

การพยากรณ์จำนวนนักท่องเที่ยวและการจัดการข้อมูลอนุกรมเวลาที่ไม่ต่อเนื่อง เพื่อพัฒนาการท่องเที่ยวอย่างยั่งยืนในอุทยานแห่งชาติทางทะเล ภาคตะวันออกของประเทศไทย

Forecasting the Number of Tourists and Managing Interrupted Time Series Data for Sustainable Tourism Development in Thailand's Eastern Marine National Parks

จันทร์จิรา พิลาดัง, ภาคนันท์ ชุมประไพ และ นพรัตน์ แป้นงาม*

Janjira Piladaeng, Pakkanan Chumpraphai and Nopparat Panngam*

สาขาวิชาสถิติ ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ประเทศไทย

Statistics Major, Department of Mathematics, Faculty of Science, Burapha University, Thailand

Received : 19 May 2025, Received in revised form : 27 October 2025, Accepted : 28 October 2025

Available online : 7 November 2025

บทคัดย่อ

วัตถุประสงค์และที่มา : ประเทศไทยได้รับการยอมรับในระดับสากลว่าเป็นหนึ่งในจุดหมายปลายทางด้านการท่องเที่ยวที่สำคัญของโลก และการท่องเที่ยวยังถือเป็นปัจจัยสำคัญในการขับเคลื่อนเศรษฐกิจของประเทศอีกด้วย โดยอุตสาหกรรมการท่องเที่ยวไม่เพียงแต่เป็นแหล่งรายได้หลักของประเทศเท่านั้น แต่ยังเป็นกลไกสำคัญในการขับเคลื่อนการจ้างงานและเศรษฐกิจโดยรวม อย่างไรก็ตาม จากสถานการณ์การแพร่ระบาดของโควิด-19 ในช่วงปลายปี พ.ศ. 2562 ส่งผลให้อุตสาหกรรมการท่องเที่ยวภายในประเทศได้รับผลกระทบจากการใช้มาตรการควบคุมทางสาธารณสุขที่เข้มงวด เช่น การล็อกดาวน์ และการจำกัดการเดินทาง ส่งผลให้ไม่มีนักท่องเที่ยวต่างชาติเข้ามาในประเทศไทยตั้งแต่เดือนเมษายน พ.ศ. 2563 เป็นต้นมา ทั้งนี้ แหล่งท่องเที่ยวทางทะเลที่สำคัญของภาคตะวันออกที่ได้รับความนิยมอย่างอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ก็ได้รับผลกระทบจากการประกาศปิดการท่องเที่ยวในเขตอุทยานแห่งชาติและวนอุทยานเช่นเดียวกัน ภายหลังการเปิดประเทศในช่วงปลายปี พ.ศ. 2564 อุตสาหกรรมการท่องเที่ยวเริ่มมีการฟื้นตัวอย่างต่อเนื่อง ส่งผลให้จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่ง เพิ่มขึ้นอย่างชัดเจน ดังนั้น การพยากรณ์จำนวนนักท่องเที่ยวที่มีความแม่นยำ จึงเป็นสิ่งสำคัญในการช่วยหน่วยงานที่เกี่ยวข้องดำเนินการบริหารจัดการการท่องเที่ยวได้อย่างยั่งยืน ทั้งนี้ การพยากรณ์สามารถให้ผลลัพธ์ที่มีความแม่นยำถ้าข้อมูลที่นำมาวิเคราะห์มีความต่อเนื่องและสม่ำเสมอ แต่การแพร่ระบาดของโควิด-19 ส่งผลให้ไม่มีนักท่องเที่ยวเดินทางมายังอุทยานแห่งชาติในช่วงเวลาดังกล่าว จึงทำให้ข้อมูลบางช่วงเป็นศูนย์และเกิดความไม่ต่อเนื่อง ซึ่งอาจส่งผลต่อประสิทธิภาพของตัวแบบการพยากรณ์ ด้วยเหตุนี้ งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด เมื่อข้อมูลได้รับผลกระทบจากสถานการณ์การแพร่ระบาดของโควิด-19 โดยใช้เทคนิค

การทดแทนค่าสูญหายมาช่วยในการปรับปรุงคุณภาพข้อมูลให้มีความต่อเนื่อง พร้อมทั้งสร้างตัวแบบการพยากรณ์ที่เหมาะสมด้วยวิธีเชิงสถิติที่เป็นที่รู้จักและใช้กันอย่างแพร่หลายอย่างวิธีบ็อกซ์-เจนกินส์

วิธีดำเนินการวิจัย : ในการศึกษาครั้งนี้ใช้ข้อมูลจำนวนนักท่องเที่ยวที่เดินทางมาท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างและอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ตั้งแต่เดือนตุลาคม พ.ศ. 2558 ถึงเดือนมีนาคม พ.ศ. 2568 รวมทั้งสิ้น 114 เดือน จากกรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช โดยกำหนดรูปแบบของข้อมูลที่ใช้ในการสร้างตัวแบบการพยากรณ์ 3 รูปแบบ คือ 1. ใช้ข้อมูลจริงที่มีค่าศูนย์ 2. กำหนดค่าศูนย์เป็นข้อมูลสูญหาย และ 3. กำหนดข้อมูลช่วงการแพร่ระบาดของโควิด-19 ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนธันวาคม พ.ศ. 2564 เป็นข้อมูลสูญหาย และใช้เทคนิคการแยกส่วนประกอบตามฤดูกาลร่วมกับการประมาณค่าด้วยวิธีค่าเฉลี่ยในการแทนค่าข้อมูล ต่อมาสร้างตัวแบบการพยากรณ์โดยใช้ข้อมูลตั้งแต่เดือนตุลาคม พ.ศ. 2558 ถึงเดือนกันยายน พ.ศ. 2566 ด้วยวิธีบ็อกซ์-เจนกินส์ โดยใช้เกณฑ์สารสนเทศของอะกะอิเกะ (AIC) ในการเลือกตัวแบบที่เหมาะสมภายใต้ตัวแบบ SARIMA หลายตัวแบบที่ใช้ข้อมูลชุดเดียวกัน และใช้รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เป็นเกณฑ์ในการพิจารณาตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวที่เข้ามาท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด จากนั้นตรวจสอบความแม่นยำของการพยากรณ์โดยใช้ข้อมูลตั้งแต่เดือนตุลาคม พ.ศ. 2566 ถึงเดือนมีนาคม พ.ศ. 2568 พร้อมทั้งประเมินความแม่นยำของการพยากรณ์ด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)

ผลการวิจัย : การประมาณค่าข้อมูลอนุกรมเวลาที่มีความไม่ต่อเนื่องของจำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างและอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด โดยใช้การทดแทนค่าสูญหายด้วยวิธีการแยกส่วนประกอบฤดูกาลร่วมกับขั้นตอนวิธีค่าเฉลี่ย ภายใต้เงื่อนไขการสูญหายของข้อมูล 3 รูปแบบ ได้ผลลัพธ์การแทนค่าที่เป็นบวก ซึ่งเป็นค่าที่สมเหตุสมผลในบริบทของจำนวนนักท่องเที่ยวที่สามารถนำไปใช้ในการสร้างตัวแบบการพยากรณ์ ตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวที่เข้ามาท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด คือตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ ตามลำดับ โดยที่ตัวแบบการพยากรณ์ทั้ง 2 ตัวแบบนี้ เป็นตัวแบบที่ได้จากชุดข้อมูลที่ทดแทนค่าสูญหายจากการกำหนดช่วงการแพร่ระบาดของโควิด-19 ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงธันวาคม พ.ศ. 2564 เป็นข้อมูลสูญหาย ทั้งนี้ ถึงแม้ว่าตัวแบบดังกล่าวจะยังไม่ใช่วิธีการที่เหมาะสมที่สุด แต่สามารถนำมาใช้ในการพยากรณ์ได้ โดยที่ตัวแบบ SARIMA(0,1,0)(0,1,1)₁₂ มีความแม่นยำในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างอยู่ในเกณฑ์ที่สามารถยอมรับได้ ซึ่งมีค่า MAPE เท่ากับ 23.79% ในขณะที่ตัวแบบ SARIMA(0,1,0)(0,1,1)₁₂ มีความแม่นยำในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ดอยู่ในเกณฑ์ ซึ่งมีค่า MAPE เท่ากับ 12.94% นอกจากนี้ ผลการพยากรณ์จำนวนนักท่องเที่ยวล่วงหน้าในเดือนเมษายน ถึงเดือนธันวาคม พ.ศ. 2568 แสดงให้เห็นว่าจำนวนนักท่องเที่ยวที่คาดว่าจะเดินทางมาท่องเที่ยวอุทยานแห่งชาติหมู่เกาะช้างและอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ค่อนข้างคงที่ตามรอบฤดูกาล

สรุปผลการวิจัย : แม้ในบางช่วงเวลาจะไม่มีนักท่องเที่ยวเข้ามาท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ซึ่งทำให้เกิดความไม่ต่อเนื่องของข้อมูล แต่เทคนิคการทดแทนค่าสูญหายช่วยให้สามารถใช้

ข้อมูลในการสร้างตัวแบบการพยากรณ์เบื้องต้นได้อย่างเหมาะสม ช่วยให้หน่วยงานที่เกี่ยวข้องสามารถวางแผนและจัดการการท่องเที่ยวได้อย่างมีประสิทธิภาพมากขึ้น อย่างไรก็ตาม ผลกระทบจากการแพร่ระบาดของโควิด-19 ยังคงมีอยู่ จนกว่าจะมีข้อมูลที่เพียงพอสำหรับการสร้างตัวแบบการพยากรณ์ที่มีความแม่นยำมากขึ้นในอนาคต

คำสำคัญ : การทดแทนค่าสูญหาย ; การแพร่ระบาดของโควิด-19 ; ข้อมูลอนุกรมเวลา ; การแยกส่วนประกอบฤดูกาล ; ตัวแบบการพยากรณ์

Abstract

Background and Objectives : Thailand is internationally recognized as one of the world's major tourist destinations. Tourism also plays an important role in driving the country's economic growth. The tourism industry is not only a primary source of national revenue, but also a key mechanism for promoting employment and overall economic development. However, the outbreak of COVID-19 in late 2019 significantly affected the domestic tourism industry due to the implementation of strict public health measures, such as lockdowns and travel restrictions. As a result, no international tourists have entered Thailand since April 2020. Additionally, important popular marine tourist attractions in the eastern region, including Mu Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park, were affected by the closure of tourism in national parks and forest parks. After the country reopened in late 2021, the tourism industry began to recover, resulting in a noticeable increase in visitors to these two eastern marine national parks. Thus, accurate forecasting of tourist numbers is important for helping related agencies manage tourism sustainably. Moreover, forecasting can provide accurate results when the data used for analysis is continuous and consistent. However, the COVID-19 outbreak resulted in no tourists visiting the national parks during that period, resulting in some data being zero and discontinuous. This may affect the performance of the forecasting model. Therefore, this research aims at forecasting the number of tourists in Mu Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park during the period affected by the COVID-19 pandemic. To improve the continuity and quality of the data, a missing value imputation technique was applied. An appropriate forecasting model was then developed using the widely well-known and commonly applied Box-Jenkins statistical method.

Methodology : This study used data on the number of tourists visiting Mu Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park from October 2015 to March 2025, covering a total of 114 months. The data were obtained from the Department of National Parks, Wildlife and Plant Conservation. Three data scenarios were defined for building the forecasting model: 1. Using actual data with zero values, 2. Treating zero values as missing data, and 3. Treating data during the COVID-19 outbreak from January 2020 to December 2021 as missing data. A seasonal decomposition technique combined with mean imputation was used to fill in the missing values.

