



ชื่อโครงการ การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย

Forecasting of Rice Product in Thailand

ชื่อนิสิต นางสาวทัศนีย์ เร่งเทียน 5933518523

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์

สาขาวิชาคณิตศาสตร์

ปีการศึกษา 2562

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย

นางสาวทัศนีย์ เร่งเทียน

โครงการนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2562

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Forecasting of Rice Product in Thailand

Miss Tussanee Rangthien

A Project Submitted in Partial Fulfillment of the Requirements
for the Degree of Bachelor of Science Program in Mathematics

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2019

Copyright of Chulalongkorn University

หัวข้อโครงการ	การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย
โดย	นางสาวทัศนีย์ เร่งเทียน
สาขาวิชา	คณิตศาสตร์
อาจารย์ที่ปรึกษาโครงการหลัก	รองศาสตราจารย์ ทิพวัลย์ สันติวิภาณนท์
อาจารย์ที่ปรึกษาโครงการร่วม	รองศาสตราจารย์ ดร.ทรงเกียรติ สุเมธกิจการ

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับโครงการฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาบัณฑิตในรายวิชา 2301499 โครงการวิทยาศาสตร์ (Senior Project)

----- หัวหน้าภาควิชาคณิตศาสตร์

(ศาสตราจารย์ ดร.กฤษณะ นิยมมณี) และวิทยาการคอมพิวเตอร์

คณะกรรมการสอบโครงการ

ทิพวัลย์ สันติวิภาณนท์

----- อาจารย์ที่ปรึกษาโครงการหลัก

(รองศาสตราจารย์ ทิพวัลย์ สันติวิภาณนท์)

ดร.ทรงเกียรติ สุเมธกิจการ

----- อาจารย์ที่ปรึกษาโครงการร่วม

(รองศาสตราจารย์ ดร.ทรงเกียรติ สุเมธกิจการ)

ดร. ชีระเดช กิตติภัตสร

----- กรรมการ

(ผู้ช่วยศาสตราจารย์ วาสนา สุขกระสานติ)

ชีระเดช กิตติภัตสร

----- กรรมการ

(อาจารย์ ดร. ชีระเดช กิตติภัตสร)

นางสาวทัศนีย์ เร่งเทียน: การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย.

(Forecasting of Rice Product in Thailand)

อ.ที่ปรึกษาโครงการหลัก : รองศาสตราจารย์ ทิพวัลย์ สันติวิภาณนท์

อ.ที่ปรึกษาโครงการร่วม : รองศาสตราจารย์ ดร.ทรงเกียรติ สุเมธกิจกร, 67 หน้า.

เราศึกษาและใช้ตัวแบบ ARMA โดยวิธีบ็อกซ์-เจนกินส์ ในการพยากรณ์ปริมาณผลผลิตข้าวของประเทศไทย จากข้อมูลปริมาณผลผลิตข้าวรายเดือนตั้งแต่เดือนมกราคมปี พ.ศ. 2546 ถึงเดือนธันวาคมปี พ.ศ. 2561 ที่รวบรวมโดยสำนักงานเศรษฐกิจการเกษตร โดยแบ่งพิจารณาเป็นข้าวนาปี (แต่ละฤดูมีผลผลิตในเดือนสิงหาคมของปีหนึ่งถึงเดือนเมษายนของปีถัดไป) และข้าวนาปรัง (แต่ละฤดูมีผลผลิตในเดือนกุมภาพันธ์ถึงเดือนตุลาคมในปีเดียวกัน) ในการใช้คำสั่ง arima ในโปรแกรม R คำนวณหาตัวแบบที่เหมาะสมจากข้อมูลปริมาณผลผลิตข้าวรายเดือนเป็นเวลา 11 ปี ปีละ 9 เดือน เพื่อพยากรณ์ปริมาณผลผลิตในปีถัดไป เราได้ตัวแบบ ARMA สำหรับพยากรณ์ปริมาณผลผลิตในปี พ.ศ. 2557, 2558, 2559, 2560 และ 2561 ซึ่งมีค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยจากข้อมูลจริงของปีนั้น ๆ ที่ค่อนข้างสูง ผลการศึกษาที่ได้แสดงให้เห็นว่าการพยากรณ์โดยวิธีบ็อกซ์-เจนกินส์ ในงานวิจัยนี้ไม่แม่นยำ เนื่องจากผลการพยากรณ์ให้ค่าแตกต่างจากข้อมูลจริงมาก

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์..... ลายมือชื่อนิสิต ทัศนีย์ เร่งเทียน

สาขาวิชาคณิตศาสตร์..... ลายมือชื่อ อ.ที่ปรึกษาโครงการหลัก ทิพวัลย์ สันติวิภาณนท์

ปีการศึกษา2562..... ลายมือชื่อ อ.ที่ปรึกษาโครงการร่วม ทรงเกียรติ สุเมธกิจกร

5933518523: MAJOR MATHEMATICS

KEYWORDS: ARMA, BOX-JENKINS

TUSSANEE RANGTHIEN: FORECASTING OF RICE PRODUCT IN THAILAND. ADVISOR: ASSOC. PROF. TIPPAWAN SANTIWIPANONT, CO-ADVISOR: ASSOC. PROF. SONGKIAT SUMETKIJAKAN, Ph.D., 67 pp.

I study and utilize the ARMA (Autoregressive Moving Average) models by Box-Jenkins method in forecasting monthly rice product in Thailand based on rice product quantitative data from January 2003 to December 2018 collected by the Office of Agricultural Economics. We separately consider in-seasonal rice (each season starts from August to April of the next year) and off-seasonal rice (each season starts from February to October). Using the R-command arima to find appropriate models of five 11-year once 9 months moving data sets in order to forecast the following 12th year, we obtain ARMA models to forecast rice product quantities of the corresponding months of the years 2014, 2015, 2016, 2017 and 2018. The root mean square errors of the forecast from the actual data are relatively high. The result shows that the forecasting by Box-Jenkins method in this project is inaccurate since the forecasting results are very different from the actual data.

Department: Mathematics and Computer Science Student's Signature Tussanee Rangthien
 Field of Study: Mathematics Advisor's Signature Tippawan S.
 Academic Year: 2019 Co-advisor's Signature SMS.

กิตติกรรมประกาศ

การศึกษาในหัวข้อเรื่อง “การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย” เล่มนี้สำเร็จลุล่วงได้ด้วยดี ด้วยความเมตตาและการสนับสนุนอย่างเต็มที่จากรองศาสตราจารย์ ทิพวัลย์ สันติวิธานนท์ และรองศาสตราจารย์ ดร.ทรงเกียรติ สุขเมธกิจกร อาจารย์ที่ปรึกษาโครงการ ที่เสียสละเวลาอันมีค่าให้คำแนะนำ คอยให้กำลังใจ และช่วยหาแนวทางแก้ไขปัญหาระหว่างการทำโครงการ ตลอดจนให้แนวคิด สอนให้องค์ความรู้แก่ผู้วิจัยด้วยดีเสมอมา ข้าพเจ้าขอกราบขอบพระคุณอาจารย์เป็นอย่างสูงมา ณ โอกาสนี้

ข้าพเจ้าขอขอบพระคุณผู้ช่วยศาสตราจารย์ วาสนา สุขกระสานดี และอาจารย์ ดร. ชีระเดช กิตติภัสสร กรรมการสอบที่ได้สละเวลาพิจารณาและรับฟังเพื่อให้คำแนะนำในการปรับปรุงให้โครงการเล่มนี้มีความสมบูรณ์มากขึ้น

สุดท้ายนี้ผู้วิจัยขอขอบคุณครอบครัว เพื่อน ๆ อาจารย์ที่เคารพรัก และผู้เกี่ยวข้องทุกท่านที่มิได้กล่าวไว้ในที่นี้ที่ให้ความช่วยเหลือและเป็นกำลังใจที่สำคัญเสมอมา ยิ่งไปกว่านั้นขอขอบคุณทางภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ได้ให้การสนับสนุนในด้านต่าง ๆ รวมถึงงบประมาณในการจัดทำโครงการ หวังว่าผลการศึกษาในโครงการนี้จะเป็นประโยชน์ในการศึกษายานที่เกี่ยวข้องต่อไป

สารบัญ

หน้า

บทคัดย่อภาษาไทย	จ
บทคัดย่อภาษาอังกฤษ	ฉ
กิตติกรรมประกาศ.....	ช
สารบัญ	ฅ
สารบัญรูปภาพ	ฉ
สารบัญตาราง	ฐ
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของการศึกษา.....	1
1.2 ข้อมูลทั่วไป.....	1
1.3 วัตถุประสงค์ของการศึกษา.....	2
1.4 ขอบเขตการศึกษา.....	2
1.5 วิธีการดำเนินงาน	3
1.6 ผลที่ได้รับ	3
1.7 ประโยชน์ที่ได้รับจากการศึกษา.....	3
บทที่ 2 งานวิจัยและทฤษฎีที่เกี่ยวข้อง.....	4
2.1 งานวิจัยที่เกี่ยวข้อง	4
2.2 ทฤษฎีที่เกี่ยวข้อง.....	5
2.2.1 ความหมายของการพยากรณ์	5
2.2.2 ขั้นตอนในการพยากรณ์.....	5
2.2.3 ประเภทของเทคนิคการพยากรณ์ทางสถิติ	6
2.2.3.1 การพยากรณ์เชิงคุณภาพ (Qualitative Forecasting Techniques)	6
2.2.3.2 การพยากรณ์เชิงปริมาณ (Quantitative Forecasting Techniques)	6

2.2.4 การวิเคราะห์อนุกรมเวลา	6
2.2.4.1 การวิเคราะห์แนวโน้ม (Trend: T).....	6
2.2.4.2 การวิเคราะห์ความแปรผันเกี่ยวกับฤดูกาล (Seasonal: S)	6
2.2.5 วิธีการบ็อกซ์-เจนกินส์ (Box-Jenkins method)	7
2.2.5.1 ตัวแบบ ARMA (Autoregressive Moving Average)	8
2.2.5.2 วิธีการพยากรณ์อนุกรมเวลาโดยวิธีบ็อกซ์-เจนกินส์.....	10
บทที่ 3 การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทยแบบรายเดือน	16
3.1 แนวทางการพยากรณ์.....	16
3.2 การพยากรณ์ปริมาณผลผลิตข้าวรายปีแบบรายเดือนในประเทศไทย.....	16
3.3 การพยากรณ์ปริมาณผลผลิตข้าวรายปีแบบรายเดือนในประเทศไทย.....	37
บทที่ 4 บทสรุปและข้อเสนอแนะ	58
4.1 บทสรุป.....	58
4.2 ข้อเสนอแนะ	59
เอกสารอ้างอิง	60
ภาคผนวก ก แบบเสนอหัวข้อโครงการ รายวิชา 2301399 Project Proposal ปีการศึกษา 2562	62
ภาคผนวก ข คำสั่งโปรแกรม R ที่ใช้ในการวิเคราะห์	66
ประวัติผู้เขียน	67

รูปที่ 3.3.11 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2558 ถึงเดือนตุลาคมปี พ.ศ. 2558	45
รูปที่ 3.3.12 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF).....	46
รูปที่ 3.3.15 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง	49
รูปที่ 3.3.16 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2559 ถึงเดือนตุลาคมปี พ.ศ. 2559	49
รูปที่ 3.3.17 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF).....	50
รูปที่ 3.3.20 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง	53
รูปที่ 3.3.21 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2560 ถึงเดือนตุลาคมปี พ.ศ. 2560	53
รูปที่ 3.3.22 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF).....	54
รูปที่ 3.3.25 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง	57
รูปที่ 3.3.26 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2561 ถึงเดือนตุลาคมปี พ.ศ. 2561	57

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของการศึกษา

ข้าวเป็นธัญพืชที่สำคัญชนิดหนึ่งของประชากรโลก และเป็นพืชเศรษฐกิจที่สำคัญของประเทศไทย ประเทศไทยมีการส่งออกข้าวมากเป็นอันดับ 2 ของโลก รองจากประเทศอินเดีย นอกจากข้าวที่เป็นพืชเศรษฐกิจและสินค้าส่งออกที่สำคัญของประเทศไทยแล้ว ประเทศไทยยังมีพืชเศรษฐกิจที่สำคัญอีก เช่น ยางพารา อ้อย มันสำปะหลัง และปาล์ม เป็นต้น โดยในปี 2562 ปริมาณการส่งออกพืชเศรษฐกิจที่สำคัญมากที่สุดคือข้าว รองลงมาคือยางพารา

ในส่วนของการผลิตข้าว ประเทศไทยมีพื้นที่ทำนาจำแนกเป็นพื้นที่นาในเขตอาศัยน้ำฝนร้อยละ 75 และเป็นพื้นที่นาในเขตชลประทานร้อยละ 25 ซึ่งในปี 2562 มีเนื้อที่เพาะปลูกทั้งข้าวนาปีและข้าวนาปรังลดลง เนื่องจากมีเกษตรกรบางส่วนสนใจปรับเปลี่ยนไปปลูกพืชชนิดอื่นตามนโยบายของรัฐบาล เช่น อ้อย มันสำปะหลัง และข้าวโพด เป็นต้น อย่างไรก็ตามในภาพรวมทั้งประเทศมีปริมาณผลผลิตต่อไร่เพิ่มขึ้น เนื่องจากพื้นที่ส่วนใหญ่มีสภาพอากาศเอื้ออำนวยต่อการเพาะปลูกข้าว ปริมาณน้ำอยู่ในเกณฑ์ดีเพียงพอต่อการเพาะปลูก

ปัจจุบันข้าวยังเป็นที่ต้องการของตลาดทั้งภายในประเทศและต่างประเทศ แต่ประเทศไทยประสบปัญหาการค้าข้าวตกต่ำ เนื่องจากปริมาณการสั่งซื้อที่ลดลงของผู้บริโภคในต่างประเทศ เนื่องจากในช่วง 10 ปีที่ผ่านมา สาธารณรัฐสังคมนิยมเวียดนามได้มีการผลิตข้าวเพิ่มมากขึ้นและมีการปรับปรุงคุณภาพของข้าวจนใกล้เคียงกับของประเทศไทยแต่มีราคาที่ถูกกว่า ส่งผลให้ปริมาณผลผลิตข้าวในประเทศมีจำนวนมากกว่าความต้องการของผู้บริโภคในประเทศ แม้ว่ารัฐบาลจะออกนโยบายแก้ไขต่าง ๆ แต่ก็ไม่สามารถช่วยแก้ปัญหาในระยะยาวได้ ซึ่งปัจจัยที่มีผลต่อราคาข้าวตกต่ำ ได้แก่ ปริมาณผลผลิตข้าวในประเทศคุณภาพต่ำ และปัญหาจากสภาพแวดล้อม ด้วยเหตุนี้ผู้วิจัยจึงมีความสนใจที่จะทำการพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย เพื่อคาดการณ์สถานการณ์ของข้าวในอนาคตพร้อมกับเตรียมการรับมือสำหรับทุกฝ่ายทั้งภาครัฐ เอกชน และประชาชน

1.2 ข้อมูลทั่วไป

ข้าวเป็นธัญญาหารหลักของชาวโลก จัดเป็นพืชสายพันธุ์เดียวกับหญ้า ข้าว นับได้ว่าเป็นหญ้าที่มีขนาดใหญ่ที่สุดในโลก และมีความหลากหลายทางชีวภาพสามารถปลูกขึ้นได้ง่ายมีความทนทานต่อทุกสภาพภูมิประเทศในโลก ไม่ว่าจะเป็นถิ่นแห้งแล้งแบบทะเลทราย พื้นที่ราบลุ่มน้ำท่วมถึง หรือแม้กระทั่งบนเทือกเขาที่หนาวเย็น ข้าวก็ยังสามารถงอกงามขึ้นมาได้

ข้าวของไทยเป็นพืชอาหารประจำชาติที่มีประวัติศาสตร์มาอย่างยาวนาน ปรากฏเป็นร่องรอยพร้อม กับอารยธรรมไทยมาไม่น้อยกว่า 5,500 ปี ซึ่งมีหลักฐานจากเมล็ดข้าวที่เป็นส่วนผสมของดินที่ใช้ทำ เครื่องปั้นดินเผาที่บ้านเชียง

ปัจจุบันการปลูกข้าวในประเทศไทยในภาคเหนือตอนบนนิยมปลูกข้าวเหนียว แต่ในภาคเหนือ ตอนล่าง ภาคกลางและภาคใต้เน้นปลูกข้าวเจ้าเป็นหลัก ภาคตะวันออกเฉียงเหนือมีพื้นที่ปลูกข้าวคิดเป็น 45 % ของพื้นที่เพาะปลูกทั้งประเทศ ส่วนใหญ่ปลูกข้าวหอมมะลิ 105 ซึ่งเป็นข้าวคุณภาพดีที่สุดของโลก ข้าวที่ปลูกในพื้นที่แถบนี้จึงมักปลูกไว้เพื่อขาย รองลงมาคือภาคกลาง และภาคเหนือที่มีพื้นที่เพาะปลูก เท่ากันประมาณ 25%

ทุกวันนี้ไทยเป็นแหล่งปลูกข้าวที่ผลิออกสู่ตลาดโลกมากที่สุด และเป็นศูนย์กลางของการ ศึกษาวิจัยพันธุ์ข้าว โดยข้าวสามารถแบ่งตามฤดูเพาะปลูก ได้แก่

1. ข้าวนาปีหรือข้าวหน้าน้ำฝน คือ การทำนาข้าวที่ทำในฤดูการทำนาปรกติ พันธุ์ข้าวนาปีจะออกดอกตาม วันและเดือนที่ค่อนข้างตายตัว เมื่อถึงวันที่จะออกดอกก็ออกพร้อมกันหมด เนื่องจากช่วงของแสงต่อ วันบังคับให้ออกดอก ข้าวประเภทนี้จึงเรียกว่า “ข้าวนาปี” หรือ “ข้าวไวแสง” ซึ่งเป็นข้าวที่ออกตาม ฤดูกาล โดยในแต่ละฤดูจะเริ่มทำการเก็บเกี่ยวตั้งแต่เดือนสิงหาคมของปีหนึ่งถึงเดือนเมษายนของปี ถัดไป
2. ข้าวนาปรัง คือ การทำนาข้าวที่ปลูกนอกฤดูการทำนาปรกติ เพราะเป็นข้าวที่ไม่ต้องใช้น้ำมากซึ่งในฤดู การทำนาปรกติจะมีน้ำมากเกินไป และข้าวที่ใช้ทำนาปรังจะเป็นข้าวที่แสงไม่มีอิทธิพลต่อการออก ดอก ซึ่งเรียกว่า “ข้าวนาปรัง” หรือ “ข้าวไม่ไวแสง” ดังนั้นข้าวนาปรังจะออกตามอายุ ไม่ว่าจะปลูก เมื่อใด พืชอายุการเก็บเกี่ยวก็สามารถเก็บเกี่ยวได้ โดยในแต่ละฤดูจะเริ่มทำการเก็บเกี่ยวตั้งแต่เดือน กุมภาพันธ์ถึงเดือนตุลาคมในปีเดียวกัน

เนื่องจากตัวแบบ ARMA ของวิธีบ็อกซ์-เจนกินส์มีความเหมาะสมในการวิเคราะห์ข้อมูลที่มี ลักษณะเป็นอนุกรมเวลา (ข้อมูลที่เก็บรวบรวมตามลำดับเวลาที่เกิดขึ้นอย่างต่อเนื่อง) โดยใช้ข้อมูลในอดีต เพื่ออธิบายลักษณะของอนุกรมเวลานั้นในอนาคต สำหรับจุดเด่นของตัวแบบ ARMA คือ มีการรวบรวม แล้วให้นำน้ำหนักความสำคัญกับข้อมูลทั้งหมดที่ใช้ในการพยากรณ์ ไม่ว่าจะเป็ข้อมูลในอดีตและ/ค่า ความคลาดเคลื่อนจากการพยากรณ์ ดังนั้น โครงานนี้จึงเลือกตัวแบบ ARMA ของวิธีบ็อกซ์-เจนกินส์ เพื่อ พยากรณ์ข้อมูลในอนาคต

1.3 วัตถุประสงค์ของการศึกษา

เพื่อใช้ตัวแบบ ARMA โดยวิธีบ็อกซ์-เจนกินส์ ในการพยากรณ์ปริมาณผลผลิตข้าวแบบรายเดือน ของประเทศไทย

1.4 ขอบเขตการศึกษา

- 1.4.1 ศึกษาข้อมูลปริมาณผลผลิตข้าวในประเทศไทยโดยแบ่งตามฤดูเพาะปลูก ได้แก่ ข้าวนาปีและข้าวนาปรัง ซึ่งใช้ข้อมูลแบบรายเดือนตั้งแต่เดือนมกราคม พ.ศ. 2546 ถึง ธันวาคม พ.ศ. 2561
- 1.4.2 เทคนิคการพยากรณ์ที่ใช้ศึกษาแบบรายเดือน คือ วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method)
- 1.4.3 ความเหมาะสมของตัวแบบพยากรณ์ วัดจากค่า Akaike Information Criterion (AIC)
- 1.4.4 ความแม่นยำของตัวแบบพยากรณ์ วัดจากค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE)
- 1.4.5 เครื่องมือที่ใช้ในการศึกษา คือ โปรแกรม R และ โปรแกรม Microsoft Excel

1.5 วิธีการดำเนินงาน

- 1.5.1 ศึกษาและค้นคว้าข้อมูลจากงานวิจัย และบทความต่าง ๆ ที่เกี่ยวข้องกับการพยากรณ์ อนุกรมเวลาและแบบจำลอง ARMA
- 1.5.2 รวบรวมข้อมูลปริมาณผลผลิตข้าวนาปีและข้าวนาปรัง
- 1.5.3 ตรวจสอบและวิเคราะห์ข้อมูลปริมาณผลผลิตข้าวนาปีและข้าวนาปรัง
- 1.5.4 ใช้วิธีบ็อกซ์-เจนกินส์ในการพยากรณ์ปริมาณผลผลิตข้าวนาปีและข้าวนาปรัง
- 1.5.5 สรุปผลงานวิจัยและจัดทำรูปเล่ม

1.6 ผลที่ได้รับ

ตัวแบบพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทยที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด เพื่อใช้ในการพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย โดยแบ่งตามฤดูเพาะปลูก

1.7 ประโยชน์ที่ได้รับจากการศึกษา

- ก. ประโยชน์ด้านความรู้และประสบการณ์ก่อนนิต
 - 1. ได้ศึกษาและสร้างตัวแบบ ARMA ของวิธีบ็อกซ์-เจนกินส์
 - 2. ได้ศึกษาการพยากรณ์ปริมาณผลผลิตข้าวนาปีและข้าวนาปรังแบบรายเดือนของประเทศไทย
- ข. ประโยชน์ที่ได้จากโครงการงานต่อผู้ใช้

เป็นประโยชน์ในการวิเคราะห์ปริมาณผลผลิตข้าวในอนาคต เพื่อช่วยให้ภาครัฐหรือหน่วยงานราชการที่เกี่ยวข้องสามารถวางแผนหาแนวทางช่วยแก้ปัญหาหาคาข้าวตกต่ำได้

บทที่ 2

งานวิจัยและทฤษฎีที่เกี่ยวข้อง

บทนี้จะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับ โครงการรวมถึงทฤษฎีที่เกี่ยวข้องและแบบจำลองที่ใช้ในการศึกษา

2.1 งานวิจัยที่เกี่ยวข้อง

ศิริเทพ จันทร์บุญแก้ว (2560) ศึกษาการพยากรณ์รายเดือนและรายวันของปริมาณสายโทรเข้าในแต่ละกลุ่มบริการของธนาคาร ภูมิศึกษา ได้แก่ กลุ่มบริการบัตรเครดิต (CD) กลุ่มบริการอิเล็กทรอนิกส์และบัตรกดเงินสด (ES) กลุ่มบริการข้อมูลบัญชีเงินฝาก (AC) กลุ่มบริการอายัดบัญชี (Suspend) และกลุ่มบริการสำหรับกลุ่มลูกค้าพิเศษ (PPS) ของปี พ.ศ. 2558 และ 2559 ในการพยากรณ์ปริมาณสายโทรเข้ารายเดือนจะทำการเปรียบเทียบตัวแบบจากวิธีต่าง ๆ ได้แก่ วิธีเฉลี่ยเคลื่อนที่ (Simple Moving Average method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing method) วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) และวิธีปัจจุบันของธนาคาร โดยใช้เกณฑ์ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) และร้อยละของค่าสัมบูรณ์ของค่าความคลาดเคลื่อนโดยเฉลี่ย (MAPE) ผลการวิจัยพบว่าวิธีบ็อกซ์-เจนกินส์ ให้ประสิทธิภาพในการพยากรณ์ที่ดีที่สุด ในแต่ละกลุ่มบริการของธนาคารทั้ง 5 กลุ่ม เนื่องจากให้ค่าความคลาดเคลื่อนจากค่าจริงน้อยที่สุด

