of (A) original cola and (B) diluted cola

Table 1. Statistic Values of SSC and pH in the Calibration and Validation Sets

parameters	data set	$n^a$	range	mean	SD <sup>b</sup>
SSC (°Brix)	calibration	180	0.4–11.6	5.997	3.964
	validation	60	0.4–11.6	5.965	3.968
	all	240	0.4–11.6	5.989	3.957
pH (pH unit)	calibration	180	2.562–3.289	2.907	0.215
	validation	60	2.568–3.289	2.909	0.218
	all	240	2.562–3.289	2.907	0.215

<sup>a</sup>No. of samples; <sup>b</sup>SD, standard deviation.

information of the spectral data into consideration. However, some latent nonlinear information existed in the spectral data and it was also helpful for the calibration. Recently, some nonlinear calibration methods were applied for the NIR quantitative analysis (14). Support vector machine (SVM) was a state-of-the-art statistical learning method proposed by Vapnik (15), which had a good theoretical foundation in statistical learning theory based on the principal of structural risk minimum (SRM). SVM had been applied in many aspects (16–18). Chauchard et al. (19) applied least squares-support vector machine (LS-SVM) and NIR spectroscopy to develop a robust and portable sensor for acidity prediction in grapes.

The objective of this study was to use VIS/NIR spectroscopy for the rapid determination of soluble solids content and pH of cola beverage based on least squares-support vector machine (LS-SVM).

## MATERIALS AND METHODS

**Preparation of Cola Samples.** In this study, four varieties of colas were used for the experiments including Coca-Cola (The Coca-Cola Company), Pepsi-Cola (PepsiCo), Wahaha (Hangzhou Wahaha Group Co., Ltd.), and Coca-Cola-Light (The Coca-Cola Company). The first three varieties were ordinary and common colas in the market, whereas the fourth one, Coca-Cola-Light, was thought to be a health-conscious soft drink that would prevent weight gain because it contains little sugar. All the cola samples were stored in the laboratory with a constant

temperature of  $20 \pm 1$  °C for more than 48 h to equalize the temperature. Some pre-experiments were done using the reference methods mentioned below. The mean values for original Coca-Cola, Pepsi-Cola, Wahaha, and Coca-Cola-Light were 11.1, 11.5, 11.1, 0.5 °Brix for SSC, and 2.594, 2.605, 2.567, 2.927 pH units for pH, respectively. The ranges of SSC and pH for the first three kinds of colas were quite similar, whereas there was a large difference with Coca-Cola-Light. To have more ranges of SSC and pH values for the calibration analysis, the cola samples of Coca-Cola, Pepsi-Cola, and Wahaha were diluted by distilled water. For each variety of Coca-Cola, Pepsi-Cola, and Wahaha, 20 samples were original cola, 20 samples were diluted with 30% (v/v) distilled water, which meant 70% cola and 30% distilled water, and 20 samples were diluted with 60% (v/v) distilled water, which meant 40% cola and 60% distilled water. The original Coca-Cola-Light was used to prepare 60 samples because of the low SSC. Before the measurement of pH, a degassing step was applied to reduce the influence of carbonation in the cola beverages. Because the pH would increase when the carbonic acid was released in the form of free carbon dioxide. The pH of cola was measured after the beverage was thoroughly degassed by 15 min of sparging with air. To obtain a stable Brix or soluble solids measurement, the cola beverage was heated to accelerate the inversion of sucrose into the acidic beverage base. The soluble solids measurement was taken after the complete inversion and no more sucrose was left to invert. The sample number for each type was 60, and a total of 240 samples were prepared for further analysis. 180 cola samples (45 samples for each variety) were randomly selected for the calibration set, whereas the remaining 60 samples (15 for each variety) were separated for the validation set.

