

**การพัฒนาและเปรียบเทียบแบบจำลองการพยากรณ์ความต้องการเม็ดพลาสติก:
กรณีศึกษาบริษัทผู้ผลิตในอุตสาหกรรมปิโตรเคมี**

ญาณิศา บุญประภาศรี^{*1} และ กฤษ วงษ์เกษม²

สาขาวิชาวิศวกรรมอุตสาหกรรม ภาควิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
กรุงเทพมหานคร 10900

Received: 11 March 2026; Revised: 28 April 2026; Accepted: 12 May 2026

บทคัดย่อ

อุตสาหกรรมปิโตรเคมีเป็นหนึ่งในอุตสาหกรรมที่มีบทบาทสำคัญต่อระบบเศรษฐกิจ เนื่องจากเป็นแหล่งผลิตวัตถุดิบพื้นฐานสำหรับอุตสาหกรรมหลากหลายประเภท โดยเฉพาะเม็ดพลาสติกซึ่งเป็นวัตถุดิบหลักที่ถูกนำไปใช้ในภาคการผลิตอย่างแพร่หลาย อย่างไรก็ตาม ความผันผวนของปัจจัยทางเศรษฐกิจและความซับซ้อนของอุปสงค์และอุปทานของเม็ดพลาสติกส่งผลให้การพยากรณ์ความต้องการมีความท้าทายสูง งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบแบบจำลองการพยากรณ์ความต้องการเม็ดพลาสติก โดยมีกรณีศึกษาเป็นผู้ผลิตในอุตสาหกรรมปิโตรเคมีรายหนึ่งในประเทศไทย สินค้าที่นำมาศึกษาเป็นสินค้าที่ได้รับการคัดเลือกโดยพิจารณาจากยอดขายรวมสูงสุดในช่วงเวลาที่ศึกษา โดยมีการวิเคราะห์ลักษณะแนวโน้มและฤดูกาลของข้อมูลด้วยวิธี Mann-Kendall Trend Test และ Friedman Test คัดเลือกตัวแปรสำหรับการพยากรณ์ด้วย Mutual Information (MI) จากนั้นจึงพัฒนาและเปรียบเทียบแบบจำลองการพยากรณ์ทั้งหมด 7 แบบ ได้แก่ ARIMA, ARIMAX, XGBoost และแบบจำลองไฮบริด ได้แก่ ARIMA-XGBoost, ARIMAX-XGBoost, XGBoost-ARIMA และ XGBoost-ARIMAX ซึ่งมีโครงสร้างของแบบจำลองที่แตกต่างกัน ในการประเมินประสิทธิภาพ ทำการประเมินด้วยตัวชี้วัด Mean Absolute Percentage Error (MAPE) ผลการศึกษาแสดงให้เห็นว่า ในกรณีศึกษานี้ แบบจำลองไฮบริดที่ใช้แบบจำลอง Machine Learning เป็นแบบจำลองหลักมีแนวโน้มให้ประสิทธิภาพการพยากรณ์ดีกว่าแบบจำลองดั้งเดิมโดยส่วนใหญ่ โดยเฉพาะแบบจำลอง XGBoost-ARIMAX ซึ่งให้ความแม่นยำในการพยากรณ์สูงที่สุดเมื่อเปรียบเทียบกับแบบจำลองทั้งหมด

คำสำคัญ: การพยากรณ์ความต้องการ, อนุกรมเวลา, การเรียนรู้ของเครื่อง, แบบจำลองไฮบริด

* Corresponding author. E-mail: yanisa.boonpra2@ku.th

¹ นักศึกษา ภาควิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

² ผู้ช่วยศาสตราจารย์ ภาควิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

Development and Comparative Analysis of Demand Forecasting Models for Plastic Resin Pellets: A Petrochemical Manufacturing Case Study

Yanisa Boonprapasri^{*1} and Kris Wonggasem²

Department of Industrial Engineering, Faculty of Engineering, Kasetsart University,
Bangkok 10900

Received: 11 March 2026; Revised: 28 April 2026; Accepted: 12 May 2026

Abstract

The petrochemical industry plays a vital role in the economy as a key supplier of raw materials for various industries, particularly plastic resin pellets used in manufacturing. However, economic fluctuations and the complexity of demand and supply make plastic resin pellets demand forecasting highly challenging. This study aims to develop and compare forecasting models for plastic resin pellets demand using a case study of a petrochemical manufacturer in Thailand. The product selected for this study was chosen based on the highest total sales during the study period. Data characteristics were examined using the Mann-Kendall Trend Test and the Friedman Test, while variables were selected using Mutual Information (MI). Seven forecasting models were developed and compared, including ARIMA, ARIMAX, XGBoost, and hybrid models comprising ARIMA-XGBoost, ARIMAX-XGBoost, XGBoost-ARIMA, and XGBoost-ARIMAX. Model performance was evaluated using the Mean Absolute Percentage Error (MAPE). The results indicate that, in this case study, hybrid models employing Machine Learning as the primary model tend to provide better forecasting performance than most traditional models. In particular, the XGBoost-ARIMAX model achieved the highest forecasting accuracy among all models considered.

Keywords: demand forecasting, time series analysis, machine learning, hybrid forecasting models

* Corresponding author. E-mail: yanisa.boonpra2@ku.th

¹ Student in Faculty of engineering, Kasetsart University

² Assistant Professor in Faculty of engineering, Kasetsart University

1. บทนำ

อุตสาหกรรมปิโตรเคมีเป็นหนึ่งในอุตสาหกรรมที่มีบทบาทสำคัญต่อการขับเคลื่อนเศรษฐกิจ เนื่องจากเป็นอุตสาหกรรมต้นน้ำของหลากหลายอุตสาหกรรม เช่น อุตสาหกรรมพลาสติกและบรรจุภัณฑ์ อุตสาหกรรมยานยนต์ อุตสาหกรรมก่อสร้าง อุตสาหกรรมเกษตรและประมง เป็นต้น อุตสาหกรรมปิโตรเคมีเป็นอุตสาหกรรมที่มีการนำวัตถุดิบจากอุตสาหกรรมปิโตรเลียม เช่น แนฟทา ซึ่งเป็นผลิตภัณฑ์จากการกลั่นน้ำมันดิบ มาผ่านกระบวนการแปรรูปจนกลายเป็นผลิตภัณฑ์ปิโตรเคมี อาทิ เม็ดพลาสติกชนิดต่าง ๆ ได้แก่ พอลิเอทิลีน (Polyethylene: PE) พอลิโพรพิลีน (Polypropylene: PP) พอลิไวนิลคลอไรด์ (Polyvinyl Chloride: PVC) และ พอลิเอทิลีน เทเรฟทาเลต (Polyethylene Terephthalate: PET) รวมทั้งเส้นใยสังเคราะห์ ยางสังเคราะห์ สารเคลือบผิว และกาว เป็นต้น [1]

ในปี พ.ศ. 2565 อุตสาหกรรมปิโตรเคมีของไทยมีรายได้รวม 836,000 ล้านบาท คิดเป็น 5.2 เปอร์เซ็นต์ของ GDP และมีมูลค่าส่งออก 486,000 ล้านบาท คิดเป็น 5.7 เปอร์เซ็นต์ของมูลค่าการส่งออกสินค้าทั้งหมด โดยเมื่อพิจารณาถึงห่วงโซ่มูลค่า ตั้งแต่การผลิตผลิตภัณฑ์ปิโตรเคมีขั้นต้น ไปจนถึงการแปรรูปเป็นผลิตภัณฑ์พลาสติก พบว่าสามารถสร้างมูลค่ารวมได้มากถึง 1,720 พันล้านบาท หรือราว 10.7 เปอร์เซ็นต์ของ GDP ของประเทศ [2] นอกจากนี้ ในอุตสาหกรรมปิโตรเคมีโลก มีการคาดการณ์ว่าในปี พ.ศ. 2567 ถึง พ.ศ. 2569 ความต้องการใช้ผลิตภัณฑ์ปิโตรเคมีจะมีอัตราการเติบโตเฉลี่ย 1.5 ถึง 2.5 เปอร์เซ็นต์ต่อปี [3] และปี พ.ศ. 2568 ถึง พ.ศ. 2570 คาดว่าความต้องการใช้พลาสติกทั่วโลกจะขยายตัวเฉลี่ย 2 ถึง 3 เปอร์เซ็นต์ต่อปี [4] จากข้อมูลดังกล่าว จึงสะท้อนให้เห็นได้ว่าอุตสาหกรรมปิโตรเคมีเป็นอุตสาหกรรมซึ่งเป็นบทบาทสำคัญในการขับเคลื่อนเศรษฐกิจและอุตสาหกรรมของประเทศ จึงเป็นอุตสาหกรรมที่ควรได้รับการให้ความสำคัญอย่างต่อเนื่อง

อย่างไรก็ตาม อุตสาหกรรมปิโตรเคมีเป็นอุตสาหกรรมที่มีความซับซ้อนสูง เนื่องจากมีความเชื่อมโยงโดยตรงกับอุตสาหกรรมปิโตรเลียม ซึ่งเป็นวัตถุดิบหลักสำหรับนำมาผลิตเป็นผลิตภัณฑ์ปิโตรเคมี ส่งผลให้ได้รับผลกระทบโดยตรงจากความผันผวนทั้งในด้านราคา และ

ปริมาณวัตถุดิบของอุตสาหกรรมปิโตรเลียม อีกทั้งยังมีปัจจัยภายนอก เช่น ภาวะเศรษฐกิจ ซึ่งจะส่งผลกระทบต่ออุตสาหกรรมปิโตรเลียม และอุตสาหกรรมปิโตรเคมีตามมา

นอกจากปัจจัยความซับซ้อนในด้านอุปทาน อุตสาหกรรมปิโตรเคมียังเผชิญกับความซับซ้อนในด้านอุปสงค์ เนื่องจากผลิตภัณฑ์ปิโตรเคมี โดยเฉพาะเม็ดพลาสติก ถือเป็นวัตถุดิบสำคัญในการผลิตของหลากหลายอุตสาหกรรม เช่น อุตสาหกรรมพลาสติกและบรรจุภัณฑ์ อุตสาหกรรมก่อสร้าง เป็นต้น ซึ่งในแต่ละอุตสาหกรรมก็จะมีปริมาณความต้องการแตกต่างกันตามช่วงเวลา ซึ่งจะส่งผลให้เกิดความผันผวนและความซับซ้อนในด้านอุปสงค์ ดังนั้น การวางแผน และการพยากรณ์ความต้องการที่แม่นยำ จึงเป็นปัจจัยสำคัญอย่างยิ่งที่จะส่งผลต่อการดำเนินธุรกิจของอุตสาหกรรมปิโตรเคมี

กรณีศึกษาในงานวิจัยนี้เป็นบริษัทผู้ผลิตเม็ดพลาสติกรายหนึ่งในประเทศไทย มีการดำเนินธุรกิจด้านปิโตรเคมีครอบคลุมทั้งกระบวนการ ตั้งแต่การแปรรูปวัตถุดิบปิโตรเลียมมาเป็นผลิตภัณฑ์ปิโตรเคมีขั้นต้น ไปจนถึงการผลิตเม็ดพลาสติกอันเป็นธุรกิจหลักของบริษัท กรณีศึกษาดังกล่าว โดยหนึ่งในผลิตภัณฑ์หลักที่มีความสำคัญต่อการดำเนินธุรกิจคือเม็ดพลาสติกประเภท พอลิเอทิลีน ความหนาแน่นสูง (High-Density Polyethylene: HDPE) ซึ่งถูกนำไปใช้ในอุตสาหกรรมหลากหลายประเภท เช่น อุตสาหกรรมบรรจุภัณฑ์ อุตสาหกรรมการก่อสร้าง และอุตสาหกรรมยานยนต์ เป็นต้น อย่างไรก็ตาม ความผันผวนของราคาวัตถุดิบและลักษณะความต้องการที่เปลี่ยนแปลงตามช่วงเวลา รวมถึงปัจจัยภายนอก เช่น ภาวะเศรษฐกิจ ทำให้การพยากรณ์ความต้องการเม็ดพลาสติกประเภท HDPE มีความท้าทายหากการพยากรณ์ไม่มีประสิทธิภาพ อาจนำไปสู่ปัญหาสินค้าคงคลังส่วนเกินหรือการขาดแคลนสินค้า ซึ่งส่งผลกระทบต่อประสิทธิภาพในการดำเนินธุรกิจ

