



รายงานการวิจัย

แบบจำลองอนุกรมเวลาการถดถอยเชิงพหุของข้อมูลน้ำท่ารายเดือน
ของกลุ่มน้ำปัตตานี

**Multiple Regression Time Series Model of Monthly Runoff
for Pattani Basin**

จุฑามาศ ลักษณะกิจ

Chuthamat Laksanakit

060964

๑๑๕.๑๓

จ ๖๒๘

๒๕๕๓

คณะวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย

หน้า ๑

ได้รับการสนับสนุนทุนวิจัยจากมหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย

งบประมาณผลประโยชน์ประจำปี พ.ศ.๒๕๕๑

กิตติกรรมประกาศ
(Acknowledgement)

งานวิจัยนี้ได้รับทุนอุดหนุนโครงการวิจัยประเภททุนอุดหนุนทั่วไปประจำปี
งบประมาณ พ.ศ.2550 จากมหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย ทำให้งานสำเร็จลุล่วงได้
ด้วยดี ขอขอบพระคุณคุณอนุรักษ์ ชีระโชติ ผู้อำนวยการ โครงการส่งน้ำและบำรุงรักษาปัตตานี
และเจ้าหน้าที่โครงการฯ ทุกท่าน ที่ได้ให้ความอนุเคราะห์ข้อมูลและอำนวยความสะดวกในการ
ตรวจสอบ รวบรวมข้อมูลภาคสนามอย่างดียิ่ง

ขอบคุณเพื่อนอาจารย์และบุคลากรทุกท่านที่มีส่วนช่วยเหลือ อำนวยความสะดวก
และให้กำลังใจจนกระทั่งรูปเล่มงานวิจัยฉบับสมบูรณ์นี้แล้วเสร็จ

คณะผู้วิจัย

กันยายน 2552



แบบจำลองอนุกรมเวลาการถดถอยเชิงพหุของข้อมูลน้ำท่ารายเดือนของกลุ่มน้ำปัตตานี

จุฑามาศ ลักษณกิจ¹

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อจำลองปริมาณน้ำท่ารายเดือนของกลุ่มน้ำปัตตานีโดยใช้แบบจำลองอนุกรมเวลาการถดถอยเชิงพหุ ซึ่งอาศัยความสัมพันธ์ของคาบย่อนเวลาของข้อมูลน้ำท่าและข้อมูลน้ำฝน ณ เวลาต่างๆ เพื่อนำแบบจำลองที่เหมาะสมไปใช้ในการพยากรณ์ปริมาณน้ำท่ารายเดือนเพื่อประโยชน์ด้านการบริหารจัดการน้ำหรือการวิเคราะห์ความเสี่ยงของระบบลุ่มน้ำ

ผลการวิจัยพบว่ารูปแบบของ MLR Model ที่มีความแม่นยำในการพยากรณ์ข้อมูลที่สุดคือ $Y_t = 1.618 + 0.325Y_{t-1} + 0.145Y_{t-2} + 0.002449Rb_t + 0.001719Rp_t$ มีค่า R^2 เท่ากับ 0.472 ค่า RMSE เท่ากับ 0.765 ค่า MAE เท่ากับ 0.524 ค่าเฉลี่ยของ Relative Error เท่ากับ 0.06 และค่าเบี่ยงเบนมาตรฐานของ Relative Error เท่ากับ 1.04 ข้อมูลปริมาณน้ำท่ารายเดือนที่ได้จากแบบจำลองดังกล่าว ส่วนใหญ่มีค่า Average Monthly Mean ต่ำกว่าค่าของข้อมูลจริงโดยเฉลี่ยตลอดปีประมาณ 15.78 เปอร์เซ็นต์ ส่วนค่า Average Monthly Standard Deviation โดยรวมมีแนวโน้มเช่นเดียวกับค่า Average Monthly Mean แต่เปอร์เซ็นต์ความแตกต่างสูงกว่าโดยเฉลี่ยตลอดปีประมาณ 42.64 เปอร์เซ็นต์

คำสำคัญ: น้ำท่ารายเดือน กลุ่มน้ำปัตตานี อนุกรมเวลา แบบจำลองการถดถอย

¹ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย อ.เมือง จ.สงขลา

Multiple Regression Time Series Model of Monthly Runoff for Pattani Basin

Chuthamat Laksanakit¹

Abstract

The monthly runoff for Pattani basin was modeled by using multiple linear regression (MLR model) analyses. The lag time of monthly runoff and rainfall were considered as the independent variables. For developing the efficient MLR model to forecast the monthly runoff that had more benefit for water resources management or risk analysis.

The results show that the slightly high accuracy MLR model was $Y_t = 1.618 + 0.325Y_{t-1} + 0.145Y_{t-2} + 0.002449Rb_t + 0.001719Rp_t$ with 0.472 of R^2 , 0.765 of RMSE, 0.524 of MAE, 0.06 and 1.04 of the average and the standard deviation of relative error, respectively. The forecasted monthly runoff was almost less than the real data. The average monthly mean was less than the actual data about 15.78 per cent. Although, the average monthly standard deviation had the same trend but different of the percentage was higher around 42.64 percent with a year consider.

Keywords: Monthly Runoff Pattani Basin Time Series Regression Model

¹ Faculty of Engineering, Rajamangala University of Technology Srivijaya, Muang, Songkhla.

สารบัญเรื่อง

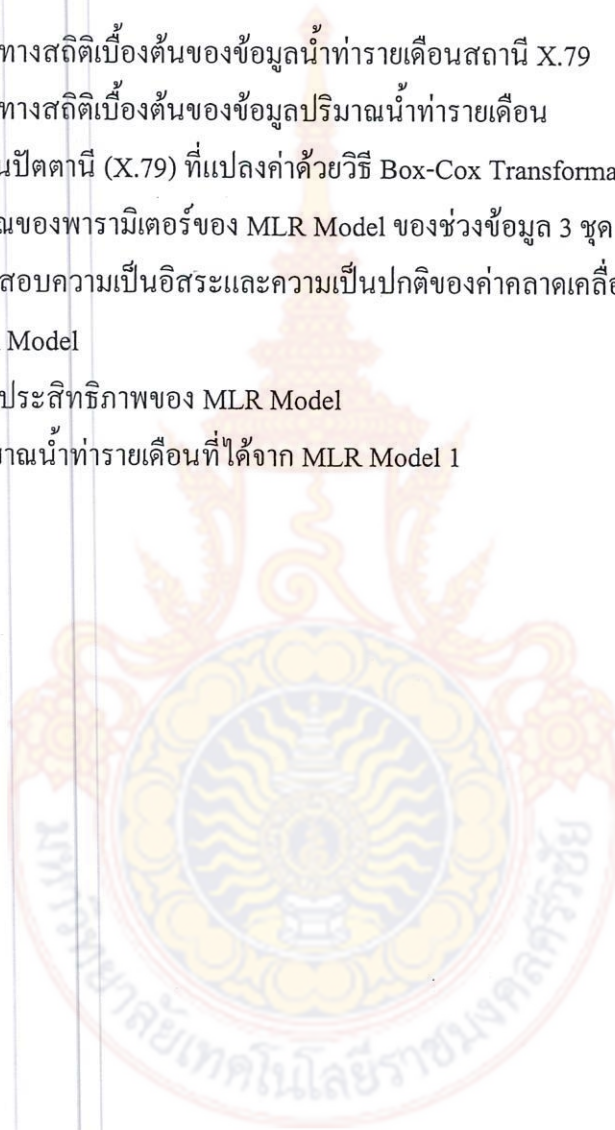
เรื่อง	หน้า
กิตติกรรมประกาศ	ข
บทคัดย่อภาษาไทย	ค
บทคัดย่อภาษาอังกฤษ	ง
สารบัญเรื่อง	จ
สารบัญตาราง	ฉ
สารบัญรูป	ช
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของโครงการวิจัย	3
1.3 ขอบเขตของการวิจัย	3
1.4 วิธีการดำเนินการวิจัยและสถานที่ทำการทดลอง/เก็บข้อมูล	4
1.5 ผลประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 ทฤษฎีและหลักการที่ใช้ในการวิจัย	5
2.1 ทฤษฎีของ Multiple Linear Regression Model	5
2.1.1 นิยามของ Multiple Linear Regression Model	5
2.1.2 การประมาณค่าพารามิเตอร์	9
2.1.3 การทดสอบสมมุติฐานและรูปแบบการถดถอยที่ดีที่สุด	10
2.2 การแปลงค่าอนุกรมเวลา	12
2.2.1 Standardization	13
2.2.2 Box-Cox Transformation	14
2.3 การตรวจสอบความเหมาะสมของรูปแบบ	15
2.3.1 การทดสอบความเป็นอิสระ	15
2.3.2 การทดสอบความเป็นปกติ	16
2.3.3 Akaike Information Criterion	17

สารบัญเรื่อง (ต่อ)

	หน้า
2.4 ค่าวัดประสิทธิภาพของ MLR Model	17
2.5 เอกสารที่เกี่ยวข้อง	18
2.6 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง	22
บทที่ 3 วิธีดำเนินการวิจัย	23
3.1 วิธีดำเนินการวิจัยของ Multiple Linear Regression Model	23
บทที่ 4 ผลการวิจัยและวิจารณ์	27
4.1 การวิเคราะห์คุณสมบัติเบื้องต้นและการเตรียมข้อมูล	27
4.1.1 การวิเคราะห์คุณสมบัติเบื้องต้นของข้อมูลปริมาณน้ำท่ารายเดือน	27
4.1.2 การวิเคราะห์คุณสมบัติเบื้องต้นของข้อมูลปริมาณน้ำฝนรายเดือน	30
4.1.3 การเตรียมข้อมูล	34
4.2 การประมาณค่าพารามิเตอร์	36
4.3 การตรวจสอบความเหมาะสมของแบบจำลอง	38
บทที่ 5 สรุปผลและเสนอแนะ	42
5.1 สรุปผลการวิจัย	42
5.2 ข้อเสนอแนะ	43
เอกสารอ้างอิง	44
ภาคผนวก	47

สารบัญตาราง

	หน้า	
ตาราง 4.1	คุณสมบัติทางสถิติเบื้องต้นของข้อมูลน้ำท่ารายเดือนสถานี X.79	28
ตาราง 4.2	คุณสมบัติทางสถิติเบื้องต้นของข้อมูลปริมาณน้ำท่ารายเดือน สถานีเขื่อนป่าสักชลสิทธิ์ (X.79) ที่แปลงค่าด้วยวิธี Box-Cox Transformation	29
ตาราง 4.3	ค่าประมาณของพารามิเตอร์ของ MLR Model ของช่วงข้อมูล 3 ชุด	38
ตาราง 4.4	ผลการทดสอบความเป็นอิสระและความเป็นปกติของค่าคลาดเคลื่อน ของ MLR Model	39
ตาราง 4.5	ค่าสถิติวัดประสิทธิภาพของ MLR Model	39
ตาราง 4.6	ข้อมูลปริมาณน้ำท่ารายเดือนที่ได้จาก MLR Model 1	41



สารบัญรูป

		หน้า
รูป 2.1	Scatter Plots Diagram	8
รูป 2.2	Correlation Matrix	8
รูป 2.3	Sample Correlograms of $\epsilon_{p,\tau}$ -Series of Daily Flows of the Oconto River	9
รูป 2.4	กราฟ Time Series Plots of Residual	11
รูป 2.5	กราฟการแจกแจงของข้อมูลอุตุ-อุทกวิทยา	13
รูป 3.1	แผนผังแสดงขั้นตอนการสร้าง Multiple Linear Regression Model	25
รูป 3.2	แผนที่ตำแหน่งสถานีวัดน้ำท่าและสถานีวัดน้ำฝนของกลุ่มน้ำป่าตानी	26
รูป 4.1	กราฟอนุกรมเวลาปริมาณน้ำท่ารายเดือนของสถานีเขื่อนป่าตानी	28
รูป 4.2	กราฟอนุกรมเวลาปริมาณน้ำท่ารายเดือนของสถานีเขื่อนป่าตानी (X.79) ที่แปลงค่าด้วยวิธี Box-Cox Transformation (ปี ค.ศ.1997-2007)	29
รูป 4.3	กราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนของสถานีเขื่อนบางลาง	30
รูป 4.4	กราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนของสถานีเมืองยะลา	31
รูป 4.5	กราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนของสถานีเขื่อนป่าตानी	31
รูป 4.6	Double Mass Curve ของสถานีวัดปริมาณน้ำฝนเขื่อนบางลาง	32
รูป 4.7	Double Mass Curve ของสถานีวัดปริมาณน้ำฝนเมืองยะลา	33
รูป 4.8	Double Mass Curve ของสถานีวัดปริมาณน้ำฝนเขื่อนป่าตानी	33
รูป 4.9	กราฟ Scatter Plots Matrix ระหว่างข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่าของสถานี X.79 และข้อมูลปริมาณน้ำฝนรายเดือนของ 3 สถานี	34
รูป 4.10	Correlation Matrix ระหว่างข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่าของสถานี X.79 และข้อมูลปริมาณน้ำฝนรายเดือนของ 3 สถานี	35
รูป 4.11	กราฟ Average Monthly Mean ของข้อมูลปริมาณน้ำท่าจริง และจาก MLR Model 1	40
รูป 4.12	กราฟ Average Monthly Standard Deviation ของข้อมูลปริมาณน้ำท่าจริง และจาก MLR Model 1	40

บทที่ 1

บทนำ

เมื่อความเจริญและอาณาเขตของสังคมเมืองได้ขยายตัวเข้าไปยังพื้นที่ในภูมิภาคต่างๆ ของประเทศ ทำให้ความต้องการปัจจัยสี่เพิ่มขึ้นตามไปด้วย โดยเฉพาะอย่างยิ่งความต้องการใช้น้ำทั้งเพื่อการอุปโภคและบริโภค ด้วยเหตุนี้ทำให้หน่วยงานที่รับผิดชอบต้องจัดเตรียมน้ำเพื่อให้เพียงพอต่อความต้องการของผู้บริโภค ดังนั้นจึงเกิดการพัฒนาโครงการทางด้านแหล่งน้ำตามมา แต่ปัญหาที่เกิดขึ้นกับการวางโครงการงานแหล่งน้ำทั้งงานขนาดเล็กและงานขนาดใหญ่ในหลายพื้นที่ที่พบบ่อยก็คือ พื้นที่นั้นไม่มีการจัดเก็บข้อมูลน้ำ หรือข้อมูลที่มีอยู่มีจำนวนน้อยไม่เพียงพอที่จะนำมาวิเคราะห์ หรือข้อมูลที่มีอยู่ขาดความน่าเชื่อถือ ปัญหาเหล่านี้มักจะเกิดกับพื้นที่ที่การตรวจวัดข้อมูลยังล่าสมัยหรือยังไม่มีการตรวจวัดข้อมูล จากสาเหตุดังกล่าวเหล่านี้ ผลักดันให้เกิดการพัฒนาการจำลองกระบวนการที่เกิดขึ้นในธรรมชาติ นั่นคือกระบวนการอุทกวิทยาด้วยแบบจำลองคณิตศาสตร์ โดยได้ผสานแนวคิดทางอุทกวิทยาเข้ากับหลักทางคณิตศาสตร์และหลักสถิติ ทำให้เกิดแบบจำลองหลากหลายรูปแบบที่สามารถนำมาใช้กับข้อมูลทางอุตุ-อุทกวิทยาได้เป็นอย่างดี ช่วยลดปัญหาในเรื่องของการจัดหาข้อมูลและด้วยประสิทธิภาพของแบบจำลองสามารถตอบสนองความต้องการในศึกษาเพื่อการออกแบบหรือวางแผนโครงการแหล่งน้ำได้ และยังช่วยให้ข้อจำกัดในด้านอื่นๆ ลดลงด้วย

1.1 ความเป็นมาและความสำคัญของปัญหา

การเปลี่ยนแปลงสภาพภูมิอากาศ (Climate Change) หรือการที่โลกร้อนขึ้น (Global Warming) ส่งผลร้ายแรงอย่างต่อเนื่องต่อระบบนิเวศน์วิทยาและสิ่งที่มีชีวิตในโลก ตัวอย่างที่เห็นได้ชัดได้แก่ ความแห้งแล้งอย่างรุนแรง ภัยพิบัติ อุทกภัย พายุฝนฟ้าคะนอง พายุทอร์นาโด แผ่นดินถล่ม และการเกิดพายุรุนแรงฉับพลัน จากภาวะอันตรายเหล่านี้พบว่า ผู้ที่อาศัยอยู่ในเขตพื้นที่ที่เสี่ยงกับการเกิดเหตุการณ์ดังกล่าวได้รับผลกระทบมากกว่าพื้นที่ส่วนอื่นๆ นอกจากนี้ ยังมีการคาดการณ์ว่า การที่อุณหภูมิของโลกสูงขึ้น เป็นเหตุให้มีการเปลี่ยนแปลงด้านการทำประมง การเพาะเลี้ยงสัตว์น้ำ และความไม่แน่นอนของฤดูกาลที่ส่งผลกระทบต่อการทำเกษตรกรรม เมื่อปริมาณผลผลิตเพื่อการบริโภคโดยรวมลดลง ย่อมส่งผลให้ภาคอุตสาหกรรมต้องอยู่ภายใต้ความผันผวนไปด้วย ในที่สุดการลงทุนหรือระบบเศรษฐกิจก็จะขาดเสถียรภาพ

การพัฒนาข้อมูลน้ำเพื่อการอุปโภค-บริโภค การเกษตร การประปาและอุตสาหกรรม การจัดการน้ำเพื่อเพิ่มประสิทธิภาพการชลประทานให้สูงขึ้น จำเป็นต้องอาศัยการวิเคราะห์ข้อมูล อุทกวิทยาและข้อมูลอุทกนิยมนิยามวิทยาประกอบการตัดสินใจ โครงการที่เกี่ยวกับการพัฒนาแหล่งน้ำในด้านต่างๆ เช่น การศึกษาปริมาตรหรือขนาดของอ่างเก็บน้ำ การจัดระเบียบการใช้ น้ำ ชีตความสามารถของลำน้ำ และการป้องกันหรือเตือนภัยจากการเกิดน้ำท่วม ที่ผู้ออกแบบจะต้องทราบคือ ปริมาณน้ำในลำน้ำ ปริมาณน้ำรายเดือนหรือรายปี อัตราการไหลต่ำสุดของลำน้ำ พายุกรรมหรือความแปรผันของปริมาณน้ำในลำน้ำ และน้ำท่วม ข้อมูลดังกล่าวเหล่านี้ได้จากการวิเคราะห์ข้อมูลน้ำท่า (Runoff) จากสถานีวัดน้ำท่าของกรมชลประทาน

เนื่องจากข้อมูลน้ำท่าเป็นข้อมูลอนุกรมเวลาที่มีความไม่แน่นอน ข้อมูลที่จะนำมาใช้ในการวิเคราะห์ต้องมีความยาวของช่วงข้อมูลหลายสิบปี ตัวอย่างข้อมูลที่บันทึกไว้เป็นระยะเวลา มากกว่า 10 ปีขึ้นไปยังนับว่าเป็นตัวอย่างข้อมูลจำนวนน้อยไม่เพียงพอที่จะใช้เป็นตัวแทนประชากรของข้อมูลน้ำท่าได้ และในทางปฏิบัติเราไม่สามารถรอเวลาเพื่อที่จะเก็บตัวอย่างข้อมูลให้มีจำนวน มากพอกับการศึกษาวิเคราะห์ได้ ด้วยข้อจำกัดทางด้านเวลานี้เอง จึงได้มีการประยุกต์วิธีการทางสถิติมาใช้ในการสร้างข้อมูลหรือพยากรณ์ปริมาณน้ำท่าที่มีคุณสมบัติทางสถิติใกล้เคียงกับข้อมูลจริง เพื่อเป็นการเพิ่มจำนวนตัวอย่างข้อมูล และเพื่อช่วยให้การศึกษาพายุกรรมหรือความแปรผันของลำน้ำ น่าเชื่อถือยิ่งขึ้น

การวิจัยครั้งนี้เป็นการศึกษาความสัมพันธ์ของข้อมูลอนุกรมเวลาน้ำท่าและน้ำฝนรายเดือน โดยใช้แบบจำลองอนุกรมเวลาการถดถอยเชิงพหุ (Multiple Regression Time Series Model) เพื่อศึกษาว่าแบบจำลองรูปแบบใดจะมีประสิทธิภาพดีกว่ากัน ในเชิงสถิติ เพื่อนำแบบจำลองที่ได้ไปใช้ประโยชน์ในการสร้างข้อมูลหรือพยากรณ์ปริมาณน้ำท่าที่จะเกิดขึ้นหากปีจ้จยหรือตัวแปรที่สำคัญคือข้อมูลฝนเปลี่ยนแปลงไป ซึ่งจะช่วยให้ผู้เกี่ยวข้องกับการบริหารจัดการแหล่งน้ำสามารถทำเลียนแบบสถานการณ์ (Simulation) ในรูปแบบต่างๆ ได้ดียิ่งขึ้น

1.2 วัตถุประสงค์ของโครงการวิจัย

1.2.1 เพื่อประยุกต์ใช้วิธีการทางสถิติในการสร้างแบบจำลอง Multiple Regression Time Series Model โดยอาศัยความสัมพันธ์ของข้อมูลน้ำฝนกับข้อมูลน้ำท่าในการหาค่าพารามิเตอร์ และหาตัวแปรอิสระของ Multiple Regression Time Series Model ที่เหมาะสมกับลักษณะข้อมูลน้ำท่ารายเดือนของกลุ่มน้ำปีตตานี

1.2.2 เปรียบเทียบว่าข้อมูลน้ำท่า (Runoff Data) และ ลอการิทึมของน้ำท่า (Log of Runoff Data) ข้อมูลรูปแบบใดจะทำให้แบบจำลองมีประสิทธิภาพมากกว่า

1.3 ขอบเขตของการวิจัย

พื้นที่ศึกษาคือกลุ่มน้ำปีตตานี ครอบคลุมพื้นที่จังหวัดยะลาและจังหวัดปีตตานี โดยมีแม่น้ำปีตตานีเป็นลำน้ำหลัก และมีแม่น้ำยะหาเป็นลำน้ำสาขา ในช่วงตอนปลายมีคลองหนองจิก แยกออกจากแม่น้ำปีตตานี ข้อมูลน้ำท่าที่นำมาใช้อยู่ในรูปของข้อมูลรายเดือนเป็นระยะเวลานานกว่า 10 ปี ส่วนข้อมูลฝนรายเดือนได้จากสถานีวัดน้ำฝนที่อยู่ในบริเวณรัศมีเดียวกันกับสถานีวัดน้ำท่า

แบบจำลองทางสถิติที่ศึกษาประกอบด้วย Multiple Regression Time Series Model โดยแบ่งออกเป็น 2 รูปแบบย่อย คือ Seasonal Model และ Nonseasonal Model

การประมาณค่าพารามิเตอร์ของ Multiple Regression Model เลือกใช้วิธี กำลังสองน้อยที่สุด (Least Square Method) และ ใช้วิธีการเลือกตัวแปรอิสระแบบขั้นตอน (Stepwise Regression) ในการเลือกสมการถดถอยที่ดีที่สุด

การทดสอบประสิทธิภาพของแบบจำลอง สำหรับ Multiple Regression Model พิจารณาจากค่า ผลรวมกำลังสองของความคลาดเคลื่อน (SSE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าสัมประสิทธิ์การตัดสินใจ (R^2)

1.4 วิธีการดำเนินการวิจัยและสถานที่ทำการทดลอง/เก็บข้อมูล

1.4.1 เลือกลำน้ำที่สนใจจะศึกษา รวบรวมข้อมูลน้ำท่ารายเดือน ข้อมูลน้ำฝนรายเดือน และข้อมูลอื่นที่เกี่ยวข้อง

1.4.2 ทำการประมาณค่าพารามิเตอร์และกำหนดตัวแปรอิสระหรือหาสมการที่ดีที่สุดของ Multiple Regression Time Series Model

1.4.3 ประยุกต์ใช้แบบจำลองในข้อ 1.5.2 ในการสร้างข้อมูลน้ำท่ารายเดือน

1.4.4 เปรียบเทียบผลการสร้างข้อมูลน้ำท่ารายเดือนของแบบจำลอง โดยใช้ค่าสถิติวัดประสิทธิภาพของการพยากรณ์เป็นเกณฑ์ในการตัดสินใจ

1.4.5 สรุปผล และข้อเสนอแนะ

1.5 ผลประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ได้แบบจำลองอนุกรมเวลาการถดถอยเชิงพหุที่สามารถจำลองความสัมพันธ์ของข้อมูลน้ำท่ารายเดือนได้ใกล้เคียงกับข้อมูลจริงมากที่สุด

1.5.2 ข้อมูลน้ำท่ารายเดือนที่ได้จากแบบจำลอง สามารถนำมาเพิ่มหรือต่อช่วงสถิติของข้อมูลน้ำท่าให้มีช่วงของข้อมูลน้ำท่าที่ยาวขึ้น ช่วยให้เห็นความเปลี่ยนแปลงหรือพฤติกรรมในด้านต่างๆ ได้ จำนวนข้อมูลที่มีความยาวเพียงพอและเหมาะสมจะช่วยให้การวิเคราะห์โดยวิธีทางสถิติซึ่งใช้ประกอบกับการออกแบบเชิงอุทกวิทยา (Hydrologic Design) หรือการวิเคราะห์ความเสี่ยงของโครงการพัฒนาแหล่งน้ำ (Risk Analysis) มีความเชื่อมั่นยิ่งขึ้น

1.5.3 ใช้ประโยชน์จากแบบจำลองอนุกรมเวลาทางสถิติในข้อ 1.5.1 ต่อขยายช่วงของข้อมูลน้ำท่ารายเดือน สำหรับลำน้ำที่มีสถานีวัดน้ำแต่เพิ่งทำการเก็บบันทึกข้อมูล หรือใช้สำหรับการประเมินเพิ่มเติมค่าข้อมูลในช่วงที่มีข้อมูลขาดหายไป

บทที่ 2

ทฤษฎีและหลักการที่ใช้ในการวิจัย

วัตถุประสงค์หลักในการสร้าง Multiple Linear Regression Model นั้น ก็เพื่อนำหลักการทางสถิติมาประยุกต์ใช้ในการจำลองกระบวนการอุทกวิทยาซึ่งเป็นกระบวนการที่เกิดขึ้นตามธรรมชาติ เป็นวิธีการหนึ่งที่สามารถช่วยให้ได้มาซึ่งข้อมูลซึ่งเกิดขึ้นจากหลักการที่เชื่อถือได้ และข้อมูลดังกล่าวจะใช้เป็นข้อมูลพื้นฐานในการดำเนินงานในส่วนของการวิเคราะห์เพื่อการวางแผนงานหรือเพื่อการออกแบบ และยังมีส่วนช่วยให้งานบรรลุวัตถุประสงค์ได้ด้วยดี ทฤษฎีและหลักการที่นำมาใช้ในการวิจัยของ Multiple Linear Regression Model ที่จะกล่าวในบทนี้มีรายละเอียดดังนี้

2.1 ทฤษฎีของ Multiple Linear Regression Model

การวิเคราะห์การถดถอยเป็นวิธีการที่ใช้อธิบายความสัมพันธ์ระหว่างตัวแปรต่างๆ ตั้งแต่ 2 ตัวขึ้นไปว่ามีความสัมพันธ์กันหรือไม่ และถ้ามีความสัมพันธ์กันจะมีลักษณะความสัมพันธ์เป็นแบบใด โดยเรียกตัวแปรที่สนใจจะวิเคราะห์ว่าตัวแปรตาม (Dependent Variable) ส่วนตัวแปรที่แทนปัจจัยต่างๆ ที่ทำให้ค่าของตัวแปรตามมีการเปลี่ยนแปลงจะเรียกว่าตัวแปรอิสระ (Independent Variables)

การประยุกต์ใช้การวิเคราะห์การถดถอยกับข้อมูลทางอุทกวิทยาโดยทั่วไปนิยมใช้ลักษณะของการถดถอยเชิงเส้น และส่วนมากจะเป็นการศึกษาความสัมพันธ์ระหว่างน้ำฝนกับน้ำท่า เนื่องจากตัวแปรทั้ง 2 ตัวนี้มีความสัมพันธ์กันอย่างเห็นได้ชัดที่สุด ทฤษฎีที่จะใช้กับการวิเคราะห์การถดถอยแบ่งออกเป็น 4 หัวข้อ โดยมีรายละเอียดของแต่ละหัวข้อดังต่อไปนี้

2.1.1 นิยามของ Multiple Linear Regression Model

ลักษณะความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระของข้อมูลเชิงปริมาณสามารถจัดได้เป็น 2 กลุ่มใหญ่ๆ คือ แบบเส้นตรง (Linear) และแบบไม่เป็นเส้นตรง (Nonlinear) เช่น แบบพหุนาม (Polynomial) แบบเอ็กซ์โปเนนเชียล (Exponential) แต่หากข้อมูลเป็นอนุกรมเวลา สามารถใช้ตัวแปรดัมมี่ (Dummy Variables) หรือตัวแปรย้อนเวลา (Lag Variables) วิเคราะห์ร่วมลักษณะความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระที่กล่าวมาแล้วได้

และควรมีจำนวนตัวอย่างอนุกรมเวลาที่มีความต่อเนื่องอย่างน้อย 10 ปี เพื่อให้การประมาณค่าพารามิเตอร์ของการถดถอยเป็นไปอย่างมีประสิทธิภาพ

ก. รูปแบบของ Multiple Linear Regression Model

เมื่อตัวแปรอิสระที่ต้องการศึกษามีมากกว่า 2 ตัวขึ้นไป จะต้องใช้การวิเคราะห์การถดถอยเชิงเส้นแบบพหุหรือ MLR Model ซึ่งรูปทั่วไปของ MLR Model ได้แสดงในสมการที่ 2.1

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_{m-1} X_{m-1,i} + \varepsilon_i \quad (2.1)$$

เมื่อนำการวิเคราะห์การถดถอยมาประยุกต์ใช้กับอนุกรมเวลา สมการของการถดถอยก็สามารถเขียนให้อยู่ในรูปแบบที่มีตัวแปรเวลาเข้ามาเกี่ยวข้อง นั่นคือสมการที่ 2.2

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_{m-1} X_{m-1,t} + \varepsilon_t \quad (2.2)$$

เมื่อ	Y_t	คือ	ตัวแปรตามที่เป็นอนุกรมเวลา
	$X_{1t}, X_{2t}, \dots, X_{m-1,t}$	คือ	ตัวแปรอิสระที่เป็นอนุกรมเวลา และมีอิทธิพลต่อตัวแปรตาม
	$\beta_0, \beta_1, \beta_2, \dots, \beta_{m-1}$	คือ	ค่าสัมประสิทธิ์การถดถอย
	ε_t	คือ	ค่าความคลาดเคลื่อนของรูปแบบการถดถอย (Residual) ที่มีการแจกแจงปกติ $N(0, \sigma_\varepsilon^2)$

ข. สมมุติฐานของการวิเคราะห์การถดถอย

- 1) ค่าความคลาดเคลื่อน ε_t เป็นตัวแปรที่มีการแจกแจงปกติ
- 2) ค่าเฉลี่ยของค่าความคลาดเคลื่อน ε_t เท่ากับศูนย์ นั่นคือ $E(\varepsilon_t) = 0$
- 3) ค่าความแปรปรวนของค่าความคลาดเคลื่อน ε_t เป็นค่าคงที่ใดๆ ที่ไม่ทราบค่า นั่นคือ $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$
- 4) ค่า ε_t และค่า ε_{t-1} เป็นอิสระต่อกัน กล่าวคือค่าความคลาดเคลื่อนจะต้องไม่มีสหสัมพันธ์ระหว่างกัน นั่นคือ $\text{COV}(\varepsilon_t, \varepsilon_{t-1}) = 0$

ก. การวิเคราะห์สหสัมพันธ์ (Correlation)

สหสัมพันธ์เป็นตัวบ่งชี้ว่าตัวแปรตามกับตัวแปรอิสระมีความสัมพันธ์กันหรือไม่ และถ้ามีความสัมพันธ์กันลักษณะของความสัมพันธ์เป็นแบบใด การวิเคราะห์สหสัมพันธ์จะเริ่มจากการทำ Scatter Plots Diagram ซึ่งเป็นการพล็อตกราฟระหว่างตัวแปรตาม Y กับกลุ่มของตัวแปรอิสระ X_1, \dots, X_{m-1} ดังตัวอย่างในรูป 2.1 เพื่อที่จะทราบได้คร่าวๆ ว่าตัวแปรตามและตัวแปรอิสระแต่ละตัวมีระดับและทิศทางของความสัมพันธ์เป็นอย่างไร

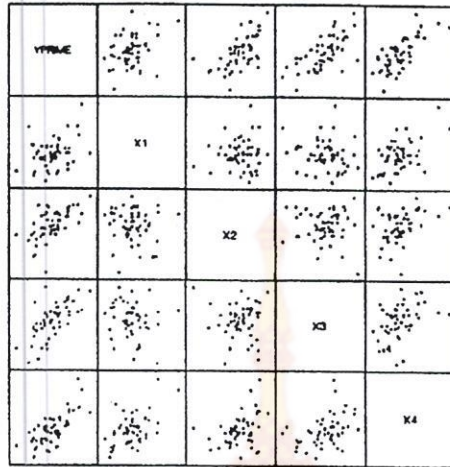
นอกจาก Scatter Plots Diagram แล้ว สามารถทราบระดับสหสัมพันธ์ได้จาก Correlation Matrix ซึ่งเป็นเมทริกซ์ที่แสดงค่าที่ใช้วัดสหสัมพันธ์ระหว่างตัวแปรตาม Y กับตัวแปรอิสระ X_1, \dots, X_{m-1} ดังตัวอย่างในรูป 2.2 ค่าที่ใช้วัดสหสัมพันธ์ระหว่างตัวแปรที่นิยมใช้มี 2 ค่า คือ ค่าสัมประสิทธิ์สหสัมพันธ์ของ Spearman r_S ซึ่งเป็นค่าวัดสหสัมพันธ์ระหว่างลำดับที่ของตัวแปร และค่าสัมประสิทธิ์สหสัมพันธ์ของ Pearson r_{YX} ที่จะใช้ในการศึกษานี้เป็นค่าวัดสหสัมพันธ์ระหว่างค่าของตัวแปร ค่า r_{YX} คำนวณได้สมการที่ 2.3

$$r_{YX_i} = \frac{\sum_{t=1}^N (Y_t - \bar{Y})(X_{i,t} - \bar{X})}{\left(\sum_{t=1}^N (Y_t - \bar{Y})^2 \sum_{t=1}^N (X_{i,t} - \bar{X})^2 \right)^{1/2}} \quad (2.3)$$

เมื่อ Y_t คือ ตัวแปรตาม

$X_{i,t}$ คือ ตัวแปรอิสระตัวที่ $i=1, \dots, m-1$

ค่าสัมประสิทธิ์สหสัมพันธ์ r_{YX_i} มีค่าอยู่ระหว่าง -1 ถึง 1 ถ้าตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กันมาก ค่า $|r_{YX_i}|$ จะมีค่าเข้าใกล้ค่า 1 ส่วนเครื่องหมายบวกหรือลบจะบอกว่าทิศทางของสหสัมพันธ์ระหว่างตัวแปรเป็นแบบใด ถ้าเป็นบวกหมายถึงมีความสัมพันธ์ไปในทางเดียวกัน แต่ถ้าเป็นลบหมายถึงมีความสัมพันธ์ในลักษณะตรงข้ามกัน



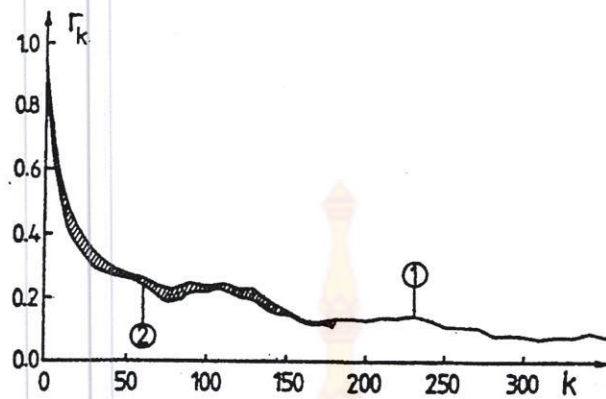
รูป 2.1 Scatter Plots Diagram

	Y'	X_1	X_2	X_3	X_4
Y'	1.000	.346	.593	.665	.726
X_1		1.000	.090	-.150	.502
X_2			1.000	-.024	.369
X_3				1.000	.416
X_4					1.000

รูป 2.2 Correlation Matrix

ง. Stationary

Stationary คือลักษณะที่อนุกรมเวลาอยู่ในภาวะสมดุลเชิงสถิติ (Statistical Equilibrium) กล่าวคือ คุณสมบัติทางสถิติของอนุกรมเวลานั้นไม่เปลี่ยนแปลงเมื่อเวลาเปลี่ยนแปลงไป ในทางปฏิบัติ การตรวจสอบว่าอนุกรมเวลามีลักษณะ Stationary หรือไม่นั้นมีวิธีการที่ยุ่งยากจึงต้องพิจารณาจากลักษณะอื่นแทน โดยทั่วไป นิยมพิจารณาจากกราฟ Correlogram (r_k) และกราฟ Partial Correlogram ($\hat{\phi}_k(k)$) ของข้อมูลอนุกรมเวลา หากกราฟคู่เข้าสู่ค่าศูนย์และไม่มีลักษณะของฤดูกาล แสดงว่าอนุกรมเวลาชุดนั้น Stationary และสามารถดูตัวอย่างกราฟ Correlogram (r_k) ของอนุกรมเวลาที่ Stationary ได้จากรูป 2.3



รูป 2.3 Sample Correlograms of $\varepsilon_{p,\tau}$ -Series of Daily Flows of the Oconto River :

- (1) Use of 10 Harmonic for the Periodicity in m_τ and S_τ ; and (2) Use of 6 Harmonic for the Periodicity in m_τ and S_τ

ที่มา : Yevjevich (1984)

2.1.2 การประมาณค่าพารามิเตอร์ (Parameter Estimation)

เพื่อให้ความคลาดเคลื่อนในการประมาณค่าของ MLR Model มีค่าต่ำสุด จะใช้วิธีที่เรียกว่า “วิธีกำลังสองน้อยที่สุด (Least Square Method)” ประมาณค่าพารามิเตอร์จาก สมการที่ 2.2 จำนวน m ตัว นั่นคือ $\beta_0, \beta_1, \dots, \beta_{m-1}$ โดยจะแทนค่าประมาณของพารามิเตอร์ ด้วยตัวแปร b_0, b_1, \dots, b_{m-1} ตามลำดับ เมื่อตัวแปรตามและตัวแปรอิสระเป็นอนุกรมเวลา ในแต่ละชุดของข้อมูลจะประกอบด้วยข้อมูลจำนวนมาก ดังนั้นจึงนำวิธีเมตริกซ์ มาช่วยเพื่อให้การคำนวณง่ายขึ้น โดยสามารถหาค่าประมาณของพารามิเตอร์ b_0, b_1, \dots, b_{m-1} ได้จากสมการที่ 2.4

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y}) \quad (2.4)$$

$m \times 1 \quad m \times m \quad m \times 1$

เมื่อ \mathbf{b} คือ เมตริกซ์ของค่าประมาณของพารามิเตอร์ที่มีขนาด $m \times 1$

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{m-1} \end{bmatrix}$$

$m \times 1$

\mathbf{X} คือ เมตริกซ์ของค่าของตัวแปรตามที่มีขนาด $N \times m$

เขตปฏิเสธสมมุติฐาน H_0

$$t > t_{1-\alpha/2; N-k-1}$$

ถ้าปฏิเสธสมมุติฐาน H_0 หมายความว่า X_j ไม่ควรอยู่ในสมการถดถอย

การตรวจสอบโดยนัยทางสถิติ โดยการประเมินคุณภาพของค่าประมาณของพารามิเตอร์ดังที่ได้กล่าวมาแล้ว เป็นขั้นตอนหนึ่งของการเลือกตัวแปรอิสระให้กับสมการซึ่งมีหลายวิธีได้แก่ วิธีเพิ่มตัวแปรอิสระ (Forward Selection) วิธีลดตัวแปรอิสระ (Backward Selection) และวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise Regression)

การตรวจสอบความสอดคล้องกับสมมุติฐานของการวิเคราะห์การถดถอย จะทำการตรวจสอบจากค่าคลาดเคลื่อน ซึ่งสมมุติฐานในแต่ละข้อก็จะมีวิธีการตรวจสอบแตกต่างกัน แต่หากข้อมูลที่น่ามาวิเคราะห์เป็นอนุกรมเวลาจะต้องทำการตรวจสอบสมมุติฐานที่ว่าค่าคลาดเคลื่อนจะต้องไม่มีสหสัมพันธ์ระหว่างกัน $COV(\varepsilon_t, \varepsilon_{t-1}) = 0$ โดยเลือกใช้วิธีการตรวจสอบหลายวิธีประกอบกัน ดังนี้

ก. Time Series Plots

ทริงศิริ (2542) กล่าวว่าไว้ว่าหากข้อมูลเป็นอนุกรมเวลา จะพิจารณาความเป็นอิสระจากกราฟอนุกรมเวลาของค่าความคลาดเคลื่อน ถ้าจุดของกราฟของค่าความคลาดเคลื่อนมีการเคลื่อนไหวอยู่ในแนวระนาบขนานกับแกนนอน กล่าวได้ว่าอนุกรมเวลามีความเป็นอิสระดังตัวอย่างในรูป 2.4



รูป 2.4 กราฟ Time Series Plots of Residual

ข. Durbin & Watson Test

การทดสอบของ Durbin และ Watson เป็นการทดสอบหาความสัมพันธ์แบบ Autocorrelation ของค่าคลาดเคลื่อน ซึ่งเป็นความสัมพันธ์ที่เกี่ยวข้องกับเวลา นิยมใช้ทดสอบกับค่าคลาดเคลื่อนของแบบจำลองการถดถอย สถิติทดสอบของ Durbin และ Watson ดังสมการที่ 2.5

$$D = \frac{\sum_{t=1}^N (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^N \varepsilon_t^2} \quad (2.5)$$

เมื่อ D คือ ค่าทดสอบของ Durbin และ Watson

ε_t คือ ค่าคลาดเคลื่อน

โดยมีสมมติฐานดังนี้

H_0 : ค่าความคลาดเคลื่อน ε_t และ ε_{t-1} เป็นอิสระต่อกัน

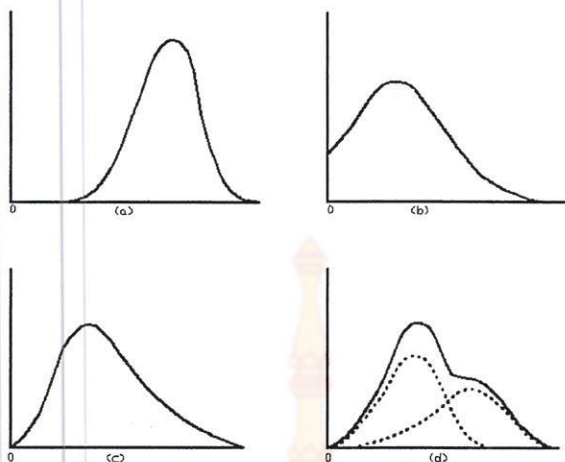
H_1 : ค่าความคลาดเคลื่อน ε_t และ ε_{t-1} ไม่เป็นอิสระต่อกัน

ค่า D ที่ได้จะมีค่าอยู่ระหว่าง 0 ถึง 4 ถ้าค่า D เข้าใกล้ 2 จะถือว่าไม่นัยสำคัญ นั่นคือ ε_t และ ε_{t-1} เป็นอิสระต่อกัน แต่ถ้าค่า D ไม่เข้าใกล้ 2 แสดงว่า ε_t และ ε_{t-1} มีความสัมพันธ์กัน

2.2 การแปลงค่าอนุกรมเวลา (Transformation of Time Series)

ทฤษฎีการแจกแจงและเทคนิคทางสถิติที่นำมาประยุกต์ใช้กับข้อมูลอุตุ-อุทกวิทยาบ่อยที่สุดคือ การสมมุติให้ตัวแปรมีการแจกแจงแบบปกติ แต่เนื่องจากกราฟการแจกแจงของข้อมูลอุตุ-อุทกวิทยาส่วนมากจะไม่สมมาตร เพราะข้อมูลที่เป็นอนุกรมเวลามีความไม่แน่นอนและมีเฉพาะค่าบวก ดังตัวอย่างกราฟการแจกแจงในรูป 2.5 ให้สังเกตว่าทุกรูปขอบเขตของกราฟจะเริ่มจากศูนย์และเป็นกราฟที่ไม่สมมาตร

โดยทั่วไป อนุกรมคาบเวลาจะมีอิทธิพลของแนวโน้ม ฤดูกาล วัฏจักร หรือเหตุการณ์ที่ผิดปกติรวมอยู่ด้วย การที่อนุกรมเวลามีอิทธิพลเหล่านี้จะทำให้ไม่สามารถสมมุติให้อนุกรมเวลาที่มีการแจกแจงแบบปกติได้ จะต้องทำการแปลงค่าอนุกรมเวลาเสียก่อนเพื่อที่จะปรับอนุกรมเวลาให้มีการแจกแจงเป็นปกติ แล้วจึงจะนำอนุกรมเวลาที่แปลงค่าแล้วไปวิเคราะห์ในขั้นตอนต่อไป วิธีการแปลงค่าที่นิยมใช้กับอนุกรมเวลาทางอุตุ-อุทกวิทยามีดังนี้



รูป 2.5 กราฟการแจกแจงของข้อมูลอุตุ-อุทกวิทยา
ที่มา : Holder (1985)

2.2.1 Standardization

อนุกรมเวลาที่ Standardized แล้วจะต้องมีค่าเฉลี่ยเท่ากับศูนย์ และค่าเบี่ยงเบนมาตรฐานเท่ากับหนึ่งเสมอ วิธีนี้จะช่วยเปลี่ยนการแจกแจงของข้อมูลจากการแจกแจงปกติ $N(\mu, \sigma^2)$ ให้เป็นการแจกแจงปกติมาตรฐาน $N(0,1)$ โดยใช้การปรับความแปรปรวนของอนุกรมเวลาหรือขจัดอิทธิพลที่เกิดจากฤดูกาลออกจากอนุกรมเวลา ค่า Standardized สำหรับอนุกรมเวลาหาได้จากสมการ 2.6

$$Z_{v,\tau} = \frac{X_{v,\tau} - \hat{\mu}_\tau}{\hat{\sigma}_\tau} \quad (2.6)$$

- เมื่อ $X_{v,\tau}$ คือ อนุกรมเวลาลำดับปีที่ v ลำดับคาบเวลาที่ τ
เมื่อ $v=1, \dots, n$ และ $\tau=1, \dots, \omega$
(อนุกรมเวลารายวัน $\omega=365$)
- $Z_{v,\tau}$ คือ ค่า Standardized ของอนุกรมเวลา
- $\hat{\mu}_\tau$ คือ ค่าประมาณของค่าเฉลี่ยของอนุกรมเวลาคาบเวลาที่ τ
- $\hat{\sigma}_\tau$ คือ ค่าประมาณของค่าเบี่ยงเบนมาตรฐานของอนุกรมเวลาคาบเวลาที่ τ

2.2.2 Box-Cox Transformation

เมื่อได้ทำการตรวจสอบค่าสัมประสิทธิ์ความเบ้ C_s ของอนุกรมเวลาแล้วพบว่าค่า C_s มีค่าไม่เข้าใกล้ค่าศูนย์ ทำให้ไม่สามารถสมมุติให้อนุกรมเวลามีการแจกแจงแบบปกติได้ เพื่อแก้ไขปัญหาที่เกิดขึ้นกับอนุกรมเวลาข้อมูลดิบ Box และ Cox (1964) ได้เสนอวิธีการแปลงค่าเพื่อช่วยปรับให้อนุกรมเวลาที่มีการแจกแจงแบบปกติได้ โดยใช้สมการที่ 2.7 และ 2.8 สำหรับงานที่ต้องการความละเอียดสูงจะใช้เกณฑ์ของค่า C_s ที่ $-0.05 < C_s < 0.05$ แต่ถ้าเป็นงานที่ไม่ต้องการความละเอียดมากนักก็จะใช้เกณฑ์ของค่า C_s ที่ $-0.1 < C_s < 0.1$

$$Y_t = (X_t + \lambda_2)^{\lambda_1} \quad \lambda_1 \neq 0 \quad (2.7)$$

$$Y_t = \ln(X_t + \lambda_2) \quad \lambda_1 = 0 \quad (2.8)$$

สำหรับทุกช่วงเวลา t ค่า $(X_t + \lambda_2) > 0$ เสมอ

เมื่อ	Y_t	คือ	อนุกรมเวลาแปลงค่า
	λ_1	คือ	เลขชี้กำลังของสมการ Box-Cox Transformation
	λ_2	คือ	ค่าคงที่ใดๆ

สำหรับสมการที่ 2.7 กรณีที่ค่า $\lambda_1 \neq 0$ การหาค่า λ_1 ที่เหมาะสมจะใช้วิธี Interaction โดยจะทำการกำหนดค่าเริ่มต้นให้ $\lambda_1 = 0$ และเพิ่มค่าขึ้นทีละน้อยไปเรื่อยๆ คือ $\lambda_1 = 0, \pm 0.5, \pm 1.0, \dots$ ในแต่ละครั้งที่เปลี่ยนค่า λ_1 ก็จะหาค่า C_s ของค่า Standardized ของอนุกรมเวลาแปลงค่า Z_t โดยจะเลือกใช้ค่า λ_1 ที่ทำให้ค่า $C_s(Z_t)$ มีค่าต่ำที่สุด

วิธีแปลงค่าในสมการที่ 2.8 มีชื่อเรียกอีกชื่อหนึ่งว่า “Logarithmic Transformation” เป็นวิธีที่นักอุทกวิทยานิยมใช้ เพราะเป็นวิธีหนึ่งที่ช่วยป้องกันไม่ให้ค่าที่ได้จากแบบจำลองติดลบ

การแปลงค่า Y_t ย้อนกลับให้เป็นค่า X_t สามารถคำนวณได้จากสมการที่ 2.9 และ 2.10 ตามลำดับ

$$X_t = Y_t^{1/\lambda_1} - \lambda_2 \quad \lambda_1 \neq 0 \quad (2.9)$$

$$X_t = \exp(Y_t) - \lambda_2 \quad \lambda_1 = 0 \quad (2.10)$$

2.3 การตรวจสอบความเหมาะสมของรูปแบบ (Diagnostic Checking)

เมื่อได้กำหนดรูปแบบให้กับข้อมูลอนุกรมเวลาที่นำมาศึกษา จะต้องทำการตรวจสอบความเหมาะสมของรูปแบบทุกครั้งก่อนที่นำรูปแบบนั้นไปใช้ วิธีการตรวจสอบความเหมาะสมของรูปแบบแบ่งออกเป็น 2 ส่วน เป็นการตรวจสอบความเหมาะสมของรูปแบบจากค่าคลาดเคลื่อน (Residual) โดยส่วนแรกเป็นการทดสอบความเป็นอิสระของค่าคลาดเคลื่อน (Test of Independence) ประกอบด้วยวิธี Autocorrelation Check วิธี Portmanteau Lack of Fit Test หรือเรียกอีกอย่างหนึ่งว่า Box-Pierce Test ส่วนที่สองเป็นการทดสอบความเป็นปกติของค่าคลาดเคลื่อน (Test of Normality) ประกอบด้วยวิธี Normal Probability Plots และวิธี Kolmogorov-Siminov Test หากอนุกรมเวลาที่นำมาวิเคราะห์มีรูปแบบความเหมาะสมมากกว่าหนึ่งรูปแบบ จะใช้ค่า Akaike Information Criterion (AIC) และการพิจารณา Model Correlogram เป็นเกณฑ์ในการตัดสินใจเลือกรูปแบบที่เหมาะสมที่สุด ดังรายละเอียดต่อไปนี้

2.3.1 การทดสอบความเป็นอิสระ (Test of Independence)

วิธีการทดสอบความเป็นอิสระที่นิยมใช้ทดสอบค่าคลาดเคลื่อนจากแบบจำลองอนุกรมเวลามีหลายวิธี แต่ละวิธีก็จะมีเงื่อนไขและความเหมาะสมกับข้อมูลแต่ละแบบ สำหรับการศึกษานี้ขอเสนอวิธี Autocorrelation Check และวิธี Portmanteau Lack of Fit Test มีรายละเอียดดังนี้

ก. Autocorrelation Check

การทดสอบความเป็นอิสระของค่าคลาดเคลื่อนจากแบบจำลอง โดยการพิจารณาจากค่าอัตโนมัติสหสัมพันธ์ของค่าคลาดเคลื่อน $r_k(\epsilon_t)$ กรณีที่ตัวอย่างเป็นอนุกรมเวลาที่มีขนาดใหญ่ ใช้การพิจารณาจากกราฟ Correlogram ของค่า $r_k(\epsilon_t)$ ของค่าคลาดเคลื่อนว่ามีค่าใกล้เคียงศูนย์หรือไม่ หรืออีกนัยหนึ่งคือ $|r_k(\epsilon_t)| \leq \frac{1.96}{\sqrt{N}}$ หรือไม่ ถ้าค่า $r_k(\epsilon_t)$ ที่ $k=1$ อยู่ภายในขอบเขต

$\pm \frac{1.96}{\sqrt{N}}$ แสดงว่าค่าคลาดเคลื่อนเป็นอิสระต่อกัน

060964

๑๑๕.๑๓

๖๖๒๘

๒๕๕๓

ข. Portmanteau Lack of Fit Test

การทดสอบความเป็นอิสระของค่าคลาดเคลื่อนจากแบบจำลอง โดยการพิจารณาว่าค่าคลาดเคลื่อนไม่มีสหสัมพันธ์ในตัวเอง เป็นหนึ่งในหลายวิธีที่นำมาใช้ตรวจสอบความเหมาะสมของรูปแบบของ AR Model ซึ่ง Box-Pierce (1970) ได้เสนอสถิติทดสอบไว้ดังสมการที่ 2.11

$$Q = (N-d) \sum_{k=1}^L r_k^2(\varepsilon_t) \quad (2.11)$$

เมื่อ $r_k(\varepsilon_t)$ คือ ค่าอัตโนมัติสหสัมพันธ์ของค่าคลาดเคลื่อน ε_t

L คือ จำนวนมากที่สุดของการย่นเวลาที่พิจารณา

2.3.2 การทดสอบความเป็นปกติ (Test of Normality)

มีวิธีการทดสอบหลายวิธีเป็นที่ยอมรับสำหรับทดสอบความเป็นปกติของค่าคลาดเคลื่อนจากแบบจำลองอนุกรมเวลา ในการศึกษาเลือกใช้วิธีการทดสอบความเป็นปกติ 2 วิธี ประกอบด้วย วิธีการกราฟฟิก คือ Normal Probability Plots และวิธีการใช้สถิติทดสอบของ Kolmogorov-Smirnov Test ตามรายละเอียดต่อไปนี้

ก. Normal Probability Plots

วิธีที่นิยมใช้ในการทดสอบความเป็นปกติของค่าคลาดเคลื่อนจากแบบจำลอง คือ การพล็อตกราฟการแจกแจงของค่าคลาดเคลื่อนลงบนกระดาษ Normal Probability แล้วพิจารณาว่าจุดที่พล็อตนั้นประมาณได้ว่าเป็นเส้นตรงหรือไม่ หรือจุดที่พล็อตนั้นอยู่ในช่วงความเชื่อมั่นหรือไม่ ถ้าจุดของกราฟอยู่ในแนวเส้นตรงหรืออยู่ในช่วงความเชื่อมั่น แสดงว่าค่าคลาดเคลื่อนจากแบบจำลองชุดนั้นมีการแจกแจงแบบปกติ

ข. Kolmogorov-Smirnov Test

Kolmogorov-Smirnov เป็นสถิติที่ใช้ทดสอบการแจกแจงของข้อมูลตัวอย่างว่าเป็นแบบปกติหรือไม่ หลักการของการทดสอบนี้คือ การเปรียบเทียบค่าฟังก์ชันการแจกแจงสะสมของข้อมูลตัวอย่างกับของประชากร ถ้าค่าความแตกต่างต่ำ แสดงว่าข้อมูลตัวอย่างนั้นมีการแจกแจงแบบปกติ สถิติทดสอบของ Kolmogorov-Smirnov คำนวณได้จากสมการที่ 2.12

$$D_N = \max |F_N(x) - F(x)| \quad (2.12)$$

เมื่อ $F_N(x)$ คือ ความน่าจะเป็นสะสมของค่าสังเกต

$F(x)$ คือ ความน่าจะเป็นสะสมภายใต้สมมุติฐาน

x_i คือ ค่าสังเกตที่ $i=1, 2, \dots, N$

ที่ระดับนัยสำคัญ α ค่าวิกฤตคือ $D_{\alpha, N}$ ในกรณีที่จำนวนข้อมูล $N > 30$

$$\text{ค่าวิกฤตจะมีค่าประมาณ } D_{\alpha, N} \approx \sqrt{\frac{-\ln\left(\frac{\alpha}{2}\right)}{2N}}$$

2.3.3 Akaike Information Criterion (AIC)

โดยทั่วไปอนุกรมเวลาหนึ่งชุดมักจะมีรูปแบบของ MLR Model ที่สามารถใช้อ้างอิงข้อมูลได้เป็นอย่างดีมากกว่า 1 รูปแบบขึ้นไป ดังนั้นในการตัดสินใจเลือกว่าจะใช้รูปแบบใดรูปแบบหนึ่งนั้นจะพิจารณาจากค่า AIC ซึ่งรูปแบบที่มีความเหมาะสมที่สุดจะต้องมีค่า AIC ต่ำที่สุด โดยสามารถหาค่า AIC ได้จากสมการที่ 2.13

$$AIC(p, q) = N \ln(\hat{\sigma}_e^2) + 2(p + q) \quad (2.13)$$

เมื่อ N คือ จำนวนข้อมูลอนุกรมเวลา

$\hat{\sigma}_e^2$ คือ ค่าความแปรปรวนของค่าความคลาดเคลื่อนที่ประมาณค่าด้วยวิธีภาวะน่าจะเป็นสูงสุด

p, q คือ ลำดับของ ARMA (p,q) Model

2.4 ค่าวัดประสิทธิภาพของ MLR Model

วัตถุประสงค์ของการหารูปแบบของแบบจำลองเพื่อพยากรณ์ข้อมูลอนุกรมก็คือ ต้องการข้อมูลพยากรณ์ที่มีคุณสมบัติทางสถิติใกล้เคียงข้อมูลจริงเท่าที่จะเป็นไปได้ ด้วยเหตุนี้ในการพยากรณ์ข้อมูลก็จะต้องมีเกณฑ์ที่ใช้สำหรับวัดประสิทธิภาพของข้อมูลพยากรณ์ ค่าวัดประสิทธิภาพที่ใช้กันทั่วไปมีหลายค่า เช่น ผลรวมกำลังสองของค่าความคลาดเคลื่อน (Sum of Square Error, SSE) ค่าเฉลี่ยกำลังสองของค่าความคลาดเคลื่อน (Mean Square Error, MSE) ค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination, R^2) และค่าสัมประสิทธิ์การตัดสินใจปรับแล้ว (Adjust Coefficient of Determination, R_a^2) ซึ่งแต่ละค่าก็จะมีข้อจำกัดแตกต่างกันไป

