

กระบวนการสโตแคสติกสำหรับแบบจำลองดัชนีคุณภาพอากาศ

STOCHASTIC PROCESSES FOR AIR QUALITY INDEX MODEL



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาดมหลังสุตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาคณิตศาสตร์ประยุกต์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2557

KMITL-2014-SC-M-001-020

กระบวนการสโตแคสติกสำหรับแบบจำลองดัชนีคุณภาพอากาศ

STOCHASTIC PROCESSES FOR AIR QUALITY INDEX MODEL



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาคณิตศาสตร์ประยุกต์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2557

KMITL – 2014 – SC – M – 001 – 020

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

STOCHASTIC PROCESSES FOR AIR QUALITY INDEX MODEL



THIS THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN APPLIED MATHEMATICS
FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2014

KMITL – 2014 – SC – M – 001 – 020

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2013

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	กระบวนการสโตแคสติกสำหรับแบบจำลองดัชนีคุณภาพอากาศของปัญหาหมอกควันในภาคเหนือ
นักศึกษา	นางสาวขวัญศิริ ศิริมงคล
รหัสประจำตัว	55650707
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	คณิตศาสตร์ประยุกต์
พ.ศ.	2556
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ดร.บุญขมาส พิมพ์พรรณชาติ

บทคัดย่อ

หมอกควันในภาคเหนือตอนบนของประเทศไทยถือเป็นปัญหามลพิษที่เกิดขึ้นต่อเนื่องมาหลายทศวรรษ ความรุนแรงปรากฏชัดในช่วงเดือนกุมภาพันธ์ถึงเมษายน งานวิจัยนี้ได้ศึกษาข้อมูลสารมลพิษและปัจจัยทางอุตุนิยมวิทยาระหว่างปีค.ศ. 2005 - 2010 เพื่อสร้างแบบจำลองที่ใช้พยากรณ์ปริมาณความเข้มข้นของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน โดยใช้อนุกรมเวลา ซึ่งทำให้มองภาพรวมของปัจจัยต่างๆ และแนวโน้มในอนาคต หลังจากนั้นได้ทำการสร้างแบบจำลองการวิเคราะห์การถดถอยโลจิสติกเพื่อพยากรณ์ระดับการเกิดหมอกควัน โดยแบ่งเป็น 3 ระดับ คือ การเกิดหมอกควันที่ไม่มีผลต่อสุขภาพ มีผลต่อสุขภาพ และมีผลต่อสุขภาพมาก นอกจากนี้ยังมีประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมสำหรับการวิเคราะห์การถดถอยพหุคูณ เพื่อพยากรณ์ปริมาณฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) ผลการวิเคราะห์อนุกรมเวลาพบว่าปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน ได้แก่ PM_{10} ก๊าซคาร์บอนไดออกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ และปริมาณน้ำฝน ซึ่งทุกปัจจัยเหล่านี้มีการเปลี่ยนแปลงตามฤดูกาลเชิงบวก เหมาะสมต่อการพยากรณ์ด้วยวิธีของโฮลต์และวินเซอร์ แบบจำลองระดับการเกิดหมอกควันให้ค่า $Pseudo R^2 = 0.596$ สามารถพยากรณ์ระดับการเกิดหมอกควันในปี 2011 ได้ถูกต้อง 93.7% คุณภาพอากาศอยู่ในเกณฑ์ดี 60% ปานกลาง 31.23% และมีผลต่อสุขภาพ 8.77% แบบจำลองการวิเคราะห์การถดถอยพหุคูณสำหรับพยากรณ์ปริมาณ PM_{10} โดยใช้ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันให้ค่า $R^2 = 0.8615$ และ $RMSE = 11.18$

คำสำคัญ: หมอกควัน ดัชนีคุณภาพอากาศ อนุกรมเวลา วิธีของโฮลต์และวินเซอร์ การวิเคราะห์การถดถอยโลจิสติก การวิเคราะห์การถดถอยพหุคูณ ขั้นตอนวิธีเชิงพันธุกรรม

Thesis Title	Stochastic Processes for Air Quality Index Model of Haze Problems in Northern Region of Thailand
Student	Khwansiri Sirimangkhal
Student ID	55650707
Degree	Master of Science
Program	Applied Mathematics
Year	2556
Thesis Advisor	Dr. Busayamas Pimpunchat

ABSTRACT

Haze and air pollutions are major environmental problems in the upper Northern region in Thailand. The severity of this problem has clearly appeared in dry season every year particular February to April. The pollutants and meteorological data in 2005 – 2010 were analyzed in this paper. The study aims to forecast associated factors of haze problem by using time series. The logistic regression was considered for levels of daily haze problem: not much health effect, more health effect and alert and serious health effect. In addition genetic algorithm is applied for multiple regressions for forecasting PM_{10} concentrations. The associated factors of monthly haze problem are PM_{10} , CO_2 , NO_2 , Relative Humidity and Rain. Each factor has an additive seasonal effect, the Holt – Winter’s method is a good method for this model. The model gives Pseudo $R^2 = 0.596$. The results of levels of daily haze problem have 93.7% accuracy for forecasting value in 2011. Most air quality is good. Multiple regression model for forecasting PM_{10} concentrations give $R^2 = 0.8615$ and $RMSE = 11.18$.

Keywords: Haze Problem, PM_{10} , Air Quality Index, AQI, Time Series, Holt – Winter’s method, Logistic Regression, Multiple Regressions, Genetic algorithm, GA

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จได้ด้วยความกรุณาจากอาจารย์ที่ปรึกษา ดร.บุษยามาส พิมพ์พรรณชาติ ที่ให้ความช่วยเหลือ ให้คำชี้แนะในการแก้ปัญหา ตลอดจนให้ความรู้และประสบการณ์ที่ดีแก่ข้าพเจ้า

ขอขอบพระคุณ ดร.สิริพร แสนนำ วินเทอร์ ดร.วรรณพร สรรประเสริฐ และ ผศ.ดร.สุชุมาส สาริกะวณิช กรรมการสอบหัวข้อและเค้าโครงร่างวิทยานิพนธ์ที่ได้กรุณาให้คำแนะนำ ตลอดจนข้อเสนอแนะต่างๆ จนทำให้วิทยานิพนธ์นี้สำเร็จลงได้ในที่สุด

สุดท้ายต้องขอขอบคุณสำนักจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อมที่ให้การสนับสนุนข้อมูลสำหรับการวิจัยครั้งนี้

สำหรับคุณงามความดีอันใดที่เกิดจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบให้กับบิดามารดา ซึ่งเป็นที่รักและเคารพยิ่ง ตลอดจนครูอาจารย์ทุกท่านที่ได้ถ่ายทอดความรู้และประสบการณ์ที่ดีแก่ข้าพเจ้า



สารบัญ

บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญรูป	VI
สารบัญตาราง	VIII
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	5
1.3 สมมติฐานของการศึกษา	5
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย	5
1.5 ขอบเขตการวิจัย	6
1.6 ขั้นตอนของการศึกษา	6
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	8
2.1 สถานการณ์หมอกควันในประเทศไทย	8
2.2 ดัชนีคุณภาพอากาศ	11
2.3 การวิเคราะห์สหสัมพันธ์	14
2.4 การวิเคราะห์อนุกรมเวลา	15
2.5 การวิเคราะห์การถดถอยโลจิสติก	19
2.6 การประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมสำหรับการวิเคราะห์การถดถอย	28
2.6 งานวิจัยที่เกี่ยวข้อง	36
บทที่ 3 วิธีการดำเนินงาน	39
3.1 ข้อมูลและการเก็บรวบรวมข้อมูล	39
3.2 ตัวแบบการพยากรณ์อนุกรมเวลา	39
3.3 ตัวแบบการวิเคราะห์การถดถอยโลจิสติก	42
3.4 ตัวแบบการวิเคราะห์การถดถอยพหุคูณ โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม	44

สารบัญ (ต่อ)

บทที่ 4 ผลการดำเนินการ	46
4.1 ตัวแบบพยากรณ์อนุกรมเวลาของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน	46
4.2 ตัวแบบพยากรณ์การเกิดหมอกควันโดยการวิเคราะห์การถดถอยโลจิสติก	50
4.3 ตัวแบบพยากรณ์ PM_{10} โดยการวิเคราะห์การถดถอยพหุคูณ	52
บทที่ 5 สรุปผลการดำเนินการ	54
เอกสารอ้างอิง	56



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

	หน้า
รูปที่ 1.1 สภาพอากาศของกรุงเทพมหานครที่เกิดจากผลกระทบของ โอโซน	3
รูปที่ 1.2 ผุ่นละอองขนาดต่างๆ	4
รูปที่ 1.3 การแพร่กระจายของผุ่นละอองขนาดต่างๆในระบบทางเดินหายใจ	4
รูปที่ 2.1 สถานการณ์หมอกควันในพื้นที่จังหวัดเชียงใหม่	9
รูปที่ 2.2 สถานีตรวจวัดคุณภาพอากาศในบรรยากาศ	10
รูปที่ 2.3 แนวโน้มยอดขายของสินค้าชนิดหนึ่ง	16
รูปที่ 2.4 ข้อมูลอนุกรมเวลาที่มีการแปรผันตามฤดูกาลเชิงบวก	17
รูปที่ 2.5 ข้อมูลอนุกรมเวลาที่มีการแปรผันตามฤดูกาลเชิงคูณ	17
รูปที่ 2.6 ข้อมูลอนุกรมเวลาที่มีการแปรผันตามวัฏจักร	17
รูปที่ 2.7 ฟังก์ชันตอบสนองโลจิสติก	21
รูปที่ 2.8 วิวัฒนาการของปากนกตามลักษณะของอาหารตามทฤษฎีของชาลส์ ดาร์วิน	32
รูปที่ 2.9 การถ่ายทอดลักษณะทางพันธุกรรมตามทฤษฎีของเกรเกอร์ เมนเดล	33
รูปที่ 2.10 หน้าจอขั้นตอนวิธีเชิงพันธุกรรมโดยโปรแกรมสำเร็จรูป	35
รูปที่ 2.11 แผนผังการทำงานของขั้นตอนวิธีเชิงพันธุกรรม	36
รูปที่ 3.1 ปริมาณ PM_{10} รายเดือนตั้งแต่ปี 2005 – 2010	40
รูปที่ 3.2 ปริมาณ CO รายเดือนตั้งแต่ปี 2005 – 2010	40
รูปที่ 3.3 ปริมาณ NO_2 รายเดือนตั้งแต่ปี 2005 – 2010	41
รูปที่ 3.4 ปริมาณความชื้นสัมพัทธ์รายเดือนตั้งแต่ปี 2005 – 2010	41
รูปที่ 3.5 ปริมาณน้ำฝนรายเดือนตั้งแต่ปี 2005 – 2010	42
รูปที่ 4.1 ปริมาณ PM_{10} ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน ปี ค.ศ.2005 – 2010	47
รูปที่ 4.2 ปริมาณ CO ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน ปี ค.ศ.2005 – 2010	47
รูปที่ 4.3 ปริมาณ NO_2 ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือนปี ค.ศ.2005 – 2010	47
รูปที่ 4.4 ปริมาณความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน	48
รูปที่ 4.5 ปริมาณน้ำฝนที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน ปี ค.ศ.2005 – 2010	48
รูปที่ 4.6 ปริมาณ PM_{10} ที่ได้จากการตรวจวัด ปี ค.ศ.2005 – 2011	49
รูปที่ 4.7 ผลการวิเคราะห์ค่าพารามิเตอร์ของแบบจำลองการวิเคราะห์การถดถอยโลจิสติก	50
รูปที่ 4.8 ผลการวิเคราะห์การเกิดหมอกควันโดยการวิเคราะห์การถดถอยโลจิสติกในปี 2011	51

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 4.9 ความเข้มข้นรายวันของ PM_{10} ที่ได้จากการตรวจวัดในปี ค.ศ. 2011	52
รูปที่ 4.10 ระดับเกณฑ์ดัชนีคุณภาพอากาศของ PM_{10} ในปี ค.ศ. 2011	52
รูปที่ 4.11 ความเข้มข้นรายวันของ PM_{10} ที่ได้จากการตรวจวัดและค่าพยากรณ์ในปี ค.ศ. 2011	53



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

	หน้า
ตารางที่ 2.1 เกณฑ์ดัชนีคุณภาพอากาศสำหรับประเทศไทย	12
ตารางที่ 2.2 ความเข้มข้นของสารมลพิษทางอากาศที่เทียบเท่ากับดัชนีคุณภาพอากาศ	12
ตารางที่ 2.3 ข้อมูลอนุกรมเวลา	15
ตารางที่ 3.1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยต่างๆและ PM_{10} (รายเดือน)	39
ตารางที่ 3.2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยต่างๆและ PM_{10} (รายวัน)	43
ตารางที่ 3.3 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระต่างๆ	43
ตารางที่ 4.1 ค่า α_1 และ α_3 ที่ทำให้ RMSE มีค่าน้อยที่สุด	46
ตารางที่ 4.2 ค่าจริงและค่าพยากรณ์ของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในปีค.ศ.2011	49



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

สถานการณ์หมอกควันในภาคเหนือตอนบนของประเทศไทยถือเป็นปัญหาสำคัญ ความรุนแรงของปัญหานี้โดยทั่วไปปรากฏชัดเจนในช่วงหน้าแล้ง ซึ่งมักพบการเพิ่มขึ้นของปริมาณฝุ่นละอองขนาดเล็ก เนื่องจากความแห้งแล้งส่งผลให้เกิดการเพิ่มขึ้นของไฟป่า และในช่วงเวลาดังกล่าวเกษตรกรจะทำการเผาเศษวัสดุเพื่อเตรียมพื้นที่สำหรับการเกษตรในช่วงฤดูฝน ประกอบกับสภาพภูมิประเทศซึ่งมีภูเขาสูงล้อมรอบ ทำให้มลพิษต่างๆถูกกักไว้และแผ่ปกคลุมทั่วบริเวณ ซึ่งสภาวะอากาศที่แห้งและนิ่งทำให้ฝุ่นละอองที่เกิดขึ้นสามารถแขวนลอยอยู่ในบรรยากาศได้นาน ประกอบกับฝนตกน้อยทำให้การชะล้างหมอกควันหรือฝุ่นที่แขวนลอยอยู่ในอากาศเป็นไปได้น้อย ทั้งนี้ ค่ามาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไปในประเทศไทยที่ทางกรมควบคุมมลพิษได้กำหนดไว้คือ ค่าเฉลี่ยปริมาณฝุ่นละอองขนาดเล็กกว่า 10 ไมครอนในเวลา 24 ชั่วโมงจะต้องไม่เกิน 120 ไมโครกรัมต่อลูกบาศก์เมตร และค่าเฉลี่ยในเวลา 1 ปีจะต้องไม่เกิน 50 ไมโครกรัมต่อลูกบาศก์เมตร ซึ่งปัญหานี้ส่งผลกระทบต่อสภาพบรรยากาศทั่วไป ต่อสุขภาพอนามัย และต่อธุรกิจการท่องเที่ยวและบริการ จากปัญหาข้างต้นทำให้มีความจำเป็นอย่างยิ่งที่จะต้องทราบแนวโน้มของปัจจัยที่มีผลต่อการเกิดหมอกควัน โดยอาศัยการวิเคราะห์หอนุกรมเวลามาใช้ในการพยากรณ์ปัจจัยต่างๆที่มีผลต่อการเกิดหมอกควัน

ดัชนีคุณภาพอากาศถือเป็นการรายงานคุณภาพอากาศสู่สาธารณะชนเพื่อให้ง่ายต่อความเข้าใจของประชาชนทั่วไป ว่าคุณภาพอากาศอยู่ในระดับใด มีผลกระทบต่อสุขภาพหรือไม่ ดัชนีชี้วัดคุณภาพอากาศสำหรับประเทศไทยคำนวณจากความเข้มข้นของสารมลพิษ 5 ปัจจัย ได้แก่ ก๊าซโอโซน ก๊าซไนโตรเจนไดออกไซด์ ก๊าซซัลเฟอร์ไดออกไซด์ ก๊าซคาร์บอนมอนนอกไซด์ และฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน ทั้งนี้ ดัชนีคุณภาพอากาศที่คำนวณได้ของปัจจัยใดมีค่าสูงที่สุดจะใช้ค่านั้นเป็นค่าบ่งชี้คุณภาพอากาศของวันนั้น สารมลพิษทั้ง 5 มีบทบาทต่อการดำเนินชีวิตของมนุษย์ และสามารถส่งผลต่อสุขภาพมนุษย์ได้ดังต่อไปนี้

1) ก๊าซคาร์บอนมอนนอกไซด์ (CO)

ก๊าซคาร์บอนมอนนอกไซด์เกิดจากการเผาไหม้ที่ไม่สมบูรณ์ของเชื้อเพลิงที่มีคาร์บอน แออสเป็นองค์ประกอบ เป็นก๊าซไม่มีสี กลิ่น และรส ทำให้ความเป็นอันตรายสูง เนื่องจากไม่สามารถรับรู้ได้ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้ว่ามีก๊าซนี้อยู่ในบรรยากาศ โดยปกติในบรรยากาศทั่วไปมีปริมาณก๊าซคาร์บอนมอนอกไซด์ประมาณ 0.1 พีพีเอ็ม (ส่วนในล้านส่วน) ในท้องถนนที่มีการจราจรหนาแน่นอาจมีค่าสูงถึง 100 – 200 พีพีเอ็ม เมื่อหายใจเอาก๊าซคาร์บอนมอนอกไซด์เข้าไป ก๊าซนี้จะลดความสามารถของเลือดในการเป็นตัวนำออกซิเจนจากปอดไปยังเนื้อเยื่อต่างๆ ซึ่งจะทำให้ร่างกายได้รับก๊าซออกซิเจนไม่เพียงพอ ทำให้หัวใจทำงานสูบน้ำเลือดมากขึ้น มีอาการมึนงง ตาพร่ามัว ปวดศีรษะ คลื่นไส้ อ่อนเพลีย โดยหากระดับความเข้มข้นสูงกว่า 1,000 ppm จะทำให้เป็นลม หมดสติ และเสียชีวิตได้

2) ก๊าซไนโตรเจนไดออกไซด์ (NO₂)

ก๊าซไนโตรเจนไดออกไซด์เกิดจากการทำปฏิกิริยาทางเคมีของไนโตรเจนกับออกซิเจนในระหว่างการเผาไหม้เชื้อเพลิงประเภทต่างๆ ที่อุณหภูมิสูงๆ เมื่อหายใจเข้าไปแล้วอาจทำให้เกิดความระคายเคืองในถุงลม ทำให้เกิดการคล้ายกับโรคหลอดลมตีบตัน โดยเฉพาะในบุคคลที่เป็นโรคหอบหืดอยู่แล้ว นอกจากนี้หากก๊าซนี้เข้าไปในปอดจะเปลี่ยนไปเป็นสารไนโตรซามีน (Nitrosamines) ซึ่งทำให้เกิดมะเร็งในปอดได้ มนุษย์เริ่มได้กลิ่นก๊าซนี้ที่ระดับ 230 ไมโครกรัมต่อลูกบาศก์เมตร หากมีความเข้มข้นเพิ่มขึ้นจะทำให้เกิดกลิ่นเร็วขึ้น ผู้ป่วยที่เป็นโรคหืดหอบอาจมีอาการเร็วขึ้นหากได้รับก๊าซนี้ที่ระดับ 190 ไมโครกรัมต่อลูกบาศก์เมตร คนทั่วไปเริ่มต้นมีอาการเมื่อได้รับก๊าซนี้ที่ระดับความเข้มข้นประมาณ 1,300 – 1,800 ไมโครกรัมต่อลูกบาศก์เมตร

3) ก๊าซซัลเฟอร์ไดออกไซด์ (SO₂)

ก๊าซซัลเฟอร์ไดออกไซด์เป็นก๊าซไม่มีสี หรืออาจมีสีเหลืองอ่อนๆ มีรส และกลิ่นที่ระดับความเข้มข้นที่สูงพอ ก๊าซนี้เกิดจากการเผาไหม้เชื้อเพลิงที่มีกำมะถัน (ซัลเฟอร์) เป็นส่วนประกอบเป็นก๊าซที่มีผลกระทบโดยตรงต่อสุขภาพ ทำให้เกิดความระคายเคืองต่อระบบทางเดินหายใจ ผิวหนัง และเยื่อเมือก ทำให้เกิดการแสบตา เมื่อเราหายใจเอาก๊าซนี้เข้าไป จะทำให้ก๊าซละลายเป็นของเหลวในระบบทางเดินหายใจ ทำให้เป็นโรคระบบทางเดินหายใจได้ง่าย และเกิดการกำเริบของโรคปอดขึ้น เช่น โรคหลอดลมอักเสบเรื้อรัง และโรคปอดอื่นๆ ความอันตรายจะรุนแรงมากขึ้นเมื่อรวมกับฝุ่นละออง ซึ่งจะทำให้เกิดความระคายเคืองต่อเนื้อเยื่อในระบบทางเดินหายใจมากขึ้น

4) ก๊าซโอโซน (O₃)

ก๊าซโอโซนเป็นก๊าซที่ไม่มีสีหรือมีสีฟ้าอ่อน มีกลิ่นฉุน ในธรรมชาตินั้นก๊าซโอโซนในชั้นบรรยากาศที่มนุษย์อยู่หรือโทรโพสเฟียร์ (Tropospheric ozone หรือ ground level zone) มีปริมาณที่เหมาะสมอยู่แล้ว ทำให้อากาศสะอาดเนื่องจากสามารถกำจัดเชื้อโรคได้ แต่เมื่อมีปริมาณสูงขึ้น จะส่งผลกระทบต่อสุขภาพอนามัยของมนุษย์ จะก่อให้เกิดการระคายเคืองตาและระคายเคืองต่อระบบทางเดินหายใจ ลดความสามารถในการทำงานของปอดลง ทำให้เหนื่อยเร็ว โดยเฉพาะในเด็ก คนชรา และคนที่ป่วยเป็นโรคปอดเรื้อรังอยู่แล้ว ก๊าซโอโซนก่อให้เกิดปรากฏการณ์หมอกควันเคมีแสง หรือโฟโตเคมีคัลสม็อก (Photochemical smog) ในชั้นบรรยากาศ ซึ่งมีลักษณะเป็นหมอกเอกลากรนี้เป็นเอกลากรที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ควันทึขาวปกคลุมไปทั่วบรรยากาศ สามารถมองเห็นได้จากตึกสูง ดังรูปที่ 1.1 ซึ่งแสดงถึงสภาพอากาศของกรุงเทพมหานครในวันที่มีก๊าซโอโซนสูง และในวันที่มีก๊าซโอโซนต่ำ ปรากฏการณ์นี้มีผลให้เกิดการสะสมของสารมลพิษทางอากาศในบริเวณนั้นมากขึ้น เนื่องจากมีการแพร่กระจายได้น้อย (ศิวพันธุ์ ชูอินทร์, 2556)



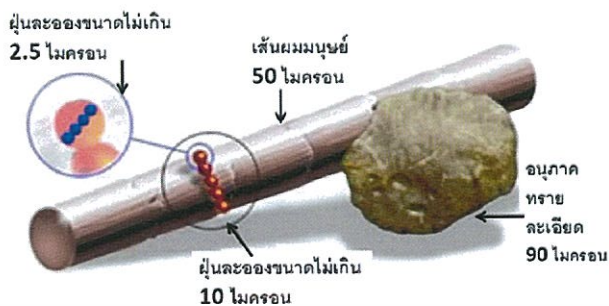
รูปที่ 1.1 สภาพอากาศของกรุงเทพมหานครที่เกิดจากผลกระทบของ โอโซน

(ที่มา http://www.prcmu.cmu.ac.th/perin_detail.php?perin_id=303)

5) ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10})

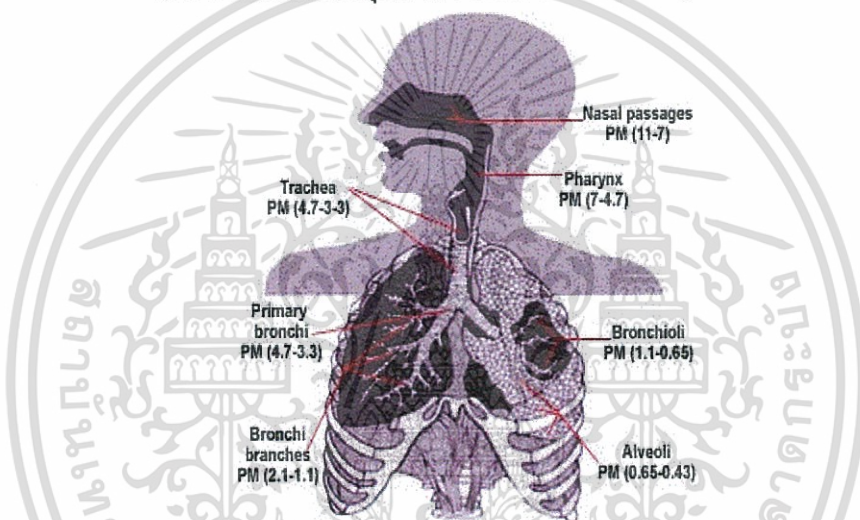
ฝุ่นละออง คือ กลุ่มของมลสารอนุภาคเล็กๆ อาจเป็นของแข็งหรือของเหลวที่กระจัดกระจายอยู่ในบรรยากาศ ฝุ่นละอองที่มีอยู่ในบรรยากาศรอบๆตัวเรา มีทั้งขนาดใหญ่กว่า 500 ไมครอน ซึ่งเป็นฝุ่นทรายขนาดใหญ่ มองเห็นได้ด้วยตาเปล่า ไปจนถึงขนาดเล็กมากตั้งแต่ 0.002 ไมครอน ซึ่งมองด้วยตาเปล่าไม่เห็น ฝุ่นละอองที่สามารถเข้าสู่ระบบทางเดินหายใจของมนุษย์ได้มีขนาดเล็กกว่า 10 ไมครอน ฝุ่นละอองเหล่านี้เมื่อเข้าสู่ระบบทางเดินหายใจจะก่อให้เกิดการระคายเคืองและทำลายเนื้อเยื่อของอวัยวะต่างๆ เช่น เนื้อเยื่อปอด ซึ่งหากได้รับในปริมาณมากหรือในระยะเวลาอันนานจะสามารถสะสมในเนื้อเยื่อปอด เกิดเป็นพังผืดหรือแผลขึ้นได้ และทำให้การทำงานของปอดเสื่อมประสิทธิภาพลง ทำให้หลอดลมอักเสบ เกิดหอบหืด ถุงลมโป่งพอง และโอกาสเกิดโรคระบบทางเดินหายใจเนื่องจากการติดเชื้อเพิ่มขึ้นได้ (นิพนธ์ ตั้งคณาภิรักษ์, 2552) ภาพฝุ่นละอองขนาดต่างๆแสดงในรูปที่ 1.2 ซึ่งแสดงถึงอนุภาคฝุ่นทรายละเอียดซึ่งมีขนาด 90 ไมครอน เส้นผมของมนุษย์ซึ่งมีขนาด 50 ไมครอน และฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน เปรียบเทียบกับ ฝุ่นละอองขนาด 10 ไมครอน รูปที่ 1.3 แสดงถึงการแพร่กระจายของฝุ่นละอองขนาดต่างๆในระบบทางเดินหายใจ จากการวิจัยพบว่าฝุ่นละอองที่ก่อให้เกิดปัญหาต่อสุขภาพจะเป็นฝุ่นละอองขนาดเล็กที่มีขนาดไม่เกิน 10 ไมครอน โดยฝุ่นละอองขนาดเล็กนี้จะสามารถเข้าไปถึงถุงลมในปอด และฝุ่นละอองจะเป็นพิษมากขึ้นหากฝุ่นละอองนั้นเกิดจากการรวมตัวของก๊าซบางชนิด เช่น ซัลเฟอร์ไดออกไซด์ ออกไซด์ของไนโตรเจน รายงานจากการศึกษาวิจัยพบว่าพื้นที่ซึ่งมีปริมาณฝุ่นละออง

ขนาดเล็กในปริมาณมากจะมีจำนวนผู้ป่วยโรคระบบทางเดินหายใจมากขึ้นด้วย (สำนักจัดการคุณภาพอากาศและเสียง, 2551)



รูปที่ 1.2 ฝุ่นละอองขนาดต่างๆ

(ที่มา: สำนักจัดการคุณภาพอากาศและเสียง, 2554)



รูปที่ 1.3 การแพร่กระจายของฝุ่นละอองขนาดต่างๆในระบบทางเดินหายใจ

(ที่มา: สำนักจัดการคุณภาพอากาศและเสียง, 2554)

เกณฑ์ดัชนีคุณภาพอากาศของประเทศไทยสามารถแบ่งเป็น 5 ระดับ คือ 0-50 อยู่ในเกณฑ์คุณภาพอากาศดี 51-100 อยู่ในเกณฑ์คุณภาพอากาศปานกลาง ไม่มีผลต่อสุขภาพ 101-200 อยู่ในเกณฑ์ที่มีผลต่อสุขภาพ กลุ่มเสี่ยง ได้แก่ ผู้ป่วยโรคทางเดินหายใจ เด็ก และผู้สูงอายุไม่ควรทำกิจกรรมภายนอกอาคารเป็นเวลานาน 201-300 อยู่ในเกณฑ์ที่มีผลต่อสุขภาพมาก กลุ่มเสี่ยงควรจำกัดการทำกิจกรรมภายนอกอาคาร และค่าดัชนีคุณภาพอากาศที่มากกว่า 300 อยู่ในเกณฑ์ที่เป็นอันตรายต่อสุขภาพ บุคคลทั่วไปควรหลีกเลี่ยงกิจกรรมภายนอกอาคาร สำหรับผู้ป่วยโรคระบบทางเดินหายใจควรอยู่ในอาคารเท่านั้น

สารมลพิษเหล่านี้เป็นอันตรายต่อสุขภาพได้ โดยเฉพาะเด็ก คนชรา และผู้ป่วยโรคระบบทางเดินหายใจ จึงมีการตรวจวัดและรายงานข้อมูลดังกล่าวผ่านดัชนีคุณภาพอากาศ เพื่อชี้วัดถึง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถานการณ์มลพิษในแต่ละพื้นที่ สำหรับในพื้นที่ที่เกิดปัญหาหมอกควันนั้น ผลกระทบต่อสุขภาพของฝุ่นละอองขนาดเล็กจะเกิดมากที่สุดซึ่งเราจะพบจำนวนผู้ป่วยด้วยโรคระบบทางเดินหายใจที่สูงกว่าเมื่อเปรียบเทียบกับพื้นที่ที่ไม่เกิดปัญหาหมอกควัน นอกจากนี้เราจะพบว่าในช่วงที่เกิดสถานการณ์หมอกควัน PM_{10} เป็นสารมลพิษที่ให้ค่าดัชนีคุณภาพอากาศสูงสุด

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

การศึกษานี้มีวัตถุประสงค์เพื่อนำความรู้ทางคณิตศาสตร์มาประยุกต์ใช้กับปัญหาการเกิดหมอกควัน โดยมีรายละเอียดดังนี้

1. ค้นหารูปแบบอนุกรมเวลาที่เหมาะสมกับ PM_{10} และปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน เพื่อใช้ตัวแบบในการพยากรณ์ปริมาณความเข้มข้นรายเดือนของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน
2. สร้างแบบจำลองการวิเคราะห์การถดถอยพหุคูณ โดยใช้ตัวแปรต้นคือปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน เพื่อพยากรณ์ปริมาณความเข้มข้นรายเดือนของ PM_{10}
3. สร้างแบบจำลองดัชนีคุณภาพอากาศสำหรับปัญหาหมอกควัน โดยใช้การวิเคราะห์การถดถอยโลจิสติกในการพยากรณ์การเกิดปัญหาหมอกควันรายวัน และวิเคราะห์เกณฑ์ดัชนีคุณภาพอากาศของฝุ่นละอองขนาดเล็ก เพื่อศึกษาว่าสภาพอากาศมีผลกระทบต่อสุขภาพหรือไม่

1.3 สมมติฐานของการศึกษา

สมมติฐานในการศึกษานี้ คือคาดว่าปัจจัยที่นำมาศึกษามีความเกี่ยวข้องกับปัญหาหมอกควัน สามารถค้นหาตัวแบบโดยการวิเคราะห์อนุกรมเวลาที่เหมาะสมในการอธิบายปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน สามารถนำการถดถอยโลจิสติกมาใช้ในการพยากรณ์ระดับการเกิดหมอกควัน และสามารถค้นหาตัวแบบการวิเคราะห์การถดถอยพหุคูณเพื่อพยากรณ์ปริมาณของ PM_{10} ในอนาคตได้อย่างมีประสิทธิภาพ

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

ในการศึกษานี้สนใจการนำกระบวนการสโตแคสติกมาอธิบายแบบจำลองคุณภาพอากาศ ผ่านการวิเคราะห์อนุกรมเวลา ซึ่งเป็นการวิเคราะห์ที่ค่าของ PM_{10} เป็นตัวแปรตาม และเวลาเป็นตัวแปรอิสระ กล่าวคือสนใจเพียงค่าของ PM_{10} ที่เปลี่ยนแปลงตามช่วงเวลาหนึ่ง การวิเคราะห์อนุกรมเวลามีรูปแบบที่เป็นไปได้หลากหลายรูปแบบ ดังนั้นจึงต้องมีการค้นหาแบบที่เหมาะสมเพื่อใช้อธิบายตัวแบบได้ดีที่สุดโดยพิจารณาจากรากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย นอกจากนั้น

เอกสารได้นำการวิเคราะห์การถดถอยโลจิสติกซึ่งเป็นการวิเคราะห์เชิงคุณภาพมาวิเคราะห์การเกิดหมอกควัน

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ควัน เพื่อใช้เป็นตัวแบบในการพยากรณ์การเกิดสถานการณ์หมอกควันในอนาคต นอกจากนี้ยังทำการพยากรณ์ปริมาณของ PM_{10} ซึ่งเป็นปัจจัยสำคัญที่บ่งบอกความรุนแรงของสถานการณ์หมอกควัน

1.5 ขอบเขตการวิจัย

งานวิจัยนี้สนใจศึกษาดัชนีคุณภาพอากาศโดยอาศัยข้อมูลสารมลพิษทางอากาศและปัจจัยทางอุตุนิยมวิทยา ของพื้นที่จังหวัดเชียงใหม่ ที่ได้รับการสนับสนุนข้อมูลจากสำนักจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ตั้งแต่เดือนมกราคม ปี ค.ศ. 2005 ถึง เดือนธันวาคม ปี ค.ศ. 2010 ในการวิเคราะห์อนุกรมเวลาใช้ข้อมูลค่าเฉลี่ยรายเดือนในการพิจารณา เพื่อหามองเห็นภาพรวมของข้อมูล ส่วนการวิเคราะห์การถดถอยโลจิสติกใช้ข้อมูลค่าเฉลี่ยรายวันในการพิจารณา เนื่องจากหมอกควันมักเกิดในช่วงระหว่างเดือนกุมภาพันธ์ถึงเมษายนเท่านั้น เหมาะที่จะพิจารณาเป็นรายวันเพื่อจะให้เห็นความชัดเจนของตัวแบบ

