บทที่ 3

แนวความคิดและระเบียบวิธีวิจัย

3.1 แนวความคิดที่ใช้ในการศึกษา

การศึกษาการประมาณค่าความผันผวนของผลตอบแทนของราคาน้ำมันดิบ ถ่านหินและก๊าซธรรมชาติ ใช้แนวคิดและทฤษฎีเกี่ยวกับการวิเคราะห์อนุกรมเวลา การทดสอบความนิ่งของข้อมูล การทดสอบ unit root แบบจำลอง Autoregressive Conditional Heteroscadasticity (ARCH) แบบจำลอง Generalized Conditional Heteroscadasticity (GARCH) แบบจำลอง GARCH-in-mean และแบบจำลอง Exponential GARCH ดังต่อไปนี้

3.1.1 การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

การวิเคราะห์อนุกรมเวลาเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านไป ลักษณะของการเปลี่ยนแปลงอาจมีหรือไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใดและสามารถพยากรณ์การเปลี่ยนแปลงข้อมูลในอนาคตได้ การวิเคราะห์ข้อมูลอนุกรมเวลานี้จะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน (ศิริลักษณ์ เล็กสมบูรณ์, 2531)

3.1.2 การทดสอบความนิ่งของข้อมูล (Stationary) และการทดสอบ Unit Root

ในการวิเคราะห์ข้อมูลทางเศรษฐศาสตร์ ซึ่งข้อมูลที่ใช้เป็นข้อมูลแบบอนุกรมเวลา (Time series data) มีความจำเป็นที่จะต้องมีการทดสอบว่า ข้อมูลที่เราใช้นั้นมีความนิ่ง (Stationary) หรือไม่ เพราะหากเราละเลยที่จะทดสอบข้อมูล แล้วข้อมูลเกิดมีความไม่นิ่ง (Nonstationary) จะทำให้เวลาที่หาสมการถดถอยระหว่างตัวแปรอนุกรมเวลาสองตัวแปรออกมา จะได้ค่า R^2 ที่สูงมาก และค่าสถิติ t จะมีนัยสำคัญ ทั้งที่ไม่มีความสัมพันธ์ของตัวแปรทั้งสองในทางเศรษฐศาสตร์เลย

(Enders, 1995: 216 and Gujarati, 1995: 709 อ้างถึงในทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542)

เหตุที่ทำให้ได้ค่า R^2 สูงเช่นนี้เป็นเพราะอนุกรมเวลาได้รับอิทธิพลจากแนวโน้ม (Trend) ไม่ใช่เนื่องจากความสัมพันธ์ที่แท้จริงระหว่างตัวแปรอนุกรมเวลาทั้งสองตัวแปร ส่วนกรณีที่ค่าสถิติ t มีนัยสำคัญเป็นเพราะอนุกรมเวลาทั้งสองมีแนวโน้มที่แข็งแรงมาก (Strong Trend) โดยสรุปแล้วการละเลยการทดสอบความนิ่งของข้อมูลอาจนำไปสู่การลงความเห็นที่ผิดพลาดได้ในที่สุด

การทดสอบว่าข้อมูลที่นำมาศึกษามีความนิ่งหรือไม่ สามารถทำได้โดยการทดสอบ Unit Root ซึ่งทำได้โดยใช้การทดสอบ DF (Dickey-Fuller Test) ซึ่งเสนอโดย Dickey และ Fuller ในปี 1981 และวิธีการทดสอบ ADF (Augmented Dickey-Fuller Test) ซึ่งเสนอโดย Said และ Dickey ในปี 1984

ข้อมูลที่มีลักษณะนิ่ง (Stationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) เท่ากันตลอดระยะเวลาที่ศึกษา

ส่วนข้อมูลที่มีลักษณะไม่นิ่ง (Nonstationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) ไม่เท่ากันตลอดระยะเวลาที่ศึกษา

ทั้งนี้การวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลาส่วนมากจะพบปัญหาความไม่นิ่งของข้อมูล ซึ่งสามารถแก้ไขได้ด้วยการทำให้ข้อมูลมีความนิ่งเสียก่อน โดยอาจใช้วิธีการหาผลต่าง (Difference) ของข้อมูล การแปลงให้อยู่ในรูป Logarithm หรือการทดสอบหาความสัมพันธ์ของตัวแปรในระยะยาว (Cointegration) เป็นต้น

วิธีการทดสอบ Unit Root (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542) สามารถทดสอบได้โดยใช้การทดสอบ DF (Dickey-Fuller (DF) Test) ซึ่งเสนอโดย Dickey และ Fuller ในปี 1981 หรือใช้การทดสอบ ADF (Augmented Dickey-Fuller (ADF) Test) ซึ่งเสนอโดย Said และ Dickey ในปี 1984 สมมุติฐานว่าง (null hypothesis) ของการทดสอบ DF (DF test) คือ $H_0 : \rho = 1$ จากสมการ

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{3.1}$$

ซึ่งเรียกว่าการทดสอบ unit root โดยถ้า $|\rho| < 1$ X_t จะมีลักษณะนิ่ง (stationary) และถ้า $\rho = 1$ X_t จะมีลักษณะไม่นิ่ง (nonstationary) อย่างไรก็ตามการทดสอบนี้สามารถทำได้อีกทางหนึ่งซึ่งเหมือนกับสมการ (3.1) กล่าวคือ

