

การวิเคราะห์รูปแบบฝนในจังหวัดขอนแก่นด้วยข้อมูลฝนจากดาวเทียม CHIRPS และวิธี Unsupervised Clustering Analyzing Rainfall Patterns in Khon Kaen Province using Satellite-based CHIRPS Rainfall Data and Unsupervised Clustering

ณัฐพล ศิลา^{1,2} ธัญธร ออวะละ² และ ชุพันธ์ ชมภูจันทร์^{2,*}

¹ สำนักบริหารโครงการ กรมชลประทาน 811 ถ.สามเสน แขวงถนนนครไชยศรี เขตดุสิต กรุงเทพมหานคร 10300

² ภาควิชาวิศวกรรมชลประทาน คณะวิศวกรรมศาสตร์ กำแพงแสน มหาวิทยาลัยเกษตรศาสตร์ จ.นครปฐม 73140

*Corresponding author; E-mail address: fengcpc@ku.ac.th

บทคัดย่อ

จังหวัดขอนแก่นเป็นแหล่งเพาะปลูกอ้อยที่สำคัญของประเทศไทย พื้นที่เพาะปลูกส่วนใหญ่อยู่นอกเขตชลประทานซึ่งมักประสบปัญหาภัยแล้งและฝนทิ้งช่วงส่งผลกระทบต่อเศรษฐกิจเติบโตและผลผลิต การทราบรูปแบบการตกของฝนในแต่ละปีจึงมีความสำคัญต่อการวางแผนการเพาะปลูกและการจัดการน้ำ การศึกษานี้มีวัตถุประสงค์เพื่อจำแนกรูปแบบการตกของฝนในจังหวัดขอนแก่น โดยใช้ข้อมูลฝนรายเดือน CHIRPS ในช่วงปี พ.ศ. 2524-2566 ทำการตรวจสอบความถูกต้องของข้อมูลกับสถานีอุตุนิยมวิทยาขอนแก่น (WMS Code 48381) พบว่ามีค่าสัมประสิทธิ์สหสัมพันธ์ (r) เท่ากับ 0.89 และค่า PBIAS เท่ากับ -5.94% อยู่ในเกณฑ์ระดับดีมาก การศึกษาใช้วิธี Silhouette Analysis กำหนดจำนวนกลุ่มที่เหมาะสมพบว่าสามารถจำแนกได้ 4 กลุ่ม และเปรียบเทียบการจัดกลุ่มด้วยวิธี K-means และ Hierarchical Clustering ผลการประเมินการจัดกลุ่มด้วยดัชนี Davies-Bouldin Index (DBI) พบว่า วิธี Hierarchical Clustering เป็นวิธีที่เหมาะสมที่สุดในการจำแนกรูปแบบฝน โดยแบ่งเป็น 4 กลุ่ม ได้แก่ กลุ่มที่ 1 มีฝนสูงต่อเนื่องช่วงกลางถึงปลายปี (13.95%) กลุ่มที่ 2 มีฝนตกหนักเฉพาะเดือนกันยายน (30.23%) กลุ่มที่ 3 มีฝนตกน้อยกระจายสม่ำเสมอตลอดฤดูฝน (23.26%) และกลุ่มที่ 4 มีฝนน้อยช่วงต้นปีแล้วเพิ่มขึ้นกลางปี (32.56%) โดยพบว่ากลุ่มที่ 4 มีความเสี่ยงต่อปริมาณฝนน้อยที่ส่งผลกระทบต่อผลผลิต และมีความเสี่ยงต่อภัยแล้งและฝนทิ้งช่วง ผลการศึกษานี้สามารถนำไปประยุกต์ใช้ในการกำหนดปฏิทินการเพาะปลูกที่เหมาะสม การวางแผนการให้น้ำเสริม และการประเมินผลผลิตตามรูปแบบการตกของฝนต่อไป

คำสำคัญ: การจัดกลุ่มข้อมูล, รูปแบบฝน, ข้อมูลฝน CHIRPS, พื้นที่เพาะปลูกอ้อย, จังหวัดขอนแก่น

Abstract

Khon Kaen Province is a significant area for sugarcane cultivation in Thailand, where most plantations rely on rainfed agriculture that often faces drought and prolonged dry spells, negatively impacting crop growth and yield. Understanding annual rainfall patterns is crucial for agricultural planning and water management. This study aimed to classify rainfall patterns in Khon Kaen Province using monthly CHIRPS rainfall data from 1981 to 2023. Data validation against the Khon Kaen Meteorological Station (WMS Code 48381) showed high accuracy, with

a correlation coefficient (r) of 0.89 and PBIAS of -5.94%. Silhouette Analysis determined the optimal clustering into four groups. The K-means and Hierarchical Clustering methods were compared, with the Davies-Bouldin Index (DBI) indicating that Hierarchical Clustering is the most suitable method. Four distinct patterns were identified: Group 1 with sustained high rainfall from mid to late year (13.95%), Group 2 with intense rainfall concentrated in September (30.23%), Group 3 with consistently low rainfall throughout the rainy season (23.26%), and Group 4 with low early-year rainfall that increases by mid-year (32.56%). Group 4 was identified as the most vulnerable to low rainfall, potentially affecting crop yield, with a high risk of drought and dry spells. These findings can be applied to optimize cropping calendars, plan supplementary irrigation, and assess yield potential based on rainfall patterns.

Keywords: Data Clustering, Rainfall Pattern, CHIRPS Rainfall Data, Sugarcane Cultivation Area, Khon Kaen Province

1. คำนำ

ภัยแล้งเป็นปัญหาที่สร้างความเสียหายอย่างมากต่อภาคเกษตรกรรมในประเทศไทย โดยเฉพาะในพื้นที่ภาคตะวันออกเฉียงเหนือของประเทศไทยซึ่งพื้นที่เพาะปลูกส่วนใหญ่อยู่นอกเขตชลประทานและต้องพึ่งพาน้ำฝนเป็นหลัก ทำให้เมื่อเกิดภาวะฝนทิ้งช่วงหรือฝนตกน้อยกว่าปกติ ส่งผลให้ความชื้นในดินลดลงและกระทบต่อการเจริญเติบโตและผลผลิตของพืช [1], [2] การรู้รูปแบบการตกของฝนรายเดือนในพื้นที่เกษตรกรรมที่อาศัยน้ำฝนเป็นหลักจึงมีความสำคัญอย่างยิ่งต่อการวางแผนการผลิตพืชอย่างมีประสิทธิภาพ ข้อมูลฝนรายเดือนช่วยให้เกษตรกรสามารถเลือกพืชที่เหมาะสมตามระดับปริมาณฝนที่คาดการณ์ไว้ รวมทั้งช่วยในการออกแบบความจุแหล่งกักเก็บน้ำ กำหนดการเตรียมแปลง เวลาเพาะปลูก ตารางการให้น้ำและใส่ปุ๋ยที่เหมาะสม [3]

