การปรับแก้ความถูกต้องของข้อมูลฝนใน ERA5 โดยใช้การเรียนรู้เชิงลึกสำหรับการคาดการณ์ปริมาณฝนในลุ่มแม่น้ำชี

Bias-Correction Rainfall Data in ERA5 Using Deep Learning for Rainfall Prediction in the Chi River Basin

สิริจันทร์ คนซื่อ[1] ชาญชัย เพชรพงศ์พันธุ์[2,*] ชัยวัฒน์ เอกวัฒน์พานิชย์[2] ดวงฤดี โฆษิตกิตติวงศ์ ก้องกิจกุล[2] และ วงศ์นรินทร์ คำพอ[2]

[1,2] ภาควิชาวิศวกรรมโยธา คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี จ.กรุงเทพมหานคร

*Corresponding author; E-mail address: chanchai.pet@kmutt.ac.th

บทคัดย่อ

ในการวิเคราะห์และประเมินผลกระทบจากภัยธรรมชาติทางด้านอุตุ-อุทกวิทยา เช่น อุทกภัย และ ภัยแล้ง ด้วยแบบจำลองทางอุทกวิทยาต้องอาศัยข้อมูลที่มีความละเอียดสูง เพื่อให้สามารถจำลองสถานการณ์ที่อาจเกิดขึ้นได้อย่างแม่นยำ แต่ข้อมูลที่มีอยู่ในปัจจุบันยังขาดความละเอียดที่เพียงพอทำให้เกิดข้อจำกัดในการวิเคราะห์และคาดการณ์ งานวิจัยนี้จึงมีเป้าหมายเพื่อประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ในการปรับปรุงความละเอียดของข้อมูลปริมาณน้ำฝนในระดับลุ่มน้ำย่อย โดยการนำตัวแปรทางด้านสภาพภูมิอากาศจาก European Centre for Medium-Range Weather Forecasts (ECMWF) และข้อมูลปริมาณฝนจากสถานีตรวจวัดของกรมอุตุนิยมวิทยาในพื้นที่ลุ่มน้ำชี มาทำการปรับแก้ความถูกต้อง (Bias-Correction) ด้วยแบบจำลองโครงข่ายประสาทแบบเวียนซ้ำ (Recurrent Neural Network หรือ RNN) ซึ่งพัฒนาให้สามารถเรียนรู้และวิเคราะห์ความสัมพันธ์เชิงเวลาและพื้นที่ระหว่างตัวแปรสภาพภูมิอากาศและข้อมูลฝนสังเกตุการณ์ ผลการศึกษาแสดงให้เห็นถึงระดับความสัมพันธ์ระหว่างปัจจัยทางด้านภูมิอากาศที่ใช้ข้อมูลระดับ Global Scale และกับปริมาณน้ำฝนจากการตรวจวัด จากผลการฝึกและทดสอบแบบจำลองในช่วงปี พ.ศ. 2538–2567 พบว่า แบบจำลองที่พัฒนาขึ้นมีความสามารถในการประมาณค่าปริมาณฝนได้อย่างแม่นยำ โดยเฉพาะอย่างยิ่งในช่วงฤดูฝน ซึ่งเป็นช่วงเวลาที่มีความแปรปรวนของสภาพอากาศสูง จากการทดสอบแบบจำลอง R^2 อยู่ในช่วง 0.77–0.93 และมีค่า RMSE อยู่ในช่วง 2.4–5.8 มม. องค์ความรู้นี้สามารถนำไปใช้ในการสนับสนุนการกำหนดนโยบายด้านการจัดการน้ำและการลดความเสี่ยงจากภัยพิบัติในพื้นที่ต่าง ๆ ได้

คำสำคัญ: การคาดการณ์ปริมาณน้ำฝน, การปรับแก้ความถูกต้อง, ปัญญาประดิษฐ์, โครงข่ายประสาทแบบเวียนซ้ำ, ECMWF

Abstract

In analyzing and assessing the impacts of hydro-climatic disasters, such as floods and droughts, hydrological models require high-resolution data to accurately simulate possible scenarios. However, the insufficient data resolution limits analysis and forecasting. Therefore, this study aims to apply the artificial intelligence technology to improve the resolution of rainfall data at the sub-basin scale. The climate variables from the European Centre for Medium-Range Weather Forecasts (ECMWF) and measured rainfall data from the Thai Meteorological Department's monitoring stations in the Chi River Basin would be downscaled by using the Recurrent Neural Network (RNN) model which is developed to learn and analyze the relationship between climate variables and measured rainfall data. Moreover, from the model training and testing results during 1995–2024, it was found that the developed model can accurately estimate rainfall, especially during the rainy season, which is a period of high weather variability. From model testing results indicated the range of R^2 was 0.77 to 0.93 and the RMSE was 2.4 to 5.8 mm. This knowledge can be applied to support water management policy formulation and disaster risk reduction in various areas

Keywords: Rainfall prediction, Bias-Correction, Artificial Intelligence, Recurrent Neural Network, ECMWF

1. บทนำ

การเปลี่ยนแปลงสภาพภูมิอากาศ (Climate Change) คือการเปลี่ยนแปลงในลักษณะเฉลี่ยหรือความแปรปรวนของภูมิอากาศในระดับโลกหรือระดับภูมิภาคเมื่อเวลาผ่านไป ซึ่งอาจเกิดขึ้นจากกระบวนการภายในของระบบภูมิอากาศโลก แรงกระตุ้นจากภายนอก เช่น การเปลี่ยนแปลงของพลังงานแสงอาทิตย์ หรือจากกิจกรรมของมนุษย์โดยตรง โดยเฉพาะในช่วงศตวรรษที่ผ่านมา ปรากฏการณ์ดังกล่าวได้ส่งผลกระทบต่อวงจรน้ำและระบบอุทกวิทยาในหลายพื้นที่ทั่วโลก [1] ในประเทศไทย พบว่าปริมาณฝนในฤดูมรสุมมีแนวโน้มเพิ่มขึ้นในหลายภูมิภาค โดยเฉพาะภาคเหนือและภาคตะวันออกเฉียงเหนือ ส่งผลให้เกิดน้ำท่วมฉับพลันและดินถล่มในพื้นที่เสี่ยงภัย นอกจากนี้ อุณหภูมิที่สูงขึ้นยังทำให้เกิดการระเหยของน้ำมากขึ้น ส่งผลต่อการลดลงของปริมาณน้ำในแหล่งน้ำต้นทุนในบางฤดูกาล ซึ่งอาจกระทบต่อภาคเกษตรกรรมและความมั่นคงด้านน้ำของประเทศ

การประเมินผลกระทบของการเปลี่ยนแปลงสภาพภูมิอากาศต่อกระบวนการอุทกวิทยาในระดับลุ่มน้ำ มักดำเนินการผ่านแบบจำลองสภาพภูมิอากาศโลก (Global Climate Models: GCMs) ซึ่งถูกนำมาใช้ในการคาดการณ์อนาคตภายใต้สถานการณ์การปล่อยก๊าซเรือนกระจกต่าง ๆ อย่างไรก็ตาม GCM มีข้อจำกัดที่สำคัญคือ ความละเอียดเชิงพื้นที่ต่ำ ซึ่งไม่เหมาะสมต่อการใช้งานในระดับพื้นที่เฉพาะ เช่น ลุ่มน้ำขนาดเล็กหรือการวางแผนโครงสร้างพื้นฐานในระดับท้องถิ่น [2] ด้วยเหตุนี้จึงจำเป็นต้องมีการลดขนาดข้อมูล (Downscaling) จาก GCM เพื่อให้ได้ชุดข้อมูลอุตุนิยมวิทยาที่มีความละเอียดสูงเพียงพอ

การลดขนาดข้อมูลสามารถดำเนินการได้สองแนวทางหลัก ได้แก่ การลดขนาดแบบไดนามิก (Dynamical Downscaling) และการลดขนาดเชิงสถิติ (Statistical Downscaling) [3] โดยในระยะหลัง การลดขนาดเชิงสถิติโดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ได้รับความนิยมเพิ่มขึ้นอย่างมาก เนื่องจากสามารถวิเคราะห์ความสัมพันธ์ที่ซับซ้อนระหว่างตัวแปรทางอุตุนิยมวิทยากับปริมาณฝนจากการตรวจวัดจริงได้อย่างมีประสิทธิภาพ [4] อย่างไรก็ตาม แม้การลดขนาดจะช่วยเพิ่มความละเอียดของข้อมูลได้ดีขึ้น แต่ข้อมูลที่ได้จาก GCM ยังคงมีอคติ (bias) อยู่เดิมในหลายกรณี เช่น ค่าปริมาณฝนที่สูงเกินจริงหรือต่ำเกินไปเมื่อเทียบกับข้อมูลตรวจวัด ซึ่งอาจส่งผลให้แบบจำลองที่นำมาใช้งานจริงเกิดความคลาดเคลื่อน