พงพิพัฒน์ คนอยู่ (2555) ศึกษาการพยากรณ์ราคาข้าวเปลือกเจ้านาปีชนิด 5% รายเดือน ด้วยวิธีบ็อกซ์-เจนกินส์ โดยใช้ข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2548 ถึงเดือนกันยายน พ.ศ. 2555 รวมทั้งสิ้น 93 เดือน และใช้โปรแกรมสำเร็จรูป MINITAB ช่วยในการพยากรณ์ จากการศึกษาเบื้องต้นพบว่าข้อมูลราคาข้าวเปลือกเจ้านาปีชนิด 5% เป็นอนุกรมเวลาไม่ Stationary จึงกำหนดตัวแบบที่เหมาะสมคือตัวแบบ ARIMA(0,1,1) เพื่อพยากรณ์ล่วงหน้าจำนวน 1 เดือน คือ เดือนตุลาคม พ.ศ. 2555 และตรวจสอบความเหมาะสมของตัวแบบ โดยใช้ค่าสถิติ Q ซึ่งมีการแจกแจงแบบไคสแควร์ ผลการวิจัยพบว่าวิธีบ็อกซ์-เจนกินส์ ให้ประสิทธิภาพในการพยากรณ์ดีที่สุด เนื่องจากให้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุด

วารงคณา กิรติวิบูลย์ (2556) ศึกษาการเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีบ็อกซ์-เจนกินส์ วิธีการทำให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีฤดูกาลอย่างง่าย และวิธีการพยากรณ์รวมระหว่างวิธีบ็อกซ์-เจนกินส์และวิธีการทำให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีฤดูกาลอย่างง่าย สำหรับการพยากรณ์อุณหภูมิเฉลี่ยต่อเดือนในเขตกรุงเทพมหานคร โดยใช้ข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2546 ถึงเดือนพฤษภาคม พ.ศ. 2555 รวมทั้งสิ้น 113 ค่า จากการศึกษาพบว่าอนุกรมเวลาของข้อมูลมีส่วนประกอบของความแปรผันตามฤดูกาลอย่างชัดเจน และตรวจสอบความเหมาะสมของตัวแบบ โดยใช้ค่าสถิติ Q ซึ่งมีการแจกแจงแบบไคสแควร์ ผลการวิจัยพบว่าวิธีการพยากรณ์รวมมีความเหมาะสมมากที่สุด เนื่องจากให้ค่าเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด

จากงานวิจัยทั้งหมดที่ได้ศึกษามาทำให้ทราบความเหมาะสมสำหรับเทคนิคที่ใช้ในการพยากรณ์แต่ละวิธี นั้นมีความเหมาะสมกับข้อมูลมากน้อยอย่างไร และพยากรณ์ได้แม่นยำใกล้เคียงกับค่าจริงมากน้อยอย่างไร โดยพบว่าวิธีบ็อกซ์-เจนกินส์ให้ค่าใกล้เคียงกับค่าจริงมากที่สุด

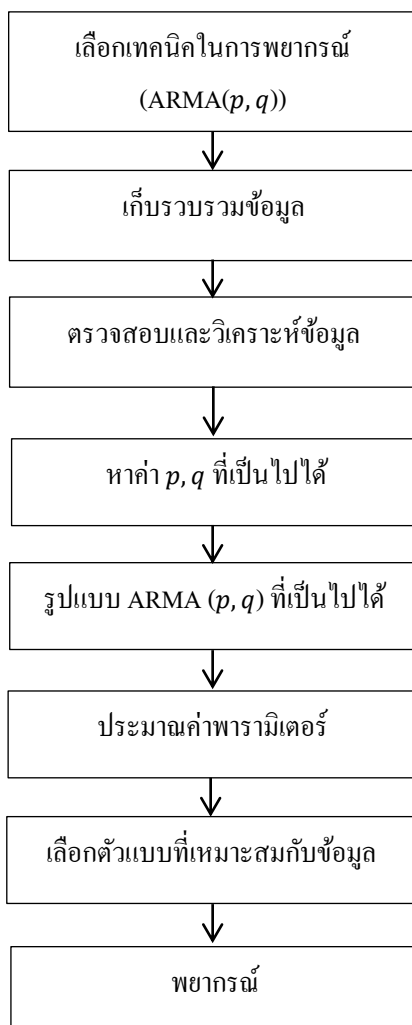
2.2 ทฤษฎีที่เกี่ยวข้อง

2.2.1 ความหมายของการพยากรณ์

การพยากรณ์ (Forecasting) หมายถึง การคาดคะเนหรือการทำนายลักษณะการเกิดของเหตุการณ์ในอนาคต โดยศึกษารูปแบบการเกิดของเหตุการณ์หรือสภาพการณ์จากข้อมูลที่เก็บรวบรวมอย่างมีระบบ และ/หรือจากความรู้ความสามารถ ประสบการณ์และวิจารณญาณของผู้พยากรณ์ (ทรงศิริ แต่สมบัติ, 2549)

2.2.2 ขั้นตอนในการพยากรณ์

กระบวนการพยากรณ์สามารถแบ่งออกเป็น 9 ขั้นตอน ดังแผนภาพต่อไปนี้



2.2.3 ประเภทของการพยากรณ์ทางสถิติ

2.2.3.1 การพยากรณ์เชิงคุณภาพ (Qualitative Forecasting Techniques)

เป็นการพยากรณ์ที่ผู้พยากรณ์ใช้ความรู้ ความสามารถ ประสบการณ์และ/หรือวิจารณ์ญาณในเรื่องที่จะพยากรณ์และเรื่องที่เกี่ยวข้องเป็นหลักในการพยากรณ์ นิยมใช้เมื่อไม่มีข้อมูลในอดีตหรือมีน้อยมาก หรือเป็นการพยากรณ์ระยะยาวที่ใช้ข้อมูลในอดีตไม่ได้ (ทรงศิริ แต่สมบัติ, 2549)

2.2.3.2 การพยากรณ์เชิงปริมาณ (Quantitative Forecasting Techniques)

เป็นการพยากรณ์ที่ใช้ข้อมูลที่มีการเก็บรวบรวมอย่างต่อเนื่อง ตั้งแต่ในอดีตจนถึงปัจจุบันมาวิเคราะห์ การพยากรณ์เชิงปริมาณสามารถแบ่งเป็น 2 วิธี ดังนี้

1. การพยากรณ์เชิงสาเหตุ (Causal Forecasting)

เป็นการพยากรณ์ที่เน้นการหาความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variables) ซึ่งเป็นตัวแทนของ “เหตุ” กับตัวแปรตาม (Dependent Variables) ซึ่งเป็นตัวแทนของ “ผล” (กรินทร์ กาญจนานนท์, 2561)

2. การพยากรณ์อนุกรมเวลา (Time Series Forecasting)

เป็นการศึกษาหารูปแบบการเคลื่อนไหวของอนุกรมเวลาที่กำหนดด้วยตัวแบบอนุกรมเวลา (Time Series Model) ระดับความถูกต้องของการพยากรณ์ขึ้นอยู่กับความเหมาะสมของตัวแบบอนุกรมเวลาที่กำหนด จึงเหมาะกับสภาพของการพยากรณ์ที่ไม่มีปัจจัยอื่น ๆ (ทรงศิริ แต่สมบัติ, 2549)

จะเห็นได้ว่าการพยากรณ์เชิงปริมาณนั้นเหมาะสมที่จะใช้ในการพยากรณ์ข้อมูล เนื่องจากทางสำนักงานเศรษฐกิจการเกษตรมีข้อมูลในอดีตมากเพียงพอสำหรับการพยากรณ์ ดังนั้นงานวิจัยชิ้นนี้จึงเลือกการพยากรณ์เชิงปริมาณ โดยใช้วิธีการพยากรณ์อนุกรมเวลาในการวิเคราะห์ข้อมูลและพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย

2.2.4 การวิเคราะห์อนุกรมเวลา

อนุกรมเวลา (Time Series) หมายถึง ข้อมูลหรือค่าสังเกตที่เก็บรวบรวมตามลำดับเวลาที่เกิดขึ้นอย่างต่อเนื่อง โดยช่วงเวลาที่เก็บรวบรวมข้อมูลอาจจะห่างเท่ากันหรือไม่เท่ากันก็ได้ ซึ่งช่วงเวลาอาจเป็นรายวัน รายสัปดาห์ รายเดือน รายไตรมาส หรือรายปี

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นการศึกษาลักษณะของข้อมูลว่ามีรูปแบบการเคลื่อนไหวอย่างไรและสร้างตัวแบบอนุกรมเวลาแทนรูปแบบการเคลื่อนไหวนั้น เพื่อพยากรณ์ค่าในอนาคตจากตัวแบบอนุกรมเวลาที่สร้างขึ้นภายใต้ข้อสมมติว่ารูปแบบการเคลื่อนไหวในอนาคตไม่ต่างจากรูปแบบการเคลื่อนไหวในอดีต

ลักษณะของอนุกรมเวลา

1. ข้อมูลที่มีลักษณะคงที่ (Stationary) คือ ข้อมูลหรือค่าสังเกตที่มีค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) คงที่ตลอดระยะเวลาที่ศึกษา
2. ข้อมูลที่มีลักษณะไม่คงที่ (Non-Stationary) คือ ข้อมูลหรือค่าสังเกตที่มีค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) ไม่คงที่ตลอดระยะเวลาที่ศึกษา โดยข้อมูลที่มีลักษณะไม่คงที่จะขึ้นกับส่วนประกอบหลัก 4 ส่วนประกอบ ดังต่อไปนี้ (ทรงศิริ แต่สมบัติ, 2549)
 - 2.1 ข้อมูลที่มีแนวโน้ม (Trend: T) เป็นข้อมูลที่มีรูปแบบการเคลื่อนไหวหรือการเปลี่ยนแปลงค่าอนุกรมเวลาในระยะยาวในลักษณะเพิ่มขึ้นหรือลดลงอย่างสม่ำเสมอตามระยะเวลา อาจจะมีแนวโน้มเป็นเส้นตรง (Linear) หรือแบบเส้นโค้งแบบควอดราติก (Quadratic) หรือแบบเอ็กซ์โปเนนเชียล (Exponential)
 - 2.2 ข้อมูลที่มีความแปรผันเกี่ยวกับฤดูกาล (Seasonal: S) เป็นข้อมูลที่มีรูปแบบการเคลื่อนไหวที่เหมือนกันเป็นช่วง ๆ เรียกแต่ละช่วงเวลาที่มีรูปแบบเหมือนกันใน 1 ปีนี้ว่าคาบของฤดูกาล เช่น สำหรับข้อมูลรายไตรมาสคาบของฤดูกาลเท่ากับ 4 เป็นต้น
 - 2.3 ข้อมูลที่มีความแปรผันตามวัฏจักร (Cycle: C) เป็นข้อมูลที่มีรูปแบบการเคลื่อนไหวขึ้น ๆ ลง ๆ คล้ายกันเป็นช่วง ๆ โดยที่แต่ละช่วงมีระยะเวลายาวนานกว่า 1 ปี
 - 2.4 ข้อมูลที่มีความแปรผันแบบผิดปกติหรือการแปรผันไม่สม่ำเสมอ (Irregular: I) เป็นข้อมูลที่มีส่วนประกอบของอนุกรมเวลาที่ปรากฏลักษณะการเคลื่อนไหวที่ไม่มีรูปแบบ การเคลื่อนไหวชนิดนี้คือส่วนที่เหลือจากแนวโน้ม ความแปรผันเกี่ยวกับฤดูกาล ความแปรผันตามวัฏจักร การเคลื่อนไหวแบบผิดปกติโดยทั่วไปเป็นผลกระทบจากเหตุการณ์ที่ไม่คาดคิด เช่น ภัยพิบัติทางธรรมชาติ สงคราม โรคระบาด เป็นต้น

2.2.5 วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method)

George E. P. Box และ Gwilym M. Jenkins ได้นำเสนอรูปแบบ ARIMA ไว้ในหนังสือชื่อ Time Series Analysis: Forecasting and Control พิมพ์เผยแพร่ในปี ค.ศ. 1970 และปรับปรุงในปี ค.ศ. 1994 ซึ่งได้รับความนิยมมากจนถึงปัจจุบัน โดยกำหนดตัวแบบ ARIMA ซึ่งเป็นตัวแบบที่อธิบายรูปแบบการเคลื่อนไหวของอนุกรมเวลาที่ค่าสังเกตมีสหสัมพันธ์เชิงอัตถคอดอย (Autoregressive) และมีลักษณะคงที่ (Stationary) แม้ว่าอนุกรมเวลาที่เก็บรวบรวมโดยทั่วไปจะมีการเคลื่อนไหวเนื่องจากแนวโน้มและ/หรือความแปรผันเกี่ยวกับฤดูกาลซึ่งเป็นลักษณะของอนุกรมเวลาที่ไม่คงที่ ดังนั้นจะต้องทำการแปลงอนุกรมเวลาที่ไม่คงที่ให้เป็นอนุกรมเวลาใหม่ที่คงที่ก่อนแล้วจึงกำหนดตัวแบบ ARIMA ให้กับอนุกรมเวลาใหม่ที่คงที่ (ทรงศิริ แต่สมบัติ, 2549)

เนื่องจากข้อมูลปริมาณผลผลิตข้าวแบบรายเดือนในประเทศไทยแบ่งตามฤดูเพาะปลูก ได้แก่ ข้าวนาปี และข้าวนาปรังที่เก็บรวบรวมข้อมูลโดยสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์เป็นข้อมูลที่มีลักษณะค่าเฉลี่ยและความแปรปรวนไม่คงที่ขึ้นลงตลอดเวลา ดังนั้นในงานวิจัยนี้เราจะสมมติให้อนุกรมเวลามีลักษณะคงที่ (Stationary)

วิธีบ็อกซ์-เจนกินส์ศึกษาความสัมพันธ์ของข้อมูลในอดีต โดยมีสมมติฐานว่าข้อมูลหรือค่าสังเกตในปัจจุบัน (Y_t) มีความสัมพันธ์เป็นกระบวนการเชิงเส้นระหว่างค่าสังเกตในอดีต (Y_{t-1}, Y_{t-2}, \dots) และค่าความคลาดเคลื่อนในอดีต ($\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$) และนำมาพยากรณ์ค่าในอนาคต (Y_{t+1})

โดยทั่วไปจะสมมติว่าค่าความคลาดเคลื่อน ($\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$) มีการแจกแจงปกติและเป็นอิสระต่อกัน มีค่าเฉลี่ยเท่ากับศูนย์ และค่าความแปรปรวนคงที่ (σ^2) หรือ $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots \sim N(0, \sigma^2)$ (ทรงศิริ แต่สมบัติ, 2549)

แนวคิดการพัฒนาเทคนิคการพยากรณ์บ็อกซ์-เจนกินส์ได้มาจากการศึกษาของกระบวนการเชิงเส้น ดังสมการต่อไปนี้ (สมเกียรติ เกตุเอี่ยม, 2548)

$$Y_t = \mu + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \dots$$

โดยที่	Y_t	แทนค่าของอนุกรมเวลา ณ เวลา t โดยที่ $t = 1, 2, \dots, n$
	μ	แทนค่าเฉลี่ยของ Y_t เมื่ออนุกรมเวลามีลักษณะ Stationary
	φ_i	แทนค่าสัมประสิทธิ์ของตัวแบบ โดยที่ $i = 1, 2, \dots, n - 1$
	n	แทนค่าข้อมูลหรือค่าสังเกตทั้งหมด
	ε_t	แทนค่าของอนุกรมเวลาของความคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระต่อกัน โดยค่าเฉลี่ยเท่ากับศูนย์ และค่าความแปรปรวนคงที่ ณ เวลา t โดยที่ $t = 1, 2, \dots, n$

2.2.5.1 ตัวแบบ ARMA (Autoregressive Moving Average)

เป็นการผสมตัวแบบย่อยทั้ง 2 ส่วน ได้แก่ Autoregressive (AR) และ Moving Average (MA) โดยใช้สัญลักษณ์เป็นรูปแบบทั่วไปดังนี้ ARMA (p, q) เมื่อ p คือ จำนวนสัมประสิทธิ์ของตัวแบบ AR และ q คือ จำนวนสัมประสิทธิ์ของตัวแบบ MA

1. ตัวแบบการถดถอยในตัวเอง (Autoregressive model: AR(p))

แนวความคิดของการใช้ตัวแบบการถดถอยในตัวเองจากหลักการถดถอยเชิงเส้นที่ใช้ตัวดำเนินการเชิงเส้นของตัวแปรอิสระต่าง ๆ เพื่อทำนายค่าของตัวแปรตาม แต่ต่างกันในกรณีของตัวแบบการถดถอยในตัวเองจะมีการใช้ตัวดำเนินการเชิงเส้นของค่าของตัวแปรที่จะทำนายในอดีตแทนตัวแปรอิสระ สำหรับตัวแบบการถดถอยในตัวเองจะบ่งชี้จำนวนอันดับ โดยใช้สัญลักษณ์ p สำหรับองค์ประกอบ AR โดยมีการให้น้ำหนักกับแต่ละข้อมูลโดยใช้ค่าสัมประสิทธิ์ถ่วงน้ำหนัก ϕ_1, \dots, ϕ_p (กรินทร์ กาญจนานนท์, 2561)

ตัวแบบการถดถอยในตัว มีสมการดังต่อไปนี้ (Hyndman, R.J., & Athanasopoulos, G., 2018)

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

โดยที่	Y_t	แทนค่าข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$
	Y_{t-k}	แทนค่าข้อมูลหรือค่าสังเกตในอดีต ณ เวลา $t - k$ โดยที่ $k = 1, 2, \dots, p$
	c	แทนค่าคงที่หรือค่าจุดตัดแกน Y
	ϕ_i	แทนค่าสัมประสิทธิ์ของตัวแบบ AR ณ เวลา $t - i$ โดยที่ $i = 1, 2, \dots, p$
	p	แทนอันดับของตัวแบบ AR
	n	แทนค่าข้อมูลหรือค่าสังเกตทั้งหมด
	ε_t	แทนค่าของอนุกรมเวลาของความคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระต่อกัน โดยมีค่าเฉลี่ยเท่ากับศูนย์ และค่าความแปรปรวนคงที่ ณ เวลา t โดยที่ $t = 1, 2, \dots, n$

หรืออาจเขียนในรูปของตัวดำเนินการย้อนหลัง (Backward-shift หรือ Lag Operator: B)

นิยามโดย $BY_t = Y_{t-1}$ จะได้สมการดังต่อไปนี้

$$Y_t = c + \phi_1 BY_t + \phi_2 B^2 Y_t + \dots + \phi_p B^p Y_t + \varepsilon_t$$

$$Y_t - \phi_1 BY_t - \phi_2 B^2 Y_t - \dots - \phi_p B^p Y_t = c + \varepsilon_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = c + \varepsilon_t$$

2. ตัวแบบค่าเฉลี่ยเคลื่อนที่ (Moving Average model: MA(q))

สำหรับตัวแบบค่าเฉลี่ยเคลื่อนที่ที่จะบ่งชี้จำนวนอันดับ โดยใช้สัญลักษณ์ q สำหรับองค์ประกอบ MA โดยมีการให้น้ำหนักกับค่าความคลาดเคลื่อน โดยใช้ค่าสัมประสิทธิ์ถ่วงน้ำหนัก $\theta_1, \dots, \theta_q$ (กรินทร์ กาญจนานนท์, 2561)

ตัวแบบค่าเฉลี่ยเคลื่อนที่ มีสมการดังต่อไปนี้ (Hyndman, R.J., & Athanasopoulos, G., 2018)

$$Y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

โดยที่	Y_t	แทนค่าข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$
	c	แทนค่าคงที่หรือค่าจุดตัดแกน Y
	θ_j	แทนค่าสัมประสิทธิ์ของตัวแบบ MA ณ เวลา $t - j$ โดยที่ $j = 1, 2, \dots, q$
	q	แทนอันดับของตัวแบบ MA
	n	แทนค่าข้อมูลหรือค่าสังเกตทั้งหมด

ε_{t-l} แทนค่าของอนุกรมเวลาของความคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระต่อกัน โดยมีค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงที่ ณ เวลา $t - l$

โดยที่ $l = 0, 1, 2, \dots, q$

หรืออาจเขียนในรูปของตัวดำเนินการย้อนหลัง (Backward-shift หรือ Lag Operator: B)

นิยามโดย $B\varepsilon_t = \varepsilon_{t-1}$ จะได้สมการดังต่อไปนี้

$$Y_t = c + (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)\varepsilon_t$$

3. ตัวแบบผสม หรือตัวแบบ Autoregressive Moving Average

ตัวแบบ Autoregressive Moving Average เป็นแบบจำลองที่ประยุกต์ใช้ตัวแบบ Autoregressive และตัวแบบ Moving Average มาผสมกัน โดยใช้ในการอธิบายอนุกรมเวลา ณ เวลา t (Y_t) จะขึ้นกับข้อมูลในอดีต สามารถเขียนในรูปทั่วไป คือ ตัวแบบ ARMA(p, q) ซึ่งสามารถเขียนแสดงตัวแบบมาตรฐานได้ดังสมการนี้ (Hyndman, R.J., & Athanasopoulos, G., 2018)

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

โดยที่ Y_t แทนค่าข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$

Y_{t-k} แทนค่าข้อมูลหรือค่าสังเกตในอดีต ณ เวลา $t - k$ โดยที่ $k = 1, 2, \dots, p$

c แทนค่าคงที่หรือค่าจุดตัดแกน Y

ϕ_i แทนค่าสัมประสิทธิ์ของตัวแบบ AR ณ เวลา $t - i$ โดยที่ $i = 1, 2, \dots, p$

θ_j แทนค่าสัมประสิทธิ์ของตัวแบบ MA ณ เวลา $t - j$ โดยที่ $j = 1, 2, \dots, q$

p แทนอันดับของตัวแบบ AR

q แทนอันดับของตัวแบบ MA

n แทนค่าข้อมูลหรือค่าสังเกตทั้งหมด

ε_{t-l} แทนค่าของอนุกรมเวลาของความคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระต่อกัน โดยมีค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงที่ ณ เวลา $t - l$

โดยที่ $l = 0, 1, 2, \dots, q$

2.2.5.2 วิธีการพยากรณ์อนุกรมเวลาโดยวิธีบ็อกซ์-เจนกินส์

1. หาอันดับของตัวแบบ AR หรือค่า p และอันดับของตัวแบบ MA หรือค่า q ที่เป็นไปได้ทั้งหมด

จากเรื่องฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function: ACF หรือ ρ_k) ซึ่งเป็นฟังก์ชันการวัดสหสัมพันธ์ระหว่างข้อมูลที่ช่วงเวลาห่างกัน k หน่วย คือ ข้อมูล ณ เวลา t (Y_t) และ $t - k$ (Y_{t-k}) โดยมีสูตรในการคำนวณดังนี้

$$\rho_k = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t-k})}}$$

โดยที่ $Cov(Y_t, Y_{t-k})$ แทนความแปรปรวนร่วมระหว่างข้อมูล ณ เวลา t กับ $t - k$

โดยที่ $k = 1, 2, \dots, n - 1$

$Var(Y_{t-k})$ แทนความแปรปรวนของข้อมูลหรือค่าสังเกต ณ เวลา $t - k$

โดยที่ $k = 1, 2, \dots, n - 1$

n แทนจำนวนข้อมูลหรือค่าสังเกตทั้งหมด

ซึ่งสามารถประมาณค่า ρ_k ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (Correlation Coefficient: r_k) โดยที่ค่า q ที่เป็นไปได้ คือ ค่า k (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) ที่ทำให้ค่า r_k ออกนอกช่วงความเชื่อมั่น 95% โดยมีสูตรในการคำนวณดังนี้ (ทรงศิริ แต่สมบัติ, 2549)

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

โดยที่ Y_t แทนค่าข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$

Y_{t-k} แทนค่าข้อมูลหรือค่าสังเกตในอดีต ณ เวลา $t - k$ โดยที่ $k = 1, 2, \dots, n - 1$

n แทนจำนวนข้อมูลหรือค่าสังเกตทั้งหมด

\bar{Y} แทนค่าเฉลี่ยของข้อมูลทั้งหมด

จากเรื่องฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial Auto-Correlation Function: PACF หรือ ρ_{kk}) ซึ่งพิจารณาความสัมพันธ์ระหว่างข้อมูลที่ช่วงเวลาห่างกัน k หน่วย คือ ข้อมูล ณ เวลา t (Y_t) และ $t - k$ (Y_{t-k}) โดยไม่มีอิทธิพลของข้อมูลระหว่างเวลา t (Y_t) และ $t - k$ (Y_{t-k}) เข้ามาเกี่ยวข้อง โดยมีสูตรในการคำนวณดังนี้

$$\rho_{kk} = \frac{Cov((Y_t | Y_{t-1}, \dots, Y_{t-(k+1)}), (Y_{t-k} | Y_{t-1}, \dots, Y_{t-(k+1)}))}{\sqrt{Var(Y_t | Y_{t-1}, \dots, Y_{t-(k+1)})} \sqrt{Var(Y_{t-k} | Y_{t-1}, \dots, Y_{t-(k+1)})}}$$

โดยที่ $Cov\left((Y_t|Y_{t-1}, \dots, Y_{t-(k+1)}), (Y_{t-k}|Y_{t-1}, \dots, Y_{t-(k+1)})\right)$ แทนค่าความแปรปรวนร่วมระหว่าง $Y_t|Y_{t-1}, \dots, Y_{t-(k+1)}$ กับ $Y_{t-k}|Y_{t-1}, \dots, Y_{t-(k+1)}$ โดยที่ $k = 1, 2, \dots, n - 1$

$Var(Y_{t-k}|Y_{t-1}, \dots, Y_{t-(k+1)})$ แทนความแปรปรวนของข้อมูลหรือค่าสังเกต ณ เวลา $t - k$

โดยที่ $k = 1, 2, \dots, n - 1$

n แทนจำนวนข้อมูลหรือค่าสังเกตทั้งหมด

ซึ่งสามารถประมาณค่า ρ_{kk} ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (Partial Correlation Coefficient: r_{kk}) โดยที่ค่า p ที่เป็นไปได้ คือ ค่า k (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) ที่ทำให้ค่า r_{kk} ออกนอกช่วงความเชื่อมั่น 95% โดยมีสูตรในการคำนวณดังนี้ (ทรงศิริ แต่สมบัติ, 2549)

$$r_{kk} = \begin{cases} r_1 & \text{เมื่อ } k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{เมื่อ } k = 2, 3, \dots, n - 1 \end{cases}$$

โดยที่ $r_{k,j} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$ สำหรับ $j = 1, 2, \dots, k - 1, k = 2, 3, \dots$ และ $k > j$ โดยค่า r_k, r_{kk} มีค่าอยู่ระหว่าง -1 ถึง 1

เนื่องจากข้อมูลมีจำนวนมากกว่า 30 จึงใช้ค่าสถิติ Z ในการคำนวณ ดังสูตรต่อไปนี้

$$Z_{cal} = (r_k - \rho_k) * \sqrt{n}$$

โดยที่ n คือ จำนวนข้อมูลหรือค่าสังเกตทั้งหมด

การหาช่วงความเชื่อมั่น $(1 - \alpha) \cdot 100\%$ ของ r_k จาก

$$P\left(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(-Z_{\frac{\alpha}{2}} < (r_k - \rho_k) * \sqrt{n} < Z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\frac{-Z_{\frac{\alpha}{2}}}{\sqrt{n}} < r_k - \rho_k < \frac{Z_{\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(\frac{-Z_{\frac{\alpha}{2}}}{\sqrt{n}} + \rho_k < r_k < \frac{Z_{\frac{\alpha}{2}}}{\sqrt{n}} + \rho_k\right) = 1 - \alpha$$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้น $Z_{\frac{\alpha}{2}}$ เท่ากับ 1.96 จะได้ว่า ช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ และหาช่วงความเชื่อมั่น 95% ของ r_{kk} ในทำนองเดียวกันกับการหาช่วงความเชื่อมั่น 95% ของ r_k

2. ประมาณค่าพารามิเตอร์ $\phi_i, i = 1, 2, \dots, p$ และ $\theta_j, j = 1, 2, \dots, q$ สำหรับแต่ละ p, q ที่เป็นไปได้ทั้งหมดจากข้อ 1

2.1 ประมาณค่าพารามิเตอร์ ϕ_i ของตัวแบบ AR โดยวิธีกำลังสองน้อยที่สุด คือ การคำนวณค่าสัมประสิทธิ์การถดถอยเชิงเส้น ดังสมการต่อไปนี้ (ทรงศิริ แต่สมบัติ, 2549)

$$\phi_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y}_t)(Y_{t-k} - \bar{Y}_{t-k})}{\sum_{t=k+1}^n (Y_{t-k} - \bar{Y}_{t-k})^2}$$

โดยที่ Y_t แทนค่าข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$
 Y_{t-k} แทนค่าข้อมูลหรือค่าสังเกตในอดีต ณ เวลา $t - k$ โดยที่ $k = 1, 2, \dots, p$
 \bar{Y}_t แทนค่าเฉลี่ยของข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$
 \bar{Y}_{t-k} แทนค่าเฉลี่ยของข้อมูลหรือค่าสังเกต ณ เวลา $t - k$ โดยที่ $k = 1, \dots, p$
 k คือ ช่วงเวลาที่ข้อมูลหรือค่าสังเกตอยู่ห่างกัน
 n คือ จำนวนข้อมูลหรือค่าสังเกตทั้งหมด
 p แทนอันดับของตัวแบบ AR

2.2 ประมาณค่าพารามิเตอร์ θ_j ของตัวแบบ MA โดยใช้หลักการของโมเมนต์ (Method of Moments) เมื่อทราบค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (r_k) ดังสมการต่อไปนี้

$$r_k = \frac{\sum_{j=0}^{q-k} \theta_j \theta_{k+j}}{\sum_{j=0}^q \theta_j^2}; k = 1, \dots, q$$

โดยที่ θ_j แทนค่าสัมประสิทธิ์ของตัวแบบ MA ณ เวลา $t - j$ โดยที่ $j = 1, 2, \dots, q$
 และ $\theta_0 = 1$
 k คือ ช่วงเวลาที่ข้อมูลหรือค่าสังเกตอยู่ห่างกัน
 q แทนอันดับของตัวแบบ MA

2.3 ค่าจุดตัดแกน Y หรือค่า c คำนวณได้จากสมการต่อไปนี้

$$c = \bar{Y}_t - \phi_1 \bar{Y}_{t-1} - \dots - \phi_p \bar{Y}_{t-p}$$

โดยที่ \bar{Y}_t แทนค่าเฉลี่ยของข้อมูลหรือค่าสังเกต ณ เวลา t โดยที่ $t = 1, 2, \dots, n$
 \bar{Y}_{t-k} แทนค่าเฉลี่ยของข้อมูลหรือค่าสังเกต ณ เวลาที่ $t - k$ โดยที่ $k = 1, \dots, p$

- k คือ ช่วงเวลาที่ข้อมูลหรือค่าสังเกตอยู่ห่างกัน
- ϕ_i แทนค่าสัมประสิทธิ์ของตัวแบบ AR ณ เวลา $t - i$ โดยที่ $i = 1, 2, \dots, p$
- p แทนอันดับของตัวแบบ AR

3. เลือกตัวแบบ ARMA (p, q) ที่เหมาะสมกับข้อมูล

โดยพิจารณาจากค่าสถิติที่ใช้เป็นเกณฑ์ คือ ค่า Akaike Information Criterion (AIC) ในปี ค.ศ. 1973 อาไคเคะได้เสนอเกณฑ์ในการคัดเลือกตัวแบบเพื่อใช้หาตัวแบบที่ให้ค่าพยากรณ์แม่นยำที่สุด เมื่อขนาดตัวอย่างใหญ่ สมการถดถอยที่ได้จากตัวแบบที่มีค่าเอไอซีต่ำที่สุดจะเป็นตัวแบบที่เหมาะสมที่สุด โดยเกณฑ์เอไอซีมีนิยามดังนี้ (Hyndman, R.J., & Athanasopoulos, G., 2018)

$$AIC = -2 \log(L) + 2(p + q + k + 1)$$

โดยที่ L คือ ค่า Likelihood function ของข้อมูล

p แทนอันดับของตัวแบบ AR

q แทนอันดับของตัวแบบ MA

c แทนค่าคงที่หรือค่าจุดตัดแกน Y

$k = 1$ ถ้า $c \neq 0$ และ $k = 0$ ถ้า $c = 0$

เนื่องจากอนุกรมเวลาเป็นตัวแปรสุ่ม (Y_i) ที่มีการแจกแจงปกติ (Normal distribution) ที่มีพารามิเตอร์ $\mu \in \mathbb{R}$ และ $\sigma > 0$ และเขียนแทนด้วย $Y_i \sim \mathcal{N}(\mu, \sigma^2)$ ดังนั้นฟังก์ชันความหนาแน่นความน่าจะเป็นสามารถเขียนได้ดังนี้

$$f(Y_i; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(Y_i - \mu)^2}{2\sigma^2}}$$

เพราะฉะนั้น ค่า Likelihood function (L) สามารถคำนวณได้จาก

$$\begin{aligned} L(\theta) &= \prod_{i=1}^n (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(Y_i - \mu)^2}{\sigma^2}\right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \mu)^2\right) \end{aligned}$$

โดยที่ μ แทนค่าเฉลี่ยของข้อมูลหรือค่าสังเกต

σ^2 แทนค่าความแปรปรวนของข้อมูลหรือค่าสังเกต

n แทนจำนวนข้อมูลหรือค่าสังเกตทั้งหมด

จากเกณฑ์ในการเลือกตัวแบบ ARMA(p, q) คือ ค่า Akaike Information Criterion (AIC) เราจะพิจารณาเลือกตัวแบบ ARMA(p, q) ที่เหมาะสมที่ให้ค่า AIC ต่ำที่สุด

4. ได้สมการพยากรณ์และผลการพยากรณ์

เมื่อเราทำการเลือกตัวแบบ ARMA (p, q) หรืออีกความหมายหนึ่งก็คือ เราจะได้ p, q ที่เหมาะสมกับข้อมูลที่สุด หลังจากนั้นจะทำการสร้างสมการพยากรณ์ โดยมีสมการรูปทั่วไปดังนี้

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = c + (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t$$

เมื่อแทนค่าประมาณพารามิเตอร์ทั้งหมดและค่าจุดตัดแกน Y (c) ที่ได้จากข้อ 2 แล้วหลังจากนั้น จะ ได้ ข้อมูลหรือค่าสังเกตในอดีต (Y_{t-1}, Y_{t-2}, \dots) เพื่อพยากรณ์แบบจุดที่ค่า Y_t

เกณฑ์ที่นิยมใช้ในการวัดค่าความแม่นยำ (Accuracy) สามารถดูได้จากค่าความคลาดเคลื่อนจากการพยากรณ์ (Forecast Errors) หรือค่าส่วนเหลือ (Residuals) ณ เวลา t ใด ๆ ซึ่งก็คือค่าความแตกต่างระหว่างค่าที่ได้จากค่าสังเกต (Y_t) และค่าพยากรณ์ (F_t) โดยสามารถแสดงเป็นสมการได้ดังนี้

$$\text{ค่าความคลาดเคลื่อนจากการพยากรณ์} = F_t - Y_t$$

ตามทฤษฎี ถ้าเลือกตัวแบบที่เหมาะสมในการพยากรณ์แล้วค่าความคลาดเคลื่อนจากการพยากรณ์จะต้องมีค่าน้อยที่สุด ดังนั้นผลรวมของค่าความคลาดเคลื่อนจากการพยากรณ์ควรมีค่าเท่ากับศูนย์หรือมีค่าต่ำที่สุด สำหรับวิธีการคำนวณค่าความคลาดเคลื่อนตลอดช่วงเวลาของการพยากรณ์สามารถพิจารณาได้ดังรูปแบบต่าง ๆ ดังนี้

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) (ศิริเทพ จันทร์บุญแก้ว, 2560)

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}}$$

โดยที่ F_t แทนค่าพยากรณ์ของอนุกรมเวลา ณ เวลา t โดยที่ $t = 1, 2, \dots, n$

Y_t แทนค่าจริงของอนุกรมเวลา ณ เวลา t โดยที่ $t = 1, 2, \dots, n$

n แทนจำนวนข้อมูลหรือค่าสังเกตทั้งหมด

บทที่ 3

การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทยแบบรายเดือน

ในงานวิจัยนี้จะทำการพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทยแบบรายเดือน โดยแบ่งตามฤดูเพาะปลูก ได้แก่ ข้าวนาปีและข้าวนาปรังของปี พ.ศ. 2557, 2558, 2559, 2560 และ 2561 ในบทนี้จะทำการหาตัวแบบในการพยากรณ์แต่ละเดือนที่เหมาะสมกับข้อมูลทั้ง 2 ฤดูเพาะปลูกที่สุด โดยใช้ตัวแบบ ARMA ของวิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) และใช้ค่า Akaike Information Criterion (AIC) เป็นเกณฑ์ในการเลือกตัวแบบ ARMA ที่เหมาะสมที่สุดให้กับข้อมูล

3.1 แนวทางการพยากรณ์

การจะพยากรณ์แบบรายเดือนได้นั้นจะเริ่มจากการเลือกวิธีที่จะใช้ในการพยากรณ์ โดยในงานวิจัยนี้จะใช้วิธีบ็อกซ์-เจนกินส์ ในการพยากรณ์โดยวิธีนี้เป็นวิธีที่สร้างสมการพยากรณ์โดยอาศัยความสัมพันธ์ของข้อมูลซึ่งข้อมูลที่จะนำมาสร้างตัวแบบได้นั้นจะต้องมีลักษณะคงที่ ขั้นตอนต่อไปจะทำการหาตัวแบบที่เป็นไปได้ทั้งหมดจากกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) หลังจากนั้นจะทำการประมาณค่าพารามิเตอร์ของตัวแบบที่เป็นไปได้ทั้งหมด จากนั้นทำการเลือกตัวแบบ ARMA ที่เหมาะสมที่สุดกับข้อมูล โดยใช้ค่า AIC เป็นเกณฑ์ในการตัดสินใจต่อไปจะเป็นการพยากรณ์แบบรายเดือนโดยวิธีบ็อกซ์-เจนกินส์ ซึ่งแบ่งเป็นการพยากรณ์ปริมาณผลผลิตข้าวนาปีและการพยากรณ์ปริมาณผลผลิตข้าวนาปรัง โดยให้ผลการพยากรณ์ดังต่อไปนี้

3.2 การพยากรณ์ปริมาณผลผลิตข้าวนาปีแบบรายเดือนในประเทศไทย



รูปที่ 3.2.1 กราฟแสดงปริมาณผลผลิตข้าวนาปีแบบรายเดือนในประเทศไทยตั้งแต่ปี พ.ศ. 2546-2561

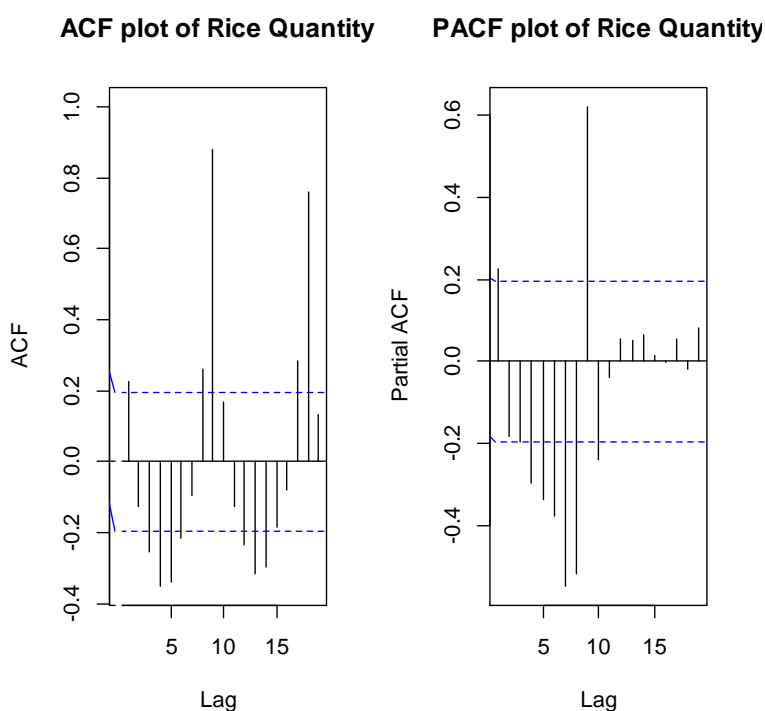
เนื่องจากวิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) เหมาะกับข้อมูลที่มีลักษณะคงที่ (Stationary) และเหมาะสมกับการใช้ข้อมูลในการพยากรณ์ประมาณ 50 ค่าขึ้นไป ดังนั้น จำนวนข้อมูลที่ใช้ในการพยากรณ์ปริมาณผลผลิตข้าวนาปีในประเทศไทยจะใช้ข้อมูลแบบรายเดือนจำนวน 9 เดือนต่อปี ได้แก่ เดือนมกราคม กุมภาพันธ์ มีนาคม เมษายน สิงหาคม กันยายน ตุลาคม พฤศจิกายน และธันวาคม เนื่องจากข้าวนาปีจะเริ่มเก็บเกี่ยวตั้งแต่เดือนสิงหาคมของปีหนึ่งจนถึงเดือนเมษายนของปีถัดไป ซึ่งรวมเก็บข้อมูลจำนวนทั้งสิ้น 99 เดือน หรือ 11 ปี เพื่อพยากรณ์ปี พ.ศ. 2557 (ตั้งแต่ปี พ.ศ. 2546-2556), ปี พ.ศ. 2558 (ตั้งแต่ปี พ.ศ. 2547-2557), ปี พ.ศ. 2559 (ตั้งแต่ปี พ.ศ. 2548-2558), ปี พ.ศ. 2560 (ตั้งแต่ปี พ.ศ. 2549-2559) และปี พ.ศ. 2561 (ตั้งแต่ปี พ.ศ. 2550-2560)

การพยากรณ์ปริมาณผลผลิตข้าวนาปีแบบรายเดือนในประเทศไทยปี พ.ศ. 2557 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.2.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปีแบบรายเดือนตั้งแต่ปี พ.ศ. 2546-2556 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.2.2 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.2.2 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ (-0.1970,0.1970) ตัวสุดท้ายคือ 18 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 18 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ (-0.1970,0.1970) ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,3) สามารถแสดงได้ดังตารางที่ 3.2.3 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.2.3 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$p \backslash q$	0	1	2	3	4	5	6	7	8	9	10
0	-	3298.25	3296.89	3295.07	3287.49	3276.65	3260.73	3217.42	NA	NA	NA
1	3296.83	3298.70	3269.96	3265.49	3255.64	3246.29	3230.44	3200.23	NA	NA	NA
2	3298.45	3279.28	3227.37	3226.19	3207.42	3189.59	3232.33	3125.10	NA	NA	NA
3	3278.33	3278.88	3272.31	3225.72	3218.60	3242.11	3174.79	3120.80	NA	NA	3018.57
4	3261.75	3258.81	3256.70	3258.51	3261.32	3206.42	3183.92	NA	NA	NA	3020.20
5	3267.27	3268.37	3236.12	3230.74	3191.89	NA	3170.98	NA	NA	NA	NA
6	3251.20	3249.38	3250.49	3189.04	3173.48	NA	NA	3093.40	NA	NA	NA
7	3247.36	3220.21	3220.41	3216.62	3185.77	NA	3128.63	3071.95	3039.64	3021.52	3023.33
8	3235.21	3229.61	3216.31	3215.37	3153.90	3164.27	NA	3055.24	3021.61	3022.18	3022.17
9	3200.40	3201.02	3201.92	3201.98	3202.06	3156.41	3155.64	3070.27	3019.67	NA	3019.51
10	3200.35	3195.48	3195.85	3175.91	3150.03	3149.13	3112.06	3061.30	3021.77	3020.54	NA
11	3196.58	3196.92	NA	3135.43	3165.16	3148.02	3141.38	3060.46	3022.88	NA	3024.36
12	3196.50	3199.27	3178.39	3132.64	3135.28	3139.70	3127.55	3061.73	3023.89	NA	NA
13	3197.97	3194.45	3180.96	3181.27	3182.39	NA	3122.20	3056.32	3025.35	NA	NA
14	3179.53	3191.94	3192.94	3182.03	3116.00	3122.87	3110.47	3050.93	3028.88	NA	3028.23
15	3196.97	3188.79	3153.17	3153.06	3120.29	3119.16	3124.12	3068.62	3030.66	NA	NA
16	3190.52	3191.79	3155.05	3154.64	3134.54	3112.82	3134.00	3050.23	3030.15	NA	NA
17	3172.93	3183.92	3156.56	3157.66	3112.84	NA	3080.12	3056.39	3028.58	NA	NA
18	3158.41	3156.68	3157.46	3157.49	3147.51	3107.34	3105.83	3050.68	3029.69	NA	NA

หมายเหตุ : NA = not available

จากตารางที่ 3.2.3 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 - 0.1464B^1 + 0.0191B^2 + 0.0175B^3 + 0.0285B^4 + 0.0226B^5 + 0.0243B^6 + 0.0372B^7 - 0.0424B^8 - 0.9359B^9 + 0.1932B^{10})Y_t = 2,766,884.5 + (1 - 0.6167B^1 + 0.1671B^2 - 0.1124B^3)\varepsilon_t$$

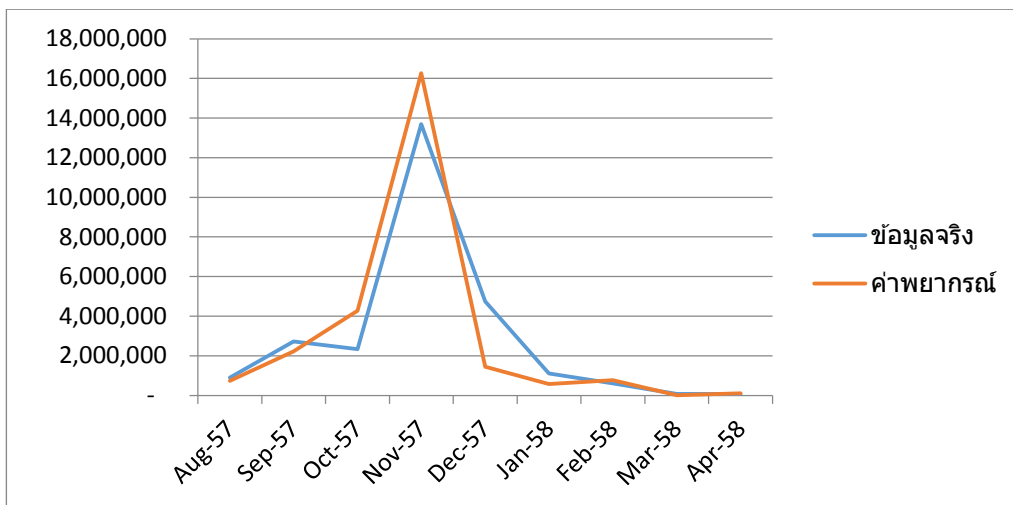
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,3) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.2.4 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปีจริง ดังรูปที่ 3.2.5 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2557 ถึงเดือนเมษายนปี พ.ศ. 2558 ดังรูปที่ 3.2.6

ตารางที่ 3.2.4 แสดงปริมาณผลผลิตข้าวนาปีจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปี

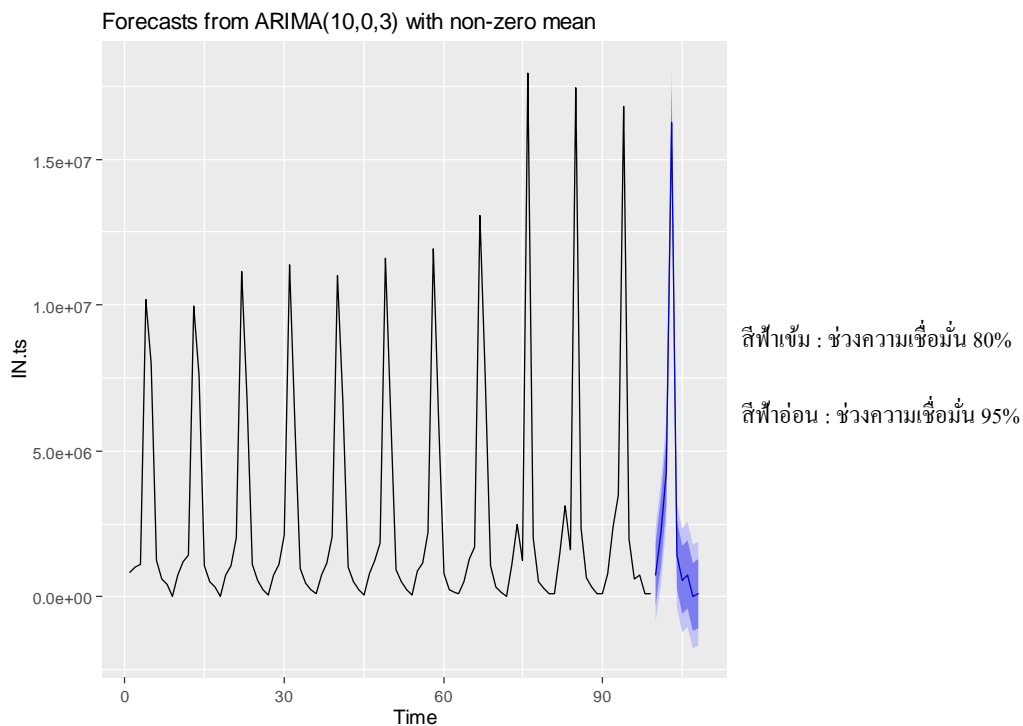
วันที่	ปริมาณผลผลิตข้าวนาปีจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
สิงหาคม 2557	898,201	737,757	-160,444	-17.86%
กันยายน 2557	2,718,980	2,218,610	-500,370	-18.40%
ตุลาคม 2557	2,340,417	4,278,128	1,937,711	82.79%
พฤศจิกายน 2557	13,698,095	16,259,758	2,561,663	18.70%
ธันวาคม 2557	4,745,259	1,441,269	-3,303,990	-69.63%
มกราคม 2558	1,111,467	582,323	-529,144	-47.61%
กุมภาพันธ์ 2558	608,903	767,228	158,325	26.00%
มีนาคม 2558	68,335	13,816	-54,519	-79.78%
เมษายน 2558	80,307	113,842	33,535	41.76%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\ &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\ &= 1,557,008 \end{aligned}$$



รูปที่ 3.2.5 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปีจริง



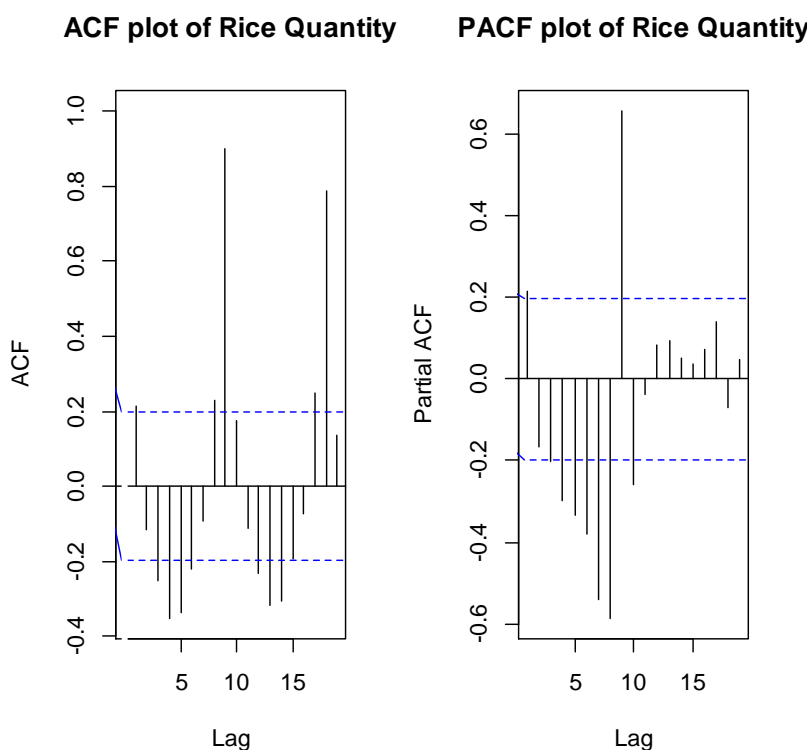
รูปที่ 3.2.6 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2557 ถึงเดือนเมษายนปี พ.ศ. 2558

การพยากรณ์ปริมาณผลผลิตข้าวนาปีแบบรายเดือนในประเทศไทยปี พ.ศ. 2558 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.2.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปีแบบรายเดือนตั้งแต่ปี พ.ศ. 2547-2557 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.2.7 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.2.7 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 18 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 18 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,2) สามารถแสดงได้ดังตารางที่ 3.2.8 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.2.8 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	3300.88	3300.13	3298.03	3290.47	3279.91	3264.55	3225.99	NA	3045.77	NA
1	3299.77	3301.68	3273.55	3268.10	3258.36	3248.86	3234.45	3206.90	3101.66	NA	NA
2	3301.50	3283.05	3229.92	3228.32	3214.24	3197.11	3234.49	3144.38	3080.60	NA	3041.76
3	3281.71	3282.52	3276.27	3227.29	3219.58	3246.46	3182.97	3132.39	3069.93	NA	NA
4	3261.96	3261.21	3260.02	3258.88	3260.60	3209.00	3189.98	3120.97	3059.09	NA	NA
5	3271.23	3272.95	3253.94	3236.12	3197.93	NA	3177.59	3116.74	3060.20	NA	NA
6	3239.92	3240.89	3252.19	3185.89	3176.15	3194.93	3171.08	3131.43	3062.19	NA	NA
7	3240.08	3222.83	3220.12	3234.66	3195.55	NA	3142.43	3092.33	3064.45	3049.07	3050.82
8	3218.91	3220.34	3232.39	3226.65	3167.18	3195.89	3129.23	3100.76	3049.28	3048.81	3052.29
9	3204.29	3206.27	3204.80	3204.63	3170.69	3152.13	3125.56	3096.07	3048.50	3046.60	3048.31
10	3206.25	3206.26	3194.28	3179.17	3156.29	3152.73	3126.91	3090.65	3053.02	3048.16	3048.87
11	3198.02	3199.39	3148.91	3197.69	3167.38	3154.33	3134.85	3084.76	3052.97	3045.37	3047.95
12	3198.37	3200.33	NA	3152.86	3144.12	3147.02	3108.44	3081.99	3052.29	3049.46	3049.20
13	3200.18	3200.73	3185.22	3186.78	3129.47	3125.10	3132.71	3073.21	3055.38	3049.76	3054.06
14	3180.91	3179.45	3175.69	3183.04	3149.38	3124.69	3124.42	3075.99	3055.56	3051.07	3054.47
15	3179.28	3194.11	3195.73	3197.83	3136.29	3126.42	3128.21	3086.16	3058.89	3054.06	3055.78
16	3179.75	3174.38	3166.40	3167.18	3132.95	3128.09	3095.66	3073.89	3058.66	3054.32	3056.56
17	3189.55	3188.29	3167.68	3170.20	3129.69	3129.52	3113.47	3082.44	3056.88	3056.30	3057.49
18	3154.56	3154.76	3156.64	3157.36	3157.99	3122.00	3127.03	3077.99	3058.21	3057.07	3059.26

หมายเหตุ : NA = not available

จากตารางที่ 3.2.8 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 - 0.6213B^1 + 0.0117B^2 + 0.0049B^3 + 0.0140B^4 + 0.0067B^5 + 0.0103B^6 + 0.0152B^7 - 0.0461B^8 - 0.9124B^9 + 0.6002B^{10})Y_t = 2,761,832.2 + (1 - 1.0757B^1 + 0.2209B^2)\varepsilon_t$$

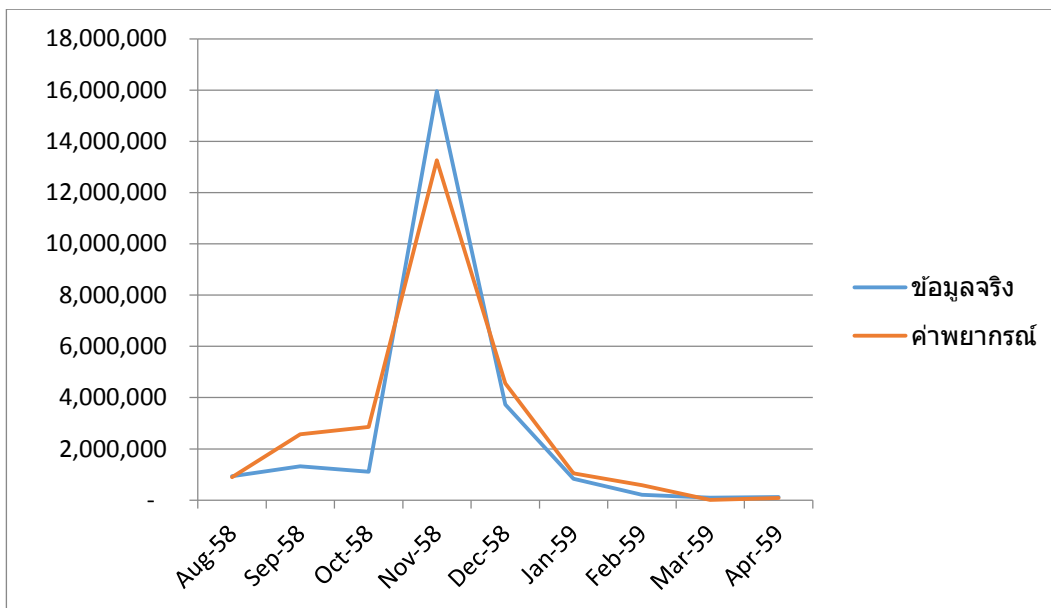
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,2) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.2.9 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวในปีจริง ดังรูปที่ 3.2.10 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2558 ถึงเดือนเมษายนปี พ.ศ. 2559 ดังรูปที่ 3.2.11

ตารางที่ 3.2.9 แสดงปริมาณผลผลิตข้าวนาปีจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปี

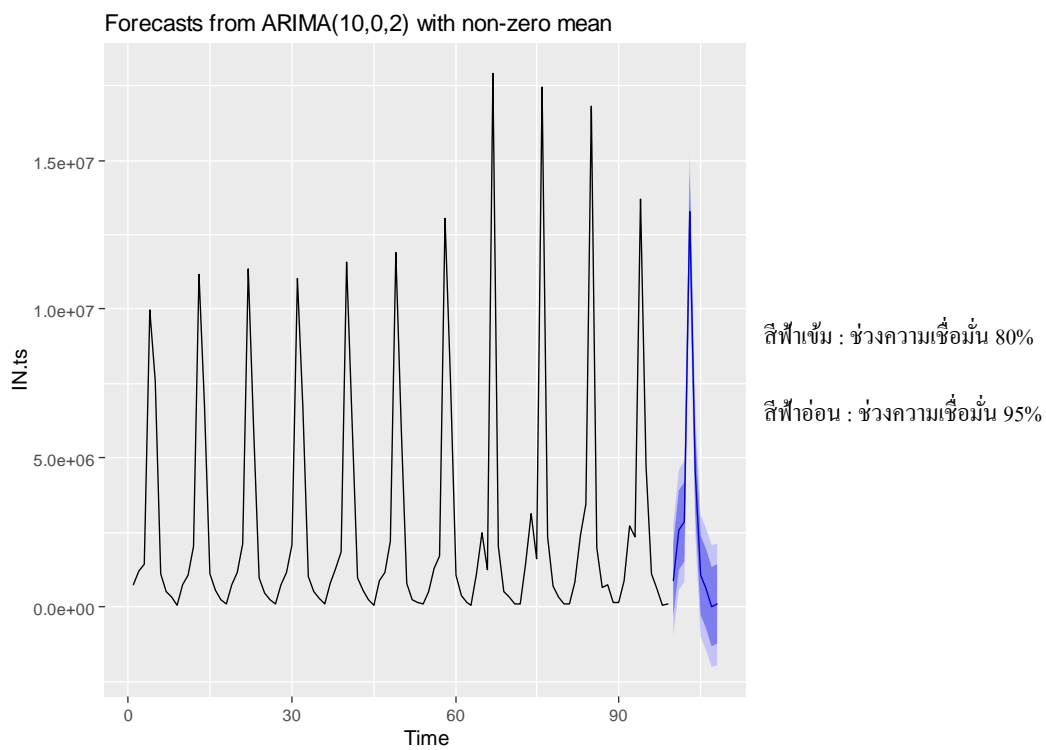
วันที่	ปริมาณผลผลิตข้าวนาปีจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
สิงหาคม 2558	926,997	890,753	-36,244	-3.91%
กันยายน 2558	1,324,350	2,570,460	1,246,110	94.09%
ตุลาคม 2558	1,110,660	2,861,747	1,751,087	157.66%
พฤศจิกายน 2558	15,956,560	13,260,364	-2,696,196	-16.90%
ธันวาคม 2558	3,729,766	4,558,597	828,831	22.22%
มกราคม 2559	832,625	1,043,501	210,876	25.33%
กุมภาพันธ์ 2559	204,064	584,247	380,183	186.31%
มีนาคม 2559	101,899	14,137	-87,762	-86.13%
เมษายน 2559	124,618	79,111	-45,507	-36.52%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned}
 \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\
 &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\
 &= 1,191,433
 \end{aligned}$$



รูปที่ 3.2.10 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวในปีจริง



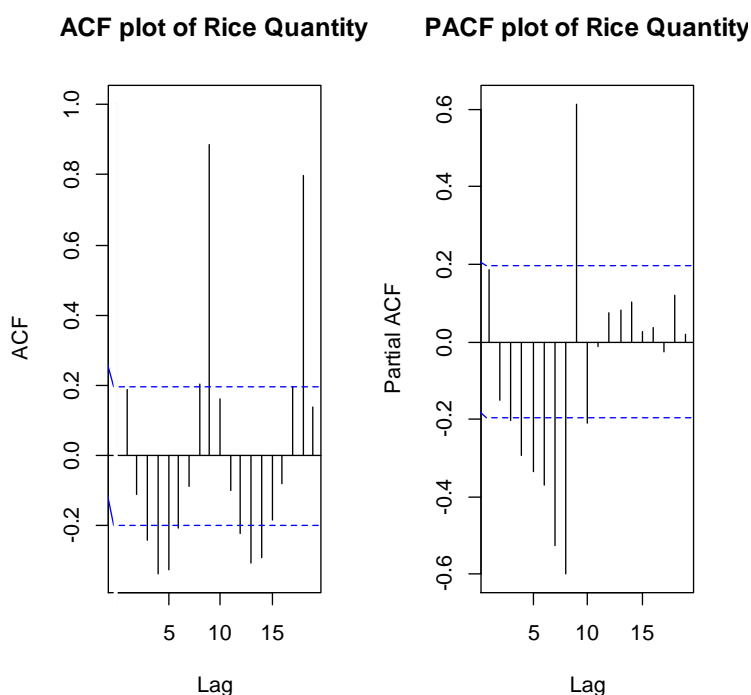
รูปที่ 3.2.11 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวในปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2558 ถึงเดือนเมษายนปี พ.ศ. 2559

การพยากรณ์ปริมาณผลผลิตข้าวในปีแบบรายเดือนในประเทศไทยปี พ.ศ. 2559 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.2.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวในปีแบบรายเดือนตั้งแต่ปี พ.ศ. 2548-2558 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.2.12 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.2.12 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ (-0.1970, 0.1970) ตัวสุดท้ายคือ 18 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 18 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ (-0.1970, 0.1970) ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,3) สามารถแสดงได้ดังตารางที่ 3.2.13 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.2.13 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	3307.69	3307.44	3305.36	3298.13	3287.45	3272.47	3235.05	3143.28	NA	NA
1	3306.89	3308.79	3280.57	3274.61	3264.94	3255.23	3241.16	3213.29	3102.40	NA	NA
2	3308.51	3288.91	3237.01	3234.23	3221.91	3207.62	3241.24	3151.37	NA	NA	NA
3	3287.81	3281.10	3283.10	3236.90	3227.55	3251.36	3193.05	NA	NA	NA	3053.08
4	3276.78	3272.42	3269.16	3271.00	3271.89	3216.78	3194.28	3132.34	NA	NA	NA
5	3278.66	3278.46	3265.93	3245.08	3206.90	NA	3191.52	NA	NA	NA	NA
6	3251.14	3252.81	3263.31	3264.04	3186.33	NA	3179.65	3138.93	NA	NA	3059.05
7	3257.65	3232.90	3233.26	3233.17	3196.94	NA	NA	3099.06	3069.38	NA	NA
8	3265.68	3238.14	3227.51	3228.79	3163.11	3165.04	NA	3092.87	3055.94	NA	NA
9	3215.42	3213.61	3215.23	3215.43	3215.66	3159.67	3165.5	3105.11	3054.17	NA	NA
10	3212.30	3206.90	3208.31	3202.74	3176.73	3161.85	3163.05	3104.58	3056.25	3054.51	3055.39
11	3208.24	3208.63	3166.86	3154.48	3186.01	3163.60	3157.21	3092.61	3057.65	3054.26	3056.74
12	3207.44	3209.38	3199.74	3149.82	3151.11	3151.94	3116.82	3091.64	3059.20	3056.14	3058.26
13	3209.20	3209.23	3201.14	3198.31	3200.28	3134.53	3137.84	3086.51	3060.91	3057.73	3059.91
14	3195.75	3191.40	3206.14	3194.70	3155.78	3133.61	3127.79	3088.55	3062.49	3060.79	3062.79
15	3189.34	3199.15	3203.16	3175.78	3141.12	3134.57	3104.39	3095.85	3063.24	3062.56	3063.35
16	3203.71	3205.70	3177.60	3175.65	3157.85	3132.05	3147.55	3079.73	3066.62	3063.10	3065.34
17	3186.42	3197.47	3176.31	3178.01	3133.54	3132.17	3115.76	3081.21	3065.85	3063.51	3067.05
18	3166.53	3168.52	3168.62	3165.18	3170.30	3125.93	3140.87	3084.10	3066.25	3066.30	3069.33

หมายเหตุ : NA = not available

จากตารางที่ 3.2.13 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1+0.0632B^1 + 0.0438B^2 + 0.0461B^3 + 0.0497B^4 + 0.0474B^5 + 0.0443B^6 + 0.0574B^7 + 0.0053B^8 - 0.9170B^9 - 0.0139B^{10})Y_t = 2,753,746.86 + (1 - 0.4025B^1 - 0.0744B^2 - 0.0382B^3)\varepsilon_t$$

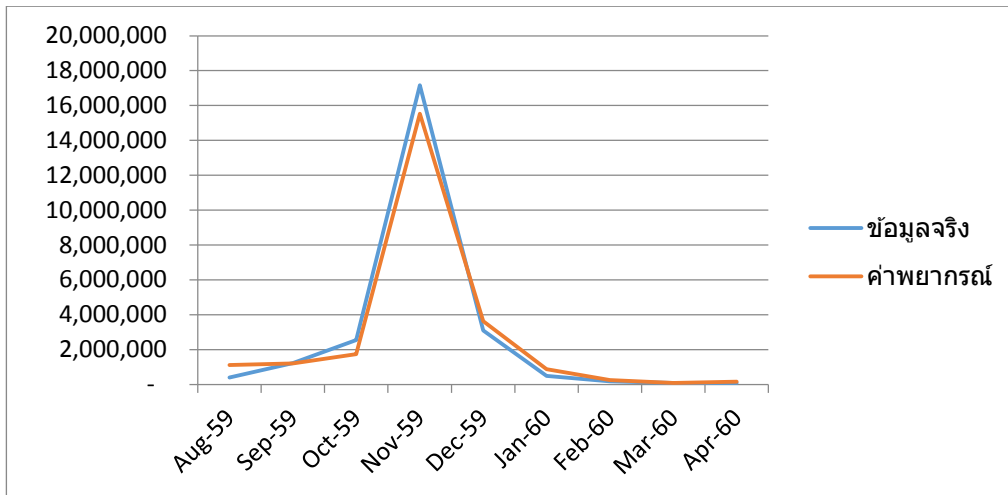
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,3) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.2.14 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปีจริง ดังรูปที่ 3.2.15 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2559 ถึงเดือนเมษายนปี พ.ศ. 2560 ดังรูปที่ 3.2.16

ตารางที่ 3.2.14 แสดงปริมาณผลผลิตข้าวนาปีจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปี

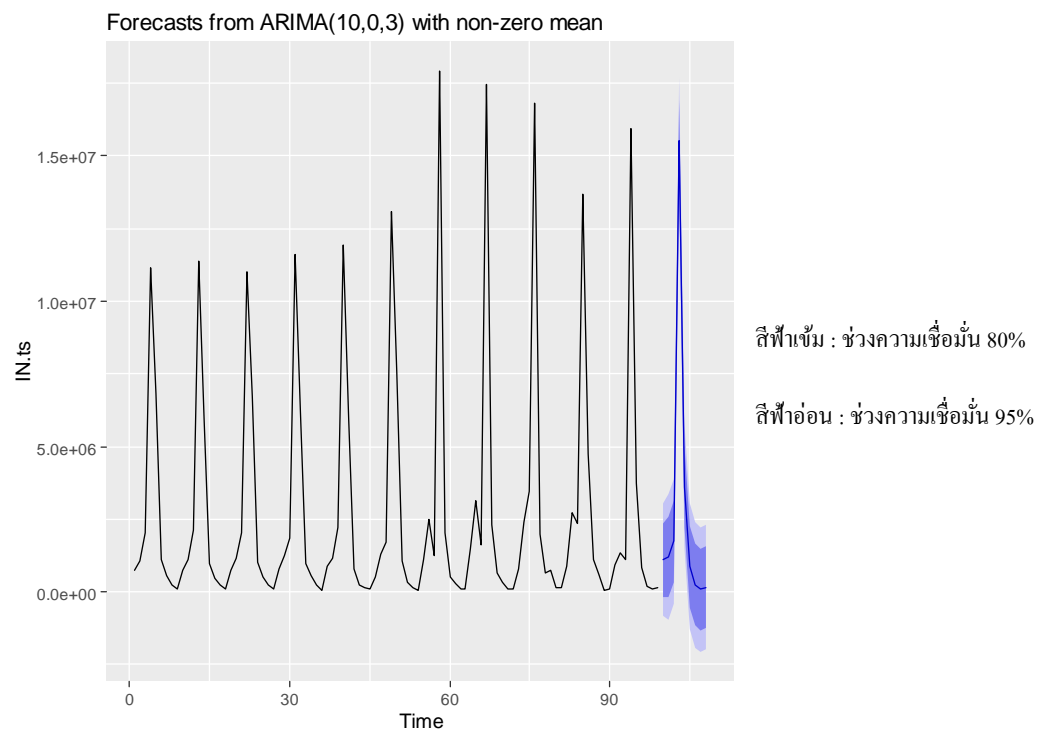
วันที่	ปริมาณผลผลิตข้าวนาปีจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
สิงหาคม 2559	389,749	1,108,669	718,920	184.46%
กันยายน 2559	1,225,567	1,205,355	-20,212	-1.65%
ตุลาคม 2559	2,538,734	1,736,068	-802,666	-31.62%
พฤศจิกายน 2559	17,160,004	15,520,800	-1,639,204	-9.55%
ธันวาคม 2559	3,092,658	3,621,624	528,966	17.10%
มกราคม 2560	485,072	875,132	390,060	80.41%
กุมภาพันธ์ 2560	176,861	242,260	65,399	36.98%
มีนาคม 2560	78,588	78,113	-475	-0.60%
เมษายน 2560	89,112	165,733	76,621	85.98%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\ &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\ &= 690,460 \end{aligned}$$



รูปที่ 3.2.15 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวในปีจริง



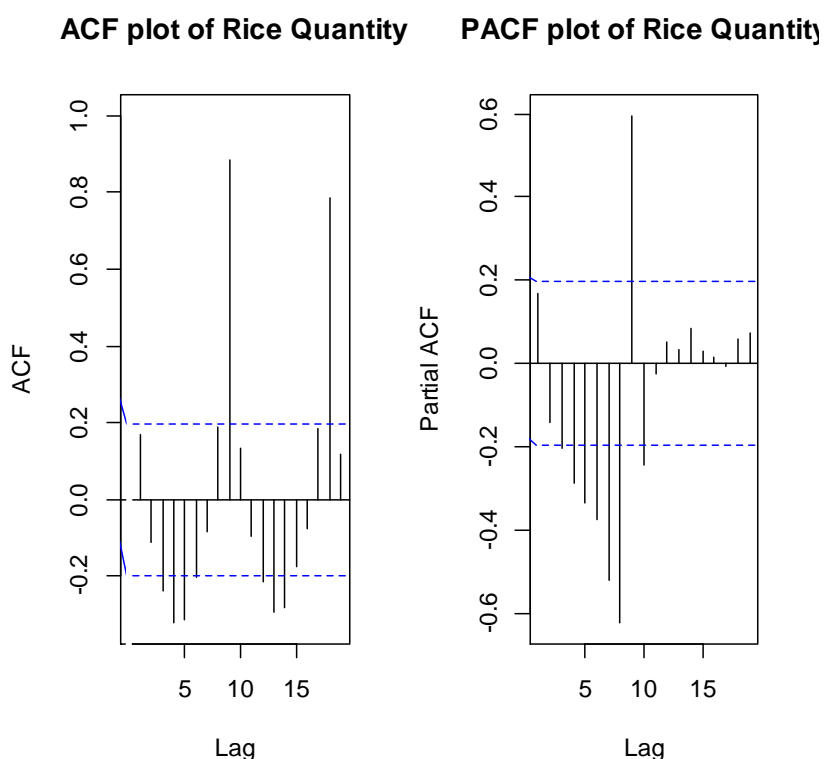
รูปที่ 3.2.16 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวในปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2559 ถึงเดือนเมษายนปี พ.ศ. 2560

