



รายงานผลการวิจัยฉบับสมบูรณ์

กระบวนการสโตแคสติก สำหรับการวิเคราะห์ตัวแบบอนุกรมเวลา
โดยประยุกต์ใช้กับข้อมูลราคาจริง และราคาในตลาดล่วงหน้าของ
ข้าวในประเทศไทย

**Stochastic Processes for Time Series Models with an
Application to Data of Cash and Future Prices of Rice in
Thailand**

ดร. พิษณุ ทองขาว

ผู้ช่วยศาสตราจารย์ นิตยา บุญลิตธิ์

งานวิจัยนี้ได้รับทุนสนับสนุนจากงบประมาณเงินรายจ่าย ประจำปีงบประมาณพ.ศ. 2558

มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

กิตติกรรมประกาศ

การวิจัยเรื่องนี้สำเร็จลงด้วยดี ด้วยความช่วยเหลือจากหลายท่าน คณะผู้วิจัยขอขอบพระคุณ รศ. สุภัตรา โกสไศยกานนท์ อธิการบดีมหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร ผศ. ดร. วิโรจน์ ฤทธิทอง คณบดีคณะวิศวกรรมศาสตร์ ดร.สิงห์แก้ว ปือกเท็ง ผู้อำนวยการสถาบันวิจัยและพัฒนา ที่ให้การสนับสนุนการทำงานวิจัยของอาจารย์ในมหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร มาตั้งแต่เริ่มต้น ทำนี้คณะผู้วิจัยขอขอบคุณมหาวิทยาลัยเทคโนโลยีราชมงคลพระนครที่ได้ให้ทุนสนับสนุน จนกระทั่งงานวิจัยฉบับนี้สำเร็จลงด้วยดี

คณะผู้วิจัย



ชื่อเรื่อง : กระบวนการสโตแคสติกสำหรับการวิเคราะห์ตัวแบบอนุกรมเวลาโดยประยุกต์ใช้กับ
ข้อมูลราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทย

ผู้วิจัย : ดร. พิษณุ ทองขาว และ ผศ. นิตยา บุญสิทธิ์

พ.ศ. : 2558

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อใช้หลักการของกระบวนการสโตแคสติกสำหรับสร้างตัวแบบและประมาณค่าพารามิเตอร์ต่างๆด้วยวิธีของเบย์ กับข้อมูลราคาจริง และราคาในตลาดล่วงหน้าของข้าวในประเทศไทย ซึ่งจากการวิเคราะห์ข้อมูลราคาข้าวชนิดต่างๆที่ใช้ในงานวิจัยนี้ได้แก่ ราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ย จำนวน 120 เดือน ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 รายเดือนเฉลี่ย จะเป็นข้อมูลอนุกรมเวลาโดยมีส่วนประกอบของค่าอัตโนมัติ สัมพันธ์ แนวโน้ม ค่าฤดูกาล และค่าผิดปกติ โดยเขียนอัลกอริทึมและแปลงเป็นโปรแกรมใน OpenBUGS และประเมินประสิทธิภาพของตัวแบบต่างๆจากการเขียนโปรแกรมจำลองสถานการณ์ในโปรแกรม R หลังจากนั้นนำตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มซึ่งได้แก่ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเวย์บูล ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล และตัวแบบการพยากรณ์แบบปรับเรียบเอ็กซ์โปเนนเชียลมาเปรียบเทียบกัน โดยดูจากค่า RMSE MAPE และ MAE ผลการวิจัยพบว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของราคาข้าว 2 ชนิด ได้แก่ ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 รายเดือนเฉลี่ย ทั้งการหาตัวแบบที่เหมาะสม (Model Fitting) และการตรวจสอบความถูกต้องของตัวแบบ (Model Validation) ส่วนตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเวย์บูลให้ค่า RMSE MAPE และ MAE ต่ำสุด คือ ราคาปิดของข้าวขาว 5% ทั้งการหาตัวแบบที่เหมาะสม และการตรวจสอบความถูกต้องของตัวแบบ

คำสำคัญ: กระบวนการสโตแคสติก, ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้ม, อนุกรมเวลา, อัตตสหสัมพันธ์, วิธีการของเบย์

Title : Stochastic Processes for Time Series Models with an Application to Data of
Cash and Future Prices of Rice in Thailand

Researcher : Dr. Pitsanu Tongkhaw and Assist. Prof. Nittaya Boonsith

Year : 2015

Abstract

The objective of this research is to use a stochastic process for building a model using Bayesian parameter estimation. The model is applied to the of crash prices and future prices of rice in Thailand for 120 moths including monthly closing average price of white rice 5%, non-glutinous paddy 15%, and Hom Mali non-glutinous paddy 105. The data are time series data containing autocorrelation, trend, seasonality, and outliers. The algorithms are turned into codes in OpenBugs and the performance of the model is evaluated using simulation in R. The model with a Weibull cumulative distribution function trend, the model with an exponential cumulative distribution function trend, and the exponential smoothing model are compared. The root mean squared error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE) are used as the criteria. The results show that the model with the model with an exponential cumulative distribution function trend gives smallest RMSE, MAPE and MAE for the crash prices of non-glutinous paddy 15% and Hom Mali non-glutinous paddy 105 for both model fitting part and model validation part. The model with a Weibull cumulative distribution function trend gives the smallest RMSE, MAPE and MAE for the closing average price of white rice 5% for both model fitting part and model validation part.

Keyword: Stochastic process, Cumulative distribution function for trend, Time series, Autocorrelation, Bayesian method

สารบัญ

	หน้า
กิตติกรรมประกาศ	ก
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
สารบัญ	ง
บทที่ 1 บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์ของการวิจัย	3
สมมุติฐานในการวิจัย	3
ขอบเขตของการวิจัย	4
ประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	5
ทฤษฎีที่เกี่ยวข้อง	5
งานวิจัยที่เกี่ยวข้อง	55
บทที่ 3 วิธีดำเนินการวิจัย	59
ข้อมูลและแหล่งข้อมูล	59
ขอบเขตการวิจัย	59
ขั้นตอนการดำเนินการวิจัย	60
สถานที่ทำวิจัย	65
ระยะเวลาในการทำวิจัย	65
บทที่ 4 ผลการวิเคราะห์ข้อมูล	66
ผลการวิเคราะห์ลักษณะของข้อมูล	66
ผลการใช้หลักการทางกระบวนการสโตแคสติกสร้างตัวแบบ	
เพื่อวิเคราะห์ข้อมูลอนุกรมเวลา	70
ผลของการประเมินประสิทธิภาพของตัวแบบ	73
ผลของการประมาณค่าพารามิเตอร์	76
ผลของการเปรียบเทียบตัวแบบ	81

สารบัญ (ต่อ)

	หน้า
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ	85
สรุปผล	85
อภิปรายผล	87
ข้อเสนอแนะในการวิจัยครั้งต่อไป	88
บรรณานุกรม	89
ประวัติคณะผู้วิจัย	94



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวิเคราะห์โจทย์ปัญหาในชีวิตจริงทางวิทยาศาสตร์นั้น จะใช้ตัวแบบทางคณิตศาสตร์อยู่สองชนิดคือ ตัวแบบที่กำหนดให้ค่าพารามิเตอร์ และตัวแปรต่างๆเป็นแบบคงที่ (Deterministic model) ซึ่งตัวแบบประเภทนี้ส่วนใหญ่จะสามารถหาคำตอบได้อย่างง่าย และมีความถูกต้องและแน่นอน สำหรับการวิเคราะห์ตัวแบบประเภทนี้จะมีโปรแกรมสำเร็จรูปมากมายเพื่อช่วยวิเคราะห์ข้อมูลเพื่อทำการหาคำตอบ ข้อเสียของตัวแบบที่เป็นแบบ Deterministic ค่าพารามิเตอร์และค่าของตัวแปรต่างๆที่กำหนดให้เป็นค่าคงที่นั้นซึ่งในความเป็นจริงค่าต่างๆส่วนใหญ่จะไม่เป็นค่าที่มีความไม่แน่นอนในความเป็นจริงในชีวิตประจำวัน เช่นค่าของราคาข้าวของต่างๆ ภูมิอากาศ ดอกเบี้ยธนาคาร ค่าของเงินต่างๆในโลก เป็นต้น แต่มีตัวแบบอีกชนิดหนึ่งที่สามารถตอบโจทย์ให้ใกล้เคียงกับความเป็นจริง ซึ่งจะกำหนดให้ค่าพารามิเตอร์ และตัวแปรต่างๆเป็นแบบไม่คงที่ตัวแบบเรียกตัวแบบนี้ว่าตัวแบบที่เป็นแบบสโตแคสติก (Stochastic model)

สำหรับตัวแบบสโตแคสติก ค่าพารามิเตอร์ และตัวแปรต่างๆในตัวแบบจะอยู่ในรูปแบบของฟังก์ชันความน่าจะเป็น จะไม่สามารถคาดเดาคำตอบหรือผลลัพธ์ว่าจะออกมาเป็นเท่าใด แต่จะสามารถอธิบายความน่าจะเป็นของคำตอบหรือผลลัพธ์นั้นๆโดยใช้หลักการคาดเดาต่างๆหรือหลักการของทฤษฎีความน่าจะเป็นเข้ามาช่วยในการประมาณ หรือการจำลองสถานการณ์เพื่อทำการหาคำตอบ สำหรับงานวิจัยนี้ผู้วิจัยจึงสนใจในหลักการของกระบวนการสโตแคสติกเพื่อสร้างตัวแบบวิเคราะห์ข้อมูลให้มีความเหมือนกับความเป็นจริงกับปัจจุบันให้มากที่สุดและนำตัวแบบที่ได้ไปประยุกต์ใช้กับข้อมูลราคาจริงของข้าวและราคาซื้อขายข้าวในตลาดล่วงหน้า ซึ่งราคาข้าวทั้งสองชนิดนั้นจะเป็นข้อมูลอนุกรมเวลา ซึ่งในตัวแบบสำหรับการวิเคราะห์ข้อมูลจะมีส่วนประกอบที่สำคัญได้แก่ แนวโน้ม (trend) อัตตสหสัมพันธ์ (autocorrelation) ฤดูกาล (seasonal) วัฏจักร (cycle) และเหตุการณ์ผิดปกติ (irregular variation or outlier) ในส่วนของแนวโน้ม จะกำหนดฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของข้อมูลอนุกรมเวลาที่เหมาะสมกับข้อมูลที่มีความผันผวน ซึ่งจะ

อยู่ในรูปของตัวแปรสุ่มเพื่อจะช่วยให้การวิเคราะห์ข้อมูลอนุกรมเวลาให้มีความแม่นยำขึ้น โดยจะต้องมีการกำหนดฟังก์ชันต่างๆของการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มโดยปรับให้พอดี (fitting) กับข้อมูลให้มากที่สุด หรือมีความผิดพลาดน้อยที่สุดเมื่อเปรียบเทียบกับข้อมูลจริง

ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มที่นิยมใช้กันอยู่ทั่วไป เช่น ฟังก์ชันการแจกแจงความน่าจะเป็นสะสมแบบเวย์บูล (Weibull cumulative distribution function) แบบเอ็กซ์โพเนนเชียล (exponential cumulative distribution function) เป็นต้น ในส่วนของอัตถศาสตร์สัมพันธ์ (Autoregression หรือ AR) เช่น AR(1) AR(2) AR(3) เป็นต้น สำหรับในส่วน of ฤดูกาล วัฏจักร และเหตุการณ์ผิดปกติ ก็จะมีฟังก์ชันการแจกแจงของความน่าจะเป็นต่างๆ เหมือนกัน แต่ปัญหาที่ตามมาก็คือจะทำให้ตัวแบบสำหรับการวิเคราะห์ข้อมูลอนุกรมเวลาจะมีความซับซ้อนมากขึ้น การหาคำตอบด้วยวิธีทางคณิตศาสตร์ต่างๆไปไม่สามารถทำการประมาณพารามิเตอร์ได้ ซึ่งจะต้องประมาณพารามิเตอร์โดยการจำลองสถานการณ์ และถ้าใส่ฟังก์ชันครบทุกส่วนของข้อมูลอนุกรมเวลาจากที่ได้กล่าวมาข้างต้นวิธีการจำลองสถานการณ์ก็ยังมีขีดจำกัดในการประมาณพารามิเตอร์เพื่อให้ใกล้เคียงกับข้อมูลจริงมากที่สุด

วิธีการหนึ่งซึ่งสามารถประมาณค่าพารามิเตอร์โดยใช้การจำลองสถานการณ์กับตัวแบบที่มีฟังก์ชันหลายๆฟังก์ชันรวมอยู่ด้วยกันซึ่งเรียกว่าวิธีการของเบย์ สำหรับวิธีการของเบย์ ในประเทศไทยยังถือว่ามีการใช้กันอยู่น้อยมาก และที่มีการใช้ส่วนใหญ่ ก็ยังเป็นตัวแบบที่วิเคราะห์กันทั่วๆไป ส่วนใหญ่ก็ยังใช้กันอยู่ในตัวแบบที่มีข้อจำกัดที่เป็นตัวแบบง่ายๆไม่สามารถวิเคราะห์ข้อมูลที่มีส่วนประกอบของฟังก์ชันต่างๆหลายฟังก์ชันรวมกันได้ ซึ่งก็จะทำให้ได้ผลการวิเคราะห์ไม่แตกต่างจากการวิเคราะห์ด้วยตัวแบบอื่นๆที่นิยมใช้กันอยู่ในปัจจุบัน ข้อดีของวิธีการของเบย์คือสามารถนำไปประยุกต์ใช้ในการประมาณค่าพารามิเตอร์ในตัวแบบที่มีความซับซ้อนได้ดี และการที่พารามิเตอร์มีได้หลายระดับ และมองเป็นฟังก์ชันต่างๆย่อมสอดคล้องกับสภาพความเป็นจริงในชีวิตประจำวันที่ว่าทุกสิ่งทุกอย่างย่อมขึ้นอยู่กับกันและกัน และไม่เป็นอิสระกัน ซึ่งก็จะตรงกับความเป็นจริงในสถานการณ์ในปัจจุบันก็คือส่วนใหญ่มักจะมีความไม่แน่นอน ซึ่งความไม่แน่นอนของทุกสิ่งทุกอย่างต้องมีการขึ้นอยู่กับกันและกันอย่างเป็นขั้นๆ ดังที่กล่าวมา เช่น ราคาขึ้นอยู่กับความต้องการสินค้าอากาศ ภัยธรรมชาติ นโยบายทางการบริหารประเทศเป็นต้น และความต้องการก็ยังขึ้นอยู่กับปัจจัยทางด้านอื่นๆอีก ทำให้เกิดการขึ้นอยู่กับสิ่งต่างๆหลายๆชั้นจึงทำให้เกิดความซับซ้อนมากขึ้นไปอีกสำหรับการวิเคราะห์ ดังนั้นตัวแบบทางคณิตศาสตร์ที่ใช้ในการวิเคราะห์จึงมีความซับซ้อนตามไปด้วย และต้องบรรจุฟังก์ชันของความน่าจะเป็นต่างๆที่เหมาะสมกับข้อมูลในการประมาณ

ค่าพารามิเตอร์หรือตัวแปรไว้ในตัวแบบ ซึ่งตัวแบบประเภทนี้เรียกว่าตัวแบบเชิงสโตแคสติก วิธีของเบย์กับตัวแบบที่วิเคราะห์ข้อมูลเชิงลึกหลายชั้นในปัจจุบันจึงมีผู้นำไปประยุกต์ใช้ในงานวิจัยด้านต่างๆกันอย่างแพร่หลายส่วนใหญ่อยู่นานต่างประเทศ

ด้วยเหตุนี้ผู้วิจัยจึงสนใจที่จะศึกษา และสร้างตัวแบบโดยใช้กระบวนการสโตแคสติกด้วยวิธีของเบย์มาทำการวิเคราะห์ข้อมูลเชิงลึกที่มีส่วนประกอบหลายๆส่วนดังที่ได้กล่าวมาข้างต้นซึ่งจะมีฟังก์ชันความน่าจะเป็นต่างๆหลายๆฟังก์ชัน ได้แก่ฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวนอน ฟังก์ชันของอัตราส่วนสัมพันธ์ และฟังก์ชันของค่าผิดปกติ รวมอยู่ในตัวแบบด้วยเพื่อให้สอดคล้องกับสภาพความเป็นจริงในชีวิตประจำวัน ด้วยกระบวนการสโตแคสติกโดยใช้วิธีการประมาณค่าพารามิเตอร์แบบเบย์ หลังจากนั้นนำไปประยุกต์ใช้กับข้อมูลราคาข้าว และเมื่อเราสร้างตัวแบบได้แล้วจะต้องมีการประเมินค่าความน่าเชื่อถือของตัวแบบโดยใช้วิธีจำลองสถานการณ์และเขียนโปรแกรมในโปรแกรม R ซึ่งเหตุจูงใจที่เลือกใช้กับราคาข้าวก็เพราะว่าข้าวเป็นพืชเศรษฐกิจที่สำคัญของโลก และมีเหตุการณ์ต่างๆเกิดขึ้น เช่น สถานะทางเศรษฐกิจ นโยบายทางการเมืองภายในประเทศ หรือปัจจัยต่างๆที่เกิดจากในและนอกประเทศ ได้ตลอด และตัวแบบที่ได้จากงานวิจัยนี้ก็สามารถนำไปใช้ในการประมาณการเพื่อพยากรณ์ข้อมูลอนุกรมเวลาอื่นๆได้เช่นกัน

วัตถุประสงค์ของการวิจัย

ใช้หลักการของกระบวนการสโตแคสติกเพื่อสร้างตัวแบบสำหรับประมาณค่าพารามิเตอร์ต่างๆ กับข้อมูลราคาจริง และราคาในตลาดล่วงหน้าของข้าวในประเทศไทย

สมมุติฐานของการวิจัย

ราคาข้าว มีความสัมพันธ์กับ แนวโน้มที่มีรูปแบบเป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมแบบเอกโพเนนเชียล และแนวโน้มที่มีรูปแบบเป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมแบบเวย์บูล

ขอบเขตของโครงการวิจัย

ข้อมูลราคาปีคของข้าวขาว5% เฉลี่ยรายเดือน ตั้งแต่ปี 2548-2557 จำนวน 120 เดือน จากตลาดสินค้าเกษตรล่วงหน้าแห่งประเทศไทย ข้อมูลราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ตั้งแต่ปี 2540-2557 จำนวน 216 เดือนจากสำนักงานสถิติการเกษตร

ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถใช้หลักการประมาณพารามิเตอร์เชิงสโตแคสติกสำหรับการวิเคราะห์เชิงลึกกับตัวแบบที่มีความซับซ้อนซึ่งเหมือนกับสภาพความเป็นจริง
2. ได้ตัวแบบการวิเคราะห์ และพยากรณ์ข้อมูลอนุกรมเวลาที่มีความเหมาะสมกับราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทย ที่ตรงกับสภาพความเป็นจริงให้มีความแม่นยำมากยิ่งขึ้น
3. เป็นแนวทางในการประมาณค่าพารามิเตอร์ในตัวแบบอื่นๆ ได้ และไม่จำเป็นต้องตัวแบบด้านการพยากรณ์เท่านั้น แต่ยังสามารถนำหลักการนี้ไปใช้ในการประมาณค่าตัวแปรในตัวแบบทางด้านการวิจัยดำเนินงานเพื่อหาค่าเหมาะที่สุดได้เช่นกัน
4. ค่าพารามิเตอร์ และค่าพยากรณ์ที่ได้สามารถนำไปใช้สำหรับวางแผนและตัดสินใจต่างๆ เช่น ภาครัฐใช้วางแผนทำการเกษตรที่เหมาะสมให้กับเกษตรกรได้
5. นำตัวแบบการพยากรณ์ที่ได้ไปประยุกต์ใช้กับการพยากรณ์ด้านอื่นๆ ได้
6. ใช้เป็นพื้นฐานในการพัฒนาตัวแบบต่อไป

บทที่ 2

เอกสาร และงานวิจัยที่เกี่ยวข้อง

กระบวนการสโตแคสติกสำหรับการวิเคราะห์ตัวแบบอนุกรมเวลาโดยประยุกต์ใช้กับข้อมูลราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทยในครั้งนี้ผู้ศึกษาได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องตามหัวข้อดังต่อไปนี้

ทฤษฎีที่เกี่ยวข้อง
งานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1. สโตแคสติก (Stochastic) (Dembo, 2013)

สโตแคสติกจะเกี่ยวข้องกับตัวแปรสุ่ม (Random Variable) เช่น การประมาณค่าพารามิเตอร์ในแบบจำลองการพยากรณ์ การวิเคราะห์ระบบสินค้าคงคลัง โดยที่ค่าพารามิเตอร์ในแบบนั้นเป็นตัวแปรสุ่ม ซึ่งอยู่ในภาพของฟังก์ชันของความน่าจะเป็น วิธีการประมาณค่าที่นิยมใช้กันอยู่ทั่วไป เช่น การประมาณค่าความควรจะเป็นสูงสุด (Maximum Likelihood Estimation) วิธีนี้นิยมใช้ประมาณค่าพารามิเตอร์ที่เป็นตัวแปรสุ่มโดยที่ตัวแปรสุ่มแต่ละตัวมีค่าของ prior distribution แยกเดี่ยวๆหรือมีหลายระดับชั้น ฟังก์ชันต้องไม่มีความซับซ้อนเท่าไรจึงจะคำนวณหาพารามิเตอร์มาได้ แต่ถ้าตัวแปรสุ่มแต่ละตัวมีค่าของฟังก์ชันความน่าจะเป็นหรือ prior distribution หลายระดับชั้นซึ่งจะเรียกว่า hyperprior distribution และเมื่อหา joint distribution มาแล้ว ได้ฟังก์ชันที่มีความซับซ้อนมาก เราก็จะนิยมใช้วิธีการประมาณค่าที่สามารถประมาณฟังก์ชันที่มีความซับซ้อนได้ เช่น วิธีการประมาณแบบเบย์ (Bayesian Estimation) เป็นต้น

จากที่กล่าวมาข้างต้นจะเรียกว่าวิธีกระบวนการสโตคาสติก (Stochastic Process) เพราะกระบวนการประมาณค่าพารามิเตอร์นอกจากจะมีตัวแปรสุ่มแล้ว ยังมีเวลาหรือจำนวนการวนซ้ำมาเกี่ยวข้องในกระบวนการด้วย วิธีของเบย์จะใช้หลักการของห่วงโซ่มาร์คอฟ (Markov Chain) มาช่วยในการประมาณ ซึ่งวิธีการของห่วงโซ่มาร์คอฟ ถือว่าเป็นวิธีการของกระบวนการสโตคาสติกที่ได้รับความนิยมมากพอสมควร และถูกนำไปประยุกต์ใช้ในหลายหลายสาขาวิชา ในที่นี้จะขอลำกระบวนการสโตคาสติกคร่าวๆ โดยมี S เป็นเซตของสถานที่เป็นไปได้ทั้งหมด (State Space) ก็คือเซตของตัวแปรสุ่ม $\{X_t, t \in T\}$ โดยที่ X_t จะมีค่าความน่าจะเป็น และสามารถหาค่าความน่าจะเป็นของมันได้ โดยจะเหมือนกับเซตของตัวแปรสุ่มที่เราใช้กันอยู่ทั่วไป เช่น $f(x)$, $p(x)$ เป็นต้น เซต T จะถูกเรียกว่าเซตพารามิเตอร์ของกระบวนการ ถ้า $T = \mathbb{N} = \{0, 1, 2, \dots\}$ กระบวนการนี้จะเรียกว่ากระบวนการหนึ่งของการประมาณพารามิเตอร์แบบไม่ต่อเนื่อง (A discrete parameter process) แต่ถ้า T ไม่สามารถนับได้ กระบวนการนี้เรียกว่ากระบวนการหนึ่งของการประมาณพารามิเตอร์แบบต่อเนื่อง (A continuous parameter process) โดยทั่วไป T จะเป็นจำนวนจริงบวกโดยที่ $T = \mathbb{R}_+ = (0, \infty]$ และ $T = [a, b] \subset \mathbb{R}$ เมื่อค่า t คือเวลาและเมื่ออยู่ในภาพของ X_t ก็จะเรียกว่า สถานะ หรือ ตำแหน่ง หรือ กระบวนการ ณ เวลา t ถ้า S อยู่ใน $\mathbb{N}$ ก็จะเรียกว่าเซตของจำนวนเต็มทั้งหมด หรือ เซตจำกัดเซตหนึ่งสำหรับค่าคงที่ $\omega \in \Omega$ การจับคู่จะได้ดังนี้

$$t \rightarrow X_t(\omega)$$

สมการข้างบนได้กำหนดบนเซตของพารามิเตอร์ T ซึ่งจะถูกรเรียกว่า A realization หรือ trajectory เป็นเส้นทางตัวอย่าง หรือฟังก์ชันตัวอย่างของกระบวนการ ให้ $\{X_t, t \in T\}$ เป็น กระบวนการสโตคาสติกแบบที่มีค่าเป็นจำนวนจริงกระบวนการหนึ่ง (real-valued) และ $\{t_1 < \dots < t_n\} \subset T$ ดังนั้น ฟังก์ชันความน่าจะเป็น $P_{t_1}, \dots, P_{t_n} o (X_{t_1}, \dots, X_{t_n})^{-1}$ ของเวกเตอร์ที่เป็นตัวแปรสุ่ม

ถ้าให้ X คือตัวแปรสุ่ม (Random Variable) โดยมี $f(x)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น (Probability Distribution Function) และ X มีค่าอยู่ระหว่างค่าขีดจำกัดล่างและขีดจำกัดบน หรือ lower $x < x < \text{upper } x$ เซตของค่าสังเกต $\mathbf{x}_t = \{x_1, x_2, \dots, x_n\}$ เมื่อ t คือ มิติของเวลาหรือระยะทางหรืออื่นๆ

ถ้า t คือ เวลา $\mathbf{x}_t$ คือ อนุกรมเวลา (Time Series) เซตของค่าสังเกตที่มาจากกระบวนการสโตแคสติก เรียกว่า Realization of Stochastic Process

กระบวนการสโตแคสติก คือ เซตของตัวแปรสุ่ม $\{X_t : X_1, X_2, \dots, X_n\}$

X_1 = ตัวแปรสุ่ม ซึ่งมี $f(x_1)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น

X_2 = ตัวแปรสุ่ม ซึ่งมี $f(x_2)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น

X_n = ตัวแปรสุ่ม ซึ่งมี $f(x_n)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น

$f(x_1, x_2, \dots, x_n)$ = ฟังก์ชันการแจกแจงความน่าจะเป็นร่วมของหลายตัวแปรสุ่ม (Joint Probability Distribution Function)

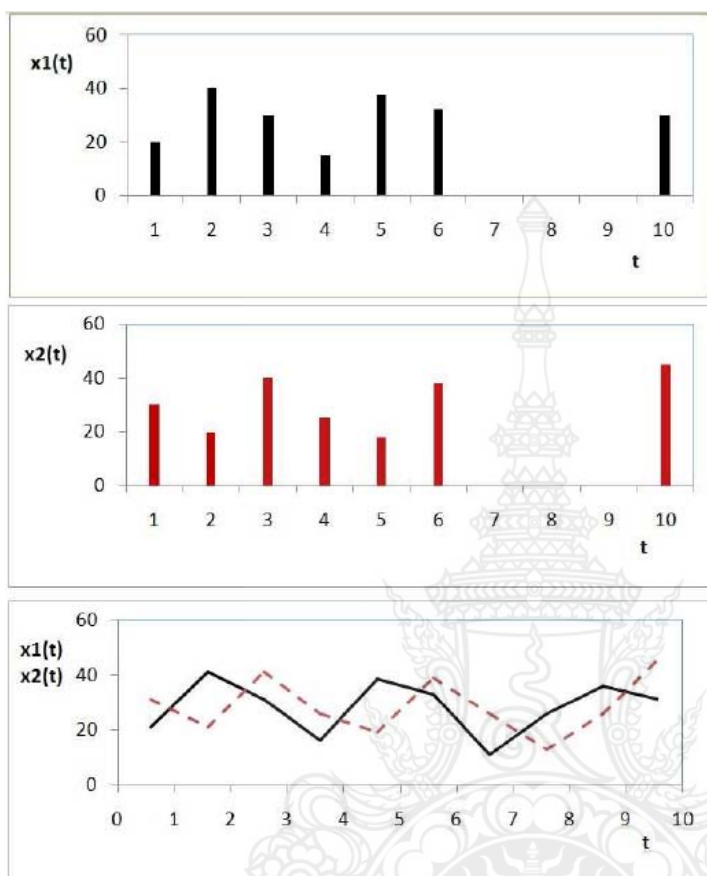
แบบจำลองซึ่งอธิบายโครงสร้างความน่าจะเป็นของเซตของค่าสังเกต $\{x_1, x_2, \dots, x_n\}$ เรียกว่า แบบจำลองสโตแคสติก (Stochastic Model)