การศึกษาก่อนและรูปแบบการตกของฝนในพื้นที่ จำเป็นต้องอาศัยข้อมูลตรวจวัดที่ต่อเนื่องยาวนานเพื่อประเมินความผันผวนของปริมาณน้ำฝน ตรวจสอบรูปแบบการเปลี่ยนแปลงของสภาพภัยแล้ง [4] ในประเทศไทยมักประสบปัญหาข้อจำกัดของข้อมูลสถานีตรวจวัดน้ำฝนเนื่องจากการกระจายตัวของสถานีไม่ครอบคลุมพื้นที่และข้อมูลฝนที่ไม่มีความต่อเนื่อง ปัจจุบันมีชุดข้อมูลฝนดาวเทียมเป็นทางเลือกที่มีศักยภาพทดแทนกรณีข้อมูล

สถานีตรวจวัดน้ำฝนไม่เพียงพอ อาทิ ข้อมูล CHIRPS (Climate Hazards Group Infrared Precipitation with Stations) ซึ่งมีความละเอียดเชิงพื้นที่สูง (0.05°) และมีข้อมูลระยะยาว (ตั้งแต่ปี ค.ศ. 1981) ทำให้สามารถประเมินลักษณะของปริมาณน้ำฝนและภัยแล้งได้อย่างละเอียด อีกทั้งมีความแม่นยำสูงเมื่อเปรียบเทียบกับข้อมูลตรวจวัดภาคพื้นดิน [5]

จังหวัดขอนแก่นมีพื้นที่เกษตรกรรมส่วนใหญ่นอกเขตชลประทาน ทำให้การทำเกษตรต้องพึ่งพาน้ำฝนเป็นหลัก การศึกษารูปแบบการตกของฝนในจังหวัดจึงมีความสำคัญอย่างยิ่งต่อการวางแผนเพาะปลูกพืชเศรษฐกิจ โดยเฉพาะอย่างยิ่งเป็นพืชเศรษฐกิจสำคัญของจังหวัดที่จำเป็นต้องอาศัยปริมาณน้ำฝนที่เพียงพอต่อการเจริญเติบโตและให้ผลผลิตที่ดี งานวิจัยนี้เลือกใช้ข้อมูลฝนรายเดือนจากผลิตภัณฑ์ CHIRPS มาวิเคราะห์รูปแบบการตกของฝนรายเดือนในจังหวัดขอนแก่น โดยใช้วิธีการจัดกลุ่มแบบ Unsupervised Clustering เพื่อจำแนกรูปแบบการตกของฝนที่แตกต่างกัน ซึ่งจะช่วยให้เข้าใจลักษณะการเกิดภัยแล้งและฝนทิ้งช่วงในพื้นที่ได้ดียิ่งขึ้น ผลลัพธ์จากการศึกษานี้จะมีประโยชน์อย่างยิ่งต่อการกำหนดปฏิทินการเพาะปลูกที่เหมาะสม การวางแผนการให้น้ำเสริม และการประเมินผลผลิตตามรูปแบบการตกของฝน อันจะนำไปสู่การเพิ่มประสิทธิภาพในการผลิตและการลดความเสียหายจากภัยแล้งได้อย่างยั่งยืน

2. ระเบียบวิธีการศึกษา

2.1 พื้นที่ศึกษา

จังหวัดขอนแก่นมีพื้นที่ประมาณ 10,886 ตารางกิโลเมตร อยู่ในเขตลุ่มน้ำมูลและลุ่มน้ำชี โดยมีแม่น้ำสายสำคัญไหลผ่าน คือ แม่น้ำชีและแม่น้ำพอง ลักษณะภูมิประเทศของจังหวัดขอนแก่นส่วนใหญ่ (มากกว่า 80%) เป็นพื้นที่ราบถึงพื้นที่ลาดเอียงเล็กน้อย อยู่สูงจากระดับน้ำทะเลปานกลาง 100-200 เมตร และมีที่ราบลุ่มกว้างตามแนวแม่น้ำชี ทางทิศตะวันตกเป็นพื้นที่ภูเขาสูงบริเวณเขตภูกระดึงและเทือกเขาเพชรบูรณ์ ส่วนทางทิศตะวันออกและทิศตะวันออกเฉียงใต้มีลักษณะเป็นพื้นที่ลูกคลื่นสูงต่ำสลับกันลาดเทไปทางทิศตะวันออกและทิศใต้ของจังหวัด (Figure 1) สภาพภูมิอากาศของจังหวัดขอนแก่นมีฤดูฝนตั้งแต่เดือนพฤษภาคมถึงตุลาคม และส่วนเดือนอื่นๆ มีฝนตกน้อยมากการกระจายตัวของฝนไม่สม่ำเสมอทั้งปีเดียวกันและระหว่างปี โดยมีปริมาณฝนเฉลี่ยต่อปี 1,290 มิลลิเมตร จังหวัดขอนแก่นมีพื้นที่เพาะปลูกทั้งหมด 4,369,044 ไร่ (คิดเป็น 64.19% ของพื้นที่ทั้งหมด) แต่มีเพียง 614,681 ไร่ (14%) ที่อยู่ในเขตชลประทาน พืชเศรษฐกิจหลักคือข้าว ซึ่งมีพื้นที่ปลูก 58.9% ของพื้นที่เกษตรกรรมทั้งหมด รองลงมาคือถั่วและมันสำปะหลัง ซึ่งมีพื้นที่ปลูก 12.9% และ 4.6% ของพื้นที่เกษตรกรรมตามลำดับ [6]

2.2 การรวบรวมและตรวจสอบข้อมูลน้ำฝน

การศึกษานี้ใช้ข้อมูลฝนรายเดือนจากผลิตภัณฑ์ CHIRPS (Climate Hazards Group Infrared Precipitation with Stations) ซึ่งเป็นชุดข้อมูลที่พัฒนาโดยสำนักงานสำรวจธรณีวิทยาแห่งสหรัฐอเมริกา (USGS) และ Climate Hazards Center มหาวิทยาลัยแคลิฟอร์เนีย ซานตาบาร์บารา ข้อมูล CHIRPS ได้จากการแปลผลภาพถ่ายดาวเทียมในช่วงคลื่นอินฟราเรดความร้อนและผ่านการปรับแก้ความเอนเอียงด้วยข้อมูลฝนจากสถานีตรวจวัดภาคพื้นดิน มีความต่อเนื่องตั้งแต่ปี ค.ศ. 1981 จนถึงปัจจุบัน และมีความละเอียดเชิงพื้นที่ 0.05° × 0.05° (ประมาณ 5 กิโลเมตร) [7] จากการซ้อนทับขอบเขตจังหวัดขอนแก่นกับชั้นข้อมูล CHIRPS ในโปรแกรม QGIS พบว่ามีจำนวนจุดภาพ (pixel) ทั้งหมด 357 pixel ที่อยู่ในพื้นที่ศึกษา

การตรวจสอบความถูกต้องของข้อมูลฝน CHIRPS ทำโดยเปรียบเทียบกับข้อมูลปริมาณน้ำฝนรายเดือนระหว่างปี พ.ศ. 2524-2566 จากสถานี

