การประยุกต์การวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณแบบภาวะน่าจะเป็นสูงสุดและแบบแกร่งของเมทริกซ์ความแปรปรวนร่วมในการจำแนกกลุ่มของดิน ตามองค์ประกอบทางเคมี : กรณีศึกษาภูมิภาคตะวันตกของประเทศไทย

โดย

นายวัชรพงษ์ ว่องนิยมเกษตร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติประยุกต์

ภาควิชาสถิติ

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2552

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การประยุกต์การวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณแบบภาวะน่าจะเป็นสูงสุดและแบบแกร่งของเมทริกซ์ความแปรปรวนร่วมในการจำแนกกลุ่มของดินตามองค์ประกอบทางเคมี: กรณีศึกษาภูมิภาคตะวันตกของประเทศไทย

โดย
นายวัชรพงษ์ ว่องนิยมเกษตร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2552
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

APPLICATIONS OF PCA BASED ON MAXIMUM LIKELIHOOD AND ROBUST ESTIMATIONS OF COVARIANCE MATRIX TO SOIL CHEMICAL COMPOSITION FOR CLUSTERING PATTERNS: CASE STUDY FROM THE WESTERN REGION OF THAILAND

BY

Watcharaphong Wongniyomkaset

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Statistics

Graduate School

SILPAKORN UNIVERSITY

2009

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “การประยุกต์การวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณแบบภาวะน่าจะเป็นสูงสุดและแบบแกร่งของเมทริกซ์แปรปรวนร่วมในการจำแนกกลุ่มของดินตามองค์ประกอบทางเคมี: กรณีศึกษาภูมิภาคตะวันตกของประเทศไทย” เสนอโดย นายวัชรพงษ์ ว่องนิยมเกษตร เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาสถิติประยุกต์

.....

(รองศาสตราจารย์ ดร.ศิริชัย ชินะตังกูร)

คณบดีบัณฑิตวิทยาลัย

วันที่.....เดือน.....พ.ศ. .....

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ผู้ช่วยศาสตราจารย์ ดร.กมลชนก พานิชการ

คณะกรรมการตรวจสอบวิทยานิพนธ์

.....ประธานกรรมการ

(รองศาสตราจารย์ไพบูลย์ รัตนประเสริฐ)

...../ ..... / .....

.....กรรมการ	.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.บุญอ้อม โฉมที)	(รองศาสตราจารย์วีรานันท์ พงศาภักดี)
...../ ..... /	...../ ..... / .....
.....กรรมการ	.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.กุศยา ปลั่งพงษ์พันธ์)	(ผู้ช่วยศาสตราจารย์ ดร.กมลชนก พานิชการ)
...../ ..... /	...../ ..... / .....

47304201: สาขาวิชาสถิติประยุกต์
คำสำคัญ : การวิเคราะห์องค์ประกอบหลักแบบแกร่ง/ตัวประมาณค่าแบบเครื่องหมาย/ตัวประมาณค่าแบบอันดับ/องค์ประกอบทางเคมี

วัชรพงษ์ ว่องนิยมเกษตร: การประยุกต์การวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณแบบภาวะน่าจะเป็นสูงสุดและแบบแกร่งของเมทริกซ์ความแปรปรวนร่วมในการจำแนกกลุ่มของดินตามองค์ประกอบทางเคมี: กรณีศึกษาภูมิภาคตะวันตกของประเทศไทย. อาจารย์ที่ปรึกษาวิทยานิพนธ์: ผศ.ดร. กมลชนก พานิชการ. 97 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาการวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณเมทริกซ์ความแปรปรวนร่วม 3 วิธีคือการประมาณแบบภาวะน่าจะเป็นสูงสุด การประมาณแบบแกร่งด้วยตัวประมาณค่าแบบเครื่องหมายและการประมาณแบบแกร่งด้วยตัวประมาณค่าแบบอันดับ บนข้อมูล 2 ลักษณะคือ ข้อมูลดิบและข้อมูลแปลง โดยใช้กรณีศึกษาจากตัวอย่างดิน 58 ตัวอย่างที่ถูกเก็บมาจาก 3 จังหวัดได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม ซึ่งถูกวัดโดยใช้เทคนิคทางกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope, SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์เรย์ที่วัดได้จากตัวอย่างวัตถุด้วยหัวตรวจวัด Energy Dispersive X-ray Micro-analysis (EDX) ประกอบด้วยตัวแปร 10 ตัวคือออกซิเจน แมกนีเซียม อลูมิเนียม ซิลิคอน โพแทสเซียม แคลเซียม เหล็ก ไทเทเนียม คาร์บอน และโซเดียม

ผลการวิจัยพบว่า

(1) ปริมาณองค์ประกอบทางเคมีที่พบในตัวอย่างของดินคือ O > Si > Al > Fe > C > K > Mg > Ca > Ti > Na

(2) การวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณแบบภาวะน่าจะเป็นสูงสุดสามารถอธิบายความแปรปรวนของข้อมูลได้มากที่สุดทั้งบนข้อมูลดิบและข้อมูลแปลง

(3) การแปลงข้อมูลด้วยวิธีของ Aitchison (1986) สามารถทำให้เห็นการจำแนกกลุ่มของข้อมูลดินค่อนข้างชัดเจน แต่ลักษณะการแบ่งกลุ่มไม่สอดคล้องกับชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาคือดินเหนียว ดินร่วนเหนียว ดินร่วน ดินร่วนปนทรายแป้ง และดินร่วนปนทราย

ดังนั้นจึงอาจพิจารณาปัจจัยทางสิ่งแวดล้อมอื่น เช่น ลักษณะทางกายภาพของดิน (สี ค่า pH) ปริมาณโลหะหนัก หรือปริมาณจุลธาตุอาหารในดิน

ลายมือชื่อนักศึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์.....

47304204: MAJOR: APPLIED STATISTICS
KEY WORDS: ROBUST PRINCIPAL COMPONENT ANALYSIS/THE SIGN ESTIMATOR/THE RANK ESTIMATOR/THE CHEMICAL COMPOSITION

WATCHARAPHONG WONGNIYOMKASET: APPLICATIONS OF PCA BASED ON MAXIMUM LIKELIHOOD AND ROBUST ESTIMATIONS OF COVARIANCE MATRIX TO SOIL CHEMICAL COMPOSITION FOR CLUSTERING PATTERNS: CASE STUDY FROM THE WESTERN REGION OF THAILAND. THESIS ADVISORS: ASST. PROF. KAMOLCHANOK PANISHKAN, Ph.D. 97 pp.

The purpose of this research is to study principal component analysis (PCA) based on three estimations of covariance matrix; maximum likelihood estimation, robust estimation using the rank estimator and robust estimation using the sign estimator. The analyses were applied to the fifty eight raw and transformed data collected from three agricultural provinces in the western region of Thailand which are Nakhon Pathom, Samut Sakhon and Samut Songkham. Soil chemical compositions were measured by Scanning Electron Microscope (SEM) and Energy Dispersive X-ray Microanalysis (EDX). There are ten component variables; Oxygen, Magnesium, Aluminum, Silicon, Potassium, Calcium, Iron, Titanium, Carbon and Sodium.

The results of the study are as follows.

(1) The basic result for soil indicate decreased amount of O, Si, Al, Fe, C, K, Mg, Ca, Ti and Na.

(2) Principal component analysis based on maximum likelihood estimation can explain the most percentage of the total variation of both raw and transformed data.

(3) The transformed data with the method of Aitchison (1986) indicated the apparent classification of the soil samples. However, the classification is not consistent to the five textures of soil; clay, clay loam, medium loam, silty loam and sand loam.

So the other environmental factors such as the physical characteristics of soil (color, pH), the amount of metal and micronutrient can be considered.

Department of Statistics Graduate School , Silpakorn University Academic Year 2009
Student's signature .....
Thesis Advisor's signature .....

