การสกัดดัชนีสำหรับการเกาะกลุ่มข้าวพันธุ์ไทยตามการตอบสนองทางสรีรวิทยาต่อความเครียดจากความเค็ม

นางสาวพรชนิตว์ ทรัพย์วิไล

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์และวิทยาการคณนา ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2558
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Index Extraction for Thai Rice Cultivars Clustering according to Physiological Responses to Salt Stress

Miss Pornchanit Supwilai

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Applied Mathematics and Computational Science
Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University
Academic Year 2015

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การสกัดดัชนีสำหรับการเกาะกลุ่มข้าวพันธุ์ไทยตามการตอบสนองทางสรีรวิทยาต่อความเครียดจากความเค็ม
โดย	นางสาวพรชนิตว์ ทรัพย์วิไล
สาขาวิชา	คณิตศาสตร์ประยุกต์และวิทยาการคณนา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	อาจารย์ ดร.กิติพร พลายมาศ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	รองศาสตราจารย์ ดร.ศุภจิตรา ชัชวาลย์ อาจารย์ ดร.มนนัทธ์ พงษ์พานิช

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิทยาศาสตร์
(รองศาสตราจารย์ ดร.พลกฤษณ์ แสงวณิช)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.กรุง สินอภิรมย์สราญ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(อาจารย์ ดร.กิติพร พลายมาศ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(รองศาสตราจารย์ ดร.ศุภจิตรา ชัชวาลย์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(อาจารย์ ดร.มนนัทธ์ พงษ์พานิช)

.....กรรมการ
(รองศาสตราจารย์ ดร.ชัชวิทย์ อาภรณ์เทวัญ)

.....กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.ดวงใจ สุริยาอรุณโรจน์)

พรชนิตว์ ทรัพย์วิไล : การสกัดดัชนีสำหรับการเกาะกลุ่มข้าวพันธุ์ไทยตามการตอบสนองทางสรีรวิทยาต่อความเครียดจากความเค็ม (Index Extraction for Thai Rice Cultivars Clustering according to Physiological Responses to Salt Stress) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: อ. ดร.กิติพร พลายมาศ, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: รศ. ดร.ศุภจิตรา ชัชวาลย์, อ. ดร.มนนัทธ์ พงษ์พานิช, 81 หน้า.

ข้าวเป็นพืชที่มีความสำคัญทางเศรษฐกิจของประเทศไทย หนึ่งในปัจจัยที่มีผลกระทบต่อผลผลิตข้าวคือดินเค็ม ข้อมูลจากการเปลี่ยนแปลงทางสรีรวิทยาของข้าวพื้นเมืองพันธุ์ต่าง ๆ ต่อสภาวะเค็มจึงเป็นประโยชน์ต่อการจัดกลุ่มเพื่อบ่งชี้พันธุ์ข้าวทนเค็มและไม่ทนเค็ม รวมถึงการหาดัชนีการแปลงข้อมูลดิบเพื่อเป็นตัววัดสำหรับบอกสัดส่วนการเปลี่ยนแปลงทางสรีรวิทยาก็มีความสำคัญสำหรับการวิเคราะห์ข้อมูลและการจัดกลุ่มเป็นอย่างมาก งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อ (1) หาวิธีการแปลงข้อมูลเพื่อสกัดดัชนีที่ดีที่สุดตามการตอบสนองทางสรีรวิทยาของข้าวต่อความเครียดจากความเค็มในดิน (2) จัดกลุ่มข้าวพันธุ์ต่าง ๆ ตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์ โดยศึกษาดัชนีทั้งหมด 9 ดัชนี ได้แก่ 1. Tolerance (TOL) 2. Mean productivity (MP) 3. Geometric mean productivity (GMP) 4. Stress tolerance index (STI) 5. Yield index (YI) 6. Yield stability index (YSI) 7. Harmonic mean (HM) 8. Tolerance2 (TOL2) 9. Tolerance3 (TOL3) แต่ละดัชนีจะนำมาใช้ในการแปลงข้อมูลในการจัดกลุ่มข้าวพันธุ์ต่าง ๆ ตามลักษณะความทนเค็มของข้าว และประเมินผลการจัดกลุ่มด้วยค่า Pakhira-Bandyopadhyay-Maulik (PBM) จากการศึกษาในครั้งนี้ พบว่า MP GMP และ STI เป็นดัชนีที่เหมาะสมของค่าสรีรวิทยาน้ำหนัก STI เป็นดัชนีที่เหมาะสมของค่าสรีรวิทยา RWC และ YSI เป็นดัชนีที่เหมาะสมของค่าสรีรวิทยา CMS ผลการจัดกลุ่มข้าวยังแสดงให้เห็นกลุ่มพันธุ์ข้าวไทยต่อสภาวะทนเค็มด้วย

ภาควิชา	คณิตศาสตร์และวิทยาการคอมพิวเตอร์	ลายมือชื่อนิสิต .....
		ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
สาขาวิชา	คณิตศาสตร์ประยุกต์และวิทยาการคณนา	ลายมือชื่อ อ.ที่ปรึกษาร่วม .....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม .....
ปีการศึกษา	2558	

5672027823 : MAJOR APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE
KEYWORDS: HIERARCHICAL CLUSTERING / SALT TOLERANT RICE / INDEX MEASURE

PORNCHANIT SUPWILAI: Index Extraction for Thai Rice Cultivars Clustering according to Physiological Responses to Salt Stress. ADVISOR: KITIPORN PLAIMAS, Dr.rer.nat., CO-ADVISOR: ASSOC. PROF. SUPACHITRA CHADCHAWAN, Ph.D., MONNAT PONGPANICH, Ph.D., 81 pp.

Rice is the economically important crop in Thailand. One of the factors that impact on rice productivity is saline soil. The data based on the physiological changes of local Thai rice cultivars due to salt stress are useful for clustering salt-tolerant and non-salt-tolerant rice. Additionally, finding an index to transform raw data in order to measure the fraction of changes is an important task in data analysis and clustering. The objectives of this research are: (1) to find the data transformation method to extract the best index according to physiological responses to soil salinity stress; (2) to cluster Thai rice cultivars according to salt stress response phenotype in order to evaluate salinity tolerance level in various rice cultivars. There were in total 9 indices studied: 1. Tolerance (TOL) 2. Mean productivity (MP) 3. Geometric mean productivity (GMP) 4. Stress tolerance index (STI) 5. Yield index (YI) 6. Yield stability index (YSI) 7. Harmonic mean (HM) 8. Tolerance2 (TOL2) 9. Tolerance3 (TOL3). Each index was applied to the physiological data and then, hierarchical clustering technique was applied to group rice cultivars. Then, Pakhira-Bandyopadhyay-Maulik (PBM) values were applied to appraise the goodness of a clustering structure. Overall, MP GMP and STI were found to be the best index for the weight of rice, STI was found to be the best index for the relative water content and YSI was found to be the best index for the cell membrane stability in clustering the studied rice cultivars. Additionally, the clustering result showed a group of salt-tolerant rice as well.

Department:	Mathematics and Computer Science	Student's Signature	.....
		Advisor's Signature	.....
Field of Study:	Applied Mathematics and Computational Science	Co-Advisor's Signature	.....
		Co-Advisor's Signature	.....
Academic Year:	2015		

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงได้ด้วยความกรุณาและความช่วยเหลืออย่างสูงยิ่งจาก อาจารย์ ดร.กิติพร พลายมาศ อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ รองศาสตราจารย์ ศุภจิตรา ชัชวาล อาจารย์ที่ปรึกษาร่วม และ อาจารย์ ดร. มนนัทธ์ พงษ์พานิช อาจารย์ที่ปรึกษาร่วม ที่กรุณาให้คำปรึกษา และให้ความรู้ ตลอดจนตรวจสอบและแก้ไขวิทยานิพนธ์นี้ ดิฉันขอขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณเหล่าคณาจารย์ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ สาขาคณิตศาสตร์ประยุกต์และวิทยาการคณนาที่ประสิทธิ์ประสาทวิชาให้อย่างเต็มที่

ขอขอบคุณ นายนพคุณ คุณผลวัฒนา และนิสิตปริญญาโท ภาควิชาพฤกษศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ความรู้ ให้คำแนะนำ และให้คำปรึกษาเกี่ยวกับข้อมูลที่ใช้ในวิทยานิพนธ์นี้

ขอขอบคุณครอบครัวที่คอยให้กำลังใจ ขอบคุณเพื่อน ๆ พี่ ๆ น้อง ๆ สาขาคณิตศาสตร์ประยุกต์และวิทยาการคณนาที่คอยช่วยเหลือและให้คำแนะนำต่าง ๆ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฎ
สารบัญรูปภาพ ..... ฑ
บทที่ 1 บทนำ 1
1.1 ที่มาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ของงานวิจัย 1
1.3 ขอบเขตของงานวิจัย 2
1.4 วิธีการดำเนินงานวิจัย 3
1.5 ประโยชน์ที่คาดหวังว่าจะได้รับ 4
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 5
บทที่ 3 ข้อมูลสรีรวิทยาต้นข้าว 7
3.1 การเก็บข้อมูลของข้าวชุดควบคุม 8
3.2 การเก็บข้อมูลของข้าวที่มีการให้สารละลายโซเดียมคลอไรด์ลงในดิน 8
บทที่ 4 วิธีการ 9
4.1 กระบวนการในการสกัดดัชนีที่เหมาะสม 9
4.2 ดัชนีสำหรับการแปลงข้อมูล 13
4.3 ค่าผิดปกติ (Outlier) 16
4.3.1 ค่าผิดปกติของข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 16
4.3.2 ค่าผิดปกติของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 17

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

หน้า

4.4 การแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด (Min-Max Normalization) 18
4.4 การวิเคราะห์การเกาะกลุ่ม (Clustering analysis) 19
4.4.1 การจัดกลุ่มแบบลำดับชั้น (Hierarchical clustering) 19
4.5 การประเมินจำนวนกลุ่มในการจัดกลุ่ม 21
4.5.1 การประเมินจำนวนกลุ่มในการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 21
4.5.2 การประเมินจำนวนกลุ่มในการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีโดยการใช้ Pakhira-Bandyopadhyay-Maulik (PBM) 22
4.6 การประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยา 24
4.6.1 ประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยาจากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 24
4.6.2 ประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยาจากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 24
4.7 การทดสอบแมนน์-วิทนีย์ ยู (Mann – Whitney U Test) 26
4.8 เกณฑ์การบ่งชี้ลักษณะความทนเค็มของกลุ่มจากการจัดกลุ่มข้อมูลสรีรวิทยาหกลักษณะ ที่ได้รับการแปลงด้วยดัชนีด้วยดัชนีที่เหมาะสม 27
บทที่ 5 ผลการวิจัย 28
5.1 ข้อมูลชุดที่ 1 28
5.1.1 การเลือกวันที่นำมาใช้ทดลอง 28
5.1.2 การหาค่าผิดปกติ (Outliers) 30
5.1.3 การหาดัชนีที่ดีในการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 33
5.1.4 การหาดัชนีที่เหมาะสมในการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 35

หน้า

5.1.5 การทดสอบการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม 37
5.1.6 ผลการจัดกลุ่ม 39
5.2 ข้อมูลชุดที่ 2 46
5.2.1 การเลือกวันที่นำมาใช้ทดลอง 46
5.2.2 การหาค่าผิดปกติ (Outliers) 48
5.2.3 การหาดัชนีที่ดีในการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 50
5.2.4 การหาดัชนีที่เหมาะสมในการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี 52
5.2.5 การทดสอบการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม 54
5.2.6 ผลการจัดกลุ่ม 55
5.3 ข้อเสนอแนะ 62
5.4 สรุป 63
รายการอ้างอิง 66
ภาคผนวก ก พันธุ์ข้าวชุดที่ 1 68
ภาคผนวก ข พันธุ์ข้าวชุดที่ 2 69
ภาคผนวก ค ค่าผิดปกติของข้อมูลชุดที่ 1 70
ภาคผนวก ง ค่าผิดปกติของข้อมูลชุดที่ 2 73
ภาคผนวก จ ผลการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 1 76
ภาคผนวก ฉ ผลการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 2 79

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

หน้า

ประวัติผู้เขียนวิทยานิพนธ์ 81

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

ตารางที่ 4.1 ตารางแสดงลักษณะของดัชนีต่าง ๆ 15

ตารางที่ 5.1 ตารางแสดงค่า p-value ของการทดสอบทางสถิติด้วย Paired t - test เพื่อเปรียบเทียบข้อมูลสรีรวิทยาของข้อมูลชุดที่ 1 ระหว่าง Y_c กับ Y_s ของ SFW SDW RFW RDW RWC และ CMS 28

ตารางที่ 5.2 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ เป็นค่าผิดปกติซ้ำกันมากที่สุด 31

ตารางที่ 5.3 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนี ที่เหมือนกันเป็นค่าผิดปกติซ้ำกันมากที่สุด 32

ตารางที่ 5.4 ตารางแสดงดัชนีที่ให้ค่าความแปรปรวนของค่าเฉลี่ยสมาชิกในกลุ่มที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าว ด้วยข้อมูลสรีวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีมีค่ามากที่สุด 3 อันดับ ของข้อมูลชุดที่ 1 35

ตารางที่ 5.5 ตารางแสดงดัชนีที่ให้ค่า PBM ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าวด้วยข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่เหมือนกัน แล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดมีค่ามากที่สุด 3 อันดับของข้อมูลชุดที่ 1 37

ตารางที่ 5.6 ตารางแสดงค่า p-value ของการทดสอบทางสถิติของการจัดกลุ่มโดยทดสอบแต่ละคู่กลุ่มด้วย (a) ข้อมูลสรีรวิทยา SDW ที่ได้รับการแปลงด้วยดัชนี STI (b) ข้อมูลสรีรวิทยา SDW ที่ได้รับการแปลงด้วยดัชนี TOL3 และ (c) RDW ที่ได้รับการแปลงด้วยดัชนี STI (d) ข้อมูลสรีรวิทยา RDW ที่ได้รับการแปลงด้วยดัชนี HM ของข้อมูลชุดที่ 1 38

ตารางที่ 5.7 ตารางแสดงค่าเฉลี่ยของแต่ละข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 1 45

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ 5.8 ตารางแสดงค่า p-value ของการทดสอบทางสถิติ Paired t - test เพื่อเปรียบเทียบข้อมูลสรีรวิทยาของข้อมูลชุดที่ 2 ระหว่าง Y_c กับ Y_s ของ SFW SDW RFW RDW RWC และ CMS 46

ตารางที่ 5.9 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ เป็นค่าผิดปกติซ้ำกันมากที่สุด 49

ตารางที่ 5.10 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนี ที่เหมือนกันเป็นค่าผิดปกติซ้ำกันมากที่สุด 50

ตารางที่ 5.11 ตารางแสดงดัชนีที่ให้ค่าความแปรปรวนของค่าเฉลี่ยสมาชิกในกลุ่มที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าว ด้วยข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีมีค่ามากที่สุด 3 อันดับ ของข้อมูลชุดที่ 2 52

ตารางที่ 5.12 ตารางแสดงดัชนีที่ให้ค่า PBM ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าวด้วยข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่เหมือนกัน แล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดมีค่ามากที่สุด 3 อันดับของข้อมูลชุดที่ 2 54

ตารางที่ 5.13 ตารางแสดงค่า p-value ของการทดสอบทางสถิติของการจัดกลุ่มโดยทดสอบแต่ละคู่กลุ่มด้วย (a) ข้อมูลสรีรวิทยา SFW ที่ได้รับการแปลงด้วยดัชนี TOL2 และ (b) ข้อมูลสรีรวิทยา RFW ที่ได้รับการแปลงด้วยดัชนี YI ของชุดที่ 2 55

ตารางที่ 5.14 ตารางแสดงค่าเฉลี่ยของแต่ละข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 2 61

ตารางที่ ก.1 ตารางแสดงพันธุ์ข้าวชุดที่ 1 68

ตารางที่ ข.1 ตารางแสดงพันธุ์ข้าวชุดที่ 2 69

ตารางที่ ค.1 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่สูงกว่าปกติ 70

ตารางที่ ค.2 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่ต่ำกว่าปกติ 71

ตารางที่ ค.3 ตารางแสดงพันธุ์ข้าวข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีที่เหมือนกันและดัชนีที่ทำให้ค่า PBM มากที่สุด แล้วให้ค่าระยะทางมาฮาลาโนบิสที่สูงกว่าปกติ 72

ตารางที่ ง.1 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่สูงกว่าปกติ 73

ตารางที่ ง.2 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่ต่ำกว่าปกติ 74

ตารางที่ ง.3 ตารางแสดงพันธุ์ข้าวข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีที่เหมือนกันและดัชนีที่ทำให้ค่า PBM มากที่สุด แล้วให้ค่าระยะทางมาฮาลาโนบิสที่สูงกว่าปกติ 75

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญรูปภาพ

รูปที่ 4.1 แผนภาพกระบวนการในการสกัดดัชนีที่เหมาะสมกับสรีรวิทยา และจัดกลุ่มพันธุ์ข้าวที่มาจากดัชนีที่เหมาะสมกับสรีรวิทยา 12

รูปที่ 4.2 บ็อกซ์พล็อตที่แสดงค่าผิดปกติของข้อมูล 17

รูปที่ 4.3 พงศาวลีจากการจัดกลุ่มแบบลำดับชั้น 20

รูปที่ 4.4 กราฟแสดงค่าความแปรปรวนภายในกลุ่มและระหว่างกลุ่มของการจัดกลุ่มตั้งแต่ 2 ถึง 15 กลุ่ม 22

รูปที่ 5.1 แผนภูมิแท่งเปรียบเทียบความแตกต่างของค่าเฉลี่ยทางสรีรวิทยาของข้อมูลชุดที่ 1 (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ระหว่าง Y_c กับ Y_s 29

รูปที่ 5.2 แผนภูมิแสดงการเปรียบเทียบค่าความแปรปรวนระหว่างกลุ่มจากการเลือกจำนวนกลุ่มที่เหมาะสมของแต่ละดัชนีของ (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ของข้อมูลชุดที่ 1 34

รูปที่ 5.3 แผนภูมิแสดงการเปรียบเทียบค่า PBM จากการเลือกจำนวนกลุ่มที่เหมาะสมของ Salt TOL MP GMP STI YI YSI HM TOL2 TOL3 Best1 และ Best6 ของข้อมูลชุดที่ 1 36

รูปที่ 5.4 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลสรีรวิทยา SFW SDW RFW RDW RWC และ CMS ด้วยโดยดัชนี GMP GMP MP MP STI และ YSI ตามลำดับ 41

รูปที่ 5.5 Density histogram, Density Curve และค่าเฉลี่ย ของกลุ่มที่ 1 2 และ 4 ของ (a) SFW (b) SDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 1 42

รูปที่ 5.6 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 2 และ 4 ของ (c) RFW (d) RDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 1 43

รูปที่ 5.7 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 2 และ 4 ของ (e) RWC (f) CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 1 44

รูปที่ 5.8 แผนภูมิแท่งเปรียบเทียบความแตกต่างของค่าเฉลี่ยทางสรีรวิทยาของข้อมูลชุดที่ 2 (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ระหว่าง Y_c กับ Y_s 47

รูปที่ 5.9 แผนภูมิแสดงการเปรียบเทียบค่าความแปรปรวนระหว่างกลุ่มจากการเลือกจำนวนกลุ่มที่เหมาะสมของแต่ละดัชนีของ (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ของข้อมูลชุดที่ 2 51

รูปที่ 5.10 แผนภูมิแสดงการเปรียบเทียบค่า PBM จากการเลือกจำนวนกลุ่มที่เหมาะสมของ Salt TOL MP GMP STI YI YSI HM TOL2 TOL3 Best1 และ Best6 ของข้อมูลชุดที่ 2..... 53

รูปที่ 5.11 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลสรีรวิทยา SFW SDW RFW RDW RWC และ CMS ด้วยโดยดัชนี STI MP MP MP STI และ YSI ตามลำดับ..... 57

รูปที่ 5.12 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 และ 2 ของ (a) SFW (b) SDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 2 58

รูปที่ 5.13 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 และ 2 ของ (c) RFW (d) RDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 2 59

รูปที่ 5.14 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 และ 2 ของ (e) RWC (f) CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 2..... 60

รูปที่ จ.1 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า SDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี STI..... 76

รูปที่ จ.2 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า SDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี TOL3 77

รูปที่ จ.3 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า RDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี STI 77

รูปที่ จ.4 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า RDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี HM..... 78

รูปที่ ฉ.1 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า SFW ของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี TOL2 79

รูปที่ ฉ.2 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า RFW ของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี YI 80

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ข้าวเป็นอาหารหลักของคนไทย และข้าวยังเป็นพืชที่มีความสำคัญทางเศรษฐกิจของประเทศที่มีการส่งออกเป็นอันดับต้น ๆ สิ่งแวดล้อมที่มีการเปลี่ยนแปลงอย่างรวดเร็วในปัจจุบัน เช่น ฝนไม่ตกต้องตามฤดูกาล ดินขาดความอุดมสมบูรณ์หรือดินเค็ม เป็นต้น ส่งผลกระทบต่อผลผลิตข้าวทำให้ผลผลิตข้าวลดลง

ดินเค็ม คือ ดินที่มีเกลือชนิดต่าง ๆ ซึ่งละลายน้ำได้อยู่ในดินสูงจนทำให้เป็นอันตรายต่อต้นพืช [1] เนื่องจากดินเป็นทรัพยากรที่สำคัญโดยเป็นที่ยึดเหนี่ยวของต้นข้าว และเป็นแหล่งธาตุอาหารของต้นข้าว ถ้าหากดินเค็มแล้วต้นข้าวที่ไม่ทนทานต่อความเค็ม จะเจริญเติบโตได้ไม่ดี ได้รับความเสียหาย และอาจตายได้ หลายประเทศประสบกับปัญหาดินเค็มและประเทศไทยเป็นประเทศหนึ่งที่ประสบปัญหาดินเค็มด้วย โดยเฉพาะในภาคตะวันออกเฉียงเหนือซึ่งเป็นพื้นที่ที่สำคัญในการปลูกข้าวและเป็นพื้นที่ดินเค็มโดยส่วนใหญ่ซึ่งมีเกลืออยู่ในรูปของโซเดียมคลอไรด์ (NaCl) [2] การลดความเค็มของดินมีหลายวิธี เช่น การล้างเกลือออกจากดิน การใช้ปุ๋ยอินทรีย์ เป็นต้น แต่ละวิธีก็มีค่าใช้จ่ายค่อนข้างสูงและใช้เวลานาน [2] ดังนั้นการปลูกข้าวทนเค็มจึงเป็นวิธีหนึ่งที่นิยมเพราะมีข้าวพันธุ์ที่ทนทานต่อความเค็มในระดับหนึ่ง

1.2 วัตถุประสงค์ของงานวิจัย

1. เพื่อหาการแปลงข้อมูลสำหรับกำหนดดัชนีที่เหมาะสมทางสรีรวิทยา
2. เพื่อจัดกลุ่มข้าวพันธุ์ไทยจากข้อมูลสรีรวิทยาตามการตอบสนองต่อความเครียดจากความเค็มด้วยดัชนีที่หาได้

1.3 ขอบเขตของงานวิจัย

1. ศึกษาข้อมูลการเจริญเติบโตของข้าวพันธุ์ไทย 2 ชุด จำนวน 53 และ 54 พันธุ์ในระยะต้นกล้า
2. ศึกษาข้อมูลผลกระทบของข้าวพันธุ์ไทย ตามการตอบสนองทางสรีรวิทยาของข้าวต่อความเครียดจากความเค็มของเกลือโซเดียมคลอไรด์ที่มีค่าการนำไฟฟ้า 9 เดซิซีเมนต่อเมตร
3. ใช้ข้อมูลสรีรวิทยาซึ่งประกอบด้วย น้ำหนักสดต้น น้ำหนักแห้งต้น น้ำหนักสดราก น้ำหนักแห้งราก ปริมาณน้ำสัมพัทธ์ในใบ และค่าความเสถียรของเยื่อหุ้มเซลล์เพื่อศึกษาการเจริญโตของต้นข้าว
4. ศึกษาการแปลงข้อมูลด้วยดัชนีซึ่งประกอบด้วย Tolerance, Mean productivity, Geometric mean productivity, Stress tolerance index, Yield index, Yield stability index และ Harmonic mean และพัฒนาการแปลงข้อมูลเพื่อให้ดัชนีใหม่สื่อความหมายการทนทานต่อความเครียดได้ดีขึ้น
5. จัดกลุ่มข้าวพันธุ์ต่าง ๆ โดยใช้ข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนีและใช้วิธีการจัดกลุ่มแบบลำดับชั้น (Hierarchical clustering) ซึ่งวิธีการที่ใช้ในการคำนวณระยะทางระหว่างกลุ่มที่ใช้คือวิธีของวาร์ด (Ward)
6. ใช้ค่าความแปรปรวนระหว่างกลุ่ม เป็นตัวประเมินในการเลือกจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่ม ที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะเท่านั้น ใช้ค่า PBM เป็นตัวประเมินการเลือกจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่ม ที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะเข้าร่วมกัน
7. หาดัชนีที่เหมาะสม ในการแปลงข้อมูลจากการจัดกลุ่มข้อมูลที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ และหกคุณลักษณะเข้าร่วมกัน
8. การทดสอบทางสถิติ ที่ใช้ในการประเมินผลการจัดกลุ่มข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ คือ แมนน์-วิทนีย์ ยู (Mann – Whitney U Test)

1.4 วิธีการดำเนินงานวิจัย

1. ศึกษาผลกระทบจากความเค็มของดินที่มีผลต่อข้าว รวมถึงข้อมูลการตอบสนองทางสรีรวิทยาของข้าวต่อความเครียดจากความเค็ม และวิเคราะห์ข้อมูล
2. ศึกษาการแปลงข้อมูลด้วยดัชนีและพัฒนาการแปลงข้อมูลเพื่อให้ได้ดัชนีใหม่
3. แปลงข้อมูลด้วยดัชนีต่าง ๆ
4. วิเคราะห์ข้อมูลที่ได้รับการแปลงข้อมูล
5. ตัดตัวอย่างข้อมูลที่ให้ค่าผิดปกติ (Outliers)
6. สำหรับข้อมูลที่จะนำไปจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีแล้วจะแปลงข้อมูล ให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด (Min-Max Normalization) สำหรับข้อมูลที่จะนำไปจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีแล้วจะแปลงข้อมูล ให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดในการประเมินดัชนีที่เหมาะสมเท่านั้น
7. ทำการจัดกลุ่มข้าวพันธุ์ต่าง ๆ ด้วยวิธีจัดกลุ่มแบบลำดับชั้นจากข้อมูลที่ได้รับการแปลงด้วยดัชนี
8. เลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่ม ที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะและข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี
9. คัดเลือกดัชนีที่เหมาะสมโดยการประเมินผลการจัดกลุ่ม ดัชนีที่เหมาะสมของแต่ละสรีรวิทยาของการจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ มาจากดัชนีที่ทำให้การจัดกลุ่มด้วยจำนวนกลุ่มที่เหมาะสม (จากขั้นที่ 8) มีค่าความแปรปรวนระหว่างกลุ่มมากอย่างผิดปกติ ดัชนีที่เหมาะสมของการจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะมาจากดัชนีที่ทำให้การจัดกลุ่มด้วยจำนวนกลุ่มที่เหมาะสม (จากขั้นที่ 8) มีค่า PBM ที่คำนวณจากข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีแล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่กำหนดด้วยการกำหนดค่าต่ำสุดและสูงสุดมีค่ามากที่สุด
10. ใช้วิธีการทางสถิติในการทดสอบการจัดกลุ่ม
11. สรุปผลขั้นสุดท้ายที่ได้ และนำมาเรียบเรียงเขียนรายงาน

