บทที่ 3

วิธีดำเนินการวิจัย

งานวิจัยนี้มุ่งที่จะสร้างรูปแบบในการประมาณค่าการแผ่รังสีของดวงอาทิตย์ เพื่อนำไปสร้างแผนที่การแผ่รังสีดวงอาทิตย์ของประเทศไทย ซึ่งเป็นประโยชน์สำหรับผู้ที่ทำงานในด้านการนำพลังงานแสงอาทิตย์มาใช้ให้เป็นประโยชน์ โดยมีลำดับขั้นตอนการดำเนินการวิจัยดังนี้

3.1 การเก็บรวบรวมข้อมูลรังสีดวงอาทิตย์และความยาวนานแสงแดด

ข้อมูลของทุก ๆ สถานี อุตุนิยมวิทยาที่นำมาวิเคราะห์เป็นข้อมูลทุติยภูมิ คัดลอกจากฝ่ายข้อมูลกองภูมิอากาศ กรมอุตุนิยมวิทยา และเอกสารสถิติอุทกวิทยา สำนักงานพลังงานแห่งชาติ โดยมีลำดับขั้นการเก็บข้อมูลคือ

จัดทำตารางบันทึก วัน เดือน ปี สถานที่ตั้งของสถานีอุตุนิยมวิทยา เส้นรุ้ง (Latitude) เส้นแวง (Longtitude)

บันทึกค่าการแผ่รังสีของดวงอาทิตย์รวมเป็นรายวัน (Cal cm^{-2} day^{-1}) และค่าความยาวนานแสงแดด (ชั่วโมง) ใน 1 วัน

3.2 การวิเคราะห์ข้อมูล

3.2.1 การแปลง (Transformation) ข้อมูลดิบให้เป็นข้อมูลไดเมนชั่นเลสจะต้องคำนวนหาค่ารังสีนอกบรรยากาศของโลก (Extraterrestrial Radiation; H_m) โดยหาได้จากสูตร

$$H_m = \frac{24}{\pi} \cdot I_{sc} \{1 + 0.033 \cos (\frac{360n}{365})\}\{\cos \phi \cos\delta \sin W_s + \frac{2\pi W_s}{360} \sin\phi\sin\delta\} \quad (3\text{-}1)$$

$$I_{sc} = \text{Solar constant} = \frac{4.871}{4.1868} \times 100 \text{ Cal } cm^{-2} \text{ } hr^{-1} \quad (3\text{-}2)$$

ϕ = เส้นรุ้งของตำแหน่งที่ต้องการทราบค่ารังสี

n = วันที่ของปี $n = 1,2,3,.....,365$

δ = Solar declination angle $= 23.45 \sin (360.\frac{(284+n)}{365})$ (3-3)

W_s = Sunrise hours angle $= \cos^{-1} (-\tan\phi\tan\delta)$ (3-4)

คำนวนค่าความยาวของวัน (S_m)

$$S_m = \frac{2}{15} \cos^{-1} (-\tan\phi\tan\delta) \quad (3\text{-}5)$$

ค่า H_m และ S_m ของทั้ง 365 วัน จะแตกต่างกันและต่างกันในสถานีที่อยู่ ณ ตำแหน่งเส้นรุ้งที่ต่างกันด้วย

การหาค่า H_m และ S_m เพื่อที่จะให้ได้ค่าไดเมนชั่นเลสของการแผ่รังสี (H/H_m) และค่าไดเมนชั่นเลสความยาวนานแสงแดด (S/S_m) คำนวนด้วยเครื่องคอมพิวเตอร์ โปรแกรมการคำนวน H_m และ S_m ดังแสดงในภาคผนวก ก. และค่าของ H_m และ S_m ที่หาได้ของทั้ง 33 สถานีอุตุนิยมวิทยา ดังรายละเอียดที่แสดงไว้ในตารางที่ ก.1-ตารางที่ ก.66 ตามลำดับ

H คือค่าการแผ่รังสีของดวงอาทิตย์รวมที่วัดได้ (แคลอรี่/ตารางเซนติเมตร-วัน)