กิตติกรรมประกาศ

ในการศึกษาระดับบัณฑิตศึกษา สิ่งสำคัญประกาศหนึ่งคือ การเขียนและเรียบเรียงวิทยานิพนธ์ ซึ่งเป็นงานวิจัยที่ต้องอาศัยความวิริยะและความพากเพียร ผู้วิจัยผ่านความยากลำบากด้วยความกรุณาของคณาจารย์ผู้ประสาทความรู้ ในโอกาสนี้ผู้วิจัยกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.กมลชนก พานิชการ ผู้ที่เสียสละเวลาทุกเมื่อเพื่อให้งานวิจัยเสร็จสมบูรณ์ คอยให้ความช่วยเหลือในการเรียบเรียงวิทยานิพนธ์ แก้ปัญหาที่เกิดขึ้นในระยะเวลาที่ผ่านมา และตรวจสอบแก้ไขด้วยดีมาตลอดอย่างสุดซึ้ง ผู้ช่วยศาสตราจารย์ ดร.ปราณี นิลกรณ์ สำหรับข้อแนะนำในการเขียนโปรแกรม SAS สำหรับวิเคราะห์ข้อมูลอันเป็นประโยชน์ยิ่งต่อผู้วิจัย และผู้ช่วยศาสตราจารย์ ดร.มยุวา อารีกิจเสรี สำหรับข้อมูลที่ใช้ในงานวิจัยและความรู้ในการวิเคราะห์องค์ประกอบทางเคมีของดินที่เป็นส่วนสำคัญยิ่ง

กราบขอบพระคุณ รองศาสตราจารย์ไพบูลย์ รัตนประเสริฐ ประธานกรรมการตรวจสอบวิทยานิพนธ์ รองศาสตราจารย์วีรานันท์ พงศาภักดี ผู้ช่วยศาสตราจารย์ ดร.กุศยา ปลั่งพงษ์พันธ์ และผู้ช่วยศาสตราจารย์ ดร.บุญอ้อม โฉมที กรรมการผู้ทรงคุณวุฒิตรวจสอบวิทยานิพนธ์ที่คอยให้คำแนะนำที่เป็นประโยชน์ยิ่งต่อผู้วิจัย ตลอดจนประสบการณ์ในการเป็นผู้ช่วยสอนการใช้โปรแกรม Minitab สำหรับนักศึกษาระดับปริญญาตรี

กราบขอบพระคุณอาจารย์ทุกท่านในภาควิชาสถิติที่คอยประสิทธิ์ประสาทวิชาความรู้ อีกทั้งกำลังใจและคำแนะนำในการดำเนินชีวิตต่อไปในอนาคต

สุดท้ายขอกราบขอบพระคุณ คุณแม่อันเป็นที่รักยิ่ง และพี่ชายกับน้องสาวในความรัก ความอบอุ่น ความช่วยเหลือที่มีให้ตลอดมา อีกทั้งยังเป็นกำลังใจและสนับสนุนอย่างดีมาโดยตลอด

ผู้วิจัยขออำนาจคุณพระศรีรัตนตรัย ดลบันดาลให้ทุกท่านจงมีแต่ความสุขความเจริญ ประสบแต่ความสำเร็จ ไร้โรคภัยทั้งในโลกนี้และโลกหน้าชั่วกาลนานเทอญ

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญตาราง ..... ฌ
สารบัญภาพ ..... ฎ
บทที่
1 บทนำ 1
ความเป็นมาและความสำคัญของปัญหา 1
วัตถุประสงค์ของการวิจัย 4
ขอบเขตของการวิจัย 4
ประโยชน์ของการวิจัย 5
นิยามศัพท์เฉพาะ 5
2 ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง 7
ทฤษฎีของดิน 7
แหล่งที่มาของตัวอย่างดิน 16
การวิเคราะห์องค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุด 23
การวิเคราะห์องค์ประกอบหลักโดยใช้ตัวประมาณค่าแบบแกร่ง 28
วรรณกรรมที่เกี่ยวข้อง 31
3 วิธีดำเนินการวิจัย 36
4 ผลการวิจัย 38
ผลการตรวจสอบข้อมูลเบื้องต้นและสหสัมพันธ์ขององค์ประกอบทางเคมี
ของดิน 38
ผลการวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณ
แบบภาวะน่าจะเป็นสูงสุด 44
ผลการวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณ
แบบเครื่องหมาย 56

บทที่ หน้า
ผลการวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ 67
5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ 79
สรุปผลการวิจัย 79
อภิปรายผล 86
ข้อเสนอแนะจากงานวิจัย 88

บรรณานุกรม 90

ภาคผนวก 92

ประวัติผู้วิจัย 97

สารบัญตาราง

ตารางที่		หน้า
1	ธาตุอาหารที่จำเป็นต่อพืชชั้นสูงทั่วไป และความเข้มข้นในเนื้อเยื่อพืชซึ่งจัดว่าเพียงพอ	13
2	หน้าที่สำคัญของธาตุอาหารพืช และอาการขาดแคลนธาตุอาหารพืช	14
3	ค่าสถิติเบื้องต้นขององค์ประกอบทางเคมีของดิน	39
4	ค่าสถิติเบื้องต้นขององค์ประกอบทางเคมีของดินโดยแยกตามชนิดของดิน	40
5	ค่าสถิติสำหรับการทดสอบการแจกแจงปกติ	41
6	เมตริกซ์สหสัมพันธ์ของเพียร์สันขององค์ประกอบทางเคมีของดิน	42
7	การจำแนกตัวอย่างของดินจำนวน 58 ตัวอย่างตามชนิดของดิน	44
8	ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ	45
9	ค่าสัมประสิทธิ์ขององค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ	46
10	ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ	48
11	ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง	51
12	ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง	52
13	ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง	53
14	ความแปรปรวน และความแปรปรวนร่วมของธาตุ ๆ ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ	56
15	ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ	57
16	ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ	59
17	ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง	62

ตารางที่ หน้า

18 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง 63
19 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง 64
20 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ 68
21 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ 68
22 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ 70
23 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง 73
24 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง 74
25 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง 75
26 ตัวแปรที่มีค่าสัมประสิทธิ์สูงในแต่ละองค์ประกอบหลักที่ 1-3 โดยจำแนกตามตัวประมาณค่าเมทริกซ์ความแปรปรวนร่วมบนข้อมูลดิบ และข้อมูลแปลงด้วยวิธีของ Aitchison 86
27 ข้อมูลองค์ประกอบทางเคมีของดินขนาด 65 ไมครอนจำนวน 58 ตัวอย่าง 93

สารบัญภาพ

ภาพที่ หน้า

1 ส่วนประกอบของดินในปริมาตรที่เหมาะสมต่อการเพาะปลูก 8
2 ภาพขั้นตอนการอัดผงดิน 17
3 ภาพสะท้อนของอิเล็กตรอนทุติยภูมิ และอะตอมของธาตุตัวอย่างที่ปลดปล่อยพลังงาน 18
4 ภาพจาก SEM ใช้หัวตรวจวัด EDX ทำการศึกษาด้วยโปรแกรม Smiling programs 20
5 ตัวอย่างของผลที่ได้จากการวิเคราะห์ด้วย SEM/EDX 21
6 Boxplot ของตัวแปรที่เป็นองค์ประกอบทางเคมีของดิน 43
7 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ 47
8 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ 49
9 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ 50
10 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดค่าจากข้อมูลแปลง 53
11 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง 54
12 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง 55
13 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ 58
14 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ 60
15 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ 61
16 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง 64

ภาพที่ หน้า

17 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบเครื่องหมาย จากข้อมูลแปลง 65

18 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบเครื่องหมาย จากข้อมูลแปลง 67

19 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ 70

20 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบอันดับจาก ข้อมูลดิบ 71

21 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบอันดับจาก ข้อมูลดิบ 72

22 Scree plot ของค่าไอเกนองค์ประกอบหลักด้วยค่าประมาณแบบอันดับจาก ข้อมูลแปลง 75

23 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบอันดับจาก ข้อมูลแปลง 76

24 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบอันดับจาก ข้อมูลแปลง 77

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันมีการประยุกต์การวิเคราะห์ตัวแปรพหุโดยใช้เทคนิคการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis, PCA) กันอย่างกว้างขวางในหลายศาสตร์ รวมทั้งในเรื่องของดินและการปรับปรุงคุณภาพดิน โดยใช้กรณีศึกษาในหลาย ๆ ประเทศ (กมลชนก พานิชการ 2550) ประเทศไทยของเรานั้นเป็นประเทศเกษตรกรรมต้องอาศัยดินที่มีคุณภาพในการเพาะปลูก การศึกษาคุณสมบัติด้านต่าง ๆ ของดินเพื่อทำการปรับปรุงให้มีคุณภาพเหมาะสมจึงมีความสำคัญเป็นอย่างยิ่ง รวมถึงการศึกษาเข้าลึกถึงความสัมพันธ์ทั้งทางด้านกายภาพและเคมีก็เป็นสิ่งที่ควรศึกษาเพิ่มเติม

การทราบลักษณะและองค์ประกอบทางเคมีของดินช่วยทำให้เราทราบว่าตัวอย่างของดินที่นำมาทำการศึกษานี้ประกอบด้วยแร่ธาตุอาหารใดบ้างและมีความเพียงพอเหมาะสมที่จะทำการเกษตร เพื่อให้เกิดผลผลิตที่คุ้มค่าหรือไม่ ในการศึกษาเทคนิคการวิเคราะห์องค์ประกอบหลัก PCA เป็นเทคนิคที่ใช้อธิบายโครงสร้างของความแปรปรวนและความแปรปรวนร่วมโดยใช้ผลบวกเชิงเส้นของตัวแปรเดิมโดยมีวัตถุประสงค์เพื่อลดข้อมูล (มิติ) (Data reduction) และเพื่อให้ตีความข้อมูลง่ายขึ้นโดยสูญเสียสารสนเทศที่มีอยู่ในข้อมูลน้อยที่สุด ความแปรผันรวม (Total variability) ที่เกิดจากตัวแปรทุกตัวที่ศึกษานั้นอาจเกิดขึ้นจากองค์ประกอบหลักๆ ไม่กี่องค์ประกอบ จึงอาจอธิบายความแปรผันรวมของตัวแปรทั้งหมดด้วยองค์ประกอบเพียงไม่กี่องค์ประกอบดังกล่าว แต่ในการวิเคราะห์องค์ประกอบเหล่านั้น ถ้าข้อมูลตัวอย่างที่ทำการศึกษามีค่านอกกลุ่มปะปนอยู่จะส่งผลต่อเมทริกซ์ความแปรปรวนร่วมซึ่งจะมีความไวต่อค่านอกกลุ่ม และอาจทำให้ค่าประมาณองค์ประกอบหลักที่ได้จากการคำนวณเมทริกซ์ความแปรปรวนร่วมดังกล่าวด้วยวิธีการวิเคราะห์ PCA อธิบายความผันแปรของข้อมูลได้ไม่ดีเท่าที่ควร สำหรับในกรณีค่าตอบสนองที่ถูกวัดขึ้นมาจากตัวแปรที่ถูกนำมาศึกษาเป็นค่าเปอร์เซ็นต์ ซึ่งเป็นข้อมูลในลักษณะ Compositional Data มีลักษณะไม่เหมาะกับการวิเคราะห์โดยตรงบนข้อมูลดิบ เนื่องจากเงื่อนไขว่าผลรวมของตัวแปรมีค่าเป็น 100% Aitchison (1986) ได้เสนอการแปลงข้อมูลเพื่อให้มีการแจกแจงใกล้เคียงการแจกแจงปกติ ดังนี้

$$y_{ij} = \ln(x_{ij}) - p^{-1}\sum_{i=1}^{p}\ln(x_{ij}) \qquad (1)$$

โดย y_{ij} คือค่าที่ได้จากการแปลงค่าสังเกต x_{ij}

x_{ij} คือค่าสังเกตของตัวแปรที่ i บนหน่วยสังเกตที่ j

p คือจำนวนตัวแปรที่ศึกษา

เมื่อ $i = 1, 2, ..., p, j = 1, 2, ..., n$

ก่อนการนำไปวิเคราะห์ PCA ภายใต้เมทริกซ์ความแปรปรวนร่วม Baxter (1991) ได้นำการแปลงข้อมูลนี้ ไปใช้กับข้อมูลองค์ประกอบทางเคมีของแก้ว และประสบปัญหาว่าสมการการแปลงข้อมูลของ Aitchison ไม่ได้นิยามเมื่อค่า x_{ij} เท่ากับศูนย์ จึงกำหนดค่าศูนย์ให้มีค่าน้อยกว่าค่าที่น้อยที่สุดเล็กน้อย เขาสรุปว่าการวิเคราะห์ PCA โดยใช้ข้อมูลดิบหรือข้อมูลที่แปลงแล้วให้สารสนเทศที่เป็นประโยชน์ แต่การใช้ PCA บนข้อมูลที่แปลงแล้วควรใช้ด้วยความระมัดระวังเพราะจะมีความไวต่อค่าที่เล็ก ผลที่ได้อาจถูกกำหนดโดยตัวแปรที่มีค่าน้อย Baxter (1992) ได้เสนอให้การแปลงข้อมูลของ Aitchison แต่นำข้อมูลที่แปลงแล้วมาวิเคราะห์ PCA ภายใต้เมทริกซ์สหสัมพันธ์ และแสดงผลที่ได้น่าเชื่อถือกว่าการวิเคราะห์ภายใต้เมทริกซ์ความแปรปรวนร่วม Baxter (1995) ได้เสนอให้ใช้การแปลงข้อมูลแบบอันดับ (rank transformation) ได้เปรียบเทียบการแปลงข้อมูลหลายแบบบนชุดข้อมูลองค์ประกอบทางเคมีของแก้ว Marden (1999) ได้เสนอการวิเคราะห์องค์ประกอบหลักอย่างแกร่ง (Robust Principal Component Analysis, RPCA) แทนการวิเคราะห์องค์ประกอบหลักแบบปกติ (PCA) ในกรณีที่ตัวอย่างของข้อมูลที่ทำการศึกษามีค่านอกกลุ่มโดยที่ค่านอกกลุ่มไม่มีความผิดปกติอย่างรุนแรง โดยใช้ตัวประมาณค่าแบบแกร่งโดยแบ่งออกเป็น 2 วิธี คือ เครื่องหมายของตัวแปรพหุ (multivariate signs) และอันดับของตัวแปรพหุ (multivariate rank) ต่อจากนั้นก็จะหาค่าของเวกเตอร์ไอเกนของเมทริกซ์ความแปรปรวนร่วมของตัวอย่างของข้อมูลตัวแปรที่ได้ทำการแปลงดังกล่าวแล้ว แล้วเขาก็ทำการประมาณค่าเมทริกซ์ความแปรปรวนร่วมด้วยตัวประมาณค่าคือ ตัวประมาณค่าแบบเครื่องหมาย (the sign estimator, $\hat{\sum}_{Sn}$) ตัวประมาณค่าแบบอันดับ (the rank estimator, $\hat{\sum}_{Rn}$) และตัวประมาณค่าแบบวงรีต่ำสุด (the minimum-volume-ellipsoid estimator, $\hat{\sum}_{MVE}$) เปรียบเทียบกับตัวประมาณค่าแบบปกติ (the regular estimator, $\hat{\sum}_{n}$) ที่ได้จากวิธีวิเคราะห์องค์ประกอบหลักแบบปกติ (PCA) โดยข้อมูลที่นำมาทำการศึกษาได้มาจากการจำลองข้อมูล 1,000 ครั้งให้มีการแจงแจกแบบปกติสองตัวแปรที่ขนาดตัวอย่างต่างๆ การแจกแจงแบบโคชีสองตัวแปร การแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียลสองตัวแปร การแจกแจงแบบยูนิฟอร์มสองตัวแปรและจากการจำลองข้อมูลที่มีข้อมูลผิดปกติปนอยู่ด้วย (contaminate) ที่ระดับต่างๆ จากการศึกษาพบว่าโดยทั่วไปแล้วตัวประมาณค่าแบบเครื่องหมายและตัวประมาณแบบอันดับจะเป็นตัวประมาณค่าที่ดีที่สุด ยกเว้นในกรณีที่ข้อมูลผิดปกติในระดับสูงตัวประมาณค่าแบบวงรีต่ำสุดจะเป็นตัวประมาณค่าที่ดีที่สุด ส่วนตัวประมาณค่าแบบปกติจะให้ผลใช้ได้ในกรณีที่ข้อมูลมีการแจงแจกแบบปกติสอง