ทรงศิริ (2539) กล่าวว่า หากต้องการเปรียบเทียบวิธีการพยากรณ์ข้อมูล โดยแบบจำลองหลายรูปแบบเมื่อใช้อนุกรมเวลาชุดเดียวกัน ให้ใช้ค่า Relative Error เป็นค่าวัดประสิทธิภาพของแบบจำลอง รูปแบบที่เหมาะสมกว่าคือรูปแบบที่มีค่าเฉลี่ยของค่า Relative Error ต่ำกว่า ค่า Relative Error สามารถคำนวณได้จากสมการที่ 2.14

$$\text{Relative Error} = \frac{(X_t - X_t^*)}{X_t} \quad (2.14)$$

เมื่อ	X_t	คือ	ค่าสถิติจากอนุกรมเวลาข้อมูลดิบ ณ เวลา t
	X_t^*	คือ	ค่าสถิติจากอนุกรมเวลาข้อมูลพยากรณ์ ณ เวลา t
	t	คือ	ลำดับเวลาที่ t=1, 2, ..., N

เนื่องจากค่าวัดประสิทธิภาพมีหลายค่า นักวิจัยส่วนใหญ่จะทำการพิจารณาจากหลายค่ารวมกัน แต่สำหรับการศึกษานี้จะใช้เฉพาะค่า Relative Error เป็นค่าวัดประสิทธิภาพของข้อมูลพยากรณ์จาก MLR Model เพราะเป็นค่าวัดประสิทธิภาพที่สอดคล้องกับวัตถุประสงค์ของการศึกษาครั้งนี้ที่สุด

2.5 เอกสารที่เกี่ยวข้อง

การศึกษาและการวิเคราะห์คุณสมบัติต่างของข้อมูลน้ำท่าซึ่งเป็นข้อมูลอนุกรมเวลาก็เพื่อที่จะช่วยให้สามารถจำลองกระบวนการของน้ำท่าให้ใกล้เคียงกับสภาพธรรมชาติให้มากที่สุดเท่าที่จะเป็นไปได้ โดยที่แบบจำลองที่นิยมใช้กับข้อมูลอนุกรมเวลาทางอุตุ-อุทกวิทยาก็คือแบบจำลองตามทฤษฎีของ Box & Jenkins (1970) มีรูปแบบหลักคือ ARIMA (p,d,q) Model นอกจากทฤษฎีดังกล่าวแล้ว ทฤษฎีของการถดถอยก็ได้รับความนิยมไม่น้อยไปกว่ากัน เพราะทั้งสองทฤษฎีนี้สามารถใช้ข้อมูลอนุกรมเวลาทางอุตุ-อุทกวิทยาได้เช่นเดียวกัน ในอดีตที่ผ่านมามีการนำทฤษฎีการสร้างแบบจำลองอนุกรมเวลาทั้งสองรูปแบบมาประยุกต์ใช้กับข้อมูลน้ำท่าและข้อมูลอื่นๆหลายครั้ง ทั้งแบบที่เป็นข้อมูลอนุกรมเวลารายวัน รายสัปดาห์ หรือรายปี ซึ่งแต่ละลักษณะของอนุกรมเวลาจะมีวิธีการวิเคราะห์แตกต่างกันออกไป ดังนั้นผู้ศึกษาจึงได้ทำการรวบรวมรายละเอียดของการสร้างแบบจำลองจากแหล่งต่างๆ ดังนี้

Quimpo (1967) ได้ศึกษาการสร้างแบบจำลองสำหรับอนุกรมเวลาการไหลรายวันของกลุ่มน้ำทั้งหมด 17 แห่งทั่วสหรัฐอเมริกา เพื่อหาโครงสร้างของแบบจำลองและแยกส่วนที่เป็น Deterministic Component ออกจาก Stochastic Component ของอนุกรมเวลา โดยก่อนจะนำ

อนุกรมเวลาการไหลรายวันไปวิเคราะห์ด้วย Autoregressive Model (AR Model) ได้ทำการ Standardized อนุกรมเวลาการไหลรายวัน เพื่อลดผลกระทบจากความแปรปรวนของอนุกรมเวลาการไหลรายวัน ผลจากการวิเคราะห์ด้วย AR Model พบว่า ค่าความคลาดเคลื่อนของอนุกรมเวลาการไหลรายวันจากการพยากรณ์ของ AR (2) Model ให้ผลเป็นที่น่าพอใจ

Yevjevich (1975) ทำการศึกษาการสังเคราะห์ข้อมูลอุทกวิทยาโดยใช้ Autoregressive Model (AR Model) โดยนำข้อมูล Monthly Net Basin Supplies (NBS) ของทะเลสาบ Michigan-Huron ซึ่งตั้งอยู่ระหว่างชายแดนของสหรัฐอเมริกาและแคนาดามาเป็นกรณีศึกษา จากการศึกษาพบว่า การประมาณค่าพารามิเตอร์ของ AR Model โดยใช้วิธี Moments และวิธี Maximum Likelihood ในกรณีที่ตัวอย่างข้อมูลมีขนาดใหญ่จะทำให้ผลลัพธ์ที่ไม่ต่างกัน ผลการทดสอบความเหมาะสมของรูปแบบตั้งแต่ AR (1) Model – AR (3) Model โดยการพิจารณารูป Residual Correlograms และการใช้ค่า AIC เป็นเกณฑ์ พบว่ารูปแบบที่เหมาะสมกับการสังเคราะห์ข้อมูล NBS ที่สุดคือ AR (2) Model เพราะเป็นรูปแบบที่ค่า Autocorrelation of Residual อยู่ในช่วงความเชื่อมั่นที่ 95 % และเป็นรูปแบบมีค่า AIC ต่ำสุด

Hipel et al. (1977) ทำการศึกษาและพัฒนาเทคนิคการสร้างแบบจำลองสำหรับอนุกรมเวลาตามวิธีของ Box-Jenkins ให้เข้าใจได้ง่ายขึ้น โดยได้เสนอทฤษฎีและเทคนิคขั้นสูงที่จะใช้เป็นเครื่องช่วยในการพิจารณาเลือกรูปแบบที่เหมาะสมของแบบจำลอง (Identification) การประมาณค่าพารามิเตอร์ (Estimation) และการตรวจสอบประสิทธิภาพของ ARIMA Model และได้ทำการพิสูจน์แล้วว่าค่า Inverse Autocorrelation Function (IACF) และค่า Inverse Partial Autocorrelation Function (IPACF) เป็นตัวช่วยในการตัดสินใจเลือกรูปแบบที่เหมาะสมให้กับแบบจำลองสามารถใช้ได้กับทั้ง 2 แบบ คือ Non-Seasonal และ Seasonal Model ในขั้นตอนของการประมาณค่าพารามิเตอร์ หากอนุกรมเวลาที่นำมาวิเคราะห์เป็นอนุกรมเวลาทั่วไปที่ยังไม่ผ่านการแปลงค่าวิธี Modified Sum of Square จะประมาณค่าพารามิเตอร์ได้มีประสิทธิภาพมากขึ้น แต่หากข้อมูลผ่านการแปลงค่า (Transforms) โดยวิธีของ Box-Cox Power Transformation มาแล้วควรใช้วิธี Maximum Likelihood ประมาณค่าพารามิเตอร์ของแบบจำลอง

McLeod et al. (1977) ได้ใช้เทคนิคการสร้างแบบจำลองวิธีใหม่ที่ได้คิดค้นและพัฒนาต่อจากวิธีของ Hipel et al. (1977) มาใช้ในการสร้างแบบจำลอง ARIMA (Autoregressive Integrated Moving Average) Model เพื่อจำลองการไหลรายปีเฉลี่ยของแม่น้ำ Saint Lawrence River ในสหรัฐอเมริกา พบว่าแบบจำลองที่สร้างขึ้นมีประสิทธิภาพกว่าแบบจำลองที่วิเคราะห์โดย Carlson et al. (1970)

Miller et al. (1981) ทำการศึกษาการจำลองอนุกรมเวลาน้ำท่ารายวันเฉลี่ยของ Minomonee River ในสหรัฐอเมริกา โดยใช้ Univariate Time Series Model สร้างอนุกรมเวลาน้ำท่ารายวันสังเคราะห์ เพื่อเปรียบเทียบกับการสร้างอนุกรมเวลาน้ำท่ารายวันสังเคราะห์ของ Regression Model ซึ่งได้นำเอาความสัมพันธ์ของน้ำฝนกับการไหลในลำน้ำมาวิเคราะห์ร่วมกันในแบบจำลอง โดยมีตัวแปรอิสระเป็นค่าลอการิทึมของน้ำท่าของช่วงเวลาก่อนหน้า 1 คาบเวลาและ 2 คาบเวลา ค่ากำลังสองของน้ำฝนในปัจจุบัน ค่าน้ำฝนของช่วงเวลาก่อนหน้า 1 คาบเวลา และ 2 คาบเวลา พบว่าการสร้างอนุกรมเวลาของน้ำท่ารายวันสังเคราะห์ของ Regression Model มีประสิทธิภาพกว่าการสร้างข้อมูลของ Univariate Time Series Model และได้แนะนำให้ใช้เทคนิคการประมาณค่าพารามิเตอร์ของ Random Coefficient Model แทนการประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุด

Seckler (1982) ทำการศึกษาระเหยของทะเลสาบ Hefner ในรัฐ Oklahoma และที่ Fort Collins ในรัฐ Colorado โดยใช้ Time Series Multiple Linear Regression Models ในการประมาณค่าพารามิเตอร์ Constant (C), Mass Transfer Coefficient (N_t), เลขชี้กำลัง m และ p ของสมการ Mass Transfer Evaporation, $E_t = CN_t u_{z,t}^m \Delta e_{z,t}^p$ เมื่อ $u_{z,t}$ คือ ความเร็วลมที่ระดับความสูง z เหนือระดับผิวน้ำ ณ เวลา t และ $\Delta e_{z,t}$ คือ ค่าความแตกต่างของความดันไอที่ระดับความสูง z เหนือระดับผิวน้ำ ณ เวลา t จากการศึกษาพบว่าสมการการระเหยของทะเลสาบ Hefner คือ $E_t = 0.42 * 10^{-2} (k' / \ln(z/z_0))^{0.80} u_{z,t} \Delta e_{z,t}$ และสมการการระเหยของ Fort Collins คือ $E_t = 0.0045 u_{z,t}^{0.50} \Delta e_{z,t}$ เมื่อนำสมการไปใช้พยากรณ์การระเหยพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ระหว่างค่าการระเหยสังเกตกับค่าการระเหยพยากรณ์ของอนุกรมเวลาจากทะเลสาบ Hefner และ Fort Collins ที่คำนวณได้มีค่าเท่ากับ 0.77 และ 0.95 ตามลำดับ จากนั้นได้นำสมการการระเหยของทะเลสาบ Hefner ไปประยุกต์ใช้กับอนุกรมเวลาการระเหยจากอ่างเก็บน้ำ Elephant Butter Reservoir ใน New Mexico พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าการระเหยสังเกตกับค่าการระเหยพยากรณ์ที่คำนวณได้มีค่าเท่ากับ 0.95

Saw Suat Kheng (1983) ได้ใช้เทคนิคและวิธีการวิเคราะห์อนุกรมเวลาของ Box-Jenkins เพื่อหารูปแบบที่จะสามารถจำลองและสังเคราะห์ข้อมูล Hourly Wind Speed and Direction ของประเทศมาเลเซีย จากการศึกษาพบว่ารูปแบบ ARIMA(1,0,1) x (0, 1, 1)₂₄ เป็นรูปแบบที่เหมาะสม นอกจากนี้ยังพบว่า การผสมเทคนิควิธี Differencing วิธี Harmonic Analysis และการ Standardization เข้าด้วยกันยังไม่สามารถจัดอิทธิพลของความแปรปรวนประจำวันออกได้ทั้งหมด อย่างไรก็ตาม เงื่อนไขในการกำหนดความเหมาะสมของแบบจำลองยังคงใช้ค่า Absolute Error ของค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐาน เป็นเกณฑ์

Saadat Ayub Khan (2006) สร้างแบบจำลองน้ำฝน-น้ำท่า (Rainfall-Runoff Modeling) โดยวิธีการทางสถิติและวิธี Data Driven เพื่อพยากรณ์ระดับของลุ่มน้ำแม่น้ำ Ouse ทางตอนเหนือของประเทศอังกฤษ รูปแบบที่ศึกษาประกอบด้วย Multiple Linear Regression (MLR) และเทคนิคการจำลอง Data Driven ที่แตกต่างกันอีก 3 วิธี คือ Neural Network (NN) Evolutionary Neural Networks (Evol-NN) และ M5 Model Trees ช่วงเวลาที่จะพยากรณ์ค่าลุ่มน้ำคือ ทุกๆ 6 ชั่วโมง และทุกๆ 24 ชั่วโมง โดยเลือกใช้ค่าสถิติวัดความเหมาะสมของรูปแบบ (Goodness-of-Fit Statistics) ดังนี้ ค่า Root Mean Squared (RMSE) ค่า Mean Absolute Error (MAE) ค่า Coefficient of Efficiency และค่า Coefficient of Determination (R^2) ผลการศึกษาพบว่าโดยทั่วไปวิธี Data Driven มีประสิทธิภาพมากกว่าวิธีการทางสถิติ และ M5 Model Trees เป็นแบบจำลองที่มีศักยภาพสูงสุดสำหรับการสร้างการพยากรณ์ระดับน้ำที่มีลักษณะค่อนข้างชัดเจน

Sinnakaudan et al. (2006) ใช้แบบจำลอง Multiple Linear Regression ในการพยากรณ์ปริมาณตะกอนที่ท้องน้ำของแม่น้ำต่างๆ ในประเทศมาเลเซีย ข้อมูลทางชลศาสตร์และข้อมูลตะกอนที่นำมาศึกษามีจำนวน 364 ตัวอย่าง โดยรวบรวมจากแม่น้ำตามธรรมชาติและแม่น้ำที่มนุษย์สร้างขึ้นจำนวน 9 สาย ซึ่งเป็นแม่น้ำที่อยู่ลุ่มน้ำที่มีคุณลักษณะแตกต่างกันไป ตัวแปรต่างๆ ที่นำมาใช้กับแบบจำลองผ่านการคัดเลือกอย่างระมัดระวัง โดยอาศัยหลักการจากการสำรวจงานวิจัยที่ผ่านมา และจากประสบการณ์ภาคสนาม จากการทดสอบและจัดกลุ่มตัวแปรสามารถแบ่งได้เป็น 5 พารามิเตอร์ คือ ความสามารถในการเคลื่อนไหว (Mobility) การเคลื่อนย้าย (Transport) การตกตะกอน (Sediment) รูปร่าง (Shape) และการต้านทานการไหล (Flow Resistance) พารามิเตอร์ที่มีอิทธิพลมากที่สุดจากแต่ละกลุ่มจะถูกเลือกเข้าไปในแบบจำลองโดยวิธีการที่เป็นไปได้ เงื่อนไขทางสถิติที่ใช้ในการคัดเลือกแบบจำลองที่เหมาะสมที่สุด คือค่า R-square ค่า Adjusted R-square ค่า Mean Square Error และค่า Mallows' C_p ความแม่นยำของแบบจำลองหลากหลายรูปแบบ

ตัดสินจากค่า Discrepancy Ratio ซึ่งเป็นสัดส่วนระหว่างค่าคำนวณและค่าจริง แบบจำลองที่ใช้ งานได้ดีที่สุดซึ่งหมายถึงแบบจำลองที่มีเปอร์เซ็นต์การพยากรณ์ข้อมูลที่ใช้งานได้แม่นยำสูงสุด โดยปกติแบบจำลองที่เหมาะสมที่สุดสำหรับแม่น้ำที่มีการกระจายอย่างสม่ำเสมอของตะกอนที่อง น้ำขนาด d_{50} อยู่ในช่วง 0.37-0.4 mm และใช้การได้ดีกว่าสมการตะกอนที่องน้ำของ Graf สมการ ของ Yang และสมการของ Ackers-White

2.6 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง

ลักษณะของงานวิจัยที่ผ่านมา ส่วนใหญ่จะเป็นการนำข้อมูลอนุกรมเวลาทางอุตุ-อุทก วิทยามาวิเคราะห์ เพื่อหาแบบจำลองที่จะสามารถจำลองและสร้างข้อมูลสังเคราะห์ให้กับ แหล่งข้อมูลที่ต้องการได้ ข้อมูลอนุกรมเวลาที่นำมาศึกษาสามารถแบ่งได้เป็น อนุกรมเวลารายปี และอนุกรมคาบเวลา เช่น รายวัน รายสัปดาห์ รายเดือน เมื่อข้อมูลอนุกรมเวลามีคาบเวลาที่ ต่างกัน วิธีการและเทคนิคการวิเคราะห์ย่อมแตกต่างกัน ดังนั้น การสรุปสาระสำคัญหรือเทคนิค วิธีของการวิเคราะห์หอนุกรมคาบเวลาของงานวิจัยดังกล่าว สามารถนำมาประยุกต์ใช้กับการศึกษานี้ ที่ทำการวิเคราะห์และหาแบบจำลองโดยใช้ข้อมูลอนุกรมคาบเวลาเช่นเดียวกัน มีรายละเอียดดังนี้

- การแปลงค่าข้อมูลด้วยวิธี Box-Cox Transformation จะช่วยลดความแปรปรวนของ อนุกรมคาบเวลาได้ดี
- การขจัดอิทธิพลของฤดูกาลออกจากข้อมูลอนุกรมคาบเวลาโดยการ Standardization จะทำให้แบบจำลองที่วิเคราะห์ได้มีประสิทธิภาพยิ่งขึ้น
- การทดสอบความเป็นอิสระของค่าคลาดเคลื่อน โดยการพิจารณากราฟ Correlogram of Residual เป็นวิธีที่เหมาะสมกับอนุกรมเวลารายวันที่มีจำนวนตัวอย่างมาก
- การใช้ค่า AIC เป็นเกณฑ์ในการเลือกรูปแบบที่เหมาะสม มีข้อดีหลายอย่าง คือ ไม่ ยุ่งยากในการคำนวณ สะดวกต่อการนำไปใช้ และผลที่ได้จากการวิเคราะห์ก็ไม่แตกต่างจากการ เลือกรูปแบบโดยใช้เกณฑ์อื่นๆ

บทที่ 3

วิธีดำเนินการวิจัย

การดำเนินการวิจัยในส่วนของ Multiple Linear Regression Model (MLR Model) จะดำเนินการตามแผนผังในรูป 3.1 โดยมีรายละเอียดดังนี้

3.1 วิธีดำเนินการวิจัยของ Multiple Linear Regression Model

3.1.1 ข้อมูลที่จะต้องรวบรวมเพิ่มเติมจากข้อมูลน้ำท่ารายเดือน คือ ข้อมูลน้ำฝนรายเดือนของสถานีอุตุนิยมวิทยาที่มีความสัมพันธ์ที่ดี สามารถนำมาประเมินค่าของข้อมูลน้ำท่าของสถานีตัวอย่างได้ จากการรวบรวมข้อมูลหลักทั้ง 2 อย่างผลปรากฏว่า ข้อมูลน้ำฝนที่นำมาใช้วิเคราะห์ร่วมกับข้อมูลน้ำท่าสถานีเขื่อนปัตตานีเป็นข้อมูลน้ำฝนจากสถานีเขื่อนบาง สถานีเมืองยะลา และสถานีเขื่อนปัตตานี ซึ่งตรวจวัดปริมาณน้ำฝนจากบริเวณใกล้เคียงกับสถานีวัดน้ำท่าที่สุดเท่าที่หาได้ ตำแหน่งของสถานีแสดงในรูป 3.2

