NEAR-INFRARED SPECTROSCOPY AND HYPERSPECTRAL IMAGE FOR EVALUATION OF TEXTURE OF PARBOILED RICE AND KHAO DAWK MALI 105 (KDML 105) RICE AUTHENTICATION

JIRAPORN ONMANKHONG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

DOCTOR OF ENGINEERING IN AGRICULTURAL ENGINEERING

SCHOOL OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2021

KMITL-2020-EN-D-108-080

เทคนิคเนียร์อินฟราเรดสเปกโทรสโกปีและไฮเปอร์สเปกตรัลอิมเมจสำหรับการประเมินเนื้อสัมผัสของข้าวนึ่งและการพิสูจน์ความแท้ของข้าวขาวดอกมะลิ 105

ธนวัต สันหนองเมือง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชาวิศวกรรมเกษตร
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2564
KMITL-2021-EN-D-108-080

หัวข้อวิทยานิพนธ์	เทคนิคเนียร์อินฟราเรดสเปกโทรสโกปีและไฮเปอร์สเปกตรัลอิมเมจสำหรับการประเมินเนื้อสัมผัสของข้าวนึ่งและการพิสูจน์ความแท้ของข้าวขาวดอกมะลิ 105
นักศึกษา	นางสาวจิราพร อ่อนมั่นคง
รหัสประจำตัว	60601172
ปริญญา	วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมเกษตร
พ.ศ.	2021
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ศ.ดร.ปานมนัส ศิริสมบูรณ์

บทคัดย่อ

งานวิจัยนี้ได้ศึกษาความเป็นไปได้ในการใช้เทคนิคเนียร์อินฟราเรดสเปกโทรสโกปีและเทคนิคไฮเปอร์สเปกตรัลอิมเมจ เพื่อประเมินเนื้อสัมผัส (ความแข็งและความเหนียว) ของข้าวนึ่งหุงสุก และการพิสูจน์ความแท้ของข้าวขาวดอกมะลิ105 ตามลำดับ วัตถุประสงค์ของงานวิจัยนี้แบ่งออกเป็น 3 ส่วนหลัก วัตถุประสงค์แรก ได้แก่การศึกษาเบื้องต้นเกี่ยวกับผลกระทบขององค์ประกอบทางเคมี (อะไมโลสและไขมัน) ต่อสภาวะการผลิตข้าวนึ่งที่แตกต่างกัน ซึ่งส่งผลกระทบต่อเนื้อสัมผัสของข้าวนึ่งหุงสุก เพื่อพิจารณาความเป็นไปได้ในการเทคนิคเนียร์อินฟราเรดสเปกโทรสโกปีในการประเมินเนื้อสัมผัสของข้าวนึ่งหุงสุก จากการทดลองพบว่าปริมาณอะไมโลสมีความสัมพันธ์ทางลบกับความแข็ง (ค่าสัมประสิทธิ์สหสัมพันธ์ (r)=-0.52) ของข้าวนึ่งหุงสุกอย่างมีนัยสำคัญ เนื่องจากการละลายของปริมาณอะไมโลสในน้ำระหว่างกระบวนการแช่ทำให้ปริมาณอะไมโลสลดลง สำหรับความแข็งที่เพิ่มขึ้นอาจเกิดจากการเปลี่ยนแปลงของสตาร์ชในกระบวนการการเจลาติไนเซชันและการเกิดผลึกใหม่ ทำให้ความแข็งของข้าวนึ่งหุงสุกเพิ่มขึ้นถึงแม้ว่าอะไมโลสจะลดลง ปริมาณอะไมโลสมีความสัมพันธ์ทางลบกับความเหนียว (r=-0.38) ของข้าวนึ่งหุงสุก สำหรับปริมาณไขมันพบว่ามีความสัมพันธ์ในเชิงบวกกับความแข็ง (r=0.20) และความเหนียว (r=0.12) ของข้าวนึ่งหุงสุกในระดับต่ำ วัตถุประสงค์ที่สองของงานวิจัยนี้ คือการพัฒนาแบบจำลองการสอบเทียบที่ดีที่สุดของการทำนายเนื้อสัมผัสของข้าวนึ่งหุงสุก โดยใช้สเปกตรัลของข้าวนึ่งดิบจากฟูเรียร์ทรานฟอร์มเนียร์อินฟราเรดสเปกโตรมิเตอร์ที่ความยาวคลื่น 12,500-4,000 cm^{-1} วิธีอ้างอิงมาตรฐานที่ใช้ในการวัดเนื้อสัมผัสของข้างหุงสุก คือวิธี ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking Method แบบจำลองการสอบเทียบความแข็งที่ดีที่สุดถูกพัฒนาด้วยสมการการถดถอยกำลังสองน้อยที่สุดบางส่วน โดยใช้การจัดการสเปกตรัลด้วยวิธีการปรับค่าเฉลี่ยให้เรียบ โดยให้

สัมประสิทธิ์การตัดสินใจของชุดการทำนาย ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยของการทำนาย และอัตราส่วนของการทำนายต่อค่าเบี่ยงเบน เท่ากับ 0.70 7.21 N และ 1.93 ตามลำดับ แบบจำลองการสอบเทียบความเหนียวที่ดีที่สุดถูกพัฒนาด้วยสมการการถดถอยองค์ประกอบหลัก โดยใช้การจัดการสเปกตรัลด้วยวิธีการปรับเป็นค่ามาตรฐานเฉลี่ยโดยให้สัมประสิทธิ์การตัดสินใจของชุดการทำนาย ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยของการทำนาย และอัตราส่วนของการทำนายต่อค่าเบี่ยงเบน เท่ากับ 0.66 38.00 Nmm และ 1.75 ตามลำดับ เกณฑ์อัตราส่วนของการทำนายต่อค่าเบี่ยงเบนทั้งสองแบบจำลองเหมาะสำหรับแอปพลิเคชันการทำนาย วัตถุประสงค์ที่สาม คือการพัฒนาแบบจำลองการสอบเทียบของการพิสูจน์ความแท้ของข้าวขาวดอกมะลิ105 โดยใช้เนียร์อินฟราเรดไฮเปอร์สเปกตรัลอิมเมจจำแนกพันธุ์ข้าว 3 สายพันธุ์ ได้แก่ พันธุ์ขาวดอกมะลิ105 ปทุมธานี1 และพิษณุโลก2 ซึ่งได้ทดสอบข้าวสองประเภทได้แก่ ข้าวกล้องและข้าวสาร ภาพไฮเปอร์สเปกตรัลอิมเมจได้มาจากการสแกนจากระบบ push-broom line scanning ที่ความยาวคลื่น 913 ถึง 2519 nm แบบจำลองการจำแนกพันธุ์ข้าวที่ดีที่สุดสำหรับข้าวกล้อง คือแบบจำลองที่ถูกพัฒนาด้วยซัพพอรตเวคเตอรแมชชีน โดยใช้สเปกตรัลเนียร์อินฟราเรดเฉลี่ย ซึ่งค่าความถูกต้องในการจำแนกประเภทของชุดทดสอบมีค่าเท่ากับ 93.0% สำหรับแบบจำลองการจำแนกพันธุ์ข้าวที่ดีที่สุดสำหรับข้าวสาร คือแบบจำลองที่ถูกพัฒนาด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันจากข้อมูลภาพเนียร์อินฟราเรดไฮเปอร์สเปกตรัลอิมเมจ ซึ่งค่าความถูกต้องในการจำแนกประเภทของชุดทดสอบมีค่า 95.2% จากผลการวิจัยนี้สนับสนุนว่าเทคนิคเนียร์อินฟราเรดไฮเปอร์สเปกตรัลอิมเมจเป็นเครื่องมือที่รวดเร็วและไม่ทำลายตัวอย่างสามารถป้องกันการปลอมปนของข้าวหอมมะลิไทยเมื่อเทียบกับวิธีดั้งเดิม (1-2 วัน)

Thesis	Near-infrared spectroscopy and hyperspectral image for evaluation of texture of parboiled rice and Khao Dawk Mali 105 (KDML105) rice authentication
Student	Miss Jiraporn Onmankhong
Student ID.	60601172
Degree	Doctor of Engineering
Program	Agricultural Engineering
Year	2021
Thesis Advisor	Prof. Dr. Panmanas Sirisomboon

ABSTRACT

This research studied the feasibility of using Near-infrared (NIR) spectroscopic technique to evaluate the texture properties (i.e., hardness and toughness) of cooked parboiled rice and the Near-infrared hyperspectral image (NIR-HSI) for Khao Dawk Mali 105 (KDML105) rice authentication. This thesis has three objectives. The first objective was to conduct a preliminary study of the effect of chemical components (i.e. amylose and fat content) on different parboiled rice process conditions which affect the texture properties of cooked parboiled rice to determine the feasibility of applying NIR spectroscopy in the evaluation of the texture properties. It was found that the amylose content had a significantly negative correlation with the hardness of cooked parboiled rice (Correlation Coefficient (r)=-0.52), and a negative correlation with its toughness (r=-0.38). The amylose content decreased with the increasing hardness of cooked parboiled rice. This might be because the amylose content leached out into the water during the soaking process. The increasing hardness for cooked parboiled rice may be due to changes in starch during gelatinization and recrystallization processes. Therefore, hardness of cooked parboiled rice increases although the amylose decreases. The fat content had a low correlation with the hardness (r=0.20) and the toughness (r=0.12) of texture properties of cooked parboiled rice. The second objective was to develop the best calibration model of texture properties prediction of cooked parboiled rice from spectra of milled parboiled rice using the NIR spectroscopy. The

Fourier-transform NIR (FT-NIR) spectrometer (MPA, Bruker Ltd., Ettlingen, Germany) at a wavelength 12,500-4,000 cm^{-1} (800-2,500 nm) was used in this part. The ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking Method was used as a reference test. The Partial Least Squares Regression (PLSR) was the optimal calibration model of hardness with moving average smoothing pre-processing with the coefficient of determination of prediction set (r^2), root mean square error of prediction (RMSEP) and ratio of prediction to deviation (RPD) of 0.70, 7.24 N and 1.93, respectively. The Principal Components Regression (PCR) was the best model for toughness using mean normalization preprocessing provided r^2, RMSEP and RPD of 0.66, 38.00 Nmm and 1.75, respectively. The RPD threshold of both models was fair for prediction application. The third objective was to develop the calibration classification model of KDML105 rice authentication using the NIR-HSI, which was evaluated to classify three rice varieties including KDML105 (Thai Jasmine rice), PTT1 (Pathum Thani1), and PSL2 (Phitsanulok 2) at both brown and milled rice. The push-broom line scanning system (Compovision, Sumitomo Electric Industries, Ltd., Osaka, Japan) was applied to acquire NIR-HSI at the wavelength from 913 to 2519 nm. The optimal model for brown rice was the Support Vector Machine (SVM) model based on the averaged NIR spectra, of which the classification accuracy of the test set was 93.0%. For the optimal model for milled rice was the Convolutional Neural Network (CNN) model based on the NIR-HSI data, of which the classification accuracy of the test set was 95.2%. The results support that the HSI method works as a rapid and non-destructive tool for preventing the adulteration of Thai Jasmine rice compared with traditional methods (1-2 days).

Acknowledgements

I wish to express my deepest and sincere gratitude to my advisors, Professor Dr. Panmanas Sirisomboon for her valuable supervision, advice, supporting, immense knowledge as well as encouragement throughout the course of my study. Without her persistent help, the aim of this thesis would not have been realized. I would like to express to Assistant Professor Dr. Jiraporn Sripinyowanich Jongyingcharoen for her help recommendations. Particularly, my doctorate program had the financial support from Royal Golden Jubilee scholarship (RGJ) PhD program (Grant numbers PHD/0013/2560) from National Research Council of Thailand (NRCT) for tuition fee, research expenses, monthly expenses, as well as international research experience. In addition, I would like to express my sincere gratitude to Professor Satoru Tsuchikawa, Associate Professor Tetsuya Inagaki and Assistant Professor Te Ma for allowing me to join in their laboratory at Nagoya University, Japan, and their invaluable help, constant encouragement and immense knowledge throughout the international research experience.

I wish to thank the Near Infrared Spectroscopy Research Center for Agricultural Products and Food (www.nirsresearch.com) at King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand and Graduate School of Bio Agricultural Sciences, Nagoya University, Nagoya, Japan for laboratory space and instruments. In addition, I wish to thank the Capital Chainat Rice Mill company limited, in Sapphaya District, Chainat Province, Thailand for providing parboiled rice samples.

I express my gratitude to all faculties of the department of agricultural engineering at King Mongkut's Institute of Technology Ladkrabang (KMITL), not only the research methodologies but also everything for living, motivation and support. I would like to thank my friends for their wonderful collaboration, you supported experimental me greatly and were always willing to help me.

Finally, I would like to express appreciation to my family for all their support both physically and mentally throughout the course of my study.

Jiraporn Onmankhong

Table of contents

Page

Table of contents (continued)

Page

Table of contents (continued)

Table of contents (continued)

Page

List of tables

Table **Page**

List of figures

List of figures

Figure **Page**

Abbreviations

ANN	Artificial neural networks
Bias	Average error
CNN	Convolutional neural network
FC	Fully Connected
FN	Number of false negatives
FP	Number of false positives
FT-NIR	Fourier transform near-infrared
HSI	Hyperspectral image
ISO	International organization for Standardization
KDML105	Khao Dawk Mali 105
LLM	Linear learning machine
LSVM	Linear support vector machines
MLP	Multiple-layer perception
MLR	Multiple linear regressions
MSC	Multiple scatter correlation
NIR	Near Infrared
PC	Principal components
PCA	Principal component analysis
PCR	Principal component regression
PLS	Partial least squares regression
PSL2	Phitsanulok 2
PTT1	Pathum Thani1

Abbreviations

R^2	Coefficient of determination
ReLU	Rectified linear unit
Rep	Repeatability
RMSEC	Root mean square error of calibration
RMSECV	Root mean square error of cross validation
RMSEP	Root mean square error of prediction
ROI	Regions of interest
RPD	Standard deviation to standard error of prediction
SD	Standard deviation
SEC	Standard deviation of error of calibration
SEP	standard deviation of error of prediction
SIMCA	Soft independent modeling of class analogies
SNV	Standard normal Variate
SVM	Support vector machines
TN	Number of true negatives
TP	Number of true positives
TPA	Texture profile analysis

Chapter 1

Introduction

1.1 Statement and significance of the problems

Thailand is one of the main producers of rice (*Oryza sativa L.*) (Chaijan and Panpipat 2020) which is essential food for approximately half of the world's population (Oli et al. 2014). The total amount of rice exported by Thailand was 5,724,680 metric tons, generating 115,915 million baht to the country (Thai Rice Exporters Association 2020). Parboiled rice and Thai jasmine rice, which are the most important rice exports of Thailand, were studied in this research.

Parboiled rice is one important product of Thailand rice exports (Thai Rice Exporters Association 2020). The advantages of parboiled rice include various nutrients, high rice yield, and delayed digestion rate all of which generate interest for the purchaser and the rice industry (Li et al. 2021). In addition, the parboiled rice can extend shelf life and decrease broken kernels (Jannasch et al. 2020). There are three steps in the process of parboiling rice: soaking, steaming, and drying of paddy (Jannasch et al. 2020).The soaking and steaming process affect the texture properties of parboiled rice product (Miao et al. 2016; Villanova et al. 2017). The texture properties (hardness and toughness) of cooked parboiled rice influence the consumer acceptability (Direct communication with the parboiled rice factory). Therefore, the texture of cooked parboiled rice is considered an important factor of parboiled product for consumers. Also, different parboiled rice production conditions affect the physical and chemical (amylose and fat content) properties of rice (Patindol et al. 2008). The common instrument that has been used to measure the texture of cooked rice samples has been the texture analyzer (Miao et al. 2016). Texture profile analysis (TPA) test has been used to determine the average values of the hardness, adhesiveness, cohesiveness and springiness of cooked rice (Miao et al. 2016).

Khao Dawk Mali 105 (KDML105) rice variety is the most famous of Thai Jasmine rice because of its longgrain, exclusive aroma, pure white color, soft texture, and the good flavor after being cooked (Siriphollakul et al. 2017; Kukusamude and Kongsri. 2018; Qingyu et al. 2020). Thailand is a leading manufacturer and exporter of premium

grade KDML105 rice with an expensive price in the global market (Boling et al. 2011). However, KDML 105 rice can be produced in a limited amount per year, as the planting can only be done in the rainy season from June to November in Northeastern and Northern Thailand (Boling et al. 2011; Korinsak et al. 2016). KDML105 rice has been found to be adulterated with other cheaper varieties. The popular rice variety mixed with KDML105 includes Pathum Thani1 (PTT1) and Phitsanulok 2 (PSL2). The price of rice in 2020 ranges as follows: KDML105 priced at 26 baht per kg, while PTT1 and PSL2 were priced at 19 and 14 baht per kg (Information received on April 7, 2021 from the Rice Department). As shown, the price difference is approximately between 7 and 12 baht per kg. The PTT1 and PSL2 which are mixed with KDML105 have a similar appearance that is difficult to distinguish by the human eye. This is a problem that affects the reputation and credibility of Thailand's export, as well as the prices of KDML105. Moreover, the physical and chemical properties are specific characteristics of each variety (Attaviroj et al. 2011) including the milling quality, grain appearance, cooking characteristics, and eating quality (Attaviroj and Noomhorm 2014). Therefore, rice quality is changed when the rice is contaminated with other varieties (Yadav and Khatkar 2007; Attaviroj et al. 2011). Presently, the Thailand Ministry of Commerce has enacted stipulations on the Thai Jasmine rice standard, which has indicated that pure KDML105 must not comprise less than 92% of the KDML105 (Ministry of Commerce. 2016). The standard methods to evaluate the authentication of Thai Jasmine rice include (1) test method for determination of amylose content; (2) test method for Alkali spreading value; (3) test method for checking cooked rice kernels boiled in water; (4) test method by iodine staining, which was specified by the Ministry of Commerce, Thailand (Ministry of Commerce 2016).

However, the common method is not a rapid or safe way to check texture properties of the cooked parboiled rice and pure KDML105 rice. Near infrared (NIR) spectroscopy (800-2500 nm) is a fast and non-destructive technique for determining the quality of agricultural and food products, which is related to the stretching and bending vibrations among the hydrogen bonds of the products such as C-H, O-H, and N-H, which have dominant overtones (Ma et al. 2020). The absorbance and reflectance spectrum of NIR can identify the different chemicals. Also, NIR hyperspectral imaging (HSI) is a rapid and non-destructive technique, an advanced technique that provides a

NIR spectral image at each wavelength with a high spatial resolution. The acquisition of NIR spectral image makes HSI technique suitable for kernel-by-kernel classifier. Therefore, the NIR and NIR-HSI techniques could be applied to the evaluation of the texture properties and the classification of rice varieties respectively because of the different chemical components that affect the different NIR spectral.

Research problems regarding parboiled rice texture and rice authentication were found. However, the possibility of using the NIR and NIR-HSI techniques to solve is very high. Both techniques have the NIR principle possibilities which are related to the stretching and bending vibrations among the hydrogen bonds described above The chemical component is found to have an important factor for NIR prediction. In the case of the texture of parboiled rice, the different conditions of the parboiled rice process affect the physical and chemical properties of the rice (Oli et al. 2014). Longer soaking time led to reduced amylose content in the hot water soaking (Patindol et al. 2008). The amylose content increased with the gelatinization temperature (Varavinit et al. 2003). Also, the fat content was decreased at high soaking temperatures (Sareepuang et al. 2008) whereas an increased steaming time increases the fat content (Ibukun 2008). In addition, the decreased hardness of cooked parboiled rice is due to the decreasing amylose content of the rice, and the changed stickiness of parboiled rice is due to the different proportions of amylose and amylopectin (Sarangapani et al. 2015). Moreover, the hardness of cooked rice is affected by fat content (Kim et al. 1986) which is a negative correlation (Cameron and Wang 2005). Therefore, it is possible to use NIR technique to predict the texture of cooked parboiled rice. However, this research was a preliminary study conducted to demonstrate the possibility of a NIR technique to predict the texture of cooked parboiled rice. In case of rice variety (i.e. KMDL105, PTT1, PSL2) classification, the different varieties of rice were different in the amylose and protein content (Division of Rice Research and Development 2016), followed by KDML105 (12-17% and 9.23%), PTT1(15-19% and 9.60%) and PSL2 (28% and 9.81%), respectively. Regarding the different chemical compositions of the rice, it is possible to use the NIR-HSI technique to classify rice variety, and the NIR HIS is most appropriate for both authentication and classification application (Manley 2014). A previous study identified that starch and protein were key for the classification of various kinds of rice on whole grain samples, including KDML105, PTT1, PSL2, Suphan

Buri60, and Chainat 1, using a Fourier-transform NIR (FT-NIR) spectroscopy (Attaviroj et al. 2011). However, the NIR spectral data were collected from several samples (approximately 70 ± 2 g). It is highly desired to achieve a single-sample measurement, especially for the simultaneous detection of multiple targets. In addition, Rittiron et al. (2005) have studied the detection of variety contamination in milled Japanese rice using a single kernel NIR technique and found that protein distribution was key for the varieties classification of Koshihikari (pure) and Akitakomachi (contamination) variety. These studies would have less capability if the average protein content was close to the pure variety (Rittiron et al. 2005). This is difficult to achieve by conventional NIR spectrometers on single grain. However, the main advantage of NIR-HIS is a spectrum collected as each pixel in the image which shows the distribution of different chemical components in a sample. It is, therefore, most suitable for the analysis of heterogeneous samples (Manley. 2014). For the above HSI-NIR advantage, the rice variety classification using HSI-NIR may have high accuracy than NIR. Therefore, this research opted for the NIR-HSI technique for KDML105 rice authentication due to more information gained in the average spectrum of single grain than offered by NIR spectroscopy.

1.2 Goal and objective

The overall objective of this thesis was to investigate the possibility of using NIR and NIR-HIS techniques for evaluating the texture (i.e. hardness and toughness) properties of parboiled rice (rice variety including Suphan Buri 1, Suphan Buri 90, RD41, RD47, Chai Nat 1, Phitsanulok 2 and so on) and KDML105 rice authentication from PTT1, PSL2. The research has three specific objectives as follows:

- To preliminarily study the influence of chemical components (i.e., amylose and fat content) on different parboiled rice process conditions to determine the possibility of applying NIR spectroscopy to evaluate the hardness and toughness of the texture properties of parboiled rice.
- To establish the best calibration model of texture properties prediction for parboiled rice using the NIR spectroscopic and to examine the performance of the model.

- To develop the calibration classification model of KDML105 rice authentication using the NIR-HSI and assess the accuracy percentage of the calibration and to validate the model using unknown samples.

1.3 Hypothesis to be tested

The amylose and fat content in milled parboiled rice have a correlation with the texture of cooked parboiled rice. Also, the texture has a correlation with the NIR spectral characteristic of whole grain parboiled rice since the texture has a correlation with the amylose and fat content. In addition, an NIR hyperspectral characteristic of KDML105 milled rice indicated the authenticity of KMDL105 milled rice.

1.4 Scope or limitation of the study

In this thesis, the Suphan Buri 1 rice variety was used to conduct a preliminary study on the influence of chemical components on different parboiled rice process conditions to determine the possibility of applying NIR spectroscopy to evaluate the texture properties of parboiled rice. The high amylose group (amylose more than 25%) including the Suphan Buri 1, Suphan Buri 90, RD41, RD47, Chai Nat 1, Phitsanulok 2, and so on (used for parboiled rice production in parboiled rice factory) are used for a model development of texture properties prediction. The KDML105, PTT1, and PSL2 were used in KDML105 rice authentication.

For the measurement, the FT-NIR spectrometer (MPA, Bruker Ltd., Ettlingen, Germany) at a wavelength 12,500-4,000 cm^{-1} (800-2,500 nm) with 16 cm^{-1} resolution was used in case of texture properties prediction. Moreover, the push-broom line scanning system (Compovision, Sumitomo Electric Industries, Ltd., Osaka, Japan) was used to acquire NIR hyperspectral diffused reflectance images for rice variety classification. This system comprised the HSI camera possessing a spectroscope and a two-dimensional photosensitive element [256 pixels (wavelength from 913 to 2519 nm with a 6.2 nm spectral resolution) × 320 pixels (element of detectors)].

1.5 Process of the study

This thesis is organized into seven chapters as follows:

Chapter 1 presents the research background and statement of the problem including the research objectives of this thesis.

Chapter 2 shows the details of theories and literature reviews including the NIR and NIR-HSI technique, the physiology of rice (parboiled rice, Khao Dawk Mali 105, Pathum Thani 1, and Phitsanulok 2), texture properties of parboiled rice and model development using principal component regression, principal component analysis, partial least squares regression, and support vector machine regression.

Chapter 3 illustrates the NIR scanning precision test for texture characteristics of parboiled rice. This chapter illustrates the precise FT spectrometer and shows the precision analysis for reference laboratory of texture properties. The purpose of this chapter is to ensure that the NIR spectral recording from FT-NIR and standard reference of texture properties are precise, and thus suitable for creating a model.

Chapter 4 presents the influence of processing parameters of parboiled rice on its physicochemical and texture. This chapter illustrates the study of how different parboiled rice conditions affect the chemical composition of parboiled rice. The outcome comfirms that NIR spectroscopy can be applied to the different spectrum (different chemical compositions) to evaluate the texture properties of parboiled rice.

Chapter 5 illustrates the texture evaluation of cooked parboiled rice using nondestructive milled whole grain NIR spectroscopy. The calibration model of texture properties was constructed in this chapter.

Chapter 6 displays the rapid and non-destructive method to distinguish genuine Thai Jasmine rice (Khao Dawk Mali 105) from Pathum Thani 1 and Phitsanulok 2 using NIR-HSI combined with machine learning and deep learning methods. This chapter shows the optimal model of KDML105 rice authentication. Furthermore, a validated model of KDML105 rice authentication to predict unknown samples is also displayed in this chapter.

Chapter 7 presents the overall main conclusion and also recommends further study from the studies in chapters 3, 4, 5, and 6 of this thesis.

Chapter 2
Literature review

2.1 Overview of Parboiled rice

Rice (*Oryza sativa L.*) is an essential food for over half of the world's population. It is estimated that more than half of the total population consume rice (Pallegedara. 2020). Approximately 20% of the world's rice is used to produce parboiled rice (Zhu et al. 2020). Parboiled rice production of Thailand exported to foreign countries amounted to 1,441,924 metric tons in 2020, which is worth 21,026 million bath (Thai Rice Exporters Association. 2020). The counties of parboiled rice imported from Thailand are Benin, South Africa, Bangladesh, Cameroon, Yemen, Russian Federation, Algeria, U. Arab Emirates, Nigeria, Spain, and others (Napasintuwong. 2019). The parboiled rice factories in Thailand prefer to use rice with high amylose content to produce rice for export because this increases the value of rice (Direct communication with the parboiled rice factory). Parboiled rice process eliminates the problem of cracks and incomplete grain filling and leads to many favourable changes including easy shelling, higher head rice yield, fewer broken grain, increased resistance to insect and nutrient retention (Kwofie and Ngadi. 2017).

2.1.1 Parboiling process

The condition of parboiling process changes the chemical and physical properties of rice (Patindol et al. 2008). Parboiled rice is produced in three step before milling starting with soaking, steaming and drying of paddy (Li et al. 2021). Each process has different details as follows:

2.1.1.1 Soaking process

The soaking process involves soaking paddy in the hot water, which is suitable (moisture level 24-30% wb) for completing the gelatinization of the steaming process (Buggenhout et al. 2014; Oli et al. 2014). With a 6- hour soaking time at 65°C condition, the moisture content of paddy was about 30% which results in a good quality production (Islam et al. 2002). The hot water promotes heat transmission from the surface of the hull to the middle of the endosperm of paddy (Luh. 1991). Moreover, soaking in hot water at high temperatures reduces the processing times which also

helps avoid fermentation which affects the smell of parboiled rice products when occurred (Luh. 1991).

2.1.1.2 Steaming process

Steam heating must be applied uniformly to keep the heat regularly touching the rice (Luh. 1991). Gelatinisation and re-crystallisation are the main changes in rice starch during parboiling (Oli et al. 2014). The starch is gelatinized, filling the void and cementing the cracks to reduce the breakage (Thurber et al. 2014; Kwo and Ngadi. 2017). The time of steaming must be long enough to completely gelatinize the whole kernel (Luh. 1991). The steaming temperature of 100°C at 20 min resulted in the absence of a chalky kernel, which would identify the complete gelatinization of parboiled rice (Kwofie and Ngadi. 2017).

2.1.1.3 Drying

The aims of drying parboiled rice are to reduce the level of moisture content to be suitable for milling and storage and to obtain the maximum milling yield (Luh. 1991). After the soaking and steaming process, the paddy is sun-dried till the moisture content is reduced to 14%, which is considered optimum level (Luh. 1991; Miah et al. 2002).

2.1.2 Texture properties of parboiled rice after cooking

Cereal is consumed mainly in the form of whole (milled) kernel after cooking (Ramesh et al. 2000). Therefore, the texture properties of cooked kernel is the main factor consideration of consumer acceptability and palatability as well as moisture to touch (Ramesh et al. 2000). Rice is the main cereal which is largely consumed in the form of whole grains after cooking (Li and Gilbert. 2018). Cooked rice texture is a multi-parameter with hardness or firmness and stickiness or adhesiveness as the most commonly determined parameters (Ramesh et al. 2000; Li et al. 2016). In addition, the different countries and regions of rice customers have different preferences of cooked rice (Li and Gilbert. 2018).

Cooked parboiled rice is more fluffy, non-sticky and free-flowing than the non-parboiled (Oli et al. 2014).The parboiled rice process affects the texture properties of cooked rice (Villanova et al. 2017). The hardness and toughness parameters of cooked

parboiled rice affect the consumer acceptability, and texture properties of cooked rice which change during parboiling (Miao et al. 2016). Different physical and chemical properties of the parboiled rice occurred at different conditions during parboiling (Patindol et al. 2008). However, texture properties of cooked rice are influenced by a wide range of factors, such as the starch (Oli et al. 2014), protein (Oli et al. 2014), and lipids (Saleh and Meullenet. 2007), postharvest processing (Champagne et al. 1998), the milling ratio (Park et al. 2001), variety (Oli et al. 2014), storage duration (ageing) (Oli et al. 2014), and the cooking method (Leelayuthsoontorn and Thipayarat. 2006). However, the major factor controlling the texture of cooked rice is amylose content; the high-amylose rice becomes flaky and dry upon cooking whereas low-amylose rice is sticky and moist (Juliano et al. 1981; Oli et al. 2014). This thesis focuses on the chemical properties of parboiled rice which affect the texture properties of cooked parboiled rice.

2.1.3 Changes in chemical properties of rice during parboiling

The changing of starch, protein, and lipids occurs in the parboiled rice process (Luh. 1991). Different parboiled condition processes lead to different chemical compositions (Oli et al. 2014).

2.1.3.1 Starch

Starch has a multi-scale structural model where the granules are made from the alternating growth rings of amorphous and crystalline structure (Oli et al. 2014). The starch granules comprise almost entirely of two major polysaccharides, namely amylose and amylopectin (Biswas and Juliano. 1988; Li and Gilbert. 2018).

During the steaming process, swollen starch granules are dissolved and the phase transition of starch occurs where the fully gelatinized starch turns into a viscous mass (Oli et al. 2014).

During the drying process, the starch increases in crystallinity including the formation of crystalline amylose-lipid complexes (Derycke. 2007). The crystallinity is a gradual change from a viscous-amorphous state to a glassy state which affects the physical state of diffusion of the water in rice kernel (Oli et al. 2014).

The changing of the texture of cooked parboiled rice is controlled by 1) the gelatinisation and thermal breakdown of starch and 2) the recrystallisation (retrogradation) of the starch with some lipid–amylose inclusion complexes (Oli et al. 2014).

Moreover, Biswas and Juliano (1988) and Patindol et al. (2008) reported a mean 1% decrease in apparent amylose content during the soaking process because of the leaching out into water according to the severity of parboiling. Normally, the reduced hardness of cooked parboiled rice is due to the decrease in amylose content of the rice, and the changed stickiness of cooked parboiled rice is due to the different proportions of amylose and amylopectin (Sarangapani et al. 2015).

2.1.3.2 Protein

Proteins are highly concentrated on the outward surface of the rice kernel and in the bran (Champagne et al. 2004).

During the steaming process, surface proteins may regulate the water diffusion into the starch granule and control the granule swelling during gelatinization (Matveev et al. 2000). Although the protein bodies in the kernel are ruptured during steaming process (Oli et al. 2014). The protein is hydrolysed leading to the raised disulphide bonds which increased the viscosity and hardness in the parboiled rice (Oli et al. 2014).

2.1.3.3 lipids

Lipids in rice are divided into two types: non-starch and starch lipid where 60% of non-starch lipids is contained in rice bran (aleurone layer and embryo) (Godber and Juliano. 2004; Oli et al. 2014). Starch lipids are rather low concentrations and main in complex with amylose (Oli et al. 2014).

During parboiling, the lipid bodies of the non-starch lipids are broken and fat is released from the surface of the kernel which is diffused outwards and hence the bran of parboiled rice is more oily (Mahadevappa and Desikachar. 1968). The lipids offer a barrier to water diffusion because of their hydrophobic nature (Oli et al. 2014). The hydrophobic nature of fat might contribute to the reduced water absorption of full-fat bran (Kaur et al. 2011; Billiris et al. 2012). Moreover, Saleh and Meullenet. 2007 reported that low surface lipids content decreased the firmness and increased the stickiness of cooked rice.

2.1.4 Changes in physical properties of rice during parboiling

2.1.4.1 color

The color of parboiled rice is different from that of the non-parboiled rice (Oli et al. 2014). In addition, the color of grain was evaluated for consumer acceptability (Direct communication with the parboiled rice factory).

During the parboiling process, the color of parboiled rice is changed depending on the of severity of soaking and steaming condition (Sareepuang et al. 2008; Lamberts et al. 2006). Parboiled rice has an amber color due to the diffusion of husk color into the endosperm during parboiling process (Lamberts et al. 2006). Another reason has been reported that increased levels of reducing sugar and free α-amino nitrogen (FAN) and isomerisation of glucose to fructose exhibit the likelihood of non-enzymatic Maillard type of browning for color change in parboiled rice (Lamberts et al. 2008).

2.2 Thai Jasmine rice (KDML105) authentication

In 2020, Thailand's export quantities of Thai jasmine rice to many parts of the world were as high as 40,999 million tons, and worth 1,189,752 million (Thai Rice Exporters Association. 2020). Thai Jasmine or Thai Hom Mali rice or KDML105, which is grown in Thailand, is respected to be the optimal aromatic rice variety because of its long-grain, unique aroma, mild taste, white color, good flavor and soft texture after cooking (Phanchaisri et al. 2007; Siriphollakul et al. 2017; Kukusamude and Kongsri. 2018; Udomkun et al. 2018; Qingyu et al. 2020). The key chemical compound that elicits this aroma is 2-acetyl-1-pyrroline (2AP) found in KDML 105 rice (Kong-ngern et al. 2011; Mahatheeranont et al. 2001; Tulyathan and Leeharatanaluk; 2007). Moreover, the rice aroma intensity mainly depended on the 2AP content which varied among rice cultivars, storage and harvest (Hu et al. 2020). The KDML105 is transparent and contains very few chalky kernels (Wangcharoen et al. 2016). Thailand is the main producer and exporter of KDML105 with a high price in the global market (Boling et al. 2011). However, the KDML105 rice is only grown in rainy conditions (June to November) and in Northeastern and Northern Thailand (Boling et al. 2011; Korinsak et al. 2016;

Udomkun et al. 2018). Because the KDML105 is harvested once a year, there is only a limited amount of rice yield annually. As a result, the market demand is not met (Phanchaisri et al. 2007; Pitiphunpong et al. 2011) Due to the high price of pure KDML105, some traders may adulterate the Jasmine rice by mixing it with low grade, low price, or low nutritious rice to sell at a high price (Timsorn et al. 2017). The uncontrolled mix of rice results in lower cooked-rice quality and unfair trade (Pitiphunpong et al. 2011).

Popular rice varieties to be contaminated with KDML105 include Pathum Thani1 (PTT1) and Phitsanulok 2 (PSL2). Generally, the difference in chemical and physical characteristics (Table 2.1) depends on the variety and the quality of the starch especially its cooking properties (Table 2.1) (Juliano. 1985; Qiu et al. 2018; Jiamyangyuen et al. 2019). The amylose content of the rice is an important characteristic which determines eating and cooking quality (Jiamyangyuen et al. 2019). The amylose content has a parameter to predict the texture properties of cooked rice which is classified into the following: amylose content as waxy (0-5%), very low (5–12%), low (12–20%), intermediate (20–25%), and high (25–33%) of cooked rice (Li and Gilbert. 2018; Jiamyangyuen et al. 2019). However, physical characteristics (shape and size) of other kinds of rice used for mixing are quite similar to those of KDML105, making it difficult to differentiate by naked-eye (Table 2.2) (Attaviroj and Noomhorm. 2014; Timsorn et al. 2017). For the PTT1 variety especially, it is very difficult to distinguish between KDML105 and PTT1 varieties through physical properties and amylose content (Attaviroj and Noomhorm. 2014). In addition, Because the PTT1 variety brings an offseason rice crop, its yearly production is infinitely higher than KDML105 variety which is an in-season rice crop (Attaviroj and Noomhorm. 2014) In addition, the effect by alkali spreading value difference of KDML105 and PTT1 was 6.4 and 7.0, respectively (Srisawas and Jindal. 2007). The gel consistency of KDML105 and PTT1 was 75 and 64 mm, respectively (Srisawas and Jindal. 2007). Also, the DNA finger prints method is used to inspect the KDML105 (Department of Foreign Trade, Ministry of Commerce. 2016). Genetic diversity can be evaluated with morphological traits, seed proteins, isozymes, and DNA markers (Ashu and RS. 2015). Molecular marker

technology is a powerful tool for determining genetic variation in rice varieties, which can reveal abundant differences among genotypes at the DNA level providing a more direct and reliable consumer (Ashu and RS. 2015).

Table 2.1 The chemical and physical properties of KDML105, PTT1, and PSL2 variety (Attaviroj and Noomhorm. 2014; Division of Rice Research and Development. 2016; Boonpermpol et al. 2018).

Rice variety	Amylose content (%)	Protein content (%)	Size (long × width × thick) (mm)	Texture properties after cooking
Khao Dawk Mali 105 (KDML105)	15.91	9.23	7.5×2.1×1.8	Softness, Natural fragrant
Pathum Thani 1 (PTT1)	15.42	9.60	7.6×2.1×1.7	Softness, Stickiness, Mild fragrance
Phitsanulok 2 (PSL2)	26.65	9.81	7.9×2.1×1.6	Crumbly, Hardness

2.3 Instrumental measurement of texture

The textural properties of parboiled rice after cooking have previously been presented in 2.1.2. The important parameters of cooked rice are hardness and toughness, both of which are the indicators for consumer acceptance.

The measuring of the textural characteristics of cooked rice is important for consumer acceptance. Highly labor-intensive sensory analysis has inevitably led to the development of instrument methods which are designed to measure food properties that correlate to sensory characteristics (Kress-Rogers and Brimelow. 2001). These methods have been separated in many ways, according to the type of measurement and the type of food (Kress-Rogers and Brimelow. 2001).

This thesis focuses on the extrusion test, which was used to assess the textural properties of cooked parboiled rice according to the ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after the Cooking method (ISO. 2012).

Table 2.2 The RGB photo of KDML105, PTT1, and PSL2 varieties for brown and milled rice (The rice grains in this table are the seeds used in the experiment).

Rice varieties	Khao Dawk Mali 105 (KDML105)	Pathum Thani 1 (PTT1)	Phitsanulok 2 (PSL2)
Brown rice			
Milled rice			

*This figure was taken by the camera (Fujifilm, X-T20, Philippines) with 24.3 pixel megapixel (MP) resolution at room temperature.

2.3.1 Extrusion test

The extrusion test was used for assessing the texture of cooked rice made from non-parboiled rice and parboiled rice. A 4:7.6 rice-to-water ratio by weight was used for cooking (ISO. 2012). 5 g of cooked parboiled rice was placed in the testing cell of the rig. The test speed was 1.67 mm/s (ISO. 2012). The Mini Ottawa cell (HDP-MK05 model, stable Micro systems, Surrey, UK) was used for ISO 11747 (Figure 2.1).

Figure. 2.1 Rice extrusion test

The rice was pushed down by a plunger of a similar cross-section to the cell. Hence, the rice was compressed before being extruded through the holes in a perforated plate which, until a distance of 15 mm, caused the rice to fall through the holes. Two parameters of textural properties from the extrusion test are hardness and toughness, which are obtained from the maximum force (N) and area under the graph from initial to maximum force point (Nmm) of the force-distance curve according to ISO 11747 (Figure 2.2). The graph of the extrusion test in Figure 2.2 was used to analyze the parboiled rice texture using a texture analyzer (TA.HD. plus, Stable Micro Systems, UK).

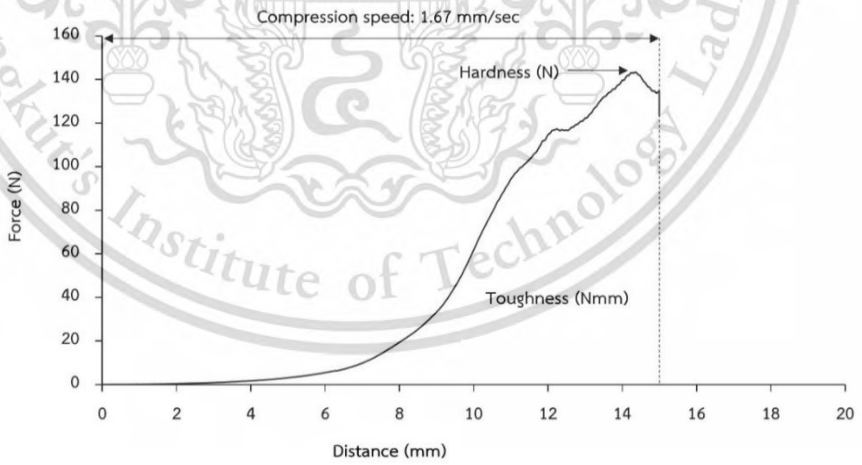

Figure. 2.2 The extrusion test of 5-g cooked parboiled rice samples. TPA text parameters: force-distance test, Mini Ottawa cell test speed is 1.67 mm/s to travel 15 mm (ISO. 2012). (This graph was obtained from our experiment).

2.3.2 Texture Profile Analysis (TPA)

TPA is a texture measurement method of cooked rice, and is a popular method (Deshmukh et al., 2015; Xu et al. 2019; Ye et al. 2019; Zhu et al. 2020). The TPA is an imitative test that endeavors to mimic mastication via machine. These machine instruments provide stress and/or strain during the test. The load is removed from the rice and allowed to relax after the first bite (Rosenthal. 1999). In the second bite, the rice is compressed again and then allowed to relax for a second time (Rosenthal. 1999).

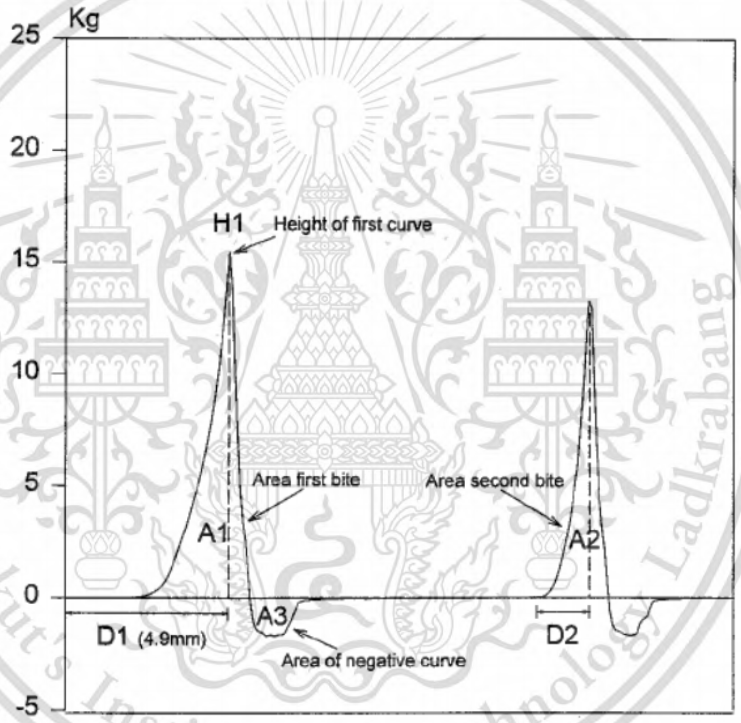

Figure. 2.3 Typical texture profile analysis (TPA) curve of 1 g rice samples. TPA text parameters: force-distance test, compression plate set at 5 mm to travel 4.9 mm at 1 mm/sec (Champagne et al. 1999).

Champagne et al. (1999) performed an experiment to evaluate the texture of cooked rice using a texture analyzer (TA.XT2, Texture Technologies Crop., Scarsdale, NY). Samples of short, medium, and long grain cultivars (N=76) were obtained in Louisiana, Arkansas, Texas, and California. For similar cooked rice types, the rice-to-

water weight ratios appropriate for three cooked types were based on amylose content (0%, 1:1; 10–19%, 1:1.4; 20–25%, 1:1.7). 1 gram of warm cooked rice was placed in a single-grain layer on the base plate. A compression plate was set at 5 mm above the base. A two-cycle compression using the force-versus-distance program was used to allow the plate to travel 4.9 mm, return, and repeat, for which the test speed was 1 mm/sec. The test was repeated on two duplicate samples (n=3). Figure 2.3 shows parameters the recorded from test curves as follow (Champagne et al. 1999) :

1) H1, hardness (kg) is the measurement of force at the peak of first curve.
2) A3, adhesiveness is the area of negative force curve representing work to separate the plunger from the sample on an upstroke after the first curve.
3) A2/A1, cohesiveness is the ratio of area under curves A2/A1, where A1 is the area under the first curve and A2 is the area under the second curve.
4) D2/D1, springiness is the ratio of D2 to D1, where D1 is total distance (4.9 mm) traveled by the plunger on a downstroke and D2 is distance traveled on a downstroke by the plunger from the point of sample contact to the end of the downstroke.
5) A5/A4, resilience, where A4 is measured from the first data point to the first probe reversal point and A5 is measured from the first probe reversal point to the point where force returns to zero.
6) Gumminess was calculated by multiplying hardness by cohesiveness.
7) Chewiness was obtained by multiplying gumminess by springiness.

Moreover, Rosenthal (1999) reported the sensory definitions of each of the parameters of TPA for cooked rice as shown in Table 2.3.

2.4 Near-infrared spectroscopy

2.4.1 Principles of NIR spectral

Infrared energy is the electromagnetic energy of molecular vibration which is a defined within the energy bands of 780 to 2,500 nm (12821 to 4000 cm^{-1}) (Workman and Weyer. 2012). The absorption of NIR is based on whether the different chemical bonds in organic matter absorb or emit light of different wavelengths when the sample

is irradiated (Zareef et al. 2020). Organic matter in samples has different spectra because of the specific vibrational frequencies of chemical bonds, determined by the shape of the molecule, the mass of the constituent atoms, the stiffness of the bonds, and the periods of the associated vibrational coupling (Zareef et al. 2020). The most prominent absorption bands in the NIR region are associated with molecular overtone and combination vibrations of fundamental vibrations of –CH, –NH, –OH (and –SH) functional groups (Reich. 2005; Zareef et al. 2020). Also, the NIR spectrum was comprised of information about physical properties such as particle size or crystallinity of the sample (Blanco et al. 1999).

Table 2.3 Parameters Measured by Texture Profile Analysis (Rosenthal. 1999).

Parameter	**Sensory Definition**
Hardness	Force required to compress a food between the molars.
Elasticity	The extent to which a compressed food returns to its original size when the load is removed.
Adhesiveness	The work required to pull the food away from a surface.
Cohesiveness	The strength of the internal bonds making up the food.
Chewiness	The energy required to chew a solid food until it is ready for swallowing.
Gumminess	The energy required to disintegrate a semisolid food so that it is ready for swallowing.

Spectroscopy is a technique which presents the characteristic molecular vibrations inside materials, showing matter that interacts with electromagnetic radiation at different wavelengths and declaring a signature known as a spectrum. The amplitude spectra at any wavelength correlate to absorptivity and the number of molecules assumed to correspond to a concentration of the sample. Therefore, NIR spectroscopy

is a technique that measures the NIR spectrum of a material that is radiated by NIR light (Workman and Weyer. 2012). NIR spectroscopy is a non-destructive, non-invasive, rapid, and effective technique for which the advantages are that there is little or no need for sample preparation, it is fast, has lower cost than conventional techniques, and has the capability to analyze a wide range of products.

2.4.2 Sample presentation devices

When NIR radiation hits a sample, the incident radiation may be transmitted absorbed, or reflected, and the relative contribution of each phenomenon depends on the chemical constitution and physical parameters of the sample (Nicolaï et al. 2007). The type of sample presentation is mainly divided into “transmission,” “reflection,” “transfection,” and “interaction” modes, in which transfection mode is the combination of transmission and reflection (Tsuchikawa. 2007). Moreover, interaction mode is a hybrid of transmittance and reflectance (Osborne. 2006).

In the case of the transmission mode, the incident light radiates vertically to one side of the sample and the transmitted light is detected from the opposite side; this mode is widely used in instances with a cuvette for liquids without scattering or in low-scattering conditions (Tsuchikawa. 2007). Beer’s Law is an equation that has great practical appeal due to its simple linear relationship between a readily measured quantity (I/I_0 is the fraction of the incident light that penetrated the sample and was detected) and the concentration (c) of the absorber, which is usually what is to be deduced in spectroscopy (Dahm and Dahm. 2021).

In the case of the reflection mode, the Kubelka–Munk equation is applicable (Tsuchikawa. 2007). Most of the radiation is reflected from the surface by regular or specular reflection, and no absorption takes place for a smooth surface. This situation is called diffuse reflectance because most of the incident radiation is reflected (Osborne. 2006). If, however, some of the radiation penetrates the surface when it reaches each particle then it can be reflected, absorbed or transmitted (Osborne. 2006). Therefore, the absorbance (A) values relative to a standard reference material are measured according to log 1/ R and log 1/T for reflectance and transmittance spectra, respectively (Reich. 2005). In addition, diffuse reflection mode has been used

for milled whole grain (Wu and Shi. 2004). Furthermore, Fu et al. (2007) found that reflection measurements are easier to obtain and the light levels of the reflected radiation are much higher than those of transmitted radiation; however, variations in superficial and surface properties of fruit may influence calibrations.

In the case of the transfection mode, the radiation is transmitted through the sample, reflected from the ceramic and then transmitted back through the sample before finally reaching the detector (Osborne. 2006). NIR spectra from small volumes of the sample can be clearly measured (Tsuchikawa. 2007). The transfection mode was recommended to avoid stray light, which is an important spectral acquisition (Sandak et al. 2016).

In the case of interaction, an interaction probe with a concentric outer ring of illuminator and an inner portion of receptor is usually used. The end of the probe is in contact with the surface of the sample. Therefore, only the light transmitted through the sample can be detected (Tsuchikawa. 2007). Fundamentally, the relative merit of interaction mode is that it provides a compromise between reflection and transmission modes (Magwaza et al. 2012). It is normally accomplished using a fiber-optic probe in which one set of fiber-optic bundles carries the incident radiation and another carries the reflected radiation, which is particularly useful for large samples such as intact fruit (Osborne. 2006).

2.4.3 Near-infrared (NIR) instrumentation

Generally, a NIR spectrometer is composed of a light source, a monochromator (Grating), a sample holder, or a sample presentation interface, and a detector, allowing for transmittance or reflectance measurements (Reich. 2005). The main NIR transmission spectroscopy components are as follows (Zareef et al. 2020): (1) source of light; (2) beam-splitter system; (3) reflector; (4) detector inlet valve/sample chamber; (5) detector for diffuse reflection; (6) detector for transmission; and (7) data processing and control analyzed system which is shown in Figure 2.4. The beam-splitter system plays a significant role due to its translation of multicolor into single color light such as an interferometer, light filter, and gratings (Zareef et al. 2020). The light source is

normally a tungsten halogen lamp because it is small and durable (Kawata. 2002). The monochromator is used to disperse or spread out radiation according to wavelength, utilizing prisms and gratings (Manickavasagan and Jayasuriya. 2014). Detector types are silicon, lead sulfide (PbS) and indium gallium arsenide (InGaAs) (Kawata. 2002).

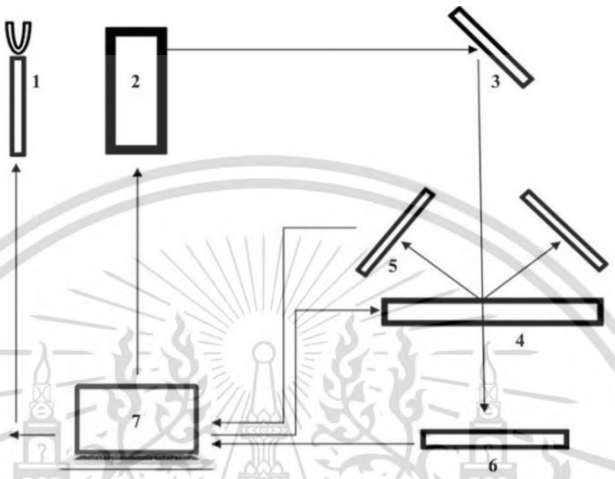

Figure. 2.4 Components of the NIR spectroscopy (Zareef et al. 2020).

The analysis of the NIR information procedure generally consists of the following steps: collection of spectral data; spectral data pre-processing; calibration models developed with a set of samples and models, validated with a set of independent samples which is detailed in the topic of modeling and performance analysis.

In this thesis, Interferometer or Fourier transform (FT) NIR spectroscopy was used. The main advantage of FT-NIR spectroscopy is the high signal-to-noise ratio and scan speed. FT-NIR spectroscopy collects the absorption and reflectance spectrum of a sample in a wide range of spectra. These instruments have a light source that illuminates NIR radiation towards the interferometer, and consist of a beam splitter and two mirrors, of which one is fixed, and the other is moving at a constant velocity. A simple form of interferometer is the Michelson interferometer, that includes two mirrors placed mutually perpendicular to each other, and a beam splitter (Figure 2.5). When the NIR Illuminates the interferometer, the beam splitter partially reflects half of the radiation to one mirror and transmits another half of the radiation to the other mirror. The beams reflected back from both mirrors are recombined at the beam

splitter and directed to the sample. Information in the time domain, from the data collected during the motion of the moving mirror, contains the spectral information of the sample, which is retrieved by Fourier transformation (Manickavasagan and Jayasuriya. 2014).

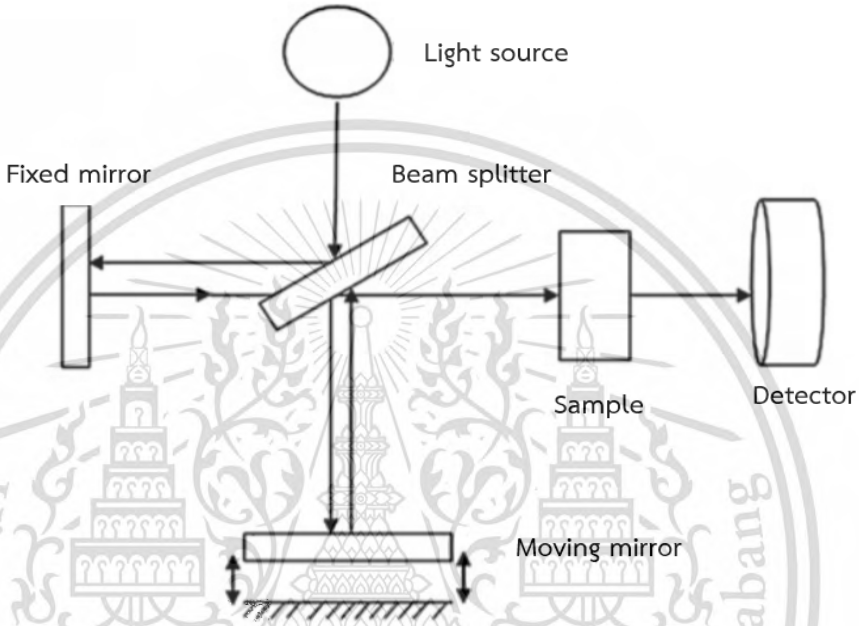

Figure. 2.5 Schematic view of Michelson interferometer (Manickavasagan and Jayasuriya. 2014).

2.4.4 Near-infrared spectroscopic applications of rice quality

Rice is a staple commodity of most people in Asian countries (Jiamyangyuen et al. 2019). There have been papers published on NIR applications for the prediction of rice quality traits, including the texture properties of cooked rice and the chemicals of rice, as shown in Table 2.4. Many researchers have conducted the NIR spectroscopic technique to determine rice quality and have mentioned that the technique was usable to evaluate the quality of rice. Most of the research had higher accuracy, which indicated that NIR spectroscopy was an ideal technique for predicting rice quality.

Table 2.4 NIR Applications predicting rice quantified.

Wavelength (nm)	Sample form	Sample container	Measurement mode	Prediction accuracy	Reference
400–2,500	Whole grain rice	Transport cell	Reflectance	Slickness (R^2=0.81), Stickiness (R^2=0.71), Hardness (R^2=0.85), Cohesiveness (R^2=0.96)	Champagne et al. 2001
400–2,500	Whole grain rice	Rectangular transport cell	Reflectance	Adhesion to lips (R^2=0.88), Hardness (R^2=0.79), Cohesiveness (R^2=0.79), Toothpack (R^2=0.85)	Meullenet et al. 2002
1100-2,498	Milled rice powder	Small ring cup	Diffuse reflectance	Acid amino (R^2=0.98)	Wu et al. 2002
1100-2,498	Rice flour	Spinning cup	Reflectance	Amylose content (R^2=0.99) Protein content (R^2=0.98)	Sohn et al. 2004
1100-2,500	Whole grain rice	Small ring cup	Diffuse reflectance	Amylose content (R^2=0.85)	Wu and Shi. 2004

R^2 is coefficient of determination, R is correlation coefficient.

Table 2.4 NIR Applications predicting rice quality (continued).

Wavelength (nm)	Sample form	Sample container	Measurement mode	Prediction accuracy	Reference
833–2,500	Brown rice grain, Brown rice flour, Milled rice grain, Milled rice flour	Sample cup	Reflectance	Fat content of brown rice grain (R^2=0.79) Fat content of brown rice flour (R^2=0.84) Fat content of milled rice grain (R^2=0.89) Fat content of milled rice flour (R^2=0.91)	Wang et al. 2006
940-2,222	Whole grain rice	A holder which was modified from grain counting plate	Tranmission	Springiness (R^2=0.61), Resilience (R^2=0.86), Deformation (R^2=0.87), Cohesiveness (R^2=0.91)	Siriphollakul et al. 2017
833–2,500	Rice flour	Circular sample cup	Transflection	Amylose content (R=0.94)	Sousa et al. 2018

R^2 is coefficient of determination, R is correlation coefficient.

2.5 Hyperspectral imaging

2.5.1 Fundamentals of hyperspectral imaging and components of hyperspectral imaging systems

NIR spectroscopy has gained attention as a non-destructive assessor of food quality, as well as for its ability to obtain information about sample components based on the light absorption of a sample. While it is not easy to know the position location information, it is easy to judge the position of certain features by naked eye or computer vision systems, although it is difficult to conduct a quantitative analysis of a component. Therefore, hyperspectral imaging technique is a combination of the strong and weak points of visible/near-infrared spectroscopic techniques and vision techniques (ElMasry et al. 2012).

The composition of images and the specifications of the system utilized to produce the images are important for hyperspectral images. A hyperspectral image could be generated in three conventional types: whiskbroom (point-to-point spectral scanning), pushbroom (a line-by-line spatial scanning), and tunable filter or staring imaging (an area scanning) according to the characteristics of image spectrographs (ElMasry and Nakauchi. 2016). The main difference between them is that whiskbroom scanning scans the sample in the spatial domain by moving the sample point-by-point, while pushbroom scanning moves the sample line by line (Xiong et al. 2014). On the other hand, the tunable filter obtains images one wavelength after another, and so it is also called wavelength scanning. This method is different from whiskbroom and pushbroom scanning, as it keeps the sample fixed (Xiong et al. 2014). However, the pushbroom hyperspectral imaging system is most widely used in food inspection experiments because of its consistency for on-line applications (ElMasry et al. 2012).

The pushbroom hyperspectral imaging system comprising an illumination lamp, spectrograph (a light dispersive device), a camera, and a computer equipped with image acquisition software is presented in Figure 2.6 (a) (ElMasry and Nakauchi. 2016). The resulting hyperspectral image of a food sample is a stack of sub-images of this sample at discrete, contiguous, spectral narrow bands such that a complete reflectance spectrum can be obtained for any region of the sample being imaged (Nakauchi et al. 2012; ElMasry and Nakauchi. 2016)

Although it is possible to acquire hyperspectral images in reflectance, transmittance, or interacting modes, acquiring images using reflectance configuration (in

which light source and camera are orientated in the same direction) is common and the reflectance measurements could be converted to absorbance values (using Absorbance = log(1/Reflectance)) to fit Lambertian modeling (ElMasry and Nakauchi. 2015).

Hyperspectral imaging systems produce three-dimensional (3D) hypercubes consisting of two-dimensional spatial data and one-dimension spectral data, forming a stack of images at continuous wavelengths, and every image at a specific wavelength is called a sub-image (Figure 2.6b and 2.6c) and could be assigned as I (x, y, λ), i.e., a spatial image I (x, y) at different wavelengths (λ) (Xiong et al. 2014; Özdoğan et al. 2021). Hence, a food sample may look brighter at certain wavelengths, having high reflectance (low absorbance) at these particular wavelengths. Moreover, the same sample might appear darker at certain wavelengths if their reflectance values at these particular wavelengths were very low (high absorbance) as shown in Fig. 2.6c. Normally, sub-images at two adjacent wavelengths are very similar, while sub-images at distant wavelengths can be much less similar and may have independent information (ElMasry and Nakauchi. 2016). Furthermore, no single sub-image at one wavelength has sufficient data to completely explain the object's properties, which explains why hyperspectral imaging is useful in the analysis of food materials (Kim et al. 2004).

Hyperspectral information has a hypercube that includes a considerable amount of data; therefore, various procedures are performed to both protect necessary data and to reduce redundant information and the dimensionality of the data (Özdoğan et al. 2021). There are three main steps in the processing of hypercubes: image acquisition and pre-processing, extraction of data and treatments, data modeling, and post-processing (Özdoğan et al. 2021), which are detailed in the topic of modeling and performance analysis.

Furthermore, many researchers conducted the HSI-NIR technique to classify rice varieties. In addition, the research had higher accuracy, which indicated that HSI-NIR spectroscopy could classify rice varieties (Kong et al. 2013; Wang et al. 2015; Qiu et al. 2018; He et al. 2019; Weng et al. 2020).

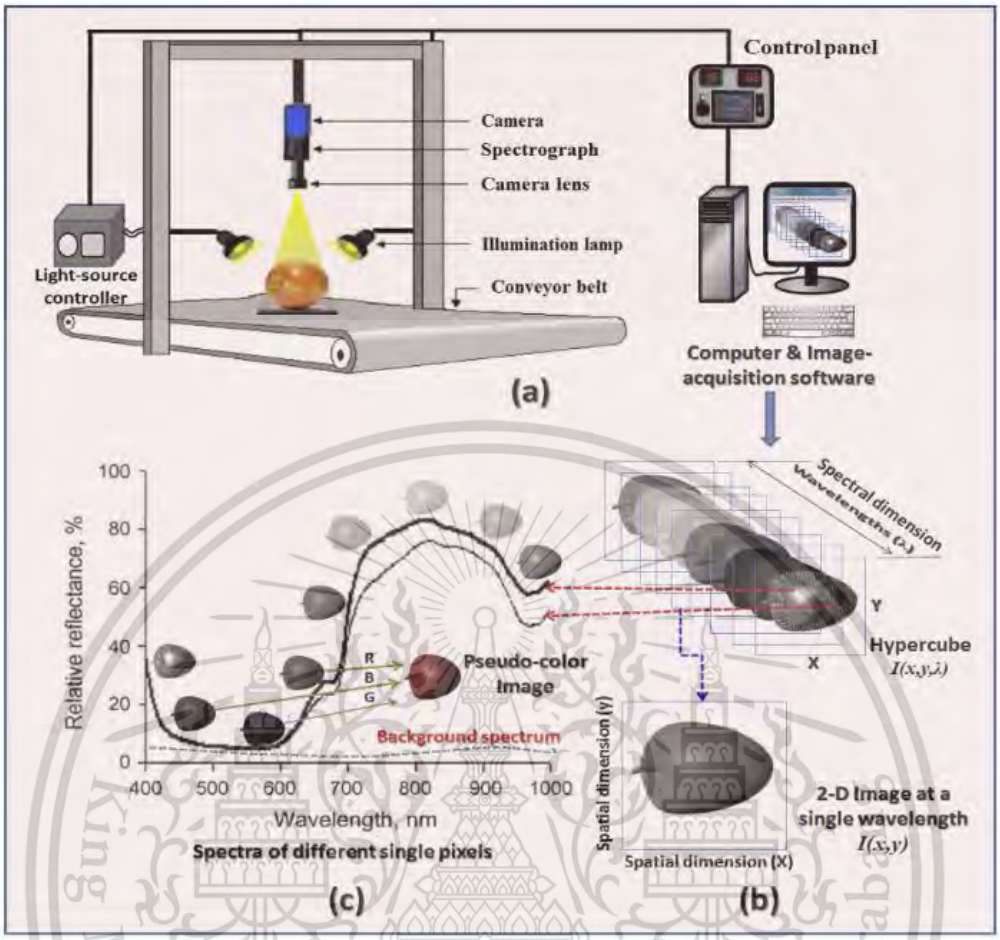

Figure. 2.6 (a) The basic configuration of pushbroom hyperspectral imaging for acquiring hyperspectral data structure 'Hypercube'. Hypercube could be displayed either as individual 2D sub-image I(x,y) at any given wavelength (b) or as spectra I(λ) at any given pixels in the image (c) (ElMasry and Nakauchi. 2016).

Moreover, the hypercube obtained from NIR-HSI systems contains a mass of information with large dimensionality, hence the main purpose of hyperspectral data analysis is to reduce the dimensionality and retain the useful data for discrimination or measurement analysis of food quality and safety (Sergi, 2015).

However, hyperspectral images need to undergo image preprocessing to extract the NIR spectra for chemometrics. After acquiring a hyperspectral image for the tested sample, this image should be calibrated with the help of white and dark hyperspectral

images (Sun. 2010). This is due to the fact that the raw spectral image collected using conventional NIR or hyperspectral imaging shows detector signal intensity (Sergi. 2015). The differences in camera quantum efficiency and different physical configuration of hyperspectral imaging systems, as well as the uncorrected radiance for different hyperspectral imaging systems, might not be the same even when imaging the same target under the same imaging conditions (ElMasry and Nakauchi 2016). In order to correct the acquired images, it is necessary to use two reference images, called the 'dark' and 'white' images, to obtain a relative reflectance (ElMasry and Nakauchi. 2016).

All of the collected spectral images were converted to relative reflectance values for the following analysis (Ma et al. 2020) using Eq. (2.1):

$$R_{\lambda,n} = \frac{S_{\lambda,n} - D_{\lambda,n}}{B_{\lambda,n} - D_{\lambda,n}} \tag{2.1}$$

where $R_{\lambda,n}$ is the relative reflectance value, and where λ and n represent the wavelength and pixel index variables, respectively. S is the sample image. B and D are the white reference and dark image. The dark reference image was acquired by turning off the light source and completely covering the lens with its cap (Ma et al. 2020).

After that, the spectral data of relative reflectance were extracted from different regions of interest (ROIs) that presented different quality features in the relative reflectance image (Sun. 2010). For each ROI, the average relative reflectance spectrum was calculated across all ROI pixels from all areas of the images. Finally, only one spectrum per sample was obtained (Sergi. 2015).

NIR spectroscopy only provides the mean spectrum (average measurement) of a sample, irrespective of the area of the sample scanned. This is because the spectra collected are averaged to provide a single spectrum, and the information on the spatial distribution of constituents within the sample is thus lost (Manley. 2014). Spectral and spatial features combined in hyperspectral images allow simultaneous characterization of different chemical attributes in the sample, as well the uniformity and gradient distributions of these attributes (ElMasry et al. 2012). Grayscale or color mapping with intensity scaling is commonly used to display compositional contrast between pixels in an

image (Sun. 2010). The final results of these calculations are used to develop key quantitative image parameters to characterize various traits in the tested samples in different categories by performing classification, identification, mapping, and/or visualization (Sun. 2010).

2.5.2 Comparison between Hyperspectral image and RGB image

The differences between HSI and RGB are illustrated in Figure 2.7, which depicts the light reflectance curve of a single pixel from an arbitrary sample imaged using both of these methods (Mehta et al. 2018). A conventional imaging system with an ordinary color camera exhibits relatively broad spectral responses due to only three different spectra corresponding to the visual primary colors (i.e. red (R), green (G), and blue (B)) (ElMasry and Nakauchi. 2016). The RGB image comprises 3 bands which are contained on the visible band. The results that can be obtained are limited. On the other hand, the hyperspectral image consists of hundreds of bands (Ozdemir and Polat. 2020). The traditional imaging is most effective when quality attributes of a sample are related to its extrinsic characteristics, but it becomes less effective, or ineffective when quality attributes are mainly determined by the intrinsic properties of the sample, such as composition and internal physical characteristics, which are not readily detectable at the surface (ElMasry and Nakauchi. 2016). The spectral measurements of the hyperspectral images at each pixel provide the collected spectrum of the materials within the pixel (Manolakis et al. 2003).

2.5.3 Application of NIR-HSI to authentication of agricultural products and foods

Presently, authentication of agricultural products and food has become a necessity throughout the world (Mendez et al. 2019). The main objective of authentication is to detect fraud, which is a contraband activity for economic purposes involving the adulteration of false information on food labels (Danezis et al. 2016; Barreto et al. 2018); Cases of fraud have been shown to cause distrust among consumers, large economic losses to companies or governments, destruction of brands, and a devaluing of the market value of products (Mendez et al. 2019). Previous studies have presented the applicability of NIR-HIS to agriculture and food authentication. The previous research indicated the

ability to estimate agriculture and food authenticity using NIR-HSI. Therefore, there are possibilities to apply this technique to the authentication of rice varieties.

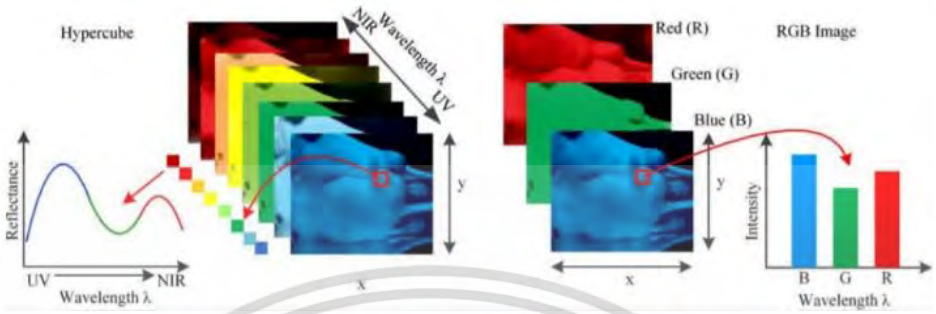

Figure. 2.7 Hyperspectral and RGB image components (Mehta et al. 2018; Ozdemir and Polat. 2020)

2.6 Modeling and performance analysis

Chemometrics is the science of relating measurements made through chemical systems or statistical methods (Cen and He. 2007) to design or select optimal measurement procedures and experiments, and to provide maximum chemical information or to extract relevant information from chemical data (Otto. 2017). Without computing capabilities and multivariate methods, NIR spectroscopy applications would not be possible (Agelet and Hurburgh. 2010). Chemometrics makes it possible to use NIR to resolve highly overlapped and broad peaks, high sensitivity to sample physical characteristics, and high information redundancy (Agelet and Hurburgh. 2010). Also, Multivariate data analysis or chemometrics is thus required to extract suitable information from spectra that would correlate with the measured property (e.g. protein, fat, moisture), for which mathematical procedures are used to remove unwanted information (such as spectral noise or effect of particle size) without losing important or required information (Manley. 2014).

Chemometrics in NIR analysis includes three steps, as follows: (1) spectral data pre-processing; (2) building calibration models; (3) model testing and transfer (Cen and He. 2007; Teye et al. 2020).

Table 2.5 Application of NIR-HSI on authentication of agricultural product and food.

Wavelength (nm)	Sample classification	Algorithm	Accuracy classification	Reference
1,039-1,612	The classification of rice seed varieties including Zhongzheyou No.1, Zhongzheyou No.5, Zhongzheyou No.8, Zhongzheyou No.86	Random Forest (RF)	Both calibration and prediction sets were 100%.	Kong et al. 2013
1,069 1,126, 1,189 1,243, 1,413	The classification of oats, dehulled barley, wheat and rye.	Partial Least Squares Discriminant Analysis (PLS-DA)	95 and 99% of oats and dehulled oats prediction, respectively.	Erkinbaev et al. 2017
924-1,657	The classification of maize seed varieties from two kinds of parental inbred lines including 14 varieties (A1, A2, A3, A4, A5, A6, and 7 as well as B1, B2, B3, B4, B5, B6, and B7)	Least square support vector machine (LSSVM)	Training and test sets were 99.8 and 98.2%, respectively.	Yang et al. 2017
874–1,734	The classification of rice seed varieties including Japonica rice: Xiushui 134, Zhejing 99; Indica rice:Zhongjiazao 17 and Zhongzao 39	Convolutional Neural Network (CNN)	Training and test sets were 89.6 and 87.0%, respectively.	Qiu et al. 2018
874–1,734	The classification of coffee bean varieties including Typic Arabica, Catimor Arabice, Fushan Robusta, and Xinglong Robusta coffee.	Support Vector Machine (SVM)	Calibration and prediction sets were 89.9 and 97.5%, respectively.	Zhang et al. 2018

Table 2.5 Application of NIR-HSI on authentication of agricultural product and food (continued)

Wavelength (nm)	Sample classification	Algorithm	Accuracy classification	Reference
874–1,734	The classification of coffee bean varieties including Typic Arabica, Catimor Arabice, Fushan Robusta, and Xinglong Robusta coffee.	Support Vector Machine (SVM)	Calibration and prediction sets were 89.9 and 97.5%, respectively.	Zhang et al. 2018
1,001 1,025 1,116 1,132 1,156 1,180 1,204 1,220 1,254 1,289 1,301 1,351 1,375 1,405 1,460 1,470 1,564 1,588 1,625	The classification of common maize seeds of four varieties including Datian387 (DT387), Quchen8 (QC8), Quchen11 (QC11), and Quchen13 (QC13). The classification for silage maize seeds of four varieties including Quchen9 (QC9), Quchen19 (QC19), Quchen29 (QC29) and Quchen513 (QC513)	Radial basis function neural network (RBFNN) Support Vector Machine (SVM)	The classification accuracy of common maize seeds was over 97%. The classification accuracy of silage maize seeds was over 98%.	Bai et al. 2020
908-1,735	The classification for black tea quality including seven grades.	Decision tree algorithms (fine tree)	Calibration set and prediction set were 94.22 and 93.13%, respectively.	Ren et al. 2020
874-1,734	The classification for varieties of coated maize kernels including Hongfeng707 (HF707), Jidan50 (JD50) and Jidan (JD53).	Logistic regression (LR)	Calibration set and prediction set were 97.6 and 96.7%, respectively.	Zhang et al. 2020

2.6.1 Spectral data pre-processing

Given the complexity of NIR spectra and small spectral differences between different samples, multivariate data analysis is essential for the effective use of NIR spectroscopy as an analysis technique (Manley. 2014). The pre-processing technique is extremely important as it removes noise and corrects signal weaknesses as a result of the background and physical characteristics of the sample, resulting in uncontrolled variations in the baseline (Mees et al. 2018; Teye et al. 2020). The pre-processing methods are explained as follows:

2.6.1.1 Smoothing

A moving average is used to reduce the effect of noise in the spectra (Lamb and Hurburgh. 1991). The moving average smoothing is calculated using the surrounding data points to recalculate a value for the central point (Andrade-Garda. 2013). Each point is replaced by the average of itself and the points before and after (Andrade-Garda. 2013).

Savizy Golay spectral pre-processing is employed to remove noise while retrieving the original signal structure (Acharya et al. 2016). The Savitzky-Golay method works similarly to the moving average but performs a local polynomial regression, which means that each time the window is moved, a new fit is performed to the points into the window (Andrade-Garda. 2013). The value of the polynomial in the center of the window will substitute for the noisy value (Andrade-Garda. 2013). The degree of the polynomial is fixed before running the filter and most often a polynomial of degree 2 (quadratic) or 3 (cubic) provides satisfactory results (Andrade-Garda. 2013).

Median filter smoothing is used as a preprocessing technique to remove noise from a signal or image (Al-mbaideen. 2018). The basic principle is to use the median value to replace the value of a neighborhood (Li et al. 2019). It replaces each point in the filtered signal with the median of a running window through the signal (Al-mbaideen. 2018). The median value of the window is the middle value of the window (Al-mbaideen. 2018).

2.6.1.2 Normalization

Spectral normalization is applied to reduce scattering influence due to particle size (Fu et al. 2014). Normalization preprocessing serves to fit spectral data to a similar range (0-1) to compensate for differences due to sampling quantity and path length (Mukasa et al. 2019). The Spectral normalization is calculated from the absorbance value of each variable divided by the value of the mean, range (max-min), and the area under the curve for the observation (which is the same for all samples) of spectra data in which there are forms of normalization (Cho. 2013).

2.6.1.3 Baseline

The baseline offset, the value of the lowest point in the spectrum of each sample, is subtracted from all the variables. The result of this is that the minimum value is set as 0 and the rest are positive values. The amount subtracted is different for each peak, since it depends on the offset used to correct the baseline shifted spectra affected by the different particle size and inhomogeneity of the sample (CAMO. 2015).

Linear baseline correction transforms a sloped baseline into a horizontal baseline. The technique is used to point out two variables that should define the new baseline. These are both defined as 0, and the rest of the variables are transformed according to this, with linear interpolation/extrapolation (CAMO. 2015). In multivariate analysis, the procedure is the same in order to start all atomic peaks in the same position, in general at zero scales (Andrade-Garda. 2013).

2.6.1.4 Standard Normal Variate

Standard Normal Variate (SNV) is used to correct shifts in the spectral data due to scattering and differences in the particle size (Candolfi et al. 1999, Sun. 2009). The SNV spectrum is calculated from the average intensity value and subsequently by subtracting this value from each spectrum. After that, the sum of the squared intensities is calculated and the spectrum is divided by the standard deviation of the spectrum (Sandak et al. 2016).

2.6.1.5 Multiple scatter correlation

Multiple scatter correlation (MSC) is used to correct the scatter level of the spectra (Berg and Engelsen. 2009; Mukasa et al. 2019). The concept of MSC is undesirable for scatter effect that will be removed from the data matrix prior to data modeling (Berg and Engelsen. 2009). The multiplicative of scatter is affected by the differences in particle sizes between samples (Rinnan et al. 2009). Both MSC and SNV reduce the spectral variability because of scatter and baseline shifts (Wu and Sun. 2013). The outcome of MSC is very similar to SNV in many cases (Huang et al. 2010). However, many spectroscopists prefer SNV over MSC, since SNV corrects each spectrum individually and does not need the entire data set (Huang et al. 2010). The most important difference between MSC and SNV is that MSC requires a reference spectrum to start the calculation (Pelliccia. 2018). The reference spectrum is ideally a spectrum free of scattering effects; however, the collected unwanted spectrum scattering effects are not easy to remove (Pelliccia. 2018). For this reason, if the data is reasonably well behaved, then it is possible to take the average spectrum as a close approximation to the ideal spectrum (Pelliccia. 2018). MSC pre-processing is done in two steps: (1) find the regression of each spectrum X_i against mean spectrum $\mathbf{X}_m$, which is done by ordinary least squares regression: $\mathbf{X}_i \approx \mathbf{a}_i + \mathbf{b}_i \mathbf{X}_m$; (2) Then calculate the corrected spectrum $\mathbf{X}_i^{msc} = (\mathbf{X}_i - \mathbf{a}_i) / \mathbf{b}_i$ (Berg and Engelsen. 2009; Pelliccia. 2018).

2.6.1.6 Derivative

The first and second derivatives are most often used in practice rather than higher-order derivatives (Huang et al. 2010). Higher order derivatives are not recommended due to the fact that they increase noise in the signal and reduce its magnitude (Romìa and Bernàrdez. 2009). Derivatives are used to resolve peak overlap (or enhance resolution) and eliminate constant and linear baseline drift between samples (Huang et al. 2010). The basic method is as follows: the first derivative is estimated as the difference between two subsequent spectral measurement points; the second-order derivative is calculated by the difference between two successive points of the first-order derivative spectra (Berg and Engelsen. 2009).

Table 2.6 The pre-processing with various software.

Pre-processing	Software	Reference
Smoothing (Savitzky–Golay), SNV, MSC, 1st Derivative	Unscrambler 9.7 (CAMO AS, Oslo, Norway)	Li et al. 2018
Smoothing (Savitzky–Golay, Median Filter), Normalization, SNV, MSC, 1st Derivative, 2nd Derivative	ParLeS software, Version 3.1	Mohammadi-Moghaddam et al. 2018
Smoothing (Savitzky–Golay), Normalize, Baseline correction, SNV, MSC, 1st Derivative, 2nd Derivative	Unscrambler software Version 10.5 (Camo, Oslo, Norway).	Puertas and Vázquez. 2019
Smoothing (Savitzky–Golay), SNV, MSC	Unscrambler X (version 10.4, CAMO Software AS, Norway)	Abu-Khalaf and Hmidat. 2020
SNV, MSC , 1st Derivative	OPUS 7.2 (Bruker, Ettlingen, Germany) software	Benes et al. 2020
Smoothing (Savitzky–Golay), SNV, MSC, 1st Derivative, 2nd Derivative	Python software (version 3.7.0; Python Software Foundation License) using a code developed by the research group	Badaró et al. 2020
SNV, 1st Derivative, 2nd Derivative	MATLAB (v.9.3, R2015b; The Mathworks, Natick, MA) by in-house written functions.	Firmani et al. 2020
SNV, MSC, 1st Derivative, 2nd Derivative	MATLAB (R2010a; The Mathworks, Natick, MA) using PLS_Toolbox8 (Eigenvector Research, Inc)	Fonseca et al. 2020

Table 2.6 The pre-processing with various software (continued).

Pre-processing	Software	Reference
Baseline offset, Linear baseline correction, SNV, MSC	Matlab 2019a (Mathworks, USA), following all chemometric procedures is well-described in Fernandes et al. (2019)	Pereira et al. 2020
MSC	Matlab 2015a (Mathworks, Inc; Massachusetts, USA) according to classification toolbox (version 2.0), developed by Milano Chemometrics and QSAR Research Group (http//michem. disat.unimib.it/chm).	Sampaio et al. 2020
MSC	Unscrambler X 10.1 software package (CAMO AS, Oslo, Norway)	Ren et al. 2020
Smoothing (Savitzky–Golay), Linear baseline, Correction Baseline offset, SNV, MSC	Matlab 2019a (Mathworks, USA) (chemometric procedures is not specified)	Elainy et al., 2021

2.6.2 Calibration models

Pattern recognition or grouping of analytical data could be divided into two main groups comprising unsupervised and supervised methods. The unsupervised method is a strategy to classify data with no supervisor, such as principal component analysis (PCA) and cluster analysis. Conversely, if the membership of objects is known then this is a supervised method called pattern recognition, for example, linear learning machine (LLM), discriminant analysis, the soft independent modeling of class analogies (SIMCA), and support vector machines (SVM) (Otto. 2017).

However, Linear SVM analysis was used for model classification in this thesis, both in concept and theory. Also, a convolutional neural network (CNN) of deep learning models was used for model classification. Moreover, PCA analysis was applied for dimensionality reduction before SVM and CNN analysis in this thesis. In the case of regression, Principal component regression (PCR), Partial least squares regression (PLS), and Support vector machines (SVM) were focused during this thesis.

2.6.2.1 Regression calibration model

1) Principal Component Analysis (PCA)

PCA is used widely in data processing and dimensionality reduction (Zou et al. 2014). The key idea of PCA is to approximate the original matrix $\mathbf{X}$ by a product of two small matrices, with the score ($\mathbf{T}$) and loading ($\mathbf{L}$) matrices, according to equation 2.2 (Otto. 2017):

$$\mathbf{X} = \mathbf{T}\mathbf{L}^{\mathrm{T}} \tag{2.2}$$

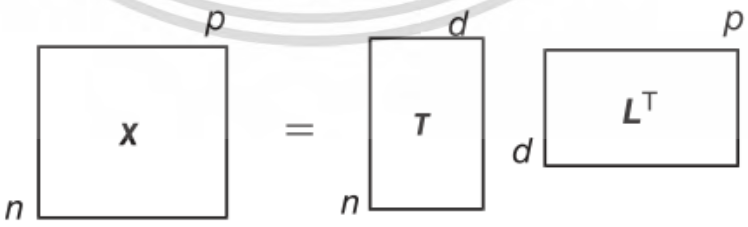

Figure. 2.8 Matrix form of equation 2.2 (Otto. 2017)

where $\mathbf{X}$ is the original data matrix including $\mathbf{n}$ rows (observations) and $\mathbf{p}$ columns (variables). $\mathbf{T}$ is the scores matrix with $\mathbf{n}$ rows and $\mathbf{d}$ columns (number of principal components (PCs)). $\mathbf{L}$ is the loading matrix with $\mathbf{d}$ columns and $\mathbf{p}$ rows, and $\mathbf{L}^{\mathrm{T}}$ is the transposition of a score matrix.

Most of the variance of the data is comprised in the first PC. For the second PC, there is more information than in the third one and so on.

For this, Eq. (2.3) is to be converted to the scores on the left side by

$$\mathbf{T} = \mathbf{XL} \tag{2.3}$$

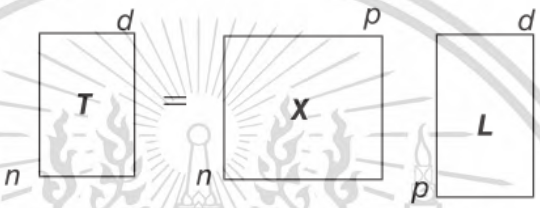

Figure. 2.9 Matrix form of equation 2.3 (Otto. 2017)

$\mathbf{t}$ is called "new spectral variable" and is calculated as

$$\begin{aligned} t_{11} &= x_{11}l_{11} + x_{12}l_{21} + x_{13}l_{31} + \ldots x_{1p}l_{p1} \\ t_{21} &= x_{21}l_{11} + x_{22}l_{21} + x_{23}l_{31} + \ldots x_{2p}l_{p1} \\ &\vdots \\ t_{n1} &= x_{n1}l_{11} + x_{n2}l_{21} + x_{n3}l_{31} + \ldots x_{np}l_{p1} \end{aligned} \tag{2.4}$$

where X_{np} are original spectra and l_{p1} are the weight or loading. Therefore, the elements of the first PC read as t_{11}

2) Principal component regression (PCR)

The principle of PCR is a combination of PCA and Multiple linear regressions (MLR) (Zou et al. 2014). MLR is called a multiple regression, and can be calculated using the following equation (2.5):

$$\hat{y}_i = b_0 + b_1x_1 + b_2x_2 + \ldots + b_ix_i \tag{2.5}$$

where y_i and $\hat{y}_i$ are the measured value and predicted value, respectively. Moreover, the regression coefficient matrix $\mathbf{B}$ of MLR is found by solving:

$$\mathbf{B} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \tag{2.6}$$

The main idea of PCR is to replace the $\mathbf{n} \times \mathbf{p}$ matrix of $\mathbf{X}$ raw data with a smaller $\mathbf{n} \times \mathbf{d}$ matrix of $\mathbf{T}$ scores matrix from PCA (Dunn. 2021) as present in Figure 2.10.

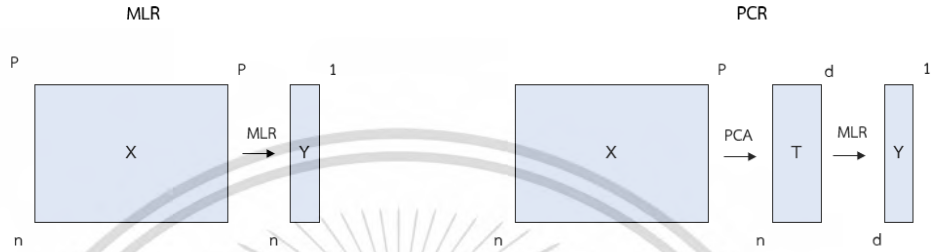

Figure. 2.10 Matrix form of MLR and PCR (Dunn. 2021).

Then, the $\mathbf{T}$ scores matrix is related to the $\mathbf{Y}$ reference value (Dunn. 2021). From equations 2.3 from PCA, the Matrix of referenced value can be calculated by regressing $\mathbf{Y}$ against $\mathbf{T}$ (Romìa and Bernàrdez. 2009) :

$$\hat{\mathbf{Y}} = \mathbf{TB} \tag{2.7}$$

where $\mathbf{B}$ is regression coefficient matrix, which is obtained from least square regression by multiple linear regressions (Romìa and Bernàrdez. 2009), calculated as:

$$\mathbf{B} = (\mathbf{T}^{\mathrm{T}}\mathbf{T})^{-1}(\mathbf{T}^{\mathrm{T}}\mathbf{Y}) \tag{2.8}$$

3) Partial least squares regression (PLS)

PLS regression combines the PCA and MLR (Wold et al. 2001). PLS is built from the relationship of the two matrices $\mathbf{X}$ and $\mathbf{Y}$ by score vectors and is based on the basic latent component decomposition (Boulesteix and Strimmer. 2007). Both the $\mathbf{X}$ and $\mathbf{Y}$ matrices are decomposed into smaller matrices according to:

For $\mathbf{X}$ matrix: the same meanings as given in Eq. (2.2).

For $\mathbf{Y}$ matrix:

$$\mathbf{Y} = \mathbf{TQ}^{\mathrm{T}} \tag{2.9}$$

where $\mathbf{T}$ and $\mathbf{Q}$ are the scores and loading matrix, respectively, of the Y matrix, where $\mathbf{T}$ can be calculated from Eq. (2.3).

The regression coefficient matrix $\mathbf{B}$ will be calculated, and then the predicted valued is estimated as:

$$\hat{\mathbf{Y}} = \mathbf{XB} \tag{2.10}$$

If data set is any dimension, $\mathbf{Q}^{\mathbf{T}}$will be determined as:

$$\mathbf{Q}^{\mathrm{T}} = (\mathbf{T}^{\mathrm{T}}\mathbf{T})^{-1}(\mathbf{T}^{\mathrm{T}}\mathbf{Y}) \tag{2.11}$$

From eq. (2.10) and (2.11):

$$\mathbf{XLQ}^{\mathrm{T}} = \mathbf{XB}$$

Regression coefficient $\mathbf{B}$ matrix can be defined as:

$$\mathbf{B} = \mathbf{LQ}^{\mathrm{T}} \tag{2.12}$$

$$\mathbf{B} = \mathbf{L}(\mathbf{T}^{\mathrm{T}}\mathbf{T})^{-1}(\mathbf{T}^{\mathrm{T}}\mathbf{Y}) \tag{2.13}$$

Therefore, the predicted valued matrix is calculated as:

$$\hat{\mathbf{Y}} = \mathbf{XL}(\mathbf{T}^{\mathrm{T}}\mathbf{T})^{-1}(\mathbf{T}^{\mathrm{T}}\mathbf{Y}) = \mathbf{X}_{\mathrm{pre}}\mathbf{B} \tag{2.14}$$

2.6.2.2 Classification calibration model

1) Linear Support Vector Machines (LSVM)

Kotsiantis et al. (2006) presented that SVMs revolve around the notion of a "margin" either side of a hyperplane that separates two data classes. In the case of linearly separable data, once the optimum separating hyperplane is found, data points that lie on its margin are known as support vector points and the solution is represented as a linear combination of only these points. Other data points are ignored. Hence, the model complexity of an SVM is not affected by the number of features encountered in the training data. Therefore, SVMs are well suited to dealing with learning tasks where the number of features is large with respect to the number of training instances.

Let $S=\{(x_k, y_k), k=1,2,...,n\}$ be a linearly separable set of data, where x_k is the input data of the feature space and $y_k=\{-1,+1\}$ is the class label. As shown in Figure 2.11, the SVM seeks a hyperplane to divide the data into two classes, as follows (Sun and Huang. 2011).

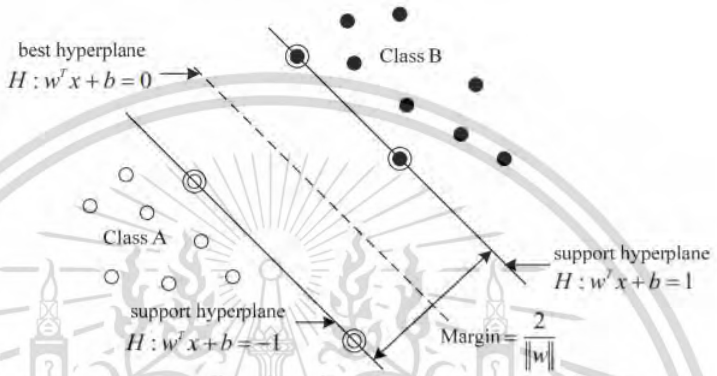

Figure. 2.11 Graphical representation of SVM principle (Sun and Huang. 2011).

Martínez-Ramón and Christodoulou (2006) explained that the expression of the separating hyperplane is then $W^TX=0$. If the data is not centered around the origin, the hyperplane must be made biased by it, so its expression is given by $W^TX+b=0$ (Martínez-Ramón and Christodoulou. 2006). To find the maximum margin between two classes, two support planes are determined (Sathyanarayana and Amarappa. 2014), which demonstrate that the positive class support plane is denoted by $W^TX+b=+1$ and negative class support plane is denoted by $W^TX+b=-1$ for lies in positive class ($y_k=+1$) and negative class ($y_k=-1$), respectively (Sathyanarayana and Amarappa. 2014). One can reformulate the SVM criterion as: maximize the distance between the separating hyperplane and the nearest samples subject to the constraints (Martínez-Ramón and Christodoulou. 2006).

$$y_k[W^TX_i+b]\geq 1 \tag{2.15}$$

where $y_k \in \{+1,-1\}$ is the label associated to the sample X_k.

Class interval is equal to $\frac{2}{\|w\|}$. Maximum class interval is equal to $\|W\|^2$ minimum margin (Liu et al. 2010). Optimal separating surface is able to make $\frac{1}{\|W\|^2}$ minimum subject contain $y_k[W^T X_k + b] \geq 1$ (Liu et al. 2010).

In practical situations, the samples are not linearly separable, so constraint cannot be satisfied (Martínez-Ramón and Christodoulou. 2006). For that reason, slack variables must be introduced to account for the non-separable samples (Martínez-Ramón and Christodoulou 2006). Then, the optimization criterium consist of minimizing the (primal) functional (Burges. 1998) follow as:

$$L_p = \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{N}\xi_i \tag{2.16}$$

subject to the constraints

$$y_i[W^T X_i + b] \geq 1 - \xi_i \tag{2.17}$$

The values of ξ_i must be constrained to be nonnegative. If the sample X_i is correctly classified by the hyperplane and it is out of the margin, its corresponding slack variable is $\xi_i = 0$. If it is well classified but it is within the margin, $0 < \xi_i < 1$. If the sample is misclassified, then $\xi_i > 1$. The value of C is a trade-off between the maximization of the margin and the minimization of the errors (Burges. 1998).

Only two classes exist and binary classification by SVM is enough for many problems that involve classification of multiple classes (Wu et al. 2018). For multi-classification by SVM, the one-against-one method was used in this thesis. This method constructs $k(k-1)/2$ binary classifiers for a problem with k classes, and each binary classifier is trained to distinguish the samples from one class with the other class in each pair of classes (Hsu and Lin. 2002). For example, three classes are represented by red, blue and yellow classes, where three binary classifiers are constructed with hyperplanes and corresponding support vectors formed between the red and blue classes, red and yellow classes, and blue and yellow classes, respectively (Wu et al.2018). After the classification by all the three binary classifiers, the testing data sample

is classified to the class that wins the maximum voting process (Wu et al. 2017).

2) Convolutional Neural Networks (CNN)

The most established algorithms of various deep learning models are convolutional neural networks (CNN), which have been a dominant the method in computer vision tasks since the astonishing results that were shared during the object recognition competition (Patil and Rane. 2021). CNN is most often used to examine visual imagery (Song et al. 2021).

Generally, a CNN consists of three main neural layers, which are convolutional layers, pooling layers, and fully connected layers. Different kinds of layers play different roles (Guo et al. 2016). In Fig. 2.10, a general CNN architecture for image classification (Zhang et al. 2020) is shown layer by layer. A CNN is a hierarchical neural network whose convolutional layers alternate with pooling layers, followed by some fully connected layers (see Figure 2.12) (Guo et al. 2016).

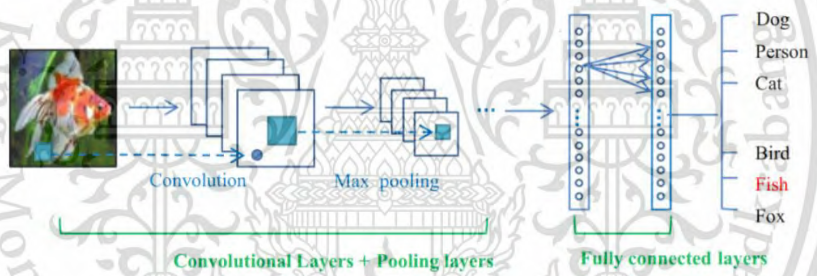

Figure. 2.12 The pipeline of the general CNN architecture (Guo et al. 2016).

The input x of each layer in a CNN model is collected in three dimensions including height (m), width (m), and depth (r), or m × m × r, where the height (m) is equal to the width. The depth is referred to as the channel number. For example, the depth (r) in an RGB image is equal to three (R channel, G channel, B channel) (Alzubaidi et al. 2021).

2.1) Convolutional layer

In CNN architecture, the most important component is the convolutional layer. It consists of a collection of convolutional filters (so-called kernels) (Alzubaidi et al. 2021).

In terms of Kernel definition, Alzubaidi et al. (2021) explained that a grid of discrete numbers or values describes the kernel. Each value is called the kernel weight. The convolutional layers are the foundation of CNN, as they contain the learned kernels (weights), which extract features that distinguish different images from one another. Wang et al. (2020) reported that the padding, kernel size, and stride are hyper-parameters specified by the designers of the network architecture. The convolutional neuron performs an elementwise dot product with a unique kernel learned and the output of the previous layer's corresponding neuron (Wang et al. 2020).

Kattenborn et al. (2021) illustrated that having a large number of parameters in deep networks makes them prone to overfitting and, therefore, the regularization aims to facilitate a network's ability to generalize. Rectified Linear Unit (ReLU) regularizes the network by reducing the parameters of the model as it ignores values < 0. These values are, in theory, not activated. The reduction of parameters also greatly decreases the computing time (Kattenborn et al. 2021). The output weight of the ReLU function equals the weighted sum of the inputs as long as this sum is > 0 (values < 0 are ignored). The output weight of the ReLU function equals the weighted sum of the inputs as long as this sum is > 0 (values < 0 are ignored) (Kattenborn et al. 2021).

The ReLU activation function is a one-to-one mathematical operation (Wang et al. 2020) as shown in Figure 2.13.

2.2) Pooling Layers

This approach shrinks large-size feature maps to create smaller feature maps, maintaining the majority of the dominant information (or features) in every step of the pooling stage (Alzubaidi et al. 2021). There are many types of pooling layers in different CNN architectures, but they all have same the purpose (Wang et al. 2020). The maximum and average pooling are illustrated in Figure 2.14.

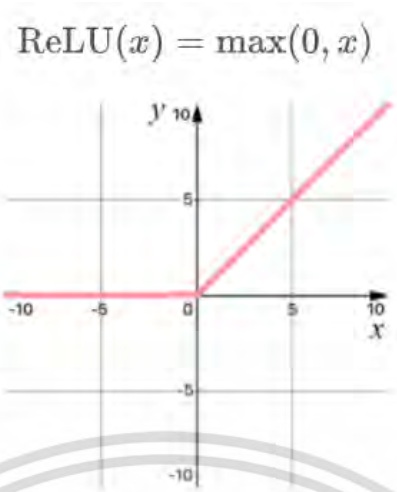

Figure. 2.13 The ReLU activation function, which disregards all negative data (Wang et al. 2020)

Figure. 2.14 These three pooling operations including the average pooling and maximum pooling (Bhatnagar et al. 2020)

2.3) Full Connection layer

Alzubaidi et al. (2021) explained that a full connection layer is located at the end of each CNN architecture and inside this layer, each neuron is connected to all neurons of the previous layer. This was termed the Fully Connected (FC) approach, and is utilized as the CNN classifier. It follows the basic method of the conventional multiple-layer perception (MLP) neural network, as it is a type of feed-forward Artificial Neural Network (ANN) (Alzubaidi et al. 2021). The input of

the FC layer comes from the last pooling or convolutional layer. This input is in the form of a vector, which is created from the feature maps after flattening (Alzubaidi et al. 2021).

Flattening a layer involves converting a three-dimensional layer (last pooling or convolutional layer) in the network into a one-dimensional vector to fit the input of a fully-connected layer for classification (Wang et al. 2020).

MLP consists of a minimum of three layers of nodes, and utilizes the backpropagation technique for its training which is part of the supervised learning method (Savalia and Emamian 2018). Schaijk (2019) demonstrated that for the number of inputs X_1, X_2, ..., X_n for a flattened layer, each value is multiplied by the weight value. All the resulting values are then summed up to a single value including bias (like the constant in linear regression) (Schaijk. 2019). Finally, MLP results have to be turned into an output, which is done by feeding the result to an activation function (Schaijk. 2019).

Wang et al. (2020) explained that the activation function called Softmax performs a normalization on unscaled scalar values, known as logits, to yield output class scores that sum to one. The Softmax function is necessary to calculate the classification score which explains how the logits result from the previous lead to the final classification. This is calculated (Wang et al. 2020) as follows:

$$\text{Soft}\max(X_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \tag{2.18}$$

where X_i represents the values from the neurons of the output layer (MLP results). The exponential acts as the non-linear function. Later, these values are divided by the sum of exponential value X_j in order to normalize and then convert them into probabilities. These are the probability values that a data point belonging to the respective classes (Wang et al. 2020). The outputs of the Softmax function are always in the range of [0, 1] and add up to 1; therefore, they form a probability distribution for each class (Zhou. 2019).

2.6.3 Model testing

2.6.3.1 Performance analysis for regression model

The calibration equation is calculated, as it is required to determine the performance of the prediction model. There are two types of validation comprising internal validation or cross-validation and external validation or test set validation.

The samples are divided into calibration and validation set by the ratio of, for example, 50:50, 60:40, 70:30, 80:20 and so on. Conzen (2006) suggested that when the number of samples is less than 50, the model should be validated by cross validation.

1) Cross-validation

The steps of cross-validation are as follows (Conzen 2006) :

Step 1: Sample is removed from a calibration set.

Step 2: A Calibration model is calculated from the remaining samples.

Step 3: The error of the prediction for a removing sample is calculated and analyzed from $y_{cv,i} - \hat{y}_{cv,i}$ where $y_{cv,i}$ and $\hat{y}_{i,cv}$ are the measured and predicted values of the removed sample's number i respectively.

Step 4: The removed sample is returned to the calibration set. A new sample is then removed, a new model is calculated and a new sample predicted.

Step 5: Step 4 is repeated for all samples of the calibration data set that were removed.

Then root mean square error of cross validation (RMSECV) is calculated following as:

$$\mathrm{RMSECV} = \frac{\sum_{i=1}^{N} (\hat{y}_{cv,i} - y_{cv,i})^2}{N_t} \qquad (2.19)$$

where N_t is number of the total samples.

Krishni (2018) illustrated that the K-fold cross validation method is one means of cross validation. A calibration set is divided into a K number of sections/folds where each fold is used as a validation set at some point. For example, for 5-fold cross-validation (K=5), the data is separated into five folds; the first fold is used to test the model and the remaining are used to train the model. In the second step, 2nd fold is

used as the testing set while the remaining serve as the training set. The process is repeated till each fold of the 5 folds has been used as the testing set.

2) Test set

Conzen (2006) explained that the external validation or test set will be built when the number of samples is more than 50 samples. The calibration model will be built by calibration set and then will be validated by validation set. It is essential to know the performance of the model by the coefficient of determination (R^2), root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), the ratio of prediction to deviation (standard error of prediction to the standard deviation, RPD) and bias.

Moreover, the repeatability (Rep) is used to demonstrate the precision of the reference test for reference value. The Rep is the standard deviation of the difference between duplicates of reference data. The Rep was used to calculate the maximum coefficient of determination (R^2_{max}) by following formula (2.20) (Dardenne. 2010):

$$R^2_{max} = \frac{(SD_y^2 - Rep^2)}{SD_y^2} \quad (2.20)$$

where Rep is the repeatability which is measured by the standard deviation of the difference between duplication, and $\mathbf{SD}_y$ is the standard deviation of the calibration data set. If there is no error in the spectra or model, the coefficient of determination (R^2) will be approximately equal to R^2_{max}. $\mathbf{SD}_y$ and Rep can specify whether a reference laboratory analysis is precise or not (Dardenne. 2010).

3) Coefficient of determination (R^2)

Williams (2019) explained that the showing of the percentage of variance in $\mathbf{X}$ data can be explained by the variance in the $\mathbf{Y}$ data. The unexplained variance is attributable to another factor that cannot be controlled such as sample preparation, reference testing and so on.

The coefficient of determination (R^2) can be calculated via a formula (Carpita. 1996). Moreover, the guidelines for the interpretation of R^2 are shown in Table 2.5

$$R^2 = 1 - \frac{SSR}{TSS} = 1 - \frac{(y_i - \hat{y})^2}{(y_i - \bar{y})^2} \quad (2.21)$$

when SSR and TSS, are residue and total sum square, respectively. The y_i, $\hat{y}_i$ and $\bar{y}_i$ are measured value, predicted value, and mean of measured value, respectively.

Table 2.7 Guidelines for the interpretation of R^2. (Williams et al. 2019).

R^2	Interpretation
Up to 0.25	Not usable in NIR spectroscopy calibration
0.26-0.49	Poor correlation, reason should be researched
0.50-0.64	OK for rough screening
0.66-0.81	OK for screening and some other approximate calibrations
0.83-0.90	Usable with caution for most applications, including research
0.92-0.96	Usable in most applications, including quality assurance
0.98+	Excellent, usable in any application

4) Root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP)

RMSEC and RMSEP are defined via the same calculation process. The difference is a change to a different set of data. The RMSEC was created from the calibration set, while the RMSEP was computed from the prediction set which can be defined as (Williams. 2007; Magwaza et al. 2012):

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{N_c} (\hat{y}_{ic} - y_{ic})^2}{N_c}} \quad (2.22)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{N_p} (\hat{y}_{ip} - y_{ip})^2}{N_p}} \quad (2.23)$$

5) Bias

Bias is an average of residual or the mean difference between measured data and predicted data in a prediction set. It indicates the overall accuracy of the calibration model, which can be calculated (Williams. 2007; Magwaza et al. 2012) as:

$$\text{Bias} = \frac{\sum_{i=1}^{N_p} \hat{y}_{ip} - y_{ip}}{N_p} \quad (2.24)$$

6) Standard deviation of error of calibration (SEC) and standard deviation of error of prediction (SEP).

SEC and SEP are the standard deviation of error of calibration and prediction set which are calculated from the standard deviation of the difference between measured value and predicted value of calibration sets and prediction sets, respectively. It can be calculated as (Williams. 2019):

$$\text{SEC} = \sqrt{\frac{\sum_{i=1}^{N_c} (\hat{y}_{ic} - y_{ic} - \text{Bias}_c)^2}{N_c - 1}} \quad (2.25)$$

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^{N_p} (\hat{y}_{ip} - y_{ip} - \text{Bias}_p)^2}{N_p - 1}} \quad (2.26)$$

7) RPD

Ratio of SEP to standard deviation of measured value of prediction set (SD). It can be calculated as (Williams. 2019):

$$\text{RPD} = \frac{\text{SD}}{\text{SEP}} \quad (2.27)$$

Chang et al. (2001) set a threshold of RPD for predictive models for non-NIR absorber, especially some constituents, to indicate the performance of a model: "excellent models (RPD is greater than 2), fair models (RPD ranges from 1.4 to 2),

and non-reliable models (RPD <1.4." These RPD criteria are soil NIR spectroscopy. Soil is semi-organic and semi-inorganic materials which are difficult to create a NIR model similar to the modeling of grain texture, which is a physical property.

A good model should have a lower RMSEC, and RMSEP or RMSECV, and a higher R or R^2 (Magwaza et al. 2012). Other performance parameters that illustrate a good model are a low Bias (average difference between predicted and measured values) and a small difference between RMSEC and RMSEP (Magwaza et al. 2012).

2.6.3.2 Performance analysis for classification model

The basic concept is the fault in the evaluation of classification models. The total error number in the observed set can be a guide for work for a classifier (Novakovic et al. 2017).

$$\text{Accuracy} = \frac{\text{number of correctly classified examples}}{\text{total number of cases}} \quad (2.28)$$

However, the disadvantages of this accuracy are as follows: (1) it ignores the differences between the types of errors; (2) it is dependent on the distribution of class in the dataset.

The confusion matrix (Koavi and Provost. 1998) contains information about actual and predicted classifications. A confusion matrix of size n x n that is related to a classifier presents the predicted and actual classification, where n is the number of different classes (Visa et al. 2011).