ดังนั้น งานวิจัยนี้จึงมุ่งเน้นการพัฒนารูปแบบการพยากรณ์ความต้องการเม็ดพลาสติก เพื่อสนับสนุนการวางแผนการผลิต การบริหารสินค้าคงคลัง และการดำเนินงานด้านซัพพลายเชนให้มีความเหมาะสมและมีประสิทธิภาพมากยิ่งขึ้น

2. การทบทวนวรรณกรรม

งานวิจัยที่เกี่ยวข้องกับการพยากรณ์ความต้องการในอุตสาหกรรมปิโตรเคมีและอุตสาหกรรมพลาสติกได้รับการศึกษาอย่างต่อเนื่อง เนื่องจากลักษณะของอุตสาหกรรมที่มีความผันผวนทั้งในด้านอุปสงค์และอุปทาน Udom and Phumchusri [5] ได้ทำการพยากรณ์ยอดขายเม็ดพลาสติกของผู้จัดจำหน่ายรายหนึ่ง โดยเสนอและเปรียบเทียบแบบจำลองเชิงอนุกรมเวลาหลายวิธี พบว่าแบบจำลอง ARIMA ให้ความแม่นยำในการพยากรณ์สูงกว่าวิธีพื้นฐานอื่น ๆ เช่น วิธี Moving Average และวิธี Holt's and Winter Exponential นอกจากนี้ Siregar et al. [6] ได้นำเสนอการประยุกต์ใช้แบบจำลอง ARIMA สำหรับการพยากรณ์ความต้องการวัตถุดิบในการผลิตผลิตภัณฑ์พลาสติกของโรงงานแห่งหนึ่ง ขณะเดียวกัน ในด้านการพยากรณ์ราคา Kantanantha and Pattaraumpornchai [7] ได้ทำการพัฒนาและเปรียบเทียบแบบจำลองการพยากรณ์ราคาของเม็ดพลาสติกประเภทโพลิโพรพิลีน (PP) โดยใช้วิธี Machine Learning ได้แก่ Support Vector Regression (SVR) และ eXtreme Gradient Boosting (XGBoost) รวมถึงโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) ร่วมกับการนำปัจจัยภายนอกมาใช้เป็นตัวแปรในการพยากรณ์ นอกจากนี้ ในบริบทของอุตสาหกรรมเคมี Zhao et al. [8] ได้เสนอการใช้แบบจำลอง SARIMA ในการพยากรณ์ยอดขายรายเดือนของผลิตภัณฑ์เคมี โดยงานวิจัยดังกล่าวยังเสนอแนะว่าการพยากรณ์ในอนาคต ควรพิจารณาการพัฒนาแบบจำลองแบบผสม เพื่อเพิ่มความแม่นยำในการพยากรณ์

ขณะเดียวกัน ยังมีงานวิจัยจำนวนมากในอุตสาหกรรมอื่น ๆ ที่มุ่งเน้นการพัฒนาและเปรียบเทียบแบบจำลองการพยากรณ์ ซึ่งสามารถสะท้อนแนวคิดและวิธีการที่เป็นประโยชน์ต่อการประยุกต์ใช้ในการพยากรณ์ในอุตสาหกรรมปิโตรเคมีได้ Rabbi et al. [9] ได้ทำการประยุกต์ใช้แบบจำลอง ARIMAX เพื่อพยากรณ์ความต้องการใช้ไฟฟ้า โดยมีการรวมปัจจัยภายนอกเข้ามาในแบบจำลอง ผลการศึกษาพบว่าแบบจำลอง ARIMAX สามารถให้ผลการพยากรณ์ที่มีประสิทธิภาพสูงได้ นอกจากนี้ Daramola et al. [10] ได้พัฒนาแบบจำลอง

การพยากรณ์ปริมาณขยะที่ถูกสร้างและถูกรีไซเคิล โดยนำเสนอทั้งแบบจำลองทางสถิติ แบบจำลอง Machine Learning และแบบจำลองไฮบริด เช่น ARIMAX, XGBoost และ ARIMAX-XGBoost ผลการศึกษาพบว่าแบบจำลองไฮบริด ARIMAX-XGBoost ให้ความแม่นยำสูงที่สุด ในส่วนของบริบทการศึกษาด้านทรัพยากรน้ำ George et al. [11] ได้นำเสนอแบบจำลองไฮบริด ARIMA-XGBoost ในการพยากรณ์ระดับน้ำใต้ดิน และทำการเปรียบเทียบแบบจำลองไฮบริดกับแบบจำลองเดี่ยวอื่น ๆ เช่น ARIMA, XGBoost และ Random Forest เป็นต้น ผลการศึกษาพบว่าแบบจำลองไฮบริด ARIMA-XGBoost ให้ค่าความคลาดเคลื่อนต่ำกว่าแบบจำลองเดี่ยว สะท้อนถึงศักยภาพของแบบจำลองไฮบริดในการพยากรณ์ ขณะเดียวกัน Nurhayati et al. [12] ได้พัฒนาแบบจำลองไฮบริด SARIMAX-XGBoost เพื่อเพิ่มความสามารถของแบบจำลองในการอธิบายทั้งความสัมพันธ์เชิงเส้นและไม่เชิงเส้นของข้อมูล สำหรับนำมาใช้ในการพยากรณ์จำนวนผู้ติดเชื้อ HIV รายใหม่ในประเทศอินโดนีเซีย ผลการศึกษาพบว่าแบบจำลองไฮบริด SARIMAX-XGBoost ให้ความแม่นยำสูงและให้ผลลัพธ์ที่ดีกว่าแบบจำลองเดี่ยวอย่างมีนัยสำคัญ ในการพยากรณ์การไหลของการจราจร Işlak et al. [13] ได้นำเสนอและเปรียบเทียบประสิทธิภาพของแบบจำลองไฮบริดกับแบบจำลองดั้งเดิม เช่น Naïve, ARIMA และ XGBoost-ARIMA เป็นต้น ผลการศึกษาพบว่าแบบจำลองไฮบริดให้ผลการพยากรณ์ที่มีความแม่นยำสูงกว่าแบบจำลองดั้งเดิม

โดยสรุป งานวิจัยในอุตสาหกรรมอื่นเหล่านี้แสดงให้เห็นถึงศักยภาพของแบบจำลองไฮบริดในการเพิ่มความแม่นยำของการพยากรณ์ โดยเฉพาะกับในกรณีที่ข้อมูลมีความซับซ้อนสูง ซึ่งสามารถนำแนวคิดดังกล่าวมาประยุกต์ใช้กับการพยากรณ์ความต้องการในอุตสาหกรรมปิโตรเคมีได้อย่างเหมาะสม อย่างไรก็ตาม จากการทบทวนงานวิจัยที่ผ่านมา พบว่าส่วนใหญ่ยังพัฒนาแบบจำลองไฮบริดโดยใช้แบบจำลองเชิงสถิติเป็นแบบจำลองหลัก และใช้แบบจำลองด้าน Machine Learning ในขั้นตอนถัดไปเพื่อพยากรณ์หรือปรับแก้ค่า Residual งานวิจัยนี้จึงมุ่งนำเสนอแนวทางเพิ่มเติม โดยสลับบทบาทของแบบจำลองดังกล่าว ให้แบบจำลอง Machine Learning ทำหน้าที่เป็นแบบจำลอง

หลัก และใช้แบบจำลองเชิงสถิติในการพยากรณ์ค่า Residual ควบคู่กับการนำเสนอเปรียบเทียบกับแบบจำลองไฮบริดในรูปแบบดั้งเดิม และแบบจำลองแบบเดี่ยว

3. ระเบียบวิธีวิจัย

3.1 การรวบรวมข้อมูล

งานวิจัยนี้ได้มีการรวบรวมข้อมูลปริมาณการสั่งซื้อของเม็ดพลาสติกประเภท HDPE จากบริษัทกรณีศึกษาซึ่งเป็นผู้ผลิตในอุตสาหกรรมปิโตรเคมี รวมถึงได้มีการรวบรวมข้อมูลปัจจัยอื่น ๆ ที่อาจส่งผลต่อปริมาณการสั่งซื้อเม็ดพลาสติกในแต่ละเดือน ดังตารางที่ 1 เพื่อสะท้อนสภาพแวดล้อมทางเศรษฐกิจและอุตสาหกรรมที่มีผลต่อความผันผวนของความต้องการสินค้า

ตารางที่ 1 รายละเอียดปัจจัยที่เกี่ยวข้องและแหล่งที่มาของข้อมูล

ตัวแปร	รายละเอียดตัวแปร	แหล่งที่มาของข้อมูล
V1	สัดส่วนยอดขายตามกลุ่มลูกค้า	บริษัทกรณีศึกษา
V2	สัดส่วนยอดขายตามภูมิภาค	บริษัทกรณีศึกษา
V3	ราคาน้ำมันดิบ (Brent)	U.S. Energy Information Administration (EIA)
V4	ดัชนีอุตสาหกรรมของ Polyethylene Resin (PE)	สำนักงานเศรษฐกิจอุตสาหกรรม
V5	ดัชนีอุตสาหกรรมของถุงพลาสติก	สำนักงานเศรษฐกิจอุตสาหกรรม
V6	ดัชนีอุตสาหกรรมของแผ่นฟิล์มพลาสติก	สำนักงานเศรษฐกิจอุตสาหกรรม
V7	ดัชนีอุตสาหกรรมของบรรจุภัณฑ์พลาสติกอื่น ๆ	สำนักงานเศรษฐกิจอุตสาหกรรม
V8	ดัชนีราคาผู้บริโภค (CPI)	สำนักงานนโยบายและยุทธศาสตร์การค้า
V9	ดัชนีราคาผู้ผลิต (PPI)	สำนักงานนโยบายและยุทธศาสตร์การค้า
V10	ปริมาณการส่งออก HDPE	กรมศุลกากร
V11	อัตราแลกเปลี่ยนระหว่างดอลลาร์สหรัฐและบาทไทย	ธนาคารแห่งประเทศไทย

ข้อมูลที่ใช้ในการศึกษาครอบคลุมช่วงระยะเวลาตั้งแต่เดือนมกราคม พ.ศ. 2565 ถึงเดือนกันยายน พ.ศ. 2568 โดยข้อมูลทั้งหมดถูกจัดให้อยู่ในรูปแบบข้อมูลรายเดือน และข้อมูลในช่วงต้นบางส่วนถูกนำไปใช้สำหรับการสร้างตัวแปรตัวแปรเชิงคุณลักษณะ ส่งผลให้ข้อมูลที่สามารถนำมาใช้พัฒนาแบบจำลองมีจำนวนทั้งสิ้น 43 เดือน โดยข้อมูลถูกแบ่งออกเป็นชุดข้อมูลฝึกและชุดข้อมูลทดสอบตามลำดับเวลาดังนี้

- ชุดข้อมูลฝึก ครอบคลุมช่วงระยะเวลาตั้งแต่เดือนมีนาคม พ.ศ. 2565 ถึงเดือนกันยายน พ.ศ. 2567

- ชุดข้อมูลทดสอบ ครอบคลุมช่วงระยะเวลาตั้งแต่เดือนตุลาคม พ.ศ. 2567 ถึงเดือนกันยายน พ.ศ. 2568

