บทที่ 2

กรอบแนวคิดทางทฤษฎีและเอกสารที่เกี่ยวข้อง

2.1 แนวคิดทฤษฎีที่เกี่ยวข้อง

2.1.1 ความรู้พื้นฐานเกี่ยวกับมลพิษทางอากาศ

มลพิษทางอากาศหมายถึงภาวะของอากาศที่มีการเจือปนของสารพิษในปริมาณความเข้มข้นสูงกว่าปกติเป็นเวลานานพอที่จะทำให้เกิดอันตรายแก่มนุษย์สัตว์พืชหรือทรัพย์สินต่างๆเช่นอาคารบ้านเรือนโบราณสถานโบราณวัตถุภาชนะเครื่องใช้เครื่องจักรกลที่เป็นโลหะยานพาหนะต่างๆให้เกิดความสกปรกและเกิดการกัดกร่อนผุพังทรุดโทรมจนใช้การไม่ได้อาจเกิดขึ้นเองตามธรรมชาติเช่นฝุ่นละอองจากลมพายุภูเขาไฟระเบิดแผ่นดินไหวไฟไหม้ป่าซึ่งปกติมลพิษอากาศจากธรรมชาติเป็นอันตรายต่อมนุษย์น้อยมากเพราะแหล่งกำเนิดอยู่ไกลและปริมาณที่เข้าสู่สภาพแวดล้อมของมนุษย์และสัตว์มีน้อยสำหรับกรณีเกิดจากการกระทำของมนุษย์เช่นมลพิษจากท่อไอเสียของรถยนต์โรงงานอุตสาหกรรมกระบวนการผลิตกิจกรรมด้านการเกษตรการระเหยของก๊าซบางชนิดเป็นต้น (เอกสารประกอบการสอนรายวิชา Air & Noise Pollution and Control)

ปัจจุบันอากาศทั่วโลกซึ่งรวมประเทศไทยด้วยนั้นมีมลพิษปนเปื้อนอยู่มากทั้งอากาศที่ใช้หายใจภายนอกอาคาร (Outdoor Air Pollution) และภายในอาคาร(Indoor Air Pollution) เป็นสาเหตุทำให้เกิดการเจ็บป่วยโดยเฉพาะอย่างยิ่งทำให้เกิดโรคระบบทางเดินหายใจซึ่งทำให้สูญเสียค่าใช้จ่ายในการรักษาพยาบาลเป็นอย่างมากแหล่งของสารปนเปื้อนเหล่านี้มาจากการประกอบกิจกรรมต่างๆในสังคมเป็นส่วนใหญ่นอกจากนี้ยังเกิดจากแหล่งตามธรรมชาติอีกด้วยซึ่งอาจจำแนกได้ดังนี้

1) แหล่งก่อมลพิษทางอากาศจากการประกอบกิจกรรมในสังคมได้แก่การประกอบอาชีพในการประกอบการอุตสาหกรรมเกษตรกรรมการบริการการก่อสร้างการจราจรการเผาศพและพฤติกรรมของบุคคลในสังคมได้แก่การสูบบุหรี่การทิ้งและการเผาขยะการเผาป่าเป็นต้น

2) แหล่งก่อมลพิษทางอากาศจากธรรมชาติได้แก่ฝุ่นละอองก๊าซและรังสีที่เกิดขึ้นตามธรรมชาติได้แก่ฝุ่นละอองก๊าซและรังสีจากพื้นผิวโลกจากการระเบิดของภูเขาไฟแหล่งก๊าซเรดอนจากพื้นผิวโลก

ตารางที่ 2.1 ตารางชนิดของมลพิษทางอากาศ

ประเภทของสาร	**ชนิดของสาร**	**ตัวอย่างสารมลพิษ**
แก๊สที่เป็นอนินทรียสาร	ออกไซด์ของไนโตรเจน	NO, NO2
	ออกไซด์ของกำมะถัน	SO2, SO3
	ออกไซด์ของคาร์บอน	CO, CO2
	แก๊สอนินทรีย์อื่นๆ	H2S, HF, NH3, Cl2
แก๊สที่เป็นอินทรียสาร	ไฮโดรคาร์บอน	มีเทน, บิวเทน, ออกเทน, เบนซิน ,อะเซทิลีน, เอทิลีน, บิวตะไดอีน
	แอลดีไฮด์และคีโทน	ฟอร์มาลดีไฮด์, อะซีโทน
	อินทรียสารอื่นๆ	แอลกอฮอล์, กรด, อินทรีคลอริเนเตต, ไฮโดรคาร์บอน, เบนโซไพริน
อนุภาคสาร	อนุภาคสารที่เป็นของแข็ง	ควัน, เขม่า, ฝุ่น, ขี้เถ้า, คาร์บอน, ตะกั่ว, ใยหิน
	อนุภาคสารที่เป็นของเหลว	สเปรย์, กรดต่างๆ, ละอองน้ำ, น้ำมัน

ที่มา: พิมล เรียนวัฒนาและชัยวัฒน์ เจนวาณิชย์(2525)

2.1.2 ผลกระทบของสารมลพิษทางอากาศ

เกิดภาวะการเพิ่มขึ้นของอุณหภูมิโลก (Global Warming)

เมื่อประมาณ 10,000 ปีที่ผ่านมาปรากฏการณ์เรือนกระจกเป็นเหตุให้อุณหภูมิของโลกสูงขึ้น 1 - 2 องศาเซลเซียสนับแต่ปีพ.ศ. 2403 เป็นต้นมาพบว่าอุณหภูมของโลกสูงขึ้นอีกประมาณ0.5 องศาเซลเซียสคณะกรรมการระหว่างชาติว่าด้วยความเปลี่ยนแปลงของภูมิอากาศสรุปว่าถ้าหากแก้ปัญหานี้ไม่ได้อุณหภูมิเฉลี่ยของโลกจะเพิ่มขึ้น 0.2 - 0.5 องศาเซลเซียสทุก 10 ปีทำให้เกิดความแห้งแล้งรุนแรงภาวะฝนทิ้งช่วงยาวนานกว่าปกติและเกิดปัญหาอื่นตามมา

ระดับน้ำทะเลสูงขึ้นและเกิดน้ำท่วมรุนแรงกว่าเดิม

นักวิทยาศาสตร์คำนวณว่าถ้าอุณหภูมิของโลกเพิ่ม 1.5 - 4.5 องศาเซลเซียสน้ำแข็งขั้วโลกจะละลายเป็นผลให้น้ำทะเลสูงขึ้น 20 - 140 เซนติเมตรโดยคาดว่าน้ำทะเลจะสูงขึ้นอย่างมากใน พ.ศ.2573 ศตวรรษที่แล้วระดับสูงกว่าเดิม 10 - 15 เซนติเมตรปัจจุบันสูงขึ้นปีละ 1.2 มิลลิเมตรIPCC ประมาณว่าในพ.ศ. 2573 น้ำทะเลจะสูงขึ้น 20 เซนติเมตร พ.ศ. 2633 สูงเพิ่มอีก 60 เซนติเมตรและ พ.ศ. 2683 จะสูงกว่าเดิมถึง 1 เมตรถ้าน้ำทะเลสูงขึ้นเพียง 50 เซนติเมตรเมืองสำคัญและท่าเรือจะจมน้ำใต้ผิวน้ำคนจำนวนมากต้องอพยพและเกิดปัญหาสังคมมากมายเช่นกรุงเทพมหานครมะนิลาโตเกียวกัลกัตตานิวยอร์กบัวโนส/ไอเรสภาคใต้ของประเทศบังคลาเทศมัลดีฟส์เนเธอร์แลนด์พื้นที่ทางใต้และตะวันออกของสหราชอาณาจักรและชายฝั่งด้านตะวันออกเฉียงใต้ของสหรัฐอเมริกา

ผลกระทบต่อเกษตรกรรม

ทำให้ขยายการเกษตรไปทางขั้วโลกถ้าอุณหภูมิเพิ่มขึ้น 1 องศาเซลเซียสจะสามารถปลูกธัญพืชสูงขึ้นไปทางขั้วโลกเหนือได้ 150 - 200 กิโลเมตรและปลูกในพื้นที่สูงขึ้นอีก 100 - 200 เมตรพืชที่ปลูกตามขอบทะเลทรายจะเสียหายเพราะทะเลทรายขยายตัวการนำพืชไปปลูกถิ่นอื่นต้องปรับสภาพดินและน้ำวัชพืชและพืชจะโตเร็วและมีขนาดใหญ่กว่าเดิมเนื่องจากได้รับคาร์บอนไดออกไซด์เพิ่มขึ้นแต่ดินจะเสื่อมเร็วเพราะแร่ธาตุจะถูกนำไปใช้มากพืชจะขาดไนโตรเจนความต้านทานโรคและแมลงลดลงผลผลิตพืชมีแนวโน้มสูงขึ้นโดยพืชใช้คาร์บอนไดออกไซด์ในการสังเคราะห์แสงได้ดีกว่าจะให้ผลผลิตมากกว่าเช่นพืชที่ใช้คาร์บอน 3 อะตอม (พวกถั่วมันสำปะหลังกล้วยมันฝรั่งหัวผักกาดหวานและข้าวสาลี) จะมีผลผลิตสูงกว่าพืชที่ใช้คาร์บอน 4 อะตอม (พวกข้าวโพดข้าวฟ่างอ้อยและลูกเดือย) ผลผลิตในหลายแหล่งเช่นสหรัฐอเมริกายุโรปและญี่ปุ่นจะมากเกินความต้องการทำให้ราคาตกต่ำซึ่งส่งผลกระทบต่อเศรษฐกิจและสังคมโลกเป็นเหตุให้ต้องเปลี่ยนแปลงการผลิตและการใช้ดินต้องปรับปรุงพันธุ์พืชให้มีความต้านทานโรคแมลงและอากาศที่แห้งแล้งขึ้น

ผลกระทบต่อสุขภาพของชุมชนได้แก่

1) มีผลเสียต่ออารมณ์ร่างกายและการปฏิบัติกิจกรรมโดยอากาศร้อนทำให้คนรู้สึกหงุดหงิดฉุนเฉียวง่ายเหนื่อยง่ายและประสิทธิภาพการทำงานต่ำ

2) มีอันตรายต่อผิวหนังอุณหภูมิที่สูงมากจะทำให้เหงื่อออกมากโดยเฉพาะตามง่ามเท้ารักแร้และข้อพับทำให้ผิวหนังเปื่อยเกิดผดผื่นคันหรือถูกเชื้อราหรือแบคทีเรียทำให้อักเสบได้ง่าย

3) ทำให้โรคเขตร้อนระบาดได้มากขึ้นเช่นโรคไข้สาซึ่งเกิดจากเชื้อไวรัสโดยยุงเป็นพาหะมีอาการโรคไข้เลือดออกต่อมน้ำเหลืองอักเสบบวมปวดกล้ามเนื้อและข้ออาจเสียชีวิตได้ไม่มีวัคซีน

และยาที่ใช้รักษาเฉพาะเมื่อพ.ศ. 2540 ระบาดในประเทศบราซิลมีผู้ป่วยไม่ต่ำกว่า 24,000 คนและในเวเนซุเอลา 32,000 คนเสียชีวิต 40 คนหากอุณหภูมิสูงขึ้นจะทำให้โรคนี้ระบาดทั่วแถบร้อนของโรคได้

4) เป็นอันตรายต่อเด็กและคนชราโดยจะทำให้มีโอกาสเสียชีวิตจากคลื่นความเย็นและคลื่นความร้อนมากขึ้น

ผลกระทบทางเศรษฐกิจและสังคม

1) ใช้ทรัพยากรธรรมชาติให้มากขึ้นเพราะอากาศร้อนจะทำให้มีการใช้เครื่องปรับอากาศและแร่เชื้อเพลิงเพิ่มขึ้นโดยเฉพาะในชุมชนเมืองซึ่งจะมีอุณหภูมิสูงกว่าชนบท

2) ราคาพืชผลการเกษตรตกต่ำทั่วโลกเพราะประเทศที่มีกำลังซื้อพืชผลได้เกินความต้องการทำให้เกิดการเปลี่ยนแปลงด้านรูปแบบการค้าและสินค้าเกษตรกรรม

3) เกษตรกรจะเสียต้นทุนการผลิตมากขึ้นเพราะดินเสื่อมความอุดมสมบูรณ์เร็วศัตรูพืชเพิ่มขึ้นความต้านทานของพืชลดลงขณะเดียวกันก็ต้องลดรายจ่ายลงเช่นลดการจ้างงานเป็นต้น

4) ประเทศที่ยากจนจะขาดแคลนอาหารมากขึ้นเนื่องจากการปลูกพืชในบางแห่งได้ผลน้อยทะเลทรายเพิ่มขนาดและพืชหลักของท้องถิ่นซึ่งได้แก่ข้าวโพดข้าวฟ่างอ้อยและลูกเดือยมีอัตราเพิ่มของผลผลิตน้อยลง

5) แหล่งท่องเที่ยวชายหาดจะถูกน้ำทะเลท่วมดินจะพังทลายทำให้เสียงบประมาณเพื่อการปรับปรุงจำนวนมาก