3.1.2 ตรวจสอบว่ามีข้อมูลน้ำฝนครบทุกวันหรือไม่ ผลการตรวจสอบ พบว่าทุกสถานีมีข้อมูลครบถ้วน

3.1.3 กำหนดตัวแปรตามและตัวแปรอิสระ แล้วตรวจสอบระดับและทิศทางของความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระโดยใช้กราฟ Scatter Plots Diagram ประกอบกับ Correlation Matrix ในเบื้องต้นให้ใช้ค่าข้อมูลปริมาณน้ำท่า ณ ปัจจุบันแปลงค่าเป็นตัวแปรตาม ตัวแปรอิสระประกอบด้วยข้อมูลปริมาณน้ำท่าแปลงค่าที่ย้อนเวลา 1 คาบเวลา 2 คาบเวลา ค่าข้อมูลปริมาณน้ำฝน ณ ปัจจุบัน และ ณ เวลาต่างๆ ย้อนไปประมาณ 2 คาบเวลา

3.1.4 ประมาณค่าพารามิเตอร์ของ MLR Model โดยใช้วิธีกำลังสองน้อยที่สุด

3.1.5 ใช้วิธี Stepwise เลือกตัวแปรอิสระให้กับรูปแบบ MLR Model ที่ดีที่สุด

3.1.6 ตรวจสอบความเหมาะสมของรูปแบบจากค่าคลาดเคลื่อนที่ได้จากแบบจำลอง $\hat{\epsilon}_t$ ตามสมมติฐานที่กำหนดไว้ของ MLR Model กล่าวคือ ค่าคลาดเคลื่อน เป็นตัวแปรที่มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับศูนย์ ค่าความแปรปรวนเป็นค่าคงที่ใดๆ และค่าคลาดเคลื่อนเป็นอิสระต่อกัน หากคุณสมบัติของค่าคลาดเคลื่อนจากแบบจำลอง $\hat{\epsilon}_t$ ไม่ตรงกับสมมติฐานที่กำหนด

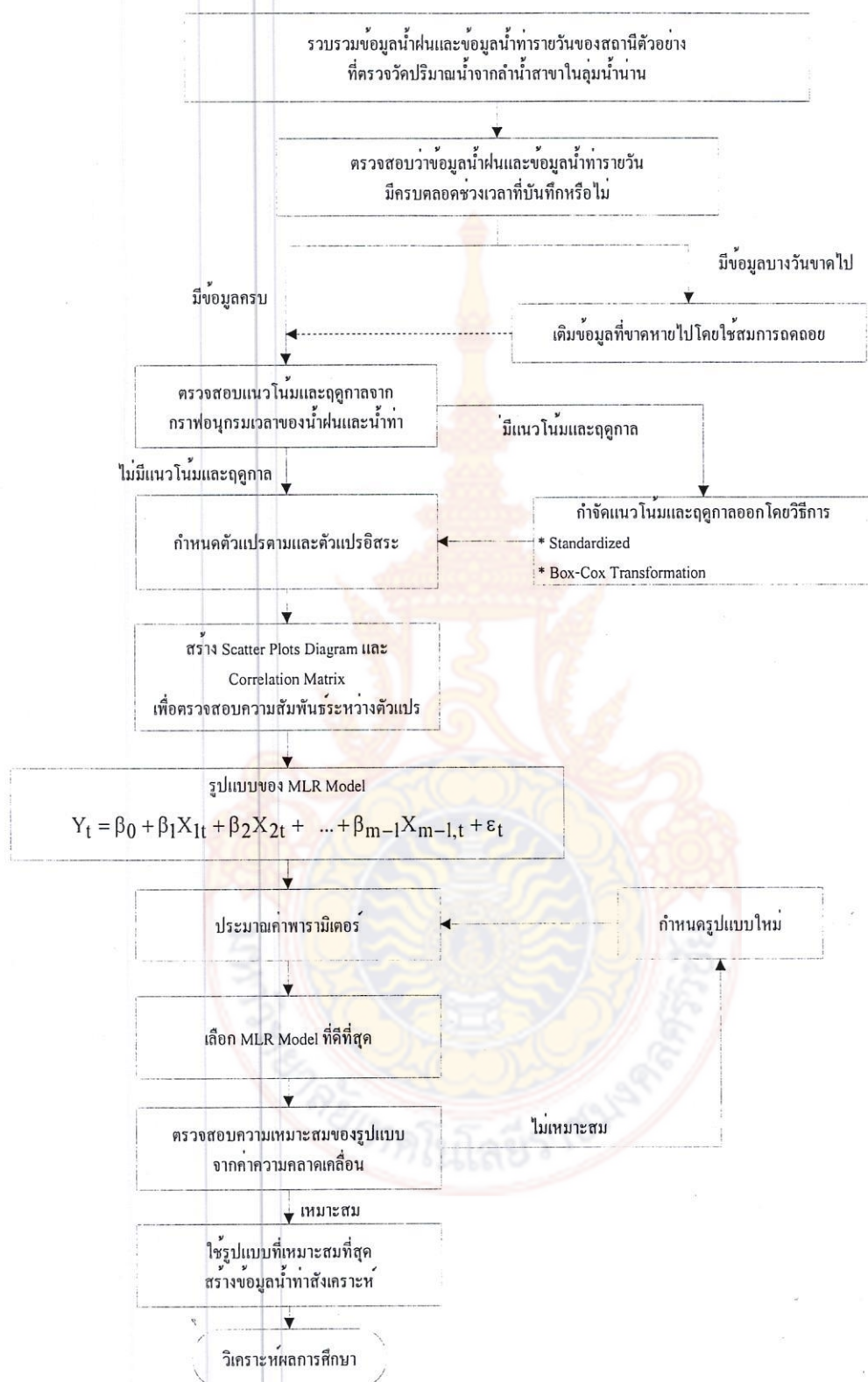
ไว้ให้แปลงค่าตัวแปรตามและตัวแปรอิสระ โดยใช้วิธี Box-Cox Transformation จากนั้นจึงกลับไปทำขั้นตอนที่ 3.1.3

3.1.7 ทำการพยากรณ์ข้อมูลปริมาณน้ำท่าโดยใช้รูปแบบ MLR Model ที่เหมาะสมที่สุดจากข้อมูลทั้ง 3 ชุด จากนั้นนำข้อมูลน้ำท่าพยากรณ์มาคำนวณหาค่าสถิติเบื้องต้น

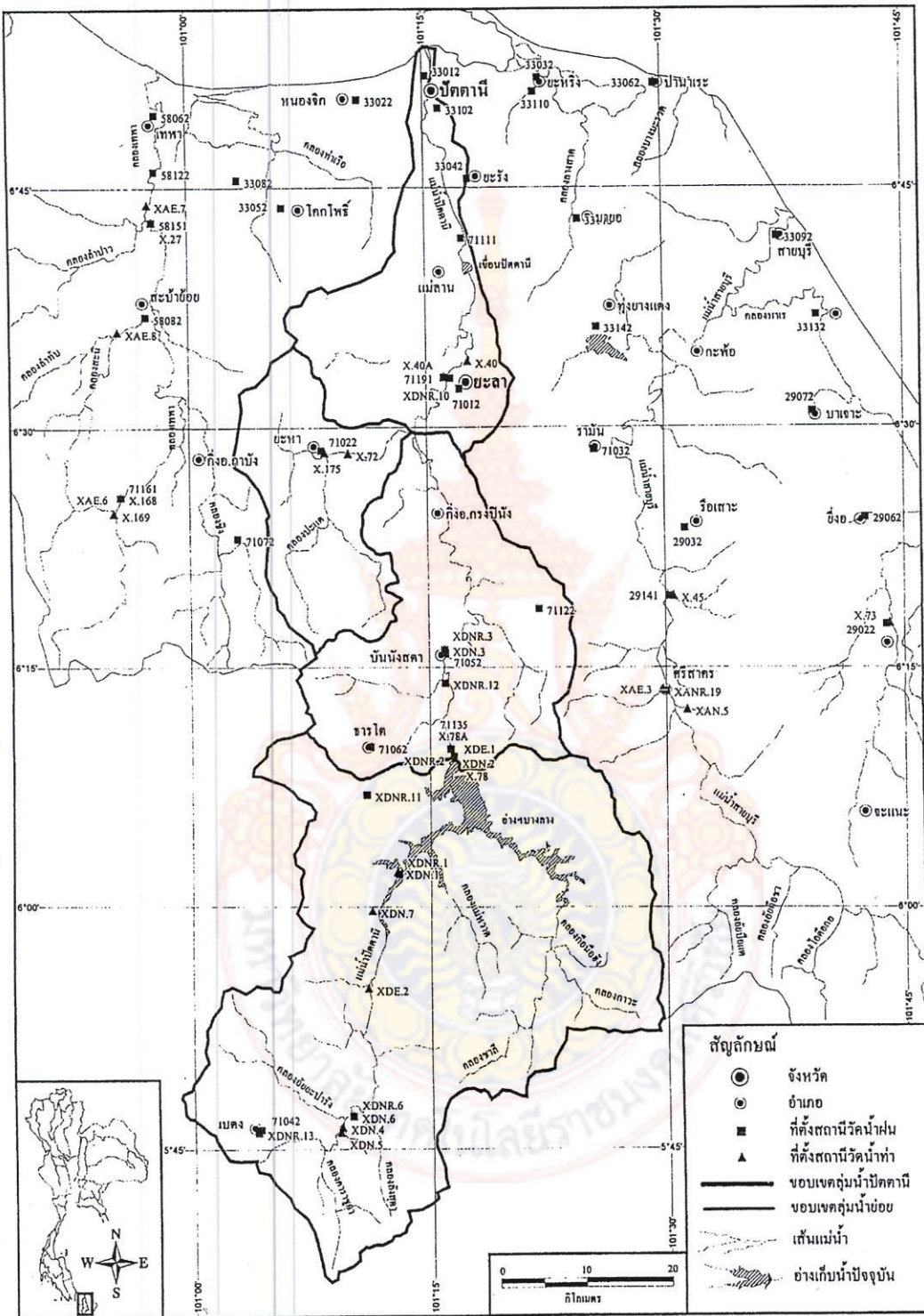
3.1.8 ทำการเปรียบเทียบค่าสถิติเพื่อการตัดสินใจ (Goodness-of-Fits Statistics) เพื่อหาข้อสรุปว่าแบบจำลองใดมีประสิทธิภาพเหนือกว่า

3.1.9 สรุปผลและข้อเสนอแนะ





รูป 3.1 แผนผังแสดงขั้นตอนการสร้าง Multiple Linear Regression Model



รูป 3.2 แผนที่ตำแหน่งสถานีวัดน้ำท่าและสถานีวัดน้ำฝนของกลุ่มน้ำปัตตานี

ที่มา : โครงการส่งและบำรุงรักษาปัตตานี

บทที่ 4

ผลการวิจัยและวิจารณ์

ในบทนี้จะนำเสนอผลการหาแบบจำลองที่เหมาะสมกับข้อมูลน้ำท่ารายเดือน ซึ่งได้แบ่งผลการวิจัยออกเป็น 3 ส่วนคือ ส่วนแรกเป็นผลการวิเคราะห์คุณสมบัติเบื้องต้นและการเตรียมข้อมูล ส่วนที่สองเป็นผลการประมาณค่าพารามิเตอร์ของแบบจำลอง และส่วนสุดท้ายเป็นการตรวจสอบความเหมาะสมของแบบจำลอง โดยจะกล่าวรายละเอียดดังต่อไปนี้

4.1 การวิเคราะห์คุณสมบัติเบื้องต้นและการเตรียมข้อมูล

(Preliminary Analysis and Data Preparation)

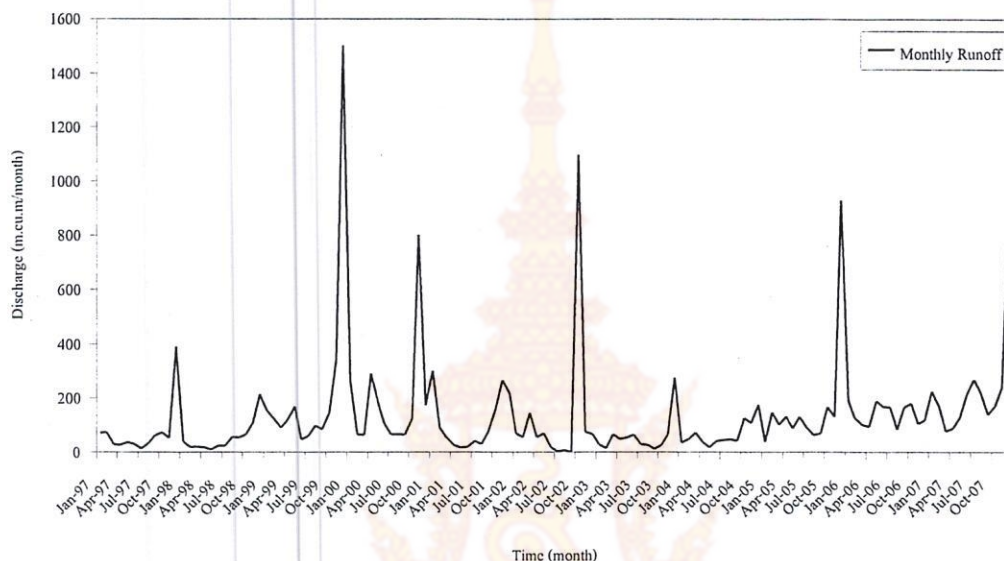
ข้อมูลที่นำมาศึกษาในครั้งนี้เป็นข้อมูลปริมาณน้ำท่ารายเดือนสถานีวัดน้ำเขื่อนปัตตานี (X.79) จำนวน 1 สถานี และสถานีวัดน้ำฝนรายเดือนในพื้นที่ลุ่มน้ำปัตตานีจำนวน 3 สถานี ประกอบด้วย สถานีวัดปริมาณน้ำฝนเขื่อนบางลาง (71135) สถานีวัดปริมาณน้ำฝนเมืองยะลา (71191) และสถานีวัดปริมาณน้ำฝนเขื่อนปัตตานี (7111) ช่วงเวลาของข้อมูลทั้งสิ้น 11 ปี เริ่มจากปี ค.ศ. 1997 - 2007 ข้อมูลปริมาณน้ำท่าและข้อมูลปริมาณน้ำฝนรายเดือนจัดอยู่ในประเภท “อนุกรมคาบเวลา” โดยทั่วไปจะแทนด้วย $X_{v,\tau}$ ที่มีตัวแปรย่อยบอกคาบเวลาอย่างชัดเจน เมื่อ v คือ ลำดับปีที่ $v=1, \dots, n$ และ τ คือ ลำดับคาบเวลาที่ $\tau=1, \dots, \omega$ (สำหรับอนุกรมเวลารายเดือน $\omega=12$) รายละเอียดของข้อมูลแสดงไว้ในภาคผนวก ก

4.1.1 การวิเคราะห์คุณสมบัติเบื้องต้นของข้อมูลปริมาณน้ำท่ารายเดือน

วัตถุประสงค์ของการวิเคราะห์คุณสมบัติทางสถิติเบื้องต้นของข้อมูลดิบ เช่น ค่าเฉลี่ย แทนด้วยสัญลักษณ์ \bar{x} ค่าเบี่ยงเบนมาตรฐาน แทนด้วยสัญลักษณ์ S หรือค่าสัมประสิทธิ์ความเบ้ แทนด้วยสัญลักษณ์ C_s เพื่อตรวจสอบว่าสามารถใช้ข้อมูลดิบในการหารูปแบบของ MLR Model ได้โดยตรงหรือต้องทำการแปลงค่าข้อมูลดิบก่อนนำไปวิเคราะห์ต่อไป

กราฟอนุกรมเวลาปริมาณน้ำท่ารายเดือนสถานีเขื่อนปัตตานี (X.79) ในรูป 4.1 เห็นชัดว่ามีลักษณะของฤดูกาล เพราะการเคลื่อนไหวของกราฟมีลักษณะซ้ำๆ กันทุกปีซึ่งเป็นธรรมชาติของข้อมูลอนุกรมเวลา ช่วงเวลาที่อัตราการไหลอยู่ในเกณฑ์สูงกว่า 800 ล้านลูกบาศก์

เมตรต่อเดือน มี 5 ช่วง คือ ธันวาคม ปี ค.ศ.1999 พฤศจิกายน ปี ค.ศ.2000 ตุลาคม ปี ค.ศ.2002 ธันวาคม ปี ค.ศ.2005 และ ปี ค.ศ.2007 ตามลำดับ สำหรับช่วงอื่นๆ อัตราการไหลอยู่ในเกณฑ์ปกติ



รูป 4.1 กราฟอนุกรมเวลาปริมาณน้ำท่ารายเดือนของสถานีเขื่อนปัตตานี (ปี ค.ศ.1997-2007)

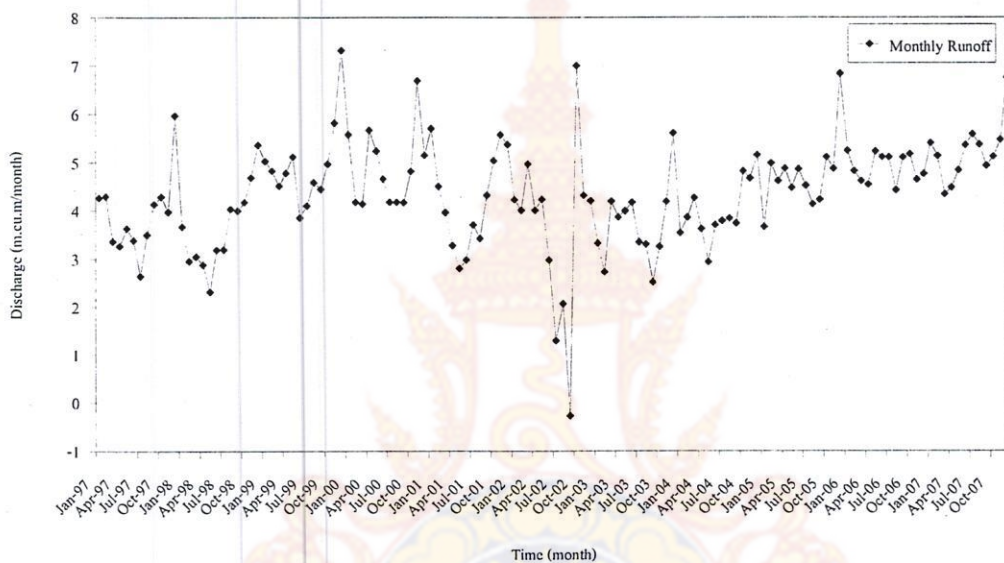
นอกจากนี้ข้อมูลปริมาณน้ำท่ารายเดือนยังเป็นข้อมูลที่มีความแปรปรวนของค่าสถิติสูงดังจะเห็นจากคุณสมบัติทางสถิติเบื้องต้นในตาราง 4.1 ค่าเฉลี่ยของข้อมูลปริมาณน้ำท่ารายเดือนเฉลี่ย $avg.\bar{x}_T$ มีค่าเท่ากับ 134.39 ค่าเฉลี่ยของส่วนเบี่ยงเบนมาตรฐานรายเดือนเฉลี่ย $avg.S_T$ มีค่าเท่ากับ 129.61 และค่าเฉลี่ยของค่าสัมประสิทธิ์ความเบ้รายเดือนเฉลี่ย $avg.Cs_T$ มีค่าเท่ากับ 1.38 เมื่อค่า $avg.Cs_T$ ไม่เข้าใกล้ค่าศูนย์ก็ไม่อาจยอมรับได้ว่าข้อมูลปริมาณน้ำท่ารายเดือนมีการแจกแจงแบบปกติ