Ruuska et al. (2018) reported that for a binary confusion matrix (see Figure 2.18), observations classified correctly into the positive class are called true positives and observations classified correctly into the negative class are called true negatives. Instances of the positive class classified falsely as negative are called false negatives and instances of negative class classified falsely as positive are called false positives (Ruuska et al. 2018). Numbers of true positive, false positive, true negative and false negative observations are notated by TP, FP, TN and FN (Ruuska et al. 2018).

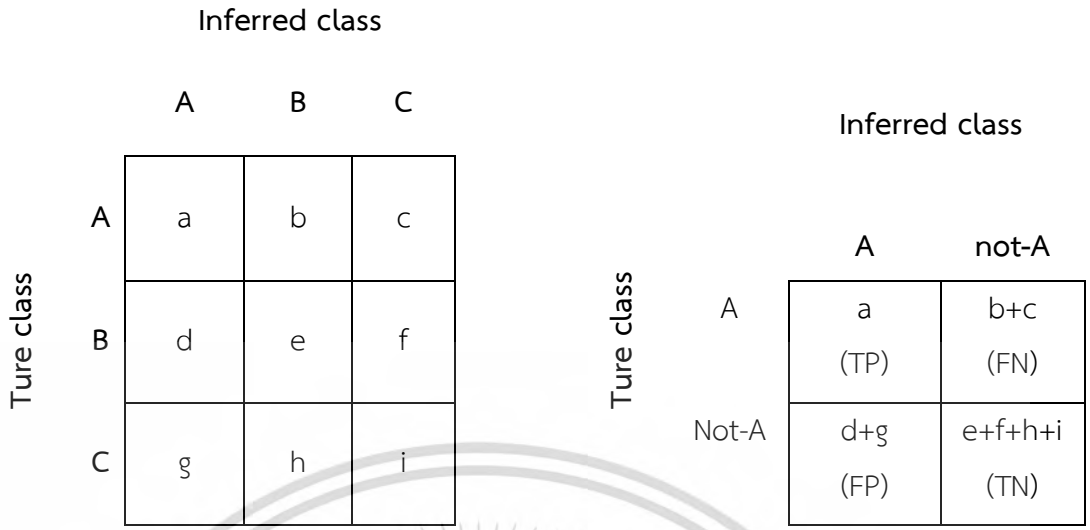

Figure. 2.15 An example of a 3x3 confusion matrix for classes A, B and C (left) and the corresponding binary confusion matrix for class A (right) (Ruuska et al. 2018)

Chapter 3

Near Infrared Scanning Precision Test for Texture Characteristics of Parboiled Rice

This chapter illustrates the precision of FT-NIR spectrometers in assessing milled parboiled rice, as well as the precision analysis of reference laboratories for textural properties of cooked parboiled rice. The aim of this chapter is to ensure that the NIR spectral recording from an FT-NIR spectrometer and standard reference test of textural properties by our laboratory are precise, and are suitable for creating a model for predicting the texture of cooked parboiled rice.

* This chapter constituted the by publication article: Onmankhong, J., and Sirisomboon, P. 2018. "Near Infrared Scanning Precision Test for Texture Characteristics of Parboiled Rice." **The 11th Thai Society of Agricultural Engineering International Conference: TSAE 2018,** 26-27 April 2018, 144-147.

3.1 Introduction

Rice is the most widely consumed staple food for nearly half of the world's population for centuries (Ghasemi et al. 2009). About 20% of the world rice production is processed as parboiled rice. The parboiling process improves structural and nutritional properties of rice. (Villanova et al. 2017). In 2017, Thailand exported over 10 million ton of rice (Thai Rice Exporters Association 2017) including the parboiled rice. The parboiling process consists of three additional steps to conventional rice processing, which are: soaking, pressure steaming and drying (Leethanapanich et al. 2016; Paiva et al. 2016; Sarangapani et al. 2016; Villanova et al. 2017). Soaking and steaming regimes affected texture cooked rice of parboiled rice (Graham-Acquaah et al. 2015). Among multiple texture attributes of the cooked rice that may affect consumer acceptability including hardness and toughness (Juliano 1981). Therefore, the texture of cooked rice is considered important factors for the consumer. The instruments have been used to measure the texture of cooked rice including Texture Profile Analysis (TPA). Another instrumental method that can assess rice properties, including cooked rice texture, is near infrared spectroscopy (NIRS) (Siriphollahul et al. 2017). The NIR is the nondestructive technique. This method has been used for extremely precise apparent amylose content (Villareal et al. 1994), protein content (Shuso et al. 2003), amino acid content (Wu 2002), gelatinization temperature, gel consistency (Bao et al. 2001), and rapid visco analysis (RVA) parameters (Frederick and Franklin 2002; Sirihollakul et al. 2015). Most of the rice quality prediction models have high precision and accuracy (Siriphollakul et al. 2017). It is interesting to make the NIR spectroscopy model development for prediction the texture of parboiled rice. Before NIR spectroscopy models development, it is needed to check for both repeatability and reproducibility of the spectral data and the repeatability of reference test. The objective of this research was to evaluate the overall precision of FT-NIR spectrometer scanning and that of reference test for measuring the texture characteristics of parboiled rice.

3.2 Materials and methods

3.2.1 Samples

Parboiled rice samples were collected from Capital Chainat Rice Mill company limited, in Saphaya District, Chainat Province, Thailand. A total of 7 samples were collected on 9, 10, 11, 12, 14, 15 and 16-Oct-2017.

3.2.2 NIR scanning

FT-NIR spectrometer (MPA, Bruker, Germany) in the wavenumber range of 12500 - 4000 cm^{-1} were used for scanning of parboiled rice in absorbance mode at room temperature of 25±2 ºC, where the reference material was gold. Parboiled rice sample was scanned in a quartz cup with the diameter and height of 9.7 cm and 9 cm, respectively.

3.2.3 Repeatability and reproducibility of NIR scanning

In the NIR repeatability of scanning test, the samples were loaded and scanned 10 times in same location to check the difference of parboiled rice spectrum. For the reproducibility, the samples were scanning the sample for 10 times reloaded and rescanned every time to check the homogeneity of parboiled rice sample. The calculating was done by first selected 3 wavenumbers at 6897, 6711 and 5154.6 cm^{-1} which were the vibration bands of starch (the first two wavenumbers) and water of parboiled rice samples. Then the standard deviation (SD) of absorption at each wavenumber was calculated and averaged. The SD was repeatability or reproducibility depended on how the spectra was obtained.

3.2.4 Reference method

The reference method used was ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking, using texture analyser (TA.HD.plus, Stable Micro Systems, UK). The texture of cooked parboiled rice including hardness, toughness distance and toughness were evaluated.

3.2.5 Repeatability of reference test and maximum coefficient of determination

The repeatability (Rep) of reference test was the standard deviation of the different between the duplicates of the test of 7 samples. Then the repeatability was

used to determine the maximum coefficient of determination (R^2_{max}), which was calculated using following formula (Dardenne 2010) using equation (3.1)

$$R^2_{max} = \frac{SD_y^2 - Rep^2}{SD_y^2} \quad (3.1)$$

where SD_y is the standard deviation of data of measured value in 7 samples. According to Dardenne 2010, the maximum R^2 is possible only when there are no error in the spectra or the model and SD_y and Rep can indicate that the range in value for the samples is too narrow, and/or reference method is not sufficiently precise.

3.3 Result and Discussion

Table 3.1 show the standard deviation of repeatability and reproducibility for scanning of parboiled rice using FT-NIR spectrometer at 3 wavenumbers were 6896.6, 6711.45, 5153.16 cm^{-1} that were selected. The absorbance band at 6896.6 cm^{-1} (1450 nm) and 6711.45 cm^{-1} (1490 nm) region were the absorption bands of starch (Workman and Weyer 2008). The absorbance band at 5153.16 cm^{-1} (1940 nm) region was of water band (Figure 3.1). The scanning repeatability and reproducibility were 0.002769 and 0.017151 respectively, which indicated acceptable. Low repeatability value indicated highly precise scanning instrument.

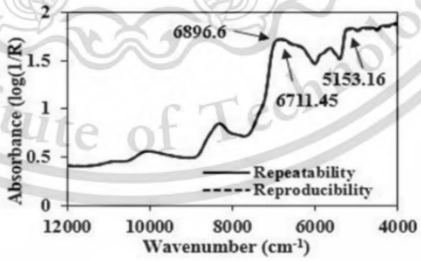

Figure. 3.1 Average spectra of FT-NIR spectrometer for parboiled rice.

Table 3.1 Repeatability and reproducibility of NIR scanning using FT-NIR spectrometer for parboiled rice.

Samples	Repeatability				Reproducibility			
	Wavenumbers (cm^{-1})			Average	Wavenumbers (cm^{-1})			Average
	6896.6	**6711.45**	**5153.16**		**6896.6**	**6711.45**	**5153.16**	
1	0.00277	0.00091	0.00119	0.00163	0.01899	0.01945	0.02057	0.01967
2	0.00257	0.00753	0.00124	0.00378	0.01310	0.01326	0.01402	0.01346
3	0.02053	0.00071	0.00093	0.00739	0.00803	0.00830	0.00852	0.00830
4	0.00350	0.00287	0.00231	0.00290	0.01979	0.01953	0.02069	0.02000
5	0.00082	0.00071	0.00071	0.00074	0.01282	0.01298	0.01331	0.01304
6	0.00147	0.00161	0.00125	0.00143	0.00956	0.01037	0.01210	0.01068
7	0.00143	0.00156	0.00158	0.00152	0.09444	0.00437	0.00595	0.03492
		Average		0.00277		Average		0.01715

Table 3.2 shows the precision for the reference method i.e. ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking. The repeatability were 4.45 N, 0.20 mm and 61.56 N mm for hardness, toughness distance and toughness, respectively. The R^2_{max} of hardness, toughness distance and toughness were 0.94, 0.84 and 0.81 respectively which were calculated from equation (3.1). The R^2_{max} was possible only when there is no error in the spectra or model (Dardenne 2010). The error from reference method was 6, 16 and 19 % for hardness, toughness distance and toughness, respectively.

Table 3.2 The repeatability of hardness, toughness distance and toughness respectively by the reference method i.e. ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking.

Texture	Repeatability
hardness	4.45 N
toughness distance	0.20 mm
toughness	61.56 N mm

3.4 Conclusions

From the results presented in this study, the repeatability value was better than the reproducibility value of the NIR scanning using FT-NIR spectrometer for parboiled rice which may be due to non homogeneous of samples. However, the values indicated that the FT-NIR spectrometer for parboiled rice was precise. The precision analysis for reference laboratory indicates that good model can be developed for hardness but fair model for toughness distance and toughness.

Chapter 4

The influence of processing parameters of parboiled rice on its physiochemical and texture properties

The aim of this chapter is to study how different parboiled rice conditions affect the chemical composition of parboiled rice. The relationship between chemical and textural properties were analyzed to ensure that different spectral information i.e. different chemical compositions can be applied to NIR spectroscopy to evaluate the textural properties of parboiled rice. Moreover, this chapter shows how different conditions of the parboiled process influence the physicality (whiteness and yellowness) of milled parboiled rice.

* This chapter constituted the publication article: Onmankhong, J., Jongyingcharoen, J. S., and Sirisomboon, P. 2021. "The influence of processing parameters of parboiled rice on its physiochemical and texture properties." **Journal of Texture Studies**. 52(2): 219-227.

4.1 Introduction

The staple food for almost half of the world is rice (*Oryza sativa* L.) (Oli et al. 2014), and 20% of the world's rice is used to produce parboiled rice (Zhu et al. 2020). Parboiled rice is also a major Thai export, and is exported from Thailand to other parts of the world. Parboiled rice is preferred in Africa, particularly West Africa, in other parts of Asia, and in Caribbean countries, particularly amongst Indian emigre (Akhter et al. 2014). Each country requires different attributes to signify quality in parboiled rice products (Calingacion et al. 2014). The parboiling process includes soaking raw rice in hot water, steaming for finishing the gelatinization, and drying (Islam et al. 2002; Parnsakhorn and Noomhorm. 2008; Akhter et al. 2014). The advantages of the parboiled rice process are the strengthening of grain integrity, an increasing in milling restoration and a decreasing in cooking losses (Bhattacharya. 1985; Bello et al. 2006).

The raw material used to create parboiled rice is paddy. The physicochemical properties of paddy have been found to change after three months of storage (Saikrishna et al. 2018). During storage, the texture properties have also been found to change, such as hardness and cohesiveness (Saikrishna et al. 2018). When aged rice is used in the production process of parboiled rice, the rice has a dark color that is required by consumers in some countries. Ageing is known to aid in higher volume expansion and water absorption upon cooking, and in obtaining a harder and less sticky cooked rice (Saikrishna et al. 2018). Parboiled rice has more structural and nutritional properties than milled rice (Villanova et al. 2017). However, the parboiled rice process negatively influences the color of grain (Kwofie and Ngadi. 2017), and the soaking process effects the discoloration of the grain (Villanova et al. 2017). Moreover, the color changing of parboiled rice may occur due to enzymatic reactions in the soaking process (Villanova et al. 2017). For the exportation of parboiled rice, the whiteness and yellowness of grain was evaluated for consumer acceptability (direct communication with the parboiled rice factory). The criteria of each customer are different; therefore, factories need to consider the parboiled rice production conditions.

Also, the texture properties of cooked rice are extremely important to its eating quality, which effects the acceptance by consumers (Li et al. 2016). The texture properties of cooked rice can be evaluated by a sensory test using a panelist who has been trained. This method has disadvantages in the sense that high labour intensity is required. Another method involves texture properties including the hardness and stickiness of cooked rice being measured by 90% deformation using a texture analyzer (Diao et al. 2019). The texture profile analysis (TPA) method has been used to measure texture properties including hardness, adhesiveness, cohesiveness, and chewiness of cooked rice (Xu et al. 2019; Ye et al. 2019). The hardness and adhesiveness of cooked rice has been measured using texture profile analysis (Sarangapani et al. 2015; Zhu et al. 2020) and the stickiness has also been measured by texture profile analysis (Zhu et al. 2020).

However, the hardness of rice depends on the amylose content, which is a significant difference between milled rice and parboiled rice (Sarangapani et al. 2015). The reduced hardness of cooked parboiled rice is due to the decreasing of amylose content of the rice, and the changed stickiness of parboiled rice is due to the different proportions of amylose and amylopectin (Sarangapani et al. 2015). Also, the hardness of cooked rice is affected by fat content (Kim et al. 1986), as the fat content of cooked rice has been found to have a negative impact on the hardness of the rice (Cameron and Wang. 2005). The different conditions of the parboiled rice process change the physical and chemical properties of the rice (Oli et al. 2014).

The goal of this study was to investigate the effect of the different conditions of the parboiled rice process on the physical properties (whiteness and yellowness) chemical properties (amylose and fat content) of rice. In addition, the relationship between chemical and textural properties (hardness and toughness) were analyzed. The hardness and toughness of the cooked rice were the parameters of the texture standard of parboiled rice, defined by ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after the Cooking method (International Organization for Standardization [ISO]. 2012). The findings would be

useful for the consideration of suitable conditions for the parboiled rice process in the parboiled rice products industry.

4.2 Materials and methods

4.2.1 Rice Materials

Paddy samples of the Suphanburi 1variety were collected from Ratchaburi Rice Seed Center, Ratchaburi province, Thailand. The samples were harvested in January of 2018. Before parboiling, the sample was aged for 6 months. The storage temperature for keeping the rice for aging was approximately at 28 °C. The 500 g sample was randomly taken from a 50 kg sized sack, and there were 25 samples in a total. The average moisture content of the samples was 12±1 %wb, measured by moisture meter (Model TA-5 OGA Electric Co., Ltd., Tokyo, Japan).

4.2.2 Parboiling Process

The three steps involved in the parboiling process are soaking in water, steaming and drying. In this study, the experiment was divided into two parts. The first part was the study of the effect of the soaking conditions, including soaking time (two, three and six hours) and soaking temperature (65 and 75°C), on the physical and chemical properties. The soaking temperatures of 65 °C with the soaking time of 135 min was suitable for completing the gelatinization of parboiled rice in the steaming process (Buggenhout et al. 2014). The soaking process was to prepare the starch enough amount of water for gelatinization in the steaming process. This condition of soaking provided the moisture content in grain to be around 34% (Buggenhout et al. 2014). The soaking temperature of 75 °C and longer time than 135 min was tried in this experiment for more severe soaking condition, which was expected to gain more water content in grain to assure the completion of gelatinization of starch in grain. In this part, the steaming condition was fixed at 100 °C for 20 minutes to ensure the absence of white belly, which would indicate complete gelatinization of parboiled rice (Kwofie, and Ngadi. 2017). The second part of the experiment had to do with a soaking temperature and time fixed at 65 °C for 6 hours, under which the moisture content of soaked paddy was about 30%, as recommended by Islam et al. 2002. The effects of

steaming conditions in this section included the investigation of steaming time (10 and 20 minutes) and steaming temperature (90 and 100 °C). Factorial experimentation with a Completely Randomized Design (CRD) was conducted for both experimental parts.

During the parboiled rice process, a paddy sample of 500 g was deposited into a stainless-steel basket and soaked in a water bath (WB350, Memmert, Germany) at the predetermined soaking conditions. The ratio of rice to water in this experiment was 4: 7. 6 by weight following the ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after the Cooking method (International Organization for Standardization [ISO]. 2012). After soaking, the sample was taken out of the water and placed on a perforated plate above the level of the water for steaming at the predetermined conditions. Afterwards, the sample was placed on a stainless-steel tray for the drying step. The drying sample was tempered at room temperature (25±1°C) in an oven (UF260, Memmert, Germany) as long as the moisture content was 13±1%wb. After drying, the parboiled paddy was permitted to stand for a time (48 hours) before milling. The parboiled rice samples were shelled with a rice husker (THU, SATAKE TESTER, Japan) and then were milled for 30 seconds by a polisher (TM05, SATAKE, Japan).

Table 4.1 The experiment conditions of parboiled rice process.

Process	Soaking effect		Steaming effect	
	Temperature (°C)	Time (hours)	Temperature (°C)	Time (min)
Soaking	65 and 75	2,3 and 6	100	20
Steaming	65	6	90 and 100	10 and 20

4.2.3 Physical properties (Whiteness and yellowness)

The whiteness and yellowness (b*) of the milled non-parboiled rice and parboiled rice samples were acquired using a whiteness meter (MM1D, SATAKE, Japan) and a spectrophotometer (MiniScan EZ, HunterLab, USA), respectively. For the spectrophotometer, the D65 illuminant and 10° observer angle were used.

4.2.4 Chemical properties (Amylose and fat content)

The milled non-parboiled rice and parboiled rice samples were evaluated for amylose and fat content at the Food and Nutrition Laboratory at the Institute of Nutrition in Mahidol University, Nakhon Pathom province, Thailand. The amylose content of the milled non-parboiled rice and parboiled rice was determined by method of Juliano, B.O.1971 (Juliano. 1971). The fat content of milled parboiled rice was determined using the AOAC (2016) 945.16 method (Association of official analytical chemists [AOAC]. 2012).

4.2.5 Parboiled rice cooking and texture measurement

Parboiled rice cooking and texture measurements were referenced by an ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after the Cooking method (International Organization for Standardization [ISO]. 2012). Before cooking, defect parboiled milled rice kernels were selected and removed by hand to avoid influencing the textural properties. 20 g of milled parboiled rice was placed in a beaker (100 ml size), and then 38 ml of distilled water was added (4: 7.6 rice-to-water ratio by weight). It was stirred and then covered with watch-glasses. The beaker was placed on a sieve in the cooking container, which was above the water level and in contact with the steam, for 20 minutes. Then the heat source was turned off and the sample beaker was left in the cooking container for 10 minutes. After that, the sample beakers were removed from the cooking container and then put upside down on the watch-glasses immediately. The sample beakers were left at room temperature at 20-25°C for one hour. The 5 g of cooked parboiled rice samples were put in a plastic box for the texture properties test.

The extrusion test was used for the cooked milled non- and parboiled rice texture using a texture analyser (TA.HD. plus, Stable Micro Systems, UK) according to the ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking method. Two texture property parameters were obtained from force-distance curve: hardness (maximum force; N) and toughness (area under the graph from initial to maximum force point; Nmm). The average values of three replicates were reported as the hardness and toughness of each sample.

4.2.6 Statistical Analysis

The properties of milled non- and parboiled rice measured in this research were tested in triplicate. In this study, the two-way ANOVA and Duncan's multiple range test were performed to analyze the effect of temperature and period of soaking and steaming processes on the whiteness, yellowness, amylose content and fat content. The correlations between the chemical and texture properties were analyzed by Pearson, Spearman and Kendall's tau test. The data were analyzed using IBM SPSS statistics 26 (IBM, New York, USA).

4.3 Results

4.3.1 The influence of the soaking and steaming processes on the physical, chemical and textural properties of parboiled rice.

4.3.1.1 The influence of the production process on the textural properties of cooked parboiled rice.

As there was no interaction effect between the soaking time and soaking temperature, the different textural properties of parboiled rice with varied soaking times and varied soaking temperatures are shown in Table 4.2. It was found that the hardness and toughness increased with increasing time of soaking process. However, the hardness and toughness were not significantly ($p>0.05$) changed with temperatures of soaking process. In terms of steaming conditions, an interaction effect between time and temperature was observed (Table 4.3). It was found that the hardness and toughness increased with increasing steaming time. The gelatinization of rice occurred when the moisture content of the paddy was adequate (24-30%wb) (Oli et al. 2014). When the moisture content reached the equilibrium point, the increasing of soaking time combined with soaking temperature might account for the gelatinization during the soaking process. This therefore increased the hardness and toughness of the cooked parboiled rice with increasing time of soaking process.

4.3.1.2 The influence of the production process on whiteness and yellowness (b*)

The whiteness of parboiled rice has an influence on the price of exportation

(Parnsakhorn and Noomhorm. 2008). The whiteness of non-parboiled rice was 44.08% and it was decreased when the rice was parboiled. From two-way ANOVA, there was an interaction effect between soaking time and soaking temperature. Table 4.4 shows the average value of whiteness, yellowness, amylose and fat content of parboiled rice effected by the process conditions. The whiteness was decreased with the increasing duration of the soaking process when the soaking temperature was constant ($p<0.05$). Also, the whiteness was reduced with the increasing temperatures of soaking and steaming when the treatment time was constant. On the other hand, the whiteness of the parboiled rice was not significantly changed with steaming time ($p>0.05$), even if there was an interaction effect between the steaming conditions.

The yellowness of parboiled milled rice effects consumer acceptance of rice exports. The yellowness of non-parboiled rice was 22.45, which was less than that of the parboiled rice. The yellowness increased with the increasing of the temperature and time of the soaking process (Table 4.4). The different time and temperature of the steaming process did not affect ($p>0.05$) the yellowness of parboiled rice (Table 4.4).

Table 4.2 Effect of period and temperature of soaking conditions on the texture properties of cooked parboiled rice.

	Hardness (N)	Toughness (Nmm)
Soaking time (hr)		
2	126.84^{b}	552.12^{b}
3	130.66^{ab}	590.19^{b}
6	143.38^{a}	646.87^{a}
Soaking temperature (°C)		
65	133.05^{ns}	589.08^{ns}
75	134.21^{ns}	603.70^{ns}

The lowercase letters in the same column of each parameter are different significantly different ($p\leq0.05$), The uppercase letters in the same row of each parameter are different significantly ($p\leq0.05$). ns mean no significant difference.

Table 4.3 Effect of period and temperature of steaming conditions on the texture properties of cooked parboiled rice.

	Hardness (N)		Toughness (Nmm)	
	90 °C	**100 °C**	**90°C**	**100°C**
10 min	$116.53^{b,B}$	$123.99^{b,A}$	$575.52^{a,NS}$	$518.77^{b,NS}$
20 min	$137.5^{a,NS}$	$146.20^{a,NS}$	$440.92^{b,B}$	$634.76^{a,A}$

The lowercase letters in the same column of each parameter are different significantly different ($p \leq 0.05$), The uppercase letters in the same row of each parameter are different significantly ($p \leq 0.05$). NS mean no significant difference.

4.3.1.3 The influence of the production process on amylose and fat content.

The amylose and fat content of non-parboiled rice was 35.73% and 2.06%. The amylose content of the parboiled rice was influenced by the altering conditions. The amylose content was reduced with the increase in time of the soaking process (Table 4.4). However, the different temperatures of soaking did not significantly ($p > 0.05$) affect the amylose content. As displayed in Table 4.4, the temperature of the steaming process did not have a significant ($p>0.05$) effect, while the time of the steaming process showed significant ($p<0.05$) influence on the amylose content of parboiled rice. The amylose content significantly increased with the increasing time at 90℃ of the steaming process. However, the steaming conditions at 100 ℃ revealed that the amylose decreased with increasing time.

Table 4.4 also shows the effect of different parboiled conditions on the fat content of the parboiled rice. In this study, the fat content was increased with increasing time of the soaking process, while it was decreased with increasing temperature. The fat content was not significantly ($p>0.05$) changed at 90℃ with increasing of steaming time, but the fat content was increased when increasing steaming time at 100℃ steaming temperature. The fat content at 10 min of steaming time was decreased with increasing the steaming temperature. On the other hand, the fat content at 20 min was increased with increasing the steaming temperature.

Table 4.4 Effect of parboiled conditions on physical and chemical properties.

Time	Temperature							
	Whiteness (%)		Yellowness		Amylose (%)		Fat (%)	
Non-parboiled rice								
	44.08		22.45		35.73		2.06	
Soaking condition with 100 °C of steaming temperature for 20 minutes								
	65 °C	75 °C	65 °C	75 °C	65 °C	75 °C	65 °C	75 °C
2 hr	$37.37^{a,A}$	$29.79^{a,B}$	$26.85^{b,B}$	$28.74^{b,A}$	$37.08^{a,NS}$	$37.18^{a,NS}$	$1.6^{c,A}$	$1.3^{c,B}$
3 hr	$35.43^{b,A}$	$29.93^{a,B}$	$27.21^{b,B}$	$28.18^{b,A}$	$35.78^{b,NS}$	$35.84^{b,NS}$	$2.11^{b,A}$	$2.05^{b,B}$
6 hr	$29.80^{c,A}$	$28.17^{b,B}$	$30.61^{a,B}$	$31.62^{a,A}$	$34.69^{c,NS}$	$34.66^{c,NS}$	$2.52^{a,A}$	$2.19^{a,B}$
Steaming condition with 65 °C of soaking temperature for 6 hours								
	90 °C	100 °C	90 °C	100 °C	90 °C	100 °C	90 °C	100 °C
10 min	$33.03^{ns,A}$	$29.81^{ns,B}$	$27.43^{ns,NS}$	$29.11^{ns,NS}$	$35.11^{b,NS}$	$35.73^{a,NS}$	$2.40^{ns,A}$	$2.01^{b,B}$
20 min	$33.13^{ns,A}$	$29.80^{ns,B}$	$28.18^{ns,NS}$	$30.61^{ns,NS}$	$35.61^{a,NS}$	$34.69^{b,NS}$	$2.40^{ns,B}$	$2.52^{a,A}$

The lowercase letters in the same column of each parameter are different significantly ($p \leq 0.05$). The uppercase letters in the same row of each parameter are different significantly ($p \leq 0.05$). ns or NS mean no significant difference

4.4 Discussion

The whiteness of non-parboiled rice was more than that of parboiled rice. In addition, the dark color of parboiled rice was changed by high temperatures and long steaming times (Bhattacharya. 1985). The lower temperatures and periods for the steaming and soaking process resulted in good quality production (Islam et al. 2002). Parnsakhorn and Noomhorm (2008) reported that the whiteness decreases following an increased soaking period, soaking temperature and steaming period. The results of this research correspond to those of the above research, except that the steaming time did not significantly ($p>0.05$) affect the whiteness, which might be altered by different rice varieties. In addition, this study showed that an increased steaming temperature causes a decrease in whiteness.

In this study, the increasing yellowness of parboiled rice production depended on increasing the period and temperature of the soaking process. Concordantly, Bhattacharya and Subba. 1966 reported that the period and temperature of soaking of different parboiled brown rice samples affected the adjustment of color from white to yellow. Furthermore, Parnsakhorn and Noomhorm (2008) reported that the yellowness of parboiled rice was also affected by the soaking temperature, soaking period and steaming period, as a dark yellowness appeared at high soaking temperatures and periods. Similarly, Leethanapanich et al. (2016) reported that the yellowness increased with increasing soaking temperature and steaming time. Moreover, no significant effect of the steaming process on the yellowness ($p>0.05$) was found in this study. However, no reports showed that the yellowness was not affected by the steaming process. The dark amber colored parboiled rice was possibly caused as a result of the distribution of husk color toward the endosperm (Lamberts et al. 2006). Other reasons for the color change were that the parboiled rice process improved the level of reducing sugar and free α -amino nitrogen (FAN), and that the isomerisation of glucose to fructose suggests the possibility of non-enzymatic Maillard type of browning (Lamberts et al. 2008). On the other hand, the yellowness in this research was quite high because of ageing rice. The diffusion of husk and bran pigments and Maillard reaction caused color change in the parboiling process (Oli et al. 2014). Husk and bran pigments were diffused into endosperm in the process, as aged (six months post-harvest storage) paddy was used in this study. The longer period of storage of paddy caused the dark color of the husk

and bran. Therefore, the dark color of the husk and bran were the reason for the increase of yellowness. This indicated that the different conditions of the parboiled process influenced the degree of color, which corresponds to the studies of Jayanayanan (1965), Pillaiyar and Mohandoss (1981), Tolaba and Su (2001), Lv et al (2009).

In this study, soaking temperature had no significant effect on the amylose content of parboiled rice ($p>0.05$). However, the amylose content was significantly decreased when longer periods of soaking process of parboiled rice were used ($p<0.05$). As compared to the non-parboiled sample, of which the amylose content was 35.73%, soaking for 2 hours provided the parboiled rice with higher amylose content while lower amylose contents were observed in the 3 and 6 hour soaked parboiled rice. Higher amylose content of the 2-hour soaked parboiled sample could be due to the formation of insoluble amylose-lipid complex during the first period of parboiling process. Milasinovic et al. (2010) also explained that during heating the rice, amylopectin was debranched and resulted in linear chain amylose. In contrast, soaking rice for more than 3 hours caused different effect on amylose content of the parboiled rice. Longer soaking time led to discharge of amylose content into the soaking hot water and lower amylose content was observed (Patindol et al. 2008).

For the gelatinization during the steaming process, swollen starch granules are dissolved and then change the crystallinity of starch (Oli et al. 2014). The amylose content in this study did not vary with the steaming temperature (90°C and 100°C). The amylose content increased with the gelatinization temperature, as reported by Varavinit et al. (2003) (68.05-76.53°C for Thai rice varieties) and by Park et al. (2007) (59.6-71.9°C for rice varieties from California and Texas states, USA). However, the steaming temperature did not affect the amylose content. This might be because the steaming temperature exceeded the gelatinized temperature. For an Iranian variety of long rice paddy (Fajr) (Taghinezhad et al. 2016), with steaming time, the degree of starch gelatinization increased gradually and gelatinization enthalpy (ΔH) of parboiled rice decreased gradually. Gelatinization enthalpies tend to have an inverse relationship to amylose content (Morrison and Azudin. 1987). This indicates that when the steaming time increases, the amylose content should decrease, which corresponds to our results at the steaming temperature of 100°C. However, at a steaming temperature of 90°C,

the amylose content increased with the steaming time. The mild-parboiling treatment demonstrated the formation of crystalline amylose-lipid complexes (Liu et al. 2020) which might be due to the migration of lipids from bran to endosperm and the interaction of them with starch during parboiling (Cheng et al. 2019). This phenomenon might prevent the amylose content from leaking to the outer layer of endosperm, hence the increase of amylose content.

Lipids in rice are classified into two types: non-starch and starch lipids. The non-starch lipid- and starch lipid contents are 2.9–3.4%db and 0.66–0.76%db, respectively (Oli et al. 2014). The major lipids stay in the aleurone layer at approximately 60% of the non-starch lipid (Godber and Juliano. 2004; Oli et al., 2014). During parboiling process, the fat moved out from the rice kernel but water soluble nutrient diffused into the endosperm of kernel (Roy et al. 2011). Research has found that the external movement of oil is expedited by high temperatures (Chakravarty et al. 1966) and long periods of soaking (Vasan et al. 1971). Fat moves from the surface kernel because of the broken bodies of the non-starch lipids (Oli et al. 2014). Therefore, more fat in the aleurone layer of the bran causes difficulty in paddy milling (Patindol et al. 2008). Moreover, the fat content was decreases at high soaking temperatures (Sareepuang et al. 2008), whereas an increased steaming period increases the fat content (Ibukun. 2008). Similarly, the results of fat content in this study corresponded with the research above. The fat content was increased when increasing the steaming time at a 100°C steaming temperature. This corresponds to the result of Hasbullah et al. (2016) and Fonseca et al. (2014) which found that the fat content increases with the steaming duration at 100°C and at 120°C, respectively. In our experiment, when the steaming temperature was 90°C the fat content was not different significantly ($p>0.05$) at 90°C steaming temperature with an increase of steaming time. This might be because the temperature was not as high as the temperature used by Hasbullah et al. (2016) and Fonseca et al. (2014). The fat content at 10 min of steaming time decreased when increasing the steaming temperature (90°C to 100°C) and at 20 min of steaming time, the fat content increased while increasing the steaming temperature. In our experiment, the dried milled parboiled rice samples were subjected to total fat evaluation. The fat content of the samples at 10 min steaming time was decreased, which corresponds to Oli et al. (2014)'s findings that during parboiling, the lipid bodies

of the non-starch lipids were broken and fat was released from the surface of the kernel where the lipids were diffused outwards, and hence the bran of parboiled rice was more oily. However, when the steaming time was increased, the fat content was increased which might be due to the severe temperature effect following the result of Hasbullah et al. (2016) and Fonseca et al. (2014).

For texture properties, the hardness and toughness of cooked parboiled rice was investigated. This study found that only the hardness increased with the decreasing of amylose content of parboiled rice. This might have been caused by the gelatinization and recrystallization process of starch polymorphisms during parboiled rice production (Oli et al. 2014). The starch polymorphisms occur during the recrystallization process which affects the textural properties of cooked parboiled rice (Oli et al. 2014).

4.4.1 The relationship between the chemical and textural properties

The correlation coefficient between the hardness and toughness of cooked parboiled rice was 0.69. As shown in Table 4.5, there was a significant negative correlation between hardness and amylose content of the rice (r=-0.52), identified using the Pearson correlation test. Negative correlation was also observed between toughness and amylose content as well (r=-0.38). These results indicate that hardness and toughness of cooked parboiled rice increases with decreasing amylose content. This is in contrast to cooked non-parboiled rice, for which the higher the amylose content, the harder the cooked rice (Lu et al. 2013). The hardness of parboiled rice in this study increased, when it should be decreased according to the amylose content that was lower. Oli et al. (2014) reported that the final texture of cooked parboiled rice was mainly changed by the gelatinization and recrystallisation of starch during parboiling. It might be due to the higher crystallization after the gelatinization in the steaming process where the higher crystallization allowed water to disperse less into rice grains, which caused the hardness to increase, even though the amylose content reduced.

The fat content result in this study indicated a significant positive correlation with hardness (r=0.20) and toughness (r=0.12) using the Spearman correlation test (Table 4.5). The hydrophobic nature of fat causes reduced water absorption in brown rice (full-fat bran) (Oli et al. 2014). Therefore, the fat contents might affect the hardness

properties (Oli et al. 2014). In the same way, Li et al. (2016) (Li et al. 2016)reported that the lipids of rice composition correlate with hardness and stickiness, with a low correlation coefficient (r=0.364 and 0.231, respectively) using Spearman correlation tests. The texture properties were also likely defined by other factors such as the crystallinity structure of cooked parboiled rice (Oli et al. 2014). The X-ray diffractometer test could inspect the degree of crystallinity of rice starch (Saniso et al. 2020). X-ray patterns was depending on parboiled rice process condition (Ong and Blanshard. 1995; Saniso et al. 2020). The degree of crystallites depended on the parboiled rice process conditions and amylose content (Lamberts et al. 2009). The starch gelatinization temperature and gelatinization enthalpy were defined by a differential scanning calorimetry test which leads to the degree of gelatinization of starch (Du et al. 2019). Moreover, Islam et al. (2002) reported that the hardness of parboiled rice correlates with the degree of starch gelatinization.

Table 4.5 Relationship between the physical, chemical and texture properties of parboiled rice.