ตัวแปรที่ขนาดตัวอย่างต่างๆ การแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียลสองตัวแปร และการแจกแจงแบบยูนิฟอร์มสองตัวแปร นอกจากนั้นเขาได้ประยุกต์กับข้อมูลจริงของตัวอย่างของรถยนต์ 111 ตัวอย่างซึ่งประกอบด้วยตัวแปร p = 11 ตัวโดยตัวแปรเหล่านี้ได้มาจากการวัดค่าต่างๆ ของรถยนต์ พบว่าเมื่อทำการวิเคราะห์การประมาณค่าองค์ประกอบหลักในกรณีที่มีข้อมูลสมบูรณ์ไม่ตัดค่านอกกลุ่มออก และกรณีที่ข้อมูลตัดค่านอกกลุ่มออก พบว่าค่าประมาณองค์ประกอบหลักที่ 1-2 ที่ได้จากการประมาณค่าเมทริกซ์ความแปรปรวนร่วมด้วยตัวประมาณค่าแบบปกติ ค่าประมาณองค์ประกอบหลักทั้งสองที่ได้จากการวิเคราะห์ข้อมูลที่ตัดค่านอกกลุ่มแล้วจะมีค่ามากกว่าในกรณีที่ข้อมูลยังไม่ตัดค่านอกกลุ่มออกแสดงให้เห็นว่า ค่าประมาณองค์ประกอบหลักที่ได้จะไม่คงที่และเมื่อเปรียบเทียบกับค่าประมาณองค์ประกอบหลักที่ 1-2 ที่ได้จากการประมาณค่าเมทริกซ์ความแปรปรวนร่วมอีก 3 ตัวคือตัวประมาณค่าแบบเครื่องหมาย ตัวประมาณค่าแบบอันดับและตัวประมาณค่าแบบวงรีต่ำสุดที่ทำการวิเคราะห์ข้อมูลที่ตัดค่านอกกลุ่มออกและที่ยังไม่ตัดค่านอกกลุ่มออกค่าประองค์ประกอบหลักที่ได้จะมีค่าใกล้เคียงและแตกต่างกันเพียงเล็กน้อยแสดงให้เห็นว่าค่าประมาณองค์ประกอบหลักที่ได้จะค่อนข้างคงที่ ดังนั้นจึงกล่าวได้ว่าการตัดค่านอกกลุ่มออกมีผลต่อค่าประมาณองค์ประกอบหลักที่ได้จากการประมาณค่าเมทริกซ์ความแปรปรวนร่วมด้วยตัวประมาณค่าแบบปกติ (สอดคล้องกับความจริงที่ว่าเมทริกซ์ความแปรปรวนร่วมจะมีความไวต่อค่านอกกลุ่ม) ภาณุพงษ์ พนมวัน (2547) ศึกษาการจัดกลุ่มของลูกปัดแก้วที่ได้จากการขุดค้นแหล่งโบราณคดีพรหมทินใต้ จังหวัดลพบุรี พิจารณาจากองค์ประกอบทางเคมีภายใต้การแปลงข้อมูลแบบต่างๆ สามารถจัดกลุ่มของลูกปักแก้วได้ 4 กลุ่ม สอดคล้องกับสีของลูกปัด แต่พบว่าการแปลงข้อมูลแบบอันดับไม่สามารถจัดกลุ่มของลูกปัดแก้วได้อย่างชัดเจน

ดังนั้นผู้วิจัยสนใจศึกษาการประยุกต์การวิเคราะห์องค์ประกอบหลักแบบปกติ (PCA) กับการวิเคราะห์องค์ประกอบหลักแบบแกร่ง (RPCA) บนข้อมูลแบบ Compositional Data เพื่อดูความสามารถในการจำแนกกลุ่มของข้อมูลโดยข้อมูลที่ใช้เป็นองค์ประกอบทางเคมีของดินซึ่งทำการวัดโดยการใช้เทคนิคการถ่ายภาพด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope, SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์ที่วัดได้จากตัวอย่างดินด้วยหัวตรวจวัด Energy Dispersive X-ray Spectrometer (EDX) ช่วยให้สามารถอนุมานได้ว่าตัวอย่างของดินที่นำมาทำการศึกษานั้นประกอบด้วยองค์ประกอบทางเคมีของแร่ธาตุอาหารใด ในการศึกษาองค์ประกอบทางเคมีของดิน จะมีองค์ประกอบหลักคือ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) ไทเทเนียม (Ti) คาร์บอน(C) และโซเดียม (Na) และเพื่อหาตัวแปรที่มีบทบาทสำคัญในการกำหนดความผันแปรของข้อมูลในแต่ละกลุ่มโดยคาดหวังว่าโครงสร้างหรือลักษณะการจัดเข้ากลุ่มอาจจะมีความสัมพันธ์หรือสอดคล้องกับปัจจัย

ต่างๆ เช่น ชนิดของดิน ลักษณะทางกายภาพของดิน (สี ค่า pH) ปริมาณโลหะหนัก หรือปริมาณจุลธาตุอาหารในดินรวมถึงลักษณะการใช้ประโยชน์ของดินโดยรวม เพื่อเป็นประโยชน์ในการปรับปรุงคุณภาพดินเพื่อการเกษตรกรรม

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาการประยุกต์วิธีวิเคราะห์องค์ประกอบหลัก (PCA) กับวิธีวิเคราะห์องค์ประกอบหลักอย่างแกร่ง (RPCA)

2. เพื่อศึกษาลักษณะของการแปลงข้อมูลด้วยวิธีของ Aitchison (1986) สำหรับข้อมูลองค์ประกอบทางเคมีของดินซึ่งมีลักษณะเป็น compositional data (ค่าเปอร์เซ็นต์)

ขอบเขตของการวิจัย

1. ข้อมูลที่นำมาศึกษาเป็นองค์ประกอบทางเคมีของดิน ซึ่งถูกวัดโดยการใช้เทคนิคทางกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope, SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์เรย์ที่วัดได้จากตัวอย่างวัตถุด้วยหัวตรวจวัด Energy Dispersive X-ray Micro-analysis (EDX) จากตัวอย่างดิน 58 ตัวอย่างที่ถูกเก็บมาจาก 3 จังหวัด ได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม นำมาผึ่งให้แห้งและแยกขนาดของดินให้มีขนาดที่ 65 ไมครอน ซึ่งเป็นขนาดที่ดูดซับโลหะมากที่สุด ก่อนนำไปวิเคราะห์ด้วยเทคนิคทางกล้องจุลทรรศน์โดยใช้หัวตรวจวัด Secondary Electron Image Detector กับ Energy Dispersive X-ray Detector (SEM/EDX)

2. ตัวแปรที่ถูกนำมาศึกษาได้แก่ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) ไทเทเนียม (Ti) คาร์บอน(C) และโซเดียม (Na)

3. วิธีการวิเคราะห์องค์ประกอบหลักในการจำแนกกลุ่มของดินมี 3 วิธี ได้แก่

3.1 การวิเคราะห์องค์ประกอบหลักแบบปกติ (PCA) โดยใช้ตัวประมาณค่าแบบภาวะน่าจะเป็นสูงสุด (the maximum likelihood estimator: $\hat{\sum}_n$) ประมาณค่าเมทริกซ์ความแปรปรวนร่วม

3.2 การวิเคราะห์องค์ประกอบหลักอย่างแกร่ง (RPCA) โดยใช้ตัวประมาณค่าแบบอันดับ (the rank estimator: $\hat{\sum}_{Rn}$) ประมาณค่าเมทริกซ์ความแปรปรวนร่วม

3.3 การวิเคราะห์องค์ประกอบหลักอย่างแกร่ง (RPCA) โดยใช้ตัวประมาณค่าแบบเครื่องหมาย(the sign estimator: $\hat{\sum}_{Sn}$)ประมาณค่าเมทริกซ์ความแปรปรวนร่วม

4. ข้อมูลที่นำมาวิเคราะห์มี 2 ลักษณะ คือ

4.1 ข้อมูลดิบ (raw data)

4.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison (1986) ดังในสมการที่ (1)

ประโยชน์ของการวิจัย

1. สามารถนำระเบียบวิธีทางสถิติในการวิเคราะห์องค์ประกอบหลักมาประยุกต์ใช้กับข้อมูลองค์ประกอบทางเคมีของดินและเลือกวิธีการประมาณค่าเมทริกซ์ความแปรปรวนร่วมที่เหมาะสมกับข้อมูลที่มีค่านอกกลุ่ม ในกรณีที่ค่านอกกลุ่มเป็นค่าที่ผิดปกติไม่รุนแรง

2. ผลจากการวิเคราะห์เชิงสถิติสามารถจำแนกกลุ่มของดิน ตามองค์ประกอบทางเคมี และช่วยให้ทราบแนวทางในการปรับปรุงคุณภาพของดินเพื่อให้เหมาะสมในการทำเกษตรกรรม

นิยามศัพท์เฉพาะ

1. ดิน การที่จะให้คำนิยามของคำว่าดินนั้น ขึ้นอยู่กับวัตถุประสงค์และประโยชน์ที่จะได้รับจากการศึกษาดิน โดยทั่วไปการศึกษาในเรื่องดินแบ่งออกเป็น 2 แนวทางหลักด้วยกันคือ

ปฐพีวิทยาธรรมชาติ (pedology) แนวทางหลักนี้ศึกษาเกี่ยวกับการกำเนิดดิน การจำแนกและการตรวจลักษณะดิน โดยเน้นดินในสภาพเทหวัตถุธรรมชาติมากกว่าใช้ดินเพื่อการปลูกพืช ความรู้ในแนวทางการศึกษานี้จะเกิดประโยชน์โดยตรงต่อวิศวกรมากกว่าเกษตรกร เป็นต้น (pedon เป็นภาษากรีก หมายถึง soil หรือ earth) ดังนั้น ในแนวทางเช่นนี้ ดินจึงหมายถึงเทหวัตถุธรรมชาติ (natural body) ที่ปกคลุมผิวโลกอยู่บางๆ เกิดขึ้นจากผลของการแปรสภาพหรือผุพังของหินและแร่และอินทรียวัตถุผสมคลุกเคล้ากัน

ปฐพีวิทยาสัมพันธ์ (edaphology) แนวทางหลักนี้ศึกษาถึงความสัมพันธ์ระหว่างดินกับสิ่งมีชีวิต โดยเฉพาะพืช การใช้ที่ดินเพื่อการปลูกพืช รวมทั้งสมบัติของดินที่มีผลต่อการให้ผลผลิตของพืช (edaphos เป็นคำในภาษากรีก หมายถึง soil หรือ ground) โดยแนวทางเช่นนี้ ความหมายหรือคำจำกัดความของดินคือ เทหวัตถุที่เกิดขึ้นตามธรรมชาติรวมกันเป็นชั้น (profile) จากส่วนผสมของแร่ธาตุต่างๆ ที่สลายตัวเป็นชิ้นเล็กชิ้นน้อยกับอินทรียวัตถุที่เปื่อยผุพัง อยู่รวมกันเป็นชั้นบางๆ ห่อหุ้มผิวโลก และเมื่อมีอากาศและน้ำเป็นปริมาณที่เหมาะสมแล้วจะช่วยค้ำจุนพร้อมทั้งช่วยในการยังชีพและการเจริญเติบโตของพืช

2. การจำแนกดิน (Soil Classification) หมายถึง การแบ่งดินออกเป็นหมวดหมู่ในระดับต่างๆ ของความคล้ายคลึงกันของสมบัติของดิน และลักษณะการเกิดของดินตามวัตถุประสงค์ที่ต้องการ โดยใช้ข้อมูลจากการศึกษาดินเป็นพื้นฐาน

3. องค์ประกอบทางเคมีของดิน หมายถึง การศึกษาคุณลักษณะย่อยทางเคมีในแง่ของธาตุอาหารที่มีผลต่อการเจริญเติบโตของพืชที่มีอยู่ในตัวอย่างดิน ที่ได้จากการวิเคราะห์โดยใช้เทคนิคต่างๆ (ดูรายละเอียดบทที่ 2)

4. ค่านอกกลุ่ม (outliers) ในการวิจัยนี้จะพิจารณาโดยใช้ Boxplot แบ่งค่านอกกลุ่มออกเป็น 2 ระดับ คือ ค่านอกกลุ่มปานกลาง (mild outliers) และค่านอกกลุ่มสุดปลาย (extreme outliers) โดยค่านอกกลุ่มปานกลางคือค่าสังเกตที่อยู่ในช่วง $(Q_1 - 3IQR, Q_1 - 1.5IQR)$ หรือ $(Q_1 + 1.5IQR, Q_1 + 3IQR)$ ซึ่งการพลอตค่าสังเกตเหล่านี้แทนด้วยสัญลักษณ์ o และค่านอกกลุ่มสุดปลายคือค่าสังเกตที่มีค่าน้อยกว่า $Q_1 - 3IQR$ หรือ ค่าสังเกตที่มีค่ามากกว่า $Q_1 + 3IQR$ ซึ่งการพลอตค่านอกกลุ่มสุดปลายนี้แทนด้วยสัญลักษณ์ $*$ (เมื่อ Q_1 คือควอร์ไทล์ที่ 1 และ IQR คือพิสัยระหว่างควอร์ไทล์ที่ 3 และควอร์ไทล์ที่ 1)

บทที่ 2
ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องกับงานวิจัยมีอยู่ 2 ส่วน คือ ส่วนแรกเป็นทฤษฎีของดิน ตลอดจนธาตุอาหารที่มีความจำเป็นในการเจริญเติบโตของพืชที่มีอยู่ในดินและปัจจัยอื่นๆ ที่เกี่ยวข้อง ส่วนที่ 2 เป็นทฤษฎีทางสถิติที่ใช้ในการจัดกลุ่ม และวรรณกรรมที่เกี่ยวข้องเป็นตัวอย่างของงานวิจัยที่ใช้เทคนิคการประมาณค่า องค์ประกอบหลักแบบปกติเปรียบเทียบกับการประมาณค่าองค์ประกอบหลักอย่างแกร่งในกรณีที่ข้อมูลมีค่านอกกลุ่ม

1. ทฤษฎีของดิน (มหาวิทยาลัยเกษตรศาสตร์ 2544: 1-6)

1.1 ความสำคัญของดินกับสิ่งมีชีวิต

ดินเป็นระบบนิเวศ (ecological system) มีพลวัต (dynamic) ที่มีความสำคัญดังนี้

1. ดินเป็นแหล่งผลิตปัจจัยทั้ง 4 ของมนุษย์ อันได้แก่ อาหาร เครื่องนุ่งห่ม ที่อยู่อาศัย และยารักษาโรค ซึ่งอาจได้มาจากดินทั้งทางตรงและทางอ้อม

2. ดินเป็นเครื่องกรองที่มีชีวิต จึงมีผู้ใช้กำจัดของเสียทั้งของแข็งและของเหลว แล้วกักไม่ให้สารมลพิษ (pollution) ตลอดจนเชื้อโรคลงไปปนเปื้อนน้ำใต้ดิน

3. ดินทำหน้าที่เป็นที่เกาะยึด (anchorage) ของรากพืช เพื่อยึดลำต้นให้แน่น ไม่ให้ล้มเอียง เป็นที่เก็บน้ำแก่พืช ให้อาหารแก่รากพืช ในการหายใจ และให้ธาตุอาหารแก่พืช เพื่อการเจริญเติบโต ทนทานต่อโรคแมลงและภัยธรรมชาติ

ในด้านธาตุอาหารพืช มีกฎที่ควรเข้าใจเป็นพื้นฐานเบื้องต้น ซึ่งนักเคมีชาวเยอรมันชื่อ Justus von Liebig เป็นผู้ตั้งกฎไว้ชื่อว่า Law of the minimum หรือ Law of limiting factors มีใจความว่า “ในบรรดาปัจจัยต่างๆ ที่ควบคุมการเจริญเติบโตของพืชนั้น ปัจจัยที่มีอยู่น้อยที่สุดเป็นปัจจัยที่จำกัดการเจริญเติบโต และการให้ผลผลิตของพืช” ซึ่งหมายความว่า ปัจจัยต่างๆ ที่จำเป็นต่อการเจริญเติบโต พืชจะขาดปัจจัยใดปัจจัยหนึ่งไม่ได้และหากปัจจัยไม่พอเพียงเท่าปริมาณที่พืชต้องการ ปัจจัยนั้นจะเป็นตัวจำกัดการเจริญเติบโตและผลผลิตของพืช

ดิน ในที่นี้หมายถึง เทหวัตถุที่เกิดขึ้นตามธรรมชาติรวมกันเป็นชั้น (profile) จากส่วนผสมของแร่ธาตุต่างๆ ที่สลายตัวเป็นชิ้นเล็กชิ้นน้อยกับอินทรียวัตถุที่เปื่อยผุพัง อยู่รวมกันเป็น

ชั้นบางๆ ห่อหุ้มผิวโลก และเมื่อมีอากาศและน้ำเป็นปริมาณที่เหมาะสมแล้ว จะช่วยค้ำจุนพร้อมทั้งช่วยในการยังชีพและการเจริญเติบโตของพืช

ส่วนประกอบของดิน (soil component) จะแบ่งส่วนประกอบออกตามความสำคัญที่เกี่ยวข้องกับการเจริญเติบโตของพืชได้เป็น 4 ส่วนใหญ่ๆ คือ

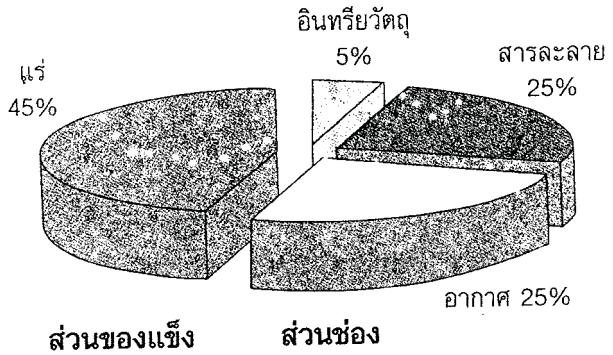
1. อนินทรียวัตถุ (mineral matter) เป็นส่วนที่เกิดจากชิ้นเล็กชิ้นน้อยของแร่และหินต่างๆ ที่สลายตัวโดยทางเคมี ทางฟิสิกส์ และทางชีวเคมี

2. อินทรียวัตถุ (organic matter) ได้แก่ส่วนที่เกิดจากการเน่าเปื่อยผุพังหรือการสลายตัวของเศษเหลือของพืชและสัตว์ที่ทับถมกันอยู่บนดิน

3. น้ำที่อยู่ในดินนั้นพบอยู่ในช่องระหว่างเม็ดดิน (aggregate) หรืออนุภาคดิน (particle) ที่เรียกช่องหรือที่ว่างนี้ว่า pore space

4. อากาศ ที่ว่างในดินระหว่างก้อนดินหรืออนุภาคดินนั้นมีอากาศอยู่ แก๊สที่พบโดยทั่วไปในอากาศในดินมีไนโตรเจน ออกซิเจน และคาร์บอนไดออกไซด์

ปริมาตรของแต่ละส่วนประกอบของดินที่เหมาะแก่การเพาะปลูก (ภาพที่ 1) โดยทั่วไปแล้วจะมีส่วนประกอบที่เป็นของแข็งประมาณ 50% โดยปริมาตร (อนินทรียวัตถุ ประมาณ 45%) โดยปริมาตร และอินทรียวัตถุประมาณ 5% โดยปริมาตร) และส่วนประกอบที่เป็นช่องว่างและน้ำ 50% โดยปริมาตร (ซึ่งมีอากาศ 25% โดยปริมาตร และน้ำประมาณ 25% โดยปริมาตร)



ภาพที่ 1 ส่วนประกอบของดินในปริมาตรที่เหมาะสมต่อการเพาะปลูก

ที่มา: มหาวิทยาลัยเกษตรศาสตร์, ปฐพีเบื้องต้น, พิมพ์ครั้งที่ 9 (กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์, 2544), 5.

ดินบน (surface soil) ซึ่งมีเนื้อดินประเภท silt loam มักจะมีส่วนประกอบโดยปริมาตรดังกล่าวแล้ว ดังนั้นดินชนิดนี้จึงนับได้ว่าเป็นตัวแทนของดินทั่วไป แต่อย่างไรก็ตามจะต้องมีปัจจัยอื่นๆ นอกจากดินเหมาะสมกับการเจริญเติบโตของพืชด้วย

ดินทราย (sandy soil) มักจะมีส่วนเป็นช่องน้อย แต่กลับมีส่วนที่เป็นของแข็งมากกว่าดินทั่วไป

ดินเหนียว (clayey soil) มักจะมีส่วนที่เป็นช่องมาก แต่มีส่วนที่เป็นของแข็งน้อยกว่าดินทั่วๆ ไป

ส่วนปริมาณของอินทรียวัตถุจะมากหรือน้อยนั้นขึ้นอยู่กับการใช้ดินหรือแหล่งที่ให้กำเนิดแก่ดินนั้นๆ

หน้าที่ของแต่ละส่วนของดิน แต่ละส่วนของดินมีหน้าที่ต่อการเจริญเติบโตของพืชและสิ่งมีชีวิตในดินแตกต่างกันดังนี้คือ

1. อนินทรียวัตถุ

1.1 เป็นแหล่งกำเนิดของธาตุอาหารพืช และเป็นแหล่งอาหารของจุลินทรีย์ดิน

1.2 เป็นส่วนที่ควบคุมเนื้อดิน (soil texture)

1.3 ส่วนของอนุภาคดินเหนียว (clay fraction) เป็นส่วนที่สำคัญที่สุดในการเกิดกระบวนการทางเคมีต่างๆ ในดิน

2. อินทรียวัตถุ

2.1 เป็นแหล่งกำเนิดธาตุอาหารของพืชและของจุลินทรีย์ดินโดยเฉพาะอย่างยิ่งไนโตรเจน ฟอสฟอรัส และกำมะถัน

2.2 ให้พลังงานแก่จุลินทรีย์ดิน

2.3 ควบคุมสมบัติทางกายภาพสมบูรณ์ของดิน (soil tilt) เช่น โครงสร้างดิน ความร่วนซุย การระบายน้ำและการแลกเปลี่ยนอากาศของดิน

3. น้ำ

3.1 ให้น้ำแก่พืช

3.2 ช่วยในการละลายธาตุอาหารต่างๆ ในดิน และในการดูดและขนย้ายอาหารพืช

4. อากาศ

4.1 ให้ออกซิเจนแก่รากพืชและจุลินทรีย์ในการหายใจ

4.2 ให้คาร์บอนไดออกไซด์ ซึ่งเมื่อรวมกับน้ำจะให้กรดคาร์บอนิค เป็นกรดที่มีความสำคัญยิ่งในกระบวนการทางเคมีในดินและยังเป็นแหล่งให้คาร์บอนแก่จุลินทรีย์บางชนิดในดินด้วย

4.3 ให้แก๊สไนโตรเจน ซึ่งเป็นแหล่งของไนโตรเจนแก่จุลินทรีย์บางชนิด

ค่า pH (พีเอช) เป็นค่าที่แสดงความเป็นกรดจากปฏิกิริยาของอิออนของไฮโดรเจน (H^+) โดยค่าความเป็นกรดเป็นด่างของดินจะวัดจากปฏิกิริยาดินโดยที่

ปฏิกิริยาดิน หมายถึง ระดับชั้นของสภาพกรด หรือสภาพด่างของดิน ซึ่งแสดงด้วยค่า pH ของดิน ดังนี้

pH	สภาพกรดหรือสภาพด่างของดิน
<3.5	กรดรุนแรงมากที่สุด (Ultra acid)
3.5-4.5	กรดรุนแรงมาก (extremely acid)
4.6-5.0	กรดจัดมาก (very strongly acid)
5.1-5.5	กรดจัด (stongly acid)
5.6-6.0	กรดปานกลาง (moderately acid)
6.1-6.5	กรดเล็กน้อย (slightly acid)
6.6-7.3	กลาง (neutral)
7.4-7.8	ด่างเล็กน้อย (slightly alkaline)
7.9-8.4	ด่างปานกลาง (moderately alkaline)
8.5-9.0	ด่างจัด (strongly alkaline)
>9.0	ด่างจัดมาก (very strongly alkaline)

1.2 เนื้อดิน (soil texture)

เนื้อดินเป็นสมบัติทางฟิสิกส์ขั้นมูลฐาน ซึ่งจะมีผลควบคุมสมบัติทางฟิสิกส์อื่นๆ ของดิน เนื้อดินสื่อความหมายด้านขนาดหรือความหยาบ ละเอียดของอนุภาคอนินทรีย์ (inorganic particles) ที่เป็นองค์ประกอบของดินนั้น

ในด้านปฐพีวิทยา เนื้อดินถูกจำแนกเป็นหลายประเภท สิ่งที่กำหนดประเภทของเนื้อดินคือสัดส่วนโดยมวลของอนุภาคอนินทรีย์ 3 กลุ่มขนาด (soil separates) คือ

