Journal of Cereal Science 89 (2019) 102795

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Adaptation of visible and short wave Near Infrared (VIS-SW-NIR) common PLS model for quantifying paddy hardness

M.G.G. Awanthi[a,*], B.M.S. Jinendra[a], S.B. Navaratne[b], C.M. Navaratne[a]

[a] Department of Agricultural Engineering, Faculty of Agriculture, University of Ruhuna, Sri Lanka
[b] Department of Food Science and Technology, Faculty of Applied Science, University of Jayawardhenepura, Sri Lanka

ARTICLE INFO

Keywords:
Hardness
Near infrared
Non destructive
Paddy
Spectroscopy

ABSTRACT

Hardness of paddy is an indicator of preserved quality of rice as it maintains the integrity of grain to avoid falling-off during milling. Most of the prevailing methods carter to determine the hardness are destructive and time consuming. Though, use of Near Infrared (NIR) is nondestructive, its conventional approach implies separate Partial Least Square (PLS) model for each new paddy type to be predicted successfully. Therefore, this study evaluates NIR spectroscopy (588–1091 nm) to be used as a common PLS model for quantifying a range of paddy varieties and processing conditions in rapid and non-destructive mode. The acquired NIR spectra were used to construct PLS calibration models by using Pirouette 4.5 software. As a result, the lowest standard error of validation (SEV = 1.711) and highest correlation coefficient of validation (R^2 = 0.936) were obtained. The research has successfully demonstrated that the potential possibility of NIR spectroscopy to be used as a common PLS model for quantifying hardness of paddy from new varieties and conditions.

1. Introduction

Quality of rice depends upon large number of factors in a paddy production and processing line-up from farm to table. Particularly, these factors are ranked based on their importance to the end users (Cuevas et al., 2016). The most important parameter related to global human nutrition is the milling quality which can be expressed as the quantities of whole and broken kernels that remain after hulling and milling paddy rice (Lyman et al., 2013). Furthermore, the market demand of processed paddy with broken grain is much lesser than that of whole grains due to poor quality attributes of rice after cooking (Lai et al., 1983). Therefore, grains should be resistance to breakage for reducing broken rice yield (Jongkaewwattana and Geng, 2001). It can be measured as hardness of grains which is the foremost determinant of the grain quality associated with milling recovery (Juliano, 1979). Hardness can define as the maximum force applied on a grain at any time during the first cycle of compression (Harris, 2012). It was determined on rough, brown and milled rice forms using different methods and it is important to many facets of the rice industry (Pomeranz and Webb, 1985). Besides processing and milling grain breakage, it is also an essential parameter when dealing with storage changes and aging, drying and handling, kernel appearance and translucency and resistance to insects (Webb et al., 1986). In addition, hardness is a key parameter to let the consumer to decide the quality of parboiled rice in the market because, texture of the rice change significantly due to gelatinization and retro-gradation of starch during parboiling (Poritosh et al., 2006).

A standardized method has not been recognized to measure grain hardness because each method showed particular drawbacks related to the equipment and measurement principles are being used. Hardness can be tested as single kernel or bulk sample. The force required to indent single kernel can be measured using different hardness testers including Leitz Minilsad Hardness Tester, Barcol Impressor and Miag Microhardness Tester (Bakhella et al., 1992). Similarly, the force needed to crush the grain can be measured using Instron® universal testing machine and continuous automated single-kernel hardness tester (Lai et al., 1985). Limitation of using these penetrometers type single gain testing methods except time consuming is that hardness value can be highly varied due to the level of moisture content and uniformity among the grains. The laser light-scattering method is another method for testing single kernel hardness. It measures the mean volume diameter of grains. However, reports revealed that overlapping between different grain classes when measuring the hardness of wheat grains using laser light-scattering methods (Bakhella et al., 1992).

Parameters used to measure hardness by using bulk samples testing technique are the grinding resistance, grinding time, volume of milled material and the ratio of coarse to fine particles either by volume or by weight (Cauvain and Young, 2009). Work required to grind the rice

* Corresponding author.
E-mail address: mggawanthi@gmail.com (M.G.G. Awanthi).

https://doi.org/10.1016/j.jcs.2019.102795
Received 12 February 2019; Received in revised form 1 June 2019; Accepted 28 June 2019
Available online 29 June 2019
0733-5210/

grains can be measured in grinding resistance technique using Stevensen hardness tester. Miller et al. (1984) determined the hardness of rice indirectly by measuring the time spend to grind 4 g of rice with the Brabender automatic micro hardness tester. Although these two methods were not affected by kernel size, grinding time method was affected by temperature (Miller et al., l98l) while grinding resistance method was influenced by moisture content (Stenvert, 1974). Based on sieving and weighing of ground grains, hardness can be expressed as pearling index or particle size index. It was found that pearling index values were influenced by both genetic and environmental factors and it couldn't rank different grain classes due to bran properties. Even though particle size index technique was relatively quick, reliable and less expensive, major disadvantage of both these techniques is the lack of complete standardization process to evaluate grain hardness (Bakhella et al., 1992).

