



บทความวิจัย

การเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและปริมาณการนำเข้าไฟฟ้า

บุษราพรรณ กันธธรรม* และ อัชฌา อระวีพร

ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 08 7188 1782 อีเมล: boosraphan@gmail.com DOI: 10.14416/j.kmutnb.2023.03.001

รับเมื่อ 26 ธันวาคม 2564 แก้ไขเมื่อ 9 พฤษภาคม 2565 ตอรับเมื่อ 12 พฤษภาคม 2565 เผยแพร่ออนไลน์ 3 มีนาคม 2566

© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

บทคัดย่อ

ในการวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการถดถอยไม่อิงพารามิเตอร์ 3 วิธี สำหรับข้อมูลอนุกรมเวลา ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมดา และวิธีเส้นโค้งบี โดยพิจารณาประสิทธิภาพของวิธีการถดถอยไม่อิงพารามิเตอร์จากค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ซึ่งวิธีเหล่านี้จะใช้ค่าพารามิเตอร์ปรับให้เรียบเพื่อควบคุมประสิทธิภาพการปรับให้เรียบของเส้นโค้ง โดยวิธีการแบ่งข้อมูลเป็นหลายส่วน และยังมีการกำหนดค่านอตเพื่อให้เส้นโค้งใกล้เคียงกับข้อมูลมากที่สุด โดยใช้อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา ตั้งแต่วันที่ 10 พฤศจิกายน 2563 ถึงวันที่ 9 กันยายน 2564 เป็นข้อมูลรายวันจำนวน 200 วัน และปริมาณการนำเข้าไฟฟ้า (กิโลวัตต์ชั่วโมง) ตั้งแต่เดือนเมษายน 2547 ถึงเดือนธันวาคม 2563 เป็นข้อมูลรายเดือนจำนวน 200 เดือน และนำตัวแบบการถดถอยไม่อิงพารามิเตอร์ที่มีประสิทธิภาพไปใช้ในการพยากรณ์ล่วงหน้า โดยมีเกณฑ์ คือค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด จากผลการประมาณค่าและพยากรณ์ทั้ง 2 ข้อมูล พบว่า ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยของวิธีเส้นโค้งกำลังสามธรรมดามีค่าต่ำสุด จึงเป็นวิธีที่มีประสิทธิภาพมากที่สุด และเมื่อพิจารณาค่านอตพบว่ายิ่งกำหนดค่านอตมากจะทำให้ค่าความคลาดเคลื่อนกำลังสองมีค่าลดลง นั่นหมายความว่าค่านอตมีผลต่อประสิทธิภาพของวิธีสถิติไม่อิงพารามิเตอร์

คำสำคัญ: การถดถอยไม่อิงพารามิเตอร์ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมดา วิธีเส้นโค้งบี

การอ้างอิงบทความ: บุษราพรรณ กันธธรรม และ อัชฌา อระวีพร, “การเปรียบเทียบวิธีการถดถอยไม่อิงพารามิเตอร์สำหรับข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและปริมาณการนำเข้าไฟฟ้า,” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 34, ฉบับที่ 1, หน้า 1-11, เลขที่บทความ 241-185709, ม.ค.-มี.ค. 2567.



Comparing Nonparametric Regression Methods for Exchange Rate of Thai Baht against the US Dollar and Volume Import of Electricity Data

Boosaraphan Kanthom* and Autcha Araveeporn

Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

* Corresponding Author, Tel. 08 7188 1782, E-mail: boosaraphan@gmail.com DOI: 10.14416/j.kmutnb.2023.03.001

Received 26 December 2021; Revised 9 May 2022; Accepted 12 May 2022; Published online: 3 March 2023

© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

Abstract

This research aims to compare the efficiency of three nonparametric regression methods for time series data, namely smoothing spline, natural cubic spline, and B-spline methods. The efficiency of the nonparametric regression method was determined by the lowest mean squared error. Smoothing parameter selection through cross-validation was conducted to control the smoothing performance of the curves. The number of knots were defined for fitting curves to measured data. We employed daily exchange rate of Thai Baht (THB) to US dollar (USD) from November 10, 2020 to September 9, 2021, for a total of 200 days. We also employed monthly electricity import volume (GWH) from April 2004 to December 2020, 200 months in total. Then, the efficient nonparametric regression model was applied to predict future values using the lowest mean absolute percentage error criterion. As results, the mean square error and mean absolute percentage error of the natural cubic spline method presented the lowest values, making it the most efficient predictive model. For the number of knots, the mean square error decreased as we increased the number of knots. This implied the effects of the number of knots on nonparametric regression performance.

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

Please cite this article as: B. Kanthom and A. Araveeporn, "Comparing nonparametric regression methods for exchange rate of Thai Baht against the US Dollar and volume import of electricity data," *The Journal of KMUTNB*, vol. 34, no. 1, pp. 1-11, ID. 241-185709, Jan.-Mar. 2024 (in Thai).

1. บทนำ

การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลเพื่อหาตัวแปรตาม (Dependent Variable) แทนด้วยสัญลักษณ์ y และตัวแปรอิสระ (Independent Variable) แทนด้วยสัญลักษณ์ x โดยวัตถุประสงค์เพื่อจะประมาณ หรือพยากรณ์ตัวแปรตามจากตัวแปรที่สร้างขึ้นที่เรียกว่าตัวแบบการถดถอย (Regression Model) หรือตัวแบบการถดถอยอิงพารามิเตอร์ (Parametric Regression Model) ในกระบวนการสร้างตัวแบบการถดถอย จะมีข้อสมมติเบื้องต้นคือ ตัวแปรอิสระอยู่ในมาตราวัดนามบัญญัติ (Nominal Scale) ขึ้นไป และตัวแปรตามอยู่ในมาตราวัดอันดับ (Interval Scale) ขึ้นไป โดยมีความสัมพันธ์กันเป็นเชิงเส้นตรง และมีการแจกแจงปกติ ความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่มีการแจกแจงปกติ และตัวแปรตามไม่มีความสัมพันธ์กันเอง (Non-Autocorrelation) นอกจากนี้ตัวแปรอิสระไม่เกิดความสัมพันธ์เชิงเส้นพหุ (Multicollinearity) แต่ในบางครั้งข้อมูลที่ได้อาจไม่เป็นไปตามข้อสมมติเบื้องต้น เช่น ตัวแปรตามและตัวแปรอิสระเกิดความสัมพันธ์ไม่เชิงเส้น (Non-Linear) หรือเป็นข้อมูลอนุกรมเวลา (Time Series) ที่มีรูปแบบที่ไม่แน่นอน ไม่เป็นเชิงเส้น และอาจเป็นเส้นโค้ง ทำให้ค่าความคลาดเคลื่อนสุ่มเกิดความสัมพันธ์กันเอง การใช้ตัวแบบการถดถอยอิงพารามิเตอร์อาจทำให้เกิดความคลาดเคลื่อนสูงในการประมาณค่าหรือพยากรณ์ตัวแปรตามได้ เพื่อแก้ปัญหาการวิเคราะห์การถดถอยที่ไม่ตรงตามข้อสมมติเบื้องต้น จึงมีการนำเอาวิธีการถดถอยไม่อิงพารามิเตอร์ (Nonparametric Regression) มาใช้ในการพยากรณ์ค่าของตัวแปรตาม

