Analysis on Spectral Matching of Biochemical Component for Rice

Tianyu Tang, Yanlin Tang and Xiaonan Wei

College of Physics, Guizhou University, Guiyang, Guizhou, 550025, P.R. China
Email: tangtianyu1992@163.com, tylgzu@163.com

Abstract. The protein, starch and amylase are three important indexes to characterize rice quality. The starch, crude protein and amylase of rice were extracted by experiments. The hyperspectral reflectance of starch, crude protein and amylase, their mixed samples and rice samples were determined by a ASD FieldSpec Pro FR™. The spectral characteristics of starch, crude protein, amylase and their correlation with component content were analyzed by using spectral matching technique and multivariate statistical method. The results showed that the spectra of three biochemical components were significantly different, and the spectral peaks and valleys of the mixed samples showed “red shift” or “blue shift”. The contents (%) of crude protein, starch and amylase in rice flour were significantly related to the absorption area S between 2020nm and 2235nm on their spectral curve. The results showed that the hyperspectral method could be used to estimate the contents of crude protein, starch and amylase content in rice, and then to detect rice quality.

Keywords. Spectral matching, Crude protein, Starch, Amylase; Rice.

1. Introduction
In 60-70s of the 20th century, the researchers from the United States Department of Agriculture (USDA) measured and analyzed the spectra of dried and pounded leaves of various plants in detail, and obtained about 42 absorption characteristics of certain biochemical components in the range of 0.4μm to 2.4μm [1]. It is possible to use remote sensing technology to estimate the biochemical components of plants. The biochemical components of plants mainly refer to various pigments (such as chlorophyll, lutein, carotenoid, etc.), various nutrients (such as N, P, K), cellulose, hemicellulose, lignin, starch, sugar, protein and fat, etc. Previous experiments showed that the absorption of biochemical components in crops within the range of 0.4μm to 2.4μm has some obvious characteristics, and the central absorption wavelength is respectively 970nm, 1180nm and 1450nm for water, 480nm and 660nm for chlorophyll, 1730nm, 2100nm and 2300nm for cellulose, 760nm, 1640nm and 2100nm for nitrogen, 1730nm and 2300nm for lignin and 990nm, 1450nm, 1530nm, 1900nm and 2250nm for starch [1]. There have been many reports to study the chemical components of living organisms by reflection spectroscopy. Jin et al. estimated leaf biochemical contents from hyperspectral reflectance by different PLS regressions [2].
Card et al. showed that the chemical component content of leaves could be estimated by using the variables \( \log (1/R) \) of spectral reflectance \( R \) and the first and second derivatives of \( \log(1/R) \) [3]. Yang et al. established different regression algorithms for estimating nitrogen concentration in paddy rice leaf by the first-derivative fluorescence spectrum [4]. Kokaly et al. showed that the nitrogen content of rice leaves could be estimated by using the spectral absorption characteristics of 2054nm and 2172nm [5]. LaCapra et al. estimated nitrogen and lignin contents in rice leaves by imaging spectrometer, and found that lignin had a poor correlation with \( \log(1/R) \) (\( R^2 =0.44 \)), while nitrogen had a good correlation with \( \log(1/R) \) (\( R^2 =0.74 \) [6]. Xie et al. studied the prediction model of the crude protein content, amylose content of rice and actual yield under high temperature stress based on hyper-spectral remote sensing [7]. Japanese scientist Asaka et al. used satellite remote sensing to estimate of protein contents in winter wheat grain [8]. Curran, Kokaly et al. estimated the biochemical components in the dry leaves by stepwise regression according to the reflection spectrum, and the results showed that the estimation accuracy from high to low was total chlorophyll, chlorophyll a, chlorophyll b, nitrogen, sugar, cellulose, lignin, water, phosphorus, protein, amino acid, starch, respectively [9, 10].