ดังนั้น ขั้นตอนสำคัญที่มักถูกดำเนินการก่อนการนำข้อมูลจำลองมาใช้ในระดับลุ่มน้ำ คือ การปรับแก้ความเอนเอียงของข้อมูล (Bias Correction) ซึ่งสามารถดำเนินการได้หลากหลายวิธี เช่น เทคนิคการปรับเชิงสถิติ (Statistical Bias Correction) ที่ทำให้ข้อมูลจำลองมีลักษณะการแจกแจงใกล้เคียงกับข้อมูลสังเกตการณ์ในอดีต เพื่อให้ผลการวิเคราะห์หรือการพยากรณ์ในขั้นตอนถัดไปมีความแม่นยำและสอดคล้องกับสภาพความเป็นจริงมากยิ่งขึ้น โดยเฉพาะเมื่อนำข้อมูลนั้นไปใช้เป็นข้อมูลนำเข้าสำหรับการพัฒนาแบบจำลองเชิงลึกด้วย Machine Learning [10] ด้วยเหตุนี้ การประยุกต์ใช้เทคนิค Bidirectional Long Short-Term Memory (Bi-LSTM) จึงมีศักยภาพในการลดข้อจำกัดดังกล่าว เนื่องจาก Bi-LSTM สามารถเรียนรู้ลำดับข้อมูลย้อนหลังและปัจจุบันได้พร้อมกัน ทำให้สามารถเรียนรู้การวิเคราะห์ข้อมูลเชิงเวลา (Time Series Analysis) ได้อย่างมีประสิทธิภาพ อีกทั้งยังเหมาะสำหรับลุ่มน้ำที่มีความซับซ้อน เช่น ลุ่มแม่น้ำชี ซึ่งเป็นลุ่มน้ำที่มีลักษณะน้ำท่าในแม่น้ำจะไม่เพิ่มขึ้นทันทีหลังจากเกิดฝนตก แต่จะค่อย ๆ เพิ่มขึ้นในภายหลัง ซึ่งสะท้อนให้เห็นถึงการตอบสนองของลุ่มน้ำที่ค่อนข้างล่าช้า และการที่น้ำฝนส่วนหนึ่งถูกดูดซับและกักเก็บไว้ในดินก่อนจะไหลลงสู่แม่น้ำ ทำให้กระบวนการเกิดน้ำท่ามีความล่าช้าจากเวลาที่ฝนตกจริง ซึ่งสะท้อนถึงพฤติกรรมของฝนที่ซับซ้อนและองค์ประกอบน้ำท่าที่เกิดจากน้ำฝนสะสม [9]

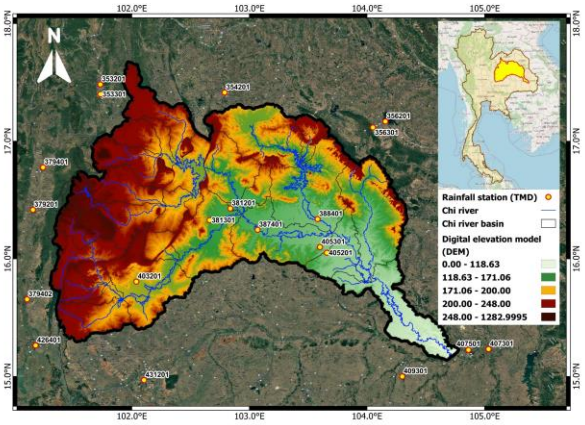
จากประเด็นที่กล่าวมาข้างต้นงานวิจัยนี้จึงมีวัตถุประสงค์หลักคือเพื่อประยุกต์ใช้เทคนิค Machine Learning ในการปรับแก้ความถูกต้องปริมาณฝน โดยใช้แบบจำลองโครงข่ายประสาทเทียมแบบเวียนซ้ำ (Recurrent Neural Network: RNN) ในการประเมินความสัมพันธ์ของข้อมูลทางภูมิอากาศจากศูนย์ ECMWF (European Centre for Medium-Range Weather Forecasts) กับข้อมูลฝนสังเกตการณ์จากสถานีตรวจวัดของกรมอุตุนิยมวิทยาในพื้นที่ลุ่มน้ำชี โดยมีวัตถุประสงค์ย่อยดังต่อไปนี้:

- เพื่อพัฒนาแบบจำลองที่สามารถคาดการณ์ปริมาณฝนได้อย่างแม่นยำ ทั้งในเชิงพื้นที่และเชิงเวลา
- เพื่อเปรียบเทียบปริมาณฝนรายวันที่ผ่านการปรับแก้ข้อมูลด้วยแบบจำลอง RNN กับข้อมูลฝนสังเกตการณ์
- เพื่อเปรียบเทียบค่าความเข้มของฝนเฉลี่ยรายเดือนระหว่างผลจากแบบจำลองกับข้อมูลสังเกตการณ์

ผลลัพธ์จากงานวิจัยนี้มีเป้าหมายเพื่อสนับสนุนการบริหารจัดการทรัพยากรน้ำในพื้นที่ลุ่มน้ำชีให้มีประสิทธิภาพมากยิ่งขึ้นในอนาคต

2. พื้นที่ศึกษา

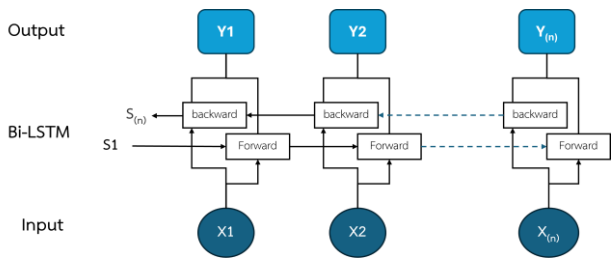
ลุ่มน้ำชีตั้งอยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทย มีพื้นที่รวมประมาณ 49,131.92 ตารางกิโลเมตร ครอบคลุมพื้นที่ใน 14 จังหวัด ได้แก่ ชัยภูมิ ขอนแก่น หนองบัวลำภู อุดรธานี มหาสารคาม นครราชสีมา เลย เพชรบูรณ์ กาฬสินธุ์ ร้อยเอ็ด ยโสธร อุบลราชธานี ศรีสะเกษ และมุกดาหาร พื้นที่ดังกล่าวมีภูมิประเทศที่หลากหลาย ตั้งแต่แนวเทือกเขาดงพญาเย็นทางตะวันตกซึ่งเป็นต้นน้ำของแม่น้ำชี ไปจนถึงพื้นที่ราบและลูกคลื่นลอนในตอนกลางของลุ่มน้ำ ซึ่งมีแม่น้ำชีและลำน้ำสาขาที่สำคัญ เช่น น้ำพรม น้ำพอง น้ำเชิญ ลำปาว และน้ำยัง เป็นแหล่งน้ำหลักของภูมิภาค อย่างไรก็ตาม พื้นที่ลุ่มน้ำชีประสบปัญหาด้านอุทกวิทยาหลายประการ โดยเฉพาะน้ำท่วมในช่วงฤดูฝนและภัยแล้งในฤดูแล้ง อันมีสาเหตุจากความแปรปรวนของปริมาณฝนที่เกิดขึ้นจากการเปลี่ยนแปลงสภาพภูมิอากาศ โดยมีการติดตั้งสถานีวัดน้ำฝนกว่า 80 แห่งในพื้นที่เพื่อประเมินปริมาณฝนในฤดูต่าง ๆ พบว่าฤดูฝนมีค่าเฉลี่ยสูงถึง 179.25 มิลลิเมตรต่อเดือน ขณะที่ฤดูแล้งมีปริมาณฝนลดลงอย่างมาก [5]