Subsequently, the forecasting models were constructed using data from October 2015 to September 2023 using the Box-Jenkins method. The Akaike Information Criterion (AIC) was employed to select the appropriate model among several SARIMA models using the same dataset. The root mean square error (RMSE) was used as a criterion to determine the suitable model for forecasting tourist arrivals in Mu Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park. The accuracy of the forecasts was then validated using data from October 2023 to March 2025, with the mean absolute percentage error (MAPE) used to assess forecasting performance.

Main Results : The estimation of time series data with interrupted tourist numbers in Mu Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park was conducted using a seasonal decomposition approach combined with the mean algorithm under three patterns of missing data conditions. The imputed results yielded positive values, which are reasonable within the context of tourist numbers and can be further utilized in forecasting models. The suitable models for forecasting the number of tourists visiting Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park are SARIMA(2,1,0)(0,1,1)₁₂ and SARIMA(0,1,0)(0,1,1)₁₂, respectively. Both models were developed using data in which the period of the COVID-19 outbreak from January 2020 to December 2021 was treated as missing and replaced using imputation techniques. Although these models may not be the most suitable, they can still be used for forecasting. The SARIMA(2,1,0)(0,1,1)₁₂ model provides acceptable accuracy for forecasting tourist numbers at Mu Ko Chang National Park with a mean absolute percentage error of 23.79%, while the SARIMA(0,1,0)(0,1,1)₁₂ model shows good accuracy for forecasting visitors at Khao Laem Ya – Mu Ko Samet National Park with a mean absolute percentage error of 12.94%. Additionally, the forecast of tourist numbers from April to December 2025 indicates that the number of tourist arrivals to Mu Ko Chang National Park and Khao Laem Ya–Mu Ko Samet National Park remains relatively stable, following seasonal patterns.

Conclusions : Although there were periods with no tourist visits to Mu Ko Chang National Park and Khao Laem Ya – Mu Ko Samet National Park, resulting in data discontinuity, the missing data imputation technique enabled the use of available data to construct an appropriate preliminary forecasting model. This approach helps relevant agencies improve the planning and management of tourism more effectively. However, the impact of the COVID-19 pandemic still remains until sufficient additional data becomes available to develop more accurate forecasting models in the future.

Keywords : missing value imputation, COVID-19 pandemic, time series data, seasonal decomposition, forecasting model

*Corresponding author. E-mail : noparatk@go.buu.ac.th

Introduction

ประเทศไทยได้รับการยอมรับในระดับสากลว่าเป็นหนึ่งในจุดหมายปลายทางด้านการท่องเที่ยวที่สำคัญของโลก โดยผลสำรวจเกี่ยวกับแผนการท่องเที่ยวระดับโลกประจำปี พ.ศ. 2564 ของวีซ่า ได้จัดอันดับให้ประเทศไทยเป็นประเทศที่นักท่องเที่ยวต้องการเดินทางมาเยือนมากที่สุดเป็นลำดับที่สี่ของโลก ทั้งนี้ ในปี พ.ศ. 2562 ประเทศไทยต้อนรับนักท่องเที่ยวชาวต่างชาติกว่า 39.79 ล้านคน และนักท่องเที่ยวชาวไทยเดินทางภายในประเทศรวม 166 ล้านคน-ครั้ง (Ministry of Tourism and Sports, 2021) นอกจากนี้ ภาคการท่องเที่ยวยังมีบทบาทสำคัญในการจ้างงานที่เกี่ยวข้องกับการท่องเที่ยวมากกว่า 7 ล้านคน ซึ่งคิดเป็น 20% ของการจ้างงานทั้งหมดในประเทศ (Mahavanakul & Prasongsak, 2023) ซึ่งสะท้อนให้เห็นถึงความสำคัญของอุตสาหกรรมท่องเที่ยวที่ไม่เพียงเป็นแหล่งรายได้หลักของประเทศเท่านั้น แต่ยังเป็นกลไกสำคัญในการขับเคลื่อนการจ้างงานและเศรษฐกิจโดยรวมอีกด้วย แต่การแพร่ระบาดของโรคอุบัติใหม่อย่างโควิด-19 ตั้งแต่ช่วงปลายปี พ.ศ. 2562 ส่งผลให้หลายประเทศใช้มาตรการการควบคุมที่เข้มงวด เช่น การล็อกดาวน์ และการเว้นระยะห่างทางกายภาพ ซึ่งส่งผลกระทบต่ออุตสาหกรรมการบินและการท่องเที่ยวอย่างรุนแรง รวมถึงประเทศไทยที่ได้รับผลกระทบจากสถานการณ์ดังกล่าวโดยตรง ส่งผลให้ไม่มีนักท่องเที่ยวต่างชาติเข้ามาในประเทศไทยตั้งแต่เดือนเมษายน พ.ศ. 2563 เป็นต้นมา (Ministry of Tourism and Sports, 2021) เนื่องจากประเทศไทยมีการพึ่งพารายได้จากภาคการท่องเที่ยวในสัดส่วนที่สูง การแพร่ระบาดของโควิด-19 จึงส่งผลกระทบต่ออย่างรุนแรงต่อเศรษฐกิจ โดยเฉพาะจากการลดลงของนักท่องเที่ยวต่างชาติที่เป็นแหล่งรายได้หลักของอุตสาหกรรมท่องเที่ยวไทย อย่างไรก็ตาม ในปี พ.ศ. 2564 สถานการณ์การแพร่ระบาดเริ่มดีขึ้นจากความคืบหน้าของการฉีดวัคซีนป้องกันโรคโควิด-19 ส่งผลให้มาตรการการควบคุมโรคระบาดไม่เข้มงวดเหมือนในปี พ.ศ. 2563 แต่มีการกำหนดระดับของพื้นที่ควบคุมให้เหมาะสมกับสถานการณ์ และเริ่มเปิดประเทศในช่วงปลายปี พ.ศ. 2564 ทำให้บรรยากาศการเดินทางท่องเที่ยวกลับมาคึกคักอีกครั้ง ทั้งนี้ ในปี พ.ศ. 2565 การฟื้นตัวของภาคการท่องเที่ยวเริ่มเห็นได้ชัดเจนขึ้น กรมการท่องเที่ยวจึงได้ดำเนินมาตรการฟื้นฟูภาคการท่องเที่ยวไทยให้กลับมาแข็งแกร่งอีกครั้ง ทำให้ในช่วงไตรมาสแรกของปี พ.ศ. 2566 จำนวนนักท่องเที่ยวกลับมาใกล้เคียงกับช่วงก่อนการแพร่ระบาดของโควิด-19 อย่างไรก็ตาม การฟื้นตัวของการท่องเที่ยวนั้นมีความแตกต่างกันในแต่ละพื้นที่ โดยภาคตะวันออกเป็นหนึ่งในภูมิภาคที่มีการพึ่งพานักท่องเที่ยวต่างชาติ ทำให้ยังคงฟื้นตัวได้ช้ากว่าภูมิภาคอื่น ๆ ที่พึ่งพานักท่องเที่ยวชาวไทยเป็นหลัก (Mahavanakul & Prasongsak, 2023)

การท่องเที่ยวถือเป็นปัจจัยสำคัญในการขับเคลื่อนเศรษฐกิจ โดยเฉพาะแหล่งท่องเที่ยวทางธรรมชาติอย่างอุทยานแห่งชาติ ที่ดึงดูดนักท่องเที่ยวทั้งชาวไทยและชาวต่างชาติ อุทยานแห่งชาติหมู่เกาะช้างและอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด เป็นแหล่งท่องเที่ยวทางทะเลที่สำคัญของภาคตะวันออกที่ได้รับความนิยมอย่างต่อเนื่อง อย่างไรก็ตาม การแพร่ระบาดของโควิด-19 ส่งผลกระทบต่ออย่างมากต่อการท่องเที่ยวในพื้นที่ดังกล่าว โดยในปี พ.ศ. 2563 การแพร่ระบาดระลอกแรกทำให้กรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช ประกาศปิดการท่องเที่ยวในเขตอุทยานแห่งชาติและวนอุทยานทุกแห่งทั่วประเทศ ตั้งแต่วันที่ 25 มีนาคม ส่งผลให้ตั้งแต่เดือนเมษายนถึงมิถุนายนในปีดังกล่าวไม่มีนักท่องเที่ยวเข้าไปในพื้นที่ (ดังแสดงใน Figure 1 และ 2) ก่อนจะกลับมาเปิดให้บริการนักท่องเที่ยวอีกครั้งในเดือนกรกฎาคม ต่อมาในช่วงกลางปี พ.ศ. 2564 การแพร่ระบาดของโควิด-19 สายพันธุ์เดลต้า ทำให้มีการปิดอุทยานแห่งชาติหมู่เกาะช้างอีกครั้งตั้งแต่เดือนมิถุนายนถึงเดือน

สิงหาคม ส่งผลให้จำนวนนักท่องเที่ยวลดลงเหลือศูนย์ (Figure 1) ในขณะที่อุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด มีการปิดเฉพาะบางแหล่งท่องเที่ยวในช่วงเวลาเดียวกัน ส่งผลให้จำนวนนักท่องเที่ยวลดลงอย่างเห็นได้ชัด (Figure 2)

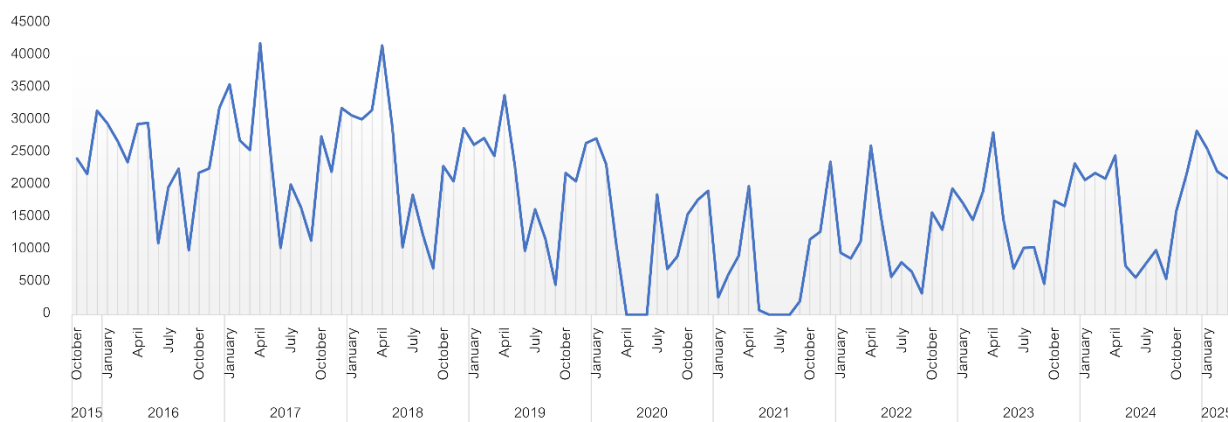


Figure 1 Number of Tourists in Mu Ko Chang National Park from October 2015 to March 2025

