

การเพิ่มประสิทธิภาพการควบคุมระบบตัวควบคุมแบบสัดส่วนร่วมกับปริพันธ์
และอนุพันธ์ ด้วยการประยุกต์เทคนิคโมเดลพยากรณ์

ENHANCING PROPORTIONAL-INTEGRAL-DERIVATIVE CONTROL WITH
APPLYING MODEL PREDICTIVE TECHNIQUES



การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์
คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2567

ลิขสิทธิ์ของคณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ENHANCING PROPORTIONAL-INTEGRAL-DERIVATIVE CONTROL WITH
APPLYING MODEL PREDICTIVE TECHNIQUES



AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE PROGRAM IN DATA SCIENCE AND
ANALYTICS

SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2024

COPYRIGHT OF SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

ชื่อเรื่อง	การเพิ่มประสิทธิภาพการควบคุมระบบตัวควบคุมแบบสัดส่วนร่วมกับ ปริพันธ์ และอนุพันธ์ ด้วยการประยุกต์เทคนิคโมเดลพยากรณ์
นักศึกษา	สุทธิพงษ์ สัจจรัมย์
รหัสประจำตัว	66056084
ปริญญา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์
ปีการศึกษา	2567
อาจารย์ที่ปรึกษาหลัก	ผู้ช่วยศาสตราจารย์ ดร.ณัฐพร ชื่นเจริญ

บทคัดย่อ

การศึกษาค้นคว้ามุ่งเน้นการปรับปรุงระบบควบคุม PID แบบวงจรมอด (Close Loop) โดยประยุกต์ใช้โมเดลพยากรณ์ ซึ่งใช้ข้อมูลจากโรงงานอุตสาหกรรมแห่งหนึ่งในประเทศไทย ซึ่งเป็นข้อมูลทุติยภูมิ (Secondary Data) ที่จัดเก็บข้อมูลแบบเรียลไทม์ (Real-Time) บนฐานข้อมูล (Database) โดยผู้ศึกษาใช้ข้อมูลทุกๆ 1 วินาที จากวันที่ 1 มีนาคม 2568 เวลา 00:00 ถึงวันที่ 7 มีนาคม 2568 เวลา 00:00 ซึ่งมีทั้งสิ้น 518,401 ตัวอย่าง จำนวน 6 ตัวแปร ได้แก่ 1.ค่าอัตราไหลของน้ำ (PV) 2.ค่าเป้าหมาย (SV) 3.ค่าควบคุม (MV) 4.ค่า Proportional (P) 5.ค่า Integral (I) 6.ค่า Derivative (D) และทำการการสร้างตัวแปรใหม่ (Feature Engineering) ขึ้นมา คือ ค่าผิดพลาด $e(t)$ ซึ่งได้มาจากการนำค่า SV ลบด้วยค่า PV และค่าPIDเอาต์พุต $u(t)$ ซึ่งได้มาจากการนำค่า $e(t)$ ไปคำนวณในสมการมาตรฐาน PID จากนั้นสร้างแบบจำลองการถดถอยเชิงเส้น (Linear Regression) ทั้งหมด 5 วิธี คือ OLS, OLS-HAC, GLS, GLSAR, WLS และสุดท้ายผู้ศึกษาเลือกใช้งานแบบจำลอง OLS-HAC ในการปรับปรุงระบบควบคุม PID โดยที่แบบจำลอง OLS-HAC มีค่า RMSE เท่ากับ 0.727 และค่า R^2 เท่ากับ 0.842 สามารถสร้างสมการเส้นตรงได้คือ $Y = 519.9999 - 0.562X$

คำสำคัญ: ค่าควบคุม, ค่าผิดพลาด, ค่าเป้าหมาย, ค่าPIDเอาต์พุต, ค่าอัตราไหลของน้ำ

Title ENHANCING PROPORTIONAL-INTEGRAL-DERIVATIVE CONTROL
WITH APPLYING MODEL PREDICTIVE TECHNIQUES

Student SUTTHIPONG SANGRAM

Student ID 66056084

Degree Master of Science Program in Data Science and Analytics

Academic Year 2024

Advisor Assistant Professor Dr. NATTAPORN CHUENJARERN

ABSTRACT

This study focuses on improving a closed loop PID control system by applying a forecasting model using data from an industrial factory in Thailand. Which is secondary data that is stored in real-time on the database. The researcher used data every 1 second from March 1, 2025, time 00:00 to March 7, 2025, time 00:00, which total 518,401 samples with 6 variables. Including 1. Flow Rate (PV) 2. Setpoint (SV) 3. Manipulated Variable (MV) 4. Proportional value (P) 5. Integral value (I) 6. Derivative value (D) and create a new variable by feature engineering method, which are Error value $e(t)$ that is obtained by taking SV value and subtracting PV value and PIDoutput value $u(t)$ that is obtained by calculating $e(t)$ in standard PID equation. Then, create a linear regression model total 5 statistical method. There are several regression estimators, including OLS, OLS-HAC, GLS, GLSAR, and WLS. Finally, the researcher chose to use the OLS-HAC model to improve the PID control system, where the OLS-HAC model has an RMSE value of 0.727 and an R^2 value of 0.842, which can create a linear equation as $Y = 519.9999 - 0.562X_1$

Keywords: Control Value, Error Value, Setpoint Value, PIDoutput value, Water Flow Rate Value

กิตติกรรมประกาศ

การศึกษาค้นคว้าอิสระฉบับนี้สำเร็จลุล่วงด้วยดี มีเนื้อหาครบถ้วนและถูกต้องสมบูรณ์ได้ด้วยความอนุเคราะห์อย่างยิ่งจาก ผู้ช่วยศาสตราจารย์ ดร.ณัฐพร ชื่นเจริญ อาจารย์ที่ปรึกษา ที่ได้กรุณาให้คำแนะนำ ชี้แนะแนวทาง และติดตามความก้าวหน้าของการศึกษาค้นคว้าอิสระเรื่อง "การเพิ่มประสิทธิภาพการควบคุมระบบตัวควบคุมแบบสัดส่วนร่วมกับปริพันธ์และอนุพันธ์ ด้วยการประยุกต์เทคนิคโมเดลพยากรณ์" มาโดยตลอด รวมถึงการตรวจแก้ไขความถูกต้องจนกระทั่งโครงการนี้เสร็จสมบูรณ์ ผู้ศึกษาขอกราบขอบพระคุณท่านด้วยความเคารพอย่างสูงมา ณ โอกาสนี้

ผู้ศึกษาขอกราบขอบพระคุณเป็นอย่างสูงต่อ ผู้ช่วยศาสตราจารย์ ดร.จิรภัทร์ หยกรัตนศักดิ์ และ ผู้ช่วยศาสตราจารย์ ดร.ยุวดี กล่อมวิเศษ คณะกรรมการสอบการศึกษาค้นคว้าอิสระ ที่ได้กรุณาสละเวลาอันมีค่าเพื่อให้คำแนะนำและข้อเสนอแนะอันเป็นประโยชน์ยิ่ง ซึ่งเป็นส่วนสำคัญที่ช่วยให้งานศึกษาค้นคว้าอิสระครั้งนี้สำเร็จลุล่วงไปได้ด้วยดี

ผู้ศึกษาขอกราบขอบพระคุณเป็นอย่างสูงต่อคณาจารย์วิทยาลัยมหาบัณฑิต สาขาวิทยาการข้อมูลและการวิเคราะห์ทุกท่าน ที่ได้มอบวิชาความรู้อันทรงคุณค่า ซึ่งเป็นพื้นฐานสำคัญให้การศึกษาค้นคว้าอิสระครั้งนี้สำเร็จลุล่วงไปได้ด้วยดี

สุดท้ายนี้ ผู้ศึกษาขอกราบขอบพระคุณบิดา มารดา ครอบครัว และเพื่อนทุกท่าน ที่ได้มอบการสนับสนุน กำลังใจสำคัญตลอดระยะเวลาการศึกษาค้นคว้าอิสระในครั้งนี้

สุทธิพงษ์ สัจจรัสมย์

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูปภาพ.....	ณ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	1
1.3 ขอบเขตของงานวิจัย.....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	2
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	3
2.1 ระบบควบคุม PID.....	3
2.1.1 ค่าสัดส่วน (Proportional).....	3
2.1.2 ค่าปริพันธ์ (Integral).....	3
2.1.3 ค่าอนุพันธ์ (Derivative).....	4
2.2 การถดถอยเชิงเส้น (Linear Regression).....	5
2.2.1 แบบจำลอง OLS (Ordinary Least Squares).....	6
2.2.2 แบบจำลอง GLS (Generalized Least Squares).....	11

สารบัญ (ต่อ)

	หน้า
2.2.3 แบบจำลอง GLSAR (GLS with Autoregressive).....	12
2.2.4 แบบจำลอง WLS (Weighted Least Squares).....	14
2.2.5 การทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอย	15
2.2.6 การประเมินผลประสิทธิภาพตัวแบบ.....	16
2.3 งานวิจัยที่เกี่ยวข้อง	17
บทที่ 3 วิธีการดำเนินงานวิจัย.....	19
3.1 ศึกษาการทำงานของระบบ.....	19
3.2 แนวคิดในการใช้แบบจำลองพยากรณ์ในระบบควบคุม PID.....	21
3.3 การรวบรวมข้อมูล.....	22
3.4 การจัดการข้อมูล.....	23
3.4.1 การสำรวจข้อมูล (Data Exploration)	23
3.4.2 การทำความสะอาดข้อมูล (Data Cleaning).....	26
3.4.3 การแปลงข้อมูล (Data Transformation).....	28
3.5 เครื่องมือที่ใช้.....	32
3.6 การแบ่งชุดข้อมูลสำหรับการเรียนรู้ของเครื่อง (Data Splitting).....	33
3.6.1 ชุดฝึกฝน (Training Set).....	33
3.6.2 ชุดทดสอบ (Testing Set).....	33

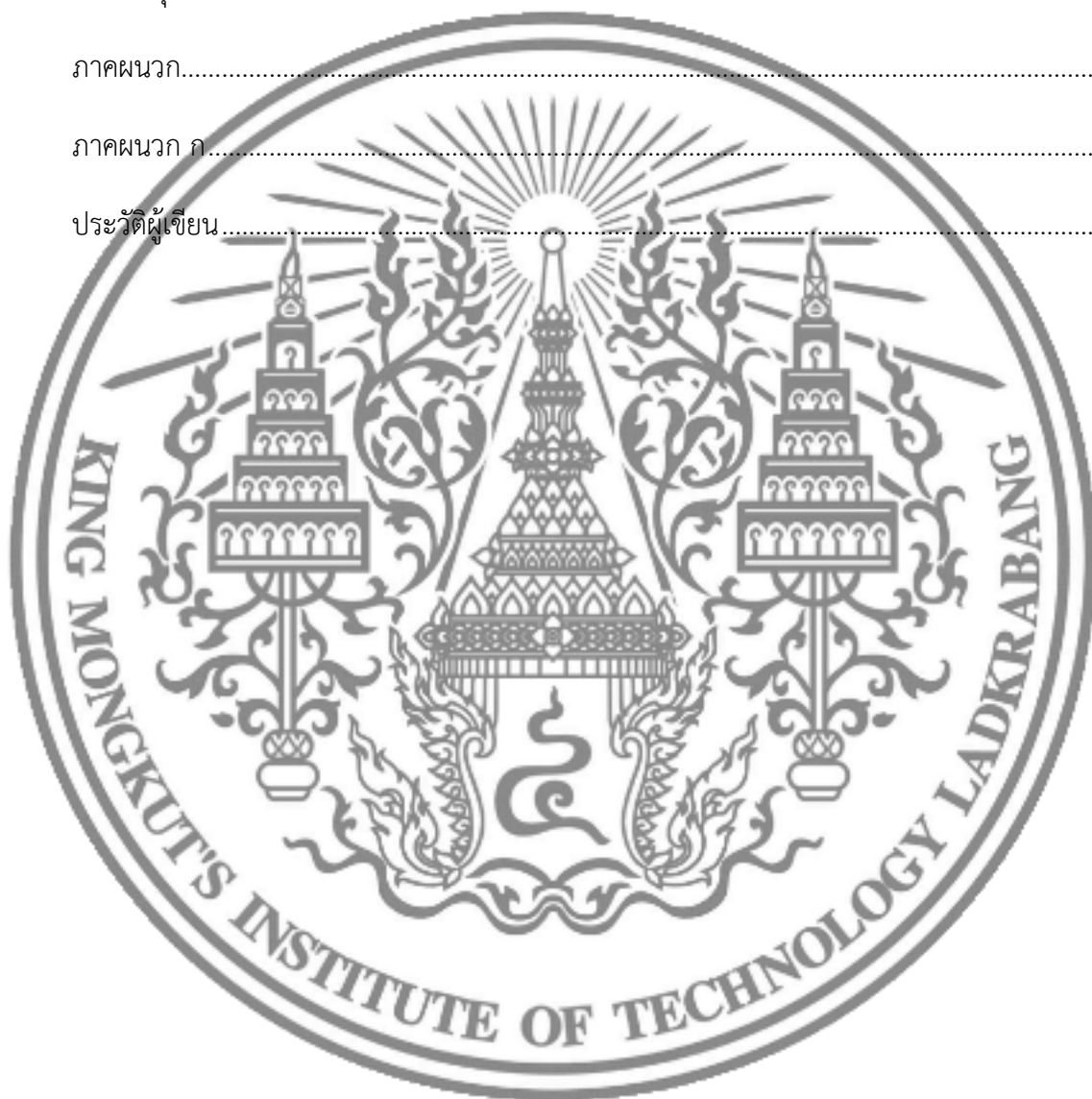
สารบัญ (ต่อ)

หน้า

บทที่ 4 ผลการวิจัยและการอภิปรายผล	34
4.1 แบบจำลอง OLS (Ordinary Least Squares).....	34
4.1.1 ตรวจสอบข้อสมมติฐานของแบบจำลองถดถอยเชิงเส้น OLS.....	35
4.2 แบบจำลอง OLS ปรับปรุงด้วย HAC.....	37
4.2.1 ตรวจสอบข้อสมมติฐานของแบบจำลองถดถอยเชิงเส้น OLS-HAC.....	38
4.3 แบบจำลอง GLS (Generalized Least Squares).....	40
4.3.1 ตรวจสอบข้อสมมติของแบบจำลองถดถอยเชิงเส้น GLS.....	41
4.4 แบบจำลอง GLSAR (GLS with Autoregressive errors).....	43
4.4.1 ตรวจสอบข้อสมมติของแบบจำลองถดถอยเชิงเส้น GLSAR.....	44
4.5 แบบจำลอง WLS (Weighted Least Squares)	46
4.5.1 ตรวจสอบข้อสมมติของแบบจำลองถดถอยเชิงเส้น WLS.....	47
4.6 การอภิปรายผล.....	49
4.6.1 ผลของแบบจำลอง OLS, OLS-HAC, GLS, GLSAR, WLS.....	49
4.6.2 เลือกแบบจำลองและประเมินประสิทธิภาพ (Select and Evaluation).....	51
4.6.3 ปัญหาที่พบ	52
4.6.4 ผลการนำตัวแบบไปใช้งานจริง	53
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	55
5.1 สรุปผลการวิจัย.....	55
5.2 ข้อเสนอแนะ.....	56

สารบัญ (ต่อ)

	หน้า
บรรณานุกรม.....	57
ภาคผนวก.....	60
ภาคผนวก ก.....	61
ประวัติผู้เขียน.....	62



สารบัญตาราง

	หน้า
ตารางที่ 3.1 คุณลักษณะข้อมูลดิบ	23
ตารางที่ 3.2 สถิติข้อมูลดิบ	24
ตารางที่ 3.3 สถิติข้อมูลดิบ (หลังขจัดค่าผิดปกติ)	27
ตารางที่ 3.4 คุณลักษณะข้อมูลผ่านการแปลงแล้ว	28
ตารางที่ 3.5 สถิติข้อมูลผ่านการแปลงแล้ว	29
ตารางที่ 3.6 สถิติข้อมูลผ่านการแปลงแล้ว (หลังขจัดค่าผิดปกติ)	30
ตารางที่ 3.7 เครื่องมือที่ใช้งาน	32
ตารางที่ 4.1 ผลลัพธ์แบบจำลอง OLS	34
ตารางที่ 4.2 ผลลัพธ์แบบจำลอง OLS-HAC	37
ตารางที่ 4.3 เปรียบเทียบแบบจำลอง OLS และแบบจำลอง OLS ที่ปรับปรุงด้วย HAC	39
ตารางที่ 4.4 ผลลัพธ์แบบจำลอง GLS	40
ตารางที่ 4.5 ผลลัพธ์แบบจำลอง GLSAR	43
ตารางที่ 4.6 ผลลัพธ์แบบจำลอง WLS	46
ตารางที่ 4.7 เปรียบเทียบค่าสถิติของแบบจำลอง OLS, OLS-HAC, GLS, GLSAR, WLS	50

สารบัญรูปภาพ

หน้า

รูปที่ 2.1 แนวคิดการหาเส้นตรงที่ดีที่สุด.....	6
รูปที่ 3.1 กระบวนการควบคุมอัตราไหลของน้ำ	20
รูปที่ 3.2 บล็อกไดอะแกรมแสดงการทำงานภายในของระบบ DCS	20
รูปที่ 3.3 บล็อกไดอะแกรมแนวคิดในการประยุกต์แบบจำลองพยากรณ์.....	21
รูปที่ 3.4 ผังการทำงานของกระบวนการ Predictive-Optimize.....	22
รูปที่ 3.5 ตัวอย่างข้อมูลดิบ.....	23
รูปที่ 3.6 ตรวจสอบข้อมูลหายในแต่ละคอลัมน์ของข้อมูลดิบ	24
รูปที่ 3.7 การกระจายตัวของข้อมูลดิบ.....	25
รูปที่ 3.8 ความสัมพันธ์ระหว่างตัวแปรของข้อมูลดิบ.....	25
รูปที่ 3.9 แผนภาพกล่องของข้อมูลดิบ	26
รูปที่ 3.10 แผนภาพกล่องของข้อมูลดิบ (หลังจัดค่าผิดปกติ)	26
รูปที่ 3.11 การกระจายตัวของข้อมูลดิบ (หลังจัดค่าผิดปกติ).....	27
รูปที่ 3.12 ตัวอย่างข้อมูลที่ผ่านการแปลงแล้ว	28
รูปที่ 3.13 แผนภาพกล่องของข้อมูลที่ผ่านการแปลงแล้ว.....	29
รูปที่ 3.14 แผนภาพกล่องของข้อมูลที่ผ่านการแปลงแล้ว (หลังจัดค่าผิดปกติ).....	30
รูปที่ 3.15 การกระจายตัวของข้อมูลที่ผ่านการแปลงแล้ว (หลังจัดค่าผิดปกติ).....	31
รูปที่ 3.16 ความสัมพันธ์ระหว่างตัวแปรของข้อมูลที่ผ่านการแปลงแล้ว	31
รูปที่ 4.1 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน OLS.....	35
รูปที่ 4.2 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง OLS	36

สารบัญรูปภาพ (ต่อ)

	หน้า
รูปที่ 4.3 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน GLS.....	41
รูปที่ 4.4 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง GLS	42
รูปที่ 4.5 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน GLSAR	44
รูปที่ 4.6 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง GLSAR.....	45
รูปที่ 4.7 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน WLS.....	47
รูปที่ 4.8 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง WLS	48
รูปที่ 4.9 ประเมินประสิทธิภาพแบบจำลอง OLS-HAC.....	51
รูปที่ 4.10 เปรียบเทียบค่าอัตราไหลของน้ำที่วัดได้จริงกับค่าทำนายด้วยแบบจำลอง OLS-HAC.....	52
รูปที่ 4.11 จำลองการใช้งาน Predictive-Optimize ในระบบควบคุม PID	53
รูปที่ 4.12 แผนภาพกล่องเปรียบเทียบก่อนและหลังใช้งานแบบจำลอง OLS-HAC.....	54



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันอุตสาหกรรมต่างๆในประเทศไทยได้ใช้ระบบควบคุมแบบอัตโนมัติอย่างแพร่หลายมากขึ้น โดยการใช้คอมพิวเตอร์หรือโปรแกรมต่างๆเพื่อสั่งงานให้เครื่องจักรหรืออุปกรณ์ในกระบวนการผลิตนั้นๆทำงานตามที่ออกแบบไว้ ซึ่งการสั่งงานจะถูกส่งผ่านไปยังระบบควบคุมอีกที่

ระบบควบคุมที่ได้รับความนิยมอย่างสูงในอุตสาหกรรมต่างๆคือระบบควบคุมแบบ Proportional-Integral-Derivative (PID) เนื่องจากมีการออกแบบที่เรียบง่ายและสามารถประยุกต์ใช้ได้อย่างมีประสิทธิภาพในหลากหลายรูปแบบ

อย่างไรก็ตามระบบควบคุมแบบ PID ยังมีข้อจำกัดในบางกรณี เช่น ไม่สามารถทำงานได้อย่างมีประสิทธิภาพในลักษณะงานที่มีการเปลี่ยนแปลงตามเวลา (Time-varying) อย่างรวดเร็ว หรือการปรับค่าพารามิเตอร์ K_p , K_i , K_d ของระบบควบคุม PID ไม่เหมาะสมกับลักษณะของงานที่ต้องการควบคุม ทำให้ระบบอัตโนมัติหรืออุปกรณ์ในกระบวนการผลิตทำงานได้ไม่ดีเท่าที่ควร ดังนั้นการเพิ่มประสิทธิภาพการควบคุมแบบ PID จึงเป็นสิ่งสำคัญอย่างยิ่งที่ต้องดำเนินการ

การศึกษานี้มุ่งเน้นการประยุกต์ใช้เทคนิคโมเดลพยากรณ์ (Predictive Model) เพื่อปรับปรุงการควบคุมแบบ PID ให้ตอบสนองต่อการเปลี่ยนแปลงตามเวลาได้อย่างรวดเร็ว และลดความผันผวนค่าที่ต้องการควบคุม

1.2 วัตถุประสงค์ของงานวิจัย

- 1) เพื่อเปรียบเทียบแบบจำลองการถดถอยเชิงเส้น (Linear Regression) ทั้งหมด 5 วิธี ได้แก่ OLS, OLS-HAC , GLS, GLSAR, WLS
- 2) เพื่อเลือกแบบจำลองที่เหมาะสมสำหรับระบบควบคุมแบบ PID
- 3) นำแบบจำลองไปใช้งานร่วมกับระบบควบคุมแบบ PID

