

Applying singular value decomposition technique for quantifying the insects in commercial Thai Hommali Rice from NIR Spectrum

Puttinun Jarruwat* and Prasan Choomjaihan[†]
*Department of Agricultural Engineering,
Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang
Ladkrabang, Bangkok, Thailand, 10520
*putjar2001@yahoo.com
[†]prasan.ch@kmitl.ac.th*

Received 17 February 2016

Accepted 19 June 2016

Published 10 August 2016

Insect infestation in rice stock is a significant issue in rice exporting business, resulting in the loss of product quality, nutrient as well as the economic losses. However, detecting the insect contamination with the traditional sorting techniques were destructive, inaccurate, time consuming and unable to detect the internal insect infestation. This study used near infrared (NIR) spectroscopy for obtaining the absorbent spectra from the insect contamination in two kinds of rice samples, Milled Hommali rice (MHR) and Brown Hommali rice (BHR). The mathematical methods of partial least squares (PLSs) regression and singular value decomposition (SVD) were employed to construct the predicting model. The statistical analysis results, R^2 , RMSEP, RPD and bias, concluded that the predictive models from PLS for MHR and BHR were 0.95 and 0.90, 0.014 and 0.019, 4.79 and 3.11, as well as -0.007 and -0.008 , respectively; while the statistical analysis results from SVD for MHR and BHR were 0.97 and 0.96, 0.012 and 0.013, 5.71 and 5.39, as well as -0.003 and 0.002, respectively. It showed that SVD technique performed better than PLS technique which shows that using the advantage of SVD technique required less amounts of wave numbers for predicting and was possible to construct the low cost handheld equipment for detecting the insects in rice samples.

Keywords: Insect; rice; NIR; SVD.

[†]Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.

1. Introduction

Rice is one of the major food sources with rich nutrition for mankind living and also becomes the most important agricultural product in the Asian economy. Thailand has been nominated as one of the top largest rice production countries with about 22% of Thailand exporting goods.¹ Before exporting, rice grain should be stored strategically for economical purposes according to the time frame. Generally, about 5–10% of grain stocks are significantly deteriorated due to insect infestation during storage.² The damage to grains by insects causes loss of weight, nutrients, germination ability and increased susceptibility to contamination by fungi and downgrades the grain resulting in lowering its market value.³ The damaged grains are not only resulting in an economic loss in terms of quantity and quality losses but could also be turned into political issue in the exporting country. Rejection from the buyer countries can occur, if a large proportion of rice grains are damaged or infested with insects or both. Internationally, standards of rice products impurity have been announced by FAO and WHO called CODEX ALIMENTARIUS. They have been used as guidelines and codes of practice to contribute to the safety, quality and fairness of the international food trade. Certainly, the standard for exporting rice is included in CODEX STAN 198-1995 and has set the amount of filth (impurities of animal origin, including dead insects) which must not exceed 0.1%.⁴ Domestically, the ministry of commerce of Thailand has specified that the foreign material allowance in premium grade of Thai rice product must not exceed 0.2% for domestic trade.⁵ According to the report of FAO, Thailand has encountered 10% of rice grain damage from insect infestation yearly²; particularly, the damage from rice weevil. Traditionally, insect detection methods used the stick trier or the sleeve-type trier to stab into the package and to collect the rice sample for examination by visual inspection, sieving and floatation methods. These methods are widely used for insect contamination in grain stock. Most of these methods tend to have one or more disadvantages such as being subjective, destructive, inaccurate, time consuming, especially, sometimes unable to detect internal insect infestation. However, chemical and electrical detections of insect infestation, such as carbon dioxide, uric acid

measurement, X-ray detection, electronic nose, and acoustic impact emissions^{6–10} have become alternative examinations. The problems of some methods were their inability to detect low-intensity of insect infestation and some of them lacked potential when being examined automatically.

The near infrared (NIR) spectroscopy can solve these problems with the real-time inspection, which is more reliable and more accurate. The NIR spectroscopy technique measures the chemical composition of biological materials using diffuse reflectance or transmittance of the samples at several wavelengths including detection of insect or piece of insect in whole single grain, ground grain and bulk grain samples.^{11–17} However, NIR spectroscopy detection method also has weaknesses such as the complex development of robust calibration models, and inconsistency across several individual instruments. Most of the commercial NIR instruments are inconvenient for practical use in the field. The suggested suitable predicting ranges of wavelengths of NIR were calculated from statistically complex mathematic methods. Sometime the selected wavelengths were not related to the chemical origin of measuring materials or unable to explain the source of chemical terms. The chemical compound profiles of insect and of rice are used, related to their active wavelengths resulting from NIR spectroscopy, then, the mathematics algorithm is applied to solve the proportion of the insect infestation which is a more reliable and trustful method. The selected active wavelengths can be used more numbers of wavelength for rice weevil and rice than one wavelength for each. Therefore, the set of equation becomes greater than two equations from two wavelengths for solving two unknowns (the equation becomes greater than a square 2×2 matrix; i.e., 3×2 , 4×2 and so on). Singular valued decomposition (SVD) is a powerful mathematical technique to deal with the solution of nonsquare sets of equations or matrices based on the theorem of linear algebra¹⁸ and it has been used with several applications such as measuring bubble size distribution in bread doughs, and identifying distribution of wheat compositions throughout sizes from the First Break milling process.^{19,20} Furthermore, SVD has been widely applied in many areas of research, for instance, for mapping the X-ray signal to measure the distribution of acrylic acid in microporous polypropylene,²¹ for signal enhancement and noise suppression of seismic

data²² and for detecting the defect using the magnetostrictive guided wave technique.²³

In this study, Fourier transform near-infrared (FT-NIR) spectral technique was performed for carrying out the full spectral data. Therefore, the objective of this study is to apply SVD for calculating the proportion of insect infestation in rice stock from FT-NIR spectral data.

2. Material and Methods

2.1. Sample preparation and NIR data acquisition

Brown Hom mali rice (BHR) and Milled Hom mali rice (MHR) samples at approximately 14% moisture content (wet basis) were used throughout this study and collected directly from four different rice mill plants. Moisture content was determined using a standard oven method (AOAC, 2000). The broken kernels were removed leaving only the whole healthy kernel for examination. Each rice sample was mixed with dead rice weevils (DRW) by proportion of 0%, 0.018%, 0.037%, 0.055%, 0.074%, 0.093%, 0.111%, 0.130%, 0.148%, 0.167%, 0.185%, 0.204% and 0.222% by weight which covered the standard allowance range of CODEX STAN 198-1995 and of the standard of Thai rice product.

A multi-purpose analyzer FT-NIR spectrometer (Bruker Corporation, Germany) with the wave number range of 12,500–4,000 cm^{-1} (800–2500 nm) was used throughout this experiment. The diffuse reflectance mode at 16 cm^{-1} was set for measuring sample. The spectrum of the sample was derived from the average of 64 scans. The sample holder is cylinder-shaped with 8.7 cm in diameter (inside), 9.0 cm in height, and 0.5 cm in thickness. The base of the sample holder is made of a transparent crystal quartz and the side is made of tin-plated metal. The top of the container is an opening channel for filling and emptying samples.

100 g of uniformly mixed rice samples were filled into the sample holder; then, were placed on the measuring station of the multi-purpose analyzer FT-NIR. The experiments were carried out with four rice sources of two types of rice samples (MHR and BHR) and 13 levels of contamination (10 measuring for each mixture), giving rise to a total of 1040 experiments. The mathematical spectral treatments (i.e., standard normal variate

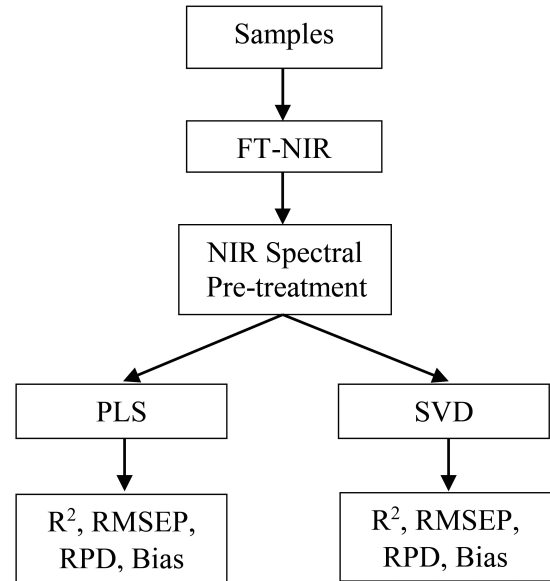


Fig. 1. Flow diagram of the experiment.

(SNV), min-max normalization, Based line offset, multiplicative scattering correction (MSC), first derivative, second derivative, smoothing, straight line subtraction, SNV + first derivative and MSC + first derivative) were then calculated from NIR raw absorbent spectra. Pretreated spectra were used for further data analysis, PLS and SVD. The flow diagram throughout this experiment is illustrated in Fig. 1.

2.2. Data analysis by PLS regression for NIR

The pretreated spectra and the proportion of rice weevils were then analyzed using a partial least squares (PLSs) regression of the OPUS 7.0 (Bruker Corporation, Germany) program in order to establish the prediction model. The optimal number of PLS factors was determined based on the least residual sum of square using the OPUS 7.0 program whereby suitable ranges of wavelengths were automatically identified. The best mathematical pretreatment for prediction of NIR spectra was indicated using statistical information as expressed by coefficients of determination (R^2_{val}), root mean square error of prediction (RMSEP), ratio of standard deviation of reference data in the validation set to standard error of prediction (RPD), and the average of the differences between the actual value and the NIR value (i.e., bias).

2.3. Applying SVD technique for NIR

The SVD approach was used as the tool for generating proportional results of rice weevil versus rice sample. The SVD calculation was performed by MATLAB program version 7.11.0. SVD method uses the fundamental of matrices calculation based on the following theorem of linear algebra.¹⁸ The SVD structure was composed by three related matrices as $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, where \mathbf{A} describes a linear mapping from the vector space \mathbf{x} to the vector space \mathbf{b} . If the set of equations has m linear equation with n coefficients, the size of matrix \mathbf{A} , \mathbf{x} and \mathbf{b} can be defined as $m \times n$, $n \times 1$, and $m \times 1$, respectively. \mathbf{A} and \mathbf{b} are known; the equation is solved to find the unknown \mathbf{x} .

In terms of square matrix, computing the \mathbf{x} , the inverse of \mathbf{A} must be calculated then multiplied with \mathbf{b} as follows: $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$. In contrast, the SVD method has three composition matrices of the matrix \mathbf{A} , as $\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T$. An $m \times n$ matrix \mathbf{A} is the product of an $m \times n$ column-orthogonal matrix \mathbf{U} , and $n \times n$ diagonal matrix \mathbf{W} with positive or zero elements, and the transpose of an $n \times n$ orthogonal matrix \mathbf{V} . Then $\mathbf{A}^{-1} = \mathbf{V} \cdot [\text{diag}(1/w_j)] \cdot \mathbf{U}^T$; where $[\text{diag}(1/w_j)]$ is the inverse of diagonal element of matrix \mathbf{W} (if $w_j = 0$, the inverse of w_j is replaced with zero). Therefore, with the substitution of \mathbf{A}^{-1} , the matrix \mathbf{x} can be calculated as $\mathbf{x} = \mathbf{V} \cdot [\text{diag}(1/w_j)] \cdot \mathbf{U}^T \cdot \mathbf{b}$.

Accordingly, knowing matrix \mathbf{A} and vector \mathbf{b} , the unknown vector \mathbf{x} can be calculated by a best fit solution that minimizes the error. In the current study, the matrix \mathbf{A} was produced by NIR absorbent intensity of the particular wave numbers which were pretreated by previous explained approach. The selected wave numbers were the predominant wave numbers indicating the important chemical compositions of rice sample, BHR and MHR, and of DRW. In this study, selected wave numbers were three wave numbers for rice samples and three wave numbers for rice weevil which were six wave numbers in total. The matrix \mathbf{b} was produced by pretreating raw NIR absorbent spectra with same six selective wave numbers as matrix \mathbf{A} in unknown mixture sample. From these two matrices, the matrix \mathbf{x} was calculated, giving the proportion of rice weevil and rice sample in the unknown sample. The sizes of matrices \mathbf{A} , \mathbf{x} and \mathbf{b} were, therefore, 6×2 , 2×1 and 6×1 , respectively and composed

as follows:

$$\begin{bmatrix} A_{WL1,DRW} & A_{WL1,rs} \\ A_{WL2,DRW} & A_{WL2,rs} \\ A_{WL3,DRW} & A_{WL3,rs} \\ A_{WL4,DRW} & A_{WL4,rs} \\ A_{WL5,DRW} & A_{WL5,rs} \\ A_{WL6,DRW} & A_{WL6,rs} \end{bmatrix} \cdot \begin{bmatrix} x_{DRW,i,1} \\ x_{rs,i,1} \end{bmatrix} = \begin{bmatrix} b_{WL1,i,1} \\ b_{WL2,i,1} \\ b_{WL3,i,1} \\ b_{WL4,i,1} \\ b_{WL5,i,1} \\ b_{WL6,i,1} \end{bmatrix} \quad (1)$$

where

- (1) $\mathbf{A}_{m,n}$ is the pretreated NIR spectroscopy absorbent intensity of the particular wave number m of material n .
- (2) $\mathbf{x}_{n,1}$ is the proportion of n in unknown sample i .
- (3) $\mathbf{b}_{m,1}$ is the pretreated NIR spectroscopy absorbent intensity of the particular wave number m of unknown sample i .
- (4) m refers to the particular wave number (WL1toWL6) of DRW and rs (dead rice weevil and rice sample).
- (5) n refers to DRW and rs .

The performance of SVD technique was evaluated by statistical values, i.e., coefficients of determination (R_{val}^2) and root mean square error of prediction (RMSEP), ratio of standard deviation of reference data in the validation set to standard error of prediction (RPD), and the average of the differences between the actual value and NIR value (i.e., bias).

3. Results and Discussions

3.1. FT-NIR spectral analysis for rice and rice weevil

The original spectra for BHR, MHR and DRW from the multi-purpose analyzer FT-NIR spectrometer were individually averaged and shown in Fig. 2. It shows that at the wave number longer than about 8500 cm^{-1} , the NIR absorbance of DRW was higher than the rice samples, while, at the shorter wave number range, the NIR absorbance of rice samples as greater than DRW.

This figure demonstrates the predominant peaks of BHR and MHR were at $10,098$, 8370 , 6895 , 5685 , 5207 , 4806 , 4438 and 4019 cm^{-1} . The peaks of $10,098$, 6895 , 4806 , 4438 and 4019 cm^{-1} were related to the chemical structure of starchy material

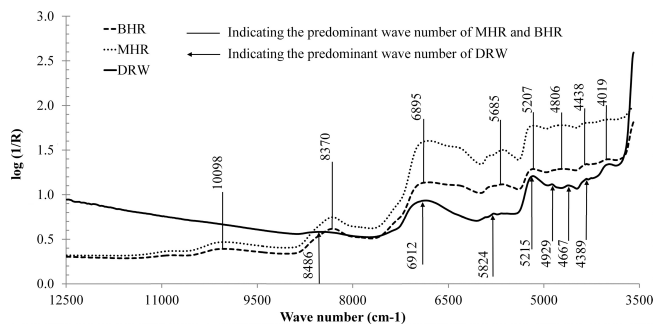


Fig. 2. Initial NIR spectra of BHR, MHR and DRW obtained using an FT-NIR spectrometer over the wave number region 12,500–4,000 cm^{-1} .

and the peaks of 8370, 5685 and 5207 were related to the chemical structure of CH_3 , CH_2 and CONH , respectively, regarding the report of Osborne *et al.*,²⁴ Furthermore, the predominant peaks for DRW were 8486, 6912, 5824, 5215, 4929, 4667 and 4389 cm^{-1} . The peaks were related to the chemical structure of $\text{HC}=\text{CH}$, aromatic, CH_2 , CONH , CONH_2 , $\text{HC}=\text{CH}$ and CH_3 , respectively.²⁴

Figures 3 and 4 show the second derivative spectra between 12500 cm^{-1} and 3500 cm^{-1} of Hommali rice samples (BHR and MHR) and the insect (DRW), respectively. In Fig. 3, the important peaks of Hommali rice samples appeared at the wave number 7043, 5886, 5616, 5261 and 4397 cm^{-1} , where the chemical structure components of these wave numbers were Phenolic or ArOH , CH_3 , cellulose and starch, respectively. Meanwhile, Fig. 4 shows the significant peaks of DRW, these appeared at the wave number 7113, 5793, 5246, 4868, 4629, 4343 and 4050 cm^{-1} , most of the illustrated peaks were related to the O–H, C–H, N–H and amide, which were respectively ROH, CH_2 , POH (Phosphate), Protein, CONHR, CH_2 and CH_3 .²⁴ Additionally, some researchers^{13,25}

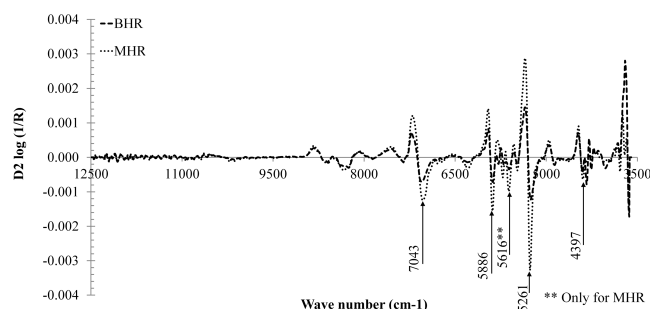


Fig. 3. Average second-derivative of NIR spectra of BHR, MHR obtained using an FT-NIR spectrometer over the wave number region 12,500–4,000 cm^{-1} .

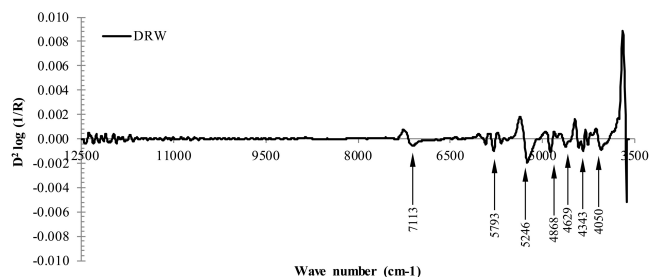


Fig. 4. Average second-derivative of NIR spectra of DRW obtained using an FT-NIR spectrometer over the wave number region 12,500–4,000 cm^{-1} .

mentioned that the wave number of 8489 cm^{-1} (1178 nm) is related to the chitin which is the majority composition in the insect cuticles and it is close to the wave number of 8486 cm^{-1} on the initial NIR spectra of DRW. It could be assumed that this wave number was related to chitin. Nevertheless, this wave number did not appear dominantly on the second derivative of NIR spectra when compared with other wave numbers as mentioned earlier.

3.2. Development of PLS models

The model for predicting DRW was developed by applying the PLS factors in the wavenumber range of 9407–5447 cm^{-1} and 4746–4421 cm^{-1} for predicting the percentage of DRW in MHR and in the wave number range of 9407–4244 cm^{-1} of predicting the percentage of DRW in BHR as demonstrated in Table 1. The PLS factors of PLS model for MHR and BHR were selected using the lowest value of RMSEP as shown in Fig. 5. The best model for predicting DRW in MHR sample was established using nine PLS factors with standard normal variate (SNV) pretreatment method. The vibration bands used in the model were mainly starch, lipid and cellulose. The MHR model expressed the statistical values of R^2 , RMSEP, RPD and bias of 0.95%, 0.014%, 4.79% and -0.007% , respectively. Meanwhile, the model predicting DRW in BHR sample was also established. The model employed 10 PLS factors from the straight line subtraction pretreatment method. The vibration bands used in the model were similar to the model of MHR, as well as, the vibration band of sucrose was also included. The BHR model, together, expressed the statistical values of R^2 , RMSEP, RPD and bias of 0.90, 0.019%, 3.11% and -0.008% , respectively. Considering the selected NIR spectrum pretreating

Table 1. Statistics of prediction of DRW (%) in MHR and BHR by partial least square model.

Type of rice	Wave number ranges	Pre-treatment	PLS Factors	R^2_{val}	RMSEP	RPD	Bias
MHR	9407–5447 cm^{-1} (1063–1836 nm) 4746–4421 cm^{-1} (2107–2262 nm)	SNV	9	0.95	0.014	4.79	-0.007
BHR	9407–4244 cm^{-1} (1063–2356 nm)	Straight line subtraction	10	0.90	0.019	3.11	-0.008

Notes: MHR — Milled Hommali rice.

BHR — Brown Hommali rice.

DRW — Dead rice weevils.

R^2_{val} — The coefficient of determination of validation set.

RMSEP — Root mean square error of prediction.

RPD — Ratio of standard deviation of reference data in validation set to standard error of prediction.

Bias — The average of difference between actual value and NIR value.

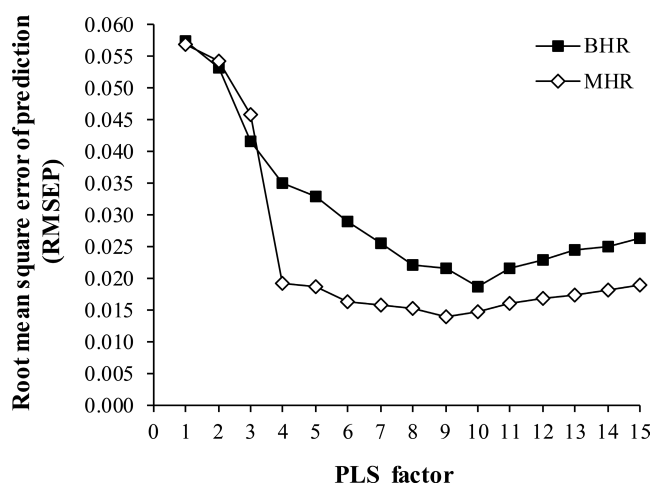


Fig. 5. Root mean square error of prediction versus PLS factors for BHR and MHR.

process, the suitable predicting models of MHR and BHR picked SNV and straight line subtraction, respectively. Even though both the pretreated processes are for adjusting the baseline shifting, tilting and scattering during the change of path length of the NIR signals, the SNV is focused on scattering correction and straight line subtraction is aimed on baseline correction.²⁶ In this case, the effect of straight line subtraction pretreating process of BHR could cause a variations in the chemical compositions of brown rice, e.g., oil, fiber, starch and etc., which affects the path length of NIR spectrum. While, the effect of SNV pretreating process of MHR could cause the surface characteristic of milled rice to have as more reflection on its surface than brown rice. Additionally, many researchers reported that the baseline shift and scattering correction

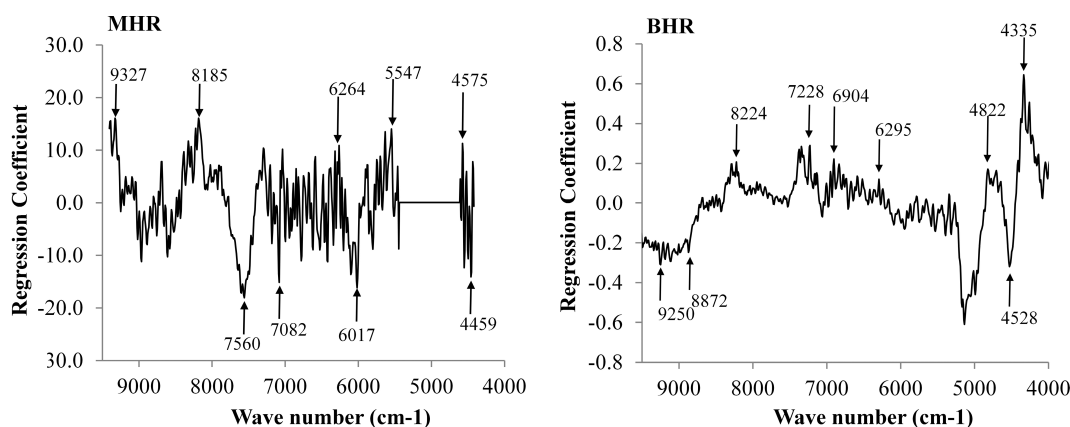


Fig. 6. Regression coefficient plots of the validated models for DRW prediction in MHR (left) and BHR (right) samples.

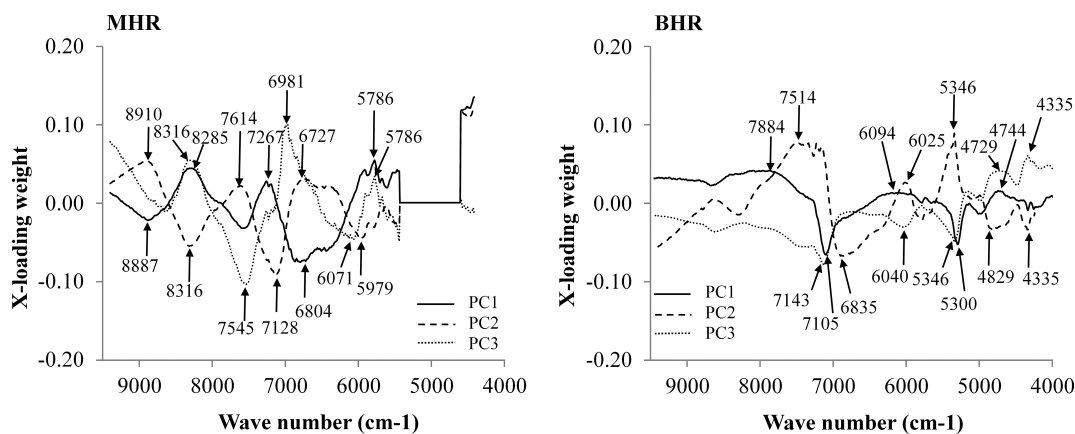


Fig. 7. X-Loading weight plots of the validated models for DRW prediction in MHR (left) and BHR (right) samples.

were needed for milled rice and brown rice materials.^{26–29}

Considering the statistical value, it has been suggested that an R^2 of 0.83–0.90 indicates that the model is usable with caution for most applications, including research; whereas, an R^2 of 0.92–0.96 implies that this model is usable in most

applications in the quality assurance level. Additionally, an RPD in the range of 3.1–4.9 has indicated that the model is for fairly screening level.³⁰

Figures 6 and 7 show the regression coefficient plots of validated model for DRW prediction and the X-loading weight of the PLS factor plots of the best validated model of prediction set using FT-NIR

Table 2. The absorption bands (from regression coefficient and X-loading weight) with predominantly affected the prediction model of DRW (%) in MHR using partial least square regression.

Wave number (cm ⁻¹)	Wavelength (nm)	Plot	Band vibration ²¹	Structure ²¹
9327	1072	RC	2 x C–H stretching + 2 x C–C stretching	Benzene
8910–8887	1122–1125	PC1,PC2	C–H stretching second overtone	Aromatic
8316–8285	1202–1207	PC1,PC2,PC3	C–H stretching second overtone	CH ₂
8185	1222	RC	C–H stretching second overtone	CH
7614–7545	1313–1325	PC2,RC,PC3	2 x C–H stretching + C–H deformation	CH ₃
7267	1376	PC1	2 x C–H stretching + C–H deformation	CH ₂
7128–7082	1403–1412	PC2,RC	O–H stretching first overtone	ROH
6981	1432	PC3	N–H stretching first overtone	CONH ₂
6804	1470	PC1	N–H stretching first overtone	CONHR
6727	1487	PC2	N–H stretching first overtone	CONH ₂
6071	1647	PC3	C–H stretching first overtone	R-CH-CH
6017	1662	RC	C–H stretching first overtone	cis-RCH=CHR
5979	1673	PC2	C–H stretching first overtone	Aromatic
5786	1728	PC1,PC3	C–H stretching first overtone	CH ₂
5547	1803	RC	C–H stretching + 2 x C–O stretching	Cellulose
4575	2186	RC	CH ₂ asym.stretching+C=stretching	HC=CH
4459	2243	RC	N–H stretching + NH ₃ ⁺ deformation	Amino acid

Notes: MHR — Milled Hommali rice.

DRW — Dead rice weevils.

RC = Regression coefficient.

PC1 = PLS factor1.

PC2 = PLS factor2.

PC3 = PLS factor3.

Table 3. The absorption bands (from regression coefficient and X-loading weight) with predominantly affected the prediction model of DRW (%) in BHR using partial least square regression.

Wave number (cm ⁻¹)	Wavelength (nm)	Plot	Band vibration ²¹	Structure ²¹
9250	1081	RC	2 x C-H stretching + 2 x C-C stretching	Benzene
8872	1127	RC	C-H stretching second overtone	Aromatic
8224	1216	RC	C-H stretching second overtone	CH ₂
7884–7514	1268–1331	PC1,PC2	2 x C-H stretching + C-H deformation	CH ₃
7228	1383	RC	2 x C-H stretching + C-H deformation	CH ₂
7143–7105	1400–1407	PC3,PC1	O-H stretching first overtone	ROH
6904	1448	RC	2 x C-H stretching + C-H deformation	Aromatic
6835	1463	PC2	N-H stretching first overtone	CONH ₂
6094	1641	PC1	C-H stretching first overtone	
6040–6025	1655–1660	PC3,PC2	C-H stretching first overtone	<i>cis</i> -RCH=CHR
5346–5300	1870–1887	PC2,PC3,PC1	O-H stretching + 2 x C-O stretching	Starch
5145	1940	RC	O-H stretching + O-H deformation	H ₂ O
4829–4822	2071–2074	PC2,RC	O-H stretching + O-H deformation	ROH, Sucrose, Starch
4744–4729	2108–2115	PC1,PC3	N-H sym. stretching + amide III	CONH ₂ , CONHR
4528	2208	RC	N-H stretching + C=O stretching	CONH ₂ , Protein
4335	2307	RC,PC2,PC3	C-H stretchini + C-H deformation	CH ²

Notes: BHR — Brown Hom mali rice.

DRW — Dead rice weevils.

RC = Regression coefficient.

PC1 = PLS factor1.

PC2 = PLS factor2.

PC3 = PLS factor3.

spectrometer. The distinctive values of regression coefficient and X-loading affected the bond vibration on the DRW prediction which was demonstrated in the high peaks in the plots. The peaks of MHR and BHR are reported in Tables 2 and 3, respectively. The dominant wave numbers used in validated model for MHR and BHR could be referred as the peak vibration bonds that responded to the structure of CH₂, CH₃, H₂O, starch, aromatic, cellulose, amino acid and so on. These were the fundamental chemical compounds of rice and weevil. However, the main constituent of rice sample, starch, did not predominantly appear on MHR. Correspondingly, the main component of DRW, chitin (6667 cm⁻¹ and 8489 cm⁻¹), was also not presented. It could be concluded that these components, starch and chitin, did not absorb well the electromagnetic in the range of NIR compared with other constituents. Considering the validated model for BHR, the wave number of H₂O is 5145 cm⁻¹, appeared with high peak. Additionally, from Table 1 of MHR model, the OPUS program selected wave number ranging from 9407–5447 cm⁻¹ and 4746–4421 cm⁻¹ for producing PLS model, which mean it abandoned wave number

band of 5446–4747 cm⁻¹. The relinquished wave number band was mainly the absorption band of starch (shown in Table 3). This confirms the earlier statement that the suitable predicting model of NIR was calculated from statistically complex mathematical methods in which sometimes those selected wavelengths were not related to the chemical origin of measuring materials or unable to explain the source of chemical terms. Therefore, this model can be interfered with the moisture content from two objects, BHR and DRW, and also some absorption bands of starch were not selected to perform the PLS model. This model was, then, not a robust model for predicting.

The scatter plots of the best equation (shown in Table 1) for predicting percentage of DRW in MHR and BHR samples versus the actual percentage of DRW in MHR and BHR samples are shown in Fig. 8. The NIR spectra for PLS model prediction of DRW percentage in MHR were pretreated by SNV technique and in BHR were pretreated by straight line subtraction technique. Most of the predicting samples from validation set were gathered nearby the target line, which implied that using PLS

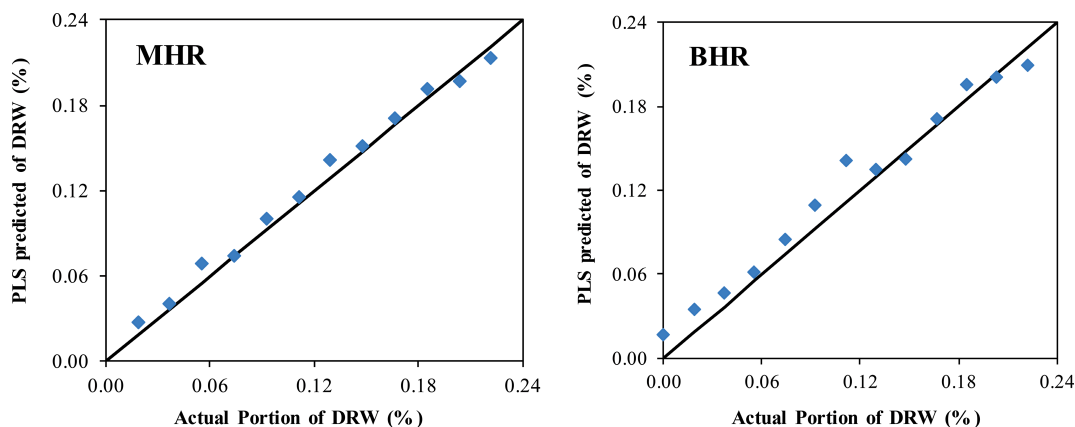


Fig. 8. Scatter plots of prediction percentage of DRW obtained from PLS validation set versus the actual percentage of DRW in MHR and BHR.

Table 4. The predominant wave numbers affect the chemical structures of rice samples and DRW.

Rice samples		DRW	
Wave number (cm ⁻¹)	Structure ²⁴	Wave number (cm ⁻¹)	Structure ^{13,25}
10101	Starch	8850	Lipid
6944	Starch, Sucrose	8489	Chitin
6545	Starch	6667	Chitin
6329	Starch, Glucose	5988	Lipid
5263	Starch		
5000	Starch		
4808	Starch, ROH, Sucrose		
4762	Starch		

Note: DRW — Dead rice weevils.

Table 5. Statistics of prediction of DRW (%) in MHR and BHR by SVD model.

Type of rice	Wave number (cm ⁻¹)	Pre-treatment	R^2_{val}	RMSEP	RPD	Bias
MHR	Rice : 5263, 5000, 4808 DRW : 8850, 8489, 5988	Base line offset	0.97	0.012	5.71	-0.003
BHR	Rice : 5263, 5000, 4808 DRW : 8850, 8489, 6667	SNV	0.96	0.013	5.39	0.002

Notes: MHR — Milled Hommali rice.

BHR — Brown Hommali rice.

DRW — Dead rice weevils.

R^2_{val} — The coefficient of determination of validation set.

RMSEP — Root mean square error of prediction.

RPD — Ratio of standard deviation of reference data in validation set to standard error of prediction.

Bias — The average of difference between actual value and NIR value.

method for predicting DRW in rice sample was possible. The coefficient of determination, R^2 , for MHR and BHR model, therefore, was 0.95 and 0.90 respectively, as reported earlier in Table 1.

3.3. SVD from selected wave number

The wave numbers for substituting their pretreated NIR spectra absorbance in Eq. (1) were selected according to the predominant wave numbers which

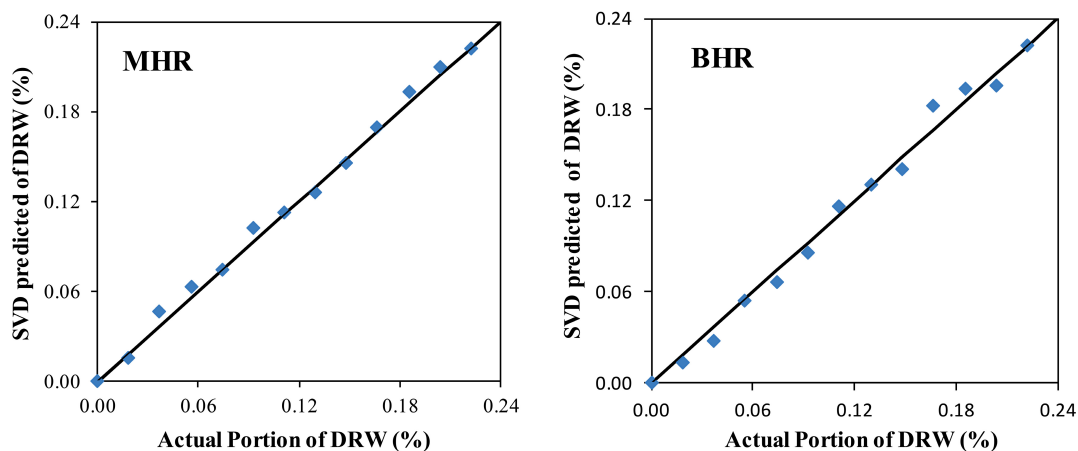


Fig. 9. Scatter plots of prediction percentage of DRW obtained from SVD analysis versus the actual percentage of DRW in MHR and BHR.

affect the band vibration of rice and DRW structures. Rice kernel mainly composes of starch; therefore, the selected wave numbers must be related to the starch molecular structure which was 10,101, 6944, 6545, 6329, 5263, 5000, 4808 and 4762 cm^{-1} . Rice weevil is an arthropod having exoskeleton and called cuticle made of chitin.³¹ The basic arthropod body structure consists of heart, gut, brain, nerve cord and eye, which principally are lipids.³² Consequently, the selected wave numbers for DRW were 8850, 8489, 6667 and 5988 cm^{-1} . Table 4 shows the selected wave numbers of rice sample and DRW are related to their chemical structures. The NIR spectra of individual selected wave numbers were pretreated by the approaches explained in Sec. 2.1 The pretreated NIR spectra from three selected wave numbers of rice and of DRW, six wave numbers in total, were randomly chosen for substituting in matrices $[A]$ and $[b]$ to calculate the proportion of DRW in rice sample matrix $[x]$ by (Eq. (1)). The best prediction models of MHR and BHR were considered from statistical values, R^2 , RMSEP, RPD and bias. Table 5 shows that the best set of wave numbers for predicting percentage of DRW in MHR were 8850, 8489, 5988, 5263, 5000 and 4808 cm^{-1} with base line offset pretreatment. It gave R^2 , RMSEP, RPD and bias of 0.97%, 0.012%, 5.71% and -0.003% respectively. For predicting percentage of DRW in BHR, the best set of wave numbers was 8850, 8489, 6667, 5263, 5000 and 4808 cm^{-1} with SNV pretreatment. The statistical values of R^2 , RMSEP, RPD and bias were 0.96, 0.013, 5.39 and 0.002, respectively. The results showed that the statistical values of the SVD

method were better than that of the PLS methods but not significantly different in terms of statistical ranges. Correspondingly, the scatter plots of the best equation for predicting percentage of DRW in MHR and BHR samples versus the actual percentage of DRW in MHR and BHR samples are shown in Fig. 9. The predicting data were more gathered nearby the target line compared with the PLS method, nevertheless, they indicated that the predicted percentage of DRW in both rice samples was not different from the actual percentage of DRW in both rice samples. Additionally, the potential of using the SVD method, selected only the related wave number with the chemical constituents of rice and DRW, could be easier and cheaper for constructing the handheld equipment for detecting DRW in rice stock because it required less number of light sources and detectors.

4. Conclusion

Prediction model of the DRW in rice sample using SVD technique from NIR signals was performed and the model's performance was compared with PLS technique based on statistical information. The wave numbers used in for SVD analysis technique were selected from the predominantly responded wave number of NIR to the major chemical components of DRW and rice samples (chitin, lipid and starch). Statistical results of employing SVD technique for predicting DRW in rice samples were better than using PLS technique. Therefore, SVD technique can be used as a tool for evaluating the DRW in rice samples with less amount of wave

numbers. This finding, therefore, provided the potential for using only six wave numbers instead of using the full range of NIR wave number for constructing the handheld equipment to operate on the working site with lower production cost when compared while using the NIR spectroscopy.

Acknowledgment

We thank the Agricultural Research Development Agency (ARDA) for financial support for this study.

References

- Office of Agricultural Economics, Thailand foreign agricultural trade statistics 2015, http://www.oae.go.th/oae_report/exportimport/export_result.php (2015).
- Food and Agriculture Organization of the United Nations, Insect pests of stored products, <http://www.fao.org/docrep/x5036e/x5036e0y.htm> (2015).
- N. D. G. White, Insects, mites, and insecticides in stored grain ecosystems, *Stored-Grain Ecosystems*, D. S. Jayas, N. D. G. White, W. E. Muir, Eds., pp. 123–168, Marcel Dekker Inc., New York (1995).
- CODEX International Food Standard, CODEX STAN 198-1995 “Standards for Rice”, <http://codexalimentarius.org> (2013).
- Ministry of commerce of Thailand, National imports-exports product standard: Standard for Thai Hom Mali Rice, <http://www.moc.go.th> (2013).
- W. A. Bruce, M. W. Street Jr., “Infrared CO₂ detection of hidden insects,” *J. Georgia Entomol. Soc.* **9**, 260 (1974).
- W. M. Lamkin, N. C. Unruh, Y. Pomeranz, “Use of fluorometry for the determination of uric acid in grain elimination of interfering fluorescence,” *J. Cereal Chem.* **68**(1), 81–86 (1991).
- C. Karunakaran, D. S. Jayas, N. D. G. White, “Detection of internal wheat seed infestation by *Rhizopertha dominica* using X-ray imaging,” *J. Stored Prod. Res.* **40**(5), 507–516 (2004).
- T. C. Pearson, A. E. Cetin, A. H. Tewfik, R. P. Haff, “Feasibility of impact-acoustic emission for detection of damaged wheat kernels,” *J. Digit. Signal Process.* **17**, 617–633 (2007).
- H. Zhang, J. Wang, “Detection of age and insect damage incurred by wheat, with an electronic nose,” *J. Stored Prod. Res.* **43**(4), 489–495 (2007).
- S. S. Kim, M. R. Phyu, M. J. Kim, S. H. Lee, “Authentication of rice using near infrared reflectance spectroscopy,” *J. Cereal Chem.* **80**(3), 346–349 (2003).
- P. Cheewapramong, “Use of Near-Infrared Spectroscopy for Qualitative and Quantitative Analyses of Grains and Cereal Products,” Dissertations & Theses in Food Science and Technology, Food Science and Technology Department, University of Nebraska (2007).
- F. E. Dowell, J. E. Throne, J. E. Baker, “Automated nondestructive detection of internal insect infestation of wheat kernels using near-infrared reflectance spectroscopy,” *J. Econ. Entomol.* **91**, 899–904 (1998).
- F. E. Dowell, J. E. Throne, D. Wang, J. E. Baker, “Identifying stored-grain insects using near-infrared spectroscopy,” *J. Econ. Entomol.* **92**, 165–169 (1999).
- J. E. Baker, F. E. Dowell, J. E. Throne, “Detection of parasitized rice weevils in wheat kernels with near-infrared spectroscopy,” *J. Biol. Control.* **16**, 88–90 (1999).
- E. B. Maghirang, F. E. Dowell, J. E. Baker, J. E. Throne, “Automated detection of single wheat kernels containing live or dead insects using near infrared reflectance spectroscopy,” *Trans. ASAE* **46**, 1277–1282 (2003).
- P. Jarruwat, P. Choomjaihan, “Feasibility study on estimation of rice weevil quantity in rice stock using near-infrared spectroscopy technique,” *J. Innov. Opt. Health. Sci.* **7**(4), 145001, [8 pages] (2014).
- W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge, (1988).
- G. M. Campbell, C. Fang, I. I. Muhamad, “On predicting roller milling performance VI Effect of kernel hardness and shape on the particle size distribution from First Break Milling of wheat,” *Trans. IchemE* **85**(C1), 7–23 (2007).
- P. Choomjaihan, “Extending the Breakage Equation for First Break milling of wheat to include particle composition,” Ph.D. thesis in the Faculty of Engineering and Physical Science, School of Chemical Engineering and Analytical Science, University of Manchester, England (2008).
- I. N. Koprinarov, A. P. Hitchcock, C. T. McCrory, R. F. Childs, “Quantitative mapping of structure polymeric systems using singular value decomposition analysis of soft X-ray image,” *J. Phys. Chem.* **106**(21), 5358–5364 (2002).
- M. Bekara, M. V. D. Baan, “Local singular value decomposition for signal enhancement of seismic data,” *Geophysics* **72**(2), V59–V65 (2007).
- M. Tang, X. Wu, M. Cong, K. Guo, “A method based on SVD for detecting the defect using the

- magnetostrictive guided wave technique,” *Mech. Syst. Signal Process.* **70–71**, 601–612 (2016).
24. B. G. Osborne, T. Fearn, P. H. Hindle, *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. John Wiley and Sons, New York (1993).
 25. J. Perez-Mendoza, J. E. Throne, F. E. Dowell, J. E. Baker, “Detection of insect fragments in wheat flour by near-infrared spectroscopy,” *J. Stored Prod. Res.* **39**, 305–312 (2003).
 26. R. J. Barnes, M. S. Dhanoa, S. J. Lister, “Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra,” *J. App. Spec.* **43**(5), 772–777 (1989).
 27. R. Lapchareonsuk, P. Sirisomboon, “Eating quality of cooked rice determination using Fourier transform near infrared spectroscopy,” *J. Innov. Opt. Health. Sci.* **7**(6), 8 (2014), doi: 10.1142/S1793545814500035.
 28. K. Kaewson, P. Sirisomboon, “Study on evaluation of gamma oryzanol of germinated brown rice by near infrared spectroscopy,” *J. Innov. Opt. Health. Sci.* **7**(4), 7 (2014), doi: 10.1142/S1793545814500023.
 29. F. Xianshu, Y. Xiaoping, Y. Zihong, C. Haifeng, “Analysis of Antioxidant Activity of Chinese Brown Rice by Fourier-Transformed Near Infrared Spectroscopy and Chemometrics,” *J. Chem.* **2015**, 5 (2015) Article ID 379327.
 30. P. Williams, *Near-Infrared Technology Getting the Best Out of Light*, pp. 5–16, PDK Grain Nanaimo, British Columbia, and Winnipeg, Manitoba, Canada (2007).
 31. M. Locke, “The Wigglesworth Lecture: Insects for studying fundamental problems in biology,” *J. Insect Physiol.* **47**(4–5), 495–507 (2001).
 32. I. A. Sani, S. S. Ahmed, *Arthropodes & Classification of Class Insect*, 1st Edition, p.85, Lulu publisher, US (2015).