บทที่ 3
วิธีดำเนินการวิจัย

1. วัตถุดิบ สารเคมี และอุปกรณ์

1.1 วัตถุดิบ

ตัวอย่างข้าว จำนวน 60 ตัวอย่าง ประกอบด้วยข้าวเปลือกหอมมะลิ จำนวน 50 ตัวอย่าง ข้าวกลุ่มปริมาณแอมิโลสต่ำ 3 ตัวอย่าง ได้แก่ ข้าวพันธุ์หอมคลองหลวง พันธุ์ปทุมธานี 1 และพันธุ์พิษณุโลก 1 ข้าวกลุ่มปริมาณแอมิโลสปานกลาง จำนวน 2 ตัวอย่าง ได้แก่ ข้าวพันธุ์สุพรรณบุรี 2 และ สุพรรณบุรี 60 และข้าวกลุ่มแอมิโลสสูง จำนวน 5 ตัวอย่าง ประกอบด้วย ข้าวพันธุ์ชัยนาท 1 พันธุ์ชัยนาท 2 พันธุ์เหลืองประทิว 123 พันธุ์พิษณุโลก 2 และพันธุ์ปทุมธานี 60 โดยตัวอย่างข้าวทั้งหมดเพาะปลูกในปี 2552 จากศูนย์วิจัยข้าว และศูนย์เมล็ดพันธุ์ข้าวต่าง ๆ ในประเทศไทย แสดงดังตารางที่ 3.1

ตารางที่ 3.1 ตัวอย่างข้าวที่เก็บรวบรวมสำหรับใช้ในการทดลอง

กลุ่มข้าว	พันธุ์ข้าว	อักษรย่อ	แหล่งเพาะปลูก	จำนวนตัวอย่าง
ข้าวหอมมะลิ	1. ข้าวดอกมะลิ 105	1. KDML105	ศูนย์วิจัยข้าว และศูนย์เมล็ดพันธุ์ข้าว	39
	2. กข15	2. RD15	ในภาคตะวันออกเฉียงเหนือ ภาคเหนือ	11
			ตอนล่างและภาคกลาง	
ปริมาณแอมิโลสต่ำ	1. หอมคลองหลวง	1. HKLG	1. ศูนย์วิจัยข้าวพิษณุโลก	1
	2. ปทุมธานี 1	2. PTN1	2. ศูนย์วิจัยข้าวปทุมธานี	1
	3. พิษณุโลก 1	3. PL1	3. ศูนย์วิจัยข้าวพิษณุโลก	1
ปริมาณแอมิโลส	1. สุพรรณบุรี 2	1. SPR2	1. ศูนย์วิจัยข้าวสุพรรณบุรี	1
ปานกลาง	2. สุพรรณบุรี 60	2. SPR60	2. ศูนย์วิจัยข้าวสุพรรณบุรี	1
ปริมาณแอมิโลสสูง	1. ชัยนาท 1	1. CNT1	1. ศูนย์วิจัยข้าวชัยนาท	1
	2. ชัยนาท 2	2. CNT2	2. ศูนย์วิจัยข้าวชัยนาท	1
	3. เหลืองประทิว 123	3. LPT123	3. ศูนย์วิจัยข้าวพิษณุโลก	1
	4. พิษณุโลก 2	4. PL2	4. ศูนย์วิจัยข้าวพิษณุโลก	1
	5. ปทุมธานี 60	5. PTN60	5. ศูนย์เมล็ดพันธุ์ข้าวนครสวรรค์	1
			จำนวนตัวอย่างทั้งหมด	60

จากตารางที่ 3.1 จะพบว่าตัวอย่างข้าวเปลือกหอมมะลิซึ่งประกอบด้วยข้าวพันธุ์ขาวดอกมะลิ 105 และข้าวพันธุ์กข 15 จะมีจำนวนตัวอย่างมากที่สุดนั้น เนื่องจากในการวิจัยครั้งนี้ต้องการศึกษาความเป็นไปในการสร้างสมการสำหรับทำนายการจำแนกข้าวหอมมะลิออกข้าวกลุ่มที่มีปริมาณแอมิโลสต่ำ ข้าวกลุ่มที่มีปริมาณแอมิโลสปานกลาง และข้าวกลุ่มที่มีปริมาณแอมิโลสสูง คณะวิจัยจึงได้เก็บรวบรวมตัวอย่างข้าวหอมมะลิจากศูนย์วิจัยข้าว และศูนย์เมล็ดข้าวต่าง ๆ ในประเทศไทย ได้แก่ในเขตภาคตะวันออกเฉียงเหนือ ภาคเหนือตอนล่าง และภาคกลาง เพื่อให้ตัวอย่างข้าวหอมมะลิมีความหลากหลาย นอกจากนี้การรวบรวมตัวอย่างของข้าวกลุ่มอื่น ๆ จะช่วยขยายช่วง (Range) คุณสมบัติทางเคมีกายภาพของข้าวกล้อง เมื่อนำมาใช้ในการสร้างสมการสำหรับทำนายคุณสมบัติทางเคมีกายภาพของข้าวกล้องได้อีกด้วย

1.2 สารเคมี

- 1.2.1 เอทานอล 95% (Merck, Germany)
- 1.2.2 กรดอะซิติค (1.0 N) (Merck, Germany)
- 1.2.3 กรดไฮโดรคลอริก (Merck, Germany)
- 1.2.4 โพแทสเซียมไอโอไดด์ (0.2%) (Ajax, Australia)
- 1.2.5 โพแทสเซียมไฮดรอกไซด์ (2.0%) (Ajax, Australia)
- 1.2.6 โซเดียมไฮดรอกไซด์ (1.0 N และ 0.09 N) (Merck, Germany)
- 1.2.7 โพเทโทแอมิโลสมาตรฐาน (Fluka, Switzerland)
- 1.2.8 แอมิโลเพกทินมาตรฐาน (Fluka, Switzerland)
- 1.2.9 Ethylene diamine tetraacetic acid (EDTA) (Merck, Germany)

1.3 เครื่องมือที่ใช้ในการวิจัย

1.3.1	เครื่องกะเทาะเปลือก	Satake, Model SB	ญี่ปุ่น
1.3.2	เครื่องขัดสี	Satake, Model SKB	ญี่ปุ่น
1.3.3	เครื่องคัดขนาดเมล็ดข้าว	Satake, TRG05A	ญี่ปุ่น
1.3.4	เครื่องวัดค่าการดูดกลืนแสง	Shimadzu, UV-1601	ญี่ปุ่น
1.3.5	เครื่องสเปคโตรสโคปีอินฟราเรดย่านใกล้ (FT-NIR spectroscopy)	Buchi NIRLab N-200	สวิตเซอร์แลนด์
1.3.6	เครื่องวิเคราะห์ไนโตรเจนแบบดูมาส (Nitrogen combustion analyzer)	Leco model FP-2000	อเมริกา

1.3.7 เครื่องวัดความหนืดอย่างรวดเร็ว (Rapid visco analyzer; RVA)

		Model 3-D	ออสเตรเลีย
1.3.8	ชุดสกัดไขมัน	Foss, model 1043	สวีเดน
1.3.9	เครื่องชั่งชนิดละเอียด	Mettler, AE 50	ญี่ปุ่น
1.3.10	ตู้อบลมร้อน (Hot air oven)	Memmert	เยอรมัน
1.3.11	เครื่องบด (Udy Miller)	Udy, Cyclotec	อเมริกา
1.3.12	โถดูดความชื้น (Air-tight desicator)	Memmert	เยอรมันนี
1.3.13	เครื่องแก้วพื้นฐาน		
1.3.14	เครื่องทำความเย็น	Sanyo	ญี่ปุ่น