ตรวจวัดน้ำฝนของกรมอุตุนิยมวิทยาขอนแก่น (WMS Code 48381) วิเคราะห์ความสัมพันธ์ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ (r) และค่าเปอร์เซ็นต์ความเอนเอียง (PBIAS) ดังสมการที่ (1) และ (2) ค่า r ที่เข้าใกล้ +1.0 แสดงถึงความสัมพันธ์โดยตรงอย่างมากในทิศทางเดียวกัน ส่วนค่า PBIAS ที่เข้าใกล้ 0 บ่งชี้ว่าข้อมูล CHIRPS มีความน่าเชื่อถือสูง การประเมินผลใช้เกณฑ์ตาม Table 1 ซึ่งแสดงระดับคุณภาพของข้อมูลตามค่า r และ PBIAS

$$r = \frac{\sum_{i=1}^n (D_i - \bar{D})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^n (D_i - \bar{D})^2} \cdot \sqrt{\sum_{i=1}^n (G_i - \bar{G})^2}} \quad (1)$$

$$PBIAS = \frac{\sum_{i=1}^n (D_i - G_i)}{\sum_{i=1}^n G_i} \times 100\% \quad (2)$$

โดย

n คือ จำนวนสถานีตรวจวัดน้ำฝน

D_i คือ ปริมาณน้ำฝนจากชุดข้อมูลฝน (Precipitation Dataset)

G_i คือ ปริมาณน้ำฝนจากสถานีตรวจวัดน้ำฝน (Rain Gauge Station)

$\bar{D}$ คือ ค่าเฉลี่ยปริมาณน้ำฝนจากชุดข้อมูลฝน

$\bar{G}$ คือ ค่าเฉลี่ยปริมาณน้ำฝนจากสถานีตรวจวัดน้ำฝน

Table 1 Performance rating as a function of r and PBIAS [8]

r	Rating	PBIAS	Rating
0.0-0.1	Very Low		
0.1-0.3	Low	PBIAS ≤ ±10%	Very Good
0.3-0.5	Moderate	±10% < PBIAS ≤ ±15%	Good
0.5-0.7	High	±15% < PBIAS ≤ ±25%	Satisfactory
0.7-0.9	Very High	PBIAS ≥ ±25%	Unsatisfactory
0.9-1.0	Almost Perfect		

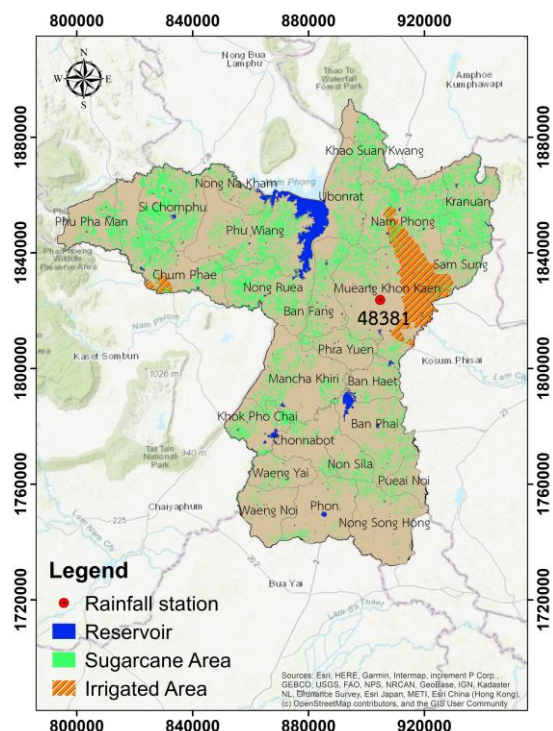


Figure 1 Study area and Rainfall stations

2.3 วิธีการจัดกลุ่ม

วิธีการจัดกลุ่มข้อมูล (Clustering Methods) เป็นเทคนิคการเรียนรู้แบบ Unsupervised Learning ที่มีวัตถุประสงค์เพื่อแบ่งข้อมูลออกเป็นกลุ่มที่มีลักษณะคล้ายคลึงกันภายในกลุ่มเดียวกัน แต่แตกต่างกันระหว่างกลุ่ม การศึกษาใช้วิธีการจัดกลุ่ม 2 วิธี คือ Hierarchical Clustering และ K-means Clustering เพื่อจำแนกรูปแบบการตกของฝนในจังหวัดขอนแก่น และเปรียบเทียบประสิทธิภาพของทั้งสองวิธี

วิธีการจัดกลุ่มแบบ Hierarchical Clustering ใช้แนวคิดการรวมกลุ่มแบบขั้นบันได (Agglomerative) โดยเริ่มต้นจากการถือว่าแต่ละจุดข้อมูลเป็นกลุ่มของตัวเอง จากนั้นจะทำการรวมกลุ่มที่มีความคล้ายคลึงกันมากที่สุดเข้าด้วยกันทีละคู่ตามเกณฑ์การเชื่อมโยง (Linkage Criterion) ที่เลือก จนกระทั่งเหลือเพียงกลุ่มเดียว กระบวนการนี้จะสร้างแผนภูมิต้นไม้ (Dendrogram) ที่แสดงลำดับการรวมกลุ่ม ดังแสดงตัวอย่างใน Figure 2 โดย Dendrogram เริ่มต้นจากข้อมูล 10 หน่วย และใช้เส้นประแนวนอนเป็นจุดตัด (Cutoff Point) เพื่อกำหนดจำนวนกลุ่มที่เหมาะสม ซึ่งในที่นี้ได้ 3 กลุ่ม ทั้งนี้ ระยะห่างในแนวตั้งของแผนภูมิแสดงถึงระดับความแตกต่างระหว่างกลุ่ม [9] สำหรับการศึกษาใช้เกณฑ์การเชื่อมโยงใช้วิธี Ward Linkage หรือเรียกว่าวัดความแปรปรวนต่ำสุด โดยหลักการคือการเลือกรวมคู่กลุ่มที่ทำให้ความแปรปรวนภายในกลุ่ม (Within-Cluster Variance) หรือผลรวมกำลังสองของความคลาดเคลื่อน (Sum of Squared Error, SSE) เพิ่มขึ้นน้อยที่สุด ซึ่งจะทำให้ได้กลุ่มข้อมูลที่มีลักษณะใกล้เคียงกันมากภายในกลุ่มเดียวกัน ตามสมการ:

$$\Delta(C_i, C_j) = \min \frac{n_i n_j}{n_i + n_j} \|u_i - u_j\|^2 \quad (3)$$

โดย

$\Delta(C_i, C_j)$ คือการเพิ่มขึ้นของความแปรปรวนเมื่อรวมกลุ่ม C_i และ C_j
 n_i และ n_j คือจำนวนข้อมูลในกลุ่ม C_i และ C_j ตามลำดับ
 u_i และ u_j คือจุดศูนย์กลางของกลุ่ม C_i และ C_j ตามลำดับ
 $\|u_i - u_j\|^2$ คือระยะห่างแบบยูคลิดกำลังสองระหว่างจุดศูนย์กลางของสองกลุ่ม