1.1 รูปแบบของกระบวนการสโตแคสติก (Types of Stochastic Process) (วรารุช, 2554)

รูปแบบของกระบวนการสโตแคสติก สามารถแบ่งออกได้เป็นหลายลักษณะดังนี้

1) กระบวนการสโตแคสติกแบบ Discrete vs. Continuous

Discrete Process คือ กระบวนการสโตแคสติกที่มีค่าไม่ต่อเนื่อง เช่น จำนวนวันที่ฝนตกในแต่ละปี (Number of Rainy Days per Year) ขณะที่ Continuous Process คือ กระบวนการสโตแคสติกที่มีค่าต่อเนื่อง เช่น ปริมาณฝนรายปี เป็นต้น โดยทั่วไปข้อมูลอุทกวิทยาจะมีการบันทึกในช่วงเวลาที่ไม่ต่อเนื่อง (Discrete Time) เช่น รายปี รายเดือน แต่สามารถนำมาพล็อตเป็นกราฟได้ทั้งแบบต่อเนื่องและไม่ต่อเนื่อง ดังภาพที่ 1 ถือว่าเป็นข้อมูลจากกระบวนการสโตแคสติกต่อเนื่อง (Continuous Process)



ภาพที่ 1 Two realizations of stochastic processes

2) Independent (Purely Random) vs. Dependent Process

กระบวนการสุโตเคสติกแบบอิสระ (Independent Process) และแบบไม่อิสระ (Dependent Process) จะพิจารณาได้จากคุณสมบัติของฟังก์ชันการแจกแจงความน่าจะเป็นร่วมของหลายตัวแปรสุ่ม (Joint Probability Distribution Function) ดังนี้

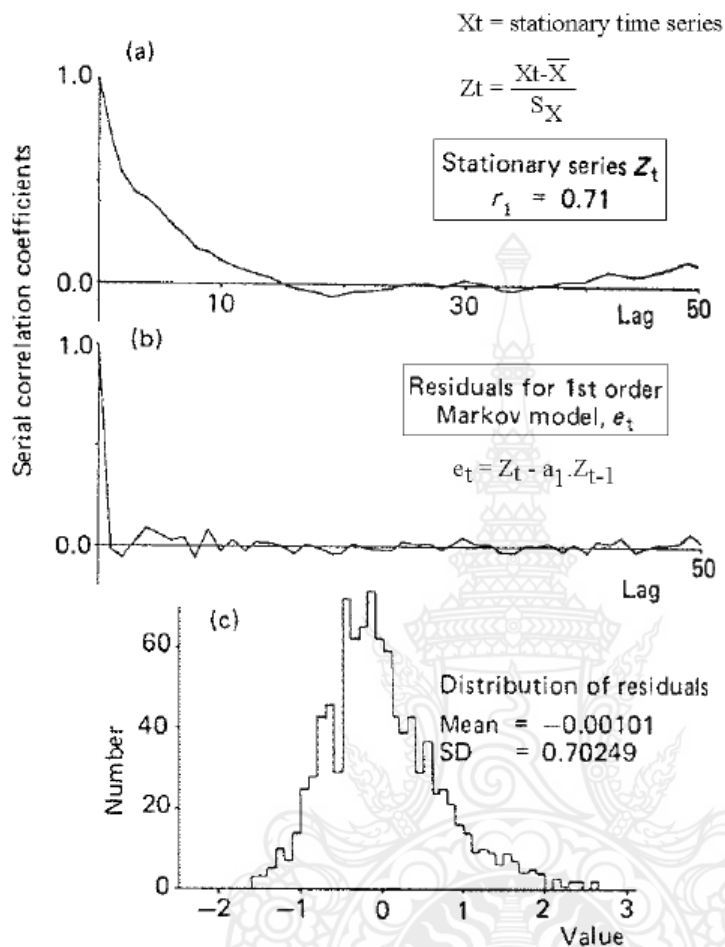
ถ้า $f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2)\dots f(x_n)$ หรือ ฟังก์ชันความน่าจะเป็นร่วมของหลายตัวแปรสุ่มเท่ากับผลคูณของ Marginal Distribution แสดงว่าเป็น Independent Process

ถ้า $f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2 | x_1)f(x_3 | x_1, x_2) \dots f(x_n | x_1, x_2, \dots, x_{n-1})$ หรือ
ฟังก์ชันความน่าจะเป็นร่วมของหลายตัวแปรสัมพันธ์กับผลคูณของ Conditional Distribution
แสดงว่าเป็น Dependent Process ซึ่งกรณีของ Dependent Process สามารถแบ่งออกได้เป็น 2 แบบ คือ

Serial Dependence Stochastic Process (Time Dependence) หรือ กระบวนการทางสถิติ
สถิติที่ขึ้นอยู่กับเวลา

Spatially Dependence Stochastic Process (Area Dependence) หรือ กระบวนการทางสถิติ
สถิติที่ขึ้นอยู่กับพื้นที่

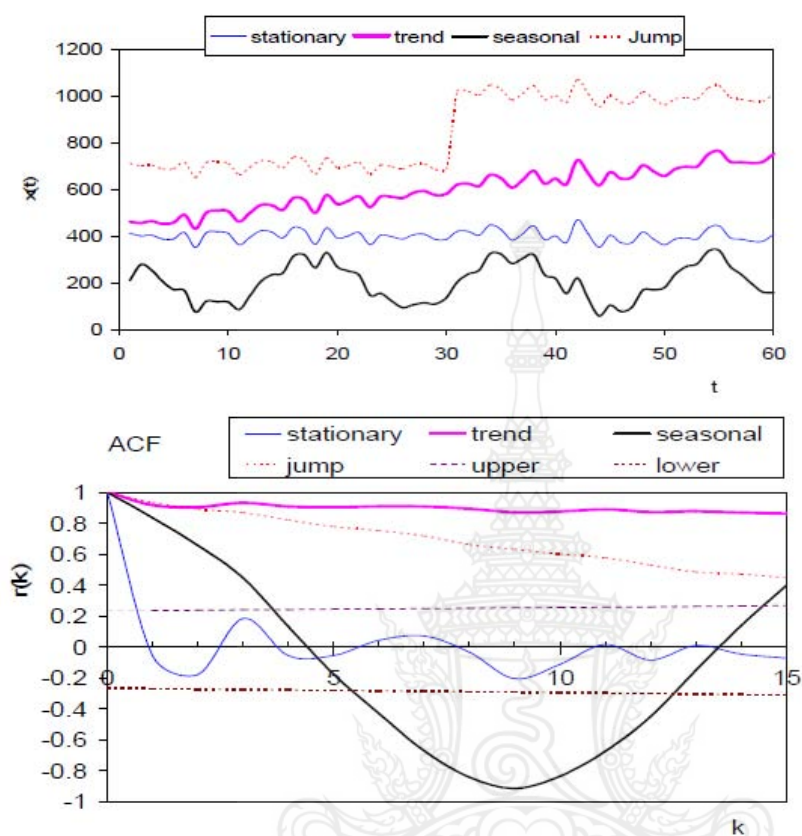
คุณสมบัติทางสถิติที่สามารถแสดงให้เห็นว่าอนุกรมเวลาเป็นแบบอิสระหรือไม่ คือ ฟังก์ชัน
สหสัมพันธ์ในตัวเอง (Autocorrelation Function: ACF) ตัวอย่าง ACF ของ Stationary Dependent and
Independent Series แสดงอยู่ในภาพที่ 2



ภาพที่ 2 Autocorrelation Function of Dependent Series (z_t) and Independent Series (ε_t)

3) Stationary vs. Non-Stationary Process

กระบวนการแบบ Stationary คือ กระบวนการสโตแคสติกซึ่งมีพารามิเตอร์ เช่น μ, σ^2, p ไม่เปลี่ยนแปลงไปตามเวลา ในทางกลับกันกระบวนการแบบ Non-Stationary คือ กระบวนการซึ่งพารามิเตอร์เปลี่ยนแปลงไปตามเวลา ปรัชการณณ์ เช่น Trend, Jump หรือ Seasonal Variation คือ องค์ประกอบของ Non-Stationary ที่สามารถเห็นได้ในอนุกรมเวลา ดังภาพที่ 3



ภาพที่ 3 Stationary and Non-Stationary Time Series

1.2 องค์ประกอบของอนุกรมเวลา (Components of Time Series)

อนุกรมเวลาประกอบด้วยองค์ประกอบที่สำคัญ 4 องค์ประกอบ คือ Trend, Periodic, Catastrophic และ Random ดังภาพที่ 4 และสามารถเขียนสมการอนุกรมเวลาได้ดังนี้

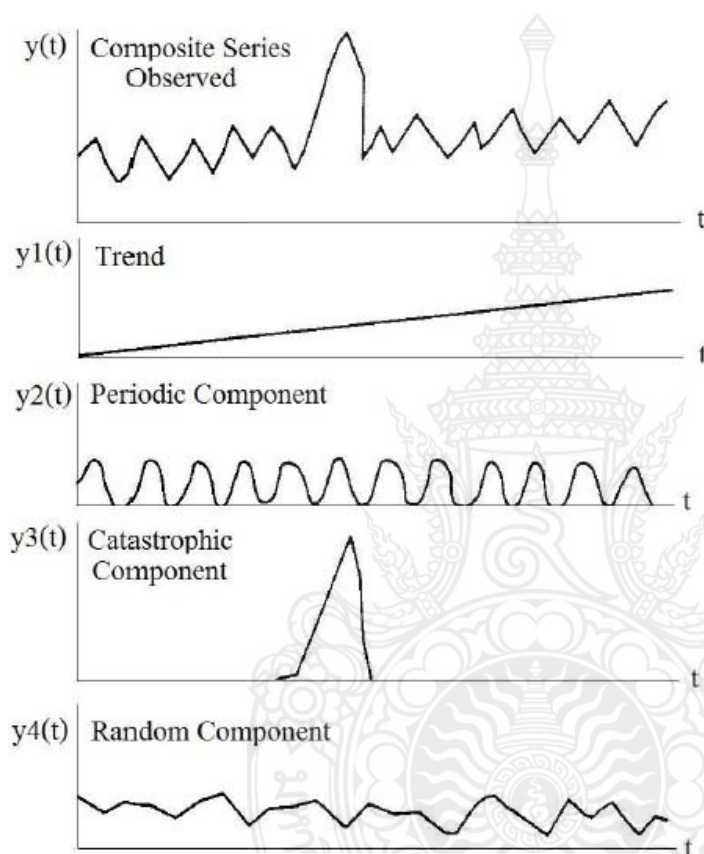
$$Y(t) = Y_1(t) + Y_2(t) + Y_3(t) + Y_4(t) \quad (1)$$

$$Y_1(t) = \text{Trend}$$

$$Y_2(t) = \text{Periodic Component}$$

$Y_3(t)$ = Catastrophic Component

$Y_4(t)$ = Random Component

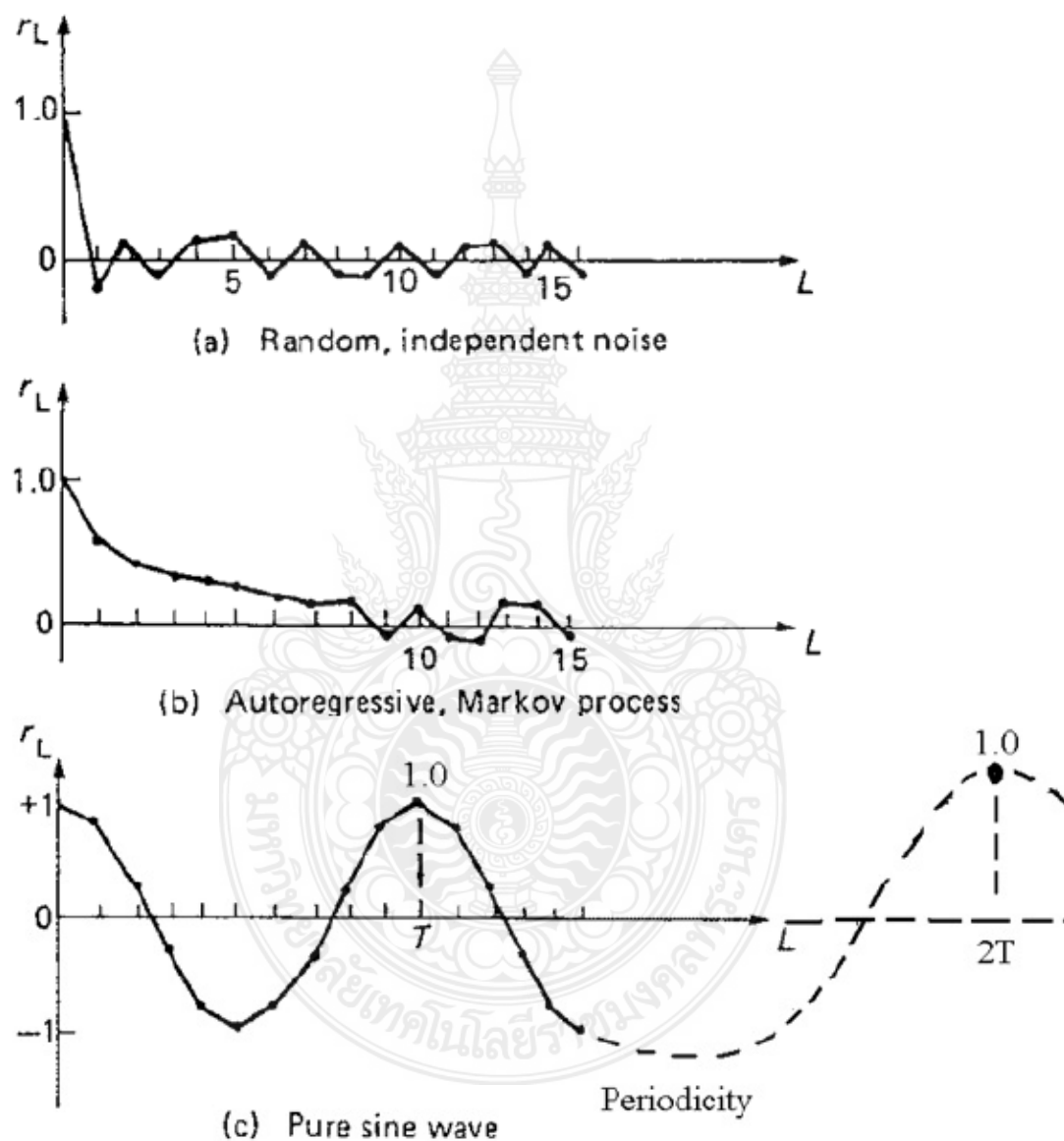


ภาพที่ 4 Components of Time Series

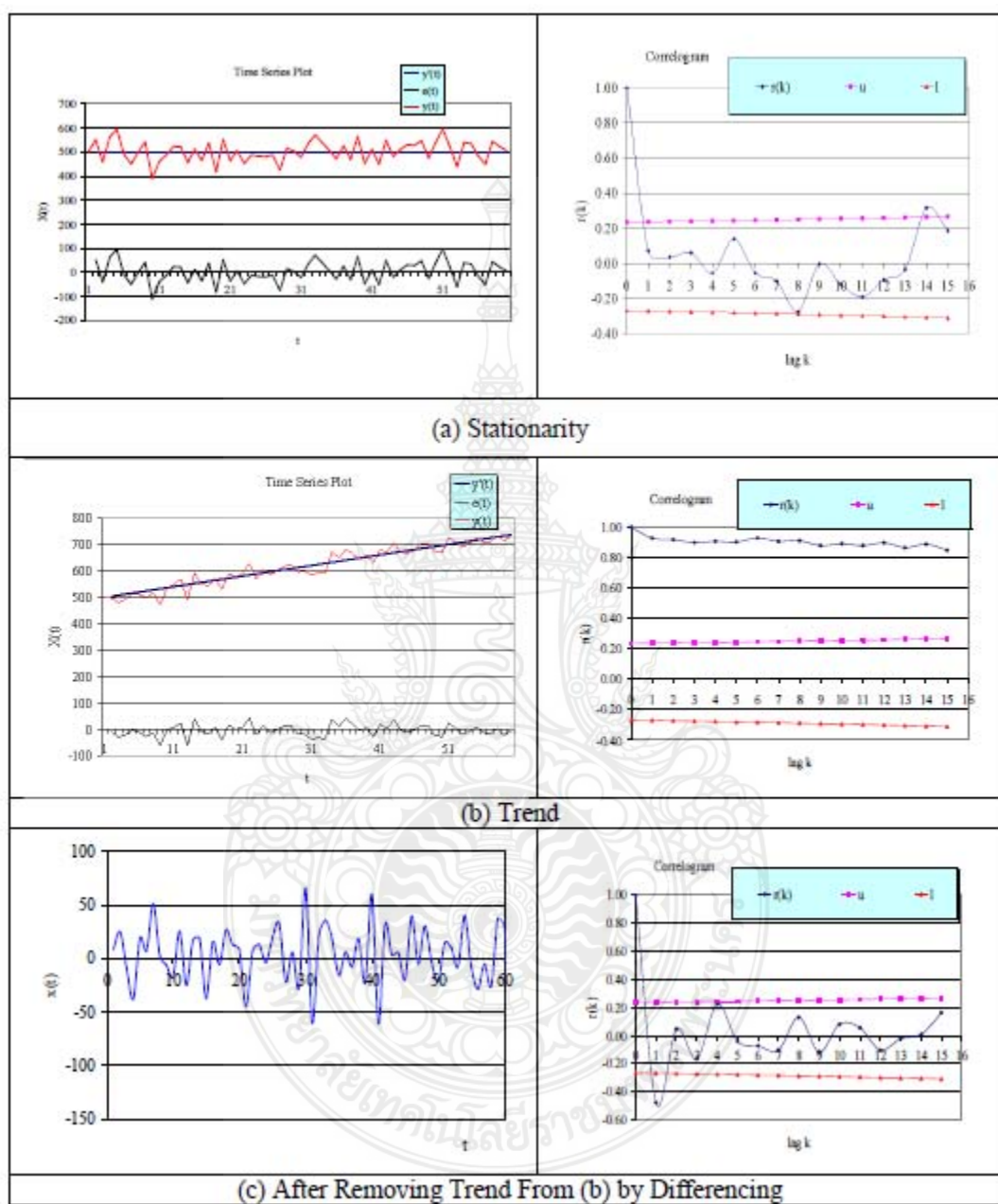
Autocorrelation Function (ACF) หรือเรียกอีกอย่างหนึ่ง คือ Correlogram ของกระบวนการสโตแคสติกแบบต่างๆ แสดงอยู่ในภาพที่ 5 ส่วนภาพที่ 6 แสดงการพล็อตเปรียบเทียบระหว่างอนุกรมเวลาและ ACF ของอนุกรมเวลาแบบต่างๆ ACF คือคุณสมบัติที่สำคัญของ Periodic Time Series

เมื่อวิเคราะห์อนุกรมเวลาใน Time Domain ส่วน Spectral Density Function คือ คุณสมบัติที่สำคัญของ Periodic Time Series เมื่อวิเคราะห์อนุกรมเวลาใน Frequency Domain ภาพที่ 7 แสดง ACF

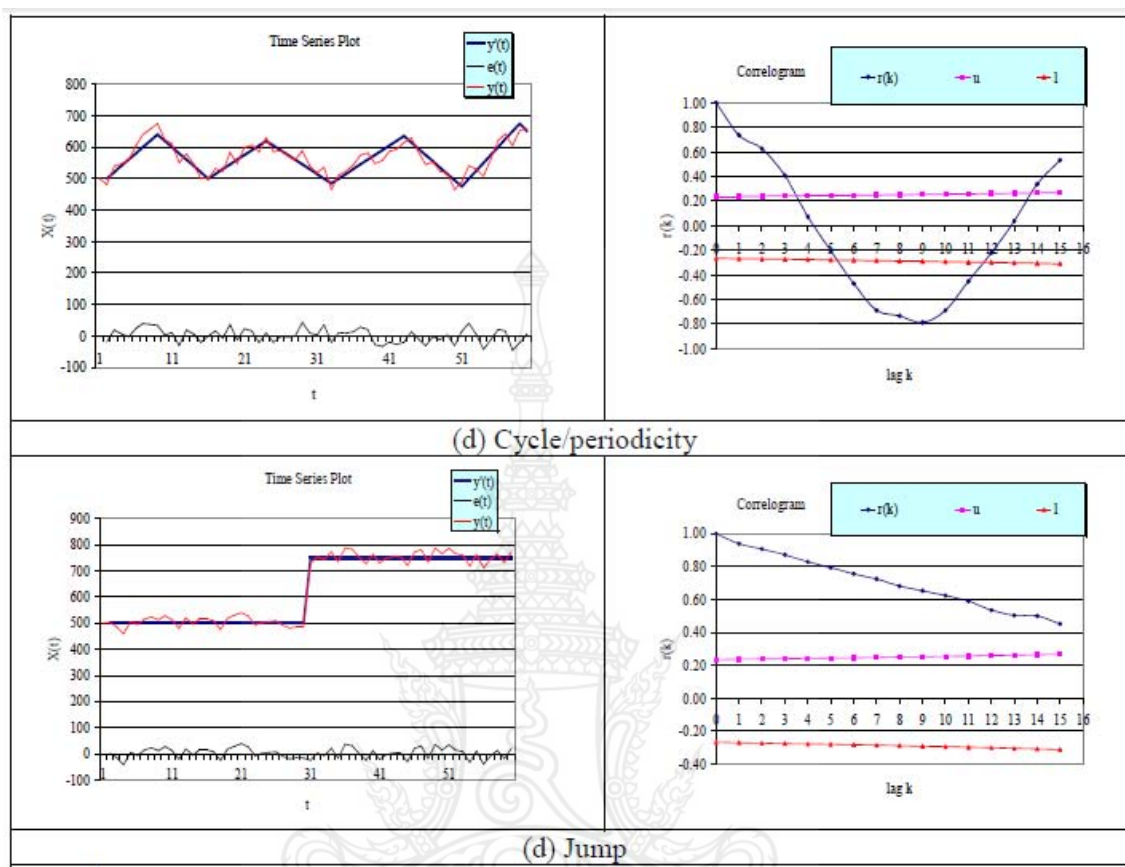
และ Spectral Density Function ของ Periodic Time Series ที่มีการ Remove Periodic Components ที่มี Frequency ต่างๆ



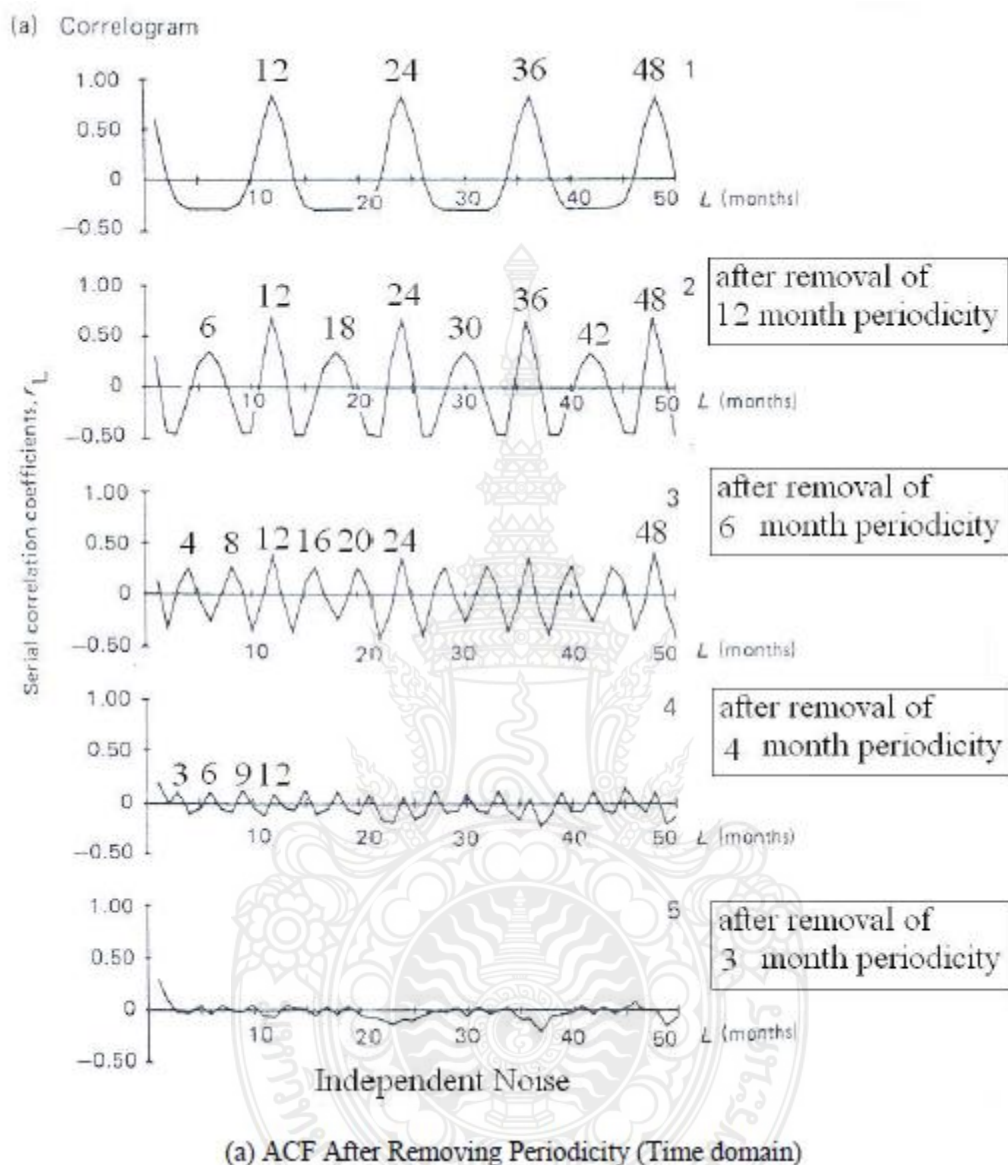
ภาพที่ 5 Correlogram or ACF Showing Independent, Markov Process and Periodicity



ภาพที่ 6 Characteristics of Time Series and Correlogram for Different Types of Time Series



ภาพที่ 6 (ต่อ)



ภาพที่ 7 ACF and Spectral Density Function after Removing Periodicity

1.3 คุณสมบัติพื้นฐานของกระบวนการสุ่ม (Basic Properties of Stochastic Process)

ให้ $\{X_t\}$ = เซตของตัวแปรสุ่ม หรือ Stochastic Process

$$E(X_t) = \mu_t \text{ โดยที่ } t=1,2,\dots,N$$

$$\text{VAR}(X_t) = \sigma_t^2 \text{ โดยที่ } t=1,2,\dots,N$$

$$\text{Cov}(X_t, X_{t-k}) = \sigma_{X_t, X_{t-k}} \text{ โดยที่ } t=1,2,\dots,N \text{ เมื่อ } \sigma_{X_t, X_{t-k}} \text{ เป็น Lag } k \text{ Autocovariance}$$

$$\rho_k = \frac{\text{Cov}(X_t, X_{t-k})}{\sqrt{\text{VAR}(X_t)\text{VAR}(X_{t-k})}}$$

เมื่อ ρ_k คือ Lag k ACF

ACF หรือ Correlogram หรือ Serial Correlation Coefficients คือ พารามิเตอร์ที่ใช้วัดความสัมพันธ์เชิงเส้น (Linear Dependence) ของกระบวนการสโตแคสติก

ตามที่ได้กล่าวมาแล้ว กระบวนการสโตแคสติกแบบ Stationary คือ กระบวนการที่พารามิเตอร์ μ, σ^2, p ไม่เปลี่ยนแปลงไปตามเวลา ซึ่งสามารถแบ่งออกได้เป็น 2 แบบตามคุณสมบัติของพารามิเตอร์ คือ First Order Stationary และ Second Order Stationary ดังตารางที่ 1

ตารางที่ 1 Different Types of Stationary Processes

First Order Stationary Process	$\mu_t = \mu = \text{constant}$
Second Order Stationary Process	$\mu_t = \mu = \text{constant}$ and $\sigma_t^2 = \sigma^2 = \text{constant}$
- Weak Stationary (weak sense)	... second order stationary
- Strong Stationary (strong sense)	... second order stationary + other parameters such as correlation coefficient are not varied with time

กระบวนการสโตแคสติกอาจเป็น Stationary เมื่อพิจารณาจากพารามิเตอร์ตัวหนึ่ง แต่อาจเป็น Non-Stationary เมื่อพิจารณาจากพารามิเตอร์ตัวอื่น

สรุป กระบวนการสโตแคสติกสามารถแบ่งออกได้เป็นประเภทต่างๆ ดังตารางที่ 2

ตารางที่ 2 Classification of Stochastic Process

Deterministic	vs.	Stochastic
Discrete	vs.	Continuous
Independent	vs.	Dependent
Stationary	vs.	Non-Stationary

1.4 แบบจำลองอนุกรมเวลาแบบง่าย (Simple Time Series Model)

ให้ x_t = Independent Time Series with $f(x; \theta)$

$$x_t = \mu + \sigma \varepsilon_t ; t = 1, 2, \dots \quad (2)$$

ε_t = Independent Series

โดยที่ $\varepsilon_t = \frac{x_t - \mu}{\sigma}$ ซึ่งมี Mean = 0 และ Standard Deviation = 1

μ, σ = Constant parameters (do not vary with time)

แบบจำลองจะเป็น Stationary parameter ถ้าพารามิเตอร์มีค่าคงที่

x_t = Dependent Time Series

ε_t = Dependent Series ซึ่งมี Mean = 0 และ Standard Deviation = 1

$$\varepsilon_t = \frac{x_t - \mu}{\sigma}$$

$$\varepsilon_t = \phi \varepsilon_{t-1} + \zeta_t \quad (3)$$

ζ_t = Independent Series ซึ่งมี Mean = 0 และ Variance = $(1 - \phi^2)$

สมการที่ 3 คือ ภาพแบบจำลองสโตแคสติกแบบ Autoregressive ซึ่งเป็นแบบจำลองแบบง่าย ดู Yevjevich (1963), Fiering (1967)

แบบจำลองสโตแคสติกที่ซับซ้อนมากขึ้น คือ Autoregressive and Moving Average (ARMA) ดู O'Connor (1976)

Joint Physical and Statistical Analysis ของแบบจำลอง

AR(1)-First Order Autoregressive Model

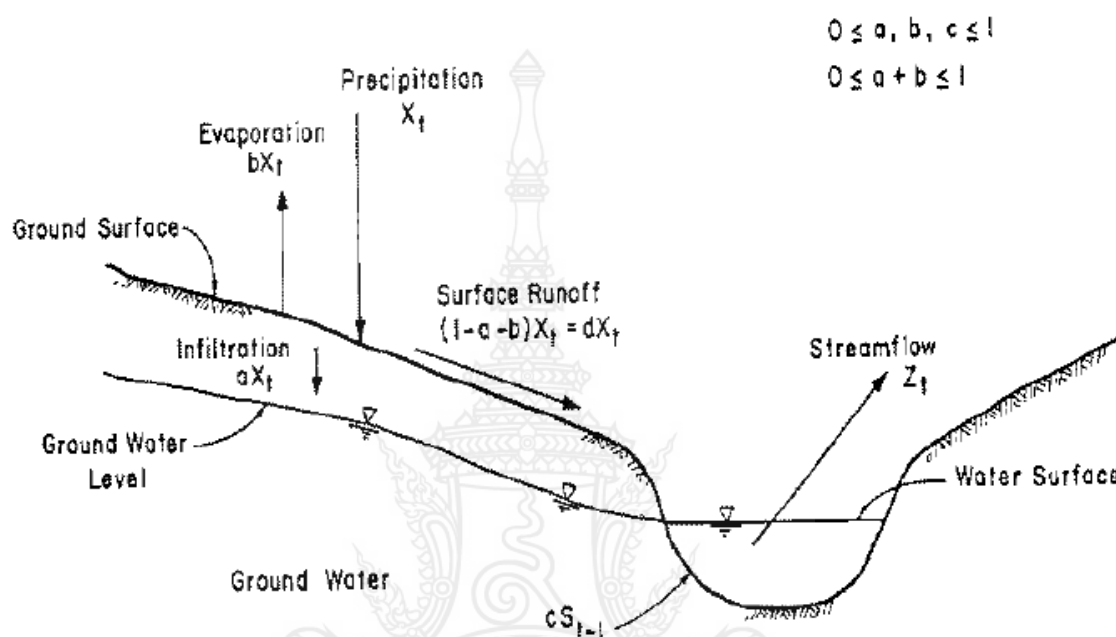
ให้ Y_t = Time Dependent Annual Runoff Series

$$Y_t - \bar{Y} = \phi(Y_t - \bar{Y}) + \varepsilon_t$$

ϕ = Autoregressive Coefficient

ε_t = Independent Statistical Component or White Noise

Firring (1968) ได้พิสูจน์ให้เห็นว่า อัตราการไหลของน้ำในทางน้ำสามารถอธิบายได้ด้วยแบบจำลอง ARMA(1,1,0) ดังภาพที่ 8



ภาพที่ 8 Conceptual Representation of the Precipitation-Stream Flow Process
(Salas and Smith. 1980a)

1.5 ค่าสถิติหลักของแบบจำลองสโตแคสติกแบบง่าย (Main Statistics of Simple Stochastic Model)

1) Mean (μ) Less uncertain

2) Standard Deviation (σ) Less uncertain

3) Skewness Coefficient (γ) High uncertain

4) Autocorrelation Coefficient (ρ_k) Very uncertain for small N

5) Other statistics of hydrologic time series

range (storage capacity)

run (drought)

rescale range

1.6 วิธีการจำลองสถานการณ์ทางสโตแคสติก (Stochastic Simulations)

จะกล่าวถึงแบบจำลองของอนุกรมเวลาทั้ง Annual Time Series และ Periodic Tim Series
 ดังแสดงในตารางที่ 3

ตารางที่ 3 Important Stochastic Models

1) AR(p)
2) ARMA(p,q)
3) ARIMA(p,d,q), Seasonal ARIMA(P,D,Q) _w , Multiplicative ARIMA(p,d,q)x(P,D,Q) _w
4) Multivariate AR(p)
5) Disaggregate

ข้อจำกัดของแบบจำลอง (Limitation of Model)

- 1) ไม่สามารถ Reproduce Short Term Dependence
- 2) ไม่สามารถ Reproduce Long Term Dependence

3) การประมาณค่าพารามิเตอร์มีความยุ่งยาก

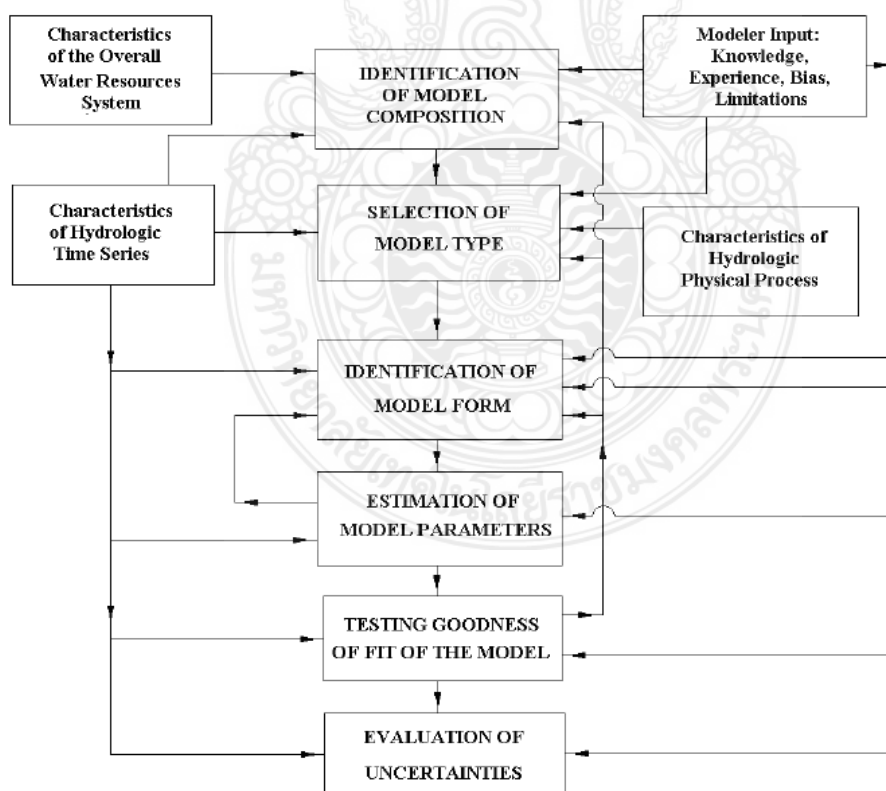
4) มีข้อจำกัดในการ Generate ข้อมูลใหม่จำนวนมากๆ

5) ขาดพื้นฐานทางด้านกายภาพ (Physical Basis)

6) มีพารามิเตอร์มากเกินไป

1.6.1 วิธีทางสมมาตร (Systematic Approach)

สำหรับการจำลองอนุกรมเวลาทางอุทกวิทยา Systematic Approach สำหรับ
การจำลองอนุกรมเวลาทางอุทกวิทยา ประกอบด้วย 6 ขั้นตอนดังนี้ (ดูภาพที่ 9)



ภาพที่ 9 Systematic Approach for Time Series Modeling

1) การหาส่วนประกอบของแบบจำลอง (Identification of Model Composition)

- a. Univariate หรือ Multivariate Model
- b. Combination of Univariate and Disaggregation
- c. Combination of Multivariate and Disaggregation
- d. ส่วนประกอบของแบบจำลองขึ้นอยู่กับคุณลักษณะของระบบแหล่งน้ำ คุณลักษณะของอนุกรมเวลา และความเห็นของผู้สร้างแบบจำลอง (Modeler's Input)
- e. ยกตัวอย่าง เช่น ถ้าต้องการสังเคราะห์อนุกรมเวลาของปริมาณน้ำที่ไหลเข้าระบบอ่างเก็บน้ำรายเดือน ทางเลือกในการสร้างแบบจำลองอาจเป็น

(1) จำลองปริมาณ น้ำที่ไหลเข้าระบบอ่างเก็บน้ำรายเดือนด้วย Multivariate Model หรือ

(2) จำลองปริมาณน้ำที่ไหลเข้าระบบอ่างเก็บน้ำรายปีเป็นรายเดือน หรือ

(3) ใช้ Univariate Model จำลองปริมาณน้ำที่ไหลเข้าแต่ละอ่างเก็บน้ำรายเดือนแบบเป็นอิสระแก่กัน กรณีเช่นนี้จะใช้เมื่อค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างอ่างเก็บน้ำ (Cross Correlation Coefficient) ไม่มีนัยสำคัญทางสถิติ

2) การเลือกประเภทแบบจำลอง (Selection of Model Type)

- a. AR

b. ARMA....เหมาะกับอนุกรมเวลาซึ่ง ACF ที่ค่าลดลงอย่างช้าๆ (Low Decaying Correlation)

c. ARIMA

d. อื่นๆ

3) การหาพหุลักษณะของแบบจำลอง (Identification of Model Form)

a. หาลำดับ (Orders) ของแบบจำลอง เช่น ค่า p และ q

b. ตรวจสอบว่าอนุกรมเวลามีลักษณะเบ้ (Skewed) หรือไม่ ถ้าเบ้ต้องตรวจสอบต่อไปว่าอนุกรมเวลาเป็นแบบ Constant หรือ Periodic ถ้าเป็นแบบ Periodic ต้องดูว่าควรใช้ Fourier Series วิเคราะห์หรือไม่

4) การประมาณค่าพารามิเตอร์แบบจำลอง (Estimation of Model Parameters)

a. Method of Moments

b. Method of Maximum Likelihood

5) การทดสอบความเหมาะสมของแบบจำลอง (Testing Goodness of Fit of the Model)

a. ตรวจสอบพิสูจน์ (Verify) ว่าแบบจำลองเป็นไปตามสมมติฐานที่กำหนดหรือไม่ เช่น Residual (ζ_t) เป็นอิสระและมีการแจกแจงปกติ (Independent and Normality)

b. ตรวจสอบพิสูจน์ (Verify) ว่าแบบจำลองเป็นตัวแทนที่เหมาะสมของอนุกรมเวลาของค่าสังเกต (Historical Time Series)

(1) ตรวจสอบ Model Correlogram เปรียบเทียบกับ Historical Correlogram

(2) ตรวจสอบค่าสถิติของอนุกรมเวลาที่สังเคราะห์ขึ้นแล้วเปรียบเทียบกับค่าสถิติกับอนุกรมเวลาของค่าสังเกต เช่น ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ ค่าสหสัมพันธ์ ค่าสถิติที่เกี่ยวข้องกับปริมาณน้ำในอ่างเก็บน้ำ (Storage Related Statistics) และค่าสถิติที่เกี่ยวข้องกับความแห้งแล้ง (Drought Related Statistics)

c. ถ้าการตรวจสอบความเหมาะสมไม่ผ่านเกณฑ์ในข้อ a และ b ต้องปรับเปลี่ยนภาพลักษณะ และชนิดของแบบจำลอง

6) การประเมินความไม่แน่นอน (Evaluation of Uncertainties)

a. ความไม่แน่นอนของแบบจำลอง (Model Uncertainty)

(1) ทดสอบความแตกต่างระหว่างค่าสถิติที่คำนวณจากอนุกรมเวลาที่สังเคราะห์ขึ้นจากแบบจำลองทางเลือกต่างๆ

b. ความไม่แน่นอนของแบบพารามิเตอร์ (Parameter Uncertainty)

หาการแจกแจงความน่าจะเป็นของพารามิเตอร์ และใช้แบบจำลองกับพารามิเตอร์ที่สุ่มจากการแจกแจงนั้น (วรารุช, 2554)

วิธีการประมาณค่าพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method: MLM)

เมื่อเราต้องการวัดค่าที่แท้จริงของปริมาณบางอย่าง (x_T)

- ทำได้โดยการวัดค่าของปริมาณนั้น ซ้ำๆ กันหลายครั้ง $\{x_1, x_2, \dots, x_n\}$
- วิธีที่นิยมใช้ในการประมาณค่า x_T จากค่าที่วัดได้ คือ การนำมาหาค่าเฉลี่ย ดังนี้

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i$$

- กำหนดให้ $x_T = \mu_x$
- วิธีการนี้ เป็นวิธีที่สมเหตุสมผลหรือไม่
- MLM เป็นวิธีการทั่วไปที่ใช้เพื่อประมาณค่าพารามิเตอร์ของสิ่งที่สนใจจากชุด

ข้อมูลที่มี

ข้อตกลงเบื้องต้นของวิธีการน่าจะเป็นสูงสุด

- กำหนดให้กลุ่มตัวอย่างค่าวัด $\{x_1, x_2, \dots, x_n\}$ จากข้อมูลทั้งหมด N
- กำหนดให้ $f(x, \alpha)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น (Probability Distribution Function) ของตัวแปรสุ่ม X
- วิธี MLM จะหาค่า α ที่ทำให้ความน่าจะเป็นของค่าวัด x_i ที่ได้มีค่ามากที่สุด

เราจะใช้วิธี MLM ได้อย่างไร

- กำหนดให้ความน่าจะเป็นของค่าวัด x_1 คือ $f(x_1, \alpha)dx$
- กำหนดให้ความน่าจะเป็นของค่าวัด x_2 คือ $f(x_2, \alpha)dx$

- กำหนดให้ความน่าจะเป็นของค่าวัด x_n คือ $f(x_n, \alpha)dx$
- ถ้าค่าวัดแต่ละค่าอิสระต่อกัน ความน่าจะเป็นร่วมของค่าวัดที่ได้ คือ

$$L = f(x_1, \alpha)dx \cdot f(x_2, \alpha)dx \cdot \dots \cdot f(x_n, \alpha)dx = f(x_1, \alpha) \cdot f(x_2, \alpha) \cdot \dots \cdot f(x_n, \alpha)dx^n$$

- ถ้าเราดึงเทอม dx^n ออก จะเหลือเฉพาะค่าคงที่ซึ่งเป็นสัดส่วน (proportionality constant) เรียกว่า ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) $L = \prod_{i=1}^N f(x_i, \alpha)$

- เมื่อต้องการหาค่า α ที่ทำให้ได้ค่า L ที่มีค่าสูงสุด ให้ทำดังนี้

$$\left. \frac{\partial L}{\partial \alpha} \right|_{\alpha=\alpha^*} = 0$$

- สำหรับค่า L และ $\ln L$ จะให้ค่าสูงสุดที่ตำแหน่งเดียวกัน การหาค่าสูงสุดของ $\ln L$ จะทำได้ง่ายกว่า เพราะ $\ln L$ สามารถเปลี่ยนให้อยู่ในรูปของผลบวกได้ จึงทำให้ง่ายต่อการคำนวณ $\ln L = \sum_{i=1}^N \ln f(x_i, \alpha)$

- เงื่อนไขการหาค่าสูงสุดตัวใหม่จะกลายเป็น

$$\left. \frac{\partial L}{\partial \alpha} \right|_{\alpha=\alpha^*} = \sum_{i=1}^N \left. \frac{\partial}{\partial \alpha} \ln f(x_i, \alpha) \right|_{\alpha=\alpha^*} = 0$$

- ค่า α อาจจะเป็นชุดของค่าพารามิเตอร์ (เช่น เป็นความชันและระยะแกน) หรือเป็นค่าเดียวก็ได้
- สมการที่ใช้คำนวณหาช่วงของค่า α อาจได้จากสมการเชิงเส้นอย่างง่ายจับคู่กับสมการไม่เป็นเชิงเส้น

ตัวอย่าง

- กำหนดให้ $f(x, \alpha)$ มีการแจกแจงปกติ
- ให้ $\alpha = \mu$ เป็นค่าเฉลี่ยของการแจกแจงปกติ
- เราต้องการประมาณค่าที่ดีที่สุดของ α จากชุดข้อมูลที่เราวัดได้ คือ $\{x_1, x_2, \dots, x_n\}$
- ให้ข้อกำหนดเบื้องต้น ข้อมูลที่ได้มีการกระจายเท่ากัน คือมีส่วนเบี่ยงเบนมาตรฐานเท่ากัน เป็น σ ดังนั้น ข้อมูลแต่ละตัวมีการแจกแจงดังนี้

$$f(x_i, \alpha) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i - \alpha)^2}{2\sigma^2}}$$

- ฟังก์ชันภavnน่าจะเป็นสูงสุดของปัญหานี้ คือ

$$L = \prod_{i=1}^N f(x_i, \alpha) = \prod_{i=1}^N \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i - \alpha)^2}{2\sigma^2}} = \left[\frac{1}{\sigma\sqrt{2\pi}} \right]^n e^{-\frac{(x_1 - \alpha)^2}{2\sigma^2}} e^{-\frac{(x_2 - \alpha)^2}{2\sigma^2}} \dots e^{-\frac{(x_n - \alpha)^2}{2\sigma^2}}$$

$$= \left[\frac{1}{\sigma\sqrt{2\pi}} \right]^n e^{-\sum_{i=1}^n \frac{(x_i - \alpha)^2}{2\sigma^2}}$$

- หาค่า α ที่ให้ค่า Log Likelihood function มีค่ามากที่สุด ดังนี้

$$\frac{\partial \ln L}{\partial \alpha} = \frac{\partial}{\partial \alpha} \left[n \ln \left(\frac{1}{\sigma \sqrt{2\pi}} \right) - \sum_{i=1}^n \frac{(x_i - \alpha)^2}{2\sigma^2} \right] = 0$$

$$\frac{\partial}{\partial \alpha} \sum_{i=1}^n (x_i - \alpha)^2 = 0$$

$$2 \sum_{i=1}^n (x_i - \alpha)(-1) = 0$$

$$\sum_{i=1}^n x_i = n\alpha$$

จะได้ค่าเฉลี่ย $\alpha = \frac{\sum_{i=1}^n x_i}{n}$

- ถ้า σ ของข้อมูลแต่ละค่ามีค่าแตกต่างกัน

- ค่า α จะเป็นค่าเฉลี่ยถ่วงน้ำหนัก (Weighted average) ดังนี้

$$\alpha = \frac{\sum_{i=1}^n \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

1.6.2 วิธีการพยากรณ์แบบ เบย์ (Bayesian) (Robert, 2001), (Congdon, 2006), West, 1997)

ตัวแบบเบย์สร้างจาก Likelihood, $p(\mathbf{Y} | \boldsymbol{\theta})$, และ Prior, $\pi(\boldsymbol{\theta})$, เมื่อ $\mathbf{Y}$ คือตัวแปรสุ่มที่สังเกตค่าได้ และ $\boldsymbol{\theta}$ คือค่าพารามิเตอร์ที่สังเกตค่าไม่ได้ การแจกแจงร่วม (Joint Distribution) ของ $\boldsymbol{\theta}$ กับ $\mathbf{Y}$ สามารถเขียนอยู่ในรูป

$\pi(\boldsymbol{\theta}, \mathbf{Y}) = p(\mathbf{Y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})$ และ Posterior ที่สร้างจากกฎของเบย์คือ

$$\pi(\boldsymbol{\theta} | \mathbf{Y}) = \frac{p(\boldsymbol{\theta}, \mathbf{Y})}{p(\mathbf{Y})} = \frac{p(\mathbf{Y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})}{p(\mathbf{Y})} \quad (4)$$

โดยที่

$p(\mathbf{Y}) = \sum_{\boldsymbol{\theta}} p(\mathbf{Y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})$ เมื่อ $\boldsymbol{\theta}$ เป็นตัวแปรสุ่มชนิดไม่ต่อเนื่อง (Discrete) และ

$p(\mathbf{Y}) = \int p(\mathbf{Y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ เมื่อ $\boldsymbol{\theta}$ เป็นตัวแปรสุ่มชนิดต่อเนื่อง (Continuous)

เนื่องจาก $p(\mathbf{Y})$ เป็นฟังก์ชันของ $\mathbf{Y}$ ซึ่งไม่ขึ้นอยู่กับ $\boldsymbol{\theta}$ จึงถูกพิจารณาว่าเป็นค่าคงที่และสามารถเขียน $\pi(\boldsymbol{\theta} | \mathbf{Y})$ อยู่ในภาพ $\pi(\boldsymbol{\theta} | \mathbf{Y}) \propto p(\mathbf{Y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})$ นั่นคือ $\pi(\boldsymbol{\theta} | \mathbf{Y})$ เป็นสัดส่วนกับผลคูณของ Likelihood กับ Prior

ตัวแบบที่ซับซ้อนสามารถใช้ตัวแบบเบย์แก้ปัญหาได้ เช่น ใช้ตัวแบบเบย์ที่มี 3 ชั้น ได้แก่ ขั้นตอนที่ 1 ระบุการแจกแจงของตัวแปรสุ่มที่สังเกตค่าได้เมื่อกำหนดพารามิเตอร์ให้ ขั้นตอนที่ 2 ระบุการแจกแจงของพารามิเตอร์เมื่อกำหนดไฮเปอร์พารามิเตอร์ให้ และขั้นตอนที่ 3 ระบุการแจกแจงของไฮเปอร์พารามิเตอร์ในทำนองเดียวกัน จำนวนขั้นตอนอาจมีมากกว่า 3 ได้ ตัวแบบเบย์สามารถเพิ่มความแกร่ง (Robustness) ให้กับตัวประมาณแบบเบย์ได้ เนื่องจากความไม่แน่นอน (Uncertainty) ถูกนำมาคิดไว้ในขั้นตอนของการแจกแจงของ Prior นอกจากนี้วิธีการของเบย์ยังทำให้การประมาณค่าพารามิเตอร์ใน Posterior ง่ายขึ้น โดยใช้การจำลองสถานการณ์ (Simulation) การจำลองสถานการณ์ที่ใช้กันอย่างแพร่หลายคือ วิธีเชิ่งตัวเลข Markov Chain Monte Carlo (MCMC)

สำหรับตัวอย่างเพื่อให้เห็นภาพรวมของวิธีเบย์ จะขอยกตัวอย่างตัวแบบที่มีความซับซ้อนจึงใช้วิธีการของเบย์ในการแก้ปัญหา ตัวแบบคือ

$$Y_{it} = \beta_0 + \beta_1 Z_{it.1} + \beta_2 Z_{it.2} + \dots + \beta_p Z_{it.p} + \gamma_i \Delta W(t | \alpha_i, \delta_i) + \gamma_i X_{it} + \varepsilon_{it}$$

เมื่อ Y_{it} แทนราคา หรือปริมาณผลผลิตของพืชชนิดที่ i ในช่วงเวลา t , $i=1,...,m$ และ $t=1,...,T_i$ โดยที่

$$\varepsilon_{it} \sim N\left(0, [\gamma_i(1+3\zeta_{it})\sigma_\varepsilon]^2\right)$$

$$Y_{it} \sim N\left(\beta_0 + \beta_1 Z_{it,1} + \beta_2 Z_{it,2} + \dots + \beta_p Z_{it,p} + \gamma_i \Delta W(t | \alpha_i, \delta_i) + \gamma_i X_{it}, [\gamma_i(1+3\zeta_{it})\sigma_\varepsilon]^2\right)$$

Prior (มีหลายระดับ) คือ

$$1) p(\sigma_\varepsilon) \propto Unif(0, \infty), \quad p(\beta_i) \propto \text{constant}$$

2) ข้อมูลผิดปกติ (Outliners)

$$\zeta_{it} \sim \text{Bern}(0.05)$$

3) ราคาหรือปริมาณผลผลิตรวมทุกช่วงเวลา

$$\gamma_i \sim N(g_i, \sigma_\gamma^2), \quad p(\sigma_\gamma) \propto Unif(0, \infty)$$

$$g_j \sim N(\mu_g, \sigma_g^2), \quad p(\mu_g) \propto 1, \quad p(\sigma_g) \propto Unif(0, \infty)$$

$$S_i \sim N(\gamma_i, [0.2\gamma_i]^2)$$

4) Autoregression ที่ซ่อนเร้นอยู่ (Latent Autoregression)

$$X_{it} \sim N(\lambda_{11} X_{it-1} + \lambda_{12} X_{it-2}, \sigma_x^2), \quad \sigma_x = 0.8\sigma_\varepsilon$$

$$(\lambda_{11}, \lambda_{12})^T \sim N(\boldsymbol{\mu}_\lambda, \boldsymbol{\Sigma}_\lambda), p(\boldsymbol{\mu}_\lambda, \boldsymbol{\Sigma}_\lambda) \propto |\boldsymbol{\Sigma}_\lambda|^{-2}$$

$$(X_{i0}, X_{i,-1})^T \sim N(\boldsymbol{\mu}_{x_0}, \boldsymbol{\Sigma}_{x_0}), \boldsymbol{\mu}_{x_0} = (0, 0)^T, \boldsymbol{\Sigma}_{x_0} = \text{diag}(2, 2).$$