It is clear that most of the prevailing methods for detecting paddy hardness are destructive, time consuming and not standardized. To overcome these drawbacks, Near Infrared (NIR) spectroscopic technique can be used as a powerful tool to assess grain hardness (Bakhella et al., 1992). Previous research found that NIR spectroscopy of wheat and rice grain has a number of well-known advantages namely speed of analysis, no sample preparation required, environmental friendly and non-destructive nature, low cost per test and concurrent analysis of multiple constituents (Srivastava et al., 2018; Mishra et al., 2018). Despite its advantages, this technique still struggles with initial cost of the instrument. However, NIR technique can be successfully employed in quantifying various physiochemical properties of rice such as, whiteness, milling degree, cooking quality, protein, amylase, moisture content, total phenolic contents and antioxidant capacity in rapid and non-destructive mode (Zhang et al., 2008). Recently, Srivastava et al. (2018) have successfully predicted the quality of the stored rice in terms of protein, ash, fat, fiber, carbohydrate, hardness, moisture content, amylose content, 1000 kernel weight, cooking time, solubility weight loss, bulk, and true density by using Fourier Transform infrared (FTNIR) spectroscopy as well. Therefore, high cost of the instrument is not an obstruction as it can evaluate large number of quality parameters of rice rather than hardness.

Since absorption of NIR energy increase in overlapping overtones and combinations of rotation and vibration of different chemical bonds including C–H, O–H and N–H, chemical composition of the organic molecules can be determined using NIR (Williams, 2003). Similarly, hardness of rice could be influenced by composition of several molecules like amylose, amylopectin, compact arrangement of starch granules, protein and lipid (Zhou et al., 2002). Hence, there should be a possibility to detect hardness of rice using NIR technique. According to this principle, hardness was determined in rough, brown and milled rice forms by using NIR spectroscopy at 1680 nm (Webb et al., 1986). It was found that NIR reflectance and hardness value measured using particle size index was significantly correlated in all rice forms. Similarly, simple models by PLS from the NIRS spectra for predicting hardness of brown rice was developed under a wavelength range of 1400–2400 nm. Its correlation coefficient was ranging from 0.39 to 0.79 (Siriphollakul et al., 2013). It is noticeable that accuracy of the previously reported results regarding to the hardness of paddy is not high enough. Therefore, it should be developed a strong model based on NIR spectroscopy for accurate evaluation of paddy hardness. Furthermore, a recent study has obtained the best model for hardness of infested and fresh stored rice grains with lowest root mean square error of cross validation (RMSECV) values 0.02 and maximum correlation coefficient (R^2) 99.82 using FTNIR spectroscopic method within a wave number range of 12,000–4000 cm^{-1} (Srivastava et al., 2018). However, neither of those reports having the information for using their PLS model performance for quantifying wide range of other new types of paddy seed samples and nor attempts for evaluating their model performance for samples prepared by different processing techniques.

Although VIS-SW-NIR spectroscopy joined with multivariate analysis are used to measure hardness of grains, no reports on their adaptation to a wide range of new sample types. On the other hand, well developed model should be able to predict hardness of wide range of rice grain such as diverse rice varieties, parboiled rice, and rice stored under different conditions. Otherwise, there must be a time consuming conventional NIR model building procedure to take place for each and every single variety type. That does not bring the NIR for its full advantage for material analysis. Therefore, the main objective of this study was to develop a universal VIS-SW-NIR spectroscopic model to be able to quantify the hardness of wide range of new paddy sample types that those varieties have not been used in the model building process. The samples were used in this study were obtained from two varieties and further maintained them under various perturbed conditions to accumulate wide range of hardness attributes to be in the PLS model from which we expect to compatible with the hardness attributes of wide range of prospective new samples from new varieties to be predicted.

2. Methodology

2.1. Sample preparation

In the conventional approach, a successful model building of NIR required a wide range of spectral database available in the model. This is therefore a labor intensive and time consuming task. However, the required amount of diversity to be in the data set for a successful model is depending upon the sample type. The spectroscopic attributes of the sample i.e. molecular structure and physical properties of particle that make the changes of light matter interaction will be the leading. Therefore, in this research**,** the model building was done using raw rice variety AT 362 for calibration and then the model validation was done using Kuruluthuda, At-353 and At-401 varieties. Then, the model was further upgraded by adding wide range of data types to get it up to global scale model i.e. adding the spectra of different paddy varieties, then parboiled paddy of different varieties, the raw and parboiled paddy kept in a different temperature levels, etc. Therefore, different types of paddy samples have been prepared for the experiment at different stages of the model development as follows.

2.1.1. Preparing of raw paddy samples

AT-362 and Kuruluthuda rice varieties were selected for the experiment at different stages of the model development. Fresh paddy was harvested in *Maha* season (2016/2017), then cleaned and kept in dried in an open yard for sun drying until the moisture content reaches to 14%.

2.1.2. Preparing of the parboiled paddy samples

Portion of At-362 and Kuruluthuda paddy were parboiled by adapting following processing steps to obtain wide range of hardness values. Firstly, two portions of each variety were soaked in cold water for 48 h s while replacing existing water with fresh in every 12 h to avoid the development of microbes that may lead foul smell for the parboiled rice. The soaked water was drained off and steamed for about 4–6 min until few grains were split open over the surface of the paddy. The parboiled paddy was dried in an open sun drying yard as a thin layer (Thickness approximately 1 cm) until moisture content reaches to 18%. Thereafter, this paddy was heaped up and kept for a while (2–3 h) for uniform moisture diffusion with a view to prevent the case-hardening of the grains. Finally, paddy was spread over the sun drying yard again to get the final moisture content around 14%.