6) การพัฒนาประเทศทำได้ล่าช้าเนื่องจากต้องใช้งบประมาณเพื่อแก้ไขปัญหาที่เกิดขึ้น

2.1.3 ดัชนีชี้วัดผลกระทบต่อสุขภาพของมลพิษทางอากาศในประเทศไทย

ตัวชี้วัดมลพิษทางอากาศ(Indicator of Air Pollution) หมายถึงสิ่งที่แสดงนัยสำคัญและแนวโน้มของคุณลักษณะและปริมาณของอากาศ

ในกรณีของดัชนีคุณภาพอากาศนั้นสำนักจัดการคุณภาพอากาศและเสียงกรมควบคุมมลพิษได้กำหนดเกณฑ์ของดัชนีคุณภาพอากาศสำหรับประเทศไทยโดยมีการคำนวณเทียบมาตรฐานคุณภาพบรรยากาศทั่วไปของสารมลพิษทางอากาศ 5 ประเภทได้แก่ก๊าซโอโซนเฉลี่ย 1 ชั่วโมงก๊าซไนโตรเจนไดออกไซด์เฉลี่ย 1 ชั่วโมงก๊าซคาร์บอนมอนนอกไซด์เฉลี่ย 8 ชั่วโมงก๊าซซัลเฟอร์ไดออกไซด์เฉลี่ย 24 ชั่วโมงและฝุ่นละอองขนาดเล็กกว่า 10 ไมครอนเฉลี่ย 24 ชั่วโมงทั้งนี้ดัชนีคุณภาพอากาศที่คำนวณได้ของสารมลพิษทางอากาศประเภทใดมีค่าสูงสุดจะใช้เป็นดัชนีคุณภาพอากาศของวันนั้นดัชนีคุณภาพอากาศของประเทศไทยแบ่งออกเป็น 5 ระดับคือตั้งแต่ 0 ถึงมากกว่า 300 ซึ่งแต่ละระดับจะใช้สีเป็นสัญลักษณ์เปรียบเทียบระดับของผลกระทบต่อสุขภาพอนามัยโดยดัชนีคุณภาพ

อากาศ 100 จะมีค่าเทียบเท่ามาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไปหากดัชนีคุณภาพอากาศสูงเกินกว่า 100 แสดงว่าค่าความเข้มข้นของมลพิษทางอากาศมีค่าเกินมาตรฐานและคุณภาพอากาศในวันนั้นจะเริ่มมีผลกระทบต่อสุขภาพอนามัยของประชาชนดังแสดงในตารางที่ 1 ซึ่งดัชนีเปรียบเทียบคุณภาพอากาศในระดับต่างๆนั้นได้แสดงไว้ในตารางที่ 2 ส่วนกรณีของรังสีทิศทางลมความเร็วลมปริมาณน้ำฝนความชื้นสัมพัทธ์นั้นไม่ได้ทำการตรวจวัดอย่างต่อเนื่องมีการตรวจวัดเป็นบางกรณีและไม่ได้เผยแพร่ข้อมูลโดยทั่วไป

ตารางที่ 2.2 ตารางเกณฑ์ของดัชนีคุณภาพอากาศสำหรับประเทศไทย

AQI	**ความหมาย**	**สีที่ใช้**	**แนวทางการป้องกันผลกระทบ**
0-50	คุณภาพดี	ฟ้า	ไม่มีผลกระทบต่อสุขภาพ
51-100	คุณภาพปานกลาง	เขียว	ไม่มีผลกระทบต่อสุขภาพ
101-200	มีผลกระทบต่อสุขภาพ	เหลือง	ผู้ป่วยโรคระบบทางเดินหายใจควรหลีกเลี่ยงการออกกำลังภายนอกอาคาร บุคคลทั่วไป โดยเฉพาะเด็กและผู้สูงอายุไม่ควรทำกิจกรรมภายนอกอาคารเป็นเวลานาน
201-300	มีผลกระทบต่อสุขภาพมาก	ส้ม	ผู้ป่วยโรคระบบทางเดินหายใจควรหลีกเลี่ยงกิจกรรมภายนอกอาคาร บุคคลทั่วไป โดยเฉพาะเด็กและผู้สูงอายุควรจำกัดการออกกำลังภายนอกอาคาร
มากกว่า 300	อันตราย	แดง	บุคคลทั่วไป ควรหลีกเลี่ยงการออกกำลังภายนอกอาคาร

ที่มา:กรมควบคุมมลพิษ(2555)

การคำนวณดัชนีคุณภาพอากาศรายวันของสารมลพิษทางอากาศแต่ละประเภท (i)

จะคำนวณจากค่าความเข้มข้นของสารมลพิษทางอากาศจากข้อมูลผลการตรวจวัดคุณภาพอากาศโดยแต่ละระดับของค่าความเข้มข้นของสารมลพิษทางอากาศเทียบเท่ากับค่าดัชนีคุณภาพอากาศที่ระดับต่างๆดังแสดงไว้ในตารางที่ 2และมีสูตรการคำนวณดังนี้

$$I_i = \frac{I_{ij+1} - I_{ij}}{X_{ij+1} - X_{ij}} (X_i - X_{ij}) + I_{ij}$$

กำหนดให้ Xi=ความเข้มข้นของสารมลพิษทางอากาศจากผลการตรวจวัด

Xij=ความเข้มข้นของสารมลพิษทางอากาศที่เป็นค่าต่ำสุดของช่วงพิสัยที่มีค่า Xi นั้น

Xij+1=ความเข้มข้นของสารมลพิษทางอากาศที่เป็นค่าสูงสุดของช่วงพิสัยที่มีค่าXi นั้น

Ii=ค่าดัชนีย่อยคุณภาพอากาศ

Iij=ค่าดัชนีย่อยคุณภาพอากาศที่เป็นค่าต่ำสุดของช่วงพิสัยที่มีค่าIi นั้น

Iij+1=ค่าดัชนีย่อยคุณภาพอากาศที่เป็นค่าสูงสุดของช่วงพิสัยที่มีค่าIi นั้น

AQI=ค่าดัชนีคุณภาพอากาศ

ตารางที่2.3 ตารางค่าความเข้มข้นของสารมลพิษทางอากาศที่เทียบเท่ากับค่าดัชนีคุณภาพอากาศ

AQI	**PM_{10} (24 ชม.)**	**O_3 (1 ชม.)**		**SO_2 (24 ชม.)**		**NO_2 (1 ชม.)**		**CO (8 ชม.)**	
	มคก./ลบม.	มคก./ลบม.	ppb	มคก./ลบม.	ppb	มคก./ลบม.	ppb	มคก./ลบม.	ppb
50	40	100	51	65	25	160	85	5.13	4.48
100	120	200	100	300	120	320	170	10.26	9
200	350	400	203	800	305	1,130	600	17	14.84
300	420	800	405	1,600	610	2,260	1,202	34	29.69
400	500	1,000	509	2,100	802	3,000	1,594	46	40.17
500	600	1,200	611	2,620	1,000	3,750	1,993	57.5	50.21

ที่มา:กรมควบคุมมลพิษ (2555)

2.1.4 อนุกรมเวลาแบบบอกซ์และเจนกินส์(Box – Jenkins)

การพยากรณ์อนุกรมเวลาแบบบอกซ์และเจนกินส์จะใช้อนุกรมเวลาในอดีตเพียงอย่างเดียวในการพยากรณ์อนุกรมเวลาในอนาคตลักษณะที่สำคัญของการพยากรณ์แบบบอกซ์และเจนกินส์

การพยากรณ์ระยะสั้นการพยากรณ์แบบบอกซ์และเจนกินส์จะใช้การพยากรณ์ในระยะสั้น เพราะรูปแบบของการพยากรณ์จะให้ความสำคัญกับอนุกรมเวลาที่อยู่ใกล้เวลาพยากรณ์มากกว่าอนุกรมเวลาที่อยู่ไกลเวลาพยากรณ์ดังนั้นการพยากรณ์จากข้อมูลระยะยาวอาจทำให้เชื่อถือได้น้อยกว่าการพยากรณ์จากข้อมูลระยะสั้น

ชนิดของอนุกรมเวลาอนุกรมเวลาที่ใช้พยากรณ์จะเป็นจำนวนจริงและอนุกรมเวลาจะต้องเกิดขึ้นในช่วงเวลาที่เท่ากัน

ขนาดของอนุกรมเวลาควรจะใช้อนุกรมเวลาอย่างน้อย50 ตัวสำหรับอนุกรมเวลาที่เป็นฤดูกาล ควรใช้จำนวนมากๆ โดยอนุกรมเวลาจะต้องเกิดขึ้นในช่วงเวลาที่เท่ากัน

อนุกรมเวลาที่มีลักษณะนิ่งอนุกรมเวลาที่จะใช้ในการพยากรณ์แบบบอกซ์และเจนกินส์จะต้องเป็นอนุกรมเวลาที่มีลักษณะนิ่งถ้าอนุกรมเวลามีลักษณะไม่นิ่งจะต้องหาผลต่างหรือแปลงอนุกรมเวลาเพื่อเปลี่ยนให้เป็นอนุกรมเวลาที่มีลักษณะนิ่ง

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง หมายถึง การที่ข้อมูลอนุกรมเวลาอยู่ในสภาพของการสมดุลเชิงสถิติ (Statistical Equilibrium) ซึ่งหมายถึง การที่ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงถึงแม้เวลาจะเปลี่ยนไป แสดงได้ดังนี้โดยกำหนดให้

$X_t, X_{t+1}, X_{t+2}, \dots, X_{t+k}$ และ $X_{t+m}, X_{t+m+1}, X_{t+m+2}, \dots, X_{t+m+k}$เป็นข้อมูลอนุกรมเวลาที่เวลา t, t+1, t+2, …, t+kและ เวลา t+m, t+m+1, t+m+2, …, t+m+kตามลำดับ

$P(X_t, X_{t+1}, X_{t+2}, \dots, X_{t+k})$และ$\mathrm{P}(\mathrm{X}_{t+m}, \mathrm{X}_{t+m+1}, \mathrm{X}_{t+m+2}, \dots, \mathrm{X}_{t+m+k})$เป็นการแจกแจงความน่าจะเป็นร่วมของ $\mathrm{X}_t, \mathrm{X}_{t+1}, \mathrm{X}_{t+2}, \dots, \mathrm{X}_{t+k}$ และ $\mathrm{X}_{t+m}, \mathrm{X}_{t+m+1}, \mathrm{X}_{t+m+2}, \dots, \mathrm{X}_{t+m+k}$ ตามลำดับกล่าวได้ว่า X_t เป็นข้อมูลอนุกรมเวลาที่มีลักษณะนิ่งแบบเข้ม (Strictly Stationary) เมื่อ

$\mathrm{P}(\mathrm{X}_t, \mathrm{X}_{t+1}, \mathrm{X}_{t+2}, \dots, \mathrm{X}_{t+k} = \mathrm{P}(\mathrm{X}_{t+m}, \mathrm{X}_{t+m+1}, \mathrm{X}_{t+m+2}, \dots, \mathrm{X}_{t+m+k})$

ถ้า$\mathrm{P}(\mathrm{X}_t, \mathrm{X}_{t+1}, \mathrm{X}_{t+2}, \dots, \mathrm{X}_{t+k}) \neq \mathrm{P}(\mathrm{X}_{t+m}, \mathrm{X}_{t+m+1}, \mathrm{X}_{t+m+2}, \dots, \mathrm{X}_{t+m+k})$

สรุปได้ว่าข้อมูลอนุกรมเวลาดังกล่าวมีลักษณะไม่นิ่งแบบเข้ม

จะเรียกอนุกรมเวลาว่ามีลักษณะนิ่งแบบอ่อนเมื่อลักษณะของการแจกแจงความน่าจะเป็นร่วมไม่เปลี่ยนแปลงเมื่อเวลาเปลี่ยนแปลงเช่นค่าเฉลี่ย(Mean) ความแปรปรวน(Variance) ความแปรปรวนร่วม(Covariance)มีค่าเท่ากันทุกช่วงเวลาสำหรับอนุกรมเวลาที่มีแนวโน้มหรือมีฤดูกาลจะมีค่าเฉลี่ยไม่เท่ากันส่วนอนุกรมเวลาที่มีความแปรผันสูงจะเป็นลักษณะของข้อมูลที่มีความแปรปรวนไม่เท่ากันซึ่งอนุกรมเวลาทั้งสองดังกล่าวจะเป็นอนุกรมเวลาที่มีลักษณะไม่นิ่งแบบอ่อน

ดังนั้นในการทดสอบอนุกรมเวลาว่านิ่งหรือไม่นั้นจึงทดสอบการมีแนวโน้มหรือมีฤดูกาลแทนการทดสอบค่าเฉลี่ยโดยตรงเพราะไม่สะดวกในการแบ่งช่วงเวลาซึ่งการทดสอบแนวโน้มและฤดูกาลมีสถิติที่ใช้ในการทดสอบหลายตัวในที่นี้จะเสนอเพียงวิธีของบอกซ์และเจนกินส์โดยพิจารณาจากคอริโรแกรม(Correlogram)ที่ได้จากการเขียนกราฟสัมประสิทธิ์สหสัมพันธ์ในตัวเอง(ρ_k)ในแต่ละช่วงห่างของเวลากับช่วงห่างkช่วงเวลา