วิธีการถดถอยไม่อิงพารามิเตอร์ หรือเรียกว่าวิธีการปรับให้เรียบ (Smoothing Method) เป็นอีกทางเลือกหนึ่งในการแสดงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระในลักษณะไม่เชิงเส้น อาจแสดงในลักษณะเส้นโค้งโดยจะทำการสร้างเส้นโค้งให้เรียบไปตามการกระจายของข้อมูล เรียกว่าแผนภาพการกระจายปรับให้เรียบ (Scatterplot Smoothing) โดยวิธีการเหล่านี้พิจารณาการกระจายข้อมูลของตัวแปร

ตามเมื่อกำหนดตัวแปรอิสระเพื่อให้ได้การประมาณเส้นโค้งที่ใกล้เคียงกับการกระจายมากที่สุด [1] ซึ่งข้อมูลอาจจะมีลักษณะการกระจายไม่มีรูปแบบหรือเป็นโค้งหลายๆ โค้ง จึงต้องพิจารณาค่าพารามิเตอร์ปรับให้เรียบ (Smoothing Parameter) ซึ่งเป็นค่าที่ใช้ในการปรับเส้นโค้งโดยจะจำกัดหรือลดการเปลี่ยนแปลงขึ้นลงของข้อมูลเพื่อให้การประมาณค่าแทรกไปตามเส้นโค้งของข้อมูลและปรับให้เรียบขึ้น การปรับเส้นโค้งที่เกิดจากการเชื่อมต่อของเส้นโค้งต่างๆ ด้วยข้อต่อที่เรียกว่า นอต (Knot) ซึ่งเป็นเทคนิคของการปรับข้อมูลให้เรียบ ซึ่งวิธีการปรับให้เรียบมีหลายวิธีที่นิยมใช้ในปัจจุบันเช่น วิธีการถดถอยเคอร์เนล (Kernel Regression Method) [2], [3] วิธีเส้นโค้งเรียบ (Smoothing Spline Method) [4] วิธีเส้นโค้งที่ถูกลงโทษ (Penalized Spline Method) [5], [6] และวิธีการกรองแนวโน้ม (Trend Filtering Method) [7]

Wahba [4] ได้นำเสนอวิธีเส้นโค้งเรียบ (Smoothing Spline Method) เป็นวิธีที่นิยมใช้ในการปรับข้อมูลให้เรียบ ซึ่งใช้ประมาณเส้นโค้งพหุนามธรรมชาติ โดยการลดผลรวมกำลังสองของฟังก์ชันที่ถูกลงโทษ (Penalized) บนพารามิเตอร์ปรับให้เรียบ และมีนอตที่ไม่ซ้ำกันในแต่ละค่าของตัวแปรอิสระ เพื่อควบคุมการปรับข้อมูลให้เหมาะสม โดยใช้ค่าประมาณกำลังสองน้อยสุด (Least Squares Estimation) ในการลงโทษความหยาบของข้อมูล และยังมีการพัฒนาวิธีนี้โดย Green และ Silverman [8] เรียกว่าวิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method) โดยเป็นการพิจารณาการประมาณค่าจากความคลาดเคลื่อนกำลังสองต่ำสุดบนปริพันธ์ของการอนุพันธ์ลำดับที่สองยกกำลังสอง นอกจากนี้ยังมีวิธีเส้นโค้งอีกวิธีที่นิยมใช้คือวิธีเส้นโค้งบี (B-Spline Method) [9] เป็นวิธีการใช้ฟังก์ชันที่เรียกว่า ฟังก์ชันเส้นโค้งบี (B-Spline Function) สำหรับการประมาณค่าตัวแบบการถดถอยไม่อิงพารามิเตอร์

ในงานวิจัยนี้ผู้วิจัยสนใจศึกษาข้อมูลในรูปแบบอนุกรมเวลา เนื่องจากข้อมูล หรือค่าสังเกตมีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านไป ลักษณะของการเปลี่ยนแปลงอาจมีหรือ

ไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านไปในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใด โดยความสัมพันธ์ของข้อมูลได้จากแผนภาพการกระจายของข้อมูลที่ใช้ฉากแสดงในรูปแบบแนวโน้ม (Trend) วัฏจักร (Cycle) การเดินแบบสุ่ม (Random Walk) และข้อมูลที่ได้มาอาจเกิดสหสัมพันธ์ในตัว (Autocorrelation) เป็นภาวะที่ค่าสังเกตของอนุกรมเวลา ณ จุดเวลาที่อยู่ห่างกัน 2 จุด มีความสัมพันธ์กัน ก่อให้เกิดความคลาดเคลื่อนในการประมาณได้ ซึ่งการใช้วิธีการถดถอยไม่อิงพารามิเตอร์จะช่วยให้ความคลาดเคลื่อนในการพยากรณ์ค่าในอนาคตน้อยลงได้