1.6 ขั้นตอนของการศึกษา

วิธีการดำเนินการ โดยสรุปดังนี้

- 1) รวบรวมข้อมูลรายชั่วโมงของความเข้มข้นของสารมลพิษทางอากาศ และปัจจัยทางอุตุนิยมวิทยา
- 2) พิจารณาข้อมูลค่าเฉลี่ยรายเดือนของทุกปัจจัยเพื่อค้นหาปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน
- 3) ค้นหารูปแบบอนุกรมเวลาที่เหมาะสมในการอธิบายการเปลี่ยนแปลงของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน โดยใช้ข้อมูลค่าเฉลี่ยรายเดือนของปี ค.ศ. 2005 – 2010
- 4) พยากรณ์ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในปี ค.ศ. 2011 และตรวจสอบความแม่นยำของตัวแบบ โดยพิจารณาจากรากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าสัมประสิทธิ์การตัดสินใจ
- 5) พิจารณาข้อมูลค่าเฉลี่ยรายวันของข้อมูลทั้งหมด เพื่อค้นหาปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน
- 6) ค้นหารูปแบบการวิเคราะห์การถดถอยโลจิสติกที่เหมาะสมในการอธิบายการเกิดปัญหาหมอกควันรายวัน
- 7) พยากรณ์ระดับการเกิดหมอกควัน และตรวจสอบความแม่นยำของตัวแบบ โดยพิจารณาจากรากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าสัมประสิทธิ์การตัดสินใจ
- 8) นำปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันเพื่อพยากรณ์ปริมาณ PM_{10} รายเดือน โดยใช้ขั้นตอนวิธีเชิงพันธุกรรมหาค่าพารามิเตอร์ในสมการการวิเคราะห์การถดถอยพหุคูณ
- 9) พยากรณ์ปริมาณ PM_{10} และตรวจสอบความแม่นยำของตัวแบบ โดยพิจารณาจากรากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าสัมประสิทธิ์การตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 10) เปรียบเทียบการพยากรณ์ปริมาณ PM_{10} ระหว่างการวิเคราะห์อนุกรมเวลากับการวิเคราะห์การถดถอยพหุคูณ โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม
- 11) สรุปผลการวิจัย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงสถานการณ์หมอกควันในประเทศไทย ความหมายของดัชนีคุณภาพอากาศ การคำนวณค่าดัชนีคุณภาพอากาศและเกณฑ์ดัชนีคุณภาพอากาศ พร้อมทั้งกล่าวถึงงานวิจัยที่เกี่ยวข้องในเรื่องของการวิเคราะห์การถดถอยโลจิสติก ขั้นตอนวิธีเชิงพันธุกรรม และการพยากรณ์อนุกรมเวลาโดยวิธีของโฮลต์และวินเนอร์

2.1 สถานการณ์หมอกควันในประเทศไทย

หมอก (fog) เกิดจากการกลั่นตัวของไอน้ำเมื่ออากาศหนาวเย็น ส่วนใหญ่จะเกิดบริเวณที่มีอากาศเย็นท้องฟ้าแจ่มใสไม่มีลม อาจทำให้เกิดน้ำค้าง น้ำค้างแข็งบริเวณใกล้พื้นดิน และหมอกมักเกิดในช่วงเวลากลางคืน หรือเช้ามืด หมอกนั้นไม่จัดว่าเป็นมลพิษทางอากาศ แต่จัดเป็นมลภาวะทางทัศนียภาพของผู้ที่ขับขี่ยานพาหนะ (www.hiso.or.th/hiso/health_event/ghealth_event5.php)

ควัน (smoke) คืออนุภาคเล็กๆ ที่มีขนาด 0.5 – 10 ไมครอน ลอยปะปนอยู่ในอากาศ เกิดจากการเผาไหม้ไม่สมบูรณ์ของเชื้อเพลิงหรือสารอินทรีย์ที่ประกอบด้วยคาร์บอนและออกไซด์ของคาร์บอน

ส่วนหมอกควัน (smog) หมายถึง การผสมของควันหรือฝุ่นในอากาศ ส่วนใหญ่เกิดจากเผาเศษวัสดุเหลือใช้ทางการเกษตร และไฟฟ้า หมอกควันจัดเป็นมลพิษทางอากาศอย่างหนึ่งในบรรดาสารต่างๆ ที่ปะปนอยู่ในอากาศ อาทิเช่น ก๊าซซัลเฟอร์ไดออกไซด์ สารตะกั่ว ก๊าซคาร์บอนมอนอกไซด์ ไนโตรเจนไดออกไซด์ และก๊าซโอโซน เป็นต้น สำหรับหมอกควันจัดเป็นฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน ซึ่งมีผลกระทบต่อสุขภาพของประชาชน แต่การที่เราจะดูว่าฝุ่นละอองนั้นจะเป็นอันตรายต่อสุขภาพหรือไม่ ไม่ได้ขึ้นอยู่กับขนาดของฝุ่นเท่านั้น แต่ยังขึ้นอยู่กับปริมาณฝุ่นที่ได้รับด้วย และเพื่อให้ประชาชนโดยทั่วไปเข้าใจได้ง่ายจึงได้มีการคิดค่าดัชนีที่สามารถบอกถึงระดับคุณภาพอากาศที่อาจเป็นอันตรายต่อสุขภาพในระดับต่างๆ ไว้ด้วย เรียกว่าดัชนีคุณภาพอากาศ ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอนจัดเป็นฝุ่นขนาดเล็ก จะอยู่ในอากาศได้นาน เนื่องจากมีความเร็วในการตกตัวต่ำ และสามารถแขวนลอยอยู่ในอากาศได้นานมากขึ้น หากมีแรงกระทำจากภายนอกเข้ามาเกี่ยวข้อง เช่น การไหลเวียนทางอากาศ กระแสลม เป็นต้น

สถานการณ์หมอกควันในประเทศไทยสำหรับภาคเหนือตอนบนประสบปัญหา 8 จังหวัด ได้แก่ เชียงใหม่ เชียงราย น่าน แพร่ ลำพูน ลำปาง แม่ฮ่องสอน และพะเยา เกิดขึ้นเป็นประจำทุกปีระหว่างเดือนกุมภาพันธ์ถึงเมษายน เนื่องจากช่วงเวลาดังกล่าวมีสภาพอากาศที่แห้งและนิ่ง ทำให้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่บนสื่อออนไลน์
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารมลพิษสามารถแขวนลอยอยู่ในอากาศได้เป็นเวลานาน ทำให้การเกิดมลพิษเพิ่มขึ้นโดยเฉพาะฝุ่นละอองขนาดเล็กที่มีผลกระทบต่อระบบทางเดินหายใจของมนุษย์ สาเหตุหลักเกิดจากไฟฟ้า การเผาเพื่อเตรียมพื้นที่สำหรับการเกษตร หรืออาจเป็นมลพิษหมอกควันข้ามแดนมาจากประเทศเพื่อนบ้าน สำหรับประเทศไทยได้กำหนดค่ามาตรฐานความเข้มข้นเฉลี่ยรายวันของ PM_{10} ไม่เกิน 120 ไมโครกรัมต่อลูกบาศก์เมตร



รูปที่ 2.1 สถานการณ์หมอกควันในพื้นที่จังหวัดเชียงใหม่

(ที่มา: <http://www.riskcomddc.com/th/news/newspaper-detail.php?id=23655>)

ปัญหาหมอกควันถือเป็นปัญหาสำคัญในภาคเหนือตอนบนของประเทศ งานวิจัยตั้งแต่ปี ค.ศ.2008 – 2011 โดยการสนับสนุนการวิจัยของสำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.) พบว่าในช่วงหน้าแล้งของภาคเหนือตอนบนมักเกิดมลพิษทางอากาศจากหมอกควัน โดยพบว่า ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) เป็นสารมลพิษหลัก แล้วยังมีงานวิจัยอีกว่า ปัญหาหมอกควันในเขตนอกเมืองจะมีความรุนแรงสูงกว่า ทั้งนี้เพราะสาเหตุหลักที่ทำให้เกิดปัญหาหมอกควันคือ ไฟป่าและการเผาเศษวัสดุเพื่อเตรียมพื้นที่สำหรับการเกษตรของเกษตรกร ซึ่งเป็นพื้นที่นอกเมืองเสียส่วนมาก ประกอบกับสภาพภูมิประเทศซึ่งมีภูเขาสูงล้อมรอบ ทำให้มลพิษต่างๆ ถูกกักไว้และแผ่ปกคลุมทั่วเมือง ซึ่งสภาวะอากาศที่แห้งและนิ่งทำให้ฝุ่นละอองที่เกิดขึ้นสามารถแขวนลอยอยู่ในบรรยากาศได้นาน นอกจากนี้ปริมาณฝนที่ตกน้อยทำให้การชะล้างหมอกควันหรือฝุ่นที่แขวนลอยอยู่ในอากาศเป็นไปได้น้อย ทั้งนี้ค่ามาตรฐานฝุ่นละอองขนาดเล็กกว่า 10 ไมครอนในบรรยากาศโดยทั่วไปในประเทศไทยที่ทางกรมควบคุมมลพิษได้กำหนดไว้คือ ค่าเฉลี่ยในเวลา 24 ชั่วโมงจะต้องไม่เกิน 120 ไมโครกรัมต่อลูกบาศก์เมตร และค่าเฉลี่ยในเวลา 1 ปี จะต้องไม่เกิน 50 ไมโครกรัมต่อลูกบาศก์เมตร ซึ่งปัญหานี้ส่งผลกระทบต่อสภาพบรรยากาศทั่วไป ต่อสุขภาพอนามัยต่อธุรกิจการท่องเที่ยวและบริการในหลายพื้นที่ (สำนักจัดการคุณภาพอากาศและเสียง, 2554) นอกจากนี้ยังส่งผลกระทบต่อทัศนวิสัยดังแสดงในรูปที่ 2.1

การป้องกันและแก้ไขปัญหานี้ จะต้องมีการเฝ้าระวังและติดตามตรวจสอบปริมาณฝุ่นละอองขนาดเล็กในพื้นที่ และรายงานข้อมูลให้หน่วยงานที่เกี่ยวข้องรับทราบ ดังนั้นการติดตามตรวจสอบและคาดการณ์สถานการณ์การเกิดปัญหาหมอกควันนี้ ซึ่งมีส่วนสำคัญเพื่อการเตรียมพร้อมรับมือสถานการณ์ที่จะเกิดขึ้นได้อย่างมีประสิทธิภาพ

ระบบเครือข่ายการติดตามตรวจสอบคุณภาพอากาศในบรรยากาศของประเทศไทยจัดตั้งขึ้นโดยมีวัตถุประสงค์เพื่อติดตามตรวจสอบสภาพปัญหามลพิษที่เกิดขึ้นในพื้นที่ต่างๆของประเทศไทย เพื่อเฝ้าระวังสภาพปัญหา ศึกษาและคาดการณ์การเปลี่ยนแปลงของคุณภาพอากาศที่เกิดขึ้นในพื้นที่ต่างๆ ซึ่งเป็นข้อมูลสำคัญที่จะนำไปใช้กำหนดมาตรการและแนวทางการแก้ไขปัญหา การลดและป้องกันปัญหาที่อาจจะเกิดขึ้นได้ โดยได้เริ่มดำเนินการตั้งแต่ปี พ.ศ. 2520 เป็นต้นมา ระบบเครือข่ายการติดตามตรวจสอบคุณภาพอากาศในบรรยากาศของประเทศไทย แบ่งเป็น 2 ส่วน คือ

1) สถานีตรวจวัดคุณภาพอากาศในบรรยากาศ (Ambient Air Monitoring Station)

ประกอบด้วยสถานีตรวจวัดแบบกึ่งถาวร (สามารถเคลื่อนย้ายได้ถ้าหากมีความจำเป็น) และหน่วยตรวจวัดคุณภาพอากาศแบบเคลื่อนที่ (Mobile Unit) โดยติดตั้งอุปกรณ์เชื่อมต่อข้อมูลไปยังกรมควบคุมมลพิษ



รูปที่ 2.2 สถานีตรวจวัดคุณภาพอากาศในบรรยากาศ
(ที่มา: สำนักจัดการคุณภาพอากาศและเสียง, 2554)

2) ระบบจัดเก็บและวิเคราะห์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นระบบคอมพิวเตอร์ที่ไทยพัฒนาซอฟต์แวร์ขึ้นเอง โดยมีพื้นฐานมาจากซอฟต์แวร์ของรัฐบาลสวีเดน สามารถเรียกใช้ข้อมูล ระบบฐานข้อมูลทางภูมิศาสตร์ แหล่งกำเนิดมลพิษ และแบบจำลองทางคณิตศาสตร์สำหรับประเมินการแพร่กระจายของสารมลพิษทางอากาศ โดยมีศูนย์ควบคุมส่วนกลางตั้งอยู่ที่ กรมควบคุมมลพิษ กรุงเทพมหานคร และมีสถานีตรวจวัดคุณภาพอากาศรวมทั้งสิ้น 57 แห่งทั่วประเทศไทย รวมทั้งสถานีตรวจวัดสภาพอากาศทางอุตุนิยมวิทยาอีก 5 แห่ง ทำให้ทราบสถานการณ์คุณภาพอากาศและภาวะมลพิษทางอากาศได้อย่างต่อเนื่อง ในปัจจุบันมีการรายงานข้อมูลสารมลพิษอากาศรายชั่วโมงของสถานีตรวจวัดต่างๆ ในสื่อออนไลน์ผ่าน www.aqmthai.com และสามารถเรียกดูข้อมูลย้อนหลังได้อีกด้วย

2.2 ดัชนีคุณภาพอากาศ (Air Quality Index)

ดัชนีคุณภาพอากาศ เป็นการรายงานข้อมูลคุณภาพอากาศในรูปแบบที่ง่ายต่อความเข้าใจของประชาชนทั่วไป เพื่อเผยแพร่ประชาสัมพันธ์ให้สาธารณชนได้ทราบถึงสถานการณ์มลพิษทางอากาศในแต่ละพื้นที่ว่าอยู่ในระดับใด มีผลกระทบต่อสุขภาพอนามัยหรือไม่ ซึ่งดัชนีคุณภาพอากาศเป็นรูปแบบสากลที่ใช้กันอย่างแพร่หลายในหลายประเทศ เช่น สหรัฐอเมริกา ออสเตรเลีย สิงคโปร์ มาเลเซีย ประเทศไทย เป็นต้น

ดัชนีคุณภาพอากาศที่ใช้อยู่ในประเทศไทยมีการคำนวณเทียบจากมาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไปของสารมลพิษทางอากาศ 5 ประเภท ได้แก่

1. ก๊าซโอโซน (O_3) เฉลี่ย 1 ชั่วโมง
2. ก๊าซไนโตรเจนไดออกไซด์ (NO_2) เฉลี่ย 1 ชั่วโมง
3. ก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) เฉลี่ย 24 ชั่วโมง
4. ก๊าซคาร์บอนมอนอกไซด์ (CO) เฉลี่ย 8 ชั่วโมง
5. ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) เฉลี่ย 24 ชั่วโมง

สำนักจัดการคุณภาพอากาศและเสียง ในสังกัดกรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม เป็นหน่วยงานที่รับผิดชอบตรวจวัดสารมลพิษทางอากาศดังกล่าว คำนวณค่าดัชนีคุณภาพอากาศ และรายงานค่าดัชนีคุณภาพอากาศรายวันในแต่ละพื้นที่ โดยค่าที่คำนวณได้ของสารมลพิษทางอากาศประเภทใดมีค่าสูงสุด จะใช้เป็นค่าชี้วัดดัชนีคุณภาพอากาศของวันนั้น (วนิดา จินศาสตร์, 2551)

สำนักจัดการคุณภาพอากาศและเสียง มีสถานีตรวจวัดอากาศทั้งในกรุงเทพมหานคร ปริมณฑล และจังหวัดในภูมิภาค รวมทั้งสิ้น 19 จังหวัด ได้แก่ สถานีตรวจวัดอากาศจังหวัด เอกสกร จังหวัดกรุงเทพมหานคร สมุทรปราการ สมุทรสาคร ปทุมธานี นนทบุรี เชียงใหม่ ลำปาง นครสวรรค์ คาร์ค้ำไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สระบุรี นครราชสีมา ขอนแก่น อุดรธานี ราชบุรี ชลบุรี ระยอง ฉะเชิงเทรา สุราษฎร์ธานี ภูเก็ต และ สงขลา โดยมีการตรวจวัดคุณภาพอากาศจากพารามิเตอร์หลายชนิด ได้แก่ ความเข้มข้นของก๊าซ ฝุ่นละออง (จำแนกตามขนาด) อุณหภูมิ ทิศทางลม ความเร็วลม ปริมาณน้ำฝน และความชื้นสัมพัทธ์ แต่ละประเทศมีการกำหนดเกณฑ์ดัชนีคุณภาพอากาศและความเข้มข้นของสารมลพิษทางอากาศที่เทียบเท่าดัชนีคุณภาพอากาศตามตารางที่ 2.1 และ 2.2 ตามลำดับ แตกต่างกันไป

ตารางที่ 2.1 เกณฑ์ดัชนีคุณภาพอากาศสำหรับประเทศไทย

AQI	ความหมาย	สีที่ใช้	ระดับของผลกระทบต่อสุขภาพอนามัย
0-50	คุณภาพดี	ฟ้า	ไม่มีผลกระทบต่อสุขภาพ
51-100	คุณภาพปานกลาง	เขียว	ไม่มีผลกระทบต่อสุขภาพ
101-200	มีผลกระทบต่อสุขภาพ	เหลือง	ผู้ป่วยโรคระบบทางเดินหายใจควรหลีกเลี่ยงการออกกำลังกายนอกอาคาร บุคคลทั่วไป โดยเฉพาะเด็กและผู้สูงอายุไม่ควรทำกิจกรรมภายนอกอาคารเป็นเวลานาน
201-300	มีผลกระทบต่อสุขภาพมาก	ส้ม	ผู้ป่วยโรคระบบทางเดินหายใจควรหลีกเลี่ยงกิจกรรมภายนอกอาคาร บุคคลทั่วไป โดยเฉพาะเด็กและผู้สูงอายุควรจำกัดการออกกำลังกายนอกอาคาร
มากกว่า 300	อันตราย	แดง	บุคคลทั่วไป ควรหลีกเลี่ยงการออกกำลังกายนอกอาคาร สำหรับผู้ป่วยโรคระบบทางเดินหายใจ ควรอยู่ในอาคาร

ตารางที่ 2.2 ความเข้มข้นของสารมลพิษทางอากาศที่เทียบเท่ากับดัชนีคุณภาพอากาศ

AQI	PM ₁₀ (24 hr.)	O ₃ (1 hr.)		SO ₂ (24 hr.)		NO ₂ (1 hr.)		CO (8 hr.)	
	µg./m ³	µg./m ³	ppb	µg./m ³	ppb	µg./m ³	ppb	µg./m ³	ppb
50	40	100	51	65	25	160	85	5.13	4.48
100	120	200	100	300	120	320	170	10.26	9.00
200	350	400	203	800	305	1,130	600	17.00	14.84
300	420	800	405	1,600	610	2,260	1,202	34.00	29.69
400	500	1,000	509	2,100	802	3,000	1,594	46.00	40.17
500	600	1,200	611	2,620	1,000	3,750	1,993	57.50	50.21

ที่มา United States Environmental Protection Agency, July 1999, Guideline for Reporting of Daily Air Quality - Air Quality Index (AQI), 40 CFR Part 58, Appendix G.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การคำนวณดัชนีคุณภาพอากาศรายวันของสารมลพิษทางอากาศแต่ละประเภท

การคำนวณดัชนีคุณภาพอากาศรายวัน คำนวณจากค่าความเข้มข้นของสารมลพิษทางอากาศจากข้อมูลผลการตรวจวัดคุณภาพอากาศ โดยแต่ละระดับของค่าความเข้มข้นของสารมลพิษทางอากาศเทียบเท่ากับค่าดัชนีคุณภาพอากาศที่ระดับต่างๆดังตารางที่ 2.2

(http://www.pcd.go.th/info_serv/air_aqi.htm)

สูตรการคำนวณมีดังนี้

$$I_i = \frac{I_{ij+1} - I_{ij}}{X_{ij+1} - X_{ij}} (X_i - X_{ij}) + I_{ij} \quad (2.1)$$

กำหนดให้

- X_i = ความเข้มข้นของสารมลพิษทางอากาศจากผลการตรวจวัด
 X_{ij} = ความเข้มข้นของสารมลพิษทางอากาศที่เป็นค่าต่ำสุดของช่วงพิสัยที่มีค่า X_i นั้น
 X_{ij+1} = ความเข้มข้นของสารมลพิษทางอากาศที่เป็นค่าสูงสุดของช่วงพิสัยที่มีค่า X_i นั้น
 I_i = ค่าดัชนีย่อยคุณภาพอากาศ
 I_{ij} = ค่าดัชนีย่อยคุณภาพอากาศที่เป็นค่าต่ำสุดของช่วงพิสัยที่มีค่า I_i นั้น
 I_{ij+1} = ค่าดัชนีย่อยคุณภาพอากาศที่เป็นค่าสูงสุดของช่วงพิสัยที่มีค่า I_i นั้น

ปัจจัยที่มีผลต่อการรายงานดัชนีคุณภาพอากาศ ได้แก่ พื้นที่ครอบคลุม ระยะเวลา ประเภทของสารมลพิษ เช่น ในพื้นที่ภาคเหนือตอนบนจะเกิดสถานการณ์หมอกควันในช่วงระหว่างเดือนกุมภาพันธ์ถึงเมษายนของทุกปี ซึ่งจะมีการรายงานดัชนีคุณภาพอากาศรายวันสม่ำเสมอในระยะเวลาดังกล่าว โดยเฉพาะอย่างยิ่งการรายงานค่า PM_{10} ซึ่งเป็นสารมลพิษที่สำคัญที่ใช้ในการบ่งบอกถึงระดับความรุนแรงของสถานการณ์หมอกควัน

ตัวอย่างการคำนวณดัชนีคุณภาพอากาศ

จงคำนวณค่าดัชนีคุณภาพอากาศของข้อมูลการตรวจวัดของก๊าซ CO ที่ความเข้มข้น 8 ชั่วโมงเฉลี่ยเท่ากับ 0.29 ppb ก๊าซ O_3 ที่ความเข้มข้น 1 ชั่วโมงเฉลี่ยเท่ากับ 69 ppb ก๊าซ NO_2 ที่ความเข้มข้น 1 ชั่วโมงเฉลี่ยเท่ากับ 12 ppb ก๊าซ SO_2 ที่ความเข้มข้น 24 ชั่วโมงเฉลี่ยเท่ากับ 40 ppb และ PM_{10} ที่ความเข้มข้น 24 ชั่วโมงเฉลี่ยเท่ากับ $41 \mu g/m^3$

วิธีทำ นำค่าที่ตรวจวัดได้และค่าในตารางที่ 2.2 แทนค่าลงในสมการ (2.1)

$$I_{CO} = \frac{(50 - 0)(0.29 - 0)}{(4.48 - 0)} + 0 = 3.24$$

$$I_{O_3} = \frac{(100 - 50)(69 - 51)}{(100 - 51)} + 50 = 68.37$$

$$I_{NO_2} = \frac{(50 - 0)(12 - 0)}{(85 - 0)} + 0 = 7.06$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับ (85=0) งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$I_{SO_2} = \frac{(50 - 0)(27 - 0)}{(25 - 0)} + 0 = 54$$

$$I_{PM_{10}} = \frac{(100 - 50)(41 - 40)}{(120 - 40)} + 50 = 50.625$$

ดัชนีคุณภาพอากาศของก๊าซ O_3 มีค่ามากที่สุด คือ 68.37 เราจะใช้ค่า 68.37 เป็นค่าดัชนีคุณภาพอากาศของวันนี้ ซึ่งถืออยู่ในเกณฑ์คุณภาพอากาศปานกลาง ไม่มีผลต่อสุขภาพ แทนด้วยสัญลักษณ์สีเขียว

สำหรับในช่วงเวลาที่เกิดสถานการณ์หมอกควันนั้น เราจะพบว่า PM_{10} เป็นสารมลพิษหลักที่ทำให้ค่าดัชนีคุณภาพอากาศสูงสุด

2.3 การวิเคราะห์สหสัมพันธ์

การวิเคราะห์สหสัมพันธ์ (Correlation Analysis) เป็นวิธีการที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว โดยสัมประสิทธิ์สหสัมพันธ์ (r) มีค่าตั้งแต่ -1 ถึง 1 ค่าลบแสดงความสัมพันธ์ทางตรงกันข้าม ค่าบวกแสดงความสัมพันธ์ทางเดียวกัน (กัลยา วานิชย์บัญชา, 2548)

$r = 0.50$ ถึง 1.00 หรือ $r = -0.50$ ถึง -1.00 ถือว่าข้อมูลมีความสัมพันธ์ในระดับสูง

$r = 0.30$ ถึง 0.49 หรือ $r = -0.30$ ถึง -0.49 ถือว่าข้อมูลมีความสัมพันธ์ในระดับปานกลาง

$r = 0.10$ ถึง 0.29 หรือ $r = -0.10$ ถึง -0.29 ถือว่าข้อมูลมีความสัมพันธ์ในระดับต่ำ

$r = 0.00$ ถือว่าข้อมูลไม่มีความสัมพันธ์กัน

โดยคำนวณได้จากสมการ

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \quad (2.2)$$

ปัจจัยที่ใช้ในการศึกษาครั้งนี้ประกอบด้วย 12 ปัจจัย ได้แก่ ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) ก๊าซคาร์บอนมอนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO_2) ความชื้นสัมพัทธ์ (Relative humidity : RH) ความดัน (Pressure) ปริมาณน้ำฝน (Rain) ก๊าซไนโตรเจนมอนอกไซด์ (NO) ทิศทางลม (Wind direction) ความเร็วลม (Wind Speed) อุณหภูมิ (Temperature) ก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) และก๊าซโอโซน (O_3)

ในการศึกษาครั้งนี้ได้ทำการค้นหาปัจจัยที่เกี่ยวข้องกับการเกิดหมอกควัน โดยใช้สัมประสิทธิ์สหสัมพันธ์มาช่วยในการตัดสินใจเพื่อหาปัจจัยที่มีอิทธิพลต่อการเกิดหมอกควันนั้นคือ หาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยต่างๆกับ PM_{10} เนื่องจาก PM_{10} เป็นตัวแปรสำคัญที่ใช้บ่งชี้ว่าสถานการณ์หมอกควันในพื้นที่ที่มีความรุนแรงมากน้อยเพียงใด

2.4 การวิเคราะห์อนุกรมเวลา (Time series)

การวิเคราะห์การถดถอยเป็นการพิจารณาค่าของข้อมูลในช่วงเวลาใดเวลาหนึ่ง โดยไม่คำนึงถึงการเปลี่ยนแปลงของเวลา และพิจารณาการเปลี่ยนแปลงของตัวแปรตามซึ่งเป็นผลมาจากตัวแปรอิสระ แต่การวิเคราะห์อนุกรมเวลานั้นเป็นการพิจารณาค่าของข้อมูลที่เกิดการเปลี่ยนแปลงอันเนื่องมาจากการเปลี่ยนแปลงของเวลาเป็นสำคัญ

ข้อมูลอนุกรมเวลา (Time Series Data) หมายถึง ข้อมูลหรือค่าสังเกตที่เปลี่ยนแปลงไปตามลำดับของเวลา โดยข้อมูลที่สังเกตจะเป็นข้อมูลที่อยู่ในช่วงเวลาที่สนใจของผู้วิเคราะห์ ซึ่งจะให้เห็นรูปแบบของการเปลี่ยนแปลงค่าสังเกตในช่วงเวลาที่ผ่านมา เพื่อใช้ในการพยากรณ์ค่าสังเกตดังกล่าวในอนาคต

วิธีการพยากรณ์อนุกรมเวลา (Time Series Forecasting Method) เป็นการวิเคราะห์พฤติกรรมในอดีตของตัวแปรอนุกรมเวลา เพื่อพยากรณ์พฤติกรรมในอนาคต โดยถ้าสามารถค้นพบพฤติกรรมที่เป็นระบบบางอย่างในตัวแปรอนุกรมเวลา ผู้วิเคราะห์ก็จะสามารถสร้างแบบจำลองของพฤติกรรมของตัวแปรตาม แล้วนำมาใช้ในการพยากรณ์พฤติกรรมของตัวแปรเหล่านั้นในอนาคต เช่น สมมติว่ามีข้อมูลอนุกรมเวลาเรียงตามระยะเวลา $t-n$ จนถึง t ดังตารางที่ 2.3

ตารางที่ 2.3 ข้อมูลอนุกรมเวลา

ช่วงเวลาที่	ค่าของตัวแปร
$t-n$	Y_{t-n}
\vdots	\vdots
$t-1$	Y_{t-1}
t	Y_t

เทคนิคที่ใช้วิเคราะห์พฤติกรรมของตัวแปรอนุกรมเวลาเพื่อพยากรณ์ค่าตัวแปรในอนาคต คือ การใช้แบบจำลอง Extrapolation ซึ่งมีรูปแบบทางคณิตศาสตร์โดยทั่วไป ดังนี้

$$\hat{Y}_{t+1} = f(Y_t, Y_{t-1}, Y_{t-2}, \dots, Y_{t-n}) \quad (2.3)$$

โดย \hat{Y}_{t+1} คือ ค่าตัวแปรอนุกรมเวลาที่พยากรณ์ได้ในช่วงเวลา $t+1$

Y_t คือ ค่าตัวแปรอนุกรมเวลาในช่วงเวลา t

Y_{t-1} คือ ค่าตัวแปรอนุกรมเวลาในช่วงเวลา $t-1$

จุดประสงค์ของแบบจำลอง Extrapolation คือ การระบุฟังก์ชัน $f()$ ตามสมการ (2.3)

เพื่อใช้พยากรณ์ค่าของตัวแปรอนุกรมเวลาในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

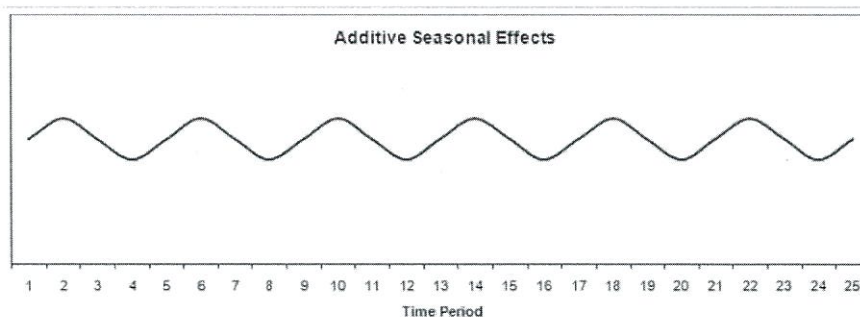
ข้อมูลอนุกรมเวลาที่เก็บรวบรวมได้ในช่วงระยะเวลาที่เท่ากันจะมีค่าเปลี่ยนแปลงซึ่งจะเป็นตัวแปรตาม ส่วนระยะเวลาที่เปลี่ยนแปลงจะเป็นตัวแปรอิสระ สาเหตุที่ทำให้อนุกรมเวลาเปลี่ยนแปลงจะเรียกว่า ส่วนประกอบของอนุกรมเวลา (Components of a time series) ซึ่งจะประกอบด้วยสาเหตุ 4 ชนิด (มุกดา แม้นมินทร์, 2549) ดังนี้

1. แนวโน้ม (Trend : T) เป็นสาเหตุที่ทำให้ข้อมูลอนุกรมเวลาเปลี่ยนแปลงในระยะเวลา มากกว่าหนึ่งปี การเปลี่ยนแปลงอาจจะเพิ่มขึ้นหรือลดลง การเปลี่ยนแปลงอาจจะเป็นไปอย่างรวดเร็วหรือช้าๆ ลักษณะแนวโน้มของการเปลี่ยนแปลงอาจจะเป็นเชิงเส้น หรือเส้นโค้งพาราโบลา หรือเส้นโค้งเอ็กโปเนนเชียล รูปที่ 2.3 แสดงยอดขายของสินค้าชนิดหนึ่งซึ่งมีแนวโน้มสูงขึ้นเมื่อเวลาผ่านไป

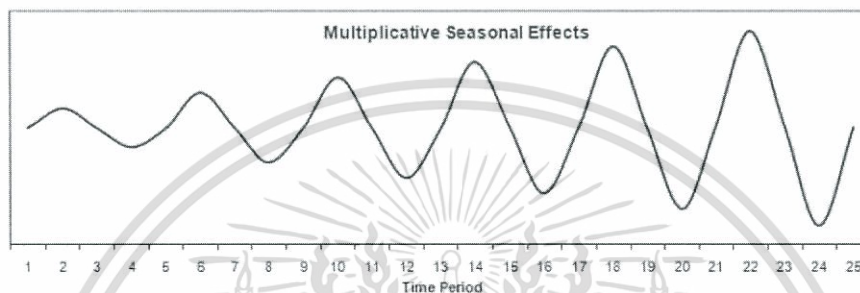


รูปที่ 2.3 แนวโน้มยอดขายของสินค้าชนิดหนึ่ง

2. การแปรผันตามฤดูกาล (Seasonal Variation : S) เป็นสาเหตุที่ทำให้ข้อมูลอนุกรมเวลาเปลี่ยนแปลงในระยะเวลาสั้นๆ อาจจะเป็นรายวัน รายสัปดาห์ รายเดือน หรือรายไตรมาส จะเรียกช่วงระยะเวลานี้ว่า ฤดูกาล การเปลี่ยนแปลงอาจจะเพิ่มขึ้นหรือลดลงเกิดขึ้นซ้ำๆกัน ในรอบของช่วงเวลาที่ไม่เกินหนึ่งปี ค่าที่วัดการเปลี่ยนแปลงตามฤดูกาล เรียกว่า ดัชนีฤดูกาล (Seasonal Index) สามารถแบ่งเป็น 2 ลักษณะ คือ การแปรผันตามฤดูกาลเชิงบวก และการแปรผันตามฤดูกาลเชิงลบ ดังรูปที่ 2.4 และรูปที่ 2.5 ตามลำดับ การแปรผันตามฤดูกาลเชิงบวกเป็นการเปลี่ยนแปลงที่มีฤดูกาลโดยมีค่าซ้ำๆ ไม่ต่างจากฤดูกาลก่อนหน้ามากนัก ส่วนการแปรผันตามฤดูกาลเชิงลบเป็นการเปลี่ยนแปลงที่เพิ่มขึ้นเป็นสัดส่วนกับเวลา