1. Sand หรืออนุภาคทราย จัดเป็นกลุ่มขนาดโตที่สุดในดิน
2. Silt หรืออนุภาคทรายตะกอนหรืออนุภาคทรายแป้ง จัดเป็นกลุ่มขนาดปานกลาง
3. Clay หรืออนุภาคดินเหนียว จัดเป็นกลุ่มขนาดเล็กที่สุดในดิน

การใช้ดินเชิงปฏิบัติสำหรับเพาะปลูกโดยทั่วไป ไม่จำเป็นต้องทราบเนื้อดินที่แน่นอน เกษตรอาจจำแนกประเภทเนื้อดินออกเป็น 3 กลุ่มใหญ่ๆ ดังนี้ กล่าวคือ

1. กลุ่มดินเนื้อละเอียด (fine-textured soils) ซึ่งประกอบด้วย 5 ประเภท คือ

1.1 ดินเหนียว (clay)

1.2 ดินเหนียวปนทรายแป้ง (silty clay)

1.3 ดินเหนียวปนทราย (sandy clay)

1.4 ดินร่วนเหนียว (clay loam)

1.5 ดินร่วนเหนียวปนทรายแป้ง (silty clay loam)

2. กลุ่มดินเนื้อปานกลาง (medium-textured soils) ประกอบด้วย 4 ประเภทคือ

2.1 ดินร่วน (loam)

2.2 ดินร่วนปนทรายแป้ง (silt loam)

2.3 ดินทรายแป้ง (silt)

3. กลุ่มดินเนื้อหยาบ (coarse-textured soils) ประกอบด้วยดิน 3 ประเภท คือ

3.1 ดินทราย (sand)

3.2 ดินทรายร่วน (loamy sand)

3.3 ดินร่วนทราย (sandy loam)

1.3 ธาตุอาหารที่จำเป็น (essential nutrient elements) สำหรับพืช

เนื่องจากธาตุที่พบในพืชไม่น้อยกว่า 60 ธาตุนั้น มิได้จำเป็นต่อการดำรงชีพของพืชเสียทั้งหมด การที่จะกำหนดว่าธาตุใดเป็นธาตุอาหารที่จำเป็นต่อพืช (essential nutrient elements) นั้นมีหลักพิจารณา 2 ประการคือ ประการแรก ธาตุหนึ่งเป็นธาตุที่จำเป็นต่อพืชเมื่อพืชไม่อาจดำรงชีวิตจนครบชีพจักรจากปราศจากธาตุนั้นโดยสิ้นเชิง ประการที่สอง ธาตุหนึ่งเป็นธาตุที่จำเป็นต่อพืช เมื่อพิสูจน์ได้ว่าธาตุนั้นเป็นองค์ประกอบของสารที่จำเป็นต่อการดำรงชีพของพืช เช่น ไนโตรเจนโปรตีน และแมกนีเซียมในคลอโรฟิลล์ ธาตุดังกล่าวต้องเกี่ยวข้องโดยตรงกับเมแทบอลิซึมของพืชในทางใดทางหนึ่ง

ในการพิจารณานี้หากพิสูจน์ได้ว่าเข้าหลักเกณฑ์ข้อใดข้อหนึ่งก็คือ ธาตุนั้นเป็นธาตุอาหารพืชได้แล้ว แต่ทุกธาตุที่ยอมรับกันในปัจจุบันเข้าเกณฑ์ทั้งสองข้อ อย่างไรก็ตามนักวิชาการด้านธาตุอาหารพืชในอดีตมักยึดหลักการข้อแรกเป็นสำคัญ

ธาตุอาหารของพืชชั้นสูงที่ยอมรับกันอย่างกว้างขวางมีอยู่ 16 ธาตุ ซึ่งแบ่งได้เป็น 2 กลุ่ม คือ

1. ธาตุอาหารมหัตภาคหรือมหาธาตุ (macronutrients หรือ major elements) หมายถึงธาตุที่พืชต้องการปริมาณมากและสะสมในเนื้อเยื่อพืชในความเข้มข้นสูงกว่า 500 มิลลิกรัม/กิโลกรัม (พืชแห้ง) มี 9 ธาตุ ได้แก่ ไฮโดรเจน คาร์บอน ออกซิเจน ไนโตรเจน ฟอสฟอรัส โพแทสเซียม แคลเซียม แมกนีเซียม และกำมะถัน สำหรับไนโตรเจน ฟอสฟอรัส และโพแทสเซียม เรียกรวมกันว่า ธาตุอาหารหลัก (primary nutrient elements) หรือธาตุปุ๋ย (fertilizer elements)

เนื่องจากพืชต้องการในปริมาณมากแต่พืชได้รับจากดินไม่ค่อยเพียงพอ จึงมีการใช้ปุ๋ยประกอบด้วยธาตุทั้งสาม ส่วนแคลเซียม แมกนีเซียม และกำมะถัน เรียกว่า ธาตุอาหารรอง (secondary nutrient elements) เพราะไม่มีปัญหาความขาดแคลนในดินทั่วๆ ไปเหมือนสามธาตุแรก แคลเซียมและแมกนีเซียมนั้นเรียกว่าธาตุปูน (lime elements) เนื่องจากปูนที่ใช้เพื่อแก้ไขดินกรดมีสองธาตุนี้เป็นหลัก

พืชชั้นสูงได้รับธาตุคาร์บอน และธาตุออกซิเจนที่ต้องการจากอากาศโดยตรง คาร์บอนเข้าสู่พืชในรูปของก๊าซคาร์บอนไดออกไซด์ทางปากใบ (stomata) ส่วนออกซิเจนเข้าสู่พืชในรูปก๊าซออกซิเจนทั้งทางปากใบและทางผิวของราก แม้ว่าแร่ในดินจะมีออกซิเจนเป็นองค์ประกอบสำคัญ แต่พืชไม่ได้ใช้ออกซิเจนส่วนนี้เลย ไฮโดรเจนนั้นพืชได้จากไฮโดรเจนอะตอมที่เป็นส่วนประกอบในโมเลกุลของน้ำ เมื่อน้ำเข้าร่วมในการสังเคราะห์แสง นอกจากนี้ธาตุทั้งสองในโมเลกุลของน้ำยังเข้าร่วมเป็นองค์ประกอบของสารอินทรียสาร ในกระบวนการแยกสลายด้วยน้ำ (hydrolysis) อีกด้วย ธาตุอาหารมหัตภาคที่เหลืออีก 6 ธาตุ พืชได้มาจากดินทั้งสิ้น ยกเว้นธาตุไนโตรเจนบางส่วนที่พืชตระกูลถั่วอาจได้มาจากการตรึงก๊าซไนโตรเจนจากอากาศของไรโซเบียมอาศัยที่รากประมาณร้อยละ 96 ของน้ำหนักแห้งของพืชประกอบด้วยคาร์บอน ไฮโดรเจนและออกซิเจน เพียงร้อยละ 4 เท่านั้น ที่เป็นธาตุอื่นๆ และได้มาจากดิน (ตารางที่ 1) แม้กระนั้นธาตุอาหารมหัตภาคเพียง 6 ธาตุ และธาตุอาหารจุลภาคอีก 7 ธาตุ ที่ได้จากดิน อันเป็นองค์ประกอบราวร้อยละ 4 ของพืชแห้งนี้เอง ที่เป็นตัวกำหนดการเจริญเติบโตของพืช ทั้งนี้เพราะถ้ามีการจัดการเรื่องน้ำและการระบายอากาศของดินอย่างเหมาะสมแล้ว จะไม่ปรากฏว่าพืชขาดคาร์บอน ไฮโดรเจน และออกซิเจน กระทบกระเทือนต่อการเจริญเติบโตเลย

ตารางที่ 1 ธาตุอาหารที่จำเป็นต่อพืชชั้นสูงทั่วไป และความเข้มข้นในเนื้อเยื่อพืชซึ่งจัดว่าเพียงพอ

