

การพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ จังหวัดเพชรบูรณ์

ธิดิมา กิ่งกระโทก, ศ.ม.¹

วัฒนา ชยธวัช, ปร.ค.²

1. กลุ่มงานควบคุมโรคติดต่อ สำนักงานสาธารณสุขจังหวัดเพชรบูรณ์ จังหวัดเพชรบูรณ์,

2. คณะสหเวชศาสตร์ มหาวิทยาลัยปทุมธานี จังหวัดปทุมธานี

วันที่รับบทความ (Received), 20 พฤศจิกายน 2566 วันที่แก้ไขบทความ (Revised), 13 ธันวาคม 2566

วันที่ตอบรับบทความ (Accepted), 3 มกราคม 2566

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบรายเดือน ปี พ.ศ. 2566 ของจังหวัดเพชรบูรณ์ โดยใช้ข้อมูลจำนวนผู้ป่วยโรคปอดอักเสบรายเดือนจากสำนักระบาดวิทยา กรมควบคุมโรค ปี พ.ศ. 2546 ถึง 2565 ด้วยวิธีบอซซ์และเจนกินส์ แบบจำลองฤดูกาล ARIMA หรือ SARIMA (Seasonal Autoregressive Integrated Moving Average) วิธีการแยกตัวประกอบแบบจำลองการคูณที่คำนวณดัชนีฤดูกาลตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนที่ (MMRMA) และแบบจำลองตามทฤษฎีระบบเกรย์แล้วแจกแจง ค่าพยากรณ์รายเดือนด้วยดัชนีฤดูกาลตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนที่ (GREYRMA) ผลการวิจัยพบว่า จำนวนผู้ป่วยมีจำนวนค่อนข้างสม่ำเสมอในช่วงต้นปีถึงกลางปี และเพิ่มสูงขึ้นในช่วงเดือนกันยายนและตุลาคม และแต่ละปีมีแนวโน้มเพิ่มขึ้น การเลือกแบบจำลองจากเกณฑ์พิจารณาความแม่นยำด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error, MAPE) เมื่อพิจารณาค่าจริงเทียบกับ ค่าพยากรณ์ในอดีต 240 เดือน แบบจำลอง SARIMA มี MAPE ร้อยละ 21.11 ต่ำกว่า MMRMA ร้อยละ 10.19 แต่เมื่อพิจารณาค่าพยากรณ์รายเดือน ปี พ.ศ. 2566 เปรียบเทียบกับค่าจริงเดือน มกราคม ถึง ตุลาคม พบว่า มีค่า MAPE ของวิธี SARIMA, GREYRMA และ MMRMA ร้อยละ 63.32, 82.81, และ, 41.74 ตามลำดับ การพยากรณ์ด้วยแบบจำลอง MMRMA มีความใกล้เคียงกับข้อมูลจริงในปี พ.ศ. 2566 มากกว่า จึงเป็นค่าพยากรณ์ที่น่าจะยอมรับได้สำหรับเดือน พฤศจิกายน และธันวาคมที่เหลืออยู่

คำสำคัญ: การพยากรณ์อนุกรมเวลา, จำนวนผู้ป่วยโรคปอดอักเสบ, จังหวัดเพชรบูรณ์

Forecasting Pneumonia Incidents: Phetchabun Province

Thitima Kingkratoke, M.P.H.¹

Vadhana Jayathavaj, Ph.D.²

1. Communicable Disease Control Unit, Phetchabun Provincial Public Health Office., Phetchabun Province., 2. Faculty of Allied Health Sciences, Pathumthani University, Pathum Thani Province.

Abstract

This research aims to forecast the number of monthly pneumonia patients in 2023 in Phetchabun Province. The number of monthly pneumonia patients from 2003 to 2022 came from the Bureau of Epidemiology, Department of Disease Control, using the Box and Jenkins method (Seasonal ARIMA or SARIMA (Seasonal Autoregressive Integrated Moving Average), the decomposition method (the multiplicative model) (MMRMA), and the model based on Gray Systems Theory and distributing monthly forecast values with the seasonal index according to the ratio to moving average method (GREYRMA). The results showed that the number of patients was consistent from the beginning to the middle of the year and increased during September and October. In each year, there was a trend of increasing. Model selection was based on the accuracy criteria of mean absolute percentage error (MAPE). When considering actual values against forecast values over the past 240 months, the SARIMA model had a MAPE of 21.11 percent, lower than MMRMA by 10.19 percent. But when considering the monthly forecast values for the year 2023 compared to the actual values from January to October, it was found that the MAPE values of the SARIMA, GRAYRMA, and MMRMA methods were 63.32%, 82.81%, and 41.74%, respectively. Forecasting with the MMRMA model is closer to actual data for 2023 and therefore is an acceptable forecast for the remaining months of November and December.

Keywords: Time Series Forecasting, Pneumonia Incidents, Phetchabun Province

1. บทนำ

ปอดอักเสบ เป็นโรคที่พบได้ประมาณร้อยละ 8-10 ของผู้ป่วยที่มีการติดเชื้อระบบหายใจเฉียบพลัน เป็นสาเหตุการตายอันดับหนึ่งของโรคติดเชื้อในเด็กอายุต่ำกว่า 5 ปี เกิดจากสาเหตุหลัก 2 กลุ่ม คือ ปอดอักเสบที่เกิดจากการติดเชื้อและปอดอักเสบที่ไม่ได้เกิดจากการติดเชื้อ โดยทั่วไปพบปอดอักเสบที่เกิดจากการติดเชื้อมากกว่าชนิดของปอดอักเสบ จำแนกได้หลายแบบ ปัจจุบันนิยมจำแนกตามสภาพแวดล้อมที่เกิดปอดอักเสบเป็น ปอดอักเสบในชุมชน (community-acquired pneumonia - CAP) และปอดอักเสบในโรงพยาบาล (nosocomial pneumonia หรือ hospital-acquired pneumonia -HAP)⁽¹⁾

