

Application of near-infrared spectroscopy for screening the potato flour content in Chinese steamed bread

Hui Wang² · Du Lv² · Nan Dong² · Sijie Wang³ · Jia Liu^{1,2}

Received: 6 May 2018 / Revised: 20 December 2018 / Accepted: 28 December 2018 / Published online: 23 February 2019
© The Korean Society of Food Science and Technology 2019

Abstract Near-infrared (NIR) spectroscopy combined with chemometrics was used as a technique to predict the potato flour content in Chinese steamed bread (CSB). The inner core of CSB was chosen as the measuring position for acquiring the NIR spectra. Spectra between 4000 and 10,000 cm^{-1} were analysed using a partial least-squares regression. The coefficient of determination (R_{CV}^2) and the root mean square error of cross-validation in the calibration set were found to be 0.7940–0.8955 and 4.22–5.93, depending on the pre-treatment of the spectra. The external validation set gave an R^2 and a ratio to performance deviation of 0.8865 and 3.07. Reasonable recovery

(93.1–102.5%) and good intra-assay (3.3–8.3%) and inter-assay (7.6–17.2%) precision illustrated the feasibility of this method. The result of this study reveals that NIR spectroscopy could be used as rapid tool to determine the potato flour content in CSB (> 20%).

Keywords Chinese steamed bread · Near-infrared · Partial least squares regression · Potato flour · Potato staple food

Introduction

Potato (*Solanum tuberosum* L.) is an important food crop and the fourth largest staple crop in China after rice, wheat and maize (Wang et al., 2011). Their higher yield per unit area and nutritional value have led to an increase in potato production over the past years (Wang et al., 2015). In the past, the majority of potatoes in China were consumed in Chinese cuisine and freshly fried or roasted. Incorporating potatoes into food production and processing is still in its infancy. However, potatoes can provide humans with a large proportion of their dietary daily energy intake compared with other crops. Thus, potato is a good supplier of many nutrients, including carbohydrates, protein, minerals and vitamins (Tian et al., 2016). In 2015, recognizing the importance of the potato industry, the Ministry of Agriculture of the People's Republic of China implemented a plan to boost potato production and make the tuber one of the nation's staple foods. According to the plan, China will have more than 6.67 million hectares of potato planting area, and at least 30% could be processed into staple foods. Thus, the government and research institutes have been designing and planning potato plantings in accordance with local resources, diversifying the variety of potato-incorporated staple foods (PISF) based on market demand,

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10068-018-00552-x>) contains supplementary material, which is available to authorized users.

✉ Jia Liu
mcgrady456@163.com

Hui Wang
450155684@qq.com

Du Lv
1002651607@qq.com

Nan Dong
460288583@qq.com

Sijie Wang
394344127@qq.com

¹ National Engineering Research Center of Seafood, School of Food Science and Technology, Dalian Polytechnic University, Dalian 116034, People's Republic of China

² Institute of Food Processing Technology, Guizhou Academy of Agricultural Science, Guiyang 550006, People's Republic of China

³ School of Liquor and Food Engineering, Guizhou University, Guiyang 550025, People's Republic of China

improving consumer demand for PISF production, modernizing food processing and manufacturing PISFs.

Potato flour has been incorporated into wheat and rice flour-based foods for nutrition fortification because potato flour has a higher nutritional value than either wheat or rice flour (Please see Supplementary Material 1). The first generation of PISFs, including noodles and Chinese steamed bread (CSB), are already on the market. These products were prepared with a high content of potato flour (> 20%). Generally, the higher the potato flour content, the higher the price of the PISF. However, there are no significant chemical differences between potato and wheat or rice flours. They all share the same major chemical constituent, namely starch, which makes it difficult to determine the real potato flour contents in PISFs. Substituting a lower potato flour content for a higher one, without disclosure, can be very lucrative for the manufacturer or supplier of raw materials because their profit margins can be greatly increased. Therefore, it is imperative to develop a rapid determination method for the content of potato flour in PISFs to prevent major economic fraud and maintain market order.

Near-infrared (NIR) spectroscopy has been frequently used for the analysis of agricultural products and other foods for the past several years (Rady and Guyer, 2015). NIR covers a broad range of the electromagnetic spectrum between 780 and 2500 nm and is associated with the overtones and combination of fundamental O–H, C–H and N–H bonds, which are the primary structural components of organic molecules, such as protein, starch and water (Futami et al., 2016). However, overlapping and broad absorption peaks make it difficult to complete quantitative analyses using NIR spectra. Modern multivariate statistical techniques, including partial least-squares (PLS) regressions and principal components analyses (PCAs), aid in rapidly analyzing a large number of data values acquired using different analytical techniques. The combination of NIR and chemometric methods has been successfully used in the classification and prediction of potato properties (Ni et al., 2011), assessment of potato dry matter (Bernhard et al., 2016; Helgerud et al., 2012; 2015), phytochemical determination and classification in purple- and red-fleshed potato tubers (Tierno et al., 2015), detection of potato resistant starch content (Ayvaz and Rodriguez-Saona, 2015), screening of acrylamide content in commercial potato chips, discrimination of purple sweet potatoes and their samples adulterated with the white sweet potato flour (Ding et al., 2015), quality control of potato chips (Shiroma and Rodriguez-Saona, 2009) and on-line monitoring of fat, dry matter and acrylamide contents in potato chips (Ade-dipe et al., 2016; Pedreschi et al. 2010). However, there have been limited reports on methods for determining the potato flour contents of PISFs. Thus, this study aimed to

investigate the practicality of using NIR spectroscopy as a rapid method for the quantitative analysis of the potato flour content in a PISF, using CSB as the model.

Materials and methods

Materials

Wheat flour with a high gluten strength (Wudeli Flour Co., Ltd., Hebei, China and Luwang Co., Ltd., Shandong, China) and potato flour (Linkage Potato Co., Ltd., Inner Mongolia, China; Yanbei Potato Co., Ltd., Zhangjiakou, China and Sanlai Food Co., Ltd., Shanxi, China) were purchased. A commercial instant dried yeast (Angel Yeast Co., Ltd., Hubei, China), flour improving agent (Angel Yeast Co., Ltd., Hubei, China), food-grade sucrose and salt were purchased from a local supermarket.

Preparation of CSB

CSB was prepared with 0–42% potato flour and the basic potato flour-incorporated dough constituents as follows: Wheat flour (58–100%), potato flour (0–42%) and water (31–37%) were combined to form a dough according to the relative ratios. Then, yeast (0.9 g), sucrose (2.5 g), salt (0.3 g) and a steamed bread flour improving agent (0.6 g) were added to doughs having different ratios of potato flour. After all of the ingredients were mixed and kneaded to form dough, the dough was subsequently fermented, sheeted, rolled, divided, proofed, steamed and cooled (Shiau et al., 2015).

Instrumental analysis

The NIR spectra of CSB and potato flour-incorporated CSB (24 h after CSB preparation) were measured on a NIR equipment (Antaris II, Thermo Fisher Scientific Inc., Massachusetts, USA). An internal gold background, as the reference, and all of the spectra were measured using an InGaAs detector. The background was corrected every hour to eliminate the influence of baseline shifts. The NIR spectra of samples were recorded at 4 cm⁻¹ intervals over a spectral range from 4000 to 10,000 cm⁻¹ for the multivariate analysis. To improve the signal-to-noise ratio significantly, the average of 64 scans was chosen for each spectrum. The temperature during the test was maintained at room temperature (25 ± 1 °C), and the humidity was kept at a stable level by a dehumidifier.

Chemometric software for data processing and statistical techniques

The data were processed using the Unscrambler (Version 9.7; Camo Inc., Trondheim, Norway). The chemometric procedures and test arrangements are discussed below.

Calibration and prediction sample sets

The potato flour (potato granules) content in the CSB ranged from 0 to 42%. The potato flour contents in the 132 samples of the calibration set were 0%, 2%, 4%, 6%, 8%, 10%, 12%, 14%, 16%, 18%, 20%, 22%, 24%, 26%, 28%, 30%, 32%, 34%, 36%, 38%, 40% and 42% (w/w). The calibration set was divided into two groups, wheat flour (Wudeli Flour Co., Ltd.) with potato flour (Linkage Potato Co., Ltd.) incorporated and wheat flour (Wudeli Flour Co., Ltd.) with potato flour (Sanlai Food Co., Ltd.) incorporated. The potato flour content in 66 samples from the validation set also ranged from 0 to 42% (w/w) in 2% intervals. The wheat flour (Luwang Co., Ltd.) in the external validation was incorporated with potato flour (Yanbei Potato Co., Ltd.). In addition, the samples in the calibration and external validation sets were prepared by different members of the laboratory to ensure that there was no association between the calibration and external validation sets.

The confirmation of measuring position

To determine the best position for acquiring a stable and credible spectrum of the CSB, three parts (the top, inner core and the bottom) of the CSB were evaluated by comparing the parameters, namely the root mean square error of calibration (RMSEC), in the calibration model, and the coefficient of determination between the predicted and the measured parameters in the calibration (R_c^2). The potato flour contents in 30 samples (6 samples for each potato flour content) were set as 0%, 5%, 10%, 15% and 20%. These CSBs were prepared using wheat flour (Wudeli Flour Co., Ltd.) and potato flour (Linkage Potato Co., Ltd.).

Calibration step: cross validation

In the present study, the linear prediction models between the spectral data and the potato flour content of the CSB were constructed using a PLS analysis. The calibration models in this study were developed using PLS regressions with leave-one-out cross validation, which creates sample subsets that are missing one sample and the analysis predicts the missing sample, iteratively, until every sample in the calibration set has been left out once (Porfire et al., 2012). First, one sample was randomly chosen by the

software and removed from the calibration set, while the rest of the spectra were used to construct a PLS model. Then, this procedure was repeated over again until every sample in the calibration set had been left out.

The leverage was computed and used to detect outlying samples in the original data set, when the model calculation was complete. Outliers are samples with high leverage and studentized residuals, which compromise the precision of the model (Adedipe et al., 2016). The parameters, such as RMSEC, the root mean square error of cross-validation (RMSECV), RMSEP, the root mean square error of prediction (RMSEP), R_c^2 and the coefficient of the cross-validation (R_{cv}^2) were used to evaluate the performance of the PLS models. In addition, the ratio to performance deviation (RPD) was introduced to measure the precision of the PLS model based on the root mean square values (Williams and Sobering, 1993):

$$RPD = \delta / RMSEP \quad (1)$$

where δ represents the standard deviation of the validation set. An RPD value of < 1 indicates that the established model has no predictive power; $1 \leq RPD < 2$ indicates that the established model can only separate the lower from higher values; $2 \leq RPD < 3$ indicates that a reasonable model is already established; and $3 < RPD$ value indicates that an excellent model already established (Landsberg and Waring, 1997).

To improve the reliability, accuracy and stability of the calibration models, data were pre-processed before analyses. The steps used are as follows (Krepper et al., 2017; Sunoj et al., 2016):

- (1) Automatic baseline correction, automatic smooth processing (Savitzky–Golay, three data points) and normalization;
- (2) First derivative (Savitzky–Golay, polynomial = 2) was calculated by different data point gaps (5, 9 and 13 pts); and
- (3) Second derivative (Savitzky–Golay, polynomial = 2) was calculated by different data point gaps (5, 9 and 13 pts).

The optimum number of PLS factors was determined by plotting the RMSECV against the number of factors and determining the minimum of the plot (Winkler-Moser et al., 2015). Additionally, the best model for the determination of the potato flour content should have the highest RPD (Mantanus et al., 2010).

External validation step

The external validation of the method was checked by adding the NIR spectra of 198 new samples (potato flour content: 2–40%) to the optimal model (Bin et al., 2016).

These samples were prepared with wheat flour and potato flour from different sources (different brands). Finally, the RMSEP, the coefficient of determination (R^2) for the validation and the RPD of the model were evaluated.

Analytical characteristics test

The analytical characteristics of this method were evaluated according to our previous report (Liu et al., 2013).

$$\text{Recovery}(\%) = (\text{P-NIR}/\text{P-SS}) \times 100\% \quad (2)$$

where P-NIR represents proportion of potato flour in sample detected by NIR spectroscopy, and P-SS represents proportion of potato flour in sample spiked by laboratory personnel. The intra-assay relative standard deviation (RSD) was determined for each proportion in three replicates. The inter-assay RSD was evaluated for each proportion prepared by 20 laboratory personnel.

PCA analysis

PCA analysis was carried out using the Unscrambler (Version 9.7, Camo Inc., Trondheim, Norway). To determine the number of principal components that are statistically significant in capturing the underlying features in data sets, a leave-one-out cross validation method (Wold et al., 2001) was performed.

Results and discussion

NIR spectra of potato and wheat flours

The raw NIR spectra of potato and wheat flours are presented in Fig. 1A. The NIR spectroscopy showed that the peak at approximately 4015 cm^{-1} belongs to a combination of the C–H and C–C stretching vibrations (Xu et al., 2015). The peak at 4308 cm^{-1} could also be ascribed to a combination of the C–H and C–C stretching vibrations (Xu et al., 2015). The peaks at approximately 4747 cm^{-1} and 5168 cm^{-1} represented the combination of the C–O stretch and O–H deformation and a combination of the baseband of the O–H stretch and the first overtone of C–O deformation (Li et al., 2012), respectively. A few bands between $5418\text{--}5931 \text{ cm}^{-1}$ and $5982\text{--}6811 \text{ cm}^{-1}$ belonged to the first overtones of the C–H stretch of various groups and the first overtone of the O–H stretch or N–H stretch, respectively. The last peak at approximately 8320 cm^{-1} represented the second overtones of the C–H stretches of various groups (Vagnini et al., 2009). The raw NIR spectra of the potato and wheat flours showed insignificant differences, which meant that the determination of potato flour in CSB was obstructed without the aid of chemometrics.

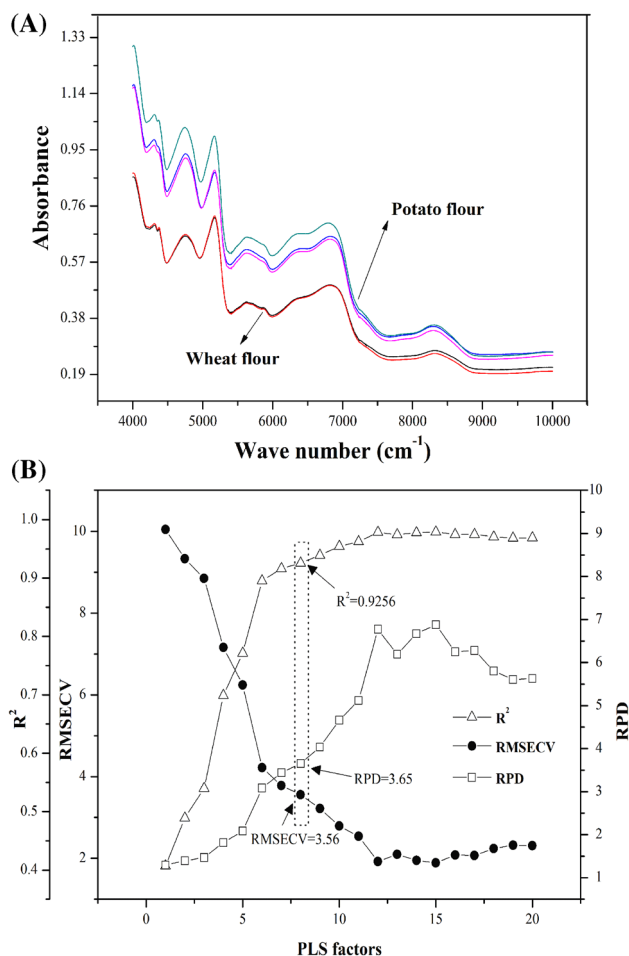


Fig. 1 NIR spectra of potato and wheat flour (A), RMSECV (root-mean-square error of cross-validation), R^2_{CV} (coefficient of determination of cross-validation) and RPD (the ratio to performance deviation) plotted as a function of the number of factors used in the PLS (partial least squares) model (B)

The confirmation of measuring position

Owing to different water evaporation rates, three positions, the top, inner core and bottom layer of CSB, exhibited different surface layer textures. Thus, it was necessary to determine the part that was most suitable for acquiring the NIR spectra. For each position, the PLS model was calculated, and the results are summarized in Table 1. According to the observed RMSECV and R^2_{cv} values, the NIR spectra from the top layer and inner core showed good performance. Additionally, the inner core produced the most accurate determination of potato flour in CSB. Thus, the inner core of CSB was chosen as the position for acquiring the NIR spectra.

Table 1 Results of calibration and cross-validation of the PLS models at the different measuring position of CSB

Test position	PLS factors	Calibration		Cross-validation	
		R_c^2	RMSEC(%)	R_{cv}^2	RMSECV(%)
The top layer	10	0.9855	0.8588	0.9531	1.5443
Inner core	10	0.9943	0.5323	0.9615	1.4432
The bottom layer	10	0.9719	1.1858	0.8122	3.1074

^aCoefficient of determination in calibration^bRoot-mean-square error of calibration^cCoefficient of determination in cross-validation^dRoot-mean-square error of cross-validation

Calibration and validation of the PCA and PLS procedures

Optimization of the calibration model

In the present study, the full spectral range of 4000–10,000 cm^{-1} was used to determine the potato flour content in CSB. The spectral pre-treatments and the number of PLS factors were investigated in this study to determine their affects. Before calibration, outliers should be removed. Therefore, the leverage values of the PCA were carefully checked. After the removal of the outliers, the PLS models were recalculated for each pre-treatment. Various spectral pre-treatments for the determination of the potato flour content in CSB are summarized in Table 2.

Prediction results based on the spectral data sets after treatments with different pre-processing methods were compared with the ones using raw spectra. The worst prediction, with $R_{CV}^2 = 0.794$ and $\text{RMSECV} = 5.93$, was

obtained by directly using the raw spectra. However, the calibration results from different spectral treatments were very close. Positive correlations were found between the values predicted with the NIR models and the proportion of potato flour additions, which produced desirable R_{CV}^2 and RMSECV values of 0.8592–0.8955 and 4.22–5.09, respectively. Regression models based on the first and second derivatives of the NIR spectra resulted in improved PLS prediction model performances. In general, the NIR-reflectance measurement of a sample measures the diffusively reflected (the primary source of information in the NIR spectra) and specularly reflected (mirror-like reflections that do not contain any chemical information) radiation. The selected pre-processing methods can remove some, but not all, of the undesired scatter or particle-size information in the spectra. The most basic method for derivation is finite differences, in which the first derivative is estimated as the difference between two subsequent spectral measurement points. The second order derivative

Table 2 Results of calibration and validation of the PLS models on the raw spectra and spectra with various pretreatments

Mathematical pretreatment	Outliner	PLS factors	Calibration		Cross-validation	
			R_c^{2a}	RMSEC ^b	R_{CV}^{2c}	RMSECV ^d
Raw spectra	4	6	0.8277	5.38	0.7940	5.93
Spectra pre-treatment ^f	3	6	0.8870	4.36	0.8592	4.90
First derivative (5pts) ^{f,g}	2	6	0.9210	3.64	0.8955	4.22
First derivative (9 pts) ^{f,g}	2	6	0.9106	3.87	0.8721	4.67
First derivative (13 pts) ^{f,g}	3	6	0.8853	4.39	0.8482	5.09
Second derivative (5 pts) ^{f,g}	2	6	0.9220	3.62	0.8922	4.29
Second derivative (9 pts) ^{f,g}	2	6	0.9131	3.82	0.8930	4.31
Second derivative (13pts) ^{f,g}	2	6	0.8986	4.12	0.8737	4.64

^aCoefficient of determination in calibration^bRoot-mean-square error of calibration^cCoefficient of determination in cross-validation^dRoot-mean-square error of cross-validation^eStandard deviation ratio of cross-validation^fPre-treatments of spectra include automatic baseline correction, automatic smooth processing and normalisation^gPts, different data point gaps

is then estimated by calculating the difference between two successive points of the first-order derivative spectra. Both derivation techniques can improve the prediction behavior of the model by increasing the signal-to-noise ratio. Thus, the prediction behavior of the model was improved after the pre-treatment (Li et al., 2010).

The pre-processing method of the first derivative (5 pts) was selected as the spectral pre-treatment method that resulted in the best prediction performance, having the lowest RMSECV of 4.22 and the highest R_{CV}^2 of 0.8955, for the PLS prediction model. Furthermore, 20 PLS factors were introduced into the PLS model to optimize of the number of PLS factors. The RMSECV, R_{CV}^2 and RPD values were recorded as functions of the PLS factors for determining the potato flour content in CSB using the pre-processing method of the first derivative (5 pts) as presented in Fig. 1B. The RMSECV value decreased sharply within the initial 12 PLS factors. Adding more PLS factors to the model improved its linearity, resulting in R_{CV}^2 and RPD increasing and RMSECV decreasing as the number of PLS factors increased (Rinnan et al., 2009). However, too many PLS factors (> 8) led to the model running the risk of over-fitting (Feng et al., 2015; Hacisalihoglu et al., 2016). The RPD values were relatively high (3.65–6.77) when there were 8–20 PLS factors. To balance the advantages and disadvantages, in this study, the optimal number of PLS factors was set as eight.

PCA

To produce a clear 2D map, not all of the samples were used in the PCA. The score plot of the first two PCs of the PCA was carried out by analyzing 36 samples (0%, 6%, 10%, 20%, 30% and 40%) with the pre-treatment of the first derivative (5 pts), as can be seen in Fig. 2A. PC1 and PC2 account for 96.54% and 2.84% of the total variation among these samples, respectively. CSBs with different potato flour contents were located at different parts of the 2D map. The whole wheat CSB lay on the positive side of both PCs (top right corner), and the 40% potato flour-incorporated CSB lay mostly along the negative value area of PC1 and positive value area of PC2 (top left corner). However, others did not produce apparent regional distributions. The X-loadings plot indicates how well a variable was taken into account by the model components, and it was used to understand how much each variable contributed to meaningful variation in the data and to interpret variable relationships (Wójcicki et al., 2015). The PCA loading spectra of the first two factors are shown in Fig. 2B. They accounted for 99.38% of the total variation and revealed characteristics among the samples. The highest negative loading in PC1 was found at approximately 5299 cm^{-1} , as presented in Fig. 2B. Positive

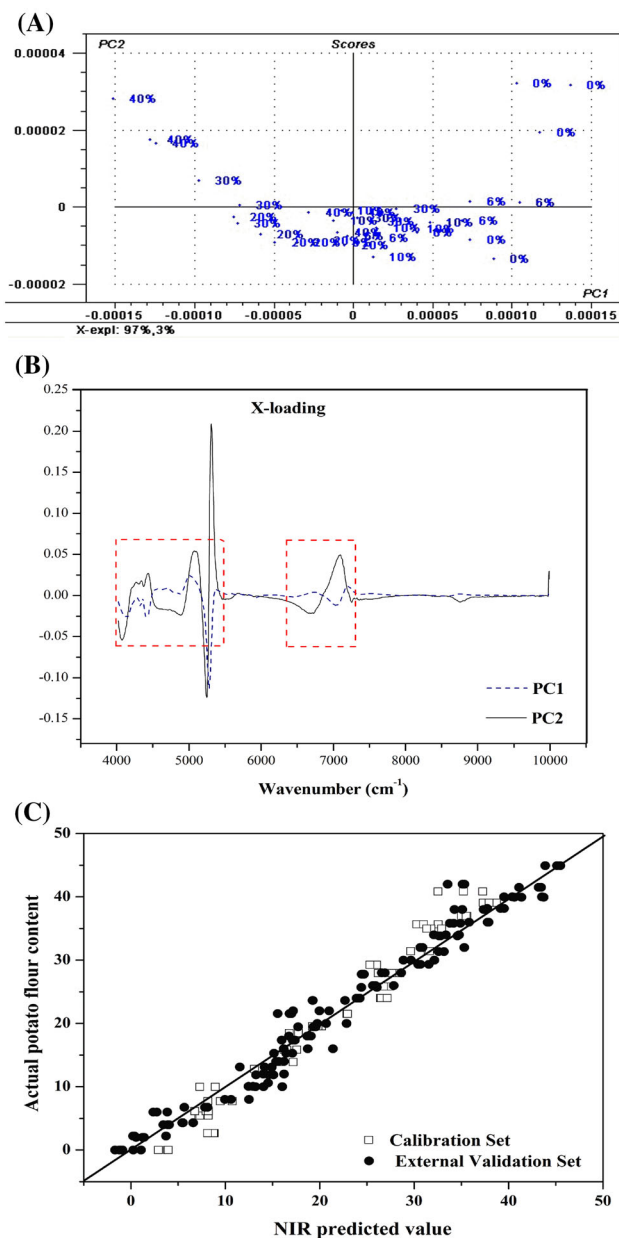


Fig. 2 Two-D scores plots (A) and the loading plots (B) of PCA results (total 36 samples, 0%, 6%, 10%, 20%, 30% and 40%) and NIR predicted value vs. Actual potato flour content (C) in the calibration set (open square) and external validation set (filled circle)

loadings on PC 2 were found in the range of $4169\text{--}5546\text{ cm}^{-1}$ and $6807\text{--}7262\text{ cm}^{-1}$. These regions could help us to understand the differences between the potato and wheat flours, which requires intensive study.

External validation

An external validation was carried out to check the predictive ability of the optimal PLS model by retraining the NIR spectroscopy of 198 samples in the external validation

Table 3 Analytical characteristics of the proposed method

Parameters	Potato flour content (%)				
	12.9	21.4	23.3	33.1	40.2
Recovery (%) ^a	93.1 ± 3.8	98.5 ± 4.2	99.3 ± 3.1	101.2 ± 2.5	102.5 ± 1.4
Intra-assay RSD (%) ^b	8.3	3.6	6.2	4.5	3.3
Inter-assay RSD (%) ^c	17.2	9.3	8.4	7.6	9.8

^aRecovery was calculated from following equation: (proportion of potato flour in sample detected by NIR spectroscopy/proportion of potato flour in sample spiked by laboratory personnel) × 100%

^bIntra-assay relative standard deviation (RSD) was measured on each potato flour content in three replicates

^cInter-assay relative standard deviation (RSD) was evaluated on each proportion prepared by 20 laboratory personnel

Table 4 Influence of moisture content on the stability of the proposed method

Moisture content (%)	Potato flour content (%)			
	13.5	22.6	30.5	38.2
30–33	12.6 ± 1.3	21.3 ± 1.7	29.2 ± 0.8	37.5 ± 2.3
24–28	11.3 ± 0.9	20.4 ± 2.3	28.5 ± 2.5	35.8 ± 2.8
18–22	10.2 ± 1.1	19.8 ± 1.9	27.5 ± 2.1	33.2 ± 2.6

set, and the results are shown in Fig. 2C. The optimal PLS prediction model was performed with eight PLS factors and produced satisfactory R^2 and RPD values of 0.8865 and 3.07, respectively, in the external validation. NIR technology provides a potential analysis tool for the prediction of the potato flour content in CSB.

Analytical characteristics of the PLS method

To better understand the practicality of using NIR spectra for determining the potato flour content in CSB, recovery and both intra- and inter-assay RSDs of this model were evaluated. As demonstrated in Table 3, recoveries of between 93.1 and 102.5% were obtained in tests between the actual potato flour contents and NIR-predicted values. The intra-assay RSD for each proportion, prepared in triplicate, ranged from 3.3 to 8.3%, and the inter-assay RSD for each proportion, prepared by different laboratory personnel, ranged from 7.6 to 17.2%. Thus, the NIR technology exhibited satisfactory sensitivity and repeatability levels for samples with potato flour contents greater than 20%, thereby providing a feasible way of distinguishing CSBs with high (> 20%) potato flour contents. Table 4 showed the effect of moisture content on the potato flour content in CSB which was predicted by the NIR model. Samples with a potato flour content ranged from 13.5 to 38.2% were thermally dried to 3 level of moisture content. As we can see from the table, the predicted values from the NIR model was decreased as the decrease of moisture content. The moisture content had a significant influence on the NIR models for prediction of composition

of materials. The reason may ascribe to the change in intensity and positions of absorbance of –OH influenced by the interaction between water and other components (Williams, 2009).

The ability of NIR technology combined with PLS regression to determine the potato flour content in CSB was investigated. The ideal calibration model with eight PLS factors ($R_{CV}^2 = 0.9256$, RPD = 3.65) for the determination of potato flour content in CSB was developed using NIR spectra from 4000 to 10,000 cm^{-1} and a pre-processing combination of baseline correction, smooth processing, normalization, and a first derivative with five data point gaps. CSB prepared using potato and wheat flour of different brands were used for the external validation, and the optimal calibration showed a good prediction performance ($R^2 = 0.8865$, RPD = 3.07). This is the first application of NIR spectroscopy combined with the chemometric method to determine the potato content in a PISF. Moreover, this work described a quick and non-destructive analysis method compared with the tedious routine analyses, such as gel electrophoresis and PCR.

To make the application of this NIR methodology more representative and practical, it is advisable to consider more factors, such as the storage time and method (refrigeration and freezing) of the CSB, the raw material form (mashed potato and potato flour prepared using different processing methods), the brand (the more brands that are considered in the preparation, the more accurate the calibration) and other (moisture, starch retrogradation, and adulteration of potato starch). In addition, more research should be performed to confirm the general applicability of

NIR for potato flour content determination in other staple foods.

Acknowledgements This research was conducted by financial supported from Scientific and Technological Fund of Guizhou Province ([2017] 1179), Science Technology Foundation of Guizhou Province ([2015] 7072), Construction Project of Innovative Talents Base of Guizhou Province ([2016] 22), and the Scientific and Technological Project of Guizhou Province ([2014] 6016). We also thank the Analytical and Testing Centre in Guizhou Normal University (Guiyang, China) for providing the NIR instrument.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Adedipe OE, Johanningsmeier SD, Truong VD, Yencho GC. Development and validation of a near-infrared spectroscopy method for the prediction of acrylamide content in French-fried potato. *J. Agric. Food Chem.* 64: 1850–1860 (2016)
- Ayvaz H, Rodriguez-Saona LE. Application of handheld and portable spectrometers for screening acrylamide content in commercial potato chips. *Food Chem.* 174: 154–162 (2015)
- Bin J, Ai FF, Fan W, Zhou JH, Yun YH, Liang YZ. A modified random forest approach to improve multi-class classification performance of tobacco leaf grades coupled with NIR spectroscopy. *RSC Adv.* 6: 30353–30361 (2016)
- Bernhard T, Truberg B, Friedt W, Snowdon R, Wittkop B. Development of near-infrared reflection spectroscopy calibrations for crude protein and dry matter content in fresh and dried potato tuber samples. *Potato Res.* 59: 149–165 (2016)
- Ding X, Ni Y, Kokot S. NIR spectroscopy and chemometrics for the discrimination of pure, powdered, purple sweet potatoes and their samples adulterated with the white sweet potato flour. *Chemometr. Intell. Lab.* 144: 17–23 (2015)
- Feng YZ, Downey G, Sun DW, Wals HD, Xu JL. Towards improvement in classification of *Escherichia coli*, *Listeria innocua* and their strains in isolated systems based on chemometric analysis of visible and near-infrared spectroscopic data. *J. Food Eng.* 149: 87–96 (2015)
- Futami Y, Ozaki Y, Ozaki Y. Absorption intensity changes and frequency shifts of fundamental and first overtone bands for OH stretching vibration of methanol upon methanol–pyridine complex formation in CCl₄: analysis by NIR/IR spectroscopy and DFT calculations. *Phys. Chem.* 18: 5580–5586 (2016)
- Hacisalihoglu G, Gustin JL, Louisma J, Armstrong P, Peter GF, Walker AR, Settles AM. Enhanced single seed trait predictions in soybean (*Glycine max*) and robust calibration model transfer with near-infrared reflectance spectroscopy. *J. Agric. Food Chem.* 64: 1079–1086 (2016)
- Helgerud T, Segtnan VH, Wold JP, Ballance S, Knutsen SH, Rukke EO, Afseth NK. Near-infrared spectroscopy for rapid estimation of dry matter content in whole unpeeled potato tubers. *J. Food Res.* 1: 55–65 (2012)
- Helgerud T, Wold JP, Pedersen MB, Liland KH, Ballance S, Knutsen SH, Rukke EO, Afseth NK. Towards on-line prediction of dry matter content in whole unpeeled potatoes using near-infrared spectroscopy. *Talanta* 143: 138–144 (2015)
- Krepper G, Romeoa F, Fernandes DDS, Dinizc PHGD, de Araujo MCU, Di Nezio MS, Pistonesi MF, Centurión ME. Determination of fat content in chicken hamburgers using NIR spectroscopy and the Successive Projections Algorithm for interval selection in PLS regression (iSPA-PLS). *SPECTROCHIM ACTA A* 189: 300–306 (2017)
- Landsberg JJ, Waring RH. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecol. Manag.* 95: 209–228 (1997)
- Liu J, Wen Y, Dong N, Lai C, Zhao GH. Authentication of lotus root powder adulterated with potato starch and/or sweet potato starch using Fourier transform mid-infrared spectroscopy. *Food Chem.* 141: 3103–3109 (2013)
- Li B, Wei Y, Duan H, Xi L, Wu X. Discrimination of the geographical origin of *Codonopsis pilosula* using near infrared diffuse reflection spectroscopy coupled with random forests and *k*-nearest neighbor methods. *Vib. Spectrosc.* 62: 17–22 (2012)
- Li WL, Xing LH, Fang LM, Wang J, Qu HB. Application of near infrared spectroscopy for rapid analysis of intermediates of Tanreqing injection. *J. Pharmaceut. Biomed.* 53 350–358 (2010)
- Mantanus J, Ziemons E, Lebrun P, Rozet E, Klinkenberg R, Streel B, Evrard B, Hubert P. Active content determination of non-coated pharmaceutical pellets by near infrared spectroscopy: method development, validation and reliability evaluation. *Talanta* 80: 1750–1757 (2010)
- Ni Y, Mei M, Kokot S. Analysis of complex, processed substances with the use of NIR spectroscopy and chemometrics: Classification and prediction of properties - The potato crisps example. *Chemometr. Intell. Lab.* 105: 147–156 (2011)
- Porfire A, Rus L, Vonica AL, Tomuta I. High-throughput NIR-chemometric methods for determination of drug content and pharmaceutical properties of indapamide powder blends for tableting. *J. Pharmaceut. Biomed.* 70: 301–309 (2012)
- Pedreschi F, Segtnan VH, Knutsen SH. On-line monitoring of fat, dry matter and acrylamide contents in potato chips using near infrared intertance and visual reflectance imaging. *Food Chem* 121: 616–620 (2010)
- Rinnan Å, van den Berg F, Engelsen SB. Review of the most common pre-processing techniques for near-infrared spectra. *Trac-Trend Anal. Chem.* 28: 1201–1222 (2009)
- Rady AM, Guyer DE. Rapid and/or nondestructive quality evaluation methods for potatoes: A review. *Comput. Electron. Agr.* 117: 31–48 (2015)
- Sunoj S, Ighathinathane C, Visvanathan R. Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. *Comput. Electron. Agr.* 124: 234–242 (2016)
- Shiroma C, Rodriguez-Saona L. Application of NIR and MIR spectroscopy in quality control of potato chips. *J. Food Compos. Anal.* 22: 596–605 (2009)
- Shiau SY, Wu MY, Liu YL. The effect of pineapple core fiber on dough rheology and the quality of mantou. *J. Food Drug Anal.* 23: 493–500 (2015)
- Tian JH, Chen JC, Ye XQ, Chen SG. Health benefits of the potato affected by domestic cooking: A review. *Food Chem.* 202: 165–175 (2016)
- Tierno R, López A, Riga P, Arazuri S, Jaren C, Benedicto I, de Galarreta JIR. Phytochemicals determination and classification in purple and red fleshed potato tubers by analytical methods and near infrared spectroscopy. *J. Sci. Food Agr.* 96: 1888–1899 (2015)
- Vagnini M, Miliani C, Cartechini L, Rocchi P, Brunetti BG, Sgamellotti A. FT-NIR spectroscopy for non-invasive identification of natural polymers and resins in easel paintings. *Anal. Bioanal. Chem.* 395: 2107–2118 (2009)

- Wang B, Ma YL, Zhang ZB, Wu ZM, Wu YF, Wang QH, Li MF. Potato viruses in China. *Crop. Prot.* 30: 1117–1123 (2011)
- Wang CL, Shen SH, Zhang SY, Li QZ, Yao YB. Adaptation of potato production to climate change by optimizing sowing date in the Loess Plateau of central Gansu, China, *J. Integr. Agr.* 14 (2015) 398–409
- Williams PC, Sobering DC. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spec.* 1: 25–32 (1993)
- Williams P. Influence of water on prediction of composition and quality factors: the Aquaphotomics of low moisture agricultural materials. *J. Near Infrared Spec.* 17:315–328 (2009).
- Winkler-Moser JK, Singh M, Rennick KA, Bakota EL, Jham G, Liu SX, Vaughn SF. Detection of corn adulteration in Brazilian coffee (*Coffea arabica*) by tocopherol profiling and near-infrared (NIR) spectroscopy. *J. Agric. Food Chem.* 63: 10662–10668 (2015)
- Wójcicki K, Khmelinskii I, Sikorski M, Sikorska E. Near and mid infrared spectroscopy and multivariate data analysis in studies of oxidation of edible oils. *Food Chem.* 187: 416–423 (2015)
- Wold S, Sjostrom M, Eriksson L. PLS-regression: A basic tool of chemometrics. *Chemometr Intell Lab.* 58: 109–130 (2001)
- Xu L, Shi W, Cai CB, Zhong W, Tu K. Rapid and nondestructive detection of multiple adulterants in kudzu starch by near infrared (NIR) spectroscopy and chemometrics, *LWT-Food Sci. Technol.* 61: 590–595 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.