*IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)*
*e-ISSN: 2319-2380, p-ISSN: 2319-2372. Volume 14, Issue 8 Ser. I (August 2021), PP 39-47*
*www.iosrjournals.org*

# Starch Content Prediction of Potato Flour Noodles Used Hyperspectral Imaging Technology

Zhang Jing, Wang Sihua, Peng Huihui, Guo Zhen, Yue Minghui, Ren Zhishang, Ma Chengye*

*(School of Agricultural Engineering and Food Science, Shandong University of Technology, China*
*Corresponding author: Ma Chengye)*

***Abstract***
*The main goal of this study was used the hyperspectral imaging (HSI) technology in the range of 901-2251 nm was investigated to examine the starch content of potato flour noodles. Different spectral pre-treatment methods were used to improve the stability of the partial least squares regression (PLSR) and principal component regression (PCR) models. Compared the effects of other models, the result shows that PLSR is the best model. Based on the established SNV-CARS-PLSR model, the calibration performance of the dataset with $R^2c$ of 0.9490 and RMSEC of 0.1337%, and the validation performance of the dataset with $R^2p$ of 0.8946 and RMSEP of 0.1811%. The changes of starch contents of potato flour noodles during comprehensive comparison selected the optimal models were successfully visualized.*

---

Date of Submission: 03-08-2021 Date of Acceptance: 17-08-2021

---

## I. Introduction

Noodles have been industrialized in China and have a huge consumption in China's convenience food (Lambrecht, Rombouts, Nivelle, & Delcour, 2017). Potato flour noodles chemical compositions such as starch content have significant impacts on the quality attributes. Potato is the fourth major grain besides rice, wheat and corn. It will be processed into staple food such as steamed bread, noodles and rice flour, which will promote the rapid development of potato related industries(Pu et al., 2017). In 2015, our country officially launched the potato as a major grain strategic goals. The purpose is to process potatoes into Chinese traditional steamed bread, noodles and other staple food(Zhang, Xu, Wu, Hu, & Dai, 2017). In the potato staple food strategy, production is the basis, processing technology is the means, and product quality inspection is the key. The quality of potato has an important influence on the quality of processed products(Fernández-Ahumada, Garrido-Varo, Guerrero-Ginel, Wubbels, Sluis, & Meer, 2006).

At present, the potato products on the market have potato noodles, potato rice flour, potato mixed with rice, mashed potato porridge and potato mixed with a series of oat powder and so on. Among them, potato noodles have potato fresh noodles, potato half dry noodles, potato dry noodles. In this study we mainly used the potato whole flour and wheat flour mixed into the potato flour noodles. Considering the prominence of potato flour noodles constitution regarding economical or health related aspects, there is a need to analyse the chemical composition such as the starch content of potato flour noodles to ensure that the consumers and processors get the right quality of potato flour noodles from their suppliers(Jiang, Fang, Wang, & Fan, 2015). Detection technology potato product quality is also one of the important factors affecting the development of potato industry. While the traditional chemical detection methods of potato flour noodles are time-consuming and laborious, and some use toxic reagents, which is harmful to the health of detection personnel. However, the sensory evaluation of noodles mainly depends on the taste of people and is influenced by various factors such as experimentors, personal preferences, regions and so on(Barbin, ElMasry, Sun, & Allen, 2013). The subjective factors are large and cannot reflect the edible quality of noodles very well. Thus, the food industry needs a cost-effective, non-pollution and non- destructive technology to rapidly and efficiently measure the quality of potato flour noodles(Haoping Huang et al., 2021; C. Ma, Ren, Zhang, Du, Jin, & Yin, 2021).

In recent years, non-destructive testing technology has been widely used in the field of agricultural products, and has achieved good research results(Kang, Lee, & Son, 2008). Studies have shown that machine vision technology can detect the external quality of potato flour noodles, but can not detect its internal quality; NIR spectroscopy can detect the internal quality information of potato flour noodles, but can not detect the external quality; HSI technology, as a new, green non-destructive testing technology, takes into account the advantages of machine vision technology and NIR spectroscopy technology(Lu, Saeys, Kim, Peng, & Lu, 2020). It can detect the internal and external quality of potato flour noodles at the same time, has the characteristics of high resolution and non-destructive, etc. It has become a research hotspot in the field of potato flour noodles

non-destructive testing. However, there are still some problems in the application of HSI technology to detect the starch in potato flour noodles. Such as Hyperspectral data has a large number of bands and data, a lot of redundant information, and the calculation workload of data processing is difficult, which affects the modeling speed(Kamruzzaman, ElMasry, Sun, & Allen, 2012). And the processing method of hyperspectral data will affect the prediction accuracy of modeling. At the same time, the further popularization of HSI technology for non-destructive testing of the starch in potato flour noodles needs to improve the modeling speed and prediction accuracy.

## II. Material And Method

### 2.1 Sample preparation

Potato flour was obtained from Gansu Zhengyang Modern Agriculture Service Co., Ltd., wheat flour was purchased from Zhejiang Cereals, Oils and Foodstuffs Import and Export Co., Ltd.

Preparation of potato flour noodles: the potato whole flour and wheat flour were mixed in a random proportion to make noodle samples with the mass fraction of potato whole flour between 0 ~ 35%. The potato flour noodles were packed in to air tight containers for further tests.