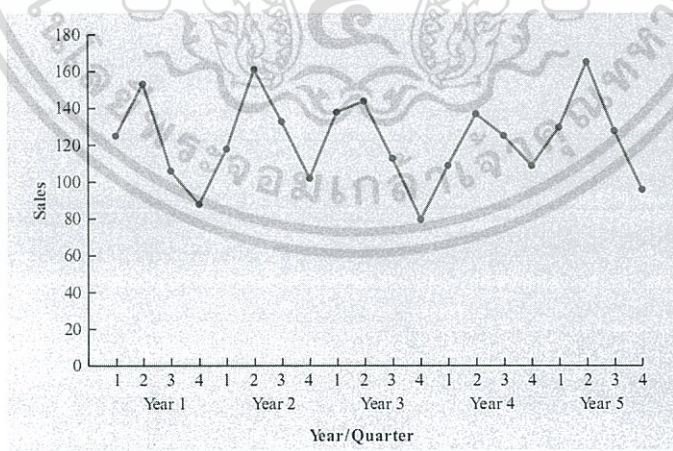


รูปที่ 2.4 ข้อมูลอนุกรมเวลาที่มีการแปรผันตามฤดูกาลเชิงบวก



รูปที่ 2.5 ข้อมูลอนุกรมเวลาที่มีการแปรผันตามฤดูกาลเชิงคูณ

3. การแปรผันตามวัฏจักร (Cyclical Variation : C) เป็นสาเหตุที่ทำให้ข้อมูลอนุกรมเวลาเปลี่ยนแปลงขึ้นและลงซ้ำๆกัน ในช่วงระยะเวลาที่ยาวนานกว่าการแปรผันตามฤดูกาล ช่วงระยะเวลาการเปลี่ยนแปลงไม่แน่นอน รูปที่ 2.6 แสดงปริมาณการขายสินค้าหนึ่งที่มีการแปรผันตามวัฏจักร



รูปที่ 2.6 ข้อมูลอนุกรมเวลาที่มีการแปรผันตามวัฏจักรของปริมาณการขายสินค้าหนึ่ง

4. การแปรผันเนื่องจากเหตุการณ์ไม่ปกติ (Irregular Variation: I) เป็นสาเหตุที่ทำให้ข้อมูลอนุกรมเวลาเปลี่ยนแปลงขึ้นและลงในระยะเวลาที่ไม่แน่นอน มักจะเกิดในช่วงระยะเวลาสั้นๆ ซึ่งจะเกิดจากเหตุการณ์ที่ไม่สามารถคาดคะเนได้ล่วงหน้า เช่น เหตุการณ์ที่สหรัฐอเมริกาเกิดเอกลีออนเป็นเอกลีออนที่ส่งมอบเวลาให้กับบริการเชิงงานเพื่อการศึกษาเท่านั้น ไม่นับเข้าตาเห็นไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วินาศกรรมทำให้สถานะเศรษฐกิจของสหรัฐอเมริกาถดถอย เหตุการณ์น้ำท่วมพื้นที่สวนผักทำให้ผักมีราคาสูงขึ้น เหตุการณ์สงครามในประเทศผลิตน้ำมันทำให้ราคาน้ำมันในตลาดโลกสูงขึ้น เป็นต้น

ข้อมูลอนุกรมเวลาอาจได้รับอิทธิพลของปัจจัยที่เป็นส่วนประกอบของอนุกรมเวลาทั้ง 4 ปัจจัยหรือเพียงปัจจัยใดปัจจัยหนึ่งเท่านั้น สำหรับข้อมูลที่ได้ศึกษาในครั้งนี้พบว่า เป็นข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงตามฤดูกาลเชิงบวก ซึ่งสามารถพยากรณ์ได้ด้วยวิธีของโฮลต์และวินเทอร์ (Holt – Winter’s Method) สำหรับการศึกษาค้างนี้ การพยากรณ์อนุกรมเวลาที่มีการเปลี่ยนแปลงตามฤดูกาลเชิงบวกโดยวิธีของโฮลต์และวินเทอร์ (Holt-Winter’s method for additive seasonal effects) ซึ่งอธิบายได้ด้วยตัวแบบ

$$Z_t = \beta_0 + \beta_1 t + S_t + \alpha_t$$

เมื่อ β_0 แสดงระดับของแนวโน้ม

β_1 แทนความชันของแนวโน้มเชิงเส้น

S_t แทนการแปรผันตามฤดูกาล

α_t เป็นตัวแปรสุ่มอิสระที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวน σ_a^2

การพยากรณ์ตามตัวแบบนี้เป็นการที่เราสามารถปรับค่าประมาณได้ทุกหน่วยเวลาจากค่าประมาณเริ่มต้นเมื่อเวลา t_0 และค่าคงที่ในการทำให้เรียบที่เลือกอย่างเหมาะสม ซึ่งสามารถอธิบายการปรับค่าประมาณของ β_0 , β_1 และ S_t สำหรับเวลา n ใดๆ ได้ดังนี้

1. เมื่อมีค่าสังเกต Z_n ที่เวลา n ใดๆ ค่าประมาณของ β_0 อาจปรับค่าได้ในรูป

$$\hat{\beta}_0(n) = \alpha_1(Z_n - \hat{S}_{n-s}) + (1 - \alpha_1)[\hat{\beta}_0(n-1) + \hat{\beta}_1(n-1)] \quad (2.4)$$

เมื่อ $0 < \alpha_1 < 1$ เป็นค่าคงที่ในการทำให้เรียบ

2. ค่าประมาณของ β_1 อาจปรับค่าได้ในรูป

$$\hat{\beta}_1(n) = \alpha_2[\hat{\beta}_0(n) - \hat{\beta}_0(n-1)] + (1 - \alpha_2)\hat{\beta}_1(n-1) \quad (2.5)$$

เมื่อ $0 < \alpha_2 < 1$ เป็นค่าคงที่ในการทำให้เรียบ

3. ค่าที่แสดงการแปรผันตามฤดูกาลที่เป็นค่าสังเกต $Z_n - \hat{\beta}_0(n)$ ในรูป

$$\hat{S}_n = \alpha_3[Z_n - \hat{\beta}_0(n)] + (1 - \alpha_3)\hat{S}_{n-s} \quad (2.6)$$

เมื่อ $0 < \alpha_3 < 1$ เป็นค่าคงที่ในการทำให้เรียบ

4. ค่าพยากรณ์ที่เวลา $n+l$ โดยวิธีของวินเทอร์แสดงได้ในรูป

$$\hat{Z}_n(l) = \hat{\beta}_0(n) + \hat{\beta}_1(n)l + \hat{S}_{n+l-s} \quad (2.7)$$

ในกรณีนี้ค่าคงที่ในการทำให้เรียบที่ต้องกำหนดคือ α_1 , α_2 และ α_3 สามารถหาได้จาก

โปรแกรมสำเร็จรูป (Microsoft office excel) ซึ่งจะให้ผลลัพธ์ α_1 , α_2 และ α_3 ที่ให้ค่าความ

เอกสารคัดค้านน้อยที่สุด (พนิดา พานิชกุล, 2546) การศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression Analysis)

การวิเคราะห์การถดถอยโลจิสติก เป็นเทคนิคการวิเคราะห์เชิงคุณภาพ (Qualitative statistical techniques) ที่ต่างจากเทคนิคการวิเคราะห์เชิงปริมาณในเรื่องของข้อมูลที่เป็นตัวแปรตามเป็นตัวแปรเชิงคุณภาพหรือตัวแปรเชิงกลุ่ม การวิเคราะห์การถดถอยโลจิสติกถูกนำมาใช้เพื่อทำนายว่าจะเกิดเหตุการณ์หนึ่งขึ้นหรือไม่ หรือมีโอกาสเกิดขึ้นมากน้อยเพียงใด โดยมีการกำหนดค่าของตัวแปรอิสระหนึ่งตัวหรือหลายตัวที่คาดว่าจะมีผลต่อการเกิดเหตุการณ์นั้นๆ ซึ่งจะช่วยให้เราเข้าใจสาเหตุของการเกิดเหตุการณ์นั้นๆ ได้ในที่สุด (กัลยา วานิชบัญญัติ, 2548)

2.5.1 ประเภทของการวิเคราะห์การถดถอยโลจิสติก

ประเภทของการวิเคราะห์การถดถอยโลจิสติกมีอยู่หลายประเภทขึ้นอยู่กับเกณฑ์ในการแบ่ง เช่น

1. ใช้จำนวนตัวแปรอิสระเป็นเกณฑ์ในการแบ่ง สามารถแบ่งได้ 2 ประเภท คือ
 - 1.1 การวิเคราะห์การถดถอยโลจิสติกอย่างง่าย (Simple logistic regression analysis) เป็นการวิเคราะห์ที่ประกอบด้วยตัวแปรอิสระเพียง 1 ตัวเท่านั้น
 - 1.2 การวิเคราะห์การถดถอยโลจิสติกเชิงพหุ (Multiple logistic regression analysis) คือเป็นการวิเคราะห์ที่ประกอบด้วยตัวแปรอิสระตั้งแต่ 2 ตัวขึ้นไป
2. ใช้จำนวนกลุ่มของตัวแปรตามเป็นเกณฑ์ในการแบ่ง สามารถแบ่งได้ 2 ประเภท คือ
 - 2.1 การวิเคราะห์การถดถอยโลจิสติกแบบ 2 กลุ่ม (Binary logistic regression analysis)
 - 2.2 การวิเคราะห์การถดถอยโลจิสติกแบบหลายกลุ่ม (Multinomial logistic regression)

2.5.2 วัตถุประสงค์ของการวิเคราะห์การถดถอยโลจิสติก

1. เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม สามารถระบุได้ว่าตัวแปรอิสระใดบ้างที่สามารถใช้อธิบายโอกาสการเกิดเหตุการณ์หรือการไม่เกิดเหตุการณ์ที่สนใจ
2. เพื่อทำนายโอกาสที่จะเกิดเหตุการณ์ที่สนใจจากสมการ โลจิสติกที่เหมาะสม โดยเลือกตัวแปรที่เหมาะสม เพื่อให้เปอร์เซ็นต์ของความถูกต้องในการทำนายมีค่าสูงสุด (ศิริชัย กาญจนวาสี, 2551)

2.5.3 ข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยโลจิสติก

การวิเคราะห์การถดถอยโลจิสติกมีข้อตกลงเบื้องต้น ดังนี้

1. ตัวแปรอิสระอาจเป็นข้อมูลชนิดที่มีได้ 2 ค่า (Dichotomous variable) หรือเป็นสเกลอัตราภาค (Interval scale) และสเกลอัตราส่วนก็ได้ (Ratio scale)
2. ค่าคาดหวังของค่าความคลาดเคลื่อนเป็นศูนย์ หรือ $E(e) = 0$

3 ตัวแปรอิสระไม่ควรมีความสัมพันธ์กันมากเกินไป (multicollinearity) โดยถ้าใช้เกณฑ์ของ Burns and Grove (1993) กำหนดค่าสัมประสิทธิ์สหสัมพันธ์ไม่เกิน 0.65 และถ้าใช้เกณฑ์ของ Stevens (1996) กำหนดค่าสัมประสิทธิ์สหสัมพันธ์ไม่เกิน 0.80

4 การวิเคราะห์การถดถอยโลจิสติกจะต้องใช้ขนาดตัวอย่าง n มากกว่า การวิเคราะห์การถดถอยแบบปกติ โดยทั่วไป $n \geq 30p$ โดยที่ p เป็นจำนวนตัวแปรอิสระ

สำหรับการวิเคราะห์การถดถอยแบบปกติ นอกจากจะมีเงื่อนไขทั้ง 5 ข้อข้างต้นจะต้องเพิ่มเงื่อนไขอีก 2 ข้อ คือ

- 1 ค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ
- 2 ค่าแปรปรวนของค่าคลาดเคลื่อนคงที่

2.5.4 โมเดลการวิเคราะห์การถดถอยโลจิสติกอย่างง่าย (Simple logistic regression analysis model)

ในกรณีนี้ตัวแปรตามมีค่าได้เพียง 2 ค่า ซึ่งมักแทนด้วย 0 และ 1 เช่นเมื่อเราสนใจการเกิดเหตุการณ์ของ Y จะได้ 1 แทนการเกิดเหตุการณ์ และ 0 แทนการไม่เกิดเหตุการณ์ ดังนั้นตัวแปรตามจะมีการแจกแจงแบบเบอร์นูลลี (Bernoulli distribution)

$$P\{Y = y\} = p^y(1-p)^{1-y} \quad ; \quad y = 0, 1 \quad (2.8)$$

สำหรับตัวอย่างหน่วยที่ i จะได้ว่า

$$P\{Y_i = y_i\} = p^{y_i}(1-p)^{1-y_i} \quad ; \quad y_i = 0, 1 \quad (2.9)$$

จากสมการที่ (2.9) เมื่อ $y_i = 0$ จะได้

$$P\{Y_i = 0\} = p^0(1-p)^{1-0} = 1-p$$

จากสมการที่ (2.9) เมื่อ $y_i = 1$ จะได้

$$P\{Y_i = 1\} = p^1(1-p)^{1-1} = p$$

$$E\{Y_i\} = \sum y_i P\{Y_i = y_i\}$$

$$= 0 \cdot P\{Y_i = 0\} + 1 \cdot P\{Y_i = 1\}$$

$$= 0 \cdot (1-p) + 1 \cdot (p)$$

$$E\{Y_i\} = p \quad (2.10)$$

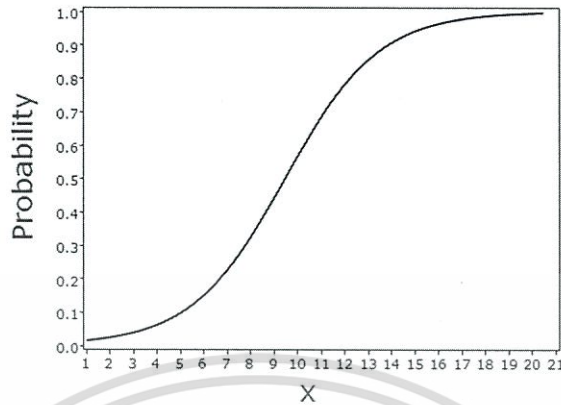
ซึ่งทำให้ $0 \leq E\{Y_i\} \leq 1$

เนื่องจาก Y มีค่าได้เพียง 2 ค่า คือ 0 และ 1 จึงทำให้ความสัมพันธ์ระหว่าง X และ Y ไม่ได้อยู่รูปเชิงเส้น แต่จะอยู่ในรูป

$$E\{Y\} = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (2.11)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยเรียกสมการที่ (2.11) ว่าฟังก์ชันการตอบสนองโลจิสติก (Logistic Response Function) ซึ่งกราฟของฟังก์ชันแสดงได้ดังรูปที่ 2.7 โดยที่ $0 \leq E\{Y\} \leq 1$



รูปที่ 2.7 ฟังก์ชันการตอบสนองโลจิสติก (Logistic Response Function)

จากสมการที่ (2.10) และ (2.11) จะได้

$$E\{Y\} = p = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (2.12)$$

โดยที่ $P\{Y = 1\} = p$

$$P\{Y = 0\} = 1 - p$$

ดังนั้นสมการที่ (2.12) จะได้ว่า ความน่าจะเป็นของการเกิดเหตุการณ์ คือ

$$P\{Y = 1\} = p = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (2.13)$$

เป้าหมายของการวิเคราะห์การถดถอยโลจิสติกอย่างง่าย คือการประมาณค่า β_0 และ β_1 เนื่องจาก Y_i มีการแจกแจงแบบเบอร์นูลลีซึ่งมีการแจกแจงดังแสดงในสมการที่ (2.8) และ (2.9) ในกรณีที่ใช้ข้อมูลตัวอย่าง และข้อมูลตัวอย่าง n หน่วยเป็นอิสระกัน ฟังก์ชันความควรจะเป็น (Likelihood function) (กัลยา วานิชปัญญา, 2548)

$$L = \prod_{i=1}^n p^{Y_i} (1 - p)^{1 - Y_i}$$

$$L = \prod_{i=1}^n \left\{ \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \right\}^{Y_i} \left\{ 1 - \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \right\}^{1 - Y_i} \quad (2.14)$$

$$P\{Y = 0\} = 1 - p = 1 - \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^w}$$

โดยที่ $w = \beta_0 + \beta_1 x$

สมการที่ (2.14) จะเป็น

$$L = \prod_{i=1}^n \left\{ \frac{e^w}{1+e^w} \right\}^{Y_i} \left\{ 1 - \frac{e^w}{1+e^w} \right\}^{1-Y_i} \quad (2.15)$$

หาค่า log ของสมการที่ (2.14) จะได้ $\log_e(L)$ ซึ่งเรียกว่า log-likelihood function

$$\begin{aligned} \log_e(L) &= \ln L = \sum_{i=1}^n (Y_i \ln[P(Y_i)] + (1 - Y_i) \ln[1 - P(Y_i)]) \\ &= \sum_{i=1}^n (Y_i (\ln[P(Y_i)] - \ln[1 - P(Y_i)]) + \ln[1 - P(Y_i)]) \\ &= \sum_{i=1}^n \left(Y_i \ln \left(\frac{P(Y_i)}{1 - P(Y_i)} \right) + \ln[1 - P(Y_i)] \right) \\ &= \sum_{i=1}^n \left(Y_i \ln \left(\frac{e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}}{1 + e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}} \right) + \ln \left(1 - \frac{e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}}{1 + e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}} \right) \right) \\ &= \sum_{i=1}^n (Y_i \ln(e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}) + \ln 1 - \ln(1 + e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}})) \\ \ln L &= \sum_{i=1}^n (Y_i (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}) - \ln(1 + e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}})) \end{aligned} \quad (2.16)$$

เนื่องจากสมการที่ (2.16) เป็นสมการไม่เชิงเส้น จึงใช้วิธีนิวตันราฟสัน (Newton - Raphson Method) ซึ่งเป็นวิธีการแก้สมการไม่เชิงเส้น (nonlinear equations) โดยมีกระบวนการทำซ้ำโดยการหาอนุพันธ์เทียบกับค่าพารามิเตอร์ของฟังก์ชันความควรจะเป็นสูงสุด

กำหนดให้ $\ln L = L(B)$ ขั้นตอนการประมาณค่า มีดังนี้

- 1) กำหนดค่าพารามิเตอร์เริ่มต้น $B_0 = [\beta_0 \ \beta_1 \ \dots \ \beta_k]^T$
- 2) หาอนุพันธ์ย่อยของ $\ln L$ เทียบกับ $\beta_0, \beta_1, \dots, \beta_k$ แล้วนำมาสร้างเป็นเวกเตอร์

$$U(B) \text{ โดยที่ } U(B) = \left[\frac{\partial L(B)}{\partial \beta_0} \quad \frac{\partial L(B)}{\partial \beta_1} \quad \dots \quad \frac{\partial L(B)}{\partial \beta_k} \right]^T$$

- 3) หาเมทริกซ์ $H(B)$ มิติ $(k+1) \times (k+1)$ โดยที่สมาชิกของเมทริกซ์ได้จากการหาอนุพันธ์ย่อยอันดับที่สอง (Second partial derivative) ของ $L(B)$ โดยที่สมาชิกตัวที่ (i, j) คือ $\frac{\partial^2 L(B)}{\partial \beta_i \partial \beta_j}$; $i, j = 0, 1, 2, \dots, k$ เรียกเมทริกซ์ $H(B)$ ว่า Hessian

matrix

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 4) คำนวณค่าประมาณของพารามิเตอร์ (B) จาก $\hat{B} = B^* - H^{-1}(B^*)U(B^*)$ ซึ่งเวกเตอร์ $U(\hat{B})$ เป็นเวกเตอร์ Score ของ B ที่ประมาณด้วยความควรจะเป็นสูงสุดโดยใช้ Taylor Series กระจาย $U(B)$ รอบ B^* ซึ่ง B^* อยู่ใกล้ๆ \hat{B} จะได้
- $$U(\hat{B}) \approx U(B^*) + H(B^*)(\hat{B} - B^*)$$

จากนิยามของตัวประมาณความควรจะเป็นสูงสุดของ B จะได้ว่า $\frac{\partial L(B)}{\partial \beta_j}$

กำหนดให้ $\hat{\beta} = 0$ สำหรับ $j = 0, 1, 2, \dots, k$ ดังนั้น $U(\hat{B}) = 0$ จะได้

$$0 \approx U(B^*) + H(B^*)(\hat{B} - B^*)$$

$$\hat{B}H(B^*) = -U(B^*) + B^*H(B^*)$$

$$\hat{B} = B^* - H^{-1}(B^*)U(B^*)$$

- 5) ประมาณค่า \hat{B} โดยมีการคำนวณซ้ำๆ ซึ่งค่าประมาณ \hat{B} ณ รอบที่ $k + 1$ คือ

$$\hat{B}_{k+1} = \hat{B}_k - H^{-1}(\hat{B}_k)U(\hat{B}_k)$$

สำหรับ $k = 0, 1, 2, \dots$ ซึ่ง \hat{B}_0 เป็นเวกเตอร์ของตัวประมาณพารามิเตอร์เริ่มต้น โดยค่า \hat{B}_{k+1} ที่ยอมรับเป็นค่าประมาณพารามิเตอร์จะพิจารณาจากผลต่างระหว่าง \hat{B} ในรอบที่ k กับรอบที่ $k + 1$ มีค่าน้อยมากจนถือว่าไม่แตกต่าง ซึ่งในที่นี้ กำหนดเกณฑ์ คือ $|\hat{B}_{k+1} - \hat{B}_k| < 0.0000001$

2.5.5 การปรับรูปแบบความสัมพันธ์ให้อยู่ในรูปเชิงเส้น

จากสมการของการวิเคราะห์การถดถอยโลจิสติกในสมการที่ (2.13) หรืออยู่ในรูป $\ln L$ ในสมการที่ (2.16) จะพบว่าความสัมพันธ์ไม่ได้อยู่ในรูปเชิงเส้น สามารถปรับให้อยู่ในรูปเชิงเส้นได้ดังนี้

กำหนดให้ Odds Ratio (OR) เป็นอัตราส่วนระหว่างโอกาสที่จะเกิดกับโอกาสที่เหตุการณ์จะไม่เกิด จะได้ว่า

$$OR = \frac{P\{Y = 1\}}{P\{Y = 0\}} = \frac{p}{1 - p} = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \cdot \frac{1}{1 + e^{\beta_0 + \beta_1 x}}$$

นั่นคือ $OR = e^{\beta_0 + \beta_1 x}$ (2.17)

ดังนั้น OR เป็นค่าที่แสดงว่าโอกาสที่จะเกิดเหตุการณ์เป็นกี่เท่าของโอกาสที่จะไม่เกิดเหตุการณ์ ค่า OR สามารถแบ่งได้เป็น 3 กรณี ดังนี้

- ถ้าค่า OR มากกว่า 1 แสดงว่า โอกาสที่จะเกิดเหตุการณ์มากกว่าโอกาสที่จะไม่เกิดเหตุการณ์
- ถ้าค่า $OR = 1$ แสดงว่า โอกาสที่จะเกิดเหตุการณ์และโอกาสที่จะไม่เกิดเหตุการณ์เท่ากัน
- ถ้าค่า OR น้อยกว่า 1 แสดงว่า โอกาสที่จะเกิดเหตุการณ์ต่ำกว่าโอกาสที่จะไม่เกิดเหตุการณ์

นำสมการที่ (2.17) มาหาค่า $\log_e(OR)$

$$\ln(OR) = \log_e(e^{\beta_0 + \beta_1 x})$$

$$\ln(OR) = \beta_0 + \beta_1 x \quad (2.18)$$

ซึ่งสมการที่ (2.18) อยู่ในรูปเชิงเส้น โดยเรียกสมการที่ (2.18) ว่า ฟังก์ชันตอบสนองโลจิท (Logit Response function) หรือกรณีที่ใช้ข้อมูลตัวอย่าง สมการที่ (2.18) จะกลายเป็น

$$\ln(OR') = b_0 + b_1 x \quad (2.19)$$

2.5.6 โมเดลการวิเคราะห์การถดถอยโลจิสติกเชิงพหุ (Multiple logistic regression analysis model)

การวิเคราะห์การถดถอยโลจิสติกเชิงพหุเป็นการวิเคราะห์ความถดถอยโลจิสติกที่มีตัวแปรอิสระมากกว่า 1 ตัวแปร ในรูปทั่วไปกำหนดให้มี m ตัว คือ X_1, X_2, \dots, X_m ส่วนตัวแปรตาม Y มีได้ 2 ค่า คือ $Y = 1$ ถ้าเกิดเหตุการณ์ และ $Y = 0$ ถ้าไม่เกิดเหตุการณ์

Y ยังคงมีการแจกแจงแบบเบอร์นูลลีที่มีความน่าจะเป็น p

$$p = E(Y) = P\{Y = 1\} = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}} \quad (2.20)$$

หรือ
$$P\{Y = 1\} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m)}}$$

$$P\{Y = 0\} = 1 - P\{Y = 1\} = 1 - p = \frac{1}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}}$$

$$\text{Odd Ratio} = OR = \frac{p}{1 - p} = e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}$$

ฟังก์ชันตอบสนองโลจิท คือ

$$\ln(OR) = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m \quad (2.21)$$

หรือกรณีที่ใช้ข้อมูลตัวอย่าง สมการที่ (2.21) จะกลายเป็น

$$\ln(OR') = b_0 + b_1 x_1 + \dots + b_m x_m \quad (2.22)$$

2.5.7 การวิเคราะห์การถดถอยโลจิสติกแบบหลายกลุ่ม (Multinomial logistic regression)

ในหัวข้อ 2.5.4 - 2.5.6 ได้กล่าวถึงการวิเคราะห์การถดถอยแบบสองกลุ่ม ซึ่งเป็นการวิเคราะห์เมื่อตัวแปรตามมีค่าได้เพียง 2 ค่าเท่านั้น ในกรณีที่ตัวแปรตามมีค่าตั้งแต่ 3 ค่าขึ้นไป เราจะเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใช้การวิเคราะห์การถดถอยแบบหลายกลุ่ม เช่น ตัวแปรตามแทนอาการเจ็บป่วย ซึ่งแบ่งเป็นสามระดับ คือ รุนแรงมาก ปานกลาง และเล็กน้อย ซึ่งเราจะแทนอาการเจ็บป่วยด้วยตัวเลข 0, 1 และ 2 การวิเคราะห์การถดถอยแบบหลายกลุ่มการขยายแนวความคิดต่างจากเดิมคือ ในการวิเคราะห์การถดถอยแบบสองกลุ่ม เราสนใจโอกาสที่จะเกิดเหตุการณ์ ซึ่งแทนด้วย p และโอกาสที่จะไม่เกิดเหตุการณ์แทนด้วย $1 - p$ แต่การวิเคราะห์การถดถอยแบบหลายกลุ่มจะมีความน่าจะเป็นที่จะเกิดเหตุการณ์ซึ่งแทนด้วย p_1, p_2, \dots, p_k สำหรับการแบ่ง k กลุ่ม จะต้องมีการกำหนดฐานในการเปรียบเทียบ

กำหนดให้ p_j เป็นกลุ่มอ้างอิง (หรือ ฐาน) ในกรณีที่มีตัวแปรอิสระ m ตัวแปรตัวแบบโลจิท (Logit Model) คือ

$$\ln \frac{p_i}{p_j} = \beta_{0j} + \beta_{1j}x_{1i} + \beta_{2j}x_{2i} + \dots + \beta_{mj}x_{mi} \quad (2.23)$$

โดยที่ $i = 1, 2, \dots, m$ และ $j = 1, 2, \dots, k-1$

ในกรณีที่ตัวแปรตามมี 3 กลุ่ม หรือมีค่าเป็นรหัส 3 ค่า และมีตัวแปรอิสระ 3 ตัวแปร

กำหนดให้ p_3 เป็นกลุ่มอ้างอิง ($Y = 3$) ตัวแบบโลจิท (Logit Model) คือ

$$\ln \frac{p_1}{p_3} = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3 \quad (2.24)$$

$$\ln \frac{p_2}{p_3} = \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3 \quad (2.25)$$

$$p_1 + p_2 + p_3 = 1 \quad (2.26)$$

พิจารณาสมการ (2.24)

$$\ln p_1 - \ln p_3 = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3 \quad (2.27)$$

พิจารณาสมการ (2.25)

$$\ln p_2 - \ln p_3 = \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3 \quad (2.28)$$

(2.27) - (2.28) จะได้

$$\ln p_1 - \ln p_3 = (\beta_{01} - \beta_{02}) + (\beta_{11} - \beta_{12})x_1 + (\beta_{21} - \beta_{22})x_2 + (\beta_{31} - \beta_{32})x_3$$

$$\text{หรือ } \ln \frac{p_1}{p_3} = (\beta_{01} - \beta_{02}) + (\beta_{11} - \beta_{12})x_1 + (\beta_{21} - \beta_{22})x_2 + (\beta_{31} - \beta_{32})x_3 \quad (2.29)$$

นำ -1 คูณตลอดสมการ (2.25) จะได้

$$\ln \frac{p_3}{p_2} = -\beta_{02} - \beta_{12}x_1 - \beta_{22}x_2 - \beta_{32}x_3 \quad (2.30)$$

ถ้าเราประมาณค่าตัวแปรไม่ทราบค่าในสมการ (2.29) และ (2.30) จะได้ค่าประมาณของ $\beta_{01}, \beta_{02}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \beta_{31}$ และ β_{32} และเมื่อแทนค่าลงในสมการ (2.29) และ (2.30) จะได้ $\ln \frac{p_1}{p_3}$ และ $\ln \frac{p_2}{p_3}$

จากสมการ (2.24) สมมติว่าประมาณค่า β_{31} ได้ 0.3 ซึ่งก็คือ 0.3/1 หมายความว่า เมื่อ x_3 มีค่าเพิ่มขึ้น 1 จะมีผลให้ $\ln \frac{p_1}{p_3}$ เพิ่มขึ้น 0.3 ซึ่งสรุปได้ยากว่า p_1 และ p_3 มีแนวโน้มอย่างไร

ดังนั้น เราจึงเปลี่ยนสมการ (2.24) – (2.26) ให้อยู่ในรูปความน่าจะเป็น p_1, p_2 และ p_3 ซึ่งทำให้ง่ายต่อการแปลผล

จากสมการ (2.24)

$$p_1 = p_3 * \exp(\beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3) \quad (2.31)$$

$$p_2 = p_3 * \exp(\beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3) \quad (2.32)$$

$$p_3 = p_3 \quad (2.33)$$

นำสมการ (2.31) + (2.32) + (2.33) จะได้

$$1 = p_3 \sum_{i=1}^2 \exp(\beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \beta_{3i}x_3) + p_3$$

$$1 = p_3 \left(\sum_{i=1}^2 \exp(\beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \beta_{3i}x_3) + 1 \right)$$

$$p_3 = \frac{1}{\left(\sum_{i=1}^2 \exp(\beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \beta_{3i}x_3) + 1 \right)} \quad (2.34)$$

เมื่อแทนสมการ (2.34) ลงในสมการ (2.31)

$$p_1 = \frac{\exp(\beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3)}{1 + \left(\sum_{i=1}^2 \exp(\beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \beta_{3i}x_3) + 1 \right)} \quad (2.35)$$

เมื่อแทนสมการ (2.34) ลงในสมการ (2.32)

$$p_2 = \frac{\exp(\beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3)}{1 + \left(\sum_{i=1}^2 \exp(\beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \beta_{3i}x_3) + 1 \right)} \quad (2.36)$$

สมการ (2.34), (2.35) และ (2.36) คือสมการที่ใช้คาดหมายค่าความน่าจะเป็นที่จะเกิด $Y = 1, Y = 2$ และ $Y = 3$ ตามลำดับ

สำหรับกรณีทั่วไป เมื่อตัวแปรตามแบ่งเป็น k กลุ่ม เมื่อ p_j เป็นกลุ่มอ้างอิง

$$p_i = \frac{\exp(\beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \dots + \beta_{mi}x_m)}{\left(\sum_{n=1}^{k-1} \exp(\beta_{0n} + \beta_{1n}x_1 + \beta_{2n}x_2 + \beta_{mn}x_m) + 1\right)} \quad (2.37)$$

เมื่อ $i = 1, 2, \dots, k-1$

$$p_j = \frac{1}{\left(\sum_{n=1}^{k-1} \exp(\beta_{0n} + \beta_{1n}x_1 + \beta_{2n}x_2 + \beta_{mn}x_m) + 1\right)} \quad (2.38)$$

2.5.8 การตรวจสอบความเหมาะสมของสมการการถดถอยโลจิสติก

การตรวจสอบความเหมาะสมหรือความคล่องจง (Goodness of fit test) ของสมการการวิเคราะห์การถดถอยโลจิสติกดังแสดงในสมการที่ (2.18) กรณีที่มีตัวแปรอิสระ 1 ตัว หรือสมการที่ (2.21) ซึ่งมีตัวแปรอิสระ m ตัว

การใช้สถิติทดสอบไคกำลังสอง

การตรวจสอบความเหมาะสมหรือความคล่องจง (goodness of fit) ของรูปแบบหรือสมการการวิเคราะห์ความถดถอยโลจิสติก จะใช้สถิติทดสอบไคกำลังสองของเพียร์สัน (Pearson Chi-Square) หรือสถิติไคกำลังสองความควรจะเป็น (Likelihood Ratio Chi-Square)

$$H_0 : \text{mod el} : P\{Y = 1\} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \text{ เหมาะสมกับข้อมูล}$$

$$H_1 : \text{mod el} : P\{Y = 1\} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \text{ ไม่เหมาะสมกับข้อมูล}$$

โดยที่สถิติไคกำลังสองของเพียร์สัน (χ^2)

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

หรือ สถิติไคกำลังสองความควรจะเป็น (G^2)

$$G^2 = 2 \sum_i \sum_j O_{ij} \log \left(\frac{O_{ij}}{E_{ij}} \right)$$

โดยที่ O_{ij} = ความถี่หรือจำนวนใน cell (i, j) ที่ได้จากข้อมูล

E_{ij} = ความถี่หรือจำนวนใน cell (i, j) ที่คาดว่าจะเกิดเมื่อสมมติฐาน H_0 จริง

การใช้ฟังก์ชันความควรจะเป็น

จากสมการที่ (2.16) ซึ่งเป็นสมการ ln ของฟังก์ชันความควรจะเป็น หรือ Log likelihood function (LL) กรณีที่มีตัวแปรอิสระ 1 ตัว จะได้

$$\ln L = LL = \sum_{i=1}^n y_i \left\{ \frac{e^w}{1 + e^w} \right\} + \sum_{i=1}^n (1 - y_i) \left\{ \frac{1}{1 + e^w} \right\} \quad (2.39)$$

โดยที่ $w = \beta_0 + \beta_1 x$ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือกรณีที่มีตัวแปรอิสระ m ตัว

$$\ln L = LL = \sum_{i=1}^n y_i \left\{ \frac{e^z}{1 + e^z} \right\} + \sum_{i=1}^n (1 - y_i) \left\{ \frac{1}{1 + e^z} \right\} \quad (2.40)$$

โดยที่ $z = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m$

เนื่องจากฟังก์ชันความควรจะเป็น L มีค่าระหว่าง 0 ถึง 1 ดังนั้นจึงศึกษาค่า $-2LL$ หรือ $-2 \ln L$ โดยจะใช้ค่า $-2LL$ วัดความเหมาะสมของสมการโลจิสติก ถ้าสมการโลจิสติกเหมาะสม ค่า $-2LL$ จะต่ำ นั่นคือ ถ้าสมการโลจิสติกเหมาะสมที่สุด ค่า $-2LL$ จะเป็นศูนย์

ค่าที่ใช้ประเมินวิธีการพยากรณ์

เนื่องจากการพยากรณ์นั้นมีหลายวิธี โดยแต่ละวิธีมีตัวแบบหรือแบบจำลองทางคณิตศาสตร์ เพื่อพยากรณ์ที่แตกต่างกันออกไป ในการตัดสินใจเลือกการพยากรณ์ที่เหมาะสม และถูกต้องนั้นทำได้โดยเลือกใช้การประเมินวิธีต่าง ได้แก่

- 1) ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Mean Square Error: MSE)
- 2) รากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Root Mean Square Error: RMSE)
- 3) ค่าเฉลี่ยของค่าเบี่ยงเบนสัมบูรณ์ (Mean Absolute Deviation: MAD)
- 4) ค่าเฉลี่ยของความคลาดเคลื่อนร้อยละสัมบูรณ์ (Mean Absolute Percent Error: MAPE)

โดยแต่ละวิธีมีสูตรการคำนวณดังต่อไปนี้

$$MSE = \sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}$$

$$RMSE = \sqrt{MSE}$$

$$MAD = \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{n}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{(Y_i - \hat{Y}_i)}{n} \right|$$

เมื่อ Y_i คือ ค่าที่แท้จริงของค่าสังเกตตัวที่ i ในอนุกรมเวลา

เมื่อ \hat{Y}_i คือ ค่าที่ของตัวแปรพยากรณ์จากค่าสังเกตตัวที่ i

2.6 การประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมสำหรับปัญหาการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอยเป็นการวิเคราะห์ข้อมูลเพื่อการพยากรณ์ที่ใช้กันอย่างกว้างขวางในหลายวงการ เช่น วงการธุรกิจ วงการการศึกษาและวิทยาศาสตร์ เป็นต้น โดยการใช้เทคนิคในการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม โดยตัวแปรอิสระคือตัวแปรที่ผู้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิเคราะห์ข้อมูลเห็นว่าเป็นตัวแปรที่เป็นตัวกำหนดความผันแปรของตัวแปรตาม หรืออาจกล่าวได้ว่าตัวแปรอิสระเป็นตัวแปรที่มีอิทธิพลต่อตัวแปรตามนั่นเอง ส่วนตัวแปรตามคือตัวแปรที่มีค่าเปลี่ยนแปลงไปตามการเปลี่ยนแปลงของตัวแปรอิสระ

การวิเคราะห์การถดถอย เป็นการศึกษาเกี่ยวกับความสัมพันธ์ระหว่างตัวแปร 2 ตัวหรือลักษณะที่สนใจศึกษา 2 ลักษณะ คือ ตัวแปรอิสระ (Independent variable) ซึ่งเป็นตัวแปรที่ไม่ได้ถูกกำหนดโดยตัวแปรอื่นๆที่เกี่ยวข้องในเรื่องที่ศึกษาอยู่ แต่อาจเกิดขึ้นจากปัจจัยอื่นๆที่ไม่เกี่ยวข้อง กับเรื่องที่ศึกษาหรือเป็นข้อมูลพื้นฐานทั่วไป ตัวแปรตาม (Dependent variable) เป็นตัวแปรที่เปลี่ยนแปลงไปตามตัวแปรอิสระ ความสัมพันธ์ระหว่างตัวแปร 2 ตัว อาจมีรูปแบบของความสัมพันธ์ใน 2 แบบ คือความสัมพันธ์ที่เป็นเส้นตรง (Linear) และความสัมพันธ์ที่ไม่เป็นเส้นตรง (Non-linear) แต่ในที่นี้จะกล่าวถึงเฉพาะความสัมพันธ์ที่เป็นเส้นตรง โดยการวิเคราะห์ความสัมพันธ์ที่มีตัวแปรอิสระเพียงตัวแปรเดียวเป็นตัวบ่งบอกอิทธิพลของตัวแปรตาม เรียกว่า การวิเคราะห์การถดถอยอย่างง่าย (Simple regression analysis) ส่วนการวิเคราะห์ความสัมพันธ์ที่มีตัวแปรอิสระหลายตัว X_1, X_2, \dots, X_k เพื่อการอธิบายตัวแปรตามเรียกว่า การวิเคราะห์การถดถอยพหุคูณ (Multiple regression analysis) จุดประสงค์ของการวิเคราะห์ คือ เพื่อวิเคราะห์ความเหมาะสมของตัวแบบและหาอิทธิพลโดยรวมของตัวแปรต้น (X_k) ที่มีต่อตัวแปรตาม และเพื่อสร้างสมการเชิงทำนายและวิเคราะห์หาอิทธิพลของตัวแปรต้น (X_k) ที่มีต่อ Y สำหรับงานวิจัยชิ้นนี้สนใจศึกษาการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

สมการการวิเคราะห์การถดถอยเชิงเส้นพหุคูณสำหรับข้อมูล n ตัวอย่าง คือ

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + e_i \quad (2.41)$$

ซึ่ง $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ เป็นตัวแปรไม่ทราบค่า จึงต้องมีการประมาณค่า โดยต้องใช้ข้อมูลตัวอย่างของตัวแปร Y, X_1, X_2, \dots, X_k แทนค่าลงในสมการ (2.41) ทำให้ประมาณค่า Y หรือประมาณสมการที่ (2.41) ด้วยสมการที่ (2.42)

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki} \quad (2.42)$$

หรือ
$$\hat{Y}_i = a + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki} \quad (2.43)$$

โดยที่
$$\hat{\beta}_0 = a, \hat{\beta}_1 = b_1, \hat{\beta}_2 = b_2, \dots, \hat{\beta}_k = b_k$$

ดังนั้นค่าคลาดเคลื่อนในการประมาณค่า Y_i ด้วย \hat{Y}_i คือ $Y_i - \hat{Y}_i = e_i$ (สมการที่(2.43) - (2.41))

โดยทั่วไปการประมาณค่า $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ ด้วยค่า a, b_1, b_2, \dots, b_k ตามลำดับนั้นยังคงมีเป้าหมาย คือ เพื่อให้ผลบวกของค่าคลาดเคลื่อนยกกำลังสองมีค่าน้อยที่สุด โดยใช้วิธีกำลังสองน้อยที่สุด นั่นคือหาค่า a, b_1, b_2, \dots, b_k ที่ทำให้ $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ มีค่าต่ำที่สุด

ในกรณีที่มีตัวแปรอิสระ 2 ตัว (X_1, X_2) ที่มีความสัมพันธ์กับ Y สมการถดถอย คือ

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i \quad (2.44)$$

ค่าประมาณของ Y_i คือ

$$\hat{Y}_i = a + b_1 X_{1i} + b_2 X_{2i} \quad (2.45)$$

และความคลาดเคลื่อน $e_i = Y_i - \hat{Y}_i$

ต้องการ $\min \sum_{i=1}^n e_i^2 = \min \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ จึงใช้อนุพันธ์เชิงส่วน (Partial Derivative) เทียบกับ a, b_1 และ b_2 แล้วให้เท่ากับศูนย์ดังนี้

$$\begin{aligned} \frac{\partial}{\partial a} \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] &= \frac{\partial}{\partial a} \left[\sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i})^2 \right] \\ &= -2 \sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i}) = 0 \end{aligned}$$

$$\text{หรือ} \quad -2 \sum_{i=1}^n Y_i + 2na + 2b_1 \sum_{i=1}^n X_{1i} + 2b_2 \sum_{i=1}^n X_{2i} = 0$$

$$na + b_1 \sum_{i=1}^n X_{1i} + b_2 \sum_{i=1}^n X_{2i} = \sum_{i=1}^n Y_i \quad (2.46)$$

$$\begin{aligned} \frac{\partial}{\partial b_1} \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] &= \frac{\partial}{\partial b_1} \left[\sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i})^2 \right] \\ &= -2 \sum_{i=1}^n X_{1i} (Y_i - a - b_1 X_{1i} - b_2 X_{2i}) = 0 \end{aligned}$$

$$\text{หรือ} \quad -2 \sum_{i=1}^n X_{1i} Y_i + 2a \sum_{i=1}^n X_{1i} + 2b_1 \sum_{i=1}^n X_{1i}^2 + 2b_2 \sum_{i=1}^n X_{1i} X_{2i} = 0$$

$$a \sum_{i=1}^n X_{1i} + b_1 \sum_{i=1}^n X_{1i}^2 + b_2 \sum_{i=1}^n X_{1i} X_{2i} = \sum_{i=1}^n X_{1i} Y_i \quad (2.47)$$

ในทำนองเดียวกัน $\frac{\partial}{\partial b_2} \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] = 0$ จะได้

$$-2 \sum_{i=1}^n X_{2i} Y_i + 2a \sum_{i=1}^n X_{2i} + 2b_1 \sum_{i=1}^n X_{1i} X_{2i} + 2b_2 \sum_{i=1}^n X_{2i}^2 = 0$$

$$a \sum_{i=1}^n X_{2i} + b_1 \sum_{i=1}^n X_{1i} X_{2i} + b_2 \sum_{i=1}^n X_{2i}^2 = \sum_{i=1}^n X_{2i} Y_i \quad (2.48)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้ไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมการ (6), (7) และ (8) เรียกว่าชุดของสมการปกติ ดังนี้

$$\left. \begin{aligned} na + b_1 \sum_{i=1}^n X_{1i} + b_2 \sum_{i=1}^n X_{2i} &= \sum_{i=1}^n Y_i \\ a \sum_{i=1}^n X_{1i} + b_1 \sum_{i=1}^n X_{1i}^2 + b_2 \sum_{i=1}^n X_{1i} X_{2i} &= \sum_{i=1}^n X_{1i} Y_i \\ a \sum_{i=1}^n X_{2i} + b_1 \sum_{i=1}^n X_{1i} X_{2i} + b_2 \sum_{i=1}^n X_{2i}^2 &= \sum_{i=1}^n X_{2i} Y_i \end{aligned} \right\} \quad (2.49)$$

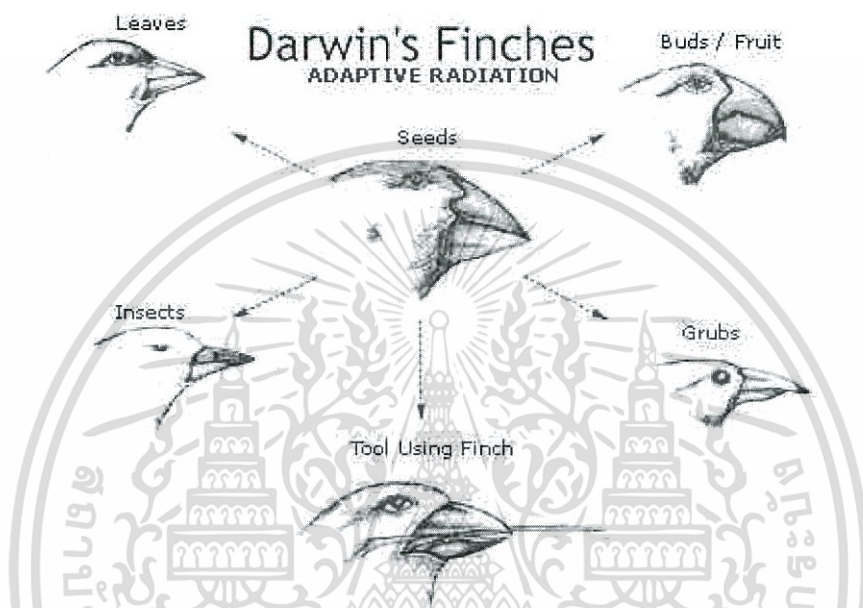
จากงานวิจัยของสิริกัลยา ประมวล แสดงให้เห็นว่า ขั้นตอนวิธีเชิงพันธุกรรมให้ประสิทธิภาพในการหาค่าของ a, b_1, b_2, \dots, b_k ได้ดีกว่าวิธีกำลังสองน้อยที่สุด ดังนั้นในงานวิจัยชิ้นนี้สนใจหาค่าของ a, b_1, b_2, \dots, b_k โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm – GA) เป็นเทคนิคทางปัญญาประดิษฐ์ (Artificial Intelligence) อย่างหนึ่งที่ใช้ในการค้นหา การเพิ่มประสิทธิภาพ และการเรียนรู้ (Search Optimization and Learning) ด้วยการเลียนแบบทฤษฎีการวิวัฒนาการทางธรรมชาติ โดยขั้นตอนวิธีเชิงพันธุกรรมมีจุดเด่นในด้านความทนทานต่อความผิดพลาดในการค้นหาคำตอบจากแหล่งข้อมูลที่มีความซับซ้อนและยากที่จะสร้างแบบจำลองด้วยสมการคณิตศาสตร์ เนื่องจากเป็นกระบวนการค้นหาที่ไม่มีความเฉพาะเจาะจงกับแบบจำลองหรือลักษณะเฉพาะของข้อมูลแบบใดแบบหนึ่ง ด้วยเหตุนี้ขั้นตอนวิธีเชิงพันธุกรรมจึงถูกนำมาประยุกต์ใช้ในการแก้ปัญหาได้หลากหลายรูปแบบ ตั้งแต่การจัดตารางเวลา (Timetable Scheduling) การออกแบบระบบควบคุมอัตโนมัติ (Control System Design) การออกแบบเพื่อเพิ่มประสิทธิภาพของระบบท่อส่งก๊าซ (Gas Pipeline Optimization) และการพัฒนาระบบปัญญาประดิษฐ์ที่สามารถเรียนรู้จากสภาพแวดล้อมได้ (Genetic Based Machine Learning) เป็นต้น โดยหลักการของขั้นตอนวิธีเชิงพันธุกรรมเป็นการเลียนแบบกระบวนการวิวัฒนาการตามธรรมชาติ เพื่อพัฒนาหรือทำการ “วิวัฒนาการ” คำตอบที่ดีที่สุดในการแก้ปัญหา (<http://kapitaennem0.wordpress.com/2013/07/17/genetic-algorithm>)

2.6.1 ทฤษฎีวิวัฒนาการของชาลส์ ดาร์วิน

ทฤษฎีของชาลส์ ดาร์วิน (Charles Darwin) ได้เสนอทฤษฎีวิวัฒนาการตามธรรมชาติ กล่าวถึงความหลากหลายตามธรรมชาติและ ความแตกต่างของสายพันธุ์สิ่งมีชีวิต ว่าสายพันธุ์ของสิ่งมีชีวิตต่างๆ มีที่มาร่วมกัน แต่อิทธิพลจากสภาพแวดล้อมและการกลายพันธุ์ตามธรรมชาติอย่างต่อเนื่อง ทำให้เกิดความหลากหลายของสายพันธุ์ต่างๆ ขึ้นในที่สุด ซึ่งสายพันธุ์ที่มีความเหมาะสมกับสภาพแวดล้อมจะมีโอกาสอยู่รอดได้มากและมีโอกาสถ่ายทอดลักษณะทางพันธุกรรมไปสู่ลูกหลานรุ่นต่อไป ในขณะที่สายพันธุ์ที่ไม่เหมาะสมกับสภาพแวดล้อมจะมีโอกาสอยู่รอดได้น้อย และมีโอกาสสูญพันธุ์ไปในที่สุด โดยกระบวนการคัดเลือกสายพันธุ์ที่เหมาะสมนี้เรียกว่า กระบวนการคัดเลือกทางธรรมชาติ (Natural Selection) ซึ่งการเปลี่ยนแปลงของสภาพแวดล้อมย่อมเอกลักษณะให้เกิดการเปลี่ยนแปลงลักษณะและจำนวนประชากรของสายพันธุ์ที่เหมาะสมไปด้วย รวมถึงการคัดเลือกสายพันธุ์ที่ไม่เหมาะสมให้ลดจำนวนลงหรือสูญพันธุ์ไปเลย

สายพันธุ์ที่ไม่สามารถปรับตัวให้เข้ากับสภาพแวดล้อมได้ทันอาจสูญพันธุ์ไป ตัวอย่างเช่น การศึกษานกที่อาศัยอยู่ในแหล่งอาหารที่แตกต่างกันจะพบว่าการวิวัฒนาการลักษณะของปาก เพื่อให้เหมาะสมกับลักษณะของอาหารตามสภาพแวดล้อมของนกสายพันธุ์นั้น ดังรูปที่ 2.8 กล่าวโดยสรุปได้ว่าทฤษฎีวิวัฒนาการของชาลส์ ดาร์วิน คือการเปลี่ยนแปลงอย่างต่อเนื่องของสายพันธุ์และลักษณะทางพันธุกรรมเพื่อให้เหมาะสมกับการเปลี่ยนแปลงสภาพแวดล้อม (<http://kapitaennem0.wordpress.com/2013/07/17/genetic-algorithm>)

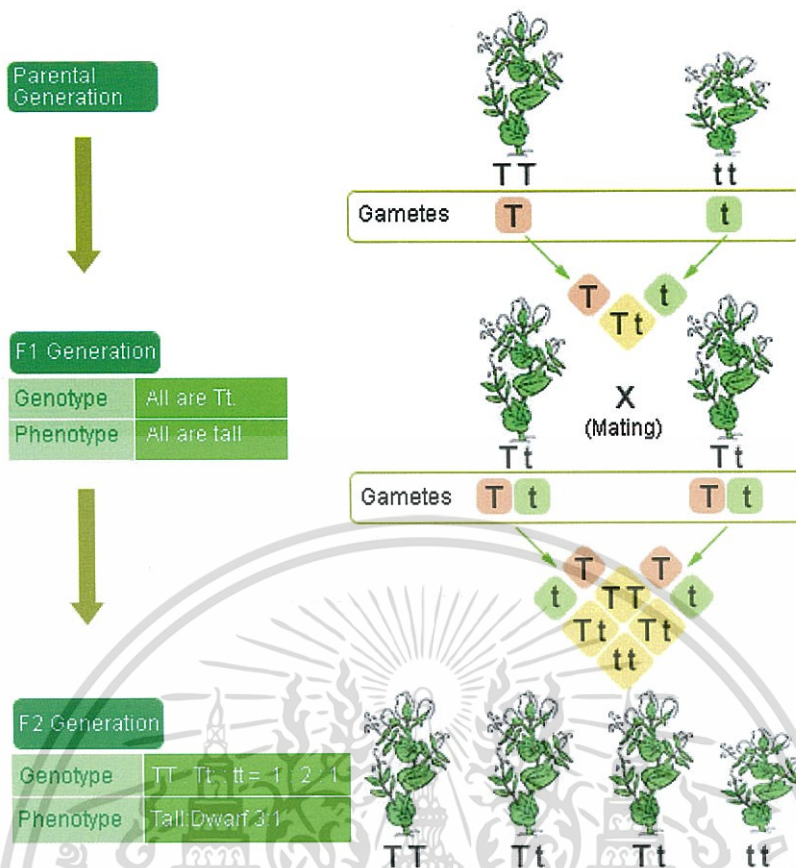


รูปที่ 2.8 วิวัฒนาการของปากนกตามลักษณะของอาหารตามทฤษฎีของชาลส์ ดาร์วิน
(ที่มา: <http://kapitaennem0.wordpress.com/2013/07/17/genetic-algorithm>)

2.6.2 ทฤษฎีการถ่ายทอดลักษณะทางพันธุกรรมของเกรเกอร์ เมนเดล

ทฤษฎีของดาร์วิน ได้กล่าวถึงการถ่ายทอดลักษณะทางพันธุกรรม แต่ยังไม่สามารถอธิบายกลไกการถ่ายทอดได้อย่างชัดเจน เกรเกอร์ เมนเดล (Gregor Mendel) ได้เสนอทฤษฎีการถ่ายทอดลักษณะทางพันธุกรรม จากการศึกษาความแตกต่างในสายพันธุ์ของพืช การเกิดลักษณะพันธุ์ทาง (Hybrids) จากการผสมพันธุ์พืชสายพันธุ์เดียวกันและการถ่ายทอดลักษณะทางพันธุกรรมในการเพาะต้นถั่ว จากการศึกษาพบว่าลักษณะทางพันธุกรรมที่ถ่ายทอดไปยังรุ่นลูกหลาน เช่น รูปร่างและสีของดอกไม้ สามารถแบ่งออกได้เป็น 2 ลักษณะ ได้แก่ ลักษณะเด่น (Dominant Characters) กับลักษณะด้อย (Recessive Characters) ซึ่งการเกิดลักษณะทางพันธุกรรมในรุ่นลูกหลานเป็นผลมาจากการถ่ายทอดลักษณะทางพันธุกรรมของรุ่นพ่อแม่ผ่านกระบวนการสืบพันธุ์ โดยมีอัตราการเกิดลักษณะต่างๆ ที่คงที่และสามารถคำนวณโอกาสในการเกิดลักษณะนั้นล่วงหน้าได้ ตัวอย่างเช่น การผสมพันธุ์พืชสายพันธุ์เดียวกันที่มีลักษณะสูงกับลักษณะแคระจะให้โอกาสในการเกิดต้นแคระ 1

ใน 4 ในลูกหลานรุ่นที่ 2 ดังรูปที่ 2.9 ใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 การถ่ายทอดลักษณะทางพันธุกรรมตามทฤษฎีของเกรเกอร์ เมนเดล
(ที่มา: <http://kapitaenem0.wordpress.com/2013/07/17/genetic-algorithm>)

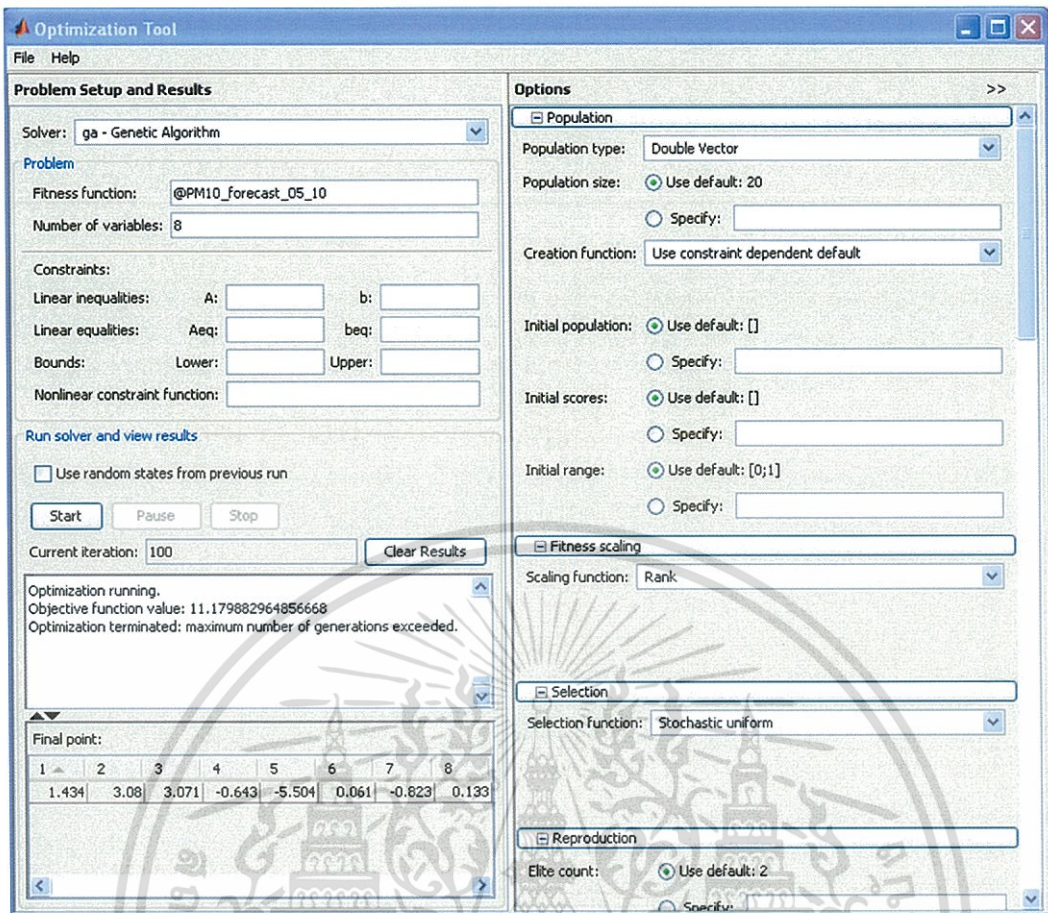
ทฤษฎีการถ่ายทอดลักษณะทางพันธุกรรมของเกรเกอร์ เมนเดล เป็นที่มาของการค้นพบหน่วยพันธุกรรม หรือ ยีน (Gene) ซึ่งเป็นหน่วยย่อยของโครโมโซม (Chromosome) ในการกำหนดลักษณะทางพันธุกรรมของสิ่งมีชีวิต และทำหน้าที่ถ่ายทอดลักษณะทางพันธุกรรมดังกล่าวไปยังรุ่นลูกหลาน ทั้งในส่วนองลักษณะที่มองเห็นได้ภายนอก (Phenotype) และลักษณะทางชีวเคมีภายใน (Genotype) โดยทฤษฎีของเกรเกอร์ เมนเดล สามารถสรุปได้เป็นกฎสำคัญ 2 ข้อ กล่าวคือ กฎการแยกตัว (Law of Segregation) ที่กล่าวว่า โดยปกติแล้วยีนที่กำหนดลักษณะทางพันธุกรรมแต่ละชนิดจะมีลักษณะเป็นคู่ของอัลลีล (Allele) โดยในการสืบพันธุ์คู่ของอัลลีลของพ่อแม่จะถูกแยกออกจากกัน และลูกจะได้รับอัลลีลที่เป็นคู่จากการสุ่มผสมอัลลีลแต่ละข้างของพ่อแม่ กับกฎการจัดเรียงอย่างอิสระ (Law of Independent Assortment) ที่กล่าวว่า ยีนที่กำหนดลักษณะทางพันธุกรรมแต่ละชนิดจะถูกถ่ายทอดไปยังรุ่นลูกโดยไม่มี ความเกี่ยวข้องกับยีนของลักษณะอื่น ตัวอย่างเช่น การถ่ายทอดยีนที่เป็นลักษณะของสีดอกไม้ จะเป็นไปโดยอิสระไม่มีความเกี่ยวข้องกับ การถ่ายทอดยีนที่เป็นลักษณะของรูปร่างดอกไม้

2.6.3 หลักการของขั้นตอนวิธีเชิงพันธุกรรม

ความสำเร็จของกระบวนการวิวัฒนาการในธรรมชาติในการคัดเลือกสิ่งมีชีวิตที่เหมาะสมกับสภาพแวดล้อม การกลายพันธุ์เพื่อเปิดโอกาสในการพัฒนาสายพันธุ์ และการอยู่รอดของเผ่าพันธุ์ ด้วยการปรับตัวให้เข้ากับสิ่งแวดล้อมที่เปลี่ยนแปลงไป รวมถึงการส่งต่อลักษณะทางพันธุกรรมของเผ่าพันธุ์ที่เหมาะสมในลักษณะของการถ่ายทอดและผสมยีนของรุ่นพ่อแม่ไปยังรุ่นลูกหลาน ได้กลายเป็นแรงบันดาลใจของมนุษย์ในการค้นหาคำตอบที่เหมาะสมที่สุดในการแก้ปัญหา ขั้นตอนวิธีเชิงพันธุกรรมเป็นแนวคิดที่ริเริ่มโดยจอห์น ฮอลแลนด์ (John Holland) ในการค้นหาคำตอบในการแก้ปัญหาด้วยการเลียนแบบกระบวนการวิวัฒนาการทางธรรมชาติ

ขั้นตอนวิธีเชิงพันธุกรรม เป็นแนวคิดใหม่ (Novel Approach) ในการแก้ปัญหาและค้นหาคำตอบด้วยการเลียนแบบกระบวนการทางธรรมชาติ ขั้นตอนวิธีเชิงพันธุกรรมใช้การผสมชุดคำตอบที่มีความเหมาะสมและคัดเลือกคำตอบที่ดีกว่าเพื่อสร้างชุดคำตอบที่ดียิ่งขึ้นแทนการเปรียบเทียบคำตอบข้างเคียงเพื่อปรับทิศทาง การค้นหา ทำให้ขั้นตอนวิธีเชิงพันธุกรรมมีจุดแข็งในการที่สามารถใช้ค้นหาคำตอบจากชุดข้อมูลที่ไม่ต่อเนื่องได้ และการใช้ชุดคำตอบจำนวนหนึ่งแทนการเปรียบเทียบระหว่างคำตอบแต่ละคำตอบทำให้การค้นหาด้วยขั้นตอนวิธีเชิงพันธุกรรมสามารถหาคำตอบที่ดีที่สุด (Global Minima / Global Maxima) ได้โดยไม่ติดอยู่กับคำตอบที่ดีเฉพาะที่ (Local Minima / Local Maxima) อย่างไรก็ตาม ขั้นตอนวิธีเชิงพันธุกรรมเป็นวิธีที่ใช้การคำนวณค่อนข้างมากและใช้เวลานาน แต่ความก้าวหน้าของเทคโนโลยีคอมพิวเตอร์ในปัจจุบันทำให้สามารถลดเวลาในการดำเนินการตามกระบวนการของขั้นตอนวิธีเชิงพันธุกรรมได้เป็นอย่างมาก จนเป็นที่สนใจของนักวิจัยในปัจจุบัน

กระบวนการของขั้นตอนวิธีเชิงพันธุกรรม เป็นการเลียนแบบกระบวนการวิวัฒนาการและการถ่ายทอดลักษณะทางพันธุกรรมตามธรรมชาติ โดยเริ่มต้นจากการกำหนดปัญหาในรูปของยีนและโครโมโซม และการกำหนดฟังก์ชันความเหมาะสม (Fitness Function) เพื่อใช้เป็นพื้นฐานในกระบวนการวิวัฒนาการชุดคำตอบ จากนั้นจะกำหนดชุดคำตอบชุดแรก (Initial Generation) ในรูปของโครโมโซมด้วยการสุ่ม และนำชุดคำตอบนั้นเข้าสู่กระบวนการวิวัฒนาการ ซึ่งเป็นกระบวนการต่อเนื่องที่ประกอบด้วยตัวดำเนินการ (Operator) ได้แก่ การสืบพันธุ์ (Reproduction) การผสมยีน (Crossover) กับการกลายพันธุ์ (Mutation) และนำไปประเมินความเหมาะสมด้วยฟังก์ชันความเหมาะสม (Fitness Function) สำหรับเงื่อนไขในการหยุดทำงานของขั้นตอนวิธีเชิงพันธุกรรม (Stop Condition) สามารถกำหนดได้หลากหลายรูปแบบ เช่น ครบรอบการทำงานที่ได้กำหนดไว้ พบเป้าหมายหรือคำตอบที่ต้องการ หรือพบว่าคำตอบที่ได้เริ่มลู่เข้าสู่คำตอบที่ดีที่สุด เช่น คำตอบที่ได้จากประชากรแต่ละรุ่นไม่มีการเปลี่ยนแปลงหรือคงที่เป็นจำนวนที่ติดต่อกัน



รูปที่ 2.10 หน้าจอขั้นตอนวิธีเชิงพันธุกรรมโดยโปรแกรมสำเร็จรูป

จะเห็นได้ว่าขั้นตอนวิธีเชิงพันธุกรรม อาศัยการหาคำตอบแบบสุ่มจากพื้นที่ในการหาคำตอบ และหวังว่าคำตอบและโครโมโซมที่เลือกมานั้นจะถูกนำไปสู่คำตอบที่ดีที่สุด ดังนั้นขั้นตอนวิธีเชิงพันธุกรรมจึงมีขั้นตอนวิธีค้นหาคำตอบทางคอมพิวเตอร์ที่ไม่ยุ่งยากซับซ้อนและสามารถนำไปประยุกต์ใช้ในการหาคำตอบของปัญหาการตัดสินใจต่างๆ ได้โดยง่าย

2.6.4 การตั้งปัญหาและแสดงผลโดยโปรแกรมสำเร็จรูป

งานวิจัยนี้ใช้โปรแกรม MATLAB Version 7.10.0.499 (R 2010a)

1. ฟังก์ชันวัตถุประสงค์ (Fitness function)

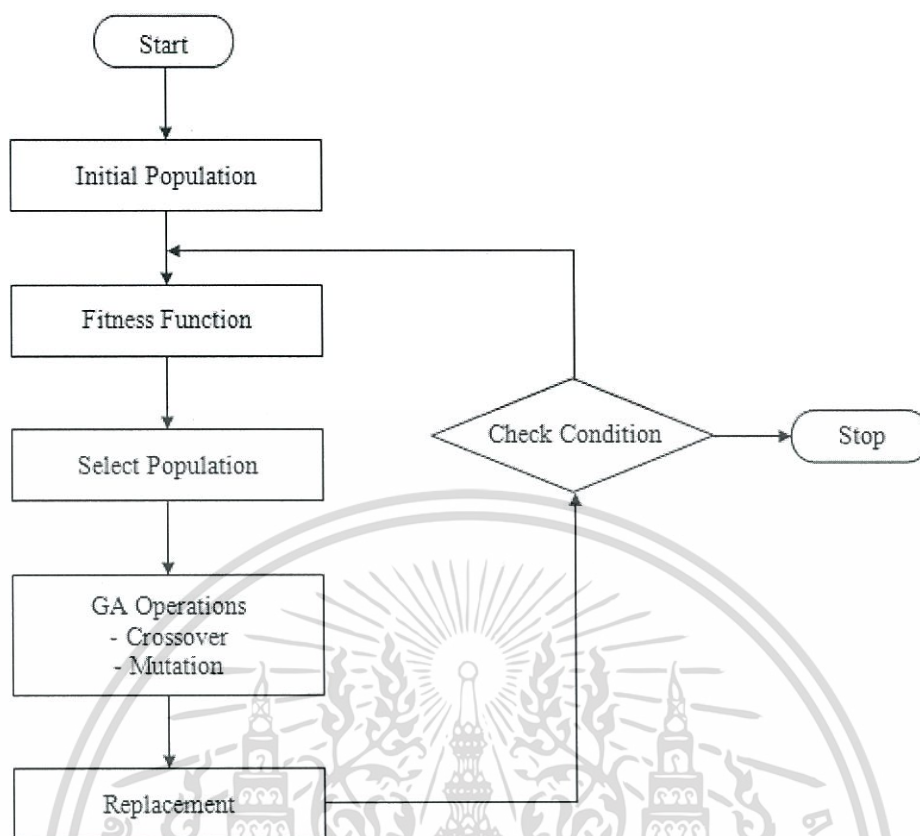
ฟังก์ชันวัตถุประสงค์ที่ต้องการหาค่าที่เหมาะสมที่สุด

2. จำนวนของตัวแปร (Number of variables)

คือจำนวนตัวแปรอิสระที่ใช้ในฟังก์ชันวัตถุประสงค์

3. การประมวลผลและสรุปผล (Run solver and view results)

การเริ่มประมวลผลนั้นให้กดปุ่มเริ่มต้น (Start) และเมื่อกระบวนการประมวลผลสิ้นสุดลงจะมีหน้าจอสถานะและสรุปผลออกมา และมีค่าที่ดีที่สุดแสดงออกมา



รูปที่ 2.11 แผนผังการทำงานของขั้นตอนวิธีเชิงพันธุกรรม

การประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมในงานวิจัยชิ้นนี้ ฟังก์ชันวัตถุประสงค์ที่ต้องการหาค่าที่เหมาะสมที่สุดกำหนดให้เป็นผลบวกของค่าตลาดเคลื่อนยอกำลังสองมีค่าน้อยที่สุด เพื่อให้ได้มาซึ่งสัมประสิทธิ์การถดถอยพหุคูณดังสมการที่ (2.43)

2.6 งานวิจัยที่เกี่ยวข้อง

มงคล ราชะนาคร และคณะ ได้กล่าวว่าคุณภาพอากาศมีความสำคัญมากในการดำรงชีวิต เนื่องจากมีผลกระทบต่อคุณภาพชีวิตและสุขภาพอนามัยของประชากร การวิจัยในเมืองต่างๆทั่วโลกแสดงให้เห็นว่าระดับของฝุ่นละอองในอากาศมีความสัมพันธ์ต่ออัตราการตายก่อนวัยอันควรและการเจ็บป่วยด้วยโรคที่เกี่ยวข้องกับระบบทางเดินหายใจ เช่น โรคภูมิแพ้ โรคหอบหืด โดยเฉพาะผู้ที่มีภูมิคุ้มกันต่ำ เช่น เด็กเล็กและผู้สูงอายุ จังหวัดเชียงใหม่และลำพูนเป็นพื้นที่ที่มีลักษณะเป็นแอ่งกระทะ ประกอบด้วยพื้นที่ที่มีภูเขาล้อมรอบ มีสภาพอากาศที่แห้งและนิ่งเป็นเวลานาน ก่อให้เกิดการสะสมของมลพิษที่ไม่สามารถแพร่กระจายออกสู่ภายนอกได้ ส่งผลให้มีการตรวจพบปริมาณฝุ่นขนาดเล็กสูงเกินมาตรฐาน นอกจากสาเหตุทางด้านภูมิศาสตร์แล้ว การคมนาคมขนส่งที่เพิ่มมากขึ้น การก่อสร้าง และการเผาในที่โล่ง ล้วนแต่เป็นสาเหตุของมลพิษดังกล่าว เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจุบันยังไม่มีแนวโน้มที่จะลดลงแต่อย่างใด ผู้คนในพื้นที่ที่ตกอยู่ในสภาวะเสี่ยงต่อการเกิดโรคที่เกี่ยวข้องกับระบบทางเดินหายใจ ปริมาณฝุ่น PM_{10} รวมถึง $PM_{2.5}$ มีปริมาณสูง ก่อให้เกิดอันตรายต่อสุขภาพ เนื่องจากอนุภาคฝุ่นดังกล่าวสามารถเข้าสู่ปอดได้ อันตรายที่เกิดจาก PM_{10} คือ ความเป็นพิษ (Toxicity) การก่อให้เกิดมะเร็ง (Carcinogenicity) และความเป็นกรด (Acidity)

ยุทธ ไกยวรรณ ได้ศึกษาหลักการและการใช้การวิเคราะห์การถดถอยโลจิสติกสำหรับการวิจัย พบว่าการวิเคราะห์การถดถอยโลจิสติกเป็นการวิเคราะห์ที่มีเป้าหมายเพื่อพยากรณ์โอกาสที่จะเกิดเหตุการณ์ที่สนใจ โดยอาศัยสมการโลจิสติกที่สร้างขึ้นจากชุดตัวแปรอิสระที่เป็นตัวแปรที่มีข้อมูลอยู่ในระดับช่วงเป็นอย่างน้อย โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระจะต้องมีความสัมพันธ์กันต่ำ ตัวแปรตามเป็นตัวแปรเชิงกลุ่ม และในการวิเคราะห์จะต้องใช้ขนาดตัวอย่างไม่ต่ำกว่า 30 ตัว เนื่องจากตัวแปรตามเป็นตัวแปรเชิงกลุ่ม ทำให้ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามไม่อยู่ในรูปแบบเชิงเส้น ในการวิเคราะห์จะต้องมีการปรับความสัมพันธ์อยู่ในรูปแบบเชิงเส้น ในรูปของ odds และในการเขียนโมเดลโลจิสติกจะต้องเขียนให้อยู่ในรูป log ของ odds เรียกว่า logit การวิเคราะห์การถดถอยโลจิสติกใช้ในงานที่สนใจพยากรณ์โอกาสที่จะเกิดเหตุการณ์ที่สนใจ โดยจำเป็นต้องพยากรณ์ค่าเชิงปริมาณของตัวแปรตามนั้น เช่น นักวิจัยต้องการศึกษาว่าเวลาที่ใช้ในการทำข้อสอบมีความสัมพันธ์กับผลการสอบของนักศึกษาหรือไม่ ตัวแปรอิสระคือเวลาที่ใช้ในการทำข้อสอบ และตัวแปรตามคือผลการสอบซึ่งแบ่งเป็น 2 กรณี คือ ผ่านและไม่ผ่าน

วนิดา รัตนมณี และ ศุภชัย ปทุมนากุล ได้กล่าวว่าการทำงานในองค์กรต่างๆ ล้วนแล้วแต่ต้องประสบปัญหาในการทำงาน ปัญหาส่วนใหญ่เป็นปัญหาที่ต้องการคำตอบที่ดีที่สุด หรือคำตอบที่ให้ความพึงพอใจสูงสุด (Optimization problem) ตัวอย่างปัญหาเหล่านี้ได้แก่ ปัญหาด้านการวางแผนการผลิต การจัดผังการทำงาน การกำหนดเส้นทางทางการขนย้ายสินค้า หรือการวางแผนการเดินทางเดียว เป็นต้น ซึ่งคำตอบของปัญหาต่างๆ เหล่านี้มีคำตอบที่เป็นไปได้อยู่เป็นจำนวนมาก สำหรับปัญหาที่มีขนาดใหญ่ วิธีการที่จะใช้ในการหาคำตอบที่เหมาะสมที่สุด คือ วิธีการแบบฮิวริสติก ซึ่งเป็นวิธีการที่ใช้หาคำตอบที่ดีที่สุดจากกลุ่มคำตอบที่สุ่มมาได้ (search space) บางส่วน วิธีการหาคำตอบแบบฮิวริสติกมีอยู่หลายวิธี เช่น การค้นหาแบบตาบู่ (Tabu search) ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm) เป็นต้น ขั้นตอนวิธีเชิงพันธุกรรมเป็นวิธีการหนึ่งที่สามารถหาคำตอบแบบฮิวริสติกได้ผลดี และมีกระบวนการทำงานที่ไม่ซับซ้อน เมื่อเปรียบเทียบกับเทคนิคอื่น ขั้นตอนวิธีเชิงพันธุกรรม เป็นวิธีการที่พยายามหาคำตอบที่เป็นไปได้ให้มากที่สุดและมีการพัฒนาคำตอบไปเรื่อยๆ เพื่อการได้มาซึ่งคำตอบที่น่าพึงพอใจที่สุด และใกล้เคียงกับคำตอบที่ดีที่สุด วิธีการเช่นนี้เป็น

หลักการง่ายๆ แต่สามารถใช้ได้ผลดีกับปัญหาที่มีขนาดใหญ่โดยไม่ต้องหาค่าตัวแปรที่เป็นไปได้ทั้งหมดแล้วเลือกคำตอบที่ดีที่สุดมา นอกจากนี้ยังสามารถหลีกเลี่ยงปัญหาที่จะได้รับคำตอบที่ผิดพลาดซึ่งสามารถเกิดขึ้นได้หากใช้การค้นหาด้วยวิธีอื่นๆ

สิริกัลยา ประมวล ได้ทำการศึกษาเปรียบเทียบวิธีการสำหรับการพยากรณ์ปริมาณอินทรีย์วัตถุในดินสามวิธี คือ วิธีการถดถอยเชิงเส้นพหุคูณ วิธี โครงข่ายประสาทเทียมแบบเชื่อมโยงไปข้างหน้า และวิธีการจีเนติกอัลกอริทึม โดยใช้ตัวแปรพยากรณ์ทั้งหมด 17 ตัวแปร ผลที่ได้จากงานวิจัยพบว่า ประสิทธิภาพของการพยากรณ์ที่ได้จากวิธีการจีเนติกอัลกอริทึมดีกว่า การพยากรณ์ที่ได้จากวิธีการถดถอยเชิงเส้นพหุคูณ และวิธี โครงข่ายประสาทเทียมแบบเชื่อมโยงไปข้างหน้า



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงาน

ในบทนี้จะกล่าวถึงข้อมูลที่ใช้ในการวิเคราะห์ ตัวแบบการพยากรณ์อนุกรมเวลาเพื่อพยากรณ์ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน เพื่อศึกษาแนวโน้มในการเปลี่ยนแปลงรายเดือนของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน และตัวแบบการวิเคราะห์การถดถอยโลจิสติก เพื่อสร้างตัวแบบสำหรับทำนายการเกิดหมอกควันรายวัน โดยอาศัยข้อมูลรายวันปัจจัยที่เกี่ยวข้องกับการเกิดหมอกควันในการพิจารณา

3.1 ข้อมูลและการเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้ในการศึกษาครั้งนี้ ได้แก่ ข้อมูลความเข้มข้นรายเดือนของสารมลพิษทางอากาศ 5 ปัจจัย ได้แก่ O_3 , NO_2 , SO_2 , CO และ PM_{10} โดยระบบเครือข่ายการติดตามตรวจสอบคุณภาพอากาศในบรรยากาศในพื้นที่จังหวัดเชียงใหม่ ในสังกัดสำนักจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ ซึ่งเป็นหน่วยงานที่เก็บรวบรวมข้อมูลสารมลพิษต่างๆ ตั้งแต่ 1 มกราคม ค.ศ. 2005 ถึง 31 ธันวาคม ค.ศ. 2010

3.2 ตัวแบบการพยากรณ์อนุกรมเวลา

สำหรับตัวแบบการพยากรณ์อนุกรมเวลาของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน เราพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์เพื่อวัดความสัมพันธ์ของปัจจัยต่างๆ ที่มีต่อ PM_{10} โดยข้อมูลอนุกรมเวลาที่นำมาพิจารณาเป็นข้อมูลค่าเฉลี่ยรายเดือน คำนวณตามสมการ (2.2)

ตารางที่ 3.1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยต่างๆ และ PM_{10} (รายเดือน)

Factor	Correlation	Factor	Correlation
CO	0.74	Temperature	-0.02
NO	0.35	Relative Humidity	-0.80
NO_2	0.87	Pressure	0.49
SO_2	0.40	Wind speed	-0.08
O_3	0.15	Rain	-0.62
Wind direction	-0.10		

จากตารางที่ 3.1 เราสามารถสรุปได้ว่า

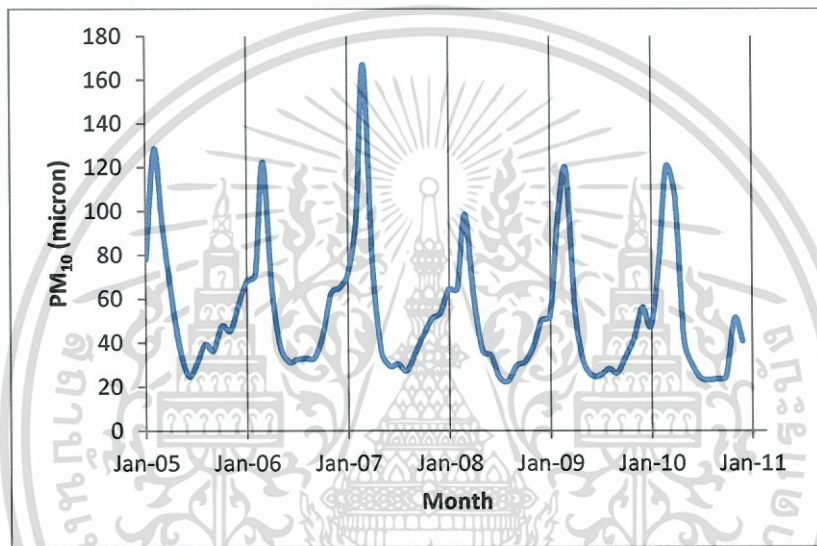
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจัยที่มีความสัมพันธ์กับ PM_{10} สูง ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ และปริมาณน้ำฝน

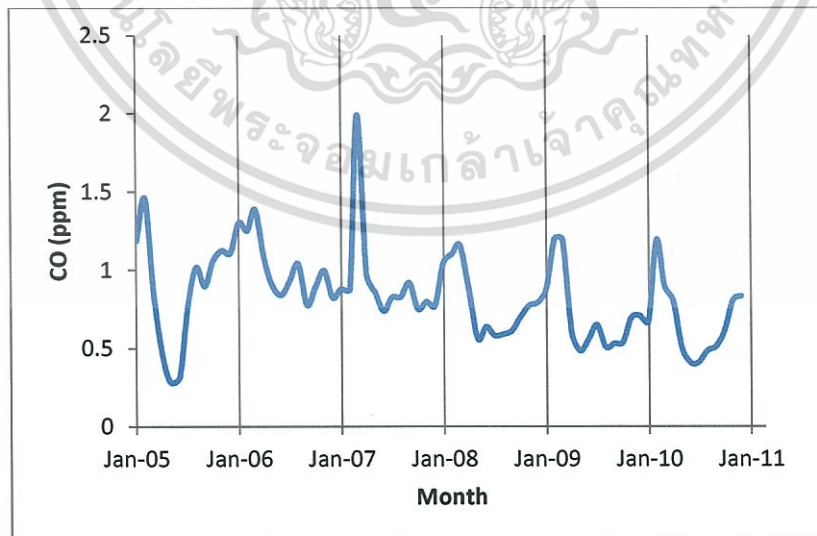
ปัจจัยที่มีความสัมพันธ์กับ PM_{10} ปานกลาง ได้แก่ ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์

ปัจจัยที่มีความสัมพันธ์กับ PM_{10} ต่ำ ได้แก่ ความเร็วลม ทิศทางลม อุณหภูมิ และก๊าซโอโซน

ดังนั้นปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันนั้นขึ้นอยู่กับปัจจัยที่มีความสัมพันธ์กับ PM_{10} สูง ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO_2) ความชื้นสัมพัทธ์ (RH) และปริมาณน้ำฝน (Rain)

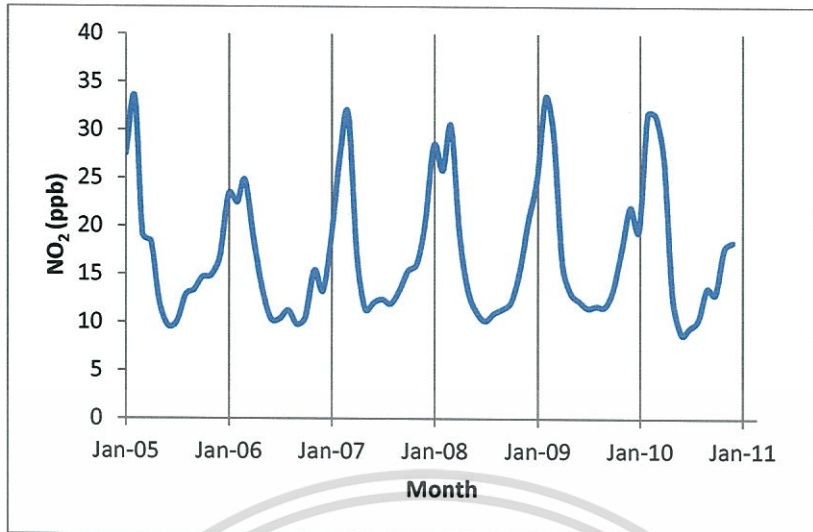


รูปที่ 3.1 ปริมาณฝุ่นละอองขนาดเล็กกว่า 10 ไมครอนรายเดือนตั้งแต่ปี 2005 – 2010



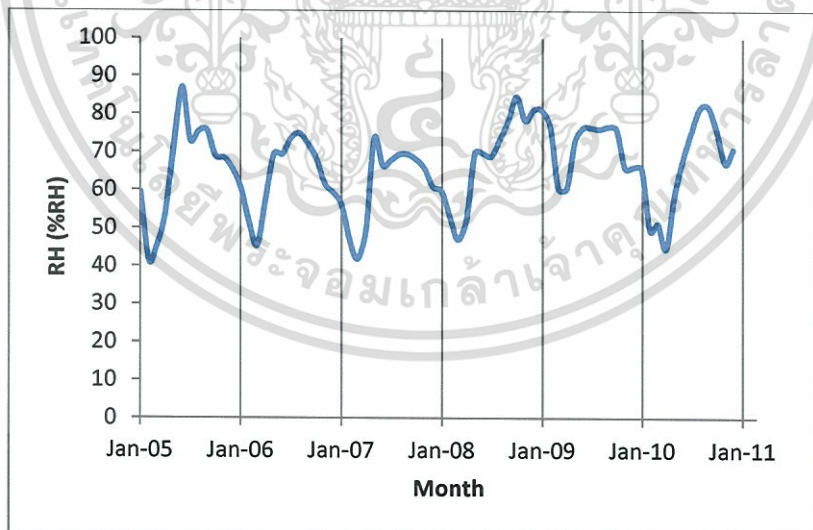
รูปที่ 3.2 ปริมาณก๊าซคาร์บอนมอนอกไซด์รายเดือนตั้งแต่ปี 2005 – 2010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



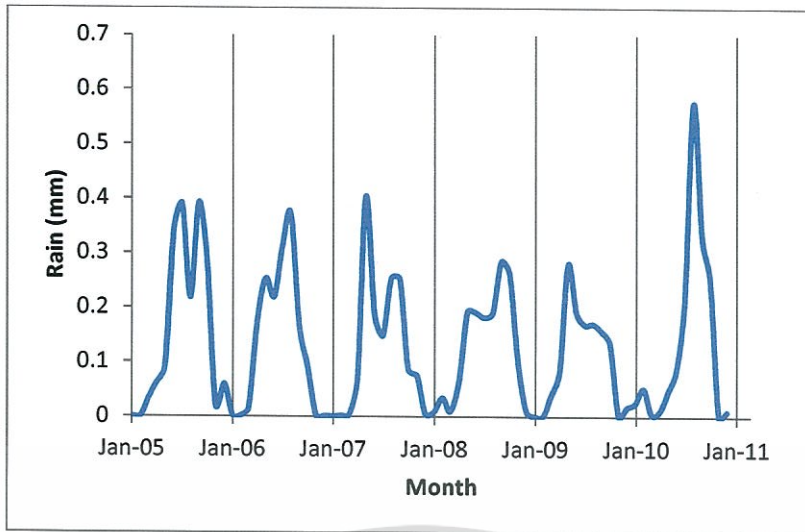
รูปที่ 3.3 ปริมาณก๊าซไนโตรเจนไดออกไซด์รายเดือนตั้งแต่ปี 2005 – 2010

ค่าของข้อมูลที่นำมาพิจารณา มีการเปลี่ยนแปลงซ้ำทุกๆ 12 เดือน จึงสรุปได้ว่าข้อมูลมีการเปลี่ยนแปลงตามฤดูกาล เนื่องจากค่าในแต่ละปีมีการเปลี่ยนแปลงที่อยู่ในระดับใกล้เคียงกัน นั้นหมายความว่าข้อมูลมีการเปลี่ยนแปลงเชิงบวกด้วยเช่นกัน รูปที่ 3.1 – 3.5 แสดงปริมาณของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันที่ได้รับการตรวจวัดในพื้นที่จังหวัดเชียงใหม่ พบว่าแต่ละปัจจัยมีการเปลี่ยนแปลงตามฤดูกาล โดยมีคาบฤดูกาลเป็น 12 เดือน ซึ่งต่อไปเราจะทำการพยากรณ์ข้อมูลอนุกรมเวลาของปัจจัยดังกล่าวนี้โดยใช้การวิเคราะห์อนุกรมเวลา จากสมการ (2.4) – (2.7)



รูปที่ 3.4 ปริมาณความชื้นสัมพัทธ์รายเดือนตั้งแต่ปี 2005 – 2010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 ปริมาณน้ำฝนรายเดือนตั้งแต่ปี 2005 – 2010

เนื่องจากอนุกรมเวลาชุดนี้ไม่มีแนวโน้ม และมีคาบของฤดูกาลเท่ากับ 12 เดือน เราเลือกวิเคราะห์ด้วยวิธีวินเทอร์กรณีสฤดูกาลเชิงบวกชนิดที่ไม่มีแนวโน้ม ฉะนั้นตัวแบบที่ใช้ในการพยากรณ์เมื่อ $n = 72$ และ $s = 12$ คือ

$$\hat{Z}_{72}(l) = \hat{\beta}_0(72) + \hat{S}_{60+l}, \quad l = 1, 2, \dots \quad (3.1)$$

โดยที่ค่า $\hat{\beta}_0(72)$ และ \hat{S}_{60+l} ได้มาจากสมการ

$$\hat{\beta}_0(n) = \alpha_1(Z_n - \hat{S}_{n-12}) + (1 - \alpha_1)\hat{\beta}_0(n-1) \quad (3.2)$$

$$\hat{S}_n = \alpha_3[Z_n - \hat{\beta}_0(n)] + (1 - \alpha_3)\hat{S}_{n-12} \quad (3.3)$$

เนื่องจากข้อมูลมีการเปลี่ยนแปลงตามฤดูกาลเชิงบวกชนิดที่ไม่มีแนวโน้ม ทำให้ไม่มีพจน์ β_1 หรือสมการ (2.5) ในตัวแบบ ในกรณีนี้ค่าคงที่ในการทำให้เรียบที่ต้องกำหนดคือ α_1 และ α_3 ซึ่งการศึกษาครั้งนี้ใช้การวิเคราะห์และตัดสินใจโดยใช้โปรแกรมสำเร็จรูปช่วยในการหาค่าตอบ

3.3 ตัวแบบการวิเคราะห์การถดถอยโลจิสติก

สถานการณ์หมอกควันมักปรากฏชัดในช่วงเดือนกุมภาพันธ์ถึงเมษายนเท่านั้น เนื่องจากระยะเวลาดังกล่าวถืออยู่ในหน้าแล้ง สภาพอากาศนิ่ง ทำให้สารมลพิษสามารถอยู่ในชั้นบรรยากาศได้เป็นเวลานาน ประกอบกับพื้นที่ภาคเหนือตอนบนมีลักษณะเป็นแอ่งกระทะ ทำให้สารมลพิษทางอากาศกระจายตัวได้น้อย สำหรับตัวแบบพยากรณ์การเกิดหมอกควัน เราพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์เพื่อวัดความสัมพันธ์ของปัจจัยต่างๆ ที่มีต่อ PM_{10} โดยข้อมูลอนุกรมเวลาที่นำมาพิจารณาเป็นข้อมูลค่าเฉลี่ยรายวัน คำนวณตามสมการ (2.2)

ตารางที่ 3.2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยต่างๆและ PM₁₀ (รายวัน)

Factor	Correlation	Factor	Correlation
CO	0.6676	Temperature	0.0160
NO	0.3191	Relative Humidity (RH)	-0.6633
NO ₂	0.8845	Pressure	0.3278
SO ₂	0.4719	Wind speed	-0.0642
O ₃	0.5576	Rain	-0.2234
Wind direction	-0.0865		

จากตารางที่ 3.2 เราสามารถสรุปได้ว่า

ปัจจัยที่มีความสัมพันธ์กับ PM₁₀ สูง ได้แก่ ก๊าซคาร์บอนมอนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO₂) ความชื้นสัมพัทธ์ (RH) และก๊าซโอโซน (O₃)

ปัจจัยที่มีความสัมพันธ์กับ PM₁₀ ปานกลาง ได้แก่ ความดัน (Pressure) ก๊าซไนโตรเจนมอนอกไซด์ (NO) และก๊าซซัลเฟอร์ไดออกไซด์ (SO₂)

ปัจจัยที่มีความสัมพันธ์กับ PM₁₀ ต่ำ ได้แก่ ความเร็วลม (Wind speed) ทิศทางลม (Wind direction) อุณหภูมิ (Temperature) และปริมาณน้ำฝน (Rain)

ตัวแบบการวิเคราะห์การถดถอยโลจิสติกสำหรับตัวแบบการเกิดหมอกควันใช้ปัจจัยที่มีความสัมพันธ์กับ PM₁₀ สูงและปานกลางในการสร้างตัวแบบ นั่นคือปัจจัยที่นำมาพิจารณาทั้งสิ้น 7 ปัจจัย ได้แก่ ก๊าซคาร์บอนมอนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์

ในการตรวจสอบว่าตัวแปรอิสระมีความสัมพันธ์กันมากเกินไปหรือไม่นั้น พิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ ตารางที่ 3.3 แสดงให้เห็นว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่มีค่าไม่เกิน 0.65 ตามเกณฑ์ของ Burns and Grove (1993) นั่นคือ ไม่เกิดปัญหาที่ตัวแปรอิสระมีความสัมพันธ์กันมาก (multicollinearity) ตัวแปรอิสระทั้ง 7 สามารถนำไปสร้างแบบจำลองการวิเคราะห์การถดถอยโลจิสติกได้

ตารางที่ 3.3 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระต่างๆ

Factor	CO	NO ₂	RH	O ₃	Pressure	NO	SO ₂
CO	1						
NO ₂	0.6502	1					
RH	-0.3872	-0.5324	1				
O ₃	0.1688	0.4853	-0.6115	1			
Pressure	0.3202	0.3192	-0.2803	0.1237	1		
NO	0.4283	0.5380	-0.0213	-0.1483	0.1653	1	
SO ₂	0.4087	0.3577	-0.2446	0.0715	0.3183	0.2296	1

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการศึกษาวิจัยเท่านั้น ไม่สามารถนำข้อมูลไปใช้ในการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยทั่วไปแล้ว กรมควบคุมมลพิษจะรายงานความรุนแรงของปัญหาหมอกควันเมื่อตรวจวัดค่า PM_{10} ได้สูงถึง 120 ไมครอน แต่เมื่อค่า PM_{10} มีค่าเกิน 100 ก็ควรเฝ้าระวัง เริ่มมีผลกระทบต่อสุขภาพของคนในพื้นที่ โดยเฉพาะผู้ที่ภูมิคุ้มกันต่ำ หรือผู้มีโรคประจำตัวเกี่ยวกับระบบทางเดินหายใจ ในการศึกษาครั้งนี้แบ่งระดับการเกิดหมอกควันเป็น 3 ระดับ คือ ค่า PM_{10} ที่อยู่ระหว่าง 0 – 99 ไมครอน หมายถึง “ไม่มีผลกระทบต่อสุขภาพ” ค่า PM_{10} ที่อยู่ระหว่าง 100 – 119 ไมครอน หมายถึง “มีผลกระทบต่อสุขภาพ” และ ค่า PM_{10} ที่เกินกว่า 120 ไมครอน หมายถึง “มีผลกระทบต่อสุขภาพมาก”

กำหนดให้ p_1 แทน ความน่าจะเป็นที่ PM_{10} มีค่าอยู่ระหว่าง 0 – 99 ไมครอน

p_2 แทน ความน่าจะเป็นที่ PM_{10} มีค่าอยู่ระหว่าง 100 – 199 ไมครอน

p_3 แทน ความน่าจะเป็นที่ PM_{10} มีค่ามากกว่า 120 ไมครอน

ให้ p_1 เป็นกลุ่มอ้างอิง ตัวแบบโลจิท (Logit Model) คือ

$$\ln \frac{p_2}{p_1} = \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \dots + \beta_{72}x_7 \quad (3.4)$$

$$\ln \frac{p_3}{p_1} = \beta_{03} + \beta_{13}x_1 + \beta_{23}x_2 + \dots + \beta_{73}x_7 \quad (3.5)$$

$$p_1 + p_2 + p_3 = 1 \quad (3.6)$$

ความน่าจะเป็นสามารถสรุปได้ดังนี้

$$p_1 = \frac{1}{\left(\sum_{i=2}^3 \exp(\beta_{0i} + \beta_{1i}x_1 + \dots + \beta_{7i}x_7) + 1 \right)} \quad (3.7)$$

$$p_2 = \frac{\exp(\beta_{02} + \beta_{12}x_1 + \dots + \beta_{72}x_7)}{1 + \left(\sum_{i=2}^3 \exp(\beta_{0i} + \beta_{1i}x_1 + \dots + \beta_{7i}x_7) + 1 \right)} \quad (3.8)$$

$$p_3 = \frac{\exp(\beta_{03} + \beta_{13}x_1 + \dots + \beta_{73}x_7)}{1 + \left(\sum_{i=2}^3 \exp(\beta_{0i} + \beta_{1i}x_1 + \dots + \beta_{7i}x_7) + 1 \right)} \quad (3.9)$$

โดยที่ $x_1, x_2, x_3, \dots, x_7$ คือค่าของก๊าซคาร์บอนมอนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์ ตามลำดับ ค่าพารามิเตอร์ β หาโดยวิธีนิวตันราฟสัน

3.4 ตัวแบบการวิเคราะห์การถดถอยพหุคูณโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

จากการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ระหว่างข้อมูลปัจจัยรายเดือนเปรียบเทียบกับ PM_{10} ดังแสดงในตารางที่ 3.1 โดยปัจจัยที่นำมาสร้างตัวแบบนั้นพิจารณาจากปัจจัยที่มีผลต่อ PM_{10} สูงและปานกลาง ซึ่งมีทั้งสิ้น 7 ปัจจัยได้แก่ ก๊าซคาร์บอนมอนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์
สมการการวิเคราะห์การถดถอยพหุคูณเขียนได้เป็น

$$\hat{Y}_i = a + b_1X_{1i} + b_2X_{2i} + \dots + b_7X_{7i} \quad (3.10)$$

โดยที่ X_1, X_2, \dots, X_7 คือค่าของก๊าซคาร์บอนมอนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์
ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์
ตามลำดับ การประมาณค่าพารามิเตอร์ a, b_1, \dots, b_7 หาโดยขั้นตอนวิธีเชิงพันธุกรรม โดยกำหนด
ฟังก์ชันวัตถุประสงค์เป็นค่า RMSE



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการดำเนินการ

ในส่วนของผลจากการดำเนินการวิจัยแสดง ตัวแบบพยากรณ์อนุกรมเวลาของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน และตัวแบบพยากรณ์การเกิดหมอกควัน โดยการวิเคราะห์การถดถอยโลจิสติกที่ได้จากการวิเคราะห์

4.1 ตัวแบบพยากรณ์อนุกรมเวลาของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน

ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในการวิเคราะห์อนุกรมเวลานี้ เลือกจากปัจจัยที่มีความสัมพันธ์กับ PM_{10} สูง ซึ่งในตารางที่ 3.1 แสดงปัจจัยที่มีความสัมพันธ์กับ PM_{10} สูง ทั้งสิ้น 4 ปัจจัย ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ และก๊าซโอโซน ผลการหาค่า α_1 และ α_3 ที่ทำให้ RMSE มีค่าน้อยที่สุดของแต่ละตัวแปรแสดงได้ดังนี้

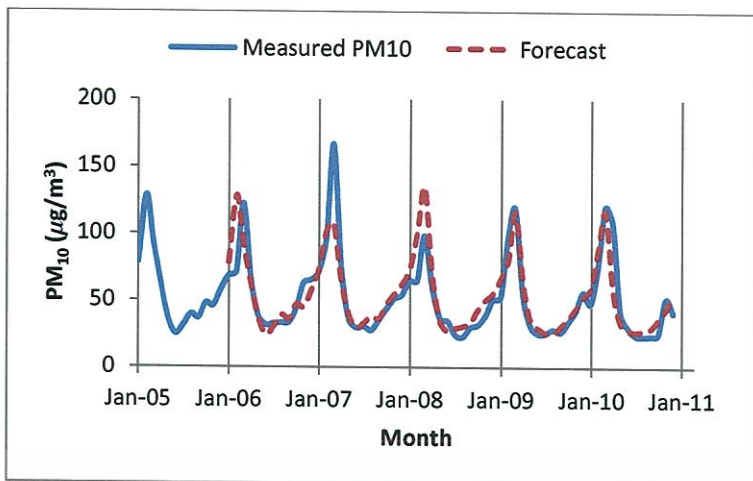
ตารางที่ 4.1 ค่า α_1 และ α_3 ที่ทำให้ RMSE มีค่าน้อยที่สุด

ตัวแปร	α_1	α_3	RMSE
PM_{10}	0.0242	0.4715	16.105
CO	0.0792	0.7383	0.244
NO_2	0.0211	0.5426	3.411
RH	0.4522	1	6.033
Rain	0.3921	0.8520	0.100

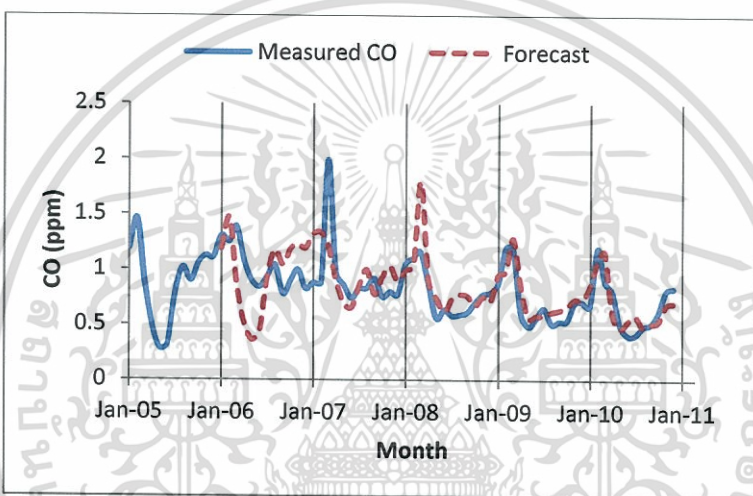
จากข้อมูลอนุกรมเวลาทำให้เราได้กราฟความสัมพันธ์ระหว่างปัจจัยต่างๆที่เราสนใจและเวลา โดยกราฟที่แสดงแกนนอน หมายถึงเดือนนับตั้งแต่ ค.ศ. 2005 - ค.ศ. 2010 กราฟเส้นตรงคือค่าที่ได้จากการตรวจวัดโดยกรมควบคุมมลพิษ ส่วนกราฟเส้นประคือค่าที่ได้จากการพยากรณ์ ตัวแบบทางคณิตศาสตร์ที่ได้มีดังนี้

จากรูปที่ (4.1) – (4.5) เราจะเห็นว่า PM_{10} และปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันนั้นมีการเปลี่ยนแปลงตามฤดูกาล และไม่มีแนวโน้ม ตัวแบบนี้เป็นตัวแบบที่เหมาะสมสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเชิงบวก ซึ่งหากข้อมูลไม่มีเหตุการณ์ที่ผิดปกติจะทำค่าคลาดเคลื่อนที่ได้จากตัวแบบมีค่าน้อย แต่ถ้าข้อมูลมีเหตุการณ์ผิดปกติ สังเกตว่าค่าจริงและค่าพยากรณ์แตกต่างกันมาก จะทำให้เกิดค่าคลาดเคลื่อนมาก และค่าพยากรณ์ในแต่ละเดือนของปีถัดมาจะไม่แตกต่างจากค่าพยากรณ์ในแต่ละเดือนของปีก่อนหน้ามากนัก

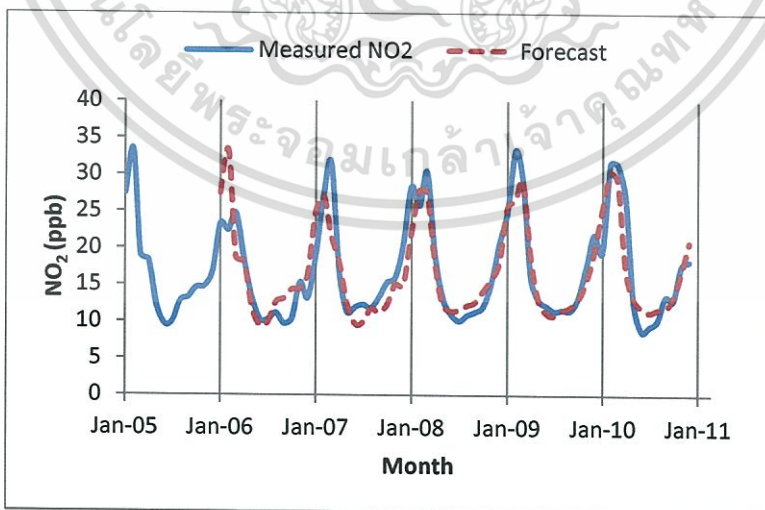
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 ปริมาณ PM₁₀ ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน ปี ค.ศ.2005 – 2010

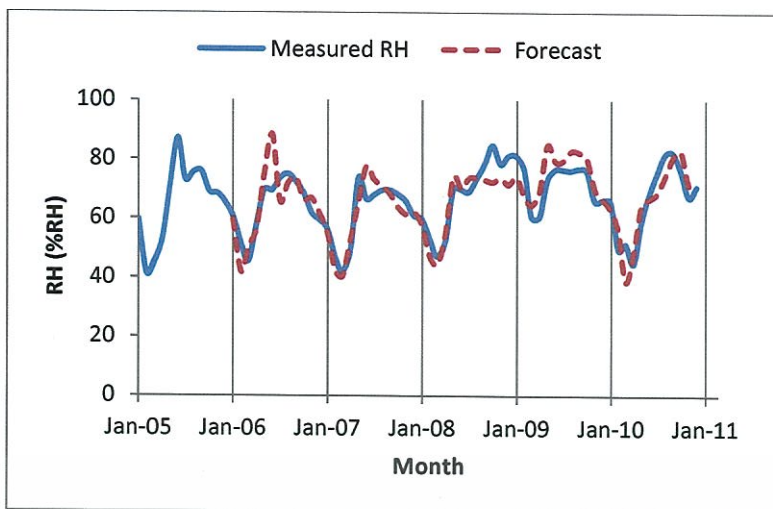


รูปที่ 4.2 ปริมาณ CO ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน ปี ค.ศ.2005 - 2010

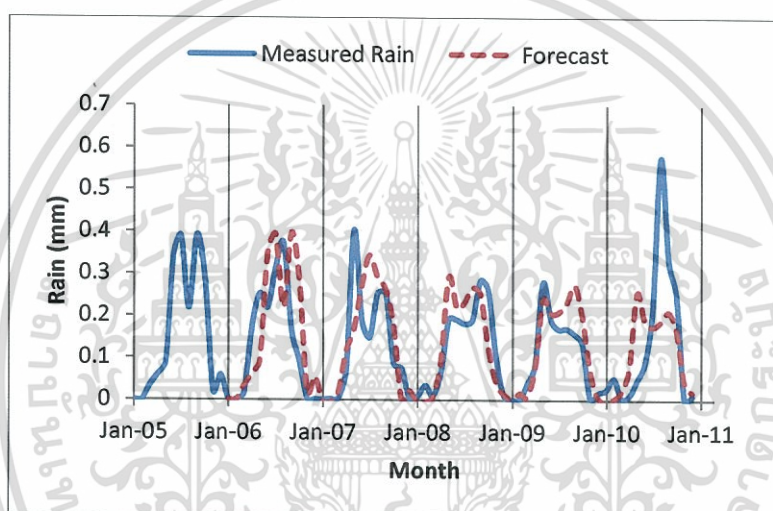


รูปที่ 4.3 ปริมาณ NO₂ ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือนปี ค.ศ.2005 - 2010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 ปริมาณความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน



รูปที่ 4.5 ปริมาณน้ำฝนที่ได้จากการตรวจวัดและค่าพยากรณ์รายเดือน ปี ค.ศ.2005 – 2010

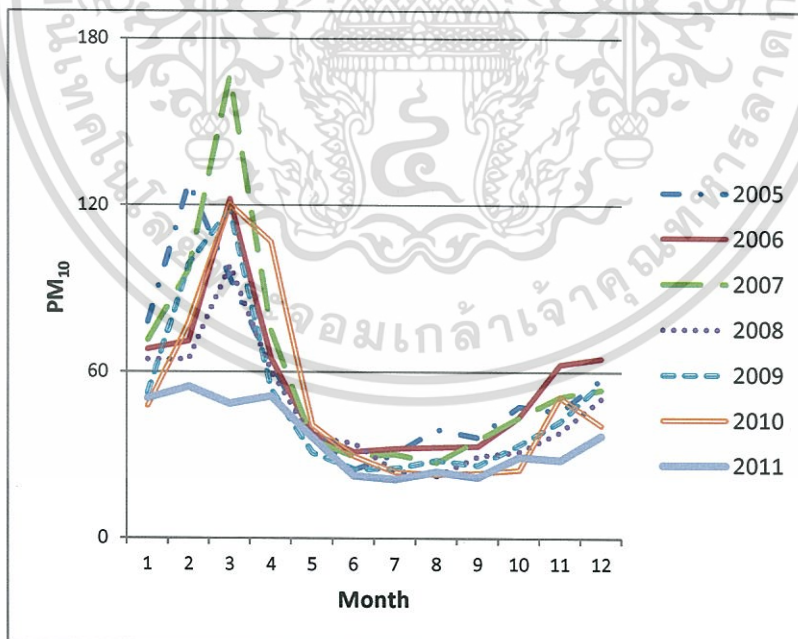
ตารางที่ 4.3 แสดงค่าที่ได้จากการตรวจวัดโดยกรมควบคุมมลพิษ และค่าพยากรณ์ที่ได้จากตัวแบบพยากรณ์ของไฮลด์และวินเนอร์ ให้ค่าความคลาดเคลื่อนน้อย ยกเว้นของ PM_{10} ซึ่งเมื่อสังเกตจากตาราง พบว่าในช่วงเดือน 2 – 4 ค่าที่พยากรณ์ได้มีความแตกต่างจากค่าที่ได้จากการตรวจวัด โดยเฉพาะเดือนที่ 3 ค่าที่พยากรณ์ได้คือ 118.73 แต่ค่าที่ตรวจวัดได้คือ 48.83 ซึ่งเป็นค่าที่แตกต่างกันมาก ทำให้เกิดค่าคลาดเคลื่อนสูง หากเราสังเกตค่า PM_{10} ที่ได้จากการตรวจวัดอยู่ในช่วง 90 – 120 ไมครอน แสดงว่าในปี 2011 ความรุนแรงของสถานการณ์หมอกควันในบริเวณที่ทำการศึกษามิสูงนัก ทำให้ PM_{10} ถูกพบได้น้อยกว่าปี 2005 - 2010

จากตารางที่ 4.3 ค่าพยากรณ์ PM_{10} ที่ได้จากตัวแบบให้ความคลาดเคลื่อนสูง เราจึงศึกษาปริมาณ PM_{10} ที่ได้จากการตรวจวัด ปี ค.ศ.2005 – 2011 ดังรูปที่ 4.11 พบว่าสำหรับปีค.ศ.2005 –

2010 อื่นๆ PM_{10} จะมีแนวโน้มสูงกว่าช่วงอื่นๆของปี แต่ปริมาณ PM_{10} ที่ได้จากการตรวจวัดในปี 2011 ช่วงเวลาดังกล่าวมี ปริมาณ PM_{10} น้อยกว่าปกติ

ตารางที่ 4.2 ค่าจริงและค่าพยากรณ์ของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในปีค.ศ.2011

Month	PM_{10}		CO		NO_2		RH		Rain	
	Real	Forecast	Real	Forecast	Real	Forecast	Real	Forecast	Real	Forecast
1	50.36	54.23	0.83	0.72	21.32	22.26	66.93	68.49	21.32	22.26
2	54.59	84.11	0.82	1.17	24.88	31.14	54.25	54.89	24.88	31.14
3	48.83	118.73	0.87	0.98	18.94	30.58	65.25	51.04	18.94	30.58
4	51.41	81.00	0.70	0.76	16.57	22.39	69.34	45.96	16.57	22.39
5	36.71	36.29	0.52	0.50	12.52	12.74	79.00	62.45	12.52	12.74
6	22.69	28.03	0.46	0.45	11.09	10.09	77.89	71.26	11.09	10.09
7	21.34	24.47	0.48	0.48	8.96	10.21	76.28	75.19	8.96	10.21
8	24.03	24.79	0.58	0.51	8.54	10.88	81.40	78.18	8.54	10.88
9	22.09	26.09	0.61	0.54	8.10	12.71	81.26	78.01	8.1	12.71
10	29.36	30.27	0.58	0.61	10.76	13.10	79.76	74.84	10.76	13.1
11	28.06	46.76	0.77	0.79	15.38	17.20	68.98	67.73	15.38	17.2
12	37.06	47.89	0.72	0.80	17.69	19.44	70.41	70.83	17.69	19.44
RMSE : 24.446		RMSE : 0.118		RMSE : 4.570		RMSE : 9.650		RMSE : 0.117		



รูปที่ 4.6 ปริมาณ PM_{10} ที่ได้จากการตรวจวัด ปี ค.ศ.2005 - 2011

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ตัวแบบพยากรณ์การเกิดหมอกควันโดยการวิเคราะห์การถดถอยโลจิสติก

การศึกษาครั้งนี้แบ่งระดับการเกิดหมอกควันเป็น 3 ระดับ คือ ค่า PM_{10} ที่อยู่ระหว่าง 0 – 99 ไมครอน หมายถึง “ไม่มีผลต่อสุขภาพ” ค่า PM_{10} ที่อยู่ระหว่าง 100 – 119 ไมครอน หมายถึง “มีผลต่อสุขภาพ” และ ค่า PM_{10} ที่เกินกว่า 120 ไมครอน หมายถึง “มีผลต่อสุขภาพมาก” กำหนดให้การเกิดหมอกควันระดับที่ไม่มีผลต่อสุขภาพเป็นกลุ่มอ้างอิง

```
. mlogit out co no2 rh o3 pressure no so2
```

```
Iteration 0: log likelihood = -749.87965
Iteration 1: log likelihood = -580.02889
Iteration 2: log likelihood = -399.03679
Iteration 3: log likelihood = -333.00859
Iteration 4: log likelihood = -312.14938
Iteration 5: log likelihood = -304.21215
Iteration 6: log likelihood = -302.97569
Iteration 7: log likelihood = -302.9412
Iteration 8: log likelihood = -302.94117
```

```
Multinomial logistic regression
```

```
Log likelihood = -302.94117
```

```
Number of obs = 2099
LR chi2(14) = 893.88
Prob > chi2 = 0.0000
Pseudo R2 = 0.5960
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
2					
co	.8957	.4771868	1.88	0.061	-.0395689 1.830969
no2	.2051789	.0335445	6.12	0.000	.1394329 .2709249
rh	-.0518598	.0160146	-3.24	0.001	-.0832479 -.0204717
o3	.0680068	.0227069	2.99	0.003	.023502 .1125116
pressure	.1133021	.0718298	1.58	0.115	-.0274817 .2540859
no	-.1139172	.0421187	-2.70	0.007	-.1964682 -.0313661
so2	.2559123	.1436128	1.78	0.075	-.0255637 .5373883
_cons	-89.63859	52.73306	-1.70	0.089	-192.9935 13.71632
3					
co	2.231597	.4368017	4.58	0.000	1.277483 3.185711
no2	.3623642	.0398866	9.08	0.000	.284183 .4405405
rh	-.1028791	.018717	-5.50	0.000	-.1395638 -.0661944
o3	.1358265	.0252098	5.39	0.000	.0864161 .1852368
pressure	.2418897	.0826411	2.93	0.003	.0799161 .4038634
no	-.1296898	.0392339	-3.31	0.001	-.2065868 -.0527928
so2	-.0596997	.155681	-0.38	0.701	-.3648289 .2454296
_cons	-188.3436	60.94121	-3.09	0.002	-307.7862 -68.901

(out=1 is the base outcome)

รูปที่ 4.7 ผลการวิเคราะห์แบบจำลองการวิเคราะห์การถดถอยโลจิสติก

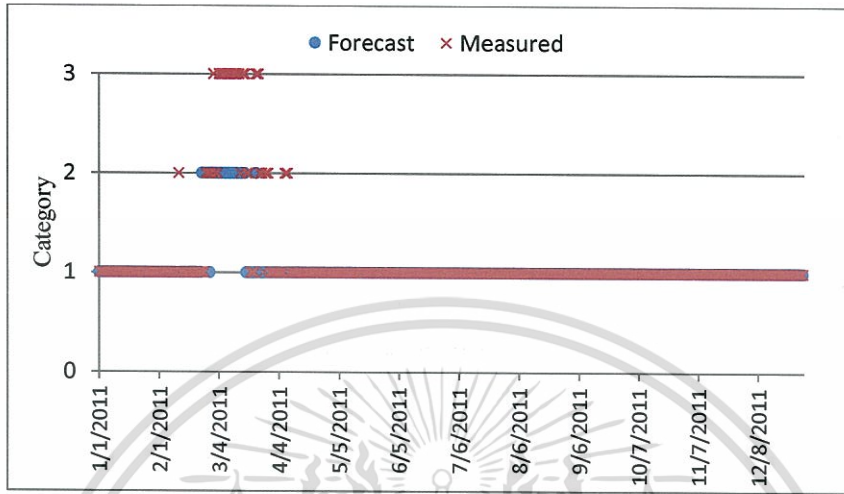
จากรูปที่ 4.7 ตัวแบบโลจิสติกที่ได้คือ

$$\ln \frac{p_2}{p_1} = 0.9x_1 + 0.21x_2 - 0.05x_3 + 0.07x_4 + 0.11x_5 - 0.11x_6 + 0.26x_7 - 89.64 \quad (4.1)$$

$$\ln \frac{p_3}{p_1} = 2.23x_1 + 0.36x_2 - 0.1x_3 + 0.14x_4 + 0.24x_5 - 1.13x_6 - 0.06x_7 - 188.34 \quad (4.2)$$

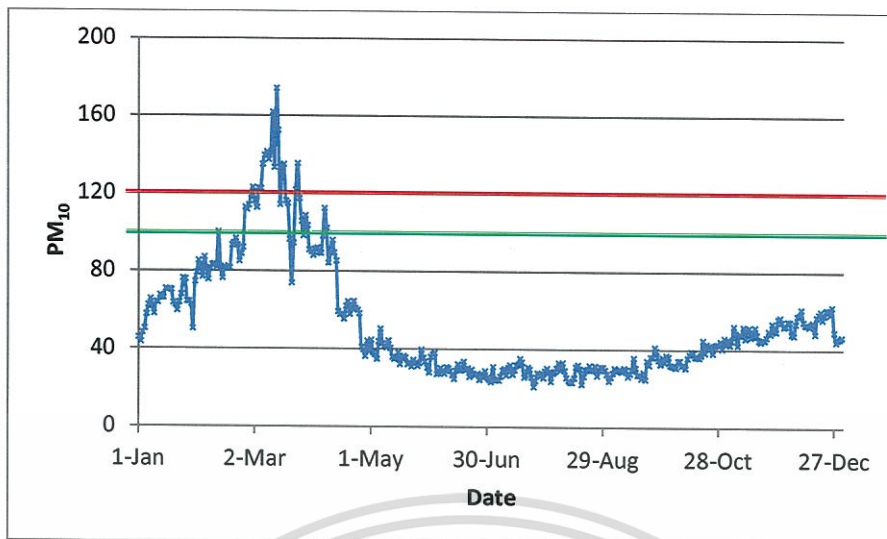
เมื่อ X_1, X_2, \dots, X_7 หมายถึงข้อมูลรายวันของก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนนอกไซด์ และก๊าซซัลเฟอร์

ได้ออกไซด์ตามลำดับ ในการหาค่าพารามิเตอร์ด้วยวิธีของนิวตันราฟสันให้ค่าความควรจะเป็นสูงสุดเป็น -302.94 ซึ่งเกิดจากการคำนวณซ้ำ 8 ครั้ง ความแปรปรวนของตัวแปรตามที่สามารถอธิบายได้ด้วยตัวแปรอิสระมีค่าเป็น 59.60 ($Pseudo R^2 = 0.5960$)

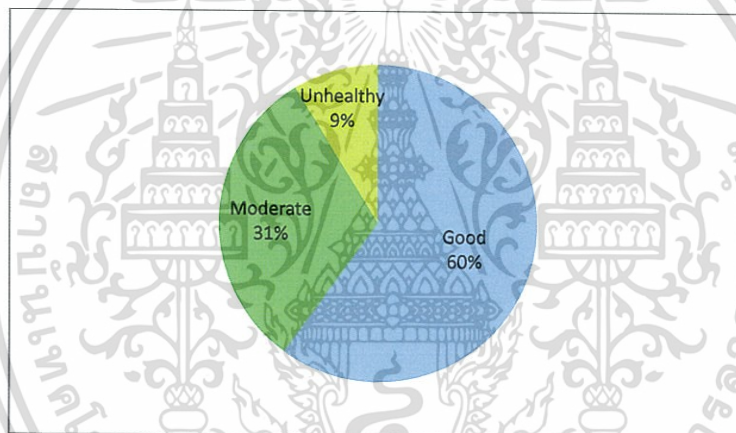


รูปที่ 4.8 ผลการวิเคราะห์การเกิดหมอกควันโดยการวิเคราะห์การถดถอยโลจิสติกในปี 2011

รูปที่ 4.8 แสดงค่าที่พยากรณ์ระดับของการเกิดหมอกควันในปี 2011 ที่ได้จากตัวแบบการวิเคราะห์การถดถอยโลจิสติกเปรียบเทียบกับระดับของการเกิดหมอกควันที่ได้จากการตรวจวัด โดยตัวแบบพยากรณ์ในปี 2011 สร้างจากการค่าเฉลี่ยของข้อมูลรายวันของ PM_{10} ตั้งแต่ปี 2005 - 2010 จากการวิเคราะห์ข้อมูลที่ได้จากการตรวจวัดพบว่า PM_{10} มีค่าเกินมาตรฐานในช่วงมีนาคมถึงเมษายน ดังแสดงในรูปที่ 4.10 และเมื่อทำการวิเคราะห์เกณฑ์ดัชนีคุณภาพอากาศของ PM_{10} ในปี 2011 พบว่า PM_{10} อยู่ในคุณภาพอากาศดี 60% (219/365) คุณภาพอากาศปานกลาง 31% (114/365) และคุณภาพอากาศที่มีผลต่อสุขภาพ 9% (32/365) ซึ่งสอดคล้องกับรูปที่ 4.9 ซึ่งแสดงว่า PM_{10} มีค่าเกินมาตรฐานในช่วง 1 เดือนหรือประมาณ 30 วันนั่นเอง



รูปที่ 4.9 ความเข้มข้นรายวันของ PM10 ที่ได้จากการตรวจวัดในปี ค.ศ. 2011



รูปที่ 4.10 ระดับเกณฑ์ดัชนีคุณภาพอากาศของ PM₁₀ ในปี ค.ศ. 2011

แม้ว่าเกณฑ์ดัชนีคุณภาพอากาศส่วนใหญ่อยู่เกณฑ์คุณภาพอากาศดีถึงร้อยละ 60 แต่หากพิจารณาในช่วงที่เกิดปัญหาหมอกควันนั้น แทบจะไม่พบเกณฑ์ดัชนีคุณภาพอากาศดี ในช่วงที่เกิดปัญหาหมอกควันคุณภาพอากาศอยู่ในช่วงปานกลางประมาณร้อยละ 50 และคุณภาพอากาศที่มีผลต่อสุขภาพประมาณร้อยละ 50 เช่นกัน

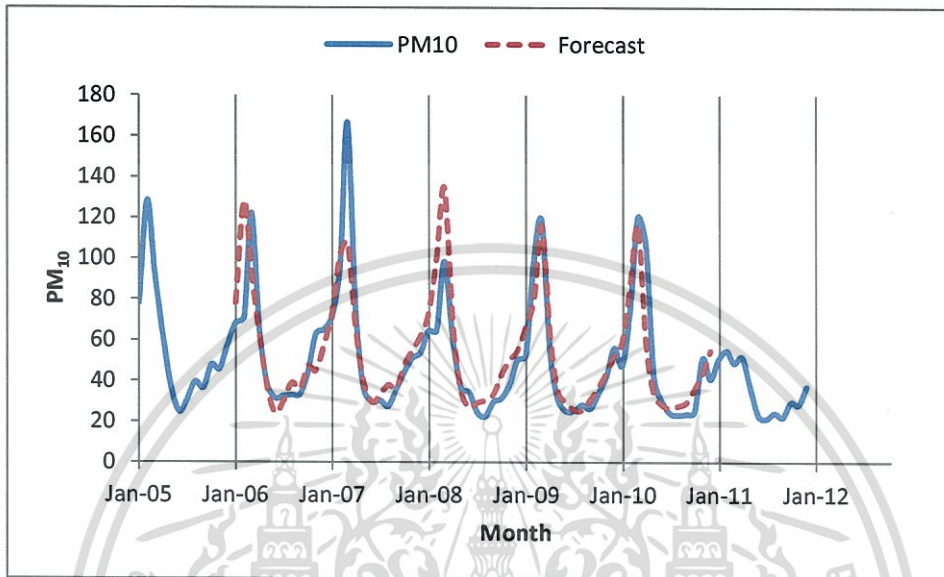
4.3 ตัวแบบพยากรณ์ PM₁₀ โดยการวิเคราะห์การถดถอยพหุคูณ

สมการการวิเคราะห์การถดถอยพหุคูณที่ได้จากขั้นตอนวิธีเชิงพันธุกรรมคือ

$$PM_{10} = 1.43 + 3.08CO + 3.07NO_2 - 0.64RH - 5.5Rain + 0.06Pressure - 0.82NO + 0.13SO_2$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยตัวแบบให้ค่า $R^2 = 0.8615$ และ $RMSE = 11.18$ ซึ่งหากเปรียบเทียบการพยากรณ์ PM_{10} ด้วยการวิเคราะห์อนุกรมเวลา พบว่าตัวแบบพยากรณ์ PM_{10} โดยการวิเคราะห์การถดถอยพหุคูณให้ค่า $RMSE$ ดีกว่า ซึ่งหมายความว่าหากต้องการพยากรณ์ PM_{10} อย่างมีประสิทธิภาพควรอาศัยปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันด้วย



รูปที่ 4.11 ความเข้มข้นรายวันของ PM_{10} ที่ได้จากการตรวจวัดและค่าพยากรณ์ในปี ค.ศ. 2011 ของตัวแบบการวิเคราะห์การถดถอยพหุคูณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรุปผลการดำเนินการ

ในงานวิจัยชิ้นนี้ได้ทำการวิเคราะห์หอนุกรมเวลาเพื่อพยากรณ์ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน ซึ่งเป็นการอธิบายภาพรวมของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน โดยใช้ข้อมูลรายเดือนของปัจจัยต่างๆ นอกจากนี้ยังได้ทำการวิเคราะห์การถดถอยโลจิสติกเพื่อสร้างตัวแบบการเกิดปัญหาหมอกควัน ซึ่งเป็นการวิเคราะห์เชิงลึก เนื่องจากการเกิดหมอกควันไม่ได้เกิดขึ้นตลอดทั้งปี แต่เป็นช่วงเวลาหนึ่งเท่านั้น ข้อมูลที่นำมาพิจารณาจึงเป็นข้อมูลรายวัน เพื่อศึกษาการเกิดหมอกควันในแต่ละวัน โดยอาศัยปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันเช่นเดียวกับการวิเคราะห์หอนุกรมเวลา แต่แตกต่างกันที่ใช้ข้อมูลรายวันในการพิจารณา

ในการวิเคราะห์หอนุกรมเวลาเพื่อพยากรณ์ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน พบว่าปัจจัยที่มีผลต่อการเกิดหมอกควันของจังหวัดเชียงใหม่ได้แก่ ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) ก๊าซคาร์บอนมอนนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO_2) ความชื้นสัมพัทธ์ (Relative humidity : RH) และปริมาณน้ำฝน (Rain) ซึ่งเมื่อวิเคราะห์ข้อมูลหอนุกรมเวลาของปัจจัยดังกล่าว พบว่าเป็นข้อมูลหอนุกรมเวลาที่มีฤดูกาลการเปลี่ยนแปลงตามฤดูกาลในเชิงบวก (Additive seasonal effect) ซึ่งสามารถหาค่าพยากรณ์โดยวิธีของวินเตอร์ โดยหากข้อมูลไม่มีเหตุการณ์ที่ผิดปกติจะทำค่าคลาดเคลื่อนที่ได้จากตัวแบบมีค่าน้อย แต่ถ้าข้อมูลมีเหตุการณ์ผิดปกติจะทำให้เกิดค่าคลาดเคลื่อนมาก และค่าพยากรณ์ในแต่ละเดือนของปีถัดมาจะไม่แตกต่างจากค่าพยากรณ์ในแต่ละเดือนของปีก่อนหน้ามากนัก

ปัญหาหมอกควันในภาคเหนือตอนบนของประเทศไทยครอบคลุมพื้นที่ 8 จังหวัด ได้แก่ จังหวัดเชียงใหม่ เชียงราย ลำพูน ลำปาง แม่ฮ่องสอน พะเยา แพร่ และน่าน โดยความรุนแรงปรากฏชัดในช่วงหน้าแล้งของทุกปี สาเหตุเกิดจากสภาวะอากาศที่แห้งและนิ่ง ทำให้ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) ที่เกิดขึ้นสามารถแขวงลอยอยู่ในบรรยากาศได้นาน รวมทั้งปริมาณฝุ่นละอองที่เพิ่มขึ้น เนื่องจากสภาวะอากาศดังกล่าวส่งผลให้การเกิดไฟป่าเพิ่มขึ้น การเผาที่เกิดจากฝีมือมนุษย์ หากว่าค่าเฉลี่ยรายวันของ PM_{10} มีค่าเกินกว่า 120 ไมโครกรัมต่อลูกบาศก์เมตรถือว่าอยู่ในเกณฑ์ที่มีผลต่อสุขภาพ งานวิจัยนี้มีวัตถุประสงค์เพื่อติดตามตรวจสอบการเกิดปัญหาหมอกควันโดยใช้การวิเคราะห์การถดถอยโลจิสติกเพื่อพยากรณ์การเกิดปัญหาหมอกควันในเชียงใหม่ โดยแบ่งการเกิดหมอกควันเป็น 3 ระดับ ตามระดับความเป็นอันตรายต่อสุขภาพ กลุ่มตัวอย่างคือข้อมูลของปัจจัยที่มีผลต่อการเกิดปัญหาหมอกควันตั้งแต่ปี ค.ศ. 2005 – 2010 ของพื้นที่จังหวัดเชียงใหม่ เพื่อสร้างสมการพยากรณ์ และพยากรณ์การเกิดหมอกควันในปี ค.ศ. 2011 ปัจจัยที่มีผลต่อการเกิด

หมอกควันที่ได้จากการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อมูลรายวันของ PM_{10} กับปัจจัยต่างๆ ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์ แบบจำลองการวิเคราะห์การถดถอยโลจิสติกอธิบายความแปรปรวนระหว่างกลุ่มตัวแปรพยากรณ์ได้ 56.29% ทำนายโอกาสในการเกิดหมอกควันได้ถูกต้อง 97.26% การวิเคราะห์สหสัมพันธ์ของปัจจัยต่างๆกับ PM_{10} พบว่า ก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ และความชื้นสัมพัทธ์ เป็นปัจจัยสำคัญที่พบทั้งในการวิเคราะห์ข้อมูลทั้งรายวันและรายเดือน

สำหรับแบบจำลองการวิเคราะห์การถดถอยพหุคูณ โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม เพื่อพยากรณ์ PM_{10} โดยอาศัยปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันสูงและปานกลาง ซึ่งมีทั้งสิ้น 7 ปัจจัย ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ ก๊าซโอโซน ความดัน ก๊าซไนโตรเจนมอนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์ ตัวแบบให้ค่าสัมประสิทธิ์การตัดสินใจ (R^2) 0.8615 และค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยเป็น 11.18 ซึ่งหากเปรียบเทียบกับพยากรณ์ PM_{10} โดยตัวแบบการวิเคราะห์หอนุกรมเวลาซึ่งให้ค่าสัมประสิทธิ์การตัดสินใจ (R^2) 0.7226 และค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยเป็น 16.105 จะเห็นได้ว่าแบบจำลองการวิเคราะห์การถดถอยพหุคูณให้ค่าสัมประสิทธิ์การตัดสินใจสูงกว่า และค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าตัวแบบการวิเคราะห์หอนุกรมเวลา นั่นคือแบบจำลองการวิเคราะห์การถดถอยพหุคูณมีความเหมาะสมที่จะใช้เป็นตัวแบบพยากรณ์ PM_{10} มากกว่าตัวแบบการวิเคราะห์หอนุกรมเวลา

เอกสารอ้างอิง

- กัลยา วานิชย์บัญชา. 2548. การวิเคราะห์ข้อมูลหลายตัวแปร. กรุงเทพฯ : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ขวัญศิริ ศิริมงคล และบุษยามาสพิมพ์พรรณชาติ. 2555. อนุกรมเวลาเพื่อพยากรณ์ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในพื้นที่จังหวัดเชียงใหม่. การประชุมวิชาการทางคณิตศาสตร์ครั้งที่ 18. พิทลุง : มหาวิทยาลัยทักษิณ.
- วนิดา จีนศาสตร์. 2551. มลพิษอากาศและการจัดการคุณภาพอากาศ. กรุงเทพฯ : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- นิพนธ์ ตังคณานุรักษ์ และคณิตา ตังคณานุรักษ์. 2552. เคมีบรรยากาศ. กรุงเทพฯ : สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.
- นภาพร พานิช และแสงสันดี พานิช. 2544. แบบจำลองทางคณิตศาสตร์ด้านคุณภาพอากาศ. กรุงเทพฯ : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- นันทพร ฉะอึ้งรัมย์ และคณะ. 2552. การหาค่าพารามิเตอร์ที่เหมาะสมด้วยวิธีเงินเนติกอัลกอริทึมในการพยากรณ์ข้อมูลคุณภาพน้ำ. ปัญหาพิเศษ สาขาคณิตศาสตร์ประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- ยุทธ ไกยวรรณ. 2555. หลักการและการใช้การวิเคราะห์การถดถอยโลจิสติกสำหรับการวิจัย. วารสารวิจัยมหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย 4(1) : 1-12.
- พนิดา พานิชกุล และยุทธภูมิ วงศ์วัฒนฤกษ์. 2546. คัมภีร์การวิเคราะห์และตัดสินใจปัญหาเชิงธุรกิจโดยใช้ Excel. กรุงเทพฯ : เคทีพี คอมพ์ แอนด์ คอนซัลท์.
- มงคล ราชนาคร และคณะ. 2550. โครงการการวิเคราะห์เพื่อหามลพิษทางอากาศในอนุภาคฝุ่นในจังหวัดเชียงใหม่และลำพูน. สำนักงานกองทุนสนับสนุนการวิจัย.
- มุกดา แม่นมินทร์. 2549. อนุกรมเวลาและการพยากรณ์. กรุงเทพฯ : ปรกาศพริก.
- อดิศัย โทวิชา. 2551. การวิเคราะห์การถดถอย. กรุงเทพฯ : โครงการตำราวิชาการเฉลิมพระเกียรติ 80 พรรษา มหาวิทยาลัยราชภัฏสวนสุนันทา.
- วนิดา รัตนมณี และ สุภชัย ปทุมนากุล. 2546. การหาคำตอบที่น่าพึงพอใจโดยเทคนิคเงินเนติกอัลกอริทึม. วิศวกรรมสาร มหาวิทยาลัยขอนแก่น 30(4) : 319-336.
- สิริกัลยา ประมวล. การเปรียบเทียบวิธีการถดถอยเชิงเส้นพหุคูณ วิธีโครงข่ายประสาทเทียมและวิธีเงินเนติกอัลกอริทึม โดยใช้องค์ประกอบหลักในการพยากรณ์อินทรีย์วัตถุในดิน. วิทยานิพนธ์ สาขาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ. 2554. **รู้รอบทิศ มลพิษทางอากาศ.**
กรุงเทพฯ : กชกร พับลิชชิ่ง.

Goldberg, D.E. 1989. **Genetic Algorithms in Search, Optimization, and Machine Learning.**
Massachusetts : Addison-Wesley Publishing.

Russel, Stuart and Norvig, Peter. 1995. **Artificial Intelligence: A Modern Approach.** New
Jersey : Prentice – Hall.

Davis, L. D., and Mitchell. M. 1991. **Handbook of Genetic Algorithms.** New York : Van
Nostrand Reinhold.

Phoothiwut, S. and Junyapoon, S. **Size Distribution of Atmospheric Particulate-bound
Polycyclic Aromatic Hydrocarbons and Characteristics of PAHs during Haze Period
in Lampang Province, Northern Thailand.** Air Quality, Atmosphere & Health, 6(2),
pp.397-405, 2013.

Chantara, S., Sillapapiromsuk, S., Wiriya, W., **Atmospheric Pollutants in Chiang Mai
(Thailand) over a Five-year Period (2005-2009), their Possible Sources and Relation
to Air Mass Movement.** Atmospheric Environment, 60, pp.88-98, 2012.

Wiriya, W., Sillapapiromsuk, S., Chantara, S., **PM₁₀-bound polycyclic aromatic hydrocarbons
in Chiang Mai (Thailand): Seasonal variations, source identification, health risk
assessment and their relationship to air-mass movement.** Atmospheric Research, 124,
pp.109-122, 2013.

PCD (Pollution Control Department). **Maryland Department of the Environment, and Chiang
Mai Municipality. Chiang Mai Emission Inventory in Municipality and
Neighborhood Area.** Report of Ability of Federal and Local Government Official,
Chaing Mai, 2002.

Pengchai, P., Chantara, S., Sopajaree, K., Wangkarn, S., Tengcharoenkul, U. and Rayanakorn,
M., **Seasonal Variation, Risk Assessment and Source Estimation of PM₁₀ and PM₁₀-
bound PAHs in the Ambient Air of Chiang Mai and Lumphun, Thailand.** Environ.
Monit. Assess., 154, pp. 197–218, 2009.

Pollution Control Department, **Manual measurement of dust in ambient,** 2003.

- Rayanakorn, M., **Haze and air pollution in Chiang Mai**, Login design work, 9-14, 2010.
- Jeremy, C., **Air pollution: an introduction**, E&FN Spon, 1997.
- World Health Organization, **WHO air quality guidelines global update 2005**, Report on working group meeting, Bonn, Germany, 18-20 October 2005.
- Kleinbaum, D.G., **Logistic regression: a self – learning text**, 3rd ed., Springer, 2010.
- Hair, J. F. and others, **Multivariate data analysis**, 7th ed., Pearson Prentice Hall, pp.319-348, 2006
- Hilbe, J.M., **Logistic Regression Models**, CRC Press, 2009.
- Hosmer, D.W., **Applied logistic regression**, 2nd ed., John Wiley, 2000.
- Shaodong, X. and others, **Characteristics of PM₁₀, SO₂, NO_x and O₃ in ambient air during the dust storm period in Beijing**, 345, pp.153-164, 2005.
- William R.D. & Matthew G., **Multivariate analysis: methods and applications**, John Wiley, pp.391-399, 1984.
- Chatterjee, S. & Hadi, A.S., **Regression Analysis by Example**, 4th ed., John Wiley & Sons Publisher, pp.317-335, 2006.
- Sophia, R. and Brian, S.E., **A handbook of statistical analyses using Stata**, 4th, Chapman & Hall, pp.111-129, 2006.
- Hamilton, L.C., **Statistics with Stata**, Brooks/Cole Publishing Company, pp.137-145, 1990.
- Burn, N. and Grove, S.K. 1993. **The Practice of Nursing Research: Conduct, Critique and Utilization**. W.B.Saunders Company, Philadelphia.
- Steven, J. 1996. **Applied multivariate statistics for the social science**. Lawrence Erlbaum Associate, Inc., Mahwah, New Jersey.
- [Online]. PCD (Pollution Control Department), <http://www.aqnis.pcd.go.th>, (6 July 2012).
- [Online]. กัปปัตน์ นี โม (นามแฝง) : <http://kapitaennem0.wordpress.com/2013/07/17/genetic-algorithm>.
- [Online]. กรมควบคุมมลพิษ กระทรวงทรัพยากรและสิ่งแวดล้อม : www.pcd.go.th

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตารางมาตรฐานคุณภาพอากาศของประเทศไทยเปรียบเทียบกับประเทศอื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มาตรฐานคุณภาพอากาศของประเทศไทยเปรียบเทียบกับประเทศอื่นๆ

ประเทศ	SO ₂			NO ₂			CO		O ₃		PM ₁₀	
	1 ชม.	24 ชม.	รายปี	1 ชม.	24 ชม.	รายปี	1 ชม.	8 ชม.	1 ชม.	8 ชม.	24 ชม.	รายปี
สหรัฐอเมริกา	-	0.37	-	-	-	0.10	0.10	10.0	0.24	-	0.15	0.05
ญี่ปุ่น	0.26	0.11	-	-	0.08	-	-	22.8	0.12	-	-	-
เนเธอร์แลนด์	0.76	0.23	-	0.18	-	-	40.0	6.0	0.12	-	-	-
ออสเตรเลีย	0.44	0.16	-	0.30	0.12	-	34.3	11.4	0.24	0.10	-	-
เม็กซิโก	-	-	-	0.40	-	-	-	15.0	-	-	-	-
ไต้หวัน	0.78	0.26	-	-	0.10	-	22.9	-	-	-	-	-
แคนาดา	0.82	0.27	-	0.40	0.20	-	15.0	6.0	0.10	-	-	-
เยอรมันนี	0-	0.27	-	0.20	-	0.80	30.0	-	-	-	-	-
WHO	0.35	0.13	0.05	0.40	0.15	-	30.0	10.0	0.15	0.10	-	-
ประเทศไทย	0.78	0.30	0.04	0.32	-	-	34.2	10.3	0.20	-	0.12	0.05

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The seal of the National Library of Thailand is a circular emblem. It features a central five-tiered umbrella (parasol) with a sunburst above it. On either side of the umbrella are two smaller, three-tiered umbrellas. The entire emblem is surrounded by a decorative border. The Thai text 'กรมหอสมุดแห่งชาติ' (National Library) is written along the top inner edge, and 'พระจอมเกล้าเจ้าคุณทหารลาดกระบัง' (King Chulalongkornrajavidyalaya University) is written along the bottom inner edge.

ภาคผนวก ข

ข้อมูลจากการตรวจวัดและค่าพยากรณ์ที่ได้จากการวิเคราะห์อนุกรมเวลา
โดยวิธีของโฮลต์และวินเซอร์แบบมีฤดูกาลเชิงบวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล PM₁₀ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา

Month-Year	Measured PM ₁₀	Base Level	Trend	Seasonal Factor	Forecasted PM ₁₀
Jan-05	77.872			21.088	
Feb-05	128.168			71.383	
Mar-05	93.817			37.032	
Apr-05	62.825			6.041	
May-05	37.225			-19.560	
Jun-05	24.679			-32.106	
Jul-05	30.383			-26.402	
Aug-05	39.320			-17.465	
Sep-05	36.444			-20.340	
Oct-05	47.577			-9.207	
Nov-05	45.620			-11.164	
Dec-05	57.484	56.785	0.000	0.700	
Jan-06	68.119	56.549	0.000	16.600	77.872
Feb-06	71.079	55.172	0.000	45.226	127.932
Mar-06	122.116	55.896	0.000	50.795	92.204
Apr-06	64.909	55.968	0.000	7.408	61.937
May-06	37.812	56.002	0.000	-18.914	36.408
Jun-06	31.297	56.181	0.000	-28.701	23.896
Jul-06	32.406	56.245	0.000	-25.193	29.779
Aug-06	32.891	56.102	0.000	-20.174	38.780
Sep-06	33.210	56.041	0.000	-21.515	35.762
Oct-06	43.626	55.963	0.000	-10.683	46.833
Nov-06	62.626	56.395	0.000	-2.962	44.799
Dec-06	64.864	56.583	0.000	4.275	57.094
Jan-07	71.315	56.537	0.000	15.740	73.183
Feb-07	96.825	56.418	0.000	42.954	101.763
Mar-07	166.503	57.853	0.000	78.074	107.213
Apr-07	74.856	58.086	0.000	11.822	65.262
May-07	35.744	58.003	0.000	-20.491	39.172
Jun-07	29.400	58.005	0.000	-28.656	29.302

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล PM₁₀ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์หอนุกรมเวลา (ต่อ)

Month-Year	Measured PM ₁₀	Base Level	Trend	Seasonal Factor	Forecasted PM ₁₀
Jul-07	30.283	57.944	0.000	-26.357	32.812
Aug-07	27.329	57.691	0.000	-24.978	37.770
Sep-07	35.437	57.673	0.000	-21.855	36.176
Oct-07	43.598	57.591	0.000	-12.244	46.990
Nov-07	51.021	57.504	0.000	-4.622	54.629
Dec-07	53.585	57.305	0.000	0.505	61.778
Jan-08	64.386	57.096	0.000	11.756	73.046
Feb-08	64.515	56.235	0.000	26.604	100.049
Mar-08	98.275	55.363	0.000	61.495	134.310
Apr-08	60.177	55.193	0.000	8.598	67.185
May-08	36.398	55.234	0.000	-19.711	34.702
Jun-08	34.287	55.421	0.000	-25.109	26.579
Jul-08	24.367	55.307	0.000	-28.518	29.064
Aug-08	22.491	55.117	0.000	-28.584	30.329
Sep-08	29.550	55.027	0.000	-23.563	33.263
Oct-08	31.303	54.749	0.000	-17.526	42.784
Nov-08	38.302	54.463	0.000	-10.063	50.127
Dec-08	50.374	54.352	0.000	-1.609	54.968
Jan-09	52.251	54.016	0.000	5.380	66.108
Feb-09	98.918	54.459	0.000	35.023	80.620
Mar-09	118.338	54.517	0.000	62.592	115.954
Apr-09	53.844	54.293	0.000	4.332	63.115
May-09	30.839	54.202	0.000	-21.433	34.582
Jun-09	25.018	54.103	0.000	-26.984	29.093
Jul-09	25.248	54.095	0.000	-28.673	25.586
Aug-09	28.096	54.158	0.000	-27.395	25.511
Sep-09	26.287	54.053	0.000	-25.545	30.595
Oct-09	33.681	53.985	0.000	-18.836	36.528
Nov-09	42.171	53.942	0.000	-10.868	43.922
Dec-09	55.981	54.030	0.000	0.070	52.333

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล PM₁₀ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์หอนุกรมเวลา (ต่อ)

Month-Year	Measured PM ₁₀	Base Level	Trend	Seasonal Factor	Forecasted PM ₁₀
Jan-10	47.687	53.747	0.000	-0.014	59.411
Feb-10	77.568	53.475	0.000	29.869	88.769
Mar-10	120.184	53.575	0.000	64.486	116.067
Apr-10	106.644	54.755	0.000	26.756	57.908
May-10	40.877	54.938	0.000	-17.957	33.322
Jun-10	29.621	54.978	0.000	-26.217	27.954
Jul-10	23.915	54.921	0.000	-29.773	26.305
Aug-10	23.047	54.812	0.000	-29.455	27.526
Sep-10	23.602	54.675	0.000	-28.152	29.267
Oct-10	24.666	54.404	0.000	-23.976	35.839
Nov-10	50.893	54.583	0.000	-7.483	43.536
Dec-10	40.694	54.245	0.000	-6.352	54.652

ค่าพยากรณ์ PM₁₀ ในปี 2011 ที่ได้จากการวิเคราะห์หอนุกรมเวลา และค่าที่ได้จากการตรวจวัด

Month-Year	Measured PM ₁₀	Forecasted PM ₁₀
Jan-11	50.360	54.231
Feb-11	54.594	84.114
Mar-11	48.597	118.730
Apr-11	51.320	81.001
May-11	36.535	36.288
Jun-11	22.658	28.0277
Jul-11	21.317	24.472
Aug-11	24.038	24.789
Sep-11	22.093	26.093
Oct-11	29.362	30.268
Nov-11	28.221	46.761
Dec-11	37.047	47.892

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล CO ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา

Month-Year	Measured CO	Base Level	Trend	Seasonal Factor	Forecasted CO
Jan-05	1.1840			0.304	
Feb-05	1.4563			0.576	
Mar-05	0.8771			-0.003	
Apr-05	0.4689			-0.411	
May-05	0.2806			-0.600	
Jun-05	0.3181			-0.562	
Jul-05	0.7707			-0.110	
Aug-05	1.0182			0.138	
Sep-05	0.8965			0.016	
Oct-05	1.0591			0.179	
Nov-05	1.1245			0.244	
Dec-05	1.1090	0.880	0.000	0.229	
Jan-06	1.3018	0.890	0.000	0.384	1.184
Feb-06	1.2516	0.873	0.000	0.431	1.466
Mar-06	1.3863	0.914	0.000	0.348	0.870
Apr-06	1.0751	0.959	0.000	-0.022	0.502
May-06	0.8955	1.001	0.000	-0.235	0.359
Jun-06	0.8412	1.033	0.000	-0.289	0.439
Jul-06	0.9307	1.034	0.000	-0.105	0.924
Aug-06	1.0409	1.023	0.000	0.049	1.172
Sep-06	0.7767	1.003	0.000	-0.163	1.040
Oct-06	0.8947	0.980	0.000	-0.016	1.182
Nov-06	0.9980	0.962	0.000	0.090	1.224
Dec-06	0.8248	0.933	0.000	-0.020	1.191
Jan-07	0.8807	0.898	0.000	0.087	1.317
Feb-07	0.8792	0.863	0.000	0.125	1.329
Mar-07	1.9893	0.924	0.000	0.877	1.211
Apr-07	0.9749	0.930	0.000	0.027	0.903
May-07	0.8604	0.943	0.000	-0.123	0.695
Jun-07	0.7376	0.950	0.000	-0.232	0.654

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล CO ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured CO	Base Level	Trend	Seasonal Factor	Forecasted CO
Jul-07	0.8259	0.948	0.000	-0.118	0.845
Aug-07	0.8290	0.935	0.000	-0.066	0.997
Sep-07	0.9186	0.947	0.000	-0.063	0.772
Oct-07	0.7507	0.932	0.000	-0.138	0.930
Nov-07	0.7984	0.915	0.000	-0.062	1.023
Dec-07	0.7717	0.905	0.000	-0.104	0.895
Jan-08	1.0468	0.909	0.000	0.124	0.992
Feb-08	1.1029	0.915	0.000	0.172	1.034
Mar-08	1.1575	0.864	0.000	0.446	1.792
Apr-08	0.8699	0.863	0.000	0.012	0.892
May-08	0.5600	0.848	0.000	-0.245	0.740
Jun-08	0.6366	0.850	0.000	-0.218	0.616
Jul-08	0.5815	0.838	0.000	-0.220	0.732
Aug-08	0.5897	0.824	0.000	-0.190	0.773
Sep-08	0.6124	0.812	0.000	-0.164	0.760
Oct-08	0.6986	0.814	0.000	-0.121	0.674
Nov-08	0.7739	0.816	0.000	-0.047	0.752
Dec-08	0.7915	0.822	0.000	-0.050	0.712
Jan-09	0.8746	0.816	0.000	0.076	0.946
Feb-09	1.2006	0.833	0.000	0.316	0.988
Mar-09	1.1924	0.826	0.000	0.387	1.279
Apr-09	0.5983	0.807	0.000	-0.151	0.839
May-09	0.4829	0.801	0.000	-0.299	0.562
Jun-09	0.5601	0.799	0.000	-0.234	0.583
Jul-09	0.6489	0.805	0.000	-0.173	0.579
Aug-09	0.5069	0.796	0.000	-0.263	0.615
Sep-09	0.5302	0.788	0.000	-0.233	0.632
Oct-09	0.5366	0.778	0.000	-0.210	0.667
Nov-09	0.6970	0.775	0.000	-0.070	0.731
Dec-09	0.7086	0.774	0.000	-0.061	0.726

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล CO ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์หอนุกรมเวลา (ต่อ)

Month-Year	Measured CO	Base Level	Trend	Seasonal Factor	Forecasted CO
Jan-10	0.6704	0.760	0.000	-0.046	0.849
Feb-10	1.1922	0.769	0.000	0.395	1.076
Mar-10	0.8994	0.748	0.000	0.213	1.156
Apr-10	0.8000	0.764	0.000	-0.013	0.597
May-10	0.5004	0.767	0.000	-0.275	0.466
Jun-10	0.4069	0.757	0.000	-0.320	0.534
Jul-10	0.4100	0.743	0.000	-0.291	0.585
Aug-10	0.4840	0.744	0.000	-0.261	0.480
Sep-10	0.5122	0.744	0.000	-0.232	0.510
Oct-10	0.6120	0.750	0.000	-0.157	0.534
Nov-10	0.8084	0.760	0.000	0.017	0.680
Dec-10	0.8338	0.771	0.000	0.031	0.699

ค่าพยากรณ์ CO ในปี 2011 ที่ได้จากการวิเคราะห์หอนุกรมเวลา และค่าที่ได้จากการตรวจวัด

Month-Year	Measured CO	Forecasted CO
Jan-11	0.8236	0.7248
Feb-11	0.8180	1.1662
Mar-11	0.8690	0.9835
Apr-11	0.7020	0.7576
May-11	0.5173	0.4956
Jun-11	0.4652	0.4511
Jul-11	0.4790	0.4796
Aug-11	0.5798	0.5102
Sep-11	0.6056	0.5388
Oct-11	0.5786	0.6141
Nov-11	0.7671	0.7881
Dec-11	0.7234	0.8014

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล NO₂ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา

Month-Year	Measured NO ₂	Base Level	Trend	Seasonal Factor	Forecasted NO ₂
Jan-05	27.4394			10.605	
Feb-05	33.3477			16.513	
Mar-05	19.1019			2.267	
Apr-05	18.2981			1.463	
May-05	12.0171			-4.818	
Jun-05	9.5573			-7.277	
Jul-05	10.0171			-6.818	
Aug-05	12.7723			-4.062	
Sep-05	13.3000			-3.535	
Oct-05	14.6096			-2.225	
Nov-05	14.8055			-2.029	
Dec-05	16.7511	16.835	0.000	-0.084	
Jan-06	23.3206	16.748	0.000	8.417	27.439
Feb-06	22.4098	16.518	0.000	10.750	33.261
Mar-06	24.7025	16.643	0.000	5.410	18.785
Apr-06	18.4041	16.650	0.000	1.621	18.107
May-06	13.2702	16.680	0.000	-4.054	11.832
Jun-06	10.2459	16.698	0.000	-6.830	9.403
Jul-06	10.2774	16.706	0.000	-6.607	9.880
Aug-06	11.1618	16.675	0.000	-4.850	12.644
Sep-06	9.7058	16.602	0.000	-5.359	13.140
Oct-06	10.4079	16.518	0.000	-4.333	14.377
Nov-06	15.3192	16.536	0.000	-1.588	14.489
Dec-06	13.1936	16.467	0.000	-1.814	16.452
Jan-07	18.8762	16.340	0.000	5.226	24.884
Feb-07	27.3595	16.346	0.000	10.893	27.090
Mar-07	31.6009	16.554	0.000	10.638	21.755
Apr-07	16.9585	16.528	0.000	0.975	18.175
May-07	11.2884	16.503	0.000	-4.684	12.474
Jun-07	11.9763	16.552	0.000	-5.607	9.674

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ข้อมูล NO₂ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล NO₂ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured NO ₂	Base Level	Trend	Seasonal Factor	Forecasted NO ₂
Jul-07	12.3090	16.602	0.000	-5.351	9.945
Aug-07	11.8715	16.604	0.000	-4.786	11.752
Sep-07	13.2862	16.647	0.000	-4.275	11.246
Oct-07	15.3019	16.711	0.000	-2.747	12.314
Nov-07	16.0059	16.729	0.000	-1.119	15.122
Dec-07	20.3577	16.844	0.000	1.076	14.915
Jan-08	28.3646	16.977	0.000	8.569	22.071
Feb-08	25.7813	16.933	0.000	9.784	27.871
Mar-08	30.4153	16.993	0.000	12.149	27.572
Apr-08	19.1985	17.019	0.000	1.628	17.968
May-08	13.2162	17.038	0.000	-4.216	12.336
Jun-08	10.9388	17.028	0.000	-5.868	11.431
Jul-08	10.0933	16.994	0.000	-6.192	11.676
Aug-08	10.8748	16.966	0.000	-5.494	12.208
Sep-08	11.3197	16.937	0.000	-5.003	12.691
Oct-08	12.0564	16.892	0.000	-3.880	14.190
Nov-08	15.2362	16.880	0.000	-1.404	15.773
Dec-08	20.4507	16.933	0.000	2.401	17.957
Jan-09	24.8803	16.920	0.000	8.239	25.502
Feb-09	33.4109	17.062	0.000	13.346	26.704
Mar-09	29.4626	17.067	0.000	12.283	29.210
Apr-09	16.0029	17.010	0.000	0.198	18.696
May-09	12.9579	17.014	0.000	-4.129	12.794
Jun-09	12.1263	17.034	0.000	-5.347	11.145
Jul-09	11.4340	17.047	0.000	-5.878	10.842
Aug-09	11.6362	17.049	0.000	-5.450	11.553
Sep-09	11.5887	17.039	0.000	-5.246	12.045
Oct-09	13.3975	17.044	0.000	-3.753	13.159
Nov-09	17.6052	17.086	0.000	-0.360	15.640
Dec-09	21.9283	17.137	0.000	3.698	19.486

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล NO₂ ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured NO ₂	Base Level	Trend	Seasonal Factor	Forecasted NO ₂
Jan-10	19.5288	17.014	0.000	5.133	25.376
Feb-10	31.6238	17.040	0.000	14.017	30.359
Mar-10	31.5401	17.087	0.000	13.460	29.323
Apr-10	26.8227	17.289	0.000	5.263	17.286
May-10	12.6864	17.279	0.000	-4.380	13.160
Jun-10	8.7516	17.212	0.000	-7.036	11.932
Jul-10	9.3867	17.170	0.000	-6.912	11.334
Aug-10	10.2189	17.139	0.000	-6.247	11.721
Sep-10	13.4531	17.172	0.000	-4.417	11.893
Oct-10	12.9151	17.161	0.000	-4.021	13.418
Nov-10	17.6266	17.179	0.000	0.078	16.801
Dec-10	18.2809	17.124	0.000	2.319	20.876

ค่าพยากรณ์ NO₂ ในปี 2011 ที่ได้จากการวิเคราะห์อนุกรมเวลา และค่าที่ได้จากการตรวจวัด

Month-Year	Measured NO ₂	Forecasted NO ₂
Jan-11	21.3211	22.2570
Feb-11	24.8760	31.1409
Mar-11	18.9423	30.5837
Apr-11	16.5685	22.3870
May-11	12.5198	12.7432
Jun-11	11.0948	10.0875
Jul-11	8.9641	10.2117
Aug-11	8.5413	10.8763
Sep-11	8.0988	12.7064
Oct-11	10.7618	13.1030
Nov-11	15.3804	17.2020
Dec-11	17.6930	19.4430

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา

Month-Year	Measured RH	Base Level	Trend	Seasonal Factor	Forecasted RH
Jan-05	59.5941			-5.807	
Feb-05	41.2485			-24.153	
Mar-05	45.0996			-20.302	
Apr-05	53.4155			-11.986	
May-05	71.2356			5.834	
Jun-05	87.0675			21.666	
Jul-05	72.9965			7.595	
Aug-05	75.4724			10.071	
Sep-05	75.7595			10.358	
Oct-05	68.9420			3.541	
Nov-05	68.4250			3.024	
Dec-05	65.5585	65.401	0.000	0.157	
Jan-06	60.5787	65.847	0.000	-5.268	59.594
Feb-06	51.3055	70.193	0.000	-18.888	41.694
Mar-06	45.4024	68.163	0.000	-22.761	49.891
Apr-06	57.8755	68.931	0.000	-11.056	56.177
May-06	69.5459	66.571	0.000	2.975	74.765
Jun-06	69.4000	58.052	0.000	11.348	88.237
Jul-06	73.4906	61.599	0.000	11.892	65.648
Aug-06	74.9596	63.087	0.000	11.873	71.670
Sep-06	72.3833	62.607	0.000	9.777	73.445
Oct-06	68.3513	63.603	0.000	4.748	66.147
Nov-06	61.6986	61.375	0.000	0.324	66.627
Dec-06	59.4333	60.426	0.000	-0.992	61.532
Jan-07	56.3193	60.951	0.000	-4.632	55.158
Feb-07	46.5380	62.974	0.000	-16.436	42.063
Mar-07	41.8595	63.719	0.000	-21.859	40.214
Apr-07	49.4733	62.276	0.000	-12.803	52.663
May-07	73.5972	66.050	0.000	7.547	65.251
Jun-07	66.4824	61.114	0.000	5.368	77.398

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ข้อมูลความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured RH	Base Level	Trend	Seasonal Factor	Forecasted RH
Jul-07	67.8639	58.789	0.000	9.075	73.006
Aug-07	69.4159	58.225	0.000	11.190	70.662
Sep-07	69.4069	58.861	0.000	10.546	68.002
Oct-07	68.0834	60.884	0.000	7.199	63.609
Nov-07	65.6801	62.906	0.000	2.774	61.208
Dec-07	60.8466	62.424	0.000	-1.577	61.914
Jan-08	59.7769	63.321	0.000	-3.544	57.792
Feb-08	52.9134	66.047	0.000	-13.134	46.885
Mar-08	46.9874	67.313	0.000	-20.326	44.188
Apr-08	52.2179	66.277	0.000	-14.059	54.510
May-08	69.9283	64.515	0.000	5.413	73.824
Jun-08	69.6412	64.406	0.000	5.236	69.883
Jul-08	68.9748	62.368	0.000	6.607	73.480
Aug-08	73.5087	62.346	0.000	11.163	73.558
Sep-08	78.5599	64.909	0.000	13.651	72.892
Oct-08	84.6895	70.598	0.000	14.091	72.108
Nov-08	78.3111	72.832	0.000	5.479	73.372
Dec-08	81.1438	77.304	0.000	3.840	71.255
Jan-09	81.0363	80.594	0.000	0.442	73.759
Feb-09	76.4182	84.645	0.000	-8.227	67.460
Mar-09	60.5484	82.940	0.000	-22.391	64.319
Apr-09	60.5424	79.169	0.000	-18.627	68.881
May-09	72.9664	73.916	0.000	-0.950	84.582
Jun-09	76.5236	72.728	0.000	3.796	79.152
Jul-09	76.4530	71.425	0.000	5.028	79.334
Aug-09	76.1573	68.517	0.000	7.641	82.588
Sep-09	76.7806	66.081	0.000	10.700	82.168
Oct-09	76.1346	64.255	0.000	11.880	80.172
Nov-09	65.9181	62.529	0.000	3.389	69.734
Dec-09	66.1344	62.423	0.000	3.711	66.369

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured RH	Base Level	Trend	Seasonal Factor	Forecasted RH
Jan-10	65.7473	63.726	0.000	2.021	62.865
Feb-10	49.3955	60.966	0.000	-11.571	55.499
Mar-10	51.3011	66.721	0.000	-15.420	38.575
Apr-10	44.6625	65.169	0.000	-20.507	48.095
May-10	58.6205	62.637	0.000	-4.017	64.219
Jun-10	68.2651	63.466	0.000	4.799	66.433
Jul-10	75.2429	66.518	0.000	8.725	68.494
Aug-10	81.5886	69.878	0.000	11.711	74.158
Sep-10	82.1167	70.574	0.000	11.543	80.578
Oct-10	76.0511	67.678	0.000	8.373	82.453
Nov-10	67.1899	65.925	0.000	1.265	71.067
Dec-10	70.8306	66.465	0.000	4.366	69.636

ค่าพยากรณ์ความชื้นสัมพัทธ์ ในปี 2011 ที่ได้จากการวิเคราะห์อนุกรมเวลา และค่าที่ได้จากการตรวจวัด

Month-Year	Measured RH	Forecasted RH
Jan-11	66.9328	68.4859
Feb-11	54.2545	54.8944
Mar-11	65.2527	51.0450
Apr-11	69.3444	45.9584
May-11	78.9973	62.4483
Jun-11	77.8917	71.2645
Jul-11	76.2782	75.1904
Aug-11	81.3992	78.1759
Sep-11	81.2569	78.0081
Oct-11	79.7618	74.8377
Nov-11	68.9792	67.7299
Dec-11	70.4073	70.8306

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับราชการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ทำในเชิงพาณิชย์ด้านการค้า
ข้อมูลความชื้นสัมพัทธ์ที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลปริมาณน้ำฝนที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา

Month-Year	Measured Rain	Base Level	Trend	Seasonal Factor	Forecasted Rain
Jan-05	0.000			-0.158	
Feb-05	0.000			-0.158	
Mar-05	0.032			-0.126	
Apr-05	0.061			-0.096	
May-05	0.089			-0.068	
Jun-05	0.343			0.185	
Jul-05	0.389			0.232	
Aug-05	0.217			0.060	
Sep-05	0.391			0.234	
Oct-05	0.288			0.131	
Nov-05	0.020			-0.138	
Dec-05	0.059	0.158	0.000	-0.098	
Jan-06	0.001	0.158	0.000	-0.157	0.000
Feb-06	0.001	0.158	0.000	-0.157	0.000
Mar-06	0.015	0.157	0.000	-0.135	0.032
Apr-06	0.173	0.160	0.000	-0.036	0.061
May-06	0.252	0.163	0.000	0.017	0.091
Jun-06	0.219	0.160	0.000	0.117	0.348
Jul-06	0.312	0.159	0.000	0.189	0.392
Aug-06	0.371	0.162	0.000	0.141	0.218
Sep-06	0.167	0.157	0.000	0.112	0.396
Oct-06	0.093	0.153	0.000	0.027	0.288
Nov-06	0.001	0.153	0.000	-0.145	0.015
Dec-06	0.001	0.151	0.000	-0.127	0.054
Jan-07	0.000	0.152	0.000	-0.154	-0.006
Feb-07	0.001	0.152	0.000	-0.153	-0.005
Mar-07	0.003	0.151	0.000	-0.142	0.017
Apr-07	0.070	0.150	0.000	-0.060	0.115
May-07	0.403	0.155	0.000	0.142	0.168
Jun-07	0.188	0.154	0.000	0.072	0.272

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลปริมาณน้ำฝนที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured Rain	Base Level	Trend	Seasonal Factor	Forecasted Rain
Jul-07	0.148	0.150	0.000	0.086	0.343
Aug-07	0.254	0.149	0.000	0.121	0.290
Sep-07	0.251	0.149	0.000	0.107	0.261
Oct-07	0.087	0.147	0.000	-0.020	0.176
Nov-07	0.075	0.148	0.000	-0.106	0.002
Dec-07	0.005	0.148	0.000	-0.136	0.022
Jan-08	0.008	0.148	0.000	-0.146	-0.006
Feb-08	0.034	0.149	0.000	-0.132	-0.005
Mar-08	0.009	0.149	0.000	-0.141	0.007
Apr-08	0.064	0.149	0.000	-0.073	0.089
May-08	0.192	0.146	0.000	0.090	0.291
Jun-08	0.190	0.146	0.000	0.057	0.218
Jul-08	0.181	0.145	0.000	0.059	0.231
Aug-08	0.190	0.143	0.000	0.081	0.266
Sep-08	0.284	0.144	0.000	0.125	0.250
Oct-08	0.262	0.147	0.000	0.053	0.124
Nov-08	0.098	0.148	0.000	-0.075	0.041
Dec-08	0.009	0.148	0.000	-0.138	0.012
Jan-09	0.000	0.148	0.000	-0.147	0.002
Feb-09	0.001	0.148	0.000	-0.140	0.016
Mar-09	0.038	0.148	0.000	-0.124	0.007
Apr-09	0.083	0.148	0.000	-0.069	0.075
May-09	0.279	0.149	0.000	0.111	0.238
Jun-09	0.190	0.149	0.000	0.048	0.206
Jul-09	0.168	0.148	0.000	0.038	0.208
Aug-09	0.170	0.147	0.000	0.049	0.229
Sep-09	0.156	0.144	0.000	0.064	0.272
Oct-09	0.133	0.143	0.000	0.019	0.198
Nov-09	0.001	0.142	0.000	-0.111	0.068
Dec-09	0.017	0.142	0.000	-0.131	0.004

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลปริมาณน้ำฝนที่ได้จากการตรวจวัดและค่าพยากรณ์จากการวิเคราะห์อนุกรมเวลา (ต่อ)

Month-Year	Measured Rain	Base Level	Trend	Seasonal Factor	Forecasted Rain
Jan-10	0.024	0.143	0.000	-0.132	-0.005
Feb-10	0.051	0.144	0.000	-0.114	0.002
Mar-10	0.005	0.143	0.000	-0.132	0.019
Apr-10	0.010	0.142	0.000	-0.103	0.074
May-10	0.047	0.138	0.000	0.002	0.253
Jun-10	0.088	0.135	0.000	-0.004	0.186
Jul-10	0.217	0.136	0.000	0.061	0.173
Aug-10	0.574	0.145	0.000	0.256	0.186
Sep-10	0.332	0.147	0.000	0.129	0.208
Oct-10	0.249	0.149	0.000	0.063	0.166
Nov-10	0.000	0.148	0.000	-0.131	0.038
Dec-10	0.009	0.148	0.000	-0.135	0.017

ค่าพยากรณ์ปริมาณน้ำฝน ในปี 2011ที่ได้จากการวิเคราะห์อนุกรมเวลา
และค่าที่ได้จากการตรวจวัด

Month-Year	Measured Rain	Forecasted Rain
Jan-11	0.010	0.016
Feb-11	0.001	0.034
Mar-11	0.084	0.016
Apr-11	0.104	0.045
May-11	0.495	0.150
Jun-11	0.299	0.144
Jul-11	0.218	0.209
Aug-11	0.440	0.404
Sep-11	0.358	0.277
Oct-11	0.132	0.211
Nov-11	0.001	0.017
Dec-11	0.003	0.013

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก ค

ข้อมูลจากการตรวจวัดและค่าพยากรณ์ที่ได้จากแบบจำลองการวิเคราะห์
การถดถอยพหุคูณโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล PM₁₀ ที่ได้จากการตรวจวัดและค่าที่ได้จากตัวแบบการวิเคราะห์การถดถอยพหุคูณ
โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

Month-Year	CO	NO ₂	RH	Rain	Pressure	NO	SO ₂	PM ₁₀	PM ₁₀ (GA)
Jan-05	1.18	27.44	59.59	0.00	734.56	11.25	20.26	77.87	89.59
Feb-05	1.46	33.35	41.25	0.00	733.40	13.83	23.81	128.17	118.66
Mar-05	0.88	19.10	45.10	0.03	733.42	7.46	26.98	93.82	76.13
Apr-05	0.47	18.30	53.42	0.06	732.14	5.01	27.53	62.83	68.90
May-05	0.28	12.02	71.24	0.09	729.89	2.97	24.02	37.23	38.48
Jun-05	0.32	9.56	87.07	0.34	728.56	3.75	11.31	24.68	17.06
Jul-05	0.77	10.02	73.00	0.39	728.97	5.63	9.67	30.38	26.92
Aug-05	1.02	12.77	75.47	0.22	727.88	9.96	7.72	39.32	31.61
Sep-05	0.90	13.30	75.76	0.39	730.83	7.15	13.00	36.44	34.90
Oct-05	1.06	14.61	68.94	0.29	733.97	6.37	15.34	47.58	45.53
Nov-05	1.12	14.81	68.43	0.02	733.90	9.84	9.36	45.62	44.49
Dec-05	1.11	16.75	65.56	0.06	734.51	9.44	12.51	57.48	52.83
Jan-06	1.30	23.32	60.58	0.00	733.39	10.37	18.71	68.12	77.11
Feb-06	1.25	22.41	51.31	0.00	733.86	7.57	19.03	71.08	82.50
Mar-06	1.39	24.70	45.40	0.01	731.75	5.85	25.64	122.12	95.84
Apr-06	1.08	18.40	57.88	0.17	731.19	4.51	22.43	64.91	67.28
May-06	0.90	13.27	69.55	0.25	731.09	3.95	18.97	37.81	43.02
Jun-06	0.84	10.25	69.40	0.22	730.16	3.52	15.15	31.30	33.63
Jul-06	0.93	10.28	73.49	0.31	728.83	5.20	12.20	32.41	29.00
Aug-06	1.04	11.16	74.96	0.37	729.21	7.76	10.35	32.89	28.46
Sep-06	0.78	9.71	72.38	0.17	731.86	6.95	10.70	33.21	26.83
Oct-06	0.89	10.41	68.35	0.09	734.03	4.65	10.26	43.63	34.32
Nov-06	1.00	15.32	61.70	0.00	734.02	6.14	12.24	62.63	53.54
Dec-06	0.82	13.19	59.43	0.00	734.96	7.62	10.43	64.86	46.54
Jan-07	0.88	18.88	56.32	0.00	734.74	9.26	14.53	71.31	65.35
Feb-07	0.88	27.36	46.54	0.00	733.73	10.41	20.57	96.83	97.48
Mar-07	1.99	31.60	41.86	0.00	732.30	11.52	31.08	166.50	117.32
Apr-07	0.97	16.96	49.47	0.07	733.35	3.00	30.84	74.86	71.00
May-07	0.86	11.29	73.60	0.40	731.21	2.63	16.14	35.74	34.11
Jun-07	0.74	11.98	66.48	0.19	728.00	4.82	13.99	29.40	39.32

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล PM₁₀ ที่ได้จากการตรวจวัดและค่าที่ได้จากตัวแบบการวิเคราะห์การถดถอยพหุคูณ
โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (ต่อ)

Month-Year	CO	NO ₂	RH	Rain	Pressure	NO	SO ₂	PM ₁₀	PM ₁₀ (GA)
Jul-07	0.83	12.31	67.86	0.15	727.60	7.50	10.19	30.28	37.21
Aug-07	0.83	11.87	69.42	0.25	727.23	5.86	9.03	27.33	35.47
Sep-07	0.92	13.29	69.41	0.25	728.55	6.08	10.51	35.44	40.20
Oct-07	0.75	15.30	68.08	0.09	731.14	7.17	9.93	43.60	46.82
Nov-07	0.80	16.01	65.68	0.07	732.71	9.87	12.94	51.02	49.01
Dec-07	0.77	20.36	60.85	0.00	732.51	11.06	16.02	53.58	65.21
Jan-08	1.05	28.36	59.78	0.01	732.27	12.10	19.20	64.39	90.87
Feb-08	1.10	25.78	52.91	0.03	731.74	11.63	25.22	64.51	88.54
Mar-08	1.16	30.42	46.99	0.01	730.85	7.40	33.62	98.28	111.42
Apr-08	0.87	19.20	52.22	0.06	729.64	3.85	35.77	60.18	75.55
May-08	0.56	13.22	69.93	0.19	727.60	4.41	17.92	36.40	41.17
Jun-08	0.64	10.94	69.64	0.19	728.16	6.35	15.32	34.29	32.70
Jul-08	0.58	10.09	68.97	0.18	727.72	6.21	11.27	24.37	29.96
Aug-08	0.59	10.87	73.51	0.19	728.38	7.69	11.38	22.49	28.26
Sep-08	0.61	11.32	78.56	0.28	729.30	8.52	10.58	29.55	25.20
Oct-08	0.70	12.06	84.69	0.26	731.47	8.57	10.31	31.30	23.96
Nov-08	0.77	15.24	78.31	0.10	732.69	12.03	11.77	38.30	36.39
Dec-08	0.79	20.45	81.14	0.01	733.00	13.34	15.00	50.37	50.49
Jan-09	0.87	24.88	81.04	0.00	733.17	15.54	1.37	52.25	60.86
Feb-09	1.20	33.41	76.42	0.00	730.33	15.12	1.13	98.92	91.18
Mar-09	1.19	29.46	60.55	0.04	730.08	10.39	0.67	118.34	92.84
Apr-09	0.60	16.00	60.54	0.08	729.19	5.34	0.52	53.84	53.52
May-09	0.48	12.96	72.97	0.28	728.51	4.24	0.07	30.84	35.54
Jun-09	0.56	12.13	76.52	0.19	727.10	5.44	0.16	25.02	30.36
Jul-09	0.65	11.43	76.45	0.17	727.04	7.07	0.25	25.25	27.34
Aug-09	0.51	11.64	76.16	0.17	727.98	8.04	0.29	28.10	26.97
Sep-09	0.53	11.59	76.78	0.16	728.66	7.31	0.30	26.29	27.21
Oct-09	0.54	13.40	76.13	0.13	730.50	8.28	0.30	33.68	32.64
Nov-09	0.70	17.61	65.92	0.00	732.16	10.85	0.52	42.17	51.38
Dec-09	0.71	21.93	66.13	0.02	732.10	14.43	0.49	55.98	61.51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล PM₁₀ ที่ได้จากการตรวจวัดและค่าที่ได้จากตัวแบบการวิเคราะห์การถดถอยพหุคูณ
โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (ต่อ)

Month-Year	CO	NO ₂	RH	Rain	Pressure	NO	SO ₂	PM ₁₀	PM ₁₀ (GA)
Jan-10	0.67	19.53	65.75	0.02	732.34	13.40	0.46	47.69	55.09
Feb-10	1.19	31.62	49.40	0.05	730.64	20.77	0.87	77.57	98.11
Mar-10	0.90	31.54	51.30	0.00	730.52	9.81	0.91	120.18	104.99
Apr-10	0.80	26.82	44.66	0.01	729.38	7.68	1.58	106.64	96.21
May-10	0.50	12.69	58.62	0.05	727.62	4.84	0.42	40.88	44.76
Jun-10	0.41	8.75	68.27	0.09	727.66	6.14	0.18	29.62	24.86
Jul-10	0.41	9.39	75.24	0.22	727.48	7.56	0.56	23.92	20.49
Aug-10	0.48	10.22	81.59	0.57	727.37	10.51	0.90	23.05	14.84
Sep-10	0.51	13.45	82.12	0.33	728.61	8.79	0.59	23.60	27.30
Oct-10	0.61	12.92	76.05	0.25	729.44	6.57	0.33	24.67	32.16
Nov-10	0.81	17.63	67.19	0.00	731.35	9.97	1.33	50.89	51.75
Dec-10	0.83	18.28	70.83	0.01	730.01	12.61	1.20	40.69	49.18

ค่าพยากรณ์ PM₁₀ ในปี 2011 ที่ได้จากแบบจำลองการวิเคราะห์การถดถอยพหุคูณ
โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม และค่าที่ได้จากการตรวจวัด

Month-Year	CO	NO ₂	RH	Rain	Pressure	NO	SO ₂	PM ₁₀	PM ₁₀ (GA)
Jan-11	0.99	23.73	63.84	0.01	733.41	11.99	12.42	50.36	73.15
Feb-11	1.18	28.99	52.97	0.01	732.28	13.22	15.11	54.59	96.08
Mar-11	1.25	27.80	48.53	0.02	731.49	8.74	19.82	48.60	99.76
Apr-11	0.80	19.28	53.03	0.08	730.82	4.90	19.78	51.32	72.08
May-11	0.60	12.57	69.32	0.21	729.32	3.84	12.92	36.53	39.51
Jun-11	0.58	10.60	72.90	0.20	728.27	5.00	9.35	22.66	29.65
Jul-11	0.69	10.59	72.50	0.24	727.94	6.53	7.36	21.32	28.49
Aug-11	0.74	11.42	75.18	0.30	728.01	8.30	6.61	24.04	27.60
Sep-11	0.71	12.11	75.83	0.26	729.63	7.47	7.61	22.09	30.27
Oct-11	0.76	13.11	73.71	0.19	731.76	6.94	7.75	29.36	35.91
Nov-11	0.87	16.10	67.87	0.03	732.81	9.78	8.02	28.22	47.76
Dec-11	0.84	18.49	67.32	0.02	732.85	11.42	9.27	37.05	54.29

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ ห้ามมิให้ใช้ในงานที่ออกจำหน่ายโดยไม่ขออนุญาต ทั้งนี้ภายใต้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อนุกรมเวลาเพื่อพยากรณ์ปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควัน
ในพื้นที่จังหวัดเชียงใหม่

ขวัญศิริ สิริมงคล^{1*} และ บุญยมาศ พิมพ์พรรณชาติ²

หน่วยวิจัยคณิตศาสตร์อุตสาหกรรมและสาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีเจ้าคุณทหารลาดกระบัง ถนนคลองกรุง เขตลาดกระบัง กรุงเทพมหานคร 10520 ประเทศไทย
¹Khwansiri.1@gmail.com, ²knbusaya@kmitl.ac.th

บทคัดย่อ : หมอกควันในภาคเหนือของประเทศไทย เป็นปัญหามลพิษที่เกิดขึ้นเป็นประจำทุกปี เริ่มตั้งแต่เดือนพฤศจิกายนและมีปริมาณสูงสุดในเดือนมีนาคม สาเหตุหลักเกิดจากไฟป่า ฝุ่นละอองจากถนน การก่อสร้าง และเขม่าจากน้ำมันเชื้อเพลิง ทำให้คุณภาพอากาศแย่ง ประกอบกับสภาพภูมิประเทศซึ่งมีภูเขาล้อมรอบ ทำให้มลพิษต่าง ๆ ถูกกักไว้และแผ่ปกคลุมทั่วเมือง งานวิจัยนี้ได้ทำการศึกษาและพยากรณ์ปัจจัยที่มีผลต่อการเกิดหมอกควันในอากาศของจังหวัดเชียงใหม่โดยวิธีของวินเดอร์ ซึ่งอาศัยข้อมูลค่าเฉลี่ยรายเดือนจากปัจจัยที่มีผลต่อการเกิดปัญหาหมอกควันในช่วงระหว่างปี ค.ศ. 2005 – 2010 โดยปัจจัยหลักที่มีผลต่อการเกิดปัญหาหมอกควันที่ได้จากการวิเคราะห์สหสัมพันธ์ได้แก่ ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) ก๊าซคาร์บอนมอนนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO_2) ความชื้นสัมพัทธ์ (Relative humidity : RH) และปริมาณน้ำฝน (Rain) เพื่อดูแนวโน้มและพยากรณ์ปริมาณปัจจัยดังกล่าวในปี 2011 โดยหากข้อมูลไม่มีเหตุการณ์ที่ผิดปกติจะทำให้ค่าคลาดเคลื่อนที่ได้จากตัวแบบมีค่าน้อย แต่ถ้าข้อมูลมีเหตุการณ์ผิดปกติจะทำให้เกิดค่าคลาดเคลื่อนมาก และค่าพยากรณ์ในแต่ละเดือนของปี 2011 มีแนวโน้มไม่แตกต่างจากค่าพยากรณ์ในแต่ละเดือนของปีก่อนมากนัก

คำสำคัญ : อนุกรมเวลา; ปัญหาหมอกควัน; วิธีของวินเดอร์; การพยากรณ์

*Corresponding author e-mail : Khwansiri.1@gmail.com

1 บทนำ

ปัญหาหมอกควันในภาคเหนือตอนบนของประเทศไทยถือเป็นปัญหาสำคัญ ความรุนแรงของปัญหานี้โดยทั่วไปปรากฏชัดเจนในช่วงหน้าแล้ง ซึ่งมักพบการเพิ่มสูงขึ้นของปริมาณฝุ่นละอองขนาดเล็ก เนื่องจากความแห้งแล้งส่งผลให้เกิดการเพิ่มขึ้นของไฟป่า และในช่วงเวลาดังกล่าว เกษตรกรจะทำการเผาเศษวัสดุเพื่อเตรียมพื้นที่สำหรับทำการเกษตรในช่วงฤดูฝน ประกอบกับสภาพภูมิประเทศซึ่งมีภูเขาสูงล้อมรอบ ทำให้มลพิษต่าง ๆ ถูกกักไว้และแผ่ปกคลุมทั่วเมือง ซึ่งสภาวะอากาศที่แห้งและนิ่งทำให้ฝุ่นละอองที่เกิดขึ้นสามารถแขวนลอยอยู่ในบรรยากาศได้นาน ประกอบกับฝนตกน้อยทำให้การชะล้างหมอกควันหรือฝุ่นที่แขวนลอยอยู่ในอากาศเป็นไปได้น้อย ทั้งนี้ ค่ามาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไปในประเทศไทยที่ทางกรมควบคุมมลพิษได้กำหนดไว้คือ ค่าเฉลี่ยปริมาณฝุ่นละอองขนาดเล็กขนาดไม่เกิน 10 ไมครอน ซึ่งตามมาตรฐานได้กำหนดไว้ว่าค่าเฉลี่ยในเวลา 24 ชั่วโมง จะต้องไม่เกิน 120 ไมโครกรัมต่อลูกบาศก์เมตร และค่าเฉลี่ยในเวลา 1 ปี จะต้องไม่เกิน 50 ไมโครกรัมต่อลูกบาศก์เมตร ซึ่งปัญหานี้ส่งผลกระทบต่อสภาพบรรยากาศทั่วไป ต่อสุขภาพอนามัย และต่อธุรกิจการท่องเที่ยวและบริการ จากปัญหาข้างต้นทำให้เราจำเป็นต้องหาวิธีที่จะต้องทราบแนวโน้มของปัจจัยที่มีผลต่อการเกิดหมอกควัน โดยอาศัยการวิเคราะห์หอนุกรมเวลา มาใช้ในการพยากรณ์ค่าของปริมาณปัจจัยต่างๆที่มีผลต่อการเกิดปัญหาหมอกควัน

2 วิธีการดำเนินงาน

จากข้อมูลของกรมควบคุมมลพิษประกอบด้วย 12 ปัจจัย ได้แก่ ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) ก๊าซคาร์บอนมอนนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO_2) ความชื้นสัมพัทธ์ (Relative humidity : RH) ความดัน (Pressure) ปริมาณน้ำฝน (Rain) ก๊าซไนโตรเจนมอนอกไซด์ (NO) ทิศทางลม (Wind direction) ความเร็วลม (Wind Speed) อุณหภูมิ (Temperature) ก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) และก๊าซโอโซน (O_3)

งานวิจัยนี้ได้ทำการศึกษาปัจจัยที่เกี่ยวข้องกับการเกิดหมอกควัน โดยใช้สัมประสิทธิ์สหสัมพันธ์มาช่วยในการตัดสินใจเพื่อหาปัจจัยที่มีอิทธิพลต่อการเกิดหมอกควัน นั่นคือหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยต่างๆกับค่า PM_{10} เนื่องจาก PM_{10} เป็นตัวแปรสำคัญที่ใช้บ่งชี้ว่าสถานการณ์หมอกควันในพื้นที่ว่ามีความรุนแรงมากน้อยเพียงใด

2.1 ปัจจัยที่เกี่ยวข้องกับการเกิดหมอกควัน

การวิเคราะห์สหสัมพันธ์ (Correlation Analysis) เป็นวิธีการที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว โดยสัมประสิทธิ์สหสัมพันธ์ (r) มีค่าตั้งแต่ -1 ถึง 1 ค่า r เป็นลบแสดงความสัมพันธ์ในทิศทางตรงกันข้าม ค่า r เป็นบวกแสดงความสัมพันธ์ในทิศทางเดียวกัน

- r = 0.50 ถึง 1.00 หรือ r = -0.50 ถึง -1.00 ถือว่าข้อมูลมีความสัมพันธ์กันในระดับสูง
- r = 0.30 ถึง 0.49 หรือ r = -0.30 ถึง -0.49 ถือว่าข้อมูลมีความสัมพันธ์กันในระดับปานกลาง
- r = 0.10 ถึง 0.29 หรือ r = -0.10 ถึง -0.29 ถือว่าข้อมูลมีความสัมพันธ์กันในระดับต่ำ
- r = 0.00 ถือว่าข้อมูลไม่มีความสัมพันธ์กัน

ทำการวิเคราะห์ว่าแต่ละปัจจัยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่าไร และมีความสัมพันธ์กับค่า PM₁₀ อยู่ในระดับใด ดังแสดงในตารางต่อไปนี้

ตารางที่ 1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง PM₁₀ และปัจจัยทางอุตุนิยมวิทยา

Factor	CO	NO ₂	RH	Rain	Pressure	NO	Wind dir	Wind spe	Temp	SO ₂	O ₃
r	0.74	0.87	-0.80	-0.62	0.49	0.35	-0.10	0.08	-0.02	0.40	0.15

ปัจจัยที่มีความสัมพันธ์กับค่า PM₁₀ สูง ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ ก๊าซไนโตรเจนไดออกไซด์ ความชื้นสัมพัทธ์ และปริมาณน้ำฝน

ปัจจัยที่มีความสัมพันธ์กับค่า PM₁₀ ปานกลาง ได้แก่ ความดัน ก๊าซไนโตรเจนมอนนอกไซด์ และก๊าซซัลเฟอร์ไดออกไซด์

ปัจจัยที่มีความสัมพันธ์กับค่า PM₁₀ ต่ำ ได้แก่ ความเร็วลม ทิศทางลม อุณหภูมิ และก๊าซโอโซน ดังนั้นปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในงานวิจัยนี้จะพิจารณาจากปัจจัยที่มีความสัมพันธ์กับ PM₁₀ สูงทั้ง 4 ค่า ได้แก่ ก๊าซคาร์บอนมอนนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO₂) ความชื้นสัมพัทธ์ (RH) และปริมาณน้ำฝน (Rain)

ซึ่งต่อไปเราจะทำการพยากรณ์ข้อมูลอนุกรมเวลาของปัจจัยดังกล่าวนี้ โดยใช้การวิเคราะห์อนุกรมเวลา

2.2 ข้อมูลอนุกรมเวลา

ข้อมูลอนุกรมเวลาเป็นข้อมูลที่เก็บต่อเนื่องตั้งแต่ต้นจนถึงสิ้นสุดระยะเวลาที่ระบุ เพื่อนำมาใช้ในการพยากรณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยอนุกรมเวลาหมายถึงค่าของข้อมูลหรือค่าสังเกตที่เปลี่ยนไปตามลำดับของเวลาที่เกิดขึ้น การวิเคราะห์ลักษณะของข้อมูลดังกล่าวจะเรียกว่าการวิเคราะห์อนุกรมเวลา (Analysis of Time Series) วัตถุประสงค์ที่สำคัญของการวิเคราะห์อนุกรมเวลาก็คือเพื่อหารูปแบบการเปลี่ยนแปลงของตัวแปรที่เราสนใจ

ส่วนประกอบของอนุกรมเวลา

1. ค่าแนวโน้ม (Long Term Trend: T) ค่าแนวโน้มเป็นการแสดงถึงการเคลื่อนไหวหรือเปลี่ยนแปลงของข้อมูลในระยะยาว เช่น ปริมาณการใช้ไฟฟ้าของประเทศไทย ปริมาณการนำเข้าน้ำมันดิบ เป็นต้น

2. ค่าการผันแปรตามฤดูกาล (Seasonal Variation: S) หมายถึงการเปลี่ยนแปลงตามฤดูกาล โดยเกิดขึ้นซ้ำๆ กันในรอบ 1 ปี จนกลายเป็นแบบแผนเดียวกัน เช่นผลผลิตข้าวจะสูงในช่วงไตรมาสแรกของปี

3. ค่าการผันแปรตามวัฏจักร (Cyclical Variation: C) หมายถึงการเคลื่อนไหวที่เป็นไปตามวัฏจักรซึ่งการเคลื่อนไหวตามวัฏจักรนี้จะมีลักษณะคล้ายกับการผันแปรตามฤดูกาล แต่จะมีระยะเวลาที่ยาวนานกว่า

4. การผันแปรเนื่องจากเหตุการณ์ไม่ปกติ (Irregular Variation: I) การผันแปรชนิดนี้ไม่แน่นอน ไม่สามารถคาดการณ์ได้ล่วงหน้า เช่น ภัยธรรมชาติ สงคราม การนัดหยุดงาน เป็นต้น

การได้รับอิทธิพลของปัจจัยที่เป็นส่วนประกอบของอนุกรมเวลาอาจมีทั้ง 4 ปัจจัยหรือมีเพียงปัจจัยใดปัจจัยหนึ่งเท่านั้น สำหรับข้อมูลที่ได้ศึกษาในครั้งนี้พบว่า เป็นข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงตามฤดูกาลเชิงบวก ซึ่งเป็นลักษณะข้อมูลที่ไม่มีแนวโน้ม

2.3 ตัวแบบพยากรณ์

การพยากรณ์อนุกรมเวลาที่มีการเปลี่ยนแปลงตามฤดูกาลเชิงบวกโดยวิธีของวินเตอร์ (Holt-Winter's method for additive seasonal effects) อธิบายได้ด้วยตัวแบบ

$$Z_t = \beta_0 + \beta_1 t + S_t + a_t$$

เมื่อ β_0 แสดงระดับของแนวโน้ม

β_1 แทนความชันของแนวโน้มเชิงเส้น

S_t แทนการแปรผันตามฤดูกาล

a_t เป็นตัวแปรสุ่มอิสระที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวน σ_a^2

การพยากรณ์ตามตัวแบบนี้สามารถปรับค่าประมาณได้ทุกหน่วยเวลาจากค่าประมาณเริ่มต้นเมื่อเวลา t_0 และจากค่าคงที่ในการทำให้เรียบที่เลือกอย่างเหมาะสม ซึ่งสามารถอธิบายการปรับค่าประมาณของ β_0 β_1 และ S_t สำหรับเวลา n ใดๆ ได้ดังนี้

1. เมื่อมีค่าสังเกต Z_n ที่เวลา n ใดๆ ค่าประมาณของ β_0 ปรับค่าได้ในรูป

$$\hat{\beta}_0(n) = \alpha_1(Z_n - \hat{S}_{n-s}) + (1 - \alpha_1)[\hat{\beta}_0(n-1) + \hat{\beta}_1(n-1)]$$

เมื่อ $0 < \alpha_1 < 1$ เป็นค่าคงที่ในการทำให้เรียบ

2. ค่าประมาณของ β_1 ปรับค่าได้ในรูป

$$\hat{\beta}_1(n) = \alpha_2[\hat{\beta}_0(n) - \hat{\beta}_0(n-1)] + (1 - \alpha_2)\hat{\beta}_1(n-1)$$

เมื่อ $0 < \alpha_2 < 1$ เป็นค่าคงที่ในการทำให้เรียบ

3. ค่าที่แสดงการแปรผันตามฤดูกาลที่เป็นค่าสังเกต $Z_n - \hat{\beta}_0(n)$ ในรูป

$$\hat{S}_n = \alpha_3[Z_n - \hat{\beta}_0(n)] + (1 - \alpha_3)\hat{S}_{n-s}$$

เมื่อ $0 < \alpha_3 < 1$ เป็นค่าคงที่ในการทำให้เรียบ

4. ค่าพยากรณ์ที่เวลา $n+l$ โดยวิธีของวินเดอร์แสดงได้ในรูป

$$\hat{Z}_n(l) = \hat{\beta}_0(n) + \hat{\beta}_1(n)l + \hat{S}_{n+l-s}$$

3 การพยากรณ์

เนื่องจากอนุกรมเวลาชุดนี้ไม่มีแนวโน้ม และมีคาบของฤดูกาลเท่ากับ 12 เดือน เราสามารถเลือกวิเคราะห์ด้วยวิธีวินเดอร์สำหรับกรณีฤดูกาลเชิงบวกชนิดที่ไม่มีแนวโน้ม ฉะนั้นตัวแบบที่ใช้ในการพยากรณ์เมื่อ $n = 72$ และ $s = 12$ คือ

$$\hat{Z}_{72}(l) = \hat{\beta}_0(72) + \hat{S}_{60+l}, \quad l = 1, 2, \dots$$

โดยที่ค่า $\hat{\beta}_0(72)$ และ \hat{S}_{60+l} ได้มาจากสมการ

$$\hat{\beta}_0(n) = \alpha_1(Z_n - \hat{S}_{n-12}) + (1 - \alpha_1)\hat{\beta}_0(n-1)$$

$$\hat{S}_n = \alpha_3[Z_n - \hat{\beta}_0(n)] + (1 - \alpha_3)\hat{S}_{n-12}$$

ในกรณีนี้ค่าคงที่ในการทำให้เรียบที่ต้องกำหนดคือ α_1 และ α_3 ซึ่งการศึกษาค้นคว้าครั้งนี้ใช้การวิเคราะห์และตัดสินใจโดยใช้โปรแกรมสำเร็จรูปช่วยในการหาคำตอบ

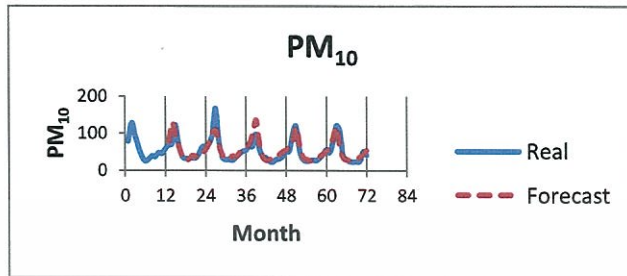
4 ผลการดำเนินงาน

ผลการหาค่า α_1 และ α_3 ที่ทำให้ RMSE มีค่าน้อยที่สุดของแต่ละตัวแปรแสดงได้ดังนี้

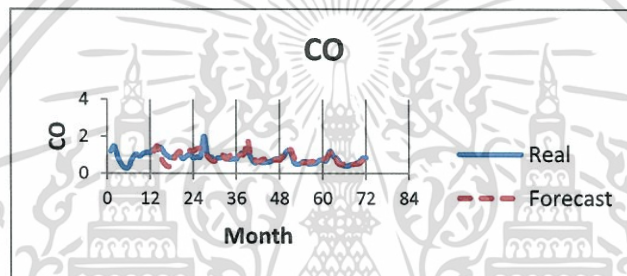
ตารางที่ 2 α_1 และ α_3 ที่ทำให้ RMSE มีค่าน้อยที่สุด

ตัวแปร	α_1	α_3	RMSE
PM ₁₀	0.02421	0.47152	16.105
CO	0.07923	0.738348	0.244
NO ₂	0.02114	0.542558	3.411
RH	0.45221	1	6.033
Rain	0.39209	0.85203	0.100

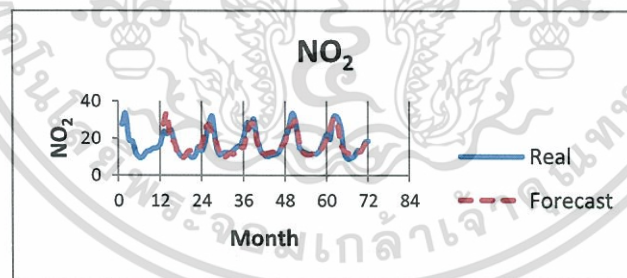
จากข้อมูลอนุกรมเวลาทำให้เราได้กราฟความสัมพันธ์ระหว่างปัจจัยต่าง ๆ ที่เราสนใจและเวลา โดยกราฟที่แสดงในแนวนอน หมายถึงเดือนโดยเริ่มตั้งแต่ ม.ค. 2005 – ธ.ค. 2010 กราฟเส้นทึบสีฟ้าคือค่าข้อมูลจริงที่ได้จากกรมควบคุมมลพิษ ส่วนกราฟเส้นประสีแดงคือค่าที่ได้จากการพยากรณ์จากตัวแบบทางคณิตศาสตร์ที่ได้ครั้งนี้



รูปที่ 1 ปริมาณฝุ่นละอองขนาดเล็กกว่า 10 ไมครอนรายเดือนตั้งแต่ปี 2005 – 2010

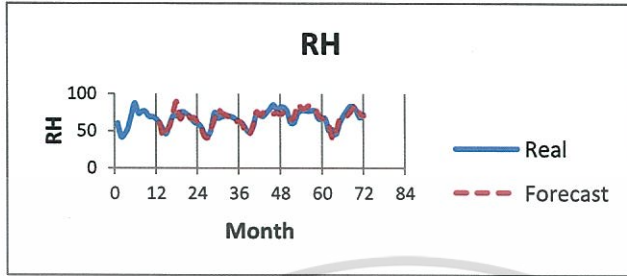


รูปที่ 2 ปริมาณก๊าซคาร์บอนมอนอกไซด์รายเดือนตั้งแต่ปี 2005 – 2010

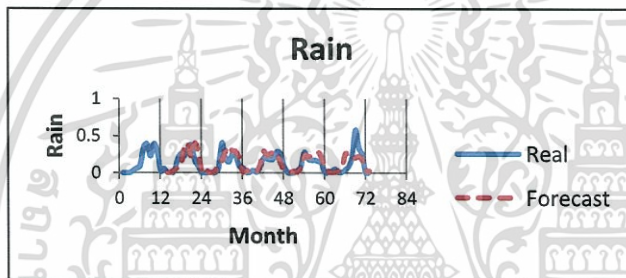


รูปที่ 3 ปริมาณก๊าซไนโตรเจนไดออกไซด์รายเดือนตั้งแต่ปี 2005 – 2010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4 ปริมาณความชื้นสัมพัทธ์รายเดือนตั้งแต่ปี 2005 – 2010



รูปที่ 5 ปริมาณน้ำฝนรายเดือนตั้งแต่ปี 2005 – 2010

ตัวแบบนี้เป็นตัวแบบที่เหมาะสมสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเชิงบวก ซึ่งหากข้อมูลไม่มีเหตุการณ์ที่ผิดปกติค่าคลาดเคลื่อนที่ได้จากตัวแบบจะมีค่าน้อย แต่ถ้าข้อมูลมีเหตุการณ์ผิดปกติ ค่าจริงและค่าพยากรณ์จะแตกต่างกันมาก ทำให้เกิดค่าคลาดเคลื่อนมาก และสำหรับค่าพยากรณ์ที่ได้จากแบบจำลองสำหรับแต่ละเดือนของปี 2011 จะมีแนวโน้มไม่แตกต่างจากปีก่อนหน้ามากนัก

ตาราง 3 ค่าจริงและค่าพยากรณ์ของปัจจัยที่เกี่ยวข้องกับปัญหาหมอกควันในปี ค.ศ. 2011

Month	PM ₁₀		CO		NO ₂		RH		Rain	
	Real	Forecast	Real	Real	Real	Forecast	Real	Forecast	Real	Forecast
1	50.36	54.23	0.83	66.93	21.32	22.26	66.93	68.49	21.32	22.26
2	54.59	84.11	0.82	54.25	24.88	31.14	54.25	54.89	24.88	31.14
3	48.83	118.73	0.87	65.25	18.94	30.58	65.25	51.04	18.94	30.58
4	51.41	81	0.7	69.34	16.57	22.39	69.34	45.96	16.57	22.39
5	36.71	36.29	0.52	79.00	12.52	12.74	79.00	62.45	12.52	12.74
6	22.69	28.03	0.46	77.89	11.09	10.09	77.89	71.26	11.09	10.09
7	21.34	24.47	0.48	76.28	8.96	10.21	76.28	75.19	8.96	10.21
8	24.03	24.79	0.58	81.4	8.54	10.88	81.40	78.18	8.54	10.88
9	22.09	26.09	0.61	81.26	8.10	12.71	81.26	78.01	8.1	12.71
10	29.36	30.27	0.58	79.76	10.76	13.10	79.76	74.84	10.76	13.1
11	28.06	46.76	0.77	68.98	15.38	17.20	68.98	67.73	15.38	17.2
12	37.06	47.89	0.72	70.41	17.69	19.44	70.41	70.83	17.69	19.44
RMSE : 24.446		RMSE : 0.118		RMSE : 4.570		RMSE : 9.650		RMSE : 0.117		

5 บทสรุป

ปัจจัยหลักที่ได้จากการวิจัยนี้มีผลต่อการเกิดหมอกควันของจังหวัดเชียงใหม่ได้แก่ ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM₁₀) ก๊าซคาร์บอนมอนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO₂) ความชื้นสัมพัทธ์ (Relative humidity : RH) และปริมาณน้ำฝน (Rain) ซึ่งเมื่อวิเคราะห์ข้อมูลอนุกรมเวลาของปัจจัยดังกล่าว พบว่าเป็นข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงตามฤดูกาลในเชิงบวก (Additive seasonal effect) ซึ่งสามารถหาค่าพยากรณ์โดยวิธีของวินเตอร์ โดยหากข้อมูลไม่มีเหตุการณ์ที่ผิดปกติจะทำค่าคลาดเคลื่อนที่ได้จากตัวแบบมีค่าน้อย แต่ถ้าข้อมูลมีเหตุการณ์ผิดปกติจะทำให้เกิดค่าคลาดเคลื่อนมาก และค่าพยากรณ์ในแต่ละเดือนของปี 2011 จะมีแนวโน้มไม่แตกต่างจากค่าในแต่ละเดือนของปีก่อนมากนัก

เอกสารอ้างอิง

- [1] S. Xie, Tong Yu, Y. Zhang, L. Qi and X. Tang, *Characteristics of PM₁₀, SO₂, NO_x and O₃ in ambient air during the dust storm period in Beijing*, Science of the Total Environment, 345 (2005), 153-164.
- [2] M. Douglas, *Forecasting and time series analysis*, McGraw-Hill, New York (1990).
- [3] R. Cliff, *Spreadsheet modeling and decision analysis : a practical introduction to management science*, South-Western, Mason (2004).
- [4] กัลยา วาณิชชัญญา, *การวิเคราะห์ข้อมูลทางสถิติด้วย Excel*, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ (2552).
- [5] พนิดา พานิชกุล และยุทธภูมิ วงศ์วัฒนฤกษ์, *ค้นภึร์ การวิเคราะห์และตัดสินใจปัญหาเชิงธุรกิจโดยใช้ Excel*, เคทีพี คอมพ์ แอนด์ คอนซัลท์, กรุงเทพฯ. (2546).
- [6] *มลพิษหมอกควันในประเทศไทย*, สืบค้นจาก http://www.pcd.go.th/info_serv/air_thaihaze.html
- [7] [Online] <http://www.statsoft.com/textbook/time-series-analysis/?button=3#census>
- [8] [Online] http://en.wikipedia.org/wiki/Time_series

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้