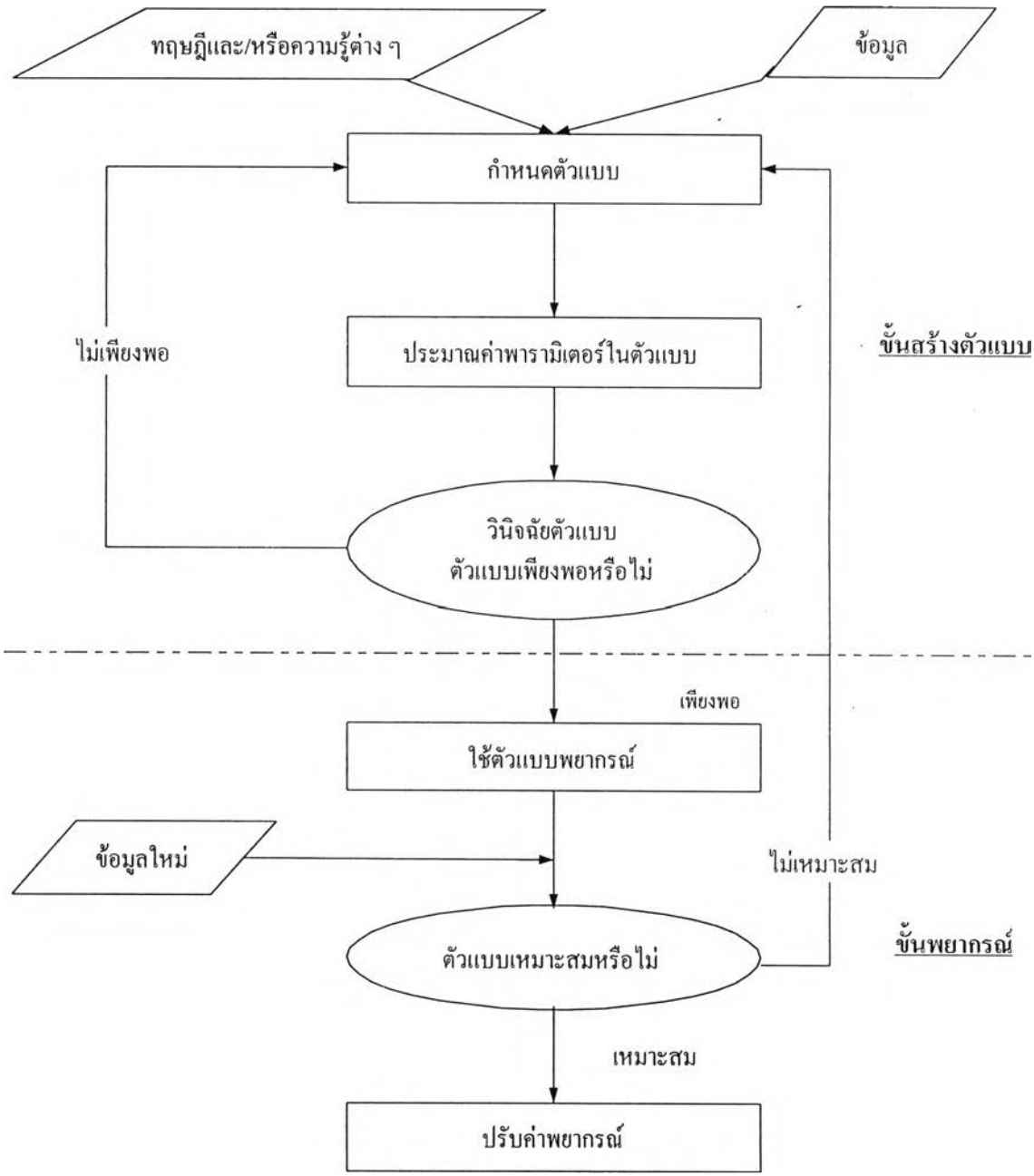
บทที่ 3
การวิเคราะห์ข้อมูล

ระบบงานพยากรณ์ของการวิจัยเรื่องการพยากรณ์ปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของกรมชลประทานแบ่งออกเป็น 2 ขั้นตอนใหญ่ ๆ คือขั้นตอนของการสร้างตัวแบบพยากรณ์และขั้นตอนของการพยากรณ์โดยมีรายละเอียดขั้นตอนดังแผนภูมิรูปภาพที่ 3.1

