

FTIR spectroscopy coupled with machine learning approaches as a rapid tool for identification and quantification of artificial sweeteners

Yu-Tang Wang^{a,b}, Bin Li^b, Xiao-Juan Xu^b, Hai-Bin Ren^b, Jia-Yi Yin^b, Hao Zhu^c, Ying-Hua Zhang^{a,b,*}

^a Key Laboratory of Dairy Science, Ministry of Education, Northeast Agricultural University, Harbin 150030, PR China

^b Department of Food Science, Northeast Agricultural University, Harbin 150030, PR China

^c The Rutgers Center for Computational and Integrative Biology, Camden, NJ 08102, United States



ARTICLE INFO

Keywords:

Mid-infrared spectroscopy
Chemometrics
Artificial sweeteners
Identification
Quantification

ABSTRACT

Fourier transform infrared (FTIR) spectroscopy calibrations were developed to simultaneously determine the multianalytes of five artificial sweeteners, including sodium cyclamate, sucralose, sodium saccharin, acesulfame-K and aspartame. By combining the pretreatment of the spectrum and principal component analysis, 131 feature wavenumbers were extracted from the full spectral range for modelling to qualitative and quantitative analysis. Compared to random forest, k nearest neighbour and linear discriminant analysis, support vector machine model had better predictivity, indicating the most effective identification performance. Furthermore, multivariate calibration models based on partial least squares regression were constructed for quantifying any combinations of the five artificial sweeteners, and validated by prediction data sets. As shown by the good agreement between the proposed method and the reference HPLC for the determination of the sweeteners in beverage samples, a promising and rapid tool based on FTIR spectroscopy, coupled with chemometrics, has been performed to identify and objectively quantify artificial sweeteners.

1. Introduction

Artificial high-intensity sweeteners provide the sensation of sweetness, but with little or no intake of food energy. Sweeteners as food additives are commonly used in bakery, beverage and confectionery products, which form an important class of non-nutritive components in food. Artificially-sweetened products are intensely promoted by the food industry because low-calorie food can help to avoid obesity, control body weight and manage diabetes (Edwards, Rossi, Corpe, Butterworth, & Ellis, 2016; Mooradian, Smith, & Tokuda, 2017). Consumption of sweeteners is increasing worldwide, and is expected to rise continually due to public health efforts to lower sugar intake. However, artificial sweeteners have caused controversy owing to suspicions of adverse health effects. The role of artificial sweeteners in the management of obesity is also controversial (Roberts, 2015). There are positive correlations between artificial sweeteners consumption with thyroid, reproductive systems, bone, and adipose tissue (Rother, Conway, & Sylvetsky, 2018; Setti et al., 2018).

For consumer safety, it is necessary to control the amount of sweeteners. The maximum level of sweeteners in foodstuffs is limited by specific regulations in many countries to ensure consumer health. In

order to achieve better management for the use of sweeteners, reliable and fast quantitative methods of analysis are required to measure levels of sweeteners. A number of analytical methods based on different principles have been developed for the analysis of artificial sweeteners (Zygler, Wasik, & Namieśnik, 2009). Of the great variety of methods to analyze artificial sweeteners, the chromatographic technique has received worldwide acceptance. Currently, the most popular method is high-performance liquid chromatography (HPLC). An analytical procedure involving the use of HPLC-tandem mass spectrometry has been developed for the determination of nine high-intensity sweeteners (Ordoñez, Rodil, Quintana, & Cela, 2015). Sucralose and acesulfame in well water were determined by ultra-HPLC-tandem mass spectrometry (Wu et al., 2014). The common sweeteners including aspartame, cyclamate, saccharin and acesulfame K, were determined by capillary electrophoresis with contactless conductivity detection (Stojkovic, Mai, & Hauser, 2013). The chromatographic separation and analysis of aspartame, saccharin and sucralose in water samples was developed using liquid chromatography/TOF mass spectrometry (Ferrer & Thurman, 2010). Considering the need for sophisticated equipment and specially trained personnel, it is necessary to develop a simple, rapid and reproducible analytical method to determine the artificial sweeteners.

* Corresponding author at: Key Laboratory of Dairy Science, Ministry of Education, Northeast Agricultural University, Harbin 150030, PR China.
E-mail address: yinghuazhang@neau.edu.cn (Y.-H. Zhang).

