https://doi.org/10.48196/016.02.2020.05

Non-Destructive Prediction of Paddy Seed Quality using Near Infrared Spectroscopy

Fidelina T. Flores[1], Joanne P. Foliente[2], Kevin F. Yaptenco[3], and Paul R. Armstrong[4]

[1]Assistant Professor 3, Agricultural, Food & Bioprocess Engineering Division , Institute of Agricultural and Biosystems Engineering , College of Engineering and Agro-Industrial Technology, University of the Philippines Los Baños, 4031 College, Laguna, Philippines (Author for correspondence email: ftflores@up.edu.ph)

[2]University Research Associate, Agricultural, Food & Bioprocess Engineering Division, Institute of Agricultural and Biosystems Engineering , College of Engineering and Agro-Industrial Technology, University of the Philippines Los Baños, 4031 College, Laguna, Philippines

[3]Professor 9, Agricultural, Food & Bioprocess Engineering Division , Institute of Agricultural and Biosystems Engineering , College of Engineering and Agro-Industrial Technology, University of the Philippines Los Baños, 4031 College, Laguna, Philippines

[4]Research Agricultural Engineer, Center for Grain and Animal Health Research (CGAHR), United States Department of Agriculture-Agricultural Research Service (USDA-ARS), Manhattan, Kansas, 66502, USA

ABSTRACT

The need for a rapid and non-destructive method for measuring paddy seed quality is important since it could help save time and effort of the breeders and farmers. Near-infrared (NIR) spectroscopy which is a non-destructive analytical technique was used in this study. It was used to predict the germination rate and vigor of hybrid paddy seeds and classify the seeds according to the level of germination rate. A total of 71 seed lots with varying germination rate and vigor were used in the study. Paddy seeds were classified according to their germination rate (very high >90%, high 80-90%, moderate 70-79% and low <70%). Each seed lots were scanned at five random locations. Reflectance values of paddy seeds were scanned using NIR spectrometers. After scanning, the germination rate was determined using standard methods in the laboratory. Pre-processing techniques were done to the spectral data gathered. The preprocessed spectral data were subjected to principal component analysis and partial least square regression (PLSR) to predict the paddy seed quality parameters. Performance of PLSR using the PCA scores from the preprocessed spectral data of paddy seeds showed that the calibration model for predicting the germination rate had a computed R^2 value of 0.626-0.804 and the RMSE was 11.716-8.489. The R^2 and RMSE of the model for seed vigor was 0.670-0.787 and 14.095-11.328, respectively. The model for predicting the seed class had a computed R^2 value of 0.803 and the RMSE was 0.514 for the NIR spectrometer and the R^2 and RMSE was 0.648 and 0.687, respectively, for the vis NIR spectrometer. Based on the RMSE values and using the guidelines by Williams and Norris (2001) the models for seed class can be used with caution in most applications like in research. Results showed the potential of NIR spectroscopy as a non-destructive method of predicting seed quality parameters and classifying paddy seeds according to the level of germination rate. The results of the study could be used to predict seed germination rate and vigor and classify sort seed lots quickly based on the values of germination rate and seed vigor.

Keywords— seed classification, germination rate, seed vigor, NIR, spectroscopy

INTRODUCTION

There is a need to have a faster, cost-effective and reliable test method for seed quality control monitoring. The usual method of quality measurement of hybrid rice seeds is time-consuming as the results will usually be available after 7- 10 days. During peak season of operations, a faster way of quality control monitoring of seed assessment will improve the product reliability and the this will enable to make quick decision for traders. Seed quality parameters that require a quick assessment are seed germination rate and vigor.

Germination test is done to determine the percentage of seeds that are alive in each seed lot. It provides a measure of the potential of the crop in the field. A viable seed that germinates should be able to absorb moisture within two days and produce a root and the first leaf within four days (IRRI Rice Knowledge Bank, 2019). Evaluation of seed quality parameters is time-consuming, laborious, requires experience (Rahman and Cho 2016)

Near-infrared spectroscopy (NIRS) uses the near-infrared (NIR) radiation in the wavelength range of 780-2500 nm to assess the response of molecular bonds (specifically O-H, C-H, and N-H) when stimulated. NIR light hitting a sample can be reflected, absorbed or transmitted; the amount of light apportioned to each phenomenon is a function of the chemical makeup and physical structure of the material (Burns and Ciurczak, 2007). NIRS have been used to discriminate between viable and non-viable tomato seeds (Shrestha et al 2017) and predict germination rate of hybrid rice seeds (Li et al 2008).