1.3 ขอบเขตของงานวิจัย

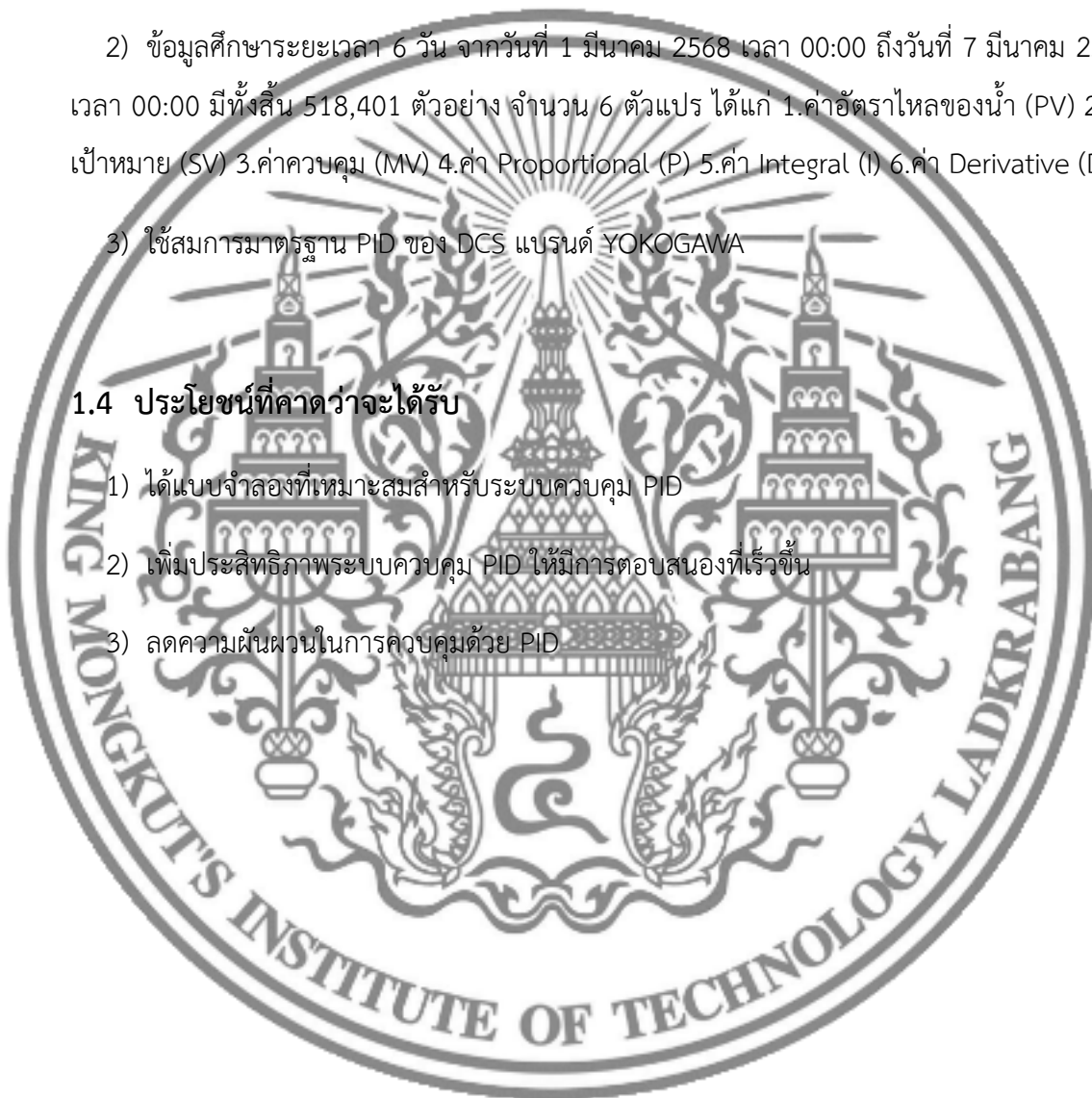
1) ใช้ข้อมูลจากการควบคุมอัตราไหลของน้ำด้วยระบบควบคุม PID แบบวงปิด (Closed loop) โดยมี feedback control มาจากเครื่องมือวัดอัตราไหล (Flow meter)

2) ข้อมูลศึกษาระยะเวลา 6 วัน จากวันที่ 1 มีนาคม 2568 เวลา 00:00 ถึงวันที่ 7 มีนาคม 2568 เวลา 00:00 มีทั้งสิ้น 518,401 ตัวอย่าง จำนวน 6 ตัวแปร ได้แก่ 1.ค่าอัตราไหลของน้ำ (PV) 2.ค่าเป้าหมาย (SV) 3.ค่าควบคุม (MV) 4.ค่า Proportional (P) 5.ค่า Integral (I) 6.ค่า Derivative (D)

3) ใช้สมการมาตรฐาน PID ของ DCS แปรนด์ YOKOGAWA

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ได้แบบจำลองที่เหมาะสมสำหรับระบบควบคุม PID
- 2) เพิ่มประสิทธิภาพระบบควบคุม PID ให้มีการตอบสนองที่เร็วขึ้น
- 3) ลดความผันผวนในการควบคุมด้วย PID



บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบควบคุม PID

ระบบควบคุม PID มีองค์ประกอบทางคณิตศาสตร์ 3 อย่าง คือ 1.สัดส่วน (Proportional) 2.ปริพันธ์ (Integral) และ 3.อนุพันธ์ (Derivative) เป็นระบบควบคุมที่ใช้กันอย่างแพร่หลายในงานอุตสาหกรรมต่างๆ อาทิ เช่น กระบวนการผลิตที่ต้องควบคุมด้วยระบบอัตโนมัติหรือระบบที่ต้องการควบคุมให้ได้ผลลัพธ์ตามเป้าหมาย ระบบนี้มีความสามารถควบคุมได้อย่างแม่นยำและตอบสนองได้ดี โดยจะใช้งานแบบระบบปิด (Closed loop) กล่าวคือ จะต้องมีสัญญาณย้อนกลับ (Feedback Control) จากค่าที่วัดได้จริง (Process Variable) นำมาเปรียบเทียบกับค่าเป้าหมาย (Set Point) เพื่อคำนวณหาค่าผิดพลาด $e(t)$ ในชวงเวลานั้นๆ หลังจากนั้นจะส่งค่าผิดพลาดดังกล่าวไปยังกระบวนการทางคณิตศาสตร์ Proportional(P)-Integral(I)-Derivative(D) เพื่อคำนวณหาค่า PIDเอาต์พุต $u(t)$ จะได้สมการที่ (2.1)

$$u(t) = P_{out} + I_{out} + D_{out} \quad (2.1)$$

ส่วนแบบระบบเปิด (Open loop) จะไม่นิยมนำมาใช้งานควบคุม เนื่องจากไม่มีกระบวนการลดค่าผิดพลาด ทำให้ไม่สามารถควบคุมค่าได้ตามที่ต้องการ

2.1.1 ค่าสัดส่วน (Proportional)

จะเป็นการหาค่าสัดส่วนโดยตรงของค่าผิดพลาด $e(t)$ ซึ่งขนาดของค่าสัดส่วนจะถูกกำหนดโดยค่าคงที่อัตราขยายสัดส่วน (K_p) หน้าที่ช่วยให้ระบบตอบสนองได้รวดเร็ว แต่ถ้าปรับค่าสูงเกินไปอาจทำให้ระบบสั่นได้ง่าย และถ้าใช้งานค่าสัดส่วนเพียงอย่างเดียว จะไม่สามารถลดค่าผิดพลาดให้เป็นศูนย์ได้ จะได้สมการที่ (2.2)

$$P_{out} = K_p e(t) \quad (2.2)$$

2.1.2 ค่าปริพันธ์ (Integral)

จะเป็นการหาค่าผลรวมของความคลาดเคลื่อน $e(t)$ ในทุกช่วงเวลาก่อนหน้า จนถึงช่วงเวลาปัจจุบัน ซึ่งขนาดของค่าปริพันธ์จะถูกกำหนดโดยค่าคงที่อัตราขยายสัดส่วน (K_i) มีหน้าที่ช่วยให้ระบบ

ลดค่าผิดพลาดให้เข้าใกล้ศูนย์ได้ แต่ถ้าปรับค่าสูงเกินไปอาจทำให้ระบบเกิดโอเวอร์ชูต (Overshoot) หรือลดค่าผิดพลาดให้เข้าใกล้ศูนย์ได้ช้า จะได้สมการที่ (2.3)

$$I_{out} = K_i \int_0^t e(t) dt$$

(2.3)

2.1.3 ค่าอนุพันธ์ (Derivative)

จะเป็นการหาค่าอัตราการเปลี่ยนแปลงของค่าผิดพลาด $e(t)$ ในทุกๆช่วงเวลา หรือก็คือการหาค่าอนุพันธ์อันดับหนึ่งของค่าผิดพลาด ที่สัมพันธ์กับเวลาในช่วงนั้นๆ ซึ่งขนาดของค่าอนุพันธ์จะถูกกำหนดโดยค่าคงที่อัตราขยายสัดส่วน (K_d) หน้าที่ช่วยให้ระบบชะลออัตราการเปลี่ยนแปลง สามารถลดการเกิดโอเวอร์ชูต (Overshoot) แต่ถ้าปรับค่าสูงเกินไปอาจทำให้ระบบมีความไวต่อสัญญาณรบกวน (Noise) ได้ง่าย เกิดความไม่เสถียรในระบบได้ จะได้สมการที่ (2.4)

$$D_{out} = K_d \frac{de(t)}{dt}$$

(2.4)

จะได้สมการพื้นฐาน ดังสมการที่ (2.5)

$$u(t) = K_p e(t) + K_i \int_0^t e(t) dt + K_d \frac{de(t)}{dt}$$

(2.5)

โดยที่

$u(t)$ คือ ค่าPIDเอาต์พุต

$e(t)$ คือ ค่าผิดพลาด

K_p, K_i, K_d คือ ค่าคงที่สัดส่วน (Proportional Gain), ค่าคงที่อินทิกรัล (Integral Gain),

ค่าคงที่ดิฟเฟอเรนเชียล (Derivative Gain) ตามลำดับ

จากสมการพื้นฐานจึงได้มีการปรับปรุงรูปแบบสมการให้อยู่ในรูปแบบมาตรฐาน โดยรวมค่าผิดพลาดทั้งหมดให้อยู่ด้วยกัน เพื่อให้สามารถนำไปใช้งานได้ง่าย จะได้สมการมาตรฐาน ดังสมการที่ (2.6)

$$u(t) = K_p \left(e(t) + \frac{1}{T_i} \int_0^t e(t) dt + T_d \frac{de(t)}{dt} \right) \quad (2.6)$$

โดยที่

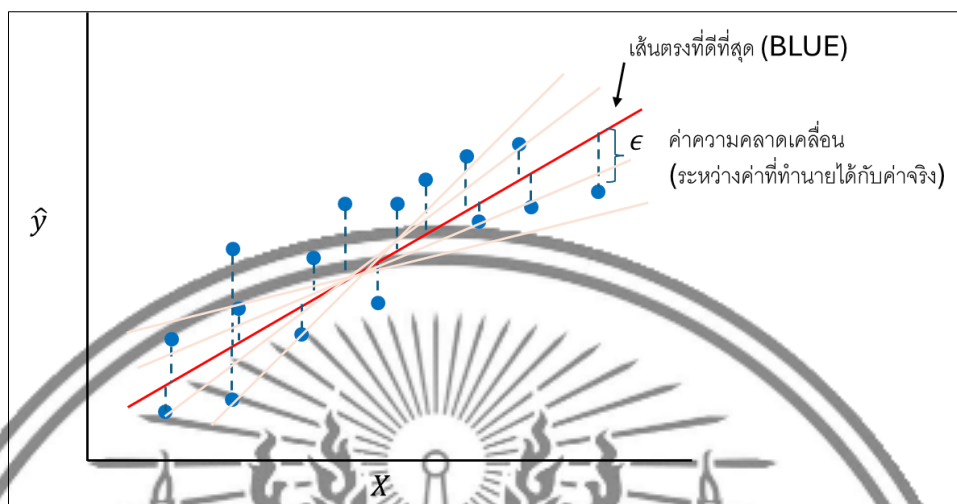
$$T_i = \frac{K_p}{K_i} \quad \text{คือ ระยะเวลาที่ใช้คำนวณค่าผิดพลาดในอดีตหรือใช้สำหรับ reset ระบบ}$$

$$T_d = \frac{K_d}{K_p} \quad \text{คือ ระยะเวลาที่ใช้คำนวณค่าผิดพลาดในอนาคต}$$

2.2 การถดถอยเชิงเส้น (Linear Regression)

การวิเคราะห์การถดถอยเชิงเส้น เป็นเทคนิคทางสถิติที่ใช้ในการศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวแปรขึ้นไป โดยแบ่งเป็นตัวแปรอิสระ (Independent Variable) และตัวแปรตาม (Dependent Variable) ซึ่งการวิเคราะห์การถดถอยเชิงเส้นมี 2 แบบ คือ 1.การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression) 2.การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ (Multiple Linear Regression)

แนวคิดในการวิเคราะห์การถดถอยเชิงเส้น จะเป็นการนำตัวแปรมาหาความสัมพันธ์ที่สามารถอธิบายแนวโน้มด้วยสมการเส้นตรง โดยจะมีเส้นตรงเพียงเส้นเดียวเรียกว่าเส้นตรงที่ดีที่สุด (Best Linear Unbiased Estimator, BLUE) เป็นเส้นตรงที่ใช้แทนความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามได้ดีที่สุดตามเกณฑ์ทางสถิติ หมายความว่าเส้นตรงเส้นนี้จะมีผลรวมของกำลังสองของค่าความคลาดเคลื่อน (Sum of Squared Residuals - SSR) น้อยที่สุด ดังรูปที่ 2.1



รูปที่ 2.1 แนวคิดการหาเส้นตรงที่ดีที่สุด

2.2.1 แบบจำลอง OLS (Ordinary Least Squares)

เป็นวิธีการประมาณค่าในแบบจำลองการถดถอยเชิงเส้น โดยการสร้างสมการเส้นตรงระหว่างตัวแปรอิสระ (Independent Variable) กับตัวแปรตาม (Dependent Variable) ใช้ได้กับตัวแปรอิสระตั้งแต่ 1 ตัวแปรขึ้นไป ในการศึกษาี้จะมีตัวแปรอิสระ 1 ตัวแปร จากนั้นคำนวณหาค่า β_0 และ β_1 ที่ทำให้ผลรวมกำลังสองของค่าความคลาดเคลื่อน (Sum of Squared Residuals, SSR) มีค่าน้อยที่สุด [2,5] ดังสมการที่ (2.7)

$$y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

(2.7)

โดยที่

y_i คือ ตัวแปรตามที่ i

X_i คือ ตัวแปรอิสระที่ i

β_0 คือ ค่าคงที่ (Intercept) หรือจุดตัดแกน y เมื่อ $X_j, \varepsilon = 0$

β_1 คือ สัมประสิทธิ์ความชัน (Slope coefficient)

ε_i คือ ค่าความคลาดเคลื่อน (Error term) ที่ i

ข้อสมมติฐานของวิธีการประมาณค่าด้วยวิธี OLS (Ordinary Least Squares) มีดังนี้ [2,5]

1) ความเป็นเส้นตรงของแบบจำลอง (Linearity) คือ ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามต้องเป็นเชิงเส้น ทดสอบได้ด้วยการสร้างแผนภาพการกระจายตัวระหว่างค่าความคลาดเคลื่อน (residuals) และค่าทำนาย (fitted values) ถ้ามีการกระจายแบบสุ่ม ไม่มีรูปแบบโค้ง แสดงว่ามีความเป็นเส้นตรง (Linearity)

2) ค่าความคลาดเคลื่อนเป็นอิสระต่อกัน (No Autocorrelation) คือ ค่าความคลาดเคลื่อนของแต่ละข้อมูลต้องไม่สัมพันธ์กัน

$$Cov(\varepsilon_i, \varepsilon_j) = 0 \text{ เมื่อ } i \neq j$$

ทดสอบได้ด้วยวิธีต่างๆ เช่น วิธี Durbin-Watson โดยกำหนดเกณฑ์ในการวัดค่า ดังนี้

- มีค่าอยู่ในช่วง 1.5 – 2.5 แสดงว่ามีความเป็นอิสระ (No Autocorrelation)
- มีค่าอยู่ในช่วง 2.6 – 4.0 แสดงว่ามีสหสัมพันธ์ในตัวเองเชิงลบ
- มีค่าอยู่ในช่วง 0 – 1.4 แสดงว่ามีสหสัมพันธ์ในตัวเองเชิงบวก

3) ค่าความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity) คือ ค่าความคลาดเคลื่อนต้องมีความแปรปรวนคงที่ตลอดช่วงของข้อมูล

$$Var(\varepsilon_i) = \sigma^2$$

ทดสอบได้ด้วยวิธีต่างๆ เช่น การสร้างแผนภาพการกระจายของค่าความคลาดเคลื่อน (residuals) กับค่าทำนาย (fitted values) หรือทดสอบทางสถิติด้วย Breusch-Pagan เพื่อดูว่าความแปรปรวนของ residuals มีความสัมพันธ์กับค่าพยากรณ์หรือไม่ โดยตั้งสมมติฐาน H_0 : Homoscedasticity ถ้า p-value น้อยกว่า 0.05 ให้ปฏิเสธ H_0 หมายความว่า ค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroscedasticity)

4) ค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติ (Normality of Errors) คือ ความคลาดเคลื่อนต้องมีการแจกแจงเป็นโค้งปกติ ทดสอบได้ด้วยวิธีต่างๆ เช่น การสร้างแผนภาพการกระจาย Q-Q Plot (Quantile-Quantile Plot) ถ้าค่าความคลาดเคลื่อน (residuals) เรียงตัวกันเป็นเส้นตรง

แสดงว่าค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติ (Normality) หรือทดสอบทางสถิติด้วย Shapiro-Wilk ในตัวอย่างขนาดไม่เกิน 5,000 ตัวอย่าง เพื่อดูว่าค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติหรือไม่ โดยตั้งสมมติฐาน H_0 : Normality ถ้า p-value น้อยกว่า 0.05 ให้ปฏิเสธ H_0 หมายความว่า ค่าความคลาดเคลื่อนมีการแจกแจงแบบไม่ปกติ (Not Normality)

การสร้างแบบจำลอง OLS มีดังนี้ [2,5]

- 1) กำหนดผลรวมของกำลังสองของค่าความคลาดเคลื่อน (SSR) เป็นฟังก์ชันเป้าหมาย
- 2) หาค่าของ β_0 และ β_1 ที่ทำให้ SSR มีค่าน้อยที่สุด จากอนุพันธ์ย่อยของฟังก์ชัน SSR

อนุพันธ์ย่อยของฟังก์ชัน SSR เทียบกับ β_0 [2,5] จะได้สมการที่ (2.8)

$$\frac{\partial SSR}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 X_i)$$

(2.8)

อนุพันธ์ย่อยของฟังก์ชัน SSR เทียบกับ β_1 [2,5] จะได้สมการที่ (2.9)

$$\frac{\partial SSR}{\partial \beta_1} = -2 \sum_{i=1}^n X_i (y_i - \beta_0 - \beta_1 X_i)$$

(2.9)

- 3) กำหนดจุดต่ำสุดของฟังก์ชัน SSR ของอนุพันธ์ย่อยแต่ละตัว ให้เท่ากับศูนย์
- 4) แก้ระบบสมการเชิงเส้นสองตัวแปรของ β_0 และ β_1

ประมาณค่า $\hat{\beta}_1$ [2,5] ได้จากสมการ (2.10)

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(y_i - \bar{y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (2.10)$$

โดยที่ \bar{X}, \bar{y} คือ ค่าเฉลี่ยของ X , ค่าเฉลี่ยของ y ตามลำดับ

ประมาณค่า $\hat{\beta}_0$ [2,5] ได้จากสมการ (2.11)

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{X} \quad (2.11)$$

5) สร้างสมการเส้นตรงจากค่า $\hat{\beta}_0$ และ $\hat{\beta}_1$ [2,5] ได้จากสมการ (2.12)

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X \quad (2.12)$$

6) หาค่าความคลาดเคลื่อน (Residual) [2,5] ได้จากสมการ (2.13)

$$\varepsilon = y - \hat{y} = y - (\hat{\beta}_0 + \hat{\beta}_1 X) \quad (2.13)$$

7) หาค่าผลรวมของกำลังสองของค่าความคลาดเคลื่อนที่น้อยที่สุด [2,5] ได้จากสมการ (2.14)

$$SSR = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 X)^2 \quad (2.14)$$

การแก้ไขปัญหา Heteroskedasticity และ Autocorrelation ใน OLS

เลือกใช้วิธี Heteroskedasticity and Autocorrelation Consistent (HAC) ซึ่งจะคำนวณค่าความคลาดเคลื่อนมาตรฐานของแบบจำลอง OLS ให้ใหม่ ทนต่อปัญหาค่าความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ (Heteroskedasticity) และทนต่อปัญหาค่าความคลาดเคลื่อนไม่เป็นอิสระ (Autocorrelation) เพื่อช่วยให้การทดสอบสมมติฐานและช่วงความเชื่อมั่น (Confidence Intervals) มีความน่าเชื่อถือมากขึ้น โดยที่ตัวประมาณค่าสัมประสิทธิ์การถดถอยในแบบจำลอง OLS ยังคงเหมือนเดิม

ขั้นตอนการคำนวณแบบ HAC มีดังนี้ [5]

1) กำหนดแบบจำลอง และคำนวณหาค่า u_t ด้วยวิธี OLS [5] จากสมการที่ (2.15)

$$y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + u_t \quad (2.15)$$

2) กำหนดสมการช่วย และคำนวณหาค่า r_t ด้วยวิธี OLS [5] จากสมการที่ (2.16)

$$x_{t1} = \delta_0 + \delta_2 x_{t2} + \dots + \delta_k x_{tk} + r_t \quad (2.16)$$

3) คำนวณหาค่า \hat{a}_t [5] จากสมการที่ (2.17)

$$\hat{a}_t = \hat{r}_t \cdot \hat{u}_t \quad (2.17)$$

4) คำนวณหาค่า \hat{v} [5] จากสมการที่ (2.18)

$$\hat{v} = \sum_{t=1}^n \hat{a}_t^2 + 2 \sum_{h=1}^g \left(1 - \frac{h}{g+1}\right) \sum_{t=h+1}^n \hat{a}_t \hat{a}_{t-h} \quad (2.18)$$