ตาราง 4.1 คุณสมบัติทางสถิติเบื้องต้นของข้อมูลน้ำท่ารายเดือนสถานี X.79

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average
Mean, \bar{x}_T	128.27	84.62	74.09	84.74	87.76	78.24	73.23	65.98	69.90	210.91	207.86	447.02	134.39
Standard Deviation, S_T	101.72	51.34	34.87	78.55	69.45	66.76	78.71	56.42	48.46	297.91	217.21	453.88	129.61
Skewness, Cs_T	0.58	0.27	0.37	2.01	0.48	1.07	1.86	2.13	0.72	3.15	2.37	1.56	1.38

เมื่อนำข้อมูลปริมาณน้ำท่ารายเดือนมาแปลงค่าด้วยวิธี Box-Cox Transformation แล้วนำมาพล็อตกราฟอนุกรมเวลาดังรูป 4.2 พบว่าอิทธิพลของฤดูกาลดังที่ปรากฏในรูป 4.1 ถูกขจัดออก จากนั้นหาค่าคุณสมบัติทางสถิติเบื้องต้นดังตาราง 4.2 พบว่า ค่าเฉลี่ยของข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่าเฉลี่ย $avg.\bar{x}_T$ มีค่าเท่ากับ 4.34 ค่าเฉลี่ยของส่วนเบี่ยงเบนมาตรฐานราย

เดือนเฉลี่ย $avg.S_T$ มีค่าเท่ากับ 0.93 และค่าเฉลี่ยของค่าสัมประสิทธิ์ความเบ้รายเดือนเฉลี่ย $avg.Cs_T$ มีค่าเท่ากับ -0.26 เมื่อค่า $avg.Cs_T$ จากคุณสมบัติทางสถิติเบื้องต้นที่ปรากฏอาจยอมรับได้ว่าข้อมูลน้ำท่าแปลงค่ามีการแจกแจงแบบปกติ ช่วยทำให้สามารถประมาณค่าพารามิเตอร์ของ MLR Model ในขั้นตอนต่อไปได้อย่างมีประสิทธิภาพมากขึ้น ซึ่งสอดคล้องกับการศึกษาของ Hipel et al. (1977) และ Miller et al. (1981)



รูป 4.2 กราฟอนุกรมเวลาปริมาณน้ำท่ารายเดือนของสถานีเขื่อนปัตตานี (X.79) ที่แปลงค่าด้วยวิธี Box-Cox Transformation (ปี ค.ศ.1997-2007)

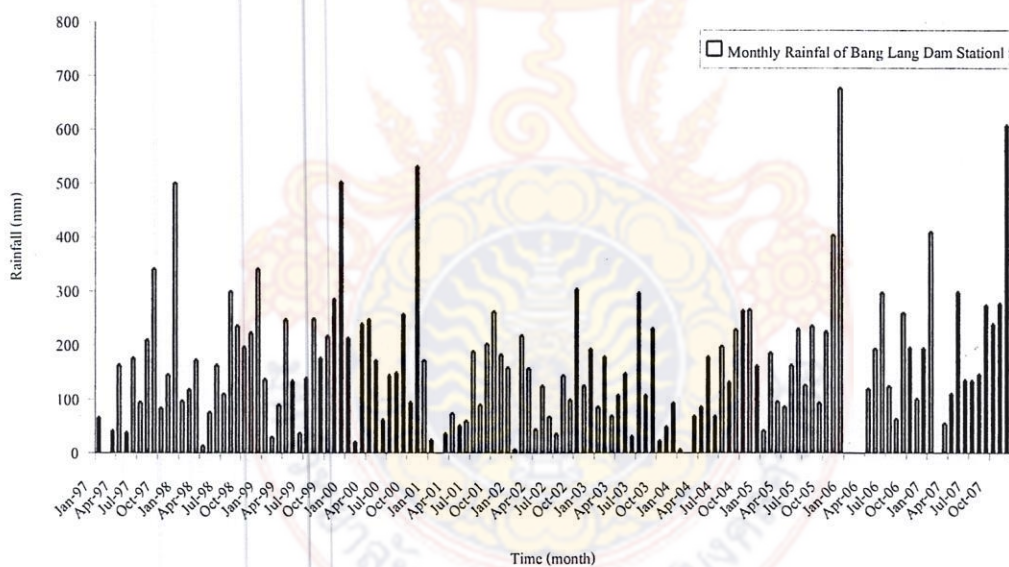
ตาราง 4.2 คุณสมบัติทางสถิติเบื้องต้นของข้อมูลปริมาณน้ำท่ารายเดือนสถานีเขื่อนปัตตานี (X.79) ที่แปลงค่าด้วยวิธี Box-Cox Transformation

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average
Mean, x_T	4.51	4.21	4.18	4.11	4.07	4.02	3.77	3.89	3.75	4.89	4.99	5.68	4.34
Standard Deviation, S_T	0.90	0.79	0.56	0.84	1.04	0.87	1.18	0.85	1.50	0.91	0.81	0.95	0.93
Skewness, Cs_T	0.04	-0.86	-0.92	0.30	-0.46	0.27	-0.59	-0.54	-2.27	0.73	0.85	0.39	-0.26

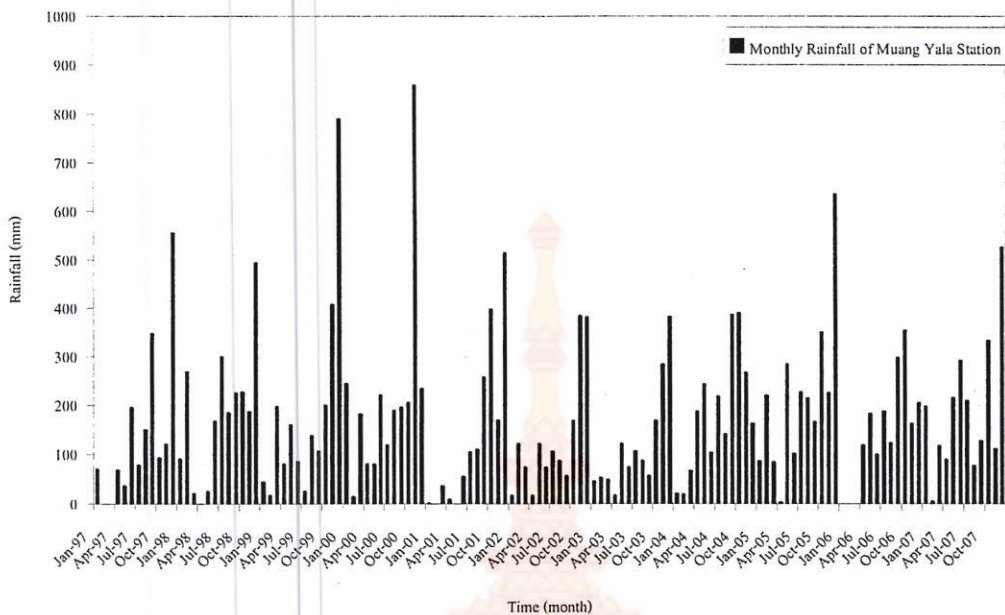
4.1.2 การวิเคราะห์คุณสมบัติเบื้องต้นของข้อมูลปริมาณน้ำฝนรายเดือน

ข้อมูลปริมาณน้ำฝนรายเดือนที่ใช้ในการศึกษานี้มีทั้งสิ้น 3 สถานี ประกอบด้วย สถานีวัดปริมาณน้ำฝนเขื่อนบางลาง (71135) สถานีวัดปริมาณน้ำฝนเมืองยะลา (71191) และสถานีวัดปริมาณน้ำฝนเขื่อนปัตตานี (7111) การวิเคราะห์คุณสมบัติทางสถิติเบื้องต้นของข้อมูลปริมาณน้ำฝนรายเดือนมีความซับซ้อนน้อยกว่าการวิเคราะห์คุณสมบัติทางสถิติเบื้องต้นของข้อมูลปริมาณน้ำท่ารายเดือน ดังรายละเอียดต่อไปนี้

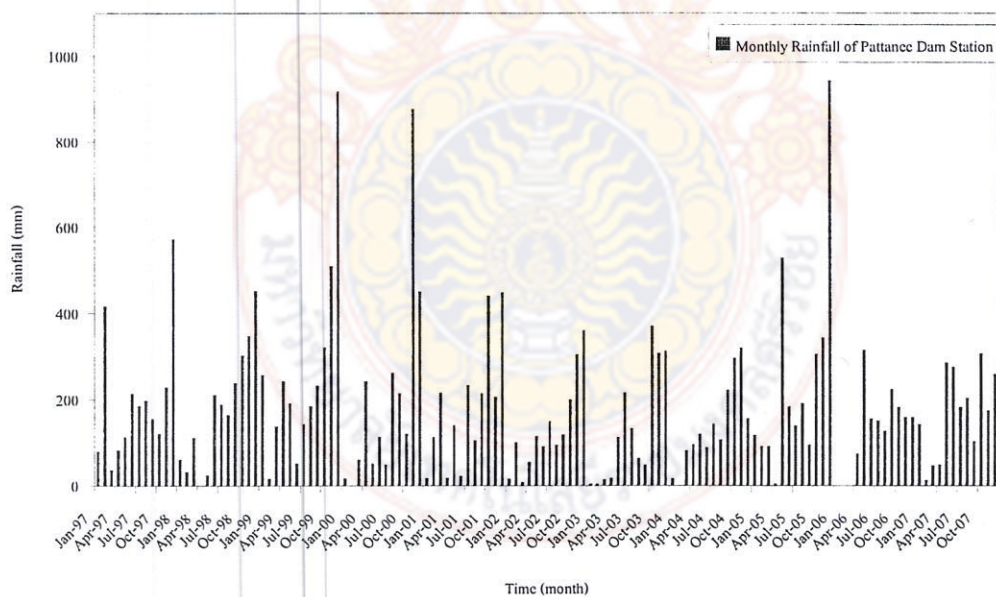
ในเบื้องต้นให้ตรวจสอบว่าข้อมูลปริมาณน้ำฝนรายเดือนของทั้ง 3 สถานี มีจำนวนครบถ้วนและสอดคล้องกับข้อมูลน้ำท่ารายเดือนหรือไม่ โดยสามารถพิจารณาได้จากกราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนในรูป 4.3 ถึงรูป 4-5 เทียบกับกราฟอนุกรมเวลาปริมาณน้ำท่ารายเดือนในรูป 4.1



รูป 4.3 กราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนของสถานีเขื่อนบางลาง (ปี ค.ศ.1997-2007)

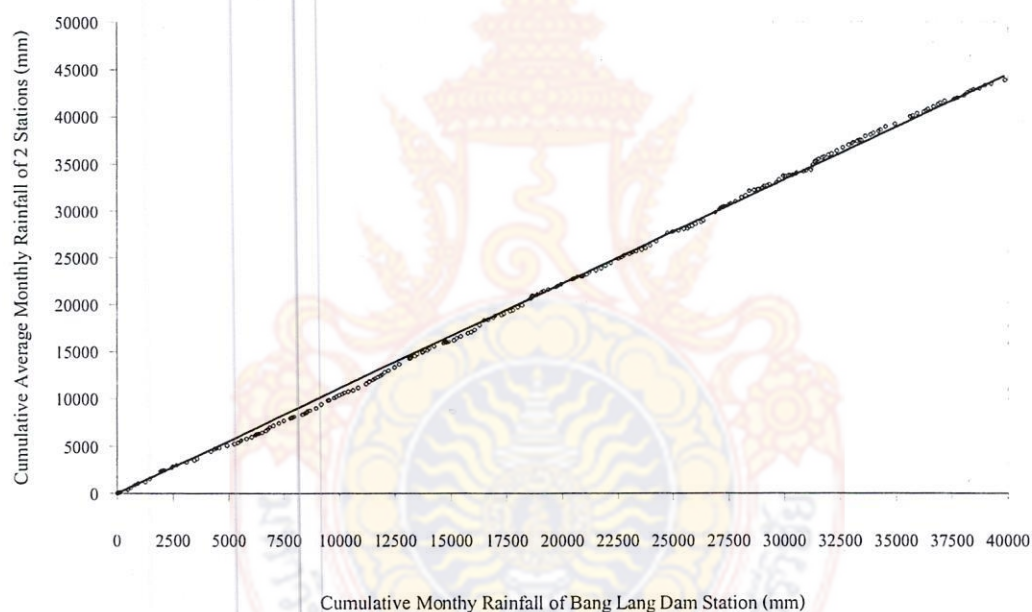


รูป 4.4 กราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนของสถานีเมืองยะลา (ปี ค.ศ.1997-2007)

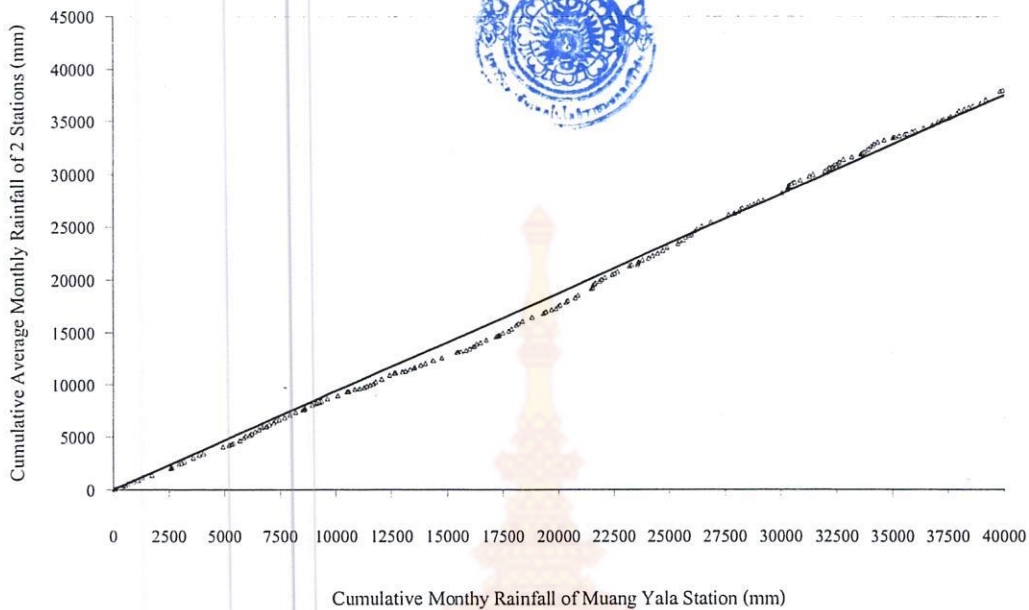


รูป 4.5 กราฟอนุกรมเวลาปริมาณน้ำฝนรายเดือนของสถานีเขื่อนปัตตานี (ปี ค.ศ.1997-2007)

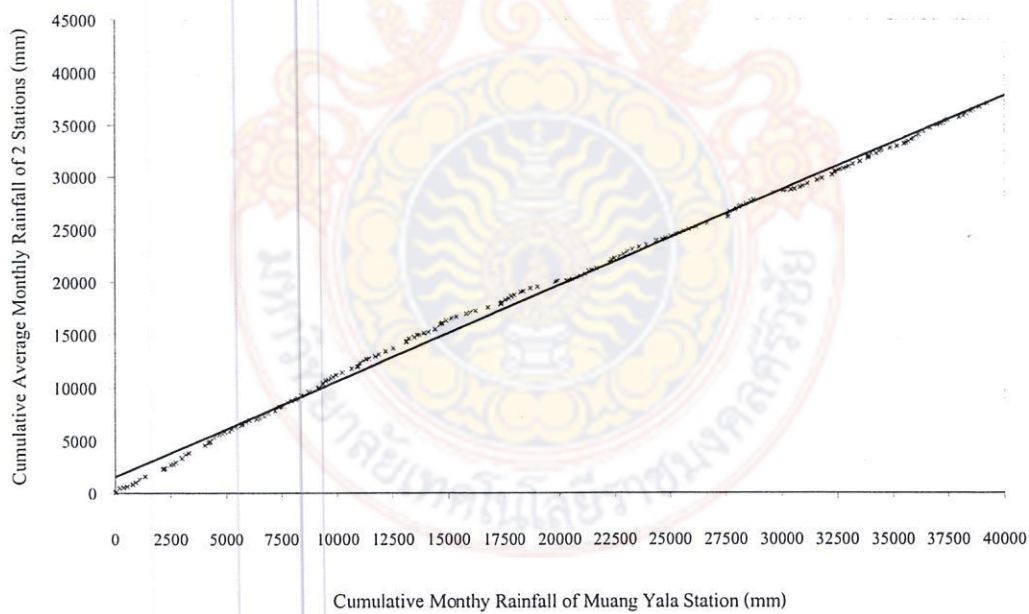
จากนั้นเป็นการตรวจสอบความน่าเชื่อถือของข้อมูลปริมาณน้ำฝนรายเดือน (Test of Consistency) โดยวิธี Double Mass Curve เนื่องจากเป็นวิธีซึ่งเป็นที่ยอมรับและนิยมใช้อย่างแพร่หลาย เริ่มจากนำข้อมูลน้ำฝนรายเดือนของสถานีวัดน้ำฝนเขื่อนบางลาง (71135) มาหาค่าสะสมและนำไปพล็อตกราฟร่วมกับข้อมูลน้ำฝนรายเดือนเฉลี่ย 2 สถานี (สถานีเมืองยะลาและสถานีเขื่อนปัตตานี) สะสมดังตัวอย่างในรูป 4.6 เมื่อพิจารณาแล้วพบว่ากราฟมีความชันเพียงความชันเดียวซึ่งหมายความว่าข้อมูลปริมาณน้ำฝนรายเดือนของสถานีวัดปริมาณน้ำฝนเขื่อนบางลาง (71135) มีความน่าเชื่อถือ สำหรับกราฟ Double Mass Curve ของสถานีที่เหลืออีก 2 สถานี แสดงในรูป 4.7 และรูป 4.8



รูป 4.6 Double Mass Curve ของสถานีวัดปริมาณน้ำฝนเขื่อนบางลาง



รูป 4.7 Double Mass Curve ของสถานีวัดปริมาณน้ำฝนเมืองยะลา

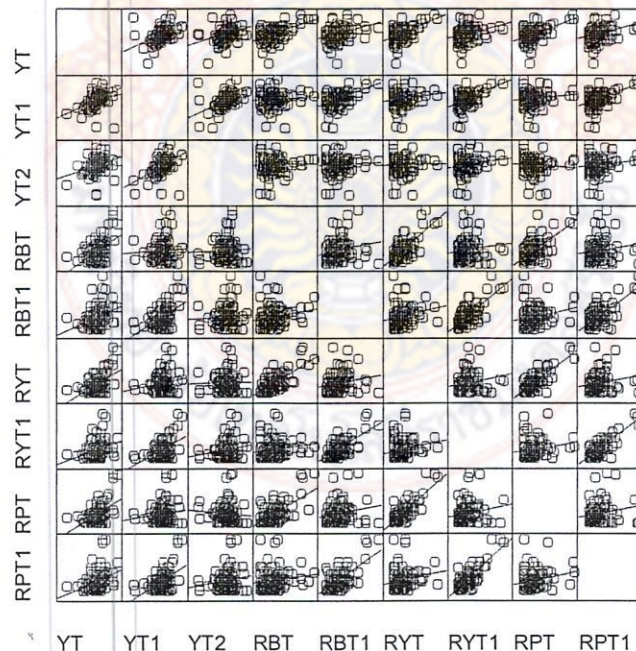


รูป 4.8 Double Mass Curve ของสถานีวัดปริมาณน้ำฝนเขื่อนปัตตานี

เมื่อข้อมูลปริมาณน้ำฝนผ่านการตรวจสอบแล้วพบว่ามีความสอดคล้องกับข้อมูลปริมาณน้ำท่าและมีความน่าเชื่อถือ นั้นหมายความว่าข้อมูลชุดดังกล่าวพร้อมที่จะนำไปวิเคราะห์ในขั้นตอนต่อไป