การพยากรณ์ปริมาณผลผลิตข้าวในปีแบบรายเดือนในประเทศไทยปี พ.ศ. 2560 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.2.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวในปีแบบรายเดือนตั้งแต่ปี พ.ศ. 2549-2559 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.2.17 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.2.17 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 18 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 18 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(9,9) สามารถแสดงได้ดังตารางที่ 3.2.18 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.2.18 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$p \backslash q$	0	1	2	3	4	5	6	7	8	9	10
0	-	3315.01	3315.02	3312.96	3306.20	3295.70	3280.60	3244.07	NA	NA	NA
1	3314.39	3316.27	3286.64	3280.93	3271.90	3262.04	3247.11	3219.23	NA	NA	NA
2	3302.91	3294.38	3246.75	3242.45	3230.83	3214.96	3246.74	3160.41	NA	NA	NA
3	3292.73	3293.77	3225.20	3245.21	3234.76	3257.33	3199.52	NA	NA	NA	NA
4	3281.00	3276.51	3274.31	3274.36	3276.28	3223.10	3200.24	3137.23	NA	NA	NA
5	3284.31	3284.75	3247.37	3262.87	3212.32	3209.27	3187.92	3127.94	NA	NA	NA
6	3254.83	3256.25	3250.72	3207.26	3192.30	NA	3185.77	3146.87	NA	NA	3063.52
7	3249.32	3250.57	3238.03	3227.48	3230.52	3199.65	3154.98	3104.20	NA	3062.21	3064.16
8	3249.95	3245.94	3232.79	3234.14	3172.17	3173.37	3177.12	3119.30	3073.02	NA	NA
9	3221.19	3222.98	3221.98	3221.61	3222.57	3166.77	3137.81	3110.58	3056.47	3055.86	3057.90
10	3222.76	3222.67	3213.24	3192.18	3176.27	3164.62	3126.80	3125.08	3058.14	3058.52	3059.31
11	3215.45	3218.16	3170.75	3163.37	3188.50	3160.06	3164.50	3096.99	3059.88	3057.81	3059.65
12	3216.03	3218.00	3158.32	3173.99	3166.25	3156.99	3129.70	3095.83	3061.61	3059.65	3061.61
13	3217.95	3211.21	3209.07	3206.61	3208.60	NA	3127.34	3090.47	3063.64	3061.95	3067.51
14	3207.83	3208.36	3187.64	3199.93	3160.11	3139.36	3133.47	3092.57	3065.51	3064.22	3066.30
15	3196.72	3203.62	3211.27	3158.28	3156.36	3140.86	3112.48	3104.64	3067.03	3066.09	3067.89
16	3193.45	3195.08	3185.44	3180.26	3161.74	3140.71	3146.07	3085.45	3070.75	3068.12	3068.96
17	3194.94	3204.47	3174.09	3186.84	3138.65	3143.02	3104.85	3094.31	3072.90	3067.35	NA
18	3171.33	3172.56	3172.92	3173.41	3175.09	3131.79	3106.32	3090.29	3069.00	3070.72	3072.9

หมายเหตุ : NA = not available

จากตารางที่ 3.2.18 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 + 0.0580B^1 + 0.0509B^2 + 0.0460B^3 + 0.0483B^4 + 0.0501B^5 + 0.0463B^6 + 0.0574B^7 + 0.0189B^8 - 0.9332B^9)Y_t$$

$$= 2,772,800.80 + (1 - 0.3986B^1 + 0.0901B^2 - 0.0677B^3 - 0.0760B^4 - 0.0427B^5 - 0.0405B^6 + 0.1734B^7 + 0.1660B^8 - 0.3126B^9)\varepsilon_t$$

ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(9,9) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.2.19 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปีจริง ดังรูปที่ 3.2.20 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2560 ถึงเดือนเมษายนปี พ.ศ. 2561 ดังรูปที่ 3.2.21

ตารางที่ 3.2.19 แสดงปริมาณผลผลิตข้าวนาปีจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปี

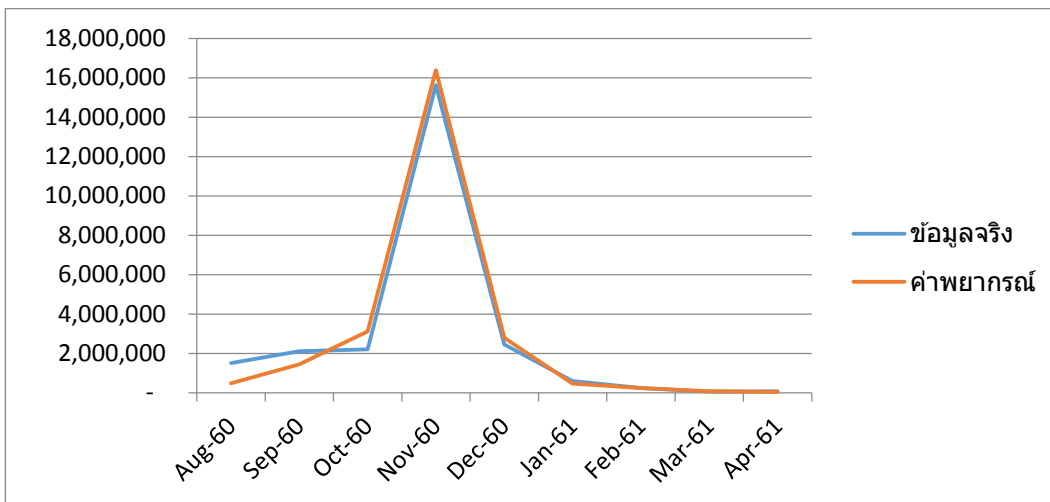
วันที่	ปริมาณผลผลิตข้าวนาปีจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
สิงหาคม 2560	1,519,254	487,726	-1,031,528	-67.90%
กันยายน 2560	2,117,251	1,455,645	-661,606	-31.25%
ตุลาคม 2560	2,216,210	3,129,639	913,429	41.22%
พฤศจิกายน 2560	15,627,757	16,380,229	752,472	4.81%
ธันวาคม 2560	2,451,808	2,794,020	342,212	13.96%
มกราคม 2561	602,030	469,761	-132,269	-21.97%
กุมภาพันธ์ 2561	248,298	249,100	802	0.32%
มีนาคม 2561	72,339	90,330	17,991	24.87%
เมษายน 2561	79,402	57,994	-21,408	-26.96%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

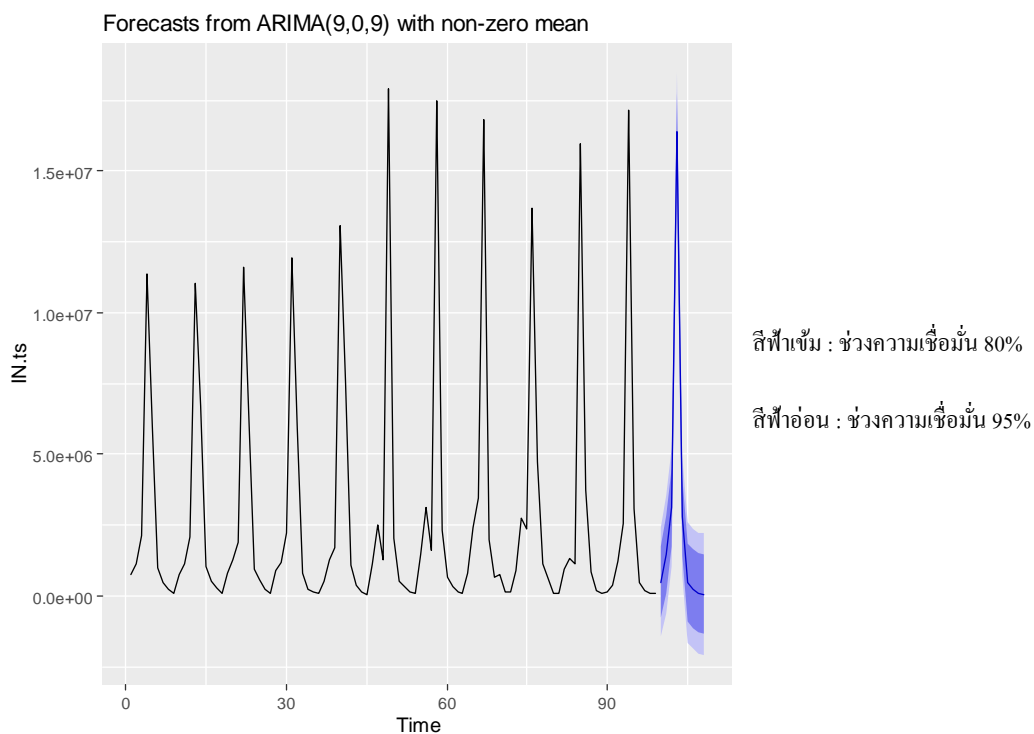
$$RMSE = \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}}$$

$$= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}}$$

$$= 580,969$$



รูปที่ 3.2.20 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวในปีจริง



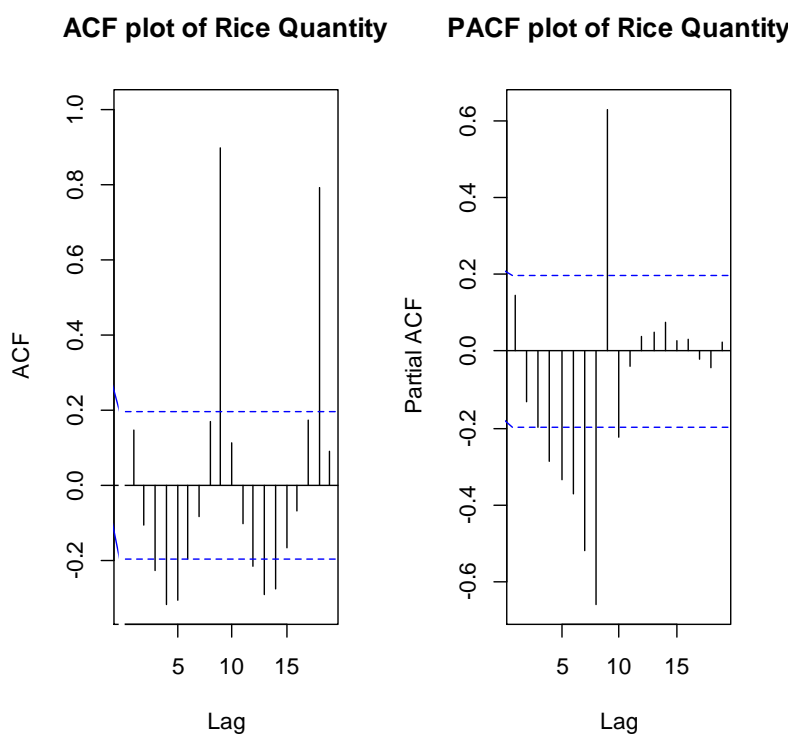
รูปที่ 3.2.21 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวในปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2560 ถึงเดือนเมษายนปี พ.ศ. 2561

การพยากรณ์ปริมาณผลผลิตข้าวแบบรายเดือนในประเทศไทยปี พ.ศ. 2561 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.2.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวแบบรายเดือนตั้งแต่ปี พ.ศ. 2550-2560 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.2.22 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.2.22 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ (-0.1970,0.1970) ตัวสุดท้ายคือ 18 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 18 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ (-0.1970,0.1970) ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(9,1) สามารถแสดงได้ดังตารางที่ 3.2.23 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.2.23 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	3319.65	3319.99	3318.03	3311.25	3300.58	3285.94	3250.26	3129.32	NA	NA
1	3319.19	3321.07	3288.62	3282.40	3274.21	3264.17	3249.62	3220.93	3091.53	3055.48	NA
2	3303.54	3297.28	3249.59	3244.73	3233.43	3219.77	3248.43	3166.74	NA	3056.55	3058.08
3	3295.19	3296.68	3299.05	3249.10	3237.06	3257.91	3204.23	NA	3073.05	3058.32	NA
4	3285.75	3280.05	3275.68	3276.95	3278.19	3222.51	3201.94	3140.64	3066.83	3060.32	NA
5	3287.49	3286.54	3247.33	3248.21	3218.32	3210.16	3187.59	3131.69	NA	NA	NA
6	3260.96	3262.62	3262.87	3223.66	3194.60	3213.78	3185.69	3149.36	NA	NA	3066.16
7	3253.17	3246.01	3244.26	3239.78	3197.62	3198.88	3156.90	3108.07	NA	3065.53	3067.47
8	3251.89	3254.76	3236.87	3236.98	3172.56	3188.93	3177.08	3125.45	3071.93	3065.60	NA
9	3226.20	3224.94	3227.03	3226.03	3207.72	3183.07	3186.40	3165.88	3057.33	3058.56	3060.44
10	3220.51	3212.77	3214.40	3202.80	3186.08	3169.58	3155.34	3115.11	3058.99	NA	3061.78
11	3212.93	3214.60	3177.58	3168.10	3179.86	3172.21	3166.12	3116.29	3061.09	NA	3064.43
12	3213.49	3215.48	3204.79	3207.76	3166.92	3162.98	3141.13	3101.74	3062.78	3064.36	3064.57
13	3215.27	3206.51	3206.01	3201.45	3203.43	3145.04	3148.44	3098.16	3065.51	3066.35	3068.39
14	3205.27	3206.90	3205.50	3205.11	3138.84	3146.93	3144.62	3096.37	3066.16	3066.6	3070.28
15	3197.75	3198.73	3177.26	3178.89	3147.76	3136.53	3122.98	3102.20	3067.16	3068.91	3070.98
16	3197.44	3198.28	3192.92	3194.22	3148.57	3143.66	3125.19	3089.55	3070.27	3070.72	3072.78
17	3197.06	3194.74	3178.60	3189.21	3132.58	3133.68	3128.17	3100.65	3071.83	3071.36	3074.38
18	3168.78	3170.31	3177.70	3171.97	3173.63	3150.47	3153.30	3096.29	3073.44	3073.44	3075.42

หมายเหตุ : NA = not available

จากตารางที่ 3.2.23 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 + 0.1188B^1 + 0.1137B^2 + 0.1136B^3 + 0.1142B^4 + 0.1141B^5 + 0.1095B^6 + 0.1164B^7 + 0.0676B^8 - 0.8536B^9)Y_t = 2,808,669.31 + (1 - 0.3218B^1)\varepsilon_t$$

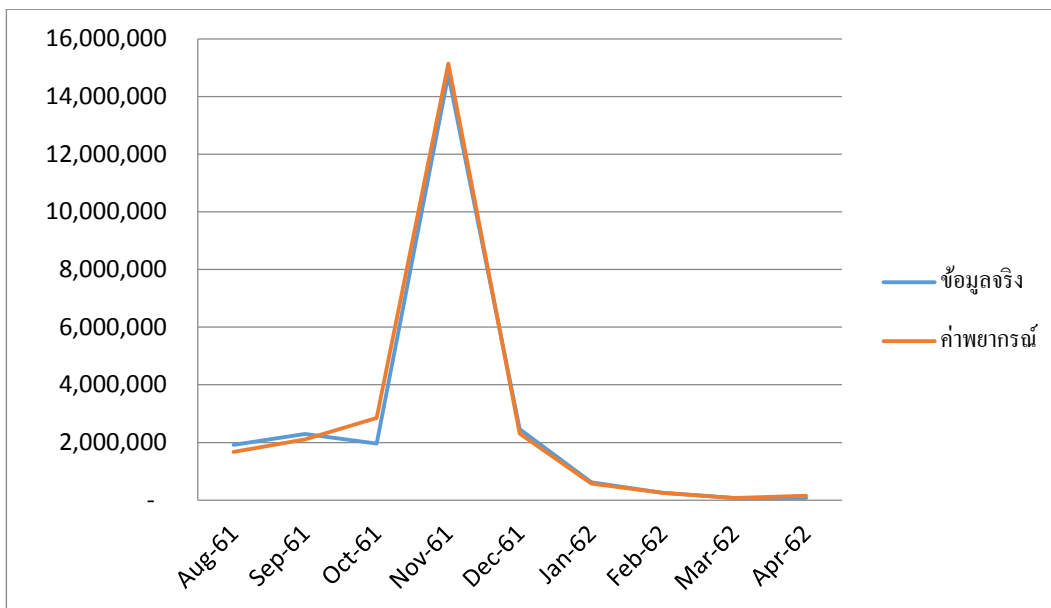
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(9,1) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.2.24 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวในปีจริง ดังรูปที่ 3.2.25 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2561 ถึงเดือนเมษายนปี พ.ศ. 2562 ดังรูปที่ 3.2.26

ตารางที่ 3.2.24 แสดงปริมาณผลผลิตข้าวนาปีจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปี

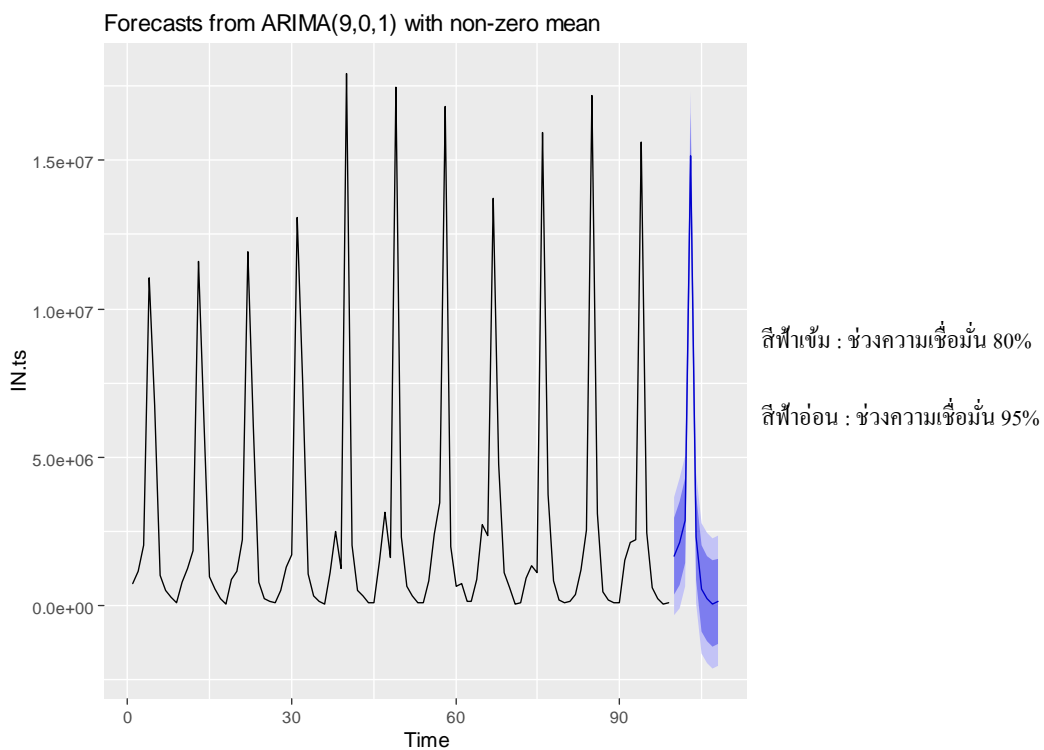
วันที่	ปริมาณผลผลิตข้าวนาปีจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
สิงหาคม 2561	1,914,032	1,668,168	-245,864	-12.85%
กันยายน 2561	2,291,788	2,106,467	-185,321	-8.09%
ตุลาคม 2561	1,957,363	2,847,833	890,470	45.49%
พฤศจิกายน 2561	14,756,199	15,140,934	384,735	2.61%
ธันวาคม 2561	2,458,479	2,302,781	-155,698	-6.33%
มกราคม 2562	623,643	579,057	-44,586	-7.15%
กุมภาพันธ์ 2562	253,087	247,188	-5,899	-2.33%
มีนาคม 2562	71,019	69,697	-1,322	-1.86%
เมษายน 2562	73,135	153,673	80,538	110.12%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned}
 \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\
 &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\
 &= 344,561
 \end{aligned}$$

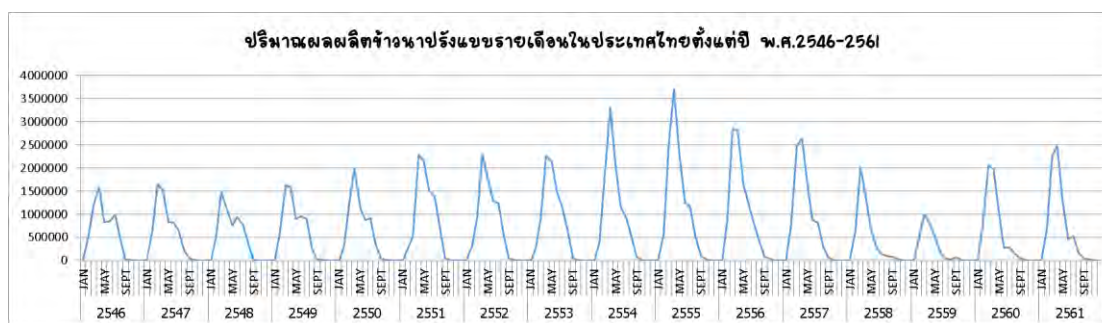


รูปที่ 3.2.25 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวในปีจริง



รูปที่ 3.2.26 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวในปีตั้งแต่เดือนสิงหาคมปี พ.ศ. 2561 ถึงเดือนเมษายนปี พ.ศ. 2562

3.3 การพยากรณ์ปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทย



รูปที่ 3.3.1 กราฟแสดงปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทยตั้งแต่ปี พ.ศ. 2546-2561

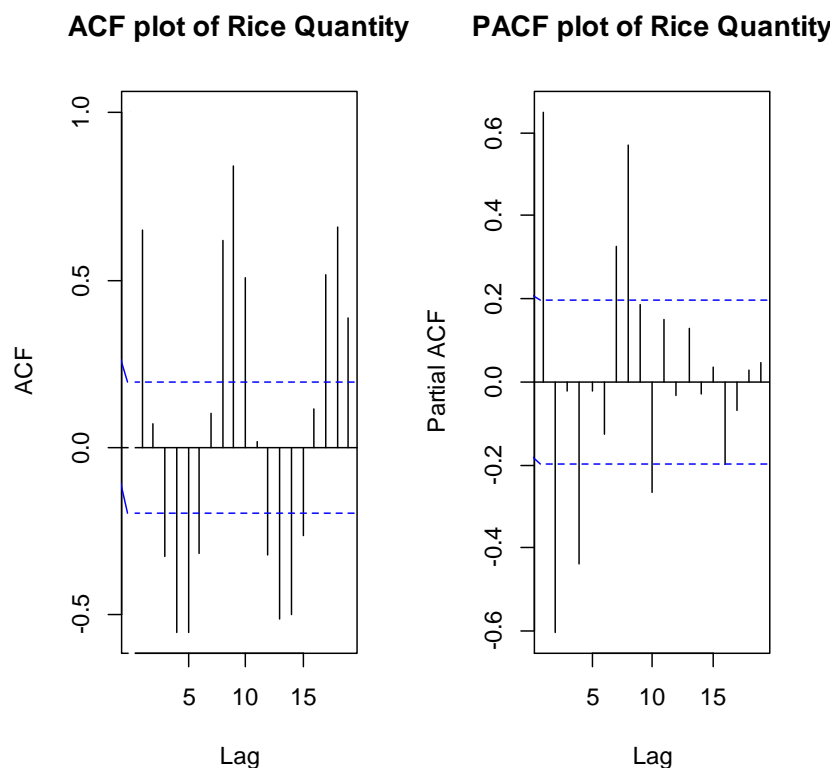
เนื่องจากวิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) เหมาะกับข้อมูลที่มีลักษณะคงที่ (Stationary) และเหมาะกับการใช้ข้อมูลในการพยากรณ์ประมาณ 50 ค่าขึ้นไป ดังนั้น จำนวนข้อมูลที่ใช้ในการพยากรณ์ปริมาณผลผลิตข้าวนาปรังในประเทศไทยจะใช้ข้อมูลแบบรายเดือนจำนวน 9 เดือนต่อปี ได้แก่ เดือนกุมภาพันธ์ มีนาคม เมษายน พฤษภาคม มิถุนายน กรกฎาคม สิงหาคม กันยายน และตุลาคม เนื่องจากข้าวนาปรังจะเริ่มเก็บเกี่ยวตั้งแต่เดือนกุมภาพันธ์จนถึงเดือนตุลาคม ซึ่งรวมเก็บข้อมูลจำนวนทั้งสิ้น 99 เดือน หรือ 11 ปี เพื่อพยากรณ์ปี พ.ศ. 2557 (ตั้งแต่ปี พ.ศ. 2546-2556), ปี พ.ศ. 2558 (ตั้งแต่ปี พ.ศ. 2547-2557), ปี พ.ศ. 2559 (ตั้งแต่ปี พ.ศ. 2548-2558), ปี พ.ศ. 2560 (ตั้งแต่ปี พ.ศ. 2549-2559) และปี พ.ศ. 2561 (ตั้งแต่ปี พ.ศ. 2550-2560)