รูปที่ 1 พื้นที่ศึกษาลุ่มน้ำชี

3. แบบจำลองกลุ่มโครงข่ายระบบประสาทแบบย้อนกลับ

ในการศึกษานี้ได้เลือกใช้โครงข่ายประสาทเทียมแบบ Bidirectional Long Short-Term Memory (Bidirectional LSTM) ซึ่งเป็นรูปแบบหนึ่งของ LSTM ที่พัฒนาต่อมาจากโครงข่ายประสาทเทียมแบบอนุกรมเวลา (Recurrent Neural Network: RNN) โดยพื้นฐานแล้ว RNN ถูกออกแบบมาเพื่อประมวลผลข้อมูลที่อยู่ในรูปแบบของลำดับเวลา (time series) เช่น ข้อมูลสภาพภูมิอากาศหรือปริมาณฝนด้วยคุณสมบัติของการเชื่อมต่อแบบวนกลับในแต่ละหน่วยทำให้สามารถจดจำบริบทของข้อมูลจากอดีตมาส่งต่อในการทำนายค่าลำดับถัดไปได้ อย่างไรก็ตาม RNN แบบดั้งเดิมมีข้อจำกัดในด้านการเรียนรู้ข้อมูลที่มีระยะห่างกันมากในลำดับเวลาซึ่งมักนำไปสู่ปัญหาการลืมข้อมูลในอดีต ดังนั้นจึงมีการพัฒนา LSTM (Long Short-Term Memory) ขึ้นมาโดยเสริมโครงสร้างด้วยหน่วยความจำภายใน (memory cell) และกลไกควบคุมการไหลของข้อมูล ได้แก่ forget gate, input gate และ output gate ซึ่งช่วยให้โมเดลสามารถเก็บรักษาและเลือกใช้ข้อมูลที่สำคัญได้อย่างมีประสิทธิภาพมากขึ้น เพื่อเพิ่มศักยภาพในการเรียนรู้บริบทของข้อมูลอย่างรอบด้าน งานวิจัยนี้จึงได้นำ Bidirectional LSTM มาใช้ซึ่งเป็นรูปแบบที่ขยายความสามารถของ LSTM ให้สามารถเรียนรู้จากข้อมูลในทั้งอดีต และอนาคต พร้อมกันโดยโมเดลจะมีทิศทางการเรียนรู้สองด้านคือ ด้านหนึ่งเรียนรู้จากลำดับเวลาไปข้างหน้า (forward sequence) และอีกด้านหนึ่งเรียนรู้ย้อนกลับจากอนาคตไปยังอดีต (backward sequence) โดยแสดงกระบวนการทำงานดัง**รูปที่ 2** จึงสามารถเข้าใจความสัมพันธ์ของข้อมูลภูมิอากาศที่เปลี่ยนแปลงอย่างต่อเนื่องและมีความซับซ้อนได้ดียิ่งขึ้น



รูปที่ 2 กระบวนการทำงานของ Bidirectional LSTM (Bi-LSTM)

4. ระเบียบวิธีวิจัย

4.1 รวบรวมข้อมูล

งานวิจัยนี้ใช้ข้อมูลด้านสภาพภูมิอากาศที่เกี่ยวข้องกับลุ่มน้ำชี โดยประกอบด้วยข้อมูลปริมาณฝนย้อนหลัง 30 ปี ตั้งแต่ปี พ.ศ. 2538 – 2567 จากกรมอุตุนิยมวิทยาประเทศไทย ซึ่งเป็นข้อมูลปริมาณฝนรายวัน มีความสำคัญอย่างยิ่งในการศึกษาความผันแปรของปริมาณฝนในระยะยาว นอกจากนี้ยังได้นำข้อมูลจากชุดข้อมูล ERA5 ซึ่งเป็นข้อมูลวิเคราะห์ซ้ำ (reanalysis) รุ่นที่ 5 ของศูนย์พยากรณ์อากาศระยะกลางแห่งยุโรป (European Centre for Medium-Range Weather Forecasts: ECMWF) มาใช้ประกอบการวิเคราะห์ โดยข้อมูล ERA5 เป็นข้อมูลรายชั่วโมงในรูปแบบ กริด (Gridded Data) ซึ่งจัดเรียงตามโครงสร้างพิกัดภูมิศาสตร์แบบละติจูด–ลองจิจูด (Regular latitude-longitude grid) โดยมีความครอบคลุมในระดับโลก (Global coverage) โดยมีความละเอียดในแนวนอนเท่ากับ 0.25 องศา x 0.25 องศา หรือประมาณ 27.75 กิโลเมตร x 27.75 กิโลเมตร ต่อช่องกริด จึงสามารถนำมาใช้ประเมินแนวโน้มของสภาพภูมิอากาศในระดับภูมิภาคได้ สำหรับในการศึกษานี้ ข้อมูล ERA5 ถูกนำมาใช้วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรทางภูมิสภาพอากาศกับปริมาณน้ำฝนตรวจวัดและระดับน้ำในลุ่มน้ำชี โดยเลือกใช้ตัวแปรทางอุตุนิยมวิทยาทั้งหมด 13 ตัวแปร ที่ระดับความกดอากาศ 850 hPa และ 500 hPa **ดังตารางที่ 1** ซึ่งเป็นระดับที่ใช้ในงานวิเคราะห์พลศาสตร์ของบรรยากาศ เนื่องจากสามารถสะท้อนกลไกสำคัญของการเกิดฝนในเขตร้อนได้อย่างชัดเจน โดยระดับ 850 hPa (ประมาณ 1,500 เมตรจากผิวโลก) เป็นชั้นอากาศต่ำที่แสดงถึงลมมรสุม ความชื้น และการยกตัวของอากาศ ส่วนระดับ 500 hPa (ประมาณ 5,500 เมตร) เป็นชั้นบรรยากาศระดับกลางที่ใช้วิเคราะห์การเคลื่อนที่ของคลื่นบรรยากาศขนาดใหญ่ ซึ่งส่งผลโดยตรงต่อการก่อตัวของฝน [6]

ตารางที่ 1 ตัวแปรทางสภาพภูมิอากาศ ERA5 (ECMWF 2568)

พารามิเตอร์	ตัวย่อ	ชื่อภาษาไทย	หน่วย
Specific humidity	q	ความชื้นจำเพาะ	kg kg-1
Relative humidity	r	ความชื้นสัมพัทธ์	%
U-component of wind	u	องค์ประกอบลมทิศตะวันตก-ออก	m s-1
V-component of wind	v	องค์ประกอบลมทิศเหนือ-ใต้	m s-1
Vertical velocity	w	ความเร็วแนวดิ่ง	Pa s-1
Divergence	d	การแผ่กระจายของลม	s-1
Vorticity (relative)	vo	การหมุนวนสัมพัทธ์	s-1
Potential vorticity	pv	ศักย์การหมุนวน	K m2 kg-1 s-1
Geopotential	z	ศักย์จีโอพอเทนเชียล	m2 s-2
Fraction of cloud cover	tcc	สัดส่วนการปกคลุมของเมฆ	Dimensionless
Specific cloud liquid water content	clwc	ปริมาณน้ำของเหลวในเมฆเฉพาะมวล	kg kg-1
Specific rain water content	crwc	ปริมาณน้ำฝนเฉพาะมวล	kg kg-1
Ozone mass mixing ratio	O3	สัดส่วนผสมของมวลโอโซน	kg kg-1

4.2 พัฒนาแบบจำลองเพื่อปรับแก้ความถูกต้องฝน

การพัฒนาแบบจำลองเพื่อปรับแก้ความถูกต้องฝนประกอบด้วยกระบวนการหลัก 5 ขั้นตอน โดยเริ่มจาก การเตรียมข้อมูล ซึ่งเป็นขั้นตอนสำคัญที่ใช้ในการประมวลผลและจัดรูปแบบข้อมูลจากแหล่งต่าง ๆ ให้สามารถนำมาใช้ร่วมกันได้อย่างถูกต้องและสอดคล้อง โดยในขั้นตอนนี้ได้นำข้อมูลปริมาณฝนจากสถานีวัดฝนรายวันของกรมอุตุนิยมวิทยา มาทำการเชื่อมโยงกับชุดข้อมูล ERA5 ของ ECMWF ซึ่งเป็นข้อมูลรายชั่วโมง จึงได้นำมาคิดเป็นปริมาณฝนรายวันโดยพิจารณาจากปริมาณสะสมในแต่ละวัน เพื่อให้สอดคล้องกับความละเอียดเวลาของข้อมูลฝนที่สังเกตการณ์ หลังจากนั้นได้ทำการ ปรับช่วงเวลาและจับคู่ตำแหน่งเชิงพื้นที่ โดยเลือกจุดกริดของข้อมูล ERA5 ที่ใกล้เคียงกับพิกัดของสถานีวัดฝนมากที่สุด เพื่อให้ข้อมูลจากทั้งสองแหล่งอยู่ในรูปแบบที่ตรงกันทั้งในเชิงเวลาและเชิงพื้นที่ อีกทั้งยังมีการคัดกรองเฉพาะช่วงเวลาที่ข้อมูลมีความครบถ้วนและสมบูรณ์เท่านั้น เพื่อนำมาใช้ในการวิเคราะห์และฝึกแบบจำลองต่อไป