ในบทความนี้ได้นำเสนอวิธีการปรับข้อมูลให้เรียบ 3 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ และวิธีเส้นโค้งบี เนื่องจากวิธีการเหล่านี้มีการใช้ค่าพารามิเตอร์ปรับให้เรียบ และค่าอนตมาใช้ในการประมาณค่าตัวแปรการถดถอยไม่อิงพารามิเตอร์ โดยกำหนดให้ตัวแปรอิสระคือลำดับข้อมูลของอนุกรมเวลา และตัวแปรตามคือข้อมูลอนุกรมเวลา โดยทำการเปรียบเทียบประสิทธิภาพของการประมาณค่า ด้วยค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error; MSE) และเปรียบเทียบประสิทธิภาพการพยากรณ์ ด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error; MAPE)

2. วัสดุ อุปกรณ์และวิธีการวิจัย

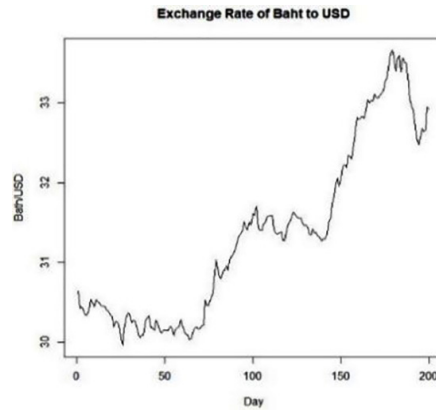
2.1 ขอบเขตของการศึกษา

การกำหนดขอบเขตของการศึกษา ดังนี้

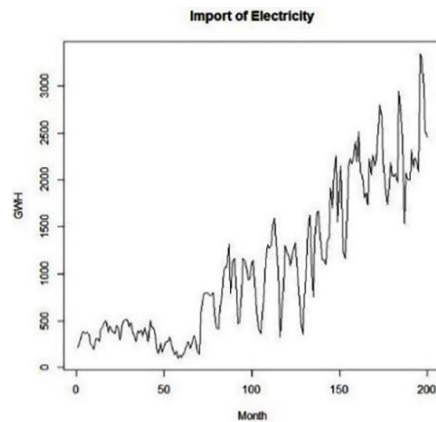
2.1.1 กำหนดตัวแปรอิสระ (x) ให้เป็นลำดับของข้อมูลอนุกรมเวลาเริ่มตั้งแต่ $1, 2, \dots, n$ เมื่อ n คือจำนวนข้อมูลทั้งหมด

2.1.2 พิจารณาตัวแปรตาม (y) จากข้อมูลจริง 2 ชุด ในรูปแบบข้อมูลอนุกรมที่มีลักษณะเป็นแนวโน้ม (Trend) ไม่เชิงเส้น (Non-Linear) เป็นฤดูกาล (Seasonal) และข้อมูลมีลักษณะไม่คงที่ (Non-Stationary) หรือการเดินแบบสุ่ม (Random Walk) ได้แก่

- อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา [10] ซึ่งเป็นข้อมูลรายวันทั้งหมด 200 วัน



รูปที่ 1 อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา รายวัน



รูปที่ 2 ปริมาณการนำเข้าไฟฟ้าของประเทศไทยรายเดือน

โดยแบ่งข้อมูลในการประมาณค่า 170 วัน เนื่องจากข้อมูลเป็นรายวันจึงทำการแบ่งข้อมูลสำหรับพยากรณ์ในช่วงเวลา 30 วัน หรือ 1 เดือน เพื่อดูแนวโน้มของข้อมูลที่มีลักษณะที่ผันแปรตลอดเวลาหรือเรียกว่ามีการเดินแบบสุ่มที่ข้อมูลมีลักษณะการเคลื่อนไหวที่ไม่สามารถคาดเดาได้ แสดงดังรูปที่ 1

- ปริมาณการนำเข้าไฟฟ้าของประเทศไทย (กิกะวัตต์ ชั่วโมง) [11] ซึ่งเป็นข้อมูลรายเดือนทั้งหมด 200 เดือน โดยแบ่งข้อมูลในการประมาณค่า 188 เดือน เนื่องจากข้อมูลเป็นรายเดือนจึงทำการแบ่งข้อมูลสำหรับการพยากรณ์ในช่วง 12 เดือน หรือ 1 ปี เพื่อดูแนวโน้มและฤดูกาลของข้อมูลที่อาจเกิดเป็นประจำ โดยพิจารณาจากที่ช่วงหน้าร้อนของทุกปีจะมีการนำเข้าไฟฟ้าในปริมาณมากกว่าช่วงเวลาอื่น แสดงดังรูปที่ 2

2.1.3 การกำหนดจำนวนค่าออก เพื่อให้เกิดข้อต่อในการสร้างเส้นโค้งเชื่อมต่อๆ กันให้ใกล้ข้อมูลมากที่สุด โดยพิจารณาขอบเขตของนอตตามขนาดตัวอย่าง เนื่องจากการกำหนดค่าออกน้อย จะทำให้เส้นโค้งไม่ปรับไปตามข้อมูลหรือมากเกินไปจนทำให้ใช้ระยะเวลาในการประมวลผลนานเกินความจำเป็น

2.1.4 ประมวลค่าพารามิเตอร์ปรับให้เรียบ (λ) เพื่อใช้ในการปรับเส้นโค้ง โดยจะจำกัด หรือลดการเปลี่ยนแปลงขึ้นลงของข้อมูลในฟังก์ชัน เพื่อให้การประมาณค่าแทรกไปตามเส้นโค้งของข้อมูลและปรับให้เรียบขึ้น โดยใช้วิธีการแบ่งข้อมูลเป็นหลายส่วน (Cross-validation Method) [12]

2.1.5 วิเคราะห์ข้อมูลจากวิธีการของสถิติไม่อิงพารามิเตอร์ 3 วิธี ได้แก่ วิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมชาติ และวิธีเส้นโค้งบี โดยใช้โปรแกรมอาร์ เวอร์ชัน 4.1.1

2.1.6 เกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าของสถิติไม่อิงพารามิเตอร์ คือค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และประสิทธิภาพของการพยากรณ์โดยวิธีการของสถิติไม่อิงพารามิเตอร์ ด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ดังนี้

- ค่าความคลาดเคลื่อนกำลังสองเฉลียดังสมการที่ (1)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, i = 1, 2, \dots, n \quad (1)$$