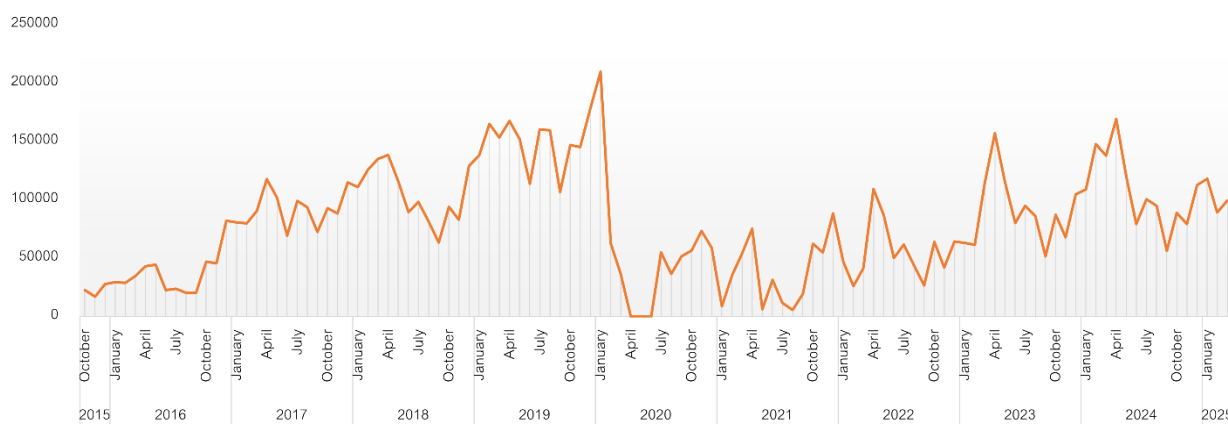


Figure 2 Number of Tourists in Khao Laem Ya – Mu Ko Samet National Park from October 2015 to March 2025

ภายหลังการเปิดประเทศในช่วงปลายปี พ.ศ. 2564 อุตสาหกรรมการท่องเที่ยวมีการฟื้นตัวอย่างต่อเนื่อง ทำให้จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลทั้ง 2 แห่ง เพิ่มขึ้นอย่างชัดเจน (Figure 1 และ 2) โดยอุทยานแห่งชาติหมู่เกาะช้างมีจำนวนนักท่องเที่ยวอย่างสม่ำเสมอตามรอบฤดูกาล ในขณะที่อุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด มีอัตราการเติบโตของนักท่องเที่ยวสูงขึ้นอย่างต่อเนื่องและขยายตัวเพิ่มขึ้นในแต่ละรอบฤดูกาล เนื่องจากพื้นที่อุทยานแห่งชาติทางทะเลทั้ง 2 แห่ง เป็นพื้นที่ที่มีความหลากหลายทางชีวภาพ อีกทั้งยังเป็นแหล่งท่องเที่ยวสำคัญที่ส่งเสริมเศรษฐกิจของประเทศ ด้วยเหตุนี้ การพยากรณ์จำนวนนักท่องเที่ยวที่มีความแม่นยำจึงเป็นสิ่งสำคัญในการช่วยหน่วยงานที่เกี่ยวข้อง

ดำเนินการในด้านต่าง ๆ เช่น การกำหนดนโยบายการท่องเที่ยวที่เหมาะสม ที่มีการคำนึงถึงขีดความสามารถในการรองรับของพื้นที่ การป้องกันการใช้ทรัพยากรจนเกินขีดจำกัด รวมถึงการสร้างสมดุลระหว่างเศรษฐกิจกับการอนุรักษ์ธรรมชาติ เพื่อให้การพัฒนาทางเศรษฐกิจจากการท่องเที่ยวเป็นไปอย่างยั่งยืน ดังนั้น การบริหารจัดการการท่องเที่ยวโดยใช้ข้อมูลพยากรณ์ที่ถูกต้องและแม่นยำ จะช่วยให้การดำเนินการในอุทยานแห่งชาติทางทะเลทั้ง 2 แห่ง เป็นไปอย่างยั่งยืน ไม่ทำลายทรัพยากรธรรมชาติและระบบนิเวศ ซึ่งสอดคล้องกับเป้าหมายการพัฒนาที่ยั่งยืน (Sustainable Development Goals; SDGs) เป้าหมายที่ 14 ที่มุ่งเน้นการอนุรักษ์และใช้ประโยชน์จากมหาสมุทร ทะเล และทรัพยากรทางทะเลอย่างยั่งยืน เพื่อการพัฒนาที่ยั่งยืน

ในประเทศไทย มีงานวิจัยหลายงานที่ได้ศึกษาเกี่ยวกับการพยากรณ์จำนวนนักท่องเที่ยวทั้งในระดับประเทศ เช่น การพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมายังประเทศไทย (Arttaweekul, 2013; Keerativibool, 2013; Saothayanun *et al.*, 2014; Lake, 2016; Mukma *et al.*, 2018; Anekunnawut, 2020; Intarapak *et al.*, 2022) และในระดับภูมิภาค เช่น การพยากรณ์จำนวนนักท่องเที่ยวที่มาท่องเที่ยวในภาคเหนือ (Kummong & Homdok, 2019) หรือแหล่งท่องเที่ยวเฉพาะอย่าง อุทยานแห่งชาติภูหินร่องกล้า (Boonlha, 2019) โดยวิธีการที่ใช้ในการพยากรณ์นั้นมีความหลากหลาย ไม่ว่าจะเป็นวิธีเชิงสถิติ ที่นิยมใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา เช่น วิธีแยกส่วนประกอบ วิธีการทำให้เรียบแบบเลขชี้กำลังของโฮลท์-วินเทอร์ วิธีบ็อกซ์-เจนกินส์ และวิธีการพยากรณ์รวม เป็นต้น (Arttaweekul, 2013; Keerativibool, 2013; Saothayanun *et al.*, 2014; Mukma *et al.*, 2018; Intarapak *et al.*, 2022) นอกจากนี้ ยังมีการใช้เทคนิคที่ทันสมัยขึ้นเพื่อเพิ่มความแม่นยำในการพยากรณ์ เช่น วิธีการเรียนรู้ของเครื่อง (Lake, 2016) และวิธีการเรียนรู้เชิงลึก (Kummong & Homdok, 2019; Anekunnawut, 2020) เป็นต้น ซึ่งงานวิจัยเหล่านี้สะท้อนให้เห็นถึงแนวทางที่หลากหลายในการพยากรณ์จำนวนนักท่องเที่ยว โดยสามารถนำไปประยุกต์ใช้ให้เหมาะสมกับแต่ละพื้นที่ โดยทั่วไป การพยากรณ์สามารถให้ผลลัพธ์ที่แม่นยำเมื่อข้อมูลที่นำมาใช้ในการวิเคราะห์มีความต่อเนื่องและสม่ำเสมอ อย่างไรก็ตาม เหตุการณ์ผิดปกติอย่างการแพร่ระบาดของโควิด-19 นำไปสู่มาตรการการล็อกดาวน์และข้อจำกัดในการเดินทาง ส่งผลให้ไม่มีนักท่องเที่ยวเดินทางมายังอุทยานแห่งชาติในช่วงเวลาดังกล่าว ดังนั้น จึงทำให้ข้อมูลเกิดความไม่ต่อเนื่อง และอาจส่งผลกระทบต่อประสิทธิภาพของตัวแบบการพยากรณ์ ทั้งนี้ ในการศึกษาที่ผ่านมาของ Hyndman & Rostami-Tabar (2024) ได้มีการเสนอวิธีการจัดการกับข้อมูลอนุกรมเวลาที่หยุดชะงักจากสถานการณ์การแพร่ระบาดของโควิด-19 ไว้หลายแนวทาง เช่น การใช้ตัวแบบที่มีความยืดหยุ่นสูง (Highly Adaptive Model) เพื่อปรับค่าพยากรณ์ตามการเปลี่ยนแปลงของลักษณะเฉพาะของอนุกรมเวลาในแต่ละช่วงเวลา การใช้ตัวแบบการแทรกแซง (Intervention Model) ที่มีตัวแปรร่วมที่อธิบายการหยุดชะงักอย่างชัดเจน รวมถึงการกำหนดค่าข้อมูลในช่วงที่หยุดชะงักหรือไม่ต่อเนื่องให้เป็นค่าสูญหายแล้วทำการประมาณค่าใหม่ นอกจากนี้ ยังมีวิธีการปรับลดน้ำหนักของข้อมูลในช่วงเวลาที่เกิดการหยุดชะงัก และการใช้ตัวแบบผสมผสาน (Ensemble Model) เพื่อรวมผลลัพธ์จากหลายวิธี นอกจากนี้ ในการศึกษาของ Mueller & Sobreira (2024) ได้ศึกษาผลกระทบของโควิด-19 ต่อความสามารถของตัวแบบอนุกรมเวลาในการพยากรณ์ความต้องการการท่องเที่ยวในประเทศโปรตุเกส รวมถึงพิจารณาแนวทางในการลดอิทธิพลของค่าสังเกตแบบสุดขีด (Extreme Observation) ที่เกิดขึ้นในช่วงการแพร่ระบาดของโควิด-19 โดยมีแนวทางในการจัดการข้อมูล 2 วิธี คือ การตัดข้อมูลที่ผิดปกติออกทั้งหมด โดยแบ่ง

การศึกษาออกเป็น 2 กรณี คือ การตัดข้อมูลในปี พ.ศ. 2563 และการตัดข้อมูลในช่วงปี พ.ศ. 2563 ถึง 2564 ในขณะที่อีกแนวทางหนึ่งที่ใช้ในการจัดการข้อมูล คือ การแทนค่าข้อมูลในช่วงที่ได้รับผลกระทบด้วยค่าเฉลี่ยจากช่วงเวลาก่อนและหลังการแพร่ระบาดของโควิด-19 โดยแบ่งการแทนค่าข้อมูลออกเป็น 2 กรณี เช่นเดียวกัน นั่นคือ การแทนค่าข้อมูลในปี พ.ศ. 2563 และการแทนค่าข้อมูลในช่วงปี พ.ศ. 2563 ถึง 2564 ด้วยวิธีการประมาณค่าในช่วงเชิงเส้น (Linear Interpolation) ซึ่งเป็นการหาค่าเฉลี่ยข้อมูลแต่ละเดือนของปี พ.ศ. 2562 และ 2564 และของปี พ.ศ. 2562 และ 2565 ตามลำดับ

จากแนวทางข้างต้น การศึกษานี้จึงมุ่งเน้นการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลทั้ง 2 แห่งที่ได้รับผลกระทบจากสถานการณ์การแพร่ระบาดของโควิด-19 ที่ทำให้ข้อมูลบางช่วงเวลาเป็นศูนย์ เพื่อจัดการกับปัญหาดังกล่าว จึงได้มีการกำหนดค่าศูนย์ให้เป็นข้อมูลสูญหาย และใช้เทคนิคการประมาณค่าสูญหายมาช่วยในการปรับปรุงคุณภาพข้อมูลให้มีความต่อเนื่อง จากนั้นจึงเลือกใช้วิธีเชิงสถิติที่เป็นที่รู้จักและใช้กันอย่างแพร่หลายอย่างวิธีบ็อกซ์-เจนกินส์ มาสร้างตัวแบบการพยากรณ์ เนื่องจากเป็นวิธีที่ไม่จำเป็นต้องกำหนดรูปแบบที่ตายตัวก่อนทำการวิเคราะห์ และสามารถปรับรูปแบบเพื่อให้ได้ตัวแบบที่เหมาะสมสำหรับการพยากรณ์ได้

Methodology

ในหัวข้อนี้จะกล่าวถึงข้อมูลที่ใช้ในการศึกษา การจัดการข้อมูลจนกระทั่งเกิดความไม่ต่อเนื่อง การพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ และการประเมินตัวแบบ ดังมีรายละเอียดต่อไปนี้

