บทที่ 2

เอกสาร และงานวิจัยที่เกี่ยวข้อง

กระบวนการสโตแคสติกสำหรับการวิเคราะห์ตัวแบบอนุกรมเวลาโดยประยุกต์ใช้กับข้อมูลราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทยในครั้งนี้ผู้ศึกษาได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องตามหัวข้อดังต่อไปนี้

ทฤษฎีที่เกี่ยวข้อง
งานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1. สโตคาสติก (Stochastic) (Dembo, 2013)

สโตคาสติกจะเกี่ยวข้องกับตัวแปรสุ่ม (Random Variable) เช่น การประมาณค่าพารามิเตอร์ในตัวแบบการพยากรณ์ การวิเคราะห์ระบบสินค้าคงคลัง โดยที่ค่าพารามิเตอร์ในตัวแบบนั้นเป็นตัวแปรสุ่ม ซึ่งอยู่ในภาพของฟังก์ชันของความน่าจะเป็น วิธีการประมาณค่าที่นิยมใช้กันอยู่ทั่วไป เช่น การประมาณค่าความควรจะเป็นสูงสุด (Maximum Likelihood Estimation) วิธีนี้นิยมใช้ประมาณค่าพารามิเตอร์ที่เป็นตัวแปรสุ่มโดยที่ตัวแปรสุ่มแต่ละตัวมีค่าของ prior distribution แค่ระดับเดียวหรือมีหลายระดับชั้น ฟังก์ชั่นต้องไม่มีความซับซ้อนเท่าไรจึงจะคำนวณหาค่าพารามิเตอร์มาได้ แต่ถ้าตัวแปรสุ่มแต่ละตัวมีค่าของฟังก์ชันความน่าจะเป็นหรือ prior distribution หลายระดับชั้นซึ่งจะเรียกว่า hyperprior distribution และเมื่อหา joint distribution มาแล้ว ได้ฟังก์ชันที่มีความซับซ้อนมาก เราก็จะนิยมใช้วิธีการประมาณค่าที่สามารถประมาณฟังก์ชันที่มีความซับซ้อนได้ เช่น วิธีการประมาณแบบเบย์ (Bayesian Estimation) เป็นต้น

จากที่กล่าวมาข้างต้นจะเรียกว่าวิธีกระบวนการสโตคาสติก (Stochastic Process) เพราะกระบวนการประมาณค่าพารามิเตอร์นอกจากจะมีตัวแปรสุ่มแล้ว ยังมีเวลาหรือจำนวนการวนซ้ำมาเข้าเกี่ยวข้องในกระบวนการด้วย วิธีของเบย์จะใช้หลักการของห่วงโซ่มาร์คอฟ (Markov Chain) มาช่วยในการประมาณ ซึ่งวิธีการของห่วงโซ่มาร์คอฟ ถือว่าเป็นวิธีการของกระบวนการสโตคาสติกที่ได้รับนิยมมากพอสมควร และถูกนำไปประยุกต์ใช้ในหลายหลายสาขาวิชา ในที่นี้จะขอกล่าวกระบวนการสโตคาสติกคร่าวๆ โดยมี S เป็นเซตของสถานที่เป็นไปได้ทั้งหมด (State Space) ก็คือเซตของตัวแปรสุ่ม $\{X_t, t \in \mathrm{T}\}$ โดยที่ X_t จะมีค่าความน่าจะเป็น และสามารถหาค่าความน่าจะเป็นของมันได้ โดยจะเหมือนกับเซตของตัวแปรสุ่มที่เราใช้กันอยู่ทั่วๆไป เช่น f(x), p(x) เป็นต้น เซต T จะถูกเรียกว่าเซตพารามิเตอร์ของกระบวนการ ถ้า $T = \mathbb{N} = \{0,1,2,...\}$ กระบวนการนี้จะเรียกว่ากระบวนการหนึ่งของการประมาณพารามิเตอร์แบบไม่ต่อเนื่อง (A discrete parameter process) แต่ถ้า T ไม่สามารถนับได้ กระบวนการนี้เรียกว่ากระบวนการหนึ่งของการประมาณพารามิเตอร์แบบต่อเนื่อง (A continuous parameter process) โดยทั่วๆไป T จะเป็นจำนวนจริงบวกโดยที่ $T = \mathbb{R}_+ = (0,\infty]$ และ $T = [a,b] \subset \mathbb{R}$ เมื่อค่า t คือเวลา และเมื่ออยู่ในภาพของ X_t ก็จะเรียกว่า สถานะ หรือ ตำแหน่ง หรือกระบวนการ ณ เวลา t ถ้า S อยู่ใน $\mathbb{N}$ ก็จะเรียกว่าเซตของจำนวนเต็มทั้งหมด หรือ เซตจำกัดเซตหนึ่ง สำหรับค่าคงที่ $\omega \in \Omega$ การจับคู่จะได้ดังนี้

$$t \to X_t(\omega)$$

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

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

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

กระบวนการสโตแคสติก คือ เซทของตัวแปรสุ่ม $\{X_t : X_1, X_2, \ldots, X_n\}$

X_1 = ตัวแปรสุ่ม ซึ่งมี $f(x_1)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น

X_2 = ตัวแปรสุ่ม ซึ่งมี $f(x_2)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น

X_n = ตัวแปรสุ่ม ซึ่งมี $f(x_n)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็น

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

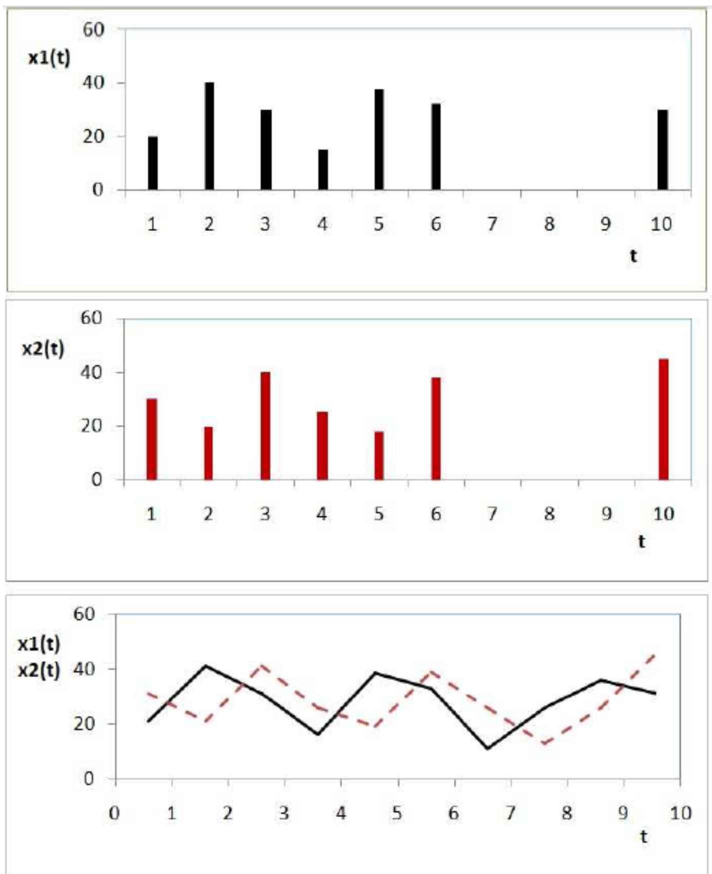
แบบจำลองซึ่งอธิบายโครงสร้างความน่าจะเป็นของเซทของค่าสังเกต $\{x_1, x_2, \ldots, x_n\}$ เรียกว่า แบบจำลองสโตแคสติก (Stochastic Model)

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

รูปแบบของกระบวนการสโตแคสติก สามารถแบ่งออกได้เป็นหลายลักษณะดังนี้

1) กระบวนการสโตแคสติกแบบ Discrete vs. Continuous

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



ภาพที่ 1 Two realizations of stochastic processes

2) Independent (Purely Random) vs. Dependent Process

กระบวนการสโตแคสติกแบบอิสระ (Independent Process) และแบบไม่อิสระ (Dependent Process) จะพิจารณาจาได้จากคุณสมบัติของฟังก์ชันการแจกแจงความน่าจะเป็นร่วมของหลายตัวแปรสุ่ม (Joint Probability Distribution Function) ดังนี้