เมื่อ y_i คือ ค่าจริงของตัวแปรตามที่ i

\hat{y}_i คือ ค่าประมาณของตัวแปรตามที่ i ที่ได้จากตัวแบบการถดถอยไม่อิงพารามิเตอร์

วิธีที่มีประสิทธิภาพในการประมาณค่าที่ดีที่สุดจะให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดในแต่ละวิธี

โดยข้อมูลชุดที่ 1 $n = 170$ และข้อมูลชุดที่ 2 $n = 188$

- ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลียดังสมการที่ (2)

$$MAPE = \frac{1}{m} \sum_{j=1}^m \frac{|y_j - \hat{y}_j|}{y_j} \times 100, j = 1, 2, \dots, m \quad (2)$$

เมื่อ y_j คือ ค่าจริงของตัวแปรตามที่ j

\hat{y}_j คือ ค่าพยากรณ์ของตัวแปรตามที่ j ที่ได้จากตัวแบบการถดถอยไม่อิงพารามิเตอร์

วิธีที่มีประสิทธิภาพในการพยากรณ์ที่ดีที่สุดจะให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุดในแต่ละวิธี

โดยชุดข้อมูลที่ 1 $m = 30$ และข้อมูลชุดที่ 2 $m = 12$

2.2 วัสดุ อุปกรณ์และวิธีการวิจัย

การถดถอยไม่อิงพารามิเตอร์พิจารณารูปแบบความสัมพันธ์จากฟังก์ชันตัวแปรอิสระ ($f(x_i)$) และตัวแปรตาม (y) โดยมีตัวแบบการถดถอยไม่อิงพารามิเตอร์ดังสมการที่ (3)

$$y_i = f(x_i) + \varepsilon_i, i = 1, \dots, n \quad (3)$$

เมื่อ ε_i คือ ค่าความคลาดเคลื่อนสุ่มของค่าสังเกต

โดยในงานวิจัยนี้ต้องการประมาณค่าของฟังก์ชันของตัวแปรอิสระ 3 วิธี ได้แก่

2.2.1 วิธีเส้นโค้งเรียบ (Smoothing Spline Method)

การประมาณค่าของวิธีเส้นโค้งเรียบ เป็นการประมาณค่าจากการทำให้เกิดค่าต่ำสุดของฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุด (Penalized Least Square; PLS) ให้เหมาะสมกับฟังก์ชันของตัวแปรอิสระ ($f(x_i)$) ดังสมการที่ (4)

$$PLS = \sum_{i=1}^n \{y_i - f(x_i)\}^2 + \lambda \int_a^b \{f''(x_i)\}^2 dx \quad (4)$$

เมื่อ $RSS = \sum_{i=1}^n \{y_i - f(x_i)\}^2$ คือ ผลบวกกำลังสองของส่วนเหลือและ $\lambda \int_a^b \{f''(x_i)\}^2 dx$ คือ ความหยาบที่ถูกลงโทษ (Roughness Penalty) ในช่วง $[a, b]$ โดย λ คือ ค่าพารามิเตอร์ปรับให้เรียบหรือ λ การอนุพันธ์อันดับสองของความหยาบที่ถูกลงโทษ เรียกว่าเส้นโค้งเรียบกำลังสาม (Cubic Smoothing Spline) [2]

Wu and Zhang [13] แสดงการเขียนความหยาบที่ถูกลงโทษในรูปแบบเมทริกซ์ ได้ดังสมการที่ (5)

$$\lambda \int_a^b \{f''(x_i)\}^2 dx = \mathbf{f}^T \mathbf{G} \mathbf{f} \quad (5)$$



โดยกำหนดให้ $\mathbf{f} = (f_1, \dots, f_K)^T$

เมื่อ $f_r = f(\tau_r), r = 1, 2, \dots, K$ โดยทั่วไป K คือจำนวนนอต และ τ_1, \dots, τ_K เป็นนอตทั้งหมดของเส้นโค้งเรียบที่สามารถจัดเรียงตามลำดับที่เพิ่มขึ้นเป็น

$$-\infty < a < \tau_1 < \tau_2 < \dots < \tau_K < b < \infty$$

$$\text{กำหนดให้ } h_r = \tau_{r-1} - \tau_r, \quad r = 1, 2, \dots, K-1$$

โดยเมทริกซ์ G มีขนาด $K \times K$ ดังสมการที่ (6)

$$\mathbf{G} = \mathbf{A}\mathbf{B}^{-1}\mathbf{A}^T \quad (6)$$

เมื่อเมทริกซ์ $\mathbf{A} = (a_{rs})$ มีขนาด $K \times (K-2)$ โดยค่าอื่น ๆ จะมีค่าเป็น 0 ทั้งหมดยกเว้นที่ $r = 1, 2, \dots, K-2, a_{rr} = h_r^{-1}$ 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 $r = 1, 2, \dots, K-2, b_{r,r+1} = h_{(r+1)}/6, b_{r+1,r+1} = (h_{(r+1)} + h_{(r+2)})/3, b_{r+2,r+1} = h_{(r+2)}/6$ และ $b_{K-3,K-2} = h_{(K-2)}/6, b_{K-2,K-2} = (h_{(K-2)} + h_{(K-1)}/3$ ซึ่งเป็นไปตามเกณฑ์ฟังก์ชันที่ถูกกล่องโทซกำลังสองน้อยที่สุดจากสมการที่ (4) สามารถเขียนในรูปเมทริกซ์ได้ดังสมการที่ (7)

$$\|\mathbf{y} - \mathbf{W}\mathbf{f}\|^2 + \lambda \mathbf{f}^T \mathbf{G} \mathbf{f} \quad (7)$$

เมื่อ $\mathbf{y} = (y_1, \dots, y_n)^T$ คือ เวกเตอร์ของตัวแปรตาม $\mathbf{W} = (w_{ir})$ มีขนาด $n \times K$ เป็นเมทริกซ์ตกกระทบ (Incidence Matrix) กับ $w_{ir} = 1$ ถ้า $x_i = \tau_r$ และ $\|\mathbf{y} - \mathbf{W}\mathbf{f}\|^2 = \sum_{i=1}^n \{y_i - f(x_i)\}^2$ ดังนั้นฟังก์ชันเส้นโค้งเรียบ (\hat{f}_λ) จากนอต $\tau_r, r = 1, \dots, K$ ได้ดังสมการที่ (8)

$$\hat{f}_\lambda = (\mathbf{W}^T \mathbf{W} + \lambda \mathbf{G})^{-1} \mathbf{W}^T \mathbf{y} \quad (8)$$