1.5 ประโยชน์ที่คาดหวังว่าจะได้รับ

1. ได้การแปลงค่าข้อมูลสำหรับกำหนดดัชนีที่เหมาะสมทางสรีรวิทยา
2. ได้ผลการจัดกลุ่มข้าวพันธุ์ไทยจากข้อมูลสรีรวิทยาตามการตอบสนองต่อความเครียดจากความเค็มด้วยดัชนีที่เหมาะสม

วิทยานิพนธ์เล่มนี้ แบ่งออกเป็น 5 บท ในบทที่ 2 กล่าวถึงงานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์ในส่วนของการหาการแปลงค่าข้อมูลเพื่อระบุพันธุ์ที่ทนต่อความเครียดในสภาวะต่าง ๆ บทที่ 3 อธิบายข้อมูลสรีรวิทยาต้นข้าวที่นำมาใช้ บทที่ 4 นำเสนอวิธีการที่ใช้ในวิทยานิพนธ์ และบทสุดท้าย กล่าวถึง ผลงานวิจัย ข้อเสนอแนะของงานวิจัย และบทสรุป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในปี 2008 Goudarzi และ Pakniyat [3] ได้ใช้ดัชนี Stress susceptibility index (SSI) และ Stress tolerance index (STI) เพื่อแปลงข้อมูลในการคัดเลือกพันธุ์ที่ทนทานต่อความเค็ม จุดประสงค์ของงานวิจัยนี้คือประเมินความสามารถของดัชนี SSI และ STI เพื่อแปลงข้อมูลในการคัดเลือกพันธุ์ที่ทนทานต่อความเค็ม ในงานวิจัยนี้ทำการทดลองกับข้าวสาลีจำนวน 15 พันธุ์ โดยมีชุดทดลอง 3 ชุด ชุดที่ 1 ปลูกในดินสภาวะปกติ ชุดที่ 2 และ 3 ปลูกในดินที่ได้รับความเครียดจากความเค็ม 6.8 และ 13.8 เดซิซีเมนต่อเมตร ตามลำดับ จากการทดลองพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนี SSI กับค่าปริมาณโซเดียมของต้นเป็นบวกอย่างมีนัยสำคัญ และ ค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนี STI กับค่าปริมาณโซเดียมของต้นเป็นลบอย่างมีนัยสำคัญ นั่นคือดัชนี SSI และ STI มีประสิทธิภาพในการระบุพันธุ์ที่ทนทานต่อความเค็ม

ในปี 2009 Talebi และคณะ [4] ได้นำเสนอเกณฑ์ในการคัดเลือกดัชนีบ่งชี้ความทนแล้งของข้าวสาลีดูลัม โดยคัดเลือกจากดัชนี 7 ดัชนี คือดัชนี 1. Stress susceptibility index (SSI) 2. Stress tolerance index (STI) 3. Tolerance (TOL) 4. Yield index (YI), Yield stability index (YSI), Mean productivity (MP), และ Geometric mean productivity (GMP) จุดประสงค์ของงานวิจัยนี้คือประเมินความสามารถของดัชนีทั้ง 7 เพื่อระบุพันธุ์ทนแล้งภายใต้ความหลากหลายของสภาพแวดล้อม การทดลองนี้ทำกับข้าวสาลีดูลัมจำนวน 24 พันธุ์ โดยมี 2 ชุดการทดลอง ชุดที่ 1 ให้อยู่ในภาวะขาดน้ำและชุดที่ 2 อยู่ภายใต้สภาพแวดล้อมปกติ จากการทดลองพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนี MP GMP และ STI กับผลผลิตที่อยู่ภายใต้สภาพแวดล้อมปกติและผลผลิตที่ได้จากดินที่อยู่ในภาวะขาดน้ำเป็นบวกอย่างมีนัยสำคัญ นั่นคือดัชนีเหล่านี้มีประสิทธิภาพในการระบุพันธุ์ที่ให้ผลผลิตสูงมากอย่างมีนัยสำคัญกว่าดัชนีอื่น ๆ ภายใต้สภาพแวดล้อมที่ไม่ได้รับน้ำ

ในปี 2014 Kamyab-Talesh และคณะ [5] ได้หาการแปลงข้อมูลที่ดีที่สุดโดยใช้ดัชนีเพื่อบ่งชี้พันธุ์ข้าวที่ทนทานต่อความเค็ม โดยคัดเลือกจากดัชนี 8 ดัชนี คือดัชนี 1.Stress susceptibility index (SSI) 2. Stress tolerance index (STI) 3. Tolerance (TOL) 4. Yield index (YI) 5. Yield

stability index (YSI) 6. Mean productivity (MP) 7. Geometric mean productivity (GMP) 8.Harmonic mean (HM) จุดประสงค์ของงานวิจัยนี้คือประเมินความสามารถของดัชนีทั้ง 8 เพื่อระบุพันธุ์ข้าวที่ทนทานต่อความเค็มภายใต้ความหลากหลายของสภาพแวดล้อม ข้าวจำนวน 8 พันธุ์ถูกนำมาทดลองโดยแบ่งเป็น 4 ชุด ชุดที่ 1 อยู่ในภาวะปกติ ชุดที่ 2 3 และ 4 ได้รับความเครียดจากความเค็มที่ระดับ 2 4 และ 6 เดซิซีเมนต่อเมตร ตามลำดับ จากการทดลองพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนี MP GMP STI และ HM กับผลผลิตที่อยู่ในภาวะปกติและผลผลิตที่ได้จากดินที่ได้รับความเครียดจากความเค็มทั้ง 3 ระดับ เป็นบวกอย่างมีนัยสำคัญ นั่นคือดัชนีเหล่านี้มีประสิทธิภาพในการระบุพันธุ์ที่ให้ผลผลิตสูงมากอย่างมีนัยสำคัญกว่าดัชนีอื่น ๆ ภายใต้สภาวะดินที่ได้รับความเครียดจากความเค็ม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 3
ข้อมูลสรีรวิทยาต้นข้าว

ข้อมูลที่ใช้วิเคราะห์ในวิทยานิพนธ์นี้ เป็นข้อมูลที่ได้มาจากงานวิจัยของภาควิชาพฤกษศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยเป็นข้อมูลที่ได้มาจากการทดลองที่มีแผนแบบสุ่มบล็อกสมบูรณ์ (Randomized complete block design ; RCBD) จำนวน 4 ซ้ำ ในระยะต้นกล้า ประกอบด้วยข้อมูลชุดควบคุม และข้อมูลที่ได้รับความเครียดจากความเค็มจากดิน ข้อมูลสรีรวิทยาต้นข้าวเป็นค่าเฉลี่ยของต้นข้าว 4 ต้น โดยข้อมูลสรีรวิทยาที่เก็บคือ น้ำหนักสดของต้น (SFW) น้ำหนักแห้งของต้น (SDW) น้ำหนักสดของราก (RFW) น้ำหนักแห้งของราก (RDW) ปริมาณน้ำสัมพัทธ์ในใบ (RWC) ซึ่งเป็นเกณฑ์ในการวัดปริมาณน้ำในพืช และค่าความเสถียรของเยื่อหุ้มเซลล์ (CMS) ซึ่งเป็นค่าที่คำนวณจากปริมาณการรั่วไหลของสารอิเล็กโทรไลท์ ต่อไปชื่อข้อมูลสรีรวิทยาจะใช้อักษรย่อแทน ในวิทยานิพนธ์นี้ใช้ข้อมูลสรีรวิทยาต้นข้าวพันธุ์ไทย 2 ชุด เก็บข้อมูลโดย นพคุณ คุณผลวัฒนา และคณะ รายชื่อพันธุ์ข้าวทั้งหมดแสดงในภาคผนวก ก ตารางที่ ก.1 และ ภาคผนวก ข ตารางที่ ข.1

<u>ชุดที่ 1</u> จำนวน 53 พันธุ์ ซึ่งมี Pokkali เป็นพันธุ์ทนทานต่อความเค็ม (Salt tolerance ; ST) และ IR29 เป็นพันธุ์อ่อนแอต่อความเค็ม (Salt susceptible ; SS) ปลูกที่โรงเรือนปลูกข้าว ศูนย์วิจัยข้าวนครราชสีมา สำนักวิจัยและพัฒนาข้าว กรมการข้าว นครราชสีมา ในเดือนเมษายน 2556

<u>ชุดที่ 2</u> จำนวน 54 พันธุ์ ซึ่งมี KDML105 และ CSSL11 เป็นพันธุ์ทนทานต่อความเค็มระดับปานกลาง (Moderately salt tolerance ; MT) ปทุมธานี1 เป็นพันธุ์อ่อนแอต่อความเค็ม (Salt susceptible ; SS) ปลูกที่โรงเรือนปลูกพืชทดลอง ศูนย์เชี่ยวชาญเฉพาะทางสิ่งแวดล้อมและสรีรวิทยาของพืช คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในเดือนกุมภาพันธ์ 2557

3.1 การเก็บข้อมูลของข้าวชุดควบคุม

ชุดที่ 1 เก็บข้อมูลสรีรวิทยาของต้นข้าว 4 ต้น จำนวน 3 ครั้ง เก็บข้อมูลครั้งที่ 1 เมื่อต้นข้าวอายุ 14 วัน ครั้งที่ 2 เมื่อต้นข้าวมีอายุ 17 วัน ครั้งที่ 3 เมื่อต้นข้าวมีอายุ 20 วัน

ชุดที่ 2 เก็บข้อมูลสรีรวิทยาของต้นข้าว 4 ต้น จำนวน 4 ครั้ง เก็บข้อมูลครั้งที่ 1 เมื่อต้นข้าวอายุ 14 วัน ครั้งที่ 2 เมื่อต้นข้าวมีอายุ 17 วัน ครั้งที่ 3 เมื่อต้นข้าวมีอายุ 20 วัน ครั้งที่ 4 เมื่อต้นข้าวมีอายุ 23 วัน

3.2 การเก็บข้อมูลของข้าวที่มีการให้สารละลายโซเดียมคลอไรด์ลงในดิน

เมื่อต้นข้าวอายุ 14 วัน จึงให้สารละลายโซเดียมคลอไรด์ความเข้มข้น 115 มิลลิโมลาร์ จนท่วมกระถางซึ่งทำให้มีความเค็ม 9 เดซิซีเมนต่อเมตรซึ่งอยู่ในระดับความเค็มมาก รักษาความเค็มให้คงที่ตลอดการทดลองโดยการรักษาระดับน้ำและตรวจวัดค่าการนำไฟฟ้าของดินทุก 3 วัน

ชุดที่ 1 เก็บข้อมูลสรีรวิทยาของต้นข้าว 4 ต้น จำนวน 2 ครั้ง โดยเก็บหลังจากมีการให้สารละลายโซเดียมคลอไรด์เป็นเวลา 3 และ 6 วัน ซึ่งมีการเก็บข้อมูลของข้าวที่ปลูกในภาวะปกติด้วย

ชุดที่ 2 เก็บข้อมูลสรีรวิทยาของต้นข้าว 4 ต้น จำนวน 3 ครั้ง โดยเก็บหลังจากมีการให้สารละลายโซเดียมคลอไรด์เป็นเวลา 3 6 และ 9 วัน ซึ่งมีการเก็บข้อมูลของข้าวที่ปลูกในภาวะปกติด้วย

บทที่ 4
วิธีการ

4.1 กระบวนการในการสกัดดัชนีที่เหมาะสม

กระบวนการในการสกัดดัชนีที่เหมาะสมดังที่แสดงในรูปที่ 4.1 มีกระบวนการดังนี้

1. เลือกข้อมูลสรีรวิทยาต้นข้าวที่นำมาใช้ โดยเลือกข้อมูลสรีรวิทยาต้นข้าวในช่วงเวลาที่ข้อมูลสรีรวิทยาชุดควบคุม และที่ได้รับความเครียดจากความเค็มมีความแตกต่างอย่างมีนัยสำคัญ

2. แปลงข้อมูลด้วยดัชนี 9 ดัชนี 1) Tolerance (TOL) 2) Mean productivity (MP) 3) Geometric mean productivity (GMP) 4) Stress tolerance index (STI) 5) Yield index (YI) 6) Yield stability index (YSI) 7) Harmonic mean (HM) 8) Tolerance2 (TOL2) 9) Tolerance3 (TOL3)

3. สร้างชุดข้อมูลสรีรวิทยาต้นข้าวที่ได้รับการแปลงด้วยดัชนี 2 แบบ 1) ชุดข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี ซึ่งแต่ละข้อมูลสรีรวิทยาจะประกอบด้วยชุดข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนี 9 ชุด 2) ชุดข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี ซึ่งจะประกอบด้วยข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่คละกันไป ซึ่งมีทั้งหมด 9^6 = 531,441 ชุด

<u>ชุดข้อมูลสรีรวิทยาต้นข้าวเพียงหนึ่งลักษณะที่ได้รับการแปลงด้วยดัชนี</u>

4. ตัดตัวอย่างข้อมูลที่ให้ค่าผิดปกติ (Outliers) ที่อยู่นอกช่วง $Q_1 - 1.5\text{IQR}$ ถึง $Q_3 + 1.5\text{IQR}$
5. แปลงข้อมูลที่ได้รับการแปลงด้วยดัชนี ให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด (Min-Max Normalization)
6. ทำการจัดกลุ่มข้าวพันธุ์ต่าง ๆ ด้วยวิธีจัดกลุ่มแบบลำดับชั้น (Hierarchical clustering) จากข้อมูลที่ได้รับการแปลงด้วยดัชนี
7. เลือกจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่มสรีรวิทยาวิทยาข้าวที่ได้รับการแปลงด้วยดัชนีจากจำนวนกลุ่มที่ทำให้การจัดกลุ่มมีค่าความแปรปรวนระหว่างกลุ่มมากที่สุด
8. ประเมินดัชนีที่เหมาะสม โดยดัชนีที่เหมาะสมของแต่ละสรีรวิทยามาจากดัชนีที่ทำให้การจัดกลุ่มด้วยจำนวนกลุ่มที่เหมาะสม (จากขั้นที่ 7) มีค่าความแปรปรวนระหว่างกลุ่มมากอย่างผิดปกติ
9. ใช้วิธีการทางสถิติในการทดสอบการจัดกลุ่มที่มาจากดัชนีที่เหมาะสมของแต่ละสรีรวิทยาด้วยการทดสอบ แมนน์-วิทนีย์ ยู (Mann – Whitney U Test)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

<u>ชุดข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี</u>

10. ตัดตัวอย่างข้อมูลที่ให้ค่าผิดปกติ (Outliers) ที่อยู่นอกช่วง $Q_1 - 1.5\text{IQR}$ ถึง $Q_3 + 1.5\text{IQR}$

11. ทำการจัดกลุ่มข้าวพันธุ์ต่าง ๆ ด้วยวิธีจัดกลุ่มแบบลำดับชั้น (Hierarchical clustering) จากข้อมูลที่ได้รับการแปลงด้วยดัชนี

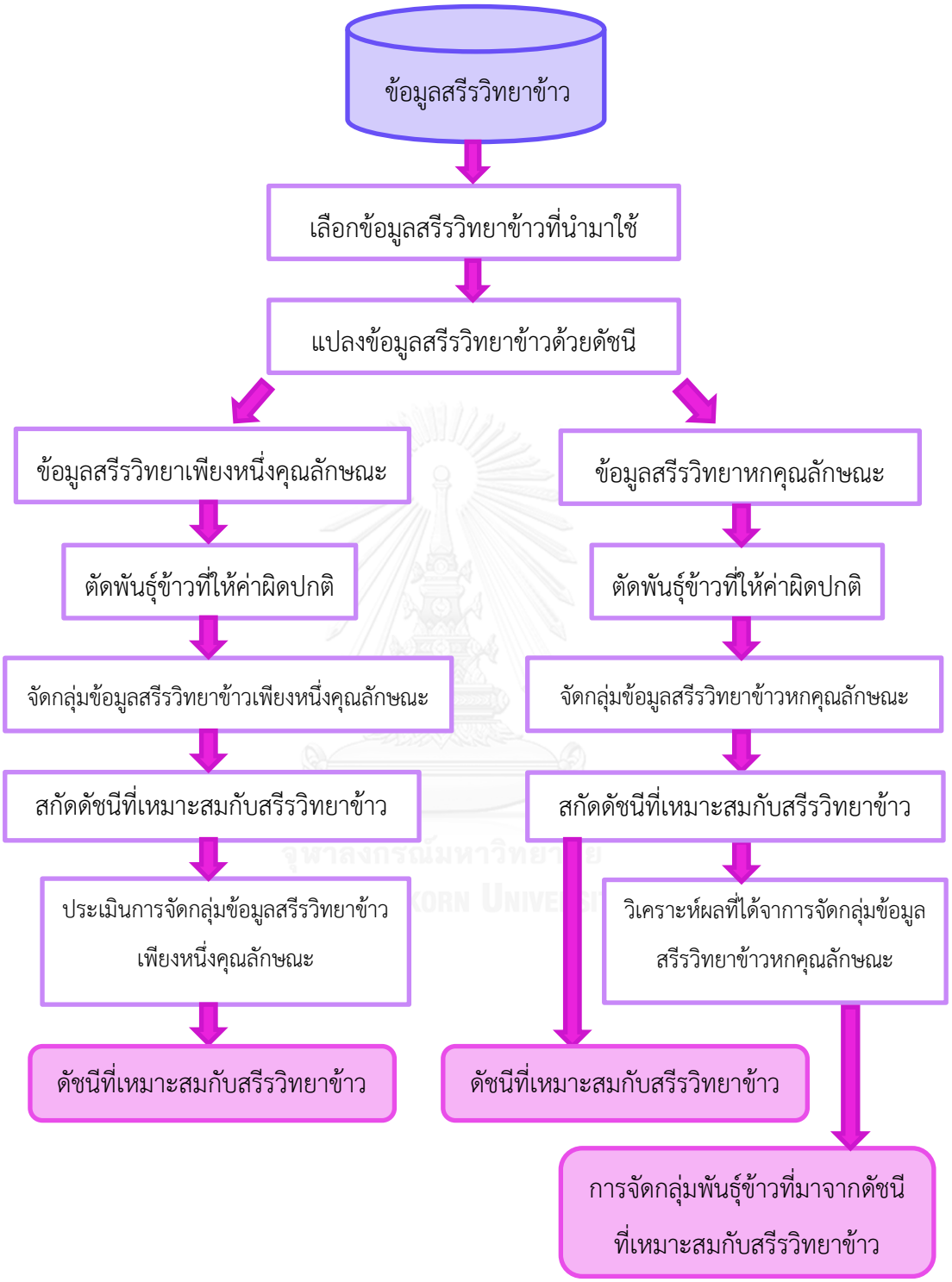
12. เลือกจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่มสรีรวิทยาวิทยาข้าวที่ได้รับการแปลงด้วยดัชนีจากจำนวนกลุ่มที่ทำให้การจัดกลุ่มมีค่า PBM มากที่สุด

13. แปลงข้อมูลที่ได้รับการแปลงด้วยดัชนี ให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด (Min-Max Normalization) ในการประเมินดัชนีที่ที่สุด

14. ประเมินดัชนีที่เหมาะสม โดยดัชนีที่เหมาะสมของแต่ละสรีรวิทยามาจากดัชนีที่ทำให้การจัดกลุ่มด้วยจำนวนกลุ่มที่เหมาะสม (จากขั้นที่ 12) มีค่า PBM จากข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนี แล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด (จากขั้นที่ 13) มีค่ามากที่สุด

15. วิเคราะห์ผลการจัดกลุ่มสรีรวิทยาหกคุณลักษณะที่เกิดจากดัชนีที่เหมาะสม โดยวิเคราะห์ข้อมูลสรีรวิทยาหกคุณลักษณะของแต่ละกลุ่มที่จัดได้ และประเมินลักษณะความทนเค็มของกลุ่มแต่ละกลุ่มจากพันธุ์ข้าวที่ระบุลักษณะความทนเค็มแล้ว

จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.1 แผนภาพกระบวนการในการสกัดดัชนีที่เหมาะสมกับสรีรวิทยา และจัดกลุ่มพันธุ์ข้าวที่มาจากดัชนีที่เหมาะสมกับสรีรวิทยา

4.2 ดัชนีสำหรับการแปลงข้อมูล

ดัชนีการแปลงข้อมูลนิยมใช้เพื่อเป็นตัววัดสำหรับบอกสัดส่วน การเปลี่ยนแปลงทางสรีรวิทยา โดยในวิทยานิพนธ์นี้ได้ทำการแปลงข้อมูลของทุกข้อมูลสรีรวิทยาด้วยดัชนี 7 ดัชนี จาก Kamyab-Talesh และคณะ [5] และดัชนีที่พัฒนาใหม่ 2 ดัชนี โดยมีตารางที่ 4.1 เป็นตารางแสดงลักษณะของดัชนีต่างๆ ดัชนีที่ใช้มีดังนี้

กำหนดให้ Y_s และ Y_c เป็นข้อมูลสรีรวิทยาของพันธุ์ที่ถูกคัดเลือกภายใต้ความเครียดและสภาวะปกติตามลำดับ และ $\bar{Y}_s$ และ $\bar{Y}_c$ เป็นค่าเฉลี่ยของข้อมูลสรีรวิทยาของทุกพันธุ์ภายใต้ความเครียดและสภาวะปกติตามลำดับ

Tolerance (TOL) เป็นดัชนีที่แสดงถึงข้อมูลสรีรวิทยาที่ลดลง โดยเปรียบเทียบข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียด และค่าที่น้อยจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{TOL} = Y_c - Y_s \tag{4.1}$$

Mean Productivity (MP) เป็นดัชนีที่แสดงค่าเฉลี่ยเลขคณิตของข้อมูลสรีรวิทยาภายใต้ทั้งสภาวะปกติและสภาวะที่ได้รับความเครียด โดยค่าเฉลี่ยชนิดนี้ไม่เหมาะกับข้อมูลที่มีข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียดมีค่าแตกต่างกันมาก [6] และค่าที่มากจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{MP} = (Y_c + Y_s)/2 \tag{4.2}$$

Geometric Mean Productivity (GMP) เป็นดัชนีที่แสดงค่าเฉลี่ยเรขาคณิตของข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียด ถ้าข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียดมีค่าแตกต่างกันมาก ค่าที่ได้จากการแปลงข้อมูลโดย GMP จะดีกว่า MP [6] และค่าที่มากจะบ่งบอกถึงความทนต่อความเครียดได้ดี โดยดัชนี GMP คำนวณได้ดังนี้

$$\mathrm{GMP} = \sqrt{Y_c Y_s} \tag{4.3}$$

Stress Tolerance Index (STI) เป็นดัชนีที่สามารถบ่งชี้พันธุ์ที่มีข้อมูลสรีรวิทยาสูงภายใต้ทั้งสภาวะปกติและสภาวะที่ได้รับความเครียด [7] โดยเป็นดัชนีที่ดัดแปลงมาจาก GMP และค่าที่มากจะบ่งบอกถึงความทนต่อความเครียดได้ดี สามารถคำนวณได้ดังนี้

$$\mathrm{STI} = Y_c Y_s / \bar{Y}_c^{\,2} \tag{4.4}$$

Yield Index (YI) เป็นดัชนีที่ใช้แค่ข้อมูลสรีรวิทยาภายใต้สภาวะที่ได้รับความเครียด ถ้ามีค่ามากกว่าหรือเท่ากับ 1 แสดงว่าข้อมูลสรีรวิทยาของพันธุ์นั้นมีค่ามากกว่าหรือเท่ากับค่าเฉลี่ยของทุกพันธุ์ ในทางตรงกันข้ามถ้ามีค่าน้อยกว่า 1 แสดงว่าข้อมูลสรีรวิทยาของพันธุ์นั้นมีค่าน้อยกว่าค่าเฉลี่ยของทุกพันธุ์ และค่าที่มากจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{YI} = Y_s / \bar{Y}_s \tag{4.5}$$

Yield Stability Index (YSI) เป็นดัชนีที่หาสัดส่วนข้อมูลสรีรวิทยาภายใต้สภาวะที่ได้รับความเครียดจากความเค็มและสภาวะปกติ และค่าที่มากจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{YSI} = Y_s / Y_c \tag{4.6}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Harmonic Mean (HM) เป็นดัชนีที่แสดงค่าเฉลี่ยฮาร์โมนิคของข้อมูลสรีรวิทยาภายใต้ทั้งสภาวะปกติและสภาวะที่ได้รับความเครียด โดยค่าเฉลี่ยชนิดนี้เหมาะสมกับข้อมูลที่เป็นอัตราส่วน ถ้าข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียดมีค่าผิดปกติโดดมากจนเกินไป ค่าที่ได้จากการแปลงข้อมูลโดย HM จะดีที่สุด [6] และค่าที่มากจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{HM} = 2Y_c Y_s / (Y_c + Y_s) \tag{4.7}$$

Tolerance2 (TOL2) เป็นดัชนีที่พัฒนาจากดัชนี TOL โดยเปลี่ยนจากการหาผลต่างของข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียดโดยตรง เป็นการหาผลต่างของสัดส่วนข้อมูลสรีรวิทยากับค่าเฉลี่ยของข้อมูลสรีรวิทยาของสภาวะปกติ และข้อมูลสรีรวิทยากับค่าเฉลี่ยข้อมูล

สรีรวิทยาของสภาวะที่ได้รับความเครียด และค่าที่น้อยจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{TOL2} = Y_c/\bar{Y}_c - Y_s/\bar{Y}_s \tag{4.8}$$