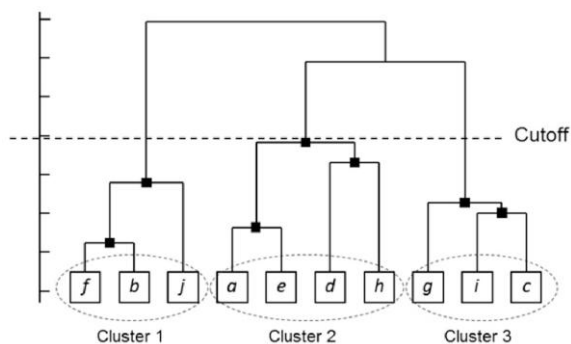


Figure 2 Hierarchical clustering dendrogram showing 10 data units grouped into 3 clusters at the cutoff point. [9]

วิธีการจัดกลุ่มแบบ K-means Clustering เป็นอัลกอริทึมแบบวนซ้ำ (Iterative Algorithm) มีเป้าหมายเพื่อแบ่งข้อมูล n จุดออกเป็น k กลุ่ม โดยจำนวน k จะต้องกำหนดไว้ล่วงหน้า หลักการทำงานเริ่มจากการกำหนดจุดศูนย์กลาง (Centroid) ของแต่ละกลุ่มแบบสุ่ม จากนั้นจะจัดข้อมูลแต่ละ

จุดให้อยู่ในกลุ่มที่มีจุดศูนย์กลางที่ใกล้ที่สุด และคำนวณจุดศูนย์กลางใหม่จากค่าเฉลี่ยของข้อมูลในแต่ละกลุ่ม กระบวนการนี้จะทำซ้ำไปเรื่อย ๆ จนกว่าตำแหน่งของจุดศูนย์กลางจะไม่เปลี่ยนแปลง (Figure 3) โดยมีฟังก์ชันวัตถุประสงค์คือการลดผลรวมกำลังสองของระยะทางภายในกลุ่ม (Within-Cluster Sum of Squares, WCSS) หรือผลรวมของระยะทางแบบยูคลิดยกกำลังสอง (Squared Euclidean Distances) ระหว่างจุดข้อมูลแต่ละจุดกับจุดศูนย์กลาง (Centroid) ของกลุ่มให้มีค่าน้อยที่สุด ซึ่งจะทำให้ข้อมูลในกลุ่มเดียวกันมีความคล้ายคลึงกันมากที่สุด [10] ตามสมการ:

$$WCSS = \min_c \sum_{i=1}^K \sum_{x_j \in C_i} |x_j - u_i|^2 \quad (4)$$

โดย

$|x_j - u_i|^2$ คือระยะทางระหว่างจุดข้อมูลกับจุดศูนย์กลางของกลุ่ม

x_j คือตำแหน่งของจุดข้อมูล

u_i คือตำแหน่งของจุดศูนย์กลางของกลุ่มที่ i โดย $i=1, \dots, K$

K คือจำนวนกลุ่มทั้งหมด

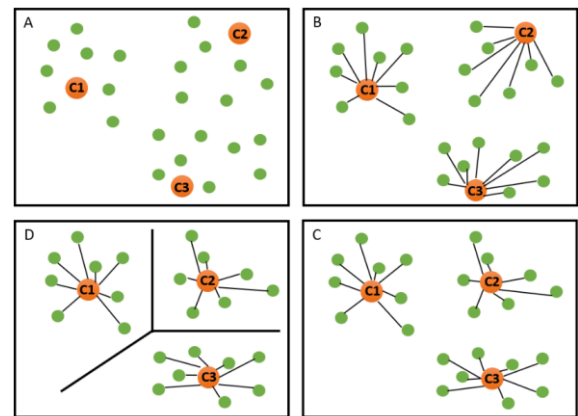


Figure 3 K-means clustering process: initialization (A), first assignments (B), final assignment (C), and final partition (D). [10]

2.4 การกำหนดจำนวนกลุ่มและการประเมินประสิทธิภาพการจัดกลุ่ม

การศึกษานี้ใช้วิธี Silhouette Analysis เพื่อกำหนดจำนวนกลุ่มที่เหมาะสม วิธีนี้คำนวณค่า Silhouette Score โดยหาค่าเฉลี่ยของ Silhouette Coefficient ของข้อมูลทั้งหมด โดย Silhouette Coefficient คำนวณจากระยะห่างเฉลี่ยภายในกลุ่ม (a) และระยะห่างเฉลี่ยไปยังกลุ่มที่ใกล้ที่สุด (b) สำหรับแต่ละจุดข้อมูล [11] ตามสมการ:

$$S = \frac{b-a}{\max(a,b)} \quad (5)$$

ค่า Silhouette Score มีช่วงตั้งแต่ -1 ถึง 1 โดยค่าที่ใกล้ 1 หมายถึงกลุ่มมีการแยกตัวออกจากกลุ่มอื่นอย่างชัดเจน และจุดข้อมูลในกลุ่มมีการกระจุกตัวดี ค่าที่ใกล้ 0 หมายถึง กลุ่มต่าง ๆ มีขอบเขตที่ทับซ้อนกันและไม่แยกจากกันชัดเจน ส่วนค่าที่ใกล้ -1 หมายถึง จุดข้อมูลถูกจัดอยู่ในกลุ่มที่ไม่เหมาะสม

หลังจากกำหนดจำนวนกลุ่มที่เหมาะสมแล้ว การศึกษานี้ได้เปรียบเทียบประสิทธิภาพของวิธี Hierarchical Clustering และ K-means Clustering โดยใช้ Davies-Bouldin Index (DBI) เป็นเกณฑ์ในการประเมิน DBI วัดคุณภาพของการจัดกลุ่ม แนวคิดพื้นฐานคือการประเมินการแยกตัวระหว่าง

กลุ่มที่ i และกลุ่มที่ j ซึ่งควรมีระยะห่างระหว่างกลุ่มมากที่สุดและมีการกระจายตัวภายในกลุ่มน้อยที่สุด DBI มีความสัมพันธ์เชิงบวกกับการกระจายตัวภายในกลุ่ม และมีความสัมพันธ์เชิงลบกับระยะห่างระหว่างกลุ่ม สมการในการคำนวณค่า Davies-Bouldin Index คือ [12]:

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{j \neq i} \left\{ \frac{S_i + S_j}{d_{i,j}} \right\} \quad (6)$$

โดย

$S_i = \frac{1}{|C_i|} \sum_{x_j \in C_i} \|x_j - v_i\|$ คือค่าวัดการกระจายตัวภายในกลุ่มที่ i

K คือจำนวนกลุ่ม

x_j คือเวกเตอร์คุณลักษณะ n มิติที่ถูกจัดอยู่ในกลุ่มที่ i

v_i คือจุดศูนย์กลางของกลุ่มที่ i

C_i คือกลุ่มที่ i

$\|x_j - v_i\|$ คือระยะทางแบบยูคลิด

$d_{i,j} = \|v_i - v_j\|$ คือระยะห่างระหว่างจุดศูนย์กลางของกลุ่มที่ i และ j

กระบวนการวิเคราะห์ในการศึกษานี้เริ่มจากการใช้ Silhouette Analysis เพื่อกำหนดจำนวนกลุ่มที่เหมาะสม จากนั้นทำการจัดกลุ่มรูปแบบด้วยวิธี Hierarchical Clustering และ K-means Clustering และเปรียบเทียบประสิทธิภาพของทั้งสองวิธีด้วยค่า DBI วิธีที่ให้ค่า DBI ต่ำกว่าจะถูกเลือกเป็นวิธีที่เหมาะสมที่สุดในการจำแนกรูปแบบฝนในจังหวัดขอนแก่น

