Rapid quantification of rice (*Oryza sativa*) qualities based on adaptive near infrared spectroscopy

R Hayati*, A A Munawar², A Marliah¹

¹Department of Agrotechnology, Universitas Syiah Kuala, Jalan Tgk. Hasan Krueng Kalee 3, Banda Aceh 23111, Indonesia
²Department of Agricultural Engineering, Universitas Syiah Kuala, Jalan Tgk. Hasan Krueng Kalee 3, Banda Aceh 23111, Indonesia

*Email: rita@unsyiah.ac.id

Abstract. Determination of rice quality parameters is the key factor affecting sustainable agriculture practices. The main purpose of this present study is to develop prediction models based on adaptive near infrared spectroscopy (NIRS) for rapid quantification of rice qualities in form of protein content. Rice samples were obtained from several paddy field in Aceh province with different cultivars. Near infrared spectral data of rice samples were acquired and in wavelength range from 1000 to 2500 nm and recorded as diffuse reflectance spectrum. Prediction models were established using principal component analysis (PCA), principal component analysis (PCR) and partial least square regression (PLSR). The results showed that NIRS combined with PCA can classify rice samples based on their cultivars. Moreover, this approach with PCR and PLSR can also predicted and determined protein contents with satisfactory performance achieving maximum correlation coefficient ($r$) of 0.81 and ratio prediction to deviation (RPD) index of 2.84 for PCR and $r$ of 0.90 and RPD of 3.19 for PLSR respectively. Based on achieved results, it may conclude that adaptive NIRS approach can be used to quantify rice qualities rapidly and non-destructively.

1. Introduction
Rice (*Oryza sativa* L.) is one of the food crops that serves as a staple crop other than wheat and corn in Indonesia. Rice contains nutrients that the body needs such as carbohydrates (46.45%), protein (2.09%), water (49.15%) and fat (2.05%). Indonesia is an agrarian country because its population lives from the agricultural sector. In 2020 the production of dry milled grain (GKG) in Indonesia was 92.41 million tons. Then there was an increase of 4.62% in 2019 so that GKG was obtained as much as 90.38 million tons. The increase in rice production in 2019 was caused by an increase in harvested area of 2.63 million hectares (4.17%), while in 2018 the harvested area was 21.35 million hectares and decreased productivity by 1.81 quintals/hectare (-1.55%) from 52.36 quintals/hectare to 62.55 quintals/hectare [1].

The average use of rice for consumption is 169 kg/capita/year with a total rice demand of 32.66 million tons/year. The increase in population causes the need for rice to also increase. To meet the needs, it is necessary to make new varieties that can improve the quality of rice in the future. Based on this, a breeding technology is needed that can meet these criteria [2,3]. *Sanbei* rice originating from Simeulu is one of the local varieties from Aceh that has been irradiated with gamma rays at the isotope and radiation technology application center.
In addition, the Sanbei rice line also had better M1 vegetative growth compared to several other varieties that had been irradiated. The yield potential and physicochemical properties of the Sanbei rice line from the M3 mutation and it is found that the yield potential of the Sanbei genotype is superior, and has better physicochemical and organoleptic properties. However, it is necessary to study the next generation of M3 to determine the effect of given gamma ray irradiation so that quality stability can be known. This is due to mutations/radiation having random traits, so the results obtained are positive (superior traits, expected traits) and negative (undesirable traits).

Rice quality can be measured using several test parameters, namely physical quality: yield, grain size and shape and liming, cooking quality like gelatinization temperature and amylose content and taste quality, aroma and texture [4–6]. However, these methods are based on standard laboratory procedures with solvent extraction and involved chemical materials. Thus, unsuitable for rapid application for quantifying rice qualities [7,8].

In recent decades, near infrared spectroscopy or known as NIRS has several advantages over conventional chemical analysis because it is non-destructive, rapid and can analyze samples to determine several quality parameters simultaneously [8–11]. Therefore, the main objective of this present study is to develop prediction models based on adaptive near infrared spectroscopy (NIRS) for rapid quantification of rice qualities in form of protein and fat content.

2. Materials and methods

2.1. Rice samples
Rice samples were collected and harvested from different paddy field in several areas of Aceh Province. Some of them are also from mutant rice genetically modified and irradiated using X-rays.