2. วิธีดำเนินการวิจัย

2.1 การรวบรวมตัวอย่างข้าว

ข้าวเปลือกที่ใช้ในการทดลองเป็นข้าวเปลือกที่เพาะปลูกในปี 2552 รวบรวมจากศูนย์วิจัยข้าว และศูนย์เมล็ดข้าวต่าง ๆ ในประเทศไทย ปริมาณความชื้นข้าวเปลือกเริ่มต้น 13-14% ประกอบด้วยข้าวหอมมะลิ ข้าวกลุ่มที่มีปริมาณแอมิโลสต่ำ ข้าวกลุ่มที่มีแอมิโลสปานกลาง และข้าวกลุ่มที่มีแอมิโลสสูง ข้าวเปลือกแต่ละตัวอย่างจะถูกนำมาทำความสะอาดด้วยเครื่องทำความสะอาด (Rice pre-cleaner equipment) เพื่อกำจัดข้าวเมล็ดลีบ เศษฟาง เศษหญ้า และสิ่งเจือปนต่าง ๆ จากนั้นนำตัวอย่างข้าวเปลือกมาบรรจุในถุงพลาสติกชนิดโพลีโพรพิลีน (Polypropylene) ความหนา 70 ไมโครเมตร ถุงละ 5 กิโลกรัม ปิดปากถุงให้สนิทด้วยเครื่องปิดผนึกพลาสติก นำถุงพลาสติกบรรจุข้าวเปลือกไปเก็บรักษาที่เครื่องทำความเย็นอุณหภูมิประมาณ $10^{o}C$ จนกว่าจะนำมาทดลอง

2.2 การเตรียมวัตถุดิบข้าวกล้องขนาดเต็มเมล็ด

นำตัวอย่างข้าวเปลือกแต่ละตัวอย่าง มาตั้งทิ้งไว้ที่อุณหภูมิห้อง นาน 24 ชั่วโมง เพื่อปรับความชื้นให้สมดุลย์ จากนั้นจึงนำข้าวเปลือกมากะเทาะเปลือกจะได้ข้าวกล้อง นำข้าวกล้องมาคัดแยกข้าวหักและปลายข้าวออกด้วยครื่องคัดขนาดเมล็ดข้าวแบบตะแกรงกลมจนได้เฉพาะข้าวกล้องขนาดเต็มเมล็ด นำข้าวกล้องขนาดเต็มเมล็ดที่เตรียมได้แต่ละตัวอย่างบรรจุแบบสุญญากาศในซองลามิเนตใสชนิด NY/LLDPE มาชั่งน้ำหนักตัวอย่างละ 5 กิโลกรัม จากนั้นมาบรรจุแบบสุญญากาศ และนำมาเก็บรักษาที่อุณหภูมิห้อง จนกว่าจะนำมาวัดค่าการดูดกลืนแสงด้วยเครื่องสเปคโตรสโคปีอินฟราเรดย่านใกล้

2.3 การวัดค่าการดูดกลืนแสงด้วยเทคนิคสเปคโตรสโคปีอินฟราเรดย่านใกล้

นำตัวอย่างข้าวกล้องขนาดเต็มเมล็ดทั้ง 60 ตัวอย่าง ที่เตรียมไว้ในข้อ 2.2 มาชั่งน้ำหนักในจานเพาะเชื้อ (Petri dish) ประมาณ 80 กรัม และวัดค่าการดูดกลืนแสง และวัดค่าการดูดกลืนแสงด้วยเครื่องสเปคโตรสโคปีอินฟราเรดย่านใกล้ (FT-NIR spectroscopy, Buchi NIRLab N-200) ช่วงจำนวนคลื่นที่ 10,000-4,000 cm^{-1} (ความยาวคลื่น 1,100-2,500 นาโนเมตร) ทำการวัดด้วยหลักการสะท้อนแสง (Reflectance) ที่อุณหภูมิ 25^{o}C โดยเริ่มที่ความยาวคลื่นต่ำที่สุดและเพิ่มขึ้นช่วงละ 2 นาโนเมตร ของแต่ละหน่วยตัวอย่าง ค่าการดูดกลืนแสงที่วัดได้จำนวน 701 ค่า จะถูกเก็บข้อมูลไว้ในรูปของลอการิทึม (log 1/reflectance) และบันทึกด้วยโปรแกรม NIRCal รุ่น 5.21 โดยทำการทดลองตัวอย่างละ 2 ซ้ำ

2.4 การวิเคราะห์คุณสมบัติทางเคมีกายภาพของข้าวกล้อง

นำตัวอย่างข้าวกล้องแต่ละตัวอย่างที่ผ่านการวัดค่าการดูดกลืนแสงด้วยเครื่องสเปคโตรสโคปีอินฟราเรดย่านใกล้มาบดด้วยเครื่องบด ขนาด 100 เมช (Mesh) ร่อนผ่านตะแกรง และนำมาบรรจุในถุงอลูมิเนียมฟอยด์ เก็บรักษาที่อุณหภูมิห้อง เพื่อนำมาใช้ในการวิเคราะห์คุณสมบัติทางเคมีกายภาพของข้าวกล้อง ซึ่งจะทำการวิเคราะห์คุณสมบัติทางเคมีกายภาพที่สำคัญ ได้แก่

2.4.1 ปริมาณแอมิโลส ด้วยวิธี AACC (1999) แสดงดังภาคผนวก ก

2.4.2 ปริมาณโปรตีนด้วยวิธีดูมาส (Nitrogen combustion method) ด้วยวิธี AACC (2000a) แสดงดังภาคผนวก ข

2.4.3 ค่าความคงตัวของแป้งสุก ตามวิธีของ Cagampang, Perez & Juliano (1973) แสดงดังภาคผนวก ค

2.4.4 ค่าการสลายเมล็ดข้าวในด่าง ตามวิธีของ Little, Hilder & Dawan (1958) แสดงดังภาคผนวก ง

2.4.5 คุณสมบัติด้านความหนืดของข้าวด้วยเครื่องวัดความหนืดอย่างรวดเร็วด้วยวิธี AACC (2000b) แสดงดังภาคผนวก จ

2.5 การวิเคราะห์ข้อมูลและการสร้างสมการที่เหมาะสมในการทำนายคุณสมบัติทางเคมีกายภาพของข้าวกล้อง

2.5.1 การวิเคราะห์ข้อมูล

คุณสมบัติทางเคมีกายภาพของข้าวกล้องแต่ละพารามิเตอร์ จะนำมาคำนวณหาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (ค่าเฉลี่ย $\pm$ ส่วนเบี่ยงเบนมาตรฐาน) นำข้อมูลที่ได้จากการทดลองมาวิเคราะห์ความแปรปรวนโดยใช้โปรแกรมสำเร็จรูป SPSS version 11.0 และ

เปรียบเทียบความแตกต่างของค่าเฉลี่ยด้วยวิธี Duncan's New Multiple Range Test (DMRT) ที่ระดับความเชื่อมั่น 95%

2.5.2 การสร้างสมการที่เหมาะสมในการทำนายคุณสมบัติทางเคมีกายภาพของข้าวกล้อง

2.5.2.1 การตรวจสอบข้อมูลผิดปกติ (Outlier)

ข้อมูลค่าทางเคมีของคุณสมบัติทางเคมีกายภาพของข้าวกล้องจะถูกนำมาตรวจสอบหาข้อมูลที่ผิดปกติ ซึ่งอาจหาตัวอย่างที่มีค่าทางเคมีสูงหรือต่ำจากการกระจายค่าทางเคมีเป็นแบบปกติ (Normal distribution) โดยจะใช้ค่าทางสถิติคะแนน t (t-score) โดยแปลงค่าทางเคมีเป็นค่าทางสถิติ t ตามวิธีของ Sirisomboon, Tanaka, Fujita & Kojima (2007) และ ฤทธิรณ (2552a) ดังสมการ

$$t_i = \frac{x_i - \bar{x}}{SD}$$

เมื่อ t_i คือ ค่าทางสถิติ t ของตัวอย่าง i

x_i คือ ค่าทางเคมีของตัวอย่าง i

$\bar{x}$ คือ ค่าเฉลี่ยทางเคมี

SD คือ ค่าเบี่ยงเบนมาตรฐานของค่าทางเคมี

หากค่าสัมบูรณ์ของค่าทางสถิติคะแนน t ของตัวอย่างใดมีค่ามากกว่าหรือเท่ากับ 3 จะหมายถึงตัวอย่างนั้นอยู่นอกกลุ่มประชากรตัวอย่างที่สนใจ ที่ระดับความเชื่อมั่น 95% ให้พิจารณาตัดตัวอย่างนั้นทิ้ง

2.5.2.2 การปรับแต่งสเปคตรัม

นำข้อมูลสเปคตรัมอินฟราเรดย่านใกล้ (NIR spectrum) ที่บันทึกไว้ด้วยโปรแกรม NIRCal ไปวิเคราะห์ข้อมูลด้วยโปรแกรม Unscrambler รุ่น X 10.1 (CAMO software, ASA, Norway) ด้วยไฟล์ข้อมูลในรูปแบบ JCAMP-DX ทำการเฉลี่ยสเปคตรัมอินฟราเรดย่านใกล้ จากนั้นนำสเปคตรัมอินฟราเรดย่านใกล้ที่หาค่าเฉลี่ยแล้วมาปรับแต่งสเปคตรัมก่อนการวิเคราะห์ด้วยวิธีการแปลงค่าอนุพันธ์เป็นอันดับสองโดยใช้วิธีการปรับเรียบแบบซาวิตซ์กีโกเลย์ทุกช่วงความยาวคลื่น 10 นาโนเมตร (Savitzky-Golay algorithm : left and right side in 10 nm gap size of the second derivative) วิธีการปรับแก้การกระเจิงแบบผลคูณ (MSC) และวิธีการปรับความแปรปรวนให้เป็นมาตรฐาน (SNV) นอกจากนี้ยังปรับแต่งสเปคตรัมอินฟราเรดย่านใกล้ด้วยวิธีการปรับแก้การกระเจิงแบบผลคูณร่วมกับวิธีการแปลงค่าอนุพันธ์เป็น

อันดับสองโดยใช้วิธีการปรับเรียบแบบซาวิตซ์กีโกเลย์ทุกช่วงความยาวคลื่น 10 นาโนเมตร (MSC + Sativzky-Golay second derivative : 10 nm averaging for left and right side) และวิธีการปรับความแปรปรวนให้เป็นมาตรฐานร่วมกับวิธีการแปลงค่าอนุพันธ์เป็นอันดับสองโดยใช้วิธีการปรับเรียบแบบซาวิตซ์กีโกเลย์ทุกช่วงความยาวคลื่น 10 นาโนเมตร (SNV + Sativzky-Golay second derivative : 10 nm averaging for left and right side) อีกด้วย

2.5.2.3 การสร้างสมการเพื่อทำนายที่เหมาะสมในการทำนายคุณสมบัติทางเคมีภายภาพของข้าวกล้องขนาดเต็มเมล็ดด้วยวิธีการถดถอยกำลังสองน้อยที่สุดบางส่วน หรือ PLSR

การสร้างสมการ PLSR เพื่อทำนายคุณสมบัติทางเคมีกายภาพของข้าวกล้องขนาดกล้องเต็มเมล็ด ได้แก่ ปริมาณแอมิโลส ปริมาณโปรตีน ค่าความคงตัวของเจล ค่าการสลายเมล็ดข้าวในสารละลายด่าง และคุณสมบัติด้านความหนืด จำนวน 120 ตัวอย่างจากการวิเคราะห์ จำนวน 2 ซ้ำนั้น ได้กำหนดให้ค่าเคมีของคุณสมบัติทางเคมีกายภาพของข้าวกล้องเป็นตัวแปรตาม (ตัวแปร Y) และสเปคตรัมอินฟราเรดย่านใกล้เป็นตัวแปรอิสระ (ตัวแปร X) จากนั้นให้เรียงลำดับค่าทางเคมีหรือตัวแปรตาม (ตัวแปร Y) จากค่าน้อยไปมาก และแบ่งข้อมูลออกเป็น 2 กลุ่ม ได้แก่ กลุ่มสร้างสมการ (Calibration) และกลุ่มตรวจสอบความถูกต้อง (Validation) ด้วยอัตราส่วนตัวอย่างข้าวกล้องกลุ่มสร้างสมการต่อตัวอย่างข้าวกล้องกลุ่มตรวจสอบความถูกต้องเท่ากับ 2:1 (หมายถึงกำหนดให้ 2 ตัวอย่างแรกอยู่ในกลุ่มสร้างสมการ และตัวอย่างที่ 3 อยู่ในกลุ่มตรวจสอบความถูกต้อง ตัวอย่างที่ 4-5 อยู่ในกลุ่มสร้างสมการ และตัวอย่างที่ 6 อยู่ในกลุ่มตรวจสอบความถูกต้อง กำหนดอย่างนี้ไปเรื่อยจนครบจำนวนตัวอย่าง) หลังจากนั้นทำการตรวจสอบสเปคตรัมอินฟราเรดย่านใกล้ที่ผิดปกติ (Outlier) ด้วยวิธีการวิเคราะห์องค์ประกอบหลัก หรือ PCA หากตัวอย่างอยู่นอกกลุ่มตัวอย่างส่วนใหญ่ให้พิจารณาตัดสเปคตรัมอินฟราเรดย่านใกล้นั้นทิ้ง

เมื่อตรวจสอบข้อมูลที่ผิดปกติทั้งจากค่าทางเคมีที่วิเคราะห์ด้วยวิธีมาตรฐานและสเปคตรัมอินฟราเรดย่านใกล้แล้ว นำข้อมูลกลุ่มตัวแปรอิสระ (ตัวแปร X) และตัวแปรตาม (ตัวแปร Y) มาหาความสัมพันธ์ด้วยการสร้างสมการโดยใช้วิธี PLSR ที่ช่วงความยาวคลื่น 1,100-2,500 นาโนเมตร และทดสอบสมการด้วยวิธีการทดสอบความแม่นยำภายในกลุ่มทั้งหมดหรือฟูลครอสแวลิเดชัน (Full cross validation) โดยใช้ตัวอย่างทั้งหมด เพื่อตรวจสอบการปรับแต่งสเปคตรัม และช่วงความยาวคลื่นที่เหมาะสมในการสร้างสมการแคลิเบรชัน ทำการพิจารณาความเหมาะสมของสมการ PLSR ที่สร้างขึ้นจากค่าสัมประสิทธิ์สหสัมพันธ์ (R) และค่า

ความคลาดเคลื่อนของสมการ PLSR (SEC) โดยสมการที่ดีต้องมีค่าสัมประสิทธ์สหสัมพันธ์สูงและค่าความคลาดเคลื่อนของสมการ PLSR ต่ำ (Williams & Norris, 2001)

การตรวจสอบความถูกต้องของสมการที่สร้างขึ้น ทำได้โดยนำค่าที่วิเคราะห์ได้จริงทางเคมีของตัวอย่างกลุ่มตรวจสอบความถูกต้อง (Validation) มาทำการเปรียบเทียบค่าทางเคมีที่ได้จากการทำนายด้วยสมการ PLSR ที่สร้างขึ้นจากสเปคตรัมอินฟราเรดย่านใกล้ (Predicted value) กับค่าทางเคมีที่วิเคราะห์ได้จริง (Actual value) การตรวจสอบความถูกต้องของสมการจะพิจารณาจากค่าความคลาดเคลื่อนจากการทำนายด้วยตัวอย่างกลุ่มตรวจสอบความถูกต้อง (Standard Error of Prediction; SEP) ค่าความแตกต่าง (Bias) ระหว่างค่าที่ทำนายได้จากสมการกับค่าทางเคมีที่วิเคราะห์ได้จริงของแต่ละสมการ และค่า RPD (Ratio of standard error of Performance to standard Deviation) ซึ่งเป็นสัดส่วนระหว่างค่าเบี่ยงเบนมาตรฐานของค่าทางเคมีและค่า SEP ของตัวอย่างในกลุ่มที่ใช้ในการทดสอบความแม่นยำของสมการแคลิเบรชันและค่าสัมประสิทธิ์สหสัมพันธ์ สมการที่สร้างขึ้นที่มีความแม่นยำควรมีค่า SEP และ Bias ต่ำ แต่ควรมีค่า RPD สูง (William & Norris, 2001; รณฤทธิ์ ฤทธิรณ, 2552a, หน้า 7-1 ถึง 7-16)

2.5.3 การสร้างสมการและการตรวจสอบความแม่นยำของสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิตามคุณสมบัติทางเคมีกายภาพ

การสร้างสมการและการตรวจสอบความแม่นยำของสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิจากข้าวกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิตามคุณสมบัติทางเคมีกายภาพจะนำข้อมูลค่าการดูดกลืนแสงที่ได้จากการวัดค่าด้วยเครื่องสเปคโตรสโคปีอินฟราเรดย่านใกล้หรือสเปคตรัมอินฟราเรดย่านใกล้ จำนวน 120 ตัวอย่าง มาเป็นตัวแปรจำแนกกลุ่ม เพื่อสร้างสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิโดยใช้เทคนิคทางสถิติวิธี PCA และ PLS-DA

2.5.3.1 การสร้างสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิด้วยเทคนิคทางสถิติวิธี PCA

นำสเปคตรัมอินฟราเรดย่านใกล้ที่ผ่านการหาค่าเฉลี่ย การปรับแต่งสเปคตรัม และการตรวจสเปคตรัมที่ผิดปกติแล้ว มาสร้างสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิด้วยเทคนิคทางสถิติวิธี PCA จากนั้นพิจารณาความสัมพันธ์ระหว่างค่าน้ำหนักปัจจัย (PCA loading) และคะแนนปัจจัย (PCA

score) บนแกนองค์ประกอบหลัก (Principal component; PC) โดยค่าน้ำหนักปัจจัยจะใช้ในการอธิบายโครงสร้างของข้อมูลในรูปแบบความสัมพันธ์ของตัวแปร และค่าคะแนนปัจจัยจะอธิบายโครงสร้างของข้อมูลในรูปแบบความสัมพันธ์ของตัวอย่าง ซึ่งจะแสดงถึงความคล้ายหรือความแตกต่างของตัวอย่าง โดยแต่ละตัวอย่างจะมีค่าคะแนน (Score) ในแต่ละ PC ซึ่งค่าเหล่านี้จะแสดงให้เห็นตำแหน่งของตัวอย่างตามแกน PC โดยตัวอย่างที่มีค่าคะแนนใกล้เคียงกันในแกน PC เดียวกัน จะมีความคล้ายกัน ในทางตรงกันข้ามตัวอย่างที่มีค่าคะแนนแตกต่างกันมากจะมีค่าคะแนนของตัวแปรเดิมนั้นแตกต่างกันด้วย (ธงชัย สุวรรณสิชณน์ และปิติพร ฤทธิเรืองเดช, 2552, หน้า 6-1 ถึง 6-33)

2.5.3.2 การสร้างสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิด้วยเทคนิคทางสถิติวิธี PLS-DA

การสร้างสมการทำนายสำหรับแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิด้วยเทคนิค PLS-DA ตามคุณสมบัติทางเคมีกายภาพนั้น จะต้องมีการกำหนดตัวแปรหุ่นให้มีค่าเท่ากับ 1 หรือ 0 ในกรณีนี้สเปคตรัมอินฟราเรดย่านใกล้ของตัวอย่างข้าวกล้องหอมมะลิจะจัดให้อยู่ในกลุ่มที่มีค่าเท่ากับ 1 และสเปคตรัมอินฟราเรดย่านใกล้ของตัวอย่างข้าวกล้องที่ไม่ใช่ข้าวหอมมะลิจะจัดให้อยู่ในกลุ่มที่มีค่าเท่ากับ 0 แสดงดังตารางที่ 3.2 สเปคตรัมอินฟราเรดย่านใกล้ที่ผ่านการหาค่าเฉลี่ย การปรับแต่งสเปคตรัม และการตรวจสเปคตรัมที่ผิดปกติแล้วจะถูกนำมาสร้างสมการ PLS-DA เพื่อทำนายการแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิ การพิจารณาความถูกต้องของการแบ่งกลุ่มจะพิจารณาจากค่าอ้างอิงที่กำหนดขึ้น กล่าวคือ การแบ่งกลุ่มข้าวกล้องหอมมะลิจะถูกต้องเมื่อมีค่าอ้างอิงอยู่ในช่วง 0.5 ถึง 1.5 และในทางตรงกันข้ามถ้าค่าอ้างอิงอยู่ในช่วง -0.5 ถึง 0.5 แสดงว่าตัวอย่างข้าวกล้องกลุ่มนั้นเป็นข้าวกล้องอื่นที่ไม่ใช่ข้าวหอมมะลิ

ตารางที่ 3.2 การกำหนดตัวแปรหุ่นของข้าวกล้องหอมมะลิและข้าวกล้องกลุ่มอื่นที่ไม่ใช่ข้าวหอมมะลิ

	ข้าวหอมมะลิ	**ข้าวแอมิโลสต่ำ**	**ข้าวแอมิโลสปานกลาง**	**ข้าวแอมิโลสสูง**
ข้าวหอมมะลิ	1	0	0	0
ข้าวแอมิโลสต่ำ	0	1	0	0
ข้าวแอมิโลสปานกลาง	0	0	1	0
ข้าวแอมิโลสสูง	0	0	0	1

การสร้างสมการ PLS-DA สำหรับการทำนายแบ่งกลุ่มข้าวกล้องหอมมะลิออกจากตัวอย่างข้าวกล้องกลุ่มอื่นนั้น จะแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม คือ กลุ่มสำหรับสร้างสมการแคลิเบรชัน และกลุ่มตรวจสอบความถูกต้อง ที่ช่วงความยาวคลื่น 1,100-2,500 นาโนเมตร และทดสอบสมการด้วยวิธีการทดสอบความแม่นยำภายในกลุ่มทั้งหมด โดยสมการที่ดีต้องมีค่า RMSE (Root mean square of the standard error) และค่าสัมประสิทธิ์สหสัมพันธ์สูง การตรวจสอบประสิทธิภาพการแบ่งกลุ่ม (Non-error rate; NER หรือ Classification rate) จะสามารถตรวจสอบได้ดังสมการ (Ballabio & Todeschini, 2009, pp. 83-103)

$$\text{ประสิทธิภาพการแบ่งกลุ่ม} = \frac{\sum_{g=1}^{G} n_{gg}}{n}$$

เมื่อ n_{gg} คือ จำนวนตัวอย่างที่จัดอยู่ในกลุ่มที่ถูกต้อง

n คือ จำนวนตัวอย่างทั้งหมด