3. ผลการศึกษาและวิจารณ์ผล

3.1 ผลการประเมินความถูกต้องของข้อมูลฝน CHIRPS

การประเมินปริมาณฝนรายเดือนจากข้อมูล CHIRPS เทียบกับข้อมูลจากสถานีตรวจวัดน้ำฝนของกรมอุตุนิยมวิทยา (WMS Code 48381) พบว่าข้อมูลฝน CHIRPS มีความสัมพันธ์กันโดยตรงอย่างมากกับข้อมูลสถานีตรวจวัดโดยมีค่าสัมประสิทธิ์สหสัมพันธ์ (r) เท่ากับ 0.89 ซึ่งอยู่ในเกณฑ์ความเที่ยงของข้อมูลระดับสูงมาก ตามที่แสดงดัง Figure 4 และค่าเปอร์เซ็นต์ความเอนเอียง (PBIAS) เท่ากับ -5.94% ซึ่งถือว่า PBIAS อยู่ในช่วง $\pm 10\%$ อยู่ในเกณฑ์ความแม่นยำของข้อมูลระดับดีมาก

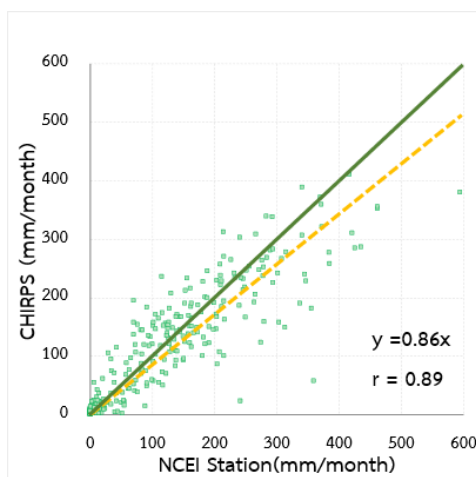


Figure 4 Correlation between CHIRPS rainfall data and Khon Kaen station observations (WMS Code 48381)

3.2 การกำหนดจำนวนกลุ่มที่เหมาะสมและการเปรียบเทียบประสิทธิภาพจัดกลุ่ม

จากการวิเคราะห์ด้วยวิธี Silhouette Analysis เพื่อหาจำนวนกลุ่มที่เหมาะสมสำหรับการจัดกลุ่มรูปแบบฝน พบว่าค่า Silhouette Score ที่ดีที่สุดเท่ากับ 0.205 คือจำนวน 2 กลุ่ม แต่เนื่องจากผู้วิจัยต้องการจำแนกลักษณะการเกิดภัยแล้งและฝนทิ้งช่วงให้ละเอียดยิ่งขึ้น จึงได้พิจารณาเลือกจำนวน 4 กลุ่ม ซึ่งมีค่า Silhouette Score เท่ากับ 0.191 ซึ่งเป็นค่าที่มากเป็นอันดับที่ 2 ดังแสดงใน Figure 5

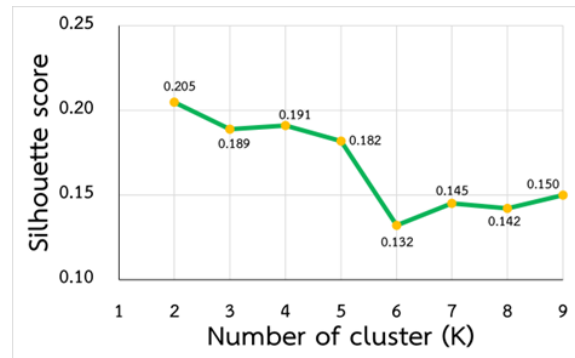


Figure 5 Silhouette analysis for optimal number of clusters

เมื่อได้กำหนดจำนวนกลุ่มแล้ว ได้ทำการเปรียบเทียบประสิทธิภาพระหว่างวิธีการจัดกลุ่มแบบ Hierarchical Clustering และ K-means Clustering โดยใช้ค่า Davies-Bouldin Index (DBI) เป็นเกณฑ์ในการประเมิน ผลการวิเคราะห์พบว่าวิธี Hierarchical Clustering มีค่า DBI เท่ากับ 1.764 ซึ่งต่ำกว่าวิธี K-means ที่มีค่า DBI เท่ากับ 1.971 แสดงว่าวิธี Hierarchical Clustering มีประสิทธิภาพในการจำแนกรูปแบบฝนได้ดีกว่า ดังนั้นจึงเลือกใช้ผลลัพธ์จากวิธี Hierarchical Clustering ในการวิเคราะห์รูปแบบฝนในขั้นตอนต่อไป ผลการจัดกลุ่มด้วยวิธี Hierarchical Clustering ใช้ข้อมูลรายปี (แกน X) และวัดความแตกต่างด้วยค่า Distance (แกน Y) โดยปีที่มีรูปแบบของข้อมูลฝนคล้ายคลึงกันจะมีค่า Distance ต่ำและถูกรวมเป็นกลุ่มเดียวกัน ซึ่งการรวมกลุ่มจะทำต่อเนื่องจนได้แผนภาพ dendrogram และมีการตัดกลุ่มที่ค่า Distance เท่ากับ 370.59 เพื่อแบ่งข้อมูลออกเป็น 4 กลุ่มหลัก คือ H.1, H.2, H.3 และ H.4 ดังแสดงใน Figure 6

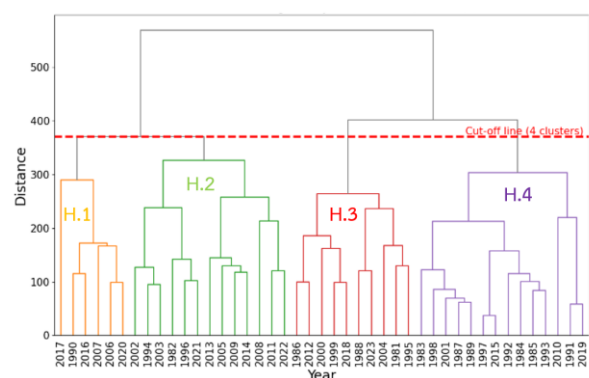


Figure 6 Hierarchical clustering dendrogram showing four rainfall pattern groups

3.3 การวิเคราะห์รูปแบบการตกของฝน

เมื่อวิเคราะห์รูปแบบการตกของฝนจากผลการจัดกลุ่มด้วยวิธี Hierarchical Clustering ทั้ง 4 กลุ่ม สามารถอธิบายลักษณะเฉพาะของแต่ละกลุ่มได้ดังนี้ (Figure 7)

