บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการศึกษาวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ และเพื่อเปรียบเทียบความถูกต้องของการพยากรณ์จากตัวแบบที่ได้จากวิธีต่างๆ ซึ่งวิธีการคัดเลือกสมการถดถอยที่ได้นำมาใช้ในการสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณมี 3 วิธีดังนี้

1) วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ(BMA_{SVT})

2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด(OPM)

3) วิธีการถดถอยแบบขั้นบันได(SR)

สองวิธีแรกเป็นวิธีการภายใต้แนวทางของเบส์ ส่วนวิธีที่สามเป็นวิธีพื้นฐานซึ่งนิยมใช้กันอย่างแพร่หลายและมีประสิทธิภาพดีในการคัดเลือกสมการถดถอยที่ดีที่สุดเมื่อการวิเคราะห์ความถดถอยเป็นไปตามข้อตกลงเบื้องต้น ซึ่งผู้วิจัยได้นำมาเปรียบเทียบกับวิธีการภายใต้แนวทางของเบส์เพื่อศึกษาดูแนวโน้มว่าวิธีการคัดเลือกตัวแบบภายใต้แนวทางของเบส์ จะมีความเหมาะสมมากกว่าวิธีพื้นฐานหรือไม่เมื่อใช้หลักเกณฑ์คู่สังยุคเบตา ดังนั้นถ้าผลการศึกษาวิจัยพบว่าวิธีการใดมีประสิทธิภาพในการพยากรณ์มากที่สุดก็ควรจะนำไปใช้ในการคัดเลือกตัวแบบการถดถอยเชิงเส้นพหุคูณที่ดีที่สุดต่อไป ซึ่งมีการกำหนดสถานการณ์ต่างๆ ในการวิจัยครั้งนี้ไว้ดังนี้

1) การแจกแจงของค่าคลาดเคลื่อนสุ่มเป็นแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ
2) กำหนดขนาดตัวอย่างที่ศึกษา คือ 10 30 50 และ 100
3) จำนวนตัวแปรอิสระที่ศึกษาคือ 3 5 10 และ 15 ตามลำดับ
4) ค่าคงที่สำหรับวิธี BMA_{SVT} และวิธีOPM มี 4 ระดับ คือ (1, 5) (1,10) (10 , 100) และ (10,500)

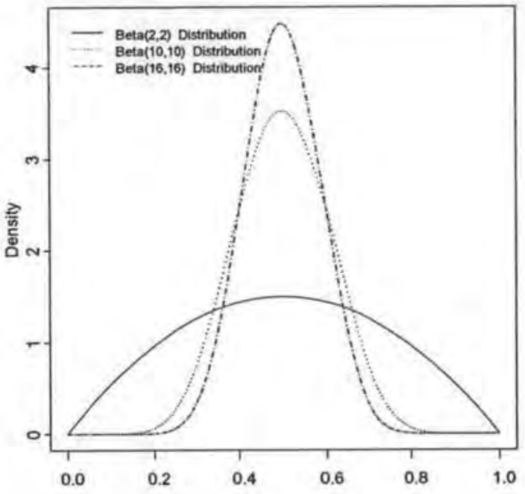
เกณฑ์ที่ใช้ในการตัดสินใจว่าวิธีการคัดเลือกตัวแบบการถดถอยวิธีใดจะมีประสิทธิภาพมากที่สุดจะพิจารณาจากค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Average of Mean Square Error (AMSE)) และใช้ค่าอัตราส่วนผลต่างของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Ratio of Different Average Mean Square Error (RDAMSE)) เพื่อประกอบการเปรียบเทียบประสิทธิภาพของวิธีการข้างต้นซึ่งวิธีใดให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่า

อัตราส่วนผลต่างของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดจะเป็นวิธีที่ดีที่สุด ผลการวิจัยได้ข้อสรุปดังนี้

5.1 <u>สรุปผลการวิจัย</u>

5.1.1 <u>ผลการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตาในกรณีที่</u> $\alpha = \beta = 2$, $\alpha = \beta = 10$ <u>และ</u> $\alpha = \beta = 16$

จากการเปรียบเทียบของทั้ง 3 กรณี พบว่าในกรณี[1]ที่ $\alpha = \beta = 2$ ให้ค่าAMSE ของแต่ละวิธีน้อยที่สุด รองลงมาได้แก่ กรณีที่ $\alpha = \beta = 10$ และกรณีที่ $\alpha = \beta = 16$ เนื่องจากการแจกแจงเบตาหลายตัวแปรอยู่ในรูปแบบสมมาตรซึ่งมีพารามิเตอร์เป็นค่าคงที่เท่ากันทุกค่า ดังนั้นเมื่อให้ค่า $\alpha = \beta = 2$ จึงเป็นค่าที่มีรูปแบบระฆังคว่ำที่มีพื้นที่มากที่สุดและมีลักษณะใกล้เคียงการแจกแจงปกติมากที่สุด แสดงได้ดังรูปที่ 5.1



<u>รูปที่ 5.1</u> แสดงการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตาในกรณีที่ $\alpha = \beta = 2$, $\alpha = \beta = 10$ และ $\alpha = \beta = 16$

[1] Antti Honkela, "Dirichlet distribution," http://www.cis.hut.fi/ahonkela/dippa/node95.html#fig:dirichletpdf (May 2001).

5.1.2 ผลการเปรียบเทียบค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย

จากการเปรียบเทียบค่า AMSE ของทั้ง 3 วิธี พบว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่วิธี BMA_{SVT} OPM และSR ตามลำดับ แสดงว่าวิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากที่สุดในการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตา วิธี OPM นั้นมีความสามารถในการคัดเลือกสมการถดถอยใกล้เคียงกับวิธี BMA_{SVT} ซึ่งจะสังเกตได้จากค่า AMSE ของวิธี OPM จะสูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อยเท่านั้น ส่วนวิธี SR มีค่า AMSE สูงกว่าวิธี BMA_{SVT} และวิธีOPM อย่างชัดเจนในทุกสถานการณ์ ซึ่งจากผลการวิจัยที่ได้นี้สรุปได้ว่าวิธี BMA_{SVT} มีความสามารถในการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตาได้ดีที่สุด และอาจกล่าวได้ว่าวิธีการภายใต้แนวทางของเบส์มีความสามารถในการคัดเลือกสมการถดถอยเชิงเส้นพหุคูณดีกว่าวิธีการพื้นฐานซึ่งในที่นี้คือวิธีการถดถอยแบบขั้นบันได โดยสังเกตได้จากวิธีการภายใต้แนวทางของเบส์จะให้ค่า AMSE ต่ำกว่าวิธีการพื้นฐานอย่างชัดเจนทุกสถานการณ์ ส่วนเรื่องเกี่ยวกับหลักการความไม่แน่นอนของตัวแบบและการแปลงที่เหมาะสมของตัวแปรอิสระจะสรุปได้ว่าวิธี BMA_{SVT} ซึ่งคำนึงถึงหลักดังกล่าวจะให้ค่าพยากรณ์ที่ได้มีความถูกต้องและแม่นยำมากยิ่งขึ้น

5.1.3 ปัจจัยที่มีผลต่อค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี BMA_{SVT} และวิธี OPM

1) ขนาดตัวอย่าง (n)

เมื่อขนาดตัวอย่างเพิ่ม ค่า AMSE ของทุกวิธีมีแนวโน้มลดลง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้ และเมื่อขนาดตัวอย่างเพิ่มขึ้นวิธี BMA_{SVT} และวิธี OPM จะมีค่า AMSE ต่ำกว่าวิธี SR อย่างชัดเจน

2) ค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่ม (σ)

เมื่อค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่าเพิ่มขึ้น ค่าAMSEของทุกวิธีมีแนวโน้มเพิ่มขึ้น ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

3) จำนวนตัวแปรอิสระ

เมื่อจำนวนตัวแปรอิสระในตัวแบบการถดถอยเพิ่มขึ้นจะส่งผลให้ค่า AMSEของทุกวิธีมีแนวโน้มสูงขึ้นเพราะในงานวิจัยนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูป เมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้น ทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มสูงขึ้น โดยเฉพาะอย่างยิ่งวิธีภายใต้แนวทางของเบส์ทั้ง 2 วิธีคือวิธีBMA_{SVT}และวิธีOPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการ

พิจารณาถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT}จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น