2.1.4.1 ฟังก์ชันสหสัมพันธ์ในตัวเอง(Autocorrelation Function : ACF)

ถ้าอนุกรมเวลามีลักษณะนิ่ง จะได้ความแปรปรวนร่วมในตัวเอง (Auto Covariance) ของอนุกรมเวลาที่มีช่วงห่างเท่ากันจะไม่เปลี่ยนแปลง ซึ่งความแปรปรวนร่วมในตัวเองของ X_t และ X_{t+k}ที่ห่างกัน k หน่วย ใช้สัญลักษณ์ γ_k

โดยที่$\gamma_k = E[(X_t-\mu)(X_{t+k}-\mu)]$

X_tเป็นอนุกรมเวลา ณ เวลาt

μเป็นค่าเฉลี่ยของอนุกรมเวลา

ให้ρ เป็นสัมประสิทธิ์สหสัมพันธ์ในตัวเอง(Autocorrelation Coefficients) โดยที่$\rho_k = \frac{\gamma_k}{\gamma_0}$

ρ_k ; k = 0, 1, 2, …เรียกว่า ฟังก์ชันสหสัมพันธ์ในตัวเองซึ่งจะมีค่าตั้งแต่ -1ถึง ในทางปฏิบัติจะประมาณค่าของρ_kจากอนุกรมเวลา$X_1, X_2, \ldots, X_n$ โดยตัวประมาณของρ_kจะแทนด้วย r_k โดยที่

$$R_k = \frac{\sum_{t=1}^{n-k}(X_t-\bar{X})(X_{t+k}-\bar{X})}{\sum_{t=1}^{n}(X_t-\bar{X})^2} \text{ หรือ } r_k = \frac{C_k}{C_0}$$

$$\text{เมื่อ} \quad C_k = \frac{\sum_{t=1}^{n-k}(X_t-\bar{X})(X_{t+k}-\bar{X})}{n} \text{ และ } C_0 = \frac{\sum_{t=1}^{n}(X_t-\bar{X})^2}{n}$$

โดยที่$\bar{X}$เป็นค่าเฉลี่ยของอนุกรมเวลา$X_1, X_2, \ldots, X_n$ซึ่ง$\bar{X} = \frac{\sum_{t=1}^{n} X_t}{n}$

ความคลาดเคลื่อนมาตรฐานของฟังก์ชันสหสัมพันธ์ในตัวเอง (Standard Error of Autocorrelation Function)ในการทดสอบนัยสำคัญของสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (r_k)Bartlett (1946)ได้ประมาณค่าความแปรปรวนของr_k ; k = 1, 2, 3, … ของอนุกรมเวลาคงที่ดังนี้

$$Var(r_k) = \frac{1 + 2\sum_{j=1}^{k-1} r_j^2}{n}$$

ดังนั้น ค่าความคลาดเคลื่อนมาตรฐานของ r_k มีค่าเท่ากับ

$$SE(r_k) = \sqrt{\frac{1 + 2\sum_{j=1}^{k-1} r_j^2}{n}}$$

ซึ่งจะใช้ในการทดสอบนัยสำคัญของ r_k นั่นคือ ที่ระดับนัยสำคัญ 0.05

ถ้า $|r_k| \geq 1.96SE(r_k)$แล้วr_kจะมีนัยสำคัญทางสถิติ นั่นคือ $\rho_k \neq 0$ในทางตรงข้าม ρ_k จะมีค่าเท่ากับ 0 ก็ต่อเมื่อ $|r_k| < 1.96SE(r_k)$

2.1.4.2 ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน(Partial Autocorrelation Function : PACF)

สหสัมพันธ์ในตัวเองบางส่วนใช้สัญลักษณ์ $\emptyset_{kk}$ การหาสหสัมพันธ์ในตัวเองบางส่วนทำได้โดยอาศัยสมการยูล – วอคเกอร์ (Yule – Walker Equation) ซึ่งสามารถเขียนให้อยู่ในรูปเมตริกซ์ ได้ดังนี้

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{j-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{j-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{j-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{j-1} & \rho_{j-2} & \rho_{j-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \emptyset_{k1} \\ \emptyset_{k2} \\ \emptyset_{k3} \\ \vdots \\ \emptyset_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{bmatrix}$$

ในทางปฏิบัติค่า $\emptyset_{kk}$จะถูกประมาณด้วย$\emptyset_{kk}$และค่าρ_jจะถูกประมาณด้วย r_j; j = 1, 2, 3, …, kดังนั้น ตัวประมาณหาได้ในรูปของr_k เช่นเมื่อ j =1 ;$\hat{\emptyset}_{11} = r_1$

$$j = 2 ; \begin{bmatrix} 1 & r_1 \\ r_1 & 1 \end{bmatrix} \begin{bmatrix} \hat{\emptyset}_{21} \\ \hat{\emptyset}_{22} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\emptyset}_{21} \\ \hat{\emptyset}_{22} \end{bmatrix} = \frac{\begin{bmatrix} 1 & r_1 \\ r_1 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}}{1 - r_1^2}$$

$$\hat{\emptyset}_{21} = \frac{r_1 - r_2 r_1}{1 - r_1^2} = \frac{r_{1(1-r_2)}}{1 - r_1^2}, \hat{\emptyset}_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

ความคลาดเคลื่อนมาตรฐานของฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน(Standard Error of Partial Autocorrelation Function)ในการทดสอบนัยสำคัญของสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน$(\emptyset_{kk})$ Quenouille (1949) ได้ประมาณค่าความแปรปรวนของ$\widehat{\emptyset}_{kk}$ดังนี้

$Var(\widehat{\emptyset}_{kk}) = \frac{1}{n}$; k = 1, 2, 3, …

ดังนั้นค่าความคลาดเคลื่อนมาตรฐานของ $\emptyset_{kk}$มีค่าเท่ากับ

$$SE(\widehat{\emptyset}_{kk}) = \sqrt{\frac{1}{n}}$$

ซึ่งจะใช้ในการทดสอบนัยสำคัญของ$\emptyset_{kk}$นั่นคือ ที่ระดับนัยสำคัญ0.05

ถ้า $|\widehat{\emptyset}_{kk}| \geq 1.96SE(\widehat{\emptyset}_{kk})$*แล้ว* $\emptyset_{kk}$จะมีนัยสำคัญทางสถิติ นั่นคือ$\emptyset_{kk} \neq 0$และ $\emptyset_{kk}$จะมีค่าเท่ากับ 0ก็ต่อเมื่อ $|\widehat{\emptyset}_{kk}| < 1.96SE(\widehat{\emptyset}_{kk})$

2.1.4.3 รูปแบบอนุกรมเวลาแบบบอกซ์– เจนกินส์

2.1.4.3.1 รูปแบบถดถอยในตัวเอง *(Autoregressive Model : AR)* เป็นรูปแบบของอนุกรมเวลาที่มีลักษณะนิ่ง โดยที่สามารถกระจายอยู่ในรูปของอนุกรมที่ผ่านมาและค่าความคลาดเคลื่อนให้ $X_t, X_{t-1}, X_{t-2}, \ldots$เป็นอนุกรมเวลา

$\tilde{X}_t, \tilde{X}_{t-1}, \tilde{X}_{t-2}, \ldots$ เป็นอนุกรมเวลาที่แสดงถึงการเบี่ยงเบน (Deviation) ไปจากค่าเฉลี่ย μ เพราะฉะนั้น $\tilde{X}_t = X_t - \mu$ซึ่งจะใช้รูปแบบการถดถอยในตัวเอง (AR(p))เป็นดังนี้

$$\tilde{X}_t = \emptyset_1\tilde{X}_{t-1} + \emptyset_2\tilde{X}_{t-2} + \ldots + \emptyset_p\tilde{X}_{t-p} + a_t \quad \ldots (1)$$

ซึ่งมีพารามิเตอร์ p+2 ตัว คือ $\mu, \emptyset_1, \emptyset_2, \ldots, \emptyset_p, \sigma_a^2$

เมื่อ a_tคือ ค่าคลาดเคลื่อนของรูปแบบ หรือเรียกว่าสิ่งรบกวนอย่างสุ่ม (white Noise) ซึ่งจะมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_a^2

2.1.4.3.2 รูปแบบเฉลี่ยเคลื่อนที่(Moving Average Model : MA)เป็นรูปแบบของอนุกรมเวลาที่มีลักษณะนิ่ง โดยสามารถกระจายอยู่ในรูปของ

$$\tilde{X}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \quad \ldots (2)$$

รูปแบบนี้เรียกว่า กระบวนการเฉลี่ยเคลื่อนที่ลำดับq ใช้สัญลักษณ์MA(q) ซึ่งมีพารามิเตอร์ไม่ทราบค่าq+2 ตัว คือ$\mu, \theta_1, \theta_2, \ldots, \theta_q, \sigma_a^2$

2.1.4.3.3 รูปแบบผสม(Mixed Autoregressive Moving Average Model : ARMA)เป็นรูปแบบของอนุกรมเวลาที่มีลักษณะนิ่ง โดย X_tสามารถกระจายให้อยู่ในรูปอนุกรมเวลาที่ผ่านมา และ a_tเป็นรูปแบบผสมของ AR และ MA ดังนี้

$$\tilde{X}_t = \emptyset_1\tilde{X}_{t-1} + \emptyset_2\tilde{X}_{t-2} + \ldots + \emptyset_p\tilde{X}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \quad \ldots (3)$$

รูปแบบนี้เรียกว่า กระบวนการผสมการถดถอยในตัวเองและการเฉลี่ยเคลื่อนลำดับที่ (p,q) ใช้สัญลักษณ์ARMA(p,q)ซึ่งมีพารามิเตอร์ p+q+2ตัว คือ$\mu, \emptyset_1, \emptyset_2, \ldots, \emptyset_p, \theta_1, \theta_2, \ldots, \theta_q, \sigma_a^2$

2.1.4.3.4 รูปแบบการถดถอยในตัวเองรวมเฉลี่ยเคลื่อนที่(Autoregressive Integrated Moving Average Model : ARIMA)

เป็นรูปแบบของอนุกรมเวลาที่มีลักษณะไม่นิ่ง ในกรณีที่ X_tไม่นิ่ง ก็จะต้องใช้ผลต่างเพื่อเปลี่ยนให้ X_tเป็นอนุกรมเวลาที่มีลักษณะนิ่ง สมมติว่าอนุกรมเวลาที่มีลักษณะนิ่ง (W_t)เป็นผลต่างครั้งที่ d ของอนุกรมเวลาที่มีลักษณะไม่นิ่ง (X_t)

ดังนั้น$W_t = \nabla^d X_t = (1-B)^d X_t$ถ้า W_tสามารถกระจายอยู่ในรูปของ Wที่ผ่านมา เมื่อ $\widetilde{W}_t = W_t - \mu_w$ได้ดังนี้

$$\widetilde{W}_t = \emptyset_1\widetilde{W}_{t-1} + \emptyset_2\widetilde{W}_{t-2} + \ldots + \emptyset_p\widetilde{W}_{t-p} + a_t \quad \ldots (4)$$

รูปแบบนี้เรียกว่า กระบวนการถดถอยในตัวเองลำดับที่(p,d) (Autoregressive Integrated Process of Order (p,d)) ใช้สัญลักษณ์ARI(p,d) ถ้า W_tสามารถกระจาย ได้ดังนี้

$$\widetilde{W}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \quad \ldots (5)$$

รูปแบบนี้เรียกว่า กระบวนการรวมการเฉลี่ยเคลื่อนที่ลำดับที่(d,q) (Integrated Moving Average Process of Order (d,q)) ใช้สัญลักษณ์IMA(d,q)

และถ้า W_tสามารถกระจายอยู่ในรูปของWที่ผ่านมา และ aที่ผ่านมา ได้ดังนี้

$$\widetilde{W}_t = \emptyset_1\widetilde{W}_{t-1} + \ldots + \emptyset_p\widetilde{W}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \quad \ldots (6)$$

รูปแบบนี้เรียกว่า กระบวนการถดถอยในตัวเองรวมการเฉลี่ยเคลื่อนที่ลำดับที่ (p,d,q)ใช้สัญลักษณ์ ARIMA(p,d,q)

2.1.4.4 การวิเคราะห์อนุกรมเวลาที่มีลักษณะไม่นิ่ง (Non – Stationary Analysis)

อนุกรมเวลาที่มีลักษณะไม่นิ่ง หมายถึง อนุกรมเวลาที่ไม่อยู่ในสภาวะดุลเชิงสถิติ (Non – Statistical Equilibrium) กล่าวคือ อนุกรมเวลาที่มีลักษณะไม่นิ่งนั้น จะมีค่าเฉลี่ย $(E(X_t))$ความแปรปรวน$(V(X_t))$และคุณสมบัติอื่นๆ เปลี่ยนแปลง เมื่อเวลาเปลี่ยนไป ก่อนการกำหนดรูปแบบให้กับอนุกรมเวลาในทฤษฎีของบอกซ์ เจนกินส์ ต้องปรับอนุกรมเวลาที่มีลักษณะไม่นิ่ง ให้เป็นอนุกรมเวลาที่มีลักษณะนิ่งก่อนเสมอด้วยการหาผลต่างของแต่ละช่วงเวลาหรือผลต่างฤดูกาลจอง