### 2.2 Compositional analysis

Iodine test represents a practical reference method for evaluating the index of starch content of potato flour noodles. Its principle is that after the sample was removed fat and soluble sugar, that the starch is hydrolyzed to monosaccharide and will turn blue-black in contact with iodine solution(Menesatti, Zanella, D'Andrea, Costa, Paglia, & Pallottino, 2009). In order to improve the accuracy of the measurement, each sample was measured three times, and the average value was taken as the chemical value of the starch content.

### 2.3 Hyperspectral image acquisition and pre-processing

#### 2.3.1 Hyperspectral imaging system

The hyperspectral system used in this paper includes a camera obscura，mobile platform，a line-scan spectrograph ( Isuzu Optics Corp., Taiwan, China, with a wavelength range of 900-2500 nm), a CCD camera with high resolution, two halogen lamps, a conveyor belt and a computer (contain data acquisition software and an image processor).

#### 2.3.2 Image acquisition and calibration

In this study, we used the wavelength range of 900-2500 nm HSI system to collect the hyperspectral images of the potato flour noodles, and then the collected hyperspectral images are corrected in black and white.

In order to reduce the influence of camera dark current. The white reference image was obtained by using a uniform white ceramic sheet with 99% reflectance, while the dark reference image with 0% reflectance was collected by turning off the light source and tightly covering the objective lens with its opaque cap. Then, the relative reflectance ($R_T$) of the pixel of the raw spectral image of the sample is calculated using the two reference images, and the calculation formula is as follows (J. Ma, Cheng, Sun, & Liu, 2019).

$$R_T = \frac{I_{\text{raw}} - I_{\text{dark}}}{I_{\text{white}} - I_{\text{dark}}}$$

Where: $R_T$: corrected image;

$I_{raw}$: Raw hyperspectral image;

$I_{dark}$: blackboard calibration image;

$I_{white}$: white board calibration image.

This equation converts the reflectance values of all pixels in the raw hyperspectral image with absolute reflectance values (arbitrary reflectance values) into relative reflectance values (unitless) with the help of standard reference data.

#### 2.3.3 ROI selection

In order to automatically obtain the spectrum of the samples, we recognize the image at the highest reflectance of each sample, and then segment it according to a fixed threshold, because the reflectance value of the background is close to 0 at each wavelength. In this way, the region of interest (ROI) composed of pixels of the potato flour noodles samples can be easily obtained. Finally, the spectrum of the ROI was extracted and the average spectrum of ROI was calculated.

### 2.4 Multivariate analysis

#### 2.4.1 Sample set partition method

In this study, the calibration set and the prediction set models were strictly built using the SPXY algorithm. SPXY algorithm was first proposed by Galvao (Galvão, Araujo, José, Pontes, Silva, & Saldanha,

2005). The principle is to calculate the distance between samples with spectral value and sugar value as characteristic parameters, so as to ensure the difference and representativeness between samples. Effectively cover the multi-dimensional vector space to avoid over fitting or poor prediction effect of the prediction model caused by too small or the same difference between samples, so as to improve the stability and accuracy of the model(Xu, Riccioli, & Sun, 2016).

**2.4.2 Outliers elimination method**

During the calibration, outliers were identified and elimination. In the research, even a single outlier will have a greater and more harmful impact on the model than the normal samples，so that leading to misleading results(Kamruzzaman et al., 2012). In this paper, the Monte Carlo cross-validation method is used to eliminate outliers，this method not only checks the physical and chemical reference values of outliers, but also tests the advantages of outliers and spectral outliers(Wu, Shi, Wang, He, Bao, & Liu, 2012).

**2.4.3 Pre-processing algorithms**

In order to better evaluate the ability of hyperspectral imaging to predict the starch content of potato flour noodles, the absorption spectrum was also considered. In addition, spectral pre-processing of HSI is also needed to reduce the negative effects of HSI in the acquisition process, including random noise, optical path length change and useless light scattering(Wu & Sun, 2013). In this study, Smoothing, Baseline, Standard normal variate (SNV), De-trending, Orthogonal signal correction (OSC), Moving average, Savitzky-Golay were carried out.

**2.4.4 Feature-related wavelengths selection**

The acquired HSI of each potato flour noodles samples were characterised as high dimensionality with multicollinearity amongst straight wavelength bands. Selecting optimal wavelengths from the entire wavelength could simplify the calculations and improve the prediction efficiency of protein content detection (Khulal, Zhao, Hu, & Chen, 2016). In this study, we mainly used four measures to select the feature-related wavelengths, including Regression Coefficient (RC), Successive Projections Algorithm (SPA), Competitive adaptive reweighting algorithm (CARS), Elimination of uninformative variables (UVE). In this study, a full wavelength model was established, then the simplified models were established using the optimal wavelengths selected by these three methods. Finally, the optimal combine would be selected by comparing the model performance.