1. ข้อมูลที่ใช้ในการศึกษา

ข้อมูลที่ใช้ในการศึกษาค้างนี้เป็นข้อมูลจำนวนนักท่องเที่ยวที่เดินทางมาท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างและอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ซึ่งอยู่ภายใต้การกำกับของสำนักบริหารพื้นที่อนุรักษ์ที่ 2 (ศรีราชา) ตั้งแต่เดือนตุลาคม พ.ศ. 2558 ถึงเดือนมีนาคม พ.ศ. 2568 รวมทั้งหมด 114 เดือน ทั้งนี้ ข้อมูลดังกล่าวได้มาจากเว็บไซต์ของกรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช (Department of National Parks, Wildlife and Plant Conservation, 2021) โดยในการศึกษาครั้งนี้ได้ทำการแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลที่ใช้สำหรับสร้างตัวแบบการพยากรณ์ซึ่งประกอบไปด้วยข้อมูลตั้งแต่เดือนตุลาคม พ.ศ. 2558 ถึงเดือนกันยายน พ.ศ. 2566 รวมทั้งสิ้น 96 เดือน และข้อมูลที่ใช้สำหรับตรวจสอบความแม่นยำของการพยากรณ์ ซึ่งประกอบไปด้วยข้อมูลตั้งแต่เดือนตุลาคม พ.ศ. 2566 ถึงเดือนมีนาคม พ.ศ. 2568 รวมทั้งสิ้น 18 เดือน

2. การจัดการข้อมูลจนกระทั่งเกิดความไม่ต่อเนื่อง

การแพร่ระบาดของโควิด-19 ได้สร้างความท้าทายอย่างมากต่อการพยากรณ์ข้อมูลจนกระทั่งเกิดความไม่ต่อเนื่องทางสถิติที่เคยมีความเสถียรจากข้อมูลในอดีต กลับต้องเผชิญกับการเปลี่ยนแปลงที่เกิดขึ้นอย่างกะทันหัน อันเป็นผลมาจากมาตรการควบคุมต่าง ๆ ซึ่งส่งผลให้ข้อมูลที่นำมาใช้ในการพยากรณ์ไม่สามารถสะท้อนถึงแนวโน้มและความต่อเนื่องในอดีตได้ ดังนั้น การจัดการกับข้อมูลจนกระทั่งเกิดความไม่ต่อเนื่องจึงเป็นเรื่องที่สำคัญอย่างยิ่ง เนื่องจากข้อมูลที่เกิดการหยุดชะงักหรือได้รับผลกระทบจากเหตุการณ์ที่ผิดปกติ สามารถส่งผลกระทบต่อความแม่นยำของการพยากรณ์ เพื่อให้การพยากรณ์มีประสิทธิภาพและแม่นยำยิ่งขึ้น การเลือกใช้แนวทางที่เหมาะสมเพื่อลดผลกระทบจากข้อมูลที่ผิดปกติจึงเป็นสิ่งที่

จำเป็น ในการศึกษาของ Hyndman & Rostami-Tabar (2024) ได้มีการเสนอวิธีการรับมือกับข้อมูลอนุกรมเวลาที่ถูกรบกวน (Interrupted Time Series) หลากหลายแนวทาง อย่างไรก็ตาม ในการศึกษาครั้งนี้เลือกใช้วิธีการกำหนดให้ค่าของข้อมูลในช่วงที่ถูกรบกวนเป็นค่าสูญหาย พร้อมทั้งทำการประมาณค่าสูญหายดังกล่าว

ข้อมูลสูญหาย (Missing Data) หรือค่าสูญหาย (Missing Value) หมายถึง ค่าข้อมูลที่ไม่ได้ถูกบันทึกในชุดข้อมูลที่ศึกษา ซึ่งปัญหานี้พบได้บ่อยในงานวิจัยเกือบทุกประเภท และอาจส่งผลกระทบต่อความถูกต้องของผลการวิเคราะห์และข้อสรุปที่ได้จากข้อมูล (Kang, 2013) เพื่อแก้ไขปัญหาดังกล่าว หนึ่งในแนวทางที่นิยมใช้คือการทดแทนค่าสูญหาย (Missing Value Imputation) ในกรณีที่ข้อมูลมีลักษณะเป็นอนุกรมเวลา และแสดงรูปแบบที่มีการเปลี่ยนแปลงตามฤดูกาลอย่างชัดเจน วิธีการทดแทนค่าสูญหายโดยการแยกส่วนประกอบฤดูกาล (Seasonally Decomposed Missing Value Imputation Method) จึงเป็นวิธีการหนึ่งที่เหมาะสมในการทดแทนข้อมูลที่สูญหาย โดยไม่ทำให้รูปแบบของข้อมูลผิดไปจากความแปรผันตามฤดูกาล ทั้งนี้ สามารถใช้ชุดคำสั่ง imputeTS ในโปรแกรม R มาช่วยในการจัดการกับข้อมูลอนุกรมเวลาที่มีค่าสูญหายดังกล่าวได้ (Moritz & Bartz-Beielstein, 2017; Moritz *et al.*, 2022) โดยเริ่มต้นจากการแยกส่วนประกอบฤดูกาลของอนุกรมเวลาด้วยวิธี LOESS (Seasonal Decomposition of Time Series by Loess) (Cleveland *et al.*, 1990) จากนั้นลบส่วนประกอบฤดูกาลออกจากข้อมูลดั้งเดิม เพื่อให้ได้ชุดข้อมูลที่ไม่มีผลกระทบจากฤดูกาล และทำการทดแทนค่าสูญหายในข้อมูลที่ไม่มีส่วนประกอบฤดูกาลด้วยขั้นตอนวิธี (Algorithm) การทดแทนค่าสูญหาย ซึ่งมีหลายขั้นตอนวิธีให้เลือกใช้ ได้แก่ การประมาณค่าในช่วง (Interpolation) การใช้ตัวแบบการปรับให้เรียบของ Kalman และ State Space (Kalman Smoothing and State Space Models) การใช้ค่าสังเกตครั้งสุดท้าย (Last Observation Carried Forward) การใช้วิธีค่าเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก (Weighted Moving Average) และการใช้ค่าเฉลี่ย (Mean Value) เป็นต้น จากนั้น ทำการเพิ่มส่วนประกอบฤดูกาลกลับเข้าไปในข้อมูล เพื่อให้ได้ข้อมูลอนุกรมเวลาที่สมบูรณ์และสามารถนำไปใช้ในการพยากรณ์ต่อไป

การจัดการกับข้อมูลอนุกรมเวลาที่เกิดความไม่ต่อเนื่องนั้น ถือเป็นขั้นตอนสำคัญก่อนการสร้างตัวแบบการพยากรณ์ โดยเฉพาะเมื่อข้อมูลถูกรบกวนจากสถานการณ์ต่าง ๆ อย่างสถานการณ์การแพร่ระบาดของโควิด-19 และในการศึกษาครั้งนี้มีขั้นตอนในการดำเนินการ ดังนี้

1. ตรวจสอบข้อมูลเบื้องต้น โดยการค้นหาค่าที่เป็นศูนย์ในชุดข้อมูลจำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลทั้ง 2 แห่ง ที่ใช้สำหรับสร้างตัวแบบการพยากรณ์
2. กำหนดรูปแบบการสูญหายของข้อมูลที่ใช้สำหรับสร้างตัวแบบการพยากรณ์ ภายใต้การสูญหายของข้อมูลรูปแบบต่าง ๆ 3 รูปแบบ คือ
 - รูปแบบที่ 1 ใช้ข้อมูลจริงที่มีค่าศูนย์: ในกรณีนี้ ไม่กำหนดให้ค่าศูนย์เป็นข้อมูลสูญหาย หมายความว่าข้อมูลสมบูรณ์ 100% (หรือไม่มีค่าสูญหาย)
 - รูปแบบที่ 2 กำหนดค่าศูนย์เป็นข้อมูลสูญหาย : ในกรณีนี้ ข้อมูลจำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างมีค่าสูญหาย 6.250% ในขณะที่ข้อมูลจำนวนนักท่องเที่ยวในอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ดมีค่าสูญหาย 3.125%

- รูปแบบที่ 3 กำหนดข้อมูลช่วงการแพร่ระบาดของโควิด-19 เป็นข้อมูลสูญหาย : ให้ข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนธันวาคม พ.ศ. 2564 เป็นข้อมูลสูญหาย (รวมทั้งสิ้น 24 เดือน) โดยคิดเป็นค่าสูญหาย 25%
3. ประมาณค่าสูญหายด้วยการแยกส่วนประกอบฤดูกาลร่วมกับการประมาณค่าด้วยขั้นตอนวิธีค่าเฉลี่ย ภายใต้การสูญหายของข้อมูล 3 รูปแบบ โดยมีขั้นตอน ดังนี้
- แยกส่วนประกอบของอนุกรมเวลาด้วยวิธี STL (Seasonal-Trend Decomposition Procedure Based on Loess) ซึ่งเป็นการแยกส่วนประกอบแนวโน้มและฤดูกาลโดยใช้เทคนิค LOESS (Locally Estimated Scatterplot Smoothing) ในรูปแบบของการแยกส่วนประกอบแบบบวก (Additive Decomposition) โดยข้อมูลจะถูกแยกออกเป็น 3 ส่วนประกอบหลัก ได้แก่ ส่วนประกอบแนวโน้ม (Trend Component) ส่วนประกอบฤดูกาล (Seasonal Component) และส่วนประกอบที่เหลือ (Remainder Component) (Cleveland *et al.*, 1990)
 - ลบส่วนประกอบฤดูกาล (Deseasonalization) ออกจากชุดข้อมูลเดิม เพื่อขจัดอิทธิพลของฤดูกาลต่อการประมาณค่าสูญหาย ทำให้ค่าที่ประมาณได้ไม่ถูกรบกวนและสะท้อนการเปลี่ยนแปลงของข้อมูลโดยปราศจากอิทธิพลของฤดูกาล
 - ทดแทนค่าที่สูญหายในชุดข้อมูลที่ลบฤดูกาลออกแล้ว โดยใช้การประมาณค่าด้วยขั้นตอนวิธีค่าเฉลี่ย ซึ่งขั้นตอนวิธีนี้เหมาะสมต่อการประมาณค่าเนื่องจากให้ผลลัพธ์ที่เป็นบวก จึงช่วยป้องกันค่าประมาณที่ไม่สมเหตุสมผล เช่น ค่าติดลบ ซึ่งไม่สามารถเกิดขึ้นได้ในบริบทของจำนวนนักท่องเที่ยว
 - นำส่วนประกอบฤดูกาลกลับมารวมกับชุดข้อมูลที่ทำการทดแทนค่าที่สูญหายแล้ว เพื่อให้ได้ชุดข้อมูลที่สมบูรณ์และคงลักษณะเชิงฤดูกาลตามเดิม

3. การพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์

วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins Method) เป็นเทคนิคการพยากรณ์เชิงปริมาณที่อาศัยพฤติกรรมของข้อมูลในอดีตเพื่อคาดการณ์อนาคต โดยไม่จำเป็นต้องกำหนดรูปแบบที่ตายตัวของตัวแบบล่วงหน้า ซึ่งต่างจากวิธีการพยากรณ์อื่น ๆ ที่ต้องระบุโครงสร้างของตัวแบบก่อนวิเคราะห์ ทำให้วิธีนี้สามารถใช้กับข้อมูลที่ไม่มีแนวโน้มหรือไม่มีฤดูกาลที่ชัดเจนได้ กระบวนการพยากรณ์ด้วยวิธีนี้จะอาศัยแนวคิดที่ว่า ข้อมูลอนุกรมเวลาควรมีคุณสมบัติหนึ่ง (Stationary) หากข้อมูลมีคุณสมบัติหนึ่งอยู่แล้วจะใช้ตัวแบบการเฉลี่ยเคลื่อนที่ถดถอยในตัว (Autoregressive Moving Average; ARMA) แต่หากข้อมูลยังไม่มีคุณสมบัติหนึ่ง (Non-Stationary) จะต้องปรับข้อมูลให้มีคุณสมบัติหนึ่งก่อน โดยการหาผลต่างอันดับที่ 1 เพื่อลดแนวโน้มของข้อมูล ทำให้สามารถนำไปใช้ในการพยากรณ์ได้อย่างมีประสิทธิภาพ ซึ่งตัวแบบหลังจากทำการปรับข้อมูลแล้วจะเรียกว่า ตัวแบบรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ (Autoregressive Integrated Moving Average; ARIMA)

สำหรับข้อมูลอนุกรมเวลาที่มีลักษณะตามฤดูกาล ตัวแบบรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ตามฤดูกาล (Seasonal ARIMA; SARIMA) ได้ถูกออกแบบมาเพื่อรองรับรูปแบบดังกล่าว โดยใช้การหาผลต่างตามฤดูกาล (Seasonal

Differencing) เพื่อจับความผันผวนที่เกิดขึ้นเป็นรอบ ๆ ทั้งนี้ ตัวแบบ SARIMA จะถูกกำหนดโดยอันดับ p, d, q, P, D, Q และ s แทน ความยาวฤดูกาล หรือสามารถเขียนให้อยู่ในรูปของตัวแบบ SARIMA(p, d, q)(P, D, Q)_s และแสดงให้อยู่ในรูปของสมการได้ ดังนี้

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \delta + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (1)$$

เมื่อ Y_t แทน ค่าที่สังเกตได้ของข้อมูล ณ เวลา t และ ε_t แทน ค่าคลาดเคลื่อน ณ เวลา t ในขณะที่ B เป็นตัวดำเนินการย้อนหลัง (Backshift Operator) และ δ เป็นค่าคงที่ นอกจากนี้ ϕ_p, Φ_P, θ_q , และ Θ_Q เป็นพหามิตอรีในส่วนของตัวแบบ AR, SAR, MA, และ SMA ตามลำดับ ซึ่งเป็นองค์ประกอบของตัวแบบ SARIMA

จากลักษณะของข้อมูลจำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่ง ที่มีการเปลี่ยนแปลงต่อเนื่องและเกิดรูปแบบซ้ำตามฤดูกาลในแต่ละปี (Figure 1 และ 2) ตัวแบบ SARIMA จึงมีความเหมาะสมสำหรับการพยากรณ์ เนื่องจากสามารถจัดการรูปแบบตามฤดูกาลได้โดยตรงและไม่ต้องกำหนดรูปแบบล่วงหน้า ทำให้สามารถปรับตัวแบบให้สอดคล้องกับลักษณะข้อมูลจริง ส่งผลให้การพยากรณ์มีความแม่นยำและสะท้อนลักษณะการเปลี่ยนแปลงของจำนวนนักท่องเที่ยวได้อย่างเหมาะสม ทั้งนี้ การพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีบ็อก-เจนกินส์ มีขั้นตอนในการสร้างตัวแบบ (Lorchirachoonkul & Jitthavech, 2005; Manmin, 2006) ดังนี้

1. ตรวจสอบคุณสมบัติเบื้องต้นของอนุกรมเวลา โดยการพิจารณาผ่านกราฟของอนุกรมเวลาที่พล็อตเทียบกับเวลารวมถึงการพิจารณาจากกราฟของฟังก์ชันสหสัมพันธ์ในตัว (Autocorrelation Function; ACF) และฟังก์ชันสหสัมพันธ์ในตัวบางส่วน (Partial Autocorrelation Function; PACF) หากพบว่าอนุกรมเวลายังไม่มีคุณสมบัติเบื้องต้น จำเป็นต้องแปลงให้อยู่ในรูปแบบที่นิ่งก่อน
2. กำหนดตัวแบบเบื้องต้นโดยการพิจารณาจากรูปแบบของกราฟ ACF และ PACF หรือใช้ซอฟต์แวร์สำหรับการวิเคราะห์ทางสถิติในการกำหนดตัวแบบโดยไม่จำเป็นต้องตรวจสอบคุณสมบัติเบื้องต้นในข้อ 1
3. ประเมินค่าสัมประสิทธิ์ของตัวแบบที่เลือกโดยใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method) หากพบว่าสัมประสิทธิ์ใดไม่มีนัยสำคัญทางสถิติ ควรทำการปรับปรุงตัวแบบใหม่จนกว่าสัมประสิทธิ์ทุกตัวจะมีนัยสำคัญทางสถิติ รวมถึงมีค่าของเกณฑ์สารสนเทศของอะกะอิเกะที่ต่ำที่สุด
4. ตรวจสอบความเหมาะสมของตัวแบบ โดยการตรวจสอบข้อสมมุติเบื้องต้นของค่าคลาดเคลื่อน นั่นคือ ค่าคลาดเคลื่อนควรมีการแจกแจงปรกติ มีค่าเฉลี่ยเป็นศูนย์ ไม่มีความสัมพันธ์กัน และมีความแปรปรวนคงที่ตลอดเวลา หากตัวแบบไม่เป็นไปตามเงื่อนไขเหล่านี้ จำเป็นต้องกำหนดตัวแบบในข้อ 2 ใหม่ จนกว่าจะได้ตัวแบบที่เหมาะสม

4. การประเมินตัวแบบ

ในงานวิจัยนี้ จะประเมินประสิทธิภาพของตัวแบบด้วยตัวชี้วัด 3 ตัว ได้แก่ เกณฑ์สารสนเทศของอะกะอิเกะ รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ซึ่งมีรายละเอียด ดังนี้

4.1 เกณฑ์สารสนเทศของอะกะอิเกะ (Akaike's information criterion; AIC)

เกณฑ์สารสนเทศของอะกะอิเกะ เป็นเกณฑ์คัดเลือกตัวแบบที่เหมาะสมที่สุดภายใต้ตัวแบบที่มีการใช้ข้อมูลชุดเดียวกัน แต่มีจำนวนพารามิเตอร์ที่แตกต่างกัน โดยมีค่า AIC ต่ำที่สุด จากการประมาณค่าพารามิเตอร์ของตัวแบบด้วยวิธีภาวะน่าจะเป็นสูงสุดสามารถคำนวณหาค่า AIC ได้จาก

$$AIC = -2 \ln(L) + 2p \quad (2)$$

โดยที่ L แทน ฟังก์ชันภาวะน่าจะเป็นสูงสุด ในขณะที่ p แทน จำนวนพารามิเตอร์ในตัวแบบ

4.2 รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error; RMSE)

รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย เป็นค่าวัดความแตกต่างระหว่างค่าจริงและค่าพยากรณ์ โดยค่า RMSE ที่ต่ำหมายถึงค่าพยากรณ์มีความแม่นยำและใกล้เคียงกับค่าจริง สามารถคำนวณได้ตามสมการต่อไปนี้

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (3)$$

เมื่อ n แทน จำนวนข้อมูล ในขณะที่ Y_t และ $\hat{Y}_t$ แทน ข้อมูลจริงและค่าพยากรณ์ ณ เวลา t ตามลำดับ

4.3 ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error; MAPE)

ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย เป็นค่าวัดที่ใช้ประเมินความถูกต้องของการพยากรณ์ โดยตัวแบบที่ให้ค่า MAPE ต่ำกว่าถือว่ามีความแม่นยำสูงกว่า ซึ่งสามารถคำนวณได้ตามสมการต่อไปนี้

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100 \quad (4)$$

เมื่อ n แทน จำนวนข้อมูล ในขณะที่ Y_t และ $\hat{Y}_t$ แทน ข้อมูลจริงและค่าพยากรณ์ ณ เวลา t ตามลำดับ นอกจากนี้ ค่า MAPE ยังสามารถแบ่งระดับความแม่นยำในการพยากรณ์ (Lewis, 1982) ได้ดัง Table 1

Table 1 Typical interpretation of MAPE values

MAPE Criteria	Interpretation
MAPE $\leq$ 10%	Highly Accurate
10% < MAPE $\leq$ 20%	Good Accuracy
20% < MAPE $\leq$ 50%	Acceptable Accuracy
MAPE > 50%	Inaccurate

Results

ในหัวข้อนี้จะนำเสนอผลลัพธ์จากการวิเคราะห์และพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด โดยใช้วิธีบ็อกซ์-เจนกินส์ โดยแบ่งการนำเสนอออกเป็น 3 ส่วน คือ การประมาณค่าข้อมูลอนุกรมเวลาที่เกิดความไม่ต่อเนื่อง การสร้างตัวแบบการพยากรณ์ และการประเมินความแม่นยำของการพยากรณ์

1. การประมาณค่าข้อมูลอนุกรมเวลาที่เกิดความไม่ต่อเนื่อง

การประมาณค่าข้อมูลอนุกรมเวลาที่มีความไม่ต่อเนื่องของจำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ได้ทำการแทนค่าที่เป็นศูนย์โดยใช้การทดแทนค่าสูญหายด้วยวิธีการแยกส่วนประกอบฤดูกาลร่วมกับขั้นตอนวิธีค่าเฉลี่ย ภายใต้เงื่อนไขการสูญหายของข้อมูล 3 รูปแบบ และผลการทดแทนค่าสูญหายสำหรับข้อมูลที่ใช้ในการสร้างตัวแบบการพยากรณ์แสดงไว้ใน Table 2

2. การสร้างตัวแบบการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออก

ในการกำหนดตัวแบบการพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ ขั้นตอนแรกคือการตรวจสอบคุณสมบัติเบื้องต้นของข้อมูลก่อนที่จะกำหนดตัวแบบ หากข้อมูลไม่มีคุณสมบัติเบื้องต้น จะต้องทำให้ข้อมูลมีคุณสมบัติเบื้องต้น จากนั้นจึงจะสามารถกำหนดตัวแบบที่เหมาะสมได้ เพื่อให้ง่ายต่อการดำเนินการ การศึกษาได้ใช้ฟังก์ชัน `auto.arima()` ในโปรแกรม R ช่วยในการกำหนดตัวแบบซึ่งฟังก์ชันนี้ไม่จำเป็นต้องตรวจสอบคุณสมบัติเบื้องต้นของข้อมูลก่อนการวิเคราะห์ โดยพิจารณาจากค่า AIC ร่วมกับผลการทดสอบสมมติฐานค่าพารามิเตอร์ แล้วเลือกตัวแบบที่พารามิเตอร์ทุกตัวมีนัยสำคัญทางสถิติพร้อมทั้งมีค่า AIC ต่ำที่สุด ทั้งนี้ ผลลัพธ์จากการเลือกตัวแบบของแต่ละชุดข้อมูล สำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง สรุปไว้ในตารางที่ 3 และสำหรับอุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด สรุปไว้ใน Table 4

จาก Table 3 และ 4 จะเห็นว่าตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด เป็นตัวแบบที่ได้จากข้อมูลชุดที่ 3 ภายใต้การกำหนดให้ข้อมูลช่วงการแพร่ระบาดของโควิด-19 เป็นข้อมูลสูญหาย เนื่องจากมีค่า RMSE ต่ำที่สุด นั่นคือ ตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ ตามลำดับ

Table 2 Missing Value Imputation for Data Used in Building the Forecasting Model

Year	Month	Mu Ko Chang National Park			Khao Laem Ya – Mu Ko Samet National Park		
		Dataset 1	Dataset 2	Dataset 3	Dataset 1	Dataset 2	Dataset 3
2020	January	27,218	27,218	24,737	209,630	209,630	90,406
	February	23,254	23,254	24,572	62,450	62,450	98,085
	March	10,777	10,777	23,470	36,574	36,574	100,541
	April	0	32,340	35,134	0	116,757	108,973
	May	0	20,046	23,023	0	97,730	96,728
	June	0	8,047	9,304	0	61,013	73,358
	July	18,582	18,582	15,472	54,779	54,779	86,246
	August	7,052	7,052	11,998	36,470	36,470	78,792
	September	9,040	9,040	6,498	51,237	51,237	49,234
	October	15,518	15,518	21,673	56,327	56,327	78,217
	November	17,739	17,739	19,690	73,055	73,055	71,020
	December	19,097	19,097	26,259	58,669	58,669	90,247
2021	January	2,700	2,700	24,559	8,683	8,683	89,089
	February	6,228	6,228	24,391	35,143	35,143	96,707
	March	9,093	9,093	23,476	53,744	53,744	101,183
	April	19,815	19,815	34,807	75,024	75,024	108,900
	May	686	686	22,863	6,015	6,015	97,872
	June	0	8,391	9,752	31,395	31,395	74,371
	July	0	13,098	15,362	11,521	11,521	86,676
	August	0	10,530	12,289	5,568	5,568	78,931
	September	2,069	2,069	6,827	19,054	19,054	49,148
	October	11,657	11,657	21,710	61,993	61,993	78,825
	November	12,827	12,827	19,754	54,738	54,738	70,834
	December	23,590	23,590	26,072	88,084	88,084	89,004

Note: Dataset 1 – Real data with zero values ; Dataset 2 – Zero values treated as missing data

Dataset 3 – COVID-19 pandemic period treated as missing data

Table 3 Models and Parameter Estimates for Forecasting the Number of Tourists in Mu Ko Chang National Park for Each Dataset

Dataset	Model	AIC	RMSE	Parameter	Estimate	Std. Error	t-statistic	p-value
1	SARIMA(1,0,0)(0,1,1) ₁₂	1,705.17	5,382.99	ϕ_1	0.7013	0.0835	8.3994	< 0.001***
				Θ_1	-0.6120	0.1360	-4.4992	< 0.001***
2	SARIMA(1,0,1)(0,1,1) ₁₂	1,673.75	4,422.48	ϕ_1	0.9193	0.0676	13.6066	< 0.001***
				θ_1	-0.5498	0.1553	-3.5405	< 0.001***
				Θ_1	-0.6094	0.1278	-4.7671	< 0.001***
3	SARIMA(2,1,0)(0,1,1) ₁₂	1,597.20	3,132.69	ϕ_1	-0.3460	0.1059	-3.2677	0.002**
				ϕ_2	-0.2468	0.1052	-2.3461	0.021*
				Θ_1	-0.5901	0.1123	-5.2541	< 0.001***

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4 Models and Parameter Estimates for Forecasting the Number of Tourists in Khao Laem Ya – Mu Ko Samet National Park for Each Dataset

Dataset	Model	AIC	RMSE	Parameter	Estimate	Std. Error	t-statistic	p-value
1	SARIMA(1,0,0)(0,1,1) ₁₂	1,973.16	23,600.81	ϕ_1	0.8535	0.0587	14.5393	< 0.001***
				Θ_1	-1.0000	0.1747	-5.7237	< 0.001***
2	SARIMA(1,0,0)(0,1,1) ₁₂	1,968.17	22,913.24	ϕ_1	0.8431	0.0605	13.9400	< 0.001***
				Θ_1	-1.0000	0.1963	-5.0952	< 0.001***
3	SARIMA(0,1,0)(0,1,1) ₁₂	1,869.92	16,171.22	Θ_1	-0.7449	0.1613	-4.6189	< 0.001***

*** $p < .001$

นอกจากนี้ ค่าพารามิเตอร์ของตัวแบบดังกล่าวยังมีนัยสำคัญทางสถิติทุกตัว แสดงว่าสามารถนำค่าประมาณพารามิเตอร์ไปใช้ในการสร้างสมการพยากรณ์ได้ เมื่อแทนค่าประมาณพารามิเตอร์ที่ได้จากตัวแบบที่เหมาะสม จะได้สมการสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่ง ดังนี้

1. สมการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง ที่ได้จากตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ เมื่อไม่มีพจน์ของค่าคงที่ คือ

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.3460(Y_{t-1} - Y_{t-2} - Y_{t-13} + Y_{t-14}) - 0.2468(Y_{t-2} - Y_{t-3} - Y_{t-14} + Y_{t-15}) + 0.5901e_{t-12} \quad (5)$$

2. สมการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ที่ได้จากตัวแบบ SARIMA(0,1,0)(0,1,1)₁₂ เมื่อไม่มีพจน์ของค่าคงที่ คือ

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + 0.7449e_{t-12} \quad (6)$$

เพื่อตรวจสอบว่าตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ มีความเหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ดหรือไม่ จะต้องทำการตรวจสอบข้อสมมติเบื้องต้นของตัวแบบดังกล่าวก่อน โดยตัวแบบที่เหมาะสมนั้น ค่าคลาดเคลื่อนจะต้องมีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับศูนย์ มีความแปรปรวนคงที่ และเป็นอิสระต่อกัน ซึ่งผลการตรวจสอบข้อสมมติดังกล่าวสำหรับตัวแบบการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด จะแสดงใน Table 5

Table 5 Model Assumption Validation for Forecasting the Number of Tourists in Mu Ko Chang and Khao Laem Ya – Mu Ko Samet National Parks

Model (National Park)	Hypothesis Testing	Statistical Test	Statistic	p-value
SARIMA(2,1,0)(0,1,1) ₁₂ (Mu Ko Chang)	Normality	Anderson-Darling	2.9786	< 0.001***
	Zero Mean	t	-1.2979	0.197
	Homogeneity of Variance	Levene's	1.5992	0.112
	Independence	Ljung-Box (Lag = 24)	24.4610	0.436
SARIMA(0,1,0)(0,1,1) ₁₂ (Khao Laem Ya – Mu Ko Samet)	Normality	Anderson-Darling	3.2652	< 0.001***
	Zero Mean	t	1.4833	0.141
	Homogeneity of Variance	Levene's	1.3204	0.226
	Independence	Ljung-Box (Lag = 24)	23.6440	0.482

*** $p < .001$

จาก Table 5 พบว่า ตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ ไม่เป็นไปตามข้อสมมุติเบื้องต้นในเรื่องของการแจกแจงปกติ นั่นคือ ค่าคลาดเคลื่อนเบี่ยงเบนไปจากการแจกแจงปกติ ($p < 0.001$) ในขณะที่ผลการตรวจสอบข้อสมมุติเบื้องต้นอื่น ๆ ของค่าคลาดเคลื่อนทั้ง 2 ตัวแบบ เป็นไปตามข้อสมมุติเบื้องต้นทั้งหมดนั่นคือค่าคลาดเคลื่อนมีค่าเฉลี่ยเป็นศูนย์ มีความแปรปรวนคงที่ และเป็นอิสระต่อกัน ($p > 0.05$) ทั้งนี้ ผลการตรวจสอบความเหมาะสมของตัวแบบแสดงให้เห็นว่าผลการประมาณค่าอนุกรมเวลาที่มีความไม่ต่อเนื่อง ยังได้รับผลกระทบจากการแพร่ระบาดของโควิด-19 เนื่องจากตัวแบบการพยากรณ์ทั้ง 2 ตัวแบบ ไม่มีการแจกแจงปกติ จึงทำให้ตัวแบบดังกล่าวยังไม่ใช่วิธีการที่เหมาะสมที่สุดแต่อย่างไรก็ตาม ยังคงสามารถนำทั้ง 2 ตัวแบบ ไปใช้ในการพยากรณ์พยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่งได้

3. การประเมินความแม่นยำของการพยากรณ์

ผลการสร้างตัวแบบการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออก พบว่า ตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด คือ ตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ ตามลำดับ ถึงแม้ว่าตัวแบบดังกล่าวจะยังไม่ใช่วิธีการที่เหมาะสมที่สุด แต่ยังสามารถใช้ในการพยากรณ์ข้อมูลได้ โดยค่าจริงและค่าพยากรณ์ตั้งแต่เดือนตุลาคม พ.ศ. 2566 ถึงเดือนมีนาคม พ.ศ. 2568 จากทั้ง 2 ตัวแบบ แสดงไว้ใน Table 6 และดังแสดงใน Figure 3 และ 4

จากการเปรียบเทียบค่าจริงกับค่าพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างด้วยตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ ตามข้อมูลใน Table 6 และ Figure 3 พบว่า ค่าพยากรณ์มีความคลาดเคลื่อนจากค่าจริงค่อนข้างมากในช่วงเดือนมกราคม ถึงเดือนพฤษภาคม พ.ศ. 2567 และในช่วงเดือนพฤศจิกายน พ.ศ. 2567 ถึงเดือนมีนาคม พ.ศ. 2568 โดยตัวแบบดังกล่าวให้ค่า MAPE เท่ากับ 23.79 แสดงให้เห็นว่ามีความแม่นยำในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างอยู่ในเกณฑ์ที่สามารถยอมรับได้ ในทำนองเดียวกัน เมื่อพิจารณาค่าจริงกับค่าพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ด้วยตัวแบบ SARIMA(0,1,0)(0,1,1)₁₂ ตาม Table 6 และ Figure 4 จะเห็นได้ว่าค่าพยากรณ์ส่วนใหญ่สอดคล้องกับค่าจริง ยกเว้นในช่วงเดือนมกราคม ถึงเดือนเมษายน พ.ศ. 2567 และเดือนมกราคม พ.ศ. 2568 ที่ค่าจริงและค่าพยากรณ์แตกต่างกันอย่างชัดเจน โดยตัวแบบนี้ให้ค่า MAPE เท่ากับ 12.94 ซึ่งบ่งชี้ว่ามีความแม่นยำในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด อยู่ในเกณฑ์

เพื่อคาดการณ์แนวโน้มของนักท่องเที่ยวที่จะเดินทางมาท่องเที่ยวอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ดในอนาคต จึงได้ทำการพยากรณ์ข้อมูลล่วงหน้าตั้งแต่เดือนเมษายน ถึงเดือนธันวาคม พ.ศ. 2568 เพิ่มเติม ด้วยตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ ตามลำดับ โดยค่าพยากรณ์ที่ได้จากทั้ง 2 ตัวแบบ แสดงไว้ใน Table 7 และดังแสดงใน Figure 5 และ 6 ทั้งนี้ ผลการพยากรณ์แสดงให้เห็นว่าแนวโน้มของนักท่องเที่ยวที่เดินทางมายังอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่ง ในช่วงเวลาดังกล่าวค่อนข้างคงที่ตามรอบฤดูกาล

Table 6 Actual and Forecasted Values for the Number of Tourists in Mu Ko Chang and Khao Laem Ya – Mu Ko Samet National Parks from October 2023 to March 2025

Year	Month	Model (National Park)			
		SARIMA(2,1,0)(0,1,1) ₁₂ (Mu Ko Chang)		SARIMA(0,1,0)(0,1,1) ₁₂ (Khao Laem Ya – Mu Ko Samet)	
		Actual Value	Forecasted Value	Actual Value	Forecasted Value
2023	October	17,579	18,515	87,003	83,333
	November	16,763	16,571	67,645	73,041
	December	23,348	23,210	104,414	98,031
2024	January	20,808	17,723	108,827	79,755
	February	21,863	16,223	147,656	80,707
	March	21,027	18,269	137,520	98,953
	April	24,578	29,422	169,070	129,213
	May	7,570	17,178	119,095	106,920
	June	5,722	6,943	79,058	75,897
	July	7,894	11,054	100,221	91,939
	August	9,980	9,552	94,528	82,425
	September	5,503	4,346	56,275	54,885
	October	16,074	18,339	88,695	86,817
	November	21,741	16,203	79,050	76,524
	December	28,368	22,849	112,491	101,515
2025	January	25,686	17,407	117,803	83,239
	February	22,104	15,889	88,950	84,191
	March	21,032	17,930	99,047	102,437

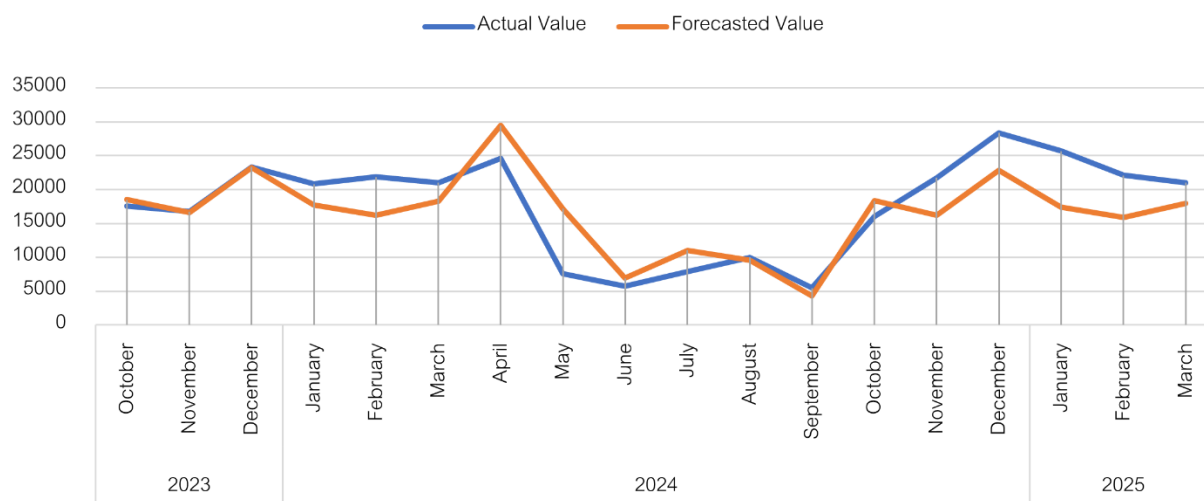


Figure 3 Actual and Forecasted Values for the Number of Tourists in Mu Ko Chang National Park from October 2023 to March 2025

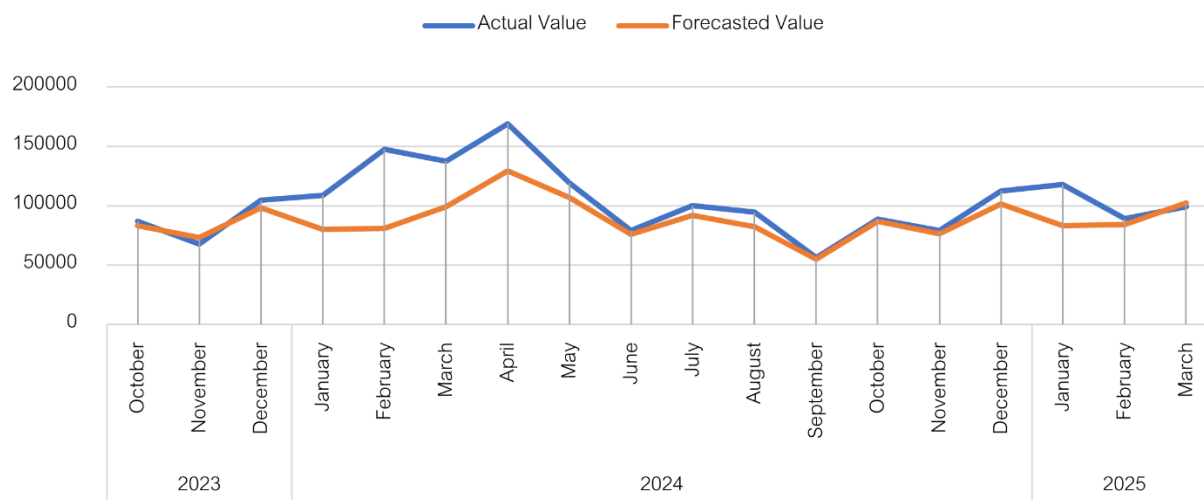


Figure 4 Actual and Forecasted Values for the Number of Tourists in Khao Laem Ya – Mu Ko Samet National Park from October 2023 to March 2025

Table 7 Forecasted Values for the Number of Tourists in Mu Ko Chang and Khao Laem Ya – Mu Ko Samet National Parks from April 2025 to December 2025

Year	Month	Model (National Park)	
		SARIMA(2,1,0)(0,1,1) ₁₂ (Mu Ko Chang)	SARIMA(0,1,0)(0,1,1) ₁₂ (Khao Laem Ya – Mu Ko Samet)
2025	April	29,089	132,697
	May	16,844	110,404
	June	6,609	79,381
	July	10,720	95,423
	August	9,218	85,908
	September	4,012	58,368
	October	18,005	90,301
	November	15,868	80,008
	December	22,515	104,998

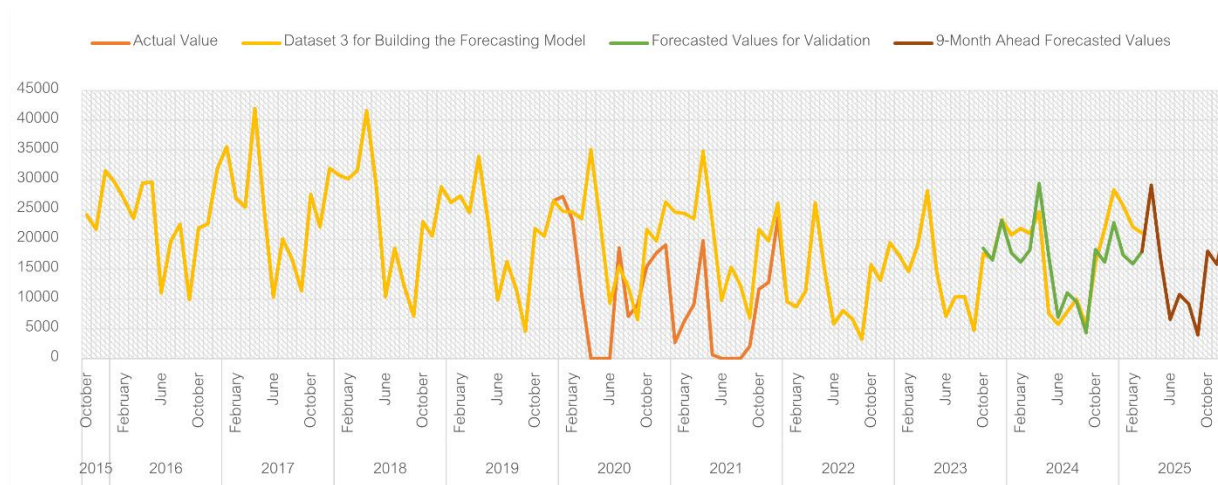


Figure 5 Actual Value, Dataset 3 for Building the Forecasting Model, Forecasted Values for Validation, and 9-Month Ahead Forecasted Values for the Number of Tourists in Mu Ko Chang National Park

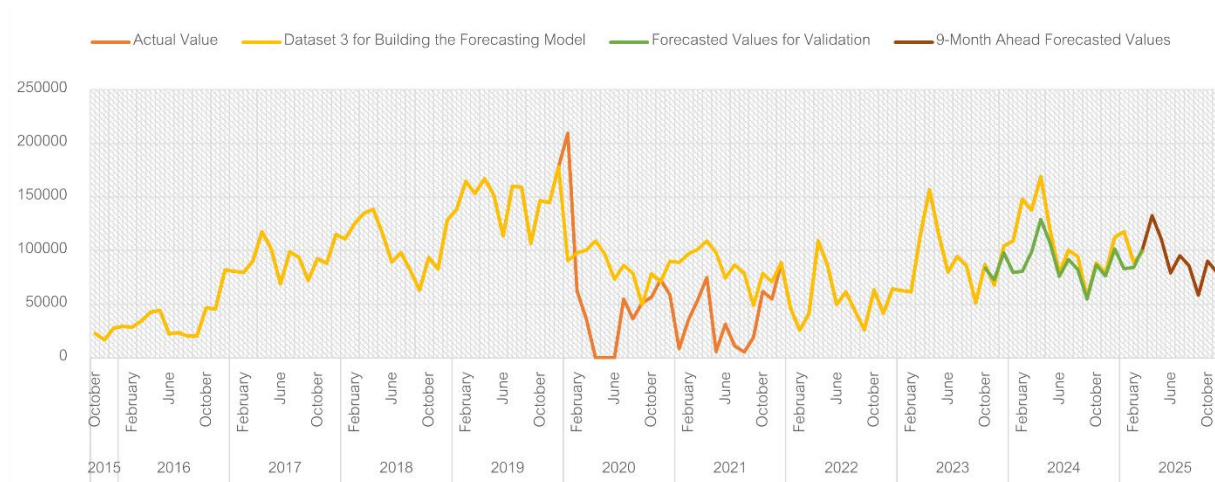


Figure 6 Actual Value, Dataset 3 for Building the Forecasting Model, Forecasted Values for Validation, and 9 Month Ahead Forecasted Values for the Number of Tourists in Khao Laem Ya – Mu Ko Samet National Park

Discussion

จากผลการศึกษาที่ได้ ถึงแม้ว่าในช่วงสถานการณ์การแพร่ระบาดของโควิด-19 จะทำให้ไม่มีนักท่องเที่ยวมาเยี่ยมชมอุทยานแห่งชาติทางทะเลทั้ง 2 แห่ง หรือข้อมูลจำนวนนักท่องเที่ยวมีค่าเป็นศูนย์ในบางช่วงเวลา แต่วิธีการประมาณค่าสูญหายช่วยให้ข้อมูลดังกล่าวมีความต่อเนื่อง และสามารถนำไปใช้ในการสร้างตัวแบบการพยากรณ์ได้อย่างมีประสิทธิภาพ การปรับปรุงข้อมูลในลักษณะนี้จึงเป็นขั้นตอนสำคัญที่ช่วยให้หน่วยงานสามารถใช้ผลพยากรณ์เพื่อการตัดสินใจเชิงนโยบาย และการจัดการในภาคปฏิบัติได้อย่างมีประสิทธิภาพมากขึ้น โดยเฉพาะในด้านการวางแผนล่วงหน้าและการจัดสรรทรัพยากรตามฤดูกาลท่องเที่ยว อย่างไรก็ตาม วิธีการนี้ยังมีข้อจำกัดหรือความเสี่ยงที่อาจทำให้เกิดความเอนเอียง หากค่าที่ประมาณได้มีความแตกต่างจากจำนวนนักท่องเที่ยวที่นำเข้ามาเยี่ยมชมในอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่ง ในกรณีที่ไม่มีกรณีการแพร่ระบาดของโควิด-19 (Hyndman & Rostami-Tabar, 2024; Mueller & Sobreira, 2024) อย่างไรก็ตาม ผลกระทบจากสถานการณ์ดังกล่าวยังคงมีอยู่ จนกว่าจะมีข้อมูลเพียงพอสำหรับการสร้างตัวแบบการพยากรณ์ที่มีความแม่นยำมากขึ้น

นอกจากนี้ ตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทั้ง 2 แห่ง ยังเป็นตัวแบบที่ได้จากการประมาณค่าสูญหายรูปแบบที่ 3 นั่นคือ กำหนดให้ข้อมูลช่วงการแพร่ระบาดของโควิด-19 ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนธันวาคม พ.ศ. 2564 เป็นข้อมูลสูญหาย ซึ่งสอดคล้องกับการศึกษาของ Mueller & Sobreira (2024) ที่พบว่า การกำหนดแนวทางในการแทนค่าข้อมูลในช่วง 2 ปี ดังกล่าว ช่วยลดผลกระทบจากสถานการณ์การแพร่ระบาดของโควิด-19 ได้ดีกว่ารูปแบบอื่น ๆ อีกทั้งวิธีการนี้ยังให้ค่าประมาณที่มีรูปแบบการเพิ่มขึ้นและลดลงสอดคล้องกับพฤติกรรมเดิมของข้อมูล ส่งผลให้ตัวแบบการพยากรณ์ที่สร้างขึ้นจากข้อมูลดังกล่าวสามารถรักษาน้ำหนักและโครงสร้างเชิงฤดูกาลของข้อมูลไว้ได้ อย่างไรก็ตาม หากค่าประมาณที่ได้ไม่สอดคล้องกับความเป็นจริงหรือเกิดความเอนเอียง ตัวแบบอาจให้ค่าพยากรณ์ที่

คลาดเคลื่อนได้ สำหรับรูปแบบการสูญหายของข้อมูลรูปแบบอื่น ๆ ได้แก่ รูปแบบที่ 1 การใช้ข้อมูลจริงที่มีค่าเป็นศูนย์ (ไม่มีข้อมูลสูญหาย) จะได้รับผลกระทบจากค่าผิดปกติในช่วงสถานการณ์การแพร่ระบาดของโควิด-19 ทำให้แนวโน้มและรูปแบบเชิงฤดูกาลเปลี่ยนแปลงไป ทำให้ไม่สะท้อนพฤติกรรมปกติของนักท่องเที่ยว ส่วนรูปแบบที่ 2 ที่กำหนดให้ค่าศูนย์เป็นข้อมูลสูญหาย ช่วยทำให้ลดผลกระทบจากค่าผิดปกติได้ แต่ยังไม่สามารถรักษารูปแบบความผันผวนตามฤดูกาลได้อย่างครบถ้วน ดังนั้น การเลือกใช้รูปแบบการประมาณค่าสูญหายที่เหมาะสม จึงขึ้นอยู่กับลักษณะของข้อมูลและสถานการณ์ของการกำหนดค่าสูญหาย โดยในการพัฒนางานวิจัยในอนาคต การปรับปรุงวิธีประมาณค่าสูญหายให้มีความยืดหยุ่นและสามารถตอบสนองต่อการเปลี่ยนแปลงของข้อมูลได้ดียิ่งขึ้น จะช่วยเพิ่มความแม่นยำและความน่าเชื่อถือของผลการพยากรณ์

Conclusions

ในการศึกษาการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด ได้กำหนดรูปแบบการสูญหายของข้อมูลสำหรับการสร้างตัวแบบการพยากรณ์ออกเป็น 3 รูปแบบ คือ 1) ใช้ข้อมูลจริงที่มีค่าเป็นศูนย์ 2) กำหนดให้ข้อมูลที่เป็นศูนย์เป็นข้อมูลสูญหาย และ 3) กำหนดให้ข้อมูลช่วงโควิด-19 ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนธันวาคม พ.ศ. 2564 เป็นข้อมูลสูญหาย พร้อมทั้งทำการทดแทนค่าสูญหายดังกล่าวด้วยการแยกส่วนประกอบฤดูกาลร่วมกับขั้นตอนวิธีค่าเฉลี่ย เพื่อให้ค่าประมาณจำนวนนักท่องเที่ยวที่ได้ไม่ติดลบ หรือหลีกเลี่ยงการได้ค่าประมาณที่ไม่สมเหตุผล จากนั้นสร้างตัวแบบการพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ โดยเลือกตัวแบบที่พารามิเตอร์ทุกตัวมีนัยสำคัญทางสถิติและมีค่าเกณฑ์สารสนเทศอะกะอิเกะ (AIC) ต่ำที่สุดในแต่ละชุดข้อมูลที่ได้ทำการประมาณค่าสูญหาย ต่อมา พิจารณาเลือกตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติทั้ง 2 แห่ง โดยใช้รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ที่ต่ำที่สุด พร้อมทั้งตรวจสอบความแม่นยำของการพยากรณ์ของตัวแบบที่เหมาะสม ด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)

ผลการศึกษาการสร้างตัวแบบการพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ พบว่า ตัวแบบที่มีความเหมาะสมในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้าง และอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด คือ ตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ และ SARIMA(0,1,0)(0,1,1)₁₂ ตามลำดับ ซึ่งเป็นตัวแบบที่ได้จากการประมาณค่าข้อมูลสูญหายในช่วงการแพร่ระบาดของโควิด-19 และจากการเปรียบเทียบค่าจริงกับค่าพยากรณ์ พบว่า ตัวแบบ SARIMA(2,1,0)(0,1,1)₁₂ มีความแม่นยำในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติหมู่เกาะช้างอยู่ในเกณฑ์ที่พอยอมรับได้ ในขณะที่ตัวแบบ SARIMA(0,1,0)(0,1,1)₁₂ มีความแม่นยำในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติเขาแหลมหญ้า – หมู่เกาะเสม็ด อยู่ในเกณฑ์ดี ทั้งนี้ เมื่อทำการพยากรณ์ข้อมูลล่วงหน้าตั้งแต่เดือนเมษายน ถึงเดือนธันวาคม พ.ศ. 2568 พบว่าจำนวนนักท่องเที่ยวในอุทยานแห่งชาติทางทะเลภาคตะวันออกทั้ง 2 แห่ง ค่อนข้างคงที่ตามรอบฤดูกาล

แม้ว่าตัวแบบการพยากรณ์จำนวนนักท่องเที่ยวที่ใช้ในงานวิจัยนี้อาจจะยังไม่ใช่วิธีที่เหมาะสมที่สุด แต่หลังจากการปรับปรุงคุณภาพของข้อมูลให้มีความต่อเนื่อง ตัวแบบที่ได้สามารถนำไปใช้ในการพยากรณ์เบื้องต้นได้อย่างมีประสิทธิภาพ ซึ่งจะช่วยให้หน่วยงานที่เกี่ยวข้องสามารถวางแผนและบริหารจัดการการท่องเที่ยวได้ดีขึ้น โดยอาจจะเป็นการพัฒนากลยุทธ์

การจัดสรรบุคลากร การจัดสรรงบประมาณ การวางแผนดูแลทรัพยากรธรรมชาติ รวมถึงการวางแผนกิจกรรมส่งเสริมการท่องเที่ยวให้สอดคล้องกับฤดูกาลและแนวโน้มของนักท่องเที่ยว เพื่อสนับสนุนการพัฒนาการท่องเที่ยวอย่างยั่งยืนในระยะยาว นอกจากนี้ เพื่อเพิ่มประสิทธิภาพและความแม่นยำของการพยากรณ์ในอนาคต การปรับปรุงตัวแบบถือเป็นปัจจัยสำคัญที่จะช่วยให้ตัวแบบทันสมัยและตอบสนองต่อการเปลี่ยนแปลงของข้อมูลได้อย่างเหมาะสม อีกทั้ง การศึกษาเพิ่มเติมเกี่ยวกับการพยากรณ์ด้วยเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) หรือการเรียนรู้เชิงลึก (Deep Learning) อาจช่วยยกระดับความเหมาะสมและความแม่นยำของตัวแบบให้สูงขึ้น

References

- Anekkunawut, P. (2020). *Forecasting International Tourist Arrivals to Thailand using Time Series and Artificial Neuron Network* [Master's thesis, Chulalongkorn University]. Chula DigiVerse.
doi.org/10.58837/ CHULA.THE.2020.1171 (in Thai)
- Arttaweekul, S. (2013). *Forecasting Amount of Foreign Tourists and Influence of Economy's Indicator to the Forecasting Amount of Foreign Tourists* [Master's thesis, Thammasat University]. TU Digital Collection.
Retrieved from https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:95936 (in Thai)
- Boonlha, K. (2019). Forecasting Numbers of Tourists in the Phu Hin Rong Kla National Park. *Science and Technology Nakhon Sawan Rajabhat University Journal*, 11(14), 1-12. (in Thai)
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 9(1), 3-73.
- Department of National Parks, Wildlife and Plant Conservation. (2021). *Tourist Data in National Parks*. April 7, 2025, Retrieved from <https://catalog.dnp.go.th/dataset/stat-tourism> (in Thai)
- Hyndman, R. J., & Rostami-Tabar, B. (2024). Forecasting Interrupted Time Series. *Journal of the Operational Research Society*, 1-14. doi.org/10.1080/01605682.2024.2395315
- Intarapak, S., Supapakorn, T., & Vuthipongse, W. (2022). Classical Forecasting of International Tourist Arrivals to Thailand. *Journal of Statistical Theory and Applications*, 21(2), 31-43.
doi.org/10.1007/s44199-022-00041-5

- Kang, H. (2013). The Prevention and Handling of the Missing Data. *Korean Journal of Anesthesiology*, 64(5), 402-406. doi.org/10.4097/kjae.2013.64.5.402
- Keerativibool, W. (2013). Forecasting Model for the Number of International Tourist Arrivals to Thailand. *Srinakharinwirot Science Journal*, 29(2), 9-26. (in Thai)
- Kummong, R., & Homdok, S. (2019). Forecasting Foreign Tourists in Northern Thailand Using Artificial Neural Network. *Academic Journal of Science and Applied Science*, 3(6), 15-31. (in Thai)
- Lake, P. (2016). Forecasting International Tourist Arrival to Thailand by Nationality Using Time Series Data Mining Techniques. In *KU SRC 1st National Conference 2016*. (pp. 453-463). Thailand: Chon Buri. (in Thai)
- Lewis, C. D. (1982). *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*. London: Butterworth Scientific.
- Lorchirachoonkul, V., & Jitthavech, J. (2005). *Forecasting Techniques*. Bangkok: National Institute of Development Administration. (in Thai)
- Mahavanakul, W., & Prasongsak, N. (2023, June 13). How Has Thai Tourism Recovered Spatially?. *Krungthep Thurakit*, 1-4. (in Thai)
- Manmin, M. (2006). *Time Series and Forecasting*. Bangkok: Foreprinting (in Thai)
- Ministry of Tourism and Sports. (2021). *Annual Report 2019*. Nonthaburi: Parbpim. (in Thai)
- Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time Series Missing Value Imputation in R. *R Journal*, 9(1), 207-218. doi.org/10.32614/RJ-2017-009
- Moritz, S., Gatscha, S., Wang, E., & Hause, R. (2022). *Package 'imputeTS'*. [Computer software]. CRAN - R Project. Retrieved from <https://cran.rstudio.com/>
- Mueller, R., & Sobreira, N. (2024). Tourism Forecasts After COVID-19: Evidence of Portugal. *Annals of Tourism Research Empirical Insights*, 5(1), 1-14. doi.org/10.1016/j.annale.2024.100127

Mukma, M., Chomtee, B., & Payakkapong, P. (2018). Forecasting the Number of Chinese Tourists in Thailand.

Thai Science and Technology Journal, 26(3), 417-428. (in Thai)

Saothayanun, L., Taweesakulvatchara, S., Kanjanasakda, Y., & Somrang, B. (2014). A Forecasting Methods for the Number of International Tourists in Thailand: Box-Jenkins Method and Winter's Method.

Thai Science and Technology Journal, 22(1), 89-98. (in Thai)