4.1.3 การเตรียมข้อมูล

การวิเคราะห์ระดับและทิศทางความสัมพันธ์ระหว่างตัวแปรตามและกลุ่มของตัวแปรอิสระสำหรับแบบจำลองการถดถอยเชิงพหุ (Multiple Linear Regression, MLR Model) ถือเป็นขั้นตอนสำคัญของการเตรียมข้อมูล เพราะถ้าหากข้อมูลที่นำมาใช้ใน MLR Model ไม่สามารถระบุระดับและทิศทางความสัมพันธ์ได้ชัดเจน ข้อมูลดังกล่าวก็จัดได้ว่าไม่มีความเหมาะสม อาจต้องดำเนินการจัดเตรียมข้อมูลชุดใหม่หรือเตรียมวิธีการแปลงค่าข้อมูลเพื่อให้สามารถวิเคราะห์ผลได้ วิธีการวัดระดับและทิศทางความสัมพันธ์ระหว่างตัวแปรต่างๆ ที่ทำได้ง่ายและมีประสิทธิภาพ คือ Scatter Plots Matrix ดังรูป 4.9 ซึ่งเป็นกราฟของข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่า ณ ปัจจุบัน (Box Cox Transformed) แทนด้วย Y_t (เมื่อ $t=1, \dots, N$) นำมาพล็อตกราฟ Scatter Plots Matrix กับกลุ่มของตัวแปรอิสระซึ่งประกอบด้วย ข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่าย้อนเวลา 1 ช่วงเวลา Y_{t-1} และ 2 ช่วงเวลา Y_{t-2} และข้อมูลปริมาณน้ำฝนรายเดือนของสถานีวัดน้ำเขื่อนบาง ๓ ช่วงเวลาต่างๆ ซึ่งจะแทนด้วย Rb_{t-k} (เมื่อ $t=1, \dots, N$ และ $k=0, 1, \dots$) ข้อมูลปริมาณน้ำฝนรายเดือนของสถานีเมืองยะลา ซึ่งจะแทนด้วย Ry_{t-k} และข้อมูลปริมาณน้ำฝนรายเดือนของสถานีเขื่อนปัตตานี ซึ่งจะแทนด้วย Rp_{t-k}



รูป 4.9 กราฟ Scatter Plots Matrix ระหว่างข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่าของสถานี X.79 และข้อมูลปริมาณน้ำฝนรายเดือนของ 3 สถานี

อย่างไรก็ตามการพิจารณาจากกราฟ Scatter Plots Matrix ควบคู่กับเมตริกซ์ของค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Matrix) ยิ่งช่วยเพิ่มระดับความเชื่อมั่นในความสัมพันธ์ของตัวแปรตามและกลุ่มของตัวแปรอิสระได้ดียิ่งขึ้น ดังจะเห็นได้จากรูป 4.10 ว่าระดับความสัมพันธ์ระหว่างข้อมูลปริมาณน้ำฝนรายเดือนของ 3 สถานี จากค่าสัมประสิทธิ์สหสัมพันธ์มีค่าอยู่ระหว่าง 0.035 ถึง 0.506 ตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (ปริมาณน้ำท่ารายเดือนสถานีเขื่อนปัตตานี ณ เวลาปัจจุบัน) สูงสุด 3 ลำดับแรกคือ ข้อมูลปริมาณน้ำฝนรายเดือนของสถานีเมืองยะลา ณ ปัจจุบัน Ry_t ด้วยระดับความสัมพันธ์ 0.506 ข้อมูลปริมาณน้ำฝนรายเดือนของสถานีเมืองเขื่อนบางลาง ณ ปัจจุบัน Rb_t ด้วยระดับความสัมพันธ์ 0.498 และข้อมูลปริมาณน้ำฝนรายเดือนของสถานีเขื่อนปัตตานี ณ ปัจจุบัน Rp_t ด้วยระดับความสัมพันธ์ 0.485 ตามลำดับ

	Y_t	Y_{t-1}	Y_{t-2}	Rb_t	Rb_{t-1}	Ry_t	Ry_{t-1}	Rp_t	Rp_{t-1}
Y_t	1.000	0.475	0.326	0.498	0.347	0.506	0.305	0.485	0.336
Y_{t-1}		1.000	0.525	0.054	0.477	0.161	0.487	0.126	0.488
Y_{t-2}			1.000	0.030	0.061	0.007	0.168	-0.084	0.146
Rb_t				1.000	0.184	0.729	0.150	0.662	0.152
Rb_{t-1}					1.000	0.207	0.719	0.203	0.689
Ry_t						1.000	0.206	0.813	0.190
Ry_{t-1}							1.000	0.191	0.824
Rp_t								1.000	0.196
Rp_{t-1}									1.000

รูป 4.10 Correlation Matrix ระหว่างข้อมูลปริมาณน้ำท่ารายเดือนแปลงค่าของสถานี X.79 และข้อมูลปริมาณน้ำฝนรายเดือนของ 3 สถานี

เพื่อให้แน่ใจว่ารูปแบบของ MLR Model มีความสอดคล้องเหมาะสมกับลักษณะของข้อมูล ในกรณีที่ไม่สามารถหารูปแบบของ MLR Model จากข้อมูลดิบได้ ก็ต้องเตรียมการแปลงค่าข้อมูลด้วยวิธีการที่เชื่อถือได้ ซึ่งจากผลการวิเคราะห์คุณสมบัติทางสถิติเบื้องต้นทั้งจากข้อมูลปริมาณน้ำท่ารายเดือนและข้อมูลปริมาณน้ำฝนรายเดือน จำเป็นต้องแปลงค่าข้อมูลปริมาณน้ำท่ารายเดือนซึ่งจะใช้แทนตัวแปรตามด้วยวิธี Box-Cox Transformation สำหรับข้อมูลปริมาณน้ำฝนรายเดือนไม่จำเป็นต้องทำการแปลงค่าแต่อย่างใด ข้อมูลเหล่านี้พร้อมที่จะนำไปประมาณค่าพารามิเตอร์ของ MLR Model เพื่อหารูปแบบที่เหมาะสมที่สุดไปใช้งาน

4.2 การประมาณค่าพารามิเตอร์ (Parameter Estimation)

การประมาณค่าพารามิเตอร์ อาทิ ค่ารายเดือนเฉลี่ย \bar{x}_T ส่วนเบี่ยงเบนมาตรฐานรายเดือนเฉลี่ย S_T และค่าสัมประสิทธิ์ความแปรปรายเดือนเฉลี่ย Cs_T ของข้อมูลปริมาณน้ำท่าและข้อมูลปริมาณน้ำฝนรายเดือนของแต่ละสถานีซึ่งมีจำนวนอย่างละ 132 ตัวอย่างด้วยวิธีโมเมนต์ เพราะ Yevjevich (1975) ซึ่งทำการศึกษาการสังเคราะห์ข้อมูล Monthly Net Basin Supply (NBS) ของทะเลสาบ Michigan-Huron ซึ่งใช้ข้อมูลจำนวน 520 ตัวอย่าง พบว่าหากตัวอย่างมีขนาดใหญ่ ($N>30$) การประมาณค่าพารามิเตอร์ด้วยวิธีโมเมนต์ (Moment Method) หรือวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood) ให้ผลลัพธ์ไม่ต่างกัน

สำหรับการประมาณค่าพารามิเตอร์ของ MLR Model เลือกใช้วิธีกำลังสองน้อยที่สุด (Least Square Method) เนื่องจากเป็นวิธีที่ทำให้ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์มีค่าต่ำสุด โดยได้แบ่งชุดข้อมูลที่จะนำมาใช้ประมาณค่าพารามิเตอร์ออกเป็น 3 ชุด ประกอบด้วย ชุดที่หนึ่งเป็นช่วงข้อมูลของปี ค.ศ.1997-2001 ชุดที่สองเป็นช่วงข้อมูลของปี ค.ศ.2001-2007 และชุดที่สามเป็นช่วงข้อมูลของปี ค.ศ.1997-2007 เพื่อหาช่วงข้อมูลช่วงใดที่ให้ค่าพารามิเตอร์ของ MLR Model มีประสิทธิภาพมากที่สุด ตัวแปรตามคือ ข้อมูลปริมาณน้ำท่ารายเดือนของสถานีเขื่อนปัตตานีแปลงค่า ณ เวลาปัจจุบันซึ่งแทนด้วย Y_t และกลุ่มของตัวแปรอิสระของข้อมูลทั้ง 3 ชุดที่นำมาประมาณค่าพารามิเตอร์ประกอบด้วย ข้อมูลปริมาณน้ำท่ารายเดือนของสถานีเขื่อนปัตตานีแปลงค่า ณ ช่วงเวลาก่อนหน้า 2 ช่วงเวลา คือ Y_{t-1} และ Y_{t-2} ข้อมูลปริมาณน้ำฝนของสถานีเขื่อนบางลาง ณ เวลาปัจจุบัน แทนด้วย Rb_t และช่วงเวลาก่อนหน้า 1 ช่วงเวลา Rb_{t-1} ข้อมูลปริมาณน้ำฝนของสถานีเมืองยะลา ณ เวลาปัจจุบัน แทนด้วย Ry_t และช่วงเวลาก่อนหน้า 1 ช่วงเวลา Ry_{t-1} และข้อมูลปริมาณน้ำฝนของสถานีเขื่อนปัตตานี ณ เวลาปัจจุบัน แทนด้วย Rp_t และช่วงเวลาก่อนหน้า 1 ช่วงเวลา Rp_{t-1} รวมจำนวนตัวแปรอิสระทั้งสิ้น 8 ตัวแปร

การคัดเลือกตัวแปรอิสระเข้าในสมการถดถอยนั้น ใช้เทคนิคการคัดเลือกตัวแปรโดยวิธี Stepwise ต่างจากวิธีของ Sinnakaudan et al. (2006) ซึ่งใช้แบบจำลอง Multiple Linear Regression ในการพยากรณ์ปริมาณตะกอนที่ท้องน้ำของแม่น้ำต่างๆ ในประเทศมาเลเซีย เนื่องจาก การศึกษานี้ใช้ข้อมูลปริมาณน้ำท่าและข้อมูลปริมาณน้ำฝนรายเดือนซึ่งเป็นข้อมูลอนุกรมเวลามีคุณลักษณะแตกต่างจากข้อมูลการเคลื่อนย้าย (Transport) การตกตะกอน (Sediment) รูปร่าง (Shape) และการต้านทานการไหล (Flow Resistance) ที่ใช้ในการศึกษาของ Sinnakaudan et al. (2006)

วิธี Stepwise ทำงานโดยคัดเลือกตัวแปรอิสระเข้าในสมการถดถอยครั้งละ 1 ตัวแปร โดยเริ่มจากตัวแปรอิสระที่มีสหสัมพันธ์กับตัวแปรตามมากที่สุดไปน้อยสุด ในแต่ละครั้งจะทำการประมาณค่าพารามิเตอร์ β_j (เมื่อ $j=0, \dots, m$) ด้วยวิธีกำลังสองน้อยที่สุด โดยจะใช้ b_j แทนค่าประมาณของค่าพารามิเตอร์เหล่านั้น และทำการตรวจสอบนัยสำคัญของตัวแปรอิสระจากค่าความแปรปรวนของค่าประมาณของค่าพารามิเตอร์ $\text{Var}(b_j)$ โดยการทดสอบ t-Test หรือ F-Test และวัดจากค่า R^2 หากตัวแปรอิสระตัวไหนตรวจสอบแล้วว่าไม่มีนัยสำคัญกับตัวแปรตามหรือเป็นตัวแปรที่ไม่ได้ทำให้ค่า R^2 ของสมการถดถอยเพิ่มก็จะถูกคัดออกจากสมการ

ผลปรากฏว่าเมื่อใช้ข้อมูลชุดที่ 1 ซึ่งช่วงเวลาของข้อมูลคือ ปี ค.ศ.1997-2001 ประมาณค่าพารามิเตอร์ได้รูปแบบของ MLR Model ที่เหมาะสม ดังสมการ 4.1

$$Y_t = 1.618 + 0.325Y_{t-1} + 0.145Y_{t-2} + 0.002449Rb_t + 0.001719Rp_t \quad \text{-----}(4.1)$$

เช่นเดียวกันเมื่อใช้ข้อมูลชุดที่ 2 ซึ่งช่วงเวลาของข้อมูลคือ ปี ค.ศ.2002-2007 ประมาณค่าพารามิเตอร์ได้รูปแบบของ MLR Model ที่เหมาะสม ดังสมการ 4.2

$$Y_t = 2.035 + 0.395Y_{t-1} + 0.003863Rb_t \quad \text{-----}(4.2)$$

เมื่อใช้ข้อมูลชุดที่ 3 ซึ่งช่วงเวลาของข้อมูลคือ ปี ค.ศ.1997-2007 ประมาณค่าพารามิเตอร์ได้รูปแบบของ MLR Model ที่เหมาะสม ดังสมการ 4.3

$$Y_t = 1.883 + 0.409Y_{t-1} + 0.002924Rb_t + 0.001357Rb_{t-1} \quad \text{-----}(4.3)$$

เมื่อพิจารณารูปแบบของ MLR Model จากสมการ 4.1 ถึง 4.3 และตาราง 4.3 สังเกตได้ว่าการประมาณค่าพารามิเตอร์ของชุดข้อมูลทั้ง 3 ชุด ถึงแม้ว่าจะใช้จำนวนตัวแปรอิสระเท่ากันทุกชุด แต่การใช้วิธี Stepwise เลือکتัวแปรอิสระเข้าในสมการถดถอยนั้นเป็นวิธีการที่ช่วยขจัดปัญหาตัวแปรอิสระมีความสัมพันธ์กันเอง (Collinearity) ซึ่งหากใช้วิธีการอื่นจะได้แบบจำลองที่มีตัวแปรอิสระซ้ำซ้อน ดังนั้นรูปแบบของ MLR Model ของข้อมูลแต่ละชุดจึงมีจำนวนตัวแปรอิสระแตกต่างกันไป

ตาราง 4.3 ค่าประมาณของพารามิเตอร์ของ MLR Model ของช่วงข้อมูล 3 ชุด

Variables	Estimated Parameters of MLR Model		
	Series 1 (Year 1997-2001)	Series 2 (Year 2002-2007)	Series 3 (Year 1997-2007)
Constant	1.61800	2.03500	1.88300
Y_{t-1}	0.32500	0.39500	0.40900
Y_{t-2}	0.14500	-	-
Rb_t	0.00245	0.00386	0.002924
Rb_{t-1}	-	-	-
Ry_t	-	-	0.001357
Ry_{t-1}	-	-	-
Rp_t	-0.00172	-	-
Rp_{t-1}	-	-	-

4.3 การตรวจสอบความเหมาะสมของแบบจำลอง (Diagnostic Checking)

การตรวจสอบความเหมาะสมของแบบจำลองแบ่งออกเป็น 3 ส่วน คือ การทดสอบความเป็นอิสระของค่าคลาดเคลื่อน (Test of Independence) การตรวจสอบความเป็นปกติของค่าคลาดเคลื่อน (Test of Normality) และการวัดประสิทธิภาพของแบบจำลองจากค่าสถิติต่างๆ (Goodness-of-Fit Statistics) ประกอบด้วย รากที่สองของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสอง (Root Mean Square Error) ค่าเฉลี่ยของค่าความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Error, MSE) สัมประสิทธิ์วัดประสิทธิภาพ (Coefficient of Efficiency) ค่า Relative Error และค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination, R^2)

การศึกษานี้ได้ตรวจสอบความเป็นอิสระของค่าคลาดเคลื่อนโดยการพิจารณาจากค่าอัตโนมัติสัมพันธ์ของค่าคลาดเคลื่อน $r_k(\epsilon_t)$ และวิธี Durbin & Watson Test สำหรับการทดสอบความเป็นปกติของค่าคลาดเคลื่อนใช้วิธี Kolmogorov-Smirnov และวัดประสิทธิภาพของ MLR Model โดยใช้ค่าสถิติต่างๆ ดังที่กล่าวมาแล้วเป็นเกณฑ์

จากตาราง 4.4 เชื่อได้ว่า MLR Model 1 ถึง MLR Model 3 มีค่าคลาดเคลื่อนที่เป็นอิสระต่อกัน เนื่องจากค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation) มีค่าเท่ากับ 0 และ -0.004 ซึ่งหมายความว่าค่าคลาดเคลื่อนส่วนใหญ่ไม่มีความสัมพันธ์กันเลย และค่า Durbin-Watson ของแต่ละแบบจำลองมีค่าเท่ากับ 1.933 2.286 และ 2.056 ตามลำดับ ซึ่งอยู่ในเกณฑ์คือ ประมาณ 2.0

ตาราง 4.4 ผลการทดสอบความเป็นอิสระและความเป็นปกติของค่าคลาดเคลื่อนของ MLR Model

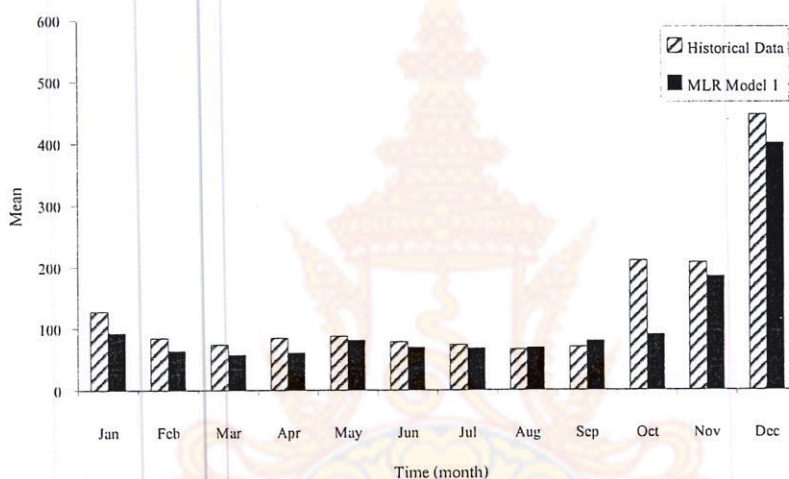
Model	Correlation	Test of Independence				Test of Normality				
		sig.	nsig.	Durbin-Watson	sig.	nsig.	Kolmogorov-Smirnov		sig.	nsig.
							D_N	$D_{\alpha,N}$		
MLR Model 1	0.000	✓	-	1.933	✓	-	0.931	0.351	✓	-
MLR Model 2	0.000	✓	-	2.286	✓	-	0.892	0.403	✓	-
MLR Model 3	-0.004	✓	-	2.056	✓	-	1.026	0.243	✓	-

เมื่อพิจารณาค่าสถิติวัดประสิทธิภาพของ MLR Model ในตาราง 4.5 พบว่าค่าคลาดเคลื่อน MLR Model 1 มีค่า RMSE เท่ากับ 0.765 ค่า MAE เท่ากับ 0.524 ค่า R^2 เท่ากับ 0.472 ค่าเฉลี่ยของ Relative Error เท่ากับ 0.06 และส่วนเบี่ยงเบนมาตรฐานของ Relative Error เท่ากับ 1.04 ซึ่งน้อยที่สุดเมื่อเทียบกับรูปแบบอื่นๆ ดังนั้นจึงสามารถสรุปได้ว่าการใช้ช่วงข้อมูลปี ค.ศ. 1997-2001 มีศักยภาพในประมาณค่าพารามิเตอร์ทำให้ MLR Model 1 เป็นรูปแบบที่เหมาะสมที่สุดสำหรับพยากรณ์ข้อมูลปริมาณน้ำท่ารายเดือนให้กับสถานีเขื่อนปัตตานีต่อไป