At present, near-infrared (NIR) spectroscopy has been widely used to analyze the quality of grain and feed, but it is generally based on absorption spectrum, and there are few reports on the use of reflection spectrum. Two important indexes to measure rice quality are protein content and amylose content of rice [11]. Wu, Yadav et al. measured proteins and amino acids in rice by NIR reflectance spectrometer (1100-2500nm) [12, 13]. Barton et al. used a NIR spectrometer (400-2498nm) to analyze the quality of rice (apparent amylopectin, protein, ash, etc.) [14]. There were many researchers using NIR spectroscopy to detect the quality of rice [15-22], and the detection results had the same accuracy as those obtained by chemical analysis method, Fourier transform infrared absorption spectroscopy and Raman spectroscopy. An automatic rice quality detection system based on spectr um has been developed in Japan [23]. ASD FieldSpec Pro FR™ spectrometers work the same way as a visible- near-infrared spectrometer. Tang et al. used hyperspectral to estimate the content of crude protein and starch of rice ears and grains [24], which indicated that the content of biochemical components in rice could be detected by hyper-spectrometer. The purpose of this paper was to analyze the spectral characteristics and matching effect of rice protein, amylose and starch (including amylose and amylopectin) by using hyper-spectrometer, so as to provide the basis for the detection of rice quality by hyperspectral technology and the monitoring of rice quality by using hyperspectral remote sensing method.

2. Instruments and Measurements

2.1. Instruments

The FieldSpec Pro FR™ spectrometer was produced by ASD Inc, USA, and the spectral value was 350–2500nm, of which the spectral sampling interval (band width) of 350–1000nm was 1.4nm and the spectral resolution was 3nm, and the spectral sampling interval (band width) of 1000–2500nm was 2nm and the spectral resolution was 10nm.

2.2. Test Method

2.2.1. Test Samples. The rice planted with different nitrogen levels was harvested and threshed, and the
randomly sampled grains were dried at 45°C under constant temperature. The dried rice was shelled to obtain the polished rice, which was crushed through 200 mesh sieve to make rice flour and then put into the drying oven for experiment. The starch, crude protein and amylose of rice were extracted with white rice. The extraction method of rice crude protein followed the method of reference [25]. The extraction of starch followed the method of reference [26]. Amylose was extracted according to the method of reference [27].

2.2.2. *Spectral Measurement*. The sample was placed in a black box with a height of 2cm and a diameter of about 5cm. The box is full of samples. The spectrometer probe was fixed with a tripod, and the spectrometer probe was positioned vertically down to the middle of the sample, 0.15m away from the sample surface. The angle of view of the spectrometer is 8° (the field of view diameter of the sample surface is about 2.1cm). The light source is a 50W halogen lamp carried by the spectrometer, with a distance of 0.45m from the sample surface and a azimuth of 70° (angle between the sample surface). Standard whiteboard calibration is performed before each data collection. When measuring spectral data, 10 sampling spectra are taken as one spectrum, 10 spectra are recorded each time, and then the average is calculated. The data analysis software is SPSS22.0.

2.2.3. *The Contents of Starch, Crude Protein and Amylose*. The crude nitrogen content was measured by Kjeldahl method. According to the principle that the general rice protein contains about 16.8% nitrogen, the measured total nitrogen was multiplied by the conversion factor K=5.95 to calculate the crude protein content. The starch was measured by Anthranone method, and the amylose was measured by reference [27].

3. Results and Discussion

3.1. *Spectrum of Starch, Crude Protein and Amylose in Rice*

The spectra of starch, crude protein and amylose extracted from brown rice are shown in figure 1 (a). Here A, P and S stand for amylose, crude protein and starch respectively. It can be seen from the figure that the spectral curves of protein, amylose and starch are quite different. The reflectivity of crude starch is significantly higher than that of crude protein. Above 1200nm, the location and number of reflection peaks and valleys and the relative depth of reflection valleys are significantly different. The reflectance spectra of amylose and starch are basically the same shape. This is because the main absorption factors affecting the curve shape of protein reflection spectrum are the vibration, bending and deformation of C-H, O-H, N-H and N=H bonds, and the main absorption factors affecting the curve shape of the reflection spectra of amylose and starch are the vibration, bending and deformation of the C-H, CH₂, C-O and O-H bonds among them in the range of 350nm-2400nm according to the principle of physics and chemistry. Therefore, the reflection spectrum curve of protein is significantly different from that of starch in shape, absorption peak position and number, which is mainly reflected in the following aspects: protein has three small absorption peaks around 410nm, 2060nm and 2180nm, while starch does not. The main reflection peaks and valleys differ by more than 10nm.

Figure 1 (b) shows the reflection spectra of samples with different ratios (mass ratios) of starch and crude protein extracted from rice. When the content of protein in the mixture is less than 50%, the three
small absorption peaks of protein around 410nm, 2060nm and 2180nm in the mixture spectrum were no longer obvious. The peak and valley positions of the reflection spectrum of the mixed samples show red shift and blue shift with the different mixing ratio of starch and crude protein. Theoretically, the contents of starch, crude protein and amylose in mixed samples can be estimated by extracting spectral reflectance values and curve shape features (such as derivative spectrum, area and width of reflection valley, value of "blue shift" and "red shift", SAI, etc.) [28].

![Spectrum of starch, crude protein and amylose in rice.](image)

**Figure 1.** Spectrum of starch, crude protein and amylose in rice.

The relation between the absorption area S (The region surrounded by the reflected spectral line and the non-absorptive reference line is between 2020nm and 2235nm) and the relative content of crude starch and crude protein was analyzed for the spectral curve of mixtures with different ratios of starch and crude protein near 2100nm (2020~2235nm) (see figure 2). The results showed that there was a significant negative correlation between the relative content of crude protein and area S, and a significant positive correlation between the relative content of starch and area S, with a determination coefficient of $R^2$ above 0.997.

![Comparison of absorption peaks of starch and crude protein.](image)

**Figure 2.** Comparison of absorption peaks of starch and crude protein.

The infrared absorption produced by the chain end group, side chain and the absorbed water of the
biopolymer often adds up to an indistinguishable absorption peak, but derivative spectroscopy can be used to separate these absorption characteristics. Figure 3 is the first derivative spectrum of rice starch and crude protein mixture. As can be seen from Figure 3, the peak and valley positions of spectral reflection of the original mixed sample (the first derivative is zero) change with the change of component proportion. The first derivative value of the same wave band also varies with the proportion of components, and the difference is obvious.

**Figure 3.** First derivative spectra of rice starch and crude protein mixture.

### 3.2 Application of Spectral Matching in Hyperspectral Characteristic Analysis of Rice

The spectrum of rice samples from different varieties (japonica rice and indica rice) was measured in the experiment, as shown in figure 4. It showed that the spectral curves of various rice samples were similar.

The absorption area S between 2020 and 2235nm of spectral curves of different rice samples was calculated respectively. The results showed that the correlation between the content of starch, crude protein, amylose and absorption area S in rice reached an extremely significant level. Based on the stepwise regression method, the prediction models of content of starch, crude protein and amylose of rice are as follows (n=20):

\[
P_C = -0.5496S + 17.165 \quad R^2=0.639^{**} \\
S_C = 3.0594S + 39.551 \quad R^2=0.884^{**} \\
A_C = -0.6608S + 26.441 \quad R^2=0.671^{**}
\]

Through the test of the above prediction models, it is found that the correlation between the measured value and the predicted value is very high, reaching the extremely significant test level (n=20), and the test determination coefficient $R^2$ is 0.897, 0.908 and 0.83, respectively.
4. Conclusions
From the above analysis, it can be seen that the absorption peaks of starch and crude protein of rice are mainly located in the NIR region, and their absorption peaks overlap with each other, as well as those of other biochemical components (such as lignin and cellulose) [29]. Derivative spectrum can be used to excavate the variation trend of these reflection valleys, which is helpful to separate the absorption characteristics of different biochemical components. The results showed that the high content biochemical components in the sample had better estimation effect than the low content biochemical components when using spectral method.

In this paper, the spectral characteristics of rice starch, crude protein and amylose and the spectral characteristics of different proportions of rice starch and crude protein were analyzed. The results show that:

(1) The spectra of protein were significantly different from those of amylose and starch. The reflectance spectra of amylose and starch are basically the same because amylose and amyllopectin have the same chemical groups.

(2) When different proportions of starch and crude protein are mixed, the reflection spectrum of the mixture will show a "blue shift" or "red shift" phenomenon compared to the reflection spectrum of pure material, which is exactly the result of spectrum matching.

(3) The contents of starch, crude protein and amylose in rice are significantly correlated with the absorption area S between 2020 and 2235nm of its spectral curve.

In this paper, the spectral matching of starch, crude protein and amylose in rice was studied, which may provide an idea for monitoring rice quality in field with remote sensing technology.

Acknowledgments
Project supported by the National Natural Science Foundation of China (No. 11164004, 61835003) and the Guizhou Provincial Photon Science and Technology Innovation Talent Team (No. 20154017).

References
[1] Pu R and Gong P. 2000 *Hyperspectral Remote Sensing and Its Application*, Higher Education Press, Beijing.
[2] Jin J and Wang Q. 2019 Remote Sensing 11 197.
[3] Card D, Peterson D and Matson P. 1988 Remote Sensing of Environment 26 123-147.
[4] Yang J, Du L, Cheng Y, Shi S, Xiang C, Sun J and Chen B. 2020 Optics Express 28 (13) 18728.
[5] Kokaly R. 2001 Remote Sensing of Environment, 75 153-161.
[6] LaCapra V., Melack J, Gastil M and Valeriano D. 1996 Remote Sensing of Environment 55 50-58.
[7] Xie X, Zhang Y, Li R, Shen S and Bao Y. 2019 Quality assurance and safety for Crops & Foods, 11 (6) 517-527.
[8] Asaka D, Hayashi T and Shiga H. 2006 Japanese Society of Soil Science and Plant Nutrition 77 (3) 317-320.
[9] Curran P, Dungan J and Gholz H. 2001 Remote Sensing of Environment 76 349-359.
[10] Kokaly R and Clark R. 1999 Remote Sensing of Environment 67 267-287.
[11] Chen N, Luo Y, Zhu Z, Zgang B, Zheng Y and Xie L. 1997 Chinese Journal of Rice Science 11 (2) 70-76.
[12] Yadav B and JindaI V. 2001 Computers and electronics in agriculture 33 19-33.
[13] Wu J, Shi C and Zhang X. 2002 Field Crops Research 75 1-7.
[14] Barton F, Himmelsbach D and McClung A. 2000 Cereal Chemistry 77 (5) 669-672.
[15] Sun D, Cen H and Wang H. 2019 Plant Methods 15 54.
[16] Lapchareonsuk R and Sirisomboon P. 2015 International Journal of Food Properties 18 1128–1138.
[17] Lu X, Sun J, Mao H, Wu X and Gao H. 2017 International Journal of Food Properties 20 (1) 1037–1044.
[18] Zhang J, Li M, Pan T, Yao L and Chen J. 2019 Computers and Electronics in Agriculture 164 104882.
[19] Lin L, He Y, Xiao Z, Zhao K, Dong T and Nie P. 2019 Applied Science 9 1654.
[20] Lin P, Zhang H, He J, Zou Z and Chen Y. 2020 Spectroscopy and Spectral Analysis 40 (1) 233–238.
[21] Ernest T, Charles L, Terry M and Christopher E. 2019 Spectrochimica Acta-Part A: Molecular and Biomolecular Spectroscopy 217 147-154.
[22] Nicola C, Martin B and Ian D. 2018 Applied Spectroscopy Reviews 53 (8) 667–687.
[23] Shuso K, Motoyasu N and Kazuhiro T. 2003 Computers and Electronics in Agriculture 40 115-126.
[24] Tang Y, Huang J and Wang R. 2004 Chinese Scientia Agricultura Sinic 37 (9) 1282-1287.
[25] Liu M, Li X, Liu Y, Huang J and Tang Y. 2014 Spectroscopy Letters 47 101–106.
[26] Yi C and Yao H. 2004 Food & Machinery 20 (6) 18-21.
[27] Hong Y, Gu Z and Liu X. 2004 Science and Technology of Food Industry 25 (4) 86-88.
[28] Wang J, Zhang B, Liu J, Tong Q and Zheng L. 1999 Chinese Journal of Image and Graphics 4 (11) 957-964.
[29] Curran P. 1989 Remote Sensing of Environment 30 271-278.