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \tag{3.2}$$

ซึ่งก็คือ $X_t = (1+\theta)X_{t-1} + \varepsilon_t$ ซึ่งคือสมการที่ (1) นั่นเอง โดยที่ $\rho = (1+\theta)$ ถ้า θ ในสมการ (2) มีค่าเป็นลบ จะได้ว่า ρ ในสมการ (1) จะมีค่าน้อยกว่า 1 ดังนั้นสามารถจะสรุปได้ว่า การปฏิเสธ $H_0 : \theta = 0$ ซึ่งเป็นการยอมรับ $H_a : \theta < 0$ หมายความว่า $\rho < 1$ และ X_t มี integration of order zero นั่นคือ X_t มีลักษณะนิ่ง (stationary) และถ้าเราไม่สามารถปฏิเสธ $H_0 : \theta = 0$ ได้ ก็จะหมายความว่า X_t มีลักษณะไม่นิ่ง (nonstationary)

ถ้า X_t เป็นแนวเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (random walk with drift) สามารถจะเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t \tag{3.3}$$

และถ้า X_t เป็นแนวเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (random walk with drift) และมีแนวโน้มตามเวลาเชิงเส้น (linear time trend) สามารถเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t \tag{3.4}$$

โดยที่ t = เวลา ซึ่งก็จะทำการทดสอบ $H_0 : \theta = 0$ โดยมี $H_a : \theta < 0$ เช่นเดียวกับที่กล่าวมาข้างต้น โดยสรุปแล้ว Dickey and Fuller ได้พิจารณาสมการถดถอย 3 รูปแบบที่แตกต่างกันในการทดสอบว่ามี unit root หรือไม่ ซึ่ง 3 สมการดังกล่าว ได้แก่

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t$$
$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t$$
$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t$$

โดยตัวพารามิเตอร์ที่อยู่ในความสนใจในทุกสมการ คือ θ นั่นคือ ถ้า $\theta = 0$; X_t จะมี unit root โดยการเปรียบเทียบค่าสถิติ t (t–statistic) ที่คำนวณได้กับค่าที่เหมาะสมที่อยู่ในตาราง Dickey-Fuller (Dickey–Fuller tables) หรือกับ ค่าวิกฤติ MacKinnon (MacKinnon critical values) (Enders, 1995: 221 and Gujarati, 1995: 769 อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542)

อย่างไรก็ตามค่าวิกฤติ (Critical values) จะไม่เปลี่ยนแปลง ถ้าสมการ (3.2), (3.3), (3.4) ถูกแทนที่โดยกระบวนการเชิงอัตถดถอย (autoregressive processes)

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{p} \phi_i \Delta X_{t-i} + \varepsilon_t \quad (3.5)$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{p} \phi_i \Delta X_{t-i} + \varepsilon_t \quad (3.6)$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \sum_{i=1}^{p} \phi_i \Delta X_{t-i} + \varepsilon_t \quad (3.7)$$

จำนวนของ lagged difference terms ที่จะนำเข้ามารวมในสมการนั้นจะต้องมีมากพอที่จะทำให้พจน์ค่าความคลาดเคลื่อน (error terms) มีลักษณะเป็น serially independent และเมื่อนำเอาการทดสอบ DF (Dickey – Fuller (DF) test) มาใช้กับสมการ (3.5) – (3.7) เราจะเรียกว่าการทดสอบ ADF (augmented Dickey – Fuller (ADF) test) ค่าสถิติทดสอบ ADF (ADF test statistic) มีการแจกแจงเชิงเส้นกำกับ (asymptotic distribution) เหมือนกับสถิติ DF (DF statistic) ดังนั้นก็สามารถใช้ค่าวิกฤติ (critical values) แบบเดียวกัน (Enders, 1995: 221 and Gujarati, 1995: 720)

3.1.3 การเลือกแบบจำลองที่เหมาะสมจากการทดสอบ Unit Root โดยการทดสอบสัมประสิทธิ์การถดถอย (Deterministic Regressors)

เป็นการทดสอบว่า แบบจำลองใดเป็นแบบจำลองที่มีความเหมาะสมที่สุดระหว่างกรณีของแบบจำลองที่ไม่มีค่าคงที่และแนวโน้มเวลา(None) แบบจำลองที่มีค่าคงที่ (Intercept) และแบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา (Trend and Intercept) โดยการทดสอบความมีนัยสำคัญทางสถิติของสัมประสิทธิ์ของตัวถดถอย (ค่าคงที่หรือค่าแนวโน้มเวลา) โดยมีขั้นตอนการทดสอบดังนี้

ขั้นตอนที่ 1 เริ่มการทดสอบจากแบบจำลองกรณีที่มีทั้งค่าคงที่และแนวโน้มเวลาตามสมการ (3.8)

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum \beta_i \Delta y_{t-1} + \varepsilon_t \quad (3.8)$$

