Nondestructive evaluation of Bakwan paddy grains moisture content by means of spectrophotometry

M Makky1, R E Putry1, K Nakano2 and Santosa1

1Dept. of Agricultural Engineering, Andalas University, Kampus UNAND Limau Manis Padang 25163 West Sumatera, Indonesia
2United Graduate School of Agricultural Science, Gifu University, Yanagido 501-1193, Gifu Prefecture, Japan

E-mail: muh_makky@yahoo.com

Abstract. Paddy grains moisture content (MC) strongly correlated to the physical properties of rice after being milled. Incorrect MC will cause higher percentage of broken rice and prompts the grains to be more fragile. In general, paddy grains with 13 – 14% MC are ideal for post-harvest processing. The objective of this study is to measure the MC of intact paddy grain from CV. Bakwan by means of non-destructive evaluation using NIR spectral assessment. Paddy grains samples with identical MC were put into 30 mm tube glass and measured using NIR spectrophotometer. The electromagnetic radiation absorbance under consideration upon spectral measurement fell between 1000 and 2500 nm. The grains’ actual MC was then measured by primary method, based on weight measurement i.e. oven method. In this study, the spectral data of the grains was then processed by means of Principal Component Analysis (PCA) before correlated with its MCs by Partial Least Square (PLS) method. The model calibration obtained correlation (r) of 0.983 and RMSEC of 1.684. Moreover, model validation produced correlation (r) of 0.973, RMSEP of 2.095, and bias of 0.2, indicating that the MC of paddy grains can be precisely identified by non-destructive evaluation using spectral analysis.

1. Introduction
Rice is the main staple food and commodities in Indonesia. It is consumed mainly in the middle and western part of the country with more than 45 million tons of rice annual consumption nationwide [1]. As the main crops and food, rice become the crucial factor for ensuring food security and independence of this young republic.

Directly after being harvested, the paddy were subsequently reaped and passed through several processes including stacking, handling, threshing, cleaning, and hauling [2] before stored or consumed. Before stored, the grain’s moisture content (MC) should be reduced to a safe level, viz. 9 to 14% wet based (wb) [3] through natural or mechanical drying. The final edible white or brown kernel depends on the phases of the milling processes [2]. High MC will increase the product loss upon milling due to kernel break up or damage [4]. Normally, the MC of grains should be maintained at less than 14% before being milled into rice [3]. Beside losses, the MC of grain dictates the quality of rice produced after milling [5]. The percentage of kernel with cracks or damage correlates with its MC [6]. Higher MC increase yield losses and consequently reduces its retail price [7, 8].

Accurate measurement of the MC of the grains is commonly performed by means of weight methods [9, 10]. This method uses random samples from the bulk, and due to the huge resources
required it only considered as representative of the overall grains. On the other hand, many non-destructive evaluation methods for food and other agricultural products had been established. These methods offer better accuracy, lower operating costs, more rapid measurements, and repeatability when used for determining the internal properties [11] of the samples. Various methods had successfully employed to determine the features of oil palm fruits [12 – 21], tomato leaf [22], citrus [23], star fruits [24], pomegranate [25], and other agricultural commodities [26 – 28], including grains [29]. Grains properties such as starch [30], protein [31], mineral [32], and MC [29] can be determined non-destructively, by means of machine vision and spectroscopy. The MC influences the rice spectrum properties when observed by near-infrared (NIR) electromagnetic radiation [33]. The radiation is most notable when the reflection of incandescent light from the probe falls between 400 nm and 1050 nm. At a certain frequency, the MC causes variation in the absorption [33]. This phenomenon can be used to correlate the MC of rice kernel with its optical properties.

Water absorbs wide range of electromagnetic radiation and the nature is heavily influenced by the condition and the state of the water. In liquid form, water absorbs electromagnetic radiation at wavelength of 698, 718, 810, and 935 nm [34 – 38], which is observable when measured using VIS/NIR spectrophotometer. According to these facts, the water (MC) inside the paddy grain is possible to be evaluated nondestructively using Vis/NIR spectroscopy.