ในขั้นตอนของการสร้างตัวแบบจะเริ่มด้วยการกำหนดตัวแบบทดลองเป็นตัวแบบเริ่มต้นที่คาดว่าจะเป็นตัวแบบที่ใช้ได้โดยอาศัยความรู้และทฤษฎีต่าง ๆ ที่เกี่ยวข้อง รวมถึงการวิเคราะห์ข้อมูลใช้เทคนิคการพยากรณ์เชิงปริมาณประกอบไปด้วย การวิเคราะห์การถดถอย และการวิเคราะห์ข้อมูลอนุกรมเวลาเพื่อพยากรณ์ ใช้วิธีการวิเคราะห์ต่อไปนี้ วิธีการปรับให้เรียบ วิธีการแยกองค์ประกอบ และวิธีการของบ็อกซ์และเจนกินส์ จากนั้นทำการประมาณค่าพารามิเตอร์ในตัวแบบและทำการตรวจสอบความเหมาะสมในเชิงสถิติ โดยทำการตรวจสอบข้อสมมติหรือคุณสมบัติต่าง ๆ ในเชิงสถิติรวมทั้งรูปแบบความสัมพันธ์ของตัวแบบ โดยพิจารณาจากค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์ เป็นค่าที่ใช้พิจารณาว่าวิธีการใดให้ค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบรูณ์ต่ำสุด จะเป็นวิธีที่เหมาะสมกับข้อมูลอนุกรมเวลาชุดนั้น ๆ แต่ถ้าพบว่าตัวแบบที่กำหนดขึ้นในครั้งแรกนั้นยังไม่สอดคล้องกับข้อสมมติหรือขาดคุณสมบัติในเชิงสถิติ หรือยังมีรูปแบบที่ไม่เหมาะสมจะทำการปรับแก้ตัวแบบใหม่ และประมาณค่าพารามิเตอร์ของตัวแบบใหม่ตลอดจนตรวจสอบความเหมาะสมและรูปแบบของตัวแบบใหม่อีก กรรมวิธีนี้จะวนเวียนเช่นนี้จนกว่าจะพบว่าตัวแบบพยากรณ์ผ่านการทดสอบว่ามีความเหมาะสมในเชิงสถิติ เมื่อผ่านขั้นตอนนี้แล้วก็จะสามารถใช้ตัวแบบพยากรณ์ค่าที่ต้องการ ซึ่งจะนำไปสู่ขั้นตอนของการพยากรณ์

ในงานขั้นพยากรณ์โดยทั่วไปนักพยากรณ์จะพยากรณ์ค่าในอนาคตมากกว่า 1 คาบเวลา เช่น ในงานวิจัยนี้จะพยากรณ์ค่ารายเดือนใน 12 เดือนข้างหน้า ฉะนั้นเมื่อเวลาผ่านไปในแต่ละคาบเวลาอาจมีความเป็นไปได้ที่จะปรับแก้ค่าพยากรณ์ในคาบเวลาที่เหลือ โดยใช้ค่าของข้อมูลจริงที่เกิดขึ้นมาใช้ประโยชน์ในการปรับแก้ค่าพยากรณ์ที่เหลือใหม่ ทั้งนี้ก่อนที่จะปรับค่าพยากรณ์เหล่านั้น ควรตรวจสอบก่อนว่าตัวแบบพยากรณ์ที่ใช้อยู่นั้นยังมีความเหมาะสมหรือไม่ โดยใช้ข้อมูลใหม่ที่เกิดขึ้นร่วมด้วย ในการตรวจสอบซึ่งอาจกล่าวได้ว่างานส่วนนี้เป็นงานควบคุมตัวแบบพยากรณ์ ถ้าพบว่าตัวแบบไม่เหมาะสมควรจะปรับแก้ตัวแบบใหม่ โดยจะกลับเข้าสู่งานขั้นสร้างตัวแบบที่กล่าวมาในขั้นตอนที่ 1 ในส่วนของข้อมูลได้จากการเก็บรวบรวมของฝ่ายจัดสรรน้ำ