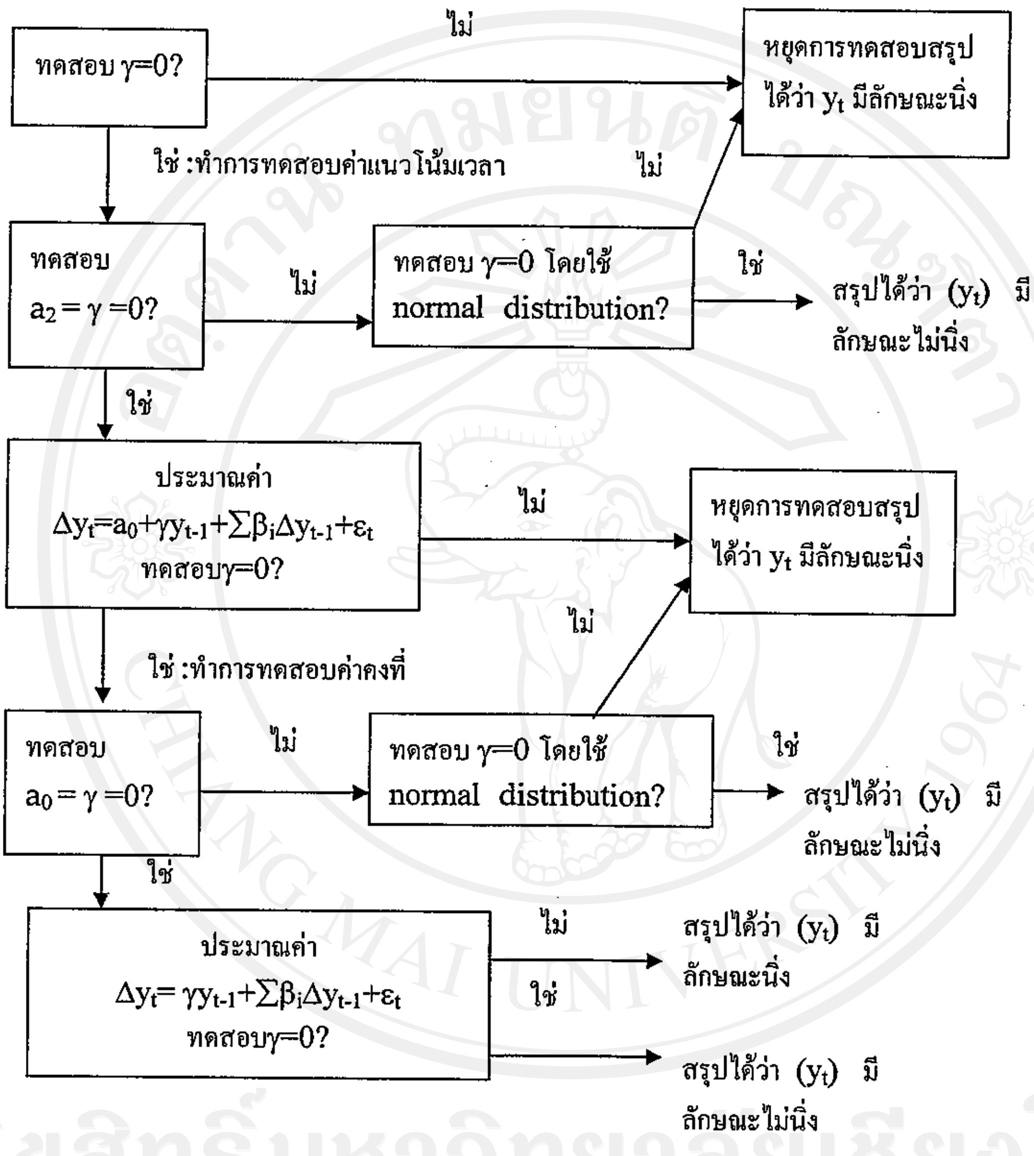
ทำการทดสอบสมมติฐานว่าง H_0: $\gamma=0$ โดยใช้ τ_τ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่าง แสดงว่า ข้อมูล y_t มีลักษณะนิ่งแล้ว และเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา

ขั้นตอนที่ 2 ถ้าเกิดการยอมรับสมมติฐานว่างในขั้นตอนที่ 1 แสดงว่าในแบบจำลองดังกล่าวมีตัวถดถอยที่ไม่จำเป็นอยู่ในสมการ ซึ่งอาจทำให้อำนาจการทดสอบของสมการลดลง ดังนั้นจึงต้องมีการทดสอบความมีนัยสำคัญทางสถิติของค่าแนวโน้ม (a_2t)ที่อยู่ในสมการ โดยการทดสอบสมมติฐานว่าง H_0: $a_2=\gamma=0$โดยใช้ ϕ_3 statistic ถ้าค่าสัมประสิทธิ์ของค่าแนวโน้มไม่มีนัยสำคัญทางสถติให้ข้ามไปขั้นตอนที่ 3 อย่างไรก็ตามถ้าค่าสัมประสิทธิ์ของค่าแนวโน้มมีนัยสำคัญทางสถติให้ทำการทดสอบความไม่นิ่งของข้อมูลอีกครั้งโดยใช้ การแจกแจงแบบปกติมาตรฐาน(standardized normal distribution) ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่า ข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลาแต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล y_t มีลักษณะไม่นิ่ง

ขั้นตอนที่ 3 ทำการประมาณค่าแบบจำลองตามสมการ(3.8) ที่ปราศจากค่าแนวโน้มเวลา และทดสอบ unit root โดยใช้τ_μ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่าข้อมูล y_tมีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่ปราศจากค่าแนวโน้มเวลา แต่ถ้าเกิดการยอมรับสมมติฐานว่างให้ทำการทดสอบความมีนัยสำคัญทางสถิติของค่าคงที่ โดยมีสมมติฐานว่าง H_0: $a_0=\gamma=0$ โดยใช้ ϕ_1 statistic ถ้าค่าคงที่ไม่มีนัยสำคัญให้ข้ามไปขั้นตอนที่ 4 อย่างไรก็ตามถ้าค่าคงที่มีนัยสำคัญทางสถติ ให้ทำการทดสอบความไม่นิ่งของข้อมูลอีกครั้งโดยใช้ การแจกแจงแบบปกติมาตรฐาน (standardized normal distribution) ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่า ข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลาแต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล y_t มีลักษณะไม่นิ่ง

ขั้นตอนที่ 4 ทำการประมาณค่าแบบจำลองตามสมการ (3.8) ที่ปราศจากค่าแนวโน้มเวลา และค่าคงที่และทดสอบ unit root โดยใช้ τ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่าข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่ปราศจากค่าแนวโน้มเวลาและค่าคงที่แต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล y_t มีลักษณะไม่นิ่ง

รูปที่ 3.1 ขั้นตอนการเลือกแบบจำลองที่เหมาะสม



ที่มา: Enders (1995)