One important advantage of NIRS is immediate availability of results unlike the traditional tests that would usually take days. This will allow for quick decisions to be made regarding disposal of seed lots for the traders and farmers. The study aimed to establish the feasibility of using NIR to classify hybrid paddy seeds according to the level of germination rate and to develop calibration models using near-infrared spectroscopy to discriminate between paddy seeds according to seed germination rates and validate the calibration models.

METHODOLOGY

Seed Sampling

A total of 71 seed lots at different germination rate were used in the study. The average moisture content of the seeds is 9.19%. Seeds were categorized according to the germination rate: 90% and above are considered very high, 80-89% are high, 70-79% are moderate and low for less than 70%. Seed samples (Figure 1) were placed in glass petri dish that serve as sample holders; the amount of seed was sufficient to completely cover the bottom prior to scanning. Each lot was scanned five times and after scanning the samples were subjected to the standard method of determining the germination rate. The five spectral data were divided into calibration set (4) and validation set (1).

In the standard method of determining the germination rate of seeds, 100 seeds were placed in a moist germination paper and the relative humidity and temperature to where the samples were exposed was monitored. After 5 days, the number of seeds that germinated were counted (seed vigor) and after another 2 days the total number of seeds (germination rate) that germinate was counted.

Figure 1. Paddy seeds in petri dishes ready for NIR scanning.

Development of Calibration and Validation Models

Seed lots were scanned using the Ocean Optics Inc. spectrometers in reflectance mode. USB 4000 with a wavelength range of 300-1,000 nm and NIR Quest 512 with wavelength range of 900-1,700 nm were used. An Ocean Optics HL-2000 with tungsten-halogen lamp and light range of 360 nm to 2400 nm was used as a light source. A white Labsphere's Spectralon disc was used as a standard reference and the spectrometer was calibrated before scanning samples. Each sample dish was scanned at random points to produce five spectra. The spectrometer had a dedicated bifurcated fiber optic reflectance probe (Ocean Optics QR 600 series) to transmit light from the light source. Figure 2 shows the setup for spectral data acquisition.

SpectraSuite Spectroscopy Software by Ocean Optics, Inc. was used for acquisition of sample spectra; each spectral scan was the average of 10 multiple scans, with a boxcar width of 10 nm. Each paddy seed lot was scanned at five random points. Nicolai et al (2007) recommended a detector response of 50% of the saturation level to determine integration time. Using this guide, a gap of 7 mm between the surface of the sample and the probe tip was maintained using a probe holder, while integration time for the NIRQuest 512 and USB 4000 spectrometers was set at 12 ms and 8 ms, respectively. The fiber optic probe was mounted perpendicular to the sample surface to obtain reflectance spectra.

Preliminary test was done to check the effect of varying the acquisition parameters on the spectral data using the two spectrometers. This step will also avoid and minimize computer lagging during data acquisition and decreasing the resolution of the spectral data. The white 'reference' and 'dark' spectra was also taken for calibration before after setting the optimum parameters and before data acquisition.

Preprocessing and transformation of spectral data was performed to reduce the effect of light scattering, data noise and peak/valley overlap. The preprocessing techniques mean centering and mean scatter correction were used to develop calibration models based on the coefficient of determination (R^2) and root mean square error (RMSE). Mean centering (MC) is the first preprocessing technique done to the data. This step guarantees that the results will be to consider the variation within the mean and is recommended for all data types. The next transformation is the multiplicative scatter correction (MSC) which is a normalization technique. This step is used to compensate for additive (baseline shift) and multiplicative (tilt) effects present in the spectral data. The step removes the effects of scattering by linearizing each spectrum to some 'ideal' spectrum of the sample (Nicolai et al., 2007).

Acquired NIR spectra were subjected to data preprocessing by mean centering (MC) and mean scatter correction (MSC) using an open source platform, R Software version 4.0.2 to reduce the effect of light scattering, data noise and peak/ valley overlap. Transformed spectral data were subjected to a combination of principal component analysis (PCA) and partial least squares regression (PLSR). Principal component (PC) scores from PCA were then used as a new set of independent variables for PLSR against seed class.