แนวทางของ Newey-West แนะนำให้ใช้ค่า g คือ จำนวนเต็มของ $4 \cdot \left(\frac{n}{100}\right)^{\frac{2}{9}}$

5) คำนวณหาค่า Standard Error ใหม่ [5] จากสมการที่ (2.19)

$$se(\hat{\beta}_1) = \left[\frac{se(\hat{\beta}_1)}{\hat{\sigma}} \right]^2 \sqrt{\hat{v}} \quad (2.19)$$

2.2.2 แบบจำลอง GLS (Generalized Least Squares)

เป็นวิธีการประมาณค่าที่พัฒนามาจาก OLS โดยมีวัตถุประสงค์เพื่อขจัดปัญหาค่าความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ (Heteroskedasticity) หรือค่าความคลาดเคลื่อนไม่อิสระต่อกัน (Autocorrelation) ซึ่งอาจทำให้การประมาณค่าด้วย OLS ขาดความน่าเชื่อถือ

GLS ใช้การประมาณหรือกำหนดโครงสร้างของเมทริกซ์ความแปรปรวนร่วมของค่าความคลาดเคลื่อน (Covariance Matrix, Ω) เพื่อแปลงสมการของ OLS ให้เป็นสมการใหม่ โดยจะมีค่าความคลาดเคลื่อนมาตรฐานใหม่ที่มีคุณสมบัติสอดคล้องกับข้อสมมติฐานของ OLS คือ มีความแปรปรวนคงที่ และเป็นอิสระต่อกัน [2,3,4,5] ดังสมการที่ (2.20)

$$(\hat{\beta}_{GLS}) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$$

(2.20)

โดยที่

$\hat{\beta}_{GLS}$ คือ เวกเตอร์ประมาณค่าพารามิเตอร์ $\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$

Ω คือ เมทริกซ์ความแปรปรวนร่วมของค่าความคลาดเคลื่อน ε

กรณี Heteroskedasticity ค่า σ_i^2 ในแต่ละข้อมูลบนแนวเส้นทแยงมุมจะไม่เท่ากัน

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}$$

กรณี Autocorrelation จะมีค่านอกแนวเส้นทแยงมุมไม่เท่ากับ 0

$$\Omega = \begin{bmatrix} \sigma_1^2 & p & p^2 & \dots & p^{T-1} \\ p & \sigma_1^2 & p & \dots & p^{T-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ p^{T-1} & p^{T-2} & p^{T-3} & \dots & \sigma_1^2 \end{bmatrix}$$

2.2.3 แบบจำลอง GLSAR (GLS with Autoregressive)

เป็นวิธีการประมาณค่าที่พัฒนามาจาก GLS เพื่อขจัดปัญหาค่าความคลาดเคลื่อนไม่เป็นอิสระต่อกัน (Autocorrelation) ของข้อมูลอนุกรมเวลา โดยใช้วิธีการกำหนดเมทริกซ์ความแปรปรวนร่วมของค่าความคลาดเคลื่อน (Ω) ให้มีโครงสร้างแบบ Autoregressive แล้วนำไปผ่านกระบวนการแปลงข้อมูล ซึ่งจะไม่มีการแสดงในสมการ เนื่องจากข้อมูลผ่านกระบวนการแปลงในขั้นตอน Autoregressive เรียบร้อยแล้ว [2,3,4,5] ดังสมการที่ (2.21)

$$(\hat{\beta}_{GLSAR}) = (X^{*'} X^*)^{-1} X^{*'} Y^* \quad (2.21)$$

โดยที่

$\hat{\beta}_{GLSAR}$ คือ เวกเตอร์ประมาณค่าพารามิเตอร์ $\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$

X^* คือ เมทริกซ์ X^* ที่ผ่านการแปลงข้อมูลจากขั้นตอน Autoregressive (AR)

Y^* คือ เวกเตอร์ Y^* ที่ผ่านการแปลงข้อมูลจากขั้นตอน Autoregressive (AR)

ขั้นตอนการแปลง Autoregressive (AR) [2,3,4,5]

1) กำหนดแบบจำลอง

$$y_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + u_t \quad \text{โดยที่ } u_t = \rho u_{t-1} + e_t$$

2) ใช้ OLS ประมาณค่าแบบจำลองในข้อ 1 จะได้ความคลาดเคลื่อน u_t

3) ใช้ OLS ประมาณค่า $\hat{\rho}$ จาก

$$\hat{u}_t = \hat{\rho} \hat{u}_{t-1} + e_t$$

4) ใช้ $\hat{\rho}$ แปลงข้อมูลด้วยวิธี Prais-Winsten

เมื่อข้อมูล Time Series ที่ $t = 1$

$$x_{j1}^* = \sqrt{1 - \hat{\rho}^2} \cdot x_{j1} \quad \text{เมื่อ } j=1,2,\dots,k$$

$$y_1^* = \sqrt{1 - \hat{\rho}^2} \cdot y_1$$

เมื่อข้อมูล Time Series ที่ $t \geq 2$

$$x_{jt}^* = x_{jt} - \hat{\rho} x_{j,t-1} \quad \text{เมื่อ } j=1,2,\dots,k$$

$$y_t^* = y_t - \hat{\rho} y_{t-1}$$

ดังนั้น จะได้

$$X^* = \begin{bmatrix} \sqrt{1 - \hat{p}^2} & \sqrt{1 - \hat{p}^2} \cdot x_{11} & \cdots & \sqrt{1 - \hat{p}^2} \cdot x_{1k} \\ 1 & x_{21} - \hat{p}x_{11} & \cdots & x_{2k} - \hat{p}x_{1k} \\ 1 & x_{31} - \hat{p}x_{21} & \cdots & x_{3k} - \hat{p}x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} - \hat{p}x_{n-1,1} & \cdots & x_{nk} - \hat{p}x_{n-1,k} \end{bmatrix}$$

สำหรับ $t \geq 2$ คอลัมน์แรกคือค่าคงที่ (intercept) จะได้เท่ากับ 1

$$Y^* = \begin{bmatrix} \sqrt{1 - \hat{p}^2} \cdot y_1 \\ y_2 - \hat{p}y_1 \\ y_3 - \hat{p}y_2 \\ \vdots \\ y_n - \hat{p}y_{n-1} \end{bmatrix}$$

2.2.4 แบบจำลอง WLS (Weighted Least Squares)

เป็นวิธีการประมาณค่าที่พัฒนาจาก GLS เพื่อขจัดปัญหาค่าความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ (Heteroskedasticity) โดยใช้หลักการถ่วงน้ำหนักข้อมูลแต่ละจุด แนวคิดคือ ข้อมูลที่มีค่าความคลาดเคลื่อนน้อย จะได้รับการถ่วงน้ำหนักที่สูง [2,3,4,5] ดังสมการที่ (2.22)

$$(\hat{\beta}_{WLS}) = (X'WX)^{-1}X'WY \quad (2.22)$$

โดยที่

$\hat{\beta}_{WLS}$ คือ เวกเตอร์ประมาณค่าพารามิเตอร์ $\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$

W คือ เมทริกซ์น้ำหนักแบบแนวทแยง ขนาด $n \times n$ ที่มีน้ำหนัก w_i อยู่บนแนวทแยง

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_i \end{bmatrix}$$

โดยทั่วไปแล้ว น้ำหนัก (w_i) ที่ใช้ในการถ่วงน้ำหนักคือ ค่าผกผันของความแปรปรวนของค่าความคลาดเคลื่อน (σ_i^2) ของแต่ละข้อมูล [4,5] ดังสมการที่ (2.23)

$$w_i = \frac{1}{\sigma_i^2} \quad (2.23)$$

ปัญหาสำคัญในการใช้ WLS คือ ไม่ทราบค่า σ_i^2 ของแต่ละข้อมูล ดังนั้นต้องทำการหาค่าประมาณ σ_i^2 จากการคำนวณค่าความคลาดเคลื่อน residuals ด้วย OLS Regression ก่อน จากนั้นคำนวณหาค่า w_i และนำไปคำนวณหาค่า β_{WLS} ต่อ ในสมการที่ (2.22)

2.2.5 การทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอย

เป็นกระบวนการทางสถิติเพื่อตรวจสอบว่าตัวแปรอิสระแต่ละตัวในแบบจำลองการถดถอยมีความสัมพันธ์ที่มีนัยสำคัญทางสถิติกับตัวแปรตามหรือไม่ สมมติฐานที่ใช้ในการทดสอบ

$$H_0: \beta_n = 0$$

$$H_1: \beta_n \neq 0$$

สมมติฐานหลัก (Null Hypothesis, H_0) สัมประสิทธิ์การถดถอยมีค่าเท่ากับศูนย์ หมายความว่าตัวแปรอิสระนั้นไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตาม

สมมติฐานทางเลือก (Alternative Hypothesis, H_1) สัมประสิทธิ์การถดถอยมีค่าไม่เท่ากับศูนย์ หมายความว่าตัวแปรอิสระนั้นมีความสัมพันธ์เชิงเส้นกับตัวแปรตาม

โดยพิจารณาตัดสินใจจาก

- 1) ค่า p-value น้อยกว่าระดับนัยสำคัญ (α) จะปฏิเสธสมมติฐาน H_0 ยอมรับ H_1 หมายความว่า สัมประสิทธิ์การถดถอยมีค่าไม่เท่ากับศูนย์ ตัวแปรอิสระนั้นมีความสัมพันธ์เชิงเส้นกับตัวแปรตาม
- 2) ค่า p-value มากกว่าระดับนัยสำคัญ (α) จะยอมรับ H_0 หมายความว่า สัมประสิทธิ์การถดถอยมีค่าเท่ากับศูนย์ ตัวแปรอิสระนั้นไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตาม

2.2.6 การประเมินผลประสิทธิภาพตัวแบบ

ในการศึกษาค้นคว้าอิสระในครั้งนี้ จะใช้วิธีการ Root Mean Squared Error (RMSE) และ R-squared (R^2) เพื่อวัดผลประสิทธิภาพของตัวแบบจำลอง โดยมีตัวแปรอิสระคือค่า PIDเอาต์พุต $u(t)$ และค่าตัวแปรตามคือค่าอัตราไหลของน้ำ (PV)

Root Mean Squared Error (RMSE) หรือรากที่สองของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน นิยมใช้ในการประเมินความแม่นยำของแบบจำลองการถดถอย โดยพื้นฐานแล้ว RMSE จะวัดขนาดเฉลี่ยของความแตกต่างระหว่างค่าที่ทำนายได้กับค่าจริง [2] ดังสมการที่ (2.24)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (2.24)$$

โดยที่

\hat{y} คือ ค่าที่ทำนายได้จากตัวแบบจำลอง

y คือ ค่าที่วัดได้จริง

R-squared (R^2) เป็นค่าทางสถิติที่ใช้วัดสัดส่วนของความแปรปรวนในตัวแปรตาม (Dependent Variable) ที่สามารถอธิบายได้ด้วยตัวแปรอิสระ (Independent Variable) ในแบบจำลองการถดถอยเชิงเส้น [2] ดังสมการที่ (2.25)

ถ้าเข้าใกล้ 1 (หรือใกล้ 100%) แสดงว่าแบบจำลองการถดถอยสามารถอธิบายความแปรปรวนในตัวแปรตามได้ หรือตัวแปรอิสระสามารถทำนายการเปลี่ยนแปลงของตัวแปรตามได้ดี

ถ้าเข้าใกล้ 0 (หรือใกล้ 0%) แสดงว่าแบบจำลองการถดถอยไม่สามารถอธิบายความแปรปรวนในตัวแปรตามได้ หรือตัวแปรอิสระไม่สามารถทำนายการเปลี่ยนแปลงของตัวแปรตามได้ดี

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (2.25)$$

โดยที่

SS_{res} คือ ผลรวมของกำลังสองของค่าความคลาดเคลื่อน ดังสมการที่ (2.26)

$$SS_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.26)$$

SS_{tot} คือ ผลรวมของกำลังสองทั้งหมด ดังสมการที่ (2.27)

$$SS_{tot} = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (2.27)$$

2.3 งานวิจัยที่เกี่ยวข้อง

2.3.1 Virgantari, F., Widyastiti, M., & Lifa, N. M. [5] ได้ทำการนำเสนอเกี่ยวกับการเลือกค่าน้ำหนักที่ดีที่สุด ใน WLS ระหว่าง $1/X$, $1/\sqrt{X}$, $1/Y$, และ $1/\sigma^2$ โดยใช้ข้อมูลจากดัชนีการพัฒนามนุษย์ (HDI) ของอินโดนีเซีย

2.3.2 Yildiz, K. M. [8] ได้ทำการนำเสนอการออกแบบตัวควบคุม PID แบบปรับตัวได้โดยใช้โมเดลเครือข่ายประสาทเทียม (NeuralPID) เพื่อควบคุมระบบกันสะเทือนของรถยนต์ โดยทำนายพารามิเตอร์ Proportional (P) เพื่อให้ได้การตอบสนองที่เหมาะสมกับสภาพถนนที่แตกต่างกัน

2.3.3 Yang, Q., Chen, G., Guo, M., Chen, T., Luo, L., & Sun, L. [10] ได้ทำการนำเสนอการควบคุมแบบไฮบริด โดยใช้แบบจำลองการทำนายกับตัวควบคุม PID เพื่อตอบสนองการเปลี่ยนแปลงโหลดที่รวดเร็วและสภาพแวดล้อมการทำงานที่ซับซ้อนของเครื่องผลิตพลังงานแบบวิกฤตยิ่งยวด

2.3.4 Agnihotri, S. P., & Waghmare, L. M. [11] ได้ทำการนำเสนอแบบจำลองการถดถอย สำหรับการปรับจูนตัวควบคุม PID ด้วยระบบที่มีลำดับเศษส่วนและหน่วงเวลา (Fractional Order Time Delay System)

2.3.5 Thaker, M. K., & Patel, V. P. [12] ได้ทำการนำเสนอวิธีการปรับจูนตัวควบคุมสัดส่วน-ปริพันธ์-อนุพันธ์ (PID) โดยใช้แนวทางการเรียนรู้ของเครื่อง (Machine Learning) และการหาค่าที่เหมาะสม (Optimization) โดยใช้อัลกอริธึม Multiple Linear Regression จาก Machine Learning และอัลกอริธึมการหาค่าที่เหมาะสมแบบ Gradient Descent เพื่อให้ได้ค่าพารามิเตอร์ PID



บทที่ 3

วิธีการดำเนินงานวิจัย

การศึกษานี้ทำการศึกษาการเพิ่มประสิทธิภาพการควบคุมระบบตัวควบคุมแบบสัดส่วนร่วมกับปริพันธ์และอนุพันธ์ ด้วยการประยุกต์เทคนิคโมเดลพยากรณ์ มีขั้นตอนการศึกษาดังนี้

3.1 ศึกษาการทำงานของระบบ

ในการศึกษานี้ใช้สมการมาตรฐาน PID ที่มาจากระบบควบคุม DCS (Distributed Control System) ยี่ห้อ YOKOGAWA รุ่น CentumVP เนื่องจากผู้ศึกษาจะนำแบบจำลองไปประยุกต์ใช้งานกับ DCS ยี่ห้อนี้โดยเฉพาะ

สมการมาตรฐาน PID ของยี่ห้อ YOKOGAWA [1]

$$MV = \frac{100}{P} \left(e(t) + \frac{1}{I} \int_0^t e(t) dt + D \frac{de(t)}{dt} \right)$$

โดยที่

MV คือ ค่าควบคุม

$e(t)$ คือ ค่าผิดพลาด ณ เวลา t

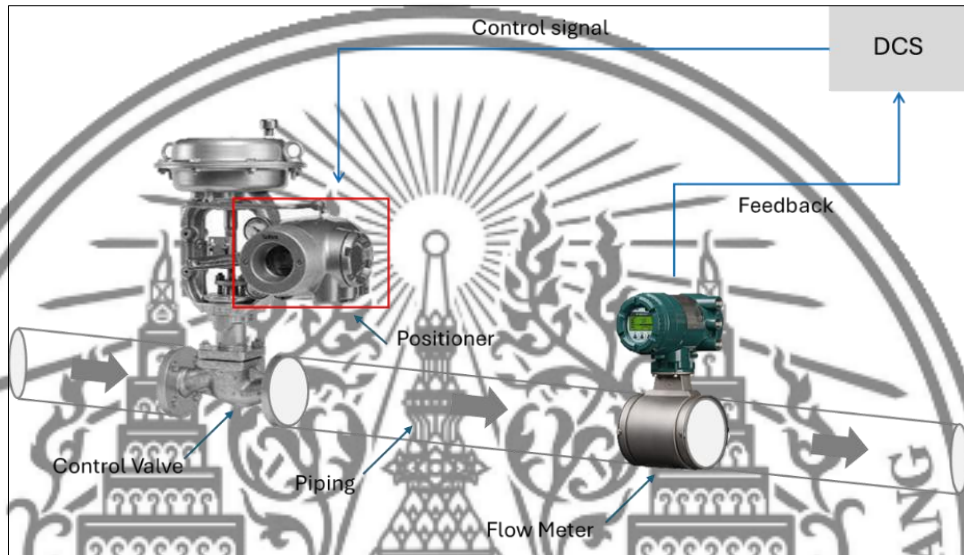
P คือ ค่าคงที่สัดส่วน

I คือ ค่าคงที่ระยะเวลาที่ใช้คำนวณค่าผิดพลาดในอดีต หรือใช้สำหรับ reset ระบบ

D คือ ค่าคงที่ระยะเวลาที่ใช้คำนวณค่าผิดพลาดในอนาคต

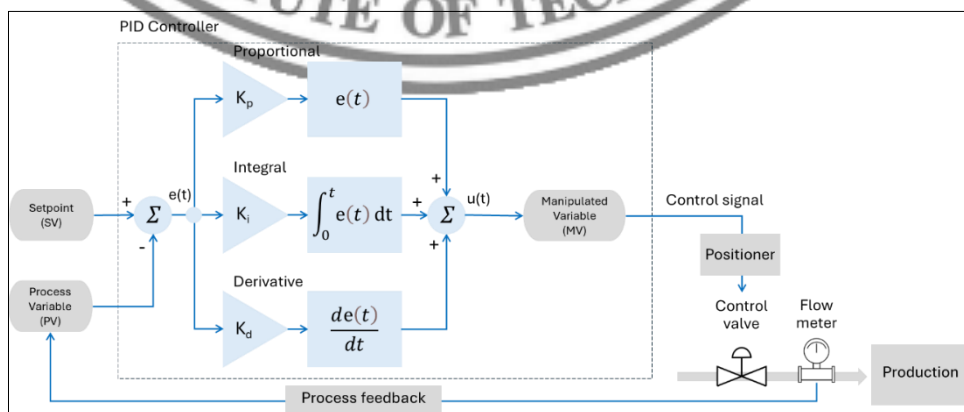
ใช้กระบวนการควบคุมอัตราไหลของน้ำเป็นต้นแบบในการศึกษา ดังรูปที่ 3.1 โดยมีอุปกรณ์ที่เรียกว่า คอนโทรลวาล์ว (Control Valve) เป็นตัวควบคุมอัตราไหลของน้ำ และทำการวัดค่าด้วยอุปกรณ์ที่เรียกว่า เครื่องมือวัดอัตราไหล (Flow Meter) ค่าอัตราไหลของน้ำ (PV) จะถูกส่งกลับไปยัง

(Feedback) ระบบควบคุม DCS จากนั้นระบบ DCS จะส่งค่าควบคุม (MV) ไปยังคอนโทรลวาล์ว (Control Valve)



รูปที่ 3.1 กระบวนการควบคุมอัตราไหลของน้ำ

ภายในของระบบ DCS จะคำนวณค่า PID เอาต์พุต $u(t)$ จากระบบควบคุม PID โดยมีวัตถุประสงค์เพื่อให้ค่า PV ใกล้เคียงค่าเป้าหมาย (SV) จากนั้นจะแปลงเป็นค่าควบคุม (MV) และส่งสัญญาณในรูปแบบทางไฟฟ้า เช่น 4-20 mA , 0-10Vdc ไปยังโพซิชั่นเนอร์ (Positioner) เพื่อให้คอนโทรลวาล์ว (Control Valve) ทำงานตามค่าควบคุมที่ตั้งไว้ในระบบ DCS ดังรูปที่ 3.2

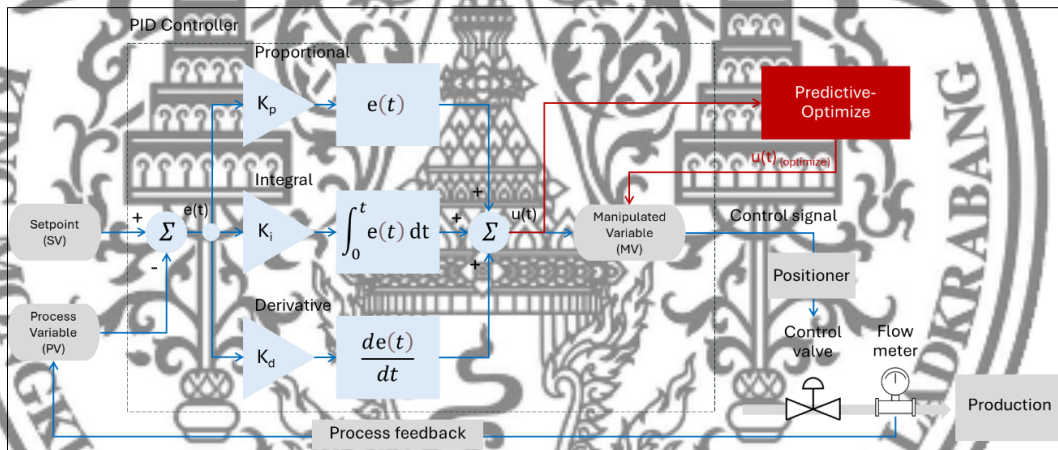


รูปที่ 3.2 บล็อกไดอะแกรมแสดงการทำงานภายในของระบบ DCS

3.2 แนวคิดในการใช้แบบจำลองพยากรณ์ในระบบควบคุม PID

อัตราไหลของน้ำที่ตั้งค่าเป้าหมายไว้ที่ 520 ลบ.ม./ชม. เมื่อใช้ระบบควบคุม PID พบว่าอัตราไหลของน้ำที่วัดได้มีค่าผันผวน (Fluctuation) อยู่ในช่วงประมาณ 514 - 526 ลบ.ม./ชม. ดังตารางที่ 3.3 ดังนั้นผู้ศึกษาจึงมีแนวคิดในการประยุกต์ใช้เทคนิคโมเดลพยากรณ์ เพื่อลดความผันผวนนี้ โดยมีตัวแปรอิสระคือค่า PIDเอาต์พุต $u(t)$ และค่าตัวแปรตามคือค่าอัตราไหลของน้ำ (PV)