กลุ่มที่ 1 (H.1 - เส้นสีเหลือง) มีปริมาณฝนสะสมเฉลี่ยรายปี 1,417.81 มิลลิเมตร ซึ่งสูงที่สุดในบรรดาทั้ง 4 กลุ่ม ในช่วงต้นฤดูฝน (พฤษภาคม) มีปริมาณฝนปานกลาง ช่วงฝนทิ้งช่วง (มิถุนายน-กรกฎาคม) มีปริมาณฝนลดลง ส่วนช่วงปลายฤดูฝน (สิงหาคม-ตุลาคม) มีปริมาณฝนสูงอย่างต่อเนื่อง โดยเฉพาะเดือนตุลาคมยังคงมีฝนมากกว่ากลุ่มอื่นๆ อย่างชัดเจน

กลุ่มที่ 2 (H.2 - เส้นสีเขียว) มีปริมาณฝนสะสมเฉลี่ยรายปี 1,318.84 มิลลิเมตร ช่วงต้นฤดูฝน (พฤษภาคม) มีปริมาณฝนปานกลาง ช่วงฝนทิ้งช่วง (มิถุนายน-กรกฎาคม) มีปริมาณฝนลดลงชัดเจนแล้วค่อยๆ เพิ่มขึ้น ช่วงปลายฤดูฝน (สิงหาคม-ตุลาคม) มีลักษณะเด่นคือปริมาณฝนสูงมากในเดือนกันยายน (สูงกว่า 300 มิลลิเมตร) ซึ่งสูงที่สุดในบรรดาทั้ง 4 กลุ่ม

กลุ่มที่ 3 (H.3 - เส้นสีแดง) มีปริมาณฝนสะสมเฉลี่ยรายปี 1,197.46 มิลลิเมตร มีลักษณะเด่นที่ช่วงต้นฤดูฝน (พฤษภาคม) มีปริมาณฝนสูงที่สุดในบรรดาทั้ง 4 กลุ่ม ช่วงฝนทิ้งช่วง (มิถุนายน-กรกฎาคม) มีปริมาณฝนลดลงแล้วเพิ่มขึ้นในระดับที่สูงกว่ากลุ่มอื่นๆ ส่วนช่วงปลายฤดูฝน (สิงหาคม-ตุลาคม) กลับมีปริมาณฝนต่ำกว่ากลุ่มอื่นๆ และค่อนข้างคงที่ไม่ผันผวนมาก

กลุ่มที่ 4 (H.4 - เส้นสีม่วง) มีปริมาณฝนสะสมเฉลี่ยรายปี 1,075.48 มิลลิเมตร ซึ่งต่ำที่สุดในบรรดาทั้ง 4 กลุ่ม ช่วงต้นฤดูฝน (พฤษภาคม) มีปริมาณฝนต่ำที่สุดเมื่อเทียบกับกลุ่มอื่นๆ ช่วงฝนทิ้งช่วง (มิถุนายน-กรกฎาคม) มีปริมาณฝนต่ำและค่อยๆ เพิ่มขึ้นเล็กน้อย ส่วนช่วงปลายฤดูฝน (สิงหาคม-ตุลาคม) มีการเพิ่มขึ้นอย่างรวดเร็วของปริมาณฝนในเดือนสิงหาคมและกันยายน โดยมีค่าสูงสุดในเดือนสิงหาคม แล้วค่อยๆ ลดลงในเดือนตุลาคม

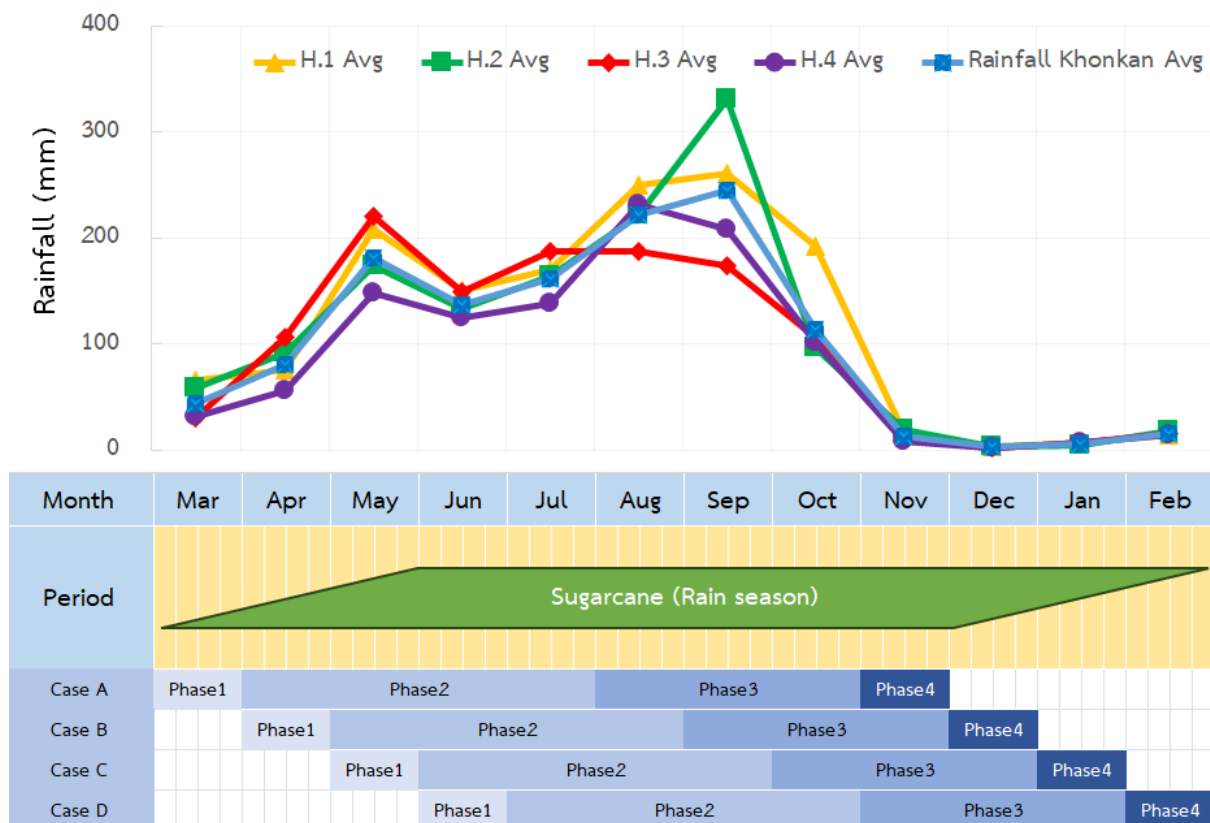


Figure 7 Rainfall patterns with sugarcane cultivation schedule

การศึกษานี้ได้วิเคราะห์ผลกระทบของรูปแบบฝนทั้ง 4 กลุ่มต่อการเพาะปลูกอ้อยในจังหวัดขอนแก่น โดยพิจารณาช่วงเวลาเริ่มต้นปลูกที่แตกต่างกัน 4 ช่วงเวลา ได้แก่: Case A: เริ่มเพาะปลูกเดือนมีนาคม, Case B: เริ่มเพาะปลูกเดือนเมษายน, Case C: เริ่มเพาะปลูกเดือนพฤษภาคม, Case D: เริ่มเพาะปลูกเดือนมิถุนายน โดยแต่ละกรณีจะมีระยะการเจริญเติบโตของอ้อยแตกต่างกันตามช่วงเวลา โดยแบ่งเป็น 4 ระยะ ได้แก่ Phase 1 (การตั้งตัว) Phase 2 (เติบโตทางลำต้น) Phase 3 (สร้างน้ำตาล) และ Phase 4 (สุกแก่) ดังแสดงใน Figure 7