2.1.3. Preparation of paddy samples under different storage conditions

Since the hardness values represent a wide data range to develop a strong model, multiple variables were selected as given in Table 1. About 5 kg of Parboiled and raw paddy of each variety (At-362 and Kuruluthuda) were packed (48 bags) in Polysack® bags

Table 1
Variables and their levels of the experiment.

Variables	Levels (with 3 replicates)
Temperature levels	26, 30, 34 & 38 °C
Rice varieties	AT-362, Kuruluthuda
Processing technologies	Raw, Parboiled
Storage time	30,60,120,150 and 180 day

(0.45 m × 0.3 m) and stored at 26, 30, 34 and 38 °C in temperature-controlled chambers for 6 months. Samples were collected monthly as each treatment represent 3 replicates. Total number of collected samples were 288.

2.1.4. *Preparation of paddy samples under different moisture content*

At-353 and At-401 paddy varieties were soaked in water for 12 h and collected the samples at 2 day intervals to get range of hardens data for prediction sets. Total number of samples collected for the prediction set was 60.

2.2. *Data collection*

All individual paddy samples were first used to obtain the NIR spectra to be used in the model building process and prediction process and then the same samples were used to measure their hardiness values by the reference hardness test.

2.2.1. *Acquiring of NIR data*

Reflectance spectra from the wavelength range of 588 nm up to 1091 nm was acquired using handheld type NIR Spectrometer FQA-NIR Gun (588–1091 Shizuoka Shibuya Seiki, Hamamatsu, Japan). A black rubber probe extension was fixed into the sensor head of the NIR instrument to focus the radiation to the samples. This avoids the samples been exposed to possible outside light conditions that may not be constant all the time. The detail dimensions of the probe extension including its improved performance has been previously published (Jinendra et al., 2010). A custom made plastic cup with 4.5 cm in diameter and 1 cm in depth was constructed to hold the paddy samples. The plastic sample cup was filled with paddy sample and the excess sample surface was trimmed with a spatula (Plate 1[a]). A quartz glass screen attached to a plastic plate was kept on the top of the paddy samples to place the spectrometer probe securely and similarly at each spectra acquisition attempt (Plate 1[b]). Acquisition of all the NIR spectral data from all the paddy samples throughout the study were executed under the set optimum sensor exposer integration time of 150 ms (Plate 1[c]). Four spectra were acquired from four directions in each sample to get more variation as marked in plate 1[b]. In the model building process, the same average hardness value taken from reference hardness test was similarly assigned to all 4 spectra belonging to that particular sample in the data table.

2.2.2. *Measuring reference data*

Digital force gauge (Model-500B, SUNDOO, China) with external load cell sensor (SH–2N–500N) equipped in an electric vertical stand (Model-SJX-2KV, SUNDOO, China) was used to measure the hardness from paddy grains. A 10 mm diameter stainless steel V model convex probe at a cross head speed 150 mm/min was used to break a single grain. Serial output (RS- 232C) of the instrument was connected to the computer to display the test force curve and test process. Hardness was recorded as the peak force (N) applied on a grain at any time during the first cycle of the force curve.

The test was repeated for 20 individual paddy grains from the same sample lot. Thereafter, extreme and unmatched odd values were removed and calculated the mean of the selected hardness value of a sample.

2.3. *Data analysis*

2.3.1. *Preprocessing of the spectra*

Spectral data at initial end wavelength regions (588–613 nm and 1070–1091 nm) were excluded due to low signal-to-noise ratios in the instrument. Then, spectral bands between 615 nm and 1068 nm were used for analyses. The acquired spectra were first evaluated by Principal Component Analysis (PCA) to remove possible outliers. A substantial baseline variation between the spectra was presented even after removing the outliers. This baseline effect is usually presented due to sample presentation variations and environmental noise which is undesirable for data analysis. Thus, different combinations of algorithm options and math transformations were applied to the outlier removed spectra to reduce the baseline effect. All chemo-metric procedures were performed on commercially available software Pirouette (Pirouette ver. 4.5, Infometrix, Woodin-ville, WA).

2.3.2. *PLS calibration*

The spectra removing after the outliers as shown in Table 2 were applied in the Partial Least Square (PLS) calibration models in Pirouette 4.5 software and models were developed to predict the hardness of the paddy. All the model configurations including number of factors, data preprocessing, spectra mathematical transformations, derivative points etc. were applied in their various possible combinations to be in the PLS algorithm. The model performance of each model parameter combination was observed by means of standard error of calibration (SEC) and coefficient of correlation (R^2) the results using the trial and error method.

2.3.3. *Model validation*

To validate the constructed PLS model, validation data sets were prepared using collected NIR and reference data according to the requirement of each model. After developing model to predict paddy hardness for same variety, new validation set was prepared with 60 paddy samples from two other paddy varieties (At-353 and At-401). Two hundred and forty NIR spectra were recorded from 60 paddy samples as 4 spectra per sample (one spectrum per one direction). Then, model set and validation set were used to evaluate the model. Validation model performance was assessed in terms of coefficient of determination (R^2) and standard error of validation (SEV).

3. Result and discussion

3.1. *Results of primary NIR model for same variety hardness prediction*

Step 1; Model set variety AT- 362, prediction set variety AT- 362,

The NIR spectra obtained from 30 samples of raw paddy variety AT-362, as four spectra per sample (n = 120) were used to prepare the primary model set. Results, obtained from predicting the spectra taken from an independent set of the same variety spectra (n = 40) under different combinations of model configurations in the model algorithm options i.e. preprocessing and math transformations in PLS model are shown in Table 3.