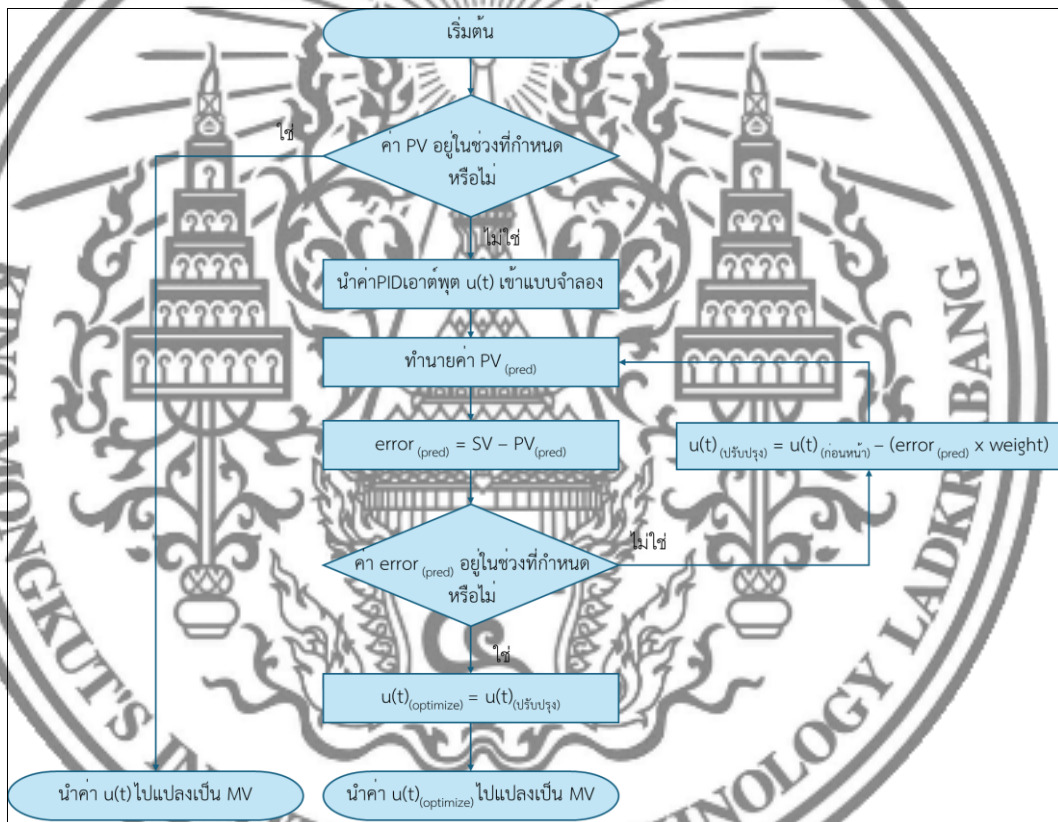
แนวคิดนี้จะนำค่า PIDเอาต์พุต $u(t)$ มาทำนายค่าอัตราไหลของน้ำ (PV) ถ้าหากพบว่าค่าทำนายมีค่าที่สูงหรือน้อยกว่าเกณฑ์ที่ตั้งไว้ จะทำการปรับปรุงค่า PIDเอาต์พุต $u(t)$ ซึ่งจะได้ค่าใหม่คือ $u(t)_{(optimize)}$ แล้วนำไปแปลงเป็นค่า MV ส่งไปควบคุมคอนโทรลลวาล์วต่อไป ดังรูปที่ 3.3



รูปที่ 3.3 บล็อกไดอะแกรมแนวคิดในการประยุกต์แบบจำลองพยากรณ์

การทำงานของกระบวนการ Predictive-Optimize จะทำงานลักษณะ Feedforward control แบบไม่ต่อเนื่อง เมื่อใดที่การทำงานของกระบวนการ Predictive-Optimize จะทำงานเมื่อพบว่าค่าอัตราไหลของน้ำ (PV) ไม่อยู่ในช่วงควบคุมที่กำหนดไว้ แต่ถ้าหากค่า PV ยังอยู่ควบคุมอยู่ในช่วง กระบวนการ Predictive-Optimize จะไม่ทำงาน ทำให้ระบบ PID ยังคงทำงานเป็นระบบหลักในการควบคุมเหมือนเดิม เพราะต้องการลดความเสี่ยงในการควบคุมล้มเหลวจากกระบวนการทำนายค่าควบคุม

ในขณะที่กระบวนการ Predictive-Optimize ทำงาน จะนำค่า PID เอาต์พุต $u(t)$ ไปทำนายค่า $PV_{(pred)}$ และคำนวณหาค่า $error_{(pred)}$ เทียบกับค่าเป้าหมาย(SV) หากพบว่าค่า $error_{(pred)}$ ยังไม่อยู่ในช่วงที่กำหนดไว้ จะทำการปรับปรุงค่า $u(t)$ ไปเรื่อยๆ จนกว่าค่า $error_{(pred)}$ จะอยู่ในช่วงที่กำหนดไว้ เมื่อได้ค่า $error_{(pred)}$ อยู่ในช่วงที่กำหนดแล้ว จะนำค่า $u(t)$ ดังกล่าว ส่งไปแปลงเป็นค่า MV เพื่อใช้ควบคุมต่อไป กระบวนการ Predictive-Optimize มีแนวคิดตามผังการทำงาน ดังรูปที่ 3.4



รูปที่ 3.4 ผังการทำงานของกระบวนการ Predictive-Optimize

3.3 การรวบรวมข้อมูล

การศึกษานี้ได้ใช้ข้อมูลจากโรงงานอุตสาหกรรมแห่งหนึ่งในประเทศไทย ซึ่งมีการจัดเก็บข้อมูลแบบเรียลไทม์ (Real-Time Data) บันทึกลงในฐานข้อมูล (Database) โดยผู้ศึกษาจะใช้ข้อมูลทุกๆ 1 วินาที จากวันที่ 1 มีนาคม 2568 เวลา 00:00 ถึงวันที่ 7 มีนาคม 2568 เวลา 00:00 โดยจะมีข้อมูลทั้งสิ้น 518,401 ตัวอย่างและมี 6 ตัวแปร ดังตารางที่ 3.1

ตารางที่ 3.1 คุณลักษณะข้อมูลดิบ

ลำดับที่	คอลัมน์	ลำดับที่	คอลัมน์
1	ค่าอัตราไหลของน้ำ (PV)	4	ค่า Proportional (P)
2	ค่าเป้าหมาย (SV)	5	ค่า Integral (I)
3	ค่าควบคุม (MV)	6	ค่า Derivative (D)

3.4 การจัดการข้อมูล

3.4.1 การสำรวจข้อมูล (Data Exploration)

1) ข้อมูลดิบที่ใช้ศึกษามีทั้งสิ้น 518,401 ตัวอย่าง และมี 6 ตัวแปร ดังรูปที่ 3.5

DateTime	PV	SV	MV	P	I	D
01-Mar-25 00:00:01	517.86	520.00	52.85	100	0	20
01-Mar-25 00:00:02	517.87	520.00	52.97	100	0	20
01-Mar-25 00:00:03	517.95	520.00	53.12	100	0	20
01-Mar-25 00:00:04	518.02	520.00	53.26	100	0	20
01-Mar-25 00:00:05	518.10	520.00	53.40	100	0	20
01-Mar-25 00:00:06	518.18	520.00	53.54	100	0	20
01-Mar-25 00:00:07	518.26	520.00	53.55	100	0	20
01-Mar-25 00:00:08	518.34	520.00	53.53	100	0	20
01-Mar-25 00:00:09	518.42	520.00	53.51	100	0	20
01-Mar-25 00:00:10	518.50	520.00	53.49	100	0	20
...
...
...
06-Mar-25 23:59:51	522.85	520.00	56.56	100	0	20
06-Mar-25 23:59:52	522.48	520.00	56.58	100	0	20
06-Mar-25 23:59:53	522.11	520.00	56.60	100	0	20
06-Mar-25 23:59:54	521.74	520.00	56.63	100	0	20
06-Mar-25 23:59:55	521.37	520.00	56.65	100	0	20
06-Mar-25 23:59:56	521.11	520.00	56.65	100	0	20
06-Mar-25 23:59:57	521.23	520.00	56.56	100	0	20
06-Mar-25 23:59:58	521.34	520.00	56.48	100	0	20
06-Mar-25 23:59:59	521.45	520.00	56.39	100	0	20
07-Mar-25 00:00:00	521.57	520.00	56.30	100	0	20

รูปที่ 3.5 ตัวอย่างข้อมูลดิบ

ตารางที่ 3.2 สถิติข้อมูลดิบ

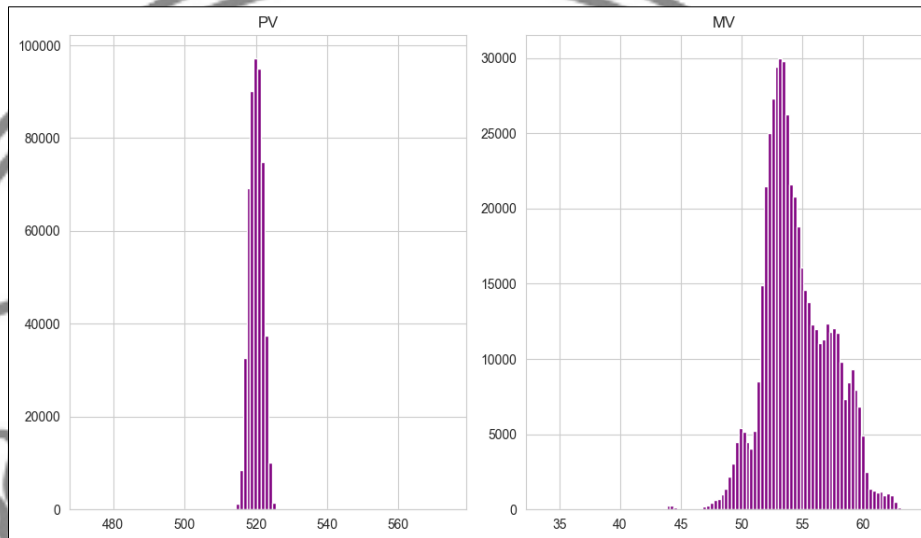
ลำดับ	รายการ	ค่าอัตรา ไหลของ น้ำ (PV)	ค่า เป้าหมาย (SV)	ค่า ควบคุม (MV)	ค่า P	ค่า I	ค่า D
1	ค่าเฉลี่ย	520.006	520.0	54.522	100.0	20.0	0.0
2	ค่าเบี่ยงเบน มาตรฐาน	1.875	0.0	2.754	0.0	0.0	0.0
3	ค่าต่ำสุด	472.864	520.0	33.716	100.0	20.0	0.0
4	ค่าควอร์ไทล์ 25%	518.623	520.0	52.643	100.0	20.0	0.0
5	ค่าควอร์ไทล์ 50%	520.023	520.0	53.998	100.0	20.0	0.0
6	ค่าควอร์ไทล์ 75%	521.385	520.0	56.455	100.0	20.0	0.0
7	ค่าสูงสุด	574.080	520.0	63.408	100.0	20.0	0.0

2) ไม่พบข้อมูลสูญหายใดๆ ในแต่ละคอลัมน์ของข้อมูลดิบ

DateTime	PV	SV	MV	P	I	D
No missing values found.						

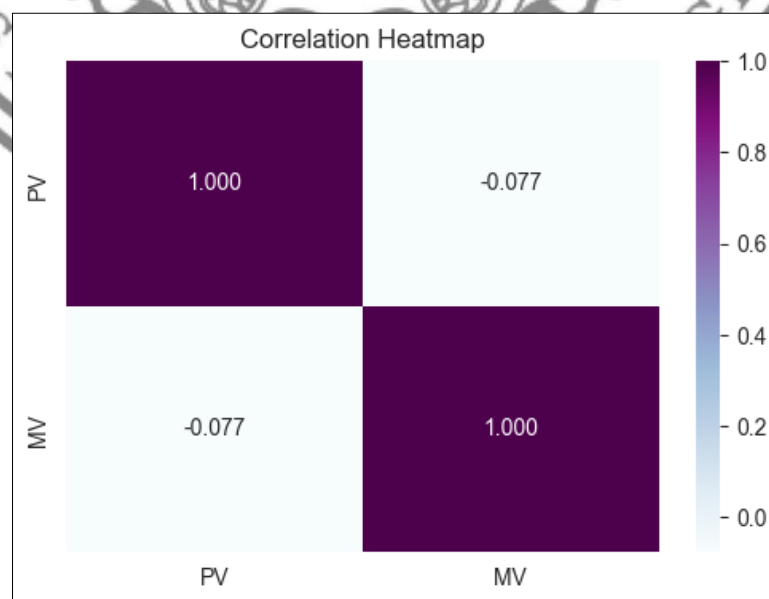
รูปที่ 3.6 ตรวจสอบข้อมูลหายในแต่ละคอลัมน์ของข้อมูลดิบ

3) สำรวจการกระจายตัวของข้อมูล ความสัมพันธ์ระหว่างตัวแปรของข้อมูล และค่าผิดปกติ เนื่องจากค่าเป้าหมาย (SV), ค่า P, ค่า I และค่า D เป็นค่าคงที่ ไม่มีการเปลี่ยนแปลง ดังนั้น ผู้ศึกษาจะไม่แสดงการสำรวจข้อมูลในส่วนนี้



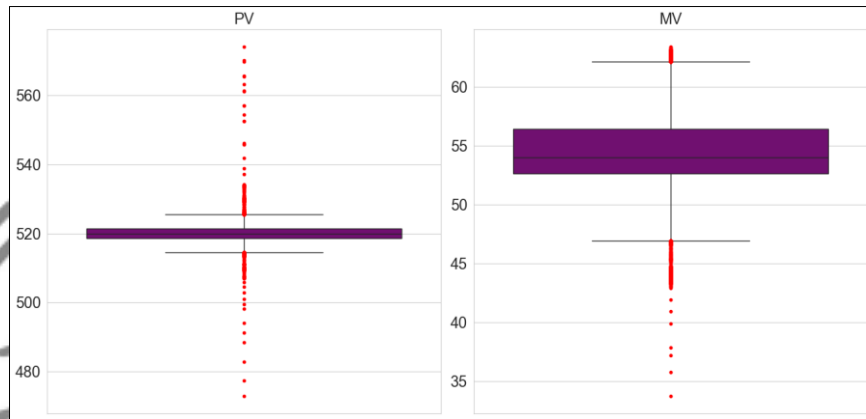
รูปที่ 3.7 การกระจายตัวของข้อมูลดิบ

เมื่อตรวจสอบความสัมพันธ์ระหว่างตัวแปร พบว่าตัวแปรไม่มีความสัมพันธ์กัน ดังรูปที่ 3.8



รูปที่ 3.8 ความสัมพันธ์ระหว่างตัวแปรของข้อมูลดิบ

เมื่อตรวจสอบข้อมูลดิบด้วยแผนภาพกล่องและใช้วิธีการพิสัยควอไทล์เพื่อตรวจสอบค่าผิดปกติ พบว่ามีค่าผิดปกติที่อยู่นอกขอบเขตบนและขอบเขตล่าง ดังรูปที่ 3.9



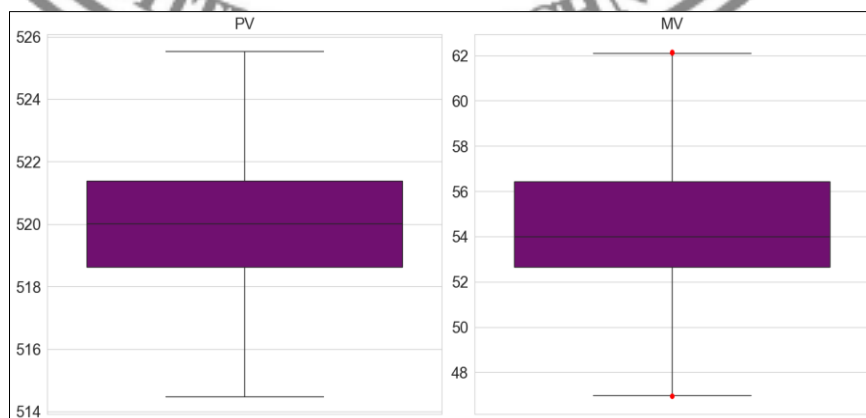
รูปที่ 3.9 แผนภาพกล่องของข้อมูลดิบ

3.4.2 การทำความสะอาดข้อมูล (Data Cleaning)

การจัดค่าผิดปกติเลือกใช้วิธีการพิสัยควอไทล์ (IQR Method) โดยค่าขอบเขตบนคำนวณจาก $Q3 + (IQR \cdot 1.5)$ และขอบเขตล่างคำนวณจาก $Q1 - (IQR \cdot 1.5)$ จะได้ขอบเขตดังนี้

- 1) ค่า PV มีขอบเขตบนคือ 525.529 และขอบเขตล่างคือ 514.479
- 2) ค่า MV มีขอบเขตบนคือ 62.176 และขอบเขตล่างคือ 46.925

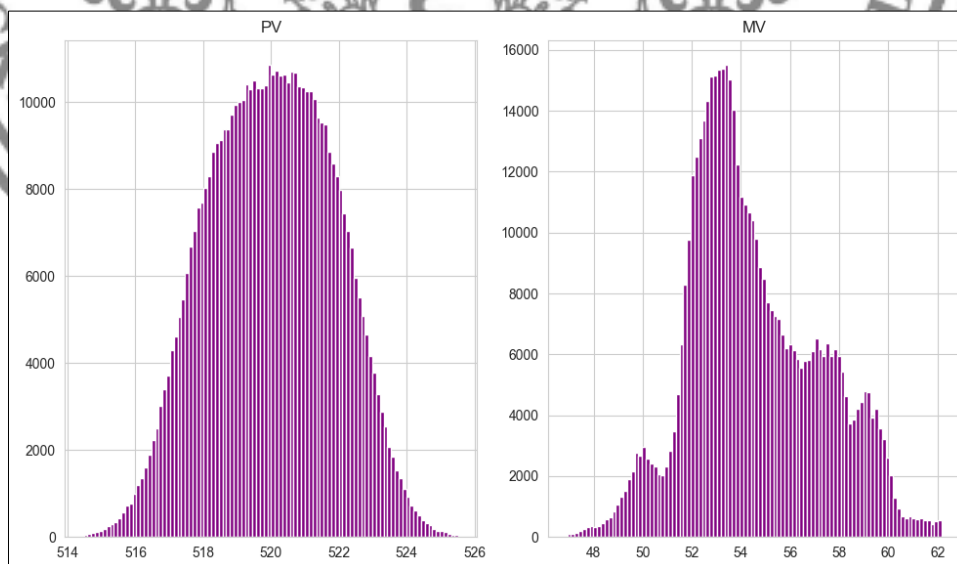
พบว่าจากเดิมมีข้อมูล 518,401 แถว หลังทำความสะอาดข้อมูลจะเหลือ 515,111 แถว (ลดลง 0.63%) จะไม่พบค่าผิดปกติในข้อมูลดิบ ดังรูปที่ 3.10 และค่าสถิติข้อมูลดิบ ดังตารางที่ 3.3



รูปที่ 3.10 แผนภาพกล่องของข้อมูลดิบ (หลังจัดค่าผิดปกติ)

ตารางที่ 3.3 สถิติข้อมูลดิบ (หลังขจัดค่าผิดปกติ)

ลำดับ	รายการ	ค่าอัตรา ไหลของ น้ำ (PV)	ค่า เป้าหมาย (SV)	ค่า ควบคุม (MV)	ค่า P	ค่า I	ค่า D
1	ค่าเฉลี่ย	520.005	520.0	54.518	100.0	20.0	0.0
2	ค่าเบี่ยงเบน มาตรฐาน	1.842	0.0	2.682	0.0	0.0	0.0
3	ค่าต่ำสุด	514.480	520.0	46.927	100.0	20.0	0.0
4	ค่าควอร์ไทล์ 25%	518.623	520.0	52.648	100.0	20.0	0.0
5	ค่าควอร์ไทล์ 50%	520.024	520.0	53.995	100.0	20.0	0.0
6	ค่าควอร์ไทล์ 75%	521.386	520.0	56.426	100.0	20.0	0.0
7	ค่าสูงสุด	525.524	520.0	62.175	100.0	20.0	0.0



รูปที่ 3.11 การกระจายตัวของข้อมูลดิบ (หลังขจัดค่าผิดปกติ)

3.4.3 การแปลงข้อมูล (Data Transformation)

ทำการการสร้างตัวแปรใหม่ (Feature Engineering) จากตัวแปรที่มีอยู่ ได้ดังนี้ 1.ค่าผิดพลาด $e(t)$ ซึ่งมาจากการนำค่า SV ลบด้วยค่า PV และ 2.ค่า PIDเอาต์พุต $u(t)$ ซึ่งมาจากการคำนวณจากสมการมาตรฐาน (3.1) จะได้ข้อมูลชุดใหม่ ดังรูปที่ 3.12

DateTime	PV	SV	MV	et	ut
2025-03-01 00:00:19	519.22	520.00	53.33	0.78	2.28
2025-03-01 00:00:20	519.30	520.00	53.32	0.70	2.16
2025-03-01 00:00:21	519.38	520.00	53.30	0.62	2.00
2025-03-01 00:00:22	519.46	520.00	53.28	0.54	1.84
2025-03-01 00:00:23	519.53	520.00	53.26	0.47	1.69
2025-03-01 00:00:24	519.61	520.00	53.25	0.39	1.53
2025-03-01 00:00:25	519.69	520.00	53.23	0.31	1.37
2025-03-01 00:00:26	519.77	520.00	53.21	0.23	1.21
2025-03-01 00:00:27	519.92	520.00	53.19	0.08	0.98
2025-03-01 00:00:28	520.07	520.00	53.18	-0.07	0.74
...
...
...
2025-03-06 23:59:51	522.85	520.00	56.56	-2.85	-4.52
2025-03-06 23:59:52	522.48	520.00	56.58	-2.48	-4.36
2025-03-06 23:59:53	522.11	520.00	56.60	-2.11	-4.17
2025-03-06 23:59:54	521.74	520.00	56.63	-1.74	-3.96
2025-03-06 23:59:55	521.37	520.00	56.65	-1.37	-3.72
2025-03-06 23:59:56	521.11	520.00	56.65	-1.11	-3.57
2025-03-06 23:59:57	521.23	520.00	56.56	-1.23	-3.76
2025-03-06 23:59:58	521.34	520.00	56.48	-1.34	-3.93
2025-03-06 23:59:59	521.45	520.00	56.39	-1.45	-4.08
2025-03-07 00:00:00	521.57	520.00	56.30	-1.57	-4.20

