


การประมาณค่าการระเหยน้ำโดยใช้ตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่

ชัญญาภรณ์ แสงประสาน

ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
พฤษภาคม 2558
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

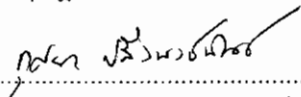
คณะกรรมการควบคุมคุณวุฒินิพนธ์และคณะกรรมการสอบคุณวุฒินิพนธ์ ได้พิจารณา
คุณวุฒินิพนธ์ของ ชนัญญาญจน์ แสงประสาน ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพา
ได้


คณะกรรมการควบคุมคุณวุฒินิพนธ์

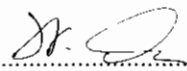

..... อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร.เสรี ชัดเข้ม)

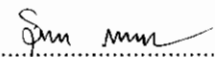

..... อาจารย์ที่ปรึกษาร่วม
(ดร.พัชรี วงษ์เกษม)

คณะกรรมการสอบคุณวุฒินิพนธ์


..... ประธาน
(ผู้ช่วยศาสตราจารย์ ดร.กศยา ปลั่งพงษ์พันธ์)


..... กรรมการ
(รองศาสตราจารย์ ดร.เสรี ชัดเข้ม)


..... กรรมการ
(ดร.พัชรี วงษ์เกษม)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุชาดา กรเพชรปानी)

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญาอนุมัติให้รับคุณวุฒินิพนธ์ฉบับนี้เป็นส่วนหนึ่งของ
การศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา
ของมหาวิทยาลัยบูรพา


..... คณบดีวิทยาลัยวิทยาการวิจัย
และวิทยาการปัญญา

(ผู้ช่วยศาสตราจารย์ ดร.สุชาดา กรเพชรปानी)
วันที่ 16 เดือน กรกฎาคม พ.ศ. 2558

ประกาศคุณูปการ

คุณุภินิพนธ์ฉบับนี้สำเร็จได้ด้วยความกรุณาจากคณะกรรมการควบคุมคุณุภินิพนธ์ รองศาสตราจารย์ ดร.เสรี ชัดรัมย์ อาจารย์ที่ปรึกษาหลัก และ ดร.พัชรี วงษ์เกษม อาจารย์ที่ปรึกษา ร่วม ตลอดจนคณะกรรมการสอบคุณุภินิพนธ์ทุกท่าน ที่กรุณาให้คำปรึกษาแนะนำแนวทางที่ถูกต้อง แก้ไขข้อบกพร่องต่าง ๆ ด้วยความละเอียดถี่ถ้วน เอาใจใส่ด้วยดีเสมอมา ผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างยิ่ง จึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.กฤษยา ปลั่งพงษ์พันธ์ และ ผู้ช่วยศาสตราจารย์ ดร.สุชาดา กรเพชรปภาณี ที่กรุณาเป็นประธานและกรรมการสอบคุณุภินิพนธ์ พร้อมทั้งให้คำแนะนำแนวทางที่ถูกต้อง ทำให้คุณุภินิพนธ์ฉบับนี้มีความสมบูรณ์ยิ่งขึ้น

ผู้วิจัยขอขอบคุณ มหาวิทยาลัยราชภัฏสกลนครที่สนับสนุนทุนการศึกษา สำหรับ ปีการศึกษาแรกในการศึกษาระดับปริญญาเอก ขอขอบคุณสำนักงานคณะกรรมการการอุดมศึกษา ที่ได้มอบทุนการศึกษาและทุนการทำวิจัยในระดับปริญญาเอกจนสำเร็จการศึกษา ภายใต้โครงการ ส่งเสริมการวิจัยในอุดมศึกษา และขอขอบคุณสำนักงานอุดมศึกษาจังหวัดสกลนคร ที่ให้ความอนุเคราะห์ข้อมูลอุดมศึกษาที่สำคัญและเป็นประโยชน์ต่อการวิจัยนี้

ผู้วิจัยขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ศิริลักษณ์ เจริญจิตต์พรชัย อาจารย์ประจำ ภาควิชาคณิตศาสตร์และสถิติ มหาวิทยาลัยมหาสารคาม และคุณประเสริฐ หวังสันติตระกูล วิศวกร โยธาประจำกรมโยธาธิการและผังเมือง ที่ได้กรุณาให้คำปรึกษาด้านสถิติ แนะนำการเขียนโปรแกรม ด้วยความห่วงใยใส่ใจเสมอมา ขอขอบคุณอาจารย์เอกรัตน์ อ่อนน้อมที่เอื้อเฟื้อสถานที่และเครื่องพิมพ์ และขอขอบคุณรุ่นพี่ รุ่นน้อง และเพื่อน ๆ ทุกคน ซึ่งผู้วิจัยไม่อาจกล่าวนามได้ครบในที่คอยให้ กำลังใจและให้ความช่วยเหลือกันด้วยดีเสมอมา

สุดท้ายนี้ ขอกราบขอบพระคุณ คุณพ่อสอน คุณแม่ยิ้ม ที่เป็นกำลังใจสำคัญ ที่ทำให้ คุณุภินิพนธ์ฉบับนี้สำเร็จได้ด้วยดี และขอขอบคุณสมาชิกในครอบครัวทุกคน ที่คอยให้กำลังใจ ห่วงใย รวมทั้งยังให้การสนับสนุนแก่ผู้วิจัยด้วยดีตลอดมา ประโยชน์อันเกิดจากคุณุภินิพนธ์ฉบับนี้ ผู้วิจัย ขอมอบเป็นเครื่องบูชาพระคุณพ่อ แม่ ครูบาอาจารย์ ตลอดจนผู้มีพระคุณทุกท่านที่ได้อบรมเลี้ยงดู ส่งเสริม ประสทธิประสาทวิชาความรู้ และปรารถนาดีต่อผู้วิจัยเสมอมา

ชัญญาภรณ์ แสงประสาน

52810063: สาขาวิชา: การวิจัยและสถิติทางวิทยาการปัญญา;

ปร.ด. (การวิจัยและสถิติทางวิทยาการปัญญา)

คำสำคัญ: การระเหยน้ำ/ ตัวประมาณค่าแบบอัตราส่วน/ วิธีกำลังสองตัดขอบน้อยสุด/ ตัวแปรช่วย

ชัญญกาญจน์ แสงประสาน: การประมาณค่าการระเหยน้ำโดยใช้ตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่ (AN EVAPORATION ESTIMATION USING THE MODIFIED RATIO ESTIMATOR) คณะกรรมการควบคุมคุณภาพ: เสรี ชัดเข้ม, ค.ด., พิชรี วงษ์เกษม, ปร.ด. 133 หน้า. ปี พ.ศ. 2558.

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อใช้ตัวแปรช่วยสองตัวและข้อมูลประชากรมีค่าผิดปกติ 2) เปรียบเทียบประสิทธิภาพตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) ภายใต้ 12 สถานการณ์ ของ 2 เงื่อนไข คือ ก) ขนาดตัวอย่างสุ่ม และ ข) ร้อยละของค่าผิดปกติ เกณฑ์เปรียบเทียบประสิทธิภาพ คือ เกณฑ์ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: *MSE*) และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Mean Percent Relative Error: *MPRE*) ต่ำสุด และ 3) ประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร เดือนพฤษภาคม ถึง ธันวาคม พ.ศ. 2556 ด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ และใช้ข้อมูลการระเหยน้ำ อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ ซึ่งเก็บบันทึกข้อมูลโดยสถานีอุตุนิยมวิทยาสกลนครจังหวัดสกลนคร กำหนดความแม่นยำของการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนครด้วยค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ไม่เกิน 10%

ผลการวิจัยปรากฏว่า

1. ตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่

$$(\bar{y}_M = \{ \bar{y} + b_1^* (\bar{X} - \bar{x}) + b_2^* (\bar{Z} - \bar{z}) \} \frac{\hat{e}_X + \beta_{1(x)} \frac{\hat{e}_Z}{\hat{e}_X} + \beta_{1(z)} \frac{\hat{e}_Z}{\hat{e}_X}}{\hat{e}_X + \beta_{1(x)} \frac{\hat{e}_Z}{\hat{e}_X} + \beta_{1(z)} \frac{\hat{e}_Z}{\hat{e}_X}})$$

มากกว่าตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007)

2. การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ไม่เกิน 10%

52810063: MAJOR: RESEARCH AND STATISTICS IN COGNITIVE SCIENCE;

Ph.D. (RESEARCH AND STATISTICS IN COGNITIVE SCIENCES)

KEYWORDS: RATIO ESTIMATOR/ LEAST TRIMMED SQUARE/ AUXILIARY VARIABLES/
EVAPORATION

CHANANKARN SAENGPRASAN: AN EVAPORATION ESTIMATION USING THE
MODIFIED RATIO ESTIMATOR. ADVISORY COMMITTEE: SEREE CHADCHAM, Ph.D.,
PATCHAREE WONGKASEM, Ph.D. 133 P. 2015.

The objectives of this research were (1) to modify the ratio population mean estimator (\bar{y}_M) that uses two auxiliary variables in a population that has outliers, (2) to compare the efficiency of the modified ratio population mean estimator with Gupta and Shabbir's (2007) estimator, $\bar{y}_{m(SD)}$, using the minimum Mean Square Error (*MSE*) and the Mean Percent Relative error (*MPRE*) as criterions under 12 simulation conditions that vary in sample size and the percentage of outliers, and (3) to estimate the monthly average evaporation, from May to December 2013, in Sakon Nakhon province using the \bar{y}_M estimator. Monthly average climatic data, consisting of pan evaporation, air temperature, and relative humidity from Sakon Nakhon meteorological station were used to estimate monthly evaporation. In this study, monthly evaporation estimation is considered to be acceptable if the mean percent relative error (*MPRE*) is less than 10%.

The results were as follows:

1. The modified ratio estimator,

$$\bar{y}_M = \left\{ \bar{y} + b_1^* (\bar{X} - \bar{x}) + b_2^* (\bar{Z} - \bar{z}) \right\} \frac{\bar{e}^{\bar{X}} + \beta_{1(x)} \frac{\bar{e}^{\bar{Z}}}{\bar{e}^{\bar{x}}} + \beta_{1(z)} \frac{\bar{e}^{\bar{y}}}{\bar{e}^{\bar{x}}}}{\bar{e}^{\bar{X}} + \beta_{1(x)} \frac{\bar{e}^{\bar{Z}}}{\bar{e}^{\bar{x}}} + \beta_{1(z)} \frac{\bar{e}^{\bar{y}}}{\bar{e}^{\bar{x}}}}$$

was more efficient than Gupta and Shabbir's (2007) estimator, $\bar{y}_{m(SD)}$.

2. The modified ratio estimator (\bar{y}_M) was acceptable for the estimation of monthly average pan evaporations in Sakon Nakhon province since the mean percent relative error (*MPRE*) was less than 10%.

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญภาพ.....	ณ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการวิจัย.....	5
สมมติฐานของการวิจัย.....	7
ประโยชน์ที่คาดว่าจะได้รับการวิจัย.....	7
ขอบเขตการวิจัย.....	8
นิยามศัพท์เฉพาะ.....	10
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	12
ตอนที่ 1 ตัวประมาณค่าแบบอัตราส่วนและงานวิจัยที่เกี่ยวข้อง.....	12
ตอนที่ 2 การระเหยน้ำ ข้อมูลอุตุนิยมวิทยาและงานวิจัยที่เกี่ยวข้อง.....	34
3 วิธีดำเนินการวิจัย.....	47
ตอนที่ 1 การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่.....	47
ตอนที่ 2 การจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่า แบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบ อัตราส่วน $\bar{y}_{m(SD)}$	57
ตอนที่ 3 การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่.....	61
4 ผลการวิจัย.....	64
ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ ปรับใหม่.....	65
ตอนที่ 2 ผลการจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่า แบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบ อัตราส่วน $\bar{y}_{m(SD)}$	73

สารบัญ (ต่อ)

บทที่	หน้า
4 ผลการวิจัย (ต่อ) ตอนที่ 3 ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่.....	85
5 สรุปและอภิปรายผล	92
สรุปผลการวิจัย.....	92
อภิปรายผล.....	95
ข้อเสนอแนะ.....	96
บรรณานุกรม	98
ภาคผนวก	102
ภาคผนวก ก ผลการทดสอบการแจกแจงของการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และ ความชื้นสัมพัทธ์.....	103
ภาคผนวก ข ผลการวิเคราะห์การถดถอยพหุคูณเบื้องต้นของการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์.....	113
ภาคผนวก ค ข้อมูลการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ ตั้งแต่เดือนพฤษภาคม 2556 ถึงเดือนธันวาคม 2556.....	115
ภาคผนวก ง การเขียนโปรแกรมจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของ ตัวประมาณค่า \bar{y}_M กับตัวประมาณค่า $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007).....	120
ภาคผนวก จ การเขียนโปรแกรมชุดคำสั่งเพื่อประมาณค่าการระเหยน้ำโดยเฉลี่ย รายเดือนในจังหวัดสกลนคร.....	129
ประวัติย่อของผู้วิจัย	133

สารบัญตาราง

ตารางที่	หน้า
4-1 ลักษณะข้อมูลเบื้องต้นของข้อมูลอุตุนิยมหาวิทยาลัยรายเดือนจำนวน 76 เดือน.....	73
4-2 ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างตัวแปรช่วยกับการระเหยน้ำแบบภาค.....	75
4-3 ผลการทดสอบการแจกแจงของข้อมูลด้วยสถิติ Anderson-Darling test.....	78
4-4 ผลการทดสอบการแจกแจงปกติสำหรับข้อมูลที่แปลงค่าแล้ว ด้วยสถิติ Anderson-Darling test.....	79
4-5 ค่าพารามิเตอร์ของข้อมูลประชากรจำลอง เมื่อ $N = 10,000$	80
4-6 ผลการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่า $\bar{y}_{m(SD)}$	81
4-7 ค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนของเดือนพฤษภาคม 2550 ถึงเดือนธันวาคม 2555.....	86
4-8 ค่าพารามิเตอร์ของข้อมูลประชากรรายเดือนพฤษภาคมถึงเดือนธันวาคม.....	86
4-9 ค่าสถิติของข้อมูลตัวอย่างรายเดือนพฤษภาคมถึงเดือนธันวาคม พ.ศ. 2556.....	88
4-10 ความแม่นยำของค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนพฤษภาคม ถึงธันวาคม 2556 ด้วยตัวประมาณค่า \bar{y}_M	90

สารบัญภาพ

ภาพที่	หน้า
1-1 กรอบแนวคิดในการวิจัยสำหรับการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่.....	7
2-1 กราฟการแจกแจงปกติที่มีพารามิเตอร์ μ และ σ	16
2-2 กราฟการแจกแจงแกมมาที่มีพารามิเตอร์ $\alpha > 0$ และ $\beta > 0$	17
2-3 กราฟของการแจกแจงไวบูลล์ที่มีพารามิเตอร์ $\alpha = 1$ และ $\beta > 0$	18
2-4 การใช้แผนภาพกล่องเพื่อตรวจสอบค่าผิดปกติ.....	22
2-5 เครื่องวัดการระเหยแบบถาด (Evaporation Pan) ชนิด American Class A Pan.....	37
3-1 แผนผังของการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่.....	50
3-2 กรอบแนวคิดของขั้นตอนการวิจัย สำหรับวัตถุประสงค์การวิจัยข้อ 2) และข้อ 3).....	51
3-3 การเขียนโปรแกรมเพื่อจำลองข้อมูลประชากร.....	57
3-4 การเขียนโปรแกรมเพื่อคัดแยกข้อมูลประชากรเป็นกลุ่มข้อมูลปกติและกลุ่มข้อมูลผิดปกติ.....	58
3-5 แผนผังการเขียนโปรแกรม 12 สถานการณ์ ยกตัวอย่างกรณี $n = 20$ มีค่าผิดปกติร้อยละ 10.....	60
3-6 สรุปขั้นตอนการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน โดยใช้ตัวประมาณค่า \bar{y}_M	63
4-1 การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M).....	66
4-2 แผนภาพการกระจายระหว่งการระเหยน้ำแบบถาดกับกับตัวแปรอุณหภูมิตัวแปรอื่น.....	74
4-3 แผนภาพกล่องแสดงการกระจายของข้อมูล ก) การระเหยน้ำแบบถาด ข) อุณหภูมิอากาศเฉลี่ย และ ค) ความชื้นสัมพัทธ์.....	76
4-4 กราฟฮิสโทแกรมของข้อมูลเมื่อเทียบกับการแจกแจงปกติ การแจกแจงแกมมาและการแจกแจงไวบูลล์.....	77
4-5 สรุปสถานการณ์ 12 สถานการณ์ที่ตัวประมาณ \bar{y}_M มีความเที่ยงตรงสูงกว่าตัวประมาณ $\bar{y}_{m(SD)}$	83
4-6 สรุปสถานการณ์ 12 สถานการณ์ที่ตัวประมาณ \bar{y}_M มีความแม่นยำสูงกว่าตัวประมาณ $\bar{y}_{m(SD)}$	84
4-7 การเปลี่ยนแปลงของค่าเฉลี่ยของการระเหยน้ำ (มิลลิเมตร) อุณหภูมิอากาศเฉลี่ย (°C) และความชื้นสัมพัทธ์ จากเดือนพฤษภาคม ถึงเดือนธันวาคม 2556.....	89
4-8 ค่าพารามิเตอร์และค่าประมาณของการระเหยน้ำโดยเฉลี่ยรายเดือนพฤษภาคมถึงธันวาคม 2556 ด้วยตัวประมาณค่า \bar{y}_M	91

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การระเหยน้ำเป็นการเปลี่ยนแปลงเคลื่อนที่ของน้ำจากสถานะของเหลวไปเป็นไอ ซึ่งเป็นส่วนหนึ่งในวัฏจักรน้ำ (Hydrologic Cycle) ในระบบน้ำบรรยากาศ ที่ประกอบด้วย ฝน การดักค้าง การระเหยและการคายน้ำ ทั้งหมดล้วนมีความสัมพันธ์เกี่ยวเนื่องซึ่งกันและกัน ดังนั้นการศึกษาสมดุลของน้ำเพื่อประโยชน์ทางเกษตรกรรมหรือชลศาสตร์ต่าง ๆ ในเรื่องการคาดคะเนการระเหยน้ำ โดยเฉพาะการระเหยน้ำที่พื้นดินให้ได้ค่าที่ถูกต้องทำได้ยาก ขึ้นอยู่กับขนาดของพื้นที่ศึกษา แต่อย่างไรก็ตาม การคาดคะเนการระเหยน้ำในพื้นที่เล็ก ๆ ย่อมมีความละเอียดกว่าการศึกษาในพื้นที่ขนาดใหญ่ ๆ และในระเบียบวิธีอุทกวิทยา เราต้องตัดทอนความซับซ้อนออกบ้าง เพื่อให้สามารถคำนวณได้ และให้ได้คำตอบที่ไม่ห่างจากความจริงเกินกว่าจะยอมรับได้ (วิชัย ศรีบุญลือ, 2552, หน้า 8) เนื่องจากการระเหยน้ำไม่สามารถวัดได้โดยตรง แต่อาศัยเทคโนโลยีใหม่ ๆ ที่เกิดขึ้น วัดการเปลี่ยนแปลงไอน้ำจากพื้นผิวน้ำหรือพื้นผิวดิน เพื่อประมาณค่าการระเหยน้ำ (Evaporation) ซึ่งมี 3 วิธีการ ได้แก่ ภาควัดการระเหยน้ำ (Evaporation Pans) สมการความต่อเนื่องหรือสมการสมดุล (Continuity Equations) และสูตรเอมพิริคัล (Empirical Formula) (Thompson, 1999, p. 91)

สำหรับประเทศไทย สถานีตรวจอากาศภาคพื้นดินของกรมอุตุนิยมวิทยาในแต่ละสถานี ประมาณค่าการระเหยน้ำจากภาควัดการระเหยแบบ Class A Pan ซึ่งการระเหยน้ำในประเทศไทย คาบ 10 ปีในช่วง พ.ศ 2544 – 2553 ปรากฏว่า การระเหยน้ำมีความแตกต่างกันออกไปในแต่ละพื้นที่ และช่วงที่มีการระเหยน้ำสูงอยู่ในช่วงเดือนมีนาคม เมษายน และพฤษภาคม โดยเดือนเมษายน มีปริมาณการระเหยน้ำสูงที่สุดในรอบปี (พิศ คงบริรักษ์ และกฤติกา สืบศักดิ์, 2554, หน้า 30-38) ค่าการระเหยน้ำมีความสำคัญ ใช้เป็นค่าเริ่มต้นในการหาปริมาณการระเหยน้ำในอ่างเก็บน้ำ การหาปริมาณการใช้น้ำของพืช (Crop Evapotranspiration) หรือการใช้น้ำของพืชอ้างอิงในแต่ละพื้นที่ ดังนั้นการศึกษการระเหยน้ำจึงเป็นประโยชน์ต่อการเกษตรกรรมหรือพื้นที่ใช้ประโยชน์ทางการเกษตร ซึ่งสถิติการใช้ที่ดินรายภาคภูมิศาสตร์นับตั้งแต่ พ.ศ. 2552 ถึง พ.ศ. 2555 จะเห็นว่า ภาคตะวันออกเฉียงเหนือเป็นภาคที่มีจำนวนผู้ถือครองเนื้อที่ทำการเกษตรและเนื้อที่ถือครองทำการเกษตรโดยรวมมากที่สุด โดยเฉพาะด้านการเพาะปลูกพืช (สำนักงานเศรษฐกิจการเกษตร, 2557, หน้า 168)

จังหวัดสกลนครเป็นจังหวัดหนึ่งในสี่สิบจังหวัดของภาคตะวันออกเฉียงเหนือตอนบน ตั้งอยู่ระหว่างเส้นละติจูดที่ 16 องศา 45 ลิปดา ถึง 18 องศา 15 ลิปดา เหนือ และเส้นลองจิจูดที่ 103 องศา 15 ลิปดา ถึง 104 องศา 30 ลิปดา ตะวันออก มีเนื้อที่ประมาณ 10,553.7 ตารางกิโลเมตร มีขนาดใหญ่เป็นอันดับ 9 ของจังหวัดในภาคตะวันออกเฉียงเหนือ มีอ่างเก็บน้ำชลประทานในจังหวัดสกลนคร เป็นโครงการชลประทานขนาดใหญ่ 1 แห่ง โครงการชลประทานขนาดกลาง 41 แห่ง และโครงการชลประทานขนาดเล็ก 192 แห่ง (กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม, 2555, หน้า 13-14) นับเป็นจังหวัดที่มีโครงการชลประทานมากกว่าหลายจังหวัดในภาคตะวันออกเฉียงเหนือ และมีอ่างเก็บน้ำที่สำคัญถึง 38 แห่ง การศึกษการประมาณค่าการระเหยน้ำในจังหวัดสกลนคร ทำให้ได้

ข้อมูลที่สามารถเป็นประโยชน์ต่อการวางแผนจัดการทรัพยากรน้ำเพื่อการเกษตรกรรมได้ ซึ่งปัจจุบัน การประมาณค่าการระเหยน้ำในจังหวัดสกลนคร จะใช้เครื่องวัดการระเหยแบบถาด (Evaporation Pan) ชนิด American Class A Pan และบันทึกข้อมูลโดยสถานีอุตุนิยมวิทยาสกลนคร

สำหรับการศึกษาการพยากรณ์การระเหยน้ำในพื้นที่ใกล้เคียงอื่น ๆ มีนักวิจัยหลายคนได้ ศึกษาการพยากรณ์การระเหยน้ำหรือการคายระเหยน้ำ ด้วยวิธีการวิเคราะห์เชิงปริมาณสองกลุ่มหลัก ๆ ได้แก่ วิธีการที่ 1) พยากรณ์การระเหยน้ำโดยอาศัยความสัมพันธ์ระหว่างการระเหยน้ำกับข้อมูลทางอุตุนิยมวิทยาอื่น ๆ เช่น ความชื้นสัมพัทธ์ อุณหภูมิ ปริมาณแสงแดด ปริมาณรังสีดวงอาทิตย์ ความเร็วลมเฉลี่ย และความกดอากาศ เทคนิควิธีที่ใช้ในการพยากรณ์ในกลุ่มนี้ ได้แก่ การวิเคราะห์ การถดถอยพหุคูณ การประยุกต์โครงข่ายประสาทเทียมแบบต่าง ๆ รวมทั้งวิธีของ Penman วิธีของ Priestley-Taylor และวิธีของ Stephens-Stewart (Kim, Shiri, Kisi, & Singh, 2013; Kisi, 2006; Shirsath, & Singh, 2010; Goel, 2009; Irmak, & Haman, 2003) และวิธีการที่ 2) พยากรณ์ การระเหยน้ำโดยอาศัยความสัมพันธ์ระหว่างการระเหยน้ำในปัจจุบันกับในอดีต เทคนิควิธีที่ใช้ในการพยากรณ์กลุ่มนี้ คือ การวิเคราะห์อนุกรมเวลาแบบ Box-Jenkins (Dabral, Jhajharia, Mishra, Hangshing, & Doley, 2014; Meshram, Gorantiwar, Kulkarni, & Hangaarkekar, 2013; Dodangeh, Abedi, & Gohari, 2012; Moshrik, Ahmed, & Zeyad, 2008) ผลจากการศึกษา เปรียบเทียบประสิทธิภาพของการพยากรณ์การระเหยน้ำระหว่างวิธีการวิเคราะห์การถดถอยพหุคูณ กับวิธีโครงข่ายประสาทเทียม เมื่อใช้ทุก ๆ ตัวแปรอุตุนิยมวิทยาที่ศึกษาสร้างตัวแบบปรากฏว่า วิธีโครงข่ายประสาทเทียมสามารถพยากรณ์ได้แม่นยำกว่าวิธีการวิเคราะห์การถดถอยพหุคูณเพียง เล็กน้อย (Shirsath, & Singh, 2010; Goel, 2009) แต่วิธีการวิเคราะห์การถดถอยพหุคูณสามารถ พยากรณ์การระเหยน้ำรายวันได้แม่นยำกว่าวิธีของ Penman วิธีของ Priestley-Taylor และวิธีของ Stephens-Stewart และกรณีที่ใช้ตัวแปรอุตุนิยมวิทยาเพียงสามตัวสร้างตัวแบบพยากรณ์การระเหย น้ำรายวันปรากฏว่า วิธีการวิเคราะห์การถดถอยพหุคูณมีความแม่นยำใกล้เคียงกับวิธีโครงข่าย ประสาทเทียมอย่างมาก (Shirsath, & Singh, 2010) นอกจากนี้การศึกษาของ Saengprasan, Chadcham, and Jermjitpornchai (2014) ได้เปรียบเทียบประสิทธิภาพการพยากรณ์การระเหยน้ำ ในจังหวัดสกลนครระหว่างการวิเคราะห์การถดถอยพหุคูณกับการวิเคราะห์อนุกรมเวลาแบบ Box-Jenkins ปรากฏว่า การพยากรณ์การระเหยน้ำรายเดือนด้วยวิธีการวิเคราะห์การถดถอยพหุคูณมี ความแม่นยำกว่าการพยากรณ์ด้วยตัวแบบ Box-Jenkins อย่างไรก็ตาม ข้อมูลการระเหยน้ำเป็น ข้อมูลที่มีลักษณะของฤดูกาลร่วมด้วย โดยเฉพาะในช่วงเดือนมีนาคมถึงเมษายนที่มีการระเหยน้ำ สูงสุดในรอบปี ถือได้ว่าเป็นค่าสูงผิดปกติในรอบปี ดังนั้นในการปรับปรุงวิธีการพยากรณ์ด้วย การวิเคราะห์การถดถอยพหุคูณ หากนำค่าผิดปกติ (Outliers) ของข้อมูล มารวมเพื่อปรับปรุงตัวแบบ การพยากรณ์ ย่อมจะได้ความแม่นยำในการพยากรณ์สูงขึ้น ทั้งนี้ตัวแบบพยากรณ์ที่ดี ควรเป็นแบบ ทางคณิตศาสตร์ (Mathematical Model) ที่ประหยัด แต่มีความแม่นยำในการพยากรณ์สูง เมื่อ เทียบกับตัวแบบอื่น ๆ

ส่วนใหญ่การศึกษาเชิงสำรวจ มีจุดมุ่งหมายเพื่อประมาณค่าพารามิเตอร์ของตัวแปรที่สนใจ ศึกษาซึ่งไม่ทราบค่า เช่น ค่าเฉลี่ยประชากร ค่าสัดส่วนประชากร หรือค่าความแปรปรวนประชากร อีกทั้งการศึกษาเชิงสำรวจ มีข้อจำกัดด้านระยะเวลาของการศึกษาและงบประมาณ จึงต้องใช้วิธีสุ่ม

ตัวอย่างเพื่อให้ได้ตัวแทนที่ดีของประชากร เพื่อใช้ในการประมาณค่าพารามิเตอร์ที่สนใจเหล่านั้น ทั้งนี้เพื่อให้ได้ตัวประมาณค่าพารามิเตอร์ที่มีความแม่นยำหรือลดความคลาดเคลื่อนของตัวประมาณค่าให้ได้มากที่สุด ในบางครั้งตัวแปรที่สนใจศึกษาอาจมีความสัมพันธ์กับตัวแปรอื่น ๆ ที่เรียกว่า ตัวแปรช่วย (Auxiliary Variables: X) ดังนั้น ถ้าหากใช้ตัวแปรช่วยที่มีความสัมพันธ์กับตัวแปรที่สนใจศึกษาเหล่านี้ มาช่วยในการประมาณค่าพารามิเตอร์ ก็อาจจะทำให้ได้การประมาณที่มีความถูกต้องหรือมีประสิทธิภาพสูงขึ้น (สุรินทร์ นียมางกูร, 2546, หน้า 125)

วิธีการประมาณค่าเฉลี่ยประชากรโดยใช้สารสนเทศจากตัวแปรช่วยวิธีการหนึ่ง คือ การใช้ตัวประมาณค่าแบบอัตราส่วน (Ratio Estimator: \bar{y}_r) ที่ใช้ประมาณค่าเฉลี่ยประชากร \bar{Y} กล่าวคือ $\bar{y}_r = [\bar{y}/\bar{x}]\bar{X} = \hat{R}\bar{X}$ เมื่อ \bar{X} คือ ค่าเฉลี่ยประชากรของตัวแปร X ที่ทราบค่า \bar{x} คือ ค่าเฉลี่ยตัวอย่างของ X และ \bar{y} คือ ค่าเฉลี่ยตัวอย่างของ Y โดยค่าคลาดเคลื่อนกำลังสองเฉลี่ยของ \bar{y}_r คือ $MSE(\bar{y}_r) @ [(1-f)/n] [R^2 S_x^2 - 2RS_{xy} + S_y^2]$ ถ้าตัวแปร Y และ X มีความสัมพันธ์กันสูงหรืออัตราส่วน $r_i = y_i/x_i$ มีค่าเกือบคงที่ทุกค่า $i = 1, 2, \dots, n$ แล้ว ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรจะมีความแม่นยำสูง (ประชุม สุวัตถิ, 2552, หน้า 53)

ในทางวิชาการมีนักสถิติและนักวิชาการหลายคนได้พยายามพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร โดยอาศัยข้อมูลจากตัวแปรช่วย ให้มีความแม่นยำมากขึ้น หรือทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) มีค่าน้อยลง เริ่มจาก Sisodia and Dwivedi (1981) ใช้ค่าสัมประสิทธิ์ความแปรผันของตัวแปรช่วย (C_x) ที่ทราบค่า มาปรับตัวประมาณค่าแบบอัตราส่วน \bar{y}_r ต่อมา Singh and Kakran (1993 cited in Oral, 2010) ได้นำสัมประสิทธิ์ความโค้งของตัวแปร X (Coefficient of kurtosis of X : $\beta_{2(x)}$) มาปรับตัวประมาณค่าของ Sisodia and Dwivedi ที่เสนอไว้ในปี ค.ศ. 1981 ต่อมา Kadilar and Cingi (2004) ได้นำตัวประมาณค่าแบบอัตราส่วนที่ใช้หลักการวิเคราะห์การถดถอย $\bar{y}_{pr} = [\bar{y} + b(\bar{X} - \bar{x})/\bar{x}]\bar{X} = \hat{R}_{pr}\bar{X}$ มาร่วมในการปรับตัวประมาณค่าแบบอัตราส่วนของ Sisodia and Dwivedi ที่เสนอไว้ในปี ค.ศ. 1981 ได้ตัวประมาณค่า \bar{y}_{psD} และปรับตัวประมาณค่าแบบอัตราส่วนของ Singh and Kakran ที่เสนอไว้ในปี ค.ศ. 1993 ได้ตัวประมาณค่า \bar{y}_{psK}

จากการศึกษาของ Sisodia and Dwivedi (1981), Singh and Kakran (1993) และ Kadilar and Cingi (2004) จะเห็นแนวทางการพัฒนาตัวประมาณค่าแบบอัตราส่วนโดยใช้ตัวแปรช่วยเพียงหนึ่งตัว แต่ในบางครั้งตัวแปรที่ศึกษา (Y) อาจมีความสัมพันธ์กับตัวแปรช่วยมากกว่า 1 ตัว และการใช้สารสนเทศจากตัวแปรช่วยทั้งหมด ย่อมเป็นการเพิ่มความแม่นยำในการพยากรณ์ค่า Y มากขึ้น ดังนั้นนักสถิติหลายคนจึงได้พยายามพัฒนาตัวประมาณค่าแบบอัตราส่วน โดยใช้ตัวแปรช่วยสองตัว เช่น การศึกษาของ Abu-Dayyeh, Ahmed, Ahmed, and Muttalak (2003), Kadilar and Cingi (2005), Gupta and Shabbir (2007) ที่ได้เสนอตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากรโดยใช้สารสนเทศจากตัวแปรช่วย 2 ตัว ซึ่ง Gupta and Shabbir (2007) ได้เสนอกุ่มของตัวประมาณค่าแบบอัตราส่วน โดยใช้ตัวแปรช่วย 2 ตัวที่มีความสัมพันธ์ทางบวกกับตัวแปรที่ศึกษา และใช้หลักการแปลงค่าตัวแปรช่วย นำไปปรับตัวประมาณค่าที่ Sisodia and Dwivedi เสนอ

ไว้ในปี ค.ศ. 1981 ได้ตัวประมาณค่า $\bar{y}_{m(SD)}$ และปรับตัวประมาณค่าที่ Upadhyaya and Singh เสนอไว้ในปี ค.ศ. 1999 ได้ตัวประมาณค่า $\bar{y}_{m(US2)}$ เมื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าปรากฏว่า ตัวประมาณค่า $\bar{y}_{m(SD)}$ และตัวประมาณค่า $\bar{y}_{m(US2)}$ มีประสิทธิภาพดีกว่าตัวประมาณค่าดั้งเดิม (\bar{y}_r) แต่ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าใหม่กับตัวประมาณค่าดั้งเดิมนั้นไม่ได้คำนึงถึงเรื่องการใช้จำนวนตัวแปรช่วยที่ไม่เท่ากัน นอกจากนี้การศึกษาของ Gupta and Shabbir (2007) ก็ไม่ได้ศึกษาในกรณีที่ข้อมูลประชากรมีค่าผิดปกติ (Outliers) รวมอยู่ด้วย

สำหรับข้อมูลประชากรที่มีการแจกแจงไม่สมมาตรหรือมีความเบ้ (Skewness) แล้ว การสุ่มตัวอย่างข้อมูลที่เก็บรวบรวมมา ควรมีการตรวจสอบเพื่อหาค่าผิดปกติ ถ้าหากข้อมูลตัวอย่างมีค่าผิดปกติร่วมอยู่ด้วย ผลการวิเคราะห์ข้อมูลย่อมมีความผิดพลาด ฉะนั้นเมื่อปรากฏค่าผิดปกติในข้อมูล นักวิจัยสามารถเลือกที่จะลบข้อมูลนั้นทิ้ง หรือไม่ก็หาวิธีการปรับให้เหมาะสม หากนักวิจัยเลือกวิธีการปรับให้เหมาะสม ก็ควรเลือกใช้สถิติที่มีความแกร่ง (Yarmohammadi & Mahmoudvand, 2010) ดังนั้น นักวิจัยพยายามพัฒนาตัวประมาณค่าที่มีความเหมาะสมเพื่อใช้กับข้อมูลที่มีค่าผิดปกติ ดังการศึกษาของ Kadilar, Candan and Cingi (2007) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากร อาศัยการวิเคราะห์การถดถอยแกร่ง (Robust Regression) มาปรับตัวประมาณค่าอัตราส่วนเดิมของ Kadilar and Cingi ที่นำเสนอในปี ค.ศ. 2004 โดยใช้ตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธี Huber-M Estimator และการศึกษาของภทรรรณ แสงนวกิจ (2553) ที่ปรับตัวประมาณค่าของ Kadilar and Cingi (2004) และตัวประมาณค่าของ Kadilar et al. (2007) ด้วยหลักการวิเคราะห์การถดถอยแกร่ง เมื่อประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุด (Least Trimmed Squares: LTS) ปรับตัวประมาณค่าเดิม ทำให้ได้ตัวประมาณค่าแบบอัตราส่วนที่มีประสิทธิภาพสูงกว่าตัวประมาณค่าเดิม สำหรับการวิเคราะห์การถดถอยแกร่งที่มีการประมาณค่าสัมประสิทธิ์การถดถอยหลายวิธีนั้น Yarmohammadi and Mahmoudvand (2010) ได้เปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยที่มีความแกร่งพบว่า ประสิทธิภาพของตัวประมาณค่าด้วยวิธี LTS ใกล้เคียงกับตัวประมาณค่าด้วยวิธี Least Median Square (LMS) และตัวประมาณค่าทั้งสองมีประสิทธิภาพสูงกว่าตัวประมาณค่าแบบวิธี Huber-M

จากผลการศึกษาเกี่ยวกับการพัฒนาตัวประมาณค่าแบบอัตราส่วนด้วยการใช้สถิติถดถอยแกร่งข้างต้น ร่วมกับทิศทางการใช้สารสนเทศของตัวแปรช่วยแบบใหม่ ที่ใช้สัมประสิทธิ์ความเบ้ของข้อมูลที่มีลักษณะการแจกแจงไม่สมมาตร เมื่อใช้ตัวแปรช่วยหนึ่งตัว (Yan & Tian, 2010; ภทรรรณ แสงนวกิจ, 2553) และผลการศึกษาของ Shirsath and Singh (2010), Goel (2009) และการศึกษาของ Saengprasan et al. (2014) เกี่ยวกับประสิทธิภาพของการพยากรณ์การระเหยน้ำในจังหวัดสกลนครด้วยการวิเคราะห์การถดถอยพหุคูณ ที่แสดงให้เห็นว่า หลักการวิเคราะห์การถดถอยพหุคูณสามารถใช้ในการพยากรณ์หรือประมาณค่าการระเหยน้ำได้แม่นยำ อีกทั้งยังไม่ปรากฏการพัฒนาตัวประมาณค่าแบบอัตราส่วนที่ใช้การวิเคราะห์การถดถอยแกร่งร่วมกับสัมประสิทธิ์ความเบ้ กรณีที่มีตัวแปรช่วยสองตัว ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาตัวประมาณค่าแบบอัตราส่วนขึ้นใหม่ สำหรับกรณีที่มีตัวแปรช่วยสองตัว โดยใช้สัมประสิทธิ์ความเบ้ (Coefficient of Skewness) และหลักการวิเคราะห์การถดถอยแกร่งที่มีการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุด (LTS)

มาปรับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ที่ Gupta and Shabbir (2007) ได้นำเสนอไว้ เพื่อให้ได้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ที่สามารถประมาณค่าการระเหยน้ำรายเดือนโดยเฉลี่ยในจังหวัดสกลนคร โดยใช้สารสนเทศจากตัวแปรช่วยสองตัว นอกจากนี้ยังสามารถนำตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กรณีใช้ตัวแปรช่วยสองตัวนี้ ไปใช้ประมาณค่าเฉลี่ยประชากรของชุดข้อมูลอื่น ๆ ที่มีลักษณะไม่สมมาตรหรือที่มีค่าผิดปกติปะปนอยู่ได้

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ โดยใช้ตัวแปรช่วยสองตัว ในกรณีข้อมูลประชากรมีค่าผิดปกติ
2. เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) ภายใต้สถานการณ์ 12 สถานการณ์ (3×4) ประกอบด้วย 1) ขนาดตัวอย่างสุ่ม 3 เงื่อนไข และ 2) ร้อยละของค่าผิดปกติ 4 เงื่อนไข ด้วยการจำลองข้อมูลโดยเทคนิคมอนติคาร์โล
3. เพื่อประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

กรอบแนวคิดในการวิจัย

การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร กรณีใช้ตัวแปรช่วยสองตัว ที่ Gupta and Shabbir (2007) ได้นำเสนอไว้ในรูปแบบ Ratio-Cum-Product Type คือ

$$\bar{y}_{m(SD)} = \bar{y} \frac{\frac{\sum x + C_x \frac{\sum z}{\sum z} + C_z \frac{\sum y}{\sum y}}{\sum x + C_x \frac{\sum z}{\sum z} + C_z \frac{\sum y}{\sum y}}}{\frac{\sum x + C_x \frac{\sum z}{\sum z} + C_z \frac{\sum y}{\sum y}}{\sum x + C_x \frac{\sum z}{\sum z} + C_z \frac{\sum y}{\sum y}}} \quad (1.1)$$

ตามสมการ (1.1) ไม่ได้คำนึงถึงการวิเคราะห์ข้อมูลที่มีค่าผิดปกติ (Outliers) รวมอยู่ด้วย ถ้าข้อมูลชุดใดมีค่าผิดปกติอยู่ ผลลัพธ์การประมาณค่าเฉลี่ยอาจบิดเบือนได้ ดังนั้นนักวิจัยหลายคนได้พยายามพัฒนาตัวประมาณค่าแบบอัตราส่วนที่มีความแกร่ง ดังการศึกษาของ Kadilar et al. (2007) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากร ด้วยการวิเคราะห์การถดถอยแกร่ง (Robust Regression) ที่มีการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธี Huber-M Estimator แต่ Rousseeuw (1984) ได้เสนอตัวประมาณค่าสัมประสิทธิ์การถดถอยแบบแกร่งไว้หลายวิธี ดังเช่นวิธี Huber-M Estimator วิธี Least Median Square (LMS) และวิธีกำลังสองตัดขอบน้อยสุด (Least Trimmed Squares: LTS) ต่อมา Yarmohammadi and Mahmoudvand (2010) ได้เปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยแบบแกร่งพบว่า ประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธี LTS ดีกว่าตัวประมาณค่าสัมประสิทธิ์การถดถอย

แบบ Huber-M ด้วยเหตุนี้ ผู้วิจัยจึงมีแนวคิดที่จะปรับเทอม \bar{y} ในสมการ (1.1) ด้วยการวิเคราะห์การถดถอยแปร่งที่ใช้ตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุด (LTS)

จากแนวคิดการวิเคราะห์การถดถอยแปร่งที่ใช้ตัวประมาณค่าสัมประสิทธิ์การถดถอยวิธีกำลังสองตัดขอบน้อยสุด (LTS) ร่วมกับการใช้สารสนเทศสัมประสิทธิ์ความเบ้ของตัวแปรช่วย ซึ่งเป็นทิศทางการใช้สารสนเทศของตัวแปรช่วยแบบใหม่ (Yan & Tian, 2010) นั้น ผู้วิจัยมุ่งจะปรับตัวประมาณค่า $\bar{y}_{m(SD)}$ ที่ Gupta and Shabbir เสนอไว้ในสมการที่ (1.1) โดยพัฒนาเป็นตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ที่ใช้ตัวแปรช่วยสองตัวเพื่อให้ได้ตัวประมาณค่าที่มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยลง และมีความเหมาะสมมากขึ้นในการประมาณค่าเฉลี่ยของการระเหยน้ำ ซึ่งเป็นข้อมูลที่มีค่าผิดปกติด้านบวกพร้อมด้วย ดังนี้

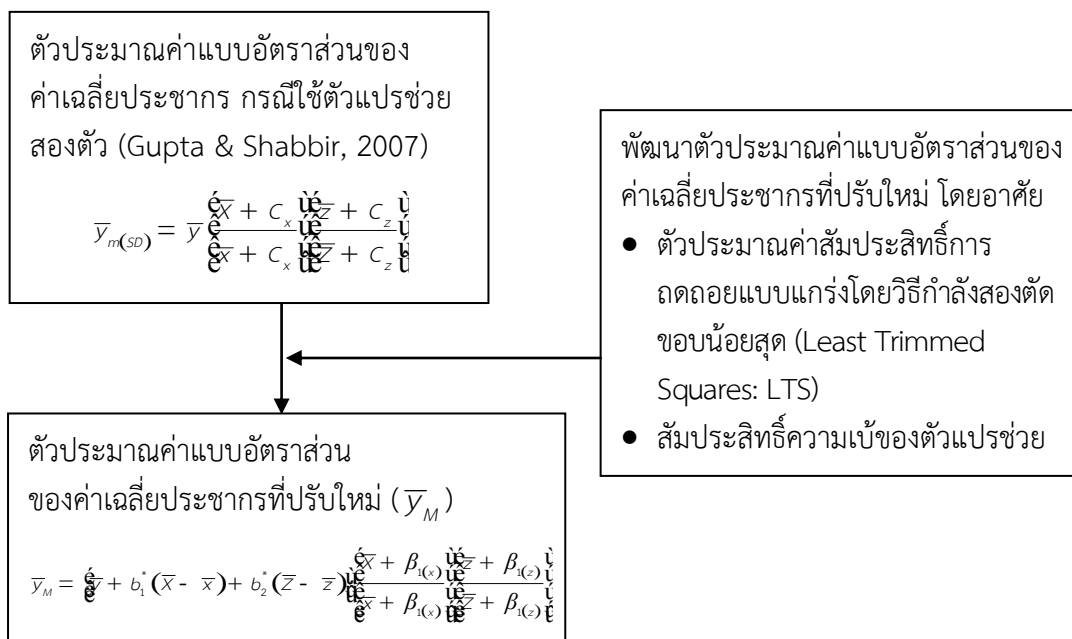
$$\bar{y}_M = \frac{\bar{y} + b_1^* (\bar{X} - \bar{x}) + b_2^* (\bar{Z} - \bar{z})}{\frac{\bar{x} + \beta_{1(x)} \bar{z} + \beta_{1(z)} \bar{y}}{\bar{x} + \beta_{1(x)} \bar{z} + \beta_{1(z)} \bar{y}}} \quad (1.2)$$

เมื่อ $\beta_{1(x)}, \beta_{1(z)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย X, Z ตามลำดับ

b_1^*, b_2^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุด

และใช้สูตรหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) เพื่อเปรียบเทียบประสิทธิภาพกับตัวประมาณค่า $\bar{y}_{m(SD)}$ และประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนครด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ที่ผู้วิจัยนำเสนอ

จากที่กล่าวข้างต้น กรอบแนวคิดในการวิจัยสำหรับการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กรณีใช้ตัวแปรช่วยสองตัว ดังภาพที่ 1-1



ภาพที่ 1-1 กรอบแนวคิดในการวิจัยสำหรับการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

สมมติฐานของการวิจัย

1. ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ที่พัฒนาขึ้น ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) น้อยกว่าตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) ภายใต้ 12 สถานการณ์ (3×4) ซึ่งเป็นผลประกอบของเงื่อนไขย่อย 2 เงื่อนไข ดังนี้

- 1.1 ขนาดตัวอย่างสุ่ม (n) 3 เงื่อนไข คือ n เท่ากับ 20, 40 และ 100 หน่วยตัวอย่าง
- 1.2 ร้อยละของค่าผิดปกติ (o) 4 เงื่อนไข คือ o เท่ากับ ร้อยละ 5, ร้อยละ 10, ร้อยละ 15 และร้อยละ 20

2. การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Percent Relative Error: PRE) โดยเฉลี่ย ไม่เกิน 10%

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เพื่อให้ได้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ที่มีความแม่นยำในการประมาณค่าเฉลี่ยประชากรข้อมูลที่มีค่าผิดปกติด้านบวก กรณีใช้ตัวแปรช่วยสองตัว

2. เพื่อทราบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อเทียบกับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ กรณีใช้ตัวแปรช่วยสองตัว ภายใต้แต่ละสถานการณ์ ที่เป็นผลประกอบของ 2 เงื่อนไข ได้แก่ 1) ขนาดตัวอย่างสุ่ม และ 2) ร้อยละของค่าผิดปกติ

3. เพื่อสามารถประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนครได้แม่นยำยิ่งขึ้น เมื่อใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ขอบเขตของการวิจัย

การประมาณค่าการระเหยน้ำโดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กรณีที่ใช้ตัวแปรช่วยสองตัว มีขอบเขตการวิจัยตามวัตถุประสงค์การวิจัย ดังนี้

ขอบเขตของการวิจัยสำหรับวัตถุประสงค์การวิจัยข้อ 1 และข้อ 2

1. ศึกษาเฉพาะตัวประมาณค่าแบบอัตราส่วนที่อยู่ในรูปแบบอัตราส่วน (Ratio-Cum-Product Type) เท่านั้น
2. ศึกษาภายใต้การชักตัวอย่างแบบสุ่มเชิงเดียวชนิดไม่คืนที่ (SRS without Replacement) โดยการสุ่มตัวอย่างขนาด n จากประชากรขนาด N
3. สำหรับการจำลองข้อมูล ผู้วิจัยศึกษาลักษณะเบื้องต้นของข้อมูลอุตุนิยมหาวิทยาลัยรายเดือนก่อน เพื่อนำคุณลักษณะเบื้องต้นเหล่านั้นไปใช้ในการจำลองข้อมูล ในที่นี้ข้อมูลอุตุนิยมหาวิทยาลัยรายเดือนประกอบด้วยการระเหยน้ำ (มิลลิเมตร) อุณหภูมิอากาศ ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%) จากเดือนมกราคม พ.ศ. 2550 ถึงเดือนเมษายน พ.ศ. 2556 จำนวน 76 เดือน เก็บบันทึกข้อมูลโดยสถานีอุตุนิยมหาวิทยาลัยสกลนคร จังหวัดสกลนคร ตั้งอยู่สูงจากระดับน้ำทะเลปานกลาง 171 เมตร

ในบรรดาตัวแปรอุตุนิยมหาวิทยาลัยรายเดือน 5 ตัวแปร (อุณหภูมิอากาศ ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในช่วงกลางวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%)) ผู้วิจัยทำการคัดเลือกตัวแปรช่วยจำนวน 2 ตัว โดยพิจารณาเลือกตัวแปรสองตัวแรกที่มีความสัมพันธ์สูงสุดกับการระเหยน้ำรายเดือนในจังหวัดสกลนคร และตัวแปรช่วยสองตัวนั้นจะต้องมีค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันไม่เกิน 0.7 และค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรสองตัวนั้น จะต้องมิต่ำกว่าความสัมพันธ์ระหว่างตัวแปรช่วยใด ๆ กับการระเหยน้ำรายเดือน ทั้งนี้เพื่อป้องกันการเกิดปัญหาตัวแปรช่วยมีความสัมพันธ์กันเอง หรือที่เรียกว่าปัญหา Collinearity (Hair, Black, Babin, & Anderson, 2010, p. 205)

4. จำลองข้อมูลเพื่อเปรียบเทียบการประมาณค่าเฉลี่ยประชากรระหว่างตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ภายใต้สถานการณ์ 12 สถานการณ์ ด้วยโปรแกรม Matlab และมีการทดลองซ้ำ 1,000 ครั้ง โดยที่

สถานการณ์ที่ศึกษาทั้งหมด 12 สถานการณ์ (3×4) นั้น แต่ละสถานการณ์เป็นผลประกอบระหว่างเงื่อนไขย่อย 2 เงื่อนไข ดังนี้

4.1 ขนาดตัวอย่างสุ่ม (n) 3 เงื่อนไข ได้แก่ 20, 40 และ 100 ตัวอย่าง

4.2 ร้อยละของค่าผิดปกติ (o) 4 เงื่อนไข ได้แก่ ร้อยละ 5, ร้อยละ 10, ร้อยละ 15 และ ร้อยละ 20

5. เกณฑ์ในการเปรียบเทียบประสิทธิภาพของตัวประมาณค่า พิจารณาจากสมบัติสองประการของตัวประมาณค่า (Cochran, 1977, p. 16) ดังนี้

5.1 ความเที่ยงตรง (Precision) ของตัวประมาณค่า พิจารณาจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของตัวประมาณค่า โดยที่ตัวประมาณค่า $\hat{\theta}_1$ มีความเที่ยงตรงสูงกว่า $\hat{\theta}_2$ ก็ต่อเมื่อ $MSE(\hat{\theta}_1) < MSE(\hat{\theta}_2)$

5.2 ความแม่นยำ (Accuracy) ของตัวประมาณค่า พิจารณาจากค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Mean Percent Relative Error: $MPRE$) โดยที่ตัวประมาณค่า $\hat{\theta}_1$ มีความแม่นยำสูงกว่า $\hat{\theta}_2$ ก็ต่อเมื่อ $MPRE(\hat{\theta}_1) < MPRE(\hat{\theta}_2)$ กล่าวคือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ยของตัวประมาณค่าใด มีค่าต่ำสุด แสดงว่า ตัวประมาณค่านั้นมีความแม่นยำสูงกว่า

ขอบเขตของการวิจัยสำหรับวัตถุประสงค์การวิจัยข้อ 3

ศึกษาผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) โดยประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 จำนวน 8 เดือน มีรายละเอียดของตัวแปรที่ศึกษาดังนี้

1. ตัวแปรที่ศึกษาสำหรับตัวประมาณค่าแบบอัตราส่วน

1.1 ตัวแปรที่ศึกษา (Study Variable) คือ ปริมาณการระเหยน้ำจากเดือนพฤษภาคม พ.ศ. 2556 ถึงเดือนธันวาคม พ.ศ. 2556 รวมทั้งสิ้น 8 เดือน

1.2 ตัวแปรช่วย (Auxiliary Variables) ได้แก่ ตัวแปรอุณหภูมิรายเดือนสองตัวแรกที่มีความสัมพันธ์สูงสุดกับการระเหยน้ำรายเดือนในจังหวัดสกลนคร ที่ผ่านการคัดเลือกตามขอบเขตการวิจัยของการจำลองข้อมูลในวัตถุประสงค์การวิจัยข้อ 2 ซึ่งตัวแปรช่วยในที่นี้ ได้แก่ 1) อุณหภูมิอากาศ และ 2) ความชื้นสัมพัทธ์ จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556

2. การพิจารณาผลการประมาณค่าการระเหยน้ำรายเดือน

เนื่องจากค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในชั้นบรรยากาศไม่สามารถวัดได้โดยตรง และการวัดค่าให้ถูกต้องนั้นทำได้ยาก (Thompson, 1999, p. 91; วิชัย ศรีบุญลือ, 2552, หน้า 8) จึงอาศัยแนวทางการศึกษาของ Dabral et al. (2014) และการศึกษาของ Irmak and Haman (2003) โดยกำหนดค่าพารามิเตอร์การระเหยน้ำจากค่าเฉลี่ยของข้อมูลการระเหยน้ำย้อนหลัง ซึ่ง Irmak and Haman (2003) ได้ประมาณค่าการระเหยน้ำแบบภาคในรัฐฟลอริดาด้วยวิธีการประมาณค่าการระเหยน้ำ 5 วิธี พบว่าทั้ง 5 วิธีนั้นให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์อยู่ระหว่าง 6%

ถึง 32% ดังนั้น การประมาณค่าการระเหยน้ำรายเดือนในการวิจัยนี้ จึงกำหนดให้ค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนในแต่ละเดือน เท่ากับค่าเฉลี่ยเฉพาะเดือนนั้นจากข้อมูลอุตุนิยมหาวิทยาลัยรายเดือนย้อนหลัง 6 ปี (พฤษภาคม พ.ศ. 2550 ถึงเดือนธันวาคม พ.ศ. 2555) และผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) พิจารณาจากค่าเฉลี่ยของเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Mean Percent Relative Error: MPRE) โดย MPRE ไม่เกิน 10% เป็นเกณฑ์ที่ยอมรับได้ในการประมาณค่าการระเหยน้ำ

นิยามศัพท์เฉพาะ

การระเหยน้ำ (Evaporation) หมายถึง ปริมาณน้ำระเหย (ลูกบาศก์เซนติเมตร) ภายในระยะเวลา 24 ชั่วโมง จากเครื่องวัดการระเหยแบบถาด (Evaporation Pan) ชนิด American Class A Pan ในที่นี้ ปริมาณน้ำระเหยจะเท่ากับปริมาณน้ำที่วัดได้จากเครื่องวัดการระเหยแบบถาดในวันปัจจุบัน ลบด้วยปริมาณน้ำที่วัดได้ในวันก่อนหน้า

อุณหภูมิอากาศ (Temperature) หมายถึง อุณหภูมิอากาศที่อ่านจากเครื่องเทอร์โมมิเตอร์ที่ติดตั้งอยู่ภายในเรือนเทอร์โมมิเตอร์ (Thermometer Screen) ในสนามอุตุนิยมหาวิทยาลัย สำนักงานอุตุนิยมหาวิทยาลัยจังหวัดสกลนคร มีหน่วยวัดเป็นองศาเซลเซียส ($^{\circ}\text{C}$)

อุณหภูมิอากาศเฉลี่ยรายวัน (Daily Mean Temperature) หมายถึง อุณหภูมิอากาศเฉลี่ยในแต่ละวัน โดยเจ้าพนักงานอุตุนิยมหาวิทยาลัยจะวัดอุณหภูมิอากาศ 8 ช่วงเวลา ได้แก่ เวลา 01.00 น. 04.00 น. 07.00 น. 10.00 น. 13.00 น. 16.00 น. 19.00 น. และ 22.00 น. แล้วนำค่าอุณหภูมิอากาศทั้ง 8 ช่วงเวลามาเฉลี่ยเป็นอุณหภูมิอากาศเฉลี่ยรายวันของวันนั้น ๆ

อุณหภูมิอากาศเฉลี่ยรายเดือน (Monthly Mean Temperature) หมายถึง อุณหภูมิอากาศเฉลี่ยในแต่ละเดือน เมื่อสิ้นเดือนใด ๆ อุณหภูมิอากาศเฉลี่ยรายวันจะถูกนำมารวมกัน แล้วหารด้วยจำนวนวันในเดือนนั้น ๆ เป็นอุณหภูมิอากาศเฉลี่ยรายเดือน

ความชื้นสัมพัทธ์ (Relative Humidity) หมายถึง อัตราส่วนระหว่างความดันไอน้ำจริงต่อความดันไอน้ำอิ่มตัวที่อุณหภูมิเดียวกัน มีหน่วยเป็นร้อยละ

ตัวแปรช่วย (Auxiliary Variable) หมายถึง ตัวแปรเชิงปริมาณอื่นที่ไม่ใช่ตัวแปรที่สนใจศึกษา แต่มีความสัมพันธ์กับตัวแปรที่สนใจศึกษา

ค่าผิดปกติ (Outlier) หมายถึง ค่าข้อมูลตัวแปรที่สนใจศึกษา (Y) ที่อยู่ห่างไกลจากข้อมูลส่วนใหญ่ ๆ ที่อยู่ในกลุ่ม ทั้งนี้พิจารณาได้จากค่าข้อมูลที่แปลงเป็นคะแนนมาตรฐานแล้ว จะมีค่าคะแนนมาตรฐานมากกว่า 3 ขึ้นไป หรือเป็นข้อมูลที่มีค่ามากกว่า 1.5 เท่าของความกว้างของกล่องในแผนภาพกล่อง แต่ไม่เกิน 3 เท่าของความกว้างของกล่องในแผนภาพกล่อง นั่นคือ ข้อมูลที่มีค่ามากกว่า $Q_3 + 1.5IQR$ แต่ไม่เกิน $Q_3 + 3IQR$ หรือข้อมูลที่มีค่าน้อยกว่า $Q_1 - 1.5IQR$ แต่ไม่น้อยกว่า $Q_1 - 3IQR$

รูปแบบการแจกแจงของตัวแปรช่วย (Distribution of Auxiliary Variable) หมายถึง การแจกแจงทางสถิติของตัวแปรช่วย ซึ่งรูปแบบการแจกแจงของตัวแปรช่วยนี้ จะแสดงถึงลักษณะ

การเกิดของข้อมูล ค่าพารามิเตอร์ ตลอดจนความสมมาตรหรือไม่สมมาตรของข้อมูลตัวแปรช่วย ทั้งนี้ การระบุรูปแบบการแจกแจงของตัวแปรช่วยจะทดสอบด้วยสถิติทดสอบ Anderson-Darling Test สัมประสิทธิ์ความเบ้ (Coefficient of Skewness) หมายถึง ค่าตัวเลขที่ใช้วัดความสมมาตรของเส้นโค้งการแจกแจงข้อมูล โดยที่เส้นโค้งการแจกแจงข้อมูลใดมีความสมมาตรแล้ว สัมประสิทธิ์ความเบ้จะมีค่าเท่ากับศูนย์

ตัวประมาณค่าด้วยวิธีกำลังสองตัดขอบน้อยสุด (Least Trimmed Squares: LTS) หมายถึง ตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ทำให้ผลรวมกำลังสองของค่าความคลาดเคลื่อน (Error) ของข้อมูล h จำนวน h มีค่าน้อยสุด วิธีกำลังสองตัดขอบน้อยสุดนี้เสนอโดยครั้งแรกโดย Rousseeuw ในปี ค.ศ. 1983

ความแม่นยำ (Accuracy) หมายถึง สมบัติของตัวประมาณค่าใดมีค่าประมาณที่เข้าใกล้หรือเท่ากับค่าพารามิเตอร์ หากค่าประมาณมีค่าเท่ากับค่าพารามิเตอร์ ถือว่าค่าประมาณนั้นมีความแม่นยำสูงสุด การพิจารณาความแม่นยำของตัวประมาณค่าในที่นี้ พิจารณาจากค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ หากตัวประมาณค่าใด ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ต่ำสุด ถือว่าเป็นตัวประมาณค่าที่มีความแม่นยำสูงกว่า

ความเที่ยงตรง (Precision) หมายถึง สมบัติของตัวประมาณค่าที่แสดงถึงความเที่ยงตรงในการวัดค่าซ้ำหลาย ๆ ครั้งแล้ว ค่าที่ได้จะมีค่าใกล้เคียงกัน หรือค่าประมาณที่ได้เหล่านั้นมีการกระจายน้อย มากน้อยเพียงใด ถ้าหากตัวประมาณค่าใด ให้ค่าประมาณในการทดลองซ้ำ ๆ หลายครั้งแล้ว ไม่มีการกระจายเลย หรือให้ค่าประมาณที่มีค่าเท่ากันทุกครั้งแล้ว ถือว่าตัวประมาณค่านี้มีความเที่ยงตรงสูงสุด ในที่นี้พิจารณาความเที่ยงตรงของตัวประมาณค่าจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย

ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) หมายถึง ค่าที่ใช้วัดความมีประสิทธิภาพของตัวประมาณค่า ซึ่งคำนวณได้จากผลรวมของผลต่างของค่าจริงกับค่าประมาณยกกำลังสอง แล้วหารด้วยจำนวนข้อมูล แต่สำหรับการหาค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า ใช้นิยาม $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$ โดยที่ตัวประมาณค่าใดมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ถือเป็นตัวประมาณค่าที่มีความเที่ยงตรงมากที่สุด

ประสิทธิภาพของตัวประมาณค่า (Efficiency of Estimator) หมายถึง สมบัติของตัวประมาณค่าที่ดี มีค่าประมาณที่ใกล้เคียงกับค่าพารามิเตอร์มากที่สุด และมีความคลาดเคลื่อนน้อยที่สุด ในที่นี้พิจารณาประสิทธิภาพของตัวประมาณค่าจากความเที่ยงตรงสูงสุดและความแม่นยำสูงสุด

ค่าคลาดเคลื่อนสัมพัทธ์ (Relative Error) หมายถึง อัตราส่วนของความคลาดเคลื่อนระหว่างค่าจริงกับค่าประมาณ เมื่อเทียบกับค่าจริง ซึ่งใช้วัดความแม่นยำในการประมาณค่าของตัวประมาณค่า สามารถคำนวณได้จากค่าสัมบูรณ์ของผลต่างระหว่างค่าจริงกับค่าประมาณ หารด้วยค่าจริง

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การศึกษาการประมาณค่าการระเหยน้ำโดยใช้ตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่ เมื่อใช้ตัวแปรช่วยสองตัว ผู้วิจัยนำเสนอทฤษฎี แนวคิดและงานวิจัยที่เกี่ยวข้อง ดังนี้

ตอนที่ 1 ตัวประมาณค่าแบบอัตราส่วนและงานวิจัยที่เกี่ยวข้อง

ตอนที่ 2 การระเหยน้ำ ข้อมูลอุตุนิยมวิทยาและงานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 ตัวประมาณค่าแบบอัตราส่วนและงานวิจัยที่เกี่ยวข้อง

นิยาม ทฤษฎี และเอกสารที่เกี่ยวข้องกับตัวประมาณค่าแบบอัตราส่วน มีดังนี้

1. นิยามและทฤษฎีที่เกี่ยวข้องกับการสำรวจตัวอย่าง

1.1 ประชากร (Population)

ประชากร หมายถึง เซตของบุคคล สิ่งของ หรือเหตุการณ์ทั้งหมดที่ต้องการศึกษา สามารถแบ่งประชากรเป็น 2 ชนิด ได้แก่

1.1.1 ประชากรจำกัด (Finite Population) คือ ประชากรที่สามารถนับจำนวนหน่วยตัวอย่างที่เป็นสมาชิกของประชากรได้

1.1.2 ประชากรอนันต์ (Infinite Population) คือ ประชากรที่ไม่สามารถนับจำนวนหน่วยตัวอย่างที่เป็นสมาชิกของประชากรได้

1.2 ตัวอย่าง (Sample)

ตัวอย่าง หมายถึง เซตย่อยของประชากร ซึ่งเป็นกลุ่มของหน่วยตัวอย่างที่เลือกจากประชากร (Dunn & Clark, 2009, pp. 13-14) เพื่อใช้เป็นตัวแทนของประชากรในการศึกษาเกี่ยวกับลักษณะของประชากรที่ต้องการศึกษา

1.3 พารามิเตอร์ (Parameters)

พารามิเตอร์ หมายถึง ค่าคงตัวที่แสดงคุณลักษณะของประชากร และพารามิเตอร์เป็นฟังก์ชันของค่าของหน่วยข้อมูลทั้งหมดในประชากร ยกตัวอย่างพารามิเตอร์ที่สำคัญ เช่น

ยอดรวมประชากร (Population Total) คือ $T_Y = \sum_{i=1}^N Y_i$

ค่าเฉลี่ยประชากร (Population Mean) คือ $\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N}$

สัดส่วนประชากร (Population Proportion) คือ $P = A/N$

ความแปรปรวนประชากร (Population Variance) คือ

$$\sigma^2 = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N}$$

อัตราส่วนของประชากร (Population Ratio) คือ $R = \frac{T_Y}{T_X} = \bar{Y}/\bar{X}$ (Cochran, 1977, pp. 20-21) สำหรับกรณีการเก็บรวบรวมข้อมูลสองตัวแปร (Y_i, X_i) ที่มีความสัมพันธ์กันจากหน่วยตัวอย่างแต่ละหน่วยในประชากร

เมื่อ N คือ ขนาดประชากร (Population Size)

Y_i คือ ค่าข้อมูลของตัวแปรที่หนึ่งของหน่วยตัวอย่างที่ i ในประชากร
ซึ่ง $i = 1, 2, \dots, N$

X_i คือ ค่าข้อมูลของตัวแปรที่สองของหน่วยตัวอย่างที่ i ในประชากร

A คือ จำนวนหน่วยตัวอย่างของประชากรที่มีลักษณะที่สนใจ

1.4 ตัวสถิติ (Statistic)

ตัวสถิติ หมายถึง ฟังก์ชันของตัวอย่างสุ่ม (Random Sample) ซึ่งฟังก์ชันดังกล่าวจะไม่มีตัวพารามิเตอร์อื่นใดที่ยังไม่ทราบค่าติดอยู่เลย และค่าที่คำนวณได้เรียกว่า ค่าสถิติ ตัวอย่างตัวสถิติที่สำคัญ เช่น

ยอดรวมตัวอย่าง (Sample Total) คือ $\sum_{i=1}^n y_i$

ค่าเฉลี่ยตัวอย่าง (Sample Mean) คือ $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$

สัดส่วนตัวอย่าง (Sample Proportion) คือ $p = \frac{a}{n}$

ความแปรปรวนตัวอย่าง (Sample Variance) คือ $s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{(n-1)}$

อัตราส่วนของตัวอย่าง (Sample Ratio) คือ $\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{\bar{y}}{\bar{x}}$

(Cochran, 1977, pp. 20-21) สำหรับกรณีการเก็บรวบรวมข้อมูลสองตัวแปร (y_i, x_i) ที่มีความสัมพันธ์กันจากหน่วยตัวอย่างแต่ละหน่วย

เมื่อ n คือ ขนาดตัวอย่าง (Sample Size)

y_i คือ ค่าข้อมูลของตัวแปรที่หนึ่งของหน่วยตัวอย่างที่ i ในตัวอย่าง
ซึ่ง $i = 1, 2, \dots, n$

x_i คือ ค่าข้อมูลของตัวแปรที่สองของหน่วยตัวอย่างที่ i ในตัวอย่าง

a คือ จำนวนหน่วยตัวอย่างที่มีลักษณะที่สนใจในตัวอย่าง

1.5 ตัวประมาณค่า (Estimator)

ตัวประมาณค่า หมายถึง ตัวสถิติที่ใช้ประมาณค่าพารามิเตอร์ ค่าที่คำนวณได้จากตัวประมาณค่า เรียกว่า ค่าประมาณ (Estimate) สำหรับการประมาณค่าพารามิเตอร์ θ นั้น ต้องการ

ให้การประมาณค่าพารามิเตอร์มีความถูกต้องและแม่นยำมากที่สุด ฉะนั้นมักใช้ตัวประมาณค่า $\hat{\theta}$ ที่มีสมบัติที่ดี ดังนี้

1.5.1 ความไม่เอนเอียง (Unbiasness)

ตัวประมาณค่า $\hat{\theta}$ ใด ๆ ของพารามิเตอร์ θ จะมีสมบัติเป็นตัวประมาณค่าไม่เอนเอียง (Unbiased estimator) ก็ต่อเมื่อค่าเฉลี่ยของ $\hat{\theta}$ เท่ากับ θ หรือ $E(\hat{\theta}) = \theta$ ทั้งนี้ค่าความเอนเอียงจะลดลง เมื่อขนาดตัวอย่างเพิ่มมากขึ้น

1.5.2 ความต้องกัน (Consistency)

ถ้าตัวประมาณค่า $\hat{\theta}$ เป็นตัวประมาณค่าของพารามิเตอร์ θ ที่มีค่าขึ้นอยู่กับขนาดตัวอย่าง n แล้ว ตัวประมาณค่า $\hat{\theta}$ มีสมบัติความต้องกัน หมายความว่า ถ้าเพิ่มขนาดตัวอย่างมากขึ้นเรื่อย ๆ ตัวประมาณค่า $\hat{\theta}$ นั้น จะมีค่าเข้าใกล้พารามิเตอร์ θ มากขึ้น (Cochran, 1977, p. 21)

1.5.3 ความมีประสิทธิภาพ (Efficiency)

ถ้าตัวประมาณค่า $\hat{\theta}_1$ และตัวประมาณค่า $\hat{\theta}_2$ เป็นตัวประมาณค่าไม่เอนเอียงของพารามิเตอร์ θ แล้ว เรียกตัวประมาณค่า $\hat{\theta}_1$ ว่าเป็นตัวประมาณค่าที่มีประสิทธิภาพสูงกว่า $\hat{\theta}_2$ ก็ต่อเมื่อ $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ แต่ถ้าหากตัวประมาณค่า $\hat{\theta}_1$ และตัวประมาณค่า $\hat{\theta}_2$ เป็นตัวประมาณค่าเอนเอียงของพารามิเตอร์ θ จะพิจารณาประสิทธิภาพตัวประมาณค่าด้วยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) และเรียกตัวประมาณค่า $\hat{\theta}_1$ ว่าเป็น ตัวประมาณค่าที่มีประสิทธิภาพสูงกว่าตัวประมาณค่า $\hat{\theta}_2$ ก็ต่อเมื่อ $MSE(\hat{\theta}_1) < MSE(\hat{\theta}_2)$ ซึ่งค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า $MSE(\hat{\theta})$ (Cochran, 1977, pp. 15-16) นิยามได้ดังนี้

$$\begin{aligned} MSE(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 \\ &= E\left(\hat{\theta} - E(\hat{\theta}) + (E(\hat{\theta}) - \theta)\right)^2 \\ &= E\left(\hat{\theta} - E(\hat{\theta})\right)^2 + E\left(E(\hat{\theta}) - \theta\right)^2 + 2E\left(\hat{\theta} - E(\hat{\theta})\right)\left(E(\hat{\theta}) - \theta\right) \\ &= V(\hat{\theta}) + Bias(\hat{\theta})^2 \end{aligned}$$

เมื่อ $Bias(\hat{\theta})$ คือ ค่าความเอนเอียงของ $\hat{\theta}$

2. ตัวประมาณค่าแบบอัตราส่วน

สำหรับการสำรวจตัวอย่าง เพื่อประมาณค่าพารามิเตอร์ของตัวแปรที่ศึกษา Y เมื่อมีตัวแปรช่วย X ที่มีความสัมพันธ์กับตัวแปรที่ศึกษา ในการประมาณค่าพารามิเตอร์ของ Y โดยใช้สารสนเทศจากตัวแปรช่วย X ย่อมทำให้ได้ค่าประมาณที่แม่นยำมากขึ้น วิธีประมาณค่าเฉลี่ยประชากรโดยใช้สารสนเทศจากตัวแปรช่วย วิธีหนึ่งก็คือ ตัวประมาณค่าแบบอัตราส่วน (Ratio Estimator)

2.1 ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร

เมื่อสุ่มตัวอย่างขนาด n จากประชากรขนาด N ด้วยวิธีการชักตัวอย่างแบบสุ่มเชิงเดียว ชนิดไม่คืนที่ (Simple Random Sampling Without Replacement: SRSWOR)

ให้ (y_i, x_i) เป็นค่าสังเกตของหน่วยตัวอย่างที่ i ของตัวแปรที่ศึกษา Y และตัวแปรช่วย X ซึ่งมีความสัมพันธ์กับตัวแปรที่ศึกษา โดยที่ทราบค่าเฉลี่ยประชากรของตัวแปรช่วย จะได้ตัวประมาณค่าแบบอัตราส่วน (Ratio estimator) ของค่าเฉลี่ยประชากร (\bar{Y}) ดังนี้

$$\bar{y}_r = \frac{\bar{y}}{\bar{x}} \bar{X} = R\bar{X} \quad (2.1)$$

เมื่อ \bar{y} คือ ค่าเฉลี่ยตัวอย่างของตัวแปร Y และ
 \bar{x} คือ ค่าเฉลี่ยตัวอย่างของตัวแปร X

2.2 ประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร

เนื่องจากตัวประมาณค่าแบบอัตราส่วน \bar{y}_r เป็นตัวประมาณค่าเอนเอียง ดังนั้นประสิทธิภาพของตัวประมาณค่า จะพิจารณาด้วยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ซึ่งค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า \bar{y}_r คือ

$$MSE(\bar{y}_r) @ \frac{1-f}{n} (R^2 S_x^2 - 2RS_{xy} + S_y^2) \quad (2.2)$$

เมื่อ R คือ อัตราส่วนของค่าเฉลี่ยของประชากรของตัวแปร Y กับ X หรือ $R = \frac{\bar{Y}}{\bar{X}}$

S_x^2 คือ ค่าความแปรปรวนของประชากรของตัวแปร X

S_y^2 คือ ค่าความแปรปรวนของประชากรของตัวแปร Y

S_{xy} คือ ค่าความแปรปรวนร่วมของประชากรของตัวแปร X และ Y

ถ้าตัวแปร Y และ X มีความสัมพันธ์กันสูง หรืออัตราส่วน $r_i = y_i/x_i$ มีค่าเกือบคงที่ทุกค่า $i = 1, 2, \dots, n$ แล้ว ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรจะมีความแม่นยำสูง นอกจากนี้ Cochran (1977, p. 157) ได้เปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วน \bar{y}_r กับตัวประมาณค่า \bar{y} เมื่อทำการชักตัวอย่างแบบสุ่มเชิงเดียวด้วยตัวอย่างขนาดใหญ่ พบว่า ถ้า R มีค่าเป็นบวกแล้ว ตัวประมาณค่าแบบอัตราส่วน \bar{y}_r มีประสิทธิภาพสูงกว่าตัวประมาณค่า \bar{y} ภายใต้เงื่อนไข

$$\rho_{xy} > \frac{1 - \frac{S_x^2}{\bar{X}^2}}{2 - \frac{S_y^2}{\bar{Y}^2}} \quad (2.3)$$

เมื่อ ρ_{xy} คือ สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ของประชากรของตัวแปร X กับ Y

3. ทฤษฎีสถิติที่เกี่ยวข้อง

ในการวิจัยนี้จะถือว่าข้อมูลที่ต้องการศึกษามีสมบัติเป็นตัวแปรสุ่ม (Random Variable) ดังนั้นทฤษฎีสถิติที่เกี่ยวข้องจะกล่าวถึงรูปแบบการเกิดของข้อมูลตัวแปรที่ศึกษา นั่นคือ การแจกแจงของข้อมูล (Distribution) ซึ่งอธิบายด้วยฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function) ตลอดจนสมบัติที่สำคัญของตัวแปรสุ่มที่สามารถคำนวณได้จากฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่มนั้น

3.1 การแจกแจงของตัวแปรสุ่ม (Distribution of Random Variables)

การแจกแจงของตัวแปรสุ่มที่นำมาใช้ในการวิจัย ประกอบด้วย การแจกแจงปกติ (Normal Distribution) การแจกแจงแกมมา (Gamma Distribution) และการแจกแจงไวบูลล์ (Weibull Distribution) รายละเอียดดังนี้

3.1.1 การแจกแจงปกติ (Normal Distribution)

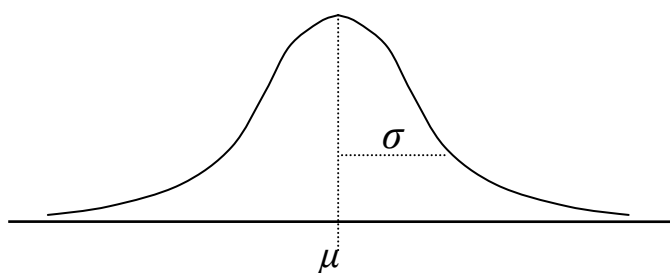
การแจกแจงปกติเป็นการแจกแจงที่สำคัญที่สุดในบรรดาศาสตร์ทางความน่าจะเป็นและสถิติ ซึ่งหลาย ๆ ประชากรที่ศึกษามีการแจกแจงที่เข้าใกล้โค้งปกติ (Normal Curve) ยกตัวอย่างเช่น ส่วนสูง น้ำหนัก เวลาตอบสนองทางการทดลองทางจิตวิทยา การวัดความคลาดเคลื่อนในการทดลองทางวิทยาศาสตร์ หรือลักษณะทางกายภาพอื่น ๆ ถึงแม้ว่า บางครั้งตัวแปรอาจจะไม่มีการแจกแจงปกติ แต่ค่าเฉลี่ยและผลรวมของตัวแปรนั้น ยังคงมีการแจกแจงเข้าใกล้การแจกแจงปกติ โดยอาศัยทฤษฎีบทขีดจำกัดกลาง (Central Limit Theorem) (Devore, 2012, pp. 152-153)

ถ้าตัวแปรสุ่ม X มีการแจกแจงปกติที่มีพารามิเตอร์ $-\infty < \mu < \infty$, $\sigma > 0$ แล้วฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่ม X นิยามโดย

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2} \frac{(x-\mu)^2}{\sigma^2}}; \quad -\infty < x < \infty \quad (2.4)$$

เขียนแทนด้วย $X : N(\mu, \sigma^2)$

ในที่นี้ μ เป็นพารามิเตอร์บอกตำแหน่ง (Location Parameter) ส่วน σ เป็นพารามิเตอร์บอกมาตราส่วน (Scale Parameter) และกราฟของฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงปกติที่มีพารามิเตอร์ μ และ σ ดังภาพที่ 2-1



ภาพที่ 2-1 กราฟการแจกแจงปกติที่มีพารามิเตอร์ μ และ σ

จากภาพที่ 2-1 จะเห็นว่า กราฟของฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงปกติ มีลักษณะเป็นโค้งระฆังคว่ำ ที่สมมาตรที่ $x = \mu$ ดังนั้น ณ จุดกึ่งกลางของโค้งระฆังคว่ำนี้จะเป็นทั้งค่าเฉลี่ย ค่ามัธยฐาน และฐานนิยม

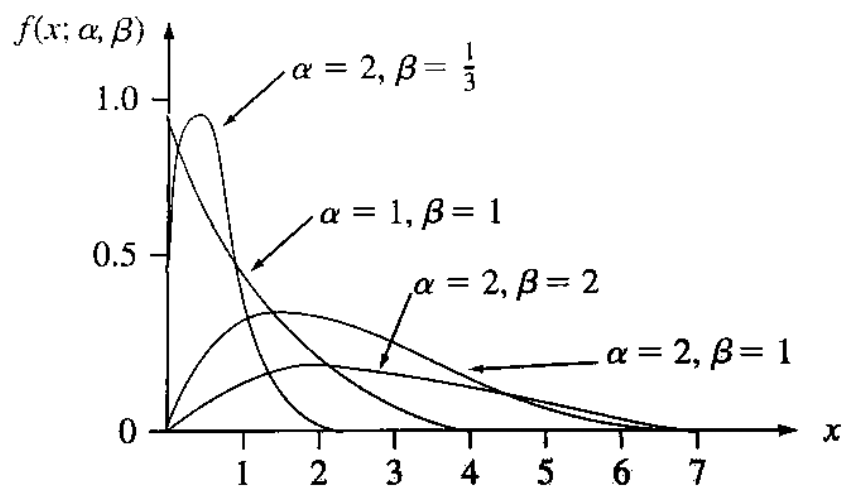
3.1.2 การแจกแจงแกมมา (Gamma Distribution)

ถ้าตัวแปรสุ่ม X มีการแจกแจงแกมมา ที่มีพารามิเตอร์ $\alpha > 0$, $\beta > 0$ แล้ว ฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่ม X นิยามโดย

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & ; x > 0 \\ 0 & ; x \leq 0 \end{cases} \quad (2.5)$$

เขียนแทนด้วย X : $Gamma(\alpha, \beta)$

ในที่นี้ α เป็นพารามิเตอร์บอกรูปร่าง (Shape Parameter) ส่วน β เป็นพารามิเตอร์บอกมาตราส่วน (Scale Parameter) และกำหนดให้ $\Gamma(\alpha)$ เป็นฟังก์ชันแกมมา โดยนิยาม $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy$, $\alpha > 0$ กราฟของฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแกมมา ส่วนใหญ่จะมีลักษณะเบ้ขวา แต่ถ้าพารามิเตอร์ α มีค่าเพิ่ม และพารามิเตอร์ β มีค่าลดลง แล้วกราฟดังกล่าวจะลู่เข้าใกล้การแจกแจงปกติ (Devore, 2012, pp. 167-168) ดังภาพที่ 2-2



ภาพที่ 2-2 กราฟของการแจกแจงแกมมาที่มีพารามิเตอร์ $\alpha > 0$ และ $\beta > 0$

3.1.3 การแจกแจงไวบูลล์ (Weibull Distribution)

ถ้าตัวแปรสุ่ม X มีการแจกแจงไวบูลล์ ที่มีพารามิเตอร์ $\alpha > 0$, $\beta > 0$ แล้วฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่ม X นิยามโดย

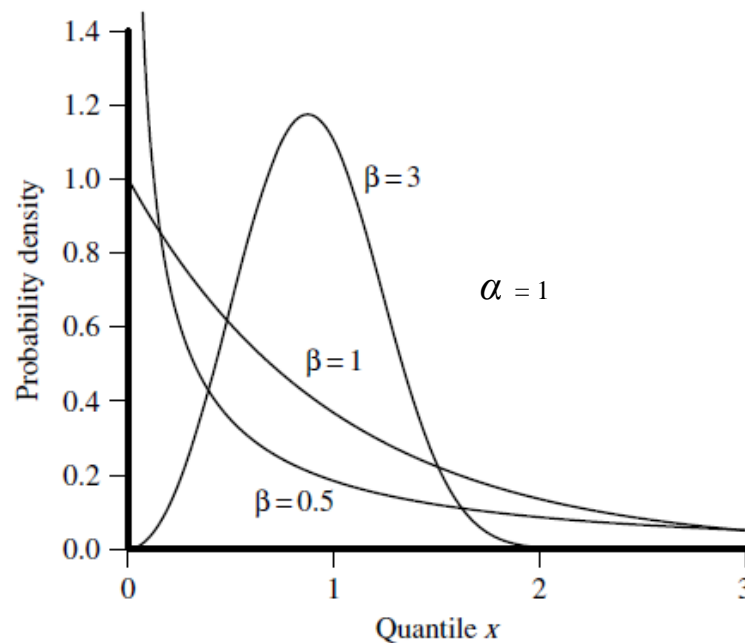
$$f(x; \alpha, \beta) = \begin{cases} \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta} & ; x > 0 \\ 0 & ; x \leq 0 \end{cases} \quad (2.6)$$

เขียนแทนด้วย $X : Weibull(\alpha, \beta)$

ในที่นี้ α เป็นพารามิเตอร์บอกมาตราส่วน (Scale Parameter) ส่วน β เป็นพารามิเตอร์บอกรูปร่าง (Shape Parameter) และกำหนดให้ $\Gamma(\alpha)$ เป็นฟังก์ชันแกมมา โดยนิยาม

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy, \quad \alpha > 0 \text{ เมื่อ } \beta \text{ เท่ากับ } 1 \text{ การแจกแจงไวบูลล์จะลดรูปกลายเป็น}$$

การแจกแจงแบบเอ็กโปเนนเชียล การแจกแจงไวบูลล์มักถูกนำมาใช้ในการจำลองแรงดึงความแข็งแรงของวัสดุ โดยเฉพาะวัสดุที่เปราะ เช่น โบรอนและคาร์บอน กราฟของฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงไวบูลล์ ที่มีพารามิเตอร์ต่าง ๆ (Ryan, 2007, p. 102; Forbes, Evans, Hastings, & Peacock, 2011, pp. 193-194) ดังภาพที่ 2-3



ภาพที่ 2-3 กราฟของการแจกแจงไวบูลล์ที่มีพารามิเตอร์ $\alpha = 1$ และ $\beta > 0$

3.2 สมบัติที่สำคัญของตัวแปรสุ่ม (The Important Property of Random Variable)

สมบัติที่สำคัญของตัวแปรสุ่มที่มักจะศึกษา ได้แก่ การหาค่าคาดหวังหรือค่าเฉลี่ย การวัดการกระจายหรือการหาค่าความแปรปรวน รวมทั้งความเบ้ (Skewness) และความโด่ง (Kurtosis) อันเป็นลักษณะรูปแบบการเกิดของข้อมูลตัวแปรสุ่มนั้น สำหรับการอธิบายสมบัติที่สำคัญข้างต้น จะอธิบายด้วยนิยามของโมเมนต์ (Ryan, 2007, pp. 106-107) ดังนี้

3.2.1 นิยามโมเมนต์ (Definition of Moments)

นิยาม ให้ X เป็นตัวแปรสุ่มใด ๆ ให้ c เป็นค่าคงตัว และ k เป็นจำนวนเต็มบวก ถ้า $E(|X - c|^k) < \infty$ เราเรียก $E(|X - c|^k)$ ว่า โมเมนต์อันดับที่ k ของ X รอบจุด c

3.2.2 ค่าคาดหวัง (Expected Value)

การหาค่าคาดหวังหรือค่าเฉลี่ยของตัวแปรสุ่ม X (μ_x) เมื่อ X เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็น $f(x)$ สามารถคำนวณได้จากโมเมนต์อันดับที่ 1 ของ X รอบจุดกำเนิด (ศูนย์) ดังนี้

$$\mu_x = E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (2.7)$$

3.2.3 ความแปรปรวน (Variance)

ถ้าตัวแปรสุ่ม X เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็น $f(x)$ สามารถวัดการกระจายได้ด้วยความแปรปรวน (Variance) ของตัวแปรสุ่ม X เขียนแทนด้วย $V(X)$ ได้จากโมเมนต์อันดับที่ 2 ของ X รอบค่าเฉลี่ย (μ_x) ดังนี้

$$V(X) = E\left[\sum_{-\infty}^{\infty} (X - \mu_x)^2 f(x)dx\right] \quad (2.8)$$

3.2.4 ความเบ้ (Skewness)

สำหรับตัวแปรสุ่ม X ใด ๆ ที่หาโมเมนต์อันดับที่ 3 รอบค่าเฉลี่ย (μ_3) ได้ ความเบ้หรือสัมประสิทธิ์ความเบ้ของการแจกแจงของตัวแปรสุ่ม X เขียนแทนด้วย γ_1 นิยามได้ดังนี้

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad (2.9)$$

เนื่องจากโมเมนต์รอบค่าเฉลี่ยที่มีอันดับเป็นคี่ จะมีค่าเท่ากับศูนย์ทั้งหมด ดังนั้น การพิจารณาความเบ้ของการแจกแจงของตัวแปรสุ่มใด ๆ ถ้าสัมประสิทธิ์ความเบ้ γ_1 เท่ากับศูนย์แล้ว กราฟของการแจกแจงของตัวแปรสุ่มจะมีลักษณะสมมาตร แต่ถ้าสัมประสิทธิ์ความเบ้ γ_1 มีค่าติดลบ กราฟของการแจกแจงของตัวแปรสุ่มจะมีลักษณะเบ้ทางลบหรือเบ้ซ้าย ในทางตรงกันข้าม

ถ้าสัมประสิทธิ์ความเบ้ γ_1 มีค่าบวก กราฟของการแจกแจงของตัวแปรสุ่มจะมีลักษณะเบ้ทางบวก หรือเบ้ขวา

3.2.5 ความโด่ง (Kurtosis)

สำหรับตัวแปรสุ่ม X ใด ๆ ที่หาโมเมนต์อันดับที่ 4 รอบค่าเฉลี่ย (μ_4) ได้ ความโด่ง หรือสัมประสิทธิ์ความโด่งของการแจกแจงของตัวแปรสุ่ม X เขียนแทนด้วย γ_2 นิยามได้ดังนี้

$$\gamma_2 = \frac{\mu_4}{\sigma^4} \quad (2.10)$$

การพิจารณาความโด่งของการแจกแจงของตัวแปรสุ่มใด ๆ นิยมเปรียบเทียบกับ ความโด่งของการแจกแจงปกติเป็นฐานในการเปรียบเทียบ

4. ค่าผิดปกติ

ค่าผิดปกติ (Outlier) เป็นข้อมูลหรือค่าสังเกตที่มีค่าแตกต่าง ออกจากข้อมูลค่าสังเกต ส่วนใหญ่ที่อยู่ในชุดข้อมูลเดียวกัน ค่าผิดปกติส่งผลกระทบต่อค่าเฉลี่ยและความแปรปรวนของการแจกแจงของข้อมูลที่ศึกษา ในกรณีที่มีค่าผิดปกติในตัวอย่างขนาดเล็ก ก็สามารถบิดเบือนผลของการวิเคราะห์ได้ แต่สำหรับตัวอย่างขนาดใหญ่แล้ว ค่าผิดปกติจะส่งผลกระทบต่อค่าเฉลี่ยเล็กน้อย แต่จะทำให้ความแปรปรวนเพิ่มขึ้น ซึ่งสามารถเปลี่ยนแปลงนัยสำคัญทางสถิติของค่าประมาณสัมประสิทธิ์การถดถอยได้ (Hendra & Staum, 2010, p. 1) เมื่อเก็บรวบรวมข้อมูลตามแบบแผนที่กำหนดไว้แล้ว นักวิจัยควรตรวจสอบความถูกต้องหรือความสมบูรณ์ของข้อมูลก่อนวิเคราะห์ทางสถิติ

วิธีการตรวจสอบค้นหาค่าผิดปกติ อาจทำได้หลายวิธี ดังนี้

1) การใช้แผนภาพฮิสโทแกรม (Histogram) ตรวจสอบความสมมาตรของข้อมูลเบื้องต้น ทั้งนี้เมื่อข้อมูลมีลักษณะไม่สมมาตรแล้ว แผนภาพฮิสโทแกรมสามารถแสดงให้เห็นข้อมูลที่อยู่ไกลนอกกลุ่มข้อมูลอื่น ๆ ซึ่งคาดว่า ข้อมูลนั้นอาจจะเป็นค่าผิดปกติก็ได้

Wilcox (2009, p. 37) แนะนำให้ใช้วิธีอื่นในการตรวจสอบค่าผิดปกติ เช่น การใช้แผนภาพกล่อง (Box Plot) เพื่อให้ได้ผลที่น่าเชื่อถือมากขึ้น

2) การใช้ค่ามาตรฐาน โดยแปลงค่าข้อมูลดิบเป็นค่ามาตรฐานด้วยสมการ $z = \frac{x - \bar{x}}{s}$

ในการพิจารณาค่ามาตรฐานของข้อมูลใด ๆ นั้น Wilcox (2009, pp. 23-24) แนะนำให้ยึดเกณฑ์ค่ามาตรฐานเกิน $|2|$ ก็ถือว่าข้อมูลนั้นเป็นค่าผิดปกติ นอกจากนี้ Wilcox ยังได้แสดงตัวอย่างการหาค่าผิดปกติจากชุดข้อมูลด้วยการใช้ค่ามาตรฐาน ซึ่งให้ผลที่คลุมเครือ ดังนี้

ชุดข้อมูล ($n=10$) ประกอบด้วย 2, 2, 3, 3, 4, 4, 4, 100,000, 100,000 จะได้ค่าเฉลี่ย $\bar{x} = 20,002.5$ และ $s = 42,162.38$ แต่พิจารณาค่า $x_9, x_{10} = 100,000$ เมื่อแปลงเป็นค่ามาตรฐาน จะได้

$$z_{10} = \frac{|100,000 - 20,002.5|}{42,162.38} = 1.897 \quad (2.11)$$

ผลลัพธ์นี้แสดงให้เห็นว่า การใช้ค่ามาตรฐานไม่สามารถระบุค่า 100,000 เป็นค่าผิดปกติ ทั้ง ๆ ที่ข้อมูลค่านี้แตกต่างจากข้อมูลที่เหลือ 8 ค่า อย่างชัดเจน ในกรณีเช่นนี้ Wilcox แนะนำให้ใช้แผนภาพกล่อง (Box Plot) ตรวจสอบค่าผิดปกติ เพื่อให้ผลที่ชัดเจนมากขึ้น

3) การใช้แผนภาพกล่อง (Box Plot) นักสถิติหลายคนแนะนำวิธีใช้แผนภาพกล่อง ตรวจสอบค่าผิดปกติในทิศทางเดียวกัน ดังเช่น กัลยา วานิชย์บัญชา (2554, หน้า 74) และ Wilcox (2009, pp. 42-43) แนะนำวิธีการใช้แผนภาพกล่อง เริ่มจากสร้างแผนภาพกล่องจากค่าสถิติ 5 ค่า ประกอบด้วย

3.1) ค่าต่ำสุดของข้อมูลที่ยังไม่ต่ำผิดปกติ คือ ค่าต่ำสุดของข้อมูลชุดนั้น ๆ ที่มีค่าไม่ต่ำกว่า $Q_1 - 1.5/IQR$ เมื่อ IQR คือ ค่าพิสัยควอร์ไทล์ (Interquartile Range) ซึ่งเท่ากับ $Q_3 - Q_1$

3.2) ควอร์ไทล์ที่ 1: Q_1

3.3) ควอร์ไทล์ที่ 2: Q_2

3.4) ควอร์ไทล์ที่ 3: Q_3

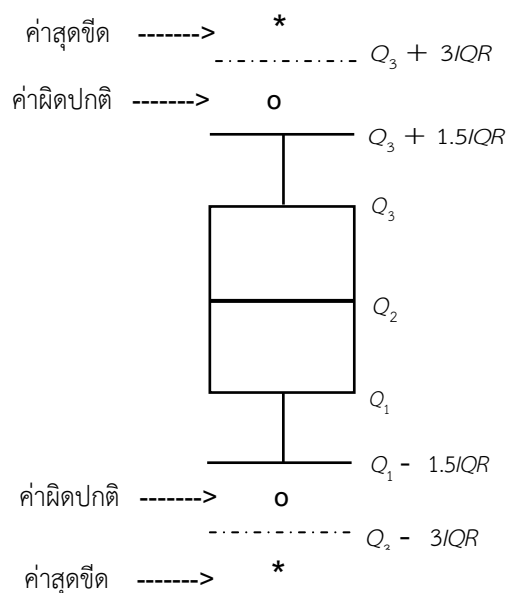
3.5) ค่าสูงสุดของข้อมูลที่ยังไม่สูงผิดปกติ คือ ค่าสูงสุดของข้อมูลชุดนั้น ๆ ที่มีค่าไม่เกิน $Q_3 + 1.5/IQR$

ทั้งนี้แผนภาพกล่องสามารถแสดงค่าผิดปกติได้ 2 ลักษณะ ได้แก่

1) ค่าผิดปกติ (Outlier) หมายถึง ข้อมูลที่มีค่าระหว่าง 1.5 ถึง 3 เท่าของความกว้างของกล่อง นั่นคือ ข้อมูลที่มีค่ามากกว่า $Q_3 + 1.5/IQR$ แต่ไม่เกิน $Q_3 + 3/IQR$ หรือข้อมูลที่มีค่าน้อยกว่า $Q_1 - 1.5/IQR$ แต่ไม่น้อยกว่า $Q_1 - 3/IQR$ จะแสดงค่าผิดปกติด้วยเครื่องหมายวงกลม (o)

2) ค่าสุดขีด (Extreme) หมายถึง ข้อมูลที่มีค่ามากกว่า 3 เท่าของความกว้างของกล่อง นั่นคือ ข้อมูลที่มีค่ามากกว่า $Q_3 + 3/IQR$ หรือข้อมูลที่มีค่าน้อยกว่า $Q_1 - 3/IQR$ จะแสดงค่าสุดขีดด้วยเครื่องหมายดอกจัน (*)

สรุปการใช้แผนภาพกล่องเพื่อตรวจสอบค่าผิดปกติ ได้ดังภาพที่ 2-4



ภาพที่ 2-4 การใช้แผนภาพกล่องเพื่อตรวจสอบค่าผิดปกติ

ในกรณีที่มีค่าผิดปกติอยู่ในชุดข้อมูลที่ศึกษา นักวิจัยควรตรวจสอบค่าผิดปกติว่าเกิดขึ้นเนื่องจากสาเหตุใด หากเป็นสาเหตุที่ชัดเจน สามารถแก้ไขได้ นักวิจัยก็ควรปรับแก้ค่าผิดปกติให้ถูกต้อง หากไม่สามารถหาเหตุผลหลักฐานที่อธิบายการเกิดค่าผิดปกติได้แล้ว Hendra and Staum (2010, p. 9) ได้แนะนำให้ นักวิจัยตัดสินใจระหว่างไม่นำค่าผิดปกติมารวมวิเคราะห์ หรือจะเลือกวิธีปรับให้เหมาะสมกับสถานการณ์ เช่น อาจจะใช้สถิติที่มีความแกร่ง

5. การวิเคราะห์การถดถอยแกร่ง

การศึกษาในเรื่องการวิเคราะห์การถดถอยแกร่งจำนวนมาก โดย The Princeton Robustness Study (Andrews et al., 1972 cited in Ryan, 2009, p. 348) ข้อค้นพบที่ได้ คือ ตัวประมาณค่าโดยวิธีกำลังสองน้อยที่สุด (Ordinary Least Squares: OLS) มีประสิทธิภาพต่ำกว่าตัวประมาณค่าแบบอื่น ๆ เมื่อการแจกแจงของความคลาดเคลื่อนมีลักษณะเบ้ ฉะนั้นเมื่อชุดข้อมูลมีค่าผิดปกติรวมอยู่ วิธีการประมาณค่าด้วยวิธีกำลังสองน้อยที่สุด อาจทำให้ได้ค่าประมาณที่ไม่มีความเสถียร (Stable) ดังนั้นการวิจัยนี้จึงเลือกการวิเคราะห์การถดถอยแกร่งมาประยุกต์ในการประมาณค่าเฉลี่ยประชากร

กำหนดตัวแบบการถดถอยเชิงเส้นพหุ (Multiple Linear Regression Model) ดังนี้

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i; \quad i = 1, 2, \dots, n \quad (2.12)$$

เมื่อ Y คือ ตัวแปรตาม

X_j คือ ตัวแปรอิสระตัวที่ j เมื่อ $j = 1, 2, \dots, k$

β_0 คือ ค่าตัดแกน Y

β_j คือ สัมประสิทธิ์การถดถอยบางส่วนตัวที่ j เมื่อ $j = 1, 2, \dots, k$

ε_i คือ ความคลาดเคลื่อนสุ่ม ซึ่งมีข้อตกลงเบื้องต้น $\varepsilon_i : NID(0, \sigma^2)$

สำหรับการประมาณค่าความคลาดเคลื่อน (e_i) คำนวณได้จากความแตกต่างระหว่างค่าจริงกับค่าประมาณ แทนด้วยสมการ $e_i = Y_i - \hat{Y}_i; i = 1, 2, \dots, n$ โดยที่ $\hat{Y}_i = b_0 + b_1X_{1i} + \dots + b_kX_{ki}$ จากที่กล่าวไว้ข้างต้นว่า การวิเคราะห์การถดถอยแกร่งนั้น มีผู้วิจัยศึกษาเกี่ยวกับวิธีการประมาณค่าสัมประสิทธิ์การถดถอยไว้หลายวิธี ยกตัวอย่างเช่น วิธีกำลังสองมัธยฐานน้อยที่สุด (Least Median Square: LMS) วิธีกำลังสองตัดขอบน้อยที่สุด (Least Trimmed Squares: LTS) วิธี M-Estimation (ME) และวิธี MM-Estimation (MME) เป็นต้น แต่วิธีการเหล่านี้ยังมีประสิทธิภาพที่แตกต่างกันบ้าง ในประเด็นที่ข้อมูลมีค่าผิดปกติ ทั้งในด้านตำแหน่งค่าผิดปกติและปริมาณค่าผิดปกติที่ปะปนอยู่กับชุดข้อมูล ดังการศึกษาของ Yarmohammadi and Mahmoudvand (2010) ที่ว่า ค่าผิดปกติส่งผลต่อประสิทธิภาพการประมาณค่าพารามิเตอร์การถดถอย เมื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าชนิดแกร่งเหล่านั้น พบว่า ประสิทธิภาพของตัวประมาณค่าด้วยวิธี LMS กับตัวประมาณค่าด้วยวิธี LTS มีประสิทธิภาพใกล้เคียงกัน (ซึ่งหลาย ๆ สถานการณ์จำลองที่ศึกษานั้น ตัวประมาณค่าด้วยวิธี LTS มีประสิทธิภาพสูงกว่าตัวประมาณแบบอื่น ๆ) นอกจากนี้ยังพบว่า ประสิทธิภาพของตัวประมาณค่า จะมีค่าเพิ่มขึ้นตามขนาดกลุ่มตัวอย่างและร้อยละของจำนวนค่าผิดปกติ สำหรับการวิจัยนี้เลือกวิธีการประมาณค่าสัมประสิทธิ์ด้วยวิธีกำลังสองตัดขอบน้อยที่สุด (LTS) มาใช้ในปรับตัวประมาณค่าแบบอัตราส่วนโดยใช้ตัวแปรช่วย 2 ตัว

การวิเคราะห์การถดถอยแกร่งเมื่อประมาณสัมประสิทธิ์ด้วยวิธีกำลังสองตัดขอบน้อยที่สุด (LTS) มีหลักการประมาณค่าพารามิเตอร์ พิจารณาว่าค่าสัมประสิทธิ์การถดถอยที่ทำให้ผลรวมของค่าคลาดเคลื่อนกำลังสองของข้อมูลที่ปรับแต่ง (Trimmed Data) h จำนวน มีค่าต่ำที่สุด กล่าวคือ ถ้าข้อมูล n ค่า สามารถปรับแต่งข้อมูลได้ โดยเรียงลำดับค่าคลาดเคลื่อนกำลังสองของข้อมูลแต่ละค่าจากน้อยที่สุดไปหามากที่สุด (Rousseeuw & Leroy, 2003, p. 132) ดังสมการ (2.13)

$$\sum_{i=1}^h (e^2)_{1:n} \quad \sum_{i=2}^h (e^2)_{2:n} \quad \sum_{i=n-h+1}^n (e^2)_{n:n} \quad (2.13)$$

จากนั้นเลือกข้อมูลที่ปรับแต่ง (Trimmed Data) h จำนวน ที่ทำให้ผลรวมของค่าคลาดเคลื่อนกำลังสอง มีค่าน้อยที่สุด ดังสมการ (2.13)

$$\text{Minimize } \sum_{i=1}^h (e^2)_{i:n} \quad (2.14)$$

โดยที่ $(e^2)_{i:n}$ คือ สถิติลำดับของค่าคลาดเคลื่อนกำลังสอง

n คือ ขนาดตัวอย่าง

h คือ จำนวนข้อมูล h ค่า ที่ปรับแต่ง (Trimmed) จากข้อมูลทั้งหมด n ค่า

Rousseeuw and Leroy (2003, p. 132) แนะนำให้กำหนด $h = \lfloor n/2 \rfloor + \lfloor (p+1)/2 \rfloor$

โดย p คือ จำนวนพารามิเตอร์ที่ประมาณค่าในเส้นถดถอย เพื่อให้ตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธี LTS มีโอกาสที่จะได้ค่า Breakdown Point ที่มีค่าสูงสุดได้ถึง 50%

6. วิธีอนุกรมเทย์เลอร์

วิธีอนุกรมเทย์เลอร์ เป็นอนุกรมที่ใช้ประมาณค่าฟังก์ชันที่สามารถหาอนุพันธ์ได้อย่างต่อเนื่อง (Continuously Differentiable Function) ซึ่งนิยามดังนี้

ให้ $f(x)$ เป็นฟังก์ชันที่หาอนุพันธ์ทุกอันดับได้ที่จุด x_0 แล้ว

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ &= f(x_0) + \frac{f'(x_0)}{1!} (x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \\ &\quad \frac{f'''(x_0)}{3!} (x - x_0)^3 + \dots \end{aligned}$$

จะเรียกว่า อนุกรมเทย์เลอร์ (Taylor Series) โดยที่ x_0 เป็นค่าคงตัว เรียกว่า จุดศูนย์กลางของอนุกรม (Burden, & Faires, 2011, pp. 10-11)

สำหรับงานวิจัยนี้จะใช้อนุกรมเทย์เลอร์ดีกรีอันดับหนึ่งที่กำหนดไว้ ดังสมการ (2.15)

$$\begin{aligned} h(\bar{x}, \bar{z}, \bar{y}) &\cong h(\bar{X}, \bar{Z}, \bar{Y}) + \left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}} (\bar{x} - \bar{X}) \\ &\quad + \left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{z}} \right|_{\bar{x}, \bar{z}, \bar{y}} (\bar{z} - \bar{Z}) + \left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{y}} \right|_{\bar{x}, \bar{z}, \bar{y}} (\bar{y} - \bar{Y}) \end{aligned} \quad (2.15)$$

เมื่อ $\bar{y}_M = h(\bar{x}, \bar{z}, \bar{y})$ และ $\bar{y} = h(\bar{X}, \bar{Z}, \bar{Y})$

7. งานวิจัยที่เกี่ยวกับตัวประมาณค่าแบบอัตราส่วน

ภทวรรณ แสงนวกิจ (2553) ได้ศึกษาการประมาณค่าเฉลี่ยด้วยตัวประมาณค่าแบบอัตราส่วน เมื่อข้อมูลมีค่าผิดปกติ โดยเสนอตัวประมาณค่าที่ปรับใหม่ด้วยการใช้สัมประสิทธิ์การถดถอยจากวิธีกำลังสองตัดขอบน้อยที่สุดมาใช้ และได้เปรียบเทียบประสิทธิภาพระหว่างตัวประมาณค่าที่นำเสนอกับตัวประมาณค่าของ Kadilar and Cingi ที่นำเสนอในปี ค.ศ. 2004 และปี ค.ศ. 2007 โดยพิจารณาจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า พบว่าตัวประมาณใหม่ที่นำเสนอมีประสิทธิภาพสูงกว่าตัวประมาณค่าของ Kadilar and Cingi อย่างมีเงื่อนไข จากนั้นได้จำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณค่าที่นำเสนอกับตัวประมาณค่าของ Kadilar and Cingi (2004, 2007) ในกรณีที่ตัวแปรที่ต้องการศึกษามีค่าผิดปกติ ในสถานการณ์ต่าง ๆ ผลการเปรียบเทียบ พบว่า เมื่อข้อมูลมีค่าผิดปกติ ตัวประมาณค่าใหม่ทั้งสามตัวมีค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณของ Kadilar and Cingi ในทุกสถานการณ์

Abu-Dayyeh et al. (2003) ได้เสนอตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากรโดยใช้สารสนเทศจากตัวแปรช่วย 2 ตัว ได้แก่

$$\bar{y}_a^* = \bar{y} \frac{a_{x_1} \frac{\sigma_1^2}{C_{x_1}}}{\frac{\sigma_1^2}{C_{x_1}} + \frac{a_{x_2} \frac{\sigma_2^2}{C_{x_2}}}{\frac{\sigma_2^2}{C_{x_2}}}} \quad (2.16)$$

$$\bar{y}_w^* = \bar{y} \omega_1 \frac{a_{x_1} \frac{\sigma_1^2}{C_{x_1}}}{\frac{\sigma_1^2}{C_{x_1}} + \frac{a_{x_2} \frac{\sigma_2^2}{C_{x_2}}}{\frac{\sigma_2^2}{C_{x_2}}}} + \bar{y} \omega_2 \frac{a_{x_2} \frac{\sigma_2^2}{C_{x_2}}}{\frac{\sigma_1^2}{C_{x_1}} + \frac{a_{x_2} \frac{\sigma_2^2}{C_{x_2}}}{\frac{\sigma_2^2}{C_{x_2}}}} \quad (2.17)$$

$$\text{โดยที่ } a_1^* = - \frac{c_y b_{y1x}}{c_{x_1}}, \quad a_2^* = - \frac{c_y b_{y2x}}{c_{x_2}}, \quad w_1^* + w_2^* = 1,$$

$$w_1^* = \frac{a_2^2 C_{x_2}^2 + a_2 \rho_{yx_2} C_y C_{x_2} - a_1 \rho_{yx_1} C_y C_{x_1} - a_1 a_2 \rho_{x_1 x_2} C_{x_1} C_{x_2}}{a_1^2 C_{x_1}^2 + a_2^2 C_{x_2}^2 - 2 a_1 a_2 \rho_{x_1 x_2} C_{x_1} C_{x_2}} \quad \text{ทำให้ MSE น้อยสุด}$$

$$\text{ดังนั้น } MSE_{\min}(\bar{y}_a^*) = \bar{y}^2 C_y^2 \frac{f}{n} (1 - \rho_{yx_2}^2) \quad \text{โดยที่ } \rho_{yx_2}^2 = \frac{\rho_{yx_1}^2 + \rho_{yx_2}^2 - 2 \rho_{yx_1} \rho_{yx_2} \rho_{x_1 x_2}}{1 - \rho_{x_1 x_2}^2}$$

$$MSE_{\min}(\bar{y}_w^*) = \bar{y}^2 \frac{f}{n} \left(\frac{c_y^2}{C_y} + w_1^{*2} (a_1^2 C_{x_1}^2 - a_2^2 C_{x_2}^2 - 2 a_1 a_2 \rho_{x_1 x_2} C_{x_1} C_{x_2}) + a_2^2 C_{x_2}^2 + a_2 \rho_{yx_2} C_y C_{x_2} - 2 w_1^* (a_2^2 C_{x_2}^2 + a_2 \rho_{yx_2} C_y C_{x_2} - a_1 \rho_{yx_1} C_y C_{x_1} - a_1 a_2 \rho_{x_1 x_2} C_{x_1} C_{x_2}) \right)$$

จากการแสดงเชิงตัวเลขกับข้อมูลจริงและสถานการณ์จำลอง พบว่า ตัวประมาณค่า \bar{y}_a^* และ \bar{y}_w^* มีประสิทธิภาพสูงอันดับ 1 และอันดับ 2 ตามลำดับ เมื่อเปรียบเทียบกับตัวประมาณค่า \bar{y} และตัวประมาณค่าอื่น ๆ

Kadilar and Cingi (2004) ได้นำตัวประมาณค่าแบบอัตราส่วนที่ใช้หลักการวิเคราะห์

$$\text{การถดถอย } \bar{y}_{pr} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X} = \hat{R}_{pr} \bar{X} \quad \text{มาร่วมในการปรับตัวประมาณค่าแบบ}$$

อัตราส่วนของ Sisodia and Dwivedi ที่เสนอไว้ในปี ค.ศ. 1981 และตัวประมาณค่าแบบอัตราส่วนของ Singh and Kakran ที่เสนอไว้ในปี ค.ศ. 1993 ทำให้ได้ตัวประมาณค่าใหม่สองตัว ได้แก่

$$\bar{y}_{psd} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + C_x} (\bar{X} + C_x) = \hat{R}_{psd} \bar{X}_{SD} \quad (2.18)$$

$$\bar{y}_{psk} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + \beta_{2(x)}} (\bar{X} + \beta_{2(x)}) = \hat{R}_{psk} \bar{X}_{SK} \quad (2.19)$$

ใช้อุกรมเทอร์เลอร์หาสมการ MSE ได้ $MSE(\bar{y}_{pSD}) @ \frac{1-f}{n} (R_{SD}^2 S_X^2 + S_Y^2 (1-\rho^2))$ และ $MSE(\bar{y}_{pSK}) @ \frac{1-f}{n} (R_{SK}^2 S_X^2 + S_Y^2 (1-\rho^2))$ เมื่อทำการเปรียบเทียบประสิทธิภาพพบว่า \bar{y}_{pSD} เป็นตัวประมาณค่าที่ดีกว่าตัวประมาณค่า \bar{y}_{pr} แต่ตัวประมาณค่า \bar{y}_{pr} จะมีประสิทธิภาพสูงกว่าตัวประมาณค่า \bar{y}_{pSK} ภายใต้เงื่อนไข $|\bar{X} + \beta_{2(x)}| < |\bar{X}|$

Kadilar and Cingi (2005) ได้นำเสนอตัวประมาณค่าแบบอัตราส่วนแบบใหม่เมื่อใช้ตัวแปรช่วยสองตัว โดยตัวประมาณค่าเฉลี่ยแบบใหม่ที่นำเสนอ ได้พัฒนาจากตัวประมาณค่าแบบ

อัตราส่วน $\bar{y}_{ratio} = \bar{y} \frac{\alpha_1 \bar{X}_1^{\alpha_1} \alpha_2 \bar{X}_2^{\alpha_2}}{\alpha_1 \bar{X}_1^{\alpha_1} \alpha_2 \bar{X}_2^{\alpha_2}}$ ของ Abu-Dayyeh et al. (2003) ร่วมกับตัวประมาณค่า

แบบถดถอย $\bar{y}_{reg2} = \bar{y} + b_1(\bar{X}_1 - \bar{x}_1) + b_2(\bar{X}_2 - \bar{x}_2)$ ทำให้ได้ตัวประมาณค่าใหม่ที่

Kadilar and Cingi นำเสนอ คือ

$$\bar{y}_{pr} = \bar{y} \frac{\alpha_1 \bar{X}_1^{\alpha_1} \alpha_2 \bar{X}_2^{\alpha_2}}{\alpha_1 \bar{X}_1^{\alpha_1} \alpha_2 \bar{X}_2^{\alpha_2}} + b_1(\bar{X}_1 - \bar{x}_1) + b_2(\bar{X}_2 - \bar{x}_2) \quad (2.20)$$

แล้วใช้อุกรมเทอร์เลอร์หาสมการของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ได้ดังนี้

$$\begin{aligned} MSE(\bar{y}_{pr}) @ \frac{1-f}{n} & \times S_y^2 + (\alpha_1 R_1 + B_1)^2 S_{x_1}^2 + (\alpha_2 R_2 + B_2)^2 S_{x_2}^2 \\ & - 2(\alpha_1 R_1 + B_1) S_{yx_1} - 2(\alpha_2 R_2 + B_2) S_{yx_2} \\ & + 2(\alpha_1 R_1 + B_1)(\alpha_2 R_2 + B_2) S_{x_1 x_2} \end{aligned} \quad (2.21)$$

$$\text{โดยที่ } R_1 = \frac{\bar{Y}}{\bar{X}_1}, R_2 = \frac{\bar{Y}}{\bar{X}_2}, B_1 = \frac{S_{yx_1}}{S_{x_1}^2}, B_2 = \frac{S_{yx_2}}{S_{x_2}^2}$$

$$\alpha_1^* = \frac{S_y}{R_1 S_{x_1}} \rho_1^*, \alpha_2^* = \frac{S_y}{R_2 S_{x_2}} \rho_2^*$$

$$\rho_1^* = \frac{\rho_{x_1 x_2} (\rho_{yx_1} \rho_{x_1 x_2} - \rho_{yx_2})}{1 - \rho_{x_1 x_2}^2}, \rho_2^* = \frac{\rho_{x_1 x_2} (\rho_{yx_2} \rho_{x_1 x_2} - \rho_{yx_1})}{1 - \rho_{x_1 x_2}^2} \text{ เป็นค่า}$$

α_1 และ α_2 ที่เหมาะสม ที่ทำให้ค่า $MSE(\bar{y}_{pr})$ มีค่าน้อยสุด ผลการเปรียบเทียบประสิทธิภาพของ

ตัวประมาณค่าใหม่ \bar{y}_{pr} ที่ Kadilar and Cingi (2005) นำเสนอกับตัวประมาณค่าแบบถดถอย \bar{y}_{reg2} พบว่า ตัวประมาณค่า \bar{y}_{pr} มีประสิทธิภาพสูงกว่าตัวประมาณค่า \bar{y}_{reg2} ภายใต้เงื่อนไข

$$\frac{1-f}{n} S_y^2 (1 - \rho_{yx_1}^2 - \rho_{yx_2}^2 + 2\rho_{yx_1}\rho_{yx_2}\rho_{x_1x_2}) (\rho_1^{*2} + \rho_2^{*2} + 2\rho_{x_1x_2} (\rho_1^*\rho_2^* + \rho_1^*\rho_{yx_2} + \rho_2^*\rho_{yx_1})) < 0$$

จากนั้น Kadilar and Cingi (2005) ได้ประมาณผลผลิตแอปเปิ้ลทั้งหมดในปี ค.ศ. 1999 จากหมู่บ้าน 204 หมู่บ้านในเขต Black Sea ประเทศตุรกี ใช้ตัวแปรช่วย 2 ตัว คือ จำนวนต้นแอปเปิ้ลในปี ค.ศ. 1999 และปริมาณผลผลิตแอปเปิ้ลในปี ค.ศ. 1998 พบว่า ตัวประมาณค่าใหม่ \bar{y}_{pr} ที่นำเสนอมีประสิทธิภาพสูงกว่าตัวประมาณแบบถดถอย \bar{y}_{reg2}

Perri (2005) ได้นำเสนอตัวประมาณค่าแบบอัตราส่วนชนิด Ratio-Cum-Product แบบใหม่ เมื่อใช้การรวมตัวแปรช่วยสองตัว เพื่อให้ได้ตัวประมาณค่าแบบอัตราส่วนที่ใช้ประมาณค่าเฉลี่ยประชากรหรือยอดรวมประชากร โดยตัวประมาณค่าแบบใหม่ที่น่าสนใจ ได้พัฒนาจากตัวประมาณค่าแบบอัตราส่วนที่ Singh ได้นำเสนอไว้ในปี ค.ศ. 1965 และปี ค.ศ. 1967 ดังนี้

$$\hat{Y}_{R1} = \bar{y} \frac{\bar{X}_1 \bar{X}_2}{\bar{X}_1 \bar{X}_2}, \hat{Y}_{R2} = \bar{y} \frac{\bar{X}_1 \bar{X}_2}{\bar{X}_1 \bar{X}_2}, \hat{Y}_{P1} = \bar{y} \frac{\bar{X}_1 \bar{X}_2}{\bar{X}_1 \bar{X}_2} \text{ และ } \hat{Y}_{P2} = \bar{y} \frac{\bar{X}_1 \bar{X}_2}{\bar{X}_1 \bar{X}_2} \quad (2.22)$$

เพื่อให้ได้ตัวประมาณค่าใหม่ที่มีประสิทธิภาพดีกว่าตัวประมาณค่าใน (2.22) Perri จึงแทนค่าเฉลี่ยตัวอย่างของตัวแปรช่วย \bar{x}_i , $i = 1, 2$ ด้วยตัวประมาณค่าที่ไม่เอนเอียง

$\hat{t}_i = \bar{x}_i + \alpha_i (\bar{X}_i - \bar{x}_i)$ เมื่อ α_i เป็นค่าคงตัวที่เหมาะสม จะได้ตัวประมาณค่าแบบใหม่ที่น่าสนใจดังนี้

$$\hat{Y}_{R1}^\alpha = \bar{y} \frac{\hat{t}_2 \bar{X}_1}{\hat{t}_1 \bar{X}_2}, \hat{Y}_{R2}^\alpha = \bar{y} \frac{\bar{X}_1 \bar{X}_2}{\hat{t}_1 \hat{t}_2}, \hat{Y}_{P1}^\alpha = \bar{y} \frac{\hat{t}_1 \hat{t}_2}{\bar{X}_1 \bar{X}_2} \text{ และ } \hat{Y}_{P2}^\alpha = \bar{y} \frac{\hat{t}_1 \bar{X}_2}{\hat{t}_2 \bar{X}_1} \quad (2.23)$$

เกณฑ์ในการประเมินความเที่ยงตรงของตัวประมาณเหล่านั้น พิจารณาจากค่า MSE จึงใช้การกระจายอนุกรมเทย์เลอร์อันดับหนึ่งเพื่อหาสมการ MSE ของตัวประมาณค่าใน (2.23) ดังนี้

$$MSE(\hat{Y}_{R1}^\alpha) B \frac{1-f}{n} \hat{\epsilon}_0^2 + \gamma_1^2 + \gamma_2^2 - 2(\gamma_{01} - \gamma_{02} - \gamma_{12}) \hat{\epsilon}_1 \quad (2.24)$$

$$MSE(\hat{Y}_{R2}^\alpha) B \frac{1-f}{n} \hat{\epsilon}_0^2 + \gamma_1^2 + \gamma_2^2 - 2(\gamma_{01} + \gamma_{02} - \gamma_{12}) \hat{\epsilon}_1 \quad (2.25)$$

$$MSE(\hat{Y}_{P1}^\alpha) B \frac{1-f}{n} \hat{\epsilon}_0^2 + \gamma_1^2 + \gamma_2^2 + 2(\gamma_{01} + \gamma_{02} + \gamma_{12}) \hat{\epsilon}_1 \quad (2.26)$$

$$MSE(\hat{Y}_{P2}^\alpha) B \frac{1-f}{n} \hat{\epsilon}_0^2 + \gamma_1^2 + \gamma_2^2 + 2(\gamma_{01} - \gamma_{02} - \gamma_{12}) \hat{\epsilon}_1 \quad (2.27)$$

โดยที่ $f = \frac{n}{N}$, $\gamma_i^2 = R_i S_i (1 - \alpha_i)$ และ $\gamma_{0i} = R_i S_{0i} (1 - \alpha_i)$

เมื่อ S_0^2, S_1^2, S_2^2 คือ ความแปรปรวนของตัวแปร Y, X_1 และ X_2 ตามลำดับ

S_{01}, S_{02}, S_{12} คือ ความแปรปรวนร่วมของตัวแปร Y กับ X_1 , ตัวแปร Y กับ X_2
และตัวแปร X_1 กับ X_2 ตามลำดับ

จากนั้นทำการสถิติเชิงตัวเลขเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าที่นำเสนอ กับตัวประมาณค่าเดิม โดยประมาณค่าปริมาณขยะรีไซเคิลทั้งหมดของประเทศอิตาลีในปี ค.ศ. 2003 (Y) เมื่อตัวแปรช่วยสองตัว คือ ปริมาณขยะรีไซเคิลทั้งหมดของประเทศอิตาลีในปี ค.ศ. 2002 (X_1) และจำนวนชาวบ้านทั้งหมดในปี ค.ศ. 2003 (X_2) ผลการเปรียบเทียบประสิทธิภาพ พบว่า ตัวประมาณค่าที่นำเสนอมีประสิทธิภาพสูงกว่าทุก ๆ ตัวประมาณค่าเดิม

Kadilar and Cingi (2006) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ย ประชากรชั้นใหม่ 4 ตัว จากตัวประมาณค่า $\bar{y}_{KC1}, \bar{y}_{KC2}, \bar{y}_{KC3}, \bar{y}_{KC4}$ และ \bar{y}_{KC5} ที่ Kadilar and Cingi ได้เสนอไว้ในปี ค.ศ. 2004 โดยอาศัยหลักการนำตัวประมาณค่ามารวมกันแบบถ่วงน้ำหนัก ได้ ตัวประมาณค่าใหม่ 4 ตัว ดังนี้

$$\bar{y}_{pr1} = \omega_1 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X} + \omega_2 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + C_x} (\bar{X} + C_x) \quad (2.28)$$

$$\bar{y}_{pr2} = \omega_1 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X} + \omega_2 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + \beta_{2(x)}} (\bar{X} + \beta_{2(x)}) \quad (2.29)$$

$$\bar{y}_{pr3} = \omega_1 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X} + \omega_2 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} \beta_{2(x)} + C_x} (\bar{x} \beta_{2(x)} + C_x) \quad (2.30)$$

$$\bar{y}_{pr4} = \omega_1 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X} + \omega_2 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} C_x + \beta_{2(x)}} (\bar{x} C_x + \beta_{2(x)}) \quad (2.31)$$

โดยที่ $\omega_1 + \omega_2 = 1$

จากนั้นใช้อินทิกรัลเทอร์เลอร์หาสมการของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของ ตัวประมาณค่าใหม่ ($i = 1, 2, 3, 4$) ได้ดังสมการ (2.32)

$$MSE(\bar{y}_{pr_i}) = \frac{1-f}{n} (S_y^2 - 2\eta S_{xy} + \eta^2 S_x^2) \quad (2.32)$$

เมื่อ $\eta = \omega_1 (B + R) + \omega_2 (B + R_{KC_i})$

โดยค่า ω_1 และ ω_2 ที่เหมาะสม คือ $\omega_1^* = \frac{R_{KC_i}}{R_{KC_i} - R}$ และ $\omega_2^* = \frac{R}{R - R_{KC_i}}$ จะ

ได้ $\eta = \beta$ และทำให้ MSE มีค่าต่ำสุดสำหรับทุกตัวประมาณค่าใหม่ที่เสนอดังสมการ (2.33)

$$MSE_{\min}(\bar{y}_{pr}) = \frac{1-f}{n} (S_y^2 - 2\beta S_{xy} + \beta^2 S_x^2) \quad (2.33)$$

ผลการเปรียบเทียบประสิทธิภาพ พบว่า ตัวประมาณค่าใหม่ทั้งสี่ตัวมีประสิทธิภาพสูงกว่าทุกตัวประมาณค่าเดิม

Gupta and Shabbir (2007) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากร ที่ใช้สารสนเทศจากตัวแปรช่วย 2 ตัว โดยนำตัวประมาณค่า $\bar{y}_{SD} = \bar{y} \frac{\bar{X} + C_x}{\bar{x} + C_x}$ ของ

Sisodia and Dwivedi (1981), ตัวประมาณค่า $\bar{y}_{SK} = \bar{y} \frac{\bar{X} + \beta_{2(x)}}{\bar{x} + \beta_{2(x)}}$ ของ Singh and Kakran

(1993), ตัวประมาณค่า $\bar{y}_{US1} = \bar{y} \frac{\frac{\bar{x}}{e^{\bar{x}}} \beta_{2(x)} + C_x}{\frac{\bar{x}}{e^{\bar{x}}} \beta_{2(x)} + C_x}$ และ $\bar{y}_{US2} = \bar{y} \frac{\frac{\bar{x} C_x}{e^{\bar{x}}} + \beta_{2(x)}}{\frac{\bar{x} C_x}{e^{\bar{x}}} + \beta_{2(x)}}$ ของ

Upadhyaya and Singh (1999) มาพัฒนาเป็นตัวประมาณค่าใหม่ด้วยการเพิ่มพจน์ของตัวแปรช่วยตัวที่สอง นั่นคือเพิ่มตัวแปรช่วย Z เข้าในตัวประมาณค่า \bar{y}_{SD} , \bar{y}_{SK} , \bar{y}_{US1} และ \bar{y}_{US2} ซึ่ง Gupta and Shabbir (2007) ได้พัฒนาเป็นกลุ่มของตัวประมาณค่าใหม่ ที่ใช้ตัวแปรช่วยสองตัว คือ

$$\bar{y}_m = \bar{y} [f_i(x)]^1 [f_i(z)]^2, \quad i = 1, 2, 3, 4 \quad (2.34)$$

ดังนั้นตัวประมาณค่าใหม่ที่พัฒนาจากตัวประมาณค่า \bar{y}_{SD} , \bar{y}_{SK} , \bar{y}_{US1} และ \bar{y}_{US2} ได้แก่

$$\bar{y}_{m(SD)} = \bar{y} \frac{\frac{\bar{x} + C_x}{e^{\bar{x}}} \bar{u}^1 \frac{\bar{z} + C_z}{e^{\bar{z}}} \bar{u}^2}{\frac{\bar{x} + C_x}{e^{\bar{x}}} \bar{u}^1 \frac{\bar{z} + C_z}{e^{\bar{z}}} \bar{u}^2} \quad (2.35)$$

$$\bar{y}_{m(SK)} = \bar{y} \frac{\frac{\bar{x} + \beta_{2(x)}}{e^{\bar{x}}} \bar{u}^1 \frac{\bar{z} + \beta_{2(z)}}{e^{\bar{z}}} \bar{u}^2}{\frac{\bar{x} + \beta_{2(x)}}{e^{\bar{x}}} \bar{u}^1 \frac{\bar{z} + \beta_{2(z)}}{e^{\bar{z}}} \bar{u}^2} \quad (2.36)$$

$$\bar{y}_{m(US1)} = \bar{y} \frac{\frac{\frac{\bar{x}}{e^{\bar{x}}} \beta_{2(x)} + C_x}{e^{\bar{x}}} \bar{u}^1 \frac{\frac{\bar{z}}{e^{\bar{z}}} \beta_{2(z)} + C_z}{e^{\bar{z}}} \bar{u}^2}{\frac{\frac{\bar{x}}{e^{\bar{x}}} \beta_{2(x)} + C_x}{e^{\bar{x}}} \bar{u}^1 \frac{\frac{\bar{z}}{e^{\bar{z}}} \beta_{2(z)} + C_z}{e^{\bar{z}}} \bar{u}^2} \quad (2.37)$$

$$\bar{y}_{m(US2)} = \bar{y} \frac{\bar{C}_x + \beta_{2(x)} \frac{\bar{C}_z + \beta_{2(z)}}{\bar{C}_z + \beta_{2(z)}}}{\bar{C}_x + \beta_{2(x)}} \quad (2.38)$$

ทั้งนี้ จากสมการ (2.34) ถ้า $(J_1, J_2) = (1, 1), (-1, -1), (1, -1), (-1, 1)$ แล้ว ตัวประมาณค่าแบบอัตราส่วนเหล่านั้น จะเป็นตัวประมาณค่าแบบอัตราส่วนชนิด Ratio Type, Product Type, Ratio-Cum-Product Type และ Product-Cum-Ratio Type ตามลำดับ ยกตัวอย่างสมการของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า $\bar{y}_{m(US2)}$ ดังนี้

$$MSE(\bar{y}_{m(US2)}) \approx \bar{y}^2 \lambda^2 C_y^2 + \theta_1^2 J_1^2 C_x^2 + \theta_2^2 J_2^2 C_z^2 - 2\theta_1 J_1 \rho_{yx} C_y C_x - 2\theta_2 J_2 \rho_{yz} C_y C_z + 2\theta_1 \theta_2 J_1 J_2 \rho_{xz} C_x C_z \quad (2.39)$$

$$\text{เมื่อ } \lambda = \frac{N-n}{Nn}, \theta_1 = \frac{\bar{C}_x}{\bar{C}_x + \beta_{2(x)}}, \theta_2 = \frac{\bar{C}_z}{\bar{C}_z + \beta_{2(z)}}$$

โดยที่ $\beta_{2(x)}, \beta_{2(z)}$ เป็นสัมประสิทธิ์ความโค้งของตัวแปร X และตัวแปร Z ตามลำดับ

$$\text{กำหนดค่าเหมาะสม } J_1 = \frac{C_y(\rho_{yx} - \rho_{yz}\rho_{xz})}{\theta_1 C_x(1 - \rho_{xz}^2)} \text{ และ } J_2 = \frac{C_y(\rho_{yz} - \rho_{yx}\rho_{xz})}{\theta_2 C_z(1 - \rho_{xz}^2)}$$

$$\text{ทำให้ได้ค่า } MSE \text{ ต่ำสุด เท่ากับ } MSE_0(\bar{y}_{m(US2)}) \approx \lambda^2 \bar{y}^2 C_y^2 \frac{\rho_{yx}^2 + \rho_{yz}^2 - 2\rho_{yx}\rho_{yz}\rho_{xz}}{1 - \rho_{xz}^2}$$

สำหรับตัวประมาณค่า $\bar{y}_{m(SD)}, \bar{y}_{m(SK)}$ และ $\bar{y}_{m(US1)}$ ต่างก็มีค่า J_1 และ J_2 ที่เหมาะสมแตกต่างกัน เพื่อให้ MSE ต่ำสุด แต่ท้ายสุดก็ได้ค่า MSE ต่ำสุดเท่ากัน คือเท่ากับ $MSE_0(\bar{y}_{m(US2)})$ ผลการเปรียบเทียบประสิทธิภาพ พบว่า ตัวประมาณค่าใหม่ทุกตัวมีประสิทธิภาพสูงกว่าตัวประมาณค่า $\bar{y}, \bar{y}_R, \bar{y}_{SD}, \bar{y}_{SK}, \bar{y}_{US1}$ และ \bar{y}_{US2}

Kadilar et al. (2007) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากร โดยใช้การวิเคราะห์การถดถอยแกร่ง (Robust regression) มาปรับตัวประมาณค่าแบบอัตราส่วน ของ Kadilar and Cingi ที่นำเสนอในปี ค.ศ. 2004 ได้ตัวประมาณค่าใหม่ 5 ตัวดังนี้

$$\bar{y}_{pr1} = \frac{\bar{y} + b_{rob}(\bar{X} - \bar{x})}{\bar{x}} \quad (2.40)$$

$$\bar{y}_{pr2} = \frac{\bar{y} + b_{rob}(\bar{X} - \bar{x})}{\bar{x} + C_x}(\bar{X} + C_x) \quad (2.41)$$

$$\bar{y}_{pr3} = \frac{\bar{y} + b_{rob}(\bar{X} - \bar{x})}{\bar{x} + \beta_{2(x)}} \frac{\bar{X}}{\beta_{2(x)}} + \beta_{2(x)} \frac{\bar{y}}{\beta_{2(x)}} \quad (2.42)$$

$$\bar{y}_{pr4} = \frac{\bar{y} + b_{rob}(\bar{X} - \bar{x})}{\bar{x}\beta_{2(x)} + C_x} \frac{\bar{X}}{\beta_{2(x)}} + C_x \frac{\bar{y}}{\beta_{2(x)}} \quad (2.43)$$

$$\bar{y}_{pr5} = \frac{\bar{y} + b_{rob}(\bar{X} - \bar{x})}{\bar{x}C_x + \beta_{2(x)}} \frac{\bar{X}}{C_x} + \beta_{2(x)} \frac{\bar{y}}{C_x} \quad (2.44)$$

โดยที่ b_{rob} คือ ค่าสัมประสิทธิ์การถดถอยที่ประมาณด้วยวิธี Huber-M ในการวิเคราะห์การถดถอยแกร่ง จากนั้นเปรียบเทียบประสิทธิภาพ พบว่า ตัวประมาณค่าใหม่จะมีประสิทธิภาพสูงกว่าตัวประมาณค่าเดิม ภายใต้เงื่อนไข $-2R_{KC_i} < B_{rob} - B < 0$, $i = 1, 2, 3, 4$ และ 5

Sharma and Tailor (2010) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนชนิด Ratio-Cum-Dual แบบใหม่ จากแนวคิดที่ว่า เมื่อความสัมพันธ์ระหว่างตัวแปรที่ศึกษากับตัวแปรช่วยมีค่าเป็นบวกในระดับมาก ควรใช้ตัวประมาณค่าแบบอัตราส่วนชนิด Ratio Type แต่ถ้าในทางตรงกันข้าม เมื่อความสัมพันธ์ระหว่างตัวแปรที่ศึกษากับตัวแปรช่วยมีค่าเป็นลบ ก็ควรใช้ตัวประมาณค่าแบบอัตราส่วนชนิด Product Type ดังนั้น Sharma and Tailor จึงพัฒนาตัวประมาณค่าแบบอัตราส่วน

Ratio-Cum-Dual แบบใหม่ จากตัวประมาณค่าแบบอัตราส่วน $\bar{y}_R = \bar{y} \frac{\bar{X} \bar{y}}{C_x \bar{X}}$ และตัวประมาณค่า

แบบอัตราส่วนของ Srivenkataramana ที่นำเสนอในปี ค.ศ. 1980 คือ $\bar{y}_R^{(d)} = \bar{y} \frac{\bar{X}^* \bar{y}}{C_x \bar{X}}$

อาศัยการแปลงค่า $\bar{x}_i^* = (N\bar{X} - nx_i)/(N - n)$ โดยนำเสนอตัวประมาณค่าแบบอัตราส่วนชนิด Ratio-Cum-Dual แบบใหม่ ดังนี้

$$\hat{Y}_{bk1} = \bar{y} \frac{\alpha \bar{X} \bar{y}}{C_x \bar{X}} + (1 - \alpha) \frac{\bar{X}^* \bar{y}}{C_x \bar{X}} \quad (2.45)$$

เมื่อ α เป็นค่าคงตัวที่เหมาะสม ซึ่งถ้า $\alpha = 1$ และ $\alpha = 0$ แล้วตัวประมาณค่า \hat{Y}_{bk1} จะลดรูปเป็นตัวประมาณค่า \bar{y}_R และ $\bar{y}_R^{(d)}$ ตามลำดับ เมื่อใช้การประมาณค่าอันดับหนึ่ง จะได้

$$B(\hat{Y}_{bk1}) = \frac{(1 - f)}{n} \bar{Y} C_x^2 \left\{ \alpha - \{(1 - g)\alpha + g\} K \right\} \quad (2.46)$$

$$MSE(\hat{Y}_{bk1}) = \frac{(1-f)}{n} \bar{Y} \left\{ C_y^2 + [(1-g)\alpha + g] C_x^2 [(1-g)\alpha + g - 2K] \right\} \quad (2.47)$$

เมื่อเปรียบเทียบประสิทธิภาพพบว่า ตัวประมาณค่า \hat{Y}_{bk1} มีประสิทธิภาพสูงกว่าตัวประมาณค่า \bar{y}_R ภายใต้เงื่อนไข $\frac{2K - g - 1}{1 - g} < \alpha < 1$ หรือ $1 < \alpha < \frac{2K - g - 1}{1 - g}$ และตัวประมาณค่า \hat{Y}_{bk1} มีประสิทธิภาพสูงกว่าตัวประมาณค่า $\bar{y}_R^{(d)}$ ภายใต้เงื่อนไข $\frac{2(K - g)}{1 - g} < \alpha < \infty$

Yarmohammadi and Mahmoudvand (2010) ได้ศึกษาเปรียบเทียบประสิทธิภาพของค่าผิดปกติที่มีต่อความต้านทาน (Resistant) และความแกร่ง (Robust) ของสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination) ในการวิเคราะห์การถดถอย โดยเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Square: OLS) กับตัวประมาณค่าชนิดแกร่ง ประกอบด้วยวิธี Least Median Square: LMS, วิธี Least Trimmed Square: LTS, วิธี M-Estimation: M-E และวิธี MME ผลการศึกษาพบว่า ค่าผิดปกติส่งผลต่อประสิทธิภาพการประมาณค่าสัมประสิทธิ์การถดถอย เมื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าชนิดแกร่งเหล่านั้น พบว่า ประสิทธิภาพของตัวประมาณด้วยวิธี LMS กับตัวประมาณด้วยวิธี LTS ใกล้เคียงกัน (ซึ่งหลาย ๆ สถานการณ์จำลองที่ศึกษานั้น ตัวประมาณด้วยวิธี LTS มีประสิทธิภาพสูงกว่าตัวประมาณค่าอื่น ๆ) นอกจากนี้ยังพบว่า ประสิทธิภาพของตัวประมาณด้วยวิธี M-E และวิธี MME จะมีค่าเพิ่มขึ้นตามขนาดตัวอย่างและร้อยละของจำนวนค่าผิดปกติ

Yan and Tian (2010) ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากร โดยใช้สัมประสิทธิ์ความเบ้ของตัวแปรช่วย มาปรับตัวประมาณค่าแบบอัตราส่วนของ Kadilar and Cingi ที่นำเสนอในปี ค.ศ. 2004 ได้ตัวประมาณค่าใหม่ 6 ตัวดังนี้

$$t_1 = \bar{y} [(\bar{x} + \beta_1) / (\bar{x} + \beta_1)] \quad (2.48)$$

$$t_2 = \bar{y} [(\beta_2 \bar{x} + \beta_1) / (\beta_2 \bar{x} + \beta_1)] \quad (2.49)$$

$$t_3 = \bar{y} [(\beta_1 \bar{x} + \beta_2) / (\beta_1 \bar{x} + \beta_2)] \quad (2.50)$$

$$t_4 = \bar{y} [(C_x \bar{x} + \beta_1) / (C_x \bar{x} + \beta_1)] \quad (2.51)$$

$$t_5 = [\bar{y} - b(\bar{x} - \bar{x})] [(\bar{x} + \beta_1) / (\bar{x} + \beta_1)] \quad (2.52)$$

$$t_6 = [\bar{y} - b(\bar{x} - \bar{x})] [(\beta_1 \bar{x} + \beta_2) / (\beta_1 \bar{x} + \beta_1)] \quad (2.53)$$

โดยที่ β_1 คือ ค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วย จากนั้นใช้อุกรมเทอร์เลอร์ติกรีอันดับหนึ่งหาสมการของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่าใหม่ และเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณค่าที่นำเสนอกับตัวประมาณค่าของ Kadilar and Cingi

ด้วยการคำนวณเชิงตัวเลข พบว่า ตัวประมาณค่าใหม่ตัวที่ห้าและตัวที่หก มีประสิทธิภาพใกล้เคียงกับตัวประมาณค่าเดิมของ Kadilar and Cingi

สรุปผลการวิจัยที่เกี่ยวข้องข้างต้น จะเห็นทิศทางการพัฒนาตัวประมาณค่าแบบอัตราส่วน อยู่สองกรณี กล่าวคือ กรณีที่หนึ่ง ใช้ตัวแปรช่วยหนึ่งตัว ดังการศึกษาของ Kadilar and Cingi (2004, 2006) และ Kadilar et al. (2007) เห็นได้ว่า นักวิจัยได้ใช้สารสนเทศของตัวแปรช่วย ไม่ว่าจะมีความโด่งหรือค่าสัมประสิทธิ์การแปรผันของตัวแปรช่วย มาปรับตัวประมาณค่าแบบอัตราส่วนเดิม และยังใช้หลักการวิเคราะห์การถดถอยเชิงเดียวและการวิเคราะห์การถดถอยแกลง มาปรับทอม \bar{y} ในตัวประมาณค่าแบบอัตราส่วนเดิม เพื่อให้ได้ตัวประมาณค่าที่มีประสิทธิภาพดีขึ้น ในระยะต่อมา เริ่มมีทิศทางใหม่ในการเลือกใช้สารสนเทศของตัวแปรช่วย นั่นคือ นักวิจัยเริ่มนำค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วย มาปรับตัวประมาณค่าที่มีผู้นำเสนอไว้ก่อนหน้า ดังการศึกษาของ Yan and Tian (2010) และการศึกษาของภทวรรณ แสงนวกิจ (2553) ที่ใช้ค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วยหนึ่งตัวร่วมกับหลักการวิเคราะห์การถดถอยแกลง มาพัฒนาตัวประมาณค่าที่ปรับใหม่ ทำให้มีประสิทธิภาพที่สูงกว่าตัวประมาณค่าเดิม ส่วนกรณีที่สอง จะใช้ตัวแปรช่วยสองตัว ในสถานการณ์ที่ตัวแปรศึกษา (Y) อาจมีความสัมพันธ์กับตัวแปรช่วยมากกว่า 1 ตัว และการใช้สารสนเทศจากตัวแปรช่วยทั้งหมด ย่อมเป็นการเพิ่มความแม่นยำในการพยากรณ์ค่ามากขึ้น จะเห็นได้จากการศึกษาของนักวิจัยหลายคนดังการศึกษาของ Abu-Dayyeh et al. (2003), Kadilar and Cingi (2005) และ Gupta and Shabbir (2007) ที่ได้พัฒนาตัวประมาณค่าแบบอัตราส่วนโดยใช้ตัวแปรช่วยสองตัว ซึ่งตัวประมาณค่าแบบอัตราส่วนใหม่ที่นักวิจัยนำเสนอ มีประสิทธิภาพสูงกว่าตัวประมาณค่าแบบอัตราส่วนเดิมที่ใช้ตัวแปรช่วยหนึ่งตัว ภายใต้เงื่อนไขที่เหมาะสม และนอกจากนี้บางผลการวิจัยยังเปรียบเทียบประสิทธิภาพของตัวประมาณค่าใหม่กับตัวประมาณค่าเดิม โดยไม่ได้คำนึงถึงการใช้จ่ายจำนวนตัวแปรช่วยที่ไม่เท่ากัน ดังการศึกษาของ Gupta and Shabbir (2007) และยังไม่ปรากฏว่ามีการนำเสนอตัวประมาณค่าแบบอัตราส่วน ที่เหมาะสมกับกรณีที่ข้อมูลประชากรมีค่าผิดปกติรวมอยู่ด้วย

หากนักวิจัยต้องการพัฒนาตัวประมาณค่าแบบอัตราส่วน กรณีที่ใช้ตัวแปรช่วยสองตัว ให้มีความเหมาะสมที่จะใช้ประมาณค่าเฉลี่ยประชากร เมื่อข้อมูลมีค่าผิดปกติรวมอยู่ หรือข้อมูลมีการแจกแจงที่มีความเบ้ สามารถใช้หลักการของการวิเคราะห์การถดถอยแกลง จากการศึกษาของ Yarmohammadi and Mahmoudvand (2010) ที่พบว่า ประสิทธิภาพของตัวประมาณค่าด้วยวิธี LTS ใกล้เคียงกับตัวประมาณค่าด้วยวิธี Least median Square (LMS) และหลาย ๆ สถานการณ์จำลองนั้น ตัวประมาณค่าด้วยวิธี LTS มีประสิทธิภาพสูงกว่าตัวประมาณค่าอื่น ๆ และจากการศึกษาของ Yan and Tian (2010) และการศึกษาของภทวรรณ แสงนวกิจ (2553) ที่ใช้ค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วย มาปรับตัวประมาณค่าแบบอัตราส่วน จากการศึกษาวรรณกรรมทั้งหมดข้างต้นนั้น ยังไม่ปรากฏการพัฒนาตัวประมาณค่าแบบอัตราส่วน เมื่อใช้ตัวแปรช่วยสองตัว โดยอาศัยหลักการของการวิเคราะห์การถดถอยแกลงร่วมกับการใช้ค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วย ดังนั้น การพัฒนาตัวประมาณค่าแบบอัตราส่วน กรณีที่ใช้ตัวแปรช่วยสองตัว และมีความเหมาะสมที่จะใช้ประมาณค่าเฉลี่ยประชากร เมื่อข้อมูลมีค่าผิดปกติ จึงเป็นแนวทางที่สามารถศึกษาต่อไปได้

ตอนที่ 2 การระเหยน้ำ ข้อมูลอุตุนิยมวิทยาและงานวิจัยที่เกี่ยวข้อง

1. การระเหยน้ำ

ในวัฏจักรน้ำ เริ่มจากพระอาทิตย์ส่องแสงในช่วงกลางวัน ความร้อนจากแสงอาทิตย์ทำให้พื้นดิน มหาสมุทร ทะเลสาบ หรือแหล่งน้ำอื่น ๆ ของโลกมีอุณหภูมิอากาศสูงขึ้น ซึ่งทำให้น้ำในแหล่งน้ำหรือน้ำในพื้นดินเกิดการระเหยกลายเป็นไอน้ำ กระบวนการที่น้ำในสถานะของเหลวเปลี่ยนสถานะโดยธรรมชาติเป็นแก๊สคือไอน้ำ (Water Vapor) เรียกว่า การระเหยน้ำ (Evaporation) โดยทั่วไปสามารถรับรู้ถึงการระเหยได้ โดยดูจากน้ำที่ค่อย ๆ หายไปที่ละน้อย เมื่อมันกลายเป็นไอน้ำ การที่โมเลกุลของของเหลวจะระเหยได้ จะต้องเป็นโมเลกุลที่อยู่บนพื้นผิว อยู่ในทิศทางที่เหมาะสม และมีพลังงานจลน์ภายในมากกว่าแรงยึดเหนี่ยวระหว่างโมเลกุล ที่จะเปลี่ยนสภาวะจากของเหลวเป็นแก๊ส หรือไอ วิชัย ศรีบุญลือ (2552, หน้า 213) กล่าวว่า การระเหยเป็นขั้นตอนที่สำคัญขั้นตอนหนึ่งในวัฏจักรน้ำ ในการออกแบบและจัดการเกี่ยวกับน้ำ การระเหยเป็นการสูญเสียน้ำจากอ่างที่สำคัญที่สุด การสูญเสียความชื้นของดินในพื้นที่การเกษตรย่อมเป็นผลจากการระเหย ดังนั้นการคำนวณค่าการระเหยจึงมีความสำคัญยิ่งในด้านอุทกวิทยา สาเหตุที่สำคัญในการระเหย คือ พลังงานที่ทำให้โมเลกุลของน้ำวิ่งรุนแรงขึ้น จนกระจายออกจากกันและเคลื่อนที่ขึ้นสู่บรรยากาศ ส่วนปัจจัยเสริมให้เกิดการระเหย คือ การแพร่ของไอน้ำออกจากชั้นที่ติดกับผิวน้ำที่กำลังเกิดการระเหย นั่นคือ นอกจากพลังงานแล้ว การแพร่ก็สำคัญมากเช่นเดียวกัน ทั้งนี้เพราะว่า ถ้าไม่มีการแพร่ น้ำเมื่อระเหยเป็นไอน้ำแล้ว ก็จะควบแน่น (Condensation) กลับเป็นน้ำอีกเช่นเดิม

2. การวัดการระเหยน้ำ

ในแต่ละวันน้ำในแหล่งน้ำต่าง ๆ บนโลก รวมทั้งน้ำในดิน เมื่อได้รับแสงแดด ก็เกิดการระเหยกลายเป็นไอน้ำในบรรยากาศ ซึ่งการระเหยน้ำนี้มีปริมาณแตกต่างกันไปตามพื้นที่และสภาพภูมิอากาศ ดังนั้น การวัดการระเหยน้ำให้ได้ค่าถูกต้องเป็นไปได้ยากมาก และไม่สามารถวัดได้โดยตรง ถ้าวันใดในเวลากลางวันท้องฟ้าโปร่งใส ปราศจากเมฆบดบัง แสงอาทิตย์ก็จะส่องผ่านมายังโลกได้มาก วันนั้นก็ทำให้มีการระเหยน้ำมากตามไปด้วย หรือถ้าวันใดในอากาศมีความชื้นมาก การระเหยน้ำก็จะน้อยตามไปด้วยเช่นกัน ดังนั้นองค์การอุตุนิยมวิทยาโลกจึงได้ให้สถานีอุตุนิยมวิทยาวัดค่าการระเหยน้ำ โดยให้ค่าของการระเหยจากผิวดินและการระเหยจากพืช (Transpiration) การวัดการระเหยจากผิวดินและผิวน้ำ มีความสำคัญมากสำหรับวิศวกรอุทก (Water Engineers) จำนวนน้ำที่ระเหยไปจากผิวดินโลกมากเท่าใด จำนวนน้ำฟ้า (Precipitation) ที่ตกลงมายังผิวโลกก็มากเท่านั้น วัตถุประสงค์ในการวัดอัตราการระเหยน้ำ ก็เพื่อที่จะได้ทราบรายละเอียดของกรรมวิธีเหล่านี้ รวมทั้งใช้เป็นข้อมูลในอ้างอิงสำหรับการคำนวณทางอุทกวิทยา และเพื่อประโยชน์ในทางเกษตรกรรมด้วย (ส่วนจัดสรรน้ำ สำนักบริหารจัดการน้ำ, 2553)

2.1 หน่วยที่ใช้ในการวัด

อัตราการระเหยจากผิวดิน วัดเป็นปริมาตรของน้ำ (ของเหลว) ซึ่งระเหยไปต่อหน่วยพื้นที่ต่อหน่วยเวลา ซึ่งมีค่าเท่ากับความลึกของน้ำที่หายไปจากพื้นที่ทั้งหมด ทั้งนี้ หน่วยของเวลาอาจใช้เป็นตัวหรือชั่วโมงก็ได้ หน่วยของความลึกอาจใช้เป็นตัวเมตร หรือ มิลลิเมตร หรือ นิ้วก็ได้

การวัดการระเหยน้ำนับว่ายุ่งยากกว่าการวัดปริมาณน้ำฝนหรือหยาดน้ำฟ้า และเป็นที่ยอมรับกันว่า เรายังไม่ได้รับค่าที่ถูกต้องแท้จริงของการระเหยน้ำจากผิวพื้นโลกเหนือบริเวณใด ๆ

2.2 หลักการทั่วไปของการวัดการระเหย

การวัดการระเหยโดยทั่วไปไม่สามารถวัดการระเหยน้ำได้ในธรรมชาติ จึงจำเป็นต้องวัดจากปริมาณน้ำที่ระเหยจากเครื่องมือ โดยการเทียบเคียงกับค่าระดับน้ำหรือน้ำหนักของน้ำที่หายไปในแต่ละวัน เครื่องมือที่ใช้วัดการระเหยน้ำแบ่งออกได้เป็น 4 ประเภท ดังนี้

2.2.1 แบบใช้ถังขนาดใหญ่ (Evaporation Tanks) เป็นถังเหล็กขนาดใหญ่ หรืออาจหล่อเป็นถังคอนกรีตก็ได้ ฝังลงไปบนดิน หรือถ้าในทะเลสาบจะใช้ถังลอย (Pontoon) ก็ได้ โดยกำหนดให้ระดับน้ำในถังนั้นใกล้เคียงกับระดับผิวพื้นโดยรอบ ต้องวัดความลึกของน้ำในถังตามช่วงระยะเวลาเป็นประจำ (ปกติวันละ 1 ครั้ง) และช่าง ๆ ถังต้องมีเครื่องวัดฝนติดตั้งไว้ด้วย เพื่อจะได้ทราบค่าของปริมาณฝนที่ตกลงมา ถ้าหากมีฝนตก แล้วจึงคิดคำนวณหาอัตราการระเหยน้ำได้

2.2.2 แบบใช้แผ่นผ้าแบนเล็ก (Small Evaporation Pans) มีทั้งแบบที่อ่านค่าของจำนวนน้ำที่ระเหยไป หรือการเปลี่ยนแปลงของระดับน้ำหรือน้ำหนักของน้ำที่หายไป

2.2.3 แบบที่ใช้พอร์ซเฟสเลน (Porous Porcelain Bodies) วิธีนี้ใช้วัดจำนวนของน้ำที่หายไป โดยพอร์ซเฟสเลนทรงกลม ทรงกระบอก หรือเป็นแผ่นซึ่งชุ่มน้ำอยู่ตลอดเวลา

2.2.4 แบบที่ใช้กระดาษชุ่มน้ำ (Instruments using Wet Paper Surfaces) วิธีนี้ใช้วัดน้ำที่หายไป โดยการใส่กระดาษกรอง (Filter Paper) ที่ชุ่มน้ำ ปล่อยให้ถูกกับอากาศในตู้สุกริน ปัจจุบันนิยมใช้แบบที่ 1 กันแพร่หลาย เพราะได้ผลใกล้เคียงธรรมชาติ ส่วนวิธีที่ 2 3 และ 4 ส่วนมากเลิกใช้กันแล้ว เพราะค่าที่ได้เพียงหยาบ ๆ เท่านั้น

2.3 เครื่องวัดน้ำระเหยแบบใช้ถังขนาดใหญ่ (Evaporation Tanks)

แบบใช้ถังมาตรฐานเป็นถังสี่เหลี่ยมจัตุรัส ขนาดกว้าง 6 ฟุต และลึก 2 ฟุต สร้างด้วยเหล็กแผ่นยัดให้ติดกัน รอบ ๆ ด้านนอก 3 นิ้ว นับจากขอบบนทำเป็นขอบรองรับขนาดกว้าง 1 นิ้ว ยึดติดอยู่ตั้งนั้นขอบบนของถัง จึงอยู่สูงจากพื้นดิน 3 นิ้ว การติดตั้งต้องคำนึงถึงกรณีต่าง ๆ ดังนี้

2.3.1 การกระเซ็นของน้ำออกจากถังเมื่อมีลมแรง

2.3.2 การกระเซ็นของน้ำฝนจากพื้นดินโดยรอบเข้ามาในถัง

2.3.3 การติดตั้งถังและเครื่องวัดฝนต้องเป็นแบบเดียวกัน เพื่อว่า เมื่อมีฝนตกจะได้คำนวณระดับน้ำที่เพิ่มขึ้นจากฝนได้

2.3.4 ปากถังต้องเปิดไว้ให้สัมผัสกับอากาศ เช่นเดียวกับการระเหยของสิ่งแวดล้อมโดยรอบ ถ้ามีสิ่งกีดขวางมากเกินไป จะไม่ได้รับรายงานของการระเหยน้ำที่เกิดขึ้นในบริเวณที่ไม่มีอะไรกีดขวาง เนื่องจากการเปลี่ยนแปลงของความแรงของลม

ในข้อ 1 ข้อ 3 และข้อ 4 เมื่อคิดหักชดเชยซึ่งกันและกันแล้ว นับว่าหมดไปพอดี ควรพิจารณาจัดให้มีเครื่องกีดขวางได้พอประมาณ เช่น ทำเป็นตาข่ายรูปสี่เหลี่ยม หรือตาข่ายไนลอนขนาดใหญ่คลุมปากถังไว้ เพื่อป้องกันสัตว์ที่จะเข้าไปกินน้ำในถัง และการติดตั้งเครื่องวัดฝน ก็ต้องพิจารณาให้เหมาะสมตามหลักการมาตรฐานการติดตั้งเครื่องวัดฝน ควรให้เครื่องวัดฝนอยู่ห่างจากถังน้ำในระยะ 3 – 10 ฟุต ในข้อ 2 การป้องกันไม่ให้น้ำกระเซ็นจากพื้นโดยรอบเข้าไปในถัง ทำได้โดยการปลูก

หญ້ารอบ ๆ ถัง และตัดให้สั้นอยู่เสมอ ดังนั้นขอบบนของถังจะอยู่สูงอย่างน้อย 2 นิ้วจากยอดหญ້า โดยที่ถังน้ำมีบริเวณกว้าง ย่อมทำให้น้ำเป็นปลิวหรือกระเพื่อมเมื่อมีลมพัด ทำให้ยากต่อการวัดระดับน้ำ จึงควรทำค้ำยกคอกสำหรับกั้น (Still-Water Pond) เป็นรูปคล้ายหีบสี่เหลี่ยม ด้านบนเปิด ขนาด 4x4 นิ้ว ลึก 12 นิ้ว กับมีรูรอบ ๆ ตรงกลางของฐาน บัดกรียึดติดอยู่กับตรงกลางของด้านใดด้านหนึ่งของถัง เมื่อติดเข้าที่แล้ว ขอบบนอยู่ระดับเดียวกับขอบของถังพอดี

การวัดระดับน้ำในถังใช้วัดด้วย “Index” Evaporation Gauge ซึ่งติดไว้ตรงมุมของถัง (หรือจะใช้ Hook gauge ติดไว้ที่ Still- Water Pond ตรงกลางของด้านใดด้านหนึ่งก็ได้) Index Gauge นี้ประกอบด้วยลูกกลอยทองแดง เลื่อนขึ้นลงได้ในทรงกระบอกทองแดง ทำหน้าที่เป็น Still-Water Pond โช้ทองเหลืองติดอยู่กับลูกกลอย ติดอยู่กับล้อหมุนจะเลื่อนไปตามอาการของโช้เข็มชี้หมุนไปตามขอบที่มีรัศมี 9 ½ นิ้ว สเกลแบ่งไว้จาก 0 - 4 นิ้ว และแบ่งละเอียดลงไปถึง 0.02 นิ้ว

เครื่องวัดการระเหยน้ำแบบนี้นิยมใช้ในสหราชอาณาจักรและประเทศออสเตรเลีย แต่เนื่องจากต้องเสียค่าใช้จ่ายมาก จึงไม่นิยมแพร่หลายนักในประเทศอื่น ๆ อย่างไรก็ดี ผลที่ได้รับจากการวัดการระเหยน้ำด้วยวิธีนี้นับว่าใกล้เคียงธรรมชาติมากที่สุด สำหรับประเทศไทยใช้เครื่องวัดการระเหยน้ำแบบภาคใหญ่ของอเมริกา หรือที่กรมอุตุนิยมวิทยาเรียกว่า “เครื่องวัดการระเหยแบบภาค” (Evaporation Pan) ชนิด American Class A Pan ดังนี้

2.4 เครื่องวัดการระเหยแบบภาค (Evaporation Pan) ชนิด American Class

A Pan

เครื่องแบบนี้นิยมใช้กันแพร่หลาย ประกอบด้วย

2.4.1 ภาคน้ำ (Pan) เป็นภาครูปทรงกลม ขนาดลึก 10 นิ้ว และเส้นผ่าศูนย์กลางปากถัง 48 นิ้ว ทำด้วยเหล็กเคลือบสังกะสีหรือโลหะผสมอย่างเบา (Monel Metal) ถ้าในบางพื้นที่ซึ่งน้ำมีสารที่ทำให้เกิดสนิมมาก จึงนิยมใช้โลหะผสมอย่างเบามากกว่า ตัวภาคน้ำตั้งอยู่บนฐานไม้สำหรับรองรับสูงจากพื้นดิน 6 นิ้ว ควรถมดินยกพื้นระดับให้สูงจากระดับเดิมเล็กน้อย ต้องหมั่นตรวจดูภาคน้ำว่ามีรูรั่วหรือมีรอยสนิม อย่างน้อยเดือนละครั้ง ถ้ามีสนิมให้ขัดออกด้วยแปรงทองเหลืองแล้วทาสีทับภาคน้ำต้องสะอาดอยู่เสมอ และน้ำต้องให้สะอาด อย่าให้มีผงตะกอนและฝ้าน้ำมันจับ เพราะฝ้าน้ำมันทำให้ลดอัตราการระเหยลง

2.4.2 ขอวัดระดับน้ำ (Micrometer Hook Gauge) ทำปลายเป็นรูปขอกคล้ายเบ็ด มีสเกลแบ่งไว้เป็นนิ้ว จาก 0 - 4 นิ้ว แบ่งทุก ๆ 0.1 นิ้ว ตรงฐานแบ่งละเอียดลงไปถึง 0.01 นิ้ว มีขา 3 ขา สำหรับวางบนปากที่รองรับ ขอวัดนี้บางครั้งก็แบ่งสเกลเป็นมิลลิเมตร ตั้งแต่ 0 - 10 เซนติเมตร แบ่งทุก ๆ 1 มิลลิเมตร และที่ฐานแบ่งละเอียดลงไปให้อ่านได้ถึง 0.1 มิลลิเมตร

2.4.3 ที่รองรับขอวัดระดับน้ำ (Stilling Well) เป็นรูปทรงกระบอก ใช้เพื่อป้องกันอาการปลิว หรือกระเพื่อมของน้ำในภาคน้ำ และเป็นที่ยางขอวัดด้วย ทำเป็นทรงกระบอกทองเหลืองสูง 8 นิ้ว ขนาดเส้นผ่าศูนย์กลาง 3 ½ นิ้ว ตั้งอยู่บนฐานรูปสามเหลี่ยม มีสกรูยึด และมีเกลียวสำหรับแต่งระดับ 3 ตัว การติดตั้งให้ตั้งไว้ในภาคน้ำ ห่างจากด้านเหนือของขอบภาคน้ำ 1 ฟุต แต่งระดับปากที่รองรับขอวัดด้วยเกลียว 3 ตัว ให้ได้ระดับตามอนจริง ๆ

2.4.4 เครื่องวัดความเร็วลม (Anemometer) ใช้แบบวัดความเร็วรวม (Totalize Wind Velocity) ติดตั้งไว้บนฐานไม้รองรับภาคน้ำระเหย ให้ตัวถ้วยอยู่สูงจากปากถังเล็กน้อย

2.4.5 เทอร์โมมิเตอร์ลอยน้ำ (Floating Thermometer) เป็นเทอร์โมมิเตอร์รูปตัวยู (U) ข้างหนึ่งเป็นเทอร์โมมิเตอร์สูงสุด อีกข้างหนึ่งเป็นเทอร์โมมิเตอร์ต่ำสุด ติดอยู่กับฟลุตลอยน้ำ เมื่อเอาเทอร์โมมิเตอร์นี้ลอยอยู่ในภาต ตัวเทอร์โมมิเตอร์ต้องอยู่ต่ำกว่าระดับผิวน้ำประมาณ 1 นิ้ว การอ่านเมื่อได้ค่าของอุณหภูมิอากาศสูงสุดและต่ำสุดแล้ว เอามาหาค่าเฉลี่ย จะได้ค่าเฉลี่ยของอุณหภูมิอากาศของน้ำในภาต

2.4.6 ถังเก็บน้ำ (Water-Storage Tank) ใช้เฉพาะในที่กั้นตารน้ำและห่างไกลจากการคมนาคม และให้มีความจุประมาณ 30 แกลลอน แต่ที่นิยมควรเป็นถังน้ำสี่เหลี่ยม 400 แกลลอน เครื่องวัดการระเหยแบบภาต (Evaporation Pan) ชนิด American Class A Pan หรือบางครั้งเรียกว่า ภาตวัดการระเหยน้ำแบบอเมริกัน มีลักษณะดังภาพที่ 2-5



ภาพที่ 2-5 เครื่องวัดการระเหยแบบภาต (Evaporation Pan) ชนิด American Class A Pan

2.5 วิธีการวัดระดับน้ำในภาตระเหย

ขั้นแรก ให้เลื่อนปลายขอวัดลงไปใต้ระดับน้ำเล็กน้อย แล้วจึงค่อย ๆ เลื่อนขึ้น จนปลายขอวัดแตะระดับน้ำพอดี ขณะทำการวัดระดับน้ำนี้ แสงสะท้อนจากท้องฟ้าจะช่วยให้แต่งระดับดียิ่งขึ้น เสร็จแล้วก็ยกขอวัดออกจากที่รองรับ แล้วจึงอ่านค่าที่สเกล การอ่านต้องให้ละเอียดให้ได้ค่าใกล้เคียงทศนิยม 2 ตำแหน่งของหลักมิลลิเมตร หรือเศษร้อยของนิ้ว การวัดค่าการระเหยน้ำ โดยปกติใช้วัดในช่วง 24 ชั่วโมง คือ ทำการตรวจวัดในเวลา 07.00 น. ทุกวัน การใช้ขอวัดให้ปฏิบัติดังนี้

2.5.1 เมื่อเติมน้ำให้ระดับสูงประมาณ 2 นิ้ว ใต้ปากถัง แล้วตั้งปลายขอวัด ให้ขอวัดแตะกับระดับน้ำพอดี จดค่าสเกลไว้ (พยายามให้ใกล้เคียงทศนิยม 2 ตำแหน่งของหลักมิลลิเมตร)

2.5.2 ภายหลัง 24 ชั่วโมง (คือเวลา 07.00 น. ของวันรุ่งขึ้น) น้ำระเหยไป ระดับน้ำจะลดลง ให้เลื่อนขอวัดลงไป จนปลายขอวัดแตะกับระดับน้ำใหม่ ผลต่างที่ได้จากการอ่านครั้งแรก และครั้งหลัง คือ ค่าของการระเหยน้ำ การวัดดังกล่าวนี้เป็นารวัดการระเหยน้ำตามปกติ เมื่อไม่มีฝนตก

2.5.3 ในกรณีที่มีฝนตก ค่าของการระเหยน้ำ คือ (จำนวนน้ำฝน + ระดับน้ำเดิม) – ระดับน้ำใหม่

3. ข้อมูลอุตุนิยมวิทยาที่เกี่ยวข้อง

สำหรับการประมาณค่าการระเหยน้ำรายเดือนในจังหวัดสกลนคร โดยใช้ตัวแปรช่วยสองตัวนี้ จะพิจารณาคัดเลือกตัวแปรช่วยจากข้อมูลอุตุนิยมวิทยาที่เกี่ยวข้อง ซึ่งเป็นข้อมูลที่ได้จากการตรวจอากาศผิวพื้นหรือการตรวจจาดูประกอบอุตุนิยมวิทยาต่าง ๆ ในระดับผิวพื้นที่สูงขึ้นไปจากพื้นดินไม่เกิน 10 เมตร ณ สถานีอุตุนิยมวิทยาสกลนคร ในแต่ละวันสถานีอุตุนิยมวิทยาสกลนคร ทำการตรวจอากาศผิวพื้น 8 ช่วงเวลาหรือทุก ๆ 3 ชั่วโมง ได้แก่ เวลา 01.00 น. 04.00 น. 07.00 น. 10.00 น. 13.00 น. 16.00 น. 19.00 น. และ 22.00 น. จากนั้นข้อมูลอุตุนิยมวิทยาที่วัดได้จาก 8 ช่วงเวลา จะนำมาเฉลี่ยเป็นข้อมูลอุตุนิยมวิทยาเฉลี่ยรายวัน และทุก ๆ เดือน จะนำข้อมูลอุตุนิยมวิทยาเฉลี่ยรายวันเหล่านั้น มาเฉลี่ยเป็นข้อมูลอุตุนิยมวิทยารายเดือนต่อไป

การวิจัยนี้ใช้ข้อมูลอุตุนิยมวิทยาที่เกี่ยวข้องกับการระเหยน้ำ เป็นข้อมูลอุตุนิยมวิทยาโดยเฉลี่ยรายเดือน ประกอบด้วย 1) อุณหภูมิอากาศเฉลี่ย 2) ความกดอากาศเฉลี่ย 3) ปริมาณฝนโดยเฉลี่ย 4) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย และ 5) ความชื้นสัมพัทธ์ จากเดือนมกราคม พ.ศ. 2550 ถึงเดือนเมษายน พ.ศ. 2556 จำนวน 76 เดือน เก็บบันทึกข้อมูลโดยสถานีอุตุนิยมวิทยา สกลนคร จังหวัดสกลนคร ตั้งอยู่สูงจากระดับน้ำทะเลปานกลาง 171 เมตร ซึ่งสถานีอุตุนิยมวิทยา สกลนคร (2556) ได้อธิบายการเก็บบันทึกข้อมูลอุตุนิยมวิทยาข้างต้นไว้ดังนี้

3.1 อุณหภูมิอากาศเฉลี่ย

การตรวจวัดอุณหภูมิอากาศ จะอ่านจากเครื่องเทอร์โมมิเตอร์ที่ติดตั้งอยู่ภายในเรือนเทอร์โมมิเตอร์ (Thermometer Screen) ในสนามอุตุนิยมวิทยา โดยปกติเครื่องเทอร์โมมิเตอร์นี้จะถูกติดตั้งอยู่ในสถานที่ที่มีอากาศถ่ายเทตามธรรมชาติ และอยู่ในร่มเงาไม่ถูกรังสีจากดวงอาทิตย์ โดยตรง ทางอุตุนิยมวิทยาจะเรียกค่าอุณหภูมิอากาศที่อ่านได้จากเครื่องเทอร์โมมิเตอร์นี้ว่า อุณหภูมิอากาศตุ้มแห้ง (Dry-Bulb Temperature) มีหน่วยวัดเป็นองศาเซลเซียส ($^{\circ}\text{C}$) ในแต่ละวันเจ้าพนักงานอุตุนิยมวิทยาจะวัดอุณหภูมิอากาศ 8 ช่วงเวลา ได้แก่ เวลา 01.00 น. 04.00 น. 07.00 น. 10.00 น. 13.00 น. 16.00 น. 19.00 น. และ 22.00 น. แล้วนำค่าอุณหภูมิอากาศทั้ง 8 ช่วงเวลามาเฉลี่ยเป็นอุณหภูมิอากาศเฉลี่ยของวันนั้น ๆ จากนั้นเมื่อสิ้นเดือนใด ๆ อุณหภูมิอากาศเฉลี่ยรายวันจะถูกนำมารวมกัน แล้วหารด้วยจำนวนวันในเดือนนั้น ๆ เป็นอุณหภูมิอากาศเฉลี่ยรายเดือน

3.2 ความกดอากาศเฉลี่ย

ความกดอากาศ หมายถึง แรงที่เกิดจากน้ำหนักของมวลอากาศกดลงมาบนพื้นดินหนึ่งหน่วยพื้นที่ โดยทั่วไป บริเวณใกล้พื้นโลกจะมีมวลอากาศหนาแน่นกว่าบริเวณที่สูงขึ้นไปจากพื้นโลก ดังนั้น บริเวณใกล้พื้นโลกจะมีความกดอากาศสูงกว่าบริเวณที่สูงขึ้นไปจากพื้นโลก เครื่องวัดความกดอากาศ คือ เครื่องบาริโอมิเตอร์ปรอท ภายในบรรจุปรอทซึ่งมีคุณสมบัติเหมาะสมดีกว่าสารชนิดอื่น เครื่องบาริโอมิเตอร์ปรอทต้องติดตั้งภายในอาคาร หรือภายในห้องที่มีสภาพอากาศปกติ กล่าวคือ ไม่ถูกรบกวนจากแสงแดดความร้อนหรือความเย็นจากเครื่องปรับอากาศ หรือจากความร้อนจากแหล่งอื่นที่ไม่ได้เกิดจากธรรมชาติ ต้องติดตั้งอยู่ในที่ที่มั่นคง แข็งแรง ไม่ถูกรบกวนจากการสั่นสะเทือน อยู่สูงจาก

ระดับน้ำทะเลปานกลางได้ถูกต้อง ความกดอากาศมีหน่วยการวัดเป็น มิลลิเมตรปรอท (mm.Hg) หรือ มิลลิบาร์ (mbs) แต่ในปัจจุบันนักอุตุนิยมวิทยานิยมใช้หน่วยวัดความกดอากาศ เป็นหน่วยวัด เฮกโตพาสคาล (Hecto Pascal เขียนย่อว่า hPa) ซึ่ง 1 hPa จะเท่ากับ 1 มิลลิบาร์ เนื่องจากในแต่ละวันจะมีการอ่านค่าความกดอากาศจากเครื่องบาริโอมิเตอร์ปรอททุก ๆ 3 ชั่วโมง ค่าความกดอากาศที่อ่านได้ในแต่ละครั้งของแต่ละสถานที่ที่ใช้ เพื่อการวิเคราะห์สภาพลมฟ้าอากาศของภูมิภาคหรือทวีปใด ๆ จำเป็นต้องปรับให้เป็นมาตรฐานเดียวกันดังนี้

3.2.1 อัตราผิดของเครื่อง (Index Error) ค่าความผิดประจำเครื่องแต่ละเครื่องได้จากการสอบเทียบเครื่องนั้น ๆ กับเครื่องมาตรฐาน ซึ่งแสดงไว้ในใบรับรองการสอบเทียบประจำเครื่องบาริโอมิเตอร์ เครื่องบาริโอมิเตอร์ที่ดีต้องมีอัตราผิดประจำเครื่องไม่เกิน ± 0.3 มิลลิเมตรปรอท

3.2.2 การปรับแก้อุณหภูมิอากาศ (Temperature Correction) บาริโอมิเตอร์ทุกเครื่องจะต้องมีเทอร์โมมิเตอร์ประจำเครื่อง เพื่อให้ทราบค่าอุณหภูมิอากาศของเครื่องขณะอ่านบาริโอมิเตอร์ อัตราผิดเนื่องจากอุณหภูมิอากาศเกิดจากการขยายตัวของปรอทและสเกลบนเครื่อง ดังนั้นจึงต้องปรับแก้อัตราผิดเข้าหาอุณหภูมิอากาศมาตรฐานที่ 0°C

3.2.3 ความถ่วงตามละติจูดของที่ตั้งของสถานี (Gravity Correction) ค่าความกดอากาศที่อ่านได้จากบาริโอมิเตอร์ที่ความกดและอุณหภูมิอากาศใด ๆ ขึ้นอยู่กับค่าของอัตราเร่งความถ่วง ซึ่งเปลี่ยนแปลงตามละติจูดและความสูงของพื้นที่นั้น ๆ ฉะนั้น ค่าความกดอากาศที่อ่านได้บนพื้นที่ใด ๆ ต้องปรับแก้เข้าสู่ภาวะมาตรฐานที่มีอัตราเร่งของความถ่วง 9.80665 เมตร/วินาที²

3.2.4 การปรับแก้ระดับความสูงของสถานี (Altitude Correction) สถานีที่ตั้งของสถานีตรวจอากาศต่าง ๆ ตั้งอยู่บนพื้นที่ที่มีความสูงแตกต่างกัน เพื่อให้ค่าความกดอากาศที่อ่านได้จากแต่ละสถานีมีมาตรฐานเดียวกันนั้น ต้องปรับแก้ให้เป็นความกดอากาศที่ระดับความสูงจากระดับน้ำทะเลปานกลาง โดยพื้นที่ที่อยู่ใกล้ระดับน้ำทะเลปานกลางจะมีค่าความกดอากาศลดลงหนึ่งมิลลิบาร์เมื่อความสูงเพิ่มขึ้น 8.5 เมตร แต่ที่ระดับความสูง 5.5 กิโลเมตร ความกดอากาศลดลงหนึ่งมิลลิบาร์ ทุกช่วงความสูงที่เพิ่มขึ้น 15 เมตร (ริงสรรค์ อาภาศัพท์กุล, 2547, หน้า 61-62) และความกดอากาศปกติที่ระดับน้ำทะเลปานกลาง จะมีค่าเท่ากับ 760 mm.Hg หรือเท่ากับ 1,013.2 mbs สำหรับความกดอากาศเฉลี่ยในที่นี้ จะเป็นการนำค่าความกดอากาศที่อ่าน ณ เวลา 8 ช่วงเวลาในแต่ละวัน มาเฉลี่ยกันเป็นความกดอากาศเฉลี่ย มีหน่วยเป็นมิลลิบาร์

3.3 ปริมาณฝนโดยเฉลี่ย

ปริมาณฝนที่วัดในแต่ละวัน จะวัดด้วยเครื่องวัดฝนที่มีลักษณะเป็นรูปทรงกระบอก 2 ชั้นวางซ้อนกัน และเครื่องวัดฝนต้องติดตั้งในที่โล่งแจ้ง ระยะห่างของเครื่องวัดต้องอยู่ห่างจากสิ่งกีดขวางแวดล้อม อย่างน้อยสองเท่าของความสูงของสิ่งนั้น ถังวัดฝนส่วนใหญ่มักนิยมปากถังที่มีขนาดเส้นผ่าศูนย์กลาง 8 นิ้ว สูงจากพื้นดิน 1 เมตร มีกรวยรับน้ำฝน เพื่อให้ไหลรวมลงสู่ถังเก็บด้านใน เพื่อป้องกันการระเหยน้ำฝน การวัดน้ำฝนแต่น้ำน้ำฝนที่อยู่ในถังรองด้านในเทลงใส่กระบอกแก้วดวงวัดฝน แล้วอ่านสเกลที่อยู่ด้านข้างกระบอกแก้ว มีหน่วยเป็น มิลลิเมตร (มม.) หรือ นิ้ว (25.4 มม. เท่ากับ 1 นิ้ว) ฝนที่วัดได้เป็นความสูง (ความหนา) ของน้ำฝนจากพื้นดิน โดยอนุโลมว่าน้ำฝนนั้น ไม่มีการระเหยหรือไหลซึมลงสู่ดิน ทั้งนี้การวัดปริมาณน้ำฝนในแต่ละวัน จะรวมน้ำฝนตั้งแต่วันที่ 07.00

น. ของวันปัจจุบัน จนถึงเวลา 07.00 น. ของวันรุ่งขึ้น สำหรับปริมาณฝนโดยเฉลี่ยรายเดือน จะคำนวณจากค่าเฉลี่ยของปริมาณฝนรายวันในเดือนนั้น ๆ

3.4 ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย

การวัดระยะเวลาที่มีแสงแดดในแต่ละวัน จะใช้เครื่องวัดแสงแดดแบบลูกแก้ว ซึ่งลูกแก้วทรงกลมจะรวมแสงแดดให้เป็นจุดโฟกัส จนเกิดความร้อน เผาไหม้ที่แผ่นกระดาษกราฟที่มีคุณสมบัติไวต่อความร้อนและทนน้ำ โดยที่กระดาษกราฟมีสเกลบอกระยะเวลาเป็นชั่วโมง ดังนั้น การติดตั้งเครื่องวัดแสงแดดแบบลูกแก้ว ต้องติดตั้งให้แกนกลางของเครื่องชี้ไปทางทิศเหนือ เพื่อให้กระดาษกราฟที่รองรับจุดโฟกัสของแสงแดดอยู่ในแนวตะวันออก - ตะวันตก ตามเส้นทางโคจรของดวงอาทิตย์ ระยะเวลาที่มีแสงแดดในแต่ละวันมีหน่วยที่วัดเป็นชั่วโมง และระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ยของแต่ละเดือน จะนำระยะเวลาที่มีแสงแดดในแต่ละวันในเดือนนั้นมาเฉลี่ยกัน

3.5 ความชื้นสัมพัทธ์

ความชื้นสัมพัทธ์ (Relative Humidity) หมายถึง อัตราส่วนระหว่างความดันไอน้ำจริงต่อความดันไอน้ำอิ่มตัวที่อุณหภูมิอากาศเดียวกัน มีหน่วยเป็นร้อยละ ไฮโครมิเตอร์แบบกระเปาะเปียกและแห้ง (Wet and Dry Bulb Psychrometer) เป็นเครื่องมือสำหรับวัดความชื้นอากาศ ประกอบด้วยเทอร์โมมิเตอร์ชนิดเดียวกันสองอัน ติดตั้งไว้เคียงกัน อันหนึ่งสำหรับวัดอุณหภูมิอากาศของอากาศธรรมดา เทอร์โมมิเตอร์อันนี้เรียกว่า “เทอร์โมมิเตอร์ตุ้มแห้ง” (Dry Bulb Thermometer) เทอร์โมมิเตอร์อีกอันหนึ่งถูกหุ้มด้วยผ้าฝ้ายที่เปียกน้ำ หุ้มตุ้มปรอท เทอร์โมมิเตอร์นี้เรียกว่า “เทอร์โมมิเตอร์ตุ้มเปียก” (Wet Bulb Thermometer) ถ้าอากาศร้อนและมีความชื้นต่ำ จะทำให้พื้นที่กระเปาะเปียกกระเหยมากขึ้น เทอร์โมมิเตอร์กระเปาะเปียกจะแสดงค่าอุณหภูมิอากาศต่ำลง และมีค่าแตกต่างจากอุณหภูมิอากาศของเทอร์โมมิเตอร์กระเปาะแห้งมากขึ้น การหาค่าความชื้นสัมพัทธ์ จะนำค่าอุณหภูมิอากาศตุ้มแห้งและอุณหภูมิอากาศตุ้มเปียก ไปหาค่าผลต่างแล้ว นำค่าผลต่างไปเปิดตารางสำเร็จเพื่อหาค่าความชื้นสัมพัทธ์ (รังสรรค์ อากาศศึกษา, 2547, หน้า 134-136)

4. งานวิจัยที่เกี่ยวกับการระเหยน้ำหรือการคายระเหยน้ำ

พิศ คงบริรักษ์ และกฤติกา สืบศักดิ์ (2554) ได้ศึกษาเกี่ยวกับการระเหยน้ำในประเทศไทย คาบ 10 ปีในช่วง พ.ศ 2544-2553 เพื่อวิเคราะห์การระเหยน้ำในแต่ละพื้นที่ โดยรวบรวมข้อมูลการระเหยน้ำ (Evaporation) จากสถานีของกรมอุตุนิยมวิทยาในช่วง 10 ปีที่ผ่านมา ตั้งแต่ พ.ศ. 2544-2553 และใช้การนำข้อมูลเข้าโปรแกรม Surfer ซึ่งได้ข้อมูลเชิงพื้นที่ในลักษณะของจุด (Point) แล้ว Interpolate ข้อมูลด้วยวิธี Kriging Interpolate นำเสนอในรูปแบบแผนที่เชิงตัวเลขแสดงค่าปริมาณน้ำระเหยเฉลี่ยในคาบ 10 ปี ผลการศึกษาพบว่า การระเหยน้ำมีความแตกต่างกันออกไปในแต่ละพื้นที่ ซึ่งภาคตะวันออกเฉียงเหนือมีการระเหยสูงสุดในเดือนเมษายนและลดลงเป็นลำดับในเดือนมีนาคมและพฤษภาคม และในภาพรวมทั้งประเทศจะมีการระเหยน้ำสูงอยู่ในช่วงเดือนมีนาคมเมษายน และพฤษภาคม โดยเดือนเมษายนมีปริมาณการระเหยน้ำสูงที่สุดในรอบปี

Irmak and Haman (2003) ได้ประเมินวิธีการประมาณค่าการระเหยน้ำแบบภาคด้วยตัวแปรทางอุตุนิยมวิทยาในรัฐฟลอริดา ข้อมูลที่ใช้ คือ ข้อมูลทางอุตุนิยมวิทยารายวัน จำนวน 8,395 วัน ในช่วงเวลา 23 ปี ตั้งแต่วันที่ 1 มกราคม ค.ศ. 1978 ถึงวันที่ 30 กันยายน ค.ศ. 2000 เก็บบันทึกโดย

สถานีอุตุนิยมวิทยา GAARC ข้อมูลอุตุนิยมวิทยารายวันที่ใช้ประกอบด้วย ปริมาณน้ำฝน อุณหภูมิ อากาศสูงสุด อุณหภูมิอากาศต่ำสุด ความชื้นสัมพัทธ์ ความเร็วลมและทิศทาง รังสีดวงอาทิตย์ และการระเหยน้ำแบบลาด ตามลำดับ ทำการศึกษาเปรียบเทียบประสิทธิภาพของการประมาณค่า การระเหยน้ำ 5 วิธี ได้แก่ 1) วิธี Penman 2) วิธี Kohloer-Nordenson-Fox (KNF) 3) วิธี Christiansen 4) วิธี Priestley-Taylor (PT) และ 5) วิธี Linacre ใช้เกณฑ์รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: *RMSE*) เป็นเกณฑ์เปรียบเทียบความแม่นยำ (Accuracy) และความน่าเชื่อถือได้ (Reliability) ของวิธีการประมาณค่าทั้ง 5 วิธี โดยเกณฑ์ความคลาดเคลื่อนที่ยอมรับได้ (Acceptable Error) คือ รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยจะต้องมีค่าไม่เกิน 0.5 ($RMSE \leq 0.5$) จากนั้นนำเสนอค่าประมาณการระเหยน้ำ ทั้งในเทอมค่าเฉลี่ยรายวัน ค่าเฉลี่ยรายเดือน และค่าเฉลี่ยรายปี สำหรับประสิทธิภาพ (Performance) ของค่าประมาณการระเหยน้ำรายเดือน จะอธิบายด้วยค่าร้อยละของความคลาดเคลื่อน (Percentage Error: %E) ซึ่งกำหนดให้ค่าการระเหยน้ำเฉลี่ยรายเดือนที่แท้จริงเป็นค่าเฉลี่ยมาจากค่าการระเหยน้ำที่วัดได้ (Measured Vales) ในเดือนนั้น ๆ ในช่วงเวลา 23 ปี ผลการศึกษาพบว่า วิธี KNF มีความแม่นยำในการประมาณค่าการระเหยน้ำแบบลาด มากที่สุด รองลงมาคือ วิธี Penman, วิธี Priestley-Taylor, วิธี Christiansen และวิธี Linacre ตามลำดับ เมื่อประมาณค่าการระเหยน้ำแบบลาดด้วยวิธี KNF พบว่า เดือนพฤษภาคมจะมีการระเหยน้ำเฉลี่ยรายเดือนสูงสุด สำหรับผลการประมาณค่าการระเหยน้ำเฉลี่ยรายเดือนด้วยวิธีทั้ง 5 วิธีนั้น พบว่า วิธี KNF, วิธี Penman, วิธี Priestley-Taylor, วิธี Christiansen และวิธี Linacre มีค่าร้อยละของความคลาดเคลื่อน (%E) เท่ากับ 6%, 17%, -24%, -27% และ 32% ตามลำดับ (เมื่อเรียงตามขนาดของ %E) นอกจากนี้เมื่อคัดเลือกข้อมูลอุตุนิยมวิทยาในปี ค.ศ. 1981 ที่มีลักษณะอากาศแบบแห้งหรือความชื้นต่ำ (Dry Year) และปี ค.ศ. 1983 ที่มีลักษณะอากาศแบบเปียกหรือความชื้นสูง (Wet Year) มาทดสอบการประมาณค่าการระเหยน้ำเฉลี่ยรายเดือนด้วยวิธี KNF พบว่า ค่าประมาณการระเหยน้ำเฉลี่ยรายเดือนในปี ค.ศ. 1981 และปี ค.ศ. 1983 มีค่า %E เท่ากับ 9% และ 5% ตามลำดับ ซึ่งในปี ค.ศ. 1981 จะได้ค่าประมาณการระเหยน้ำเฉลี่ยรายเดือนต่ำกว่าค่าจริง (Underestimation) เป็นอย่างมาก โดยเฉพาะเดือนในช่วงฤดูร้อน อาจจะเนื่องจากการนำพาความร้อนสัมผัส (Sensible Heat Advection) ไม่สามารถอธิบายในวิธีของ KNF จึงทำให้เกิดการประมาณค่าที่ต่ำกว่าความเป็นจริง

Kisi (2006) ได้ศึกษาเกี่ยวกับการประมาณตัวแบบการระเหยน้ำแบบลาดรายวันด้วยวิธีการรวมวิธีโครงข่ายประสาทเทียม (Neural Network) และการใช้งานตรรกศาสตร์คลุมเครือ (Fuzzy) หรือที่เรียกว่า A Neuro-Fuzzy Computing Technique ข้อมูลทางอุตุนิยมวิทยารายวันที่ใช้ประกอบด้วย อุณหภูมิอากาศ รังสีดวงอาทิตย์ ความเร็วลม ความกดอากาศ ความชื้นสัมพัทธ์ และการระเหยน้ำ ตามลำดับ ซึ่งเก็บบันทึกโดยสถานีอุตุนิยมวิทยา Arcata-Eureka และสถานีอุตุนิยมวิทยา Daggett จากปี 1987 — 1990 ใช้ข้อมูลสามปีแรกในการฝึกสอนเพื่อสร้างตัวแบบ และใช้ข้อมูลที่เหลือในการทดสอบตัวแบบพยากรณ์การระเหยน้ำ ทำการเปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์การระเหยน้ำระหว่างวิธีโครงข่ายประสาทเทียม วิธี Neuro-Fuzzy และวิธีของ Stephens-Stewart ผลการศึกษาพบว่า การพยากรณ์การระเหยน้ำทั้งในสถานี Arcata-Eureka

และสถานี Daggett ด้วยวิธีโครงข่ายประสาทเทียมและวิธี Neuro-Fuzzy ได้ความแม่นยำมากกว่าวิธีของ Stephens-Stewart

Moshrik et al. (2008) ได้ศึกษาตัวแบบอนุกรมเวลาเพื่อพยากรณ์การคายระเหยน้ำรายเดือนของพืชอ้างอิงในแ่งประเทศจอร์แดน โดยใช้ข้อมูลการระเหยน้ำที่วัดด้วยถาดวัด โดยใช้ข้อมูลจากปี 1973-2002 ทำการสร้างตัวแบบการพยากรณ์ด้วยตัวแบบ Seasonal Autoregressive and Moving Average (ARIMA) และเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์ที่ได้กับค่าการคายระเหยน้ำของพืชอ้างอิงที่คำนวณจากข้อมูลการระเหยน้ำ ด้วยรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: *RMSE*) ค่าร้อยละของค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error) และค่าร้อยละของค่าสูงสุดความคลาดเคลื่อนสัมบูรณ์ (Maximun Absolute Percentage Error) ผลการสร้างตัวแบบสามารถกำหนดตัวแบบที่เหมาะสมได้สามตัวแบบ แต่ตัวแบบที่เหมาะสมในการพยากรณ์การคายระเหยน้ำของพืชอ้างอิงมากที่สุด คือ ตัวแบบ ARIMA(1,0,0)(0,1,1)₁₂

Goel (2009) ได้ศึกษาประเมินประสิทธิภาพตัวแบบจากวิธีโครงข่ายประสาทเทียมเพื่อพยากรณ์การระเหยน้ำในอ่างเก็บน้ำ ซึ่งมีวัตถุประสงค์เปรียบเทียบตัวแบบวิธีโครงข่ายประสาทเทียมกับวิธีการวิเคราะห์การถดถอยพหุคูณ เพื่อการพยากรณ์การระเหยน้ำของอ่างเก็บน้ำในบริเวณ Anand Sagar เมือง Shegaon ประเทศอินเดีย โดยใช้ข้อมูลทางอุตุนิยมวิทยาเฉลี่ยในรายสัปดาห์ของปี 1990 ถึงปี 2004 จากสถานีอุตุนิยมวิทยาเมือง Manasgaon ข้อมูลนำเข้าหรือตัวแปรอิสระประกอบด้วยอุณหภูมิอากาศเฉลี่ย ระยะเวลาที่มีแสงแดด ความชื้นสัมพัทธ์เฉลี่ย และความเร็วลม ทำการพยากรณ์การระเหยน้ำในอ่างเก็บน้ำด้วยวิธีโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (Multi-Layer Perceptron Neural Network) และวิธีการวิเคราะห์การถดถอยพหุคูณห้ากรณี โดยกรณีหนึ่งถึงกรณีที่สี่ ใช้ข้อมูลนำเข้าเพียงหนึ่งตัว ได้แก่ อุณหภูมิอากาศเฉลี่ย ความเร็วลม ระยะเวลาที่มีแสงแดด และความชื้นสัมพัทธ์เฉลี่ย ตามลำดับ ส่วนกรณีห้าจะใช้ทั้งสี่ตัวแปรเป็นข้อมูลนำเข้าทั้งหมด เปรียบเทียบประสิทธิภาพการพยากรณ์ด้วยค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ผลการศึกษาการพยากรณ์การระเหยน้ำในอ่างเก็บน้ำกรณีหนึ่งซึ่งใช้ตัวแปรอุณหภูมิอากาศเป็นข้อมูลนำเข้าตัวเดียวพบว่า วิธีโครงข่ายประสาทเทียมจะมีประสิทธิภาพสูงกว่าวิธีการวิเคราะห์การถดถอยพหุคูณอย่างชัดเจน ส่วนกรณีอื่น ๆ ที่เหลือพบว่าวิธีโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้นมีประสิทธิภาพดีกว่าวิธีการวิเคราะห์การถดถอยพหุคูณเพียงเล็กน้อย

Shirsath and Singh (2010) ได้ศึกษาเปรียบเทียบการประมาณการระเหยน้ำแบบถาดรายวันระหว่างวิธีโครงข่ายประสาทเทียม วิธีการวิเคราะห์การถดถอยพหุคูณ และตัวแบบทางอุตุนิยมวิทยาหรือสูตรอย่างง่ายสามวิธี ได้แก่ วิธี Penman วิธี Priestley—Taylor และ วิธี Stephens—Stewart เพื่อประมาณการระเหยน้ำแบบถาดรายวันในกรุงนิวเดลี อินเดีย โดยใช้ข้อมูลทางอุตุนิยมวิทยารายวัน 6 ตัวแปร ประกอบด้วย อุณหภูมิอากาศสูงสุด อุณหภูมิอากาศต่ำสุด ระยะเวลาที่มีแสงแดดในแต่ละวัน ความเร็วลม ความชื้นสัมพัทธ์ซึ่งวัด ณ เวลา 7.21 น. และ 14.21 น. (ตามเวลามาตรฐานของอินเดีย) ตามลำดับ ซึ่งเก็บบันทึกโดย Indian Agricultural Research

Institute (IARI) จากปี 2004 — 2007 ทำการสุ่มข้อมูลจำนวน 1,096 วัน (จากข้อมูลทั้งหมด) เพื่อสร้างตัวแบบ และสุ่มข้อมูลอีกส่วนจำนวน 365 วัน เพื่อประเมินประสิทธิภาพของตัวแบบที่ได้ โดยพิจารณาประสิทธิภาพของตัวแบบจากสัมประสิทธิ์ตัวกำหนด (Coefficient of Determination: R^2) และรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: $RMSE$)

ผลการศึกษาของ Shirsath and Singh (2010) พบว่า ทุก ๆ ตัวแบบการระเหยน้ำแบบ ภาตรายวันที่สร้างด้วยวิธีโครงข่ายประสาทเทียมและวิธีการวิเคราะห์การถดถอยพหุคูณ มีความแม่นยำมากกว่าวิธี Penman วิธี Priestley—Taylor และ วิธี Stephens และ Stewart ทั้งนี้ พิจารณาความแม่นยำจากค่า $RMSE$ ที่ต่ำสุดจากชุดข้อมูลที่ใช้ประเมินตัวแบบ ส่วนตัวแบบที่ได้จากวิธีการวิเคราะห์การถดถอยพหุคูณพบว่า ตัวแปรอุณหภูมิอากาศสูงสุด เป็นตัวแปรที่มีความสัมพันธ์กับการระเหยน้ำมากที่สุด รองลงมา คือ ความเร็วลมและความชื้นสัมพัทธ์ ณ เวลา 7.21 น. ตามลำดับ ประสิทธิภาพของตัวแบบที่สร้างด้วยวิธีการวิเคราะห์การถดถอยพหุคูณเมื่อใช้ตัวแปรอุณหภูมิตั้งหกตัวเป็นตัวแปรอิสระ ปรากฏว่า ตัวแบบที่ได้สามารถอธิบายความผันแปรของการระเหยน้ำรายวันได้ 61% แต่เมื่อลดตัวแปรอิสระลงในตัวแบบประหยัด เหลือเพียงตัวแปรอุณหภูมิอากาศสูงสุด ความเร็วลมและความชื้นสัมพัทธ์ ณ เวลา 7.21 น. พบว่า ตัวแบบประหยัดยังคงสามารถอธิบายความผันแปรของการระเหยน้ำรายวันได้ 61% เท่าเดิม สำหรับตัวแบบที่ได้จากวิธีโครงข่ายประสาทเทียม ให้นำเข้าตัวแปรอิสระทั้งหมด และกรณีนำเข้าตัวแปรอิสระเพียงสามตัว (อุณหภูมิอากาศสูงสุด ความเร็วลมและความชื้นสัมพัทธ์ ณ เวลา 7.21 น.) พบว่า ตัวแบบที่ได้สามารถอธิบายความผันแปรของการระเหยน้ำรายวันได้ 69% และ 68% ตามลำดับ จะสังเกตได้ว่า ทั้งกรณีที่ใช้ตัวแปรอิสระทั้งหมดและกรณีที่ใช้ตัวแปรอิสระสามตัวเท่านั้น ตัวแบบที่ได้จากวิธีโครงข่ายประสาทเทียมแม่นยำกว่าตัวแบบที่ได้จากวิธีวิเคราะห์การถดถอยพหุคูณเพียงเล็กน้อย

Dodangeh et al. (2012) ได้ประยุกต์ตัวแบบอนุกรมเวลาเพื่อพยากรณ์แนวโน้มตัวแปรทางอุตุนิยมวิทยาในอนาคต เพื่อประโยชน์ในการจัดการทรัพยากรน้ำ โดยสร้างตัวแบบอนุกรมเวลาของข้อมูลทางอุตุนิยมวิทยา ได้แก่ ความชื้นสัมพัทธ์ การระเหยน้ำ อุณหภูมิอากาศ ความเร็วลม และระยะเวลาที่มีแสงแดดในแต่ละวัน ทำการตรวจสอบคุณสมบัติสหสัมพันธ์ก่อนกำหนดตัวแบบอนุกรมเวลาโดยพิจารณาลำดับของอนุกรมเวลาจากฟังก์ชันสหสัมพันธ์ในตนเอง (Autocorrelation Function: ACF) และสหสัมพันธ์บางส่วน (Partial Autocorrelation Function: PACF) ผลการศึกษาพบว่า ตัวแบบอนุกรมเวลาที่เหมาะสมในการพยากรณ์ความชื้นสัมพัทธ์ การระเหยน้ำ อุณหภูมิอากาศ ความเร็วลม และระยะเวลาที่มีแสงแดดในแต่ละวัน คือ ตัวแบบ ARIMA(0,0,1)(0,0,1) ตัวแบบ ARIMA(2,0,4)(1,1,0) ตัวแบบ ARIMA(4,0,0)(0,1,1) ตัวแบบ ARIMA(1,0,1)(0,1,1) และตัวแบบ ARIMA(1,0,0)(0,1,1) ตามลำดับ

Kim et al. (2013) ศึกษาเกี่ยวกับการประมาณการระเหยน้ำแบบภาตรายวันในประเทศเกาหลีใต้ด้วยวิธีที่ใช้ข้อมูลในการขับเคลื่อน (Different Data-Driven Methods) และข้อมูล ณ เวลาที่เล็ก (Lag-Time Patterns) ข้อมูลทางอุตุนิยมวิทยารายวันที่ใช้ประกอบด้วย อุณหภูมิอากาศเฉลี่ย ความเร็วลมเฉลี่ย ระยะเวลาที่มีแสงแดดในแต่ละวัน ความชื้นสัมพัทธ์เฉลี่ย การแผ่รังสีออกชั้นบรรยากาศโลก และการระเหยน้ำ ตามลำดับ ซึ่งเก็บบันทึกโดยสถานีอุตุนิยมวิทยา Daegu และสถานี

อุตุนิยมวิทยา Ulsan จากปี 1985 – 1990 ทำการเปรียบเทียบการพยากรณ์การระเหยน้ำในแต่ละสถานีอุตุนิยมวิทยาด้วยวิธีโครงข่ายประสาทเทียมสามวิธี (Multilayer Perceptron-Neural Networks Model: MLP-NNM, Generalized Regression Neural Networks Model: GRNNM, Adaptive Neuro-Fuzzy Inference System: ANFIS) กับวิธีการวิเคราะห์การถดถอยพหุคูณ สถิติที่ใช้เปรียบเทียบประสิทธิภาพของการพยากรณ์ ได้แก่ ค่าสัมประสิทธิ์สหสัมพันธ์ (CC) รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ค่าสัมประสิทธิ์ของ Nash-Sutcliffe และค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE)

จากการพยากรณ์การระเหยน้ำด้วยวิธี MLP-NNM วิธี GRNNM และวิธี ANFIS จะใช้ข้อมูลนำเข้าจากกลุ่มตัวแปรที่เกี่ยวข้องกับอุณหภูมิอากาศ (TEM2, TEM1) กลุ่มตัวแปรที่เกี่ยวข้องกับรังสี (RAD2, RAD1) กลุ่มตัวแปรที่เกี่ยวข้องกับระยะเวลาที่มีแสงแดด (SUN1) และกลุ่มที่ใช้เป็นตัวแปรข้อมูลนำเข้าทุกตัว (Merged) พบว่า กรณีที่ใช้ตัวแปรข้อมูลนำเข้าทุกตัวสามารถพยากรณ์การระเหยน้ำในทั้งสองสถานีวัดได้ดีกว่าทุก ๆ ตัวแบบในกรณี ๆ อื่น สำหรับการพยากรณ์การระเหยน้ำในทั้งสองสถานีวัดก็พบเช่นเดียวกันว่า กรณีที่ใช้ตัวแปรข้อมูลนำเข้าหรือตัวแปรอิสระทุกตัวสามารถพยากรณ์ได้ดีกว่าตัวแบบกรณีอื่น ๆ และเมื่อนำตัวแบบที่ดีที่สุดในการพยากรณ์การระเหยน้ำในสถานี Daegu ด้วยวิธี MLP-NNM วิธี GRNNM วิธี ANFIS และวิธีการวิเคราะห์การถดถอยพหุคูณ มาเปรียบเทียบประสิทธิภาพด้วยค่าสัมประสิทธิ์สหสัมพันธ์พบว่า ตัวแบบจากวิธี MLP-NNM วิธี GRNNM วิธี ANFIS และวิธีการวิเคราะห์การถดถอยพหุคูณ มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.832, 0.849, 0.830, 0.817 ตามลำดับ จะเห็นว่าตัวแบบจากวิธี MLP-NNM วิธี GRNNM วิธี ANFIS ดีกว่าวิธีการวิเคราะห์การถดถอยพหุคูณเล็กน้อย สำหรับการพยากรณ์การระเหยน้ำในสถานี Ulsan พบว่าวิธี MLP-NNM วิธี GRNNM วิธี ANFIS และวิธีการวิเคราะห์การถดถอยพหุคูณ มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.806, 0.832, 0.827, 0.786 ตามลำดับ ทำให้เห็นได้ว่า ตัวแบบจากวิธี MLP-NNM วิธี GRNNM วิธี ANFIS ดีกว่าวิธีการวิเคราะห์การถดถอยพหุคูณ

Meshram et al. (2013) ได้พยากรณ์การระเหยน้ำแบบภาคจากอ่างเก็บน้ำ Maknni ประเทศอินเดีย ใช้ข้อมูลการระเหยน้ำแบบภาคเฉลี่ยรายสัปดาห์จากปี 1998-2012 จำนวน 728 คาบเวลา ทำการสร้างตัวแบบการพยากรณ์ด้วยตัวแบบ Seasonal Autoregressive and Moving Average (SARIMA) กำหนดตัวแบบด้วยการพิจารณาลำดับของอนุกรมเวลาจากฟังก์ชันสหสัมพันธ์ในตนเอง (Autocorrelation Function: ACF) และสหสัมพันธ์บางส่วน (Partial Autocorrelation Function: PACF) ประมาณค่าพารามิเตอร์ของตัวแบบด้วยวิธี Maximum likelihood Method และเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์ที่ได้ด้วยรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) ผลการศึกษาพบว่า ตัวแบบอนุกรมเวลาที่เหมาะสมในการพยากรณ์การระเหยน้ำแบบภาคเฉลี่ยรายสัปดาห์ คือ ตัวแบบ SARIMA(1,0,0)(1,1,1)₅₂

Dabral et al. (2014) ได้ศึกษาตัวแบบอนุกรมเวลาของการระเหยน้ำแบบภาค กรณีศึกษาในภาคตะวันออกเฉียงเหนือ ประเทศอินเดีย โดยใช้ข้อมูลการระเหยน้ำแบบภาคเฉลี่ยรายเดือนจากปี ค.ศ. 1965-2005 บันทึกข้อมูลโดยสมาคมวิจัยชา (Tea Research Association) เมือง Jorhat รัฐอัสสัม ประเทศอินเดีย ข้อมูลที่ใช้ในการวิจัยแบ่งเป็นสองชุด ข้อมูลชุดที่ 1 คือ ข้อมูลการระเหยน้ำจากปี ค.ศ. 1965 – 2000 ใช้ในการสร้างตัวแบบอนุกรมเวลา ข้อมูลชุดที่ 2 คือ ข้อมูลการระเหยน้ำ

จากปี ค.ศ. 2001 – 2005 ใช้ในการประเมินความถูกต้องของตัวแบบอนุกรมเวลาที่ได้จากการวิจัยนี้ คณะผู้วิจัยได้กำหนดว่า อนุกรมเวลาของการระเหยน้ำนี้ มีส่วนประกอบของทั้ง Deterministic และ Stochastic ขั้นแรกทำการทดสอบจุดวกกลับ (Turning Point Test) เพื่อหาส่วนประกอบของ แนวโน้ม (Trend) และกำจัดส่วนประกอบของแนวโน้มออกจากชุดข้อมูล จากนั้นวิเคราะห์หา ส่วนประกอบของวัฏจักรด้วยการวิเคราะห์อนุกรมฟูเรียร์ ผลการวิเคราะห์ตัวแบบในส่วน Stochastic พบว่า ตัวแบบอนุกรมเวลาของการระเหยน้ำแบบภาคที่เหมาะสม คือ ตัวแบบ MA อันดับที่สอง ซึ่ง คัดเลือกด้วยเกณฑ์ Bayesian Information Criterion (BIC) ต่ำสุด

Saengprasan et al. (2014) ได้เปรียบเทียบประสิทธิภาพการพยากรณ์การระเหยน้ำใน จังหวัดสกลนครระหว่างการวิเคราะห์การถดถอยพหุคูณกับการวิเคราะห์อนุกรมเวลาแบบ Box-Jenkins ใช้ข้อมูลอุตุนิยมวิทยารายเดือน ประกอบด้วยการระเหยน้ำแบบภาค (มม.) อุณหภูมิอากาศ ($^{\circ}\text{C}$) ความกดอากาศ (hPa) ความชื้นสัมพัทธ์ (%) ปริมาณฝน (มม.) และปริมาณแสงแดด (ชม.) จากเดือน มกราคม พ.ศ. 2550 ถึงเดือนธันวาคม พ.ศ. 2556 เก็บบันทึกข้อมูลโดยสถานีอุตุนิยมวิทยาสกลนคร จังหวัดสกลนคร ทำการพยากรณ์การระเหยน้ำแบบภาคด้วยวิธีการวิเคราะห์การถดถอยพหุคูณ เมื่อ คัดเลือกตัวแปรแบบขั้นบันได (Stepwise Multiple Regression) และวิธีบอกซ์-เจนกินส์ (Box-Jenkins Method) เปรียบเทียบประสิทธิภาพการพยากรณ์ด้วยเกณฑ์ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (Mean Absolute Deviation: MAD) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ที่ต่ำที่สุด พบว่า การพยากรณ์การระเหยน้ำรายเดือนด้วยวิธีการวิเคราะห์การถดถอยพหุคูณมีความแม่นยำกว่าการพยากรณ์ด้วยตัวแบบ Box-Jenkins ซึ่งที่จริงข้อมูลการระเหยน้ำเป็นข้อมูลที่มี ลักษณะของฤดูกาลร่วมด้วย โดยเฉพาะในช่วงเดือนมีนาคมถึงเมษายนที่มีการระเหยน้ำสูงสุดในรอบปี ถือได้ว่าเป็นค่าสูงผิดปกติในรอบปี ฉะนั้นในการปรับปรุงวิธีการพยากรณ์ด้วยการวิเคราะห์การถดถอย พหุคูณ หากนำค่าผิดปกติของข้อมูล (Outliers) มารวมปรับปรุงตัวแบบการพยากรณ์ ย่อมจะได้ความ แม่นยำในการพยากรณ์มากขึ้น

สรุปผลการวิจัยที่เกี่ยวข้องข้างต้น จะเห็นทิศทางวิธีการเชิงปริมาณที่นักวิจัยใช้ในประมาณ ค่าการระเหยน้ำหรือการคายระเหยน้ำ สามารถแบ่งออกเป็น 2 วิธีหลัก ๆ กล่าวคือ วิธีที่หนึ่ง พยากรณ์การระเหยน้ำโดยอาศัยความสัมพันธ์ระหว่างการระเหยน้ำกับข้อมูลทางอุตุนิยมวิทยาอื่น ๆ ยกตัวอย่างวิธีการ เช่น การวิเคราะห์การถดถอยพหุคูณ การประยุกต์ของโครงข่ายประสาทเทียมแบบ ต่าง ๆ หรือการใช้สูตรอย่างง่าย (Empirical Formula) ดังการศึกษาของ Kisi (2006), Goel (2009), Shirsath and Singh (2010) และ Kim et al. (2013) ส่วนวิธีที่สอง พยากรณ์การระเหยน้ำโดย อาศัยความสัมพันธ์ระหว่างการระเหยน้ำในปัจจุบันกับในอดีต ด้วยการวิเคราะห์อนุกรมเวลาแบบ Box-Jenkins ดังการศึกษาของ Moshrik et al. (2008), Dodangeh et al. (2012), Meshram et al. (2013) และ Dabral et al. (2014)

สรุปผลการศึกษารายละเอียดของกลุ่มนักวิจัยที่เลือกใช้วิธีการเชิงปริมาณในกลุ่มวิธีที่หนึ่ง สอดคล้องกันว่า เมื่อใช้ทุก ๆ ตัวแปรอุตุนิยมวิทยาที่ศึกษาสร้างตัวแบบ วิธีโครงข่ายประสาทเทียมมี ความแม่นยำกว่าวิธีการวิเคราะห์การถดถอยพหุคูณเพียงเล็กน้อย แต่ผู้วิจัยพบข้อสังเกตว่า ค่า สัมประสิทธิ์ที่ตัวกำหนด (Coefficient of Determination: R^2) ของตัวแบบจากวิธีโครงข่ายประสาท เทียม มีค่าสูงกว่า R^2 ของตัวแบบการวิเคราะห์การถดถอยพหุคูณเพียงเล็กน้อย หรือเกือบจะใกล้เคียง

กัน (Goel, 2009; Shirsath & Singh, 2010) ดังนั้น หลักการวิเคราะห์การถดถอยพหุคูณจึงเป็นวิธีการที่น่าสนใจที่นำไปใช้ประมาณค่าการระเหยน้ำ เนื่องจากมีขั้นตอนการวิเคราะห์ที่ไม่ซับซ้อนเหมือนวิธีโครงข่ายประสาทเทียมแบบต่าง ๆ และสอดคล้องกับผลการศึกษาเบื้องต้นของ Saengprasan et al. (2014) ที่ว่า การวิเคราะห์การถดถอยพหุคูณสามารถประมาณค่าการระเหยน้ำในจังหวัดสกลนครได้ดีกว่าการวิเคราะห์อนุกรมเวลาแบบ Box-Jenkins ดังนั้น แนวทางการพัฒนาตัวประมาณค่าแบบอัตราส่วน กรณีที่ใช้ตัวแปรช่วยสองตัว สามารถนำการวิเคราะห์การถดถอยแกร่งไปประยุกต์หรือพัฒนาให้ได้ตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่ ให้เหมาะกับการประมาณค่าการระเหยน้ำโดยเฉลี่ยหรือข้อมูลอื่น ๆ ที่มีค่าผิดปกติได้

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัย เรื่อง การประมาณค่าการระเหยน้ำด้วยตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่ โดยใช้ตัวแปรช่วยสองตัว มีวิธีดำเนินการวิจัย ดังนี้

ตอนที่ 1 การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ตอนที่ 2 การจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$

ตอนที่ 3 การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ตอนที่ 1 การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ขั้นตอนนี้เป็นการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อใช้ตัวแปรช่วยสองตัว และข้อมูลประชากรมีค่าผิดปกติ มีขั้นตอนดังนี้

1. ศึกษาวรรณกรรมที่เกี่ยวข้องกับการพัฒนาตัวประมาณค่าแบบอัตราส่วน (Ratio Estimator) ต่าง ๆ ที่มีผู้นำเสนอไว้ก่อนหน้านี้ ตลอดจนการปรับปรุงตัวประมาณค่าแบบอัตราส่วน โดยใช้สารสนเทศของตัวแปรช่วย เช่น การใช้ค่าสัมประสิทธิ์การแปรผัน ค่าสัมประสิทธิ์ความโค้งหรือค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วย เป็นต้น งานวิจัยนี้ได้ศึกษาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร โดยใช้ตัวแปรช่วยสองตัวของ Gupta and Shabbir (2007) เป็นตัวประมาณค่าเริ่มต้น (Initial Estimator) เพื่อพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ทั้งนี้ตัวประมาณค่าแบบอัตราส่วนของ Gupta and Shabbir (2007) ที่นำเสนอไว้ในรูปแบบ Ratio-Cum-Product Type ดังสมการ (3.1)

$$\bar{y}_{m(SD)} = \bar{y} \frac{\frac{C_x \bar{X} + C_z \bar{Z}}{C_x \bar{x} + C_z \bar{z}}}{\frac{C_x \bar{X} + C_z \bar{Z}}{C_x \bar{x} + C_z \bar{z}}} \quad (3.1)$$

เมื่อ C_x คือ สัมประสิทธิ์การแปรผันของประชากรของตัวแปรช่วย X

C_z คือ สัมประสิทธิ์การแปรผันของประชากรของตัวแปรช่วย Z

\bar{X} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย X

\bar{Z} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย Z

\bar{x} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย X

\bar{z} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย Z

และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ดังสมการ (3.2)

$$MSE(\bar{y}_{m(SD)}) \approx \bar{y}^2 \lambda \left[\frac{C_y^2}{C_x^2} + \alpha_1^2 C_x^2 + \alpha_2^2 C_z^2 - 2\alpha_1 \rho_{yx} C_y C_x + 2\alpha_2 \rho_{yz} C_y C_z - 2\alpha_1 \alpha_2 \rho_{xz} C_x C_z \right] \quad (3.2)$$

เมื่อ ρ_{yx} คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร Y กับตัวแปรช่วย X

ρ_{yz} คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร Y กับตัวแปรช่วย Z

ρ_{xz} คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรช่วย X กับตัวแปรช่วย Z

$$\text{และ } \alpha_1 = \frac{\bar{X}}{\bar{X} + C_x}, \alpha_2 = \frac{\bar{Z}}{\bar{Z} + C_z}$$

2. ศึกษาวรรณกรรมที่เกี่ยวข้องการวิเคราะห์การถดถอยแกร่ง (Robust Regression) ในด้านประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีต่าง ๆ ทั้งนี้ Rousseeuw and Leroy (2003, p. 9, pp. 12-15) ได้อธิบายการวิเคราะห์การถดถอยแกร่งที่มีตัวประมาณค่าสัมประสิทธิ์การถดถอยไว้หลายวิธี เช่น วิธีค่าสัมบูรณ์น้อยสุด (Least Absolute Values) วิธี M-Estimator วิธี GM-Estimator วิธี LMS และวิธีกำลังสองตัดขอบน้อยสุด (LTS) เป็นต้น นอกจากนี้ยังมีวิธี Huber-M Estimator ซึ่ง Kadilar, Candan and Cingi (2007) ใช้ประมาณค่าสัมประสิทธิ์การถดถอยด้วย และนำแนวคิดการวิเคราะห์การถดถอยแกร่งร่วมกับการใช้ค่าสัมประสิทธิ์ความเบ้ของตัวแปรช่วย มาพัฒนาตัวประมาณค่าของ Gupta and Shabbir (2007) ให้มีความเหมาะสมสำหรับใช้วิเคราะห์ข้อมูลที่มีค่าผิดปกติ ด้วยแนวคิดดังกล่าว ผู้วิจัยจึงนำเสนอตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ดังนี้

$$\bar{y}_M = \frac{\bar{y} + b_1^*(\bar{X} - \bar{x}) + b_2^*(\bar{Z} - \bar{z})}{\frac{\bar{x} + \beta_{1(x)}\bar{z}}{\bar{x} + \beta_{1(x)}} + \frac{\bar{z} + \beta_{1(z)}\bar{x}}{\bar{z} + \beta_{1(z)}}} \quad (3.3)$$

เมื่อ $\beta_{1(x)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย X

$\beta_{1(z)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย Z

b_1^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย X

b_2^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย Z

\bar{X} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย X

\bar{Z} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย Z

\bar{x} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย X

\bar{z} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย Z

3. ศึกษาแนวคิดทางการวิเคราะห์เชิงตัวเลขและทฤษฎีบทเทย์เลอร์ (Burden & Faires, 2011, pp. 10-11) ที่ใช้หาค่าประมาณของฟังก์ชันที่หาอนุพันธ์ได้อย่างต่อเนื่อง ทฤษฎีบทเทย์เลอร์เป็นดังนี้

ให้ $f(x)$ เป็นฟังก์ชันที่หาอนุพันธ์ทุกอันดับได้ที่จุด x_0 แล้ว

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

$$\begin{aligned}
&= f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \\
&\quad \frac{f'''(x_0)}{3!}(x - x_0)^3 + \dots
\end{aligned} \tag{3.4}$$

ทฤษฎีบทเทย์เลอร์ดังสมการ (3.4) จะเรียกว่า อนุกรมเทย์เลอร์ สามารถนำไปกระจายเทอมโพลิโนเมียล เพื่อใช้หาสมการประมาณค่า MSE ของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M)

4. ใช้วิธีการกระจายเทอมโพลิโนเมียลของอนุกรมเทย์เลอร์ดีกรีอันดับหนึ่ง (ภทรวรรณ แสงนวกิจ, 2553; Abu-Dayyeh et al., 2003; Kadilar & Cingi, 2004; 2005; 2006; 2007; Gupta & Shabbir, 2007) เพื่อหาสมการ MSE ของตัวประมาณค่าที่ปรับใหม่ ดังนี้

กำหนดให้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) และค่าเฉลี่ยประชากรที่แท้จริงหรือค่าพารามิเตอร์ (\bar{Y}) เป็นฟังก์ชันของตัวแปรศึกษา (Y) ตัวแปรช่วย X และตัวแปรช่วย Z ได้ดังสมการ (3.5) และ (3.6)

$$\bar{y}_M = h(\bar{x}, \bar{z}, \bar{y}) \tag{3.5}$$

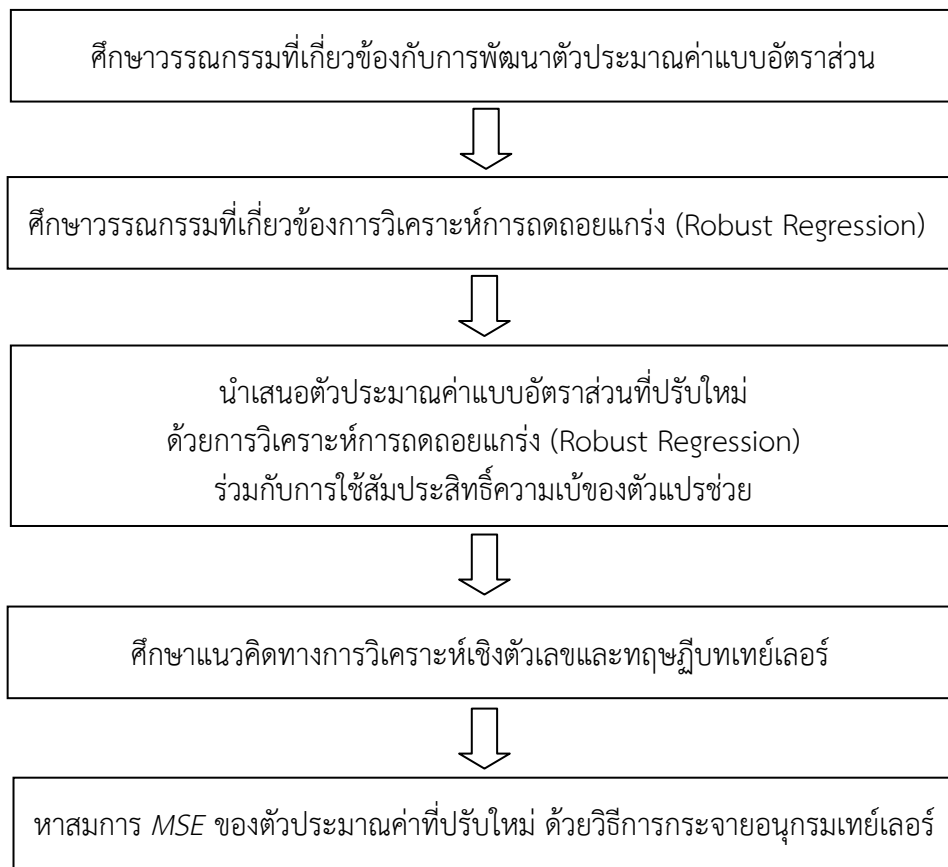
$$\bar{Y} = h(\bar{X}, \bar{Z}, \bar{Y}) \tag{3.6}$$

แล้วกระจายเทอมโพลิโนเมียลของอนุกรมเทย์เลอร์ใน (3.4) ถึงพจน์ดีกรีอันดับหนึ่ง จะได้

$$\begin{aligned}
h(\bar{x}, \bar{z}, \bar{y}) &\cong h(\bar{X}, \bar{Z}, \bar{Y}) + \frac{\left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}}} (\bar{x} - \bar{X}) \\
&\quad + \frac{\left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{z}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{z}} \right|_{\bar{x}, \bar{z}, \bar{y}}} (\bar{z} - \bar{Z}) + \frac{\left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{y}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left. \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{y}} \right|_{\bar{x}, \bar{z}, \bar{y}}} (\bar{y} - \bar{Y})
\end{aligned} \tag{3.7}$$

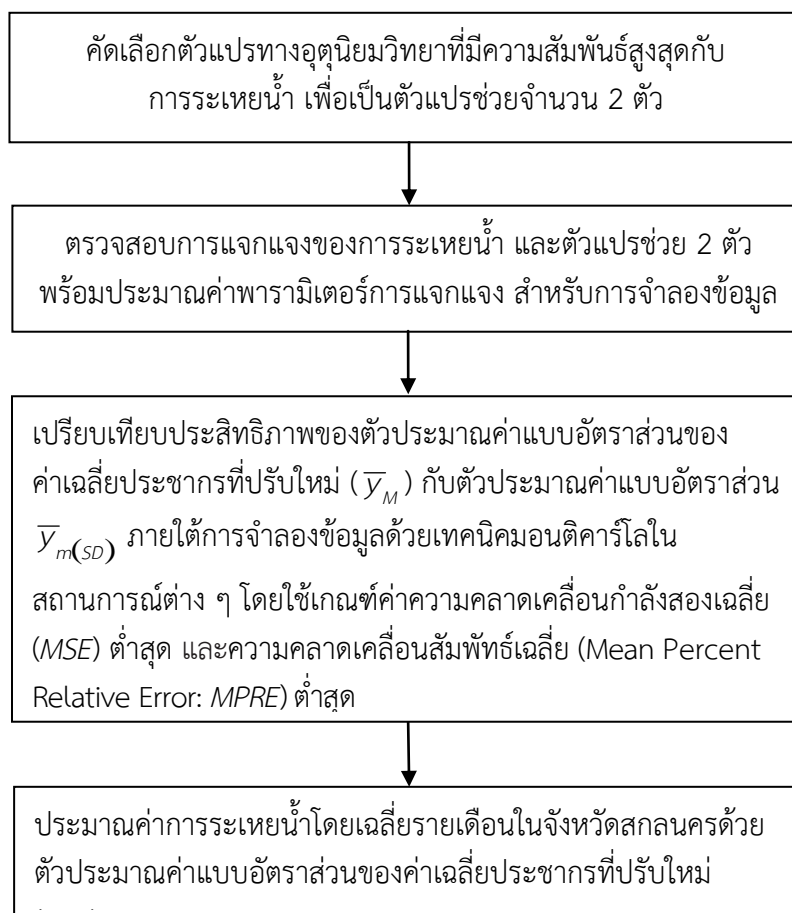
สมการ (3.7) ต้องจัดเทอมสมการให้อยู่ในรูป $E(\bar{y}_M - \bar{Y})^2$ ซึ่ง $MSE(\bar{y}_M) = E(\bar{y}_M - \bar{Y})^2$

จากที่กล่าวมาข้างต้น สามารถสรุปแผนผังของการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อใช้สารสนเทศตัวแปรช่วยสองตัว และข้อมูลประชากรมีค่าผิดปกติ ดังภาพที่ 3-1



ภาพที่ 3-1 แผนผังของการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

เมื่อผู้วิจัยพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ พร้อมทั้งหาสมการ MSE เรียบร้อยแล้ว จึงดำเนินการตามวัตถุประสงค์การวิจัยข้อ 2) ต่อไปนี้ ซึ่งกรอบแนวคิดของขั้นตอนการวิจัย สำหรับวัตถุประสงค์การวิจัยข้อ 2) เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{Y}_{m(SD)}$ และวัตถุประสงค์การวิจัยข้อ 3) เพื่อประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ดังภาพที่ 3-2



ภาพที่ 3-2 กรอบแนวคิดของขั้นตอนการวิจัย สำหรับวัตถุประสงค์การวิจัยข้อ 2) และข้อ 3)

จากภาพที่ 3-2 จะเห็นว่า ขั้นตอนแรก ต้องศึกษาเบื้องต้น (Pilot Study) เกี่ยวกับลักษณะของข้อมูลอุตุนิยมวิทยาทั้งหมด ได้แก่ การหาค่าสูงสุด ค่าต่ำสุด ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน สัมประสิทธิ์ความเบ้ และสัมประสิทธิ์ความโด่ง ตลอดจนจรรยาบรรณความสัมพันธ์รายคู่ระหว่างการระเหยน้ำกับข้อมูลอุตุนิยมวิทยาเหล่านั้น เพื่อคัดเลือกตัวแปรทางอุตุนิยมวิทยาที่มีความสัมพันธ์สูงสุดกับการระเหยน้ำ เพื่อเป็นตัวแปรช่วยจำนวน 2 ตัว จากนั้นจึงตรวจสอบการแจกแจงข้อมูลทางสถิติ (Statistical Distribution) ของข้อมูลที่คัดเลือกได้ เพื่อทำการจำลองข้อมูลต่อไป

ตอนที่ 2 การจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$

ขั้นตอนนี้เป็นการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) จากการทบทวนวรรณกรรมงานวิจัยที่เกี่ยวกับการประมาณค่าการระเหยน้ำหลาย ๆ เรื่อง (Irmak & Haman, 2003; Kişi, 2006; Goel, 2009; Shirsath & Singh, 2010; Kim et al., 2013; Saengprasan et al., 2014) พบว่า ตัวแปรที่ใช้ในการประมาณค่าการระเหยน้ำ คือ ข้อมูลอุตุนิยมวิทยาที่เกี่ยวข้องกับการระเหยน้ำ ประกอบด้วย อุณหภูมิอากาศเฉลี่ย อุณหภูมิอากาศสูงสุด อุณหภูมิอากาศต่ำสุด ความเร็วลมเฉลี่ย ระยะเวลาที่มีแสงแดดในแต่ละวัน ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย ความชื้นสัมพัทธ์เฉลี่ย การแผ่รังสีนอกชั้นบรรยากาศโลก ปริมาณรังสีดวงอาทิตย์ ความกดอากาศเฉลี่ย และปริมาณฝนโดยเฉลี่ย ดังนั้น งานวิจัยนี้จึงคัดเลือกตัวแปรอุตุนิยมวิทยาที่เกี่ยวข้องกับการระเหยน้ำ จากวรรณกรรมข้างต้น ซึ่งใช้ตัวแปรอุตุนิยมวิทยาที่มีความสอดคล้องกันตรงกัน จำนวน 5 ตัวแปรได้แก่ อุณหภูมิอากาศ ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%)

ก่อนการจำลองข้อมูล จะทำศึกษาลักษณะเบื้องต้นของการระเหยน้ำและข้อมูลอุตุนิยมวิทยาที่เกี่ยวข้องหรือมีความสัมพันธ์กับการระเหยน้ำก่อน เพื่อให้ทราบลักษณะการเกิดข้อมูลการระเหยน้ำและปัจจัยอื่นที่เกี่ยวข้อง จากนั้นจึงทำการแสดงเชิงตัวเลขด้วยวิธีการจำลองข้อมูลโดยเทคนิคมอนติคาร์โล เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ รายละเอียดดังนี้

1. การศึกษาลักษณะข้อมูลเบื้องต้น (Pilot Study)

เนื่องจากการวิจัยนี้ใช้ข้อมูลอุตุนิยมวิทยาเฉลี่ยรายเดือน ประกอบด้วยการระเหยน้ำ (มิลลิเมตร) อุณหภูมิอากาศ ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%) จากเดือนมกราคม 2550 ถึงเดือนเมษายน 2556 จำนวน 76 เดือน ที่เก็บบันทึกโดยสำนักงานอุตุนิยมวิทยาจังหวัดสกลนคร ขั้นตอนการศึกษาลักษณะข้อมูลเบื้องต้นโดยใช้สถิติพื้นฐานและสถิติอนุมาน ดังนี้

1.1 สถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด ใช้ในการอธิบายลักษณะข้อมูลการระเหยน้ำและข้อมูลอุตุนิยมวิทยาอย่างคร่าว ๆ ซึ่งผู้วิจัยจะใช้ค่าสถิติพื้นฐานเหล่านี้ ในการกำหนดลักษณะการแจกแจงของข้อมูลในขั้นตอนการจำลองข้อมูล ใช้แผนภาพการกระจาย (Scatter Plot) เพื่ออธิบายลักษณะความสัมพันธ์ระหว่างการระเหยน้ำกับตัวแปรช่วยแต่ละตัว และใช้แผนภาพกล่องในการตรวจสอบค่าผิดปกติของการระเหยน้ำ

1.2 สถิติอนุมาน ได้แก่

1.2.1 สถิติทดสอบ t ใช้ทดสอบความมีนัยสำคัญของสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างการระเหยน้ำกับข้อมูลอุตุนิยมวิทยา ทีละรายคู่ เพื่อคัดเลือกตัวแปรช่วยสองตัวแรกที่มีความสัมพันธ์กับการระเหยน้ำมากที่สุด โดยพิจารณาเลือกตัวแปรสองตัวแรกที่มีความสัมพันธ์สูงที่สุดกับการระเหยน้ำรายเดือนในจังหวัดสกลนคร และตัวแปรช่วยสองตัวนั้นจะต้องมีค่าสัมประสิทธิ์

สหสัมพันธ์เพียร์สันไม่เกิน 0.7 และค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรสองตัวนั้น จะต้องมีย่านน้อยกว่าความสัมพันธ์ระหว่างตัวแปรช่วยใด ๆ กับการระเหยน้ำรายเดือน ทั้งนี้เพื่อป้องกันการเกิดปัญหาตัวแปรช่วยมีความสัมพันธ์กันเอง หรือที่เรียกว่าปัญหา Collinearity (Hair, Black, Babin, & Anderson, 2010, p. 205)

1.2.2 สถิติทดสอบ Anderson-Darling Test ใช้ในการทดสอบการแจกแจงของตัวแปรการระเหยน้ำและตัวแปรช่วย เนื่องจากสถิติ Anderson-Darling Test เป็นสถิติที่มีความเหมาะสมในการทดสอบการแจกแจงข้อมูลที่มีความเบ้และการแจกแจงข้อมูลที่มีมาตราแบบหางยาว (Yap & Sim, 2011) ทั้งนี้การทดสอบการแจกแจงของข้อมูลด้วยสถิติ Anderson-Darling Test วิเคราะห์ด้วยโปรแกรม Minitab

การทดสอบการแจกแจงข้อมูลตัวแปรการระเหยน้ำและตัวแปรช่วย ตั้งสมมติฐานการทดสอบ (จุฑาพร เนียมวงศ์ และธนภฤต เตยานุรักษ์, 2558, หน้า 1-8) ดังนี้

H_0 : ประชากรของตัวอย่างสุ่มมีการแจกแจงตามที่คาดหวัง

H_1 : ประชากรของตัวอย่างสุ่มไม่มีการแจกแจงตามที่คาดหวัง

สถิติทดสอบ Anderson-Darling Test (AD) คือ

$$AD = -n - \frac{1}{n} \sum_{i=1}^n [(2i-1)\log u_i + (2n+1-i)\log(1-u_i)] \quad (3.8)$$

เมื่อ u_i คือ ความถี่สะสมของค่าสังเกตลำดับที่ i เมื่อ $i = 1, 2, \dots, n$

n คือ จำนวนข้อมูลทั้งหมด

ในกรณีการทดสอบ Anderson-Darling Test ให้ผลการยอมรับสมมติฐานหลักมากกว่า 1 การแจกแจงขึ้นไป จะทำการคำนวณหาค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เพื่อหาการแจกแจงที่ใกล้เคียงกับข้อมูลมากที่สุด โดยค่า RMSE คำนวณได้ดังนี้

$$RMSE = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - X_i^a)^2} \quad (3.9)$$

เมื่อ X_i คือ ค่าจริงของข้อมูล

X_i^a คือ ค่าประมาณของข้อมูล

n คือ จำนวนข้อมูลทั้งหมด

หากการทดสอบ Anderson-Darling Test ให้ผลปฏิเสธสมมติฐานหลักทุก ๆ การแจกแจงที่ทำการทดสอบ นั่นคือ ไม่สามารถระบุการแจกแจงของข้อมูลที่เหมาะสมได้ จะทำการแปลงข้อมูลด้วยวิธี Johnson Transformation และทดสอบการแจกแจงปกติสำหรับข้อมูลที่แปลงค่าแล้วด้วยสถิติ Anderson-Darling Test

เมื่อระบุการแจกแจงของข้อมูลที่เหมาะสมได้เรียบร้อยแล้ว จึงนำค่าประมาณพารามิเตอร์ที่ได้ และสมการแปลงค่าข้อมูลที่ได้ (ถ้ามี) ไปทำการจำลองข้อมูลต่อไป

2. การจำลองข้อมูล (Simulation)

หลังจากคัดเลือกตัวแปรช่วยสองตัวที่มีความสัมพันธ์สูงสุดกับการระเหยน้ำรายเดือนในจังหวัดสกลนคร ดังข้อ 1.2.1 แล้ว ในที่นี้ผลการคัดเลือกตัวแปรช่วย ได้แก่ อุณหภูมิอากาศ (X) และความชื้นสัมพัทธ์ (Z) จากนั้นจึงนำค่าสถิติพื้นฐานที่ได้จากข้อ 1.1 และผลการทดสอบการแจกแจงตัวแปรการระเหยน้ำและตัวแปรช่วย (ซึ่งผลการทดสอบการแจกแจงข้อมูลในข้อ 1.2.2 จะได้ว่า การทดสอบการแจกแจงข้อมูลเบื้องต้นของอุณหภูมิอากาศและความชื้นสัมพัทธ์ ไม่สามารถระบุการแจกแจงใด ๆ ที่เหมาะสมกับข้อมูลได้ จึงต้องทำการแปลงค่าข้อมูลก่อน จึงสามารถทดสอบได้ว่า ข้อมูลที่แปลงค่าแล้วมีการแจกแจงปกติ รายละเอียดดังภาคผนวก ก) มาทำการจำลองข้อมูล (Simulation) ด้วยโปรแกรม Matlab

ขั้นตอนการจำลองข้อมูลเพื่อสร้างประชากรและขั้นตอนการทดสอบประสิทธิภาพของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ดังนี้

2.1 ขั้นตอนการจำลองข้อมูลประชากร

กำหนดประชากรขนาด 10,000 หน่วย การสร้างประชากรสำหรับการจำลองข้อมูล จะอาศัยผลลัพธ์จากข้อ 1.2.2 ที่ว่า อุณหภูมิอากาศ (X) และ ความชื้นสัมพัทธ์ (Z) ที่แปลงค่าข้อมูลแล้วมีการแจกแจงปกติ ซึ่งมีค่าประมาณพารามิเตอร์ของการแจกแจงปกติ ดังภาคผนวก ก ขั้นตอนการสร้างประชากรสำหรับการจำลองข้อมูล เป็นดังนี้

2.1.1 การสร้างข้อมูลตัวแปร X

เนื่องจากอุณหภูมิอากาศ (X) ทำการแปลงข้อมูล (Transform) ด้วยสมการ (3.10)

$$X_t = 1.2251 + 0.9295 \operatorname{Asinh} \frac{X_t - 1.2251}{0.9622} \quad (3.10)$$

เมื่อ Asinh คือ ฟังก์ชันไฮเพอร์โบลิกไซน์ผกผัน

จะได้ว่า $X_t : N(0.03141, 0.9168)$

ดังนั้นการสร้างข้อมูลตัวแปรอุณหภูมิอากาศ จะสุ่มตัวแปร

$X_t : N(0.03141, 0.9168)$ ก่อน จากนั้นจึงทำการแปลงค่า X_t ย้อนกลับเพื่อหาค่าตัวแปร X ด้วยสมการย้อนกลับของสมการ (3.10) ดังสมการ (3.11)

$$X = 1.2251 + 0.9622 \operatorname{sinh} \frac{X_t - 1.2251}{0.9295} \quad (3.11)$$

เมื่อ sinh คือ ฟังก์ชันไฮเพอร์โบลิกไซน์

2.1.2 การสร้างข้อมูลตัวแปร Z

เนื่องจากความชื้นสัมพัทธ์ (Z) ทำการแปลงข้อมูล (Transform) ด้วยสมการ (3.12)

$$Z_t = -0.2517 + 0.7113 \ln \frac{Z_t - 54.7732}{86.2413 - Z_t} \quad (3.12)$$

เมื่อ \ln คือ ฟังก์ชันผกผันของฟังก์ชันเอ็กโปเนนเชียล ที่ส่งกลับค่าลอการิทึมธรรมชาติของตัวเลข

$$\text{จะได้ว่า } Z_t : N(0.0148, 0.9821)$$

ดังนั้นการสร้างข้อมูลตัวแปรความชื้นสัมพัทธ์ จะสุ่มตัวแปร

$Z_t : N(0.0148, 0.9821)$ ก่อน จากนั้นจึงทำการแปลงค่า Z_t ย้อนกลับเพื่อหาค่าตัวแปร Z ด้วยสมการย้อนกลับของสมการ (3.12) ดังสมการ (3.13)

$$Z = \frac{86.2413e^{\frac{\alpha Z_t + 0.2517 \frac{0}{\sigma}}{0.7113}} + 54.7732}{1 + e^{\frac{\alpha Z_t + 0.2517 \frac{0}{\sigma}}{0.7113}}} \quad (3.13)$$

เมื่อ e คือ ฟังก์ชันเอ็กโปเนนเชียล

2.1.3 สร้างตัวแปร Y หรือการระเหยน้ำที่มีความสัมพันธ์กับ X และ Z ในรูป

$Y_i = \alpha + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i$ อาศัยผลการวิเคราะห์การถดถอยพหุคูณจากข้อมูล อุตุนิยมวิทยาเฉลี่ยรายเดือนของการระเหยน้ำ อุณหภูมิอากาศ และความชื้นสัมพัทธ์ จากเดือนมกราคม พ.ศ. 2550 ถึงเดือนเมษายน พ.ศ. 2556 จำนวน 76 เดือน (ดังภาคผนวก ข) ทำให้กำหนดสมการการถดถอย คือ $Y_i = 3.877 + .199X_i - .067Z_i + \varepsilon_i$ โดยที่ $\varepsilon_i : N(0,1)$

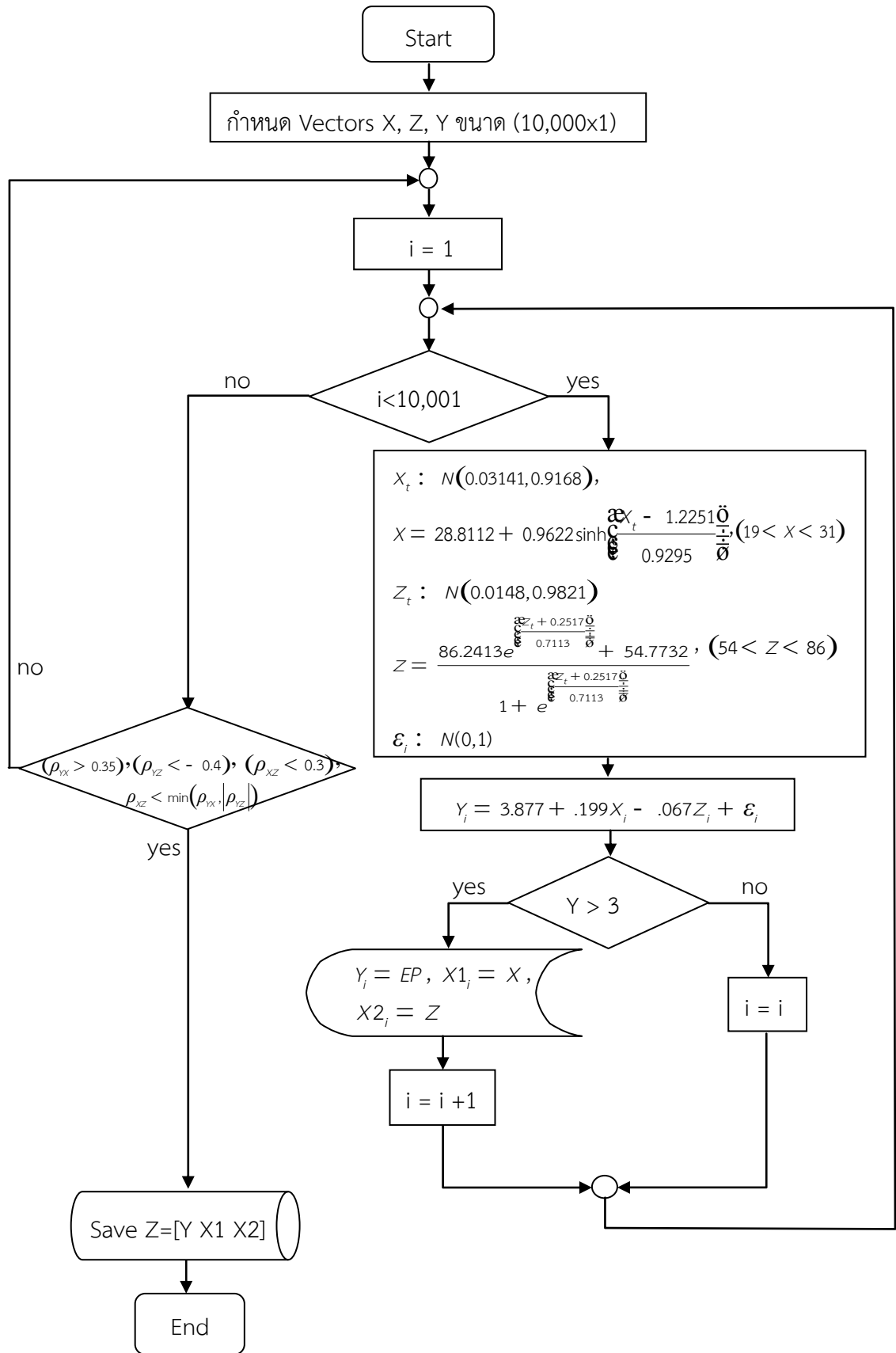
2.1.4 การคัดเลือกค่าข้อมูลตัวแปร Y , X และ Z ที่ได้จากข้อ 2.1.1 ถึงข้อ 2.1.3 หรืออันดับ (Y, X, Z) ที่เหมาะสมในการสร้างประชากร ซึ่งกำหนดให้สอดคล้องกับลักษณะข้อมูลเบื้องต้นที่แท้จริงของข้อมูลการระเหยน้ำในจังหวัดสกลนคร (Y) และตัวแปรช่วย อุณหภูมิอากาศ (X) และความชื้นสัมพัทธ์ (Z) ตามค่าสถิติพื้นฐานที่ได้จากข้อ 1.1 ดังนี้

- อุณหภูมิอากาศมีค่าอยู่ระหว่าง 19°C ถึง 31°C ($19 < X < 31$)
 - ความชื้นสัมพัทธ์มีค่าอยู่ระหว่าง 54 ถึง 86 ($54 < Z < 86$)
 - การระเหยน้ำมีค่ามากกว่า 3 มิลลิเมตรขึ้นไป ($Y > 3$)
 - ความสัมพันธ์ระหว่างการระเหยน้ำกับอุณหภูมิอากาศมีทิศทางเดียวกันและอยู่ในระดับปานกลาง ($\rho_{YX} > 0.35$)
 - ความสัมพันธ์ระหว่างการระเหยน้ำกับความชื้นสัมพัทธ์มีทิศทางผกผันและอยู่ในระดับปานกลาง ($\rho_{YZ} < -0.4$)
 - ความสัมพันธ์ระหว่างอุณหภูมิอากาศกับความชื้นสัมพัทธ์มีทิศทางเดียวกันและอยู่ในระดับน้อย ($\rho_{XZ} < 0.3$) แต่ ρ_{XZ} จะต้องไม่มีค่ามากกว่า ρ_{YX} หรือ ρ_{YZ}
- ในขั้นตอน 2.1.4 นี้จะอาศัยค่าพิสัยของข้อมูลในการกำหนดขอบเขตค่าข้อมูล และทำการวัดระดับความสัมพันธ์รายคู่ ρ_{YX} , ρ_{YZ} และ ρ_{XZ} ด้วยสัมประสิทธิ์สหสัมพันธ์เพียร์สันพร้อมกับทดสอบความมีนัยสำคัญของความสัมพันธ์เหล่านั้นด้วยสถิติ t-Test หากผลการทดสอบความสัมพันธ์รายคู่ ρ_{YX} , ρ_{YZ} และ ρ_{XZ} ไม่มีนัยสำคัญหรือไม่เป็นตามเงื่อนไขที่กำหนด จะทำการสุ่ม

ข้อมูล (Y, X, Z) จากข้อ 2.1.1 ถึงข้อ 2.1.3 ใหม่ จนกระทั่งได้ชุดข้อมูลประชากร (Y, X, Z) ที่มีลักษณะสอดคล้องกับเกณฑ์คัดเลือกข้อมูลในข้อ 2.1.4

2.1.5 บันทึกข้อมูลประชากรในรูปอันดับ (Y, X, Z) จำนวน 10,000 ค่าที่ผ่านการคัดเลือกในข้อ 2.1.4

สำหรับการเขียนโปรแกรมเพื่อจำลองข้อมูลประชากรในรูปอันดับ (Y, X, Z) จากข้อ 2.1.1 ถึงข้อ 2.1.5 สรุปได้ดังภาพที่ 3-3



ภาพที่ 3-3 การเขียนโปรแกรมเพื่อจำลองข้อมูลประชากร

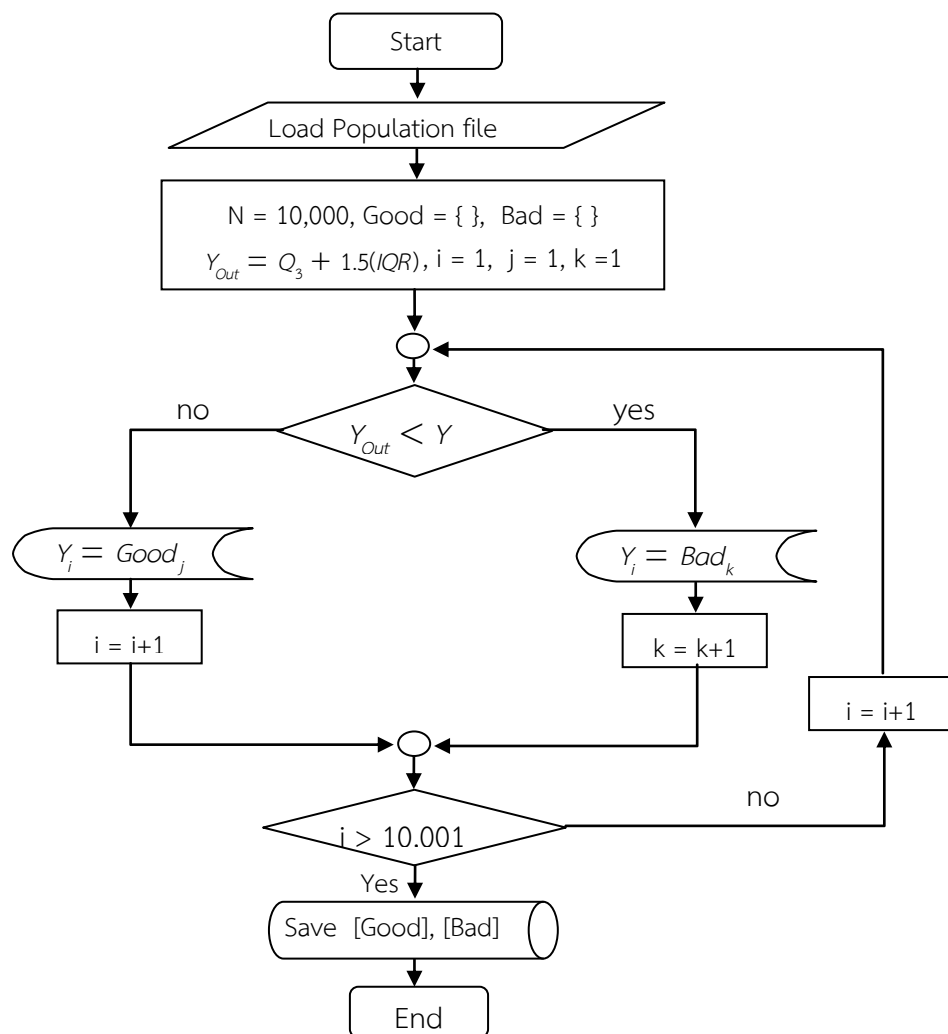
2.2 ขั้นตอนการทดสอบประสิทธิภาพของตัวประมาณค่า

เมื่อสร้างจำลองข้อมูลประชากรแล้ว ทำการทดสอบประสิทธิภาพของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ภายใต้ 12 สถานการณ์และสุ่มข้อมูลตัวอย่างด้วยวิธีการชักตัวอย่างแบบสุ่มเชิงเดียว ดังนี้

2.2.1 ทำการดึงข้อมูล (Load) ข้อมูลประชากร

2.2.2 ทำการตรวจสอบค่าผิดปกติ เพื่อคัดแยกข้อมูลประชากรออกเป็นสองกลุ่มย่อย คือ 1) กลุ่มข้อมูลที่มีค่าปกติ (Non Outliers) กับ 2) กลุ่มข้อมูลที่เป็นค่าผิดปกติ (Outliers) ทั้งนี้อาศัยค่าพิสัยควอร์ไทล์ (Interquartile Range: IQR) และเกณฑ์ในการคัดแยกข้อมูลประชากร คือ ข้อมูลการระเหยน้ำใดมีค่ามากกว่า $Q_3 + 1.5(IQR)$ ถือว่าเป็นข้อมูลผิดปกติ

สำหรับการเขียนโปรแกรมเพื่อคัดแยกข้อมูลประชากรออกเป็นสองกลุ่มย่อย คือ 1) กลุ่มข้อมูลที่มีค่าปกติ ((Non Outliers) กับ 2) กลุ่มข้อมูลที่เป็นค่าผิดปกติ (Outliers) แสดงขั้นตอนได้ดังภาพที่ 3-4



ภาพที่ 3-4 การเขียนโปรแกรมเพื่อคัดแยกข้อมูลประชากรเป็นกลุ่มข้อมูลปกติและกลุ่มข้อมูลผิดปกติ

2.2.3 กำหนดเงื่อนไขย่อยของแต่ละสถานการณ์การจำลอง ได้แก่

- กำหนดขนาดตัวอย่างสุ่ม 3 เงื่อนไข คือ ขนาดตัวอย่าง เท่ากับ 20 40 และ 100 หน่วยตัวอย่าง

- กำหนดร้อยละของค่าผิดปกติ ให้มีค่าผิดปกติเกิดขึ้น 4 เงื่อนไข คือ ร้อยละของค่าผิดปกติ เท่ากับ ร้อยละ 5 ร้อยละ 10 ร้อยละ 15 และ ร้อยละ 20

2.2.4 กำหนดรอบการทำซ้ำ รอบที่ i เมื่อ $i = 1, 2, 3, \dots, 1,000$

2.2.5 ทำการสุ่มตัวอย่างตัวอย่างขนาด n ตามเงื่อนไขร้อยละของค่าผิดปกติ จะได้สถานการณ์ทั้งหมดที่ศึกษา เท่ากับ 12 สถานการณ์ (ขนาดตัวอย่างสุ่ม 3 เงื่อนไข \times ร้อยละของค่าผิดปกติ 4 เงื่อนไข) สำหรับการสุ่มตัวอย่างเพื่อให้ได้ขนาดตัวอย่างที่มีร้อยละของค่าผิดปกติตามเงื่อนไขกำหนดนั้น มีแนวทางดังนี้ สมมติสถานการณ์เมื่อ $n = 20$ และมีค่าผิดปกติ ร้อยละ 5 นั่นคือตัวอย่างชุดนี้จะมีข้อมูลที่เป็นค่าผิดปกติ จำนวน 1 ค่า ให้สุ่มตัวอย่างที่เป็นข้อมูลที่เป็นค่าปกติจำนวน 19 ค่าจากกลุ่มข้อมูลที่มีค่าปกติ (Good) ที่คัดแยกไว้ในข้อ 2.2.2 และสุ่มตัวอย่างที่เป็นข้อมูลค่าผิดปกติจำนวน 1 ค่าจากกลุ่มข้อมูลที่เป็นค่าผิดปกติ (Bad) ที่คัดแยกไว้ในข้อ 2.2.2 ซึ่งสถานการณ์การสุ่มตัวอย่างในเงื่อนไขอื่น ๆ ก็สุ่มตัวอย่างในแนวทางเดียวกันนี้

2.2.6 ทำการจำลองข้อมูลในแต่ละสถานการณ์ ตามข้อ 2.2.5 ทำซ้ำจำนวน 1,000 ครั้ง

2.2.7 คำนวณค่าเฉลี่ยจากตัวประมาณค่าแบบอัตราส่วนตัวใหม่ที่น่าเสนอ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) จากข้อมูลที่ได้จากข้อ 2.2.6 จำนวนซ้ำ 1,000 ครั้ง จากนั้นคำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) จากสูตร

$$MSE(\bar{y}) = \frac{1}{1,000} \sum_{i=1}^{1,000} (\bar{y}_i - \bar{Y})^2$$

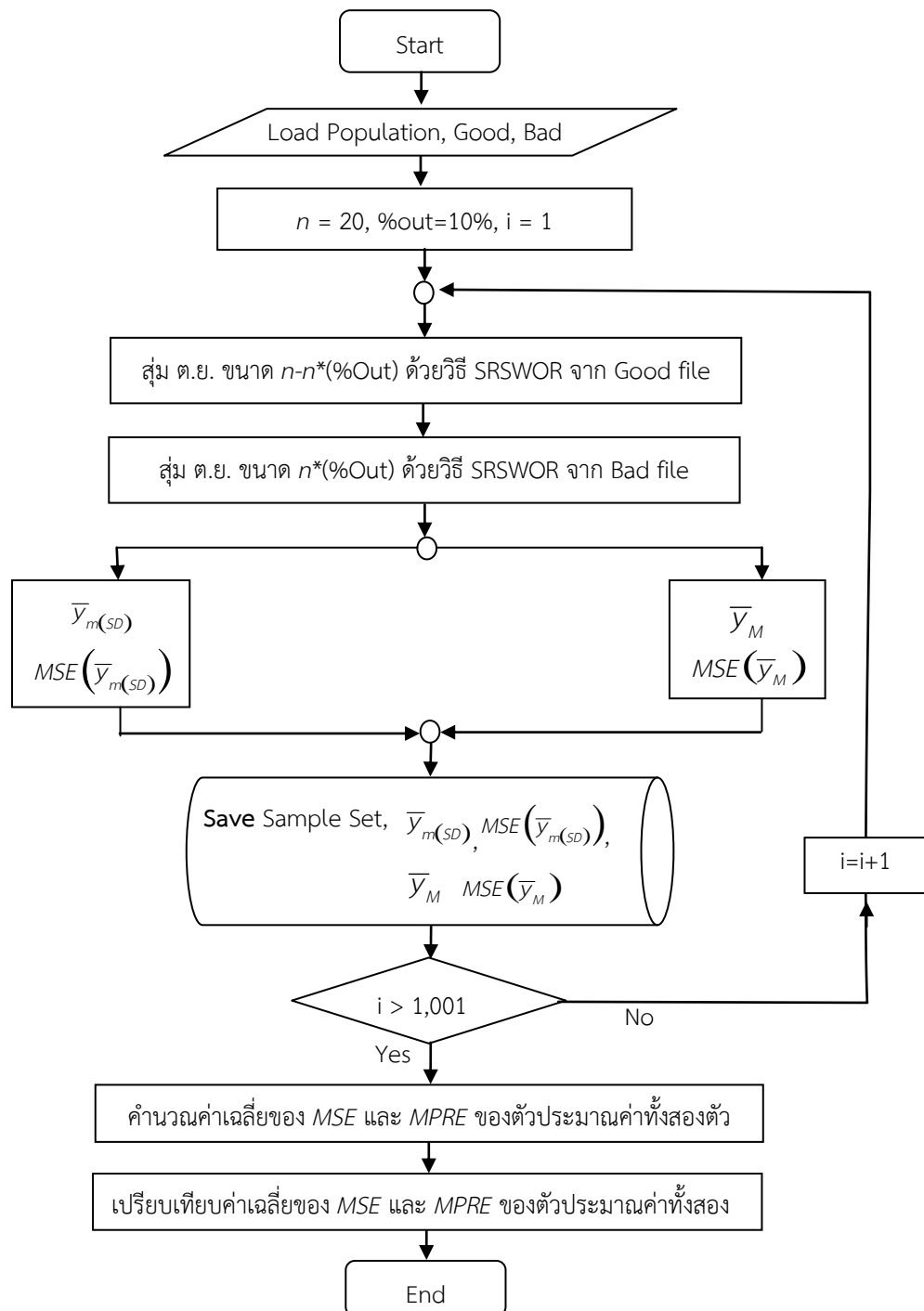
และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์

(MPRE) ของตัวประมาณค่าทั้งสองตัว

สำหรับการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธี LTS ของตัวแปรช่วย X และ Z ในสูตรของตัวประมาณค่า \bar{y}_M นั้น ผู้วิจัยวิเคราะห์โดยใช้คำสั่งฟังก์ชัน ltsregres รันบนโปรแกรม Matlab ซึ่งฟังก์ชันนี้พัฒนาเริ่มแรกโดย Driessen และ Brenkers ในปี ค.ศ. 2004 ต่อมาฟังก์ชันนี้ได้ถูกปรับปรุงพัฒนาขึ้นใหม่โดย Verboven, Engelen and Smets (2006)

2.2.8 ทำการเปรียบเทียบประสิทธิภาพของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่า $\bar{y}_{m(SD)}$ โดยพิจารณาความเที่ยงตรง (Precision) ของตัวประมาณค่า จากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของตัวประมาณค่า (Cochran, 1977, p. 16) และพิจารณาความแม่นยำ (Accuracy) ของตัวประมาณค่า จากค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Mean Percent Relative Error: MPRE) หากค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าใดมีค่าต่ำสุด ก็ถือว่าเป็นตัวประมาณค่าที่มีความเที่ยงตรงสูงกว่า และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ของตัวประมาณค่าใดมีค่าต่ำสุด ก็ถือว่าเป็นตัวประมาณค่าที่มีความแม่นยำสูงกว่า

สำหรับการเขียนโปรแกรมเพื่อทดสอบประสิทธิภาพของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ภายใต้ 12 สถานการณ์ ในแต่ละสถานการณ์จะมีการทำงานคล้ายคลึงกัน ตามข้อ 2.2.1 ถึงข้อ 2.2.8 ดังภาพที่ 3-5



ภาพที่ 3-5 แผนผังการเขียนโปรแกรม 12 สถานการณ์ ยกตัวอย่างกรณี $n = 20$ มีค่าผิดปกติร้อยละ 10

ตอนที่ 3 การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ขั้นตอนนี้เป็นการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อใช้ตัวแปรช่วยสองตัว สำหรับการวิจัยนี้ประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 จำนวน 8 เดือน อาศัยข้อมูลอุตุนิมวิทยารายวันในแต่ละเดือนไปประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน รายละเอียดดังนี้

1. ประชากรที่ศึกษา

ประชากรที่ศึกษาหรือตัวแปรที่ศึกษา (Study Variable) เป็นการระเหยน้ำรายเดือนในจังหวัดสกลนคร จากเดือนพฤษภาคมถึงเดือนธันวาคม พ.ศ. 2556 จำนวน 8 เดือน

ผู้วิจัยประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนทั้ง 8 เดือน ด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ โดยใช้ตัวแปรช่วย (Auxiliary Variables) ที่ผ่านการคัดเลือกตามเกณฑ์ข้อ 1.2.1 ในตอนที่ 2 แล้ว ได้แก่ ตัวแปรอุณหภูมิอากาศและความชื้นสัมพัทธ์

สำหรับการกำหนดค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนนั้น เนื่องด้วยค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในชั้นบรรยากาศไม่สามารถวัดได้โดยตรง และวัดค่าให้ถูกต้องทำได้ยาก (Thompson, 1999, p. 91; วิชัย ศรีบุญลือ, 2552, หน้า 8) จึงอาศัยแนวทางการศึกษาของ Irmak and Haman (2003) และการศึกษาของ Dabral et al. (2014) กำหนดค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนในแต่ละเดือนด้วยค่าเฉลี่ยเฉพาะเดือนนั้น จากข้อมูลอุตุนิมวิทยารายเดือนย้อนหลัง 6 ปี นับเริ่มจากพฤษภาคม พ.ศ. 2550 ถึงเดือนธันวาคม พ.ศ. 2555

2. กลุ่มตัวอย่าง

กลุ่มตัวอย่าง เป็นข้อมูลอุตุนิมวิทยารายวันของจังหวัดสกลนคร ประกอบด้วย การระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยขอความอนุเคราะห์ใช้ข้อมูลอุตุนิมวิทยารายวัน วัดค่าและเก็บบันทึกข้อมูลโดยสำนักงานอุตุนิมวิทยาสกลนคร ข้อมูลอุตุนิมวิทยารายวันในที่นี้เป็นข้อมูลที่เฉลี่ยจาก 8 ช่วงเวลาในแต่ละวัน ประกอบด้วยการระเหยน้ำ (มิลลิเมตร) อุณหภูมิอากาศ ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%)

เมื่อคัดเลือกตัวแปรช่วยตามเกณฑ์ข้อ 1.2.1 ในตอนที่ 2 แล้ว ข้อมูลอุตุนิมวิทยารายวันที่ใช้ในการประมาณค่าการระเหยน้ำโดยเฉลี่ย ประกอบด้วย อุณหภูมิอากาศและความชื้นสัมพัทธ์

4. ขั้นตอนการวิเคราะห์ข้อมูล

หลังจากกำหนดค่าพารามิเตอร์ค่าเฉลี่ยของการระเหยน้ำรายเดือน จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 ด้วยค่าเฉลี่ยเฉพาะเดือนนั้นจากข้อมูลอุตุนิมวิทยารายเดือนย้อนหลัง 6 ปีแล้ว จึงประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน ของเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 จากข้อมูลอุตุนิมวิทยารายวันของแต่ละเดือน โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ที่นำเสนอ (\bar{y}_M) ดัง (3.14)

$$\bar{y}_{M_i} = \hat{y}_i + b_{1_i}^* (\bar{X}_i - \bar{x}_i) + b_{2_i}^* (\bar{Z}_i - \bar{z}_i) \frac{\hat{X}_i + \beta_{1(x)} \hat{Z}_i + \beta_{1(z)} \hat{y}_i}{\bar{X}_i + \beta_{1(x)} \bar{Z}_i + \beta_{1(z)} \bar{y}_i} \quad (3.14)$$

เมื่อ \bar{y}_i คือ ค่าเฉลี่ยการระเหยน้ำของเดือนที่ i

เมื่อ $i = 5, 6, \dots, 12$ หมายถึงเดือนพฤษภาคม มิถุนายน กรกฎาคม สิงหาคม กันยายน ตุลาคม พฤศจิกายน ธันวาคม พ.ศ. 2556 ตามลำดับ

$b_{1_i}^*$ คือ สัมประสิทธิ์การถดถอยด้วยวิธี LTS ของตัวแปรอุณหภูมิอากาศ ของเดือนที่ i

$b_{2_i}^*$ คือ สัมประสิทธิ์การถดถอยด้วยวิธี LTS ของตัวแปรความชื้นสัมพัทธ์ ของเดือนที่ i

\bar{X}_i คือ ค่าเฉลี่ยอุณหภูมิอากาศในระดับประชากร ของเดือนที่ i

\bar{x}_i คือ ค่าเฉลี่ยอุณหภูมิอากาศในระดับตัวอย่าง ของเดือนที่ i

\bar{Z}_i คือ ค่าเฉลี่ยความชื้นสัมพัทธ์ในระดับประชากร ของเดือนที่ i

\bar{z}_i คือ ค่าเฉลี่ยความชื้นสัมพัทธ์ในระดับตัวอย่าง ของเดือนที่ i

$\beta_{1(x)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรอุณหภูมิอากาศ ของเดือนที่ i

$\beta_{1(z)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรความชื้นสัมพัทธ์ ของเดือนที่ i

เมื่อได้ค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนของเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 แล้ว จึงศึกษาความแม่นยำในการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Percent Relative Error: *PRE*) ของค่าประมาณการระเหยน้ำของเดือนใด ๆ สูตรดัง (3.15) และงานวิจัยนี้กำหนดเกณฑ์ความคลาดเคลื่อนที่ยอมรับได้ คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย (Mean Percent Relative Error: *MPRE*) ของค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนทั้ง 8 เดือน มีค่าไม่เกิน 10% (Irmak and Haman, 2003) ซึ่งการคำนวณค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย ดัง (3.16)

$$PRE_i = \left(\left| \bar{y}_{M_i} - \bar{y}_i \right| / \bar{y}_i \right) \cdot 100\% \quad (3.15)$$

$$MPRE = \frac{1}{8} \sum_{i=5}^{12} PRE_i \quad (3.16)$$

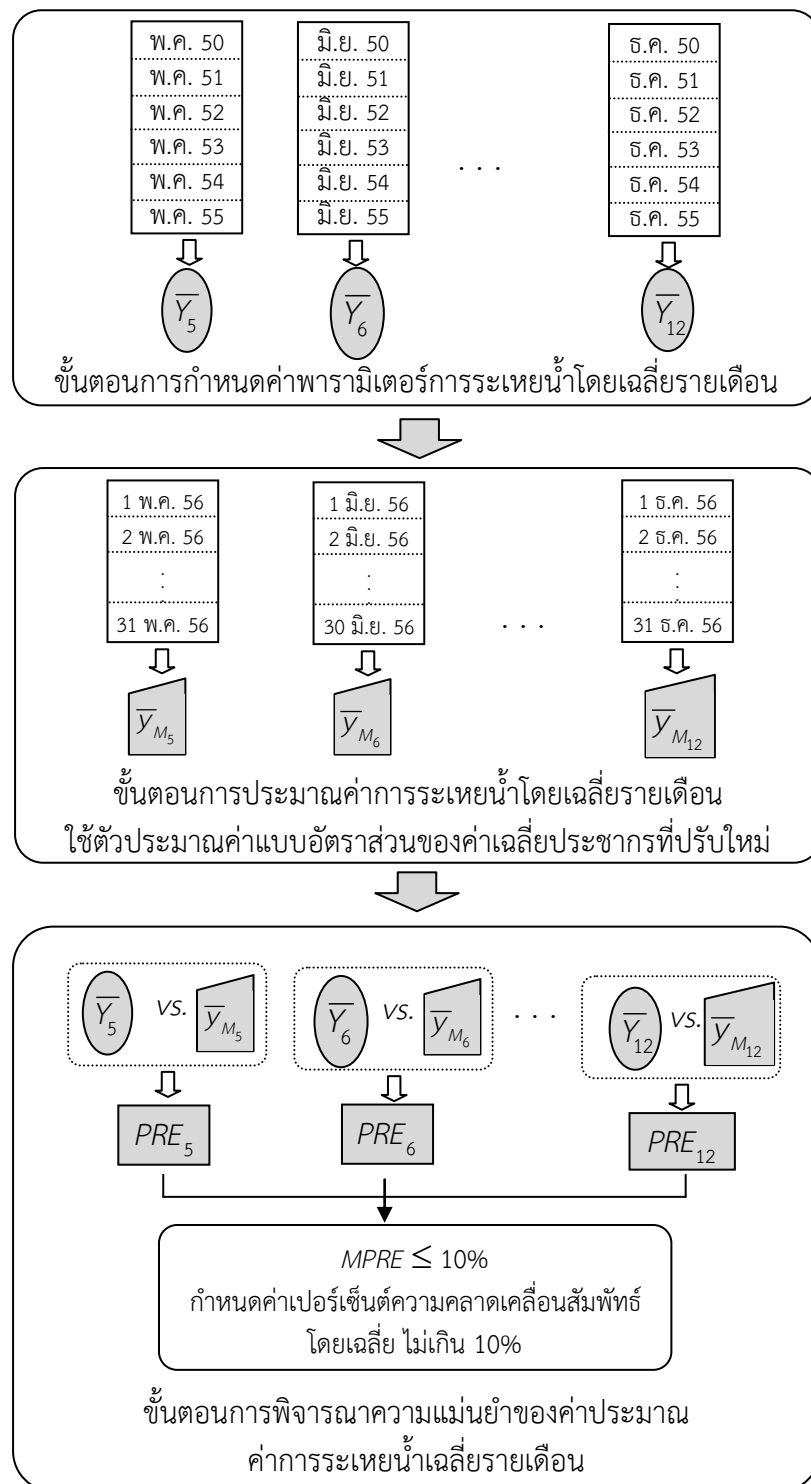
เมื่อ PRE_i คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ของค่าประมาณการระเหยน้ำเดือนที่ i

\bar{y}_{M_i} คือ ค่าประมาณการระเหยน้ำเฉลี่ยรายเดือนของเดือนที่ i ด้วยตัวประมาณค่า \bar{y}_M

\bar{y}_i คือ ค่าพารามิเตอร์ของการระเหยน้ำโดยเฉลี่ยรายเดือนของเดือนที่ i

MPRE คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย

สรุปขั้นตอนการดำเนินการวิจัยในตอนต้นที่ 3 ดังภาพที่ 3-6



ภาพที่ 3-6 สรุปขั้นตอนการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน โดยใช้ตัวประมาณค่า \bar{y}_M

บทที่ 4 ผลการวิจัย

การวิจัย เรื่อง การประมาณค่าการระเหยน้ำโดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ขอเสนอผลการวิเคราะห์ข้อมูลเป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ตอนที่ 2 ผลการจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{Y}_{m(SD)}$

ตอนที่ 3 ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ในการนำเสนอผลส่วนนี้ กำหนดความหมายของคำย่อและสัญลักษณ์ ดังนี้

Min. คือ ค่าต่ำสุดข้อมูลชุดใด ๆ

Max. คือ ค่ามากสุดในข้อมูลชุดใด ๆ

Mean คือ ค่าเฉลี่ยของข้อมูลชุดใด ๆ

SD. คือ ค่าเบี่ยงเบนมาตรฐานของข้อมูลชุดใด ๆ

Skewness คือ ค่าความเบ้ของข้อมูลชุดใด ๆ

Kurtosis คือ ค่าความโด่งของข้อมูลชุดใด ๆ

CV คือ ค่าสัมประสิทธิ์ความแปรผัน

r คือ ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน

\bar{Y}_M คือ ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

$\bar{Y}_{m(SD)}$ คือ ตัวประมาณค่าแบบอัตราส่วนของ Gupta and Shabbir ที่นำเสนอใน

ปี ค.ศ. 2007

b^* คือ ค่าสัมประสิทธิ์การถดถอยแบบด้วยวิธี LTS

MSE คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error)

$MSE(\bar{Y}_M)$ คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า \bar{Y}_M

$MSE(\bar{Y}_{m(SD)})$ คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า $\bar{Y}_{m(SD)}$

p-value คือ ความน่าจะเป็นที่ค่าสถิติทดสอบมีค่ามากหรือน้อยกว่าค่าสถิติทดสอบที่คำนวณได้จากข้อมูล ภายใต้สมมติฐานหลักเป็นจริง

PRE คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Relative Error)

MPRE คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Relative Error)

ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

1. ผลการเสนอตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ผลการศึกษาแนวทางการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อใช้สารสนเทศของตัวแปรช่วยสองตัว เช่น การใช้ค่าสัมประสิทธิ์การแปรผัน ค่าความโค้งหรือค่าความเบ้ของตัวแปรช่วย เป็นต้น จะพบว่า ตัวประมาณค่าแบบอัตราส่วนของ Gupta and Shabbir (2007) นำเสนอไว้ มีประสิทธิภาพมากกว่าตัวประมาณค่าแบบอัตราส่วนของ Sisodia and Dwivedi (1981), Singh and Kakran (1993) และ Upadhyaya and Singh (1999) ซึ่ง Gupta and Shabbir (2007) ได้นำเสนอตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร โดยใช้ตัวแปรช่วยสองตัวดังสมการ (4.1) และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ดังสมการ (4.2)

$$\bar{y}_{m(SD)} = \bar{y} \frac{\frac{\bar{c}_x + C_x \frac{\bar{c}_z + C_z \bar{c}_y}{\bar{c}_y}}{\bar{c}_x + C_x \frac{\bar{c}_z + C_z \bar{c}_y}{\bar{c}_y}}}{\bar{c}_x + C_x \frac{\bar{c}_z + C_z \bar{c}_y}{\bar{c}_y}} \quad (4.1)$$

$$MSE(\bar{y}_{m(SD)}) \gg \bar{y}^2 \lambda \left\{ \frac{\bar{c}_y^2}{\bar{c}_y} + \alpha_1^2 C_x^2 + \alpha_2^2 C_z^2 - 2\alpha_1 \rho_{yx} C_y C_x + 2\alpha_2 \rho_{yz} C_y C_z - 2\alpha_1 \alpha_2 \rho_{xz} C_x C_z \right\} \quad (4.2)$$

แต่ Gupta and Shabbir (2007) เปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนที่นำเสนอกับตัวประมาณค่าของ Sisodia and Dwivedi (1981), Singh and Kakran (1993) และ Upadhyaya and Singh (1999) โดยที่ตัวประมาณค่าแบบอัตราส่วนเหล่านั้นใช้ตัวแปรช่วยไม่เท่ากัน และประกอบกับตัวประมาณค่าแบบอัตราส่วนของ Gupta and Shabbir (2007) ไม่ได้คำนึงการมีค่าผิดปกติปะปนในข้อมูล ดังนั้น งานวิจัยนี้จึงใช้ตัวประมาณค่าแบบอัตราส่วนของ Gupta and Shabbir (2007) เป็นตัวประมาณค่าเริ่มต้น (Initial Estimator) เพื่อพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ให้มีความเหมาะสมกับการประมาณค่าเฉลี่ยประชากรเมื่อข้อมูลมีค่าผิดปกติ ได้ดียิ่งขึ้น

การปรับพจน์ในตัวประมาณค่าแบบอัตราส่วน (4.1) เพื่อพัฒนาให้เป็นตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ที่มีความเหมาะสมกับการวิเคราะห์ข้อมูลที่มีค่าผิดปกติรวมอยู่ ตามที่ Yarmohammadi and Mahmoudvand (2010) ได้แนะนำให้ใช้สถิติการถดถอยแกร่ง เมื่อข้อมูลมีค่าผิดปกติปะปนอยู่ และจำลองข้อมูลเพื่อแสดงให้เห็นว่า ประสิทธิภาพของตัวประมาณค่าด้วยวิธี LTS ใกล้เคียงกับตัวประมาณค่าด้วยวิธี Least Median Square (LMS) และตัวประมาณค่าด้วยวิธี LTS กับวิธี LMS มีประสิทธิภาพสูงกว่าตัวประมาณค่าแบบวิธี Huber-M ดังนั้น งานวิจัยนี้จึงอาศัยการวิเคราะห์การถดถอยแกร่ง (Robust Regression) ที่ใช้ตัวประมาณค่าสัมประสิทธิ์การถดถอยวิธีกำลังสองตัดขอบน้อยสุด (LTS) มาปรับพจน์ \bar{y} ในสมการ (4.1) ด้วยพจน์ $\bar{y} + b_1^* (\bar{X} - \bar{x}) + b_2^* (\bar{Z} - \bar{z})$ และอาศัยทิศทางการใช้สารสนเทศของตัวแปรช่วยแบบใหม่ มาปรับสูตรให้เหมาะสมกับข้อมูลที่มีความเบ้ ดังการศึกษาของ ภทรรรณ แสงนวกิจ (2553) และ Yan and Tian (2010) นั่นคือ การใช้สัมประสิทธิ์ความเบ้ของตัวแปรช่วยปรับสูตร (4.1) โดยแทนพจน์

สัมประสิทธิ์การแปรผันของตัวแปรช่วย X และตัวแปรช่วย Z (C_x, C_z) ด้วยสัมประสิทธิ์ความเบ้ของตัวแปรช่วย X และตัวแปรช่วย Z ($\beta_{1(x)}, \beta_{1(z)}$) ตามลำดับ

ด้วยเหตุผลที่กล่าวข้างต้น ผู้วิจัยจึงนำเสนอตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ดังสมการ (4.3)

$$\bar{y}_M = \frac{\bar{y} + b_1^*(\bar{X} - \bar{x}) + b_2^*(\bar{Z} - \bar{z})}{\frac{\bar{y}}{\bar{y}} + \beta_{1(x)} \frac{\bar{X} - \bar{x}}{\bar{X}} + \beta_{1(z)} \frac{\bar{Z} - \bar{z}}{\bar{Z}}} \quad (4.3)$$

เมื่อ $\beta_{1(x)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย X

$\beta_{1(z)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย Z

b_1^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย X

b_2^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย Z

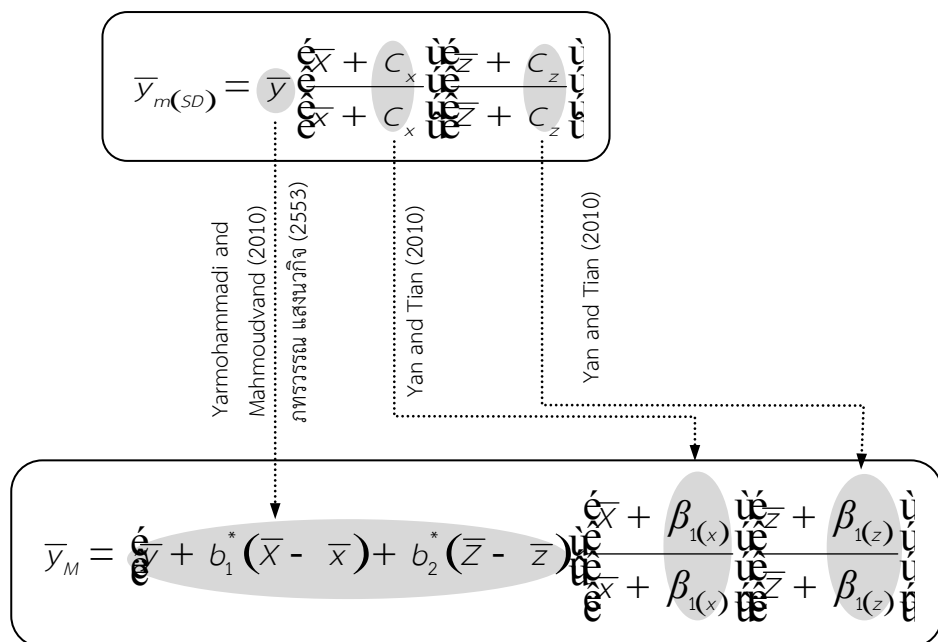
\bar{X} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย X

\bar{Z} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย Z

\bar{x} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย X

\bar{z} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย Z

สรุปการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) จากตัวประมาณค่าเริ่มต้นของ Gupta and Shabbir (2007) ดังภาพที่ 4-1



ภาพที่ 4-1 การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M)

2. ผลการหาสมการ MSE ของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ผู้วิจัยนำเสนอตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ดังสมการ 4.3 และอาศัยทฤษฎีบทเทอร์เลอร์ โดยการกระจายพจน์โพลิโนเมียลของอนุกรมเทอร์เลอร์ดีกรีอันดับหนึ่ง ดังการศึกษาของภทรวรรณ แสงนวกิจ (2553), Abu-Dayyeh et al. (2003), Kadilar and Cingi (2004; 2005; 2006; 2007), และ Gupta and Shabbir (2007) เพื่อหาสมการ MSE ของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ดังนี้

กำหนดให้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) และค่าเฉลี่ยประชากรที่แท้จริงหรือค่าพารามิเตอร์ (\bar{Y}) เป็นฟังก์ชันของตัวแปรศึกษา (Y) ตัวแปรช่วย X และตัวแปรช่วย Z ได้ดังสมการ (4.4) และ (4.5)

$$\begin{aligned}\bar{y}_M &= h(\bar{x}, \bar{z}, \bar{y}) \\ &= \frac{\bar{y}}{\bar{x}} + b_1^* (\bar{x} - \bar{x}) + b_2^* (\bar{z} - \bar{z}) \frac{\bar{x} + \beta_{1(x)} \frac{\bar{z}}{\bar{x}} + \beta_{1(z)} \frac{\bar{y}}{\bar{x}}}{\bar{x} + \beta_{1(x)} \frac{\bar{z}}{\bar{x}} + \beta_{1(z)} \frac{\bar{y}}{\bar{x}}}\end{aligned}\quad (4.4)$$

$$\bar{Y} = h(\bar{X}, \bar{Z}, \bar{Y}) \quad (4.5)$$

พิจารณากระจายอนุกรมเทอร์เลอร์ดีกรีอันดับหนึ่ง เพื่อหาสมการ MSE ของตัวประมาณค่า \bar{y}_M จะได้การกระจายอนุกรมเทอร์เลอร์ดีกรีอันดับหนึ่งของสมการ (4.4) ดังนี้

$$\begin{aligned}h(\bar{x}, \bar{z}, \bar{y}) &\cong h(\bar{X}, \bar{Z}, \bar{Y}) + \frac{\left| \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left| \bar{x} \right|} (\bar{x} - \bar{X}) \\ &+ \frac{\left| \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{z}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left| \bar{z} \right|} (\bar{z} - \bar{Z}) + \frac{\left| \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{y}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left| \bar{y} \right|} (\bar{y} - \bar{Y})\end{aligned}\quad (4.6)$$

จาก (4.6) ต้องจัดพจน์สมการให้อยู่ในรูป $E(\bar{y}_M - \bar{Y})^2$ ซึ่ง $MSE(\bar{y}_M) = E(\bar{y}_M - \bar{Y})^2$

ดังนั้น พิจารณาค่าจากพจน์ $\frac{\left| \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left| \bar{x} \right|} (\bar{x} - \bar{X})$ ในสมการ (4.6) จะได้ว่า

$$\begin{aligned}&\frac{\left| \frac{\partial h(\bar{x}, \bar{z}, \bar{y})}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left| \bar{x} \right|} (\bar{x} - \bar{X}) \\ &= \frac{\left| \frac{\partial}{\partial \bar{x}} \left[\frac{\bar{y}}{\bar{x}} + b_1^* (\bar{x} - \bar{x}) + b_2^* (\bar{z} - \bar{z}) \frac{\bar{x} + \beta_{1(x)} \frac{\bar{z}}{\bar{x}} + \beta_{1(z)} \frac{\bar{y}}{\bar{x}}}{\bar{x} + \beta_{1(x)} \frac{\bar{z}}{\bar{x}} + \beta_{1(z)} \frac{\bar{y}}{\bar{x}}} \right]}{\partial \bar{x}} \right|_{\bar{x}, \bar{z}, \bar{y}}}{\left| \bar{x} \right|} (\bar{x} - \bar{X})\end{aligned}$$

$$\begin{aligned}
 &= \frac{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{y} + b_1^* (\bar{x} - \bar{x}) + b_2^* (\bar{z} - \bar{z})}{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} \frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{y} \\
 &\quad + \frac{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{y} + b_1^* (\bar{x} - \bar{x}) + b_2^* (\bar{z} - \bar{z})}{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} (\bar{x} - \bar{x}) \\
 &= \frac{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{y} + b_1^* (\bar{x} - \bar{x}) + b_2^* (\bar{z} - \bar{z})}{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} \frac{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{y}}{\left(\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{x} \right)^2} \\
 &\quad + \frac{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{y} - b_1^*}{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} (\bar{x} - \bar{x}) \\
 &= (1) \frac{\frac{\partial}{\partial \bar{z}} \bar{y} + b_1^* (0) + b_2^* (0)}{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} - 1 \frac{0}{\left(\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{x} \right)^2} + (-b_1^*) (\bar{x} - \bar{x}) \\
 &= \frac{\frac{\partial}{\partial \bar{z}} \bar{y} - \bar{y}}{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} - b_1^* (\bar{x} - \bar{x}) \tag{4.7}
 \end{aligned}$$

พิจารณาค่าพจน์ $\frac{\partial}{\partial \bar{z}} h(\bar{x}, \bar{z}, \bar{y}) \Big|_{\bar{x}, \bar{z}, \bar{y}}$ ($\bar{z} - \bar{z}$) ในสมการ (4.6) จะได้ว่า

$$\begin{aligned}
 &\frac{\partial}{\partial \bar{z}} h(\bar{x}, \bar{z}, \bar{y}) \Big|_{\bar{x}, \bar{z}, \bar{y}} (\bar{z} - \bar{z}) \\
 &= \frac{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{y} + b_1^* (\bar{x} - \bar{x}) + b_2^* (\bar{z} - \bar{z})}{\frac{\partial}{\partial \bar{z}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(z)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} \frac{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{y}}{\frac{\partial}{\partial \bar{x}} \left(\frac{\partial}{\partial \bar{z}} + \beta_{1(x)} \frac{\partial}{\partial \bar{x}} \right) \bar{x}} (\bar{z} - \bar{z})
 \end{aligned}$$

$$\begin{aligned}
&= \frac{\frac{\sum_{i=1}^n \hat{e}_i \bar{x}_i + \beta_{1(x)} \sum_{i=1}^n \hat{e}_i \bar{z}_i + \beta_{1(z)} \sum_{i=1}^n \hat{e}_i \bar{y}_i}{\sum_{i=1}^n \hat{e}_i \bar{x}_i + \beta_{1(x)} \sum_{i=1}^n \hat{e}_i \bar{z}_i + \beta_{1(z)} \sum_{i=1}^n \hat{e}_i \bar{y}_i}}{\left. \vphantom{\frac{\sum_{i=1}^n \hat{e}_i \bar{x}_i + \beta_{1(x)} \sum_{i=1}^n \hat{e}_i \bar{z}_i + \beta_{1(z)} \sum_{i=1}^n \hat{e}_i \bar{y}_i}{\sum_{i=1}^n \hat{e}_i \bar{x}_i + \beta_{1(x)} \sum_{i=1}^n \hat{e}_i \bar{z}_i + \beta_{1(z)} \sum_{i=1}^n \hat{e}_i \bar{y}_i}} \right|_{\bar{x}, \bar{z}, \bar{y}}} (\bar{y} - \bar{Y}) \\
&= (\bar{y} - \bar{Y}) \tag{4.9}
\end{aligned}$$

นำค่าที่ได้จากสมการ (4.7)–(4.9) แทนค่าในสมการ (4.4) จะได้ว่า

$$\begin{aligned}
h(\bar{x}, \bar{z}, \bar{y}) &\cong h(\bar{X}, \bar{Z}, \bar{Y}) + \frac{\sum_{i=1}^n \hat{e}_i \bar{x}_i - \bar{Y}}{\sum_{i=1}^n \hat{e}_i \bar{x}_i + \beta_{1(x)}} - b_1^* \frac{\sum_{i=1}^n \hat{e}_i \bar{x}_i}{\sum_{i=1}^n \hat{e}_i \bar{x}_i} (\bar{x} - \bar{X}) \\
&\quad + \frac{\sum_{i=1}^n \hat{e}_i \bar{z}_i - \bar{Y}}{\sum_{i=1}^n \hat{e}_i \bar{z}_i + \beta_{1(z)}} - b_2^* \frac{\sum_{i=1}^n \hat{e}_i \bar{z}_i}{\sum_{i=1}^n \hat{e}_i \bar{z}_i} (\bar{z} - \bar{Z}) + (\bar{y} - \bar{Y}) \tag{4.10}
\end{aligned}$$

จึงจัดพจน์ในสมการ (4.10) ให้มีความกระชับในการนำไปใช้งานจริงมากขึ้น

โดยกำหนดให้ $R_1 = \frac{\bar{Y}}{\bar{X} + \beta_{1(x)}}$ และ $R_2 = \frac{\bar{Y}}{\bar{Z} + \beta_{1(z)}}$ แทนค่าใน (4.10) จะได้

$$h(\bar{x}, \bar{z}, \bar{y}) \cong h(\bar{X}, \bar{Z}, \bar{Y}) - (b_1^* + R_1)(\bar{x} - \bar{X}) - (b_2^* - R_2)(\bar{z} - \bar{Z}) + (\bar{y} - \bar{Y}) \tag{4.11}$$

จากนั้นแทนค่า $\bar{y}_M = h(\bar{x}, \bar{z}, \bar{y})$ และ $\bar{Y} = h(\bar{X}, \bar{Z}, \bar{Y})$ ใน (4.11) แล้วจัดสมการ (4.11) ให้อยู่ในรูป $E(\bar{y}_M - \bar{Y})^2$ เพื่อหาสมการ $MSE(\bar{y}_M) = E(\bar{y}_M - \bar{Y})^2$ ได้ดังนี้

$$\begin{aligned}
\bar{y}_M &\cong \bar{Y} - (b_1^* + R_1)(\bar{x} - \bar{X}) - (b_2^* - R_2)(\bar{z} - \bar{Z}) + (\bar{y} - \bar{Y}) \\
\bar{y}_M - \bar{Y} &\cong (\bar{y} - \bar{Y}) - (b_1^* + R_1)(\bar{x} - \bar{X}) - (b_2^* - R_2)(\bar{z} - \bar{Z}) \\
(\bar{y}_M - \bar{Y})^2 &\cong \{(\bar{y} - \bar{Y}) - (b_1^* + R_1)(\bar{x} - \bar{X}) - (b_2^* - R_2)(\bar{z} - \bar{Z})\}^2 \\
&\cong (\bar{y} - \bar{Y})^2 + (b_1^* + R_1)^2 (\bar{x} - \bar{X})^2 + (b_2^* - R_2)^2 (\bar{z} - \bar{Z})^2 \\
&\quad - 2(b_1^* + R_1)(\bar{x} - \bar{X})(\bar{y} - \bar{Y}) - 2(b_2^* - R_2)(\bar{y} - \bar{Y})(\bar{z} - \bar{Z}) \\
&\quad + 2(b_1^* + R_1)(b_2^* - R_2)(\bar{x} - \bar{X})(\bar{z} - \bar{Z})
\end{aligned}$$

$$\begin{aligned}
E(\bar{y}_M - \bar{Y})^2 &\cong E(\bar{y} - \bar{Y})^2 + E(b_1^* + R_1)^2 (\bar{x} - \bar{X})^2 + E(b_2^* - R_2)^2 (\bar{z} - \bar{Z})^2 \\
&\quad - 2E(b_1^* + R_1)(\bar{x} - \bar{X})(\bar{y} - \bar{Y}) - 2E(b_2^* - R_2)(\bar{y} - \bar{Y})(\bar{z} - \bar{Z}) \\
&\quad + 2E(b_1^* + R_1)(b_2^* - R_2)(\bar{x} - \bar{X})(\bar{z} - \bar{Z})
\end{aligned}$$

เมื่อสุ่มตัวอย่างขนาดใหญ่ จึงกำหนดให้ $E(b_1^*) = B_1^*$ และ $E(b_2^*) = B_2^*$ ซึ่งจะทำให้ความแปรปรวนของตัวประมาณค่ามีค่าต่ำสุด (Cochran, 1977, pp. 193-194) จะได้ว่า

$$\begin{aligned}
E(\bar{y}_M - \bar{Y})^2 &\cong E(\bar{y} - \bar{Y})^2 + (B_1^* + R_1)^2 E(\bar{x} - \bar{X})^2 + (B_2^* - R_2)^2 E(\bar{z} - \bar{Z})^2 \\
&\quad - 2(B_1^* + R_1)E[(\bar{x} - \bar{X})(\bar{y} - \bar{Y})] - 2(B_2^* - R_2)E[(\bar{y} - \bar{Y})(\bar{z} - \bar{Z})] \\
&\quad + 2(B_1^* + R_1)(B_2^* - R_2)E[(\bar{x} - \bar{X})(\bar{z} - \bar{Z})]
\end{aligned}$$

$$\begin{aligned}
E(\bar{y}_M - \bar{Y})^2 &\cong V(\bar{y}) + (B_1^* + R_1)^2 V(\bar{x}) + (B_2^* - R_2)^2 V(\bar{z}) \\
&\quad - 2(B_1^* + R_1)\text{Cov}(\bar{x}, \bar{y}) - 2(B_2^* - R_2)\text{Cov}(\bar{y}, \bar{z}) \\
&\quad + 2(B_1^* + R_1)(B_2^* - R_2)\text{Cov}(\bar{x}, \bar{z}) \\
&\cong \frac{1-f}{n} S_y^2 + (B_1^* + R_1)^2 \frac{1-f}{n} S_x^2 + (B_2^* - R_2)^2 \frac{1-f}{n} S_z^2 \\
&\quad - 2(B_1^* + R_1) \frac{1-f}{n} S_{yx} - 2(B_2^* - R_2) \frac{1-f}{n} S_{yz} \\
&\quad + 2(B_1^* + R_1)(B_2^* - R_2) \frac{1-f}{n} S_{xz} \\
&\cong \frac{1-f}{n} S_y^2 + (B_1^* + R_1)^2 S_x^2 + (B_2^* - R_2)^2 S_z^2 \\
&\quad - 2(B_1^* + R_1)S_{yx} - 2(B_2^* - R_2)S_{yz} \\
&\quad + 2(B_1^* + R_1)(B_2^* - R_2)S_{xz}
\end{aligned}$$

ดังนั้น สมการ MSE ของตัวประมาณค่า \bar{y}_M คือ

$$\begin{aligned}
MSE(\bar{y}_M) &\cong \frac{1-f}{n} S_y^2 + (B_1^* + R_1)^2 S_x^2 + (B_2^* - R_2)^2 S_z^2 \\
&\quad - 2(B_1^* + R_1)S_{yx} - 2(B_2^* - R_2)S_{yz} + 2(B_1^* + R_1)(B_2^* - R_2)S_{xz} \quad (4.12)
\end{aligned}$$

สรุปผลการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ได้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ดังสมการ (4.13)

$$\bar{y}_M = \frac{\sum_{i=1}^N \bar{y}_i + \beta_{1(x)} \sum_{i=1}^N \bar{x}_i + \beta_{1(z)} \sum_{i=1}^N \bar{z}_i}{\sum_{i=1}^N \bar{y}_i + \beta_{1(x)} \sum_{i=1}^N \bar{x}_i + \beta_{1(z)} \sum_{i=1}^N \bar{z}_i} \quad (4.13)$$

เมื่อ $\beta_{1(x)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย X

$\beta_{1(z)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย Z

b_1^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย X

b_2^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย Z

\bar{X} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย X

\bar{Z} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย Z

\bar{x} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย X

\bar{z} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย Z

และมีสมการ MSE ดังสมการ (4.14)

$$MSE(\bar{y}_M) \cong \frac{1-f}{n} \sum_{i=1}^N \bar{e}_i^2 + (B_1^* + R_1)^2 S_x^2 + (B_2^* - R_2)^2 S_z^2 - 2(B_1^* + R_1)S_{yx} - 2(B_2^* - R_2)S_{yz} + 2(B_1^* + R_1)(B_2^* - R_2)S_{xz} \quad (4.14)$$

โดยที่ $R_1 = \frac{\bar{y}}{\bar{x} + \beta_{1(x)}}, R_2 = \frac{\bar{y}}{\bar{z} + \beta_{1(z)}}, f = \frac{n}{N},$

$$S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2,$$

$$S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2,$$

$$S_z^2 = \frac{1}{N-1} \sum_{i=1}^N (Z_i - \bar{Z})^2,$$

$$S_{xy} = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y}),$$

$$S_{zy} = \frac{1}{N-1} \sum_{i=1}^N (Z_i - \bar{Z})(Y_i - \bar{Y}),$$

$$S_{zx} = \frac{1}{N-1} \sum_{i=1}^N (Z_i - \bar{Z})(X_i - \bar{X}) \text{ ตามลำดับ}$$

ตอนที่ 2 ผลการจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$

ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน ($\bar{y}_{m(SD)}$) ของ Gupta and Shabbir (2007) นั้น ขอนำเสนอผลการศึกษาลักษณะเบื้องต้นของข้อมูลอุตุนิยมหาวิทยาลัยที่เกี่ยวข้องหรือมีความสัมพันธ์กับการระเหยน้ำแบบลาดก่อน เพื่อให้ทราบลักษณะการเกิดข้อมูลการระเหยน้ำและปัจจัยอื่นที่เกี่ยวข้อง จากนั้นจึงนำเสนอผลการแสดงเชิงตัวเลขด้วยวิธีการจำลองข้อมูลโดยเทคนิคมอนติคาร์โล เพื่อเปรียบเทียบประสิทธิภาพของการประมาณค่าเฉลี่ยประชากรระหว่างตัวประมาณค่าแบบอัตราส่วนที่ปรับใหม่กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ภายใต้สถานการณ์ 12 สถานการณ์

1. ผลการศึกษาลักษณะข้อมูลเบื้องต้น (Pilot Study)

การวิจัยนี้ใช้ข้อมูลอุตุนิยมหาวิทยาลัยรายเดือน ประกอบด้วยการระเหยน้ำแบบลาด (มิลลิเมตร) อุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%) จากเดือนมกราคม 2550 ถึงเดือนเมษายน 2556 จำนวน 76 เดือน ที่เก็บบันทึกข้อมูลโดยสำนักงานอุตุนิยมหาวิทยาลัยจังหวัดสกลนคร ข้อมูลอุตุนิยมหาวิทยาลัยรายเดือนมีลักษณะข้อมูลเบื้องต้น อธิบายด้วยค่าสถิติพื้นฐาน แสดงดังตารางที่ 4-1

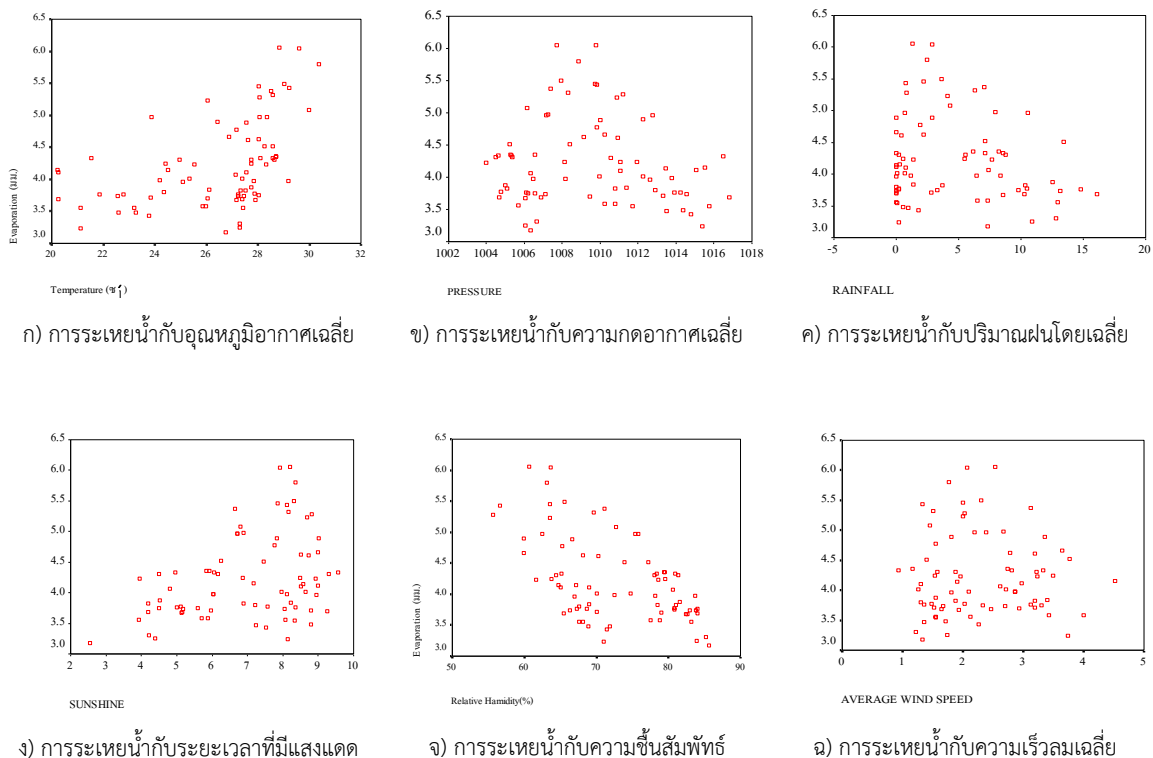
ตารางที่ 4-1 ลักษณะข้อมูลเบื้องต้นของข้อมูลอุตุนิยมหาวิทยาลัยรายเดือนจำนวน 76 เดือน

ข้อมูลอุตุนิยมหาวิทยาลัยรายเดือน (หน่วยวัด)	ลักษณะข้อมูลเบื้องต้น					
	Min.	Max.	Mean	SD.	Skewness	Kurtosis
1) การระเหยน้ำแบบลาด (มิลลิเมตร)	3.17	6.05	4.23	0.68	0.86	0.11
2) อุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$)	20.22	30.38	26.38	2.52	-0.96	0.04
3) ความกดอากาศเฉลี่ย (hPa)	1004.00	1016.80	1009.55	4.52	0.28	-1.03
4) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร)	0.00	16.06	4.56	4.52	0.73	-0.62
5) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง)	2.56	9.55	6.98	1.69	-0.52	-0.80
6) ความชื้นสัมพัทธ์ (%)	55.62	85.67	72.54	7.94	-0.06	-1.15

จากตารางที่ 4.1 พบว่า การระเหยน้ำแบบลาดมีค่าอยู่ระหว่าง 3.17 ถึง 6.05 มิลลิเมตร มีค่าเฉลี่ย 4.23 มิลลิเมตร ค่าเบี่ยงเบนมาตรฐาน 0.68 มิลลิเมตร มีความเบ้เท่ากับ 0.86 ซึ่งเป็น

ลักษณะเบาเล็กน้อย มีความโด่งเท่ากับ 0.11 อุณหภูมิอากาศเฉลี่ยมีค่าอยู่ระหว่าง 20.22 ถึง 30.38 °C มีค่าเฉลี่ย 26.38 °C ค่าเบี่ยงเบนมาตรฐาน 2.52 °C ความกดอากาศเฉลี่ยมีค่าอยู่ระหว่าง 1004.00 ถึง 1016.80 hPa มีค่าเฉลี่ย 1009.55 hPa ค่าเบี่ยงเบนมาตรฐาน 4.52 hPa ปริมาณฝนโดยเฉลี่ยมีค่าไม่เกิน 16.06 มิลลิเมตร มีค่าเฉลี่ย 4.56 มิลลิเมตร ค่าเบี่ยงเบนมาตรฐาน 4.52 มิลลิเมตร ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ยอยู่ในช่วง 2.56 ถึง 9.55 ชั่วโมง ค่าเฉลี่ย 6.98 ชั่วโมง ค่าเบี่ยงเบนมาตรฐาน 1.69 ชั่วโมง ส่วนความชื้นสัมพัทธ์มีค่าอยู่ในช่วง 55.62% ถึง 85.67% มีค่าเฉลี่ย 7.94% ค่าเบี่ยงเบนมาตรฐาน 7.94%

จากนั้นพิจารณาความสัมพันธ์รายคู่คร่าว ๆ ด้วยแผนภาพการกระจาย (Scatter Plot) ระหว่างการระเหยน้ำแบบภาคกับตัวแปรอุตุนิยมวิทยาอื่น จำนวน 5 ตัว ซึ่งเป็นตัวแปรที่จะทำการคัดเลือกเพื่อเป็นตัวแปรช่วยในการประมาณค่าการระเหยน้ำต่อไป ดังภาพที่ 4-2



ภาพที่ 4-2 แผนภาพการกระจายระหว่างการระเหยน้ำแบบภาคกับตัวแปรอุตุนิยมวิทยาอื่น

ก) อุณหภูมิอากาศเฉลี่ย ข) ความกดอากาศเฉลี่ย ค) ปริมาณฝนโดยเฉลี่ย

ง) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย จ) ความชื้นสัมพัทธ์ และ ฉ) ความเร็วลมเฉลี่ย

จากภาพที่ 4-2 เมื่อพิจารณาแผนภาพการกระจายระหว่างการระเหยน้ำแบบภาคกับตัวแปรอุตุนิยมวิทยาอื่นทั้ง 5 ตัว พบว่า อุณหภูมิอากาศเฉลี่ยกับความชื้นสัมพัทธ์ มีความสัมพันธ์เชิงเส้น กว่าการระเหยน้ำชัดเจนกว่าตัวแปรอื่น ๆ จึงคัดเลือกตัวแปรช่วยสองตัวแรกที่มีความสัมพันธ์กับ

การระเหยน้ำแบบลาดมากที่สุด ในที่นี้ พิจารณาเลือกตัวแปรช่วยจากตัวแปรอุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$) ความกดอากาศเฉลี่ย (hPa) ปริมาณฝนโดยเฉลี่ย (มิลลิเมตร) ระยะเวลาที่มีแสงแดดในวันโดยเฉลี่ย (ชั่วโมง) และความชื้นสัมพัทธ์ (%) โดยวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างตัวแปรช่วยใด ๆ กับการระเหยน้ำแบบลาด ผลดังตารางที่ 4-2

ตารางที่ 4-2 ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างตัวแปรช่วยกับการระเหยน้ำแบบลาด

ข้อมูลอุตุนิยมหาวิทยาลัย รายเดือน	ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน					
	EVAPO	TEMP	PRESSURE	RAINFALL	SUNSHINE	HUMIDITY
1) การระเหยน้ำแบบลาด (EVAPO)	1.00	0.48**	-0.13	-0.18	0.37**	-0.55**
2) อุณหภูมิอากาศเฉลี่ย (TEMP)		1.00	-0.84**	0.52**	-0.40**	0.32*
3) ความกดอากาศเฉลี่ย (PRESSURE)			1.00	-0.80**	0.76**	-0.67**
4) ปริมาณฝนโดยเฉลี่ย (RAINFALL)				1.00	-0.82**	0.80**
5) ระยะเวลาที่มีแสงแดด ในวันโดยเฉลี่ย (SUNSHINE)					1.00	-0.87**
6) ความชื้นสัมพัทธ์ (HUMIDITY)						1.00

หมายเหตุ ** หมายถึง มีนัยสำคัญทางสถิติที่ระดับ 0.01 เมื่อทดสอบด้วยสถิติ t-Test

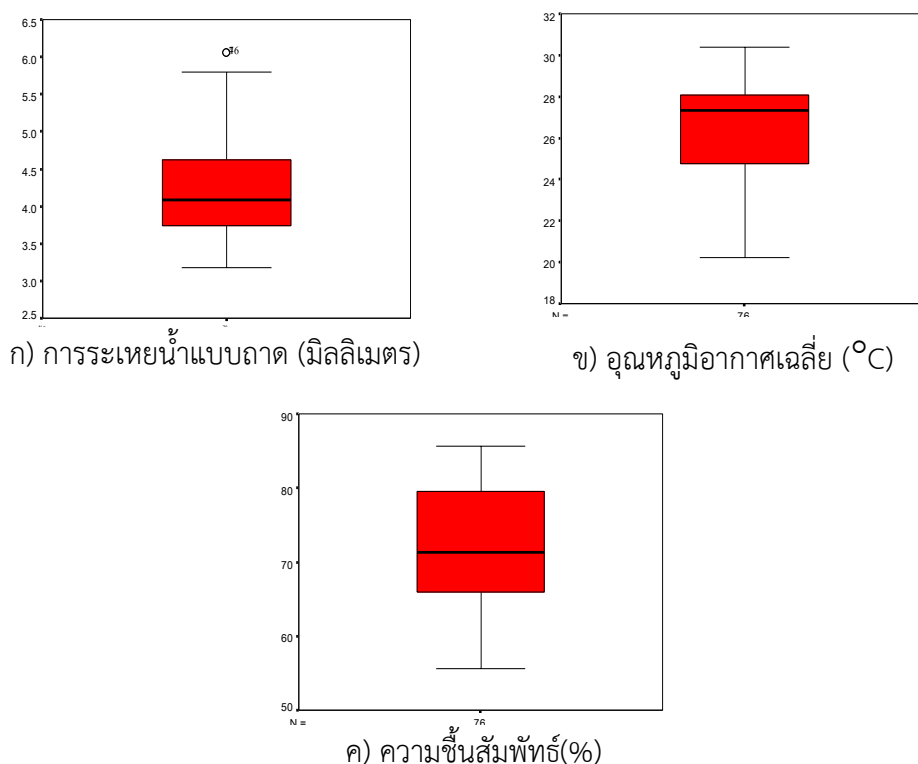
* หมายถึง มีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อทดสอบด้วยสถิติ t-Test

จากตารางที่ 4-2 พบว่า ตัวแปรสองตัวแรกที่มีความสัมพันธ์สูงสุดกับการระเหยน้ำแบบลาด ได้แก่ ความชื้นสัมพัทธ์ (HUMIDITY) ที่มีค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างความชื้นสัมพัทธ์กับการระเหยน้ำแบบลาด เท่ากับ -0.55 รองลงมาคือ อุณหภูมิอากาศเฉลี่ย (TEMP) ที่มีค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างอุณหภูมิอากาศเฉลี่ยกับการระเหยน้ำแบบลาด เท่ากับ 0.48 และพบว่า ตัวแปรความชื้นสัมพัทธ์กับอุณหภูมิอากาศเฉลี่ยมีความสัมพันธ์กัน วัดค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันได้เท่ากับ 0.32

ถึงแม้ว่า ตัวแปรความชื้นสัมพัทธ์กับอุณหภูมิอากาศเฉลี่ยจะมีความสัมพันธ์กันเอง แต่ความสัมพันธ์ระหว่างความชื้นสัมพัทธ์กับอุณหภูมิอากาศเฉลี่ยนั้นมีค่าเท่ากับ 0.32 ซึ่งไม่เกิน 0.7 และค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างตัวแปรช่วยสองตัวนั้นมีค่าน้อยกว่าความสัมพันธ์ระหว่างความชื้นสัมพัทธ์กับการระเหยน้ำรายเดือน ($r = -0.55$) และความสัมพันธ์ระหว่างอุณหภูมิอากาศเฉลี่ยกับการระเหยน้ำรายเดือน ($r = 0.48$) จึงเป็นไปตามเกณฑ์การป้องกันการเกิดปัญหาตัวแปรช่วยมีความสัมพันธ์กันเอง หรือที่เรียกว่าปัญหา Collinearity (Hair et al., 2010, p. 205)

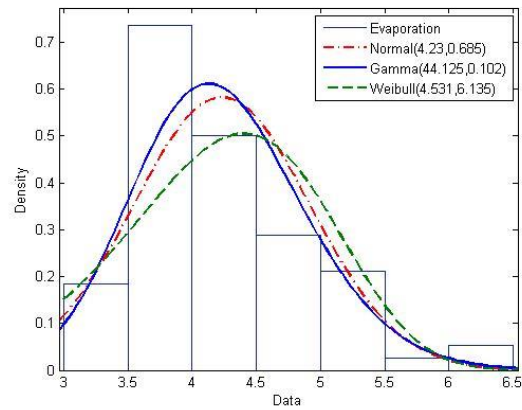
ดังนั้น งานวิจัยนี้จึงคัดเลือกตัวแปรอุณหภูมิอากาศเฉลี่ย (TEMP) และความชื้นสัมพัทธ์ (HUMIDITY) เป็นตัวแปรช่วยในการประมาณค่าการระเหยน้ำแบบถาวร

เมื่อคัดเลือกอุณหภูมิอากาศเฉลี่ย (TEMP) และความชื้นสัมพัทธ์ (HUMIDITY) เป็นตัวแปรช่วยสองตัวในการประมาณค่าการระเหยน้ำแล้ว จึงทำการตรวจสอบค่าผิดปกติและลักษณะการแจกแจงของข้อมูลการระเหยน้ำแบบถาวร อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ ด้วยแผนภาพกล่อง (Boxplot) ดังภาพที่ 4-3

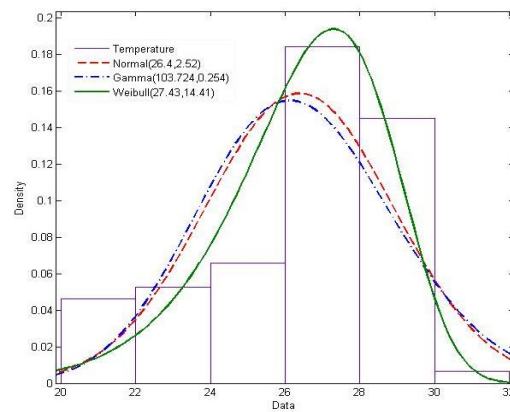


ภาพที่ 4-3 แผนภาพกล่องแสดงการกระจายของข้อมูล ก) การระเหยน้ำแบบถาวร ข) อุณหภูมิอากาศเฉลี่ย และ ค) ความชื้นสัมพัทธ์

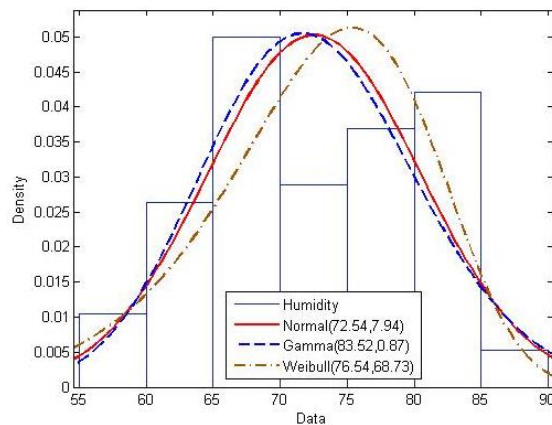
จากภาพที่ 4-3 จะเห็นว่าข้อมูลการระเหยน้ำแบบถาวร มีค่าเฉลี่ย 4.23 มิลลิเมตร (ต่อหน่วยเวลา) ค่าเบี่ยงเบนมาตรฐาน 0.68 มิลลิเมตร (ต่อหน่วยเวลา) มีลักษณะเบ้ขวา และมีค่ามากผิดปกติจำนวน 2 ค่า สำหรับข้อมูลอุณหภูมิอากาศ ค่อนข้างมีลักษณะเบ้ซ้ายเล็กน้อย และข้อมูลความชื้นสัมพัทธ์มีลักษณะเกือบสมมาตรเนื่องจากเบ้ขวาเพียงเล็กน้อย เมื่อนำข้อมูลการระเหยน้ำแบบถาวร อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ สร้างกราฟฮิสโทแกรมเพื่อเปรียบเทียบการแจกแจงความน่าจะเป็นของข้อมูลกับการแจกแจงปกติ การแจกแจงแกมมา และการแจกแจงไวบูลล์ ดังภาพที่ 4-4



ก) การระเหยน้ำแบบสภาพ (มิลลิเมตร)



ข) อุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$)



ค) ความชื้นสัมพัทธ์(%)

ภาพที่ 4-4 กราฟฮิสโทแกรมของข้อมูลเมื่อเทียบกับการแจกแจงปกติ การแจกแจงแกมมาและการแจกแจงไวบูลส์

จากภาพที่ 4-4 จึงทดสอบลักษณะการแจกแจงของข้อมูลด้วยสถิติทดสอบ Anderson-Darling Test ผลดังตารางที่ 4-3

ตารางที่ 4-3 ผลการทดสอบการแจกแจงของข้อมูลด้วยสถิติ Anderson-Darling Test

ข้อมูล อุตุนิยมวิทยา	การแจก แจงที่ ทดสอบ	ค่าประมาณพารามิเตอร์ของการแจกแจง			ค่าสถิติ Anderson- Darling	p-value
		พารามิเตอร์ บอกตำแหน่ง	พารามิเตอร์ บอกรูปร่าง	พารามิเตอร์ บอกขนาด		
1) การระเหย น้ำแบบภาค	ปกติ	4.2314	-	0.6848	1.856 [*]	< 0.005
	แกมมา	-	41.2508	0.1026	1.346 [*]	< 0.005
	ไวบูลล์	-	6.13151	4.5315	2.666 [*]	< 0.010
2) อุณหภูมิ อากาศเฉลี่ย	ปกติ	26.3830	-	2.5172	3.070 [*]	< 0.005
	แกมมา	-	103.7236	0.2544	3.505 [*]	< 0.005
	ไวบูลล์	-	14.4100	27.4312	1.711 [*]	< 0.010
3) ความชื้น สัมพัทธ์	ปกติ	72.5394	-	7.9367	1.335 [*]	< 0.005
	แกมมา	-	83.5184	0.8685	1.307 [*]	< 0.005
	ไวบูลล์	-	10.5596	76.0810	1.575 [*]	< 0.010

หมายเหตุ * หมายถึง มีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 4-3 ผลการทดสอบการแจกแจงของการระเหยน้ำแบบภาคกับการแจกแจงปกติ การแจกแจงแกมมา และการแจกแจงไวบูลล์ พบว่า ปฏิเสธสมมติฐานหลักทุก ๆ สมมติฐาน นั่นคือ ข้อมูลการระเหยน้ำแบบภาคไม่สามารถระบุได้ว่าการแจกแจงปกติ การแจกแจงแกมมาและการแจกแจงไวบูลล์ ที่ระดับนัยสำคัญทางสถิติที่ระดับ 0.05 ในทำนองเดียวกัน ผลการทดสอบการแจกแจงของอุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ พบว่า ปฏิเสธสมมติฐานหลักทุก ๆ สมมติฐาน นั่นคือ ข้อมูลอุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ ไม่สามารถระบุได้ว่าการแจกแจงปกติ การแจกแจงแกมมาและการแจกแจงไวบูลล์ ที่ระดับนัยสำคัญทางสถิติที่ระดับ 0.05

เมื่อไม่สามารถระบุการแจกแจงของข้อมูลการระเหยน้ำแบบภาค อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ได้ จึงทำการแปลงข้อมูลด้วยวิธี Johnson Transformation และทดสอบการแจกแจงปกติสำหรับข้อมูลที่แปลงค่าแล้วด้วยสถิติ Anderson-Darling Test ผลดังตารางที่ 4-4

ตารางที่ 4-4 ผลทดสอบการแจกแจงปกติสำหรับข้อมูลที่แปลงค่าแล้วด้วยสถิติ Anderson-Darling Test

ข้อมูล อุตุนิยมวิทยา	สมการแปลงค่าข้อมูล	ค่าประมาณพารามิเตอร์ ของการแจกแจงปกติ		ค่าสถิติ Anderson- Darling	p-value
		พารามิเตอร์ บอกตำแหน่ง	พารามิเตอร์ บอกขนาด		
1) การระเหย น้ำแบบลาด	$Y_t = -2.0488 + 1.5685 \cdot \operatorname{Asinh} \frac{\frac{\alpha Y - 3.2696}{\sigma}}{0.4359}$	0.0812	1.0196	0.343 ^{ns}	0.481
2) อุณหภูมิ อากาศ เฉลี่ย	$X_t = 1.2251 + 0.9295 \operatorname{Asinh} \frac{\frac{\alpha X - 28.8112}{\sigma}}{0.9622}$	0.0314	0.9168	0.491 ^{ns}	0.214
3) ความชื้น สัมพัทธ์	$Z_t = -0.2517 + 0.7113 \ln \frac{\frac{\alpha Z - 54.7732}{\sigma}}{86.2413 - Z}$	0.0148	0.9821	0.627 ^{ns}	0.099

หมายเหตุ ns หมายถึง ไม่มีนัยสำคัญทางสถิติ

จากตารางที่ 4-4 พบว่า เมื่อแปลงค่าข้อมูลการระเหยน้ำแบบลาด อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ ด้วยสมการ (4.15) (4.16) และ (4.17) ตามลำดับ ดังนี้

$$Y_t = -2.0488 + 1.5685 \cdot \operatorname{Asinh} \frac{\frac{\alpha Y - 3.2696}{\sigma}}{0.4359} \quad (4.15)$$

$$X_t = 1.2251 + 0.9295 \operatorname{Asinh} \frac{\frac{\alpha X - 28.8112}{\sigma}}{0.9622} \quad (4.16)$$

$$Z_t = -0.2517 + 0.7113 \ln \frac{\frac{\alpha Z - 54.7732}{\sigma}}{86.2413 - Z} \quad (4.17)$$

เมื่อนำข้อมูลการระเหยน้ำแบบลาด อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ ที่แปลงค่าแล้ว มาทดสอบการแจกแจงปกติด้วยสถิติ Anderson-Darling พบว่า ยอมรับสมมติฐานหลักทุก ๆ สมมติฐาน นั่นคือ ข้อมูลการระเหยน้ำแบบลาด อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์ที่แปลงค่าแล้ว มีการแจกแจงแบบปกติที่ระดับนัยสำคัญทางสถิติที่ระดับ 0.05 (รายละเอียดดังภาคผนวก ก) เมื่อทราบสมการแปลงค่าข้อมูลที่ทำให้การแจกแจงของตัวแปรช่วยความชื้นสัมพัทธ์กับอุณหภูมิอากาศเฉลี่ยเป็นการแจกแจงปกติแล้ว จึงนำสมการแปลงค่าข้อมูลและค่าประมาณพารามิเตอร์ของการแจกแจง ไปทำการจำลองข้อมูลต่อไป

2. ผลการจำลองข้อมูล (Simulation)

ผลการจำลองข้อมูล แบ่งเป็น 2 ส่วน คือ ผลการจำลองข้อมูลประชากรและผลการทดสอบประสิทธิภาพของตัวประมาณค่า ดังนี้

2.1 ผลการจำลองข้อมูลประชากร

ผลการสร้างข้อมูลประชากรโดยใช้คุณหมุมืออากาศ (X) และความชื้นสัมพัทธ์ (Z) ที่แปลงค่าข้อมูลแล้วมีการแจกแจงปกติ สร้างตัวแปรการระเหยน้ำแบบถาด (Y) ที่มีความสัมพันธ์กับ X และ Z ในรูป $Y_i = 3.877 + .199X_i - .067Z_i + \varepsilon_i$ โดยที่ $\varepsilon_i : N(0,1)$ ทำการคัดเลือกชุดข้อมูลอันดับ (Y, X, Z) ที่มีความเหมาะสมกับลักษณะของข้อมูลจริง จำนวน 10,000 อันดับ ซึ่งข้อมูลประชากรจำลองนี้ มีค่าพารามิเตอร์ที่สำคัญที่ใช้ในการเปรียบเทียบประสิทธิภาพของประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ดังตารางที่ 4-5

ตารางที่ 4-5 ค่าพารามิเตอร์ของข้อมูลประชากรจำลอง เมื่อ $N = 10,000$

ตัวแปร	ค่าพารามิเตอร์					
	Min.	Max.	Mean	SD.	Skewness	CV
1) การระเหยน้ำแบบถาด (Y)	3.0000	7.7187	4.0834	0.7776	0.8642	0.1904
2) อุณหภูมิอากาศ (X)	19.0149	30.9967	27.1859	1.7790	-1.3043	0.0654
3) ความชื้นสัมพัทธ์ (Z)	59.4955	86.0000	85.3354	1.5393	-6.4401	0.0180
ความสัมพันธ์รายคู่ :	$\rho_{YX} = 0.37$		$\rho_{YZ} = -0.40$		$\rho_{XZ} = 0.03$	

จากตารางที่ 4-5 พบว่า ข้อมูลประชากรจำลองขนาด $N = 10,000$ หน่วย ที่มีลักษณะสอดคล้องกับข้อมูลจริงแล้ว จึงทำการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ต่อไป

2.2 ผลการทดสอบประสิทธิภาพของตัวประมาณค่า

ผลการทดสอบประสิทธิภาพของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ภายใต้ 12 สถานการณ์ เมื่อกำหนดประชากรขนาด 10,000 หน่วย ทำการสุ่มตัวอย่างด้วยวิธีการชักตัวอย่างแบบสุ่มเชิงเดียว และจำลองข้อมูลผิดปกติตามสถานการณ์ที่กำหนดไว้ นั่นคือ กำหนดขนาดตัวอย่างสุ่ม (n) เท่ากับ 20 40 และ 100 หน่วยตัวอย่าง และร้อยละของจำนวนค่าผิดปกติ เท่ากับ ร้อยละ 5 ร้อยละ 10 ร้อยละ 15 และ ร้อยละ 20 ตามลำดับ ซึ่งได้การจำลองข้อมูล 12 (3×4) สถานการณ์ และแต่ละสถานการณ์ ทำซ้ำ 1,000 ครั้ง ใช้เกณฑ์ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยจากนิยาม $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$ และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความ

คลาดเคลื่อนสัมพัทธ์ (*MPRE*) ของตัวประมาณค่าทั้งสองตัว เป็นเกณฑ์เปรียบเทียบประสิทธิภาพของตัวประมาณค่า ทั้งนี้ประสิทธิภาพของตัวประมาณค่าพิจารณาในด้านความเที่ยงตรง (*Precision*) และความแม่นยำ (*Accuracy*) โดยพิจารณาความเที่ยงตรงของตัวประมาณค่า จากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (*Mean Square Error: MSE*) ของตัวประมาณค่า (Cochran, 1977, p. 16) และพิจารณาความแม่นยำของตัวประมาณค่า จากค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (*Mean Percent Relative Error: MPRE*) หากค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าใดมีค่าต่ำสุด ก็ถือว่าเป็นตัวประมาณค่าที่มีความเที่ยงตรงสูงกว่า และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ของตัวประมาณค่าใดมีค่าต่ำสุด ก็ถือว่าเป็นตัวประมาณค่าที่มีความแม่นยำสูงกว่า

ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) ด้วยเกณฑ์ค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด และเกณฑ์ค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ต่ำสุด ดังตารางที่ 4-6

ตารางที่ 4-6 ผลการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่า $\bar{y}_{m(SD)}$

ขนาดตัวอย่าง	ร้อยละของค่าผิดพลาด	ตัวประมาณค่าเริ่มต้น ($\bar{y}_{m(SD)}$)			ตัวประมาณค่าใหม่ (\bar{y}_M)		
		ค่าเฉลี่ย $\bar{y}_{m(SD)}$	<i>MSE</i>	<i>MPRE</i>	ค่าเฉลี่ย \bar{y}_M	<i>MSE</i>	<i>MPRE</i>
n = 20	5	4.1727	0.1787	3.4268	4.1406	0.1761	3.3273
	10	4.2841	0.2533	5.1652	4.2295	0.2252	4.3968
	15	4.4058	0.3559	7.9325	4.3039	0.2856	5.9308
	20	4.5324	0.4736	10.9950	4.3856	0.3719	7.8994
n = 40	5	4.1719	0.1415	2.7857	4.1513	0.1348	2.6476
	10	4.2896	0.2311	5.0678	4.2422	0.1946	4.0577
	15	4.4018	0.3354	7.7974	4.3236	0.2683	5.9610
	20	4.5349	0.4639	11.0553	4.3986	0.3428	7.7899
n = 100	5	4.1668	0.1077	2.2097	4.1500	0.0951	1.8958
	10	4.2900	0.2168	5.0585	4.2484	0.1788	4.0511
	15	4.4124	0.3353	8.0555	4.3446	0.2704	6.3948
	20	4.5346	6.4461	11.0496	4.4211	4.8869	8.2703

หมายเหตุ ตัวหนา หมายถึง ค่า *MSE* หรือค่า *MPRE* ที่มีค่าน้อยที่สุดในแนวแถว

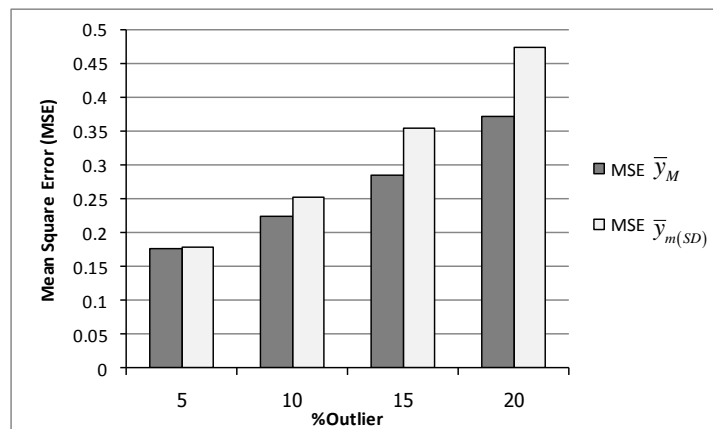
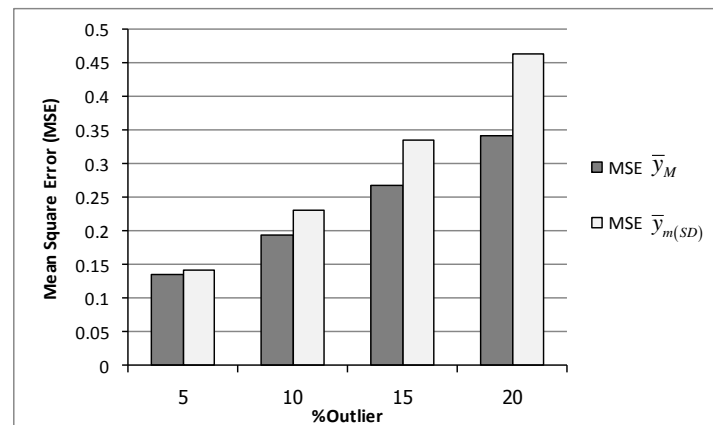
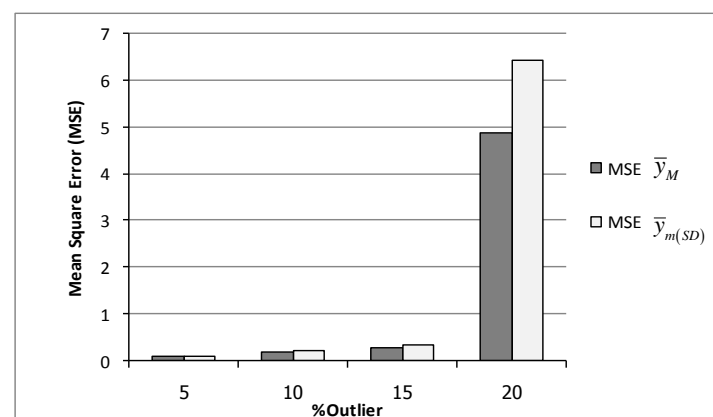
จากตารางที่ 4-6 ผลการจำลองทั้งหมด 12 สถานการณ์ พิจารณาประสิทธิภาพของตัวประมาณค่าทั้งสองตัว ในด้านความเที่ยงตรง (Precision) และความแม่นยำ (Accuracy) จากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ($MPRE$) ผลดังนี้

1) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่า \bar{y}_M มีค่าน้อยกว่าค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุกสถานการณ์ สรุปได้ว่าตัวประมาณค่า \bar{y}_M มีความเที่ยงตรงมากกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุกสถานการณ์ทั้ง 12 สถานการณ์

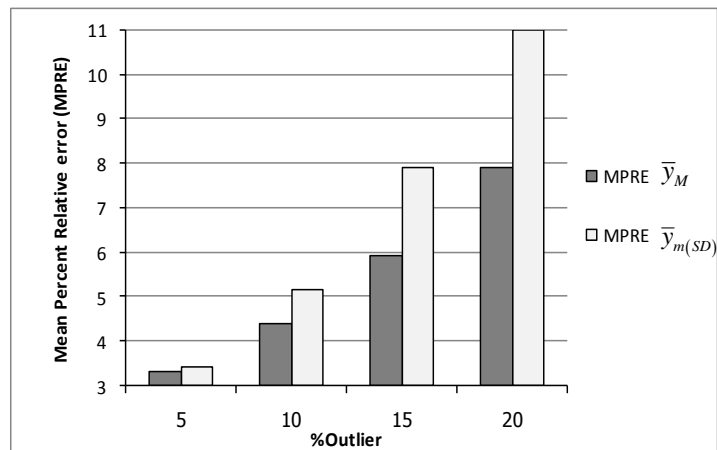
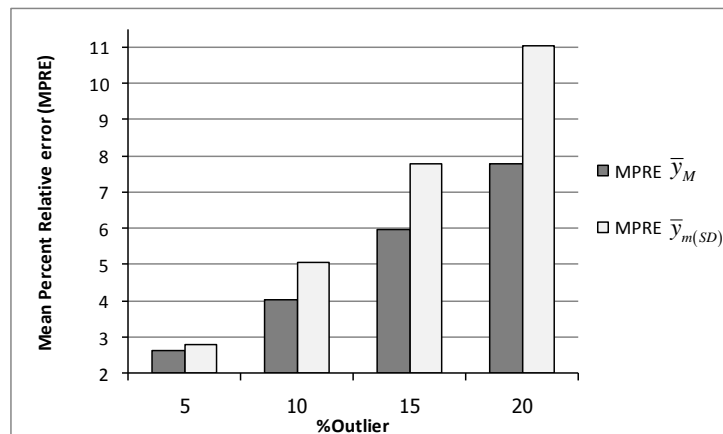
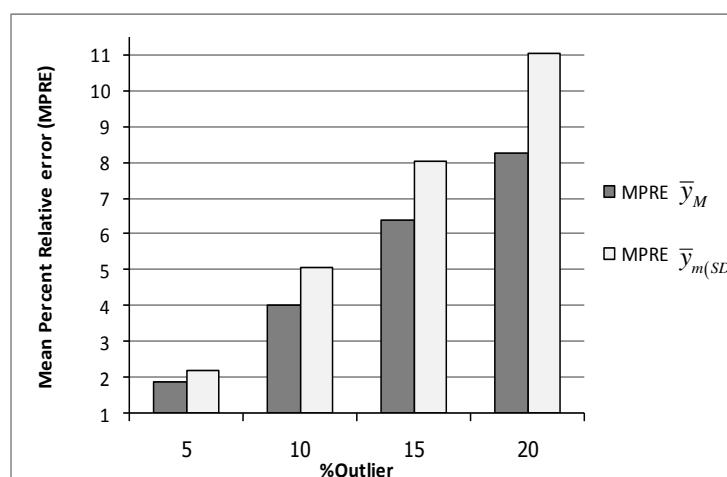
2) ค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ($MPRE$) ของตัวประมาณค่า \bar{y}_M มีค่าน้อยกว่าค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุกสถานการณ์ สรุปได้ว่า ตัวประมาณค่า \bar{y}_M มีความแม่นยำมากกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุกสถานการณ์ทั้ง 12 สถานการณ์

นอกจากนี้ ยังพบข้อสังเกตที่ว่า ในทุกกรณีการสุ่มตัวอย่างขนาด $n = 20, 40$ และ 100 นั้น เมื่อร้อยละของจำนวนค่าผิดปกติมีค่าเพิ่มขึ้น จากร้อยละ 5 เพิ่มเป็นร้อยละ 10 ร้อยละ 15 และร้อยละ 20 ตามลำดับ จะพบว่า ค่า MSE และค่า $MPRE$ ของตัวประมาณค่า \bar{y}_M มีค่าเพิ่มขึ้น นั่นคือเมื่อข้อมูลมีค่าผิดปกติมากขึ้น ตัวประมาณค่า \bar{y}_M อาจจะมีค่าความเที่ยงตรงและความแม่นยำลดลงบ้าง แต่ตัวประมาณค่า \bar{y}_M ยังมีความเที่ยงตรงและความแม่นยำสูงกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ โดยเฉพาะเมื่อชุดข้อมูลมีจำนวนค่าผิดปกติปะปนในชุดข้อมูลมากขึ้น ตัวประมาณค่า \bar{y}_M ยิ่งมีความเที่ยงตรงและความแม่นยำสูงกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ อย่างชัดเจนมากขึ้นด้วย

สรุปสถานการณ์ทั้งหมดที่ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า \bar{y}_M น้อยกว่าค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า $\bar{y}_{m(SD)}$ และค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ($MPRE$) ของตัวประมาณค่า \bar{y}_M มีค่าน้อยกว่าค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า $\bar{y}_{m(SD)}$ ซึ่งสอดคล้องเป็นไปตามสมมติฐานการวิจัย จำนวน 12 สถานการณ์ ดังภาพที่ 4-5 และภาพที่ 4-6

ก) $n = 20$ ข) $n = 40$ ค) $n = 100$

ภาพที่ 4-5 สรุปสถานการณ์ 12 สถานการณ์ที่ตัวประมาณ \bar{y}_M มีความเที่ยงตรงสูงกว่าตัวประมาณ $\bar{y}_{m(SD)}$

ก) $n = 20$ ข) $n = 40$ ค) $n = 100$

ภาพที่ 4-6 สรุปสถานการณ์ 12 สถานการณ์ที่ตัวประมาณ \bar{y}_M มีความแม่นยำสูงกว่าตัวประมาณ $\bar{y}_{m(SD)}$

กล่าวโดยสรุปผลการจำลองข้อมูลเพื่อการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน ด้วยตัวประมาณค่า \bar{y}_M กับตัวประมาณค่า $\bar{y}_{m(SD)}$ ในกรณีตัวแปรช่วย คือ อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ พบว่าการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าที่นำเสนอ \bar{y}_M จะมีประสิทธิภาพดีกว่าตัวประมาณ $\bar{y}_{m(SD)}$ ทั้งในด้านความเที่ยงตรงและความแม่นยำของตัวประมาณค่า ไม่ว่าจะสุ่มตัวอย่างขนาดใดก็ตาม และเมื่อข้อมูลมีค่าผิดปกติมากขึ้น ตัวประมาณค่า \bar{y}_M อาจจะมีความเที่ยงตรงและความแม่นยำลดลงบ้าง เมื่อเทียบกับกรณีที่มีค่าผิดปกติเล็กน้อย แต่ตัวประมาณค่า \bar{y}_M ยังมีความเที่ยงตรงและความแม่นยำสูงกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ โดยเฉพาะเมื่อชุดข้อมูลมีจำนวนค่าผิดปกติปะปนในชุดข้อมูลยิ่งมากขึ้นเท่าใด ตัวประมาณค่า \bar{y}_M ยังมีความเที่ยงตรงและความแม่นยำสูงกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ อย่างชัดเจนมากขึ้นด้วย

ตอนที่ 3 ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรแบบใหม่ที่นำเสนอ (\bar{y}_M) กรณีใช้ตัวแปรช่วยสองตัว โดยนำข้อมูลอุณหภูมิมิถุนายนรายวันในแต่ละเดือนมาประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน เนื่องจากค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในชั้นบรรยากาศไม่สามารถวัดได้โดยตรงและวัดค่าให้ถูกต้องทำได้ยาก (Thompson, 1999, p. 91; วิชัย ศรีบุญลือ, 2552, หน้า 8) จึงอาศัยแนวทางการศึกษาของ Irmak and Haman (2003) และการศึกษาของ Dabral et al. (2014) กำหนดค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนในแต่ละเดือนด้วยค่าเฉลี่ยเฉพาะเดือนนั้น จากข้อมูลอุณหภูมิมิถุนายนรายเดือนย้อนหลัง 6 ปี (พฤษภาคม พ.ศ. 2550 ถึงเดือนธันวาคม พ.ศ. 2555) ผลดังนี้

1. ผลการวิเคราะห์ค่าพารามิเตอร์ประชากรที่ศึกษา

การระเหยน้ำรายเดือนในจังหวัดสกลนคร จากเดือนพฤษภาคม 2550 ถึงเดือนธันวาคม 2555 ภายในระยะเวลา 6 ปีย้อนหลัง รายละเอียดดังตารางที่ 4-7

ตารางที่ 4-7 ค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนของเดือนพฤษภาคม 2550 ถึงเดือน ธันวาคม 2555

ปี พ.ศ.	เดือน							
	พฤษภาคม	มิถุนายน	กรกฎาคม	สิงหาคม	กันยายน	ตุลาคม	พฤศจิกายน	ธันวาคม
2550	5.2035	4.2077	4.0332	3.3316	3.4360	3.4671	3.7680	3.7116
2551	3.9368	3.3707	3.4452	3.0384	3.3643	3.7061	3.9820	3.2390
2552	4.9765	4.2160	3.5032	3.5539	3.7467	4.0135	4.1420	3.4848
2553	4.9152	3.8420	4.3561	2.9887	3.8420	3.7058	3.8017	3.5506
2554	4.9658	4.1923	3.5313	3.4023	3.1713	3.3523	4.0150	3.5548
2555	4.1968	4.2293	4.0319	3.7042	3.9593	4.6084	4.1040	3.9674
ค่าพารามิเตอร์การระเหยน้ำ* (\bar{Y}_i)	4.6991	4.0097	3.8168	3.3365	3.5866	3.8089	3.9688	3.5847
	\bar{Y}_5	\bar{Y}_6	\bar{Y}_7	\bar{Y}_8	\bar{Y}_9	\bar{Y}_{10}	\bar{Y}_{11}	\bar{Y}_{12}

หมายเหตุ ค่าพารามิเตอร์การระเหยน้ำ หมายถึง ปริมาณการระเหยน้ำโดยเฉลี่ยรายเดือน (มิลลิเมตร)

จากตารางที่ 4-7 จะได้ค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือนของเดือนพฤษภาคมถึงเดือนธันวาคม พ.ศ. 2556 เท่ากับ 4.6991, 4.0097, 3.8168, 3.3365, 3.5866, 3.8089, 3.9688 และ 3.5847 มิลลิเมตร ตามลำดับ

สำหรับค่าพารามิเตอร์อื่น ๆ เช่น ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน สัมประสิทธิ์ความแปรปรวน สำหรับประชากรตัวแปรช่วยอุณหภูมิอากาศและความชื้นสัมพัทธ์ ของเดือนพฤษภาคมถึงเดือนธันวาคม จากข้อมูล 6 ปีย้อนหลัง รายละเอียดดังตารางที่ 4-8

ตารางที่ 4-8 ค่าพารามิเตอร์ของข้อมูลประชากรรายเดือนพฤษภาคมถึงเดือนธันวาคม

อันดับเดือน/ พ.ศ. 2550 - 2555	ข้อมูลอุตุนิยมวิทยา	ค่าพารามิเตอร์			
		N	Mean	SD.	Skewness
5) พฤษภาคม	1) การระเหยน้ำแบบลาด (Y)	186	4.6991	1.9001	0.0395
	2) อุณหภูมิอากาศ (X)	186	28.4380	1.6765	0.1775
	3) ความชื้นสัมพัทธ์ (Z)	186	75.6559	8.8797	-0.2822
6) มิถุนายน	1) การระเหยน้ำแบบลาด (Y)	180	4.0097	1.7622	-0.2702
	2) อุณหภูมิอากาศ (X)	180	28.3924	1.3696	0.0819
	3) ความชื้นสัมพัทธ์ (Z)	180	79.5396	6.3979	-0.1339

ตารางที่ 4-8 ค่าพารามิเตอร์ของข้อมูลประชากรรายเดือนพฤษภาคมถึงเดือนธันวาคม (ต่อ)

อันดับเดือน/ พ.ศ. 2550 - 2555	ข้อมูลอุตุนิยมวิทยา	ค่าพารามิเตอร์			
		N	Mean	SD.	Skewness
7) กรกฎาคม	1) การระเหยน้ำแบบลาด (Y)	186	3.8168	1.7720	-0.1967
	2) อุณหภูมิอากาศ (X)	186	28.0200	1.3688	-0.1183
	3) ความชื้นสัมพัทธ์ (Z)	186	80.7345	5.8625	-0.1914
8) สิงหาคม	1) การระเหยน้ำแบบลาด (Y)	186	3.3365	1.6117	-0.0482
	2) อุณหภูมิอากาศ (X)	186	27.4357	1.1696	-0.2618
	3) ความชื้นสัมพัทธ์ (Z)	186	83.4543	4.9478	0.1493
9) กันยายน	1) การระเหยน้ำแบบลาด (Y)	180	3.5866	1.5254	-0.1266
	2) อุณหภูมิอากาศ (X)	180	27.4626	1.1801	0.2632
	3) ความชื้นสัมพัทธ์ (Z)	180	82.7026	5.7914	-0.4133
10) ตุลาคม	1) การระเหยน้ำแบบลาด (Y)	186	3.8089	1.3345	-0.4898
	2) อุณหภูมิอากาศ (X)	186	26.7039	1.3891	-1.2216
	3) ความชื้นสัมพัทธ์ (Z)	186	76.4556	7.2647	0.4974
11) พฤศจิกายน	1) การระเหยน้ำแบบลาด (Y)	180	3.9688	1.0562	-0.1583
	2) อุณหภูมิอากาศ (X)	180	24.7847	2.4376	-0.2949
	3) ความชื้นสัมพัทธ์ (Z)	180	69.1507	5.1449	1.1990
12) ธันวาคม	1) การระเหยน้ำแบบลาด (Y)	186	3.5847	0.9390	-0.0235
	2) อุณหภูมิอากาศ (X)	186	22.9376	2.4671	-0.1860
	3) ความชื้นสัมพัทธ์ (Z)	186	68.7991	3.7384	0.9735

จากตารางที่ 4-8 พบว่าข้อมูลการระเหยน้ำแบบลาด อุณหภูมิอากาศและความชื้นสัมพัทธ์ ในระดับประชากรของเดือนพฤษภาคม ถึงเดือนธันวาคม มีลักษณะเบ้เล็กน้อย

2. ผลการวิเคราะห์ค่าสถิติจากกลุ่มตัวอย่าง

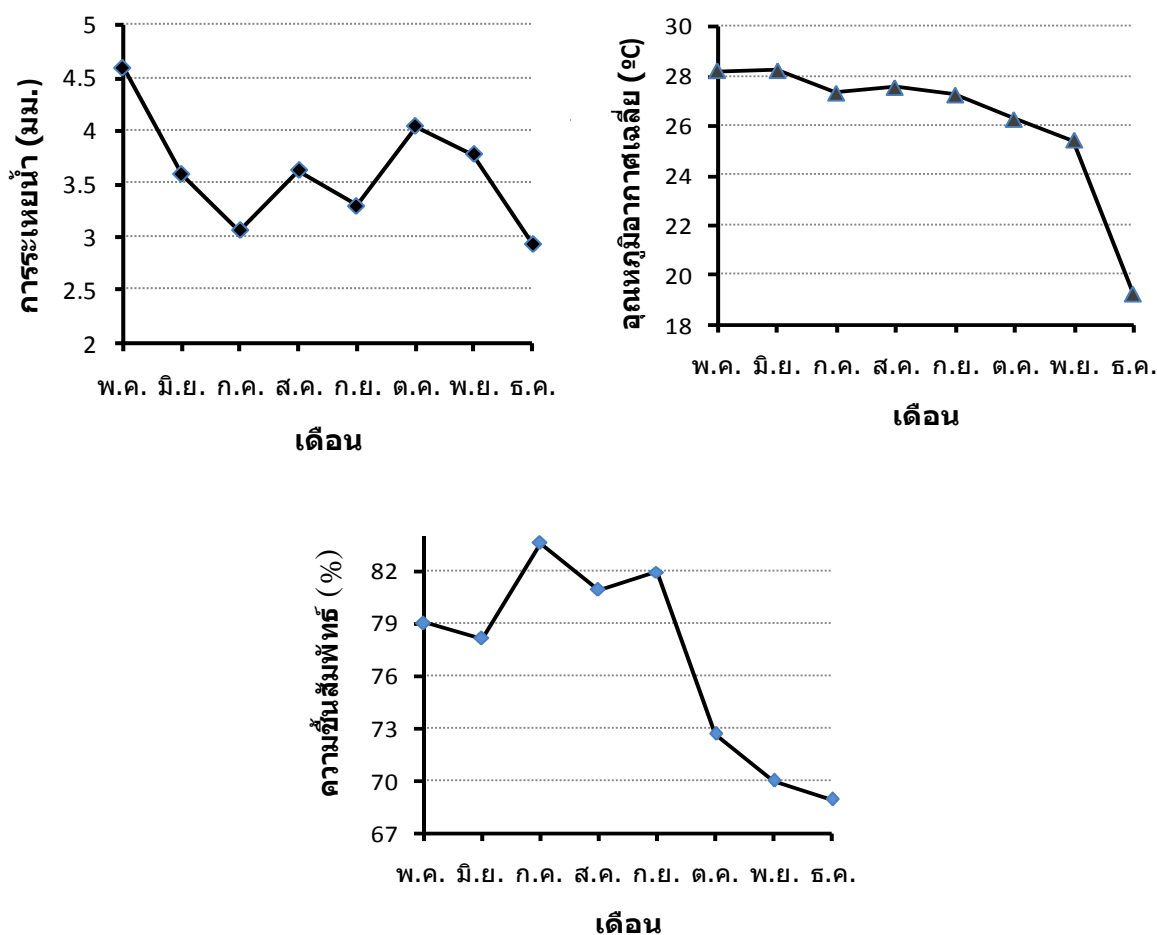
กลุ่มตัวอย่างข้อมูลอุตุนิยมวิทยารายวันของจังหวัดสกลนคร ประกอบด้วย การระเหยน้ำ (มิลลิเมตร) และตัวแปรช่วยอุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$) และความชื้นสัมพัทธ์ จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 มีค่าสถิติพื้นฐานเบื้องต้นพร้อมค่าสัมประสิทธิ์การถดถอยแบบด้วยวิธี LTS (b^*) ของตัวแปรช่วยอุณหภูมิอากาศ (X) และความชื้นสัมพัทธ์ (Z) ดังตารางที่ 4.9

ตารางที่ 4-9 ค่าสถิติของข้อมูลตัวอย่างรายเดือน พฤษภาคมถึงเดือนธันวาคม พ.ศ. 2556

อันดับเดือน/ พ.ศ. 2556	ข้อมูลอุตุนิยมวิทยา	ค่าสถิติ				
		<i>n</i>	Mean	<i>SD.</i>	Skewness	<i>b</i> *
5) พฤษภาคม	1) การระเหยน้ำแบบลาด (Y)	31	4.5974	1.7410	-0.8952	-
	2) อุณหภูมิอากาศ (X)	31	28.1981	1.7803	0.5543	0.3432
	3) ความชื้นสัมพัทธ์ (Z)	31	79.0148	6.4630	-0.0126	-0.0686
6) มิถุนายน	1) การระเหยน้ำแบบลาด (Y)	30	3.5813	1.9395	-0.1728	-
	2) อุณหภูมิอากาศ (X)	30	28.2217	1.5330	-0.3314	0.3750
	3) ความชื้นสัมพัทธ์ (Z)	30	78.1270	7.6419	-0.0514	-0.0841
7) กรกฎาคม	1) การระเหยน้ำแบบลาด (Y)	31	3.0594	1.9963	-0.3455	-
	2) อุณหภูมิอากาศ (X)	31	27.3242	1.3106	-0.1874	0.3716
	3) ความชื้นสัมพัทธ์ (Z)	31	83.5674	5.3278	0.0511	-0.0803
8) สิงหาคม	1) การระเหยน้ำแบบลาด (Y)	31	3.6229	2.0965	0.2741	-
	2) อุณหภูมิอากาศ (X)	31	27.5697	1.1045	-0.3339	0.3760
	3) ความชื้นสัมพัทธ์ (Z)	31	80.8855	5.6706	-0.7938	-0.0828
9) กันยายน	1) การระเหยน้ำแบบลาด (Y)	30	3.2943	1.7359	-0.2361	-
	2) อุณหภูมิอากาศ (X)	30	27.2360	0.8763	-0.4107	0.3756
	3) ความชื้นสัมพัทธ์ (Z)	30	81.9480	4.7717	-0.3140	-0.0814
10) ตุลาคม	1) การระเหยน้ำแบบลาด (Y)	31	4.0419	1.0201	-0.5758	-
	2) อุณหภูมิอากาศ (X)	31	26.2832	1.3316	-0.8154	0.3478
	3) ความชื้นสัมพัทธ์ (Z)	31	72.6719	5.1057	1.8876	-0.0712
11) พฤศจิกายน	1) การระเหยน้ำแบบลาด (Y)	30	3.7707	1.0024	0.1562	-
	2) อุณหภูมิอากาศ (X)	30	25.4077	1.7091	-0.5298	0.1844
	3) ความชื้นสัมพัทธ์ (Z)	30	70.0197	3.1371	1.2009	-0.0094
12) ธันวาคม	1) การระเหยน้ำแบบลาด (Y)	31	2.9361	0.8739	-0.1399	-
	2) อุณหภูมิอากาศ (X)	31	19.2074	2.7091	0.3190	0.1307
	3) ความชื้นสัมพัทธ์ (Z)	31	68.9381	5.2633	3.1662	0.0077

จากตารางที่ 4-9 จะเห็นได้ว่า การระเหยน้ำแบบลาดโดยเฉลี่ยรายเดือน มีค่าสูงสุดในเดือนพฤษภาคม ซึ่งอยู่ในฤดูร้อน ดังนั้นอุณหภูมิอากาศเฉลี่ยของเดือนพฤษภาคมและมิถุนายนมีค่าสูงกว่าเดือนอื่น ๆ ส่วนเดือนที่มีการระเหยน้ำแบบลาดโดยเฉลี่ยรายเดือน มีค่าต่ำสุด คือ เดือนธันวาคม ซึ่งอยู่ในฤดูหนาว ก็สอดคล้องกับความชื้นสัมพัทธ์ในเดือนที่อยู่ในช่วงฤดูหนาว จะมีค่าสูงกว่าเดือนในฤดูอื่น ๆ อุณหภูมิอากาศและความชื้นสัมพัทธ์ ต่างก็มีความสัมพันธ์กับการระเหยน้ำในระดับปานกลาง โดยที่อุณหภูมิอากาศมีความสัมพันธ์ทางบวกกับการระเหยน้ำ ส่วนความชื้นสัมพัทธ์มีความสัมพันธ์ทางลบกับการระเหยน้ำ (รายละเอียดดังภาคผนวก ค)

สรุปการเปลี่ยนแปลงของค่าเฉลี่ยของการระเหยน้ำ (มิลลิเมตร) และตัวแปรช่วยอุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$) และความชื้นสัมพัทธ์ จากเดือนพฤษภาคม ถึงเดือนธันวาคม พ.ศ. 2556 ดังภาพที่ 4-7



ภาพที่ 4-7 การเปลี่ยนแปลงของค่าเฉลี่ยของการระเหยน้ำ (มิลลิเมตร) อุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$) และความชื้นสัมพัทธ์ จากเดือนพฤษภาคม ถึงเดือนธันวาคม 2556

3. ผลการประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่า \bar{y}_M

ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กรณีใช้ตัวแปรช่วยสองตัว ได้แก่ อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ โดยประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน ของเดือน พฤษภาคมถึงเดือนธันวาคม 2556 เป็นจำนวน 8 เดือน จากข้อมูลอุตุนิยมหาวิทยาลัยราชภัฏวรินทร์ในแต่ละเดือนนั้น ๆ (รายละเอียดข้อมูลอุตุนิยมหาวิทยาลัยราชภัฏวรินทร์ ดังภาคผนวก ค) นำข้อมูลตัวอย่างเข้าโปรแกรมชุดคำสั่ง Matlab ที่เขียนเพื่อประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน ของเดือน พฤษภาคมถึง ธันวาคม 2556 ด้วยตัวประมาณค่า \bar{y}_M ดังสมการ 4.18 (โปรแกรมชุดคำสั่ง ดังภาคผนวก จ)

$$\bar{y}_M = \frac{\bar{y} + b_1^*(\bar{X} - \bar{x}) + b_2^*(\bar{Z} - \bar{z})}{\frac{\bar{y}_x + \beta_{1(x)} \bar{z} + \beta_{1(z)} \bar{x}}{\bar{y}_x + \beta_{1(x)} \bar{z} + \beta_{1(z)} \bar{x}}} \quad (4.18)$$

หลังจากประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนของเดือน พฤษภาคมถึงเดือนธันวาคม 2556 โดยใช้ตัวประมาณค่า \bar{y}_M แล้ว จึงพิจารณาความแม่นยำของการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย (Mean Percent Relative Error: MPRE) ผลดังตารางที่ 4-10

ตารางที่ 4-10 ความแม่นยำของค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือน พฤษภาคมถึงธันวาคม 2556 ด้วยตัวประมาณค่า \bar{y}_M

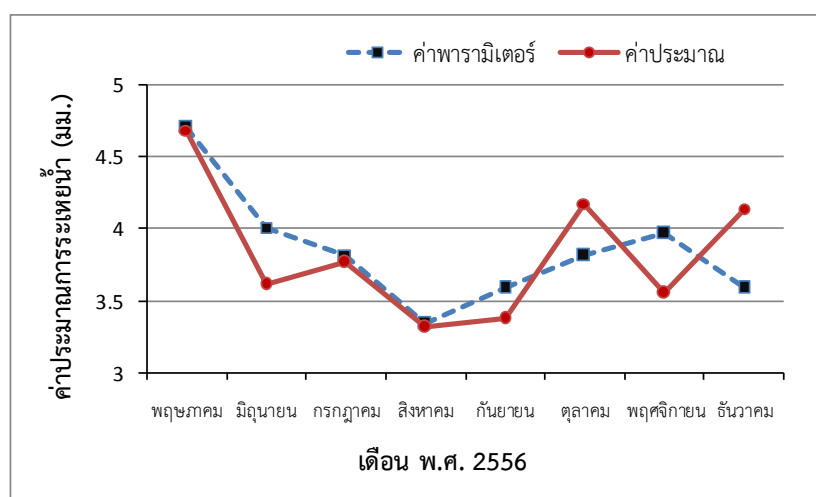
เดือน	ค่าพารามิเตอร์ \bar{y}_i	ค่าประมาณ \bar{y}_{M_i}	ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (PRE)
พฤษภาคม	4.6991	4.6853	0.2937
มิถุนายน	4.0097	3.6147	9.8511
กรกฎาคม	3.8168	3.7657	1.3388
สิงหาคม	3.3365	3.3169	0.5874
กันยายน	3.5866	3.3716	5.9945
ตุลาคม	3.8089	4.1744	9.5959
พฤศจิกายน	3.9688	3.5493	10.5699
ธันวาคม	3.5847	4.1398	15.4853
ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย (PMRE)			6.7146

จากตารางที่ 4-10 พบว่า ค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) มีค่าใกล้เคียงกับค่าพารามิเตอร์มากที่สุด ใน 3 อันดับแรก ได้แก่ เดือน พฤษภาคม สิงหาคม และกรกฎาคม ตามลำดับ เมื่อพิจารณาค่าความคลาดเคลื่อนสัมพัทธ์ในรายเดือนนั้น พบว่า ค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนด้วย

ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ในเดือนพฤษภาคม มิถุนายน กรกฎาคม สิงหาคม กันยายน และตุลาคม ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ไม่เกิน 10% ยกเว้นเดือน พฤศจิกายนและธันวาคมมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์เกิน 10% เล็กน้อย

เมื่อพิจารณาในภาพรวม จะเห็นว่า การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วย ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย เท่ากับ 6.7146% (ไม่เกิน 10%) ซึ่งสอดคล้องเป็นไปตามสมมติฐานการวิจัย

สรุปค่าพารามิเตอร์และค่าประมาณของการระเหยน้ำโดยเฉลี่ยรายเดือนพฤษภาคมถึง ธันวาคม 2556 ด้วยตัวประมาณค่า \bar{y}_M ดังภาพที่ 4-8



ภาพที่ 4-8 ค่าพารามิเตอร์และค่าประมาณของการระเหยน้ำโดยเฉลี่ยรายเดือนพฤษภาคมถึง ธันวาคม 2556 ด้วยตัวประมาณค่า \bar{y}_M

จากภาพที่ 4-8 พบว่า ค่าประมาณของการระเหยน้ำโดยเฉลี่ยรายเดือนพฤษภาคม กรกฎาคม และสิงหาคม มีค่าใกล้เคียงกับค่าพารามิเตอร์ ส่วนค่าประมาณของการระเหยน้ำโดยเฉลี่ย เดือนมิถุนายน กันยายน พฤศจิกายน จะได้ค่าประมาณที่ต่ำกว่าค่าพารามิเตอร์ (Underestimation) แต่ค่าประมาณของการระเหยน้ำโดยเฉลี่ยเดือนตุลาคมและธันวาคม จะได้ค่าประมาณที่สูงเกินค่าพารามิเตอร์ (Overestimation) ถึงอย่างไรก็ตาม ตัวประมาณค่า \bar{y}_M ก็ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ยไม่เกิน 10% ซึ่งสอดคล้องเป็นไปตามสมมติฐานการวิจัย จึงถือว่าตัวประมาณค่า \bar{y}_M สามารถใช้ประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนได้

สรุปได้ว่าการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร ซึ่งเป็นข้อมูลที่มีความผันแปรหรืออาจมีค่าผิดปกติรวมอยู่ด้วย หากสามารถใช้ตัวแปรช่วยเพื่อเพิ่มแม่นยำในการประมาณค่าข้อมูลได้แล้ว ตัวประมาณค่า \bar{y}_M มีความเหมาะสมที่จะใช้ในการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนครได้

บทที่ 5

สรุปและอภิปรายผล

การวิจัยนี้เป็นการประมาณค่าการระเหยน้ำโดยใช้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ซึ่งใช้ตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) เป็นตัวประมาณค่าเริ่มต้น (Initial Estimator) ทำการปรับสูตรโดยอาศัยการวิเคราะห์ถดถอยແກ່ງที่ประมาณค่าสัมประสิทธิ์สหสัมพันธ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุด ร่วมกับการใช้สัมประสิทธิ์ความเบ้ของตัวแปรช่วย มาพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กรณีใช้ตัวแปรช่วยสองตัว เพื่อให้ได้ตัวประมาณค่าแบบอัตราส่วนที่มีความเหมาะสมกับการประมาณค่าเฉลี่ยประชากร เมื่อข้อมูลเดิมอาจปะปนด้วยข้อมูลค่าผิดปกติ โดยมีวัตถุประสงค์เพื่อ

- 1) พัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ โดยใช้ตัวแปรช่วยสองตัว ในกรณีข้อมูลประชากรมีค่าผิดปกติ
- 2) เปรียบเทียบประสิทธิภาพตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) ภายใต้สถานการณ์ 12 สถานการณ์ ซึ่งเป็นผลประกอบของเงื่อนไขขนาดตัวอย่าง 3 สถานการณ์และร้อยละของค่าผิดปกติ 4 สถานการณ์ โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล พิจารณาความเที่ยงตรง (Precision) และความแม่นยำ (Accuracy) ของตัวประมาณค่า เป็นเกณฑ์เปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ และ
- 3) ประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร เดือนพฤษภาคม ถึง ธันวาคม 2556 โดยใช้ข้อมูลอุตุนิยมหาวิทยาลัยรายวันในแต่ละเดือนใด ๆ มาประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน ด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กำหนดเกณฑ์ที่ใช้ในการยอมรับความคลาดเคลื่อนจากการประมาณค่า (Acceptable Error) สำหรับการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่า \bar{y}_M คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย (Mean Percent Relative Error: MPRE) ไม่เกิน 10%

สรุปผลการวิจัย

1. ผลการพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่

การพัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ เมื่อใช้ตัวแปรช่วยสองตัวนี้ ได้พัฒนามาจากตัวประมาณค่าแบบอัตราส่วนของ Gupta and Shabbir (2007) โดยอาศัยการวิเคราะห์การถดถอยແກ່ງที่ใช้ตัวประมาณค่าสัมประสิทธิ์การถดถอยวิธีกำลังสองตัดขอบน้อยสุด (LTS) ร่วมกับการใช้สารสนเทศสัมประสิทธิ์ความเบ้ของข้อมูล ทำให้ได้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ดังนี้

$$\bar{y}_M = \frac{\sum y}{n} + b_1^* (\bar{X} - \bar{x}) + b_2^* (\bar{Z} - \bar{z}) \quad (5.1)$$

เมื่อ $\beta_{1(x)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย X

$\beta_{1(z)}$ คือ สัมประสิทธิ์ความเบ้ของตัวแปรช่วย Z

b_1^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย X

b_2^* คือ สัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดของตัวแปรช่วย Z

\bar{X} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย X

\bar{Z} คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย Z

\bar{x} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย X

\bar{z} คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย Z

จากนั้นใช้วิธีการกระจายอนุกรมเทย์เลอร์ดีกรีอันดับหนึ่ง หาสมการ MSE ของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ดังนี้

$$MSE(\bar{y}_M) \cong \frac{1-f}{n} \left[\frac{\sum y^2}{n} + (B_1^* + R_1)^2 S_x^2 + (B_2^* - R_2)^2 S_z^2 - 2(B_1^* + R_1)S_{yx} - 2(B_2^* - R_2)S_{yz} + 2(B_1^* + R_1)(B_2^* - R_2)S_{xz} \right] \quad (5.2)$$

$$\text{เมื่อ } f = \frac{n}{N}, R_1 = \frac{\bar{y}}{\bar{x} + \beta_{1(x)}} \text{ และ } R_2 = \frac{\bar{y}}{\bar{z} + \beta_{1(z)}} \text{ ตามลำดับ}$$

2. ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$

การเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน ($\bar{y}_{m(SD)}$) ของ Gupta and Shabbir (2007) ด้วยวิธีการจำลองข้อมูลโดยเทคนิคมอนติคาร์โล เริ่มจากการศึกษาลักษณะเบื้องต้นของข้อมูลอุตุนิยมวิทยาที่มีความสัมพันธ์กับการระเหยน้ำแบบภาค สามารถคัดเลือกตัวแปรช่วยสองตัวที่มีความสัมพันธ์กับการระเหยน้ำแบบภาคได้ คือ อุณหภูมิอากาศเฉลี่ย ($^{\circ}\text{C}$) (Temp) และความชื้นสัมพัทธ์ (%) (Humidity) พร้อมทั้งทดสอบรูปแบบการแจกแจงลักษณะการเกิดข้อมูลการระเหยน้ำและตัวแปรช่วยสองตัวนั้น จึงกำหนดรูปแบบการแจกแจงของตัวแปรช่วยสองตัวที่ผ่านการแปลงค่าข้อมูลแล้วสามารถระบุการแจกแจงเป็นการแจกแจงปกติได้ จากนั้นจึงจำลองข้อมูลโดยเทคนิคมอนติคาร์โลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ภายใต้สถานการณ์ 12 สถานการณ์ ซึ่งประกอบด้วย

ขนาดตัวอย่างสุ่ม (n) เท่ากับ 20 40 และ 100 หน่วยตัวอย่างและร้อยละของค่าผิดปกติ (o) เท่ากับ ร้อยละ 5 ร้อยละ 10 ร้อยละ 15 และ ร้อยละ 20 ทั้งนี้แต่ละสถานการณ์ ทำซ้ำ 1,000 ครั้ง และใช้ เกณฑ์ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) กับค่าเฉลี่ยของค่า เปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ (Mean Percent Relative Error: $MPRE$) ต่ำสุด เป็นเกณฑ์ เปรียบเทียบประสิทธิภาพของตัวประมาณค่าในด้านความเที่ยงตรง (Precision) และความแม่นยำ (Accuracy) ของตัวประมาณค่า ตามลำดับ

ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ ปรับใหม่ (\bar{y}_M) กับตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากร ($\bar{y}_{m(SD)}$) โดยใช้เกณฑ์ค่าความคลาดเคลื่อนกำลัง สองเฉลี่ย (MSE) ต่ำสุด และเกณฑ์ค่าเฉลี่ยของค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์ ($MPRE$) ต่ำสุด สรุปว่า ตัวประมาณค่า \bar{y}_M มีความเที่ยงตรงและความแม่นยำมากกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุกสถานการณ์ทั้ง 12 สถานการณ์ ซึ่งสอดคล้องเป็นไปตามทุก ๆ สมมติฐานการวิจัย จำนวน 12 สถานการณ์ นอกจากนี้ยังพบว่า เมื่อข้อมูลมีร้อยละของค่าผิดปกติมากขึ้น ตัวประมาณค่า \bar{y}_M อาจจะมีค่าความเที่ยงตรงและความแม่นยำลดลงบ้าง แต่ตัวประมาณค่า \bar{y}_M ยังมีความเที่ยงตรง และความแม่นยำสูงกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ โดยเฉพาะเมื่อชุดข้อมูลมีค่าผิดปกติปะปนในชุด ข้อมูลเป็นจำนวนมาก ๆ ตัวประมาณค่า \bar{y}_M ยังมีความเที่ยงตรงและความแม่นยำสูงกว่าตัวประมาณ ค่า $\bar{y}_{m(SD)}$ อย่างชัดเจนมากขึ้นด้วย

3. ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัว ประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M)

ผลการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร โดยใช้ตัวประมาณ ค่า \bar{y}_M และใช้ตัวแปรอุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์เป็นตัวแปรช่วยในการประมาณค่า การระเหยน้ำรายเดือนในจังหวัดสกลนคร จากเดือนพฤษภาคมถึงธันวาคม พ.ศ. 2556 จำนวน 8 เดือน และใช้ข้อมูลอุตุนิยมวิทยารายวันในแต่ละเดือนประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน เนื่องจากค่าการระเหยน้ำในชั้นบรรยากาศ ไม่สามารถวัดได้โดยตรงและวัดค่าให้ถูกต้องนั้นทำได้ยาก จึงกำหนดค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือน ของเดือนพฤษภาคมถึงเดือนธันวาคม พ.ศ. 2556 ด้วยค่าเฉลี่ยเฉพาะเดือนนั้นจากข้อมูลอุตุนิยมวิทยารายเดือนย้อนหลัง 6 ปี เริ่มจากเดือน พฤษภาคม พ.ศ. 2550 ถึงธันวาคม พ.ศ. 2555 จะได้ค่าพารามิเตอร์การระเหยน้ำโดยเฉลี่ยรายเดือน ของเดือนพฤษภาคมถึงธันวาคม พ.ศ. 2556 เท่ากับ เท่ากับ 4.8005, 4.0097, 3.8168, 3.3365, 3.7590, 3.8089, 3.9688 และ 3.5847 มิลลิเมตร ตามลำดับ

หลังจากนั้นจึงประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือน โดยใช้ตัวประมาณค่าที่ \bar{y}_M นำเสนอ ผลปรากฏว่า ค่าประมาณการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วน ของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) มีค่าใกล้เคียงกับค่าพารามิเตอร์มากที่สุดในเดือนพฤษภาคม สิงหาคม กรกฎาคมและ ตามลำดับ

เมื่อพิจารณาความแม่นยำในการประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนของจังหวัด สกลนครในภาพรวม ด้วยเกณฑ์ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย ($MPRE$) ไม่เกิน

10% ผลปรากฏว่า การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ย เท่ากับ 6.71% (ไม่เกิน 10%) ซึ่งสอดคล้องเป็นไปตามสมมติฐานการวิจัย สรุปว่าตัวประมาณค่า \bar{y}_M สามารถใช้ประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนครได้

อภิปรายผล

การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนจากข้อมูลอุตุนิยมหาวิทยาลัยรายวันนั้น ในบางวันค่าการระเหยน้ำแบบภาคที่วัดได้ อาจจะมีค่าสูงมากกว่าวันอื่น ๆ ซึ่งอาจเป็นค่าผิดปกติในชุดข้อมูลรายเดือน การประมาณค่าการระเหยน้ำโดยเฉลี่ยด้วยค่าเฉลี่ยเลขคณิตอาจไม่เหมาะสม หากข้อมูลนั้นมีค่าผิดปกติปะปนอยู่ ดังนั้นผู้วิจัยจึงได้พัฒนาตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ขึ้นจากตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) โดยใช้การวิเคราะห์ถดถอยแรงที่ประมาณค่าสัมประสิทธิ์สหสัมพันธ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดและสัมประสิทธิ์ความเบ้ของตัวแปรช่วย เพื่อให้ได้ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ที่มีความเหมาะสมในการประมาณค่าเฉลี่ยประชากรเมื่อข้อมูลมีค่าผิดปกติปะปนอยู่

ผลการจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพในการประมาณค่าเฉลี่ยประชากรเมื่อใช้ตัวแปรช่วยสองตัว ภายใต้ 12 สถานการณ์ ทั้งนี้ผู้วิจัยกำหนดลักษณะข้อมูลการจำลองให้สอดคล้องกับลักษณะจริงของข้อมูลการระเหยน้ำแบบภาคกับตัวแปรช่วยสองตัวที่ผ่านการคัดเลือก นั่นคือ อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ โดยตัวแปรช่วยสองตัวนี้มีความสัมพันธ์กับการระเหยน้ำแบบภาคอยู่ในระดับปานกลาง ซึ่งผลการคัดเลือกตัวแปรช่วยนี้มีความสอดคล้องกับการเลือกตัวแปรนำเข้าสมการถดถอยพหุคูณของการวิจัยของ Shirsath and Singh (2010) ที่ได้คัดเลือกตัวแปรอิสระที่สัมพันธ์กับการระเหยน้ำแบบภาค สามอันดับแรก ได้แก่ อุณหภูมิสูงสุด ความเร็วลม และความชื้นสัมพัทธ์ ตามลำดับ แต่ในการวิจัยนี้ผู้วิจัยเลือกใช้อุณหภูมิอากาศเฉลี่ย ซึ่งก็มีความเกี่ยวข้องกับอุณหภูมิสูงสุด และยังสอดคล้องกับการเลือกตัวแปรนำเข้าตัวแบบเพื่อประมาณค่าการระเหยน้ำแบบภาครายวันในประเทศเกาหลีใต้ของ Kim et al. (2013) ที่ว่า การใช้ตัวแปรนำเข้าหรือตัวแปรอิสระทุก ๆ ตัวเข้าตัวแบบ สามารถอธิบายความผันแปรของการระเหยน้ำแบบภาคได้มากถึงร้อยละ 80 ขึ้นไป แต่หากลดตัวแปรนำเข้าลงเหลือเพียงตัวแปรอิสระที่เป็นกลุ่มตัวแปรที่เกี่ยวข้องกับอุณหภูมิเท่านั้น ตัวแบบประหยัดที่ได้ก็ยังสามารถอธิบายความผันแปรของการระเหยน้ำแบบภาคได้มากถึงร้อยละ 80 ขึ้นไป เช่นเดียวกับตัวแบบที่ใช้ทุก ๆ ตัวแปรนำเข้าตัวแบบ ฉะนั้นการคัดเลือกตัวแปรช่วยสองตัวนี้จึงมีความเหมาะสมมากกว่าตัวแปรอุตุนิยมหาวิทยาลัยอื่น ๆ

การทดสอบรูปแบบการแจกแจงของตัวแปรช่วยอุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์สามารถสรุปได้ว่า อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ ต้องทำการแปลงค่าข้อมูลก่อน จึงสามารถทดสอบการแจกแจงปกติและระบุเป็นการแจกแจงปกติได้ ข้อค้นพบที่ได้ในการจำลองข้อมูลพบว่า ไม่ว่าจะสุ่มตัวอย่างขนาดเท่าใดก็ตาม ร้อยละของค่าผิดปกติที่เพิ่มขึ้น จากร้อยละ 5 เพิ่มเป็นร้อยละ 10 ร้อยละ 15 และร้อยละ 20 ตามลำดับ จะพบว่า ค่า MSE และค่า $MPRE$ ของทั้ง

ตัวประมาณค่า \bar{y}_M และตัวประมาณค่า $\bar{y}_{m(SD)}$ มีค่าเพิ่มขึ้น ทั้งของตัวประมาณค่า \bar{y}_M และตัวประมาณค่า $\bar{y}_{m(SD)}$ ซึ่งสอดคล้องกับผลการศึกษาของภทรวรรณ แสงนวกิจ (2553) ดังนั้น สรุปได้ว่า ร้อยละของค่าผิดพลาดที่ปะปนในชุดข้อมูล ส่งผลกระทบต่อประสิทธิภาพของตัวประมาณค่า ทั้งในด้านความเที่ยงตรง (Precision) และความแม่นยำ (Accuracy) ทำให้ความเที่ยงตรงและความแม่นยำของค่าประมาณมีค่าลดลง

ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าในด้านความเที่ยงตรง (Precision) และความแม่นยำ (Accuracy) ของตัวประมาณค่า \bar{y}_M กับตัวประมาณค่า $\bar{y}_{m(SD)}$ ในสถานการณ์จำลอง 12 สถานการณ์ ได้ข้อค้นพบว่า ค่า *MSE* และค่า *MPRE* ของตัวประมาณค่า \bar{y}_M มีค่าน้อยกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุก ๆ สถานการณ์ นั่นคือ ตัวประมาณค่า \bar{y}_M มีประสิทธิภาพทั้งในด้านความเที่ยงตรงและความแม่นยำดีกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ ในทุกสถานการณ์ ซึ่งสอดคล้องกับผลการใช้สถิติแกร่งร่วมในตัวประมาณค่าแบบอัตราส่วน ดังผลการศึกษาของ Yarmohammadi and Mahmoudvand (2010), Kadilar, Candan and Cingi (2007) และภทรวรรณ แสงนวกิจ (2553) จึงกล่าวได้ว่า ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) ซึ่งพัฒนาจากตัวประมาณค่าแบบอัตราส่วน $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007) โดยใช้การวิเคราะห์ถดถอยแกร่งที่ประมาณค่าสัมประสิทธิ์สหสัมพันธ์การถดถอยด้วยวิธีกำลังสองตัดขอบน้อยสุดและสัมประสิทธิ์ความเบ้ของตัวแปรช่วย สามารถประมาณค่าเฉลี่ยของข้อมูลที่มีค่าผิดพลาดปะปนอยู่ได้แม่นยำและเที่ยงตรงมากขึ้น จึงทำให้ตัวประมาณค่า \bar{y}_M สามารถประมาณค่าการระเหยน้ำแบบภาคโดยเฉลี่ยได้ดีกว่าตัวประมาณค่าแบบ $\bar{y}_{m(SD)}$

ตามข้อค้นพบกล่าวไว้ข้างต้น จึงสรุปได้ว่า การประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) มีความเหมาะสมมากกว่าตัวประมาณค่า $\bar{y}_{m(SD)}$ เมื่อประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนด้วยตัวประมาณค่า \bar{y}_M จะได้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมพัทธ์โดยเฉลี่ยไม่เกิน 10% ซึ่งสอดคล้องเป็นไปตามสมมติฐานการวิจัย ถือว่า ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ (\bar{y}_M) มีความคลาดเคลื่อนอยู่ในเกณฑ์ยอมรับได้ มีความเหมาะสมที่ใช้ประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนได้

ข้อเสนอแนะ

ข้อเสนอแนะในการนำไปใช้

1. ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กรณีที่ใช้ตัวแปรช่วยสองตัวที่มีความสัมพันธ์กับตัวแปรที่ศึกษานี้ สามารถใช้ประมาณค่าเฉลี่ยประชากรที่มีค่าผิดพลาดปะปนอยู่ ภายใต้เงื่อนไข ความสัมพันธ์ของตัวแปรช่วยตัวหนึ่งต้องมีทิศทางเดียวกัน ($\rho_{yx} > 0$) และความสัมพันธ์ของตัวแปรช่วยตัวที่สองต้องมีทิศทางผกผัน ($\rho_{yz} < 0$) กัน

2. ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่นี้ สามารถประมาณค่าเฉลี่ยประชากรข้อมูลอื่น ๆ ที่อาจจะมีค่าผิดปกติปะปนอยู่ได้ ยกตัวอย่าง เช่น การประมาณค่าความสูงของบุคคลจากความยาวของการก้าวเดินปกติกับความยาวของเท้า การประมาณค่าความสูงของบุคคลจากความยาวของมือและความยาวของเท้า หรือการประมาณค่าอายุของเด็กจากปริมาณฮีโมโกลบิน A (Hemoglobin A) และปริมาณฮีโมโกลบิน F (Hemoglobin F) ซึ่งมีประโยชน์อย่างมากต่อการพิสูจน์เอกลักษณ์บุคคลหรือการระบุตัวบุคคล ในกระบวนการทางนิติวิทยาศาสตร์ หรือการประมาณค่าปริมาณขยะทั้งหมดในเขตพื้นที่ที่ศึกษา จากปริมาณขยะทั้งหมดในปีที่ผ่านมาและจำนวนครัวเรือนทั้งหมดในเขตพื้นที่ที่ศึกษาในปัจจุบัน ซึ่งมีประโยชน์ต่อการวางแผนจัดสรรงบประมาณ ตลอดจนวางแผนบริหารจัดการการกำจัดขยะ เพื่อลดมลพิษที่เกิดจากขยะ ส่งผลให้ประชาชนได้อาศัยในชุมชนที่มีสภาพแวดล้อมที่ดี เป็นต้น แต่อย่างไรก็ตาม การนำตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่นี้ไปใช้ได้ต้องมีประสิทธิภาพ ผู้ที่สนใจนำตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่นี้ไปใช้ ควรพิจารณาทิศทางของความสัมพันธ์ระหว่างตัวแปรศึกษากับตัวแปรช่วย และปริมาณค่าผิดปกติที่ปะปนอยู่ในชุดข้อมูลก่อน

ข้อเสนอแนะในการศึกษาต่อไป

1. ตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ กรณีที่ใช้ตัวแปรช่วยสองตัวที่มีความสัมพันธ์กับตัวแปรที่ศึกษานี้ มีปัจจัยที่คาดว่าจะส่งผลต่อประสิทธิภาพของตัวประมาณค่า นั่นคือ ระดับความสัมพันธ์ระหว่างตัวแปรช่วยกับตัวแปรที่ศึกษา ซึ่งในการวิจัยนี้ ตัวแปรช่วยมีความสัมพันธ์ในระดับปานกลางกับตัวแปรที่ศึกษา จึงอาจเป็นสาเหตุให้ประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่ ลดลง เมื่อสุ่มตัวอย่างขนาดใหญ่ ดังนั้นในการศึกษาครั้งต่อไป ในส่วนของการจำลองข้อมูล ควรศึกษากรณีที่ตัวแปรช่วยมีความสัมพันธ์กับตัวแปรที่ศึกษาในหลาย ๆ ระดับ เช่น ระดับกลางหรือระดับสูงมาก ($0.5 < \rho < 0.9$)

2. ในการประมาณค่าการระเหยน้ำรายเดือนด้วยตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยประชากรที่ปรับใหม่นี้ ใช้ตัวแปรช่วย คือ อุณหภูมิอากาศเฉลี่ยและความชื้นสัมพัทธ์ ถึงแม้ว่าตัวแปรช่วยสองตัวนี้จะมีความสัมพันธ์กับการระเหยน้ำมากกว่าข้อมูลอุตุนิยมนิวทียาอื่น ๆ แต่ก็มี ความสัมพันธ์อยู่ในระดับปานกลางเท่านั้น ดังนั้นในการศึกษาครั้งต่อไป ในทางปฏิบัติอาจจะเป็น การศึกษาในพื้นที่อื่น ควรพิจารณาเพิ่มตัวแปรอุตุนิยมนิวทียาอื่น ๆ เพื่อใช้คัดเลือกตัวแปรช่วยที่คาดว่าจะมีความสัมพันธ์กับตัวแปรที่ศึกษาในระดับมาก เช่น อุณหภูมิสูงสุดในแต่ละวัน ความเร็วลม เป็นต้น

บรรณานุกรม

- กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม. (2555). *รายงานสถานการณ์คุณภาพสิ่งแวดล้อมปี 2554 ภาคที่ 9 ลุ่มน้ำโขง เลย อุดรธานี หนองคาย บึงกาฬ สกลนคร นครพนม. อุดรธานี: สำนักงานสิ่งแวดล้อมภาคที่ 9.*
- กัลยา วาณิชย์บัญชา. (2554). *หลักสถิติ* (พิมพ์ครั้งที่ 13). กรุงเทพฯ: บริษัทธรรมสาร จำกัด.
- จุฑาทพร เนียมวงศ์ และธนภุต เตยานุรักษ์. (2558). การประมาณค่าเฉลี่ยปริมาณก๊าซไอโซน เขตอุตสาหกรรมแหลมฉบัง จ.ชลบุรี. ใน *การประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 7* (หน้า 1-8). พิษณุโลก: คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร.
- ประชุม สุวัตถิ. (2552). *การสำรวจด้วยตัวอย่าง: การชักตัวอย่างและการวิเคราะห์*. กรุงเทพฯ: สถาบันบัณฑิตพัฒนบริหารศาสตร์.
- พิศ คงบริรักษ์ และภฤติกา สืบศักดิ์. (2554). การระเหยของน้ำในประเทศไทยคาบ 10 ปี (พ.ศ. 2544–2553). ใน *เอกสารผลงานวิชาการอุดรนิยมวิทยา 2554*, (หน้า 30–38). กรุงเทพฯ: กรมอุตุนิยมวิทยา.
- ภทรวรรณ แสงนวกิจ. (2553). การประมาณค่าเฉลี่ยด้วยตัวประมาณค่าแบบอัตราส่วน เมื่อข้อมูลมีค่าผิดปกติ. ใน *การประชุมวิชาการสถิติและสถิติประยุกต์ ครั้งที่ 11 ประจำปี 2553*. เชียงใหม่: ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่.
- รังสรรค์ อาภาศัพท์กุล. (2547). *อุตุนิยมวิทยาเบื้องต้น*. กรุงเทพฯ: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- วิชัย ศรีบุญลือ. (2552). *อุทกวิทยา*. ขอนแก่น: โรงพิมพ์มหาวิทยาลัยขอนแก่น.
- สถานีอุตุนิยมวิทยา. (2556). *ความรู้ด้านอุตุนิยมวิทยา*. วันที่ค้นข้อมูล 26 พฤษภาคม 2556, เข้าถึงได้จาก <http://www.metsakon.tmd.go.th/bunyay.htm>.
- ส่วนจัดสรรน้ำ สำนักบริหารจัดการน้ำ. (2553). *การวัดการระเหยน้ำ*. วันที่ค้นข้อมูล 17 ธันวาคม 2555, เข้าถึงได้จาก <http://202.129.59.73/wm/Water/water4.1/mean.pdf>.
- สำนักงานเศรษฐกิจการเกษตร. (2557). *สถิติการเกษตรของประเทศไทย ปี 2556*. นนทบุรี: ชุมชนสหกรณ์การเกษตรแห่งประเทศไทย จำกัด สาขา 4.
- สุรินทร์ นิยมางกูร. (2546). *เทคนิคการสุ่มตัวอย่าง* (พิมพ์ครั้งที่ 5). กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.
- Abu-Dayyeh, W. A., Ahmed, M. S., Ahmed, R. A., & Muttalak, H. A. (2003). Some estimators of a finite population mean using auxiliary information. *Applied Mathematics and Computation*, 139: 287–298.
- Andrew, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H., & Tukey, J. W. (1972). *Robust Estimates of Location: Survey and Advances*. Princeton, NJ: Princeton University Press.
- Burden, R. L. & Faires, J. D. (2011). *Numerical Analysis* (9th ed.). Canada: Cengage Learning.

- Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: Wiley.
- Dabral P. P., Jhajharia D., Mishra P., Hangshing L., & Doley, B. J. (2014). Time series modeling of pan evaporation: A case study in the northeast India. *Global NEST journal*, 16(2); 280-292.
- Devore, J. L. (2012). *Probability and Statistics for Engineering and the Sciences* (8th ed.). Canada: Cengage Learning.
- Dodangeh, S., Abedi, K. J., & Gohari, S. A. (2012). Application of time series modeling to investigate future climatic parameters trend for water resources management purposes. *JWSS-Isfahan University of Technology*, 16(59): 59-74.
- Dunn, O. J., & Clark, V. A. (2009). *Basic Statistics: A Primer for Biomedical Sciences* (4th ed.). New Jersey: John Wiley & Sons Inc.
- Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical Distributions* (4th ed.). New York: John Wiley & Sons, Inc.
- Goel, A. (2009). ANN based modeling for prediction of evaporation in reservoirs. *IJE Transactions A: Basics*, 22(4); 351-8.
- Gupta, S., & Shabbir, J. (2007). On the use of transformed auxiliary variables in estimating population mean by using two auxiliary variables. *Journal of Statistical Planning and Inference*, 137; 1606–1611.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). New Jersey: Pearson Prentice Hall.
- Hendra, R., & Staum, P. W. (2010). A SAS® application to identify and evaluate outliers. *Proceeding of the 23rd Annual Conference of NorthEast SAS Users Group*. Baltimore, Maryland.
- Irmak, S., & Haman, D. Z. (2003). Evaluation of five methods for estimating class A pan evaporation in a humid climate. *HortTechnology Journal*, 13(3); 500-508.
- Kadilar, C., & Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, 151(3); 893–902.
- Kadilar, C., & Cingi, H. (2005). A new estimators using two auxiliary variables, *Applied Mathematics and Computation*, 162; 901–908.
- Kadilar, C., & Cingi, H. (2006). Improved in estimating the population mean in simple random sampling. *Applied Mathematics Letters*, 19; 75–79.
- Kadilar, C., Candan, M., & Cingi, H. (2007). Ratio estimators using robust regression. *Hacettepe Journal of Mathematics and Statistics*, 36(2); 181–188.

- Kim, S., Shiri, J., Kisi, O., & Singh, V. P. (2013). Estimating daily pan evaporation using different data-driven methods and lag-time patterns. *Water Resources Management*, 27(7); 2267-86.
- Kişi, Ö. (2006). Daily pan evaporation modeling using a neuro-fuzzy computing technique. *J. Hydrology*, 329; 636– 46.
- Meshram D. T., Gorantiwwar S. D., Kkulkarni A. D., & Hangaargekar P. A. (2013). Forecasting for Evaporation for Manki reservoir in Osmanabad district of Maharashtra, India. *Inter. J. of Advanced Technology in Civil Engineer*, 2(2); 19-23.
- Moshrik, R. H., Ahmed, N. B., & Zeyad, S. T. (2008). Developing reference crop evapotranspiration time series simulation model using class a pan: A case study for the Jordan Valley/Jordan. *Jordan J. Earth and Envir. Sci*, 1(1); 33-44.
- Oral, E. (2010). Improving efficiency of ratio-type estimators through order statistics. *Section on Survey Research Methods of the 2010 Joint Statistical Meetings*, Vancouver, British Columbia: July 31 - August 5.
- Perri, P. F. (2005). Combining two auxiliary variables in ratio-cum-product type estimator. In *Proceeding of Italian Statistical Society, Intermediate meeting on Statistics and Environment, Messina, 21-23 September, 2005* (pp. 193-196). Italy: Italian Statistical Society.
- Rousseeuw, P. J., & Leroy, A. M. (2003). *Robust Regression and Outlier Detection*. New Jersey: John Wiley & Sons, Inc.
- Ryan, T. P. (2007). *Modern Engineering Statistics*. New York: John Wiley & Sons, Inc.
- Ryan, T. P. (2008). *Modern Regression Methods* (2nd ed.). New York: Wiley-Interscience.
- Saengprasan, C., Chadcham, S., & Jermjitpornchai, S. (2014). On the comparison of efficiency between Multiple Regression and Box-Jenkins Methods for evaporation forecasting. In *Burapha University International Conference 2014 (BUU2014), July 3-4, 2014, Pattaya Chonburi* (pp. 341-348). Thailand: Burapha University.
- Sharma, B., & Tailor, R. (2010). A new ratio-cum-dual to ratio estimator of finite population mean in simple random sampling. *Global Journal of Science Frone r Research*, 10(1); 27–31.
- Shirsath, P. B., & Singh, A. K. (2010). A comparative study of daily pan evaporation estimation using ANN, regression and climate based models. *Water Resources Management*, 24; 1571–81.

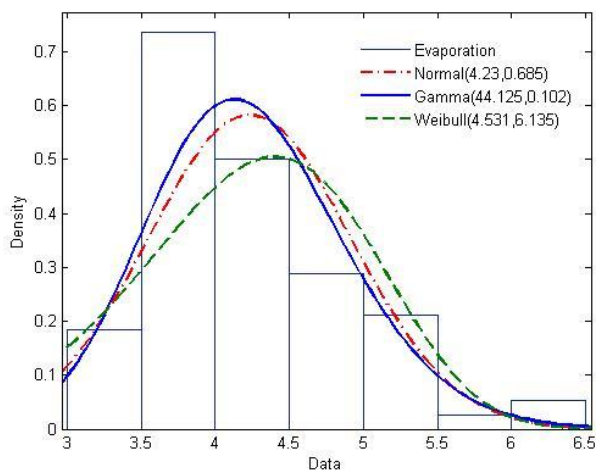
- Singh, H. P., & Kakran, M. S. (1993). *A modified ratio estimator using known coefficient of kurtosis of an auxiliary character*. unpublished paper.
- Sisodia, B. V. S., & Dwivedi, V. K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable. *Journal of Indian Society Agricultural Statistics*, 33(2); 13–18.
- Thompson, S. A. (1999). *Hydrology for water management*. Netherlands: A.A.Balkema.
- Upadhyaya, L. N., & Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical J.*, 41(5); 627–636.
- Verboven, S., Engelen, S., & Smets, N. (2006). *LIBRA: the Matlab Library for Robust Analysis*. Retrieved March 15, 2013, from <http://wis.kuleuven.be/stat/robust.html>.
- Wilcox, R. R. (2009). *Basic Statistics: Understanding Conventional Methods and Modern Insights*. New York: Oxford University Press, Inc.
- Yan, Z., & Tian, B. (2010). Ratio method to the mean estimation using coefficient of skewness of auxiliary variables. *In International Conference on Information Computer Application (ICICA 2010) Information Computing and Applications Part II*, October 15-18, 2010, Tangshan, China.
- Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, 81(12); 2141-2155, DOI:10.1080/00949655.2010.520163
- Yarmohammadi, M., & Mahmoudvand, R. (2010). The effect of outliers on robust and resistant coefficient of determination in the linear regression models. *International Journal of Academic Research*, 2(3); 133–138.

ภาคผนวก

ภาคผนวก ก

ผลการทดสอบการแจกแจงของการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์

ภาคผนวก ก-1 ผลการทดสอบการแจกแจงของการระเหยน้ำด้วยสถิติ Anderson-Darling test



Distribution ID Plot for Evaporation

Descriptive Statistics

N	N*	Mean	StDev	Median	Minimum	Maximum	Skewness	Kurtosis
76	0	4.23144	0.684751	4.086	3.1713	6.0513	0.864756	0.107413

Box-Cox transformation: Lambda = -2

Johnson transformation function:

$$-2.04883 + 1.56854 * \text{Asinh} \left(\frac{X - 3.26962}{0.435861} \right)$$

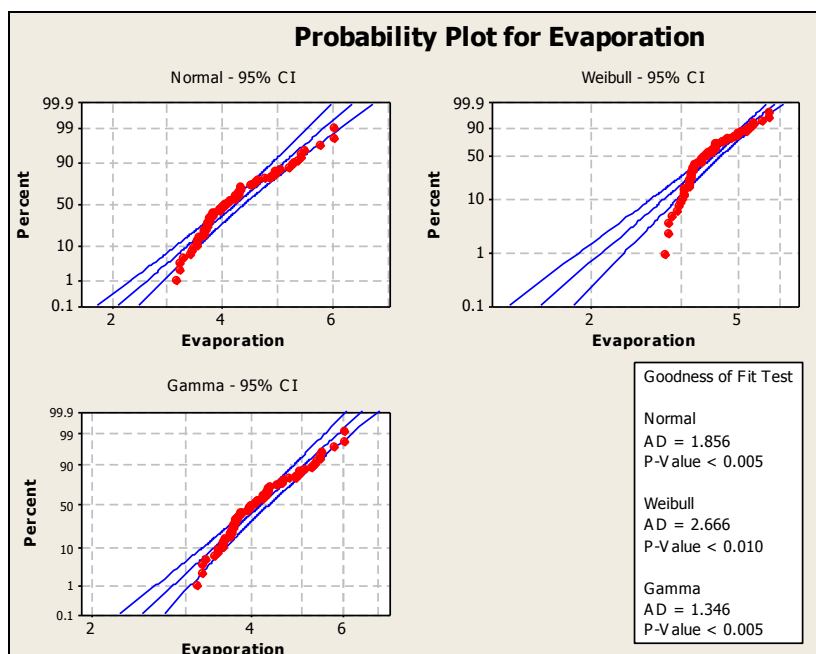
Goodness of Fit Test

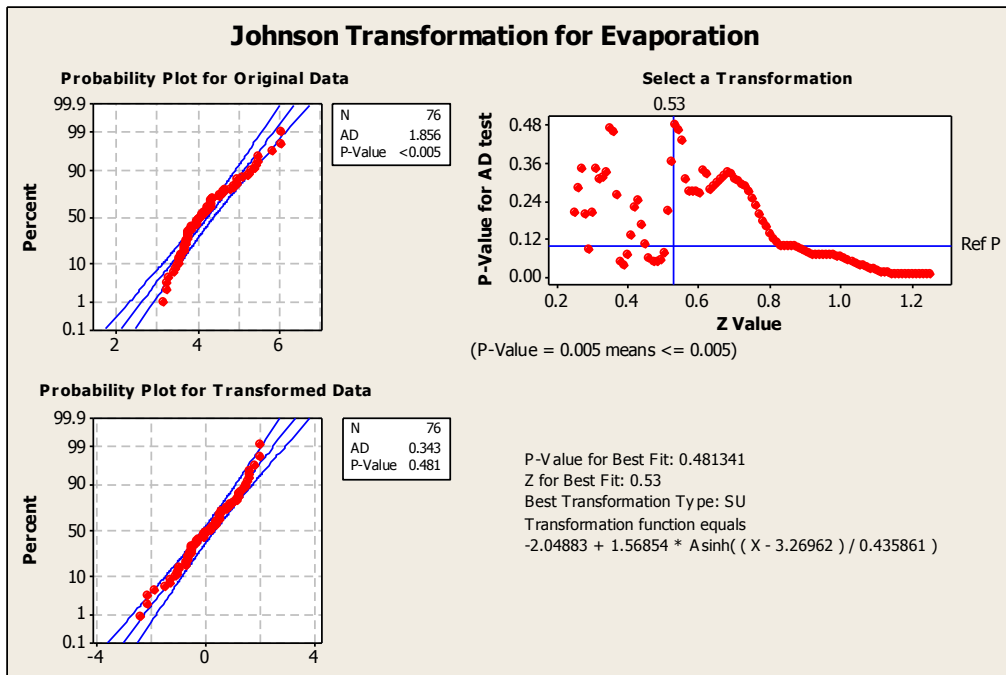
Distribution	AD	P	LRT	P
Normal	1.856	<0.005		
Box-Cox Transformation	0.406	0.342		
Lognormal	1.109	0.006		
3-Parameter Lognormal	0.369	*	0.008	
Exponential	25.210	<0.003		
2-Parameter Exponential	4.497	<0.010	0.000	
Weibull	2.666	<0.010		
3-Parameter Weibull	0.482	0.239	0.000	
Smallest Extreme Value	3.783	<0.010		
Largest Extreme Value	0.540	0.178		
Gamma	1.346	<0.005		
3-Parameter Gamma	0.529	*	0.003	
Logistic	1.515	<0.005		
Loglogistic	1.008	<0.005		
3-Parameter Loglogistic	0.395	*	0.008	
Johnson Transformation	0.343	0.481		

ML Estimates of Distribution Parameters

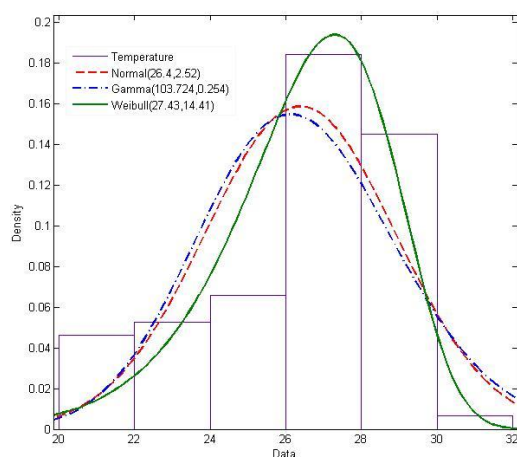
Distribution	Location	Shape	Scale	Threshold
Normal*	4.23144		0.68475	
Box-Cox Transformation*	0.05984		0.01715	
Lognormal*	1.43037		0.15494	
3-Parameter Lognormal	0.33351		0.44038	2.69571
Exponential			4.23144	
2-Parameter Exponential			1.07428	3.15716
Weibull		6.13511	4.53148	
3-Parameter Weibull		1.65533	1.22499	3.13647
Smallest Extreme Value	4.59619		0.76623	
Largest Extreme Value	3.92211		0.51785	
Gamma		41.25076	0.10258	
3-Parameter Gamma		4.45581	0.31595	2.82358
Logistic	4.15750		0.38359	
Loglogistic	1.41862		0.08887	
3-Parameter Loglogistic	0.17992		0.29841	2.87963
Johnson Transformation*	0.08124		1.01960	

* Scale: Adjusted ML estimate





ภาคผนวก ก-2 ผลการทดสอบการแจกแจงของอุณหภูมิอากาศเฉลี่ยด้วยสถิติ Anderson-Darling test



Distribution ID Plot for Temperature

Descriptive Statistics

N	N*	Mean	StDev	Median	Minimum	Maximum	Skewness	Kurtosis
76	0	26.3830	2.51720	27.32	20.2151	30.3817	-0.959972	0.0350835

Box-Cox transformation: Lambda = 5.00000

Johnson transformation function:

$1.22505 + 0.929455 * \text{Asinh} \left(\frac{X - 28.8112}{0.962180} \right)$

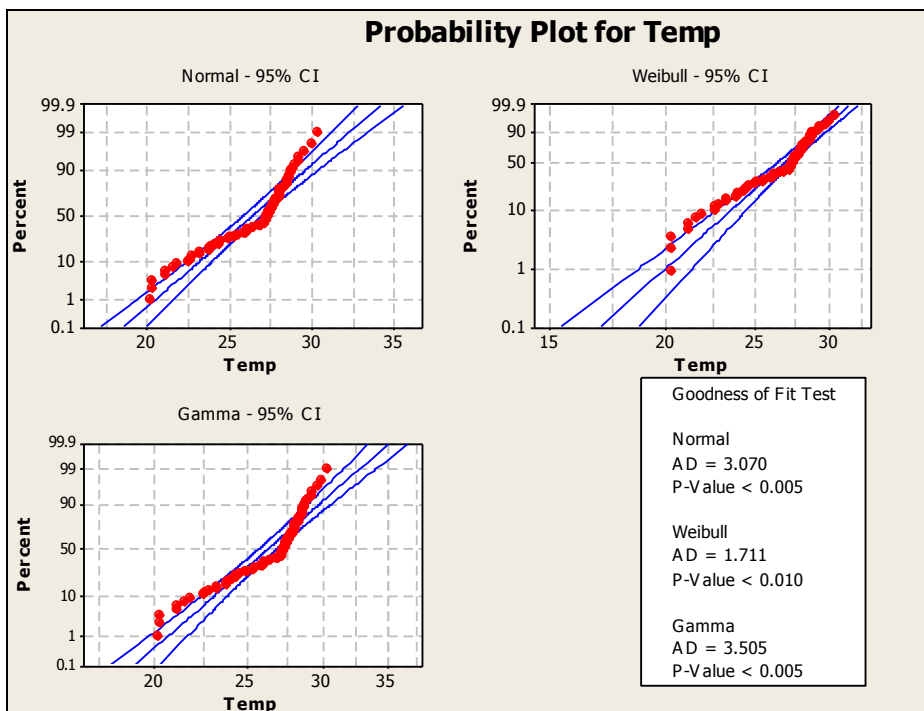
Goodness of Fit Test

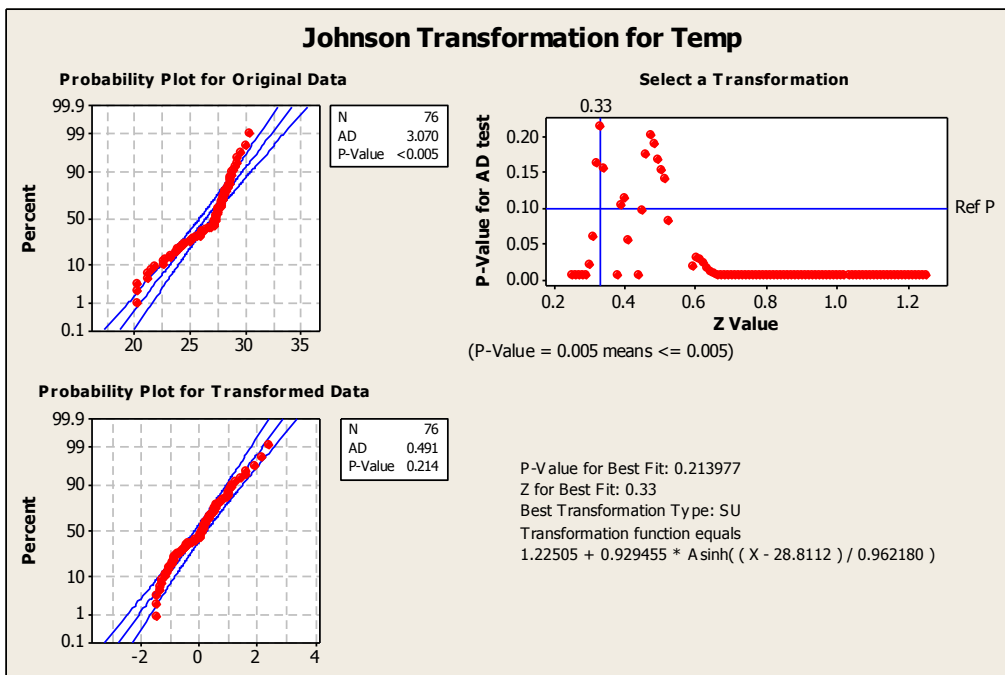
Distribution	AD	P	LRT P
Normal	3.070	<0.005	
Box-Cox Transformation	1.335	<0.005	
Lognormal	3.699	<0.005	
3-Parameter Lognormal	3.094	*	0.006
Exponential	28.899	<0.003	
2-Parameter Exponential	12.621	<0.010	0.000
Weibull	1.711	<0.010	
3-Parameter Weibull	1.297	<0.005	0.109
Smallest Extreme Value	1.296	<0.010	
Largest Extreme Value	4.852	<0.010	
Gamma	3.505	<0.005	
3-Parameter Gamma	7.435	*	1.000
Logistic	2.616	<0.005	
Loglogistic	3.064	<0.005	
3-Parameter Loglogistic	2.617	*	0.016
Johnson Transformation	0.491	0.214	

ML Estimates of Distribution Parameters

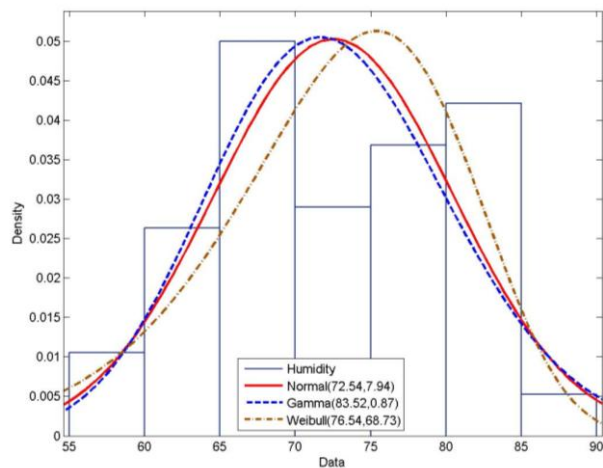
Distribution	Location	Shape	Scale	Threshold
Normal*	26.38296		2.51720	
Box-Cox Transformation*	1.38432E+07		5.41469E+06	
Lognormal*	3.26789		0.10078	
3-Parameter Lognormal	9.24873		0.00024	-1.03650E+04
Exponential			26.38296	
2-Parameter Exponential			6.25009	20.13286
Weibull		14.40995	27.43124	
3-Parameter Weibull		6239.06811	11383.33577	-1.13558E+04
Smallest Extreme Value	27.50095		1.82437	
Largest Extreme Value	25.03221		2.82066	
Gamma		103.72355	0.25436	
3-Parameter Gamma		924.68551	0.08313	-51.13271
Logistic	26.71386		1.40051	
Loglogistic	3.28270		0.05494	
3-Parameter Loglogistic	9.25481		0.00013	-1.04280E+04
Johnson Transformation*	0.03141		0.91680	

* Scale: Adjusted ML estimate





ภาคผนวก ก-3 ผลการทดสอบการแจกแจงของความชื้นสัมพัทธ์ด้วยสถิติ Anderson-Darling test



Distribution ID Plot for Humidity

Descriptive Statistics

N	N*	Mean	StDev	Median	Minimum	Maximum	Skewness	Kurtosis
76	0	72.5394	7.93660	71.2661	55.6169	85.6708	-0.0574576	-1.14708

Box-Cox transformation: Lambda = 1

Johnson transformation function:

$$-0.251741 + 0.711345 * \ln((x - 54.7732) / (86.2413 - x))$$

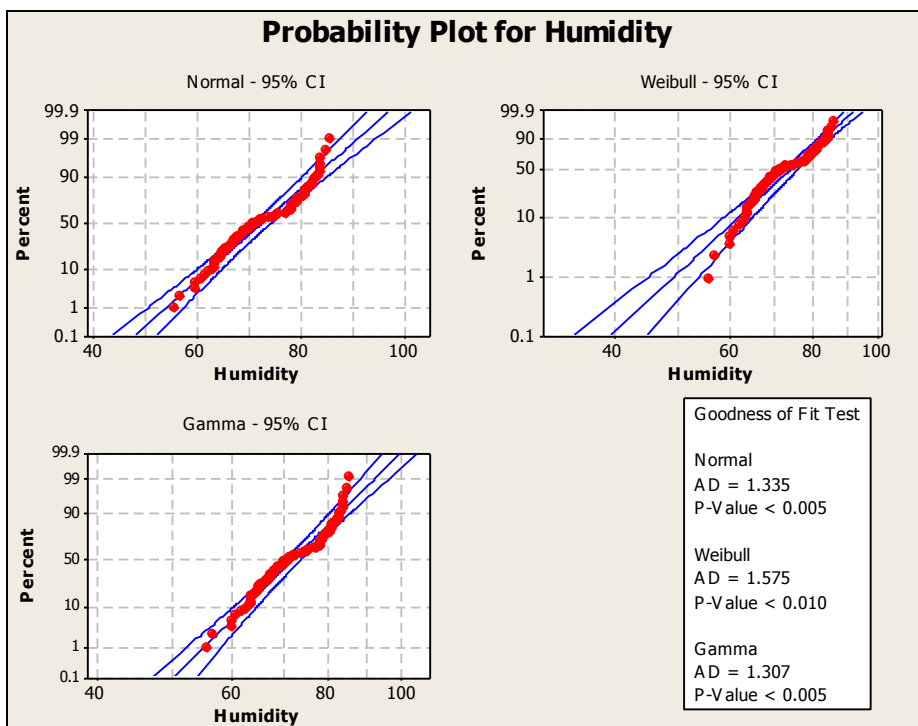
Goodness of Fit Test

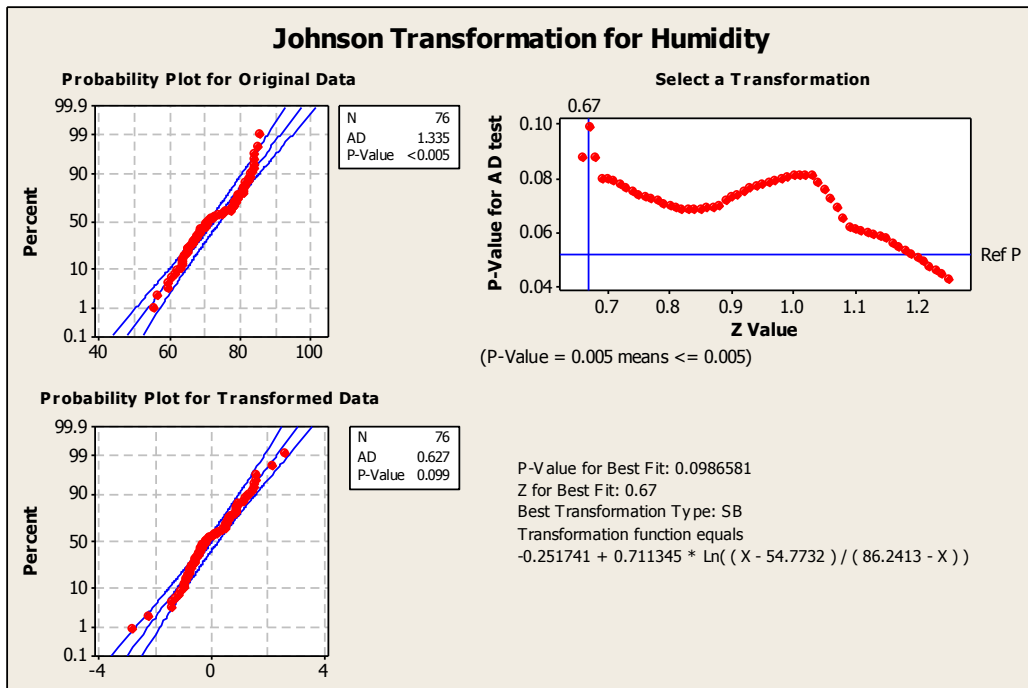
Distribution	AD	P	LRT P
Normal	1.335	<0.005	
Box-Cox Transformation	1.335	<0.005	
Lognormal	1.254	<0.005	
3-Parameter Lognormal	1.372	*	0.339
Exponential	27.750	<0.003	
2-Parameter Exponential	9.121	<0.010	0.000
Weibull	1.575	<0.010	
3-Parameter Weibull	1.278	<0.005	0.083
Smallest Extreme Value	1.733	<0.010	
Largest Extreme Value	1.212	<0.010	
Gamma	1.307	<0.005	
3-Parameter Gamma	1.373	*	1.000
Logistic	1.497	<0.005	
Loglogistic	1.418	<0.005	
3-Parameter Loglogistic	1.483	*	0.619
Johnson Transformation	0.627	0.099	

ML Estimates of Distribution Parameters

Distribution	Location	Shape	Scale	Threshold
Normal*	72.53936		7.93660	
Box-Cox Transformation*	72.53936		7.93660	
Lognormal*	4.27813		0.11074	
3-Parameter Lognormal	10.25762		0.00028	-2.84264E+04
Exponential			72.53936	
2-Parameter Exponential			17.14805	55.39127
Weibull		10.55955	76.08396	
3-Parameter Weibull		3.26892	25.52286	49.71975
Smallest Extreme Value	76.44409		7.00555	
Largest Extreme Value	68.60007		7.34572	
Gamma		83.51839	0.86854	
3-Parameter Gamma		114.41989	0.74087	-12.49360
Logistic	72.52387		4.80763	
Loglogistic	4.27999		0.06662	
3-Parameter Loglogistic	6.08774		0.01091	-367.93771
Johnson Transformation*	0.01475		0.98213	

* Scale: Adjusted ML estimate





ภาคผนวก ข

ผลการวิเคราะห์การถดถอยพหุคูณเบื้องต้น
ของการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์
(เพื่อกำหนดค่าสัมประสิทธิ์การถดถอยในสมการการจำลองข้อมูลประชากร)

ภาคผนวก ข-1 ผลการวิเคราะห์การถดถอยพหุคูณเบื้องต้น ของการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	ความชื้น (%), อุณหภูมิอากาศ Temperature (ซ)	.	Enter

a. All requested variables entered.

b. Dependent Variable: น้ำระเหยแบบถาด Evaporation (มม.)

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.885 ^a	.783	.777	.3235840	1.474

a. Predictors: (Constant), ความชื้น (%), อุณหภูมิอากาศ Temperature (ซ)

b. Dependent Variable: น้ำระเหยแบบถาด Evaporation (มม.)

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27.522	2	13.761	131.427	.000 ^a
	Residual	7.644	73	.105		
	Total	35.166	75			

a. Predictors: (Constant), ความชื้น (%), อุณหภูมิอากาศ Temperature (ซ)

b. Dependent Variable: น้ำระเหยแบบถาด Evaporation (มม.)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.877	.455		8.518	.000
	อุณหภูมิอากาศ Temperature (ซ)	.199	.016	.731	12.705	.000
	ความชื้น (%)	-.067	.005	-.782	-13.585	.000

a. Dependent Variable: น้ำระเหยแบบถาด Evaporation (มม.)

ภาคผนวก ค

ข้อมูลการระเหยน้ำ อุณหภูมิอากาศเฉลี่ย และความชื้นสัมพัทธ์
ตั้งแต่เดือนพฤษภาคม 2556 ถึงเดือนธันวาคม 2556

ตารางที่ ค-1 ค่าสถิติของข้อมูลอุตุวิทยารายวันเดือนพฤษภาคมและมิถุนายน 2556

วันที่	พฤษภาคม			มิถุนายน		
	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์
1	4.470	26.950	75.125	5.060	30.525	71.000
2	6.840	27.938	72.000	5.070	28.900	74.250
3	6.030	26.613	75.125	2.930	27.413	81.875
4	2.810	24.313	88.750	1.960	27.600	85.750
5	5.310	27.313	76.750	0.000	26.463	87.750
6	6.480	28.400	75.250	2.850	27.163	84.750
7	3.700	26.225	81.625	3.950	28.738	83.500
8	5.200	27.400	81.375	5.230	30.063	73.500
9	0.000	27.675	80.000	4.550	31.163	68.000
10	6.130	28.513	80.500	0.000	29.975	74.500
11	4.000	27.538	77.875	2.110	25.138	86.875
12	2.550	27.425	81.250	2.640	26.738	76.000
13	5.580	29.900	78.875	7.120	28.413	65.125
14	1.190	30.363	75.000	4.370	27.938	68.250
15	5.950	31.863	72.500	6.440	28.563	63.750
16	6.490	32.750	64.500	3.910	29.300	68.125
17	5.660	30.313	70.375	2.250	28.963	73.375
18	5.990	28.875	76.000	5.430	29.438	77.500
19	5.400	28.075	82.000	5.720	29.950	74.250
20	1.480	27.338	85.000	2.700	28.425	81.375
21	4.910	28.225	81.375	3.840	27.325	85.125
22	3.000	26.588	90.625	3.650	27.300	81.750
23	3.520	26.525	89.625	4.420	25.463	91.875
24	3.620	26.238	90.125	1.350	27.625	78.625
25	4.930	26.950	86.125	3.540	27.488	86.500
26	4.910	29.638	72.375	1.370	25.013	91.375
27	6.390	30.538	68.875	3.100	28.713	76.000
28	6.550	29.013	81.375	0.000	28.088	81.375
29	4.950	27.175	85.000	4.990	29.663	72.125
30	2.900	27.963	79.875	6.890	29.100	79.500
31	5.580	29.475	74.125			
Correlation*	-	0.400	-0.486	-	0.422	-0.511

หมายเหตุ* หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างการระเหยน้ำกับตัวแปรช่วย

ตารางที่ ค-2 ค่าสถิติของข้อมูลอุตุวิทยารายวันเดือนกรกฎาคมและสิงหาคม 2556

วันที่	กรกฎาคม			สิงหาคม		
	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์
1	4.680	30.025	75.750	5.790	27.913	83.750
2	4.310	29.250	74.000	3.930	27.450	81.250
3	0.000	28.388	80.750	1.840	27.638	84.375
4	1.830	27.838	81.875	0.000	26.700	82.625
5	5.760	29.413	73.750	3.740	27.375	83.000
6	4.810	28.888	77.625	6.640	29.350	79.625
7	4.180	28.213	79.125	1.230	26.063	86.875
8	4.690	27.950	77.625	1.510	26.613	83.750
9	4.420	27.000	81.250	2.550	26.313	88.875
10	3.170	26.638	83.125	0.000	27.488	84.500
11	1.600	27.363	79.500	5.960	26.838	85.500
12	6.110	27.575	81.875	8.550	27.925	77.375
13	5.700	28.625	78.750	4.270	28.038	80.625
14	2.970	27.825	83.250	3.200	27.938	80.125
15	3.550	27.988	84.625	5.880	29.113	73.750
16	5.050	27.350	85.625	6.690	29.025	72.250
17	0.510	26.988	87.000	6.360	28.650	68.250
18	5.190	28.600	79.125	3.000	28.825	68.750
19	0.000	27.013	86.625	5.620	29.250	68.875
20	0.000	26.013	91.375	3.470	27.613	78.000
21	4.380	26.500	88.250	1.850	25.738	88.125
22	3.070	27.288	85.625	3.230	24.938	90.375
23	2.730	27.563	83.375	2.320	26.050	84.375
24	0.000	25.713	91.875	1.900	26.600	82.625
25	1.710	26.175	86.250	0.730	27.413	82.000
26	5.090	26.763	83.500	2.080	26.913	85.625
27	2.900	25.925	93.375	4.730	28.200	82.250
28	0.000	24.313	91.375	3.520	28.763	80.000
29	0.000	24.688	92.125	3.660	28.150	81.625
30	3.090	26.100	86.500	3.740	27.513	80.500
31	3.340	27.063	85.625	4.320	28.263	77.750
Correlation*	-	0.604	-0.631	-	0.583	-0.523

หมายเหตุ* หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างการระเหยน้ำกับตัวแปรช่วย

ตารางที่ ค-3 ค่าสถิติของข้อมูลอุตุณิยมวิทยารายวันเดือนกันยายนและตุลาคม 2556

วันที่	กันยายน			ตุลาคม		
	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์
1	4.130	28.675	79.250	2.820	25.800	81.375
2	1.810	28.138	82.500	4.200	27.475	77.375
3	3.180	28.588	80.875	3.620	27.650	73.500
4	4.060	28.688	78.250	4.860	27.400	68.500
5	1.220	26.925	86.000	3.530	26.925	70.250
6	3.700	27.025	81.250	4.480	27.100	72.375
7	2.750	27.550	80.000	4.400	26.575	70.250
8	2.880	26.413	83.500	4.460	26.188	63.500
9	4.820	28.475	75.750	5.240	26.388	66.000
10	5.680	28.288	75.875	4.490	27.125	69.500
11	5.940	27.800	78.250	4.940	27.125	73.625
12	1.910	27.325	84.000	2.500	27.550	74.625
13	0.000	25.688	90.125	5.160	28.238	71.750
14	5.040	27.063	85.500	4.710	28.275	71.500
15	0.000	27.225	84.625	4.910	26.950	70.625
16	6.580	27.738	82.250	1.760	24.113	92.625
17	2.770	27.213	84.875	5.860	26.788	78.000
18	1.870	26.950	87.750	4.960	26.850	74.875
19	3.630	26.200	90.750	2.920	26.563	72.250
20	4.650	26.763	85.500	3.360	25.863	72.875
21	3.730	26.575	86.375	4.490	25.600	73.875
22	2.250	26.563	88.125	3.920	25.725	75.750
23	5.250	27.625	82.875	4.000	26.113	72.875
24	2.050	27.513	80.375	5.280	25.675	69.375
25	2.560	27.450	82.000	1.750	23.213	71.375
26	4.870	27.050	82.625	4.760	23.725	66.625
27	3.490	27.013	77.125	3.630	23.563	69.750
28	3.440	26.900	76.875	3.310	24.363	72.125
29	4.570	26.825	69.750	3.120	25.600	72.000
30	0.000	24.800	75.375	4.500	26.750	72.750
31				3.360	27.475	70.875
Correlation*	-	0.477	-0.319	-	0.436	-0.433

หมายเหตุ* หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างการระเหยน้ำกับตัวแปรช่วย

ตารางที่ ค-4 ค่าสถิติของข้อมูลอุตุนิยมวิทยารายวันเดือนพฤศจิกายนถึงธันวาคม 2556

วันที่	พฤศจิกายน			ธันวาคม		
	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์	การระเหยน้ำ	อุณหภูมิอากาศเฉลี่ย	ความชื้นสัมพัทธ์
1	5.520	26.875	69.500	4.380	21.600	64.625
2	3.880	26.600	66.750	2.270	19.700	62.750
3	4.300	25.625	66.875	4.020	20.375	65.250
4	4.020	25.050	68.250	3.030	21.500	67.875
5	3.630	25.725	68.125	3.870	22.450	66.875
6	3.450	26.175	70.125	2.910	20.800	69.625
7	1.550	24.975	80.750	2.860	20.113	67.000
8	4.080	27.063	71.500	3.640	20.438	69.125
9	3.060	26.675	67.625	1.670	21.375	70.375
10	4.220	26.738	68.125	4.330	21.438	69.000
11	2.690	28.213	72.000	2.880	22.538	67.625
12	4.720	28.513	73.250	1.770	23.025	71.000
13	6.290	27.913	68.750	3.600	24.125	70.375
14	4.180	26.000	68.000	3.950	23.988	67.875
15	3.230	24.825	69.625	4.070	22.613	73.250
16	2.890	24.375	73.625	3.020	15.938	93.000
17	3.200	24.088	73.375	1.870	16.538	75.875
18	4.000	24.025	72.625	3.090	16.788	64.125
19	2.030	24.063	69.125	3.550	16.525	65.750
20	3.560	23.400	70.875	2.250	16.113	70.375
21	4.320	24.125	69.875	3.500	16.088	66.500
22	2.580	25.100	70.125	1.850	16.950	66.875
23	4.510	25.500	74.000	2.460	18.150	70.500
24	2.730	26.538	71.625	2.960	18.675	66.875
25	4.810	26.000	70.125	2.420	16.700	68.750
26	3.690	25.125	69.875	2.060	16.838	68.250
27	3.690	26.038	68.750	3.630	16.738	64.000
28	5.080	24.238	66.125	1.160	16.250	70.000
29	3.080	21.388	64.250	2.580	16.788	69.375
30	4.130	21.213	66.875	3.550	16.838	69.125
31				1.820	17.388	65.000
Correlation*	-	0.269	-0.358		0.398	-0.119

หมายเหตุ* หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างการระเหยน้ำกับตัวแปรช่วย

ภาคผนวก ง

การเขียนโปรแกรมจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่า \bar{y}_M
กับตัวประมาณค่า $\bar{y}_{m(SD)}$ ของ Gupta and Shabbir (2007)

```

=====
                          Generate Population:  $X_t \sim \text{Normal}$  and  $Z_t \sim \text{Normal}$ 
=====

clear all
tic
clc
N = 0.1;
k = 0;
while mod(N,k)~=0
N = input('Population Size : ');
k = input('Jumping rate : ');
end
t = 0;
for t = 1:N/k
PVAL1 = .99;
PVAL2 = .99;
RHO1 = 0;
RHO2 = 0;
RHO3 = 0;
while PVAL1 >= 0.05
i = 1;
X1 = zeros(k,1);
X2 = zeros(k,1);
Y = zeros(k,1);
    while i < k+1
        X = 0;
        while X<19 || X>31
            Xt = normrnd(0.03141,0.9168);
            X = 28.8112+0.9622*sinh((Xt-1.2251)/0.9295);
        end
        Z = 0;
        while Z<54 || Z>86
            Zt = normrnd(6.9828,1.6851);
            A = exp((Zt+0.2517)/0.7113);
            Z = (86.2413*A+54.7732)/(1+A);
        end
        E = normrnd(0,1);
        EP = 3.877+0.199*X-0.067*Z+E;
        if EP > 3

```

```

        Y(i,1) = EP;
        X1(i,1) = X;
        X2(i,1) = Z;
        i = i+1;
    else
        i = i;
    end
end
end
[RHO1,PVAL1] = corr(Y,X1,'tail','right')
[RHO2,PVAL2] = corr(Y,X2,'tail','left')
if (RHO1 > 0.45) && (RHO2 < -0.7 )
    [RHO3,PVAL3] = corr(X1,X2,'tail','right')
    if (RHO3 >0) && (RHO3 < min(RHO1,abs(RHO2)))
        PVAL1 = 0;
    else
        PVAL1 = 0.99;
    end
else
    PVAL1 = 0.99;
end
end
if RHO3 < 0
    PVAL1 = 0.99
end
end
end
PopKP =[Y X1 X2];
if t == 1
    Pop = PopKP
else
    Pop = [Pop; PopKP]
    disp(['          sample size (n) = ',length(Pop)])
    if mod(length(Pop),50) == 0
        save('PopTH_tran23','Pop','-TABS','-double')
    end
end
end
end
save('PopTH_tran23','Pop','-TABS','-double');
corr(Pop)
[RHO4,PVAL4] = corr(Pop(:,1),Pop(:,3),'tail','left');
disp(['PHO YX2 = ',num2str(PVAL4,'%10.4f')])

```

```

[RHO5,PVAL5] = corr(Pop(:,1),Pop(:,2),'tail','right');
disp(['PHO YX1= ',num2str(PVAL5,'%10.4F')])
time = toc;
datestr(time/86400,'HH:MM:SS.FFF')
clear E EP N PVAL1 PVAL2 PopKP RHO1 RHO2 X X1 X2 Y Z k t i Pop

```

```
=====
```

```

                        Create Good Data and Bad Data files

```

```
=====
```

```

load PopTH_tran;
[n,p] = size(Pop);
mu = mean(Pop)
sigma = std(Pop)
upper_out = quantile(Pop,0.75)+1.5*iqr(Pop)
%=====
%      check outiler
up_out = repmat(upper_out,n,1);
outliers = Pop > up_out;
nout = sum(outliers)
%=====
Outlier = zeros(nout(1),3)
PopGood = zeros(10000-nout(1),3)
i = 1;
j = 1;
k = 1;
for i = 1:n
    if outliers(i,1) == 1
        i
        Outlier(j,:) = Pop(i,:);
        j = j+1
    else
        PopGood(k,:) = Pop(i,:);
        k = k+1
    end
end

```

```

end

save('PopGood','PopGood','-TABS','-double');
save('Outlier','Outlier','-TABS','-double');
clear all

=====
                        Create Sample Data for 12 situations
                        For example: n = 20 & %Outlier = 5%
=====

clear all

clc
tic

load('PopGood.mat')
load('Outlier.mat')

n = input('sample Size (20, 40 ,100): ');
PercentOut = input('Percent of outliers (.05, .10, .15, .20): ');
LocOut = 1;
k = 1;
samp02005 = cell(1000,3)
out = zeros(1,2);
while k < 1001
    while out(1,1)~= 0 || out(1,2)~= n*PercentOut
        sampGood = datasample(PopGood,n-n*PercentOut,'Replace',false);
        sampOut = datasample(Outlier,n*PercentOut,'Replace',false);
        sampyOut = [sampGood; sampOut];
        [out(1,1), out(1,2)] = count_outlier02(sampyOut);
    end
    samp02005{k,1} = sampyOut
    k = k+1
    out = zeros(1,2);
end
save('samp02005','samp02005','-TABS','-double');
clear all

toc

```

```

=====
                        Function for counting the number of outliers
=====
function [n_extreme n_out] = count_outall(sampyOut)
n_extreme = 0;
n_out = 0;
out_upp =quantile(sampyOut,0.75)+3*(quantile(sampyOut,0.75)-
quantile(sampyOut,0.25))
out_low =quantile(sampyOut,0.75)+1.5*(quantile(sampyOut,0.75)-
quantile(sampyOut,0.25))
for i = 1:length(sampyOut)
    if sampyOut(i,1)> out_upp(1)
        n_extreme = n_extreme + 1
    else
        if sampyOut(i,1) > out_low(1)
            n_out = n_out + 1
        end
    end
end
disp(['number of extremes: ',num2str(n_extreme)]);
disp(['number of outliers: ',num2str(n_out)]);
end
=====
                        Compute and Compare proposed Statistic vs. Gupta's statistic
                        For example: n = 20 & %Outlier = 5%
=====
clear all
clc
load PopTH_tran;
N = length(Pop(:,1));
Ybar = mean(Pop(:,1));
Xbar = mean(Pop(:,2));
Zbar = mean(Pop(:,3));

```

```

Cy = std(Pop(:,1))/Ybar;
Cx = std(Pop(:,2))/Xbar;
Cz = std(Pop(:,3))/Zbar;
Yskew = skewness(Pop(:,1));
Xskew = skewness(Pop(:,2));
Zskew = skewness(Pop(:,3));
clear Pop;
load samp02005;
n = length(samp02005{1,1}(:,1));
t = 1000;
lamda = (N-n)/(N*n);
i = 1;
while i < t+1
    S = samp02005{i,1};
    x = samp02005{i,1}(:,2:3);
    y = samp02005{i,1}(:,1);
    %=====
    %compute least trimmed square
    result=ltsregres(x,y,'plots',0,'intercept',0);
    samp02005{i,2} = [result.slope; result.rsquared]
    %=====
    %compute all statistics form sample
    meanS = mean(S);
    cov = cov(S);
    cv = std(S)./mean(S);
    skew = skewness(S);
    alpha1 = Xbar/(Xbar+cv(2));
    alpha2 = Zbar/(Zbar+cv(3));
    r1 = Ybar/(Xbar+skew(2));
    r2 = Ybar/(Zbar+skew(3));
    cor = corr(S);
    %=====
    %compute 2 y-bar and MSE of them

```

```

y_barSD =
meanS(1)*((Xbar+Cx)/(meanS(2)+Cx))*((Zbar+Cz)/(meanS(3)+Cz))

MSE1 = (Ybar^2)*lamda*(cv(1)^2+alpha1^2*cv(2)^2+alpha2^2*cv(3)^2-...
        2*alpha1*cor(1,2)*cv(1)*cv(2)+2*alpha2*cor(1,3)*cv(1)*cv(3)...
        -2*alpha1*alpha2*cor(2,3)*cv(2)*cv(3))

%=====

y_barM = (meanS(1)+result.slope(1,1)*(Xbar-
meanS(2))+result.slope(2,1)*(Zbar-meanS(3)))*...
        ((Xbar+Xskew)/(meanS(2)+Xskew))*((meanS(3)+Zskew)/(Zbar+Zskew))

MSE2 = lamda*(cov(1,1)+cov(2,2)*(r1+result.slope(1,1))^2+...
        cov(3,3)*(r2-result.slope(2,1))^2-
2*(r1+result.slope(1,1))*cov(1,2)...
        +2*(r2-result.slope(2,1))*cov(1,3)...
        +2*(r1+result.slope(1,1))*(r2-result.slope(2,1))*cov(2,3))

%=====

Result{i,1} = S;
Result{i,2} = y_barSD;
Result{i,3} = MSE1;
Result{i,4} = (abs(y_barSD-Ybar)/Ybar)*100;
Result{i,5} = y_barM;
Result{i,6} = MSE2;
Result{i,7} = (abs(y_barM-Ybar)/Ybar)*100;
Result{i,8} = (y_barSD-Ybar)^2;
Result{i,9} = (y_barM-Ybar)^2;
i = i+1

clear S x y result meanS cov cv skew alpha1 alpha2 r1 r2 cor
end

clear sample

Result{i,1} = Ybar;
Result{i,2}= mean([Result{:,2}]);
Result{i,3}= mean([Result{:,3}]);
Result{i,4}= mean([Result{:,4}]);
Result{i,5}= mean([Result{:,5}]);
Result{i,6}= mean([Result{:,6}]);

```



```

Result{i,7}= mean([Result{:,7}]);
Result{i,8} = sqrt(sum([Result{:,8}])/1000);
Result{i,9} = sqrt(sum([Result{:,9}])/1000);
%Result;
save('ResultTH_tran02005skew','Result');
disp([' ybarSD = ',num2str(Result{i,2}, '%.4f')]);
disp(['mean of %RE ybarSD = ',num2str(Result{i,4}, '%.4f')]);
disp(['RMSE ybarSD = ',num2str(Result{i,8}, '%12.4f')]);
disp(['    ']);
disp([' ybarM = ',num2str(Result{i,5}, '%.4f')]);
disp(['mean of %RE ybarM = ',num2str(Result{i,7}, '%.4f')]);
disp(['RMSE ybarM = ',num2str(Result{i,9}, '%12.4f')]);
disp(['    ']);
if abs(Result{i,5}-Ybar)< abs(Result{i,2}-Ybar)
    disp(['ybarM is close to Ybar'])
else
    disp(['ybarSD is close to Ybar'])
end
disp(['    '])
disp(['-----'])
%disp(['    '])
if Result{i,7} < Result{i,4}
    disp(['    %RE ybarM < %RE ybarSD =====> Good'])
else
    disp(['    %RE ybarM > %RE ybarSD =====> Fail'])
end
disp(['-----'])
%disp(['    '])
if Result{i,9} < Result{i,8}
    disp(['    RMSE ybarM < RMSE ybarSD =====> Good'])
else
    disp(['    RMSE ybarM > RMSE ybarSD =====> Fail'])
end

```

ภาคผนวก จ

การเขียนโปรแกรมชุดคำสั่งเพื่อประมาณค่าการระเหยน้ำโดยเฉลี่ยรายเดือนในจังหวัดสกลนคร

```

=====
                        Estimated Monthly Average Evaporation
                        *for example: May
=====

clear all
clc
load A5MayP
N = length(A5MayP(:,1));
Ybar = mean(A5MayP(:,1));
Xbar = mean(A5MayP(:,2));
Zbar = mean(A5MayP(:,3));
Cy = std(A5MayP(:,1))/Ybar;
Cx = std(A5MayP(:,2))/Xbar;
Cz = std(A5MayP(:,3))/Zbar;
Yskew = skewness(A5MayP(:,1));
Xskew = skewness(A5MayP(:,2));
Zskew = skewness(A5MayP(:,3));
Cor = corr(A5MayP);
Cov = cov(A5MayP);
%clear A5MayP;
May = {};
May{1,1} = Ybar;

load A5May;
S = A5May;
May{1,2} = A5May;
n = length(A5May(:,1));
lamda = (N-n)/(N*n);

x = A5MayP(:,2:3);
y = A5MayP(:,1);
%compute least trimmed square
result=ltsregres(x,y,'plots',0,'intercept',0);
May{1,3} = [result.slope; result.rsquared]
%compute all statistics form sample
meanS = mean(S);
cov = cov(S);
cv = std(S)./mean(S);

```

```

skew = skewness(S);
alpha1 = Xbar/(Xbar+cv(2));
alpha2 = Zbar/(Zbar+cv(3));
r1 = Ybar/(Xbar+Xskew);
r2 = Ybar/(Zbar+Zskew);
cor = corr(S);

%compute 2 y-bar and MSE of them
y_barSD =
meanS(1)*((Xbar+Cx)/(meanS(2)+Cx))*((meanS(3)+Cz)/(Zbar+Cz))
MSE1 = (Ybar^2)*lamda*(Cy^2+alpha1^2*Cx^2+alpha2^2*Cz^2-...
        2*alpha1*Cor(1,2)*Cy*Cx-2*alpha2*Cor(1,3)*Cy*Cz...
        +2*alpha1*alpha2*Cor(2,3)*Cx*Cz)
%=====
y_barM = (meanS(1)+result.slope(1,1)*(Xbar-
meanS(2))+result.slope(2,1)*(Zbar-meanS(3)))*...
        ((Xbar+Xskew)/(meanS(2)+Xskew))*((Zbar+Zskew)/(meanS(3)+Zskew))
MSE2 = lamda*(Cov(1,1)+Cov(2,2)*(r1+result.slope(1,1))^2+...
        Cov(3,3)*(-r2+result.slope(2,1))^2 ...
        -2*(r1+result.slope(1,1))*Cov(1,2) ...
        -2*(-r2+result.slope(2,1))*Cov(1,3) ...
        +2*(r1+result.slope(1,1))*(-r2+result.slope(2,1))*Cov(2,3))
%=====
May{2,1} = y_barSD;
May{2,2} = y_barM;
May{2,3} = mean(x(:,1))*result.slope(1,1)+ ...
        mean(x(:,2))*result.slope(2,1);
May{3,1} = MSE1;
May{3,2} = MSE2;

clear S x y meanS cov cv skew alpha1 alpha2 r1 r2 cor
%show output
May
save('ResultTH_RP_May5','May')
disp(['-----'])
disp(['Ybar = ',num2str(Ybar,'%6.4f')])
disp(['YbarSD = ',num2str(May{2,1},'%6.4f')])
disp(['YbarM = ',num2str(May{2,2},'%6.4f')])

```

```
disp(['MSE YbarSD = ', num2str(May{3,1}, '%6.4f')])
disp(['MSE YbarM = ', num2str(May{3,2}, '%6.4f')])
RE1 = (abs(Ybar-May{2,1})/Ybar)*100;
RE2 = (abs(Ybar-May{2,2})/Ybar)*100;
%C = abs(Ybar-May{2,3});
disp([' %RE of ybarSD = ', num2str(RE1, '%6.4f')])
disp([' %RE of ybarM = ', num2str(RE2, '%6.4f')])
```