บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ระบบสารสนเทศภูมิศาสตร์ (GIS)

ระบบสารสนเทศภูมิศาสตร์ (Geographic Information System) หมายถึง เครื่องมือที่ใช้ระบบคอมพิวเตอร์ หรือใช้ในการนำเข้า จัดเก็บ จัดเตรียม ดัดแปลง แก้ไข จัดการ และวิเคราะห์ พร้อมทั้งแสดงผลข้อมูลเชิงพื้นที่ ตามวัตถุประสงค์ต่าง ๆ ที่ได้กำหนดไว้ ดังนั้น GIS จึงเป็นเครื่องมือที่มีประโยชน์เพื่อใช้ในการจัดการ และบริหารการใช้ทรัพยากรธรรมชาติและสิ่งแวดล้อม และสามารถติดตามการเปลี่ยนแปลงข้อมูลด้านพื้นที่ ให้เป็นไปอย่างมีประสิทธิภาพเนื่องจากแหล่งข้อมูลต่าง ๆ เช่น ข้อมูลปฐมภูมิ (primary data) หรือข้อมูลทุติยภูมิ (secondary data) เพื่อให้เป็นข่าวสารที่มีคุณค่าข้อมูลที่จะนำเข้าสู่ควรเป็นข้อมูลเฉพาะเรื่อง และเป็นข้อมูลที่สามารถนำมาใช้ตอบคำถามต่าง ๆ ได้ตรงตามวัตถุประสงค์ เป็นข้อมูลที่มีความถูกต้องและเชื่อถือได้และเป็นปัจจุบันมากที่สุด อนึ่ง ข้อมูลสารสนเทศสามารถแบ่งออกได้เป็น 2 ประเภท คือ 1. ข้อมูลที่มีลักษณะเชิงพื้นที่ (Raster Data) และ 2. ข้อมูลอธิบายพื้นที่ (non-spatial data or attribute data) ข้อมูลเชิงพื้นที่ เป็นข้อมูลที่แสดงตำแหน่งที่ตั้งทางภูมิศาสตร์ (geo-reference data) และข้อมูลแสดงเป็นตารางกริด (raster data) ข้อมูลที่มีทิศทาง ประกอบด้วยลักษณะ 3 อย่างคือ

ข้อมูลจุด (point) เช่น ที่ตั้งหมู่บ้าน โรงเรียน เป็นต้น

ข้อมูลเชิงเส้น (arc or line) เช่น ถนน แม่น้ำ ท่อประปา เป็นต้น

ข้อมูลพื้นที่หรือเส้นรอบรูป (polygon) เช่น พื้นที่ป่าไม้ ตัวเมือง เป็นต้น

ข้อมูลที่มีลักษณะเป็นกริด (raster data) จะเป็นลักษณะตารางสี่เหลี่ยมเล็ก ๆ เท่ากันและต่อเนื่องกัน ซึ่งสามารถอ้างอิงค่าพิกัดทางภูมิศาสตร์ ขนาดของตารางกริด ความละเอียดในการเก็บข้อมูล จะใหญ่หรือเล็กขึ้นอยู่กับการจัดจำนวนแถว จำนวนคอลัมน์ ตัวอย่างข้อมูลที่ใช้โดยตารางกริด เช่น ภาพถ่ายดาวเทียม Landsat หรือ ข้อมูลระดับค่าความสูง (digital elevation model : DEM) เป็นต้น

โปรแกรม ArcView GIS

Arc View เป็นโปรแกรม GIS โปรแกรมหนึ่ง ที่ได้รับการพัฒนามาจาก บริษัทEnvironmental Systems Research Institute Inc. (ESRI) เพื่อใช้งานในการนำเสนอข้อมูล และเรียกค้นข้อมูลจากโปรแกรม Arc/Info หรือโปรแกรมอื่น ที่สามารถใช้งานได้ง่ายและมีประสิทธิภาพเนื่องจาก

ทำงานบนระบบปฏิบัติการ Windows System ซึ่งเมนู ต่าง ๆ แสดงบนหน้าจอ และสามารถเปิดได้หลาย ๆ หน้าต่างในระหว่างการทำงาน

โปรแกรม Arc view โปรแกรมแรกคือ Arc View 1.0 สามรถใช้งานได้เฉพาะการนำเสนองานรูปแบบแผนที่เท่านั้น แต่โปรแกรมได้มีการพัฒนาเรื่อยมา จนถึง version 3.2 ที่มีประสิทธิภาพมากยิ่งขึ้นใกล้เคียงกับโปรแกรม PC Arc/Info กล่าวคือนอกจากจะใช้งานในการนำเสนอข้อมูลและเรียกค้นข้อมูลตามเงื่อนไขต่าง ๆ และใช้ในการผลิตแผนที่ได้เป็นอย่างดีแล้ว ยังสามารถสร้างและแก้ไขข้อมูล ทั้งที่เป็นข้อมูลพื้นที่ และตารางฐานข้อมูลได้ด้วย และยังสามารถรับข้อมูลที่จัดเก็บในรูปแบบต่าง ๆ เช่น Autocad (dwg), Image (tiff,.bmp, etc.) และยังสามารถใช้ในการวิเคราะห์ข้อมูลเชิงพื้นที่ (special analysis) ได้ด้วย โดยการเขียนชุดคำสั่งหรือใช้โปรแกรมประยุกต์ (ชุดคำสั่งสำเร็จรูป) ที่ได้จัดเขียนไว้โดยผู้เชี่ยวชาญ

โปรแกรม Arc View ประกอบไปด้วยหน้าต่างที่สำคัญ 5 หน้าต่าง คือ Project Window, View Window, Table Window, Chart Window, Layout Window ดังภาพ 1

Project Window คือแฟ้มข้อมูลที่ Arc View สร้างขึ้นเพื่อใช้ในการจัดระบบในการทำงานทั้งหมดใน Project หนึ่งซึ่งจะรวมองค์ประกอบทั้งหมดให้อยู่ในแฟ้มเดียวกัน แต่ Project file ที่มีนามสกุลเป็น .apr ซึ่งแฟ้มข้อมูลดังกล่าวจะไม่มีข้อมูลพื้นที่และตารางฐาน แต่จะใช้ในการเรียกค้นข้อมูลจากแหล่งต่าง ๆ ใน Project หนึ่ง ๆ จะประกอบด้วยหน้าต่างย่อย หรือ องค์ประกอบหลัก 4 หน้าต่าง คือ View, Table, Charts, Layouts แต่ Arc View จะทำงานครั้งละ 1 Project เท่านั้น หากต้องการดูลายละเอียดใน Project อื่น ต้องปิด Project ที่กำลังทำงานอยู่ก่อน

View Window เป็นองค์ประกอบหนึ่งของ Project ที่ใช้ในการนำเสนอ (display) ข้อมูลแผนที่ หรือ เรียกว่า theme การเรียกค้น (query) การย่อ-ขยาย พื้นที่ที่น่าสนใจ (explore) และการวิเคราะห์ต่าง ๆ แต่ไม่ได้เปลี่ยนแปลงข้อมูลเดิม นอกจากจะมีการบันทึกเพิ่มเติม

Table Window เป็นอีกองค์ประกอบหนึ่งของ Project ที่ใช้ในการแสดงฐานข้อมูลของแผนที่หรือฐานข้อมูลอื่น ๆ ที่จะเก็บโดยใช้ dBase และ Arc View สามารถรับข้อมูลจากแหล่งอื่น ๆ ที่บันทึกไว้ได้

Chart Window เป็นการนำเสนอข้อมูลในตารางฐานข้อมูลทั้งที่เป็นตารางที่ติดมากับข้อมูลพื้นที่ หรือตารางฐานข้อมูลอื่น ซึ่งมีความสัมพันธ์กับข้อมูลเชิงพื้นที่ ทั้งนี้ข้อมูลที่นำมาเสนออาจให้ความกระจ่างมากกว่าข้อมูลที่เสนอในรูปแบบตาราง

Layout Window คือแผนที่ที่ประกอบด้วยข้อมูลต่าง ๆ เช่น ชั้นข้อมูล (Themes) แผนภูมิ (Chart) หรือสัญลักษณ์ต่าง ๆ ทั้งที่ทำโดย Arc View หรือนำเข้าจากแหล่งข้อมูลอื่น

ภาพ 1 แสดงตัวอย่างรูปแสดง Window ในโปรแกรม Arc view

โปรแกรม SPSS

1. ความหมายของสถิติ

คำว่า สถิติ (Statistics) มาจากภาษาเยอรมันว่า Statistik มีรากศัพท์มากจาก Stat หมายถึง ข้อมูล หรือสารสนเทศ ซึ่งจะอำนวยประโยชน์ต่อการบริหารประเทศในด้านต่าง ๆ เช่น การทำสำมะโนครัว เพื่อตะทราบจำนวนพลเมืองในประเทศทั้งหมด ในสมัยต่อมา คำว่า สถิติ ได้หมายถึงตัวเลขหรือข้อมูลที่ได้จากการเก็บรวบรวม เช่น จำนวนผู้ประสบอุบัติเหตุบนท้องถนน อัตราการเกิดของเด็กทารก ปริมาณน้ำฝนในแต่ละปี เป็นต้น สถิติในความหมายนี้เป็นทั้ง วิทยาศาสตร์ และศิลปะศาสตร์ เรียกว่าสถิติศาสตร์

2. ประเภทของสถิติ

สถิติแบ่งออกเป็น 2 ประเภทคือ

2.1 สถิติพรรณนา (Descriptive Statistics) เป็นสถิติที่ใช้อธิบายคุณลักษณะของสิ่งที่ต้องการศึกษากลุ่มใดกลุ่มหนึ่ง ไม่สามารถอ้างอิงไปยังกลุ่มอื่น ๆ ได้ สถิติที่อยู่ในประเภทนี้ เช่น ค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม ส่วนเบี่ยงเบนมาตรฐาน ค่าพิสัย ฯลฯ

2.2 สถิติอ้างอิง (Inferential Statistics) เป็นสถิติที่ใช้อธิบายคุณลักษณะของสิ่งที่ต้องการศึกษากลุ่มใดกลุ่มหนึ่ง หรือหลายกลุ่ม แล้วสามารถอ้างอิงไปยังกลุ่มประชากรได้ โดยกลุ่มที่นำมาศึกษาจะต้องเป็นตัวแทนที่ดีของประชากร ตัวแทนที่ดีของประชากรได้มาโดยวิธีสุ่มตัวอย่าง และตัวแทนที่ดีของประชากรเรียกว่ากลุ่มตัวอย่าง สถิติอ้างอิงสามารถแบ่งออกได้เป็น 2 ประเภท คือ

2.2.1 สถิติมีพารามิเตอร์ (Parametric Statistics) เป็นวิธีการทางสถิติที่จะต้องเป็นไปตามข้อตกลงเบื้องต้น 3 ประการดังนี้

1) ข้อมูลที่เก็บรวบรวมได้จะต้องอยู่ในระดับช่วงขึ้นไป (Interval Scale)

2) ข้อมูลที่เก็บรวบรวมได้จากกลุ่มตัวอย่างจะต้องมีการแจกแจงเป็นโค้งปกติ

3) กลุ่มประชากรแต่ละกลุ่มที่นำมาศึกษาจะต้องมีความแปรปรวนไม่เท่ากันสถิติที่อยู่ในประเภทนี้ เช่น t-test, Z-test, ANOVA, Regression ฯลฯ

2.2.2 สถิติไร้พารามิเตอร์ (Nonparametric Statistics) เป็นวิธีการทางสถิติที่สามารถนำมาใช้ได้โดยปราศจากข้อตกลงเบื้องต้นทั้ง 3 ประการข้างต้น สถิติที่อยู่ในประเภทนี้ เช่น ไคสแควร์ Median Test, Sign test ฯลฯ

โดยปกติแล้วนักวิจัยมักนิยมใช้สถิติมีพารามิเตอร์ทั้งนี้เพราะผลลัพธ์ที่ได้จากการใช้สถิติมีพารามิเตอร์มีอำนาจการทดสอบ (Power of Test) สูงกว่าการใช้สถิติไร้พารามิเตอร์ สถิติมีพารามิเตอร์เป็นการทดสอบที่ได้มาตรฐาน มีขั้นตอนต่างๆที่สมบูรณ์ ดังนั้นเมื่อมีข้อมูลมีคุณสมบัติที่สอดคล้องกับข้อตกลงเบื้องต้นสามประการในการใช้สถิติมีพารามิเตอร์จึงไม่มีผู้ใดคิดที่จะหันไปใช้สถิติไร้พารามิเตอร์ในการทดสอบสมมติฐาน

3. **โปรแกรมสำเร็จรูปทางสถิติ**

ในการวิเคราะห์ข้อมูลโดยใช้สถิตินั้น หากข้อมูลมีปริมาณน้อยเราสามารถคำนวณด้วยเครื่องคิดเลขขนาดเล็ก ๆ ได้ แต่ในความเป็นจริงแล้ว การทำวิจัยเรื่องใด ๆ ก็ตาม ปริมาณข้อมูลจะมีมากการคำนวณด้วยเครื่องคิดเลขขนาดเล็กจึงเป็นการเสียเวลาและแรงงานมาก ซ้ำยังอาจเกิดความผิดพลาดได้ง่าย จึงจำเป็นที่จะต้องนำปัญญาประดิษฐ์หรือคอมพิวเตอร์เข้ามาช่วยเหลือในการวิเคราะห์ข้อมูล เพราะการใช้คอมพิวเตอร์จะช่วยประหยัดเวลา แรงงาน และค่าใช้จ่ายลงได้มากทั้งยังมีความถูกต้องสูงด้วย

นักวิจัยส่วนใหญ่นิยมใช้โปรแกรมสำเร็จรูปในการวิเคราะห์ข้อมูลมากกว่าที่จะเขียนโปรแกรมขึ้นมาใช้เอง เนื่องจากสามารถเรียนรู้ได้รวดเร็วและไม่ต้องเสียเวลาในการเขียนโปรแกรม

ในระยะแรกโปรแกรมสำเร็จรูปสามารถวิเคราะห์ข้อมูลทางสถิติบางอย่างได้ เช่นโปรแกรมจำพวกกระดาษทด (Spread Sheet) ได้แก่โปรแกรม Lotus 1-2-3 เป็นต้น ในการวิเคราะห์สถิติขั้นสูงโปรแกรมเหล่านี้จะไม่สามารถวิเคราะห์ได้ แต่สามารถสร้างแผนภูมิหรือกราฟได้อย่างสวยงาม ต่อมาจึงได้มีการพัฒนาโปรแกรมวิเคราะห์ข้อมูลทางสถิติโดยเฉพาะ ได้แก่ SAS (Statistical Analysis System), MINITAB Student ware, SP (Statistical Package), ISP (Interactive Statistical Programs), SPSS/PC+ (Statistical Package for the Social Sciences) เป็นต้น แต่โปรแกรมทางสถิติเหล่านี้ ไม่สามารถสร้างแผนภูมิหรือกราฟได้อย่างสวยงาม จึงต้องพึ่งพาอาศัยกันระหว่างโปรแกรม 2 ประเภทนี้ต่อมาจึงได้มีผู้คิดค้นโปรแกรมสำเร็จรูปที่สามารถวิเคราะห์สถิติขั้นสูงได้ และยังสามารถสร้างแผนภูมิหรือกราฟได้หลายรูปแบบและสวยงาม โปรแกรมนี้มีชื่อว่า SPSS for Windows

4. **การวิเคราะห์การถดถอย**

การวิเคราะห์การถดถอยเป็นกระบวนการทางสถิติเพื่อให้ได้สมการถดถอยสำหรับทำนายปรากฏการณ์ต่าง ๆ สิ่งที่ถูกทำนายเรียกว่าตัวแปรเกณฑ์หรือตัวแปรตาม ตัวแปรทำนาย คือตัวแปรอิสระในกระบวนการนี้ ตัวแปรตามหรือตัวแปรเกณฑ์จะมีเพียงตัวเดียว ส่วนตัวแปรอิสระหรือตัวแปรทำนายจะมีกี่ตัวก็ได้ ถ้าหากมีตัวเดียว จะเรียกว่าการถดถอยอย่างง่าย หากมีตั้งแต่ 2 ตัวขึ้นไป จะเรียกว่า การถดถอยพหุคูณ

4.1 การวิเคราะห์การถดถอยอย่างง่าย

การวิเคราะห์การถดถอยอย่างง่ายจะประกอบไปด้วยตัวแปรเกณฑ์ 1 ตัวและตัวแปรทำนาย 1 ตัว มีสมการเขียนอยู่ในรูปของค่าพารามิเตอร์ว่า

$$Y_i = a + bX + e_i$$

เมื่อ Y_i = คะแนนของบุคคลที่ i; a = ค่าเฉลี่ยของประชากรเมื่อค่า X = 0 หรือจุดตัดแกน Y; b = สัมประสิทธิ์การถดถอยในประชากร หรือความชันของเส้นถดถอย; e_i = ความคลาดเคลื่อนของบุคคลที่ i

สัมประสิทธิ์การถดถอย (b) เป็นตัวบ่งชี้อิทธิพลของตัวแปรอิสระบนตัวแปรตาม อธิบายได้ง่ายๆว่า เมื่อ X เปลี่ยนแปลงไป 1 หน่วยแล้ว Y จะเปลี่ยนแปลงไป b หน่วย

ในความเป็นจริงเราไม่สามารถศึกษาจากประชากรได้ทั้งหมดเนื่องจากมีปริมาณมาก เราจึงต้องศึกษากับกลุ่มตัวอย่างที่สามารถอ้างอิงไปสู่ประชากรได้ ดังนั้นสมการถดถอยที่เขียนอยู่ในรูปของสถิติมีว่า

$$Y_i = a + bX + e_i$$

เมื่อ a = ตัวประมาณค่าของ a ; b คือตัวประมาณค่าของ b ; และ e คือตัวประมาณค่าของ e ; แต่เนื่องจาก $e = Y - Y'$ เมื่อค่า Y = คะแนนที่สังเกตได้ และ Y' = คะแนนที่ได้จากสมการทำนาย ดังนั้น $\sum(Y - Y')^2$ จะต้องมีค่าน้อยที่สุด (least – squares solution) และ $\sum e = 0$ ดังนั้นสมการจึงเหลือเพียงแค่ค่า a และ b เท่านั้น

$$Y_i = a + bX$$

เมื่อ $$b = \frac{\sum xy}{\sum x^2}$$

$$a = \bar{Y} - b\bar{X}$$

4.2 การวิเคราะห์การถดถอยพหุคูณ

จะประกอบไปด้วยตัวแปรเกณฑ์ 1 ตัวและตัวแปรทำนายตั้งแต่ 2 ตัวขึ้นไป สามารถเขียนเป็นสมการถดถอยได้ว่า

$$Y' = a + b_1X_1 + b_2X_2 + \ldots + b_kX_k$$

ถ้าหากมีตัวแปรทำนาย 2 ตัว สามารถเขียนสมการได้ว่า

$$Y' = a + b_1X_1 + b_2X_2$$

คำนวณค่า b ด้วยสูตร

$$b_1 = \frac{(\sum x_2^2)(\sum x_1 y) - (\sum x_1 x_2)(\sum x_2 Y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$b_2 = \frac{(\sum x_1^2)(\sum x_{21} y) - (\sum x_1 x_2)(\sum x_1 Y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

ด้วยสูตร

$$a = \overline{Y} + b_1 \overline{X}_1 + b_2 \overline{X}_2$$

4.3 การทดสอบนัยสำคัญ

จะทดสอบส่วนต่าง ๆ ของสมการถดถอยดังนี้

4.3.1 ทดสอบสัมประสิทธิ์การทำนาย (R^2)

สมมติฐาน

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

สถิติทดสอบ

$$F = \frac{R^2/k}{(1-R^2)/(N-k-1)}$$

เมื่อ df = k และ (N - k - 1)

4.3.2 ทดสอบสัมประสิทธิ์การถดถอย (b)

สมมติฐาน

$H_0 : \beta = 0$

$H_1 : \beta \neq 0$

สถิติทดสอบ

$$t = \frac{b_i}{S_{b_i}}$$

เมื่อ df = (N - k - 1)

4.4 วิธีการคัดเลือกตัวแปร

วิธีการคัดเลือกตัวแปรเข้าสมการเพื่อให้สมการสามารถทำนายตัวแปรเกณฑ์ได้สูงสุดมีวิธีการคัดเลือกตัวแปรดังนี้

4.4.1 วิธีการเลือกแบบก้าวหน้า (Forward Selection)

วิธีการนี้จะเป็นการเลือกตัวแปรทำนายที่มีสหสัมพันธ์กับตัวแปรตามสูงที่สุดเข้าสมการก่อนส่วนตัวแปรที่เหลือจะคำนวณหาสหสัมพันธ์แบบแยกส่วน (partial correlation) โดยเป็นความสัมพันธ์เฉพาะตัวแปรที่เหลือตัวนั้นกับตัวแปรตามโดยขจัดอิทธิพลของตัวแปรอื่น ๆ ออก ถ้าตัวแปรใดมีความสัมพันธ์กันสูงอย่างมีนัยสำคัญทางสถิติก็จะนำเข้าสมการต่อไป จะทำแบบนี้จนกระทั่งสหสัมพันธ์แบบแยกส่วนระหว่างตัวแปรอิสระที่ไม่ได้นำเข้าสมการแต่ละตัวกับตัวแปรตามมีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติ ก็จะหยุดการคัดเลือกและได้สมการที่มีสัมประสิทธิ์การทำนายสูง

4.4.2 วิธีการเลือกแบบถอยหลัง (Backward Selection)

วิธีการนี้เป็นการนำตัวแปรทำนายทั้งหมดเข้าสมการ จากนั้นก็จะค่อย ๆ ขจัดตัวแปรทำนายออกทีละตัว โดยจะหาสหสัมพันธ์ระหว่างตัวแปรทำนายที่อยู่ในสมการแต่ละตัวกับตัวแปรตามเมื่อขจัดตัวแปรทำนายอื่น ๆ ออกแล้วหากทดสอบค่าสหสัมพันธ์แล้วพบว่าไม่มีนัยสำคัญทางสถิติ ก็จะขจัดออกจากสมการ แล้วดำเนินการทดสอบตัวแปรที่เหลืออยู่ในสมการต่อไป จนกระทั่งสหสัมพันธ์ระหว่างตัวแปรทำนายแต่ละตัวกับตัวแปรตามเมื่อขจัดตัวแปรอิสระอื่น ออกแล้วพบว่ามีนัยสำคัญทางสถิติ ก็จะหยุดการคัดเลือก และได้สมการการทดสอบที่มีสัมประสิทธิ์การทำนายสูงสุด

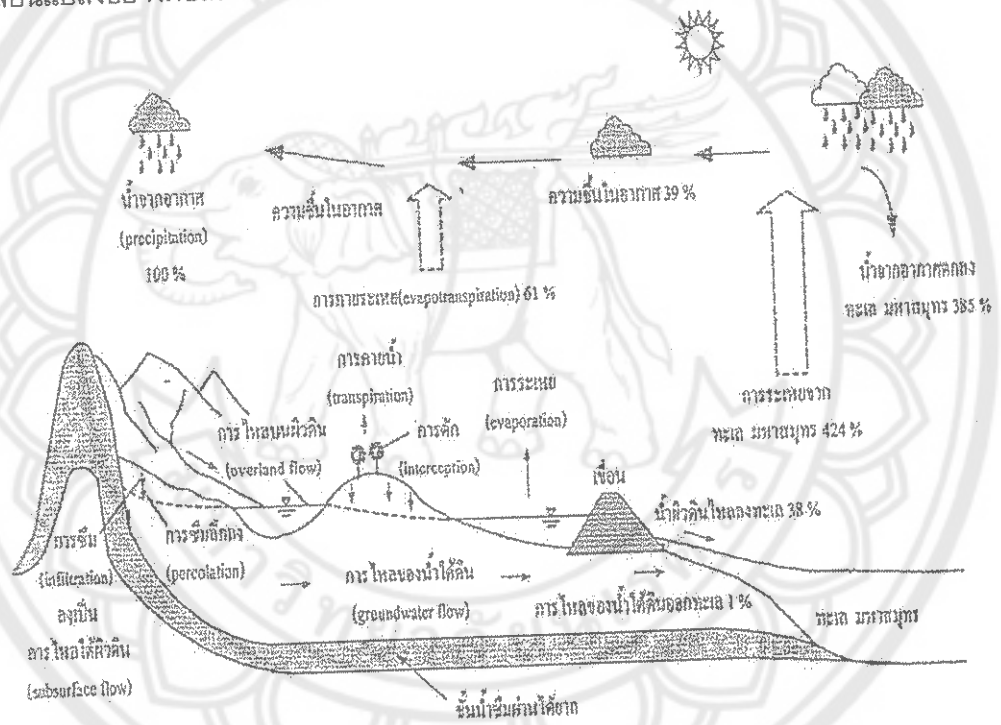
4.4.3 วิธีการเลือกแบบลำดับขั้น (Stepwise Selection)

การคัดเลือกแบบนี้เป็นการผสมผสานระหว่างวิธีการคัดเลือกตัวแปรทำนายทั้งสองวิธีที่กล่าวมาแล้วเข้าด้วยกัน ในขั้นแรกจะเลือกตัวแปรทำนายที่มีสหสัมพันธ์กับตัวแปรตามสูงที่สุดเข้าสมการก่อนจากนั้นก็จะทดสอบตัวแปรที่ไม่ได้อยู่ในสมการว่าจะมีตัวทำนายตัวใดบ้างมีสิทธิ์เข้ามาอยู่ในสมการด้วยวิธีการคัดเลือกแบบก้าวหน้า (Forward Selection) และขณะเดียวกันก็จะทดสอบตัวแปรที่อยู่ในสมการด้วยว่าตัวแปรทำนายที่อยู่ในสมการตัวแปรใดมีโอกาสที่จะถูกขจัดออกจากสมการด้วยวิธีการคัดเลือกแบบถอยหลัง (Backward Selection) โดยจะกระทำการคัดเลือกผสมทั้งสองวิธีนี้ในทุกขั้นตอนจนกระทั่งไม่มีตัวแปรใดที่ถูกคัดออกจากสมการ และไม่มีตัวแปรใดที่จะถูกนำเข้ามาในสมการ กระบวนการก็จะยุติและได้สมการถดถอยที่มีสัมประสิทธิ์การทำนายสูงสุด

มีโอกาสที่จะถูกขจัดออกจากสมการด้วยวิธีการคัดเลือกแบบถอยหลัง (Backward Selection) โดยจะกระทำการคัดเลือกผสมทั้งสองวิธีนี้ในทุกขั้นตอนจนกระทั่งไม่มีตัวแปรใดที่ถูกคัดออกจากสมการ และไม่มีตัวแปรใดที่จะถูกนำเข้ามาในสมการ กระบวนการก็จะยุติและได้สมการถดถอยที่มีสัมประสิทธิ์การทำนายสูงสุด

วงจรอุทกวิทยา (Hydrologic Cycle)

วงจรอุทกวิยาเป็นศูนย์รวมในการศึกษาอุทกวิทยา โดยวงจรอุทกวิทยาเป็นวงจรที่ไม่มีจุดเริ่มต้นและจุดสุดท้ายของกระบวนการเปลี่ยนแปลงของปริมาณน้ำในโลกเพราะมีการเปลี่ยนแปลงอย่างต่อเนื่องตลอดเวลา ซึ่งภาพรวมของวงจรอุทกวิทยาสามารถแสดงได้ดังภาพ 2

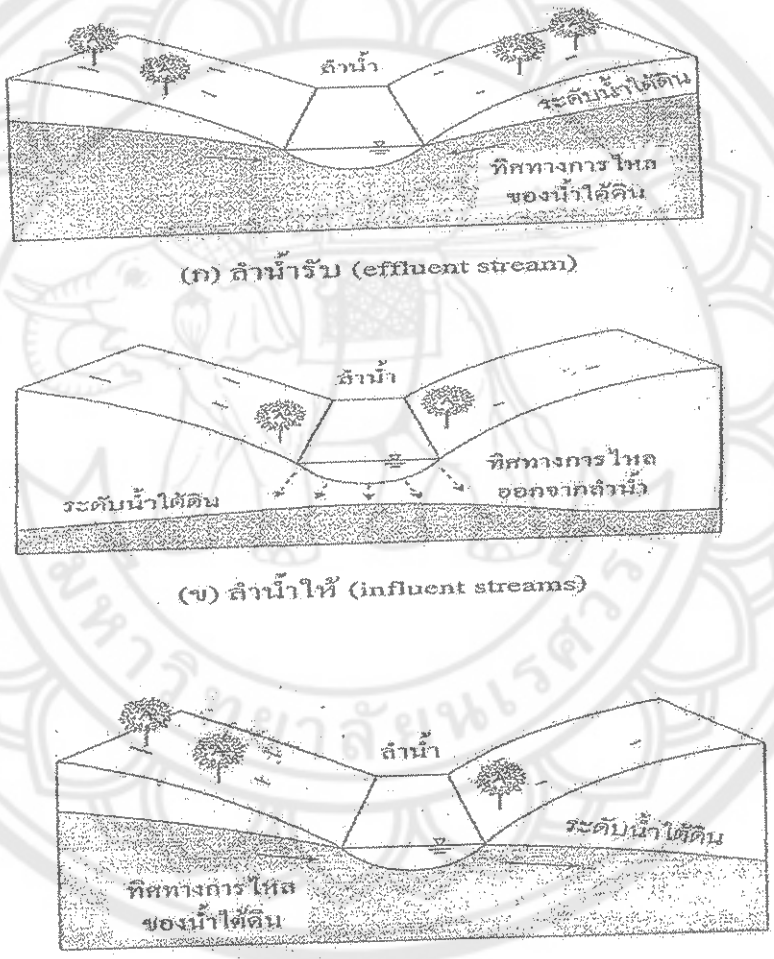


ภาพ 2 แสดงวงจรอุทกวิทยา

ที่มา: วิโรจน์ ชัยธรรม, 2536

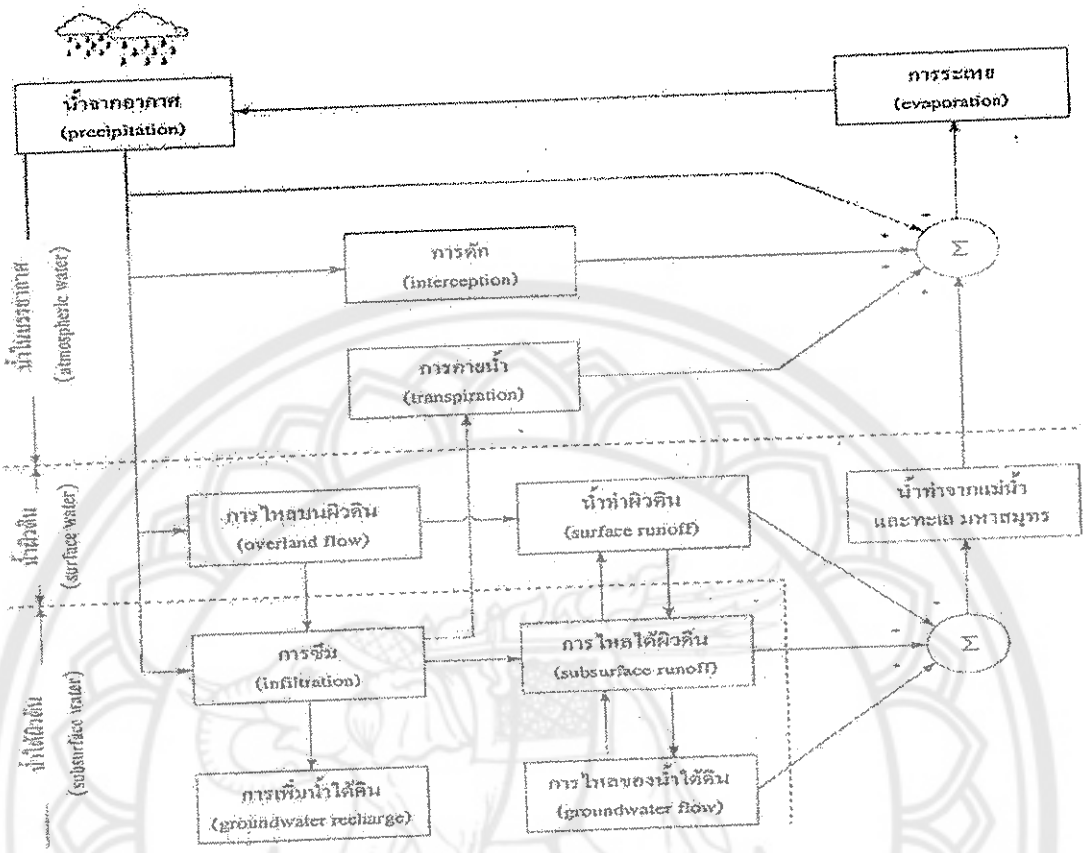
เมื่อพิจารณาภาพ 2 จะเห็นได้ว่าเมื่อน้ำมีการระเหย (evaporation) จากทะเลมหาสมุทรและน้ำที่สะสมอยู่บนพื้นแผ่นดิน เช่น อ่างเก็บน้ำ ห้วย หนอง คลอง บึง หรือจากน้ำใต้ผิวดินบางส่วนขึ้นสู่บรรยากาศเป็นไอน้ำ (water vapor) ซึ่งมีการลอยตัวขึ้นไปสะสมจนกระทั่งเกิดกระบวนการควบแน่นและกลั่นตัวกลายเป็นไอน้ำจากอากาศ (precipitation) ตกลงมาสู่ทะเล

ระหว่างช่องว่างของเม็ดดินหรือรอยหินแตกลงไปในน้ำใต้ดิน (groundwater) ซึ่งถ้าน้ำมีการให้น้ำแก่น้ำใต้ดินจะเรียกว่าลำน้ำให้และลำน้ำรับ (effluent stream) ดังภาพ 3 (ก) ถ้าลำน้ำมีการให้น้ำแก่น้ำใต้ดินจะเรียกว่า ลำน้ำให้ (influent stream) ดังภาพ 3 (ข) นอกจากนี้ยังมีลำน้ำบางแห่งที่เป็นลำน้ำรับและลำน้ำให้ดังภาพ 3 (ค) ซึ่งท้ายที่สุดแล้วน้ำใต้ดินมักจะมีแนวการไหลซึมออกสู่แหล่งน้ำหรือทะเลมหาสมุทร แล้วเกิดการระเหยกลับสู่บรรยากาศหมุนเวียนอย่างต่อเนื่องเป็นวงจรอุทกวิทยาอีกครั้ง



ภาพ 3 แสดงลำน้ำรับและลำน้ำให้

ที่มา: วิโรจน์ ชัยธรรม, 2536



ภาพ 4 แสดงแผนผังวงจรอุทกวิทยา

ที่มา: วีระพล แต้สมบัติ, 2528

การวางโครงการเพื่อป้องกันอุทกภัยเบื้องต้น

การวางโครงการ หมายถึง การวางแผนการใช้ทรัพยากรและแผนปฏิบัติอย่างมีระบบ เพื่อก่อให้เกิดการประหยัดและได้ประโยชน์ตามวัตถุประสงค์ของโครงการ โดยจะเลือกดำเนินการโครงการที่ใช้ทรัพยากรธรรมชาติอย่างประหยัด แต่เกิดประโยชน์และสามารถตอบสนองวัตถุประสงค์ของโครงการได้ ในการศึกษาวางโครงการจึงต้องนำองค์ประกอบต่าง ๆ มาทำการวิเคราะห์ทั้งระบบ (System Analysis) เพื่อหาแนวทางเลือกในกรณีต่าง ๆ (Alternative cases) ที่เป็นไปได้ โดยการวางโครงการในงานวิจัยนี้มุ่งเน้นไปที่การป้องกันอุทกภัยเป็นหลัก ยังไม่ครอบคลุมถึงการใช้ประโยชน์ของอ่างเก็บน้ำเพื่อการเกษตร, ผลกระทบทางสิ่งแวดล้อม และด้านอื่น ๆ เพื่อหาแนวทางที่ดีและเหมาะสมที่สุด การจัดทำรายงานวางโครงการ มีขั้นตอนและวิธีการดำเนินการดังนี้

1. **การกำหนดจุดที่ตั้งหัวงาน**

การพิจารณาถึงความต้องการและปัญหาที่เกิดขึ้นในเขตพื้นที่โครงการ ทำให้ทราบถึงขอบเขตโครงการและรายละเอียดเบื้องต้นของโครงการ ประกอบด้วย พิกัดจุดที่ตั้งหัวงาน ชื่อลำน้ำ และบริเวณที่จะได้รับประโยชน์จากโครงการ ซึ่งดูได้จากแผนที่มาตราส่วน 1:50,000 หลักเกณฑ์ในการพิจารณาเลือกจุดที่ตั้งหัวงาน มีดังนี้

พิจารณาบริเวณที่มีสภาพภูมิประเทศเหมาะสม (Topo Suitability)

1.1 ถ้าเป็นหัวงานประเภทอ่างเก็บน้ำ ควรเป็นบริเวณเขาแคบ หรือมีสันเนินอยู่ใกล้กัน บริเวณที่เก็บกักน้ำมีลักษณะเป็นแอ่งกว้าง ทำให้ไม่ต้องสร้างทำนบดินยาวเกินไปหรือสูงเกินไป

1.2 ถ้าเป็นหัวงานประเภทฝายหรือเขื่อนทดน้ำ ควรเลือกทำเลที่มีช่องลัด แนวลำน้ำบริเวณเหนือน้ำและท้ายน้ำค่อนข้างตรง มีตลิ่งสูงและมั่นคง ลำน้ำมีน้ำไหลตลอดปี ไม่คดเคี้ยวมากนัก และเป็นบริเวณที่น้ำไม่กัดเซาะตลิ่ง

2. **การรวบรวมสถิติข้อมูลต่าง ๆที่ใช้ในการประกอบพิจารณาโครงการ**

สถิติข้อมูลต่าง ๆ ที่ใช้ประกอบการพิจารณาวางโครงการ มีดังนี้

2.1 แผนที่ขนาดต่าง ๆ ที่ครอบคลุมเขตพื้นที่โครงการ เช่น แผนที่มาตราส่วน 1:50,000 แผนที่มาตราส่วน 1:250,000

2.2 สถิติข้อมูลทางด้านอุตินิยมวิทยาและอุทกวิทยาในเขตพื้นที่โครงการ

2.2.1 สถิติข้อมูลทางด้านอุตุนิยมวิทยา ได้แก่ ปริมาณน้ำฝนเฉลี่ยรายเดือน อุณหภูมิเฉลี่ยรายเดือน อัตราการระเหย ความชื้น ความเร็วและทิศทางของลม ฯลฯ

2.2.2 สถิติข้อมูลทางด้านอุทกวิทยา (ภายในลุ่มน้ำเดียวกันหรือลุ่มน้ำใกล้เคียง) ได้แก่ ระดับน้ำเฉลี่ยรายเดือน และปริมาณน้ำท่ารายเดือน

3. **การวิเคราะห์ปริมาณน้ำฝน** (Runoff Analysis)

ข้อมูลปริมาณน้ำฝนได้จากสถานีตรวจวัดของหน่วยงานต่าง ๆ เช่น กรมอุตุนิยมวิทยา กรมชลประทาน สถานีวิจัยทดลองทางด้านการเกษตร หรือสถานีหน่วยจัดการต้นน้ำต่าง ๆ ของกรมป่าไม้ โดยจะพิจารณาข้อมูลจากสถานีที่อยู่ใกล้เคียงกับพื้นที่หัวงานโครงการ การวิเคราะห์ปริมาณฝนจะแยกออกได้เป็นสภาพฝนโดยทั่วไปในเขตพื้นที่ลุ่มน้ำของโครงการที่ศึกษา ซึ่งจะต้องทำการวิเคราะห์สภาพฝนรายปี รูปแบบการผันแปรตามฤดูกาล การแพร่กระจายตามพื้นที่ ปริมาณฝนรายเดือน และเปอร์เซ็นต์การแพร่กระจายเป็นรายเดือนของฝน จำนวนวันที่ฝนตกเฉลี่ยรายเดือน/รายปี สำหรับการวิเคราะห์ปริมาณน้ำฝนของสถานีหลัก จำเป็นต้องมีการตรวจสอบความเชื่อถือได้ของข้อมูลก่อนที่จะนำไปใช้ในการศึกษา ประโยชน์ของการวิเคราะห์ปริมาณน้ำฝนที่สามารถนำไปใช้งานได้ มีดังนี้

3.1 นำไปใช้ในการศึกษาเกี่ยวกับการประเมินปริมาณน้ำท่าในกรณีที่ข้อมูลมีไม่เพียงพอ

3.2 นำไปใช้คำนวณหาปริมาณฝนใช้การ และคำนวณปริมาณน้ำชลประทานที่พืชต้องการ

3.3 ใช้สร้างกราฟความสัมพันธ์ระหว่างความเข้ม ช่วงเวลา-ความถี่ของฝน เพื่อคำนวณหาปริมาณน้ำนองสูงสุดสำหรับการออกแบบรายละเอียดอาคารหัวงาน และอาคารประกอบอื่น ๆ

3.4 ใช้วิเคราะห์ความถี่ของการเกิด (Frequency) ของปริมาณฝนสูงสุดในคาบปี (Return Period) ต่าง ๆ เพื่อคำนวณหากราฟน้ำนองสูงสุดในรอบปีการเกิดซ้ำต่าง ๆ สำหรับการออกแบบอาคารระบายน้ำล้นและอาคารประกอบอื่น ๆ

4. **การวิเคราะห์ปริมาณน้ำท่า** (Runoff Analysis)

น้ำท่า (Runoff) หมายถึง น้ำฝนส่วนที่ตกลงบนผิวดินแล้วไหลไปตามผิวดินลงสู่ลำน้ำ หลังจากที่บางส่วนได้ระเหยและรั่วซึมลงไปในดินแล้ว ในระหว่างที่น้ำไหลไปตามผิวดินเรียกว่า Overland Flow เมื่อไหลลงลำน้ำแล้วเรียกว่า Stream Flow ตามปกติปริมาณน้ำส่วนที่ไหลลงลำน้ำจะมีค่าประมาณ 15%-35% ของปริมาณฝนที่วัดได้ ทั้งนี้ขึ้นอยู่กับสภาพภูมิอากาศ ดิน ทางน้ำ ลักษณะของลุ่มน้ำ สภาพพื้นที่ และสภาพป่าไม้ในเขตลุ่มน้ำ ฯลฯ

ในการพิจารณาวางโครงการชลประทานนั้น ปริมาณน้ำนับว่าเป็นองค์ประกอบสำคัญในการพิจารณาว่าควรจะทำโครงการชลประทานหรือไม่ และควรเป็นโครงการชลประทานประเภทใด กล่าวคือ ถ้าเป็นลำน้ำที่มีปริมาณน้ำไหลสม่ำเสมอเกือบตลอดปี ก็อาจพิจารณาโครงการประเภทฝายทดน้ำ หากเป็นลำน้ำที่มีปริมาณน้ำไหลมากเฉพาะในช่วงฤดูฝน ถ้าฝนไม่ตกจะมีปริมาณน้ำเพียงเล็กน้อยหรือไม่มีเลย ในกรณีนี้อาจพิจารณาโครงการในลักษณะอ่างเก็บน้ำ นอกจากนี้การแผ่กระจายของน้ำในแต่ละเดือนจะเป็นเครื่องกำหนดขนาดของพื้นที่โครงการด้วย

5. **การคำนวณปริมาณน้ำท่า** (Runoff Estimation)

การคำนวณปริมาณน้ำท่า มีขั้นตอนดำเนินงานดังนี้

5.1 กำหนดจุดที่ตั้งหัวงาน จากแผนที่มาตราส่วน 1:50,000

5.2 หาพื้นที่รับน้ำฝน (Drainage Area) เป็นพื้นที่ลุ่มน้ำเหนือแนวที่ตั้งหัวงาน ซึ่งมีอาณาเขตล้อมรอบบรรจบกันเป็นวงปิดด้วยแนวสันปันน้ำ (Watershed) หรือสันเนินสูงสุด ภายในพื้นที่รับน้ำนี้หากมีฝนตกจนเกินน้ำไหลนองแล้ว น้ำทั้งหมดจะไหลลงมายังที่ตั้งหัวงาน การลากเส้นสันปันน้ำจากจุดที่ตั้งหัวงาน ต้องตรวจดูว่าลำน้ำนั้นมีขอบเขตของลำน้ำ และลำน้ำห้วยสาขาครอบคลุมพื้นที่ถึงบริเวณใด เลือกจุดสูงสุด (บริเวณต้นน้ำ) ลากเส้นตามแนวสันเนินลงมายังจุด

ที่ตั้งหัวงาน สังเกตลักษณะแนวสันเนินซึ่งจะตรงกันข้ามกับร่องน้ำให้ดี (เส้น Contour จะมีลักษณะแหลมหรือมน สวนทิศทางการไหลของน้ำ) โดยเฉพาะเส้นชั้น Contour บริเวณที่เป็นภูเขาสลับซับซ้อนจะมีปัญหาการลากเส้นสันปันน้ำมาก ต้องสังเกตจุดแสดงระดับความสูง ขนาดและลักษณะของพื้นที่รับน้ำฝนดังกล่าว มีความสัมพันธ์กับปริมาณน้ำต้นทุนที่ไหลผ่านหัวงานตลอดจนอัตราน้ำหลากสูงสุดที่หัวงานจะได้รับ นอกจากนี้ยังมีองค์ประกอบอื่นที่เกี่ยวข้องอีก เช่น รูปร่างของลุ่มน้ำ ความลาดชันของลำน้ำ ทิศทางรับน้ำฝน ปริมาณน้ำฝนที่ตกในลุ่มน้ำ ชนิดของดิน ระบบลำน้ำ ชนิดของพืชที่ปกคลุม ฯลฯ

5.2.1 เลือกหาสถานีน้ำฝนที่อยู่ใกล้เคียงและคาดว่าจะอาศัยเป็นตัวแทนของฝนที่ตกในเขตลุ่มน้ำที่ตั้งหัวงานได้ ในกรณีที่สถิติน้ำฝนของสถานีที่เลือกมีสถิติการเก็บข้อมูลติดต่อกันน้อยกว่า 15 ปี ให้เลือกใช้ข้อมูลจากสถานีอื่นในบริเวณใกล้เคียงที่มีการเก็บข้อมูลติดต่อกันอย่างน้อย 15 ปี มาใช้แทนได้

5.2.2 การคำนวณหาปริมาณน้ำท่าที่ไหลผ่านหัวงานหาได้โดยวิธี Terrain โดยใช้กราฟ Runoff Coefficient Chart ช่วยประกอบการพิจารณาโดยดูจากสภาพฝน สภาพลุ่มน้ำ สภาพดิน สภาพภูมิประเทศ ความลาดเทของลำน้ำและชนิดของพืชที่ปกคลุมเป็นเกณฑ์ สามารถแบ่งได้เป็น 5 เกณฑ์พร้อมสมการคำนวณ Runoff Coefficient แสดงดังต่อไปนี้

Type A= Steep Mountainous area, no paddy field.

$Y_A = 0.1305X + 8.501$

Type B = Rather steep area, open forest.

$Y_B = 0.1295X + 5.711$

Type C = Rolling area, open forest, some paddy field.

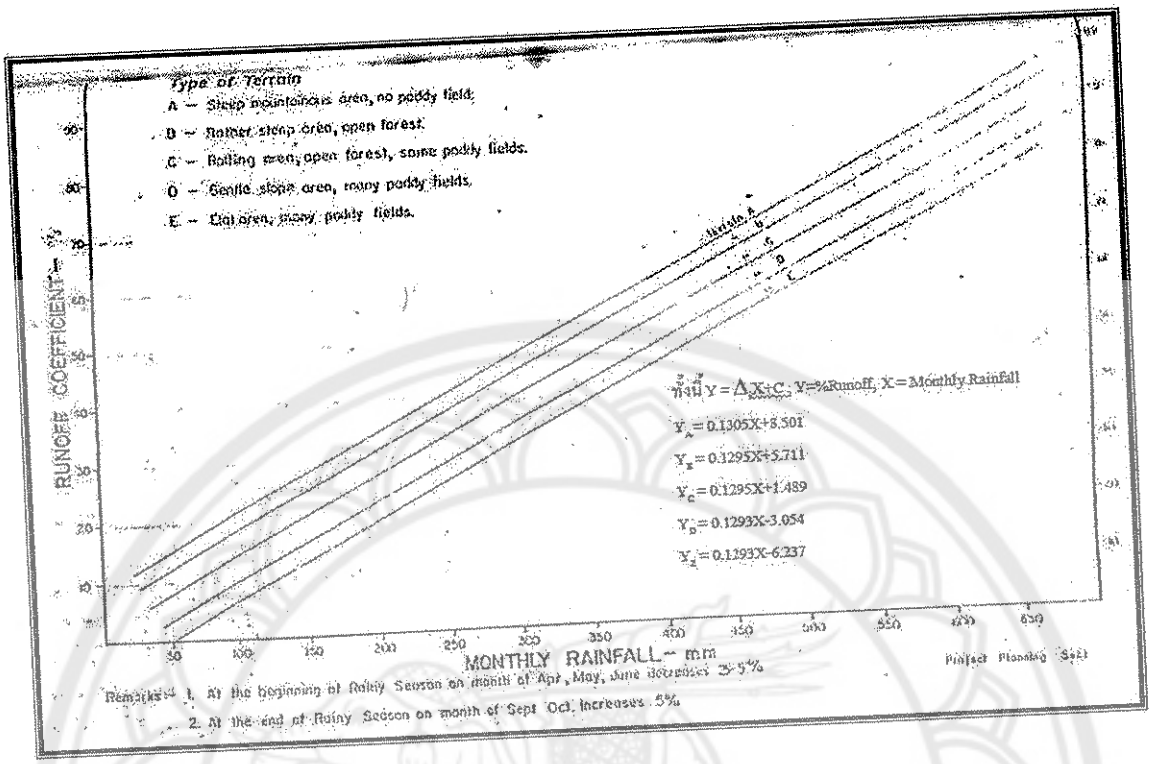
$Y_C = 0.1295X + 1.489$

Type D = Gentle slope area, many paddy field.

$Y_D = 0.1293X - 3.054$

Type E = Flat area, many paddy field.

$Y_E = 0.1293X - 6.237$



ภาพ 5 แสดง Runoff Estimation Chart สำหรับพื้นที่ไม่เกิน 2,000 ตร.กม.

ที่มา: พิมาย จันทะชิต, 2548.

วิธีการวิเคราะห์ด้วย Surrogate Worth Trade-off

วิธีการวิเคราะห์ Surrogate Worth Trade-offเป็นการหาความสัมพันธ์ของตัวแปรตาม Y สองตัวหรือมากกว่า กับ ตัวแปรอิสระ X สองตัวแปรหรือมากกว่า การถดถอยพหุคูณ(Multiple regression) ก็เป็นอีกวิธีการหนึ่งที่จะศึกษาความสัมพันธ์ระหว่างตัวแปรตาม Y และตัวแปรอิสระ X สองตัวหรือมากกว่า การวิเคราะห์การถดถอยเชิงเส้นพหุคูณคือ การศึกษาอิทธิพลของตัวแปรอิสระหลายตัวรวมกันมีผลกระทบต่อตัวแปรตาม สมการถดถอยพหุคูณจะชี้ให้เห็นความสัมพันธ์เฉลี่ยระหว่างตัวแปรอิสระเหล่านี้ที่มีต่อตัวแปรตาม ทำให้เราสามารถใช้ความสัมพันธ์นี้พยากรณ์ตัวแปรตามได้

รูปแบบการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i \ , i = 1,2, \ldots, n$$

เมื่อ Y_i = ตัวแปรตาม

$X_{i1}, X_{i2}, \ldots, X_{ik}$ = ตัวแปรอิสระ

$\beta_0, \beta_1, \ldots, \beta_k$ = พารามิเตอร์

ε_i = ค่าความคลาดเคลื่อนของค่าสังเกตที่ i

ข้อสมมติของรูปแบบการถดถอยเชิงเส้นพหุคูณมีดังนี้

1. ค่า X สำหรับตัวแปรอิสระแต่ละตัวเป็นค่าที่ทราบค่า นั่นคือ X ไม่ใช่ตัวสุ่ม
2. ความคลาดเคลื่อน ε_i เป็นตัวแปรสุ่มที่เป็นอิสระ มีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับศูนย์และความแปรปรวนเท่ากับ $\sigma 2$
3. ที่แต่ละเซตของค่า $X_1, X_2, \ldots, X_k$ Y เป็นอิสระมีการแจกแจงปกติที่มีค่าเฉลี่ย $\mu y(X_1, X_2, \ldots, X_k)$ เท่ากับ $\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}$ และค่าความแปรปรวนเท่ากับ $\sigma^2 Y|X_1, X_2, \ldots, X_k$ เมื่อ $\beta_0, \beta_1, \ldots, \beta_k$ และ $\sigma^2 Y|X_1, X_2, \ldots, X_k$ เป็นพารามิเตอร์ที่ไม่ทราบค่า
4. ที่แต่ละค่า $X_1, X_2, \ldots, X_k$ ความแปรปรวนของ Y ที่ $X_1, X_2, \ldots, X_k$ มีค่าเท่ากันนั่นคือ $\sigma^2 Y|X_1, X_2, \ldots, X_k = \sigma^2$ สำหรับทุกเซตของค่า $X_1, X_2, \ldots, X_k$
5. ไม่มีความสัมพันธ์อย่างสมบูรณ์ระหว่างตัวแปร $X_1, X_2, \ldots, X_k$ หรือไม่มีความสัมพันธ์กันเองเชิงเส้น (Multicollinearity)

ฟังก์ชันถดถอย

เนื่องจาก $E\{\varepsilon\} = 0$ จะได้ฟังก์ชันถดถอยหรือสมการถดถอยของประชากรรูปแบบฟังก์ชันถดถอย

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

หรือ

$$\mu_y(X_1, X_2, \ldots, X_k) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

นั่นคือ ค่าเฉลี่ยของ Y สำหรับทุกเซตของ $X_1, X_2, \ldots, X_k$ เท่ากับ $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$ พารามิเตอร์ β_j , j = 0,1,2,...,k เรียกสัมประสิทธิ์ถดถอย (regression coefficients) β_0 แทนระยะตัดแกน Y หรือ β_0 คือค่าเฉลี่ยของการแจกแจง Y เมื่อ $X_1 = X_2 = \ldots = X_k = 0$

β_j , j = 0,1,2,...,k แทนความชันของผิวการถดถอยของประชากร ที่มีตัวแปรอิสระที่ j β_j จะแสดงการเปลี่ยนแปลงของ Y เมื่อ X_j เปลี่ยนไปหนึ่งหน่วย ขณะที่ตัวแปรอิสระอื่นๆที่เหลืออยู่

ในรูปแบบ ค่าคงที่ ด้วยเหตุนี้จึงมักเรียก พารามิเตอร์ β_j , j = 0,1,2,...,k ว่าสัมประสิทธิ์การถดถอยบางส่วน (partial regression coefficients)

ในกรณีตัวแปรอิสระมี 2 ตัว คือ X_1 และ X_2 รูปแบบการถดถอย คือ

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i \ , i = 1,2,\ldots,n$$

และฟังก์ชันถดถอยหรือสมการถดถอยของประชากร คือ

$$\mu_y(X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยที่สุด

เราจะประมาณค่า $\beta_0,\beta_1,\beta_2,\ldots,\beta_k$ ซึ่งเป็นพารามิเตอร์ที่ไม่ทราบค่าของสมการถดถอย โดยใช้ข้อมูลของตัวอย่าง ดังนี้ ให้สมการถดถอยเชิงเส้นพหุคูณของตัวอย่าง (หรือที่ประมาณ) ที่มีตัวแปรอิสระ k ตัว คือ

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

เมื่อ $\hat{Y}$ คือค่าประมาณของค่าเฉลี่ยของ Y ซึ่งกำหนด $X_1,X_2,\ldots,X_k$ มาให้ หรือ $\hat{Y}$ คือค่าประมาณค่าของ $\mu_y(X_1, X_2, \ldots, X_k)$

b_0 คือค่าประมาณ β_0

b_i คือค่าประมาณ β_j

ความคลาดเคลื่อนของตัวอย่าง เขียนแทนด้วย e_i กำหนดโดย

$$e_i = Y_i - [b_0 + b_1 X_{i1} + b_2 X_{i2} + \ldots + b_k X_{ik}] \ , i = 1,2,\ldots,n$$

เรียก e_i ว่า residua

เช่นเดียวกับการถดถอยเชิงเส้นอย่างง่าย การประมาณค่า $\beta_0,\beta_1,\beta_2,\ldots,\beta_k$ โดยวิธีกำลังสองน้อยที่สุด (OLS) คือการหาค่า $b_0,b_1,b_2,\ldots,b_k$ ที่ทำให้ผลบวกกำลังสองของ residual มีค่าน้อยที่สุด

จากคุณสมบัติที่ว่า $\sum e_i^2$ มีค่าน้อยที่สุด หาสมการปกติได้ดังนี้

ให้ $S = \sum e_i^2$

$= \sum(Y_i - b_o - \sum_{j=1}^{k} b_j X_{ij})^2$

ทำให้ S มีค่าน้อยที่สุดโดยการอนุพันธ์ย่อยเทียบกับ b0,b1,b2,...,bk แล้วให้เท่ากับ 0

$$\frac{\partial s}{\partial b_o} = -2\sum_{i=1}^{n}(Y_i - b_0 - \sum_{j=1}^{k} b_j X_{ij}) = 0$$

และ

$$\frac{\partial s}{\partial b_j} = -2\sum_{i=1}^{n}(Y_i - b_0 - \sum_{j=1}^{k} b_j X_{ij}) = 0\ ,\ j = 1,2,\dots,k$$

จัดสมการใหม่ จะได้สมการปกติ ดังนี้

$$nb_0 + b_1\sum_{i=1}^{n} X_{i1} + b_2\sum_{i=1}^{n} X_{i2} + \dots + b_k\sum_{i=1}^{n} X_{ik} = \sum_{i=1}^{n} Y_i$$

$$b_0\sum_{i=1}^{n} X_{i1} + b_1\sum_{i=1}^{n} X_{i1}^2 + b_2\sum_{i=1}^{n} X_{i1}X_{i2} + \dots + b_k\sum_{i=1}^{n} X_{i1}X_{ik} = \sum_{i=1}^{n} X_{i1}Y_i$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$b_0\sum_{i=1}^{n} X_{i1} + b_1\sum_{i=1}^{n} X_{ik}X_{i1} + b_2\sum_{i=1}^{n} X_{ik}X_{i2} + \dots + b_k\sum_{i=1}^{n} X_{ik}^2 = \sum_{i=1}^{n} X_{ik}Y_i$$

ทั้งหมดมี k+1 สมการ และในกรณีสมการถดถอยมีตัวแปรอิสระ2ตัว สมการปกติจะมี 3 สมการดังนี้

$$nb_0 + b_1\sum X_1 + b_2\sum X_2 = \sum Y$$

$$b_0\sum X_1 + b_1\sum X_1^2 + b_2\sum X_1X_2 = \sum X_1Y$$

$$b_0\sum X_2 + b_1\sum X_1X_2 + b_2\sum X_2^2 = \sum X_2Y$$

สัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple Determination)

การแบ่งส่วนผลบวกกำลังสองในการวิเคราะห์การถดถอยพหุคูณ กล่าวคือ ผลบวกทั้งหมดของความเบี่ยงเบนกำลังสองของตัวแปร Y จากค่าเฉลี่ย $\bar{Y}$ หรือ total sum of squared deviation เขียนแทนด้วย SST ถูกแบ่งเป็น 2 ส่วน คือ sum of square due to regression เขียนแทนด้วย SSR และ sum of squares of residuals หรือ deviation due to error เขียนแทนด้วย SSE นั่นคือ

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma(\hat{Y}_i - \bar{Y})^2 + \Sigma(Y_i - \hat{Y}_i)^2$$

SST = SSR + SSE

SST มีองศาแห่งความเป็นอิสระเท่ากับ n-1 SSR มีองศาแห่งความเป็นอิสระเท่ากับ k SSE มีองศาแห่งความเป็นอิสระเท่ากับ n-k-1 และส่วนแบ่งขององศาแห่งความเป็นอิสระ ดังนี้

$$(n - 1) = (k) + (n - k - 1)$$

เช่นเดียวกับกรณีการถดถอยเชิงเส้นอย่างง่าย ค่าสัมประสิทธิ์การตัดสินใจของการถดถอยพหุคูณ เท่ากับอัตราส่วนระหว่าง SSR และ SST นั่นคือ

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

การคำนวณค่า R^2 จากการแบ่งส่วนผลบวกกำลังสอง

SST = SSR +SSE

หรือ

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma(\hat{Y}_i - \bar{Y})^2 + \Sigma(Y_i - \hat{Y}_i)^2$$

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma Y^2 - \frac{(\Sigma Y^2)}{n} = Y'Y - \frac{(\Sigma Y^2)}{n} = SST$$

$$SSE = Y'Y - b'X'Y$$

$$SSR = \Sigma(Y_i - \bar{Y})^2 - (\hat{Y}_i - \bar{Y})^2$$

$$= Y'Y - \frac{(\Sigma Y^2)}{n} - Y'Y - b'X'Y$$

$$= b'X'Y - \frac{(\Sigma Y^2)}{n}$$

$$R^2 = \frac{b'X'Y - \frac{(\Sigma Y^2)}{n}}{Y'Y - \frac{(\Sigma Y^2)}{n}}$$

เมื่อ $b'X'Y = b_0\Sigma Y + b_1\Sigma X_1 Y + b_2\Sigma X_2 Y + \ldots + b_k\Sigma X_k Y$ สัมประสิทธิ์การตัดสินใจพหุคูณ R^2 คือสัดส่วนความแปรผันทั้งหมดของ Y ที่อธิบายโดยความแปรผันในตัวแปร

อิสระ$X_1, X_2, \ldots, X_k$ หรือสัดส่วนของความแปรผันทั้งหมดของ Y ที่อธิบายโดยความสัมพันธ์ระหว่าง Y และ $X_1, X_2, \ldots, X_k$ ความหมายของ R^2 ในการถดถอยอย่างง่ายจะสมมูล R^2

ในการถดถอยพหุคูณค่า R^2 อยู่ระหว่าง 0 และ 1 ถ้าไม่มีความสัมพันธ์กันระหว่างตัวแปรตาม Y และตัวแปรอิสระที่ใช้ในสมการถดถอยแล้ว $R^2 = 0$ ถ้าสมการถดถอยที่ประมาณได้เหมาะสมกับข้อมูลอย่างสมบูรณ์ และสามารถใช้ประมาณหรือพยากรณ์ค่า Y ได้ถูกต้องโดยแท้แล้ว $R^2 = 1$

โดยทั่วไปค่า R^2 จะสูงขึ้น เมื่อมีตัวแปรอิสระเพิ่มขึ้นเข้ามาในสมการถดถอยพหุคูณ เหตุผลคือเมื่อเพิ่มตัวแปรอิสระในสมการถดถอย ค่า SST ไม่เปลี่ยนแปลง แต่โดยทั่วไป SSR จะสูง และ SSE จะต่ำลงดังนั้น R^2 จึงมีค่าสูงขึ้น การเพิ่มตัวแปรอิสระอาจไม่ช่วยให้มีนัยสำคัญที่จะอธิบายตัวแปรตาม Y การเพิ่มตัวแปรอิสระในสมการถดถอยสำหรับจุดประสงค์เพื่อให้ R^2 มีค่าสูงขึ้น มักมีผลทำให้มีตัวแปรอิสระมีจำนวนมากเกินไปในสมการการถดถอย และอาจทำให้สมการถดถอยมีรูปแบบเลวลงแทนที่จะดีขึ้น ดังนั้นจึงอาจพบว่าการเลือกรูปแบบที่มีค่า R^2 ต่ำกว่าเล็กน้อยอาจได้รูปแบบที่ดีกว่า

เพื่อป้องกันปัญหานี้ เราสามารถใช้ Adjusted R^2 เป็นค่าวัดค่ารูปแบบเหมาะสมกับข้อมูลอย่างไร โดยการนำองศาแห่งความเป็นอิสระมาพิจารณาด้วย คำนวณจากสูตร

$$\text{Adjusted } R^2 = 1 - \frac{(SSE)/(n-k-1)}{(SST)/(n-1)}$$

เมื่อ n ใหญ่ ค่า R^2 จะใกล้เคียงกับค่า Adjusted R^2 ค่า Adjusted R^2 อาจลดลงเมื่อเพิ่มตัวแปรอิสระในสมการถดถอย ทั้งนี้เพราะแม้ว่า SSE จะลดลงเสมอ แต่การลดลงอาจชดเชยด้วยองศาแห่งความเป็นอิสระ (n – k - 1)

หมายเหตุ สัมประสิทธิ์การตัดสินใจพหุคูณ อาจเขียนแทนด้วย

$$r_y^2(X_1, \ldots, X_K) = \frac{SST - SSE(X_1, \ldots, X_k)}{SST}$$

และสัมประสิทธิ์การตัดสินใจพหุคูณที่ปรับ อาจเขียนแทนด้วย

$$\text{Adj } r_y^2(X_1, \ldots, X_K) = \frac{MST - MSE(X_1, \ldots, X_k)}{MST}$$

สัมประสิทธิ์สหสัมพันธ์พหุคูณ (Coefficient of Multiple Correlation)

สัมประสิทธิ์สหสัมพันธ์พหุคูณ เขียนแทนด้วย R คือ ค่าวัดระดับความสัมพันธ์ระหว่างตัวแปรตาม Y และตัวแปรอิระทั้งหมดรวมกัน มีค่าเท่ากับรากที่สองของสัมประสิทธิ์การตัดสินใจพหุคูณ

$$R = \sqrt{R^2}$$

R มีค่าเป็นบวกเสมอ คือ $0 \leq R \leq 1$ ถ้า $R = 1$ หมายถึงมีความสัมพันธ์อย่างสมบูรณ์ และ 0 หมายถึงไม่มีความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระทั้งหมด

ตัวอย่างกรณีศึกษา ภูเขาไฟ Mount Pinatubo ในประเทศฟิลิปปินส์ได้ปะทุขึ้นในวันที่ 15 มิถุนายน1991 ทำให้มีพื้นที่เสียหายมากกว่า 5,000,000 ไร่ จำนวนผู้บาดเจ็บและเสียชีวิตมากกว่า 700 คน และได้ทำลายบ้านเรือนไปมากกว่า 200,000 หลัง โดยรัฐบาลได้มีนโยบายที่จะป้องกันภัยธรรมชาติของการไหลของโคลนร้อน (Lahar Flow) โดยการตั้งถิ่นที่อยู่อาศัยใหม่ แต่ในทางกลับกันวิศวกรเชื่อว่าสามารถควบคุมการไหลของโคลนร้อน (Lahar Flow) โดยทำการก่อสร้างโครงการก่อสร้างคล้ายกับเขื่อน โดยรัฐบาลได้จัดสรรงบประมาณทั้งหมด 370 ล้านเหรียญ สำหรับเหตุการณ์ภูเขาไฟ Mount Pinatubo ระเบิด จากเงินงบประมาณทั้งหมดประมาณ 92 ล้านเหรียญใช้สำหรับย้ายถิ่นที่อยู่อาศัยใหม่ และอีกประมาณ 156 ล้านเหรียญใช้สำหรับสร้างโครงสร้างพื้นฐานสำหรับควบคุม Lahar Flow โดยความเสียหายต่อทรัพย์สินมีมูลค่าประมาณ 463 ล้านเหรียญ และมีผู้บาดเจ็บและผู้เสียชีวิตประมาณ 700 คน ซึ่งจากเงินลงทุนสามารถบรรเทาทรัพย์สินที่เสียหายและจำนวนผู้เสียชีวิตได้ประมาณ 250 ล้านเหรียญ และ 5,000 คนตามลำดับ การจัดสรรงบประมาณเพื่อลดความเสี่ยงต่างๆตามนโยบายของรัฐบาลนั้น เป็นปัญหาที่ต้องใช้การตัดสินใจหลายตัวแปรเพราะว่าหากรัฐบาลมีเงินลงทุนที่จำกัด ก็จะทำให้ไม่สามารถตัดสินใจในการลงทุนนั้นได้ ดังนั้นจึงได้ใช้ วิธีประเมินการตัดสินใจบนพื้นฐานพหุปัจจัย เข้ามาช่วยในการตัดสินใจในการลงทุนในการป้องกันภัยธรรมชาติของรัฐบาล โดยให้ความสัมพันธ์ระหว่างจำนวนเงินลงทุนในการตั้งถิ่นที่อยู่อาศัยใหม่กับจำนวนเงินลงทุนในการสร้างโครงสร้างพื้นฐาน โดยแปรผันตรงกับจำนวนเสียชีวิต และมูลค่าทรัพย์สินที่เสียหาย โดยได้สมการความสัมพันธ์ดังสมการตัวอย่างต่อไปนี้

$f_1(x_1, x_2) =$ 170 - 89.5 x1 + 0.588 x21 + 120 x2 – 0.442 x22

$f_2(x_1, x_2) =$ -1.3 – 0.91 x1 + 0.0134 x21 + 1.24 x2 + 0.00118 x22

โดยที่

x_1 = จำนวนเงินลงทุนที่ใช้ในโครงการตั้งถิ่นที่อยู่อาศัยใหม่

x_2 = จำนวนเงินลงทุนที่ใช้สร้างโครงสร้างพื้นฐานเพื่อใช้ในการควบคุมโคลนที่ไหลมาจากภูเขาไฟระเบิด (Lahar)

f_1 = จำนวนผู้รอดชีวิตที่สามารถช่วยได้ โดยขึ้นอยู่กับ x_1 และ x_2

f_2 = มูลค่าทรัพย์สินที่ลดความเสียหายได้ โดยขึ้นอยู่กับ x_1 และ x_2

ตาราง 1 แสดงสถานการณ์การลงทุนที่ต่างกัน สำหรับ x_1 และ x_2 ซึ่งมีผลกระทบต่อ f_1 และ f_2

x_1 (ล้าน $)	x_2 (ล้าน $)	f_1 (จำนวนคน)	f_2 (ล้าน $)
0	0	200	0.05
50	50	2000	50
50	75	3500	80
75	100	4500	150
75	150	4800	220
100	150	5000	250

งานวิจัยที่เกี่ยวข้อง

นันทวัฒน์ วีระยุทธ และณชล ไชยรัตนะ (2546) ได้ทำการศึกษาขั้นตอนวิธีเชิงพันธุกรรมแบบหลายเป็นหมายใช้ ในการแก้ปัญหาค่าเหมาะสมที่สุด ซึ่งปัญหานี้จะเกี่ยวข้องกับการเลือกค่าขอบเขตของแรงบิดใช้งานที่เหมาะสมของแอคชูเอเตอร์และวางแผนเลือกเส้นทางในการเคลื่อนที่ของ End-effector ของแขนกล เพื่อใช้เป็นอินพุทให้กับขั้นตอนวิธีควบคุมการเคลื่อนที่โดยใช้เวลาน้อยที่สุด และอยู่ภายใต้เงื่อนไขบังคับของเวลาที่ใช้ในการเคลื่อนที่และค่าคลาดเคลื่อนของการติดตาม โดยใช้รูปแบบการเข้ารหัสโครโมโซมที่แตกต่างกัน 3 รูปแบบคือการเข้ารหัสแบบเลขฐานสอง การเข้ารหัสแบบเลขจำนวนเต็ม และการเข้ารหัสแบบเลขจำนวนจริง นอกจากนี้ได้

ป
GD
5008
.T5
ฉบาก
2553
153453138
24 มิ.ย. 2554
สำนักหอสมุด

ศึกษาถึงการใช้การควบคุมความแตกต่างในปัญหาค่าเหมาะสมที่สุดแบบหลายเป้าหมายกับโครโมโซมที่มีการเข้ารหัสที่แตกต่างกัน ซึ่งผลลัพท์ที่ได้คือโครโมโซมที่เข้ารหัสแบบเลขจำนวนเต็มมีความเหมาะสมมากกว่าโครโมโซมรูปแบบอื่น

ธนัดชัย กุลวรวานิชพงษ์ และบัณฑิต เอื้ออาภรณ์ (2542) ได้ทำการศึกษาวิธีการแก้ปัญหา ออปติไมเซชันที่มีการพิจารณาฟังก์ชันวัตถุประสงค์มากกว่าหนึ่งฟังก์ชัน มาใช้แก้ปัญหาออปติมัลเพาเวอร์โฟลว์ ที่มีการพิจารณา ต้นทุนการผลิตโดยรวมของระบบ ผลกระทบทางด้านสิ่งแวดล้อม และความมั่นคงในการส่งจ่ายกำลังไฟฟ้า ของระบบเป็นฟังก์ชันวัตถุประสงค์ วิธีการดังกล่าวนี้ได้ทดสอบกับระบบทดสอบ [1] จากนั้นได้นำผลที่ได้เปรียบเทียบ กับการแก้ปัญหาออปติมัลเพาเวอร์โฟลว์แบบดั้งเดิม ที่ใช้ฟังก์ชันวัตถุประสงค์เพียงหนึ่งฟังก์ชัน พบว่าการใช้ฟังก์ชันวัตถุประสงค์เพียงหนึ่งฟังก์ชันนั้น จุดทำงานที่ได้อาจจะทำให้ ฟังก์ชันวัตถุประสงค์ตัวอื่นมีค่าสูงขึ้นในบางกรณี อาจจะมีค่าสูงจนทำให้เกิดความเสียหายต่อระบบได้ จากผลการทดสอบได้แสดงให้เห็นถึงผลกระทบ ที่อาจเกิดขึ้นเมื่อเลือกจุดการทำงาน โดยใช้ฟังก์ชันวัตถุประสงค์ เป็นฟังก์ชันดังที่ได้กล่าวมาแล้วข้างต้นแยกกันครั้งละหนึ่งฟังก์ชัน เปรียบเทียบกับการพิจารณา ฟังก์ชันวัตถุประสงค์ทั้งสามพร้อมกัน พบว่าเมื่อระบบที่พิจารณามีความคลาดเคลื่อน อันเนื่องมาจากการพยากรณ์โหลด จะทำให้เกิดผลกระทบต่อการไหลของกำลังไฟฟ้าในระบบ ซึ่งหากเลือกจุดการทำงาน โดยคำนึงถึงผลของฟังก์ชันต้นทุนการผลิตเพียงอย่างเดียว ก็อาจส่งผลให้สายส่งบางเส้นเกิดโอเวอร์โหลดได้ทำให้ระบบมีความเสี่ยงต่อความมั่นคงในการส่งกำลังไฟฟ้าผ่านสายส่ง

ชลทิศ กิตติคุณ ได้ทำการศึกษาสาเหตุและสภาพการเกิดน้ำท่วมในลุ่มน้ำคลองตะกั่วป่า จังหวัดพังงาเพื่อจัดทำฐานข้อมูลใช้ประกอบเตือนภัยน้ำท่วมสำหรับพื้นที่ชุมชนการศึกษาประกอบด้วยการวิเคราะห์ลักษณะคุณสมบัติของฝน กราฟน้ำหลากจากพื้นที่ลุ่มน้ำย่อยระดับน้ำทะเล ซึ่งเป็นปัจจัยสำคัญที่ทำให้เกิดน้ำท่วมเพื่อนำมากำหนดสถานการณ์สำหรับจำลองสภาพน้ำท่วมจากนั้นจึงทำการจำลองสภาพการเคลื่อนตัวของน้ำหลากในคลองตะกั่วป่าที่เกิดขึ้นในสถานการณ์ต่าง ๆ โดยใช้แบบจำลอง HEC-RAS โดยใช้ข้อมูลระดับน้ำและอัตราการไหลในช่วงน้ำท่วมของวันที่ 21 สิงหาคม 2550 และในช่วงปริมาณน้ำปกติระหว่างเดือนพฤษภาคม 2548 ค่าระดับน้ำที่คำนวณได้จากแบบจำลอง HEC-RAS นำไปใช้จัดทำแผนที่น้ำท่วมโดยใช้โปรแกรม HEC-GEORAS EXTENTION ผลการศึกษาพบว่าพื้นที่ชุมชนในเขตเทศบาลเมืองตะกั่วป่าจะประสบปัญหาน้ำท่วมเป็นบริเวณกว้างเมื่อปริมาณน้ำหลากในคลองตะกั่วป่ามีค่าเกินกว่า 550.0 ลูกบาศก์เมตรต่อวินาที ซึ่งส่วนใหญ่จะเกิดขึ้นเมื่อมีปริมาณฝนตกในพื้นที่ลุ่มน้ำตอนบนเกินกว่า 110.0 มิลลิเมตรใน 1 วัน ระดับน้ำท่วมสูงสุดในเขตพื้นที่ชุมชนมีความสัมพันธ์ที่ดีกับระดับน้ำสูงสุดที่สถานีวัดน้ำท่า X.188 ซึ่งอยู่เหนือน้ำจากพื้นที่ชุมชนประมาณ 24 กิโลเมตร และการเคลื่อนตัวของคลื่นน้ำหลากจากสถานี X.188 มาถึงพื้นที่ชุมชนใช้เวลาประมาณ 13 ชั่วโมง ถึง 30 ชั่วโมง