สำนักอุทกวิทยาและบริหารน้ำ กรมชลประทาน สำหรับหัวข้อ 3.1 เป็นส่วนที่เกี่ยวกับการวิเคราะห์ข้อมูล ซึ่งจะแสดงให้เห็นถึงขั้นตอนในการวิเคราะห์ข้อมูล



รูปภาพที่ 3.1 แผนภูมิแสดงโครงสร้างของระบบงานพยากรณ์

3.1 การวิเคราะห์ข้อมูล

ข้อมูลอนุกรมเวลาที่ใช้ในการวิจัยประกอบด้วยข้อมูลปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของ 24 เขื่อน ซึ่งแต่ละเขื่อนจะมีข้อมูลที่แตกต่างกันไปตามระยะเวลาที่เริ่มใช้งานของแต่ละอ่างเก็บน้ำของเขื่อนต่าง ๆ โดยรวบรวมข้อมูลปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่เป็นรายเดือนระหว่างปี พ. ศ. 2530 - 2540 จากฝ่ายจัดสรรน้ำ สำนักอุทกวิทยาและบริหารน้ำ กรมชลประทาน ข้อมูลประกอบด้วย (ข้อมูลอยู่ในภาคผนวก ก หน้า 262 ถึง 285)

1. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนภูมิพล
2. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนสิริกิติ์
3. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนแม่งัดสมบูรณ์ชล
4. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของเขื่อนกิ่วลม
5. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของเขื่อนแม่กวงอุดมธารา
6. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนอุบลรัตน์
7. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนสิรินธร
8. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนจุฬาภรณ์
9. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนลำปาว
10. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนลำตะคอง
11. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนลำพระเพลิง
12. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนน้ำอูน
13. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนห้วยหลวง
14. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนลำนางรอง
15. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนศรีนครินทร์

16. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนเขาแหลม

17. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนแก่งกระจาน

18. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนปราณบุรี

19. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนกระเสียว

20. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนทับเสลา

21. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนบางพระ

22. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนหนองปลาไหล

23. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนรัชชประภา

24. ข้อมูลปริมาณน้ำในอ่างเก็บน้ำของเขื่อนบางลาง

เมื่อได้ข้อมูลจะทำการแบ่งข้อมูลออกเป็น 2 ส่วน ส่วนที่ 1 นำข้อมูลที่เป็นรายเดือนเริ่มจากเดือนมกราคม ปี พ. ศ. 2530 หรือเริ่มจากปีที่ทำการใช้งานของอ่างเก็บน้ำในบางตัวแบบ จนถึงเดือนธันวาคม ปี พ. ศ. 2540 วิเคราะห์หาตัวแบบที่เหมาะสม โดยจะทำการหาตัวแบบที่ใช้พยากรณ์ข้อมูล แล้วทำการพยากรณ์ข้อมูลล่วงหน้า 12 เดือนของปี พ. ศ. 2541 จากนั้นนำข้อมูลที่พยากรณ์ได้ทั้ง 12 เดือน เปรียบเทียบกับข้อมูลจริงของปี พ. ศ. 2541 (12 เดือน) โดยพิจารณาจากค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์ต่ำสุด เพื่อหาวิธีการพยากรณ์ที่เหมาะสม เมื่อได้วิธีการที่เหมาะสมที่สุดจะนำข้อมูลทั้งหมดรวมกับข้อมูล ปี พ. ศ. 2541 (เป็นส่วนที่ 2) หาค่าสัมประสิทธิ์ใหม่และ/หรือตัวแบบใหม่ของค่ารายเดือนในตัวแบบพยากรณ์ของวิธีที่เหมาะสมนั้น และพยากรณ์ค่ารายเดือนใน ปี พ. ศ. 2542 และ ปี พ. ศ. 2543 เทคนิคที่นำมาสร้างตัวแบบพยากรณ์เป็นเทคนิคการพยากรณ์เชิงปริมาณ มีดังต่อไปนี้