**2.4.5 Model development and performance evaluation**

Model validation is an important step in multivariate data analysis. Chemometric methods commonly used in spectral quantitative analysis mainly includes Partial least squares regression (PLSR), Principal component regression (PCA), Multiple linear regression (MLR), Support vector machine regression (SVMR). As the multivariate data analysis methods, the PLSR and PCA models are widely used to predict the internal components of food and agricultural products (Feng, Zhu, Liu, He, & Zhang, 2019; Haibo Huang, Yu, Xu, & Ying, 2008; Nicolaï et al., 2007). The performance of model was evaluated by the correlation coefficient of calibration ($R^2_c$) and prediction ($R^2_p$), root mean square error of calibration (RMSEC) and prediction (RMSEP). Generally, the model that has better predictive power and accuracy corresponds to higher $R^2_c$ , $R^2_p$ and lower RMSEC, RMSEP.

**2.5 Visualization of starch distribution**

Recognizing the starch distribution is significant useful to understand the potato flour noodles. Chemical imaging is a practical method to visualize this phenomenon. In the current study, the best optimized models were used to visualize and map each pixel of the hyperspectral images into the chemical images for predicting starch distribution of the examined potato flour noodles. Component variation was presented in colours, where the colours rank the composition according to a colour bar displayed alongside the map (Su & Sun, 2018). In the current method, by checking the color variation in the developed map, we can easily assess the distribution of starch contents within the potato flour noodles, which can be convinent and useful for the industry and humanity to choose suitable products.

## III. Results And Discussion

**3.1 Overview of spectral profiles**

The average reflectance spectra extracted from 120 potato flour noodles with different strach contents in Fig.1. The spectral image (384 × 288 pixels) obtained from each scanned sample were combined into the spectral image of the sample, the spectral image of the edge of the sample unit and the spectral image of the background. Therefore, the spectral data of the background and the edge of the sample cell and the noise must be removed before analysis. In this paper, selects the band with hyperspectral wavelength range of 901 ~ 2251

nm and spectral reflectance range of 0.1 ~ 0.7 for follow-up work.

For the potato flour noodles, there were main peaks at 963 nm, 1221 nm, 1296 nm and 1470 nm that is corresponding to the peaks of starch. The spectra of the potato flour noodles showed two main peaks of starch and water at around 1220 nm and 1470 nm，which were related to the second overtone of C-H stretch at 1220 nm and the first overtone of O-H stretch at 1470 nm, respectively (Ozaki, Mcclure, & Christy, 2006; Zhao, Wang, Ni, Chu, Li, & Sun, 2018).

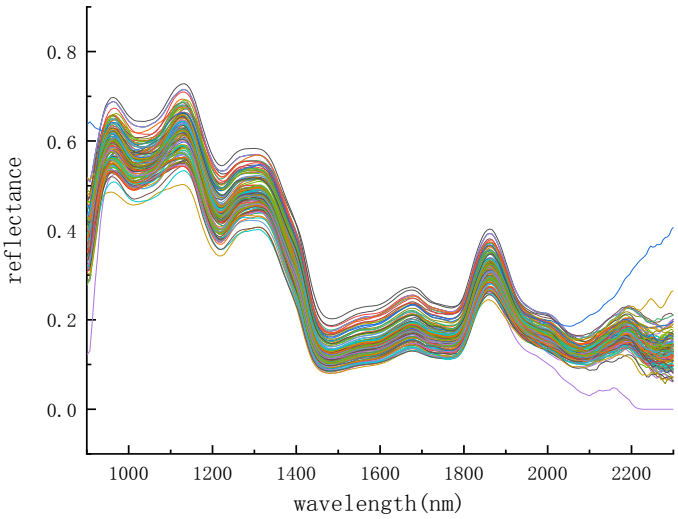


Fig. 1 Average reflectance spectra extracted from 120 potato flour noodles

**3.2 Excluding outliers**

Results of Monte Carlo abnormal samples detection of starch content of potato flour noodles are shown in Fig. 2(Du, Kasemsumran, Maruo, Nakagawa, & Ozaki, 2006). According to the mean value of prediction error and the standard deviation of prediction error, a total of 7 samples exceeded the threshold range, and these samples were considered as abnormal samples. These samples are screened out one by one, and PLSR models are constructed respectively. The models without abnormal samples are compared with the models with abnormal samples. If the model performance improves, the samples are considered as abnormal samples and shall be removed. The results are shown in Table 1.

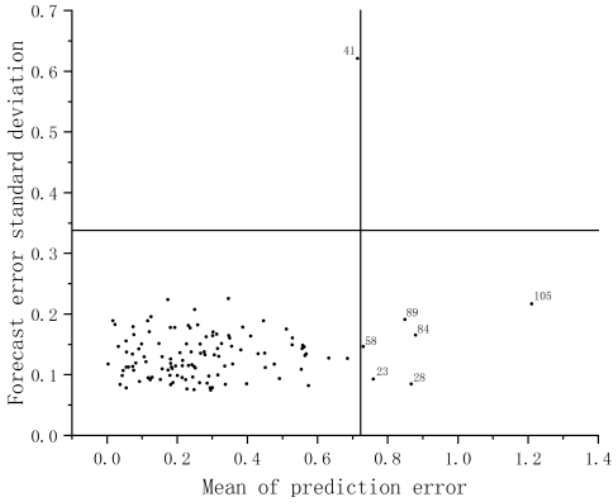


Fig 2. Monte Carlo eliminates starch outliers

As it was shownin Table 1, after elimination samples No. 105, 41, 84, and 28, the $R^2_C$ of the model is improved, while after elimination samples No. 89, 23 and 58, the $R^2_C$ of the model decreases. Therefore, No. 105, 41, 84, and 28 were determined as abnormal samples, and the remaining 116 samples after the elimination were carried ou for follow-up work.