อนุกรมเวลา หรือปรับโดยการแปลงข้อมูล (Transformation Data) ตามรูปแบบทางคณิตศาสตร์ เช่น แปลงในรูปลอการิทึม รากที่ n ของอนุกรมเวลา หรือในรูปเอกซ์โพเนนเชียล เป็นต้น แต่โดยส่วนมากจะปรับอนุกรมเวลาที่มีลักษณะไม่นิ่ง ให้เป็นอนุกรมเวลาที่มีลักษณะนิ่ง โดยการหาผลต่างครั้งที่ dการหาผลต่างครั้งที่ dของอนุกรมเวลา แทนสัญลักษณ์∇ เป็นผลต่างครั้งที่ 1∇^2 เป็นผลต่างครั้งที่ 2และ ∇^dเป็นผลต่างครั้งที่ dเมื่อ $\nabla\tilde{X}_t = \tilde{X}_t - \tilde{X}_{t-1}$ถ้าเขียนในรูปถอยหลัง(Back ward shift Operater)จะแทนด้วย$B\tilde{X}_t = \tilde{X}_{t-1}$และ$B^m\tilde{X}_t = \tilde{X}_{t-m}$ดังนั้น

$$\nabla^d = (1-B)^d \qquad \ldots(7)$$

ให้ $\tilde{W}_t = (1-B)^d\tilde{X}_t$ เป็นอนุกรมเวลาที่มีลักษณะนิ่ง โดยการหาผลต่างครั้งที่ dของอนุกรมเวลาที่มีลักษณะไม่นิ่ง (X_t)ในการกำหนดรูปแบบให้กับอนุกรม W_tก็ทำเช่นเดียวกับอนุกรมเวลาที่มีลักษณะนิ่ง ซึ่งเรียกรูปแบบของ W_tว่า ARIMA(p,d,q)

รูปแบบ ARIMA(p,d,q) เขียนได้เป็น

$$\tilde{W}_t = \emptyset_1\tilde{W}_{t-1} + \ldots + \emptyset_p\tilde{W}_{t-p} + a_t - \theta_1 a_{t-1} - a_{t-2} - \ldots - \theta_q a_{t-q} \qquad \ldots(8)$$

สามารถกระจายให้อยู่ในรูปอนุกรม X_t ได้เป็น

$$\tilde{W}_t - \emptyset_1 B\tilde{W}_t - \ldots - \emptyset_p B^p\tilde{W}_t = a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \ldots - \theta_q B^q a_t \qquad \ldots(9)$$

$$(1 - \emptyset_1 B - \ldots - \emptyset_p B^p)\tilde{W}_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)a_t \qquad \ldots(10)$$

$$(1 - \emptyset_1 B - \ldots - \emptyset_p B^p)(1-B)^d\tilde{X}_t = (1 - \theta_1 B - \ldots - \theta_q B^q)a_t \qquad \ldots(11)$$

$$\Phi(B)(1-B)^d\tilde{X}_t = \Theta(B)a_t \qquad \ldots(12)$$

การพิจารณาอนุกรมเวลาว่ามีลักษณะนิ่งหรือไม่นิ่งนั้น พิจารณาได้จากวิธีต่างๆ ดังนี้

วิธีที่1. นำอนุกรมเวลามาเขียนกราฟ พิจารณาลักษณะของกราฟว่ามีแนวโน้มและมีฤดูกาลปรากฏหรือไม่ ถ้ามีแสดงว่าเป็นอนุกรมเวลาที่มีลักษณะไม่นิ่ง

วิธีที่2. แบ่งอนุกรมเวลาเป็นช่วงๆ แล้วนำอนุกรมเวลาของแต่ละช่วงไปทดสอบค่าเฉลี่ย ความแปรปรวนและความแปรปรวนร่วมตามทฤษฎีทางสถิติ

วิธีที่3.พิจารณาคลอเรลโรแกรมของ r_k ในกรณีที่อนุกรมเวลามีลักษณะนิ่ง (Stationary) r_k จะมีค่าลดลงอย่างรวดเร็วแบบแกซ์โพเนนเชียล เมื่อ kมีค่าเพิ่มขึ้น ส่วนกรณีอนุกรมเวลาที่มีลักษณะไม่นิ่ง (Non – Stationary)r_k จะมีค่าลดลงค่อนข้างช้า สังเกตได้ว่าอนุกรมเวลานั้นมีแนวโน้มและจะมีค่าเฉลี่ยไม่เท่ากัน

ถ้า r_k ของอนุกรมเวลามีค่าลดลงอย่างช้าๆ เมื่อ kมีค่าเพิ่มขึ้น และ r_k มีค่าสูง เมื่อ k = L, 2L, 3L, …เมื่อ Lเป็นคาบฤดูกาล แสดงข้อมูลมีแนวโน้มและฤดูกาล แต่ถ้า r_k มีค่าสูงเป็นลูกคลื่นครบรอบ L ช่วงเวลา แสดงว่าอนุกรมเวลามีฤดูกาล

2.1.4.5 การกำหนดรูปแบบ (Model Identification)

ในการหารูปแบบจะพิจารณาจากฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function : ACF) และฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial Autocorrelation Function : PACF) ที่มีลักษณะลดลงหรือมีค่าเป็น 0 หลังจากช่วงห่าง p หรือ q หน่วยเวลา โดยใช้ตัวสถิติ t ในการทดสอบสมมติฐานว่ามีค่าสหสัมพันธ์เป็นศูนย์หรือไม่ รูปแบบที่สำคัญในการวิเคราะห์อนุกรมเวลามี 5 รูปแบบ ดังนี้

ARIMA	รูปแบบที่ 1 (1,d,0)	รูปแบบที่ 2 (0,d,1)
ลักษณะของ$\boldsymbol{\rho_k}$	มีค่าลดลงแบบเอกซ์โพเนนเชียล 1.ในเครื่องหมายเดียวกัน ถ้า $\emptyset_1 > 0$ 2.สลับเครื่องหมาย ลบ บวก โดยเริ่มลบก่อน ถ้า$\emptyset_1 < 0$	มีเพียงค่าเดียว คือ $\rho_1 \neq 0$ ส่วน $\rho_k = 0; k \geq 2$ 1.$\rho_1 > 0, \theta_1 < 0$ 2.$\rho_1 < 0, \theta_1 > 0$
ลักษณะของ $\emptyset_{kk}$	มีเพียงค่าเดียว คือ $\emptyset_{11} \neq 0$ ส่วน $\emptyset_{kk} = 0\,; k \geq 2$ 1.$\emptyset > 0$ถ้า$\emptyset_1 > 0$ 2.$\emptyset_{11} < 0$ถ้า$\emptyset_1 < 0$	มีค่าลดลงแบบเอกซ์โพเนนเชียล 1.สลับเครื่องหมาย บวก ลบ ถ้า $\theta_1 < 0$ 2.เครื่องหมายลบ ถ้า$\theta_1 > 0$
ค่าประมาณพารามิเตอร์	$\emptyset_1 = \rho_1$	$\rho_1 = \dfrac{-\theta_1}{1+\theta_1^2}$
ลักษณะนิ่ง	$-1 < \emptyset_1 < 1$	$-1 < \theta_1 < 1$

ARIMA	รูปแบบที่ 3 (2,d,0)	รูปแบบที่ 4 (0,d,2)
ลักษณะของ$\boldsymbol{\rho_k}$	มีค่าลดลงแบบเอกซ์โพเนนเชียล ในเครื่องหมายเดียวกันหรือต่างเครื่องหมาย	มี2ค่า คือ $\rho_1 \neq 0$และ $\rho_2 \neq 0$ ส่วน$\rho_k = 0; k \geq 3$
ลักษณะของ $\emptyset_{kk}$	มีค่า2ค่า คือ$\emptyset_{11} \neq 0, \emptyset_{22} \neq 0$ ส่วน$\emptyset_{11} = 0\,; k \geq 3$	มีค่าลดลงแบบเอกซ์โพเนนเชียลในเครื่องหมายเดียวกันหรือต่างเครื่องหมาย
ค่าประมาณพารามิเตอร์	$\emptyset_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2}$ $\emptyset_2 = \frac{\rho_2(1-\rho_1^2)}{1-\rho_1^2}$	$\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$ $\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$
ลักษณะนิ่ง	$-1 < \emptyset_2 < 1$ $\emptyset_2 + \emptyset_1 < 1$ $\emptyset_2 - \emptyset_1 < 1$	$-1 < \theta_2 < 1$ $\theta_2 + \theta_1 < 1$ $\theta_2 - \theta_1 < 1$

ARIMA	รูปแบบที่ *5 (1,d,1)*
ลักษณะของ $\boldsymbol{\rho_k}$	*มีค่าลดลงแบบเอกซ์โพเนนเชียลหลังจากช่วงห่าง 1หน่วยเวลา* 1.*ในเครื่องหมายของ* ρ_1*ถ้า* $\rho_1 = \emptyset_1 - \theta_1$ 2.*ในเครื่องหมายเดียวกัน ถ้า* $\emptyset_1 > 0$ 3.*สลับเครื่องหมาย ถ้า* $\emptyset_1 < 0$
ลักษณะของ $\emptyset_{kk}$	*มีค่าลดลงแบบเอกซ์โพเนนเชียลหลังจากช่วงห่าง 1 หน่วยเวลา* 1.$\emptyset_{11} = \rho_1$ 2.*ในเครื่องหมายเดียวกัน ถ้า* $\theta_1 > 0$ 3.*สลับเครื่องหมาย ถ้า*$\theta_1 < 0$
ค่าประมาณพารามิเตอร์	$\rho_1 = \frac{(1-\theta_1\emptyset_1)(\emptyset_1-\theta_1)}{1+\theta_1^2+2\emptyset_1\theta_1}$ $\rho_2 = \emptyset_1\rho_1$
ลักษณะนิ่ง	$-1 < \emptyset_1 < 1$ $-1 < \theta_1 < 1$

2.1.4.6 การประมาณค่าพารามิเตอร์(Parameters Estimator)

การประมาณค่าพารามิเตอร์ของรูปแบบที่นิยมใช้กันอย่างกว้างขวาง คือ วิธีแม็คซิมัมไลลิฮูด (Maximum Likelihood) แต่ในกรณีที่จำนวนตัวอย่างมีขนาดใหญ่ การใช้วิธีแม็คซิมัมไลลิฮูดจะใช้เวลาในการคำนวณมาก การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยที่สุด ก็จะได้ผลใกล้เคียงกันการคำนวณ มี 2 วิธี ดังนี้

2.1.4.6.1 การคำนวณแบบมีเงื่อนไข

ถ้าอนุกรมเวลามีรูปแบบเป็น ARIMA(p,d,q) จากอนุกรมเวลา X_tเมื่อทำการหาผลต่าง d ครั้ง จะได้อนุกรมเวลW_tเป็น $W_1, W_2, W_3, \dots, W_n (W_t = \nabla^d X_t)$ และจะมีรูปแบบเป็น

$$\widetilde{W}_t = \emptyset_1 \widetilde{W}_{t-1} + \emptyset_2 \widetilde{W}_{t-2} + \dots + \emptyset_p \widetilde{W}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad \dots(13)$$

$$a_t = \widetilde{W}_t - \emptyset_1 \widetilde{W}_{t-1} - \emptyset_2 \widetilde{W}_{t-2} - \dots + \emptyset_p \widetilde{W}_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad \dots(14)$$

จากสมการข้างต้น จะคำนวณหาค่า a_t ได้ จะต้องทราบค่า W จำนวน p ตัว และค่า a จำนวน q ตัวกำหนดให้ $W_1, W_2, W_3, \dots, W_p = W_*$; $a_1, a_2, a_3, \dots, a_q = A_*$ $\emptyset_1, \emptyset_2, \emptyset_3, \dots, \emptyset_p = \Phi$; $\theta_1, \theta_2, \theta_3, \dots, \theta_q = \Theta$

ดังนั้นเมื่อทราบค่า W_*และ A_*และสามารถประมาณค่า Φและ Θค่าของ a_t ก็สามารถหาค่าได้ โดยใช้สัญลักษณ์ $a_t(\Phi, \Theta, | W_*, A_*, W)$เมื่อ t = 1, 2, 3,…, n

กำหนดให้ $S_*(\Phi, \Theta) = \sum_{t=1}^{n} a_t^2 (\Phi, \Theta, | W_*, A_*, W)$

$S_*(\Phi, \Theta)$เรียกว่าผลบวกกำลังสองของค่าคลาดเคลื่อน (Sum of Squared Residual) ในการคำนวณ a_tแบบมีเงื่อนไขนั้นจะกำหนดให้ $W_* = 0$และ $A_* = 0$