3.1.4 แบบจำลอง Autoregressive Conditional Heteroscedasticity (ARCH)

ในการวิเคราะห์อนุกรมเวลาส่วนใหญ่แล้วจะมีการกำหนด stochastic variable ให้มีความแปรปรวนคงที่ (homoscedastic) ซึ่งในการประยุกต์ใช้กับบางข้อมูลนั้นค่าความ

แปรปรวนของค่าความคลาดเคลื่อน (error term) จะไม่ใช่ฟังก์ชั่นของตัวแปรอิสระแต่มีค่าเปลี่ยนแปลงไปตามช่วงเวลาขึ้นอยู่กับขนาดของความคลาดเคลื่อนที่เกิดขึ้นในอดีต และในบางงานศึกษา เช่น แบบจำลองของเงินเฟ้อ อัตราดอกเบี้ยหรือผลตอบแทนจากตลาดหลักทรัพย์ ในบางคาบเวลาจะมีค่าความผันผวน (volatility) สูง (และความคลาดเคลื่อนขนาดใหญ่) ตามด้วยคาบเวลาที่มีค่าความผันผวน (volatility) ต่ำ (และมีค่าความคลาดเคลื่อนที่มีขนาดเล็ก) สรุปได้ว่าค่าความแปรปรวนของค่าความคลาดเคลื่อนจากการถดถอยจะขึ้นอยู่กับค่าความผันผวน (volatility) ของความคลาดเคลื่อนในอดีตที่ผ่านมา (สธนพล วิเชียรรัตนพันธ์, 2547: 29 อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2547: 641)

ความเป็นไปได้ในการหาค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาไปพร้อมกันนั้น ในขั้นต้นการพยากรณ์อย่างมีเงื่อนไขจะมีความแม่นยำเหนือกว่าการพยากรณ์อย่างไม่มีเงื่อนไขมาก ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) ซึ่งแสดงดังสมการ

$$X_t = a_0 + a_1 X_{t-1} + \varepsilon_t \tag{3.9}$$

และการพยากรณ์อย่างมีเงื่อนไขของ X_{t+1} ดังนี้คือ

$$E_t X_{t+1} = a_0 + a_1 X_t \tag{3.10}$$

ถ้าเราใช้ค่าเฉลี่ยอย่างมีเงื่อนไขในการพยากรณ์ X_{t+1} ค่าความคลาดเคลื่อนของความแปรปรวนอย่างมีเงื่อนไขที่พยากรณ์ได้ดังนี้คือ

$$E_t\left[\left(X_{t+1} - a_0 - a_1 X_t\right)^2\right] = E_t \varepsilon_{t+1}^2 = \sigma^2 \tag{3.11}$$

ถ้าเปลี่ยนไปใช้การพยากรณ์อย่างไม่มีเงื่อนไขแล้ว ผลที่จะใช้เป็นค่าเฉลี่ยในช่วง Long-Run ของลำดับ $\{X_t\}$ ซึ่งเท่ากับ $\dfrac{a_0}{(1-a_1)}$ จะได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไขตามสมการดังนี้

$$E\left\{\left(X_{t-1} - \frac{a_0}{1-a_1}\right)^2\right\} = E\left[\left(\varepsilon_{t+1} + a_1\varepsilon_1 + a^2{}_1\varepsilon_{t-1} + a_1^3\varepsilon_{t-2} + \ldots\right)^2\right] \tag{3.12}$$

เนื่องจาก $\frac{1}{(1-a_1)^2} > 1$ เพราะฉะนั้นความแปรปรวน (variance) จากการพยากรณ์แบบไม่มีเงื่อนไข (unconditional variance) จึงมีค่าสูงกว่าความแปรปรวน ของการพยากรณ์แบบมีเงื่อนไข ในลักษณะเดียวกันถ้าความแปรปรวนของ $\{\varepsilon_t\}$ ไม่เป็นค่าคงที่ จะสามารถประมาณค่าแนวโน้มของการเปลี่ยนแปลงความแปรปรวนโดยใช้ ARMA Model ความแปรปรวนแบบมีเงื่อนไขของ X_{t+1} คือ

$$Var(X_{t+1}|X_t) = E\left[(X_{t+1} - a_0 - a_1 X_t)^2\right] = E_t \varepsilon_{t+1}^2 \tag{3.13}$$

และจากที่ให้ $E_t \varepsilon_{t-1}^2 = \sigma_{t+1}^2$ จึงแสดงว่าความแปรปรวนอย่างมีเงื่อนไขไม่ใช่ค่าคงที่และจะได้แบบจำลองในการประมาณค่าส่วนที่เหลือ (residuals) ออกมาดังสมการ

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \ldots.. + \alpha_q \hat{\varepsilon}_{t-q}^2 + V_t \tag{3.14}$$

เมื่อ V_t = white noise process

ถ้าค่าของ $\alpha_1, \alpha_2, \ldots\ldots, \alpha_q$ มีค่าเท่ากับศูนย์ ความแปรปรวนที่ประมาณค่ามาได้ (estimated variance) จะมีค่าคงที่หรือคงตัว (constant variance) α_0 อีกนัยหนึ่ง คือ ค่าความแปรปรวนอย่างมีเงื่อนไขของ Xt จะมีการเปลี่ยนแปลงสอดคล้องกับในสมการ (14) และค่าพยากรณ์สามารถเขียนได้ดังนี้

$$E\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-1}^2 + \ldots.. + \alpha_q \hat{\varepsilon}_{t+1-q}^2 \tag{3.15}$$

สมการที่มีลักษณะเช่นสมการ (3.14) เราเรียกว่า an autoregressive conditional heteroscedastic (ARCH) model และสมการ (3.15) เป็น ARCH(q) โดยค่า $E_t \hat{\varepsilon}_{t+1}^2$ หรือ σ_{t+1}^2 จะประกอบด้วย 2 องค์ประกอบ คือค่าคงที่และความผันผวน (volatility) ในคาบเวลาที่ผ่านมา ซึ่งเขียนได้เป็นส่วนเหลือกำลังสองของคาบในอดีต (ARCH term) ส่วนค่าสัมประสิทธิ์ ($\alpha_0, \alpha_1, \ldots\ldots, \alpha_q$) สามารถหาค่าได้โดยใช้วิธี Maximum likelihood

3.1.5 แบบจำลอง Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

แบบจำลอง ARCH ของ Engle, Robert F. ได้มีการพัฒนาต่อโดย Bollerslev ในปี 1986 ด้วยการให้ความแปรปรวนแบบมีเงื่อนไข (conditional variance) มีลักษณะเป็น ARMA process โดยที่ให้ error process มีลักษณะดังนี้

$$\varepsilon_t = V_t\sqrt{C_t}$$

$$\text{และ} \quad C_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i C_{t-i} \tag{3.16}$$

เมื่อ $\{V_t\}$ คือ white noise process ที่เป็นค่าอิสระจากเหตุการณ์ในอดีต (ε_{t-1}) ค่าเฉลี่ยเคลื่อนที่อย่างมีเงื่อนไขและไม่มีเงื่อนไขของ ε_t จะมาจาก C_t ในสมการ (3.16)

GARCH(p,q) นั้นใช้กระบวนการ Autoregressive และ Moving Average ในการหา Heteroscedastic Variance ได้ดังนี้

$$E_{t-1}\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-1}^2 + \sum_{i=1}^{p}\beta_i C_{t-1}$$

ถ้ากำหนดให้ค่า p=0 และ q=1 จะได้เป็น ARCH (1) หรือถ้าค่า β_I ทั้งหมดมีค่าเป็นศูนย์แบบจำลอง GARCH (p,q) จะเทียบเท่ากับแบบจำลอง ARCH (q) คุณสมบัติที่สำคัญของแบบจำลอง GARCH คือค่าความแปรปรวนอย่างมีเงื่อนไขของ disturbances ของค่า X_t สร้างขึ้นมาจากกระบวนการ ARMA จึงสามารถคาดได้ว่าส่วนเหลือจากการทำ ARMA จะแสดงถึงรูปแบบคุณลักษณะเดียวกัน เช่น ถ้าการประมาณค่า $\{X_t\}$ ด้วยกระบวนการ ARMA ค่า Autocorrelation Function (ACF) ซึ่งเป็นสหสัมพันธ์ระหว่างตัวแปรสุ่มที่หน่วยเวลาห่างกันของกระบวนการเดียวกันและ Partial Autocorrelation Function (PACF) ของส่วนที่เหลือ (residuals) ควรจะบ่งถึงกระบวนการ white noise และ ACF ของกำลังสองของส่วนเหลือ (aquared residuals) นำมาช่วยในการระบุถึงลำดับ (order) ของกระบวนการ GARCH (สธนพล วิเชียรรัตนพันธ์, 2547: 30 อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2547: 648)

3.1.6 แบบจำลอง GARCH-in-mean (GARCH-M)

จากแบบจำลอง GARCH (p , q) ในสมการ (3.16) ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนที่เหลือกำลังสอง (squared residual) ของกระบวนการนี้ โดย Engle, Robert F. ขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชั่นของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักกันในชื่อ GARCH-in-mean หรือ GARCH-M (Baillie and Bollerslve, 1992 อ้างถึงใน สธนพล วิเชียรรัตนพันธุ์, 2547: 34)

ผลตอบแทนจากหลักทรัพย์สามารถแสดงในรูปของแบบจำลอง GARCH (p , q) – M ดังสมการ (3.17) , (3.18) และ (3.19)

$$X_t = \mu_t + \delta_1 C_t^{\frac{1}{2}} + \varepsilon_t \tag{3.17}$$

$$\varepsilon_t \,/\, \Psi_{t-1} \sim N\,(0, C_t) \tag{3.18}$$

$$C_t = V + \sum_{i=1}^{p} \beta_i C_{t-1} + \sum_{i=1}^{q} \alpha_i \left(\varepsilon_{t-1}\right)^2 \tag{3.19}$$

เมื่อ X_t คือผลตอบแทนจากหลักทรัพย์

μ_t คือค่าเฉลี่ย X_t อย่างมีเงื่อนไขในอดีต (φ_{t-1}) และตามสมการข้อจำกัด $\varphi > 0$, $\alpha_1 > 0$ และ $\beta_1 \geq 0$ เพื่อให้แน่ใจว่าค่าความแปรปรวนอย่างมีเงื่อนไข (C_t) นั้นเป็นบวก

$C_t^{\frac{1}{2}}$ ในสมการ (3.17) เพื่อใช้แสดงความสัมพันธ์โดยตรงถึง trade off ระหว่างความเสี่ยงและผลตอบแทนที่คาดหวัง

อิทธิพลอย่างมีนัยสำคัญของความผันผวนในผลตอบแทนจากหลักทรัพย์ถูกตรวจจับด้วยสัมประสิทธิ์ $C_t^{\frac{1}{2}}$ (δ_1) ในสมการ (3.17) ซึ่งอธิบายแทนดัชนีของความสัมพันธ์ของการหลีกเลี่ยงความเสี่ยง ค่าสัมประสิทธิ์ δ_1 ที่เป็นบวกอย่างมีนัยสำคัญบอกถึงผู้ลงทุนในตลาดหลักทรัพย์เมื่อเผชิญกับระดับความเสี่ยงที่สูงขึ้นก็ต้องการค่าชดเชยความเสี่ยงที่มากขึ้นตามไปด้วย

3.1.7 แบบจำลอง Exponential GARCH (EGARCH)

EGARCH หรือ Exponential GARCH model ถูกเสนอโดย Nelson (1991) EGARCH model ให้ความแปรปรวนอย่างมีเงื่อนไขมีค่าดังสมการ (3.20)

$$\log\left(\sigma_t^2\right) = \omega + \beta \log\left(\sigma_{t-1}^2\right) + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{3.20}$$

ข้อแตกต่าง EGARCH model ระหว่างโปรแกรม Eviews กับโมเดลเดิมของ Nelson มีอยู่ 2 ข้อคือ ข้อแรก Nelson ได้ตั้งสมมติฐานว่า error มีการแจกแจงแบบ general distribution ส่วน Eview ได้ตั้งสมมติฐานว่า error มีการแจกแจงแบบ normal distribution ข้อที่สอง ค่า log ของความแปรปรวนอย่างมีเงื่อนไขของ Nelson แตกต่างกับของ Eview เล็กน้อยดังสมการ (3.21)

$$\log\left(\sigma_t^2\right) = \omega + \beta \log\left(\sigma_{t-1}^2\right) + \alpha\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| - \sqrt{\frac{2}{\pi}}\right) + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{3.21}$$

การประมาณค่าสมการภายใต้สมมติฐานที่ error มีการแจกแจงแบบปกติจะให้ค่าที่เหมือนกันยกเว้นค่า intercept term ω ที่แตกต่างกันเท่ากับ $\alpha\sqrt{\frac{2}{\pi}}$

และการประมาณค่า EGARCH Model โดยโปรแกรม Eviews ได้ดังสมการที่ (3.22)

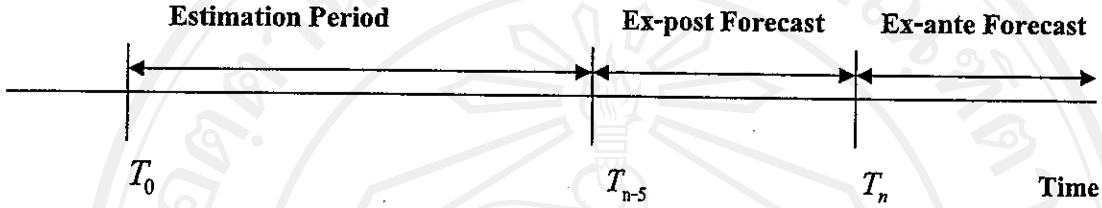
$$\log\left(\sigma_t^2\right) = \omega + \sum_{j=1}^{p}\beta_j \log\left(\sigma_{t-1}^2\right) + \sum_{i=1}^{q}\left(\alpha_i\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) \tag{3.22}$$

3.1.8 การพยากรณ์ (Forecasting)

การศึกษานี้ได้แบ่งการพยากรณ์ออกเป็น 3 ช่วงคือ Historicsl Forecast Ex-post Forecast และ Ex-ante Forecast โดย Historicsl Forecast คือ การพยากรณ์ข้อมูลในอดีตจนถึงช่วงเวลาที่พิจารณา และเนื่องจากการพยากรณ์ข้อมูลล่วงหน้านั้นจะทำให้เกิดข้อจำกัดที่ว่าความแม่นยำของข้อมูลที่ได้จากการพยากรณ์นั้น มีความน่าเชื่อถือมากน้อยเพียงใด และแบบจำลองอารีมา เหมาะสำหรับการพยากรณ์ในระยะสั้น ดังนั้นเพื่อที่จะทราบว่าแบบจำลองที่ประมาณขึ้นมานั้นสามารถที่จะพยากรณ์ได้ถูกต้องแม่นยำเพียงใด จึงได้ใช้การพยากรณ์แบบ Ex-post Forecast กล่าวคือ เป็นการพยากรณ์ข้อมูล ณ ช่วงเวลาที่มีข้อมูลจริงเกิดขึ้นแล้ว ยกตัวอย่างเช่น จะลดจำนวนค่าสังเกตการของอนุกรมเวลาลงจากข้อมูลที่มีทั้งหมด n ข้อมูล เหลือ n-5 ข้อมูล แล้วทำการถดถอยข้อมูลใหม่เพื่อดูค่า RMSE (Root Mean Squared Error) และทำการพยากรณ์ล่วงหน้า (Ex-post

Forecast) จำนวน 5 ข้อมูล เพื่อเปรียบเทียบกับข้อมูลจริงที่มีอยู่ และก็จะได้ค่า RMSE แล้วใช้ค่าสถิติดังกล่าวประกอบการพิจารณาเลือกแบบจำลองที่มีความเหมาะสม เมื่อได้แบบจำลองที่เหมาะสมภายหลังจากการวิเคราะห์ความถูกต้องแล้ว ก็สามารถนำแบบจำลองดังกล่าวใช้ในการพยากรณ์

รูปที่ 3.2 แสดงช่วงเวลาที่ใช้ในการพยากรณ์



ที่มา: Pindyck and Robinfeld (1998)

ภายหลังจากที่เลือกแบบจำลองที่ใช้เป็นตัวแทนอนุกรมเวลาข้อมูลได้แล้ว จะทำการพยากรณ์ล่วงหน้า (Ex-ante Forecast) กล่าวคือเป็นการพยากรณ์ในช่วงเวลาที่ยังไม่มีข้อมูลจริงเกิดขึ้นต่อไป

3.1.9 การตรวจสอบรูปแบบ (Diagnostic Checking)

การสร้างสมการพร้อมทั้งประมาณค่าพารามิเตอร์แล้วนั้น จะต้องทำกาตรวจสอบรูปแบบว่าสมการที่ได้มามีความเหมาะสมหรือไม่และรูปแบบใดของสมการที่ดีที่สุด โดยใช้การทดสอบต่าง ๆ ดังนี้

1) การทดสอบ Ljung-Box Q-Statistic

การทดสอบ Ljung-Box Q-Statistic เป็นการทดสอบว่าสหสัมพันธ์ในตัวเองในส่วนเหลือทุกช่วงเวลาที่ห่างกัน k มีความเป็นอิสระกันหรือไม่ โดยมีสมมติฐานดังนี้

$$H_0\ \rho(a_t)= \rho(a_t)=\ldots= \rho(a_t)=0$$

$$H_1\text{: } \rho(a_t \neq \rho(a_t) \neq \ldots \neq \rho(a_t) \neq 0$$

คำนวณตามสมการที่ (3.23) คือ

$$Q_{LB} = T\ (T+2) \sum (r_j^2 / T\text{-}j\ ) \qquad (3.23)$$

เมื่อ r_j คือสหสัมพันธ์ในตัวเองลำดับที่ j โดยที่ j=1,…,k

T คือจำนวนค่าสังเกต

ภายใต้ส่วนเหลือจากการประมาณด้วยแบบจำลอง ARIMA ค่าQ_{LB} มีการแจกแจงแบบไคสแควร์ (χ^2) ด้วยระดับความเป็นอิสระ (Degree of Freedom) เท่ากับจำนวนของสหสัมพันธ์ในตัวเองลบด้วยจำนวนของพารามิเตอร์ Autoregressive (AR) และ Moving Average (MA) ที่ได้มาจากการประมาณหรือ k-m

จะยอมรับสมมติฐานหลักเมื่อ$Q_{LB} \leq \chi^2_{\alpha,k-m}$ คือส่วนเหลือเป็นอิสระต่อกันที่ความล่า k และถ้าปฏิเสธสมมติฐานหลักเมื่อ $Q_{LB} > \chi^2_{\alpha,k-m}$ คือเกิดสหสัมพันธ์ในตัวเองอย่างน้อยหนึ่งค่าใรส่วนเหลือที่ไม่เท่ากับศูนย์

2) เกณฑ์การเลือกรูปแบบของแบบจำลองที่ดีที่สุด (Information criteria)

ในการหารูปแบบของแบบจำลอง เมื่อได้รูปแบบของแบบจำลองที่เหมาะสมหลายรูปแบบต้องมีแนวทางในการเลือกรูปแบบของแบบจำลองที่ดีที่สุด โดยพิจารณาจากค่า Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) รูปแบบของแบบจำลองที่ให้ค่า AIC และSC น้อยที่สุดจะเป็นรูปแบบที่ดีที่สุด โดย Akaike Information Criterion (AIC) สามารถคำนวณได้ตามสมการที่ (3.24) และ Schwartz Criterion (SC) คำนวณตามสมการที่ (3.25)

Akaike Information Criterion (AIC) $\quad -2l/\eta+2k/\eta \qquad (3.24)$

Schwartz Criterion (SC) $\quad -2l/\eta+ k\log\eta/\eta \qquad (3.25)$

โดยที่ k เป็นจำนวนของพารามิเตอร์ที่ทำการประมาณค่า

η เป็นจำนวนของค่าสังเกต

l เป็นค่าของ log likelihood function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว

3.2.0 การทดสอบความแม่นยำของผลการพยากรณ์ที่ได้

ในการศึกษาครั้งนี้จะทำการประเมินผลด้วย RMSE (Root Mean Square Error) ซึ่งมีสูตรการคำนวณตามลำดับ ดังนี้

RMSE คือ การวัดค่าความแตกต่างระหว่างค่าจริง และค่าที่ประมาณได้จากแบบจำลอง หาก RMSE มีค่าน้อย แสดงว่าแบบจำลองสามารถประมาณค่าประมาณได้ใกล้เคียงกับค่าจริง (Pindyck and Rubinfeld, 1998) ดังนั้นหากนี้มีค่าเท่ากับศูนย์แล้ว จะหมายความว่า ไม่เกิดความคลาดเคลื่อนในแบบจำลองนี้เลย RMSE คำนวณได้ดังนี้

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \tag{3.26}$$

โดยกำหนด Y_t^s = ค่าประมาณจากแบบจำลอง

Y_t^a = ค่าที่แท้จริง

T = จำนวนคาบเวลาที่ใช้ในการประมาณแบบจำลอง

3.2วิธีการวิจัย

1) ดำเนินการปรับข้อมูลให้อยู่ในรูปผลตอบแทนราคาของพลังงานแต่ละชนิดโดยใช้ log (relative price) ซึ่งมีสูตรในการคำนวณดังนี้

$$R_t = \text{In}\ (P_t/P_{t-1})$$

โดยที่ R_t คือ ผลตอบแทนของราคา (Price Return)

P_t คือ ราคาพลังงานที่สนใจในคาบเวลาปัจจุบัน

P_{t-1} คือ ราคาพลังงานที่สนใจในคาบเวลาที่ผ่านมา

2) นำข้อมูลผลตอบแทนราคา ซึ่งเป็นข้อมูลทุติยภูมิ (Secondary Data) ของราคาน้ำมันดิบ ราคาถ่านหิน และราคาก๊าซธรรมชาติ ซึ่งเป็นข้อมูลลักษณะอนุกรมเวลา (Time Series Data) มาตรวจสอบความนิ่งของข้อมูล โดยวิธี Unit Root Test ดังนี้

ทดสอบความนิ่งของตัวแปรที่นำมาทำการศึกษาโดยวิธี Dickey – Fuller (DF) หรือ Augmented Dickey – Fuller (ADF) ซึ่งมีสมการในการทดสอบดังนี้

$$R_t = \rho R_{t-1} + \varepsilon_t \qquad (3.27)$$

กำหนดให้ R_t คือ ตัวแปรที่เราทำการศึกษา ได้แก่ ผลตอบแทนของราคาพลังงานที่ทำการศึกษา ซึ่งประกอบด้วย ราคาน้ำมันดิบ ราคาถ่านหิน และราคาก๊าซธรรมชาติ โดยที่

α, ρ คือ ค่าคงที่

t คือ แนวโน้มเวลา

ε_t คือ ตัวแปรสุ่ม โดยมีการแจกแจงแบบปกติที่เป็นอิสระต่อกันและเหมือนกัน

โดยสมการ (3.28) ถึง (3.30) เป็นสมการที่ใช้ในการทดสอบตามวิธี DF

$$\Delta R_t = \theta R_{t-1} + \varepsilon_t \qquad (3.28)$$

$$\Delta R_t = \alpha + \theta R_{t-1} + \varepsilon_t \qquad (3.29)$$

$$\Delta R_t = \alpha + \beta_t + \theta R_{t-1} + \varepsilon_t \qquad (3.30)$$

โดยสมการ (3.31) ถึง (3.33) เป็นสมการที่ใช้ในการทดสอบตามวิธี ADF

$$\Delta R_t = \theta R_{t-1} + \sum_{j=1}^{p} \lambda_j \Delta R_{t-j} + \varepsilon_t \qquad (3.31)$$

$$\Delta R_t = \alpha + \theta R_{t-1} + \sum_{j=1}^{p} \lambda_j \Delta R_{t-j} + \varepsilon_t \qquad (3.32)$$

$$\Delta R_t = \alpha + \beta t + \theta R_{t-1} + \sum_{j=1}^{p} \lambda_j \Delta R_{t-} + \varepsilon_t \qquad (3.33)$$

การทดสอบ Unit root ทั้ง มีขั้นตอนดังนี้ทั้ง 2 วิธี คือ DF และ ADF มีขั้นตอนดังนี้

- ตั้งสมมุติฐานในการทดสอบ คือ $H_0 : \theta = 0$ และ $H_1 : \theta \neq 0$
- ทำการเปรียบเทียบค่าสถิติทีที่คำนวณได้กับค่าในตาราง Dickey – Fuller เปรียบเทียบกับค่าวิกฤติ Mackinnon แบ่งได้เป็น 2 กรณี

  - ถ้าไม่สามารถปฏิเสธ H_0 หรือยอมรับ แสดงว่าตัวแปรที่นำมาทดสอบมี Unit rootหรือมีลักษณะไม่นิ่ง ต้องมีการทำ Differencing ตัวแปรไปเรื่อย ๆ จนสามารถปฏิเสธ H_0 ได้
  - ปฏิเสธ ทำให้ทราบ Order of Integration

3) นำค่าผลตอบแทนราคาที่มีลักษณะนิ่งแล้ว มาสร้างแบบจำลองที่ดีที่สุด เพื่อประมาณการความผันผวนของราคาในอนาคต โดยมีขั้นตอนในการสร้างและประมาณค่าแบบจำลองดังนี้

3.1 สร้าง Correlograme แสดง ACF และ PACF เพื่อใช้ในการพิจารณาเลือกรูปแบบที่เหมาะสมของอนุกรมเวลา RAMA (p,q)

3.2 ประมาณค่าสมการค่าเฉลี่ยโดยเลือกใช้ lag pและq ที่ได้จากการวิเคราะห์ Correlograme

3.3 ทดลองเลือก pและq สำหรับรูปแบบที่เหมาะสมของกระบวนการGARCH (p,q)

3.4 ทำการประมาณค่าพารามิเตอร์ของแบบจำลองที่ได้จากการทดลองเลือกตามข้อ 3.2 และ3.3 และพิจารณาว่าค่าพารามิเตอร์ที่ได้มีความแตกต่างจากศูนย์อน่างมีนัยสำคัญหรือไม่โดยทดสอบค่า t-statistic และตรวจสอบเงื่อนไขสเตชั่นนารี (Stationary) ของแบบจำลอง ARMA ถ้าค่าที่ได้ไม่ตรงตามเงื่อนไขให้ทดลองเปลี่ยนค่า pและq อื่น ๆ แทน

3.5 ตรวจสอบรูปแบบที่เหมาะสมเพื่อพิจารณาว่าส่วนที่เหลือ (Residual) ไม่กิด Serial Correlation กัน โดยทำการทดสอบค่า Q_{LB}-Statistic โดยถ้ายอมรับสมมติฐานหลักแสดงว่าแบบจำลองมีความเหมาะสมแล้ว

3.6 เลือกรูปแบบที่ดีที่สุดให้กับแบบจำลอง ARMA-EGARCH ARIMA-GARCH-M และARIMA- GARCH โดยพิจารณาค่า Akaike Information Criterion (AIC) และ Schwarzt Criterion (SC) ที่น้อยที่สุดจะเป็นรูปแบบที่ดีที่สุด

4) นำแบบจำลองที่ดีที่สุดจากแต่ละแนวคิดมาพยากรณ์ผลตอบแทนของราคาในอนาคต และนำค่าที่ได้มาเปรียบเทียบกับข้อมูลจริงเพื่อหาแนวคิดที่ดีที่สุดในการพยากรณ์ผลตอบแทนเพื่อประมาณการความผันผวนของผลตอบแทนของราคาพลังงานแต่ละชนิดโดยใช้เกณฑ์ RMSE (Root Mean Square Error) ที่ต่ำที่สุด ซึ่งแสดงถึงความสามารถในการพยากรณ์ที่สูงกว่า