

การสร้างตัวแบบการถดถอยเพื่อพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี
REGRESSION MODELING FOR NO_2 PREDICTION IN CHON BURI PROVINCE

ศุทธิณี คุ่มครองทรัพย์

ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาสถิติ ปีการศึกษา 2555
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

ชื่อเรื่อง

การสร้างตัวแบบการถดถอยเพื่อพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี

REGRESSION MODELING FOR NO_2 PREDICTION IN CHON BURI PROVINCE

ชื่อนิสิต ศุทธิณี คุ่มครองทรัพย์

รหัสประจำตัวนิสิต 52030032

อาจารย์ที่ปรึกษาหลัก อาจารย์ยวนิดา พงษ์ศักดิ์ชาติ

ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ
ปีการศึกษา 2555

คณะกรรมการควบคุมปัญหาพิเศษ

..........อาจารย์ที่ปรึกษาหลัก
(อาจารย์ยวนิดา พงษ์ศักดิ์ชาติ)

คณะกรรมการสอบปัญหาพิเศษ

..........ประธานกรรมการ
(อาจารย์อภิศักดิ์ ไชยโรจน์วัฒนา)

..........กรรมการ
(อาจารย์จุฑาพร เนียมวงศ์)

..........กรรมการ
(อาจารย์ยวนิดา พงษ์ศักดิ์ชาติ)

คณะกรรมการสอบปัญหาพิเศษอนุมัติให้ปัญหาพิเศษฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรบัณฑิต สาขาวิชาสถิติ ของมหาวิทยาลัยบูรพา

..........ประธานหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาสถิติ

(ผู้ช่วยศาสตราจารย์ปรียารัตน์ นาคสุวรรณ)

วันที่.....เดือน.....พ.ศ.....

..........รักษาการหัวหน้าภาควิชาคณิตศาสตร์

(ผู้ช่วยศาสตราจารย์อภิศักดิ์ ภคพงศ์พันธุ์)

วันที่ 25 เดือน ธันวาคม พ.ศ. 2556

ประกาศคุณูปการ

ปัญหาพิเศษฉบับนี้สำเร็จลงได้ด้วยความกรุณาจาก อาจารย์วันดา พงษ์ศักดิ์ชาติ อาจารย์ที่ปรึกษา ปัญหาพิเศษที่กรุณาให้คำปรึกษา คำแนะนำแนวทางที่ถูกต้อง ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ ด้วยความละเอียดถี่ถ้วนและเอาใจใส่ด้วยดีเสมอมาผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างยิ่งจึงขอขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

นอกจากนี้ผู้วิจัยขอขอบพระคุณกรรมการสอบปัญหาพิเศษในครั้งนี้ ซึ่งประกอบด้วย ดร.จุฑาพร จินดาสวัสดิ์ และอาจารย์อภิศักดิ์ ไชยโรจน์วัฒนา ที่ได้สละเวลาและกรุณาช่วยตรวจแก้ไขและวิจารณ์ผลงาน ทำให้ปัญหาพิเศษฉบับนี้มีความสมบูรณ์ถูกต้องยิ่งขึ้น

และสุดท้ายนี้คุณค่าและประโยชน์ของปัญหาพิเศษฉบับนี้ ผู้วิจัยขอมอบเป็นกตัญญูตเวทิตาแด่ บพการี บุรพาจารย์ ผู้มีพระคุณทุกท่านทั้งในอดีตและปัจจุบัน ทำให้ข้าพเจ้าเป็นผู้มีการศึกษา และประสบความสำเร็จมาจนตราบนานเท่านานนี้รวมทั้งขอบคุณเพื่อน ๆ ทุกคน ที่เป็นกำลังใจและให้คำแนะนำในการทำ ปัญหาพิเศษจนสำเร็จลุล่วงด้วยดี

ศุทธิณี คุ่มครองทรัพย์

52030032: สาขาวิชา: สถิติ; วท.บ. (สถิติ)

คำสำคัญ: การสร้างตัวแบบ NO_2 / MLR

ศุทธิณี คุ่มครองทรัพย์: การสร้างตัวแบบการถดถอยเพื่อพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี (REGRESSION MODELING FOR NO_2 PREDICTION IN CHONBURI PROVINCE)
คณะกรรมการควบคุมปัญหาพิเศษ: อาจารย์วนิดา พงษ์ศักดิ์ชาติ, Ph.D 24 หน้า. ปี พ.ศ. 2555

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อสร้างตัวแบบการถดถอยสำหรับการพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี โดยใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression : MLR) และมีตัวแปรอิสระที่สนใจเป็นตัวแปรทางด้านอุตุนิยมวิทยา ได้แก่ ปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และทิศทางลม โดยใช้ข้อมูล NO_2 จากกรมควบคุมมลพิษและใช้ข้อมูลตัวแปรอิสระจากกรมอุตุนิยมวิทยา โดยใช้ข้อมูลตั้งแต่ปี พ.ศ. 2549–2554 รวม 6 ปี โดยแบ่งข้อมูลออกเป็นสองชุดคือ ข้อมูลตั้งแต่ปี พ.ศ. 2549–2553 ใช้สำหรับสร้างตัวแบบสำหรับการพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี และใช้ข้อมูลปี พ.ศ. 2554 สำหรับประเมินความสามารถในการพยากรณ์ตัวแบบที่ได้

ผลการศึกษาพบว่า ตัวแบบการถดถอยที่เหมาะสมสำหรับการพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี ประกอบด้วยตัวแปรอิสระทั้งหมด 5 ตัวแปร ได้แก่ ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลมและ $NO_{2(t-1)}$ โดยสมการถดถอยที่ได้มีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 67.23 สัมประสิทธิ์การตัดสินใจเท่ากับ 67.33 และค่าคลาดเคลื่อนมาตรฐานของการประมาณเท่ากับ 0.240 ppb เมื่อนำสมการถดถอยที่ได้มาพยากรณ์ค่าความเข้มข้นของ NO_2 ในปี พ.ศ. 2554 ได้ค่ารากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ 3.583 ppb ซึ่งถือได้ว่าตัวแบบการถดถอยที่ได้ดีมีความน่าเชื่อถืออยู่ในระดับปานกลาง

52030032: MAJOR: STATISTICS; B.Sc. (STATISTICS)
KEYWORDS: REGRESSION MODELING/ MLR'S METHOD
REGRESSION MODELING FOR NO_2 PREDICTION IN CHON BURI PROVINCE
ADVISOR: ASSISTANT PONGSAKCHAT VANIDA, Ph.D 24 P. ACADEMIC YEARS 2012

ABSTRACT

The objective of this study was to develop the regression model for predicting NO_2 concentration in Chon Buri Province. Multiple linear regression analysis was used and independent variables of interest were the meteorological variables including rain, relative humidity, temperature, barometric pressure, wind speed and wind direction. NO_2 data was from the Pollution Control Department and the data of independent variables was from the Department of Meteorology. The data during 2506–2010 were used and divided into 2 sets. The data from 2006–2010 was used for model building and the data in 2011 was used for testing the model.

The results showed that the final regression model for predicting NO_2 concentration in Chon Buri province included five independent variables; relative humidity, temperature, barometric pressure, wind speed and $NO_{2(t-1)}$. Adjusted coefficient of determination and the standard error of the estimate were 67.23 and 0.240 ppb respectively. For testing this model using data in 2011, RMSE was 3.583 ppb. These statistics indicated that the final regression model had moderate quality in predicting NO_2 concentration.

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ซ
สารบัญรูปภาพ.....	ช
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการศึกษา.....	2
1.3 ประโยชน์ที่คาดว่าจะได้รับการศึกษา.....	2
1.4 ขอบเขตของการศึกษา.....	2
1.5 นิยามศัพท์เฉพาะ.....	2
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	3
2.1 ทฤษฎี.....	3
2.1.1 มลพิษทางอากาศ.....	3
2.1.2 การวิเคราะห์การถดถอย.....	4
2.2 งานวิจัยที่เกี่ยวข้อง.....	6
บทที่ 3 วิธีการดำเนินการศึกษา.....	8
3.1 ข้อมูลที่นำมาศึกษา.....	8
3.2 การสร้างตัวแบบ.....	8
3.3 ความสามารถในการพยากรณ์ของตัวแบบการถดถอย.....	9
บทที่ 4 ผลการศึกษา.....	10
4.1 ข้อมูลที่นำมาศึกษา.....	10
4.2 การสร้างตัวแบบ.....	10
4.3 ความสามารถในการพยากรณ์ของตัวแบบการถดถอย.....	17
บทที่ 5 สรุปและอภิปรายผลการศึกษา.....	18
5.1 สรุปและอภิปรายผลการศึกษา.....	18
5.2 ข้อเสนอแนะ.....	18
บรรณานุกรม.....	19
ภาคผนวก.....	21

สารบัญตาราง

ตารางที่	หน้า
1. ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน มัธยฐานของ NO_2 และตัวแปรอิสระแต่ละตัว.....	10
2. ค่า VIF ของตัวแปรอิสระแต่ละตัว.....	15
3. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ.....	15

สารบัญภาพ

ภาพที่	หน้า
1. กราฟฮิสโทแกรมของตัวแปรต่าง ๆ.....	11
2. กราฟการกระจายระหว่าง $\ln y$ และตัวแปรอิสระแต่ละตัว	13
3. กราฟส่วนตกค้าง ($\hat{\varepsilon}_i$) ของสมการถดถอยที่ (4).....	15
4. กราฟส่วนตกค้างที่พล็อตระหว่าง $\hat{\varepsilon}_i$ กับ ลำดับเวลาการเก็บข้อมูลของสมการถดถอยที่ (5)..	16

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการพยากรณ์ได้เข้ามามีบทบาทในการวิจัยสาขาต่าง ๆ โดยวิธีการพยากรณ์นั้นจะใช้วิธีการวิเคราะห์การถดถอย (Regression Analysis) เป็นส่วนใหญ่ ซึ่งการวิเคราะห์การถดถอยเป็นวิธีวิเคราะห์ทางสถิติที่ใช้หารูปแบบความสัมพันธ์ระหว่างตัวแปร 2 ชนิดหรือมากกว่า เพื่อนำไปสู่การพยากรณ์ค่าจริง

การเจริญเติบโตทางเศรษฐกิจอย่างรวดเร็วของประเทศไทยที่เปลี่ยนจากภาคเกษตรกรรมมาเป็นภาคอุตสาหกรรมทำให้มีปัญหามากมายที่ตามมาคือ ปัญหาการจราจรหนาแน่น ซึ่งการจราจรที่หนาแน่นทำให้รถเคลื่อนตัวได้ด้วยความเร็วต่ำมีการหยุดและออกตัวบ่อยครั้งขึ้น น้ำมันถูกเผาผลาญมากขึ้น การสันดาปของน้ำมันเชื้อเพลิงไม่สมบูรณ์ และมีการระบายสารมลพิษทางท่อไอเสียในสัดส่วนที่เพิ่มมากขึ้น ดังนั้นบริเวณที่ใกล้ถนนที่มีการจราจรหนาแน่น จะมีปัญหามลพิษทางอากาศที่รุนแรงกว่าในบริเวณที่มีการจราจรคล่องตัว มลพิษทางอากาศ เป็นภาวะอากาศที่มีสารเจือปนอยู่ในปริมาณที่สูงกว่าระดับปกติเป็นเวลานานพอที่จะทำให้เกิดอันตรายแก่มนุษย์ สัตว์ พืช หรือทรัพย์สินต่าง ๆ อาจเกิดขึ้นเองตามธรรมชาติและเกิดจากการกระทำของมนุษย์ มลพิษทางอากาศที่เกิดจากการเผาไหม้เชื้อเพลิงชนิดต่าง ๆ ได้แก่ ฝุ่นละออง (PM_{10}) คาร์บอนมอนอกไซด์ (CO) ซัลเฟอร์ไดออกไซด์ (SO_2) และออกไซด์ของไนโตรเจน (NO_x) ซึ่งพบว่ามีปริมาณการระบายออกสู่ชั้นบรรยากาศเพิ่มมากขึ้นทุกปีตามปริมาณการใช้เชื้อเพลิงที่เพิ่มขึ้น ในการศึกษาจะศึกษามลพิษทางอากาศที่สำคัญ คือ ไนโตรเจนไดออกไซด์ (NO_2)