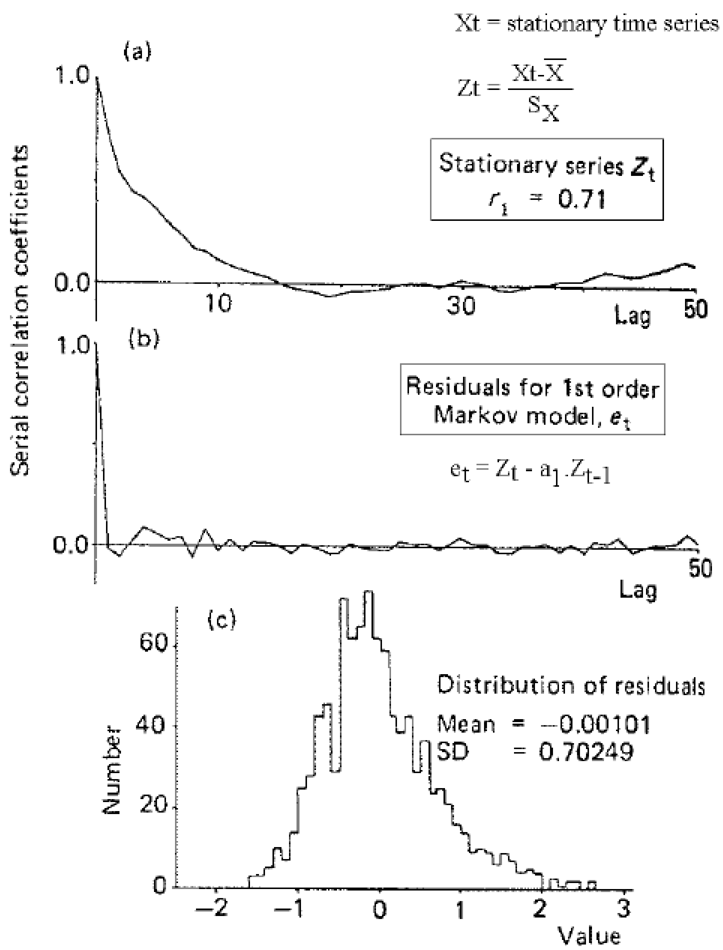
ถ้า $f(x_1, x_2, \ldots, x_n) = f(x_1) f(x_2) \ldots f(x_n)$ หรือ ฟังก์ชันความน่าจะเป็นร่วมของหลายตัวแปรสุ่มเท่ากับผลคูณของ Marginal Distribution แสดงว่าเป็น Independent Process

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

Serial Dependence Stochastic Process (Time Dependence) หรือ กระบวนการทางสโตแคสติกที่ขึ้นอยู่กับเวลา

Spartially Dependence Stochastic Process (Area Dependence) หรือ กระบวนการทางสโตแคสติกที่ขึ้นอยู่กับพื้นที่

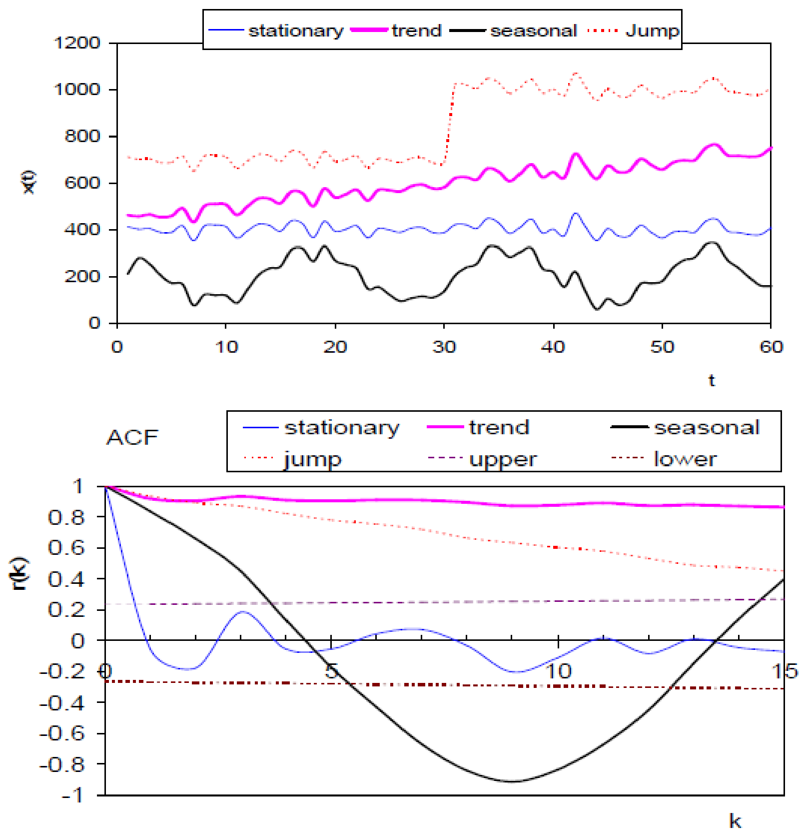
คุณสมบัติทางสถิติที่สามารถแสดงให้เห็นว่าอนุกรมเวลาเป็นแบบอิสระหรือไม่ คือ ฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function: ACF) ตัวอย่าง ACF ของ Stationary Dependent and Independent Series แสดงอยู่ในภาพที่ 2



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 กระบวนการซึ่งพารามิเตอร์เปลี่ยนไปตามเวลา ปรากฏการณ์ 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

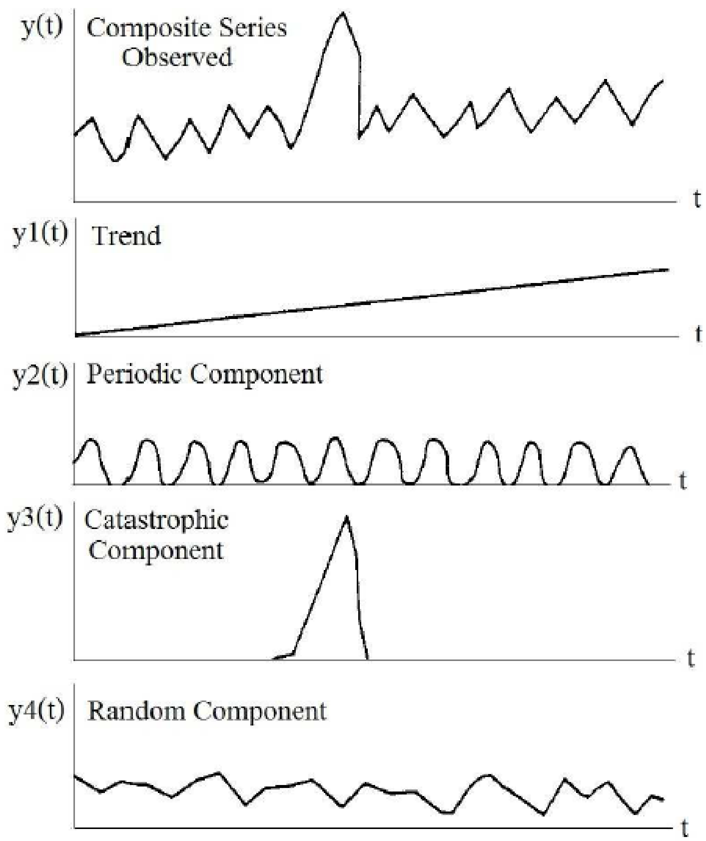
$$Y(t) = Y_1(t) + Y_2(t) + Y_3(t) + Y_4(t) \qquad (1)$$