3.2 การวิเคราะห์ลักษณะข้อมูล

ในการวิเคราะห์ลักษณะของข้อมูล จะทำการตรวจสอบทั้งแนวโน้มและความเป็นฤดูกาลของข้อมูลปริมาณการสั่งซื้อ โดยในการทดสอบแนวโน้มจะใช้วิธีการทดสอบ Mann-Kendall Trend Test ซึ่งเป็นวิธีการทดสอบทางสถิติแบบไม่อาศัยพารามิเตอร์ที่ใช้เพื่อตรวจสอบว่าข้อมูลอนุกรมเวลาแสดงแนวโน้มการเปลี่ยนแปลงในทิศทางเพิ่มขึ้นหรือลดลงอย่างต่อเนื่องตลอดช่วงเวลาที่พิจารณาหรือไม่ [14] ส่วนการทดสอบความเป็นฤดูกาลจะใช้การทดสอบ Friedman Test ซึ่งอาศัยการจัดอันดับข้อมูล เพื่อตรวจสอบว่าข้อมูลอนุกรมเวลาแสดงลักษณะของคาบเวลาหรือไม่ [15]

3.3 การคัดเลือกตัวแปรโดยใช้ Mutual Information

Mutual Information (MI) เป็นตัวชี้วัดที่ใช้บ่งชี้ระดับความพึ่งพากันระหว่างตัวแปรสุ่มสองตัว โดยสามารถสะท้อนว่าการทราบค่าของตัวแปรหนึ่งสามารถช่วยลดความไม่แน่นอนของอีกตัวแปรหนึ่งได้มากเพียงใด ในงานวิจัยด้านการพยากรณ์อนุกรมเวลาที่มีความซับซ้อนและเกี่ยวข้องกับตัวแปรจำนวนมาก ขั้นตอนการคัดเลือกตัวแปร (Feature Selection) ถือเป็นขั้นตอนที่มีความสำคัญ และพบว่าการใช้ MI ในขั้นตอนการคัดเลือกตัวแปรสามารถช่วยเพิ่มประสิทธิภาพและความแม่นยำของแบบจำลองการพยากรณ์ได้ [16] จากแนวคิดดังกล่าว งานวิจัยนี้จึงนำ MI มาใช้ใน

การคัดเลือกตัวแปรสำหรับนำไปใช้ในการพัฒนาแบบจำลองการพยากรณ์

3.4 Autoregressive Integrated Moving Average (ARIMA)

แบบจำลอง Autoregressive Integrated Moving Average (ARIMA) เป็นแบบจำลองเชิงสถิติที่ใช้กันอย่างแพร่หลายในการพยากรณ์อนุกรมเวลา โดยแบบจำลอง ARIMA ประกอบไปด้วยองค์ประกอบหลักสามส่วน ได้แก่ ส่วน Autoregressive (AR) แทนด้วยพารามิเตอร์ p ซึ่งเป็นค่าข้อมูลในอดีต ส่วน Integrated (I) แทนด้วยพารามิเตอร์ d เกี่ยวข้องกับการแปลงข้อมูลด้วยการทำอนุพันธ์เพื่อให้ข้อมูลมีคุณสมบัติคงที่ และส่วน Moving Average (MA) แทนด้วยพารามิเตอร์ q อธิบายอิทธิพลของค่าความคลาดเคลื่อนจากการพยากรณ์ในอดีต องค์ประกอบทั้งสามส่วนนี้ถูกนำมารวมกันเพื่อสร้างแบบจำลองเชิงเส้นสำหรับการพยากรณ์ [17] การพัฒนาแบบจำลอง ARIMA มักอาศัยแนวทาง Box-Jenkins ซึ่งประกอบด้วยขั้นตอนการระบุแบบจำลอง การประมาณค่า การตรวจสอบความเหมาะสมและการพยากรณ์ [8] ในงานวิจัยนี้ จึงจะมุ่งเน้นพัฒนาแบบจำลอง ARIMA ตามแนวทางของ Box-Jenkins

3.5 Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX)

แบบจำลอง Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) เป็นแบบจำลองที่พัฒนาต่อยอดจากแบบจำลอง ARIMA โดยมีการนำตัวแปรภายนอกเข้ามาเป็นส่วนหนึ่งในการวิเคราะห์ ทำให้แบบจำลองสามารถพิจารณาอิทธิพลของปัจจัยดังกล่าวต่อการเปลี่ยนแปลงของข้อมูลได้ [18] ในงานวิจัยนี้ การพัฒนาแบบจำลอง ARIMAX ดำเนินการตามแนวทางเดียวกับการสร้างแบบจำลอง ARIMA ตามกระบวนการ Box-Jenkins และมีการเพิ่มตัวแปรที่เกี่ยวข้องเข้ามาในขั้นตอนการสร้างแบบจำลอง

3.6 eXtreme Gradient Boosting (XGBoost)

แบบจำลอง eXtreme Gradient Boosting (XGBoost) เป็นอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning)

ที่พัฒนาขึ้นจากแนวคิดของ Gradient Boosting Decision Tree ซึ่งได้รับการออกแบบให้มีประสิทธิภาพสูงทั้งในด้านความแม่นยำและความเสถียร โดยแบบจำลอง XGBoost จัดอยู่ในกลุ่มอัลกอริทึมแบบ Tree-based มีกระบวนการเรียนรู้ที่ดำเนินไปในลักษณะเป็นลำดับขั้น คือ แบบจำลองในแต่ละรอบจะถูกสร้างขึ้นเพื่อปรับปรุงและลดข้อผิดพลาดจากแบบจำลองในรอบก่อนหน้า ส่งผลให้ความสามารถในการพยากรณ์ของแบบจำลองเพิ่มขึ้นอย่างต่อเนื่อง [7]

สำหรับแบบจำลอง XGBoost งานวิจัยนี้ใช้ชุดตัวแปรที่ได้จากการพิจารณาค่า MI ในการพยากรณ์ และทำการปรับจูนไฮเปอร์พารามิเตอร์ด้วยวิธี Grid Search ดังแสดงในตารางที่ 2 โดยการกำหนดช่วงค่าของไฮเปอร์พารามิเตอร์ดังกล่าวพิจารณาจากการทดลองเบื้องต้น เพื่อระบุขอบเขตที่ให้ประสิทธิภาพการพยากรณ์เหมาะสมภายใต้ลักษณะและขนาดของชุดข้อมูลที่ศึกษา อีกทั้งยังเป็นการควบคุมความซับซ้อนของแบบจำลองเพื่อลดความเสี่ยงของการเกิดปัญหา Overfitting งานวิจัยนี้ดำเนินการปรับจูนร่วมกับการทำ 5-fold cross-validation ภายใต้การแบ่งข้อมูลแบบ TimeSeriesSplit เพื่อคงลำดับเชิงเวลาของข้อมูล และคัดเลือกชุดไฮเปอร์พารามิเตอร์ที่เหมาะสมที่สุดสำหรับการสร้างแบบจำลอง

ตารางที่ 2 การกำหนดช่วงค่าไฮเปอร์พารามิเตอร์ของแบบจำลอง XGBoost

ไฮเปอร์พารามิเตอร์	ค่า
n_estimators	150, 250
max_depth	2, 3
learning_rate	0.03, 0.05
subsample	0.6, 0.8
colsample_bytree	0.5, 0.7
min_child_weight	5, 7, 9
reg_lambda	5, 10, 20
reg_alpha	0, 0.1, 1.0

3.7 แบบจำลองไฮบริด (Hybrid Model)

แบบจำลองไฮบริด เป็นแนวคิดที่มุ่งเน้นการแยกบทบาทของการพยากรณ์ออกเป็นสองส่วน โดยเริ่มจากการ

ใช้แบบจำลองหลักสร้างค่าพยากรณ์เบื้องต้น จากนั้นจะนำค่าความคลาดเคลื่อนระหว่างค่าจริงและค่าพยากรณ์ไปสร้างแบบจำลองเพิ่มเติมเพื่อเรียนรู้รูปแบบที่แบบจำลองหลักไม่สามารถอธิบายได้ และผลการพยากรณ์ขั้นสุดท้ายจะได้จากการรวมค่าพยากรณ์ของทั้งสองส่วนเข้าด้วยกัน การใช้แบบจำลองไฮบริดช่วยเพิ่มความสามารถในการอธิบายข้อมูลอนุกรมเวลาที่มีลักษณะพฤติกรรมทั้งในส่วนที่เป็นเชิงเส้นและไม่เป็นเชิงเส้นได้ [13]

ในงานวิจัยนี้ แนวคิดของแบบจำลองไฮบริดถูกนำมาประยุกต์ใช้ โดยนำแบบจำลองเดี่ยว ได้แก่ ARIMA, ARIMAX และ XGBoost ที่ได้พัฒนาขึ้นก่อนหน้านี้ มาสร้างเป็นแบบจำลองไฮบริดในรูปแบบต่าง ๆ โดยสำหรับแบบจำลองหลัก จะใช้ค่าพารามิเตอร์ที่ได้รับการคัดเลือกแล้วว่าเป็นค่าที่เหมาะสมที่สุดจากการพัฒนาแบบจำลองเดี่ยวนำมาใช้ต่อในแบบจำลองไฮบริด และสำหรับแบบจำลองที่ใช้พยากรณ์ค่า Residual จะมีการปรับจูนค่าพารามิเตอร์ด้วยวิธี Grid Search เพื่อทำการคัดเลือกค่าที่เหมาะสมที่สุด ซึ่งในงานวิจัยนี้ ได้ทำการพัฒนาแบบจำลองไฮบริดทั้งหมด 4 แบบจำลอง ดังนี้

3.7.1 ARIMA-XGBoost

แบบจำลองไฮบริด ARIMA-XGBoost ถูกพัฒนาโดยกำหนดให้ ARIMA ทำหน้าที่เป็นแบบจำลองหลักในการพยากรณ์ค่าความต้องการจากข้อมูลอนุกรมเวลาแบบตัวแปรเดียว จากนั้นนำค่าความคลาดเคลื่อนระหว่างค่าจริงและค่าที่ ARIMA พยากรณ์ได้ไปใช้เป็นข้อมูลสำหรับแบบจำลอง XGBoost เพื่อพยากรณ์ค่า Residual ที่ ARIMA ไม่สามารถอธิบายได้ โดยแบบจำลอง XGBoost ได้รับการปรับจูนไฮเปอร์พารามิเตอร์ด้วยวิธี Grid Search และใช้ช่วงค่าเดียวกับแบบจำลอง XGBoost เดี่ยว ดังแสดงในตารางที่ 2 โดยผลการพยากรณ์ขั้นสุดท้ายได้จากการรวมค่าพยากรณ์ของ ARIMA และค่าพยากรณ์ Residual จาก XGBoost เข้าด้วยกัน

3.7.2 ARIMAX-XGBoost

แบบจำลองไฮบริด ARIMAX-XGBoost ใช้แนวคิดเดียวกับ ARIMA-XGBoost โดยเปลี่ยนแบบจำลองหลักจาก

ARIMA เป็น ARIMAX เพื่อพยากรณ์ข้อมูล จากนั้นใช้ XGBoost ซึ่งได้รับการปรับจูนไฮเปอร์พารามิเตอร์ด้วย Grid Search ตามช่วงค่าที่กำหนดไว้ในตารางที่ 2 ในการพยากรณ์ค่าความคลาดเคลื่อนจากผลการพยากรณ์ของ ARIMAX ก่อนนำผลการพยากรณ์มารวมกันเป็นค่าพยากรณ์ขั้นสุดท้าย

3.7.3 XGBoost-ARIMA

แบบจำลองไฮบริด XGBoost-ARIMA ถูกพัฒนาภายใต้แนวคิดตรงกันข้ามกับแบบจำลองไฮบริดก่อนหน้านี้ โดยกำหนดให้ XGBoost ทำหน้าที่เป็นแบบจำลองหลักสำหรับการพยากรณ์ค่าความต้องการจากข้อมูล โดยอาศัยความสามารถในการเรียนรู้ความสัมพันธ์ที่ไม่เป็นเชิงเส้นของตัวแปร จากนั้นนำค่าความคลาดเคลื่อนที่เกิดจากการพยากรณ์ของ XGBoost ซึ่งสะท้อนส่วนที่แบบจำลองหลักยังไม่สามารถอธิบายได้ มาใช้เป็นข้อมูลสำหรับแบบจำลอง ARIMA เพื่อพยากรณ์ค่า Residual และอธิบายโครงสร้างเชิงเวลาที่หลงเหลืออยู่ในข้อมูล โดยแบบจำลอง ARIMA จะถูกกำหนดโครงสร้างตามขั้นตอนของ Box-Jenkins และปรับจูนพารามิเตอร์ของแบบจำลองด้วยวิธี Grid Search สุดท้ายนำผลการพยากรณ์จากทั้งสองแบบจำลองมารวมกันเป็นค่าพยากรณ์ขั้นสุดท้าย