Calibration models of germination rate, seed vigor and seed class were developed using calibration data set Calibration models were then validated

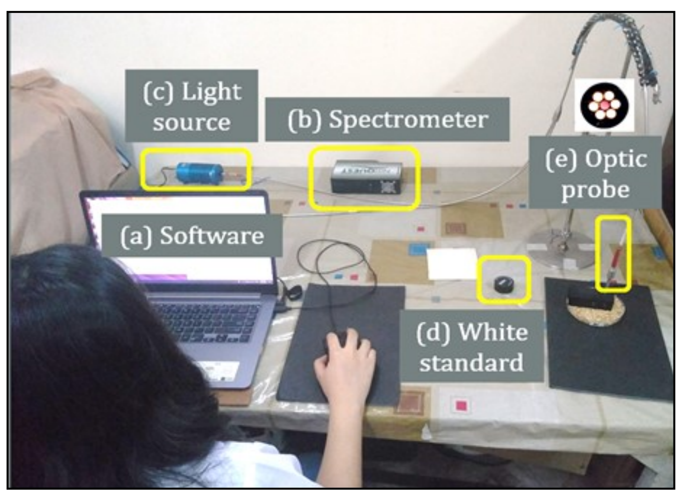

Figure 2. NIR laboratory setup.

using the prediction data set; R^2 and RMSE were computed as described previously.

RESULTS AND DISCUSSION

Seed Samples

The 71 seed lots are divided into 24 lots for very high germination rate, 21 lots for high germination rate, 15 lots for moderate germination rate and 11 seed lots for low germination rate. Seed lots were scanned five times, resulting in 355 spectral data divided into 284 spectral data for a calibration set and 71 spectral data for an independent prediction set.

Data Preprocessing and Pretreatment.

After data filtering, spectra were preprocessed using R Software version 4.0.2. Spectral pre-processing is a critical step before creating a model. Samples have spectral variations due to the scattering of light and difference in effective path length. Multiplicative scatter correction (MSC) and mean centering (MC) are the techniques used for data preprocessing and pretreatment to correct the raw spectral data for any light scattering effect.

Preprocessing and pretreatment of these spectral data is important to minimize errors that is present in the raw data and it will also aid in determining the exact location or peaks of these chemical bonds to have a basis for classifying and predicting seed quality parameters.

Raw spectral data of the reflectance values of the seed samples is shown in Figures 3 and 4 for NIR spectrometer and the vis NIR spectrometers, respectively. The plots in Figures 3 and 4 are divided according to the level of germination rate. There are common peaks and valleys at different wavelength regions in the NIR such as in the 1150 nm-1250 nm, 1400 nm-1500 nm and 1600 nm-1650 nm and in the vis NIR the points are in

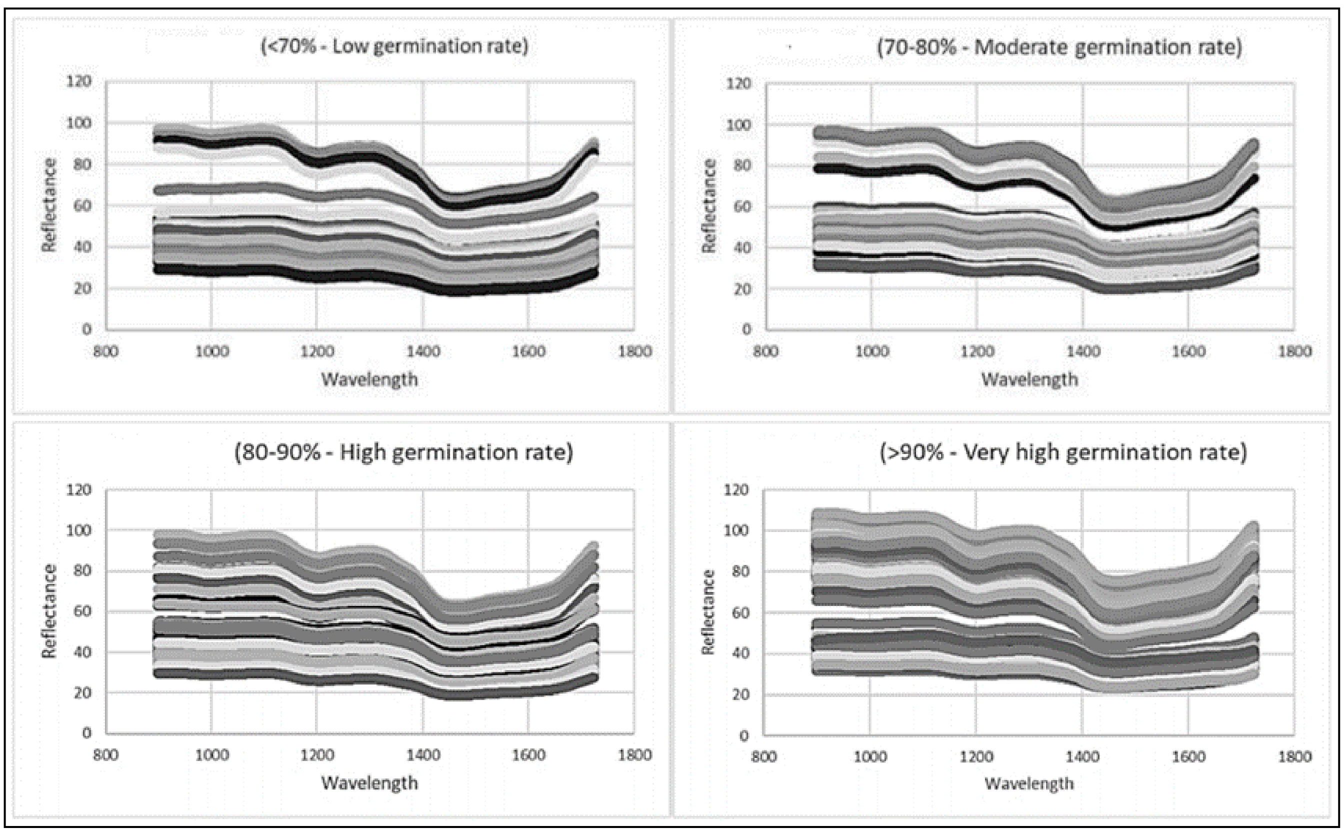

Figure 3. Raw spectral data of paddy seeds at NIR region under different levels of germination rate.

between 375 nm-450 nm and 850 nm-950 nm regions. Reflectance values in these regions varies based on the level of germination rate for each variety.

After data preprocessing, PCA-PLSR analysis is done on the preprocessed spectral data. Table 1 shows the result of the performance of the models developed using the NIR and vis-NIR spectrometers. The parameters used to evaluate the performance of the models are R^2 and RMSE.

Based on the results using the NIRQuest 512 spectrometer, the calibration model for seed class has an R^2 value of 0.803 and RMSE of 0.514. For the results using the USB 4000 spectrometer, the calibration model has an R^2 of 0.648 and a RMSE is 0.687. Based on the RMSE value and using the guidelines by Williams and Norris (2001) the model can be used with caution in most applications like in research.

Based on the results using the two spectrometers, between the two quantitative parameters, seed vigor has the highest R^2 (0.670-0.787) and RMSE values (14.095-11.328) compared to germination rated (R^2 range 0.626-0.804 and RMSE range 11.716-8.489). Based on the RMSE value and using the guidelines by Williams and Norris (2001) the model can be used with caution in most screening applications like in research. Figures 5 and 6 show predicted values of seed germination rate, vigor and seed class using the calibration model and the prediction model.

The seed germination and vigor of artificially aged paddy seeds was classified using NIR hyperspectral imaging and the models have an accuracy ranging from 87.5% - 93.67% (He, et.al 2019). Soybean seed quality was also predicted using a spectrometer with a range from 950nm to 1650nm and the R^2 of the model ranged from 0.4428 to 0.4631. Results also showed that qualitative prediction of germination rate, vigor

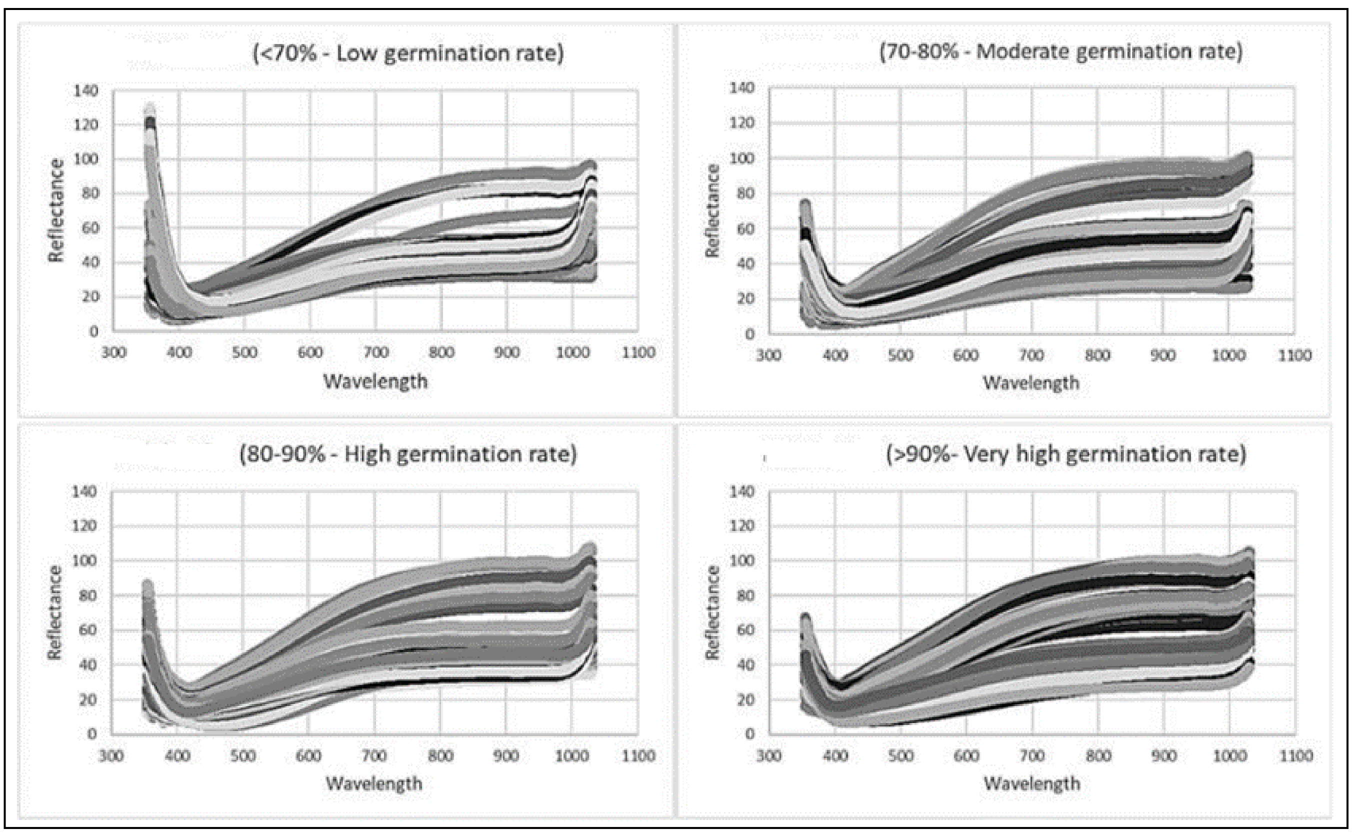

Figure 4. Raw spectral data of paddy seeds at visible region under different levels of germination rate

Table 1. Performance of PLSR model for seed class using PCA scores using the two spectrometers

PARAMETER		SPECTROMETER MODEL			
		NIR QUEST 512		USB 4000	
		(900nm-170nm)		(300nm-1000nm)	
		Calibration	Prediction	Calibration	Prediction
Germination Rate	R^2	0.804	0.294	0.626	0.250
	RMSE	8.489	18.796	11.716	17.159
	MAE	6.675	14.812	8.573	12.738
Seed Vigor	R^2	0.787	0.313	0.670	0.249
	RMSE	11.328	23.460	14.095	22.230
	MAE	9.032	18.935	10.792	16.984
Seed Class	R^2	0.803	0.370	0.648	0.287
	RMSE	0.514	1.031	0.687	1.036
	MAE	0.417	0.812	0.551	0.824

was possible using 2-category models (e.g. low or high) and that quantitative prediction of germination was less accurate (Al-Amery et.al 2018).

A similar study that predicts the percent germination rate of hybrid seeds developed a model with an R^2 of 0.964 using a spectral range of 600 nm-1100 nm. The samples had undergone an accelerated aging process and the range of germination rate is between 65-90% (Li, et.al 2014). Another study predicted the storage age (one to four years) of paddy seeds and the model developed was 97.5% accurate in predicting age (1 year to 4 years of storage) using artificial a neural network (ANN) model. The study indicated that paddy seed stored for different periods could be classified based on visible NIR spectroscopy (Li, et. al 2008). Variations in the study might be due to the preprocessing and transformation techniques, multivariate analysis and the difference in the rice variety and spectrometer used.

Based on the results of the models using the two spectrometers, it shows that both the qualitative models (seed class) and the quantitative models (seed vigor and germination rate) developed was able to predict seed quality parameters. The results can then be used as a basis in designing a portable NIR instrument that could serve for on-site screening to easily identify the seed class of the samples and decide if it will be accepted for further processing or for disposal. The portable NIR instrument will use the significant wavelength that is attributed with the seed quality parameters.

SUMMARY AND CONCLUSION

The study evaluated the potential of using NIR spectroscopy to predict seed quality parameters and classify hybrid paddy seeds according to the level of germination rate. The reflectance values of the paddy seed samples at two spectrometers, which worked at the range 300 nm – 1000 nm and 900 nm – 1700 nm, were gathered. Data transformation and multivariate analysis were employed to differentiate the raw spectral data depending on the seed quality parameters. The preprocessed spectral data were then modelled using multivariate analysis. Principal component analysis (PCA) of the preprocessed spectral data was performed and the scores based on the number of useful principal components were used in the partial least square regression. PLSR was done to develop regression models that would predict hybrid seed quality parameters. The performance of the models were described by the R^2 and RMSE values.

Results showed that the four-level qualitative and quantitative models could predict seed quality parameters like vigor and germination rate and classify seeds according to seed class. The models using developed in the 900-1700 nm region (R^2 0.787-0.804) performed better compared to those in the region 300-1000 nm (R^2 0.626-0.670).These showed that the NIR spectral data of paddy seeds had information that could be used in predicting the germination rate and seed vigor and classifying samples according to their level of seed germination rate. Thus, it can be concluded that the use of NIR spectroscopy with multivariate analysis is an effective method to predict seed quality parameters and classify them.

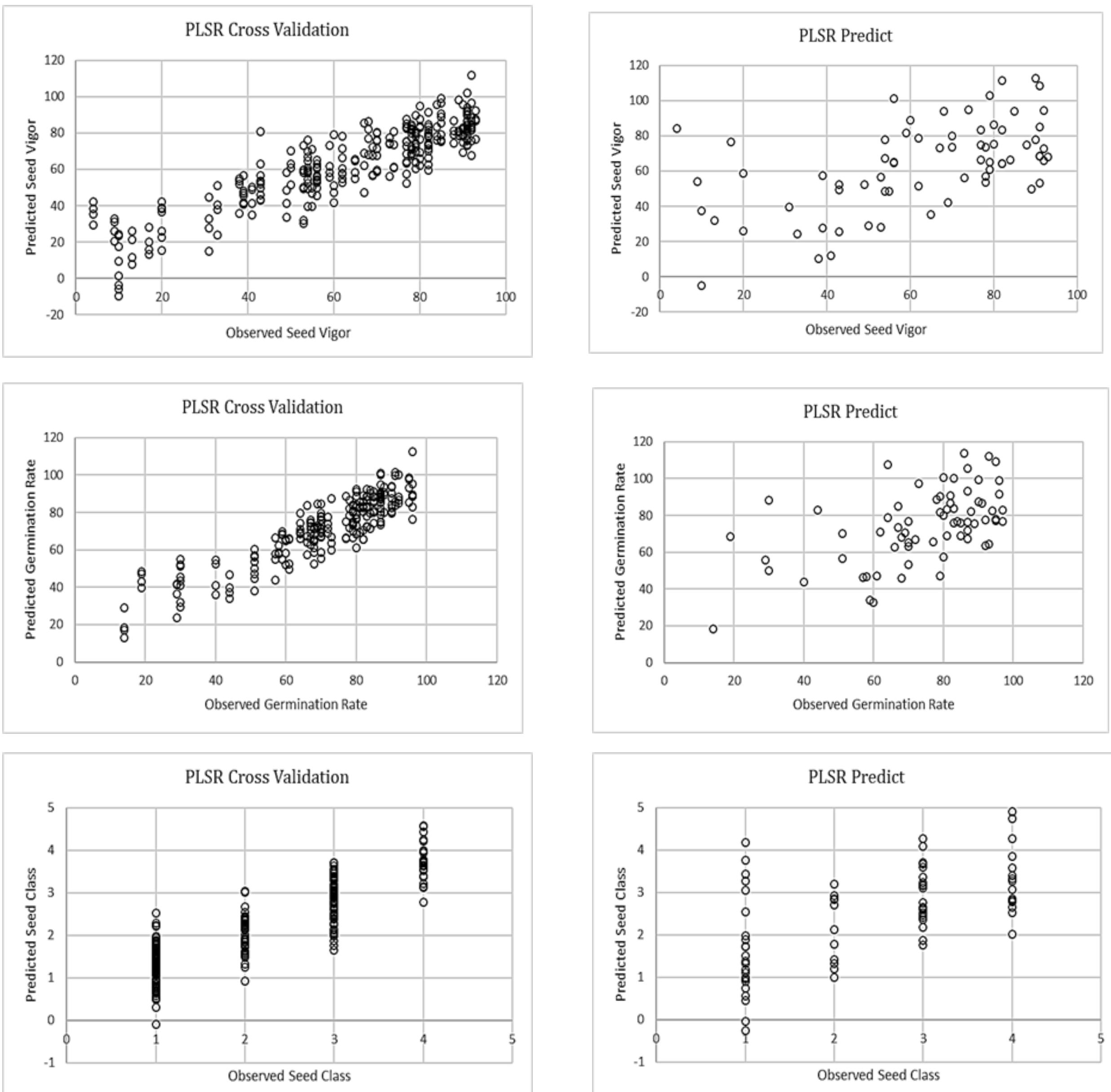

Figure 5. PLSR Cross validation and PLSR Prediction models of paddy seeds using NIR Quest 512 spectrometer at near infrared region.