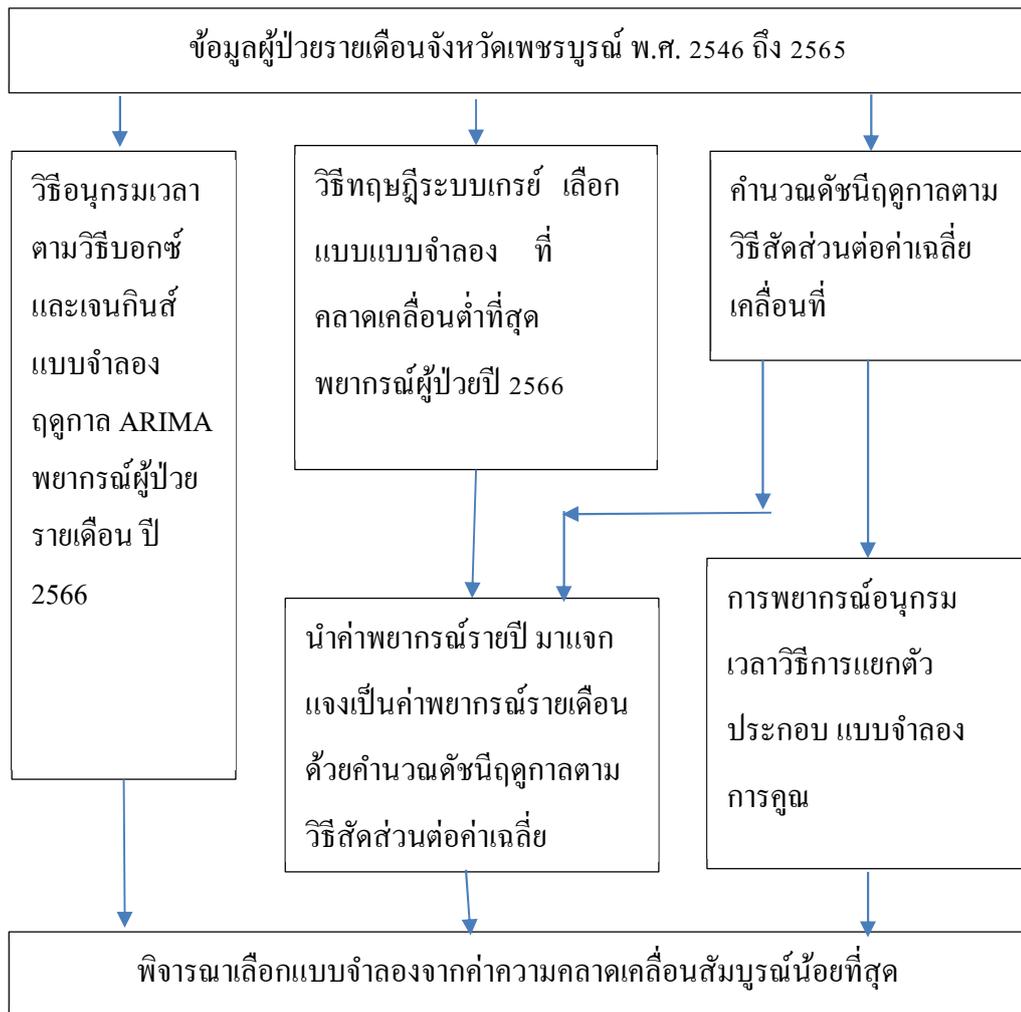
การพยากรณ์โรคปอดอักเสบ ปี พ.ศ. 2565 ได้ใช้ข้อมูลผู้ป่วยโรคปอดอักเสบรายเดือน ระหว่างปี พ.ศ. 2554 ถึง 2564 ด้วยแบบจำลอง ARIMA (2,0,0) (0,1,1)₁₂ มีค่าความคลาดเคลื่อนของการพยากรณ์ร้อยละ 11.70 คาดการณ์ว่าในปี พ.ศ. 2565 ประเทศไทยจะพบผู้ป่วยโรคปอดอักเสบประมาณ 191,814 คน ซึ่งค่าการพยากรณ์ที่ได้จะพบจำนวนผู้ป่วยสูง 2 ช่วง คือ ช่วงที่ 1 ช่วงต้นปี (เดือนมกราคม-มีนาคม) จะพบผู้ป่วยประมาณ 16,000 คนต่อเดือน และช่วงที่ 2 (เดือนสิงหาคม-พฤศจิกายน) จะพบผู้ป่วยประมาณ 20,000 คนต่อเดือน⁽²⁾ ข้อมูลจากระบบรายงานการเฝ้าระวังโรค 506 ปี พ.ศ. 2565 มีผู้ป่วย 234,196 คน เสียชีวิต 284 คน⁽³⁾ ข้อมูลจำนวนผู้ป่วย สัปดาห์ที่ 43 ระหว่างวันที่ 22-28 ตุลาคม 2566⁽⁴⁾ ปี พ.ศ. 2566 มีผู้ป่วย 244,353 คน เสียชีวิต 235 คน⁽⁵⁾

การศึกษาการพยากรณ์เปรียบเทียบเกี่ยวกับจำนวนผู้ป่วยโรคปอดอักเสบในประเทศไทยได้ใช้วิธีการแยกส่วนประกอบ วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลของวินเตอร์เชิงผลคูณ และวิธีบอซซ์และเจนกินส์โดยใช้ข้อมูลผู้ป่วยรายไตรมาส ตั้งแต่ พ.ศ. 2551 ถึง พ.ศ. 2561 เปรียบเทียบตัวแบบพยากรณ์ที่เหมาะสมภายใต้เกณฑ์ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (mean absolute percentage error, MAPE) ต่ำที่สุด พบว่าวิธีบอซซ์-เจนกินส์ให้ค่า MAPE ต่ำที่สุด ตัวแบบที่เหมาะสมในการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ คือ ตัวแบบรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่แบบมีฤดูกาล ARIMA (0,1,1) (0,1,1)₄⁽⁶⁾

การศึกษาการทำนายในพื้นที่จังหวัดเพชรบูรณ์พบ การพยากรณ์การเกิดโรคไข้เลือดออก ในพื้นที่จังหวัดเพชรบูรณ์ พ.ศ. 2558⁽⁷⁾ และปัจจัยทำนายพฤติกรรมส่งเสริมสุขภาพของผู้สูงอายุตำบลนางัว อำเภอเมืองจังหวัดเพชรบูรณ์⁽⁸⁾ แต่ยังไม่มีการศึกษาการพยากรณ์เกี่ยวกับจำนวนผู้ป่วยโรคปอดอักเสบทั้ง ๆ ที่มีรายงานจำนวนผู้ป่วยโรคปอดอักเสบ เป็นอันดับ 1 ของเขตสุขภาพที่ 3 ตั้งแต่ปี พ.ศ. 2561 เป็นต้นมา ระหว่างปี พ.ศ. 2561 ถึง 2565 มีผู้ป่วยน้อยที่สุดในปี พ.ศ. 2563 จำนวน 3,365 คน สูงที่สุดในปี พ.ศ. 2565 จำนวน 6,118 คน ปีอื่น ๆ มีผู้ป่วยไม่ต่ำกว่าจะ 4,000 คน ต่อปี⁽⁹⁾

