

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

1) การวิเคราะห์แนวโน้ม (Time Series or Trend Method) เป็นวิธีที่ง่ายที่สุด ด้วยวิธีการแสดงความสัมพันธ์ของการใช้ไฟฟ้ากับค่าของเวลาช่วงนั้นๆ โดยอาศัยหลักการที่ว่า การใช้ไฟฟ้าจะเพิ่มขึ้นตลอดเวลาที่ผ่านมา เพราะฉะนั้น การวิเคราะห์แนวโน้มของการใช้ไฟฟ้าที่ผ่านมาในอดีตก็สมมติว่าแนวโน้มจะเป็นไปเช่นเดียวกับในอนาคตวิธีนี้จะได้ผลในกรณีที่ข้อมูลในอดีตระยะเวลาพอสมควร และสภาวะทางเศรษฐกิจของประเทศค่อนข้างมั่นคง และมีเสถียรภาพ ทั้งในอดีตและอนาคตข้างหน้า

หาสมการที่เหมาะสมที่สุดสำหรับการพยากรณ์โดยวิธี Time Series Approach จากสมการต่อไปนี้

$$\begin{aligned} Y &= a_0 + a_1X \\ Y &= a_0 + a_1X_1 + a_2X_2 \\ Y &= ab^x \\ \text{Lny} &= k + ab^x \\ 1/y &= k + a\text{Ln}^x \\ \text{Lny} &= a_0 + a_1\text{Ln}^x \end{aligned}$$

2) แบบจำลองทางเศรษฐศาสตร์ (Econometric Model) อาศัยหลักความจริงที่ว่าความต้องการใช้ไฟฟ้ามีความสัมพันธ์อย่างใกล้ชิดกับความเจริญเติบโตทางเศรษฐกิจ ดังนั้นความต้องการใช้ไฟฟ้าจึงขึ้นอยู่กับข้อมูลหรือตัวแปรทางเศรษฐกิจ (Economic Variables and/or Demographic Variables) เช่น ผลิตภัณฑ์มวลรวมประชาชาติ (Gross Domestic Product : GDP) รายได้ประชาชาติ (National Income) จำนวนประชากร เป็นต้น การพยากรณ์วิธีการนี้คือ การสร้างสมการแสดงความสัมพันธ์ระหว่างความต้องการใช้ไฟฟ้ากับข้อมูลทางเศรษฐกิจ จากสมการเมื่อทราบค่าของข้อมูลทางเศรษฐกิจ โดยอาศัยข้อมูลจากแผนพัฒนาเศรษฐกิจแห่งชาติซึ่งในแผนงานนี้ใช้การวิเคราะห์ความถดถอยแบบพหุคูณ (Multiple Regression Analysis)

การพยากรณ์วิธี Regression Approach โดยมีสมการดังนี้

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2$$

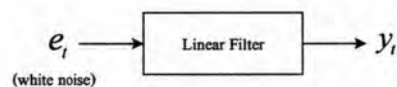
Y	=	ตัวแปรตาม(Dependent Variable)
X_1	=	ตัวแปรอิสระตัวที่ 1 (Independent Variable of 1)
X_2	=	ตัวแปรอิสระตัวที่ 2 (Independent Variable of 2)
β_0	=	ค่าคงที่(Constant)
β_1	=	ค่าสัมประสิทธิ์ของ X_1 (Coefficient of X_1)
β_2	=	ค่าสัมประสิทธิ์ของ X_2 (Coefficient of X_2)

3) ทฤษฎีการผลิต การผลิต หมายถึง กระบวนการเปลี่ยนแปลงปัจจัยการผลิต(Input) ที่ใส่เข้าไปในกระบวนการผลิตออกมาเป็นผลผลิต(Output) ในการพิจารณากระบวนการผลิต สิ่งที่น่าสนใจที่จะศึกษาก็คือ ลักษณะความสัมพันธ์ระหว่างปัจจัยการผลิตที่ใส่เข้าไปในกระบวนการผลิต และผลผลิตที่รับออกมา ความสัมพันธ์เช่นนี้เรียกว่าฟังก์ชันการผลิต(Production Function) เพื่อที่จะหาว่าความต้องการไฟฟ้าจะมากหรือน้อยขึ้นอยู่กับอะไร สามารถหาได้จากฟังก์ชันการผลิต ซึ่งการผลิตต้องใช้ปัจจัยต่างๆ

4)แบบจำลอง $ARMA(p,q)$ และเทคนิคการพยากรณ์ Box-Jenkins Method

(Auto Regressive Moving Average Models and Box-Jenkins Forecasting Techniques)

พิจารณา $y(t)$, Stochastic Time Series ซึ่งคือผลตอบสนองของวงจรกรองเชิงเส้น (Linear Filter) ที่มี white noise, $e(t)$ เป็นสัญญาณเข้า โดยสัญญาณเข้ามีการกระจายแบบ Gaussian Distributions



$ARMA(p,q)$, Auto-Regressive Moving Average : เป็น process ซึ่งสัญญาณขาออกเกิดจากผลรวมเชิงเส้นในเทอมของค่าสังเกตที่ y_{t-1}, \dots, y_{t-p} และในเทอมของ White Noise e_{t-1}, \dots, e_{t-q} โดยที่ p, q คืออันดับของ AR และ MA ตามลำดับ โดยมีความสัมพันธ์ระหว่างสัญญาณเข้ากับสัญญาณออกคือ

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

สำหรับ Box-Genkins Method คือกระบวนการในการพยากรณ์ค่า y_t บนพื้นฐานของกระบวนการ *ARMA*

Stationary Random Sequences (อนุกรมเวลาชนิดสแตชันนารี) และวิธีของ Box และ Jenkins

วิธีการของ Box และ Jenkins เป็นวิธีการพยากรณ์ค่าในอนาคตที่พัฒนาโดย George E.P. Box และ Gwilym M. Jenkins วิธีการนี้ให้ค่าพยากรณ์ที่มีค่าความถูกต้องสูงแต่มีขั้นตอนที่ยุ่งยากในการวิเคราะห์เพราะต้องใช้ความรู้เรื่องตัวแบบ *ARIMA(p, d, q)* รูปแบบที่จะกำหนดให้กับอนุกรมเวลาจะเป็นรูปแบบในกลุ่มของรูปแบบ *ARMA(p, q)* (Auto-Regressive and Moving Average order p and q) models ซึ่งเป็นการรวมส่วนของรูปแบบ *AR(p)* และรูปแบบ *MA(q)* เข้าด้วยกัน รูปแบบ *AR(p)* หมายถึงรูปแบบที่แสดงว่าค่าสังเกต Y_t จะขึ้นอยู่กับค่าของ Y_{t-1}, \dots, Y_{t-p} หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ *MA(q)* หมายถึงรูปแบบที่แสดงว่าค่าสังเกต Y_t จะขึ้นอยู่กับค่าของความคลาดเคลื่อน $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ หรือความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า เช่น รูปแบบ

$$AR(1) ; Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

$$MA(1) ; Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

$$ARMA(1,1) ; Y_t = \theta_0 + \phi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

การกำหนดรูปแบบ *ARMA(p, q)* ที่เหมาะสมให้กับอนุกรมเวลาจะได้ออกจากการพิจารณาว่าอนุกรมเวลามีค่าลักษณะบางค่าได้แก่ Auto-Correlation Function (ค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติ) และ Partial Auto-Correlation Function (ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติ) ของตัวอย่าง (r_k, r_{kk}) ซึ่ง k หมายถึง คาบเวลาห่างระหว่างอนุกรม และเรียกคาบเวลานี้ว่าแล็ก k (Lag k) สอดคล้องกับค่าลักษณะ ได้แก่ Auto-Correlation Function และ Partial Auto-Correlation Function (ρ_k, ρ_{kk}) ของรูปแบบ *ARMA(p, q)* ไค ขนาดของอนุกรมเวลาจะต้องใหญ่พอสมควร นั่นคืออนุกรมเวลาควรมีขนาดเกิน 30 เพราะจะต้องหาค่า r_k และ r_{kk} หลายค่าเพื่อจะได้นำมาเปรียบเทียบ ρ_k และ ρ_{kk} วิธีการของ Box และ Jenkins แบ่งอนุกรมเวลาออกเป็น 2 ประเภท ดังนี้

1) Stationary Times Series : เป็นอนุกรมเวลา $\{Y_t\}$ ที่มีค่าเฉลี่ยและค่าความแปรปรวนของ Y_t คงที่ นั่นคือค่าเฉลี่ย $E(Y_t)$ และค่าความแปรปรวน $V(Y_t)$ มีค่าคงที่สำหรับแต่ละเวลา t ซึ่งอนุกรมเวลาที่มีแนวโน้ม และหรือฤดูกาล จะมี $E(Y_t)$ ไม่คงที่ ส่วนอนุกรมเวลาที่ความผันแปรของ Y_t สูง

จะเป็นลักษณะของอนุกรมเวลาที่ $V(Y_t)$ ไม่คงที่ จะเรียกว่า Non-Stationary Time Series นอกจากนี้จะเป็นอนุกรมเวลาที่มีค่าเฉลี่ยและค่าความแปรปรวนคงที่แล้วอนุกรมเวลานี้จะต้องมี Auto-Correlation Function ที่ lag k ขึ้นอยู่กับค่า k อย่างเดียว อนุกรมเวลาที่จะกำหนดรูปแบบ $ARMA(p, q)$ จะต้องเป็น Non-Stationary Time Series เท่านั้น

2) Non-Stationary Times Series : เป็นอนุกรมเวลาที่ไม่มีความสมบัติเป็น Stationary Times Series จะหารูปแบบ $ARMA(p, q)$ ให้กับอนุกรมเวลาดังกล่าวไม่ได้ จะต้องแปลงอนุกรมเวลานี้ให้เป็นอนุกรมเวลาใหม่ที่มีความสมบัติเป็น Stationary Times Series จึงจะสามารถหารูปแบบ $ARMA(p, q)$ ให้กับอนุกรมเวลาดังกล่าวได้ สำหรับรายละเอียดการแปลง (transform) Non-Stationary Times Series ให้เป็น Stationary Times Series จะอธิบายในตอนต่อไป ขั้นตอนการพยากรณ์โดยวิธีการของ Box และ Jenkins มี 4 ขั้นตอนดังนี้

ขั้นตอนที่ 1 Identification : เป็นการหารูปแบบ $ARMA(p, q)$ ที่คาดว่าจะเหมาะสมกับอนุกรมเวลาโดยพิจารณาจากการเปรียบเทียบค่า r_k และ r_{kk} ของอนุกรมเวลากับค่า ρ_k และ ρ_{kk} ของแต่ละรูปแบบ เนื่องจากต้องพิจารณา $r_k, r_{kk}, \rho_k, \rho_{kk}$ พร้อมกันหลาย ๆ ค่าจึงมักจะพิจารณาจาก Correlogram ที่ได้จากการแปรค่า lag k ต่าง ๆ นั่นคือการเปรียบเทียบจะเป็นการเปรียบเทียบ Correlogram ของ r_k กับ ρ_k และ Correlogram ของ r_{kk} กับ ρ_{kk} สำหรับแต่ละรูปแบบจะมี Correlogram ของ ρ_k และ ρ_{kk} ต่างกัน อนุกรมเวลาที่จะนำมาหารูปแบบจะต้องเป็น Stationary Times Series เท่านั้น

ขั้นตอนที่ 2 Estimation : จะทำได้โดยการหาค่าประมาณแบบง่ายหรือค่าประมาณที่ได้จากการวิเคราะห์เชิงตัวเลข (Numerical Method) สำหรับค่าประมาณแบบง่ายจะทำได้โดยการสร้างสมการที่มาจากความสัมพันธ์ระหว่าง ρ_k และพารามิเตอร์ โดยสมการที่สร้างขึ้นจะมีจำนวนเท่ากับจำนวนพารามิเตอร์ที่ต้องการประมาณ ส่วนค่าประมาณที่ได้จากการวิเคราะห์เชิงตัวเลขจะทำได้จากการแก้สมการที่สร้างขึ้นโดยวิธีกำลังสองน้อยที่สุด โดยทั่วไปจะใช้ค่าที่ได้จากการประมาณแบบง่ายนำมาเป็นค่าเริ่มต้นในการวิเคราะห์เชิงตัวเลข

ขั้นตอนที่ 3 Diagnostic Checking : เมื่อได้กำหนดรูปแบบและประมาณค่าพารามิเตอร์แล้วจะต้องตรวจสอบว่ารูปแบบที่กำหนดนั้นมีความเหมาะสมหรือไม่ ซึ่งตรวจสอบได้หลายวิธี เช่น การพิจารณา Correlogram ของ r_k ของค่า Residual เป็นต้น

ขั้นตอนที่ 4 Forecasting : จะทำได้ทั้งการพยากรณ์แบบ Point Forecasting และ Interval Forecasting โดยการพยากรณ์จะใช้สมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนดและผ่านการตรวจสอบในขั้นตอนที่ผ่านมาแล้ว