2.2.2 วิธีเส้นโค้งกำลังสามธรรมชาติ (Natural Cubic Spline Method)

Hastie [1] ศึกษาตัวแบบเชิงบวกน้อยทั่วไป (Generalized Additive Model) เป็นวิธีที่สร้างตัวแบบเชิงบวกน้อยทั่วไป

โดยประยุกต์กับฟังก์ชันปรับให้เรียบซึ่งจะประมาณค่าโดยเส้นโค้งการถดถอยสำหรับการประมาณค่าเส้นโค้งกำลังสามธรรมชาติ

วิธีการสร้างเส้นโค้งกำลังสามธรรมชาติสร้างขึ้นเพื่อลดความไม่แน่นอนนี้ในการวางตัวเป็นเส้นตรงในขอบเขตของนอต โดยมีการกำหนดเงื่อนไข

$$\begin{aligned} N_1(x_i) &= 1, \quad N_1(x_i) = x_i \\ N_{k+2}(x_i) &= d_k(x_i) - d_{k-1}(x_i), \quad r = 1, \dots, K-2 \\ d_k(x_i) &= \frac{(x_i - \tau_r)_+^3 - (x_i - \tau_K)_+^3}{\tau_r - \tau_K} \end{aligned}$$

โดยสัญลักษณ์ w_+^k เมื่อ k แทนลำดับการยกกำลังของ w และ $w_+ = \max(0, w)$ ซึ่งสามารถเขียนฟังก์ชันของตัวแปรอิสระได้ดังสมการที่ (9)

$$f(x_i) = \sum_{r=1}^K N_r(x_i) \theta_r \quad (9)$$

เมื่อ $N_r(x_i)$ เป็นชุดฟังก์ชันขนาด N ของเส้นโค้งกำลังสามธรรมชาติ โดยที่ $\theta = (\theta_1, \dots, \theta_K)^T$ คือ เวกเตอร์สัมประสิทธิ์ของเส้นโค้งกำลังสามธรรมชาติ $N_r(x_i)$ กับจำนวนนอต ซึ่งเป็นค่าที่มีความเหมาะสมของฟังก์ชันเส้นโค้งกำลังสามธรรมชาติที่ค่านอต $\tau_r, r = 1, \dots, K$ ได้ดังสมการที่ (10)

$$\hat{f}(x_i) = \sum_{r=1}^K N_r(x_i) \hat{\theta}_r \quad (10)$$

ค่าสัมประสิทธิ์สามารถหาได้จากเกณฑ์เกิดค่าต่ำสุดของฟังก์ชันที่ถูกกล่องโทซกำลังสองน้อยที่สุดจากสมการที่ (4) ดังนี้

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

เมื่อ $\mathbf{y} = (y_1, \dots, y_n)^T, \{\mathbf{N}\}_{ir} = N_r(x_i)$ และ

$$\{\boldsymbol{\Omega}_K\}_{ir} = \int N_r^*(x) N_r^*(x) dx$$

ดังนั้นตัวประมาณของฟังก์ชันเส้นโค้งกำลังสามธรรมชาติ (\hat{f}_λ) ที่ค่านอต $\tau_r, r = 1, \dots, K$ โดยขึ้นอยู่กับการมีเตอร์ปรับให้เรียบดังสมการที่ (11)

$$\hat{f}_\lambda = (\mathbf{N}^T \mathbf{N} + \lambda \boldsymbol{\Omega}_K)^{-1} \mathbf{N}^T \mathbf{y} \quad (11)$$

2.2.3 วิธีเส้นโค้งบี (B-Spline Method)

สไปลน์ (Spline) เป็นตัวแบบพหุนามแบบแบ่งส่วน หรือเรียกว่าพหุนามทีละส่วน (Piecewise Polynomial) ซึ่งเป็นชิ้นส่วนของพหุนามที่มีคุณสมบัติแบ่งส่วนเป็นช่วงบนจุดนอต [9] โดยจุดนอต คือจุดที่บอกถึงการเปลี่ยนแปลงของข้อมูลในช่วงย่อยๆ นั้น การให้ชุดของนอต K เป็นฟังก์ชันพื้นฐานเส้นโค้งบีซ้ำๆ สามารถกำหนดได้โดยแสดงดังสมการที่ (12)

$$B_r^m(x) = \begin{cases} 1 & , \tau_r < x < \tau_{r+1} \\ 0 & \text{อื่นๆ} \end{cases} \quad (12)$$

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

De Boor [9] ได้แนะนำอัลกอริทึมในการคำนวณเส้นโค้งบีของฟังก์ชันพหุนามทีละส่วน ลำดับที่ m ของเส้นโค้งบีจะคำนวณจากลำดับที่ $(m-1)$ ดังสมการที่ (13)

$$B_r^m(x_i) = \frac{x_i - \tau_r}{\tau_{r+m-1} - \tau_r} B_r^{m-1} + \frac{\tau_{r+m} - x_i}{\tau_{r+m} - \tau_{r+1}} B_{r+1}^{m-1} \quad (13)$$

เมื่อลำดับที่ m พิจารณาจากเซตของ $\{B_r^m \mid r=1, \dots, K+m+1\}$ ดังนั้น ฟังก์ชันของตัวแปรอิสระของวิธีเส้นโค้งบีสามารถเขียนได้ดังสมการที่ (14)

$$f(x_i) = \sum_{r=1}^K B_r^m(x_i) \theta_r \quad (14)$$

ซึ่งค่าที่มีความเหมาะสมของฟังก์ชันเส้นโค้งบีประมาณที่ค่านอต $\tau_r, r=1, \dots, K$ ดังสมการที่ (15)

$$\hat{f}(x_i) = \sum_{r=1}^K B_r^m(x_i) \hat{\theta}_r \quad (15)$$

ค่าสัมประสิทธิ์สามารถหาได้จากเกณฑ์เกิดค่าต่ำสุดของฟังก์ชันที่ถูกลดโทษกำลังสองน้อยที่สุดจากสมการที่ (4) ดังนี้