NO_2 เป็นก๊าซมีสีน้ำตาลแดง มีกลิ่นฉุนเฉพาะตัว น้ำหนักโมเลกุล 46.01 g/mol ความหนาแน่น 1.58 kg/m^3 และจุดเดือด 21.2 องศาเซลเซียส ละลายน้ำได้เล็กน้อย เมื่อทำปฏิกิริยากับน้ำได้กรดไนตริกและกรดไนตริก จะมีฤทธิ์กัดกร่อนสูง เกิดขึ้นได้เองตามธรรมชาติแต่เกิดในปริมาณไม่มาก สาเหตุหลักมาจากการเผาไหม้เชื้อเพลิงชนิดต่าง ๆ ในอุณหภูมิสูง เช่น การเผาไหม้เชื้อเพลิงเครื่องยนต์ รถยนต์ โรงไฟฟ้า อุตสาหกรรมต่าง ๆ การเผาไหม้เชื้อเพลิงในครัวเรือน โดยสัดส่วนการเกิดจากรถยนต์มีสูงถึงร้อยละ 55 ของการเกิดทั้งหมด เมื่อ NO_2 เข้าสู่ร่างกายโดยการหายใจโดยตรง หรือในรูปไอระเหยของละอองของกรดไนตริกหรือไนตริก หลังจากที่ยาซึมรวมตัวกับละอองน้ำหรือความชื้นแล้วทำให้เกิดการระคายเคืองอย่างรุนแรงต่อระบบทางเดินหายใจ ตา จมูก เยื่อเมือก และผิวหนังที่สัมผัส การสัมผัส NO_2 ที่ความเข้มข้นต่ำ เป็นระยะเวลานาน อาจทำให้เกิดสีเหลืองที่ผิวหนังและฟัน มีอาการปวดบวมน้ำและเกิดโรคเรื้อรังได้ ถ้าได้รับความเข้มข้นสูงจะทำให้ระคายเคืองที่ปอดอย่างรุนแรง ทำให้รู้สึกไม่สบายตัว ผิวหนังเป็นสีเขียวคล้ำจากการขาดออกซิเจน ไอ หายใจขัด หายใจลำบาก เป็นไข้ ปวดศีรษะ คลื่นไส้ อาเจียน และอาจถึงตายได้ คณะกรรมการสิ่งแวดล้อมแห่งชาติได้กำหนดมาตรฐานของ NO_2 ในบรรยากาศทั่วไป คือ ค่าเฉลี่ยของไนโตรเจนไดออกไซด์ในเวลา 1 ชั่วโมง จะต้องไม่เกิน 0.17 ส่วนในล้านส่วน (ppm) หรือ ไม่เกิน 0.32 มิลลิกรัมต่อลูกบาศก์เมตร (mg/m^3) (ที่มา : ประกาศคณะกรรมการสิ่งแวดล้อมแห่งชาติฉบับที่ 33 พ.ศ. 2552)

เนื่องจากจังหวัดชลบุรีเป็นแหล่งอุตสาหกรรมและการท่องเที่ยว จึงทำให้มีผู้คนจำนวนมากเดินทางเข้ามาในจังหวัดชลบุรี ทั้งเพื่อทำงานและพักผ่อนเยี่ยมชมนสถานที่ท่องเที่ยวต่าง ๆ ปัญหาที่ตามมาคือ ปัญหาการจราจรหนาแน่น ซึ่งการจราจรที่หนาแน่นเป็นสาเหตุหนึ่งที่ทำให้เกิดมลพิษทางอากาศ จึงทำให้ผู้ศึกษาสนใจในการศึกษาการสร้างความสามารถในการพยากรณ์ค่าความเข้มข้นของ NO_2 โดยใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression : MLR) และมีตัวแปรอิสระที่น่าสนใจ ได้แก่

ปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และทิศทางลม โดยเลือกจังหวัด ชลบุรีเป็นกรณีศึกษาครั้งนี้ เนื่องจากเห็นถึงปัญหามลพิษทางอากาศที่เกิดขึ้นจากการจราจร โดยใช้ข้อมูลตั้งแต่ ปี พ.ศ. 2549–2554

1.2 วัตถุประสงค์ของการศึกษา

เพื่อสร้างตัวแบบการถดถอยสำหรับการพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี

1.3 ประโยชน์ที่คาดว่าจะได้รับจากการศึกษา

1. ได้ตัวแบบการถดถอยเชิงเส้นพหุคูณที่สามารถพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี
2. สามารถนำผลที่ได้จากการพยากรณ์ไปใช้เป็นข้อมูลในการหาวิธีลดและป้องกันไม่ให้ความเข้มข้นของ NO_2 มีค่าเกินมาตรฐานที่กำหนดไว้

1.4 ขอบเขตของการศึกษา

1. พื้นที่ในการศึกษา คือ จังหวัดชลบุรี โดยสร้างตัวแบบการถดถอยเพื่อพยากรณ์ค่าความเข้มข้นของ NO_2 ซึ่งเป็นมลพิษที่เกิดขึ้นจากการจราจร
2. ใช้ข้อมูลความเข้มข้นของ NO_2 เป็นค่าเฉลี่ยรายวันจากกรมควบคุมมลพิษตั้งแต่ปี พ.ศ. 2549–2554 รวม 6 ปี จากสถานีสำนักงานสามัญศึกษาชลบุรี และใช้ข้อมูลปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และทิศทางลม เป็นค่าเฉลี่ยรายวันจากกรมอุตุนิยมวิทยาตั้งแต่ปี พ.ศ. 2549–2554 รวม 6 ปี จากสถานีชลบุรี

1.5 นิยามศัพท์เฉพาะ

ไนโตรเจนไดออกไซด์ (NO_2) หมายถึง มลพิษทางอากาศที่เกิดขึ้นจากการจราจร มีหน่วยเป็น ppb
 ความเร็วลม หมายถึง การเคลื่อนที่ของอากาศที่ทำให้เกิดแรง หรือความกดที่ผ่านจุดที่กำหนดให้บนพื้นผิวโลก และแรงหรือความกดเป็นสัดส่วนกับกำลัง 2 ของความเร็วลม มีหน่วยเป็น กิโลเมตรต่อชั่วโมง
 ทิศทางลม หมายถึง การวัดทิศลมนิยมวัดทิศลมตามเข็มทิศ และวัดเป็นองศา ถ้าวัดทิศลมด้วยเข็มทิศ เข็มทิศจะถูกแบ่งออกเป็น ทิศใหญ่ ๆ 4 ทิศ คือ ทิศเหนือ ทิศใต้ ทิศตะวันออก ซึ่งทิศทั้ง 4 ทิศ เมื่อแบ่งย่อยอีกจะเป็น 8 ทิศ โดยจะเพิ่มทิศตะวันออกเฉียงเหนือ ทิศตะวันออกเฉียงใต้ ทิศตะวันตกเฉียงเหนือ และการวัดทิศลมที่เป็นองศาบอกมุมของลมจากทิศจริง ในลักษณะที่เวียนไปตามเข็มนาฬิกา ใช้สเกลจาก 0 องศา ไปจนถึง 360 องศา มีหน่วยเป็น องศา

ความดันบรรยากาศ หมายถึง ความกดดันอยู่จุดใดหนึ่งของชั้นบรรยากาศของโลก โดยทั่วไปความกดอากาศจะประมาณเท่ากับความกดดันที่เกิดขึ้นย้อนหน้าของอากาศอยู่บนจุดนั้น ๆ ซึ่งหมายความว่า จุดที่มีความกดอากาศต่ำจะมีอากาศที่มีมวลสารต่ำกว่าจะอยู่ข้างบน มีหน่วยเป็น เฮกโตปาสกาล

ความชื้นสัมพัทธ์ หมายถึง อัตราส่วนระหว่างปริมาณความชื้น (ไอน้ำ) ที่มีอยู่จริงในอากาศกับปริมาณความชื้น (ไอน้ำ) ที่อากาศขณะนั้น มีหน่วยเป็นร้อยละ

อุณหภูมิ หมายถึง ปริมาณที่บอกถึงความร้อนหรือเย็นของวัตถุ มีหน่วยเป็น องศาเซลเซียส

ปริมาณน้ำฝน หมายถึง ระดับความลึกของน้ำฝนในภาชนะที่รองรับน้ำฝน ทั้งนี้ภาชนะที่รองรับน้ำฝนจะต้องตั้งอยู่ในแนวระดับ และวัดในช่วงเวลาที่กำหนด มีหน่วยเป็น มิลลิเมตร

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1. ทฤษฎี

2.1.1 มลพิษทางอากาศ

ความหมายของมลพิษทางอากาศ

มลพิษทางอากาศ หมายถึง ภาวะอากาศที่มีสารเจือปนอยู่ในปริมาณที่สูงกว่าระดับปกติเป็นเวลานานพอที่จะทำให้เกิดอันตรายแก่มนุษย์ สัตว์ พืช หรือทรัพย์สินต่าง ๆ อาจเกิดขึ้นเองตามธรรมชาติ เช่น ฝุ่นละอองจากลมพายุ ภูเขาไฟระเบิด แผ่นดินไหว ไฟไหม้ป่า ก๊าซธรรมชาติอากาศเสียที่เกิดขึ้นโดยธรรมชาติเป็นอันตรายต่อมนุษย์น้อยมาก เพราะแหล่งกำเนิดอยู่ไกลและปริมาณที่เข้าสู่สภาพแวดล้อมของมนุษย์และสัตว์มีน้อย หรือเกิดจากการกระทำของมนุษย์ ได้แก่ มลพิษจากท่อไอเสียของรถยนต์ จากโรงงานอุตสาหกรรม จากกระบวนการผลิตต่าง ๆ จากกิจกรรมด้านการเกษตร จากการระเหยของก๊าซบางชนิดซึ่งเกิดจากขยะมูลฝอยและของเสีย เป็นต้น