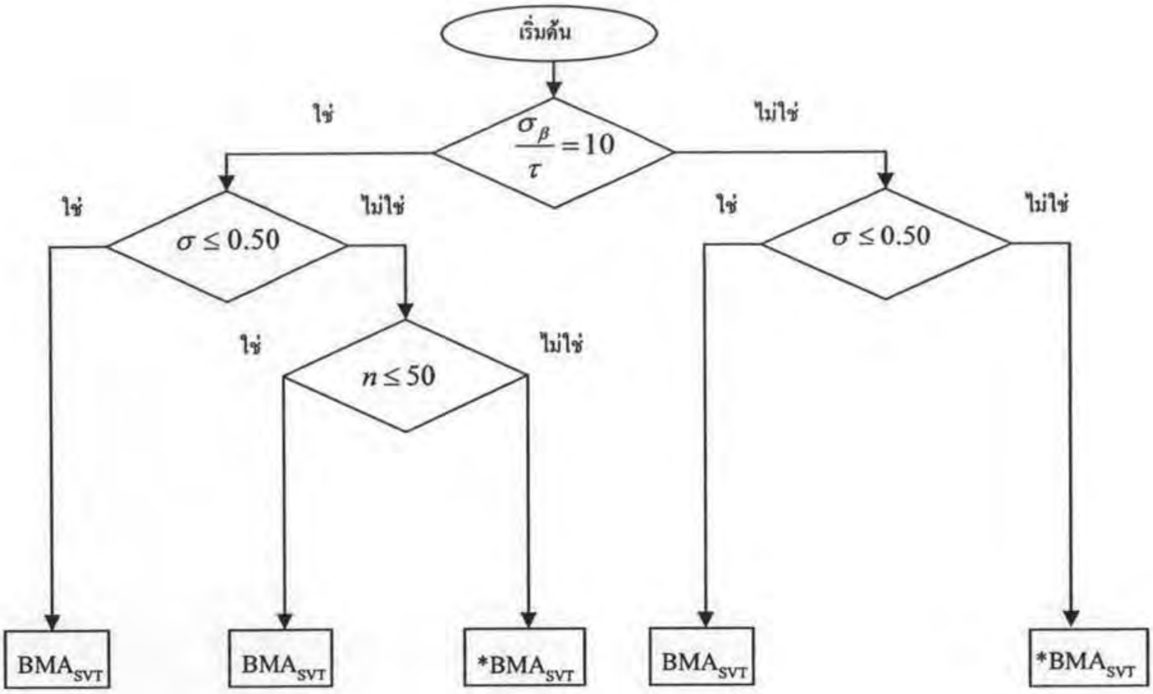
4) <u>ค่าคงที่ $\frac{\sigma_\beta}{\tau}$ และ</u> c

ค่าคงที่ $\frac{\sigma_\beta}{\tau}$ และ c จะมีผลต่อค่า AMSE ของทั้งวิธี BMA_{SVT} และวิธีOPM เท่านั้น โดยค่า AMSE จะเพิ่มขึ้นเมื่อค่าคงที่สูงขึ้น เนื่องจากค่าคงที่เป็นค่าที่กำหนดลักษณะการกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยที่สุ่มได้ ค่าคงที่ที่สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น ทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้ค่า AMSE มีค่าสูงขึ้น การเปลี่ยนแปลงของค่าคงที่นี้จะส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM มากพอควรซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE เมื่อค่าคงที่เพิ่มขึ้น

5.1.4 <u>ผลสรุปการเลือกวิธีการสร้างตัวแบบถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตา</u>

ผลสรุปการเลือกวิธีการสร้างตัวแบบถดถอยที่ดีที่สุดเชิงเบส์ในการวิจัยนี้ พบว่าการสร้างสมการถดถอยด้วยวิธี BMA_{SVT}จะส่งผลให้ได้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากที่สุด ส่วนวิธีการสร้างสมการถดถอยที่มีประสิทธิภาพในการพยากรณ์รองลงมาคือวิธี OPM และวิธี SR ตามลำดับ และจากผลการวิจัยเปรียบเทียบค่า AMSE ของทั้ง3 วิธี แม้ว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่วิธี BMA_{SVT} OPM และ SR ตามลำดับ แต่พบว่ามีบางกรณีที่วิธี BMA_{SVT} มีค่าต่ำกว่าวิธีอื่นๆค่อนข้างมาก ซึ่งแสดงให้เห็นว่าในกรณีดังกล่าวเราควรเลือกวิธี BMA_{SVT} ในการสร้างสมการถดถอยเพื่อให้ได้ค่าพยากรณ์ที่ถูกต้องและแม่นยำมากที่สุด ส่วนในกรณีอื่นๆนั้นแม้ว่าวิธี BMA_{SVT} จะให้ค่า AMSE ต่ำกว่าวิธีอื่นๆแต่ในการเลือกใช้วิธีการสร้างสมการถดถอย เราอาจคำนึงถึงปัจจัยด้านอื่นๆด้วย เช่น ระยะเวลาในการคำนวณ ซึ่งวิธี BMA_{SVT} นั้นใช้ระยะเวลาในการคำนวณค่อนข้างมากเมื่อเทียบกับวิธี OPM โดยแผนผังแสดงข้อสรุปของกรณีต่างๆที่วิธี BMA_{SVT} จะให้ค่า AMSE ต่ำกว่าวิธี OPM อย่างชัดเจนแสดงดังรูปที่ 5.2[2]

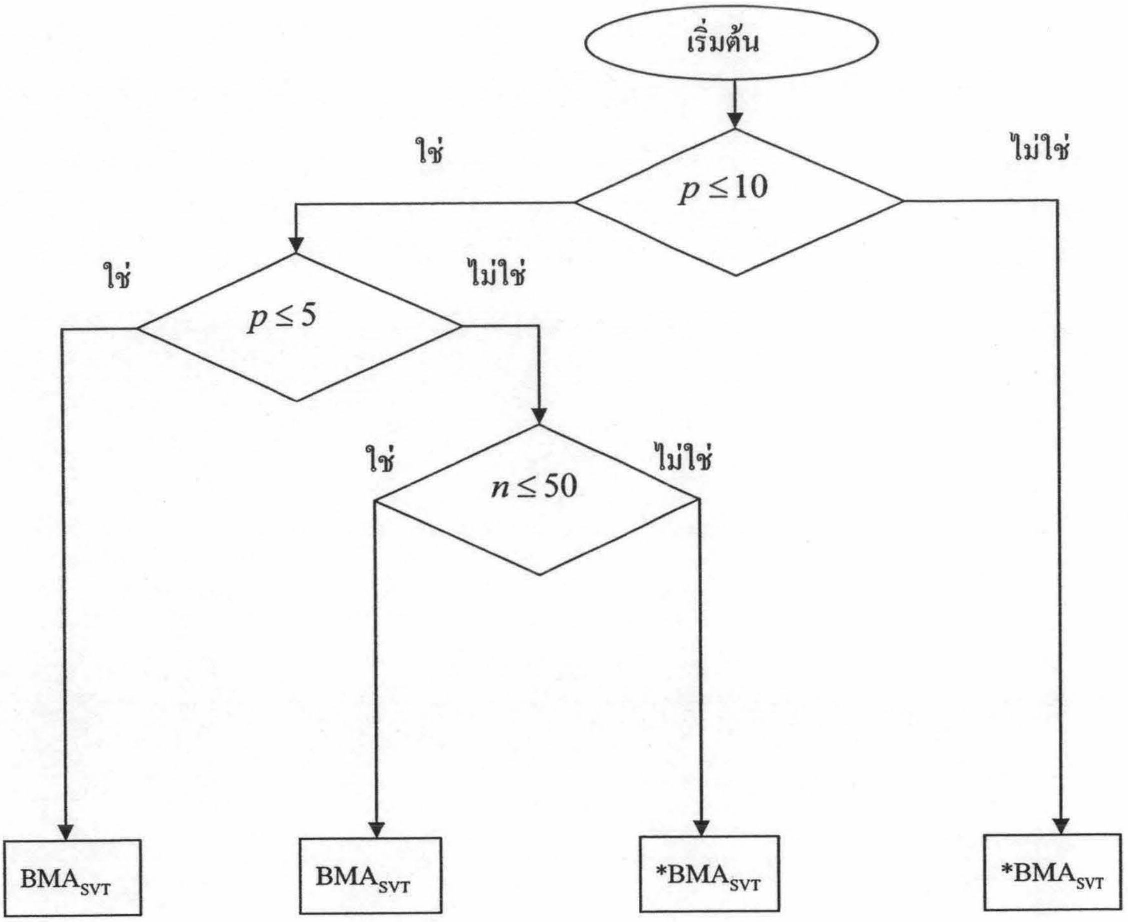
[2] จิตติมา ผสมญาติ, "การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ," (วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย,2546) หน้า 133-134.