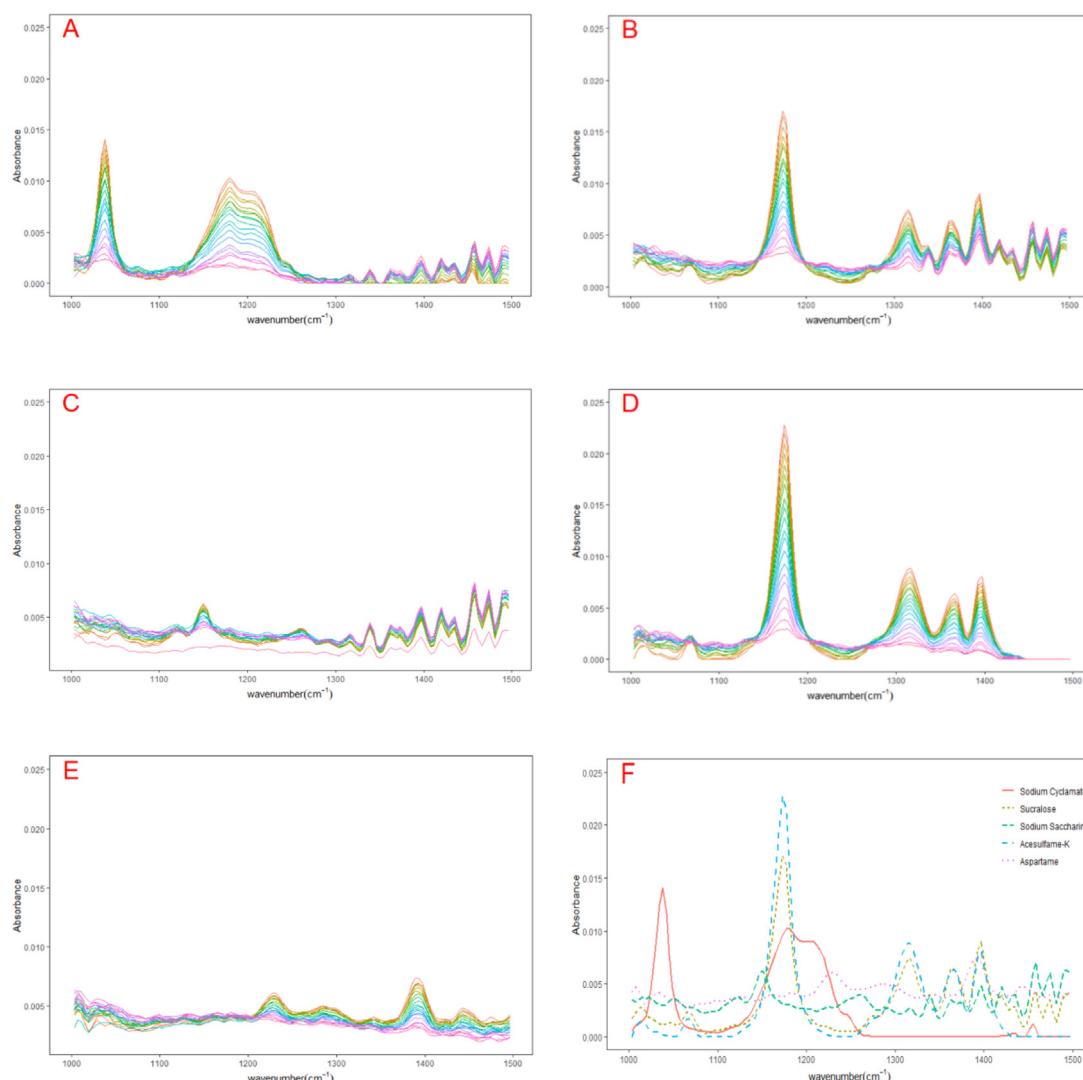


Fig. 1. Stacked FTIR spectra of five sweeteners in water at wavelengths between 1500 and 1000 cm⁻¹. A: Sodium cyclamate; B: Sucralose; C: Sodium saccharin; D: Acesulfame-K; E: Aspartame; F: Five sweeteners.

Sweeteners may be used separately or in combination with other sweeteners as mixtures to increase sweetening power due to synergistic effect (Behrens, Blank, & Meyerhof, 2017). Most of the analytical methods based on different principles have been developed for individual sweeteners. So far, the majority of the published multi-sweetener methods have focused on the determination of only few types of compounds (Zyglyer et al., 2009). There is a need to determine a mixture of several types of sweeteners in a single sample with one test.

The wide application of sweeteners has prompted a continuous search for new effective methods for the sweetener quantification. Fourier transform infrared (FTIR) spectroscopy as a non-destructive alternative to conventional methods has become an increasingly important tool in the quantitative analysis of complex components. Unlike many instrumental analytical techniques, FTIR does not require the dissolution or extraction of the samples, which substantially simplifies the analysis procedure and reduces the time needed for testing. In addition, the same sample with recorded IR spectra, can then be analyzed further by other tools. Previous studies have shown that FTIR spectroscopy has been effective in the quantification of several important factors in food samples, such as adulteration in camel milk (Mabood et al., 2017), fatty acid profile (Ferrand et al., 2011) and quick detection of pork (Abu-Ghoush et al., 2017). With advances in FTIR spectroscopy analysis technology, more data is available, providing an

opportunity to discover spectra and constituent concentration relationships by statistical modelling, which can lead to a better understanding of inputs and their impact on the analysis.

In order to present an accurate, fast and environmentally friendly analytical method for determination of artificial sweeteners that can be applied for quality control in industrial applications, this paper proposes a novel method of FTIR coupled with computational models to identify and quantify the multianalytes of artificial sweeteners. The method developed in this study, which combined qualitative and quantitative processing in the determination of compounds by multivariate analysis, can be employed to expand the application of FTIR.

2. Material and methods

2.1. Samples preparation

Five different sweeteners, sodium cyclamate ($C_6H_{12}NNaO_3S$), sucralose ($C_{12}H_{19}Cl_3O_8$), sodium saccharin ($C_7H_8NNaO_5S$), acesulfame-K ($C_4H_4KNO_4S$) and aspartame ($C_{14}H_{18}N_2O_5$), were purchased from Sigma (Saint Louis, MO, USA). Any combinations of two, three, four or five sweeteners were selected as samples. Each sweetener sample was dissolved in deionized water at 38 different concentrations. Sample concentration ranges were determined based on their standard

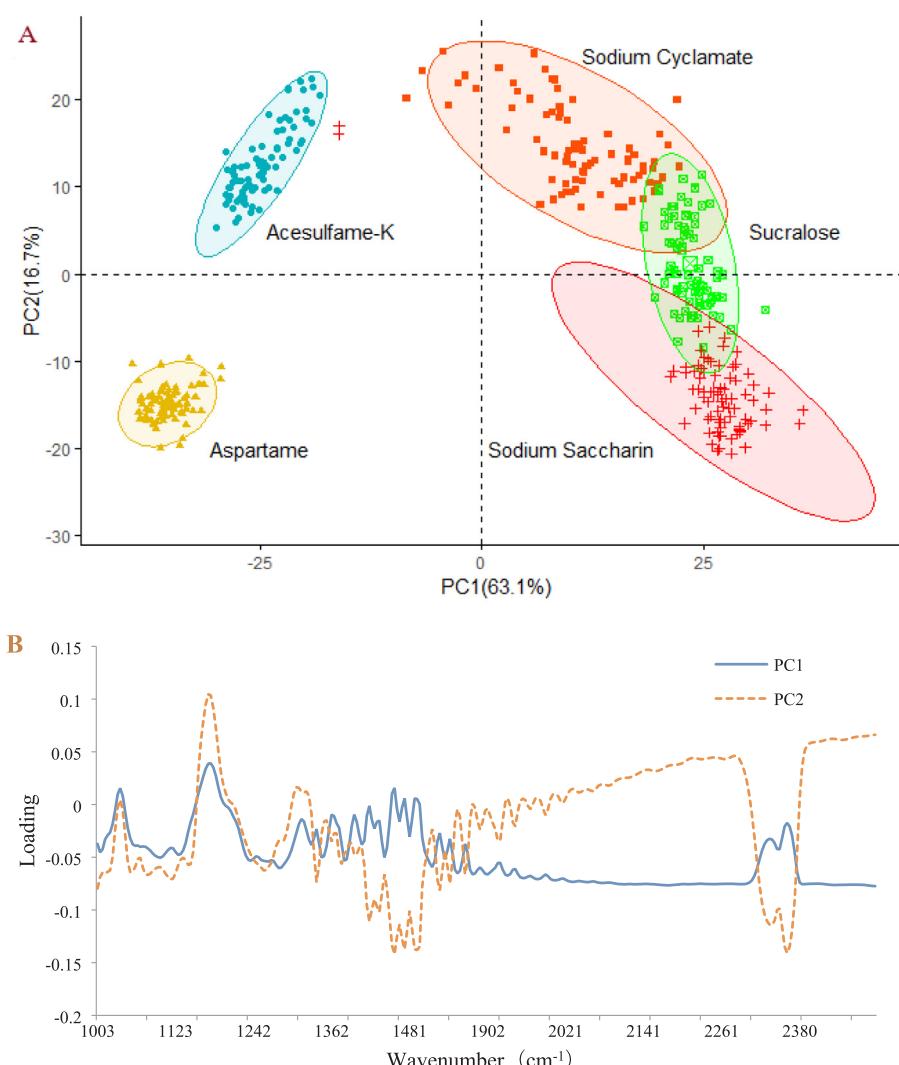


Fig. 2. (A) The PCA plot of 5 sweetener samples; (B) Loadings profiles for the PCs of the FTIR spectroscopic data (range: 2500–1000 cm^{-1}).