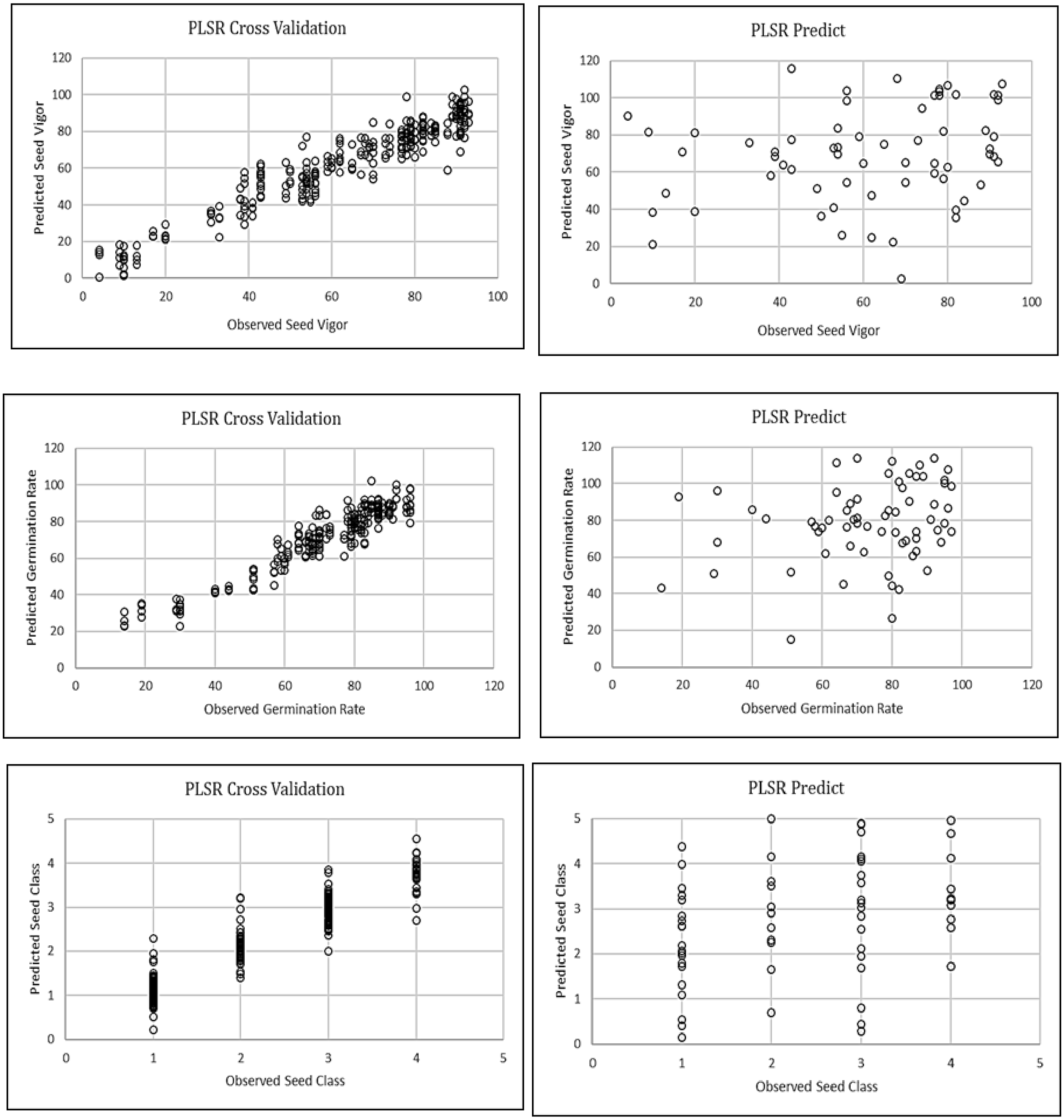

Figure 6. PLSR Cross validation and PLSR Prediction models of paddy seeds using USB 4000 spectrometer at visible region.

RECOMMENDATION

The study showed the feasibility of using NIR spectroscopy in classifying and predicting paddy seed quality parameters. With this, the following are recommended for further research to predict paddy seed quality using an FT-IR instrument or other portable type spectrometer for field use and fast identification of seed parameters and to consider reducing the seed classification into two (pass of fail based on a given value) or three (low, medium and high) class.

ACKNOWLEDGMENT

The authors acknowledge the support of the Philippine Commission on Higher Education and U.S. Department of Agriculture (USDA) – Stored Product Insect and Engineering Research Unit (SPIER) for the NIR spectrometer unit used in the study. The mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval for the exclusion of other products that may also be suitable. USDA is an equal opportunity provider and employer.

LITERATURE CITED

AL-AMERY, M., GENEVE, R.L., SANCHES, M.F., ARMSTRONG, P.R., MAGHIRANG, E.B., LEE, C., VIEIRA, R.D., AND HILDEBRAND, D.F. 2018. Near-infrared spectroscopy used to predict soybean seed germination and vigour. Seed Sci Res. 1-8.

BURNS, D.A. AND CIURCZAK, E.W. 2007. Handbook of Near-Infrared Analysis. 3rd Edition. CRC Press. 2007. 834 pages.

HE, X., FENG, X., SUN, D., LIU, F., BAO, Y., AND HE, Y. 2019. Rapid and Nondestructive Measurement of Rice Seed Vitality of Different Years Using Near-Infrared Hyperspectral Imaging. Molecules 24 (2227)

IRRI Rice Knowledge Bank, 2019. Seed Quality. (Retrieved from the http://www.knowledgebank.irri.org/step-by-step-production/pre-planting/seed-quality on April 2, 2019)

LI, X., HE Y, WU C. 2008. Non-destructive discrimination of paddy seeds of different storage age based on vis-NIR spectroscopy. J Stored Prod Res. 44: 264-268.

LI, Y., JIANG D, LIU Y, DING W, DING Q, ZHA L. 2014. Testing of Germination Rate of Hybrid Rice Seeds Based on Near-Infrared Reflectance Spectroscopy. Spectroscopy and Spectral Analysis. 34(6): 1528-1532.

NICOLAI, B.M., BEULLENS, K., BOBELYN, E., PEIRS, A., SAEYS, W., THERON, K.I., AND LAMMERTYN, J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biology Technology. 46: 99-118.

RAHMAN A, CHO BK. 2016. Assessment of seed quality using non-destructive measurement techniques: a review. Seed Sci Res. 26: 285-305.

SHRESTHA S, DELEURAN LC, GISLUM R. (2017). Separation of viable and non-viable tomato (*Solanum lycopersicum L.)*. Comp Elect Agric. 142: 348-355.

WILLIAMS, P. C., AND NORRIS, K. 2001. Implementation of near-infrared technology. In Near-Infrared Technology in Agricultural and Food Industries, eds. P. Williams, and K. Norris, 145 169. St. Paul, Minnesota: American Association of Cereal Chemists. ■