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

เมื่อ $\{B_r^m\}_{ir} = B_r^m(x_i), \{\Omega_K\}_{ir} = \int B_r^m(x) B_s^m(x) dx$ และ $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$ คือ เวกเตอร์สัมประสิทธิ์ของเส้นโค้งบี ดังนั้นตัวประมาณของฟังก์ชันเส้นโค้งบี (\hat{f}_λ) ดังสมการที่ (16)

$$\hat{f}_\lambda = (\mathbf{B}^T \mathbf{B} + \lambda \boldsymbol{\Omega}_K)^{-1} \mathbf{B}^T \mathbf{y} \quad (16)$$

3. ผลการทดลอง

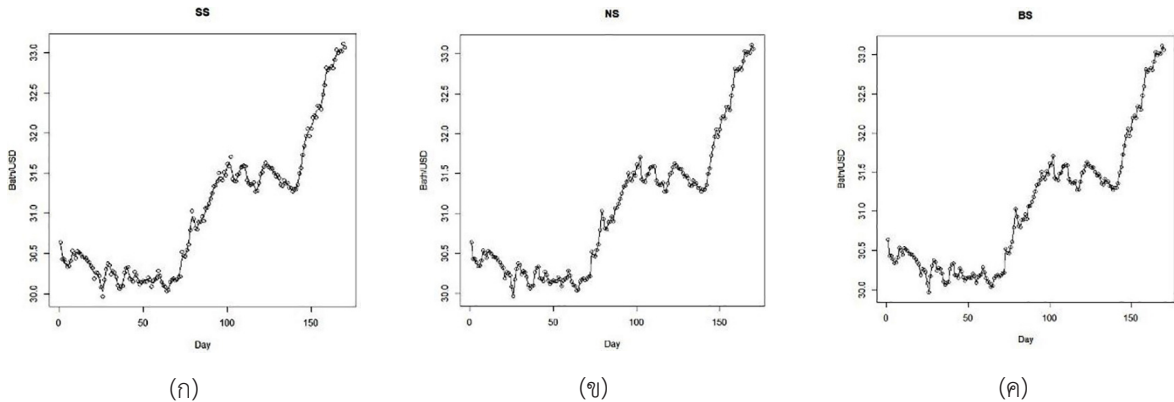
ข้อมูลจริงที่นำมาวิเคราะห์ ได้แก่ อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา ตั้งแต่วันที่ 10 พฤศจิกายน 2563 ถึงวันที่ 9 กันยายน 2564 ซึ่งเป็นข้อมูลรายวันทั้งหมด 200 วัน และปริมาณการนำเข้าไฟฟ้าของประเทศไทย (กิกะวัตต์ชั่วโมง) ตั้งแต่เดือนเมษายน 2547 ถึงเดือนธันวาคม 2563 ซึ่งเป็นข้อมูลรายเดือนทั้งหมด 200 เดือน

จากการวิเคราะห์ข้อมูลทั้ง 3 วิธี โดยให้ y เป็นอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของประเทศไทย ให้ x เป็นลำดับวันและเดือน ซึ่งได้มีการแทนค่านอตในแต่ละตัวแบบตั้งแต่ 5 จนกระทั่งได้ค่านอต และค่าพารามิเตอร์ปรับให้เรียบที่หาค่าที่ทำให้ได้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำสุดเพื่อให้ค่าเส้นโค้งที่ได้แทรกไปตามข้อมูล แสดงดังรูปที่ 3 และ 4 โดยค่าความคลาดเคลื่อนกำลังสองเฉลี่ยและค่านอตแสดงดังตารางที่ 1

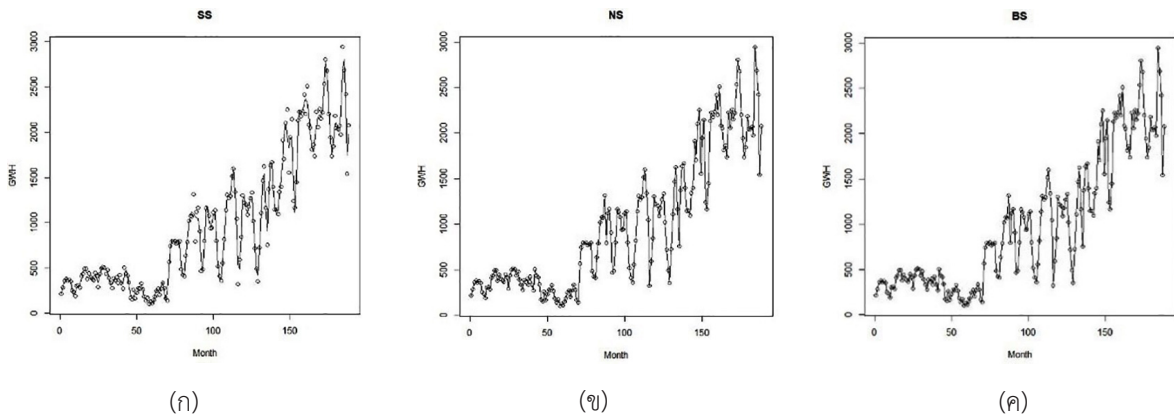
ตารางที่ 1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่านอต (Knot) ของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมชาติ (NS) และ วิธีเส้นโค้งบี (BS)

Method	Exchange Rate		Import of Electricity	
	MSE	Knot	MSE	Knot
SS	8.22×10^{-4}	89	6170	97
NS	8.27×10^{-28}	169	6.96×10^{-23}	187
BS	1.08×10^{-6}	169	37.60	187

เมื่อพิจารณาจากรูปที่ 3 และ 4 พบว่า กราฟค่าประมาณทั้งกรณีข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา (Exchange Rate) และข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย (Import of Electricity) โดยวิธีเส้นโค้งทั้ง 3 วิธี มีลักษณะคล้ายคลึงกัน ทำให้ยากต่อการสรุปผลว่าวิธีใดมีประสิทธิภาพที่สุด ดังนั้นจึงพิจารณา