Properties	Correlation coefficient					
	Pearson		Spearman		Kendall's tau	
	Hardness (N)	Toughness (Nmm)	Hardness (N)	Toughness (Nmm)	Hardness (N)	Toughness (Nmm)
Chemical properties						
Amylose content (%)	-0.52**	-0.38	-0.49*	-0.37	-0.35*	-0.24
Fat content (%)	0.18	0.04	0.20	0.12	0.11	0.07

**Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

4.5 Conclusion

The different variables of the parboiled rice process (duration and temperature for soaking and steaming) influenced the physical (whiteness and yellowness) properties of parboiled rice of the Suphanburi 1 variety. For the chemical properties, the amylose and fat content were significantly correlated with the textural properties

of cooked parboiled rice, that is, hardness and toughness. The findings are for the consideration of the parboiled rice industry, allowing producers to match the appropriate conditions of the process to the specification requirement of customers. However, it is important to conduct further analysis with the X-ray diffractometer and differential scanning calorimetry test on the non-parboiled rice and parboiled rice of the Suphanburi 1 variety to gain understanding of gelatinization enthalpy and the amorphization of starch of parboiled rice grain, which might affect the texture properties of cooked parboiled rice.

Chapter 5

Texture evaluation of cooked parboiled rice using nondestructive milled whole grain near infrared spectroscopy

As well as cooked parboiled rice ISO texture evaluation functioning as a standard in indicating important textural properties of cooked parboiled rice, one consumer acceptability criterion is hardness and toughness. Therefore, this chapter illustrates the texture evaluation of cooked parboiled rice using nondestructive milled whole grain NIR spectroscopy. The hardness and toughness calibration models were constructed in this chapter. The FT-spectrometer which was tested for precision in chapter 3 was used. Also, the ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking method was used as reference test.

* This chapter constituted the publication article: Onmankhong, J., and Sirisomboon, P. 2021. “Texture evaluation of cooked parboiled rice using nondestructive milled whole grain near infrared spectroscopy.” **Journal of Cereal Science**. 97: 103151.

5.1 Introduction

Presently, parboiled rice product is widespread in many countries in Europe, North and South America, with parboiled rice product originating in Asia (Bello et al. 2006). Rice product is exported from Thailand at more than 10 billion to 200 billion tons every year, making Thailand the leader in the world export market (Siriphollakul et al. 2017). One important rice product from Thailand is parboiled rice. The advantage of the parboiled rice process is that the head rice yield is more than that of the unparboiled rice, because the parboiling process decreases breakage of the kernel during the dehulling and milling process (Leethanapanich et al. 2016). The parboiling process is comprised of soaking, steaming and drying, where soaking and steaming affect the texture properties (Villanova et al. 2017). The hardness and toughness of cooked parboiled rice effects the consumer acceptability, and texture properties of cooked rice change during parboiling (Miao et al. 2016). The changing of physical and chemical properties of the rice occurred at different conditions during parboiling (Patindol et al. 2008). Therefore, the texture of cooked parboiled rice is considered an important factor for consumers.

The typical instrument that has been used to measure the texture of cooked rice samples was the texture analyser (Miao et al. 2016). It used texture profile analysis (TPA) test to determine the average values of the hardness, adhesiveness, cohesiveness and springiness (Miao et al. 2016) However, the above method is not a rapid way of checking texture properties. Therefore, the development of a model to predict texture properties of cooked parboiled rice from milled parboiled rice is important for the parboiled rice industry. The near infrared (NIR) spectroscopy technique can evaluate the texture properties of cooked normal rice (Siriphollakul et al. 2017). Nondestructive technique, quick results and acceptable accuracy were all provided by the NIR spectroscopy technique. Besides, the NIR spectroscopy technique has been used for high precision and accuracy in assessing the amylose content (Siriphollakul et al. 2017) and amino acid content (Wu et al. 2002) which effect the texture properties of cooked rice (Derycke. 2007), (Siriphollakul et al. 2017). Therefore, if the NIR spectroscopy technique can possibly predict the texture properties of cooked

parboiled rice, it will benefit the parboiled rice producing factories in checking the texture properties before exportation. The factories do not need to check the textural properties from cooked parboiled rice with processes that are complicated using time and labor. The objective of this research was to develop an NIR spectroscopy model to predict the hardness and toughness values of cooked parboiled rice from milled whole grain parboiled rice NIR spectra. In this study, the hardness and toughness of the cooked rice were the parameters of the texture standard of parboiled rice, defined by ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after the Cooking method (ISO. 2012).

5.2 Materials and methods

5.2.1 Parboiled rice samples

The 165 parboiled rice samples were obtained from Capital Chainat Rice Mill company limited, in Sapphaya District, Chainat Province, Thailand. The samples were taken from the conveyor belt position (90 min/sample) before sending samples to the storage tank for exportation. These samples had a moisture content of about 14%, which was determined in the factory by using a moisture meter (F-999s, Kett, Japan). All of the 165 samples were manufactured using rice with high amylose content in central Thailand. Generally, parboiled rice production in Thailand use high amylose rice varieties to produce the parboiled rice because the texture of the hard parboiled rice is preferred by customers. (direct communication with the factory). The rice variety used in this study including RD41, RD47, Chai Nat 1, Suphan Buri 90, Phitsanulok 2 and so on. The samples were obtained from the different conditions of parboiled rice processes, which were adjusted in the factory during process depending on the age of rice and quality required by customers.

5.2.2 Near infrared scanning of the parboiled milled rice

The 150 g of parboiled milled rice sample was put in a quartz cup with a diameter and height of 8.70 and 8.75 cm, respectively. The samples were scanned by a Fourier transform (FT-NIR) spectrometer (MPA, Bruker Ltd., Germany) at a wavenumber of 12,500–4,000 cm^{-1} (800-2,500 nm). The sample in a quartz cup was

scanned at the bottom position through quartz material in diffuse reflectance mode with a resolution of 16 cm^{-1}. Absorbance spectra were recorded during the rotation of the quartz cup. The sixty-four scans of the sample were averaged to one spectrum. The room temperature (25 ± 1°C) was controlled during scanning.

5.2.3 Preparation of cooked parboiled rice

Following the ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking method (ISO, 2012), the parboiled rice sample was cooked. The ratio of rice to water in this experiment was 4:7.6 by weight by following the ISO standard. 20 g of milled parboiled rice was mixed with distilled 38 ml of water in a beaker, then stirred with a glass rod and covered using watch-glasses. The water bath (W350, Memmert, Germany) was used for cooking rice. The test tube rack was placed in the water bath to support the beakers to be higher than the water surface. The sample in beaker was cooked via steaming for 20 min. The heat source was removed after 20 min and the sample beaker was allowed to stay in the water bath for 10 min. The beakers were removed from their cooking containers and then placed upside down on the watch-glasses at a temperature of between 20 °C and 25 °C for 1 h. The above describes the preparation method of cooked parboiled rice used in this experiment.

5.2.4 Texture analysis

An ISO 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after Cooking method, (ISO, 2012) using texture analyser (TA.HD. plus, Stable Micro Systems, UK) with a load cell capacity of 100 kg was used for measuring the parboiled cooked rice textural properties. The principle factor of an ISO method is the measurement of force needed to extrude cooked rice through a perforated plate. The pretest, test and post test speeds were 1, 1.67 and 10 mm/s. Two replications were performed in this experiment. The mini Ottawa cell (HDP-MK05 model, Stable Micro systems, Surrey, UK) extrusion method was used in this study. The cooked parboiled rice was compressed with increasing force until a distance of 15 mm was reached, causing the cooked rice to fall through the holes. The hardness (N) and toughness (Nmm) values were used for textural properties in this study. These values were calculated using Exponent Stable

Micro Systems software, version 6,1,13, 0 (Stable Micro Systems, UK) where the maximum force and area under the graph illustrated the hardness and toughness values, respectively. The diagram of texture analysis for hardness and toughness values in cooked parboiled rice are showed in Figure 5.1 The repeatability (Rep) was used to indicate the precision of the reference test for hardness and toughness. The Rep of reference data is the standard deviation of the different between duplicates. The Rep was used to calculate the maximum coefficient of determination ($\mathrm{R}^2_{\mathrm{max}}$) by formula (5.1) (Dardenne, 2010):

$$\mathrm{R}^2_{\mathrm{max}} = \frac{\mathrm{SD}_{\mathrm{y}}^2 - \mathrm{Rep}^2}{\mathrm{SD}_{\mathrm{y}}^2} \qquad (5.1)$$

where SD_{y} is the standard deviation of the calibration data set. If there is no error in the spectra or model, the coefficient of determination (R^2) will be approximately equal to $\mathrm{R}^2_{\mathrm{max}}$. SD_{y} and Rep can indicate whether a reference laboratory analysis is precise or not (Dardenne, 2010).

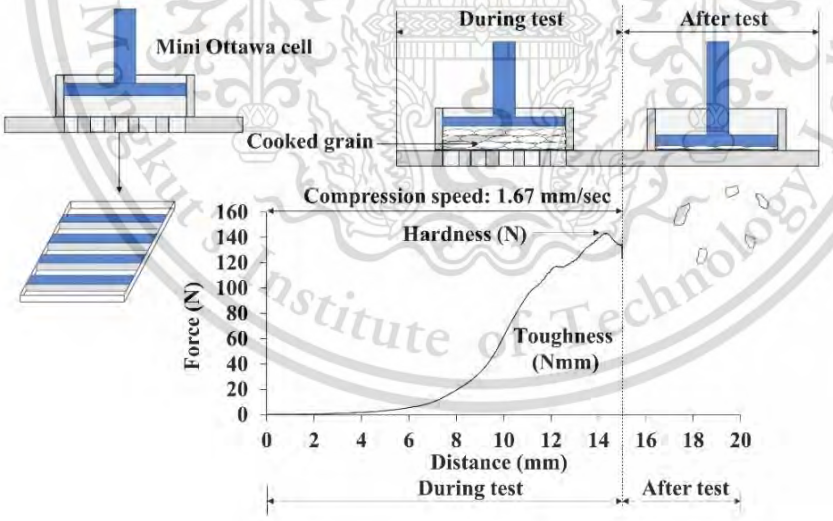

Figure. 5.1 The compression test and diagram of texture analysis for hardness and toughness of cooked parboiled rice.

5.2.5 Near infrared spectroscopy analysis

Partial least squares regression (PLSR), Principal component regression (PCR) and Support vector machine regression (SVM) using Unscrambler X10.3 (Camo, Norway) were used to develop the modeling from spectral and reference data (hardness and toughness value). Before modeling, the raw spectral data from FT-NIR, had been performed for the pre-processing including moving average, Gaussian filter, median filter and Savitzky-Golay smoothing methods (segment size 3, 5, 7, 9, 11, 13, 17 and 21), normalization (area normalization, unit vector normalization, mean normalization, maximum normalization and range normalization), baseline (baseline offset and linear baseline correction), standard normal variate, multiplicative scatter correction and derivatives (first and second derivative (gap size 3, 5, 7, 9, 11, 13, 17 and 21)) was used for developing the model. The samples were divided into two sets: 80% for the calibration set (132 samples) and the remaining 20% for the validation set (33 samples). The optimized models were selected based on model performance parameters including the coefficient of determination (calibration set; R^2, validation set; r^2), root mean square error of prediction (RMSEP), bias, and ratio of prediction to deviation (RPD).

The relative standard deviation of prediction (RSEP) calculated as follows:

$$\mathrm{RSEP} = \sqrt{\sum\left[\frac{(Y_i - Y_{pre})^2}{\sum Y_i^2}\right]} \times 100 \qquad (5.2)$$

where Y_i, and Y_{pre} are measured value and predicted value, respectively.

5.3 Results and discussion

5.3.1 Texture properties of cooked parboiled rice

In this study, the constant ratio of rice to water was used because the water content affects the texture properties of cooked rice. The statistical information of the hardness and toughness of parboiled cooked rice are presented in Table 5.1. However, in hardness and toughness analysis by texture analyzer, compression distance was set 15 mm which caused error in measured values which maximum force indicating hardness might not reached before 15 mm compression distance but the value at 15 mm was used for hardness and the

corresponding area under curve was calculated as toughness. 28 samples (21.2%) of the calibration set and 3 samples (12.12%) of the validation set in this study were indicating problem samples, which reduced the accuracy of the texture measurement. Figure 5.2 shows that hardness had a positive correlation with toughness of cooked parboiled rice. The hardness increased following the increased toughness with correlation coefficient (r) of 0.82. Ruptures occurred in the grain because of granule expansion, making the hardness value decrease. The repeatability for hardness and toughness was 6.92 N and 43.52 Nmm, and the R^2_{max} of the hardness and toughness values were 0.75 and 0.64 respectively, showing the high error of the measurement methods of texture properties of cooked parboiled rice which was 25% and 36%.

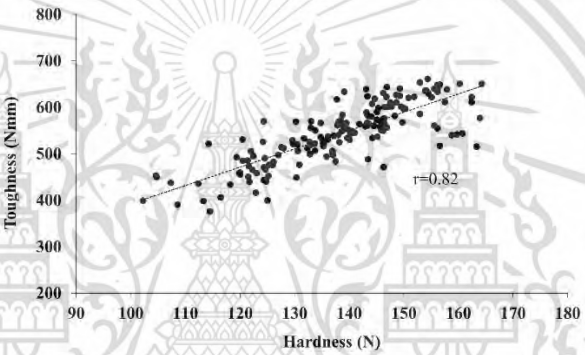

Figure. 5.2 The relationship between hardness and toughness value of cooked parboiled rice.

5.3.2 Near infrared spectroscopy models of hardness and toughness for cooked parboiled rice

Table 5.2 shows the optimum results each the preprocessing for the PLSR and PCR models for hardness and toughness of cooked parboiled rice. The total results of all models showed that the PLSR and PCR models performed better than the SVM models of hardness and toughness at the full wavelength range of 12,500–4,000 cm^{-1} (800-2,500 nm). The all r^2 of SVM models for hardness was lower than 0.49. Williams (2019) suggested that a regression model with r^2 range from up to 0.25 and 0.26–0.49 is not usable in NIR spectroscopy calibration and poor correlation where the reason should be researched, respectively. For SVM toughness models, the r^2 was lower than 0.58. The r^2 ranged from

0.50 to 0.64 indicated that the model can be acceptable for screening and some other approximate calibration (Williams. 2019).

Moving average (segment size 11) pre-processing was used the PLSR and PCR for optimum hardness models with a factor of 8 and 9, respectively. It provided an r^2 of 0.70 for both PLSR and PCR. Moreover, the RMSEP of PLSR and PCR of the models were close (7.24 N and 7.27 N, respectively). However, the bias of the PLSR model was more than that of PCR model. The pretreatment methods to reduce the noise in the spectrum, which contained small spiny peaks, used were moving average (segment size 11) smoothing for the hardness model. Moving average preprocessing consider that the random noise can be averaged out (or, at the very least, its magnitude reduced significantly) by averaging a number of consecutive signal points (Andrsde-Garda. 2009). The optimal moving average depends on the noise distribution and signal width. It is best number of different filter widths to find the optimum (Brereton. 2018). Smoothing pre-processing was used to reduce the noise on NIR spectra. However, too much smoothing, the informative signal was also reduced in intensity and resolution and too little smoothing and noise remains (Andrsde-Garda. 2009).

For toughness model, the PLSR and PCR optimum models provided r^2 0.65 and 0.66 with factors of 3. Williams (2019) suggested that the model with the r^2 range from 0.66 to 0.81 can be used for rough screening and some other approximate calibrations. Moreover, the RMSEP of PCR was lower than PLSR. The range normalization and mean normalization was used for the best models for PLSR and PCR, respectively. Normalization is the transformations where its purpose is to "scale" samples in order to get all data on approximately the same scale. It is applied in cases where the data are collected with a method (system) where the detector signal is a function of sample mass or of source power instead of sample concentration (Unscrambler X10.3, Camo, Norway). Mean normalization is the most classical case of normalization, which consists in dividing each row (each observation) of a data matrix by its average, to obtain the area under the curve becomes the same for all samples, thus neutralizing the influence of the hidden factor (Unscrambler X10.3, Camo, Norway). For range normalization, each row is divided by its range, i.e. "max value – min value". The property of range-normalized sample is the curve span becomes 1 (Unscrambler X10.3, Camo, Norway).

Table 5.1 The statistical data of hardness and toughness for cooked parboiled rice.

Textural properties	Calibration set					Validation set				
	Number of samples	Minimum	Maximum	Mean	Standard deviation	Number of samples	Minimum	Maximum	Mean	Standard deviation
Hardness (N)	132	102.24	164.25	138.60	13.54	33	107.30	162.38	137.68	13.98
Toughness (Nmm)	132	375.72	661.17	543.10	65.48	33	415.98	648.55	546.01	66.34

To obtain the optimal hardness model, it was required only smoothing pretreatment which indicated that the effect of noise along the spectrum was predominant for hardness prediction. In case of the optimal model for toughness, the normalization pretreatment was recommended which indicated that the sample mass or source power instead of sample concentration affected the toughness prediction.

The optimal models of hardness and toughness provided r^2 of 0.70 and 0.66, respectively. This indicated that 70% and 66% of the total variation could be explained by NIR spectra, which 30% (100-70) and 34% (100-66) of the total variation is unexplained variance that could not be explained by NIR spectra. The R^2_{max} were 0.75 (75%) and 0.64 (64%) for the hardness and toughness, respectively, showing that the errors of reference test were 25% (100-75) and 36% (100-64). Therefore, an error value from NIR spectra scanning and other error was 5% (100-70-25) and 1% (100-64-35) for hardness and toughness models, respectively. Moreover, because error in texture analysis indicating, in the methodology of texture properties of cooked parboiled rice, the percentage of the error can be calculated only if model developed without these error samples were compared. The hardness PLSR model and toughness PCR model showed low RMSEP with 7.24 N and 38.00 Nmm, respectively, which were lower than other models in this study. The RSEP which was the relative ratio between absolute error of prediction and reference value, showed that every predicted value could give the error of reference value approximately 0.91% and 1.21% of hardness and toughness, respectively. Chang et al. (2001) set a threshold of RPD, for predictive models for non-NIR absorber, especially some constituents, to indicate the performance of a model: “excellent models (RPD is greater than 2), fair models (RPD ranges from 1.4 to 2), and non-reliable models (RPD <1.4.”. This study was for the development of predictive model of grain texture that were not more or less associated with any classical absorber. Therefore, the RPD threshold was selected for explanation of the model performance. The RPD was 1.93 and 1.75 for the best PLSR hardness and PCR toughness model, respectively, which could be interpreted that the models were fair for prediction application.

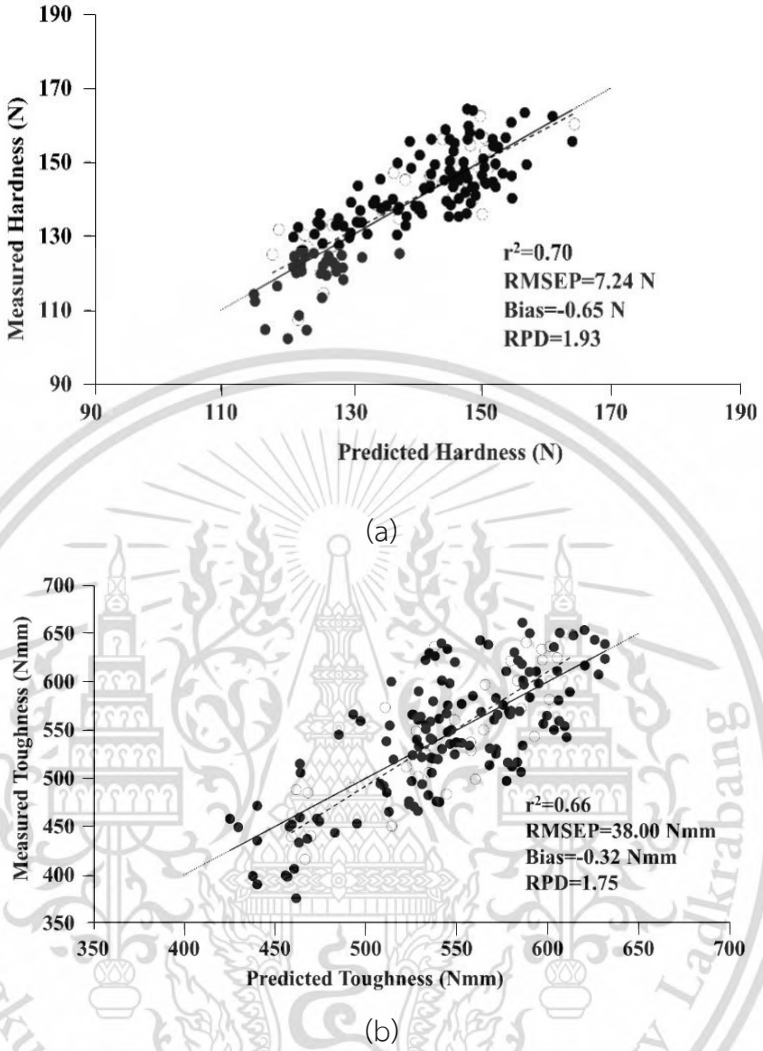

Figure. 5.3 Scatter plot represent (a) The optimal model of hardness using PLSR (b) The optimal model of toughness using PCR.Trend line of calibration Trend line of validation Target line.

Table 5.2 The results of PLSR and PCR models for prediction of hardness and toughness of cooked parboiled rice.

Textural properties	Pre-processing	Factor	Calibration set		Prediction set				
			R^2	RMSEC	r^2	RMSEP	Bias	RSEP	RPD
PLSR modeling									
Hardness (N)	Moving Average (Segment size 11)	8	0.72	7.20	0.70	7.24	-0.65	0.91	1.93
	Mean Normalization	8	0.78	6.40	0.64	7.93	-1.1	1.00	1.76
	Baseline offset	7	0.74	6.89	0.65	7.87	-1.1	0.99	1.78
	SNV	9	0.85	5.31	0.55	8.89	-1.36	1.12	1.57
	MSC	9	0.85	5.33	0.55	8.89	-1.34	1.12	1.57
	First Derivative (Gap size 13)	7	0.83	5.57	0.70	7.23	-1.39	0.91	1.93
Toughness (N)	Moving Average (Segment size 21)	4	0.58	42.42	0.65	38.97	-0.74	1.24	1.70
	Range normalization	3	0.58	42.51	0.65	38.88	-0.99	1.24	1.71
	Linear Baseline correction	3	0.62	39.95	0.54	44.41	-4.02	1.42	1.49
	SNV	3	0.57	47.75	0.48	47.10	1.54	1.50	1.41
	MSC	3	0.47	47.76	0.48	47.70	1.55	1.50	1.39
	First Derivative (Gap size 11)	6	0.71	35.4	0.58	42.21	-7.94	1.35	1.57

Table 5.2 The results of PLSR and PCR models for prediction of hardness and toughness of cooked parboiled rice (continued).

Textural properties	Pre-processing	Factor	Calibration set		Prediction set				
			R^2	RMSEC	r^2	RMSEP	Bias	RSEP	RPD
PCR modeling									
Hardness (N)	Moving Average (Segment size 11)	9	0.66	7.92	0.70	7.27	-0.09	0.91	1.92
	Area normalization	5	0.50	9.62	0.39	10.35	-0.13	1.30	1.35
	Linear Baseline correction	4	0.49	9.64	0.59	8.54	-0.22	1.07	1.64
	SNV	3	0.49	9.68	0.41	10.25	0.15	1.29	1.36
	MSC	3	0.49	9.68	0.41	10.25	0.15	1.29	1.36
	First Derivative (Gap size 11)	8	0.6	8.53	0.61	8.62	0.68	1.09	1.62
Toughness (N)	Median Filter (Segment size 21)	4	0.57	42.86	0.64	38.98	-0.88	1.34	1.70
	Mean normalization	3	0.56	43.17	0.66	38.00	-0.32	1.21	1.75
	Baseline offset	3	0.48	47.02	0.57	42.98	5.24	1.37	1.54
	SNV	4	0.46	47.74	0.46	47.89	2.31	1.53	1.39
	MSC	4	0.46	47.74	0.46	47.89	2.31	1.53	1.39
	Second Derivative (Gap size 21)	10	0.55	43.71	0.60	41.24	-5.21	1.31	1.61

5.3.3 Important wavelength for prediction of hardness and toughness cooked parboiled rice

Figure. 5.4 a and b show the regression coefficient plots for PLSR and PCR of the optimum hardness and toughness models. The high regression coefficient of each wavenumber affects the prediction model because of the particular bond vibration (Posom et al. 2016). There are high regression peaks of the PLSR hardness model where the corresponded chemical bond vibration are indicated by Osborne et al. (1993) as follows: At approximately 10144.32 cm^{-1} (985.77 nm), 5307.45 cm^{-1} (1884.14 nm) and 4813.73 cm^{-1} (2077.39 nm) vibrations were the vibration of starch. The CH bond vibration was approximately at 8169.46 cm^{-1} (1224.07 nm) (C–H str second overtone). This was the highest peak of the regression coefficient plot. The CH_3 bond vibration were approximately at 8678.60 cm^{-1} (1152.26 nm), 7575.46 cm^{-1} (1320.05 nm), 7290.03 cm^{-1} (1371.74 nm) and 5839.74 cm^{-1} (1712.41 nm), which were C–H str second overtone, 2 × C–H str + C–H def, 2 × C–H str + C–H def and C–H str first overtone vibration, respectively. The $CONH_2$ and ROH bond vibration were approximately 6973.74 cm^{-1} (1433.95 nm) and 6580.31 cm^{-1} (1519.69 nm), which were N–H str first overtone vibration. The approximately at 5538.88 cm^{-1} (1805.42 nm) was the vibration of cellulose. The bond vibration at 7128.03 cm^{-1} (1402.91 nm) was of the ROH bond which was O–H str first overtone.

For the PCR toughness model, according to Osborne et al. (1993): the starch vibration was approximately at 6881.17 cm^{-1} (1453.24 nm) and 5353.74 cm^{-1} (1867.85 nm). Williams (2009) indicated that the bands at the 1430 nm area in low moisture maize, flour, cellulose and starch are possibly associated with the proliferation of O–H groupings of the glucose molecule, which are present in much higher concentration than the O–H of water. Therefore, the vibration approximately at 6881.17 cm^{-1} (1453.24 nm) was starch vibration. The C-H str second overtone vibration was approximately at 8331.46 cm^{-1} (1200.27 nm). The CONH bond vibration was approximately at 5199.45 cm^{-1} (1923.28 nm) which was C=O str second overtone. The bond vibration at 5970.88 cm^{-1} (1674.80 nm) may be the vibration of aromatic. The CH_2 bond vibration at approximately 7174.31 cm^{-1}

(1393.86 nm) was 2×C-H str+C-H def which was the highest peak of the regression coefficient plot.

Thus, the starch in parboiled milled rice affected the hardness and toughness of cooked parboiled milled rice. A major factor effecting cooked rice texture was amylose content, which is a major part of the composition of starch granules (amylose and amylopectin) (Bertoft. 2017). Cooked rice texture with low amylose content becomes sticky and moist whereas that with high amylose content is flaky and dry (Siriphollakul et al. 2017). The amylose content was decreased owing to the hot-water soluble in the soaking process (Patindol et al. 2008) and with the severity of parboiling process (Biswas and Juliano. 1988). The parboiled rice samples in this study were taken from an export parboiled rice factory, which produced at different conditions according to customer requirements. The samples, therefore, contained different amounts of amylose content indicating the different texture characteristics. Similarly, the regression coefficients plot in this study shows that the peak of starch affects the hardness and toughness prediction models.

Besides, another main factor related to texture properties is the structure of starch (Li and Gilbert. 2018).

In this study, the hardness and toughness parameters of parboiled rice was of interest. The texture properties were controlled by various factors. One of the main factors related to texture properties is the structure of starch (Li and Gilbert. 2018). Li et al. (2016) reported that textural properties of cooked rice was affected by the starch fine structural features, particularly the fine structure of amylase. During the parboiled rice process, the recrystallization (retrogradation) of the starch coupled with some lipid-amylose complexes had occurred during gelatinization, which changes the texture properties of parboiled cooked rice (Oli et al. 2014). The hardness of cooked rice was increased in relation to starch retrogradation, because starch was converted to a more stable crystalline form (Mutters and Thompson. 2009). The amylose content was important with the texture of cooked parboiled rice. Moreover, the lipid-amylose complexes can hinder the gelatinization and swelling of starch granules (Li and Gilbert. 2018). The higher ratio and whole molecular size of amylose branches ranging from

1000 to 2000 degrees of polymerization significantly influenced the hardness of cooked rice, increasing the hardness (Li et al. 2016; Li and Gilbert. 2018). Amylose branches ranging from 1000 to 2000 degree of polymerization were found in high amylose rice varieties (Li and Gilbert. 2018). The strong physical structure made it difficult to disrupt the grain during cooking due to a high amylose content (Lu et al. 2013).

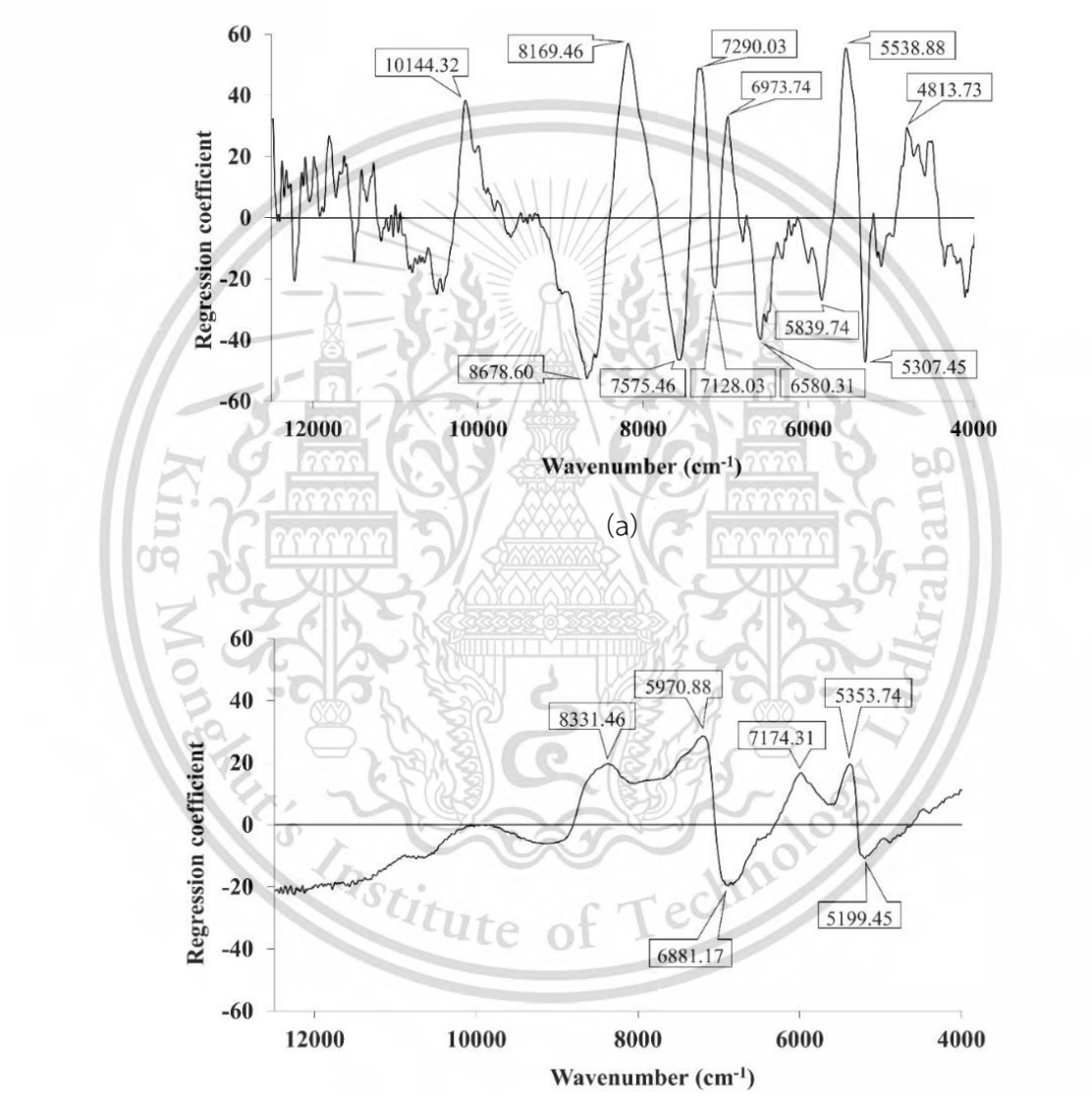

Figure. 5.4 The regression coefficients for the optimum model (a) The optimal model of hardness using PLSR (b) The optimal model of toughness using PCR.

In this study, the regression coefficient plot also identified the influence of $CONH_2$ and CONH bond vibration on the hardness and toughness models. The $CONH_2$ and CONH were the functional group of amino acid. This indicated the effects of the amino acid in parboiled milled rice on the hardness and toughness of cooked parboiled milled rice. The amino acid is in the protein composition (Hormoz. 2013). The bran area on the surface of rice grains contain high protein, and is a barrier against the spread of water towards the starch granule inside the grain during gelatinization process (Champagne et al. 2004; Oli et al. 2014). Furthermore, the swelling of starch granules during the gelatinization process was controlled by protein, owing to glass transition of protein was lower than of starch (Matveev et al. 2000). The protein in rice was a major attribute of the water absorption even though the concentration of protein was lower than that of starch (Derycke et al. 2005). As mentioned above, this is the reason why amino acid is related to the texture properties of cooked rice, because the protein controls the texture properties. In addition, the protein is hydrolysed, leading to increased disulphide bonds during parboiling that effects the hardness of cooked parboiled rice (Derycke. 2007). The different conditions of the parboiled rice process have influenced the protein, for example the increasing of steaming time decreases the protein of parboiled rice (Ibukun. 2008), which causes differentiation of the texture properties.

Absorptions at different areas of the wavelength range between 700 and 2500 nm (NIR range) are all overtones or combination bands which can be assigned to these individual functional groups such as C–H, O–H, N–H (Williams. 2019). The texture properties of parboiled cooked rice were controlled by several molecules of composition, including amylose, amylopectin, compactness of starch granules, protein and lipid, which are provided in chemical bonds as above (Awanthi et al. 2019). That is one possible cause as to why the NIR technique can detect the texture properties. Hu et al. (2018) reported that the O–H and C–H bond vibration effected the textural properties of cooked cereals, which is possibly caused by amylose, protein and water compounds.

The NIR spectrum was comprised of information about the chemical composition and also physical properties such as particle size or crystallinity of the sample (Blanco et al. 1999). Moreover, Awanthi et al. (2019) reported that the NIR technique can be used for the prediction of the quantifying hardness of paddy at a wavelength of 588–1091 nm. It was found that the highest r^2 and lowest standard error of prediction (SEP) were 0.936 and 1.711, respectively. Furthermore, Hu et al. (2018) found that PLS model was the optimal model when compared with the PCR model in all cereal texture properties, including hardness (r^2 = 0.96), springiness (r^2 = 0.87), chewiness (r^2 = 0.92), gumminess (r^2 = 0.90), cohesiveness (r^2 = 0.82) and stickiness (r^2 = 0.63) at wavelength of 1,000–2,500 nm.

5.4 Conclusions

In this research, the spectra were obtained by scanning milled parboiled rice to predict the hardness and toughness of cooked parboiled rice. The optimum models for hardness was developed by PLSR while for the toughness was from PCR. According to RPD threshold, the models obtained were fair for prediction application. The range of sample hardness and toughness of cooked parboiled rice was narrow advised by the low value of RPD (1.93 and 1.75, respectively). Therefore, the more variation of milled whole grain parboiled rice samples i.e. more number of samples from more varieties, field locations, seasons etc. of rice, is necessary for improvement of the model performance. It was observed that the repeatability of reference test (hardness and toughness) methodology was high (6.92 N and 43.52 Nmm, respectively), indicating the high error of the methodology. We investigated the source of error and found that it was from the cooked rice sample preparation before texture measurement where the moisture content of the cooked rice samples among replicates were different due to the different waiting time of the samples prior to the weighing process and texture analysis. The feasibility of the texture evaluation of cooked parboiled rice using nondestructive milled whole grain near infrared spectroscopy in this study suggested that the NIR protocol developed was applicable for real use due to the error of the NIR scanning and other unexplained errors was only 6% and 3% for hardness and toughness models, respectively.

Chapter 6

Cognitive spectroscopy for the classification of rice varieties: a comparison of machine learning and deep learning approaches in analysing long-wave near-infrared hyperspectral images of brown and milled samples

This chapter created the model classification of Thai Jasmine rice (Khao Dawk Mali 105) from Pathum Thani 1 and Phitsanulok 2 varieties using near-infrared hyperspectral imaging (NIR-HSI) combined with machine learning methods, supported by vector machines (SVM) and deep learning methods including convolutional neural networks (CNN). Both brown and milled rice were used for creating the model classification. Also, RGB imaging was used for creating the classification model to compare the classification accuracy with the NIR-HSI information. The NIR-HSI information was divided into two parts: 1) NIR spectra; 2) HSI image.

* This chapter is under preparation to submit to Food Control: Onmankhong, J., Ma, T., Inagaki, T., Sirisomboon, P. and Tsuchikawa S., 2021. “Cognitive spectroscopy for the classification of rice varieties: a comparison of machine learning and deep learning approaches in analysing long-wave near-infrared hyperspectral images of brown and milled samples”

6.1 Introduction

Rice is a staple food for almost half of the world's population (Yu et al., 2020). Thailand is the primary producer and exporter of premium-grade Khao Dawk Mali 105 (KDML105) rice with a high price in the global market (Boling et al., 2011). KDML105 is popular because it has a long grain, unique aroma, mild taste, pure white colouring, soft texture, and good flavour after cooking (Siriphollakul et al., 2017; Kukusamude and Kongsri, 2018; Qingyu et al., 2020). However, the planting season of KDML105 is only during the rainy season (June to November) in the northeastern and northern regions of Thailand (Boling et al., 2011; Korinsak et al., 2016). Types of brown and milled KDML105 rice are Thailand export products (Department of Foreign Trade Ministry of Commerce, 2016). Unfortunately, KDML105 rice on the global market has been found to be adulterated with other cheaper species (Timsorn et al. 2017), which influences consumer benefits, as well as the interests of producers. The common rice species mixed with KDML105 include Pathum Thani1 (PTT1) and Phitsanulok 2 (PSL2) (Attaviroj & Noomhorm, 2014). Such rice mixing can lead to poor cooking and eating qualities due to the amylose content of the variety (Chamarerk et al. 2016). The contamination problem will be more serious if a low-amylose-content variety is mixed with a high-amylose-content variety (Chamarerk et al., 2016), and the KDML105 variety has low amylose content (Timsorn et al. 2017). Recently, the Thailand Ministry of Commerce has enacted stipulations on the Thai Jasmine rice standard in an attempt to prevent such food fraud and has indicated that pure KDML105 must not comprise less than 92% (Department of foreign trade Ministry of Commerce, 2016).

KDML105 is generally slightly longer but smaller than the other species (Division of Rice Research and Development, 2016), but most of the physical characteristics of the rice samples will be destroyed after milling. It is challenging to distinguish them visually, even just for a small amount. However, because PSL2 generally has higher amylose and protein contents, which have been shown to be approximately 28% and 9.81%, respectively, followed by PTT1 (15-19% and 9.60%) and KDML105 (12-17% and 9.23%) (Division of Rice Research and Development, 2016), amylose content analysis could be a useful method. Other standard methods include (1) a test method for Alkali

spreading value; (2) a test method for checking cooked rice kernels boiled in water; and (3) a test method using iodine staining, which was identified by the Ministry of Commerce, Thailand (Department of Foreign Trade Ministry of Commerce, 2016). In addition, advanced studies were carried out by researchers to detect the authenticity of rice, such as the deoxyribonucleic acid (DNA)-based method (Ganopoulos et al., 2011), denaturing high-performance liquid chromatography (Peng et al., 2016), and a method using textural and pasting properties (Pitiphunpong et al., 2011). However, these methods are all time-consuming, destructive, complex, laborious, and sometimes also toxic. In the large-scale rice circulation link, the development of rapid and nondestructive methods is highly desired.

Near-infrared (NIR) spectroscopy (800-2500 nm) is a proven technique for quality analysis of agricultural and food products that utilizes the stretching and bending vibrations among the hydrogen bonds of the products, such as C-H, O-H, and N-H. A previous study indicated that starch and protein were important for the classification of various types of rice, including KDML105, PTT1, PSL2, Suphan Buri60, and Chainat 1, using Fourier transform NIR (FT-NIR) spectroscopy (Attaviroj et al., 2011). However, the NIR spectral data were collected from several kernels (approximately 70 ± 2 g). It is highly desirable to achieve a single-kernel measurement, especially for the simultaneous detection of multiple targets for such topics. This is difficult to achieve by conventional NIR spectrometers, not only because of the small size of the rice but also because of the uneven rice surface. NIR hyperspectral imaging (HSI) is an advanced technique that provides an NIR spectral image at each wavelength with a high spatial resolution. The HSI method can rapidly and nondestructively provide the spatial distribution of molecular vibration information in scanned samples (Ma et al., 2020). Recent studies have shown its possibilities for single-kernel measurement (Sun et al., 2015a; Mohan et al., 2016; Huang et al., 2016; Feng et al., 2019). Conventionally, the averaged reflectance spectrum would be extracted from each sample part on the scanned HSI images to create calibration models by machine learning methods (Weng et al., 2020). For example, Qiu et al. (2018) identified rice seed varieties, including Xiushui 134, Zhejing 99, indica rice, Zhongjiazao 17 and Zhongzao 39, and the 1D CNN

model with wavelengths of 975-1646 nm performed better than the corresponding K-nearest neighbour (KNN) and SVM models. The classification accuracies of the training and test sets were 89.6% and 87.0%, respectively (Qiu et al. 2018). Kong et al. 2013 used an NIR-HSI in the wavelength range of 1039 nm to 1612 nm to classify rice seed varieties, including Zhongzheyou No. 1, Zhongzheyou No. 5, Zhongzheyou No. 8 and Zhongzheyou No. 86. Soft independent modelling of class analogy (SIMCA), SVM, and random forest (RF) models obtained 100% classification accuracy in both the calibration and prediction sets (Kong et al. 2013). In other studies (Guerrero et al., 2012; Sun et al., 2013; Sun et al., 2015), support vector machines (SVMs) were also extensively used for classification model construction. However, information on the spatial distribution (i.e., 2D spectral image) of chemical constituents and physical structures would be lost after averaging the reflectance spectra of each sample. In addition, the non-uniform quality attributes within a seed could also cause unstable prediction accuracy (Manley, 2014). There is still room to increase the reliability of HSI classification results by simultaneously analysing the characteristics of both spectral and spatial features (Borzov & Potaturkin, 2018; Ma et al., 2020).

Deep learning (DL) is an efficient and proven method for dealing with complex classification problems and can generate high-level spatial features automatically (Weng et al., 2020). In particular, 2D convolutional neural networks (CNNs) are the most popular deep learning models (Qiu et al., 2018) and are primarily applied to image recognition and image segmentation, which consider the spatial relationship among pixels (Wu et al., 2019; Yang et al., 2018; Pang et al., 2020; Yamaguchi et al., 2021). The advantage of the CNN is that the feature representation is done automatically without additional manual extraction (Wu et al., 2019). The DL method has been utilized to identify rice samples coupled with local image descriptors (Hoai et al., 2020). In contrast, because the NIR hyperspectral images, especially at longer wavelengths, can still provide the spatial distribution of the molecular vibration information (i.e., chemical component difference) in each sample, it is expected that the DL coupled with the NIR–HSI could classify the milled seed samples by detecting the different chemical components. The main issue of the HSI method combined with the CNN

method for further commercial applications could be time-consumption, mainly due to deep CNN layers and a large amount of data with a full wavelength range. Even though such a method may provide satisfying classification accuracy on the model construction process, the developed model could have stability problems because the full wavelength range generally includes much unnecessary information (Ma et al., 2018). Various efforts, such as wavelength selection, adjustment of spatial resolution, data compression, and uncomplex CNN layers, also need to be considered.

The objective of this study is to provide a powerful method that can identify rice varieties (KDML105, PTT1, and PSL2) of both brown and milled rice using NIR-HSI imaging coupled with simple CNN layers. The specific objectives are (1) to collect the NIR images (wavelength range: 1002-1998 nm; interval: 6.2 nm) and RGB images (wavelengths: 630 nm, 530 nm, and 440 nm) of the three rice species; (2) to establish SVM models using the PC scores calculated from mean NIR spectra with either a complete or selected wavelength; (3) to develop a CNN model using PC score images calculated from the HSI data and compare its classification accuracies with the SVM method; and (4) to compare the classification accuracies earned from the HSI data and conventional RGB images by the same CNN model.

6.2 Materials and methods

6.2.1 Sample preparation

To avoid adulteration during the rice production process in the factory, the paddy samples used in this research were shelled and milled in the laboratory at King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. The KDML105 paddy was obtained from a farmer in the Surin province of Thailand to avoid adulteration from traders. The PTT1 and PSL2 paddies were received from Kla Kraeng Pohkla company, Nakhon Sawan, Thailand. First, the amylose content of 200 g of randomly selected brown rice of KDML105, PTT1, and PSL2 was tested at Intertek Testing Services (Thailand) Ltd, Bangkok, Thailand following the method of notification of the Ministry of Commerce Subject: Prescribing Thai Hom Mali Rice as a Standardized Commodity

and the Standards of Thai Hom Mali Rice (third edition) B.E.2559 of the Thai Rice Research Center. Then, the brown rice was shelled with a rice husker (THU, SATAKE TESTER, Japan) and milled by a polisher (TM05, SATAKE, Japan). For brown and milled rice samples, all 2430 kernels (i.e., 810 kernels/species) without surface damage were used in the experiment.

6.2.2 NIR hyperspectral image collection

A push- broom line scanning system (Compovision, Sumitomo Electric Industries, Ltd., Osaka, Japan) was used to acquire NIR hyperspectral diffuse reflectance images. This system comprised the HSI camera possessing a spectroscope and a two-dimensional photosensitive element [256 pixels (wavelength from 913 to 2519 nm with a 6.2 nm spectral resolution) × 320 pixels (element of detectors)]. The distance between the samples and camera was manually adapted to obtain a horizontal field of view of 50 mm (with a spatial interval of approximately 156.3 μm/pixel). The source of illumination was a tube-shaped halogen light source at 45° to the sample. Thirty kernel samples were placed once on a moving slider and scanned line by line. The movement speed of the slider was set to 23.4 mm/s, and the exposure time of the camera was set to 6 ms (frame rate: 150 frames/*s*). Each kernel sample was scanned at ten positions by rotating the kernel randomly to be aware of the chemical composition distribution of each kernel. This is also suitable for large-scale practice applications in which the sample is scanned randomly. All collected reflectance spectral images of the kernel samples were converted to relative reflectance values by Eq. 6.1:

$$R_{\lambda,n} = \frac{S_{\lambda,n} - D_{\lambda,n}}{B_{\lambda,n} - D_{\lambda,n}} \quad (6.1)$$

where $R_{\lambda,n}$ is the relative reflectance value. λ and n represent the wavelength and pixel index variables, respectively. S is the sample image. B and D represent the white reference ($BaSO_4$) and dark image. The dark reference image was acquired by turning off the light source and completely covering the lens with its cap.

Consequently, MATLAB R2020b (The MathWorks, Natick, MA, USA) was utilized for spectral extraction, preprocessing, wavelength selection, PCA, and SVM operations.

6.2.3 RGB image collection

The classification results were also benchmarked against the standard deep learning approach based on digital RGB photos to quantify the added value of the NIR-HSI approach. The same rice samples were scanned to RGB images using a printer (Brother Industries, Ltd., Batangas, Philippines) in scanner mode at King Mongkut's Institute of Technology Ladkrabang, Thailand. Each kernel sample was scanned ten times, and each time the kernel was rotated randomly. Within the scanner mode of the printer system, images of 1325 ×1880 pixels (width × length), 162 dpi and 41 × 41 pixels/seed were adjusted to fix the size of the HSI image for each rice kernel.

6.2.4 Sample area extraction and averaged NIR spectral data calculation

Figure. 6.1 shows a flowchart of the HSI and RGB image analysis processes. The HSI method offers increased spectral images to hundreds of images, which is different from the RGB image in which the RGB image was 3 images. First, the HSI (converted to relative reflectance) images and the RGB images were cropped to a single kernel per image. Then, the region of interest (ROI) was selected based on the binary image (kernels: 1; and background: 0). For the HSI data, the pixels with the same spectral information of each rice kernel were averaged to collect its mean spectrum.

6.2.5 Pre-processing and wavelength selection

After spectral data extraction, the spectral images at wavelengths under 1002 nm and over 1998 nm were found to be noise. Thus, they were removed and comprised a total of 159 wavelengths for analysis. Then, Savitzky-Golay smoothing filter (polynomial order: 2; frame length: 11) spectral preprocessing was employed to remove noise while retrieving the original signal structure (Acharya et al., 2016). A larger number for the frame length of smoothing preprocessing aimed to reduce the noise in the spectra when comparing the differences between the original spectrum at each pixel and the mean spectrum. After that, second-derivative (segment: 11; gap: 0) preprocessing was used to enhance the resolution of sharply defined features where the overlapping peaks with broad bands would be revealed (Mathian et al., 2018).

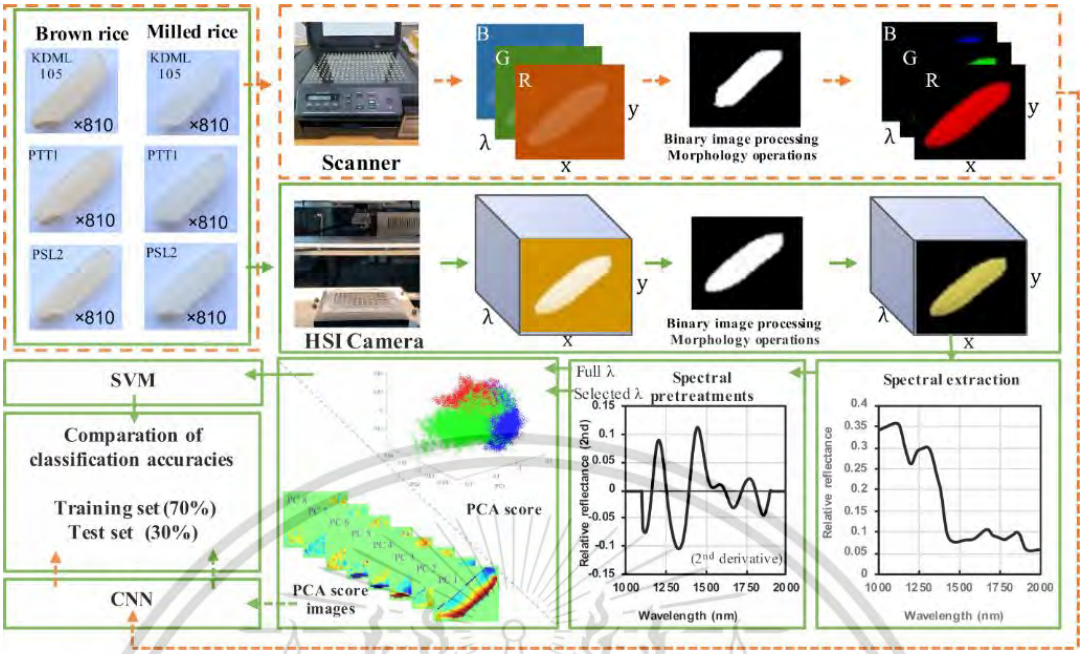

Figure. 6.1 A flowchart of the data analysis process in this study.

6.2.6 PCA

PCA was performed to reduce the high-dimensional spectral data matrix into a principal components (PCs) score matrix. It is widely applied in dimensionality reduction analysis of spectral data to remove redundant information (Zhang & Ji, 2019). In this study, NIR spectra with full wavelengths (i.e., 1002-1998 nm) and select wavelengths (i.e., selected important range of peaks from the second derivative) were subjected to compression by PCA. Their PC scores were further used to construct SVM and CNN models.

6.2.7 SVM analysis

SVM analysis was used to build classification models of the three species based on the PC scores. The SVM is a widely used classifier, and it is especially suitable for small datasets with high-dimensional feature spaces (Cortes and Vapnik. 1995; Zhang and Ji. 2019; Ma et al. 2020). During the model construction, the 70% (i.e. for 3 varieties of rice kernels, each containing 567 kernels, a total of 1701 samples × 10 rotations, a total of 17010 measurements was obtained) of total measurements for each model were randomly chosen for the training set. The remaining 30% (i.e. for 3 varieties of

rice kernels, each containing 243 kernels, a total of 729 samples × 10 rotations, a total of 7290 measurements was obtained) of total measurements were used for the test set. Additionally, a five-fold cross-validation set was applied to develop the best training classification model. The type of SVM kernel function for this study was a linear classification model. It is especially suitable for small datasets with high-dimensional feature spaces. The cross-validation results determine how well the model was generalized.

6.2.8 CNN deep learning

Generally, CNNs comprise five modules: input, convolutional layer, pooling layer, fully connected layer, and output (Pang et al., 2020). Because the pixel number of the NIR–HSI data for each sample was 41×41 in this study, it was necessary to build a sample and narrower CNN (only three layers) compared to classic ones such as the AlexNet model, VGG model, and Google Net model. It is expected that such a CNN could also be more robust for test validations because the CNN is much simpler than others, contributing to its robustness. Table 6.1 shows the customized layer structure of the neural network used for rice variety classification using HSI and RGB data of the same rice samples. For the CNN model, a customized layer structure was created using TensorFlow 1.13.1 & Keras 2.2.4 from open-source Python libraries for machine learning.

6.3 Results and discussion

6.3.1 NIR spectral characteristics of the three rice kernels

Fig. 6.2 shows the averaged reflectance spectra compiled from brown (Fig. 6.2a) and milled (Fig. 6.2b) rice from the different varieties (KDML105, PTT1, and PSL2). The peak absorption band at approximately 1202 nm was the result of the second overtone of symmetric stretching (C-H bonds) of methyl (CH_3) groups, which presented the NIR absorption of amylose in rice kernels (Sousa et al., 2018); the first overtone of the O-H band stretching vibration at a strong peak absorption band at approximately 1450 nm was associated with starch, water and O-H symmetric stretching of amylose molecules in rice kernels (Siriphollakul et al., 2017; Sousa et al., 2018); a small absorption band at 1580 nm

Table 6.1 Structure of the customized CNN network for HSI and RGB data.

Layer (type)	Output Shape		Parameters	
	HSI	RGB	HSI	RGB
Input layer	41,41,8	41,41,3		
Convolutional layer (filter:64; kernel size: 3 × 3; kernel regularizer: 0.01; bias regularizer: 0.01)	41,41,64	41,41,64	4672	1792
Pooling layer (Average pooling, pooling size: 2 × 2)	20,20,64	20,20,64		
Convolutional layer (filter:128; kernel size: 3 × 3; kernel regularizer: 0.01; bias regularizer: 0.01)			36928	36928
Pooling layer (Average pooling, pooling size: 2 × 2)	10,10,64	10,10,64		
Convolutional layer (filter:128; kernel size: 3 × 3; kernel regularizer: 0.01; bias regularizer: 0.01)			36928	36928
Pooling layer (Average pooling, pooling size: 2 × 2)	5,5,64	5,5,64		
Flatten	1600	1600		
Dense	256	256	409856	409856
Drop-out layer	256	256		
Dense	128	128	32896	32896
Dense	3	3	387	387

was assigned to the first overtone (intermolecular H-bond) from O-H stretching of starch and glucose in rice kernels (Osborne et al., 1993); and the peak absorption band at approximately 1748 nm corresponded with the first overtone of the C–H stretching vibrations related to amino acids (Biancolillo et al., 2019) when the amino acid was in the protein composition (Hormoz, 2013). After spectral preprocessing (Fig. 6.3), it is evident that the PSL2 samples could have higher amylose, moisture content (MC), and protein content than the other rice species. The important chemical in differentiating varieties of rice is amylose content, which was divided into three groups: low (12-20%), medium (21-25%), and high (26-33%) amylose content cultivars (Varavinit et al., 2003; Suwannaporn, 2019). The difference in eating quality of rice is due to the difference in amylopectin structure and protein content (Suwansri & Meullenet, 2004; Kaur et al., 2016). For example, cooked rice texture with low amylose content becomes sticky and moist, while high-amylose-content rice is flaky and dry (Juliano et al., 1981). In this study, the amylose contents of KDML105, PTT1, and PSL2 rice varieties were approximately 15.56 (low amylose), 17.91 (low amylose), and 27.79% (high amylose), respectively (see the test report in the attached supplementary material). However, the starch granules are comprised almost entirely of two major polysaccharides, namely, amylose and amylopectin (Biswas & Juliano. 1988; Li & Gilbert. 2018).

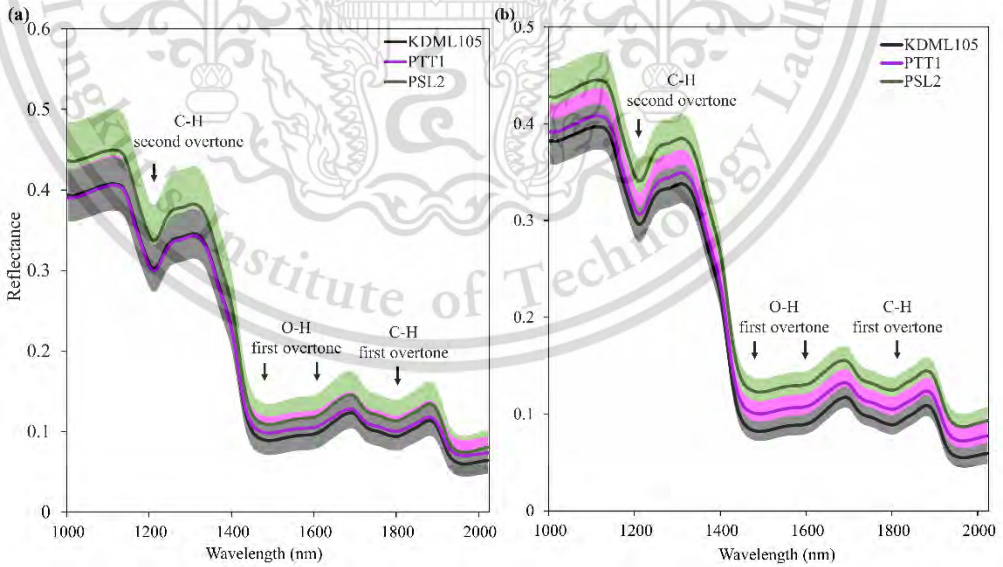

Figure. 6.2 Averaged near-infrared (NIR) reflectance spectral data (a: brown rice; b: milled rice) with their standard deviation.

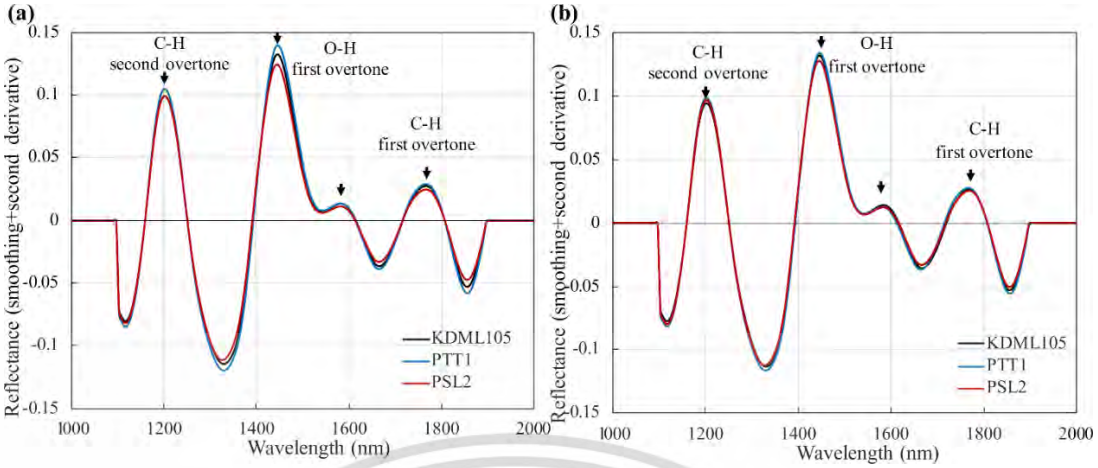

Figure. 6.3 Averaged near-infrared (NIR) reflectance spectral data (a: brown rice; b: milled rice) after smoothing and second-derivative preprocessing.

6.3.2 Preparation of PC score images from the HSI data

NIR spectra with a full wavelength range of 1002-1998 nm were first used to calculate the PC scores. In contrast, wavelengths (brown rice: 1191-1216, 1425-1533, 1564-1608, and 1734-1791 nm; milled rice: 1179-1230, 1419-1526, 1558-1614, and 1740-1791 nm) were selected to calculate the PC scores. The selected wavelength was chosen from the second-derivative (segment: 11; gap: 0) preprocessing, which selected a spectrum range from different rice varieties. Moreover, the wavelength range related to amylose is shown in detail in section 3.1. This result indicated the possibility of classifying the rice varieties by amylose content. The explained variance rate of the first three PCs in each model was 98.74%, 95.56%, 99.37%, and 98.18% for brown rice (full wavelengths), milled (full wavelengths), brown rice (selected wavelengths), and milled rice (selected wavelengths), respectively. Fig. 6.4 shows the first three PC loadings. These loadings can be understood as the weights for each variance value of spectral data, and useful information from a massive spectral dataset could be compressed into PC scores. Fig. 6.5 shows three-dimensional (3D) principal component (PC) score plots of all rice kernel models of brown and milled rice for the full and selected wavelengths. This indicates that the selected wavelength models of brown and milled rice presented the separation better than the full wavelength.

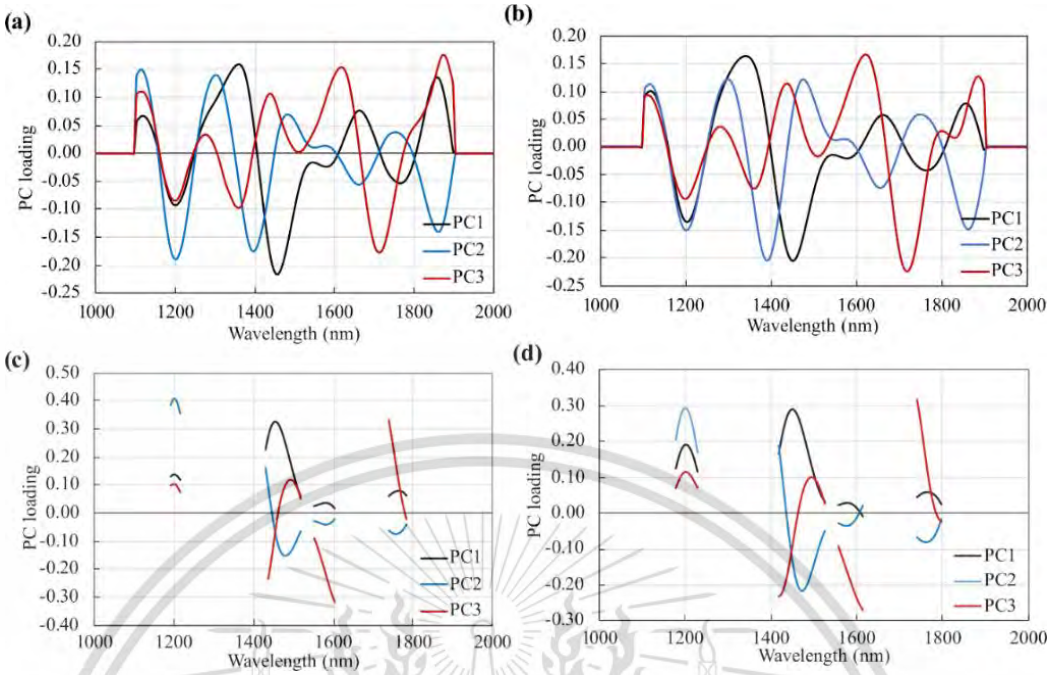

Figure. 6.4 Principal component (PC) loading plot based on three varieties of rice kernel spectral data for complete and selected wavelength models of brown (a and c, respectively) and milled rice (b and d, respectively).

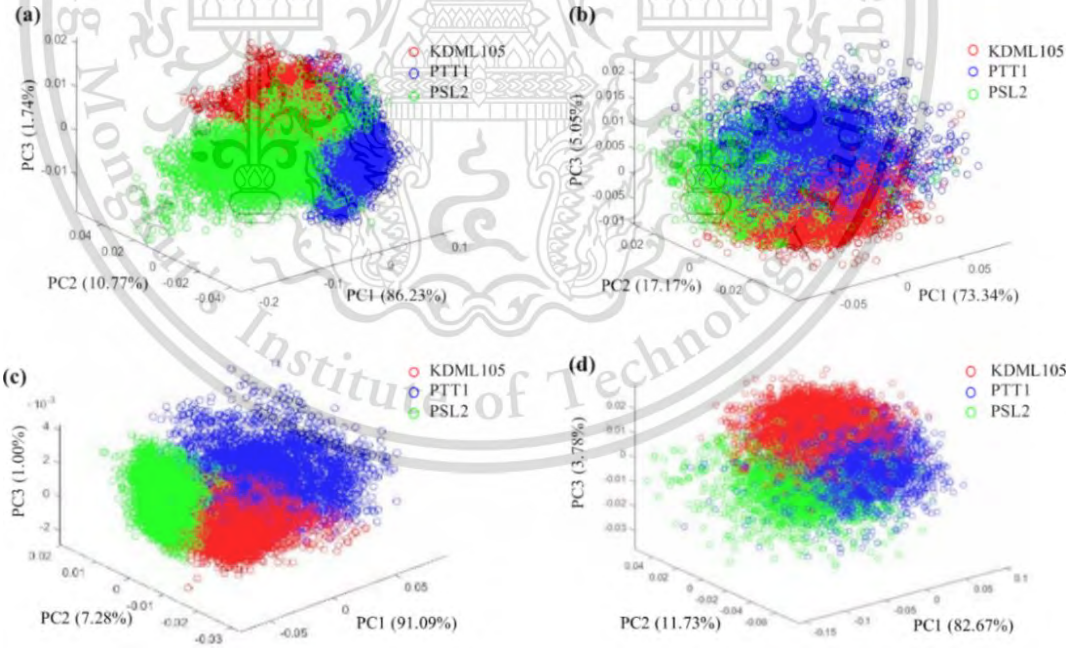

Figure. 6.5 The three-dimensional (3D) principal component (PC) score plot based on three varieties of rice kernel spectral data for complete and selected wavelength models of brown (a, c) and milled rice (b, d).

6.3.3 SVM results based on NIR-HSI spectra

Fig. 6.6 shows the confusion matrix of three variety classification results based on fivefold cross-validation by the PCA and SVM methods. The actual variety values of classification are exhibited in each row, while the prediction values are displayed in each column. The percentage in each box indicates the classification accuracy. Moreover, the box displays the number of rice kernels of each variety. The averaged accuracy of the training set for brown rice (full wavelength), milled (full wavelength), brown rice (selected wavelength), and milled rice (selected wavelength) models were 95.0%, 85.6%, 95.4%, and 86.0%, respectively. For the test set, they were 92.5%, 69.5%, 93.0%, and 85.1%, respectively. Given the optimal models of brown and milled rice expected from the selected wavelengths (i.e., light reflectance at 37 and 47 wavelengths, respectively), the averaged accuracies were higher than the full wavelength. The wavelength selection method generally improved the performance of the model during testing because it eliminated the unnecessary variables that were applied in spectral information analysis. However, the selected wavelength in this study was related to starch and protein, which is an important factor for classifying rice varieties.

6.3.4 Rice kernel variety classification accuracy based on PCA score images of HSIs using CNN

Fig. 6.7 shows the first eight PC score images (a: brown rice; b: milled), which were calculated using the same PC loadings for the SVM analysis above. For instance, the PC1 loading had a relatively large positive value at wavelengths correlating with the O¬-H band. Thus, the PC1 score image mainly contains MC and amylose content information. The PC2 loading had a relatively large positive value at wavelengths correlating with the second overtone of symmetric stretching (C-H bonds) of methyl (CH_3) groups, which suggests that the PC2 score image mainly contains amylose content information. Hence, the PC score images provide the chemical component differences, such as the amylose and protein contents, and provide their spatial distribution differences in each rice kernel.

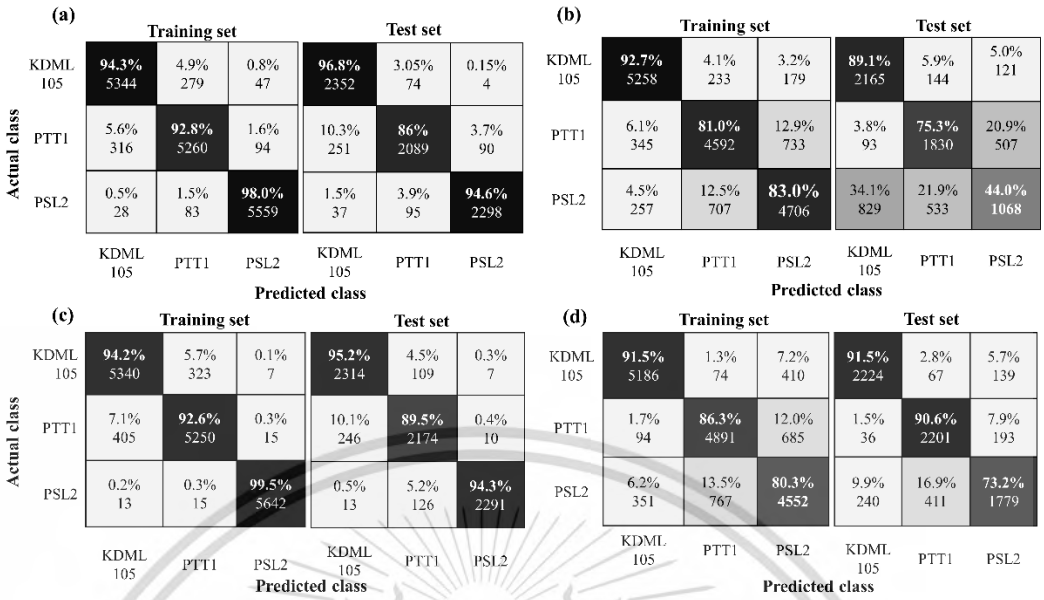

Figure. 6.6 Confusion matrix of brown rice for the SVM model using (a) full wavelengths and (c) selected wavelengths for training and test sets. The confusion matrix of milled rice for the SVM model using (b) full wavelengths and (d) selected wavelengths for training and test sets.

Fig. 6.8 shows the confusion matrix by the optimal CNN model for the test set PC score images. The average classification accuracies of the 5 tests were approximately 84.11 ± 0.52% and 95.15 ± 0.81% for brown and milled rice, respectively. The DL classification accuracy of the brown rice was almost the same as that of the SVM method, whereas for the milled rice, the accuracy was much higher. For the HSI data collected from brown rice, surface bran could affect the detected spatial distribution differences of the component differences inside the rice, which could cause the accuracy of the CNN method to not be as high as that of the SVM method. Moreover, Zhong et al. (2019) represents epochs that train the neural network once with the entire training set. If the number of epochs is too small, the fitting effect will be very poor. If the number of epochs is too large, the fitting error will no longer decrease, but the training time will continue to increase. Thus, the appropriate number of epochs should both meet the fitting precision and decrease the training time.

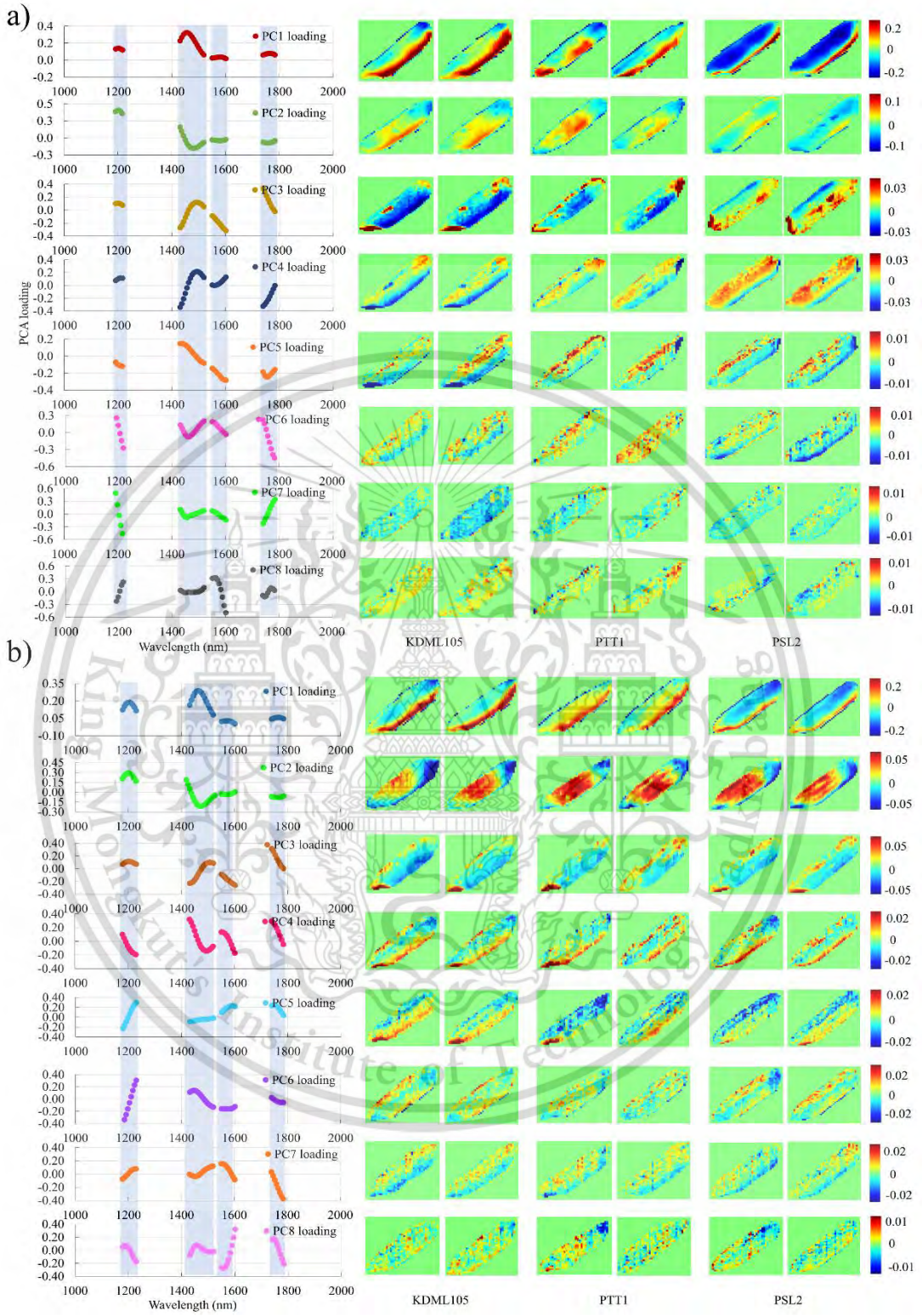

Figure. 6.7 The first eight PC score images (a: brown rice; b: milled) calculated using the same PC loadings for the SVM analysis.

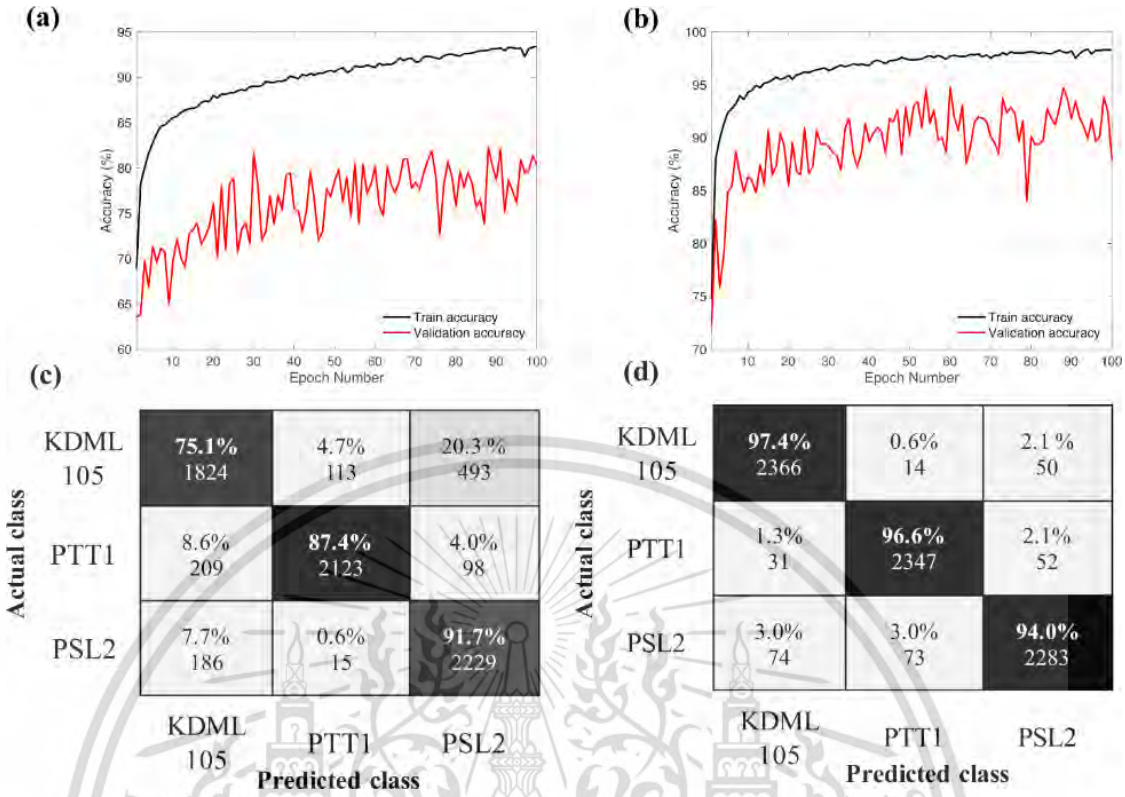

Figure. 6.8 Confusion matrix by the optimal CNN model using PC score images for brown (a: training set; c: test set) and milled (b: training set; d: test set) rice variety classification.

6.3.5 Rice variety classification accuracy based on RGB images of HSIs using CNN

Fig. 6.9 shows the confusion matrix by the optimal CNN model for the test set RGB images. The average classification accuracies of the 5 tests were approximately 85.45 ± 3.49% and 80.11 ± 4.47% for brown and milled rice, respectively. The DL classification accuracy of the brown rice was almost the same as that of the DL method using the HSI data, whereas for the milled rice, the accuracy was much lower and more unstable.

The optimal rice kernel variety classification results were obtained by the SVM and CNN models for brown and milled rice using the HSI data. It is evident that the HSI approach is better than the RGB approach. The reason could be that the NIR data contain the chemical composition differences between the three varieties. Previous

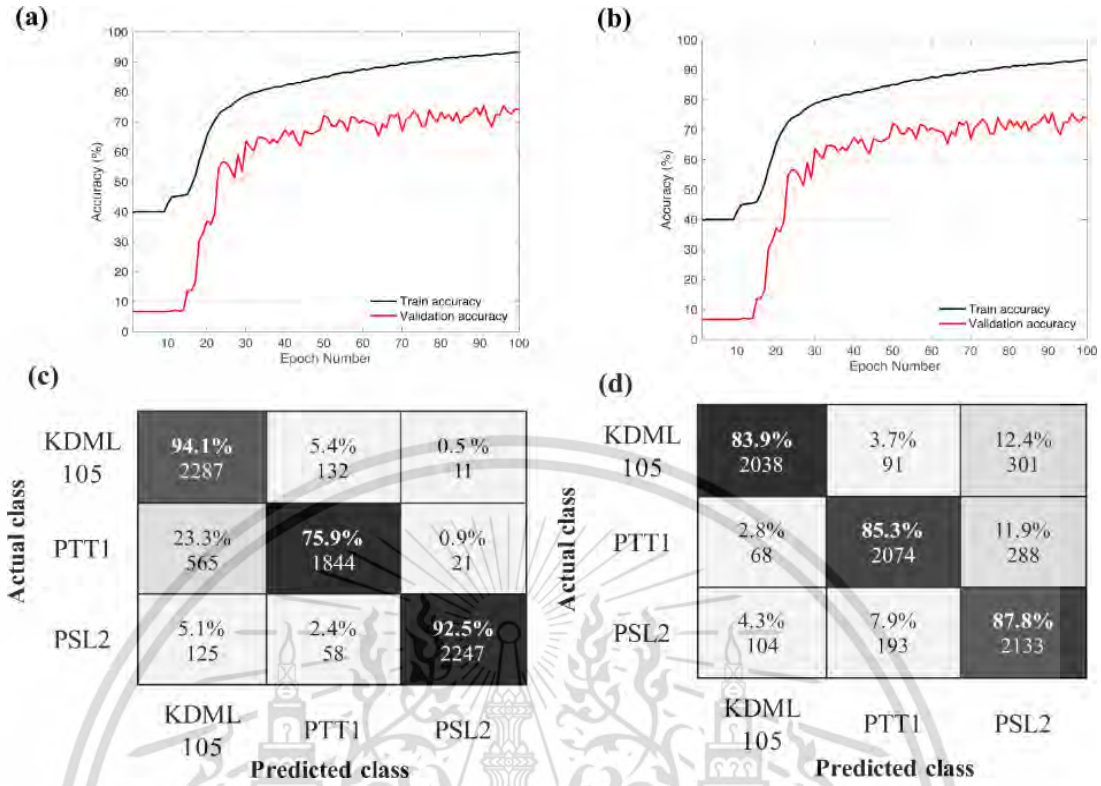

Figure. 6.9 Confusion matrix by the optimal CNN model using RGB images for brown (a: training set; c: test set) and milled (b: training set; d: test set) rice variety classification.

Table 6.2 shows the summary of the classification results.

Type of model		Percentage of average accuracy (%)	
		Brown rice	Milled rice
Machine learning	SVM model based on NIR spectral	93.0	85.1
Deep learning	CNN model based on HSI	84.1±0.5	95.2±0.8
	CNN model based on RGB image	85.5±3.5	80.1±4.5

research has investigated the ability of the CNN method to identify aphid infection in cotton leaves by HSI information and RGB imaging (Yan et al., 2021). The HSI approach

was recommended for the optimal model, which is consistent with the results of this research. This suggests that the classification results could be improved by taking advantage of spatial features and spectral features from the HSI approach. Although the physical characteristics of the three varieties are quite similar, which makes them difficult to classify by human eyes or human sensory evaluation, the classification accuracy was unexpectedly high when using the RGB images. The CNN results based on RGB images were approximately 85.5 and 80.1% for brown and milled rice, respectively. It has been shown that KDML105 is generally slightly longer but smaller than the other species (Division of Rice Research and Development, 2016), which could be the main reason that RGB data can be used for prediction. However, the physical characteristics (i.e., bran layer) of the brown rice samples were destroyed by milling, and the constructed CNN model was unstable. The experimental results suggest that CNN modelling based on HSI data is powerful for distinguishing the three species even after milling.

6.4 Conclusion

The NIR–HSI method was used to classify three brown or milled rice species (KDML105 PTT1 and PSL2). First, NIR images (wavelength range: 1002-1998 nm; interval: 6.2 nm) and RGB images (wavelengths: 630 nm, 530 nm, and 440 nm) of the three species were collected; then, the PC scores calculated from mean NIR spectra with either a complete or selected wavelength were used to construct SVM classification models, following which the PC score images of the HSI data and the conventional RGB images were used to construct CNN classification models. Overall, for brown rice, the best classification model was the SVM model based on the mean NIR spectra, and the accuracy classifications for the training and test sets were 95.4% and 93.0%, respectively. For milled rice, the best model was the CNN model based on the HSI data; the average classification accuracy for the test set was approximately 95.2%. The experimental results support the HSI method as potentially a rapid and nondestructive tool for preventing the adulteration of Thai jasmine rice compared with traditional methods (1-2 days).

The next steps would include validating and improving this approach by increasing the input of scanned HSI data for a variety of rice samples. By continuing to increase the sampling efforts, an even more thorough and effective defence against food fraud could be developed. This study is also expected to provide new insights for the classification of other targets with similar surface physical characteristics.

Chapter 7
Conclusion and recommendations

7.1 Conclusion

According to the findings of this thesis, the results can confirm that the NIR and NIR-HSI techniques were feasible to apply to evaluate the texture properties of cooked parboiled rice and KDML 105 rice authentication, respectively; however, the recommend dation recommendation is given.

The first experiment presented in chapter 3, the objective of this study of which the objective was to evaluate the precision of FT-NIR spectroscopy for measuring the texture characteristics of cooked parboiled rice. The scanning repeatability and reproducibility were 0.002769 and 0.017151 respectively, which indicated that the FT-NIR spectrometer was precise. Moreover, ISO 11747 Rice- Determination of Rice Kernel Resistance to Extrusion after Cooking method using a texture analyzer was studied the precision of the reference method.It was found that the repeatability was 4.45 N and 61. 56 N mm for hardness and toughness, respectively. The error from the reference method was 6 and 19 % for hardness, and toughness, respectively. This may be due to the transfer of cooked parboiled rice from the beaker to the plastic case in the preparation for the texture measurement. This step may affect the texture of cooked parboiled rice. Because the cooked parboiled rice that comes out from the beaker looks lumpy, it has to be taken out by spoon into a plastic box.

The preliminary study of this thesis was conducted on the influence of processing parameters of parboiled rice on its physiochemical and texture properties presented in Chapter 4. The study demonstrated that the different parboiled rice process conditions influenced chemical properties of parboiled rice. It was found that the amylose content had a significantly negative correlation with the hardness of cooked parboiled rice (Correlation Coefficient (r)=-0.52), and a negative correlation with its toughness (r=-0.38). The fat content had a low correlation with the hardness (r=0.20) and the toughness (r=0.12) of texture properties of cooked parboiled rice. Therefore,

it is possible to apply NIR spectroscopy to evaluate the texture properties of parboiled rice because the chemicals constituent studied did not correlated well with texture of cooked parboiled rice. However, accuracy of texture evaluation test was low, where the repeatability of reference test of this part was 11.18 N and 44.03 Nmm for the hardness and the toughness, respectively.

The developed calibration model suitable for evaluating the hardness and toughness of cooked parboiled rice exhibited in chapter 5 was created based on spectral of milled parboiled rice, which was scanned by a Fourier transform (FT-NIR) spectrometer (MPA, Bruker Ltd., Germany) at a wavenumber of 12,500–4,000 cm^{-1} (800-2,500 nm). The PLSR best calibration model of hardness gave coefficient of r^2, RMSEP, and RPD of 0.70, 7.24 N and 1.93, respectively. The Principal Components Regression (PCR) best model of toughness provided r^2, RMSEP and RPD of 0.66, 38.00 Nmm and 1.75, respectively. The result indicated that the hardness and toughness model can be used for rough screening and some other approximate calibrations. The R^2_{max} were 0.75 (75%) and 0.64 (64%) for the hardness and toughness, respectively, showing that the errors of reference test were 25% (100-75) and 36% (100-64). Therefore, an error value from NIR spectra scanning and other error was 5% (100-70-25) and 1% (100-64-35) for hardness and toughness models, respectively, which indicated that the NIR scanning protocol developed in this study was applicable though the reference test of hardness and toughness needs to be improved.

Chapter 6 illustrated that the NIR-HSI technique which were evaluated to classify three rice varieties (KDML105, PTT1 and PSL2) at either brown or milled conditions. The results indicated that the best model for brown rice was the SVM model based on the averaged NIR spectra, of which the classification accuracy of the test set was 93.0 %, whereas the best model for milled rice was the CNN model based on the NIR-HSI data, for which the classification accuracy of the test set was approximately 95.2%. Also, both brown and milled rice classification models were

found with the different chemical components, such as the amylose and protein contents which provide different spatial distributions in each variety.

7.2 Recommendations

The improvement of hardness and toughness model performance should involve practicing preparing samples for texture analysis to reduce the error from the reference test. Also, the model should add more variation of milled whole grain parboiled rice samples i.e., more samples from more varieties, field locations, seasons, etc. of rice. However, the relationship between chemical composition and texture properties was not high, which indicated that the chemical composition was only one part of the change in the texture of the cooked parboiled rice. Therefore, the further analysis with the X-ray diffractometer and differential scanning calorimetry test which are related to physical properties and chemical properties of the parboiled rice should be conducted, to understand additional factors i.e., gelatinization enthalpy and the amorphization of starch which might affect the texture properties of cooked parboiled rice.

In the case of KDML 105 rice authentication, the improved performance of the model classification may be done by increasing the number of varieties of rice samples such as Chai Nat 1 and RD15 and so on, which is similar to the KDML105. This helps to deter deception that leads to the misunderstanding for the premium KDML105 variety. NIR-HSI technique could be applied to the rice industry to check for adulteration of KDML105 rice, which is more reliable for export.

Bibliography

Abu-Khalaf, N., and Hmidat, M. 2020. "Visible/Near Infrared (VIS/NIR) spectroscopy as an optical sensor for evaluating olive oil quality." **Computers and Electronics in Agriculture**. 173 : 105445.

Acharya, D., Rani, A., and Agarwal, S. 2016. "Application of adaptive Savitzky-Golay filter for EEG signal processing." **Perspectives in Science**. 8 : 677–679.

Akhter, M., Ali, M. A., Haider, Z., and Muzammil, S. 2015. "Efficacy of parboiling on physico-chemical properties of some promising lines/varieties of rice (oryza satival)." Science. **Technology and Development**. 3(3) : 115–122.

Al-mbaideen, A. A. 2018. "Determination of Glucose Concentration from Near-Infrared Spectra Using PLSR Coupled with a Median Filter." **Jordan Journal of Electrical Engineering**. 4(4) : 196–206.

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. 2021. "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." **Journal of Big Data**. 8(1) : 1-74.

Andrade-Garda J.M. 2009. **Basic Chemometric Techniques in Atomic Spectroscopy**. 1ed. Cambridge : The Royal Society of Chemistry.

Andrade-Garda, J. M. 2013. **Basic Chemometric Techniques in Atomic Spectroscopy**. 2ed. Cambridge : The Royal Society of Chemistry.

Attaviroj, N., and Noomhorm, A. 2014. "Discriminant analysis of multiple physicochemical properties for Thai rough rice varietal authentication." **International Journal of Food Properties**. 17 : 1136–1149.

Attaviroj, N., and Noomhorm, A. 2014. "Discriminant analysis of multiple physicochemical properties for thai rough rice varietal authentication." **International Journal of Food Properties**. 17(5) : 1136–1149.

Attaviroj, N., and Noomhorm, A. 2014. "Discriminant analysis of multiple physicochemical properties for Thai rough rice varietal authentication." **International Journal of Food Properties**. 17 : 1136–1149.

Attaviroj, N., Kasemsumran, S., and Noomhorm, A. 2011. "Rapid variety identification of pure rough rice by fourier-transform near-infrared spectroscopy." **Cereal Chem**. 88(5) : 490–496.

Awanthi, M. G. G., Jinendra, B. M. S., Navaratne, S. B., and Navaratne, C. M. 2019. "Adaptation of visible and short wave Near Infrared (VIS-SW-NIR) common PLS model for quantifying paddy hardness." **Journal of Cereal Science**. 89 : 102795.

Badaró, A. T., Garcia-Martin, J. F., López-Barrera, M. del C., Barbin, D. F., and Alvarez-Mateos, P. 2020. "Determination of pectin content in orange peels by near infrared hyperspectral imaging." **Food Chemistry**. 323 : 126861.

Bai, X., Zhang, C., Xiao, Q., He, Y., and Bao, Y. 2020. "Application of near-infrared hyperspectral imaging to identify a variety of silage maize seeds and common maize seeds." **The Royal Society of Chemistry**. 10(20) : 11707–11715.

Bao, J.S., Cai, Y.Z., and Corke, H. 2001. "Prediction of rice starch quality parameter by near-infrared reflectance spectroscopy." **Journal of Food Science**. 66 : 936-939.

Barreto, A., Cruz-Tirado, J. P., Siche, R., and Quevedo, R. 2018. "Determination of starch content in adulterated fresh cheese using hyperspectral imaging." **Food Bioscience**. 21 : 14–19.

Bello, M., Baeza, R., and Tolaba, M. P. 2006. "Quality characteristics of milled and cooked rice affected by hydrothermal treatment." **Journal of Food Engineering**. 72 : 124–133.

Benes, E., Gere, A., and Fodor, M. 2020. "Predicting macronutrients and energy content of snack products using FT-NIR analysis and chemometric techniques." **Journal**

of Food Engineering. 280 : 109954.

Berg, F. Van Den, and Engelsen, S. B. 2009. "Review of the most common pre-processing techniques for near-infrared spectra." **Trends in Analytical Chemistry**. 28(10) : 1201–1222.

Bertoft, E. 2017. "Understanding Starch Structure : Recent Progress." **Agronomy**. 7(3) : 56.

Bhatnagar, S., Gill, L., and Ghosh, B. 2020. "Drone image segmentation using machine and deep learning for mapping raised bog vegetation communities." **Remote Sensing**. 12 : 2602.

Bhattacharya K.R., and Subba, P. V. 1966. "Processing Conditions and Milling Yield in Parboiling of Rice." **Journal of Agricultural Food Chemistry**. 14(5) : 473–475.

Biancolillo, A., Firmani, P., Bucci, R., Magri, A., and Marini, F. 2019. "Determination of insect infestation on stored rice by near infrared (NIR) spectroscopy." **Microchemical Journal**. 145 : 252–258.

Billiris, M. A., Siebenmorgen, T. J., Meullenet, J., and Mauromoustakos, A. 2012. "Rice degree of milling effects on hydration , texture , sensory and energy characteristics : Part 1 . Cooking using excess water." **Journal of Food Engineering**. 113(4) : 559–568.

Biswas, S. K., and Juliano, B. 1988. "Laboratory Parboiling Procedures and Properties of Parboiled Rice from Varieties Differing in Starch Properties." **Cereal Chem**. 65(5) : 417–423.

Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and Pages, J. 1999. "Calibration in non-linear near infrared reflectance spectroscopy: a comparison of several methods." **Analytica Chimica Acta**. 384(2) : 207-214.

Boling, A. A., Bouman, B. A. M., Tuong, T. P., Konboon, Y., and Harnpichitvitaya, D. 2011. "Yield gap analysis and the effect of nitrogen and water on photoperiod-sensitive Jasmine rice in north-east Thailand." **Wageningen Journal of Life**

Sciences. 58 : 11–19.

Boonpermpol, M., Sukprakarn, W., Thuama, S., Chanasit, U., Ananan, Y., Wongpiyachon, S., Sane, A., and Thani, P. 2018. "Effect of Rice Varieties and Compositions on Thermoplastic Starch-Based Blends." **Thai rice research journal**. 9(2) : 85–99.

Borzov, S. M., and Potaturkin, O. I. 2018. "Spectral-spatial methods for hyperspectral image classification." **Instrumentation and Data Processing**. 54(6) : 582–599.

Boulesteix, A. L., and Strimmer, K. 2007. "Partial least squares: A versatile tool for the analysis of high-dimensional genomic data." **Briefings in Bioinformatics**. 8(1) : 32–44.

Brereton R.G. 2018. **Chemometrics: data driven extraction for science**. second ed. Chennai, India : John Wiley & Sons.

Buggenhout, J., Brijs, K., and Delcour, J. A. 2014. "Soaking conditions during brown rice parboiling impact the level of breakage-susceptible rice kernels." **Cereal Chemistry**. 91(6) : 554–559.

Burges, C. J. C. 1998. "A Tutorial on Support Vector Machines for Pattern Recognition." **Data Mining and Knowledge Discovery**. 2 : 121–167.

Cameron, D. K., and Wang, Y. 2005. "A Better Understanding of Factors That Affect the Hardness and Stickiness of Long-Grain Rice." **Cereal Chemistry**. 82(2) : 113–119.

CAMO. 2015. **The Unscrambler Appecdices: Method Reference**. [Online]. Available : http://www.camo.com/TheUnscrambler/Appendices

Candolfi, A., Maesschalck, R. De, Jouan-Rimbaud, D., Hailey, P. A., and Massart, D. L. 1999. "The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra." **J. Pharm. Biomed. Anal**. 21 : 115–32.

Carpita, N. C. 1996. "Structure and biogenesis of the cell walls of grasses." **Annual Review of Plant Physiology and Plant Molecular Biology**. 47(1) : 445–76.

Cen, H., and He, Y. 2007. “Theory and application of near infrared reflectance spectroscopy in determination of food quality.” **Trends in Food Science and Technology**. 18(2) : 72–83.

Chaijan, M., and Panpipat, W. 2020. “Nutritional composition and bioactivity of germinated Thai indigenous rice extracts : A feasibility study.” **Plos one**. 15(8) : e0237844.

Chakravarty, H. B., and Ghose, T. K. 1966. **Studies on the hydration of indian paddy. III A: Efficiency of elevated temperature on composition of oil in bran**. Calcutta : Modern Tecnology of rice milling.

Champagne, E. T., Bett, K. L., Vinyard, B. T., McClung, A. M., Barton, F. E., Moldenhauer, K., Linscombe, S., and McKenzie, K. 1999. “Correlation between cooked rice texture and rapid visco analyser measurements.” **Cereal Chemistry**. 76(5) : 764–771.

Champagne, E. T., Bett-Garber, K. L., Grimm, C. C., McClung, A. M., Moldenhauer, K. A., Linscombe, S., McKenzie, K. S., and Barton, F. E. 2001. “Near-infrared reflectance analysis for prediction of cooked rice texture.” **Cereal Chemistry**. 78(3) : 358–362.

Champagne, E. T., Lyon, B. G., Min, B. K., Vinyard, B. T., Bett, K. L., Barton, F. E., Webb, B. D., McClung, A. M., Moldenhauer, K. A., Linscombe, S., McKenzie, K. S., and Kohlwey, D. E. 1998. “Effects of postharvest processing on texture profile analysis of cooked rice.” **Cereal Chemistry**. 75(2) : 181–186.

Champagne, E., Wood, D., Juliano, B., and Bechtel, D. B. 2004. “The rice grain and its gross composition.” **Rice: Chemistry and Technology.** Paul, Minnesota : American Association of Cereal Chemists.

Chang C.W., Laird D.A., Mausbach M.J., and Hurburgh C.R. 2001. “Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties.” **Soil Science Society of America Journal**. 65(2) : 480-490.

Cho, J. H. 2013. “Vibration spectroscopic techniques and application: Chemometrics tools for multivariate analysis.” **Symposium for online near infrared spectroscopy for industrial process control**. 1–29.

Conzen, J. P. 2006. **Multivariate Calibration: A practical guide for developing methods in the quantitative analytical chemistry**. Ettlingen, Germany : Bruker Optik GmbH

Cortes, C., and Vapnik, V. 1995. Support-Vector Networks. Machine learning. 20(3). 273-297.

Dahm, K. D., and Dahm, D. J. 2021. “**Near-Infrared Spectroscopy.**” Singapore : Springer Nature Singapore.

Danezis, G. P., Tsagkaris, A. S., Brusic, V., and Georgiou, C. A. 2016. “Food authentication: state of the art and prospects.” **Current Opinion in Food Science**. 10 : 22–31.

Dardenne, P. 2010. “Some considerations about NIR spectroscopy: Closing speech at NIR 2009.” **NIR news**. 21(1) : 8-14.

Department of foreign trade Ministry of Comerce. 2016. **Thailand standards for rice**. [Online]. Available : https://www.dft.go.th

Derycke, V. 2007. **Parboiling of Rice: Changes in Starch and Protein and their Relation to Cooking Properties.** Katholiek University, Leuven, Belgium.

Derycke, V., Veraverbeke, W. S., Vandeputte, G. E., Man, W. De, Hoseney, R. C., and Delcour, J. A. 2005. “Impact of Proteins on Pasting and Cooking Properties of Nonparboiled and Parboiled Rice.” **Cereal Chemistry**. 82(4) 468–474.

Division of Rice Research and Development. 2016. **Rice Knowledge Bank**. [Online]. Available : http://www.ricethailand.go.th/rkb3/Varieties.htm

Du, J., Yang, Z., Xu, X., Wang, X., and Du, X. 2019. “Effects of tea polyphenols on the structural and physicochemical properties of high-hydrostatic-pressure-gelatinized rice starch.” **Food Hydrocolloids**. 91(130) : 256–262.

Dunn, K. 2021. “Principal Component Regression (PCR).” **Process Improvement Using Data**. 371.

Elainy, V. dos S. P., Fernandes, D. D. de S., Araújo, M. C. U. de, Diniz, P. H. G. D., and Maciel, M. I. S. 2021. “In-situ authentication of goat milk in terms of its adulteration with cow milk using a low-cost portable NIR spectrophotometer.” **Microchemical Journal**. 163 : 105885.

ElMasry, G. M., and Nakauchi, S. 2016. “Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality - A comprehensive review.” **Biosystems Engineering**. 142 : 53–82.

ElMasry, G., and Nakauchi, S. 2015. “Prediction of meat spectral patterns based on optical properties and concentrations of the major constituents.” **Food Science and Nutrition**. 4(2) : 269–283.

ElMasry, G., Kamruzzaman, M., Sun, D. W., and Allen, P. 2012. “Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review.” **Critical Reviews in Food Science and Nutrition**. 52(11) : 999–1023.

Erkinbaev, C., Henderson, K., and Paliwal, J. 2017. “Discrimination of gluten-free oats from contaminants using near infrared hyperspectral imaging technique.” **Food Control**. 80 : 197–203.

Fabiyi, S. D., Member, S., Vu, H. A. I., Tachtatzis, C., Andonovic, I., Member, S., Ren, J., Member, S., and Marshall, S. 2020. “ Varietal classification of rice seeds using RGB and hyperspectral images.” **IEEE Access**. 8 : 22493–22505.

Feng, L., Zhu, S., Liu, F., He, Y., Bao, Y., and Zhang, C. 2019. “Hyperspectral imaging for seed quality and safety inspection : a review.” **Plant Methods**. 15 : 1–25.

Fernandes, D. de S., Romeo, F., and Krepper, G. 2019. “Quantification and identification of adulteration in the fat content of chicken hamburgers using digital images and chemometric tools.” **Food Science and Technology**. 100 : 20–27.

Firmani, P., La Piscopia, G., Bucci, R., Marini, F., and Biancolillo, A. 2020. "Authentication of P.G.I. Gragnano pasta by near infrared (NIR) spectroscopy and chemometrics." **Microchemical Journal**. 152 : 104339.

Fonseca Jr, F. S., Pimentel, M. F., and Leal, L. B. 2020. "Determination of ethanol in gel hand sanitizers using mid and near infrared spectroscopy." **Journal of the Brazilian Chemical Society**. 31(9) : 1759–1763.

Frederick, M., & Franklin, E. 2002. "Determination of rapid visco analyzer parameters in rice by near-infrared spectroscopy." **Cereal Chemistry**. 79 : 563-566.

Fu, X., Kim, M. S., Chao, K., Qin, J., Lim, J., Lee, H., Garrido-Varo, A., Pérez-Marín, D., and Ying, Y. 2014. "Detection of melamine in milk powders based on NIR hyperspectral imaging and spectral similarity analyses." **Journal of Food Engineering**. 124 : 97–104.

Fu, X., Ying, Y., Lu, H., and Xu, H. 2007. "Comparison of diffuse reflectance and transmission mode of visible-near infrared spectroscopy for detecting brown heart of pear." **Journal of Food Engineering**. 83(3) : 317–323.

Ganopoulos, I., Argiriou, A., and Tsaftaris, A. 2011. "Adulterations in Basmati rice detected quantitatively by combined use of microsatellite and fragrance typing with high resolution melting (HRM) analysis." Food Chemistry. 129(2) : 652–659.

Ghasemi, E., Hamed Mosavian, M. T., and Haddad Khodaparast, M. H. 2009. "Effect of stewing in cooking step on textural and morphological properties of cooked rice." **Journal of Rice Science**. 16 : 243-246.

Godber, J. S., and Juliano, B. . 2004. "Rice lipids. In; Champagne, E.T. (Ed.)." **Rice chemistry and technology, third ed. American Association of cereal chemists**, Ina., Paul, Minnesota.

Graham-Acquaah, S. Manful, J. T., Ndindeng, S. A., and Tchatcha, D. A. 2015. "Effects of soaking and steaming regimes on the quality of artisanal parboiled rice." **Journal of food processing and preservation**. 39(6) : 2286-2296.

Guerrero, J. M., Pajares, G., Montalvo, M., Romeo, J., and Guijarro, M. 2012. "Support Vector Machines for crop/weeds identification in maize fields." **Expert Systems With Applications**. 39(12) : 11149–11155.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., and Lew, M. S. 2016. "Deep learning for visual understanding: A review." **Neurocomputing**. 187 : 27–48.

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(12) : 2227.

Hoai, D. P. Van, Surinwarangkoon, T., and Hoang, V. T. 2020. "A comparative study of rice variety classification based on deep learning and hand-crafted features." **ECTI Transactions on Computer and Information Technology (ECTI-CIT)**. 14 (1) : 1–10.

Hormoz, S. 2013. "Amino acid composition of proteins reduces deleterious impact of mutations." **Scientific reports**. 3 (1) : 1–10.

Hsu, C. W., and Lin, C. J. 2002. "A comparison of methods for multiclass support vector machines." **IEEE Transactions on Neural Networks**. 13(2) : 415–425.

Hu, J., Wang, Z., Wu, Y., Liu, Y., and Ouyang, J. 2018. "Rapid determination of the texture properties of cooked cereals using near-infrared reflectance spectroscopy." **Infrared Physics & Technology**. 94 : 165-172.

Hu, X., Lu, L., Guo, Z., and Zhu, Z. 2020. "Volatile compounds, affecting factors and evaluation methods for rice aroma: A review." **Trends in Food Science and Technology**. 97 : 136–146.

Huang, J., Romero-Torres, S., and Moshgbar, M. 2010. "Practical Considerations in Data Pre-treatment for NIR and Raman Spectroscopy." **American Pharmaceutical Review**.

Huang, M., Tang, J., Yang, B., and Zhu, Q. 2016. "Classification of maize seeds of different years based on hyperspectral imaging and model updating."

Computers and Electronics in Agriculture. 122 : 139–145.

Ibukun, E. O. 2008. "Effect of prolonged parboiling duration on proximate composition of rice." **Scientific Research and Essay**. 3(7) : 323–325.

Islam, M. R., Roy, P., Shimizu, N., and Kimura, T. 2002. "Effect of Processing Conditions on Physical Properties of Parboiled Rice Effect of Processing Conditions on Physical Properties of Parboiled Rice." **Journal of Food Science and Technology Research.** 8(2) : 106-112.

ISO. 2012. "**International Organization for Standardization 11747 Rice-Determination of Rice Kernel Resistance to Extrusion after the Cooking method**." Available : https://www.iso.org/standard/50736.html?fbclid=IwAR1cA8LdcjXyhYDqNgJjW3bLRCOI2NjUVpcs3A_m7VPQtYzKkUU3RrwYDA.

Jannasch, A., Carvalho, L., Patindol, J., Wang, Y. J., and Mauromoustakos, A. 2020. "Impact of kernel thickness on parboiled rice properties." **Cereal Chemistry**. 97(4) : 755–761.

Jayanayanan, E. K. 1965. "Influence of processing condition on the browning of parboiled rice." **Rice Journal**. 68 : 16–17.

Jiamyangyuen, S., Wichaphon, J., and Boonmeejoy, J. 2019. "Classification of rice cultivars by using chemical, physicochemical, thermal, hydration properties, and cooking quality." **Food and Applied Bioscience Journal**. 7(2) : 42–62.

Juliano, B. O., Perez, C. M., Barber, S., Blakeney, A. B., Tetsuy A Iwasaki, Shibuya, N., Keneaster, K. K., Chung, S., Laignelet, B., Launay, B., Mundo, A. M. Del, Suzuki, H., Shiki, J., Tsuji, S., Tokoyama, J., Tatsumi, K. K., and Webb, B. D. 1981. "International cooperative comparison of instrument methods for cooked rice texture." **Journal of Texture Studies**. 12(1) : 17–38.

Juliano, B. 1985. Criteria and tests for rice grain qualities. **Rice Chemistry and Technology**. 2nd Edition, American Association of Cereal Chemists.

Juliano, B. O. 1971. "A Simplified Assay for Milled-Rice Amylose." **Cereal science today**. 16 : 334–340.

Juliano, B.O. 1981. "International cooperative comparison of instrument method for cooked rice texture." **Journal of texture studies**. 12(1) : 17-38.

Kattenborn, T., Leitloff, J., Schiefer, F., and Hinz, S. 2021. "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing." **ISPRS Journal of Photogrammetry and Remote Sensing**. 173 : 24–49.

Kaur, P., Pal, P., Singh, A., and Amritpal, V. 2016. "Protein and starch characteristics of milled rice from different cultivars affected by transplantation date." **Journal of Food Science and Technology**. 53(8) : 3186–3196.

Kaur, S., Sharma, S., and Nagi, H. P. S. 2011. "Functional properties and anti-nutritional factors in cereal bran." **Journal of Food and Agro-Industry**. 4(2) : 122–131.

Kawata, S. 2002. "New techniques in near-infrared spectroscopy." **Near Infrared Spectroscopy: Principles, Instruments, Applications.** Weinheim, Germany : Wiley-VCH Verlag GmbH.

Kim, I., Kim, M. S., Chen, Y. R., and Kong, S. G. 2004. "Detection of skin tumors on chicken carcasses using hyperspectral fluorescence imaging." **Transactions of the American Society of Agricultural Engineers**. 47(5) : 1785–1792.

Kim, S. M., Kim, K. O., and Kim, S. K. 1986. "Effect of defatting on gelatinization of starch and cooking properties of Akibare (Japonica)and Milyang 30 (J-indica) milled rice." **Food Science and Technology Research**. 18 : 393–397.

Koavi, R., and Provost, F. 1998. "Guest editors' introduction: On applied research in machine learning." **Machine learning**. 30(2). 127-132.

Kong, W., Zhang, C., Liu, F., Nie, P., and He, Y. 2013. "Rice Seed Cultivar Identification Using Near-Infrared Hyperspectral Imaging and Multivariate Data Analysis." **Sensors**. 13 : 8916–8927.

Kong-ngern, K., Buaphan, T., Tulaphitak, D., Phuvongpha, N., Wongpakonkul, S., and Threerakulpisut, P. 2011. "Yield, yield components, soil minerals and aroma of KDML 105 rice in Tungkularonghai, Roi-Et, Thailand." **World Academy of Science, Engineering and Technology**. 52 : 335–340.

Korinsak, S., Siangliw, M., Kotcharerk, J., Jairin, J., and Toojinda, T. 2016. "Improvement of the submergence tolerance and the brown planthopper resistance of the Thai jasmine rice cultivar KDML105 by pyramiding Sub1 and Qbph12." Field Crops Research. 188 : 105–112.

Kress-Rogers, E., and Brimelow, C. J. B. 2001. **Instrumentation and sensors for the food industry.** Woodhead Publishing.

Krishni. 2018. **Data Driven Investor**. [Online]. Available : https://medium.datadriveninvestor.com/k-fold-cross-validation-6b8518070833

Kukusamude, C., and Kongsri, S. 2018. "Elemental and isotopic profiling of Thai jasmine rice (Khao Dawk Mali 105) for discrimination of geographical origins in Thung Kula Rong Hai area , Thailand." **Food Control**. 91 : 357–364.

Kwofie, E. M., and Ngadi, M. 2017. "A review of rice parboiling systems, energy supply, and consumption." **Renewable and Sustainable Energy Reviews**. 72 : 465–472.

Lamb, D. T., and Hurburgh, C. R. 1991. "Moisture determination in single soybean seeds by near-infrared transmittance." **Transactions of the American Society of Agricultural Engineers**. 34(5) : 2123–2129.

Lamberts, L., Brijs, K., Mohamed, V., and Jan A., D. 2006. "Impact of Browning Reactions and Bran Pigments on Color of Parboiled Rice." **Journal of Agricultural Food Chemistr**y. 54(26) : 9924–9929.

Lamberts, L., Rombouts, I., Brijs, K., Gebruers, K., and Delcour, J. A. 2008. " Impact of parboiling conditions on Maillard precursors and indicators in long-grain rice cultivars." **food chemistry**. 110 : 916–922.

Leelayuthsoontorn, P., and Thipayarat, A. 2006. "Textural and morphological changes of Jasmine rice under various elevated cooking conditions." **Food Chemistry**. 96(4) : 606–613.

Leethanapanich, K., Mauromoustakos, A., and Wang, Y. 2016. "Impacts of parboiling conditions on quality characteristics of parboiled commingled rice." **Journal of Cereal Science**. 69 : 283-289.

Li, H., and Gilbert, R. G. 2018. "Starch molecular structure : The basis for an improved understanding of cooked rice texture." **Carbohydrate Polymers**. 195 : 9–17.

Li, H., Prakash, S., Nicholson, T. M., Fitzgerald, M. A., and Gilbert, R. G. 2016. "The importance of amylose and amylopectin fine structure for textural properties of cooked rice grains." **Food Chemistry**. 196 : 702–711.

Li, H., Yan, S., Yang, L., Xu, M., Ji, J., Mao, H., Song, Y., Wang, J., and Sun, B. 2021. "Starch gelatinization in the surface layer of rice grains is crucial in reducing the stickiness of parboiled rice." **Food Chemistry**. 341 : 128202.

Li, Y., Wang, T., Yuan, H., Yang, G., and Zhang, N. 2019. "Study on Application of Infrared Spectrum Technology in Power Equipment." **MATEC Web of Conferences**. 260 : 02010.

Li, Z., Li, C., Gao, Y., Ma, W., Zheng, Y., Niu, Y., Guan, Y., and Hu, J. 2018. "Identification of oil, sugar and crude fiber during tobacco (Nicotiana tabacum L.) seed development based on near infrared spectroscopy." **Biomass and Bioenergy**. 111 : 39–45.

Liu, Z., Lv, X., Liu, K., and Shi, S. 2010. "Study on SVM compared with the other text classification methods." **IEEE**. 1 : 219–222.

Lu, S., Cik, T. T., Lii, C. Y., Lai, P., and Chen, H. H. 2013. "Effect of amylose content on structure, texture and **α**-amylase reactivity of cooked rice." **Food science and technology**. 54(1) : 224-228.

Luh, B. S. 1991. **Rice**. Springer, Boston. Department of Food Science and

TechnologyUniversity of California Davis USA.

Lv, B., Li, B., Chen, S., Chen, J., and Zhu, B. 2009. "Comparison of color techniques to measure the color of parboiled rice." **Journal of Cereal Science**. 50(2) : 262–265.

Ma, T., Li, X., Inagaki, T., Yang, H., and Tsuchikawa, S. 2018. "Noncontact evaluation of soluble solids content in apples by near-infrared hyperspectral imaging." **Journal of Food Engineering**. 224 : 53–61.

Ma, T., Tsuchikawa, S., and Inagaki, T. 2020. "Rapid and non-destructive seed viability prediction using near-infrared hyperspectral imaging coupled with a deep learning approach." **Computers and Electronics in Agriculture**. 177 : 105683.

Magwaza, L. S., Opara, U. L., Nieuwoudt, H., Cronje, P. J. R., Saeys, W., and Nicolaï, B. 2012. "NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit-A Review." **Food and Bioprocess Technology**. 5(2) : 425–444.

Mahadevappa, M., and Desikachar, H. S. R. 1968. "Some observations on histology of raw and parboiled rice." **Journal of Food Science and Technology-Mysore**. 5 : 72–73.

Mahatheeranont, S., Keawsa-ard, S., and Dumri, K. 2001. "Quantification of the rice aroma compound, 2-acetyl-1-pyrroline, in uncooked Khao Dawk Mali 105 brown rice." **Journal of Agricultural and Food Chemistry**. 49(2) : 773–779.

Manickavasagan, A., and Jayasuriya, H. 2014. **Imaging with Electromagnetic Spectrum**. New York, London : Springer Heidelberg.

Manley, M. 2014. "Near-infrared spectroscopy and hyperspectral imaging: Non-destructive analysis of biological materials." **Chemical Society Reviews**. Royal Society of Chemistry. 43(24) : 8200–8214.

Manolakis, D., Marden, D., and Shaw, G. a. 2003. "Hyperspectral Image Processing for Automatic Target Detection Applications." **Lincoln Laboratory Journal**. 14(1) : 79–116.

Martínez-Ramón, M., and Christodoulou, C. 2006. *Support vector machines for antenna* array processing and electromagnetics. **Synthesis Lectures on Computational Electromagnetics**. 1(1) : 1-120.

Mathian, M., Hebert, B., Baron, F., Petit, S., Lescuyer, J., Furic, R., and Beaufort, D. 2018. "Identifying the phyllosilicate minerals of hypogene ore deposits in lateritic saprolites using the near-IR spectroscopy second derivative methodology." **Journal of Geochemical Exploration**. 186 : 298–314.

Matveev, Y. I., Grinberg, V. Y., and Tolstoguzov, V. B. 2000. "The plasticizing effect of water on proteins , polysaccharides and their mixtures. Glassy state of biopolymers, food and seeds." **Food hydrocolloids**. 14 : 425–437.

Mees, C., Souard, F., Delporte, C., Deconinck, E., Stoffelen, P., Stévigny, C., Kauffmann, J. M., and De Braekeleer, K. 2018. "Identification of coffee leaves using FT-NIR spectroscopy and SIMCA." **Talanta**. 177 : 4–11.

Mehta, N., Shaik, S., Devireddy, R., and Gartia, M. R. 2018. "Single-Cell Analysis Using Hyperspectral Imaging Modalities." **Journal of Biomechanical Engineering**. 140(2) : 020802.

Mendez, J., Mendoza, L., Cruz-Tirado, J. P., Quevedo, R., and Siche, R. 2019. "Trends in application of NIR and hyperspectral imaging for food authentication." **Scientia Agropecuaria**. 10(1) : 143–161.

Meullenet, J. F., Mauromoustakos, A., Horner, T. B., and Marks, B. P. 2002. "Prediction of texture of cooked white rice by near-infrared reflectance analysis of whole-grain milled samples." **Cereal Chemistry**. 79(1) : 52–57.

Miah, M. A. K., Haque, A., Douglass, M. P., and Clarke, B. 2002. "Parboiling of rice. Part II : Effect of hot soaking time on the degree of starch gelatinization." **International Journal of Food Science and Technology**. 37 : 539–545.

Miao, W., Lufeng, W., Xiaoyun, X., and Siyi, P. 2016. "Evaluation of cooked rice texture using a novel sampling technique." **Measurement**. 89 : 21–27.

Mohammadi-Moghaddam, T., Razavi, S. M. A., Sazgarnia, A., and Taghizadeh, M. 2018. "Predicting the moisture content and textural characteristics of roasted pistachio kernels using Vis/NIR reflectance spectroscopy and PLSR analysis." **Journal of Food Measurement and Characterization**. 12(1) : 346–355.

Mohan, L., Lohumi, S., Kim, M. S., Kang, J., and Cho, B. 2016. " Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds." **Sensors and Actuators B: Chemical**. 229 : 534–544.

Mukasa, P., Wakholi, C., Mo, C., Oh, M., Joo, H. J., Suh, H. K., and Cho, B. K. 2019. "Determination of viability of Retinispora (Hinoki cypress) seeds using FT-NIR spectroscopy." **Infrared Physics and Technology.** 98 : 62–68.

Mutters R.G., and Thompson J.F. 2009. **Rice Quality Handbook**. California : University of California.

Nakauchi, S., Nishino, K., and Yamashita, T. 2012. "Selection of optimal combinations of band-pass filters for ice detection by hyperspectral imaging." **Optics Express**. 20(2) : 986.

Napasintuwong, O. 2019. " Rice Economy of Thailand." **Agricultural and Resource Economics Working Paper**. 1–16.

Nicolaï, 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 and Technology**. 46(2) : 99–118.

Novakovic, J., Veljovi, A., Iiic, S., Papic, Z., and Tomovic, M. 2017. "Evaluation of Classification Models in Machine Learning." **Theory and Applications of Mathematics & Computer Science.** 7(1) : 39–46.

Oli, P., Ward, R., Adhikari, B., and Torley, P. 2014. "Parboiled rice : Understanding from a materials science approach." **Journal of Food Engineering**. 124 : 173–183.

Osborne B.G., Fearn T., and Hindle P.H. 1993. **Practical NIR Spectroscopy with Applications in Food and Beverage Analysis**. second ed. Singapore : Longman scientific and technical.

Osborne, B. G. 2006. **Near-Infrared Spectroscopy in Food Analysis.** Encyclopedia of Analytical Chemistry.

Otto, M. 2017. **Chemometrics: Statistics and Computer Application in Analytical Chemistry**. Weinheim, Germany : Wiley-VCH Verlag GmbH & Co. KGaA.

Ozdemir, A., and Polat, K. 2020. "Deep Learning Applications for Hyperspectral Imaging: A Systematic Review." **Journal of the Institute of Electronics and Computer**. 2(1) : 39–56.

Özdoğan, G., Lin, X., and Sun, D. W. 2021. "Rapid and noninvasive sensory analyses of food products by hyperspectral imaging: Recent application developments." **Trends in Food Science and Technology**. 111 : 151–165.

Pallegedara, A. 2020. "Preference for parboiled rice: empirical evidence from Sri Lanka." **Journal of Agribusiness in Developing and Emerging Economies**. 10(5) : 613–628.

Pang, L. E. I., Men, S. E. N., and Yan, L. E. I. 2020. "Rapid Vitality Estimation and Prediction of Corn Seeds Based on Spectra and Images Using Deep Learning and Hyperspectral Imaging Techniques." **IEEE Access**. 8 : 123026–123036.

Park, J. K., Kim, S. S., and Kim, O. K. 2001. "Effect of milling ratio on sensory properties of cooked rice and on physicochemical properties of milled and cooked rice." **Cereal Chemistry**. 78(2) : 151–156.

Parnsakhorn, S., and Noomhorm, A. 2008. "Changes in Physicochemical Properties of Parboiled Brown Rice during Heat Treatment." **Agricultural Engineering International : the CIGR**. X : 1–20.

Patil, A., and Rane, M. 2021. "Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition." **Smart Innovation, Systems and**

Technologies. 195 : 21–30.

Patindol, J., J., N., and Y.-J., W. 2008. "Functional Properties as Affected by Laboratory-Scale Parboiling of Rough Rice and Brown Rice." **Food Engineering and Physical Properties**. 73(8) : 370–377.

Pelliccia, D. 2018. **Two scatter correction techniques for NIR spectroscopy in Python**. [Online] Available : https://nirpyresearch.com/two-scatter-correction-techniques-nir-spectroscopy-python/

Peng, Z., Yuan, X., Huang, Y., Mo, J., Tan, J., Zhou, H., and Wang, L. 2016. "Application of denaturing high-performance liquid chromatography for rice variety identification and seed purity assessment." **Molecular Breeding**. 36(2) : 1–14.

Pereira, E. V. dos S., Fernandes, D. D. de S., de Araújo, M. C. U., Diniz, P. H. G. D., and Maciel, M. I. S. 2020. "Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms." **Food Science and Technology**. 127 : 109427.

Pillaiyar, P., and Mohandoss, R. 1981. "Hardness and color in parboiled rices produced at low and high temperature." **Journal of food science and technology**. 18 : 7–9.

Pitiphunpong, S., Champangern, S., and Suwannaporn, P. 2011. "The Jasmine Rice (KDML 105 Variety) adulteration detection using physico-chemical properties." **Chiang Mai J. Sci**. 38(1) : 105–115.

Posom, J., Shrestha, A., Saechua, W., and Sirisomboon, P. 2016. "Rapid non-destructive evaluation of moisture content and higher heating value of Leucaena leucocephala pellets using near infrared spectroscopy." **Energy**. 107 : 464–472.

Puertas, G., and Vázquez, M. 2019. "Cholesterol determination in egg yolk by UV-VIS-NIR spectroscopy." **Food Control**. 100 : 262–268.

Qingyu, Z., Yong, X., and Qun, S. 2020. "Changes in the major aroma-active compounds and taste components of Jasmine rice during storage." **Food Research**

International. 109160.

Qiu, Z., Chen, J., Zhao, Y., Zhu, S., and He, Y. 2018. "Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network." **Applied Sciences**. 8(212) : 1–12.

Ramesh, M., Bhattacharya, K. R., and Mitchell, J. R. 2000. " Developments in understanding the basis of cooked-rice texture." **Critical Reviews in Food Science and Nutrition**. 40(6) : 449–460.

Reich, G. 2005. "Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications." **Advanced Drug Delivery Reviews**. 57(8) : 1109–1143.

Ren, G., Wang, Y., Ning, J., and Zhang, Z. 2020. "Using near-infrared hyperspectral imaging with multiple decision tree methods to delineate black tea quality." **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**. 237 : 118407.

Rinnan, A., Norgaard, L., Berg, F. Van Den, Thygesen, J., Bro, R., and Engelsen, S. 2009. "Pre-processing." **Infrared spectroscopy for food quality analysis and control**. United States of America : Academic.

Rittiron, R., Saranwong, S., and Kawano, S. 2005. "Detection of variety contamination in milled Japanese rice using a single kernel near infrared technique in transmittance mode." **Journal of Near Infrared Spectroscopy**. 13(1) : 19–25.

Romìa, M. B., and Bernàrdez, M. A. 2009. "Multivariate Calibration for Quantitative Analysis." **Infrared spectroscopy for Food quality analysis and control**. United States of America : Academic.

Rosenthal, A. J. 1999. **Food texture measurement and Perception, Approaches to instrumenttal measurement of texture.** Gaithersburg, Maryland : An Aspen Publication.

Roy, P., Orikasa, T., Okadome, H., and Nakamura, N. 2011. "Processing Conditions , Rice

Properties , Health and Environment." **International Journal of Environmental Research and Public Health**. 8 : 1957–1976.

Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., and Mononen, J. 2018. "Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle." **Behavioural Processes**. 148 : 56–62.

Saleh, M. I., and Meullenet, J. F. 2007. "Effect of moisture content at harvest and degree of milling (based on surface lipid content) on the texture properties of cooked long-grain rice." **Cereal Chemistry**. 84(2) : 119–124.

Sampaio, P. S., Castanho, A., Almeida, A. S., Oliveira, J., and Brites, C. 2020. "Identification of rice flour types with near-infrared spectroscopy associated with PLS-DA and SVM methods." **European Food Research and Technology**. 246(3) : 527–537.

Sandak, J., Sandak, A., and Meder, R. 2016. "Assessing trees, wood and derived products with near infrared spectroscopy: Hints and tips." **Journal of Near Infrared Spectroscopy**. 24(6) : 485–505.

Saniso, E., Prachayawarakorn, S., Swasdisevi, T., and Soponronnarit, S. 2019. "Parboiled rice production without steaming by microwave-assisted hot air fluidized bed drying." **Food and Bioproducts Processing**. 120 : 8–20.

Sarangapani, C., Devi, Y., Thirundas, R., Annapure, U. S., and Deshmukh, R. R. 2015. "Effect of low-pressure plasma on physico-chemical properties of parboiled rice." **Food Science and Technology**. 63(1) : 452–460.

Sarangapani, C., Thirumdas, R., Devi, Y., Trimukhe, A., Deshmukh, R. R., and Annapure, U. S. 2016. "Effect of low-pressure plasma on physicoechemical and functional properties of parboiled rice flour." **Food Science and Technology**. 69 : 482–489.

Sareepuang, K., Siriamornpun, S., Wiset, L., and Meeso, N. 2008. "Effect of Soaking

Temperature on Physical , Chemical and Cooking Properties of Parboiled Fragrant Rice Research Unit of Drying Technology for Agricultural Products." **Journal of Agricultural Sciences**. 4(4) : 409–415.

Sathyanarayana, S., and Amarappa, S. V. 2014. "Data classification using Support vector Machine (SVM), a simplified approaCH." **International Journal of Electronics and Computer S cience Engineering**. 3(4) : 435–445.

Savalia, S., and Emamian, V. 2018. "Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks." **Bioengineering**. 5(2) : 35.

Schaijk, T. van. 2019. **Image classification: MLP vs CNN**. [Online]. **Available :** https://www.peculiar-coding-endeavours.com/2019/mlp_vs_cnn/

Sergi, F. S. 2015. "NIR Techniques and Chemometric Data Analysis Applied To Food Adulteration." **Master's degree in Enabling Technologies for the Food and Bioprocessing Industry.**

Shuso, K., Motoyasu, N., Kazuhiro, T., and Kazuhiko, I. 2003. "Development of an automatic rice-quality inspection system." **Computers and Electronics in agriculture**. 40 : 115-126.

Siriphollakul, P., Kanlayanarat, S., Rittiron, R., Wanitchang, J., Suwonsichon, T., Boonyaritthongchai, P., and Nakano, K. 2015. "Pasting properties by near-infrared reflectance analysis of whole grain paddy rice samples." **Journal of Innovative Optical Health Sciences**. 8(06) : 1550035.

Siriphollakul, P., Nakano, K., and Kanlayanarat, S. 2017. "Eating quality evaluation of Khao Dawk Mali 105 rice using near- infrared spectroscopy." **Food Science and Technology**. 79 : 70–77.

Sohn, M., Himmelsbach, D. S., and Barton, F. E. 2004. "A comparative study of fourier transform Raman and NIR spectroscopic methods for assessment of protein and apparent amylose in rice." **Cereal Chemistry**. 81(4) : 429–433

Song, H., xiu-ying Han, Montenegro-Marin, C. E., and krishnamoorthy, S. 2021. "Secure

prediction and assessment of sports injuries using deep learning based convolutional neural network." **Journal of Ambient Intelligence and Humanized Computing**. 12(3) : 3399–3410.

Sousa, P., Soares, A., Castanho, A., So, A., Oliveira, J., and Brites, C. 2018. "Optimization of rice amylose determination by NIR-spectroscopy using PLS chemometrics algorithms." **Food Chemistry**. 242 : 196–204.

Srisawas, W., and Jindal, V. K. 2007. "Sensory evaluation of cooked rice in relation to water-to-rice ratio and physicochemical properties." **Journal of Texture Studies**. 38(1) : 21–41.

Sun, D.-W. 2009. **Infrared Spectroscopy for Food Quality Analysis and Control**. United States of America : Academic.

Sun, H. C., and Huang, Y. C. 2011. "Support vector machine for vibration fault classification of steam turbine-generator sets." **Procedia Engineering**. 24 : 38–42.

Sun, J. U. N., Lu, X., Mao, H., Jin, X., and Wu, X. 2015a. "A method for rapid identification of rice origin by hyperspectral imaging technology." **Journal of food process engineering**. 40 : e12297.

Sun, J. U. N., Lu, X., Mao, H., Jin, X., Wu, X., and Al, J. S. U. N. E. T. 2015b. "A method for rapid identification of rice origin by hyperspectral imaging technology." **Journal of Food Engineering**. 1–9.

Sun, J., Jin, X., Mao, H., Wu, X., Tang, K., and Zhang, X. 2013. "Identification of lettuce leaf nitrogen level based on adaboost and hyperspectrum." **Spectroscopy and Spectral Analysis**. 33 : 3372–3376.

Sun, W. 2010. **Hyperspectral Imaging for Food Quality Analysis and Control**. Ireland, Dublin : Agriculture & Food Science Centre.

Suwannaporn, P., Pitiphunpong, S., and Champangern, S. (2007). "Classification of rice amylose content by discriminant analysis of physicochemical properties. "

Starch-Stärke. 59(3-4) : 171-177.

Suwansri, S., and Meullenet, J. F. 2004. "Physicochemical characterization and consumer acceptance by asian consumers of aromatic Jasmine rice." **Sensory and Nutritive Qualities of food**. 69 : 30–37.

Teye, E., Anyidoho, E., Agbemafle, R., Sam-Amoah, L. K., and Elliott, C. 2020. "Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review." **Infrared Physics and Technology**. 104 : 103127.

Thai Rice Exporters Association. 2017. **Export Statistics**. [Online]. Available : http://www.thairiceexporters.or.th/

Thai Rice Exporters Association. 2020. **Rice export quantity and value : 2020**. [Online]. Available : http://www.thairiceexporters.or.th/statistic_2020.html

Thurber, M. C., Phadke, H., Nagavarapu, S., Shrimali, G., and Zerriffi, H. 2014. "'Oorja' in India: Assessing a large-scale commercial distribution of advanced biomass stoves to households." **Energy for Sustainable Development**. 19 : 138-150.

Timsorn, K., Lorjaroenphon, Y., and Wongchoosuk, C. 2017. "Identification of adulteration in uncooked Jasmine rice by a portable low-cost artificial olfactory system." **Measurement**. 108 : 67–76.

Timsorn, K., Lorjaroenphon, Y., and Wongchoosuk, C. 2017. "Identification of adulteration in uncooked Jasmine rice by a portable low-cost artificial olfactory system." **Measurement**. 108 : 67–76.

Tolaba, M. P., and Su, C. 2001. "Effects of drying conditions on head rice yield and browning index of parboiled rice." **Journal of Food Engineering**. 47 : 37–41.

Tsuchikawa, S. 2007. "Sampling technique." **Near Infrared spectroscopy in food science and technology**. USA : John Wiley & Sons.

Tulyathan, V., and Leeharatanaluk, B. 2007. "Changes in quality of rice (oryza sativa l.) cv. khao dawk mali 105 during storage vanna." **Journal of Food Biochemistry**.

31(2007) : 415–425.

Udomkun, P., Innawong, B., and Niruntasuk, K. 2018. "The feasibility of using an electronic nose to identify adulteration of Pathumthani 1 in Khaw Dok Mali 105 rice during storage." **Journal of Food Measurement and Characterization**. 12(4) : 2515–2523.

Vanier, F. A., Vanier, N. L., Madruga, N. A., Pesek, J., Matyska-Pesek, M., Elias, M. C., and Oliveira, M. 2017. "Improvement of the quality of parboiled rice by using anti-browning." **Journal of Food Chemistry**. 235 : 51–57.

Varavinit, S., Shobsngob, S., Varanyanond, W., Chinachoti, P., and Naivikul, O. 2003. "Effect of amylose content on gelatinization, retrogradation and pasting properties of flours from different cultivars of Thai Rice." **Starch-Sarkestarke**. 55 : 410–415.

Vasan, B. S., Iengar, N. G. C., Subramanyam, T. V., Chandrasekaran, R., and Subrahmanyan, V. 1971. "Effect of processing on the movement of oil in the rice kernel and its reation to the oil content of paddy and the yield of oil from rice bran in terregional seminar in the in dustry Processing of rice." **Food and Agriculture Organization**. Madras, India : FAO, ECAFF.

Villanova, F. A., Vanier, N. L., Avila, Madruga, N. de, Pesek, J., Matyska-Pesek, M., Elias, M. C., and Oliveira, M. de. 2017. "Improvement of the quality of parboiled rice by using anti-browning agents during parboiling process." **Journal of Food Chemistry**. 235 : 51–57.

Villareal, C.P., Cruz, N. M., and Juliano, B.O. 1994. "Rice amylose analysis by near-infrared transmittance spectroscopy." **Cereal Chemistry**. 71(3) : 292-296.

Visa, S., Ramsay, B., Ralescu, A., and Van der Knaap, E. 2011. "Confusion Matrix-based Feature Selection." **Proceedings of the Twenty-second Midwest Artificial Intelligence and Cognitive Science Conference**. 710 : 120–127.

Wang, H. L., Wan, X. Y., Bi, J. C., Wang, J. K., Jiang, L., Chen, L. M., Zhai, H. Q., and Wan,

J. M. 2006. “Quantitative analysis of fat content in rice by near-infrared spectroscopy technique.” **Cereal Chemistry**. 83(4) : 402–406.

Wang, L., Liu, D., Pu, H., and Sun, D. 2015. “Use of hyperspectral imaging to discriminate the variety and quality of rice.” **Food Anal. Methods**. 8 : 515–523.

Wang, Z. J., Turko, R., Shaikh, O., Park, H., Das, N., Hohman, F., Kahng, M., and Polo Chau, D. H. 2020. “CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization.” **IEEE Transactions on Visualization and Computer Graphics**. 27(2) : 1396–1406.

Wangcharoen, W., Phanchaisri, C., Daengpok, W., Phuttawong, R., Hangsoongnern, T., and Phanchaisri, B. 2016. “Consumer acceptance test and some related properties of selected KDML 105 rice mutants.” **Journal of Food Science and Technology**. 53(9) : 3550–3556.

Weng, S., Tang, P., Yuan, H., Guo, B., Yu, S., Huang, L., and Xu, C. 2020. “Hyperspectral imaging for accurate determination of rice variety using a deep learning network with multi-feature fusion.” **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**. 234 : 118237.

Williams P. 2019. **Near-infrared Technology: Getting the Best Out of Light (e-book)**. African : SUN Press.

Williams, P. 2007. **Near-infrared Technology-Getting the best out of light, A Short Course in the Pratical Implementation of Near-infrared Spectroscopy for the User**. 5th ed. Nanaimo Canada : PDK Grain.

Williams, P. 2009. “Influence of water on prediction of composition and quality factors: the aquaphotomics of low moisture agricultural materials.” **Journal of Near Infrared Spectroscopy.** 17(6) : 315-328.

Wold, S., Sjöström, M., and Eriksson, L. 2001. “PLS-regression: A basic tool of chemometrics.” **Chemometrics and Intelligent Laboratory Systems**. 58(2) : 109–130.

Workman Jr., J., and Weyer, L. 2012. **Practical guide and spectral atlas for interpretive near-infrared spectrpscopy**. Boca Raton London New York : CRC Press.

Workman, J., and Weyer, L., 2008. **Practical Guide to Interpretive Near-Infrared Spectroscopy**. USA : Taylor & Francis Group.

Wu, D., and Sun, D. 2013. "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment : A review - Part I : Fundamentals." **Innovative Food Science and Emerging Technologies**. 1–14.

Wu, H., Wang, L., Guo, N., Shu, C., and Lu, C. 2017. "Brillouin Optical Time-Domain Analyzer Assisted by Support Vector Machine for Ultrafast Temperature Extraction." **Journal of Lightwave Technology**. 35(19) : 4159–4167.

Wu, H., Wang, L., Zhao, Z., Shu, C., and Lu, C. 2018. "Support Vector Machine based Differential Pulse-width Pair Brillouin Optical Time Domain Analyzer." **IEEE Photonics Journal**. 10(4) : 1.

Wu, J. G., and Shi, C. H. 2004. "Prediction of grain weight , brown rice weight and amylose content in single rice grains using near-infrared reflectance spectroscopy." **Field Crops Research**. 87 : 13–21.

Wu, J. G., Shi, C. H., and Zhang, X. M. 2002. "Estimating the amino acid composition in the milled rice powder by near-infrared reflectance spectroscopy." **Field crop research**. 75 : 1–7.

Wu, N., Zhang, Y., Na, R., Mi, C., and Zhu, S. 2019. "Variety identification of oat seeds using hyperspectral imaging : investigating the representation ability of deep convolutional neural." The Royal Society of Chemistry. 9 : 12635–12644.

Xiong, Z., Sun, D. W., Zeng, X. A., and Xie, A. 2014. "Recent developments of hyperspectral imaging systems and their applications in detecting quality attributes of red meats: A review." **Journal of Food Engineering**. 132 : 1–13.

Xu, D., Hong, Y., Gu, Z., Cheng, L., Li, Z., and Li, C. 2019. "Effect of high pressure steam

on the eating quality of cooked rice." **Food Science and Technology**. 104 : 100–108.

Yadav, R. B., and Khatkar, B. 2007. "Morphological , physicochemical and cooking properties of some Indian rice (Oryza sativa L .) cultivars." **Journal of Agricultural Technology Morphological**. 3(2) : 203–210.

Yamaguchi, T., Tanaka, Y., Imachi, Y., Yamashita, M., and Katsura, K. 2021. "Feasibility of combining deep learning and RGB images obtained by unmanned aerial vehicle for leaf area index estimation in rice." **Remote sening**. 13(84) : 1–19.

Yan, T., Xu, W., Lin, J., Duan, L., Gao, P., Zhang, C., and Costa, L. 2021. "Combining multi-dimensional convolutional neural network (cnn) with visualization method for detection of aphis gossypii glover infection in cotton leaves using hyperspectral imaging." **Frontiers in Plant Science**. 12 : 1–15.

Yang, S., Zhu, Q. B., Huang, M., and Qin, J. W. 2017. "Hyperspectral Image-Based Variety Discrimination of Maize Seeds by Using a Multi-Model Strategy Coupled with Unsupervised Joint Skewness-Based Wavelength Selection Algorithm." **Food Analytical Methods**. 10(2) : 424–433.

Yang, X., Ye, Y., Li, X., Lau, R. Y. K., and Member, S. 2018. "Hyperspectral image classification with deep learning models." **IEEE Transactions on groscience and remote sening**. 56(9) : 5408–5423.

Ye, J., Liu, C., Luo, S., Wu, J., Hu, X., and McClements, D. J. 2019. "A simulated gastrointestinal tract study of texturized rice grains: Impact of texturization on starch digestibility." **Journal of Cereal Science**. 89 : 102800.

Yu, L., SONG, C., SUN, L., LI, L., XU, Z., and TANG, C. 1. 2020. "Effects of light-emitting diodes on tissue culture plantlets and seedlings of rice (Oryza sativa L.)." **Food Control**. 19(7) : 1743–1754.

Zareef, M., Chen, Q., Hassan, M. M., Arslan, M., Hashim, M. M., Ahmad, W., Kutsanedzie, F. Y. H., and Agyekum, A. A. 2020. "An Overview on the Applications of Typical

Non-linear Algorithms Coupled With NIR Spectroscopy in Food Analysis." **Food Engineering Reviews**. 12(2) : 173–190.

Zhang, C., Liu, F., and He, Y. 2018. "Identification of coffee bean varieties using hyperspectral imaging: Influence of preprocessing methods and pixel-wise spectra analysis." **Scientific Reports.** 8(1) : 1–11.

Zhang, C., Zhao, Y., Yan, T., Bai, X., Xiao, Q., Gao, P., Li, M., Huang, W., Bao, Y., He, Y., and Liu, F. 2020. "Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning." **Infrared Physics and Technology**. 111 : 103550.

Zhang, L., and Ji, H. 2019. "Identification of wheat grain in different states based on hyperspectral imaging technology." **Spectroscopy Letters**. 52(6) : 356-366.

Zhang, Y., Gao, J., and Zhou, H. 2020. "Breeds Classification with Deep Convolutional Neural Network." **ACM International Conference Proceeding Series**. 145–151.

Zhou, V. 2019. **A Simple Explanation of the Softmax Function**. [Online]. Available : https://victorzhou.com/blog/softmax/

Zhu, L., Bi, S., Wu, G., Zhang, H., Wang, L., and Qian, H. 2020. "Comparative analysis of the texture and physicochemical properties of cooked rice based on adjustable rice cooker." **Food Science and Technology**. 130 : 109650.

Zhu, L., Bi, S., Wu, G., Zhang, H., Wang, L., and Qian, H. 2020. "Comparative analysis of the texture and physicochemical properties of cooked rice based on adjustable rice cooker." **Food Science and Technology**. 130 : 109650.

Zou, H., Hastie, T., Tibshirani, R., and Url, S. 2014. "Sparse Principal Component Analysis." **Journal of Computational and Graphical Statistics**. 15(2) : 265–286.

Appendix

Published papers

International Published papers:

1. Onmankhong, J., Jongyingcharoen, J. S., and Sirisomboon, P. 2020. "The influence of processing parameters of parboiled rice on its physiochemical and texture properties." **Journal of Texture Studies**. 52(2) : 219-227.

2. Onmankhong, J., and Sirisomboon, P. 2021. "Texture evaluation of cooked parboiled rice using nondestructive milled whole grain near infrared spectroscopy." **Journal of Cereal Science**. 97 : 103151.

Oral presentation:

1. Onmankhong, J., and Sirisomboon, P. 2018 "Near Infrared Scanning Precision Test for Texture Characteristics of Parboiled Rice." 144-147. The 11th Thai Society of Agricultural Engineering International Conference, Chulabhorn International Convention Center (Wora Wana Hua Hin Hotel & Convention) Hua Hin, Prachuap Khiri Khan, Thailand during 26-27 April 2018.

Author biography

Name-Surname	Miss Jiraporn Onmankhong
Date of birth	25 April 1993
Contact address	137/1 M.9 Banna, Mueang district, Chumphon 86190
Email address	Jiraporn.j1088@gmail.com, 60601172@kmitl.ac.th
Contact number	095-7089409

Education background

D.Eng. (Agricultural Engineering) from King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand (2017-2021, GPA: 4.00)

M.Eng. (Agricultural Engineering) from King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand (2015-2016, GPA: 3.81)

B.Eng. (Agricultural Engineering, 2nd class honor) from King Mongkut's Institute of Technology Ladkrabang, Chumphon Campus, Thailand (2011-2014, GPA: 3.25)

Training attended:

- Visiting research in Graduate School of Bio agricultural Sciences, Nagoya University, Japan (5 months 7 days, 2020)
- Joining Sakura Science Exchange Program of Japan Science and Technology Agency, Nagoya University, Japan (7 days, 2016)
- Summer training program at PTT Tank Terminal Co., LTD, Rayong, Thailand (2 month, 2013)

Scholarships:

- Royal Golden Jubilee scholarship (RGJ) PhD program from National Research Council of Thailand (NRCT) (Grant numbers PHD/0013/2560)
- KMITL Education Fund 2012

Oral presentation

- **Onmankhong, J.**, & Sirisomboon, P. (2018) "Near Infrared Scanning Precision Test for Texture Characteristics of Parboiled Rice" 11th Thai Society of Agricultural Engineering International Conference (TSAE 2018), Thailand

- **Onmankhong, J.**, & Sirisomboon, P. (2016) “Feasibility study of evaluation of ammonium laurate soap content in natural rubber latex by near infrared spectroscopy” 9th TSAE International Conference (TSAE 2016), Thailand

International Published papers:

- **Onmankhong, J.**, Jongyingcharoen, J. S., & Sirisomboon, P. (2021). The influence of processing parameters of parboiled rice on its physiochemical and texture properties. Journal of Texture Studies, 52(2), 219-227
- **Onmankhong, J.**, & Sirisomboon, P. (2021). Texture evaluation of cooked parboiled rice using nondestructive milled whole grain near infrared spectroscopy. Journal of Cereal Science, 97, 103151.

Areas of expertise and interest

- Application in near infrared spectroscopy
- Application in near infrared hyperspectral image
- Texture properties of agricultural products and food
- Non-destructive testing technique for agricultural products and food

Reference persons

- Prof. Dr. Panmanas Sirisomboon, Department of Agricultural Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Thailand Email: panmanas.si@kmitl.ac.th