แหล่งกำเนิดสารมลพิษทางอากาศ

แหล่งกำเนิดมลพิษทางอากาศที่สำคัญของประเทศไทย แบ่งเป็น 2 กลุ่มใหญ่ ๆ ดังนี้

1. แหล่งกำเนิดจากยานพาหนะ ก่อให้เกิดปัญหามลพิษทางอากาศจำกัดเฉพาะในเขตชุมชนขนาดใหญ่ เช่น กรุงเทพมหานครและปริมณฑล การจราจรที่หนาแน่นทำให้รถเคลื่อนตัวได้ด้วยความเร็วต่ำ มีการหยุดและออกตัวบ่อยครั้งขึ้น น้ำมันถูกเผาผลาญมากขึ้น การสันดาปของน้ำมันเชื้อเพลิงไม่สมบูรณ์ และมีการระบายสารมลพิษทางท่อไอเสียในสัดส่วนที่เพิ่มมากขึ้น ดังนั้น บริเวณที่ใกล้ถนนที่มีการจราจรหนาแน่น จะมีปัญหามลพิษทางอากาศที่รุนแรงกว่าในบริเวณที่มีการจราจรคล่องตัว สารมลพิษที่ระบายเข้าสู่บรรยากาศที่เกิดจากการคมนาคมขนส่ง ได้แก่ ก๊าซคาร์บอนมอนอกไซด์ ก๊าซออกไซด์ของไนโตรเจน สารประกอบไฮโดรคาร์บอน ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน สารตะกั่ว และก๊าซซัลเฟอร์ไดออกไซด์

2. แหล่งกำเนิดจากโรงงานอุตสาหกรรม ก่อให้เกิดปัญหามลพิษทางอากาศเฉพาะพื้นที่ที่โรงงานอุตสาหกรรมนั้นตั้งอยู่ ซึ่งกระจายอยู่ทั่วประเทศทั้งในเขตชนบทและเขตเมือง มลพิษทางอากาศที่มีแหล่งกำเนิดจากโรงงานอุตสาหกรรมเกิดจากการเผาไหม้เชื้อเพลิงในกระบวนการผลิต ซึ่งเป็นตัวการสำคัญที่ก่อให้เกิดผลกระทบต่อคุณภาพอากาศในบรรยากาศ และอาจส่งผลกระทบต่อสุขภาพอนามัยของประชาชนในชุมชนโดยทั่วไป หรือก่อให้เกิดความเดือดร้อนรำคาญ ซึ่งเชื้อเพลิงที่ใช้ในโรงงานอุตสาหกรรมมีอยู่ 3 ประเภทใหญ่ ๆ ด้วยกัน คือ

- เชื้อเพลิงที่เป็นของแข็ง
- เชื้อเพลิงที่เป็นของเหลว ได้แก่ น้ำมันเตา และน้ำมันดีเซล
- เชื้อเพลิงที่เป็นก๊าซ ได้แก่ ก๊าซธรรมชาติ และก๊าซ LPG

ไนโตรเจนออกไซด์ (Nitrogen Oxide: NO_2)

NO_2 หรือ กลุ่มก๊าซที่มีความไว (Highly Reactive Gases) เป็นก๊าซซึ่งมีองค์ประกอบทางเคมีของไนโตรเจนและออกซิเจนในอัตราส่วนแตกต่างกัน ไม่มีสี ไม่มีกลิ่น ยกเว้น NO_2 ที่เมื่อรวมตัวกับอนุภาคอื่น ๆ ในอากาศจะเห็นเป็นชั้นสีน้ำตาลแดง ก๊าซชนิดนี้เกิดขึ้นได้เองตามธรรมชาติ แต่เกิดขึ้นในปริมาณไม่มากนัก ซึ่งสาเหตุหลักของการเกิด NO_2 เกิดจากการเผาไหม้เชื้อเพลิงชนิดต่าง ๆ ในอุณหภูมิสูง เช่น การเผาไหม้เชื้อเพลิง ของเครื่องยนต์ในรถยนต์ โรงไฟฟ้า โรงงานอุตสาหกรรมต่าง ๆ และการเผาไหม้เชื้อเพลิงในครัวเรือน โดย

สัดส่วนการเกิดก๊าซชนิดนี้จากรถยนต์มีสูงถึงร้อยละ 55 ของทั้งหมด ค่ามาตรฐาน NO_2 ในบรรยากาศทั่วไปโดยปัจจัยที่มีผลต่อ NO_2 คือ การเผาไหม้ของเชื้อเพลิงที่อุณหภูมิสูง

คุณลักษณะของ NO_2

NO_2 เป็นก๊าซชนิดหนึ่งของออกไซด์ของไนโตรเจน (NO_x) มีลักษณะเป็นสีน้ำตาลแกมแดงและมีกลิ่นฉุนเกิดจากกระบวนการเผาไหม้เชื้อเพลิงที่อุณหภูมิสูง เช่น ยานพาหนะ โรงไฟฟ้า เป็นต้น ออกไซด์ของไนโตรเจนประกอบด้วยไนตรัสออกไซด์ (N_2O) ไนตริกออกไซด์ (NO) ไนโตรเจนไดออกไซด์ (NO_2) ไนโตรเจนไตรออกไซด์ (NO_3) ไดไนโตรเจนไตรออกไซด์ (N_2O_3) ไดไนโตรเจนเตตราออกไซด์ (N_2O_4) และไดไนโตรเจนเพนตาออกไซด์ (N_2O_5) แต่มีเพียง NO และ NO_2 ที่เป็นสารมลพิษทางอากาศที่สำคัญ N_2O เป็นก๊าซเรือนกระจกและเป็นสารมลพิษพื้นฐาน (Criteria Pollutant) ที่บ่งชี้สถานการณ์คุณภาพอากาศ เนื่องจากหลักฐานทางวิชาการระบุอย่างชัดเจนว่าการได้รับ NO_2 จะส่งผลกระทบต่อสุขภาพอนามัยโดย NO_2 ในบรรยากาศจะถูกปล่อยจากการเผาไหม้โดยตรง และเกิดจากปฏิกิริยาโฟโตเคมีคัลของ NO นอกจากนี้ทั้ง NO และ NO_2 ยังเป็นสารตั้งต้นของการเกิดสารมลพิษทุติยภูมิหลายชนิด เช่น ก๊าซโอโซน (O_3) ในบรรยากาศชั้นโทรโพสเฟียร์ กรดไนตริก (HNO_3) และอนุภาคแขวนลอยในรูปของไนเตรต (Nitrate Aerosol) ซึ่งก่อให้เกิดการตกสะสมของกรด เป็นต้น

ผลกระทบต่อสุขภาพเมื่อได้รับ NO_2

1. ในผู้ที่เป็โรคหอบหืด (Asthmatics) เมื่อได้รับ NO_2 ที่ระดับความเข้มข้น 260 ppb เป็นเวลา 15–30 นาที จะเสริมฤทธิ์การตอบสนองต่อสารก่อภูมิแพ้ เช่น ไรฝุ่น และเกสรดอกไม้ เป็นต้น ซึ่งส่งผลต่อการอักเสบของทางเดินอากาศหายใจเพิ่มขึ้น (Airway Inflammatory) และหากผู้ที่เป็โรคนี้ได้รับ NO_2 ที่ระดับความเข้มข้น 200–300 ppb เป็นเวลา 30 นาที ถึง 2 ชั่วโมง ส่งผลต่อการเพิ่มขึ้นของการถูกกระตุ้นของทางเดินอากาศหายใจ (Airway Reactivity) และอาจส่งผลต่อการเพิ่มขึ้นของอาการโรคหอบหืดและการใช้ยารักษาโรค

2. ในผู้ที่มีสุขภาพดี (Healthy Subjects) ไม่พบหลักฐานของผลกระทบต่อสมรรถภาพการทำงานของปอดหรือการอักเสบของทางเดินอากาศหายใจเมื่อได้รับ NO_2 ต่ำกว่า 1,000 ppb

2.1.2 การวิเคราะห์การถดถอย

ตัวแบบการถดถอย

การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ เป็นวิธีการทางสถิติที่มีวัตถุประสงค์เพื่อสร้างตัวแบบการถดถอยที่เหมาะสมสำหรับพยากรณ์ค่าของตัวแปรตาม (Dependent Variable) ซึ่งเป็นตัวแปรที่สนใจศึกษา โดยอาศัยค่าของตัวแปรอิสระ (Independent Variables) หรือตัวแปรพยากรณ์ (Predictors) โดยตัวแบบการถดถอยเชิงเส้นพหุคูณมีลักษณะดังนี้

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i ; \quad \begin{matrix} i = 1, 2, \dots, n \\ j = 1, 2, \dots, k \end{matrix}$$

เมื่อ y_i คือ ค่าสังเกตที่ i ของตัวแปรตาม

$\beta_0, \beta_1, \dots, \beta_k$ คือ พารามิเตอร์ (สัมประสิทธิ์การถดถอย)

$x_{i1}, x_{i2}, \dots, x_{ik}$ คือ ค่าสังเกตที่ i ของตัวแปรอิสระแต่ละตัว

ε_i คือ ค่าคลาดเคลื่อนของค่าสังเกตที่ i

n คือ จำนวนค่าสังเกตทั้งหมด

k คือ จำนวนตัวแปรอิสระ

โดยตัวแบบการถดถอยนี้มีข้อสมมุติ ดังนี้

1. ตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กันในเชิงเส้นตรง
2. ค่าคลาดเคลื่อน (ε_i) เป็นตัวแปรสุ่มที่เป็นอิสระกัน มีการแจกแจงปกติ โดยมีค่าเฉลี่ยเท่ากับ 0

และความแปรปรวนคงที่เท่ากับ σ^2 ในทุกค่าของตัวแปรอิสระ

3. ตัวแปรอิสระในตัวแบบการถดถอยต้องไม่มีความสัมพันธ์กัน

ข้อสมมุติของตัวแบบการถดถอย

ในการวิเคราะห์การถดถอย ตัวแบบที่ได้จะเป็นตัวแบบที่เหมาะสมหากข้อสมมุติในข้างต้นเป็นไปตามที่กำหนดไว้ทุกประการ ดังนั้น จึงต้องมีการตรวจสอบข้อสมมุติเหล่านี้ของตัวแบบที่ได้ โดยการตรวจสอบข้อสมมุติแต่ละข้อมีวิธีการ ดังนี้

1. การตรวจสอบลักษณะของความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ พิจารณาได้จากกราฟการกระจายระหว่างตัวแปรตามและตัวแปรอิสระแต่ละตัวแปร หากกราฟที่ได้มีลักษณะการกระจายใกล้เคียงเส้นตรง แสดงว่าตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กันในเชิงเส้นตรง

2. การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน พิจารณาได้จาก Normal Probability Plot ของค่าส่วนตกค้าง (Residual) หากกราฟมีการกระจายใกล้เคียงเส้นตรง แสดงว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ นอกจากนั้น ยังสามารถใช้วิธีการทดสอบทางสถิติ เช่น การทดสอบ Kolmogorov-Smirnov ในการตรวจสอบข้อสมมุตินี้ได้อีกด้วย

3. การตรวจสอบความคงที่ของความแปรปรวนของค่าคลาดเคลื่อน พิจารณาได้จากกราฟการกระจายระหว่างค่าส่วนตกค้างและค่าพยากรณ์ (Predicted Value, $\hat{y}_i$) หรือกราฟการกระจายระหว่างค่าส่วนตกค้างและตัวแปรอิสระแต่ละตัวแปร หากความแปรปรวนของค่าคลาดเคลื่อนมีค่าคงที่กราฟที่ได้จะมีการกระจายแบบสุ่มรอบศูนย์ แต่หากความแปรปรวนของค่าคลาดเคลื่อนไม่คงที่ กราฟจะมีการกระจายเป็นรูปกรวย นอกจากนั้น ยังสามารถใช้วิธีการทดสอบทางสถิติ เช่น การทดสอบ Breusch-Pagan ตรวจสอบข้อสมมุตินี้ได้ด้วย

4. การตรวจสอบการเป็นอิสระกันของค่าคลาดเคลื่อน ในกรณีที่มีการเก็บรวบรวมข้อมูลเป็นไปตามลำดับของเวลา หรือลำดับใด ๆ อาจทำให้ค่าคลาดเคลื่อนไม่เป็นอิสระกัน ซึ่งเรียกว่าการเกิดอัตสหสัมพันธ์ (Autocorrelation) การตรวจสอบข้อสมมุตินี้พิจารณาได้จากกราฟการกระจายระหว่างค่าส่วนตกค้างและลำดับเวลาของการเก็บข้อมูล หากค่าคลาดเคลื่อนเป็นอิสระกันกราฟจะมีการกระจายโดยสุ่มรอบศูนย์ แต่หากกราฟมีลักษณะการกระจายมีรูปแบบ เช่น มีลักษณะเป็นเส้นโค้งวัฏจักร (Cycle) แสดงให้เห็นว่าค่าคลาดเคลื่อนไม่เป็นอิสระกัน หรือเกิดปัญหาอัตสหสัมพันธ์ นอกจากนั้น ยังสามารถใช้สถิติทดสอบ Durbin - Watson ในการตรวจสอบข้อสมมุตินี้ได้ด้วย

5. การตรวจสอบความเป็นอิสระกันของตัวแปรอิสระในตัวแบบการถดถอย ในกรณีที่มีตัวแปรอิสระมากกว่าหนึ่งตัวแปรในตัวแบบการถดถอย อาจก่อให้เกิดปัญหาที่ตัวแปรอิสระเหล่านี้มีความสัมพันธ์กันเอง ซึ่งเรียกว่าการเกิดพหุสัมพันธ์ (Multicollinearity) ความรุนแรงของปัญหาที่เกิดขึ้น ๆ อยู่กับขนาดของความสัมพันธ์ระหว่างตัวแปรอิสระเหล่านี้ หากความสัมพันธ์อยู่ในระดับต่ำปัญหาจะไม่รุนแรงมากนัก แต่ถ้าความสัมพันธ์ อยู่ในระดับสูงจะส่งผลกระทบต่อตัวประมาณค่าของสัมประสิทธิ์การถดถอยในตัวแบบการถดถอย ซึ่งปัญหานี้สามารถตรวจสอบได้โดยพิจารณาจากกราฟการกระจายระหว่างตัวแปรอิสระแต่ละคู่ และพิจารณาจากค่า Variance Inflation Factor (VIF) หากค่า VIF มีค่ามากกว่า 10 ชี้ให้เห็นว่าเกิดปัญหาพหุสัมพันธ์อย่างรุนแรง

การคัดเลือกตัวแบบ

หลังจากที่ได้สร้างตัวแบบการถดถอยในเบื้องต้น และตรวจสอบข้อสมมุติแล้ว จะพบว่า มีตัวแปรอิสระบางตัวแปรที่อาจจะไม่มีความสำคัญ และไม่มีความสัมพันธ์กับตัวแปรตาม จึงต้องทำการคัดเลือกตัวแปรอิสระเพื่อให้ได้ตัวแบบที่เหมาะสมที่ประกอบด้วยตัวแปรอิสระที่สำคัญ และมีความสัมพันธ์กับตัวแปรตามเท่านั้น ซึ่งวิธีการคัดเลือกตัวแบบที่เหมาะสมนั้นมีอยู่ด้วยกันหลายวิธี ได้แก่ การเลือกตัวแปรโดยวิธีเพิ่มตัวแปร (Forward Selection) การเลือกตัวแปรโดยวิธีลดตัวแปร (Backward Elimination) และการเลือกตัวแปรโดยวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise Regression)

2.2 งานวิจัยที่เกี่ยวข้อง

Mavko, Tang and George (2008) ได้สร้างตัวแบบการถดถอยเชิงเส้นพหุคูณสำหรับพยากรณ์ค่าความเข้มข้นของ NO_2 ในเมืองพอร์ตแลนด์ ตัวแปรอิสระที่ใช้ คือ ความหนาแน่นของถนน ทางรถไฟ ปริมาณการจราจร และตัวแบบที่ได้มีสัมประสิทธิ์การตัดสินใจเท่ากับ 0.66

Chen et al. (2010) ศึกษาการใช้ตัวแบบ Land Use Regression (LUR) เพื่อพยากรณ์ความเข้มข้นของ NO_2 และ PM_{10} ในฤดูกาลที่แตกต่างกันในเทียนจินประเทศจีน โดยมีตัวแปรที่สำคัญคือ NO_2 และ PM_{10} ผลการศึกษาพบว่า การใช้ตัวแบบ LUR ในพื้นที่ศึกษาสำหรับ PM_{10} มีประสิทธิภาพดีในช่วงฤดูร้อนและให้ผลที่ดีกว่า NO_2

Molter, Lindley, Vocht, Simposn and Agius (2010) ได้สร้างตัวแบบสำหรับการพยากรณ์ค่าความเข้มข้นของ NO_2 และ PM_{10} ของเมือง Greater Manchester โดยใช้ตัวแบบ Land Use Regression มีตัวแปรอิสระคือตัวแปรเกี่ยวกับความหนาแน่นของการจราจร มลพิษทางอากาศ การใช้พื้นที่และตัวแปรทางด้านภูมิศาสตร์ ตัวแบบที่ได้มีค่าสัมประสิทธิ์การตัดสินใจเท่ากับ 0.71 และมีค่ารากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ค่อนข้างต่ำ

Wilton, Szpiro, Gould and Larson (2010) ได้สร้างตัวแบบ Land Use Regression (LUR) สำหรับการพยากรณ์ค่าความเข้มข้นของไนโตรเจนออกไซด์ไฮบริดใน Los Angeles, CA และซีแอตเทิล มีตัวแปรอิสระคือ NO_x และ NO_2 จากการศึกษาพบว่า ตัวแบบ LUR สามารถพยากรณ์ค่าความเข้มข้นของไนโตรเจนออกไซด์ไฮบริดได้ดี

Allen, Amramb, Wheeler and Brauer (2011) ศึกษาการใช้ตัวแบบ Land Use Regression (LUR) สำหรับการพยากรณ์ค่าความเข้มข้นของ NO และ NO_2 ระหว่างเมือง Winnipeg Manitoba และ Edmonton ตามการวัด 50 สถานีในช่วง 14 วัน ผลการศึกษาพบว่า ตัวแบบการ LUR มีประสิทธิภาพดี

Vlachogianni, Kassomenos, Karppinen, Karakitsios and Kukkonen (2011) ศึกษาการใช้ตัวแบบการถดถอยเชิงเส้นพหุคูณ (MLR) สำหรับการพยากรณ์ค่าความเข้มข้นของ NO_x และ PM_{10} ในกรุงเอเธนส์และเฮลซิงกิ มีตัวแปรอิสระดังนี้ คือ NO NO_2 NO_x CO O_3 $PM_{2.5}$ และ PM_{10} และตัวแปรทางอุตุนิยมวิทยา คือ อุณหภูมิห้อง ความเร็วลม ทิศทางและความชื้น ผลการศึกษาพบว่า ตัวแบบ MLR สามารถพยากรณ์ค่าความเข้มข้นของ NO_x และ PM_{10} ได้ดี

Basagana et al. (2012) ศึกษาตัวเลขของผลกระทบการวัดขนาดโดยใช้ตัวแบบ Land Use Regression (LUR) สำหรับพยากรณ์พื้นที่มลพิษทางอากาศ โดยวัด NO_2 จากทั้งหมด 148 สถานี ในเขตเมืองของเจโรนา ประเทศสเปน ผลการศึกษาพบว่า ตัวแบบ LUR สามารถอธิบายระดับมลพิษทางอากาศในพื้นที่ได้ โดยจะขึ้นอยู่กับตัวเลขที่มีขนาดใหญ่ของสถานีวัดมลพิษ

Chen, Wang, Li et al. (2012) ศึกษาการสร้างตัวแบบ Land Use Regression (LUR) จากข้อมูลอุตสาหกรรมที่เทียนจิน ประเทศจีน มีตัวแปรที่สนใจศึกษา คือ SO_2 NO_2 และ PM_{10} และมีตัวแปรอิสระ คือ ปริมาณการจราจร เครือข่ายถนน ข้อมูลการใช้พื้นที่ ความหนาแน่นของประชากร สภาพอุตุนิยมวิทยา สภาพทางกายภาพ ผลการศึกษาพบว่า ตัวแบบมีประสิทธิภาพดี

Gonzales et al. (2012) ศึกษาการประเมินผลของตัวแบบ Land Use Regression (LUR) สำหรับ NO_2 ใน El Paso, Texas, USA โดยจะประเมินแบบผสมเพื่ออธิบายการไล่ระดับสีทางด้านมิติสัมพันธ์ของ NO_2 ใน El Paso County Texas ตามการวัดเก็บรวบรวมในระหว่างฤดูกาลที่เย็น และอบอุ่นใน 2006-2007 และยังประเมินผลการดำเนินงานทั่วไปของสารเติมแต่งรูปแบบ (GAM) ที่พัฒนาขึ้นสำหรับเอลพาโซในปี 1999 เพื่อประเมินการไล่ระดับสีเชิงพื้นที่ที่ท่วมฉุกล ในปี 2006-2007 ในการศึกษาปัจจุบันเมื่อเทียบกับ 1999 การศึกษาผ่านแบบจำลอง LUR แสดงให้เห็นว่าตัวแปรทำนายมีอิทธิพลมากที่สุด

Li, Wu, Wilhelm and Ritz (2012) ศึกษาการปรับปรุงแบบจำลองการถดถอยเชิงเส้นพหุคูณสำหรับสารเติมแต่งมลพิษที่เกี่ยวข้องกับการจราจรโดยใช้แบบจำลองทั่วไป (GAM) รวมกับ cokriging ของเศษเหลือใช้อวกาศ โดยพัฒนาแบบจำลองการถดถอยเชิงเส้นสำหรับ NO_2 และไนโตรเจนออกไซด์ความเข้มข้น (NO_x) ในแคลิฟอร์เนีย ผลการศึกษาพบว่า GAM รวมกับ cokriging เมื่อเทียบกับอีก 4 วิธี ไม่สามารถปรับให้เหมาะสมกับรูปแบบและตัวแปรร่วมน้อยลงในตัวพยากรณ์ได้

Wang, Henderson, Sbihi, Allen and Brauer (2012) ได้สร้างตัวแบบสำหรับพยากรณ์มลพิษทางอากาศที่เกิดจากการจราจร โดยใช้ตัวแบบ Land Use Regression (LUR) สำหรับพยากรณ์ค่า NO และ NO_2 ผลการศึกษาพบว่า ตัวแบบ LUR สามารถที่จะให้ข้อมูลการประมาณที่เชื่อถือได้มากกว่า 7 ปี ในเมโทรแวนคูเวอร์

บทที่ 3

วิธีการดำเนินการศึกษา

การศึกษานี้มีวัตถุประสงค์เพื่อสร้างตัวแบบเพื่อพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี โดยมีวิธีการดำเนินการศึกษาตามขั้นตอนดังนี้

3.1 ข้อมูลที่นำมาศึกษา

ส่วนที่ 1 ข้อมูลความเข้มข้นของ NO_2 เฉลี่ยรายวัน ตั้งแต่ปี พ.ศ. 2549–2554 โดยได้รับความอนุเคราะห์จากกรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม

ส่วนที่ 2 ข้อมูลปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และทิศทางลม ตั้งแต่ปี พ.ศ. 2549–2554 เฉลี่ยรายวัน ได้รับความอนุเคราะห์จากกรมอุตุนิยมวิทยา กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร

หลังจากที่ได้ข้อมูลมาแล้ว มีวิธีจัดการข้อมูลดังนี้

1. เนื่องจากข้อมูลความเข้มข้นของ NO_2 ที่ได้มาเก็บข้อมูลเป็นรายชั่วโมง และความดันบรรยากาศ ความเร็วลม ทิศทางลม เก็บข้อมูลเป็นราย 3 ชั่วโมง จึงต้องนำข้อมูลมาเฉลี่ยเป็นข้อมูลรายวันก่อน ส่วนปริมาณน้ำฝน ความชื้น และอุณหภูมิ ที่ได้มาเก็บข้อมูลเป็นรายวันจึงไม่ต้องทำการปรับค่า
2. คัดเลือกข้อมูลของตัวแปรต่าง ๆ ให้มีวันที่เก็บข้อมูลตรงกัน
3. คัดเลือกข้อมูลโดยแบ่งเป็น 2 ชุด คือ ข้อมูลชุดที่ 1 ปี พ.ศ. 2549–2553 จะนำมาใช้สำหรับสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณ ข้อมูลชุดที่ 2 ปี พ.ศ. 2554 จะนำมาใช้เพื่อประเมินความสามารถในการพยากรณ์ตัวแบบการถดถอยเชิงเส้นพหุคูณที่ได้

3.2 การสร้างตัวแบบ

1. ศึกษาลักษณะของข้อมูลเบื้องต้นโดยสร้างฮิสโทแกรม และหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่ามัธยฐาน และค่าส่วนเบี่ยงเบนมาตรฐาน
2. ศึกษาลักษณะแนวโน้มความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระแต่ละตัว โดยพิจารณาจากกราฟการกระจาย
3. สร้างตัวแบบการถดถอยเชิงเส้นพหุคูณ โดยใช้ข้อมูลของปี พ.ศ. 2549–2553 โดยได้ตัวแบบการถดถอยดังนี้

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \varepsilon_i \quad (1)$$

เมื่อ y_i คือ ค่าสังเกตที่ i ของความเข้มข้นของ NO_2 (part per billion : ppb)

x_{i1} คือ ค่าสังเกตที่ i ของปริมาณน้ำฝน (มิลลิเมตร)

x_{i2} คือ ค่าสังเกตที่ i ของความชื้นสัมพัทธ์ (เปอร์เซ็นต์)

x_{i3} คือ ค่าสังเกตที่ i ของอุณหภูมิ (องศาเซลเซียส)

x_{i4} คือ ค่าสังเกตที่ i ของความดันบรรยากาศ (เฮกโตปาสกาล)

x_{i5} คือ ค่าสังเกตที่ i ของความเร็วลม (กิโลเมตรต่อชั่วโมง)

x_{i6} คือ ค่าสังเกตที่ i ของทิศทางลม (องศา)

ε_i คือ ค่าคลาดเคลื่อนของค่าสังเกตที่ i

4. ตรวจสอบข้อสมมุติของตัวแบบการถดถอย

1. ค่าคลาดเคลื่อน (ε_i) ต้องมีการแจกแจงปกติ ตรวจสอบโดยใช้ Normal Probability Plot ของค่าส่วนตกค้าง ($\hat{\varepsilon}_i$) และค่าสถิติทดสอบ Kolmogorov-Smirnov
2. ค่าคลาดเคลื่อน (ε_i) ต้องมีความแปรปรวนคงที่ ตรวจสอบโดยพิจารณาจากกราฟที่พล็อตระหว่าง $\hat{\varepsilon}_i$ กับ ค่าพยากรณ์ ($\hat{y}_i$) หรือ $\hat{\varepsilon}_i$ กับ x_j และสถิติทดสอบ Breusch-Pagan
3. ค่าคลาดเคลื่อน (ε_i) ต้องเป็นอิสระกัน ตรวจสอบโดยพิจารณาจากกราฟที่พล็อตระหว่าง $\hat{\varepsilon}_i$ กับ ลำดับเวลาของการเก็บข้อมูล และค่าสถิติทดสอบ Durbin Watson
4. ตัวแปรอิสระในตัวแบบการถดถอยต้องเป็นอิสระกัน ตรวจสอบโดยพิจารณาจากกราฟการกระจาย ที่พล็อตระหว่างตัวแปรอิสระแต่ละคู่ และพิจารณาจากค่า VIF
5. คัดเลือกตัวแบบด้วยวิธีลดตัวแปร (Backward Elimination)
6. ประเมินคุณภาพของตัวแบบที่ได้โดยพิจารณาจากค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้ว (R_a^2) สัมประสิทธิ์การตัดสินใจ (R^2) ค่าคลาดเคลื่อนมาตรฐานของการประมาณ (s)

3.3 ความสามารถในการพยากรณ์ของตัวแบบการถดถอย

จากตัวแบบการถดถอยที่ได้ นำมาพยากรณ์ค่าความเข้มข้นของ NO_2 ในปี พ.ศ. 2554 โดยใช้ข้อมูลของตัวแปรอิสระ คือ ปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และทิศทางลม ในปี พ.ศ. 2554 นำค่าพยากรณ์ที่ได้มาเปรียบเทียบกับค่าที่แท้จริง และพิจารณาความสามารถในการพยากรณ์ของตัวแบบการถดถอยโดยพิจารณาจากค่ารากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error : RMSE) สามารถคำนวณได้ดังสมการ

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

เมื่อ y_i คือ ค่าสังเกตที่ i ของตัวแปรตาม

$\hat{y}_i$ คือ ค่าพยากรณ์ที่ i ของตัวแปรตาม

n คือ จำนวนค่าสังเกตทั้งหมด

บทที่ 4

ผลการศึกษา

ในการศึกษาการสร้างตัวแบบเพื่อพยากรณ์ค่าความเข้มข้นของ NO_2 โดยการสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณ โดยมีตัวแปรตาม คือ $NO_2(y)$ และตัวแปรอิสระ คือ ปริมาณน้ำฝน (x_1) ความชื้นสัมพัทธ์ (x_2) อุณหภูมิ (x_3) ความดันบรรยากาศ (x_4) ความเร็วลม (x_5) และทิศทางลม (x_6) ได้ผลการศึกษาดังนี้

4.1 ข้อมูลที่นำมาศึกษา

1. ข้อมูลความเข้มข้นของ NO_2 ตั้งแต่ปี พ.ศ. 2549–2554 มีทั้งหมด 1,960 ค่าสังเกต และตัวแปรอิสระ คือ ปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และทิศทางลม ตั้งแต่ปี พ.ศ. 2549–2554 มีทั้งหมด 2,188 ค่าสังเกต
2. จากการคัดเลือกข้อมูลของตัวแปรต่าง ๆ ให้มีวันที่เก็บข้อมูลตรงกัน มีทั้งหมด 1,960 ค่าสังเกต
3. ข้อมูลชุดที่ 1 ปี พ.ศ. 2549–2553 มีทั้งหมด 1,627 ค่าสังเกต สำหรับการสร้างตัวแบบการถดถอย ข้อมูลชุดที่ 2 ปี พ.ศ. 2554 มีทั้งหมด 333 ค่าสังเกต สำหรับประเมินความสามารถการพยากรณ์ตัวแบบที่ได้

4.2 การสร้างตัวแบบ

1. การศึกษาลักษณะของข้อมูลเบื้องต้นโดยการหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่ามัธยฐาน และค่าส่วนเบี่ยงเบนมาตรฐาน และสร้างฮิสโทแกรม ได้ผลการศึกษาดังนี้

ตารางที่ 1 ค่าเฉลี่ย ค่ามัธยฐาน และค่าเบี่ยงเบนมาตรฐาน ของ NO_2 และตัวแปรอิสระแต่ละตัว

ตัวแปร	ค่าเฉลี่ย	มัธยฐาน	ค่าส่วนเบี่ยงเบนมาตรฐาน
NO_2	15.35	13.61	6.89
ปริมาณน้ำฝน	3.32	0.00	9.56
ความชื้น	71.45	72.00	8.41
อุณหภูมิ	28.83	29.00	1.61
ความดันบรรยากาศ	1009.10	1008.80	2.60
ความเร็วลม	1.67	1.00	1.16
ทิศทางลม	71.48	63.75	44.00

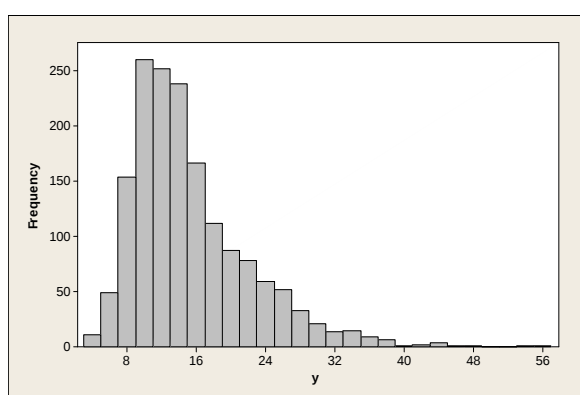
จากตารางที่ 1 พบว่าค่าความเข้มข้นของ NO_2 มีค่าเฉลี่ย เท่ากับ 15.35 มัธยฐาน เท่ากับ 13.61 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.89 ปริมาณน้ำฝน มีค่าเฉลี่ย เท่ากับ 3.32 มัธยฐาน เท่ากับ 0.00 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 9.56 ความชื้น มีค่าเฉลี่ย เท่ากับ 71.45 มัธยฐาน เท่ากับ 72.00 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 8.41 อุณหภูมิ มีค่าเฉลี่ย เท่ากับ 28.83 มัธยฐาน เท่ากับ 29.00 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.61 ความดันบรรยากาศ มีค่าเฉลี่ย เท่ากับ 1009.10 มัธยฐาน เท่ากับ 1008.80 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.60 ความเร็วลม มีค่าเฉลี่ย เท่ากับ 1.67 มัธยฐาน เท่ากับ 1.00 และส่วน

เบี่ยงเบนมาตรฐาน เท่ากับ 1.16 และทิศทางลม มีค่าเฉลี่ย เท่ากับ 71.48 มัธยฐาน เท่ากับ 63.75 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 44.00

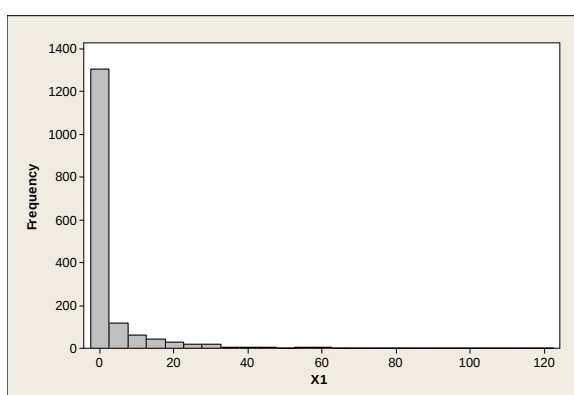
เมื่อพิจารณากราฟฮิสโทแกรมจากภาพที่ 1 พบว่า ค่าความเข้มข้นของ NO_2 และตัวแปรอิสระแต่ละตัว ไม่มีลักษณะของการแจกแจงปกติ โดยจะพบว่าการแจกแจงความน่าจะเป็นความเข้มข้นของ NO_2 ปริมาณน้ำฝน ความดันบรรยากาศ ความเร็วลม และทิศทางลม มีลักษณะการแจกแจงความน่าจะเป็นแบบเบ้ขวา และความชื้นสัมพัทธ์ อุณหภูมิ มีลักษณะการแจกแจงความน่าจะเป็นแบบเบ้ซ้าย

เนื่องจากค่าความเข้มข้นของ NO_2 ไม่ได้มีการแจกแจงปกติ ซึ่งไม่เป็นไปตามข้อสมมุติของตัวแบบการถดถอย จึงทำการแปลงค่าความเข้มข้นของ NO_2 ให้อยู่ในรูปของลอการิทึมธรรมชาติ โดยกำหนดให้เป็น

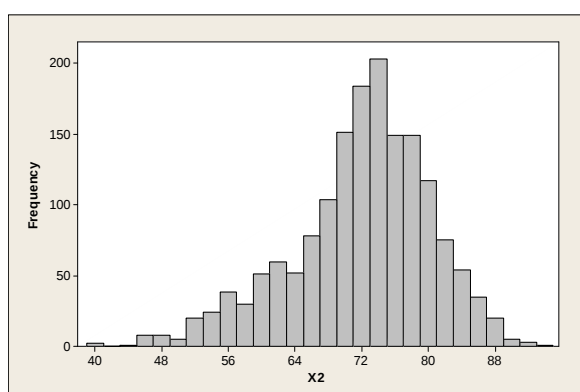
$$y^* = \ln NO_2$$



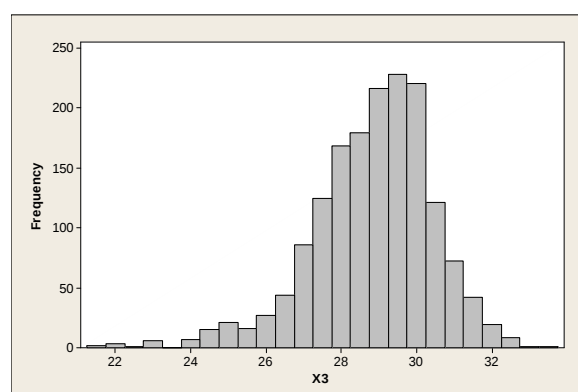
(ก) ความเข้มข้นของ NO_2



(ข) ปริมาณน้ำฝน

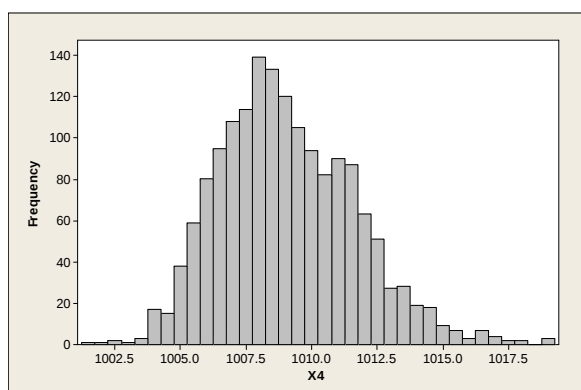


(ค) ความชื้นสัมพัทธ์

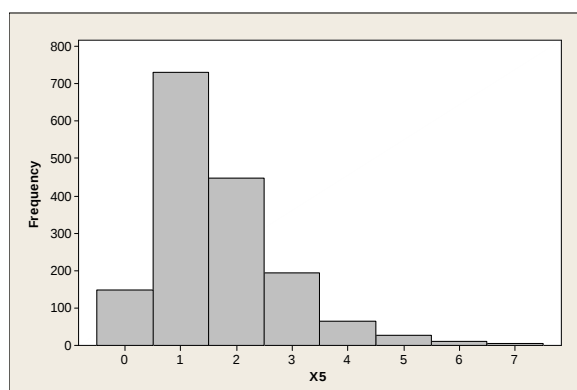


(ง) อุณหภูมิ

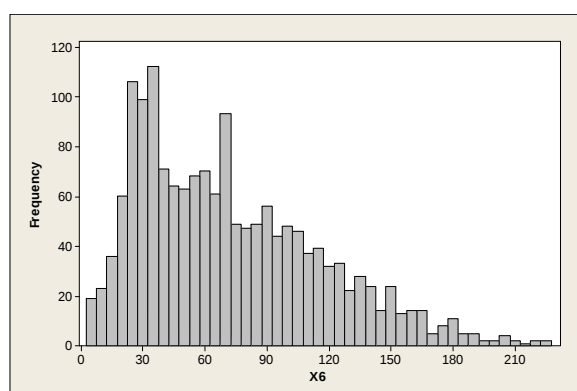
ภาพที่ 1 กราฟฮิสโทแกรมของตัวแปรต่าง ๆ



(จ) ความดันบรรยากาศ



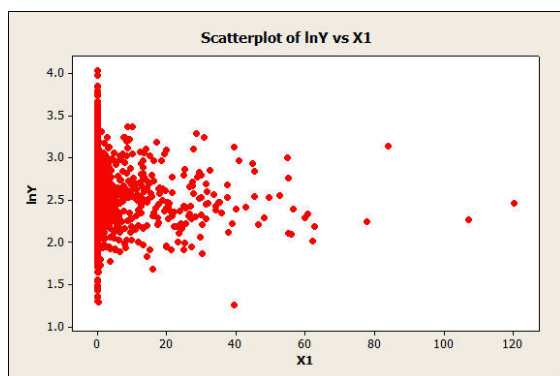
(ฉ) ความเร็วลม



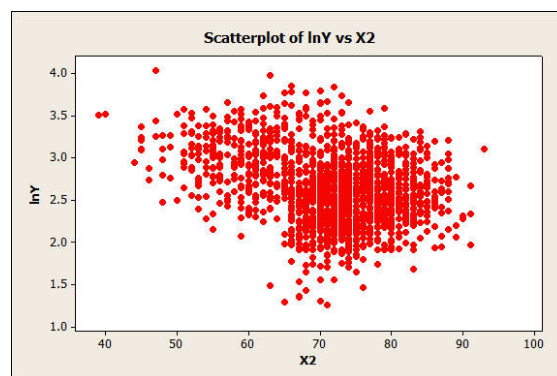
(ช) ทิศทางลม

ภาพที่ 1(ต่อ) กราฟฮิสโทแกรมของตัวแปรต่าง ๆ

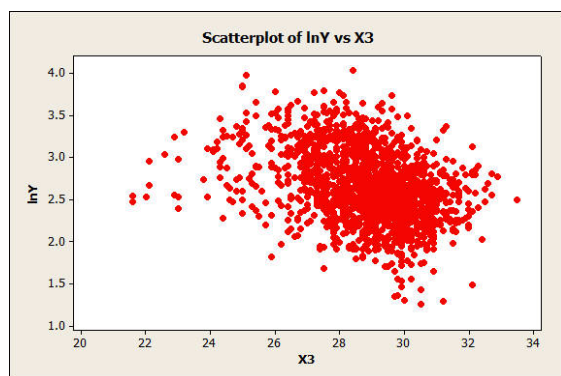
2. ศึกษาลักษณะแนวโน้มของความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระแต่ละตัว
เมื่อพิจารณากราฟการกระจาย จากภาพที่ 2 พบว่า $\ln NO_2$ มีความสัมพันธ์ที่มีแนวโน้มเชิงเส้นตรง
กับ ความชื้นสัมพัทธ์ อุณหภูมิ และความดันบรรยากาศ



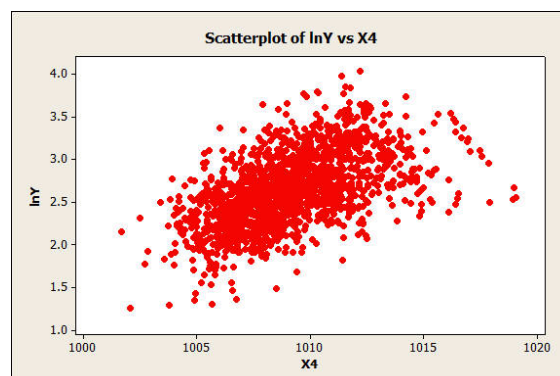
(ก) $\ln NO_2$ และ ปริมาณน้ำฝน



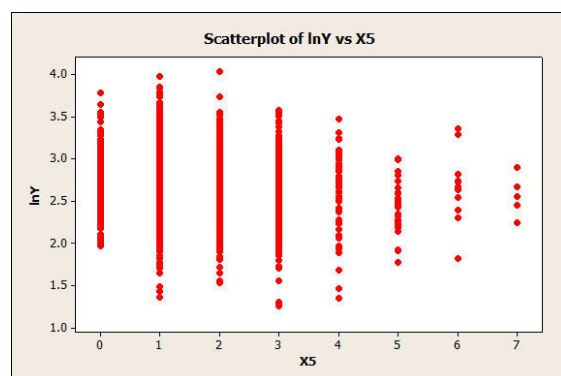
(ข) $\ln NO_2$ และ ความชื้นสัมพัทธ์



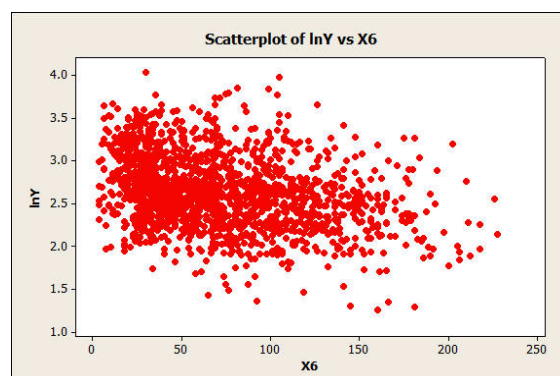
(ค) $\ln NO_2$ และ อุณหภูมิ



(ง) $\ln NO_2$ และ ความดันบรรยากาศ



(จ) $\ln NO_2$ และ ความเร็วลม



(ฉ) $\ln NO_2$ และทิศทางลม

ภาพที่ 2 กราฟการกระจายระหว่าง $\ln y$ และตัวแปรอิสระแต่ละตัว

3. สร้างตัวแบบการถดถอยเชิงเส้นพหุคูณโดยใช้ข้อมูลของปี พ.ศ. 2549–2553 ได้ตัวแบบการถดถอยดังนี้

$$y_i^* = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \varepsilon_i \quad (3)$$

และได้สมการถดถอยที่ประมาณค่าได้ จากข้อมูลในปี พ.ศ. 2549-2553 คือ

$$\begin{aligned} \hat{y}_i^* = & -68.2 - 0.00161x_{i1} - 0.0113x_{i2} - 0.0373x_{i3} - 0.0723x_{i4} - 0.123x_{i5} \\ & + 0.000024x_{i6} \end{aligned} \quad (4)$$

โดยสมการถดถอยนี้มีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 38.70 สัมประสิทธิ์การตัดสินใจเท่ากับ 38.90 ค่าคลาดเคลื่อนมาตรฐานของการถดถอยเท่ากับ 1.388 ppb

4. ตรวจสอบข้อสมมุติของตัวแบบการถดถอย

1. การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน

เมื่อพิจารณาจากภาพที่ 3 (ก) พบว่าค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ เนื่องจาก Normal Probability Plot ของ ε_i มีลักษณะใกล้เคียงเส้นตรง ซึ่งสอดคล้องกับค่าสถิติทดสอบ Kolmogorov-Smirnov มีค่า P-Value มากกว่า 0.150 ซึ่งมากกว่าระดับนัยสำคัญ 0.05

2. การตรวจสอบความคงที่ของความแปรปรวนของค่าคลาดเคลื่อน

เมื่อพิจารณาจากภาพที่ 3 (ข) พบว่าค่าคลาดเคลื่อนมีความแปรปรวนคงที่ เนื่องจากกราฟการกระจายระหว่าง ε_i กับ $\hat{y}_i^*$ มีการกระจายอย่างสม่ำเสมอรอบศูนย์ ซึ่งสอดคล้องกับการทดสอบโดยวิธีของ Breusch-Pagan (ดูรายละเอียดได้ในภาคผนวก หน้า 22)

3. การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

เมื่อพิจารณาจากภาพที่ 3 (ค) พบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกัน เนื่องจากกราฟการกระจายระหว่าง ε_i กับ ลำดับการเวลาเก็บข้อมูล มีลักษณะการกระจายแบบมีรูปแบบเป็นแบบวัฏจักร (Cycle) ซึ่งสอดคล้องกับค่าสถิติทดสอบ Durbin-Watson ซึ่งมีค่าเท่ากับ 0.710 ซึ่งน้อยกว่าค่าวิกฤติ 1.910 (ดูรายละเอียดได้ในภาคผนวก หน้า 23)

4. การตรวจความสัมพันธ์ของตัวแปรอิสระ

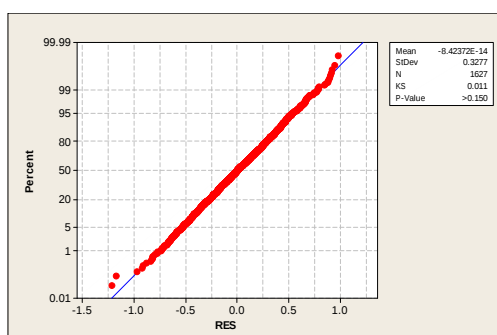
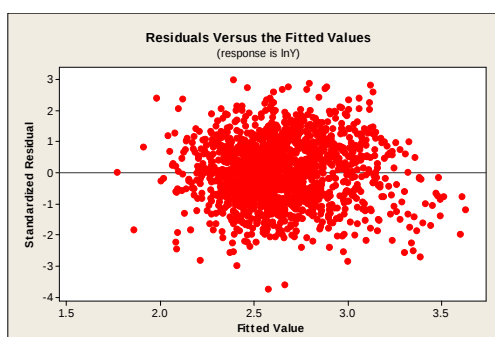
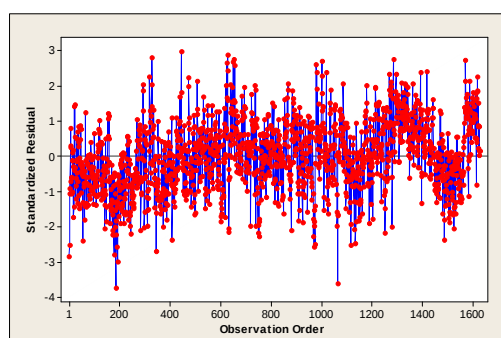
เมื่อพิจารณาจากตารางที่ 2 และตารางที่ 3 พบว่าตัวแปรอิสระมีความสัมพันธ์กันในระดับต่ำ เนื่องจากค่า VIF ไม่มีค่าใดที่มากกว่า 10 และค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระแต่ละคู่ มีค่าน้อยกว่า 0.5

ตารางที่ 2 ค่า VIF ของตัวแปรอิสระแต่ละตัว

ตัวแปรอิสระ	VIF
ปริมาณน้ำฝน	1.1
ความชื้นสัมพัทธ์	1.7
อุณหภูมิ	1.8
ความดันบรรยากาศ	2.1
ความเร็วลม	1.7
ทิศทางลม	1.5

ตารางที่ 3 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ

ตัวแปรอิสระ	x_1	x_2	x_3	x_4	x_5
x_2	0.317				
x_3	-0.049	0.115			
x_4	-0.173	-0.459	-0.578		
x_5	-0.146	-0.462	-0.213	0.287	
x_6	-0.022	0.134	0.368	-0.376	0.243

(ก) Normal Probability Plot ของ $\hat{\varepsilon}_i$ (ข) $\hat{\varepsilon}_i$ กับ $\hat{y}_i^*$ (ค) $\hat{\varepsilon}_i$ กับ ลำดับเวลาการเก็บข้อมูลภาพที่ 3 กราฟส่วนตกค้าง ($\hat{\varepsilon}_i$) ของสมการถดถอยที่ (4)

จากการตรวจสอบข้อสมมุติข้างต้นสรุปได้ว่า ค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ และมีค่าความแปรปรวนคงที่ โดยที่ตัวแปรอิสระในแบบจำลองการถดถอยมีความสัมพันธ์กันในระดับต่ำ แต่พบว่า ค่าคลาดเคลื่อนไม่เป็นอิสระกัน ดังนั้น จึงทำการแก้ไขปัญหานี้ด้วยการเพิ่มตัวแปรอิสระเข้าไปในแบบจำลองถดถอยอีกหนึ่งตัว คือ ค่าความเข้มข้นของ NO_2 ย้อนหลังไปหนึ่งหน่วยเวลา $NO_{2(t-1)}$ เป็น x_7 และแปลงรูป $NO_{2(t-1)}$ ให้อยู่ในรูปของลอการิทึมธรรมชาติ โดยกำหนดให้เป็น

$$x_7^* = \ln NO_{2(t-1)}$$

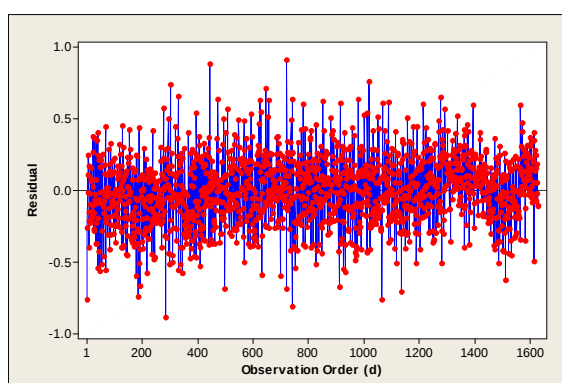
ทำการสร้างแบบจำลองถดถอยใหม่อีกครั้ง
ได้สมการการถดถอยที่ประมาณค่าได้จากข้อมูลในปี พ.ศ. 2549-2553 คือ

$$\begin{aligned} \hat{y}_i^* = & -22.9 - 0.000261x_{i1} - 0.00570x_{i2} - 0.0263x_{i3} + 0.0249x_{i4} - 0.0665x_{i5} \\ & + 0.000053x_{i6} + 0.644x_{i7}^* \end{aligned} \quad (5)$$

โดยสมการถดถอยนี้มีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 67.20 สัมประสิทธิ์การตัดสินใจเท่ากับ 67.30 ค่าคลาดเคลื่อนมาตรฐานของการถดถอยเท่ากับ 1.271 ppb

โดยทำการตรวจสอบข้อสมมุติความเป็นอิสระกันของค่าคลาดเคลื่อนใหม่อีกครั้ง

เมื่อพิจารณาจากภาพที่ 4 พบว่าค่าคลาดเคลื่อนเป็นอิสระกัน เนื่องจากกราฟการกระจายระหว่าง $\hat{\epsilon}_i$ กับ ลำดับการเวลาเก็บข้อมูล มีลักษณะการกระจายสุ่มรอบศูนย์ ซึ่งสอดคล้องกับค่าสถิติทดสอบ Durbin-Watson ที่มีค่าเท่ากับ 1.990 ซึ่งน้อยกว่าค่าวิกฤติ 1.910 (ดูรายละเอียดได้ในภาคผนวก หน้า 24)



ภาพที่ 4 กราฟส่วนตกค้างที่พล็อตระหว่าง $\hat{\epsilon}_i$ กับ ลำดับเวลาการเก็บข้อมูล ของสมการถดถอยที่ (5)

5. การคัดเลือกตัวแบบด้วยวิธีลดตัวแปร (Backward Elimination)

จากการคัดเลือกตัวแบบด้วยวิธีลดตัวแปร พบว่าตัวแปรอิสระที่เหมาะสมสำหรับตัวแบบที่ใช้ในการพยากรณ์ค่าความเข้มข้นของ NO_2 มี 5 ตัวแปรอิสระ คือ ความชื้นสัมพัทธ์ (x_2) อุณหภูมิ (x_3) ความดันบรรยากาศ (x_4) ความเร็วลม (x_5) และ $NO_{2(t-1)}$ (x_7^*)

โดยมีสมการถดถอยที่ประมาณค่าได้ คือ

$$\hat{y}_i^* = -22.8 - 0.00571x_{i2} - 0.0256x_{i3} + 0.0248x_{i4} - 0.0655x_{i5} + 0.645x_{i7}^* \quad (6)$$

โดยสมการถดถอยนี้มีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้ว เท่ากับ 67.20 สัมประสิทธิ์การตัดสินใจ เท่ากับ 67.30 ค่าคลาดเคลื่อนมาตรฐานของการถดถอย เท่ากับ 1.271 ppb

6. การศึกษาประสิทธิภาพของตัวแบบ

จากตัวแบบการถดถอยที่ได้มีตัวแปรอิสระทั้งหมด 5 ตัวแปร คือ ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และ $NO_2(t-1)$ โดยสมการถดถอยนี้มีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้ว เท่ากับ 67.20 สัมประสิทธิ์การตัดสินใจเท่ากับ 67.30 ค่าคลาดเคลื่อนมาตรฐานของการถดถอยเท่ากับ 1.271 ppb

4.3 ความสามารถในการพยากรณ์ของตัวแบบการถดถอย

จากตัวแบบการถดถอยที่ได้มีตัวแปรอิสระทั้งหมด 5 ตัวแปร คือ ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และ $NO_2(t-1)$ และได้สมการถดถอยดังสมการที่ (6) นำสมการถดถอยที่ (6) มาพยากรณ์ค่าความเข้มข้น NO_2 ในปี พ.ศ. 2554 โดยใช้ค่าของตัวแปรอิสระทั้ง 5 ตัวแปร ในปี พ.ศ. 2554 ได้ค่า RMSE เท่ากับ 3.544 ppb

บทที่ 5

สรุปและอภิปรายผลการศึกษา

ในการศึกษาการสร้างตัวแบบสำหรับพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี โดยการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ สามารถสรุปผลการศึกษาดังนี้

5.1 สรุปและอภิปรายผลการศึกษา

จากการสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณ เพื่อการพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี โดยใช้ข้อมูลในปี พ.ศ. 2549 – 2554 ณ สถานีสำนักงานสามัญศึกษาชลบุรี ของกรมควบคุมมลพิษ และสถานีชลบุรีของกรมอุตุนิยมวิทยา โดยมีแปรรูปเป็นตัวแปรทางด้านอุตุนิยมวิทยา ได้ตัวแบบการถดถอยที่เหมาะสมที่ประกอบด้วยตัวแปรอิสระทั้งหมด 5 ตัวแปร ได้แก่ ความชื้นสัมพัทธ์ อุณหภูมิ ความดันบรรยากาศ ความเร็วลม และ $NO_2(t-1)$ โดยสมการถดถอยที่ประมาณได้มีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 67.20 สัมประสิทธิ์การตัดสินใจเท่ากับ 67.30 ค่าคลาดเคลื่อนมาตรฐานของการประมาณเท่ากับ 1.271 ppb และเมื่อนำสมการถดถอยที่ได้ไปพยากรณ์ค่าความเข้มข้นของ NO_2 ในจังหวัดชลบุรี ในปี พ.ศ. 2554 โดยใช้ค่าของตัวแปรอิสระที่อยู่ในตัวแบบที่ได้ ในปี พ.ศ. 2554 ได้ค่า RMSE เท่ากับ 3.544 ppb ซึ่งสอดคล้องกับการศึกษาของ (Vlachogianni et al., 2011) ศึกษาการใช้ตัวแบบการถดถอยเชิงเส้นพหุคูณ (MLR) สำหรับการพยากรณ์ค่าความเข้มข้นของ NO_x และ PM_{10} ในกรุงเอเธนส์และเฮลซิงกิ ได้ตัวแปรอิสระที่เหมาะสมคือ ความเร็วลม และมีค่าสัมประสิทธิ์การตัดสินใจอยู่ระหว่าง 0.77 ถึง 0.84

จะเห็นว่าตัวแบบการถดถอยเชิงเส้นตรงพหุคูณที่ได้นี้ยังเป็นตัวแบบที่ไม่มีความน่าเชื่อถือมากนัก เนื่องจากมีค่าสัมประสิทธิ์การตัดสินใจค่อนข้างต่ำ และมีค่า RMSE ค่อนข้างสูง ซึ่งอาจมีสาเหตุมาจากการที่ตัวแบบการถดถอยขาดตัวแปรอิสระที่มีความสำคัญบางตัวแปรไป และอาจเกิดจากการที่ข้อมูลที่นำมาศึกษาในครั้งนี้ได้มาจากแหล่งที่มาที่แตกต่างกัน โดยที่ NO_2 นั้นได้จากสถานีสำนักงานสามัญศึกษาชลบุรี ของกรมควบคุมมลพิษ ในขณะที่ตัวแปรอิสระทั้งหมดได้มาจากสถานีชลบุรี ของกรมอุตุนิยมวิทยา ซึ่งมีที่ตั้งอยู่ในพื้นที่ที่ต่างกัน นอกจากนี้ ข้อมูลที่ศึกษาเป็นข้อมูลรายวัน ซึ่งควรมีการวัดค่าทุกวันในหนึ่งปี แต่ผู้วิจัยพบว่ามีข้อมูลขาดหายไปจำนวนหนึ่งในหนึ่งปี ทำให้ข้อมูลที่ศึกษาขาดความสมบูรณ์และต่อเนื่อง

5.2 ข้อเสนอแนะ

1. แนวคิดในการสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณในการพยากรณ์เกี่ยวกับมลพิษทางอากาศนี้สามารถนำไปประยุกต์กับมลพิษทางอากาศชนิดอื่น ๆ และใช้ในพื้นที่อื่น ๆ ของประเทศไทยได้
2. ควรมีการศึกษาถึงตัวแปรอื่น ๆ ที่อาจมีความสัมพันธ์กับมลพิษทางอากาศ เช่น ตัวแปรทางด้านภูมิศาสตร์ ตัวแปรทางด้านมลพิษทางอากาศอื่น ๆ และตัวแปรทางด้านอุตุนิยมวิทยา ที่ยังขาดไปในตัวแบบการถดถอย เพื่อเพิ่มความสามารถในการพยากรณ์ของตัวแบบการถดถอย

บรรณานุกรม

- กองอนามัยสิ่งแวดล้อม. (2550). *มลพิษทางอากาศ*. วันที่ค้นข้อมูล 13 พฤศจิกายน 2555, เข้าถึงได้จาก <http://www.school.net.th/library/snet6/envi3/monpit-a/monpit-a.htm>
- กรมควบคุมมลพิษ. (2555). *ข้อมูลคุณภาพอากาศรายวันในภาคต่างๆของประเทศไทย*. วันที่ค้นข้อมูล 13 พฤศจิกายน 2555, เข้าถึงได้จาก <http://www.pcd.go.th/airquality/regional/defaultthai.cfm>
- กรมควบคุมมลพิษ. (2555). *ประกาศคณะกรรมการสิ่งแวดล้อมแห่งชาติฉบับที่ 33 พ.ศ. 2552*. วันที่ค้นข้อมูล 11 มีนาคม 2556, เข้าถึงได้จาก <http://www.pcd.go.th/download/regulation.cfm?task=s2>
- กรมอุตุนิยมวิทยา. (2555). *พยากรณ์อากาศรายวัน*. วันที่ค้นข้อมูล 13 กุมภาพันธ์ 2556, เข้าถึงได้จาก <http://www.tmd.go.th/index.php>
- การกำหนดค่ามาตรฐานก๊าซไนโตรเจนไดออกไซด์ในบรรยากาศทั่วไป ในเวลา 1 ปี. (2550). วันที่ค้นข้อมูล 13 พฤศจิกายน 2555, เข้าถึงได้จาก http://s3.amazonaws.com/zanran_storage/www.aqnis.pcd.go.th/ContentPages/47292663.pdf
- ทรงศิริ แต่สมบัติ. (2548). *การวิเคราะห์การถดถอย (พิมพ์ครั้งที่ 3)*. กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์
- ศรัลย์ ปานศรีพงษ์. (2550). *การพัฒนาแบบจำลองโครงข่ายประสาทประดิษฐ์และแบบจำลองความถดถอยเชิงพหุ เพื่อทำนายความเข้มข้นของ PM₁₀ ในพื้นที่กรุงเทพมหานครและเทศบาลนครราชสีมา*. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิศวกรรมสิ่งแวดล้อม, มหาวิทยาลัยเทคโนโลยีสุรนารี.
- Allen, R. W., Amram, O., Wheeler, A. J., & Brauer, M. (2011). The transferability of *NO* and *NO₂* land use regression models between cities and pollutants. *Atmospheric Environment*, 45, 369-378.
- Chen, L., Bai, Z., Kong, S., Han, Bin., You, Y., Ding, X., Du, S., & Liu, A. (2010). A land use regression for predicting *NO₂* and *PM₁₀* concentrations in different seasons in Tianjin region, China. *Journal of Environmental Sciences*, 22(9), 1364-1373.
- Chen, L., Wang, Y., Li, P., Ji, Y., Kong, S., Li, Z., & Bai, Z. (2010). A land use regression model incorporating data on industrial point source pollution. *Journal of Environmental Sciences*, 24(7), 1251-1258.
- Fuller, G. W., & Green, D. (2006). Evidence for increasing concentrations of primary *PM₁₀* in London. *Atmospheric Environment*, 40, 6134-6145.
- Gonzales, M., Myers, O., Smith, L., Olvera, H. A., Mukerjee, S., Li W. W., Pingitore, N., Amaya, M., Burchiel, S., & Berwick, M. (2012). Evaluation of land use regression models for *NO₂* in El Paso, Texas, USA. *Science of the total environment*, 432, 135-142.
- Li, L., Wu, J., Wilhelm, M., & Ritz, B. (2012). Use of generalized additive models and cokriging of spatial residuals to improve land-use regression estimates of nitrogen oxides in Southern California. *Atmospheric Environment*, 55, 220-228.
- Mavko, M. E., Tang, B., & George, L. A. (2008). A sub-neighborhood scale land use regression model for predicting *NO₂*. *Science of the total environment*, 398, 68-75.

- Vlachogianni, A., Kassomenos, P., Karppinen, A., Karakitsios S. & Jaakko, K. (2011). Evaluation of a multiple regression model for the forecasting of the concentrations of NO_x and PM_{10} in Athens and Helsinki. *Science of the Total Environment*, 409, 1559–1571.
- Wang, R., Henderson, S. B., Sbihi, H., Allen, R. W., & Brauer, M., (2013). Temporal stability of land use regression models for traffic-related air pollution. *Atmospheric Environment*, 64, 312-319.
- Wilton, D., Szpiro, A., Gould, T., & Larson, T. (2010). Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression in Los Angeles, CA and Seattle, WA. *Science of the total environment*, 408, 1120–1130.

ภาคผนวก

การทดสอบค่าคงที่ของความแปรปรวนของค่าคลาดเคลื่อน โดย Breusch-Pagan Test

Anova Table NO_2 กับตัวแปรอิสระ

Source	DF	SS	MS	F	P
Regression	6	111.380	18.563	172.24	0.000
Residual Error	1620	174.595	0.108		
Total	1626	285.975			

Anova Table $\hat{\varepsilon}^2$ กับตัวแปรอิสระ

Source	DF	SS	MS	F	P
Regression	6	0.78763	0.13127	5.92	0.000
Residual Error	1620	35.93734	0.02218		
Total	1626	36.72498			

สมมติฐาน

H_0 : ค่าความแปรปรวนของค่าคลาดเคลื่อนคงที่

H_a : ค่าความแปรปรวนของค่าคลาดเคลื่อนไม่คงที่

กำหนดระดับนัยสำคัญ 0.05

$$\chi_{BP}^2 = \frac{SSR^*}{8} \div \left(\frac{SSE}{n} \right)^2$$

$$= \frac{0.78763}{7} \div \left(\frac{174.595}{1627} \right)^2$$

$$\chi_{BP}^2 = 9.7709$$

$$\chi_{0.05,7}^2 = 14.067$$

เนื่องจาก $\chi_{BP}^2 = 9.7709$ น้อยกว่า $\chi_{0.05,7}^2 = 14.067$ จึงยอมรับ H_0

ดังนั้น ความแปรปรวนของค่าคลาดเคลื่อนมีค่าคงที่ ที่ระดับนัยสำคัญ 0.05

การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

1. ใช้ข้อมูล $\ln NO_2$ กับ ตัวแปรอิสระแต่ละตัว

สมมติฐาน

H_0 : ความคลาดเคลื่อนไม่มีความสัมพันธ์กัน

H_a : ความคลาดเคลื่อนมีความสัมพันธ์กันในทางบวก

กำหนดระดับนัยสำคัญ 0.05

กำหนดให้ $N = 1627$, $dl = 1.91298$, $du = 1.92514$, $d = 0.710$

เนื่องจาก $d = 0.710 < dl = 1.91298$ จึงปฏิเสธ H_0

สรุป ความคลาดเคลื่อนมีความสัมพันธ์กัน

การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

2. ใช้ข้อมูล $\ln NO_2$ กับ ตัวแปรอิสระแต่ละตัว โดยเพิ่ม $\ln NO_{2(t-1)}$

สมมติฐาน

H_0 : ความคลาดเคลื่อนไม่มีความสัมพันธ์กัน

H_a : ความคลาดเคลื่อนมีความสัมพันธ์กันในทางบวก

กำหนดระดับนัยสำคัญ 0.05

กำหนดให้ $N = 1627$, $dl = 1.91298$, $du = 1.92514$, $d = 0.710$

เนื่องจาก $d = 1.990 > dl = 1.91298$ จึงยอมรับ H_0

สรุป ความคลาดเคลื่อนไม่มีความสัมพันธ์กัน