ในขั้นตอนที่สองเป็นขั้นตอนการสร้างฟีเจอร์และการปรับสเกลข้อมูล โดยในขั้นตอนนี้ได้มีการสร้างตัวแปรทางอนุกรมเวลาจากข้อมูล ERA5 เพื่อเสริมศักยภาพในการเรียนรู้ของแบบจำลอง เช่น ตัวแปรค่าล่าช้า (lag features) ที่สะท้อนผลกระทบของค่าฝนในอดีต ค่าทางสถิติเชิงเวลา ได้แก่ ค่าเฉลี่ยย้อนหลัง (moving average) และค่าเฉลี่ยถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียล (exponential moving average) ซึ่งช่วยให้แบบจำลองสามารถเรียนรู้แนวโน้มในช่วงสั้นและช่วงยาว นอกจากนี้ยังได้มีการปรับขนาดข้อมูลโดยใช้ฟังก์ชั่น StandardScaler เพื่อลดความเบี่ยงเบนของค่าตัวแปรให้อยู่ในสเกลที่เหมาะสมต่อการเรียนรู้ของโครงข่ายประสาทเทียม

โครงสร้างของแบบจำลองในงานวิจัยนี้เลือกใช้ Bidirectional LSTM (Bi-LSTM) เพื่อให้แบบจำลองสามารถเรียนรู้ลำดับข้อมูลจากทั้งสองทิศทาง คือจากอดีตไปสู่ปัจจุบัน (forward pass) และจากอนาคตย้อนกลับไปยังอดีต (backward pass) ในลำดับเวลาเดียวกัน ซึ่งช่วยเพิ่มศักยภาพในการประเมินบริบทและความสัมพันธ์ที่ซับซ้อนในข้อมูลเชิงเวลาได้ดียิ่งขึ้น โดยเฉพาะในกรณีที่ปัจจัยทางภูมิอากาศมีการเปลี่ยนแปลงแบบ nonlinear และ nonstationary ซึ่งเป็นการแปลี่ยนแปลงที่ขึ้นอยู่กับลำดับก่อนหน้าและหลังจากนั้นแบบไม่เป็นเชิงเส้น

สำหรับขั้นตอนการฝึกแบบจำลอง ได้ดำเนินการจัดเตรียมชุดข้อมูลโดยแบ่งข้อมูลออกเป็นสองส่วน ได้แก่ ชุดข้อมูลสำหรับฝึกตั้งแต่ปี พ.ศ. 2538-2561 และ ชุดข้อมูลสำหรับทดสอบตั้งแต่ปี พ.ศ. 2562-2567 ในอัตราส่วน 80:20 โดยการแบ่งข้อมูลในลักษณะนี้เป็นการแบ่งตามลำดับเวลา (time-based split) ซึ่งเหมาะกับข้อมูลแบบ time series เนื่องจากสามารถรักษาลำดับเหตุการณ์และโครงสร้างความสัมพันธ์เชิงเวลา (temporal dependency) ที่มีอยู่ในข้อมูลเดิมได้ [11] ต่อมาข้อมูลจากชุด ERA5 ได้ถูกนำมาใช้เป็น ข้อมูลฟีเจอร์ (Features) ขณะที่ข้อมูลปริมาณฝนจากสถานีตรวจวัดถูกกำหนดให้เป็น ข้อมูลเป้าหมาย (Target) สำหรับการเรียนรู้ของโมเดลได้มีการนำเทคนิค EarlyStopping มาใช้ในการฝึกโมเดล ซึ่งเป็นเทคนิคที่ใช้เพื่อหยุดการฝึกแบบจำลองก่อนที่จะเกิดปัญหา Overfitting หรือการที่แบบจำลองเรียนรู้ข้อมูลฝึกมากเกินไปจนสูญเสียความสามารถในการทำนายข้อมูลใหม่ กล่าวคือแบบจำลองจะมีความแม่นยำสูงกับข้อมูลฝึก แต่มีความแม่นยำน้อยเมื่อนำไปใช้กับข้อมูลทดสอบ [7] โดยกำหนดเงื่อนไขให้หยุดการฝึกเมื่อค่าความคลาดเคลื่อน (Validation loss) ไม่มีการลดลงติดต่อกันเป็นเวลา 15 epoch ซึ่งช่วยป้องกันการฝึกโมเดลมากเกินไป และลดระยะเวลาในการประมวลผล ซึ่งการปรับแก้ค่าความคาดเคลื่อน เพื่อให้แบบจำลองมีความเหมาะสมกับข้อมูลนำเข้า และช่วยเพิ่มประสิทธภาพในการลดขนาดฝนในพื้นที่ลุ่มย่อย โดยรายละเอียดของวิธีการปรับแก้ความคลาดเคลื่อนแสดงใน**ตารางที่ 2**

ตารางที่ 2 ชั้นประมวลผลในการปรับแก้ความคาดเคลื่อนของแบบจำลอง RNN

ชั้นประมวลผล	หน่วย	ฟังก์ชันกระตุ้น
Bidirectional LSTM	128	tanh
Bidirectional LSTM	64	tanh
Bidirectional LSTM	32	tanh
Fully Connected (Dense)	64	ReLU
Fully Connected (Dense)	32	ReLU
Output Layer	1	Linear

ขั้นตอนสุดท้ายเป็นกระบวนการทดสอบแบบจำลอง โดยทำการคาดการณ์ปริมาณน้ำฝนรายวันจากชุดข้อมูลทดสอบ โดยใช้แบบจำลองที่ผ่านการฝึกและปรับแก้ความคลาดเคลื่อนแล้ว การทดสอบแบบจำลองครอบคลุมช่วงระยะเวลาตั้งแต่ปี พ.ศ. 2562 ถึง 2567 โดยผลที่ได้ถูกนำมาเปรียบเทียบกับข้อมูลปริมาณฝนจากสถานีตรวจวัด จำนวน 23 แห่ง ซึ่งตำแหน่งของแต่ละสถานีแสดงไว้ใน**รูปที่ 1**

สำหรับสถานีฝนทั้ง 23 แห่งที่ใช้ในการเปรียบเทียบนี้ ได้มาจากการคัดเลือกสถานีตรวจวัดปริมาณน้ำฝนที่ตั้งอยู่ในพื้นที่ที่ครอบคลุมลุ่มแม่น้ำชี ตั้งแต่บริเวณตอนบนจนถึงตอนล่างของลุ่มน้ำ เพื่อให้การประเมินแบบจำลองสามารถสะท้อนความหลากหลายของลักษณะการเกิดฝนในแต่ละพื้นที่ได้อย่างครอบคลุม โดยพิจารณาทั้งในแง่ของการกระจายเชิงพื้นที่และความสมบูรณ์ของข้อมูลย้อนหลังในช่วงเวลาที่ศึกษา

4.3 วิเคราะห์ผลการคาดการณ์ปริมาณฝน

ในการวิเคราะห์ผลการคาดการณ์ปริมาณฝน ได้ประเมินประสิทธิภาพของแบบจำลองด้วยการใช้ดัชนีชี้วัดสามรายการ ได้แก่ Root Mean Square Error (RMSE) และ Coefficient of Determination (R^2) สุดท้ายคือค่า Nash-Sutcliffe Efficiency (NSE) ซึ่งเป็นดัชนีที่ใช้ในการวิเคราะห์และประเมินผลความแม่นยำของแบบจำลองในด้านอุทกวิทยา โดยเฉพาะในบริบทของการจำลองและพยากรณ์ปริมาณฝนที่มีลักษณะเป็นข้อมูลอนุกรมเวลา [8] ดังแสดงในสมการ 1 2 และ 3 ตามลำดับ

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\left\{\frac{1}{n}\sum_{\{i=1\}}^{\{n\}}(Pi - Oi)^2\right\}} \quad (1)$$

Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\left\{\sum_{\{i=1\}}^{\{n\}}(O_i - P_i)^2\right\}}{\left\{\sum_{\{i=1\}}^{\{n\}}(O_i - \sigma)^2\right\}} \quad (2)$$

Nash–Sutcliffe Efficiency (NSE):

$$NSE = 1 - \frac{\left\{\sum_{\{i=1\}}^{\{n\}}(O_i - P_i)^2\right\}}{\left\{\sum_{\{i=1\}}^{\{n\}}(O_i - \bar{O})^2\right\}} \quad (3)$$