3.1.1 วิธีการถดถอย

วิธีการถดถอยจะสร้างตัวแบบเชิงคณิตศาสตร์ที่เรียกว่า "ตัวแบบการถดถอย" เป็นสมการแสดงความสัมพันธ์ระหว่างตัวแปรตาม Y เป็นตัวแปรที่ต้องการพยากรณ์ค่ากับตัวแปรอิสระ X_1, X_2, . . . , X_k ($k \geq 1$) เราจำแนกตัวแบบการถดถอยในเทอมของพารามิเตอร์ เป็นตัวแบบการ

ถดถอยเชิงเส้น (Linear Regression Models) โดยวิธีการหาสมการที่สามารถพยากรณ์ตัวแปรตามจากตัวแปรอิสระซึ่งมีรูปแบบของฟังก์ชันเป็นดังนี้

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_{1,t-1} + \beta_3 X_{2,t-1} + \varepsilon_t$$

Y_t = ปริมาณน้ำในอ่างเก็บน้ำ ณ ต้นเดือน t

Y_{t-1} = ปริมาณน้ำในอ่างเก็บน้ำ ในเดือน t - 1

$X_{1,t-1}$ = ปริมาณน้ำไหลลงอ่างเก็บน้ำ ในเดือน t - 1

$X_{2,t-1}$ = ปริมาณน้ำระบายท้ายเขื่อน ในเดือน t - 1

ε_t = ค่าคลาดเคลื่อนสุ่ม

ขั้นตอนวิธีการพยากรณ์และสร้างตัวแบบการถดถอยเชิงเส้นมีขั้นตอนดำเนินการดังนี้

1. กำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรตาม ซึ่งในที่นี้คือ ปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ ณ ต้นเดือน t (Y_t) และตัวแปรอิสระ Y_{t-1} เป็นปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ในเดือน t, $X_{1,t-1}$ เป็นปริมาณน้ำไหลลงอ่างเก็บน้ำขนาดใหญ่ ในเดือน t - 1 และ $X_{2,t-1}$ เป็นปริมาณน้ำระบายท้ายเขื่อน ในเดือน t - 1 โดยผู้วิจัยจะพิจารณารูปแบบความสัมพันธ์ของตัวแปรตามและตัวแปรอิสระ ซึ่งสามารถพิจารณาโดยการเขียนกราฟระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว เพื่อพิจารณากำหนดรูปแบบความสัมพันธ์เป็นคู่ ๆ ระหว่างตัวแปรตามกับตัวแปรอิสระ

2. ประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยของรูปแบบที่ได้

3. วินิจฉัยความเหมาะสมเชิงสถิติของตัวแบบการถดถอย โดยตรวจสอบและทำการเปรียบเทียบคัดเลือกตัวแบบพยากรณ์ ถ้าพบว่าเป็นตัวแบบที่เหมาะสมกับอนุกรมเวลาชุดนั้น ๆ จริงหรือไม่ โดยพิจารณาคุณสมบัติของความคลาดเคลื่อน (e_t) ; t = 1, 2, 3, . . . , T ที่คำนวณได้แล้วทำการตรวจสอบคุณสมบัติของความคลาดเคลื่อน

4. เปรียบเทียบตัวแบบพยากรณ์ โดยทำการคัดเลือกตัวแบบด้วยการเปรียบเทียบค่าความถูกต้องของการพยากรณ์ ซึ่งในที่นี้จะใช้ค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์

5. พยากรณ์ค่าของข้อมูลปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของกรมชลประทานล่วงหน้าโดยใช้หลักการประมาณปริมาณน้ำไหลลงอ่างเก็บน้ำ