The rate of electromagnetic radiation absorption by grains can be modeled to its MC using statistical analysis. Different established model such as the Multiple Linear Regressions (MLR) [29] and Partial Least Square Regression (PLSR) [33] had been used for MC predictions in food products. In addition, Principal Component Analysis (PCA) [19] can be used to reduce the variables for explaining the correlation. While each model individually may be used for model creation, the combination of two methods, in particular the PCA and PLS, offer better accuracy and more acceptable results. Employing these methods for evaluating MC of the paddy grain will explain its internal properties and constituent without necessarily damaged the samples.

The objective of this study is to establish model for MC determination of the paddy grains from CV. Bakwan. The model was created based on the spectral data of the samples when measured using NIR spectrophotometer. The absorption data were then subsequently correlated with the samples’ MC, as measured by primary method. The model was established using PCA and PLS statistical analyses. The calibrated model was then verified for predicting the MC of the remaining grains samples. The results will offer non-destructive evaluation nature, while the performance of the model will be evaluated by means of Receiver Operating Characteristic (ROC).

2. Materials and methods
In this study, the samples’ MC is determined through electromagnetic radiation absorbance, measured using NIR spectrophotometer. 10 kg paddy grains of CV. Bakwan were harvested by manual picking from the stem. The unhusked grains were cleaned and placed into 30 ml glass tubes. The samples were compacted by placing the tubes on vortex for 10 seconds, at 80% speed, to minimized space between grains [29]. More grains were added and procedure repeated until the tubes were fully filled.

The sample tubes were then placed into NIR spectrophotometer and their spectral absorbances (1000 – 2500 nm) were recorded at normal scan speed using average of 10 scans method. The absorbance measurements were taken at two points by rotating the tube to 180° and the average of the two data were calculated. Recorded signals were pre-processed with spectroscopy software (Unscrambler® X, Camo, USA), and all measurements were replicated five times to minimize errors.

The actual samples’ MC was determined by primary method (oven) at 130 °C. The MC was calculated on wet basis. The MC measurements were replicated five times to minimize error. All measurements were following Makky et al. [29].

The absorbance data was then analyzed using engineering statistical software (Unscrambler® X, Camo, USA) to correlate with its MC. PCA method was used to identify significant correlation between certain electromagnetic radiation wavelengths with sample’s MC. Selected principal
component(s) (PCs) from the analyses were then considered as determining factor. A model to
determine \( MC \) of the samples according to its electromagnetic radiation absorbance was developed
using PLS. The electromagnetic radiation absorbance was considered as covariates in the analyses, the
samples \( MC \) as dependent variables, and the PCs from PCA analysis as factors. Data was partitioned
into two groups; the first 2/3 of data was used to train the model, and the rest of 1/3 data was used to
validate the model [19]. Model performance was analyzed using Sum of Squares Error (SSE), Relative
Error (RE), and ROC Curve [16, 17, 19].

3. Results and discussion
In this study, determination of samples’ \( MC \) was done by primary (oven) method. The sample initial
\( MC \) was measured at 11.8, 15.7, 19.5, 21.9, 24, 26.5, and 27.5% wet based (wb). The difference of \( MC \)
is caused by the variation on the weather and harvest condition upon samples collection, as well as
post-harvest handling. Raw spectral measurements data showed that samples with different \( MC \)
produce different absorbance value when measured by the spectrophotometer (figure 1). In correlation
with the \( MC \), the absorbance value of the samples increased along with the \( MC \) level. The absorbance
is influenced by sample \( MC \) throughout the electromagnetic radiation recording, between 1000 – 2500
nm. The differences were most notably in NIR spectrum of 1150 – 1800 nm (figure 1). In other
spectrum range, the difference of absorbance of the electromagnetic radiation of the samples is
observable, though the differences were minor.

The absorbance data of the sample were adjusted-normalized to set the value scaled from 0 to 1.
PCA analysis was done to distinguish which wavelengths significantly differ according to sample \( MC \).
Using KMO and Bartlett’s test of sphericity, the principal components (PC) were extracted by
determining the Eigen values greater than 1 through limited number of iterations until all spectral data
reached convergence and all data (100%) were explained. To better distinguish each principal
component, a rotation technique was applied based on Varimax method [39], in which each principal
component will be less dependent towards each other. Three PCs were obtained from the extraction
which explained more than 99% of absorbance data. Of the three PCs obtained from Varimax
extraction method using Eigen values greater than 1, the first two PCs, which explained the majority
of the electromagnetic radiation absorbance data from the samples, are presented in figure 2.