5) พารามิเตอร์อื่นๆ

$$\alpha_i \sim N(a_i, \sigma_\alpha^2), p(\sigma_\alpha) \propto \text{Unif}(0, \infty)$$

$$\delta_i \sim N(d_i, \sigma_\delta^2), p(\sigma_\delta) \propto \text{Unif}(0, \infty)$$

$$a_j \sim N(\mu_a, \sigma_a^2), p(\mu_a) \propto 1, p(\sigma_a) \propto \text{Unif}(0, \infty)$$

$$d_j \sim N(\mu_d, \sigma_d^2), p(\mu_d) \propto 1, p(\sigma_d) \propto \text{Unif}(0, \infty)$$

การประมาณค่าพารามิเตอร์ในตัวแทนดังกล่าวมีขั้นตอนดังนี้

a. สร้าง Likelihood จากการแจกแจงของ Y_{it}

b. สร้าง Posterior จากผลคูณของ Likelihood กับ prior ทุกตัว

c. จำลองสถานการณ์ด้วยวิธีการ MCMC โดยใช้การเขียนโปรแกรมใน Open bugs และ R หรือโปรแกรมคณิตศาสตร์ต่างๆ

Markov Chain Monte Carlo (MCMC) และ Gibbs sampling (Robert, 2004)

MCMC เป็นวิธีเชิงตัวเลขที่ใช้สำหรับสร้างข้อมูลจากการแจกแจงที่มีมิติขนาดใหญ่ ในตัวแบบเบย์ เป้าหมายหลักคือการสร้าง $\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \dots$ ของ Posterior จากห่วงโซ่มาร์คอฟ (Markov Chain) โดยเริ่มจาก Initial state $\theta^{(0)}$ และเมื่อห่วงโซ่ที่ในการวนซ้ำรอบที่ T ชุดของ $\theta^{(0)}, \dots, \theta^{(T)}$ จะถูกตัดทิ้ง เรียกว่า ช่วงของการ burn-in และ $\theta^{(T+1)}, \theta^{(T+2)}, \theta^{(T+3)}, \dots$ เป็นห่วงโซ่ที่คงที่ (Stationary) แล้ว ที่สร้างมาจาก Posterior มีหลายวิธีในการสร้าง MCMC แต่วิธีที่ใช้กันอย่างแพร่หลายคือ Gibbs sampling

Gibbs sampling (Geman and Geman, 1984)

เป็นวิธีการสร้าง MCMC จากการสุ่มตัวอย่างแบบวนซ้ำจากการแจกแจงแบบมีเงื่อนไขของแต่ละพารามิเตอร์เมื่อรู้พารามิเตอร์ที่เหลือทั้งหมดและข้อมูล สมมติว่า Posterior คือ $\pi(\theta | Y)$ ที่มีมิติขนาด k โดยที่ Y แทนข้อมูลที่สังเกตค่าได้ และสำหรับแต่ละ θ_i ของ θ การแจกแจงแบบมีเงื่อนไขของแต่ละพารามิเตอร์เมื่อรู้พารามิเตอร์ที่เหลือทั้งหมดและข้อมูลคือ $\pi(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k, Y) = \pi(\theta_i | \theta_{-i}, Y)$ Gibbs sampling เป็นกระบวนการวนซ้ำมีขั้นตอนดังนี้

- กำหนดค่าเริ่มต้น $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$ (5)

- รอบที่ i จะเป็นการเปลี่ยนสถานะจาก $\theta^{(i)}$ ไปเป็น $\theta^{(i+1)}$ มีขั้นตอนดังนี้

1. สุ่ม $\theta_1^{(i)}$ จาก $\pi(\theta_1 | \theta_2^{(i-1)}, \theta_3^{(i-1)}, \dots, \theta_k^{(i-1)}, Y)$

2. สุ่ม $\theta_2^{(i)}$ จาก $\pi(\theta_2 | \theta_1^{(i)}, \theta_3^{(i-1)}, \dots, \theta_k^{(i-1)}, Y)$

⋮
⋮

3. สุ่ม $\theta_k^{(i)}$ จาก $\pi(\theta_k | \theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_{k-1}^{(i)}, Y)$

ลำดับของการสุ่ม $\theta_1^{(1)}, \theta_2^{(2)}, \dots, \theta^{(T)}$ เป็นสถานะต่อเนื่องกันของ Markov Chain

การวัดความถูกต้องแม่นยำของการจำลองสถานการณ์ (Najafi and Tarazkar,2006)

การวัดความถูกต้องแม่นยำของการพยากรณ์นั้นเป็นการเลือกวิธีการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาในแต่ละชุด วิธีการวัดความถูกต้องแม่นยำของการพยากรณ์นั้นมีหลายวิธีแต่วิธีที่ใช้กันมากก็คือ เราจะใช้การพิจารณาจากค่าวัดความถูกต้อง ซึ่งต่างเป็นฟังก์ชันของค่าความคลาดเคลื่อน e_t โดยที่ e_t เป็นผลต่างของค่าจริง (Y_t) กับค่าพยากรณ์ ($\hat{Y}_t$) ณ เวลา t ดังนี้

1) Mean Squared Error (MSE)

$$\frac{\sum_{t=1}^n e_t^2}{n} = \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n} \quad (6)$$

วิธี MSE เป็นวิธีที่ใช้กันทั่วไป ข้อเสียของวิธีนี้คือไม่มีฐานการเปรียบเทียบ และถ้า MSE มีค่าสูงอาจเป็นเพราะมีความคลาดเคลื่อนสูง หรือขึ้นอยู่กับขนาดของข้อมูล

2) Mean Absolute Percentage Error (MAPE)

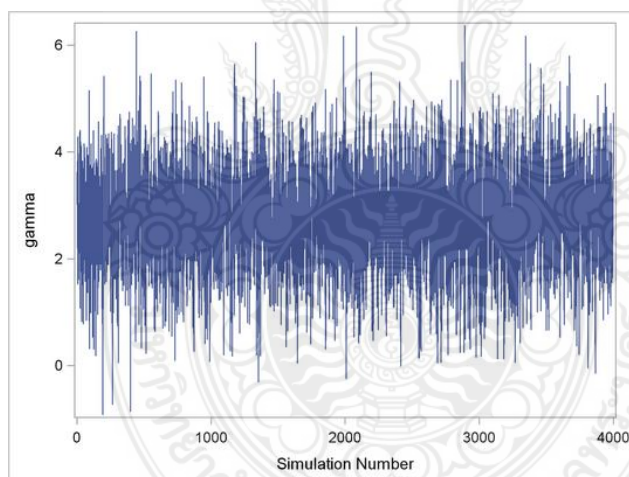
$$MAPE = \frac{\sum_{t=1}^n |e_t / Y_t|}{n} \times 100 \quad (7)$$

วิธี MAPE เป็นหนึ่งในวิธีที่ถูกยอมรับ และที่ใช้ในการเปรียบเทียบมากที่สุดสำหรับอนุกรมเวลา

3) Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{t=1}^n |e_t|}{n} = \frac{\sum_{t=1}^n |Y_t - \hat{Y}_t|}{n} \quad (8)$$

เมื่อค่า MSE (Mean Squared Error) MAPE (Mean Absolute Percentage Error) และ MAE (Mean Absolute Error) มีค่าต่ำ แสดงถึง วิธีการพยากรณ์นั้นมีความถูกต้องมาก สำหรับการตรวจสอบการลู่เข้าของ MCMC โดยดูที่ Trace plots Trace plots เป็นตัวชี้วัดตัวหนึ่งที่สามารถตรวจสอบการลู่เข้าของ MCMC ซึ่งมันจะ ช่วยบอกว่า chain ของพารามิเตอร์แต่ละตัวลู่หรือยังไม่ลู่เข้าสู่ stationary distribution และมันยังเป็นตัวช่วยบอกว่าอีกนานเท่าไรมันถึงจะลู่เข้า trace plots ยังสามารถบอกเราว่ามันลู่เข้าดีหรือยังไม่ดีอีกด้วย ดังตัวอย่างในภาพที่ 1 ที่แสดง trace plots ที่ลู่เข้าสู่ stationary distribution ที่ดี (SAS, 2011)



ภาพที่ 10 Trace plots ที่ลู่เข้าสู่ stationary distribution ที่ดีของ gamma

การประเมินประสิทธิภาพของตัวแบบ (Bernd, 2004)

ประสิทธิภาพของตัวประมาณจะถูกประเมินจาก 3 ตัว ที่นิยมใช้กันอยู่ทั่วไปซึ่งได้แก่ Relative Bias (RB) Mean Squared Error (MSE) และ the coverage probability (CP) ซึ่งตัวประเมินแต่ละตัวจะคำนวณจากเซตของข้อมูลที่เป็นอิสระกันจากการจำลองสถานการณ์ที่มาจากกระบวนการของ

MCMC โดยที่มี T ชุด โดยที่ $T = T_1, \dots, T_B$ และ S มีจำนวนที่ใหญ่พอ สามารถแสดงสูตรของตัวประเมินแต่ละตัวดังสมการข้างล่าง

$$mean = S^{-1} \sum_{s=1}^S T_s^{(k)} = \bar{T}^{(k)}, \quad (9)$$

$$bias = \bar{T}^{(k)} - \mu \quad (10)$$

$$SD = \sqrt{(S-1)^{-1} \sum_{s=1}^S (T_s^{(k)} - \bar{T}^{(k)})^2}, \quad (11)$$

$$MSE = S^{-1} \sum_{s=1}^S (T_s^{(k)} - \mu)^2 \approx SD^2 + bias^2 \quad (12)$$

$$RB(\hat{\mu}_Y) = \frac{1}{B} \sum_{b=1}^B \frac{\hat{\mu}_Y^{(b)} - \mu}{\mu} \quad (13)$$

$$MSE(\hat{\mu}_Y) = \frac{1}{B} \sum_{b=1}^B (\hat{\mu}_Y^{(b)} - \mu)^2 \quad (14)$$

$$CP(\hat{\mu}_Y) = \frac{1}{B} \sum_{b=1}^B I(\hat{\mu}_L^{(b)} < \mu < \hat{\mu}_U^{(b)}) \quad (15)$$

1.7 แบบจำลองต่างๆ

1) แบบจำลองการถดถอยเชิงพหุ (Multiple Regression Model) (Yan, 2009)

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \varepsilon_t \quad (16)$$

เมื่อ t แทนเวลา โดยที่ $t = 1, \dots, n$

Y_t แทน ตัวแปรตาม ณ เวลา t และ X_{it} แทนตัวแปรอิสระ ณ เวลา t โดยที่ $i = 1, \dots, k$

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ แทน ค่าคงที่ และ ε_t แทน ค่าความผิดพลาด ณ เวลา t และเป็นอิสระกัน โดย ε_t จะมีการแจกแจงแบบปกติ มีค่าเฉลี่ย เท่ากับ 0 และค่าความแปรปรวน คือ σ^2 สำหรับการพยากรณ์ ณ เวลา t จะได้

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_{1t} + \hat{\beta}_2 X_{2t} + \dots + \hat{\beta}_k X_{kt} + \varepsilon_t \quad (17)$$

เมื่อ $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ ได้จากการประมาณค่าโดยวิธี maximum likelihood

2) แบบจำลองค่าเฉลี่ยเคลื่อนที่ (Moving Average (MA) Model) (Bisgaard, 2011)

$$F_{t+1} = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t+1-n}}{N} \quad (18)$$

เมื่อ N = ขนาดของ moving average โดยที่ $N = 1, \dots, n$

ตัวอย่างเช่น

$$F_4 = \frac{y_3 + y_2 + y_1}{3}$$

$$F_5 = \frac{y_4 + y_3 + y_2 + y_1}{4}$$

เมื่อ

$$N = 1 \rightarrow F_2 = y_1, F_3 = y_2, \dots$$

$$N = n \rightarrow F_{n+1} = \bar{y}$$

ถ้า N เป็นฤดูกาล ดังนั้น จะสามารถกำจัดฤดูกาลไปได้ในตัว

สำหรับวิธี Moving Average ที่ได้กล่าวมาแล้ว ยังมีวิธี double MA, triple MA และ centered MA อีก เช่น 2x2 MA, 3x5 MA , 3x3x5 MA เป็นต้น

กรณีมี trend ในข้อมูล

ให้ใช้วิธี MA และ double MA โดยกำหนดให้มี moving average length ที่เท่ากันดังนี้

$$F_{t+m} = a_t + b_t m_t \quad \text{โดยที่ } m = 1, 2, 3, \dots \quad (19)$$

เมื่อ

$$a_t = 2S'_t - S''_t$$

$$b_t = \frac{2}{N-1} (S'_t - S''_t)$$

$$S'_t = \frac{1}{N} (y_t + y_{t-1} + y_{t-2} + \dots + y_{t+1-n})$$

$$S''_t = \frac{1}{N} (S'_t + S'_{t-1} + S'_{t-2} + \dots + S'_{t+1-n})$$

3) แบบจำลอง Exponential Smoothing (EXPS) Model (Montgomery, 2008)

Exponential smoothing เป็นการพยากรณ์โดยกำหนดน้ำหนักให้ค่าสังเกตในปัจจุบันมีค่ามากกว่าน้ำหนักของค่าสังเกตก่อนหน้านี้ มี 3 แบบคือ

a. Single exponential smoothing

ตัวแบบการพยากรณ์แบบ Single exponential smoothing นิยามดังนี้

ให้ S_t แทนค่าพยากรณ์ของค่าสังเกต Y_t ณ เวลา t เมื่อ $t=1, \dots, n$

$$S_t = \alpha Y_t + (1 - \alpha) S_{t-1}, \quad 0 \leq \alpha \leq 1 \quad (20)$$

ค่าพยากรณ์ S_1 ไม่เกิดขึ้นเนื่องจากไม่มีเทอม S_0 นอกเสียจากจะกำหนดค่าเริ่มต้นให้

S_0 เทคนิคนี้ใช้สำหรับข้อมูลที่ไม่มี trend และ seasonal ถ้าข้อมูลมี trend จะใช้ Double exponential smoothing และถ้ามีทั้ง trend และ seasonal จะใช้ Triple Exponential Smoothing

b. Double exponential smoothing หรือ Holt's method

ตัวแบบการพยากรณ์แบบ Double exponential smoothing ได้จากขยายตัวแบบของ Single exponential smoothing ออกไปดังนี้

$$S_t = \alpha Y_t + (1 - \alpha)(S_{t-1} + A_{t-1}), \quad 0 \leq \alpha \leq 1 \quad (21)$$

$$\text{โดยที่ } A_t = \beta(S_t - S_{t-1}) + (1 - \beta)A_{t-1}, \quad 0 \leq \beta \leq 1 \quad (22)$$

เมื่อ β คือค่าคงที่แสดง trend

c. Triple Exponential Smoothing หรือ Holt-Winters method

ตัวแบบการพยากรณ์แบบ Triple exponential smoothing ได้จากขยายตัวแบบของ Double exponential smoothing ออกไป มี 2 แบบคือ

1. Multiplicative Seasonal Model

$$S_t = \alpha(Y_t / B_{t-s}) + (1-\alpha)(S_{t-1} + A_{t-1}), 0 \leq \alpha \leq 1, \quad (23)$$

$$A_t = \beta(S_t - S_{t-1}) + (1-\beta)A_{t-1}, 0 \leq \beta \leq 1, \quad (24)$$

$$B_t = \gamma(Y_t / S_t) + (1-\gamma)B_{t-s}, 0 \leq \gamma \leq 1, \quad (25)$$

เมื่อ γ คือค่าคงที่แสดง seasonal และ s คือความยาวช่วงของ seasonal ใช้ตัวแบบการพยากรณ์นี้เมื่อข้อมูลมี seasonal เป็นแบบการคูณ (multiplicative seasonality) คือในเวลาเดียวกันของแต่ละฤดูกาล ค่าสังเกตจะเพิ่มขึ้นหรือลดลงเป็นค่าร้อยละ ตัวอย่างเช่นข้อมูลรายเดือน 5 ปี ความยาวช่วงของ seasonal คือ 12 และข้อมูลปรากฏให้เห็นว่าในเดือนธันวาคมของแต่ละปี ค่าสังเกตจะเพิ่มขึ้นร้อยละ 40 ไม่ใช่เพิ่มขึ้นหรือลดลงเป็นค่าคงที่

2. Additive Seasonal Model

$$S_t = \alpha(Y_t - B_{t-s}) + (1-\alpha)(S_{t-1} + A_{t-1}), 0 \leq \alpha \leq 1, \quad (26)$$

$$A_t = \beta(S_t - S_{t-1}) + (1-\beta)A_{t-1}, 0 \leq \beta \leq 1, \quad (27)$$

$$B_t = \gamma(Y_t - S_t) + (1-\gamma)B_{t-s}, 0 \leq \gamma \leq 1, \quad (28)$$

เมื่อ γ คือค่าคงที่แสดง seasonal และ s คือความยาวช่วงของ seasonal ใช้ตัวแบบการพยากรณ์นี้เมื่อข้อมูลมี seasonal เป็นแบบการบวก (additive seasonality) คือในเวลาเดียวกันของแต่ละฤดูกาล ค่าสังเกตจะเพิ่มขึ้นหรือลดลงเป็นค่าคงที่ ในทางปฏิบัติเราไม่ทราบว่าจะต้องกำหนดค่า moving average length เท่าไรดี จึงจะทำให้ ค่า error หรือ mean square error (MSE) หรือ ค่าวัดความผิดพลาดอื่นๆ ให้ค่าต่ำที่สุด ดังนั้นเราควรจะต้องกำหนดค่า α ให้เหมาะสม

4) แบบจำลอง Autoregressive Integrated Moving Average (ARIMA) Model (Montgomery,2008)

แบบจำลอง ARIMA เป็นแบบจำลองที่ได้รับความนิยม และเป็นวิธีที่ให้ค่าพยากรณ์ในระยะสั้นที่ดี อีกทั้งในการจัดทำสมการและการพยากรณ์ยังมีขั้นตอนที่ยุ่งยาก และซับซ้อนน้อยกว่าแบบมหภาคที่อยู่ในลักษณะระบบสมการหลายชั้น สำหรับแบบจำลอง ARIMA เป็นแบบจำลองที่พัฒนาโดย George E.P.Box และ Gwilym M. Jenkins ในปี ค.ศ. 19702 โดยพื้นฐานแล้วแบบจำลอง ARIMA เป็นวิธีที่ให้ค่าพยากรณ์ในระยะสั้นที่ดี หรือเหมาะกับการพยากรณ์ไปข้างหน้าในช่วงเวลาสั้นๆ และต้องมีช่วงของข้อมูลที่ยาวพอสมควร แบบจำลอง ARIMA(p,d,q) ประกอบด้วย 3 ส่วนหลักๆ ได้แก่ แบบจำลอง Auto Regressive (AR(p)) กระบวนการ Integrated (I(d)) และแบบจำลอง Moving Average (MA(q)) โดยรายละเอียดของแต่ละส่วนมีดังนี้

1. แบบจำลอง Auto Regressive (AR(p))

แบบจำลอง Auto Regressive เป็นรูปแบบที่แสดงว่าค่าสังเกต y_t ถูกกำหนดจากค่าของ $y_t, \dots, y_{t-p}$ หรือ ค่าสังเกตที่เกิดขึ้นก่อนหน้า p โดยกระบวนการหรือระบบ AR(p) คือกระบวนการหรือระบบ Auto Regressive ที่มีอันดับที่ p ซึ่งเขียนอยู่ในรูปสมการได้ดังนี้

$$\text{AR}(p) \quad \text{คือ} \quad x_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (29)$$

โดยที่

μ คือ ค่าคงที่ (Constant Term)

ϕ_j คือ พารามิเตอร์ตัวที่ j

ε_t คือ ความคลาดเคลื่อน ณ เวลา t

ในกรณี ของ AR(1) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t \quad (30)$$

หรือ

$$y_t - \phi_1 y_{t-1} = \mu + \varepsilon_t \quad (31)$$

หรือ

$$(1 - \phi_1 \beta) x_t = \mu + \varepsilon_t \quad (32)$$

เมื่อ β คือ backward shift operation

และในกรณี ของ AR(2) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \quad (33)$$

หรือ

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} = \mu + \varepsilon_t \quad (34)$$

หรือ

$$(1 - \phi_1 \beta - \phi_2 \beta^2) y_t = \mu + \varepsilon_t \quad (35)$$

2. แบบจำลอง Moving Average (MA(q))

แบบจำลอง Moving Average (MA) เป็นรูปแบบที่แสดงว่าค่าสังเกต y_t ถูกกำหนดจากค่าความคลาดเคลื่อน $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ หรือค่าความคลาดเคลื่อนที่อยู่ก่อนหน้า โดยกระบวนการหรือระบบ MA(q) คือกระบวนการหรือระบบ Moving Average ที่มีอันดับ q ซึ่งเขียนในรูปของ MA (q) ได้ดังนี้

$$\text{MA (q) คือ } y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (36)$$

โดยที่

μ คือ ค่าคงที่ (Constant Term)

θ_j คือ พารามิเตอร์ตัวที่ j

ε_t คือ ความคลาดเคลื่อน ณ เวลา t

ในกรณี MA(1) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (37)$$

หรือ

$$y_t = \mu + (1 - \theta_1 \beta) \varepsilon_t \quad (38)$$

และในกรณี MA(2) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (39)$$

หรือ

$$y_t = \mu + (1 - \theta_1 \beta - \theta_2 \beta^2) \varepsilon_t \quad (40)$$

3. แบบจำลอง Auto Regressive Moving Average (ARMA(p,q))

เป็นแบบจำลองที่นำเอากระบวนการ Auto Regressive และ Moving Average มาใช้รวมกัน โดยกระบวนการหรือระบบ ARMA(p,q) คือกระบวนการหรือระบบ Auto Regressive ที่มีอันดับที่ p และ Moving Average ที่มีอันดับ q ซึ่งเขียนอยู่ในรูปสมการได้ดังนี้

$$\Delta_t y_t = \delta + \phi y_{t-1} + \phi y_{t-2} + \dots + \phi y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (41)$$

โดยที่

y_t คือ ค่าสังเกตในอนุกรมเวลา ณ เวลา t

p คือ อันดับของ Autoregressive

q คือ อันดับของ Moving Average

δ คือ ค่าคงที่ (Constant Term)

t คือ เวลา

ϕ คือ พารามิเตอร์ของ Auto Regressive

θ คือ พารามิเตอร์ของ Moving Average

ε_t คือ กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t

4. กระบวนการ Integrated (I (d))

กระบวนการ Integrated (I(d)) เป็นการหาผลต่างของอนุกรมเวลาระหว่างข้อมูล ณ ปัจจุบัน กับข้อมูลถอยหลังไป d คาบเวลา โดยสาเหตุที่ต้องทำการหาผลต่างของอนุกรมเวลา เนื่องจากแบบจำลอง ARIMA ต้องใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีคุณสมบัติคงที่ (Stationary) เท่านั้น โดยในกรณีข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์มีคุณสมบัติไม่คงที่ (Nonstationary) จะต้องทำการแปลงข้อมูลดังกล่าวให้เป็นข้อมูลที่มีคุณสมบัติคงที่ก่อน โดยการหาผลต่างของข้อมูลอนุกรมเวลา ก่อนที่นำไปสร้างแบบจำลอง ARIMA ซึ่งโดยทั่วไปแล้วถ้าต้องหาผลต่างอันดับที่ d สามารถเขียนในรูปของ I(d) ได้ดังนี้

$$I(d) \text{ คือ } \Delta_d x_t = \Delta_{d-1}(x_t - x_{t-1}) \text{ หรือ } (1-\beta)^d x_t \quad (42)$$

ในกรณี I(1) สามารถเขียนรูปแบบได้ดังนี้

$$I(1) \text{ คือ } \Delta x_t = (x_t - x_{t-1}) \text{ หรือ } (1-\beta)x_t$$

ในกรณี I(2) สามารถเขียนรูปแบบได้ดังนี้

$$I(2) \text{ คือ } \Delta_2 x_t = \Delta(x_t - x_{t-1}) \text{ หรือ } (1-\beta)^2 x_t$$

โดยที่

ε_t คือ พจน์ความคลาดเคลื่อน ณ เวลา t

$(1-\beta)^d x_t$ คือ ผลต่างอันดับที่ d

β คือ Backward shift operation

จากรายละเอียดต่างๆ ที่กล่าวข้างต้นถ้านำแบบจำลอง Auto Regressive แบบจำลอง Moving Average และ กระบวนการ Integrated มาพิจารณารวมกันสามารถนำมากำหนดเป็นรูปแบบทั่วไปของแบบจำลอง ARIMA ที่ใช้ในการประมาณการคือ

$$\Delta_t y_t = \delta + \phi \Delta_d y_{t-1} + \phi \Delta_d y_{t-2} + \dots + \phi \Delta_d y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (43)$$

โดยที่

y_t คือ ค่าสังเกตในอนุกรมเวลา ณ เวลา t

d คือ จำนวนครั้งของการหาผลต่างเพื่อให้อนุกรมเวลามีคุณสมบัติคงที่ (Stationary)

p คือ อันดับของ Autoregressive

q คือ อันดับของ Moving Average

δ คือ ค่าคงที่ (Constant Term)

t คือ เวลา

Δ_d คือ ผลต่างอันดับที่ d

$\phi_1, \dots, \phi_p$ คือ พารามิเตอร์ของ Auto Regressive

$\theta_1, \dots, \theta_q$ คือ พารามิเตอร์ของ Moving Average

ε_t คือ กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t

ภายใต้ข้อสมมติที่ว่าความคลาดเคลื่อนที่กินเวลาเป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยมีการแจกแจงปกติที่มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนคงที่

ดังนั้นเพื่อให้ง่ายขึ้นสำหรับการหาค่าพยากรณ์ Y_t ของ ARIMA สามารถหาได้จากนิยามที่สรุปไว้ต่อไปนี้

$$(1 - \phi_1\beta - \phi_2\beta^2 - \dots - \phi_p\beta^p)(1 - \beta)^d Y_t = (1 - \omega_1\beta - \omega_2\beta^2 - \dots - \omega_q\beta^q)\varepsilon_t \quad (44)$$

$$\text{เมื่อ } \beta Y_t = Y_{t-1}, \beta^2 Y_t = Y_{t-2}, \beta^3 Y_t = Y_{t-3}, \dots \quad (45)$$

$$\beta \varepsilon_t = \varepsilon_{t-1}, \beta^2 \varepsilon_t = \varepsilon_{t-2}, \beta^3 \varepsilon_t = \varepsilon_{t-3}, \dots \quad (46)$$

ตัวอย่างเช่น

ARIMA(0,1,0) คือ

$$\begin{aligned} (1 - \beta)Y_t &= \varepsilon_t \\ Y_t - \beta Y_t &= \varepsilon_t \\ Y_t - Y_{t-1} &= \varepsilon_t \\ Y_t &= Y_{t-1} + \varepsilon_t \end{aligned} \quad (47)$$

ARIMA(2,1,2) คือ

$$\begin{aligned} (1 - \phi_1\beta - \phi_2\beta^2)(1 - \beta)Y_t &= (1 - \omega_1\beta - \omega_2\beta^2)\varepsilon_t \\ (1 - \phi_1\beta - \phi_2\beta^2)(Y_t - \beta Y_t) &= \varepsilon_t - \omega_1\beta\varepsilon_t - \omega_2\beta^2\varepsilon_t \\ (Y_t - \phi_1\beta Y_t - \phi_2\beta^2 Y_t) - (\beta Y_t - \phi_1\beta^2 Y_t - \phi_2\beta^3 Y_t) &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\ Y_t - \phi_1\beta Y_t - \phi_2\beta^2 Y_t - \beta Y_t + \phi_1\beta^2 Y_t + \phi_2\beta^3 Y_t &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\ Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - Y_{t-1} + \phi_1 Y_{t-2} + \phi_2 Y_{t-3} &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\ Y_t - \phi_1 Y_{t-1} - Y_{t-1} - \phi_2 Y_{t-2} + \phi_1 Y_{t-2} + \phi_2 Y_{t-3} &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\ Y_t - (\phi_1 - 1)Y_{t-1} - (\phi_2 - \phi_1)Y_{t-2} + \phi_2 Y_{t-3} &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \end{aligned}$$

$$Y_t = (\phi_1 - 1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} - \phi_2 Y_{t-3} + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} \quad (48)$$

ตัวแบบ ARIMA ใช้วิเคราะห์ข้อมูลได้ทั้งแบบ stationary และ nonstationary ข้อมูล stationary คือข้อมูลที่ mean และ variance ไม่เปลี่ยนแปลงไปตามเวลา ข้อมูล nonstationary คือข้อมูลที่ mean หรือ variance เปลี่ยนแปลงไปตามเวลา ค่า variance ทำให้ stationary ได้โดยการแปลงข้อมูลด้วย ฟังก์ชันลอการิทึม (Log transform) ซึ่งต้องทำก่อนใช้ตัวแบบ A สำหรับค่า mean ทำให้ stationary ได้ โดยการทำผลต่างข้อมูล (differencing) ซึ่งสามารถทำในตัวแบบ ARIMA ได้ โดยการกำหนดค่าให้กับ d ส่วนระดับของ autocorrelation (p) และระดับของ moving average (q) พิจารณาได้จากกราฟ (Autocorrelation function) ACF และ Partial autocorrelation function (PACF) (Montgomery, 2008)

5. Autocorrelation Function (ACF)

เป็นฟังก์ชันของการวัดสหสัมพันธ์ระหว่างข้อมูล ณ เวลา t (x_t) และ ข้อมูล ณ เวลา $t-k$ (x_{t-k}) ของช่วงเวลาห่างกัน k หน่วย ซึ่งแทนด้วยสัญลักษณ์ ρ_k หรือ r_k ในกรณีสหสัมพันธ์ในตัวเองของตัวอย่าง ซึ่งสามารถคำนวณได้ดังนี้

$$\rho_k \text{ หรือ } r_k = \frac{\sum_{t=k+1}^n (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad (49)$$

$$\text{เมื่อ } \bar{x} = \frac{1}{n} \sum_{t=1}^n x_t \text{ และ } k = 0, 1, 2, 3, \dots$$

โดยความคลาดเคลื่อนมาตรฐานของ r_k (Standard Error of r_k) ซึ่งมีสูตรในการคำนวณดังนี้

$$se_{r_k} = \frac{1}{\sqrt{n}} \quad (50)$$

สหสัมพันธ์ในตัวเองของข้อมูลสุ่ม (random data) มีการแจกแจงเชิงตัวอย่างที่สามารถประมาณได้ โดยการแจกแจงปกติด้วยค่าเฉลี่ย (mean) เท่ากับศูนย์ และความคลาดเคลื่อนมาตรฐานเท่ากับ $\frac{1}{\sqrt{n}}$

ในการศึกษาจะใช้สหสัมพันธ์ในตัวเองเป็นเครื่องมือที่สำคัญสำหรับสืบค้นคุณสมบัติของข้อมูลอนุกรมเวลาเชิงประจักษ์ โดยมี 2 วิธีสำหรับทดสอบว่าค่า r_k มีค่าแตกต่างไปจากศูนย์หรือไม่โดยใช้การแจกแจงปกติมาตรฐาน (Standard Normal Distribution) หรือ ใช้ค่าสถิติ Box-Pierce Q statistic ซึ่งมีรูปแบบดังนี้

การแจกแจงปกติมาตรฐาน (Standard Normal Distribution)

$$r_k \sim N(0, \frac{1}{\sqrt{n}})$$

ค่าสถิติ Box-Pierce Q statistic

$$Q = n \sum_{k=1}^m r^2 \sim \chi^2(m - p - q)$$

โดยที่ m คือค่าล่าหรือค่าล่าหลังสูงสุด (Maximum Lag) ที่พิจารณา

6. Partial Autocorrelation Function (PACF)

เป็นการพิจารณาสหสัมพันธ์ระหว่างตัวแปร x_t กับ x_{t-k} อาจเป็นไปได้ว่าสหสัมพันธ์ดังกล่าวเป็นผลเนื่องมาจากสหสัมพันธ์ระหว่าง 2 ตัวแปรนี้กับตัวแปร $x_{t-1}, \dots, x_{t-k+1}$ ดังนั้นเพื่อที่จะได้สหสัมพันธ์ระหว่าง x_t กับ x_{t-k} ที่ได้ขจัดความเกี่ยวข้องระหว่างตัวแปรทั้งสองตัวนี้กับตัวแปร

$x_{t-1}, \dots, x_{t-k+1}$ ดังกล่าว จึงต้องทำการวัดสหสัมพันธ์ของทั้งสองตัวแปรในรูปแบบของการสหสัมพันธ์แบบมีเงื่อนไข $Corr(x_t, x_{t-k} | x_{t-1}, \dots, x_{t-k+1})$ ซึ่งเรียกว่า Partial Autocorrelation โดยแทนด้วยสัญลักษณ์ ϕ_{kk} แต่ถ้านำสหสัมพันธ์ในตัวเองบางส่วนมาพิจารณาในรูปแบบฟังก์ชัน จะเรียกว่า Partial Autocorrelation Function (PACF) ซึ่ง ϕ_{kk} สามารถคำนวณได้ดังนี้

$$\phi_{kk} = \frac{Cov[(x_t - \hat{x}_t), (x_{t-k} - \hat{x}_{t-k})]}{\sqrt{Var(x_t - \hat{x}_t)} \sqrt{Var(x_{t-k} - \hat{x}_{t-k})}} \quad (51)$$

$$\text{โดยที่} \quad \hat{x}_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k+1} \quad (52)$$

5) วิธีการพยากรณ์แบบ Seasonal Autoregressive Integrated Moving Average Model (SARIMA) (Montgomery, 2008), (Bisgaard, 2011)

วิธีการพยากรณ์แบบ SARIMA ถูกพัฒนามาจากวิธีการพยากรณ์ ARIMA (p,d,q) โดยได้เพิ่ม (P,D,Q) ของ Seasonal เข้าไปอีก SARIMA หรือ Seasonal ARIMA แทนด้วย ARIMA (p,d,q)(P,D,Q) และถ้ามีตัวแปรร่วม (covariate) X_{1t} และ X_{2t} สามารถนิยามได้ดังต่อไปนี้

$$\phi(\beta) [\Delta(Y_t - c_1 X_{1t} - c_2 X_{2t}) - \mu] = \theta(\beta) \varepsilon_t \quad (53)$$

เมื่อ c_1 และ c_2 คือสัมประสิทธิ์การถดถอย

$$\phi(\beta) = \phi_p(\beta) \phi_p(\beta^s) \text{ และ } \theta(\beta) = \theta_q(\beta) \theta_q(\beta^s) \text{ โดยที่}$$

$$\phi_p(\beta) = 1 - \phi_1 \beta - \dots - \phi_p \beta^p, \theta_q(\beta) = 1 - \theta_1 \beta - \dots - \theta_q \beta^q,$$

$\phi_p(\beta^s) = 1 - \phi_1 \beta^s - \dots - \phi_p \beta^{sp}$, และ $\theta_q(\beta^s) = 1 - \theta_1 \beta^s - \dots - \theta_q \beta^{sq}$. Δ คือ differencing operator, $\Delta = (1 - \beta)^d (1 - \beta^s)^D$. β คือ backshift operator และ $\beta Y_t = Y_{t-1}$, $\beta^2 Y_t = Y_{t-2}$ เป็นต้น

แบบนี้ต่อไป S คือ seasonal lag และ ε_t คือ ความผิดพลาด (error) ที่มีการแจกแจงแบบปกติมีค่าเฉลี่ยเป็น 0 และความแปรปรวน σ^2 . ϕ 's และ θ 's คือพารามิเตอร์ autoregressive แบบ seasonal และ non-seasonal ตามลำดับ Θ 's และ θ 's คือพารามิเตอร์ moving average แบบ seasonal และ non-seasonal ตามลำดับ p และ q คือลำดับ (order) ของพารามิเตอร์ autoregressive แบบ non-seasonal และพารามิเตอร์ moving average แบบ non-seasonal ตามลำดับโดยที่ P และ Q คือ ลำดับ (order) ของพารามิเตอร์ autoregressive แบบ seasonal และ พารามิเตอร์ moving average แบบ seasonal ตามลำดับ d และ D แทน non-seasonal และ seasonal differences ตามลำดับ

6) แบบจำลองหลายตัวแปร (Multivariate Variable Model)

Multivariate AR(1)

$$Z_t = A_1 Z_{t-1} + B \varepsilon_t$$

Expanded Form

$$\begin{bmatrix} z_t^{(1)} \\ z_t^{(2)} \\ \vdots \\ z_t^{(n)} \end{bmatrix} = \begin{bmatrix} a^{11} & a^{12} & \dots & a^{1n} \\ a^{21} & a^{22} & \dots & a^{2n} \\ \vdots & \vdots & \dots & \vdots \\ a^{n1} & a^{n2} & \dots & a^{nn} \end{bmatrix} \begin{bmatrix} z_{t-1}^{(1)} \\ z_{t-1}^{(2)} \\ \vdots \\ z_{t-1}^{(n)} \end{bmatrix} + \begin{bmatrix} b^{11} & b^{12} & \dots & b^{1n} \\ b^{21} & b^{22} & \dots & b^{2n} \\ \vdots & \vdots & \dots & \vdots \\ b^{n1} & b^{n2} & \dots & b^{nn} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{(1)} \\ \varepsilon_t^{(2)} \\ \vdots \\ \varepsilon_t^{(n)} \end{bmatrix}$$

$$\text{เมื่อ } Z_t = \begin{bmatrix} Z_t^1 \\ Z_t^2 \\ \vdots \\ Z_t^n \end{bmatrix} = \text{Column Matrix ขนาด } n \times 1$$

$$A_1 = \begin{bmatrix} a^{11} & a^{12} & \dots & a^{1n} \\ a^{21} & a^{22} & \dots & a^{2n} \\ \vdots & \vdots & \dots & \vdots \\ a^{n1} & a^{n2} & \dots & a^{nn} \end{bmatrix} \text{ Matrix ขนาด } n \times n$$

$$B = \begin{bmatrix} b^{11} & b^{12} & \dots & b^{1n} \\ b^{21} & b^{22} & \dots & b^{2n} \\ \vdots & \vdots & \dots & \vdots \\ b^{n1} & b^{n2} & \dots & b^{nn} \end{bmatrix} \text{ Matrix ขนาด } n \times n$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \vdots \\ \varepsilon_t^n \end{bmatrix}$$

ε_t = Vector of independent normally distribution random variables with mean = 0 and standard deviation = 1 (uncorrelated in time and space)

ซึ่งหลักการของกระบวนการสโตคาสติกที่กล่าวมาข้างต้นได้มีการนำไปประยุกต์ใช้อย่างทั่วโลกอย่างกว้างขวางตั้งแต่อดีตมาจนถึงปัจจุบัน

2. ราคาข้าวไทย

ข้าว นับเป็นพืชเศรษฐกิจของประเทศไทย ซึ่งเกี่ยวข้องกับชีวิตความเป็นอยู่ของชาวนา หลายล้านคน แต่ละปีเราเป็นผู้ส่งออกข้าว นำรายได้เข้าประเทศ ร่วมสองแสนล้านบาท โดยไทยเป็นผู้ส่งออกรายใหญ่ที่สุดของโลก นอกจากนี้ยังมีผู้ส่งออกข้าวรายสำคัญ รองลงมา ได้แก่ เวียดนาม อินเดีย และสหรัฐ เป็นต้น ในอดีตข้าวไทยเคยเป็นสินค้าที่นำรายได้ให้ประเทศเป็นอันดับต้น ๆ ต่อมาเนื่องจากเราหันมาผลิตสินค้าอุตสาหกรรมมากขึ้น ได้แก่ พวกสิ่งทอ เครื่องนุ่งห่ม เครื่องใช้ไฟฟ้า อุปกรณ์อิเล็กทรอนิกส์ ยานยนต์ และชิ้นส่วน รวมทั้งอุปกรณ์คอมพิวเตอร์ เพื่อรองรับความต้องการในประเทศ และภายหลังก็ได้มุ่งผลิตสินค้าเหล่านี้เพื่อการส่งออกอีกด้วย สินค้าอุตสาหกรรมเหล่านี้จึงเป็นแหล่งรายได้ส่งออกลำคัญ แชนหน้าการส่งออกข้าว และพืชผลการเกษตรอื่น ข้าวถือเป็นสินค้าประเภท Commodity (โภคภัณฑ์) ที่ระดับราคาในตลาดโลกเป็นไปตามกลไกตลาด อย่างแท้จริง นั่นคือราคาตลาดโลกจะ

ปรับขึ้นหรือลงย่อมขึ้นอยู่กับอุปสงค์และอุปทาน เมื่อใดที่อุปทานมากกว่าอุปสงค์ราคาในตลาดโลกก็จะปรับลดลง และถ้าปีใดอุปสงค์มากกว่าอุปทานราคาก็จะสูงขึ้น

โดยทั่วไปตลาดข้าวในประเทศไทยจะมีความเชื่อมโยงกับตลาดโลกอย่างใกล้ชิด เมื่อใดที่สถานการณ์ตลาดโลกทำให้ราคาตลาดโลกเพิ่มขึ้น ราคาข้าวในประเทศก็มีแนวโน้มปรับสูงตามไปด้วย เมื่อใดที่ราคาข้าวในประเทศไทยตกต่ำ โดยมากจะเกิดในช่วงที่มีข้าวออกมามากพร้อมๆ กัน รัฐบาลก็จะใช้ โครงการรับจำนำข้าวเปลือก เพื่อพยุงราคาข้าวไว้ โดยรัฐบาลจะรับซื้อข้าวเปลือกในราคาที่ประกาศ (เรียกว่าราคาประกัน) เพื่ออุดหนุนข้าวออกจากตลาดซึ่งราคาประกันแต่ละปีก็จะอยู่ในระดับที่สูงกว่าราคาตลาด สำหรับการรับจำนำนี้ทำได้ 2 รูปแบบคือจำนำที่ยังคงเกษตรกร หรือจำนำที่โรงสีที่เข้าร่วมโครงการฯ แบบแรกนั้น ธกส. จะเป็นผู้ดำเนินการโดยข้าวกู้ฝากเกษตรกรไว้ ถ้าเป็นการรับจำนำที่โรงสี เกษตรกรก็ต้องขนข้าวมาที่โรงสี และก็จะได้รับเอกสารไปขึ้นเงินกับ ธกส. โรงสีก็มีหน้าที่สีข้าวส่งมอบแก่องค์การคลังสินค้า (อคส.) แต่ในความเป็นจริงข้าวเหล่านี้ก็ยังถูกเก็บไว้ที่โรงสี เพื่อรอเวลาจะระบายออกนอกประเทศ หรือในบางกรณี ก็มีการนำกลับมาใช้ในตลาดในประเทศก็ได้แล้วแต่นโยบาย ในอดีตที่ผ่านมาข้าวที่รัฐบาลเป็นเจ้าของผ่านโครงการรับจำนำ มักจะถูกเก็บไว้จนเสื่อมสภาพ หรือไม่ก็มีการล่องหนหายไปจากโกดัง ทำให้ไม่สามารถขายได้หรือถ้าขายได้ก็ได้อาหารต่ำมาก ดังนั้นโครงการรับจำนำข้าวจึงมักเกิดผลขาดทุนเสมอ การขาดทุนก็เนื่องจากซื้อแพงแต่ขายถูกหรือขายไม่ได้เลยก็มี

ในปี 2551 นี้ถือว่าเป็นปีทองของการส่งออกข้าวไทย โดยขณะนี้สถานการณ์ตลาดเป็นของผู้ขาย (ผู้ส่งออก) อย่างแท้จริงโดยความต้องการบริโภคข้าวของโลก มี 424 ล้านตัน ซึ่งสูงกว่าผลผลิตรวมของโลกที่ 420.6 ล้านตัน ภาพการณ์ของตลาดโลกขณะนี้ พบว่า ประเทศผู้ส่งออกรายอื่น รวมทั้งผู้นำเข้าส่วนใหญ่ ก็มีผลผลิตไม่เพียงพอ ซึ่งเกิดจากการขาดแคลนข้าวในประเทศตน อันเนื่องมาจากปัญหาภัยธรรมชาติ ทำให้ผลผลิตที่ได้ไม่พอกับการบริโภคภายในจึงต้องลดหรือระงับการส่งออกข้าว ขณะนี้พบว่าประเทศเวียดนาม อินเดีย และอิหร่าน ได้ประกาศใช้มาตรการห้ามการส่งออกข้าวแล้ว บางประเทศยังใช้มาตรการกำหนดราคาส่งออกขั้นต่ำ (Minimum Export Price) ควบคู่กันอีกด้วย สำหรับประเทศจีนมีการเพิ่มการอุดหนุนในประเทศ และเก็บภาษีส่งออกข้าว ดังนั้นปริมาณข้าวที่จะเข้าสู่ตลาดโลกในปีนี้น่าจะไม่เพิ่มไปจากปีก่อน หรืออาจลดลงก็ได้ ในขณะที่ความต้องการบริโภคข้าวในโลกยังขยายตัวต่อเนื่อง ดังนั้นจึงเป็นเรื่องแน่นอนที่สุดว่าราคาข้าวในตลาดโลกจะต้องสูงขึ้น และโดยที่ไทย

เป็นประเทศเดียวที่มีผลผลิตมากและสามารถส่งออกได้ตามปกติ ทุกประเทศจึงหันมาสั่งซื้อข้าวจากไทย เรียกได้ว่ามีเท่าไรก็ขายหมดแน่

สำหรับปีนี้คาดว่าประเทศไทยจะมีผลผลิตข้าวออกมารวมประมาณ 30 ล้านตัน เพิ่มขึ้นจากปีก่อนเล็กน้อย และจะมีปริมาณส่งออกข้าวราว ๆ 9 ล้านตัน โดยผลผลิตข้าวนาปีได้เริ่มออกสู่ตลาดแล้ว ตั้งแต่ปลายปีก่อนถึงต้นปี 2551 และข้าวนาปรังจะทยอยเข้าตลาดตั้งแต่ช่วงนี้จนถึงเดือนสิงหาคม และตั้งแต่ปลายปีที่แล้วจนถึงเดือนกุมภาพันธ์นี้ ไทยได้ส่งออกข้าวไปแล้วเฉลี่ยเดือนละหนึ่งล้านตัน เรื่องที่น่าดีใจก็คือระดับราคาข้าวส่งออกทุกชนิดได้ปรับสูงขึ้นอย่างมาก เช่น ข้าวหอมมะลิ (100% ชั้น 2) ราคาเพิ่มขึ้น 20% ข้าวขาว (100% ชั้น 2) ข้าวหนึ่ง ก็มีราคาเพิ่มขึ้นถึง 27% ถ้าผู้เขียนคาดการณ์ไม่ผิดปีนี้เราจะได้เห็นราคาข้าวส่งออกของไทยที่สูงถึงตันละ 500 เหรียญสหรัฐฯ ซึ่งนับว่าสูงที่สุดในประวัติศาสตร์การส่งออกข้าวของไทย

อย่างไรก็ดีพบว่าราคาข้าวในประเทศ ก็มีความผันผวนรุนแรง และเป็นไปในทิศทางเพิ่มขึ้นต่อเนื่องโดยราคาข้าวในประเทศ ได้ปรับสูงขึ้นเฉลี่ยสัปดาห์ละ 385-400 บาท ซึ่งสถานการณ์เช่นนี้ทำให้ผู้ส่งออกที่รับคำสั่งซื้อเพื่อส่งออกไว้ก่อนหน้า (ที่ระดับราคาต่ำ) ต้องประสบปัญหาเพราะต้องหาซื้อข้าวในประเทศที่ราคาสูงขึ้นมากเพื่อนำไปส่งออกให้ลูกค้าตามคำสั่งซื้อนั้น ขณะนี้มีบางรายที่ไม่ต้องการรับภาระความเสี่ยงถึงกับระงับคำสั่งซื้อใหม่ ซึ่งน่าจะส่งผลให้ปริมาณส่งออกข้าวในช่วงหลายเดือนข้างหน้าชะลอตัวลง (วโรทัย โกศลพิศิษฐ์กุล, 2551) แต่ในปี พ.ศ. 2557 อาจเป็นเพียงจุดเริ่มของวิกฤติข้าวไทย เพราะหลังเปิดเสรีการค้าอาเซียน ข้าวจากพม่า กัมพูชา และเวียดนาม จะเข้าสู่ไทยโดยเสรี ขณะที่ข้าวไทยมีต้นทุนการผลิตสูงมาก ข้าวนาปรังปี พ.ศ. 2556 ในไทยลงทุน 8,711 บาทต่อตัน ขณะที่เวียดนามอยู่ที่ 4,960 บาท หากเปิดเสรีอาเซียนคาดว่า ราคาข้าวในประเทศอาจเหลือไม่เกิน 6,000 บาทต่อตัน

โครงการรับจำนำข้าวที่มีตั้งแต่ก่อนปี พ.ศ. 2516 เป็นหนึ่งในโครงการ “ประชานิยม” ที่ทุกพรรคการเมืองซึ่งเคยเป็นรัฐบาลล้วนเคยนำมาใช้ เพราะตัวเลขจำนวนชาวนาซึ่งเป็นผู้มีสิทธิลงคะแนนเลือกตั้งเกือบ 18 ล้านคน หรือ 1 ใน 3 ของผู้มีสิทธิเลือกตั้งทั้งหมด ทำให้พรรคการเมืองที่ต้องการชนะเลือกตั้งต่างคิดค้นกลยุทธ์ด้านราคาข้าวมาชวนให้เชื่อได้ว่าจะสามารถยกระดับคุณภาพชีวิตชาวนา ไม่ว่าจะเป็นเรื่องราคาข้าว การอุดหนุนด้วยการให้กู้เงิน หรือปัจจัยการผลิต โดยต่างเป้าหมายที่จะโยก

คะแนนเสียงในการเลือกตั้งเป็นครั้ง ๆ ไป สิ่งที่เห็นน้อยหรือแทบไม่มีเลยในนโยบายข้าวที่มาจากฝ่ายการเมือง อาทิ การส่งเสริมกระบวนการเรียนรู้เพื่อเพิ่มศักยภาพและอำนาจต่อรอง การปกป้องสิทธิของชาวนา สวัสดิภาพและความปลอดภัยในชีวิตการใช้เคมีเกษตร การคุ้มครองเมล็ดพันธุ์และพื้นที่ที่ผลิตอาหารหรือนาข้าว การส่งเสริมการเรียนรู้เพื่อพึ่งตนเอง ลดการพึ่งพาภายนอก ลดต้นทุนการผลิต พัฒนาทักษะและเทคนิคทางการเกษตร ไปจนถึงการฟื้นฟูประเพณีที่เกี่ยวกับข้าว ทั้งหมดนี้ไม่ปรากฏในนโยบายการเมืองเรื่องข้าว (บำเพ็ญ ไชยรักษ์, 2557) จากที่ได้กล่าวมาข้างต้นจะเห็นได้ว่าราคาข้าวมีปัจจัยที่มีความไม่แน่นอนต่างๆแทรกเข้ามาตลอดเวลาเหนือการคาดเดาเหตุการณ์ได้ ผู้วิจัยจึงมีความสนใจใช้หลักการของกระบวนการทางสถิติศาสตร์สร้างตัวแบบวิเคราะห์ข้อมูลอนุกรมเวลาเชิงลึกของราคาข้าวไทย

งานวิจัยที่เกี่ยวข้อง

แบบจำลอง หรือตัวแบบการพยากรณ์ของไทยในปัจจุบันใช้ตัวแบบการพยากรณ์ส่วนใหญ่ยังใช้ตัวแบบการพยากรณ์แบบดั้งเดิม ทั่วๆ ไป (Classical model) เช่น Multiple Regression, Moving Average (MA), Exponential Smoothing (EXPS), ARIMA และ Seasonal ARIMA เป็นต้น ตัวแบบโครงข่ายประสาทเทียม (Artificial Neural Network) ตัวแบบ Semiparametric Multiple Regression เป็นต้น ตัวแบบที่มีความซับซ้อนมากขึ้น เช่นตัวแบบที่มีการนำความสัมพันธ์ของข้อมูลอนุกรมเวลามาพิจารณาคำนวณด้วย และตัวแบบที่กำหนดให้พารามิเตอร์ในตัวแบบมีความไม่แน่นอน (uncertainty) ยังมีน้อยมาก ส่วนใหญ่จะอยู่ในต่างประเทศ สำหรับในงานวิจัยนี้ได้ศึกษางานวิจัยที่เกี่ยวข้องที่ใช้ตัวแบบการพยากรณ์ที่นำมาใช้กับด้านอุตสาหกรรม และนำไปใช้กับงานด้านอื่นๆ ดังนี้

ตัวแบบการพยากรณ์แบบดั้งเดิม (Classical Models) ในการพยากรณ์ผลลัพธ์ที่จะเกิดขึ้นในอนาคตของเหตุการณ์ด้านต่างๆ นั้น Wright (1986) นำตัวแบบ Simple Exponential และ Holt ไปใช้กับข้อมูลอนุกรมเวลาที่ไม่สม่ำเสมอ (Irregular) Deetee (1991) ประยุกต์ใช้ตัวแบบ Box Jenkins (ARIMA) ในการพยากรณ์ราคาข้าว พบว่าตัวแบบนี้มีประสิทธิภาพพอให้เห็นได้ชัด และดีกว่าตัวแบบ Decomposition เช่นเดียวกันกับที่ Kerdsoomboon (1999) พบว่า ตัวแบบ Box Jenkins พยากรณ์ราคาข้าวได้ดีกว่าตัวแบบสถิติเบื้องต้น Cipra และคณะ (1995) ใช้ตัวแบบ Holt-Winter กับข้อมูลอนุกรมเวลาที่มีข้อมูลสูญหาย (Missing) Hyndman (2002) พบว่าการใช้ตัวแบบ Single Exponential Smoothing ใน

กรอบการทำงาน (Framework) ของตัวแบบ State-space มีประสิทธิภาพสูง โดยนำไปใช้กับข้อมูลอนุกรมเวลา M-Competition ซึ่งเป็นข้อมูลทางด้านเศรษฐศาสตร์ การเงิน อุตสาหกรรม ประชากรศาสตร์ และอื่นๆ ในขณะที่ Sangpattaranate (2005) ที่พบว่า Box Jenkins เป็นตัวแบบที่ใช้ในการพยากรณ์ราคาข้าวได้ดีกว่าตัวแบบ Holt-Winter และตัวแบบการถดถอย

Iqbal และคณะ (2005) ใช้ตัวแบบ ARIMA พยากรณ์ผลผลิตและพื้นที่เพาะปลูกข้าวสารในประเทศไทยปาลีสถานเพื่อใช้เป็นข้อมูลให้กับรัฐบาลในการกำหนดนโยบาย Mishra และ Desai (2005) ใช้ตัวแบบ SARIMA ในการพยากรณ์ภัยแล้ง (Drought) โดยใช้ข้อมูลอนุกรมเวลาเป็นค่าดัชนีมาตรฐานของหยาดน้ำที่ตกมาจากชั้นบรรยากาศ (Precipitation) สำหรับตัวแบบการพยากรณ์แบบ exponential smoothing นั้นมีการนำไปใช้อย่างแพร่หลายนั้น และมีการขยายเป็น Simple exponential smoothing, Holt, Holt-Winters และ double exponential smoothing Cipra (2006) ใช้ตัวแบบ double exponential smoothing กับข้อมูลอนุกรมเวลาที่ไม่สม่ำเสมอ

Boosarawongse และคณะ (2007) ได้ใช้วิธีการพยากรณ์ Box-Jenkins (ARIMA) กับ Artificial Neural Network สำหรับพยากรณ์ commodity prices การส่งออกข้าวของไทย 4 ชนิด ผลพบว่าวิธีการพยากรณ์ทั้งสองให้ผลที่ดี แต่วิธีการพยากรณ์ Artificial Neural Network ให้ผลการพยากรณ์ที่แม่นยำข้าวที่ดีกว่า 3 ชนิด โดยเปรียบเทียบจากค่าความคลาดเคลื่อนต่างๆ Sumer และคณะ (2009) ศึกษาการใช้ตัวแบบ ARIMA, SARIMA และ ตัวแบบการถดถอย (Regression Model) ที่มีฤดูกาล (Seasonal) เป็นตัวแปรซ่อนเร้น (Latent variable) ในการพยากรณ์ ปริมาณความต้องการกระแสไฟฟ้าพบว่าตัวแบบการถดถอยที่มีฤดูกาลเป็นตัวแปรซ่อนเร้นพยากรณ์ได้แม่นยำกว่า ARIMA และ SARIMA

