Final Report

Project Title:

Rice grain freshness measurement using rapid viscosity analyzer (RVA) and chemometrics

By Assoc. Prof. Dr. Sila Kittiwachana

May 2561

Contract No. MRG5980245

Final Report

Project Title:

Rice grain freshness measurement using rapid viscosity analyzer (RVA) and chemometrics

Researchers	**Institute**
Assoc. Prof. Dr Sila Kittiwachana	Department of Chemistry
Asst. Prof. Dr. Jaroon Jakmunee	Faculty of Science
	Chiang Mai University

This project granted by the Thailand Research Fund

Abstract

This research developed a rapid and accurate method based on the use of rapid visco analyzer (RVA) for predicting the storage time of rice grain. Freshly harvested rice samples, five waxy and five non-waxy rice grains, were stored in paddy form at ambient room temperature (28-32 °C) for one year. During storage, the RVA profiles of the rice samples were recorded every month. In addition, physicochemical properties, such as alkali spreading value (ASV), amylose content, gel consistency, stickiness and hardness, were measured. Chemometric models including partial least squares (PLS) regression and supervised self-organizing map (supervised SOM) were employed for predicting the storage time based on the use of the RVA profiles, the physicochemical parameters and both of the data sets simultaneously. In most cases, PLS outperformed supervised SOM. The PLS models established using the RVA profiles provided the best predictive results with root mean square error of cross validation (RMSECV) = 1.2, cross-validated explained variance (Q^2) = 0.90 and the ratio of prediction to deviation (RPD) = 3.2. Based on partial least squares-variable influence on projection (PLS-VIP), pasting properties, including peak viscosity (PV) and final viscosity (FV), were identified as the parameters having strong influence on the prediction models. The developed method detecting the rheological change of the stored rice samples was simple and could be performed quickly with no additional chemicals required.

Project Code : MRG5980245

Project Title : Rice grain freshness measurement using rapid viscosity analyzer (RVA) and chemometrics

Investigator : Assoc. Prof. Dr. Sila Kittiwachana

E-mail Address : silacmu@gmail.com

Project Period : 2 years

Keywords : rice freshness; aged rice; rapid visco analyzer (RVA); pasting property; chemometrics

Rice grain freshness measurement using rapid visco analyzer and chemometrics

1. Introduction

Rice is an important food for people in many countries. Many Asians consume rice as an essential part of everyday meals. However, in many parts of the tropics, rice can only be harvested once a year due to the environmental condition. Also, many rice cultivars are photosensitive plants meaning that they have a limited growing season during the year. Therefore, rice grain should be stored in a barn or warehouse after harvested. During storage, rice properties such as head rice yield, pasting, volume expansion and water adsorption could change (Keawpeng and Venkatachalam 2015). This causes alternation in the cooking quality of rice, in particular, its flavor and texture. For example, cooked rice prepared from aged rice has a harder texture (Kanlayakrit and Maweang 2013) and the ratios of stickiness to hardness are low in comparison to those of fresh rice grains (Keawpeng et al. 2015). Often, the properties of rice as a raw material play an important role in guaranteeing desirable product quality during manufacturing (Fang et al. 2015). Rice grain freshness, especially, plays a key role in product quality. Commercially, knowing the exact time that the rice has been stored is required for evaluating the price of the rice products.

Rice grain freshness can be estimated by measuring the changes in some related chemical and physical properties. Fats, such as neutral lipids, bran lipids and phospholipids, are among the important components of rice grains (Srikaeo and Panya 2013). After harvest, these fats, in particular the neutral lipids, start to decompose, resulting in free fatty acids. When the ageing process prolongs, an amount of the fatty acids could be obtained. Therefore, the storage time of rice grain could be estimated by measuring the amount of these fatty acids using an acid-base indicator, such as bromothymol blue (Srikaeo et al. 2013). Chuang et al. (2014) demonstrated the use of near infrared (NIR) technique combined with independent component analysis (ICA) for evaluating rice grain freshness as quantified by pH values. Although the acidity detection method is simple, the fatty acids in the rice grains next decompose into hexanal. With the fatty acids decreasing after prolonged storage, this could confuse the test results. Besides the measurement of the rice acidity, peroxidase, one of the important enzymes involved in reactive-oxygen species scavenging in the rice seed, can also be used as an indicator for estimating rice freshness (Srikaeo et al. 2013). Since the peroxidase enzyme deteriorates during storage, it is possible to estimate the storage time by determining the activity of this enzyme (Chen and Chen 2003). The methods that rely on measuring the presence of the end products of the deteriorating chemicals in rice may not actually reflect the status of rice quality, in particular, its texture quality. Therefore, it was recommended that some other techniques, such as texture analysis or human sensory tests, should be performed in combination when analyzing rice grain freshness (Srikaeo et al. 2013). Recently, colorimetric sensor array was used to discriminate the rice samples with different storage times based on linear discriminate analysis (LDA) model (Guan et al. 2017). It was possible to distinguish the difference between the rice samples with the storage ages of 6 and 12 months.

Rapid visco analyzer (RVA) is a rheological testing tool that continuously measures viscosity of a viscous sample under different applied temperatures (Tong et al. 2014). RVA has been utilized for a variety of agricultural products, such as rice (Vongsawasdi et al. 2009), soya (Kaur et al. 2011) and potato (Leivas et al. 2013).

Kanlayakrit et al. (2013) used RVA for observing the changes in paddy and milled rice stored in different storage conditions. In rice stored for several months, a decreasing trend in hydration and viscosity could be noticed (Teo et al. 2000). Zhou et al. (2002) also observed the changes in several pasting properties during storage, including, the increases in grain hardness and peak viscosity. Although the pasting parameters help reveal the characteristics of the rice texture, a methodology to interpret the RVA profiles in a meaningful way to reflect rice grain freshness has not been established. Chemometrics is a multivariate analysis that utilizes mathematics and statistics to extract the relevant information from chemical data (Brereton 2003). Chemometrics can be useful in a variety of situations, including data exploration, experimental design, pattern recognition/classification and multivariate calibration (Brereton 2009). Multivariate calibration methods, such as partial least squares (PLS) regression (Geladi and Kowalski 1986) and supervised self-organizing map (supervised SOM) (Kittiwachana et al. 2013) can be used to relate the values of predictive data to the studied response. It has been proven that PLS is among the most powerful multivariate linear calibrations (Funsueb et al. 2016). In contrast, supervised SOM is a nonlinear prediction. Compared to other nonlinear methods, SOMs better preserve the local data structure. Complex relationships between the samples can be revealed through the model structures and this allows the behavior of the studied parameters to be examined (Siripatrawan and Harte 2015).

This research project aimed to develop a new method to evaluate the freshness of stored rice grains based on the use of RVA and chemometrics. The rice samples, waxy and non-waxy rice, were stored at an ambient room temperature. The RVA profiles of the rice samples were recorded every month for one year. As representatives for linear and nonlinear calibration models, PLS and supervised SOM were utilized to extract the trend in the RVA profiles in relation to the storage time. In comparison, some physicochemical parameters such as alkali spreading value (ASV), amylose content, gel consistency, stickiness and hardness were also recorded.

2. Objectives

1. To investigate the relationship between RVA profile data and some rice freshness-related parameters such as storage time, hardness and stickiness of cooked rice, amylose content, ASV and gel consistency using chemometrics.

2. To develop chemometric models that can be used to estimate the rice grain freshness of the stored rice grains.

3. Materials and methods

3.1 Samples

Ten cultivars of Thai white rice were used in this study. Five of these cultivars were selected as representatives of waxy (glutinous) rice and the rests were non-waxy rice. These selected rice samples were a part of the development program within the Ubon Ratchathani Rice Research Center, Ubon Ratchathani, Thailand. All the rice plants were transplanted at the same period in July 2013 at the research area. The rice samples were harvested when most of the panicles were ripe (approximately 80% of the panicles turn golden yellow and hard) in November 2013. They were then dried in the sun until the grain moisture content was 14%. After that, the rice grains in paddy form were stored in sealed plastic containers at ambient room temperature of approximately 30 °C.

3.2 RVA analysis

Before RVA analysis, the stored paddy rice grains were dehulled and then ground into powder by Cyclotec Sample Mill (Cyclotec™ 1093, China) with a 100 mesh sieve. After that, 3.00 g of the sample (adjusted to 14% moisture content) was mixed in 25.0 ml distilled water in an aluminum RVA canister. The samples were analyzed using a RVA (RVA-TecMaster, Perten Instruments of Australia Pty Limited). The viscosity profile was recorded with a starting temperature of 50 °C, which was held for 1.12 min. After that, the temperature was raised to 95 °C in 3.40 min, kept for 2.28 min and finally cooled to 50 °C in 4.00 min, and then held for 1.10 min. The total run time was 12.30 min. The stirring speed was 960 rpm for the starting period (10 s) and 160 rpm for the rest of the test time. The RVA condition was adopted from the American Association of Cereal Chemists (AACC) International's approved method (AACC 2000) suggested for rice samples. The analysis procedure is illustrated in Fig 1. The measurement was repeated every month for one year, resulting in a total of thirteen RVA profiles for each sample. The test was done in duplicate and the average values were used.

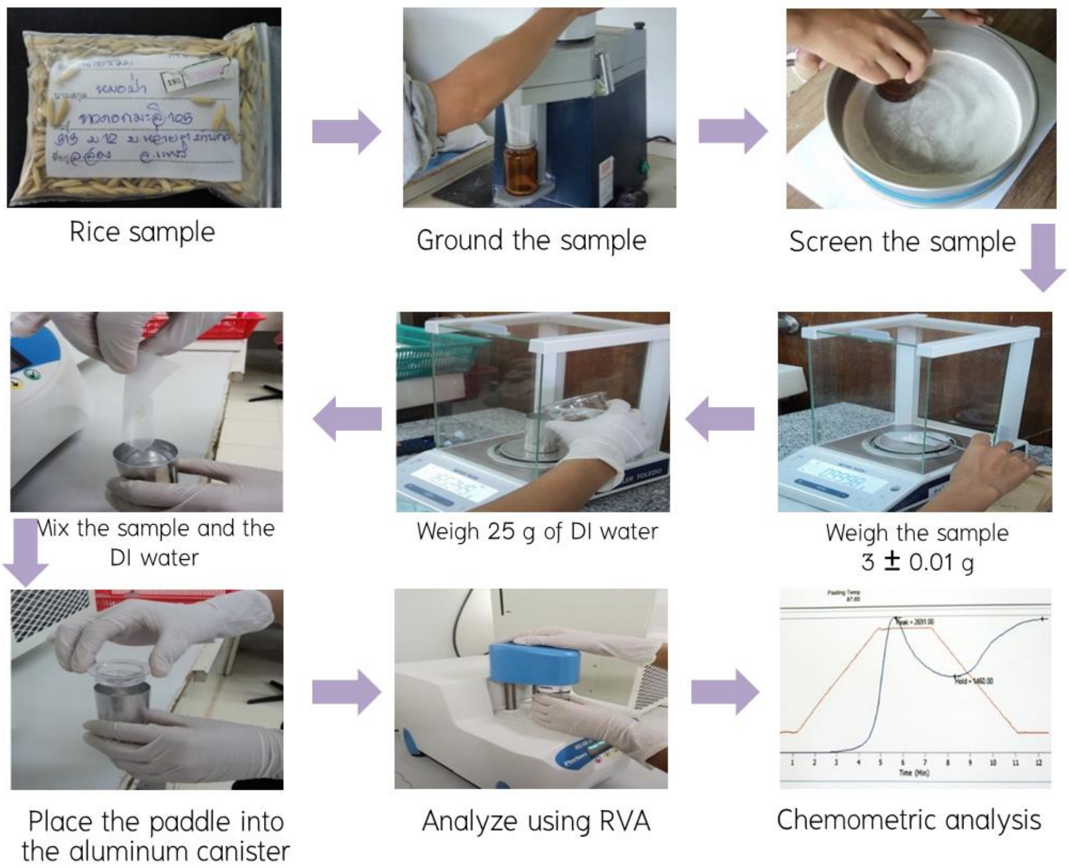

Fig. 1. Analysis procedure of the RVA measurement.

3.3 Physicochemical analyses

During the experimental period, physicochemical parameters of the stored rice, including alkali spreading value (ASV), amylose content, gel consistency, stickiness and hardness, were also recorded. ASV of the rice samples was investigated using the method presented by references (Mariotti et al. 2010; Attaviroj and Noomhorm 2014). Stickiness and hardness were measured using texture analyzer (TA.XTplus, Charpa Techcenter Co., Ltd., Bangkok, Thailand) according to the method reported by Li et al. (2016). Amylose content was determined based on the reaction between amylose and iodine solution following the procedure described by Singh et al. (2000). The gel consistency of the rice samples was also determined following the procedure of

Cagampang et al. (1973). In this research, both waxy and non-waxy rice was used. Normally, non-waxy rice has relatively higher amylose content. On the other hand, amylose content can be very low or cannot be detected in waxy rice (Wani et al. 2012). In this research, the amylose starch could not be detected in the waxy cultivars. In addition, the gel consistency test, an indirect method, can be used for screening the hardness of cooked rice. Without modification, this method is preferably suitable for soft rice with high amylose content particularly the non-waxy rice. Therefore, amylose content and gel consistency were only analyzed for the non-waxy rice samples and reported.

3.4 Chemometric analyses

In this research, the RVA profiles and the rice quality parameters, defined as $\boldsymbol{X}_{\mathrm{RVA}}$ and $\boldsymbol{X}_{\mathrm{Qual}}$, respectively, were used as predictive parameters and the storage time ($\boldsymbol{c}$) was used as response. Partial least squares (PLS) regression and supervised self-organizing map (supervised SOM) were the prediction models used in this investigation to represent linear and nonlinear multivariate calibrations.

3.4.1 Principal component analysis (PCA)

PCA is a common exploratory analysis technique in chemometrics (Brereton 2003). PCA defines new factors that effectively represent as much as possible the variation in a dataset. These defined factors are usually called principal components (PCs). Using PCA, pattern in multivariate data can be reviewed. The mathematical transformation of a data matrix $\boldsymbol{X}$ by PCA is:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}' + \boldsymbol{E}$$

where $\boldsymbol{T}$ is a scores matrix, $\boldsymbol{P}$ is a loadings matrix, and $\boldsymbol{E}$ is a residual matrix. The PC model is a product matrix of $\boldsymbol{TP'}$ containing the systematic variation. Often, the first few PCs could explain most of the important variation of the dataset. The relationship among the studied samples can be revealed using the scores ($\boldsymbol{T}$). On the other hand, the relationship or the behavior among the studied parameters can be examined using the loadings ($\boldsymbol{P}$).

3.4.1 Partial least squares (PLS) regression

Many versions of the PLS algorithms have been proposed, in this research, the PLS1 algorithm reported in Brereton (2003) was used. PLS captures variations from both the predictive ($\boldsymbol{X}$) and response ($\boldsymbol{c}$) parameters and simultaneously uses them for constructing the regression model. This can be written as two equations:

$$\boldsymbol{X} = \boldsymbol{TP}' + \boldsymbol{E}$$

$$\boldsymbol{c} = \boldsymbol{Tq}' + \boldsymbol{F}$$

Using the non-linear iterative partial least squares (NIPALS) algorithm, $\boldsymbol{X}$ is decomposed into $\boldsymbol{X}$-scores ($\boldsymbol{T}$) and $\boldsymbol{X}$-loadings ($\boldsymbol{P}$). As the same time, $\boldsymbol{c}$ is the product approximation of $\boldsymbol{T}$ and $\boldsymbol{c}$-loadings ($\boldsymbol{q}$). The aim of the PLS algorithm is to minimize the norm of $\boldsymbol{F}$. To predict the response of unknown sample $\boldsymbol{X}_{\text{test}}$, the following equation is used:

$$\hat{\boldsymbol{c}}_{test} = \boldsymbol{X}_{test}\boldsymbol{Wq}'$$

where $\boldsymbol{W}$ refers to normalized PLS weights. PLS, in most cases, could satisfactorily provide predictive results if the data follows a multivariate normal distribution. In addition to the predictive model, the importance of the predictive parameters corresponding to the studied response can be evaluated using partial least squares-variable influence on projection (PLS-VIP) (Funsueb et al. 2016). The PLS-VIP values represent the influence of $\boldsymbol{X}$ variables on $\boldsymbol{c}$ variance and can be calculated from the PLS weights obtained from the PLS model (Farrés et al. 2015). The VIP score greater than one is typically used as a criterion for variable selection (Tran et al. 2014).

3.4.2 Supervised self-organizing map (supervised SOM)

A self-organizing map (SOM) or Kohonen network is one of the most well-known artificial neural networks (Kohonen 1982). Whereas PLS assumes that data follow a multivariate normal distribution, SOM can be constructed without assuming any mathematical functions. In other words, it is a nonlinear method. Compared to other nonlinear models, SOM has an advantage in that insight into complex relationships of data can be revealed using visualization methods such as U-matrix, supervised color shading and component planes (Lloyd et al. 2008). This allows the behavior of the studied parameters to be examined. In this work, SOM, in supervised mode, was applied using the algorithm presented in reference (Kittiwachana et al. 2013). In supervised SOMs, the measurements $\boldsymbol{X}$ are used together with the response vector $\boldsymbol{c}$ in the training process by adding an additional response vector to the measurement matrix. Then, the initial weight matrix $\boldsymbol{W}$ can be trained in the same manner as for traditional unsupervised SOM learning. After that, the Euclidean distance or the dissimilarity between the unknown sample $\boldsymbol{x}_i$ and the weight vector $\boldsymbol{w}_k$ can be calculated by:

$$s_{(\boldsymbol{x}_i, \boldsymbol{w}_k)} = \sqrt{\sum_{j=1}^{J_s - 1} (x_{ij} - w_{kj})^2}$$

where J_s is the number of columns in the trained weight matrix $\boldsymbol{W}$ for the supervised SOM, which is equal to the number of the supervised component planes (J_s = J +1). It is noted that J is the number of parameters used in the study. After that, the prediction for an unknown sample $\boldsymbol{x}_i$ can be determined by identifying the best matching unit (BMU) on the trained map:

$$s_{(\mathbf{x}_i,\mathbf{w}_b)} = \min_k \{s_{(\mathbf{x}_i,\mathbf{w}_k)}\}$$

where b is the value of k having the most similar weight vector $\boldsymbol{w}_k$ to the unknown sample $\boldsymbol{x}_i$:

$$b = \arg\min_k \{s_{(\mathbf{x}_i,\mathbf{w}_k)}\}$$

Therefore, the prediction of an unknown sample $\boldsymbol{x}_i$ is

$$\hat{c}_i = w_{bJ_s}$$

where $\hat{c}_i$ is a predicted response of the unknown sample $\boldsymbol{x}_i$ which is equal to the weight value in the response plane w_{bJ_s} .

For a calibration model, different data pre-processing can strongly influence the model predictive performance and several studies have reported guidelines as to which data pre-processing should be used or avoided. For example, standardization or autoscaling can be used to ensure that all the variables are put on approximately the same scale (Brereton 2009). Row scaling or normalization can be useful if the absolute concentrations of samples are important and cannot be easily controlled (Kittiwachana et al. 2008). Element scaling such as square root and logarithmic transformations can be used to reduce asymmetry or heteroscedastic noise. They also help to ease the effect from large peaks unduly influencing the signal (Kittiwachana et al. 2010). Mean centring can be calculated by subtracting the mean of each variable from the raw data.

After mean centring, the sum of each column is zero. This allows the models to focus on the deviations from the mean and investigate variance from the data mean rather than the absolute values (Brereton 2003). Mean centring has successfully provided the optimum results for various systems (Kim et al. 2001; Postma et al. 2011; Kittiwachana et al. 2008). In this research, mean centring resulted in the best performing model and, therefore, this data pre-processing was used as a standard protocol for our dataset. However, if the RVA ($\boldsymbol{X}_{\mathrm{RVA}}$) and the rice quality ($\boldsymbol{X}_{\mathrm{Qual}}$) data were simultaneously used for the prediction, they were standardized by dividing each element by its standard deviation after mean centering to ensure that all the variables were adjusted onto the same scale. This was to guarantee that all variables had equal influence on the model.

3.4.3 Model validation (quality of prediction)

The accuracy of prediction was measured using root mean square error of calibration (RMSEC) and root mean square error of cross validation (RMSECV). RMSEC is the average difference between predicted ($\hat{c}_i$) and expected (c_i) response values in auto-prediction mode and can be calculated as:

$$\mathrm{RMSEC} = \sqrt{\frac{\sum_{i=1}^{N}(\hat{c}_i - c_i)^2}{N-1}}$$

where N is the number of samples. Using RMSEC, the establish model is tested directly on the calibration data or training samples, thus it is an internal validation. On the other hand, RMSECV calculates the error of the predicted response values ($\hat{c}_i$) based on cross validation. In this research, the calibration models were built for each rice type, waxy and non-waxy rice, where each model consisted of five cultivars. With cross validation methodology, a segment of five samples having the same storage age was selected for deletion. Then, the calibration model was developed using the remaining samples and the calibration performance was tested using the deleted samples. This process was repeated until all samples were used as test samples. The segmented criterion was designated for avoiding the overfitting problem when the only one sample was left out. The ratio of RMSECV and RMSEC ($R_{CV/auto}$) was calculated to indicate the model robustness. If this ratio is close to 1, this shows a stability of model although there is the removal of the samples from the data. The cross-validated explained variance (Q^2) was also calculated to evaluate the prediction performance of the model (Consonni et al. 2009) by:

$$Q^2 = 1 - \frac{\sum_{i=1}^{N} (\hat{c}_i - c_i)^2}{\sum_{i=1}^{N} (c_i - \bar{c})^2}$$

Values of Q^2 as close as possible to 1.0 are desired and imply the greater degree of variation within the data modelled by the calibration model. In addition, Bias or the systematic error of the average difference between the actual and predicted values and were calculated as:

$$\text{Bias} = \sum_{i=1}^{N} \frac{(\hat{c}_i - c_i)}{N}$$

In this case, $\hat{c}_i$ is the prediction of the deleted sample. From the Bias value, the bias-corrected standard error of the cross validation or (SECV) can be calculated as follows:

$$\text{SECV} = \sqrt{\frac{\sum_{i=1}^{N} (\hat{c}_i - c_i - \text{Bias})^2}{N-1}}$$

The ratio of prediction to deviation (RPD) was used to standardized the predictive accuracy and was calculated as the ratio of the standard deviation of the reference values (SD) and the root mean square of cross validation (RPD = SD/RMSECV). The relative standard deviation (RSD) was calculated according to formula:

$$\mathrm{RSD} = \frac{\mathrm{RMSECV}}{\mathrm{Mean}} \times 100$$

where Mean is the average of the storage time. In this research, the calculations of PLS, supervised SOMs and the statistical analyses were implemented using in-house MATLAB scripts (MATLAB V7.0, The Math Works Inc., Natick). The SOM map consists of 8×10 map units and the SOM parameters such as iteration number, initial learning rate and initial neighborhood width; these were set following recommended methodology (Lloyd et al. 2008).

4. Results and discussion

4.1 RVA profiles and physicochemical properties of the rice samples

Fig. 2(A) shows the RVA profiles of the rice samples used in this research. The difference between pasting curves of the waxy (cultivar 1-5) and non-waxy (cultivar 6-10) rice can be observed. The peak viscosity (PV) of the waxy rice occurred sooner than that of the non-waxy rice. This could be that pasting happened immediately after gelatinization in waxy rice, whereas, for non-waxy rice, a delay between gelatinization and pasting could be observed (Crosbie and Ross 2007). Normally, non-waxy or non-glutinous rice is high-amylose starch with a lower amount of amylopectin. On the contrary, amylopectin is the main component of waxy rice. During the retrogradation process, the recrystallization rate of amylopectin molecules was slower than that of the amylose (Wang et al. 2015). Therefore, in the end of the test, the final viscosity (FV) of waxy rice was lower than that of the non-waxy rice. The score plot of the first two principal components (PCs) is presented in Fig. 2(B). The rice samples can be clearly separated into two main clusters. The left-hand side cluster belongs to the non-waxy rice and the waxy rice is located in the right-hand side region, highlighting the difference between the RVA profiles of the waxy and non-waxy rice.

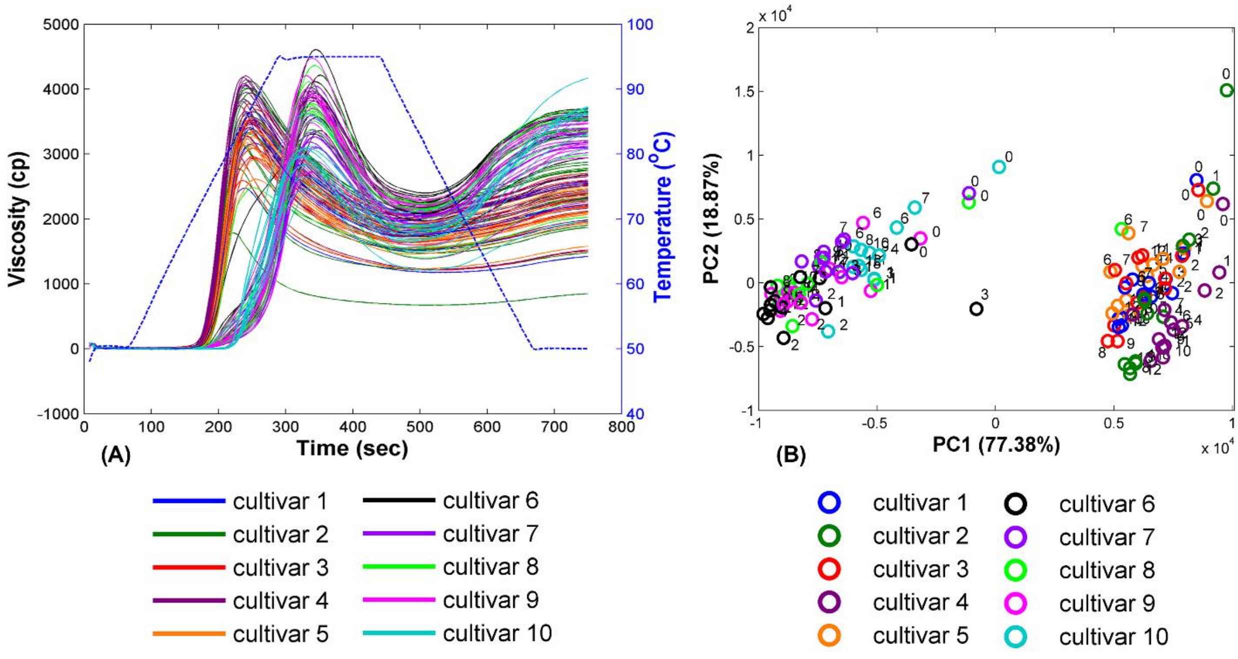

Fig. 2. (A) RVA profiles of the rice samples where the blue dotted line indicates the temperature profile and (B) score plot of the first two PCs where the numbers indicate the storage time starting from 0 (the rice before storage) to 12 (the rice with the storage time of 12 months).

In most cases, the RVA profiles from the same rice variety are located on the same region in the PCA score plot, implying that the pasting properties from the rice with the prolonged time shared the similar characteristics.

The physicochemical parameters recorded during storage, including amylose content, ASV, gel consistency, stickiness and hardness are presented in Fig. 3. In Fig. 3(A), it can be seen that the amylose content did not change significantly during storage. These results were consistent with Kanlayakrit et al. (2013). Also, the changes in ASV (Fig. 3(B)) and gel consistency (Fig. 3(C)) were not much observed. Although, some hydrolysis or degradation probably occurs during storage, leading to a proportional increase in reducing sugars and a decrease in non-reducing sugar and starch, these chemical changes were very scant (Zhou et al. 2002). Some enzyme activities could cause the changes in some main components of the storage rice such as oryzenin, amylose and amylopectin, but the changes only in their average molecular weights were mainly observed.

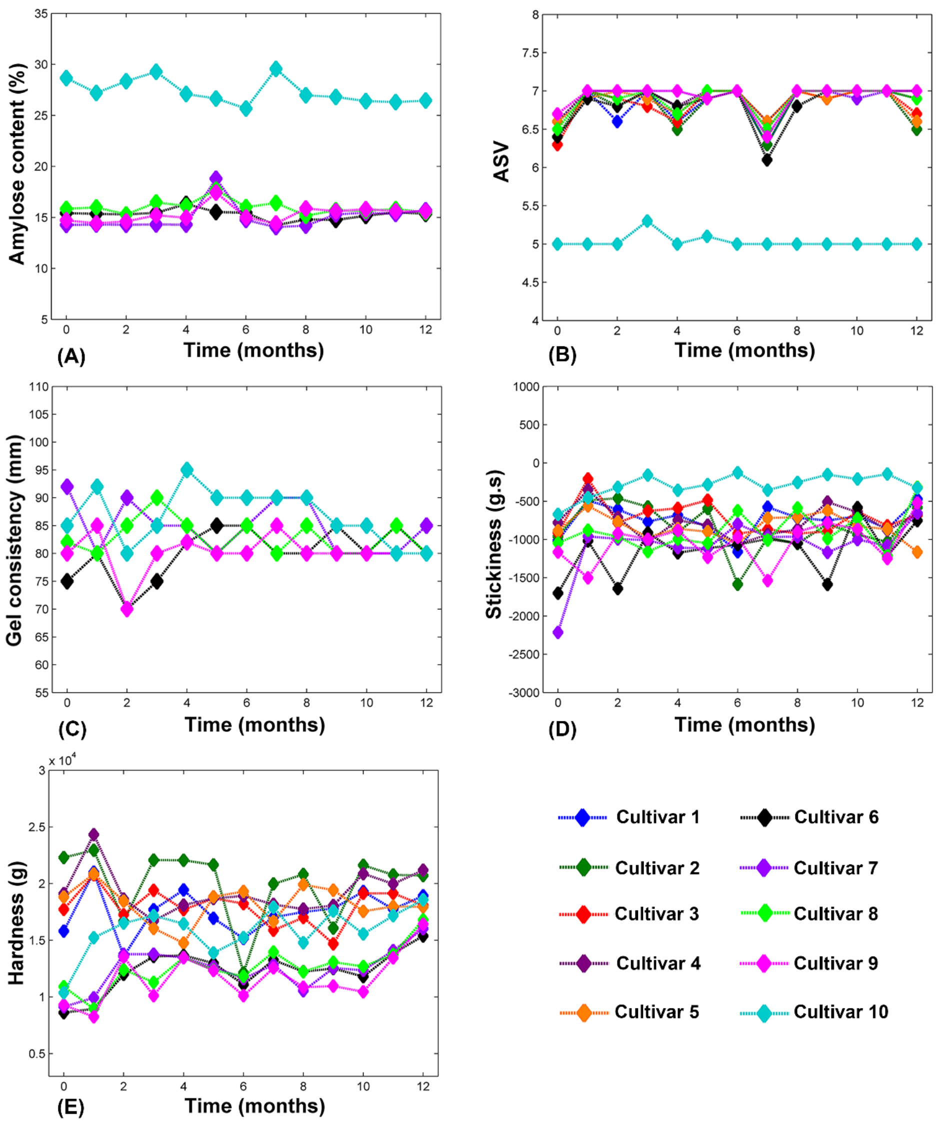

Fig. 3. (A) amylose content, (B) ASV, (C) gel consistency, (D) stickiness and (E) hardness of the rice samples recorded during storage.

Therefore, the average composition, as provided by total protein, starch and lipids, changed very little during the storage period (Chrastil 1990). Slight changes in stickiness (Fig. 3(D)) and hardness (Fig. 3(E)) could be observed in some cultivars, but the trends were inconsistent.

Rice ageing is a complex process and the different varieties might possess different changing characteristics. Although some rice properties, such as stickiness and hardness, have been reported to change during storage in relatively high temperature environments (Zhou et al. 2007), the temperature in our study was maintained at room temperature; therefore, trends in the changes of the recorded parameters were not clearly observed.

4.2 Prediction using PLS and supervised SOM

Table 1 shows the predictive results using PLS and supervised SOM established using the RVA data ($\boldsymbol{X}_{RVA}$), the quality-related parameters ($\boldsymbol{X}_{Qual}$) and both data sets ($\boldsymbol{X}_{RVA}$; $\boldsymbol{X}_{Qual}$). In this research, the optimum numbers of latent variables (LVs) for the PLS models were investigated using cross validation methodology. Due to the characteristic difference in the RVA profiles, the calibration models were built separately for waxy and non-waxy rice. Considering the models using different predictive data, the PLS model using the RVA profiles ($\boldsymbol{X}_{RVA}$) resulted in the best calibration performance with the RMSEC, RMSECV and Q^2 values of 1.1, 1.2 and 0.90, respectively. The corresponding correlation graphs between the expected and the predicted storage times for the PLS models are shown in Fig. 4. A small value of RMSEC indicated that the PLS adequately fitted the data. The RMSECV of 1.2 suggested that the model was suitable for predicting the storage age of rice. Only small difference between RMSEC and RMSECV can be observed and the value of $R_{CV/auto}$ is close to 1 implying that the model was stable and not prone to the overfitting problem. The Q^2 value of 0.90 indicated that 90% of the variation within the response parameter was modelled by the calibration model. The $R_{CV/auto}$ values from supervised SOMs is relatively higher than those from the PLS models especially when the concatenated data ($\boldsymbol{X}_{RVA}$; $\boldsymbol{X}_{Qual}$) was used. This means that the SOM models fitted very well with the training data, but not with the test samples implying that the supervised SOM models were slightly prone to overfitting.

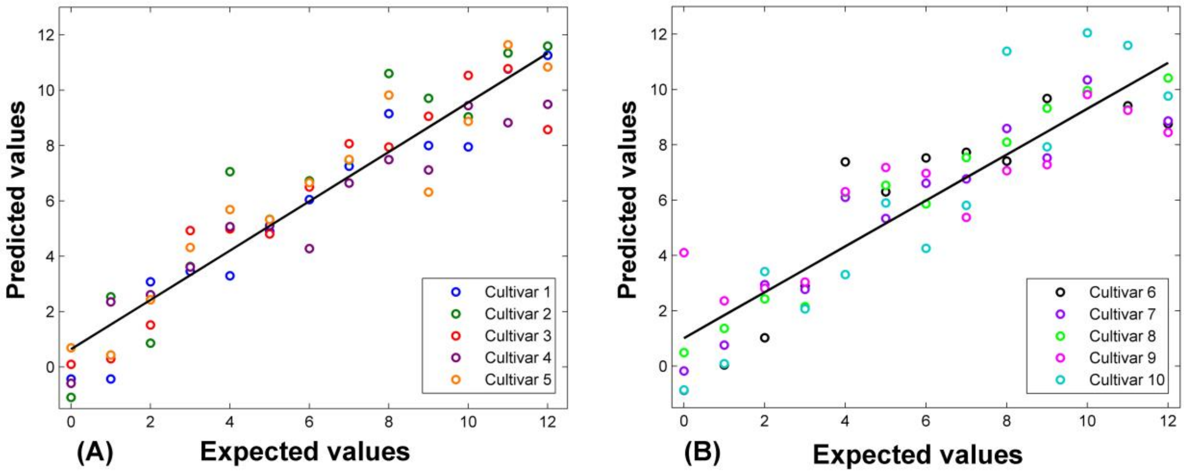

Fig. 4. Correlation graphs between expected and predicted storage times of PLS prediction of (A) waxy and (B) non-waxy rice.

The worst predictive results were obtained from the models established using the quality-related data ($\boldsymbol{X}_{Qual}$) for both the prediction using PLS and supervised SOM, resulting in relatively higher RMSECV ranging from 2.9 to 3.6. This implied that the quality-related parameters were not very much related to the storage time. The results were confirmed by the increase in the RMSECV values from the models when both data sets were concatenated and simultaneously used for the model prediction. A slightly deviation of the amylose contents on the 7th month and the ASV values on the 5th month can be observed.

Table 1 Statistical analysis of the calibration models using PLS and supervised SOM.

PLS						
Calibration statistics	Waxy rice			Non-waxy rice		
	$\boldsymbol{X}_{RVA}$	$\boldsymbol{X}_{Qual}$	$[\boldsymbol{X}_{RVA}; \boldsymbol{X}_{Qual}]$	$\boldsymbol{X}_{RVA}$	$\boldsymbol{X}_{Qual}$	$[\boldsymbol{X}_{RVA}; \boldsymbol{X}_{Qual}]$
RMSEC	1.1	3.6	1.2	1.4	2.9	1.6
LVs*	6	3	5	5	3	5
RMSECV	1.2	3.7	1.3	1.5	3.4	1.7
$\mathbf{R_{CV/auto}}$	1.1	1.0	1.1	1.1	1.2	1.1
$\mathbf{Q^2}$	0.90	0.07	0.87	0.83	0.16	0.81
SECV	1.23	3.65	1.35	1.55	3.06	1.68
Bias	0.01	0.01	0.01	0.01	0.02	0.01
RPD	3.2	1.1	2.9	2.5	1.3	2.3
RSD(%)	17.5	52.1	19.3	22.1	43.8	24.1
Supervised SOM						
Calibration statistics	Waxy rice			Non-waxy rice		
	$\boldsymbol{X}_{RVA}$	$\boldsymbol{X}_{Qual}$	$[\boldsymbol{X}_{RVA}; \boldsymbol{X}_{Qual}]$	$\boldsymbol{X}_{RVA}$	$\boldsymbol{X}_{Qual}$	$[\boldsymbol{X}_{RVA}; \boldsymbol{X}_{Qual}]$
RMSEC	1.2	3.2	1.7	1.9	2.8	1.5
RMSECV	1.4	3.1	2.4	2.2	2.9	2.2
$\mathbf{R_{CV/auto}}$	1.2	1.0	1.4	1.2	1.0	1.5
$\mathbf{Q^2}$	0.86	0.32	0.82	0.72	0.40	0.60
SECV	1.32	3.10	1.59	2.11	2.92	2.40
Bias	0.01	-0.02	-0.06	-0.65	0.14	-0.21
RPD	3.0	1.3	2.4	1.8	1.3	1.6
RSD(%)	18.9	44.3	22.6	30.2	41.8	34.2

However, investigating by plotting PCA scores of the RVA profiles with these quality-related parameters against the storage time (results not shown), no clear correlation was observed. Therefore, these matters should not interfere in the predictive results of the calibration models using the RVA profiles.

In general, RPD and RSD allow the comparison of the prediction accuracy from different models. Higher values of RPD indicate a better predictive ability corresponding to lower Bias, SECV and RSD with higher Q^2. It has been noted by several reports that RPD values greater than 2.5 are considered good enough for prediction purposes; however, in some complex systems, the calibration models with the RPD values greater than 1.75 could possess a reliable performance (Aleixandre-Tudo et al. 2015). In Table 1, using the RVA profiles, the RMSECV values from the supervised SOM, in most cases, are slightly higher. However, the predictive models were still reliable when considered that the RPD values >1.75 were good enough for prediction purposes. With the higher RPD values, it could be conducted that PLS provided the better predictive performance (RPD>2.5) than the supervised SOM models with our dataset.

4.3 Evaluation of important parameters for the model prediction

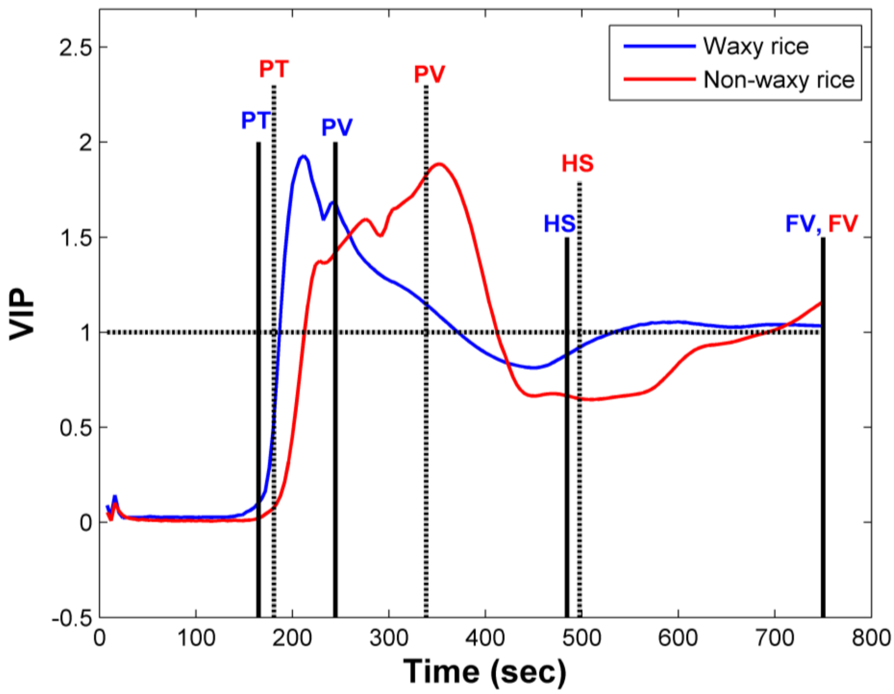

Fig. 5. PLS-VIP for the waxy and non-waxy rice samples. RVA parameters are the average values and indicated using black closed (waxy) and black dotted (non-waxy) lines.

The PLS-VIP values for the RVA profiles are shown in Fig. 5; positions of some common pasting properties, such as pasting temperature (PT), peak viscosity (PV), holding strength (HS), and final viscosity (FV) are indicated. The PLS-VIP parameters imply the importance of the parameters with respect to the prediction of the storage time. The size of the PLS-VIP parameter can be used to determine the significance or influence of the variable on the prediction model. A larger value implies that the parameter is likely to be associated with the predicted value. Normally, parameters with PLS-VIP values greater than one are considered significant. From the plots, the positions of the PT and PV parameters for the waxy and non-waxy rice are different. This result corresponds to the plots of the RVA profiles in Fig. 2(A), indicating the difference between the waxy and non-waxy types. For both waxy and non-waxy rice, PV and FV can be identified as the important parameters, since their PLS-VIP values are relatively high. In this research, a moving-window protocol based on moving window partial least squares (MWPLS) (Kasemsumran et al. 2004) was also tested with the RVA data and the regions identified as significance were more or less similar to those identified using PLS-VIP. To investigate if the model established using some selected parameters could obtain good results, common pasting properties in the RVA profiles including PT, PV, HS, FV, breakdown (BD), setback (SB) and pasting time (Pt), were used to calibrate the PLS and supervised SOM models. However, the predictive results were not better than those using the whole RVA profiles (results not shown). The changes in PV and FV were consistent with the previous report that observed noticeable changes in the RVA of the rice that had been stored for a long time (Keawpeng et al. 2015). These rheological changes could be due to natural aging process. Several mechanisms of aging have been hypothesized. Zhou et al. (2002) suggested that the alternation could be due to the changes in cell wall and proteins by enzyme activities. In addition, the interaction between starch and non-starch components (e.g. lipid, proteins and polysaccharides) in the rice could play an important role in the aging process (Patindol 2005). These ageing-induced structural changes may be attributed to the characteristics of the starch granules. For example, the starch granules of stored rice were more resistant to swelling that those of the fresh rice (Sowbhagya and Bhattacharya 2001). The removal of water from the amylose in the starch granules could lend to change in hardness and limited the granule hydration and swelling. These aging behaviors could cause the variations in the pasting properties detected by RVA.

To examine the effect of storage time on PV and FV, the component plane visualization of the PV and FV parameters is presented in Fig. 6. The numbers represented storage time ranging from 0 to 12 months. A component plane can be used to see how each variable influences the map (Kittiwachana et al. 2013). By comparing a component plane with the response plane of the supervised SOM model, it is possible to investigate their relationship. For waxy rice (Fig. 6(A)), the PV and FV values monotonically increased along with the extended storage time. The changing patterns among the response and component planes were more or less identical, implying they have a positive or direct relationship. For non-waxy rice (Fig. 6(B)), the changes in the PV and FV parameters were rather complicated. The increases in the PV and FV values were not directly related to the storage time. For example, the PV and FV values in month 2 were higher than month 7, indicating that the PV and FV values of non-waxy rice increased during early storage. After that, they decreased two to three months later. These results were consistent with the previous study by Zhou et al. (2002) reporting the change of RVA profiles during the storage time. During the first few months, the germ of fresh paddy grain was still in maturing stage and carried on with day to day metabolic processes. The foregoing suppositions could simply reflect the complexity of the rice aging process. This could be the reason that the RMSECV values from the non-waxy rice were slightly lower than those of the waxy rice.

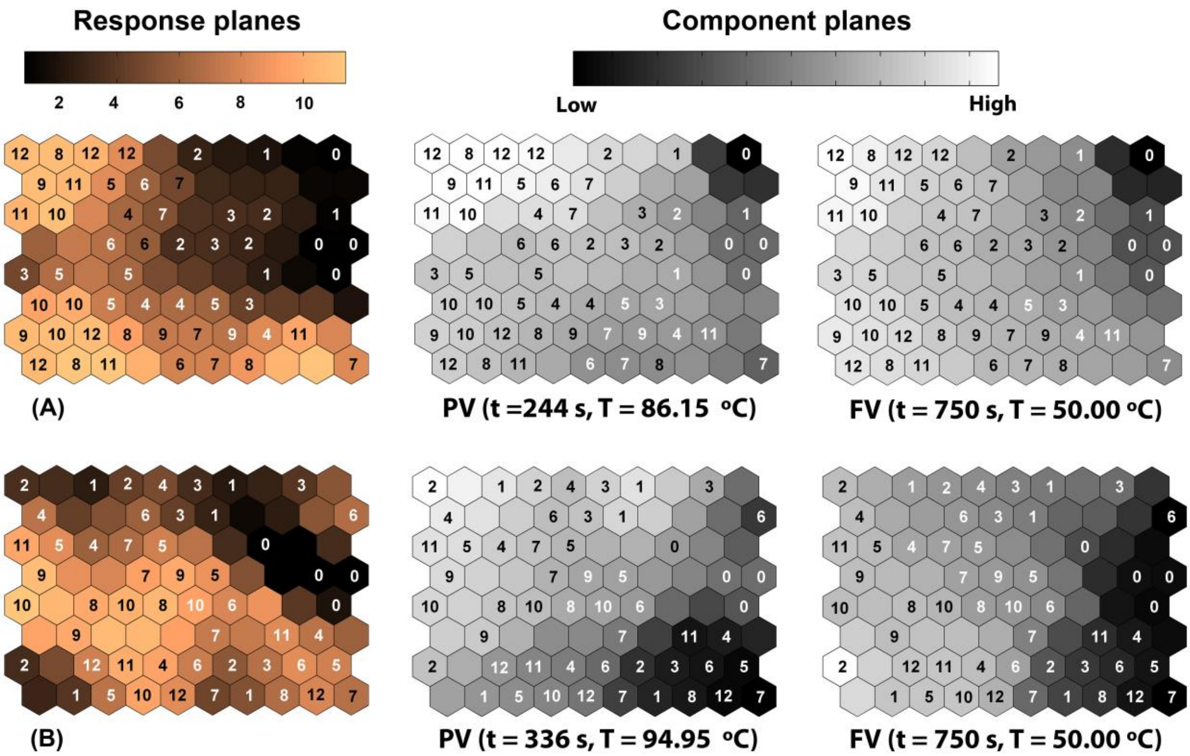

Fig. 6. Response (copper shading) and component planes (gray shading) of (A) waxy and (B) non-waxy rice. The recorded times (t) and temperatures (T) of the component planes are indicated and the numbers are the storage times starting for 0 to 12 months.

5. Conclusion

Rice grain gradually changes during storage and it is possible to observe the change using RVA. Chemometrics can extract the important information from the RVA profiles in relation to the rice storage time. Using PLS and supervised SOM, the storage time of the aged rice samples can be predicted. In our study, PLS offered better predictive results in terms of the RMSECV, Q^2, $R_{CV/auto}$, RPD and RSD values. The overall prediction performance of the waxy rice samples was better than the non-waxy. The developed method is an effective technique for determining the storage time of rice grains. It is simple and can be performed quickly with no chemicals required.

6. References

AACC (2000) Determination of the pasting properties of rice with the rapid visco analyser. Approved Methods of Analysis, 10th ed. Method 61-02.01, St. Paul, MN

Aleixandre-Tudo JL, Nieuwoudt H, Aleixandre JL, Du Toit WJ (2015) Robust ultraviolet visible (UVVis) partial least-squares (PLS) models for tannin quantification in red wine. J. Agric. Food Chem 63:1088–1098

Attaviroj N, Noomhorm A (2014) Discriminant analysis of multiple physicochemical properties for Thai rough rice varietal authentication. Int J Food Prop 17:1136-1149

Brereton RG (2003) Chemometrics: data analysis for the laboratory and chemical plant. Wiley, Chichester

Brereton RG (2009) Chemometrics for pattern recognition. Wiley, Chichester

Cagampang GB, Perez CM, Juliano BO (1973) A gel consistency test for eating quality of rice. J Sci Food Agric 24:1589-1594

Chen TF, Chen CL (2003) Analysing the freshness of intact rice grains by colour determination of peroxidase activity. J Sci Food Agric 83:1214–1218

Chrastil J (1990) Protien-starch interactions in rice grains. Influence of storage on oryzenin and starch. J Agric Food Chem 38:1804-1809

Chuang Y-K, Hu Y-P, Yang I-C, Delwiche SR, Lo YM, Tsai C-Y, Chen S (2014) Integration of independent component analysis with near infrared spectroscopy for evaluation of rice freshness. J Cereal Sci 60:238-242

Consonni V, Ballabio D, Todeschini R (2009) Comments on the definition of the Q^2 parameter for QSAR validation. J Chem Inf Model 49:1669-1678

Crosbie GB, Ross AS (2007) The RVA handbook. AACC International, U.S.A.

Fang C, Hu X, Sun C, Duan B, Xie L, Zhou P (2015) Simultaneous determination of multi rice quality parameters using image analysis method. Food Anal Methods 8:70–78

Farrés M, Platikanov S, Tsakovski S, Tauler R (2015) Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. J Chemom 29:528-536

Funsueb S, Krongchai C, Mahatheeranont S, Kittiwachana S (2016) Prediction of 2-acetyl-1-pyrroline content in grains of Thai Jasmine rice based on planting condition, plant growth and yield component data using chemometrics. Chemom Intell Lab Syst 156:203–210

Geladi P, Kowalski BR (1986) Partial least-squares regression: a tutorial. Anal Chim Acta 185:1-17

Guan B, Zhao J, Jin H, Lin H (2017) Determination of rice storage time with colorimetric sensor array. Food Anal Methods 10:1054–1062

Kanlayakrit W, Maweang M (2013) Postharvest of paddy and milled rice affected physicochemical properties using different storage conditions. Int Food Rest J 20:1359-1366

Kasemsumran S, Du YP, Murayama K, Huehne M, Ozaki Y (2004) Near-infrared spectroscopic determination of human serum albumin, γ-globulin, and glucose in a control serum solution with searching combination moving window partial least squares. Anal Chim Acta 512:223–230

Kaur M, Kaushal P, Sandhu KS (2011) Studies on physicochemical and pasting properties of Taro (*Colocasia esculenta* L.) flour in comparison with a cereal, tuber and legume flour. J Food Sci Technol 50:94–100

Keawpeng I, Venkatachalam K (2015) Effect of aging on changes in rice physical qualities. Int Food Rest J 22:2180-2187

Kim YJ, Kim S, Kim JW, Yoon G (2001) Data preprocessing and partial least squares regression analysis for reagentless determination of hemoglobin concentrations using conventional and total transmission spectroscopy, J Biomed Opt 6:177-182

Kittiwachana S, Ferreira DLS, Fido LA, Thompson DR, Escott REA, Brereton RG (2008) Dynamic analysis of on-line high-performance liquid chromatography for multivariate statistical process control. J Chromatogr A 1213:130–144

Kittiwachana S, Ferreira DLS, Fido LA, Thompson DR, Escott REA, Brereton RG (2010) Self-organizing map quality control index, Anal Chem 82:5972–5982

Kittiwachana S, Wangkarn S, Grudpan K, Brereton RG (2013) Prediction of liquid chromatographic retention behavior based on quantum chemical parameters using supervised self organizing maps. Talanta 106:229-236

Kohonen T (1982) Self-organized formation of topologically correct feature maps. Biol Cybern 43:59–69

Leivas CL, Da Costa FJOG, De Almeida RR, De Freitas RJS, Stertz SC, Schnitzler E (2013) Structural, physico-chemical, thermal and pasting properties of potato (*Solanum tuberosum* L.) flour. J Therm Anal Calorim 111:2211–2216

Li H, Prakash S, Nicholson TM, Fitzgerald MA, Gilbert RG (2016) The importance of amylose and amylopectin fine structure for textural properties of cooked rice grains. Food Chem 196:702–711

Lloyd GR, Brereton RG, Duncan JC (2008) Self organising maps for distinguishing polymer groups using thermal response curves obtained by dynamic mechanical analysis. Analyst 133:1046–1059

Mariotti M, Fongaro L, Catenacci F (2010) Alkali spreading value and image analysis. J Cereal Sci 52:227-235

Patindol J, Wang YJ, Jane JI (2005) Structure-functionality changes in starch following rough rice storage. Starch/Stärke 57:197-207

Postma GJ, Krooshof PWT, Buydens LMC (2011) Opening the kernel of kernel partial least squares and support vector machines. Anal Chim Acta 705:123– 134

Singh RK, Singh US, Khush GS (2000) Aromatic rices. Oxford & IBH, New Delhi

Siripatrawan U, Harte BR (2015) Data visualization of *Salmonella* Typhimurium contamination in packaged fresh alfalfa sprouts using a Kohonen network. Talanta 136:128–135

Sowbhagya CM, Bhattacharya KR (2001) Changes in pasting behaviour of rice during ageing. J Cereal Sci 34:115–124

Srikaeo K, Panya U (2013) Efficiencies of chemical techniques for rice grain freshness analysis. Rice Sci 20:292–297

Teo CH, Karim AAbd, Cheah PB, Norziah MH, Seow CC (2000) On the roles of protein and starch in the aging of non-waxy rice flour. Food Chem 69:229-236

Tong C, Chen Y, Tang F, Xu F, Huang Y, Chen H, Bao J (2014) Genetic diversity of amylose content and RVA pasting parameters in 20 rice accessions grown in Hainan, China. Food Chem 161:239-245

Tran TN, Afanador NL, Buydens LMC, Blanchet L (2014) Interpretation of variable importance in partial least squares with significance multivariate correlation (sMC). Chemom Intell Lab Syst 138:153–160

Vongsawasdi P, Noppharat M, Hiranyaprateep N, Tirapong N (2009) Relationships between rheological properties of rice flour and quality of vermicelli. As J Food Ag-Ind 2:102-109

Wang S, Li C, Copeland L, Niu Q, Wang S (2015) Starch retrogradation: a comprehensive review. Compr Rev Food Sci Food Saf 14:568-585

Wani AA, Singh P, Shah MA, Schweiggert-Weisz U, Gul K, Wani IA (2012) Rice starch diversity: effects on structural, morphological, thermal, and physicochemical properties-a review. Compr Rev Food Sci Food Saf 11:417-436

Zhou Z, Robards K, Helliwell S, Blanchard C (2002) Ageing of stored rice: changes in chemical and physical attributes. J Cereal Sci 35:65-78

Zhou Z, Robards K, Helliwell S, Blanchard C (2007) Effect of storage temperature on cooking behaviour of rice. Food Chem 105:491-497

Appendix

7. Suggestions for the next development

- It is suggested that the prediction for mixed rice samples should be investigated. For example, aged rice grain is mixed with fresh rice grain. The concern is that the age of the mixed rice grain can be possible or not used to identify if the rice sample has been adulterated.
- This research reported that RVA can be applied to determine the rice grain freshness. The developed method can do performed quickly without any chemical requirement. The operation cost is relative low. However, the RVA analyzer is rather expensive. The development of rapid viscosity measurement with low cost should be considered. It is not required that the developed measurement can operated with the same performance as the commercial RVA analyzer since only selected viscosity profiles are used for the prediction of the rice grain freshness.

8. Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ

   ชื่อผู้แต่ง: Sakunna Wongsaipun, Chanida Krongchai, Jaroon Jakmunee and Sila Kittiwachana

   ชื่อเรื่อง: Rice Grain Freshness Measurement Using Rapid Visco Analyzer and Chemometrics

   ชื่อวารสาร: Food Analytical Methods (Impact factor ปี 2017 = 2.038) ปี 2018 เล่มที่ 11 เลขที่ 2 หน้า 613-623

2. การนำผลงานวิจัยไปใช้ประโยชน์

   - เชิงวิชาการ: งานวิจัยนี้ถูกใช้เป็นส่วนหนึ่งในการเรียนการสอน รายวิชา 203879 “การวิเคราะห์เคโมเมทริกซ์” ภาควิชาเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ และสร้างนักวิจัยใหม่ในระดับปริญญาเอก จำนวน 1 คน คือ นางสาว กสุลณา วงค์สายปัณ รหัสประจำตัว 590551050

3. อื่นๆ

   - เสนอผลงานในที่ประชุมวิชาการ แบบบรรยาย ในการประชุม the Pure and Applied Chemistry International Conference 2016 (PACCON 2016) 9-11 ก.พ. 2559 ระหว่างวันที่ 9-11 กุมภาพันธ์ 2559 ณ ศูนย์การประชุม Bitec กรุงเทพฯ เรื่อง “A comparison of principal component analysis (PCA) and self-organizing map (SOM) for exploratory data analysis of near-infrared reflectance (NIR) spectra of rice grains”

- เสนอผลงานในที่ประชุมวิชาการ แบบบรรยาย ในการประชุม The ASIANALYSIS XIII ระหว่างวันที่ 8-11 ธ.ค. 2559 ณ The Empress International Convention Center, Chiang Mai กรุงเทพฯ เรื่อง "Application of Multiple Self-organizing Maps for Classification of Rice Samples Based on RVA profiles and Other Physico-chemical Properties"