**Table 1** PLSR model of noodle starch content before and after the elimination of abnormal samples

| Eliminate sample | Principal component number | RMSECV | $R^2_C$ | RMSEC |
|---|---|---|---|---|
| Without | 13 | 0.2346 | 0.9360 | 0.1496 |
| 105 | 13 | 0.2336 | 0.9398 | 0.1453 |
| 41 | 13 | 0.2394 | 0.9375 | 0.1484 |
| 84 | 14 | 0.2353 | 0.9406 | 0.1441 |
| 89 | 13 | 0.2317 | 0.9346 | 0.1499 |
| 28 | 13 | 0.2114 | 0.9509 | 0.1315 |
| 23 | 13 | 0.2359 | 0.9353 | 0.1499 |
| 58 | 13 | 0.2320 | 0.9346 | 0.1503 |

### 3.3 Sample set partition

According to the ratio of 2:1, the samples after removing outliers were divided into prediction set and calibration set according to SPXY method. The calibration set contains 78 samples and the prediction set contains 38 samples. Both the maximum and minimum values of starch content are in the calibration set (Table 2). This indicates that the starch content has a relatively wide range in the calibration set, which indicates that the classification method is scientific and reasonable.

**Table 2** Statistical data of noodle starch content in sample set divided by SPXY method

| sample | sample number | starch content（%） | | | |
|---|---|---|---|---|---|
| | | minimum | maximum | mean | Standard deviation |
| calibration | 78 | 61.9295 | 64.0400 | 63.0289 | 0.5964 |
| prediction | 38 | 61.9605 | 64.0145 | 63.0319 | 0.5654 |

### 3.4 Spectral pretreatment

The performance of the PLSR and PCR models established from the original spectral data and the spectral data preprocessed by different methods were shown in Table 3 and 4. From the results we know that the PLSR and PCR models for starch content of potato flour noodles with RMSE lower than 1%, the calibration datasets $R^2_C$ is 0.8863 and 0.6484 , respectively.

In the PLSR models that the result is not ideal after Normalize, Baseline, Detrend and OSC pretreatment, the effect of regression model, and the $R^2_C$ regression model is decreased compared with the original spectral data. The $R^2_C$ of the regression model pretreated by Moving Average and SNV has been improved to a certain extent. The RMSECV of Moving average is less than that of the original spectral model, while the RMSECV of SNV pretreated is increased compared with that of the original spectral model, but the difference is not significant. Among them, the regression model $R^2_C$ after SNV is 0.9001, which is the maximum value, and RMSEC is also the minimum value is 0.1873. In comparison, SNV was selected as the pretreatment method of PLSR model for the starch content of potato flour noodles. The result for PLSR model is shown in Table 3. The PCR models reported in Table 5 only the performance of the regression model established after Detrend pretreatment decreases, and $R^2_C$ (0.6223) is lower than that of the regression model established with original spectral data(Galvão et al., 2005). The $R^2_C$ of the regression model established after OSC pretreatment is still 0.6484. While the $R^2_C$ of the regression model established after the pretreatment of Moving Average, Normalize, Baseline and SNV has been improved. Normalize pretreatment is the best pretreatment method for PCR model, the $R^2_C$ is 0.8288 and the RMSEC is 0.3521. Although the pretreatment method greatly improved the performance of the PCR model, its effect was still not as good as that of the PLSR model, so only the PLSR model combined with SNV for follow-up work of starch content of potato flour noodles.

**Table 3** Starch content PLSR model of original spectral data and different pretreated spectral data

| Pretreatment method | Principal component number | RMSECV | PLSR calibration model | |
|---|---|---|---|---|
| | | | $R^2_C$ | RMSEC |
| Raw | 9 | 0.3738 | 0.8863 | 0.1998 |
| Moving average | 10 | 0.3608 | 0.8904 | 0.2962 |
| Normalize | 9 | 0.3591 | 0.8826 | 0.2030 |
| Baseline | 10 | 0.3811 | 0.8790 | 0.2061 |
| SNV | 10 | 0.3802 | 0.9001 | 0.1873 |
| Detrend | 8 | 0.3571 | 0.8699 | 0.2137 |
| OSC | 8 | 0.3779 | 0.8862 | 0.1999 |

**Table 4** PCR models for starch content of original spectral data and different pretreated spectral data

| Pretreatment method | component number | RMSECV | PCR calibration model | |
|---|---|---|---|---|
| | | | $R^2_C$ | RMSEC |
| Raw | 9 | 0.4026 | 0.6484 | 0.3514 |
| Moving average | 10 | 0.4050 | 0.6536 | 0.3487 |
| Normalize | 27 | 0.3521 | 0.8288 | 0.2451 |

| | | | | |
|---|---|---|---|---|
| Baseline | 8 | 0.3974 | 0.6609 | 0.3450 |
| SNV | 27 | 0.3762 | 0.8151 | 0.2548 |
| Detrend | 7 | 0.4014 | 0.6223 | 0.3641 |
| OSC | 8 | 0.3949 | 0.6484 | 0.3513 |

### 3.5 Extracted characteristic wavelength

#### 3.5.1 Selecting characteristic wavelengths by CARS algorithm

The CARS algorithm was used to select the characteristic wavelengths for starch content of potato flour noodles (Fig. 3)(Su et al., 2018). The variable screening process, with the change of sampling times, the total wavelength decreases gradually, then the decreasing speed gradually slowed down, that indicating the selection process of wavelength variables in Fig. 3 (a). At this time, the remaining wavelengths were correlated with the starch content of potato flour noodles , and the elimination of any wavelength might lead to the decline of the model effect, so the judgment should be made according to the change of RMSECV. Fig. 3 (b) shows the variation trend of RMSECV cross-verified in the process of variable screening. The slow decline indicates that at first the wavelengths related to the starch content of potato flour noodles are all in the model, and the irrelevant wavelengths are gradually screened out, and the model becomes more and more stable. When the sampling times reach 59 and 88 times, RMSECV obtains the minimum value. However, when the sampling times are 59 times, there are many wavelength variables. As the sampling continues, RMSECV slowly increases and then decreases until the minimum value of 0.1456 is obtained when the sampling times are 88 times. This indicates that when the sampling times are greater than 88, the wavelength related to the starch content of potato flour noodles is removed, resulting in a sharp decline in the model performance every time a relevant wavelength is removed, so the sampling stops when the sampling times are 88. Fig. 3 (c) shows the trend of wavelength variable regression coefficient in the process of sample collection and screening. At asterisk (*) position, the sampling times are 88 times, and the RMSECV reaches the minimum. Therefore, at 88 times of sampling, 33 characteristic wavelengths（1007、1013、1063、1070、1227、1296、1340、1415、1538、1544、1550、1557、1636、1660、1672、1696、1702、1732、1744、1785、1820、1913、1958、1992、2008、2014、2095、2195、2210、2215、2231、2236、2241nm）were selected, accounting for 14.4% of the total wavelength.

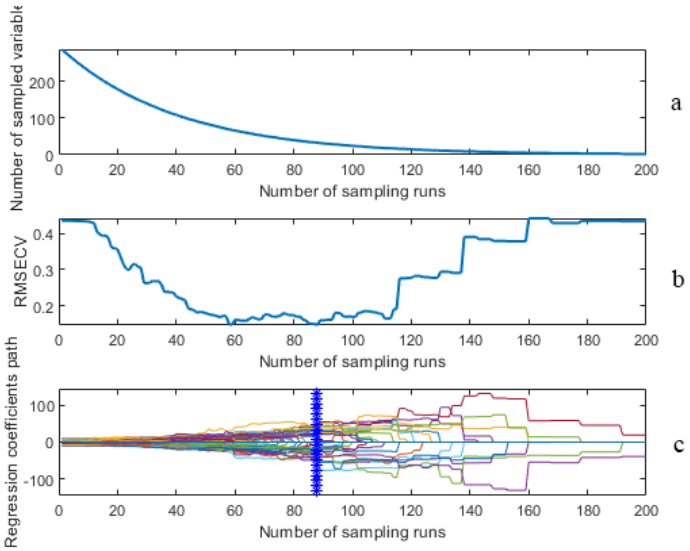


**Fig. 3** The process of screening the characteristic wavelength of noodle starch content by CARS

#### 3.5.2 SPA algorithmis used to select characteristic wavelengths

During the SPA algorithmis selecting characteristic wavelengths of starch content of potato flour noodles, it was found that the number of extracted characteristic wavelengths was too small to represent the information related to starch. However, SPA algorithmis was used to selecting the characteristic wavelength of the original spectrum, and it was found that a part of the characteristic wavelength was concentrated in the wave band after the wavelength of 2251nm(Yu, Zhao, Liu, Li, Liu, & He, 2014). Therefore, the wave band after 2251 nm was not removed when using SPA algorithmis to screen the characteristic wavelength of starch. In Fig. 4 (a), it can be seen that when the number of wavelengths is 22, RMSECV (0.31775) is the minimum value. Fig. 4 (b) shows the distribution of the selected 22 wavelengths, the abscise is the number of the wavelengths, there are 288 bands in the wavelength range of 900-2500nm. According to the number the spectral data under the corresponding wavelength are extracted from the full spectral data. The corresponding wavelengths is

respectively 901, 994, 1063, 1133, 1340, 1408, 1605, 1820, 1896, 1935, 2019, 2122, 2210, 2246, 2255, 2295, 2455, 2472, 2480, 2489, 2497, 2505nm, respectively. It accounts for 7.6% of the total band.

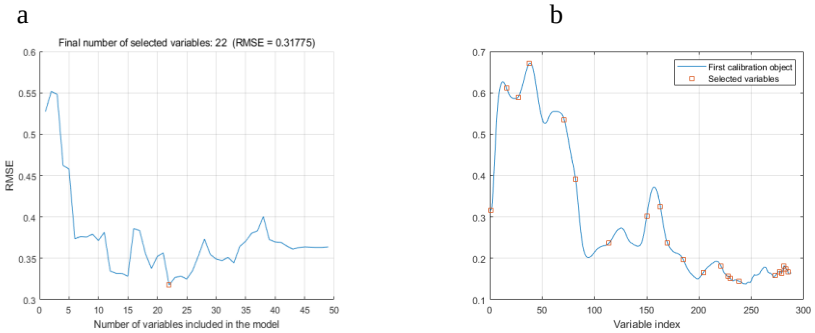


**Fig. 4** Process of SPA algorithm in screening characteristic wavelength of noodle starch content (a: RMSECV of variable, b: screened characteristic wavelength)