3.7.4 XGBoost-ARIMAX

แบบจำลองไฮบริด XGBoost-ARIMAX พัฒนาภายใต้แนวคิดเดียวกับ XGBoost-ARIMA โดยกำหนดให้ XGBoost ทำหน้าที่เป็นแบบจำลองหลัก จากนั้นนำค่าความคลาดเคลื่อนจากผลการพยากรณ์ของ XGBoost มาใช้เป็นข้อมูลสำหรับแบบจำลอง ARIMAX ซึ่งกำหนดโครงสร้างแบบจำลองตามแนวทาง Box-Jenkins และปรับจูนพารามิเตอร์ด้วยวิธี Grid Search ก่อนสรุปผลการพยากรณ์โดยนำผลลัพธ์จากแบบจำลองทั้งสองรวมกัน

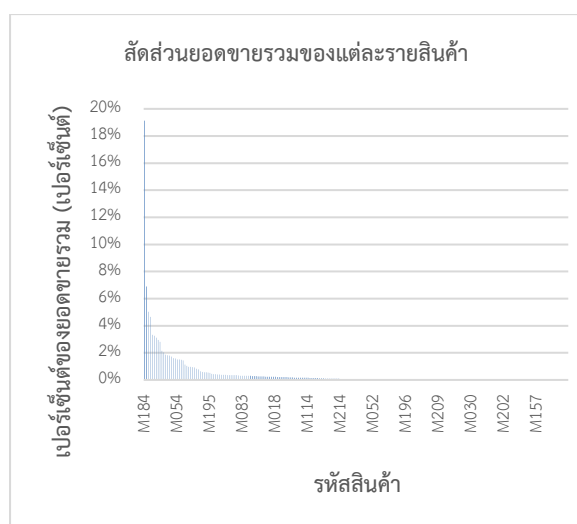
3.8 เกณฑ์การประเมินประสิทธิภาพของแบบจำลอง

ในการประเมินประสิทธิภาพของแบบจำลอง ใช้ตัวชี้วัด Mean Absolute Percentage Error (MAPE) เป็นตัวชี้วัดหลัก เนื่องจากสามารถแสดงความคลาดเคลื่อนของการพยากรณ์ในรูปของร้อยละ ทำให้ตีความผลลัพธ์ได้ง่าย

4. ผลการวิจัยและอภิปรายผล

4.1 ผลการคัดเลือกสินค้ากรณีศึกษา

เนื่องจากผลิตภัณฑ์ของบริษัทกรณีศึกษามีความหลากหลายสูงถึง 219 รายการ โดยสินค้าแต่ละรายการมีลักษณะโครงสร้างข้อมูลที่แตกต่างกัน ทั้งในด้านฐานลูกค้าและสัดส่วนการจำหน่ายทั้งในและต่างประเทศ ซึ่งส่งผลให้ปัจจัยภายนอกที่ส่งผลกระทบต่อความต้องการสินค้าแต่ละชนิดมีความแตกต่างกัน การพัฒนาแบบจำลองโดยใช้ชุดปัจจัยภายนอกเดียวกันกับสินค้าหลายรายการจึงอาจไม่สามารถสะท้อนความสัมพันธ์ที่แท้จริงของข้อมูลได้ เพื่อให้การวิเคราะห์ปัจจัยภายนอกและการพัฒนาแบบจำลองมีความแม่นยำและสอดคล้องกับลักษณะข้อมูล งานวิจัยนี้จึงใช้วิธีการเลือกตัวอย่างแบบเจาะจงโดยพิจารณาจากสินค้าที่มียอดขายรวมในช่วงระยะเวลาที่ศึกษาสูงสุด จากผลการวิเคราะห์พบว่า สินค้าที่มียอดขายรวมสูงสุด มีสัดส่วนยอดขายคิดเป็นประมาณร้อยละ 19 ของยอดขายรวมทั้งหมด โดยเมื่อเปรียบเทียบกับสินค้าอื่น ๆ ที่มียอดขายรวมเป็นอันดับรองลงมา พบว่าสินค้านี้มียอดขายรวมที่มากกว่าอย่างชัดเจนดังรูปที่ 1 ซึ่งสะท้อนถึงความสำคัญของสินค้านี้ต่อการประสิทธิภาพการดำเนินงานของบริษัท



รูปที่ 1 สัดส่วนยอดขายรวมของแต่ละรายสินค้า

ดังนั้น งานวิจัยนี้จึงเลือกสินค้านี้รหัส M184 ซึ่งเป็นสินค้าที่มียอดขายรวมสูงสุดเป็นกรณีศึกษา เพื่อใช้ในการ

พัฒนาแบบจำลองพยากรณ์ในระดับรายสินค้า โดยการมุ่งเน้นการศึกษาเชิงลึกกับสินค้าหนึ่งรายการช่วยให้สามารถวิเคราะห์และเปรียบเทียบแบบจำลองได้อย่างละเอียด ทั้งนี้กรอบวิธีการที่พัฒนาสามารถนำไปประยุกต์ใช้กับสินค้าอื่นได้ในการพัฒนาแบบจำลองในอนาคต

4.2 ผลการวิเคราะห์ลักษณะข้อมูล

จากการวิเคราะห์ข้อมูลปริมาณการสั่งซื้อรายเดือนด้วยวิธี Mann-Kendall Trend Test พบว่าค่า P-value เท่ากับ 0.18 ซึ่งมากกว่าระดับนัยสำคัญทางสถิติ ภายใต้ปริมาณข้อมูลที่ใช้ในการศึกษานี้ จึงยังไม่มีหลักฐานเพียงพอที่จะปฏิเสธสมมติฐานหลัก กล่าวคือ ข้อมูลที่มีอยู่ยังไม่แสดงให้เห็นแนวโน้มการเปลี่ยนแปลงในทิศทางเพิ่มขึ้นหรือลดลงอย่างมีนัยสำคัญทางสถิติ ในส่วนของการตรวจสอบความเป็นฤดูกาลด้วยวิธี Friedman Test พบว่าค่า P-value เท่ากับ 0.84 แสดงให้เห็นว่าภายใต้ขนาดตัวอย่างที่ใช้ในการศึกษานี้ ยังไม่มีหลักฐานเพียงพอที่จะปฏิเสธสมมติฐานหลัก กล่าวคือ ข้อมูลที่มีอยู่ยังไม่แสดงให้เห็นรูปแบบความเป็นฤดูกาลอย่างมีนัยสำคัญทางสถิติ

4.3 ผลการคัดเลือกตัวแปรโดยใช้ Mutual Information

จากกระบวนการคัดเลือกตัวแปรด้วย Mutual Information (MI) ได้ทำการรวบรวมชุดตัวแปรตั้งต้นทั้งหมดจำนวน 83 ตัวแปรที่คาดว่าจะมีความสัมพันธ์กับปริมาณการสั่งซื้อ ซึ่งได้มาจากตัวแปรที่ผ่านการสร้างคุณลักษณะ โดยประกอบด้วยตัวแปรช่วงเวลา (Lag Variables) จำนวน 50 ตัว, ตัวแปรค่าเฉลี่ยเคลื่อนที่ (Rolling Variables) จำนวน 21 ตัว, ตัวแปรความผันผวน (Volatility Variables) จำนวน 4 ตัว, ตัวแปรเชิงฤดูกาล (Seasonal Variables) จำนวน 6 ตัว, และตัวแปรปฏิสัมพันธ์ระหว่างปัจจัยภายนอก (Interaction Variables) จำนวน 2 ตัว โดยสามารถแสดงตัวแปร 15 ลำดับแรกจากการพิจารณาตามค่า MI ได้ดังตารางที่ 3

เพื่อควบคุมความซับซ้อนของแบบจำลองและลดความเสี่ยงของปัญหา Underfitting และ Overfitting งานวิจัยนี้จึงเลือกใช้ตัวแปร 5 ลำดับแรกจากการจัดลำดับตามค่า MI สำหรับนำไปพัฒนาแบบจำลอง XGBoost โดย

ตัวแปรที่ถูกเลือกประกอบด้วย Export Qty_lag2 และ Export Qty_lag1 ซึ่งแทนปริมาณการส่งออก HDPE ในช่วง 2 เดือนและ 1 เดือนก่อนหน้าตามลำดับ, PE IPI_rol3 และ Package IPI_lag2 สะท้อนระดับกิจกรรมของอุตสาหกรรม PE และอุตสาหกรรมของบรรจุภัณฑ์พลาสติกอื่น ๆ ตามลำดับ และ Oil Price_rol3 แทนราคาน้ำมันเฉลี่ยในช่วง 3 เดือนก่อนหน้า

ทั้งนี้ ตัวแปร 5 ลำดับแรกที่ถูกคัดเลือกจาก MI จะถูกนำมาใช้เฉพาะกับแบบจำลอง XGBoost เท่านั้น สำหรับแบบจำลอง ARIMAX การเลือกตัวแปรจะพิจารณาร่วมกันระหว่างผลการจัดลำดับด้วย MI และความเหมาะสมตามโครงสร้างของแบบจำลองเชิงเส้น

ตารางที่ 3 ตัวแปร 15 อันดับแรกตามค่า MI

ลำดับ	ตัวแปร	MI
1	Export Qty_lag2	0.29
2	PE IPI_rol3	0.22
3	Package IPI_lag2	0.19
4	Oil Price_rol3	0.19
5	Export Qty_lag1	0.17
6	PlasticBag IPI_rol3	0.17
7	Oil Price_lag2	0.15
8	Export Qty_rol3	0.15
9	Oil_x_PE	0.15
10	Qty_lag1	0.15
11	Qty_rol3	0.14
12	PlasticBag IPI_lag1	0.14
13	PE IPI_lag2	0.13
14	Film IPI_lag1	0.13
15	Film IPI_lag2	0.13

4.4 Autoregressive Integrated Moving Average (ARIMA)

ในการสร้างแบบจำลอง ARIMA ได้ดำเนินการตามแนวทาง Box-Jenkins โดยผลลัพธ์สามารถแสดงได้ดังนี้

1) การระบุรูปแบบของแบบจำลอง (Identification) เนื่องจากแบบจำลอง ARIMA ต้องการข้อมูลอนุกรมเวลาที่มีคุณสมบัติคงที่ (Stationary) จึงได้มีการตรวจสอบความคงที่

ของข้อมูล โดยใช้การทดสอบ Augmented Dickey-Fuller (ADF) และ Kwiatkowski-Phillips-Schmidt-Shin (KPSS) ซึ่งเป็นการทดสอบทางสถิติที่ใช้เพื่อตรวจสอบและพิจารณาความคงที่ของข้อมูลอนุกรมเวลา การทดสอบดำเนินการทั้งในรูปแบบข้อมูลดิบและข้อมูลที่ผ่านการหาความแตกต่างลำดับที่หนึ่ง (First Differencing) เพื่อใช้ประกอบการพิจารณาค่าพารามิเตอร์ d ที่เหมาะสม ผลลัพธ์พบว่าข้อมูลทั้งสองกรณีมีคุณสมบัติที่คงที่ จึงพิจารณาค่า d ให้อยู่ที่ 0 และ 1 สำหรับการพิจารณาค่าพารามิเตอร์ p และ q ได้มีการสร้างและพิจารณาจากกราฟ Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) ผลจากกราฟพบว่าค่า p และ q มีแนวโน้มอยู่ในระดับต่ำ จึงทำการกำหนดช่วงค่าพารามิเตอร์ที่เหมาะสมดังตารางที่ 4