Tolerance3 (TOL3) เป็นดัชนีที่พัฒนาจากดัชนี TOL โดยเปลี่ยนจากการหาผลต่างของข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียดจากความเค็มโดยตรง เป็นผลต่างของข้อมูลสรีรวิทยาภายใต้สภาวะปกติและสภาวะที่ได้รับความเครียดจากความเค็ม หารด้วยค่าเฉลี่ยของข้อมูลสรีรวิทยาภายใต้ทั้งสภาวะปกติและสภาวะที่ได้รับความเครียด และค่าที่น้อยจะบ่งบอกถึงความทนต่อความเครียดได้ดี ดัชนีนี้คำนวณได้ดังนี้

$$\mathrm{TOL3} = 2(Y_c - Y_s)/(Y_C + Y_s) \tag{4.9}$$

ตารางที่ 4.1 ตารางแสดงลักษณะของดัชนีต่าง ๆ

Index	ช่วงของข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนี	ลักษณะของข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนีที่บ่งบอกถึงความทนเค็ม
TOL	(-∞,∞)	น้อย
MP	[0, ∞)	มาก
GMP	[0, ∞)	มาก
STI	[0, ∞)	มาก
YI	[0, ∞)	มาก
YSI	[0, ∞)	มาก
HM	[0, ∞)	มาก
TOL2	(-∞,∞)	น้อย
TOL3	(-∞,∞)	น้อย

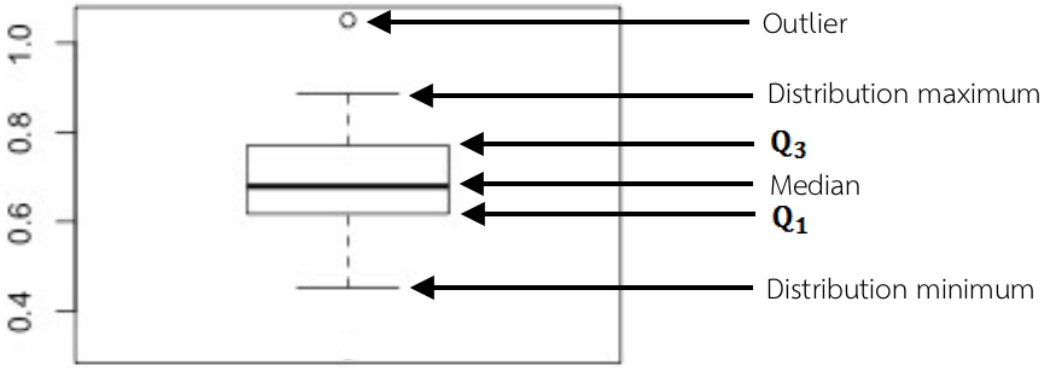
4.3 ค่าผิดปกติ (Outlier)

ข้อมูลที่ถูกแปลงข้อมูลโดยใช้ดัชนี 9 ดัชนีของแต่ละสรีรวิทยาจะถูกหาค่าผิดปกติซึ่งคือค่าที่แตกต่างจากค่าภายในกลุ่มข้อมูล ค่าผิดปกติอาจเกิดจากการเก็บข้อมูลที่ผิดพลาดหรือข้อมูลนั้นมีความแตกต่างจากกลุ่มจริง ถ้าข้อมูลมีค่าผิดปกติจะทำให้การวิเคราะห์ข้อมูลอาจเกิดความผิดพลาดได้ ดังนั้นก่อนการวิเคราะห์ข้อมูลควรตัดค่าผิดปกติออกไป วิธีการระบุว่าเป็นค่าที่มีความผิดปกติสามารถทำได้ 2 แบบดังนี้

4.3.1 ค่าผิดปกติของข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

สำหรับการหาค่าที่ผิดปกติของข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีใช้การพิจารณาจากค่าพิสัยระหว่างควอไทล์ (interquartile range : IQR) ที่ใช้วัดการกระจายตัวของข้อมูลทางสถิติ โดยสามารถคำนวณได้จาก $\mathrm{IQR} = \mathrm{Q}_3 - \mathrm{Q}_1$ โดยที่ Q_1 คือค่าข้อมูลที่อยู่ในควอไทล์ที่ 1 และ Q_3 คือค่าข้อมูลที่อยู่ในควอไทล์ที่ 3 ค่าผิดปกติ (outlier) คือค่าที่อยู่นอกช่วง $\mathrm{Q}_1 - 1.5\mathrm{IQR}$ ถึง $\mathrm{Q}_3 + 1.5\mathrm{IQR}$ ซึ่งหมายความว่าเป็นช่วงที่อยู่ห่างกันออกไปจากค่ากลางของข้อมูลอย่างสมเหตุสมผล

ในการพิจารณาการกระจายตัวของข้อมูล 1 ข้อมูล โดยพิจารณาค่ากลางทางสถิติและค่าพิสัยระหว่างควอไทล์ การวาดบ็อกซ์พล็อต (Box-plot) เป็นเครื่องหมายทางสถิติอย่างหนึ่งที่ทำให้เห็นภาพรวมของการกระจายตัวของข้อมูลได้ดี โดยบ็อกซ์พล็อตคือแผนภาพที่แสดงข้อมูลทางสถิติ ประกอบด้วย ค่ามัธยฐาน (Median) ควอไทล์ 1 (Q_1) และควอไทล์ 3 (Q_3) ค่าสูงสุดและต่ำสุดของการแจกแจง และค่าผิดปกติ ดังแสดงในรูปที่ 4.2 ซึ่งค่าผิดปกติคำนวณจากค่าที่อยู่นอกช่วง $\mathrm{Q}_1 - 1.5(\mathrm{Q}_3 - \mathrm{Q}_1)$ ถึง $\mathrm{Q}_3 + 1.5(\mathrm{Q}_3 - \mathrm{Q}_1)$ นั่นคือส่วนที่อยู่นอกเส้นประซึ่งถูกแสดงเป็นวงกลม



รูปที่ 4.2 บ็อกซ์พล็อตที่แสดงค่าผิดปกติของข้อมูล

4.3.2 ค่าผิดปกติของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

สำหรับการหาค่าผิดปกติของข้อมูลที่มีข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนี สามารถหาได้จากการคำนวณความแตกต่างระหว่างชุดข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีโดยคำนวณระยะทางมาฮาลาโนบิส (Mahalanobis distance) [8] ก่อนแล้วจึงระบุค่าระยะทางที่ผิดปกติจากค่าพิสัยระหว่างควอไทล์ได้

การหาระยะทางมาฮาลาโนบิส เป็นการหาค่าระยะทางโดยสนใจถึงความสัมพันธ์ของข้อมูลด้วยซึ่งจะต้องคำนวณเมทริกซ์ความแปรปรวนร่วม (Variance-covariance matrix) โดยเป็นค่าระยะทางระหว่างตัวอย่างข้อมูลไปยังจุดศูนย์กลางของข้อมูล สามารถคำนวณได้ดังนี้

$$\mathrm{MD}(x) = \sqrt{(x-\mu)^T C^{-1}(x-\mu)} \tag{4.10}$$

โดยที่ตัวอย่างของข้อมูล $x = (x_1, x_2, \dots, x_M)^T$ มาจากเซตของตัวอย่างข้อมูลที่มีค่าเฉลี่ย $\mu = (\mu_1, \mu_2, \dots, \mu_M)^T$ สำหรับข้อมูล M มิติ และ C คือเมทริกซ์ความแปรปรวนร่วมของตัวอย่างข้อมูล

สำหรับตัวอย่างข้อมูลที่มีค่าระยะทางมาฮาลาโนบิส ที่ห่างจากจุดศูนย์กลางอย่างผิดปกติจากค่าพิสัยระหว่างควอไทล์ ถูกกำหนดให้เป็นค่าผิดปกติของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

4.4 การแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด (Min-Max Normalization)

สำหรับการจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนี ข้อมูลที่ถูกตัดค่าผิดปกติออกจะถูกแปลงข้อมูล เพื่อให้ข้อมูลสรีรวิทยาทุกค่าอยู่ในช่วงเดียวกันในการประเมินดัชนีที่เหมาะสมคืออยู่ในช่วง 0 ถึง 1 โดยใช้การแปลงข้อมูลให้อยู่ในช่วงที่กำหนดด้วยการกำหนดค่าต่ำสุดและสูงสุด ซึ่งเป็นการแปลงข้อมูลเชิงเส้นให้ข้อมูลอยู่ในช่วงใหม่ที่ต้องการยังคงมีการกระจายตัวเดิมของข้อมูล โดยคำนวณได้จาก

$$x_{\text{new}} = \frac{x - \min(x)}{\max(x) - \min(x)} \{ \max_{\text{new}} - \min_{\text{new}} \} + \min_{\text{new}} \qquad (4.11)$$

โดยที่ $x = (x_1, x_2, \dots, x_M)^T$ เป็นข้อมูลที่ต้องการแปลง

$\min(x)$ เป็นค่าที่น้อยที่สุดของ x

$\max(x)$ เป็นค่าที่มากที่สุดของ x

x_{new} เป็นค่าที่ได้จากการแปลงข้อมูล

$\min_{\text{new}}$ เป็นค่าที่น้อยที่สุดที่ต้องการของ x_{new}

$\max_{\text{new}}$ เป็นค่าที่มากที่สุดที่ต้องการของ x_{new}

สำหรับการจัดกลุ่มข้อมูลโดยใช้ค่าทางสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีจะใช้ข้อมูลดิบในการจัดกลุ่ม การแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดกับชุดข้อมูลที่มีหลายคุณลักษณะ อาจทำให้การกระจายตัวของข้อมูลแต่ละคุณลักษณะที่อาจจะ

สอดคล้องกันหรือแตกต่างเกิดความผิดพลาดไปมากจากข้อมูลเดิม การหาการแปลงข้อมูลหรือการทำรีสเกล (rescale) ข้อมูลจึงเป็นเรื่องสำคัญและควรศึกษาต่อไป ในวิทยานิพนธ์นี้จึงใช้การแปลงข้อมูลที่ได้รับการแปรงด้วยดัชนี ให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดกับทุกข้อมูลเพื่อการประเมินดัชนีที่เหมาะสมสำหรับการจัดกลุ่มแบบมีหกคุณลักษณะรวมอยู่ด้วยกันเท่านั้น เนื่องจากตัวประเมินการจัดกลุ่มจะทำงานได้ดีสำหรับข้อมูลที่มีค่าอยู่ในช่วง 0 ถึง 1 จึงควรแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดก่อน

4.4 การวิเคราะห์การเกาะกลุ่ม (Clustering analysis)

การวิเคราะห์การเกาะกลุ่ม คือ วิธีการจัดกลุ่มวัตถุต่าง ๆ ตามคุณลักษณะ โดยวัตถุที่อยู่ในกลุ่มเดียวกันจะมีความคล้ายกันมากกว่าวัตถุที่อยู่ต่างกลุ่ม การวัดความคล้ายที่นิยมใช้ในการวิเคราะห์การจัดกลุ่มคือการวัดด้วยระยะทาง ซึ่งในวิทยานิพนธ์นี้เลือกใช้ระยะทางแบบยูคลิดและใช้การจัดกลุ่มแบบลำดับชั้น โดยที่ข้อมูลที่มีค่าใกล้กันมากจะมีระยะทางระหว่างกันน้อย

4.4.1 การจัดกลุ่มแบบลำดับชั้น (Hierarchical clustering)

การจัดกลุ่มคือการจัดตัวอย่างข้อมูลที่มีความคล้ายกันให้อยู่กลุ่มเดียวกัน โดยวิธีการที่ใช้ในการจัดกลุ่มคือ การจัดกลุ่มแบบลำดับชั้นซึ่งเป็นการแบ่งข้อมูลเป็นลำดับชั้น การจัดกลุ่มแบบลำดับชั้นเป็นการแบ่งข้อมูลเป็นลำดับชั้น โดยไม่จำเป็นต้องกำหนดจำนวนกลุ่มก่อน โดยขั้นตอนการแบ่งกลุ่มที่นิยมเป็นแบบอักโกลเมราทีฟ (Agglomerative) คือการแบ่งกลุ่มที่เริ่มจากการพิจารณาแต่ละข้อมูล ข้อมูลที่มีความคล้ายกันมากกว่า จะถูกจับรวมกันก่อนเป็นลำดับชั้น ขั้นตอนการรวมกลุ่มจะดำเนินไปเรื่อยๆ จนกระทั่งข้อมูลทั้งหมดรวมกันเป็นกลุ่มเดียว

เมื่อพิจารณาความคล้ายระหว่างกลุ่มในเชิงระยะทาง วิธีการที่ใช้ในการคำนวณระยะทางระหว่างกลุ่มที่นิยมใช้มีอยู่หลายแบบ ในวิทยานิพนธ์นี้เลือกใช้วิธีของวาร์ด (Ward) ซึ่งเป็นวิธีการที่ทำให้ค่าความแปรปรวนภายในกลุ่มมีค่าน้อยที่สุด ซึ่งคำนวณโดย

$$d(c_i, c_j) = \frac{n_i n_j}{n_i + n_j} \|c_i - c_j\|^2 \qquad (4.12)$$

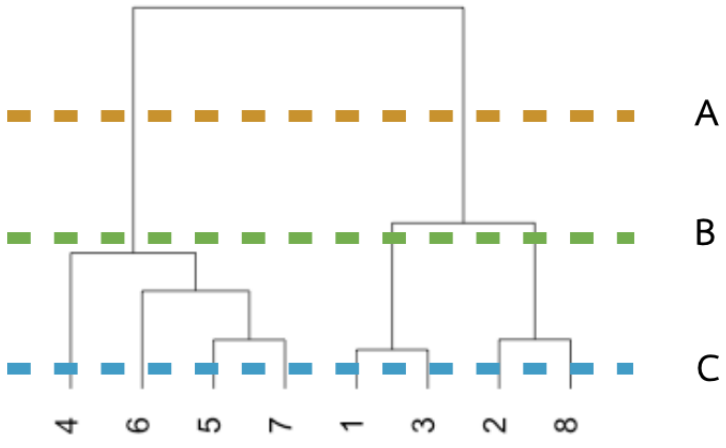
โดยที่ n_i และ n_j เป็นจำนวนสมาชิกภายในกลุ่มที่ i และ j ตามลำดับ

c_i และ c_j เป็น เซนทรอยด์ (Centroid) ของกลุ่มที่ i และ j ตามลำดับ

$\|\cdot\|$ เป็นนอร์มแบบยุคลิด (Euclidean norm)

ผลลัพธ์การแบ่งกลุ่มที่ได้ จะถูกนำเสนอในรูปของพงศาวลี (Dendrogram)

โดยการจัดกลุ่มแบบลำดับชั้นสามารถเลือกจำนวนกลุ่มที่เหมาะสม ได้จากการตัดพงศาวลี ดังแสดงรูปที่ 4.3 ตัดพงศาวลี โดยใช้เส้น A B และ C ก็จะได้ 2 กลุ่ม 3 กลุ่ม และ 8 กลุ่ม ตามลำดับ



รูปที่ 4.3 พงศาวลีจากการจัดกลุ่มแบบลำดับชั้น

การจัดกลุ่มพันธุ์ข้าวของทั้ง 9 ดัชนี โดยแบ่งเป็น 2 แบบ แบบแรกจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี แล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด แบบที่สองจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

4.5 การประเมินจำนวนกลุ่มในการจัดกลุ่ม

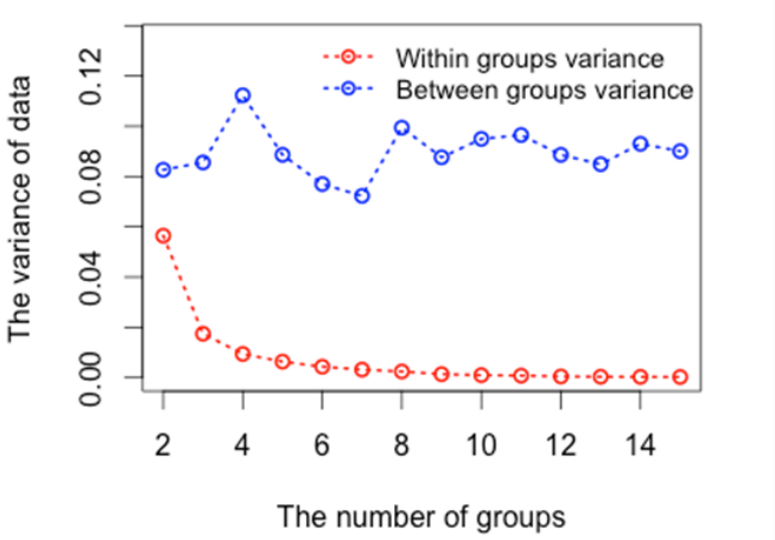
เนื่องจากการจัดกลุ่มแบบลำดับชั้นเป็นการจัดกลุ่ม ที่เลือกจำนวนกลุ่มที่เหมาะสมที่ต้องการจัดได้เอง ดังนั้นจึงจำเป็นที่จะต้องมีตัวประเมินในการกำหนดการจัดกลุ่มที่เหมาะสมซึ่งการประเมินจำนวนกลุ่มแบ่งออกเป็น 2 แบบดังนี้

4.5.1 การประเมินจำนวนกลุ่มในการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

การจัดกลุ่มที่ดีนั้นค่าความแปรปรวนภายในกลุ่มควรจะน้อย และค่าความแปรปรวนระหว่างกลุ่มควรจะมาก [9] ดังนั้นการหาจำนวนกลุ่มที่เหมาะสมเราจะหาค่าความแปรปรวนภายในกลุ่มที่น้อยที่สุด หรือค่าความแปรปรวนระหว่างกลุ่มที่มากที่สุด กราฟของค่าความแปรปรวนภายในกลุ่มและระหว่างกลุ่มของข้อมูลตัวอย่างจะถูกแสดงในรูปที่ 4.4 สังเกตว่าความแปรปรวนภายในกลุ่มจะลดลงเรื่อยๆในขณะที่จำนวนกลุ่มที่แบ่งเพิ่มขึ้นเสมอ และค่าความแปรปรวนระหว่างกลุ่มสามารถหาค่าที่สูงสุดได้

การกำหนดจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่มโดยใช้ข้อมูลสรีรวิทยา เพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีเลือกจากค่าความแปรปรวนระหว่างกลุ่ม ที่ได้จากการจัดกลุ่มที่มากที่สุด ค่าความแปรปรวนระหว่างกลุ่มยิ่งมากนั้นบ่งบอกถึงผลการจัดกลุ่มที่ได้ในแต่ละกลุ่มยิ่งมีความแตกต่างกัน



รูปที่ 4.4 กราฟแสดงค่าความแปรปรวนภายในกลุ่มและระหว่างกลุ่มของการจัดกลุ่มตั้งแต่ 2 ถึง 15 กลุ่ม

4.5.2 การประเมินจำนวนกลุ่มในการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีโดยการใช้ Pakhira-Bandyopadhyay-Maulik (PBM)

การกำหนดจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่ม โดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีใช้วิธี Pakhira-Bandyopadhyay-Maulik (PBM) ที่ถูกนำเสนอโดย Pakhira และคณะ [10] ในปี 2004 โดย PBM ใช้กำหนดจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่มซึ่งสามารถกำหนดจำนวนกลุ่มได้ดีสำหรับข้อมูลที่สร้างขึ้น และข้อมูลจริงจากฐานข้อมูล UCI ที่ข้อมูลแต่ละชุดมีการจัดประเภทกลุ่มของข้อมูล

ค่า PBM สำหรับ K กลุ่มถูกกำหนดให้เป็นดังนี้

$$\mathrm{PBM}(K) = \left(\frac{1}{K} \times \frac{E_1}{E_K} \times D_K\right)^2 \tag{4.13}$$

โดยที่

$$E_K = \sum_{k=1}^{K}\sum_{j=1}^{n} u_{kj}\|x_j - z_k\| \tag{4.14}$$

$$E_1 = \sum_{j=1}^{n}\|x_j - z_k\| \tag{4.15}$$

และ

$$D_K = \max_{i,j=1,2,\ldots,K}\|z_i - z_j\| \tag{4.16}$$

โดยที่ n เป็นจำนวนตัวอย่างของข้อมูล

m เป็นจำนวนลักษณะของข้อมูล

$X = \{x_1, x_2, \ldots, x_n\}$ เป็นตัวอย่างของข้อมูล โดยที่ $x_j \in \mathbb{R}^m, j = 1,2,\ldots,n$

$\mathrm{U}(X) = \left[u_{kj}\right]_{K\times n}$ เป็นเมทริกซ์ย่อยของชุดข้อมูล X ถ้า $x_j \in C_k$ แล้ว $u_{kj} = 1$ มิฉะนั้น $u_{kj} = 0$

z_k เป็นกึ่งกลางของกลุ่มที่ k

$\|\cdot\|$ เป็นนอร์มแบบยุคลิด (Euclidean norm)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ค่า PBM ที่มากที่สุดจะเป็นตัวกำหนดจำนวนกลุ่ม ค่า PBM ประกอบด้วย 3 ส่วน คือ $1/\mathrm{K}$, E_1/E_K และ D_K สำหรับส่วนแรกจะลดลงขณะที่ K เพิ่มขึ้น สำหรับส่วนที่สองซึ่งประกอบด้วย E_1 เป็นค่าคงที่ซึ่งไม่ขึ้นกับจำนวนกลุ่ม และ E_K ขณะที่ K เพิ่มขึ้น ระยะห่างของสมาชิกภายในกลุ่มจะน้อยลง นั่นคือสมาชิกในกลุ่มอยู่ชิดกันมากขึ้นจึงทำให้ E_K มีค่าลดลงแต่ E_1 เป็นค่าคงที่ ดังนั้นส่วนที่สองเพิ่มขึ้นขณะที่ K เพิ่มขึ้นด้วย สำหรับส่วนที่สามเป็นระยะห่างระหว่างจุดกึ่งกลางของกลุ่มข้อมูลที่มากที่สุด ซึ่งเป็นค่าที่มีขอบเขตคือค่าระยะห่างของข้อมูลที่มากที่สุด ค่าส่วนที่สองและสามเพิ่มขึ้นเมื่อ K เพิ่มขึ้นในขณะที่ส่วนที่หนึ่งลดลงไปเรื่อย ๆ เนื่องจากต้องการทำให้ได้จำนวนกลุ่มที่น้อยที่สุดเท่าที่จะเป็นไปได้ในขณะที่เพิ่มความหนาแน่นภายในกลุ่มและระยะห่างระหว่างกลุ่มมากที่สุดเท่าที่จะเป็นไปได้

4.6 การประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยา

หลังจากทำการจัดกลุ่มโดยมีจำนวนกลุ่มมาจากการประเมิน ด้วยค่าความแปรปรวนระหว่างกลุ่มสำหรับการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ และค่า PBM สำหรับการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยา 6 คุณลักษณะ จะนำค่าที่ใช้ประเมินการจัดกลุ่มเป็นตัวประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยา โดยแบ่งการประเมินดัชนีที่เหมาะสมเป็น 2 แบบ

4.6.1 ประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยาจากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

ในการประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยาจากการจัดกลุ่ม ที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี ใช้ค่าความแปรปรวนระหว่างกลุ่มจากการประเมินการจัดกลุ่ม เป็นตัวประเมินดัชนีที่เหมาะสมสำหรับของแต่ละสรีรวิทยาดัชนีที่ทำให้ค่าความแปรปรวนระหว่างกลุ่ม ที่ได้จากการประเมินการจัดกลุ่มมากผิดปกติเป็นตัวกำหนดดัชนีเหมาะสม ดังนั้นบางสรีรวิทยาอาจไม่สามารถระบุดัชนีที่เหมาะสมได้ ถ้าดัชนีทำให้ค่าความแปรปรวนระหว่างกลุ่มจากการประเมินการจัดกลุ่มไม่ปรากฏค่ามากผิดปกติ

4.6.2 ประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยาจากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

ในการประเมินดัชนีที่เหมาะสมของแต่ละสรีรวิทยา จากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่คละกันไป ดัชนีที่ใช้แปลงข้อมูลสรีรวิทยาหกคุณลักษณะซึ่งทำให้ค่า PBM จากการจัดกลุ่มด้วยจำนวนกลุ่มที่เหมาะสมที่คำนวณจากข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนี และที่ได้รับ

ความเครียดจากความเค็มโดยไม่ได้รับการแปลงด้วยดัชนี แล้วจึงแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดมีค่ามากที่สุด จะถูกกำหนดให้เป็นดัชนีเหมาะสม

ค่า PBM ของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับความเครียดจากความเค็มโดยไม่ได้รับการแปลงด้วยดัชนี และแปลงด้วยดัชนีที่คละกันไปจะถูกนำมาวิเคราะห์ดังนี้

1. ค่า PBM ของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับความเครียดจากความเค็ม โดยไม่ได้รับการแปลงด้วยดัชนี จะประเมินร่วมกับดัชนีที่ได้รับการแปลงด้วยดัชนี YI ที่ใช้แค่ข้อมูลสรีรวิทยาที่ได้รับความเครียดจากความเค็ม เพื่อตรวจสอบว่า การแปลงข้อมูลด้วยดัชนี YI ทำให้การจัดกลุ่มดีขึ้นหรือไม่

2. ค่า PBM ของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของสรีรวิทยานั้น ๆ จากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะจะประเมินร่วมกับการแปลงด้วยดัชนีที่เหมือนกัน เพื่อตรวจสอบว่าการแปลงข้อมูลลักษณะใดทำให้การจัดกลุ่มดีกว่ากัน

3. ค่า PBM ของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของสรีรวิทยานั้น ๆ จากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะจะประเมินร่วมกับการแปลงด้วยดัชนี ที่ทำให้การจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีมีค่า PBM มากที่สุด เพื่อตรวจสอบว่า การแปลงข้อมูลได้รับการแปลงด้วยดัชนีที่เหมาะสมของสรีรวิทยานั้น ๆ จากการจัดกลุ่มที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะทำให้การจัดกลุ่มดีที่สุดหรือไม่

4.7 การทดสอบแมนน์-วิทนีย์ ยู (Mann – Whitney U Test)

เนื่องจากการประเมินผลการจัดกลุ่มใช้ข้อมูล ที่มีลักษณะไม่ทราบการแจกแจงของประชากร จึงใช้การทดสอบที่ไม่ใช้พารามิเตอร์ ซึ่งการทดสอบนี้สนใจค่ามัธยฐานมากว่าค่าเฉลี่ย [11] โดยใช้สถิติทดสอบแมน-วิทนีย์ ยู ซึ่งทดสอบความเท่ากันของค่ามัธยฐานของประชากร การทดสอบแมนน์-วิทนีย์ ยู เป็นการทดสอบที่ข้อมูลที่ใช้ไม่จำเป็นต้องเป็นการแจกแจงแบบปกติ กลุ่มข้อมูลเป็นอิสระกัน และจำนวนข้อมูลของทั้งสองกลุ่มไม่จำเป็นต้องเท่ากัน โดยการทดสอบนี้มีวิธีการคือ รวมค่าทุก ๆ กลุ่มข้อมูล จัดอันดับข้อมูลทั้งหมด จากนั้นแยกข้อมูลตามเดิมซึ่งเป็นข้อมูลที่มีอันดับติดมาด้วยแล้วจึงทดสอบสมมติฐานความเท่ากันของค่ามัธยฐานของประชากร