![Figure 1. Electromagnetic radiation absorbance of CV. Bakwan paddy grains as measured by NIR spectrophotometer.](image-url)
In figure 2, the majority of the first two $PC$, as plotted, placed in separated quadrants. The differences were corresponded to the samples’ $MC$ level. All two $PC$s explained 97% of variance from the extraction sums of squared loadings of the absorbance data. Samples with 11.8, 15.7, and 19.5% $MC$ occupied similar region, while samples with $MC$ of 21.9, 26.5, and 27.5% occupied similar quadrant on other section. Only samples with $MC$ of 24% are located separately from the others in the diagram. Although the $PC$s show distinction between samples group, the PCA alone cannot be used to explain which wavelength corresponds the most to the sample $MC$. The first $PC$ according to its coefficient value for every wavelength under consideration is presented in figure 3. From the figure 3, eight distinguished wavelengths were determined from the analysis. The selected peaks are 1013 nm, 1050 nm, 1115 nm, 1147 nm, 1257 nm, 1457 nm, 1523 nm, and 1859 nm. Nonetheless, the correlation to the $MC$ of the samples cannot be established directly. It is the reason why while PCA results suggest eight absorption peaks of electromagnetic radiation wavelengths for explaining sample’s $MC$, the absorbance data of the corresponding wavelengths cannot be directly calculated to predict the $MC$ in the samples.

In order to perform the prediction, a PLS analysis was employed to develop the model based on the absorbance properties of the samples. Using statistical engineering software, the model was developed by assigning the spectral absorbance data of the samples as covariate, and rescaled the data using adjusted normalization. Samples $MC$ was assigned as dependent variable for training the model. A hyperbolic tangent was used to activate the artificial grouping in this analysis, which is not limited to several groups. Using batch training type in the software to scale conjugated gradient of optimization algorithm, the initial calibration and validation model results are explained in the figure 4a and 4b respectively.

Figure 4 shows that the samples’ $MC$ can be accurately predicted by evaluating the absorbance of the selected wavelengths. For model calibration and validation, the $r^2$ is 0.966 and 0.950 respectively, while the $RMSEC$ and $RMSEP$ is 1.684 and 2.095 respectively. The results also showed that samples with lower $MC$ can be predicted more accurately by model, as compared to the samples with higher $MC$. Higher residual of samples with more $MC$ indicated that while the model result is still within acceptable value, more errors will occur if the samples contain more moisture above normal.
Figure 3. Selected electromagnetic radiation wavelengths according to the PCA analysis of the samples’ absorbance.

This phenomenon may relate to the amount and phase of moisture in the samples, and the properties of the electromagnetic radiation itself. The absorption of electromagnetic radiation by the water can occur in the wide range of regions, and is heavily influenced by the state of the water itself, especially in the liquid phase [35 – 38]. In the UV and visible band, the absorption of electromagnetic radiation by water is very weak, in particular when the water present in liquid phase and no scattering
effect produced during measurements by spectrophotometer [35]. In contrast, water absorption in organic material is stronger in the NIR band [33, 40].

Furthermore, the reason behind the bias of the models may correlate to the variance of physical properties of the grains, in which the model cannot exactly explain. Chemical properties, such as total protein content and composition of sugars or carbohydrates in the grains could influence this analysis. Therefore, the regression analysis of the model from the samples should incorporate these features as well. Nonetheless, by computing the relative error of the model, the deviation of prediction can be reduced. This study produces better prediction model for MC determination compared to previous finding [29, 41].

In previous studies [40, 42, 43], the chemical contents of paddy grains influence certain wavelength absorption spectrum. Furthermore, wavelengths selection from the developed model suggests that the MC of paddy grain may influence the electromagnetic radiation absorption by protein in the samples. Nonetheless, no previous results indicated how strong the influence was [40, 41]. In the developed model, the preference of electromagnetic radiation within the NIR range for MC determination in paddy grains samples was in agreement to other studies. Chang et al. [41] considered three wavelengths for the absorption observation: 785, 910, and 1020 nm, for third harmonic generation, C–H stretching vibration, and second octave N–H stretching vibration respectively. Other finding [44] proposed 990 nm wavelength to observe O–H stretching vibration at the second octave. Nevertheless, both results were used to identify the protein and amylase content of rice grain while in this study, only MC of samples is considered, regardless of the amylase and proteins content of the samples.

Beside the coefficient of correlation, the model performance was evaluated by ROC. The ROC curve explained the trade-off between sensitivity and specificity. Higher sensitivity will at the same time lower the specificity of the model. The model performance as plotted in ROC curve (figure 5) is significantly higher than by chance (diagonal line), indicating high accuracy from the test results. The plot shows that the developed model obtained high true positive rate against the false positive rate for the different possible cut-points. The area under the ROC curve indicates the test accuracy for the model. The developed model, according to ROC curve, resulted in high acceptability results with high accuracy with area under the curve of 0.847.
Figure 4. (Top) PLS model calibration for samples’ MC determination according to the absorption of electromagnetic radiation in NIR region. (Below) PLS model validation for samples’ MC determination according to the absorption of electromagnetic radiation in NIR region.

Figure 5. Model performance evaluation using ROC curve.

4. Conclusions
In this study, a nondestructive evaluation method was developed to model the MC of paddy grains from CV. Bakwan. The model calculates the absorbance value of the electromagnetic radiation from
the samples and correlated the data with samples’ MC using PLS analysis. The electromagnetic radiation for the measurements is within the NIR spectrum (1000 – 2500 nm). Upon validation, the model obtained $r^2$ of 0.950 with RMSEP of 2.095. In conclusion, the model produces high accuracy and sensitivity, although some residuals were observed in samples with higher MC. This study offers a more accurate, reliable and practical solution for nondestructive evaluation of MC in paddy grains, a crucial information to prevent losses upon milling.

Acknowledgments
The authors express their gratitude to Directorate General of Higher Education (DIKTI) and Universitas Andalas for providing the fund for this research through “Klaster Riset Guru Besar” Scheme, contract number. 65/UN.16.17/PP.HGB/LPPM/2017 signed on 21 March 2017.