โดยที่

P_i = ค่าที่แบบจำลองคาดการณ์

O_i = ค่าที่สังเกตการณ์

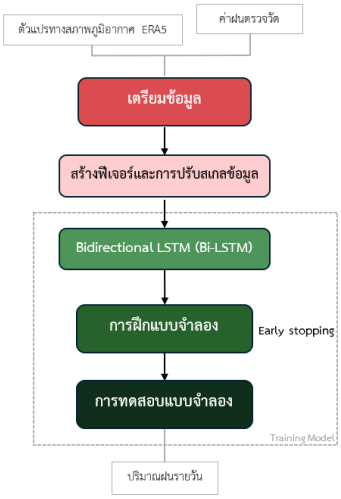
n = จำนวนข้อมูลทั้งหมด

$\bar{O}$ = ค่าเฉลี่ยของข้อมูลสังเกตการณ์

ตารางที่ 3 เกณฑ์การพิจารณาดัชนีวัดประสิทธิภาพ

ดัชนี	ค่าชี้วัด	การแปลผล
RMSE	ค่าใด ๆ > 0 (ค่ายิ่งต่ำดียิ่งดี)	ค่าใกล้ 0 หมายถึงโมเดลมีความคลาดเคลื่อนต่ำ
R^2	0.0 – 1.00	มากกว่า 0.9 หมายถึงดีมาก
NSE	-∞ ถึง 1.00	มากกว่า 0.9 หมายถึงดีมาก

ขั้นตอนการพัฒนาแบบจำลองเพื่อใช้ในการปรับแก้ความถูกต้องฝนสามารถแสดงได้ดัง**รูปที่ 3**



รูปที่ 3 แผนภาพแสดงกระบวนการทำงานของแบบจำลอง

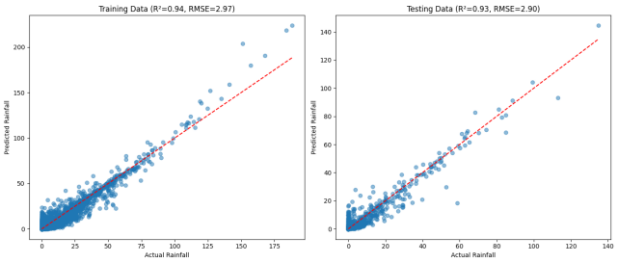
5. ผลการศึกษา

ผลการประเมินประสิทธิภาพทั้งในช่วงที่ทำการฝึกและทดสอบแบบจำลองเพื่อปรับแก้ความถูกต้องฝนในพื้นที่ลุ่มน้ำชีแสดงดัง**ตารางที่ 4** และ **รูปที่ 4** โดยจากการเปรียบเทียบผลการคาดการณ์ปริมาณฝนรายวันที่ได้จากแบบจำลองกับข้อมูลฝนตรวจวัด พบว่า ในช่วงที่ทำการฝึกแบบจำลอง ตั้งแต่ปี พ.ศ. 2538-2561 ค่า R^2 และ NSE มีค่าตั้งแต่ 0.69 ถึง 0.94 ซึ่งแสดงให้เห็นว่าแบบจำลองสามารถคาดการณ์ปริมาณฝนรายวันได้ในระดับดีถึงดีมากในหลายสถานี โดยเฉพาะสถานีที่มีค่า R^2 เท่ากับ 0.94 ได้แก่สถานี 353201, 353301, 379401, 388401 และ 405301 ซึ่งล้วนมีค่า RMSE ต่ำกว่า 3 มิลลิเมตร แสดงถึงความแม่นยำของแบบจำลองที่สูง แต่ในบางสถานี เช่น 409301, 431401 และ 407301 มีค่า R^2 ต่ำกว่า 0.75 และมีค่า RMSE อยู่ในช่วง 5–7 มิลลิเมตร ซึ่งสะท้อนถึงข้อจำกัดของแบบจำลองในการเรียนรู้รูปแบบการเกิดฝนในบางกรณีที่มีการตกหนักอย่างฉับพลันหรือไม่สม่ำเสมอ

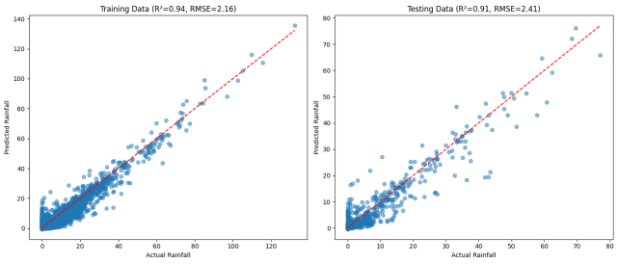
ในช่วงการทดสอบแบบจำลอง พบว่ามีค่า R^2 อยู่ในช่วง 0.77 ถึง 0.93 และ RMSE อยู่ระหว่าง 2.41 ถึง 5.82 มิลลิเมตร โดยสถานีที่แสดงให้เห็นว่าแบบจำลองมีประสิทธิภาพในการคาดการณ์ปริมาณฝนได้เป็นอย่างดี ได้แก่ สถานี 379401 (R^2= 0.91, RMSE= 2.41 มม.), 388401 (R^2= 0.93, RMSE= 2.90 มม.) และ 387401 (R^2= 0.93, RMSE= 2.88 มม.) ซึ่งสถานีเหล่านี้เป็นตำแหน่งที่แบบจำลองมีผลการคาดการณ์อยู่ในระดับดีทั้งในช่วงของการฝึกและทดสอบ เป็นการแสดงให้เห็นว่าแบบจำลองสามารถประเมินความสัมพันธ์เชิงเวลาระหว่างตัวแปรทางสภาพภูมิอากาศกับปริมาณฝนจากการตรวจวัดได้อย่างมีประสิทธิภาพ

ตารางที่ 4 ค่าดัชนีที่ได้จากการเปรียบเทียบปริมาณฝนรายวันจากแบบจำลองกับข้อมูลสังเกตการณ์

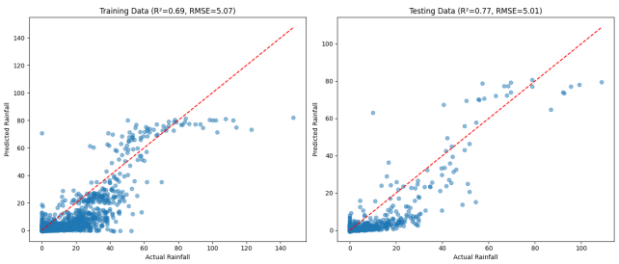
Station Rainfall	Train			Test		
	R^2	RMSE (mm)	NSE	R^2	RMSE (mm)	NSE
353201	0.94	2.60	0.94	0.93	3.07	0.93
353301	0.94	2.36	0.94	0.91	3.20	0.91
356201	0.93	3.41	0.93	0.92	3.05	0.92
379201	0.93	3.34	0.93	0.92	3.02	0.92
379401	0.94	2.16	0.94	0.91	2.41	0.91
379402	0.86	3.86	0.86	0.85	4.95	0.85
381201	0.84	4.27	0.84	0.83	4.71	0.83
381301	0.81	4.59	0.81	0.80	4.88	0.80
387401	0.83	4.77	0.83	0.93	2.88	0.93
388401	0.94	2.97	0.94	0.93	2.90	0.93
403201	0.77	4.87	0.77	0.82	4.59	0.82
405201	0.77	5.72	0.77	0.92	3.14	0.92
405301	0.94	2.95	0.94	0.93	3.19	0.93
409301	0.69	7.11	0.69	0.81	5.82	0.81
426201	0.81	4.32	0.81	0.93	2.80	0.93
426401	0.80	4.43	0.80	0.80	5.11	0.80
431201	0.83	4.03	0.83	0.80	4.84	0.80
431301	0.89	3.10	0.89	0.86	3.10	0.86
431401	0.69	5.07	0.69	0.77	5.01	0.77
354201	0.87	4.17	0.87	0.90	3.31	0.90
356301	0.86	4.47	0.86	0.90	3.81	0.90
407301	0.75	6.31	0.75	0.88	4.78	0.88
407501	0.92	3.70	0.92	0.91	4.64	0.91