### 3.5.3 UVE method is used to select characteristic wavelengths

UVE algorithm was used to select characteristic wavelengths for starch content of potato flour noodles (Fig.5). In the figure, the left side of the abscissa is the wavelength variable, the right side is the noise variable, and the ordinate is the stability coefficient of the variable. It can be seen that the absolute value of the stability coefficient of the wavelength variable is significantly higher than that of the noise variable. Eliminating the wavelength variable whose stability is less than the noise in the full wavelength variable can remove most of the influence of the noise variable. Finally, 19 characteristic wavelengths (969, 976, 1007, 1013, 1076, 1082, 1221, 1227, 1302, 1340, 1377, 1587, 1648, 1654, 1660, 1666, 1696, 1702, 2215 nm) were obtained, accounting for 8.3% of the whole band.

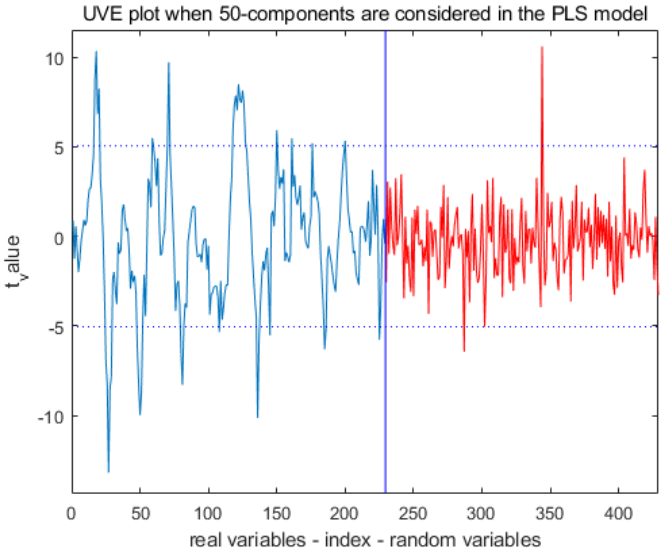


Fig. 5 UVE algorithm selection process of characteristic wavelength

The regression model of the characteristic wavelengths extracted by the three algorithms was compared with the full-wavelength model is shown in Table 5. The number of characteristic wavelengths extracted by CARS algorithm is the highest that is 33, but the number of principal components is the same as that of full-wavelength model, which is 10. However, SPA and UVE algorithms extract less characteristic wavelengths is 22 and 19, respectively, but the number of principal components is 14 and 16, respectively, which is much higher than that of full-wavelength model. From the $R^2_c$ (0.9490) of the calibration set, the characteristic wavelength models of CARS algorithm is greatly improved compared with the full-wavelength model, while SPA and UVE algorithms are not ideal, and the RMSEC (0.1337) of the CARS calibration set model is the smallest, which indicates the stability of the model. From the $R^2_p$ of the prediction set, the full-wavelength model is the largest and greater than the $R^2_c$ of the calibration set, that the model may be over fitted. Comparing the three characteristic wavelengths models, that CARS model $R^2_p$ is 0.8946 and RMSEP is 0.1811. Therefore, CARS algorithm is the best method to extract the characteristic wavelengths of starch content of potato flour

noodles.

Table 5 Comparison between the selected results of hyperspectral characteristic wavelengths of samples by different methods and the model

| Model | Pretreatment method | Extraction method | Wavelength s number | Principal component number | Calibration | | Prediction | |
|---|---|---|---|---|---|---|---|---|
| | | | | | $R^2_C$ | RMSEC | $R^2_P$ | RMSEP |
| PLSR | SNV | FS | 229 | 10 | 0.9001 | 0.3802 | 0.9335 | 0.1444 |
| | | CARS | 33 | 10 | 0.9490 | 0.1337 | 0.8946 | 0.1811 |
| | | SPA | 22 | 14 | 0.8760 | 0.2086 | 0.8833 | 0.1906 |
| | | UVE | 19 | 16 | 0.8531 | 0.2270 | 0.8912 | 0.1840 |