### Spectral Acquisition and Reference Methods for SSC and pH.

Three transmission spectra were collected for each sample by a hand-held VIS/NIR Spectroradiometers-FieldSpecHandHeld and FieldSpec Pro FR (325–1075 nm) /A110070, Trademarks of Analytical Spectral Devices, Inc. (Analytical Spectral Devices, Boulder, CO). The field-of-view (FOV) of the spectroradiometer is 25°. The light source consists of a Lowell Pro-Lam interior light source assemble/128930 with Lowell Pro-Lam 14.5V bulb/128690 tungsten halogen bulb that could be used in visible and near infrared region. The height of the light source was approximately 250 mm above the sample, and the distance between the light source and the sample depended on the energy of light source and what kind of sample was detected. Before 250 mm settled, some adjustments had been done according to the standard energy response curve. The probe was under the sample, and the distance between the sample and probe was 50 mm, and the value was fixed by the spectroradiometer fittings. The cola sample was placed in a cuvette with a 2 mm light path length. The transmission spectra from 325 to 1075 nm were measured at 1.5 nm intervals with an average reading of 30 scans for each spectrum. Three spectra were collected for each sample and the average spectrum of these three spectra was used in later analysis. All spectral data were stored in a computer and processed using the RS<sup>3</sup> software for Windows (Analytical Spectral Devices) designed with a graphical user interface.

The reference value of SSC was measured by an Abbenchttop refractometer (model: WAY-2S, Shanghai Precision & Scientific Instrument Co. Ltd., Shanghai, China). The refractive index accuracy is  $\pm 0.0002$  and the °Brix (%) range is 0–95% with temperature correction. The reference pH value was measured using a pH meter (model: PHS-4CT, Shanghai Dapu Instrument Co. Ltd., Shanghai, China), with accuracy of 0.001 pH unit.

**Preprocessing of Original Spectral Data.** In this study, the absorbance spectra were used by the transformation  $\log(1/T)$  of the transmission spectra. The average spectrum of the three absorbance spectra for each sample was employed and transformed into ASCII format by using the ASD ViewSpecPro software (Analytical Spectral Devices). Some preprocessing methods were processed before the calibration model. The smoothing way of Savitzky-Golay with three segments and zero polynomials was applied for denosing (20). Standard normal variate (SNV) was applied for light scatter correction and reducing the changes of light path length (21). The first-derivative with three segments was used to eliminate the baseline shift (22). To avoid low signal-to-noise ratio, only the region of wavelengths (400–1000

**Table 2.** Prediction Results of SSC and pH of Calibration Set by Different Models

calibration model		$R^2$	rmsEC	bias	slope	offset	RPD
SSC (°Brix)	PLS	0.967	1.001	$1.6 \times 10^{-6}$	0.936	0.384	2.4
	LS-SVM	0.996	0.376	$-1.6 \times 10^{-5}$	0.973	0.160	10.5
pH (pH units)	PLS	0.983	0.040	$-4.2 \times 10^{-8}$	0.966	0.048	3.2
	LS-SVM	0.984	0.038	$-2.2 \times 10^{-6}$	0.957	0.061	5.7

<sup>a</sup> Correlation coefficient of calibration model.

nm) were employed for the calculations. The pretreatments were implemented by "The Unscrambler V 9.6" software (CAMO PROCESS AS, Oslo, Norway).

**Partial Least Squares Analysis.** Partial least squares (PLS) regression is the most commonly used method for prediction using numerous correlated variables (such as VIS/NIR spectra). PLS analysis take the variable matrix  $Y$  (Chemical components) into consideration with variable matrix  $X$  (VIS/NIR spectra data) at the same time. Therefore, PLS analysis can be progressed to establish the regression model to predict the chemical components. To achieve this aim, the lines of matrix  $Y$  are used to calculate the components of matrix  $X$  in mathematics. Similarly at the same time, the lines of matrix  $X$  are used to predict the components of matrix  $Y$ . The details of the theory could be found in the literature (8). During the calculation, full cross-validation was used to prevent overfitting. Twenty PLS factors were calculated by PLS. Certain PLS factors were selected as input data set for the least squares-support vector machine according to their accumulative reliabilities.