2. วัตถุประสงค์ของการวิจัย

เพื่อพยากรณ์จำนวนผู้ป่วยรายเดือนในปี พ.ศ. 2566 ของจังหวัดเพชรบูรณ์ โดยใช้วิธีบอซซ์และเจนกินส์ แบบจำลองฤดูกาล ARIMA หรือ SARIMA (Seasonal Autoregressive Integrated Moving Average) วิธีการแยกตัวประกอบ แบบจำลองการคูณที่คำนวณดัชนีฤดูกาลตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนที่ และแบบจำลองตามทฤษฎีระบบเกรย์แล้วแจกแจงค่าพยากรณ์รายเดือนด้วยดัชนีฤดูกาลตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนที่ เพื่อพัฒนาเทคนิคการพยากรณ์อนุกรมเวลาในการพยากรณ์จำนวนผู้ป่วยและปรับใช้ค่าพยากรณ์ตามสถานการณ์ที่เกี่ยวข้องในอนาคตต่อไป โดยมีกรอบแนวคิดการวิจัยดังแสดงในภาพที่ 2 ดังนี้



ภาพที่ 1 กรอบแนวคิดการวิจัย

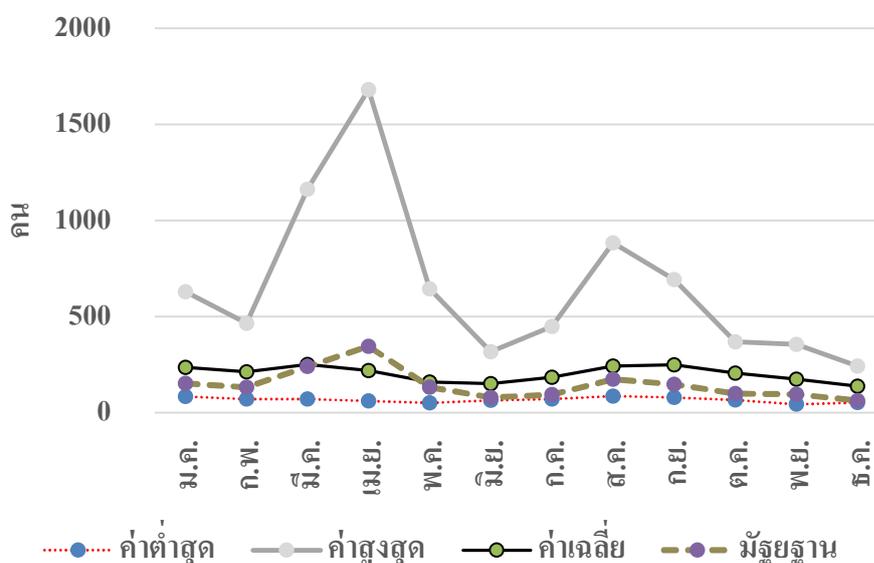
ตารางที่ 1 จำนวนผู้ป่วยโรคปอดอักเสบรายเดือน จังหวัดเพชรบูรณ์ พ.ศ. 2546 ถึง 2565

ปีที่	พ.ศ.	รวม	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
1	2546	1,051	89	70	78	64	71	69	93	154	130	91	73	69
2	2547	1,014	84	100	108	76	52	84	143	118	78	76	43	52
3	2548	1,096	94	74	110	66	71	67	159	157	103	65	69	61
4	2549	992	91	92	70	67	51	63	83	86	118	110	96	65
5	2550	1,239	124	145	118	69	79	69	70	119	155	152	77	62
6	2551	1,166	128	99	120	60	83	82	98	109	151	110	70	56
7	2552	1,366	117	122	87	69	65	93	155	169	99	145	128	117
8	2553	1,398	142	102	83	96	85	92	86	169	203	135	108	97
9	2554	1,709	164	117	136	114	96	105	133	187	216	180	113	148
10	2555	1,734	173	120	176	115	118	115	172	206	190	109	120	120
11	2556	1,855	175	134	128	86	109	131	164	183	187	183	149	226
12	2557	2,543	242	244	277	177	165	147	199	238	295	209	187	163
13	2558	2,770	260	299	264	198	173	198	173	226	255	291	240	193
14	2559	2,983	291	309	320	180	124	181	170	291	350	341	256	170
15	2560	3,351	340	260	282	178	194	250	243	322	433	362	271	216
16	2561	4,298	529	413	428	305	286	317	385	514	449	265	214	193
17	2562	4,039	462	414	416	340	283	278	245	269	367	368	355	242
18	2563	3,365	629	448	401	196	195	202	191	201	205	286	285	126
19	2564	4,155	303	206	249	223	232	181	448	883	692	313	275	150
20	2565	6,118	266	464	1163	1681	644	274	244	221	299	301	339	222
ค่าต่ำสุด			84	70	70	60	51	63	70	86	78	65	43	52
ค่าสูงสุด			629	464	1163	1681	644	317	448	883	692	368	355	242
ค่าเฉลี่ย			235	212	251	218	159	150	183	241	249	205	173	137
มัธยฐาน			151	132	239	345	132	79	94	174	147	99	96	63
ส่วนเบี่ยงเบนมาตรฐาน			174	140	156	115	114	123	167	194	204	182	139	137

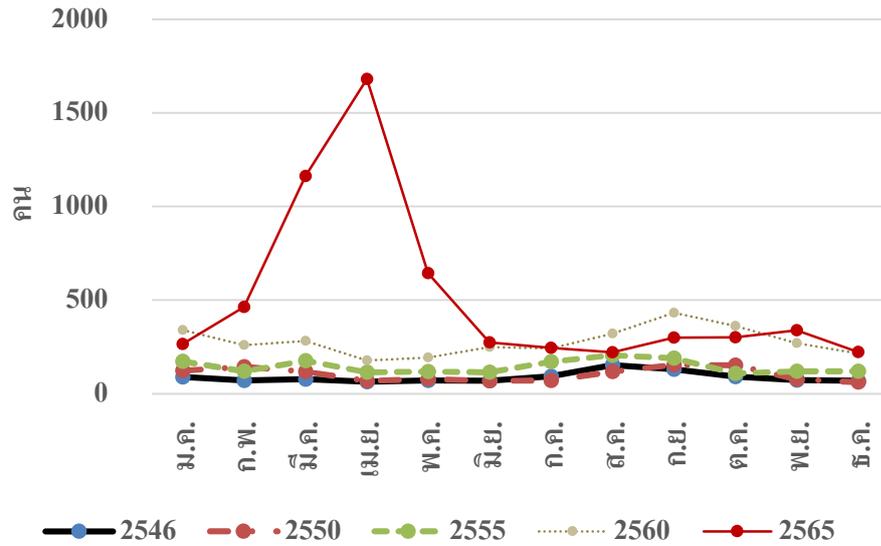
3. วิธีดำเนินการวิจัย

3.1 การรวบรวมข้อมูล

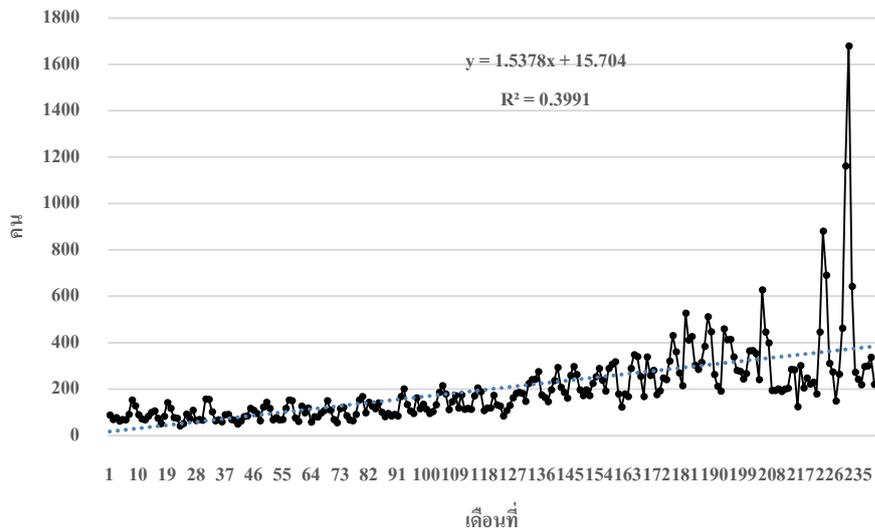
ข้อมูลที่ใช้ในการวิจัยในครั้งนี้ คือ จำนวนผู้ป่วยโรคปอดอักเสบรายเดือน ซึ่งเป็นข้อมูลทุติยภูมิ จาก สำนักระบาดวิทยา กรมควบคุมโรค พ.ศ. 2546 ถึง 2565⁽⁹⁾ ข้อมูลจำนวนผู้ป่วยโรคปอดอักเสบรายเดือน สำนักระบาดวิทยา กรมควบคุมโรค ระบบรายงานการเฝ้าระวังโรค 506 > Pneumonia > ข้อมูลย้อนหลัง > สรุปสถานการณ์รายปี พ.ศ. 2546 ถึง 2565⁽⁹⁾ และแสดงค่าสถิติของจำนวนผู้ป่วยโรคปอดอักเสบ ดังแสดงใน ตารางที่ 1 ค่าสถิติ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน จำแนกรายเดือนจากข้อมูล 20 ปี พ.ศ. 2546 ถึง 2565 ดังแสดงในภาพที่ 2



ภาพที่ 2 ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของ จำนวนผู้ป่วยโรคปอดอักเสบ พ.ศ. 2546 ถึง 2565 จำแนกรายเดือน



ภาพที่ 3 จำนวนผู้ป่วยโรคปอดอักเสบ พ.ศ. 2546 ถึง พ.ศ. 2565



ภาพที่ 4 จำนวนผู้ป่วยโรคปอดอักเสบรายเดือน 20 ปี พ.ศ. 2546 ถึง 2565

3.2 การพิจารณารูปแบบข้อมูล

จำนวนข้อมูล รูปแบบตามเวลาที่มีการเกิดซ้ำช่วงเวลาเดิม (Seasonal - S) เพิ่มขึ้นหรือลดลง (Trend - T) มีการเปลี่ยนแปลงตามวัฏจักร (Cyclical - C) และมีความเปลี่ยนแปลงเพิ่มมากขึ้นหรือลดลงอย่างผิดปกติ (Irregular - I) ทั้งนี้ เพื่อการเลือกใช้วิธีและแบบจำลองที่เหมาะสม จำนวนผู้ป่วยมีจำนวนค่อนข้างสม่ำเสมอในช่วงต้นปีถึงกลางปี และเพิ่มสูงขึ้นในช่วงเดือนกันยายนและตุลาคม ซึ่งมีลักษณะขององค์ประกอบฤดูกาล ขณะที่ปี พ.ศ. 2565 ข้อมูลรายเดือนมีลักษณะเพิ่มสูงมากช่วงเดือน มีนาคม เมษายน และพฤษภาคม ดังแสดงในภาพที่ 3 จากการเรียงข้อมูล ปี พ.ศ. 2546 ถึง 2565 (แกน x ในภาพที่ 4 เมื่อ x คือ เวลา/เดือน ที่มีค่าเป็น

1,2,3,...,235 เมื่อ 1 หมายถึง เดือนที่ 1 (ม.ค. พ.ศ. 2546) และ 235 (ธ.ค. 2565)) พบว่ามีแนวโน้มเพิ่มขึ้น (สมการถดถอยเชิงเส้นตรงในภาพที่ 4 $y=1.5378x+15.704$ เมื่อ y คือ จำนวนผู้ป่วย (คน) มีความชัน 1.5378 คนต่อเดือน) ดังแสดงในภาพที่ 4