S คือความยาวนานแสงแดด (ชั่วโมง) ใน 1 วัน

ดังนั้นข้อมูลที่ใช้วิเคราะห์ในขั้นต่อ ๆ ไป จะเป็นข้อมูลดิบที่แปลงเป็นข้อมูลแบบไดเมนชั่นเลสแล้ว

3.2.2 การวิเคราะห์หาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการแผ่รังสีของดวงอาทิตย์และความยาวนานแสงแดด

การหาค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปร 2 ตัว X และ Y คำนวนจากสูตร

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2 \sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \quad ; -1 \leq r \leq 1$$

ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ (r) มีค่าเป็นบวก แสดงว่าเมื่อตัวแปรหนึ่งมีค่าเพิ่ม ตัวแปรอีกตัวหนึ่งจะเพิ่มตาม แต่ถ้าค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเป็นลบ แสดงว่าเมื่อตัวแปรหนึ่งมีค่าเพิ่ม ตัวแปรอีกตัวหนึ่งจะมีค่าลดลง คือเป็นไปในทางตรงกันข้าม ถ้า r มีค่าเข้าใกล้ +1 หรือ -1 แสดงว่า

X และ Y มีความสัมพันธ์กันมาก ถ้า r มีค่าเข้าใกล้ 0 แสดงว่าตัวแปรทั้ง 2 มีความสัมพันธ์กันน้อย ถ้า r มีค่าเป็น 0 แสดงว่าตัวแปรทั้ง 2 ไม่มีความสัมพันธ์กัน ค่าสัมประสิทธิ์สหสัมพันธ์เป็นค่าที่บอกให้ทราบว่าตัวแปรทั้ง 2 มีความสัมพันธ์กันมากน้อยเพียงใดเพียงอย่างเดียว แต่ค่าที่จะบอกถึงความสัมพันธ์ระหว่างตัวแปรทั้ง 2 ว่า เมื่อ X เปลี่ยนไป 1 หน่วย ค่า Y จะเปลี่ยนแปลงไปเท่าไร คือค่าสัมประสิทธิ์การถดถอย ซึ่งจะคำนวนได้จากการวิเคราะห์ถดถอย

3.2.3 การวิเคราะห์หาตัวแบบการถดถอย ระหว่างการแผ่รังสีของดวงอาทิตย์และความยาวนานแสงแดด ใช้วิธีการวิเคราะห์การถดถอยเส้นโค้ง (Curvilinear Regression Analysis) โดยใช้ตัวแบบเป็น Polynomial Regression ซึ่งมีรูปแบบความสัมพันธ์ดังนี้

$$Y = \beta_0+\beta_1X+\beta_2X^2+\beta_3X^3+ \ .....+\beta_kX^k+\varepsilon$$

เมื่อ Y เป็นตัวแปรตาม

X เป็นตัวแปรอิสระ

ε เป็นค่าความคลาดเคลื่อนเชิงสุ่ม

$\beta_0,\beta_1,\beta_2, \ .....,\beta_k$ เป็นตัวคงที่ ๆ ไม่ทราบค่าซึ่งเป็นสัมประสิทธิ์ของ $X^0,X^1,X^2,..,X^k$ ตามลำดับ

สมการคาดคะเนคือ

$$Y_c = b_o+b_1X +b_2X^2+b_3X^3+ \ .....+ b_kX^k$$

เมื่อ Y_c คือค่าประมาณของ Y

$b_0,b_1,b_2,.....,b_k$ เป็นค่าประมาณของ $\beta_0,\beta_1,.....,\beta_k$

สำหรับงานวิจัยเรื่องนี้เมื่อนำข้อมูลรังสีดวงอาทิตย์และความยาวนานแสงแดดมาเขียนแผนภาพการกระจาย (scatter diagram) แล้วจากการศึกษาเบื้องต้นพบว่า ความสัมพันธ์ระหว่างรังสีดวงอาทิตย์และความยาวนานแสงแดดอยู่ในรูปของโพลีโนเมียล ดีกรีที่ห้า ซึ่งมีสมการคาดคะเนเป็น

$$Y_c= b_o+b_1X+b_2X^2+b_3X^3+b_4X^4+b_5X^5$$

การประมาณค่าสัมประสิทธิ์การถดถอยจะใช้วิธีกำลังสองน้อยที่สุด (Ordinary Least Square Method) ซึ่งคำนวนด้วยเครื่องคอมพิวเตอร์ ตามคำสั่งในโปรแกรมสำเร็จรูปสำหรับการวิจัยทางสังคมศาสตร์ (SPSS) โดยวิธี Stepwise Regression เมื่อ Y เป็นตัวแปรตาม และ X, X^2, X^3, X^4, X^5 เป็นตัวแปรอิสระ ส่วนการทดสอบสมมุติฐานโดยใช้ F-test กระทำเป็น 2 ขั้นตอนคือ

ขั้นที่ 1 ทดสอบนัยสำคัญเชิงสถิติของตัวแบบการถดถอย คือทดสอบว่าตัวแปรอิสระ (X^i) ทุกตัว มีความสัมพันธ์กับตัวแปรตาม y หรือไม่ โดยมีสมมุติฐาน H_o เป็น

$$H_o: \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

ใช้ระดับนัยสำคัญของการทดสอบเป็น 0.05 ถ้าผลการทดสอบสมมุติฐานคือ ยอมรับสมมุติฐาน H_o จะยุติการทดสอบและสรุปว่าสมการการถดถอยเป็นสมการที่ไม่เหมาะสมในการนำไปพยากรณ์ค่า Y เนื่องจาก Y และ X, X^2, X^3, X^4, X^5 ไม่มีความสัมพันธ์กัน ส่วนกรณีที่ปฏิเสธสมมุติฐาน H_o หมายความว่ามีสัมประสิทธิ์การถดถอย (β_i) อย่างน้อยหนึ่งค่าที่ไม่เท่ากับศูนย์อย่างมีนัยสำคัญ ทำการทดสอบขั้นที่ 2 ต่อไป

ขั้นที่ 2 ทดสอบนัยสำคัญเชิงสถิติของสัมประสิทธิ์การถดถอย คือทดสอบว่าตัวแปรอิสระ (X^i) แต่ละตัว มีความสัมพันธ์กับตัวแปรตาม y หรือไม่ โดยมีสมมุติฐาน H_o เป็น

$$H_o : \beta_i = 0 \; ; \quad i = 1,2,3,4,5$$

ใช้ระดับนัยสำคัญของการทดสอบเป็น 0.05 ถ้าผลการทดสอบยอมรับ สมมุติฐาน H_o อย่างมีนัยสำคัญจะตัดตัวแปรอิสระ X^i ออกจากตัวแบบการถดถอยและถ้าปฏิเสธสมมุติฐาน H_o หมายความว่าตัวแปรอิสระ X^i มีส่วนช่วยในการพยากรณ์ค่า Y จะยอมรับตัวแปรอิสระ X^i ไว้ในสมการถดถอย

3.2.4 การวิเคราะห์หาค่าสัมประสิทธิ์การถดถอยระหว่างค่า $b_0, b_1, \ldots, b_5$ ที่ได้จากข้อ 3.2.3 และเส้นรุ้ง เพื่อใช้ในการ interpolate และ extrapolate ค่า $b_0, b_1, \ldots, b_5$ ที่เส้นรุ้งต่าง ๆ โดยใช้รูปแบบทางคณิตศาสตร์ คือ รูปแบบเส้นตรงและรูปแบบเส้นโค้งพาราโบล่า

3.2.4.1 การหาค่าสัมประสิทธิ์การถดถอย โดยใช้รูปแบบเส้นตรง หาได้จากสมการ

$$b_{ij} = a_{1i} + a_{2i} N_j \; ; \quad i = 0,1,2,\ldots,5; \; j = 1,2,\ldots,7$$

เมื่อ b_{ij} เป็นค่าของสัมประสิทธิ์การถดถอยที่ i ของสถานีที่ j

a_{1i}, a_{2i} เป็นค่าสัมประสิทธิ์ที่ต้องการทราบของ b_i

N_j เป็นค่าของเส้นรุ้ง ณ. สถานีที่ j

เช่นสมการในการหาค่าสัมประสิทธิ์การถดถอยของการ interpolate และextrapolate ค่า b_o คือ

$$b_{01} = a_{10} + a_{20}N_1$$

$$b_{02} = a_{10} + a_{20}N_2$$

$$b_{03} = a_{10} + a_{20}N_3$$

$$b_{04} = a_{10} + a_{20}N_4$$

$$b_{05} = a_{10} + a_{20}N_5$$

$$b_{06} = a_{10} + a_{20}N_6$$

$$b_{07} = a_{10} + a_{20}N_7$$

เมื่อค่า $b_{01}, b_{02}, \ldots, b_{07}$ ได้จากข้อ 3.2.3 และค่า $N_1, \ldots, N_7$ เป็นค่าของเส้นรุ้งของสถานีที่ 1 ถึงสถานีที่ 7 ตามลำดับ การคำนวนค่า a_{10}, a_{20} จะใช้วิธีกำลังสองน้อยที่สุดในทำนองเดียวกัน การคำนวนหาค่าสัมประสิทธิ์การถดถอยของ $b_1, b_2, \ldots, b_5$ และเส้นรุ้ง ก็หาได้เช่นเดียวกัน สมการในการ interpolate และ extrapolate ค่า b_i ซึ่งแปรตามเส้นรุ้งเป็นดังนี้

$$\hat{b}_0 = \hat{a}_{10} + \hat{a}_{20}N$$

$$\hat{b}_1 = \hat{a}_{11} + \hat{a}_{21}N$$

$$\hat{b}_2 = \hat{a}_{12} + \hat{a}_{22}N$$

$$\hat{b}_3 = \hat{a}_{13} + \hat{a}_{23}N$$

$$\hat{b}_4 = \hat{a}_{14} + \hat{a}_{24}N$$

$$\hat{b}_5 = \hat{a}_{15} + \hat{a}_{25}N$$

007830 I 1556891X

3.2.4.2 การหาค่าสัมประสิทธิ์การถดถอย โดยใช้รูปแบบเส้นโค้ง พาราโบล่า หาได้จากสมการ

$$b_{ij} = a_{1i} + a_{2i}N_j + a_{3i}N_j^2 \quad ; \; i = 0,1,2,....5; \; j = 1,2,...,7$$

เมื่อ b_{ij} เป็นค่าของสัมประสิทธิ์การถดถอยที่ i ของสถานีที่ j

a_{1i}, a_{2i}, a_{3i} เป็นค่าสัมประสิทธิ์ที่ต้องการทราบของ b_i

N_j เป็นค่าของเส้นรุ้ง ณ. สถานีที่ j

การคำนวนหาค่า a_{1i}, a_{2i}, a_{3i} ใช้วิธีกำลังสองน้อยที่สุดเช่นเดียวกับข้อ 3.2.4.1 และจะได้สมการในการ interpolate และextrapolate ค่า b_i ซึ่งแปรตามเส้นรุ้งดังนี้

$$\hat{b}_o = \hat{a}_{10} + \hat{a}_{20}N + \hat{a}_{30}N^2$$

$$\hat{b}_1 = \hat{a}_{11} + \hat{a}_{21}N + \hat{a}_{31}N^2$$

$$\hat{b}_2 = \hat{a}_{12} + \hat{a}_{22}N + \hat{a}_{32}N^2$$

$$\hat{b}_3 = \hat{a}_{13} + \hat{a}_{23}N + \hat{a}_{33}N^2$$

$$\hat{b}_4 = \hat{a}_{14} + \hat{a}_{24}N + \hat{a}_{34}N^2$$

$$\hat{b}_5 \quad \hat{a}_{15} + \hat{a}_{25}N + \hat{a}_{35}N^2$$

3.2.5 สร้างรูปแบบ (Model) เพื่อใช้ในการประมาณค่าการแผ่รังสีดวงอาทิตย์ทั่วประเทศ

<u>รูปแบบที่ 1</u> $\hat{Y}_N = \hat{b}_o + \hat{b}_1X + \hat{b}_2X^2 + \hat{b}_3X^3 + \hat{b}_4X^4 + \hat{b}_5X^5$

<u>รูปแบบที่ 2</u> $\hat{Y}_N2 = \hat{\hat{b}}_o + \hat{\hat{b}}_1X + \hat{\hat{b}}_2X^2 + \hat{\hat{b}}_3X^3 + \hat{\hat{b}}_4X^4 + \hat{\hat{b}}_5X^5$

เมื่อ $\hat{Y}_N$, $\hat{Y}_N2$ เป็นค่าประมาณของค่าไดเมนชั่นเลสของการแผ่รังสี

$\hat{b}_o, \hat{b}_1, \ldots, \hat{b}_5$ เป็นค่าประมาณของ $b_0, b_1, \ldots, b_5$ ที่หาได้จาก ข้อ 3.2.4.1

$\hat{\hat{b}}_o, \hat{\hat{b}}_1, \ldots, \hat{\hat{b}}_5$ เป็นค่าประมาณของ $b_0, b_1, \ldots, b_5$ ที่หาได้จาก ข้อ 3.2.4.2

X เป็นค่าไดเมนชั่นเลสความยาวนานแล่งแดด

ให้ $\hat{Y}_N$ เป็นค่าประมาณของค่าไดเมนชั่นเลสของการแผ่รังสีเมื่อใช้ค่าสัมประสิทธิ์การถดถอยที่ได้จาก ข้อ 3.2.4.1

และ $\hat{Y}_N 2$ เป็นค่าประมาณของค่าไดเมนชั่นเลสของการแผ่รังสีเมื่อใช้ค่าสัมประสิทธิ์การถดถอยที่ได้จากข้อ 3.2.4.2

3.2.6 เปรียบเทียบวิธีการประมาณค่าการแผ่รังสีดวงอาทิตย์โดยการทดสอบสมมุติฐานทางสถิติ โดยใช้ t-test กล่าวคือทดสอบสมมุติฐานว่าการประมาณค่าการแผ่รังสีดวงอาทิตย์จาก 2 รูปแบบ ให้ผลไม่แตกต่างกัน โดยใช้ระดับนัยสำคัญของการทดสอบเป็น 0.05 ถ้าผลการทดสอบคือ ยอมรับสมมุติฐานแสดงว่า ไม่มีความแตกต่างระหว่างการประมาณค่าทั้ง 2 วิธี การเลือกใช้รูปแบบใดก็ขึ้นอยู่กับผู้ใช้จะตัดสินใจซึ่งก็ควรจะเลือกวิธีที่ง่าย สะดวกและประหยัดเวลาที่สุดในการคำนวน แต่ถ้าปฏิเสธสมมุติฐานแสดงว่า มีความแตกต่างกันระหว่างการประมาณทั้ง 2 วิธี การเลือกใช้รูปแบบใดควรจะเปรียบเทียบค่าผลรวมของความคลาดเคลื่อนกำลังสอง (Sum squares of residual) ซึ่งคำนวนจากสูตร

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

เมื่อ Y_i คือค่าสังเกตหรือค่าที่รวบรวมมาได้

$\hat{Y}_i$ คือค่าพยากรณ์ (predicted)

วิธีที่ควรเลือกคือวิธีที่ให้ค่าผลรวมของความคลาดเคลื่อนกำลังสอง (SSE) ต่ำที่สุด

3.2.7 การวิเคราะห์จำแนกประเภท (Discriminant Analysis) การวิเคราะห์จำแนกประเภทมีวัตถุประสงค์ที่จะคัดเลือกตัวแปรกลุ่มหนึ่งหรือชุดหนึ่งที่นักวิจัยคิดว่าตัวแปรเหล่านี้มีความสัมพันธ์กับสิ่งที่ต้องการศึกษา (ตัวแปรตาม) จนถึงขั้นที่จะสามารถแยกประชากรออกเป็นกลุ่มต่าง ๆ ได้ประชากรแต่ละกลุ่มจะมีตัวแปรที่สัมพันธ์กับกลุ่มซึ่งเป็นตัวแปรที่ใช้แบ่งแยกกลุ่มออกจากกลุ่มอื่น ๆ เช่น สมมุติว่ามีตัวแปร A,B,C, และมีกลุ่มประชากรที่ 1,2 เทคนิคการวิเคราะห์จำแนกประเภทนี้ จะชี้ให้เห็นว่าการเป็นสมาชิกของกลุ่ม 1 หรือ 2 ขึ้นกับน้ำหนักของตัวแปร A,B,C, อย่างไร สมการที่จะให้น้ำหนักของตัวแปร A,B,C ที่จะมาใช้ในการคาดประมาณ

การเป็นสมาชิกของกลุ่มเรียกว่า สมการจำแนกประเภท สมการที่ได้มีประโยชน์ 2 ประการคือ การวิเคราะห์ข้อมูลทางสถิติ และการจำแนกประเภทข้อมูล ในด้านการวิเคราะห์จะมีสถิติต่าง ๆ เกี่ยวกับตัวแปรแต่ละตัวในสมการที่ใช้ในการทดสอบความสำคัญของตัวแปร ซึ่งช่วยในการตีความหมายข้อมูล ค่าสัมประสิทธิ์ของตัวแปรต่าง ๆ ที่ได้ในสมการบ่งชี้ถึงน้ำหนักหรือความสัมพันธ์ของตัวแปรนั้นในสมการนั้นต่อการเป็นสมาชิกของกลุ่มนั้น ในลักษณะที่คล้ายคลึงกับค่าสัมประสิทธิ์ถดถอยของสมการถดถอยพหุ ในด้านการจำแนกประเภทเมื่อได้สมการมาแล้ว ก็สามารถที่จะระบุว่ากรณีใดจะเป็นสมาชิกของกลุ่มใดได้โดยอาศัยตัวแปรและค่าสัมประสิทธิ์ต่าง ๆ ที่ปรากฏอยู่ในสมการจำแนกเป็นเครื่องมือช่วยระบุจำแนกประเภท

ความสามารถของสมการในการจำแนกประเภท อาจทดสอบได้จากการจำแนกประเภทของกรณีที่ทราบว่าเป็นสมาชิกของกลุ่มใด และใช้สมการประมาณเปรียบเทียบว่าตรงกับที่เป็นจริงหรือไม่

สำหรับงานวิจัยนี้จะใช้เทคนิคการวิเคราะห์การจำแนกประเภทเพื่อศึกษาว่า ถ้าใช้ตัวแปรอิสระความยาวนานแสงแดด (X) โดยแบ่งกลุ่มออกเป็น 3 กลุ่มตามเส้นรุ้ง จะได้สมการจำแนกประเภทดีมากน้อยเพียงใด ซึ่งดูได้จากการทดสอบโดยนำสมการดังกล่าวนี้ ไปคาดคะเนความเป็นสมาชิกของกลุ่มว่าถูกต้องร้อยละเท่าไร จากข้อมูลที่เก็บรวบรวมได้ทั้งหมด การวิเคราะห์จะคำนวนด้วยเครื่องคอมพิวเตอร์ โดยคำสั่งโปรแกรมสำเร็จรูปการวิจัยทางสังคมศาสตร์ (SPSS) โดยใช้วิธีของ Mahalanobis distance ซึ่งเป็นวิธีการที่ใช้หาระยะทางระหว่างสมาชิกใด ๆ ที่ต้องการทราบว่าควรจะอยู่ในกลุ่มใด และค่าเฉลี่ยของสมาชิกในแต่ละกลุ่ม ถ้าระยะทางระหว่างสมาชิกนั้นกับค่าเฉลี่ยของสมาชิกในกลุ่มใดมีค่าน้อยที่สุด แสดงว่าสมาชิกนั้นควรอยู่ในกลุ่มดังกล่าว

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย