การพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำโดยใช้วิธีวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล

Reservoir Inflow Forecasting Using Time Series Data Mining Techniques

วีระยุทธ พิมพาภรณ์[1*] ปรียาภรณ์ พูลทอง[2] และ บุษกร แก้ววิเชียร[2]

Werayut Pimpaporn[1*] Preeyaporn Pooltong[2] and Bootsakorn Kaewwichian[2]

[1*]*อาจารย์ประจำ หลักสูตรวิทยาศาสตร์บัณฑิต ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ ศรีราชา จังหวัดชลบุรี 20230*

[2] *นักศึกษาปริญญาตรี หลักสูตรวิทยาศาสตร์บัณฑิต ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ ศรีราชา จังหวัดชลบุรี 20230*

[1*]*Lecturer in Bachelor of Science (Information Technology) in Department of Computer Science and Information of Faculty of Science at Si Racha, Kasetsart University Sriracha Campus, Chonburi, Thailand 20230*

[2] *Bachelor student in Bachelor of Science (Computer Science) in Department of Computer Science and Information of Faculty of Science at Si Racha, Kasetsart University Sriracha Campus, Chonburi, Thailand 20230*

[*]*Corresponding author, E mail: pimpaporn.w@ku.th*

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี ของโครงการส่งน้ำและบำรุงรักษาปราณบุรี จังหวัดประจวบคีรีขันธ์ โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล 3 เทคนิค ได้แก่ 1) การถดถอยเชิงเส้น 2) แบบจำลองโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น และ 3) ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย โดยทั้ง 3 เทคนิคจะถูกนำไปสร้างแบบจำลองการพยากรณ์และวัดประสิทธิภาพด้วยค่าความเคลื่อนที่กำลังสองและค่าเฉลี่ยความเคลื่อนที่กำลังสองเพื่อเลือกเทคนิคที่ดีที่สุด ข้อมูลอนุกรมเวลาที่นำมาใช้ในการศึกษาเป็นข้อมูลปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี ตั้งแต่ปี 2556 ถึง 2558

จากผลการเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีกับชุดข้อมูลในปี 2558 พบว่า แบบจำลองการพยากรณ์ด้วยเทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย มีประสิทธิภาพโดยรวมสูงที่สุด โดยมีค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ เท่ากับ 18.55% ลำดับถัดมาได้แก่ แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น โดยมีค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ เท่ากับ 33.61%

คำสำคัญ: *การพยากรณ์ ปริมาณน้ำไหลเข้า การวิเคราะห์อนุกรมเวลา เทคนิควิธีเหมืองข้อมูล*

Abstract

The purposes of this research were to develop the forecasting model of Pran Buri reservoir inflow in Prachuap Khiri Khan Province using the time series data mining techniques by three methods, as follows: 1) Linear Regression, 2) Multi-Layer Perceptron, and 3) Support Vector Machine for Regression. All of three methods were developed as the forecasting model and the effectiveness was measured by RMSE and MMRE, for selecting the best method. The time series data for study was the number of Pran Buri reservoir inflow from year 2013 - 2015.

Comparison of the effectiveness of forecasting' s model of Pran Buri reservoir inflow with time series testing data in 2015 found that the support vector machine for regression was the highest performance model for forecasting at 18.55% of MMRE, next, the multilayer perceptron forecasting was at 33.61% of MMRE.

Keywords: *Forecasting, Reservoir Inflow, Time Series Analysis, Data Mining technique*

1. บทนำ

ระบบบริหารจัดการน้ำที่มีคุณภาพจะส่งผลโดยตรงต่อการดำเนินชีวิตประจำวันของประชาชน ในปัจจุบันหลายหน่วยงานที่เกี่ยวข้องกับการบริหารจัดการน้ำได้ให้ความสำคัญในประเด็นที่เกี่ยวข้องกับสถานการณ์ภัยพิบัติด้านน้ำโดยเฉพาะอย่างยิ่งในด้านภัยแล้ง ซึ่งปัญหาดังกล่าวก่อให้เกิดผลกระทบต่อหลายภาคด้าน ทั้งด้านปัญหาการอุปโภคบริโภค ปัญหาด้านการรักษาระบบนิเวศน์และสิ่งแวดล้อม ปัญหาด้านการเกษตร เป็นต้น อย่างไรก็ตาม เพื่อเป็นการเตรียมรับมือกับสถานการณ์ดังกล่าวจึงก่อให้เกิด โครงการส่งน้ำและบำรุงรักษาปราณบุรี ซึ่งเป็นหน่วยงานหนึ่งของกรมชลประทาน ซึ่งมีหน้าที่ในการวิเคราะห์สถานการณ์น้ำในแต่ละวัน ประกอบไปด้วย ระดับน้ำทะเลเหนือน้ำ ระดับน้ำทะเลท้ายน้ำ ปริมาณน้ำในอ่าง ปริมาณน้ำไหลเข้าอ่าง ปริมาณน้ำที่ปล่อยออกสู่คลองสายใหญ่ ปริมาณน้ำฝนและจุดวัดน้ำฝนเหนืออ่างเก็บน้ำ เพื่อส่งข้อมูลให้กับศูนย์ประมวลวิเคราะห์สถานการณ์น้ำกรมชลประทาน และภารกิจหลักของโครงการส่งน้ำและบำรุงรักษาปราณบุรี คือ การจัดการทรัพยากรน้ำในพื้นที่ลุ่มน้ำเพื่อให้เพียงพอต่อความต้องการ จึงจำเป็นต้องมีการวางแผน การตัดสินใจ และการบริหารจัดสรรน้ำในหลายด้าน อาทิ เช่น รักษาระบบนิเวศน์แม่น้ำปราณบุรี อุปโภค-บริโภคสำหรับท่อน้ำดิบ 4 สาย อุปโภค-บริโภคสำหรับพื้นที่ฝ่ายส่งน้ำที่ 2 การประปาส่วนภูมิภาคสาขาปราณบุรี จากประเด็นข้างต้นจะเห็นได้ว่าการบริหารความต้องการในการใช้น้ำถือได้ว่าเป็นสิ่งที่สามารถจัดการได้และการบริหารจัดการจะมีประสิทธิภาพเพิ่มมากขึ้นเมื่อผู้บริหารสามารถทราบปริมาณน้ำหรือแนวโน้มของน้ำที่จะเข้าสู่อ่างเก็บน้ำ ซึ่งจะทำให้ผู้บริหารสามารถจัดทำแผนรับสถานการณ์ที่กำลังจะเกิดขึ้นได้อย่างถูกต้อง

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำของโครงการส่งน้ำและบำรุงรักษาปราณบุรี จังหวัดประจวบคีรีขันธ์ โดยผู้วิจัยทำการศึกษาชุดข้อมูลปริมาณน้ำไหลเข้าอ่างในช่วงระยะเวลาที่ผ่านมา ตั้งแต่ปี พ.ศ. 2556 - 2558 ทั้งนี้ปัจจัยที่ส่งผลกับระดับปริมาณน้ำไหลเข้าอ่างประกอบด้วย ปริมาณน้ำในอ่างของแต่ละวัน ปริมาณน้ำที่ปล่อยออกสู่คลองสายใหญ่ การระเหยของน้ำ อัตราการรั่วซึมและปริมาณน้ำฝนของสถานี PR-3A

หากการพยากรณ์น้ำไหลเข้าอ่างมีความแม่นยำ ก็จะส่งผลให้สามารถจัดการกับปริมาณน้ำปล่อยออกได้อย่างเหมาะสม และทำให้ทราบถึงปริมาณน้ำต้นทุนในอนาคตเพื่อทำการวางแผนจัดการปริมาณน้ำ จะทำให้ลดปัญหาการขาดแคลนน้ำในฤดูแล้งและป้องกันอุทกภัยในช่วงฤดูฝนได้

สำหรับงานวิจัยด้านการวิเคราะห์อนุกรมเวลาซึ่งได้รับความนิยมในอดีต (Classical Time Series Analysis) (Kirchgässner, Wolters, & Hassler, 2012) จะมุ่งเน้นการแยกองค์ประกอบของอนุกรมเวลา เช่น ค่าแนวโน้ม (Trend) ค่าฤดูการ (Seasonal) เป็นต้น ซึ่งในปัจจุบันได้มีการนำความรู้ด้านและเทคนิควิธีเหมืองข้อมูลมาประยุกต์ใช้ด้านชุดข้อมูลอนุกรมเวลาเพื่อสร้างแบบจำลองการทำนายที่มีประสิทธิภาพยิ่งขึ้น โดยงานวิจัยนี้เลือกใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining Techniques) 3 เทคนิค ได้แก่ การถดถอยเชิงเส้น (Linear Regression) แบบจำลองโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) และซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) โดยทั้ง 3 เทคนิคเป็นเทคนิคที่ได้รับความนิยมสำหรับการสร้างแบบจำลองการพยากรณ์ซึ่งมีค่าคำตอบเป็นตัวเลขหรือค่าปริมาณ โดยผลลัพธ์ของการประมวลผลข้อมูลอนุกรมเวลาจะอยู่ในรูปของแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำ และสามารถนำแบบจำลองที่ได้ไปทำนายอัตราปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีของโครงการส่งน้ำและบำรุงรักษาปราณบุรี เพื่อนำไปใช้ในการวางแผนบริหารและจัดการน้ำได้อย่างมีประสิทธิภาพต่อไป

2. วัตถุประสงค์

1. เพื่อเปรียบเทียบประสิทธิภาพ และเลือกเทคนิควิธีที่เหมาะสม ในการวิเคราะห์อนุกรมเวลาของปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี

2. เพื่อสร้างรูปแบบการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี โดยวิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล

3. วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มุ่งเน้นการสร้างแบบจำลองเพื่อพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีของโครงการส่งน้ำและบำรุงรักษาปราณบุรี โดยดำเนินการวิจัยตามขั้นตอน 3 ขั้นตอนดังนี้

3.1 ศึกษาปัญหาและวิเคราะห์ข้อมูล

การบริหารจัดการน้ำในอ่างเก็บน้ำเป็นภารกิจหนึ่งที่สำคัญของโครงการส่งน้ำและบำรุงรักษาปราณบุรี งานในด้านการบริหารจัดการน้ำชลประทานคือการส่งน้ำให้แก่พื้นที่ซึ่งมีความจำเป็นต่อการใช้น้ำโดยเฉพาะอย่างยิ่งการส่งน้ำเพื่อให้เกษตรกรได้รับน้ำทันตามเวลาที่กำหนด เมื่อพิจารณาประเด็นที่เกี่ยวข้องกับการบริหารจัดการน้ำในอ่างเก็บน้ำจะพบว่า น้ำในอ่างเก็บน้ำถือได้ว่าเป็นแหล่งต้นทุนน้ำที่สำคัญ จึงส่งผลให้ผู้ที่มีหน้าที่บริหารจัดการน้ำในอ่างเก็บน้ำจำเป็นอย่างยิ่งที่ต้องทราบถึงปริมาณความต้องการใช้น้ำของผู้ใช้ เพื่อใช้ในการวางแผนการจ่ายน้ำจากอ่างเก็บน้ำ แต่เมื่อพิจารณาถึงองค์ประกอบที่สำคัญอีกประการในการบริการจัดการน้ำพบว่า การที่ปริมาณน้ำจะเพียงพอต่อความต้องการของผู้ใช้ได้นั้น น้ำที่ไหลเข้าสู่อ่างเก็บน้ำยังเป็นอีกปัจจัยหนึ่งที่มีความสำคัญอย่างยิ่งกับกระบวนบริหารจัดการน้ำให้มีประสิทธิภาพ ทั้งนี้เมื่อผู้บริหารจัดการน้ำสามารถประมาณการปริมาณน้ำที่จะไหลเข้าสู่อ่างเก็บน้ำได้ จะส่งผลให้การ

บริหารจัดการน้ำสามารถทำได้มากขึ้น เช่น การเพิ่มระดับน้ำในอ่างเก็บน้ำให้สูงขึ้นเมื่อทราบว่าช่วงเวลาถัดไปจะมีน้ำเข้าสู่อ่างเก็บน้ำน้อยลง ซึ่งจะส่งผลให้ปริมาณน้ำเพียงพอต่อการส่งน้ำตลอดช่วงเวลา ในทางกลับกันเมื่อผู้บริหารจัดการน้ำประมาณการได้ว่าปริมาณน้ำไหลเข้าสู่อ่างเก็บน้ำในช่วงถัดไปจะมากขึ้นก็จะสามารถลดระดับน้ำในอ่างเก็บน้ำเพื่อรองรับน้ำที่จะเข้ามา อีกทั้งยังป้องกันการเกิดน้ำท่วมจากภาวะน้ำล้นได้อีกด้วย

จากการศึกษาและวิเคราะห์ข้อมูลปริมาณน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรีพบว่า ชุดข้อมูลมีลักษณะเป็น ข้อมูลอนุกรมเวลา (Time Series Data) (Fu, 2011) ที่มีการจัดเก็บทุกวัน ดังนั้นผู้วิจัยจึงเลือกเทคนิควิธีการวิเคราะห์อนุกรมเวลาเพื่อทำการสร้างแบบจำลองในการพยากรณ์ปริมาณน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรี สำหรับการวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของแบบจำลองในการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำ โดยวิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล ผู้วิจัยจึงเลือกชุดปริมาณน้ำไหลเข้าอ่างเก็บน้ำตั้งแต่ปี 2556 - 2558 เพื่อใช้ในกระบวนการวิจัย

3.2 การเตรียมข้อมูล (Data Preparation)

ชุดข้อมูลที่ใช้ในงานวิจัยนี้ประกอบไปด้วยปริมาณน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรี โดยข้อมูลจะมีการบันทึกเป็นรายวันตั้งแต่ปี 2556 - 2558 โดยในปี 2558 บันทึกข้อมูลทั้งหมด 42 สัปดาห์ หรือ 10 เดือน สำหรับขั้นตอนในการจัดเตรียมข้อมูลสำหรับกระบวนการวิจัย ผู้วิจัยทำการแบ่งข้อมูลออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลเรียนรู้ (Training Data set) เป็นชุดข้อมูลสำหรับสร้างแบบจำลอง (Time Series Model) การประมาณการน้ำ โดยใช้ข้อมูลในระหว่างปี 2556 - 2557 และทำการแบ่งข้อมูลที่เหลือในปี 2558 สำหรับใช้เป็นข้อมูลในการทดสอบ (Testing Data set) แบบจำลองการพยากรณ์ เพื่อเปรียบเทียบประสิทธิภาพของเทคนิควิธีเหมืองข้อมูลที่ใช้ในการสร้างแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี

สำหรับกระบวนการในการจัดเตรียมข้อมูลเพื่อนำเข้าสู่อัลกอริทึมในการสร้างแบบจำลองการพยากรณ์ปริมาณน้ำ ผู้วิจัยทำการแปลงข้อมูล (Data Transformation) จากข้อมูลที่มีการจัดเก็บรายวัน (Day) ให้อยู่ในรูปแบบของข้อมูลรายสัปดาห์ (Week) เพื่อใช้ในกระบวนการวิเคราะห์หาชุดข้อมูลย้อนหลัง (Lagged Data) จากชุดข้อมูลเรียนรู้ (Training Data set) ที่มีความเหมาะสมในการสร้างแบบจำลองการพยากรณ์ปริมาณน้ำ และแบบจำลองที่ได้จากวิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล จะถูกนำไปทดสอบหาประสิทธิภาพกับชุดข้อมูลทดสอบ (Testing Data set) และนำเสนอข้อมูลความคลาดเคลื่อนในรูปแบบของข้อมูลรายเดือน (Month)

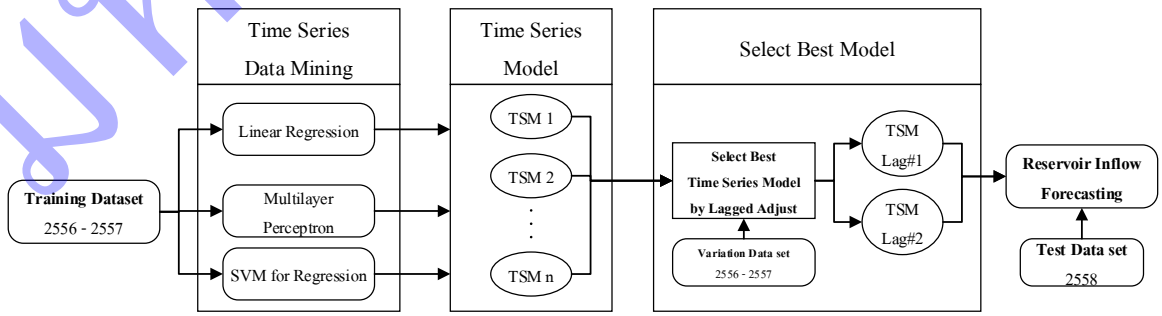
3.3 การวิเคราะห์ข้อมูล (Data Analysis)

งานวิจัยนี้มุ่งเน้นการสร้างแบบจำลองเพื่อใช้ในการพยากรณ์ปริมาณน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรี โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining techniques) 3 เทคนิค ได้แก่ การถดถอยเชิงเส้น (Linear Regression) (Cai, Hall, & others, 2006) แบบจำลองโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) (Frias-Martinez, Sanchez, & Velez, 2006) ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) (Shevade, Keerthi, Bhattacharyya, & Murthy, 2000; Smola & Schölkopf, 2004) โดยการนำแบบจำลองการพยากรณ์ที่ได้จากเทคนิคต่าง ๆ มาเปรียบเทียบประสิทธิภาพ การวิจัยนี้มีขั้นตอนในการวิเคราะห์ข้อมูลสำคัญ 2 ส่วน คือ 1) การวิเคราะห์ข้อมูลเพื่อหาจำนวนข้อมูลย้อนหลัง (Lagged) ที่มีความสามารถใน

การสร้างแบบจำลองการพยากรณ์ที่มีประสิทธิภาพดีและมีความเหมาะสมกับเทคนิควิธีเหมืองข้อมูลทั้ง 3 เทคนิค สำหรับขั้นตอนนี้ผู้วิจัยเลือกใช้วิธีการวัดรากของความเคลื่อนที่กำลังสอง (Root Mean Square Error : RMSE) และค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Error : MAE) เพื่อแสดงการเปรียบเทียบประสิทธิภาพของแบบจำลองเพื่อใช้ประมาณค่าน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรี 2) การทดสอบประสิทธิภาพกับชุดข้อมูลทดสอบ (Testing Data set) โดยใช้วิธีการประมาณการความแม่นยำในการพยากรณ์ด้วยค่าความคาดเคลื่อนสัมพัทธ์ (Magnitude of Relative Error : MRE) สำหรับทดสอบประสิทธิภาพของแบบจำลองโดยแยกแต่ละเดือนในปี 2558 และใช้ค่าเฉลี่ยความคาดเคลื่อนสัมพัทธ์ (Mean Magnitude of Relative Error : MMRE) เพื่อใช้ในการทดสอบประสิทธิภาพโดยรวมของแบบจำลองการพยากรณ์

สำหรับการสร้างแบบจำลองการพยากรณ์ผู้วิจัยเลือกใช้โปรแกรม WEKA version 3.7.13 เพื่อทำการวิเคราะห์ข้อมูลและสร้างแบบจำลองการพยากรณ์โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining techniques) วิธีการทั้งหมดจะให้ผลลัพธ์ในรูปแบบของโมเดลซึ่งถือเป็นลักษณะของการแทนความรู้ (Knowledge Representation) แบบหนึ่ง สำหรับรูปแบบการพยากรณ์จะมีลักษณะเป็นแบบ Sliding Window ซึ่งเป็นเทคนิควิธีในการจัดเรียงข้อมูลแบบอนุกรมเวลา โดยสร้างชุดข้อมูลสำหรับการสร้างแบบจำลองการพยากรณ์ผู้วิจัยจะนำข้อมูลในปี 2556 - 2557 เป็นชุดข้อมูลทดสอบแบ่งเป็นชุดข้อมูลย้อนหลัง (Lagged) ซึ่งข้อมูลย้อนหลังจะเทียบได้กับตัวแปรหรือแอตทริบิวต์ (Attribute) ที่จะนำเข้าสู่กระบวนการเหมืองข้อมูล โดยชุดข้อมูลสามารถแบ่งได้ดังนี้ 1) 13 สัปดาห์ 2) 26 สัปดาห์ 3) 39 สัปดาห์ 4) 52 สัปดาห์ 5) 65 สัปดาห์ 6) 78 สัปดาห์ และ 7) 91 สัปดาห์ โดยผลลัพธ์ในการทดสอบประสิทธิภาพของชุดข้อมูลทั้ง 7 ชุดข้อมูล จะใช้เป็นตัวชี้วัดว่าอัลกอริทึมใดเหมาะสมกับการใช้ในการสร้างแบบจำลอง สำหรับข้อมูลในปี 2558 ซึ่งมีการแบ่งข้อมูลเป็นรายเดือน กำหนดให้เป็นชุดข้อมูลทดสอบซึ่งถือเป็นเป้าหมาย (Target) ในการพยากรณ์ประสิทธิภาพของการพยากรณ์ในแต่ละอัลกอริทึมจะเป็นดัชนีชี้วัดว่าอัลกอริทึมใดมีความเหมาะสมในการพยากรณ์ปริมาณน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรีในแต่ละเดือน

จากวิธีการดำเนินการวิจัยที่ได้กล่าวมาสามารถแสดงขั้นตอนในการสร้างแบบจำลองการสร้างแบบจำลองเพื่อพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี ได้ดังรูปที่ 1



รูปที่ 1 ขั้นตอนการสร้างแบบจำลองเพื่อพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีของโครงการส่งน้ำและบำรุงรักษาปราณบุรี

4. ผลการวิจัย

จากผลการทดลองได้ทำการวัดประสิทธิภาพแบบจำลองเพื่อพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี ของโครงการส่งน้ำและบำรุงรักษาปราณบุรี โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล ตามวิธีการดำเนินงานวิจัย ด้วยใช้เทคนิควิธีเหมืองข้อมูลทั้ง 3 เทคนิค กับทดสอบชุดข้อมูลฝึกฝน (Training Data Set) และชุดข้อมูลทดสอบ (Testing Data set) จำนวน 3 ชุด ตามขั้นตอนวิธีวิจัย ผลการทดสอบสามารถแบ่งออกเป็น 2 ส่วนดังนี้

4.1 การเปรียบเทียบประสิทธิภาพในกระบวนการสร้างแบบจำลองเพื่อพยากรณ์ปริมาณน้ำ

จากชุดข้อมูลปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีที่ใช้เป็นชุดข้อมูลสำหรับการสร้างแบบจำลองการพยากรณ์ ผู้วิจัยนำข้อมูลในปี 2556 - 2557 มาใช้ชุดข้อมูลเรียนรู้เพื่อสร้างแบบจำลองการพยากรณ์ โดยลักษณะข้อมูลถูกแปลงให้อยู่ในรูปแบบข้อมูลปริมาณน้ำไหลเข้าอ่างในแต่ละสัปดาห์เรียงต่อกันจำนวน 91 สัปดาห์ โดยชุดข้อมูลทดสอบแบ่งเป็น ชุดข้อมูลย้อนหลัง (Lagged) 1) 13 สัปดาห์ 2) 26 สัปดาห์ 3) 39 สัปดาห์ 4) 52 สัปดาห์ 5) 65 สัปดาห์ 6) 78 สัปดาห์ และ 7) 91 สัปดาห์ ถูกนำมาสร้างแบบจำลองในการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำ เพื่อวิเคราะห์ประสิทธิภาพที่ได้จากเทคนิควิธีเหมืองข้อมูลทั้ง 3 เทคนิค โดยวิธีการเปรียบเทียบความคลาดเคลื่อนด้วยวิธีการ MAE และ RMSE จากผลการทดลองสามารถแสดงได้ดังตาราง

ตารางที่ 1 การเปรียบเทียบประสิทธิภาพแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำโดยใช้ชุดข้อมูลย้อนหลัง (Lagged)

Training Data Set		Time Series Data Mining Techniques					
		LinearRegression		MultilayerPerceptron		SMOreg	
Year	Week Lagged	MAE	RMSE	MAE	RMSE	MAE	RMSE
2557	Lag13	6.99	13.52	5.45	7.67	5.64	14.90
	Lag26	7.80	13.93	6.32	8.44	6.24	16.00
	Lag39	7.69	15.20	23.49	25.29	5.21	16.49
	***Lag52**	**4.01**	**6.26**	3.11	3.42	1.00	3.83
	***Lag65**	4.91	7.14	**0.91**	**1.17**	**0.11**	**0.12**
2556	Lag78	6.59	8.90	1.82	2.13	0.13	0.14
	Lag91	6.84	9.67	2.07	2.99	0.14	0.14

จากตารางที่ 1 แสดงผลการพยากรณ์ของแต่ละเทคนิควิธีกับการใช้ชุดข้อมูลโดยมีจำนวนสัปดาห์ย้อนหลัง (Lagged) ที่แตกต่างกันเมื่อใช้ค่า MAE และ RMSE เป็นดัชนีชี้วัดประสิทธิภาพ จากการทดลองแสดงให้เห็นว่า เมื่อสร้างแบบจำลองด้วยเทคนิคการถดถอยเชิงเส้น ด้วยข้อมูลย้อนหลัง 52 สัปดาห์ (Lag52) มีประสิทธิภาพสูงที่สุด (MAE = 4.01, RMSE = 6.26) สำหรับแบบจำลองที่สร้างด้วยโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น โดยใช้ข้อมูลย้อนหลัง 65 สัปดาห์ (Lag65) มีประสิทธิภาพสูงที่สุด (MAE = 0.91, RMSE = 1.17) และแบบจำลองที่สร้างด้วยซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย ด้วยชุดข้อมูลย้อนหลัง 65 วัน

(Lag65) มีประสิทธิภาพสูงที่สุด (MAE = 0.11, RMSE = 0.12)

จากผลการทดลองผู้วิจัยจึงเลือกชุดข้อมูลย้อนหลังที่ทำให้ได้ค่าประสิทธิภาพที่ดีที่สุด ไปใช้ในการสร้างแบบจำลองเพื่อการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี บนชุดข้อมูลทดสอบ (Testing Dataset) ซึ่งเป็นปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีในช่วงปี 2558

4.2 การเปรียบเทียบประสิทธิภาพแบบจำลองกับชุดข้อมูลทดสอบ

จากการนำแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีที่มีค่าประสิทธิภาพมากที่สุดในแต่ละเทคนิควิธีมาทดสอบกับชุดข้อมูลปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรีในช่วงปี 2558 โดยทำการพิจารณาประสิทธิภาพในการพยากรณ์เป็นรายเดือนโดยกำหนดให้ค่าความคลาดเคลื่อนสัมพัทธ์ (Magnitude of Relative Error) เป็นค่าดัชนีชี้วัดประสิทธิภาพของแต่ละเดือน และใช้ค่าเฉลี่ยความคาดเคลื่อนสัมพัทธ์ (Mean Magnitude of Relative Error : MMRE) เป็นดัชนีชี้วัดประสิทธิภาพโดยรวมของแบบจำลองการพยากรณ์ที่ได้จากแต่ละเทคนิควิธี ผลจากการทดลองสามารถแสดงได้ดังตาราง

ตารางที่ 2 การเปรียบเทียบประสิทธิภาพแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำในแต่ละเดือน

Test Data Set			Time Series Model					
			LinearRegression (Lag 52W)		MultilayerPerceptron (Lag 65W)		SMOreg (Lag 65W)	
Year	Month	Actual	Predict	MRE	Predict	MRE	Predict	MRE
2558	ม.ค. (W1 - W4)	14.17	13.23	0.07	11.17	0.21	14.85	**0.05**
	ก.พ. (W5 - W8)	10.39	17.71	0.70	5.97	0.43	10.53	**0.01**
	มี.ค. (W9 - W13)	3.90	12.56	2.22	3.36	0.14	3.91	**0.01**
	เม.ย. (W14 - W17)	2.18	8.23	2.77	1.86	**0.15**	4.17	0.91
	พ.ค. (W18 - W21)	1.46	7.05	3.83	0.30	0.79	1.56	**0.07**
	มิ.ย. (W22 - W26)	13.57	10.26	0.24	9.24	0.32	11.03	**0.19**
	ก.ค. (W27 - W30)	24.31	22.88	0.06	17.30	0.29	25.67	**0.06**
	ส.ค. (W31 - W34)	13.85	20.52	0.48	5.16	0.63	15.14	**0.09**
	ก.ย. (W35 - W39)	32.79	23.55	0.28	24.23	0.26	26.31	**0.20**
	ต.ค. (W40 - W42)	29.76	22.99	0.23	25.39	**0.15**	21.61	0.27
MMRE			108.90%		33.61%		**18.55%**	

เมื่อพิจารณาผลการทดลองจากตารางที่ 2 พบว่าการพยากรณ์ปริมาณน้ำที่ไหลเข้าสู่อ่างเก็บน้ำปราณบุรี โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining techniques) ในภาพรวมของปี 2558 แบบจำลองการพยากรณ์ด้วยเทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) โดยใช้ชุดข้อมูลย้อนหลัง 65 สัปดาห์ (Lag 65W) มีค่าประสิทธิภาพโดยรวมสูงที่สุด โดยมีค่า MMRE = 18.55% ลำดับถัดมาได้แก่ แบบจำลอง

การพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) โดยใช้ชุดข้อมูลย้อนหลัง 65 สัปดาห์ (Lag 65W) มีค่า MMRE = 33.61%

และเมื่อพิจารณาในรายละเอียดของผลลัพธ์การพยากรณ์แต่ละเดือนในชุดข้อมูลทดสอบปี 2558 พบว่า แบบจำลองการพยากรณ์ด้วยเทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) มีประสิทธิภาพสูงที่สุดจำนวน 8 เดือนจากชุดข้อมูลทั้งหมด 10 เดือน โดยมีค่า MRE อยู่ระหว่าง 0.01 – 0.20 และแบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) มีค่าประสิทธิภาพสูงที่สุดจำนวน 2 เดือนได้แก่ เมษายน และตุลาคม โดยมีค่า MRE อยู่ที่ 0.15 ทั้ง 2 เดือน

5. การอภิปรายผล

จากผลการวิจัยแสดงให้เห็นประเด็นข้อค้นพบสำคัญเกี่ยวกับการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี ของโครงการส่งน้ำและบำรุงรักษาปราณบุรี 2 ประเด็นได้แก่

การพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำปราณบุรี จำนวนข้อมูลย้อนหลังที่มีความเหมาะสมกับการพยากรณ์ปริมาณน้ำไหลเข้าอ่างอยู่ในช่วงระยะเวลา 52 สัปดาห์ ถึง 65 สัปดาห์ หรือ ประมาณ 1 ปี ถึง 1 ปี 4 เดือน ซึ่งจากผลดังกล่าวสามารถอภิปรายได้ว่า ในช่วงเวลาหนึ่งปีการเปลี่ยนแปลงที่เกิดขึ้นกับปริมาณน้ำที่ไหลเข้าอ่างเก็บน้ำเป็นผลโดยตรงจากฤดูการ เช่น ในช่วงฤดูร้อนปริมาณน้ำไหลเข้าอ่างจะมีปริมาณที่น้อย แต่ในช่วงฤดูฝนปริมาณน้ำไหลเข้าอ่างจะมีปริมาณที่เพิ่มมากขึ้น ดังนั้นการพยากรณ์กรปริมาณน้ำจึงต้องมีช่วงฤดูที่ครบทั้ง 3 ฤดู เพื่อให้การวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูลได้เรียนรู้ข้อมูลที่เหมาะสมต่อการสร้างแบบจำลองการพยากรณ์ ซึ่งสอดคล้องกับผลการวิจัยของ (จักรพงษ์ แต้วิจิตร, 2547) แต่อย่างไรก็ตามการที่มีข้อมูลย้อนหลังมากเกินกว่า 1 ปี 4 เดือน จะส่งผลต่อประสิทธิภาพของแบบจำลองการพยากรณ์ลดลง ดังจะเห็นได้จากผลการทดลองในตารางที่ 1

การเลือกใช้เทคนิคเหมืองข้อมูลในการสร้างแบบจำลองเพื่อใช้ในงานด้านการวิเคราะห์อนุกรมเวลาถือได้ว่าเป็นสิ่งที่จำเป็น จากผลการทดลองพบว่าการสร้างแบบจำลองการพยากรณ์ด้วยวิธี การถดถอยเชิงเส้น ให้ค่าความผิดพลาดสูงถึง 108.90% เมื่อวัดด้วยค่า MMRE โดยที่ค่าความผิดพลาดจะลดลงเมื่อเลือกใช้เทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น โดยค่า MMRE เท่ากับ 33.61% และค่าความผิดพลาดน้อยที่สุดด้วยเทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย โดยค่า MMRE เท่ากับ 18.55% อย่างไรก็ตามผู้บริหารจัดการน้ำควรกำหนดค่าความผิดพลาดที่ยอมรับได้สำหรับการนำแบบจำลองไปใช้งานจริง

6. บทสรุป

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำของโครงการส่งน้ำและบำรุงรักษาปราณบุรี จังหวัดประจวบคีรีขันธ์ โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล โดยมีขั้นตอนในการดำเนินการวิจัยแบ่งออกเป็น 3 ขั้นตอนดังนี้ 1) ศึกษาปัญหาและวิเคราะห์ข้อมูล 2) การเตรียมข้อมูลและ 3) การวิเคราะห์ข้อมูล

จากผลการทดลองพบว่าชุดข้อมูลอนุกรมเวลาประมาณน้ำไหลเข้าอ่างเก็บน้ำที่มีการแบ่งข้อมูลเป็นรายสัปดาห์ย้อนหลังจำนวน 52 สัปดาห์ (Lag52)

และย้อนหลังจำนวน 65 สัปดาห์ (Lag65) มีความเหมาะสมในการใช้เป็นชุดข้อมูลเพื่อสร้างแบบจำลองในการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำ

เมื่อพิจารณาประสิทธิภาพของแบบจำลองการพยากรณ์ปริมาณน้ำไหล่เข้าอ่างกับชุดข้อมูลในปี 2558 โดยนำเสนอข้อมูลเป็นรายเดือน พบว่าแบบจำลองการพยากรณ์ด้วยเทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย มีค่าประสิทธิภาพโดยรวมสูงที่สุด โดยมีค่า MMRE = 18.55% ลำดับถัดมาได้แก่ แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น มีค่า MMRE = 33.61%

จากผลการวิจัยแสดงให้เห็นได้ว่าการพยากรณ์ปริมาณน้ำไหลเข้าอ่างเก็บน้ำของโครงการส่งน้ำและบำรุงรักษาปราณบุรี จังหวัดประจวบคีรีขันธ์ ไม่สามารถใช้วิธีการใดวิธีการหนึ่งได้ เนื่องจากลักษณะข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงอยู่ตลอดเวลา และการประมาณค่าที่จะเกิดขึ้นในอนาคตได้อย่างมีประสิทธิภาพนั้นจำเป็นต้องคำนึงถึงรูปแบบของข้อมูลที่เกิดขึ้นในอดีตและเทคนิควิธีที่ใช้เรียนรู้ข้อมูลดังกล่าวที่มีความเหมาะสมกับข้อมูลเพื่อสร้างแบบจำลองที่มีความเหมาะสมกับชุดข้อมูล

7. กิตติกรรมประกาศ

คณะวิจัยขอขอบคุณ โครงการส่งน้ำและบำรุงรักษาปราณบุรี จังหวัดประจวบคีรีขันธ์ ที่ให้การเอื้อเฟื้อข้อมูลที่ใช้ในการวิเคราะห์ ให้คำปรึกษาเกี่ยวกับแนวทางการบริหารจัดการน้ำในอ่างเก็บน้ำปราณบุรี

8. เอกสารอ้างอิง

จักรพงษ์ แต้วิจิตร. (2547). การพยากรณ์ปริมาณน้ำไหลเข้าอ่างด้วยการวิเคราะห์อนุกรมเวลาและประยุกต์ใช้ วิธีเจเนติกอัลกอริทึ่มเพื่อหาโค้งปฏิบัติงานอ่างเก็บน้ำที่เหมาะสม: กรณีศึกษาโครงการอ่างเก็บน้ำทับเสลา จังหวัดอุทัยธานี (วิทยานิพนธ์/Thesis). มหาวิทยาลัยเกษตรศาสตร์, สำนักหอสมุด.

Cai, T. T., Hall, P., & others. (2006). Prediction in functional linear regression. *The Annals of Statistics*, *34*(5), 2159–2179.

Frias-Martinez, E., Sanchez, A., & Velez, J. (2006). Support vector machines versus multi-layer perceptrons for efficient off-line signature recognition. *Engineering Applications of Artificial Intelligence*, *19*(6), 693–704.

Fu, T. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, *24*(1), 164–181.

Kirchgässner, G., Wolters, J., & Hassler, U. (2012). *Introduction to modern time series analysis*. Springer Science & Business Media.

Shevade, S. K., Keerthi, S. S., Bhattacharyya, C., & Murthy, K. R. K. (2000). Improvements to the SMO algorithm for SVM regression. *IEEE Transactions on Neural Networks*, *11*(5), 1188–1193.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222.