การเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอนุกรมเวลา

COMPARING NONPARAMETRIC REGRESSION METHODS FOR TIME SERIES DATA

บุษราพรรณ กันธรรม

BOOSARAPHAN KANTHOM

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติและการวิเคราะห์ธุรกิจ

ภาควิชาสถิติ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2565

KMITL-2022-SC-M-050-077

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COMPARING NONPARAMETRIC REGRESSION METHODS FOR TIME SERIES DATA

BOOSARAPHAN KANTHOM

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS AND BUSINESS ANALYTICS

DEPARTMENT OF STATISTICS SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2022

KMITL-2022-SC-M-050-077

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2022

SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอนุกรมเวลา
ชื่อนักศึกษา	นางสาว บุษราพรรณ กันธรรม
รหัสประจำตัว	62605113
ปริญญา	วิทยาศาสตรมหาบัณฑิต (สถิติและการวิเคราะห์ธุรกิจ)
ภาควิชา	สถิติ
พ.ศ.	2565
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รองศาสตราจารย์ ดร.อัชฌา อระวีพร

บทคัดย่อ

ในการวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี สำหรับข้อมูลอนุกรมเวลา ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ โดยพิจารณาประสิทธิภาพของวิธีการถดถอยไม่อิงพารามิเตอร์จากค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด โดยใช้ข้อมูลจำลองที่มีลักษณะแนวโน้ม ไม่เชิงเส้น ไม่คงที่ และความแปรปรวนไม่คงที่ ที่มีส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่ากับ 1, 3, 5 และ 7 ด้วยขนาดตัวอย่าง 50, 100, 150 และ 200 จำลองข้อมูลด้วยโปรแกรมอาร์ โดยทำซ้ำ 500 รอบในแต่ละสถานการณ์ พบว่าวิธีเส้นโค้งกำลังสามธรรมชาติของข้อมูลจำลอง มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกสถานการณ์ที่ทำการศึกษา และค่านอตมีค่าใกล้เคียงกับขนาดตัวอย่าง เมื่อพิจารณาลักษณะของข้อมูลพบว่าที่ขนาดตัวอย่างเดียวกันข้อมูลความแปรปรวนไม่คงที่จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกขนาดตัวอย่าง สำหรับข้อมูลจริงศึกษา ข้อมูลรายวันของอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา จำนวน 200 วัน และข้อมูลรายเดือนของปริมาณการนำเข้าไฟฟ้า จำนวน 200 เดือน และนำตัวแบบการถดถอยไม่อิงพารามิเตอร์ที่มีประสิทธิภาพไปใช้ในการพยากรณ์ล่วงหน้า โดยมีเกณฑ์ คือค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด จากผลการประมาณค่าและพยากรณ์ พบว่าในการประมาณค่าวิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดทั้ง 2 ชุดข้อมูล ส่วนการพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกามีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยของวิธีเส้นโค้งกำลังสามธรรมชาติต่ำสุด และปริมาณการนำเข้าไฟฟ้ามีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยของวิธีเส้นโค้งบีต่ำสุด จึงสรุปได้ว่าวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีเป็นวิธีที่มีประสิทธิภาพสำหรับข้อมูลอนุกรมเวลา

คำสำคัญ : การถดถอยไม่อิงพารามิเตอร์ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี วิธีเส้นโค้งที่ถูกลงโทษ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Comparing Nonparametric Regression Methods for Time Series Data
Student Name	Boosaraphan Kanthom
Student ID	62605113
Degree	Master of Science (Statistics and Business Analytics)
Department	Statistics
Year	2022
Thesis Advisor	Assoc. Prof. Dr. Autcha Araveeporn

Abstract

This research aims to study and compare the efficiency of four nonparametric regression methods for time series data: smoothing spline, natural cubic spline, B-spline, and penalized spline methods. The efficiency of the nonparametric regression method determines by the lowest mean squared error. The data character is simulated in trend, non-linear, non-stationary, and heteroscedasticity data, while the standard deviation of error is 1, 3, 5, and 7. The R program sets the sample sizes as 50, 100, 150, and 200 and repeats 500 times in each situation. The results were founded that the natural cubic spline method played the lowest mean squared error in all cases, and the number of knots got to close the sample size. When the data character was considered, the heteroscedasticity had the lowest mean squared error in all sample sizes. For the real data, the daily exchange rate of Thai Baht against the US dollar for a total of 200 days and month of volume import of electricity of 200 months. Then, the efficiency of the nonparametric regression model is considered to predict future values depending on the lowest mean absolute percentage error criterion. From the estimating and forecasting two real datasets, it can be seen that the natural cubic spline method presented the lowest mean square error of two real data sets for estimating. For forecasting, the Thai Baht against the US dollar had the lowest mean absolute percentage error of the natural cubic spline method, and the volume import of electricity had the lowest mean absolute percentage error of the B-spline method. It can be concluded that the natural cubic spline and B-spline methods were the performance methods for time series data.

Keywords : Nonparametric Regression, Smoothing Spline Method, Natural Cubic Spline Method, B- Spline Method, Penalized Spline Method

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์เรื่อง การเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอนุกรมเวลา สำเร็จลุล่วงไปได้ด้วยดีได้ด้วยความอนุเคราะห์จาก รศ.ดร. อัชฌา อระวีพร อาจารย์ที่ปรึกษาวิทยานิพนธ์ ในการให้คำแนะนำ ปรึกษา รวมทั้งตรวจสอบงานการวิจัยในด้านความถูกต้องของเนื้อหา ตลอดจนติดตามการทำวิจัยจนทำให้การทำวิทยานิพนธ์เล่มนี้มีความสมบูรณ์มากที่สุด จึงขอกราบขอบพระคุณด้วยความเคารพอย่างสูงไว้ ณ ที่นี้ด้วย

ผู้วิจัยขอขอบพระคุณ ประธานกรรมการ รศ.ดร. วราฤทธิ์ พานิชกิจโกศลกุล อาจารย์ประจำภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ (ศูนย์รังสิต) และอาจารย์บัณฑิตประจำ ผศ.ดร. สมศรี บัณฑิตวิไล อาจารย์ประจำภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่เสียสละเวลาในการให้คำแนะนำที่เป็นประโยชน์ในการพัฒนาเนื้อหาวิทยานิพนธ์ให้มีความสมบูรณ์มากยิ่งขึ้น

นอกจากนี้ขอขอบคุณ คณาจารย์ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง รวมทั้งบุคคลอื่นๆ ที่เกี่ยวข้อง ที่ให้การช่วยเหลือในการศึกษา จนกระทั่งวิทยานิพนธ์เล่มนี้สำเร็จไปได้ด้วยดี หากมีข้อผิดพลาดประการใด ผู้วิจัยขออภัยมา ณ ที่นี้ด้วย

นางสาวบุษราพรรณ กันธรรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ก
บทคัดย่อภาษาอังกฤษ..... ข
กิติกรรมประกาศ..... ค
สารบัญ..... ง
สารบัญตาราง..... ฉ
สารบัญรูป..... ช
บทที่ 1 บทนำ..... 1
1.1 ที่มาและความสำคัญของปัญหา..... 1
1.2 วัตถุประสงค์..... 5
1.3 ขอบเขตของการศึกษา..... 6
1.4 เกณฑ์การตัดสินใจ..... 7
1.5 ประโยชน์ที่คาดว่าจะได้รับ..... 8
1.6 นิยามศัพท์..... 9
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง..... 10
2.1 ทฤษฎีที่เกี่ยวข้อง..... 10
2.1.1 การแจกแจงปรกติ..... 10
2.1.2 ตัวแบบการถดถอยไม่อิงพารามิเตอร์..... 11
2.1.3 การถดถอยไม่อิงพารามิเตอร์..... 11
2.1.3.1 วิธีเส้นโค้งเรียบ..... 12
2.1.3.2 วิธีเส้นโค้งกำลังสามธรรมชาติ..... 13
2.1.3.3 วิธีเส้นโค้งบี..... 14
2.1.3.4 วิธีเส้นโค้งที่ถูกลงโทษ..... 15
2.1.4 พารามิเตอร์ปรับให้เรียบ..... 15
2.2 งานวิจัยที่เกี่ยวข้อง..... 16
บทที่ 3 วิธีดำเนินงานวิจัย..... 18
3.1 การวางแผนการวิจัย..... 18
3.2 วิธีการดำเนินงานวิจัยสำหรับข้อมูลจำลอง..... 19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

3.3 การประยุกต์ใช้กับข้อมูลจริง 30
3.4 ขั้นตอนของโปรแกรมที่ใช้ในการวิจัย 31
3.4.1 ข้อมูลจำลอง 31
3.4.2 การประยุกต์ใช้กับข้อมูลจริง 33
บทที่ 4 ผลการวิจัย 34
4.1 ผลการวิเคราะห์ข้อมูลจำลอง 34
4.2 ผลการวิเคราะห์การประยุกต์ใช้กับข้อมูลจริง 39
4.2.1 การประมาณค่าข้อมูลจริง 41
4.2.2 การพยากรณ์ข้อมูลจริง 47
บทที่ 5 สรุปผลการวิจัย อภิปรายผลการวิจัยและข้อเสนอแนะ 50
5.1 สรุปผลการวิจัย 50
5.2 อภิปรายผลการวิจัย 51
5.3 ข้อเสนอแนะ 52
บรรณานุกรม 54
ภาคผนวก 56
ภาคผนวก ก คำสั่งโปรแกรมอาร์ที่ใช้ในงานวิจัย 57
ประวัติผู้เขียน 70

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

4.1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1..... 35

4.2 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3..... 36

4.3 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 5..... 37

4.4 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 7..... 38

4.5 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)..... 46

4.6 ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)..... 49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่		หน้า
2.1	การแจกแจงปรกติที่มีค่าเฉลี่ยเท่ากันและความแปรปรวนไม่เท่ากัน	10
3.1	ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=1	19
3.2	ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=3	20
3.3	ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=5	20
3.4	ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=7	21
3.5	ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=1	21
3.6	ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=3	22
3.7	ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=5	22
3.8	ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=7	23
3.9	ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=1	23
3.10	ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=3	24
3.11	ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=5	24
3.12	ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=7	25
3.13	ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=1	25
3.14	ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=3	26
3.15	ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=5	26
3.16	ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=7	27
3.17	ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=1	27
3.18	ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=3	28
3.19	ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=5	28
3.20	ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=7	29
3.21	แผนผังขั้นตอนการทำงานของโปรแกรมของข้อมูลจำลอง	31
3.22	แผนผังขั้นตอนการทำงานของโปรแกรมของการประยุกต์ใช้กับข้อมูลจริง	33
4.1	อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริการายวัน	40
4.2	ปริมาณการนำเข้าไฟฟ้าของประเทศไทยรายเดือน	40
4.3	การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งเรียบ (SS)	41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
4.4	การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งกำลังสามธรรมชาติ (NS)	42
4.5	การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งบี (BS)	42
4.6	การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งที่ถูกลงโทษ (PS)	43
4.7	การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งเรียบ (SS)	44
4.8	การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งกำลังสามธรรมชาติ (NS)	44
4.9	การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งบี (BS)	45
4.10	การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งที่ถูกลงโทษ (PS)	45
4.11	การพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)	47
4.12	การพยากรณ์ข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)	48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลเพื่อหาตัวแบบความสัมพันธ์ระหว่างตัวแปรตาม (Dependent Variable) แทนด้วยสัญลักษณ์ y และตัวแปรอิสระ (Independent Variable) แทนด้วยสัญลักษณ์ x โดยวัตถุประสงค์เพื่อประมาณหรือพยากรณ์ตัวแปรตามจากตัวแบบที่สร้างขึ้นที่เรียกว่าตัวแบบการถดถอย (Regression Model) หรือตัวแบบการถดถอยอิงพารามิเตอร์ (Parametric Regression Model) ในกระบวนการสร้างตัวแบบการถดถอย จะมีข้อสมมติเบื้องต้น คือ ตัวแปรอิสระอยู่ในมาตรวัดนามบัญญัติ (Nominal Scale) ขึ้นไปและตัวแปรตามอยู่ในมาตรวัดแบบช่วง (Interval Scale) ขึ้นไป โดยมีความสัมพันธ์กันในลักษณะเชิงเส้นและมีการแจกแจงปรกติ ความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่ มีการแจกแจงปรกติ และตัวแปรตามไม่มีความสัมพันธ์กันเอง (Non-Autocorrelation) นอกจากนี้ตัวแปรอิสระไม่เกิดความสัมพันธ์เชิงเส้นพหุ (Multicollinearity) แต่ในบางครั้งข้อมูลที่ได้มาไม่เป็นไปตามข้อสมมติเบื้องต้น เช่น ตัวแปรตามและตัวแปรอิสระเกิดความสัมพันธ์ไม่เชิงเส้น (Non-Linear) หรือเป็นข้อมูลอนุกรมเวลา (Time Series) ที่มีรูปแบบไม่แน่นอน ไม่เป็นเชิงเส้นและอาจเป็นเส้นโค้ง ทำให้ค่าความคลาดเคลื่อนสุ่มเกิดความสัมพันธ์กันเอง การใช้ตัวแบบการถดถอยอิงพารามิเตอร์อาจทำให้เกิดความคลาดเคลื่อนสูงในการประมาณค่าหรือพยากรณ์ตัวแปรตามได้ เพื่อแก้ปัญหาการวิเคราะห์การถดถอยที่ไม่ตรงตามข้อสมมติเบื้องต้น จึงมีการนำเอาวิธีการถดถอยไม่อิงพารามิเตอร์ (Nonparametric Regression) มาใช้ในการพยากรณ์ค่าของตัวแปรตาม

วิธีการถดถอยไม่อิงพารามิเตอร์หรือเรียกว่าวิธีการปรับให้เรียบ (Smoothing Method) เป็นอีกทางเลือกหนึ่งในการแสดงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระในลักษณะไม่เชิงเส้นอาจแสดงในลักษณะเส้นโค้ง โดยจะทำการสร้างเส้นโค้งให้เรียบไปตามการกระจายของข้อมูล เรียกว่า แผนภาพการกระจายปรับให้เรียบ (Scatterplot Smoothing) โดยวิธีการเหล่านี้พิจารณาการกระจายข้อมูลของตัวแปรตามเมื่อกำหนดตัวแปรอิสระเพื่อให้ได้การประมาณเส้นโค้งที่ใกล้เคียงกับการกระจายมากที่สุด (Hastie และ Tibshirani, 1990) ซึ่งข้อมูลอาจจะมีลักษณะการกระจายไม่มีรูปแบบหรือเป็นโค้งหลาย ๆ โค้ง จึงต้องพิจารณาค่าพารามิเตอร์ปรับให้เรียบ (Smoothing Parameter) ซึ่งเป็นค่าที่ใช้ในการปรับเส้นโค้งโดยจะกำจัดหรือลดการเปลี่ยนแปลงขึ้นลงของข้อมูลเพื่อให้การประมาณค่าแทรกไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตามเส้นโค้งของข้อมูลและปรับให้เรียบขึ้น การปรับเส้นโค้งที่เกิดจากการเชื่อมต่อของเส้นโค้งต่าง ๆ ด้วยข้อต่อที่เรียกว่า นอต (Knot) ซึ่งเป็นเทคนิคของการปรับข้อมูลให้เรียบ ซึ่งวิธีการปรับให้เรียบมีหลายวิธีที่นิยมใช้ในปัจจุบัน เช่น วิธีการถดถอยเคอร์เนล (Kernel Regression Method) (Nadaraya, 1964 และ Watson, 1964) วิธีเส้นโค้งเรียบ (Smoothing Spline Method) (Wahba, 1990) วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method) (Eilers และ Mark, 1996 และ Ruppert และคณะ, 2003) และวิธีกรองแนวโน้ม (Trend Filtering Method) (Kim และคณะ, 2009)

วิธีการถดถอยเคอร์เนล (Kernel Regression Method) เป็นวิธีดั้งเดิมที่มีการนำเสนอโดย Nadraya (1964) และ Watson (1964) สำหรับการวิเคราะห์การถดถอยไม่อิงพารามิเตอร์ชนิดหนึ่งโดยใช้ฟังก์ชันเคอร์เนลในการประมาณค่าแนวโน้มของข้อมูลร่วมกับค่าพารามิเตอร์แบบปรับให้เรียบ เป็นการประมาณค่าจากข้อมูลโดยการหาค่าเฉลี่ยรอบจุดที่ต้องการประมาณค่าจากตัวแบบและเป็นการหาค่าเฉลี่ยในบริเวณนั้นๆ เพื่อประมาณค่า โดยถ่วงน้ำหนักจุดที่ใกล้กับจุดโฟกัสด้วยฟังก์ชันเคอร์เนล (Kernel Function) ทั้งนี้บริเวณที่ใกล้จุดโฟกัสที่ทำการถ่วงน้ำหนักจากฟังก์ชันเคอร์เนลมากกว่าจุดที่ห่างออกไป ทำให้สมการถดถอยมีรูปร่างไปตามลักษณะของข้อมูล วิธีนี้เป็นการประมาณค่าแต่ละค่าของแต่ละจุดทำให้มีข้อผิดพลาดง่ายจึงมีการพัฒนาเป็น วิธีการถดถอยพหุนามเฉพาะที่ (Local Polynomial Regression) (Stone, 1977) เป็นวิธีการสร้างตัวแบบจากวิธีการถดถอยเคอร์เนลโดยพิจารณาระดับพหุนามและการกระจายของเทย์เลอร์ (Taylor) ในการประมาณค่าพารามิเตอร์

Wahba (1990) ได้นำเสนอวิธีเส้นโค้งเรียบ (Smoothing Spline Method) เป็นวิธีที่นิยมใช้ในการปรับข้อมูลให้เรียบ ซึ่งใช้ประมาณเส้นโค้งพหุนามธรรมดา โดยการลดผลรวมกำลังสองของฟังก์ชันที่ถูกลงโทษ (Penalized) บนพารามิเตอร์ปรับให้เรียบและมีนอตที่ไม่ซ้ำกันในแต่ละค่าของตัวแปรอิสระเพื่อควบคุมการปรับข้อมูลให้เหมาะสม โดยใช้ค่าประมาณกำลังสองน้อยสุด (Least Squares Estimation) ในการลงโทษความหยาบของข้อมูล และยังมีการพัฒนาวิธีนี้โดย Green และ Silverman (1994) เรียกว่าวิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method) โดยเป็นการพิจารณาการประมาณค่าจากความคลาดเคลื่อนกำลังสองต่ำสุดบนการปริพันธ์ของการอนุพันธ์ลำดับที่สองยกกำลังสอง นอกจากนี้ยังมีวิธีเส้นโค้งอีกวิธีที่นิยมใช้คือวิธีเส้นโค้งบี (B-Spline Method) (De Boor, 1978) เป็นวิธีการที่ใช้ฟังก์ชันที่เรียกว่าฟังก์ชันเส้นโค้งบี (B-Spline Function) สำหรับการประมาณค่าตัวแบบการถดถอยไม่อิงพารามิเตอร์

วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method) เป็นอีกวิธีในการปรับข้อมูลให้เรียบโดย Eilers และ Marx (1996) ได้พัฒนาฟังก์ชันเส้นโค้งที่ถูกลงโทษจากการถดถอยแบบสไปลน์ (Spline Regression) และเส้นโค้งเรียบ โดยการประมาณค่าฟังก์ชันปรับให้เรียบจะอาศัยพื้นฐานของฟังก์ชันการแบ่งค่าแบบยกกำลัง (Truncated Power Function) และเนื่องจากการปรับให้เรียบถูกควบคุมโดย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พารามิเตอร์ปรับให้เรียบที่ถูกลงโทษ จำนวนนอตและตำแหน่งของนอตจึงไม่มีความสำคัญ ต่อมา Ruppert และคณะ (2003) ได้นำเสนอการประมาณค่าพารามิเตอร์โดยวิธีการประมาณภาวะความน่าจะเป็นสูงสุดเฉพาะ (Restricted Maximum Likelihood Method) ซึ่งเป็นวิธีการที่ประมาณค่าได้ง่ายและแม่นยำกว่า

Kim และคณะ (2009) ได้นำเสนอวิธีการกรองแนวโน้ม (Trend Filtering Method) โดยแสดงวิธีการกรองในแบบต่าง ๆ พิจารณาจากเกณฑ์ต่ำสุดของฟังก์ชันกำลังสองน้อยสุดที่ถูกลงโทษ สามารถปรับตำแหน่งของนอตได้ (Tibshirani, 2013) และสร้างการประมาณแนวโน้มที่เป็นเชิงเส้นแบบช่วง (Piecewise Linear) ซึ่งช่วยในการวิเคราะห์อนุกรมเวลา (Time Series Analysis) ที่มีแนวโน้มเชิงเส้นแบบช่วง มีส่วนโค้ง (Kink) นอต (Knot) และการเปลี่ยนแปลงความชัน (Slope) ของการประมาณแนวโน้ม สามารถตีความได้ว่าเป็นการเปลี่ยนแปลงอย่างกระทันหันของข้อมูลอนุกรมเวลา การปรับข้อมูลให้เรียบขึ้นอยู่กับลักษณะของข้อมูล ตัวแบบที่ใช้ในการประมาณค่าและวิธีการปรับข้อมูล ซึ่งนักสถิติหลายท่านได้ศึกษาการประมาณค่าการถดถอยไม่อิงพารามิเตอร์โดยใช้วิธีการปรับข้อมูลให้เรียบวิธีต่าง ๆ เพื่อหาตัวแบบที่เหมาะสมกับข้อมูลมากที่สุด

นอกจากนี้ยังมีงานวิจัยที่ใช้วิธีดังกล่าวข้างต้นมาหาตัวแบบที่เหมาะสม ได้แก่ Cao และคณะ (2010) ได้ศึกษาการประมาณค่าเส้นโค้งที่ถูกลงโทษ (Penalized Spline) สำหรับตัวแบบการถดถอยสัมประสิทธิ์เชิงฟังก์ชัน เปรียบเทียบวิธีเส้นโค้งที่ถูกลงโทษและวิธีเส้นโค้งพหุนาม (Polynomial Spline) กับข้อมูลจริงการเติบโตของมูลค่าผลิตภัณฑ์มวลรวมประชาชาติที่แท้จริง (US real GNP growth) ของอเมริกา โดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนทำนายสัมบูรณ์ (Average Absolute Prediction Error: AAPE) ที่ต่ำสุดช่วยในการหาตัวแบบที่เหมาะสมกับข้อมูล ซึ่งพบว่าวิธีเส้นโค้งที่ถูกลงโทษมีค่า AAPE ต่ำสุด สามารถช่วยปรับข้อมูลที่แตกต่างกันให้ราบเรียบได้อย่างง่ายดายเหมาะสมกับข้อมูล

Tibshirani (2014) ได้เสนอการประมาณค่าพหุนามเป็นช่วงแบบปรับได้ (Adaptive Piecewise Polynomial) ผ่านวิธีกรองแนวโน้ม โดยทำการเปรียบเทียบการประมาณค่าด้วยวิธีเส้นโค้งเรียบ วิธีกรองแนวโน้ม และวิธีการถดถอยเส้นโค้งเชิงพื้นที่แบบปรับได้ (Locally Adaptive Regression Spline) จากการเปรียบเทียบทั้ง 3 วิธีในฟังก์ชันฮิลส์ (Hills Function) และฟังก์ชันดอปเพลอร์ (Doppler Function) พบว่าการประมาณค่าวิธีกรองแนวโน้มมีประสิทธิภาพดีกว่าวิธีเส้นโค้งเรียบและวิธีการถดถอยเส้นโค้งเชิงพื้นที่แบบปรับได้ เนื่องจากมีการปรับข้อมูลให้เรียบได้ดีกว่าทั้ง 2 วิธี โดยเฉพาะอย่างยิ่งในตัวแบบที่มีความซับซ้อนน้อย

Araveeporn (2015) ได้ศึกษาการเปรียบเทียบการประมาณค่าตัวแบบการถดถอยในตัวไม่เชิงเส้น (Nonlinear Autoregressive Model) โดยใช้วิธีเส้นโค้งเรียบ (Smoothing Spline) และวิธีเส้นโค้งที่ถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลงโทษ (Penalized Spline) กับข้อมูลอนุกรมเวลาที่มีลักษณะคงที่ (Stationary) และลักษณะไม่คงที่ (Non-Stationary) ซึ่งใช้เกณฑ์ค่าต่ำสุดของค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Square Error: AMSE) จากการจำลองข้อมูลพบว่าวิธีเส้นโค้งเรียบมีประสิทธิภาพดีที่สุดในทุกกรณีของข้อมูลที่มีลักษณะคงที่และไม่คงที่ เช่นเดียวกับการประยุกต์ใช้กับข้อมูลจริงของราคาทองคำ

Yang and Hong (2017) ได้เสนอเส้นโค้งที่ถูกลงโทษแบบปรับได้สำหรับข้อมูลปรับให้เรียบ โดยเปรียบเทียบวิธีการปรับให้เรียบ 6 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีกรองแนวโน้ม วิธีเส้นโค้งที่ถูกลงโทษแบบดั้งเดิม (Classical Penalized Spline) วิธีเส้นโค้งที่ถูกลงโทษโดยใช้ REML (Penalized Spline by REML) วิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้เร็วโดยใช้ REML (Fast Adaptive Penalized Spline (FAPS) by REML) และวิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้ (Adaptive Penalized Spline) ซึ่งทำการจำลองข้อมูลด้วยฟังก์ชันเม็กซิกัน แฮ็ท (Mexican Hat Function) ฟังก์ชันดอปเพลอร์ (Doppler Function) และฟังก์ชันบล็อก (Block Function) จากการเปรียบเทียบทั้ง 6 วิธี พบว่าวิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้เร็วโดยใช้ REML ไม่สามารถปรับองศาเสรี (Degrees of Freedom) ได้ จึงใช้การเปรียบเทียบทั้งหมด 5 วิธี โดยฟังก์ชันเม็กซิกัน แฮ็ท และฟังก์ชันดอปเพลอร์ ผลที่ได้คือวิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้มีความสามารถในการปรับข้อมูลได้ใกล้เคียงกับฟังก์ชันทั้งสองมากที่สุด และวิธีกรองแนวโน้มก็สามารถปรับข้อมูลได้เหมาะสมใกล้เคียงกับฟังก์ชันบล็อกได้ดีที่สุด

Araveeporn (2019) ได้ศึกษาเกี่ยวกับการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่อิงพารามิเตอร์ด้วยเทคนิคการปรับให้เรียบ 3 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งที่ถูกลงโทษ และวิธีเส้นโค้งบี วิธีการเหล่านี้ใช้พารามิเตอร์ปรับให้เรียบ (λ) เพื่อควบคุมประสิทธิภาพการปรับข้อมูลให้เรียบ จากการศึกษาพบว่าวิธีการทำให้เรียบทั้ง 3 วิธี มีประสิทธิภาพที่ดีกับข้อมูลที่ไม่เป็นเชิงเส้น (Non-Linear) จึงให้ค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ที่ต่ำสุด ช่วยในการหาตัวแบบที่เหมาะสม ซึ่งวิธีเส้นโค้งเรียบมีค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด จึงเป็นวิธีการปรับให้เรียบที่เหมาะสมกับชุดข้อมูลมากที่สุด

Airs และคณะ (2019) ได้ศึกษาการใช้ฟังก์ชันเส้นโค้งในโปรแกรมอาร์เพื่อวิเคราะห์การถดถอย ซึ่งโปรแกรมอาร์เป็นซอฟต์แวร์สถิติที่ได้รับความนิยมสูงโดยมีแพ็คเกจที่มีฟังก์ชันสไปลน์ที่ใช้ในการสร้างแบบจำลองเส้นโค้งได้หลายวิธี ได้แก่ วิธีเส้นโค้งบี วิธีเส้นโค้งกำลังสามธรรมชาติ และวิธีเส้นโค้งที่ถูกลงโทษ กับข้อมูลอายุและการวัดไขมันในร่างกาย (ความหนาของผิวหนัง) จากผู้หญิงที่มีอายุต่ำกว่า 50 ปี จำนวน 892 คน จาก 3 หมู่บ้านในแอฟริกาตะวันตก จากการศึกษาฟังก์ชันเส้นโค้งในโปรแกรมอาร์พบว่าจำนวนนอตและตำแหน่งของนอตคือสิ่งที่กำหนดความโค้งของเส้นโค้ง

จากงานวิจัยที่กล่าวมาข้างต้นมีทั้งการใช้ข้อมูลอนุกรมเวลาและข้อมูลที่มีลักษณะไม่เชิงเส้น โดยตัวแปรอิสระจะอยู่ในช่วงที่กำหนด แต่ในงานวิจัยนี้ผู้วิจัยสนใจศึกษาข้อมูลในรูปแบบอนุกรมเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านไป ลักษณะของการเปลี่ยนแปลงอาจมีหรือไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใด โดยความสัมพันธ์ของข้อมูลดูได้จากแผนภาพการกระจายของข้อมูลที่ใช้อาจแสดงในรูปแบบแนวโน้ม (Trend) วัฏจักร (Cycle) การเดินแบบสุ่ม (Random Walk) และข้อมูลที่ได้มาอาจเกิดสหสัมพันธ์ในตัว (Autocorrelation) เป็นภาวะที่ค่าสังเกตของอนุกรมเวลา ณ จุดเวลาที่อยู่ห่างกัน 2 จุดมีความสัมพันธ์กัน ก่อให้เกิดความคลาดเคลื่อนในการประมาณได้ ซึ่งการใช้วิธีการถดถอยไม่อิงพารามิเตอร์จะช่วยให้ความคลาดเคลื่อนในการพยากรณ์ค่าในอนาคตน้อยลงได้

จากงานวิจัยที่กล่าวมาข้างต้น จึงเสนอวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ วิธีเส้นโค้งเรียบ (Smoothing Spline Method) วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method) วิธีเส้นโค้งบี (B-Spline Method) และวิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method) ของข้อมูลอนุกรมเวลา เนื่องจากวิธีการเหล่านี้มีการใช้ค่าพารามิเตอร์ปรับให้เรียบและค่านอตมาใช้ในการประมาณค่าตัวแบบการถดถอยไม่อิงพารามิเตอร์ โดยกำหนดให้ตัวแปรอิสระคือลำดับข้อมูลของอนุกรมเวลา และตัวแปรตามคือข้อมูลอนุกรมเวลา โดยทำการเปรียบเทียบประสิทธิภาพของการประมาณค่า ด้วยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) และค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Square Error: AMSE) และเปรียบเทียบประสิทธิภาพการพยากรณ์ ด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) เพื่อหาตัวแบบปรับให้เรียบที่เหมาะสมกับข้อมูลมากที่สุด

1.2 วัตถุประสงค์

1.2.1 เพื่อศึกษาวิธีการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ วิธีเส้นโค้งเรียบ (Smoothing Spline Method) วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method) วิธีเส้นโค้งบี (B-Spline Method) และวิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)

1.2.2 เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ วิธีเส้นโค้งเรียบ (Smoothing Spline Method) วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method) วิธีเส้นโค้งบี (B-Spline Method) และวิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ขอบเขตของการศึกษา

การกำหนดขอบเขตของการศึกษา ดังนี้

1.3.1 กำหนดตัวแปรอิสระ (x) ให้เป็นลำดับของข้อมูลอนุกรมเวลาเริ่มตั้งแต่ 1,2,..., n เมื่อ n คือจำนวนข้อมูลทั้งหมด

1.3.2 ค่าความคลาดเคลื่อน (ξ) สุ่มจากการแจกแจงปกติ ซึ่งมีค่าเฉลี่ย $E(\xi)=\mu$ และความแปรปรวน $Var(\xi)=\sigma^2$ เขียนแทนด้วย $\xi \sim N(\mu,\sigma^2)$ โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น คือ

$$f(\varepsilon;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2}\left(\frac{\varepsilon-\mu}{\sigma}\right)^2\right] \quad ;-\infty<\varepsilon<\infty,-\infty<\mu<\infty,\ \sigma^2>0 \tag{1.1}$$

โดยกำหนดค่าเฉลี่ย (μ) ของความคลาดเคลื่อนเท่ากับ 0 และความแปรปรวน (σ^2) ของความคลาดเคลื่อนเท่ากับ 1, 9, 25 และ 49 หรือมีส่วนเบี่ยงเบนมาตรฐาน (σ) ของความคลาดเคลื่อน คือ 1, 3, 5 และ 7 เพื่อความสะดวกจะเรียกค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนว่า Standard Deviation หรือ SD

1.3.3 ตัวแบบการถดถอยไม่อิงพารามิเตอร์อย่างง่าย คือ

$$y_t=f(x_t)+\varepsilon_t \ , \ t=1,2,3,...,n \tag{1.2}$$

เมื่อ y_t คือตัวแปรตามที่ t และ $f(x_t)$ คือฟังก์ชันของตัวแปรอิสระ x_t

1.3.4 ข้อมูลของตัวแปรตามมีลักษณะเป็นข้อมูลอนุกรมเวลาทั้งในรูปแบบเชิงเส้นและไม่เชิงเส้น ดังนี้

1.3.4.1 ข้อมูลมีลักษณะเป็นแนวโน้ม (Trend) โดยตัวแปร

$$y_t=\sqrt{x_t}+\varepsilon_t \ , \ t=1,2,...,n$$

1.3.4.2 ข้อมูลมีลักษณะไม่เชิงเส้น (Non-Linear) โดยตัวแปร

$$y_t=3\sqrt{x_t}\sin\left(2\pi\left[\frac{1+30}{x_t+30}\right]\right)+\varepsilon_t \ , \ t=1,2,...,n$$

1.3.4.3 ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ (Heteroscedasticity) โดยตัวแปร

$$y_t=\sqrt{x_t}\sin\left(2\pi\left[1+\sqrt{x_t}\right]\right)+\varepsilon_t \ , \ t=1,2,...,n$$

1.3.4.4 ข้อมูลมีลักษณะไม่คงที่ (Non-Stationary) หรือการเดินแบบสุ่ม (Random Walk) โดยตัวแปร $y_t=y_{t-1}+\varepsilon_t \ , \ t=2,3,...,n$ โดยที่ $y_1=N(0,1)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3.4.5 ข้อมูลมีลักษณะคงที่ (Stationary) โดยตัวแปร

$$y_t = \exp\left[\sin\left(x_t - \frac{x_t}{1+x_t^2}\right)\right] + \varepsilon_t \ , t = 1,2,...,n$$

1.3.5 การกำหนดจำนวนนอต (Knot) เพื่อให้เกิดข้อต่อในการสร้างเส้นโค้งเชื่อมต่อ ๆ กันให้ใกล้ข้อมูลมากที่สุด โดยพิจารณาขอบเขตของนอตตามขนาดตัวอย่าง

1.3.6 ประมาณค่าพารามิเตอร์ปรับให้เรียบ (λ) เพื่อใช้ในการปรับเส้นโค้ง โดยจะกำจัดหรือลดการเปลี่ยนแปลงขึ้นลงของข้อมูลในฟังก์ชัน เพื่อให้การประมาณค่าแทรกไปตามเส้นโค้งของข้อมูลและปรับให้เรียบขึ้น โดยใช้วิธีการแบ่งข้อมูลเป็นหลายส่วน (Cross-validation Method) (Stone, 1974)

1.3.7 จำลองข้อมูลด้วยโปรแกรมอาร์ (R) เวอร์ชัน 4.1.1 โดยทำการทดลองซ้ำ 500 รอบ ในแต่ละสถานการณ์

1.3.8 กำหนดขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150 และ 200

1.3.9 ใช้ข้อมูลจริง 2 ชุด คือ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา (บาท/ดอลลาร์ สรอ.) (ธนาคารแห่งประเทศไทย, 2564) และปริมาณการนำเข้าไฟฟ้าของประเทศไทย (กิกะวัตต์ชั่วโมง) (การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย, 2564) เป็นข้อมูลอนุกรมเวลาที่มีลักษณะเป็นแนวโน้ม (Trend) ไม่เชิงเส้น (Non-Linear) เป็นฤดูกาล (Seasonal) และข้อมูลมีลักษณะไม่คงที่ (Non-Stationary) หรือการเดินแบบสุ่ม (Random Walk)

1.4 เกณฑ์การตัดสินใจ

การเปรียบเทียบประสิทธิภาพของการประมาณค่าพารามิเตอร์ปรับให้เรียบด้วยวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ จำเป็นต้องมีเกณฑ์ตัดสินใจในการเลือกวิธีการประมาณค่าพารามิเตอร์ปรับให้เรียบให้เหมาะสมที่สุดในแต่ละฟังก์ชัน แสดงรายละเอียดดังต่อไปนี้

1.4.1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) โดยในการทดลองจำนวนรอบของการหาค่า MSE เท่ากับจำนวนรอบของการทำซ้ำ (MSE_h) ดังสมการที่ (1.3)

$$MSE_h = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \ , h = 1,2,...,m; t = 1,2,...,n \qquad (1.3)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ y_t คือ ค่าจริงของตัวแปรตามที่ t

$\hat{y}_t$ คือ ค่าประมาณของตัวแปรตามที่ t ที่ได้จากตัวแบบการถดถอยไม่อิงพารามิเตอร์

n คือ ขนาดตัวอย่าง

m คือ จำนวนรอบของการทดลองซ้ำ

1.4.2 ค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Square Error: AMSE) ดังสมการที่ (1.4)

$$AMSE = \frac{1}{m}\sum_{h=1}^{m} MSE_h,\ h = 1,2,...,m \tag{1.4}$$

การประมาณค่าพารามิเตอร์ปรับให้เรียบ จากการทดลองซ้ำ 500 ครั้ง ที่ให้ค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดเมื่อเทียบกับวิธีอื่น ๆ จะถือว่าเป็นวิธีการประมาณค่าพารามิเตอร์ปรับให้เรียบที่เหมาะสมและมีประสิทธิภาพดีที่สุด

1.4.3 ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) ดังสมการที่ (1.5)

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \frac{|y_t - \hat{y}_t|}{y_t} \times 100, t = 1,2,...,n \tag{1.5}$$

การพยากรณ์ตัวแบบจากข้อมูลจริงที่ให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) มีค่าต่ำที่สุดเมื่อเทียบกับตัวแบบอื่น ๆ จะถือว่าการพยากรณ์ของข้อมูลนั้นเหมาะสมกับข้อมูลมากที่สุด

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 เพื่อทราบประสิทธิภาพของวิธีการประมาณค่าการถดถอยไม่อิงพารามิเตอร์ด้วยวิธีการปรับให้เรียบของข้อมูลอนุกรมเวลาในหลาย ๆ รูปแบบ

1.5.2. เพื่อเป็นแนวทางในการเลือกใช้วิธีการประมาณค่าการถดถอยไม่อิงพารามิเตอร์ด้วยวิธีการปรับให้เรียบของข้อมูลอนุกรมเวลาได้อย่างเหมาะสม

1.5.3 เพื่อใช้เป็นแนวทางในการศึกษาวิจัยการวิเคราะห์การถดถอยไม่อิงพารามิเตอร์ด้วยวิธีการปรับให้เรียบวิธีอื่น ๆ ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6 นิยามศัพท์

1.6.1 อนุกรมเวลา คือ เซตของข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่งตามลำดับเวลาการเกิด โดยทั่วไปมักมีระยะห่างของเวลาเท่า ๆ กัน เช่น รายวัน รายเดือน รายไตรมาส รายปี มักนำเสนอโดยใช้แผนภาพเส้น และมีองค์ประกอบ 4 ส่วน คือ แนวโน้ม (Trend) ฤดูกาล (Seasonal) วัฏจักร (Cycle) และเหตุการณ์ที่ผิดปกติหรือเหตุการณ์ความไม่แน่นอน (Irregular)

1.6.2 การทำให้เรียบ (Smoothing) คือ กระบวนการกำจัดหรือลดการเปลี่ยนแปลงขึ้นลงในข้อมูลเพื่อให้เห็นลักษณะข้อมูลที่เรียบขึ้น (พจนานุกรม ศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสภา, 2561)

1.6.3 สหสัมพันธ์ในตัว (Autocorrelation) ภาวะที่ค่าสังเกตของอนุกรมเวลา ณ จุดเวลาที่อยู่ห่างกัน 2 จุด มีความสัมพันธ์กัน (พจนานุกรม ศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสภา, 2561)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะแบ่งเนื้อหาออกเป็น 2 ส่วนหลัก ๆ คือ ส่วนของทฤษฎีที่เกี่ยวข้อง และงานวิจัยที่เกี่ยวข้อง ซึ่งเป็นแนวทางสำคัญสำหรับวิธีการดำเนินงานวิจัย แสดงรายละเอียดดังนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

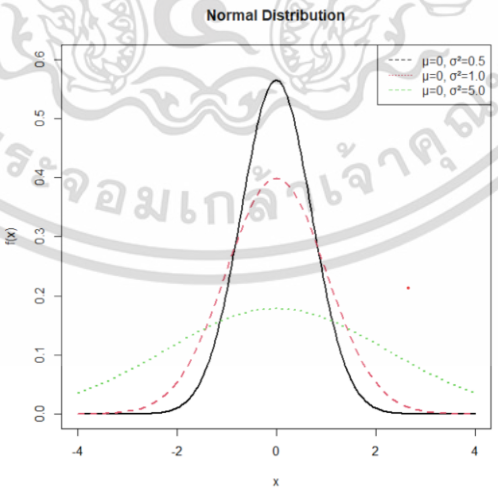
2.1.1 การแจกแจงปรกติ (Normal Distribution)

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มต่อเนื่อง X ที่มีการแจกแจงปรกติ มีลักษณะสมมาตรรอบค่าเฉลี่ย μ และความแปรปรวน σ^2

โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็นคือ

$$f(x,\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad ;-\infty<x<\infty,-\infty<\mu<\infty,\sigma^2>0 \qquad (2.1)$$

สามารถเขียนแทนด้วย $X\sim N(\mu,\sigma^2)$ เช่น การแจกแจงปรกติที่มีค่าเฉลี่ยเท่ากันและความแปรปรวนไม่เท่ากัน คือ $N(0,0.5)$, $N(0,1)$ และ $N(0,5)$ แสดงดังรูปที่ 2.1



รูปที่ 2.1 การแจกแจงปรกติที่มีค่าเฉลี่ยเท่ากันและความแปรปรวนไม่เท่ากัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 ตัวแบบการถดถอยไม่อิงพารามิเตอร์

การถดถอยไม่อิงพารามิเตอร์ คือการกำหนดความสัมพันธ์ระหว่างฟังก์ชันของตัวแปรอิสระ ($f(x_t)$) และตัวแปรตาม (y_t) Wood (2002) ได้กำหนดความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามในรูปแบบของตัวแบบการถดถอยไม่อิงพารามิเตอร์เป็น

$$y_t = f(x_t) + \varepsilon_t \ , \ t = 1,2,3,...,n \tag{2.2}$$

เมื่อ ε_t คือ ค่าความคลาดเคลื่อนสุ่มของค่าสังเกต

2.1.3 การถดถอยไม่อิงพารามิเตอร์

การถดถอยไม่อิงพารามิเตอร์นั้นใช้เทคนิคการปรับให้เรียบซึ่งจะทำให้เกิดความราบเรียบขึ้น ความราบเรียบคือเครื่องมือสำหรับการประมาณค่าฟังก์ชันของตัวแปรอิสระ ที่สามารถนำมาใช้ทำให้แนวโน้มในกราฟมีลักษณะที่ดีขึ้น

Eubank (1988, 1999) นำเสนอหลักการของการถดถอยไม่อิงพารามิเตอร์หรือการถดถอยสไปลน์ โดยพิจารณาขอบเขตในกลุ่มของค่า

$$\tau_0, \tau_1, \tau_2, ..., \tau_K, \tau_{K+1} \tag{2.3}$$

ซึ่งจะอยู่ในช่วง $[a,b]$ ที่ขอบเขต $a = \tau_0 < \tau_1 < ... < \tau_K < \tau_{K+1} < b$ โดยกำหนดค่าขอบเขต เรียกว่าค่านอต (Knot) และ $\tau_r, r = 1,2,...,K$ เรียกว่าค่านอตภายในหรือค่านอตเริ่มต้น

การถดถอยสไปลน์สามารถสร้างโดยใช้การยกกำลัง k กับค่า K นอต $\tau_1, \tau_2, ..., \tau_K$

$$\Phi_p(t) = 1, x_t, ..., x_t^k, (x_t - \tau_1)_+^k, ..., (x_t - \tau_K)_+^k, p = K + k + 1 \tag{2.4}$$

เมื่อ w_+^k คือ การยกกำลัง k ของค่าที่เป็นบวกของ w โดยที่ $w_+ = \max(0, w)$ ฟังก์ชันพื้นฐานยกกำลังของ $(k+1)$ ในสมการที่ (2.4) เป็นการใช้การยกกำลังของการแบ่งค่า ซึ่งสามารถนำไปใช้เขียนฟังก์ชันการถดถอยแบบสไปลน์ได้ ดังนี้

$$f(x_t) = \sum_{s=0}^{k} \beta_s x_t^s + \sum_{r=1}^{K} \beta_{k+r} (x_t - \tau_r)_+^k \tag{2.5}$$

เมื่อ $\beta_0, \beta_1, ..., \beta_{k+K}$ เป็นค่าสัมประสิทธิ์ที่ต้องการประมาณเพื่อให้ได้ตัวประมาณที่เหมาะสมโดยใช้วิธีทำให้ฟังก์ชันสูญเสียมีค่าต่ำสุด เทคนิคการทำให้เรียบที่เป็นที่นิยมสามารถแสดงรายละเอียดได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.3.1 วิธีเส้นโค้งเรียบ (Smoothing Spline Method)

การประมาณค่าของวิธีเส้นโค้งเรียบ เป็นการประมาณค่าจากการทำให้เกิดค่าต่ำสุดของฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุด (Penalized Least Square: PLS) ให้เหมาะสมกับฟังก์ชันของตัวแปรอิสระ $\left(f\left(x_t\right)\right)$ ดังนี้

$$PLS=\sum_{t=1}^{n}\left\{y_t-f\left(x_t\right)\right\}^2+\lambda\int_a^b\left\{f''\left(x_t\right)\right\}^2dx \tag{2.6}$$

เมื่อ $RSS=\sum_{t=1}^{n}\left\{y_t-f\left(x_t\right)\right\}^2$ คือ ผลบวกกำลังสองของส่วนเหลือและ $\lambda\int_a^b\left\{f''\left(x_t\right)\right\}^2dx$ คือ ความหยาบที่ถูกลงโทษ (Roughness Penalty) ในช่วง $[a,b]$ โดยช่วงนี้จะขึ้นอยู่เป็นช่วงจำกัดตามตัวแปรอิสระที่เกิดเส้นโค้ง โดย λ คือ ค่าพารามิเตอร์ปรับให้เรียบ หรือ λ การอนุพันธ์อันดับสองของความหยาบที่ถูกลงโทษ เรียกว่าเส้นโค้งเรียบกำลังสาม (Cubic Smoothing Spline) (Nadaraya, 1964) จะเป็นค่าคงที่เริ่มที่ศูนย์

Wu and Zhang (2006) แสดงการเขียนความหยาบที่ถูกลงโทษในรูปแบบเมทริกซ์ ได้ดังนี้

$$\lambda\int_a^b\left\{f''\left(x_t\right)\right\}^2dx=\mathbf{f}^T\mathbf{G}\mathbf{f} \tag{2.7}$$

โดยกำหนดให้ $\mathbf{f}=(f_1,...,f_K)^T$ เมื่อ $f_r=f(\tau_r),r=1,2,...,K$ โดยทั่วไป K คือจำนวนนอต และ $\tau_1,...,\tau_K$ เป็นนอตทั้งหมดของเส้นโค้งเรียบที่สามารถจัดเรียงตามลำดับที่เพิ่มขึ้นเป็น

$$-\infty\le a<\tau_1<\tau_2<...<\tau_K<b\le\infty$$

กำหนดให้ $h_r=\tau_{r-1}-\tau_r$, $r=1,2,...,K-1$ โดยเมทริกซ์ G มีขนาด $K\times K$ ดังนั้น

$$\mathbf{G}=\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^T \tag{2.8}$$

เมื่อเมทริกซ์ $\mathbf{A}=(a_{rs})$ มีขนาด $K\times(K-2)$ โดยค่าอื่น ๆ จะมีค่าเป็น 0 ทั้งหมดยกเว้นที่ $r=1,2,...,K-2,a_{rr}=h_r^{-1}$ และ $a_{r+1,r}=-(h_r^{-1}+h_{r+1}^{-1})$, $a_{r+2,r}=h_{r+1}^{-1}$

เมื่อเมทริกซ์ $\mathbf{B}=(b_{rs})$ มีขนาด $(K-2)\times(K-2)$ ค่าอื่น ๆ จะมีค่าเป็น 0 ยกเว้น $b_{11}=(h_1+h_2)/3$, $b_{21}=h_2/6$ เมื่อ $r=1,2,...,K-4,b_{r,r+1}=h_{(r+1)}/6$, $b_{r+1,r+1}=(h_{(r+1)}+h_{(r+2)})/3$ $b_{r+2,r+1}=h_{(r+2)}/6$ และ $b_{K-3,K-2}=h_{(K-2)}/6$, $b_{K-2,K-2}=(h_{(K-2)}+h_{(K-2)})/3$

ซึ่งเป็นไปตามเกณฑ์ฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุดจากสมการที่ (2.7) สามารถเขียนในรูปเมทริกซ์ได้ดังนี้

$$\left\|\mathbf{y}-\mathbf{W}\mathbf{f}\right\|^2+\lambda\mathbf{f}^T\mathbf{G}\mathbf{f} \tag{2.9}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $\mathbf{y}=(y_1,...,y_n)^T$ คือเวกเตอร์ของตัวแปรตาม $\mathbf{W}=(w_{tr})$ มีขนาด $n\times K$ เป็นเมทริกซ์ตกกระทบ (Incidence Matrix) กับ $w_{tr}=1$ ถ้า $x_t=\tau_r$ และ $\|\mathbf{y}-\mathbf{W}\mathbf{f}\|^2=\sum_{t=1}^{n}\{y_t-f(x_t)\}^2$ ดังนั้นฟังก์ชันเส้นโค้งเรียบ ($\hat{\mathbf{f}}_\lambda$) จากนอต $\tau_r, r=1,...,K$ คือ

$$\hat{\mathbf{f}}_\lambda=(\mathbf{W}^T\mathbf{W}+\lambda\mathbf{G})^{-1}\mathbf{W}^T\mathbf{y} \tag{2.10}$$

ในการประมาณค่าจะใช้ฟังก์ชัน smooth.spline ในแพ็คเกจ splines ของโปรแกรมอาร์

2.1.3.2 วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)

Hastie (1990) ศึกษาตัวแบบเชิงบวกนัยทั่วไป (Generalized Additive Model) เป็นวิธีที่สร้างตัวแบบเชิงบวกนัยทั่วไปโดยประยุกต์กับฟังก์ชันปรับให้เรียบซึ่งจะประมาณค่าโดยเส้นโค้งการถดถอย สำหรับการประมาณค่าเส้นโค้งกำลังสามธรรมชาติ

วิธีการสร้างเส้นโค้งกำลังสามธรรมชาติสร้างขึ้นเพื่อลดความไม่แน่นอนนี้ในการวางตัวเป็นเส้นตรงในขอบเขตของนอต โดยมีการกำหนดเงื่อนไข

$$N_1(x_t)=1,\ N_1(x_t)=x_t$$
$$N_{k+2}(x_t)=d_k(x_t)-d_{K-1}(x_t),\ \ r=1,...,K-2$$
$$d_k(x_t)=\frac{(x_t-\tau_r)_+^3-(x_t-\tau_K)_+^3}{\tau_r-\tau_K}$$

โดยสัญลักษณ์ w_+^k เมื่อ k แทนลำดับการยกกำลังของ w และ $w_+=\max(0,w)$ ซึ่งสามารถเขียนฟังก์ชันของตัวแปรอิสระได้ดังนี้

$$f(x_t)=\sum_{r=1}^{K}N_r(x_t)\,\theta_r \tag{2.11}$$

เมื่อ $N_r(x_t)$ เป็นชุดฟังก์ชันขนาด N ของเส้นโค้งกำลังสามธรรมชาติ โดยที่ $\boldsymbol{\theta}=(\theta_1,...,\theta_K)^T$ คือเวกเตอร์สัมประสิทธิ์ของเส้นโค้งกำลังสามธรรมชาติ $N_r(x_t)$ กับจำนวนนอต ซึ่งเป็นค่าที่มีความเหมาะสมของฟังก์ชันเส้นโค้งกำลังสามธรรมชาติ ที่ค่านอต $\tau_r, r=1,...,K$ โดยค่านอต τ_r จะแทนค่าด้วยจำนวนเต็มไปจนถึงจำนวนของขนาดตัวอย่าง ซึ่งสามารถมีได้ K ตัว ดังสมการ

$$\hat{f}(x_t)=\sum_{r=1}^{K}N_r(x_t)\,\hat{\theta}_r \tag{2.12}$$

ค่าสัมประสิทธิ์สามารถหาได้จากเกณฑ์เกิดค่าต่ำสุดของฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุดจากสมการที่ (2.6) ดังนี้

$$PLS=(\mathbf{y}-\mathbf{N}\boldsymbol{\theta})^T(\mathbf{y}-\mathbf{N}\boldsymbol{\theta})+\lambda\boldsymbol{\theta}\Omega_K\boldsymbol{\theta}$$

เมื่อ $\mathbf{y}=(y_1,...,y_n)^T, \{\mathbf{N}\}_{tr}=N_r(x_t)$ และ $\{\Omega_K\}_{tr}=\int N_t''(x)N_r''(x)dx$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นตัวประมาณของฟังก์ชันเส้นโค้งกำลังสามธรรมชาติ $\left(\hat{\mathbf{f}}_{\lambda}\right)$ ที่ค่านอต $\tau_r, r=1,2,...,K$ โดยขึ้นอยู่กับพารามิเตอร์ปรับให้เรียบดังต่อไปนี้

$$\hat{\mathbf{f}}_{\lambda}=(\mathbf{N}^T\mathbf{N}+\lambda\Omega_K)^{-1}\mathbf{N}^T\mathbf{y} \tag{2.13}$$

ในการประมาณค่าจะใช้ฟังก์ชัน ns ในแพ็คเกจ splines ของโปรแกรมอาร์

2.1.3.3 วิธีเส้นโค้งบี (B-Spline Method)

สไปลน์ (Spline) เป็นตัวแบบพหุนามแบบแบ่งส่วนหรือเรียกว่าพหุนามทีละส่วน (Piecewise Polynomial) ซึ่งเป็นชิ้นส่วนของพหุนามที่มีคุณสมบัติแบ่งส่วนเป็นช่วงบนจุดนอต (De Boor, 1978) โดยจุดนอต คือจุดที่บอกถึงการเปลี่ยนแปลงของข้อมูลในช่วงย่อย ๆ นั้น การให้ชุดของนอต K เป็นฟังก์ชันพื้นฐานเส้นโค้งบีซ้ำ ๆ สามารถกำหนดได้โดยแสดงดังนี้

$$B_r^m(x) = \begin{cases} 1 & , \tau_r < x < \tau_{r+1} \\ 0 & \text{อื่นๆ} \end{cases} \tag{2.14}$$

เมื่อ $B_r^m(x)$ คือ ลำดับที่ r ของฟังก์ชันเส้นโค้งบีของ m สำหรับค่านอต τ

De Boor (1978) ได้แนะนำอัลกอริทึมในการคำนวณเส้นโค้งบีของฟังก์ชันพหุนามทีละส่วน ลำดับที่ m ของเส้นโค้งบีจะคำนวณจากลำดับที่ $(m-1)$ ได้ดังนี้

$$B_r^m(x_t) = \frac{x_t-\tau_r}{\tau_{r+m-1}-\tau_r}B_r^{m-1}+\frac{\tau_{r+m}-x_t}{\tau_{r+m}-\tau_{r+1}}B_{r+1}^{m-1} \tag{2.15}$$

เมื่อลำดับที่ m พิจารณาจากเซตของ $\left\{B_r^m \mid r=1,...,K+m+1\right\}$ ดังนั้น ฟังก์ชันของตัวแปรอิสระของวิธีเส้นโค้งบีสามารถเขียนได้เป็น

$$f(x_t) = \sum_{r=1}^{K} B_r^m(x_t)\theta_r \tag{2.16}$$

ซึ่งค่าที่มีความเหมาะสมของฟังก์ชันเส้นโค้งบีที่ประมาณที่ค่านอต $\tau_r, r=1,2,...,K$ เป็นดังนี้

$$\hat{f}(x_t) = \sum_{r=1}^{K} B_r^m(x_t)\hat{\theta}_r \quad . \tag{2.17}$$

ค่าสัมประสิทธิ์สามารถหาได้จากเกณฑ์เกิดค่าต่ำสุดของฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุดจากสมการที่ (2.6) ดังนี้

$$PLS=(\mathbf{y}-\mathbf{B}\boldsymbol{\theta})^T(\mathbf{y}-\mathbf{B}\boldsymbol{\theta})+\lambda\boldsymbol{\theta}\Omega_K\boldsymbol{\theta}$$

เมื่อ $\{\mathbf{B}\}_{tr}=B_r^m(x_t), \{\Omega_K\}_{tr}=\int B_t^{''}(x)B_r^{''}(x)dx$ และ $\boldsymbol{\theta}=(\theta_1,...,\theta_K)^T$ คือเวกเตอร์สัมประสิทธิ์ของเส้นโค้งบี ดังนั้น ตัวประมาณของฟังก์ชันเส้นโค้งบี $\left(\hat{\mathbf{f}}_{\lambda}\right)$ คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\hat{\mathbf{f}}_{\lambda} = (\mathbf{B}^T\mathbf{B} + \lambda\Omega_K)^{-1}\mathbf{B}^T\mathbf{y} \quad (2.18)$$

ในการประมาณค่าจะใช้ฟังก์ชัน bs ในแพ็คเกจ splines ของโปรแกรมอาร์

2.1.3.4 วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)

วิธีเส้นโค้งเรียบจำเป็นต้องกำหนดอินทิกรัลเพื่อคำนวณอินทิกรัลของฟังก์ชันความหยาบ วิธีเส้นโค้งที่ถูกลงโทษได้รับการออกแบบมาเพื่อเอาชนะข้อเสียเหล่านี้โดยใช้พื้นฐานการแบ่งค่าแบบยกกำลัง (Truncated Power Basis) ในสมการ (2.4) ให้ $\mathbf{\Phi}_r(t) = (\phi_1(t),...,\phi_r(t))^T$ แสดงถึงพื้นฐานการแบ่งค่าแบบยกกำลังของระดับ k กับ K นอต $\tau_1, \tau_2, ... < \tau_K$ จากนั้นสามารถแสดง $f(x_t)$ ในสมการ (2.2) เป็น $\mathbf{\Phi}_r(t)^T\ \boldsymbol{\alpha}$ เมื่อ $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, ..., \alpha_{k+K}]^T$ เป็นเวกเตอร์สัมประสิทธิ์ที่เกี่ยวข้อง ให้ G เป็นเมทริกซ์แนวทแยง $p \times p$ ที่มีรายการแนวทแยงแรก $k+1$ เป็น 0 และรายการแนวทแยงอื่นๆ เป็น 1 นั่นคือ

$$\mathbf{G} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_K \end{pmatrix}$$

จากนั้นการปรับเส้นโค้งที่ถูกลงโทษกำหนดเป็น $\hat{\mathbf{f}}_{\lambda} = \mathbf{\Phi}_r(t)^T\ \hat{\boldsymbol{\alpha}}$ เมื่อ $\hat{\boldsymbol{\alpha}}$ โดยที่ตัวย่อของเกณฑ์การลดขนาดกำลังสองน้อยที่สุดที่ถูกลงโทษ ดังนี้

$$PLS = (\mathbf{y} - \mathbf{W}\boldsymbol{\alpha})^T(\mathbf{y} - \mathbf{W}\boldsymbol{\alpha}) + \lambda\boldsymbol{\alpha}\mathbf{G}\boldsymbol{\alpha} \quad (2.19)$$

เมื่อ $\mathbf{W} = (\mathbf{\Phi}_r(x_1),...,\mathbf{\Phi}_r(x_n))^T$ และ $\boldsymbol{\alpha}\mathbf{G}\boldsymbol{\alpha} = \sum_{r=1}^{K}\alpha_{k+r}^2$. การปรับเส้นโค้งที่ถูกลงโทษสามารถแสดงได้ดังนี้

$$\hat{\mathbf{f}}_{\lambda} = \mathbf{W}(\mathbf{W}^T\mathbf{W} + \lambda\mathbf{G})^{-1}\mathbf{W}^T\mathbf{y} \quad (2.20)$$

ในการประมาณค่าจะใช้ฟังก์ชัน spm ในแพ็คเกจ SemiPar ของโปรแกรมอาร์

2.1.4 พารามิเตอร์ปรับให้เรียบ

พารามิเตอร์ปรับให้เรียบ (λ) (Smoothing Parameter) เป็นพารามิเตอร์ที่ใช้ในการปรับเส้นโค้ง ซึ่งจะกำจัดหรือลดการเปลี่ยนแปลงขึ้นลงของข้อมูลในฟังก์ชัน ให้การประมาณค่าแทรกไปตามเส้นโค้งของข้อมูลและปรับให้ข้อมูลเรียบขึ้น โดยใช้วิธีการแบ่งข้อมูลเป็นหลายส่วน (Cross-Validation Method) (Stone, 1974) เป็นเกณฑ์ในการทดสอบประสิทธิภาพของตัวแบบให้เหมาะสมกับข้อมูล ได้ฟังก์ชันดังนี้

$$CV(\lambda) = \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{y_t - f_{\lambda}(x_t)}{1-(S_{\lambda})_{tt}}\right\}^2 \quad (2.21)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $(S_\lambda)_{tt}$ คือลำดับแนวทะแยงของ S_λ และจะเลือกพารามิเตอร์ปรับให้เรียบที่ทำให้ค่า $CV(\lambda)$ มีค่าต่ำสุด (Lee, 2003)

2.2 งานวิจัยที่เกี่ยวข้อง

การปรับข้อมูลให้เรียบขึ้นอยู่กับลักษณะของข้อมูล ตัวแบบที่ใช้ในการประมาณค่าและวิธีการปรับข้อมูล ซึ่งนักสถิติหลายท่านได้ศึกษาการประมาณค่าการถดถอยไม่อิงพารามิเตอร์โดยใช้วิธีการปรับข้อมูลให้เรียบวิธีต่าง ๆ เพื่อหาตัวแบบที่เหมาะสมกับข้อมูลมากที่สุด

Cao และคณะ (2010) ได้ศึกษาการประมาณค่าเส้นโค้งที่ถูกลงโทษ (Penalized Spline) สำหรับตัวแบบการถดถอยสัมประสิทธิ์เชิงฟังก์ชัน เปรียบเทียบวิธีเส้นโค้งที่ถูกลงโทษและวิธีเส้นโค้งพหุนาม (Polynomial Spline) กับข้อมูลจริงการเติบโตของมูลค่าผลิตภัณฑ์มวลรวมประชาชาติที่แท้จริง (US real GNP growth) ของอเมริกา โดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนทำนายสัมบูรณ์ (Average Absolute Prediction Error: AAPE) ที่ต่ำสุดช่วยในการหาตัวแบบที่เหมาะสมกับข้อมูล ซึ่งพบว่าวิธีเส้นโค้งที่ถูกลงโทษมีค่า AAPE ต่ำสุด สามารถช่วยปรับข้อมูลที่แตกต่างกันให้ราบเรียบได้อย่างง่ายดายเหมาะสมกับข้อมูล

Tibshirani (2014) ได้เสนอการประมาณค่าพหุนามเป็นช่วงแบบปรับได้ (Adaptive Piecewise Polynomial) ผ่านวิธีกรองแนวโน้ม โดยทำการเปรียบเทียบการประมาณค่าด้วยวิธีเส้นโค้งเรียบ วิธีกรองแนวโน้ม และวิธีการถดถอยเส้นโค้งเชิงพื้นที่แบบปรับได้ (Locally Adaptive Regression Spline) จากการเปรียบเทียบทั้ง 3 วิธีในฟังก์ชันฮิลส์ (Hills Function) และฟังก์ชันดอปเพลอร์ (Doppler Function) พบว่าการประมาณค่าวิธีกรองแนวโน้มมีประสิทธิภาพดีกว่าวิธีเส้นโค้งเรียบและวิธีการถดถอยเส้นโค้งเชิงพื้นที่แบบปรับได้ เนื่องจากมีการปรับข้อมูลให้เรียบได้ดีกว่าทั้ง 2 วิธี โดยเฉพาะอย่างยิ่งในตัวแบบที่มีความซับซ้อนน้อย

Araveeporn (2015) ได้ศึกษาการเปรียบเทียบการประมาณค่าตัวแบบการถดถอยในตัวไม่เชิงเส้น (Nonlinear Autoregressive Model) โดยใช้วิธีเส้นโค้งเรียบ (Smoothing Spline) และวิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline) กับข้อมูลอนุกรมเวลาที่มีลักษณะคงที่ (Stationary) และลักษณะไม่คงที่ (Non-stationary) ซึ่งใช้เกณฑ์ค่าต่ำสุดของค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Square Error: AMSE) จากการจำลองข้อมูลพบว่าวิธีเส้นโค้งเรียบมีประสิทธิภาพดีที่สุดในทุกกรณีของข้อมูลที่มีลักษณะคงที่และไม่คงที่ เช่นเดียวกับการประยุกต์ใช้กับข้อมูลจริงของราคาทองคำ

Yang and Hong (2017) ได้เสนอเส้นโค้งที่ถูกลงโทษแบบปรับได้สำหรับข้อมูลปรับให้เรียบโดยเปรียบเทียบวิธีการปรับให้เรียบ 6 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีกรองแนวโน้ม วิธีเส้นโค้งที่ถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลงโทษแบบดั้งเดิม (Classical Penalized Spline) วิธีเส้นโค้งที่ถูกลงโทษโดยใช้ REML (Penalized Spline by REML) วิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้เร็วโดยใช้ REML (Fast Adaptive Penalized Spline (FAPS) by REML) และวิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้ (Adaptive Penalized Spline) ซึ่งทำการจำลองข้อมูลด้วยฟังก์ชันเม็กซิกัน แฮ็ท (Mexican Hat Function) ฟังก์ชันดอปเพลอร์ (Doppler Function) และฟังก์ชันบล็อก (Block Function) จากการเปรียบเทียบทั้ง 6 วิธี พบว่าวิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้เร็วโดยใช้ REML ไม่สามารถปรับองศาเสรี (Degrees of Freedom) ได้ จึงใช้การเปรียบเทียบทั้งหมด 5 วิธี โดยฟังก์ชันเม็กซิกัน แฮ็ท และฟังก์ชันดอปเพลอร์ ผลที่ได้คือวิธีเส้นโค้งที่ถูกลงโทษแบบปรับได้มีความสามารถในการปรับข้อมูลได้ใกล้เคียงกับฟังก์ชันทั้งสองมากที่สุด และวิธีกรองแนวโน้มก็สามารถปรับข้อมูลได้เหมาะสมใกล้เคียงกับฟังก์ชันบล็อกได้ดีที่สุด

Araveeporn (2019) ได้ศึกษาเกี่ยวกับการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่อิงพารามิเตอร์ด้วยเทคนิคการปรับให้เรียบ 3 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งที่ถูกลงโทษ และวิธีเส้นโค้งบี วิธีการเหล่านี้ใช้พารามิเตอร์ปรับให้เรียบ (λ) เพื่อควบคุมประสิทธิภาพการปรับข้อมูลให้เรียบ จากการศึกษาพบว่าวิธีการทำให้เรียบทั้ง 3 วิธี มีประสิทธิภาพที่ดีกับข้อมูลที่ไม่เป็นเชิงเส้น (Non-linear) จึงให้ค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ที่ต่ำสุดช่วยในการหาตัวแบบที่เหมาะสม ซึ่งวิธีเส้นโค้งเรียบมีค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด จึงเป็นวิธีการปรับให้เรียบที่เหมาะสมกับชุดข้อมูลมากที่สุด

Airs และคณะ (2019) ได้ศึกษาการใช้ฟังก์ชันเส้นโค้งในโปรแกรมอาร์เพื่อวิเคราะห์การถดถอย ซึ่งโปรแกรมอาร์เป็นซอฟต์แวร์สถิติที่ได้รับความนิยมสูงโดยมีแพ็คเกจที่มีฟังก์ชันสไปลน์ที่ใช้ในการสร้างแบบจำลองเส้นโค้งได้หลายวิธี ได้แก่ วิธีเส้นโค้งบี วิธีเส้นโค้งกำลังสามธรรมชาติ และวิธีเส้นโค้งที่ถูกลงโทษ กับข้อมูลอายุและการวัดไขมันในร่างกาย (ความหนาของผิวหนัง) จากผู้หญิงที่มีอายุต่ำกว่า 50 ปี จำนวน 892 คน จาก 3 หมู่บ้านในแอฟริกาตะวันตก จากการศึกษาฟังก์ชันเส้นโค้งในโปรแกรมอาร์พบว่าจำนวนนอตและตำแหน่งของนอตคือสิ่งที่กำหนดความโค้งของเส้นโค้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงานวิจัย

ในบทนี้จะกล่าวถึงขั้นตอนที่แสดงรายละเอียดสำหรับการวัดประสิทธิภาพของการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์ โดยศึกษาวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธีคือ

1. วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
2. วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
3. วิธีเส้นโค้งบี (B-Spline Method)
4. วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)

โดยแสดงรายละเอียดการวางแผนการวิจัยและวิธีดำเนินการวิจัยในหัวข้อถัดไป

3.1 การวางแผนการวิจัย

การวิจัยในครั้งนี้ศึกษาประสิทธิภาพของการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์แต่ละวิธี โดยศึกษาในเชิงเปรียบเทียบที่ดีที่สุดเพียงวิธีเดียวเท่านั้นในแต่ละสถานการณ์ที่ศึกษาดังนี้

3.1.1 กำหนดขนาดตัวอย่าง

3.1.2 กำหนดวิธีการถดถอยไม่อิงพารามิเตอร์

3.1.3 กำหนดตัวแปรอิสระให้เป็นลำดับของข้อมูลอนุกรมเวลา

3.1.4 กำหนดข้อมูลของตัวแปรตามมีลักษณะเป็นข้อมูลอนุกรมเวลาทั้งในรูปแบบเชิงเส้นและไม่เชิงเส้น

3.1.5 กำหนดค่าส่วนเบี่ยงเบนมาตรฐานของค่าความคลาดเคลื่อนสุ่ม

3.1.6 กำหนดค่าจำนวนนอต

3.1.7 กำหนดเกณฑ์การพิจารณาประสิทธิภาพการประมาณค่า

3.1.8 จำลองข้อมูลด้วยโปรแกรมอาร์ โดยทำการทดลองซ้ำ 500 ครั้งในแต่ละสถานการณ์

3.1.9 ศึกษาข้อมูลอนุกรมเวลาจากข้อมูลจริง 2 ชุด โดยพิจารณาประสิทธิภาพของการพยากรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 วิธีการดำเนินงานวิจัยสำหรับข้อมูลจำลอง

การวิจัยในครั้งนี้ศึกษาประสิทธิภาพของการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์แต่ละวิธีของข้อมูลจำลอง โดยมีวิธีการดำเนินงานดังนี้

3.2.1 กำหนดขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150 และ 200

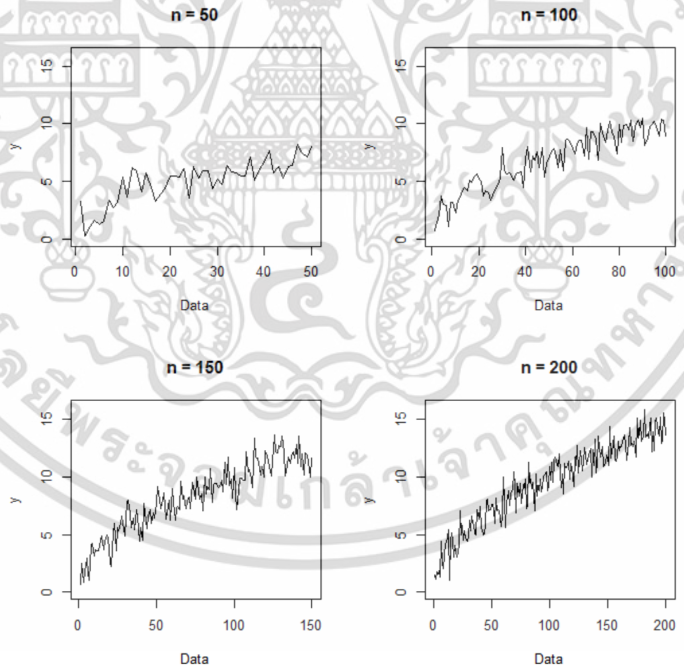
3.2.2 กำหนดตัวแปรอิสระ (x) ให้เป็นลำดับของข้อมูลอนุกรมเวลาเริ่มตั้งแต่ $1,2,...,n$ เมื่อ n คือจำนวนข้อมูลทั้งหมด

3.2.3 ข้อมูลของตัวแปรตามมีลักษณะเป็นข้อมูลอนุกรมเวลาทั้งในรูปแบบเชิงเส้นและไม่เชิงเส้น ดังนี้

3.2.3.1 ข้อมูลมีลักษณะเป็นแนวโน้ม (Trend) โดยตัวแปร

$$y_t = \sqrt{x_t} + \varepsilon_t \ , t = 1,2,...,n$$

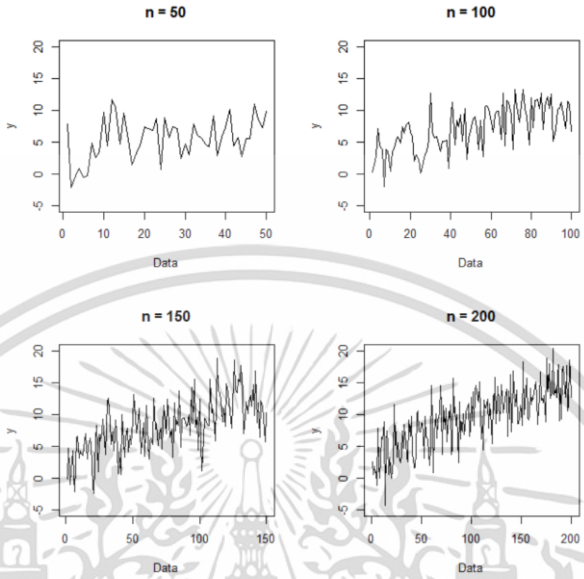
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 1 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.1



รูปที่ 3.1 ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=1

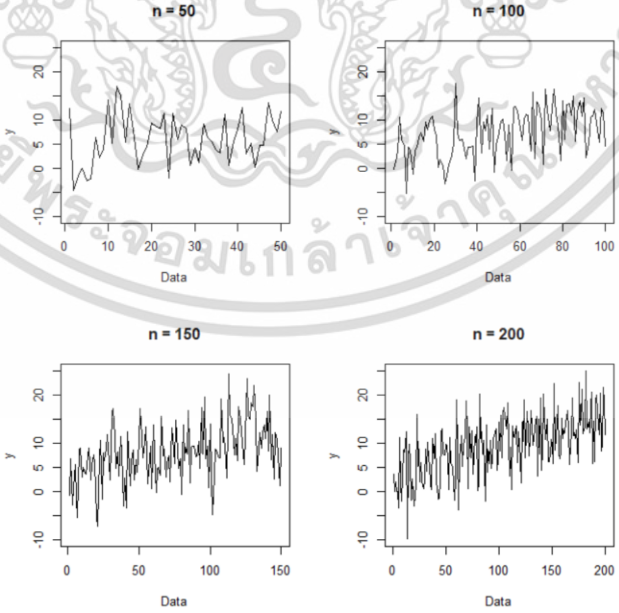
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 3 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.2



รูปที่ 3.2 ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=3

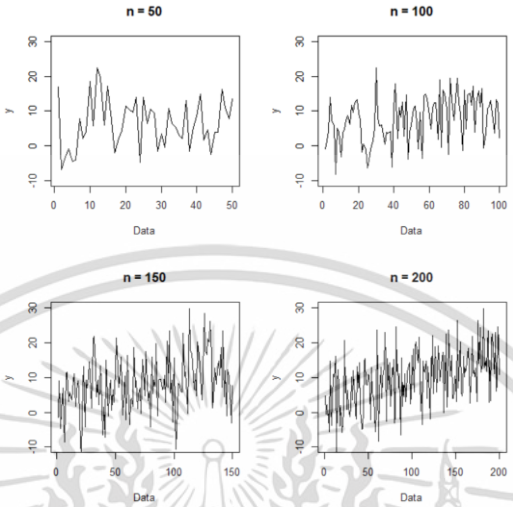
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 5 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.3



รูปที่ 3.3 ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 7 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.4

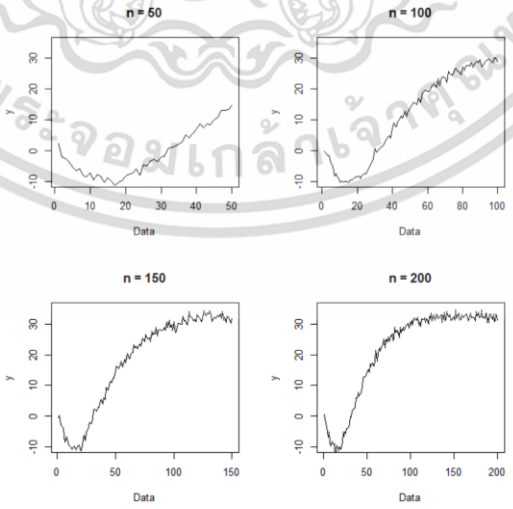


รูปที่ 3.4 ข้อมูลมีลักษณะเป็นแนวโน้ม เมื่อความคลาดเคลื่อนมีค่า SD=7

3.2.3.2 ข้อมูลมีลักษณะไม่เชิงเส้น (Non-Linear) โดยตัวแปร

$$y_t = 3\sqrt{x_t}\,\sin\left(2\pi\left[\frac{1+30}{x_t+30}\right]\right) + \varepsilon_t\,,\ t=1,2,...,n$$

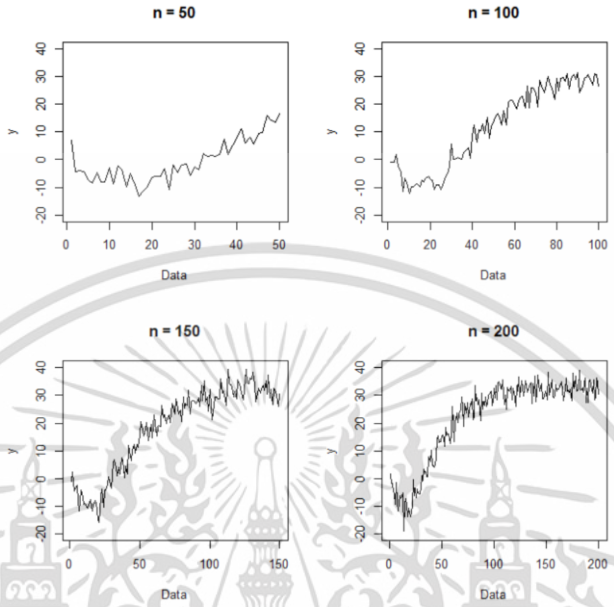
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 1 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.5



รูปที่ 3.5 ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=1

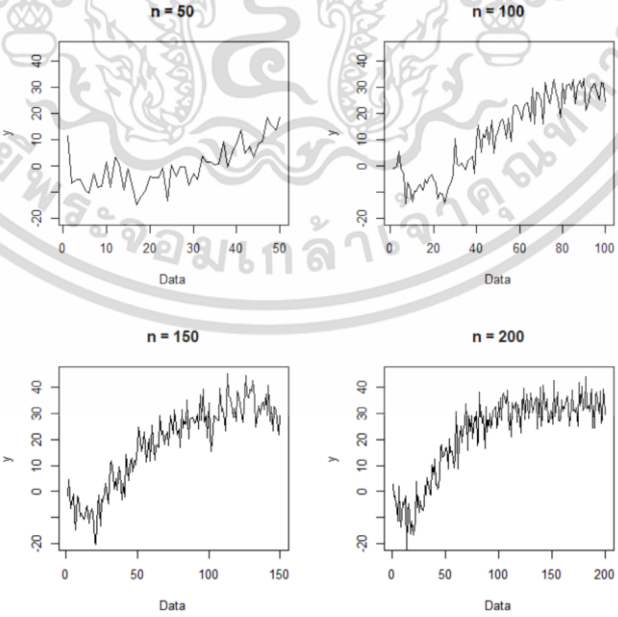
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 3 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.6



รูปที่ 3.6 ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=3

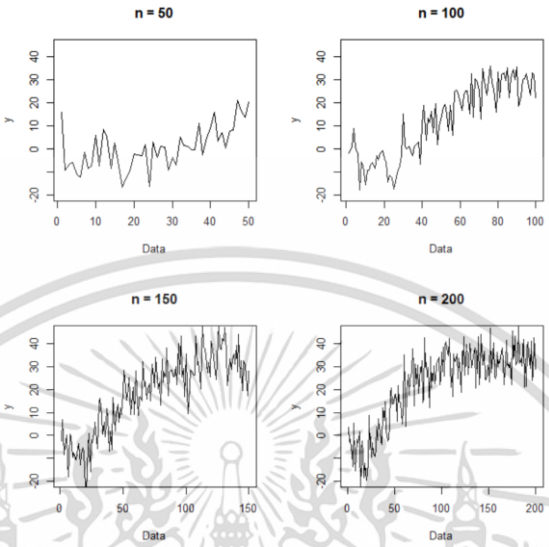
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 5 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.7



รูปที่ 3.7 ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 7 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.8

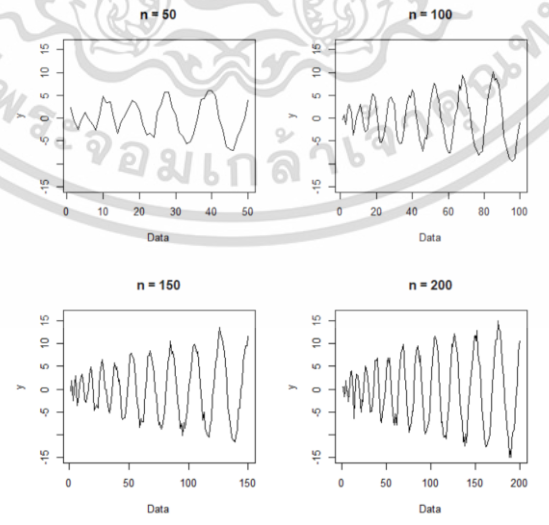


รูปที่ 3.8 ข้อมูลมีลักษณะไม่เชิงเส้น เมื่อความคลาดเคลื่อนมีค่า SD=7

3.2.3.3 ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ (Heteroscedasticity) โดยตัวแปร

$$y_t = \sqrt{x_t}\,\sin\left(2\pi\left[1+\sqrt{x_t}\right]\right)+\ \varepsilon_t\ ,\ t=1,2,...,n$$

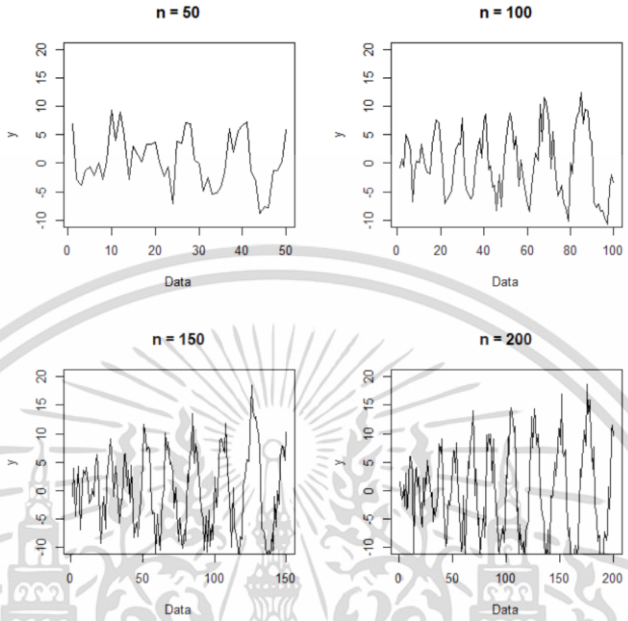
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 1 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.9



รูปที่ 3.9 ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=1

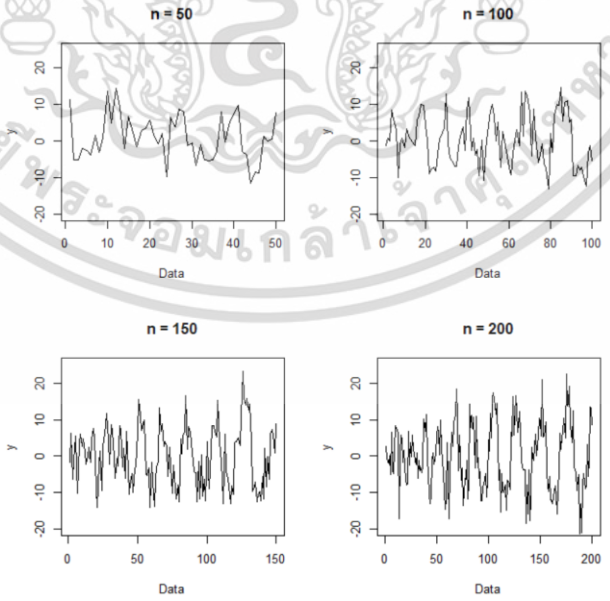
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 3 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.10



รูปที่ 3.10 ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=3

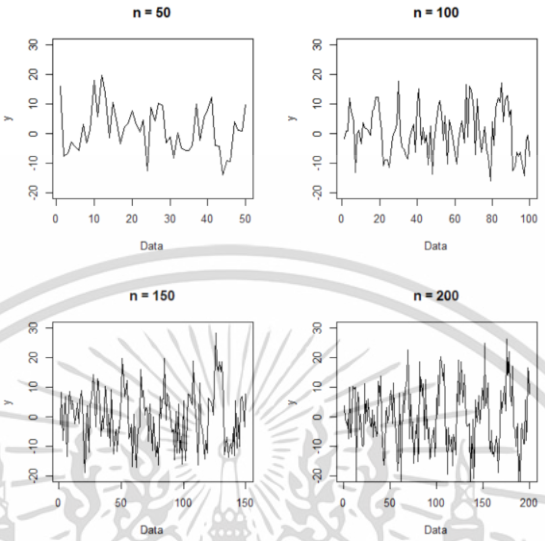
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 5 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.11



รูปที่ 3.11 ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

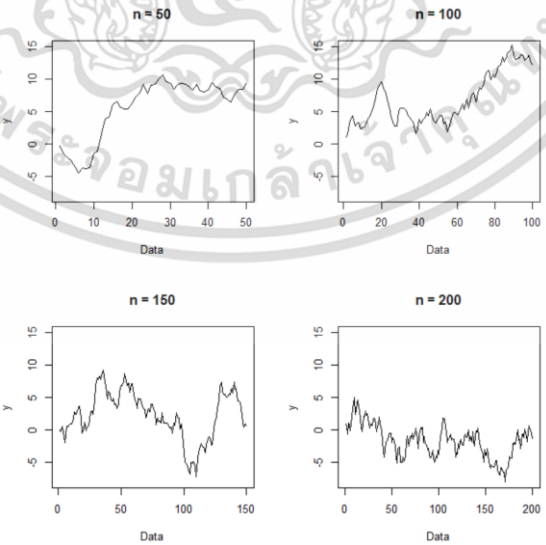
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 7 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.12



รูปที่ 3.12 ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=7

3.2.3.4 ข้อมูลมีลักษณะไม่คงที่ (Non-Stationary) หรือการเดินแบบสุ่ม (Random Walk) โดยตัวแปร $y_t = y_{t-1} + \varepsilon_t$, $t = 2,3,...,n$ โดยที่ $y_1 = N(0,1)$

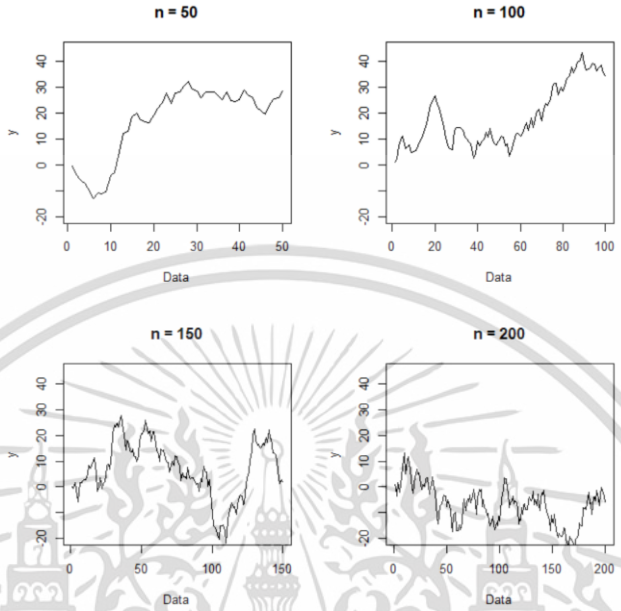
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 1 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.13



รูปที่ 3.13 ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=1

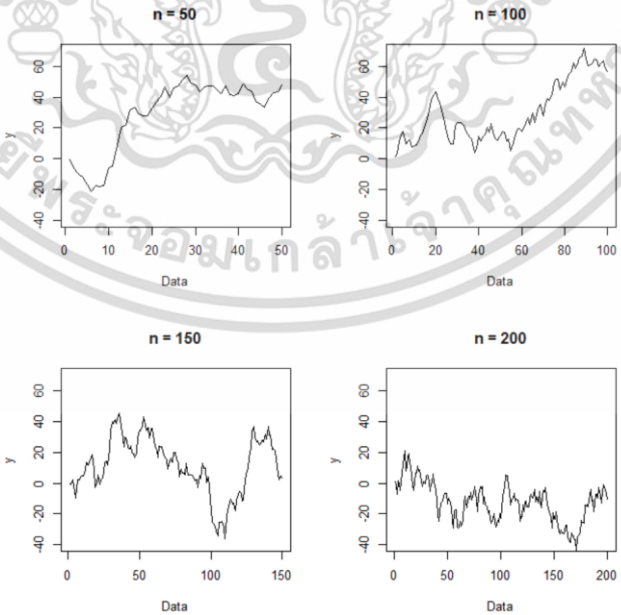
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 3 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.14



รูปที่ 3.14 ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=3

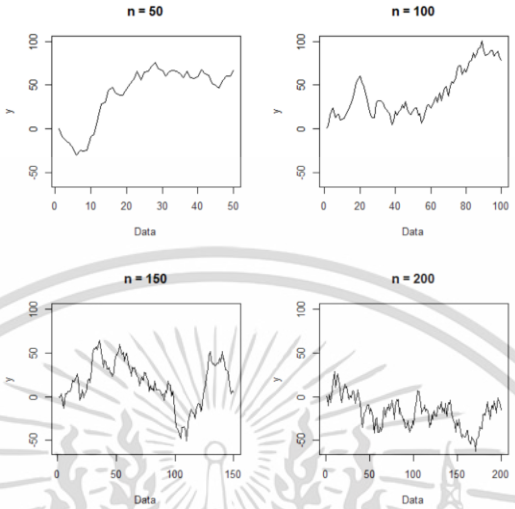
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 5 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.15



รูปที่ 3.15 ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 7 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.16

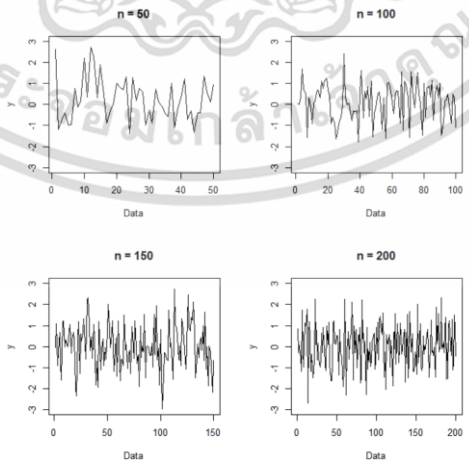


รูปที่ 3.16 ข้อมูลมีลักษณะไม่คงที่ เมื่อความคลาดเคลื่อนมีค่า SD=7

3.2.3.5 ข้อมูลมีลักษณะคงที่ (Stationary) โดยตัวแปร

$$y_t = \exp\left[\sin\left(x_t - \frac{x_t}{1+x_t^2}\right)\right] + \varepsilon_t \ , t = 1,2,...,n$$

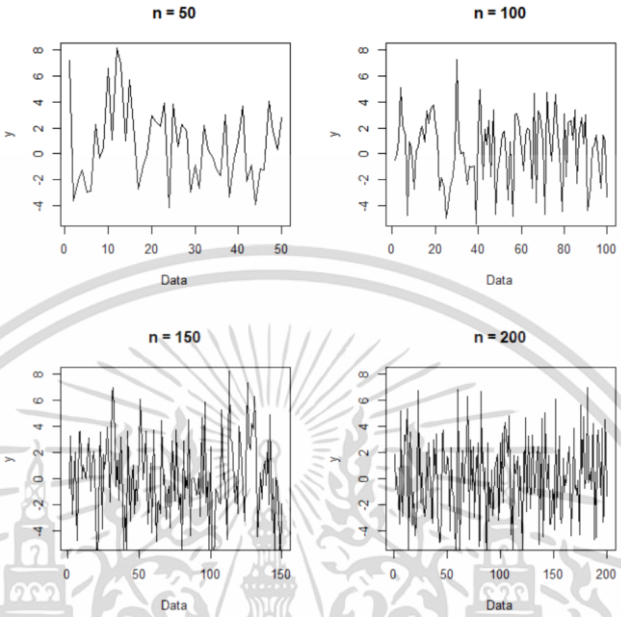
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 1 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.17



รูปที่ 3.17 ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=1

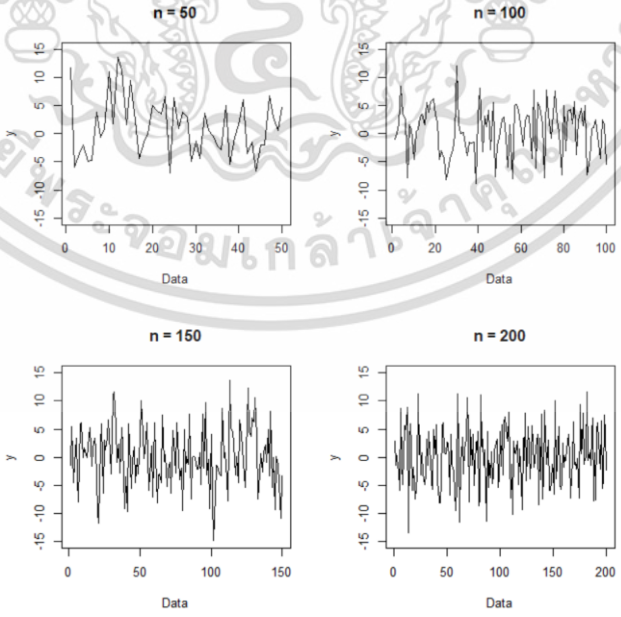
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 3 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.18



รูปที่ 3.18 ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=3

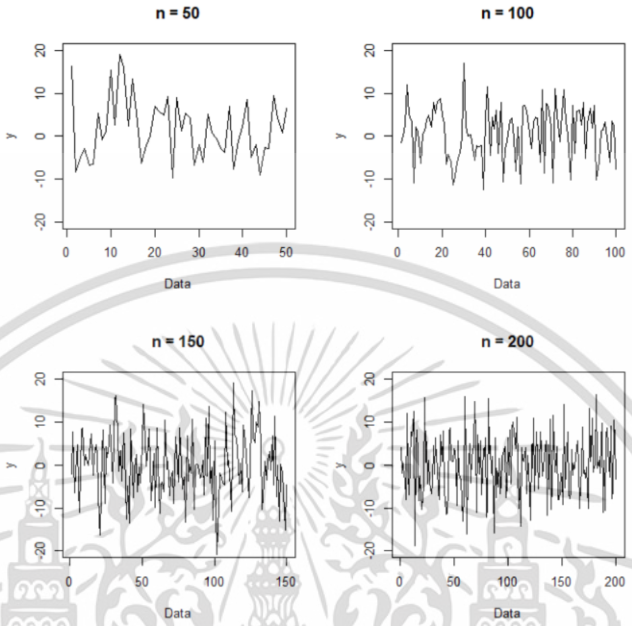
- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 5 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.19



รูปที่ 3.19 ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการสุ่มค่าความคลาดเคลื่อน (ε_t) ให้มีการแจกแจงปรกติที่ค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 7 และเมื่อขนาดตัวอย่างเพิ่มมากขึ้นแสดงลักษณะข้อมูล ดังรูปที่ 3.20



รูปที่ 3.20 ข้อมูลมีลักษณะคงที่ เมื่อความคลาดเคลื่อนมีค่า SD=7

3.2.4 กำหนดขอบเขตของจำนวนนอตตามขนาดตัวอย่าง เท่ากับ 50, 100, 150 และ 200

3.2.5 ประมาณค่าพารามิเตอร์ปรับให้เรียบ (λ) เพื่อปรับเส้นโค้งให้เรียบขึ้น โดยใช้การแบ่งข้อมูลเป็นหลายส่วน (Cross-Validation)

3.2.6 วิเคราะห์ข้อมูลจากวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี คือ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ

3.2.7 เกณฑ์การพิจารณาประสิทธิภาพการประมาณค่า คือค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด

3.2.8 จำลองข้อมูลด้วยโปรแกรมอาร์ โดยทำการทดลองซ้ำ 500 ครั้งในแต่ละสถานการณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 การประยุกต์ใช้กับข้อมูลจริง

การวิจัยในครั้งนี้ศึกษาประสิทธิภาพการพยากรณ์ของการถดถอยไม่อิงพารามิเตอร์แต่ละวิธีของข้อมูลจริง 2 ชุด โดยมีวิธีการดำเนินงานดังนี้

3.3.1 กำหนดตัวแปรอิสระ (x) ให้เป็นลำดับของข้อมูลอนุกรมเวลาเริ่มตั้งแต่ $1,2,...,n$ เมื่อ n คือจำนวนข้อมูลทั้งหมด

3.3.2 ข้อมูลของตัวแปรตามมีลักษณะเป็นข้อมูลอนุกรมเวลาทั้งในรูปแบบเชิงเส้นและไม่เชิงเส้น ดังนี้

- อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา (บาท/ดอลลาร์ สรอ.) (ธนาคารแห่งประเทศไทย, 2564) ตั้งแต่วันที่ 10 พฤศจิกายน 2563 ถึงวันที่ 9 กันยายน 2564 จำนวน 200 วัน ซึ่งเป็นข้อมูลรายวัน มีลักษณะเป็นแนวโน้มและการเดินแบบสุ่ม
- ปริมาณการนำเข้าไฟฟ้าของประเทศไทย (กิกะวัตต์ชั่วโมง) (การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย, 2564) ตั้งแต่เดือนพฤษภาคม 2547 ถึงเดือนธันวาคม 2563 จำนวน 200 เดือน ซึ่งเป็นข้อมูลรายเดือน มีลักษณะเป็นแนวโน้มและฤดูกาล

3.3.3 กำหนดขอบเขตของจำนวนนอตตามขนาดตัวอย่าง

3.3.4 ประมาณค่าพารามิเตอร์ปรับให้เรียบ (λ) เพื่อปรับเส้นโค้งให้เรียบขึ้น โดยใช้การแบ่งข้อมูลเป็นหลายส่วน (Cross-Validation)

3.3.5 วิเคราะห์ข้อมูลจากวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี คือ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ

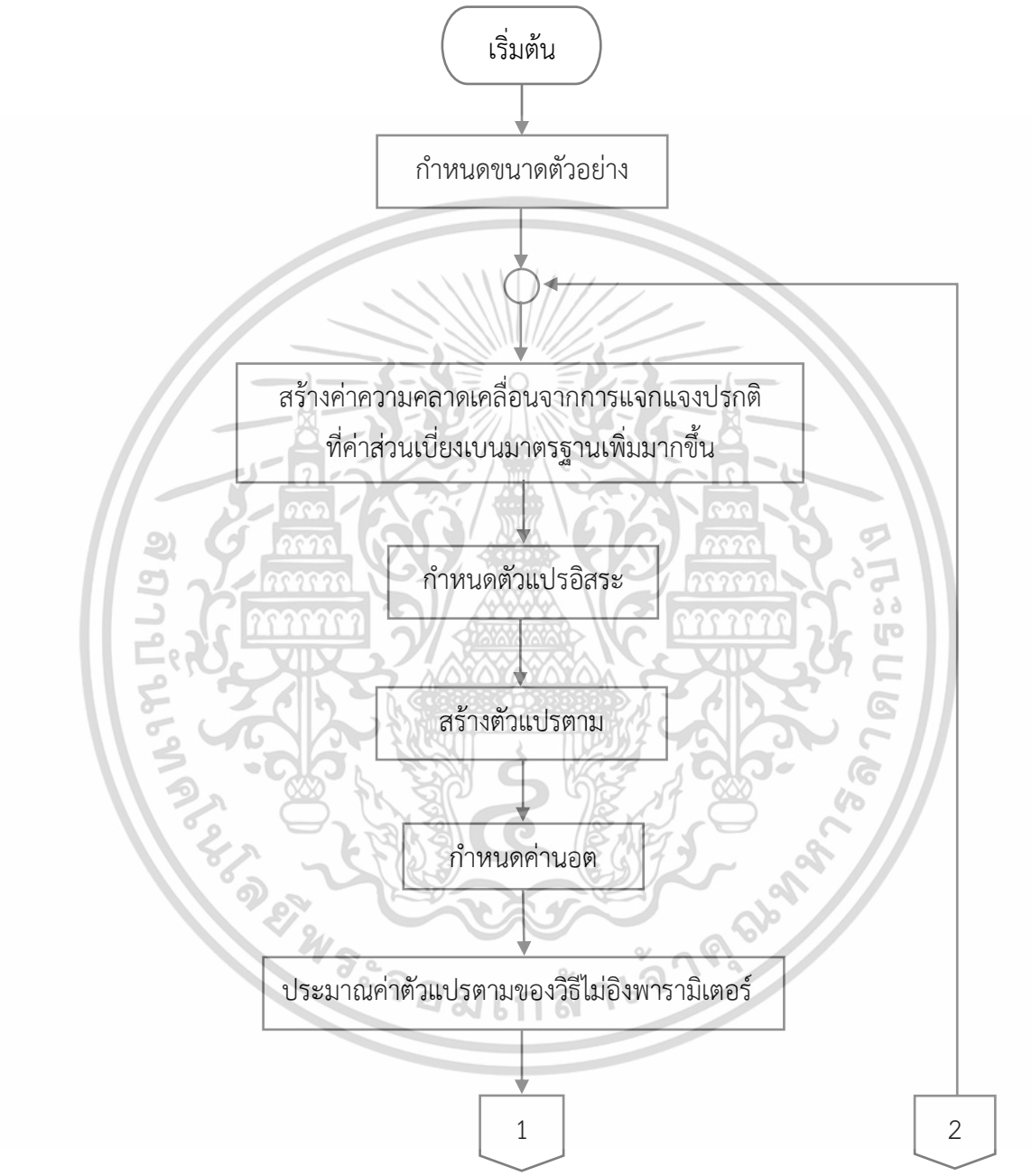
3.3.6 เกณฑ์การพิจารณาประสิทธิภาพการประมาณค่า คือค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด (MSE)

3.3.7 เกณฑ์การพิจารณาประสิทธิภาพการพยากรณ์ คือค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด (MAPE)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

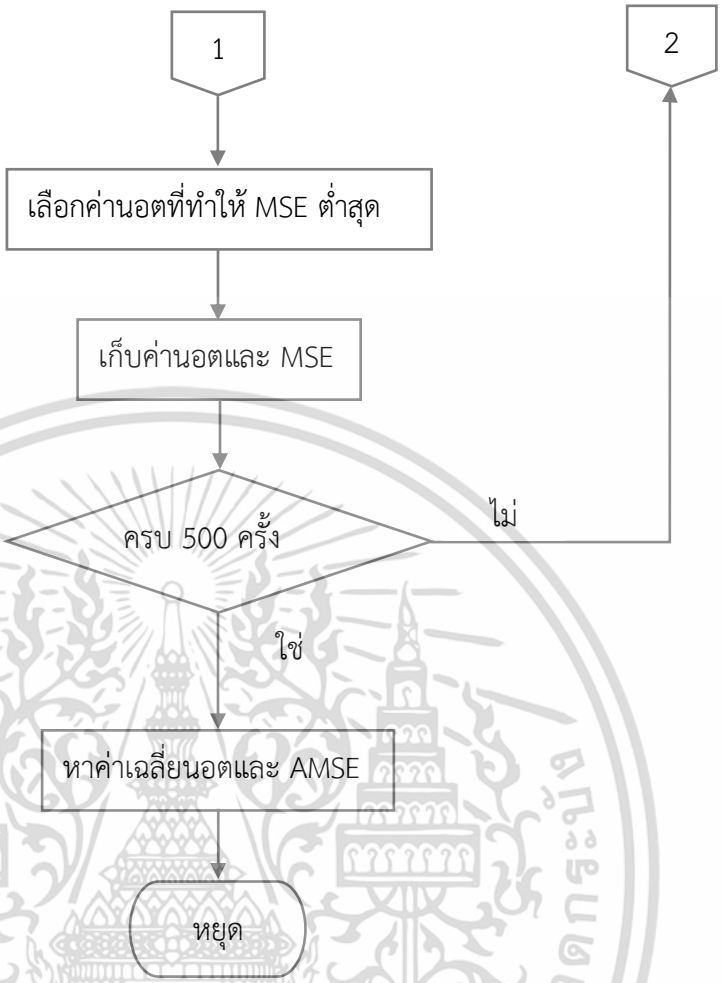
3.4 ขั้นตอนของโปรแกรมที่ใช้ในการวิจัย

3.4.1 ข้อมูลจำลอง



รูปที่ 3.21 แผนผังขั้นตอนการทำงานของโปรแกรมของข้อมูลจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.21 แผนผังขั้นตอนการทำงานของโปรแกรมของข้อมูลจำลอง (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.2 ประยุกต์ใช้กับข้อมูลจริง

รูปที่ 3.22 แผนผังขั้นตอนการทำงานของโปรแกรมของการประยุกต์ใช้กับข้อมูลจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการวิจัย

ในบทนี้จะนำเสนอผลการวิเคราะห์ข้อมูลเพื่อเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ของข้อมูลอนุกรมเวลา คือ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ ซึ่งวัดประสิทธิภาพการประมาณค่าจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) โดยใช้ข้อมูลจำลองและข้อมูลจริง 2 ชุด คือ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย เพื่อนำตัวแบบการถดถอยไม่อิงพารามิเตอร์ที่มีประสิทธิภาพไปใช้ในการพยากรณ์ข้อมูลจริงด้วยเกณฑ์ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) โดยกำหนดสัญลักษณ์และคำนิยามเพื่อใช้ประกอบการนำเสนอผลการวิเคราะห์ ดังนี้

SS	แทน	วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
NS	แทน	วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
BS	แทน	วิธีเส้นโค้งบี (B-Spline Method)
PS	แทน	วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
n	แทน	ขนาดตัวอย่าง
แนวโน้ม	แทน	ลักษณะข้อมูลเป็นแนวโน้ม (Trend)
ไม่เชิงเส้น	แทน	ลักษณะข้อมูลไม่เชิงเส้น (Non-Linear)
ไม่คงที่	แทน	ลักษณะข้อมูลไม่คงที่ (Non-Stationary)
ความแปรปรวนไม่คงที่	แทน	ลักษณะข้อมูลความแปรปรวนไม่คงที่ (Heteroscedasticity)

4.1 ผลการวิเคราะห์ข้อมูลจำลอง

การเปรียบเทียบประสิทธิภาพของการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์แต่ละวิธีของข้อมูลจำลอง พิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ในการประมาณค่าวิธีที่มีประสิทธิภาพในการประมาณค่าดีที่สุดจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ผลการวิเคราะห์ข้อมูลจำลองแสดงดังตารางที่ 4.1-4.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1

n	วิธี	แนวโน้ม		ไม่เชิงเส้น		ไม่คงที่		ความแปรปรวนไม่คงที่	
		MSE	Knot	MSE	Knot	MSE	Knot	MSE	Knot
50	SS	3.26×10^{-20}	50	3.26×10^{-20}	50	8.73×10^{-21}	50	3.42×10^{-20}	50
	NS	$\mathbf{7.92\times10^{-29}}$	49	$\mathbf{2.52\times10^{-28}}$	49	$\mathbf{1.05\times10^{-28}}$	49	$\mathbf{3.16\times10^{-29}}$	49
	BS	8.98×10^{-25}	49	9.00×10^{-25}	49	2.56×10^{-25}	49	8.97×10^{-25}	49
	PS	4.65×10^{-3}	49	4.69×10^{-3}	49	1.26×10^{-3}	49	9.69×10^{-3}	49
100	SS	0.36	65	0.36	65	0.11	65	0.36	64
	NS	$\mathbf{4.54\times10^{-28}}$	99	$\mathbf{2.35\times10^{-27}}$	99	$\mathbf{5.07\times10^{-28}}$	99	$\mathbf{1.34\times10^{-28}}$	99
	BS	2.14×10^{-17}	99	2.15×10^{-17}	99	5.23×10^{-18}	99	2.14×10^{-17}	99
	PS	3.72×10^{-11}	100	3.74×10^{-11}	100	1.12×10^{-11}	100	7.40×10^{-11}	100
150	SS	0.46	87	0.46	90	0.15	87	0.47	83
	NS	$\mathbf{6.21\times10^{-27}}$	149	$\mathbf{3.54\times10^{-26}}$	149	$\mathbf{6.93\times10^{-27}}$	149	$\mathbf{9.46\times10^{-28}}$	149
	BS	1.84×10^{-3}	148	1.84×10^{-3}	148	4.71×10^{-4}	148	1.84×10^{-3}	148
	PS	1.46×10^{-3}	149	1.46×10^{-3}	149	4.21×10^{-4}	149	2.82×10^{-3}	149
200	SS	0.49	106	0.49	109	0.16	107	0.49	103
	NS	$\mathbf{1.09\times10^{-26}}$	199	$\mathbf{7.37\times10^{-26}}$	199	$\mathbf{1.09\times10^{-26}}$	199	$\mathbf{2.86\times10^{-27}}$	199
	BS	2.31×10^{-3}	198	2.31×10^{-3}	198	6.13×10^{-4}	198	2.31×10^{-3}	198
	PS	2.07×10^{-10}	200	2.07×10^{-10}	200	6.52×10^{-11}	200	3.85×10^{-10}	200

ตารางที่ 4.2 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3

n	วิธี	แนวโน้ม		ไม่เชิงเส้น		ไม่คงที่		ความแปรปรวนไม่คงที่	
		MSE	Knot	MSE	Knot	MSE	Knot	MSE	Knot
50	SS	2.93×10^{-19}	50	2.93×10^{-19}	50	7.85×10^{-20}	50	2.95×10^{-19}	50
	NS	$\mathbf{1.66\times10^{-28}}$	49	$\mathbf{3.42\times10^{-28}}$	49	$\mathbf{9.30\times10^{-28}}$	49	$\mathbf{1.20\times10^{-28}}$	49
	BS	8.11×10^{-24}	49	8.08×10^{-24}	49	2.31×10^{-24}	49	8.10×10^{-24}	49
	PS	4.18×10^{-2}	49	4.19×10^{-2}	49	1.14×10^{-2}	49	4.76×10^{-2}	49
100	SS	3.26	64	3.26	65	0.95	65	3.26	64
	NS	$\mathbf{8.76\times10^{-28}}$	99	$\mathbf{2.74\times10^{-27}}$	99	$\mathbf{4.51\times10^{-27}}$	99	$\mathbf{5.42\times10^{-28}}$	99
	BS	1.93×10^{-16}	99	1.93×10^{-16}	99	4.70×10^{-17}	99	1.93×10^{-16}	99
	PS	3.34×10^{-10}	100	3.35×10^{-10}	100	1.01×10^{-10}	100	3.76×10^{-10}	100
150	SS	4.14	84	4.14	86	1.32	87	4.15	85
	NS	$\mathbf{8.00\times10^{-27}}$	149	$\mathbf{3.71\times10^{-26}}$	149	$\mathbf{6.23\times10^{-26}}$	149	$\mathbf{2.80\times10^{-27}}$	149
	BS	1.66×10^{-2}	148	1.66×10^{-2}	148	4.24×10^{-3}	148	1.66×10^{-2}	148
	PS	1.31×10^{-2}	149	1.31×10^{-2}	149	3.79×10^{-3}	149	1.43×10^{-2}	149
200	SS	4.43	103	4.43	107	1.45	109	4.43	102
	NS	$\mathbf{1.28\times10^{-26}}$	199	$\mathbf{7.50\times10^{-26}}$	199	$\mathbf{9.69\times10^{-26}}$	199	$\mathbf{5.22\times10^{-27}}$	199
	BS	2.08×10^{-2}	198	2.08×10^{-2}	198	5.52×10^{-3}	198	2.08×10^{-2}	198
	PS	1.86×10^{-9}	200	1.86×10^{-9}	200	5.87×10^{-10}	200	2.01×10^{-9}	200

ตารางที่ 4.3 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 5

n	วิธี	แนวโน้ม		ไม่เชิงเส้น		ไม่คงที่		ความแปรปรวนไม่คงที่	
		MSE	Knot	MSE	Knot	MSE	Knot	MSE	Knot
50	SS	8.14×10^{-19}	50	8.14×10^{-19}	50	2.18×10^{-19}	50	8.16×10^{-19}	50
	NS	$\mathbf{3.42\times10^{-28}}$	49	$\mathbf{5.19\times10^{-28}}$	49	$\mathbf{2.59\times10^{-27}}$	49	$\mathbf{2.92\times10^{-28}}$	49
	BS	2.24×10^{-23}	49	2.25×10^{-23}	49	$6..40\times10^{-24}$	49	2.25×10^{-23}	49
	PS	0.12	49	0.12	49	3.16×10^{-2}	49	0.12	49
100	SS	9.05	64	9.05	64	2.65	65	9.05	64
	NS	$\mathbf{1.68\times10^{-27}}$	99	$\mathbf{3.55\times10^{-27}}$	99	$\mathbf{1.26\times10^{-26}}$	99	$\mathbf{1.35\times10^{-27}}$	99
	BS	5.36×10^{-16}	99	5.36×10^{-16}	99	1.31×10^{-16}	99	5.35×10^{-16}	99
	PS	9.28×10^{-10}	100	9.29×10^{-10}	100	2.81×10^{-10}	100	9.74×10^{-10}	100
150	SS	11.5	83	11.5	85	3.66	87	11.5	83
	NS	$\mathbf{1.16\times10^{-26}}$	149	$\mathbf{4.07\times10^{-26}}$	149	$\mathbf{1.73\times10^{-25}}$	149	$\mathbf{6.49\times10^{-27}}$	149
	BS	4.60×10^{-2}	148	4.60×10^{-2}	148	1.18×10^{-2}	148	4.60×10^{-2}	148
	PS	3.64×10^{-2}	149	3.64×10^{-2}	149	1.05×10^{-2}	149	3.75×10^{-2}	149
200	SS	12.3	102	12.3	106	4.02	106	12.3	102
	NS	$\mathbf{1.72\times10^{-26}}$	199	$\mathbf{7.85\times10^{-26}}$	199	$\mathbf{2.69\times10^{-25}}$	199	$\mathbf{9.94\times10^{-27}}$	199
	BS	5.78×10^{-2}	198	5.78×10^{-2}	198	1.53×10^{-2}	198	5.74×10^{-2}	198
	PS	5.17×10^{-9}	200	5.17×10^{-9}	200	1.63×10^{-9}	200	5.29×10^{-9}	200

ตารางที่ 4.4 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธี SS วิธี NS วิธี BS และวิธี PS เมื่อส่วนเบี่ยงเบนมาตรฐานเท่ากับ 7

n	วิธี	แนวโน้ม		ไม่เชิงเส้น		ไม่คงที่		ความแปรปรวนไม่คงที่	
		MSE	Knot	MSE	Knot	MSE	Knot	MSE	Knot
50	SS	1.60×10^{-18}	50	1.60×10^{-18}	50	4.28×10^{-19}	50	1.60×10^{-18}	50
	NS	$\mathbf{6.04\times10^{-28}}$	49	$\mathbf{7.82\times10^{-28}}$	49	$\mathbf{5.24\times10^{-27}}$	49	$\mathbf{5.51\times10^{-28}}$	49
	BS	4.43×10^{-23}	49	4.42×10^{-23}	49	1.26×10^{-23}	49	4.42×10^{-23}	49
	PS	0.23	49	0.23	49	6.20×10^{-2}	49	0.24	49
100	SS	17.7	64	17.7	64	5.20	65	17.7	64
	NS	$\mathbf{2.92\times10^{-27}}$	99	$\mathbf{4.77\times10^{-27}}$	99	$\mathbf{2.47\times10^{-26}}$	99	$\mathbf{2.57\times10^{-27}}$	99
	BS	1.05×10^{-15}	99	1.05×10^{-15}	99	2.56×10^{-16}	99	1.05×10^{-15}	99
	PS	1.82×10^{-9}	100	1.82×10^{-9}	100	5.50×10^{-10}	100	1.87×10^{-9}	100
150	SS	22.5	82	22.5	84	7.17	88	22.5	82
	NS	$\mathbf{1.72\times10^{-26}}$	149	$\mathbf{4.61\times10^{-26}}$	149	$\mathbf{3.39\times10^{-25}}$	149	$\mathbf{1.20\times10^{-26}}$	149
	BS	9.03×10^{-2}	148	9.03×10^{-2}	148	2.31×10^{-2}	148	9.03×10^{-2}	148
	PS	7.14×10^{-2}	149	7.14×10^{-2}	149	2.06×10^{-2}	149	7.24×10^{-2}	149
200	SS	24.1	101	24.1	103	7.88	106	24.1	102
	NS	$\mathbf{2.39\times10^{-26}}$	199	$\mathbf{8.46\times10^{-26}}$	199	$\mathbf{5.28\times10^{-25}}$	199	$\mathbf{1.70\times10^{-26}}$	199
	BS	0.11	198	0.11	198	3.00×10^{-2}	198	0.11	198
	PS	1.01×10^{-8}	200	1.01×10^{-8}	200	3.19×10^{-9}	200	1.02×10^{-8}	200

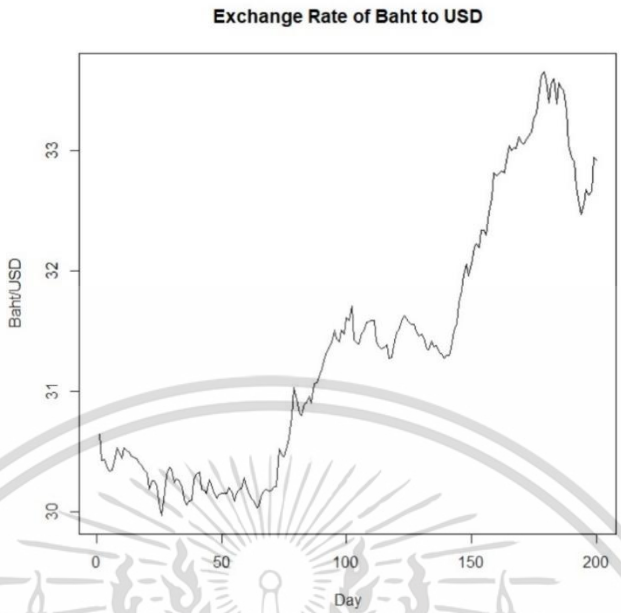
จากตารางที่ 4.1-4.4 พบว่าวิธีเส้นโค้งกำลังสามธรรมชาติ (NS) เมื่อค่าความคลาดเคลื่อนมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1, 3, 5 และ 7 ที่ขนาดตัวอย่าง 50, 100, 150 และ 200 มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกสถานการณ์ที่ทำการศึกษา และค่านอตมีค่าใกล้เคียงกับขนาดตัวอย่าง เมื่อพิจารณาลักษณะของข้อมูลพบว่าที่ขนาดตัวอย่างเดียวกันข้อมูลความแปรปรวนไม่คงที่จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกขนาดตัวอย่าง และเมื่อพิจารณาปัจจัยที่ส่งผลต่อประสิทธิภาพในการประมาณค่าพบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยเพิ่มมากขึ้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้น ตามประสิทธิภาพสัมพัทธ์เชิงเส้นกำกับ (Asymptotic Relative Efficiency) (Araveeporn, 2014, 2019)

4.2 ผลการวิเคราะห์การประยุกต์ใช้กับข้อมูลจริง

การพิจารณาประสิทธิภาพของการประมาณค่าและการพยากรณ์ของการถดถอยไม่อิงพารามิเตอร์แต่ละวิธีของข้อมูลจริง 2 ชุด คือ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย การประมาณค่าพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำที่สุดเป็นวิธีที่มีประสิทธิภาพที่สุดในการประมาณค่า การพยากรณ์พิจารณาจากค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ต่ำที่สุดเป็นวิธีที่มีประสิทธิภาพที่สุดในการพยากรณ์ ผลการวิเคราะห์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย แสดงดังรูปที่ 4.1-4.10

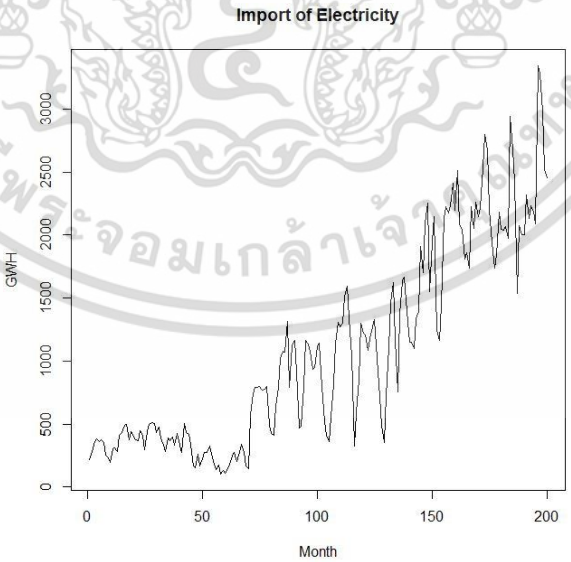
- อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา ตั้งแต่วันที่ 10 พฤศจิกายน 2563 ถึงวันที่ 9 กันยายน 2564 ซึ่งเป็นข้อมูลรายวันทั้งหมด 200 วัน โดยแบ่งข้อมูลในการประมาณค่า 170 วัน เนื่องจากข้อมูลเป็นรายวันจึงแบ่งข้อมูลสำหรับพยากรณ์ในช่วง 30 วัน หรือ 1 เดือน เพื่อดูแนวโน้มของข้อมูลที่มีลักษณะที่ผันแปรตลอดเวลาหรือเรียกว่าการเดินแบบสุ่มที่ข้อมูลมีลักษณะที่ไม่สามารถคาดเดาได้ แสดงดังรูปที่ 4.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริการายวัน

- ปริมาณการนำเข้าไฟฟ้าของประเทศไทย ตั้งแต่เดือนพฤษภาคม 2547 ถึงเดือนธันวาคม 2563 ซึ่งเป็นข้อมูลรายเดือน ซึ่งเป็นข้อมูลรายเดือนทั้งหมด 200 เดือน โดยแบ่งข้อมูลในการประมาณค่า 188 เดือน เนื่องจากข้อมูลเป็นรายเดือนจึงแบ่งข้อมูลสำหรับพยากรณ์ในช่วง 12 เดือน หรือ 1 ปี เพื่อดูแนวโน้มและฤดูกาลของข้อมูลที่อาจเกิดขึ้นเป็นประจำ แสดงดังรูปที่ 4.2

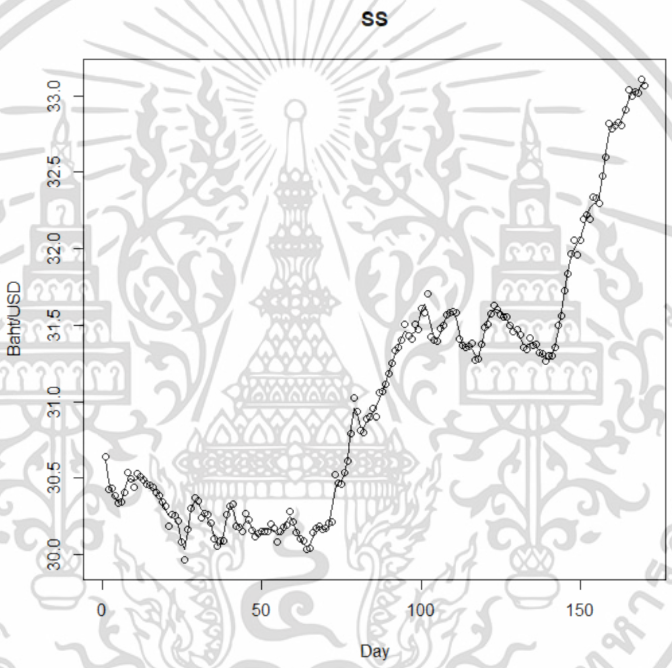


รูปที่ 4.2 ปริมาณการนำเข้าไฟฟ้าของประเทศไทยรายเดือน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

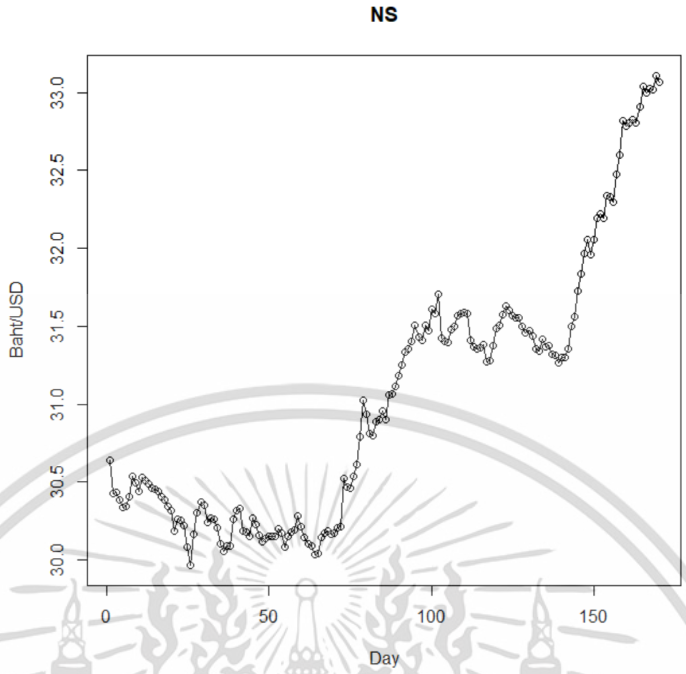
4.2.1 การประมาณค่าข้อมูลจริง

การพิจารณาประสิทธิภาพของการประมาณค่าการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ของข้อมูลอนุกรมเวลา คือ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ ของข้อมูลจริง 2 ชุด คือ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย การประมาณค่าพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำที่สุดเป็นวิธีที่มีประสิทธิภาพที่สุดในการประมาณค่า ผลการวิเคราะห์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย แสดงดังรูปที่ 4.3-4.10

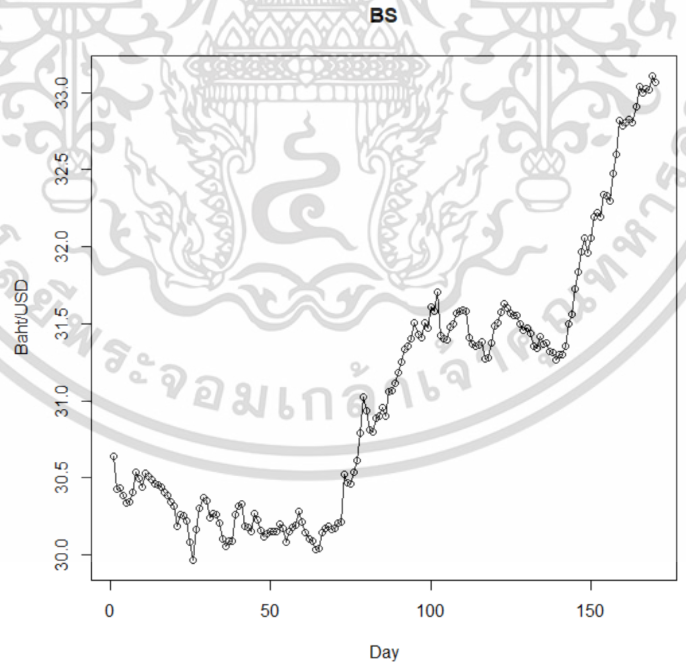


รูปที่ 4.3 การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งเรียบ (SS)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

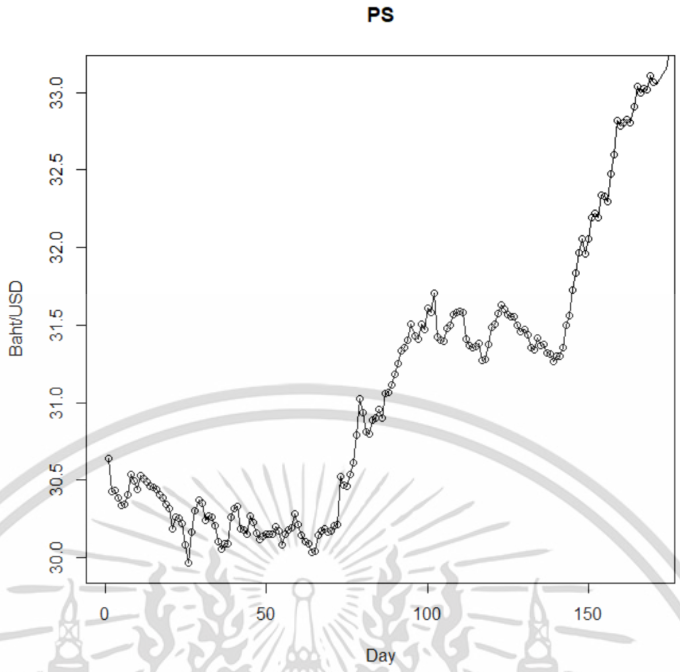


รูปที่ 4.4 การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งกำลังสามธรรมชาติ (NS)



รูปที่ 4.5 การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งบี (BS)

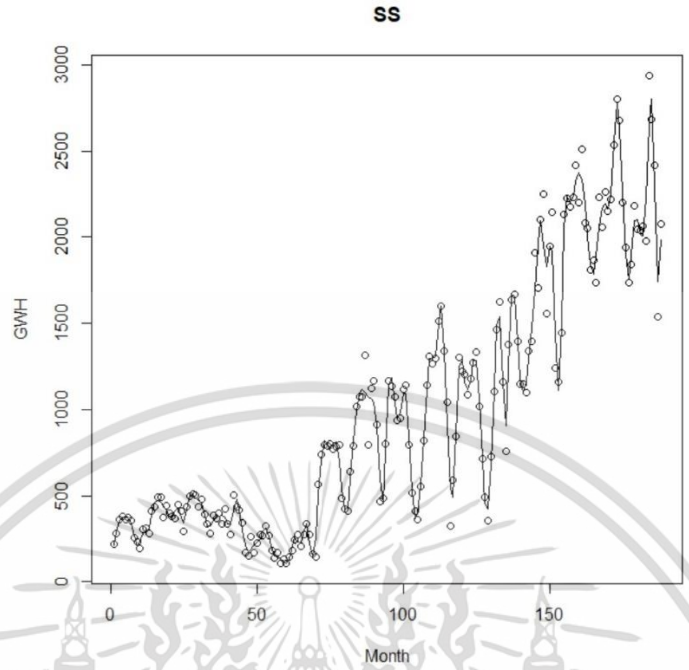
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



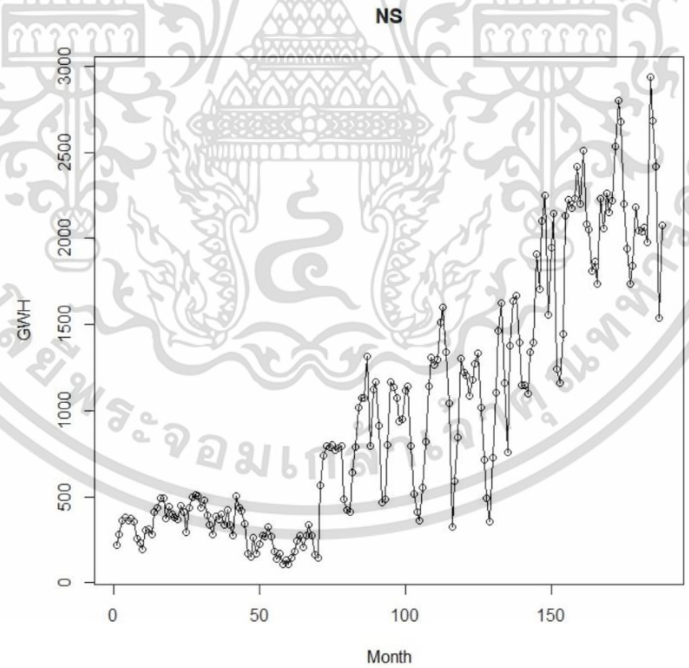
รูปที่ 4.6 การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งที่ถูกลงโทษ (PS)

จากการวิเคราะห์ข้อมูลทั้ง 4 วิธี ของข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกามีลักษณะข้อมูลเป็นแนวโน้มและการเดินแบบสุ่ม โดยให้ y เป็นอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา ให้ x เป็นลำดับวัน ซึ่งได้มีการแทนค่านอตในแต่ละตัวแบบจนกระทั่งได้ค่านอต และค่าพารามิเตอร์ปรับให้เรียบที่ให้ค่าที่ทำให้ได้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำสุด เพื่อให้เส้นโค้งที่ได้แทรกไปตามข้อมูล แสดงดังรูปที่ 4.3-4.6 โดยค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่านอต แสดงดังตาตางที่ 4.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

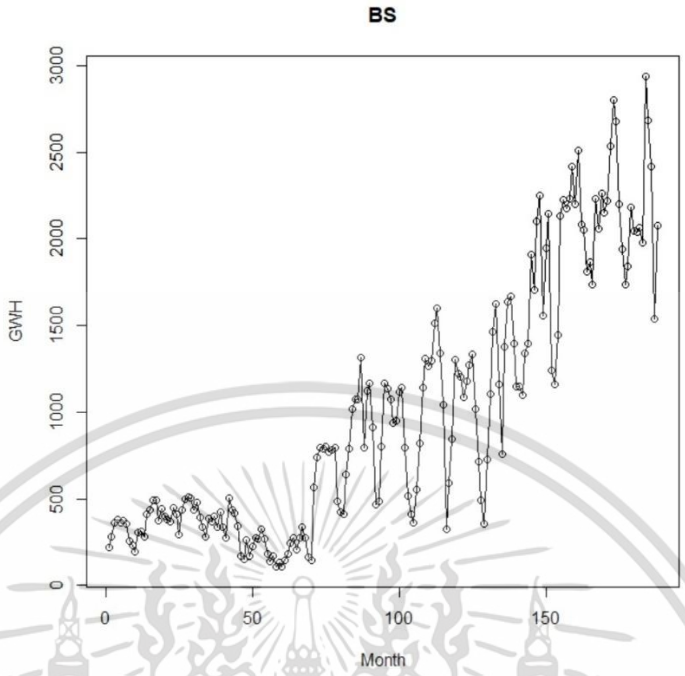


รูปที่ 4.7 การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งเรียบ (SS)

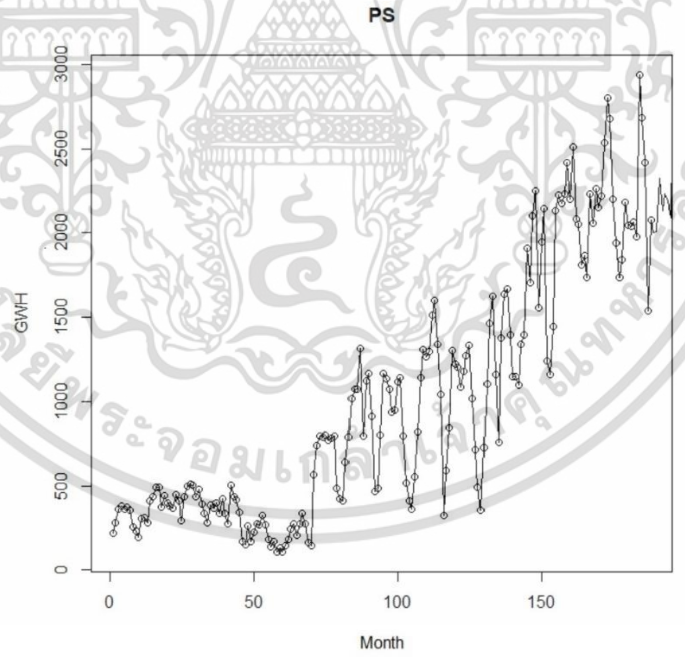


รูปที่ 4.8 การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งกำลังสามธรรมชาติ (NS)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.9 การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งบี (BS)



รูปที่ 4.10 การประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งที่ถูกลงโทษ (PS)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการวิเคราะห์ข้อมูลทั้ง 4 วิธี ของข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยมีลักษณะข้อมูลเป็นแนวโน้มและฤดูกาล โดยให้ y เป็นการประมาณค่าข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย ให้ x เป็นลำดับเดือน ซึ่งได้มีการแทนค่านอตในแต่ละตัวแบบจนกระทั่งได้ค่านอตและค่าพารามิเตอร์ปรับให้เรียบที่ให้ค่าที่ทำให้ได้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำสุด เพื่อให้เส้นโค้งที่ได้แทรกไปตามข้อมูล แสดงดังรูปที่ 4.7-4.10 โดยค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่านอตแสดงดังตาตางที่ 4.5

ตารางที่ 4.5 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)

วิธี	อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา		ปริมาณการนำเข้าไฟฟ้า	
	MSE	Knot	MSE	Knot
SS	8.22×10^{-4}	89	6175	97
NS	$\mathbf{8.27\times10^{-28}}$	169	$\mathbf{6.96\times10^{-23}}$	187
BS	1.08×10^{-6}	169	37.60	187
PS	2.61×10^{-4}	170	472.06	188

เมื่อพิจารณาจากรูปที่ 4.3-4.10 พบว่ากราฟค่าประมาณของข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย โดยวิธีเส้นโค้งทั้ง 4 วิธี มีลักษณะคล้ายคลึงกัน ทำให้ยากที่จะสรุปได้ว่าวิธีใดมีประสิทธิภาพดีที่สุด ดังนั้นจึงใช้เกณฑ์ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เป็นเกณฑ์ช่วยพิจารณา โดยพิจารณาจากตารางที่ 4.5 พบว่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา วิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด และข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย วิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด แสดงว่าวิธีเส้นโค้งกำลังสามธรรมชาติเป็นวิธีที่เหมาะสมที่สุดและมีประสิทธิภาพที่สุดในการประมาณค่าข้อมูลจริงทั้ง 2 ชุด

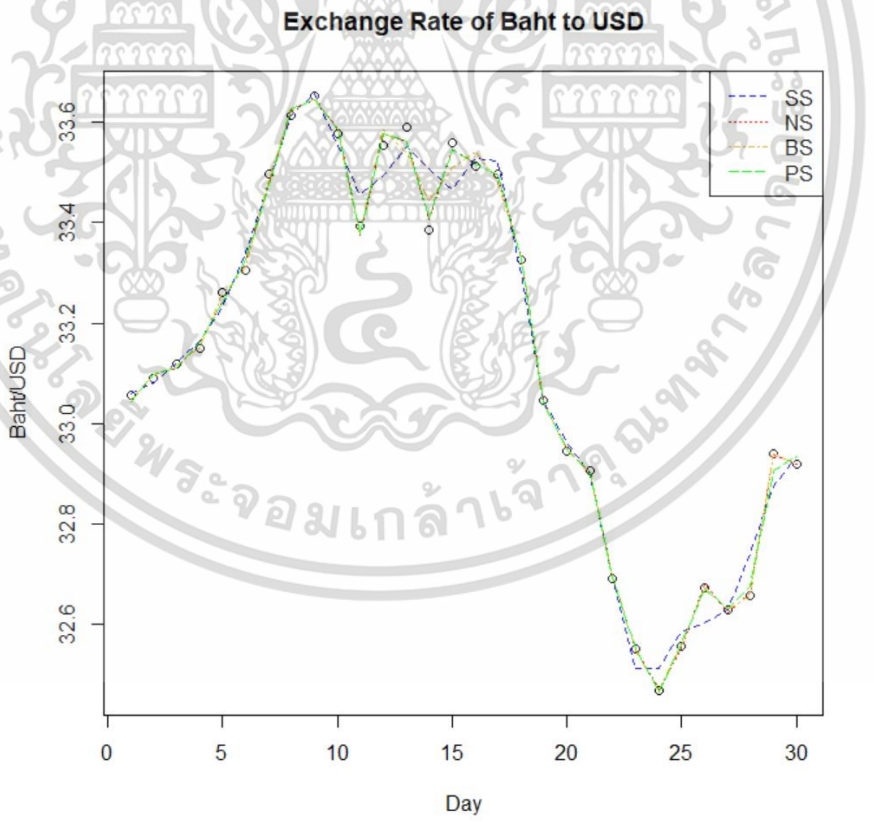
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 การพยากรณ์ข้อมูลจริง

การพิจารณาประสิทธิภาพของการพยากรณ์ของการถดถอยไม่อิงพารามิเตอร์ 4 วิธี ของข้อมูลอนุกรมเวลา คือ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ ของข้อมูลจริง 2 ชุด คือ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย การพยากรณ์พิจารณาจากค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ต่ำที่สุดเป็นวิธีที่มีประสิทธิภาพที่สุดในการพยากรณ์ ผลการวิเคราะห์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย แสดงดังรูปที่ 4.11-4.12

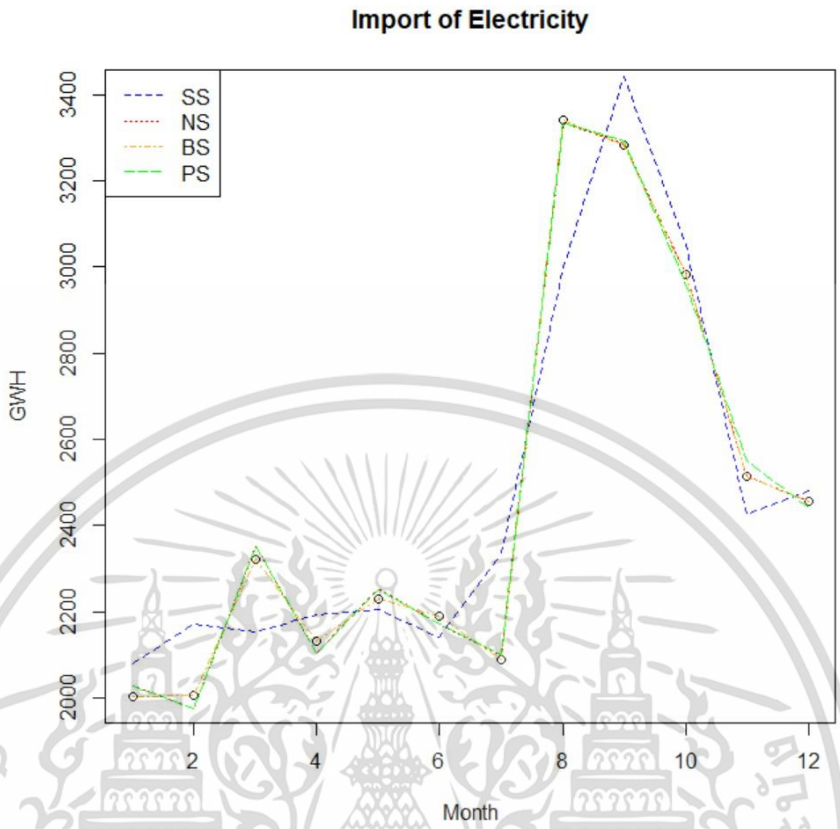
- อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา ใช้ข้อมูลสำหรับพยากรณ์ในช่วง 30 วัน หรือ 1 เดือน เพื่อดูแนวโน้มของข้อมูลที่มีลักษณะที่ผันแปรตลอดเวลาหรือเรียกว่าการเดินแบบสุ่มที่ข้อมูลมีลักษณะที่ไม่สามารถคาดเดาได้ แสดงดังรูปที่ 4.11

- ปริมาณการนำเข้าไฟฟ้าของประเทศไทย ใช้ข้อมูลสำหรับพยากรณ์ในช่วง 12 เดือน หรือ 1 ปี เพื่อดูแนวโน้มและฤดูกาลของข้อมูลที่อาจเกิดขึ้นเป็นประจำ แสดงดังรูปที่ 4.12



รูปที่ 4.11 การพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาด้วยวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 การพยากรณ์ข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยด้วยวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)

จากรูปที่ 4.11-4.12 การพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย ด้วยวิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ พบว่าวิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษมีลักษณะไม่แตกต่างกันมากนัก ทำให้ยากต่อการสรุปได้ว่าวิธีไหนมีประสิทธิภาพในการพยากรณ์ข้อมูลได้เหมาะสมที่สุด จึงใช้เกณฑ์ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) เป็นเกณฑ์ช่วยพิจารณาการพยากรณ์ โดยวิธีที่มีประสิทธิภาพในการพยากรณ์ดีที่สุดจะให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด ดังตารางที่ 4.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS)

วิธี	อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา	ปริมาณการนำเข้าไฟฟ้า
SS	9.71×10^{-2}	4.92
NS	$\mathbf{2.94\times10^{-2}}$	0.68
BS	3.50×10^{-2}	$\mathbf{1.01\times10^{-7}}$
PS	3.41×10^{-2}	0.92

จากตารางที่ 4.6 การพิจารณาประสิทธิภาพการพยากรณ์ของข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทยของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) วิธีเส้นโค้งบี (BS) และวิธีเส้นโค้งที่ถูกลงโทษ (PS) พบว่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา วิธีเส้นโค้งกำลังสามธรรมชาติมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด และข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย วิธีเส้นโค้งบีมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด แสดงว่าวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีเป็นวิธีที่มีประสิทธิภาพและเหมาะสมกับตัวแบบมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัย อภิปรายผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบวิธีการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์ 4 วิธี สำหรับข้อมูลอนุกรมเวลา ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษ ซึ่งใช้ข้อมูลจำลองที่มีลักษณะแนวโน้ม ไม่เชิงเส้น ไม่คงที่ และความแปรปรวนไม่คงที่ โดยทำการศึกษาและเปรียบเทียบในแต่ละสถานการณ์ ดังนี้

1. กำหนดขนาดตัวอย่างเท่ากับ 50, 100, 150 และ 200
2. กำหนดส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่ากับ 1, 3, 5 และ 7

เกณฑ์ที่ใช้ในการประมาณค่าหาวิธีที่มีประสิทธิภาพมากที่สุดในแต่ละสถานการณ์ จะพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) โดยใช้โปรแกรมอาร์จำลองแต่ละสถานการณ์ ทำซ้ำ 500 รอบ

นอกจากนี้ยังใช้ข้อมูลจริง 2 ชุด คือ ข้อมูลรายวันของอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา จำนวน 200 วัน และข้อมูลรายเดือนของปริมาณการนำเข้าไฟฟ้า จำนวน 200 เดือน เพื่อประมาณค่าหาวิธีที่มีประสิทธิภาพและนำไปใช้ในการพยากรณ์ โดยมีเกณฑ์ คือ ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)

5.1 สรุปผลการวิจัย

จากการศึกษาและเปรียบเทียบการประมาณค่าตัวแปรตามของการถดถอยไม่อิงพารามิเตอร์ 4 วิธี สำหรับข้อมูลอนุกรมเวลา โดยพิจารณาจากเกณฑ์ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของข้อมูลจำลองและค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยของข้อมูลจริง สามารถสรุปผลการวิจัยได้ดังนี้

- ข้อมูลจำลอง

1. ข้อมูลที่มีลักษณะแนวโน้ม ไม่เชิงเส้น ไม่คงที่ และความแปรปรวนไม่คงที่ ที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่ากับ 1, 3, 5 และ 7 ด้วยขนาดตัวอย่าง 50, 100, 150 และ 200 พบว่าวิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกสถานการณ์ และค่านอตมีขนาดใกล้เคียงกับขนาดตัวอย่าง แสดงว่าวิธีเส้นโค้งกำลังสามธรรมชาติเป็นวิธีที่มีประสิทธิภาพที่สุดในการประมาณค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. เมื่อพิจารณาจากลักษณะของข้อมูล พบว่าที่ขนาดตัวอย่างเดียวกันข้อมูลความแปรปรวนไม่คงที่จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกขนาดตัวอย่าง แสดงว่าลักษณะข้อมูลความแปรปรวนไม่คงที่มีประสิทธิภาพดีสำหรับข้อมูลอนุกรมเวลา

3. เมื่อขนาดตัวอย่างเพิ่มขึ้น พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีแนวโน้มเพิ่มมากขึ้นตามประสิทธิภาพสัมพัทธ์เชิงเส้นกำกับ (Asymptotic Relative Efficiency) (Araveeporn, 2014, 2019)

- การประยุกต์ใช้ข้อมูลจริง

1. การประมาณค่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและข้อมูลปริมาณการนำเข้าไฟฟ้า พบว่าวิธีเส้นโค้งทั้ง 4 วิธีมีลักษณะคล้ายคลึงกัน ทำให้ยากต่อการสรุปผลว่าวิธีใดมีประสิทธิภาพที่สุด จึงพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย จากผลการวิจัยวิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดและมีค่านอตใกล้เคียงกับขนาดตัวอย่างทั้ง 2 ชุดข้อมูล แสดงว่าวิธีเส้นโค้งกำลังสามธรรมชาติเป็นวิธีที่มีประสิทธิภาพที่สุดในการประมาณค่า

2. การพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและข้อมูลปริมาณการนำเข้าไฟฟ้า พบว่าวิธีเส้นโค้งกำลังสามธรรมชาติ วิธีเส้นโค้งบี และวิธีเส้นโค้งที่ถูกลงโทษมีลักษณะคล้ายกัน ทำให้ยากต่อการสรุปผลว่าวิธีใดมีประสิทธิภาพที่สุด จึงพิจารณาค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยเป็นเกณฑ์ในการหาวิธีที่มีประสิทธิภาพที่สุดในการพยากรณ์ จากผลการวิจัยพบว่าข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา วิธีเส้นโค้งกำลังสามธรรมชาติมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด และข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย วิธีเส้นโค้งบีมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด แสดงว่าวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีเป็นวิธีที่มีประสิทธิภาพที่สุดในการพยากรณ์

3. ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยของวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีต่ำกว่าวิธีอื่น แสดงว่าการพยากรณ์อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและข้อมูลปริมาณการนำเข้าไฟฟ้าของวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีมีความแม่นยำสูง เนื่องจากทั้ง 2 ชุดข้อมูลมีค่านอตสูงใกล้เคียงกับขนาดตัวอย่าง ส่งผลให้การพยากรณ์ที่ได้โค้งไปตามเส้นโค้งใกล้เคียงกับข้อมูลมากที่สุด จึงทำให้ค่าคลาดเคลื่อนน้อย

5.2 อภิปรายผลการวิจัย

จากการศึกษาและเปรียบเทียบวิธีการประมาณค่าตัวแปรตามของวิธีการถดถอยไม่อิงพารามิเตอร์ 4 วิธี สำหรับข้อมูลอนุกรมเวลา คือ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีเส้นโค้งที่ถูกลงโทษ จะพิจารณาประสิทธิภาพของวิธีการถดถอยไม่อิงพารามิเตอร์จากค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด โดยใช้ข้อมูลจำลองที่ลักษณะเป็นแนวโน้ม ไม่เชิงเส้น ไม่คงที่ และความแปรปรวนไม่คงที่ จากการประมาณค่า พบว่าวิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกสถานการณ์ ที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่ากับ 1, 3, 5 และ 7 มีขนาดตัวอย่าง 50, 100, 150 และ 200 และค่านอตมีค่าใกล้เคียงกับขนาดตัวอย่าง เมื่อพิจารณาลักษณะของข้อมูล พบว่าที่ขนาดตัวอย่างเดียวกันลักษณะข้อมูลความแปรปรวนไม่คงที่ จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดทุกขนาดตัวอย่าง

การประยุกต์ใช้ข้อมูลจริง 2 ชุด คือ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา จำนวน 200 วัน มีลักษณะข้อมูลเป็นแนวโน้มและการเดินแบบสุ่ม และข้อมูลปริมาณการนำเข้าไฟฟ้า จำนวน 200 เดือน มีลักษณะข้อมูลเป็นแนวโน้มและฤดูกาล โดยพิจารณาประสิทธิภาพการประมาณค่าและการพยากรณ์ โดยมีเกณฑ์ในการพยากรณ์ คือค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด จากการประมาณค่าและพยากรณ์ พบว่าการประมาณค่า วิธีเส้นโค้งกำลังสามธรรมชาติมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดทั้ง 2 ชุดข้อมูล สำหรับการพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา วิธีเส้นโค้งกำลังสามธรรมชาติมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด และข้อมูลปริมาณการนำเข้าไฟฟ้า วิธีเส้นโค้งบีมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด จึงสรุปได้ว่าวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีเป็นวิธีที่มีประสิทธิภาพกับลักษณะข้อมูลเป็นแนวโน้ม ไม่เชิงเส้นและฤดูกาลมากที่สุด ในส่วนของค่านอตยิ่งค่านอตมีค่าสูงจะทำให้เส้นโค้งที่สร้างขึ้นเข้าใกล้ข้อมูลมากขึ้น สอดคล้องกับการศึกษาของ Elhakeem และคณะ (2022) โดยวิธีเส้นโค้งกำลังสามธรรมชาติจะมีประสิทธิภาพดีในการประมาณค่า เมื่อข้อมูลมีลักษณะไม่เชิงเส้นและมีแนวโน้ม เนื่องจากสามารถปรับเส้นโค้งให้เข้ากับความซับซ้อนของข้อมูลได้ และ Araveeporn (2019) วิธีเส้นโค้งบีมีประสิทธิภาพกับข้อมูลไม่เชิงเส้น และกำหนดค่านอตให้เหมาะสมใกล้เคียงกับขนาดตัวอย่าง ดังนั้นในการประมาณค่าการถดถอยไม่อิงพารามิเตอร์ ถ้าทราบลักษณะของข้อมูลและกำหนดค่านอตให้เหมาะสมจะทำให้การประมาณค่ามีประสิทธิภาพดีขึ้น

5.3 ข้อเสนอแนะ

จากการวิจัยพบว่าการประมาณค่าและการพยากรณ์ข้อมูลจำลองและข้อมูลจริง โดยวิธีเส้นโค้งกำลังสามธรรมชาติเป็นวิธีที่มีประสิทธิภาพในการประมาณค่าและพยากรณ์ดีที่สุดเหมาะสมกับข้อมูลลักษณะไม่เชิงเส้น มีแนวโน้มและฤดูกาล นอกจากนี้การกำหนดค่านอต ถ้ากำหนดค่านอตน้อย จะทำให้เส้นโค้งไม่ปรับไปตามข้อมูล หรือมากเกินไปจะทำให้ใช้ระยะเวลาในการประมวลผลนานเกินความจำเป็น จึงควรกำหนดค่านอตตามขนาดตัวอย่างหรือใกล้เคียงกับขนาดตัวอย่าง เพื่อช่วยลดความคลาดเคลื่อนของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประมาณได้ และจากที่กล่าวมาข้างต้นสำหรับการศึกษาในอนาคต ควรศึกษาและเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์วิธีอื่น ๆ เช่น วิธีกรองแนวโน้ม (Trend Filtering Method) (Kim และคณะ, 2009) กับข้อมูลอนุกรมเวลาลักษณะอื่น เพื่อให้ได้วิธีที่มีประสิทธิภาพในการประมาณค่าและการพยากรณ์ที่เหมาะสมกับลักษณะของข้อมูลต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย. 2564. **Import of Electricity**. [Online]. Available : https://www.egat.co.th.

ธนาคารแห่งประเทศไทย. 2564. **Exchange Rate.** [Online]. Available : https://www.bot.or.th.

สำนักงานราชบัณฑิตยสภา. 2561. **พจนานุกรม ศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสภา.** พิมพ์ครั้งที่ 2. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

อานนท์ ศักดิ์วรวิชญ์. 2559. **ว่าด้วยการวิเคราะห์ถดถอย 33 ชนิด (On 33 kinds of Regression Analyses) พร้อมคำอธิบาย**. [Online]. Available : https://businessanalyticsnida.wordpress.com/2016/09/18/32-kinds-of-regression.

อัชฌา อระวีพร. 2555. "การประมาณค่าตัวแบบการถดถอยแบบไร้พารามิเตอร์ด้วยชุดคำสั่ง SemiPar ในโปรแกรมอาร์." *วารสารวิทยาศาสตร์ลาดกระบัง*. 21(2) : 78-86.

Ahmed, E. Rachael A., H. Kate, T. Diana L., C. Stefan A., J. Tim J., C. Alex S.F., K. Zheyuan, L. Struan F.A., G. Adam D.G., B. Babette S., Z. and Deborah A., L. 2022. "Using linear and natural cubic splines, SITAR, and latent trajectory models to characterize nonlinear longitudinal growth trajectories in cohort studies." *BMC Medical Research Methodology*. 22(1) : 1-20.

Araveeporn, A. 2012. "The Estimation of Smoothing Parameter using Smoothing Techniques on Nonparametric Regression." *Silpakorn U Science & Tech J*. 6(1) : 14-22.

Araveeporn, A. 2019. "As Estimating Parameter of Nonparametric Regression Model Based on Smoothing Techniques." *Statistical Journal of the IAOS*. 35(2) : 269-276.

Aris, P. Willi, S. Michal, A. and Matthias, S. 2019. **"A Review of Spline Function Procedures in R."** BMC Medical Research Methodology.

Cao, Y. Lin, H. Wu, Z. Wu, T. and Yu, Y. 2010. "Penalized Spline Estimation for Functional Coefficient Regression Models." *Journal of Computational Statistics and Data Analysis*. 54(4) : 891-905.

De Boor, C. 1978. **A Practical Guide to Splines.** Berlin : Springer.

Duan, L. Yu, F. Prdrycz, W. Wang, X. and Yang, X. 2018. "Time-Series Clustering Based on Linear Fuzzy Information Granules." *Applied Soft Computing*. 73 : 1053-1067.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Eilers, P.H.C. and Marx, B.D. 1996. "Flexible Smoothing with B-Splines and Penalties." *Statistical Science*. 11(2) : 89-102.

Eubank, R.L. 1988. **Spline Smoothing and Nonparametric Regression.** New York : Marcel Dekker Inc.

Eubank, R.L. 1999. **Nonparametric Regression and Spline Smoothing.** New York : Marcel Dekker Inc.

Green, P.J. and Silverman, B.W. 1994. **Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach.** London : Chapman and Hall.

Hastie , T.J. and Tibshirani , R.J. 1990. **Generalized Additive Models.** London : Chapman and Hall.

Lee, T.C.M. 2003. "Smoothing Parameter Selection for Smoothing Spline: a Simulation Study." *Computational Statistical & Data Analysis.* 42 : 139-148.

Nadaraya, A. 1964. "On estimating regression." *Theory of Probability and Its Application*. 9(1) : 141-142.

Ruppert, D. Wand, M.P. and Carroll, R.J. 2003. **Semiparametric Regression.** New York : Cambridge University Press.

Stone, M. 1974. "Cross-validation Choice and Assessment of Statistical Predictions." *Journal of the Royal Statistical Society Series B*. 36(2) : 111-147.

Wahba, G. 1990. **Spline Models for Observation Data.** Philadelphia Pennsylvania : SIAM.

Watson, G.S. 1964. "Smooth regression analysis." *Sankhya: The Indian Journal of Statistics.* 26(4) : 359-372.

Yang, L. and Hong, Y. 2017. "Adaptive Penalized Splines for Data Smoothing." *Journal of Computational Statistics and Data Analysis.* 108 : 77-83.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก
คำสั่งโปรแกรมอาร์ที่ใช้ในงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการประมาณค่าพารามิเตอร์ปรับให้เรียบด้วยวิธีการถดถอยไม่อิงพารามิเตอร์

```
#สัญลักษณ์และความหมาย
#n=ขนาดตัวอย่างเท่ากับ 50, 100, 150 และ 200
#m=จำนวนรอบที่จำลองข้อมูล 500 รอบ
#sd=ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่ากับ 1, 3, 5 และ 7
#x.t=ตัวแปรอิสระของข้อมูลอนุกรมเวลา
#y.t=ตัวแปรตามของข้อมูลอนุกรมเวลา
#err=ความคลาดเคลื่อน
#df=ค่านอต
#sse=ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE)
#df.sse=ค่านอตที่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด
#min.sse=ค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE)
#-------------------------------------------------------------------------------------------------------
set.seed(7)
#ดาวน์โหลดแพ็คเกจ splines, SemiPar
library(splines)
library(SemiPar)
#-------------------------------------------------------------------------------------------------------
#การกำหนดตัวแปร
n=50
m=500
sd=1
x.t=c(1:n)
y.t=c()
err=c()
#วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
df1=c()
sse1=c()
df.sse1=c()
min.sse1=c()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
df2=c()
sse2=c()
df.sse2=c()
min.sse2=c()
#วิธีเส้นโค้งบี (B-Spline Method)
df3=c()
sse3=c()
df.sse3=c()
min.sse3=c()
#วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
df4=c()
sse4=c()
df.sse4=c()
min.sse4=c()
#-------------------------------------------------------------------------------------------------------
#การประมาณค่าด้วยวิธีการถดถอยไม่อิงพารามิเตอร์ วนซ้ำ 500 รอบ
for(j in 1:m)
{
error=rnorm(n, mean=0, sd)
#-------------------------------------------------------------------------------------------------------
#ลักษณะข้อมูลอนุกรมเวลา
#1.ข้อมูลมีลักษณะเป็นแนวโน้ม (Trend)
y.t=(x.t)^0.5+error
#2.ข้อมูลมีลักษณะไม่เชิงเส้น (Non-linear)
#c=30
#y.t=3*(x.t)^0.5*sin(2*pi*(1+c)/(x.t+c))+error
#3.ข้อมูลมีลักษณะไม่คงที่ (Non-Stationary)
#y.t[1]<-rnorm(1, 0, 1)
#for(i in 2:n)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        #{
#y.t[i]<-y.t[i-1]+error[i]
        #}
#4.ข้อมูลมีลักษณะความแปรปรวนไม่คงที่ (Heteroscedasticity)
#y=sqrt(x.t)*sin((2*pi*(1+x.t^0.5)))
#y.t=y+error
#plot.ts(y.t, main="Time Series data")
#-------------------------------------------------------------------------------------------------------------------
#วิธีการถดถอยไม่อิงพารามิเตอร์
#1.วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
for(k in 5:n)
        {
fit1<-smooth.spline(x.t,y.t,df=k)
sse1[k-4]=sum((fitted(fit1)-y.t)^2)/n
        }
df.sse=which.min(sse1)
df.sse1[j]=df.sse+4
min.sse1[j]=min(sse1)
#2.วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
for(k in 5:n)
        {
Fit2<-lm(y.t~ns(x.t, df=k))
y.est=predict(fit2)
sse2[k-4]=mean((y.t-y.est)^2)
        }
df2[j]=which.min(sse2)
df.sse2[j]=df2[j]+4
min.sse2[j]=min(sse2)
#3.วิธีเส้นโค้งบี (B-Spline Method)
for(k in 5:n)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        {
fit3<-lm(y.t~bs(x.t, df=k))
y.est=predict(fit3)
sse3[k-4]=mean((y.t-y.est)^2)
        }
df3[j]=which.min(sse3)
df.sse3[j]=df3[j]+4
min.sse3[j]=min(sse3)
#4.วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
kn<-default.knots(x.t,99)
for(d in 5:n)
        {
Fit4<-spm(y.t~f(x.t,knots=kn,df=d))
df4[d-4]=d
residuals2=residuals(fit4)
sse4[d-4]=mean(residuals2^2)
        }
df4[j]=which.min(sse4)
df.sse4[j]=df4[j]+4
min.sse4[j]=min(sse4)
#-------------------------------------------------------------------------------------------------------
cat(c("loop:",j),fill=T)
}
#-------------------------------------------------------------------------------------------------------
#การหาค่า AMSE และค่า Knot ที่ AMSE ต่ำสุด
#1.วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
m.sse1=mean(min.sse1)
m.df1=round(mean(df.sse1))
#2.วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
m.sse2=mean(sse2)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
m.df2=round(mean(df.sse2))
#3.วิธีเส้นโค้งบี (B-Spline Method)
m.sse3=mean(min.sse3)
m.df3=round(mean(df.sse3))
#4.วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
m.sse4=mean(min.sse4)
m.df4=round(mean(df.sse4))
```

การพยากรณ์ข้อมูลจริง

```
#สัญลักษณ์และความหมาย
#n=ขนาดตัวอย่างเท่ากับ 200
#m=ขนาดตัวอย่างที่ใช้พยากรณ์
#x.t=ตัวแปรอิสระของข้อมูลอนุกรมเวลา
#y.t=ตัวแปรตามของข้อมูล ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาเท่ากับ
170 วัน และปริมาณการนำเข้าไฟฟ้าเท่ากับ 188 เดือน
#df=ค่านอต
#sse=ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE)
#df.sse=ค่านอตที่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด
#min.sse=ค่าเฉลี่ยคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE)
#MAPE=ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย
#forecast=การพยากรณ์
#x1=ตัวแปรอิสระที่ใช้ในการพยากรณ์
#y1=ตัวแปรตามที่ใช้ในการพยากรณ์ ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
เท่ากับ 30 วัน และปริมาณการนำเข้าไฟฟ้าเท่ากับ 12 เดือน
#x2=ตัวแปรอิสระที่ได้จากการพยากรณ์
#y.t_fore=ตัวแปรตามที่ได้จากการพยากรณ์
#-----------------------------------------------------------------------------------------------------------
#การนำเข้าไฟล์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
data=read.table("Exchange.txt", header=T)
y.t=as.numeric(data[,"exchange"])
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
ts.plot(y.t,xlab="Day", ylab="Baht/USD", main="Exchange Rate of Baht to USD")
#การนำเข้าไฟล์ข้อมูลปริมาณการนำเข้าไฟฟ้า
#data=read.table("Import_of_Electricity.txt", header=T)
#y.t=as.numeric(data[,"Importofelectricity"])
#ts.plot(y.t, xlab="Month", ylab="GWH", main="Import of Electricity")
#-----------------------------------------------------------------------------------------------------------------------
#ดาวน์โหลดแพ็คเกจ splines, SemiPar
library(splines)
library(SemiPar)
#-----------------------------------------------------------------------------------------------------------------------
#การกำหนดตัวแปร
n=length(y.t)
x.t=c(1:n)
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
m=n-30
#ปริมาณการนำเข้าไฟฟ้า
#m=n-12
x1=x.t[1:m]
y1=y.t[1:m]
#1.วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
sse1=c()
#2.วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
sse2=c()
#3.วิธีเส้นโค้งบี (B-Spline Method)
sse3=c()
#4.วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
sse4=c()
df4=c()
#-----------------------------------------------------------------------------------------------------------------------
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#การประมาณค่าและการพยากรณ์
#1.วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
#การประมาณค่า
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
for(k in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
#for(k in 5:188)
        {
fit1<-smooth.spline(x.t,y.t,df=k)
sse1[k-4]=sum((fitted(fit1)-y.t)^2)/n
        }
#การพยากรณ์
forecast1=fitted(fit1)[(m+1):n]
y.t_fore=y.t[(m+1):n]
MAPE1=mean(abs((y.t_fore-forecast1)/y.t_fore))*100
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
for(k in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
#for(k in 5:188)
        {
fit1<-smooth.spline(x1,y1,df=k)
sse1[k-4]=sum((fitted(fit1)-y1)^2)/m
        }
df.sse=which.min(sse1)
df.sse1=df.sse+4
min.sse1=min(sse1)
df.sse1
min.sse1
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
plot(y1, xlab="Day", ylab="Baht/USD", main ="SS")
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#ปริมาณการนำเข้าไฟฟ้า
#plot(y1, xlab="Month", ylab="GWH", main="SS")
lines(fitted(fit1))
#2.วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
#การประมาณค่า
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
for(k in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
#for(k in 5:188)
        {
fit2<-lm(y.t~ns(x.t, df=k))
y.est=predict(fit2)
sse2[k-4]=mean((y.t-y.est)^2)
        }
#การพยากรณ์
forecast2=fitted(fit2)[(m+1):n]
y.t_fore=y.t[(m+1):n]
MAPE2=mean(abs((y.t_fore-forecast2)/y.t_fore))*100
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
for(k in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
#for(k in 5:188)
        {
fit2<-lm(y1~ns(x1, df=k))
y.est=predict(fit2)
sse2[k-4]=mean((y1-y.est)^2)
        }
df.sse=which.min(sse2)
df.sse2=df.sse+4
min.sse2=min(sse2)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
df.sse2
min.sse2
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
plot(y1, xlab="Day", ylab="Baht/USD", main ="NS")
#ปริมาณการนำเข้าไฟฟ้า
#plot(y1, xlab="Month", ylab="GWH", main="NS")
lines(fitted(fit2))
#3.วิธีเส้นโค้งบี (B-Spline Method)
#การประมาณค่า
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
for(k in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
#for(k in 5:188)
        {
fit3<-lm(y.t~bs(x.t, df=k))
y.est=predict(fit3)
sse3[k-4]=mean((y.t-y.est)^2)
        }
#การพยากรณ์
forecast3=fitted(fit3)[(m+1):n]
y.t_fore=y.t[(m+1):n]
MAPE3=mean(abs((y.t_fore-forecast3)/y.t_fore))*100
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
for(k in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
#for(k in 5:188)
        {
fit3<-lm(y1~bs(x1, df=k))
y.est=predict(fit3)
sse3[k-4]=mean((y1-y.est)^2)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        }
df.sse=which.min(sse3)
df.sse3=df.sse+4
min.sse3=min(sse3)
df.sse3
min.sse3
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
plot(y1, xlab="Day", ylab="Baht/USD", main="BS")
#ปริมาณการนำเข้าไฟฟ้า
#plot(y1, xlab="Month", ylab="GWH", main="BS")
lines(fitted(fit3))
#4.วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
#การประมาณค่า
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
kn<-default.knots(x.t,169)
for(d in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
# kn<-default.knots(x.t,187)
#for(k in 5:188)
        {
fit4<-spm(y.t~f(x.t, knots=kn, df=d))
df4[d-4]=d
residuals2=residuals(fit4)
sse4[d-4]=mean(residuals2^2)
        }
#การพยากรณ์
forecast4=fitted(fit4)[(m+1):n]
y.t_fore=y.t[(m+1):n]
MAPE4=mean(abs((y.t_fore-forecast4)/y.t_fore))*100
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
kn<-default.knots(x.t,169)
for(d in 5:170)
#ปริมาณการนำเข้าไฟฟ้า
# kn<-default.knots(x.t,187)
#for(k in 5:188)
        {
fit4<-spm(y.t~f(x.t,knots=kn,df=d))
df4[d-4]=d
residuals2=residuals(fit4)
sse4[d-4]=mean(residuals2^2)
        }
df.sse=which.min(sse4)
df.sse4=df.sse+4
min.sse4=min(sse4)
df.sse4
min.sse4
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
plot(y1, xlab="Day", ylab="Baht/USD", main="PS")
#ปริมาณการนำเข้าไฟฟ้า
#plot(y1, xlab="Month", ylab="GWH", main="PS")
lines(fitted(fit4))
#-------------------------------------------------------------------------------------------------------
#การหาค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย
#1.วิธีเส้นโค้งเรียบ (Smoothing Spline Method)
MAPE1
#2.วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)
MAPE2
#3.วิธีเส้นโค้งบี (B-Spline Method)
MAPE3
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#4.วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method)
MAPE4
#-------------------------------------------------------------------------------------------------------------------
#การสร้างกราฟการพยากรณ์ข้อมูล
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
x2=seq(1,30)
plot(x2,y.t_fore, xlab="Day", ylab="Baht/USD", main="Exchange Rate of Baht to USD")
#ปริมาณการนำเข้าไฟฟ้า
#x2=seq(1,12)
#plot(x2,y.t_fore, xlab="Month", ylab="GWH", main="Import of Electricity",
ylim=c(2000,3400))
lines(x2, forecast1, lty=2, col="#0000FF")
lines(x2, forecast2, lty=3, col="#FF0000")
lines(x2, forecast3, lty=4, col="#FFA500")
lines(x2, forecast4, lty=5, col="#00FF00")
#อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา
legend("topright", c("SS","NS","BS","PS"), lty=c(2,3,4,5),
col=c("#0000FF","#FF0000","#FFA500","#00FF00"))
#ปริมาณการนำเข้าไฟฟ้า
#legend("topleft", c("SS","NS","BS","PS"), lty = c(2,3,4,5),
col=c("#0000FF","#FF0000","#FFA500","#00FF00"))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-สกุล	นางสาวบุษราพรรณ กันธรรม
วัน-เดือน-ปีเกิด	5 ตุลาคม 2531
ที่อยู่ปัจจุบัน	72 หมู่ 5 ต.แม่แรง อ.ป่าซาง จ.ลำพูน 51120
ประวัติการศึกษา	พ.ศ. 2554 วิทยาศาสตรบัณฑิต สาขาวิชาสถิติ มหาวิทยาลัยเชียงใหม่ พ.ศ. 2565 วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติและการวิเคราะห์ธุรกิจ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ผลงานทางวิชาการ	1. บุษราพรรณ กันธรรม และ อัชฌา อระวีพร. 2565. “การเปรียบเทียบวิธีเส้นโค้งกำลังสามธรรมชาติและวิธีเส้นโค้งบีของตัวแบบการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอนุกรมเวลา” เผยแพร่ใน Proceedings ของการประชุมวิชาการระดับชาติ วิทยาศาสตร์ เทคโนโลยี และนวัตกรรม มหาวิทยาลัยแม่โจ้ ครั้งที่ 3 (CSTI-MJU2022) 2. บุษราพรรณ กันธรรม และ อัชฌา อระวีพร. 2565. “การเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและปริมาณการนำเข้าไฟฟ้า” จัดพิมพ์ลงในวารสารวิชาการพระจอมเกล้าพระนครเหนือ ปีที่ 34 ฉบับที่ 2 เดือนเมษายน-มิถุนายน 2567
สถานที่ทำงาน	สำนักงานสถิติแห่งชาติ
ประสบการณ์การทำงาน	1. ธันวาคม 2564 ถึงปัจจุบัน นักวิชาการสถิติปฏิบัติการ กองสถิติสาธารณมติ สำนักงานสถิติแห่งชาติ 2. สิงหาคม 2560 - ธันวาคม 2564 นักวิชาการสถิติปฏิบัติการ ศูนย์สารสนเทศยุทธศาสตร์ภาครัฐ สำนักงานสถิติแห่งชาติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้