รูปที่ 3 การประมาณค่าข้อมูลอัตราแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา (ก) วิธีเส้นโค้งเรียบ (SS) (ข) วิธีเส้นโค้งกำลังสามธรรมดา (NS) (ค) วิธีเส้นโค้งบี (BS)



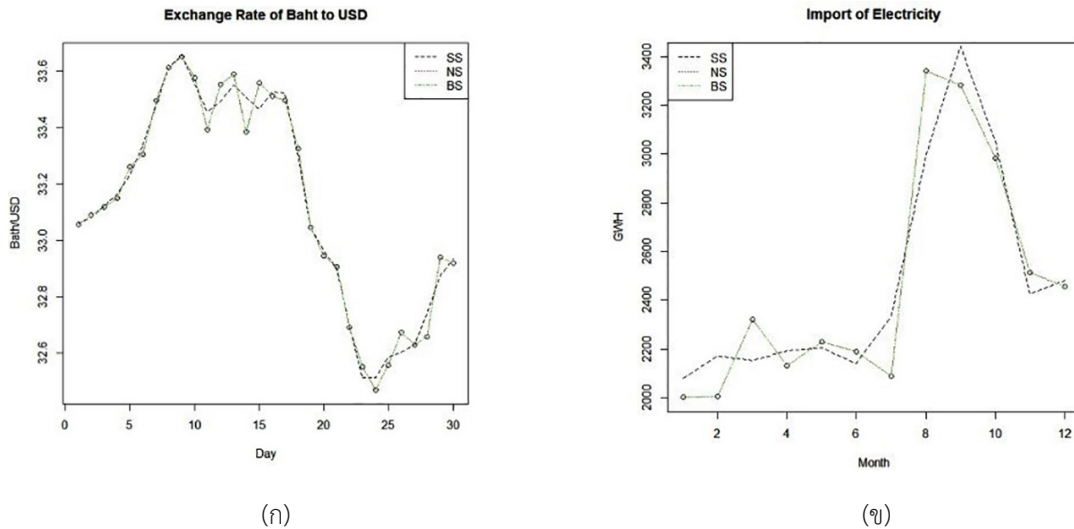
รูปที่ 4 การประมาณค่าข้อมูลการนำเข้าไฟฟ้าของประเทศไทย (ก) วิธีเส้นโค้งเรียบ (SS) (ข) วิธีเส้นโค้งกำลังสามธรรมดา (NS) (ค) วิธีเส้นโค้งบี (BS)

ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย โดยเมื่อพิจารณาจากตารางที่ 1 พบว่า วิธีเส้นโค้งกำลังสามธรรมดามีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดและมีค่าอนุตสูงสุด ทั้ง 2 ข้อมูล แสดงว่าวิธีเส้นโค้งกำลังสามธรรมดามีวิธีที่มีประสิทธิภาพที่สุดในการประมาณค่า

สำหรับข้อมูลปริมาณการนำเข้าไฟฟ้ามีความคลาดเคลื่อนกำลังสองเฉลี่ยที่แตกต่างกันค่อนข้างมาก เนื่องจากวิธีเส้นโค้งเรียบจะเน้นประมาณค่าฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุด (PLS) ให้เกิดค่าต่ำสุด วิธีเส้นโค้งกำลังสามธรรมดาที่พัฒนามาจากวิธีเส้นโค้งเรียบจะเน้นประมาณค่าจากผลรวมของฟังก์ชันปรับให้เรียบและหาค่า

ต่ำสุดของฟังก์ชันที่ถูกลงโทษกำลังสองน้อยที่สุด และวิธีเส้นโค้งบีจะเน้นประมาณค่าจากการแบ่งส่วนข้อมูลเป็นช่วงบนเส้นโค้ง นอกจากนี้วิธีเส้นโค้งเรียบและวิธีเส้นโค้งบีอาจจะไม่เหมาะสมกับลักษณะข้อมูลที่เป็นฤดูกาล จึงทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่ามาก

การพิจารณาประสิทธิภาพในการพยากรณ์ของข้อมูล ซึ่งจะเปรียบเทียบทั้งหมด 3 วิธี คือวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมดา (NS) และวิธีเส้นโค้งบี (BS) แสดงดังรูปที่ 5 โดยวิธีที่มีประสิทธิภาพในการพยากรณ์ที่สุดจะให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ต่ำที่สุด ดังตารางที่ 2



รูปที่ 5 การพยากรณ์ข้อมูล (ก) อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา (ข) ปริมาณการนำเข้าไฟฟ้าของประเทศไทย ด้วยวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมดา (NS) วิธีเส้นโค้งบี (BS)

เมื่อพิจารณาจากรูปที่ 5 พบว่า การพยากรณ์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย โดยวิธีเส้นโค้งกำลังสามธรรมดา และวิธีเส้นโค้งบีมีลักษณะคล้ายเป็นเส้นเดียวกัน ทำให้ยากต่อการสรุปผลว่าวิธีใดมีประสิทธิภาพที่สุด ดังนั้นจึงพิจารณาจากค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย โดยเมื่อพิจารณาจากตารางที่ 2 พบว่า วิธีเส้นโค้งกำลังสามธรรมดามีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุดทั้ง 2 ข้อมูล แสดงว่าวิธีเส้นโค้งกำลังสามธรรมดาเป็นวิธีที่มีประสิทธิภาพที่สุดในการพยากรณ์

เฉลี่ย (MAPE) ของวิธีเส้นโค้งกำลังสามธรรมดาต่ำกว่าวิธีอื่น แสดงว่าการพยากรณ์อัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกา และปริมาณการนำเข้าไฟฟ้าของวิธีเส้นโค้งกำลังสามธรรมดามีความแม่นยำสูง เนื่องจากทั้ง 2 ข้อมูล มีค่าอนตสูงที่สุดส่งผลให้การพยากรณ์ที่ได้โค้งไปตามเส้นโค้งใกล้เคียงกับข้อมูลมากที่สุด จึงทำให้มีค่าความคลาดเคลื่อนน้อย นอกจากนี้วิธีของเส้นโค้งกำลังสามธรรมดามีกระบวนการปรับเส้นโค้งที่ซับซ้อนกว่าวิธีอื่นๆ ทำให้วิธีนี้มีประสิทธิภาพในการประมาณและพยากรณ์มากที่สุด