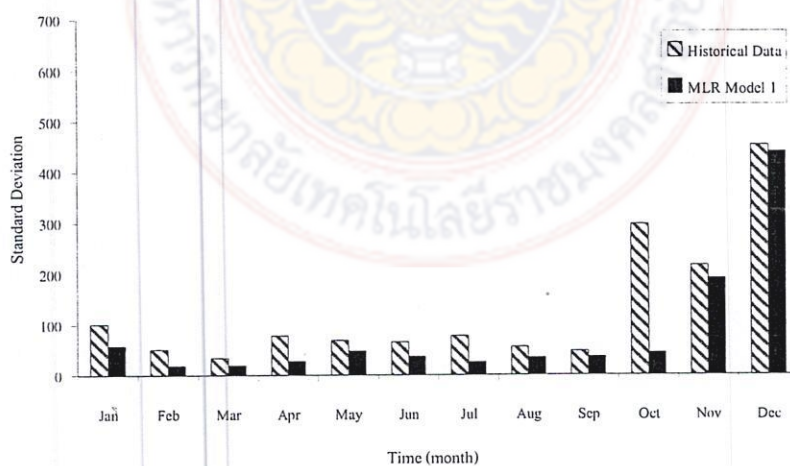
ตาราง 4.5 ค่าสถิติวัดประสิทธิภาพของ MLR Model

Model	RMSE	MAE	R^2	Relative Error	
				Mean	S.D
MLR Model 1	0.765	0.524	0.472	0.06	1.04
MLR Model 2	0.796	0.569	0.342	0.06	1.14
MLR Model 3	0.782	0.540	0.467	0.06	1.09

กราฟในรูป 4.11 และ 4.12 ซึ่งเป็นค่า Average Monthly Mean และค่า Average Monthly Standard Deviation ของข้อมูลปริมาณน้ำท่าจริงและจาก MLR Model 1 ตามลำดับ เห็นได้ชัดว่าค่า Average Monthly Mean จาก MLR Model 1 ส่วนใหญ่มีค่าต่ำกว่าค่าของข้อมูลจริงเล็กน้อย ยกเว้นค่าในเดือนตุลาคมมีค่าต่ำกว่าอย่างเห็นได้ชัด โดยประมาณ 50 เปอร์เซ็นต์ ส่วนค่า Average Standard Deviation จาก MLR Model 1 โดยภาพรวมมีแนวโน้มเช่นเดียวกับกราฟ Average Monthly Mean ในรูป 4.11 แต่เปอร์เซ็นต์ความแตกต่างจะมีมากกว่าโดยเฉพาะอย่างยิ่งในเดือนตุลาคม สำหรับรายละเอียดของข้อมูลแสดงในตาราง 4.6



รูป 4.11 กราฟ Average Monthly Mean ของข้อมูลปริมาณน้ำท่าจริง และจาก MLR Model 1



รูป 4.12 กราฟ Average Monthly Standard Deviation ของข้อมูลปริมาณน้ำท่าจริง และจาก MLR Model 1

ตาราง 4.6 ข้อมูลปริมาณน้ำท่ารายเดือนที่ได้จาก MLR Model 1

year	Discharge (million cubic meter/month)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
1997	6.82	41.31	44.71	48.42	31.76	58.65	44.51	45.77	69.64	48.70	78.65	312.13	
1998	88.03	55.89	41.61	21.60	25.20	35.04	36.10	62.56	79.85	90.48	109.85	213.03	
1999	124.12	62.45	79.67	122.82	88.88	63.88	66.94	84.69	93.39	123.03	234.25	1146.22	
2000	220.30	93.85	87.90	98.84	96.82	89.45	76.17	87.08	97.61	55.65	734.36	295.73	
2001	77.72	82.56	79.14	43.36	37.53	24.31	47.39	38.76	62.27	136.97	108.72	205.78	
2002	68.11	88.76	51.15	55.64	63.28	49.47	34.72	19.83	18.64	18.54	108.68	169.38	
2003	46.05	42.64	24.08	39.39	56.85	50.80	90.91	39.87	45.94	36.85	40.05	68.00	
2004	60.23	35.85	40.18	51.50	57.73	30.55	53.48	49.16	78.01	94.16	138.24	90.32	
2005	72.24	64.78	64.40	58.09	178.68	105.60	81.66	109.95	55.14	106.75	238.25	1372.49	
2006	94.24	74.91	51.87	69.49	119.05	145.05	100.38	81.19	123.95	111.48	96.40	102.64	
2007	162.83	59.79	72.97	62.26	138.10	104.73	111.19	137.82	152.58	166.87	145.75	436.89	Average
Mean	92.79	63.89	57.97	61.04	81.26	68.87	67.59	68.79	79.73	89.95	184.84	401.15	109.82
S.D	58.36	19.54	20.18	28.18	48.40	37.80	26.46	35.06	36.95	45.59	191.90	440.79	82.44
% Error of Mean	27.66	24.50	21.76	27.97	7.40	11.98	7.70	-4.25	-14.07	57.35	11.08	10.26	15.78
% Error of S.D	42.62	61.94	42.12	64.12	30.31	43.38	66.38	37.87	23.75	84.70	11.66	2.88	42.64

บทที่ 5

สรุปและเสนอแนะ

เนื้อหาในบทนี้แบ่งการนำเสนอออกเป็น 2 ส่วน คือ ส่วนแรกเป็นการสรุปผลการวิจัย และส่วนที่สองเป็นข้อเสนอแนะสำหรับการศึกษาในอนาคต โดยจะกล่าวรายละเอียดดังต่อไปนี้

5.1 สรุปผลการวิจัย

วัตถุประสงค์ของการศึกษานี้ คือ ประยุกต์ใช้วิธีการทางสถิติในการสร้างแบบจำลอง Multiple Regression Time Series Model, MLR Model โดยอาศัยความสัมพันธ์ของข้อมูลปริมาณน้ำฝนและข้อมูลปริมาณน้ำท่าในการหาค่าพารามิเตอร์และหาตัวแปรอิสระของ MLR Model ให้กับกลุ่มน้ำปัตตานี ผลจากการศึกษาสามารถสรุปดังต่อไปนี้

5.1.1 รูปแบบ MLR Model ที่มีประสิทธิภาพในการพยากรณ์ข้อมูลน้ำท่ารายเดือนของสถานีเขื่อนปัตตานี คือ $Y_t = 1.618 + 0.325Y_{t-1} + 0.145Y_{t-2} + 0.002449Rb_t + 0.001719Rp_t$ มีค่า R^2 เท่ากับ 0.472 ค่า RMSE เท่ากับ 0.765 ค่า MAE เท่ากับ 0.524 ค่าเฉลี่ยของ Relative Error เท่ากับ 0.06 และค่าเบี่ยงเบนมาตรฐานของ Relative Error เท่ากับ 1.04

5.1.2 การแปลงค่าข้อมูลอนุกรมเวลารายเดือนที่มีลักษณะของฤดูกาลแต่ไม่มีแนวโน้ม ด้วยวิธี Box-Cox Transformation ช่วยหาค่าพารามิเตอร์ที่ประมาณได้มีศักยภาพยิ่งขึ้น

5.1.3 การใช้ค่า R^2 ค่า RMSE ค่า MAE ค่าเฉลี่ยของ Relative Error และค่าเบี่ยงเบนมาตรฐานของ Relative Error เป็นเกณฑ์วัดประสิทธิภาพของ MLR Model ในกรณีข้อมูลรายเดือนผลลัพธ์มีทิศทางใกล้เคียงกัน

5.1.4 การใช้ค่า PACF ช่วยหาจำนวนการเลื่อนเวลาของตัวแปรอนุกรมเวลารายเดือน เป็นวิธีที่ได้ผล เนื่องจากตัวแปรอิสระซึ่งได้จากการเลื่อนเวลาของการศึกษานี้ที่อยู่นอกช่วงความเชื่อมั่นหรือผ่านเกณฑ์ของค่า PACF ส่วนใหญ่กลายเป็นตัวแปรที่เหมาะสมของ MLR Model

5.2 ข้อเสนอแนะ

จากผลการศึกษาในครั้งนี้ ทำให้ทราบถึงลักษณะและข้อจำกัดต่างๆ ของการจำลองข้อมูลอนุกรมเวลา ผู้ศึกษามีข้อเสนอแนะสำหรับการศึกษาในอนาคต เพื่อจะหารูปแบบของ MLR Model ที่มีประสิทธิภาพดีกว่าที่ได้สรุปในการศึกษานี้

5.2.1 เพื่อให้รูปแบบของ MLR Model ที่วิเคราะห์ได้ เป็นรูปแบบที่ดีที่สุดโดยปราศจากข้อโต้แย้งใดๆ ควรนำวิธีการแปลงค่าข้อมูลน้ำท่าวิธีอื่นๆ มาใช้ อาทิ วิธี Wilson-Hilferty Transformations

5.2.2 เพื่อเป็นการเพิ่มทางเลือก ควรใช้ Periodic MLR Model หรือรูปแบบอื่นในตระกูล ARMA Model ในการสร้างข้อมูลน้ำท่าสังเคราะห์ให้กับสถานีตัวอย่าง

5.2.3 เพื่อให้การนำข้อมูลน้ำฝนมาใช้เพื่ออธิบายความสัมพันธ์ที่มีต่อข้อมูลน้ำท่าได้ผลดียิ่งขึ้น การหารูปแบบของ MLR Model ควรแบ่งออกเป็น 2 รูปแบบ เป็นอย่างน้อย คือ MLR Model ที่วิเคราะห์จากข้อมูลน้ำท่าร่วมกับข้อมูลน้ำฝนในช่วงเวลาน้ำหลาก และ MLR Model สำหรับช่วงเวลาปกติ เพราะอาจช่วยทำให้ความแปรปรวนของค่าคลาดเคลื่อนจากแบบจำลองลดลง

6.2.5 ในกรณีที่มีข้อมูลน้ำฝนเพียงพอ ควรนำข้อมูลน้ำฝนจากสถานีอุตุวิทยามหาวิทยาลัยราชภัฏนครราชสีมา มาทำการเฉลี่ยตามวิธีการทางอุทกวิทยา ก่อน แล้วจึงนำข้อมูลน้ำฝนที่เฉลี่ยค่าแล้วมาวิเคราะห์หารูปแบบของ MLR Model ร่วมกับข้อมูลน้ำฝน โดยแบ่งเป็น 2 แบบ ดังรายละเอียดในข้อ 5.2.3 คือ MLR Model สำหรับช่วงน้ำหลาก และ MLR Model สำหรับช่วงเวลาปกติ

เอกสารอ้างอิง

- ทรงศิริ แต่สมบัติ. (2539). *เทคนิคการพยากรณ์เชิงปริมาณ*. กรุงเทพฯ : หจก.สำนักพิมพ์ฟิสิกส์
เซ็นเตอร์
- ทรงศิริ แต่สมบัติ. (2542). *การวิเคราะห์การถดถอย*. (พิมพ์ครั้งที่ 2). กรุงเทพฯ : สำนักพิมพ์
มหาวิทยาลัยเกษตรศาสตร์
- มนตรี พิริยะกุล. (2525). *เทคนิคการวิเคราะห์สัมพรรคถดถอย*. กรุงเทพฯ : โรงพิมพ์มหาวิทยาลัย
รามคำแหง
- Box, G.E.P., and Jenkins, G.M. (1970). *Time Series Analysis Forecasting and Control*. San
Francisco Calif : Holden-Day Inc.
- Carlson, R.F., MacCormick, A.J.A., and Watts, D.G. (1970) "Application of linear random models
to four annual streamflow series", *Water Resource and Research*, vol 6, no.4
,1070-1078
- Hipel, K.W., McLeod, A.I., and Lennox, W.C. (1977) "Advances in Box Jenkins Modeling 1,
Model Construction", *Water Resource and Research*, vol 13, no. 3, 567-575
- Holder, R.L. (1985) *Multiple Regression in Hydrology*, Wallingford : Institute of Hydrology
- Kottekoda, N.T. (1980) *Stochastic Water Resource Technology*, Hong Kong : The Macmillan
Press Ltd.
- McLeod, A.I., Hipel, K.W. and Lennox, W.C. (1977) "Advances in Box Jenkins Modelling 2,
Application", *Water Resources and Research*, vol 13, no. 3, 577-586

- Miller, R.B., Bell, W. and Ferreiro, O. and Wang, R.Y. (1981) "Modelling Daily River Flows With Precipitation Input, Water Resource and Research, vol 17, no. 1, 209-215
- Neter, J., Kutner, M.H., Nachtsheim, C.J. and Wasserman, W. (1996). *Applied Linear Regression Models*. Chicago : R. R. Donnelley & Sons Company
- Pankrats, A. (1983). *Forecasting with Univariate Box-Jenkins Models*. New York : John Wiley & Sons, Inc.
- Saadat Ayub Khan. (2006) "Rainfall-Runoff Modelling Using Data Driven and Daily River and Statistical Methods", IEEE
- Salas, J.D., Delleur, J.W., Yevjevich, V. And Lane, W.L. (1980). *Applied Modeling of Hydrologic Time Series*. Michigan : Water Resource Publications
- Saw Suat Kheng. (1983). Modeling and Simulation of Hourly Wind Speed and Direction in Malaysia. Master's Thesis. Asian Institute of Technology, Bangkok
- Seckler, J.G. El-Shaarawi, A.H. and Esterby, S.R. (1982) *Time Series Method in Hydrosiences*, Netherland : Elsevier Scientific
- Sinnakaudan, S.K. Ab Ghani, A. Ahmad, M.S.S. and Zakaria, N.A. (2006) "Multiple Linear Regression for Total Bed Material Load Prediction", *Journal of Hydraulic Engineering*, vol 132, no. 5, 521-528
- Quimpo, R.G. (1967) "Stochastic Model of Daily River Flow Sequences", *Hydrology Paper*, Colorado State University, no. 18, February

Yevjevich, V. (1975). "Generation of Hydrologic Samples : Case Study of The Great Lakes",
Hydrology Paper, Colorado State University, no.72

Yevjevich, V. (1984). *Structure of Daily Hydrologic Series*. Michigan : Water Resources
Publications



ภาคผนวก



ตาราง ผ-1 ข้อมูลปริมาณน้ำท่ารายเดือนสถานีวัดน้ำเขื่อนปัตตานี (X.79) ปี ค.ศ.1997-2007

Year	Discharge (million cubic meters)											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	70.93	73.44	28.78	26.02	37.42	29.39	14.01	32.93	62.05	72.76	52.84	390.29
1998	38.81	19.11	21.09	17.81	10.12	24.11	24.25	56.17	54.19	64.90	107.83	213.40
1999	152.93	124.72	90.90	119.24	167.62	46.92	60.22	97.68	85.01	143.42	337.98	1502.78
2000	264.31	65.06	62.28	290.60	189.23	105.02	64.86	64.86	64.38	123.67	800.97	172.54
2001	299.74	90.05	52.13	26.34	16.45	19.56	40.37	30.47	75.03	154.72	265.15	216.56
2002	68.76	54.63	143.90	54.68	69.00	19.56	3.62	7.86	0.76	1097.01	74.90	66.59
2003	27.76	15.22	66.13	47.39	54.09	64.48	28.39	27.07	12.32	25.76	65.90	274.93
2004	34.09	47.08	71.08	37.04	18.76	40.26	43.80	46.14	41.72	124.20	106.65	174.39
2005	38.58	146.02	100.95	130.83	87.62	130.03	91.57	62.23	68.43	166.17	130.83	929.07
2006	191.24	124.90	100.88	93.78	188.65	167.33	166.18	83.52	165.85	178.47	104.00	117.72
2007	223.82	170.64	76.90	88.36	126.38	213.97	268.24	216.89	139.14	168.97	239.45	859.00

ตาราง ผ-2 ข้อมูลปริมาณน้ำฝนรายเดือนที่สถานีวัดน้ำเขื่อนบางลาง ปี ค.ศ.1997-2007

Monthly Rainfall of Bang Lang Dam Station (millimeters)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	66.9	0.0	42.5	164.7	39.6	177.4	95.1	210.2	342.0	85.0	146.5	501.1
1998	97.8	119.7	174.3	14.5	76.7	164.7	111.5	300.5	236.1	197.3	223.0	342.1
1999	137.9	30.8	91.2	248.2	135.5	37.3	140.9	250.0	177.2	217.0	286.4	503.3
2000	214.3	20.6	240.2	248.2	172.8	62.2	145.5	149.5	258.1	95.1	531.2	172.8
2001	24.3	0.0	36.6	74.3	52.1	60.7	189.7	91.1	203.2	262.8	183.8	159.5
2002	7.6	219.0	158.4	45.2	126.1	68.3	36.9	145.2	100.0	305.3	126.0	194.2
2003	87.2	179.2	69.8	108.8	149.2	32.4	297.6	108.9	231.8	23.3	50.0	94.4
2004	7.3	0.0	69.2	87.2	179.2	69.8	199.0	134.0	229.5	264.7	266.0	163.0
2005	42.6	187.0	97.2	87.4	164.5	230.8	127.8	236.4	94.2	225.9	403.9	676.9
2006	0.0	0.0	0.0	120.4	193.5	297.6	125.6	63.3	259.3	195.6	102.0	195.0
2007	410.4	0.0	55.4	111.1	298.5	137.4	136.2	147.7	274.3	239.3	277.7	608.6

ตาราง ผ-3 ข้อมูลปริมาณน้ำฝนรายเดือนสถานีวัดน้ำเมืองยะลา ปี ค.ศ.1997-2007

Monthly Rainfall of Muang Yala Station (millimeters)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	70.2	0.0	0.0	68.8	37.4	196.7	78.6	151.1	348.8	94.0	120.8	556.4
1998	91.5	269.5	21.1	0.0	25.5	168.5	300.5	186.0	226.2	228.5	187.8	495.0
1999	44.5	18.0	198.0	81.0	160.7	85.5	25.5	138.7	107.4	200.5	408.5	791.1
2000	245.0	15.0	183.2	81.0	80.5	221.0	119.0	189.9	197.0	206.0	859.5	234.5
2001	1.7	0.0	36.5	10.0	0.0	55.5	105.1	110.6	258.3	398.1	169.8	515.1
2002	17.5	121.4	74.5	17.5	121.4	74.5	106.4	87.6	57.2	169.3	384.6	382.1
2003	46.0	53.5	49.2	17.5	121.4	74.5	106.4	87.6	57.2	169.3	284.6	382.1
2004	21.2	19.2	66.8	188.1	243.6	103.8	217.3	141.3	386.4	390.1	266.9	163.4
2005	86.9	219.7	84.6	4.1	285.1	102.0	226.9	213.6	166.3	349.9	225.8	635.3
2006	0.0	0.0	0.0	119.1	183.3	100.6	188.3	123.4	298.0	354.2	162.9	205.3
2007	198.2	5.7	116.9	90.3	214.7	292.1	209.5	77.7	127.4	333.0	111.3	526.3

ตาราง ผ-4 ข้อมูลปริมาณน้ำฝนรายเดือนสถานีวัดน้ำเขื่อนป่าสักชลสิทธิ์ ปี ค.ศ.1997-2007

Monthly Rainfall of Pattani Dam Station (millimeters)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	80.7	417.7	37.1	83.6	114.6	214.9	186.7	199.4	156.7	122.9	230.6	574.3
1998	61.5	33.7	112.9	0.4	25.1	212.6	189.3	165.5	239.9	303.8	349.0	453.4
1999	258.2	16.6	138.9	244.0	191.8	52.3	143.9	185.5	233.6	322.7	511.3	919.8
2000	17.7	0.0	60.7	244.0	51.9	114.6	49.9	262.9	215.2	122.0	878.7	451.9
2001	18.7	113.6	217.6	18.7	141.4	23.1	234.7	106.6	214.8	441.8	207.1	450.0
2002	16.0	102.9	8.9	55.4	116.2	93.0	150.5	95.5	119.6	200.4	306.3	361.1
2003	4.5	4.1	15.0	18.7	113.6	217.6	134.0	63.6	49.2	370.8	308.1	313.0
2004	17.2	0.0	82.8	96.4	120.3	89.6	144.2	107.4	223.1	297.0	320.8	156.7
2005	118.2	92.8	92.9	4.4	530.3	183.8	140.2	191.5	95.2	306.5	344.2	944.3
2006	0.0	0.0	0.0	74.9	315.8	156.4	151.0	127.8	225.3	182.7	159.8	159.7
2007	143.4	13.3	47.4	49.3	286.7	275.9	182.7	204.1	104.0	307.8	175.1	260.0