2.2. Near infrared spectra acquisition
Near infrared spectral data of rice samples were acquired and in wavelength range from 1000 to 2500 nm and recorded as diffuse reflectance spectrum. Spectral data were enhanced by means of baseline correction approach, while prediction models were established using principal component analysis (PCA), principal component regression (PCR) and partial least square regression (PLSR) [12,13].

2.3. Rice classification
Classification models for rice samples were established by means of principal component analysis with non-iterative partial least square or NIPALS algorithm. The models were then inspected by looking their explained variance value of the PCA score plot respectively [14].

2.4. Optical properties of rice
In order Spectral data of rice samples were transformed onto several spectra format to observe more details of rice samples optical properties. The methods used for spectra transformation were derivative and standard normal variate.

2.5. Optimum wavelengths
The most relevant and optimum wavelengths for rice classification were determined by looking the peak and valley in regression coefficient plot or loading plot.

3. Results and discussion

3.1. Spectra features of rice samples
The near infrared spectra feature of rice samples is shown in figure 1. The advantages of infrared ray technology that do not damage materials, relatively easy sample preparation, do not involve chemicals in the testing process and can predict several qualities or nutrients simultaneously, make this method widely researched and applied in many fields, including food and agriculture. In this study, we try to
examine in more depth this infrared ray technology for rapid testing of quality levels of whole rice samples, where the sample to be used is *Sigupai* Aceh rice from the M4 mutation.

**Figure 1.** Spectra features of rice samples in near infrared wavelength region

This study investigates the feasibility of near infrared technology in classifying two different types of rice based on their aroma. The results showed that the two rice cultivars could be classified and differentiated precisely with the support of principal component analysis (PCA) resulting in a total variance explained 100%. The chemical parameters of each of these differences are fibre content, carbohydrates and amylose content. The three parameters vibrated strongly at wave numbers of 4260, 7512 and 7900 cm\(^{-1}\) for carbohydrates and amylose, while for fibre content they vibrated at 5183 cm\(^{-1}\), respectively. Based on the results obtained, it can be concluded that the near infrared spectroscopic approach can classify and distinguish rice varieties based on their aroma through optical spectral properties in the near infrared region.

After the spectrum acquisition has been carried out on unmixed and mixed rice, the next step is to detect the presence of outliers using the principal component analysis as presented in Figure 2. The PCA was also combined with the *Hotelling* \(T^2\) ellipse method which will make the data more perfect. This outlier data detection method is used to inspect whether the data is outlier or leveraged data. Outlier data will appear if there is a data point outside the ellipse with a position in the middle, while leveraged data occurs when the data position is at the end of the cluster.

**Figure 2.** Spectra of rice samples in PCA and *Hotelling* \(T^2\) ellipse.

The data that is outside the elliptical line is not included in the outlier data but as leveraged data because this data does not damage other data, so there is no data that needs to be removed. Moreover, spectra
data were projected to observe more details related to sample distribution in the near infrared spectra region as shown in Figure 3.

![Figure 3. Principal component analysis of different rice samples](image)

The shape of the diffuse reflectance spectrum for the rice sample occurs due to changes in energy vibrations in the form of overtone, bending and stretching. Based on this spectrum, absorption which indicates the presence of water content occurs at wavelengths of 1460 and 1940 nm. The first overtone of the O-H molecule is clearly visible in this wavelength range, so it can be said that the band for water absorption occurs in the wavelength range of 1400 and 1900 nm. This result is also reinforced by the statement of some literatures which states that water absorption occurs at wavelengths of 1420 – 1490 nm and 1900 – 1980 nm. Furthermore, based on the results of this study, it was found that the C-H-O structure vibrates (first and second overtones) in the wavelength range of 2200 – 2300 nm [15 – 17].

The C-H-O element relates to the nutritional content of carbohydrates in rice, so it can be concluded that the sugar content or the total dissolved solids content can be predicted by the NIR method as a new method that is rapid, effective, free of chemicals, without pollution and without damaging the ingredients. While the structure of the N-H molecule vibrates (first overtone) in the wavelength range of 1500 – 1600 nm; and the molecular structure of vibrating organic acids (first overtone and bending) in the wavelength range of 1400, 1800 and 2100 nm.

The peak of the third wave occurs at a wavelength in the range of 4617 cm⁻¹ - 4864 cm⁻¹ or 2055 nm to 2165 nm. These wavelengths indicate the presence of C-O-N-H chemical compounds where these compounds are suspected to be protein content. The same thing was also proven in the research
conducted by [18] where the protein content was found at a wavelength of 4752 cm\(^{-1}\) (2104 nm). And also in accordance with the statement of [8, 19] that the presence of protein content is at a wavelength of 4739 cm\(^{-1}\). The prediction performance for protein prediction of rice samples using raw spectra data by means of principal component regression (PCR) and partial least square regression (PLSR) are presented in Figure 4 and Figure 5 respectively.

![Figure 4. Prediction performance of PCR raw spectrum for protein content determination.](image)

The peak of the fourth wave occurred at a wavelength of 5049 cm\(^{-1}\) - 5234 cm\(^{-1}\) (1910 nm - 1980 nm) indicating the presence of chemical compounds H2O where these compounds are suspected as water content. The peak of the fifth wave occurs at a wavelength in the range of 6777 cm\(^{-1}\) – 7085 cm\(^{-1}\) (1411 nm – 1475 nm). These wavelengths indicate the presence of chemical compounds C-O-N-H and H2O where these compounds are suspected to be protein and water content.

The wavelength that indicates the presence of protein and water content in this study is similar to the results of Chen et al (2008) research with a wavelength of 5170 cm\(^{-1}\) (1934 nm) indicating the presence of water content and a wavelength of 6821 cm\(^{-1}\) (1466 nm) associated with water and protein content as shown in loading plot Figure 6. This is also in accordance with the statement of [20] that there is water content at wavelengths of 5154 cm\(^{-1}\) (1940 nm) and 6896 cm\(^{-1}\) (1450 nm).
Rice as a food ingredient is composed of starch, protein, and other elements such as fat, crude fiber, minerals, vitamins, and water with uneven distribution. The outer layer of rice is rich in non-starch components such as protein, fat, fiber, ash, pentosan and lignin. While the endosperm is rich in starch. The main carbohydrate in rice is starch and only a small part is pentosan, cellulose, hemicellulose, and sugar.

![Loading plot to determine optimum wavenumbers for protein prediction](image)

**Figure 6.** Loading plot to determine optimum wavenumbers for protein prediction

Rice starch is composed of two carbohydrate polymers, namely amylose and amylopectin. The nature of the texture of rice can be seen from the comparison between the levels of amylose and amylopectin. The ratio of amylose and amylopectin can determine the taste or fluffiness of the rice texture. Amylose content determines the texture properties of rice more than other properties. Amylose content in rice ranges from 1-37%.

4. **Conclusions**

Determination of rice quality parameters is the key factor affecting sustainable agriculture practices. The main purpose of this present study is to develop prediction models based on adaptive near infrared spectroscopy (NIRS) for rapid quantification of rice qualities in form of protein content. The results showed that NIRS combined with PCA can classify rice samples based on their cultivars. Moreover, this approach with PCR and PLSR can also predicted and determined protein contents with satisfactory performance achieving maximum correlation coefficient (r) of 0.81 and ratio prediction to deviation (RPD) index of 2.84 for PCR and r of 0.90 and RPD of 3.19 for PLSR respectively. Based on achieved results, it may conclude that adaptive NIRS approach can be used to quantify rice qualities rapidly and non-destructively.

**Acknowledgement**

We sincere our gratitude and appreciation to the directorate of research and community services LPPM Universitas Syiah Kuala for funding this study through PLK scheme 2021.

**References**

[1] Munawar A A and Sabaruddin Z 2021 Fast classification of rice (Oryza sativa) cultivars based on fragrance and environmental origins by means of near infrared spectroscopy *IOP Conf. Ser. Earth Environ. Sci.* **644** 012003

[2] Hwang Y, Ryu Y, Huang Y, Kim J, Iwata H and Kang M 2020 Comprehensive assessments of
carbon dynamics in an intermittently-irrigated rice paddy Agric. For. Meteorol. 285–286 107933
[3] Moulick D, Samanta S, Sarkar S, Mukherjee A, Pattnaik B K, Saha S, Awashi J P, Bhowmick S, Ghosh D, Samal A C, Mahanta S, Mazumder M K, Choudhury S, Bramhachari K, Biswas J K and Santra S C 2021 Arsenic contamination, impact and mitigation strategies in rice agro-environment: An inclusive insight Sci. Total Environ. 800 149477
[4] Limmer M A, Mann J, Amaral D C, Vargas R and Seyfferth A L 2018 Silicon-rich amendments in rice paddies: Effects on arsenic uptake and biogeochemistry Sci. Total Environ. 624 1360–8
[5] GAO S juan, GAO J sheng, CAO W dong, ZOU C qin, HUANG J, BAI J shun and DOU F gen 2018 Effects of long-term green manure application on the content and structure of dissolved organic matter in red paddy soil J. Integr. Agric. 17 1852–60
[6] Müller A, Coradi P C, Nunes M T, Grohs M, Bressiani J, Teodoro P E, Anschau K F and Flores E M M 2021 Effects of cultivars and fertilization levels on the quality of rice milling: A diagnosis using near-infrared spectroscopy, X-ray diffraction, and scanning electron microscopy Food Res. Int. 147 110524
[7] Munawar A A, von Hörsten D, Wegener J K, Pawelzik E and Mörlein D 2016 Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics Eng. Agric. Environ. Food 9
[8] Pasquini C 2018 Near infrared spectroscopy: A mature analytical technique with new perspectives – A review Anal. Chim. Acta 1026 8–36
[9] Marques E J N, De Freitas S T, Pimentel M F and Pasquini C 2016 Rapid and non-destructive determination of quality parameters in the “Tommy Atkins” mango using a novel handheld near infrared spectrometer Food Chem. 197 1207–14
[10] Munawar A A, von Hörsten D, Wegener J K, Pawelzik E and Mörlein D 2016 Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics Eng. Agric. Environ. Food 9 208–15
[11] Nordey T, Joas J, Davrieux F, Chillet M and Léchaudel M 2017 Robust NIRS models for non-destructive prediction of mango internal quality Sci. Hortic. (Amsterdam). 216 51–7
[12] Kusumiyati, Munawar A A and Suhandy D 2021 Fast, simultaneous and contactless assessment of intact mango fruit by means of near infrared spectroscopy AIMS Agric. Food 6 172–84
[13] Munawar A A, Yunus Y, Devianti and Satriyo P 2020 Calibration models database of near infrared spectroscopy to predict agricultural soil fertility properties Data Br. 30
[14] Munawar A A, Kusumiyati and Wahyuni D 2019 Near infrared spectroscopic data for rapid and simultaneous prediction of quality attributes in intact mango fruits Data Br. 27
[15] Islam M S, Magid A S I A, Chen Y, Weng L, Ma J, Arafat M Y, Khan Z H and Li Y 2021 Effect of calcium and iron-enriched biochar on arsenic and cadmium accumulation from soil to rice paddy tissues Sci. Total Environ. 785
[16] Yang K, Ryu Y, Dechant B, Berry J A, Hwang Y, Jiang C, Kang M, Kim J, Kimm H, Kornfeld A and Yang X 2018 Sun-induced chlorophyll fluorescence is more strongly related to absorbed light than to photosynthesis at half-hourly resolution in a rice paddy Remote Sens. Environ. 216 658–73
[17] Awanthi M G G, Jinendra B M S, Navaratne S B and Navaratne C M 2019 Adaptation of visible and short wave Near Infrared (VIS-SW-NIR) common PLS model for quantifying paddy hardness J. Cereal Sci. 89 102795
[18] Deng Y, Wang Y, Zhong G and Yu X 2018 Simultaneous quantitative analysis of protein, carbohydrate and fat in nutritionally complete formulas of medical foods by near-infrared spectroscopy Infrared Phys. Technol. 93 124–9
[19] Cen H and He Y 2007 Theory and application of near infrared reflectance spectroscopy in determination of food quality Trends Food Sci. Technol. 18 72–83
[20] Cozzolino D 2014 An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals *FRIN* **60** 262–5