ตารางที่ 4 ค่าพารามิเตอร์ที่เป็นไปได้สำหรับแบบจำลอง ARIMA

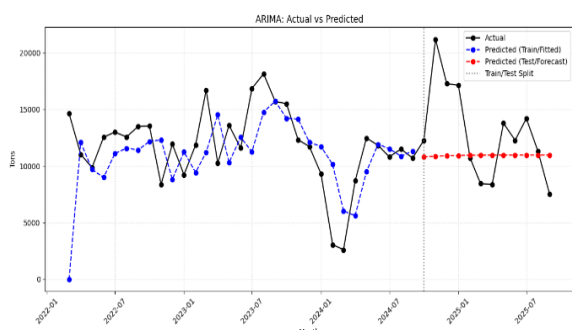
พารามิเตอร์	ค่า
p	0, 1, 2
d	0, 1
q	0, 1, 2

2) การประมาณค่า (Estimation) หลังจากกำหนดช่วงค่าที่เป็นไปได้ของพารามิเตอร์ p , d และ q แล้ว ได้มีการใช้วิธี Grid Search ภายในช่วงค่าที่กำหนดเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลอง โดยพิจารณาจากค่า Akaike Information Criterion (AIC) และ Bayesian Information Criterion (BIC) ผลการประเมินพบว่าแบบจำลอง ARIMA(1,1,1) ให้ค่า AIC และ BIC ต่ำที่สุดเมื่อเปรียบเทียบกับแบบจำลองอื่น ดังนั้นแบบจำลอง ARIMA(1,1,1) จึงถูกเลือกเป็นแบบจำลองที่เหมาะสมที่สุด

3) การตรวจสอบความเหมาะสมของแบบจำลอง (Diagnostic Checks) ทำการพิจารณากราฟของ Residuals, กราฟ Histogram แสดงการแจกแจงของ Residuals, กราฟ Normal Q-Q plot และกราฟ ACF ของ Residuals ผลการตรวจสอบพบว่ากราฟทั้งหมดแสดงลักษณะที่เหมาะสม นอกจากนี้ ผลการทดสอบ Ljung-Box

(Lag 20) ให้ค่า P-value เท่ากับ 0.99 ซึ่งภายใต้ขนาดตัวอย่างที่ใช้ในการศึกษานี้ ยังไม่พบหลักฐานเพียงพอในการปฏิเสธสมมติฐานหลักว่าเศษเหลือมีความเป็นอิสระอย่างมีนัยสำคัญ จึงสามารถสรุปได้ว่าแบบจำลอง ARIMA(1,1,1) มีความเหมาะสมในการนำไปใช้พยากรณ์

4) การพยากรณ์ (Forecasting) นำแบบจำลอง ARIMA(1,1,1) ไปใช้ในการพยากรณ์ความต้องการเม็ดพลาสติกรายเดือน พบว่าตัวชี้วัด MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 30.78 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบเท่ากับ 24.94 เปอร์เซ็นต์ โดยสามารถเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์ได้ดังรูปที่ 2 เมื่อพิจารณาจากทั้งกราฟและค่าตัวชี้วัด MAPE พบว่าแบบจำลอง ARIMA ยังมีประสิทธิภาพในการเรียนรู้รูปแบบของข้อมูลอยู่ในระดับจำกัด ทั้งในช่วงชุดข้อมูลฝึกและชุดข้อมูลทดสอบ โดยค่าพยากรณ์ในชุดข้อมูลทดสอบมีลักษณะค่อนข้างเรียบและไม่สามารถติดตามความผันผวนของค่าจริงได้อย่างเหมาะสมสะท้อนถึงข้อจำกัดของแบบจำลอง ARIMA ในการพยากรณ์ชุดข้อมูลที่มีความผันผวนสูงและมีพฤติกรรมไม่เป็นเชิงเส้น



รูปที่ 2 ผลการพยากรณ์ของแบบจำลอง ARIMA

4.5 Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX)

ในการพัฒนาแบบจำลอง ARIMAX ได้มีการคัดเลือกตัวแปรเพื่อนำมาใช้ประกอบในแบบจำลอง โดยพิจารณาจากผลการวิเคราะห์ความสำคัญของตัวแปรด้วย MI ดังแสดงในตารางที่ 3 ร่วมกับเหมาะสมของโครงสร้างแบบจำลองเชิงเส้น ผลการวิเคราะห์พบว่าตัวแปรจากข้อมูลในอดีตของปริมาณความต้องการ ได้แก่ Qty_lag1 ซึ่งสะท้อนปริมาณการสั่งซื้อในช่วงเดือนก่อนหน้า และ

Qty_roll3 ซึ่งแทนค่าเฉลี่ยของปริมาณการสั่งซื้อย้อนหลัง 3 เดือน ปรากฏอยู่ในกลุ่มตัวแปรที่มีค่า MI ค่อนข้างสูง แสดงให้เห็นว่าตัวแปรดังกล่าวมีความสัมพันธ์กับปริมาณความต้องการในระดับหนึ่ง

อีกทั้งเมื่อพิจารณาควบคู่กับลักษณะของแบบจำลอง ARIMAX ซึ่งเป็นแบบจำลองเชิงเส้น จึงเลือกใช้ตัวแปรทั้งสองเป็นตัวแปรประกอบในแบบจำลอง เนื่องจากสามารถสะท้อนโครงสร้างของข้อมูลอนุกรมเวลาได้โดยตรง ทั้งนี้ ได้มีการตรวจสอบค่า Variance Inflation Factor (VIF) ของตัวแปรที่เลือกใช้ ซึ่งพบว่ามีค่าเท่ากับ 2.43 อยู่ในระดับที่ยอมรับได้ แสดงให้เห็นว่าสามารถใช้ตัวแปรทั้งสองร่วมกันได้โดยไม่ก่อให้เกิดปัญหา Multicollinearity ที่มีนัยสำคัญ

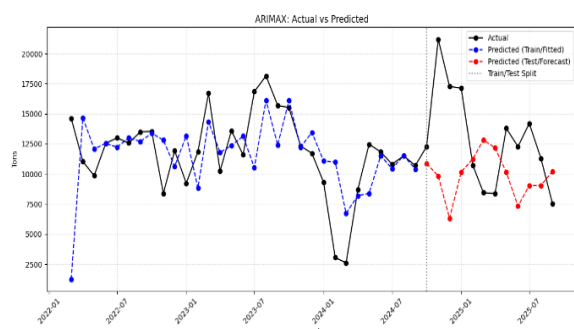
หลังจากคัดเลือกตัวแปรที่เหมาะสมแล้ว การพัฒนาแบบจำลอง ARIMAX ดำเนินการตามขั้นตอน Box-Jenkins ดังนี้

1) การระบุรูปแบบของแบบจำลอง (Identification) เนื่องจากการพัฒนาแบบจำลอง ARIMAX ดำเนินการภายใต้โครงสร้างเดียวกับแบบจำลอง ARIMA ขั้นตอนการตรวจสอบคุณสมบัติของข้อมูลและการกำหนดพารามิเตอร์จึงมีลักษณะเดียวกัน ดังตารางที่ 4

2) การประมาณค่า (Estimation) ในการเลือกค่าพารามิเตอร์ของแบบจำลองใช้วิธี Grid Search โดยพิจารณาจากค่า AIC และ BIC ซึ่งพบว่าแบบจำลอง ARIMAX(0,1,2) ให้ผลลัพธ์ที่ดีที่สุด

3) การตรวจสอบความเหมาะสมของแบบจำลอง (Diagnostic Checks) ดำเนินการโดยวิเคราะห์กราฟ Residuals, กราฟ Histogram แสดงการแจกแจงของ Residuals, กราฟ Normal Q-Q plot และกราฟ ACF ของ Residuals ซึ่งผลการตรวจสอบแสดงลักษณะที่เหมาะสมสอดคล้องกับผลการทดสอบ Ljung-Box (Lag 20) ที่ให้ค่า P-value เท่ากับ 0.99 โดยภายใต้ขนาดตัวอย่างที่ใช้ในการศึกษานี้ ยังไม่พบหลักฐานเพียงพอในการปฏิเสธสมมติฐานหลักด้านความเป็นอิสระของเศษเหลือ จึงสะท้อนว่าแบบจำลอง ARIMAX(0,1,2) มีความเหมาะสมในการพยากรณ์

4) การพยากรณ์ (Forecasting) นำแบบจำลองที่เหมาะสม ได้แก่ ARIMAX(0,1,2) ไปใช้ในการพยากรณ์พบว่าค่า MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 29.67 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบมีค่าเท่ากับ 35.86 เปอร์เซ็นต์ ผลการเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์แสดงดังรูปที่ 3 เมื่อพิจารณาจากผลลัพธ์ พบว่าค่าพยากรณ์ยังคงมีความคลาดเคลื่อนสูงเมื่อเปรียบเทียบกับค่าจริง และไม่สามารถสะท้อนความผันผวนของข้อมูลได้สะท้อนให้เห็นว่า การเพิ่มตัวแปรประกอบจากข้อมูลในอดีตของปริมาณความต้องการช่วยให้แบบจำลองมีความยืดหยุ่นของค่าพยากรณ์เพิ่มขึ้น และไม่แสดงลักษณะเรียบเช่นเดียวกับแบบจำลอง ARIMA ได้ แต่อย่างไรก็ตามแบบจำลอง ARIMAX ยังคงไม่สามารถรองรับลักษณะการเปลี่ยนแปลงของข้อมูลที่มีความผันผวนสูงและมีพฤติกรรมไม่เป็นเชิงเส้นได้อย่างเหมาะสม



รูปที่ 3 ผลการพยากรณ์ของแบบจำลอง ARIMAX

4.6 eXtreme Gradient Boosting (XGBoost)

ในการพัฒนาแบบจำลอง XGBoost ได้มีการปรับจูนไฮเปอร์พารามิเตอร์ด้วยวิธี Grid Search เพื่อคัดเลือค่าไฮเปอร์พารามิเตอร์ที่เหมาะสมที่สุด โดยค่าที่ได้จากการปรับจูนแสดงดังตารางที่ 5

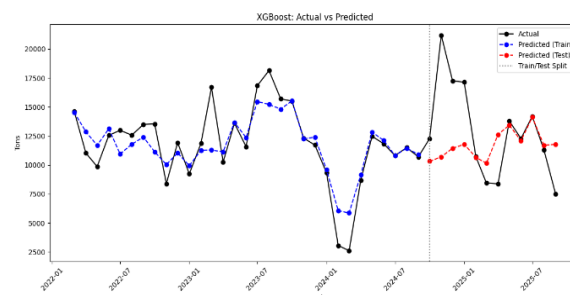
ตารางที่ 5 ผลการปรับจูนไฮเปอร์พารามิเตอร์ของแบบจำลอง XGBoost

ไฮเปอร์พารามิเตอร์	ค่า
n_estimators	250
max_depth	2
learning_rate	0.05

ตารางที่ 5 (ต่อ)

ไฮเปอร์พารามิเตอร์	ค่า
subsample	0.8
colsample_bytree	0.5
min_child_weight	7
reg_lambda	5
reg_alpha	1.0

ผลการประเมินพบว่าค่า MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 14.57 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบเท่ากับ 22.24 เปอร์เซ็นต์ โดยผลการเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์แสดงดังรูปที่ 4 เมื่อพิจารณาจากกราฟและค่า MAPE พบว่าแบบจำลอง XGBoost โดยภาพรวมให้ผลการพยากรณ์ที่มีทิศทางสอดคล้องกับค่าจริง และสามารถอธิบายลักษณะการเปลี่ยนแปลงของข้อมูลได้ในระดับที่ดีกว่าแบบจำลองเชิงสถิติ อย่างไรก็ตาม ในบางช่วงที่ค่าความต้องการมีการเปลี่ยนแปลงอย่างรวดเร็วและเกิดขึ้นในระยะเวลายาวแบบจำลอง XGBoost ยังมีข้อจำกัดในการปรับตัวให้ทันต่อการเปลี่ยนแปลงดังกล่าว ส่งผลให้ค่าพยากรณ์ในช่วงยังไม่สามารถปรับให้ถึงระดับของการเปลี่ยนแปลงที่เกิดขึ้นจริงได้



รูปที่ 4 ผลการพยากรณ์ของแบบจำลอง XGBoost

4.7 แบบจำลองไฮบริด (Hybrid Model)

สำหรับการพัฒนาแบบจำลองไฮบริด ได้ใช้แบบจำลองหลักต่อยอดจากแบบจำลองเดี่ยวที่พัฒนาขึ้นก่อนหน้านี้ ขณะที่แบบจำลองที่ใช้สำหรับพยากรณ์ค่า Residual ได้มีการปรับจูนพารามิเตอร์ใหม่ด้วยวิธี Grid Search โดยผลการคัดเลือค่าพารามิเตอร์ที่เหมาะสมที่สุดแสดงดังตารางที่ 6

ตารางที่ 6 ผลการปรับปรุงพารามิเตอร์ของแบบจำลองไฮบริดในส่วน Residual ของแบบจำลอง

แบบจำลอง	พารามิเตอร์	ค่า
ARIMA-XGBoost	n_estimators	150
	max_depth	2
	learning_rate	0.05
	subsample	0.8
	colsample_bytree	0.5
	min_child_weight	9
	reg_lambda	5
	reg_alpha	0
ARIMAX-XGBoost	n_estimators	150
	max_depth	2
	learning_rate	0.05
	subsample	0.8
	colsample_bytree	0.7
	min_child_weight	9
	reg_lambda	10
	reg_alpha	1.0
XGBoost-ARIMA	(p,d,q)	(0,1,1)
XGBoost-ARIMAX	(p,d,q)	(2,1,2)

ทั้งนี้ เพื่อประเมินว่า Residual ที่เหลืออยู่จากการพยากรณ์ด้วยแบบจำลองหลักยังคงมีโครงสร้างเชิงเวลาในรูปของแนวโน้มและความเป็นฤดูกาลหลงเหลืออยู่หรือไม่ ก่อนนำไปใช้ในขั้นตอนการพัฒนาแบบจำลองไฮบริด จึงทำการตรวจสอบลักษณะของ Residual ดังกล่าวด้วยวิธีการ Mann-Kendall Trend Test เพื่อทำการตรวจสอบแนวโน้มและใช้กราฟ ACF ในการวิเคราะห์โครงสร้างความสัมพันธ์ของข้อมูลตามช่วงเวลา และตรวจสอบรูปแบบการเกิดซ้ำในลักษณะเชิงคาบ เช่น รอบระยะเวลา 12 เดือน เพื่อใช้ประเมินความเป็นฤดูกาล ทั้งนี้เนื่องจาก Residual ที่นำมาตรวจสอบมาจากชุดข้อมูลฝึกเพียงอย่างเดียว ซึ่งมีจำนวนรอบปีไม่เพียงพอสำหรับการทดสอบ Friedman จึงใช้การวิเคราะห์กราฟ ACF แทน ผลจากการทดสอบ Mann-Kendall Trend Test พบว่า ค่า P-value ของ Residual จากแบบจำลองหลักในทุกแบบจำลองไฮบริดมีค่ามากกว่า 0.05 ดังตารางที่ 7 ซึ่งภายใต้ปริมาณข้อมูลที่ใช้ในการศึกษา นี้ แสดงให้เห็นว่ายังไม่มีหลักฐานเพียงพอที่จะปฏิเสธสมมติฐานหลัก กล่าวคือ ข้อมูลตัวอย่างที่มีอยู่ยังไม่แสดงให้เห็น

เห็นอย่างมีนัยสำคัญทางสถิติถึงความขัดแย้งกับสมมติฐานหลักที่ว่า Residual ไม่มีแนวโน้มหลงเหลืออยู่ และจากการวิเคราะห์กราฟ ACF ไม่พบรูปแบบการเกิดซ้ำในลักษณะเชิงคาบอย่างเด่นชัด จึงบ่งชี้ว่า Residual ไม่ปรากฏลักษณะความเป็นฤดูกาล

นอกจากนี้ เพื่อตรวจสอบความสอดคล้องกับสมมติฐานเบื้องต้นก่อนการพัฒนาแบบจำลองไฮบริด ได้มีการวิเคราะห์ Residual สำหรับแบบจำลองหลักทุกประเภทด้วยกราฟ ACF, การทดสอบ Ljung-Box (Lag 10) และการทดสอบ ADF ดังตารางที่ 7 โดยผลการวิเคราะห์พบว่า ผลการทดสอบ Ljung-Box ไม่สามารถปฏิเสธสมมติฐานหลักด้านความเป็นอิสระของเศษเหลือได้อย่างมีนัยสำคัญ ซึ่งสอดคล้องกับลักษณะของกราฟ ACF ขณะเดียวกัน ผลการทดสอบ ADF แสดงให้เห็นว่า Residual มีคุณสมบัติที่คงที่ จึงสามารถสรุปได้ว่า Residual มีคุณสมบัติทางสถิติที่เหมาะสมสำหรับการนำไปใช้ในขั้นตอนถัดไปของการพัฒนาแบบจำลองไฮบริด

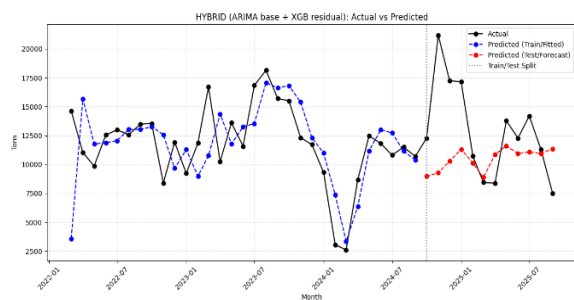
ตารางที่ 7 ค่า P-value ของแต่ละการทดสอบ

แบบจำลองหลัก	ค่า P-value ของแต่ละการทดสอบ		
	Mann-Kendall Trend Test	Ljung-Box (Lag 10)	ADF Test
ARIMA	0.15	0.98	0.00
ARIMAX	0.97	0.92	0.00
XGBoost	0.56	0.77	0.00

4.7.1 ARIMA-XGBoost

จากการประเมินประสิทธิภาพการพยากรณ์ด้วยค่า MAPE พบว่าแบบจำลองไฮบริด ARIMA-XGBoost ให้ค่า MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 22.74 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบเท่ากับ 25.05 เปอร์เซ็นต์ โดยผลการเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์แสดงดังรูปที่ 5 เมื่อพิจารณาจากผลลัพธ์ พบว่าแม้ค่า MAPE ในช่วงชุดข้อมูลทดสอบจะสูงกว่าแบบจำลอง ARIMA เล็กน้อย แต่จากกราฟ แบบจำลองไฮบริดสามารถแสดงการเปลี่ยนแปลงของค่าพยากรณ์ให้สอดคล้องกับข้อมูลจริงได้

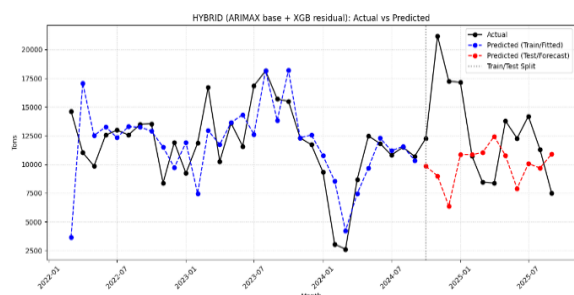
มากกว่าแบบจำลอง ARIMA เดี่ยว สะท้อนถึงการนำ XGBoost มาเรียนรู้ค่า Residual ช่วยลดลักษณะความเรียบของค่าพยากรณ์ และทำให้สามารถสะท้อนการเปลี่ยนแปลงของข้อมูลได้มากขึ้น อย่างไรก็ตาม ในบางช่วงยังพบว่าค่าพยากรณ์มีการตอบสนองต่อการเปลี่ยนแปลงของข้อมูลไม่สอดคล้องกับค่าจริงทั้งหมด เช่น มีการปรับตัวไม่ถึงระดับของการเปลี่ยนแปลงที่เกิดขึ้น



รูปที่ 5 ผลการพยากรณ์ของแบบจำลอง ARIMA-XGBoost

4.7.2 ARIMAX-XGBoost

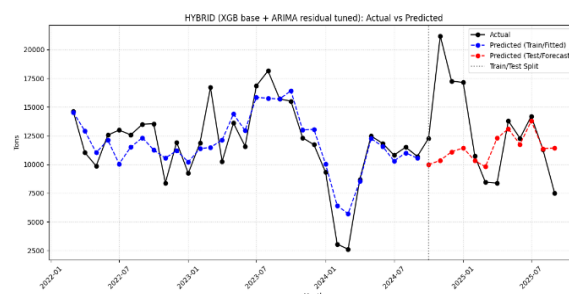
จากผลการพยากรณ์ของแบบจำลอง ARIMAX-XGBoost พบว่าตัวชี้วัด MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 23.53 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบเท่ากับ 33.64 เปอร์เซ็นต์ สามารถแสดงกราฟเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์ได้ดังรูปที่ 6 ผลลัพธ์พบว่าแบบจำลองไฮบริดสามารถลดค่า MAPE ลงจากแบบจำลอง ARIMAX เดี่ยวได้เล็กน้อยจากการนำ XGBoost มาเสริมในขั้นตอนการเรียนรู้ค่า Residual อย่างไรก็ตาม ค่าพยากรณ์ที่ได้ยังแตกต่างจากค่าจริงสูง และยังไม่สามารถอธิบายการเปลี่ยนแปลงของข้อมูลได้อย่างเพียงพอ



รูปที่ 6 ผลการพยากรณ์ของแบบจำลอง ARIMAX-XGBoost

4.7.3 XGBoost-ARIMA

จากผลการพยากรณ์ พบว่าค่า MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 16.04 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบเท่ากับ 22.54 เปอร์เซ็นต์ โดยการเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์แสดงดังรูปที่ 7 เมื่อพิจารณาผลลัพธ์โดยรวม พบว่าแบบจำลองไฮบริด XGBoost-ARIMA ให้ประสิทธิภาพใกล้เคียงกับแบบจำลอง XGBoost เดี่ยว ทั้งในแง่ของค่าตัวชี้วัด MAPE ในช่วงชุดข้อมูลทดสอบ และลักษณะของกราฟเปรียบเทียบ ซึ่งมีความแตกต่างเพียงเล็กน้อยระหว่างกัน ทั้งนี้อาจเป็นผลมาจากแบบจำลอง XGBoost สามารถอธิบายความสัมพันธ์ระหว่างปริมาณความต้องการกับปัจจัยภายนอกได้ค่อนข้างดีระดับหนึ่ง ส่งผลให้ค่า Residual ที่เหลืออยู่หลังการพยากรณ์ไม่แสดงรูปแบบเชิงเวลาที่ชัดเจนในเชิงเส้น จึงทำให้แบบจำลอง ARIMA ซึ่งพิจารณาความสัมพันธ์ของข้อมูลตามลำดับเวลาเพียงอย่างเดียว มีบทบาทในการปรับปรุงผลการพยากรณ์เพิ่มเติมค่อนข้างจำกัด และผลลัพธ์ที่ได้จึงมีลักษณะใกล้เคียงกับแบบจำลอง XGBoost เดี่ยวในภาพรวม



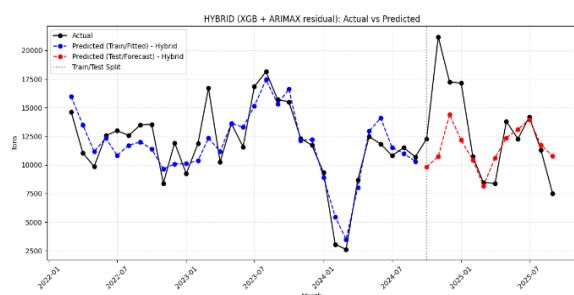
รูปที่ 7 ผลการพยากรณ์ของแบบจำลอง XGBoost-ARIMA

4.7.4 XGBoost-ARIMAX

จากผลการพยากรณ์ของแบบจำลอง XGBoost-ARIMAX พบว่าตัวชี้วัด MAPE ในช่วงชุดข้อมูลฝึกเท่ากับ 12.61 เปอร์เซ็นต์ และในช่วงชุดข้อมูลทดสอบเท่ากับ 17.83 เปอร์เซ็นต์ การเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์ดังรูปที่ 8 พบว่าค่าพยากรณ์ของแบบจำลองสามารถสะท้อนรูปแบบและทิศทางโดยรวมของข้อมูลจริงได้ค่อนข้างดี โดยเส้นพยากรณ์มีความต่อเนื่องและเคลื่อนไหวไปในทิศทางเดียวกับค่าจริงในหลายช่วงเวลา ผลลัพธ์