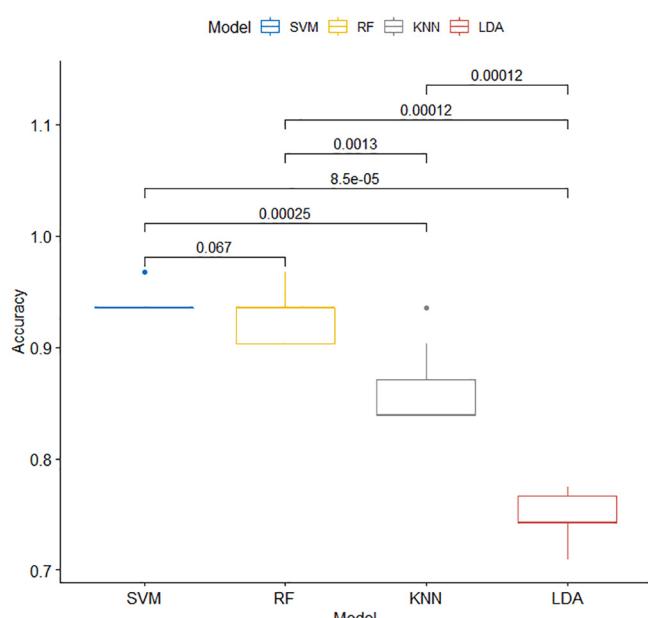


Fig. 3. Predictivity comparisons of the four classification models.

additions for sodium cyclamate as 0.05–1 g/kg, sucralose as 0.035–0.7 g/kg, sodium saccharin as 0.01–0.2 g/kg, acesulfame-K and aspartame as 0.05–1 g/kg, respectively. Total samples of different concentrations for each sweetener were prepared by gradient dilutions from the highest concentration sample.

Three diet tea drink samples containing different sweeteners were obtained from a local market.

2.2. Apparatus/spectra acquisition

FTIR spectra data of samples were collected using a PerkinElmer Frontier Optica FTIR Spectrometer (PerkinElmer, Norwalk, CT, USA). A total 20 scans per sample with a resolution of 0.4 cm^{-1} were performed in absorption mode within the wavelength range from 4000 to 400 cm^{-1} . All the samples were analyzed without any pre-treatment. Before measurement, a correction was performed to avoid atmospheric interference and reduce instrumental noise.

2.3. Spectra pre-processing

FTIR spectroscopic data were exported and processed by R (version 3.4.1) for multivariate analysis. In this work, two spectroscopic pre-processing methods, including Savitzky-Golay (SG), first derivative (1st D) and their combination were applied and compared. The number of smoothing points in our research was 3.

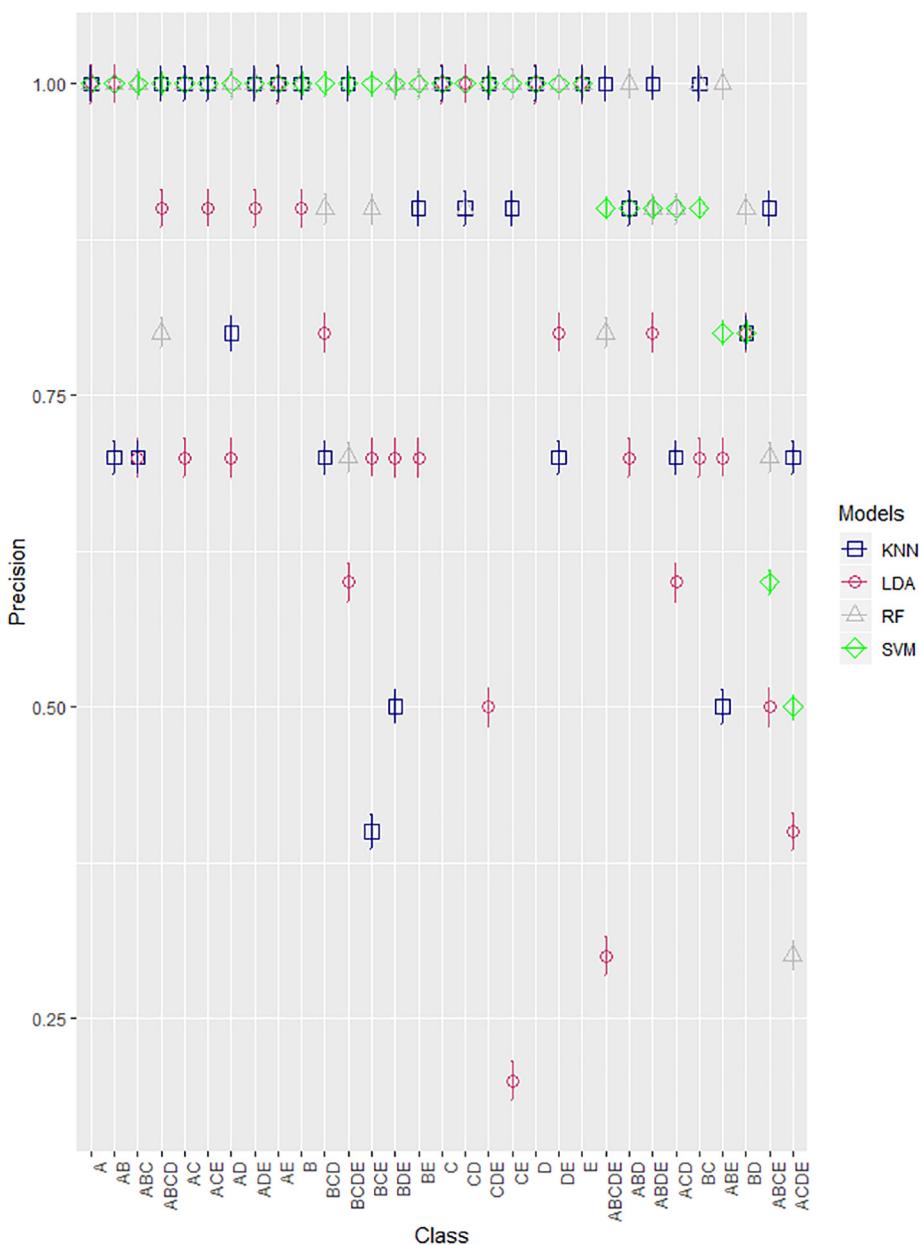


Fig. 4. The precision and recall of four classification models for the five sweeteners and their mixtures. A: Sodium cyclamate; B: Sucralose; C: Sodium saccharin; D: Acesulfame-K; E: Aspartame.

2.4. Characteristic wavelength selection

Principal component analysis (PCA) was implemented using the function “prcomp” from the “stats” R package (Śliwińska, Wiśniewska, Dymerski, Namieśnik, & Wardencki, 2014). PCA, as a common unsupervised recognition method, was applied for the initial exploration to convert a set of observations of possibly correlated variables into a reduced set of uncorrelated variables.

2.5. Classification modelling

All samples were randomly divided into training (75%) and testing (25%) sets. The training subset was then used to develop the classification models by k -Nearest Neighbour (k NN), Support Vector Machine (SVM), Random Forest (RF) and Linear Discriminant Analysis (LDA). The accuracy of models was first evaluated by using the leave-one-out cross-validation method within the training set. For each classification approach, a grid search was performed in order to identify the most

optimum parameter by examining the confusion matrix of the predictions of training set. The optimized models were then used for the models' stabilization for each machine learning technique applied. The previous steps were repeated as part of a 1000 cycle process, where at each cycle, the training and testing subset samples were randomized and reshuffled.