**Least Squares-Support Vector Machine.** Least squares-support vector machine (LS-SVM) is a new and attractive statistical learning method. LS-SVM has the capability of dealing with linear and nonlinear multivariate calibration and resolving these problems in a relatively fast way (15, 23). Moreover, SVM is capable of learning in high-dimensional feature space with fewer training data. It employs a set of linear equations instead of quadratic programming (QP) problems to obtain the support vectors (SVs). The details of LS-SVM algorithm could be found in the literature (18, 24). The LS-SVM regression model can be expressed as

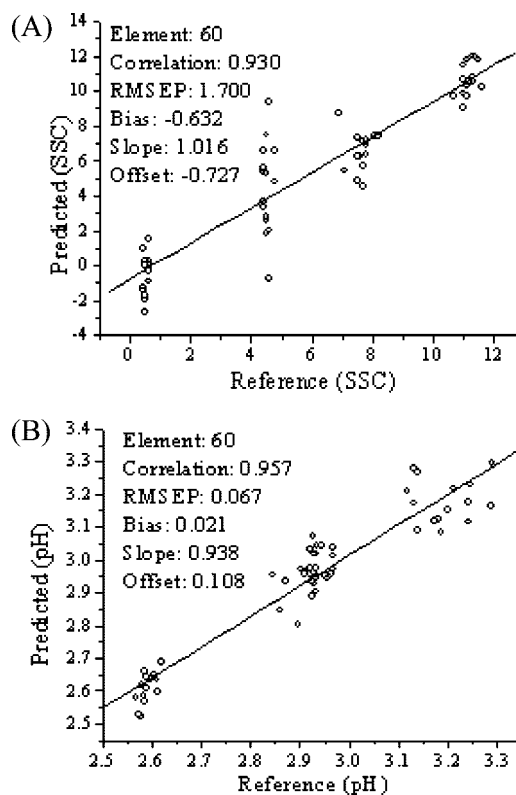
$$y(x) = \sum_{i=1}^n \alpha_i K(x, x_i) + b \quad (1)$$

Where  $K(x, x_i)$  is the kernel function. All the calculations were performed using MATLAB 7.0 (The Math Works, Natick, MA). The free LS-SVM toolbox (LS-SVM v 1.5, Suykens, Leuven, Belgium) was applied with MATLAB 7.0 to develop the calibration models.

The predictive performance of the model was evaluated by the following standards: correlation coefficient ( $r$ ), root mean square error of calibration (rmsEC) and prediction (rmsEP), and RPD which was used to evaluate how well the calibration model could predict compositional data. RPD value was the ratio of the standard deviation of the reference data (SD) to the standard error of cross-validation (SECV) or the standard error of prediction (SEP) (25). The RPD should ideally be at least three. Some other standards, such as slope and bias, should be taken into consideration for distinguishing systematic errors and studying the correlation between the reference and VIS/NIR models. A good model should have high correlation coefficient and RPD values, low rmsEC, and rmsEP. A relatively high RPD value indicated that the model was able to reliably predict the chemical components, and a stable relationship developed between the spectral data and chemical compositions.

## RESULTS AND DISCUSSION

**Overview of VIS/NIR Spectra and Statistical Values of SSC and pH.** Figure 1 shows the absorbance spectra of cola beverage in the experiments. The trends of the original spectra were quite similar (as shown in Figure 1A), but some latent differences and features existed according to the chemical components and color variance which could not be distinguished by the naked eye. With a close observation, some crossovers



**Figure 2.** Prediction results for (A) SSC and (B) pH in validation sets by PLS model

were in existence in the region 500–650 nm, and a small absorbance peak was in the region 930–1000 nm. The region 500–600 nm might have resulted from the color variance caused by the pigment, flavor, and caramel (26). The band around 930 nm might have resulted from C–H stretch third overtone form sugars, and the band around 990 nm with a small absorbance peak was produced by the O–H stretch second overtone form sugars and organic acids (27). Figure 1B shows the spectra of diluted cola. Some obvious variances were found between 400 and 600 nm because of the color difference caused by the diluting practice. To make good use of the spectra, some pretreatments aforementioned and multivariate methods should be utilized for further calculations. Table 1 summarizes the statistical values of SSC and pH in calibration and validation sets.

**PLS analysis and Extraction of PLS factors.** The PLS model was developed using the preprocessed spectral data. The calibration models were developed with 10 PLS factors for both SSC and pH. The performance was validated by the samples in validation set. The prediction results of calibration sets for SSC and pH by PLS were shown in Table 2. The RPD value of SSC for the calibration set was less than three (the RPD value was 2.4), and this value was similar with the literature (the RPD value for °Brix was 2.34) reported by Yu et al. (11). The correlation coefficient ( $r$ ), root mean square error of prediction (rmsEP), bias and RPD for prediction set were 0.930, 1.700,

**Table 3.** The Accumulative Reliabilities (AR) of PLS Factors for SSC and pH

PLS factors	1	2	3	4	5	6	7	8	9	10	11
AR(%) (SSC)	18.82	35.73	50.14	59.47	65.29	71.20	76.48	79.26	81.83	83.41	84.42
AR(%) (pH)	55.52	65.28	72.35	78.09	82.43	85.43	87.97	89.36	90.57		

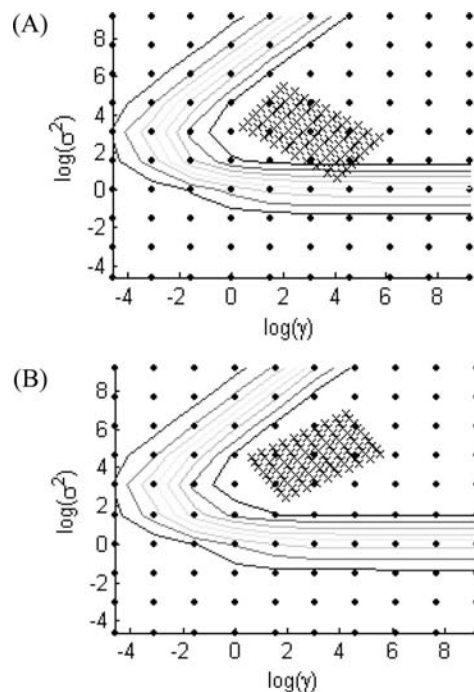
-0.632, and 2.3 for SSC, whereas 0.957, 0.067, 0.021, and 3.3 for pH, respectively (shown in **Figure 2A** for SSC and **Figure 2B** for pH). Simultaneously, the computation time for PLS calibration stage was about 120 s for both SSC and pH.

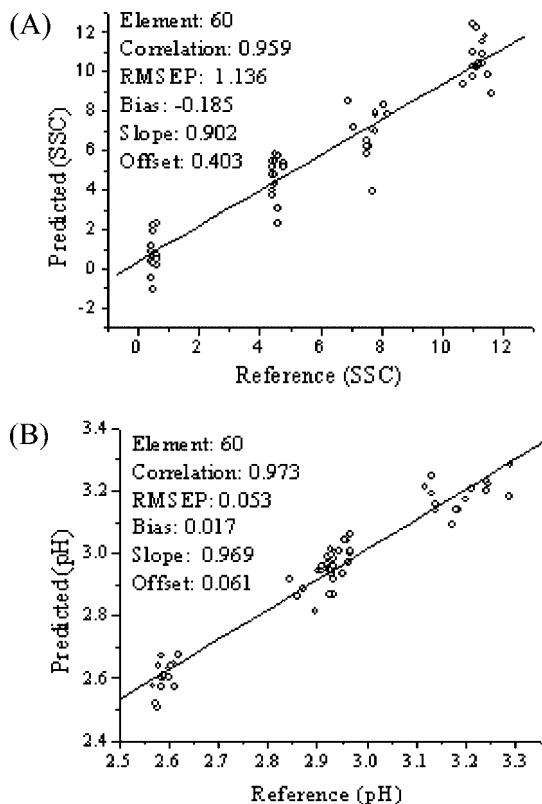
From the analysis of PLS, 20 PLS factors were calculated for all of 240 samples, and certain PLS factors would be selected as input data set of the least squares-support vector machine (LS-SVM). Although the whole spectral wavelength region (400–1000 nm) could be applied as the inputs, the training time increased with the square of the number of training samples and linearly with the number of variables (dimension of spectra) (19). Therefore, certain PLS factors were selected to improve the training speed and enhance the features of spectra and to reduce the dimensionality of the spectral data matrix and training error. The principle for the selection was that the selected PLS factors should explain the most variances and represent the main information of the raw spectra. The accumulative reliability could be perceived as such indices. The accumulative reliabilities of certain PLS factors for SSC and pH were shown in **Table 3**. It indicated that the accumulative reliabilities of the first eleven PLS factors for prediction of SSC could explain 84.423% of the total variance of the raw spectra. The contribution of first twelve PLS factors was 85.347% and the 12th PLS factor interpreted an additional 0.924%, which contributed not so much as the first 11 PLS factors. So the first 11 PLS factors were regarded as the inputs of the LS-SVM models for SSC, and the first 11 PLS factors were also suggested by the PLS analysis implemented by “The Unscrambler V9.6” software after the calculations of all the 240 samples. For the same explanation, the first nine PLS factors were applied as the inputs of LS-SVM for the pH prediction since the tenth PLS factor explained an additional 0.918% of the total variance. Therefore, the time for calibration should be considerably decreased and the precision would be improved.

**Development and Performance of LS-SVM Model.** When using LS-SVM, three crucial problems were required to be solved, including the determination of the optimal input feature subset, proper kernel function, and the optimal kernel parameters (24). Currently, there was no systematic methodology for the selection of kernel function. However, compared with other feasible kernel functions, radial basis function (RBF) as a nonlinear function was a more compacted supported kernel and able to reduce the computational complexity of the training procedure and give a good performance under general smoothness assumptions (28). RBF function could also handle the nonlinear relationships between the spectra and target attributes. Thus, RBF kernel was recommended as the kernel function of LS-SVM in this paper. In addition, proper parameter setting played a crucial role in building a good LS-SVM regression model with high prediction accuracy and stability. The regularization parameter  $\gamma$  determined the tradeoff between minimizing the training error and minimizing model complexity. The parameter  $\text{sig}2$  ( $\sigma^2$ ) of RBF kernel function was the bandwidth and implicitly defined the nonlinear mapping from input space to some high dimensional feature space. A two-step grid search technique with leave-one-out cross validation was employed to obtain the optimal parameter values.

The optimal combination of ( $\gamma$ ,  $\sigma^2$ ) within the region of ( $10^{-2}$ – $10^4$ ) was set based on experience. Grid search tried values of each parameter across the specified search range using geometric steps. Leave-one-out cross validation was used to avoid over-fitting. The first step of grid search was for a crude search with a large step size presented in the form of (•) and the second step for the specified search with a small step size presented in the form of (×). The results of the two-step grid search were shown in **Figure 3A** for SSC and **Figure 3B** for pH. The optimal combination of ( $\gamma$ ,  $\sigma^2$ ) was achieved with  $\gamma = 13.480$  and  $\sigma^2 = 17.533$  for SSC and  $\gamma = 63.319$  and  $\sigma^2 = 167.590$  for pH, respectively. The calibration model was developed using the aforementioned parameters and the prediction results for calibration set were shown in **Table 2**. The correlation coefficient ( $r$ ), root mean square error for prediction (rmsEP), bias and RPD values were used as evaluation criterions for the performance.

A total of 60 samples in the validation set were predicted by the LS-SVM model. The prediction results for SSC and pH were shown in **Figure 4A** for SSC and **Figure 4B** for pH. The correlation coefficient, rmsEP, bias and RPD were 0.959, 1.136, -0.185, and 3.5 for SSC, while 0.973, 0.053, 0.017, and 4.1 for pH by LS-SVM model, respectively. The computation time for LS-SVM calibration with two-step grid search was about 60 s for SSC and 50 s for pH. Excluding the grid search time, the computation time was only about 2 s. Once the parameters of  $\gamma$  and  $\sigma^2$  were settled, there was no need to implement the grid search step. Therefore, the computation time of LS-SVM was much less than that of PLS. Furthermore, compared with **Figure 2** and **Figure 4**, the scatter plots of LS-SVM models were much closer to the regression line than that of PLS models. Hence, the prediction results showed that LS-SVM outperformed

**Figure 3.** Two-step grid search for the optimal ( $\gamma$ ,  $\sigma^2$ ) for the prediction of (A) SSC and (B) pH



**Figure 4.** Prediction results for (A) SSC and (B) pH in validation sets by LS-SVM model

PLS model and LS-SVM model obtained an excellent prediction precision. The prediction accuracy by LS-SVM was also better than that described the similar studies. Urbano-Cuadrado et al. (4) predicted reducing sugars, pH, and other parameters in different types of wines by using NIR reflectance spectroscopy (the  $r$  values for reducing sugar and pH were 0.844 and 0.905, respectively). Apetrei et al. (29) applied electronic tongue to predict the reducing sugars ( $r = 0.843$ ) in red wines. Cen et al. (8) applied VIS/NIRS and PLS for the determination of pH value ( $r = 0.96$ ) of orange juice. Shao and He (9) applied VIS/NIR spectroscopy to measure the SSC ( $r = 0.85$ ) and pH ( $r = 0.92$ ) in bayberry juice. Yu et al. (11) used Fourier transform near infrared spectroscopy to qualify the pH value ( $r = 0.906$ ) in Chinese rice wine. The overall results demonstrated that VIS/NIR spectroscopy combined with LS-SVM could obtain an excellent precision and accuracy for the determination of SSC and pH of cola beverage.

In conclusion, the determination of soluble solids content and pH of cola beverage could be successfully performed through VIS/NIR spectroscopy based on least squares-support vector machine (LS-SVM). The smoothing way of Savitzky-Golay, Savitzky-Golay first-derivative, and standard normal variate (SNV) were used as the preprocessing methods. Certain PLS factors extracted by PLS were used as the inputs of LS-SVM models according to their accumulative reliabilities. RBF kernel was used as the kernel function and a two-step grid search technique were used for the selection of optimal combination of ( $\gamma$ ,  $\sigma^2$ ). An excellent precision and accuracy was achieved by LS-SVM models compared with PLS models. Further interpretation of the parameter optimization and variable selection would be needed to improve the calibration accuracy, stability, and generalization.

## LITERATURE CITED

- Boyce, M. C. Simultaneous determination of antioxidants, preservatives and sweeteners permitted as additives in food by mixed micellar electrokinetic chromatography. *J. Chromatogr., A* **1999**, *847*, 369–375.
- Campuzano, S.; Loaiza, Ó. A.; Pedrero, M.; De Villena, F. J. M.; Pingarrón, J. M. An integrated bienzyme glucose oxidase-fructose dehydrogenase-tetrathiafulvalene-3-mercaptopropionic acid- gold electrode for the simultaneous determination of glucose and fructose. *Bioelectrochemistry* **2004**, *63*, 199–206.
- Yan, Y. L.; Zhao, L. L.; Han, D. H.; Yang, S. M. *The Foundation and Application of Near-infrared Spectroscopy Analysis*; China Light Industry Press: Beijing, 2005; p 1–3.
- Urbano-Cuadrado, M.; De Castro, M. D. L.; Pérez-Juan, P. M.; García-Olmo, J.; Gómez-Nieto, M. A. Near infrared reflectance spectroscopy and multivariate analysis in enology determination or screening of fifteen parameters in different types of wines. *Anal. Chim. Acta* **2004**, *527*, 81–88.
- Gozzolino, D.; Smyth, H. E.; Gishen, M. Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. *J. Agric. Food Chem.* **2003**, *51*, 7703–7708.
- Contal, L.; Leon, V.; Downey, G. Detection and quantification of apple adulteration in strawberry and raspberry purees using visible and near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2002**, *10*, 289–299.
- Cozzolino, D.; Kwiatkowski, M.; Parker, M.; Gishen, M.; Damborgs, R. G.; Cynkar, W.; Herderich, M. Prediction of phenolic compounds in red wine by near infrared spectroscopy. *Anal. Chim. Acta* **2004**, *513*, 73–80.
- Cen, H. Y.; He, Y.; Huang, M. Measurement of soluble solids contents and pH in orange juice using chemometrics and VIS/NIRS. *J. Agric. Food Chem.* **2006**, *54*, 7437–7443.
- Shao, Y. N.; He, Y. Nondestructive measurement of the internal quality of bayberry juice using Vis/NIR spectroscopy. *J. Food Eng.* **2007**, *79*, 1015–1019.
- Iñón, F. A.; Garrigues, S.; De la Guardia, M. Combination of mid- and near-infrared spectroscopy for the determination of the quality properties of beers. *Anal. Chim. Acta* **2006**, *571*, 167–174.
- Yu, H. Y.; Ying, Y. B.; Fu, X. P.; Lu, H. S. Quality determination of Chinese rice wine based on Fourier transform near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2006**, *14*, 37–44.
- Luis, E. R.; Fredrick, S. F.; Michael, A. M.; Elizabeth, M. C. Rapid analysis of sugars in fruit juices by FT-NIR spectroscopy. *Carbohydr. Res.* **2001**, *336*, 63–74.
- Rambla, F. J.; Garrigues, S.; Guardia, M. D. L. PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. *Anal. Chim. Acta* **1997**, *344*, 41–53.
- Pérez-Marín, D.; Garrido-Varo, A.; Guerenro, J. E. Non-linear regression methods in NIRS quantitative analysis. *Talanta* **2007**, *72*, 28–42.
- Vapnik, V. N. *The Nature of Statistical Learning Theory*; Springer-Verlag: New York, 1995.
- Belousov, A. I.; Verzakov, S. A.; Frese, J. von. Applicational aspects of support vector machines. *J. Chemom.* **2002**, *16*, 482–489.
- Thissen, U.; Pepers, M.; Üstün, B.; Mellssen, W. J.; Buydens, L. M. C. Comparing support vector machines to PLS for spectral regression applications. *Chemom. Intell. Lab. Syst.* **2004**, *73*, 169–179.
- Guo, H.; Liu, H. P.; Wang, L. Method for selecting parameters of least squares support vector machines and application (in Chinese). *J. Syst. Simul* **2006**, *18*, 2033–2036.