3.1.2 เทคนิคการปรับให้เรียบ

วิธีการปรับให้เรียบสำหรับการพยากรณ์ข้อมูลอนุกรมเวลามีหลายวิธีด้วยกัน สำหรับในการวิจัยในครั้งนี้เทคนิคที่นำมาใช้คือ วิธีการปรับพยากรณ์ของวินเตอร์ จะใช้ในกรณีที่ข้อมูลเป็นข้อมูลอนุกรมเวลาที่มีองค์ประกอบของฤดูกาล ส่วนวิธีการปรับให้เรียบครั้งเดียวแบบเลขชี้กำลังและวิธีการพารามิเตอร์สองตัวของโฮลท์ เป็นกรณีที่ใช้กับข้อมูลที่ไม่มีองค์ประกอบของฤดูกาล ซึ่งเมื่อประมาณค่าพารามิเตอร์ของตัวแบบที่ได้ จะทำการวินิจฉัยความเหมาะสมเชิงสถิติของตัวแบบการถดถอย โดยตรวจสอบและทำการเปรียบเทียบคัดเลือกตัวแบบพยากรณ์ ถ้าพบว่าเป็นตัวแบบที่เหมาะสมกับอนุกรมเวลาชุดนั้น ๆ จริงหรือไม่ โดยพิจารณาคุณสมบัติของความคลาดเคลื่อน (e_t) ; t = 1, 2, 3, . . . , T ที่คำนวณได้ แล้วทำการตรวจสอบคุณสมบัติของความคลาดเคลื่อน วิธีใดจะเหมาะสมมากหรือน้อยผู้วิจัยจะต้องนำมาพิจารณาจากค่าพยากรณ์ที่ได้ในวิธีนั้น ๆ มาวัดค่าความถูกต้องของการพยากรณ์ นั่นคือต้องการให้ได้ค่าคลาดเคลื่อนในการพยากรณ์ต่ำสุด ซึ่งในที่นี้จะใช้ค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์

3.1.3 วิธีการแยกองค์ประกอบ

วิธีการแยกองค์ประกอบจะเป็นการแยกองค์ประกอบข้อมูลอนุกรมเวลา เพื่อพิจารณาว่าองค์ประกอบใดที่มีอิทธิพลต่อการเปลี่ยนแปลงของข้อมูล ในการวิจัยในครั้งนี้เป็นกรณีที่ข้อมูลมีระดับค่าเฉลี่ยคงที่ และมีองค์ประกอบฤดูกาล เมื่อประมาณค่าพารามิเตอร์ของรูปแบบที่ได้ จะทำการวินิจฉัยความเหมาะสมเชิงสถิติของตัวแบบ โดยตรวจสอบและทำการเปรียบเทียบคัดเลือกตัวแบบพยากรณ์ ถ้าพบว่าเป็นตัวแบบที่เหมาะสมกับอนุกรมเวลาชุดนั้น ๆ จริงหรือไม่ โดยพิจารณาคุณสมบัติของความคลาดเคลื่อน (e_t) ; t = 1, 2, 3, . . . , T ที่คำนวณได้ แล้วทำการตรวจสอบคุณสมบัติของความคลาดเคลื่อน ตลอดจนวัดค่าความถูกต้องของการพยากรณ์ นั่นคือต้องการให้ได้ค่าคลาดเคลื่อนในการพยากรณ์ต่ำสุด

3.1.4 วิธีการของบ็อกซ์ – เจนกินส์

วิธีการพยากรณ์อนุกรมเวลาโดยใช้วิธีการของบ็อกซ์ – เจนกินส์ ใช้ได้กับข้อมูลอนุกรมเวลาที่มีองค์ประกอบใด ๆ เช่น แนวโน้ม และวัฏจักร หรือฤดูกาล ซึ่งอาจจะมีองค์ประกอบมากกว่าหนึ่งองค์ประกอบในอนุกรมชุดเดียวกัน ในการวิจัยในครั้งนี้เป็นกรณีที่ข้อมูลมีองค์ประกอบฤดูกาล จึงมีความจำเป็นที่จะต้องใช้ข้อมูลขนาดใหญ่ เพราะถ้าข้อมูลมีจำนวนน้อยเกินไปอาจจะทำให้ไม่สามารถเห็นอิทธิพลหรือรูปแบบของฤดูกาลได้

