Research of Determination Method of Starch and Protein Content in Buckwheat by Mid-Infrared Spectroscopy

Fenghua Wang, Ju Yang, Hailong Zhu, Zhiyong Xi

Research of Determination Method of Starch and Protein Content in Buckwheat by Mid-infrared Spectroscopy

Fenghua Wang[1,1], Ju Yang[1], Hailong Zhu[2], Zhiyong Xi[1]

[1] Faculty of Modern Agricultural Engineering,Kunming University of Science and Technology, Kunming, Yunnan, China 650500,
wangfenghua018@163.com
[2] Engineering Training Center, Kunming University of Science and Technology, Kunming, Yunnan, China, 650500

Abstract. Starch and proteins are the vital nutrients in buckwheat, to achieve the fast detection of the buckwheat internal composition has an important theoretical significance and application value for buckwheat breeding, processing and other steps. In the paper, forty buckwheat samples from different origins have been selected. The starch and protein content of buckwheat was determined, and the mid-infrared transmission spectrum of buckwheat has been obtained using Fourier mid-infrared spectroscopy. Forty samples were randomly divided into the prediction set and validation set with 30 and 10 samples respectively. After smoothing preprocessing, the prediction models of buckwheat starch and protein content have been established using the combination method of principal component analysis and artificial neural network, finally the models have been verified. The results showed that the correlation coefficient between the prediction value and measurement value of buckwheat starch content is 0.9029, and the relative error is smaller and its mean value is 2.33%, the method of the buckwheat starch content prediction is feasible. But the prediction for buckwheat protein content is not ideal, need to be further studied.

Keywords: Mid-infrared spectroscopy, Buckwheat, Starch content, Protein content, PCA, BP Neural Network

1 Introduction

Buckwheat, also known as the triangle wheat, black wheat, belongs to the genus of dicotyledonous Polygonaceae, family of Fagopyrum Mill, mainly have two cultivar of buckwheat is Sweet Buckwheat (Fagopyrum esculentum Moench) and Tartary Buckwheat (Fagopyrum tararicum Gaerth). The buckwheat mainly located in the

[1] This study was financed by the Application Basis Surface Projects of Yunnan Province (2010ZC028), Kunming University of Science and Technology Analysis and Testing Fund (2010299)

alpine areas of the Loess plateau and the plateau mountains of Yunnan, Guizhou and Sichuan. It has a high nutritional and health value, known as the food treasures of "Medicine-Food"[1]. Therefore, it is necessary to develop the cultivation industry and further processing technology of buckwheat, to improve the quality of buckwheat has a great significance for increasing farmers' income.

Now, buckwheat is used to buckwheat flour production for making buckwheat food. The level of starch and protein content of buckwheat is an important index for the buckwheat nutritional quality evaluation [2]. Buckwheat starch and protein content measurements is still using the traditional chemical analysis methods, it has the advantage of high measurement accuracy, but the disadvantages are as follows, the detection is slow, the measurement cost is expensive, the sample is destroyed, and it should be detected in the specialize testing organization or laboratory.

With the characteristics of high resolution capability, fast scan time, large radiation flux, low stray radiation and wide spectral range, fast and accurate detection of sample components can be achieved through Fourier Mid-infrared spectroscopy combined with chemometrics. Spectroscopy has been widely used in quality inspection of agricultural products, and the good analysis results have been achieved [3-7]. However, spectroscopy based studies on buckwheat nutrients in domestic and foreign have not been reported. Therefore, eight kinds of buckwheat from different growing areas have been collected. The spectral information of buckwheat have been obtained using spectroscopic techniques, the quantitative relationship model between the spectral information and the buckwheat internal quality have been researched, and a rapid, accurate, stable, easy assessment for buckwheat quality has been established. It has a good theoretical significance and practical value in buckwheat numerous aspects such as breeding, processing and testing.

2 Experimental instrument and methods

2.1 Instruments and equipments

Spectrometer: Germany Bruker EQUINOX 55 Fourier infrared spectroscopy, CS101-2D type electric blast drying oven (Chongqing star joint venture enterprise experimental instrument Co., Ltd.), JLML rice huller machine (Shanghai Jiading Grain and Oil Detection Instrument Factory).

2.2 Experimental material

Seven kinds of tartary buckwheat and a kinds of sweet buckwheat from Yunnan, Shanxi, Sichuan province in China were selected. They are Kunming local buckwheat, Luxi tartary buckwheat, Zhaotong tartary buckwheat No. 1 and No.2, Shanxi tartary buckwheat, Xichang tartary buckwheat, Dali tartary buckwheat and Dali sweet buckwheat, a total of eight kinds of representative buckwheat varieties, each buckwheat selected five samples, a total of 40 samples. Each sample was air-dried,

eliminating Impurities, shelling, crushing over a sieve for the requirement of particle size and consistent uniformity, the diameter of the sieve is 0.2mm. All samples were bagging sealing for spare use.

2.3 Experimental methods and steps

Acquisition of the buckwheat spectral features. The transmittance of the buckwheat samples have been obtained Using Germany Brook EQUINOX55 Fourier Mid-infrared spectroscopy, set the spectral resolution of $2cm^{-1}$, the testing range of $4000 \sim 400cm^{-1}$.