2.1.4.6.2 การคำนวณแบบไม่มีเงื่อนไข

กำหนดให้$S(\Phi, \Theta) = \sum_{t=-\alpha}^{n} [a_t | \Phi, \Theta, W]^2$เป็นผลบวกของกำลังสองของค่าคลาดเคลื่อนแบบไม่มีเงื่อนไข เมื่อ$[a_t | \Phi, \Theta, W] = E[a_t | (\Phi, \Theta, W)]$

การคำนวณหา$S(\Phi, \Theta)$จะต้องทราบค่าของ W_*และ A_*การหา W_*และ A_*จะใช้วิธีการคำนวณย้อนหลังจากสมการรูปแบบ

$$\widetilde{W}_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 B) a_t$$

เปลี่ยนเป็นสมการย้อนหลัง

$$\widetilde{W}_t = (1 - \theta_1 F) e_t = e_t - \theta_1 e_{t+1} \quad \dots(15)$$

เมื่อ $FX_t = X_{t+1}$และ$F^m X_t = X_{t+m}$

ถ้า $E(W_t) = 0$จะได้$e_t = W_t + \theta_1 e_{t+1}$และ$a_t = W_t + \theta_1 a_{t+1}$
การคำนวณ e_1จะกำหนดให้ $e_{n+1} = 0$และ$e_j = 0, j = 0, -1, -2, \ldots$

2.1.4.7 การตรวจสอบรูปแบบ (Diagnostic Checks)

เมื่อได้รูปแบบและค่าพารามิเตอร์ที่ประมาณได้จะถูกนำมาตรวจสอบเพื่อดูว่ารูปแบบและค่าพารามิเตอร์ที่ได้นั้นมีความเหมาะสมในการนำไปใช้ในการพยากรณ์หรือไม่
โดยตรวจสอบสมมุติฐานของค่าความคลาดเคลื่อน โดยนำค่าความคลาดเคลื่อนมาทดสอบถึงคุณสมบัติดังต่อไปนี้

(1) ตรวจสอบสมมุติฐานของค่าความคลาดเคลื่อนว่ามีค่าเฉลี่ยเท่ากับศูนย์หรือไม่

ทดสอบ $H_0 : \mu_\varepsilon = 0$

$H_1 : \mu_\varepsilon \neq 0$

โดยใช้ตัวสถิติทดสอบ t ดังนี้

$$\mathrm{t} = \frac{\bar{e}-\mu_\varepsilon}{\frac{s}{\sqrt{n}}}$$ มีองศาความเป็นอิสระเท่ากับ n-1

เมื่อ $\bar{e}$คือค่าเฉลี่ยของค่าความคลาดเคลื่อนและมีการแจกแจงแบบปกติ

μ_εคือค่าเฉลี่ยของประชากรของค่าความคลาดเคลื่อน

s คือส่วนเบี่ยงเบนมาตรฐานของค่าความคลาดเคลื่อน

nคือจำนวนค่าความคลาดเคลื่อน

ถ้าค่าของตัวสถิติ t ที่คำนวณได้ตกอยู่ในอาณาเขตวิกฤติแสดงว่าค่าเฉลี่ยของค่าความคลาดเคลื่อนไม่เท่ากับศูนย์

(2) ตรวจสอบสมมุติฐานของค่าความคลาดเคลื่อนว่ามีค่าความแปรปรวนคงที่หรือไม่ (Richard A. Johnson, Dean W. Wichern, 2002)

ทดสอบ H_0: ค่าความคลาดเคลื่อนและค่าพยากรณ์ไม่มีความสัมพันธ์กัน

H_1: ค่าความคลาดเคลื่อนและค่าพยากรณ์มีความสัมพันธ์กัน

โดยใช้ตัวสถิติทดสอบ Z ดังนี้

$$Z = \frac{\rho - \mu_\rho}{\sigma_\rho} = \frac{\rho}{\sigma_\rho}$$

โดย $\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$ และมีการแจกแจงแบบปกติ

$n \geq 10$

$n_\rho = 0$

$$\sigma_\rho = \sqrt{\frac{1}{n-1}}$$

$$\mathrm{d_i} = \mathrm{x_i} - \mathrm{y_i}$$

เมื่อ x_iคือ ลำดับที่ Iของค่าพยากรณ์y_iคือ ลำดับที่ *i*ของค่าความคลาดเคลื่อน ถ้าค่าของตัวสถิติ Z ที่คำนวณได้ตกอยู่ในอาณาเขตวิกฤติแสดงว่าค่าความคลาดเคลื่อนและค่าพยากรณ์มีความสัมพันธ์กันนั่นคือค่าความแปรปรวนของค่าความคลาดเคลื่อนไม่คงที่

(3) ตรวจสอบสมมุติฐานของค่าคลาดเคลื่อนว่ามีความเป็นอิสระกันหรือไม่

ทดสอบ $H_0 : \rho_k(\hat{e}_t) = 0$

$H_1 : \rho_k(\hat{e}_t) \neq 0$

เป็นการทดสอบความเป็นอิสระของแต่ละช่วงห่างเวลาของค่าคลาดเคลื่อนในการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อนจะคำนวณได้ดังนี้

$$\mathrm{r_k}(\hat{\mathrm{e}}_\mathrm{t}) = \frac{\sum_{\mathrm{t=1}}^{\mathrm{n-k}}(\hat{\mathrm{e}}_\mathrm{t} - \bar{\mathrm{e}})(\hat{\mathrm{e}}_\mathrm{t+k} - \bar{\mathrm{e}})}{\sum_{\mathrm{t=1}}^{\mathrm{n-k}}(\hat{\mathrm{e}}_\mathrm{t} - \bar{\mathrm{e}})^2}$$

เมื่อ $r_\mathrm{k}(\hat{e}_t)$เป็นสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อนที่ห่างกัน k หน่วยเวลา และมีการแจกแจงแบบสติวเดนท์ที (Student's t distribution)

การทดสอบจะใช้สถิติ t ดังนี้

$$\mathrm{t} = \frac{\mathrm{r_k}(\hat{\mathrm{e}}_\mathrm{t}) - 0}{\mathrm{SE}(\mathrm{r_k}(\hat{\mathrm{e}}_\mathrm{t}))}$$ มีองศาความเป็นอิสระเท่ากับ n-1

เมื่อ $\mathrm{SE}(r_\mathrm{k}(\hat{e}_t))$เป็นส่วนเบี่ยงเบนมาตรฐานของสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อน ถ้ายอมรับ $H_0 : \rho_k(\hat{e}_t) = 0$แสดงว่าค่าคลาดเคลื่อนเป็นอิสระกัน

(4) ตรวจสอบสมมุติฐานของค่าคลาดเคลื่อนว่ามีการแจกแจงแบบปกติหรือไม่

ทดสอบH_0 :ค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ

H_1 :ค่าคลาดเคลื่อนไม่มีการแจกแจงแบบปกติ

ใช้ตัวสถิติทดสอบคือ

$$\mathrm{T_1} = \max|\mathrm{F^*(e)} - \mathrm{S(e)}|$$

เมื่อ $\mathrm{F^*(e)}$คือฟังก์ชันการแจกแจงเฉพาะของ*e*

เมื่อ $S(e)$คือการแจกแจงความถี่สะสม*e*

มีข้อตกลงเบื้องต้นคือตัวอย่างถูกเลือกมาโดยวิธีสุ่มและ$F^*(e)$ฟังก์ชันการแจกแจงเฉพาะของ e แบบต่อเนื่องให้ค่าของตัวสถิติT_1เป็นระยะทางที่มากที่สุดระหว่าง$S(e)$และ$F^*(e)$โดยจะเขียน max แทน maximum (ค่าห่างที่สุด)จะปฏิเสธH_0ที่ระดับนัยสำคัญαถ้าค่าของตัวสถิติT_1มีค่าเกินค่า $W_{1-\alpha}$

2.1.4.8 การพยากรณ์ (Forecasting)

เมื่อกำหนดรูปแบบของอนุกรมเวลาและหาค่าพารามิเตอร์ของรูปแบบที่ทำให้เกิดค่าคลาดเคลื่อนน้อยที่สุดแล้วก็จะใช้รูปแบบที่ได้ทำการพยากรณ์ข้อมูลในอนาคตซึ่งการพยากรณ์นั้นทำได้ 2 แบบคือการพยากรณ์แบบค่าเดียว (Point Forecasts or Single Numerical Values)และการพยากรณ์แบบช่วง (Interval Forecasts)

2.1.4.8.1 การพยากรณ์แบบค่าเดียว

การพยากรณ์นี้เราจะสมมติว่าเราทราบรูปแบบค่าพารามิเตอร์ค่าอนุกรมเวลาและค่าความคลาดเคลื่อนแล้วทำการพยากรณ์จากสิ่งที่ทราบเหล่านี้สมมติว่าอนุกรมเวลา t ช่วงเวลาคือ $X_t, X_{t-1}, X_{t-2}, \ldots$และทราบค่าคลาดเคลื่อน tช่วงเวลา คือ $\hat{a}_t, \hat{a}_{t-1}, \hat{a}_{t-2}, \ldots$ต้องการพยากรณ์ข้อมูลที่เวลา $t+L$เมื่อ $L \geq 1$

เวลา t เรียกว่า จุดเริ่มต้น (Origin)

Lเรียกว่า หน่วยเวลาล่วงหน้า (Lead time)

ค่าพยากรณ์ของ X_{t+L} ใช้สัญลักษณ์ $\hat{X}_t(L)$เป็นค่าคาดหวังที่มีเงื่อนไขของ X_{t+L}ดังนั้น

$\hat{X}_t(L) = E(X_{t+L}|I_t)$เมื่อ I_tเป็นอนุกรมเวลา$X_t, X_{t-1}, X_{t-2}, \ldots$

$$E(X_{t+L}) = \begin{cases} X_{t+L} & ; L \leq 0 \\ \hat{X}_t(L) & ; L \geq 1 \end{cases}$$

$$E(a_{t+L}) = \begin{cases} a_{t+L} & ; L \leq 0 \\ 0 & ; L \geq 1 \end{cases}$$

สำหรับรูปแบบการพยากรณ์ของ AR(p), MA(q), ARMA(p,q) แสดงรูปแบบพยากรณ์ดังนี้

รูปแบบ AR(1)$X_t = \emptyset_1 X_{t-L} + \theta_0 + a_t$

รูปแบบการพยากรณ์คือ

$$\hat{X}_t(L) = \hat{\emptyset}_1 X_{t-1+L} + \hat{\theta}_0; \qquad L = 1,2,3, \ldots$$

เช่น $\hat{X}_t(1) = \hat{\emptyset}_1 X_t + \hat{\theta}_0 \quad ; \qquad L = 1$

$$\hat{X}_t(L) = \hat{\emptyset}_1 X_t(L-1) + \hat{\theta}_0 ; L \geq 2$$

รูปแบบ AR(2) $\quad X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \theta_0 + a_t$

รูปแบบการพยากรณ์คือ

$$\hat{X}_t(L) = \hat{\emptyset}_1 X_{t-1+L} + \hat{\emptyset}_2 X_{t-2+L} \hat{\theta}_0 \quad ; L = 1,2,3, \ldots$$

เช่น

$$\hat{X}_t(1) = \hat{\emptyset}_1 X_t + \hat{\emptyset}_2 X_{t-1} + \hat{\theta}_0 \quad ; L = 1$$
$$\hat{X}_t(2) = \hat{\emptyset}_1 \hat{X}_t(1) + \hat{\emptyset}_2 X_t + \hat{\theta}_0 \quad ; L = 2$$
$$\hat{X}_t(L) = \hat{\emptyset}_1 \hat{X}_t(L-1) + \hat{\emptyset}_2 \hat{X}_t(L-2) + \hat{\theta}_0 \; ; L \geq 3$$

รูปแบบ MA(1) $X_t = \theta_0 + a_t - \theta_1 a_{t-1}$

รูปแบบการพยากรณ์คือ$\hat{X}_t(\mathrm{L}) = \hat{\theta}_0 + \hat{a}_{t+L} - \hat{\theta}_1 a_{t-1+L}; L = 1,2,3,\ldots$

เช่น

$$\hat{X}_t(1) = \hat{\theta}_0 - \hat{\theta}_1 a_t \quad ; \quad L = 1$$
$$\hat{X}_t(\mathrm{L}) = \hat{\theta}_0 \quad ; \quad L \geq 2$$

รูปแบบ MA(2) $X_t = \theta_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$

รูปแบบการพยากรณ์คือ

$$\hat{X}_t(\mathrm{L}) = \hat{\theta}_0 + \hat{a}_{t+L} - \hat{\theta}_1 a_{t-1+L} - \hat{\theta}_2 a_{t-2+L} \; ; L = 1,2,3,\ldots$$

เช่น

$$\hat{X}_t(1) = \hat{\theta}_0 - \hat{\theta}_1 a_t - \hat{\theta}_2 a_{t-1} \quad ; L = 1$$
$$\hat{X}_t(2) = \hat{\theta}_0 - \hat{\theta}_2 a_t \quad ; \quad L = 2$$
$$\hat{X}_t(L) = \hat{\theta}_0 \quad ; L \geq 3$$