รูปที่ 5.2 แสดงแผนผังผลสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตา

หมายเหตุ * หมายถึงกรณีที่วิธี BMA_{SVT} ให้ค่า AMSE ต่ำกว่าวิธี OPM ค่อนข้างมาก

ในทางปฏิบัตินอกจากผู้วิจัยจะต้องคำนึงถึงระยะเวลาในการคำนวณแล้วขนาดตัวอย่าง (n) ยังเป็นอีกปัจจัยหนึ่งที่ผู้วิจัยมักคำนึงถึงก่อนเสมอ ด้วยเหตุนี้ผู้วิจัยจึงขอเสนอแนวทางการพิจารณาเลือกใช้วิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตา โดยแผนผังแสดงข้อสรุปของกรณีต่างๆ ที่วิธี BMA_{SVT} จะให้ค่า AMSE ต่ำกว่าวิธี OPM ดังแสดงรูปที่ 5.3



<u>รูปที่ 5.3</u> แสดงแผนผังผลสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคเบตาในเชิงปฏิบัติ

<u>หมายเหตุ</u> * หมายถึงกรณีที่วิธี BMA_{SVT} ให้ค่า AMSE ต่ำกว่าวิธี OPM ค่อนข้างมาก

5.2 <u>ข้อเสนอแนะ</u>[3]

1) การวิจัยครั้งนี้ศึกษาเฉพาะการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณไม่รวมถึงตัวแบบความถดถอยพหุนาม ซึ่งเป็นกรณีเฉพาะที่มักพบได้บ่อยครั้งในการวิเคราะห์ความถดถอย ดังนั้นในงานวิจัยแบบเต็มรูป ควรมีการศึกษาเปรียบเทียบวิธีการคัดเลือกสมการถดถอยสำหรับตัวแบบความถดถอยพหุนามด้วย ทั้งนี้เพื่อศึกษาว่าการนำวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี ไปใช้กับการวิเคราะห์ความถดถอยพหุนามจะให้ผลสรุปเหมือนหรือแตกต่างกับการนำไปใช้การวิเคราะห์ความถดถอยเชิงเส้นพหุคูณดังเช่นในงานวิจัยครั้งนี้อย่างไร

2)วิธี BMA_{SVT}ซึ่งคำนึงถึงหลักการความไม่แน่นอนของตัวแบบเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระโดยเลือกใช้วิธีการแปลงเปลี่ยนจุด (change-point transformation) ช่วยให้ได้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากขึ้น หากมีการพิจารณาข้อมูลที่มีค่าผิดปกติ (outlier identification) ควบคู่ไปกับการพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (transformation selection) อาจทำให้วิธีการเฉลี่ยตัวแบบของเบส์มีประสิทธิภาพมากยิ่งขึ้น

3) เนื่องจากในการวิจัยครั้งนี้วิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากที่สุดอย่างคงเส้นคงวาและกรณีที่ค่าคงที่ ($\frac{\sigma_\beta}{\tau}$ และ c) ของวิธี BMA_{SVT} และวิธีOPM มีค่าต่ำๆ วิธี OPM จะให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำใกล้เคียงกับวิธี BMA_{SVT} มาก ส่วนกรณีที่ค่าคงที่ ($\frac{\sigma_\beta}{\tau}$ และ c) ของวิธี BMA_{SVT} และวิธีOPM มีค่าสูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้นทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้วิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากกว่าวิธี OPM ชัดเจนยิ่งขึ้น ดังนั้นในการพิจารณาว่าวิธีการใดจะมีประสิทธิภาพมากที่สุดสำหรับกรณีที่วิธี BMA_{SVT} ให้ค่า AMSE ต่ำกว่าวิธีอื่นๆอย่างคงเส้นคงวาจึงควรใช้การวิเคราะห์ความแปรปรวนของการทดลองแบบสุ่มโดยสมบูรณ์ภายในกลุ่ม (Randomized Completely Block Design : RCBD)เพราะการวิเคราะห์ดังกล่าวจะทำให้ได้วิธีการคัดเลือกสมการถดถอยที่มีประสิทธิภาพมากที่สุดอย่างมีนัยสำคัญทางสถิติ(ดังแสดงตัวอย่างการวิเคราะห์ในหัวข้อที่ 5 ของภาคผนวก)

---

[3] จิตติมา ผสมญาติ, "การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ," (วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย,2546) หน้า 135.