(ก) สถานีฝน 379401 รายวัน



(ข) สถานีฝน 388401 รายวัน

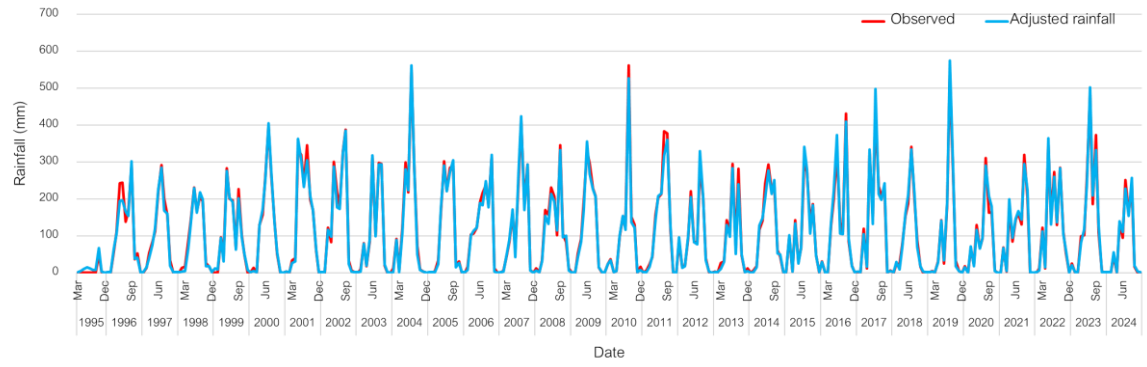


(ค) สถานีฝน 431401 รายวัน

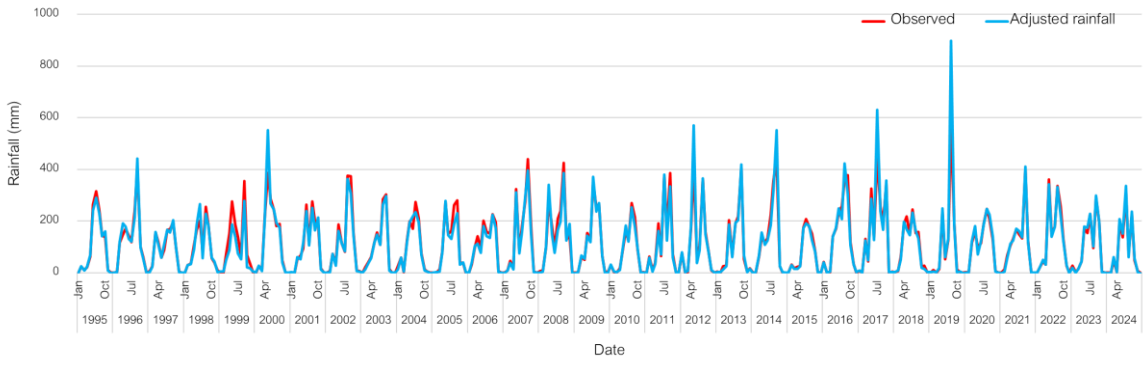
รูปที่ 4 ความสัมพันธ์ระหว่างค่าฝนที่ได้จากการปรับแก้ความถูกต้องด้วย Bi-LSTM และข้อมูลสังเกตการณ์ ภาพด้านซ้ายคือช่วงฝึก ด้านขวาคือช่วงทดสอบ

จากการเปรียบเทียบค่าความเข้มฝนรายเดือนในแต่ละปี **(รูปที่ 5)** และค่าความเข้มฝนเฉลี่ยรายเดือน **(รูปที่ 6)** ในระหว่างปี พ.ศ. 2538-2567 เพื่อประเมินความสอดคล้องของผลที่ได้จากการปรับลดขนาดด้วยแบบจำลองกับข้อมูลสังเกตุการณ์ในเชิงปริมาณและความแปรปรวนตามฤดูกาล โดยพิจารณาที่ตำแหน่งสถานีซึ่งมีค่าดัชนีชี้วัดประสิทธิภาพแบบจำลองที่ดีที่สุด 4 แห่ง ได้แก่ สถานี 388401, 387401, 426201 และ 379401 พบว่า ที่สถานี 388401 มีความสอดคล้องกันอย่างมีนัยสำคัญระหว่างค่าฝนที่ได้จากแบบจำลองกับข้อมูลที่ทำการตรวจวัด ทั้งในด้านการเปลี่ยนแปลงตามฤดูกาลและปริมาณ โดยเฉพาะในช่วงฤดูฝนตั้งแต่เดือนเดือนพฤษภาคมถึงตุลาคม และยังสามารถคาดการณ์ปริมาณฝนสูงสุดได้ใกล้เคียงกับค่าที่เกิดขึ้นจริง ขณะที่ในช่วงฤดูแล้งระหว่างเดือนพฤศจิกายนถึงเมษายนซึ่งเป็นช่วงเวลาที่มีปริมาณฝนน้อยก็สามารถคาดการณ์ได้อย่างแม่นยำ แม้กระทั่งในเดือนที่ฝนมีการขาดช่วง ที่สถานี 387401 มีการแสดงแนวโน้มการเกิดฝนในแต่ละเดือนระหว่างช่วงฤดูฝนที่มีความสม่ำเสมอและสอดคล้องกับปริมาณที่ได้จากการตรวจวัดที่สถานีเช่นกัน รวมถึงมีการคาดการณ์ช่วงที่มีปริมาณฝนสูงสุดได้อย่างแม่นยำในหลายเหตุการณ์ ถึงแม้ว่าจะมีบางช่วง เช่น ในปี พ.ศ. 2553 – 2555 ที่ผลการจำลองที่ได้มีค่าสูงกว่าปริมาณจากข้อมูลตรวจวัดเล็กน้อย ประมาณ 2 มม. ส่วนที่สถานี 426201 มีแนวโน้มของการเกิดฝนในแต่ละเดือนที่ค่อนข้างคงที่กว่าที่ตำแหน่งสถานี 388401 และ 387401 โดยมีการเกิดฝนสูงสุดที่โดนเด่นเป็นบางช่วง เช่น ในปี พ.ศ. 2550 ซึ่งแบบจำลองสามารถคาดการณ์ค่าฝนที่สูงผิดปกติได้อย่างใกล้เคียง สะท้อนให้เห็นถึงความสามารถของแบบจำลองที่พัฒนาขึ้นในการตอบสนองต่อเหตุการณ์สุดขั้ว (extreme events) นอกจากนี้ปริมาณฝนรายเดือนที่ได้ยังมีแนวโน้มของการแปรปรวนตามฤดูกาลที่สอดคล้องกับข้อมูลสังเกตุการณ์อย่างมีนัยยะสำคัญ แม้จะมีการประเมินค่าที่สูงเกินความเป็นจริง (overestimation) ในบางช่วง เช่น กลางปีของปี พ.ศ. 2559-2561 สำหรับที่ตำแหน่งสถานี 379401 นั้น แบบจำลองมีผลการคาดการณ์ที่ใกล้เคียงกับค่าที่เกิดขึ้นจริงเป็นอย่างมาก โดยเฉพาะช่วงที่มีปริมาณฝนสุงสุดในแต่ละปี เช่น ในปี พ.ศ. 2549, 2553, และ 2565 ซึ่งสามารถประเมินปริมาณและช่วงเวลาของการเกิดฝนได้อย่างแม่นยำทั้งในช่วงฤดูฝนและฤดูแล้ง แสดงให้เห็นถึงประสิทธิภาพในการเรียนรู้วัฏจักรของการเกิดฝนตามฤดูกาลได้เป็นอย่างดี ผลจากทั้ง 4 สถานีได้แสดงการเพิ่มขึ้นของปริมาณฝนอย่างต่อเนื่องเริ่มตั้งแต่เดือนพฤษภาคมและสูงสุดในช่วงเดือนกรกฎาคมถึงกันยายนจากนั้นลดลงอย่างรวดเร็วในเดือนตุลาคม ซึ่งสอดคล้องกับสภาพความเป็นจริงที่ได้จากการตรวจวัดที่สถานี โดยเฉพาะในช่วงเดือนที่มีปริมาณความเข้มฝนสูง (กรกฎาคม–กันยายน) ที่ผลการจำลองมีความใกล้เคียงกับฝนที่ตรวจวัดได้จริงเกือบทุกสถานี รวมถึงในช่วงที่ฝนมีการทิ้งช่วงหรือมีปริมาณน้อยมาก (ธันวาคม–มีนาคม) บ่งชี้ให้เห็นว่าไม่มีการ overestimate ปริมาณฝนในช่วงฤดูแล้งอย่างมีนัยยะสำคัญผลการศึกษา

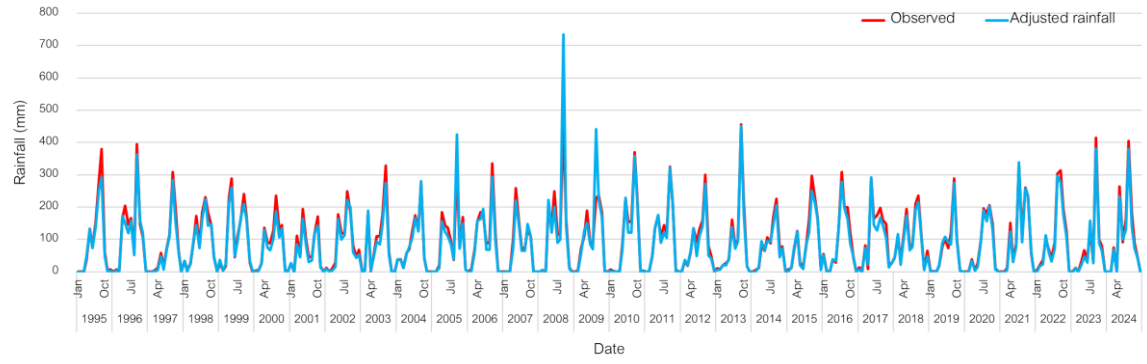
ผลการเปรียบเทียบปริมาณฝนในหลายบริบทและหลายช่วงความถี่เวลาได้แสดงให้เห็นถึงแนวทางการประยุกต์ใช้แบบจำลอง Bi-LSTM ในการปรับลดขนาดของข้อมูลฝนระดับ Global scale เพื่อคาดการณ์การเกิดฝนในระดับลุ่มน้ำได้อย่างมีประสิทธิภาพ โดยแบบจำลองที่พัฒนาขึ้นในงานวิจัยนี้สามารถทำการประเมินปริมาณฝนที่เกิดขึ้นทั้งในระยะสั้น ระยะยาว แนวโน้มการเปลี่ยนแปลงตามฤดูกาล ปริมาณสะสมรายปี ค่าฝนสูงสุดรายเดือนได้เป็นอย่างดี ไม่ว่าจะเป็นปีที่ปริมาณน้ำฝนมากหรือน้อย รวมถึงในปีที่มีความแปรปรวนของปริมาณฝนสูง ซึ่งจะเป็นเครื่องมือที่สำคัญในการประเมินการเกิดและปริมาณฝน ที่ถือเป็นข้อมูลที่มีความจำเป็นสำหรับการวางแผนบริหารจัดการน้ำ การคาดการณ์อุทกภัยและภัยแล้ง ความเพียงพอของน้ำในช่วงฤดูกาลเพาะปลูก โดยเฉพาะในบริบทของพื้นที่ลุ่มน้ำชีซึ่งมีปริมาณฝนเปลี่ยนแปลงตามฤดูกาลอย่างชัดเจน



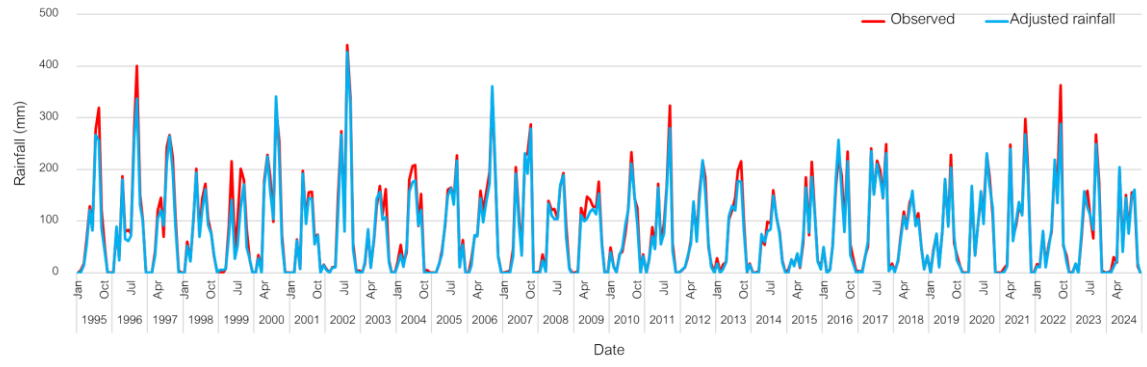
(ก) สถานีฝน 388401 รายเดือน



(ข) สถานีฝน 387401 รายเดือน

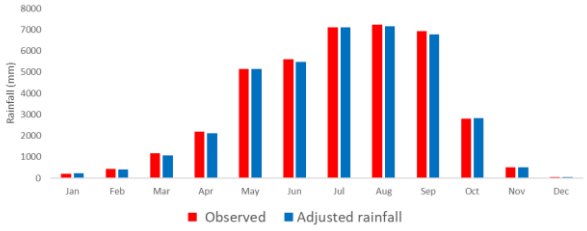


(ค) สถานีฝน 426201 รายเดือน

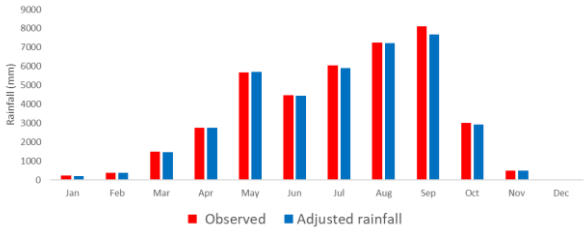


(ง) สถานีฝน 379401 รายเดือน

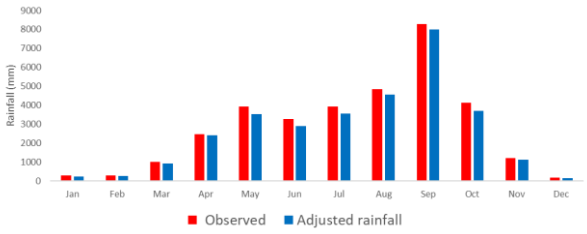
รูปที่ 5 ผลการเปรียบเทียบปริมาณฝนรายเดือนในแต่ละปีที่ได้จากการจำลองและข้อมูลสังเกตการณ์



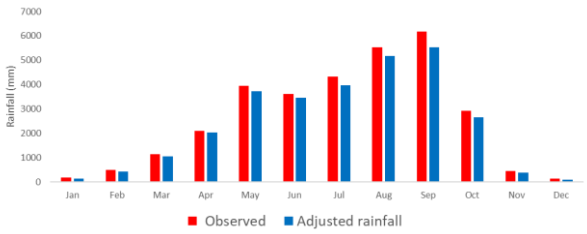
(ก) สถานีฝน 388401 เฉลี่ยรายเดือน



(ข) สถานีฝน 388401 เฉลี่ยรายเดือน



(ค) สถานีฝน 426201 เฉลี่ยรายเดือน



(ง) สถานีฝน 379401 เฉลี่ยรายเดือน

รูปที่ 6 ผลการเปรียบเทียบปริมาณฝนรายเฉลี่ยเดือนที่ได้จากการจำลองและข้อมูลสังเกตการณ์

แบบจำลองที่พัฒนาขึ้นในการศึกษานี้มีลักษณะเฉพาะที่แตกต่างจากงานวิจัยที่เกี่ยวข้อง โดยเฉพาะในด้านโครงสร้างของแบบจำลองสำหรับการลดขนาดข้อมูลฝน ซึ่งที่ผ่านมามักนิยมใช้โมเดลประเภท RNN, LSTM หรือ ANN ร่วมกับข้อมูลจาก GCM และ NCEP reanalysis เพื่อเรียนรู้ความสัมพันธ์ระหว่างตัวแปรเชิงภูมิอากาศกับข้อมูลฝนในอดีต เช่นเดียวกับงานวิจัยของ {4} ที่ใช้ LSTM ในการ downscale อุณหภูมิและปริมาณฝนทั่วประเทศจีน อย่างไรก็ตาม งานวิจัยนี้ได้เลือกใช้ Bi-LSTM ซึ่งเป็นอีกหนึ่งความสามารถของ LSTM ที่สามารถเรียนรู้ข้อมูลได้ทั้งในทิศทางจากอดีตสู่ปัจจุบันและจากปัจจุบันย้อนกลับสู่อดีตพร้อมกัน ส่งผลให้แบบจำลองสามารถเข้าใจลักษณะเชิงเวลาที่มีความไม่เป็นเชิงเส้น ได้ดียิ่งขึ้น ซึ่งเป็นประโยชน์อย่างมากในพื้นที่ที่มีพฤติกรรมฝนซับซ้อน เช่น ลุ่มน้ำชี ที่มีการตอบสนองของน้ำท่าล่าช้าต่อเหตุการณ์ฝน นอกจากนี้ยังได้มีการพัฒนาและกำหนดกระบวนการสร้างฟีเจอร์เชิงเวลา (feature engineering) เพิ่มเติม เช่น การสร้างค่าล่าช้า (lag), ค่าเฉลี่ยเคลื่อนที่ (moving average) และค่าเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก (exponential moving average: EMA) เพื่อช่วยให้แบบจำลองสามารถจับลักษณะฝนสะสมและฤดูกาลได้อย่างมีประสิทธิภาพ

6. สรุปผล

งานวิจัยนี้ได้ทำการพัฒนาแบบจำลองสำหรับการปรับปรับแก้ความถูกต้องฝนเพื่อใช้ในการคาดการณ์การเกิดฝนในระดับลุ่มน้ำย่อย ซึ่งเป็นการประยุกต์ใช้เทคโนโลยีการเรียนรู้ของเครื่องผ่านโครงข่ายประสาทเทียมแบบเวียนซ้ำ (Recurrent Neural Network: RNN) เพื่อเพิ่มความละเอียดของข้อมูลฝนจากระดับ Global scale ให้สามารถใช้งานในระดับภูมิภาคได้อย่างมีประสิทธิภาพ โดยมีพื้นที่ศึกษาเป็นพื้นที่ลุ่มน้ำชี ข้อมูลที่นำเข้าประกอบด้วยชุดตัวแปรด้านสภาพภูมิอากาศจาก ERA5 ของศูนย์ ECMWF และ ข้อมูลฝนจากสถานีตรวจวัดของกรมอุตุนิยมวิทยา ทั้งนี้ชุดข้อมูล ERA5 ได้เลือกใช้เป็นตัวแปรสภาพภูมิอากาศที่ระดับความกดอากาศ 500 และ 850 hPa จำนวน 13 ตัวแปร สำหรับการกำหนดลักษณะสภาพภูมิอากาศที่มีความเกี่ยวข้องกันการเกิดฝนในพื้นที่ศึกษา แบบจำลองได้รับการฝึกและทดสอบในช่วงปี พ.ศ. 2538–2567 และประเมินประสิทธิภาพด้วยดัชนี R^2 NSE และ RMSE

ผลการศึกษาแสดงให้เห็นว่า มีค่า R^2 ในช่วงที่ทำการทดสอบแบบจำลองอยู่ที่ 0.77–0.93 และมีค่า RMSE อยู่ในช่วง 2.4–5.8 มม. ผลการประเมินในช่วงฤดูฝนมีความใกล้เคียงกับฝนที่เกิดขึ้นจริงทั้งในเชิงปริมาณและการเปลี่ยนแปลงตามฤดูกาล โดยมีการคาดการณ์ที่สูงกว่าความเป็นจริงเล็กน้อยในบางช่วงที่มีค่าฝนสูงสุด รวมถึงคาดการณ์ได้ต่ำกว่าค่าจากการตรวจวัดในช่วงปลายฤดูร้อนจนถึงต้นฤดูฝน (เมษายน–พฤษภาคม) นอกจากนี้แบบจำลองยังสามารถประเมินค่าฝนในช่วงที่มีปริมาณฝนต่ำมากของฤดูแล้ง (ธันวาคม–มีนาคม) ได้ค่อนข้างแม่นยำ

แบบจำลอง RNN ที่พัฒนาขึ้นในงานวิจัยนี้สามารถปรับแก้ความละเอียดของข้อมูลฝนระดับ Global scale ให้มีความแม่นยำและสอดคล้องกับข้อมูลสังเกตุการณ์ในระดับลุ่มน้ำย่อยได้อย่างมีประสิทธิภาพ องค์ความรู้นี้สามารถนำไปประยุกต์ใช้เพื่อสนับสนุนการวางแผนด้านการจัดการทรัพยากรน้ำ การเตือนภัยล่วงหน้า และการลดความเสี่ยงจากภัยพิบัติในพื้นที่ที่มีข้อจำกัดด้านข้อมูลฝนสังเกตได้ในอนาคต

ในการศึกษานี้ได้ใช้ข้อมูลจากชุดวิเคราะห์ซ้ำ (Reanalysis) ซึ่งถึงแม้จะมีความแม่นยำสูงและครอบคลุมพื้นที่ทั่วโลก แต่ยังคงเป็นข้อมูลที่มีการอ้างอิงจากค่าในอดีต จึงอาจมีข้อจำกัดในการนำไปประยุกต์ใช้เพื่อคาดการณ์แนวโน้มและปริมาณฝนในอนาคตที่มีความแปรปรวนสูงและอาจอยู่นอกเหนือขอบเขตของข้อมูลเดิมที่แบบจำลองได้ทำการเรียนรู้ ดังนั้นเพื่อเป็นการเพิ่มขีดความสามารถของแบบจำลองให้สามารถรองรับกับสถานการณ์ที่ไม่เคยปรากฏมาก่อน จึงควรมีการประยุกต์ใช้ข้อมูลนำเข้าจากแหล่งอื่นเพิ่มเติม เช่น ข้อมูลจาก GCM ต่าง ๆ ภายใต้โครงการ CMIP6 ซึ่งจะช่วยให้สามารถสะท้อนลักษณะของสภาพภูมิอากาศในอนาคตภายใต้สถานการณ์การปล่อยก๊าซเรือนกระจกแบบต่าง ๆ ได้ และเป็นการนำเสนอรูปแบบหรือความสัมพันธ์ที่อาจยังไม่เคยปรากฏขึ้นในอดีต โดยการปรับเปลี่ยนหรือเพิ่มเติมชุดข้อมูลนำเข้าที่มาจากแหล่งต่าง ๆ เหล่านี้ ไม่เพียงช่วยให้แบบจำลองมีความยืดหยุ่นมากขึ้นเท่านั้น แต่ยังเป็นการ

ทดสอบความสามารถของแบบจำลองในสถานการณ์ที่มีความไม่แน่นอนสูง ซึ่งมีความสำคัญอย่างยิ่งต่อการประยุกต์ใช้ในการคาดการณ์แนวโน้มและปริมาณฝนในระยะยาวเพื่อการวางแผนจัดการน้ำในอนาคต

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับการสนับสนุนจากทุนเพชรโยธา โดยภาควิชาวิศวกรรมโยธา คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และ ทุนวิจัยพระจอมเกล้าธนบุรี ประจำปีงบประมาณ 2567 จากมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

เอกสารอ้างอิง

[1] Hattermann, F.F., Krysanova, V., Gosling, S.N., Dankers, R., Daggupati, P., Donnelly, C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., Yang, T., Müller, C., Leng, G., Tang, Q., Portmann, F.T., Hagemann, S., Gerten, D., Wada, Y., Masaki, Y., Alemayehu, T., Satoh, Y. and Samaniego, L. (2017). Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. Climatic Change, 141(3), pp. 561–576.

[2] Kim, U., Kaluarachchi, J.J. and Smakhtin, V.U. (2008). Assessing climate change impacts on water resources in the Upper Blue Nile River Basin using a regional climate model. Journal of Hydrology, 372(1–4), pp. 245–256.

[3] von Storch, H., Zorita, E. and Cubasch, U. (1993). Statistical downscaling of general circulation model output: A comparison of methods. Journal of Climate, 6(1), pp. 119–134.

[4] Wang, Q., Huang, J., Liu, R., Men, C., Guo, L., Miao, Y., Jiao, L., Wang, Y., Shoaib, M. and Xia, X. (2020). Sequence-based statistical downscaling and its application to hydrologic simulations based on machine learning and big data. Journal of Hydrology, 586, 124875.

[5] กรมชลประทาน. (2562). รายงานการศึกษาสถานการณ์น้ำในลุ่มน้ำชีมูล ปี 2562. กองอุทกวิทยา, กรมชลประทาน. เข้าถึงจาก https://water.rid.go.th/hyd/Report/Mun_2562.pdf

[6] Vu, M.T., Aribarg, T., Supratid, S., Raghavan, S.V. and Liong, S.Y. (2015). Statistical downscaling rainfall using artificial neural network: Significantly wetter Bangkok? Theoretical and Applied Climatology, [online] https://doi.org/10.1007/s00704-015-1580-1

[7] รักษ์คณา ภูสีเขียว และ สุรณพีร์ ภูมิวัฒนสาร (2565). การพยากรณ์ปริมาณน้ำฝนระยะสั้นในบริเวณพื้นที่สนามบินสุวรรณภูมิด้วยโครงข่ายระบบประสาทแบบย้อนกลับ. วารสารวิทยาการสารสนเทศและเทคโนโลยี, ปีที่ 12(1), หน้า 13–26

[8] Bun, A., Lou, Y., Wu, Y., Wang, Y., & Vaze, J. (2024). A coupled hydrologic–machine learning modelling framework to support real-time streamflow forecasting. Environmental Modelling & Software, 172, 105789. https://doi.org/10.1016/j.envsoft.2024.105789

[9] เอี่ยมอำไพ, ศ. และ รังสิวณิชพงศ์, ป. (2566). การประเมินน้ำท่าในลุ่มน้ำชีและปิงด้วยแบบจำลองโครงข่ายประสาทเทียม. การประชุมวิชาการวิศวกรรมโยธาแห่งชาติ ครั้งที่ 28, 24–26 พฤษภาคม 2566, จังหวัดภูเก็ต เข้าถึงจาก https://conference.thaince.org/index.php/ncce28/article/download/2184/1464/33391

[10] Miller, S., Ormaza-Zulueta, N., Koppa, N., & Dancer, A. (2025). Statistical downscaling differences strongly alter projected climate damages. Communications Earth & Environment, 6(145). https://doi.org/10.1038/s43247-025-02134-2

[11] APXML. (n.d.). Train-test split in time series. สืบค้นเมื่อ 15 พฤศจิกายน 2567, จาก https://apxml.com/courses/time-series-analysis-forecasting/chapter-6-model-evaluation-selection/train-test-split-time-series