การพยากรณ์ปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทยปี พ.ศ. 2557 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.3.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปรังแบบรายเดือนตั้งแต่ปี พ.ศ. 2546-2556 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.3.2 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.3.2 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 19 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 19 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(8,9) สามารถแสดงได้ดังตารางที่ 3.3.3 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.3.3 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	2927.35	2882.94	2884.93	2865.24	2867.21	2867.36	2857.07	2797.15	NA	NA
1	2891.60	2883.05	2882.21	2873.69	2867.17	2863.02	2863.89	2865.64	NA	NA	NA
2	2884.00	2884.73	2873.52	2867.36	2868.46	2864.16	2865.86	2850.55	NA	NA	NA
3	2881.62	2884.94	NA	2842.64	2863.88	2865.44	2860.24	NA	NA	NA	NA
4	2882.90	2879.53	2836.76	2874.18	2867.28	2859.91	2861.88	2842.15	NA	NA	NA
5	2873.76	2875.74	2874.60	2874.56	2833.76	NA	NA	2821.93	NA	NA	NA
6	2872.21	2866.62	NA	2840.51	2866.68	2834.80	NA	2822.50	NA	NA	NA
7	2850.27	2847.80	2849.79	2866.23	NA	NA	NA	2803.57	NA	NA	NA
8	2879.85	2856.50	2819.71	NA	NA	NA	2814.20	2790.92	NA	NA	NA
9	2853.53	2835.41	2819.86	2814.02	2802.07	2802.88	2804.88	2792.50	2784.25	NA	NA
10	2827.89	2829.46	2830.25	2823.67	2802.84	2804.84	2807.83	2807.00	2787.37	NA	NA
11	2829.52	2831.80	2807.27	2811.00	2804.84	2806.52	2808.84	2804.60	2787.50	NA	NA
12	2830.81	2830.34	2802.42	2804.94	2806.44	2802.60	2795.11	2808.09	2795.71	NA	NA
13	2833.18	2831.16	2804.38	2830.48	2828.79	2805.14	NA	2791.99	2790.73	NA	2792.51
14	2830.26	2832.11	2828.43	2833.06	2802.99	2805.68	2805.15	2807.74	2792.65	2792.53	NA
15	2832.35	2825.73	2815.78	2828.64	2804.73	2809.41	2811.09	2807.52	2797.52	2794.20	2796.35
16	2823.91	2826.58	2813.90	2823.90	2806.78	2806.58	2807.74	2809.50	2796.84	2795.75	NA
17	2825.81	2829.75	2805.16	2808.41	2810.28	2801.08	2803.75	2801.12	2794.27	2794.44	2796.41
18	2823.63	2818.09	2810.74	2806.37	2800.03	2804.40	2798.54	2802.47	2795.72	2796.42	2796.2
19	2809.46	2811.33	2813.59	2814.54	2802.81	2804.17	2802.32	2802.43	2795.91	2798.08	2795.12

หมายเหตุ : NA = not available

จากตารางที่ 3.3.3 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 - 0.9374B^1 + 0.8317B^2 - 0.6496B^3 + 0.4676B^4 - 0.1887B^5 - 0.0333B^6 + 0.3113B^7 - 0.6789B^8)Y_t \\ = 954,275.2 + (1 - 0.1743B^1 + 0.1032B^2 - 0.0304B^3 - 0.0657B^4 + 0.1773B^5 - 0.1943B^6 + 0.2465B^7 \\ - 0.3442B^8 + 0.2797B^9)\varepsilon_t$$

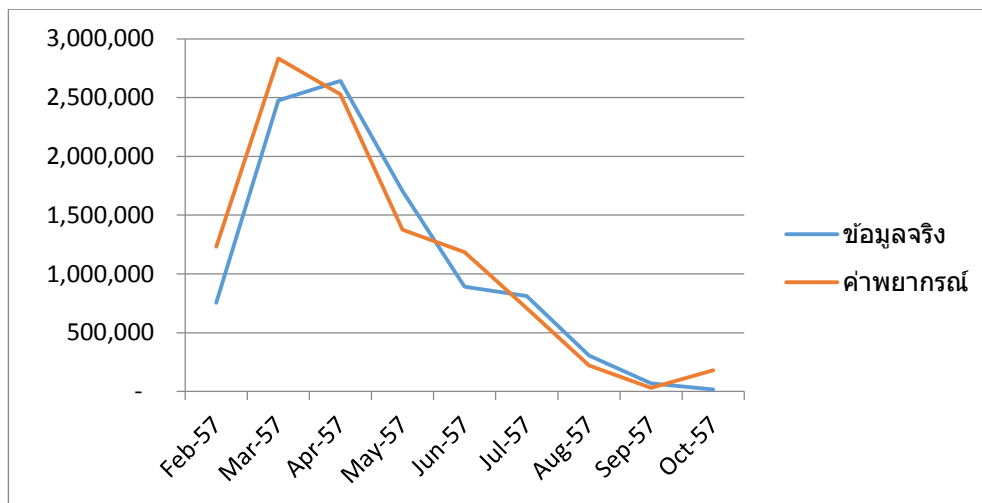
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(8,9) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.3.4 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง ดังรูปที่ 3.3.5 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2557 ถึงเดือนตุลาคมปี พ.ศ. 2557 ดังรูปที่ 3.3.6

ตารางที่ 3.3.4 แสดงปริมาณผลผลิตข้าวนาปรังจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปรัง

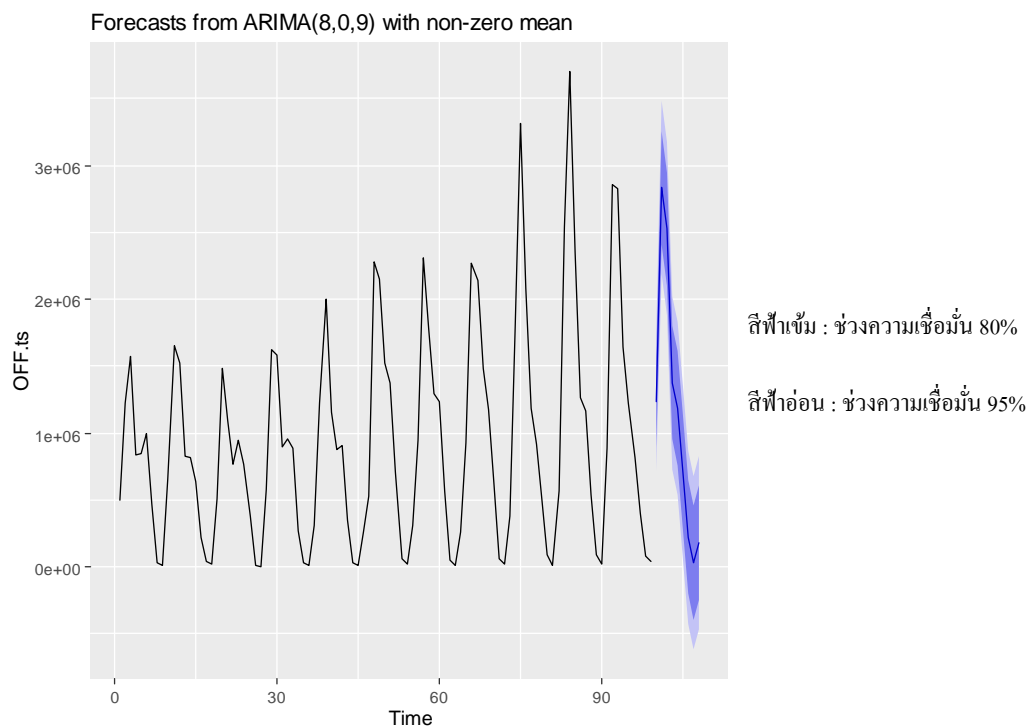
วันที่	ปริมาณผลผลิตข้าวนาปรังจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
กุมภาพันธ์ 2557	753,688	1,231,668	477,980	63.42%
มีนาคม 2557	2,476,343	2,833,657	357,314	14.43%
เมษายน 2557	2,642,566	2,527,723	-114,843	-4.35%
พฤษภาคม 2557	1,704,262	1,376,945	-327,317	-19.21%
มิถุนายน 2557	890,643	1,185,925	295,282	33.15%
กรกฎาคม 2557	811,486	707,856	-103,630	-12.77%
สิงหาคม 2557	306,648	221,755	-84,893	-27.68%
กันยายน 2557	69,657	30,532	-39,125	-56.17%
ตุลาคม 2557	16,800	180,825	164,025	976.34%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\ = \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\ = 260,349$$



รูปที่ 3.3.5 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง



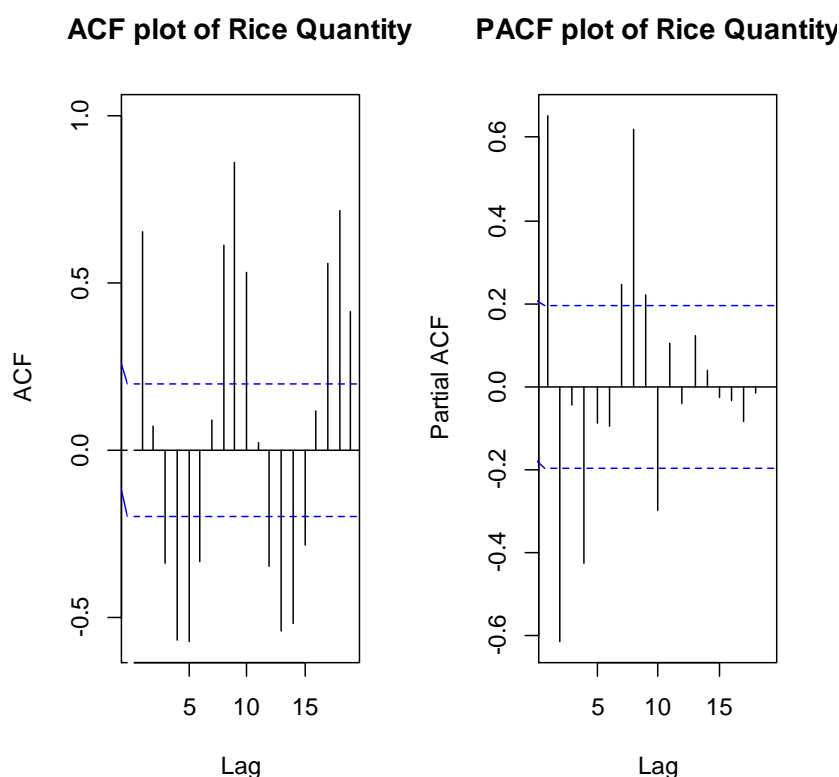
รูปที่ 3.3.6 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2557 ถึงเดือนตุลาคมปี พ.ศ. 2557

การพยากรณ์ปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทยปี พ.ศ. 2558 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.3.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปรังแบบรายเดือนตั้งแต่ปี พ.ศ. 2547-2557 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.3.7 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.3.7 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 19 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 19 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,2) สามารถแสดงได้ดังตารางที่ 3.3.8 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.3.8 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	2934.37	2887.49	2889.46	2869.79	2871.43	2871.68	2865.06	2800.61	NA	NA
1	2904.01	2892.64	2888.26	2883.77	2871.29	2872.25	2872.74	2830.92	NA	NA	NA
2	2889.83	2891.61	2878.99	2876.68	2873.29	2873.07	2874.66	2808.75	NA	NA	2776.09
3	2891.60	2889.29	NA	NA	2867.92	2869.88	2841.95	2847.59	NA	NA	2777.84
4	2892.13	2889.90	2887.16	2884.11	NA	2842.86	2844.98	NA	2802.04	NA	2779.84
5	2885.21	2887.08	2882.65	2846.75	NA	NA	2830.87	NA	NA	NA	2781.84
6	2881.09	2877.36	2872.82	2847.90	NA	2843.24	2834.07	2787.82	NA	NA	2783.84
7	2875.37	2856.79	2857.48	2859.07	2822.62	2839.80	2836.92	2810.07	2790.93	NA	NA
8	2889.89	2864.76	2825.75	2824.43	2820.00	2821.96	NA	2791.74	NA	NA	2787.38
9	2861.27	2847.21	2841.59	2839.94	2811.49	2812.93	2803.44	NA	2790.40	NA	2788.33
10	2840.13	2839.21	2816.14	2840.99	2812.93	2815.43	2813.02	2806.10	2791.10	NA	NA
11	2837.10	2839.10	2814.96	2842.95	2814.93	2816.92	2801.45	2804.02	2793.55	NA	2791.42
12	2839.10	2840.76	2810.08	2845.18	2814.46	2813.92	2815.59	2806.18	2794.82	NA	2795.37
13	2840.64	2838.90	2809.71	2839.04	2802.73	2816.96	2801.64	2803.89	2797.00	2795.31	2794.93
14	2841.84	2843.50	2809.98	2812.97	2808.51	NA	2804.18	2814.24	2800.51	2794.42	2796.57
15	2841.10	2843.09	2812.89	2810.58	2811.02	2815.02	2816.89	2813.66	2801.00	2795.95	2797.43
16	2841.63	2834.31	2835.07	2842.28	2815.85	2815.99	2816.64	2813.55	2797.24	2796.53	2800.32
17	2835.15	2833.33	2812.74	2836.84	2813.48	2810.10	2816.06	2813.82	2797.67	2798.31	NA
18	2830.49	2823.32	2820.90	2811.86	2807.90	2811.31	2807.06	2801.50	2799.35	2800.45	2802.36
19	2820.08	2822.32	2823.12	2825.03	2810.47	2809.15	2810.08	2803.72	2802.83	2802.81	2803.78

หมายเหตุ : NA = not available

จากตารางที่ 3.3.8 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 + 0.0133B^1 + 0.0182B^2 + 0.0490B^3 - 0.0145B^4 + 0.0697B^5 - 0.0169B^6 + 0.0578B^7 - 0.1301B^8 - 0.7857B^9 - 0.0459B^{10})Y_t = 939,640.3 + (1 + 0.7097B^1 + 0.0047B^2)\varepsilon_t$$

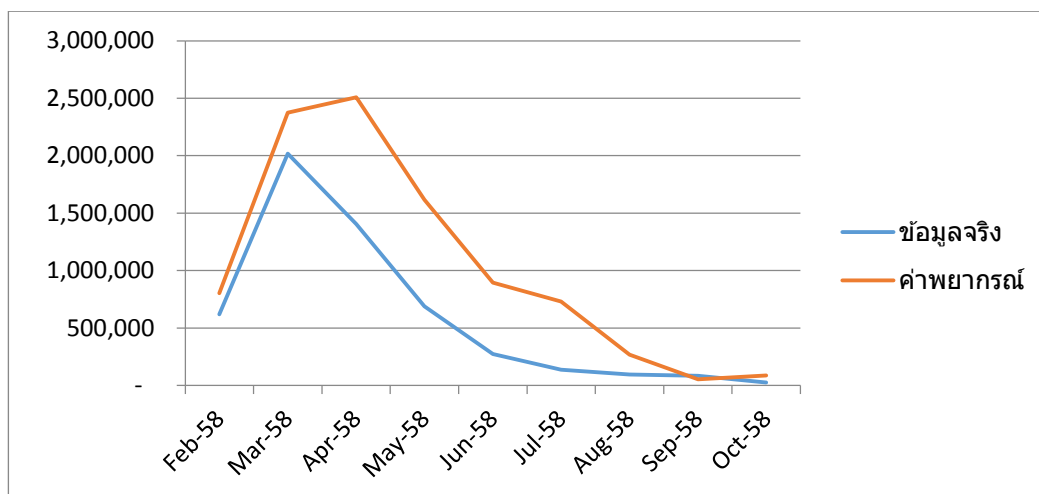
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,2) ไปหาค่าพยากรณ์และทำการเปรียบเทียบกับข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.3.9 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง ดังรูปที่ 3.3.10 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2558 ถึงเดือนตุลาคมปี พ.ศ. 2558 ดังรูปที่ 3.3.11

ตารางที่ 3.3.9 แสดงปริมาณผลผลิตข้าวนาปรังจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปรัง

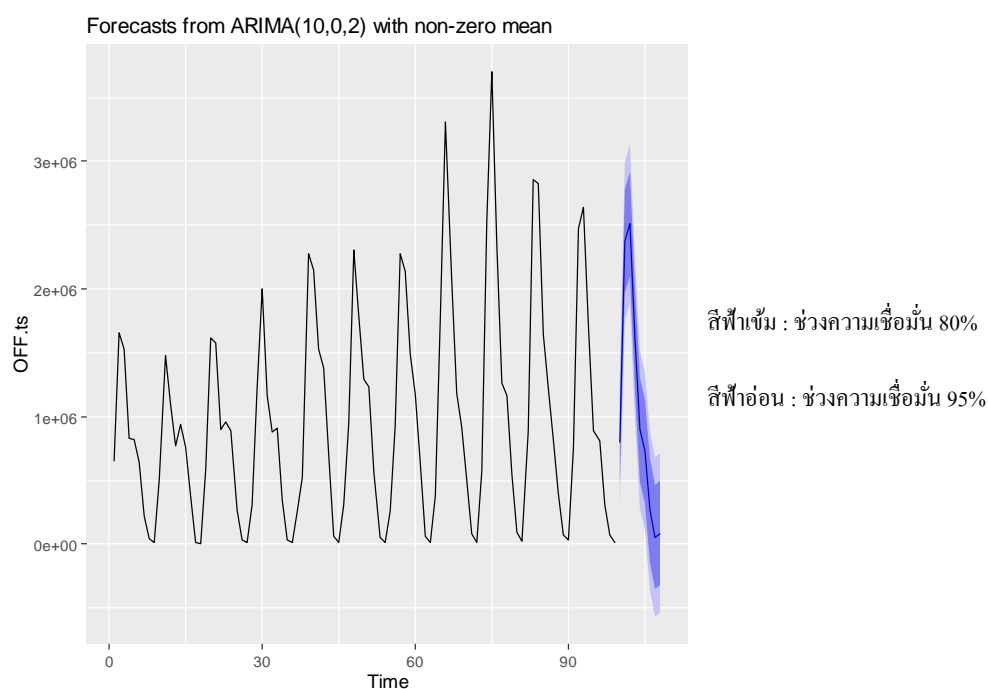
วันที่	ปริมาณผลผลิตข้าวนาปรังจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
กุมภาพันธ์ 2558	618,068	804,376	186,308	30.14%
มีนาคม 2558	2,016,508	2,375,149	358,641	17.79%
เมษายน 2558	1,405,235	2,508,338	1,103,103	78.50%
พฤษภาคม 2558	689,428	1,615,721	926,293	134.36%
มิถุนายน 2558	274,251	895,922	621,671	226.68%
กรกฎาคม 2558	138,240	730,716	592,476	428.59%
สิงหาคม 2558	95,056	267,504	172,448	181.42%
กันยายน 2558	83,932	54,514	-29,418	-35.05%
ตุลาคม 2558	26,197	87,137	60,940	232.62%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned}
 \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\
 &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\
 &= 578,313
 \end{aligned}$$



รูปที่ 3.3.10 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง



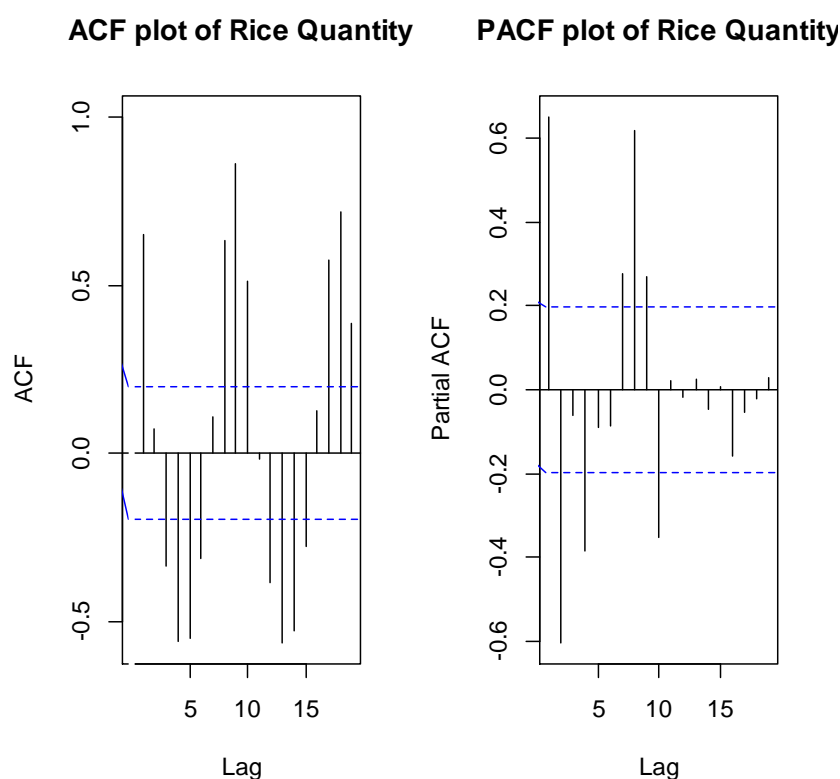
รูปที่ 3.3.11 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2558 ถึงเดือนตุลาคมปี พ.ศ. 2558

การพยากรณ์ปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทยปี พ.ศ. 2559 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.3.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปรังแบบรายเดือนตั้งแต่ปี พ.ศ. 2548-2558 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.3.12 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.3.12 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 19 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 19 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,2) สามารถแสดงได้ดังตารางที่ 3.3.13 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.3.13 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	2937.16	2891.95	2893.91	2877.17	2878.68	2879.10	2872.12	2816.58	2803.49	2790.32
1	2906.44	2897.01	2893.47	2886.98	2878.55	2877.44	2877.67	2879.32	2811.75	2788.64	2790.09
2	2895.96	2897.69	2885.08	2881.05	2882.47	2878.17	2879.61	2868.76	2801.82	2790.32	2786.94
3	2895.25	2896.41	2848.64	2850.48	2874.38	2846.11	2871.19	2855.92	2802.76	2790.23	2791.58
4	2896.84	2893.21	2890.34	2888.44	2853.55	2871.47	2849.39	2807.95	2815.26	2791.72	2792.84
5	2886.64	2888.54	2888.21	2853.14	2879.70	2848.57	2851.23	2835.92	2803.81	2793.65	2795.61
6	2888.40	2880.12	2876.26	2877.54	2882.59	2848.94	2847.15	2848.37	2807.25	2795.63	2797.60
7	2868.17	2866.73	2865.43	2867.01	2829.19	2830.42	2828.97	2831.15	2805.39	2797.62	2798.78
8	2893.79	2869.21	2856.59	2838.00	2836.20	2841.64	2824.93	2827.02	2803.25	2797.12	2800.72
9	2864.03	2846.97	2845.61	2843.30	2817.53	2816.56	2816.42	2815.79	2801.74	2798.38	2800.07
10	2843.47	2845.25	2845.31	2845.27	2848.54	2818.55	2822.21	2820.57	2804.22	2803.44	2800.92
11	2845.28	2847.22	2845.10	2848.36	2842.08	2819.65	2810.88	2811.81	2806.07	2800.94	2803.24
12	2844.87	2848.32	2848.62	2850.33	2850.54	2820.45	2823.98	2816.57	2806.87	2802.02	2806.63
13	2846.69	2846.27	2825.44	2844.59	2845.79	2848.38	2849.29	2850.09	2809.82	2803.07	2806.55
14	2850.99	2847.05	2849.93	2841.24	2830.45	2821.65	2823.89	2821.22	2811.02	2807.25	NA
15	2847.76	2842.98	2846.63	2850.14	2818.91	2819.45	2821.69	2825.45	2809.88	2803.53	2805.07
16	2837.63	2841.65	2843.61	2846.86	2848.07	2825.21	2824.15	2825.36	2809.17	2803.13	2806.38
17	2839.58	2840.49	2818.75	2844.35	2816.82	2817.46	2804.19	2822.69	2810.28	2809.39	2811.27
18	2834.85	2823.51	2822.71	2824.23	2809.82	2811.26	2804.35	2813.81	2805.94	2807.33	NA
19	2818.63	2820.33	2822.27	2824.27	2825.57	2813.81	2806.47	2812.64	2811.71	2806.54	2804.40

หมายเหตุ : NA = not available

จากตารางที่ 3.3.13 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 - 0.9808B^1 + 0.0088B^2 + 0.0315B^3 - 0.0332B^4 + 0.0385B^5 - 0.0301B^6 + 0.0565B^7 - 0.1629B^8 - 0.6880B^9 + 0.8044B^{10})Y_t = 957,029.02 + (1 - 0.2669B^1 - 0.7298B^2)\varepsilon_t$$

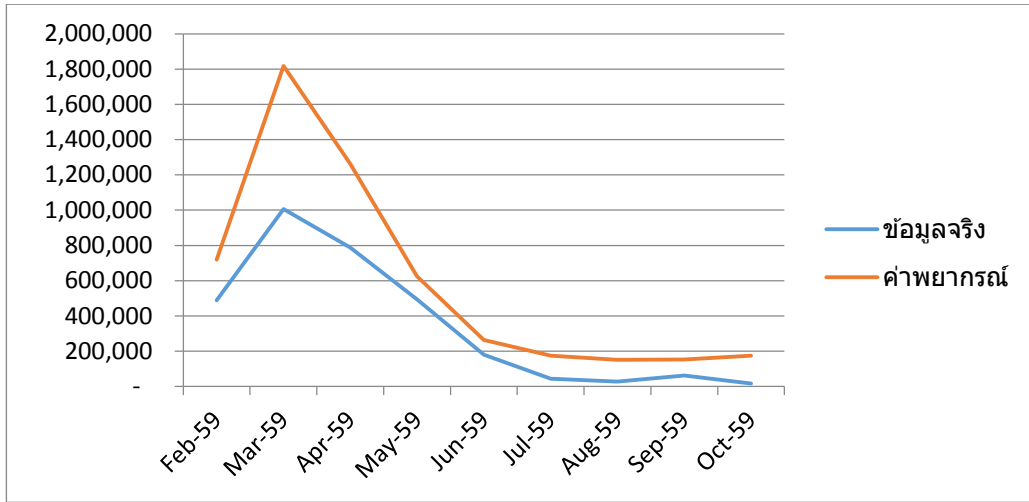
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,2) ไปหาค่าพยากรณ์และทำการเปรียบเทียบข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.3.14 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบปริมาณผลผลิตข้าวนาปรังจริง ดังรูปที่ 3.3.15 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2559 ถึงเดือนตุลาคมปี พ.ศ. 2559 ดังรูปที่ 3.3.16

ตารางที่ 3.3.14 แสดงปริมาณผลผลิตข้าวนาปรังจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปรัง

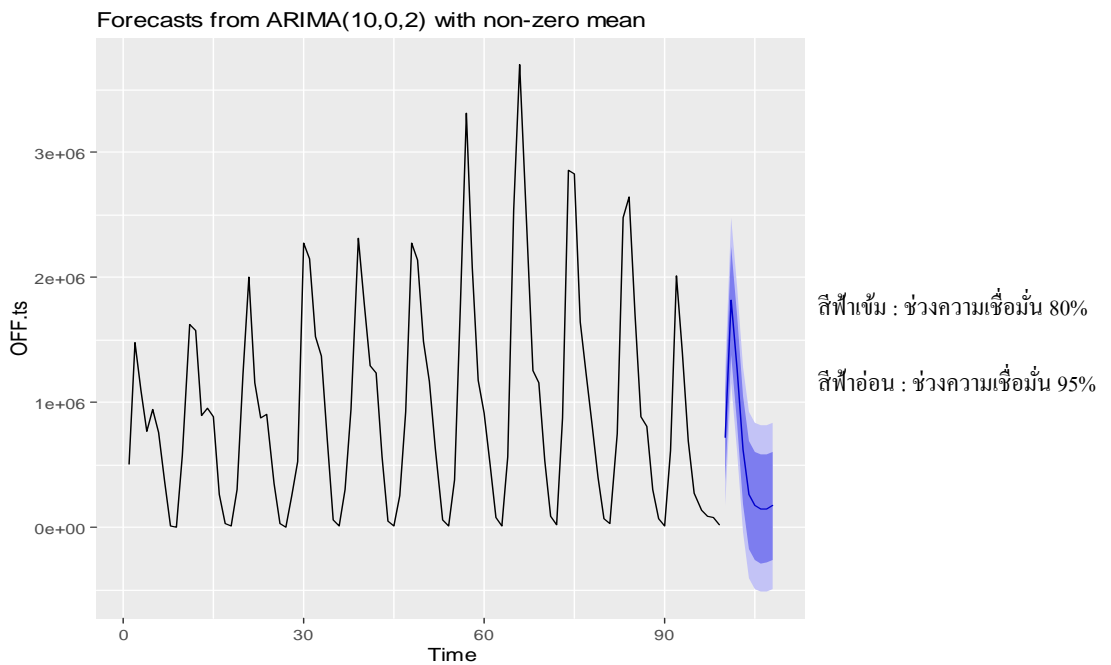
วันที่	ปริมาณผลผลิตข้าวนาปรังจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
กุมภาพันธ์ 2559	488,122	719,680	231,558	47.44%
มีนาคม 2559	1,005,526	1,817,573	812,047	80.76%
เมษายน 2559	788,791	1,263,112	474,321	60.13%
พฤษภาคม 2559	494,716	624,325	129,609	26.20%
มิถุนายน 2559	180,949	263,423	82,474	45.58%
กรกฎาคม 2559	44,121	174,501	130,380	295.51%
สิงหาคม 2559	27,478	151,886	124,408	452.75%
กันยายน 2559	62,944	153,594	90,650	144.02%
ตุลาคม 2559	16,821	174,353	157,532	936.52%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\ &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\ &= 337,825 \end{aligned}$$



รูปที่ 3.3.15 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง



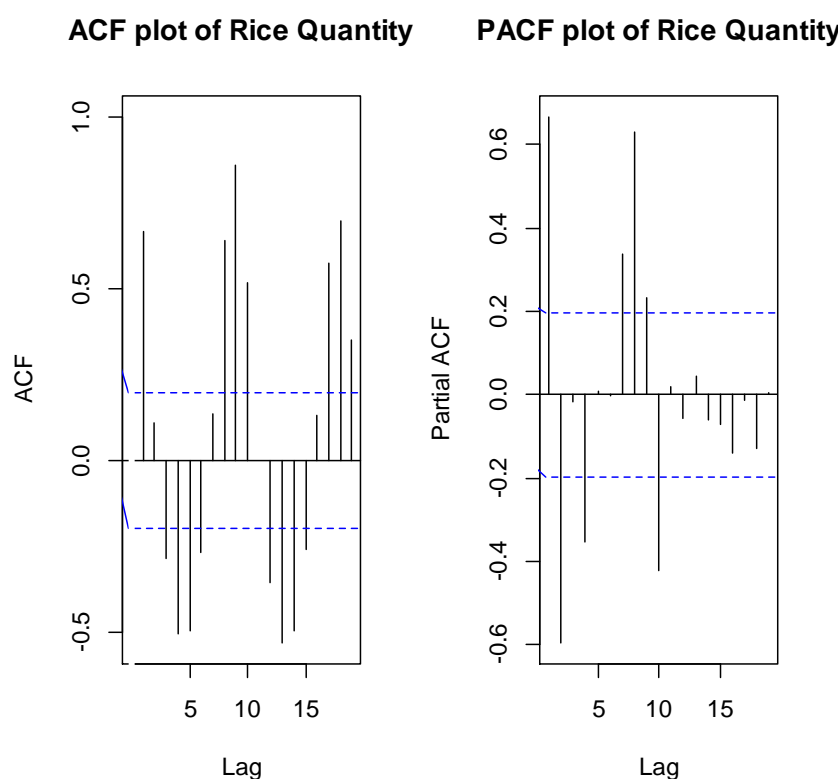
รูปที่ 3.3.16 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2559 ถึงเดือนตุลาคมปี พ.ศ. 2559

การพยากรณ์ปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทยปี พ.ศ. 2560 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.3.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปรังแบบรายเดือนตั้งแต่ปี พ.ศ. 2549-2559 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.3.17 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.3.17 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 19 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 19 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,3) สามารถแสดงได้ดังตารางที่ 3.3.18 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.3.18 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	2935.75	2892.09	2894.08	2881.68	2883.66	2885.61	2873.67	2817.86	NA	2794.45
1	2906.91	2895.88	2892.72	2886.88	2883.66	2881.8	2883.47	2885.15	NA	NA	2793.21
2	2894.70	2896.39	2887.53	2884.24	2886.02	2883.61	2879.73	2867.92	NA	NA	NA
3	2893.83	2896.99	2854.3	2856.11	2877.10	2878.36	2876.26	2870.62	2806.61	2793.83	2788.72
4	2895.69	2894.06	2891.93	2890.10	2882.77	2854.20	2854.32	2872.33	2818.08	2795.49	NA
5	2888.2	2890.05	2888.78	2888.06	2880.87	2856.21	2850.69	2838.02	2807.19	2797.46	NA
6	2888.37	2878.54	2866.59	2864.46	2851.10	2855.79	2857.24	2860.97	2812.71	2799.35	NA
7	2870.86	2866.63	2881.63	2865.40	2853.90	2849.13	2836.11	2830.20	2806.10	2800.12	2795.96
8	2866.72	2868.54	2838.65	2839.24	2837.11	2830.67	2832.38	2822.77	2806.11	2798.06	2799.66
9	2864.84	2847.31	2846.93	2844.23	2823.99	2826.87	2828.87	2825.00	2803.83	2800.02	2796.13
10	2843.40	2843.82	2845.15	2845.30	2822.22	2820.00	2822.80	2827.91	NA	2802.32	2799.98
11	2843.28	2844.60	2846.94	2845.53	2823.33	2826.54	2822.39	2819.10	2808.65	2802.66	2806.28
12	2844.13	2846.12	2844.45	2848.20	2826.78	2849.88	2826.70	2817.83	2809.29	2804.23	2806.65
13	2846.11	2846.29	2841.91	2838.19	2839.66	2830.61	2814.03	2823.48	2810.88	2804.92	2802.65
14	2847.40	2844.91	2847.06	2839.32	2841.36	2820.85	2846.16	2820.39	2814.39	2806.08	2809.73
15	2843.00	2846.61	2841.12	2841.46	2839.41	2821.29	2825.98	2824.63	NA	2805.77	2806.79
16	2835.19	2844.65	2837.98	2846.35	2838.28	2828.26	2833.64	2831.28	2814.57	2807.25	2810.72
17	2836.90	2836.00	2818.97	2835.73	2840.89	2819.48	2816.08	2819.09	2813.41	2810.65	2806.88
18	2833.57	2823.09	2820.20	2821.50	2812.06	2812.56	2818.94	2809.4	2804.90	2812.09	2808.51
19	2820.06	2821.64	2819.25	2821.65	2813.89	2814.36	2811.51	2811.16	2806.57	2806.21	2806.11

หมายเหตุ : NA = not available

จากตารางที่ 3.3.18 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 - 1.0133B^1 - 0.0030B^2 + 0.0866B^3 - 0.0744B^4 + 0.0794B^5 - 0.0730B^6 + 0.0677B^7 - 0.1831B^8 - 0.6110B^9 + 0.7692B^{10})Y_t = 946,770.04 + (1 - 0.2808B^1 - 0.7609B^2 + 0.0416B^3)\varepsilon_t$$

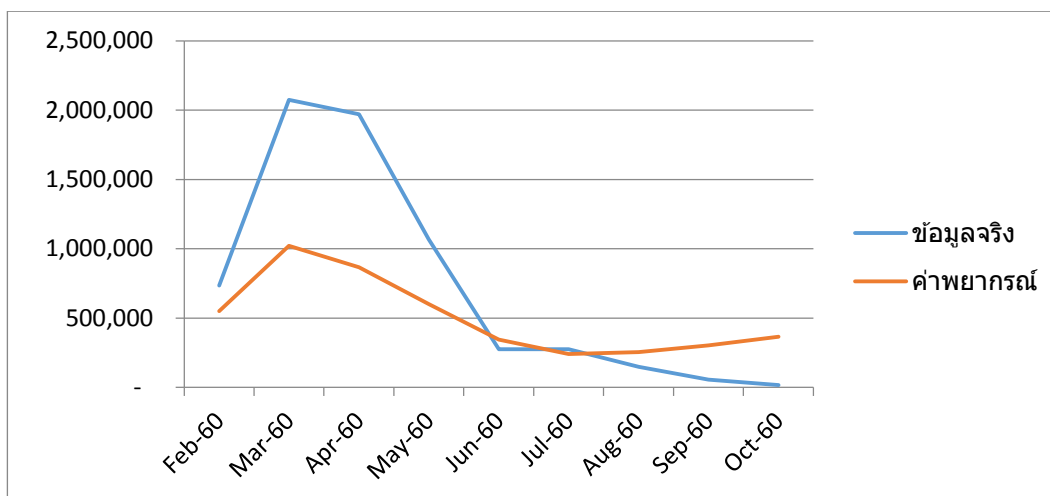
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,3) ไปหาค่าพยากรณ์และทำการเปรียบเทียบข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.3.19 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบปริมาณผลผลิตข้าวนาปรังจริง ดังรูปที่ 3.3.20 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2560 ถึงเดือนตุลาคมปี พ.ศ. 2560 ดังรูปที่ 3.3.21

ตารางที่ 3.2.19 แสดงปริมาณผลผลิตข้าวนาปรังจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปรัง

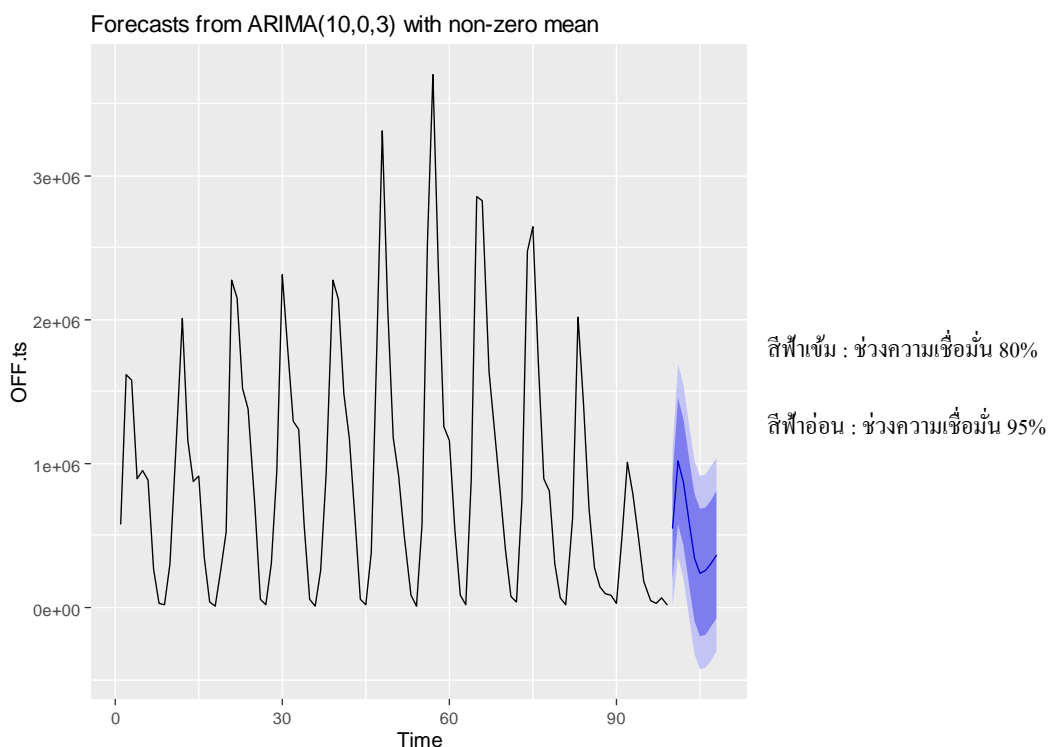
วันที่	ปริมาณผลผลิตข้าวนาปรังจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
กุมภาพันธ์ 2560	734,229	551,013	-183,216	-24.95%
มีนาคม 2560	2,075,034	1,022,045	-1,052,989	-50.75%
เมษายน 2560	1,970,519	867,383	-1,103,136	-55.98%
พฤษภาคม 2560	1,067,156	600,876	-466,280	-43.69%
มิถุนายน 2560	275,139	346,137	70,998	25.80%
กรกฎาคม 2560	275,881	242,176	-33,705	-12.22%
สิงหาคม 2560	148,382	253,963	105,581	71.15%
กันยายน 2560	57,644	304,367	246,723	428.01%
ตุลาคม 2560	16,861	365,843	348,982	2069.76%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\ &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\ &= 555,444 \end{aligned}$$



รูปที่ 3.3.20 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง



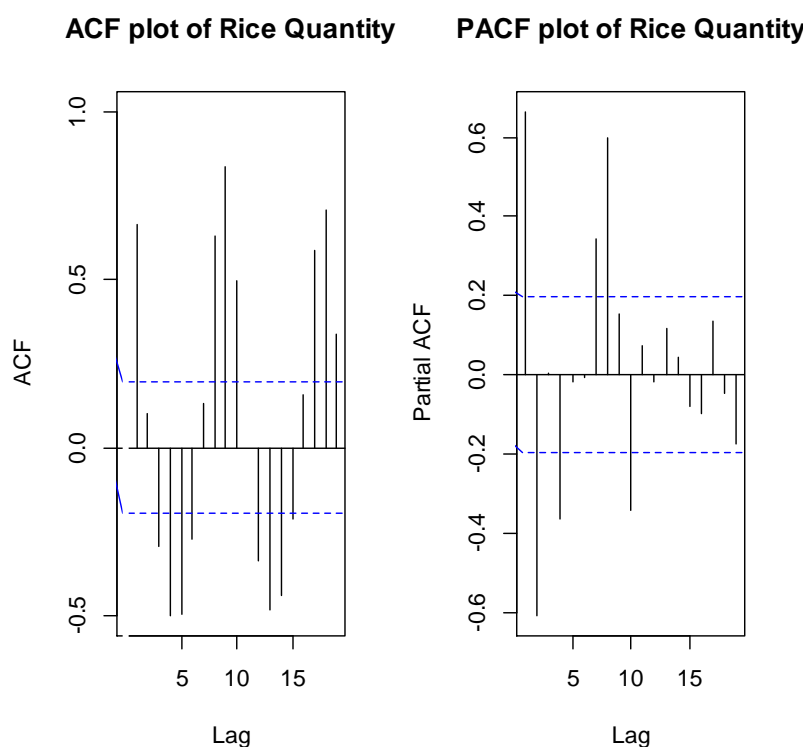
รูปที่ 3.3.21 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2560 ถึงเดือนตุลาคมปี พ.ศ. 2560

การพยากรณ์ปริมาณผลผลิตข้าวนาปรังแบบรายเดือนในประเทศไทยปี พ.ศ. 2561 มีขั้นตอนดังนี้

ขั้นตอนที่ 1 พิจารณารูปที่ 3.2.1 จะเห็นว่าข้อมูลปริมาณผลผลิตข้าวนาปรังแบบรายเดือนตั้งแต่ปี พ.ศ. 2550-2560 มีลักษณะของกราฟขึ้นลงอย่างมีนัยสำคัญทุกปี กล่าวคือข้อมูลมีส่วนประกอบของความแปรผันของฤดูกาลเป็นส่วนประกอบ

โดยในงานวิจัยนี้เราจะสมมติให้ข้อมูลมีลักษณะคงที่ (Stationary) แล้วทำขั้นตอนต่อไปซึ่งจะทำการหาค่า p และ q ที่เป็นไปได้ทั้งหมด

ขั้นตอนที่ 2 หาค่า p และ q ที่เป็นไปได้ โดยพิจารณาจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) ซึ่งจะระบุค่า q ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k (เส้นประสีน้ำเงิน) และจากลักษณะของกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) ซึ่งจะระบุค่า p ที่เป็นไปได้ทั้งหมดในตัวแบบ นั่นคือ ค่า lag (ช่วงเวลาของค่าสังเกตที่อยู่ห่างกัน) หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} (เส้นประสีน้ำเงิน)



รูปที่ 3.3.22 กราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF)

จากรูปที่ 3.3.22 พบว่ากราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_k คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 19 ดังนั้น อันดับ q จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 19 เช่นเดียวกับกราฟแสดงค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) มีค่า lag หรือแท่งที่ออกนอกช่วงความเชื่อมั่น 95% ของ r_{kk} คือ $(-0.1970, 0.1970)$ ตัวสุดท้ายคือ 10 ดังนั้น อันดับ p จะมีค่าที่เป็นไปได้ทั้งหมดตั้งแต่ 0 ถึง 10

ขั้นตอนที่ 3 เลือกตัวแบบการพยากรณ์ที่เหมาะสม ARMA(p, q) เมื่อเปรียบเทียบกับตัวแบบต่าง ๆ ที่เป็นไปได้ทั้งหมดจากโปรแกรม R โดยเลือกตัวแบบที่ให้ค่า Akaike Information Criterion (AIC) ต่ำที่สุด นั่นคือ ARMA(10,2) สามารถแสดงได้ดังตารางที่ 3.3.23 หลังจากนั้นนำตัวแบบที่เหมาะสมดังกล่าวมาประมาณค่าพารามิเตอร์

ตารางที่ 3.3.23 แสดงค่า AIC ของตัวแบบ ARMA(p, q) ที่เป็นไปได้ทั้งหมด

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10
0	-	2939.22	2893.16	2895.14	2883.46	2885.42	2887.37	2874.68	2823.50	2819.04	2802.19
1	2910.69	2898.62	2894.09	2888.89	2885.42	2887.42	2885.26	2887.15	2877.23	2801.47	2799.26
2	2896.55	2898.37	2888.89	2886.84	2859.04	2885.46	2887.25	2871.93	2812.68	2800.27	2795.90
3	2896.57	2898.56	2855.72	2863.92	2880.05	2881.36	2868.38	2872.42	2814.43	2801.38	NA
4	2898.35	2896.28	2894.31	2894.48	2887.2	2857.20	2861.11	2870.46	2820.89	2802.20	NA
5	2890.65	2892.65	2893.63	2890.13	2882.28	2851.83	2854.77	2842.72	2810.23	2804.02	NA
6	2892.65	2885.38	2872.87	2873.03	2853.00	2850.60	2852.32	2845.76	2806.17	2804.61	NA
7	2877.56	2874.06	2873.24	2875.22	2862.69	2844.97	2830.10	2837.37	2804.94	2807.03	NA
8	2868.47	2870.13	2849.73	2851.44	2847.76	2831.94	2815.40	2827.39	2806.46	NA	2799.09
9	2865.14	2852.05	2850.72	2850.79	2834.32	2825.86	2834.72	2820.36	2821.79	NA	2800.56
10	2850.54	2851.02	2847.15	2854.71	2827.46	2827.89	2838.13	2820.82	2808.30	2800.40	NA
11	2850.90	2852.86	2833.40	2851.11	2827.25	2828.47	2837.44	2821.94	2806.03	NA	2804.54
12	2853.11	2854.13	2856.05	2835.51	2844.32	2831.81	2826.37	2834.81	2807.93	2803.66	2804.74
13	2853.81	2855.70	2852.08	2853.56	NA	2828.03	2821.08	2827.26	2809.70	2804.23	2806.08
14	2854.31	2853.57	2834.53	2855.39	2824.61	NA	2825.13	2830.13	2811.62	2813.70	2803.57
15	2846.03	2850.99	2851.19	2848.34	2855.76	2828.51	2831.49	2830.37	2813.98	2809.32	NA
16	2846.77	2848.68	2850.52	2856.00	2827.71	2838.13	2843.86	2837.70	2814.73	2810.98	2809.78
17	2846.50	2847.06	2834.82	2846.24	2818.94	2826.11	2824.90	2813.63	2818.62	2809.47	2809.03
18	2839.87	2838.04	2839.64	2843.07	2843.50	2823.41	2818.19	2816.38	2819.36	2809.37	2813.16
19	2835.20	2837.35	2838.18	2836.46	2836.86	2843.35	2841.00	2829.24	2819.55	2814.69	2818.10