$Y_1(t)$ = Trend

$Y_2(t)$ = Periodic Component

$Y_3(t)$ = Catastrophic Component

$Y_4(t)$ = Random Component

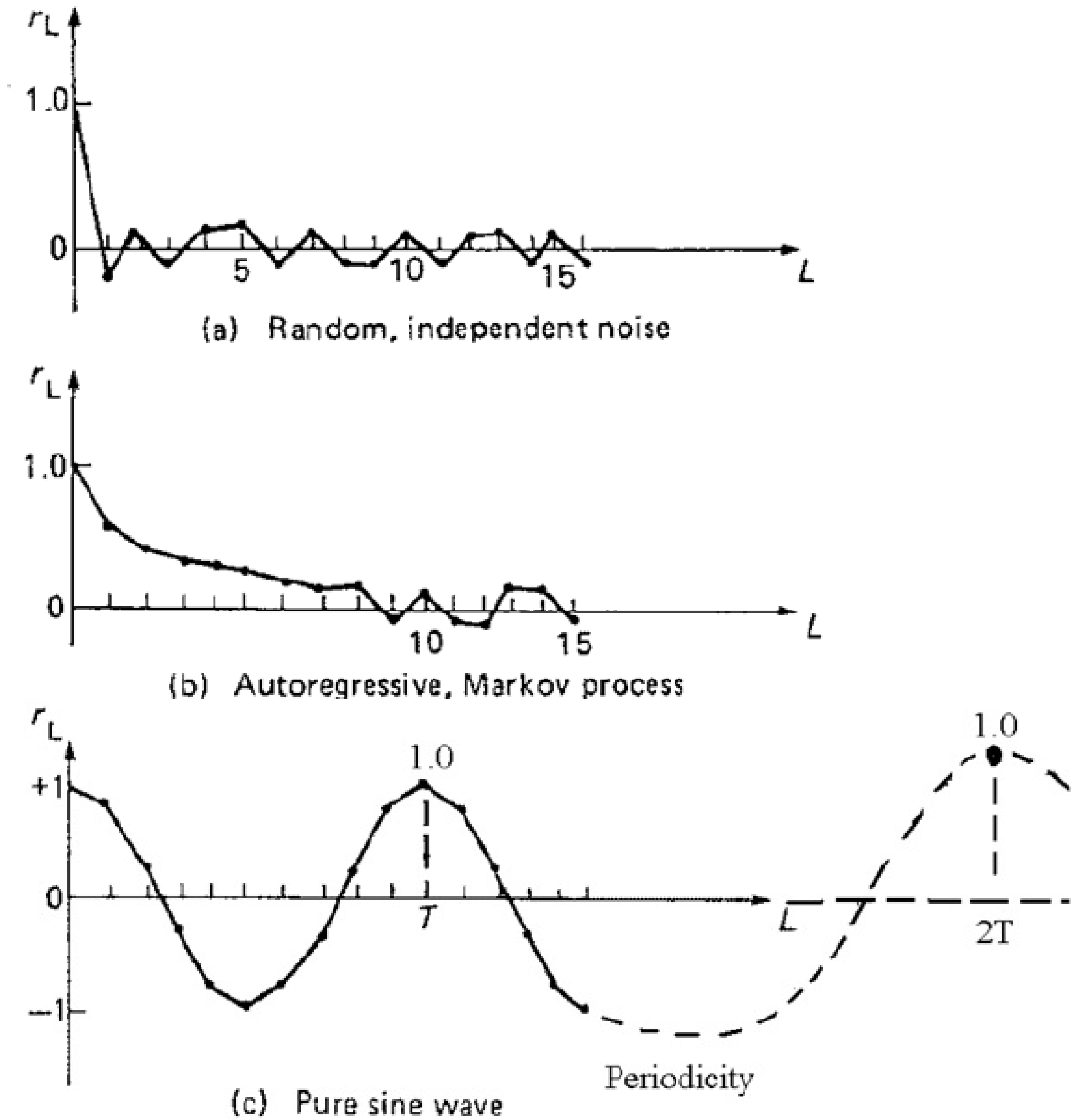


ภาพที่ 4 Components of Time Series

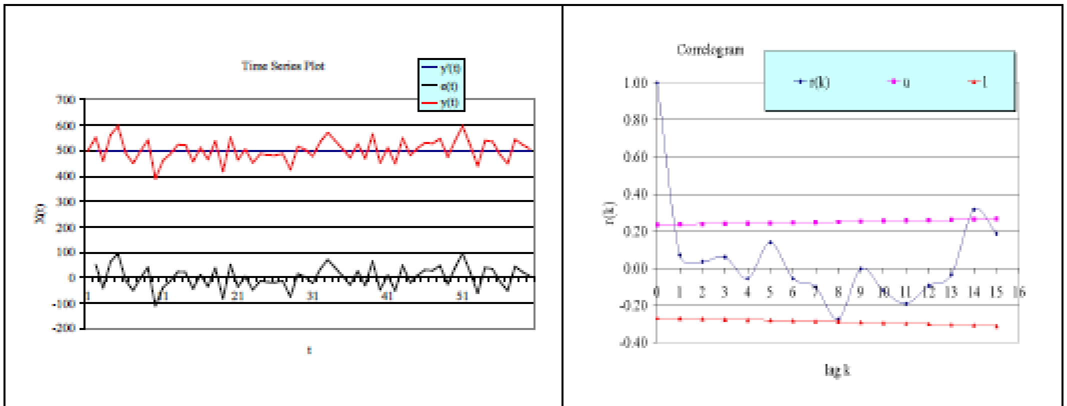
Autocorrelation Function (ACF) หรือเรียกอีกอย่างหนึ่ง คือ Correlogram ของกระบวนสโตแคสติกแบบต่างๆ แสดงอยู่ในภาพที่ 5 ส่วนภาพที่ 6 แสดงการพล๊อตเปรียบเทียบระหว่างอนุกรมเวลาและ ACF ของอนุกรมเวลาแบบต่างๆ ACF คือคุณสมบัติที่สำคัญของ Periodic Time Series

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

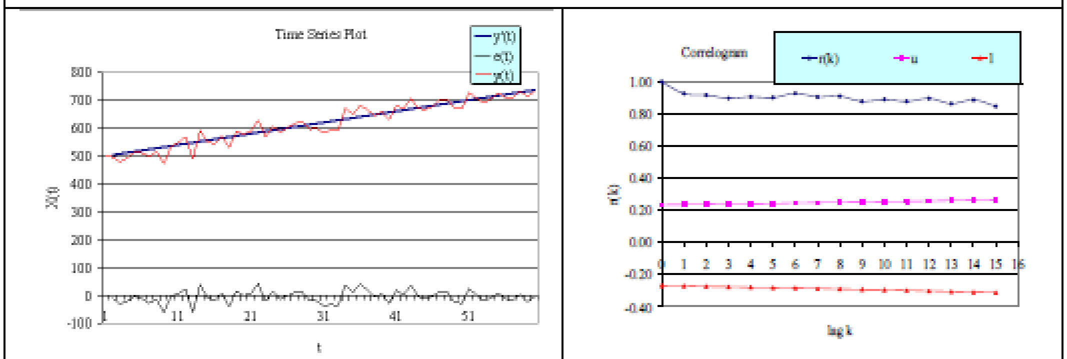
และ Spectral Density Function ของ Periodic Time Series ที่มีการ Remove Periodic Components ที่มี Frequency ต่างๆ



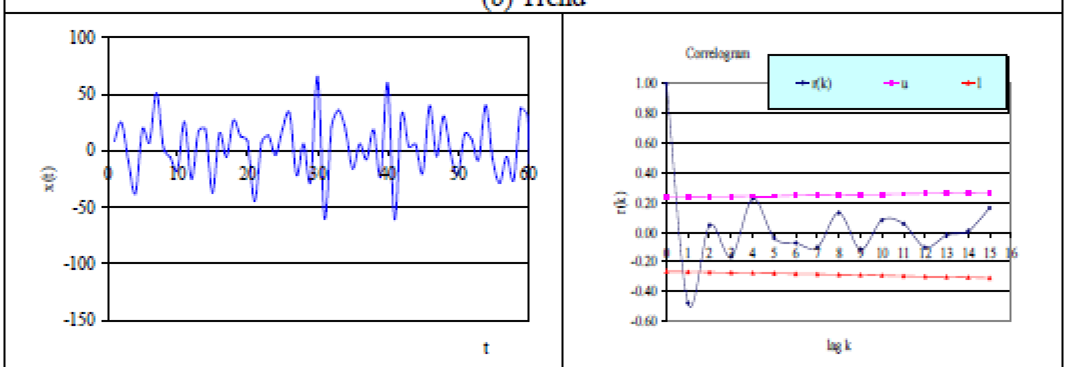
ภาพที่ 5 Correlogram or ACF Showing Independent, Markov Process and Periodicity



(a) Stationarity

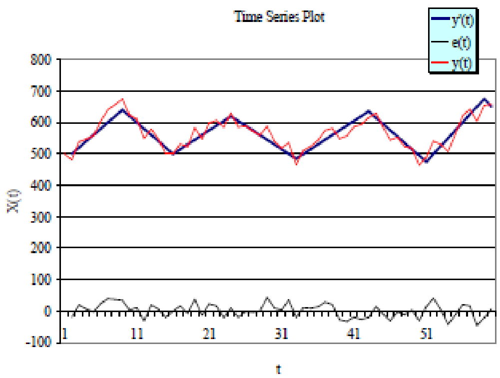


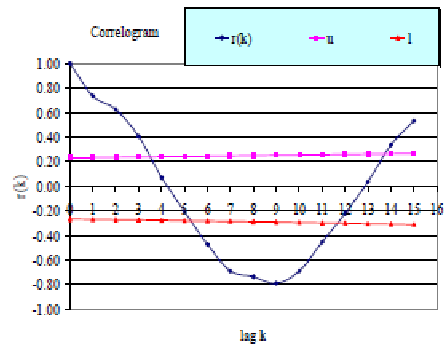
(b) Trend



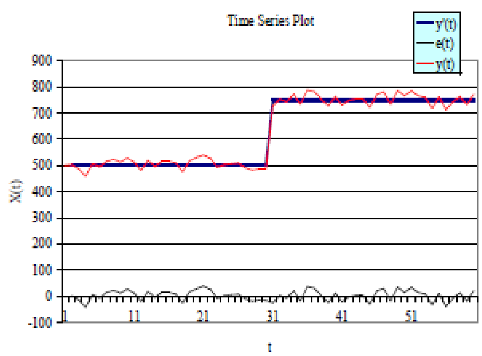
(c) After Removing Trend From (b) by Differencing

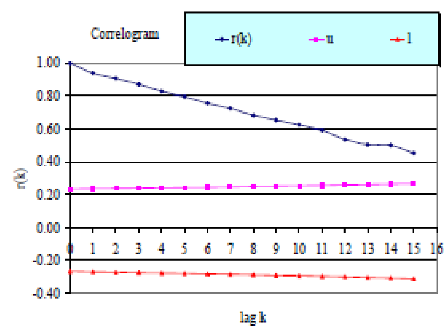
ภาพที่ 6 Characteristics of Time Series and Correlogram for Different Types of Time Series