ตารางที่ 2 ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ของวิธีเส้นโค้งเรียบ (SS) วิธีเส้นโค้งกำลังสามธรรมดา (NS) และวิธีเส้นโค้งบี (BS)

Method	Exchange Rate	Import of Electricity
SS	9.71×10^{-2}	4.92
NS	2.94×10^{-2}	6.80×10^{-9}
BS	3.50×10^{-2}	1.01×10^{-7}

จากตารางที่ 2 ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์

4. อภิปรายผลและสรุป

การวิเคราะห์ข้อมูลอัตราการแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐอเมริกาและข้อมูลปริมาณการนำเข้าไฟฟ้าของประเทศไทย ซึ่งอยู่ในรูปข้อมูลอนุกรมเวลาสำหรับการถดถอยไม่อิงพารามิเตอร์หรือเรียกว่าเทคนิคการปรับให้เรียบด้วยวิธีเส้นโค้งเรียบ วิธีเส้นโค้งกำลังสามธรรมดา และวิธีเส้นโค้งบี สำหรับการประมาณค่า และการพยากรณ์ พบว่าวิธีเส้นโค้งกำลังสามธรรมดามีประสิทธิภาพดีกว่าวิธีอื่นๆ โดยค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด และ



การพยากรณ์ด้วยวิธีเส้นโค้งกำลังสามธรรมชาติก็สามารถพยากรณ์ได้ดี โดยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด แสดงให้เห็นว่าการพยากรณ์ของวิธีเส้นโค้งกำลังสามธรรมชาติเป็นวิธีที่มีประสิทธิภาพกับข้อมูลที่มีลักษณะไม่เชิงเส้น มีแนวโน้มและฤดูกาลมากที่สุด เช่นเดียวกับการประมาณค่า ในส่วนค่านอตยถึงค่านอตยมีค่าสูงจะทำให้เส้นโค้งที่สร้างขึ้นเข้าใกล้ข้อมูลมากขึ้น สอดคล้องกับการศึกษาของ Elhakeem และคณะ [14] โดยวิธีเส้นโค้งกำลังสามธรรมชาติจะมีประสิทธิภาพดีในการประมาณค่าเมื่อข้อมูลมีลักษณะไม่เชิงเส้นและมีแนวโน้ม เนื่องจากสามารถปรับเส้นโค้งให้เข้ากับความสัมพันธ์ของข้อมูลได้ และควรกำหนดค่านอตยที่เหมาะสมใกล้เคียงกับขนาดตัวอย่าง ดังนั้นในการประมาณค่าการถดถอยไม่อิงพารามิเตอร์ถ้าทราบลักษณะข้อมูลและกำหนดค่านอตยที่เหมาะสมจะทำให้การประมาณค่ามีประสิทธิภาพดีขึ้น

นอกจากนี้การที่ข้อมูลมีลักษณะไม่เชิงเส้น และไม่ปฏิบัติตามข้อสมมติเบื้องต้นของการประมาณค่าการถดถอยนั้น การใช้ตัวแบบการถดถอยไม่อิงพารามิเตอร์จะช่วยให้การประมาณค่าได้ดีขึ้น ในครั้งต่อไปอาจศึกษาเพิ่มเติมเทคนิคการปรับให้เรียบวิธีอื่น กับข้อมูลอนุกรมเวลารูปแบบอื่น เพื่อให้ได้วิธีที่มีประสิทธิภาพในการประมาณค่าและการพยากรณ์ที่เหมาะสมกับลักษณะของข้อมูลต่อไป

5. กิตติกรรมประกาศ

ขอขอบพระคุณสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังที่ให้ทุนสนับสนุนในการวิจัยครั้งนี้

เอกสารอ้างอิง

- [1] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*, Chapman and Hall, London, 1990.
- [2] E. A. Nadaraya, "On estimating regression," *Theory of Probability and Its Application*, vol. 9, no. 1, pp. 141–142, 1964.
- [3] G. S. Watson, "Smooth regression analysis," *Sankhya: The Indian Journal of Statistics, Series A*, vol. 26, no. 4, pp. 359–372, 1964.
- [4] G. Wahba, *Spline Models for Observation Data*, SIAM, Philadelphia, Pennsylvania, 1990.
- [5] P. H. C. Eilers and B. D. Marx, "Flexible smoothing with b-splines and penalties," *Statistical Science*, vol. 11, no. 2, pp. 89–102, 1996.
- [6] D. Ruppert, M. P. Wand, and R. J. Carroll, *Semiparametric Regression*, Cambridge University Press, New York, 2003.
- [7] S. J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, " l_1 trend filtering," *SIAM Review*, vol. 51, no. 2, pp. 339–360, 2009.
- [8] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman and Hall, London, 1994.
- [9] C. De Boor, *A Practical Guide to Splines*, Springer, Berlin, 1978.
- [10] Bank of Thailand. (2021, Sep.). *Exchange Rate*. [Online] (in Thai). Available: <https://www.bot.or.th>
- [11] Electricity Generating Authority of Thailand. (2021, Sep.). *Import of Electricity*. [Online] (in Thai). Available: <https://www.egat.co.th>
- [12] M. Stone, "Cross-validation choice and assessment of statistical predictions," *Journal of the Royal Statistical Society Series B*, vol. 36, no. 2, pp. 111–147, 1974.
- [13] H. Wu and J. T. Zhang, *Nonparametric Regression Methods for Longitudinal Data Analysis*, A John Wiley & Sons, INC. Publication, New Jersey, USA, 2006.
- [14] A. Elhakeem, R. A. Hughes, K. Tilling, D. L. Cousminer,



S. A. Jackowski, T. J. Cole, A. S. F. Kwong, Z. Li, S. F. A. Grant, A. D. G. Baxter-Jones, B. S. Zemel, and D. A. Lawlor, “Using linear and natural cubic splines, SITAR, and latent

trajectory models to characterize nonlinear longitudinal growth trajectories in cohort studies,” *BMC Medical Research Methodology*, vol. 22, no.1, pp. 1–20, 2022.