SVM are inherently two-class classifiers. The way to do multiclass classification with R package e1701 is to use the ‘one-against-one’-approach, in which $k(k - 1)/2$ binary classifiers are trained and the appropriate class is found by a voting scheme (Meyer, Dimitriadou, Hornik, Weingessel, & Leisch, 2018). KNN was implemented using the function “knn” from the R package (Schliep & Hechenbichler, 2016). The best k was selected using a grid search from $k = 2$ to 10. RF was implemented with 200 trees using the “Random Forest” function from the R package (Liaw & Wiener, 2002). LDA was implemented using the “LDA” function from the “MASS” R package (Venables & Ripley, 2002).

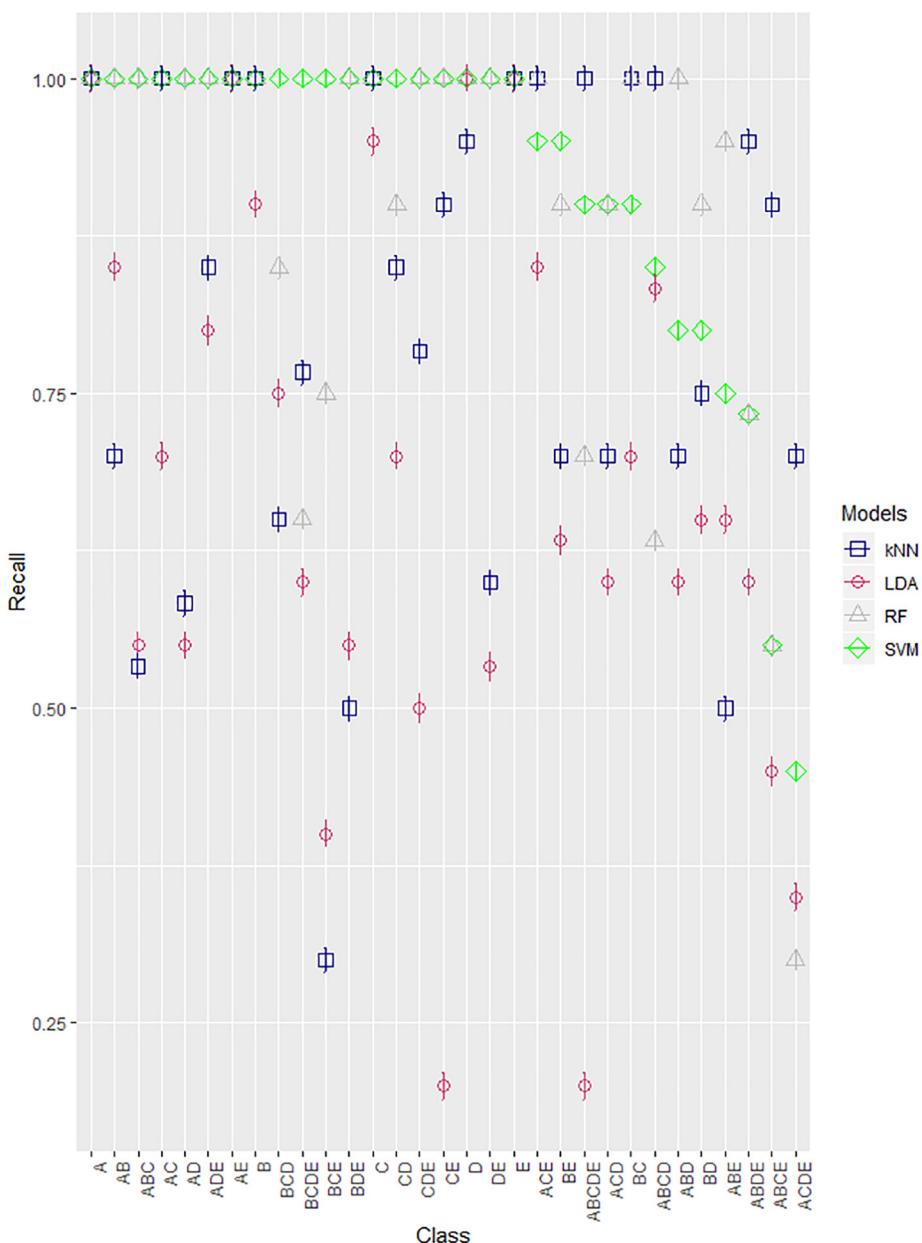


Fig. 4. (continued)

2.6. Quantification model with partial least squares regression

Partial Least Squares (PLS) regression models were built for quantitatively analyzing sweeteners according to the spectral information of samples. PLS was implemented using the “PLS” function from the R package, which is freely available from the CRAN website. Pre-treatments include SG, 1st D and their combination were applied and compared. 10-fold cross validation was applied in this modelling procedure. After the model was established, the test set was then analyzed in order to estimate the predictive capability of the established models, to minimize the risk of overfitting. Correlation coefficients values (r^2) and Root Mean Square Error for predictions (RMSEP) of the respective training and test sets were used to evaluate the predictivity of the models. The RMSEP was calculated as shown in the following equation:

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (X_{\text{obs},i} - X_{\text{model},i})^2}{n}}$$

where $\hat{y}_{t,i}$ is the percent of adulteration predicted by the model, $y_{t,i}$ is

the measured value (actual percent of adulteration), n_t is the number of samples in the test set. Average error indicated by RMSEP was expected in future predictions when applying the calibration model to the unknown samples.

2.7. Determination of sweeteners by RP-HPLC

Determination of sodium cyclamate, sucralose, sodium saccharin, acesulfame-K and aspartame in beverage samples was conducted by reverse phase-high performance liquid chromatography (RP-HPLC) in a Waters Alliance 2695 (Waters, Milford, MA, USA) equipped with a variable wavelength UV-visible detector measuring at 230 nm and a high pressure reliance cartridge guard-column (C₁₈ column, 2.5 μm 4.6 mm × 250 mm, Dalian Elite Analytical Instruments Co., Ltd., Dalian, China). All chromatograms were performed at the room temperature with the use of a 95:5 (v/v) mixture of ammonium acetate and methanol, and a flow rate of 1 ml min⁻¹ (Ni, Xiao, & Kokot, 2009).

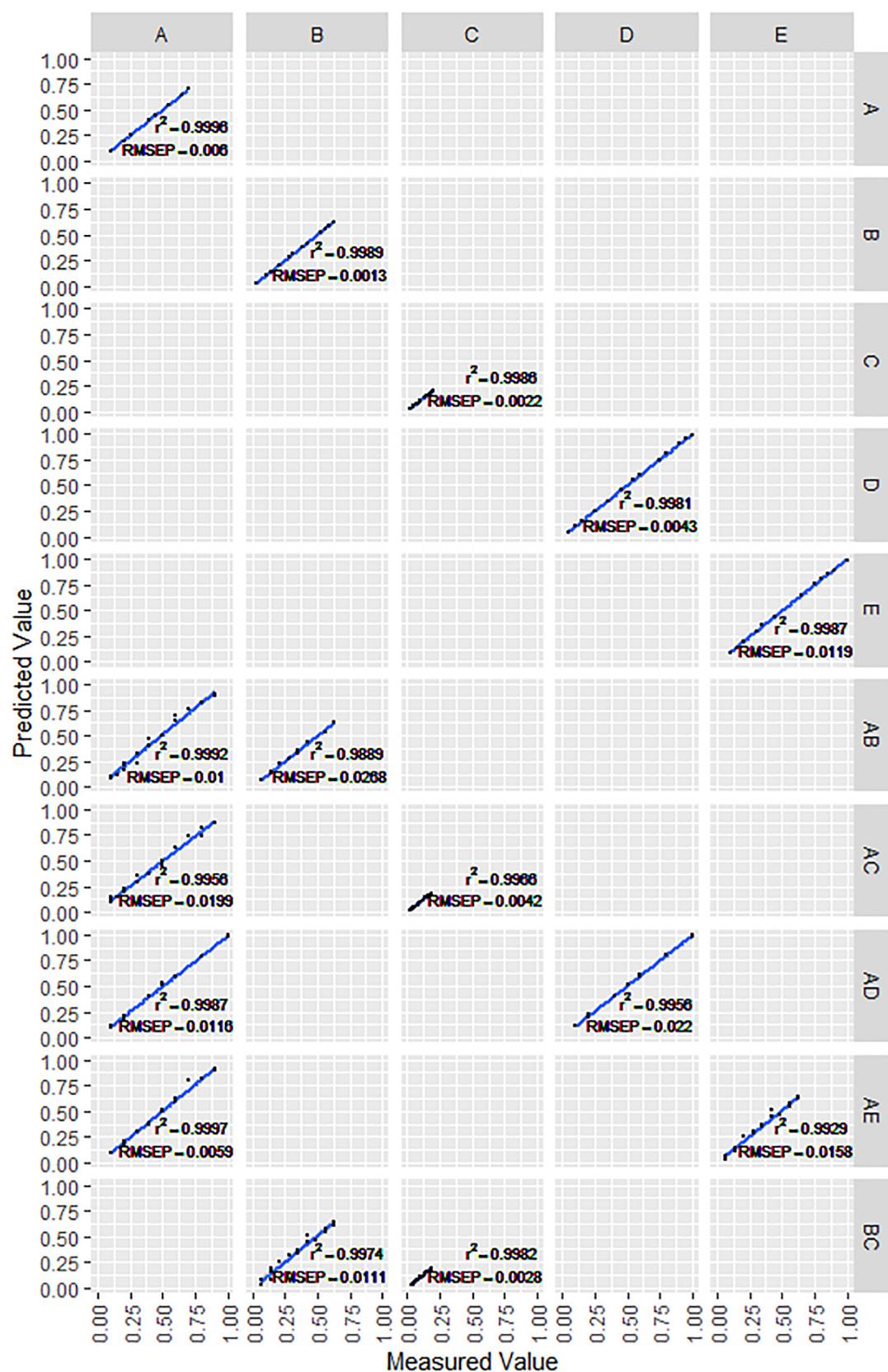


Fig. 5. Quantification models for the five sweeteners and their mixtures using PLS. A: Sodium cyclamate; B: Sucralose; C: Sodium saccharin; D: Acesulfame-K; E: Aspartame.

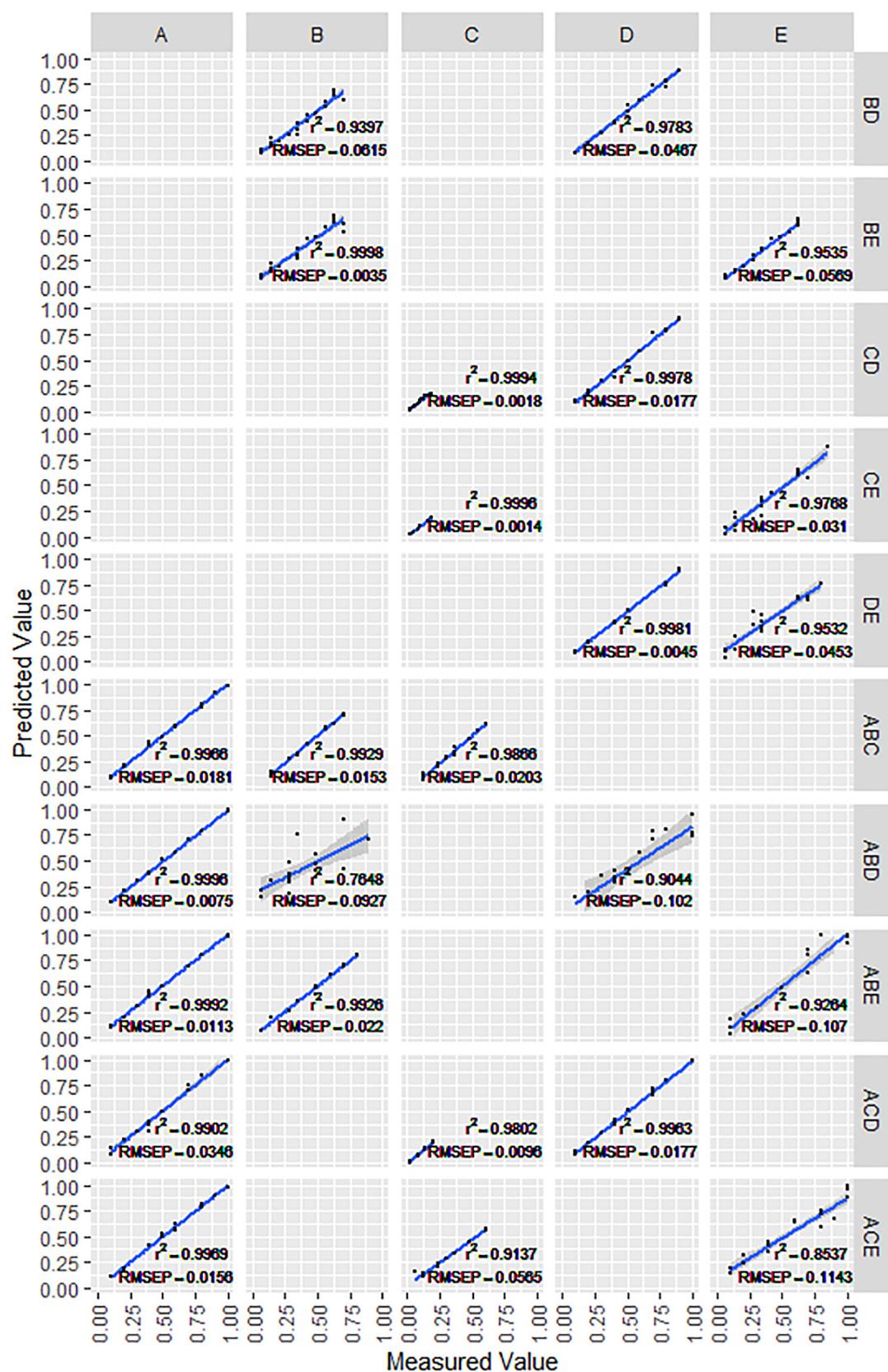


Fig. 5. (continued)

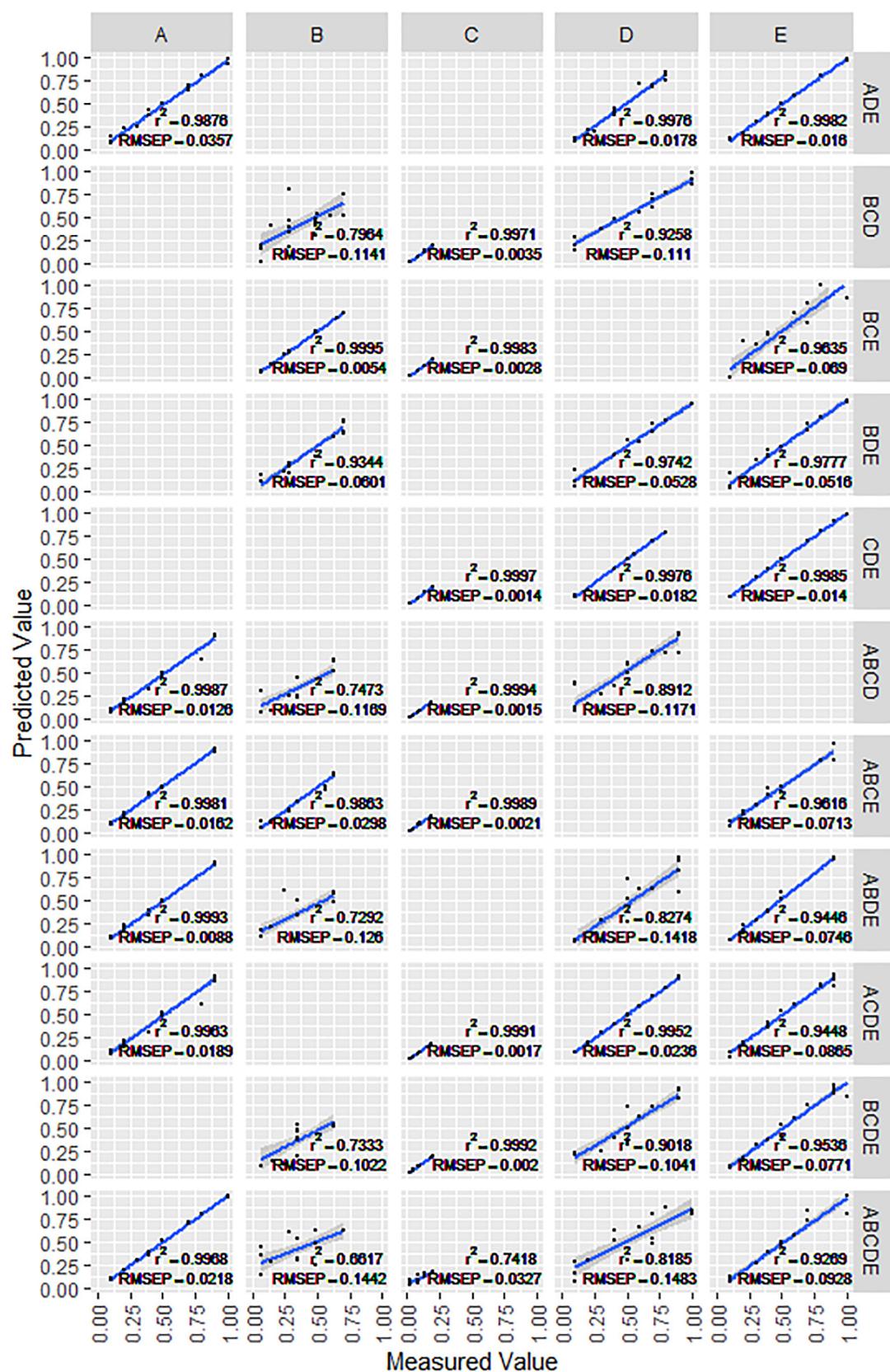


Fig. 5. (continued)

Table 1

Analytical comparison of sodium cyclamate, sucralose, sodium saccharin, acesulfame-K, and aspartame in beverages with the PLS calibration model and the reference HPLC method.

Samples	PLS (mg kg^{-1})		HPLC (mg kg^{-1})		RSE (%)
	Mean value	RSD	Mean value	RSD	
<i>Sample 1</i>					
Sodium Cyclamate	59.11	0.25	61.34	0.37	3.70
Sucralose	ND		ND		
Sodium Saccharin	ND		ND		
Acesulfame-K	111.01	1.79	114.56	0.76	3.15
Aspartame	ND		ND		
<i>Sample 2</i>					
Sodium Cyclamate	65.1	1.34	68.65	0.56	5.31
Sucralose	ND		ND		
Sodium Saccharin	55.2	1.23	51.87	0.89	6.22
Acesulfame-K	ND		ND		
Aspartame	ND		ND		
<i>Sample 3</i>					
Sodium Cyclamate	ND		ND		
Sucralose	ND		ND		
Sodium Saccharin	18.1	0.97	20.21	0.26	11.01
Acesulfame-K	ND		ND		
Aspartame	116.15	1.85	113.26	0.82	2.52

3. Results and discussion

3.1. Spectra de-noising and smoothing

Some effects including unexpected perturbations, baseline shifts and slope changes in spectra, can lead to spectrum changes that are not related to the target responses, but impair the reliability of multivariate calibration models. Such influences can be significantly reduced by mathematical pretreatment methods. Different spectral pretreatment methods were applied in this study. The computation of SG and 1st D of the developed PLS models by each pretreatment method for spectra were used to account for the resolution of overlapping peaks and removal of baseline variation. Better results were generated by using preprocessing. The most suitable data preprocessing methods for each sweetener sample were employed for sodium cyclamate as SG and 1st D, sucralose as first derivative, sodium saccharin as SG and 1st D, acesulfame-K as SG and 1st D, aspartame as SG.

3.2. Selection of characteristic infrared spectra for different sweeteners

The qualitative calibration models developed for the five sweeteners were validated using a test set of five samples for each sweetener. Fig. 1 shows the infrared spectra of the five sweeteners in water at wavelengths ranging from 1500 to 1000 cm^{-1} . As it can be seen from Fig. 1F, all spectra are different because of unique molecular structures, and exhibit a few broad absorption peaks representing their own absorptive groups.

The corrected spectral data still hold hundreds of variables that contain redundant and co-linear information. The selection of informative variables is an effective approach after spectral pretreatment to prevent the use of irrelevant bands for modelling. The screening of the qualified wavelengths for identification and quantification is of great significance as more robust and simplified calibration models are expected to be built based on the experimental results. In this study, PCA, as a wavelength selection method, has been shown to improve the resulting models.

The construction of the reliable quantitative models demands the selection of an appropriate number of principal components, or latent variables, for which the studied model reflects most informative variables. The first two principal components (PCs) from the PCA (Fig. 2A) for sweeteners were used to characterize the samples. According to the

results, the specific regions for all the sweeteners investigated in this work were obtained. The following bands can be highlighted as sodium cyclamate 1258–1119 cm^{-1} , sucralose 1196–1146 cm^{-1} and 1412–1277 cm^{-1} ; sodium saccharin 1212–1203 cm^{-1} , 1270–1245 cm^{-1} and 1328–1307 cm^{-1} ; acesulfame-K 1069–1065 cm^{-1} , 1200–1142 cm^{-1} and 1428–1269 cm^{-1} ; aspartame 1410–1330 cm^{-1} .

The factor loadings and infrared spectra of the five sweeteners are shown in Fig. 2B. The loading factors corresponding to PC1 and PC2 are more than zero, that presented strong IR absorption bands in the region. Each PC accounts for consecutively decreasing the amount of data variance, which results in the compression of significant data into a few PC variables. Each data object has a score value on each PC, and each original variable is associated with a loadings value on each PC. The highest correlation coefficient value indicated the degree of fit in the spectral region. Thus, selection of the region for the analysis of sweetener samples is fully addressable. In sodium cyclamate, asymmetrical SO_2 stretching was attributed to bands at 1215 cm^{-1} , and symmetrical SO_2 stretching to the band at 1173 cm^{-1} , in agreement with the findings of Carvalho (de Carvalho, Segato, Nunes, Novak, & Cavalheiro, 2009). The most intense band of acesulfame-K is located at 1200–1142 cm^{-1} due to NH_2 rock in amides. Other less intense absorption bands are located at 1428–1269 cm^{-1} due to C=C stretch in cyclohexenes (Armenta, Garrigues, & de la Guardia, 2004). The very strong bands of sucralose at 1196–1146 cm^{-1} can be assigned to (C=O)–O stretching modes, the weak IR bands at 1412–1277 cm^{-1} are assigned, respectively, to the scissoring, wagging or rocking modes of CH_2 (Brizuela et al., 2013). The spectrum of sodium saccharin shows the SO_2 asymmetrical stretching at 1270–1245 cm^{-1} , in agreement with the findings of Imai (Imai & Kamada, 2005). There is a strong IR band cluster for aspartame in the region of 1410–1330 cm^{-1} , which might be attributed to a high frequency single-bond stretch involving C–C and C–O bonds. Thus, selection of this region for the analysis of the five sweeteners is fully addressable.

A group of characteristic variables were chosen as the most informative wavenumbers. Relying on these wavenumbers, better detection results of identification and quantification were obtained by using the multivariate analysis models.

3.3. Classification models to identifying the sweetener

The qualitative models were developed for the sweeteners. For the discrimination of different sweeteners on the basis of the FTIR spectroscopy, the dataset was submitted for interpretation with the use of the machine learning techniques. Machine learning techniques as classifier, including kNN, LDA, SVM and RF, were used in the study. Performance of the model was evaluated according to the value leave-one-out cross-validation accuracy in the training set as well as the identification rate in the training set and validation set.

The most important standard to evaluate the classification model is the prediction accuracy of test set. As shown in Fig. 3, the predicted accuracy of the four strategies were compared and the mean of predicted accuracy of the four classified methods were 0.94 ± 0.014 , 0.94 ± 0.021 , 0.84 ± 0.034 and 0.74 ± 0.020 , respectively.

It was concluded that SVM and RF based models were superior to both the kNN and the LDA classification method, achieving a correct classification rate of over 93%. Although there was no significant difference between the nonlinear SVM and RF ($p = 0.067 > 0.05$), the standard deviation of SVM was smaller than RF. SVM model had better performance compared with RF, due to the less change of accuracy as identifying different samples. SVM as an excellent pattern recognition method was selected for categorical assignment of such complex data. SVM is a machine learning algorithm that aims to perform multivariate function estimation or non-linear classification. SVM algorithms use a set of mathematical functions that are defined as the kernel. The function of kernel is to take data as input and transform it into the

required form. Different SVM algorithms use different types of kernel functions. The polynomial, radial basis function, and sigmoid kernel functions were used to tune SVM model. The best SVM identification model was determined finally according to the highest identification rate in the training set and validation set. SVM classification model was built for identifying different sweeteners in the present study.

To further validate the models, the performance of classification model was evaluated by the precision and recall. The precision reveals the percentage of positive class labels correctly identified out of total predicted positives, and recall parameter determines the percentage of positives class labels correctly identified by the model out of the total number of actual positives. The higher precision and recall parameters represent a more robust model. As shown in Fig. 4, the sweetener species A, B, C, D, E and their mixtures are associated with the precision and recall of four classification models. The precision and recall parameter of SVM and RF were higher than the kNN and LDA. SVM is more suitable for the classification, which can be implemented in high variable dimensions with fewer training variables or samples (Liu, He, & Sun, 2009).

However, the precision of SVM was somewhat lower for the last 9 samples in the density map, and the lowest precision was 50%, while the precision of RF was 30% and kNN was 75% for the mixture of ABDE. In addition, the performance of RF was more powerful especially for the samples unsuited to the SVM and kNN classified model. The classified processing methods can be optimized according to specific sample composition.

3.4. Quantification model of each class sample

All the calibration models were performed using PLS. The PLS model could develop a linear relationship between the spectral data and the sweeteners content. The samples in the prediction set were applied to assess the prediction performance of the developed PLS models. The prediction performance was evaluated and assessed by indices as r^2 and RMSEP. Generally, a good model should have a higher r^2 value and lower RMSEP value. Prediction set was used to assess and evaluate the prediction performance of the developed model. The prediction results by PLS models were shown in Fig. 5. Every diagram in Fig. 5 indicates the predictive equation and correlation coefficient for each sweetener in different mixtures. The letters in the first line of the figure represent the specific sweetener in the mixture, and the composition of the samples was represented on the right column of the figure. The predictive equation of one sweetener is different in mixtures. As can be seen, the optimal prediction performance was achieved by a single sweetener sample with r^2 value ranged from 0.9981 to 0.9996, and RMSEP ranged from 0.012 to 0.0013. The r^2 value and RMSEP ranged from 0.9397 to 0.9998 and 0.062 to 0.0014 for the mixture of two sweeteners.

As compared with the results of single or double mixed samples, it was seen that mixtures of more types of sweeteners mostly resulted in a lower fitting degree of prediction and measured value. For example, the r^2 value and RMSEP of the three types of sweetener mixed samples were 0.7648–0.9997 and 0.11–0.0014, totally higher than those indices of the four types of sweetener mixed samples (0.7292–0.9994 and 0.13–0.0015) or the five types of sweetener mixed samples (0.6617–0.9968 and 0.14–0.022), but clearly lower than the indices resulting from one or two sweeteners mixture.

Moreover, it is worth noting that sodium cyclamate was mostly found to be more sensitive (e.g. it usually had r^2 value more than 0.9876, and RMSEP lower than 0.036) for the prediction method among the five sweeteners investigated, whereas sucralose was mostly insensitive (e.g. it usually had the lowest r^2 value as 0.6617, and RMSEP 0.14), both in single or mixed sweetener samples. The reason for a better performance of sodium cyclamate was that it has the highest concentration in the samples because of the lower sweetness, and this highest infrared correspondence improved the prediction performance.

While the concentration of sucralose in the samples was lower because of the high-potency sweetness.

PLS is a widely applied multivariate calibration method in the application of a spectroscopic technique, because it is particularly useful when the number of variables is larger than the number of data points. PLS is helpful and effective when the within group variability of the samples dominates the between-group variability and when variable reduction is needed (Barker & Rayens, 2003). This is the case for many chemometric datasets, since they often contain a lot of variables and a relatively small number of samples. For these reasons, the method is frequently applied in chemometry. The application of the PLS model in the present study was to yield perfect quantification for the five sweeteners.

3.5. Application: determination of sodium cyclamate, sucralose, sodium saccharin, acesulfame-K and aspartame in beverages

The real tea beverage samples obtained from the local market noted in Table 1 were analyzed with the use of the SVM model for identification, and PLS calibration model for quantitative analysis. These samples contained one or more of the sweeteners. The results reported in Table 1 are compared with those obtained from the reference HPLC method with the relative standard error. Such estimates between the two methods were in the range of 2.52–11.01, which supports the reliability of the proposed method.

4. Conclusion

In conclusion, FTIR equations developed using a training set of different sweeteners are applicable for the rapid determination sweetener contents. Results also showed an effect of data preprocessing methods on FTIR sweetener determination. In fact, SVM is a powerful method to deal with multivariate classification for accurate sweetener identification. Considering the variability of sweeteners, it appears that larger training sets are necessary to achieve accurate and robust calibrations for FTIR determinations of sweeteners. The low RMSEP and high r^2 of prediction and validation obtained in this study using PLS, as well as the numerous advantages of the technology, make FTIR spectroscopy a very promising tool for quality control in liquid foods containing sweeteners.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was funded by the National Natural Science Foundation of China (Grant No. 31571930). The authors wish to thank the anonymous reviewers and editors for their valuable advices.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodchem.2019.125404>.

References

- Abu-Ghoush, M., Fasfous, I., Al-Degs, Y., Al-Holy, M., Issa, A. A., Al-Reyahi, A. Y., & Alshathri, A. A. (2017). Application of mid-infrared spectroscopy and PLS-Kernel calibration for quick detection of pork in higher value meat mixes. *Journal of Food Measurement and Characterization*, 11(1), 337–346.
- Armenta, S., Garrigues, S., & de la Guardia, M. (2004). FTIR determination of aspartame and acesulfame-K in tabletop sweeteners. *Journal of Agricultural and Food Chemistry*, 52(26), 7798–7803.

- Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, 17(3), 166–173.
- Behrens, M., Blank, K., & Meyerhof, W. (2017). Blends of non-caloric sweeteners saccharin and cyclamate show reduced off-taste due to TAS2R bitter receptor inhibition. *Cell Chemical Biology*, 24(10), 1199–1204.
- Brizuela, A. B., Raschi, A. B., Castillo, M. V., Leyton, P., Romano, E., & Brandán, S. A. (2013). Theoretical structural and vibrational properties of the artificial sweetener sucralose. *Computational and Theoretical Chemistry*, 1008, 52–60.
- de Carvalho, L. C., Segato, M. P., Nunes, R. S., Novak, C., & Cavalheiro, É. T. G. (2009). Thermoanalytical studies of some sweeteners. *Journal of Thermal Analysis and Calorimetry*, 97(1), 359–365.
- Edwards, C. H., Rossi, M., Corpe, C. P., Butterworth, P. J., & Ellis, P. R. (2016). The role of sugars and sweeteners in food, diet and health: Alternatives for the future. *Trend in Food Science & Technology*, 56, 158–166.
- Ferrand, M., Huquet, B., Barbey, S., Barillet, F., Faucon, F., Larroque, H., ... Brochard, M. (2011). Determination of fatty acid profile in cow's milk using mid-infrared spectrometry: Interest of applying a variable selection by genetic algorithms before a PLS regression. *Chemometrics and Intelligent Laboratory Systems*, 106(2), 183–189.
- Ferrer, I., & Thurman, E. M. (2010). Analysis of sucralose and other sweeteners in water and beverage samples by liquid chromatography/time-of-flight mass spectrometry. *Journal of Chromatography A*, 1217(25), 4127–4134.
- Imai, Y., & Kamada, J. (2005). Vibrational spectra of saccharin nitrion and its orientation on the surface of silver metal particles. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 61(4), 711–715.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.
- Liu, F., He, Y., & Sun, G. (2009). Determination of protein content of auricularia auricular using near infrared spectroscopy combined with linear and nonlinear calibrations. *Journal of Agricultural and Food Chemistry*, 57(11), 4520–4527.
- Mabood, F., Jabeen, F., Ahmed, M., Hussain, J., Al Mashaykhi, S. A. A., Al Rubaiey, Z. M. A., ... Manzoor, S. (2017). Development of new NIR-spectroscopy method combined with multivariate analysis for detection of adulteration in camel milk with goat milk. *Food Chemistry*, 221, 746–750.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2018). e1071: Misc functions of the department of statistics, probability theory group (Formerly: E1071). TU Wien.
- Mooradian, A. D., Smith, M., & Tokuda, M. (2017). The role of artificial and natural sweeteners in reducing the consumption of table sugar: A narrative review. *Clinical Nutrition ESPEN*, 18, 1–8.
- Ni, Y., Xiao, W., & Kokot, S. (2009). A differential kinetic spectrophotometric method for determination of three sulphanilamide artificial sweeteners with the aid of chemometrics. *Food Chemistry*, 113(4), 1339–1345.
- Ordoñez, E. Y., Rodil, R., Quintana, J. B., & Cela, R. (2015). Determination of artificial sweeteners in beverages with green mobile phases and high temperature liquid chromatography-tandem mass spectrometry. *Food Chemistry*, 169, 162–168.
- Roberts, J. R. (2015). The paradox of artificial sweeteners in managing obesity. *Current Gastroenterology Reports*, 17(1).
- Rother, K. I., Conway, E. M., & Sylvetsky, A. C. (2018). How non-nutritive sweeteners influence hormones and health. *Trends in Endocrinology & Metabolism*, 29(7), 455–467.
- Schliep, K., & Hechenbichler, K. (2016). kknn: Weighted k-nearest neighbors.
- Setti, A. S., Braga, D. P. D. A., Halpern, G., Figueira, R. D. C. S., Iaconelli, A., & Borges, E. (2018). Is there an association between artificial sweetener consumption and assisted reproduction outcomes? *Reproductive Biomedicine Online*, 36(2), 145–153.
- Śliwińska, M., Wiśniewska, P., Dymerski, T., Namieśnik, J., & Wardenci, W. (2014). Food analysis using artificial senses. *Journal of Agricultural and Food Chemistry*, 62(7), 1423–1448.
- Stojkovic, M., Mai, T. D., & Hauser, P. C. (2013). Determination of artificial sweeteners by capillary electrophoresis with contactless conductivity detection optimized by hydrodynamic pumping. *Analytica Chimica ACTA*, 787, 254–259.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. New York: Springer.
- Wu, M., Qian, Y., Boyd, J. M., Hrudey, S. E., Le, X. C., & Li, X. (2014). Direct large volume injection ultra-high performance liquid chromatography-tandem mass spectrometry determination of artificial sweeteners sucralose and acesulfame in well water. *Journal of Chromatography A*, 1359, 156–161.
- Zyglar, A., Wasik, A., & Namieśnik, J. (2009). Analytical methodologies for determination of artificial sweeteners in foodstuffs. *TrAC Trends in Analytical Chemistry*, 28(9), 1082–1102.