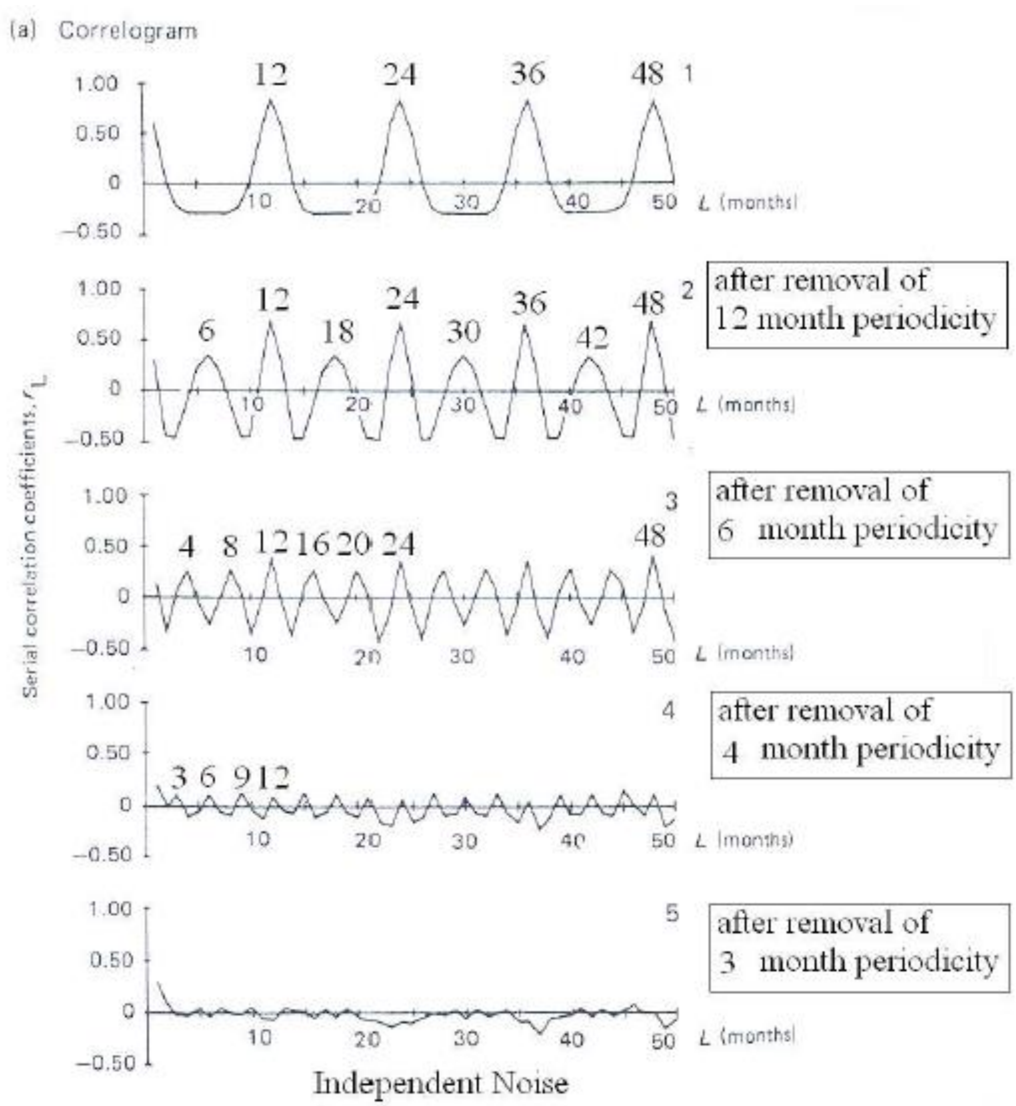
(d) Cycle/periodicity





(d) Jump

ภาพที่ 6 (ต่อ)



(a) ACF After Removing Periodicity (Time domain)

ภาพที่ 7 ACF and Spectral Density Function after Removing Periodicity

1.3 คุณสมบัติพื้นฐานของกระบวนการสโตแคสติก (Basic Properties of Stochastic Process)

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

$E(X_t) = \mu_t$ โดยที่ $t = 1, 2, \ldots, N$

$VAR(X_t) = \sigma_t^2$ โดยที่ $t = 1, 2, \ldots, N$

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

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

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 $f(x;\theta)$

$$x_t = \mu + \sigma\varepsilon_t \ ; \ t = 1, 2, \ldots \qquad (2)$$

ε_t = Independent Series

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

μ_t, 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

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

$$\varepsilon_t = \phi\varepsilon_{t-1} + \zeta_t \tag{3}$$

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

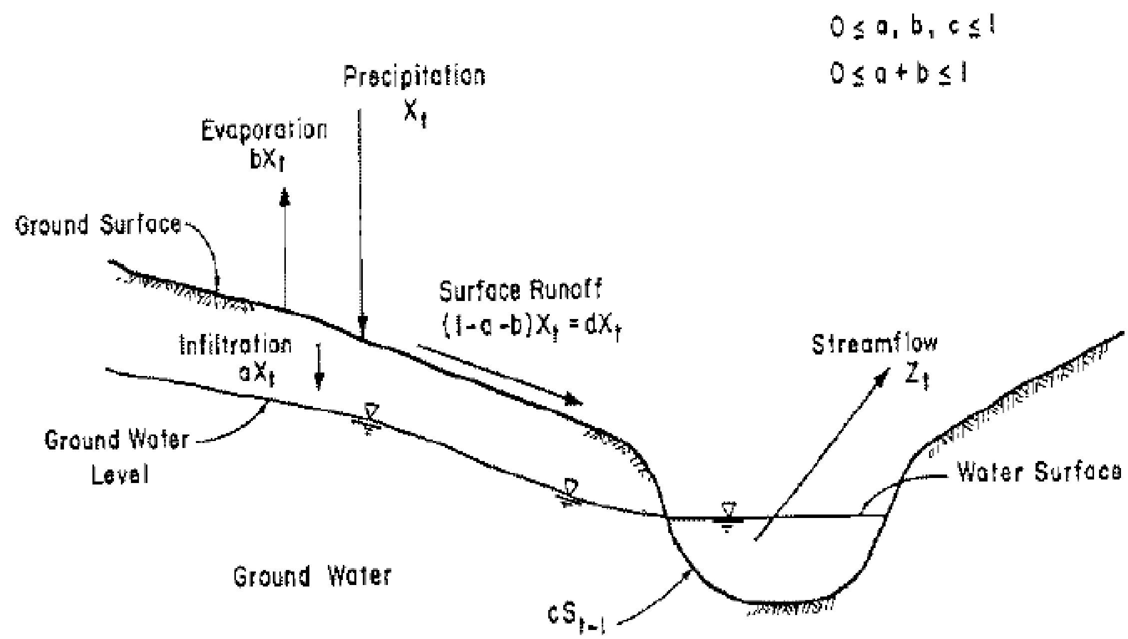
ให้ Y_t = Time Dependent Annual Runoff Series

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

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

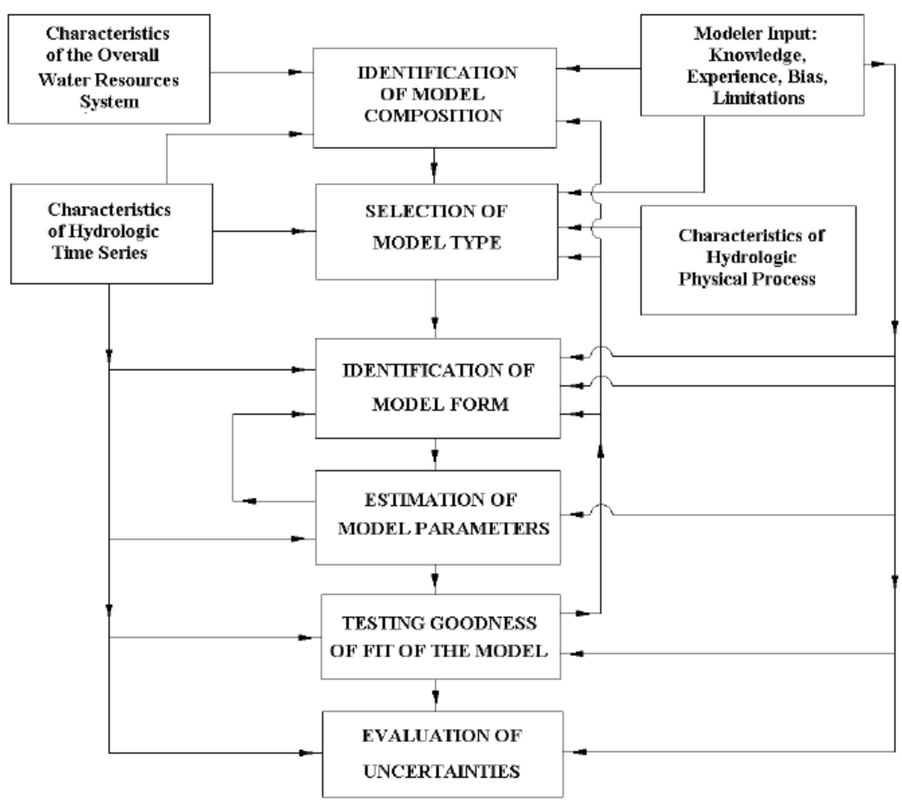
4) มีข้อจำกัดในการ Generate ข้อมูลใหม่จำนวนมากๆ

5) ขาดพื้นฐานทางด้านกายภาพ (Physical Basis)

6) มีพารามิเตอร์มากเกินไป

1.6.1 วิธีทางสมมาตร (Systematic Approach)

สำหรับการจำลองอนุกรมเวลาทางอุทกวิทยา Systematic Approach สำหรับการจำลองอนุกรมเวลาทางอุทกวิทยา 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) การหาภาพลักษณะของแบบจำลอง (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. ตรวจพิสูจน์ (Verify) ว่าแบบจำลองเป็นไปตามสมมติฐานที่กำหนดหรือไม่ เช่น Residual (ζ_t) เป็นอิสระและมีการแจกแจงปกติ (Independent and Normality)

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

(1) ตรวจสอบ Model Correlogram เปรียบเทียบกับ Historical Correlogram

(2) ตรวจสอบค่าสถิติของอนุกรมเวลาที่สังเคราะห์ขึ้นแล้วเปรียบเทียบค่าสถิติกับอนุกรมเวลาของค่าสังเกต เช่น ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ ค่าสหสัมพันธ์ ค่าสถิติที่เกี่ยวข้องกับปริมาณน้ำในอ่างเก็บน้ำ (Storage Related Statistics) และค่าสถิติที่เกี่ยวกับความแห้งแล้ง (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 (วราวุธ, 2554)

วิธีการประมาณค่าพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method: MLM)

เมื่อเราต้องการวัดค่าที่แท้จริงของปริมาณบางอย่าง (x_T)

– ทำได้โดยการวัดค่าของปริมาณนั้น ซ้ำๆ กันหลายครั้ง $\{x_1, x_2, \ldots, x_n\}$

– 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 เป็นวิธีการทั่วไปที่ใช้เพื่อประมาณค่าพารามิเตอร์ของสิ่งที่สนใจจากชุดข้อมูลที่มี

ข้อตกลงเบื้องต้นของวิธีภาวะน่าจะเป็นสูงสุด

– กำหนดให้สุ่มตัวอย่างค่าวัด $\{x_1, x_2, \ldots, x_n\}$ 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

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

– 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

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

- ค่า α อาจจะเป็นชุดของค่าพารามิเตอร์ (เช่น เป็นความชันและระยะแกน) หรือเป็นค่าเดี่ยวก็ได้

- สมการที่ใช้คำนวณหาช่วงของค่า α อาจได้จากสมการเชิงเส้นอย่างง่ายจับคู่กับสมการไม่เป็นเชิงเส้น

ตัวอย่าง

– กำหนดให้ $f(x,\alpha)$ มีการแจกแจงปกติ

– ให้ $\alpha = \mu$ เป็นค่าเฉลี่ยของการแจกแจงปกติ

– เราต้องการประมาณค่าที่ดีที่สุดของ α จากชุดข้อมูลที่เราวัดได้ คือ $\{x_1, x_2, \ldots, x_n\}$

– ให้ข้อกำหนดเบื้องต้น ข้อมูลที่ได้มีการกระจายเท่ากัน คือมีส่วนเบี่ยงเบนมาตรฐานเท่ากัน เป็น σ ดังนั้น ข้อมูลแต่ละตัวมีการแจกแจงดังนี้

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

– ฟังก์ชันภาวะน่าจะเป็นสูงสุดของปัญหานี้ คือ

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

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

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

– 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

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

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 จาก Likelihoodกับ Prior

ตัวแบบที่ซับซ้อนสามารถใช้ตัวแบบเบย์แก้ปัญหาได้ เช่นใช้ตัวแบบเบย์ที่มี 3 ขั้น ได้แก่ ขั้นตอนที่ 1 ระบุการการแจกแจงของตัวแปรสุ่มที่สังเกตค่าได้เมื่อกำหนดพารามิเตอร์ให้ขั้นตอนที่ 2 ระบุการแจกแจงของพารามิเตอร์เมื่อกำหนดไฮเปอร์พารามิเตอร์ให้ และขั้นตอนที่ 3 ระบุการแจกแจงของไฮเปอร์พารามิเตอร์ในทำนองเดียวกัน จำนวนขั้นตอนอาจมีมากกว่า 3 ได้ ตัวแบบเบย์สามารถเพิ่มความแกร่ง (Robustness) ให้กับตัวประมาณแบบเบย์ได้ เนื่องจากความไม่แน่นอน (Uncertainty) ถูกนำมาคิดไว้ในขั้นตอนของการแจกแจงของ Prior นอกจานั้นวิธีการของเบย์ยังทำให้การประมาณค่าพารามิเตอร์ใน Posterior ง่ายขึ้น โดยใช้การการจำลองสถานการณ์ (Simulation) การจำลองสถานการณ์ที่ใช้กันอย่างแพร่หลายคือ วิธีเชิงตัวเลข Markov Chain Monte Carlo (MCMC)

สำหรับตัวอย่างเพื่อให้เห็นภาพรวมของวิธีเบย์ จะขอยกตัวอย่างตัวแบบที่มีความซับซ้อนจึงใช้วิธีการของเบย์ในการแก้ปัญหา ตัวแบบคือ

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

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

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

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

Prior (มีหลายระดับ) คือ

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

2) ข้อมูลผิดปกติ (Outliners)

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

3) ราคาหรือปริมาณผลผลิตรวมทุกช่วงเวลา

$\gamma_i \sim \mathrm{N}(\mathrm{g}_i,{\sigma_\gamma}^2)$, $p(\sigma_\gamma) \propto \mathrm{U}nif(0,\infty)$

$g_j \sim \mathrm{N}(\mu_g,{\sigma_g}^2)$, $p(\mu_g) \propto 1$, $p(\sigma_g) \propto \mathrm{U}nif(0,\infty)$

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

4) Autoregression ที่ซ่อนเร้นอยู่ (Latent Autoregression)

$X_{it} \sim \mathrm{N}(\lambda_{i1}\mathrm{X}_{it-1}+\lambda_{i2}\mathrm{X}_{it-2},{\sigma_x}^2)$, $\sigma_x = 0.8\sigma_\varepsilon$

$(\lambda_{i1},\lambda_{i2})^T \sim \mathrm{N}(\boldsymbol{\mu}_{\lambda},\boldsymbol{\Sigma}_{\lambda})$, $\mathrm{p}(\boldsymbol{\mu}_{\lambda},\boldsymbol{\Sigma}_{\lambda}) \propto |\boldsymbol{\Sigma}_{\lambda}|^2$

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

5) พารามิเตอร์อื่นๆ

$\alpha_i \sim \mathrm{N}(\mathrm{a}_i,\sigma_{\alpha}^{\ 2})$, $p(\sigma_{\alpha}) \propto \mathrm{U}nif(0,\infty)$

$\delta_i \sim \mathrm{N}(\mathrm{d}_i,\sigma_{\delta}^{\ 2})$, $p(\sigma_{\delta}) \propto \mathrm{U}nif(0,\infty)$

$a_j \sim \mathrm{N}(\mu_a,\sigma_a^{\ 2})$, $p(\mu_a) \propto 1$, $p(\sigma_a) \propto \mathrm{U}nif(0,\infty)$

$d_j \sim \mathrm{N}(\mu_d,\sigma_d^{\ 2})$, $p(\mu_d) \propto 1$, $p(\sigma_d) \propto \mathrm{U}nif(0,\infty)$

การประมาณค่าพารามิเตอร์ในตัวแบบดังกล่าวมีขั้นตอนดังนี้

a. สร้าง Likelihood จากการแจกแจงของ Y_{it}

b. สร้าง Posterior จากผลคูณของ Likelihood กับ prior ทุกตัว

c. จำลองสถานการณ์ด้วยวิธีการ MCMC โดยใช้การเขียนโปรแกรม ใน Open bugs และ R หรือโปรแกรมคณิตศาสตร์ต่างๆ

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

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

Gibbs sampling (Geman and Geman, 1984)

เป็นวิธีการสร้าง MCMC จากการสุ่มตัวอย่างแบบวนซ้ำจากการแจกแจงแบบมีเงื่อนของแต่ละพารามิเตอร์เมื่อรู้พารามิเตอร์ที่เหลือทั้งหมดและข้อมูล
สมมติว่า Posterior คือ $\pi(\boldsymbol{\theta}\,|\,\mathbf{Y})$ ที่มีมิติขนาด k โดยที่ $\mathbf{Y}$ แทนข้อมูลที่สังเกตค่าได้ และสำหรับแต่ละ θ_i ของ $\boldsymbol{\theta}$ การแจกแจงแบบมีเงื่อนของแต่ละพารามิเตอร์เมื่อรู้พารามิเตอร์ที่เหลือทั้งหมดและข้อมูลคือ $\pi(\theta_i\,|\,\theta_1,...,\theta_{i-1},\theta_{i+1},...,\theta_k,\mathbf{Y})=\pi(\theta_i\,|\,\theta_{-i},\mathbf{Y})$ Gibbs sampling เป็นกระบวนการวนซ้ำมีขั้นตอนดังนี้

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

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

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

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

. .