รูปที่ 3.12 ตัวอย่างข้อมูลที่ผ่านการแปลงแล้ว

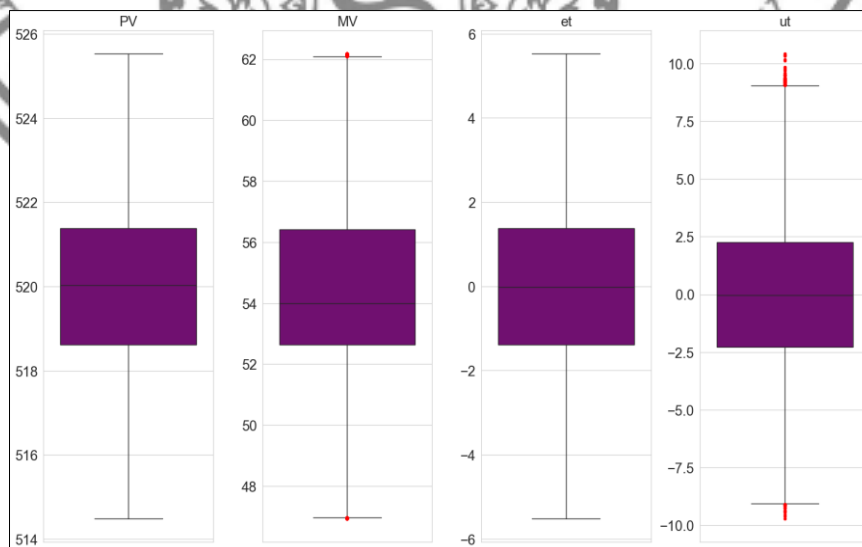
ตารางที่ 3.4 คุณลักษณะข้อมูลที่ผ่านการแปลงแล้ว

ลำดับที่	คอลัมน์	ลำดับที่	คอลัมน์
1	ค่าอัตราไหลของน้ำ (PV)	4	ค่าผิดพลาด $e(t)$
2	ค่าเป้าหมาย (SV)	5	ค่า PIDเอาต์พุต $u(t)$
3	ค่าควบคุม (MV)		

ตารางที่ 3.5 สถิติข้อมูลที่ผ่านการแปลงแล้ว

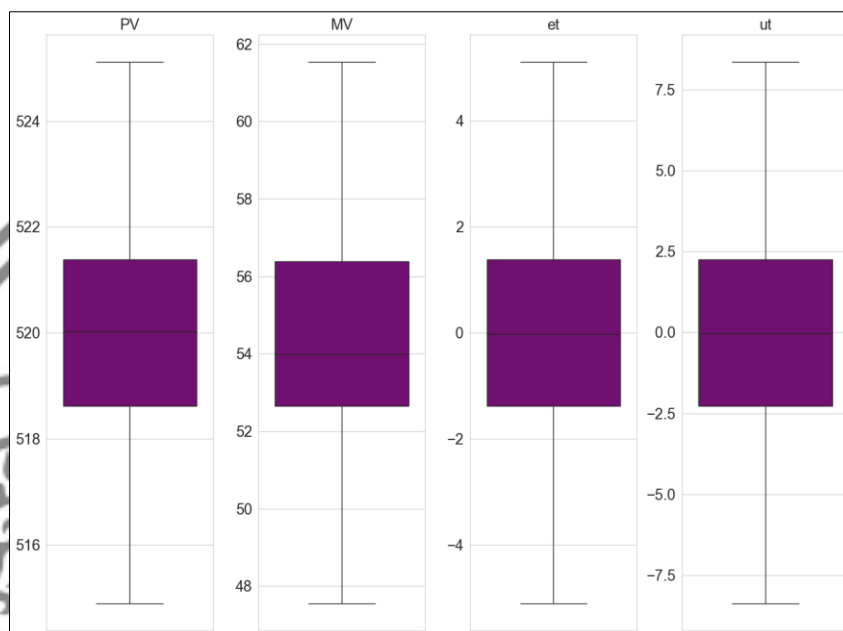
ลำดับ	รายการ	ค่าอัตราไหลของน้ำ (PV)	ค่าเป้าหมาย (SV)	ค่าควบคุม (MV)	ค่าผิดพลาด e(t)	ค่า PID เอาต์พุต u(t)
1	ค่าเฉลี่ย	520.005	520.0	54.518	-0.005	-0.009
2	ค่าเบี่ยงเบนมาตรฐาน	1.842	0.0	2.682	1.842	3.005
3	ค่าต่ำสุด	514.480	520.0	46.927	-5.524	-9.721
4	ค่าควอร์ไทล์ 25%	518.623	520.0	52.648	-1.385	-2.273
5	ค่าควอร์ไทล์ 50%	520.024	520.0	53.995	-0.024	-0.032
6	ค่าควอร์ไทล์ 75%	521.385	520.0	56.426	1.377	2.250
7	ค่าสูงสุด	525.524	520.0	62.175	5.520	10.411

เมื่อตรวจสอบข้อมูลที่ผ่านการแปลงแล้ว พบว่าค่า PID เอาต์พุต $u(t)$ มีค่าผิดปกติ ดังรูปที่ 3.13 ดังนั้น จึงทำการขจัดค่าผิดปกติ ด้วยวิธีการพิสัยควอร์ไทล์ (IQR Method) เช่นเดิม



รูปที่ 3.13 แผนภาพกล่องของข้อมูลที่ผ่านการแปลงแล้ว

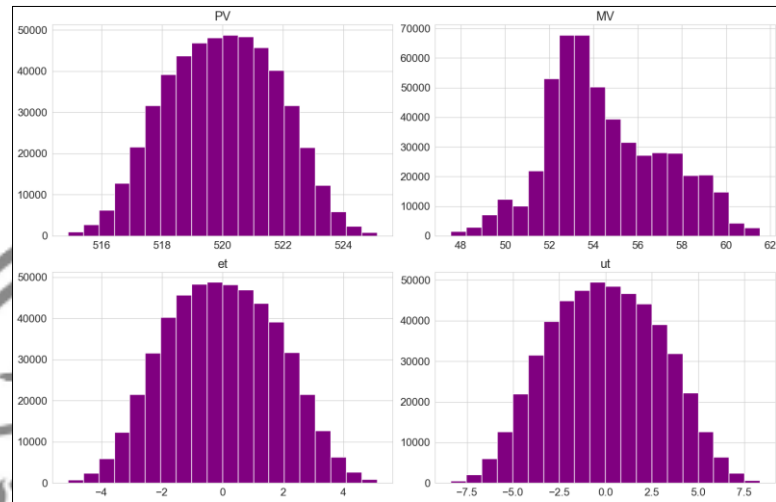
ผลจากการทำความสะอาดข้อมูลด้วยวิธีการพิสัยควอไทล์ พบว่า จำนวนข้อมูลจากเดิม 515,111 แถว ลดลงเหลือ 511,518 แถว (ลดลง 0.70%) แสดงแผนภาพกล่องข้อมูลได้ดังรูปที่ 3.14



รูปที่ 3.14 แผนภาพกล่องของข้อมูลที่ผ่านการแปลงแล้ว (หลังจัดค่าผิดปกติ)

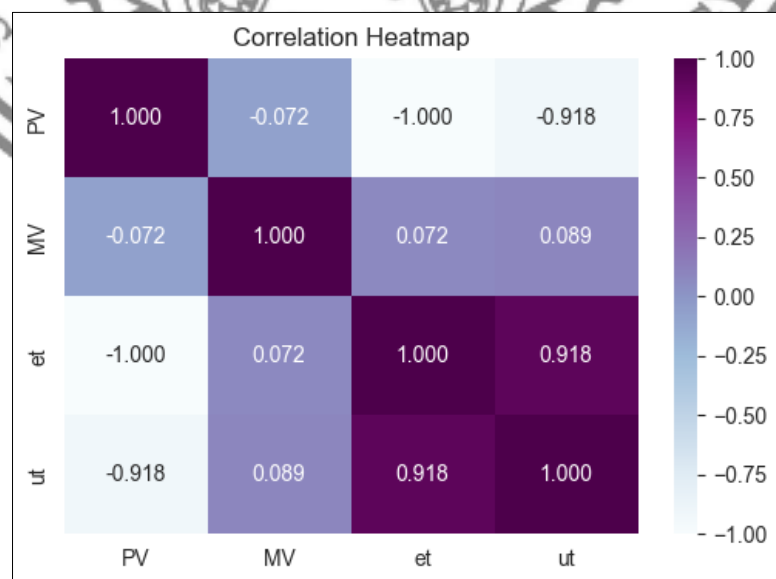
ตารางที่ 3.6 สถิติข้อมูลที่ผ่านการแปลงแล้ว (หลังจัดค่าผิดปกติ)

ลำดับ	รายการ	ค่าอัตราไหลของน้ำ (PV)	ค่าเป้าหมาย (SV)	ค่าควบคุม (MV)	ค่าผิดพลาด $e(t)$	ค่า PID เอาต์พุต $u(t)$
1	ค่าเฉลี่ย	520.004	520.0	54.499	-0.004	-0.008
2	ค่าเบี่ยงเบนมาตรฐาน	1.833	0.0	2.634	1.833	2.992
3	ค่าต่ำสุด	514.895	520.0	47.546	-5.111	-8.369
4	ค่าควอร์ไทล์ 25%	518.623	520.0	52.649	-1.383	-2.269
5	ค่าควอร์ไทล์ 50%	520.023	520.0	53.988	-0.023	-0.029
6	ค่าควอร์ไทล์ 75%	521.383	520.0	56.392	1.377	2.251
7	ค่าสูงสุด	525.111	520.0	61.534	5.105	8.345



รูปที่ 3.15 การกระจายตัวของข้อมูลที่ผ่านการแปลงแล้ว (หลังจัดค่าผิดปกติ)

เมื่อตรวจสอบความสัมพันธ์ระหว่างตัวแปรของข้อมูลที่ผ่านการแปลงแล้ว ดังรูปที่ 3.16 จะเห็นว่าค่า PIDเอาต์พุต $u(t)$ มีความสัมพันธ์กันสูงแบบแปรผกผันกับค่าอัตราไหลของน้ำ (PV) ที่ -0.918 ดังนั้น ผู้ศึกษาจึงพิจารณาเลือกการทำนายด้วยแบบจำลองการถดถอยเชิงเส้น (Linear Regression) โดยมีตัวแปรอิสระคือค่า PIDเอาต์พุต $u(t)$ และค่าตัวแปรตามคือค่าอัตราไหลของน้ำ (PV)



รูปที่ 3.16 ความสัมพันธ์ระหว่างตัวแปรของข้อมูลที่ผ่านการแปลงแล้ว

3.5 เครื่องมือที่ใช้

ตารางที่ 3.7 เครื่องมือที่ใช้งาน

ลำดับ	เครื่องมือ	คำอธิบาย
1	ซอฟต์แวร์ OSIsoft PI System	ซอฟต์แวร์สำหรับเก็บข้อมูลจากระบบ DCS, PLC, etc., รวมถึงการแสดงผลข้อมูลแบบเรียลไทม์
2	ซอฟต์แวร์ PyCharm	ซอฟต์แวร์สำหรับเขียนโปรแกรม Python
3	ซอฟต์แวร์ Microsoft Excel	ซอฟต์แวร์สำหรับจัดการข้อมูลเบื้องต้น
4	ไลบรารี Pandas ใน Python	สำหรับการจัดการและวิเคราะห์ข้อมูลที่มีโครงสร้าง (Structured Data)
5	ไลบรารี NumPy ใน Python	สำหรับการคำนวณทางคณิตศาสตร์ ดการกับอาร์เรย์ (Arrays) และเมทริกซ์ (Matrices)
6	ไลบรารี Matplotlib ใน Python	สำหรับการสร้างภาพข้อมูลแบบ 2 มิติ
7	ไลบรารี Seaborn ใน Python	สำหรับสร้างสถิติกราฟ (Statistical Graphics) ที่มีข้อมูลเชิงลึก
8	ไลบรารี Scikit-learn ใน Python	สำหรับสร้างการเรียนรู้ของเครื่อง (Machine Learning) และประเมินโมเดลการเรียนรู้ของเครื่อง
9	ไลบรารี statsmodels ใน Python	สำหรับการสร้างแบบจำลองทางสถิติ (statistical models), การทดสอบทางสถิติ (statistical tests) และการสำรวจข้อมูลเชิงสถิติ (statistical data exploration)

3.6 การแบ่งชุดข้อมูลสำหรับการเรียนรู้ของเครื่อง (Data Splitting)

3.6.1 ชุดฝึกฝน (Training Set)

แบ่งชุดข้อมูลสำหรับฝึกฝน 80% ของ 511,518 แถว ด้วยวิธีการสุ่ม จะได้ 409,214 แถว เพื่อใช้สร้างแบบจำลอง

3.6.2 ชุดทดสอบ (Testing Set)

แบ่งชุดข้อมูลสำหรับทดสอบ 20% ของ 511,518 แถว ด้วยวิธีการสุ่ม จะได้ 102,304 แถว เพื่อใช้ประเมินประสิทธิภาพของแบบจำลองที่ผ่านการเรียนรู้



บทที่ 4

ผลการวิจัยและการอภิปรายผล

ผลการศึกษการเพิ่มประสิทธิภาพการควบคุมระบบตัวควบคุมแบบสัดส่วนร่วมกับปริพันธ์ และอนุพันธ์ ด้วยการประยุกต์เทคนิคโมเดลพยากรณ์ โดยใช้ข้อมูลจากวันที่ 1 มีนาคม 2568 เวลา 00:00 ถึงวันที่ 7 มีนาคม 2568 เวลา 00:00 สร้างแบบจำลองถดถอยเชิงเส้น (Linear Regression) ด้วยวิธี OLS, OLS-HAC, GLS, GLSAR และ WLS จากนั้นเลือกแบบจำลองไปใช้ในระบบควบคุม PID ของคอนโทรลลาล์ว เพื่อควบคุมอัตราไหลของน้ำ มีผลการศึกษาดังนี้

4.1 แบบจำลอง OLS (Ordinary Least Squares)

สร้างแบบจำลองการถดถอยเชิงเส้น OLS ด้วยตัวแปรอิสระคือค่า PIDเอาต์พุต $u(t)$ และตัวแปรตามคือค่าอัตราไหลของน้ำ (PV) จะได้ผลลัพธ์แบบจำลอง OLS ดังตารางที่ 4.1

ตารางที่ 4.1 ผลลัพธ์แบบจำลอง OLS

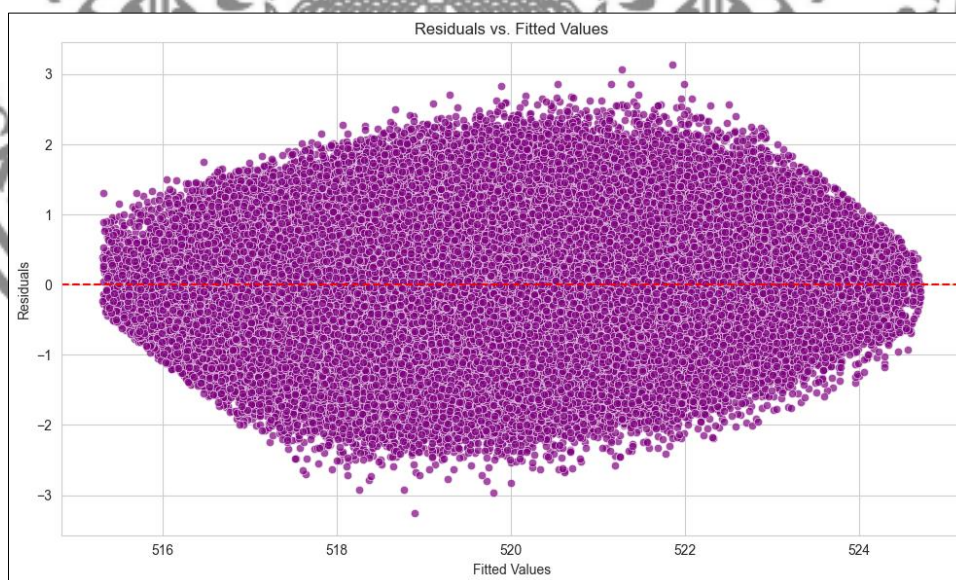
ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
1	R-squared	0.842	11	Observation	409,214
2	F-statistic	2.186×10^6	12	Covariance type	Non-robust
3	Prob(F-statistic)	0.00	13	Linearity	ใช่
4	Durbin-Watson	2.001	14	Normality	ใช่
5	Dep. Variable	PV	15	No autocorrelation	ใช่
6	Indep. Variable	ut	16	Homoscedasticity	ไม่ใช่
7	Coefficient (const)	519.9999	17	Coefficient (ut)	-0.5620

ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
8	Std.Error(const)	0.001	18	Std.Error(ut)	0.000
9	t-statistic (const)	4.57×10^5	19	t-statistic (ut)	-1478.457
10	P> t (const)	0.000	20	P> t (ut)	0.000

4.1.1 ตรวจสอบข้อสมมติฐานของแบบจำลองถดถอยเชิงเส้น OLS

1) ทดสอบความเป็นเส้นตรงของความคลาดเคลื่อน (Linearity)

เมื่อพล็อตภาพเปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนาย พบว่าค่าความคลาดเคลื่อนกระจายตัวรอบๆ เส้นศูนย์ หมายความว่า ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง ดังรูปที่ 4.1

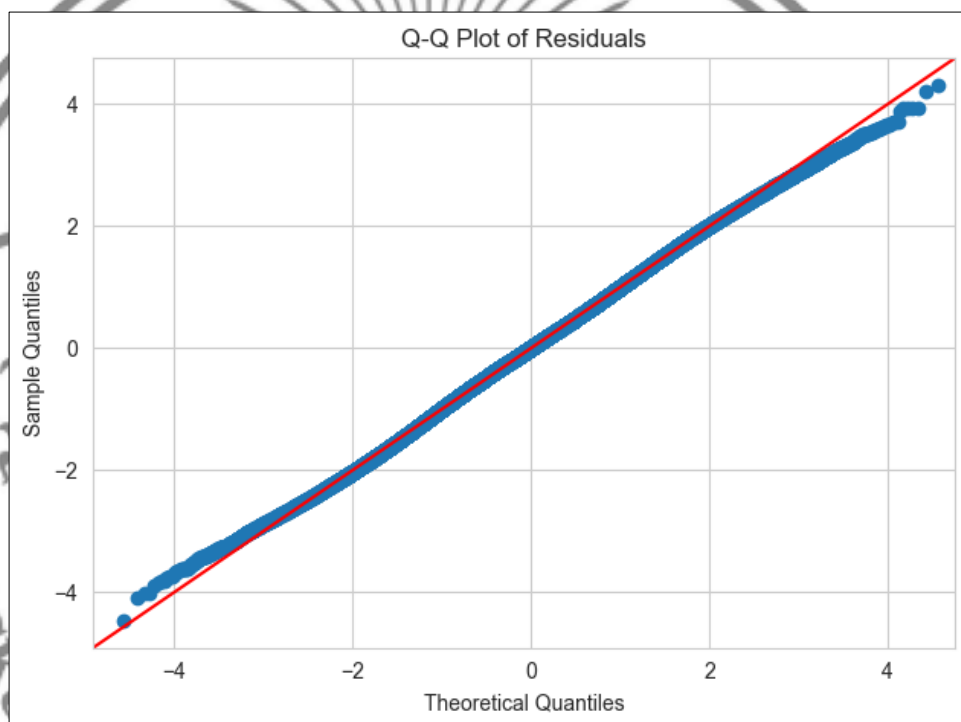


รูปที่ 4.1 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน OLS

เมื่อใช้เทคนิคทางสถิติ Harvey-Collier ในการทดสอบความเป็นเส้นตรง พบว่าค่า p-value เท่ากับ 0.9199 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง (Linearity)

2) ทดสอบการแจกแจงปกติของความคลาดเคลื่อน (Normality)

สร้างแผนภาพ Q-Q (Quantile-Quantile Plot) พบว่าค่าความคลาดเคลื่อนเรียงตัวกันเป็นเส้นตรงใกล้เคียงเส้นทแยงมุม ดังรูปที่ 4.2 ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ (Normality)



รูปที่ 4.2 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง OLS

เมื่อสุ่มตัวอย่างมา 5,000 ตัวอย่าง และใช้เทคนิคทางสถิติ Shapiro-Wilk ในการทดสอบความเป็น Normality พบว่าค่า p-value เท่ากับ 0.1275 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ (Normality)

3) ทดสอบความคลาดเคลื่อนมีความเป็นอิสระ (No autocorrelation)

จากตารางที่ 4.1 พบว่าค่าสถิติ Durbin-Watson เท่ากับ 2.001 หมายความว่า ค่าความคลาดเคลื่อนเป็นอิสระต่อกัน (No autocorrelation)

4) ทดสอบความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity)

เมื่อใช้เทคนิคทางสถิติ Breusch-Pagan พบว่าค่า p-value เท่ากับ < 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroskedasticity)

4.2 แบบจำลอง OLS ปรับปรุงด้วย HAC

จากปัญหาค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (heteroskedasticity) ที่พบในแบบจำลอง OLS จึงพิจารณาปรับปรุงแบบจำลองนี้ด้วยวิธี Heteroscedasticity and Autocorrelation Consistent (HAC) เพื่อขจัดปัญหาของค่าความคลาดเคลื่อนไม่เป็นอิสระต่อกัน (autocorrelation) และปัญหาความแปรปรวนไม่คงที่ (Heteroscedasticity) โดยไม่ทำให้ค่าประมาณของโมเดลเปลี่ยนไป แต่จะช่วยให้การทดสอบทางสถิติ เช่น t-test, F-test มีความน่าเชื่อถือมากขึ้นในกรณีที่เกิดปัญหาเหล่านี้ในแบบจำลอง จะได้ผลลัพธ์แบบจำลอง OLS-HAC ดังตารางที่ 4.2

ตารางที่ 4.2 ผลลัพธ์แบบจำลอง OLS-HAC

ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
1	R-squared	0.842	11	Observation	409,214
2	F-statistic	2.856×10^6	12	Covariance type	HAC
3	Prob(F-statistic)	0.00	13	Linearity	ไม่ทดสอบ
4	Durbin-Watson	2.001	14	Normality	ไม่ทดสอบ
5	Dep. Variable	PV	15	No autocorrelation	ไม่ทดสอบ
6	Indep. Variable	ut	16	Homoscedasticity	ไม่ทดสอบ
7	Coefficient (const)	519.9999	17	Coefficient (ut)	-0.5620

ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
8	Std.Error(const)	0.001	18	Std.Error(ut)	0.000
9	z-statistic (const)	4.57×10^5	19	z-statistic (ut)	-1690.020
10	$P > z $ (const)	0.000	20	$P > z $ (ut)	0.000

4.2.1 ตรวจสอบข้อสมมติฐานของแบบจำลองถดถอยเชิงเส้น OLS-HAC

ไม่ทำการทดสอบข้อสมมติฐาน เนื่องจากการปรับปรุงด้วย HAC ไม่ได้เปลี่ยนแปลงค่าใดๆ ในแบบจำลอง OLS เพียงแต่เป็นการเพิ่มความน่าเชื่อถือทางสถิติเท่านั้น

จากตารางที่ 4.3 สรุปได้ว่า เมื่อปรับปรุงแบบจำลอง OLS ด้วยวิธี HAC พบว่าค่าสัมประสิทธิ์การถดถอยไม่เปลี่ยนแปลง แต่ค่า F-statistic สูงขึ้นจาก 2.186×10^6 เป็น 2.856×10^6 แสดงว่าแบบจำลองมีความสามารถอธิบายความแปรปรวนของตัวแปรตามได้ดีขึ้น และค่าสถิติทดสอบของตัวแปร $u(t)$ เปลี่ยนจาก t-statistic เป็น z-statistic เนื่องจาก HAC เป็น asymptotic estimator และในเชิงค่าสัมบูรณ์มีค่าสูงขึ้น แสดงว่า ตัวแปร $u(t)$ มีอิทธิพลต่อค่าของตัวแปรตามอย่างมีนัยสำคัญ แม้ว่าจะมีปัญหา Heteroscedasticity

ตารางที่ 4.3 เปรียบเทียบแบบจำลอง OLS และแบบจำลอง OLS ที่ปรับปรุงด้วย HAC

พารามิเตอร์	แบบจำลอง OLS	แบบจำลอง OLS ปรับปรุงด้วย HAC	คำอธิบาย
R-squared	0.842	0.842	ไม่เปลี่ยนแปลง
F-statistic	2.186×10^6	2.856×10^6	ในแบบจำลอง OLS-HAC มีค่าเปลี่ยนแปลงสูงขึ้น
สัมประสิทธิ์ของค่าคงที่ const	519.9999	519.9999	ไม่เปลี่ยนแปลง
สัมประสิทธิ์ของตัวแปร u(t)	-0.5620	-0.5620	ไม่เปลี่ยนแปลง
ความคลาดเคลื่อนมาตรฐานของค่าคงที่ const	0.001	0.001	ไม่เปลี่ยนแปลง
ความคลาดเคลื่อนมาตรฐานของตัวแปร u(t)	0.000	0.000	ไม่เปลี่ยนแปลง
สถิติทดสอบของค่าคงที่ const	t = 4.57×10^5	z = 4.57×10^5	เปลี่ยนจาก t-statistic เป็น z-statistic เนื่องจาก HAC เป็น asymptotic estimator (ใช้กับตัวอย่างที่มีขนาดใหญ่มากๆ)
สถิติทดสอบของตัวแปร u(t)	t = -1478.457	z = -1690.020	เปลี่ยนจาก t-statistic เป็น z-statistic และพบว่าในเชิงค่าสัมบูรณ์ มีค่าสูงขึ้นในทิศทางเดียวกัน

4.3 แบบจำลอง GLS (Generalized Least Squares)

เนื่องจากแบบจำลองนี้ต้องใช้การคำนวณค่อนข้างมาก ดังนั้นจึงไม่เหมาะกับข้อมูลขนาดใหญ่ ซึ่งในการศึกษานี้ใช้ข้อมูลขนาดใหญ่ จึงจำเป็นต้องลดจำนวนตัวอย่างลง โดยการสุ่มมาจำนวน 30,000 ตัวอย่าง จากตัวอย่างเดิมที่มี 409,214 ตัวอย่าง เนื่องจากเป็นขนาดตัวอย่างที่คอมพิวเตอร์ของผู้ศึกษาสามารถรองรับได้ในการประมวลผล GLS จะได้ผลลัพธ์แบบจำลอง GLS ดังตารางที่ 4.4

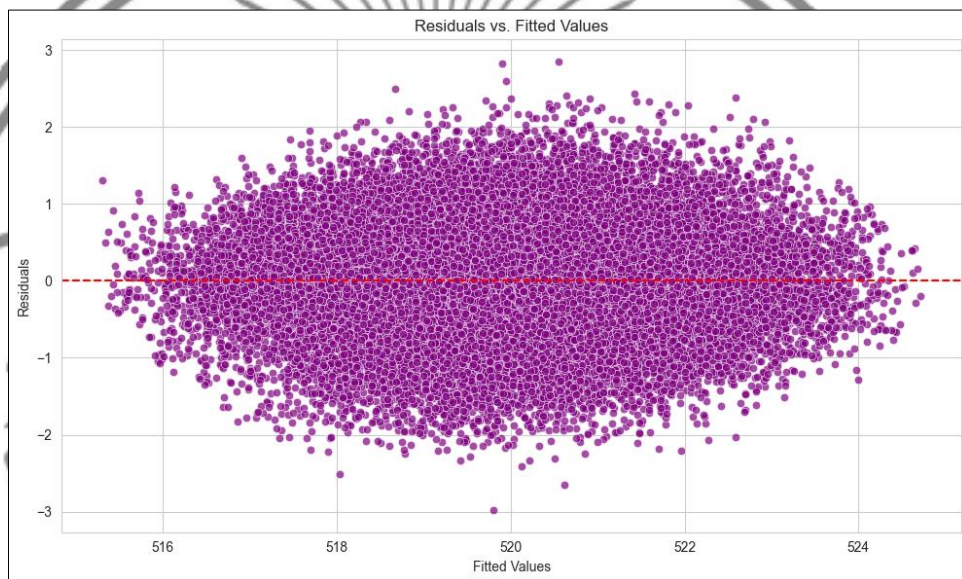
ตารางที่ 4.4 ผลลัพธ์แบบจำลอง GLS

ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
1	R-squared	0.844	11	Observation	30,000
2	F-statistic	1.627×10^5	12	Covariance type	Non-robust
3	Prob(F-statistic)	0.00	13	Linearity	ใช่
4	Durbin-Watson	1.999	14	Normality	ใช่
5	Dep. Variable	PV	15	No autocorrelation	ใช่
6	Indep. Variable	ut	16	Homoscedasticity	ใช่
7	Coefficient (const)	520.0065	17	Coefficient (ut)	-0.5637
8	Std.Error(const)	0.004	18	Std.Error(ut)	0.001
9	t-statistic (const)	1.24×10^5	19	t-statistic (ut)	-403.354
10	P> t (const)	0.000	20	P> t (ut)	0.000

4.3.1 ตรวจสอบข้อสมมติของแบบจำลองถดถอยเชิงเส้น GLS

1) ทดสอบความเป็นเส้นตรงของความคลาดเคลื่อน (Linearity)

เมื่อพล็อตภาพเปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนาย พบว่าค่าความคลาดเคลื่อนกระจายตัวรอบๆ เส้นศูนย์ หมายความว่า ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง ดังรูปที่ 4.3

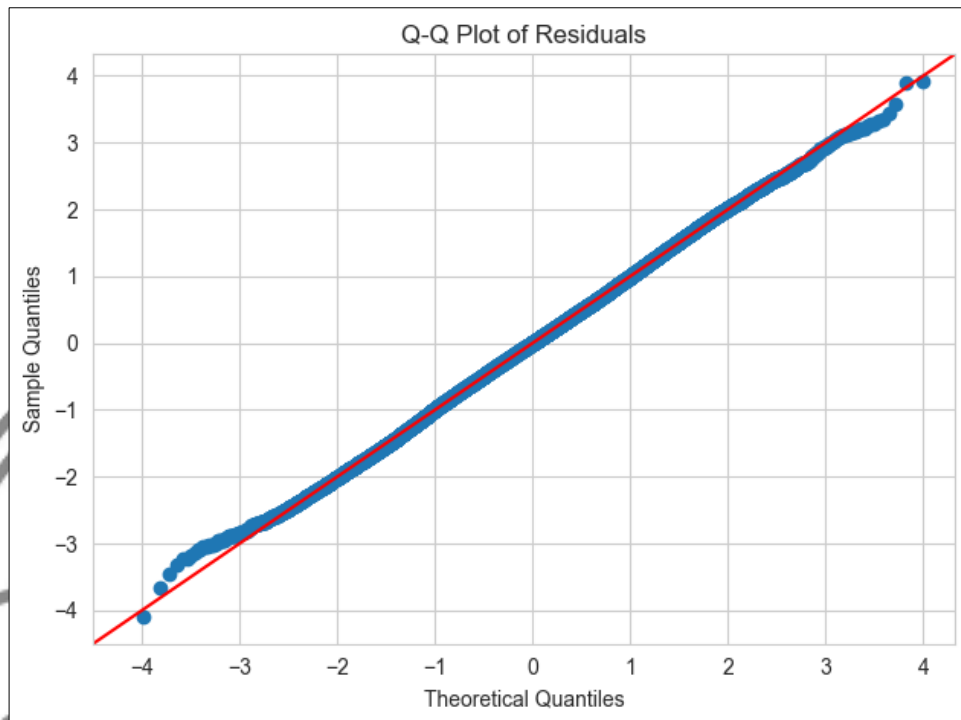


รูปที่ 4.3 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน GLS

เมื่อใช้เทคนิคทางสถิติ Harvey-Collier ในการทดสอบความเป็นเส้นตรง พบว่าค่า p-value เท่ากับ 0.2634 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง (Linearity)

2) ทดสอบการแจกแจงปกติของความคลาดเคลื่อน (Normality)

สร้างแผนภาพ Q-Q (Quantile-Quantile Plot) พบว่าค่าความคลาดเคลื่อนเรียงตัวกันเป็นเส้นตรงใกล้เคียงเส้นทแยงมุม ดังรูปที่ 4.4 ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ



รูปที่ 4.4 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง GLS

เมื่อสุ่มตัวอย่างมา 5,000 ตัวอย่าง และใช้เทคนิคทางสถิติ Shapiro-Wilk ในการทดสอบความเป็น Normality พบว่าค่า p-value เท่ากับ 0.0034 ซึ่งน้อยกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนมีแนวโน้มแจกแจงไม่ปกติ (Non-Normality) ซึ่งอาจเกิดจากการสุ่มตัวอย่างมาไม่เพียงพอ การศึกษานี้จึงพิจารณาผลจากแผนภาพ Q-Q แทน เนื่องจากครอบคลุมทุกตัวอย่างในแบบจำลอง

3) ทดสอบความคลาดเคลื่อนมีความเป็นอิสระ (No autocorrelation)

จากตารางที่ 4.4 พบว่าค่าสถิติ Durbin-Watson เท่ากับ 1.999 หมายความว่า ค่าความคลาดเคลื่อนเป็นอิสระต่อกัน (No autocorrelation)

4) ทดสอบความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity)

เมื่อใช้เทคนิคทางสถิติ Breusch-Pagan พบว่าค่า p-value เท่ากับ 0.2058 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ยอมรับสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity)

4.4 แบบจำลอง GLSAR (GLS with Autoregressive errors)

GLSAR พัฒนาต่อมาจาก GLS มีองค์ประกอบสำคัญคือ autoregressive (AR) เพื่อจัดการปัญหาค่าความคลาดเคลื่อนไม่เป็นอิสระต่อกัน (Autocorrelation) ก่อนการประมาณค่า ที่พบได้บ่อยในข้อมูลแบบ time series จะได้ผลลัพธ์แบบจำลอง GLSAR ดังตารางที่ 4.5

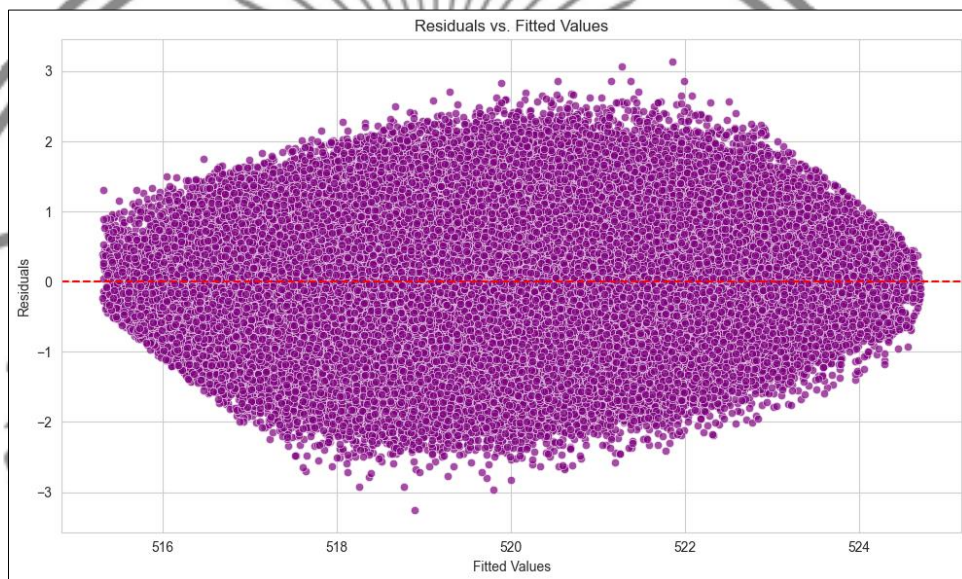
ตารางที่ 4.5 ผลลัพธ์แบบจำลอง GLSAR

ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
1	R-squared	0.842	11	Observation	409,213
2	F-statistic	2.186×10^6	12	Covariance type	Non-robust
3	Prob(F-statistic)	0.00	13	Linearity	ใช่
4	Durbin-Watson	2.000	14	Normality	ใช่
5	Dep. Variable	PV	15	No autocorrelation	ใช่
6	Indep. Variable	ut	16	Homoscedasticity	ไม่ใช่
7	Coefficient (const)	519.9999	17	Coefficient (ut)	-0.5620
8	Std.Error(const)	0.001	18	Std.Error(ut)	0.000
9	t-statistic (const)	4.58×10^5	19	t-statistic (ut)	-1478.449
10	P> t (const)	0.000	20	P> t (ut)	0.000

4.4.1 ตรวจสอบข้อสมมติของแบบจำลองถดถอยเชิงเส้น GLSAR

1) ทดสอบความเป็นเส้นตรงของความคลาดเคลื่อน (Linearity)

เมื่อพล็อตภาพเปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนาย พบว่าค่าความคลาดเคลื่อนกระจายตัวรอบๆ เส้นศูนย์ หมายความว่า ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง ดังรูปที่ 4.5

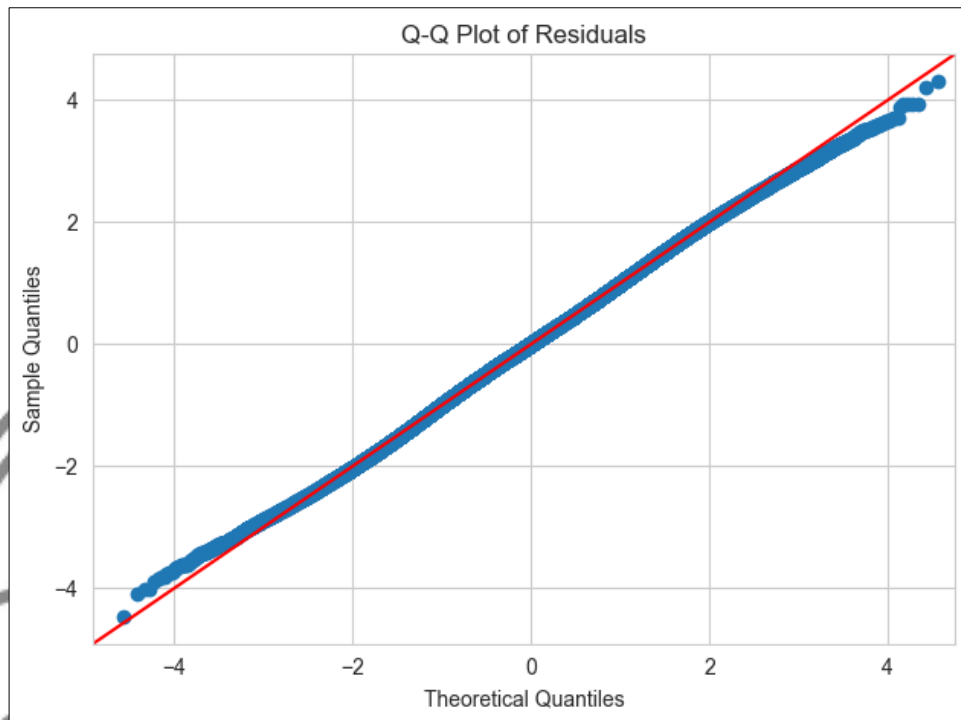


รูปที่ 4.5 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน GLSAR

เมื่อใช้เทคนิคทางสถิติ Harvey-Collier ในการทดสอบความเป็นเส้นตรง พบว่าค่า p-value เท่ากับ 0.9199 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง (Linearity)

2) ทดสอบการแจกแจงปกติของความคลาดเคลื่อน (Normality)

สร้างแผนภาพ Q-Q (Quantile-Quantile Plot) พบว่าค่าความคลาดเคลื่อนเรียงตัวกันเป็นเส้นตรงใกล้เคียงเส้นทแยงมุม ดังรูปที่ 4.6 ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ (Normality)



รูปที่ 4.6 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง GLSAR

เมื่อสุ่มตัวอย่างมา 5,000 ตัวอย่าง และใช้เทคนิคทางสถิติ Shapiro-Wilk ในการทดสอบความเป็น Normality พบว่าค่า p-value เท่ากับ 0.1275 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ (Normality)

3) ทดสอบความคลาดเคลื่อนมีความเป็นอิสระ (No autocorrelation)

จากตารางที่ 4.5 พบว่าค่าสถิติ Durbin-Watson เท่ากับ 2.000 หมายความว่า ค่าความคลาดเคลื่อนเป็นอิสระต่อกัน (No autocorrelation)

4) ทดสอบความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity)

เมื่อใช้เทคนิคทางสถิติ Breusch-Pagan พบว่าค่า p-value เท่ากับ < 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroskedasticity)

4.5 แบบจำลอง WLS (Weighted Least Squares)

การศึกษานี้จะกำหนดน้ำหนัก (weights) คือ ค่าผกผันของความแปรปรวนของค่าความคลาดเคลื่อน σ^2 [3] จะได้ผลลัพธ์แบบจำลอง WLS ดังตารางที่ 4.6

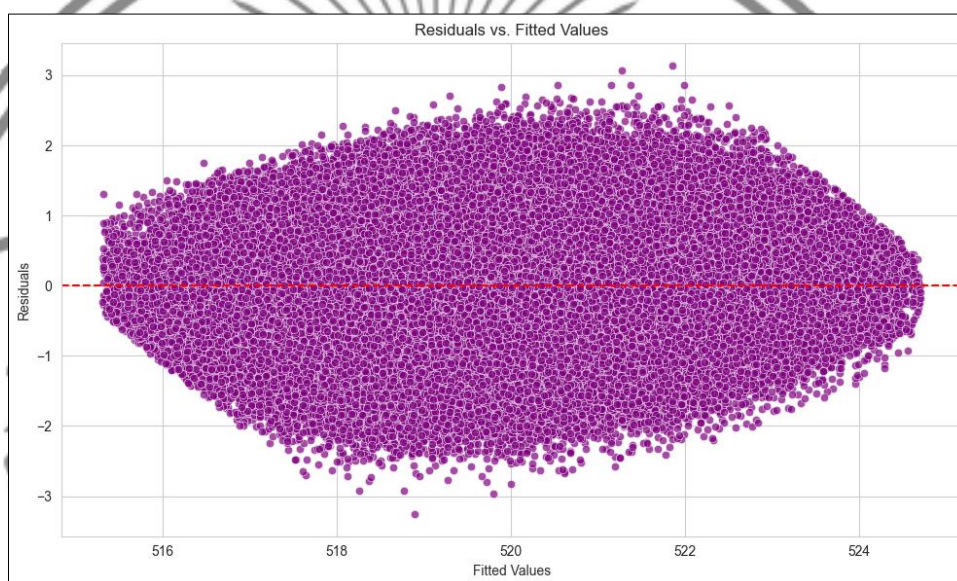
ตารางที่ 4.6 ผลลัพธ์แบบจำลอง WLS

ลำดับ	รายการ	ผลลัพธ์	ลำดับ	รายการ	ผลลัพธ์
1	R-squared	0.842	11	Observation	409,214
2	F-statistic	2.187×10^9	12	Covariance type	Non-robust
3	Prob(F-statistic)	0.00	13	Linearity	ใช่
4	Durbin-Watson	2.001	14	Normality	ใช่
5	Dep. Variable	PV	15	No autocorrelation	ใช่
6	Indep. Variable	ut	16	Homoscedasticity	ไม่ใช่
7	Coefficient (const)	519.9999	17	Coefficient (ut)	-0.5620
8	Std.Error(const)	0.001	18	Std.Error(ut)	0.000
9	t-statistic (const)	4.57×10^5	19	t-statistic (ut)	-1478.874
10	P> t (const)	0.000	20	P> t (ut)	0.000

4.5.1 ตรวจสอบข้อสมมติของแบบจำลองถดถอยเชิงเส้น WLS

1) ทดสอบความเป็นเส้นตรงของความคลาดเคลื่อน (Linearity)

เมื่อพล็อตภาพเปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนาย พบว่าค่าความคลาดเคลื่อนกระจายตัวรอบๆ เส้นศูนย์ หมายความว่า ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง ดังรูปที่ 4.7