ARMA(p, q) Models

รูปแบบ ARMA(p, q) มี p เป็น order ของ AR และ q เป็น order ของ MA จำนวนพารามิเตอร์ในรูปแบบจะเท่ากับ $p + q + 1$ ในทางปฏิบัติมักจะไม่เกิน 3 พารามิเตอร์ ตารางที่ 1 แสดงรูปแบบ ARMA ต่าง ๆ และเงื่อนไขของพารามิเตอร์ที่ปรากฏในรูปแบบโดย Y_t เป็นค่าสังเกตที่เวลา t ในอนุกรมเวลา Y_t และ ε_t เป็นค่าความคลาดเคลื่อนของการพยากรณ์ที่เวลา t โดย ε_t มีคุณสมบัติเป็น Gaussian Distribution ซึ่ง $E\{Y_t\} = \mu = 0$ และ $E\{(Y_t - \mu)^2\} = \sigma^2$ สำหรับเงื่อนไขของพารามิเตอร์ในส่วนของ AR อันดับที่ p เป็นเงื่อนไขที่กำหนดให้รูปแบบ ARMA(p, q) เป็น Stationary ส่วนเงื่อนไขของพารามิเตอร์ในส่วนของ MA ที่อันดับ q จะเป็นเงื่อนไขที่กำหนดให้รูปแบบ MA(q) เป็น Invertible

ตารางที่ 2.1 รูปแบบและเงื่อนไขของ ARMA

ARMA(p, q)	รูปแบบ	เงื่อนไข
White Noise	$Y_t = \theta_0 + \varepsilon_t$	-
AR(1)	$Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t$	$-1 < \phi_1 < 1$
AR(2)	$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$	$\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$ $-1 < \phi_2 < 1$
MA(1)	$Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t$	$-1 < \theta_1 < 1$
MA(2)	$Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} + \varepsilon_t$	$\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$ $-1 < \phi_2 < 1$
ARMA(1,1)	$Y_t = \theta_0 + \phi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t$	$-1 < \phi_1 < 1$ $-1 < \theta_1 < 1$

ในทุกรูปแบบจะประกอบด้วยค่าคงที่ θ_0 และค่าความคลาดเคลื่อน ε_t สำหรับรูปแบบ $AR(p)$ แสดงว่า Y_t ขึ้นอยู่กับค่าสังเกตที่เกิดขึ้นก่อน p ค่า หรือ Y_t ขึ้นอยู่กับ Y_{t-1}, \dots, Y_{t-p} พารามิเตอร์ใน $AR(p)$ ประกอบด้วย ϕ_1, \dots, ϕ_p ส่วนรูปแบบ $MA(q)$ แสดงว่า Y_t ขึ้นอยู่กับค่าความคลาดเคลื่อนล่าสุด q ค่า หรือ Y_t ขึ้นอยู่กับ $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ พารามิเตอร์ใน $MA(q)$ ประกอบด้วย $\theta_1, \dots, \theta_q$

Stationary and Invertible

Stationary เป็นคุณสมบัติของรูปแบบ $AR(p)$ ซึ่งเป็นคุณสมบัติที่ทำให้ $E(Y_t)$ และ $V(Y_t)$ คงที่ และ $COV(Y_t, Y_{t-k})$ จะขึ้นกับ lag k อย่างเดียว การพิจารณาว่าค่าพารามิเตอร์ ϕ_1, \dots, ϕ_p ใดที่จะทำให้รูปแบบ AR เป็น Stationary ทำได้จาก

ขั้นตอนที่ 1 จากรูปแบบ

$$AR(p): Y_t = \theta_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

$$\text{หรือ } Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \theta_0 + \varepsilon_t$$

และสามารถใช้ Backward Operator จะได้เป็น

$$(1 - \phi_1 B - \dots - \phi_p B^p) Y_t = \theta_0 + \varepsilon_t$$

$$\text{หรือ } (1 - \sum_{i=1}^p \phi_i B^i) Y_t = \theta_0 + \varepsilon_t$$

ขั้นตอนที่ 2 หากำตอบของ

$$1 - \phi_1 B - \dots - \phi_p B^p = 1 - \sum_{i=1}^p \phi_i B^i = 0 \text{ จะได้ค่า}$$

ของ B จำนวน p ค่า จะเลือก B เพียงหนึ่งค่าที่อยู่นอก Unit Circle นั่นคือ $|B| > 1$ จึงจะเป็น Stationary

Invertible เป็นคุณสมบัติของรูปแบบ $MA(q)$ ซึ่งเป็นคุณสมบัติที่ทำให้หาค่าคลาดเคลื่อนของการพยากรณ์ (ε_t) ในเทอมของ Y_t, Y_{t-1}, \dots ได้ การพิจารณาว่าค่าพารามิเตอร์ $\theta_1, \dots, \theta_q$ ใดที่จะทำให้อรูปแบบ $MA(q)$ เป็น Invertible จะทำได้โดย

1) จากรูปแบบ $MA(q)$

$$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

หรือ
$$Y_t = \theta_0 + (1 - \theta_1 B^1 - \dots - \theta_q B^q) \varepsilon_t$$

2) หาคำตอบของสมการ

$$1 - \theta_1 B^1 - \dots - \theta_q B^q = 0$$

จะได้ค่าของ B จำนวน q ค่า โดยเงื่อนไขที่เป็น Invertible คือ $|B| > 1$

ตารางที่ 2.2 ตารางคุณสมบัติของ ARMA

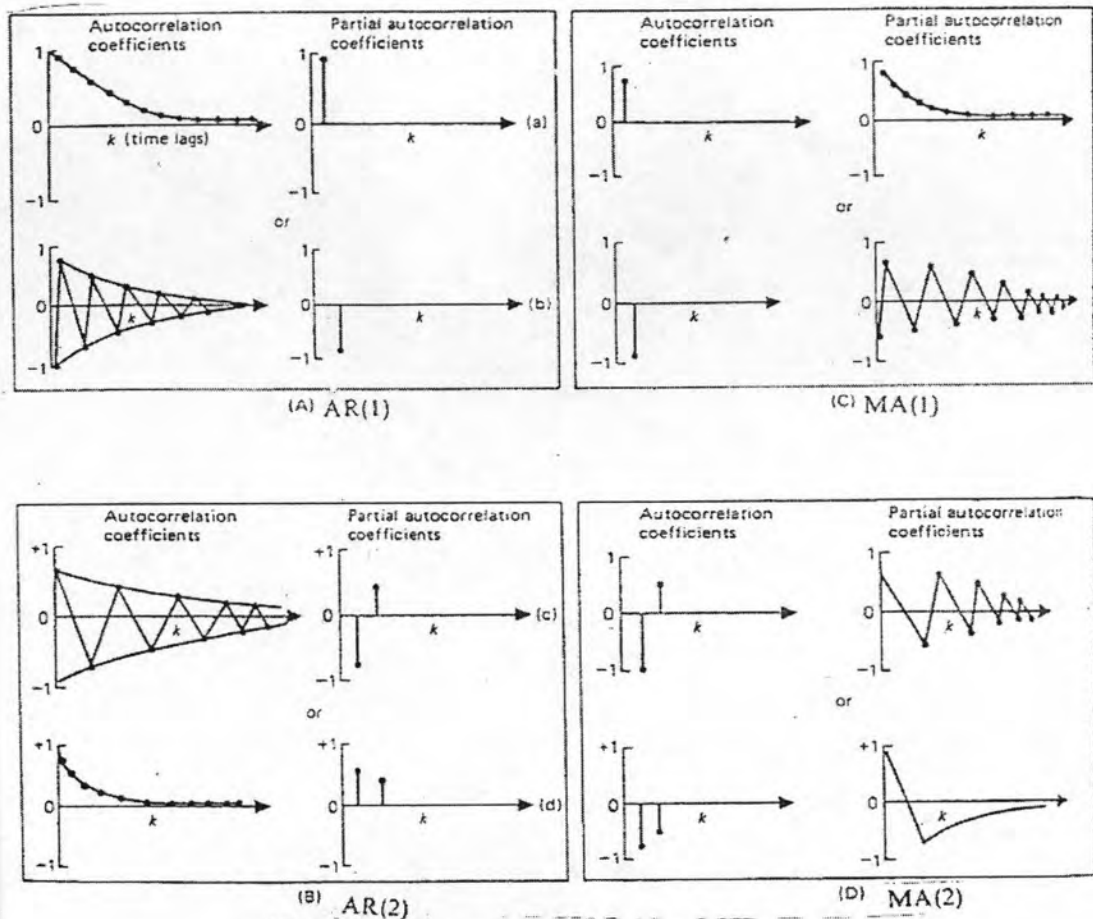
mode ls	μ_y	σ_y^2	ρ_k	ρ_{kk}
*	θ_0	σ^2	$\rho_k = 0, k \geq 1$	$\rho_{kk} = 0, k \geq 1$
AR(1)	$\frac{\theta_0}{1 - \phi_1}$	$\frac{\sigma^2}{1 - \phi_1^2}$	$\rho_k = \phi_1^k, k \geq 1$	$\rho_{11} = \phi_1,$ $\rho_{kk} = 0, k > 1$
AR(2)	$\frac{\theta_0}{1 - \phi_1 - \phi_2}$	$\frac{(1 - \phi_2)\sigma^2}{\{(1 + \phi_2)((1 - \phi_2)^2 - \phi_1^2)\}}$	$\rho_0 = 1, \rho_1 = \frac{\phi_1}{1 - \phi_2}$ $\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, k$	$\rho_{11} = \frac{\phi_1}{1 - \phi_2},$ $\rho_{22} = \phi_2,$ $\rho_{kk} = 0, k > 2$
MA(1)	θ_0	$(1 + \theta_1^2)\sigma^2$	$\rho_1 = -\frac{\theta_1}{1 + \theta_1^2}$ $\rho_k = 0, k > 1$	$\rho_{kk} = -\frac{\theta_1^k(1 - \theta_1^2)}{1 - \theta_1^{2k+2}}$
MA(2)	θ_0	$(1 + \theta_1^2 + \theta_2^2)\sigma^2$	$\rho_1 = -\frac{\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2}$ $\rho_2 = -\frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}$ $\rho_k = 0, k > 2$	ค่าลดลงรวดเร็ว

**	$\frac{\theta_0}{1-\phi_1}$	$\frac{(1-\phi_1^2-2\phi_1\theta_1)\sigma^2}{1-\phi_1^2}$	$\rho_1 = \frac{(1-\theta_1\phi_1)(\phi_1-\theta_1)}{1-\phi_1^2-2\phi_1\theta_1}$ $\rho_k = \phi_1^{k-1}\rho_1, k > 1$	ค่าลดลงรวดเร็ว
----	-----------------------------	---	---	----------------

* White Noise ** ARMA(1,1)

คุณสมบัติของ $ARMA(p,q)$ models

คุณสมบัติของ $ARMA(p,q)$ models ที่สำคัญได้แก่ $E(Y_t)$, $V(Y_t)$, Auto-Correlation Function ที่ lag k (ρ_k), Partial Auto-Correlation Function ที่ lag k (ρ_{kk}) แต่ละรูปแบบจะมีค่าวัดลักษณะต่างกัน ตารางที่ 2 แสดง $E(Y_t)$, $V(Y_t)$, ρ_k , ρ_{kk} ของ models white noise, $AR(1)$, $AR(2)$, $MA(1)$, $MA(2)$, $ARMA(1,1)$ จะเห็นว่าทั้ง $E(Y_t)$ และ $V(Y_t)$ ต่างก็เป็นฟังก์ชันของพารามิเตอร์ในรูปแบบ ซึ่งมีค่าไม่ต่างกันสำหรับทุกค่าของ t สำหรับ ρ_k , ρ_{kk} นั้นจะเป็นฟังก์ชันของพารามิเตอร์ในรูปแบบ และค่า k สำหรับค่า Auto-Correlation Function และ Partial Auto-Correlation Function ของ $AR(1)$, $AR(2)$, $MA(1)$, $MA(2)$ เป็นไปตามรูปที่ 1 และ $ARMA(1,1)$ เป็นไปตามรูปที่ 2.1



AUTOCORRELATION AND PARTIAL AUTOCORRELATION FUNCTIONS: (A) AR(1) MODELS, (B) AR(2) MODELS, (C) MA(1) MODELS, AND (D) MA(2) MODELS

รูปที่ 2.1 รูปแบบ ACF กับ PACF ของ AR(p) และ MA(q)

Auto-Correlation & Partial Auto-Correlation Function

การกำหนดรูปแบบ $ARMA(p, q)$ ให้กับอนุกรมเวลาจะทำได้โดยการเปรียบเทียบ Corelogram ของ r_k, r_{kk} กับ ρ_k, ρ_{kk} ของรูปแบบ ดังนั้นจะต้องมีการหาค่า r_k, r_{kk} สำหรับ k หลายค่า r_k เป็นค่าประมาณของ ρ_k ซึ่ง

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n-k} (y_t - \bar{y})^2} \quad k = 0, 1, 2, \dots$$

$$V(r_k) = \frac{1}{n} (1 + 2 \sum_{j=1}^{k-1} r_j^2) \quad k = 1, 2, \dots$$

ความหมายของ r_k จะเป็นคล้ายกับสัมประสิทธิ์สหสัมพันธ์ นั่นคือเป็นค่าวัดสหสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน k ช่วงเวลา หรือวัดสหสัมพันธ์ระหว่าง Y_t กับ Y_{t-k} โดย $-1 \leq r_k \leq 1$ ซึ่งหาก r_k เข้าใกล้ 1 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลา (time lag) มีสหสัมพันธ์กันสูง หาก r_k เข้าใกล้ 0 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลา (time lag) มีสหสัมพันธ์ต่ำ หาก r_k เข้าใกล้ -1 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลา (time lag) มีสหสัมพันธ์กันสูง ในทางตรงข้าม และนอกจากนี้ r_{kk} เป็นค่าโดยประมาณของ ρ_{kk} ซึ่ง

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

โดย

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad , j=1,2,\dots,k-1$$

$$V(r_{kk}) = \frac{1}{n} \quad k=1,2,\dots$$

r_k เป็นค่าวัดลักษณะของตัวอย่างที่เป็นค่าประมาณของค่าวัดลักษณะของประชากร ρ_k โดยที่ r_k เป็น Gaussian Distribution ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ_{rk}^2 หรือ $r_k \sim N(0, \sigma_{rk}^2)$ ซึ่งจะประมาณ σ_{rk}^2 ด้วย s_{rk}^2 การทราบลักษณะการแจกแจงของ r_k จะสามารถทดสอบสมมติฐานเกี่ยวกับ ρ_k โดยที่

$$H_0 : \rho_k = 0$$

$$H_1 : \rho_k \neq 0$$

และจะใช้ตัวทดสอบ Z

$$Z = \frac{r_k}{s_{rk}} = \frac{r_k}{\sqrt{\frac{1}{n}}} = r_k \sqrt{n}$$

ที่มีช่วงวิกฤติ $|Z| \geq z_{\alpha/2}$ ที่ระดับนัยสำคัญ α ในทำนองเดียวกันอาจจะใช้ตัวทดสอบสถิติ r_k ที่มีช่วงวิกฤติ $|r_k| \geq z_{\alpha/2} \sqrt{\frac{1}{n}}$ โดยทั่วไปใช้ที่ระดับนัยสำคัญเท่ากับ 0.05 จึงมีช่วงวิกฤติเป็น $|r_k| \geq \frac{1.96}{\sqrt{n}} \approx \frac{2}{\sqrt{n}}$

r_{kk} เป็นค่าวัดลักษณะของตัวอย่างที่เป็นค่าประมาณของค่าวัดลักษณะของประชากร ρ_{kk} โดยที่ r_{kk} เป็น Gaussian Distribution ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ_{rkk}^2 หรือ $r_{kk} \sim N(0, \sigma_{rkk}^2)$ ซึ่งจะประมาณ σ_{rkk}^2 ด้วย s_{rkk}^2 ดังนั้นการทดสอบสมมติฐานว่า r_{kk} มาจาก ρ_{kk} ที่เป็น 0 หรือไม่จะทดสอบโดย

$$H_0 : \rho_{kk} = 0$$

$$H_1 : \rho_{kk} \neq 0$$

จะใช้ตัวทดสอบสถิติ r_{kk} และจะปฏิเสธ H_0 ที่ระดับนัยสำคัญ 0.05 เมื่อ $|r_{kk}| \geq \frac{2}{\sqrt{n}}$

Identification (การกำหนดรูปแบบของอนุกรมเวลา)

สำหรับ Stationary Time Series จะกำหนดโดยการเปรียบเทียบ Corelogram ของ r_k, r_{kk} กับ ρ_k, ρ_{kk} ของรูปแบบต่าง ๆ โดยก่อนการพิจารณาว่ากำหนดรูปแบบให้ประมาณช่วงขอบเขตก่อน นั่นคือ $|r_k| > \frac{2}{\sqrt{n}}$ หรือไม่ และ $|r_{kk}| > \frac{2}{\sqrt{n}}$ หรือไม่

กรณีอนุกรมเวลา $\{Y_t\}$ มีค่าเฉลี่ย \bar{Y} สูงเมื่อเทียบกับ s_y ซึ่งจะทำให้ค่าของตัวทดสอบสถิติสำหรับการทดสอบสมมติฐาน $H_0 : \mu = 0$ กับ $H_1 : \mu \neq 0$ มีค่าสูงและทำให้ปฏิเสธ H_0 นั่นคือ μ จะมีค่าต่างจาก 0 เนื่องจากพารามิเตอร์ θ_0 ที่ปรากฏในทุกรูปแบบเขียนได้เป็นฟังก์ชันของ μ ดังนั้นถ้า $\mu \neq 0$ แล้ว $\theta_0 \neq 0$ แต่ถ้า $\mu = 0$ แล้ว $\theta_0 = 0$ ดังนั้นรูปแบบ $ARMA(p, q)$ ที่กำหนดจะมี θ_0 หรือไม่ขึ้นอยู่กับค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของอนุกรมเวลา $\{Y_t\}$

Estimation (การประมาณค่าพารามิเตอร์)

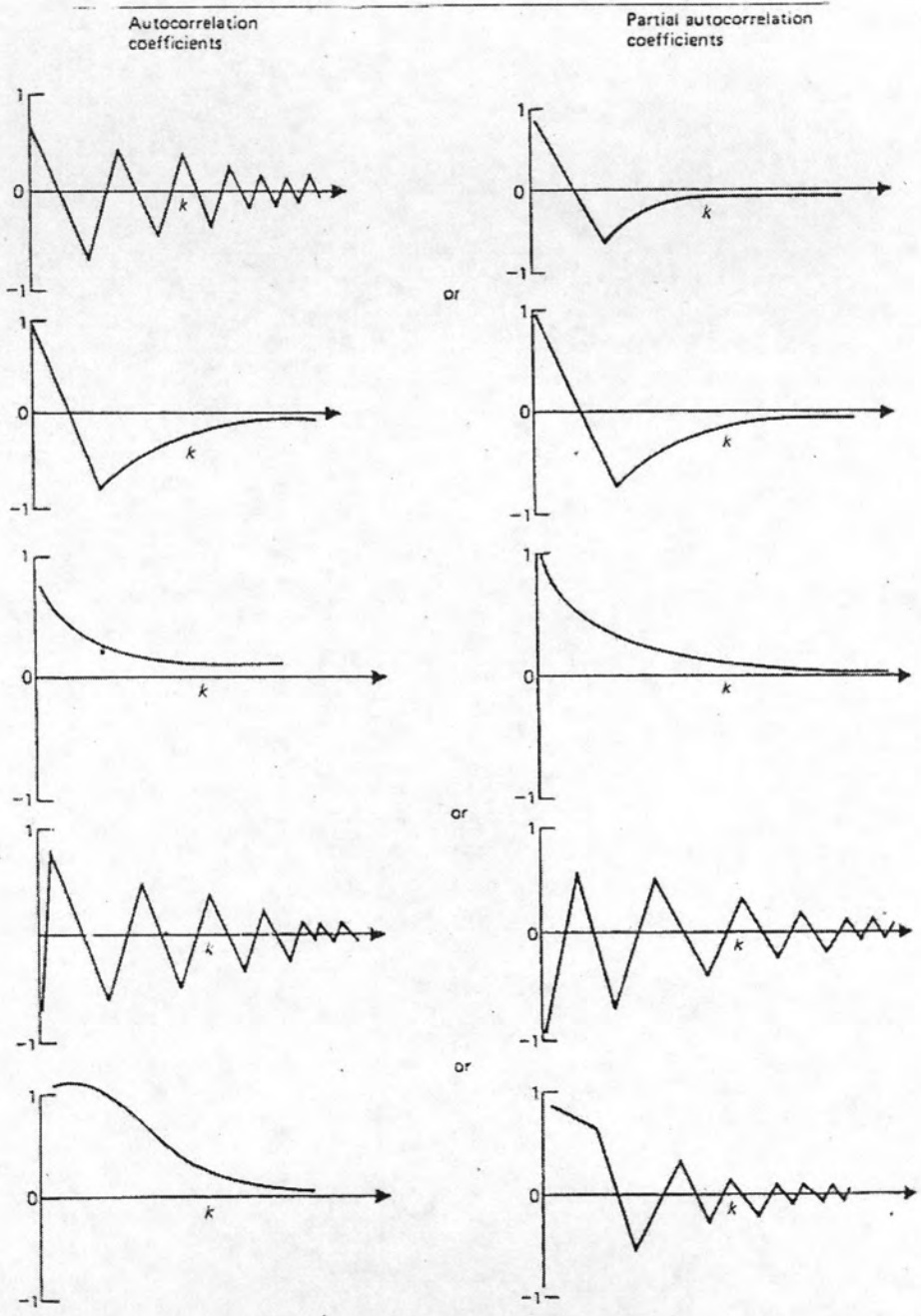
การกำหนดรูปแบบ $ARMA(p, q)$ ให้กับ Stationary Time Series ในขั้นตอน Identification เป็นขั้นตอนแรกของ Box-Jenkins Method ขั้นตอนต่อไปจะเป็นประมาณค่าพารามิเตอร์ในรูปแบบวิธีในขั้นตอนนี้มี 2 วิธี

1) วิธีการประมาณแบบง่ายจะใช้สมการที่แสดงความสัมพันธ์ระหว่าง ρ_k และพารามิเตอร์ในรูปแบบ จำนวนสมการที่ใช้จะเท่ากับอันดับที่ปรากฏในรูปแบบ เมื่อประมาณ ρ_k ด้วย r_k ในสมการความสัมพันธ์และแก้สมการจะได้ค่าประมาณของพารามิเตอร์ ส่วนค่าประมาณของ θ_0 จะได้จากสมการที่แสดงความสัมพันธ์ระหว่าง $E(Y_t)$ กับพารามิเตอร์ในรูปแบบ เช่น รูปแบบ $AR(1)$ จะหา $\hat{\phi}_1$ ได้จากการแก้สมการ $\rho_1 = \phi_1$ และ $\hat{\phi}_0$ ได้จากการแก้สมการ

$$\mu = \frac{\theta_0}{1 - \phi_1}$$

เมื่อแทน ρ_1 และ μ ด้วย r_1 และ \bar{Y} ในสมการความสัมพันธ์ทั้งสองตามลำดับ จะได้ค่าประมาณ $\hat{\phi}_0$ และ $\hat{\phi}_1$

2) การประมาณด้วยวิธีกำลังสองน้อยที่สุด ค่าประมาณของพารามิเตอร์ในรูปแบบจะเป็นค่าประมาณที่ทำให้ผลรวมกำลังสองของการประมาณน้อยที่สุด การหาค่าประมาณนี้จะใช้วิธีการวิเคราะห์เชิงตัวเลข (Numerical Analysis Method) โดยใช้ค่าที่ได้จากประมาณแบบง่ายใช้เป็นค่าเริ่มต้น สำหรับค่าคลาดเคลื่อนมาตรฐานเป็นไปตามตารางที่ 3



AUTOCORRELATION AND PARTIAL
 AUTOCORRELATION FUNCTIONS OF A MIXED
 ARMA (1, 1) MODEL

รูปที่ 2.2 รูปแบบทั่วไปของ ACF และ PACF

Diagnostic Checking (การตรวจสอบรูปแบบ)

การกำหนดรูปแบบให้กับ Stationary Time Series ในขั้นตอน Identification ซึ่งเป็นขั้นตอนแรกของ Box-Jenkins Method นั้น รูปแบบที่กำหนดอาจจะไม่ใช่รูปแบบที่เหมาะสม เนื่องจาก Corelogram ของ r_k และ r_{kk} ไม่เป็นไปตามลักษณะของ Corelogram ของ ρ_k และ ρ_{kk} ของรูปแบบใดแน่นอน ดังนั้นจึงต้องตรวจสอบความเหมาะสมของรูปแบบ ซึ่งมีหลายวิธีในการตรวจสอบ วิธีที่ตรวจสอบส่วนใหญ่จะพิจารณาที่ค่า Residual, e_t ซึ่ง

$$e_t = Y_t - \hat{Y}_t$$

คือผลต่างระหว่างค่าจริงกับค่าพยากรณ์ มาใช้เป็นหลักในการตรวจสอบความเหมาะสมของรูปแบบดังนี้

1) พิจารณาว่า Time Series $\{e_t\}$ ว่ามีการเคลื่อนไหวที่เป็นอิสระกันหรือไม่ นั่นคือเป็นการตรวจสอบข้อจำกัดของรูปแบบที่กำหนดว่า e_t จะต้องมีการแจกแจงที่เป็นอิสระกันจึงจะสรุปว่ารูปแบบที่ Identification มานั้นเหมาะสม โดยจะทดสอบสมมติฐานดังนี้

$$H_0 : \rho_k(e_t) = 0$$

$$H_1 : \rho_k(e_t) \neq 0, k = 1, 2, \dots$$

การปฏิเสธ H_0 ที่นัยสำคัญ 0.05 จะทำเมื่อ $|r_k(e_t)| \geq \frac{2}{\sqrt{n}}$ ซึ่ง n คือขนาดของอนุกรมเวลา $\{e_t\}$

2) พิจารณาค่าความแปรปรวนของค่าความคลาดเคลื่อน

$$s^2 = \sum \frac{e_t^2}{n}$$

ซึ่ง n คือจำนวนค่าความคลาดเคลื่อน รูปแบบที่เหมาะสมจะให้ s^2 ต่ำ

3) พิจารณาค่าพารามิเตอร์ในรูปแบบมีค่าเป็น 0 หรือไม่ นั่นคือเมื่อ $\theta, \hat{\theta}, s_\theta$ เป็นพารามิเตอร์ ค่าประมาณ และค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณ $\hat{\theta}$ ตามลำดับ การทดสอบสมมติฐาน $H_0 : \theta = 0$ กับ $H_1 : \theta \neq 0$ จะใช้ตัวทดสอบสถิติ

$$Z = \frac{\hat{\theta}}{s_{\hat{\theta}}}$$

การปฏิเสธ H_0 จะทำเมื่อ $|Z| \geq z_{\frac{\alpha}{2}}$ ที่ระดับนัยสำคัญ α เช่น รูปแบบ $AR(2)$ มีพารามิเตอร์ ϕ_1, ϕ_2 การทดสอบว่า ϕ_2 เป็น 0 หรือไม่นั้น จะเป็นการทดสอบว่าแทนที่จะเป็นรูปแบบ $AR(2)$ ควรจะเป็นรูปแบบ $AR(1)$ ใช่หรือไม่ นั่นคือ

$$H_0 : \phi_2 = 0$$

$$\text{หรือ } Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

$$\text{หรือ } Y_t \sim AR(1)$$

$$H_1 : \phi_2 \neq 0$$

$$\text{หรือ } Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

$$\text{หรือ } Y_t \sim AR(2)$$

4) พิจารณาว่า $\rho_k(e_t) = 0 \quad k=1, \dots, m$ หรือไม่ นั่นคือพิจารณาว่าค่าความคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน $k=1, \dots, m$ ช่วงเวลาเป็นอิสระกันหรือไม่ จะกำหนดสมมติฐานเป็น $H_0 : \rho_1(e_1) = \dots = \rho_m(e_m) = 0$ กับ $H_1 : \rho_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เป็น 0 สำหรับ $k=1, \dots, m$ การทดสอบสมมติฐานจะทำโดยใช้การทดสอบของ Box & Pierce หรือ การทดสอบของ Box & Ljung ซึ่งต่างก็เป็นการทดสอบสมมติฐานที่กำหนดสมมติฐานหลักและสมมติฐานทางเลือกเหมือนกัน มีช่วงวิกฤติเดียวกันแต่มีตัวทดสอบสถิติต่างกัน

$$4.1) \quad H_0 : \rho_1(e_1) = \dots = \rho_m(e_m) = 0$$

$$H_1 : \rho_k(e_t) \text{ อย่างน้อยหนึ่งค่าไม่เป็น 0 สำหรับ } k=1, \dots, m$$

4.2) ตัวทดสอบสถิติ Box & Pierce

$$Q = n \sum_{k=1}^m r_k^2(e_t)$$

ตัวทดสอบสถิติ Box & Ljung

$$Q_m = n(n+2) \sum_{k=1}^m \frac{r_k^2(e_t)}{n-k}$$

โดยที่ n คือขนาดของอนุกรมเวลา $\{e_t\}$

m คือช่วงเวลาห่างสูงสุดของ e_t ในอนุกรมเวลา $\{e_t\}$ ที่นำมาพิจารณา

a คือจำนวนพารามิเตอร์ทั้งหมดในรูปแบบซึ่งรวมทั้ง θ_0 ด้วย

4.3) สำหรับช่วงวิกฤติ Q

$$Q_m \geq \chi_{\alpha, m-a}^2$$

4.4) การสรุปผลทดสอบ

เมื่อปฏิเสธ H_0 จะสรุปได้ว่ามี $\rho_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เท่ากับ 0 สำหรับ $k = 1, \dots, m$ นั่นคือมีสหสัมพันธ์แบบออโตระหว่างค่าความคลาดเคลื่อนของการพยากรณ์ที่ห่างกัน k ค่า หรือสรุปได้ว่ารูปแบบที่กำหนดให้กับอนุกรมเวลายังไม่เหมาะสม

Forecasting, Prediction (การพยากรณ์)

เมื่อผ่านขั้นตอนของ Diagnostic Checking แล้วว่าการกำหนดรูปแบบให้กับ Stationary Time Series นั้นเหมาะสม ต่อไปจะเป็นขั้นตอนสุดท้ายคือการพยากรณ์ ซึ่งการพยากรณ์มี 2 แบบคือการพยากรณ์แบบจุด และการพยากรณ์แบบช่วง

1) การพยากรณ์แบบจุด : ค่าพยากรณ์ในอนาคตจะได้จากการแทนค่าสังเกตในอดีตหรือค่าความคลาดเคลื่อนในอดีตลงในสมการพยากรณ์ ในสมการพยากรณ์จะแทนค่าประมาณของพารามิเตอร์ที่ได้ประมาณจากขั้นตอน Estimation สำหรับรูปแบบ $ARMA(p, q)$ ซึ่ง $p+q \leq 2$ มีสมการพยากรณ์ $\hat{Y}_t(l)$ ตามตารางที่ 2.4

ตารางที่ 2.3 ตัวประมาณค่าและค่าคลาดเคลื่อนมาตรฐาน

รูปแบบ	ตัวประมาณ	ค่าคลาดเคลื่อนมาตรฐาน
$AR(1)$	$\hat{\phi}_1$	$\sqrt{\frac{1-\hat{\phi}_1^2}{n}}$
$AR(2)$	$\hat{\phi}_1, \hat{\phi}_2$	$\sqrt{\frac{1-\hat{\phi}_2^2}{n}}$
$MA(1)$	$\hat{\theta}_1$	$\sqrt{\frac{1-\hat{\theta}_1^2}{n}}$
$MA(2)$	$\hat{\theta}_1, \hat{\theta}_2$	$\sqrt{\frac{1-\hat{\theta}_2^2}{n}}$
$ARMA(1,1)$	$\hat{\phi}_1$	$\sqrt{\frac{(1-\hat{\phi}_1\hat{\phi}_1)(1-\hat{\theta}_1\hat{\phi}_1)}{n(\hat{\phi}_1-\hat{\theta}_1)^2}}$
	$\hat{\theta}_1$	$\sqrt{\frac{(1-\hat{\theta}_1\hat{\theta}_1)(1-\hat{\theta}_1\hat{\phi}_1)}{n(\hat{\phi}_1-\hat{\theta}_1)^2}}$

สมการพยากรณ์จะเขียนได้ทั้งในเทอมของพารามิเตอร์ μ และพารามิเตอร์อื่น หรือในเทอมของพารามิเตอร์ θ_0 และพารามิเตอร์อื่น การใช้สมการพยากรณ์จะทำได้โดยการแทนค่าประมาณของพารามิเตอร์ในรูปแบบ ลักษณะรูปแบบ $ARMA(p, q)$ ซึ่ง $p+q \leq 2$ เมื่อกำหนด $\hat{Y}_t(l)$ เป็นค่าพยากรณ์ที่เวลา $t+1$ ใช้ค่าสังเกต t ค่า โดยที่

$$\hat{Y}_t(l) = \begin{cases} Y_{t+1} & , l \leq 0 \\ \hat{Y}_t(l) & l \geq 1 \end{cases}$$

และ $\hat{e}_t(l)$ คือ Residual ของการพยากรณ์ที่เวลา $t+1$ ใช้ค่าสังเกต t ค่า โดย

ตารางที่ 2.4 รูปแบบของสมการพยากรณ์

รูปแบบ	สมการพยากรณ์ $\hat{Y}_t(l)$ (ใน เทอมของ $\hat{\mu}$)	สมการพยากรณ์ $\hat{Y}_t(l)$ (ในเทอมของ $\hat{\theta}_0$)	$\sigma_{e(l)}^2$
*	$\hat{\mu}, l \geq 1$	$\hat{\theta}_0, l \geq 1$	σ^2
AR(1)	$\hat{\mu} + \hat{\phi}_1(Y_t - \hat{\mu})$, $l \geq 1$	$\hat{\theta}_0 + \hat{\theta}_1 Y_t$, $l = 1$ $\hat{\theta}_0 + \hat{\theta}_1 \hat{Y}_t(l-1)$, $l \geq 2$	$\frac{1 - \hat{\phi}_1^{2l}}{1 - \hat{\phi}_1^2} \sigma^2$
AR(2)	$\hat{\mu} + \hat{\phi}_1(Y_t - \hat{\mu}) + \hat{\phi}_2(Y_{t-1} - \hat{\mu})$, $l = 1$ $\hat{\mu} + \hat{\phi}_1(\hat{Y}_t(l) - \hat{\mu}) + \hat{\phi}_2(Y_t - \hat{\mu})$, $l = 2$ $\hat{\mu} + \hat{\phi}_1(\hat{Y}_t(l-1) - \hat{\mu}) +$ $\hat{\phi}_2(Y_t(l-2) - \hat{\mu}), l \geq 3$	$\hat{\theta}_0 + \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1}$, $l = 1$ $\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l) + \hat{\phi}_2 Y_t$, $l = 2$ $\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1) +$ $\hat{\phi}_2 \hat{Y}_t(l-2), l \geq 3$	$(1 + \sum_{j=1}^{l-1} \hat{\psi}_j^2) \sigma^2$ $\hat{\psi}_0 = 1, \hat{\psi}_1 = \hat{\phi}_1$ $\hat{\psi}_j = \hat{\phi}_1 \hat{\psi}_{j-1} + \hat{\phi}_2 \hat{\psi}_{j-2}$, $j > 1$
MA(1)	$\hat{\mu} - \hat{\theta}_1 e_t, l = 1$ $\hat{\mu}, l \geq 2$	$\hat{\theta}_0 - \hat{\theta}_1 e_t, l = 1$ $\hat{\theta}_0, l \geq 2$	$\sigma^2, l = 1$ $(1 + \hat{\theta}_1^2) \sigma^2, l \geq 2$
MA(2)	$\hat{\mu} - \hat{\theta}_1 e_t - \hat{\theta}_2 e_{t-1}$, $l = 1$ $\hat{\mu} - \hat{\theta}_2 e_t$, $l = 2$ $\hat{\mu}$, $l \geq 3$	$\hat{\theta}_0 - \hat{\theta}_1 e_t - \hat{\theta}_2 e_{t-1}$, $l = 1$ $\hat{\theta}_0 - \hat{\theta}_2 e_t$, $l = 2$ $\hat{\theta}_0$, $l \geq 3$	σ^2 , $l = 1$ $(1 + \hat{\theta}_1^2) \sigma^2$, $l = 2$ $(1 + \hat{\theta}_1^2 + \hat{\theta}_2^2) \sigma^2$, $l \geq 3$
**		$\hat{\theta}_0 + \hat{\phi}_1 Y_t - \hat{\theta}_1 e_t$, $l = 1$	

$\hat{\mu} + \hat{\phi}_1 (Y_t - \hat{\mu}) - \hat{\theta}_1 e_t$, $l=1$	$\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1)$, $l \geq 2$	$\frac{1 + (\hat{\phi}_1 - \hat{\theta}_1)^2 (1 - \hat{\phi}_1^{2l-2})}{1 - \hat{\phi}_1^2} \sigma^2$
$\hat{\mu} + \hat{\phi}_1^l (Y_t - \hat{\mu})$, $l \geq 2$		

* White Noise, ** ARMA(1,1)

$$\hat{e}_t(l) = \begin{cases} e_{t+1}, & l \leq 0 \\ 0, & l \geq 1 \end{cases}$$

สำหรับรูปแบบ ARMA(p,q) ซึ่ง $p+q \leq 2$ มีสมการพยากรณ์ดังนี้

1) White Noise : $Y_t = \theta_0 + \varepsilon_t$ โดยที่ $\mu = \theta_0$ มีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_t(l) = \hat{\theta}_0 = \hat{\mu}, l \geq 1$$

2) AR(1) : $Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t$ โดยที่ $\mu = \frac{\theta_0}{1 - \phi_1}$ มีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_t(l) = \begin{cases} \hat{\theta}_0 + \hat{\theta}_1 Y_t, & l=1 \\ \hat{\theta}_0 + \hat{\theta}_1 \hat{Y}_t(l-1), & l \geq 2 \end{cases}$$

$$\text{หรือ } \hat{Y}_t(l) = \begin{cases} \hat{\mu} + \hat{\phi}_1 (Y_t - \hat{\mu}), & l \geq 1 \\ \hat{\mu} + \hat{\phi}_1 (\hat{Y}_t(l-1) - \hat{\mu}), & l \geq 2 \end{cases}$$

3) AR(2) : $Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$ โดยที่ $\mu = \frac{\theta_0}{1 - \phi_1 - \phi_2}$ มีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_t(l) = \begin{cases} \hat{\theta}_0 + \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1}, & l=1 \\ \hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l) + \hat{\phi}_2 Y_t, & l=2 \\ \hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1) + \hat{\phi}_2 \hat{Y}_t(l-2), & l \geq 3 \end{cases}$$

$$\text{หรือ } \hat{Y}_t(l) = \begin{cases} \hat{\mu} + \hat{\phi}_1 (Y_t - \hat{\mu}) + \hat{\phi}_2 (Y_{t-1} - \hat{\mu}), l=1 \\ \hat{\mu} + \hat{\phi}_1 (\hat{Y}_t(l) - \hat{\mu}) + \hat{\phi}_2 (Y_t - \hat{\mu}), l=2 \\ \hat{\mu} + \hat{\phi}_1 (\hat{Y}_t(l-1) - \hat{\mu}) + \hat{\phi}_2 (Y_t(l-2) - \hat{\mu}), l \geq 3 \end{cases}$$

4) $MA(1)$: $Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t$, โดยที่ $\mu = \theta_0$ มีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_t(l) = \begin{cases} \hat{\theta}_0 - \hat{\theta}_1 e_t, l=1 \\ \hat{\theta}_0, l \geq 2 \end{cases}$$

$$\text{โดยที่ } e_t = Y_t - \hat{Y}_{t-1}(l)$$

5) $MA(2)$: $Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} + \varepsilon_t$, โดยที่ $\mu = \theta_0$ มีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_t(l) = \begin{cases} \hat{\theta}_0 - \hat{\theta}_1 e_t - \hat{\theta}_2 e_{t-1}, l=1 \\ \hat{\theta}_0 - \hat{\theta}_2 e_t, l=2 \\ \hat{\theta}_0, l \geq 3 \end{cases}$$

$$\text{โดยที่ } e_t = Y_t - \hat{Y}_{t-1}(l), e_{t-1} = Y_t - \hat{Y}_{t-2}(l)$$

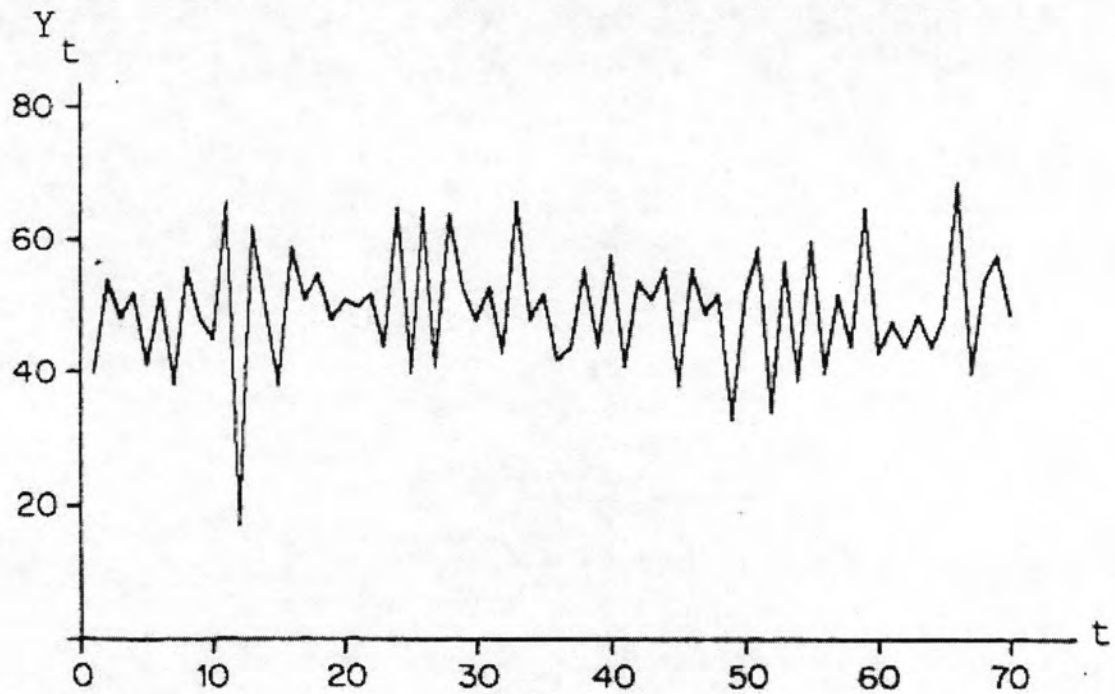
6) $ARMA(1,1)$: $Y_t = \theta_0 + \phi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t$, โดยที่ $\mu = \frac{\theta_0}{1 - \phi_1}$ มีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_t(l) = \begin{cases} \hat{\theta}_0 + \hat{\phi}_1 Y_t - \hat{\theta}_1 e_1, l=1 \\ \hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1), l \geq 2 \end{cases}$$

$$\text{หรือ } \hat{Y}_t(l) = \begin{cases} \hat{\mu} + \hat{\phi}_1 (Y_t - \hat{\mu}) - \hat{\theta}_1 e_1, l=1 \\ \hat{\mu} + \hat{\phi}_1 (Y_t - \hat{\mu}), l \geq 2 \end{cases}$$

การนำไปประยุกต์

พิจารณา Time Series ตามรูปที่ 1.3 ซึ่งเมื่อพิจารณาจาก Corelogram r_k และ r_{kk} ตามรูปที่ 4 และ 5 โดยที่ $k = 1, 2, \dots, 24$



รูปที่ 2.3 ลักษณะของข้อมูล

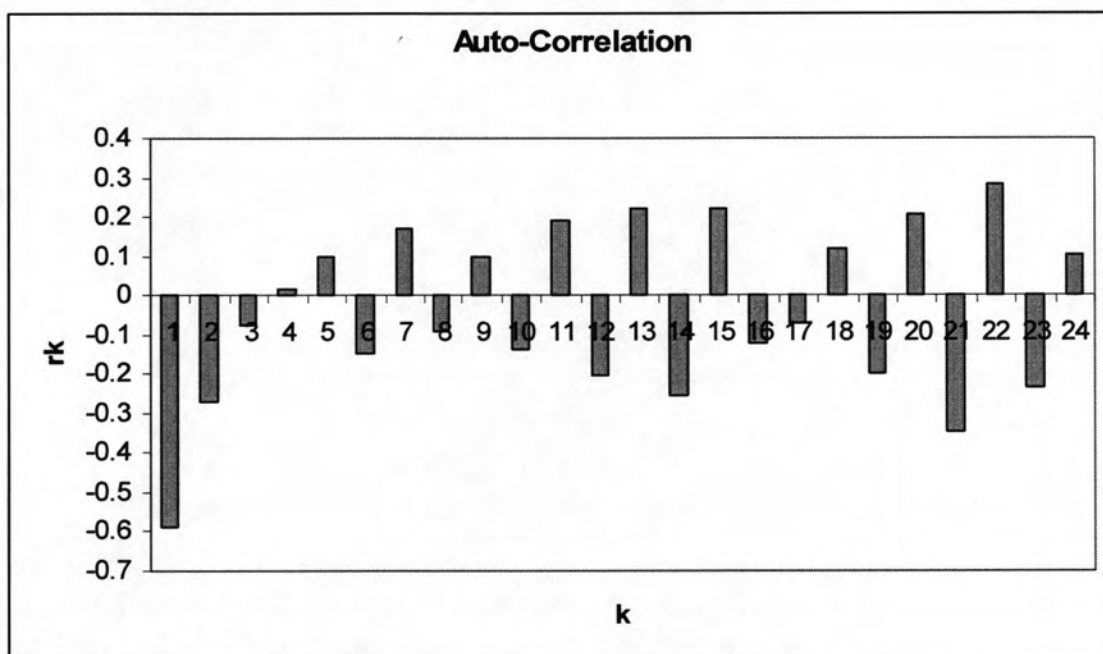
จากรูปที่ 4 และรูปที่ 5 พบว่าจะมีเพียง r_1, r_2, r_{11} เท่านั้นที่อยู่นอกช่วง $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{70}} = \pm 0.24$ และ $\frac{1}{\sqrt{n}} = 0.12$ คือค่าประมาณของ $s_{r_k}, s_{r_{kk}}$ จากการเทียบ Corelogram ของ r_k, r_{kk} กับ ρ_k, ρ_{kk} พบว่าใกล้เคียงกับ $ARMA(1,0)$ model เนื่องจาก r_{11} มีค่าสูง และ r_k มีค่าลดลงเมื่อ k มีค่าสูงขึ้น ดังนั้นจึง Identification ได้ว่า

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t, \quad -1 < \phi_1 < 1$$

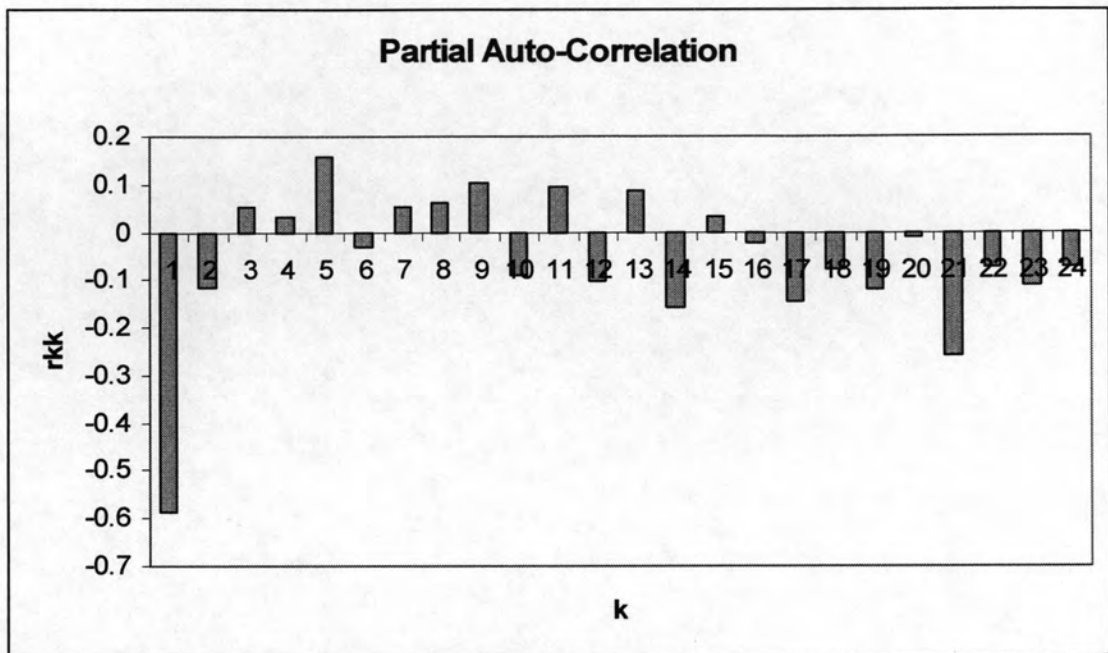
ต่อไปจะเข้าสู่ขั้นตอน Estimation ซึ่ง Stationary Time Series ชุดนี้

$\bar{Y} = 49.686$, $s = 9.2$, $r_1 = -0.588$ การประมาณ ϕ จะทำโดยวิธีกำลังสองน้อยที่สุด โดยกำหนดค่าเริ่มต้นของ ϕ เท่ากับ 0.1 ดังนั้น

$$\begin{aligned}\hat{\theta}_0 &= \hat{\mu}(1 - \hat{\phi}_1) \\ &= \bar{Y}(1 - \hat{\phi}_1) \\ &= 49.686(1 + 0.598) \\ &= 79.398\end{aligned}$$



รูปที่ 2.4 (r_k ของ Y_t)

รูปที่ 2.5 (r_{kk} ของ Y_t)

ซึ่งผลการประมาณด้วยวิธีกำลังสองน้อยที่สุดพบว่า

$$\hat{\theta}_0 = 79.398$$

$$\hat{\theta}_1 = -0.598$$

ดังนั้นผลการขั้นตอน Estimation พบว่า

$$Y_t = 79.398 - 0.598Y_{t-1} + \varepsilon_t$$

ต่อไปจะเข้าสู่ขั้นตอน Diagnostic Checking เนื่องจาก Corelogram ของ r_k, r_{kk} ไม่เป็นไปตาม ρ_k, ρ_{kk} ของรูปแบบใดรูปแบบหนึ่งแน่นอน ดังนั้นจึงต้องตรวจสอบความเหมาะสมของรูปแบบ การตรวจสอบมีหลายวิธี และควรใช้หลายวิธีพร้อม ๆ กัน การตรวจสอบทำได้ดังนี้

1) พิจารณาการเคลื่อนไหวของอนุกรมเวลา $\{e_t\}$ ดังรูปที่ 2.6 ซึ่งมีค่าความแปรปรวนของค่าความคลาดเคลื่อนเท่ากับ 54.606 ว่าเป็นอิสระต่อกันหรือไม่ โดยพิจารณาจาก $r_k(e_t)$ โดยทดสอบสมมติฐาน

$$H_0 : \rho_k(e_t) = 0$$

$$H_1 : \rho_k(e_t) \neq 0, k = 1, 2, \dots, 24$$

ใช้ตัวทดสอบสถิติ $r_k(e_t)$

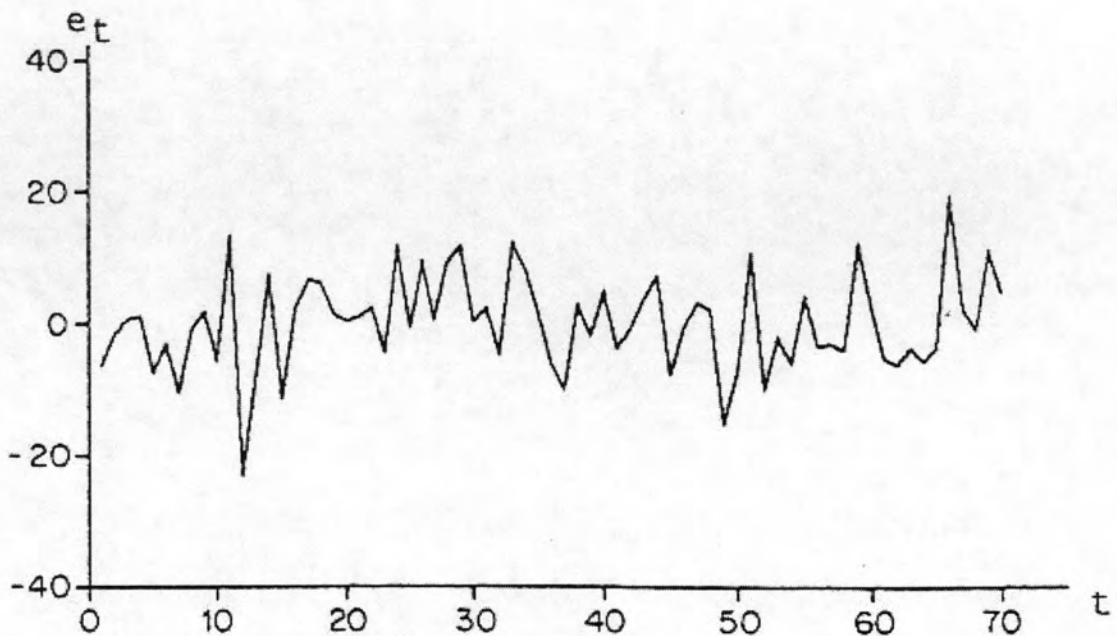
ซึ่งเป็นไปตาม Corelogram ตามรูปที่ 1.7 และ 1.8 มีช่วงวิกฤติ

$$|r_k(e_t)| \geq \frac{2}{\sqrt{69}}$$

หรือ $|r_k(e_t)| \geq 0.24$ ซึ่ง 69 คือขนาดของอนุกรมเวลา $\{e_t\}$ จะยอมรับ H_0 สรุปได้ว่า $\rho_k(e_t)$ มีค่าเป็น 0 สำหรับ $k = 1, 2, \dots, 24$ ดังนั้นอนุกรมเวลา $\{e_t\}$ มีการเคลื่อนไหวที่เป็นอิสระกัน

2) พิจารณาว่าพารามิเตอร์ที่ปรากฏในรูปแบบเป็น 0 หรือไม่ นั่นคือทดสอบสมมติฐาน

$$H_0 : \phi_1 = 0, \quad H_1 : \phi_1 \neq 0$$



รูปที่ 2.6 ลักษณะการเคลื่อนไหวของอนุกรมเวลา

มีค่าของตัวทดสอบสถิติ $Z = -6.229$ ซึ่งได้จาก

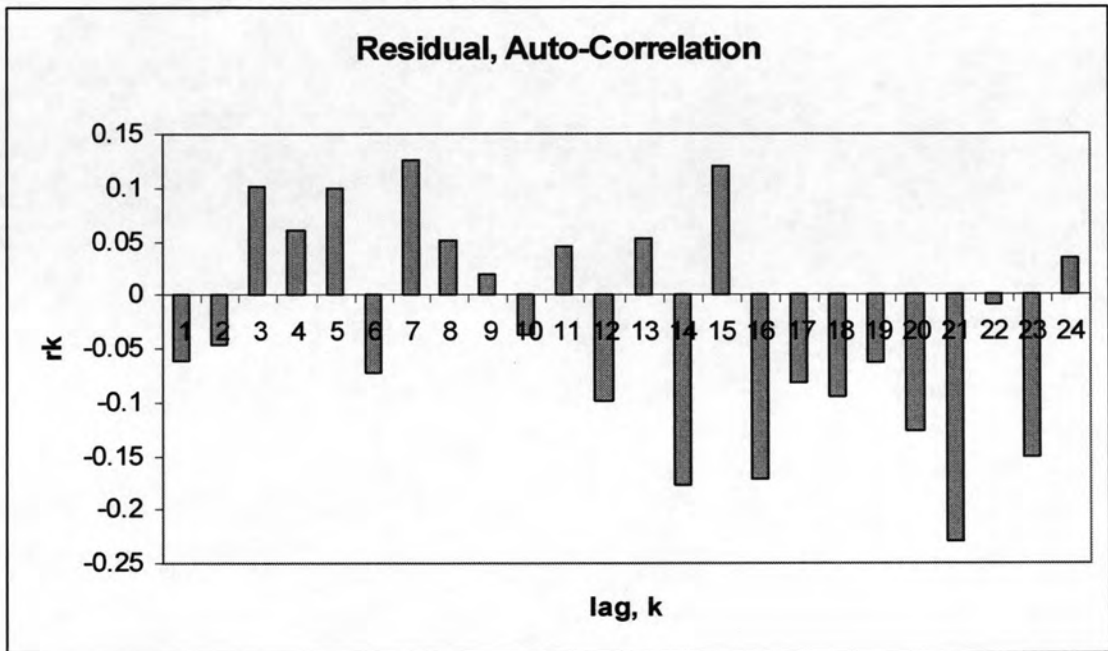
$$\hat{\phi}_1 = -0.598$$

และ
$$s_{\hat{\phi}_1} = \sqrt{\frac{1 - (\hat{\phi}_1)^2}{n}} = \sqrt{\frac{1 - (-0.598)^2}{70}} = 0.096$$

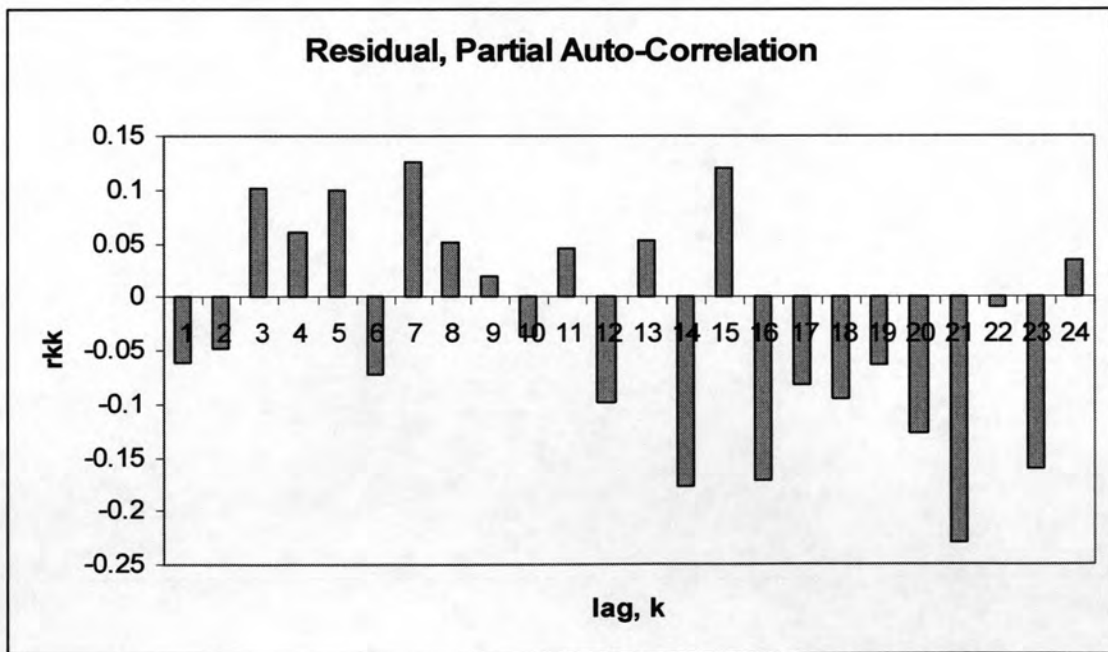
ดังนั้นสรุปได้ว่า ϕ_1 มีค่าต่างจาก 0 ที่ระดับนัยสำคัญ 0.05 นั่นคือรูปแบบ $AR(1)$ คือรูปแบบที่เหมาะสม

3) พิจารณาว่า $\rho_k(e_t) = 0$: $k = 1, \dots, 22$ มีค่าเป็น 0 พร้อมกันหรือไม่ เหตุที่กำหนด $m = 22$ เนื่องจากค่านวนค่าของตัวทดสอบสถิติ Q ที่ขึ้นแห่งความเป็นอิสระเท่ากับ 20 สำหรับการทดสอบของ Box และ Pierce และของ Box และ Ljung ที่มี

$$H_0 : \rho_1(e_t) = \dots = \rho_{22}(e_t) = 0$$



รูปที่ 2.7 (r_k ของ e_t)



รูปที่ 2.8 (r_{kk} ของ e_t)

$H_1 : \rho_k(e_t), k = 1, 2, \dots, 22$ มีอย่างน้อยหนึ่งค่าที่ไม่เท่ากับ 0 มีค่าของตัวทดสอบสถิติ
ดังนี้

3.1) Box และ Pierce

$$Q = 69 \sum_{k=1}^{22} r_k^2(e_t) = 15.776$$

3.2) Box & Ljung

$$Q_m = 69(71) \sum_{k=1}^{22} \frac{r_k^2(e_t)}{69-k} = 21.209$$

การทดสอบทั้งสองมีช่วงวิกฤติ Q หรือ $Q_m \geq \chi_{0.05, 20}^2 = 31.41$ จะยอมรับ H_0 ที่ระดับนัยสำคัญ 0.05
นั่นคือ $\rho_k(e_t)$ มีค่าเป็น 0 สำหรับ $k = 1, 2, \dots, 22$ แสดงว่าค่าความคลาดเคลื่อนของการพยากรณ์ที่ห่าง
กัน $1, 2, \dots, 22$ ช่วงเวลา มีการเคลื่อนไหวที่เป็นอิสระกัน หรือรูปแบบ $AR(1)$ เหมาะสมกับอนุกรม
เวลาแล้ว ขั้นตอนต่อไปคือการพยากรณ์ โดยมีสมการพยากรณ์เป็นดังนี้

$$\hat{Y}_{70}(l) = \begin{cases} 79.398 - 0.598Y_{70}, l = 1 \\ 79.398 - 0.568\hat{Y}_{70}(l-1), l \geq 2 \end{cases}$$

เมื่อกำหนดให้ $Y_{70} = 49$ ดังนั้น

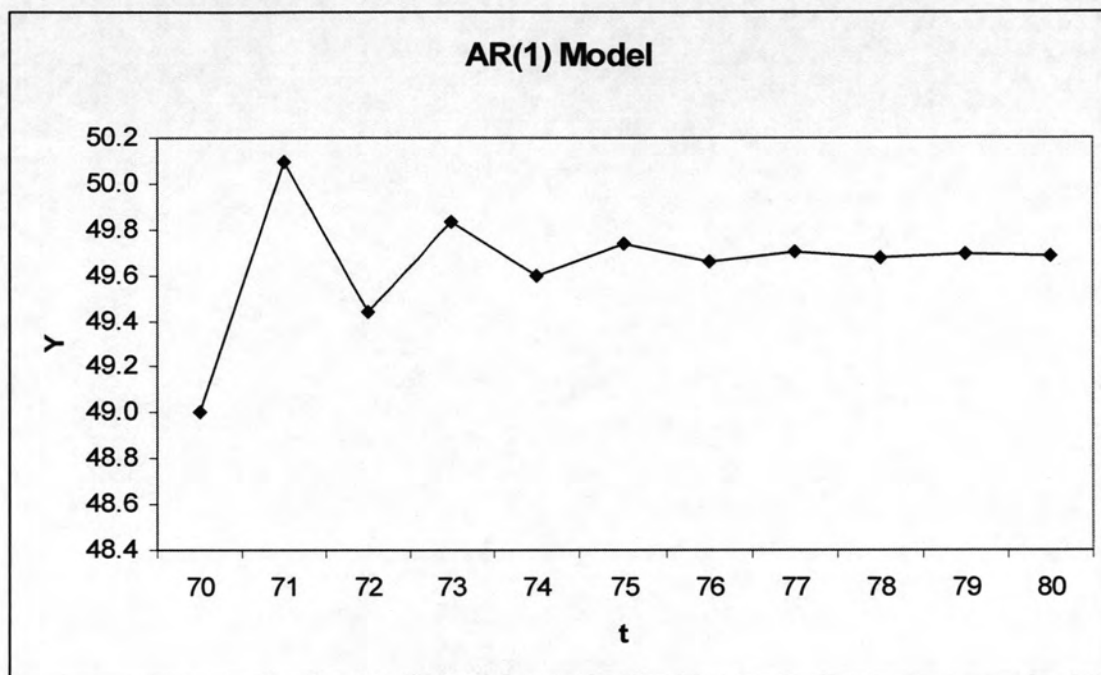
$$\hat{Y}_{71} = 79.398 - 0.598 * (49.000) = 50.096$$

$$\hat{Y}_{72} = 79.398 - 0.598 * (50.096) = 49.441$$

·
·
·

$$\hat{Y}_{80} = 79.398 - 0.598 * (49.692) = 49.682$$

ซึ่งแสดงค่าพยากรณ์ได้ตามรูปที่ 2.9



รูปที่ 2.9 ลักษณะของ AR(1)

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

พงษ์ศักดิ์ พุทธวงศ์ (2546)

วิทยานิพนธ์ฉบับนี้กล่าวถึงศึกษา ความต้องการไฟฟ้าในอุตสาหกรรมผลิต โลหะ อุตสาหกรรมผลิตปูนซีเมนต์และภาคอุตสาหกรรมในประเทศไทย โดยมีวัตถุประสงค์ เพื่อศึกษาความต้องการพลังงานไฟฟ้าของภาคอุตสาหกรรมในประเทศไทยกับปัจจัยที่มีอิทธิพลต่อความต้องการพลังงานไฟฟ้า การพยากรณ์ความต้องการพลังงานไฟฟ้าและกำลังไฟฟ้าสูงสุดระหว่างปี พ.ศ. 2545–2551 โดยใช้วิธีการวิเคราะห์ข้อมูลด้วยการวิเคราะห์สมการพหุคูณ(Multiple Regression Analysis)

โดยแบบจำลองทางสถิติ จากความสัมพันธ์ของตัวแปรในแบบจำลองทางเศรษฐศาสตร์จะได้ว่า

$$D_{es} = a + bP_e + cP_f + dR + eQ_s + U$$

เมื่อกำหนดให้

D_{es}	=	ปริมาณความต้องการใช้ไฟฟ้าของอุตสาหกรรม
P_e	=	ราคาไฟฟ้าเฉลี่ย
P_f	=	ราคาก๊าซเฉลี่ย
R	=	อัตราดอกเบี้ยเงินกู้เฉลี่ย
Q_s	=	ปริมาณผลผลิตของอุตสาหกรรม
U	=	ค่าความผิดพลาด(Disturbance Term)
a	=	ค่าคงที่ (Constant)
b,c,d,e	=	สัมประสิทธิ์ของตัวแปรต่างๆ (Coefficients)

ผลการศึกษาพบว่า ปัจจัยที่มีอิทธิพลต่ออุปสงค์ความต้องการพลังงานไฟฟ้าของอุตสาหกรรมผลิตโลหะและปูนซีเมนต์ มีดังนี้ ราคาไฟฟ้าเฉลี่ย ราคาก๊าซเฉลี่ย และปริมาณผลผลิตของอุตสาหกรรม ได้ค่าความยืดหยุ่นของอุปสงค์ต่อราคาไฟฟ้า มีค่าเท่ากับ -1.565 ของอุตสาหกรรมผลิตโลหะ และ -0.783 ของอุตสาหกรรมปูนซีเมนต์ ค่าความยืดหยุ่นของอุปสงค์ต่อผลผลิต มีค่าเท่ากับ 0.11 ของอุตสาหกรรมผลิตโลหะ และค่าความยืดหยุ่นของอุตสาหกรรมผลิตปูนซีเมนต์ มีค่าเท่ากับ 0.449 และค่าความยืดหยุ่นของอุปสงค์ต่อราคาก๊าซของอุตสาหกรรมผลิตโลหะและปูนซีเมนต์ มีค่าเท่ากับ 0.617 และ 0.332

ผลของการพยากรณ์ความต้องการพลังงานไฟฟ้าของภาคอุตสาหกรรม อุตสาหกรรมผลิตโลหะ และอุตสาหกรรมปูนซีเมนต์ ในช่วง พ.ศ. 2545-2551 มีอัตราการขยายตัวเฉลี่ยร้อยละ 4.52, 3.90 และ 3.52 ต่อปีตามลำดับ ผลของการพยากรณ์ความต้องการกำลังไฟฟ้าสูงสุดของภาคอุตสาหกรรม มีอัตราการขยายตัวเฉลี่ยประมาณ 500 เมกะวัตต์ต่อปี

วรารัตน์ สุขุมาพันธ์ (2546)

การพยากรณ์ความต้องการใช้ไฟฟ้าของประเทศไทย ซึ่งได้แยกตามประเภทผู้ใช้ไฟฟ้า ในการศึกษา มีวัตถุประสงค์เพื่อหาความสัมพันธ์ระหว่างความต้องการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าแยกตามประเภทผู้ใช้ไฟฟ้ากับตัวแปรที่สำคัญต่างๆ ที่มีผลต่อการใช้ไฟฟ้า เพื่อพยากรณ์ความต้องการใช้ไฟฟ้าของประเทศไทย แยกตามประเภทผู้ใช้ไฟฟ้า โดยใช้ข้อมูลในอดีตของการใช้ไฟฟ้าและข้อมูลตัวแปรที่สำคัญต่างๆ ที่มีผลต่อการใช้ไฟฟ้า ย้อนหลังอย่างน้อย 15 ปี

ใช้การวิเคราะห์ข้อมูลด้วยวิธี ทฤษฎีฟังก์ชันอุปสงค์ (demand function) ทฤษฎีการวิเคราะห์ความถดถอยแบบพหุคูณ (multiple regression analysis) การพยากรณ์อนุกรมเวลา (time series approach)

จากการศึกษาพบว่า การใช้ไฟฟ้าของผู้ใช้ไฟฟ้าประเภทบ้านอยู่อาศัย พบว่ามีแนวโน้มเพิ่มสูงขึ้นมาโดยตลอด ยกเว้นในช่วงปี 2541-2542 ที่การใช้ไฟฟ้าลดลงจากปีที่ผ่านมา และจำนวนรายของผู้ใช้ไฟฟ้าทุกประเภทมีแนวโน้มเพิ่มขึ้น โดยตลอดตามการขยายตัวของภาวะเศรษฐกิจ และนโยบายของภาครัฐ ที่ต้องการให้ทุกหมู่บ้านในประเทศไทยมีไฟฟ้าใช้ นอกจากนี้ยังพบว่า จำนวนผู้ใช้ไฟฟ้าของผู้ใช้ไฟฟ้าทุกประเภทไม่ได้รับผลกระทบจากภาวะวิกฤติทางเศรษฐกิจของไทย กล่าวคือ ยังคงมีการขยายตัวอยู่เกือบทุกประเภทผู้ใช้ไฟฟ้าเพียงแต่ในช่วงวิกฤติเศรษฐกิจอาจมีอัตราการเพิ่มที่ลดลงบ้าง ยกเว้น ผู้ใช้ไฟฟ้าประเภทกิจการขนาดกลางซึ่งในช่วงภาวะวิกฤติทางเศรษฐกิจในช่วงปี 2541-2542 มีจำนวนผู้ใช้ไฟฟ้าประเภทนี้ลดลง

สุมิตรา อมรรพภัคตร์ (2542)

การเปรียบเทียบวิธีการพยากรณ์ความต้องการใช้ไฟฟ้าในส่วนภูมิภาค มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสม สำหรับการพยากรณ์ข้อมูลปริมาณความต้องการใช้ไฟฟ้าในส่วนภูมิภาค และต้องการหาตัวแบบพยากรณ์ที่ให้ค่าพยากรณ์ที่มีความคลาดเคลื่อนต่ำ โดยใช้วิธีการพยากรณ์ 4 วิธี ได้แก่ 1) วิธีบอก-เจนกินส์ 2) วิธีการปรับให้เรียบแบบเลขชี้กำลัง 3) วิธีการแยก

องค์ประกอบ และ 4) วิธีการวิเคราะห์การถดถอย และนำไปเปรียบเทียบกับวิธีการพยากรณ์ของการไฟฟ้าส่วนภูมิภาค(กฟภ.) ที่ใช้อยู่ในปัจจุบัน ในการเปรียบเทียบจะใช้ค่าเฉลี่ยของเปอร์เซ็นต์ค่าความคลาดเคลื่อนสัมบูรณ์ เป็นเกณฑ์การคัดเลือกตัวแบบพยากรณ์ โดยข้อมูลที่ใช้ในการวิเคราะห์ครั้งนี้เป็นข้อมูลทศนิยม เก็บรวบรวมจากแผนกสถิติการใช้ไฟฟ้า กองเศรษฐกิจพลังไฟฟ้า กฟภ. สำหรับตัวแบบอนุกรมเวลาเป็นข้อมูลรายเดือน ในช่วงปีงบประมาณ 2536 – 2542 ส่วนตัวแบบการถดถอยเป็นข้อมูลรายปี ในช่วงปีงบประมาณ 2524 -2542

จากการศึกษาพบว่า การเปรียบเทียบข้อมูลที่วิเคราะห์ทั้ง 4 วิธี พบว่าวิธีการพยากรณ์โดยใช้บอซซ์-เจนกินส์ เหมาะสมกว่าวิธีอื่นๆ ที่นำมาศึกษาเปรียบเทียบ สำหรับข้อมูลปริมาณความต้องการใช้ไฟฟ้าทุกประเภท ซึ่งผลการพยากรณ์ในอนาคต คาดว่า ปริมาณความต้องการใช้ไฟฟ้า ในแต่ละประเภทบ้านอยู่อาศัย กิจการขนาดต่างๆ และกิจการเฉพาะอย่าง ของทุกภาค ในปี 2543 – 2545 มีแนวโน้มเพิ่มขึ้นทุกปี ส่วนปริมาณความต้องการใช้ไฟฟ้าในประเภทสูบน้ำเพื่อการเกษตรและไฟฟ้าชั่วคราวของทุกภาค ในปี 2543 -2545 จะมีแนวโน้มลดลงทุกปี สำหรับปริมาณความต้องการใช้ไฟฟ้าในประเภทส่วนราชการและองค์กรที่ไม่แสวงหากำไร ของทุกภาคยกเว้นภาคเหนือ ในปี 2543 – 2545 จะมีแนวโน้มเพิ่มขึ้นทุกปี

ธิดารัตน์ จันทวี (2539)

การพยากรณ์ความต้องการใช้ไฟฟ้าเพื่อการผลิตไฟฟ้าระยะสั้น โดยมีวัตถุประสงค์เพื่อหาวิธีการที่เหมาะสมสำหรับการพยากรณ์ข้อมูลปริมาณความต้องการใช้ไฟฟ้า เพื่อการวางแผนผลิตไฟฟ้าระยะสั้น ตลอดจนความต้องการหาตัวแบบพยากรณ์ที่ให้ค่าพยากรณ์ใกล้เคียงกับข้อมูลจริง และศึกษาเปรียบเทียบวิธีการพยากรณ์ 3 วิธี ว่าวิธีใดมีความเหมาะสมกับข้อมูลมากที่สุด โดยพิจารณาจากค่าเฉลี่ยค่าสุดของเปอร์เซ็นต์ค่าความคลาดเคลื่อนสัมบูรณ์ โดยนำเทคนิคและทฤษฎีสถิติมาช่วยในการวิเคราะห์ข้อมูล ประกอบด้วยวิธีการพยากรณ์ของบอซซ์และเจนกินส์ เทคนิคการปรับให้เรียบ และวิธีการพยากรณ์ของแผนการผลิตไฟฟ้าระยะสั้น การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย(กฟผ.) ลักษณะข้อมูลอนุกรมเวลารายเดือนของปีงบประมาณ 2533 – 2538 จำแนกตามประเภทข้อมูลทั้ง 13 ประเภท

ผลการศึกษาพบว่า จากการเปรียบเทียบข้อมูลที่วิเคราะห์ทั้ง 3 วิธี ได้ว่า วิธีการพยากรณ์ด้วยวิธีของบอซซ์และเจนกินส์ มีความเหมาะสมกว่าวิธีอื่นๆ ที่นำมาศึกษาเปรียบเทียบสำหรับข้อมูลปริมาณการใช้ไฟฟ้าทุกประเภท

Surapan Jungaroen (1970)

วิทยานิพนธ์ฉบับนี้กล่าวถึงศึกษาเกี่ยวกับอุปสงค์พลังงานไฟฟ้าของประเทศไทย โดยมีวัตถุประสงค์ที่จะหาความสัมพันธ์ระหว่างอุปสงค์พลังงานไฟฟ้าประเภทที่อยู่อาศัย และประเภทของอุตสาหกรรม กับปัจจัยทางเศรษฐกิจที่เป็นตัวกำหนดอุปสงค์พลังงานไฟฟ้า และการหาอุปสงค์พลังงานไฟฟ้าของแต่ละภาคในประเทศไทย

จากการศึกษาด้วยวิธี กำลังสองน้อยที่สุด (ordinary least square method) ใช้ข้อมูลอนุกรมเวลา 12 ปี ได้ผลการศึกษารูปได้ว่า ราคาไฟฟ้าของบ้านอยู่อาศัยและการใช้จ่ายต่อครัวเรือน มีผลต่ออุปสงค์พลังงานไฟฟ้าประเภทที่อยู่อาศัย ในขณะที่ราคาไฟฟ้าของภาคภาคอุตสาหกรรม และGDPของอุตสาหกรรม มีผลต่ออุปสงค์พลังงานไฟฟ้าประเภทภาคภาคอุตสาหกรรม ส่วนการหาอุปสงค์พลังงานไฟฟ้าในแต่ละภาค พบว่า ภาคเหนือ ภาคตะวันออกเฉียงเหนือ และเขต กปน. ตัวแปรที่มีนัยสำคัญทางสถิติพอที่จะอธิบายอุปสงค์พลังงานไฟฟ้าได้คือ ราคาไฟฟ้า และรายได้สำหรับภาคใต้ ตัวแปรที่มีความสัมพันธ์ คือ รายได้และจำนวนประชากร และภาคกลางตัวแปรที่มีความสัมพันธ์ได้แก่ ค่าไฟฟ้า และจำนวนประชากร ทั้งนี้ผู้วิจัยได้เสนอแนะไว้ตอนท้ายว่า ควรมีการศึกษาเพิ่มเติมถึงอุปสงค์พลังงานไฟฟ้าต่อหัว (per capita electricity consumption) เพื่ออธิบายพฤติกรรมของผู้บริโภคด้วย

อธยา เขียววฒนิกิ (2522)

วิทยานิพนธ์ฉบับนี้กล่าวถึงการศึกษาเรื่องอุปสงค์พลังงานไฟฟ้าของอุตสาหกรรมในเขตจำหน่ายของการไฟฟ้านครหลวง โดยมีวัตถุประสงค์เพื่อศึกษาอุปสงค์พลังงานไฟฟ้าของภาคอุตสาหกรรม ตามลักษณะการใช้งาน ทั้งนี้ผู้วิจัยได้แบ่งอุตสาหกรรมออกเป็น 3 ลักษณะคือ อุตสาหกรรม ที่ใช้ไฟฟ้าในการก่อให้เกิดพลังงานความร้อน พลังงานกล และแสงสว่าง โดยใช้อุตสาหกรรมเหล็ก อุตสาหกรรมสิ่งทอ และธุรกิจโรงแรม เป็นตัวแทนอุปสงค์พลังงานไฟฟ้าแต่ละลักษณะตามลำดับ

จากการวิเคราะห์ด้วยวิธีกำลังสองน้อยที่สุด ใช้ข้อมูลอนุกรมเวลาระหว่าง ปี 1966-1977 ได้ผลสรุปว่า อัตราค่าไฟฟ้าและปริมาณผลผลิต ของอุตสาหกรรมมีนัยสำคัญทางสถิติในการกำหนดอุปสงค์พลังงานไฟฟ้าเพื่อก่อให้เกิดพลังงานความร้อน ส่วนอัตราค่าไฟฟ้า ราคาขายปลีกน้ำมันเตา และปริมาณผลผลิตภาคอุตสาหกรรมมีนัยสำคัญทางสถิติ ในการกำหนดอุปสงค์พลังงาน

ไฟฟ้าเพื่อเป็นพลังงานกล และอัตราค่าไฟฟ้า ราคาขายปลีกน้ำมันดีเซล และจำนวนนักท่องเที่ยวมี
 นัยสำคัญทางสถิติในการกำหนดอุปสงค์พลังงานไฟฟ้าเพื่อแสงสว่าง

वरङ्गल्लिथि नामराशुर् (2530)

วิทยานิพนธ์ฉบับนี้กล่าวถึงการศึกษาเรื่องการพยากรณ์ความต้องการพลังงานไฟฟ้าในพื้นที่
 รับผิดชอบของการไฟฟ้าส่วนภูมิภาค โดยมีวัตถุประสงค์เพื่อการพยากรณ์ (forecasting) จำนวน
 ผู้ใช้ไฟฟ้าและพลังงานไฟฟ้าตามการแบ่งพื้นที่ของการไฟฟ้าส่วนภูมิภาคออกเป็น เขต ภาค และ
 พื้นที่รวมทั้งหมด

การวิเคราะห์ได้ทำใน 2 ลักษณะคือ การวิเคราะห์แบบแยกข้อมูล ซึ่งทำการวิเคราะห์การใช้
 พลังงานไฟฟ้าในแต่ละเขตของการไฟฟ้า กับจำนวนผู้ใช้ไฟฟ้า แล้วนำมารวมเป็นแต่ละภาคการใช้
 ไฟฟ้า และการคาดคะเนแบบรวมข้อมูล ซึ่งเป็นการคาดคะเนการใช้พลังงานไฟฟ้าและมีจำนวนผู้ใช้
 ไฟฟ้าจากพื้นที่รวมทั้งหมดเพื่อดูว่าการคาดคะเนทั้ง 2 ลักษณะมีความเหมาะสมหรือถูกต้องต่างกัน
 อย่างไร ในการวิเคราะห์ใช้ข้อมูลของการไฟฟ้าส่วนภูมิภาค 15 ปี (พ.ศ. 2515-2529) และการ
 วิเคราะห์ใช้วิธีการเขียนแผนภาพการกระจาย (scatter diagram) เพื่อดูลักษณะการเปลี่ยนแปลงของ
 พลังงานไฟฟ้า แล้วเลือกใช้วิธีการวิเคราะห์ความถดถอย (regression analysis) ที่เหมาะสม
 สอดคล้องกับแผนภาพการกระจายให้มากที่สุด มาทำการคาดคะเนทำนายผลจากข้อมูล ซึ่งผู้วิจัยได้
 ใช้วิธีการวิเคราะห์ความถดถอย กับข้อมูลโดยตรง

ผลการศึกษาสรุปได้ว่า การคาดคะเนจำนวนผู้ใช้ไฟฟ้าและพลังงานไฟฟ้า แบบรวมข้อมูล
 กับผลรวมของการคาดคะเนแบบแยกข้อมูล ตามการแบ่งเป็นเขต แบ่งเป็นภาค และพื้นที่รวมทั้ง
 ทั้งหมดปรากฏว่าผลจากการคาดคะเน ทั้ง 2 ลักษณะให้ผลต่างกันน้อยมาก สามารถยอมรับได้ทั้ง 2
 กรณี และการคาดคะเนจำนวนผู้ใช้ไฟฟ้าและพลังงานไฟฟ้า มีอัตราการเปลี่ยนแปลงตามสมการ
 พาราโบลา นอกจากนี้ในตอนท้ายผู้วิจัยได้เสนอแนะว่า การคาดคะเนความต้องการไฟฟ้าในพื้นที่
 ของ กฟภ. ต่อไป น่าจะทำได้เป็น model เล็ก ๆ ในแต่ละพื้นที่ โดยจัดทำแบบสำรวจสอบถามข้อมูล
 อุปกรณ์การใช้ไฟฟ้าเพื่อความสมบูรณ์ของการคาดคะเนต่อไป

พิษณุ ค้วงสุวรรณ(2535)

วิทยานิพนธ์ฉบับนี้กล่าวถึงการศึกษาเรื่องการคาดคะเนโหลดของระบบไฟฟ้าในประเทศไทย โดยอาศัยตัวเลขจากมูลค่าผลิตภัณฑ์แต่ละเขตรวม โดยมีวัตถุประสงค์ เพื่อทำการคาดคะเน (forecasting) จำนวนผู้ใช้ไฟฟ้า พลังงานไฟฟ้าตามประเภทของการใช้ไฟฟ้า พลังไฟฟ้าสูงสุด และ พลังงานไฟฟ้าทั้งหมด ผู้วิจัยได้ทำการศึกษาโดยใช้วิธีการวิเคราะห์ความถดถอย และแยกการวิเคราะห์ออกเป็น 2 ลักษณะ คือ การคาดคะเนแบบแยกข้อมูลใช้วิธีการถดถอยแบบเส้นตรงเชิงเดี่ยว (simple linear regression) กับข้อมูลการใช้ไฟฟ้า ทั้งจำนวนผู้ใช้ไฟฟ้าและพลังงานไฟฟ้า ของแต่ละประเภท และการคาดคะเนแบบรวมข้อมูล ซึ่งใช้วิธีการวิเคราะห์การถดถอยแบบเส้นตรงเชิงเดี่ยว เช่นเดียวกัน กับผลรวมของข้อมูลทั้งหมด ทั้งจำนวนผู้ใช้ไฟฟ้า และพลังงานไฟฟ้าทั้งหมด 12 เขต หรือ 4 ภาค โดยการวิเคราะห์ทั้งแบบแยกและรวมข้อมูล ได้นำข้อมูลจำนวนผู้ใช้ไฟฟ้า และ พลังงานไฟฟ้า แล้วมาหาความสัมพันธ์กับมูลค่าผลิตภัณฑ์แต่ละเขตรวม (GRP: Gross Regional Product) หรือมูลค่าผลิตภัณฑ์ภายในประเทศรวม (GDP: Gross Domestic Product)

ผลการศึกษาสรุปได้ว่า การคาดคะเนจำนวนผู้ใช้ไฟฟ้า และพลังงานไฟฟ้าทั้งแบบรวมและแบบแยกข้อมูล จะให้ผลการคาดคะเนที่ไม่แตกต่างกัน และจากการวิเคราะห์การคาดคะเนโหลดของระบบ ไฟฟ้าในประเทศไทย พบว่าการคาดคะเนความต้องการใช้ไฟฟ้าของการไฟฟ้าส่วนภูมิภาคและการไฟฟ้านครหลวง จะใช้วิธีการของ correlation คือใช้ความสัมพันธ์ของมูลค่าผลิตภัณฑ์แต่ละเขตรวมหรือ มูลค่าผลิตภัณฑ์ในประเทศรวมมาพิจารณา และใช้วิธีการวิเคราะห์ความถดถอย กับข้อมูลของการไฟฟ้าส่วนภูมิภาค และการไฟฟ้านครหลวง ส่วนผู้ใช้ไฟฟ้าโดยตรงจากการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย จะดูความเป็นไปได้ว่า ผู้ใช้ไฟฟ้ามีแผนการใช้ไฟต่อไปเป็นจำนวนเท่าใด นอกจากนี้ในตอนท้าย ผู้วิจัยได้เสนอให้มีการจัดทำแบบสำรวจและการสัมภาษณ์ผู้ใช้ไฟฟ้าเพื่อความสมบูรณ์ของการคาดคะเนเช่นเดียวกัน

อรสา ศรีรูป(2540)

วิทยานิพนธ์ฉบับนี้กล่าวถึงการศึกษาเรื่อง การคาดคะเนปริมาณความต้องการใช้ กระแสไฟฟ้าประเภทบ้านอยู่อาศัยในเขตนครหลวง โดยมีวัตถุประสงค์เพื่อศึกษาถึงปัจจัยทางเศรษฐกิจที่มีผลกระทบต่อปริมาณการใช้ไฟฟ้าประเภทบ้านอยู่อาศัยในเขตนครหลวง ใช้ข้อมูลจากการออกแบบสอบถามมาทำการวิเคราะห์ข้อมูลเชิงปริมาณ โดยใช้แบบจำลองสมการถดถอยพหุคูณ

เพื่อความสัมพันธ์ของปัจจัยต่าง ๆ กับปริมาณการใช้ไฟฟ้า นอกจากนี้ได้ศึกษารูปแบบสมการที่เหมาะสมในการพยากรณ์ปริมาณการใช้ไฟฟ้าในระยะสั้น โดยใช้ข้อมูลสถิติของปริมาณการใช้ไฟฟ้าประเภทบ้านอยู่อาศัยในเขตนครหลวง มาทำการพยากรณ์โดยใช้สถิติการพยากรณ์ แบบบอซซ์-เจนกินส์

ผลการศึกษาพบว่า จำนวนสมาชิกในครัวเรือนเป็นปัจจัยที่มีผลต่อชั่วโมงการใช้งานเครื่องใช้ไฟฟ้าเกือบทุกประเภท รายได้ของครัวเรือนเป็นปัจจัยที่มีผลต่อชั่วโมงการใช้งานของเครื่องใช้ไฟฟ้าที่มีราคาสูง เช่น เครื่องปรับอากาศ เครื่องอบผ้า และเตาไมโครเวฟ และพลังงานไฟฟ้าเฉลี่ยของเครื่องใช้ไฟฟ้าใน ครัวเรือนเป็นปัจจัยที่มีผลต่อชั่วโมงการใช้ของเครื่องใช้ไฟฟ้า ได้แก่ โทรทัศน์สี เครื่องปรับอากาศ และเตารีด นอกจากนี้ผู้วิจัยได้เสนอแนะให้มีการศึกษาการใช้ไฟฟ้าประเภทบ้านอยู่อาศัยในส่วนของภูมิภาค เพิ่มเติมเพื่อศึกษาพฤติกรรมของการใช้ไฟฟ้า เนื่องจากมีความแตกต่างของสภาพสังคมและเศรษฐกิจ ระหว่างในเขต กรุงเทพมหานคร และในส่วนภูมิภาค