นอกจากนี้วิธีการของบ็อกซ์ – เจนกินส์ ใช้กับข้อมูลอนุกรมเวลาที่อยู่ในสภาพนิ่งหรือคงที่ (Stationary Data Series) ซึ่งหมายถึงคงที่ในค่าเฉลี่ย คงที่ในค่าความแปรปรวน และคงที่ในความแปรปรวนร่วม และคงที่ในที่หมายถึงไม่แปรผันตามเวลา ซึ่งการทำให้ข้อมูลอนุกรมเวลาที่ไม่คงที่อยู่ในสภาพคงที่จะทำได้โดยการแปลงข้อมูลให้คงที่ ข้อมูลอนุกรมเวลาที่ไม่คงที่ในค่าเฉลี่ยจะใช้วิธีการทำผลต่าง โดยนำข้อมูลมาลบกันได้ข้อมูลอนุกรมชุดใหม่ที่คงที่ ในกรณีที่ข้อมูลอนุกรมเวลาที่ไม่คงที่ในความแปรปรวน วิธีการแปลงข้อมูลที่ใช้กันมากคือ ใส่ ln ในอนุกรมเวลา Y_t สำหรับอนุกรมเวลาที่มีองค์ประกอบของฤดูกาล โดยมีคาบเวลาของฤดูกาลก็ทำการหาผลต่างเช่นเดียวกันเพื่อให้อนุกรมเวลาคงที่ก่อนที่จะนำมาวิเคราะห์ตัวแบบ สำหรับวิธีการของบ็อกซ์–เจนกินส์ในการวิจัยในครั้งนี้จะมีรูปแบบของข้อมูลอนุกรมเวลาเป็นตัวแบบ ARIMA เมื่อมีองค์ประกอบของฤดูกาล

ขั้นตอนวิธีการพยากรณ์และสร้างตัวแบบ ARIMA มีขั้นตอนที่ดำเนินการดังนี้

1. นำข้อมูลอนุกรมเวลามาตรวจสอบดูว่าข้อมูลมีฤดูกาลหรือไม่ จากนั้นนำมาพิจารณาว่าข้อมูลมีลักษณะคงที่ในค่าเฉลี่ยหรือไม่ โดยดูจากค่า ACF ว่ามีลักษณะลดลงอย่างรวดเร็ว ถ้า ACF ไม่มีลักษณะดังที่กล่าวมาแสดงว่าอนุกรมเวลาไม่อยู่ในสภาพคงที่ จะต้องทำการแปลงข้อมูลอนุกรมเวลาชุดนั้น ๆ ให้เป็นอนุกรมเวลาชุดใหม่ที่อยู่ในสภาพคงที่ เมื่อกำหนดรูปแบบข้อมูลปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของกรมชลประทานให้อยู่ในสภาพคงที่แล้ว นำมาเปรียบเทียบกับรูปแบบเบื้องต้นของวิธีการของบ็อกซ์ – เจนกินส์ พิจารณาดูว่ารูปแบบที่ได้ใกล้เคียงกับรูปแบบใดของวิธีการของบ็อกซ์ – เจนกินส์ มากที่สุด

2. ประมาณค่าพารามิเตอร์ของตัวแบบที่ได้

3.ตรวจสอบความเหมาะสมของตัวแบบก่อนที่จะตัดสินใจนำตัวแบบนั้นไปใช้ในการพยากรณ์ โดยพิจารณาคุณสมบัติของความคลาดเคลื่อน (e_t) ; t = 1, 2, 3, . . . , T ที่คำนวณได้ แล้วทำการตรวจสอบคุณสมบัติของความคลาดเคลื่อน แต่ถ้าพบว่าตัวแบบที่กำหนดขึ้นยังไม่สอดคล้องกับคุณสมบัติจะทำการปรับแก้ตัวแบบใหม่ เพื่อประมาณค่าและวินิจฉัยตัวแบบกรรมวิธีเช่นนี้จะกระทำซ้ำ ๆ จนกว่าจะได้ตัวแบบที่เหมาะสมในเชิงสถิติ

4. พยากรณ์ปริมาณน้ำในอ่างเก็บน้ำขนาดใหญ่ของกรมชลประทานล่วงหน้า