Kahforoushan, Zarif and Mashahir (2010) ศึกษาการพยากรณ์ผลผลิตทางการเกษตร ซึ่งได้แก่ การปลูกพืช การเลี้ยงสัตว์ การประมง และการปลูกป่า โดยใช้วิธีการพยากรณ์ 4 วิธี ได้แก่ วิธีปรับเรียบเอ็กโปเนนเชียลโดยวิธี Holt-Winters แบบไม่มีฤดูกาล (Holt-Winters (no seasonal) Exponential Smoothing Model) วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins Model) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network Model) และ วิธี ARIMA (ARIMA Model) และใช้ ค่า MAE MSE และ MAPE เปรียบเทียบผลการพยากรณ์แต่ละวิธี ผลการศึกษาพบว่า วิธีโครงข่ายประสาทเทียมเหมาะสมในการประมาณค่าพารามิเตอร์ในตัวแบบ (Learn Stage) วิธีบ็อกซ์-เจนกินส์เหมาะสมในการประเมิน

ความถูกต้องในตัวแบบ (Model Validation) แต่วิธีปรับเรียบเอ็กโปเนนเชียลโดยวิธี Holt-Winters แบบไม่มีฤดูกาล ให้ค่า MAPE ต่ำสุดในการประมาณค่าพารามิเตอร์ในตัวแบบ (Model Fitting) และการประเมินความถูกต้องในตัวแบบ

นอกจากตัวแบบการพยากรณ์แบบเดิมที่กล่าวมาแล้วนั้น ยังมีอีกตัวแบบหนึ่งที่กำลังได้รับความนิยมเพิ่มมากขึ้นเรื่อยๆ คือตัวแบบการพยากรณ์แบบเบย์ ตัวแบบนี้เหมาะสำหรับกรณีที่กำหนดให้พารามิเตอร์ในตัวแบบมีความไม่แน่นอน (Uncertainty) และยังสามารถนำความสัมพันธ์ของข้อมูลอนุกรมเวลามาคิดคำนวณได้ด้วย ดังปรากฏในงานของ Monahan (1983) ใช้วิธีของเบย์ในการประมาณค่าพารามิเตอร์ในตัวแบบ ARMA และ Broemeling และ Land (1984) ใช้วิธีของเบย์ในการประมาณค่าพารามิเตอร์ในตัวแบบ AR(p) Liu (1994) ใช้วิธีของเบย์ในการประมาณค่าพารามิเตอร์ในตัวแบบ AR(p) ที่มีตัวแปรภายนอก (Exogenous) รวมอยู่ด้วย

Neelamegham และ Chintagunta (1999) ประยุกต์ใช้ตัวแบบเบย์ในการพยากรณ์การจำนวนผู้เข้าชมภาพยนตร์ใหม่ ที่เข้าฉายในโรงภาพยนตร์ในสัปดาห์แรกทั้งภายในประเทศและบางประเทศในต่างประเทศ ซึ่งมีประโยชน์ต่อผู้ที่เกี่ยวข้องเช่น เจ้าของโรงภาพยนตร์ ผู้แทนจำหน่าย และผู้จัดทำโฆษณา เป็นต้น จำนวนผู้เข้าชมภาพยนตร์เป็นจำนวนนับ ที่มีการแจกแจงแบบปัวซอง ผลการศึกษาพบว่า ตัวแบบเบย์พยากรณ์ได้แม่นยำในระดับประเทศ และเมื่อพิจารณาค่ารากที่สองของค่าเฉลี่ยกำลังสองของค่าความผิดพลาด (Root Mean Square Error) และค่าเฉลี่ยความผิดพลาดสัมบูรณ์ (Mean Absolute Error) พบว่ามีค่าต่ำกว่าตัวแบบของผู้วิจัยอื่นๆ คือ ตัวแบบของ Sawhney และ Eliashberg (1996) Fourth และ Woodlock (1960) และตัวแบบ Naïve (Logged) OLS และ Poisson Maximum Likelihood

De Alba และ Mendoza(2006) ศึกษาการพยากรณ์โดยใช้วิธีเบย์ เมื่อมีข้อมูลจำนวนน้อย ค่าที่พยากรณ์เป็นค่าสะสมของตัวแปรต่อเนื่องที่เป็นค่าบวกโดยทราบค่าสะสมของข้อมูลมาส่วนหนึ่งแล้ว ตัวแบบที่ถูกนำเสนอเป็นการอธิบายความสัมพันธ์ระหว่างค่ารวมทั้งหมดกับค่ารวมมาแล้วบางส่วน ของตัวแปรภายใต้อิทธิพลของฤดูกาลแบบคงที่ (Stable Seasonality) ผลการศึกษาพบว่าตัวแบบที่นำเสนอเหมาะสมเมื่อมีข้อมูลจำนวนน้อย และตัวแบบมาตรฐานทั่วไปไม่เหมาะสม Pedroza (2006) ใช้ตัวแบบเบย์ในการพยากรณ์อัตราการเสียชีวิต ของชายชาวสหรัฐอเมริกา ใช้ MCMC ในการประมาณ

ค่าพารามิเตอร์ และใช้ Gibbs sampling ในการสุ่มตัวอย่างจาก Posterior กลุ่มตัวอย่างเป็นข้อมูลการเสียชีวิตของชายชาวสหรัฐอเมริกา เป็นการพยากรณ์อัตราการเสียชีวิตในช่วงปี 1990-1999 โดยใช้ข้อมูลปี 1959-1989 การพยากรณ์นี้เมื่อนำไปเปรียบเทียบกับค่าสังเกตจริง และวิธีการของ Lee-Carter พบว่าวิธีการของเบย์เหมาะสมกว่า de Alba และ Mendoza(2007) ที่นำตัวแบบการพยากรณ์แบบเบย์ไปใช้กับข้อมูลอนุกรมเวลาที่มีอิทธิพลของฤดูกาลรวมอยู่ด้วย

Yelland (2009) ใช้วิธีของเบย์ประมาณค่าพารามิเตอร์ในตัวแบบ state-space 3 ประเภทคือ Adjusted Gaussian Dynamic Linear Model (AG), Poisson Dynamic Log-Linear Model (PL) และ Gamma-Poisson Local Level Model (GP) รวมทั้ง ตัวแบบ Climatological Baseline Model (Cm) กับข้อมูลปริมาณความต้องการซื้อสินค้า พบว่า ตัวแบบ GP ดีที่สุด Yelland (2010) นำเสนอตัวแบบเบย์ให้มีความเหมาะสมกับข้อมูลปริมาณความต้องการซื้อชิ้นส่วนอุปกรณ์คอมพิวเตอร์ และเมื่อเปรียบเทียบกับตัวแบบมาตรฐานอื่นๆ ได้แก่ Exponential smoothing (Exps) และ Judgmental Methods (Judg) พบว่าวิธีเบย์ มีความเหมาะสมมากกว่า จากที่กล่าวมาทั้งหมดนั้น Tongkhaw และ Kantanantha (2011) ได้ศึกษาตัวแบบพยากรณ์ที่นิยมใช้และมีประสิทธิภาพในปัจจุบันต่างๆ โดยนำไปประยุกต์ใช้กับข้อมูลราคาผักที่มีความผันผวนในประเทศไทยและได้สร้างตัวแบบเบย์อย่างง่ายๆ ตามหลักการของสมการการถดถอยเชิงพหุ ผลการศึกษาพบว่าวิธีการของ Exps มีความเหมาะสมมากกว่า และ SesonalARIMA มีความเหมาะสมบางค่า หลังจากนั้น Tongkhaw และ Kantanantha (2012, 2013) ขยายตัวแบบของ Yelland (2010) โดยนำไปประยุกต์ใช้กับข้อมูลราคาผักที่มีความผันผวนในประเทศไทย และนำมาเปรียบเทียบกับตัวแบบของ Yelland (2010) ARIMA SesonalARIMA และ Exps พบว่าวิธีที่นำเสนอมีความเหมาะสมมากกว่า จากที่กล่าวมาทั้งหมด

บทที่ 3

วิธีดำเนินการวิจัย

กระบวนการสโตนแคสติ้งสำหรับการวิเคราะห์ตัวแบบอนุกรมเวลาโดยประยุกต์ใช้กับข้อมูลราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทยในครั้งนี้ มีวิธีดำเนินการในแต่ละข้อต่อไปนี้

ข้อมูลและแหล่งข้อมูล

ขอบเขตการวิจัย

ขั้นตอนการดำเนินการวิจัย

สถานที่ทำวิจัย

ระยะเวลาในการทำวิจัย

ข้อมูลและแหล่งข้อมูล

ข้อมูลราคาปิดของข้าวขาว 5% เฉลี่ยรายเดือน ตั้งแต่ปี 2548-2557 จำนวน 120 เดือน จากตลาดสินค้าเกษตรล่วงหน้าแห่งประเทศไทย ข้อมูลราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ตั้งแต่ปี 2540-2557 จำนวน 216 เดือนจากสำนักงานสถิติการเกษตร

ขอบเขตการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากร คือ ราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ย ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

และกลุ่มตัวอย่าง คือ ราคาปีของข้าวขาว 5% รายเดือนเฉลี่ย จำนวน 120 เดือน ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

2. ตัวแปรสำหรับการวิจัย

ตัวแปรต้น แนวโน้ม (Trend) ฤดูกาล (Seasonal) ค่าผิดปกติ (Irregular Variation) ค่าอัตโนมัติสัมพันธ์ (Autocorrelation)

ตัวแปรตามคือ ราคาข้าวแต่ละชนิดรายเดือน

ขั้นตอนการดำเนินการวิจัย

1. ศึกษาหัวเรื่อง ต่อไปนี้

1.1 ราคาข้าวของไทย

1.2 กระบวนการสโตแคสติก ตัวแบบพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีการของเบย์ และใช้การประมาณค่าพารามิเตอร์ด้วยวิธี Markov Chain Monte Carlo (MCMC) โดยสุ่มตัวอย่างแบบ Gibbs sampling ที่มีตัวแปรตามมีค่าต่อเนื่อง จึงสมมติให้มีการแจกแจงแบบปกติ และตัวแปรต้นเป็นฟังก์ชันต่างๆ

1.3 การประมาณค่าด้วยวิธีการของเบย์

1.4 งานวิจัยที่เกี่ยวข้องกับการประยุกต์ใช้ตัวแบบทางด้านสโตแคสติก ตัวแบบพยากรณ์ต่างๆ และงานวิจัยที่เกี่ยวข้องกับข้าว

2. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลใช้โปรแกรมวิเคราะห์ข้อมูลทางสถิติ วิเคราะห์ลักษณะทั่วไปของข้อมูลจากตัวอย่าง วิเคราะห์ราคาข้าวแต่ละชนิด

2.1 การวิเคราะห์ลักษณะทั่วไปของข้อมูลจากตัวอย่าง ใช้ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน

2.2 การวิเคราะห์ราคาข้าว โดยใช้หลักการทางกระบวนการสโตแคสติก ด้วยวิธีการของเบย์ในการประมาณค่าพารามิเตอร์วิธี Markov Chain Monte Carlo (MCMC) โดยสุ่มตัวอย่างแบบ Gibbs sampling ที่มีตัวแปรตามมีค่าต่อเนื่อง จึงกำหนดให้มีการแจกแจงแบบปกติ และตัวแปรต้นเป็นฟังก์ชันต่างๆ

รายละเอียดของตัวแบบแสดงดังต่อไปนี้

$$Y_t \sim N\left(\gamma(\Delta W(t|\alpha, \delta) + A_t) + \sum_{i=1}^{s-1} \omega_i S_{it}, [\gamma(1 + \zeta_t)\sigma_Y]^2\right)$$

ซึ่ง ค่าเฉลี่ยของ Y_t คือ

$$E(Y_t) = \gamma(\Delta W(t|\alpha, \delta) + A_t) + \sum_{i=1}^{s-1} \omega_i S_{it} \quad (54)$$

ค่าความแปรปรวนของ Y_t คือ

$$\text{Var}(Y_t) = [\gamma(1 + \zeta_t)\sigma_Y]^2 \quad (55)$$

เมื่อ

γ ค่าคาดหวังของ Z ซึ่งค่า Z คือผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษา

$W(t|\alpha, \delta)$ คือการแจกแจงสะสมของแนวโน้ม ได้แก่เวย์บูล (Weibull) และ เอ็กซ์โปเนนเชียล (Exponential)

A_t ค่าอัตราสหสัมพันธ์ที่ซ่อนเร้นอยู่ ณ เวลา t

ζ_t ค่าผิดปกติ ณ เวลา t

$\omega_1, \omega_2, \dots, \omega_{s-1}$ ค่าสัมประสิทธิ์การถดถอยของตัวแปรคัมมีของฤดูกาล

$S_{1t}, \dots, S_{s-1t}$ ค่าฤดูกาล ณ เวลา t และ σ_Y^2 คือค่าความแปรปรวนโดยทั่วไปของ Y_t

กำหนดค่าการแจกแจงของ prior ในตัวแบบด้วยวิธีการของเบย์คือ

$$\omega_1, \omega_2, \dots, \omega_{s-1} \sim N(0, 1.0E06)$$

$$p(\sigma_Y^2) \sim \text{InvGamma}(0.1, 0.0001)$$

แนวโน้มคือ

$$\Delta W(t | \alpha, \delta) = W(t | \alpha, \delta) - W(t-1 | \alpha, \delta)$$

เมื่อ

$$\alpha \sim N_{[0, \infty)}(\mu_\alpha, \sigma_\alpha^2), p(\mu_\alpha) \sim N(0, 1.0E06),$$

$$p(\sigma_\alpha^2) \sim \text{InvGamma}(0.1, 0.0001)$$

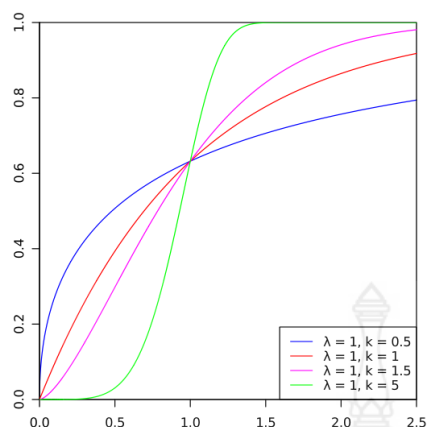
$$\delta \sim N_{[0, \infty)}(\mu_\delta, \sigma_\delta^2), p(\mu_\delta) \sim N(0, 1.0E06),$$

$$p(\sigma_\delta^2) \sim \text{InvGamma}(0.1, 0.0001)$$

$CDF(Weibul)$ ซึ่งเป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเวย์บูลคือ

$$F(y; \alpha, \delta) = \begin{cases} 1 - e^{-(y/\delta)^\alpha}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

และ แสดงด้วยภาพกราฟเมื่อกำหนดให้พารามิเตอร์มีค่าต่าง ๆ กันดังภาพที่ 11

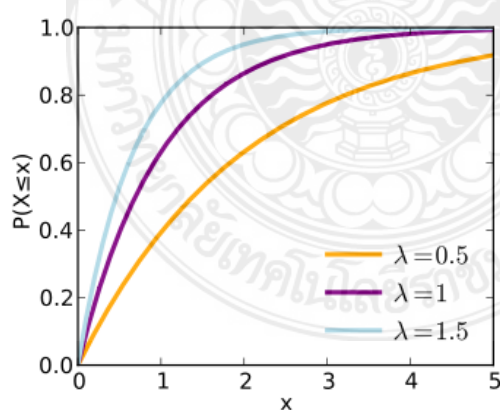


ภาพที่ 11 กราฟฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเวย์บูลเมื่อกำหนดค่าต่างๆของพารามิเตอร์ (ค่า $k = \alpha$ และ $\lambda = \delta$ สำหรับในงานวิจัยนี้)

$CDF(\exp)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเอ็กซ์โพเนนเชียล คือ

$$F(y; \alpha) = \begin{cases} 1 - e^{-\alpha y}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

และ แสดงด้วยภาพกราฟเมื่อกำหนดให้พารามิเตอร์มีค่าต่างๆกันดังภาพที่ 12



ภาพที่ 12 กราฟฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเอ็กซ์โพเนนเชียล เมื่อกำหนดค่าต่างๆของพารามิเตอร์ (ค่า $\lambda = \alpha$ สำหรับในงานวิจัยนี้)

ค่าอัตโนมัติสัมพันธ์ที่ซ่อนเร้นอยู่: $AR(i)$ คือ

$$A_t \sim N(\lambda A_{t-1}, \sigma_A^2), p(\sigma_A^2) \sim \text{InvGamma}(0.1, 0.001) \\ \lambda \sim N(0, 1.0E06), A_0 = 0$$

ค่าผิดปกติคือ

$$\zeta_t \sim \text{Bern}(0.05)$$

ค่าคาดหวังของผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษาคือ

$$\gamma \sim N_{[0, \infty)}(\mu_\gamma, \sigma_\gamma^2), p(\mu_\gamma) \sim N(0, 1.0E06) \\ p(\sigma_\gamma^2) \sim \text{InvGamma}(0.1, 0.0001)$$

ผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษา

$$Z \sim N(\gamma, \sigma_Z^2), p(\sigma_Z^2) \sim \text{InvGamma}(0.1, 0.0001)$$

สมการพยากรณ์ข้อมูลในอนาคตสามารถอธิบายในรูปของฟังก์ชันการแจกแจงของ posterior คือ

$$p(Y_{t+1} | Y_t, Y_{t-1}, \dots, Y_1) = \int \dots \int p(Y_{t+1} | \theta) p(\theta | Y_t, Y_{t-1}, \dots, Y_1) d\theta \quad (56)$$

หรือ

$$p(Y_{t+1} | Y_t, Y_{t-1}, \dots, Y_1) \propto \int \dots \int p(Y_{t+1} | \theta) p(Y_t, Y_{t-1}, \dots, Y_1 | \theta) p(\theta) d\theta \quad (57)$$

จากสมการที่ 56 และ 57 สามารถอธิบายสมการพยากรณ์ข้อมูลไปข้างหน้า โดยการอินทิเกรตซึ่งทำการประมาณค่าคำตอบโดยใช้ MCMC แบบ Gibbs sampling ซึ่งจะได้ค่าประมาณของ $\hat{Y}_{t+1}$

3. สรุปผลการวิจัย

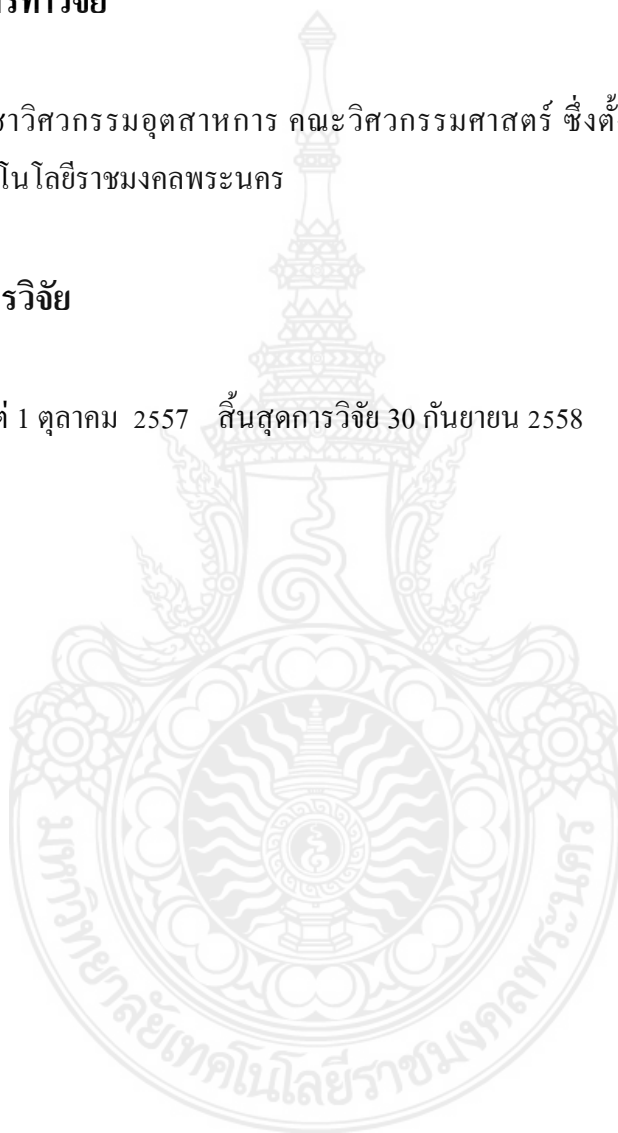
นำผลการวิเคราะห์ที่ได้ มาพิจารณาเพื่อหาข้อสรุป อภิปรายผล และข้อเสนอแนะต่อไป

สถานที่ใช้ในการทำวิจัย

สาขาวิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ ซึ่งตั้งอยู่ที่ศูนย์พระนครเหนือ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

ระยะเวลาในการวิจัย

เริ่มตั้งแต่ 1 ตุลาคม 2557 สิ้นสุดการวิจัย 30 กันยายน 2558



บทที่ 4

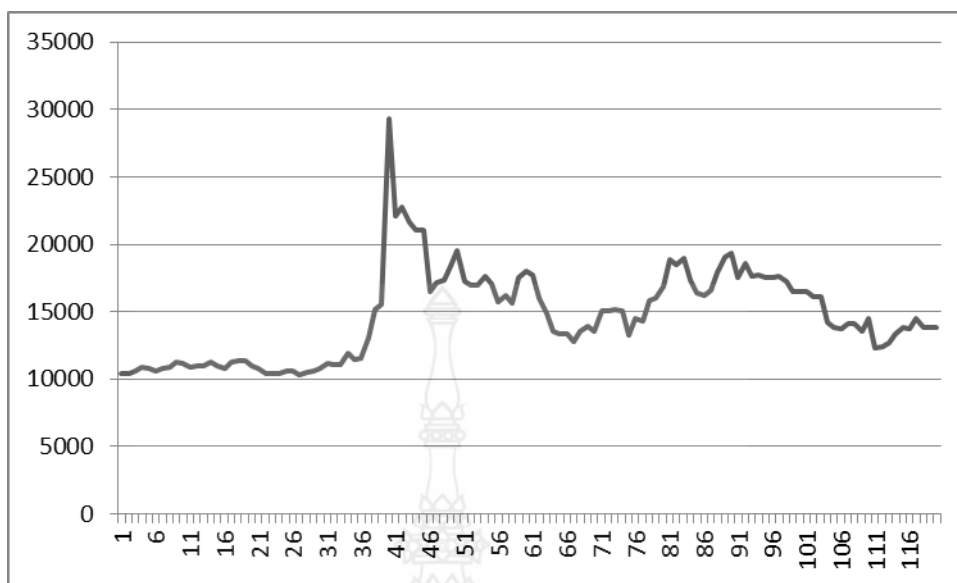
ผลการวิเคราะห์ข้อมูล

ผลการศึกษากระบวนการสโตแคสติกสำหรับการวิเคราะห์ตัวแบบอนุกรมเวลาโดยประยุกต์ใช้กับข้อมูลราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทย ในงานวิจัยนี้ได้ศึกษาข้อมูลเฉพาะข้อมูลราคาปีคของข้าวขาว 5% เฉลี่ยรายเดือน ตั้งแต่ปี 2548-2557 จำนวน 120 เดือน ข้อมูลราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ตั้งแต่ปี 2540-2557 จำนวน 216 เดือน โดยใช้หลักการทางกระบวนการสโตแคสติก ด้วยวิธีการของเบย์ในการประมาณค่าพารามิเตอร์วิธี Markov Chain Monte Carlo (MCMC) และสุ่มตัวอย่างแบบ Gibbs sampling ที่มีตัวแปรตามมีค่าต่อเนื่อง จึงกำหนดให้มีการแจกแจงแบบปกติ และตัวแปรต้นเป็นฟังก์ชันต่างๆและประมาณค่าพารามิเตอร์ดังที่ได้แสดงในบทที่ 3 เมื่อได้ค่าพารามิเตอร์ต่างๆแล้ว จึงนำมาใช้พยากรณ์ไปข้างหน้ากับข้อมูลที่เก็บไว้จำนวนอีก 12 เดือนสุดท้ายเพื่อเปรียบเทียบกับข้อมูลจริงด้วยวิธีการโดยใช้หลักการทางกระบวนการสโตแคสติก แบบเบย์ในตัวแบบการพยากรณ์แบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวกซ์บูลกับแบบเอ็กซ์โพเนนเชียล และตัวแบบการพยากรณ์แบบเอ็กซ์โพเนนเชียล ซึ่งมีรายละเอียดดังนี้

1. ผลการวิเคราะห์ลักษณะของข้อมูล

ภาพที่ 13-15 แสดงลักษณะของข้อมูลข้อมูลราคาปีคของข้าวขาว 5% เฉลี่ยรายเดือน ตั้งแต่ปี 2548-2557 จำนวน 120 เดือน ข้อมูลราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ตั้งแต่ปี 2540-2557 จำนวน 216 เดือน

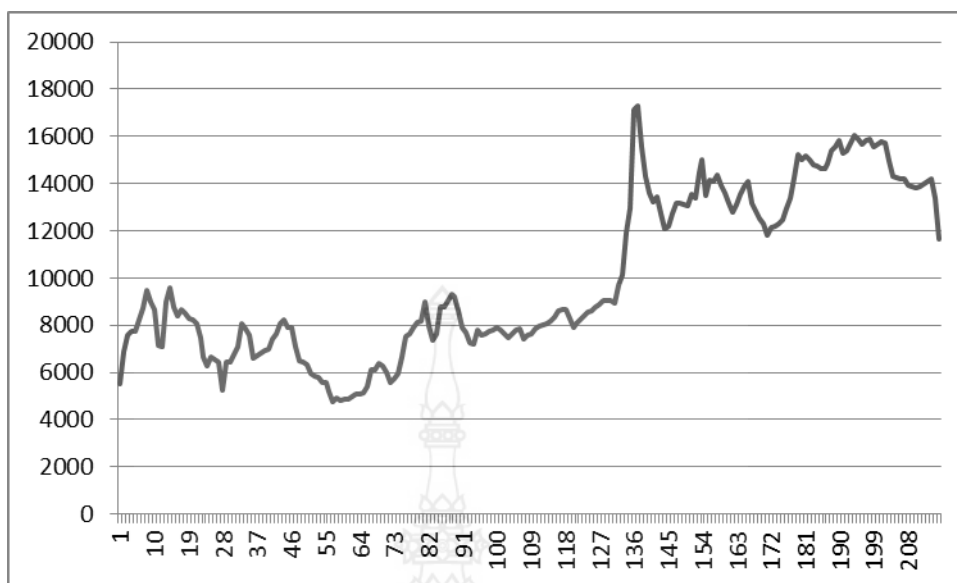
ภาพที่ 16-18 แสดงการวัดสหสัมพันธ์ระหว่างข้อมูล ด้วยกราฟ Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF)