- (19) Chauchard, F.; Cogdill, R.; Roussel, S.; Roger, J. M.; Bellon-Maurel, V. Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemom. Intell. Lab. Syst.* **2004**, *71*, 141–150.
- (20) Gorry, P. A. General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Anal. Chem.* **1990**, *62*, 570–573.
- (21) Barnes, R.; Dhanoa, M.; Lister, J. Standard normal variable transformation and detrending of near infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777.
- (22) Chu, X. L.; Yuan, H. F.; Lu, W. Z. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique (in Chinese). *Prog. Chem.* **2004**, *16*, 528–542.
- (23) Suykens, J. A. K.; Vanderwalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300.
- (24) Chen, Q. S.; Zhao, J. W.; Fang, C. H.; Wang, D. M. Feasibility study on identification of green, black and Oolong teas using near-infrared reflectance spectroscopy based on support vector machine (SVM). *Spectrochim. Acta, Part A* **2007**, *66*, 568–574.
- (25) Williams, P. C. *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd ed; Williams P. C., Norris, K., Eds.; AACC, Inc: St. Paul, MN, 2001; p 145.
- (26) Somers, C. *The Wine Spectrum: An Approach Towards Objective Definition of Wine Quality*; Winetitles: Adelaide, **1998**; p 50.
- (27) Šašić, S.; Ozaki, Y. Short-wave near-infrared spectroscopy of biological fluids. 1. quantitative analysis of fat, protein, and lactose in raw milk by partial least-squares regression and band assignment. *Anal. Chem.* **2001**, *73*, 64–71.
- (28) Wang, W. J.; Xu, Z. B.; Lu, W. Z.; Zhang, X. Y. Determination of the spread parameter in the Gaussian kernel for classification and regression. *Neurocomputing* **2003**, *55*, 643–663.
- (29) Apetrei, C.; Apetrei, I. M.; Nevares, I.; Del Alamo, M.; Parra, V.; Rodríguez-Méndez, M. L.; De Saja, J. A. Using an e-tongue based on voltammetric electrodes to discriminate among red wines aged in oak barrels of aged using alternative methods: correlation between electrochemical signals and analytical parameters. *Electrochim Acta* **2007**, *52*, 2588–2594.

---

Received for review July 10, 2007. Revised manuscript received August 31, 2007. Accepted September 4, 2007. This study was supported by National Science and Technology Support Program (2006BAD10A04, 2006BAD10A09) the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of MOE, PRC, Natural Science Foundation of China (project no: 30671213).

JF072057B