ดังกล่าวชี้ให้เห็นว่าการนำแบบจำลอง XGBoost เป็นแบบจำลองหลัก ให้เรียนรู้ความสัมพันธ์ที่ซับซ้อนและไม่เป็นเชิงเส้นของข้อมูล จากนั้นนำแบบจำลอง ARIMAX มาใช้ปรับแก้ค่า Residual ของ XGBoost โดยอาศัยตัวแปรค่าปริมาณการสั่งซื้อในช่วงเดือนก่อนหน้า (Qty_lag1) และค่าเฉลี่ยของปริมาณการสั่งซื้อย้อนหลัง 3 เดือน (Qty_roll3) ช่วยเสริมความสามารถของแบบจำลองในการพิจารณาความสัมพันธ์เชิงเวลาระยะสั้นที่เกี่ยวข้องกับข้อมูลความต้องการในอดีต และปรับแก้ความคลาดเคลื่อนที่ยังคงหลงเหลืออยู่ได้อย่างเหมาะสม ส่งผลให้ประสิทธิภาพการพยากรณ์ดีขึ้นเมื่อเปรียบเทียบกับแบบจำลอง XGBoost เดี่ยว รวมถึงแบบจำลองอื่น ๆ

อย่างไรก็ตาม ในบางช่วงเวลายังคงพบว่าค่าพยากรณ์ตอบสนองต่อการเปลี่ยนแปลงของข้อมูลจริงได้ไม่ครบถ้วน โดยเฉพาะในช่วงที่เกิดการเปลี่ยนแปลงอย่างฉับพลัน ซึ่งค่าพยากรณ์อาจปรับตัวไม่ถึงระดับของการเปลี่ยนแปลงที่เกิดขึ้นจริง ทั้งนี้ เมื่อพิจารณาในภาพรวมพบว่าแบบจำลองไฮบริด XGBoost-ARIMAX ให้ค่า MAPE ต่ำที่สุด และสามารถอธิบายพฤติกรรมของข้อมูลได้สอดคล้องกับรูปแบบจริงมากที่สุดเมื่อเปรียบเทียบกับแบบจำลองอื่น ๆ



รูปที่ 8 ผลการพยากรณ์ของแบบจำลอง XGBoost-ARIMAX

4.8 การเปรียบเทียบประสิทธิภาพการพยากรณ์ของแบบจำลอง

เพื่อประเมินและเปรียบเทียบประสิทธิภาพการพยากรณ์ของแบบจำลองที่พัฒนาขึ้น ได้ทำการเปรียบเทียบผลการพยากรณ์ของแบบจำลองต่าง ๆ กับค่าปริมาณการสั่งซื้อจริง และ Company Baseline ซึ่งเป็นวิธีการ

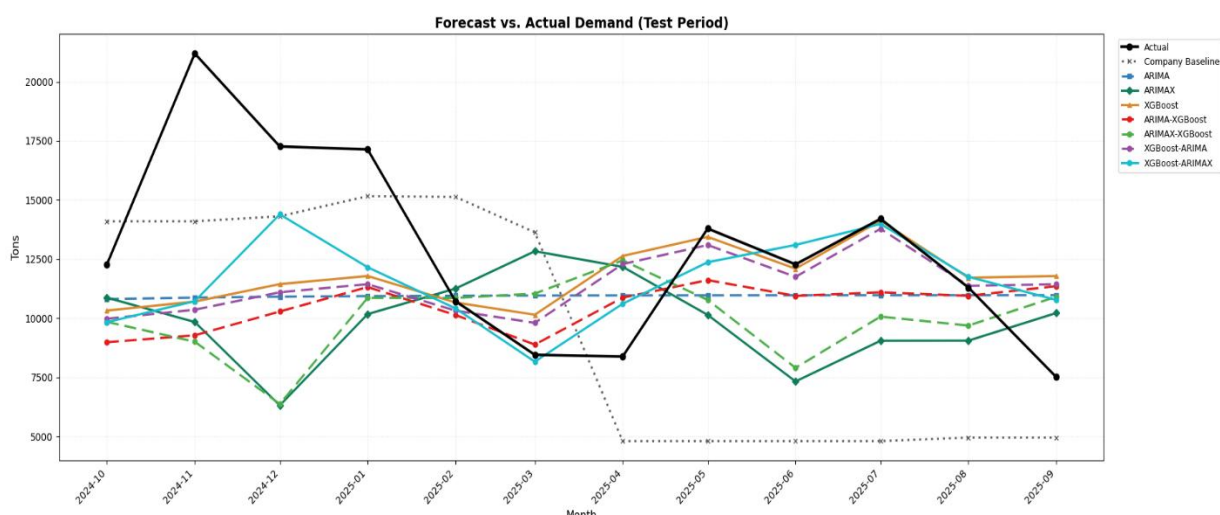
พยากรณ์ของบริษัทกรณีศึกษาที่ใช้อยู่ในปัจจุบัน สามารถแสดงตารางเปรียบเทียบตัวชี้วัดค่า MAPE ในช่วงชุดข้อมูลทดสอบดังตารางที่ 8 และกราฟเปรียบเทียบค่าพยากรณ์กับค่าจริงในช่วงชุดข้อมูลทดสอบดังรูปที่ 9

ตารางที่ 8 ค่าตัวชี้วัด MAPE ในช่วงชุดข้อมูลทดสอบ

แบบจำลองการพยากรณ์	MAPE
Company Baseline	42.06%
ARIMA	24.94%
ARIMAX	35.86%
XGBoost	22.24%
ARIMA-XGBoost	25.05%
ARIMAX-XGBoost	33.64%
XGBoost-ARIMA	22.54%
XGBoost-ARIMAX	17.83%

เมื่อพิจารณาจากกราฟการเปรียบเทียบค่าพยากรณ์กับค่าจริงในช่วงชุดข้อมูลทดสอบ พบว่าแบบจำลองเดี่ยวแต่ละแบบมีข้อจำกัดที่แตกต่างกัน โดยแบบจำลอง ARIMA มีลักษณะของค่าพยากรณ์ที่ค่อนข้างเรียบและไม่สามารถสะท้อนระดับของข้อมูลจริงได้อย่างต่อเนื่อง ขณะที่แบบจำลอง ARIMAX ค่าพยากรณ์ยังมีความคลาดเคลื่อนจากค่าจริงในหลายช่วง ส่วนแบบจำลอง XGBoost สามารถติดตามทิศทางและระดับของข้อมูลได้ดีขึ้นเมื่อเทียบกับแบบจำลองเชิงสถิติ

สำหรับแบบจำลองไฮบริด พบว่าในกรณีศึกษานี้แบบจำลองไฮบริดซึ่งใช้แบบจำลองเชิงสถิติเป็นแบบจำลองหลัก ได้แก่ ARIMA-XGBoost และ ARIMAX-XGBoost ให้ความแม่นยำในการพยากรณ์ต่ำกว่ากลุ่มแบบจำลองไฮบริดซึ่งสลับบทบาทให้แบบจำลอง Machine Learning เป็นแบบจำลองหลัก ได้แก่ XGBoost-ARIMA และ XGBoost-ARIMAX ทั้งนี้ อาจเป็นผลมาจากลักษณะของข้อมูลที่มีความซับซ้อนและผันผวนสูง ไม่มีแนวโน้มหรือความเป็นฤดูกาลอย่างมีนัยสำคัญ และมีความสัมพันธ์กับปัจจัยเศรษฐกิจภายนอกค่อนข้างมาก ซึ่งพิจารณาได้จากการที่ตัวแปรซึ่งมีค่า MI สูงส่วนใหญ่เป็นตัวแปรด้านเศรษฐกิจ



รูปที่ 9 การเปรียบเทียบค่าพยากรณ์กับค่าจริงในชุดทดสอบ

ภายนอก โดยสำหรับแบบจำลอง ARIMA-XGBoost พบว่า กราฟค่าพยากรณ์สามารถสะท้อนรูปแบบและทิศทาง โดยรวมของข้อมูลจริงได้ดีกว่าแบบจำลอง ARIMA เดี่ยว เนื่องจากการนำแบบจำลอง XGBoost สำหรับพยากรณ์ค่า Residual เข้ามาช่วยลดลักษณะความเรียบของค่าพยากรณ์ ในขณะที่ แบบจำลอง ARIMAX-XGBoost แม้ผลการพยากรณ์จะดีขึ้นกว่าแบบจำลอง ARIMAX เดี่ยว แต่ค่าพยากรณ์ก็ยังไม่สามารถอธิบายข้อมูลจริงได้อย่างเพียงพอ และยังคงมีความคลาดเคลื่อนสูง สำหรับแบบจำลอง XGBoost-ARIMA ให้ผลการพยากรณ์ที่มีประสิทธิภาพใกล้เคียงกับแบบจำลอง XGBoost เดี่ยว ซึ่งสะท้อนให้เห็นว่าแบบจำลอง ARIMA ยังไม่สามารถช่วยปรับแก้ผลการพยากรณ์ได้อย่างเด่นชัด

ในขณะที่แบบจำลอง XGBoost-ARIMAX ให้ผลการพยากรณ์ที่ดีที่สุดเมื่อเปรียบเทียบกับแบบจำลองอื่น ๆ ในกรณีศึกษา นี้ โดยสามารถสะท้อนทิศทางการเปลี่ยนแปลงของข้อมูลได้ใกล้เคียงกับค่าจริงมากขึ้น ซึ่งแสดงให้เห็นว่าการใช้ XGBoost เป็นแบบจำลองหลักที่อาศัยตัวแปรเศรษฐกิจภายนอก ร่วมกับ ARIMAX ในการปรับแก้ค่า Residual ด้วยตัวแปรความต้องการย้อนหลังระยะสั้น ช่วยให้แบบจำลองสามารถสะท้อนทั้งอิทธิพลของปัจจัยภายนอกและความสัมพันธ์เชิงเวลาในระยะสั้นของข้อมูลได้ดียิ่งขึ้น

เมื่อเปรียบเทียบกับวิธีการพยากรณ์ที่บริษัท กรณีศึกษา ใช้อยู่ในปัจจุบัน พบว่าแบบจำลองที่พัฒนาขึ้น โดยเฉพาะแบบจำลองไฮบริดกลุ่มที่ใช้แบบจำลอง Machine Learning เป็นแบบจำลองหลัก มีประสิทธิภาพในการพยากรณ์ที่ดีขึ้น และสามารถอธิบายทิศทางของข้อมูลได้เหมาะสมยิ่งขึ้น

5. สรุปผลการวิจัย

การพยากรณ์ ความต้องการเม็ดพลาสติกมีความสำคัญต่อการวางแผนการผลิตและการจัดการสินค้าคงคลังเป็นอย่างมาก เนื่องจากลักษณะของข้อมูลที่มีความผันผวนและเปลี่ยนแปลงตามสถานะทางเศรษฐกิจ งานวิจัยนี้ จึงมุ่งศึกษาการพัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ความต้องการเม็ดพลาสติก โดยใช้ข้อมูลกรณีศึกษาผู้ผลิตปิโตรเคมีรายหนึ่งในประเทศไทย

งานวิจัยนี้ ทำการศึกษาสินค้าที่ได้รับการคัดเลือกจากเกณฑ์ยอดขายรวมสูงสุดในช่วงเวลาที่ทำการศึกษา และทำการวิเคราะห์ลักษณะข้อมูลด้วยวิธีการทดสอบ Mann-Kendall Trend Test และ Friedman Test ซึ่งพบว่าข้อมูลไม่มีแนวโน้มและความเป็นฤดูกาลอย่างมีนัยสำคัญ จากนั้น มีการคัดเลือกตัวแปรด้วย Mutual Information (MI) แสดงให้เห็นว่าปริมาณความต้องการมีความสัมพันธ์กับปัจจัยเศรษฐกิจภายนอกในระดับค่อนข้างสูง จากนั้นพัฒนา

แบบจำลองการพยากรณ์ทั้งหมด 7 แบบ ได้แก่ ARIMA, ARIMAX, XGBoost, ARIMA-XGBoost, ARIMAX-XGBoost, XGBoost-ARIMA และ XGBoost-ARIMAX โดยใช้ตัวชี้วัด MAPE ในการประเมินประสิทธิภาพของแบบจำลอง

ผลการศึกษาพบว่า แบบจำลองไฮบริดที่ใช้แบบจำลอง Machine Learning เป็นแบบจำลองหลักมีประสิทธิภาพการพยากรณ์ดีกว่ากลุ่มแบบจำลองเชิงสถิติเดี่ยวและแบบจำลองไฮบริดที่ใช้แบบจำลองเชิงสถิติเป็นแบบจำลองหลัก แม้ว่าผลการพยากรณ์ของบางแบบจำลองในกลุ่มดังกล่าวจะมีประสิทธิภาพใกล้เคียง หรือด้อยกว่าแบบจำลอง Machine Learning เดี่ยวเล็กน้อย แต่แนวโน้มของผลลัพธ์โดยรวมสะท้อนให้เห็นว่า แนวทางการพัฒนาแบบจำลองไฮบริดในลักษณะดังกล่าวสามารถช่วยปรับปรุงประสิทธิภาพการพยากรณ์ได้ในกรณีศึกษาที่ โดยพบว่าแบบจำลองซึ่งให้ผลการพยากรณ์ที่ดีที่สุดเมื่อเปรียบเทียบกับแบบจำลองทั้งหมด ได้แก่ แบบจำลอง XGBoost-ARIMAX โดยมีค่า MAPE ในช่วงชุดข้อมูลทดสอบเท่ากับ 17.83 เปอร์เซนต์ ผลลัพธ์ดังกล่าวสะท้อนให้เห็นว่า สำหรับข้อมูลในกรณีศึกษาที่มีความผันผวนค่อนข้างสูง ไม่มีแนวโน้มและความเป็นฤดูกาลอย่างมีนัยสำคัญ การใช้ XGBoost เป็นแบบจำลองหลักร่วมกับ ARIMAX ในการปรับแก้ค่า Residual เพื่อพิจารณาทั้งอิทธิพลของปัจจัยเศรษฐกิจภายนอกและความสัมพันธ์เชิงเวลาระยะสั้นของข้อมูล เป็นแนวทางที่เหมาะสมในการปรับปรุงทั้งประสิทธิภาพการพยากรณ์และการอธิบายรูปแบบของข้อมูล

6. ข้อจำกัดและคำแนะนำสำหรับงานวิจัยในอนาคต

ข้อมูลที่ใช้ในการศึกษานี้เป็นข้อมูลรายเดือนรวมทั้งสิ้น 43 เดือน ซึ่งถือว่ามีความค่อนข้างจำกัดสำหรับการพัฒนาแบบจำลองในกลุ่ม Machine Learning เช่น XGBoost ข้อจำกัดด้านปริมาณข้อมูลดังกล่าวอาจส่งผลต่อความเสถียรในการเรียนรู้ของแบบจำลอง และอาจเพิ่มความเสี่ยงของปัญหา Overfitting ภายใต้ข้อจำกัดของขนาดชุดข้อมูลที่ใช้ในการศึกษาดังกล่าว รวมถึงทำให้ไม่

เหมาะสมสำหรับการพัฒนาแบบจำลอง SARIMA ที่ต้องการข้อมูลหลายรอบฤดูกาล อย่างไรก็ตาม ผู้วิจัยได้ดำเนินการทดลองเบื้องต้นในการพยากรณ์ด้วยข้อมูลระดับรายสัปดาห์เพื่อเพิ่มจำนวนตัวอย่างแล้ว แต่พบว่าข้อมูลคำสั่งซื้อในระดับดังกล่าวมีความไม่สม่ำเสมอและมีความผันผวนสูงอย่างมาก จนทำให้ไม่สามารถพยากรณ์ได้อย่างเหมาะสม ประกอบกับตัวแปรทางเศรษฐกิจที่สำคัญส่วนใหญ่ไม่มีการรายงานในระดับรายสัปดาห์ ส่งผลให้การพัฒนาแบบจำลองด้วยข้อมูลความถี่สูงมีความเสถียรต่ำ ดังนั้น ในการศึกษาครั้งต่อไป อาจพิจารณาการใช้ข้อมูลที่มีความถี่สูงขึ้น เช่น ข้อมูลรายสัปดาห์ ร่วมกับการเปรียบเทียบแนวทางการพยากรณ์อื่น ๆ ที่เหมาะสมกับลักษณะข้อมูลดังกล่าว เช่น แบบจำลองในกลุ่ม Deep Learning หรือ Neural Networks รวมถึงการพิจารณาปัจจัยเพิ่มเติมที่อาจมีผลต่อความต้องการสินค้า เพื่อประเมินความเหมาะสมและศักยภาพของแนวทางต่าง ๆ ในการพยากรณ์ในบริบทของงานวิจัยต่อไป

นอกจากนี้ เนื่องจากงานวิจัยนี้เลือกศึกษาสินค้าเพียงหนึ่งรายการ ซึ่งเป็นสินค้าที่มียอดขายรวมสูงสุดในช่วงเวลาที่ศึกษา ส่งผลให้ผลการพยากรณ์ที่ได้สะท้อนเฉพาะลักษณะของสินค้านั้น ๆ และอาจยังไม่สามารถแทนประสิทธิภาพของแบบจำลองเมื่อใช้กับสินค้าที่มีพฤติกรรมความต้องการแตกต่างกันได้ ดังนั้น งานวิจัยในอนาคตอาจพิจารณานำแบบจำลองไปพัฒนาเพิ่มเติมกับสินค้ากลุ่มอื่นที่มีรูปแบบความต้องการหลากหลาย เพื่อประเมินความเหมาะสมในการนำไปใช้งานในบริบทที่กว้างขึ้น

7. กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์จากบริษัทกรณีศึกษาที่ให้ข้อมูลและความร่วมมือเป็นอย่างดี รวมถึงผู้ที่ให้คำแนะนำและสนับสนุนตลอดกระบวนการดำเนินงานวิจัย

8. เอกสารอ้างอิง

- [1] กลุ่มอุตสาหกรรมปิโตรเคมี, “ข้อมูลและประโยชน์,” [ออนไลน์]. แหล่งที่มา: <https://www.ftipc.or.th/th/chemical/8>. [วันที่ เข้าถึง 22 สิงหาคม 2568].
- [2] กลุ่มอุตสาหกรรมปิโตรเคมี สภาอุตสาหกรรมแห่งประเทศไทย, “รายงานประจำปี 2567,” [ออนไลน์]. แหล่งที่มา: https://www.ftipc.or.th/upload_file/annual/annual_TH_12042025-18300904.pdf. [วันที่เข้าถึง 22 สิงหาคม 2568].
- [3] เทียร เทียมศักดิ์, “แนวโน้มธุรกิจ/อุตสาหกรรม ปี 2567-2569: อุตสาหกรรมปิโตรเคมี,” [ออนไลน์]. แหล่งที่มา: <https://www.krungsri.com/th/research/industry/industry-outlook/chemicals/chemicals/io/io-chemicals-2024-2026>. [วันที่เข้าถึง 25 สิงหาคม 2568].
- [4] เทียร เทียมศักดิ์, “แนวโน้มธุรกิจ/อุตสาหกรรม ปี 2568-2570: อุตสาหกรรมพลาสติก,” [ออนไลน์]. แหล่งที่มา: <https://www.krungsri.com/th/research/industry/industry-outlook/chemicals/plastics/io/io-plastics-2025-2027>. [วันที่เข้าถึง 22 สิงหาคม 2568].
- [5] P. Udom and N. Phumchusri, “A comparison study between time series model and ARIMA model for sales forecasting of distributor in plastic industry,” *IOSR Journal of Engineering*, vol. 4, no. 2, pp. 32-38, 2014, doi: 10.9790/3021-04213238.
- [6] B. Siregar, E. B. Nababan, A. Yap, U. Andayani, and Fahmi, “Forecasting of raw material needed for plastic products based in income data using ARIMA method,” in *2017 5th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*, Malang, Indonesia, 2017, pp. 135-139, doi: 10.1109/ICEEIE.2017.8328777.
- [7] N. Kantanatha and P. Pattaraumpornchai, “Machine learning-based price forecasting for polypropylene granules in Thailand,” in *2023 6th International Conference on Computational Intelligence and Intelligent Systems (CIIS '23)*, Tokyo, Japan, 2024, pp. 14-19, doi: 10.1145/3638209.3638212.
- [8] H. Zhao, X. Zuo, and P. Lin, “Sales forecasting for chemical products by using SARIMA model,” in *The 5th International Conference on Big Data and Education (ICBDE '22)*, Shanghai, China, 2022, pp. 419-427, doi: 10.1145/3524383.3524396.
- [9] F. Rabbi, S. U. Tareq, M. M. Islam, M. A. Chowdhury, and M. Abul Kashem, “A multivariate time series approach for forecasting of electricity demand in Bangladesh using ARIMAX model,” in *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, Dhaka, Bangladesh, 2020, pp. 1-5, doi: 10.1109/STI50764.2020.9350326.
- [10] I. T. Daramola, A. Garg, and D. Mehrotra, “Multivariate time series forecasting of municipal solid waste,” in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, 2024, pp. 1-5, doi: 10.1109/ICRITO61523.2024.10522312.
- [11] J. George, J. Yadav, A. M. Nair, P. M. V., B. P. Alapatt, and R. Baby, “Improving groundwater forecasting accuracy with a hybrid ARIMA-XGBoost approach,” in *2024 3rd International Conference for Advancement in Technology (ICONAT)*, Goa, India, 2024, pp. 1-7, doi: 10.1109/ICONAT61936.2024.10775242.

- [12] N. Nurhayati, A. F. A. Hakim, K. Hulliyah, and K. Nazelika, "The hybrid ARIMA-XGBoost model for forecasting performance analysis of new HIV cases in Indonesia," in *2025 13th International Conference on Cyber and IT Service Management (CITSM)*, Jakarta, Indonesia, 2025, pp. 1-6, doi: 10.1109/CITSM67730.2025.11291525.
- [13] Ü. Işlak, E. Yılmaz, İ. Arslan, and T. Çakar, "Combining similar trajectories and XGBoost via residual learning for traffic flow forecasting," in *2025 9th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Ankara, Türkiye, 2025, pp. 1-5, doi: 10.1109/ISMSIT67332.2025.11267955.
- [14] G. T. Patle, A. Libang, and S. Ahuja, "Analysis of rainfall and temperature variability and trend detection: A non-parametric Mann-Kendall test approach," in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2016, pp. 1723-1727.
- [15] A. Borucka, R. Parczewski, E. Kozłowski, and A. Swiderski, "Evaluation of air traffic in the context of the COVID-19 pandemic," *Archives of Transport.*, vol. 64, no. 4, pp. 45-57, 2022, doi: 10.5604/01.3001.0016.1048.
- [16] T. Liu, H. Wei, K. Zhang, and W. Guo, "Mutual information based feature selection for multivariate time series forecasting," in *2016 35th Chinese Control Conference (CCC)*, Chengdu, China, 2016, pp. 7110-7114, doi: 10.1109/ChiCC.2016.7554480.
- [17] C. Vithitsontorn and P. Chongstitvatana, "Demand forecasting in production planning for dairy products using machine learning and statistical method," in *2022 International Electrical Engineering Congress (iEECON)*, Khon Kaen, Thailand, 2022, pp. 1-4, doi: 10.1109/iEECON53204.2022.9741683.
- [18] J. Chang, Y. Lu, Q. Xu, and H. Zhang, "Comparative study of ARIMAX and BP neural network models in the prediction of *Spodoptera frugiperda* occurrence," in *2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Chongqing, China, 2023, pp. 1769-1774, doi: 10.1109/ITOEC57671.2023.10291455.