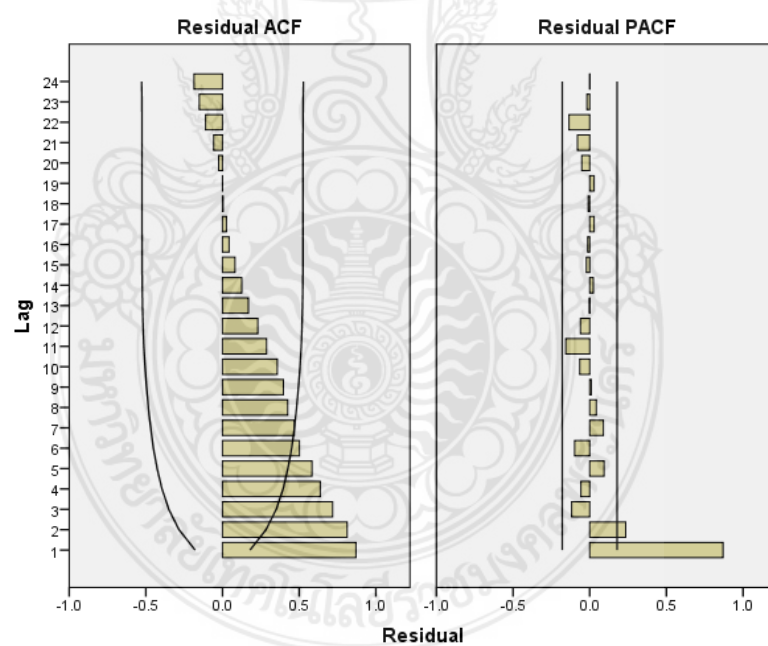
ภาพที่ 13 กราฟแสดงลักษณะของราคาปีคของข้าวขาว 5% เหลือรายเดือน



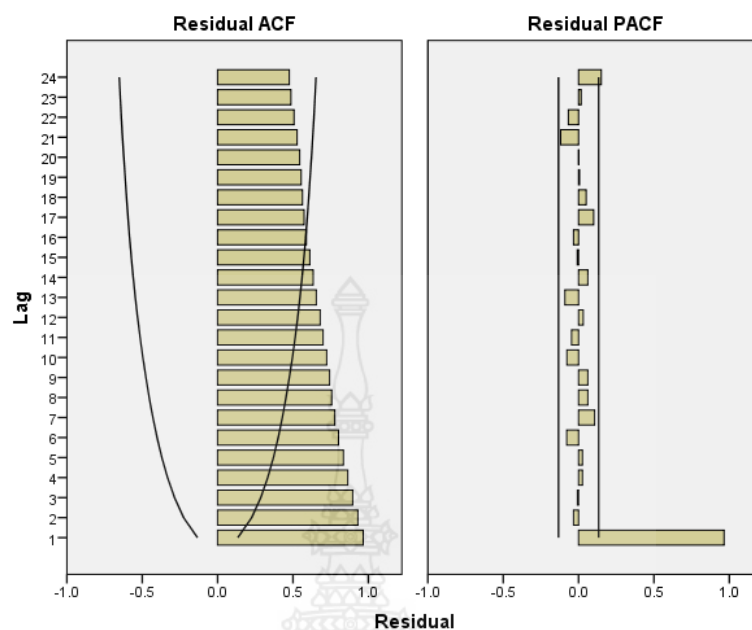
ภาพที่ 14 กราฟแสดงลักษณะของราคาข้าวเปลือกเจ้าความชื้น 15% เหลือรายเดือน



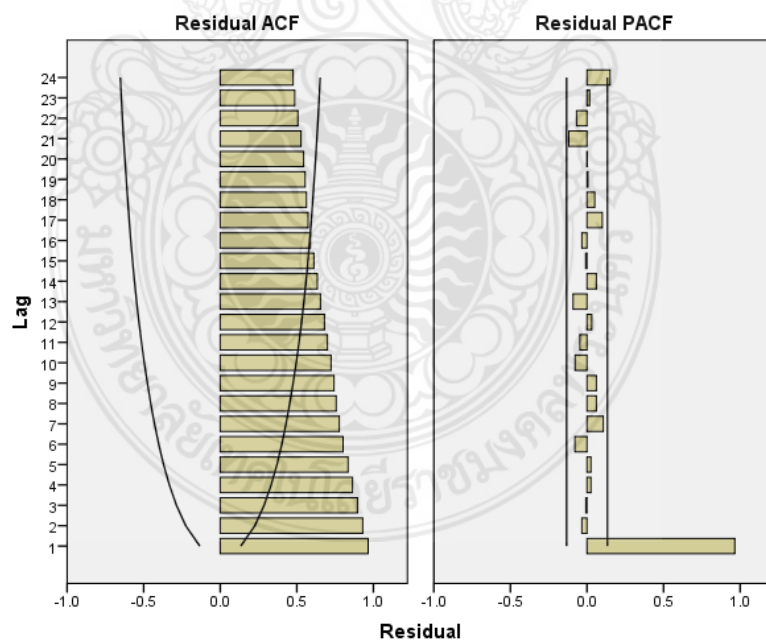
ภาพที่ 15 กราฟแสดงลักษณะของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน



ภาพที่ 16 แสดงการวัดสหสัมพันธ์ของราคาปีคของข้าวขาว 5% เฉลี่ยรายเดือน



ภาพที่ 17 แสดงการวัดสหสัมพันธ์ของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน



ภาพที่ 18 แสดงการวัดสหสัมพันธ์ของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

จากภาพที่ 13-15 แสดงการวิเคราะห์ลักษณะของราคาข้าวแต่ละตัว และแสดงให้เห็นว่า ข้อมูลมีแนวโน้ม มีข้อมูลบางตัวสูง และต่ำผิดปกติจากตัวอื่นๆของแต่ละราคาของข้าว แสดงว่า ข้อมูลมีบางตัวเป็นข้อมูลที่ผิดปกติ ภาพที่ 16-18 เมื่อดูจากกราฟของ Autocorrelation Function และ Partial Autocorrelation Function แสดงให้เห็นว่าข้อมูลมีความสัมพันธ์กันอย่างน้อย 1 lag แต่ lag 1 มีความสัมพันธ์กันมากกว่า lag อื่นๆเพราะ แท่งสูงกว่าแท่งอื่นๆของแต่ละราคาของข้าว ดังนั้นจึง นำลักษณะของข้อมูลที่ได้ไปใช้หลักการทางกระบวนการสโตแคสติก สร้างตัวแบบตามที่ได้ ออกแบบไว้ในบทที่ 3 โดยมีส่วนประกอบต่างๆดังนี้คือ มีแนวโน้ม (Trend) ฤดูกาล (Seasonal) ค่าผิดปกติ (Irregular Variation) ค่าอัตโนมัติสัมพันธ์ (Autocorrelation) สำหรับค่าอัตโนมัติสัมพันธ์ จะเลือกใช้ AR1 lag ของที่ราคาข้าว

2. ผลการใช้หลักการทางกระบวนการสโตแคสติกสร้างตัวแบบเพื่อวิเคราะห์ข้อมูลอนุกรมเวลา

ผลที่ได้จากการวิเคราะห์ข้อมูลข้างต้นมาสร้างตัวแบบเพื่อวิเคราะห์ข้อมูลข้าวแต่ละตัว มีรายละเอียดดังต่อไปนี้

$$Y_t \sim N\left(\gamma(\Delta W(t|\alpha, \delta) + A_t) + \sum_{i=1}^{s-1} \omega_i S_{it}, [\gamma(1 + \zeta_t)\sigma_Y]^2\right)$$

เมื่อ

t คือราคาเฉลี่ยของของข้าวแต่ละตัวที่เก็บมาตามเวลา โดยที่ $t=1, 2, \dots, n$ ซึ่ง ค่าเฉลี่ยของ Y_t คือ

$$E(Y_t) = \gamma(\Delta W(t|\alpha, \delta) + A_t) + \sum_{i=1}^{s-1} \omega_i S_{it}$$

ค่าความแปรปรวนของ Y_t คือ

$$\text{Var}(Y_t) = [\gamma(1 + \zeta_t)\sigma_Y]^2$$

กำหนดค่าการแจกแจงของ prior ในตัวแบบด้วยวิธีการของเบย์ คือ

$$\omega_1, \omega_2, \dots, \omega_{s-1} \sim N(0, 1.0E06)$$

$$p(\sigma_Y^2) \sim \text{InvGamma}(0.1, 0.0001)$$

แนวโน้มนคือ

$$\Delta W(t | \alpha, \delta) = W(t | \alpha, \delta) - W(t-1 | \alpha, \delta)$$

เมื่อ

$$\alpha \sim N_{[0, \infty)}(\mu_\alpha, \sigma_\alpha^2), p(\mu_\alpha) \sim N(0, 1.0E06),$$

$$p(\sigma_\alpha^2) \sim \text{InvGamma}(0.1, 0.0001)$$

$$\delta \sim N_{[0, \infty)}(\mu_\delta, \sigma_\delta^2), p(\mu_\delta) \sim N(0, 1.0E06),$$

$$p(\sigma_\delta^2) \sim \text{InvGamma}(0.1, 0.0001)$$

ค่าอัตสหสัมพันธ์ที่ซ่อนเร้นอยู่: AR(i) จะเลือกใช้ที่ lag 1 คือ

$$A_t \sim N(\lambda A_{t-1}, \sigma_A^2), p(\sigma_A^2) \sim \text{InvGamma}(0.1, 0.001)$$

$$\lambda \sim N(0, 1.0E06), A_0 = 0$$

ค่าผิดปกติคือ

$$\zeta_t \sim \text{Bern}(0.05)$$

ค่าคาดหวังของผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษาคือ

$$\gamma \sim N_{[0, \infty)}(\mu_\gamma, \sigma_\gamma^2), p(\mu_\gamma) \sim N(0, 1.0E06)$$

$$p(\sigma_\gamma^2) \sim \text{InvGamma}(0.1, 0.0001)$$

ผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษา

$$Z \sim N(\gamma, \sigma_Z^2), p(\sigma_Z^2) \sim \text{InvGamma}(0.1, 0.0001)$$

การประมาณค่าพารามิเตอร์โดยวิธี MCMC

จากตัวแบบข้างต้น และกำหนด priors ให้กับพารามิเตอร์ทั้งหมดแล้ว เราจะใช้วิธี Markov Chain Monte Carlo (MCMC) สำหรับประมาณค่าพารามิเตอร์ โดยใช้วิธีการสุ่มตัวอย่างแบบ Gibbs ซึ่งวิธีการสุ่มตัวอย่างแบบ Gibbs นี้จะเหมาะสมกับฟังก์ชันการแจกแจงแบบมีเงื่อนไข เมื่อจำนวนรอบของการสุ่มตัวอย่างมากๆ มันก็จะลู่เข้าสู่ฟังก์ชันการแจกแจงความน่าจะเป็นร่วมของ posterior ใดๆ (joint posterior distribution) สามารถเขียน likelihood ของตัวแบบได้ดังสมการ 58-60

$$\begin{aligned}
 & f(Y_1, \dots, Y_n | \gamma, w(t), \alpha, \delta, A_1, \dots, A_n, \omega_1, \dots, \omega_{s-1}, \xi_1, \dots, \xi_n, \sigma_Y^2, \\
 & \lambda, \mu_\alpha, \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2) \\
 & = \prod_{t=1}^n f(Y_t | \gamma, w(t), \alpha, \delta, A_t, \omega_1, \dots, \omega_{s-1}, \xi_t, \sigma_Y^2, \lambda, \mu_\alpha, \sigma_\alpha^2, \\
 & \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2)
 \end{aligned} \tag{58}$$

สามารถเขียนผลคูณ prior distributions ของพารามิเตอร์ทั้งหมดได้ดังนี้

$$\begin{aligned}
 & [p(\gamma | \mu_\gamma, \sigma_\gamma^2) p(\mu_\gamma) p(\sigma_\gamma^2) p(w(t) | \alpha, \delta) p(\alpha | \mu_\alpha, \sigma_\alpha^2) p(\mu_\alpha) p(\sigma_\alpha^2) p(\delta | \mu_\delta, \sigma_\delta^2) p(\mu_\delta) \\
 & p(\sigma_\delta^2) p(A_1, \dots, A_n | \lambda, \sigma_A^2) p(\lambda) p(\sigma_A^2) p(\omega_1), \dots, p(\omega_{s-1}) p(\xi_1), \dots, p(\xi_n) \\
 & p(\sigma_Y^2)]
 \end{aligned} \tag{59}$$

สามารถเขียน posterior distribution ซึ่งเกิดจากผลคูณของ likelihood กับ ผลคูณ prior distributions ของพารามิเตอร์ทั้งหมด ได้ดังนี้

$$\begin{aligned}
 & p(\gamma, w(t), \alpha, \delta, A_1, \dots, A_n, \omega_1, \dots, \omega_{s-1}, \xi_1, \dots, \xi_n, \sigma_Y^2, \lambda, \mu_\alpha, \\
 & \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2 | Y_1, \dots, Y_n) \\
 & = \prod_{t=1}^n f(Y_t | \gamma, w(t), \alpha, \delta, A_t, \omega_1, \dots, \omega_{s-1}, \xi_t, \sigma_Y^2, \lambda, \mu_\alpha, \\
 & \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2) [p(\gamma | \mu_\gamma, \sigma_\gamma^2) p(\mu_\gamma) p(\sigma_\gamma^2) p(w(t) | \alpha, \delta) \\
 & p(\alpha | \mu_\alpha, \sigma_\alpha^2) p(\mu_\alpha) p(\sigma_\alpha^2) p(\delta | \mu_\delta, \sigma_\delta^2) p(\mu_\delta) p(\sigma_\delta^2) p(A_1, \dots, A_n | \lambda, \sigma_A^2) \\
 & p(\lambda) p(\sigma_A^2) p(\omega_1), \dots, p(\omega_{s-1}) p(\xi_1), \dots, p(\xi_n) p(\sigma_Y^2)]
 \end{aligned} \tag{60}$$

สำหรับวิธีการประมาณค่าพารามิเตอร์ด้วยวิธี MCMC โดยอัลกอริทึมของการสุ่มตัวอย่างแบบ Gibbs จะทำการสร้าง The full conditional distributions ให้กับพารามิเตอร์แต่ละตัว ซึ่ง The full conditional distribution ของพารามิเตอร์แต่ละตัว เกิดจาก ผลคูณของ the likelihood กับ all priors ที่เกี่ยวข้องกับพารามิเตอร์ของตัวเอง ตัวอย่างเช่น the full conditional distributions ของพารามิเตอร์ γ และ α แสดงดังต่อไปนี้ ส่วนตัวอื่นๆก็ใช้หลักการเดียวกันกับที่แสดงดังตัวอย่าง สมการที่ 61-62 (Tongkhaw และ Kantanantha, 2013)

Full conditional distribution ของ γ คือ

$$\begin{aligned} p(\gamma | w(t), \alpha, \delta, A_1, \dots, A_n, \omega_1, \dots, \omega_{s-1}, \xi_1, \dots, \xi_n, \sigma_Y^2, \\ \lambda, \mu_\alpha, \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2, Y_1, \dots, Y_n) \\ = \prod_{t=1}^n f(Y_i | \gamma, w(t), \alpha, \delta, A_t, \omega_1, \dots, \omega_{s-1}, \xi_t, \sigma_Y^2, \\ \lambda, \mu_\alpha, \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2) \\ [p(\gamma | \mu_\gamma, \sigma_\gamma^2) p(\mu_\gamma) p(\sigma_\gamma^2)] \end{aligned} \quad (61)$$

Full conditional distribution ของ α คือ

$$\begin{aligned} p(\alpha | w(t), \gamma, \delta, A_1, \dots, A_n, \omega_1, \dots, \omega_{s-1}, \xi_1, \dots, \xi_n, \sigma_Y^2, \\ \lambda, \mu_\alpha, \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2, Y_1, \dots, Y_n) \\ = \prod_{t=1}^n f(Y_i | \gamma, w(t), \alpha, \delta, A_t, \omega_1, \dots, \omega_{s-1}, \xi_t, \sigma_Y^2, \\ \lambda, \mu_\alpha, \sigma_\alpha^2, \mu_\delta, \sigma_\delta^2, \mu_\gamma, \sigma_\gamma^2, \sigma_Z^2) \\ [p(\alpha | \mu_\alpha, \sigma_\alpha^2) p(\mu_\alpha) p(\sigma_\alpha^2)] \end{aligned} \quad (62)$$

3. ผลของการประเมินประสิทธิภาพของตัวแบบ

การประมาณค่าพารามิเตอร์ด้วยวิธี MCMC โดยใช้อัลกอริทึมของการสุ่มตัวอย่างแบบ Gibbs โดยทำการเขียนอัลกอริทึมในโปรแกรม OpenBUGS จากค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริงมาแล้วจากอัลกอริทึมในโปรแกรม OpenBUGS หลังจากนั้นจะใช้ค่าพารามิเตอร์ทุกตัวมาทำการจำลองสถานการณ์สร้างชุดข้อมูลมาใหม่อีก 1000 ชุดโดยเขียนอัลกอริทึมในโปรแกรม R และประเมินประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ที่สำคัญของแต่ละตัวของ

ฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของแนวโน้มแบบเวย์บูล และแบบเอ็กซ์โพเนนเชียล ได้ผลดังตารางที่ 4-9

ตารางที่ 4 ประเมินประสิทธิภาพของตัวแบบจากพารามิเตอร์ของราคาปีคของข้าวขาว 5%
แบบเวย์บูล

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.021	0.451	0.978	μ_γ	0.024	0.845	0.955
σ_z^2	0.028	0.542	0.968	σ_γ^2	0.021	0.375	0.945
γ	0.025	0.498	0.954	μ_α	0.028	0.954	0.987
λ	0.024	0.654	0.997	σ_α^2	0.025	0.546	0.9378
α	0.028	0.725	0.946	μ_δ	0.024	0.451	0.997
δ	0.024	0.394	0.977	σ_δ^2	0.029	0.542	0.985

ตารางที่ 5 ประเมินประสิทธิภาพของตัวแบบจากพารามิเตอร์ของราคาข้าวเปลือกเจ้าความชื้น 15%
แบบเวย์บูล

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.033	0.416	0.436	μ_γ	0.008	0.416	0.972
σ_z^2	0.031	0.533	0.789	σ_γ^2	0.006	0.429	0.971
γ	0.012	0.143	0.765	μ_α	0.021	0.326	0.968
λ	0.011	0.765	0.944	σ_α^2	0.019	0.344	0.969
α	0.008	0.908	0.958	μ_δ	0.015	0.617	0.971
δ	0.017	0.675	0.678	σ_δ^2	0.014	0.481	0.970

ตารางที่ 6 ประเมินประสิทธิภาพของตัวแบบจากพารามิเตอร์ของราคาข้าวเปลือกเจ้าหอมมะลิ 105
แบบเวย์บูล

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.011	0.446	0.988	μ_γ	0.012	0.216	0.975
σ_z^2	0.001	0.549	0.973	σ_γ^2	0.031	0.329	0.998
γ	0.011	0.156	0.962	μ_α	0.012	0.876	0.975
λ	0.017	0.264	0.920	σ_α^2	0.017	0.253	0.969
α	0.012	0.387	0.956	μ_δ	0.021	0.267	0.985
δ	0.001	0.173	0.929	σ_δ^2	0.019	0.483	0.920

ตารางที่ 7 ประเมินประสิทธิภาพของตัวแบบจากพารามิเตอร์ของราคาปีคของข้าวขาว 5% แบบ
เอ็กซ์โพเนนเชียล

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.121	0.478	0.941	μ_γ	0.112	0.789	0.921
σ_z^2	0.118	0.592	0.952	σ_γ^2	0.111	0.657	0.975
γ	0.104	0.198	0.957	μ_α	0.111	0.854	0.928
λ	0.131	0.898	0.959	σ_α^2	0.116	0.178	0.981
α	0.017	0.835	0.960	μ_δ	0.115	0.965	0.907
δ	0.018	0.779	0.952	σ_δ^2	0.019	0.456	0.935

ตารางที่ 8 ประเมินประสิทธิภาพของตัวแบบจากพารามิเตอร์ของราคาข้าวเปลือกเจ้าความชื้น 15%
แบบเอ็กซ์โพเนนเชียล

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.006	0.886	0.897	μ_γ	0.018	0.416	0.922
σ_z^2	0.021	0.483	0.955	σ_γ^2	0.026	0.429	0.912
γ	0.019	0.429	0.895	μ_α	0.021	0.011	0.901
λ	0.008	0.011	0.955	σ_α^2	0.019	0.008	0.985
α	0.617	0.008	0.936	μ_δ	0.015	0.617	0.932
δ	0.481	0.285	0.878	σ_δ^2	0.014	0.481	0.947

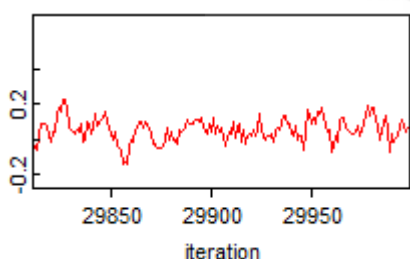
ตารางที่ 9 ประเมินประสิทธิภาพของตัวแบบจากพารามิเตอร์ของราคาข้าวเปลือกเจ้าหอมมะลิ 105 แบบเอ็กซ์โพเนนเชียล

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.025	0.846	0.949	μ_γ	0.019	0.274	0.921
σ_z^2	0.018	0.447	0.940	σ_γ^2	0.008	0.378	0.982
γ	0.009	0.357	0.972	μ_α	0.019	0.198	0.901
λ	0.017	0.224	0.947	σ_α^2	0.017	0.325	0.998
α	0.010	0.818	0.956	μ_δ	0.018	0.573	0.937
δ	0.015	0.567	0.974	σ_δ^2	0.017	0.475	0.900

จากตารางที่ 4-9 พบว่าค่า RB และค่า MSE ของพารามิเตอร์แต่ละตัวของฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของแบบเวย์บูลเอ็กซ์โพเนนเชียลมีค่าต่ำมาก และพบว่าค่า CP ของพารามิเตอร์แต่ละตัวมีค่าสูงมาก ดังนั้นจึงสรุปได้ว่าประสิทธิภาพของตัวแบบทั้งสองตัวในงานวิจัยนี้อยู่ในเกณฑ์ที่ดีมาก

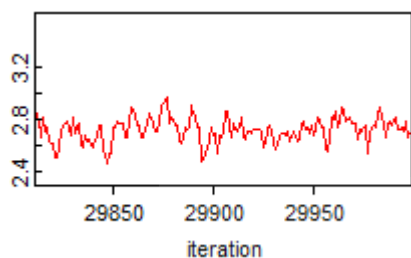
4. ผลของการประมาณพารามิเตอร์

เมื่อได้ตัวแบบที่เหมาะสมแล้วจึงเลือกใช้ค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริง ของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ย จำนวน 108 วัน ราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือนและราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนจำนวน 204 วัน ตัวอย่างผลของ trace plot ที่แสดงการการลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่ง (stationary distribution) ของพารามิเตอร์บางตัว (γ) ของราคาข้าวชนิดต่างๆของทั้งสองตัวแบบแสดงได้ดังภาพที่ 19-24 ส่วนพารามิเตอร์ตัวอื่นที่ไม่ได้นำมาแสดงก็จะมีผลการลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่งเช่นกัน

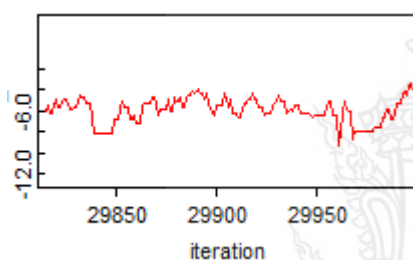


ภาพที่ 19 Trace ของ γ ของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ยของตัวแบบที่มีฟังก์ชัน

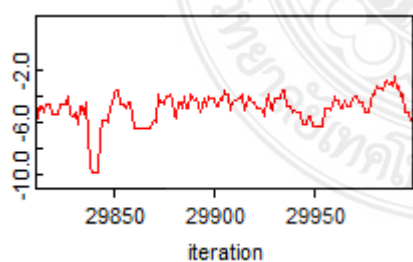
แนวน้มน้แบบเวย์บูล



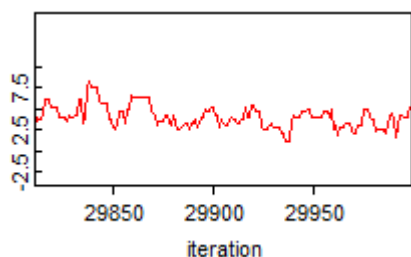
ภาพที่ 20 Trace ของ γ ของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือนของตัวแบบที่มีฟังก์ชันแนวน้มน้แบบเวย์บูล



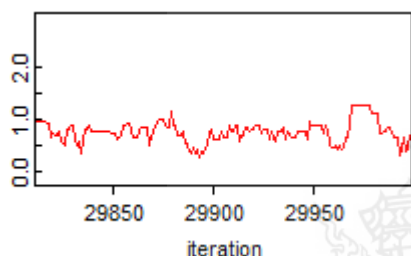
ภาพที่ 21 Trace ของ γ ของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนของตัวแบบที่มีฟังก์ชันแนวน้มน้แบบเวย์บูล



ภาพที่ 22 Trace ของ γ ของราคาปิดของข้าวขาว% รายเดือนเฉลี่ยของตัวแบบที่มีฟังก์ชันแนวน้มน้แบบเอ็กซ์โพเนนเชียล



ภาพที่ 23 Trace ของ γ ของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือนของตัวแบบที่มีฟังก์ชัน
แนวโน้มแบบเอ็กซ์โพเนนเชียล



ภาพที่ 24 Trace ของ γ ของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ของตัวแบบที่มีฟังก์ชัน
แนวโน้มแบบเอ็กซ์โพเนนเชียล

จากภาพ 19-24 แสดงการลู่เข้าของ MCMC โดยพิจารณาจากกราฟของ Trace plot พบว่า MCMC ลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่งเมื่อมีการสุ่มตัวอย่าง แบบกิบส์ จำนวน 30000 รอบ โดยตัด 10000 รอบแรกทิ้ง Trace plot ของทุกราคาข้าวไม่แสดงว่ามีแนวโน้มเกิดขึ้น แต่จะแสดงถึงการลู่เข้าของ ทุกพารามิเตอร์ ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล และฟังก์ชันแนวโน้มแบบเวย์บูล ส่วน Trace plot ตัวอื่นๆของทั้งสองตัวแบบที่ไม่ได้นำมาแสดงพบว่า MCMC ลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่งเช่นเดียวกัน ส่วนผลการประมาณค่าพารามิเตอร์แสดงได้ดังตารางที่ 10-15

ตารางที่ 10 ค่าพารามิเตอร์ของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ยของตัวแบบที่มีฟังก์ชัน
แนวโน้มแบบเวกซ์บูล

Parameter	Value	Parameter	Value
α	1243.82	μ_γ	996.00
λ	33.79	σ_Y^2	19.72
δ	281.51	σ_α^2	-141.75
γ	614.55	σ_δ^2	-23.98
μ_α	-36.58	σ_γ^2	5.55
μ_δ	557.93	σ_z^2	1362.19

ตารางที่ 11 ค่าพารามิเตอร์ของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือนของตัวแบบที่มี
ฟังก์ชันแนวโน้มแบบเวกซ์บูล

Parameter	Value	Parameter	Value
α	-997.61	μ_γ	1381.56
λ	164.83	σ_Y^2	144.23
δ	-47.74	σ_α^2	307.49
γ	-0.22	σ_δ^2	641.17
μ_α	105.93	σ_γ^2	4.17
μ_δ	-1353.10	σ_z^2	561.38

ตารางที่ 12 ค่าพารามิเตอร์ของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ของตัวแบบที่มี
ฟังก์ชันแนวโน้มแบบเวกซ์บูล

Parameter	Value	Parameter	Value
α	1507.03	μ_γ	1044.13
λ	41.84	σ_Y^2	-9.47
δ	468.03	σ_α^2	290.43
γ	597.77	σ_δ^2	573.89
μ_α	-77.33	σ_γ^2	-78.06
μ_δ	704.28	σ_z^2	497.00

ตารางที่ 13 ค่าพารามิเตอร์ของราคาปีคของข้าวขาว 5% รายเดือนเฉลี่ยของตัวแบบที่มีฟังก์ชัน
แนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	1064.60	μ_γ	937.31
λ	133.58	σ_Y^2	18.12
δ	14.79	σ_α^2	-196.11
γ	141.63	σ_δ^2	-70.19
μ_α	83.70	σ_γ^2	-18.65
μ_δ	1404.17	σ_z^2	-1315.69

ตารางที่ 14 ค่าพารามิเตอร์ของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือนของตัวแบบที่มี
ฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	1327.22	μ_γ	1064.52
λ	285.01	σ_Y^2	149.99
δ	539.16	σ_α^2	-16.51
γ	861.86	σ_δ^2	-21.53
μ_α	79.44	σ_γ^2	57.23
μ_δ	-683.20	σ_z^2	1568.52

ตารางที่ 15 ค่าพารามิเตอร์ของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนของตัวแบบที่มี
ฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	1309.86	μ_γ	1510.47
λ	17.90	σ_Y^2	-5.76
δ	-204.56	σ_α^2	625.40
γ	26.27	σ_δ^2	692.41
μ_α	112.45	σ_γ^2	104.73
μ_δ	1340.82	σ_z^2	505.28

จากตารางที่ 10-15 แสดงถึงตัวอย่างของค่าพารามิเตอร์ที่สำคัญบางตัวของทั้งสองตัวแบบที่เข้าสู่การแจกแจงใดแจกแจงหนึ่ง แต่ค่าพารามิเตอร์ตัวอื่นๆที่ไม่ได้แสดงในงานวิจัยนี้ได้เข้าสู่การแจกแจงใดแจกแจงหนึ่งเหมือนกันในภาพที่ 19-24 เช่นกัน ที่ประมาณ 30000 รอบขึ้นไป เช่นกัน ซึ่งสรุปได้ว่าค่าพารามิเตอร์แต่ละตัวจะมีค่าที่ไม่เปลี่ยนแปลงแล้วจึงยอมรับผลของการประมาณค่าพารามิเตอร์ที่ได้ในตารางที่ 10-15 และตัวอื่นๆที่ไม่ได้นำมาแสดงเพื่อนำไปใช้พยากรณ์ค่า y_t ของราคาข้าวเฉลี่ยรายเดือนแต่ละตัวในหัวข้อต่อไป

5. ผลของการเปรียบเทียบตัวแบบ

เมื่อได้ตัวแบบที่เหมาะสมแล้วจึงเลือกใช้ค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริง ของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ย จำนวน 108 วัน ราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือนและราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนจำนวน 204 วัน มาเปรียบเทียบกับวิธีการพยากรณ์ทั้งสามวิธี คือตัวแบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล ตัวแบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล และตัวแบบเอ็กซ์โปเนนเชียล เพื่อหาตัวแบบที่เหมาะสม (Fitting Model) และเปรียบเทียบเพื่อตรวจสอบความถูกต้อง ของตัวแบบ (Validation Model) จากข้อมูลที่เหลืออีก 12 เดือนของทุกราคาข้าว แสดงได้ดังตารางที่ 16-21

ตารางที่ 16 การเปรียบเทียบเพื่อหาตัวแบบที่เหมาะสมของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ย

ข้อมูล	ตัวแบบการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
108 เดือน	1. วิธีเบย์แบบเวย์บูล	0.825	4.622	1.145
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	0.921	5.111	0.941
	3.แบบปรับเรียบเอ็กซ์โปเนนเชียล	5.12	12.784	5.278

ตารางที่ 17 การเปรียบเทียบเพื่อหาตัวแบบที่เหมาะสมของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
204 เดือน	1. วิธีเบย์แบบเวย์บูล	1.591	6.830	1.530
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	1.394	5.734	1.333
	3.แบบปรับเรียบเอ็กซ์โปเนนเชียล	11.301	19.621	11.200

ตารางที่ 18 การเปรียบเทียบเพื่อหาตัวแบบที่เหมาะสมของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

ข้อมูล	ตัวแบบการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
204 เดือน	1. วิธีเบย์แบบเวย์บูล	1.411	6.777	1.431
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	1.382	5.222	1.307
	3. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	12.378	19.998	12.240

ตารางที่ 19 การเปรียบเทียบเพื่อตรวจสอบความถูกต้องของตัวแบบของราคาปีคของข้าวขาว5% รายเดือนเฉลี่ย

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
12 เดือน	1. วิธีเบย์แบบเวย์บูล	11.127	18.191	11.221
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	12.247	18.978	11.951
	3. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	12.943	19.874	12.514

ตารางที่ 20 การเปรียบเทียบเพื่อตรวจสอบความถูกต้องของตัวแบบของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
12 เดือน	1. วิธีเบย์แบบเวย์บูล	19.895	28.972	19.742
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	18.187	27.201	18.495
	3. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	19.995	29.842	19.842

ตารางที่ 21 การเปรียบเทียบเพื่อตรวจสอบความถูกต้องของตัวแบบของราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

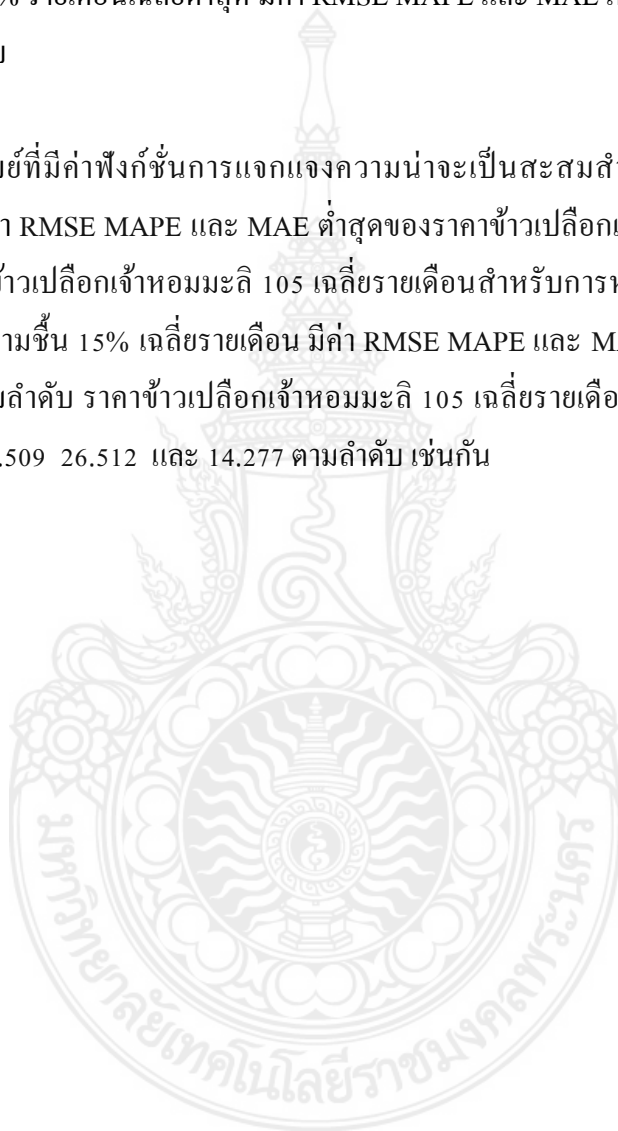
ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
12 เดือน	1. วิธีเบย์แบบเวย์บูล	15.342	26.369	15.444
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	14.509	26.512	14.277
	3. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	15.974	26.997	15.940

จากตารางที่ 16-18 ได้จากการที่นำข้อมูลจริงของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ยจำนวน 108 วัน ราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนจำนวน 204 วัน มาประมาณค่าพารามิเตอร์เพื่อการเปรียบเทียบหาตัวแบบที่เหมาะสมจากค่าพยากรณ์ของทั้งสามวิธี กับข้อมูลจริง พบว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นเหมาะสมสำหรับแนวโน้มแบบแบบเวย์บูล ให้ค่า RMSE MAPE และ MAE ของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ยต่ำสุด มีค่า RMSE MAPE และ MAE เท่ากับ 0.825 4.622 และ 1.145 ตามลำดับ

ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นเหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนสำหรับการหาตัวแบบที่เหมาะสม โดยข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน มีค่า RMSE MAPE และ MAE เท่ากับ 1.394 5.734 และ 1.333 ตามลำดับ ราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนมีค่า RMSE MAPE และ

MAE เท่ากับ 1.382 5.222 และ 1.307 ตามลำดับ ตารางที่ 19-21 ได้จากการที่นำข้อมูลจริงที่เหลืออีก จำนวน 12 เดือน มาประมาณค่าพารามิเตอร์เพื่อเปรียบเทียบการตรวจสอบความถูกต้องของตัวแบบ (Validation Model) จากค่าพยากรณ์ของทั้งสามวิธี พบว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวนอนแบบแบบเวย์บูล ให้ค่า RMSE MAPE และ MAE ของราคาปิดของข้าวขาว 5% รายเดือนเฉลี่ยต่ำสุด มีค่า RMSE MAPE และ MAE เท่ากับ 11.127 18.191 และ 11.221 ตามลำดับ

ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวนอนแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของราคาข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนสำหรับการหาตัวแบบที่เหมาะสม โดยข้าวเปลือกเจ้าความชื้น 15% เฉลี่ยรายเดือน มีค่า RMSE MAPE และ MAE เท่ากับ 18.187 27.201 และ 18.495 ตามลำดับ ราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือนมีค่า RMSE MAPE และ MAE เท่ากับ 14.509 26.512 และ 14.277 ตามลำดับ เช่นกัน



บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

สำหรับการทำงานวิจัยเรื่อง กระบวนการสโตแคสติก สำหรับการวิเคราะห์ตัวแบบอนุกรมเวลา โดยประยุกต์ใช้กับข้อมูลราคาจริง และราคาในตลาดล่วงหน้าของข้าวในประเทศไทย ในครั้งนี้ มีจุดประสงค์เพื่อ ใช้หลักการของกระบวนการสโตแคสติกเพื่อสร้างตัวแบบสำหรับประมาณค่าพารามิเตอร์ต่างๆ กับข้อมูลราคาจริง และราคาในตลาดล่วงหน้าของข้าวในประเทศไทย สรุปผล อภิปรายผล และมีข้อเสนอแนะดังต่อไปนี้

สรุปผลการวิจัย

1. ลักษณะทั่วไปของข้อมูลที่ใช้ศึกษา

ข้อมูลที่ใช้ศึกษาคือ ข้อมูลราคาปิดของข้าวขาว 5% เฉลี่ยรายเดือน ตั้งแต่ปี 2548-2557 จำนวน 120 เดือน ข้อมูลราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ตั้งแต่ปี 2540-2557 จำนวน 216 เดือน จากการพล็อตกราฟของข้อมูลตามเวลาหาความสัมพันธ์ระหว่างข้อมูล ณ เวลาปัจจุบัน กับข้อมูล ณ เวลาก่อนหน้า ด้วยกราฟ Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) พบว่าข้อมูลมีแนวโน้ม ข้อมูลมีบางตัวเป็นข้อมูลที่ผิดปกติ ข้อมูลมีความสัมพันธ์กันอย่างน้อย 1 lag แต่ lag 1 มีความสัมพันธ์กันมากกว่า lag อื่นๆ ดังนั้นจึงนำลักษณะของข้อมูลที่ได้ไปใช้หลักการทางกระบวนการสโตแคสติก โดยมีส่วนประกอบต่างๆ ดังนี้คือ มีแนวโน้ม (Trend) ฤดูกาล (Seasonal) ค่าผิดปกติ (Irregular Variation) ค่าอัตโนมัติสัมพันธ์ (Autocorrelation) สำหรับค่าอัตโนมัติสัมพันธ์ จะเลือกใช้ AR1 lag ของที่ราคาข้าว

2. หลักการของกระบวนการสโตแคสติกเพื่อสร้างตัวแบบการพยากรณ์ข้อมูลอนุกรมเวลาโดยวิธีของเบย์

การสร้างตัวแบบในงานวิจัยนี้เริ่มจากตรวจสอบลักษณะของข้อมูลจากการพล็อตกราฟของข้อมูลตามเวลา การหาค่าความสัมพันธ์ระหว่างข้อมูล ณ เวลาปัจจุบัน กับข้อมูล ณ เวลา

ก่อนหน้า ด้วยกราฟ Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) เพื่อหาลักษณะประกอบต่างๆ ของข้อมูลอนุกรมเวลา ซึ่งพบว่ามีองค์ประกอบของแนวโน้ม มีค่าผิดปกติ และอัตตสหสัมพันธ์ หลังจากนั้นนำส่วนประกอบต่างๆ ที่ไปไปออกแบบสร้างตัวแบบ โดยวิธีของเบย์ ตัวแบบที่ได้คือ

$$Y_t \sim N\left(\gamma(\Delta W(t|\alpha, \delta) + A_t) + \sum_{i=1}^{s-1} \omega_i S_{it}, [\gamma(1 + \zeta_t)\sigma_Y]^2\right)$$

เมื่อ t คือราคาเฉลี่ยของของข้าวแต่ละตัวที่เก็บมาตามเวลา โดยที่ $t = 1, 2, \dots, n$

$$\text{ค่าเฉลี่ยของ } Y_t \text{ คือ } E(Y_t) = \gamma(\Delta W(t|\alpha, \delta) + A_t) + \sum_{i=1}^{s-1} \omega_i S_{it}$$

$$\text{ค่าความแปรปรวนของ } Y_t \text{ คือ } \text{Var}(Y_t) = [\gamma(1 + \zeta_t)\sigma_Y]^2$$

$$\text{แนวโน้มคือ } \Delta W(t|\alpha, \delta) = W(t|\alpha, \delta) - W(t-1|\alpha, \delta)$$

$$\text{ค่าอัตตสหสัมพันธ์ที่ซ่อนเร้นอยู่: AR(1) จะเลือกใช้ lag 1 คือ } A_t \sim N(\lambda A_{t-1}, \sigma_A^2)$$

$$\text{ค่าผิดปกติคือ } \zeta_t \sim \text{Bern}(0.05)$$

กำหนดให้แนวโน้มมีรูปแบบต่างกันคือ มีรูปแบบเป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมแบบเวย์บูล และรูปแบบฟังก์ชันการแจกแจงความน่าจะเป็นสะสมแบบเอ็กซ์โปเนนเชียล

3. การประเมินประสิทธิภาพของตัวแบบ

การประเมินประสิทธิภาพของตัวแบบใช้ค่า RB MSE และค่า CP เมื่อได้ค่าประมาณพารามิเตอร์ที่ประมาณด้วยวิธีการของเบย์ จากข้อมูลจริงมาแล้ว จะใช้ค่าประมาณพารามิเตอร์ทุกตัวมาทำการจำลองสถานการณ์ สร้างชุดข้อมูลมาใหม่อีก 1000 ชุด ประเมินประสิทธิภาพของตัวแบบ โดยดูจากค่า RB ค่า MSE และค่า CP ของพารามิเตอร์แต่ละตัว ซึ่งค่า RB และค่า MSE มีค่าต่ำมาก และค่า CP ของพารามิเตอร์แต่ละตัวมีค่าสูงมาก สรุปได้ว่าประสิทธิภาพของตัวแบบในงานวิจัยนี้อยู่ในเกณฑ์ที่ดีมาก

4. การเปรียบเทียบตัวแบบ

ตัวแบบที่นำมาเปรียบเทียบในงานวิจัยนี้ได้แก่ ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล กับตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล โดยเปรียบเทียบจากค่า RMSE MAPE และ MAE ซึ่งทำการเปรียบเทียบสองส่วนคือ ส่วนที่ 1 คือส่วนของข้อมูลใช้สำหรับการหาตัวแบบที่เหมาะสม (Fitting Model) และส่วนที่ 2 คือส่วนของข้อมูลที่ใช้ตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ผลการวิจัยพบว่า ผลการวิจัยพบว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของราคาข้าว 2 ชนิด ได้แก่ ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 รายเดือนเฉลี่ย ทั้งการหาตัวแบบที่เหมาะสม (Model Fitting) และการตรวจสอบความถูกต้องของตัวแบบ (Model Validation) ส่วนตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูลให้ค่า RMSE MAPE และ MAE ต่ำสุดคือ ราคาปัดของข้าวขาว 5% ทั้งการหาตัวแบบที่เหมาะสม และการตรวจสอบความถูกต้องของตัวแบบ

อภิปรายผล

ตัวแบบที่นำเสนอสามารถพยากรณ์ราคาข้าวได้อย่างดี เพราะว่าตัวแบบที่นำเสนอ นำองค์ประกอบต่างๆ ที่ประกอบกันขึ้นเป็นอนุกรมเวลา มาไว้ในตัวแบบ ทั้งค่าแนวโน้ม ฤดูกาล อັตตสหสัมพันธ์ และค่าผิดปกติ ค่าแนวโน้ม เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะราบเรียบ แนวโน้มอาจมีลักษณะเป็นเส้นตรงหรือเส้นโค้งในทางเพิ่มขึ้นหรือลดลง ค่าแนวโน้มของข้อมูลเป็นการเคลื่อนไหวในช่วงระยะเวลาที่ค่อนข้างนานพอสมควร ฤดูกาล เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะการเพิ่มขึ้น หรือลดลงในลักษณะเดียวกันของรอบระยะเวลาหนึ่งที่แน่นอน เรียกว่า การเปลี่ยนแปลงตามฤดูกาล อັตตสหสัมพันธ์ เป็นค่าที่แสดงความสัมพันธ์ของข้อมูล ณ เวลาปัจจุบัน กับข้อมูล ณ เวลาก่อนหน้า และ ค่าผิดปกติ เป็นการเปลี่ยนแปลงของข้อมูลอนุกรมเวลาที่เกิดจากเหตุการณ์ที่เราไม่สามารถคาดการณ์ได้ล่วงหน้า เช่น การเกิดอุทกภัย ซึ่งเหตุการณ์เหล่านี้เป็นสิ่งที่เกิดขึ้นโดยบังเอิญไม่คาดคิดมาก่อน

การประมาณค่าพารามิเตอร์ใช้วิธีการของเบย์ ข้อดีของวิธีการประมาณค่าพารามิเตอร์แบบเบย์คือ สามารถใช้กับตัวอย่างขนาดเล็ก ในขณะที่การประมาณค่าแบบอื่นเช่น วิธี Maximum

Likelihood ตัวอย่างต้องมีขนาดใหญ่ การประมาณค่าจึงจะแม่นยำ วิธีการของเบย์ สามารถประมาณค่าพารามิเตอร์ หรือฟังก์ชันของพารามิเตอร์ได้โดยตรง โดยไม่ต้องประมาณค่าพารามิเตอร์บางตัวมาก่อน และประการสำคัญที่สุดคือวิธีการของเบย์สามารถใช้กับตัวแบบที่มีความซับซ้อน เช่นกรณีของตัวแบบที่นำเสนอ วิธีการอื่นไม่สามารถนำวิธีการอื่นมาใช้ได้

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ในการวิจัยครั้งต่อไป วิธีการของเบย์สามารถเพิ่มส่วน หรือลดประกอบต่างๆ ในข้อมูลอนุกรมเวลาได้เพื่อให้เหมาะสมกับการวิเคราะห์ข้อมูลแม่นยำเพิ่มขึ้น เช่น ฟังก์ชันของค่าผิดปกติ ฟังก์ชันของฤดูกาล รูปแบบอื่นๆ เป็นต้น
2. ในส่วนของการหาตัวแบบที่เหมาะสม (Fitting Model) จะให้ค่า Error ต่ำมาก แต่ในส่วนของการตรวจสอบความถูกต้องของตัวแบบ (Validation Model) มีค่า Error สูง ซึ่งจากการศึกษา งานวิจัยในต่างประเทศพบว่าเกิด Over fitting ดังนั้นควรจะมีการศึกษาในส่วนของ Over fitting นี้ต่อไปเพื่อให้ได้ค่าพยากรณ์ที่มีความแม่นยำมากขึ้น
3. สามารถนำวิธีการวิเคราะห์ไปประยุกต์ใช้งานข้อมูลอนุกรมเวลากับ งานทางด้านอื่นๆ ได้
4. ใช้เป็นพื้นฐานในการพัฒนาตัวแบบให้วิเคราะห์กับข้อมูลใกล้เคียงกับโลกแห่งความเป็นจริงมากที่สุด

บรรณานุกรม

บำเพ็ญ ไชยรักษ์ 2557. บทความอนาคตข้าวไทย, National geographic.

<http://www.ngthai.com/ArticleDetail.aspx?ArticleId=102#.ViheJ9IrLIU>

วราวุธ วุฒินิชย์, 2554. การวิเคราะห์ทางสถิติเคสศึกษาสำหรับระบบชลประทาน, ภาควิชาวิศวกรรม

ชลประทาน, คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเกษตรศาสตร์กำแพงแสน

วโรทัย โกศลพิศิษฐ์กุล, 2551. บทความปีทองของข้าวไทย, สำนักงานเศรษฐกิจการคลัง,

<http://www.fpo.go.th/S-I/Source/Article/Article79.htm>

Berger,D.E. **Introduction to Multiple Regression**.Claremont Graduate University.

Bernd A. Berg. 2004. **Markov Chain Monte Carlo Simulations and Their Statistical Analysis**. Singapore, World Scientific.

Bisgaard.S., Kulahci,M.2011. **Time Series Analysis and Forecasting by Example**. John Wiley & Sons. Inc.

Boonsarawongse, R. and H.C. Co. 2007. Forecasting Thailand's rice export: Statistical techniques vs artificial neural network. **Computers & industrial engineering**. 53:610-627.

Box, George E.P.,G.M.Jenkins and G. C. Reinsel.1994.**Time Series Analysis Forecasting and Contral**. 3rd ed.,Prentice-Hall,Inc.USA.

Brace,N., Kemp,R., Snelgar,R.,2006. **Spss for Psychologists: A Guide to Data Analysis Using Spss for Windows**.Palgrave Macmillan, Houndmills, UK

- Broemeling, L. and M. Land. 1984. On forecasting with univariate autoregressive process: A Bayesian approach. **Comm. Statist. Theory Methods**. 13:1305-1320.
- Cipra, T., J. Trujillo and A. Rubio. 1995. Holt-Winters method with missing observations. **Manag. Sci.** 41:174-78.
- Cipra, T. 2006. Exponential smoothing for irregular data. **Appl. Math.** 51:597-604.
- Congdon, P. 2006. **Bayesian Statistical Modelling**, 2nd ed., John Wiley & Sons: New York, pp. 1-56.
- Deetae, N.1991. **A comparative study of forecasting techniques on rice, cassava and mungbean. Farm Gate Prices**. Master Thesis, Kasetsart University, Thailand.63-64.
- Dembo, A. 2013. **Stochastic Processes**, Lecture note, Department of Statistics, Stanford University, Stanford, CA. 49-62.
- Evans, M., N. Hastings and J.B. Peacock. 2000. **Statistical Distribution**.3rd ed., John Wiley and Sons, Inc., New York.
- Gnanadesikan, R. and J.R. Kettenring. 1972. Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data. **Biometrics**. 28: 81-124.
- Hyndman, R.J., Koehle, A.B., Snyder, R.D. and Grose, S. 2002. A state space framework for automatic forecasting using exponential smoothing methods. **Int. J. Forecasting**. 18: 439-454.
- Iqbal, N., K. Bakhsh, A. Maqbool and A.S. Ahmad. 2005. Use of the ARIM Model for Forecasting Wheat Area and Production in Pakistan. **Journal of Agriculture & Social Sciences**. 1(2):120-122.

- Kahforoushan,E., Zarif,M.,and Mashahir,E.B.2010. Prediction of added value of agricultural subsectionsusing artificial neural networks: Box-Jenkins and Holt-Winters methods. **Journal of Development and Agricultural Economics** .2(4):115-121.
- Kerdsomboon, M. 1999. **Forecasting of agricultural products and prices**. Master Thesis, Chulalongkorn University, Thailand. 197-200.
- Lee, R.D. and L.R. Carter. 1992. Modeling and forecasting U.S. mortality (discussion 671-675). **Journal of the American Statistician Association**. 87: 659-671.
- Lui, S. 1994. Mutiperiod Bayesian forecasts for AR models. Ann. **Inst. Statist. Mat.** 46(3):429-452.
- Mishra, A.K. and V.R . Desai. 2005. Drought forecasting using stochastic models. **Stochastic Environmental Research and Risk Assessment**. 19(5): 326-339.
- Monahan, J.F. 1983. Fully Bayesian analysis of ARMA time series models. **J. Econometrics**. 21:307-3631.
- Montgomery,D.C.,Jennings,C.L.,Kulahci,M.2008.**Introduction to Time Series Analysis and Forecasting**A. John Wiley & Sons. Inc.
- National News Bureau of Thailand. **Thai agricultural area and path towards the kitchen of the world**. blog, 3 Jan, 2011, <http://thainews.prd.go.th/view>
- Neelamegham, R. and P. Chintagunta. 1999. A Bayesian model to forecast new product performance in domestic and international markets. **Marketing science**. 18(2): 115-136.

- Pedroza, C. 2006. A Bayesian forecasting model: predicting U.S. male mortality. **Biostatistics**. 7(4): 530-550.
- Pena, D. and F.J. Prieto. 2001. Multivariate Outlier Detection and Robust Covariance Matrix Estimation. **Technometrics**. 43: 286-300.
- Rencher, A.C. 1998. **Multivariate Statistical Inference and Applications**. John Wiley & Sons, Inc., New York.
- Robert, C. 2001. **The Bayesian Choice: From Decision-Theoretic Foundation**, 2nd ed., Springer: New York, pp. 8-31.
- Robert, C.P. and G. Casella. 2004. **Monte Carlo Statistical Methods**, 2nd ed., Springer: New York, pp. 267-291.
- Sangpattaranate, P. 2005. **Forecasting of rice prices in Thailand**. Master Thesis, Kasetsart University, Thailand. 149-153.
- SAS. 2011. SAS/STAT®9.3 User's Guide. SAS Institute Inc., North Carolina.
- Sumer, K.K., O. Goktas and A. Hepsag. 2009. The application of seasonal latent variable in forecasting electricity demand as an alternative method. **Energy Policy**. 37(4): 1317-1322.
- Tongkhaw, P. and N. Kantanantha. 2011. Appropriate forecasting models for fluctuating vegetable prices in Thailand, pp. 162-166. In IEEE Proceedings, The CET 2011. Shanghai.
- Tongkhaw, P. and N. Kantanantha. 2012. Bayesian model for time series with trend, autoregression and outliers, pp. 90-94. In Proceedings of the 10th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering). Bangkok.

- Tongkhaw, P. and N. Kantanantha. 2013. Bayesian models for time series with covariates, trend, seasonality, autoregression and outliers. *Journal of Computer Science* 9(3): 291-298.
- Tsay, R. S. **TSMODEL Algorithms**. The University of Chicago, USA.
- West, M., Harrison, J. 1997. **Bayesian forecasting and dynamic models**, 2nd ed., Springer. New York
- Wright, D.J. 1986. Forecasting data published at irregular time intervals using extension of Holt's method. **Manag. Sci.** 32: 499-510.
- Yan, X., and Su, X.G. 2009. **Linear Regression Analysis Theory and Computing**. World Scientific Publishing Company.
- Yelland, P. 2009. Bayesian forecasting for low-count time series using stat-space models: An empirical evaluation for inventory management. **International Journal of Production Economics**. 118: 95-103.
- Yelland, P.M. 2010. Bayesian Forecasting of Part Demand. **International Journal of Forecasting**. 26: 374-396.

ประวัติผู้วิจัย

ชื่อ-สกุล: นายพิษณุ ทองขาว

(Mr. Pitsanu Tongkhaw)

ตำแหน่ง: อาจารย์มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

คณะวิศวกรรมศาสตร์ สาขาวิศวกรรมอุตสาหกรรม

การศึกษา: ว.ศ.ด. วิศวกรรมอุตสาหกรรม มหาวิทยาลัยเกษตรศาสตร์

ชื่อ-สกุล: ผศ. นิตยา บุญสิทธิ

(Assist. Prof. Nittaya Boonsith)

ตำแหน่ง: ผู้ช่วยศาสตราจารย์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

คณะวิทยาศาสตร์และเทคโนโลยี สาขาวิชาคณิตศาสตร์และสถิติ

การศึกษา: ค.ม. (คณิตศาสตร์) จุฬาลงกรณ์มหาวิทยาลัย