References
[1] BPS 2015 *Official Statistical data: Paddy production in Indonesia in every province (1993-2015)* (Jakarta: Statistics Indonesia)
[2] Milling http://www.knowledgebank.irri.org/step-by-step-production/postharvest/milling
[3] Measuring moisture content in harvesting http://www.knowledgebank.irri.org/step-by-step-production/postharvest/harvesting/measuring-moisture-content
[4] Thompson J F and Mutters R G 2006 *Trans. ASABE* 49 435
[5] Siebenmorgen T J, Bautista R C and Counce P A 2007 *Appl. Eng. Agric.* 23 517
[6] Grain moisture content and grain quality http://www.knowledgebank.irri.org/step-by-step-production/postharvest/drying/drying-basics/grain-moisture-content-and-grain-quality
[7] Bautista R C and Siebenmorgen T J 2005 *Appl. Eng. Agric.* 21 637
[8] Bautista R C, Siebenmorgen T J and Mauromoustakos A 2009 *Trans. ASABE* 52 1611
[9] Methods for measuring moisture content http://www.knowledgebank.irri.org/step-by-step-production/postharvest/milling/milling-and-quality/measuring-moisture-content-in-milling
[10] IRRI moisture meter http://www.knowledgebank.irri.org/step-by-step-production/postharvest/drying/drying-basics/methods-in-measuring-moisture-content-for-drying/measuring-moisture-content-in-drying/irri-moisture-meter
[11] Manickavasagan A and Jayasuriya H 2014 *Imaging with Electromagnetic Spectrum: Applications in Food and Agriculture* 1st ed (London: Springer-Verlag)
[12] Cherie D, Herodian S, Ahmad U, Mandang T and Makky M 2015 *International Journal on Advanced Science, Engineering and Information Technology* 5 255
[13] Cherie D, Herodian S, Mandang T and Ahmad U 2015 *International Journal on Advanced Science, Engineering, and Information Technology* 5 314
[14] Makky M, Herodian S, Cherie D, Ahmad U and Mandang T. 2012 *Proc. Bogor Agricultural University Research Dissemination Seminar (Bogor)* (Bogor: Bogor Agricultural University) p 1
[15] Makky M, Soni P and Salokhe V M 2012 *Proc. Asian Forum of 2012 CSAM (Chinese Society for Agricultural Machinery) International Academic Annual Meeting (Hangzhou)* (Hangzhou: Chinese Society for Agricultural Machinery) p 1
[16] Makky M, Soni P and Salokhe V M 2014 *Int. Agrophys.* 28 319
[17] Makky M and Soni P 2013 *Comput. Electron. Agric.* 93 129
[18] Makky M and Soni P 2013 *International Journal on Advanced Science, Engineering and Information Technology* 3 1
[19] Makky M and Soni P 2014 *J. Food Eng.* 120 248
[20] Makky M 2016 *Agric. Agric. Sci. Procedia* 9 230
[21] Makky M 2016 *International Food Research Journal* 23 S81
[22] Makky M 2016 *International Journal on Advanced Science, Engineering and Information Technology* 6 349
[23] Aleixos N, Blasco J, Navarrón F and Moltó E 2002 *Comput. Electron. Agric.* 33 121
[24] Siti Mazlina M K, Abdul Ghani L A, Nur ‘Aliaa A R, Siti Aslina H and Rozita O 2008 Pertanika J. Sci. & Technol. 16 1
[25] Blasco J, Cubero S, Gómez-Sanchís J, Mira P, Moltó E 2009 J. Food Eng. 90 27
[26] Kondo N. 2009 Sens. Instrum. Food Qual. Saf. 3 81
[27] Jin H, Li L and Cheng J 2015 J. Food Anal. Methods 8 25242
[28] Williams E J and Drexler J S 1981 Peanut Science 8 134
[29] Makky M, Santosa, Putri R E and Nakano K 2017 AIP Conf. Proc. 1813 020002
[30] Ali A, Wani T A, Wani I A, Masoodi F A 2016 Journal of the Saudi Society of Agricultural Sciences 15 75
[31] Shih F F and Daigle K W 2000 J. Am. Oil Chem. Soc. 77 885
[32] Heman A and Hsieh C-L 2016 Engineering in Agriculture, Environment and Food 9 280–290
[33] Duarte F J 2016 Tunable Laser Applications 3rd ed. (New York: CRC Press)
[34] Pope R M and Fry E S 1997 Appl. Opt. 36 8710
[35] Kou L, Labrie D and Chylek P 1993 Appl. Opt. 32 3531
[36] Jonasz M 2007 Absorption coefficient of water: data sources http://www.tpdsci.com/Tpc/AbsCfOfWaterDat.php
[37] Wozniak B and Dera J 2007 Light Absorption in Sea Water. Atmospheric and Oceanographic Sciences Library (New York: Springer-Verlag)
[38] Kaiser H F 1958 Psychometrika 23 187
[39] Natsuga M and Kawamura S 2006 Trans. ASABE 49 1069
[40] Chang R K, Zhang W Y, Cui J, Wang Y H, Wei Y, and Liu Y 2009 Computer and Computing Technologies in Agriculture III 523-531
[41] Layne E 1957 Spectrophotometric and turbidimetric methods for measuring proteins Methods in Enzymology 3 eds P S Colowick and N O Kaplan (New York: Academic Press, Inc.) pp. 447-455
[42] Simonian M H and Smith J A 2006 Spectrophotometric and colorimetric determination of protein concentration Current Protocols in Molecular Biology (New York: John Wiley & Sons, Inc.) 76:10.1A:10.1.1
[43] McClure W F 2006 Introduction Near-Infrared Spectroscopy in Food Science and Technology eds Y Ozaki, W F McClure and A A Christy (New Jersey: John Wiley & Sons, Inc.)