รูปแบบ ARMA(1,1) $X_t = \emptyset_1 X_{t-1} + \theta_0 + a_t - \theta_1 a_{t-1}$

รูปแบบการพยากรณ์คือ

$$\hat{X}_t(\mathrm{L}) = \emptyset_1 X_{t-1+L} + \hat{\theta}_0 + a_{t+L} - \hat{\theta}_1 a_{t-1+L} \quad ; L = 1,2,3,\ldots$$

เช่น

$$\hat{X}_t(1) = \hat{\emptyset}_1 X_t + \hat{\theta}_0 - \hat{\theta}_1 a_t \quad ; \quad L = 1$$
$$\hat{X}_t(\mathrm{L}) = \hat{\emptyset}_1 \hat{X}_t(L-1) + \hat{\theta}_0 \quad ; \quad L \geq 2$$

2.1.4.8.2 การพยากรณ์แบบช่วง (Interval Forecasts)

จากรูปแบบ ARIMA(p,d,q) จะสามารถเปลี่ยนให้อยู่ในรูปของ MA(α) ได้

$$\emptyset(B)\widetilde{w}_t = \theta(B)a_t$$
$$\widetilde{W}_t = \emptyset^{-1}(B)\theta(B)a_t$$
$$\widetilde{X}_t = (1-B)^{-d}\emptyset^{-1}(B)\theta(B)a_t \quad \ldots(16)$$

สมมติว่ารูปแบบของ MA(α) เป็น

$$X_t = \mu + \Psi_0 a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \ldots \quad \ldots(17)$$

เมื่อ $\Psi_0 = 1$ดังนั้นจาก ARIMA(p,d,q)เมื่อ เปลี่ยนให้อยู่ในรูปแบบ(16) และเปรียบเทียบกับ(2.17)ก็สามารถหาค่า $\mu_1 \Psi_j \; : j = 0,1,2,\ldots$เช่น AR(1)

$$X_t = \emptyset_1 X_{t-1} + \theta_0 + a_t$$
$$X_t = \frac{\theta_0}{1-\emptyset_1 B} + (1-\emptyset_1 B)^{-1} a_t$$
$$= \frac{\theta_0}{1-\emptyset_1 B} + (1 + \emptyset_1 B + \emptyset_1^2 B^2 + \emptyset_1^3 B^3 + \ldots) a_t$$
$$= \frac{\theta_0}{(1-\emptyset_1)} + a_t + \emptyset_1 a_{t-1} + \emptyset_1^2 a_{t-2} + \emptyset_1^3 a_{t-3} + \ldots$$

ดังนั้น $\mu = \frac{\theta_0}{(1-\emptyset_1)}$, $\Psi_0 = 1, \Psi_j = \emptyset\{j = 1,2,3, \ldots$ …(18)

สำหรับ ARMA(1,1)

$$X_t = \emptyset_1 X_{t-1} + \theta_0 + a_t - \theta_1 a_{t-1}$$
$$X_t = \frac{\theta_0}{1-\emptyset_1 B} + (1-\theta_1 B)^{-1} a_t$$
$$= \frac{\theta_0}{1-\emptyset_1 B} + [(1+\theta_1 B)(1 + \emptyset_1 B + \emptyset_1^2 B^2 + \ldots)] a_t$$
$$X_t = \frac{\theta_0}{1-\emptyset_1 B} + [(1 + (\emptyset_1 - \theta_1) B + \emptyset_1^2 - \emptyset_1\theta_1) + (+\emptyset_1^3 - \emptyset_1^2\theta_1) B^3 + \ldots)] a_t \quad \ldots(19)$$

เมื่อเปรียบเทียบกับ MA(α) จะได้

$$\mu = \frac{\theta_0}{1-\emptyset_1}$$
$$\Psi_0 = 1$$
$$\Psi_1 = \emptyset_1 - \theta_1$$
$$\Psi_1 = \emptyset_1(\emptyset_1 - \theta_1)$$
$$\vdots$$
$$\Psi_j = \emptyset_1^{j-1}(\emptyset_1 - \theta_1)$$

ถ้า $\emptyset_1 = .62$และ $\theta_1 = -.58$

$$\Psi_0 = 1$$
$$\Psi_1 = 1.2$$
$$\Psi_1 = .74$$
$$\vdots$$

ให้ $e_t(l)$ เป็นค่าคลาดเคลื่อนของการพยากรณ์ล่วงหน้า Lหน่วยเวลา ณ ที่เวลา t

$$e_t(L) = X_{t+1} - \hat{X}_t(L)$$

จากรูปแบบของ MA(α)

$$X_t = \mu + \Psi_0 a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \ldots \quad \ldots(20)$$

เมื่อทำการพยากรณ์ล่วงหน้า Lหน่วยเวลาจะได้

$$\hat{X}_t(L) = \mu + \Psi_L a_t + \Psi_{L+1} a_{t-1} + \Psi_{L+2} a_{t-2} + \Psi_{L+3} a_{t-3} + \ldots \quad \ldots(21)$$

ดังนั้น $e_t(L) = X_{t+L} - \hat{X}_t(L)$

$$= \Psi_0 a_{t+L} + \Psi_1 a_{t+L-1} + \Psi_2 a_{t+L-2} + \cdots + \Psi_{L-1} a_{t+1} \quad \ldots(22)$$

ความแปรปรวน

$$\sigma^2[e_t(L)] = \sigma_a^2(1 + \Psi_1^2 + \Psi_2^2 + \ldots + \Psi_{l-1}^2) \qquad \ldots(23)$$

ส่วนเบี่ยงเบนมาตรฐาน

$$\sigma[e_t(L)] = \sigma_a(1 + \Psi_1^2 + \Psi_2^2 + \ldots + \Psi_{L-1}^2)^{\frac{1}{2}} \qquad \ldots(24)$$

2.1.4.9 ตัวแบบการปรับแก้ค่าคลาดเคลื่อน (Error Correction Model)

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่งหมายถึงการที่ข้อมูลอนุกรมเวลาอยู่ในสภาพของการสมดุลเชิงสถิติ (Statistical Equilibrium)ซึ่งหมายถึงการที่ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงถึงแม้เวลาจะเปลี่ยนไปแสดงได้ดังนี้กำหนดให้ $X_t, X_{t+1}, X_{t+2}, \ldots, X_{t+k}$ และ $X_{t+m}, X_{t+m+1}, X_{t+m+2}, \ldots, X_{t+m+k}$เป็นข้อมูลอนุกรมเวลาที่เวลา t, t+1, t+2, …, t+kและ เวลา t+m, t+m+1, t+m+2, …, t+m+kตามลำดับ

$P(X_t, X_{t+1}, X_{t+2}, \ldots, X_{t+k})$และ$P(X_{t+m}, X_{t+m+1}, X_{t+m+2}, \ldots, X_{t+m+k})$เป็นการแจกแจงความน่าจะเป็นร่วมขอ$X_t, X_{t+1}, X_{t+2}, \ldots, X_{t+k}$ และ$X_{t+m}, X_{t+m+1}, X_{t+m+2}, \ldots, X_{t+m+k}$ ตามลำดับ กล่าวได้ว่า X_t เป็นข้อมูลอนุกรมเวลาที่มีลักษณะนิ่งแบบเข้ม (Strictly Stationary) เมื่อ $P(X_t, X_{t+1}, X_{t+2}, \ldots, X_{t+k}) = P(X_{t+m}, X_{t+m+1}, X_{t+m+2}, \ldots, X_{t+m+k})$

ถ้า $\mathrm{P}(\mathrm{X}_t, \mathrm{X}_{t+1}, \mathrm{X}_{t+2}, \ldots, \mathrm{X}_{t+k}) \neq \mathrm{P}(\mathrm{X}_{t+m}, \mathrm{X}_{t+m+1}, \mathrm{X}_{t+m+2}, \ldots, \mathrm{X}_{t+m+k})$

สรุปได้ว่าข้อมูลอนุกรมเวลาดังกล่าวมีลักษณะไม่นิ่งแบบเข้ม

ในทางปฏิบัตินิยมใช้ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่งแบบอ่อนมากกว่านิ่งแบบเข้มเพราะสะดวกและง่ายต่อการทดสอบกล่าวคือX_tจะเป็นอนุกรมเวลาที่มีลักษณะนิ่งแบบอ่อนเมื่อ

ค่าเฉลี่ยคงที่ : $E(X_t) = \mu$

ค่าความแปรปรวนคงที่ :$Var(X_t) = \sigma^2$

ค่าความแปรปรวนร่วมคงที่: $E[(X_t - \mu)(X_{t+k} - \mu)] = \gamma_k$

ถ้าหากอนุกรมเวลาขาดคุณสมบัติข้อใดข้อหนึ่งถือได้ว่าอนุกรมเวลาดังกล่าวนั้นมีลักษณะที่ไม่นิ่งโดยการตรวจสอบข้อมูลอนุกรมเวลาว่ามีลักษณะนิ่งหรือไม่นิ่งนั้นสามารถใช้การทดสอบยูนิทรูท

2.1.4.10 การทดสอบยูนิทรูท (Unit Roots Test)

การทดสอบยูนิทรูทเป็นการตรวจสอบข้อมูลอนุกรมเวลาว่ามีลักษณะข้อมูลเป็นแบบนิ่งหรือไม่นิ่งโดยดิกกี-ฟลูเลอร์ (Dickey-Fuller) ซึ่งเป็นขั้นตอนแรกในการศึกษาตัวแบบการปรับแก้ค่าคลาดเคลื่อนโดยดิกกี-ฟลูเลอร์ได้สมมุติแบบจำลองเป็นดังนี้

$$X_t = \rho X_{t-1} + e_t \qquad \text{…(25)}$$

โดยที่ X_t, X_{t-1}คือข้อมูลอนุกรมเวลาของตัวแปรอิสระณเวลา t และ
e_tคือความคลาดเคลื่อนเชิงสุ่ม (Random Error)

ρคือสัมประสิทธิ์สหสัมพันธ์ในตัวเองซึ่ง ρมีค่าตั้งแต่−1 ถึง 1

โดยมีสมมุติฐานของการทดสอบคือ

$H_0 \ : \ |\rho| = 1$ (X_tมียูนิทรูทหรือX_tมีลักษณะไม่นิ่ง)

$H_1 \ : \ |\rho| < 1$ (X_tไม่มียูนิทรูทหรือX_tมีลักษณะนิ่ง)

โดยใช้ตัวสถิติทดสอบคือ

$$t = \frac{\hat{\rho} - 1}{S.E.(\hat{\rho})}$$

โดยที่ $\hat{\rho}$ ประมาณค่าได้จากวิธีกำลังสองน้อยที่สุด

ซึ่ง $\hat{\rho} = (X'_{t-1}X_{t-1})^{-1}X'_{t-1}X_t$และ $S.E.\hat{\rho} = (X'_{t-1}X_{t-1})^{-1}\sigma^2$

เมื่อเปรียบเทียบค่าของตัวสถิติ t ที่คำนวณได้กับค่าในตารางดิกกี-ฟลูเลอร์ถ้าค่าของตัวสถิติ t น้อยกว่าค่าวิกฤติในตารางดิกกี-ฟลูเลอร์จะปฏิเสธสมมุติฐานซึ่งกล่าวได้ว่าตัวแปรที่นำมาทดสอบมีลักษณะนิ่งอย่างไรก็ตามการทดสอบยูนิทรูทดังกล่าวข้างต้นสามารถทำได้อีกวิธีหนึ่งคือ

ให้ $\rho = (1 + \theta) \, ; \, -1 < \theta < 0$ …(26)

โดยที่ θ คือ พารามิเตอร์

จะได้ $X_t = (1 + \theta)X_{t-1} + e_t$ …(27)

$X_t = X_{t-1} + \theta X_{t-1} + e_t$ … (28)

$X_t - X_{t-1} = \theta X_{t-1} + e_t$ … (29)

$\Delta X_t = \theta X_{t-1} + e_t$ … (30)

จะได้สมมุติฐานการทดสอบของดิกกี-ฟลูเลอร์ใหม่คือ

$H_0 \ : \theta = 0$ (X_tมียูนิทรูทหรือX_tมีลักษณะไม่นิ่ง)

$H_1 : \theta = 0$ (X_tไม่มียูนิทรูทหรือX_tมีลักษณะนิ่ง)

โดยใช้ตัวสถิติทดสอบคือ

$$t = \frac{\hat{\theta}}{S.E.(\hat{\theta})}$$

โดยที่ $\hat{\theta}$ ประมาณค่าได้จากวิธีกำลังสองน้อยที่สุด

ซึ่ง $\hat{\theta} = (X'_{t-1}X_{t-1})^{-1}X'_{t-1}\Delta X_t$ และ $S.E.(\hat{\theta}) = (X'_{t-1}X_{t-1})^{-1}\sigma^2$

เมื่อเปรียบเทียบค่าของตัวสถิติ t ที่คำนวณได้กับค่าในตารางดิกกี-ฟลูเลอร์ถ้าค่าของตัวสถิติ t น้อยกว่าค่าวิกฤติในตารางดิกกี-ฟลูเลอร์จะปฏิเสธสมมุติฐานซึ่งกล่าวได้ว่าตัวแปรที่นำมาทดสอบมีลักษณะนิ่ง

เนื่องจากข้อมูลอนุกรมเวลาณเวลา t มีส่วนสัมพันธ์กับข้อมูลอนุกรมเวลาณเวลา t-1ค่าคงที่และแนวโน้มดังนั้นจะพิจารณาสมการถดถอย 3 รูปแบบที่แตกต่างกันในการทดสอบว่ามียูนิทรูทหรือไม่ซึ่งสมการทั้งสามคือ

$$\Delta X_t = \theta X_{t-1} + e_t \qquad \ldots(31)$$

$$\Delta X_t = \alpha + \theta X_{t-1} + e_t \qquad \ldots(32)$$

$$\Delta X_t = \alpha + \beta T + \theta X_{t-1} + e_t \qquad \ldots(33)$$

การตั้งสมมุติฐานของการทดสอบก็เป็นเช่นเดียวกับที่กล่าวมาข้างต้นคือถ้าการทดสอบยูนิทรูทของดิกกี-ฟลูเลอร์มีปัญหาสหสัมพันธ์ในตัวเองจะทำให้ค่าเดอร์บิน-วัตสันต่ำดังนั้นจึงมีการเสนอให้แก้ปัญหาโดยการเพิ่มขบวนการถดถอยในตัวเอง (Autoregressive Processes) เข้าไปในสมการที่ (2.16)-(2.18) โดยวิธีการนี้เรียกว่าการทดสอบออกเมนต์เทดดิกกี-ฟลูเลอร์ (Augmented Dickey-Fuller Test) ซึ่งจะได้สมการใหม่คือ

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{p} \emptyset_i \Delta X_{t-i} + e_t \qquad \ldots(34)$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{p} \emptyset_i \Delta X_{t-i} + e_t \qquad \ldots(35)$$

$$\Delta X_t = \alpha + \beta T + \theta X_{t-1} + \sum_{i=1}^{p} \emptyset_i \Delta X_{t-i} + e_t \qquad \ldots(36)$$

โดยที่ X_t, X_{t-1} คือข้อมูลอนุกรมเวลาของตัวแปรอิสระณเวลา t และ t-1

$\alpha, \beta, \theta, \emptyset$ คือค่าพารามิเตอร์

T คือค่าแนวโน้ม

e_t คือความคลาดเคลื่อนเชิงสุ่ม (Random Error)

ซึ่งวิธีการทดสอบก็เหมือนกับการทดสอบของดิกกี-ฟลูเลอร์คือทดสอบตัวพารามิเตอร์ θ โดยมีสมมุติฐานตัวสถิติทดสอบและการตัดสินใจเหมือนกับการทดสอบของดิกกี-ฟลูเลอร์

2.1.4.11 การทดสอบโคอินทิเกรชัน (Cointegration Test)

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่งสามารถนำไปใช้หาสมการถดถอยได้ส่วนอนุกรมเวลาที่มีลักษณะไม่นิ่งเมื่อนำไปใช้หาสมการถดถอยอาจได้สมการถดถอยที่ไม่แท้จริงคือค่าR^2ที่คำนวณได้มีค่าสูงแต่ค่าเดอร์บิน-วัตสันมีค่าต่ำเมื่อทราบว่าข้อมูลอนุกรมเวลามีลักษณะไม่นิ่งแล้วปัญหาสมการถดถอยไม่แท้จริงอาจไม่เกิดขึ้นถ้าหากว่าสมการถดถอยดังกล่าวมีลักษณะโคอินทิเกรชัน

โคอินทิเกรชันคือการมีความสัมพันธ์ระยะยาวระหว่างข้อมูลอนุกรมเวลาตั้งแต่ 2 ตัวแปรขึ้นไปที่มีลักษณะไม่นิ่งแต่ค่าความคลาดเคลื่อนที่ได้จากความสัมพันธ์ระยะยาวมีลักษณะนิ่งสมมุติว่าตัวแปรข้อมูลอนุกรมเวลา 2 ตัวแปรใดๆที่มีลักษณะไม่นิ่งแต่มีค่าสูงขึ้นตามไปด้วยกันทั้งคู่และมีอันดับความสัมพันธ์ของข้อมูลเหมือนกันความแตกต่างระหว่างตัวแปรทั้งสองดังกล่าวมีลักษณะนิ่งกล่าวได้ว่าข้อมูลอนุกรมเวลาดังกล่าวมีลักษณะโคอินทิเกรชัน

ดังนั้นการถดถอยร่วมกันไปด้วยกัน (Cointegration Regression) คือเทคนิคการประมาณค่าความสัมพันธ์ดุลยภาพระยะยาวระหว่างข้อมูลอนุกรมเวลาที่มีลักษณะไม่นิ่งโดยที่ค่าความคลาดเคลื่อนที่ได้จากดุลยภาพระยะยาวต้องมีลักษณะนิ่ง

วิธีการตรวจสอบว่าตัวแปรมีลักษณะโคอินทิเกรชันหรือไม่มีวิธีตรวจสอบโดยวิธีเอนเกิล- แกรนเจอร์ (Engle-Granger (EG) Test) หรือวิธีออกเมนต์เทดเอนเกิล-แกรนเจอร์ (Augmented Engle- Granger (AEG) Test)

วิธีเอนเกิล-แกรนเจอร์หรือวิธีออกเมนต์เทดเอนเกิล-แกรนเจอร์เป็นการใช้ค่าความคลาดเคลื่อนจากสมการถดถอย (Regression Equation) ที่ได้มาทำการทดสอบว่ามีโคอินทิเกรชันหรือไม่โดยการทดสอบยูนิทรูทโดยนำค่าε_tมาหาสมการถดถอยใหม่ดังนี้

$$\Delta e_t = \gamma e_{t-1} + \varepsilon_t \qquad \ldots(37)$$

โดยที่ e_t, e_{t-1}คือค่าความคลาดเคลื่อนณเวลา t และ t-1 ที่นำมาหาสมการถดถอยใหม่

γคือค่าพารามิเตอร์

ε_tคือค่าความคลาดเคลื่อนเชิงสุ่ม

สมมุติฐานคือ

$\mathrm{H}_0 \ : \ \gamma = 0$(ไม่มีโคอินทิเกรชัน)

$H_1 \ : \ \gamma < 0$ (มีโคอินทิเกรชัน)

โดยใช้สถิติทดสอบคือ

$$t = \frac{\hat{\gamma}}{S.E.(\hat{\gamma})}$$

โดยที่ $\hat{\gamma}$ประมาณค่าได้จากวิธีกำลังสองน้อยที่สุด

ซึ่ง $\hat{\gamma} = (e'_{t-1}e_{t-1})^{-1}e'_{t-1}\Delta e_t$และ$S.E.(\hat{\gamma}) = (e'_{t-1}e_{t-1})^{-1}\sigma^2$

ยอมรับสมมุติฐานH_0เมื่อค่าของตัวสถิติ t มีค่ามากกว่าค่าในตารางดิกกี-ฟลูเลอร์และจะปฏิเสธสมมุติฐานH_0เมื่อค่าของตัวสถิติ t มีค่าน้อยกว่าค่าในตารางดิกกี-ฟลูเลอร์

ถ้าการทดสอบมีปัญหาสหสัมพันธ์ในตัวเองจะทำให้ค่าเดอร์บิน-วัตสันมีค่าต่ำดังนั้นจึงมีการเสนอให้แก้ปัญหาโดยการเพิ่มขบวนการถดถอยในตัวเอง (Autoregressive Processes) เข้าไปในสมการที่ (2.22) โดยวิธีการนี้เรียกว่าการทดสอบออกเมนต์เทดดิกกี-ฟลูเลอร์ (Augmented Dickey-Fuller Test) ซึ่งจะได้สมการใหม่คือ

$$\Delta e_t = \gamma e_{t-1} \sum_{i=1}^{p} a_i \, \Delta e_{\mathrm{t-i}} + \varepsilon_{\mathrm{t}} \qquad \ldots(38)$$

โดยที่ e_t, e_{t-1}คือค่าความคลาดเคลื่อนณเวลา t และ t-1 ที่นำมาหาสมการถดถอยใหม่

γ, aคือค่าพารามิเตอร์

ε_{t}คือค่าความคลาดเคลื่อนเชิงสุ่ม

ซึ่งวิธีการทดสอบก็เหมือนกับการทดสอบของดิกกี-ฟลูเลอร์คือทดสอบตัวพารามิเตอร์γโดยมีสมมุติฐานตัวสถิติทดสอบและการตัดสินใจเหมือนกับการทดสอบของดิกกี-ฟลูเลอร์

2.1.4.11 ตัวแบบการปรับแก้ค่าคลาดเคลื่อน (Error Correction Model)

ตัวแบบการปรับแก้ค่าคลาดเคลื่อนมีรูปแบบทั่วไปดังนี้

$$\Delta \gamma_t = a_1 + a_2 \hat{e}_{\mathrm{t-i}} + a_3 \Delta X_t + e_t \qquad \ldots(39)$$

โดยที่ a_1, a_2, a_3ประมาณค่าได้จากวิธีกำลังสองน้อยที่สุด

ตัวแบบการปรับแก้ค่าคลาดเคลื่อนเป็นตัวแบบที่สามารถประมาณค่าสัมประสิทธิ์ด้วยวิธีการกำลังสองน้อยที่สุดได้โดยไม่เกิดสมการถดถอยที่ไม่แท้จริง

2.1.4.12 การทดสอบสมมุติฐานที่เกี่ยวข้องกับค่าความคลาดเคลื่อน

ตรวจสอบคุณสมบัติของความคลาดเคลื่อนเมื่อได้รูปแบบและค่าพารามิเตอร์ที่ประมาณได้จะถูกนำมาตรวจสอบเพื่อดูว่ารูปแบบและค่าพารามิเตอร์ที่ได้นั้นมีความเหมาะสมในการนำไปใช้ในการพยากรณ์หรือไม่

โดยตรวจสอบสมมุติฐานของค่าความคลาดเคลื่อน โดยนำค่าคลาดความเคลื่อนมาทดสอบดูว่ามี

(1) ตรวจสอบสมมุติฐานของค่าความคลาดเคลื่อนว่ามีค่าเฉลี่ยเท่ากับศูนย์หรือไม่

ทดสอบ $H_0 \ : \ \mu_\varepsilon = 0$

$H_1 \ : \ \mu_\varepsilon \neq 0$

โดยใช้ตัวสถิติทดสอบt ดังนี้

$$t = \frac{\bar{e} - \mu_\varepsilon}{\frac{S}{\sqrt{n}}}$$ มีองศาความเป็นอิสระเท่ากับ n-1

เมื่อ $\bar{e}$ คือค่าเฉลี่ยของค่าความคลาดเคลื่อนและมีการแจกแจงแบบปกติ

μ_{ε} คือค่าเฉลี่ยของประชากรของค่าความคลาดเคลื่อน

S คือส่วนเบี่ยงเบนมาตรฐานของค่าความคลาดเคลื่อน

n คือจำนวนค่าความคลาดเคลื่อน

ถ้าค่าของตัวสถิติ t ที่คำนวณได้ตกอยู่ในอาณาเขตวิกฤติแสดงว่าค่าเฉลี่ยของค่าความคลาดเคลื่อนไม่เท่ากับศูนย์

(2) ตรวจสอบสมมุติฐานของค่าความคลาดเคลื่อนว่ามีค่าความแปรปรวนคงที่หรือไม่

H_0: ความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่

H_1: ความแปรปรวนของความคลาดเคลื่อนมีค่าไม่คงที่

โดยใช้ตัวสถิติทดสอบคือ

White's test stat หรือ Obs - R^2 คำนวณจาก n คูณ R^2 โดยค่านี้มีการแจกแจงเป็นไคสแควร์ (Chi-square distribution) มีองศาความเป็นอิสระเท่ากับจำนวนสัมประสิทธิ์ที่อยู่ในสมการแต่ไม่รวมค่าคงที่นั่นคือเมื่อ n คือจำนวนของค่าความคลาดเคลื่อน

R^2 คือ $1 - \dfrac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$ ของสมการ

$$e_t^2 = a_1 + a_2X_1 + a_3X_2 + a_4X_3 + a_5X_1^2 + \cdots + a_7X_3^2 + v_t \quad \ldots(40)$$

กำหนดให้ $Y_t = b_1 + b_2X_1 + b_3X_2 + b_4X_3 + e_t$ เป็นสมการเริ่มต้นที่จะหาความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่

ถ้าค่าของตัวสถิติ White's test stat หรือ Obs - R^2 ที่คำนวณได้มีค่าน้อยกว่าค่าไคสแควร์ที่ระดับนัยสำคัญ α มีองศาความเป็นอิสระเท่ากับจำนวนของสัมประสิทธิ์ที่อยู่ในสมการ $e_t^2 = a_1 + a_2X_1 + a_3X_2 + a_4X_3 + a_5X_1^2 + a_6X_2^2 + a_7X_3^2 + v_t$ ไม่รวมค่าคงที่แสดงว่าค่าความคลาดเคลื่อนมีค่าความแปรปรวนคงที่

(3) ตรวจสอบสมมุติฐานของค่าคลาดเคลื่อนว่ามีความเป็นอิสระกันหรือไม่

ทำการทดสอบโดยใช้การทดสอบของเดอร์บิน-วัตสันและมีสมมุติฐานคือ

ทดสอบ $H_0 \;:\; \rho = 0$

$H_1 \;:\; \rho \neq 0$

โดยใช้ตัวสถิติทดสอบ t ดังนี้

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

การทดสอบนี้มีข้อสมมุติว่าความคลาดเคลื่อนมีรูปแบบถดถอยในตัวเองลำดับที่หนึ่ง

เมื่อ$e_t = Y_t - \hat{Y}_t, t = 1,2,3, \dots, n$ …(41)

เดอร์บิน-วัตสันได้แสดงค่าd มีขอบเขตอยู่ระหว่างd_Lและd_uโดยค่าd_Lและd_uแสดงไว้ในตารางของค่าสถิติเดอร์บิน-วัตสัน

ถ้าd<d_Lหรือ 4 −d <d_LจะปฏิเสธH_0ที่ระดับนัยสำคัญ 2α

ถ้า d >d_uและ4 −d >d_uจะยอมรับH_0ที่ระดับนัยสำคัญ 2α

ถ้านอกจากนี้ไม่สามารถสรุปผลการทดสอบได้

(4) ตรวจสอบสมมุติฐานของค่าคลาดเคลื่อนว่ามีการแจกแจงแบบปกติ

ทดสอบ H_0: ค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ

H_1: ค่าคลาดเคลื่อนไม่มีการแจกแจงแบบปกติ

ใช้ค่าสถิติทดสอบคือ

$$Jarque - Bera = \frac{n-k}{6}\left(S^2 + \frac{(K-3)^2}{4}\right)$$

โดยกำหนดให้n คือจำนวนค่าความคลาดเคลื่อน

S คือสัมประสิทธิ์ความเบ้

K คือสัมประสิทธิ์ความโด่ง

kคือจำนวนตัวเลขที่ประมาณของจำนวนค่าความคลาดเคลื่อน

ถ้าค่าJarque – Beraมีค่าน้อยกว่า$X^2_{0.99,2}$แสดงว่าค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ

2.1.4.13 การพยากรณ์ (Forecasting)

จากตัวแบบการปรับแก้ค่าคลาดเคลื่อนจะได้สมการดังนี้

$$\Delta Y_t = a_1 + a_2\hat{e}_{t-1} + a_3\Delta X_t + e_t \quad \dots(41)$$

จะหาค่าพยากรณ์ของ$\Delta\hat{X}_t$และ $\hat{e}_{t-1}$หลังจากได้ค่าพยากรณ์แล้วจะนำค่าที่ได้ไปแทนลงในตัวแบบการปรับแก้ค่าคลาดเคลื่อนเมื่อ $\Delta\hat{X}_t$พยากรณ์โดยวิธีบอกซ์และเจนกินส์$\hat{e}_{t-1}$พยากรณ์โดยนำค่า$\hat{Y}$ที่พยากรณ์โดยวิธีบอกซ์และเจนกินส์ลบกับYที่คำนวณ

จากสมการถดถอย $Y_t = a_1 + a_2X_t + e_t$ …(42)

จะได้ค่า $\Delta\hat{Y}_t$ ซึ่งค่า $\Delta\hat{Y}_t = \hat{Y}_t - \hat{Y}_{t-1}$

$$\hat{Y}_t = \Delta\hat{Y}_t + \hat{Y}_{t-1} \quad \dots(43)$$

2.1.5 การวัดความถูกต้องแม่นยำของการพยากรณ์

เนื่องจากการพยากรณ์มีหลายวิธีด้วยกันดังนั้นการที่จะตัดสินว่าวิธีการพยากรณ์แบบใดให้คำตอบที่ดีที่สุดจึงจำเป็นจะต้องมีดัชนีชี้วัดการสร้างดัชนีชี้วัดเพื่อนำมาเปรียบเทียบก็มีอยู่หลายวิธีเช่นกันแต่ละวิธีจะมีแนวทางดำเนินงานในลักษณะเดียวกันแต่สูตรต่างกันเท่านั้นในที่นี้จะนำเสนอวิธีต่างๆดังนี้

1. **MSE (Mean Square Error)**

$$MSE = \frac{\sum_{t=1}^{n} e_t^2}{n} = \frac{\sum_{t=1}^{n} \left(Y_t - \hat{Y}_t\right)^2}{n} \qquad \ldots(44)$$

2. **MAPE (Mean Absolute Percentage Error)**

$$MAPE = \frac{\sum_{t=1}^{n} \left| e_t / Y_t \right|}{n} \times 100 \qquad \ldots(45)$$

3. **MAD (Mean Absolute Deviation)**

$$MAD = \frac{\sum_{t=1}^{n} \left| e_t \right|}{n} = \frac{\sum_{t=1}^{n} \left| Y_t - \hat{Y}_t \right|}{n} \qquad \ldots(46)$$

เมื่อค่า MSE (Mean Squared Error) MAD (Mean Absolute Deviation) และ MAPE (Mean Absolute Percentage Error) มีค่าต่ำ แสดงถึงวิธีการพยากรณ์นั้นมีความถูกต้องมาก

2.2 งานวิจัยที่เกี่ยวข้อง

Otto Klemm และ **Holger Lange (1999)** ได้ทำการศึกษาเรื่องแนวโน้มของมลพิษทางอากาศในเทือกเขาฟิชเทลเกเบียร์จในบาวาเรียประเทศเยอรมนีข้อมูลที่ใช้ในการวิเคราะห์วัดจากปริมาณมลพิษทางอากาศรายชั่วโมงได้แก่ปริมาณSO_2, NO_xและO_3ซึ่งเก็บรวมรวมตั้งแต่ปี(1985-1997)จากสองแห่งคือปริมาณมลพิษที่วัดได้จากเทือกเขาและปริมาณมลพิษที่วัดจากบริเวณใกล้กับเมืองก่อนการวิเคราะห์แนวโน้มได้มีการทดสอบความเอนเอียงของข้อมูลโดยใช้autocorrelation functions, Hurst statistics, complexity analysisและ recurrence quantificationและการวิเคราะห์แนวโน้มมีการประยุกต์ใช้นอนพาราเมตริก (Mann-Kendall test) จากการศึกษาพบว่าSO_2มีแนวโน้มลดลงNO_xไม่สามารถบ่งชี้แนวโน้มได้และO_3มีแนวโน้มเพิ่มขึ้น

Shuojun Wang (2007) ได้ศึกษาหาแบบจำลองที่เหมาะสมในการทำนายข้อมูลอนุกรมเวลารายเดือนของปริมาณความเข้มข้นสูงสุดของโอโซน(O_3) 1 ชั่วโมง คาร์บอนมอนอกไซด์(CO)และก๊าซไนโตรเจนไดออกไซด์(NO_2)ในเมืองเบเกอร์ฟิลด์ในหุบเขาซานวาคีนรัฐแคลิฟอร์เนียโดยนำแบบจำลองโดยวิธีautoregressive/moving average (ARMA)และmultiple linear regression (MLR)มาเปรียบเทียบกันจากผลการศึกษาพบว่าแบบจำลอง ARMAแสดงให้เห็นว่าปริมาณมลพิษทางอากาศขึ้นอยู่กับปริมาณมลพิษทางอากาศของเดือนก่อนหน้าเดือนที่จะเกิดมลพิษทางอากาศและแบบจำลอง MLRแสดงให้เห็นว่าความเข้มข้นของปริมาณของสารพิษทางอากาศมีความสัมพันธ์กับความเข้มข้นของปริมาณมลพิษในเดือนปัจจุบันอุณหภูมิคลื่นแม่เหล็กไฟฟ้าจากแบบจำลองทั้งสองสามารถสรุปได้ว่าแบบจำลองแบบ ARMAมีความเหมาะสมในการวิเคราะห์ข้อมูลความเข้มข้นของปริมาณมลพิษจะมีความแม่นยำมากกว่าMLR

G. Saffarini และ S Odat (2008) ได้ทำศึกษาเกี่ยวข้อมูลมลพิษทางอากาศรายปีที่แตกต่างกันของเมือง Al-Hashimeyaใจกลางประเทศจอร์แดนโดยใช้การวิเคราะห์เชิงพรรณนาเพื่อดูการเปลี่ยนแปลงที่แตกต่างกันในระยะยาวของปริมาณความเข้มข้นของมลพิษทางอากาศและการวิเคราะห์อนุกรมเวลาโดยการใช้ Autoregressive, integrated, and moving average (ARIMA) modelsเครื่องมือในการทำนายปริมาณมลพิษทางอากาศในอนาคตโดยใช้ข้อมูลปริมาณมลพิษทางอากาศรายปีตั้งแต่ปี1992–2004 จากการศึกษาพบว่า NO_2,H_2S,NO,TSPมีแนวโน้มลดลงและในขณะที่PM10,NO_X, Pbมีแนวโน้ที่จะเพิ่มขึ้นส่วน SO_2ไม่มีการเปลี่ยนแปลง

A.L. Seetharam และ B.L. UdayaSimha (2009) ได้ทำการศึกษาเกี่ยวกับมลพิษทางอากาศในเขตเมืองบังกาลอร์ประเทศอินเดียการศึกษาครั้งนี้มีวัตถุประสงค์เพื่อหาแนวโน้มและพยากรณ์สารมลพิษที่สำคัญโดยการวิเคราะห์อนุกรมเวลาจากการศึกษาพบว่าSO_2มีแนวโน้มลดลงNO_xมีแนวโน้มเพิ่มขึ้นส่วนSPMมีแนวโน้มแบบผสมตัวแบบพยากรณ์จะใช้การเคลื่อนที่เฉลี่ยราย12เดือนราย24เดือนราย36เดือนพบตัวแบบพยากรณ์ที่เหมาะสมสำหรับSO_2คือMA 12 และMA 24 และ MA36 สำหรับSPM ในแบบจำลองเพื่อเปรียบเทียบกับข้อมูลจากการตรวจวัดจริงจากจุดตรวจวัดโดยใช้ค่า *Mean absolute percentage* error (MAPE) Mean absolute deviation (MAD)และ Mean squared deviation (MSD)โดยที่MSD เป็นตัววัดความคลาดเคลื่อนได้ดีกว่าMAPEและMAD ความสัมพันธ์ระหว่างตัวแบบพยากรณ์และตัวแปร SO_2อยู่ระหว่าง0.4-0.7 ความสัมพันธ์ระหว่างตัวแบบพยากรณ์และตัวแปรNO_Xอยู่ระหว่าง0.45-0.65 และความสัมพันธ์ระหว่างตัวแบบพยากรณ์และตัวแปรSPM อยู่ระหว่าง0.4-0.6

มงคลตรี กิจจานนท์ (2545) พบว่าปริมาณมลพิษการปล่อยมลพิษทางอากาศจากการผลิตไฟฟ้าทั้งหมดในประเทศไทยมีแนวโน้มที่สูงขึ้นตามอัตราการเพิ่มขึ้นของกำลังการผลิตยกเว้นก๊าซ SO_2และSPMซึ่งมีแนวโน้มคงที่และลดลงเรื่อยๆเนื่องจากการผลิตใช้ลิกไนต์เป็นเชื้อเพลิงและในช่วงปี2544-2559 การผลิตไฟฟ้าของภาครัฐจะมีแนวโน้มการปล่อยมลพิษทางอากาศลดลงส่วนภาคเอกชนมีแนวโน้มการปล่อยมลพิษทางอากาศเพิ่มขึ้นส่วนใหญ่มาจากIPPมากกว่าSPP นอกจากนี้ยังพบว่าสัดส่วนปริมาณการปล่อยมลพิษทางอากาศต่อหน่วยพลังงานไฟฟ้าจากการผลิตของภาครัฐมีค่าสูงกว่าภาคเอกชนในกรณีของก๊าซ$CO_2$$NO_x$ SO_2และSPMส่วนก๊าซCO CH_4 และ N_2O ภาคเอกชนมีค่าสูงกว่า

มงคล รานะนาคร และคณะ (2550)ได้ศึกษาวิเคราะห์เพื่อหามลพิษทางอากาศในอนุภาคฝุ่นในจังหวัดเชียงใหม่และจังหวัดลำพูนจากการศึกษาพบว่าปัจจัยที่มีผลต่อมลภาวะทางอากาศอย่างเด่นชัดคือฤดูกาลโดยองค์ประกอบของทางเคมีของฝุ่นละอองขนาดเล็กไม่เกิน10ไมครอนได้แก่สาร PAHs ไอออนโลหะธาตุต่างๆและคาร์บอนมีการผันแปรตามฤดูกาลปริมาณสารพิษโดยเฉลี่ยที่พบจากมากไปน้อยคือฤดูแล้งช่วงเปลี่ยนฤดูและฤดูฝน