. .

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

ลำดับของการสุ่ม $\boldsymbol{\theta}_1^{(1)},\boldsymbol{\theta}_2^{(2)}...,\boldsymbol{\theta}^{(T)}$ เป็นสถานะต่อเนื่องกันของ Markov Chain

การวัดความถูกต้องแม่นยำของการจำลองสถานการณ์ (Najafi and Tarazkar,2006)

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

1) Mean Squared Error (MSE)

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

วิธี MSE เป็นวิธีที่ใช้กันทั่วไป ข้อเสียของวิธีนี้คือไม่มีฐานการเปรียบเทียบ และถ้า MSE มีค่าสูงอาจเป็นเพราะมีความคลาดเคลื่อนสูง หรือขึ้นอยู่กับขนาดของข้อมูล

2) Mean Absolute Percentage Error (MAPE)

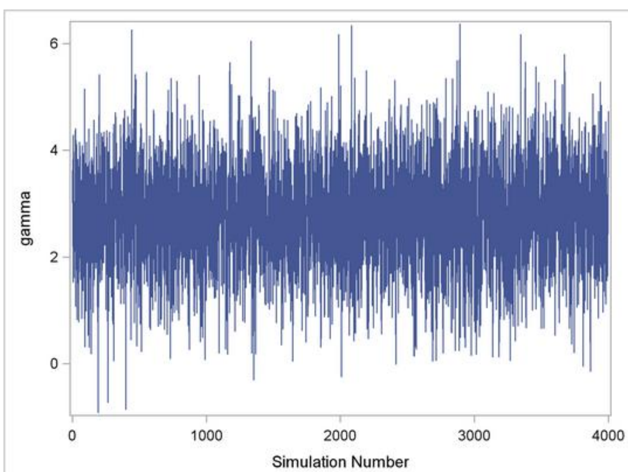
$$MAPE = \frac{\sum_{t=1}^{n} \left| e_t / Y_t \right|}{n} \times 100 \tag{7}$$

วิธี MAPE เป็นหนึ่งในวิธีที่ถูกยอมรับ และที่ใช้ในการเปรียบเทียบมากที่สุดสำหรับอนุกรมเวลา

3) Mean Absolute Error (MAE)

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

เมื่อค่า MSE (Mean Squared Error) MAPE (Mean Absolute Percentage Error) และMAE (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 โดยที่ มี T เซต โดยที่ $T = T_1,...,T_B$ และ S มีจำนวนที่ใหญ่พอ สามารถแสดงสูตรของตัวประเมินแต่ละตัวดังสมการข้างล่าง

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

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

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

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

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

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

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

1.7 แบบจำลองต่างๆ

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

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_k X_{kt} + \varepsilon_t \tag{16}$$

เมื่อ t แทนเวลา โดยที่ $t = 1,...,n$

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

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

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

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

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

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

เมื่อ N = ขนาดของ moving average โดยที่ $N = 1, ..., n$

ตัวอย่างเช่น

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

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

เมื่อ

$$N = 1 \rightarrow F_2 = y_1, F_3 = y_{2,...}$$

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

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

$F_{t+m} = a_t + b_t m_t$ โดยที่ m = 1,2,3,… (19)

เมื่อ

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

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

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

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

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

Exponential smoothing เป็นการพยากรณ์โดยกำหนดน้ำหนักให้ค่าสังเกตในปัจจุบันมีค่ามากกว่าน้ำหนักของค่าสังเกตก่อนหน้านั้น มี 3 แบบคือ

a. Single exponential smoothing

ตัวแบบการพยากรณ์แบบ Single exponential smoothing นิยามดังนี้

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

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

ค่าพยากรณ์ S_1 ไม่เกิดขึ้นเนื่องจากไม่มีเทอม S_0 นอกเสียจากจะกำหนดค่าเริ่มต้นให้

S_0 เทคนิคนี้ใช้สำหรับข้อมูลที่ ไม่มี 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

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

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

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

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

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

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

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 ไม่ใช่เพิ่มขึ้นหรือลดลงเป็นค่าคงทที่

2. Additive Seasonal Model

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

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

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

เมื่อ γ คือค่าคงที่แสดง seasonal และ s คือความยาวช่วงของ seasonal ใช้ตัวแบบการพยากรณ์นี้เมื่อข้อมูลมี seasonal เป็นแบบการบวก (additive seasonality) คือในเวลาเดียวกันของแต่ละฤดูกาล ค่าสังเกตจะเพิ่มขึ้นหรือลดลงเป็นค่าคงที 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 และซับซ้อนน้อยกว่าแบบมหาภาคที่อยู่ในลักษณะระบบสมการหลายชั้น 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 $y_t,\ldots,y_{t-p}$ หรือ ค่าสังเกตที่เกิดขึ้นก่อนหน้า p โดยกระบวนการหรือระบบ AR(p) คือกระบวนการหรือระบบ Auto Regressive ที่มีอันดับที่ p ซึ่งเขียนอยู่ในรูปสมการได้ดังนี้

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

โดยที่

μ คือ ค่าคงที่ (Constant Term)

ϕ_j คือ พารามิเตอร์ตัวที่ j

ε_t คือ ความคลาดเคลื่อน ณ เวลา t

ในกรณี ของ AR(1) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t \tag{30}$$

หรือ

$$y_t - \phi_1 y_{t-1} = \mu + \varepsilon_t \tag{31}$$

หรือ

$$(1-\phi_1\beta)x_t = \mu + \varepsilon_t \tag{32}$$

เมื่อ β คือ backward shift operation

และในกรณี ของ AR(2) สามารถเขียนรูปแบบสมการได้ดังนี้

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

หรือ

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

หรือ

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

2. แบบจำลอง Moving Average (MA(q))

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

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

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

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

$$\Delta_t y_t = \delta + \phi y_{t-1} + \phi y_{t-2} + \ldots + \phi y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} \tag{41}$$

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 ปัจจุบันกับข้อมูลถอยหลังไป d คาบเวลา โดยสาเหตุที่ต้องทำการหาผลต่างของอนุกรมเวลา เนื่องจากแบบจำลอง ARIMA ต้องใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาทีมีคุณสมบัติคงที่ (Stationary) เท่านั้น โดยในกรณีข้อมูลอนุกรรมเวลาที่ใช้ในการวิเคราะห์มีคุณสมบัติไม่คงที่ (Nonstationary) จะต้องทำการแปลงข้อมูลดังกล่าวให้เป็นข้อมูลที่มีคุณสมบัติคงที่ก่อน โดยการหาผลต่างของข้อมูลอนุกรมเวลาก่อนที่นำไปสร้างแบบจำลอง ARIMA ซึ่งโดยทั่วไปแล้วถ้าต้องหาผลต่างอันดับที่ d สามารถเขียนในรูปของ I(d) ได้ดังนี้

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

ในกรณี I (1) สามารถเขียนรูปแบบได้ดังนี้

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

ในกรณี I (2) สามารถเขียนรูปแบบได้ดังนี้

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

$$\Delta_t y_t = \delta + \phi\Delta_d y_{t-1} + \phi\Delta_d y_{t-2} + \ldots + \phi\Delta_d y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \ldots - \theta_q\varepsilon_{t-q} \tag{43}$$

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

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

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

ε_t คือ กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t

ภายใต้ข้อสมมติที่ว่าความคลาดเคลื่อนที่คนละเวลาเป็นตัวแปรสุ่มที่เป็นอิสระต่อกันโดยมีการแจง แจงปกติที่มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนคงที่

ดังนั้นเพื่อให้ง่ายขึ้นสำหรับการหาค่าพยากรณ์ Y_t ของ ARIMA สามารถหาได้จากนิยามที่สรุปไว้ต่อไปนี้

$$(1-\varnothing_1\beta-\varnothing_2\beta^2-...-\varnothing_p\beta^p)(1-\beta)^d Y_t=(1-\omega_1\beta-\omega_2\beta^2-...-\omega_q\beta^q)\varepsilon_t \tag{44}$$

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

$$\beta\varepsilon_t=\varepsilon_{t-1}\,,\ \beta^2\varepsilon_t=\varepsilon_{t-2}\,,\ \beta^3\varepsilon_t=\varepsilon_{t-3}\,,\ \ldots \tag{46}$$

ตัวอย่างเช่น

ARIMA(0,1,0) คือ

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

ARIMA(2,1,2) คือ

$$\begin{aligned}
&(1-\varnothing_1\beta-\varnothing_2\beta^2)(1-\beta)Y_t=(1-\omega_1\beta-\omega_2\beta^2)\varepsilon_t\\
&(1-\varnothing_1\beta-\varnothing_2\beta^2)(Y_t-\beta Y_t)=\varepsilon_t-\omega_1\beta\varepsilon_t-\omega_2\beta^2\varepsilon_t\\
&(Y_t-\varnothing_1\beta Y_t-\varnothing_2\beta^2 Y_t)-(\beta Y_t-\varnothing_1\beta^2 Y_t-\varnothing_2\beta^3 Y_t)=\varepsilon_t-\omega_1\varepsilon_{t-1}-\omega_2\varepsilon_{t-2}\\
&Y_t-\varnothing_1\beta Y_t-\varnothing_2\beta^2 Y_t-\beta Y_t+\varnothing_1\beta^2 Y_t+\varnothing_2\beta^3 Y_t=\varepsilon_t-\omega_1\varepsilon_{t-1}-\omega_2\varepsilon_{t-2}\\
&Y_t-\varnothing_1 Y_{t-1}-\varnothing_2 Y_{t-2}-Y_{t-1}+\varnothing_1 Y_{t-2}+\varnothing_2 Y_{t-3}=\varepsilon_t-\omega_1\varepsilon_{t-1}-\omega_2\varepsilon\\
&Y_t-\varnothing_1 Y_{t-1}-Y_{t-1}-\varnothing_2 Y_{t-2}+\varnothing_1 Y_{t-2}+\varnothing_2 Y_{t-3}=\varepsilon_t-\omega_1\varepsilon_{t-1}-\omega_2\varepsilon\\
&Y_t-(\varnothing_1-1)Y_{t-1}-(\varnothing_2-\varnothing_1)Y_{t-2}+\varnothing_2 Y_{t-3}=\varepsilon_t-\omega_1\varepsilon_{t-1}-\omega_2\varepsilon
\end{aligned}$$

$$Y_t = (\varnothing_1 - 1)Y_{t-1} + (\varnothing_2 - \varnothing_1)Y_{t-2} - \varnothing_2 Y_{t-3} + \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon \tag{48}$$

ตัวแบบ ARIMA ใช้วิเคราะห์ข้อมูลได้ทั้งแบบ stationary และ nonstationary ข้อมูล stationary คือข้อมูลที่ mean และ variance ไม่เปลี่ยนแปลงไปตามเวลา ข้อมูล nonstationary คือข้อมูลที่ mean หรือ variance เปลี่ยนแปลงไปตามเวลา ค่า variance ทำให้ stationary ได้โดยการแปลงข้อมูลด้วยฟังก์ชันลอการิทึม (Log transform) ซึ่งต้องทำก่อนใช้ตัวแบบ Aสำหรับค่า mean ทำให้ stationary ได้โดยการทำผลต่างข้อมูล (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

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

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

โดยความคลาดเคลื่อนมาตรฐานของ r_k (Standard Error of r_k) ซึ่งมีสูตรในการคำนวณดังนี้

$$se_{r_k} = \frac{1}{\sqrt{n}} \tag{50}$$

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

ในการศึกษาจะใช้สหสัมพันธ์ในตัวเองเป็นเครื่องมือที่สำคัญสำหรับสืบค้นคุณสมบัติของข้อมูลอนุกรมเวลาเชิงประจักษ์ โดยมี 2 วิธีสำหรับทดสอบว่าค่า rk 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 คือค่าล่าหรือค่าล้าหลังสูงสุด (Maximum Lag) ที่พิจารณา

6. Partial Autocorrelation Function (PACF)

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

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

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

โดยที่ $$\hat{x}_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + ... + \beta_k x_{t-k+1} \tag{52}$$

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 หรือSeasonal ARIMA แทนด้วย ARIMA $(p,d,q)(\mathrm{P,D,Q})$ และถ้ามีตัวแปรร่วม (covariate) X_{1t} และ X_{2t} สามารถนิยามได้ดังต่อไปนี้

$$\varnothing(\beta)\left[\Delta\left(Y_t - c_1 X_{1t} - c_2 X_{2t}\right) - \mu\right] = \Theta(\beta)\varepsilon_t \tag{53}$$

เมื่อ c_1 และ c_2 คือสัมประสิทธิ์การถดถอย

$\varnothing(\beta) = \phi_p(\beta)\varnothing_{\mathbf{P}}(\beta^s)$ และ $\Theta(\beta) = \theta_q(\beta)\Theta_Q(\beta^s)$ โดยที่

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

$\varnothing_{\mathbf{P}}(\beta^s) = 1 - \varnothing_1\beta^s - ... - \varnothing_P\beta^{sP}$, และ $\Theta_Q(\beta^s) = 1 - \Theta_1\beta^s - ... - \Theta_Q\beta^{sQ}$. Δ คือ differencing operator, $\Delta = (1-\beta)^d(1-\beta^s)^D$. β คือ backshift operator และ $\beta Y_t = Y_{t-1}$, $\beta^2 Y_t = Y_{t-2}$ เป็น

แบบนี้ต่อๆ ไป s คือ seasonal lag และ ε_t คือ ความผิดพลาด (error)ที่มีการแจกแจงแบบปกติมีค่าเฉลี่ยเป็น 0และความแปรปรวน σ^2. $\varnothing$'s และ ϕ's คือพารามิเตอร์autoregressiveแบบ seasonal และ non-seasonal ตามลำดับ Θ's และ θ's คือพารามิเตอร์moving average แบบ seasonal และ non-seasonal ตามลำดับ p และ q คือลำดับ (order) ของพารามิเตอร์autoregressive แบบ non-seasonal และพารามิเตอร์moving 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 differencesตามลำดับ

6) แบบจำลองหลายตัวแปร (Multivariate Variable Model)

Multivariate AR(1)

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

Expanded Form

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

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

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

ข้าวนับเป็นพืชเศรษฐกิจของประเทศไทย ซึ่งเกี่ยวข้องกับชีวิตความเป็นอยู่ของชาวนา หลายสินล้านคน แต่ละปีเราเป็นผู้ส่งออกข้าว นำรายได้เข้าประเทศ ร่วมสองแสนล้านบาท โดยไทยเป็นผู้ส่งออกรายใหญ่ที่สุดของโลก นอกจากนี้ยังมีผู้ส่งออกข้าวรายสำคัญ รองลงมา ได้แก่ เวียดนาม อินเดีย และสหรัฐ เป็นต้น ในอดีตข้าวไทยเคยเป็นสินค้าที่นำรายได้ให้ประเทศเป็นอันดับต้น ๆ ต่อมาเนื่องจากเราหันมาผลิตสินค้าอุตสาหกรรมมากขึ้นได้แก่พวกสิ่งทอเครื่องนุ่งห่ม เครื่องใช้ไฟฟ้า อุปกรณ์อีเล็คโทรนิกส์ ยานยนต์และชิ้นส่วน รวมทั้งอุปกรณ์คอมพิวเตอร์ เพื่อรองรับความต้องการในประเทศ และภายหลังก็ได้มุ่งผลิตสินค้าเหล่านี้เพื่อการส่งออกอีกด้วย สินค้าอุตสาหรรมเหล่านี้จึงเป็นแหล่งรายได้ส่งออกสำคัญ แซงหน้าการส่งออกข้าว และพืชผลการเกษตรอื่น ข้าวถือเป็นสินค้าประเภท Commodity (โภคภัณฑ์) ที่ระดับราคาในตลาดโลกเป็นไปตามกลไกตลาด อย่างแท้จริง นั่นคือราคาตลาดโลกจะ

ปรับขึ้นหรือลงย่อมขึ้นอยู่กับอุปสงค์และอุปทาน เมื่อใดที่อุปทานมากกว่าอุปสงค์ราคาในตลาดโลกก็จะปรับลดลง และถ้าปีใดอุปสงค์มากกว่าอุปทานราคาก็จะสูงขึ้น