จากการวิเคราะห์ปริมาณน้ำฝนสะสมในแต่ละระยะการเจริญเติบโตของอ้อยภายใต้รูปแบบฝนทั้ง 4 กลุ่มและเวลาเริ่มต้นปลูกที่แตกต่างกัน (Table

2) พบว่าประสิทธิภาพของการปลูกอ้อยขึ้นอยู่กับความสมดุลของปริมาณน้ำฝนในแต่ละ Phase การเจริญเติบโต โดยกรณีที่มีความเสี่ยงต่อผลผลิตมากที่สุดคือกรณี H.4 Case A ซึ่งเริ่มปลูกในเดือนมีนาคม ทั้งนี้ กรณีดังกล่าวอยู่ภายใต้รูปแบบฝน H.4 ที่แม้จะมีปริมาณน้ำฝนรวมสูงในช่วงกลางปี แต่ขาดความสมดุลในแต่ละระยะการเจริญเติบโต โดยใน Phase 1 ช่วงเดือนมีนาคม มีปริมาณน้ำฝน 31.93 มิลลิเมตร Phase 2 ระหว่างเดือนเมษายนถึงกรกฎาคม มีปริมาณน้ำฝน 468.36 มิลลิเมตร Phase 3 ระหว่างเดือนสิงหาคมถึงตุลาคม มีปริมาณน้ำฝน 542.56 มิลลิเมตร และ Phase 4 ช่วงเดือนพฤศจิกายน มีปริมาณน้ำฝน 8.79 มิลลิเมตร ทำให้ปริมาณน้ำที่อ้อยต้องการไม่เพียงพอในทุกระยะการเติบโต อาจส่งผลให้ความชื้นในดินมี

ไม่เพียงพอต่อการสะสมความหวานในลำต้นและอาจกระทบต่อการเจริญเติบโตและการสะสมสารอาหารในอ้อย จึงควรมีแผนการกักเก็บน้ำหรือจัดการให้น้ำเสริมในช่วงปลายฤดู

จากการเปรียบเทียบแนวโน้มปริมาณน้ำฝนเฉลี่ยรายเดือนของจังหวัดขอนแก่น (เส้นสีน้ำเงิน) กับรูปแบบฝนทั้ง 4 กลุ่มใน Figure 7 พบว่า เมื่อพิจารณาเฉพาะช่วงฤดูฝนตั้งแต่เดือนพฤษภาคมถึงตุลาคม รูปแบบที่มีความใกล้เคียงกับค่าเฉลี่ยฝนของจังหวัดขอนแก่นมากที่สุด คือกลุ่ม H.3 โดยลักษณะการเปลี่ยนแปลงของปริมาณน้ำฝนแสดงแนวโน้มที่สอดคล้องกันทั้งในช่วงต้นฤดูฝน (พฤษภาคม) ที่มีปริมาณน้ำฝนใกล้เคียงกัน การลดลงของปริมาณฝนในช่วงมิถุนายน และการเพิ่มขึ้นของปริมาณฝนในช่วงกรกฎาคมถึงสิงหาคมที่มีอัตราการเปลี่ยนแปลงคล้ายคลึงกัน ในขณะที่กลุ่ม H.1 มีปริมาณฝนสูงกว่าค่าเฉลี่ยเป็นอย่างมากในช่วงเดือนสิงหาคมถึงตุลาคม กลุ่ม H.2 มีความแตกต่างอย่างชัดเจนในเดือนกันยายนที่มีปริมาณฝนเพิ่มสูงขึ้นอย่างมาก และกลุ่ม H.4 มีปริมาณฝนต่ำกว่าค่าเฉลี่ยในช่วงต้นฤดูฝนและมีรูปแบบการเพิ่มขึ้นที่แตกต่างจากค่าเฉลี่ย การที่ค่าเฉลี่ยมีความใกล้เคียงกับกลุ่ม H.3 ซึ่งมีลักษณะการกระจายตัวของปริมาณฝนค่อนข้างสม่ำเสมอตลอดฤดูฝนโดยไม่มีช่วงที่ฝนตกหนักผิดปกติ สะท้อนให้เห็นถึงรูปแบบปริมาณฝนที่เกิดขึ้นจริงโดยเฉลี่ยของจังหวัดขอนแก่น

จากผลการศึกษาแบบฝน (Figure 7) และตารางการเพาะปลูกอ้อยในแต่ละกรณี (Table 3) แสดงให้เห็นว่า กรณี C และ D ในทุกรูปแบบฝนมีปริมาณน้ำฝนใน Phase 1 และ Phase 2 ที่เพียงพอต่อความต้องการน้ำของอ้อย แต่จะขาดน้ำปริมาณมากใน Phase 3 และ Phase 4 เหมือนกัน ในทุกรูปแบบฝน การคัดเลือกรูปแบบการเพาะปลูกที่เหมาะสมจึงควรคำนึงถึงความสัมพันธ์ระหว่างการกระจายตัวของปริมาณน้ำฝนและความต้องการน้ำในแต่ละ Phase การเจริญเติบโตของอ้อย เพื่อประสิทธิภาพสูงสุดในการใช้ทรัพยากรน้ำและการได้มาซึ่งผลผลิตที่มีคุณภาพ

Table 3 Cumulative rainfall by growth phase of sugarcane under different rainfall patterns and planting cases

Pattern/Case		Phase1	Phase2	Phase3	Phase4
Water Requirement (mm)		120	630	625	140
H.1	A	66.73	606.76	702.86	16.63
	B	76.45	780.24	469.56	3.09
	C	208.71	832.37	211.81	6.66
	D	151.24	873.22	26.38	15.08
H.2	A	58.91	563.87	649.92	20.01
	B	91.43	694.53	447.84	3.57
	C	174.63	851.27	120.05	4.71
	D	133.12	814.62	28.29	17.86
H.3	A	30.04	663.35	467.98	10.50
	B	106.26	744.92	290.64	2.85
	C	220.18	698.32	119.92	6.93
	D	149.05	655.84	20.27	15.81
H.4	A	31.93	468.36	542.56	8.79
	B	56.34	644.31	319.06	2.90
	C	148.86	704.19	113.21	6.79
	D	124.91	680.81	18.47	14.15

Note: Red = Insufficient rainfall, Green = Sufficient rainfall for water requirements.

4. สรุปผลการศึกษาและข้อเสนอแนะ

การศึกษานี้ได้จำแนกรูปแบบการตกของฝนในจังหวัดขอนแก่น โดยใช้ข้อมูลฝนรายเดือน CHIRPS ในช่วงปี พ.ศ. 2524-2566 ซึ่งข้อมูลมีความน่าเชื่อถือสูงเมื่อเทียบกับข้อมูลจากสถานีตรวจวัดน้ำฝนของกรมอุตุนิยมวิทยา ผลการวิเคราะห์ด้วย Silhouette Analysis และ Davies-Bouldin Index พบว่าวิธี Hierarchical Clustering เหมาะสมที่สุดในการจำแนกรูปแบบฝน โดยสามารถแบ่งออกเป็น 4 กลุ่ม กลุ่ม H.1 มีฝนสูงต่อเนื่องช่วงกลางถึงปลายปี กลุ่ม H.2 มีฝนตกหนักเฉพาะเดือนกันยายน กลุ่ม H.3 มีฝนสูงช่วงต้นฤดูและกระจายสม่ำเสมอ และกลุ่ม H.4 มีฝนน้อยช่วงต้นปีแล้วเพิ่มขึ้นกลางปี โดยกลุ่ม H.3 นั้นมีรูปแบบฝนที่มีความใกล้เคียงกับรูปแบบค่าเฉลี่ยฝนของจังหวัดขอนแก่นมากที่สุด แต่เมื่อวิเคราะห์ผลกระทบต่อการเพาะปลูกอ้อยพบว่ารูปแบบฝน H.4 โอกาสเกิดขึ้น 32.56% มีความเสี่ยงสูงต่อการปลูกอ้อยในช่วงต้นฤดู (Case A) เนื่องจากมีปริมาณฝนไม่เพียงพอในตั้งแต่ Phase 1 จนกระทั่งถึง Phase 4

ผลการศึกษาสามารถนำไปประยุกต์ใช้ในการกำหนดปฏิทินการเพาะปลูกที่เหมาะสม วางแผนการให้น้ำเสริมในช่วงวิกฤต และประเมินผลผลิตตามลักษณะรูปแบบการตกของฝนในแต่ละปี ซึ่งจะช่วยลดความเสียหายจากภัยแล้งและเพิ่มประสิทธิภาพในการผลิตทางการเกษตรในพื้นที่นอกเขตชลประทานได้อย่างยั่งยืน โดยหน่วยงานที่เกี่ยวข้องสามารถใช้ข้อมูลนี้เป็นเครื่องมือในการจัดทำนโยบายและมาตรการเชิงป้องกันสำหรับการบริหารจัดการทรัพยากรน้ำ และการส่งเสริมการใช้เทคโนโลยีการเกษตรที่เหมาะสม เช่น ระบบการให้น้ำแบบหยดหรือการปลูกพืชที่ทนแล้ง ซึ่งจะช่วยลดผลกระทบจากการเปลี่ยนแปลงสภาพภูมิอากาศและสร้างความมั่นคงทางอาหารให้กับเกษตรกรในพื้นที่ได้อย่างยั่งยืน

สำหรับการศึกษาในอนาคต ควรขยายขอบเขตการวิเคราะห์รูปแบบฝนให้ครอบคลุมพืชเศรษฐกิจชนิดอื่นในพื้นที่ โดยเฉพาะมันสำปะหลังซึ่งเป็นพืชเศรษฐกิจสำคัญอีกชนิดหนึ่งของจังหวัดขอนแก่น การวิเคราะห์ความสัมพันธ์ระหว่างรูปแบบการตกของฝนกับวงจรการเจริญเติบโตและความต้องการน้ำของมันสำปะหลังในแต่ละระยะการเติบโต จะช่วยให้เกษตรกรสามารถวางแผนการเพาะปลูกและการจัดการน้ำได้อย่างมีประสิทธิภาพมากขึ้น นอกจากนี้ ควรมีการศึกษาเปรียบเทียบความอ่อนไหวต่อภัยแล้งระหว่างอ้อยและมันสำปะหลัง เพื่อเป็นข้อมูลสำหรับเกษตรกรในการตัดสินใจเลือกชนิดพืชที่เหมาะสมกับลักษณะรูปแบบฝนในแต่ละปี ซึ่งจะช่วยลดความเสี่ยงและเพิ่มความยืดหยุ่นในการรับมือกับความแปรปรวนของสภาพภูมิอากาศในอนาคต

กิตติกรรมประกาศ

ขอขอบคุณ กรมพัฒนาที่ดิน กระทรวงเกษตรและสหกรณ์ที่ให้อาณาเขตข้อมูลการใช้ที่ดิน งานวิจัยนี้ได้รับทุนสนับสนุนจากคณะวิศวกรรมศาสตร์กำแพงแสน มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน อำเภอกำแพงแสน จังหวัดนครปฐม

เอกสารอ้างอิง

- [1] Kuntiyawichai, K. and Wongsasri, S. 2021. Assessment of drought severity and vulnerability in the Lam Phaniang River basin, Thailand, *Water*, 13(19), 2743.
- [2] Prabhakorn, S., Maskey, S., Suryadi, F. X., and de Fraiture, C. 2019. Assessment of drought hazard, exposure,

- vulnerability, and risk for rice cultivation in the Mun River Basin in Thailand, *Natural Hazards*, 97(2), 891–911.
- [3] Kumar, S. and Kumar, S. 2014. Long-term trends and probability of rainfall for crop planning in Bihar, *Journal of Agrometeorology*, 16(1), 146–148.
- [4] Paulo, A., Martins, D., and Pereira, L. S. 2016. Influence of Precipitation Changes on the SPI and Related Drought Severity. An Analysis Using Long-Term Data Series, *Water Resources Management*, 30(15), 5737–5757.
- [5] Guo, H. et al. 2017. Meteorological drought analysis in the Lower Mekong Basin using satellite-based long-term CHIRPS product, *Sustainability*, 9(6), 901.
- [6] Promkhambut, A. and Rambo, A. T. 2017. Multiple cropping after the rice harvest in rainfed rice cropping systems in Khon Kaen Province, Northeast Thailand, *Southeast Asian Studies*, 6(2), 325–338.
- [7] Geleta, C. D. and Deressa, T. A. 2021. Evaluation of Climate Hazards Group InfraRed Precipitation Station (CHIRPS) satellite-based rainfall estimates over Finchaa and Neshe Watersheds, Ethiopia, *Engineering Reports*, 3(6), 1–16.
- [8] de Oliveira-Júnior, J. F. et al. 2021. Rainfall in Brazilian Northeast via in situ data and CHELSA product: mapping, trends, and socio-environmental implications, *Environmental Monitoring and Assessment*, 193(5), 1–19.
- [9] Zhang, Y., Moges, S., and Block, P. 2016. Optimal cluster analysis for objective regionalization of seasonal precipitation in regions of high spatial-temporal variability: Application to Western Ethiopia, *Journal of Climate*, 29(10), 3697–3717.
- [10] Kandali, K., Bennis, L., and Bennis, H. 2021. A New Hybrid Routing Protocol Using a Modified K-Means Clustering Algorithm and Continuous Hopfield Network for VANET, *IEEE Access*, 9, 47169–47183.
- [11] Shahapure, K. R. and Nicholas, C. 2020. Cluster Quality Analysis Using Silhouette Score, in *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, Sydney, NSW, Australia, 2020, 747–748.
- [12] Xiao, J., Lu, J., and Li, X. 2017. Davies Bouldin Index based hierarchical initialization K-means, *Intelligent Data Analysis*, 21(6), 1327–1338.