The lowest standard error of validation (SEV = 0.918) and highest coefficient of calibration (R^2 = 0.921) value were given under the mean-centered preprocessing with align math transformation options. The result of first few attempt of the data analysis, revealed that the PLS models constructed for hardness prediction of the independent validation set of AT-362 spectra were fairly good as it records substantially high coefficient of correlation and low standard error of validation. (Fig. 1[a]).

However, previous reports revealed that possible association between hardness and different grain parameters which can be vary according to the variety including amylose, alkali spreading value, gelatinization temperature, protein and grain size and shape (Webb et al.,

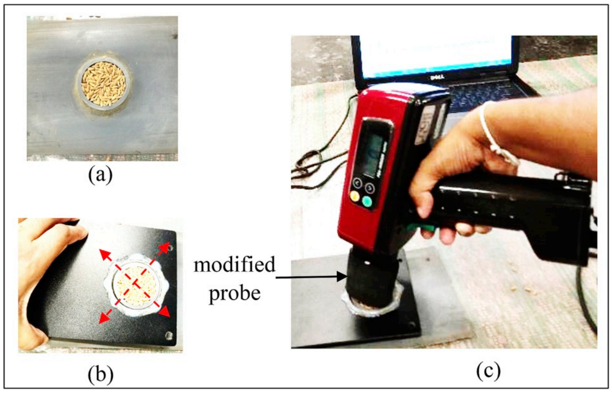

Plate 1. NIR Data acquisition.

1986; Safdar et al., 2009). Furthermore, Mir and Bosco (2013) revealed that the varietal difference in the hardness of rice grains is due to the differences in the compact arrangement of starch granules among different rice cultivars. Therefore, hardness can differ in different types of paddy variety. Then, next step was to validate the model with another variety to investigate whether validation results were performed well or not with derived model set.

3.2. Evaluation of primary model for new variety hardness prediction

Step 2; Model set variety AT- 362, prediction set variety Kuruluthuda

The NIR spectra (n = 120) obtained from 30 samples of raw paddy variety Kuruluthuda, were used in the prediction set to validate the model to assess its capabilities for a new variety to be predicted. The prediction set of spectra was evaluated by principal component analysis and remove the outliers before prediction. The PLS regression functions of measured versus predicted hardness of paddy under the best model configuration of mean-center preprocessing with align math transformation option are shown in Fig. 1(b).

The results of the model validated by new variety Kuruluthuda were found to be substantially inferior (SEV = 1.846, R^2 = 0.870) to the results of the previous primary model which is in predicting the same variety (SEV = 0.918 and R^2 = 0.921). As usually expected, this low results are due to the incompatibility of the diversity in data matrix in the two data sets corresponding to their physiochemical properties.

Moreover, after the validation of Kuruluthuda paddy, it was noticed that hardness values were not distributed normally in the population statistics and a range of hardness value was lack in the obtained distribution profiles as shown in Fig. 2(a). The missing hardness data ranges in Y fit of the PLS model for Kuruluthuda variety were between 14 and 16, 21–23 and 26–30N force range. In order to avoid the missing range and to be widen the data range, a combined PLS model was developed by both varieties. Moreover, the varieties were stored in in 4 different temperatures levels for 6 months and acquire spectra in consecutive time intervals to build a model with a wide range of hardness information available in the model for successful hardness prediction.

3.3. Development of wide range NIR model

Step 3; The model set with AT-362 and Kuruluthuda (X) four temperature levels (X) store for 6 months.

NIR spectra were recorded from 144 samples of AT-362 and Kuruluthuda stored at 26, 30, 34 and 38 °C temperatures for 6 months as four spectra per sample (n = 576) because changing storage temperature and time were positively correlated with the paddy hardness in this study. Therefore, the expected wide range of hardness data were obtained. When temperature is increasing, the main part of rice kernel consists of starch granules with semi-crystalline structure become greater due to decreasing moisture content in grains (Enevoldsen and Juliano, 1998). Therefore, resistance of kernel to breakage is increased. Furthermore, increasing storage temperature up to 37 °C enhanced the aging of the stored paddy (Pearce et al., 2001) which leads to increase hardness of the starch granules and limited the granule hydration and swelling due to starch protein interaction (Tulyathan and Leeharatanaluk, 2007). In addition, increased hardness of rice over time is associated with starch retrogradation process which convert the starch in to more stable crystalline form (Mutters and Thompson, 2009). We expected these physiochemical scenarios would help to accumulate reasonable hardness variation in our data matrix which will reasonably compatible with the sample spectra coming from prospective new variety samples. After removing possible outliers, obtained PLS regression function of measured versus predicted hardness of paddy under the mean center pre-processing with align math transformation option are shown in Fig. 1(c).

In this attempt, after accumulation of data, the statistic illustration revealed that considerable amount of missing data ranges observed in the previous model (Fig. 1(b)) were fulfilled in new prediction model (Fig. 1(c)) but, paddy hardness prediction was not improved (R^2 = 0.865 and SEV = 1.221) up to the expected level. It can be clearly identified that large number of hardness data were concentrated between 15 and 20N force range and missing hardness data were still observed between 25 and 30N force range in Y fit of the PLS model even this step. Therefore, in order to acquire comparatively higher hardness values, combined PLS model was developed with parboiled paddy as the paddy have shown substantial hardness increment after parboiling in the studied conducted in parallel during this study.

Step 4: Upgrading the combined model with parboiled paddy

Hardness of parboiled paddy is generally higher than raw paddy due to gelatinization and retro-gradation of starch during the soaking and steaming process (Poritosh et al., 2006). Since the missing data was

Table 2
Different types and number of samples used in model development in this study.

Step No	Process	Variables	No of samples	No of spectra
1	Primary model	Variety (AT-362)	30	120
2	Validation	Variety (Kuruluthuda)	30	120
3	Calibration	Varieties (AT-362, Kuruluthuda) Temperature levels (26, 30, 34,38 °C) Storage time (30,60,90,120,150,180 days)	144	576
4	Calibration	Varieties (AT-362, Kuruluthuda); Temperature levels (26, 30, 34,38 °C) Storage time (30,60,90,120,150,180 days) Processing techniques (Raw, Parboiled)	288	1152
5	Validation	Moisture content (initial and after 12hr soaking; 2,4,6,8,10-day interval) Varieties (AT-353 and AT-401)	60	240

Table 3
Prediction results of Paddy Hardness under PLS Model Configurations.

Math transformation	Preprocessing method									
	Auto Scale		Mean Center		Range scale		Variance scale		Pareto	
	SEC	R^2	SEC	R^2	SEC	R^2	SEC	R^2	SEC	R^2
Align	0.912	0.865	0.918	0.921	2.564	0.780	3.166	0.716	0.913	0.865
1st derivative	2.118	0.792	2.098	0.797	2.135	0.788	3.457	0.558	2.093	0.798
Log 10	2.178	0.779	2.181	0.778	2.497	0.706	2.198	0.774	2.180	0.778
2nd Derivative	0.909	0.835	1.936	0.830	1.932	0.831	3.195	0.616	1.913	0.835
MSC	2.119	0.792	2.148	0.786	2.140	0.787	2.139	0.787	2.133	0.789

distributing within comparatively high hardness values range, data were collected from both raw and parboiled paddy in AT-362 and Kuruluthuda stored at 26, 30, 34 and 38 °C for 6 months to increase the paddy hardness. The NIR spectra were recorded from 288 samples as four spectra per sample (n = 1152) for calibration set. After removing possible outliers, the obtained PLS regression function of measured versus predicted hardness of paddy under optimum model configuration options are shown in Fig. 1(d).

Result of this PLS analysis observed that best fitted model could be obtained with strong $R^2 = 0.946$ and low SEV = 1.204 to evaluate paddy hardness. There was no missing data range in the developed model as well. Data normally distributed as shown in Fig. 2(b) while data was ranging between 12.5 and 32.5N without showing any missing values.

3.4. Validation of the best fitted model

The developed best model was validated using new paddy samples in difference hardness ranges. At-353 and At-401 paddy varieties were soaked in water for 12 h and collected the samples at 2 day intervals to get range of hardness data for final validation set. Total number of samples collected for prediction set was 60 (n = 240)). Model validation with other 2 paddy varieties found to be yielded a better performance related to hardness between predicted values and measured