At first, 1g buckwheat sample and 1.5g potassium bromide have been mixed in the mortar, and then the mixing powder has been placed in a circular hole to tabletting. Then samples scanning have been carried on, background measurements is required before each scanning. Each sample repeat scanning for three times. The whole operation should be in dark conditions as possible to avoid potassium bromide decomposition.

Buckwheat starch and protein content measurement. Buckwheat starch content has been measured by enzyme hydrolysis method, and protein content has been measured by the Kjeldahl nitrogen method. The chemical measurement has been accomplished by Analysis and Testing Center of Yunnan Province. Table 1 is the Statistical parameters of the chemical measurement results.

Table1. Buckwheat chemical constituent's statistical parameters

	Sample number	Minimum	Maximum	Mean	Standard Deviation
Starch content (%)	40	48.97	69.49	62.22	5.73
Protein content (%)	40	8.31	13.04	10.32	1.42

3 Experimental results and analysis

3.1 Data Preprocessing

Average smoothing method was used in mid-infrared spectral pretreatment to remove the influence of high-frequency random noise, baseline drift, uneven sample and light scattering. The choice of the smoothing window size was 3. Figure 1 is the buckwheat spectral curves after pretreatment. In the figure, the abscissa is the wave number (cm^{-1}), and the ordinate is the transmission rate (%).

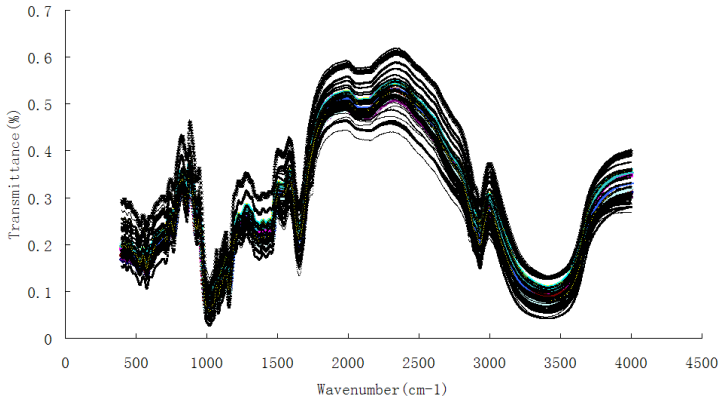

Fig.1. Transmittance of buckwheat Mid-infrared spectroscopy

3.2 Principal components extraction based on principal component analysis

Forty samples were randomly divided into the prediction set and the validation set with 30 and 10 samples respectively. Sample spectral bands from 4000 to 400 cm^{-1}, there are 1868 data in total. Computational complexity is large when using the full spectrum of computing. And due to the weak spectral information of some regional samples, the composition of the sample or the nature is lack of correlation. The principal component analysis is an effective method for data mining. It can translate the original multiple-wavelength variables into fewer new variables. These new variables, not only are not related to each other, but also can synthetically reflect the information of original multiple-wavelength variables [8, 9]. So the prediction set and the validation set were respectively for principal component analysis. The accumulative reliabilities of the principal components are shown in Table 2.

Table2. Accumulative reliabilities of the first 6 PCs

Principal component	Accumulative reliabilities of prediction set/%	Accumulative reliabilities of validation set/%
PC01	89.831	92.286
PC02	96.087	98.121
PC03	99.488	99.544
PC04	99.916	99.846
PC05	99.985	99.946
PC06	99.995	99.972

3.3 Prediction model of buckwheat starch content based on BP neural network

The score of the six principal components which extracted by PCA from the 30 prediction set is the input variables of the BP neural network, and the output layer is the measured value of buckwheat starch content. By order experiments, the best hidden layer nodes of buckwheat starch content is 6, the iteration number of network setting training is 1000, the target error setting is 0.0001, the fitting residual error of the modeling samples is 1.805×10^{-3}. The calibration model of buckwheat starch content was established to predict the 10 validation set, the correlation analysis for prediction and measured values are shown in Figure 2. The error analysis results in Table 3.

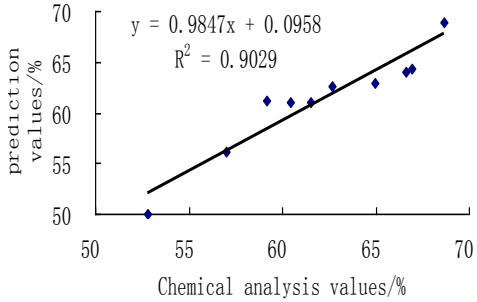

Fig.2 Relationship between chemical analysis values and prediction values of buckwheat starch content

Table 3 Prediction results for buckwheat starch by BP model

Sample NO.	Chemical analysis value(%)	Prediction value(%)	Absolute error(%)	Relative error(%)
1	60.43	61.06	0.63	1.04
2	62.68	62.61	-0.06	-0.10
3	64.97	62.97	-1.99	-3.07
4	66.62	64.08	-2.54	-3.81
5	68.69	68.94	0.25	0.36
6	66.93	64.39	-2.54	-3.79
7	59.20	61.16	1.96	3.31
8	57.03	56.07	-0.96	-1.68
9	52.80	49.99	-2.82	-5.33
10	61.55	61.08	-0.47	-0.76

3.4 Prediction model of buckwheat protein content based on BP neural network

The score of the six principal components which extracted by PCA from the 30 prediction set is the input variables of the BP neural network, and the output layer is

the measured value of buckwheat protein content. By order experiments, the best hidden layer nodes of buckwheat protein content is 6, the iterations number of network setting training is 1000, the target error setting is 0.0001, the fitting residual error of the modeling samples is 4.173×10^{-3}. The calibration model of buckwheat protein content was established to predict the 10 validation set, the correlation analysis for prediction and measured values are shown in Figure 3. The error analysis results in Table 4.

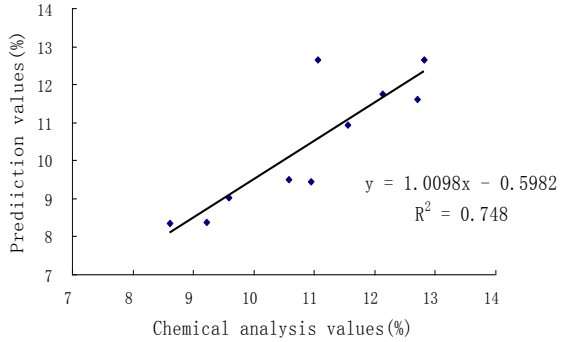

Fig.3 Relationship between chemical analysis values and prediction values of buckwheat protein

Table 4 Prediction results for buckwheat Protein by BP model

Sample NO.	Chemical analysis value(%)	Prediction value(%)	Absolute error(%)	Relative error(%)
1	12.71	11.61	-1.09	-8.61
2	11.56	10.95	-0.61	-5.25
3	10.58	9.49	-1.09	-10.34
4	9.58	9.02	-0.56	-5.83
5	8.61	8.36	-0.25	-2.92
6	9.23	8.38	-0.85	-9.20
7	11.06	12.66	1.60	14.50
8	10.95	9.45	-1.50	-13.73
9	12.82	12.65	-0.17	-1.33
10	12.13	11.74	-0.39	-3.20

4 Conclusion and discussion

(1) The neural network prediction model of buckwheat starch and protein content were established respectively using principal component analysis and artificial neural network method, and its predictions results has been verified. The results showed that the predict correlation of the artificial neural network prediction model for buckwheat

starch content on the 10 samples validation set is 0.9029, the absolute average of the relative error is 2.33%. So, the method is feasible to predict the starch content in buckwheat.

(2) The buckwheat protein content prediction model was established in our paper. For the 10 samples, the predict correlation of the model is 0. 7480, the absolute average of the relative error is 7.9%. The results showed that the prediction effect for buckwheat protein content is not ideal using infrared spectroscopy, it is necessary to increase the buckwheat varieties and the number of the samples to optimize the modeling methods.

References

1. Yuzhen Guo, Qingfu Chen, Lanying Yang et.al. : Analyses of the seed protein contents on the cultivated and wild buckwheat Fagopyrum esculentum resources. J. Genet Resour Crop Evol. 54, 1465--1472 (2007)
2. SiuMei Choi, Chingyung Ma: Structural characterization of globulin from common buckwheat (Fagopyrum esculentum Moench) using circular dichroism and Raman spectroscopy. J. Food Chemistry. 102, 150--160(2007)
3. Tu Zhenhua, Zhu Dazhou, Ji Baoping,et.al. : Progress in quality analysis of honey by infrared spectroscopy (in Chinese). J. Spectroscopy and Spectral Analysis. 30(11), 2971--2975(2010)
4. Iwona Scibisz,, Maryse Reich, Sylvie Bureau, et al.. Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. J. Food Chemistry.125, 1390--1397(2011)
5. Wang Dong, Wang Jihua, Ma Zhihong, et al.: Quantitative determination of lime additives in wheat flour by mid-infrared spectroscopy combined with attenuated total reflection(in Chinese) . J. Transactions of the CSAE. 28(4), 258--263(2012)
6. Engel R E,Long D S. : Measuring grain protein concentration with in-line near infrared reflectance spectroscopy. J. Crop Management. 10(5), 1--3(2005)
7. Song Tao, Zhang Fengping, Liu Yaomin, et al.: Rapid determination of fatty acids in soybean oils by transmission reflection-near infrared spectroscopy (in Chinese). J. Spectroscopy and Spectral Analysis. 32(8), 2100--2104(2012)
8. Zhang Shujuan, Wang Fenghua, Zhang Haihong,et al. : Detection of the Fresh Jujube Varieties and SSC by NIR Spectroscopy(in Chinese). J. Transactions of the Chinese Society of Agricultural Machinery. 40(4), 139--142(2009)
9. Zhang Shujuan, Wang Fenghua, Zhang Haihong,et al. : Near-Infrared Determination of Soil Nutrients Based on Principal component Analysis and BP Neural Network(in Chinese). J. Journal of Shanxi Agricultural University, 29(6), 483--487(2009)