สมมติฐานในการทดสอบคือ

H_0 : ค่ามัธยฐานของประชากร 2 กลุ่ม เท่ากัน

H_1 : ค่ามัธยฐานของประชากร 2 กลุ่ม แตกต่างกัน

ประเมินการจัดกลุ่มโดยใช้การทดสอบแมนน์-วิทนีย์ ยู จะถูกประเมินกลุ่มแต่ละคู่ซึ่งผลที่ได้จากการประเมินเป็นจะถูกแสดงอยู่รูปของตารางที่แสดงค่า p-value ของการประเมินกลุ่มแต่ละคู่ โดยค่า p-value ที่น้อยกว่าหรือเท่ากับ 0.05 แสดงว่าสองกลุ่มมีค่ามัธยฐานแตกต่างกันที่ระดับความเชื่อมั่น 95%

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.8 เกณฑ์การบ่งชี้ลักษณะความทนเค็มของกลุ่มจากการจัดกลุ่มข้อมูลสรีรวิทยาหกลักษณะที่ได้รับการแปลงด้วยดัชนีด้วยดัชนีที่เหมาะสม

ผลการจัดกลุ่มข้อมูลสรีรวิทยาหกลักษณะด้วยจำนวนกลุ่มและดัชนีที่เหมาะสมที่ได้ จะนำมาวิเคราะห์ความทนเค็มของกลุ่มแต่ละกลุ่มจากพันธุ์ข้าวที่ระบุลักษณะทนเค็มแล้ว วิทยานิพนธ์นี้สามารถกำหนดลักษณะความทนเค็มของกลุ่มได้เฉพาะกลุ่มที่มีข้าวที่ระบุลักษณะทนเค็มแล้ว และสามารถสรุปได้ว่าข้าวในกลุ่มนั้นจะมีลักษณะที่คล้ายคลึงกันกับข้าวที่สามารถระบุลักษณะทนเค็มที่อยู่ในกลุ่มนั้น โดยข้อมูลทั้ง 2 ชุดมีพันธุ์ข้าวที่ระบุลักษณะทนเค็มที่แตกต่างกันออกไป

ชุดที่ 1 กลุ่มที่มี Pokkali ซึ่งเป็นข้าวพันธุ์ทนทานต่อความเค็ม (Salt Tolerance ; ST) จะกำหนดให้เป็นกลุ่มที่มีลักษณะทนทานต่อความเค็ม และกลุ่มที่มี IR29 ซึ่งเป็นข้าวพันธุ์อ่อนแอต่อความเค็ม (Salt Susceptible ; SS) จะกำหนดให้เป็นกลุ่มที่มีลักษณะอ่อนแอต่อความเค็ม

ชุดที่ 2 กลุ่มที่มี KDML105 และ CSSL11 ซึ่งเป็นข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง (Moderately Salt Tolerance ; MT) จะกำหนดให้เป็นกลุ่มที่มีลักษณะทนทานต่อความเค็มระดับปานกลาง และ ปทุมธานี1 เป็นข้าวพันธุ์อ่อนแอต่อความเค็ม (Salt Susceptible ; SS) จะกำหนดให้เป็นกลุ่มที่มีลักษณะอ่อนแอต่อความเค็ม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 5
ผลการวิจัย

5.1 ข้อมูลชุดที่ 1

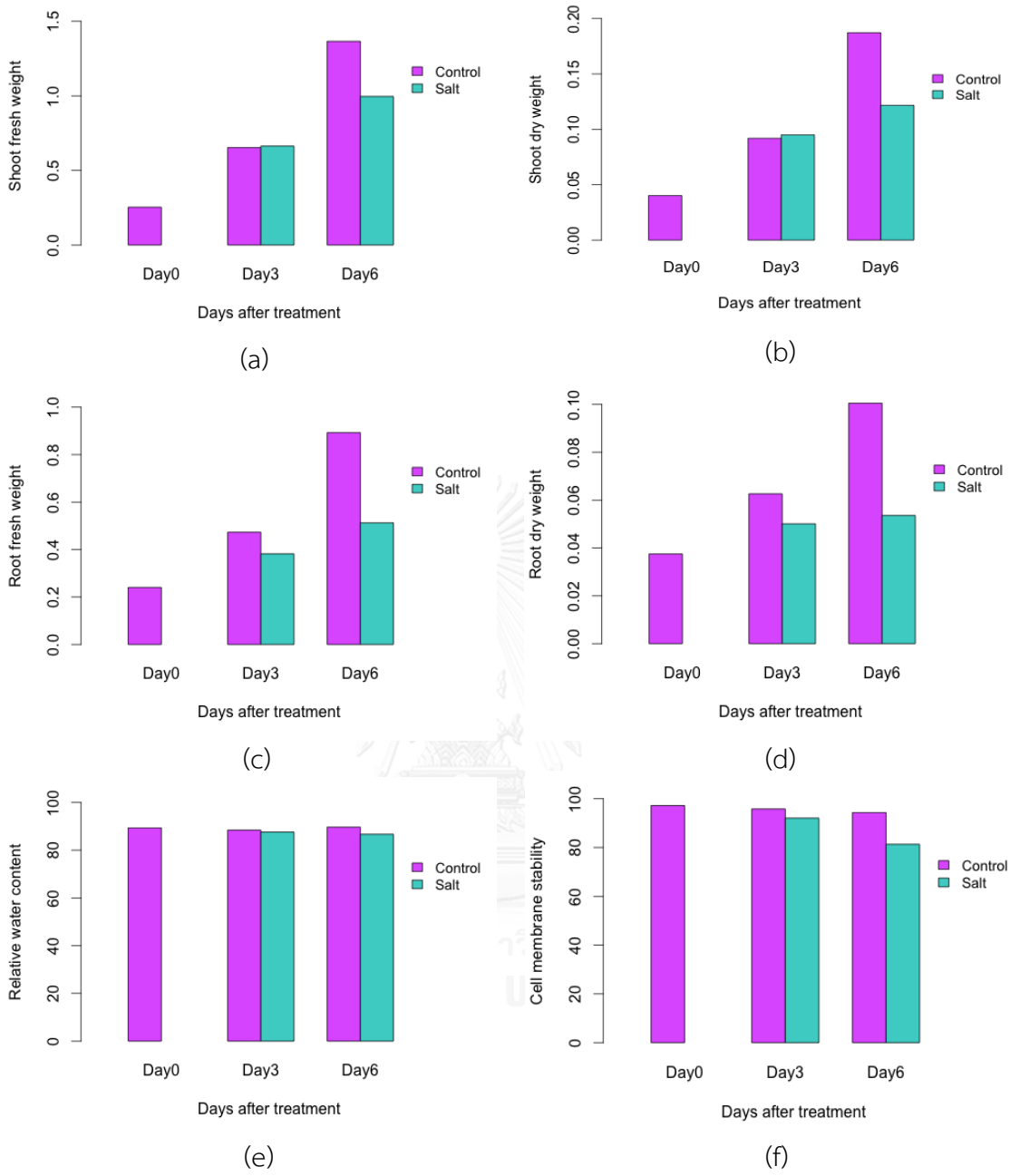
5.1.1 การเลือกวันที่นำมาใช้ทดลอง

จากการเปรียบเทียบ $\bar{Y}_c$ และ $\bar{Y}_s$ ดังแสดงในรูปที่ 5.1 พบว่าค่าเฉลี่ยของ SFW และ SDW ในช่วงวันที่ 3 มี Y_s มากกว่า Y_c เล็กน้อย ค่าเฉลี่ยของ SFW SDW RFW RDW RWC และ CMS ในช่วงวันที่ 6 มี Y_s น้อยกว่า Y_c และเริ่มเห็นความแตกต่างกันอย่างชัดเจนระหว่าง Y_c กับ Y_s จากตารางที่ 5.1 พบว่าค่าเฉลี่ยของ Y_c และ Y_s ของข้อมูลสรีรวิทยาหกคุณลักษณะ ของข้อมูลชุดที่ 1 ในช่วงวันที่ 6 มีความแตกต่างกันอย่างมีนัยสำคัญโดยมีค่า p-value ของการทดสอบทางสถิติด้วย Paired t - test น้อยกว่า 0.05 จากการพิจารณาข้างต้น ช่วงเวลาที่ถูกเลือกเพื่อนำไปใช้ในงานวิจัยนี้คือ ช่วงวันที่ 6

Chulalongkorn University

ตารางที่ 5.1 ตารางแสดงค่า p-value ของการทดสอบทางสถิติด้วย Paired t - test เพื่อเปรียบเทียบข้อมูลสรีรวิทยาของข้อมูลชุดที่ 1 ระหว่าง Y_c กับ Y_s ของ SFW SDW RFW RDW RWC และ CMS

	SFW	SDW	RFW	RDW	RWC	CMS
Day 3	0.7236	0.6542	2.124e-05	0.05375	0.2154	0.0001412
Day 6	4.891e-7	5.396e-10	6.661e-12	3.711e-11	0.004463	1.066e-9



รูปที่ 5.1 แผนภูมิแท่งเปรียบเทียบความแตกต่างของค่าเฉลี่ยทางสรีรวิทยาของข้อมูลชุดที่ 1 (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ระหว่าง Y_c กับ Y_s

5.1.2 การหาค่าผิดปกติ (Outliers)

ถ้าข้อมูลมีค่าผิดปกติจะทำให้การวิเคราะห์ข้อมูลอาจเกิดความผิดพลาดได้ ดังนั้นก่อนการจัดกลุ่มของข้อมูลทั้งแบบที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ และข้อมูลสรีรวิทยาหลายคุณลักษณะจึงควรตัดค่าผิดปกติออกไป เพื่อหลีกเลี่ยงการจัดกลุ่มที่สมาชิกในกลุ่มมีเพียงตัวเดียว สำหรับข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะค่าผิดปกติแบ่งเป็น 2 แบบ คือ ค่าที่สูงกว่าปกติและค่าที่ต่ำกว่าปกติ สำหรับข้อมูลสรีรวิทยาหกคุณลักษณะค่าผิดปกติเป็นค่าที่มีระยะทางมาฮาลาโนบิสสูงกว่าปกติ สำหรับข้อมูลทั้งสองลักษณะจะแสดงอยู่ในรูปของตาราง ในภาคผนวก ค ดังนี้ ตารางที่ ค.1 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่สูงกว่าปกติ ตารางที่ ค.2 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่ต่ำกว่าปกติ และตารางที่ ค.3 ตารางแสดงพันธุ์ข้าวข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีที่เหมือนกันและดัชนีที่ทำให้ค่า PBM มากที่สุด แล้วให้ค่าระยะทางมาฮาลาโนบิสที่สูงกว่าปกติ

พิจารณาตารางที่ 5.2 โดยรวมแล้วค่าน้ำหนักทั้ง 4 ค่าที่ได้รับการแปลงด้วยดัชนีเมื่อพิจารณาทุกดัชนี พบว่าพันธุ์ที่ให้ค่าผิดปกติซ้ำกันมากที่สุดคือ หลวงประทาน สำหรับ RWC และ CMS ที่ได้รับการแปลงด้วยดัชนีเมื่อพิจารณาทุกดัชนี พบว่าพันธุ์ที่ให้ค่าผิดปกติซ้ำกันมากที่สุดที่เหมือนกันคือ กข1

ตารางที่ 5.2 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ เป็นค่าผิดปกติซ้ำกันมากที่สุด

พันธุ์ที่ให้ค่าผิดปกติ	สภาวะปกติ		สภาวะได้รับความเครียดจากความเค็ม		แปลงด้วยดัชนีต่าง ๆ เป็นค่าผิดปกติซ้ำกันมากที่สุด
	สูง	ต่ำ	สูง	ต่ำ	
SFW	เจ้าขาว, หลวงประทาน	-	หลวงประทาน	-	หลวงประทาน (5 ดัชนี)
SDW	เจ้าขาว, หลวงประทาน	-	หลวงประทาน	-	หลวงประทาน (5 ดัชนี)
RFW	เจ้าขาว, หลวงประทาน	-	เขียวหางม้า, ทองรากไทร	-	หลวงประทาน (5 ดัชนี)
RDW	เจ้าขาว, หลวงประทาน	-	ทองรากไทร, มันวัว, เขียวหางม้า, อีขาวใหญ่	-	หลวงประทาน (4 ดัชนี)
RWC	-	ขาวอากาศ	-	กข1	กข1 (ทุกดัชนี)
CMS	-	ก้นแก้ว (6158), หางหยี71, พวงหางหมู, เจ็กเชย	-	กข1, IR29	กข1 (ทุกดัชนี), IR29 (ทุกดัชนี)

สำหรับข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่คละกันไปมีทั้งหมด 9^6 = 531,441 ชุดซึ่งข้อมูลแต่ละชุดมีโอกาสที่มีพันธุ์ที่ให้ค่าผิดปกติแตกต่างกันไป ในวิทยานิพนธ์นี้จึงเลือกใช้การหาพันธุ์ที่ให้ค่าผิดปกติโดยเลือกจากพันธุ์ที่มีระยะทางมาฮาลาโนบิสห่างจากจุดศูนย์กลางมากผิดปกติ

พิจารณาตารางที่ 5.3 สำหรับข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมือนกันโดยพิจารณาทุกดัชนี พันธุ์ที่ให้ค่าระยะทางมาฮาลาโนบิสสูงผิดปกติซ้ำกันมากที่สุดคือ เหลืองน้อย 31-1-39 (4 ดัชนี) อีขาวใหญ่ (4 ดัชนี) และ IR29 (4 ดัชนี) โดยข้อมูลสรีรวิทยาที่โดดเด่นของพันธุ์ก้นแก้ว เหลืองน้อย 31-1-39 อีขาวใหญ่ และ IR29 มีดังนี้

ก้นแก้ว ในสภาวะปกติ SFW และ SDW สูงสุดอันดับสาม RWC สูงสุดอันดับสอง และ CMS ต่ำสุดและเป็นค่าผิดปกติด้วย

เหลืองน้อย 31-1-39 ในสภาวะปกติ CMS สูงสุดอันดับสาม ในสภาวะได้รับความเครียดจากความเค็ม RWC สูงสุด

อีขาวใหญ่ ในสภาวะปกติ RWC ต่ำสุดอันดับสาม ในสภาวะได้รับความเครียดจากความเค็ม RFW สูงสุดอันดับสาม RDW สูงสุดอันดับสี่และเป็นค่าผิดปกติด้วย RWC ต่ำสุดอันดับสอง และ CMS สูงสุดอันดับสี่

IR29 ในสภาวะปกติ RWC ต่ำสุดอันดับห้า ในสภาวะได้รับความเครียดจากความเค็ม SFW ต่ำสุดอันดับสอง RFW ต่ำสุดอันดับสี่ RWC ต่ำสุดอันดับสาม CMS ต่ำสุดและเป็นค่าผิดปกติด้วย

ตารางที่ 5.3 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนี ที่เหมือนกันเป็นค่าผิดปกติซ้ำกันมากที่สุด

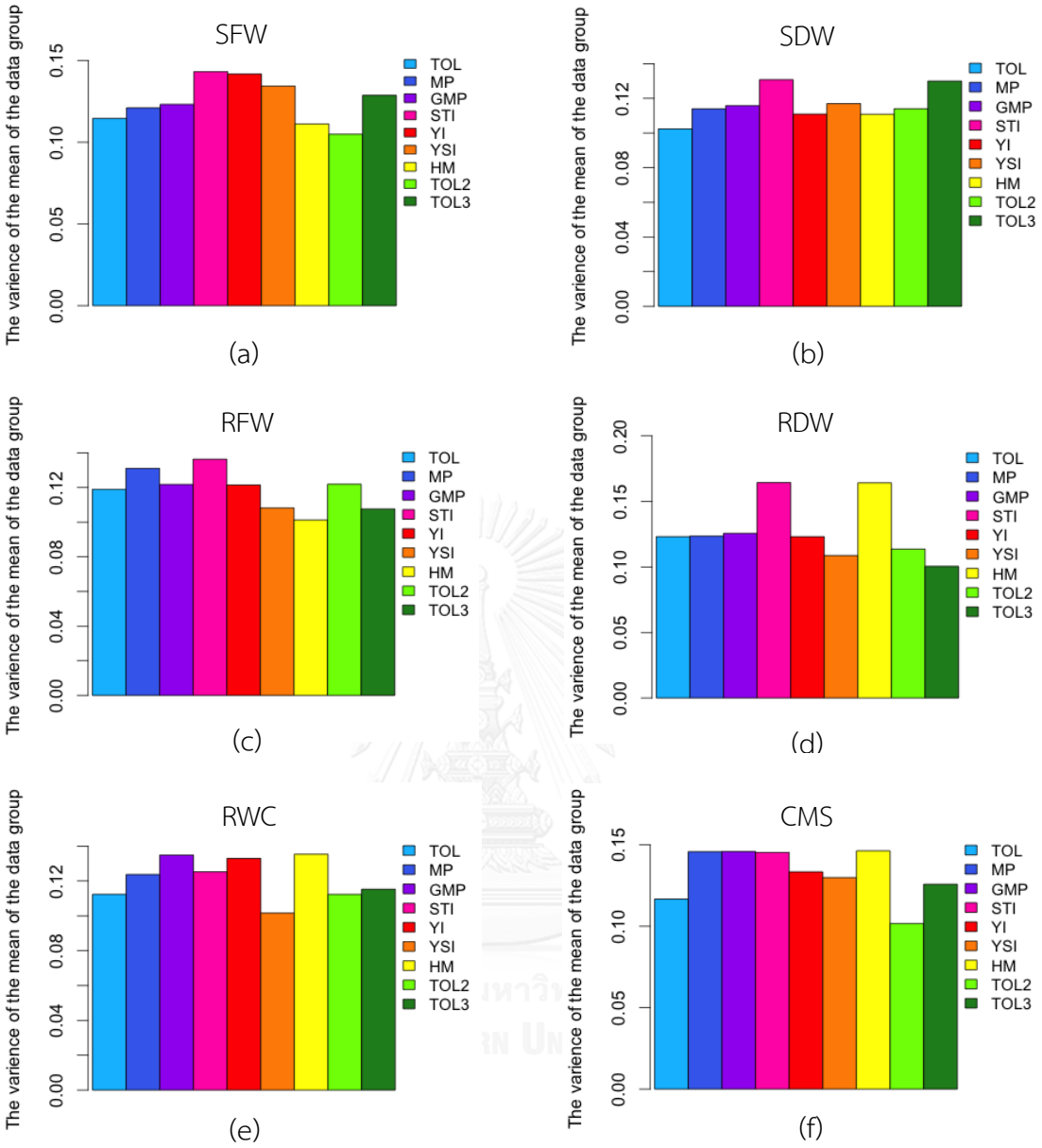
ค่าผิดปกติ	สภาวะปกติ	สภาวะได้รับความเครียดจากความเค็ม	แปลงด้วยดัชนีที่เหมือนกันเป็นค่าผิดปกติซ้ำกันมากที่สุด
ข้อมูลสรีรวิทยาหกคุณลักษณะ	ก้นแก้ว (6158), เหลืองน้อย 31-1-39	-	เหลืองน้อย 31-1-39 (4 ดัชนี), อีขาวใหญ่ (4 ดัชนี), IR29 (4 ดัชนี)

5.1.3 การหาดัชนีที่ดีในการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

สำหรับการจัดกลุ่มพันธุ์ข้าว โดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี จะใช้ข้อมูลที่ได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด ซึ่งข้อมูลสรีรวิทยาที่ได้รับความเครียดจากความเค็มและที่ได้รับการแปลงด้วยดัชนี YI เมื่อได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด มีผลทำให้ได้ค่าข้อมูลทั้งสองเหมือนกัน นั่นคือได้ผลจากการจัดกลุ่มที่เหมือนกัน การเปรียบเทียบค่าความแปรปรวนระหว่างกลุ่มจากการเลือกจำนวนกลุ่ม ที่เหมาะสมของแต่ละข้อมูลสรีรวิทยาและแต่ละดัชนี จะถูกแสดงในรูปที่ 5.2 โดยดัชนีที่ให้ค่าความแปรปรวนระหว่างกลุ่มที่มากที่สุดที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสม จากการจัดกลุ่มของแต่ละข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะมากที่สุด 3 อันดับจะถูกแสดง ในตารางที่ 5.4

การเลือกดัชนีที่เหมาะสมของแต่ละข้อมูลสรีรวิทยาจะเลือกจากค่าความแปรปรวนระหว่างกลุ่มที่มากผิดปกติ สำหรับข้อมูลชุดที่ 1 พบว่า STI และ TOL3 เป็นดัชนีที่เหมาะสมสำหรับ SDW และ STI และ HM เป็นดัชนีที่เหมาะสมสำหรับ RDW โดย SFW RFW RWC และ CMS ไม่ปรากฏดัชนีที่มีค่าความแปรปรวนระหว่างกลุ่มที่มากผิดปกติ จึงไม่สามารถสรุปดัชนีที่เหมาะสมได้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



รูปที่ 5.2 แผนภูมิแสดงการเปรียบเทียบค่าความแปรปรวนระหว่างกลุ่มจากการเลือกจำนวนกลุ่มที่เหมาะสมของแต่ละดัชนีของ (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ของข้อมูลชุดที่ 1

ตารางที่ 5.4 ตารางแสดงดัชนีที่ให้ค่าความแปรปรวนของค่าเฉลี่ยสมาชิกในกลุ่มที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าว ด้วยข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีมีค่ามากที่สุด 3 อันดับ ของข้อมูลชุดที่ 1

	อันดับที่ 1	**อันดับที่ 2**	**อันดับที่ 3**
SFW	STI	YI	YSI
SDW	STI	TOL3	YSI
RFW	STI	MP	TOL2
RDW	STI	HM	GMP
RWC	HM	GMP	YI
CMS	HM	GMP	MP

5.1.4 การหาดัชนีที่เหมาะสมในการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

เนื่องจากการประเมินจำนวนกลุ่มของข้อมูลจริงจากฐานข้อมูล UCI ที่ข้อมูลแต่ชุดมีการจัดประเภทกลุ่มของข้อมูล พบว่าการประเมินโดยค่า PBM ที่เกิดจากการจัดกลุ่มของข้อมูลดิบมีการประเมินจำนวนกลุ่มที่ดีกว่าข้อมูลที่ได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด สำหรับการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีนั้นจึงไม่ได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด แต่จะแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดเพื่อประเมินดัชนีที่เหมาะสมสำหรับการจัดกลุ่มแบบมีหกคุณลักษณะ โดยค่า PBM ที่มากที่สุดจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนี GMP GMP MP MP STI และ YSI ตามลำดับ ซึ่งไม่ปรากฏพันธุ์ที่ให้ค่าผิดปกติ โดยให้ผลการจัดกลุ่มพันธุ์ข้าวที่มีค่า PBM ที่สูงสุดแสดงในรูปที่ 5.3 ซึ่งเป็นแผนภูมิแสดงค่า PBM จากการจัดกลุ่มพันธุ์ข้าวที่ข้อมูลสรีรวิทยาหกคุณลักษณะ มาจากข้อมูลที่ได้รับความเครียดจากความเค็มโดยไม่ได้รับการแปลงด้วยดัชนี (Salt) ข้อมูลสรีรวิทยาหก

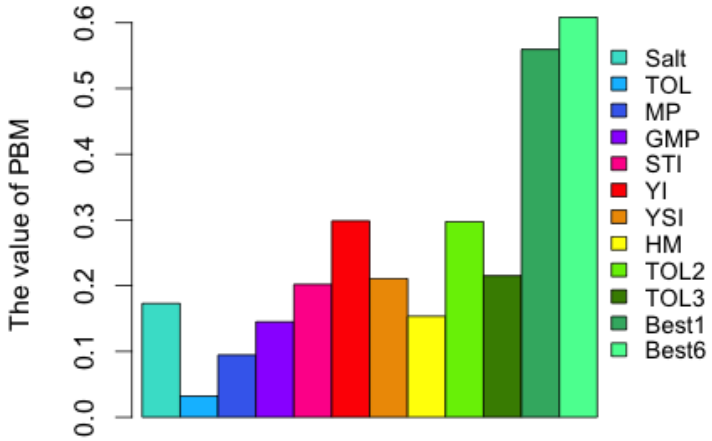
คุณลักษณะใช้ดัชนีที่เหมือนกัน (TOL MP GMP STI YI YSI HM TOL2 และ TOL3) ใช้ดัชนีที่เหมาะสมกับการจัดกลุ่มพันธุ์ข้าว ที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะเดียว (Best1) และดัชนีที่ทำให้การจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะเดียวมีค่ามากที่สุด (Best6) จากรูปนี้พบว่า

1. สำหรับจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี YI เหมือนกันให้ค่า PBM สูงกว่า Salt

2. สำหรับจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี Best1 มีค่ามากกว่า ดัชนีที่เหมือนกัน

3. สำหรับจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี Best1 น้อยกว่า Best6

ค่า PBM ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าวด้วยข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่เหมือนกันมีค่ามากที่สุด 3 อันดับจะถูกแสดงในตารางที่ 5.5



รูปที่ 5.3 แผนภูมิแสดงการเปรียบเทียบค่า PBM จากการเลือกจำนวนกลุ่มที่เหมาะสมของ Salt TOL MP GMP STI YI YSI HM TOL2 TOL3 Best1 และ Best6 ของข้อมูลชุดที่ 1

ตารางที่ 5.5 ตารางแสดงดัชนีที่ให้ค่า PBM ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าวด้วยข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่เหมือนกัน แล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดมีค่ามากที่สุด 3 อันดับของข้อมูลชุดที่ 1

	อันดับที่ 1	**อันดับที่ 2**	**อันดับที่ 3**
ทุกข้อมูลสรีรวิทยา	YI	TOL2	TOL3

5.1.5 การทดสอบการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม

จากการทดสอบทางสถิติด้วย Mann – Whitney U Test ค่า p-value ของการทดสอบทางสถิติของกลุ่มในแต่ละคู่ของข้อมูลสรีรวิทยา SDW และ RDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม จะถูกแสดงในตารางที่ 5.6 โดยจากการทดสอบพบว่า

SDW สำหรับดัชนี STI ค่า p-value ของกลุ่มที่ 5 กับกลุ่มที่ 3 และ 4 มีค่ามากกว่า 0.05 แต่น้อยกว่า 0.2 ดังนั้นพันธุ์ข้าวในกลุ่มที่ 5 มาจากประชากรที่มีค่ามัธยฐานของข้อมูลสรีรวิทยาไม่แตกต่างกับกลุ่มที่ 3 และ 4 ที่ระดับความเชื่อมั่น 95% แต่การจัดกลุ่มที่ได้กลุ่มในแต่ละคู่มีความแตกต่างระหว่างกลุ่มที่ระดับความเชื่อมั่น 80% สำหรับดัชนี TOL3 ค่า p-value ของกลุ่มที่ 4 กับกลุ่มที่ 1 2 3 5 6 และ 7 และ กลุ่มที่ 7 กับกลุ่มที่ 2 มีค่ามากกว่า 0.05 แต่น้อยกว่าหรือเท่ากับ 0.5 ดังนั้นพันธุ์ข้าวในกลุ่มที่ 4 มาจากประชากรที่มีค่ามัธยฐานของข้อมูลสรีรวิทยาไม่แตกต่างกับกลุ่มที่ 1 2 และ 3 ที่ระดับ ความเชื่อมั่น 95% แต่การจัดกลุ่มที่ได้กลุ่มในแต่ละคู่มีความแตกต่างระหว่างกลุ่มที่ระดับความเชื่อมั่น 50%

RDW สำหรับดัชนี STI และ HM ค่า p-value ของการทดสอบกลุ่มที่ 1 และ กลุ่มที่ 2 น้อยกว่า 0.05 ดังนั้นพันธุ์ข้าวในกลุ่มที่ 1 มาจากประชากรที่มีค่ามัธยฐานของข้อมูลสรีรวิทยาแตกต่างกับกลุ่มที่ 2 นั่นคือการจัดกลุ่มที่ได้มีความแตกต่างระหว่างกลุ่มที่ระดับความเชื่อมั่น 95%

ตารางที่ 5.6 ตารางแสดงค่า p-value ของการทดสอบทางสถิติของการจัดกลุ่มโดยทดสอบแต่ละคู่กลุ่มด้วย (a) ข้อมูลสรีรวิทยา SDW ที่ได้รับการแปลงด้วยดัชนี STI (b) ข้อมูลสรีรวิทยา SDW ที่ได้รับการแปลงด้วยดัชนี TOL3 และ (c) RDW ที่ได้รับการแปลงด้วยดัชนี STI (d) ข้อมูลสรีรวิทยา RDW ที่ได้รับการแปลงด้วยดัชนี HM ของข้อมูลชุดที่ 1

Cluster	2	3	4	5	6
1	1.53E-07	1.62E-04	1.47E-03	2.56E-02	8.01E-07
2		2.68E-05	4.13E-04	1.31E-02	2.95E-08
3			9.52E-03	7.14E-02	7.37E-05
4				1.33E-01	8.40E-04
5					1.90E-02

(a)

Cluster	2	3	4	5	6	7
1	2.38E-02	2.68E-05	2.86E-01	7.37E-05	4.00E-04	9.52E-03
2		2.06E-03	5.00E-01	3.57E-03	9.09E-03	5.71E-02
3			1.18E-01	2.95E-08	9.79E-07	4.13E-04
4				1.43E-01	2.00E-01	4.00E-01
5					4.02E-06	8.40E-04
6						2.80E-03

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(b)

Cluster	2	3
1	9.88E-04	5.30E-13
2		6.84E-04

(c)

Cluster	2
1	9.82E-08

(d)

5.1.6 ผลการจัดกลุ่ม

ผลการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์ โดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะ คือ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี GMP GMP MP MP STI และ YSI ตามลำดับนั้นไม่ปรากฏพันธุ์ที่ให้ค่าผิดปกติ ผลการจัดกลุ่มบอกได้เพียงว่าพันธุ์ที่อยู่ในกลุ่มเดียวกันมีลักษณะที่คล้ายคลึงกันเมื่อได้รับความเค็มเนื่องจากมีพันธุ์ที่ใช้ระบุลักษณะของพันธุ์น้อย โดยผลจากการจัดกลุ่มสามารถจากรูปที่ 5.4 พบว่า

<u>กลุ่มที่ 1</u> (ทนทานต่อความเค็ม) ประกอบด้วย Pokkali (ข้าวพันธุ์ทนทานต่อความเค็ม) ขาวกอเดียว พวงหางหมู ใบบัว หางยี71 เหนียวดำหลาย ชัยนาท1 เจ๊กเชย พ้าวฮ้าว ปิ่นแก้ว นางมล ป้องแอ้ว (21577) เม็ดมะม่วง อีมุม LPT123–TC171 มะยม (13988) พวงทอง หลวงประทาน เก้ารวง ขาวตาแห้ง17 ก้นแก้ว(1698) และเหลืองน้อย31-1-39

<u>กลุ่มที่ 2</u> ประกอบด้วย อีขาวใหญ่ อีลาย ตะเภาล่ม ดอขาว มะยม(22802) เหลืองพวงทอง ดอสามเดือน ดอดอกไม้ หมากน้ำ(22835) เจ้าขาว พลายงามปราจีนบุรี ขาวอากาศ มันวัว ก้นแก้ว(6158) ขาวก้นจุด ทองรากไทร ขาวสงวน จุดมอญ ดีสี่ เขียวหางม้า หมากยม พวงหนัก หมากน้ำ(14085) ขาวสุพรรณ จำปาจีน ประทานบ้านบุ้ง และขาวแก้ว

<u>กลุ่มที่ 3</u> มีเพียงพันธุ์เดียวคือหลวงประทาน

<u>กลุ่มที่ 4</u> (อ่อนแอต่อความเค็ม) ประกอบด้วย IR29 (ข้าวพันธุ์อ่อนแอต่อความเค็ม) กข1 และไข่มดริ้น

จากรูปที่ 5.5 ถึง รูปที่ 5.7 ซึ่งเป็น Density histogram, Density curve และค่าเฉลี่ยโดยแสดงเป็นเส้นประ ของกลุ่มที่ 1 2 และ 4 ของข้อมูลสรีรวิทยาแต่ละคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมเป็นการแสดงการกระจายตัวของข้อมูลสรีรวิทยาแต่ละคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของแต่ละกลุ่มยกเว้นกลุ่มที่ 3 เนื่องจากเป็นกลุ่มที่มีสมาชิกเพียงตัวเดียวคือ หลวงประทาน จึงไม่มีการกระจายตัวของข้อมูล

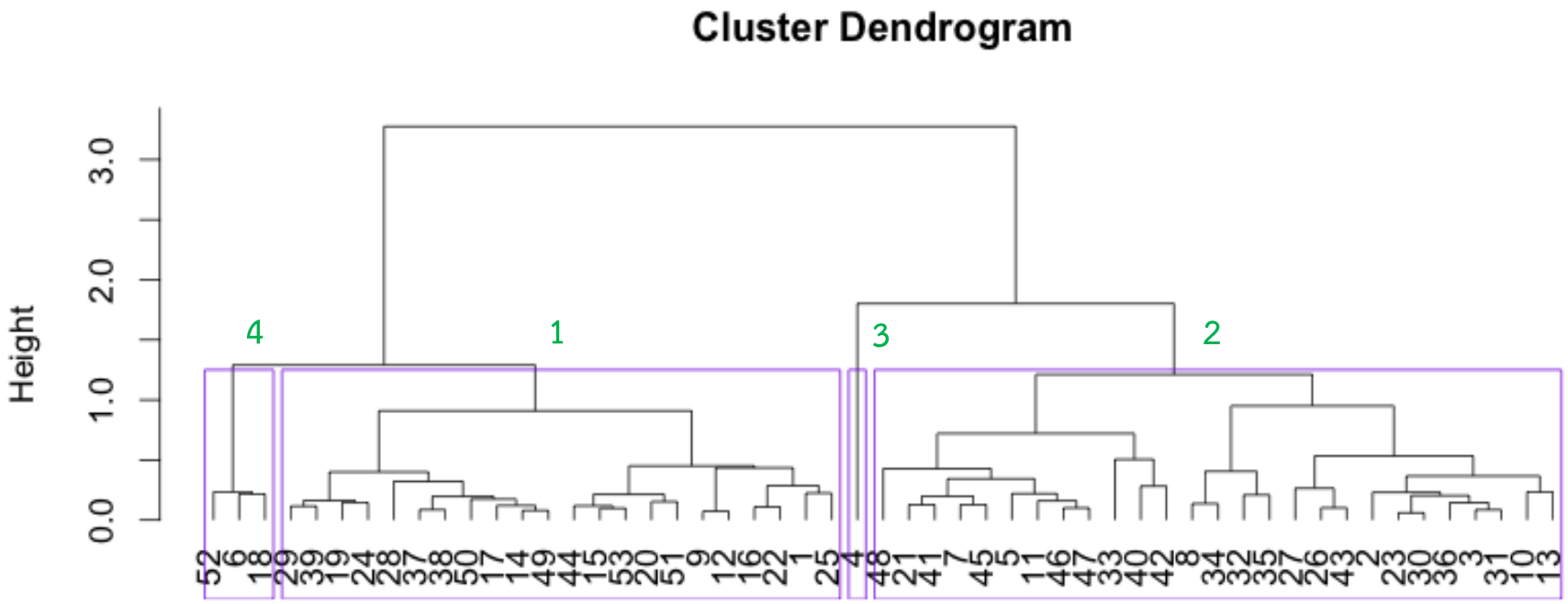
สรีรวิทยาแต่ละคุณลักษณะ และตารางที่ 5.7 แสดงค่าเฉลี่ยของกลุ่ม 1 2 3 และ 4 ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม โดยค่าเฉลี่ยของข้อมูลสรีรวิทยาแต่ละคุณลักษณะของกลุ่มที่มีค่ามากกว่าจะมีสีของช่องที่เข้มกว่า ซึ่งค่าเฉลี่ยของแต่ละกลุ่มมีลักษณะดังนี้

<u>กลุ่มที่ 1</u> เป็นกลุ่มที่ทนทานต่อความเค็มซึ่งประกอบด้วย Pokkali (ข้าวพันธุ์ทนทานต่อความเค็ม) ค่าเฉลี่ยของข้อมูลสรีรวิทยาน้ำหนัก และ CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสาม และค่าเฉลี่ยของ RWC ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสอง ดังนั้นกลุ่มที่ 1 น่าจะทนเค็มมากรองจากหลวงประทาน และกลุ่มที่ 2

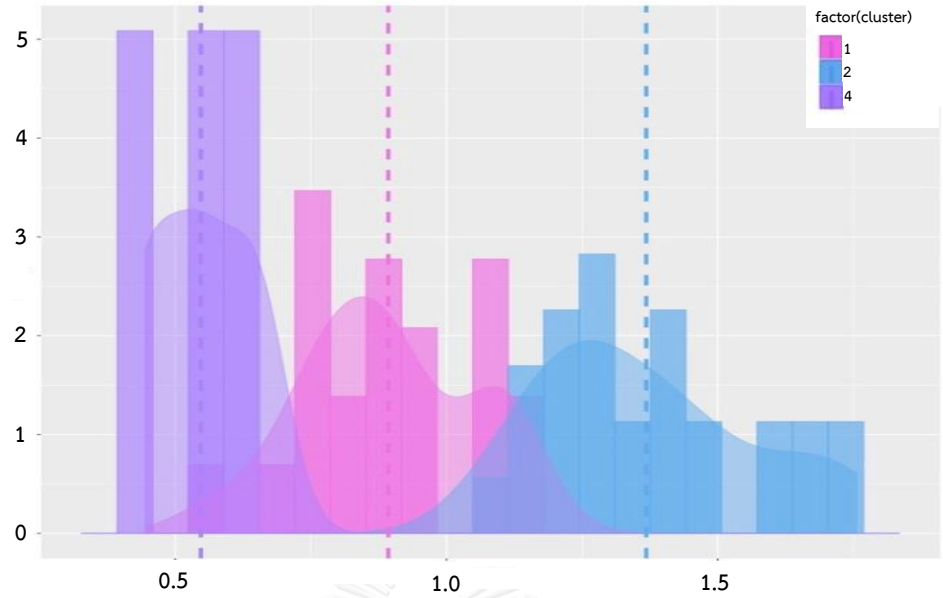
<u>กลุ่มที่ 2</u> ค่าเฉลี่ยของข้อมูลสรีรวิทยาน้ำหนัก และ CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสอง และค่าเฉลี่ยของ RWC ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสาม ดังนั้นกลุ่มที่ 2 น่าจะทนเค็มมากรองจากหลวงประทาน

<u>กลุ่มที่ 3</u> สมาชิกมีเพียงตัวเดียวคือ หลวงประทาน ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุด ดังนั้นหลวงประทานน่าจะทนเค็มมากที่สุดของข้อมูลชุดที่ 1

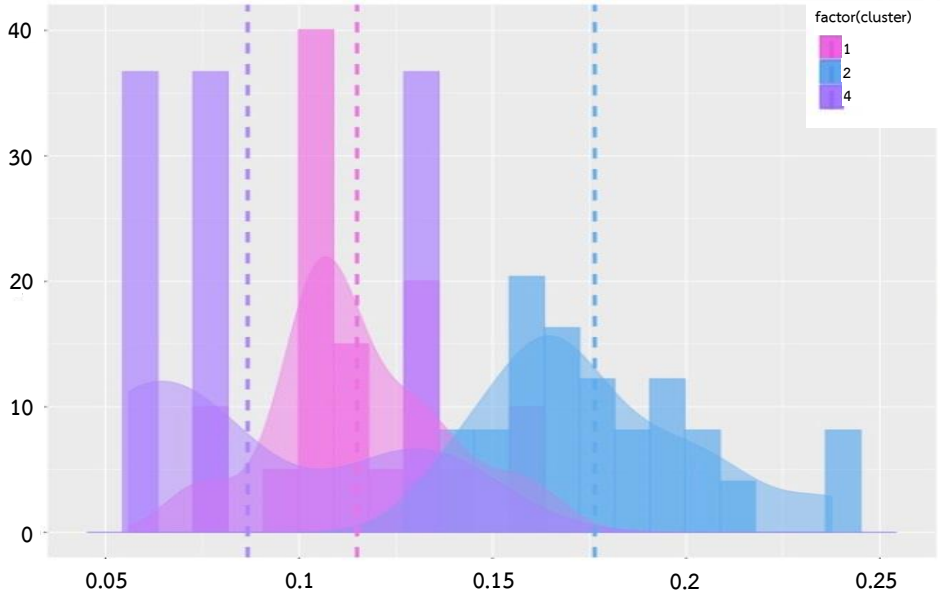
<u>กลุ่มที่ 4</u> เป็นกลุ่มที่อ่อนแอต่อความเค็มซึ่งประกอบด้วย IR29 (ข้าวพันธุ์อ่อนแอต่อความเค็ม) และมีค่าเฉลี่ยของข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่าน้อยที่สุดของข้อมูลชุดที่ 1 ดังนั้นกลุ่มที่ 4 ทนเค็มน้อยที่สุดของข้อมูลชุดที่ 1



รูปที่ 5.4 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลสรีรวิทยา SFW SDW RFW RDW RWC และ CMS ด้วยโดยดัชนี GMP GMP MP MP STI และ YSI ตามลำดับ

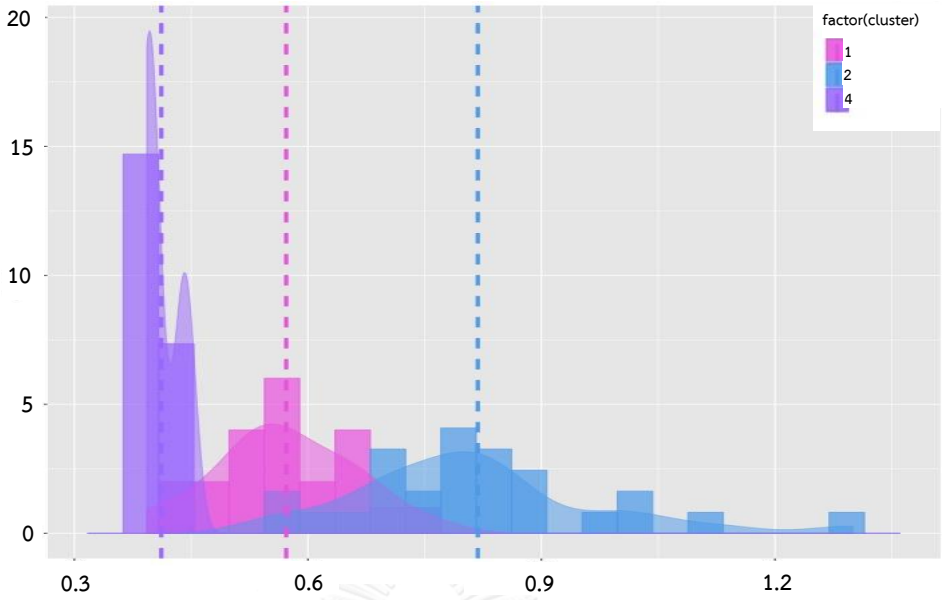


(a)

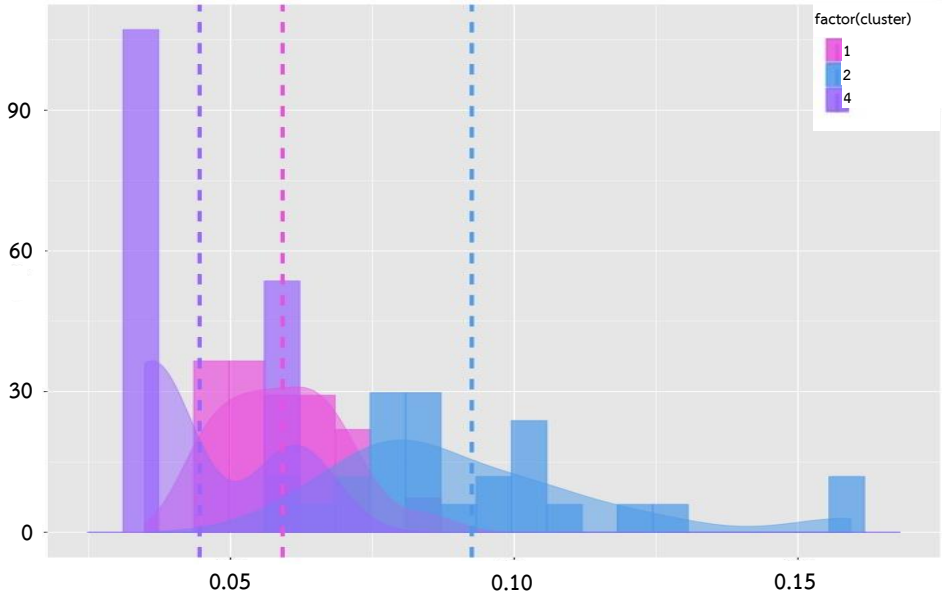


(b)

รูปที่ 5.5 Density histogram, Density Curve และค่าเฉลี่ย ของกลุ่มที่ 1 2 และ 4 ของ (a) SFW (b) SDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 1

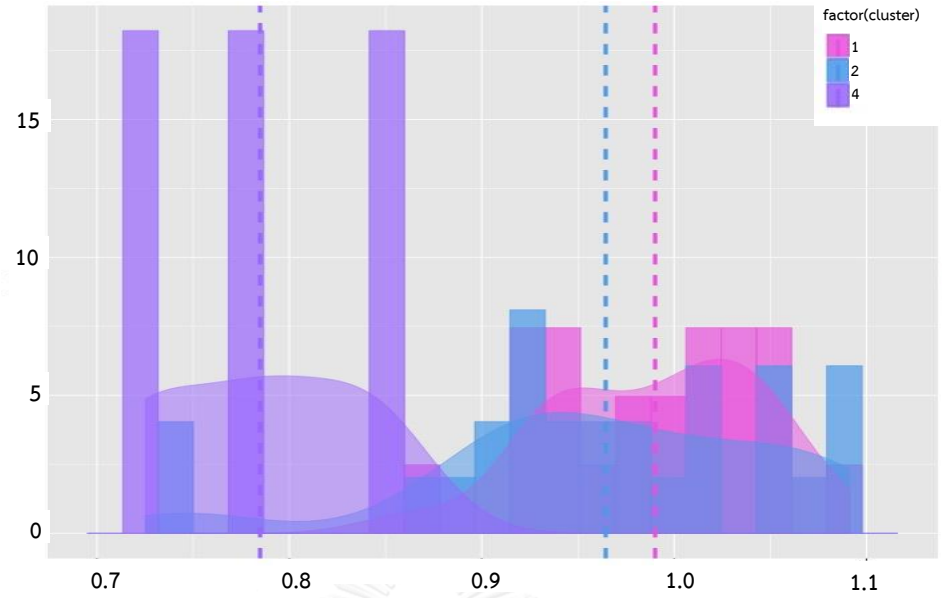


(c)

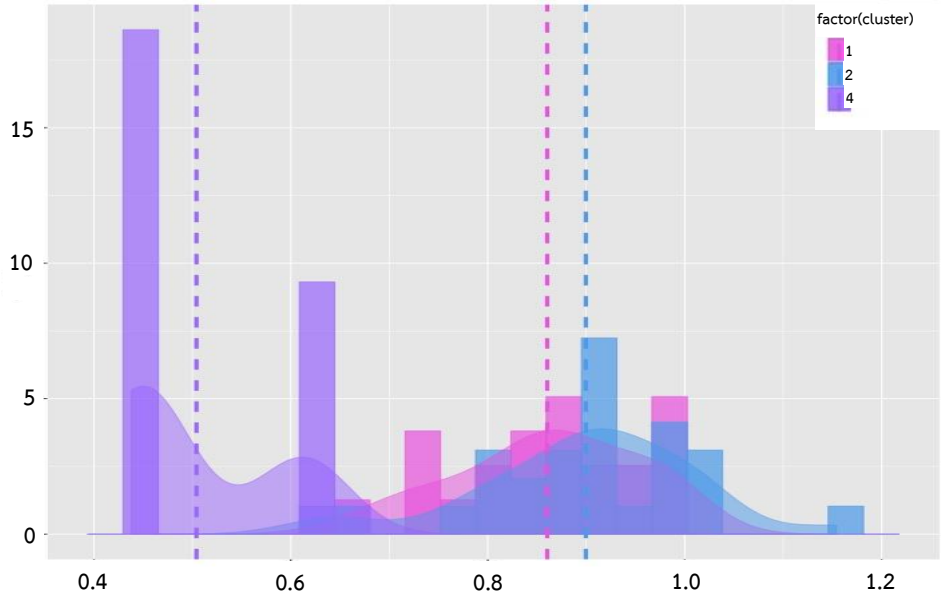


(d)

รูปที่ 5.6 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 2 และ 4 ของ (c) RFW (d) RDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 1



(e)



(f)

รูปที่ 5.7 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 2 และ 4 ของ (e) RWC (f) CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 1

ตารางที่ 5.7 ตารางแสดงค่าเฉลี่ยของแต่ละข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 1

Physiological / Cluster	SFW (GMP)	SDW (GMP)	RFW (MP)	RDW (MP)	RWC (STI)	CMS (YSI)
1 (Pokkali)	0.89273	0.11498	0.57214	0.05919	0.99015	0.86033
2	1.36816	0.17637	0.81824	0.09252	0.96443	0.89957
3* (หลวงประทาน)	2.54080	0.32069	1.32670	0.15115	1.09399	0.99580
4 (IR29)	0.54717	0.08674	0.41163	0.04457	0.78480	0.50417

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

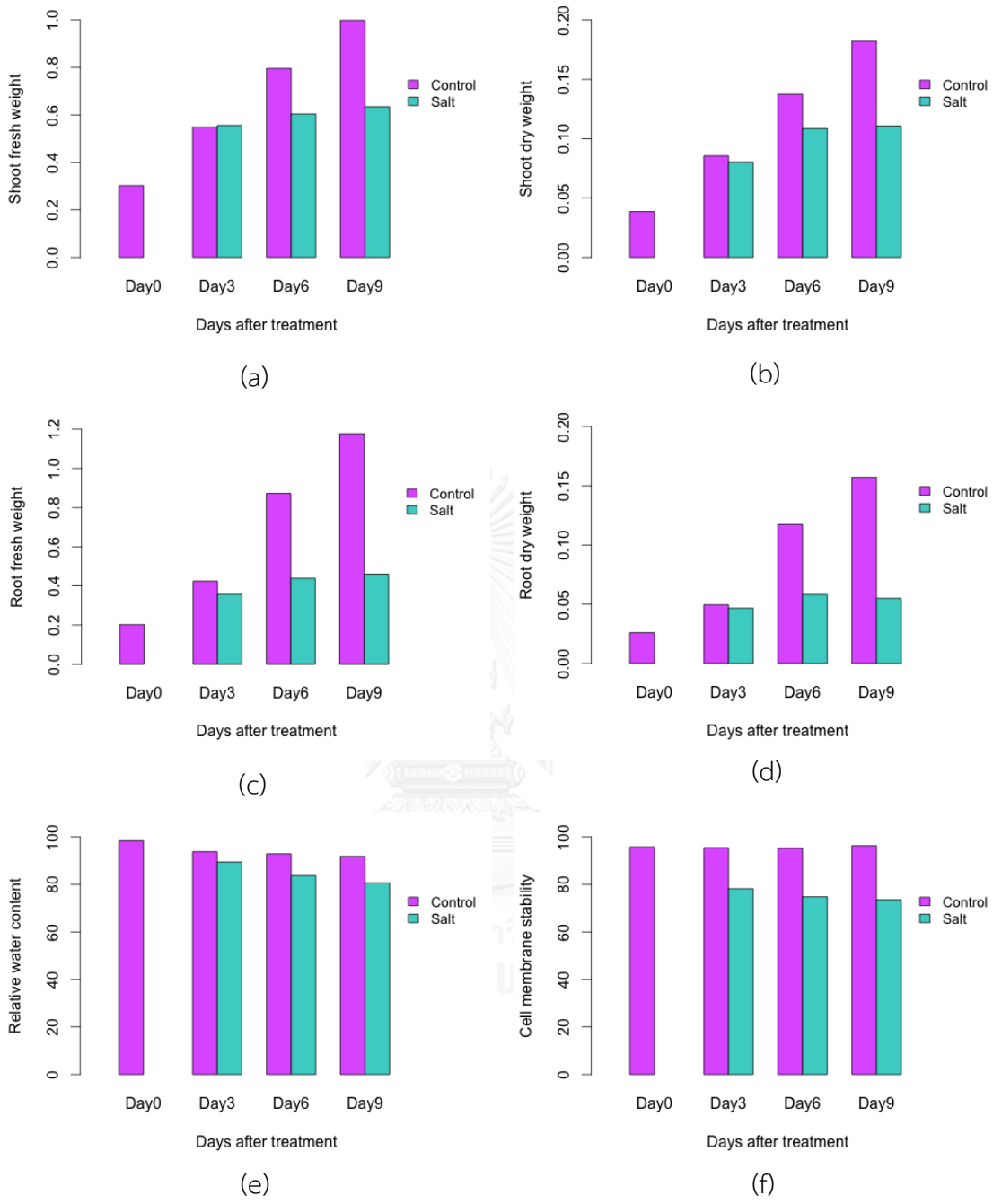
5.2 ข้อมูลชุดที่ 2

5.2.1 การเลือกวันที่นำมาใช้ทดลอง

จากการเปรียบเทียบ $\bar{Y}_c$ และ $\bar{Y}_s$ ดังแสดงในรูปที่ 5.8 พบว่า ค่าเฉลี่ยของ SFW SDW RFW RDW RWC และ CMS ในช่วงวันที่ 3 6 และ 9 มี Y_s น้อยกว่า Y_c ยกเว้น SFW ในช่วงวันที่ 3 มี Y_s มากกว่า Y_c เล็กน้อย ในช่วงวันที่ 6 เริ่มเห็นความแตกต่างกันอย่างชัดเจนระหว่าง Y_c กับ Y_s จากตารางที่ 5.8 พบว่าค่าเฉลี่ยของ Y_c และ Y_s ของข้อมูลสรีรวิทยาหกคุณลักษณะ ของข้อมูลชุดที่ 1 ในช่วงวันที่ 6 โดยมีค่า p-value ของการทดสอบทางสถิติด้วย Paired t - test น้อยกว่า 0.05 จากการพิจารณาข้างต้น ช่วงเวลาที่ถูกเลือกเพื่อนำไปใช้ในงานวิจัยนี้คือช่วงวันที่ 6

ตารางที่ 5.8 ตารางแสดงค่า p-value ของการทดสอบทางสถิติ Paired t - test เพื่อเปรียบเทียบข้อมูลสรีรวิทยาของข้อมูลชุดที่ 2 ระหว่าง Y_c กับ Y_s ของ SFW SDW RFW RDW RWC และ CMS

	SFW	SDW	RFW	RDW	RWC	CMS
Day 3	0.5939	0.0259	4.789e-06	0.1329	0.000251	2.2e-16
Day 6	< 2.2e-16	7.501e-15	< 2.2e-16	< 2.2e-16	9.351e-07	2.2e-16



รูปที่ 5.8 แผนภูมิแท่งเปรียบเทียบความแตกต่างของค่าเฉลี่ยทางสรีรวิทยาของข้อมูลชุดที่ 2 (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ระหว่าง Y_c กับ Y_s

5.2.2 การหาค่าผิดปกติ (Outliers)

ถ้าข้อมูลมีค่าผิดปกติจะทำให้การวิเคราะห์ข้อมูลอาจเกิดความผิดพลาดได้ ดังนั้นก่อนการจัดกลุ่มของข้อมูลทั้งแบบที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ และข้อมูลสรีรวิทยาหลายคุณลักษณะจึงควรตัดค่าผิดปกติออกไป เพื่อหลีกเลี่ยงการจัดกลุ่มที่สมาชิกในกลุ่มมีเพียงตัวเดียว สำหรับข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะค่าผิดปกติแบ่งเป็น 2 แบบ คือ ค่าที่สูงกว่าปกติและค่าที่ต่ำกว่าปกติ สำหรับข้อมูลสรีรวิทยาหกคุณลักษณะค่าผิดปกติเป็นค่าที่มีระยะทางมาฮาลาโนบิสสูงกว่าปกติ สำหรับข้อมูลทั้งสองลักษณะจะแสดงอยู่ในรูปของตาราง ในภาคผนวก ง ดังนี้ ตารางที่ ง.1 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่สูงกว่าปกติ ตารางที่ ง.2 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่ต่ำกว่าปกติ และตารางที่ ง.3 ตารางแสดงพันธุ์ข้าวข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีที่เหมือนกันและดัชนีที่ทำให้ค่า PBM มากที่สุด แล้วให้ค่าระยะทางมาฮาลาโนบิสที่สูงกว่าปกติ

พิจารณาตารางที่ 5.9 โดยรวมแล้วค่าน้ำหนักทั้ง 4 ค่าที่ได้รับการแปลงด้วยดัชนีเมื่อพิจารณาทุกดัชนี พบว่าพันธุ์ที่ให้ค่าผิดปกติซ้ำกันมากที่สุดคือ เล้าแตก สำหรับ RWC และ CMS ที่ได้รับการแปลงด้วยดัชนีเมื่อพิจารณาทุกดัชนี พบว่าพันธุ์ที่ให้ค่าผิดปกติซ้ำกันมากที่สุดที่เหมือนกันคือ อีพวง

ตารางที่ 5.9 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ เป็นค่าผิดปกติซ้ำกันมากที่สุด

พันธุ์ที่ให้ค่าผิดปกติ	สภาวะปกติ		สภาวะได้รับความเครียดจากความเค็ม		แปลงด้วยดัชนีต่าง ๆ เป็นค่าผิดปกติซ้ำกันมากที่สุด
	สูง	ต่ำ	สูง	ต่ำ	
SFW	-	ไอไท้	เล้าแตก	-	เล้าแตก (5 ดัชนี)
SDW	เล้าแตก	ไอไท้	เล้าแตก, อินแปง	-	เล้าแตก (5 ดัชนี)
RFW	เกษตรดอ, เล้าแตก	-	เล้าแตก	-	เล้าแตก (6 ดัชนี)
RDW	เกษตรดอ, เล้าแตก	-	เล้าแตก, อินแปง	-	เล้าแตก (6 ดัชนี)
RWC	เกษตรดอ, กข9, ไข่มดริ้น	-	-	อีพวง, หางนาค	อีพวง (ทุกดัชนี)
CMS	-	อีพวง, อู่ตะเภา, กำผาย, แม่หม้าย	-	อีพวง	อีพวง (ทุกดัชนี)

สำหรับข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่คละกันไปมีทั้งหมด 9^6 = 531,441 ชุดซึ่งข้อมูลแต่ละชุดมีโอกาสที่มีพันธุ์ที่ให้ค่าผิดปกติแตกต่างกันไป ในวิทยานิพนธ์นี้จึงเลือกใช้การหาพันธุ์ที่ให้ค่าผิดปกติโดยเลือกจากพันธุ์ที่มีระยะทางมาฮาลาโนบิสห่างจากจุดศูนย์กลางมากผิดปกติ

พิจารณาตารางที่ 5.10 สำหรับข้อมูลสรีรวิทยาหกคุณลักษณะ พันธุ์ที่ให้ค่าระยะทางมาฮาลาโนบิสสูงผิดปกติซ้ำกันมากที่สุดคือ เล้าแตก (6 ดัชนี) โดยข้อมูลสรีรวิทยาที่โดดเด่นของพันธุ์เล้าแตก มีดังนี้

<u>เล้าแตก</u> ในสภาวะปกติ SFW สูงสุด SDW สูงสุดและเป็นค่าผิดปกติด้วย RFW และ RDW สูงสุดอันดับสองและเป็นค่าผิดปกติด้วย และ CMS สูงสุดอันดับสาม ในสภาวะได้รับความเครียดจากความเค็ม ค่าน้ำหนักทั้ง 4 ค่าสูงสุดและเป็นค่าผิดปกติด้วย และ RWC ต่ำสุดอันดับสี่

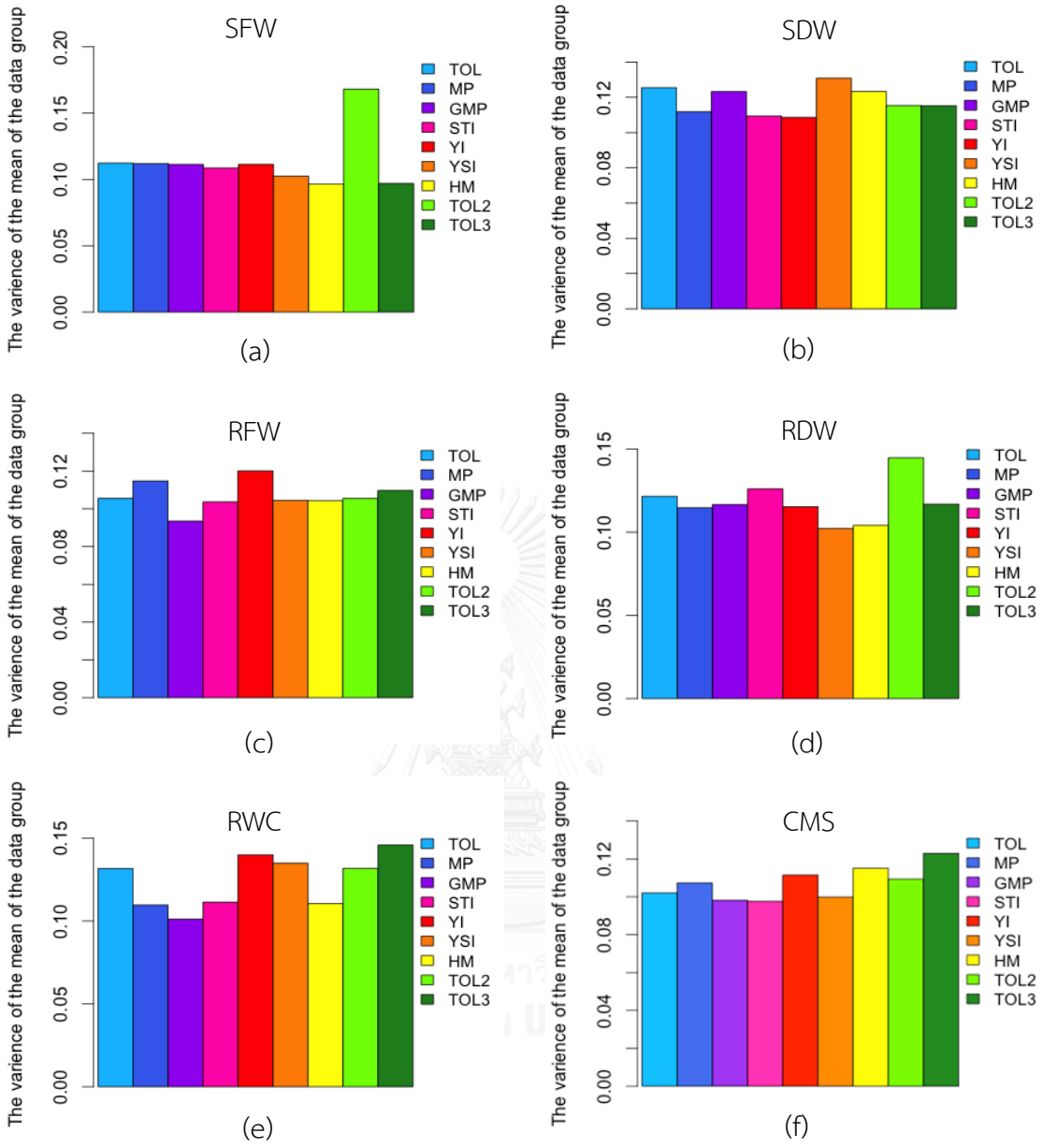
ตารางที่ 5.10 ตารางแสดงพันธุ์ข้าวของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนี ที่เหมือนกันเป็นค่าผิดปกติซ้ำกันมากที่สุด

ค่าผิดปกติ	สภาวะปกติ	สภาวะได้รับความเครียดจากความเค็ม	แปลงด้วยดัชนีที่เหมือนกันเป็นค่าผิดปกติซ้ำกันมากที่สุด
ข้อมูลสรีรวิทยาหกคุณลักษณะ	-	เล้าแตก	เล้าแตก (6 ดัชนี)

5.2.3 การหาดัชนีที่ดีในการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

สำหรับการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะ จะใช้ข้อมูลที่ได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด ซึ่งข้อมูลสรีรวิทยาที่ได้รับความเครียดจากความเค็มและที่ได้รับการแปลงด้วยดัชนี YI เมื่อได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด มีผลทำให้ได้ค่าข้อมูลทั้งสองเหมือนกัน นั่นคือได้ผลจากการจัดกลุ่มที่เหมือนกัน ค่าความแปรปรวนระหว่างกลุ่มที่มากที่สุดจากการจัดกลุ่มของแต่ละข้อมูลสรีรวิทยาจะถูกแสดงในรูปที่ 5.9 โดยดัชนีที่ให้ค่าความแปรปรวนระหว่างกลุ่มที่มากที่สุด ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมที่เหมาะสมจากการจัดกลุ่มของแต่ละข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะมากที่สุด 3 อันดับจะถูกแสดงในตารางที่ 5.11

การเลือกดัชนีที่เหมาะสมของแต่ละข้อมูลสรีรวิทยาจะเลือกจากค่าความแปรปรวนระหว่างกลุ่มที่มากผิดปกติ สำหรับข้อมูลชุดที่ 2 พบว่า TOL2 เป็นดัชนีที่เหมาะสมสำหรับ SFW และ YI เป็นดัชนีที่เหมาะสมสำหรับ RFW โดย SDW RDW RWC และ CMS ไม่ปรากฏดัชนีที่มีค่าความแปรปรวนระหว่างกลุ่มที่มากผิดปกติ จึงไม่สามารถสรุปดัชนีที่เหมาะสมได้



รูปที่ 5.9 แผนภูมิแสดงการเปรียบเทียบค่าความแปรปรวนระหว่างกลุ่มจากการเลือกจำนวนกลุ่มที่เหมาะสมของแต่ละดัชนีของ (a) SFW (b) SDW (c) RFW (d) RDW (e) RWC และ (f) CMS ของข้อมูลชุดที่ 2

ตารางที่ 5.11 ตารางแสดงดัชนีที่ให้ค่าความแปรปรวนของค่าเฉลี่ยสมาชิกในกลุ่มที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าว ด้วยข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีมีค่ามากที่สุด 3 อันดับ ของข้อมูลชุดที่ 2

	อันดับที่ 1	**อันดับที่ 2**	**อันดับที่ 3**
SFW	TOL2	TOL	MP
SDW	YSI	TOL	HM
RFW	YI	MP	TOL3
RDW	TOL2	STI	TOL
RWC	TOL3	YI	YSI
CMS	TOL3	HM	YI

5.2.4 การหาดัชนีที่เหมาะสมในการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี

เนื่องจากการประเมินจำนวนกลุ่มของข้อมูลจริงจากฐานข้อมูล UCI ที่ข้อมูลแต่ชุดที่ทดสอบมีการจัดประเภทกลุ่มของข้อมูล พบว่าการประเมินโดยค่า PBM ที่เกิดจากการจัดกลุ่มของข้อมูลดิบ มีการประเมินจำนวนกลุ่มที่ดีกว่าข้อมูลที่ได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด สำหรับการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีนั้นจึงไม่ได้รับการแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุด แต่จะแปลงให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดเพื่อประเมินดัชนี ที่เหมาะสมสำหรับการจัดกลุ่มแบบหกคุณลักษณะ ค่า PBM มากที่สุดมาจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของดัชนี STI MP MP MP STI และ YSI ตามลำดับ ซึ่งปรากฏพันธุ์ที่ให้ค่าผิดปกติคือ อินแปง โดยให้ผลการจัดกลุ่มพันธุ์ข้าวที่มีค่า PBM สูงสุดซึ่งแสดงในรูปที่ 5.10 ซึ่งเป็นแผนภูมิแสดงค่า PBM จากการจัดกลุ่มพันธุ์ข้าวที่ข้อมูลสรีรวิทยาหกคุณลักษณะมาจากข้อมูลที่ได้รับความเครียดจากความเค็มโดยไม่ได้รับการแปลงด้วยดัชนี (Salt) ข้อมูล

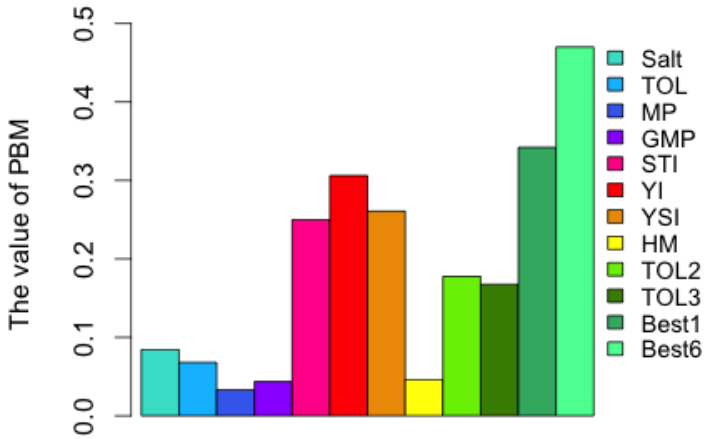
สรีรวิทยาหกคุณลักษณะใช้ดัชนีที่เหมือนกัน (TOL MP GMP STI YI YSI HM TOL2 และ TOL3) ใช้ดัชนีที่เหมาะสมกับการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะเดียว (Best1) และดัชนีที่ทำให้การจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะเดียวมีค่ามากที่สุด (Best6) จากรูปนี้พบว่า

1. สำหรับจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี YI เหมือนกันให้ค่า PBM สูงกว่า Salt

2. สำหรับจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี Best1 มีค่ามากกว่า ดัชนีที่เหมือนกัน

3. สำหรับจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี Best1 น้อยกว่า Best6

ค่า PBM ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าวด้วยข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่เหมือนกันมีค่ามากที่สุด 3 อันดับจะถูกแสดงในตารางที่ 5.12



รูปที่ 5.10 แผนภูมิแสดงการเปรียบเทียบค่า PBM จากการเลือกจำนวนกลุ่มที่เหมาะสมของ Salt TOL MP GMP STI YI YSI HM TOL2 TOL3 Best1 และ Best6 ของข้อมูลชุดที่ 2

ตารางที่ 5.12 ตารางแสดงดัชนีที่ให้ค่า PBM ที่มาจากการเลือกจำนวนกลุ่มที่เหมาะสมจากการจัดกลุ่มพันธุ์ข้าวด้วยข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ได้รับการแปลงด้วยดัชนีที่เหมือนกัน แล้วจึงแปลงข้อมูลให้อยู่ในช่วงที่ต้องการด้วยการกำหนดค่าต่ำสุดและสูงสุดมีค่ามากที่สุด 3 อันดับของข้อมูลชุดที่ 2

	อันดับที่ 1	**อันดับที่ 2**	**อันดับที่ 3**
ทุกข้อมูลสรีรวิทยา	YI	YSI	STI

5.2.5 การทดสอบการจัดกลุ่มพันธุ์ข้าวที่ใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม

จากการทดสอบทางสถิติด้วย Mann – Whitney U Test ค่า p-value ของการทดสอบทางสถิติของกลุ่มในแต่ละคู่ของข้อมูลสรีรวิทยา SFW และ RFW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมจะถูกแสดงในตารางที่ 5.13 โดยจากการทดสอบพบว่า

SDW สำหรับดัชนี TOL2 ค่า p-value ของกลุ่มที่ 3 กับกลุ่มที่ 1 และ 2 มีค่ามากกว่า 0.05 แต่น้อยกว่า 0.1 ดังนั้นพันธุ์ข้าวในกลุ่มที่ 3 มาจากประชากรที่มีค่ามัธยฐานของข้อมูลสรีรวิทยาไม่แตกต่างกับกลุ่มที่ 1 และ 2 ที่ระดับความเชื่อมั่น 95% แต่การจัดกลุ่มที่ได้กลุ่มในแต่ละคู่มีความแตกต่างระหว่างกลุ่มที่ระดับความเชื่อมั่น 90%

RFW สำหรับดัชนี YI ค่า p-value ของการทดสอบกลุ่มที่ 1 และ กลุ่มที่ 2 และ 3 น้อยกว่า 0.05 ดังนั้นพันธุ์ข้าวในกลุ่มที่ 1 มาจากประชากรที่มีค่ามัธยฐานของข้อมูลสรีรวิทยาแตกต่างกับกลุ่มที่ 2 และ 3 นั่นคือการจัดกลุ่มที่ได้มีความแตกต่างระหว่างกลุ่มที่ระดับความเชื่อมั่น 95%

ตารางที่ 5.13 ตารางแสดงค่า p-value ของการทดสอบทางสถิติของการจัดกลุ่มโดยทดสอบแต่ละคู่กลุ่มด้วย (a) ข้อมูลสรีรวิทยา SFW ที่ได้รับการแปลงด้วยดัชนี TOL2 และ (b) ข้อมูลสรีรวิทยา RFW ที่ได้รับการแปลงด้วยดัชนี YI ของชุดที่ 2

Cluster	2	3
1	9.88E-15	9.52E-02
2		5.88E-02

(a)

Cluster	2	3
1	2.25E-14	4.45E-04
2		8.70E-04

(b)

5.2.6 ผลการจัดกลุ่ม

ผลการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์ โดยใช้ทุกข้อมูลสรีรวิทยาของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลสรีรวิทยา SFW SDW RFW RDW RWC และ CMS ด้วยโดยดัชนี STI MP MP MP STI และ YSI ตามลำดับ มีพันธุ์ที่ให้ค่าผิดปกติคือ อินแปง โดยผลการจัดกลุ่มบอกได้เพียงว่าพันธุ์ที่อยู่ในกลุ่มเดียวกันมีลักษณะที่คล้ายคลึงกันเมื่อได้รับความเค็มเนื่องจากมีพันธุ์ที่ใช้ระบุลักษณะของพันธุ์น้อย จากรูปที่ 5.11 พบว่า

<u>กลุ่มที่ 1</u> ประกอบด้วย ปทุมธานี1 (ข้าวพันธุ์อ่อนแอต่อความเค็ม) CSSL11 (ข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง) ดำด่าง สายหยุด กำผาย หางนาค อีพวง กข9 สำอางค์ ไข่มดริ้น ก่ำเลี้ยว ห้ารวง เกวียนหัก พวงหางหมู ขาวประกวด UBN2123-50R-B-3 ไอไท้ เหลืองดง กข13 เหลืองงาม แม่หม้าย และเจ้าดอกขาว

<u>กลุ่มที่ 2</u> (ทนทานต่อความเค็มระดับปานกลาง) ประกอบด้วย KDML105 (ข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง) เกษตรดอ บัวน้อย ขาวกอเดียว ทูลฉลอง ชิวแม่จัน นางมล หมากบิด นางงาม สามรวงหนัก กข8 เหลืองตายัง อู่ตะเภา เจ้าดำ เหนียวขาว สามสี ป้องแอ้ว (12159) เหนียวมะลิ พวงหวาน เหลืองทอง รากแห้ง หางหมาใน ปลาซิวดำ รวงเดียว ดอกไม้ หางยี71 ม่วยหิน ขาวตาแห้ง17 ใบบัว และเจ้าแดง

<u>กลุ่มที่ 3</u> มีพันธุ์เดียวคือเล้าแตก

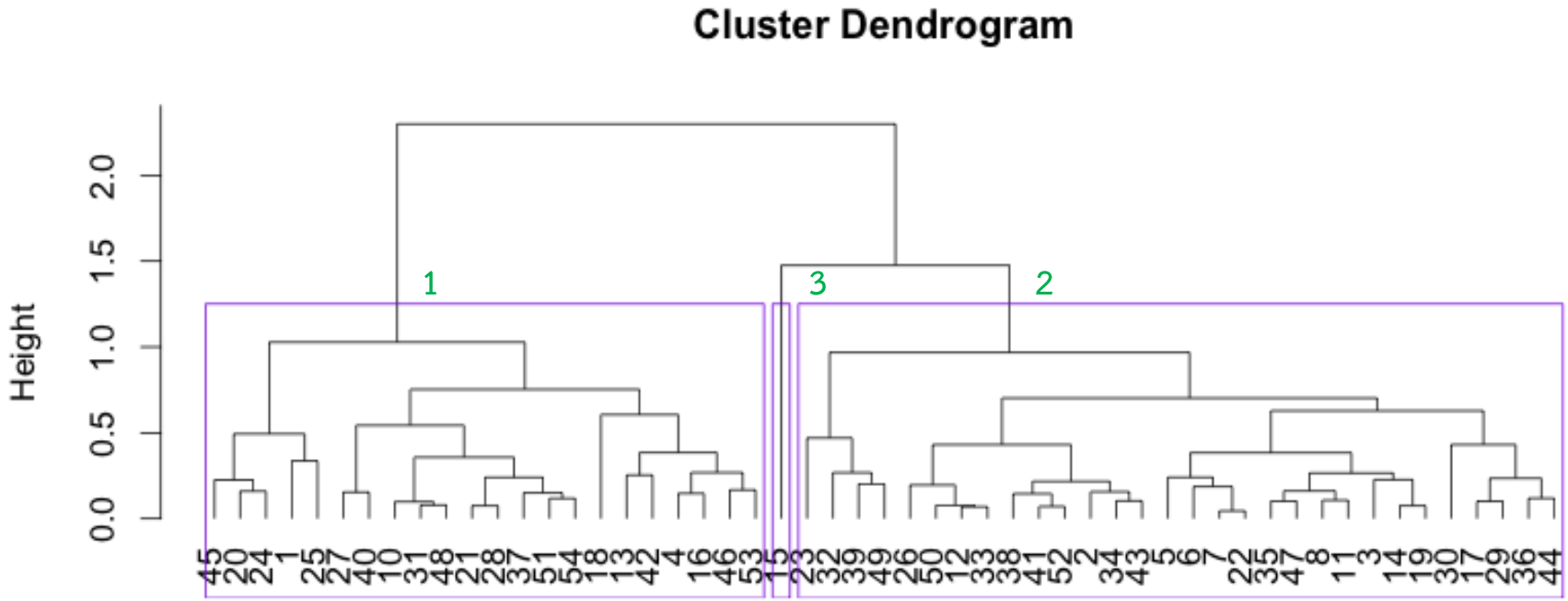
จากรูปที่ 5.12 ถึง รูปที่ 5.14 ซึ่งเป็น Density histogram, Density curve และค่าเฉลี่ยโดยแสดงเป็นเส้นประ ของกลุ่มที่ 1 และ 2 ของข้อมูลสรีรวิทยาแต่ละคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม เป็นการแสดงการกระจายตัวของข้อมูลสรีรวิทยาแต่ละคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของแต่ละกลุ่มยกเว้นกลุ่มที่ 3 เนื่องจากเป็นกลุ่มที่มีสมาชิกเพียงตัวเดียวคือ เล้าแตก จึงไม่มีการกระจายตัวของข้อมูลสรีรวิทยาแต่ละคุณลักษณะ และตารางที่ 5.14 แสดงค่าเฉลี่ยของกลุ่ม 1 2 3 และพันธุ์ที่ให้ค่าผิดปกติที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม โดยค่าเฉลี่ยของข้อมูลสรีรวิทยาแต่ละคุณลักษณะของกลุ่มที่มีค่ามากกว่าจะมีสีของช่องที่เข้มกว่า ซึ่งค่าเฉลี่ยของแต่ละกลุ่มมีลักษณะดังนี้

<u>กลุ่มที่ 1</u> ซึ่งประกอบด้วย ปทุมธานี1 (ข้าวพันธุ์อ่อนแอต่อความเค็ม) และ CSSL11 (ข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง) ค่าเฉลี่ยของข้อมูลสรีรวิทยาน้ำหนัก และ CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่าน้อยที่สุด และค่าเฉลี่ยของ RWC ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสาม ดังนั้นกลุ่มที่ 1 น่าจะทนเค็มน้อยที่สุดของข้อมูลชุดที่ 2

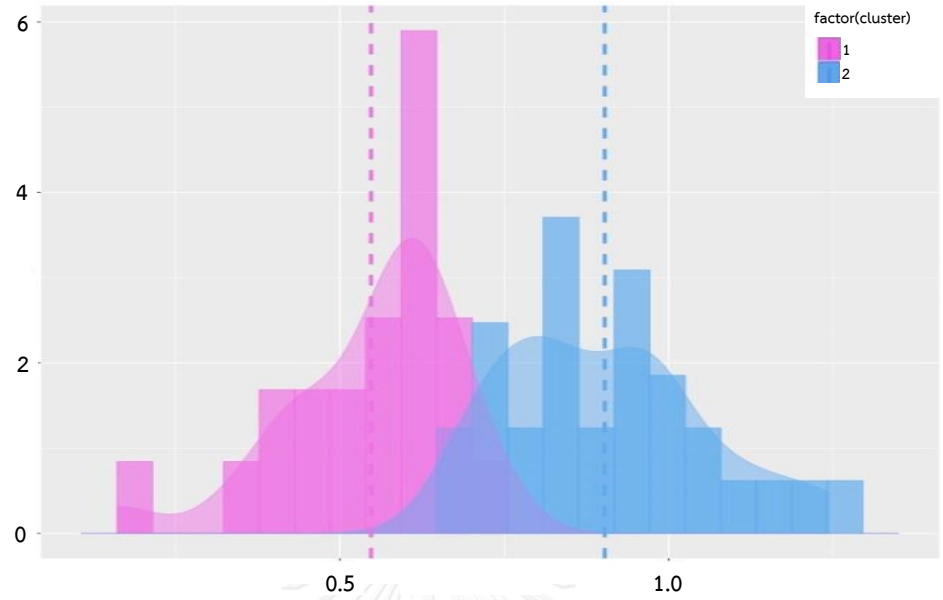
<u>กลุ่มที่ 2</u> เป็นกลุ่มที่ทนทานต่อความเค็มระดับปานกลางซึ่งประกอบด้วย KDML105 (ข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง) ค่าเฉลี่ยของข้อมูลสรีรวิทยาน้ำหนักที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสาม และค่าเฉลี่ยของ RWC และ CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากกที่สุดอันดับสอง ดังนั้นกลุ่มที่ 2 น่าจะทนเค็มปานกลางรองจากเล้าแตก และพันธุ์ที่ให้ค่าผิดปกติคืออินแปง

<u>กลุ่มที่ 3</u> สมาชิกมีเพียงตัวเดียวคือเล้าแตก ค่าเฉลี่ยของข้อมูลสรีรวิทยาน้ำหนักที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุด ค่าเฉลี่ยของ RWC ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่าน้อยที่สุด และค่าเฉลี่ยของ CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสาม ดังนั้นเล้าแตกน่าจะทนเค็มมากกว่ากลุ่ม 1 และ 2 ของข้อมูลชุดที่ 2

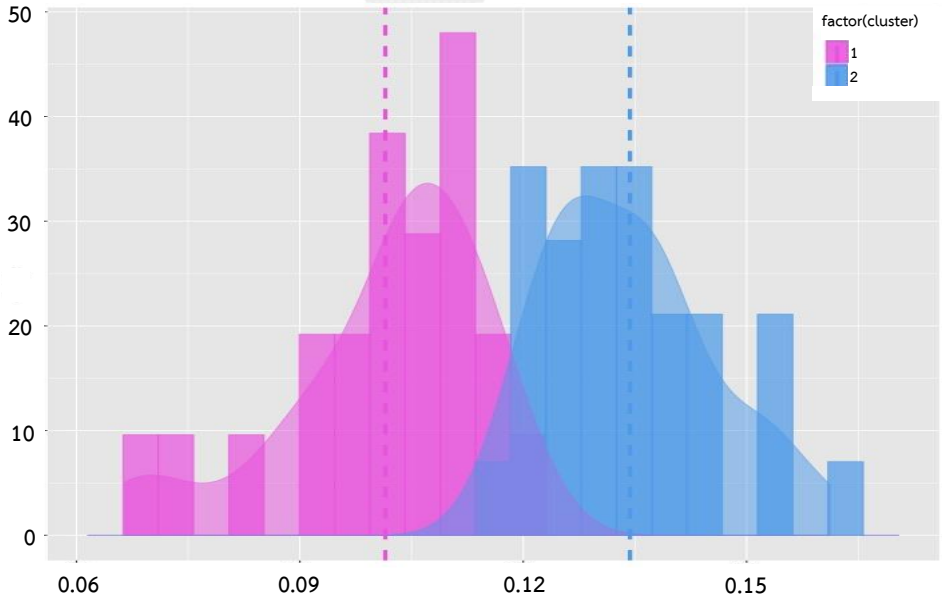
<u>กลุ่มที่ 4</u> พันธุ์ที่ให้ค่าผิดปกติคืออินแปง โดยค่าเฉลี่ยของข้อมูลสรีรวิทยาน้ำหนักที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุดอันดับสอง และค่าเฉลี่ยของ RWC และ CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมมีค่ามากที่สุด กลุ่มที่ 4 จึงน่าจะทนเค็มมากกว่า กลุ่ม 1 และ 2



รูปที่ 5.11 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลสรีรวิทยา SFW SDW RFW RDW RWC และ CMS ด้วยโดยดัชนี STI MP MP MP STI และ YSI ตามลำดับ

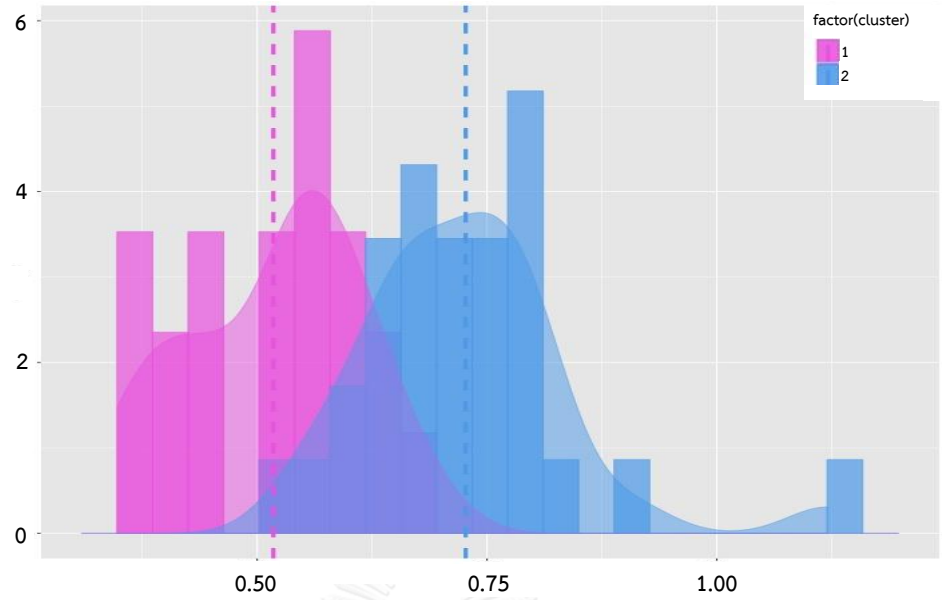


(a)

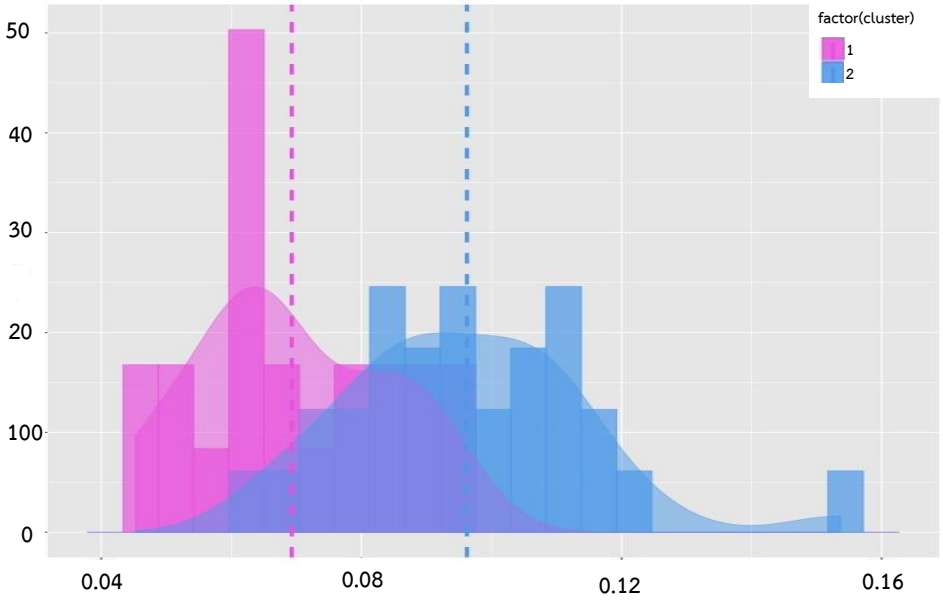


(b)

รูปที่ 5.12 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 และ 2 ของ (a) SFW (b) SDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 2

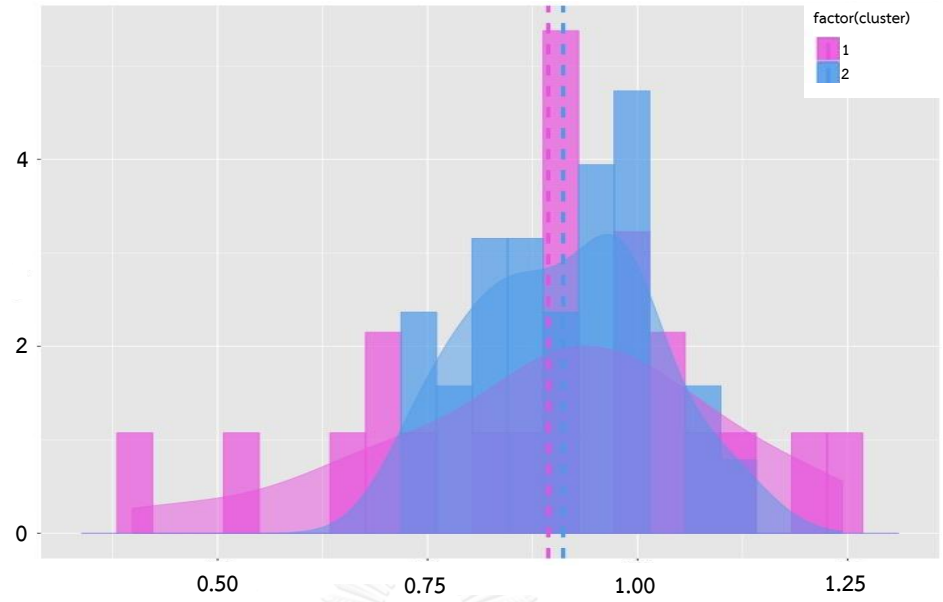


(c)

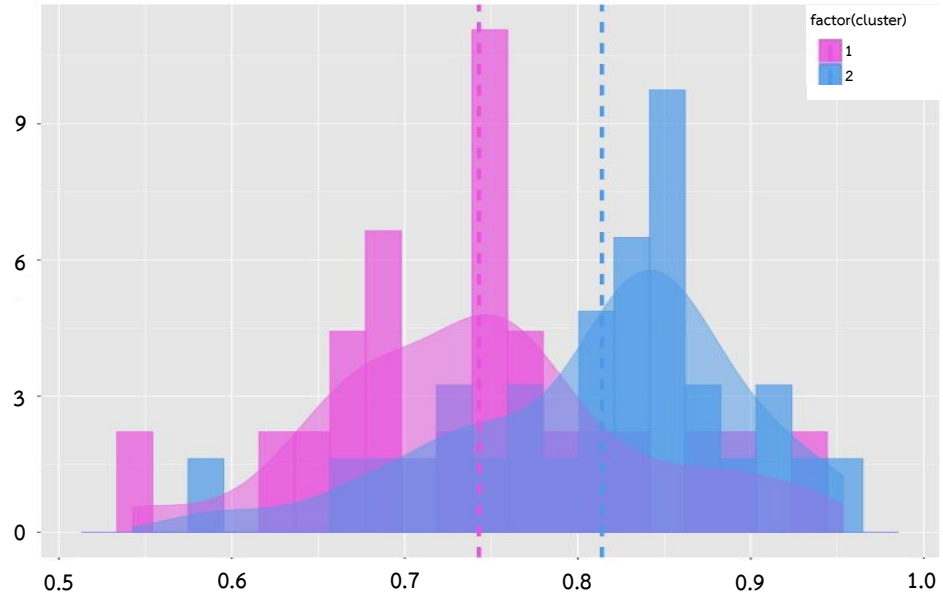


(d)

รูปที่ 5.13 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 และ 2 ของ (c) RFW (d) RDW ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 2



(e)



(f)

รูปที่ 5.14 Density histogram, Density curve และค่าเฉลี่ย ของกลุ่มที่ 1 และ 2 ของ (e) RWC (f) CMS ที่ได้รับการแปลงด้วยดัชนีที่เหมาะสม ของข้อมูลชุดที่ 2

ตารางที่ 5.14 ตารางแสดงค่าเฉลี่ยของแต่ละข้อมูลสรีรวิทยาที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 2

Physiological / Cluster	SFW (STI)	SDW (MP)	RFW (MP)	RDW (MP)	RWC (STI)	CMS (YSI)
1	0.54771	0.10147	0.51761	0.06930	0.89354	0.74290
2 (KDML105)	0.90266	0.13434	0.72684	0.09624	0.91138	0.81400
3* (เล้าแตก)	1.75311	0.20640	1.30120	0.17320	0.66549	0.79135
Outlier อินแปง	1.11853	0.16529	0.91166	0.15417	0.93983	0.85333

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5.3 ข้อเสนอแนะ

1. แต่ละดัชนีมีทั้งข้อดีและข้อด้อยแตกต่างกันออกไป โดยรวมพบว่า

   TOL เป็นดัชนีที่ไม่เหมาะในการแปลงข้อมูลเนื่องจาก ดัชนี TOL พันธุ์ที่มีค่าดัชนี TOL น้อยกว่าจะสื่อถึงพันธุ์นั้นทนเค็มได้มากกว่า แต่ถ้าข้อมูลสรีรวิทยาต้นข้าวในสภาวะปกติและได้รับความเครียดจากความเค็มน้อยทั้งคู่ จะสื่อถึงว่าข้าวพันธุ์นั้นทนเค็มได้ไม่ดี ดัชนี TOL จึงถูกพัฒนาเป็นดัชนี TOL2 และ TOl3

   MP GMP STI และ HM เนื่องจากข้อมูลที่นำมาใช้มาจากการเลือกวันที่ข้อมูลสรีรวิทยาในสภาวะปกติและได้รับความเครียดจากความเค็ม มีความแตกต่างกันอย่างมีนัยสำคัญ ดังนั้นค่าดัชนี MP GMP STI และ HM ที่มีค่าสูงอาจเกิดจากข้อมูลสรีรวิทยาในสภาวะปกติมีค่าสูง ซึ่งมีผลทำให้การวิเคราะห์ข้อมูลอาจมีความคลาดเคลื่อนได้

   YI หากพันธุ์ที่ข้อมูลสรีรวิทยาในสภาวะปกติ มากกว่า ที่ได้รับความเครียดจากความเค็มมาก และข้อมูลสรีรวิทยาได้รับความเครียดจากความเค็มสูงทำให้สื่อว่าข้าวพันธุ์นั้นไม่ทนเค็ม แต่ดัชนี YI ใช้แต่ข้อมูลสรีรวิทยาที่ได้รับความเครียดจากความเค็มอย่างเดียวทำอาจสื่อว่าข้าวพันธุ์นั้นทนเค็มได้

2. ในการหาค่าผิดปกติของข้อมูลสรีรวิทยาหกคุณลักษณะ โดยใช้การหาค่าผิดปกติของระยะทางมาฮาลาโนบิสนั้น ต้องมีการหาเมทริกซ์ผกผันของเมทริกซ์ความแปรปรวนร่วม ถ้าหากข้อมูลสรีรวิทยาแต่ละค่ามีค่าสัมประสิทธิ์สหสัมพันธ์ในระดับสูง $(r > 0.95)$ แล้วจะทำให้ไม่สามารถหาเมทริกซ์ผกผันของเมทริกซ์ความแปรปรวนร่วมได้ และการจัดกลุ่มโดยใช้หลายข้อมูลสรีรวิทยานั้นไม่ควรจะใช้ข้อมูลสรีรวิทยา ที่มีค่าสัมประสิทธิ์สหสัมพันธ์ในระดับสูงเช่นเดียวกัน เนื่องจากทำให้การจัดกลุ่มมีแนวโน้มที่เหมือนกัน ดังนั้นจึงควรตัดข้อมูลสรีรวิทยา ที่มีค่าสัมประสิทธิ์สหสัมพันธ์ในระดับสูงตัวหนึ่งก่อนนำมาใช้ในการสกัดดัชนีที่เหมาะสม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3. การจัดกลุ่มที่ใช้ข้อมูลเดียวกันให้ผลลัพธ์ที่ต่างกันออกไป เมื่อทำการเปลี่ยนวิธีในการหาระยะทางระหว่างกลุ่ม ซึ่งการเลือกวิธีการหาระยะทางระหว่างกลุ่มขึ้นกับผู้ใช้งานว่าต้องการจัดกลุ่มแบบใด

4. ข้อมูลไม่มีพันธุ์ที่ทนเค็ม และอ่อนแอต่อความเค็มมาตรฐานมากพอที่จะบ่งชี้ลักษณะในกลุ่มได้ ดังนั้นผลการจัดกลุ่มที่ได้จึงบอกได้แค่ว่าพันธุ์ที่อยู่ในกลุ่มเดียวกันมีลักษณะที่คล้ายกันเท่านั้น

5.4 สรุป

ข้อมูลชุดที่ 1

1. STI และ TOL3 เป็นดัชนีที่เหมาะสมสำหรับ SDW และ STI และ HM เป็นดัชนีที่เหมาะสมสำหรับ RDW โดย SFW RFW RWC และ CMS ไม่ปรากฏดัชนีที่มีค่าความแปรปรวนระหว่างกลุ่มที่มากผิดปกติจึงไม่สามารถสรุปดัชนีที่เหมาะสมได้

2. ค่า PBM ของข้อมูลที่ SDW เป็นค่าที่เกิดจากดัชนี STI และ TOL3 ซึ่งเป็นดัชนีที่เหมาะสมสำหรับ SDW และ RDW เป็นค่าที่เกิดจากดัชนี STI และ HM ซึ่งเป็นดัชนีที่เหมาะสมสำหรับ RDW และ SFW RFW RWC และ CMS เป็นค่าที่เกิดจากดัชนีทั้ง 9 พบว่ามีค่ามากกว่าค่า PBM ที่เกิดจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีเหมือนกัน แต่ก็ยังน้อยกว่าค่า PBM ที่มากที่สุดจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของดัชนี GMP GMP MP MP STI และ YSI ตามลำดับ ซึ่งไม่ปรากฏพันธุ์ที่ให้ค่าผิดปกติโดยให้ผลการจัดกลุ่มพันธุ์ข้าวที่มีค่า PBM สูงสุด

ผลจากการจัดกลุ่มพบว่า

<u>กลุ่มที่ 1</u> (ทนทานต่อความเค็ม) ประกอบด้วย Pokkali (ข้าวพันธุ์ทนทานต่อความเค็ม) ขาวกอเดียว พวงหางหมู ใบบัว หางยี71 เหนียวดำหลาย ชัยนาท1 เจ๊กเชย พ้าวฮ้าว ปิ่นแก้ว นางมล ป้องแอ้ว (21577) เม็ดมะม่วง อีมุม LPT123–TC171 มะยม (13988) พวงทอง หลวงประทาน เก้ารวง ขาวตาแห้ง17 ก้นแก้ว(1698) และเหลืองน้อย31-1-39

<u>กลุ่มที่ 2</u> ประกอบด้วย อีขาวใหญ่ อีลาย ตะเภาล่ม ดอขาว มะยม(22802) เหลืองพวงทอง ดอสามเดือน ดอดอกไม้ หมากน้ำ(22835) เจ้าขาว พลายงามปราจีนบุรี ขาวอากาศ มันวัว ก้นแก้ว(6158) ขาวก้นจุด ทองรากไทร ขาวสงวน จุดมอญ ดีสี่ เขียวหางม้า หมากยม พวงหนัก หมากน้ำ(14085) ขาวสุพรรณ จำปาจีน ประทานบ้านบุ้ง และขาวแก้ว

<u>กลุ่มที่ 3</u> มีเพียงพันธุ์เดียวคือหลวงประทาน

<u>กลุ่มที่ 4</u> (อ่อนแอต่อความเค็ม) ประกอบด้วย IR29 (ข้าวพันธุ์อ่อนแอต่อความเค็ม) กข1 และไข่มดริ้น

<u>ข้อมูลชุดที่ 2</u>

1. TOL2 เป็นดัชนีที่เหมาะสมสำหรับ SFW และ YI เป็นดัชนีที่เหมาะสมสำหรับ RFW โดย SDW RDW RWC และ CMS ไม่ปรากฏดัชนีที่มีค่าความแปรปรวนระหว่างกลุ่มที่มากผิดปกติ จึงไม่สามารถสรุปดัชนีที่เหมาะสมได้

2. และค่า PBM ของข้อมูลที่ SFW เป็นค่าที่เกิดจาดดัชนี TOL2 ซึ่งเป็นดัชนีที่เหมาะสมสำหรับ SFW และ RFW เป็นค่าที่เกิดจากดัชนี YI ซึ่งเป็นดัชนีที่เหมาะสมสำหรับ RFW และ SDW RDW RWC และ CMS เป็นค่าที่เกิดจากดัชนีทั้ง 9 พบว่ามีค่าน้อยกว่าค่า PBM ที่เกิดจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนีเหมือนกัน ค่า PBM มากที่สุดมาจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาหกคุณลักษณะของดัชนี STI MP MP MP STI และ YSI ตามลำดับ ซึ่งปรากฏพันธุ์ที่ให้ค่าผิดปกติคือ อินแปง โดยให้ผลการจัดกลุ่มพันธุ์ข้าวที่มีค่า PBM สูงสุด

ผลจากการจัดกลุ่มพบว่า

<u>กลุ่มที่ 1</u> (อ่อนแอต่อความเค็มของข้อมูลชุดที่ 2) ประกอบด้วย ปทุมธานี1 (ข้าวพันธุ์อ่อนแอต่อความเค็ม) CSSL11 (ข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง) ดำด่าง สายหยุด กำผาย หางนาค อีพวง กข9 สำอางค์ ไข่มดริ้น ก่ำเลี้ยว ห้ารวง เกวียนหัก พวงหางหมู ขาวประกวด UBN2123-50R-B-3 ไอไท้ เหลืองดง กข13 เหลืองงาม แม่หม้าย และเจ้าดอกขาว

<u>กลุ่มที่ 2</u> (ทนทานต่อความเค็มระดับปานกลาง) ประกอบด้วย KDML105 (ข้าวพันธุ์ทนทานต่อความเค็มระดับปานกลาง) เกษตรดอ บัวน้อย ขาวกอเดียว ทูลฉลอง ซิวแม่จัน นางมล หมากบิด นางงาม สามรวงหนัก กข8 เหลืองตายัง อู่ตะเภา เจ้าดำ เหนียวขาว สามสี ป้องแอ้ว (12159) เหนียวมะลิ พวงหวาน เหลืองทอง รากแห้ง หางหมาใน ปลาซิวดำ รวงเดียว ดอกไม้ หางยี71 ม่วยหิน ขาวตาแห้ง17 ใบบัว และเจ้าแดง

<u>กลุ่มที่ 3</u> มีพันธุ์เดียวคือเล้าแตก

<u>พันธุ์ที่ให้ค่าผิดปกติ</u> มีพันธุ์เดียวคืออินแปง

จากข้อมูลทั้งสองชุดพบว่าการจัดกลุ่มพันธุ์ข้าว ที่ใช้ข้อมูลสรีรวิทยาหกคุณลักษณะที่ได้รับการแปลงด้วยดัชนี YI เหมือนกันให้ค่า PBM สูงกว่าใช้ข้อมูลที่ได้รับความเครียดจากความเค็มโดยไม่ได้รับการแปลงด้วยดัชนี และเมื่อพิจารณาค่า PBM ที่มากที่สุด 30 อันดับแรกของข้องมูลทั้งสองชุดไม่ปรากฏดัชนี TOL ในการแปลงข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ทำให้ค่า PBM ที่มากที่สุด 30 อันดับแรกเลย แต่ปรากฏดัชนี TOL2 แลละ TOL3 ในการแปลงข้อมูลสรีรวิทยาหกคุณลักษณะ ที่ทำให้ค่า PBM ที่มากที่สุด 30 อันดับแรก นั่นคือ การพัฒนาดัชนี TOL เป็น TOL2 และ TOL3 ทำให้การจัดกลุ่มมดีขึ้น ค่า PBM ที่มากที่สุดมาจากการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาน้ำหนักทั้ง 4 ค่าเป็นดัชนีที่เกี่ยวกับค่าเฉลี่ย ซึ่งรวมถึง STI ด้วยเนื่องจากการศึกษาความสัมพันธ์กันของแต่ละดัชนี MP GMP และ STI มีความสัมพันธ์กันในระดับสูงมากอย่างมีนัยสำคัญ RWC เป็นดัชนี STI และ CMS เป็นดัชนี YSI

รายการอ้างอิง

1. กรมทรัพยากรธรณี. ดินเค็มคืออะไร [สืบค้นเมื่อ 9 ตุลาคม 2558]; เข้าถึงจาก: http://www.dmr.go.th/ewt_news.php?nid=1097&filename=saline_soil.
2. สำนักงานคณะกรรมการพิเศษเพื่อประสานงานโครงการอันเนื่องมาจากพระราชดำริ, คู่มือการจัดการดินเค็มเพื่อปลูกข้าว ศูนย์ศึกษาการพัฒนาภูพานอันเนื่องมาจากพระราชดำริ จังหวัดสกลนคร, ใน 19 ผลสำเร็จที่โดดเด่นของศูนย์ศึกษาการพัฒนาภูพานอันเนื่องมาจากพระราชดำริ. 2555, สำนักงานคณะกรรมการพิเศษเพื่อประสานงานโครงการอันเนื่องมาจากพระราชดำริ: บริษัท มูฟเม้นท์ เจน ทรี จำกัด. หน้า 16.
3. Goudarzi, M. and H. Pakniyat, *Evaluation of wheat cultivars under salinity stress based on some agronomic and physiological traits.* J. Agri. Soc. Sci., 2008. **4**(1): p. 4.
4. Talebi, R., F. Fayaz, and A.M. Naji, *Effective selection criteria for assessing drought stress tolerance in durum wheat (triticum durum desf.).* Gen. Appl. Plant Physiol., 2009. **35**: p. 11.
5. Kamyab-Talesh, F., et al., *Evaluation of some rice cultivars' response to salinity stress using resistance.* 2014. **60**: p. 12.
6. Franklin, D., *I Want the Mean, But not That One!*, in *PharmaSUG* 2016: Denver, Colorado.
7. Fernandez, G.C.J. *Effective selection criteria for assessing stress tolerance.* . in *Proceedings of the international symposium on adaptation of vegetables and other food crops in temperature and water stress*. 1992. Tainan. Taipei (Taiwan): Asian Vegetable Research and Development Center.
8. De Maesschalck, R., D. Jouan-Rimbaud, and D.L. Massart, *The Mahalanobis distance.* Chemometrics and Intelligent Laboratory Systems, 2000. **50**(1): p. 1-18.
9. Ming-Syan, C., H. Jiawei, and P.S. Yu, *Data mining: an overview from a database perspective.* IEEE Transactions on Knowledge and Data Engineering, 1996. **8**(6): p. 866-883.

10. Pakhira, M.K., S. Bandyopadhyay, and U. Maulik, *Validity index for crisp and fuzzy clusters.* Pattern Recognition, 2004. **37**(3): p. 487-501.

11. อัจฉริยา ปราบอริพ่าย, สถิติที่ไม่ใช้พารามิเตอร์. [สืบค้นเมื่อ 28 กรกฎาคม 2559]; เข้าถึงจาก: http://pirun.ku.ac.th/~faasatp/734415/data/chapter7.pdf.

ภาคผนวก ก
พันธุ์ข้าวชุดที่ 1

ตารางที่ ก.1 ตารางแสดงพันธุ์ข้าวชุดที่ 1

หมายเลข	ชื่อพันธุ์
1	ก้นแก้ว (1698)
2	เขียวหางม้า
3	ขาวสุพรรณ
4	หลวงประทาน
5	เหลืองพวงทอง
6	กข1
7	ดอขาว
8	มันวัว
9	พวงทอง
10	ประทานบ้านบุ้ง
11	ดอสามเดือน
12	หลวงประทาน
13	ขาวแก้ว
14	นางมล
15	อีมุม
16	เก้ารวง
17	ปิ่นแก้ว
18	ไข่มดริ้น
19	ใบบัว
20	มะยม
21	อีลาย
22	ขาวตาแห้ง17
23	หมากยม
24	หางยี71
25	เหลืองน้อย 31-1-39
26	จุดมอญ
27	ขาวสงวน

หมายเลข	ชื่อพันธุ์
28	เหนียวดำหลาย
29	ขาวกอเดียว
30	พวงหนัก
31	จำปาจีน
32	ขาวก้นจุด
33	เจ้าขาว
34	ก้นแก้ว (6158)
35	ทองรากไทร
36	หมากน้ำ (14085)
37	ชัยนาท1
38	เจ็กเชย
39	พวงหางหมู
40	พลายงามปราจีนบุรี
41	ตะเภาล่ม
42	ขาวอากาศ
43	ดีสี่
44	เม็ดมะม่วง
45	มะยม
46	ดอดอกไม้
47	หมากน้ำ (22835)
48	อีขาวใหญ่
49	ป้องแอ้ว (21577)
50	พ้าวฮ้าว
51	pokkali
52	IR29
53	LPT123-TC171 (LPT171)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ข
พันธุ์ข้าวชุดที่ 2

ตารางที่ ข.1 ตารางแสดงพันธุ์ข้าวชุดที่ 2

หมายเลข	ชื่อพันธุ์
1	หางนาค
2	เหลืองตายัง
3	ปลาซิวดำ
4	เหลืองงาม
5	เหนียวขาว
6	สามสี
7	ป้องแอ้ว (12159)
8	รากแห้ง
9	อินแปง
10	สำอางค์
11	หางหมาใน
12	หมากบิด
13	เหลืองดง
14	รวงเดียว
15	เล้าแตก
16	แม่หม้าย
17	ม่วยหิน
18	ไอไท้
19	ดอกไม้
20	สายหยุด
21	ห้ารวง
22	เหนียวมะลิ
23	เกษตรดอ
24	กำผาย
25	อีพวง
26	ชิวแม่จัน
27	ปทุมธานี1

หมายเลข	ชื่อพันธุ์
28	เกวียนหัก
29	ขาวตาแห้ง17
30	หางยี71
31	ไข่มดริ้น
32	บัวน้อย
33	นางงาม
34	อู่ตะเภา
35	พวงหวาน
36	ใบบัว
37	พวงหางหมู
38	สามรวงหนัก
39	ขาวกอเดียว
40	กข9
41	กข8
42	กข13
43	เจ้าดำ
44	เจ้าแดง
45	ดำด่าง
46	เจ้าดอกขาว
47	เหลืองทอง
48	ก่ำเลี้ยว
49	ทูลฉลอง
50	นางมล
51	ขาวประกวด
52	KDML105
53	CSSL 11
54	UBN02123-50R-B-3 (UBN)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ค

ค่าผิดปกติของข้อมูลชุดที่ 1

ตาราง ค.1 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่สูงกว่าปกติ

	SFW	SDW	RFW	RDW	RWC	CMS
Control	เจ้าขาว, หลวงประทาน	เจ้าขาว, หลวงประทาน	เจ้าขาว, หลวงประทาน	เจ้าขาว, หลวงประทาน	-	-
Salt	หลวงประทาน	หลวงประทาน	เขียวหางม้า, ทองรากไทร	ทองรากไทร, มันวัว, เขียวหางม้า, อีขาวใหญ่	-	-
TOL	เจ้าขาว	เจ้าขาว	เจ้าขาว, หลวงประทาน	เจ้าขาว, หลวงประทาน, ข้าวแก้ว	กข1	กข1, IR29
MP	หลวงประทาน	หลวงประทาน	หลวงประทาน, เจ้าขาว	เจ้าขาว, ทองรากไทร, หลวงประทาน	-	-
GMP	หลวงประทาน	หลวงประทาน	หลวงประทาน, ทองรากไทร	ทองรากไทร, หลวงประทาน	-	-
STI	หลวงประทาน	หลวงประทาน	หลวงประทาน, ทองรากไทร, เจ้าขาว	ทองรากไทร, หลวงประทาน, มันวัว, พลายงามปราจีนบุรี, เจ้าขาว	-	-
YI	หลวงประทาน	หลวงประทาน	เขียวหางม้า, ทองรากไทร	ทองรากไทร, มันวัว, เขียวหางม้า, อีขาวใหญ่	-	-
YSI	เหลืองพวงทอง, ประทานบ้านบุ่ง, อีมุม	-	เขียวหางม้า	ประทานบ้านบุ่ง, หางหยี71	-	-
HM	หลวงประทาน	หลวงประทาน	ทองรากไทร, หลวงประทาน	ทองรากไทร	-	-
TOL2	เจ้าขาว	เจ้าขาว	เจ้าขาว	เจ้าขาว	กข 1	IR29, กข1
TOL3	เจ้าขาว	เจ้าขาว	-	-	กข 1	IR29, กข1

ตาราง ค.2 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่ต่ำกว่าปกติ

	SFW	SDW	RFW	RDW	RWC	CMS
Control	-	-	-	-	ขาวอากาศ	ก้นแก้ว (6158), หางหยี71, พวงหางหมู, เจ็กเชย
Salt	-	-	-	-	กข1	IR29, กข1
TOL	ประทานบ้านบุ้ง, เหลืองพวงทอง	-	เขียวหางม้า	-	-	-
MP	-	-	-	-	อีขาวใหญ่,ขาวอากาศ, กข1	IR29, กข1
GMP	-	-	-	-	กข1, อีขาวใหญ่, ขาวอากาศ	IR29, กข1
STI	-	-	-	-	กข1, อีขาวใหญ่, ขาวอากาศ	IR29, กข1
YI	-	-	-	-	กข1	IR29, กข1
YSI	เจ้าขาว	-	-	-	กข1	IR29, กข1
HM	-	-	-	-	กข1, อีขาวใหญ่, ขาวอากาศ	IR29, กข1
TOL2	ประทานบ้านบุ้ง	-	เขียวหางม้า	ประทานบ้านบุ้ง, เขียวหางม้า	-	-
TOL3	เหลืองพวงทอง, ประทานบ้านบุ้ง	-	เขียวหางม้า	-	-	-

ตารางที่ ค.3 ตารางแสดงพันธุ์ข้าวข้อมูลชุดที่ 1 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีที่เหมือนกันและดัชนีที่ทำให้ค่า PBM มากที่สุด แล้วให้ค่าระยะทางมาฮาลาโนบิสที่สูงกว่าปกติ

	อันดับที่ 1	**อันดับที่ 2**	**อันดับที่ 3**	**อันดับที่ 4**	**อันดับที่ 5**
Control	ก้นแก้ว (6158)	เหลืองน้อย 31-1-39	-	-	-
Salt	-	-	-	-	-
TOL	เหลืองน้อย 31-1-39	เจ้าขาว	กข1	-	-
MP	อีขาวใหญ่	IR29	-	-	-
GMP	อีขาวใหญ่	IR29	-	-	-
STI	หลวงประทาน	ทองรากไทร	อีขาวใหญ่	-	-
YI	-	-	-	-	-
YSI	เหลืองน้อย31-1-39	เขียวหางม้า	ประทานบ้านบุ้ง	กข1	เหลืองพวงทอง
HM	อีขาวใหญ่	IR29	-	-	-
TOL2	เหลืองน้อย 31-1-39	เจ้าขาว	-	-	-
TOL3	เหลืองน้อย 31-1-39	กข1	IR29	-	-
The best index	-	-	-	-	-

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ง
ค่าผิดปกติของข้อมูลชุดที่ 2

ตาราง ง.1 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่สูงกว่าปกติ

	SFW	SDW	RFW	RDW	RWC	CMS
Control	-	เล้าแตก	เกษตรดอ, เล้าแตก	เกษตรดอ, เล้าแตก	เกษตรดอ, กข9, ไข่มดริ้น	-
Salt	เล้าแตก	เล้าแตก, อินแปง	เล้าแตก	เล้าแตก, อินแปง	-	-
TOL	ป้องแอ้ว	หางยี71	-	เกษตรดอ	อีพวง, สายหยุด	-
MP	เล้าแตก	เล้าแตก	เล้าแตก, เกษตรดอ	เล้าแตก, อินแปง, เกษตรดอ	-	-
GMP	เล้าแตก	เล้าแตก	เล้าแตก, เกษตรดอ	เล้าแตก, อินแปง	-	-
STI	เล้าแตก	เล้าแตก	เล้าแตก, เกษตรดอ	เล้าแตก, อินแปง, เกษตรดอ	-	-
YI	เล้าแตก	เล้าแตก, อินแปง	เล้าแตก	เล้าแตก, อินแปง	-	-
YSI	-	-	-	แม่หม้าย, กข9	-	-
HM	เล้าแตก	เล้าแตก	เล้าแตก, เกษตรดอ	เล้าแตก, อินแปง	-	-
TOL2	-	-	-	-	-	-
TOL3	-	อีพวง	-	-	-	อีพวง

ตาราง ง.2 ตารางแสดงพันธุ์ข้าวที่เมื่อ SFW SDW RFW RDW RWC และ CMS ของข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีต่าง ๆ แล้วเป็นค่าที่ต่ำกว่าปกติ

	SFW	SDW	RFW	RDW	RWC	CMC
Control	ไอไท้	ไอไท้	-	-	-	อีพวง, อู่ตะเภา, กำผาย, แม่หม้าย
Salt	ไอไท้	-	-	-	อีพวง, หางนาค	อีพวง
TOL	-	-	-	-	-	-
MP	ไอไท้	-	-	-	อีพวง, หางนาค	อีพวง
GMP	ไอไท้	-	-	-	อีพวง, หางนาค	อีพวง
STI	-	-	-	-	อีพวง, หางนาค	อีพวง
YI	ไอไท้	-	-	-	อีพวง, หางนาค	อีพวง
YSI	-	-	-	-	อีพวง, สายหยุด	อีพวง
HM	ไอไท้	-	-	-	อีพวง, หางนาค	อีพวง
TOL2	-	-	เล้าแตก	เล้าแตก	อีพวง, สายหยุด, เล้าแตก	-
TOL3	-	-	-	-	อีพวง, สายหยุด, เล้าแตก, ไอไท้	-

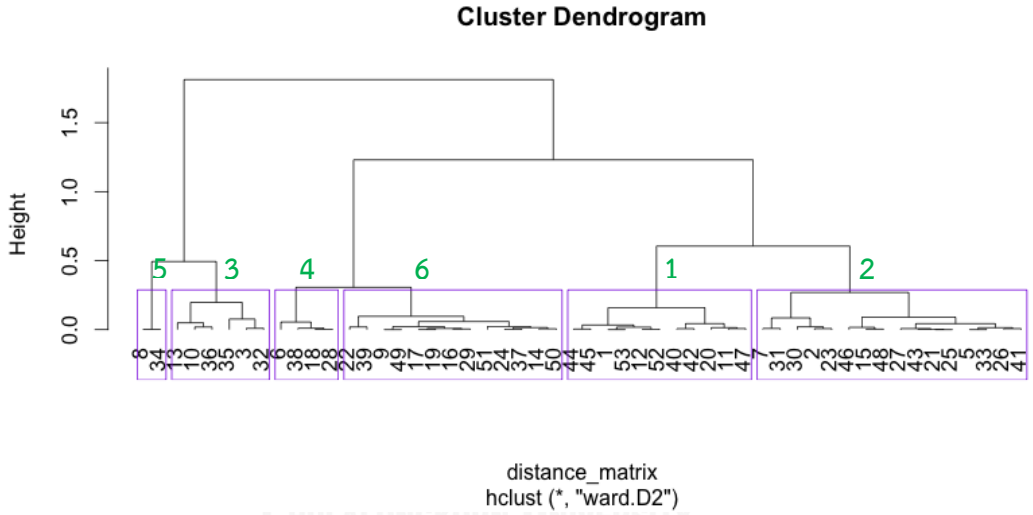
ตารางที่ ง.3 ตารางแสดงพันธุ์ข้าวข้อมูลชุดที่ 2 ที่ข้อมูลสรีรวิทยาหกคุณลักษณะที่ไม่ได้รับการแปลงด้วยดัชนี และได้รับการแปลงด้วยดัชนีที่เหมือนกันและดัชนีที่ทำให้ค่า PBM มากที่สุด แล้วให้ค่าระยะทางมาฮาลาโนบิสที่สูงกว่าปกติ

สูง	อันดับที่ 1	อันดับที่ 2	อันดับที่ 3	อันดับที่ 4
Control	-	-		
Salt	เล้าแตก	-	-	-
TOL	-	-	-	-
MP	อินแปง	เล้าแตก	อีพวง	ไอไท้
GMP	อินแปง	เล้าแตก	อีพวง	ไอไท้
STI	อินแปง	เล้าแตก	-	-
YI	เล้าแตก	-	-	-
YSI	แม่หม้าย	-	-	-
HM	เล้าแตก	อีพวง	อินแปง	ไอไท้
TOL2	เล้าแตก	-	-	-
TOL3	อีพวง	-	-	-
The best index	อินแปง	-	-	-

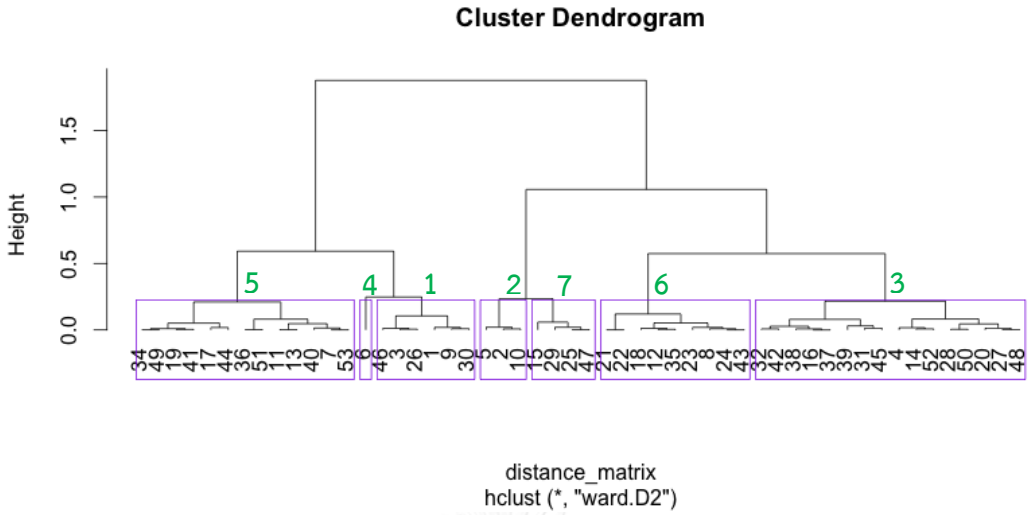
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก จ
ผลการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 1

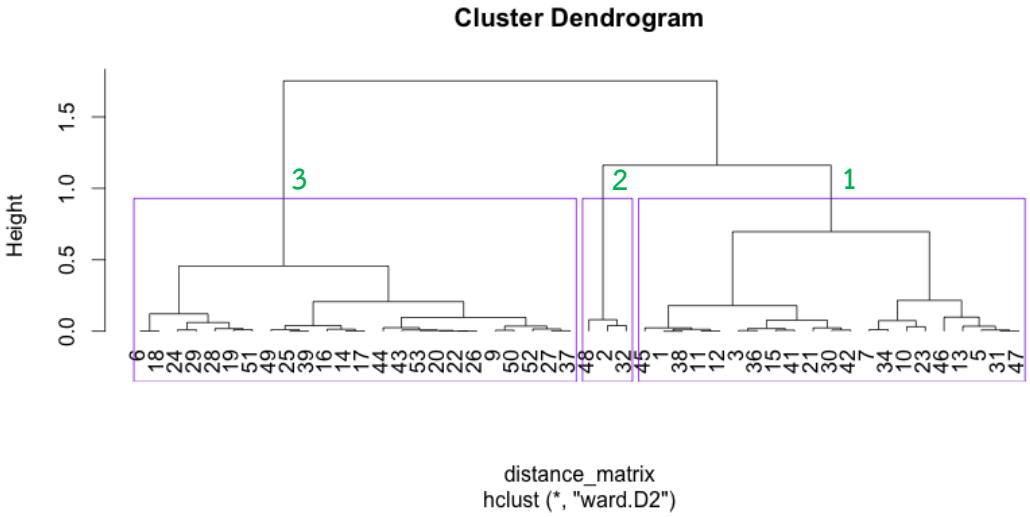
จาก รูปที่ จ.1 ถึง รูปที่ จ.4 แสดงพงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ข้อมูลสรีรวิทยา SFW SDW RFW RDW RWC และ CMS ตามลำดับ ของข้อมูลที่ได้รับการแปลงข้อมูลด้วยดัชนีที่เหมาะสม



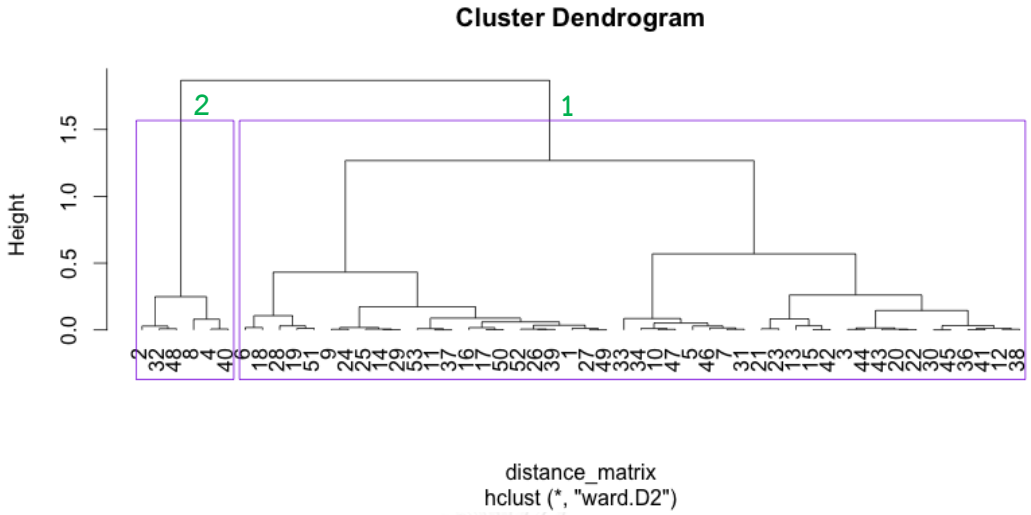
รูปที่ จ.1 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า SDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี STI



รูปที่ จ.2 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า SDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี TOL3



รูปที่ จ.3 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า RDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี STI

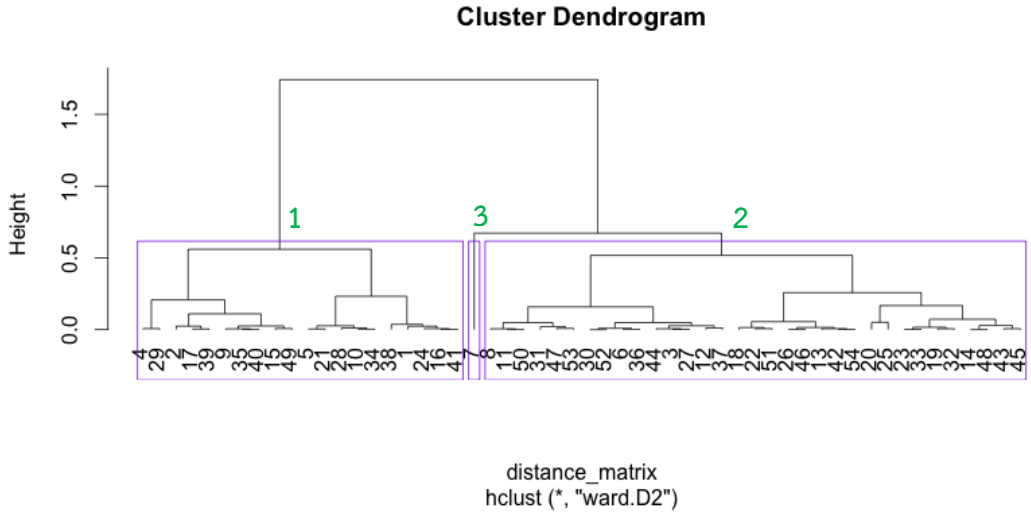


รูปที่ จ.4 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า RDW ของข้อมูลชุดที่ 1 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี HM

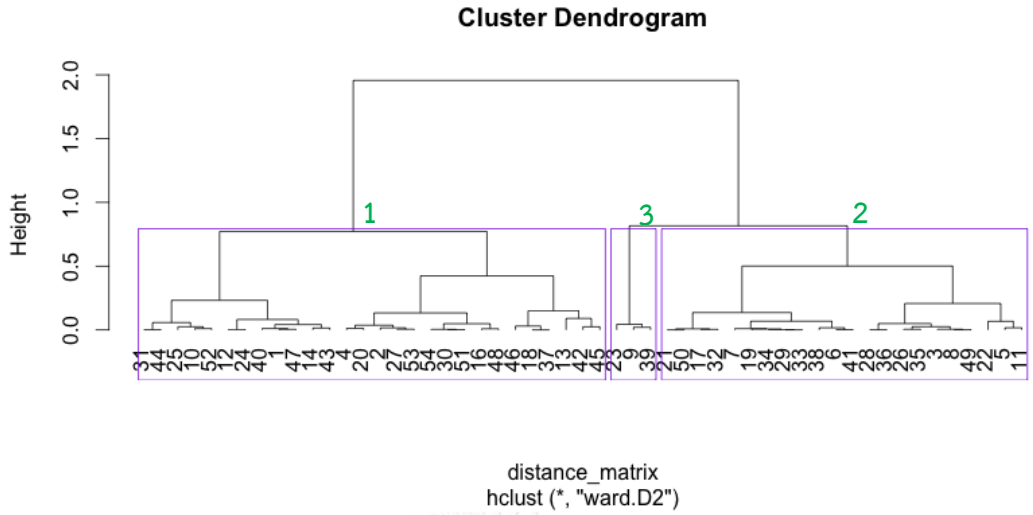
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ฉ
ผลการจัดกลุ่มพันธุ์ข้าวโดยใช้ข้อมูลสรีรวิทยาเพียงหนึ่งคุณลักษณะที่ได้รับการแปลงด้วยดัชนีที่เหมาะสมของข้อมูลชุดที่ 2

จากรูปที่ ฉ.1 ถึง รูปที่ ฉ.2 แสดงพงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ข้อมูลสรีรวิทยา SFW และ RFW ตามลำดับ ของข้อมูลที่ได้รับการแปลงข้อมูลด้วยดัชนีที่เหมาะสม



รูปที่ ฉ.1 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า SFW ของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี TOL2



รูปที่ ฉ.2 พงศาวลีจากการจัดกลุ่มข้าวตามลักษณะความทนเค็มของข้าวแต่ละพันธุ์โดยใช้ค่า RFW ของข้อมูลชุดที่ 2 ที่ได้รับการแปลงข้อมูลด้วยโดยดัชนี YI

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวพรชนิตว์ ทรัพย์วิไล เกิดเมื่อวันที่ 18 ธันวาคม พ.ศ. 2533 จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาตรี หลักสูตรวิทยาศาสตรบัณฑิต สาขาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ปี พ.ศ. 2555 และศึกษาต่อในระดับปริญญาโท สาขาคณิตศาสตร์ประยุกต์และวิทยาการคณนา ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปี พ.ศ. 2556

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University