โดยทั่วไปตลาดข้าวในประเทศไทยจะมีความเชื่อมโยงกับตลาดโลกอย่างใกล้ชิด เมื่อใดที่สถานการณ์ตลาดโลกทำให้ราคาตลาดโลกเพิ่มขึ้น ราคาข้าวในประเทศก็มีแนวโน้มปรับสูงตามไปด้วย เมื่อใดที่ราคาข้าวในประเทศไทยตกต่ำ โดยมากจะเกิดในช่วงที่มีข้าวออกมามากพร้อมๆ กัน รัฐบาลก็จะใช้ โครงการรับจำนำข้าเปลือก เพื่อพยุงราคาข้าวไว้โดยรัฐบาลจะรับซื้อข้าวเปลือกในราคาที่ประกาศ (เรียกว่าราคาประกัน) เพื่อดูดอุปทานข้าวออกจากตลาดซึ่งราคาประกันแต่ละปีก็จะอยู่ในระดับที่สูงกว่าราคาตลาด สำหรับการรับจำนำนี้ทำได้ 2 รูปแบบคือจำนำที่ยุ้งฉางเกษตรกร หรือจำนำที่โรงสีที่เข้าร่วมโครงการฯ แบบแรกนั้น ธกส. จะเป็นผู้ดำเนินการโดยข้าวก็ฝากเกษตรกรไว้ ถ้าเป็นการรับจำนำที่โรงสีเกษตรกรก็ต้องขนข้าวมาที่โรงสี และก็จะได้รับเอกสารไปขึ้นเงินกับ ธกส. โรงสีก็มีหน้าที่สีข้าวส่งมอบแก่องค์การคลังสินค้า (อคส.) แต่ในความเป็นจริงข้าวเหล่านี้ก็ยังคงถูกเก็บไว้ที่โรงสี เพื่อรอเวลาจะระบายออกนอกประเทศ หรือในบางกรณี ก็มีการนำกลับมาใช้ในตลาดในประเทศก็ได้แล้วแต่นโยบาย ในอดีตที่ผ่านมาข้าวที่รัฐบาลเป็นเจ้าของผ่านโครงการรับจำนำ มักจะถูกเก็บไว้จนเสื่อมสภาพ หรือไม่ก็มีการล่องหนหายไปจากโกดัง ทำให้ไม่สามารถขายได้หรือถ้าขายได้ก็ได้ราคาต่ำมาก ดังนั้นโครงการรับจำนำข้าวจึงมักเกิดผลขาดทุนเสมอ การขาดทุนก็เนื่องจากซื้อแพงแต่ขายถูกหรือขายไม่ได้เลยก็มี

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

เป็นประเทศเดียวที่มีผลผลิตมากและสามารถส่งออกได้ตามปกติ ทุกประเทศจึงหันมาส่งซื้อข้าวจากไทย เรียกได้ว่ามีเท่าไรก็ขายหมดแน่

สำหรับปีนี้คาดว่าประเทศไทยจะมีผลผลิตข้าวออกมารวมประมาณ 30 ล้านตัน เพิ่มขึ้นจากปีก่อนเล็กน้อย และจะมีปริมาณส่งออกข้าวราว ๆ 9 ล้านตัน โดยผลผลิตข้าวนาปีได้เริ่มออกสู่ตลาดแล้วตั้งแต่ปลายปีก่อนถึงต้นปี 2551 และข้าวนาปรังจะทยอยเข้าตลาดตั้งแต่ช่วงนี้จนถึงเดือนสิงหาคม และตั้งแต่ปลายปีที่แล้วจนถึงเดือนกุมภาพันธ์นี้ ไทยได้ส่งออกข้าวไปแล้วเฉลี่ยเดือนละหนึ่งล้านตัน เรื่องที่น่าดีใจก็คือระดับราคาข้าวส่งออกทุกชนิดได้ปรับสูงขึ้นอย่างมาก เช่น ข้าวหอมมะลิ (100% ชั้น 2) ราคาเพิ่มขึ้น 20% ข้าวขาว (100% ชั้น 2) ข้าวนึ่ง 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

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

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

งานวิจัยที่เกี่ยวข้อง

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

ตัวแบบการพยากรณ์แบบดั้งเดิม (Classical Models) ในการพยากรณ์ผลลัพธ์ที่จะเกิดขึ้นในอนาคตของเหตุการณ์ด้านต่างๆ นั้น Wright (1986) นำตัวแบบ Simple Exponential และ Holt ไปใช้กับข้อมูลอนุกรมเวลาที่ไม่สม่ำเสมอ (Irregular) Deetae (1991) ประยุกต์ใช้ตัวแบบ Box Jenkins (ARIMA) ในการพยากรณ์ราคาข้าว พบว่าตัวแบบนี้มีประสิทธิภาพอย่างเห็นได้ชัด และดีกว่าตัวแบบ Decomposition เช่นเดียวกันกับที่ Kerdsomboon (1999) พบว่า ตัวแบบ Box Jenkins พยากรณ์ราคาข้าวได้ดีกว่าตัวแบบสถิติเบื้องต้น Ciprаและคณะ (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 และอื่นๆในขณะที่ Sangpattaranate (2005) ที่พบว่า Box Jenkins เป็นตัวแบบที่ใช้ในการพยากรณ์ราคาข้าวได้ดีกว่าตัวแบบ Holt-Winter และตัวแบบการถดถอย

Iqbal และคณะ (2005) ใช้ตัวแบบ ARIMA พยากรณ์ผลผลิตและพื้นที่เพราะปลูกข้าวสารีในประเทศปากีสถานเพื่อใช้เป็นข้อมูลให้กับรัฐบาลในการกำหนดนโยบาย Mishra และ Desai (2005) ใช้ตัวแบบ SARIMA ในการพยากรณ์ภัยแล้ง (Drought) โดยใช้ข้อมูลอนุกรมเวลาเป็นค่าดัชนีมาตรฐานของหยาดน้ำที่ตกมาจากชั้นบรรยากาศ (Precipitation)สำหรับตัวแบบการพยากรณ์แบบ exponential smoothing นั้นมีการนำไปใช้อย่างแพร่หลายนั้น และมีการยืดขยายเป็น 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 ตัวแบการถดถอย (Regression Model)ที่มีฤดูกาล (Seasonal) เป็นตัวแปรซ่อนเร้น (Latent variable) ในการพยากรณ์ ปริมาณความต้องการกระแสไฟฟ้าพบว่าตัวแบการถดถอยที่มีฤดูกาลเป็นตัวแปรซ่อนเร้นพยากรณ์ได้แม่นยำกว่าARIMA และSARIMA

Kahforoushan, Zarif and Mashahir (2010) ศึกษาการพยากรณ์ผลผลิตทางด้านการเกษตร ซึ่งได้แก่ การปลูกพืช การเลี้ยงสัตว์ การประมง และการปลูกป่า โดยใช้วิธีการพยากรณ์ 4 วิธี ได้แก่ วิธีปรับเรียบเอ็กโปเนนเชียลโดยวิธี Holt-Winters แบบไม่มีฤดูกาล (Holt-Winters (no seasonal) Exponential Smoothing Model) วิธีบอกซ์-เจนกินส์ (Box-Jenkins Model) วิธีโครงข่ายปราสาทเทียม (Artificial Neural Network Model) และ วิธีARIMA (ARIMA Model) และใช้ ค่า MAE MSE และ MAPE เปรียบเทียบผลการพยากรณ์แต่ละวิธี ผลการศึกษาพบว่า วิธีโครงข่ายปราสาทเทียมเหมาะสมในการประมาณค่าพารามิเตอร์ในตัวแบบ (Learn Stage) วิธีบอกซ์-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 ที่มีการแจกแจงแบบปัวซง 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 ตัวแบบที่ถูกนำเสนอเป็นการอธิบายความสัมพันธ์ระหว่างค่ารวมทั้งหมดกับค่ารวมมาแล้วบางส่วนของตัวแปรภายใต้อิทธิพลของฤดูกาลแบบคงที่ (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) นำเสนอตัวแบบเบย์ให้มีความความเหมาะสมกับข้อมูลปริมาณความต้องการซื้อชิ้นส่วนอุปกรณ์คอมพิวเตอร์ และเมื่อเปรียบเทียบกับตัวแบบมาตรฐานอื่นๆ ได้แก่ Exponential smoothing (Exps) และ Judgmental Methods (Judg) พบว่าวิธีเบย์ มีความเหมาะสมมากกว่า จากที่กล่าวมาทั้งหมดนั้น Tongkhow และ Kantanantha (2011) ได้ศึกษาตัวแบบพยากรณ์ที่นิยมใช้และมีประสิทธิภาพในปัจจุบันต่างๆ โดยนำไปประยุกต์ใช้กับข้อมูลราคาผักที่มีความผันผวนในประเทศไทยและได้สร้างตัวแบบเบย์อย่างๆง่ายๆ ตามหลักการของสมการการถดถอยเชิงพหุ ผลการศึกษาพบว่าวิธีการของ Exps มีความเหมาะสมมากกว่า และ SesonalARIMA มีความเหมาะสมบางค่า หลัง จากนั้น Tongkhow และ Kantanantha (2012, 2013) ขยายตัวแบบของ Yelland (2010) โดยนำไปประยุกต์ใช้กับข้อมูลราคาผักที่มีความผันผวนในประเทศไทย และนำมาเปรียบเทียบกับตัวแบบของ Yelland (2010) ARIMA SesonalARIMA และ Exps พบว่าวิธีที่นำเสนอมีความเหมาะสมมากกว่า จากที่กล่าวมาทั้งหมด