ARTICLES

Exploring Near and Midinfrared Spectroscopy to Predict Trace Iron and Zinc Contents in Powdered Milk

Di Wu, Yong He,* Jiahui Shi, and Shuijuan Feng

College of Biosystems Engineering and Food Science, Zhejiang University, 268 Kaixuan Road, Hangzhou 310029, China

Near infrared (NIR) and mid-infrared (MIR) spectroscopy were investigated to predict iron and zinc contents in powdered milk. A hybrid variable selection method, namely, uninformative variable elimination (UVE) combined with successive projections algorithm (SPA), was applied to select the most effective wavenumber variables from full 2756 NIR and 3727 MIR variables, respectively. Finally, 18 NIR and 18 MIR variables were selected for iron content prediction, and 17 NIR and 12 MIR variables for zinc content prediction. The obtained effective wavenumber variables were input into partial least-squares (PLS) and least-squares-support vector machines (LS-SVM), respectively. The selected MIR variables obtained much better results than NIR to predict both iron and zinc contents in both the PLS and LS-SVM models. The iron content prediction results based on LS-SVM with 18 MIR spectra were as follows: coefficient of determination (r^2) was 0.920, residual predictive deviation (RPD) was 3.321, and root-mean-square error of prediction (RMSEP) was 1.444. The zinc content prediction results based on LS-SVM with 12 selected MIR spectra were as follows: r^2 was 0.946, RPD was 4.361, and RMSEP was 0.321. The good performance shows that UVE-SPA is a powerful variable selection tool. The overall results indicate that MIR spectroscopy incorporated to UVE-SPA-LS-SVM could be applied as an alternative fast and accurate method to determine trace mineral content in powdered milk, such as iron and zinc.

KEYWORDS: Near infrared (NIR) spectroscopy; mid-infrared (MIR) spectroscopy; uninformative variable elimination (UVE); successive projections algorithm (SPA); least-squares-support vector machine (LS-SVM); iron; zinc; powered milk

INTRODUCTION

Powdered milk is a popular manufactured dairy product. Powdered milk is high in minerals and those required in amounts less than 100 mg per day, often called trace elements, such as iron and zinc. Iron is an important mineral in powdered milk. Iron-fortified powdered milk can treat anemia. Torres verified the influence of the use of powdered milk enriched with iron and ascorbic acid as an intervention measure for treating iron deficiency anemia in children (*1*). Zinc is another essential mineral that is found in almost every cell. It stimulates the activity of approximately 100 enzymes, which are substances that promote biochemical reactions in the human body (*2, 3*). However, human milk does not provide the recommended amounts of zinc for older infants; therefore, breast-fed infants of this age should also consume age-appropriate foods containing zinc such as formula powdered milk containing zinc (*3*).

Essential elements are usually added in formula powdered milk. However, defective or excessive additions of these elements can produce detrimental effects on health. Thus, the determination of trace elements in powdered milk is important for nutritional purposes (*4, 5*). Some wet chemistry methods were applied for iron and zinc content determination in powdered milk, such as flow-injection atomic absorption spectrometry (*6*), reaction cell inductively coupled plasma mass spectrometry (*7*), electrothermal atomization atomic absorption spectrometry (*8*), and slurry sampling electrothermal vaporization inductively coupled plasma mass spectrometry (*9*). However, these methods are time-consuming, laborious, expensive, and require complex and professional laboratory operations.

Near infrared (NIR) spectroscopy and midinfrared (MIR) spectroscopy are widely employed as alternatives to wet chemistry procedures for qualitative and quantitative analysis

* To whom correspondence should be addressed. Tel: +86-571-86971143. Fax: +86-571-86971143. E-mail: yhe@zju.edu.cn (Y.H.). E-mail: ilkb8mvp@yahoo.com.cn).

in agriculture and food quality evaluation (*10, 11*). We did analyses on main component prediction using short-wave NIR, NIR, and MIR spectroscopy (*12−15*). Borin and co-workers studied the common adulterants in powdered milk by near-infrared spectroscopy (*16*).

Minerals in powdered milk probably exist in both inorganic and organic complexes. Because NIR and MIR spectroscopy measures absorption by molecular bonds, correlations between spectra and mineral contents seem unreliable. However, prediction of some minerals by NIR and MIR may be possible through their association with the organic matrix (*17*). NIR and MIR spectroscopy techniques are beginning to be applied to the analysis of minerals along with the advancement of chemometrics. Matteson and Herron concluded that minerals exhibit most of their fundamental molecular vibration modes in the midinfrared region (4000 to 400 cm^{-1}) (*18*). Researchers found that NIRS had become more and more popular and suited for detecting minerals in plant tissue, such as iron (*19, 20*) and zinc (*20*). Mesubi did an infrared study of zinc, cadmium, and lead salts of some fatty acids (*21*). From the O−C−O stretching frequencies, Mesubi deduced that the coordination of the carboxylate group to the metal ions is unsymmetric chelating bidentate. Bochmann et al. studied the infrared spectra of the dimethyl complexes of zinc, cadmium, and mercury, isolated in argon matrices (*22*). It would thus seem that prediction of mineral contents of food by NIR and MIR spectroscopy may be possible, possibly because these have some association with organic components.

The objectives of this study were to investigate the feasibility of using NIR and MIR spectroscopy to predict the iron and zinc contents in powdered milk. The prediction abilities of NIR and MIR were compared on the basis of partial least-squares (PLS) and least-square support vector machine (LS-SVM), respectively. Moreover, as there are hundreds or thousands wavenumber variables and hundreds of samples, a hybrid variable selection method, uninformative variable elimination (UVE) combined with successive projections algorithms (SPA), was applied to select effective wavenumber variables. The performance of UVE-SPA was compared with the full spectral variables or the variables only selected by UVE.

MATERIALS AND METHODS

System Setup and Data Measurement. Seven brands of powdered milk produced by the Inner Mongolia Yili Industrial Group Co., Ltd., a famous corporation in China's powdered milk market, were bought in several local supermarkets. The 7 brands were 0−6 month infant powdered milk, 6−12 month infant powdered milk, 1−3 years toddler powdered milk, 3−6 years infant powdered milk, student powdered milk with zinc, women's nutrition powdered milk, and whole powdered milk. NIR and MIR spectra were measured by JASCO Model FT/IR-4000 Fourier Transform Infrared Spectrometer (Japan). The valid range is 7800−350 cm^{-1}, and auto scan speed is 2.0 mm/s. The whole experiment was carried out at about 25 °C. The powdered milk sample was mixed with potassium bromide (KBr) at the ratio of 1:49. Each sample was scanned 40 times. These 40 data were averaged as the transmission value (% T) of this sample. The spectral data of about 60 samples were measured for each variety, and finally, the spectral data of 409 samples were obtained. Because of system imperfections, obvious scattering noises can be observed at the beginning and end of the spectral curve. Thus, the NIRS analysis was based on 4000 to 6666 cm^{-1}, while the MIRS analysis was from 400 to 4000 cm^{-1}. The smoothing way of Savitzky−Golay with three segments and zero polynomials and standard normal variate were applied for spectral pretreatment. The pretreatment methods were implemented by The Unscrambler V 9.7 (CAMO PROCESS AS, OSLO, Norway). The reference values of iron and zinc were determined by atomic absorption spectrometry (PerkinElmer AAnalyst 800 high-performance atomic absorption spectrometer). The absorbed atomic resonance line was 248.3 and 213.8 nm for iron and zinc, respectively. Each sample was measured five times, and the standard deviations of iron and zinc contents were less than 0.1 mg for each sample. Considering that both iron and zinc are trace minerals in powdered milk and that their content predictions do not need to be very accurate for consumers, the accuracy of the atomic absorption spectrometer is acceptable as the standard measurement method for the spectroscopy analysis.

Uninformative Variable Elimination. UVE is a method of variable selection based on stability analysis of the PLS regression coefficient (*23*). One advantage of UVE, compared with other variable selection methods, is that it is user independent, and therefore does not present any configuration problems (*23*). In the UVE method, a PLS regression coefficient matrix $b = [b_1...b_p]$ is calculated through a leave-one-out validation (*23*); then the reliability of each variable (wavenumber) can be quantitatively measured by its stability. The stability of variable j can be calculated as:

$$s_j = mean(\beta_j)/std(\beta_j) \qquad (1)$$

where $mean(\beta_j)$ and $std(\beta_j)$ are the mean and standard deviation, respectively, of the regression coefficients of variable j. Therefore, the larger the stability, the more important the corresponding variables. A variable whose stability is lower than a cutoff threshold is regarded as uninformative and is eliminated. In this article, the cutoff value used was 99% of the $|s_j^{random}|$. The process of UVE was executed in MATLAB 7.6 (The Math Works, Natick, USA).

Successive Projections Algorithms. SPA is to select variables whose information content is minimally redundant, in order to solve the collinearity problems. SPA employs simple projection operations in a vector space to select subsets of variables with minimum of collinearity (*24*). SPA provides more reproducible results, performing the selection in a time interval usually shorter than that with the genetic algorithm (*25*). For a detailed description of SPA, see refs *2−4*. The process of SPA was executed in MATLAB 7.6 (The Math Works, Natick, USA).

However, variables selected by SPA may be with low signal to noise ratios (S/N) or useless for multivariate calibration, which can affect model precision of prediction (*26*). Thus, UVE was first employed to select informative variables, followed by SPA to select variables that have minimum redundant information from the informative variables.

Calibration Methods. PLS analysis is a widely used calibration method in spectroscopic analysis. It is a procedure used to model the relationship between the spectral data matrix (*X*) and the target chemical properties matrix (*Y*) for predicting with the smallest number of LVs. PLS uses the *Y* information during the decomposition process so that the *X* containing higher constituent concentrations can be weighted more heavily than those with low concentrations. The PLS regression procedure has the advantage of accepting more variables than samples. In the development of the PLS model, full cross-validation was used to validate its quality and to prevent overfitting of the calibration model.

LS-SVM is an optimized algorithm based on the standard support vector machine. The RBF kernel was adopted. A Grid-search technique was applied to find out the optimal parameter values, which include the regularization parameter γ and the RBF kernel function parameter σ^2. The optimization value ranges of γ and σ^2 were $2^{-1}-2^{10}$, and $2-2^{15}$ with adequate increments. For each combination of γ and σ^2 parameters, the root-mean-square error of cross-validation (RMSECV) was calculated, and the optimum parameters were selected when producing smaller RMSECV. The details of LS-SVM description could be found in the literature (*27*).

Model Evaluation. The evaluation indices of predictive capability for all developed models were coefficient of determination (r^2), residual predictive deviation (RPD) (*28*), and root-mean-square error (RMSE) of prediction set (RMSEP). Generally, a good model should have higher r^2 and RPD values and lower RMSEP values. RMSE is calculated as follows:
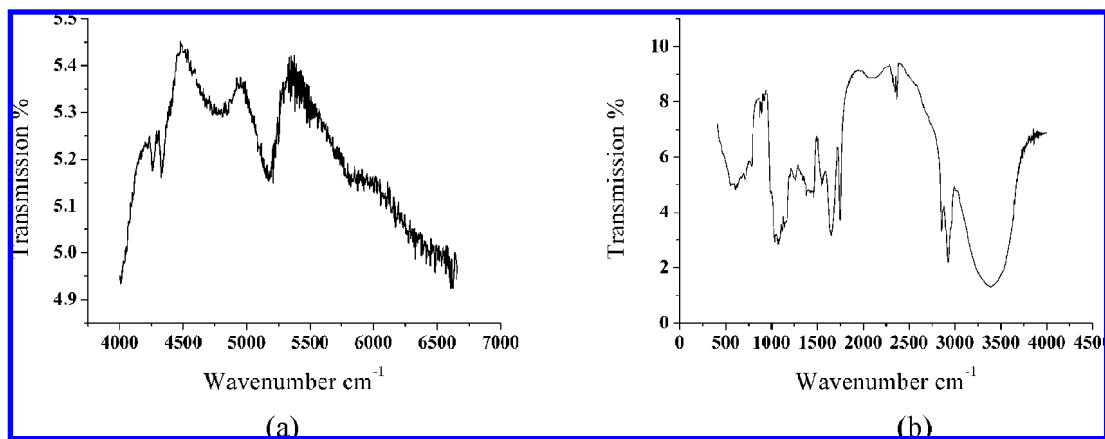
Figure 1. NIR (**a**) and MIR (**b**) transmission spectra of powdered milk samples.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \qquad (6)$$

where n is the number of samples, y_i and $\hat{y}_i$ are the reference and predicted values of the ith sample, respectively.

RESULTS AND DISCUSSION

NIR and MIR Spectral Investigation. Figure 1 shows the original transmission spectra at MIR and NIR regions. The NIR spectra contain a lot of noise. The weak peak at 4260 cm^{-1} was assigned to the CH_2 symmetric stretching vibration + CH_2 deformation vibration of HC=CHCH$_2$. The weak peak at 4330 cm^{-1} was assigned to the C—H stretching + C—H deformation vibration of CH_2. The peak at 4732 cm^{-1} was assigned to the N—H symmetric stretching vibration + amide III; 5155 cm^{-1} was H_2O absorption. In the MIRS region, there are some absorption peaks. The broad and strong peak from 3200 to 3600 cm^{-1} was due to O—H stretching vibrations. CH_2 symmetric stretching vibration showed a peak at 2854 cm^{-1}, and the CH_2 asymmetric stretching vibration shows a peak at 2925 cm^{-1}. The broad peak between 2000 and 2250 cm^{-1} was R′-C≡C-R. The C=O stretching vibration showed a peak at 1745 cm^{-1}. The C=C stretching vibration showed a peak at 1650 cm^{-1}. The peaks at 875 and 898 cm^{-1} were assigned to the C—H deformation vibration. However, it was hard to tell which peaks were related to iron or zinc absorptions. Moreover, spectral curves would be overlapped when there were many samples. Redundancy spectra from thousands of wavenumber and hundreds of samples are hard to identify and anlyze. It is important to perform feature selection in chemometrics analysis.

Statistical Values of Iron and Zinc Contents. The Kennard—Stone (KS) algorithm was executed to divide samples into calibration (209 samples) and prediction (200 samples) sets (*29*). The KS algorithm is a classic method to extract a representative set of objects from a given data set. It selects a representative data set consisting of samples separated by large Euclidean distances. Statistical values of iron and zinc in powdered milk are shown in **Table 1**. As powdered milk samples were from different brands and batches, a broad range of concentration variation was observed in the calibration and prediction sets. This situation would be helpful to build a stable and robust calibration model for the determination of iron and zinc in powdered milk. The student powdered milk with zinc contains high zinc content of about 9 mg/100 g, compared with that of the others.

Table 1. Statistical Values of Iron and Zinc in Powdered Milk in Calibration and Validation Sets

parameters	models	sample no.	range	mean	standard deviation
iron (mg/100 g)	calibration	209	4.0−25.0	14.5	5.0
	prediction	200	4.0−24.8	14.3	5.1
	all samples	409	4.0−25.0	14.4	5.0
zinc (mg/100 g)	calibration	209	3.9−9.2	5.9	1.3
	prediction	200	3.9−9.1	5.8	1.3
	all samples	409	3.9−9.2	5.8	1.3

Iron Content Determination. PLS and LS-SVM models were established to determine iron content in powdered milk using NIR and MIR spectra, respectively. **Table 2** shows prediction results of iron content in powdered milk based on full NIR and MIR spectra. For both PLS and LS-SVM models, MIR spectra obtained much better results (with $r^2 = 0.869$ and 0.904 for PLS and LS-SVM, respectively) than NIR (with $r^2 = 0.647$ and 0.675 for PLS and LS-SVM, respectively). Meanwhile, LS-SVM did a better prediction performance (with $r^2 = 0.675$ and 0.904 for NIR and MIR spectra, respectively) than PLS (with $r^2 = 0.647$ and 0.869 for NIR and MIR spectra, respectively). However, even the fact that the RPD value of MIR based on LS-SVM was less than 3 shows that it was not suitable for screening (*28*). Moreover, as a whole NIR or MIR spectra contain useless or irrelevant information, and 2756 NIR and 3727 MIR wavenumbers caused both PLS and LS-SVM processes to be complex and time consuming. Some of these variables may contain useless or irrelevant information for calibration models such as noise and background, which can worsen the predictive ability of the whole model. Eliminating these variables can predigest the calibration model and improve prediction results in terms of accuracy and robustness. Thus, uninformative wavenumbers need to be eliminated to obtain a better quantitative calibration model.

In the process of UVE, different latent variable (LV) numbers of the PLS model in UVE process were compared. The numbers of LVs were calculated from 1 to 30. Full cross-validation was used in UVE to prevent overfitting problems. **Figure 2** shows the RMSECV plot of different numbers of LVs by UVE. It could be seen that the smallest RMSECVs were obtained based on 9 and 23 LVs for NIR and MIR spectra, respectively. **Figure 3** shows the stability of each wavenumber variables in NIR (**a**) and MIR (**b**) based on 9 and 23 LVs, respectively. Wavenumber variables are at the left of the vertical line, while random variables are at the right side. Two horizontal lines show the lower and upper cutoffs. The variables whose stability is within the cutoff lines should be treated as uninformative and

Table 2. Prediction Results of Iron Content in Powdered Milk Based on NIR and MIR Spectroscopy

variable selection method	calibration method	NIR variables	r^2	RMSEP	*RPD*	MIR variables	r^2	RMSEP	*RPD*
none	PLS	2756	0.647	3.064	1.567	3727	0.869	1.919	2.501
	LS-SVM	2756	0.675	2.892	1.660	3727	0.904	1.608	2.985
UVE	PLS	298	0.653	3.225	1.488	516	0.880	1.755	2.735
	LS-SVM	298	0.693	2.810	1.714	516	0.910	1.540	3.117
UVE-SPA	PLS	18	0.511	3.421	1.403	18	0.894	1.662	2.888
	LS-SVM	18	0.593	2.933	1.636	18	0.920	1.444	3.321

be eliminated. On the basis of optimal LVs, 298 and 516 wavenumbers were selected by UVE from 2756 NIR and 3727 MIR wavenumbers, respectively.

The selected wavenumbers were set as input variables of PLS and LS-SVM. **Table 2** shows the prediction results based on 298 NIR and 516 MIR variables obtained by UVE, respectively. Although the number of selected variables for MIR was larger than NIR, the prediction results based on selected variables of MIR were much better than NIR in both PLS and LS-SVM models. Moreover, the prediction results were improved for both NIR and MIR by using selected spectral variables, compared with results predicted on the basis of the full spectrum. The RPD value of MIR up to 3 shows that it was suitable for screening (*28*).

After UVE analysis, variables with no more information for modeling than noise were eliminated, and the numbers of input variables were reduced from 2756 to 298 and from 3727 to 516 for NIR and MIR spectra, respectively. However, there are still many input variables, and further variable selection was required to obtain latent variables. SPA was carried out on

selected NIR and MIR variables from UVE, respectively. **Figure 4** shows the RMSE scree plot obtained according to the proposed variable elimination procedure by SPA for NIR and MIR, respectively. The black solid dots show the selected variable numbers. As can be seen, a sharp fall is shown in the starting part of the RMSE curve as the numbers of selected variables were from 9 to 12 and from 2 to 8 for NIR and MIR, respectively. The trends of the RMSE curve were still descending but the improvement becomes marginal with further increasing the number of selected variables. The curve tends to level off after the determination of selected variables by the SPA cutoff procedure by F-test criterion with $\alpha = 0.25$ (*25*). Thus, 18 variables (RMSE = 3.391) and 18 (RMSE = 1.593) variables were selected for NIR and MIR, respectively.

The 18 selected NIR wavenumbers by UVE-SPA are 4256, 4398, 5196, 5241, 5275, 5318, 5340, 5407, 5415, 5496, 5509, 5640, 5781, 5813, 6015, 6248, 6508, and 6616 cm^{-1}. Some retained wavenumber regions might be attributable to some specific chemical structure related to iron ion. Wavenumber 4256 cm^{-1}, which was associated to an absorbance peak in
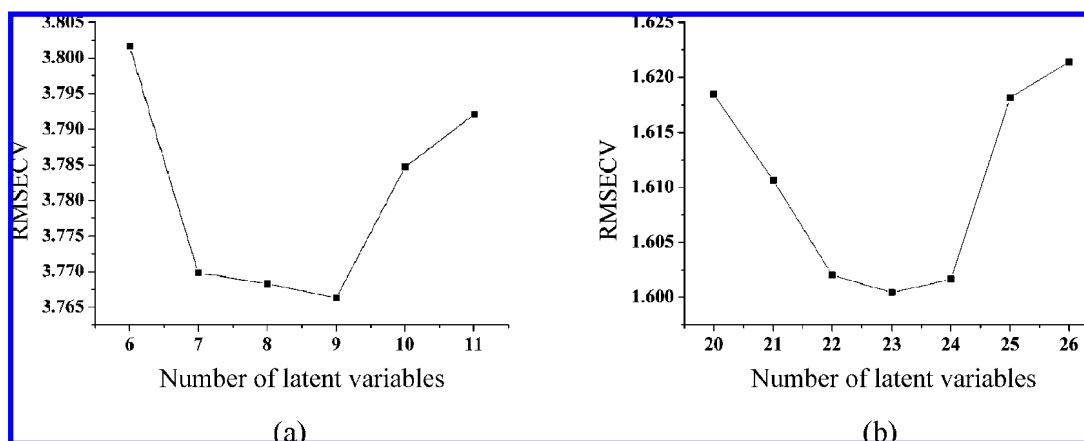

Figure 2. RMSECV plot of different numbers of latent variables in iron prediction by UVE. UVE calculations are based on full NIR spectra (**a**) and full MIR spectra (**b**).
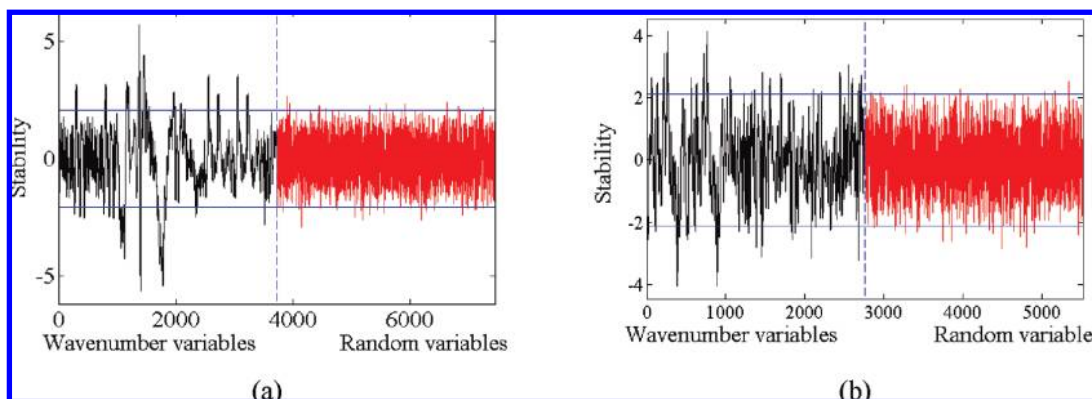

Figure 3. Stability of each variable in the NIR (**a**) and MIR (**b**) in iron content prediction by UVE-PLS of 9 and 23 LVs, respectively. The two blue horizontal lines indicate the lower and upper cutoffs.
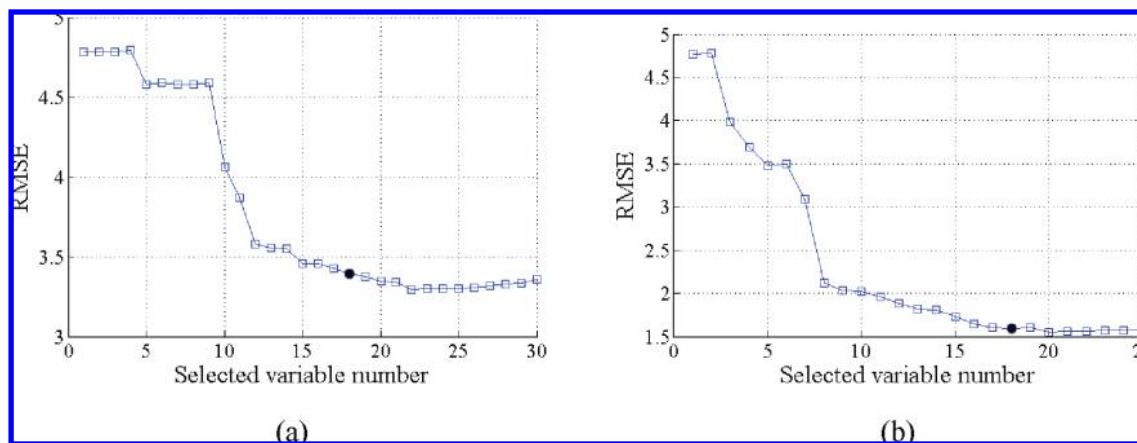
Figure 4. RMSE scree plot of SPA for NIR (**a**) and MIR (**b**) in iron content prediction.

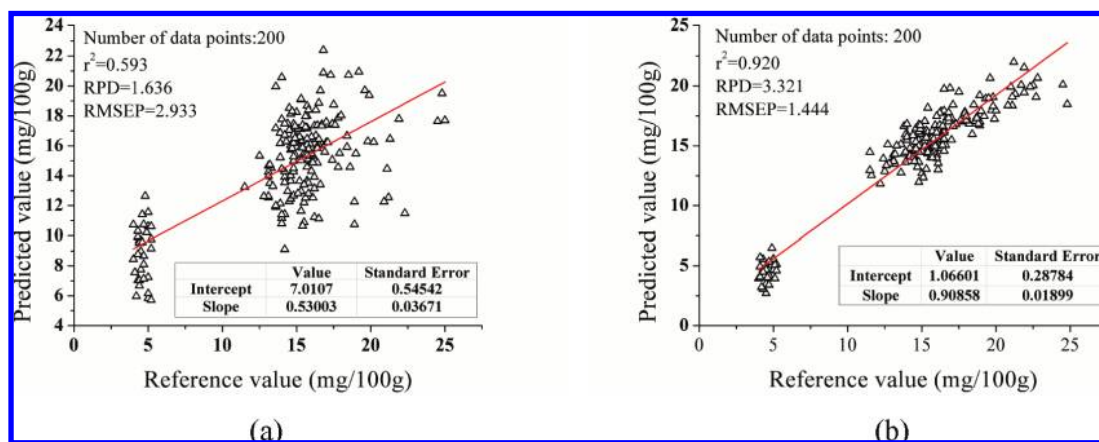**Figure 5.** Predicted vs reference values of iron content in powdered milk by LS-SVM models based on 18 NIR (**a**) and 18 MIR (**b**) wavenumbers selected by UVE-SPA.

Figure 1, was assigned to the CH_2 symmetric stretching vibration + CH_2 deformation vibration of $HC=CHCH_2$. Wavenumber 4398 cm^{-1} might be O−H stretching + C−C stretching of starch. Wavenumber 5196 cm^{-1} was assigned to 2 × C=O stretching of CONH. Wavenumber 5241 cm^{-1} was assigned to 2 × O−H stretching of ROH. Wavenumber 5275 cm^{-1} was assigned to O−H stretching + 2 × C−O stretching of starch. Wavenumber 5496 cm^{-1} was assigned to O−H stretching + 2 × C−O stretching of cellulose. Wavenumber 5640 cm^{-1} might be 2 × C−H stretching of CH_2. Wavenumber 6616 cm^{-1} might be 2 × N−H stretching of protein. Other selected wavenumbers, which might not relate to iron ion, have contribution to iron prediction model establishment. As there were many different overtone and combination bands in the NIR region, some obtained effective wavenumbers were not peaks in NIR spectra.

The 18 selected MIR wavenumbers by UVE-SPA are 666, 696, 702, 759, 830, 1168, 1454, 1478, 1494, 1541, 1721, 1729, 1747, 1903, 2110, 2397, 2683, and 2880 cm^{-1}. The broad and strong peak from 3200 to 3600 cm^{-1} due to O−H stretching vibrations were not selected. Wavenumber 2880 cm^{-1} was assigned to CH_3 symmetric stretching vibration. Wavenumbers between 2500 and 1900 cm^{-1} were assigned to triple bond stretching vibration. Wavenumber 2110 cm^{-1} was assigned to R-C≡CH. Wavenumbers between 1900 and 1200 cm^{-1} were assigned to double bond stretching vibration. C=O stretching vibration shows a peak at 1747 cm^{-1}. Wavenumbers between 1200 and 600 cm^{-1} were the fingerprint region.

The selected 18 NIR and 18 MIR wavenumbers were set as the input variables of PLS and LS-SVM models, respec-

tively. **Table 2** shows the prediction results based on selected 18 NIR and 18 MIR wavenumbers. For both PLS and LS-SVM models, MIR spectra obtained much better results (with $r^2 = 0.894$ and 0.920 for PLS and LS-SVM, respectively) than NIR (with $r^2 = 0.511$ and 0.593 for PLS and LS-SVM, respectively). LS-SVM did a better prediction performance (with $r^2 = 0.593$ and 0.920 for NIR and MIR spectra, respectively) than PLS (with $r^2 = 0.511$ and 0.894 for NIR and MIR spectra, respectively) again. The predicted versus reference values of iron content in powdered milk based on LS-SVM with 18 NIR (**a**) and 18 MIR (**b**) variables are shown in **Figure 5**. MIR spectroscopy has a good prediction ability of iron content in powdered milk than NIR. The samples in **Figure 5b** were distributed more closely to the regression line, compared with those in **Figure 5a**. For both PLS and LS-SVM models, the results based on UVE-SPA were worse than that only based on UVE in NIR analysis, while UVE-SPA improved the model's prediction ability in MIR analysis, compared with UVE. As there were only 18 wavenumbers for both NIR and MIR anaylses, the training time of both PLS and LS-SVM analysis can be saved (*30*). The UVE-SPA was more parsimonious and simple to interpret for further applications. The overall results indicate that MIR has a good prediction ability of iron content in powdered milk.

Zinc Content Determination. Similar to the analysis step in the previous section, PLS and LS-SVM models were established for the determination of zinc content in powdered milk using NIR and MIR spectra, respectively. **Table 3** shows prediction results of zinc content in powdered milk based on

Table 3. Prediction Results of Zinc Content in Powdered Milk Based on NIR and MIR Spectroscopy

variable selection method	calibration method	NIR variables	r^2	RMSEP	*RPD*	MIR variables	r^2	RMSEP	*RPD*
none	PLS	2756	0.267	1.081	1.295	3727	0.918	0.413	3.390
	LS-SVM	2756	0.460	0.911	1.536	3727	0.936	0.370	3.788
UVE	PLS	146	0.516	0.888	1.577	1322	0.940	0.353	3.966
	LS-SVM	146	0.572	0.797	1.758	1322	0.959	0.289	4.844
UVE-SPA	PLS	17	0.556	0.848	1.651	12	0.900	0.453	3.091
	LS-SVM	17	0.626	0.783	1.788	12	0.946	0.321	4.361

full NIR and MIR spectra. MIR spectra obtained much better results than NIR in both PLS and LS-SVM models. RPD values of MIR based on the LS-SVM model up to 3.7 shows a reliable ability for zinc content determination (*28*). Thus, good determination performance could be achieved by MIR, and NIR cannot show good prediction ability. However, thousands of wavenumbers in both NIR and MIR spectra were also redundant for zinc analysis, and UVE was executed to select informative wavenumbers.

UVE analysis results based on different numbers of LVs show that the smallest RMSECV were obtained based on 12 and 16 LVs for NIR and MIR spectra, respectively. On the basis of optimal LVs, 146 and 1322 wavenumbers were selected and set as input variables of PLS and LS-SVM, respectively. Prediction results are shown in **Table 3**. Although the number of selected variables for MIR was much larger than those for NIR, its prediction result was much better than NIR in both PLS and LS-SVM models. UVE improves the prediction results of both NIR and MIR, compared with results predicted based on the full spectrum. The RPD value up to 4.8 shows that MIR was suitable for screening (*28*).

SPA was carried out to further reduce NIR and MIR wavenumber variables. Seventeen variables (RMSE = 0.902) and 12 (RMSE = 0.433) variables were selected for NIR and MIR, respectively. The 17 selected NIR wavenumbers are 4111, 4218, 4373, 4378, 4383, 4470, 4741, 4755, 4887, 5163, 5183, 5779, 6014, 6133, 6363, 6430, and 6585 cm^{-1}. Some retained wavenumber regions might be attributable to some specific chemical structure related to zinc ion. Wavenumber 4383 cm^{-1} might be C−H stretching + C−H deformation vibration of CH$_3$. Wavenumber 4470 cm^{-1} might be N−H stretching + C=O stretching of amide. Wavenumber 4741 cm^{-1} might be N−H symmetric stretching + amide III of CONH$_2$, CONHR. Wavenumber 5778 cm^{-1} as assigned to 2 × C−H stretching. Other selected wavenumbers, which might not relate to the zinc ion, have contribution to zinc prediction model establishment. The 12 selected MIR wavenumbers are 750, 793, 859, 1423, 1539,

1643, 1679, 1727, 1792, 1861, 2110, and 2741 cm^{-1}. Wavenumber 2110 cm^{-1} was assigned to R-C≡CH. As many effective wavenubmers are between 1900 and 1200 cm^{-1}, which were assigned to double bond stretching vibration, C=O and C=C vibrations might be attributable to the zinc ion. Wavenumbers between 1200 and 600 cm^{-1} were the fingerprint region.

Table 3 shows the prediction results based on selected 17 NIR and 12 MIR wavenumbers. For both PLS and LS-SVM models, MIR spectra obtained much better results (with r^2 = 0.900 and 0.946 for PLS and LS-SVM, respectively) than NIR (with r^2 = 0.556 and 0.626 for PLS and LS-SVM, respectively). Moreover, LS-SVM did a better prediction performance (with r^2 = 0.626 and 0.946 for NIR and MIR spectra, respectively) than PLS (with r^2 = 0.556 and 0.900 for NIR and MIR spectra, respectively). The predicted versus reference values of zinc content in powdered milk based on LS-SVM with 17 NIR (**a**) and 12 MIR (**b**) variables are shown in **Figure 6**. MIR spectroscopy obtained a good result of zinc content prediction in powdered milk. The results in NIR analysis were improved by UVE-SPA, while prediction results in MIR analysis were worse than UVE. However, it was important to notice that the training time using LS-SVM increased with the square of the number of training samples and linearly with the number of variables (*30*); thus, the performances of UVE-SPA were still acceptable because of only 12 MIR wavenumbers, compared with 1322 wavenumbers selected by UVE. Moreover, these 1322 variables from UVE are large numbers and can cause model overfitting. Hence, the application of UVE-SPA would be helpful to reduce computation time and obtain a robust model. Hence, UVE-SPA is an efficient variable selection method and could be useful to reduce computation time. The overall results indicate that MIR has a good prediction ability of zinc content in powdered milk.

From **Tables 2** and **3**, it could be concluded that MIR spectroscopy obtained much better prediction results for trace minerals in powdered milk, such as iron and zinc content. The
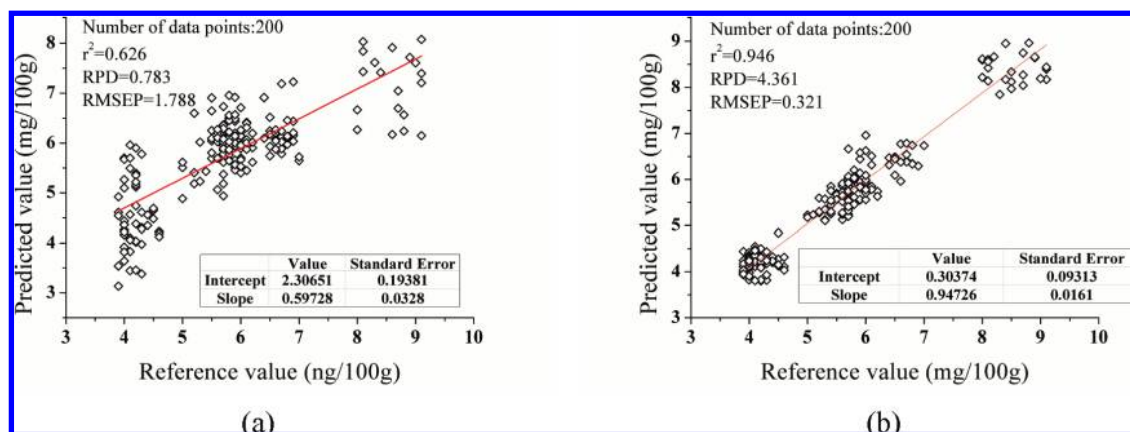

Figure 6. Predicted vs reference values of zinc content in powdered milk by LS-SVM models based on 17 NIR (**a**) and 12MIR (**b**) wavenumbers selected by UVE-SPA.

reason might be that the NIR spectrum has low sensitivity to trace minerals compared to MIR. Although some studies showed that NIR can be applied to detect minerals in plant tissue, such as iron (*19, 20*) and zinc (*20*), as the iron and zinc contents are much smaller in powdered milk, and there is lots of high frequency noise in NIR, it is hard to use NIR spectra to predict iron and zinc contents in powdered milk. In contrast, there are many fundamentals that can usually be observed in isolated positions in the MIRS region (*31*), which might contain metal ion information. MIRS can monitor low levels of key analytes and functional groups (*32*). Although it was hard to distinguish each curve with different metal ion content through naked eyes, the hybrid model of UVE-SPA-LS-SVM can extract useful information from MIR spectra to establish the calibration model and do a precise prediction of iron and zinc content in powdered milk.

Conclusions. NIR and MIR spectroscopy techniques were evaluated to determine iron and zinc content in powdered milk. UVE was first to utilized to eliminate uninformative variables from 2756 and 3727 wavenumber variables in full NIR or MIR spectra, respectively. SPA was then carried out to further select the latent variables. Finally, 18 NIR and 18 MIR variables were selected for iron content prediction, and 17 NIR and 12 MIR wavenumbers were selected for zinc content prediction. A few wavenumbers were obtained by UVE-SPA, compared with that of the full spectrum. The good performance showed a potential application using UVE-SPA to select effective NIR and MIR wavenumbers. In both PLS and LS-SVM models, the selected effective MIR wavenumbers by UVE-SPA obtained much better results than NIR for both iron and zinc content prediction. The overall results indicate that MIR spectroscopy combined with UVE-SPA-LS-SVM could be applied as a precision and rapid method for the determination of trace mineral content in powdered milk, such as iron and zinc. The iron and zinc content prediction using NIR spectroscopy needs to be further investigated. The results might be useful for the process and online monitoring of trace mineral content in other dairy products.

LITERATURE CITED

(1) Torres, M. A.; Sato, K.; Juliano, Y.; Queiroz, Sde S. Powdered milk enriched with iron and ascorbic acid as an intervention measure for treating iron deficiency anemia in children seen at a Basic Health Care Unit. *Arch. Latinoam. Nutr.* **1996**, *46*, 113–117.

(2) Sandstead, H. H. Understanding zinc: Recent observations and interpretations. *J. Lab. Clin. Med.* **1994**, *124*, 322–327.

(3) Institute of Medicine. Food and Nutrition Board. *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc*. National Academy Press: Washington, DC, 2001.

(4) Martino, F. A. R.; Sanches, M. L. F.; Sanz-Medel, A. Total determination of essential and toxic elements in milk whey by double focusing ICP-MS. *J. Anal. At. Spectrom.* **2000**, *15*, 163–168.

(5) Martino, F. A. R.; Sanches, M. L. F.; Sanz-Medel, A. The potential of double focusing-ICP-MS for studying elemental distribution patterns in whole milk, skimmed milk and milk whey of different milks. *Anal. Chim. Acta* **2001**, *442*, 191.