### 3.6 Visualization of protein content distribution of potato flour noodles

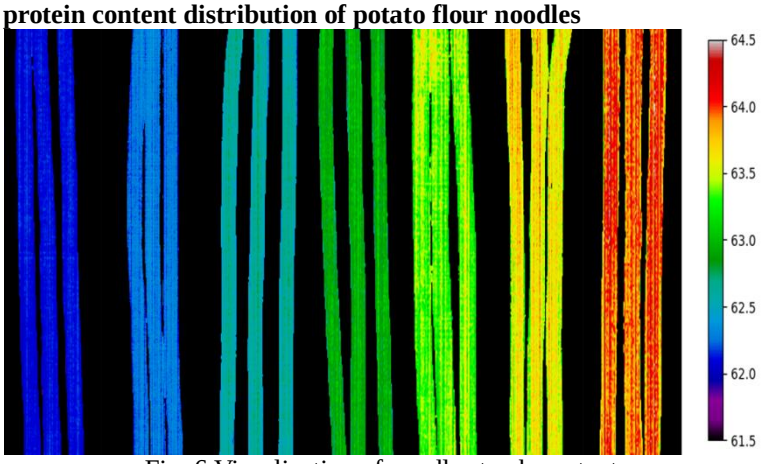


Fig. 6 Visualization of noodle starch content

HSI technology not only predict the starch content of potato flour noodles, but also realize the visualization of starch content. The SNV-CARS-PLSR prediction model was established to predict the starch content in each pixel by using the spectral data of the corresponding pixel of the HSI of potato flour noodles, and the visual distribution map was drawn in the form of false color map. As shown in Fig. 6, the distribution of starch content of potato flour noodles can be visually seen according to the color bar on the right of the image. Seven potato flour noodle samples with large difference in starch content were selected. The color difference between different potato flour noodles indicated that the starch content was different, and the difference of starch content in different positions of the same noodle could also be reflected. In Fig. 6, the starch content of potato flour noodles from left to right was 61.929%, 62.005%, 62.508%, 62.828%, 63.239%, 63.514% and 64.040%, respectively.

## IV. CONCLUSION

The results of this research demonstrated that the HSI technique operated in the NIR region (900–2500 nm) can accurately detect the starch contents of potato flour noodles. PLSR and PCR models of starch content were established, and the effects of Moving Average, Normalize, Baseline, SNV, Detrend and OSC pretreatment were compared with the original spectral models. SNV was the best pretreatment method for starch PLSR model, and Normalize was the best pretreatment method for starch PCR model, the PLSR model was better than the PCR model. The characteristic wavelength model of starch content was established, and CARS algorithm was the best method in the three models. SNV-CARS-PLSR model was the optimal model for starch, 33 characteristic wavelengths were extracted from 229 wavelengths, accounting for 14.4% of the total wavelength. And the calibration model $R^2_c$ and RMSEC were 0.9490 and 0.1337, respectively, the prediction model $R^2_p$ and RMSEP were 0.8946 and 0.1811, respectively. By substituting spectral data of pixels corresponding to potato flour noodles HSI into the established prediction model, the starch content of each pixel was predicted, and the distribution of starch content of potato flour noodles was visualized. The overall results obtained in this research showed that it was possible to employ the NIR hyperspectral imaging technique for rapidly, precisely and non-destructively predicting starch content of potato flour noodles, and are reliable enough for potential quality control programs. The results suggest that NIR spectral imaging could become a useful tool for quantifying the composition and evaluating the distribution of important components such as protein, water and fat contents.

## REFERENCES

[1]. Barbin, D. F., ElMasry, G., Sun, D.-W., & Allen, P. (2013). Non-destructive determination of chemical composition in intact and minced pork using near-infrared hyperspectral imaging. Food Chemistry, 138(2), 1162-1171. https://doi.org/https://doi.org/10.1016/j.foodchem.2012.11.120.

[2]. Du, Y. P., Kasemsumran, S., Maruo, K., Nakagawa, T., & Ozaki, Y. (2006). Ascertainment of the number of samples in the