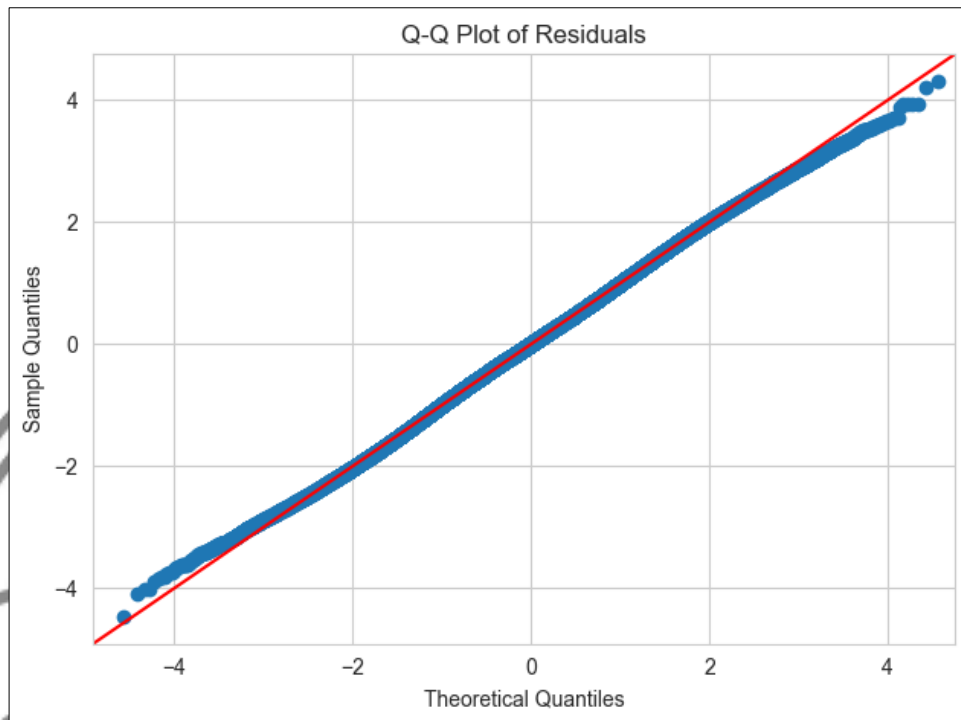


รูปที่ 4.7 เปรียบเทียบค่าความคลาดเคลื่อนกับค่าทำนายใน WLS

เมื่อใช้เทคนิคทางสถิติ Harvey-Collier ในการทดสอบความเป็นเส้นตรง พบว่าค่า p-value เท่ากับ 0.9199 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนมีความเป็นเส้นตรง (Linearity)

2) ทดสอบการแจกแจงปกติของความคลาดเคลื่อน (Normality)

สร้างแผนภาพ Q-Q (Quantile-Quantile Plot) พบว่าค่าความคลาดเคลื่อนเรียงตัวกันเป็นเส้นตรงใกล้เคียงเส้นทแยงมุม ดังรูปที่ 4.8 ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ (Normality)



รูปที่ 4.8 แผนภาพ Q-Q ความคลาดเคลื่อนของแบบจำลอง WLS

เมื่อสุ่มตัวอย่างมา 5,000 ตัวอย่าง และใช้เทคนิคทางสถิติ Shapiro-Wilk ในการทดสอบความเป็น Normality พบว่าค่า p-value เท่ากับ 0.1275 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแจกแจงปกติ (Normality)

3) ทดสอบความคลาดเคลื่อนมีความเป็นอิสระ (No autocorrelation)

จากตารางที่ 4.6 พบว่าค่าสถิติ Durbin-Watson เท่ากับ 2.001 หมายความว่า ค่าความคลาดเคลื่อนเป็นอิสระต่อกัน (No autocorrelation)

4) ทดสอบความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity)

เมื่อใช้เทคนิคทางสถิติ Breusch-Pagan พบว่าค่า p-value เท่ากับ < 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ $\alpha = 0.05$ หมายความว่า ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroskedasticity)

4.6 การอภิปรายผล

4.6.1 ผลของแบบจำลอง OLS, OLS-HAC, GLS, GLSAR, WLS

จากตารางที่ 4.7 ซึ่งแสดงผลเปรียบเทียบค่าสถิติของแบบจำลอง OLS, OLS-HAC, GLS, GLSAR, WLS พบว่า

1) ค่า R-squared (R^2) ของแบบจำลองทั้งหมด มีความสามารถอธิบายความแปรปรวนในตัวแปรตาม ด้วยตัวแปรอิสระได้สูงถึง 84.2% - 84.4% ซึ่งแบบจำลอง GLS มีค่าสูงที่สุดเพียงเล็กน้อย

2) ค่า F-statistic ของทุกแบบจำลองมีค่าที่สูง และ Prob(F-statistic) หรือ p-value น้อยกว่าระดับนัยสำคัญ $\alpha = 0.05$ บ่งชี้ว่าแบบจำลองสามารถอธิบายความแปรปรวนของตัวแปรตามจากตัวแปรอิสระได้ดี

3) แบบจำลองส่วนใหญ่มีค่าสัมประสิทธิ์ของค่าคงที่ คือ 519.9999 และค่าสัมประสิทธิ์ของตัวแปรอิสระ $u(t)$ คือ -0.5620 ยกเว้น แบบจำลอง GLS จะแตกต่างเพียงเล็กน้อย

4) ค่า p-value ของ t-statistic หรือ z-statistic ในแต่ละแบบจำลอง น้อยกว่า 0 ซึ่งบ่งชี้ว่าตัวแปรอิสระ u_t มีอิทธิพลอย่างมีนัยสำคัญต่อตัวแปรตาม PV

5) ค่า Durbin-Watson ประมาณ 2.000 ในทุกแบบจำลอง บ่งชี้ว่าค่าความคลาดเคลื่อนมีอิสระต่อกัน (No Autocorrelation)

6) การสร้างแบบจำลอง GLS จะใช้ข้อมูลแบบสุ่มจำนวน 30,000 ตัวอย่าง เนื่องจากคอมพิวเตอร์ของผู้ศึกษาไม่สามารถรองรับการประมวลผล GLS ด้วยข้อมูลที่มีขนาดใหญ่กว่านี้ได้ ส่วน GLSAR จะลดจำนวนตัวอย่าง 1 ตัวอย่างเท่านั้น เนื่องจากเป็นขั้นตอนการคำนวณใน Autoregressive (AR)

7) แบบจำลอง OLS-HAC จะไม่ทดสอบสมมติฐานต่างๆ เนื่องจากเป็นการปรับปรุง OLS ด้วย HAC ซึ่งใน HAC จะไม่มีการเปลี่ยนแปลงค่าใดๆ ใน OLS เพียงแต่คำนวณค่าความคลาดเคลื่อนมาตรฐานใหม่เท่านั้น เพื่อให้ทนต่อปัญหาค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroskedasticity) และทนต่อปัญหาค่าความคลาดเคลื่อนไม่เป็นอิสระ (Autocorrelation) ส่งผลให้ค่า F-statistic มีค่าสูงขึ้น

8) จากการทดสอบสมมติฐาน พบว่าแบบจำลองส่วนใหญ่เจอปัญหาค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroskedasticity) ยกเว้นแบบจำลอง GLS ซึ่งมีค่าความคลาดเคลื่อนแปรปรวนคงที่ (Homoscedasticity) อาจเกิดจากจำนวนตัวอย่างที่สร้างแบบจำลองน้อยเกินไป

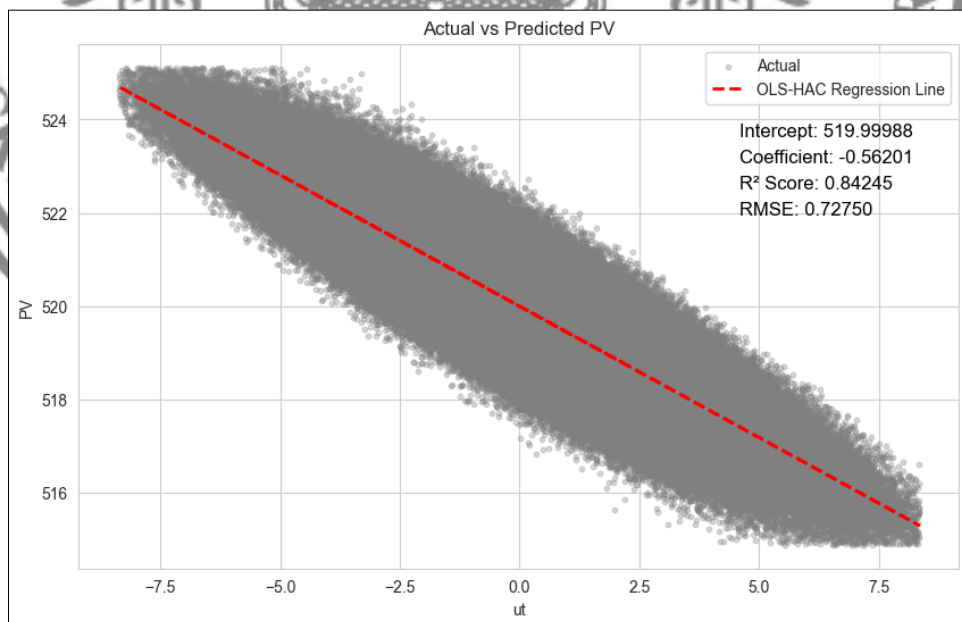
ตารางที่ 4.7 เปรียบเทียบค่าสถิติของแบบจำลอง OLS, OLS-HAC, GLS, GLSAR, WLS

รายการ	OLS	OLS-HAC	GLS	GLSAR	WLS
R-squared	0.842	0.842	0.844	0.842	0.842
F-statistic	2.186×10^6	2.856×10^6	1.627×10^5	2.186×10^6	2.187×10^6
Prob(F-statistic)	0.00	0.00	0.00	0.00	0.00
Coefficient (const)	519.9999	519.9999	520.0065	519.9999	519.9999
Coefficient (ut)	-0.5620	-0.5620	-0.5637	-0.5620	-0.5620
Std.Error(ut)	0.000	0.000	0.001	0.000	0.000
t/z-stat (ut)	-1478.457	-1690.020	-403.354	-1478.449	-1478.874
P> t / z (ut)	0.000	0.000	0.000	0.000	0.000
D.Watson	2.001	2.001	1.999	2.000	2.001
Observation	409,214	409,214	30,000	409,213	409,214
Linearity	ใช่	ไม่ทดสอบ	ใช่	ใช่	ใช่
Normality	ใช่	ไม่ทดสอบ	ใช่	ใช่	ใช่
No autocorrelation	ใช่	ไม่ทดสอบ	ใช่	ใช่	ใช่
Homoscedasticity	ไม่ใช่	ไม่ทดสอบ	ใช่	ไม่ใช่	ไม่ใช่

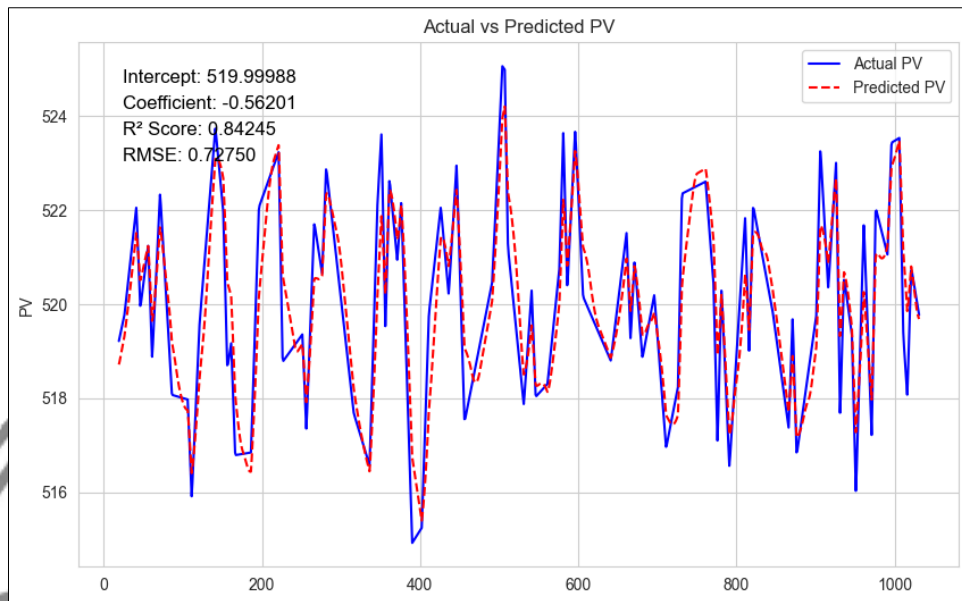
4.6.2 เลือกแบบจำลองและประเมินประสิทธิภาพ (Select and Evaluation)

ในการศึกษาครั้งนี้พิจารณาเลือกใช้แบบจำลอง OLS-HAC เป็นแบบจำลองหลักในการใช้งานกับระบบควบคุม PID เนื่องจากมีค่า F-statistic สูงที่สุด คือ 2.856×10^6 แสดงถึงความสามารถของแบบจำลองในการอธิบายความแปรปรวนของตัวแปรตามจากตัวแปรอิสระได้ดี นอกจากนี้ OLS-HAC ยังสามารถจัดการกับปัญหา Heteroscedasticity และ Autocorrelation ได้ ทำให้ค่าประมาณของพารามิเตอร์มีความน่าเชื่อถือมากขึ้น และค่า z-statistic ของตัวแปรอิสระ PID เอาต์พุต $u(t)$ ในแง่ของค่าสัมบูรณ์มีค่าสูงขึ้น คือ -1690.020 บ่งชี้ว่าตัวแปรอิสระนี้มีนัยสำคัญทางสถิติในแบบจำลอง

เมื่อประเมินประสิทธิภาพแบบจำลอง OLS-HAC พบว่ามีค่า RMSE เท่ากับ 0.727 และมีค่า R^2 เท่ากับ 0.842 หมายความว่าแบบจำลองสามารถอธิบายความแปรปรวนของตัวแปรตามได้ถึง 84.2% อย่างมีนัยสำคัญ ดังรูปที่ 4.9 สร้างสมการเส้นตรงได้คือ $Y = 519.9999 - 0.562X_1$ เมื่อนำแบบจำลอง OLS-HAC ไปทำนายค่าอัตราไหลของน้ำเปรียบเทียบกับค่าที่วัดได้จริง จะได้ดังรูปที่ 4.10



รูปที่ 4.9 ประเมินประสิทธิภาพแบบจำลอง OLS-HAC



รูปที่ 4.10 เปรียบเทียบค่าอัตราไหลของน้ำที่วัดได้จริงกับค่าทำนายด้วยแบบจำลอง OLS-HAC

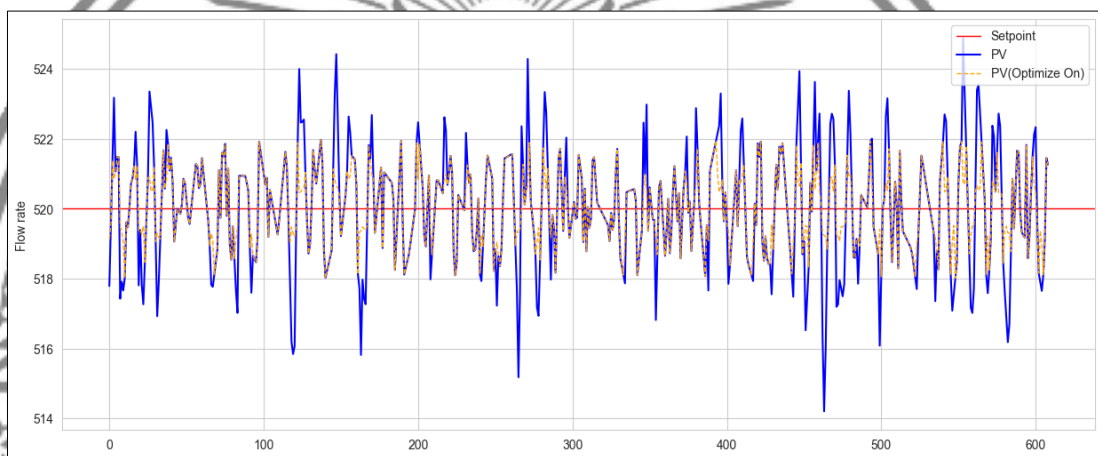
4.6.3 ปัญหาที่พบ

โดยส่วนใหญ่พบปัญหาค่าความคลาดเคลื่อนแปรปรวนไม่คงที่ (Heteroscedasticity) ในแต่ละตัวแบบจำลอง ส่งผลให้การทดสอบมีความน่าเชื่อถือลดลง จึงนำแบบจำลอง OLS มาปรับปรุงด้วยวิธี Heteroscedasticity and Autocorrelation Consistent (HAC) เพื่อให้แบบจำลองทนต่อปัญหา Heteroscedasticity และปัญหา Autocorrelation มากขึ้น โดยที่ค่า R-squared และค่าประมาณสัมประสิทธิ์การถดถอยยังคงเหมือนเดิม แต่วิธีการคำนวณค่าความคลาดเคลื่อนมาตรฐานเปลี่ยนแปลง และทำให้สถิติที่ใช้ในการทดสอบสมมติฐานมีการเปลี่ยนแปลงจาก t-statistic เป็น z-statistic เนื่องจาก ข้อมูลมีจำนวนมาก

ในการสร้างแบบจำลอง GLS พบข้อจำกัดสำคัญคือ สเปคคอมพิวเตอรืที่ใช้งานไม่เพียงพอต่อการประมวลผลข้อมูลขนาดใหญ่ ทำให้ไม่สามารถใช้ข้อมูลทั้งหมดที่มีมาสร้างแบบจำลองได้ จึงแก้ปัญหาด้วยการสุ่มตัวอย่างข้อมูลบางส่วนเพื่อนำมาสร้างแบบจำลอง GLS ซึ่งอาจทำให้เกิดความเอนเอียง (bias) และส่งผลกระทบต่อความน่าเชื่อถือของแบบจำลองได้

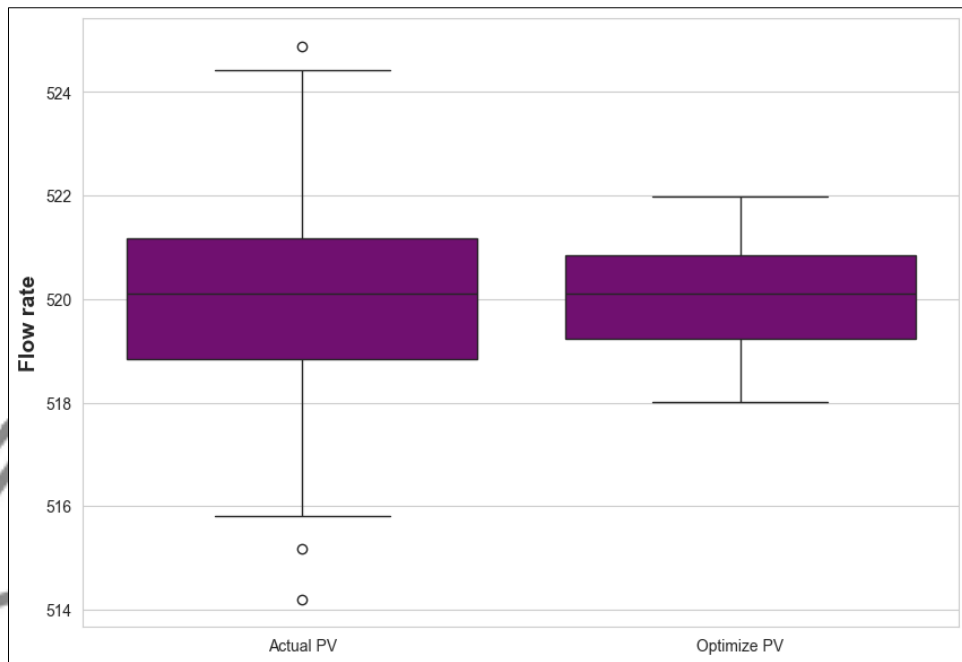
4.6.4 ผลการนำตัวแบบไปใช้งานจริง

เมื่อนำแบบจำลอง OLS-HAC ซึ่งมีค่า R^2 เท่ากับ 0.842 ไปจำลองการใช้งานกับระบบควบคุม PID แบบเรียลไทม์ ตามผังการทำงานของกระบวนการ Predictive-Optimize ดังรูปที่ 3.4 พบว่าค่าอัตราไหลที่ทำนายได้ (เส้นสีส้ม) มีค่าผันผวนลดลง เมื่อเทียบกับค่าอัตราไหลที่มาจากระบบควบคุม PID แบบดั้งเดิม ดังรูปที่ 4.11



รูปที่ 4.11 จำลองการใช้งาน Predictive-Optimize ในระบบควบคุม PID

ผลการควบคุมอัตราไหลของน้ำด้วยระบบควบคุม PID แบบดั้งเดิม จะมีค่าอยู่ในช่วงประมาณ 514 - 526 ลบ.ม./ชม. เมื่อเปรียบเทียบกับผลการจำลองการควบคุมด้วยระบบ PID ร่วมกับ Predictive-Optimize โดยใช้แบบจำลอง OLS-HAC พบว่าอัตราไหลของน้ำจะมีค่าอยู่ในช่วงประมาณ 518 - 522 ลบ.ม./ชม. เห็นได้ว่า ช่วงความผันผวนของการควบคุมอัตราไหลของน้ำมีค่าลดลง จากเดิมประมาณ 12 ลบ.ม./ชม. ลดลงเหลือประมาณ 4 ลบ.ม./ชม. ดังรูปที่ 4.12