3.3 วิธีบ็อกซ์และเจนกินส์

วิธี บ็อกซ์ และเจนกินส์ แบบจำลอง SARIMA (Seasonal AutoRegressive Integrated Moving Average) คือ ARIMA (p,d,q) (P,D,Q)_m ซึ่งหมายถึง แบบจำลอง ARIMA (p, d, q) มีส่วนประกอบที่สำคัญ 3 ส่วน ได้แก่ AutoRegressive AR (p), Intergrated (I) และ Moving Average MA (q) สำหรับ AR (p) เป็นรูปแบบที่แสดงว่า ค่าสังเกต y ณ เวลา t ขึ้นอยู่กับค่าของ y_{t-1}, \dots, y_{t-p} หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ MA (q) เป็นรูปแบบที่แสดงค่าสังเกต y ณ เวลา t ขึ้นอยู่กับค่าความคลาดเคลื่อน e_{t-1}, \dots, e_{t-q} หรือความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า ส่วน Integrated (I) เป็นการหาผลต่าง (Difference, d) ของอนุกรมเวลา เหตุผลสำคัญที่ต้องหาผลต่างของอนุกรมเวลา เนื่องจากแบบจำลอง ARIMA ใช้ในการวิเคราะห์ข้อมูล อนุกรมเวลาที่มีคุณสมบัติคงที่ (Stationary) เท่านั้น (อนุกรมเวลาที่อยู่หนึ่งมีคุณสมบัติทางสถิติ เช่น ค่าเฉลี่ย และความแปรปรวนที่ไม่แปรผันตามเวลา) ในกรณีที่ข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์สมบัติไม่คงที่ (Nonstationary) จะต้องทำการแปลงข้อมูลอนุกรมเวลาดังกล่าวให้มีคุณสมบัติคงที่ก่อน โดยการหาผลต่างของข้อมูลอนุกรมเวลา SARIMA มีพารามิเตอร์ตามฤดูกาลอีกสามรายการ (P, D, Q) พารามิเตอร์เพิ่มเติมอีกสามรายการคำนึงถึงองค์ประกอบการถดถอยอัตโนมัติ (P) องค์ประกอบผลต่าง (D) และองค์ประกอบค่าเฉลี่ยเคลื่อนที่ (Q) ในระดับฤดูกาล โดยที่ m คือจำนวนการค่าสังเกต ในแต่ละฤดูกาล (ในการศึกษาอนุกรมเวลาที่มีส่วนประกอบฤดูกาล 12 เดือน m คือ 12 ส่วนประกอบตามฤดูกาล ของแบบจำลองจะแสดงด้วยตัวพิมพ์ใหญ่ และส่วนประกอบที่ไม่ใช่ตามฤดูกาลของโมเดลจะแสดงด้วยตัวพิมพ์เล็ก⁽¹⁰⁾

การใช้โปรแกรม R⁽¹¹⁾ กับ Package Forecast^(12,13) ฟังก์ชัน auto.arima จะทำการคำนวณตัวแบบต่าง ๆ ตามแบบจำลอง SARIMA (Seasonal AutoRegressive Integrated Moving Average) เนื่องจากมีข้อมูลจำนวนผู้ป่วยรายเดือน อย่างน้อย 3 ปี ก็สามารถพยากรณ์รายเดือนล่วงหน้า 12 เดือนได้ การหาค่าพารามิเตอร์สำหรับแบบจำลอง SARIMA โดยใช้โปรแกรม R Forecast package เมื่อ install.packages ("forecast") ไว้แล้ว ก็เรียกใช้ library (forecast) เมื่อนำข้อมูลเข้าโปรแกรม R ก็สามารถเรียก ฟังก์ชัน auto.arima⁽¹⁴⁾ ทำการคำนวณจากแบบจำลอง ARIMA (p,d,q) (P,D,Q)_m ตามค่าพารามิเตอร์ p, d, q, P, D, Q, และ m ที่กำหนดจำนวนมาแล้วเลือกตัวแบบที่มีเกณฑ์ค่าพยากรณ์มีความสอดคล้องกับข้อมูลมากที่สุด หรือ Akaike Information Criterion (AIC) ที่มีค่าต่ำที่สุด⁽¹⁵⁾

3.4 วิธีทฤษฎีระบบเกรย์

แบบจำลองจากทฤษฎีระบบเกรย์ ซึ่งประกอบไปด้วยแบบจำลอง GM (1,1), GM (1,1) Error Periodic Correction (GM (1,1)EPC, Grey Verhulst Model (GVM) และ The Grey Verhulst improvement

model (GVIM) แล้วเลือกแบบจำลองที่มีความแม่นยำสูง แล้วจึงนำค่าพยากรณ์รายปีมาแจกแจงเป็นรายเดือน ด้วยดัชนีฤดูกาล เนื่องจากทฤษฎีระบบเกรย์ใช้กับข้อมูลจำนวนน้อย ทฤษฎีระบบเกรย์ใช้ในการพยากรณ์อนุกรมเวลา และสามารถใช้กับข้อมูลจำนวนน้อยที่สุดเพียง 4 คาบเวลาก็สามารถใช้ทำนายคาบเวลาที่ 5 ได้ แบบจำลองเริ่มต้นคือ GM (1,1) (first order grey model for one variable) ซึ่งจะเริ่มต้นจากการแปลงข้อมูลจริงที่ศึกษาเป็นข้อมูลสะสม แล้วใช้สมการถดถอยเชิงเส้นตรงประมาณค่าคงที่และค่าความชัน จากนั้นทำนายค่าสะสมแล้วหักออกด้วยค่าสะสม ที่ทำนายก่อนหน้าก็จะได้ค่าทำนายของแต่ละคาบเวลา ซึ่งกล่าวได้ว่าเหมาะสมกับข้อมูลอนุกรมเวลาที่มีจำนวน ค่าสังเกตเพิ่มสูงขึ้นทุกปี นอกจากนี้ยังมีแบบจำลองที่มีการแก้ไขข้อผิดพลาดเป็นระยะเพื่อให้เหมาะกับข้อมูลที่มีรูปแบบเพิ่มขึ้นและลดลงเป็นโค้งรูปตัวเอสด้วยการพิจารณาว่าแบบจำลองการทำนายสามารถทำนายได้แม่นยำเพียงใดก็สามารถพิจารณาได้จากค่าทำนายที่เบี่ยงเบนไปจากค่าจริง หรือ ค่าสหสัมพันธ์ของค่าจริงกับค่าทำนาย^(16,17)

3.5 การพยากรณ์ด้วยวิธีแยกตัวประกอบและการคำนวณดัชนีฤดูกาล

วิธีการแยกตัวประกอบ ประกอบด้วย แบบจำลองการคูณ (Multiplicative model) และแบบจำลองการบวก (Additive model) แบบจำลองการคูณ (Multiplicative model) $Y=T \times C \times S \times I$ เป็นรูปแบบที่ใช้กันอย่างแพร่หลายซึ่งถือว่าการคาดการณ์ (Y) เป็นผลงานของทั้งสี่ส่วนประกอบในช่วงเวลาหนึ่ง ๆ ผลขององค์ประกอบสี่ประการ ต่ออนุกรมเวลามีการพึ่งพาอาศัยกัน (interdependent)

แบบจำลองการบวก (Additive model) $Y=T + C + S + I$ ถือว่าองค์ประกอบทั้งสี่นี้เป็นอิสระจากกัน ซึ่งเป็นสมมติฐานที่ไม่สอดคล้องกับความเป็นจริง

การคำนวณดัชนีฤดูกาล 3 วิธี คือ วิธีค่าเฉลี่ยอย่างง่าย (Simple Moving Average) วิธีการนี้ใช้เมื่อตัวแปรอนุกรมเวลาประกอบด้วยเฉพาะส่วนประกอบตามฤดูกาลและแบบสุ่ม วิธีอัตราส่วนต่อการเฉลี่ยเคลื่อนที่ (Ratio to moving average, RMA) เป็นวิธีที่ใช้กันมากที่สุด การวัดความแปรผันตามฤดูกาล วิธีการนี้จะถือว่าการมีอยู่ของทั้งหมดสี่องค์ประกอบของอนุกรมเวลาเป็นการแยกองค์ประกอบ T, C และ I ออกจากข้อมูลอนุกรมเวลาให้เหลือ แต่ S ที่ต้องการ วิธีการอัตราส่วนต่อแนวโน้ม (Ratio to trend) วิธีการนี้ใช้เมื่อข้อมูลไม่มีความแปรปรวนของวัฏจักร^(18,19) การพยากรณ์ด้วยเทคนิคแยกตัวประกอบแบบจำลองการคูณ ดัชนีอนุกรมเวลาตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนที่⁽²⁰⁾

3.6 เกณฑ์ตัดสินความแม่นยำของแบบจำลอง

ค่าข้อมูลจริง y_i จำนวน n ค่า โดย $i = 1, 2, \dots, n$ ส่วน \hat{y}_i เป็นค่าพยากรณ์ สำหรับคาบเวลาตามข้อมูลที่นำมาใช้พยากรณ์ ส่วนค่าพยากรณ์ไปในอนาคต $i = n + 1, n + 2, \dots$

ค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (The Mean Absolute Percentage Error - MAPE)⁽²¹⁾

$$MAPE = \left(\frac{1}{n} \sum_{k=1}^n \left| \frac{y_k - \hat{y}_k}{y_k} \right| \right) \times 100\%$$

MAPE หน่วยเป็นร้อยละ ถ้าน้อยกว่า 10 มีความแม่นยำสูง ระหว่าง 10-20 ใช้พยากรณ์ได้ดี ระหว่าง 20-50 มีเหตุผลพอที่จะใช้พยากรณ์ และมากกว่า 50 ไม่มีความแม่นยำ⁽²²⁾

4. ผลการศึกษา

ตารางที่ 2 แบบจำลองตามทฤษฎีระบบเกรย์

แบบจำลอง	MAPE	ค่าพยากรณ์ 2566	การเปลี่ยนแปลงเพิ่ม/-ลดจากปี 2565 (ร้อยละ)
GM (1,1)	10.16	5,866	-4.12
GM (1,1) EPC	4.47	6,214	1.58
GVM	59.69	3,359	-45.09
GVIM	51.57	3,756	-38.61

4.1 ผลการคำนวณตามทฤษฎีระบบเกรย์

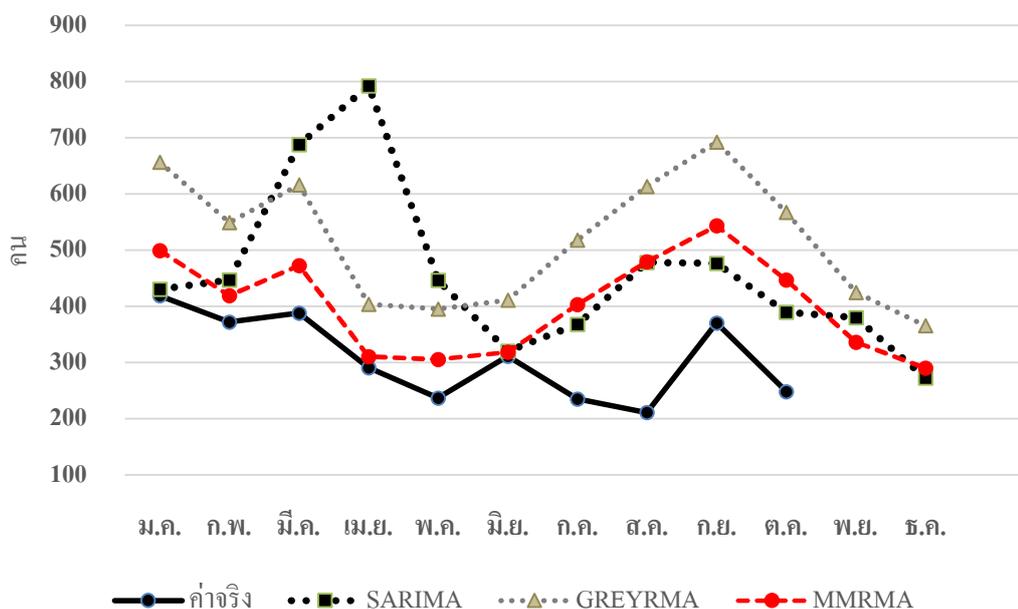
การคำนวณตามทฤษฎีระบบเกรย์ ด้วยข้อมูลจำนวนผู้ป่วยรายปี 20 ปี พ.ศ. 2546 ถึง 2565 พยากรณ์จำนวนผู้ป่วยในปี พ.ศ. 2566 พบว่า แบบจำลอง GM (1,1)EPC มีค่า MAPE ต่ำที่สุด ดังแสดงในตารางที่ 2 จึงเลือกมาใช้ในการจำแนกรายเดือนด้วยดัชนีฤดูกาลวิธีอัตราส่วนต่อค่าเฉลี่ย ซึ่งจะเรียกโดยย่อว่า GREYRMA ดังแสดงในสคมภ์ GREYRMA ตารางที่ 3

ตารางที่ 3 ค่าพยากรณ์จำนวนผู้ป่วยรายเดือน ปี พ.ศ. 2566

เดือน	ค่าจริง		ค่าพยากรณ์		ดัชนีฤดูกาล	ค่าพยากรณ์		
	2566	95% Lower	SARIMA	95% Upper		RMA	GREYRMA	MMRMA
MAPE1			21.11			4.47	31.30	
ม.ค.	419	238	431	623	1.27	657	499	
ก.พ.	372	191	447	703	1.06	549	419	
มี.ค.	388	422	688	954	1.19	617	472	
เม.ย.	291	526	792	1,058	0.78	404	311	
พ.ค.	237	179	447	714	0.76	395	305	
มิ.ย.	311	52	320	588	0.79	411	319	
ก.ค.	235	100	368	636	1.00	518	403	
ส.ค.	211	210	478	746	1.18	614	480	

ตารางที่ 3 ค่าพยากรณ์จำนวนผู้ป่วยรายเดือน ปี พ.ศ. 2566

เดือน	ค่าจริง	ค่าพยากรณ์			ดัชนีฤดูกาล	ค่าพยากรณ์	
	2566	95% Lower	SARIM A	95% Upper		RMA	GREYRM A
MAPE1			21.11			4.47	31.30
ก.ย.	370	208	477	745	1.34	693	543
ต.ค.	248	121	390	658	1.10	567	447
พ.ย.	-	111	380	649	0.82	425	336
ธ.ค.	-	4	273	541	0.71	365	290
รวม		2,364	5,490	8,616	12.00	6,214	4,824
MAPE2			63.32			82.81	41.74



ภาพที่ 5 พยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบรายเดือน พ.ศ. 2566

4.2 การคำนวณตามวิธีบอกซ์และเจนกินส์

การพยากรณ์ SARIMA โปรแกรม R แพคเกจ Forecast ฟังก์ชัน auto.arima ใช้ข้อมูลจำนวนผู้ป่วยรายเดือน 20 ปี พ.ศ. 2546 ถึง 2565 เป็นเวลา 240 เดือน ได้แบบจำลอง ARIMA (2,1,1) (0,1,1)₁₂ ซึ่งจะเรียกโดยย่อว่า SARIMA มีค่า MAPE1 ร้อยละ 21.11 (ระหว่าง 20-50 มีเหตุผลพอที่จะใช้พยากรณ์) ดังแสดงในสคมภ์ 95% Lower, SARIMA, และ 95% Upper ตารางที่ 3

4.3 การคำนวณวิธีแยกตัวประกอบและดัชนีฤดูกาล

การพยากรณ์ด้วยเทคนิคแยกตัวประกอบ⁽²⁰⁾ ได้ดัชนีอนุกรมเวลาตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนในที่นี่จะเรียกโดยย่อว่า MMRMA ดังแสดงในสคมภ์ MMRMA ตารางที่ 3

4.4 ความแม่นยำของการพยากรณ์

MAPE1 ซึ่งเป็นค่าเฉลี่ยร้อยละของความคลาดเคลื่อนของค่าพยากรณ์รายเดือนจากค่าสถิติในอดีต 240 เดือน พบว่า SARIMA และ MMRMA เท่ากับร้อยละ 21.11 และ 31.30 ตามลำดับ ส่วน GM (1,1) EPC มีค่า MAPE 20 ปี ร้อยละ 4.47

การเปรียบเทียบกับค่าสถิติในปี พ.ศ. 2566 ซึ่งมีข้อมูลจำนวน 43 สัปดาห์ มีผู้ป่วย 3,082 คน ประมาณการ ปี พ.ศ. 2566 อย่างหยาบ เท่ากับ 3,727 คน (3082*52/43) พบความคลาดเคลื่อนจากประมาณอย่างหยาบโดยวิธี SARIMA, GREYRMA และ MMRMA ร้อยละ 47.30, 66.74, และ 29.43 ตามลำดับ ดังแสดงในตารางที่ 3

การพยากรณ์ผู้ป่วยรายเดือน ปี พ.ศ. 2566

เมื่อนำค่าพยากรณ์รายเดือน ปี พ.ศ. 2566 มาทำการเปรียบเทียบกับค่าจริงรายเดือน มกราคม ถึง ตุลาคม ในปี พ.ศ. 2566 จำนวน 10 เดือน พบว่ามีค่า MAPE2 ของวิธี SARIMA, GREYRMA และ MMRMA ร้อยละ 63.32, 82.81, และ, 41.74 ตามลำดับ ดังแสดงในภาพที่ 5

5. อภิปรายผล

การพิจารณาค่าจริงเทียบกับค่าพยากรณ์ในอดีต 240 เดือน แบบจำลอง SARIMA มี MAPE ร้อยละ 21.11 ต่ำกว่า MMRMA ที่มี MAPE ร้อยละ 31.30 แต่เมื่อพิจารณาค่า MAPE แบบจำลอง SARIMA, GREYRMA และ MMRMA ร้อยละ 47.30, 66.74, และ 29.43 ตามลำดับ จากภาพที่ 5 การพยากรณ์ด้วยเทคนิคแยกตัวประกอบแบบจำลองการคูณ ดัชนีอนุกรมเวลาตามวิธีสัดส่วนต่อค่าเฉลี่ยเคลื่อนที่ หรือแบบจำลอง MMRMA มีความใกล้เคียงกับข้อมูลจริงในปี พ.ศ. 2566 มากกว่า จึงเป็นค่าพยากรณ์ที่น่าจะยอมรับได้สำหรับเดือน พฤศจิกายน และธันวาคมที่เหลืออยู่

การพิจารณาเทียบกับข้อมูลในอดีต 240 เดือนเพียงอย่างเดียว ก็จะได้ผลเช่นเดียวกับการพยากรณ์ที่ใช้ในการศึกษา 3 วิธี คือ วิธีการแยกส่วนประกอบ วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลของวินเตอร์เชิงผล

คุณ และวิธีบอกซ์และเจนกินส์โดยใช้ข้อมูลผู้ป่วยโรคปอดอักเสบในประเทศไทยรายไตรมาส ตั้งแต่ พ.ศ.2551 ถึง พ.ศ.2561 เปรียบเทียบตัวแบบพยากรณ์ที่เหมาะสมภายใต้เกณฑ์ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error, MAPE) ต่ำที่สุด ผลการศึกษาพบว่าวิธีบอกซ์-เจนกินส์ให้ค่า MAPE ต่ำที่สุดตัวแบบที่เหมาะสมในการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ คือ ตัวแบบรวมการถดถอยในตัวกับการเคลื่อนที่แบบมีฤดูกาล ARIMA (0,1,1) (0,1,1)₄⁽⁶⁾ เนื่องจากวิธีบอกซ์-เจนกินส์มีเปรียบเทียบกับแบบจำลองจำนวนมากจนได้ค่า MAPE ต่ำที่สุดแล้วจึงมาเปรียบเทียบกับวิธีอื่น

6. ข้อเสนอแนะ

การพยากรณ์ด้วยอนุกรมเวลาเป็นการพยากรณ์เชิงปริมาณที่ใช้ข้อมูลในอดีต แบบจำลองที่ได้ก็เทียบความแม่นยำกับข้อมูลในอดีตเท่านั้น การคาดการณ์อนาคตยังคงต้องใช้สิ่งที่ทราบล่วงหน้าว่าจะเกิดขึ้น ซึ่งเป็นปัจจัยที่มีผลต่อจำนวนผู้ป่วย เช่นการเปลี่ยนแปลงของภาวะอากาศ คุณภาพอากาศ การรณรงค์ลดมลภาวะ การปฏิบัติตัวของประชาชนในพื้นที่ ซึ่งสามารถนำมาปรับค่าพยากรณ์จากความเข้าใจการเกิดโรคของประชาชนในพื้นที่ก็จะทำให้การพยากรณ์มีความแม่นยำมากขึ้น

7. เอกสารอ้างอิง

1. กรมควบคุมโรค. โรคปอดบวม, ปอดอักเสบ (Pneumonia) [อินเทอร์เน็ต]. 2562 [เข้าถึงเมื่อ 16 พ.ย. 2566] เข้าถึงได้จาก: https://ddc.moph.go.th/disease_detail.php?d=21
2. กรมควบคุมโรค. คาดการณ์โรค ปี 2565 8 โรคสำคัญ [อินเทอร์เน็ต]. 2564 [เข้าถึงเมื่อ 16 พ.ย. 2566]. เข้าถึงได้จาก: <https://www.ddc.moph.go.th/uploads/publish/1305720220831091702.pdf>
3. สำนักระบาดวิทยา กรมควบคุมโรค. ระบบรายงานการเฝ้าระวังโรค 506. Pneumonia. สรุปสถานการณ์ ปี 65 รายสัปดาห์ [52] [อินเทอร์เน็ต]. 2566 [เข้าถึงเมื่อ 16 พ.ย. 2566]. เข้าถึงได้จาก: <http://doe.moph.go.th/surdata/disease.php?ds=31>
4. กลุ่มงานพัฒนายุทธศาสตร์สาธารณสุข งานข้อมูลข่าวสาร สำนักงานสาธารณสุขจังหวัดพิจิตร. สัปดาห์ระบาดวิทยา ปี 2566 [อินเทอร์เน็ต]. 2566 [เข้าถึงเมื่อ 16 พ.ย. 2566]. เข้าถึงได้จาก: <http://www.ppho.go.th/mis-new/index.php?menu=5.19>
5. สำนักระบาดวิทยา กรมควบคุมโรค. ระบบรายงานการเฝ้าระวังโรค 506. Pneumonia. สรุปสถานการณ์ ปี 66 รายสัปดาห์[43] [อินเทอร์เน็ต].2566 [เข้าถึงเมื่อ 16พ.ย.2566].เข้าถึงได้จาก: <http://doe.moph.go.th/surdata/disease.php?ds=31>
6. วราพร ต้วทอง, สวพร หิญาธิระนันท์. การเปรียบเทียบตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบในประเทศไทย. วารสารวิทยาศาสตร์และเทคโนโลยี (ววท.) 2564; 29 (3): 365-377.

7. กมล กัญญาประสิทธิ์. การพยากรณ์การเกิดโรคไข้เลือดออก ในพื้นที่ จังหวัดเพชรบูรณ์ พ.ศ. 2558. วารสารวิชาการสาธารณสุข 2559; 25 (4): 604-614.
8. จารุณี กันธ. ปัจจัยทำนายพฤติกรรมส่งเสริมสุขภาพของผู้สูงอายุตำบลนางัว อำเภอเมือง จังหวัดเพชรบูรณ์. วารสารวิชาการป้องกันควบคุมโรค สคร.2 พิษณุโลก 2565; 9 (3):19-35.
9. สำนักระบาดวิทยา กรมควบคุมโรค. ระบบรายงานการเฝ้าระวังโรค 506 > Pneumonia > ข้อมูลย้อนหลัง > สรุปสถานการณ์รายปี 2546-2565 [อินเทอร์เน็ต]. 2566 [เข้าถึงเมื่อ 16 พ.ย. 2566]. เข้าถึงได้จาก: <http://doe.moph.go.th/surdata/disease.php?dcontent=old&ds=31>
10. Santa R. Introduction to SARIMA Model [Internet]. 2023 [cited 2023 Nov 1]. Available from: <https://medium.com/@ritusantra/introduction-to-sarima-model-cbb8>
11. The R Foundation. The R Project for Statistical Computing [Internet]. 2023 [cited 2023 Nov 1]. Available from: <https://www.r-project.org/>
12. Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeeen F. forecast: Forecasting functions for time series and linear models. R package version 8.21.1 [Internet]. 2023 [cited 2023 Nov 1]. Available from: <https://pkg.robjhyndman.com/forecast/>.
13. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software 2008; 26 (3), 1–22. doi:10.18637/jss.v027.i03.
14. Wischke AF. Forecasting with R: Auto-ARIMA Step by Step [Internet]. 2018 [cited 2023 Nov 1]. Available from: <https://www.linkedin.com/pulse/forecasting-r-auto-arima-step-anton-figuerola-wischke>
15. Bevans R. March 26, 2020. Akaike Information Criterion | When & How to Use It (Example) [Internet]. Scribbr; 2023 [cited 2023 Nov 1]. Available from: <https://www.scribbr.com/statistics/akaike-information-criterion/>
16. Liu S., Lin Y. Grey systems theory and application. Verlag: Springer; 2010. pp. 106-130.
17. Zhang, Y. (2012). Improved grey derivative of grey Verhulst model and its application. International Journal of Computer Science 2012; 9 (6), 443-448.

18. Strathmore University Business School. ANALYSIS OF TIME SERIES [Internet]. 2021 [cited 2023 Nov 1]. Available from:
https://sbslearning.strathmore.edu/pluginfile.php/177213/mod_resource/content/1/Topic-4-Analysis-of-Time-Series.pdf
19. Shukla, GK., Trivedi, M. eGyanKosh > Block 4: Time Series Modelling > UNIT 14 SEASONAL COMPONENT ANALYSIS [Internet]. 2017 [cited 2023 Nov 1]. Available from:
<https://www.egyankosh.ac.in/bitstream/123456789/20805/1/Unit-14.pdf>
20. Chan B. Classical Multiplicative Decomposition – YouTube [Internet]. 2015 [cited 2023 Nov 1]. Available from: <https://www.youtube.com/watch?v=S31b-vh0Ce0>
21. Andrés D. Machine Learning Pills: Error Metrics for Time Series Forecasting [Internet]. 2023 [cited 2023 Oct 20]. Available from: <https://mlpills.dev/time-series/error-metrics-for-time-series-forecasting/>
22. Lewis CD. Industrial and business forecasting methods. London: Butterworths;1982.