validation set in Monte Carlo cross validation and the selection of model dimension with Monte Carlo cross validation. Chemometrics and Intelligent Laboratory Systems, 82(1), 83-89. https://doi.org/https://doi.org/10.1016/j.chemolab.2005.07.004.
[3]. Feng, L., Zhu, S., Liu, F., He, Y., & Zhang, C. (2019). Hyperspectral imaging for seed quality and safety inspection: a review. Plant Methods, 15(1), 91.
[4]. Fernández-Ahumada, E., Garrido-Varo, A., Guerrero-Ginel, J., Wubbels, A., Sluis, C. V. d., & Meer, J. V. d. (2006). Understanding factors affecting near infrared analysis of potato constituents. Journal of Near Infrared Spectroscopy, 14(1).
[5]. Galvão, R. K. H., Araujo, M. C. U., José, G. E., Pontes, M. J. C., Silva, E. C., & Saldanha, T. C. B. (2005). A method for calibration and validation subset partitioning. Talanta, 67(4), 736-740. https://doi.org/https://doi.org/10.1016/j.talanta.2005.03.025.
[6]. Huang, H., Hu, X., Tian, J., Jiang, X., Sun, T., Luo, H., & Huang, D. (2021). Rapid and nondestructive prediction of amylose and amylopectin contents in sorghum based on hyperspectral imaging. Food Chemistry, 359, 129954. https://doi.org/https://doi.org/10.1016/j.foodchem.2021.129954.
[7]. Huang, H., Yu, H., Xu, H., & Ying, Y. (2008). Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. Journal of Food Engineering, 87(3), 303-313. https://doi.org/https://doi.org/10.1016/j.jfoodeng.2007.12.022.
[8]. Jiang, W., Fang, J., Wang, S., & Fan, Y. (2015). Detection of starch content in potato based on hyperspectral imaging technique. International Journal of Signal Processing, Image Processing and Pattern Recognition, 8(12), 49-58. https://doi.org/10.14257/ijsip.2015.8.12.06.
[9]. Kamruzzaman, M., ElMasry, G., Sun, D.-W., & Allen, P. (2012). Non-destructive prediction and visualization of chemical composition in lamb meat using NIR hyperspectral imaging and multivariate regression. Innovative Food Science & Emerging Technologies, 16, 218-226. https://doi.org/https://doi.org/10.1016/j.ifset.2012.06.003.
[10]. Kang, S., Lee, K.-J., & Son, J.-R. (2008). On-line internal quality evaluation system for the processing potatoes. American Society of Agricultural and Biological Engineers - Food Processing Automation Conference 2008, June 28, 2008 - June 29, 2008 (pp. 68-73). Providence, RI, United states: American Society of Agricultural and Biological Engineers.
[11]. Khulal, U., Zhao, J., Hu, W., & Chen, Q. (2016). Nondestructive quantifying total volatile basic nitrogen (TVB-N) content in chicken using hyperspectral imaging (HSI) technique combined with different data dimension reduction algorithms. Food Chemistry, 197, 1191-1199. https://doi.org/https://doi.org/10.1016/j.foodchem.2015.11.084.
[12]. Lambrecht, M. A., Rombouts, I., Nivelle, M. A., & Delcour, J. A. (2017). The impact of protein characteristics on the protein network in and properties of fresh and cooked wheat-based noodles. Journal of Cereal Science, 75, 234-242. https://doi.org/https://doi.org/10.1016/j.jcs.2017.04.014.
[13]. Lu, Y., Saeys, W., Kim, M., Peng, Y., & Lu, R. (2020). Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. Postharvest Biology and Technology, 170.
[14]. Ma, C., Ren, Z., Zhang, Z., Du, J., Jin, C., & Yin, X. (2021). Development of simplified models for nondestructive testing of rice (with husk) protein content using hyperspectral imaging technology. Vibrational Spectroscopy, 114, 103230. https://doi.org/https://doi.org/10.1016/j.vibspec.2021.103230.
[15]. Ma, J., Cheng, J.-H., Sun, D.-W., & Liu, D. (2019). Mapping changes in sarcoplasmatic and myofibrillar proteins in boiled pork using hyperspectral imaging with spectral processing methods. LWT, 110, 338-345. https://doi.org/https://doi.org/10.1016/j.lwt.2019.04.095.
[16]. Menesatti, P., Zanella, A., D'Andrea, S., Costa, C., Paglia, G., & Pallottino, F. (2009). Supervised multivariate analysis of hyper-spectral NIR images to evaluate the starch index of apples. Food and Bioprocess Technology, 2(3), 308-314. https://doi.org/10.1007/s11947-008-0120-8.
[17]. Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, 46(2), 99-118. https://doi.org/https://doi.org/10.1016/j.postharvbio.2007.06.024.
[18]. Ozaki, Y., Mcclure, W. F., & Christy, A. A. (2006). Near-Infrared Spectroscopy in Food Science and Technology (Ozaki/Near-Infrared Spectroscopy in Food Science and Technology) || Other Topics. 10.1002/0470047704, 341-399.
[19]. Pu, H., Wei, J., Wang, L., Huang, J., Chen, X., Luo, C., . . . Zhang, H. (2017). Effects of potato/wheat flours ratio on mixing properties of dough and quality of noodles. Journal of Cereal Science, 76, 236-242. https://doi.org/https://doi.org/10.1016/j.jcs.2017.06.020.
[20]. Su, W.-H., & Sun, D.-W. (2018). Multispectral Imaging for Plant Food Quality Analysis and Visualization. Comprehensive Reviews in Food Science and Food Safety, 17(1), 220-239. https://doi.org/https://doi.org/10.1111/1541-4337.12317.
[21]. Wu, D., Shi, H., Wang, S., He, Y., Bao, Y., & Liu, K. (2012). Rapid prediction of moisture content of dehydrated prawns using online hyperspectral imaging system. Analytica Chimica Acta, 726, 57-66. https://doi.org/https://doi.org/10.1016/j.aca.2012.03.038.
[22]. Wu, D., & Sun, D.-W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part I: Fundamentals. Innovative Food Science & Emerging Technologies, 19, 1-14. https://doi.org/https://doi.org/10.1016/j.ifset.2013.04.014.
[23]. Xu, J.-L., Riccioli, C., & Sun, D.-W. (2016). Development of an alternative technique for rapid and accurate determination of fish caloric density based on hyperspectral imaging. Journal of Food Engineering, 190, 185-194. https://doi.org/https://doi.org/10.1016/j.jfoodeng.2016.06.007.
[24]. Yu, K.-Q., Zhao, Y.-R., Liu, Z.-Y., Li, X.-L., Liu, F., & He, Y. (2014). Application of Visible and Near-Infrared Hyperspectral Imaging for Detection of Defective Features in Loquat. Food and Bioprocess Technology, 7(11).
[25]. Zhang, H., Xu, F., Wu, Y., Hu, H.-h., & Dai, X.-f. (2017). Progress of potato staple food research and industry development in China. Journal of Integrative Agriculture, 16(12), 2924-2932. https://doi.org/https://doi.org/10.1016/S2095-3119(17)61736-2.
[26]. Zhao, X., Wang, W., Ni, X., Chu, X., Li, Y. F., & Sun, C. (2018). Evaluation of Near-Infrared Hyperspectral Imaging for Detection of Peanut and Walnut Powders in Whole Wheat Flour. Applied Sciences, 8(7).

Zhang Jing, Wang Sihua, et. al. "Starch Content Prediction of Potato Flour Noodles Used Hyperspectral Imaging Technology." *IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)*, 14(8), 2021, pp. 39-47.