รูปที่ 4.12 แผนภาพกล่องเปรียบเทียบก่อนและหลังใช้งานแบบจำลอง OLS-HAC

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การศึกษาค้นคว้าครั้งนี้มุ่งเน้นการปรับปรุงระบบควบคุม PID แบบวงจรรปิด (Close Loop) ด้วยการใช้ประยุกต์ใช้แบบจำลองพยากรณ์ โดยใช้ข้อมูลทุติยภูมิ (Secondary Data) จากฐานข้อมูล (Database) ทุกๆ 1 วินาที จากวันที่ 1 มีนาคม 2568 เวลา 00:00 ถึงวันที่ 7 มีนาคม 2568 เวลา 00:00 จะได้ข้อมูลดิบทั้งสิ้น 518,401 ตัวอย่าง และมี 6 ตัวแปร ได้แก่ 1.ค่าอัตราไหลของน้ำ (PV) 2.ค่าเป้าหมาย (SV) 3.ค่าควบคุม (MV) 4.ค่า Proportional (P) 5.ค่า Integral (I) 6.ค่า Derivative (D) จากนั้นนำข้อมูลดิบไปแปลงข้อมูล (Data Transformation) ทำให้ได้ตัวแปรใหม่ดังนี้ 7.ค่าผิดพลาด $e(t)$ และ 8.ค่า PID เอาต์พุต $u(t)$ หลังจากนั้นนำข้อมูลที่ผ่านการแปลงแล้วไปคลัสเตอร์ จะทำให้เหลือข้อมูลทั้งสิ้น 511,518 ตัวอย่าง หาค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) พบว่าค่า PID เอาต์พุต $u(t)$ และค่าอัตราไหลของน้ำ (PV) มีความสัมพันธ์กันสูงแบบแปรผกผันกันที่ -0.923 ผู้ศึกษาจึงพิจารณาสร้างแบบจำลองการถดถอยเชิงเส้น (Linear Regression) ทั้งสิ้น 5 วิธี คือ OLS, OLS-HAC, GLS, GLSAR, WLS โดยใช้ตัวแปรอิสระคือค่า PID เอาต์พุต $u(t)$ และตัวแปรตามคือค่าอัตราไหลของน้ำ (PV) และสุดท้ายได้พิจารณาเลือกแบบจำลอง OLS-HAC เป็นแบบจำลองหลักในการใช้งานกับระบบควบคุม PID เนื่องจากมีค่า F-statistic สูงที่สุด คือ 2.856×10^6 แสดงถึงความสามารถของแบบจำลองในการอธิบายความแปรปรวนของตัวแปรตามจากตัวแปรอิสระได้ดี นอกจากนี้ OLS-HAC ยังสามารถจัดการกับปัญหา Heteroscedasticity และ Autocorrelation ได้ดี ทำให้ค่าประมาณของพารามิเตอร์มีความแม่นยำมากขึ้น และค่า z-statistic ของตัวแปรอิสระ PID เอาต์พุต $u(t)$ ในแง่ของค่าสัมบูรณ์มีค่าสูงขึ้น คือ -1690.020 บ่งชี้ว่าตัวแปรอิสระนี้มีนัยสำคัญทางสถิติในแบบจำลอง

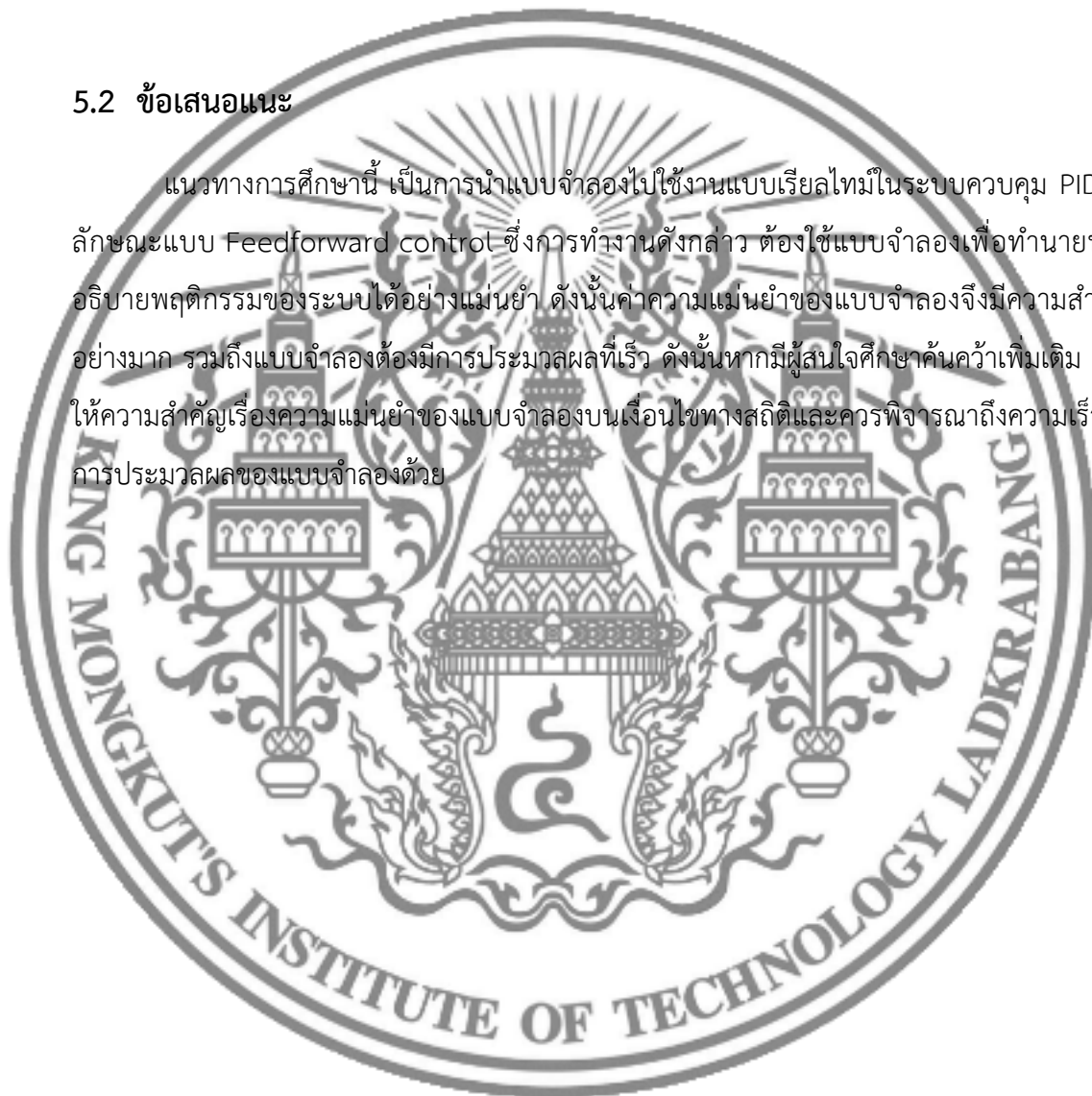
แบบจำลอง OLS-HAC มีค่า RMSE เท่ากับ 0.727 และค่า R^2 เท่ากับ 0.842 หมายความว่าแบบจำลองสามารถอธิบายความแปรปรวนของตัวแปรตามได้ถึง 84.2% อย่างมีนัยสำคัญ และสามารถสร้างสมการเส้นตรงได้คือ $Y = 519.9999 - 0.562X_1$

เมื่อจำลองการทำงานตามผังการทำงานของกระบวนการ Predictive-Optimize ดังรูปที่ 3.4 พบว่า อัตราไหลของน้ำจะมีค่าอยู่ในช่วงประมาณ 518 – 522 ลบ.ม./ชม. เมื่อเปรียบเทียบกับระบบ

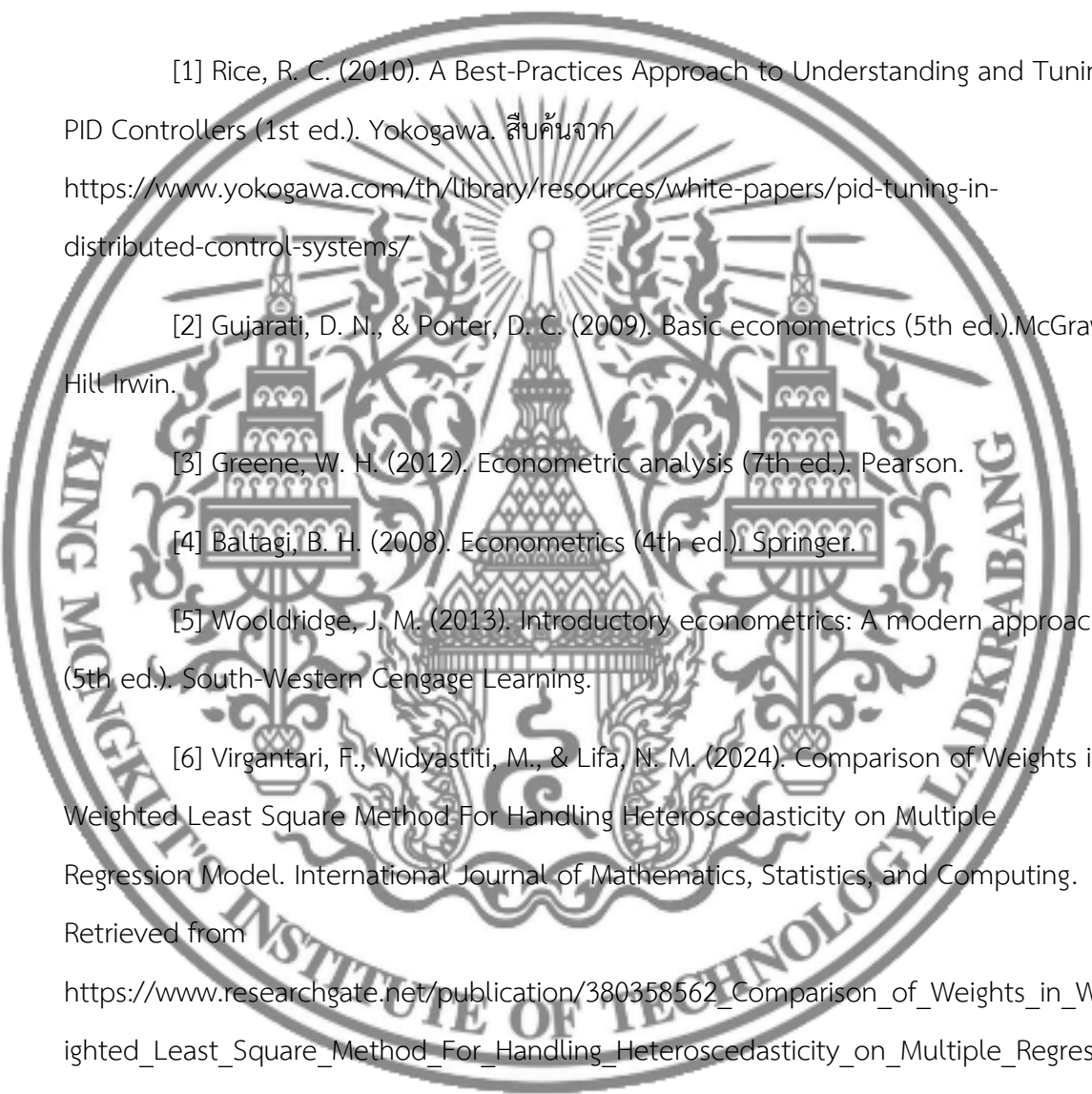
ควบคุม PID แบบดั้งเดิม ซึ่งมีค่าอยู่ในช่วงประมาณ 514 - 526 ลบ.ม./ชม. จะเห็นได้ว่า ช่วงความผันผวนอัตราไหลของน้ำจากการทำงานด้วย PID ร่วมกับ Predictive-Optimize มีค่าลดลงเหลือประมาณ 4 ลบ.ม./ชม. จากเดิมประมาณ 12 ลบ.ม./ชม. ซึ่งบ่งชี้ถึงประสิทธิภาพการควบคุมที่ดีขึ้น

5.2 ข้อเสนอแนะ

แนวทางการศึกษาี้ เป็นการนำแบบจำลองไปใช้งานแบบเรียลไทม์ในระบบควบคุม PID ในลักษณะแบบ Feedforward control ซึ่งการทำงานดังกล่าว ต้องใช้แบบจำลองเพื่อทำนายหรืออธิบายพฤติกรรมของระบบได้อย่างแม่นยำ ดังนั้นค่าความแม่นยำของแบบจำลองจึงมีความสำคัญอย่างมาก รวมถึงแบบจำลองต้องมีการประมวลผลที่เร็ว ดังนั้นหากมีผู้สนใจศึกษาค้นคว้าเพิ่มเติม ควรให้ความสำคัญเรื่องความแม่นยำของแบบจำลองบนเงื่อนไขทางสถิติและควรพิจารณาถึงความเร็วในการประมวลผลของแบบจำลองด้วย



บรรณานุกรม

- 
- [1] Rice, R. C. (2010). A Best-Practices Approach to Understanding and Tuning PID Controllers (1st ed.). Yokogawa. สืบค้นจาก <https://www.yokogawa.com/th/library/resources/white-papers/pid-tuning-in-distributed-control-systems/>
- [2] Gujarati, D. N., & Porter, D. C. (2009). Basic econometrics (5th ed.). McGraw-Hill Irwin.
- [3] Greene, W. H. (2012). Econometric analysis (7th ed.). Pearson.
- [4] Baltagi, B. H. (2008). Econometrics (4th ed.). Springer.
- [5] Wooldridge, J. M. (2013). Introductory econometrics: A modern approach (5th ed.). South-Western Cengage Learning.
- [6] Virgantari, F., Widyastiti, M., & Lifa, N. M. (2024). Comparison of Weights in Weighted Least Square Method For Handling Heteroscedasticity on Multiple Regression Model. International Journal of Mathematics, Statistics, and Computing. Retrieved from https://www.researchgate.net/publication/380358562_Comparison_of_Weights_in_Weighted_Least_Square_Method_For_Handling_Heteroscedasticity_on_Multiple_Regression_Model
- [7] Chaiyadecha, S. (2022). การทดสอบ Autocorrelation ใน Linear regression model ด้วย Python. สืบค้นจาก <https://lengyi.medium.com/autocorrelation-test-python-7dc8d8494598>

[8] Chaiyadecha, S. (2022). การทดสอบ Heteroscedasticity ใน Linear regression model ด้วย Python. สืบค้นจาก <https://lengyi.medium.com/heteroscedasticity-test-python-c9b25416d36>

[9] Yildiz, K. M. (2014). PID Parameters Prediction Using Neural Network for A Linear Quarter Car Suspension Control. Technical University.

[10] Grelewicz, P., Khuat, T. T., Czezot, J., Nowak, P., Klopot, T., & Gabrys, B. (2021). Application of Machine Learning to Performance Assessment for a class of PID-based Control Systems.

[11] Yang, Q., Chen, G., Guo, M., Chen, T., Luo, L., & Sun, L. (2022). Model Predictive Hybrid PID Control and Energy-Saving Performance Analysis of Supercritical unit. *Energies*, 17(24), 6356. Retrieved from <https://www.mdpi.com/1996-1073/17/24/6356>

[12] Agnihotri, S. P., & Waghmare, L. M. (2014). Regression Model for Tuning the PID Controller With Fractional Order Time Delay System. *Ain Shams Engineering Journal*, 5(4), 1223–1229.

[13] Thaker, M. K., & Patel, V. P. (2021). PID Tuning of FOPDT System Using Multivariate Linear Regression With Gradient Descent.

[14] Sampaio, V. (2023). Understanding Ordinary Least Squares (OLS): The Foundation of Linear Regression. Retrieved from <https://medium.com/@VitorCSampaio/understanding-ordinary-least-squares-ols-the-foundation-of-linear-regression-1d79bfc3ca35>

[15] ชิตพงษ์, กิตตินราดร. (2562). Linear Regression Algorithm. สืบค้นจาก <https://guopai.github.io/ml-blog02.html>

[16] Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59(3), 817–858. <https://doi.org/10.2307/2938229>

[17] Chaiyadecha, S. (2022). Newey–West Estimator: HAC Correction in Python. สืบค้นจาก <https://lengyi.medium.com/newey-west-estimator-hac-correction-inpython-4ad376d8fa83>

[18] Chaiyadecha, S. (2022). สอนทำ Heteroskedasticity and Autocorrelation Consistent. สืบค้นจาก <https://lengyi.medium.com/excel-tutorial-for-heteroskedasticity-and-autocorrelation-consistent-d145fccfd758>





ภาคผนวก

ภาคผนวก ก

1. ตัวอย่างข้อมูลดิบที่นำมาใช้ศึกษา มีทั้งสิ้น 518,401 แถวและมีจำนวนตัวแปร 6 ตัวแปร

DateTime	SV	PV	MV	P	I	D
01-Mar-25 00:00:01	520.00	517.86	52.85	100.00	20.00	0.00
01-Mar-25 00:00:02	520.00	517.87	52.97	100.00	20.00	0.00
01-Mar-25 00:00:03	520.00	517.95	53.12	100.00	20.00	0.00
01-Mar-25 00:00:04	520.00	518.02	53.26	100.00	20.00	0.00
01-Mar-25 00:00:05	520.00	518.10	53.40	100.00	20.00	0.00
01-Mar-25 00:00:06	520.00	518.18	53.54	100.00	20.00	0.00
01-Mar-25 00:00:07	520.00	518.26	53.55	100.00	20.00	0.00
01-Mar-25 00:00:08	520.00	518.34	53.53	100.00	20.00	0.00
01-Mar-25 00:00:09	520.00	518.42	53.51	100.00	20.00	0.00
01-Mar-25 00:00:10	520.00	518.50	53.49	100.00	20.00	0.00
01-Mar-25 00:00:11	520.00	518.58	53.48	100.00	20.00	0.00
01-Mar-25 00:00:12	520.00	518.66	53.46	100.00	20.00	0.00
01-Mar-25 00:00:13	520.00	518.74	53.44	100.00	20.00	0.00
01-Mar-25 00:00:14	520.00	518.82	53.42	100.00	20.00	0.00
01-Mar-25 00:00:15	520.00	518.90	53.40	100.00	20.00	0.00
01-Mar-25 00:00:16	520.00	518.98	53.39	100.00	20.00	0.00
01-Mar-25 00:00:17	520.00	519.06	53.37	100.00	20.00	0.00
01-Mar-25 00:00:18	520.00	519.14	53.35	100.00	20.00	0.00
01-Mar-25 00:00:19	520.00	519.22	53.33	100.00	20.00	0.00

2. ตัวอย่างข้อมูลที่ผ่านการทำความสะอาดข้อมูลแล้ว จะมีทั้งสิ้น 511,518 แถว และแปลงข้อมูลเพิ่ม

2.1) ค่าผิดพลาด $e(t)$ มาจาก SV - PV

2.2) ค่า PID เอาต์พุต $u(t)$ มาจากการคำนวณด้วยสมการ PID ของบริษัท YOKOGAWA

DateTime	SV	PV	MV	P	I	D	Error	Integral	Derivative	SumError
2025-03-01 00:00:01	520	517.86475	52.85088	100	20	0	2.1352539	3.7523193	0.5181885	2.3228699
2025-03-01 00:00:02	520	517.86591	52.973011	100	20	0	2.1340942	5.8864136	-0.00116	2.4284149
2025-03-01 00:00:03	520	517.94537	53.11544	100	20	0	2.0546265	7.94104	-0.079468	2.4516785
2025-03-01 00:00:04	520	518.02484	53.25787	100	20	0	1.9751587	9.9161987	-0.079468	2.4709686
2025-03-01 00:00:05	520	518.10431	53.400299	100	20	0	1.8956909	11.81189	-0.079468	2.4862854
2025-03-01 00:00:06	520	518.18372	53.542728	100	20	0	1.8162842	13.628174	-0.079407	2.4976929
2025-03-01 00:00:07	520	518.26318	53.545902	100	20	0	1.7368164	15.36499	-0.079468	2.5050659
2025-03-01 00:00:08	520	518.34265	53.528263	100	20	0	1.6573486	17.022339	-0.079468	2.5084656
2025-03-01 00:00:09	520	518.42212	53.510628	100	20	0	1.5778809	18.60022	-0.079468	2.5078918
2025-03-01 00:00:10	520	518.50153	53.492989	100	20	0	1.4984741	20.098694	-0.079407	2.5034088
2025-03-01 00:00:11	520	518.58099	53.475349	100	20	0	1.4190063	21.5177	-0.079468	2.4948914
2025-03-01 00:00:12	520	518.66046	53.457714	100	20	0	1.3395386	22.857239	-0.079468	2.4824005
2025-03-01 00:00:13	520	518.73993	53.440075	100	20	0	1.2600708	24.11731	-0.079468	2.4659363
2025-03-01 00:00:14	520	518.8194	53.42244	100	20	0	1.180603	25.297913	-0.079468	2.4454987
2025-03-01 00:00:15	520	518.8988	53.4048	100	20	0	1.1011963	26.399109	-0.079407	2.4211517
2025-03-01 00:00:16	520	518.97827	53.387165	100	20	0	1.0217285	27.420837	-0.079468	2.3927704
2025-03-01 00:00:17	520	519.05774	53.369526	100	20	0	0.9422607	28.363098	-0.079468	2.3604156
2025-03-01 00:00:18	520	519.13721	53.351891	100	20	0	0.862793	29.225891	-0.079468	2.3240875
2025-03-01 00:00:19	520	519.21661	53.334251	100	20	0	0.7833862	30.009277	-0.079407	2.2838501
2025-03-01 00:00:20	520	519.29608	53.316616	100	20	0	0.7039185	29.09613	-0.079468	2.158725
2025-03-01 00:00:21	520	519.37555	53.298977	100	20	0	0.6244507	27.585327	-0.079468	2.003717
2025-03-01 00:00:22	520	519.45502	53.281342	100	20	0	0.5449829	25.996216	-0.079468	1.8447937
2025-03-01 00:00:23	520	519.53442	53.263702	100	20	0	0.4655762	24.407166	-0.079407	1.6859344
2025-03-01 00:00:24	520	519.61389	53.246067	100	20	0	0.3861084	22.818115	-0.079468	1.5270142
2025-03-01 00:00:25	520	519.69336	53.228428	100	20	0	0.3066406	21.229065	-0.079468	1.3680939
2025-03-01 00:00:26	520	519.77283	53.210793	100	20	0	0.2271729	19.639954	-0.079468	1.2091705

ประวัติผู้เขียน

ชื่อ-สกุล นายสุทธิพงศ์ สังข์รัมย์

วัน เดือน ปี เกิด 23 พฤษภาคม 2533

สถานที่เกิด จังหวัดนครราชสีมา

ประวัติการศึกษา (2555) วิศวกรรมศาสตร์ สาขาไฟฟ้ากำลัง เกรดเฉลี่ย 2.24
(มหาวิทยาลัยธรรมศาสตร์)

ที่อยู่ปัจจุบัน 38/409 ถ.พหลโยธิน ต.คลองหนึ่ง อ.คลองหลวง จ.ปทุมธานี 12120

ผลงานตีพิมพ์ ไม่มี

รางวัลที่ได้รับ ไม่มี