หมายเหตุ : NA = not available

จากตารางที่ 3.3.23 จะได้สมการพยากรณ์ ดังสมการต่อไปนี้

$$(1 - 0.9826B^1 - 0.0477B^2 + 0.1300B^3 - 0.0956B^4 + 0.0713B^5 - 0.0425B^6 + 0.0548B^7 - 0.1879B^8 - 0.6274B^9 + 0.7829B^{10})Y_t = 974,372.77 + (1 - 0.2246B^1 - 0.7754B^2)\varepsilon_t$$

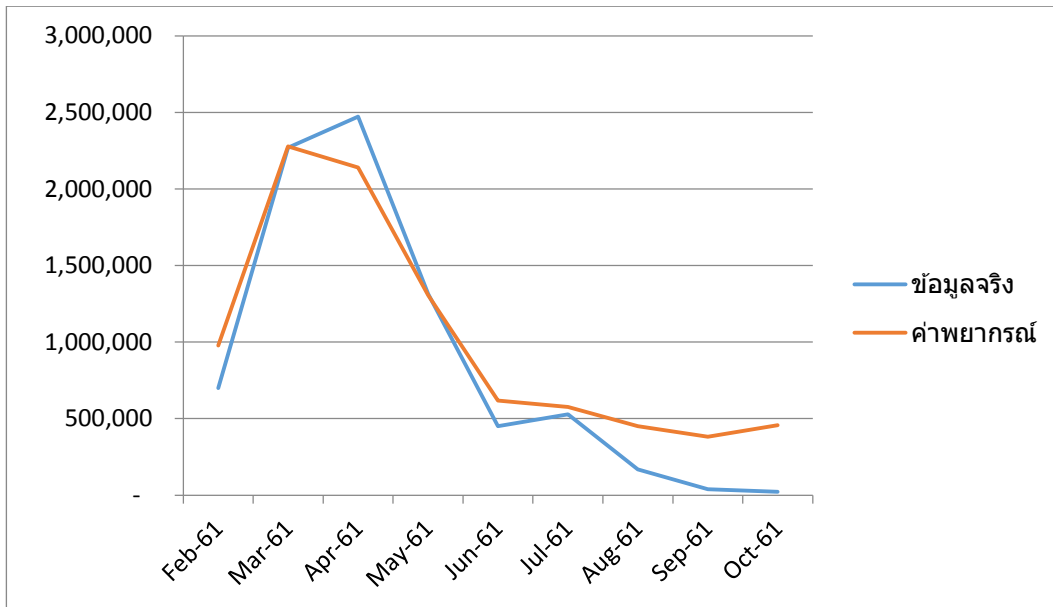
ขั้นตอนที่ 4 นำสมการพยากรณ์ที่ได้จากตัวแบบ ARMA(10,2) ไปหาค่าพยากรณ์และทำการเปรียบเทียบข้อมูลจริง หลังจากนั้นเราจะพิจารณาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เพื่อวัดความแม่นยำในการพยากรณ์ ดังตารางที่ 3.3.24 และสามารถแสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง ดังรูปที่ 3.3.25 และแสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ ปี พ.ศ. 2561 ถึงเดือนตุลาคมปี พ.ศ. 2561 ดังรูปที่ 3.3.26

ตารางที่ 3.3.24 แสดงปริมาณผลผลิตข้าวนาปรังจริงเปรียบเทียบกับค่าพยากรณ์ของปริมาณผลผลิตข้าวนาปรัง

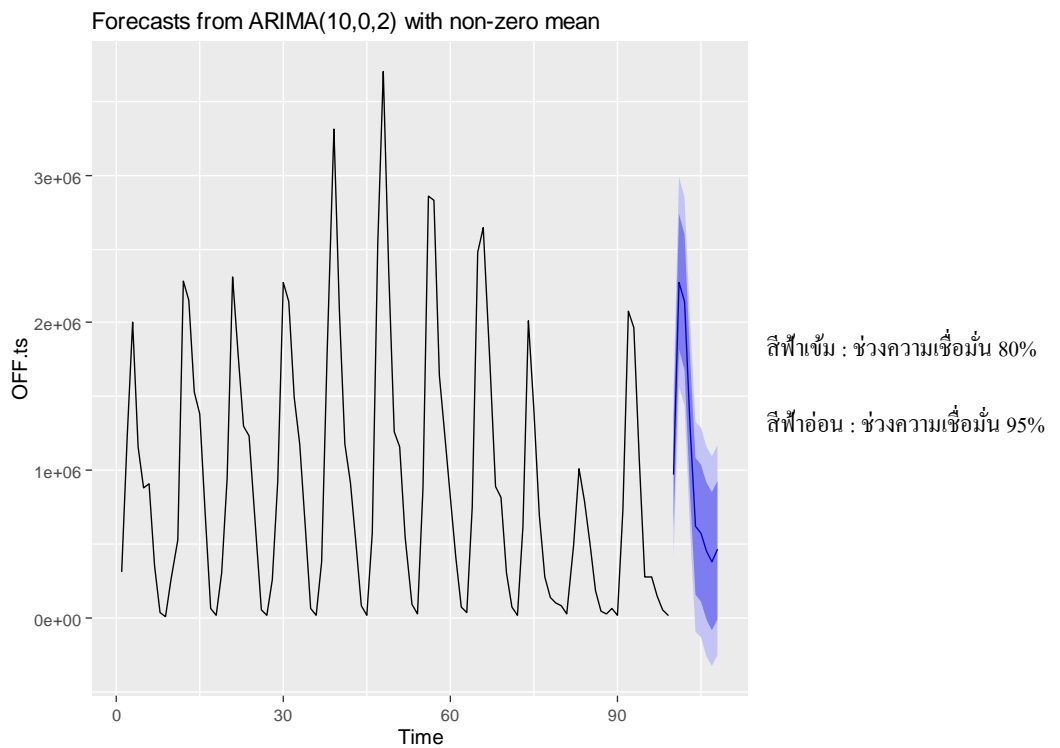
วันที่	ปริมาณผลผลิตข้าวนาปรังจริง (ตัน)	ค่าพยากรณ์ (ตัน)	ค่าความคลาดเคลื่อน (ตัน)	เปอร์เซ็นต์ความคลาดเคลื่อน
กุมภาพันธ์ 2561	699,536	976,907	277,371	39.65%
มีนาคม 2561	2,268,895	2,277,006	8,111	0.36%
เมษายน 2561	2,471,287	2,139,569	-331,718	-13.42%
พฤษภาคม 2561	1,314,711	1,305,019	-9,692	-0.74%
มิถุนายน 2561	450,818	617,619	166,801	37.00%
กรกฎาคม 2561	528,696	575,929	47,233	8.93%
สิงหาคม 2561	169,136	450,324	281,188	166.25%
กันยายน 2561	38,401	382,184	343,783	895.24%
ตุลาคม 2561	23,074	457,196	434,122	1881.43%

ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (F_t - Y_t)^2}{n}} \\ &= \sqrt{\frac{\sum_{t=1}^9 (F_t - Y_t)^2}{9}} \\ &= 258,822 \end{aligned}$$



รูปที่ 3.3.25 แสดงกราฟของค่าพยากรณ์เปรียบเทียบกับปริมาณผลผลิตข้าวนาปรังจริง



รูปที่ 3.3.26 แสดงค่าพยากรณ์ปริมาณผลผลิตข้าวนาปรังตั้งแต่เดือนกุมภาพันธ์ปี พ.ศ. 2561 ถึงเดือนตุลาคมปี พ.ศ. 2561

บทที่ 4

สรุปผลงานวิจัยและข้อเสนอแนะ

4.1 สรุปผลงานวิจัย

ในการสร้างตัวแบบ ARMA จากข้อมูลปริมาณผลผลิตข้าวในปีและนาปรังที่เก็บรวบรวมโดยสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ เพื่อพยากรณ์ปริมาณผลผลิตข้าวแบบรายเดือนในประเทศไทยของปี พ.ศ. 2557, 2558, 2559 2560 และ 2561 โดยใช้วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) และใช้จำนวนข้อมูลในการพยากรณ์ปริมาณผลผลิตข้าวจำนวน 99 เดือน หรือ 11 ปี ผลปรากฏว่าทั้ง 2 ฤดูเพาะปลูกนั้น เราสามารถเลือกตัวแบบ ARMA ที่เหมาะสมที่สุดกับข้อมูล โดยพิจารณาค่า Akaike Information Criterion (AIC) เป็นเกณฑ์ในการเลือกตัวแบบ หลังจากนั้นจะทำการประเมินความแม่นยำในการพยากรณ์ในงานวิจัยนี้ซึ่งใช้เกณฑ์ค่าความคลาดเคลื่อนของตัวแบบ โดยใช้ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) ให้ผลดังตารางที่ 4.1.1-4.1.2

ตารางที่ 4.1.1 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ของฤดูเพาะปลูกแบบนาปี

ปีที่พยากรณ์	RMSE
2557	1,557,008
2558	1,191,433
2559	690,460
2560	580,969
2561	344,561

ตารางที่ 4.1.2 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองโดยเฉลี่ย (RMSE) ของฤดูเพาะปลูกแบบนาปรัง

ปีที่พยากรณ์	RMSE
2557	260,349
2558	578,313
2559	337,825
2560	555,444
2561	258,822

จากตารางที่ 4.1.1-4.1.2 พบว่ามีค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ที่ค่อนข้างสูง จึงสามารถสรุปผลการวิจัยได้ว่าการพยากรณ์ปริมาณผลผลิตข้าวแบบรายเดือนในประเทศไทยโดยวิธีบ็อกซ์-เจนกินส์ไม่แม่นยำ เนื่องจากผลการพยากรณ์ให้ค่าแตกต่างจากข้อมูลจริงมาก

4.2 ข้อเสนอแนะ

1. โครงการนี้ได้ทำการศึกษาพยากรณ์โดยใช้วิธีบ็อกซ์-เจนกินส์เท่านั้น ซึ่งอาจจะไม่ใช่วิธีการที่จะให้ผลการพยากรณ์ที่ดีที่สุดก็ได้ ดังนั้นงานวิจัยในอนาคตควรทำการศึกษาเทคนิคที่ใช้ในการพยากรณ์เพิ่มเติมและควรมีหลายวิธีที่ใช้ในการพยากรณ์เพื่อจะได้เปรียบเทียบค่าพยากรณ์ที่ให้ค่าใกล้เคียงกับค่าจริงมากที่สุด เช่น วิธีเฉลี่ยเคลื่อนที่ (Simple Moving Average method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing method) เป็นต้น
2. ข้อมูลที่ใช้ในโครงการนี้ คือ ปริมาณผลผลิตข้าวในประเทศไทยแบบรายเดือนแบ่งตามฤดูเพาะปลูก ได้แก่ ข้าวนาปีและข้าวนาปรัง ซึ่งข้อมูลดังกล่าวเป็นข้อมูลที่มีลักษณะค่าเฉลี่ยและความแปรปรวนไม่คงที่ หรือหมายความว่าข้อมูลมีผลกระทบจากส่วนประกอบแนวโน้มและการแปรผันของฤดูกาล โดยงานวิจัยชิ้นนี้ สมมติให้ข้อมูลมีลักษณะที่คงที่ (Stationary) และทำการพยากรณ์ต่อไป ดังนั้นงานวิจัยในอนาคตก่อนจะเริ่มการพยากรณ์ควรแปลงข้อมูลให้มีลักษณะที่คงที่ (Stationary)

เอกสารอ้างอิง

- [1] กรินทร์ กาญจนานนท์. การพยากรณ์ทางสถิติ. กรุงเทพฯ : ซีเอ็ดยูเคชั่น, 2561.
- [2] ทรงศิริ แต่สมบัติ. การพยากรณ์เชิงปริมาณ. กรุงเทพฯ : สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, 2549.
- [3] ภูมิฐาน รังกุลนุวัฒน์. การวิเคราะห์อนุกรมเวลาสำหรับเศรษฐศาสตร์และธุรกิจ. พิมพ์ครั้งที่ 1, กรุงเทพฯ : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2556.
- [4] วิโรจน์ อรุณมานะกุล. สถิติและการใช้โปรแกรม R. ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
- [5] ศิริเทพ จันทร์บุญแก้ว. การพยากรณ์ปริมาณสายโทรเข้าสำหรับศูนย์บริการลูกค้าธนาคารพาณิชย์. วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมอุตสาหกรรม ภาควิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย, 2560.
- [6] สมเกียรติ เกตุเอี่ยม. เทคนิคการพยากรณ์. พิมพ์ครั้งที่ 2, การกิจเอกสารและตำรา กลุ่มงานบริการการศึกษา มหาวิทยาลัยทักษิณ, 2548.
- [7] สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์. วันที่ค้นข้อมูล 30 ตุลาคม 2562, เข้าถึงได้จาก <http://oae.go.th/view/1/ข้อมูลการผลิตสินค้าเกษตร/TH-TH>
- [8] Box, G. E., JENKINS, G. M., and REINSEL, G. C. Time series analysis: forecasting and control. 4th ed. John Wiley & Sons, 2008.
- [9] Hyndman, R.J., & Athanasopoulos, G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2., 2018.
- [10] Thanomsing, N. ภาวะน่าจะเป็นสูงสุด (Maximum Likelihood). Department of Biostatistics & Demography. Faculty of Public Health, KKU, 2558.

ภาคผนวก

ภาคผนวก ก

แบบเสนอหัวข้อโครงการ รายวิชา 2301399 Project Proposal

ปีการศึกษา 2562

ชื่อโครงการ (ภาษาไทย)	การพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย
ชื่อโครงการ (ภาษาอังกฤษ)	Forecasting of rice product in Thailand
อาจารย์ที่ปรึกษา	1. รองศาสตราจารย์ ทิพวัลย์ สันติวิภาณนท์ 2. รองศาสตราจารย์ ดร.ทรงเกียรติ สุเมธกิจการ
ผู้ดำเนินการ	นางสาวทัศนีย์ เร่งเทียน เลขประจำตัวนิสิต 5933518523 สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

หลักการและเหตุผล

ข้าวเป็นธัญพืชที่สำคัญชนิดหนึ่งของประชากรโลก และเป็นพืชเศรษฐกิจที่สำคัญของประเทศไทย ประเทศไทยมีการส่งออกข้าวมากเป็นอันดับ 2 ของโลก รองลงมาจากประเทศอินเดีย นอกจากนี้ข้าวที่เป็นพืชเศรษฐกิจและสินค้าส่งออกที่สำคัญของประเทศไทยแล้ว ประเทศไทยยังมีพืชเศรษฐกิจที่สำคัญอีก เช่น ยางพารา อ้อย มันสำปะหลัง และปาล์ม เป็นต้น โดยในปี 2562 ปริมาณการส่งออกพืชเศรษฐกิจที่สำคัญมากที่สุด คือ ข้าว รองลงมา คือ ยางพารา ซึ่งทั้งข้าวและยางพารามีปริมาณการส่งออกที่เพิ่มขึ้นจากปีที่แล้ว

ในส่วนของการผลิตข้าวประเทศไทยนั้นมีพื้นที่ทำนา จำแนกเป็น พื้นที่นาในเขตอาศัยน้ำฝนร้อยละ 75 และเป็นพื้นที่นาในเขตชลประทานร้อยละ 25 ซึ่งในปี 2562 มีเนื้อที่เพาะปลูกทั้งข้าวนาปีและข้าวนาปรังลดลง เนื่องจากมีเกษตรกรบางส่วนสนใจปรับเปลี่ยนไปปลูกพืชชนิดอื่นตามนโยบายของรัฐบาล เช่น อ้อย มันสำปะหลัง และข้าวโพดเลี้ยงสัตว์ เป็นต้น อย่างไรก็ตามในภาพรวมทั้งประเทศมีปริมาณผลผลิตต่อไร่เพิ่มขึ้น เนื่องจากพื้นที่ส่วนใหญ่มีสภาพอากาศเอื้ออำนวยต่อการเพาะปลูกข้าว ปริมาณน้ำอยู่ในเกณฑ์ดี มีน้ำเพียงพอต่อการเพาะปลูก

ปัจจุบันข้าวยังเป็นที่ต้องการของตลาดทั้งภายในประเทศและต่างประเทศ อีกทั้งสาธารณรัฐสังคมนิยมเวียดนามซึ่งเป็นประเทศคู่แข่งทางการค้าได้มีการส่งออกข้าวเป็นรองจากประเทศไทย แต่เมื่อ 10 ปีที่ผ่านมา สาธารณรัฐสังคมนิยม

เวียดนามได้มีการผลิตข้าวเพิ่มมากขึ้นและมีการปรับปรุงคุณภาพของข้าวจนใกล้เคียงกับของประเทศไทยแต่มีราคาที่สูงกว่า จึงทำให้ผู้บริโภคบางส่วนสั่งซื้อข้าวจากสาธารณรัฐสังคมนิยมเวียดนามแทนประเทศไทยที่มีราคาที่สูงกว่า ทำให้ประเทศไทยประสบปัญหาการค้าข้าวตกต่ำ เนื่องจากปริมาณการสั่งซื้อที่ลดลงของผู้บริโภคในต่างประเทศ ส่งผลให้ปริมาณผลผลิตข้าวในประเทศไทยมีจำนวนมาก ๆ กว่าความต้องการของผู้บริโภคในประเทศ แม้ว่ารัฐบาลจะออกนโยบายแก้ไขต่าง ๆ แต่ก็ไม่สามารถช่วยแก้ปัญหาในระยะยาวได้ ซึ่งปัจจัยที่มีผลต่อราคาข้าวตกต่ำ เช่น ปริมาณผลผลิตข้าวในประเทศ คุณภาพข้าว และปัญหาจากสภาพแวดล้อม เป็นต้น ด้วยเหตุนี้ผู้วิจัยจึงมีความสนใจที่จะทำการพยากรณ์ปริมาณผลผลิตข้าวในประเทศไทย

วัตถุประสงค์

เพื่อใช้ตัวแบบ ARIMA โดยวิธีบ็อกซ์-เจนกินส์ ในการพยากรณ์ปริมาณผลผลิตข้าวรายเดือนของประเทศไทย

ขอบเขตโครงการ

ศึกษาข้อมูลปริมาณผลผลิตข้าวในประเทศไทยโดยแบ่งตามฤดูเพาะปลูก ได้แก่ ข้าวนาปี โดยใช้ข้อมูลแบบรายเดือน ตั้งแต่เดือนสิงหาคม พ.ศ. 2546 ถึง เมษายน พ.ศ. 2562 และข้าวนาปรัง โดยใช้ข้อมูลแบบรายเดือน ตั้งแต่เดือนกุมภาพันธ์ พ.ศ. 2546 ถึง ตุลาคม พ.ศ. 2561

วิธีการดำเนินงาน

ก. แผนการศึกษา

1. ศึกษาและค้นคว้าข้อมูลจากงานวิจัย และบทความต่าง ๆ ที่เกี่ยวข้องเกี่ยวกับการพยากรณ์อนุกรมเวลา และแบบจำลอง ARMA
2. รวบรวมข้อมูลปริมาณผลผลิตข้าวนาปีและข้าวนาปรัง
3. ตรวจสอบและวิเคราะห์ข้อมูลปริมาณผลผลิตข้าวนาปีและข้าวนาปรัง
4. ใช้วิธีบ็อกซ์-เจนกินส์ ในการพยากรณ์ปริมาณผลผลิตข้าวนาปีและข้าวนาปรัง และทดสอบสมการ ARMA ที่ได้กับข้อมูลปริมาณผลผลิตข้าว
5. สรุปและจัดทำรูปเล่มรายงาน

ข. ระยะเวลาที่ศึกษา

ขั้นตอนการดำเนินการ	ปี 2562					ปี 2563			
	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.
1. ศึกษาและค้นคว้าข้อมูลจากงานวิจัย และบทความต่าง ๆ ที่เกี่ยวข้องเกี่ยวกับการพยากรณ์อนุกรมเวลาและแบบจำลอง ARIMA									
2. รวบรวมข้อมูลปริมาณผลผลิตข้าวในปีและข้าวนาปรัง									
3. ตรวจสอบและวิเคราะห์ข้อมูลปริมาณผลผลิตข้าวในปีและข้าวนาปรัง									
4. ใช้วิธีบ็อกซ์-เจนกินส์ ในการพยากรณ์ปริมาณผลผลิตข้าวในปีและข้าวนาปรัง และทดสอบสมการ ARIMA ที่ได้กับข้อมูลจริงของปริมาณผลผลิตข้าว									
5. สรุปและจัดทำรูปเล่มรายงาน									
6. นำเสนอโครงการงาน									

ประโยชน์ที่คาดว่าจะได้รับ

ก. ประโยชน์ด้านความรู้และประสบการณ์ต่อนิติบุคคล

1. ได้ศึกษาและการสร้างตัวแบบ ARMA ของวิธีบ็อกซ์-เจนกินส์
2. สามารถพยากรณ์ปริมาณผลผลิตข้าวในปีและข้าวนาปรังแบบรายเดือนของประเทศไทยได้

ข. ประโยชน์ที่ได้จากโครงการงานต่อผู้ใช้

เป็นประโยชน์ในการวิเคราะห์ปริมาณผลผลิตข้าวในอนาคต เพื่อช่วยให้ภาครัฐหรือหน่วยงานราชการสามารถวางแผนหาแนวทางช่วยปัญหาหาค่าข้าวตกต่ำได้

อุปกรณ์และเครื่องมือที่ใช้

1. กระดาษและเครื่องเขียน
2. คอมพิวเตอร์, โน้ตบุ๊ก
3. เครื่องพิมพ์
4. Microsoft Word

5. Microsoft Excel
6. โปรแกรม R
7. ข้อมูลปริมาณผลผลิตข้าว

งบประมาณ

รายการ	จำนวน	จำนวนเงิน (บาท)
1. กระดาษ A4	4 รีม	600
2. ถ่ายเอกสาร	1 เล่ม	100
3. ค่าหมึกเครื่องปริ้น	2 ขวด	1,000
4. หน่วยเก็บข้อมูลภายนอก (ฮาร์ดดิสก์) 2 เทระไบต์	1 ชิ้น	2,000
5. คีย์บอร์ด	1 ชิ้น	1,000
6. เม้าส์ไร้สาย	1 ชิ้น	300
รวม		5,000

เอกสารอ้างอิง

- [1] วิโรจน์ อรุณมานะกุล. สถิติและการใช้โปรแกรม R. ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
- [2] ศิริเทพ จันทน์บุญแก้ว. การพยากรณ์ปริมาณสายโทรเข้าสำหรับศูนย์บริการลูกค้าธนาคารพาณิชย์. วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมอุตสาหกรรม ภาควิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย, 2560.
- [3] สมเกียรติ เกตุเอี่ยม. เทคนิคการพยากรณ์. การกิจเอกสารและตำรา กลุ่มงานบริการการศึกษา มหาวิทยาลัยทักษิณ
- [4] สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์. วันที่ค้นข้อมูล 30 ตุลาคม 2562, เข้าถึงได้จาก <http://oae.go.th/view/1/ข้อมูลการผลิตสินค้าเกษตร/TH-TH>
- [5] Hyndman, R.J., & Athanasopoulos, G. (2018). Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia.

ภาคผนวก ข

คำสั่งโปรแกรม R ที่ใช้ในการวิเคราะห์

`library(fpp2)` : การติดตั้ง Package fpp2 ที่รวบรวมคำสั่งสำเร็จรูปที่ใช้ในการพยากรณ์

`library(openxlsx)` : การติดตั้งคำสั่งที่ใช้อ่านข้อมูลจากโปรแกรม Excel

`read.xlsx(" ",sheet =)` : คำสั่งอ่านข้อมูลจากโปรแกรม Excel

`autoplot()+xlab(" ") + ylab(" ")` : คำสั่งวาดกราฟอัตโนมัติโดยระบุชื่อแกน X และแกน Y

`ggAcf(, main = " ")` : คำสั่งวาดกราฟ ACF

`ggPacf(, main = " ")` : คำสั่งวาดกราฟ PACF

`auto.arima(, d = NA, D = NA, max.p = 10, max.q = 18, max.P = 0, max.Q = 0, max.d = 0, max.D = 0, stationary = TRUE, seasonal = FALSE)` : คำสั่งให้โปรแกรม R หาตัวแบบที่เหมาะสมกับข้อมูลอัตโนมัติ

`arima(, order = c(p,d,q))` : คำสั่งให้โปรแกรม R คำนวณค่าประมาณพารามิเตอร์และค่าสถิติต่างๆ

`forecast(, h =)` : คำสั่งให้ทำนายค่าพยากรณ์ในอนาคตโดยสามารถระบุจำนวนที่ต้องการพยากรณ์ (h)

`autoplot(forecast())` : คำสั่งวาดกราฟค่าจริงและค่าที่ได้จากการพยากรณ์

ประวัติผู้เขียน



นางสาวทัศนีย์ เร่งเทียน

รหัสนิสิต 5933518523

สาขาคณิตศาสตร์

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย