

บทที่ 4

ผลการทดลองและวิจารณ์

4.1 คุณภาพทางกายภาพและองค์ประกอบทางเคมีของเมล็ดข้าว

4.1.1 คุณภาพทางกายภาพของเมล็ดข้าว

คุณภาพทางกายภาพที่ทำการศึกษาคือ การวัดขนาดและรูปร่าง (grain dimension) ของเมล็ดข้าวสารทั้ง 5 พันธุ์ ได้แก่ ความยาว (length) ความกว้าง (width) ความหนา (thickness) และรูปร่าง (shape) ของเมล็ด พบว่า ขนาดและรูปร่างของข้าวสาร 5 พันธุ์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p \leq 0.05$) โดยเมล็ดข้าวสารพันธุ์ข้าวเจ้าหอมสุพรรณบุรีมีความยาวเฉลี่ยมากที่สุดเท่ากับ 7.56 มิลลิเมตร รองลงมาคือ พันธุ์ปทุมธานี 1, พันธุ์กข15, พันธุ์ขาวดอกมะลิ 105 และพันธุ์ชัยนาท 1 มีความยาวเฉลี่ยเท่ากับ 7.54, 7.51, 7.41 และ 7.36 มิลลิเมตร ตามลำดับ (Table 4.1) สอดคล้องกับงานชิ้น (2547) ที่พบว่า เมล็ดข้าวสารพันธุ์ข้าวเจ้าหอมสุพรรณบุรี, พันธุ์ปทุมธานี 1, พันธุ์กข15, พันธุ์ขาวดอกมะลิ 105 และพันธุ์ชัยนาท 1 มีความยาวเฉลี่ยเท่ากับ 7.5, 7.6, 7.5, 7.4 และ 7.4 มิลลิเมตร ตามลำดับ สำหรับความกว้างของเมล็ดพบว่า พันธุ์ข้าวเจ้าหอมสุพรรณบุรีมีความกว้างเฉลี่ยมากที่สุดเท่ากับ 2.30 มิลลิเมตร รองลงมาคือ พันธุ์ชัยนาท 1, พันธุ์ขาวดอกมะลิ 105, พันธุ์ปทุมธานี 1 และพันธุ์กข15 มีความกว้างเฉลี่ยเท่ากับ 2.16, 2.11, 2.10 และ 2.07 มิลลิเมตร ตามลำดับ และความหนาของเมล็ดพบว่า พันธุ์ข้าวเจ้าหอมสุพรรณบุรีมีความหนาเฉลี่ยมากที่สุดเท่ากับ 1.78 มิลลิเมตร รองลงมาคือ พันธุ์ชัยนาท 1, พันธุ์ปทุมธานี 1, พันธุ์ขาวดอกมะลิ 105 และพันธุ์กข15 มีความหนาเฉลี่ยเท่ากับ 1.76, 1.73, 1.69 และ 1.68 มิลลิเมตร ตามลำดับ เมื่อพิจารณารูปร่างเมล็ดที่จำแนกโดยใช้อัตราส่วนความยาวต่อความกว้าง พบว่า พันธุ์กข15 มีค่าอัตราส่วนความยาวต่อความกว้างเฉลี่ยมากที่สุดเท่ากับ 3.62 รองลงมาคือ พันธุ์ปทุมธานี 1, พันธุ์ขาวดอกมะลิ 105, พันธุ์ชัยนาท 1 และพันธุ์ข้าวเจ้าหอมสุพรรณบุรี มีค่าอัตราส่วนความยาวต่อความกว้างเฉลี่ยเท่ากับ 3.58, 3.51, 3.42 และ 3.29 ตามลำดับ (Table 4.1) ซึ่งเมล็ดข้าวสารแต่ละพันธุ์มีค่าอัตราส่วนความยาวต่อความกว้างเฉลี่ยมากกว่า 3 แสดงว่า ข้าวสารทั้ง 5 พันธุ์มีรูปร่างเรียวยาว (USDA, 1982) สอดคล้องกับ อรอนงค์ (2547) ที่เปรียบเทียบขนาดและรูปร่างของข้าวไทย ข้าวอิตาลี และข้าวญี่ปุ่น พบว่าข้าวไทยอยู่ในเกณฑ์เรียวยาว ขณะที่ข้าวอิตาลีอยู่ในเกณฑ์ปานกลาง และข้าวญี่ปุ่นอยู่ในเกณฑ์ป้อม จาก Table 4.1 เมื่อเปรียบเทียบขนาดและรูปร่าง พบว่า พันธุ์ข้าวเจ้าหอมสุพรรณบุรีมีเมล็ดขนาดรูปร่างเรียวยาวและกว้างที่สุด ขณะที่พันธุ์ปทุมธานี 1, พันธุ์ขาวดอกมะลิ 105 และ

พันธุ์กข15 มีเมล็ดขนาดรูปร่างเรียวยาว ส่วนพันธุ์ชัยนาท 1 มีเมล็ดขนาดรูปร่างเรียวยาวสั้น กัญญา (2547) กล่าวไว้ว่า ขนาดและรูปร่างของเมล็ดข้าวมีความแตกต่างกันขึ้นอยู่กับพันธุ์ข้าว สภาพพื้นที่ปลูก ความชื้นของเมล็ด และการดูแลรักษาทั้งหลังและก่อนการเก็บเกี่ยว ซึ่งจะมีผลต่อการจัดสีของข้าว ทำให้ข้าวสารแต่ละพันธุ์มีขนาดรูปร่างที่แตกต่าง

4.1.2 องค์ประกอบทางเคมีของข้าว

จากการวิเคราะห์องค์ประกอบทางเคมีของข้าว ได้แก่ ปริมาณอมิโลส ปริมาณโปรตีน ปริมาณไขมันโดยรวม และปริมาณความชื้น พบว่าองค์ประกอบทางเคมีแต่ละชนิดในข้าวสาร 5 พันธุ์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p \leq 0.05$) โดยพบปริมาณอมิโลสเฉลี่ยในช่วง 15.38-31.10% ซึ่งพันธุ์ชัยนาท 1 มีปริมาณอมิโลสเฉลี่ยสูงสุดเท่ากับ 31.10% รองลงมาคือ พันธุ์ปทุมธานี 1, พันธุ์ข้าวเจ้าหอมสุพรรณบุรี, พันธุ์ข้าวดอกมะลิ 105 และพันธุ์กข15 ซึ่งมีปริมาณอมิโลสเฉลี่ยเท่ากับ 18.44, 17.65, 15.41 และ 15.38% ตามลำดับ (Table 4.2) สอดคล้องกับรายงานของกรมการข้าว (2550) ที่วิเคราะห์ปริมาณอมิโลสในข้าวสารพันธุ์ชัยนาท 1, พันธุ์ปทุมธานี 1, พันธุ์ข้าวเจ้าหอมสุพรรณบุรี, พันธุ์ข้าวดอกมะลิ 105 และพันธุ์กข15 มีค่าในช่วง 26-27, 15-19, 18-19, 12-17 และ 14-17% ตามลำดับ การแบ่งประเภทพันธุ์ข้าวตามปริมาณอมิโลสในข้าวสารสามารถแบ่งกลุ่มข้าวได้ 3 กลุ่มคือ กลุ่มที่มีปริมาณอมิโลสต่ำอยู่ในช่วง 10-19% กลุ่มที่มีปริมาณอมิโลสปานกลางอยู่ในช่วง 20-25% และกลุ่มที่มีปริมาณอมิโลสสูงอยู่ในช่วง 26-34% (งามชื่น, 2547) จากการเปรียบเทียบกลุ่มดังกล่าวกับผลการวิเคราะห์ที่ได้ใน Table 4.2 พบว่าข้าวทั้ง 5 พันธุ์สามารถแบ่งออกได้เป็น 2 กลุ่มคือ กลุ่มข้าวสารที่มีปริมาณอมิโลสต่ำ ได้แก่ พันธุ์ปทุมธานี 1, พันธุ์ข้าวเจ้าหอมสุพรรณบุรี, พันธุ์ข้าวดอกมะลิ 105 และพันธุ์กข15 และกลุ่มข้าวสารที่มีปริมาณอมิโลสสูง ได้แก่ พันธุ์ชัยนาท 1

เมื่อวิเคราะห์ปริมาณโปรตีนของข้าวสาร 5 พันธุ์ พบว่ามีปริมาณโปรตีนเฉลี่ยในช่วง 7.13-8.43% โดยพันธุ์ข้าวเจ้าหอมสุพรรณบุรีมีปริมาณโปรตีนเฉลี่ยสูงสุดเท่ากับ 8.43% รองลงมาคือ พันธุ์ปทุมธานี 1, พันธุ์กข15, พันธุ์ข้าวดอกมะลิ 105 และพันธุ์ชัยนาท 1 เท่ากับ 7.77, 7.65, 7.50 และ 7.13% ตามลำดับ (Table 4.2) ซึ่งสอดคล้องกับงานวิจัยของณัฐกานต์ (2547) ที่วิเคราะห์ปริมาณโปรตีนในข้าวสาร 6 พันธุ์ พบปริมาณโปรตีนในช่วง 7.53-8.51% เช่นเดียวกับ Hirohata and Chen (1959) และ Juliano (1993) รายงานว่า ข้าวสารมีปริมาณโปรตีนในช่วง 6.5-9.6% และ 6.3-7.1% ตามลำดับ ซึ่งปริมาณโปรตีนในข้าวมีความแตกต่างกันขึ้นอยู่กับพันธุ์และสภาพแวดล้อมในการปลูก (อรอนงค์, 2547) ข้าวที่มีปริมาณโปรตีนสูงมีส่วนทำให้ระยะเวลาการหุงต้มนานขึ้น เนื่องจากโปรตีนจะเป็นตัวขัดขวางการซึมของน้ำเข้าภายในเมล็ด นอกจากนี้

ข้าวที่มีปริมาณโปรตีนสูงยังทำให้เมล็ดแกร่งขึ้น เป็นผลให้การขัดสีออกได้ยาก ทำให้ข้าวที่หุงต้มสุกจะเหนียวน้อยลงและมีสีคล้ำกว่าข้าวที่มีโปรตีนต่ำ (งามชื่น, 2547)

ส่วนการวิเคราะห์ปริมาณไขมันโดยรวมในข้าวสาร 5 พันธุ์ พบว่ามีปริมาณไขมันโดยรวมเฉลี่ยในช่วง 0.58-1.65% โดยพันธุ์กข15 มีปริมาณไขมันโดยรวมเฉลี่ยสูงสุดเท่ากับ 1.65% รองลงมาคือ พันธุ์ขาวดอกมะลิ 105, พันธุ์ข้าวเจ้าหอมสุพรรณบุรี, พันธุ์ปทุมธานี 1 และพันธุ์ชัยนาท 1 มีปริมาณไขมันโดยรวมเฉลี่ยเท่ากับ 1.21, 1.19, 0.72 และ 0.58% ตามลำดับ (Table 4.2) Wang *et al.*(2006) ตรวจวัดปริมาณไขมันในข้าวสารของกลุ่มพันธุ์อินดิกา พันธุ์จาปอนิกา และข้าวเหนียว (glutinous rice) ได้ปริมาณไขมันในช่วงที่กว้างมากคือ 0.09-1.52% ขณะที่ Hirohata and Chen (1959) และ Juliano (1993) พบปริมาณไขมันของข้าวสารในช่วง 0.3-1.1% และ 0.3-0.5% ตามลำดับ นอกจากนี้ Juliano (1993) ได้วิเคราะห์องค์ประกอบทางเคมีโดยประมาณในข้าวกล้อง พบว่า ปริมาณ โปรตีน ไขมัน เส้นใย และเถ้าส่วนใหญ่ มีปริมาณมากในชั้นเยื่อหุ้มเนื้อเมล็ด (aleurone layer) มากกว่าในส่วนเนื้อเมล็ด (inner endosperm) ที่ประกอบด้วยแป้งเป็นหลัก ดังนั้นการที่ข้าวสารมีปริมาณโปรตีนและไขมันต่ำกว่าข้าวกล้องเนื่องจากปริมาณโปรตีนและไขมัน ดังกล่าวมีอยู่ในชั้นเยื่อหุ้มเนื้อเมล็ด เมื่อทำการขัดสีข้าวกล้องปริมาณโปรตีนและไขมัน จึงหลุดตามออกไปด้วย สอดคล้องกับ Rittiron (2004) รายงานว่า ระดับการขัดสีของข้าวกล้องญี่ปุ่นที่ต่างกันทำให้ได้ปริมาณโปรตีนและปริมาณไขมันในข้าวสารที่แตกต่างกันด้วย กล่าวคือ เมื่อใช้ระยะในการขัดสีนานขึ้น ปริมาณ โปรตีนและไขมันในข้าวสารลดลง

เมื่อพิจารณาความชื้น พบว่าข้าวทั้ง 5 พันธุ์ มีปริมาณความชื้นเฉลี่ยในช่วง 9.72-10.81% โดยพันธุ์ชัยนาท 1 มีปริมาณความชื้นเฉลี่ยสูงสุดเท่ากับ 10.81% รองลงมาคือ พันธุ์ขาวดอกมะลิ 105, พันธุ์ปทุมธานี 1, พันธุ์ข้าวเจ้าหอมสุพรรณบุรี และพันธุ์กข15 มีปริมาณความชื้นเฉลี่ยเท่ากับ 10.48, 9.96, 9.90 และ 9.72% ตามลำดับ (Table 4.2) ปริมาณความชื้นหรือน้ำที่อยู่ภายในอาหารจะเป็นปัจจัยหนึ่งที่เป็นตัวกำหนดความแข็งของอาหารหรือตัวอย่างได้ เช่น ตัวอย่างที่มีความชื้นต่ำ ตัวอย่างนั้นจะมีความแข็งมาก (ณรงค์, 2538)

Table 4.1 Dimensions of five milled rice cultivars.

Cultivars	Grain milled rice dimension (millimeter)			*Shape
	Length	Width	Thickness	
CNT1	7.36 ^c	2.16 ^b	1.76 ^{ab}	3.42 ^c
PTN1	7.54 ^{ab}	2.10 ^{cd}	1.73 ^b	3.58 ^{ab}
HSP	7.56 ^a	2.30 ^a	1.78 ^a	3.29 ^d
RD15	7.50 ^{abc}	2.07 ^d	1.68 ^c	3.62 ^a
KDML105	7.41 ^{bc}	2.11 ^c	1.69 ^c	3.51 ^b
CV (%)	3.05	2.95	2.75	4.07
LSD _(0.05)	0.14	0.04	0.03	0.09

Means with the different letters in same columns indicated significant difference at 95% confidence.

*Shape = length/width (USDA, 1982)

Table 4.2 Chemical compositions of five milled rice cultivars.

Cultivars	Chemical compositions (% wet basis)			
	Amylose	Protein	Crude lipid	Moisture
CNT1	31.10 ^a	7.13 ^d	0.58 ^d	10.81 ^a
PTN1	18.44 ^b	7.77 ^b	0.72 ^c	9.96 ^c
HSP	17.65 ^c	8.43 ^a	1.19 ^b	9.90 ^c
RD15	15.38 ^d	7.65 ^b	1.65 ^a	9.72 ^d
KDML105	15.41 ^d	7.50 ^c	1.21 ^b	10.48 ^b
Number of samples	300	120	180	200
CV (%)	5.15	2.94	13.88	2.11
LSD _(0.05)	0.36	0.08	0.13	0.07

Means with the different letters in same columns indicated significant difference at 95% confidence.

CNT1 = Chai Nat 1

RD15 = Rice Department 15

PTN1 = Pathum Thani 1

KDML105 = Khao Dawk Mali 105

HSP = Khao' Jow Hawm Suphan Bure

4.2 สเปกตรัมและการแปลงข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์

4.2.1 สเปกตรัมของเมล็ดข้าวสาร 5 พันธุ์

การวัดค่าการดูดกลืนแสง NIR ของข้าวสารทั้ง 5 พันธุ์ ในช่วงความยาวคลื่น 1100-2500 นาโนเมตร พบลักษณะสเปกตรัมดั้งเดิม [original spectrum หรือ $\log(1/R)$] มีการเลื่อนตัว และมีพิทฐานกว้าง (Figure 4.1) อาจเกิดเนื่องจากขนาดและรูปร่างของเมล็ดข้าวสารในแต่ละพันธุ์มีความแตกต่างกัน ซึ่งมีผลต่อการอัดตัวและการกระจายขนาดของเมล็ดข้าวสารในเซลล์ใส่ตัวอย่าง (rotating cup) ที่แตกต่างกัน จึงเกิดการกระเจิงแสงขึ้น ทำให้สเปกตรัมที่ได้มีการเลื่อนตัว (William and Norris, 2001) โดยข้าวสารพันธุ์กข15, พันธุ์ขาวดอกมะลิ 105, พันธุ์ปทุมธานี 1 และพันธุ์ข้าวเจ้าหอมสุพรรณบุรี (กลุ่มข้าวสารที่มีปริมาณอมิโลสต่ำ) มีค่าการดูดกลืนแสงมากกว่าข้าวสารพันธุ์ชัยนาท 1 (กลุ่มข้าวสารที่มีปริมาณอมิโลสสูง) เนื่องจากข้าวสารพันธุ์ชัยนาท 1 มีขนาดและรูปร่างของเมล็ดสั้นและป้อมจะมีการอัดตัวแน่นกว่า จึงเกิดช่องว่างระหว่างเมล็ดน้อย เป็นผลให้แสงส่องผ่านเข้าไปในตัวอย่างไม่ได้น้อย ฉะนั้น ค่าการดูดกลืนแสงจึงมีค่าน้อยกว่าเมล็ดข้าวสารที่มีขนาดและรูปร่างของเมล็ดยาวและเรียวย ในการบดแป้งสาธิตให้มีขนาดอนุภาคที่แตกต่างกัน Osborne *et al.* (1993) พบว่า ค่าการดูดกลืนแสงจะลดต่ำลง เมื่ออนุภาคของตัวอย่างมีขนาดเล็กลง เนื่องจากแป้งที่อนุภาคขนาดเล็กจะดูขาวกว่าแป้งที่มีอนุภาคขนาดใหญ่ จึงมีการสะท้อนแสงกลับมาก ทำให้ค่าการดูดกลืนแสงต่ำ จากหลักการสะท้อนแสง วัตถุที่มีสีใสหรือขาว แสงจะสะท้อนออกจากวัตถุมาก หรือวัตถุดูดกลืนแสงน้อยกว่าวัตถุที่มีสีเข้ม (นิพนธ์, 2545) ในขณะที่ อนุพันธ์ (2545) กล่าวว่า ขนาดอนุภาคที่แตกต่างกันเป็นสาเหตุให้ปริมาณการกระเจิงแสงแตกต่างกันด้วย ซึ่งอนุภาคขนาดเล็กจะมีการกระเจิงแสงมาก ทำให้ระยะทางที่แสงผ่านตัวอย่างสั้นกว่าตัวอย่างที่มีอนุภาคขนาดใหญ่ ค่าการดูดกลืนแสงจึงเกิดขึ้นน้อยตามไปด้วย

จาก Figure 4.1 พบพีกเด่นชัดที่ความยาวคลื่น 1456 และ 1946 นาโนเมตร ที่มีความสัมพันธ์กับโมเลกุลน้ำ (Osborne *et al.*, 1993) สอดคล้องกับ William and Norris (2001) พบว่า สเปกตรัมของน้ำมีการดูดกลืนแสงเด่นชัดที่ความยาวคลื่น 1450 และ 1940 นาโนเมตร เนื่องจากน้ำเกิด weak band ในช่วงคลื่นแสง NIR น้ำจึงเป็นองค์ประกอบทางเคมีที่ดูดกลืนแสงที่มีประสิทธิภาพมาก จึงทำให้มีการการซ้อนทับกับพีกของโมเลกุลอื่น สเปกตรัมที่ได้จึงมีพิทฐานกว้าง เพราะฉะนั้นตัวอย่างใดที่มีความชื้นสูง จะมีสเปกตรัมที่มีการดูดกลืนแสงโดยรวมสูงกว่าตัวอย่างที่มีความชื้นต่ำ อันเป็นผลเนื่องจากน้ำมีการหักเหสูงกว่าอากาศ จึงทำให้ความแตกต่างระหว่างดัชนีหักเหขององค์ประกอบในตัวอย่างกับตัวกลางน้อยลง การหักเหของแสงจึงเกิดขึ้นน้อย ส่งผลให้การกระเจิงของแสงเกิดขึ้นน้อยด้วย แสงจึงผ่านในตัวอย่างไม่ถูกกั้น ทำให้การดูดกลืนแสงจึงมากขึ้น (Hruschka, 2001) จากการศึกษาของ Osborne *et al.* (1993) ที่วัดสเปกตรัมของข้าวสารก่อนอบ

และหลังอบ พบว่าข้าวสาลีหลังอบมีฟีกของน้ำหายไป และค่าการดูดกลืนแสงลดลงทั้งสเปกตรัม เนื่องจากตัวอย่างแห้งจะมีสีน้ำตาลอ่อน จึงสะท้อนแสงได้มากกว่าตัวอย่างเปียกที่มีสีน้ำตาลเข้ม นอกจากนี้ค่าการดูดกลืนแสงของสเปกตรัมที่แตกต่างกัน อาจเกิดจากอุณหภูมิในตัวอย่างที่มีผล โดยตรงต่อโมเลกุลของน้ำ ทำให้สเปกตรัมมีการเปลี่ยนแปลงไป (Kawano *et al.*, 1995)

4.2.2 การแปลงข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์

ลักษณะสเปกตรัมดั้งเดิมที่ได้มีฟีกที่กว้าง และมีการเลื่อนตัวของสเปกตรัม (Figure 4.1) เพราะความแตกต่างของขนาดรูปร่างเมล็ดข้าว การบรรจุข้าวในเซลล์ใส่ตัวอย่าง และความชื้นในตัวอย่าง ทำให้ยากต่อการเลือกความยาวคลื่นที่สัมพันธ์กับค่าปริมาณองค์ประกอบทางเคมีในข้าวสาร การแปลงข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์ สามารถช่วยลดอิทธิพลของปัจจัย ดังที่กล่าวได้ จึงแปลงข้อมูลสเปกตรัมดั้งเดิมด้วยวิธีอนุพันธ์อันดับสอง (second derivative) (Figure 4.2) พบว่า ลักษณะสเปกตรัมมีค่าการดูดกลืนแสงในทิศทางที่ตรงกันข้ามกับสเปกตรัมดั้งเดิมคือ มีลักษณะหวักลบลงมาด้านล่าง เพราะเป็นการหาความชันสองครั้ง เพื่อแยกฟีกออกมาให้ชัดเจนมากขึ้น (Hruschka, 2001) โดยพบฟีกชัดเจนที่ความยาวคลื่น 1200, 1426, 1702 และ 1904 นาโนเมตร ซึ่งแต่ละความยาวคลื่นมีความสัมพันธ์กับพันธะเคมีที่แตกต่างกันคือ ที่ความยาวคลื่น 1200 และ 1702 นาโนเมตร มีความสัมพันธ์กับพันธะ C-H ในไขมันและแป้ง ตามลำดับ (Osborne *et al.*, 1993) ขณะที่ Shenk *et al.* (1977) รายงานว่า ที่ความยาวคลื่น 1702 นาโนเมตร มีความสัมพันธ์กับพันธะ N-H ของปริมาณโปรตีนในอาหารสัตว์ สำหรับความยาวคลื่นที่ตำแหน่ง 1426 และ 1904 นาโนเมตร มีความสัมพันธ์กับพันธะ O-H ของน้ำ (Osborne *et al.*, 1993; William, 2007)

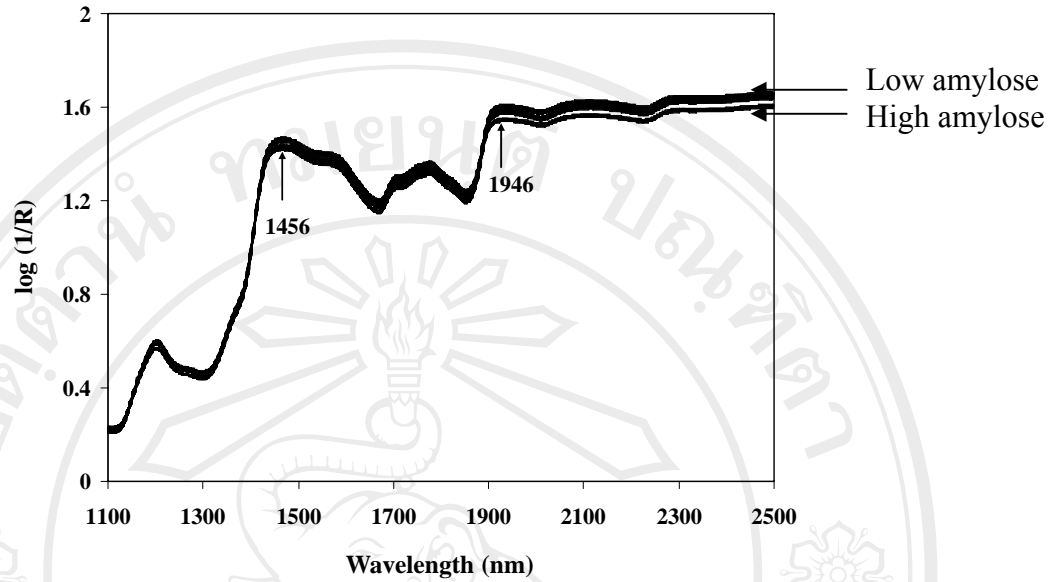


Figure 4.1 The original spectra of 5 milled rice cultivars.

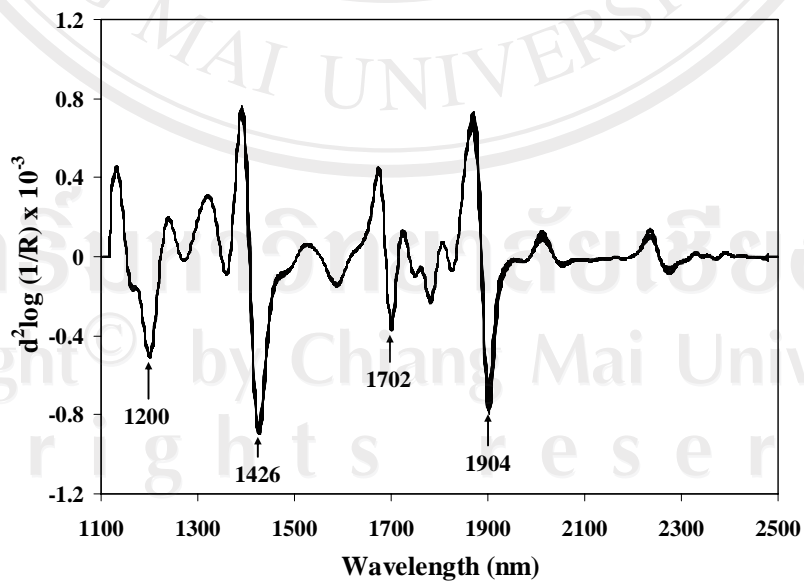
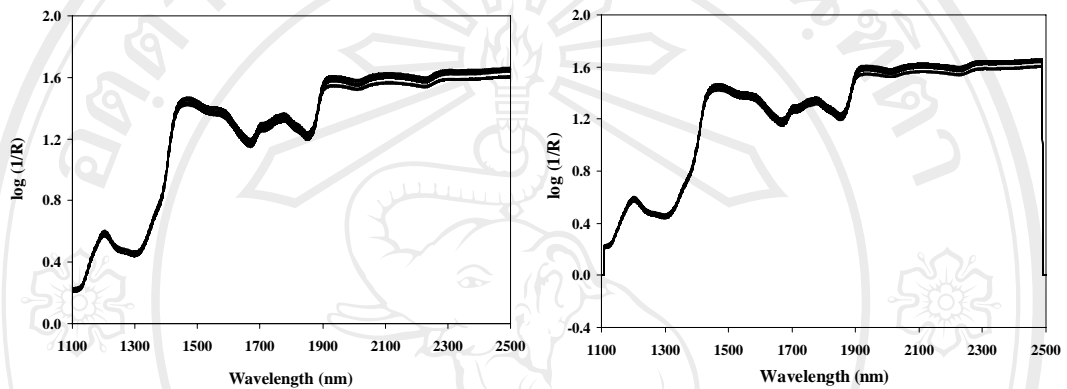


Figure 4.2 Five milled rice spectra treated with Savitzky-Golay second derivative.

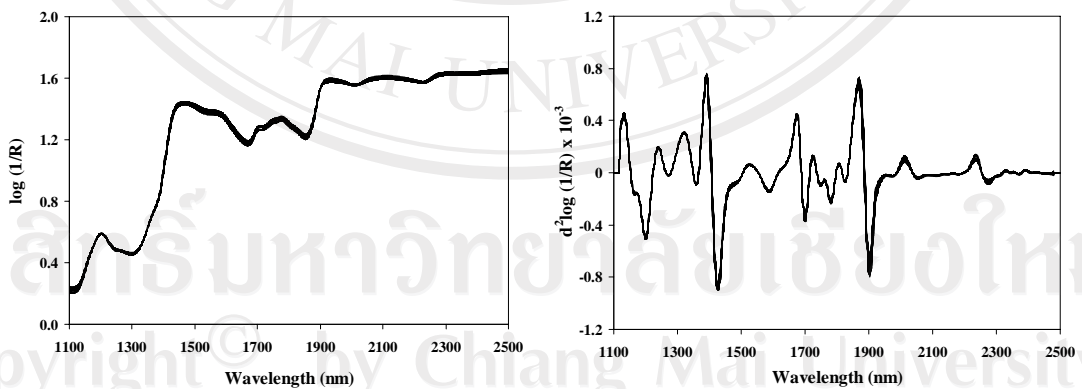
เมื่อแปลงข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์ด้วยวิธีต่างๆ พบว่า เทคนิคทางคณิตศาสตร์ที่สามารถใช้แปลงข้อมูลสเปกตรัมและให้ผลดีมี 3 วิธี คือ Savitzky-Golay second derivative, multiplicative scatter correction (MSC), และ Savitzky-Golay smoothing โดยพิจารณาจาก Figure 4.3 (a) ลักษณะสเปกตรัมดั้งเดิมที่ได้ จะสังเกตเห็นการรบกวนของสัญญาณ (noise) ในช่วงความยาวคลื่น 2050-2500 นาโนเมตร จึงแปลงข้อมูลสเปกตรัมด้วย Savitzky-Golay smoothing [Figure 4.3 (b)] เพื่อลดปัญหาการรบกวนของสัญญาณต่อค่าการดูดกลืนแสง (Katsumoto *et al.*, 2001; Siesler *et al.*, 2002) สเปกตรัมที่ได้มีลักษณะเหมือนสเปกตรัมดั้งเดิมแต่จะเรียบสม่ำเสมอมากกว่า เมื่อแปลงข้อมูลด้วย MSC [Figure 4.3 (c)] พบว่า รูปร่างของสเปกตรัมที่ได้เหมือนกับสเปกตรัมดั้งเดิม แต่เส้นสเปกตรัมในแนวแกน Y เลื่อนมาชิดกัน ส่วน Figure 4.3 (d) แปลงข้อมูลสเปกตรัมด้วย Savitzky-Golay second derivative พบสเปกตรัมลักษณะหวักลบลงมาด้านล่าง และแถบสเปกตรัมเลื่อนมาชิดกัน ทำให้ได้แถบสเปกตรัมแคบลง และการซ้อนทับของพีกแยกกันอย่างชัดเจน สำหรับ Figure 4.3 (e, f) เป็นสเปกตรัมที่แปลงข้อมูลด้วย Savitzky-Golay second derivative ร่วมกับ MSC และ MSC ร่วมกับ Savitzky-Golay second derivative สามารถลดปัญหาการซ้อนทับกันของพีก และการเลื่อนตัวของสเปกตรัม ที่เกิดจากการกระเจิงแสงได้ สอดคล้องกับรายงานของ William and Norris (2001) พบว่า การแปลงข้อมูลสเปกตรัมด้วย MSC และ second derivative สามารถลดอิทธิพลของขนาดตัวอย่างและการกระเจิงแสงได้ สำหรับการแปลงข้อมูลด้วย Savitzky-Golay second derivative, MSC ร่วมกับ Savitzky-Golay smoothing และ MSC, Savitzky-Golay second derivative ร่วมกับ Savitzky-Golay smoothing [Figure 4.3 (g, h)] สามารถลดปัญหาการซ้อนทับกันของพีก การเลื่อนตัวของสเปกตรัม และสัญญาณรบกวนต่อค่าการดูดกลืนแสงได้ โดยลักษณะของสเปกตรัมที่ได้จะคล้ายกับ Figure 4.3 (e, f) แต่จะเรียบและสม่ำเสมอกว่า การแปลงข้อมูลสเปกตรัมด้วย smoothing ที่อันดับสูงเกินไปอาจทำให้ได้พีกสเปกตรัมของ derivative มีลักษณะบิดเบือนไปจากที่ควรจะเป็นได้ (ธีรศักดิ์ และวนิดา, 2550) จะเห็นได้ว่าการแปลงข้อมูลสเปกตรัมแต่ละวิธีร่วมกันไม่ว่าก่อนหรือหลัง ให้ผลของสเปกตรัมและค่าการดูดกลืนแสงไม่ต่างกัน ดังใน Figure 4.3 (e, f) และ Figure 4.3 (g, h) อย่างไรก็ตาม การแปลงข้อมูลสเปกตรัมร่วมกันหลายๆ วิธีก็อาจทำให้ข้อมูลหายไป (Kawano and Saranwong, 2007) ในข้าวปริมาณองค์ประกอบทางเคมีแต่ละค่าประกอบด้วยโมเลกุลที่ต่างกัน การดูดกลืนแสง NIR จึงต่างกันด้วย ดังเช่น พีกของน้ำจะเห็นได้ชัดเจนมาก จึงไม่จำเป็นต้องแปลงข้อมูลด้วยเทคนิคทางคณิตศาสตร์ต่างๆ แต่กรณีของปริมาณมิโลสโปรตีน หรือไขมัน จะถูกรบกวนโดยโมเลกุลของน้ำ ที่มีประสิทธิภาพในการดูดกลืนแสง NIR ได้ดีกว่าองค์ประกอบทางเคมีเหล่านี้ จึงถูกซ้อนทับด้วยพีกของโมเลกุลน้ำ ทำให้เห็นพีกของมิโลส

โปรตีน และไขมันไม่ชัดเจน จึงต้องแปลงข้อมูลด้วยเทคนิคทางคณิตศาสตร์ด้วยวิธีต่างๆ รวมถึงการเลือกช่วงความยาวคลื่นที่เหมาะสม จะทำให้เห็นพีกของอมิโลส โปรตีน หรือไขมันในข้าวได้ชัดเจนยิ่งขึ้น



(a) original spectra

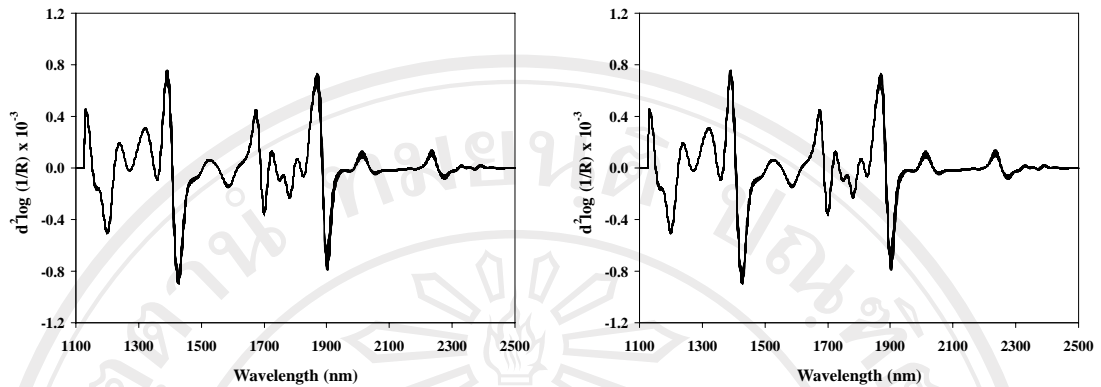
(b) Savitzky-Golay smoothing



(c) MSC

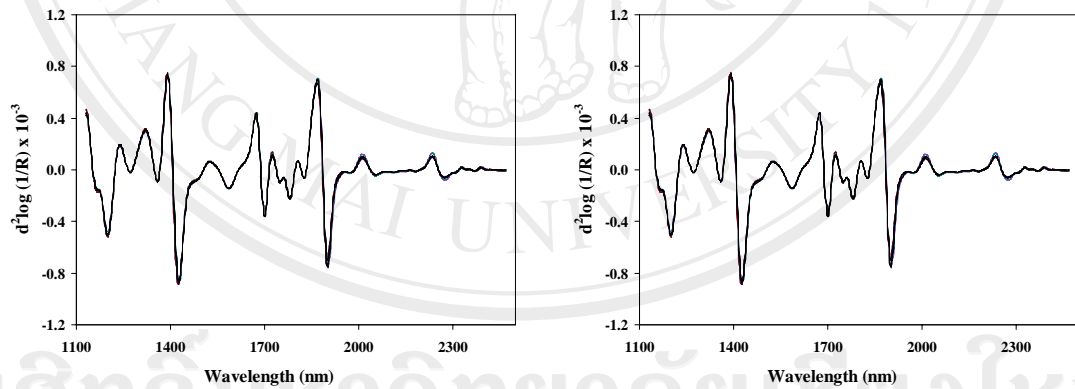
(d) Savitzky-Golay second derivative

Figure 4.3 Five milled rice cultivars spectra treated with mathematical techniques.



(e) Savitzky-Golay second derivative and MSC

(f) MSC and Savitzky-Golay second derivative



(g) Savitzky-Golay second derivative and MSC and Savitzky-Golay smoothing

(h) MSC and Savitzky-Golay second derivative and Savitzky-Golay smoothing

Figure 4.3 Five milled rice cultivars spectra treated with mathematical techniques (continued).

4.3 การสร้างสมการทำนายของค่าประกอบทางเคมีในข้าวสาร

4.3.1 การสร้างสมการทำนายปริมาณอมิโลสในข้าวสาร

การสร้างสมการทำนายปริมาณอมิโลสในข้าวสาร จำนวน 300 ตัวอย่าง โดยนำข้อมูลสเปกตรัม ในช่วงความยาวคลื่น 1100-2500 นาโนเมตร กับข้อมูลปริมาณอมิโลสในข้าวที่ได้จากการวิเคราะห์ทางเคมีมาหาความสัมพันธ์ โดยการสร้างสมการถดถอยเชิงเส้นด้วยเทคนิค PLSR แล้วใช้วิธีทดสอบสมการ (validation method) แบบ test set โดยแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มสร้างสมการทำนายปริมาณองค์ประกอบทางเคมี (calibration set) จำนวน 155 ตัวอย่าง และกลุ่มทดสอบสมการ (validation set) จำนวน 145 ตัวอย่าง ข้อมูลทางสถิติของตัวอย่างทั้งสองกลุ่มแสดงใน Table 4.3 โดยมีปริมาณอมิโลสในช่วง 13.88-33.09%

Table 4.3 Characteristics of calibration and validation sample set of five milled rice cultivars for amylose content determination by NIR.

Variable	Calibration sample set	Validation sample set
Number of samples	155	145
Amylose content (% wet basis)	13.88-33.09	14.03-32.92
Mean	19.62	19.57
Standard deviation (SD)	5.99	5.98

จากการสร้างสมการทำนายปริมาณอมิโลสของข้าวสาร (Table 4.4) เมื่อเปรียบเทียบการแปลงข้อมูลด้วยเทคนิคทางคณิตศาสตร์ พบว่าการแปลงข้อมูลสเปกตรัมทั้ง 4 วิธีให้ความแม่นยำสูง โดยมีค่า R เท่ากับ 0.96, 0.96, 0.97 และ 0.97 ตามลำดับ แต่สมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC ร่วมกับ 2nd derivative และสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC, 2nd derivative ร่วมกับ smoothing ในช่วงความยาวคลื่นที่แสดงใน Table 4.4 มีจำนวนแฟกเตอร์ (F) เท่ากับ 7 แฟกเตอร์ ซึ่งจำนวนแฟกเตอร์ที่เพิ่มขึ้นไม่ทำให้เกิด over fitting เนื่องจากค่า R เพิ่มขึ้น แต่ค่า SEC, SEP และ Bias ลดลงตามกัน เมื่อทดสอบความแตกต่างทางสถิติของค่า SEP ทั้ง 4 สมการ พบว่า สมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative และสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative ร่วมกับ smoothing มีค่า SEP สูงที่แตกต่างกับสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC ร่วมกับ 2nd derivative และสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC, 2nd derivative ร่วมกับ smoothing อย่างมีนัยสำคัญทางสถิติ ($p \leq 0.05$) โดยให้ค่า SEP เท่ากับ 1.95%, 1.97%, 1.81% และ 1.81% ตามลำดับ ซึ่งสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC ร่วมกับ 2nd derivative และสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย

MSC, 2nd derivative ร่วมกับ smoothing ให้ค่า RPD สูงเท่ากัน คือ 3.30 แต่สมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC, 2nd derivative ร่วมกับ smoothing จะให้พิกในกราฟสัมพันธ์การถดถอยที่เกี่ยวข้องกับแป้งชัดเจนกว่าสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC ร่วมกับ 2nd derivative ดังนั้น จึงเลือกเป็นสมการสำหรับใช้ในการทำนายในช่วงความยาวคลื่น 1300-2448 นาโนเมตร โดยให้ค่า F, R, SEC, SEP, Bias และ RPD เท่ากับ 7, 0.97, 1.57%, 1.81%, 0.09% และ 3.30 ตามลำดับ (Table 4.4) จากเกณฑ์การพิจารณาค่า RPD ของ William (2007) แสดงค่า RPD ที่ช่วง 3.1-4.9 สามารถแยกผลผลิตได้ ดังนั้นจะเห็นว่าสมการทำนายที่เลือกอาจจะใช้แบ่งแยกปริมาณอมิโลสในข้าวสารได้

Table 4.4 PLSR calibration results for amylose content using spectra pretreatments.

Pre-treatment	Wavelength region (nm)	F	R	SEC	SEP	Bias	RPD
2 nd derivative	1400-2458	6	0.96	1.75	1.95 ^b	0.21	3.07
2 nd derivative+smoothing	1150-2448	6	0.96	1.76	1.97 ^b	0.21	3.04
MSC+2 nd derivative	1300-2458	7	0.97	1.56	1.81 ^a	0.10	3.30
MSC+2nd derivative+Smoothing	1300-2448	7	0.97	1.57	1.81^a	0.09	3.30

F: number of factors used in the calibration equation.

R: multiple correlation coefficients.

SEC: standard error of calibration.

SEP: standard error of prediction, tested by the method of Fearn (1996).

Bias: average of difference between actual value and NIR value.

RPD: ratio of standard deviation of reference data in validation set to SEP.

Means with the different letters in same column indicate significant difference at 95% confidence.

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอย (regression coefficient) ของสมการทำนายปริมาณอมิโลสที่แปลงข้อมูลสเปกตรัมด้วย MSC, 2nd derivative และ smoothing ในช่วงความยาวคลื่น 1300-2448 นาโนเมตร (Figure 4.4) พบว่าที่ความยาวคลื่น 1524, 1582, 1772 และ 2102 นาโนเมตร มีความเกี่ยวข้องกับพันธะ C-H และ O-H ที่มีความสัมพันธ์กับแป้ง (Osborne *et al.*, 1993) ขณะที่ Shenk *et al.* (2001) พบว่าที่ความยาวคลื่น 1524 นาโนเมตรนั้นมีความเกี่ยวข้องกับพันธะ N-H งานวิจัยส่วนใหญ่พบว่าที่ความยาวคลื่น 2100 นาโนเมตร มีความสัมพันธ์กับแป้งมากที่สุด (Osborne *et al.*, 1993; William and Norris, 2001; Shenk *et al.*, 2001) แต่ในการศึกษาครั้งนี้พบพิกที่ความยาวคลื่น 2102 นาโนเมตร อันเป็นผลจากการแปลงข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์ จึงทำให้พิกเกิดการเลื่อนไปเล็กน้อย ซึ่งพันธะและโมเลกุลที่เกี่ยวข้องนั้นมีความสัมพันธ์กับโครงสร้างของอมิโลสที่เป็นองค์ประกอบส่วนหนึ่งของแป้ง

จากผลการทดลองจะเห็นว่าสมการทำนายที่ให้ผลการทำนายดีที่สุดอยู่ในช่วงความยาวคลื่น 1300-2448 นาโนเมตร ซึ่งมีข้อมูลความยาวคลื่น 2100 นาโนเมตร อยู่ด้วย

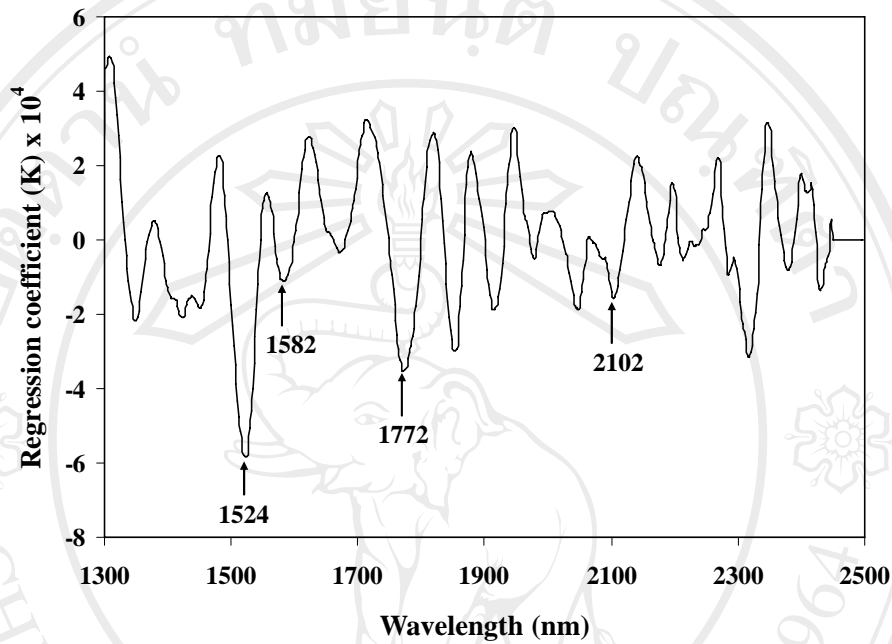


Figure 4.4 Regression coefficient plots for amylose calibration equation of five milled rice cultivars.

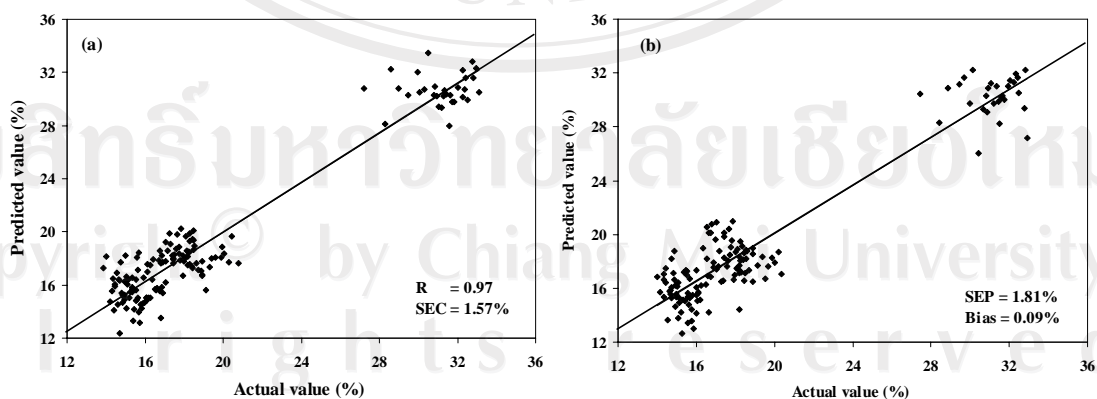


Figure 4.5 Scatter plots for predicting amylose content of (a) the calibration sample set and (b) validation sample set using amylose calibration equation developed.

จากกราฟการทำนายปริมาณอมิโลสในข้าวสาร 5 พันธุ์ (Figure 4.5) พบว่ามี การแบ่งกลุ่มพันธุ์ข้าวตามปริมาณอมิโลสออกเป็น 2 กลุ่มได้อย่างชัดเจน คือกลุ่มข้าวที่มีปริมาณ อมิโลสต่ำได้แก่ พันธุ์กข15, พันธุ์ขาวดอกมะลิ 105, พันธุ์ข้าวเจ้าหอมสุพรรณบุรี และพันธุ์ ปทุมธานี 1 และกลุ่มข้าวที่มีปริมาณอมิโลสสูง ได้แก่ พันธุ์ชัยนาท 1 ทั้งนี้อาจเป็นผลจากพันธุ์ข้าว ที่นำมาสร้างสมการทำนายมีปริมาณอมิโลสแตกต่างกันมากคือ ปริมาณอมิโลสต่ำและสูง ทำให้มี ช่วงของปริมาณอมิโลสที่นำมาสร้างสมการค่อนข้างกว้าง ค่า R ที่ได้จึงมีค่าสูง รวมทั้งค่า SEC และ SEP สูงขึ้นด้วย (William, 2001) การทำให้ค่า SEC และ SEP ลดต่ำลง อาจแก้ไขได้ โดย การสร้างสมการทำนาย 2 สมการ คือสมการทำนายสำหรับข้าวที่มีปริมาณอมิโลสต่ำและสมการ ทำนายสำหรับข้าวที่มีปริมาณอมิโลสสูง หรือหาข้าวพันธุ์ที่มีอมิโลสปานกลางมาใช้ในการพัฒนา สมการในลำดับต่อไป

4.3.2 การสร้างสมการทำนายปริมาณโปรตีนในข้าวสาร

การสร้างสมการทำนายปริมาณโปรตีนในข้าวสาร จำนวน 120 ตัวอย่าง โดยนำข้อมูล สเปกตรัม ในช่วงความยาวคลื่น 1100-2500 นาโนเมตร กับข้อมูลปริมาณโปรตีนในข้าวที่ได้จาก การวิเคราะห์ทางเคมีมาหาความสัมพันธ์ โดยการสร้างสมการถดถอยเชิงเส้นด้วยเทคนิค PLSR และตรวจสอบ outlier ด้วยวิธี leverage correction และ full cross validation พบว่าข้าวสาร ตัวอย่างที่ 20 เป็น outlier จึงตัดออก สร้างสมการและใช้วิธีทดสอบสมการ (validation method) แบบ test set โดยแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มสร้างสมการทำนายค่าทางเคมี (calibration set) จำนวน 64 ตัวอย่าง และกลุ่มทดสอบสมการ (validation set) จำนวน 55 ตัวอย่าง ข้อมูลทางสถิติ ของตัวอย่างทั้งสองกลุ่มแสดงใน Table 4.5 โดยมีปริมาณโปรตีนในช่วง 6.64-9.03%

Table 4.5 Characteristics of calibration and validation sample set of five milled rice cultivars for protein content determination by NIR.

Variable	Calibration sample set	Validation sample set
Number of samples	64	55
Protein content (% wet basis)	6.64-9.03	6.65-8.89
Mean	7.72	7.68
Standard deviation (SD)	0.50	0.47

การสร้างสมการทำนายปริมาณโปรตีนของข้าวสารทั้ง 5 พันธุ์ (Table 4.6) เมื่อเปรียบเทียบการแปลงข้อมูลด้วยเทคนิคทางคณิตศาสตร์ทั้ง 4 วิธี พบว่า มีจำนวน factor เท่ากันคือ 6 factor ส่วนความแม่นยำของสมการที่แปลงข้อมูลสเปกตรัมด้วย MSC, 2nd derivative ร่วมกับ smoothing ต่ำกว่าสมการที่แปลงข้อมูลสเปกตรัมอีก 3 วิธี มีค่า R เท่ากับ 0.95 แต่เมื่อทดสอบสมการทางสถิติของค่า SEP พบว่าค่า SEP มีค่าต่ำที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ อีกทั้งสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย MSC, 2nd derivative ร่วมกับ smoothing มีค่า RPD สูงสุดด้วย ดังนั้นจึงเลือกเป็นสมการสำหรับใช้ในการทำนายปริมาณโปรตีนในช่วงความยาวคลื่น 1130-2250 นาโนเมตร โดยให้ค่า F, R, SEC, SEP, Bias และ RPD เท่ากับ 6, 0.95, 0.15%, 0.19%, 0.04% และ 2.52 ตามลำดับ (Table 4.6)

Table 4.6 PLSR calibration results for protein content using spectra pretreatments.

Pre-treatment	Wavelength region (nm)	F	R	SEC	SEP	Bias	RPD
2 nd derivative	1110-2448	6	0.96	0.14	0.23 ^b	0.04	2.03
2 nd derivative+smoothing	1120-2250	6	0.96	0.15	0.21 ^b	0.05	2.25
MSC+2 nd derivative	1150-2448	6	0.96	0.14	0.22 ^b	0.05	2.09
MSC+2nd derivative+smoothing	1130-2250	6	0.95	0.15	0.19^a	0.04	2.52

F: number of factors used in the calibration equation.

R: multiple correlation coefficients.

SEC: standard error of calibration.

SEP: standard error of prediction, tested by the method of Fearn (1996).

Bias: average of difference between actual value and NIR value.

RPD: ratio of standard deviation of reference data in validation set to SEP.

Means with the different letters in same column indicate significant difference at 95% confidence.

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอย (regression coefficient) ของสมการทำนายปริมาณโปรตีน ที่แปลงข้อมูลสเปกตรัมด้วยวิธี MSC, 2nd derivative และ smoothing ในช่วงความยาวคลื่น 1130-2250 นาโนเมตร พบว่าที่ความยาวคลื่น 1726, 1990 และ 2188 นาโนเมตร มีความสัมพันธ์กับปริมาณโปรตีนมาก (Figure 4.6) ซึ่ง Osborne *et al.* (1993) และ Shenk *et al.* (2001) รายงานว่า ที่ความยาวคลื่น 1725 มีความสัมพันธ์กับโครงสร้าง CH₂ ขณะที่ Rittiron *et al.* (2004, 2005) พบโปรตีนในการตรวจวัดข้าวเมล็ดเดี่ยว (single kernel) ของข้าวกล้องและข้าวสารญี่ปุ่นที่ความยาวคลื่น 1735 และ 1726 นาโนเมตร ตามลำดับ ส่วนที่ความยาวคลื่น 1980 และ 2180 นาโนเมตร มีความสัมพันธ์กับโปรตีน (Osborne *et al.*, 1993; Shenk *et al.*, 2001) นอกจากนี้ Coleman *et al.* (1982) ได้รายงานช่วงความยาวคลื่น 2100-2200 นาโนเมตร มีความสัมพันธ์กับปริมาณโปรตีนในอาหารสัตว์ เช่นเดียวกับรายงานของ Osborne *et al.* (1993)

พบว่า ที่ความยาวคลื่น 2050 และ 2180 นาโนเมตร มีความสัมพันธ์กับโปรตีน สอดคล้องกับ Shenk *et al.* (2001) รายงานว่าผลผลิตทางการเกษตรส่วนใหญ่ จะพบฟีกของโปรตีนที่ความยาวคลื่น 2180 นาโนเมตร แต่ในการศึกษาครั้งนี้พบฟีกที่ความยาวคลื่น 2188 นาโนเมตร ซึ่งฟีกที่พบอาจ เลื่อนไปเล็กน้อย เนื่องจากการแปลงข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์ ดังนั้นสมการ ทำนายที่สร้างขึ้นในช่วงความยาวคลื่น 1130-2250 นาโนเมตร จึงมีความยาวคลื่นที่สัมพันธ์กับ ปริมาณโปรตีนในข้าวสาร ผลการทำนายปริมาณโปรตีนในข้าวสารโดยใช้สมการทำนายที่สร้างขึ้น จากกลุ่มตัวอย่าง calibration set และกลุ่มตัวอย่าง validation set แสดงใน Figure 4.7 ที่ให้ค่า R, SEC, SEP และ Bias เท่ากับ 0.95, 0.15%, 0.19% และ 0.04% ตามลำดับ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

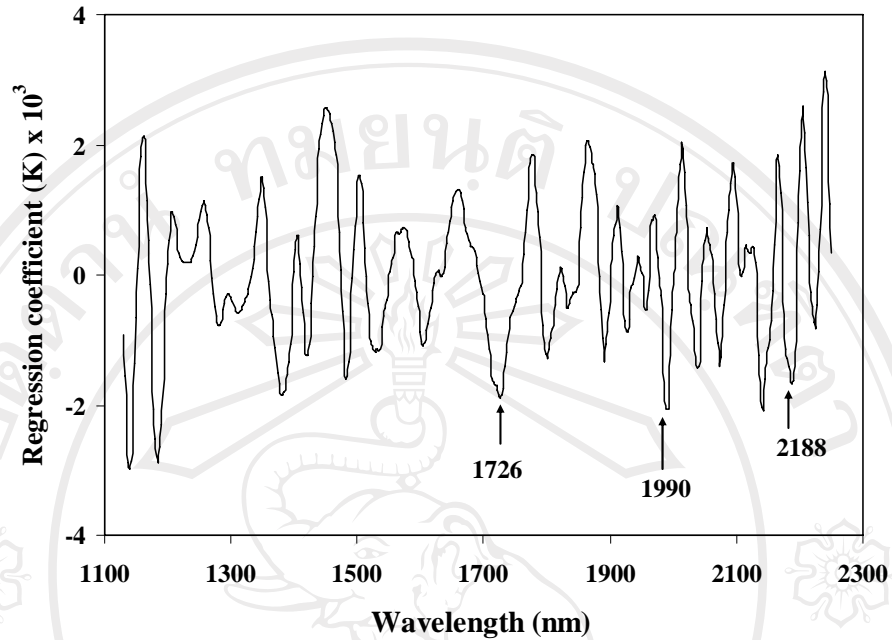


Figure 4.6 Regression coefficient plots for protein calibration equation of five milled rice cultivars.

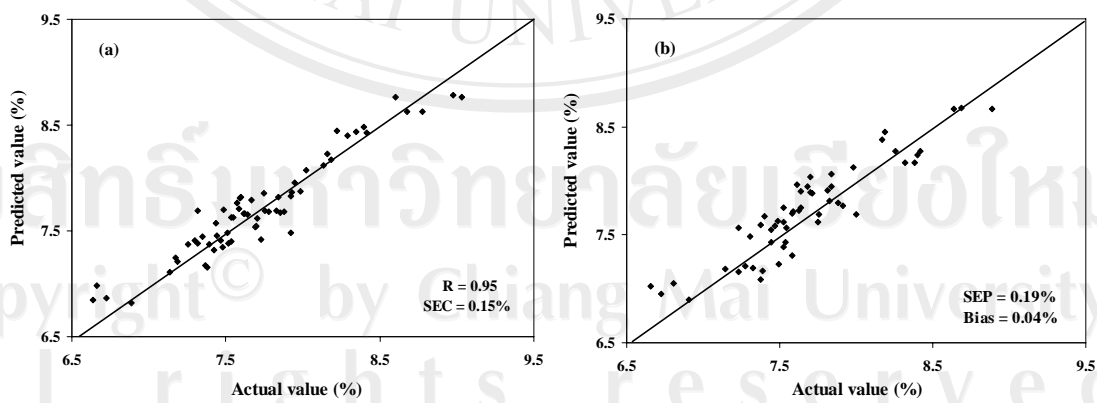


Figure 4.7 Scatter plots for predicting protein content of (a) the calibration sample set and (b) validation sample set using protein calibration equation developed.

4.3.3 การสร้างสมการทำนายปริมาณไขมันรวมในข้าวสาร

การสร้างสมการทำนายปริมาณไขมันโดยรวมในข้าวสาร จำนวน 180 ตัวอย่าง โดยนำข้อมูลสเปกตรัม ในช่วงความยาวคลื่น 1100-2500 นาโนเมตร กับข้อมูลปริมาณไขมันโดยรวมในข้าวที่ได้จากการวิเคราะห์ทางเคมีมาหาความสัมพันธ์ โดยการสร้างสมการถดถอยเชิงเส้นด้วยเทคนิค PLSR และตรวจสอบ outlier ด้วยวิธี leverage correction และ full cross validation พบว่าข้าวสารตัวอย่างที่ 58 เป็น outlier จึงตัดออก สร้างสมการและใช้วิธีทดสอบสมการ (validation method) แบบ test set โดยแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มสร้างสมการทำนายค่าทางเคมี (calibration set) จำนวน 94 ตัวอย่าง และกลุ่มทดสอบสมการ (validation set) จำนวน 85 ตัวอย่าง ข้อมูลทางสถิติของตัวอย่างทั้งสองกลุ่มแสดงใน Table 4.7 โดยมีปริมาณไขมันโดยรวมในช่วง 0.27-1.95%

Table 4.7 Characteristics of calibration and validation sample set of five milled rice cultivars for crude lipid content determination by NIR.

Variable	Calibration sample set	Validation sample set
Number of samples	94	85
Crude lipid (% wet basis)	0.27-1.95	0.41-1.89
Mean	1.08	1.07
Standard deviation (SD)	0.42	0.41

จากการสร้างสมการทำนายปริมาณไขมันโดยรวมของกลุ่มตัวอย่าง calibration set และกลุ่มตัวอย่าง validation set แสดงใน Table 4.8 เมื่อเปรียบเทียบการแปลงข้อมูลด้วยเทคนิคทางคณิตศาสตร์ทั้ง 4 วิธี พบว่า มีค่า R, SEC, SEP และ Bias ไม่แตกต่างกัน แต่สมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative ร่วมกับ MSC และสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing ให้ค่า RPD สูงสุดเท่ากับ 3.01 อย่างไรก็ตามสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing จะให้พีคในกราฟสัมประสิทธิ์ของสมการถดถอยที่เกี่ยวข้องกับไขมันชัดเจนกว่าสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative และ MSC ดังนั้น จึงเลือกสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing ในช่วงความยาวคลื่น 1500-2450 นาโนเมตร โดยให้ค่า R, SEC, SEP, Bias และ RPD เท่ากับ 0.95, 0.13%, 0.14%, 0.02% และ 3.01 ตามลำดับ

Table 4.8 PLSR calibration results for crude lipid content using spectra pretreatments.

Pre-treatment	Wavelength region (nm)	F	R	SEC	SEP	Bias	RPD
2 nd derivative	1500-2450	4	0.95	0.13	0.14	0.01	2.93
2 nd derivative+smoothing	1500-2450	4	0.95	0.13	0.14	0.01	2.92
2 nd derivative+MSC	1500-2450	4	0.95	0.13	0.14	0.02	3.01
2nd derivative +MSC+smoothing	1500-2450	4	0.95	0.13	0.14	0.02	3.01

F: number of factors used in the calibration equation.

R: multiple correlation coefficients.

SEC: standard error of calibration.

SEP: standard error of prediction, tested by the method of Fearn (1996).

Bias: average of difference between actual value and NIR value.

RPD: ratio of standard deviation of reference data in validation set to SEP.

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอย (regression coefficient) ของสมการทำนายปริมาณไขมันโดยรวมที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing ในช่วงความยาวคลื่น 1500-2450 นาโนเมตร (Figure 4.8) พบว่าที่ความยาวคลื่น 1722, 1764 และ 2378 นาโนเมตร มีความเกี่ยวข้องกับพันธะ CH (Osborne *et al.*, 1993; Shenk *et al.*, 2001; William, 2007) ซึ่งเป็นพันธะเคมีที่พบในโครงสร้างไขมัน Iwamoto *et al.* (1984) รายงานว่าที่ความยาวคลื่น 1724 นาโนเมตร มีความสัมพันธ์กับปริมาณไขมันในแป้งบริสุทธ์ ส่วนผลผลิตทางการเกษตรส่วนใหญ่จะพบพิกของไขมันที่ความยาวคลื่น 2310 นาโนเมตร (Shenk *et al.*, 2001) ดังนั้น สมการทำนายที่สร้างขึ้นในช่วงความยาวคลื่น 1500-2450 นาโนเมตร จึงมีความยาวคลื่นที่สัมพันธ์กับปริมาณไขมันในข้าวสาร ผลการทำนายปริมาณไขมันในข้าวสาร โดยใช้สมการทำนายที่เลือกแสดงใน Figure 4.9

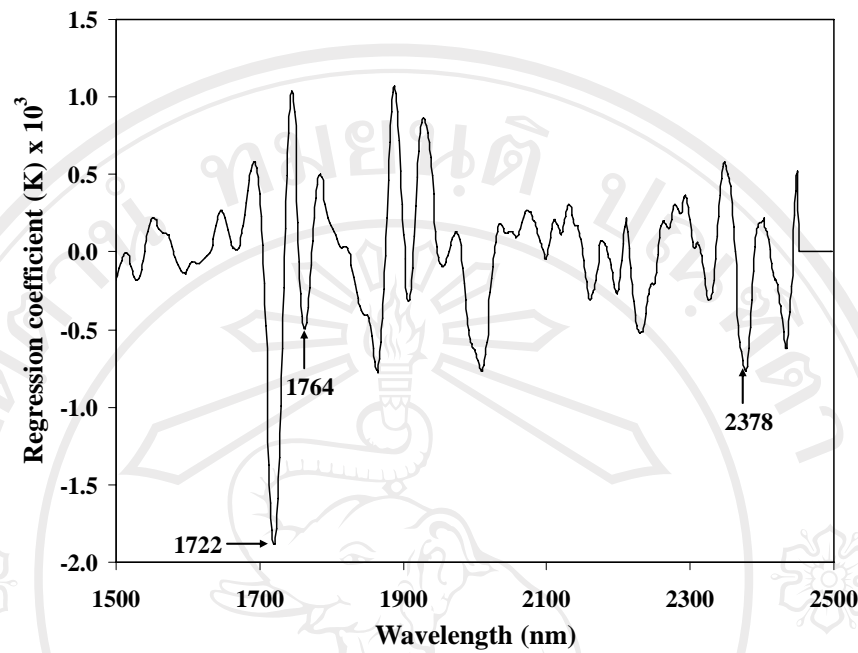


Figure 4.8 Regression coefficient plots for crude lipid calibration equation of five milled rice cultivars.

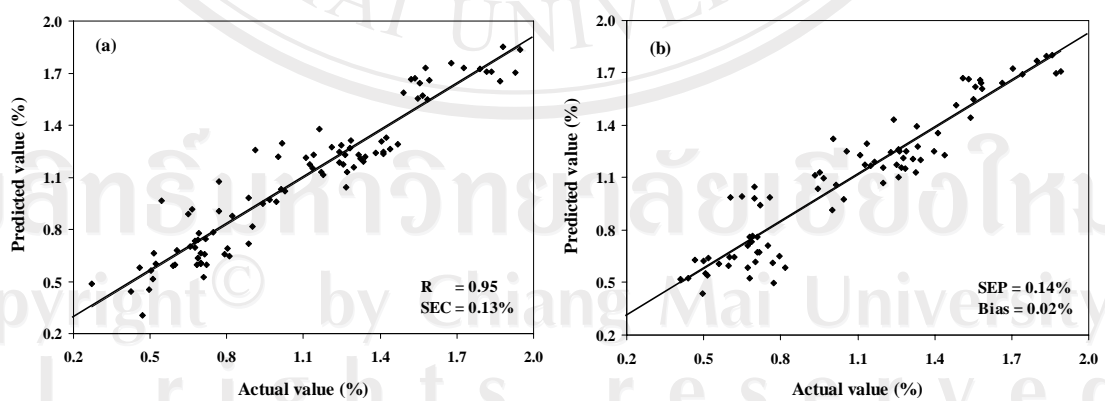


Figure 4.9 Scatter plots for predicting crude lipid content of (a) the calibration sample set and (b) validation sample set using crude lipid calibration equation developed.

4.3.4 การสร้างสมการทำนายปริมาณความชื้นในข้าวสาร

การสร้างสมการทำนายปริมาณความชื้นในข้าวสาร จำนวน 200 ตัวอย่าง โดยนำข้อมูลสเปกตรัม ในช่วงความยาวคลื่น 1100-2500 นาโนเมตร กับข้อมูลปริมาณความชื้นในข้าวที่ได้จากการวิเคราะห์ทางเคมีมาหาความสัมพันธ์ โดยการสร้างสมการถดถอยเชิงเส้นด้วยเทคนิค PLSR และตรวจสอบ outlier ด้วยวิธี leverage correction และ full cross validation พบว่าข้าวสารตัวอย่างที่ 25 เป็น outlier จึงตัดออก สร้างสมการและใช้วิธีทดสอบสมการ (validation method) แบบ test set โดยแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มสร้างสมการทำนายค่าทางเคมี (calibration set) จำนวน 104 ตัวอย่าง และกลุ่มทดสอบสมการ (validation set) จำนวน 95 ตัวอย่าง ค่าทางสถิติตัวอย่างทั้งสองกลุ่มแสดงใน Table 4.9 โดยมีปริมาณความชื้นในช่วง 9.38-10.98%

Table 4.9 Characteristics of calibration and validation sample set of five milled rice cultivars for moisture content determination by NIR.

Variable	Calibration sample set	Validation sample set
Number of samples	104	95
Moisture content (% wet basis)	9.38-10.98	9.42-10.97
Mean	10.21	10.20
Standard deviation (SD)	0.43	0.42

จาก Table 4.10 เมื่อทำการเปรียบเทียบการแปลงข้อมูลด้วยเทคนิคทางคณิตศาสตร์ทั้ง 4 วิธี พบสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative ร่วมกับ MSC และสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing ให้ค่า R สูงสุด เท่ากับ 0.93 โดยมีค่า SEP ไม่แตกต่างกัน อย่างไรก็ตามการที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing จะให้พิกในกราฟสัมพันธ์ของสมการถดถอยที่เกี่ยวข้องกับความชื้นชัดเจนกว่าสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative และ MSC ดังนั้น จึงเลือกสมการทำนายที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing ในช่วงความยาวคลื่น 1300-2400 นาโนเมตร โดยให้ค่า R, SEC, SEP, Bias และ RPD เท่ากับ 0.93, 0.16%, 0.16%, 0.02% และ 2.67 ตามลำดับ

Table 4.10 PLSR calibration results for moisture content using spectra pretreatments.

Pre-treatment	Wavelength region (nm)	F	R	SEC	SEP	Bias	RPD
2 nd derivative	1200-2400	4	0.92	0.16	0.16	0.02	2.67
2 nd derivative+smoothing	1140-2400	4	0.92	0.16	0.16	0.02	2.73
2 nd derivative+MSC	1200-2300	4	0.93	0.16	0.16	0.02	2.64
2nd derivative+MSC+smoothing	1300-2300	4	0.93	0.16	0.16	0.02	2.67

F: number of factors used in the calibration equation.

R: multiple correlation coefficients.

SEC: standard error of calibration.

SEP: standard error of prediction, tested by the method of Fearn (1996).

Bias: average of difference between actual value and NIR value.

RPD: ratio of standard deviation of reference data in validation set to SEP.

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอย (regression coefficient) ของสมการทำนาย ปริมาณความชื้น ที่แปลงข้อมูลสเปกตรัมด้วย 2nd derivative, MSC และ smoothing ในช่วง ความยาวคลื่น 1300-2300 นาโนเมตร (Figure 4.10) พบว่าพีกที่ความยาวคลื่น 1400, 1452 และ 1936 นาโนเมตร มีความเกี่ยวข้องกับพันธะ O-H (Osborne *et al.*, 1993; Shenk *et al.*, 2001) จากรายงานของ Curcio and Petty (1951) พบว่าน้ำบริสุทธิ์ที่อุณหภูมิ 20 องศาเซลเซียส จะ ดูดกลืนแสงที่ความยาวคลื่น 760, 970, 1190, 1450 และ 1940 นาโนเมตร สอดคล้องกับ William and Norris (2001) ที่พบพีกน้ำชัดเจนที่สุดที่ความยาวคลื่น 1450 และ 1940 นาโนเมตร และ Shenk *et al.* (2001) พบว่าพีกน้ำที่ชัดเจนในผลผลิตทางการเกษตรคือ ที่ความยาวคลื่น 1940 นาโนเมตร แต่ในการศึกษาครั้งนี้ พบพีกที่ความยาวคลื่น 1936 นาโนเมตร อันเป็นผลจากการแปลง ข้อมูลสเปกตรัมด้วยเทคนิคทางคณิตศาสตร์ จึงทำให้พีกเกิดการเลื่อนไปเล็กน้อย จะเห็นได้ว่าการ ดูดกลืนแสงของน้ำอาจเกิดขึ้นที่ความยาวคลื่นต่างกัน เนื่องจากน้ำเกิด weak band ในช่วงคลื่นแสง NIR จึงมีประสิทธิภาพมากในการดูดกลืนแสง รวมถึงอุณหภูมิของตัวอย่างที่มีผลกระทบโดยตรง ต่อโมเลกุลของน้ำ ที่สามารถส่งผลกระทบต่อความแม่นยำของสมการได้ (Iwamoto *et al.*, 1995) ดังนั้น สมการทำนายที่สร้างขึ้นในช่วงความยาวคลื่น 1300-2300 นาโนเมตร จึงมีความยาวคลื่นที่สัมพันธ์ กับปริมาณความชื้นในข้าวสาร ผลการทำนายปริมาณความชื้นในข้าวสาร โดยใช้สมการทำนายที่ สร้างขึ้นในกลุ่มตัวอย่าง calibration set และกลุ่มตัวอย่าง validation set แสดงใน Figure 4.11 จะพบค่า R ที่ค่อนข้างต่ำกว่าสมการทำนายปริมาณองค์ประกอบทางเคมีดังที่กล่าวมาข้างต้น อาจ เป็นเพราะปริมาณความชื้นที่วิเคราะห์ได้ไม่กระจายตัว ค่อนข้างเกาะกลุ่มกัน ทำให้การกระจายของ ข้อมูลน้อยค่าที่ทำนายได้จึงยังไม่ดีนัก (จิรวรรณ, 2545) ถ้าต้องการให้ค่า R สูงขึ้น ควรมีการขยาย ช่วง ปริมาณความชื้นให้กว้างขึ้น แต่ในการศึกษาครั้งนี้มีจำนวนตัวอย่างที่จำกัด จึงทำให้ได้ ความชื้นที่ระดับเดียว

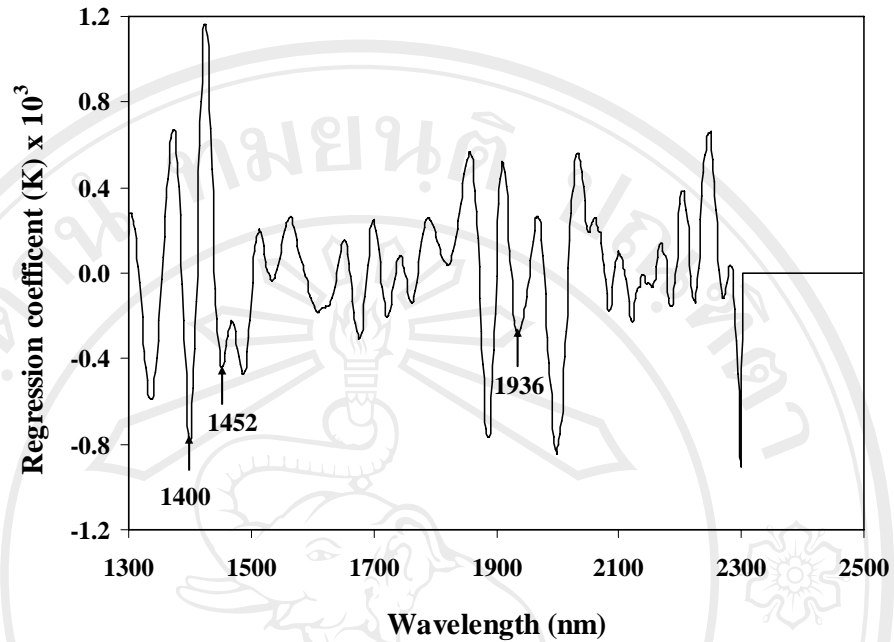


Figure 4.10 Regression coefficient plots for moisture calibration equation of five milled rice cultivars.

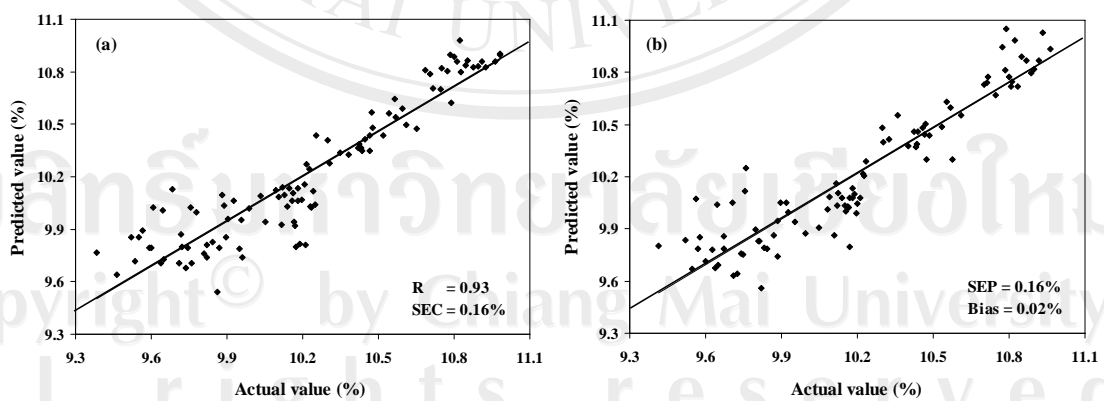


Figure 4.11 Scatter plots for predicting moisture content of (a) the calibration sample set and (b) validation sample set using moisture calibration equation developed.

4.4 การทดสอบสมการที่สร้างขึ้นด้วยข้าวจากแหล่งอื่น

นำสมการที่สร้างขึ้นในผลการทดลอง 4.3 ของปริมาณอมิโลส ปริมาณไขมันโดยรวม และปริมาณความชื้นของข้าวสาร มาทดสอบความแม่นยำ (precision) ด้วยข้าวที่ปลูกจากแหล่งอื่น ได้แก่ พันธุ์ข้าวดอกมะลิ 105 จากศูนย์วิจัยข้าวแพร่ พันธุ์ชัยนาท 1 จากเกษตรกร อำเภอตรอน จังหวัดอุดรดิตถ์ และพันธุ์ปทุมธานี 1 จากศูนย์วิจัยข้าวปทุมธานี โดยนำข้าวแต่ละพันธุ์มาวัดสเปกตรัมด้วยเครื่อง NIR พันธุ์ละ 20 ตัวอย่าง แล้วนำไปวิเคราะห์หาปริมาณอมิโลส ไขมันโดยรวม และความชื้น (Table 4.11) ซึ่งทุกขั้นตอนจะทำเช่นเดียวกับชุดที่ใช้สร้างสมการปริมาณองค์ประกอบทางเคมี แล้วนำข้อมูลไปทดสอบด้วยโปรแกรม unscrambler ซึ่งในการศึกษาครั้งนี้ จึงไม่ได้ทดสอบความแม่นยำของสมการทำนายปริมาณโปรตีน เนื่องจากต้องส่งตัวอย่างไปวิเคราะห์ทางเคมีที่ประเทศเยอรมัน

ข้อมูลทางสถิติของปริมาณอมิโลส ไขมันโดยรวม และความชื้นของตัวอย่างที่ใช้ในการทดสอบสมการ (validation unknown or prediction sample set) จำนวน 60 ตัวอย่าง พบว่าข้าวทั้ง 3 พันธุ์มีปริมาณอมิโลสอยู่ในช่วง 15.61-31.74% ปริมาณไขมันโดยรวมอยู่ในช่วง 0.35-0.91% และปริมาณความชื้นอยู่ในช่วง 10.82-13.09 % (Table 4.11)

นำสมการที่สร้างขึ้นมาทำนายปริมาณอมิโลส ไขมันโดยรวม และความชื้น ของกลุ่มตัวอย่าง unknown พบว่าให้ค่า R เท่ากับ 0.96, -0.27 และ 0.56 ตามลำดับ ค่า SEP เท่ากับ 2.07 %, 0.23% และ 0.56% ตามลำดับ และค่า bias เท่ากับ -1.17%, -0.59% และ -0.24 % ตามลำดับ (Table 4.12)

เมื่อพิจารณาค่า R ของกลุ่มตัวอย่าง validation (validation sample set) และกลุ่มตัวอย่าง unknown (prediction sample set) ของปริมาณอมิโลสมีค่าใกล้เคียงกันคือ 0.95 และ 0.96 ในส่วนของปริมาณไขมันโดยรวมและปริมาณความชื้นพบว่า ค่า R ทั้งสองกลุ่มตัวอย่างมีความแตกต่างกัน โดยค่า R ของกลุ่มตัวอย่าง validation มีค่าสูงกว่ากลุ่มตัวอย่าง unknown ซึ่งปริมาณไขมันโดยรวมมีค่า R ของกลุ่มตัวอย่าง validation เท่ากับ 0.94 และกลุ่มตัวอย่าง unknown เท่ากับ -0.27 และปริมาณความชื้นมีค่า R ของกลุ่มตัวอย่าง validation เท่ากับ 0.93 และกลุ่มตัวอย่าง unknown เท่ากับ 0.54 (Table.4.13) และเมื่อพิจารณาค่า SEP และ bias ที่ได้จากการทำนายปริมาณอมิโลส ไขมันโดยรวม และความชื้นของกลุ่มตัวอย่าง unknown เปรียบเทียบกับกลุ่มตัวอย่าง validation ของสมการ พบว่า มีค่าเพิ่มขึ้น แต่ค่า bias มีค่าเพิ่มขึ้นในทางลบ (Table.4.13) ค่า bias ที่ได้มีค่าติดลบ หมายถึง ค่าที่ได้จากการทำนายด้วยสมการมีค่าต่ำกว่าค่าทางเคมีที่วัดได้ การที่ค่า SEP และ bias ของกลุ่มตัวอย่าง unknown มีค่าเพิ่มสูงขึ้นนั้น อาจเกิดจากคุณภาพของข้าวที่นำมาทดสอบแตกต่างกัน ซึ่งพบว่าข้าวพันธุ์ปทุมธานี 1 ที่ได้จากศูนย์วิจัยข้าว

ปทุมธานี มีข้าวทองไข่มุกมากถึง 11.98% ทำให้มีผลต่อสเปกตรัมที่ได้ รวมถึงช่วงเวลาในการวัดสเปกตรัมที่แตกต่างกันด้วย สอดคล้องกับ William (2001) ที่ค่าความแตกต่างของข้อมูลที่ได้จากการทำนายด้วย NIR และข้อมูลจากวิธีอ้างอิง เกิดจากคุณภาพของตัวอย่าง, ฤดูกาล, ขั้นตอนการเตรียม, ช่วงวันเวลาในการวัด และอุณหภูมิของตัวอย่างและเครื่องมือที่มีความแตกต่างกัน โดยค่า bias ที่มีค่าต่างกันมากสามารถแก้ไขด้วยการทำ bias correction คือ นำค่า bias ขององค์ประกอบทางเคมีแต่ละค่าไปบวกกับค่าที่ทำนายด้วยสมการ สำหรับความแม่นยำของสมการทำนายที่สร้างขึ้น นอกจากจะพิจารณาจากค่า R แล้ว ยังสามารถพิจารณาจากค่า SEP ซึ่งหากค่า SEP ที่ทำนายได้ของกลุ่มตัวอย่าง unknown มีค่าใกล้เคียงกับค่า SEP ของกลุ่มตัวอย่าง validation แสดงว่าสมการที่สร้างขึ้นมีความแม่นยำ (William and Norris, 2001) เช่นเดียวกับ Kawano and Saranwong (2007) กล่าวว่า ค่า SEP ของกลุ่มตัวอย่าง unknown มีค่าไม่เกิน 1 เท่าของค่า SEP ของกลุ่มตัวอย่าง validation สามารถนำสมการที่สร้างขึ้นด้วยเทคนิค NIRS ทำนายปริมาณองค์ประกอบทางเคมีได้

จาก Table.4.13 สมการทำนายปริมาณอมิโลสมีค่า R ที่ใกล้เคียงกัน และค่า SEP ของกลุ่มตัวอย่าง unknown มีค่าไม่เกิน 1 เท่าของกลุ่มตัวอย่าง validation ถือว่าสมการทำนายปริมาณอมิโลสที่สร้างขึ้นมีความแม่นยำ และสมการทำนายปริมาณไขมันโดยรวมมีค่า SEP ของกลุ่มตัวอย่าง unknown ไม่เกิน 1 เท่าของกลุ่มตัวอย่าง validation เช่นกัน แต่มีค่า R ของกลุ่มตัวอย่าง unknown ต่ำกว่ากลุ่มตัวอย่าง validation ที่แตกต่างกัน สมการทำนายปริมาณไขมันโดยรวมที่สร้างขึ้นจึงไม่แม่นยำ อาจเกิดเนื่องจากการวิเคราะห์ทางเคมีที่มีค่าความแปรปรวนสูง และการเลือกสมการที่ไม่เหมาะสม เพื่อต้องการให้ได้ค่า R สูง, ค่า SEC และ SEP ต่ำ สำหรับสมการทำนายปริมาณความชื้นมีค่า SEP ของกลุ่มตัวอย่าง unknown เกิน 1 เท่าของกลุ่มตัวอย่าง validation และมีค่า R ของกลุ่มตัวอย่าง unknown ต่ำกว่ากลุ่มตัวอย่าง validation ที่แตกต่างกัน ดังนั้น สมการทำนายปริมาณความชื้นที่สร้างขึ้นจึงไม่มีความแม่นยำ เนื่องจากช่วงปริมาณความชื้นที่นำมาสร้างสมการอยู่ในช่วงที่แคบคือ 9.38-10.98% แต่ปริมาณความชื้นที่นำมาทำนายอยู่ในช่วง 10.82-13.09% ซึ่งมีค่าที่เกินกว่าช่วงของสมการทำนายปริมาณความชื้นที่สร้างขึ้น สอดคล้องกับ Kawano and Saranwong (2007) กล่าวว่า การที่ค่า SEP ของกลุ่มตัวอย่าง unknown เกิน 1 เท่าของค่า SEP ของกลุ่มตัวอย่าง validation เพราะจำนวนตัวอย่างที่นำมาวิเคราะห์น้อยเกินไปไม่พอเพียงที่จะใช้ทำนาย ดังนั้นจึงต้องทำการเพิ่มข้อมูลของสมการให้ครอบคลุมปริมาณองค์ประกอบทางเคมี รวมถึงการวิเคราะห์ทางเคมีต้องมีความแม่นยำของวิธีมาตรฐานด้วย (Kawano, 2002) เพื่อพัฒนาสมการให้มีความแม่นยำมากขึ้น สามารถทำนายตัวอย่างในอนาคตได้

Table 4.11 Characteristics of unknown sample for prediction sample set of amylose, crude lipid and moisture content (% wet basis) in three milled rice cultivars.

Variable	Amylose	Crude lipid	Moisture
Number of samples	60	60	60
Range	15.61-31.74	0.35-0.91	10.82-13.09
Mean	22.60	0.61	12.23
Standard deviation (SD)	5.61	0.13	0.66

Table 4.12 Result of precision test of amylose, crude lipid and moisture calibration equations developed from prediction unknown sample set.

Chemical compositions	Wavelength region (nm)	Calibration sample set			Prediction sample set		
		F	R	SEC	R	SEP	Bias
Amylose	1300-2448	7	0.97	1.57	0.96	2.07	-1.17
Crude lipid	1500-2450	4	0.95	0.13	-0.27	0.23	-0.59
Moisture	1300-2300	4	0.93	0.16	0.54	0.56	-0.24

F: number of factors used in the calibration equation.

R: multiple correlation coefficients.

SEC: standard error of calibration.

SEP: standard error of prediction.

Bias: average of difference between actual value and NIR value.

Table 4.13 A values of R, SEP and Bias of sample from validation set and prediction unknown set of amylose, crude lipid and moisture content.

Sample set	Chemical compositions								
	Amylose			Crude lipid			Moisture		
	R	SEP	Bias	R	SEP	Bias	R	SEP	Bias
Validation	0.95	1.81	0.09	0.94	0.14	0.02	0.93	0.16 ^a	0.02
Prediction	0.96	2.07	-1.17	-0.27	0.23	-0.59	0.54	0.56 ^b	-0.24

R: multiple correlation coefficients.

SEP: standard error of prediction, tested by double of values of SEP from validation set (Kawano and Saranwong, 2007).

Bias: average of difference between actual value and NIR value.