(6) Yebra, M. C.; Moreno-Cid, A.; Cespón, R.; Cancela, S. Preparation of a soluble solid sample by a continuous ultrasound assisted dissolution system for the flow-injection atomic absorption spectrometric determination of iron in powdered milk and infant formula. *Talanta* **2004**, *62*, 403–406.

(7) Chen, K. L.; Jiang, S. J. Determination of calcium, iron and zinc in powdered milk by reaction cell inductively coupled plasma mass spectrometry. *Anal. Chim. Acta* **2002**, *470*, 223–228.

(8) Campillo, N.; Viñas, P.; López-García, I.; Hernández-Córdoba, M. Direct determination of copper and zinc in cow milk, human milk and infant formula samples using electrothermal atomization atomic absorption spectrometry. *Talanta* **1998**, *46*, 615–622.

(9) Ho, C. Y.; Jiang, S. J. Determination of Cr, Zn, Cd and Pb in powdered milk by slurry sampling electrothermal vaporization inductively coupled plasma mass spectrometry. *J. Anal. At. Spectrom.* **2002**, *17*, 688–692.

(10) Nicolaï, B. M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K. I.; Lammertyn, J. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* **2007**, *46*, 99–118.

(11) Wilson, R. H.; Tapp, H. S. Mid-infrared spectroscopy for food analysis: recent new applications and relevant developments in sample presentation methods. *Trac-Trends Anal. Chem.* **1999**, *18*, 85–92.

(12) Wu, D.; Feng, S.; He, Y. Short-wave near-infrared spectroscopy of powdered milk for brand identification and component analysis. *J. Dairy Sci.* **2008**, *91*, 939–949.

(13) Wu, D.; He, Y.; Feng, S. Short-wave near-infrared spectroscopy analysis of major compounds in powdered milk and wavenumber assignment. *Anal. Chim. Acta* **2008**, *610*, 232–242.

(14) Wu, D.; Feng, S.; He, Y. Infrared spectroscopy technique for the nondestructive measurement of fat content in powdered milk. *J. Dairy Sci.* **2007**, *90*, 3613–3619.

(15) Wu, D.; He, Y.; Feng, S.; Sun, D. W. Study on infrared spectroscopy technique for fast measurement of protein content in powdered milk based on LS-SVM. *J. Food Eng.* **2008**, *84*, 124–131.

(16) Borin, A.; Ferrao, M. F.; Mello, C.; Maretto, D. A.; Poppi, R. J. Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk. *Anal. Chim. Acta* **2006**, *579*, 25–32.

(17) Givens, D. I.; Deaville, E. R. The current and future role of near infrared reflectance spectroscopy in animal nutrition: a review. *Aus. J. Agric. Res.* **1999**, *50*, 1131–1145.

(18) Matteson A.; Herron, M. M. Quantitative Mineral Analysis by Fourier Transform Infrared Spectroscopy. Presented at the SCA Conference, 1993; Paper Number 9308.

(19) Hallett, R. A.; Hornbeck, J. W.; Martin, M. E. Predicting elements in white pine and red oak foliage with visible-near infrared reflectance spectroscopy. *J. Near Infrared Spectrosc.* **1997**, *5*, 77–82.

(20) Cozzolino, D.; Moron, A. Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes. *Animal Feed Sci. Technol.* **2004**, *111*, 161–73.

(21) Mesubi, M. An infrared study of zinc, cadmium, and lead salts of some fatty acids. *J. Mol. Struct.* **1982**, *81*, 61–71.

(22) Bochmann, M.; Chesters, M. A.; Coleman, A. P.; Grinter, R.; Linder, D. R. The infrared spectra of the dimethyl complexes of zinc, cadmium and mercury. *Spectrochim. Acta, Part A* **1992**, *48*, 1173–1178.

(23) Centner, V.; Massart, D. L.; de Noord, O. E.; de Jong, S.; Vandeginste, B. M.; Sterna, C. Elimination of uninformative variables for multivariate calibration. *Anal. Chem.* **1996**, *68*, 3851.

(24) Araújo, M. C. U.; Saldanha, T. C. B.; Galvão, R.K. H.; Yoneyama, T.; Chame, H. C.; Visani, V. *Chemom. Intell. Lab. Syst.* **2001**, *57*, 65–73.

(25) Breitkreitz, M. C., Jr.; Raimundo, I. M.; Rohwedder, J. J.; Celio Pasquini, R.; Dantas Filho, H. A.; Joséb, G. E.; Araújo, M. C. U. Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration. *Analyst* **2003**, *128*, 1204–1207.

(26) Ye, S. F.; Wang, D.; Min, S. G. Successive projections algorithm combined with uninformative variable elimination for spectral variable selection. *Chemom. Intell. Lab. Syst.* **2008**, *91*, 194–199.

(27) Wu, D.; Yang, H. Q.; Chen, X. J.; He, Y.; Li, X. L. Application of image texture for the sorting of tea categories using multi-

spectral imaging technique and support vector machine. *J. Food Eng*. **2008**, *88*, 474–483.

(28) Williams, P. C. Implementation of Near-Infrared Technology. In *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P., Norris, K., Eds.; American Association of Cereal Chemists: St. Paul, MN, 2001; pp 145−169.

(29) Kennard, R. W.; Stone, L. A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148.

(30) Chauchard, F.; Cogdill, R.; Roussel, S.; Rogera, 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.

(31) Karoui, R.; De Baerdemaeker, J. A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products. *Food Chem*. **2007**, *102*, 621–640.

(32) Roychoudhury, P.; Harvey, L. M.; McNeil, B. At-line monitoring of ammonium, glucose, methyl oleate and biomass in a complex antibiotic fermentation process using attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy. *Anal. Chim. Acta* **2006**, *561*, 218–224.