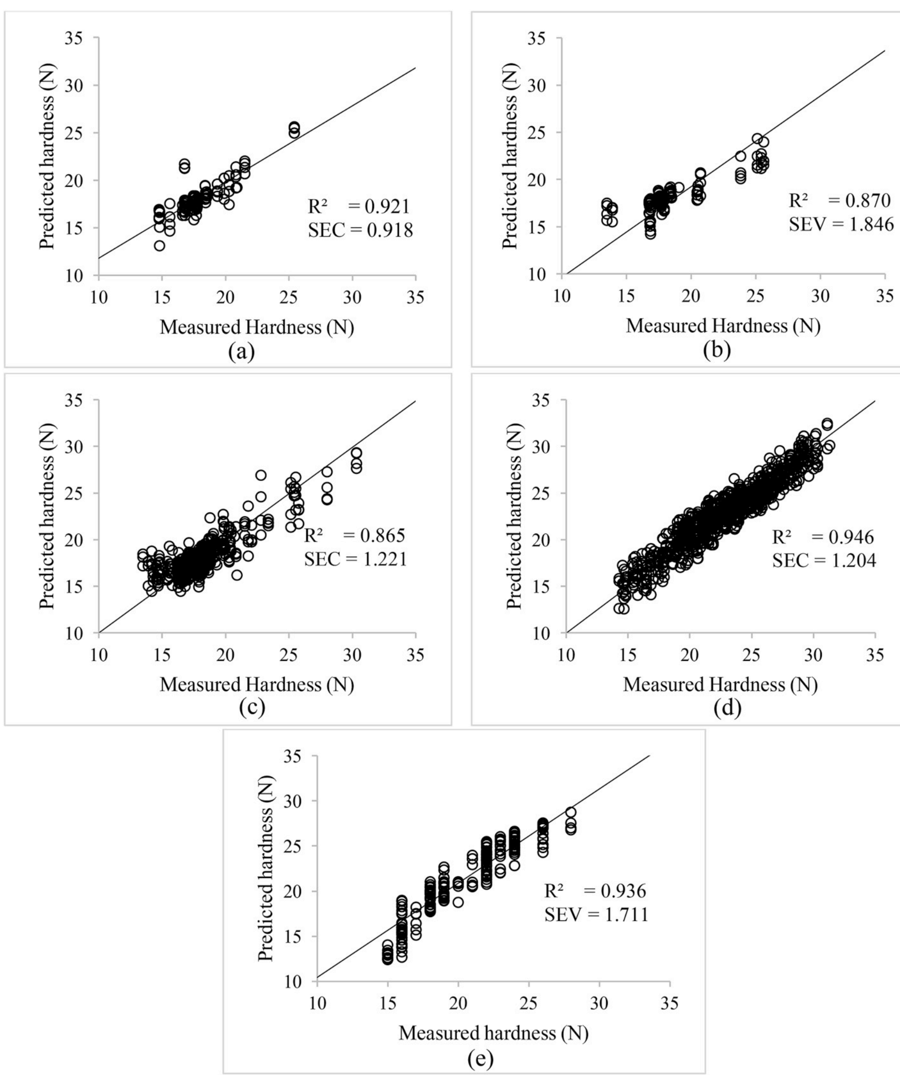

Fig. 1. PLS regression functions of predicted versus measured hardness in model development (a) Step 1; (b) Step 2; (c) Step 3; (d) Step 4; (e) Step 5.

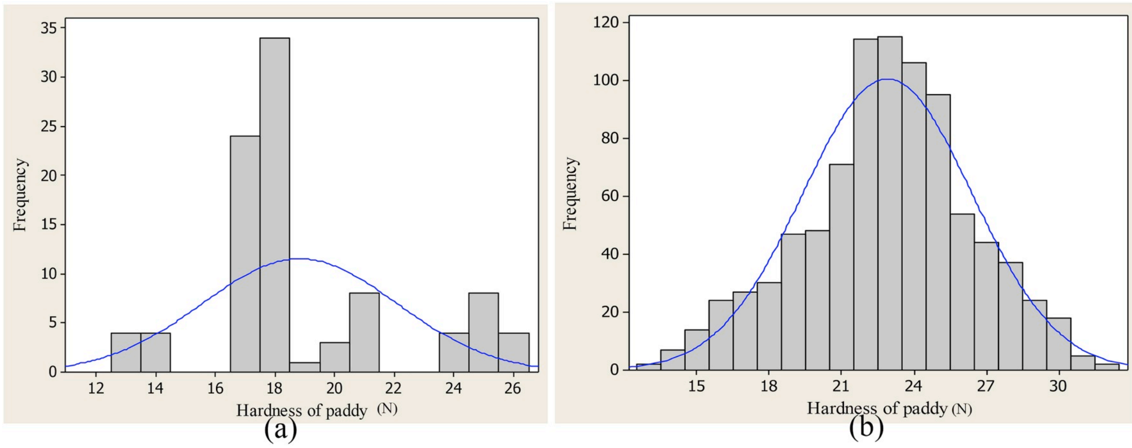

Fig. 2. Frequency of hardness data (as force in Newton) distribution of model development (a) Step 1 and (b) Step 2.

values (SEV = 1.711, R^2 = 0.936) using same pretreatments as highlighted in Figure (e). These results indicated that the paddy hardness can be rapidly and accurately predicted from NIR reflectance spectroscopy using PLS regression.

Similarly, Qingyun et al. (2007) developed a multiple linear regression model for detecting hardness of cooked rice using wavelengths of 540, 640, and 970 nm only and recorded R^2 = 0.59; SEC 0.32. Meullenet et al. (2002), and Champagne et al. (2001) also developed models for detecting hardness of cooked rice using wavelengths of 400–2500 nm and recorded R^2 values in validation were 0.59 and 0.67 respectively). However, Lapchareonsuk and Sirisomboon (2015) developed a successful model for detecting hardness of cooked rice using more similar wavelengths (600–1100 nm) which used in this experiment and reported R^2 and SEV were 0.842 and 0.364 respectively. Furthermore, similar model was developed by Siriphollakul et al. (2013) for predicting hardness of brown rice using wavelength range of 1400–2400 nm and R^2 and SEV were 0.777 and 0.185 respectively. However, our prediction values were better than that previously obtained values. It may be due to wider range of hardness values we accumulated by practicing multiple methodologies. The targeted best fitted regression line in the PLS algorithm could be obtained after only accumulating the wide range of spectral data for paddy hardens by following the new approach i.e. observing the regression line and population distribution for missing range and completing them by adding sample with induced variability by perturbation such as storing at elevated temperature, soaking, parboiling etc..

3.5. Important wavelength for detection of paddy hardness

The regression vector coefficients can be used to show the most important wavelength elaborated in the detection of hardness of paddy. In this experiment, both positive and negative regression coefficient values were obtained as shown in Fig. 3. The regression vector coefficients indicated that wave bands; 621, 669, 710, 799, 843, 1048, 1064 and 1066 nm seem to be of greatest importance in prediction of hardness.

It was found that absorption peak bands of NIR spectra for starch, protein, oil, cellulose and water were 1064, 1016, 928, 914 and 1044 nm respectively (Williams, 2003). Similarly, wavelengths showed peaks at around 1066, 1016, 928, 914 and 1044 nm in this study as highlighted on Fig. 3. Therefore, it can be proved that hardness of the paddy was affected by content of starch, protein, oil in the grain kernel and cellulose and water in the outer layer of husk. These substances including proteins comprise many N–H groups, fats and oils contain many C–H groups, water contains O–H groups, and so on. Moreover, protein, starch, oil and cellulose all contain –CH_2- groups and –OH– groups that make identification possible grain products (Williams, 2003). Lapchareonsuk and Sirisomboon (2015) found that peak of regression coefficients of textural qualities of cooked rice such as adhesiveness, hardness, and stickiness, were ranged between 935 and 990 nm. The peak at 938 nm shows the frequency characteristic of C–H stretch third overtone of CH_2 while peak at 990 nm is the vibration band of the O–H stretch second overtone of starch. In the same way, peak of regression coefficients for rice were noticed within that range (914, 927, 941, 953, 993 nm) in our study as well.

Furthermore, it can be clearly shown that the regression vector coefficients were more or less similarly distributed throughout the spectrum without concentrating to the visible region. Therefore, it can be concluded that the color of paddy grain which might be considered as a common character to be changed due to various reason have not

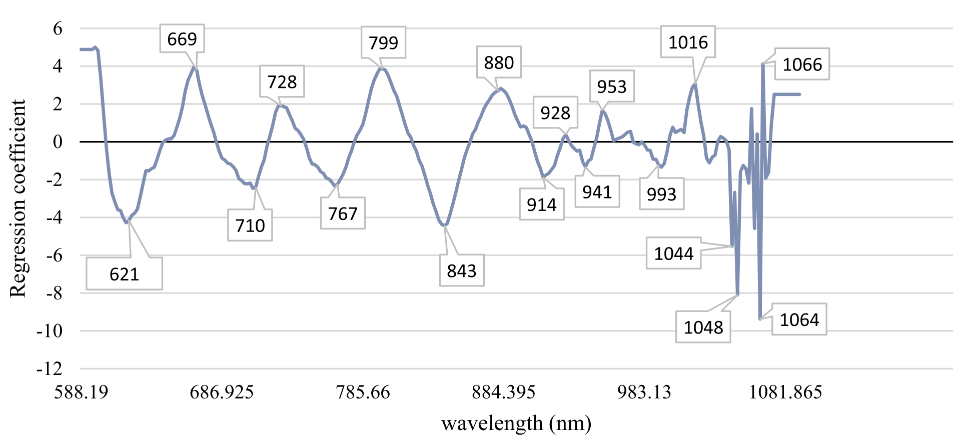

Fig. 3. Regression coefficient plot of optimum models for evaluation of hardness of paddy.

been in a prominent factor to quantify the hardness of paddy samples in the final model. This observation would be considered as an indication of reliability in the final outcome of this research.

4. Conclusions

In conventional INR approach, a successful prediction is perceived when the model and prediction sets are executed between uniform and similarity groups of materials. However, this research has successfully demonstrated that NIR can be expand its limits by accumulating wide range of variation in the spectra for the target Y variable by following a systematic approach i.e. observing the regression line and population distribution for missing range and completing them by adding sample with induced variability by perturbation such as storing at elevated temperature, soaking, parboiling etc. have been demonstrated giving substantial improvements to the prediction models leading wider acceptancy to the incoming new samples to be predicted.

The finally configured model has successfully predicted the spectra obtained from new independent varieties with the performance rates of SEV = 1.711 and coefficient of determination $R^2 = 0.936$ which could be considered as improved model performance to the existing prediction results with regards to the completely independent new variety prediction. This bring additional benefits to producers, marketers and users to evaluate the quality status range of market available paddy at whenever random sampling is needed for quality assurance and valuation without constructing separate model for each variety. This could be considered as an added meaning to the rapid nature of NIR as a tools where the users are prompt when expressing the merits of NIR as a rapid detection tool.

Acknowledgements

Authors gratefully acknowledge Shibuya Seiki Co Ltd, japan for providing the NIR spectrometers for this research. We extend our gratitude to Mr. Thusitha Saman Bandara for the technical assistance provided in installing and maintaining the incubator setup for the experiments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcs.2019.102795.

References

Bakhella, M., Moujib, M., Lookhart, G.L., Hoseney, R.C., 1992. Theories of wheat hardness and the methods of its measurement. AL AWAMIA 76, 77–97.
Cauvain, S.P., Young, L.S. (Eds.), 2009. The ICC Handbook of Cereals, Flour, Dough & Product Testing: Methods and Applications. DEStech Publications, Inc, Lancaster, Pennsylvania, USA.
Champagne, E.T., Bett-Garber, K.L., Grimm, C.C., McClung, A.M., Moldenhauer, K.A., Linscombe, S., McKenzie, K.S., Barton, F.E., 2001. Near infrared reflectance analysis for prediction of cooked rice texture. Cereal Chem. 78 (3), 358–362.
Cuevas, R.P., Pede, V.O., McKinley, J., Velarde, O., Demont, M., 2016. Rice grain quality and consumer preferences: a case study of two rural towns in the Philippines. PLoS One 11 (3), 0150345.
Enevoldsen, B.S., Juliano, B.O., 1998. Ratio of A chains to B chains in rice amylopectins. Cereal Chem. 65, 424–427.
Harris, P., 2012. Food Gels. Springer Science & Business Media, pp. 210.
Jinendra, B.M., Tamaki, K., Kuroki, S., Vassileva, M., Yoshida, S., Tsenkoa, R., 2010. Near infrared spectroscopy and aquaphotomics: novel approach for rapid in vivo diagnosis of virus infected soybean. Biochem. Biophys. Res. Commun. 397, 685–690.
Jongkaewwattana, S., Geng, S., 2001. Inter-relationships amongst grain characteristics, grain-filling parameters and rice (Oryza sativa L.) milling quality. J. Agron. Crop Sci. 187 (4), 223–229.
Juliano, B.O., 1979. The Chemical Basis of Rice Grain Quality. Chemical aspects of rice grain quality, pp. 69–90.
Lai, F.S., Afework, S., Pomeranz, Y., 1983. Determination of hardness of wheat by the Culatti micro hammer/Cutter mill. Abstract. Cereal Foods World 28, 572.
Lai, F.S., Rousser, R., Brabec, D., Pomerat, Y., 1985. Determination of hardness in wheat mixtures. II. Apparatus for automated measurement of hardness of single kernels. Cereal Chem. 62 (17), 8–184.
Lapchareonsuk, R., Sirisomboon, P., 2015. Sensory quality evaluation of rice using visible and shortwave near-infrared spectroscopy. Int. J. Food Prop. 18 (5), 1128–1138.
Lyman, N.B., Jagadish, K.S., Nalley, L.L., Dixon, B.L., Siebenmorgen, T., 2013. Neglecting rice milling yield and quality underestimates economic losses from high-temperature stress. PLoS One 8 (8), 72157.
Meullenet, J.F., Mairomoustakos, A., Horner, T.B., Marks, B.P., 2002. Prediction of texture of cooked white rice by nearinfrared reflectance analysis of whole-grain milled sample. Cereal Chem. 79 (1), 52–57.
Miller, B., Afework, S., Hughes, J.W., Pomeranz, Z., 1981l. Wheat hardness: time required to grind wheat with the Brabender Automatic Micro hardness tester. J. Food Sci. 46 1863-186.
Miller, B., Pomeranz, S.Y., Afework, S., 1984. Hardness of hardred winter weat grown in soft wheat area and of red winter wheat drown in a hard wheat area. Cereal Chem. 61, 201–203.
Mir, S.A., Bosco, S.J.D., 2013. Effect of soaking temperature on physical and functional properties of parboiled rice cultivars grown in temperate region of India. Food Nutr. Sci. 4 (03), 282.
Mishra, G., Srivastava, S., Panda, B.K., Mishra, H.N., 2018. Rapid assessment of quality change and insect infestation in stored wheat grain using FT-NIR spectroscopy and chemometrics. Food.Anal.Methods. 11 (4), 1189–1198.
Mutters, R.G., Thompson, J.F., 2009. Rice quality handbook. Postharvest technology center, department of plant sciences-mail stop 2, one shields avenue. University of California. UCANR Publications 3514.
Pearce, M.D., Marks, B.P., Meullenet, J.F., 2001. Effects of postharvest parameters on functional changes during rough rice storage. Cereal Chem. 78 (3), 354–357.
Pomeranz, Y., Webb, B.D., 1985. Rice hardness and functional properties. Cereal Foods World 30 (11), 784–790.
Poritosh, R., Shimizu, N., Shiina, T., Kimura, T., 2006. Energy consumption and cost analysis of local parboiling processes. J. Food Eng. 76, 646–655.
Qingyun, L., Yeming, C., Mikami, T., Kawano, M., Zaigui, L., 2007. Adaptability of four-samples sensory tests and prediction of visual and near-infrared reflectance spectroscopy for Chinese indica rice. J. Food Eng. 79 (4), 1445–1451.
Safdar, M.N., Naseem, K., Siddiqui, N., Amjad, M., Hameed, T., Khalil, S., 2009. Quality evaluation of different wheat varieties for the production of unleavened flat bread (Chapatti). Pakistan J. Nutr. 8 (11), 1773–1778.
Siriphollakul, P., Kanlayanarat, S., Rittiron, R., Wanitchang, J., Boonyaritthongchai, P., Nakano, K., 2013. Physical properties and correlation between NIRS spectrum of'Khao dawk Mali 105'Paddy rice. Acta.Hortic. 1088, 649–653.
Srivastava, S., Mishra, G., Mishra, H.N., 2018. FTNIR-A robust diagnostic tool for the rapid detection of rhyzopertha Dominica and *Sitophilus oryzae* infestation and quality changes in stored rice grains. Food Bioprocess Technol. 1–12.
Stenvert, N.L., 1974. Grinding resistance, a simple measure of wheat hardness. J.Flour.Anim. Feed.Milling 156 (7).
Tulyathan, V., Leeharatanaluk, B., 2007. Changes in quality of rice (Oryza sativa L.) cv. Khao Dawk Mali 105 during storage. J. Food Biochem. 31, 415–425.
Webb, B.D., Pomeranz, Y., Afework, S., Lai, F.S., Bollich, C.N., 1986. Rice grain hardness and its relation to some milling, cooking and processing characteristics. Am.Assoc.Cereal Chem. 63 (1), 27–30.
Williams, P., 2003. Near-infrared Technology–Getting the Best Out of Light. PDK Grain. pp. 8–10 Nanaimo, Canada.
Zhang, C., Shen, Y., Chen, J., Xiao, P., Bao, J., 2008. Nondestructive prediction of total phenolics, flavonoid contents, and antioxidant capacity of rice grain using near-infrared spectroscopy. J. Agric. Food Chem. 56 (18), 8268–8272.
Zhou, Z., Robards, K., Helliwell, S., Blanchard, C., 2002. Ageing of stored rice: changes in chemical and physical attributes. J. Cereal Sci. 35 (1), 65–78.