ธาตุ	สัญลักษณ์ธาตุ	รูปที่เป็นประโยชน์ต่อพืช	ความเข้มข้นในเนื้อเยื่อพืช	จำนวนอะตอมของธาตุเมื่อเทียบกับโมลิบดินัม
			(มก./กก.)	
โมลิบดินัม	Mo	$MoO_4^{=}$	0.1	1
ทองแดง	Cu	Cu^{+}, Cu^{++}	6	100
สังกะสี	Zn	Zn^{++}	20	300
แมงกานีส	Mn	Mn^{++}	50	1,000
โบรอน	B	H_3BO_3, $B_4O^{=}_7$	20	2,000
เหล็ก	Fe	Fe^{++}, Fe^{+++}	100	2,000
คลอริน	Cl	Cl^{-}	100	3,000
			(%)	
กำมะถัน	S	$SO_4^{=}$	0.1	30,000
ฟอสฟอรัส	P	$H_2PO_4^{-}$, $HPO_4^{=}$	0.2	60,000
แมกนีเซียม	Mg	Mg^{++}	0.2	80,000
แคลเซียม	Ca	Ca^{++}	0.5	125,000
โพแทสเซียม	K	K^{+}	1.0	250,000
ไนโตรเจน	N	NH_4^{+}, NO_3^{-}	1.5	1,000,000
ออกซิเจน	O	O_2, H_2O	45	30,000,000
คาร์บอน	C	CO_2	45	35,000,000
โฮโดรเจน	H	H_2O	6	60,000,000

2. ธาตุอาหารจุลภาคหรือจุลธาตุหรือธาตุอาหารเสริม (micronutrients หรือ trace elements หรือ elements) หมายถึงธาตุที่พืชต้องการปริมาณน้อยและสะสมในเนื้อเยื่อพืชในความเข้มข้นต่ำกว่า 50 มิลลิกรัม/กิโลกรัม (พืชแห้ง) ได้แก่ เหล็ก แมงกานีส ทองแดง สังกะสี โบรอน โมลิบดินัม และคลอรีน (ปัจจุบันเพิ่มนิกเกิล (Ni) อีก 1 ธาตุ จึงรวมเป็น 17 ธาตุ)

พืชได้รับธาตุอาหารจุลภาคจากดิน แม้ว่าพืชต้องการในปริมาณน้อยเมื่อเทียบกับธาตุอาหารมหัตภาค แต่ทั้งนี้ไม่ได้หมายความว่ากลุ่มธาตุทั้งสองประเภทจะมีความสำคัญต่อการเจริญเติบโตของพืชมากน้อยกว่ากัน ความจริงแล้วธาตุที่จำเป็นต่อการเจริญเติบโตของพืชทุกธาตุมีความสำคัญต่อการดำรงชีวิตของพืชเท่าๆ กัน จะต่างกันก็แต่ปริมาณที่พืชต้องการเท่านั้น

นอกจากธาตุเหล็กแล้ว จุลธาตุอาหารเหล่านี้ยังพบอยู่ในดินปริมาณที่น้อยมาก ยิ่งกว่านั้นอัตราที่จะเปลี่ยนมาสู่รูปที่เป็นประโยชน์ต่อพืชยังเป็นไปอย่างช้าๆ อีกด้วย แม้ว่าพืชจะใช้ธาตุเหล่านี้ในปริมาณที่น้อยก็ตาม แต่เมื่อปลูกพืชต่อเนื่องกันเป็นเวลานานพืชก็อาจขาดแคลน ดินที่มักจะขาดหรือมีธาตุอาหารพวกนี้ไม่พอ ได้แก่ ดินทราย ดินอินทรีย์ (organic soil) และดินด่าง (alkaline soil) ทั้งนี้เพราะในดินทรายมีธาตุพวกนี้ในปริมาณต่ำ ส่วนในดินด่างจัดและดินอินทรีย์นั้นอาจจะมีธาตุพวกนี้พอประมาณ แต่ส่วนใหญ่จะอยู่ในสภาพที่พืชใช้ไม่ได้

สำหรับหน้าที่สำคัญและอาการขาดธาตุอาหารทั้งธาตุอาหารหลัก ธาตุอาหารรอง และจุลธาตุอาหารแสดงไว้ในตารางที่ 2

ตารางที่ 2 หน้าที่สำคัญของธาตุอาหารพืช และอาการขาดแคลนธาตุอาหารพืช

ธาตุ	หน้าที่สำคัญ	อาการขาดธาตุ
ไนโตรเจน	เป็นองค์ประกอบของกรดอะมิโน โปรตีน คลอโรฟิลล์ กรดนิวคลีอิก และเอนไซม์ในพืช ส่งเสริมการเจริญเติบโตของยอดอ่อน ใบ และกิ่งก้าน	โตช้า ใบล่างมีสีเหลืองซีดทั้งแผ่นใบ ต่อมากลายเป็นสีน้ำตาลแล้วร่วงหล่น หลังจากนั้นใบบนๆ ก็มีสีเหลือง
ฟอสฟอรัส	ช่วยในการสังเคราะห์โปรตีนและสารอินทรีย์ที่สำคัญในพืช เป็นองค์ประกอบของสารที่ทำหน้าที่ถ่ายทอดพลังงานในกระบวนการต่างๆ เช่น การสังเคราะห์แสงและการหายใจ	ใบล่างเริ่มมีสีม่วงตามแผ่นใบ ต่อมาใบเป็นสีน้ำตาลและร่วงหล่น ลำต้นแกร็น ไม่ผลิดอกออกผล
โพแทสเซียม	ช่วยสังเคราะห์น้ำตาล แป้ง และโปรตีน ส่งเสริมการเคลื่อนย้ายของน้ำตาลจากใบไปยังผล ช่วยให้ผลเจริญเติบโตเร็ว พืชแข็งแรง มีความต้านทานต่อโรคบางชนิด	ใบล่างมีอาการเหลืองแล้วกลายเป็นสีน้ำตาลตามขอบใบแล้วลุกลามเข้ามาเป็นหย่อมๆ ตามแผ่นใบ อาจพบว่าแผ่นใบโค้งเล็กน้อย รากเจริญช้า ลำต้นอ่อนแอ ผลไม่เติบโต
แคลเซียม	เป็นองค์ประกอบในสารที่เชื่อมผนังเซลล์ให้ติดกัน ช่วยในการแบ่งเซลล์ การผสมเกสร การงอกของเมล็ด และช่วยให้เอนไซม์บางชนิดทำงานได้ดี	ใบที่เจริญใหม่ๆ หงิก ตายอดไม่เจริญ อาจมีจุดดำที่เส้นใบ รากสั้น ผลแตก และมีคุณภาพไม่ดี

ตารางที่ 2 (ต่อ)

ธาตุ	หน้าที่สำคัญ	อาการขาดธาตุ
แมกนีเซียม	เป็นองค์ประกอบของคลอโรฟิลล์ ช่วยสังเคราะห์กรดอะมิโน ไขมัน และน้ำตาล ทำให้สภาพกรด ด่างในเซลล์พอเหมาะ ช่วยในการงอกของเมล็ด	ใบแก่จะเหลือง ยกเว้นเส้นใบและใบร่วงหล่นเร็ว
กำมะถัน	เป็นองค์ประกอบของกรดอะมิโน โปรตีน และวิตามิน	ใบทั้งบนและล่างมีสีเหลืองซีด และต้นอ่อนแอ
โบรอน	ช่วยในการออกดอกและการผสมเกสร มีบทบาทสำคัญในการติดผลและการเคลื่อนย้ายน้ำตาลมาสู่ผล การเคลื่อนย้ายของฮอร์โมน การใช้ประโยชน์จากไนโตรเจนและการแบ่งเซลล์	ตายอดตายแล้วเริ่มมีตาข้าง แต่ตาข้างจะตายอีก ลำต้นไม่ค่อยยืดตัว กิ่งและใบจึงชิดกัน ใบเล็ก หนา โค้งและเปราะ
ทองแดง	ช่วยในการสังเคราะห์คลอโรฟิลล์ การหายใจ การใช้โปรตีนและแป้ง กระตุ้นการทำงานของเอนไซม์บางชนิด	ตายอดชะงักการเจริญเติบโตและกลายเป็นสีดำ ใบอ่อนเหลือง พืชทั้งต้นชะงักการเจริญเติบโต
คลอรีน	มีบทบาทบางประการเกี่ยวกับฮอร์โมนในพืช	พืชเหี่ยวง่าย ใบสีซีดและบางส่วนแห้งตาย
เหล็ก	ช่วยในการสังเคราะห์คลอโรฟิลล์ มีบทบาทสำคัญในการสังเคราะห์แสงและหายใจ	ใบอ่อนมีสีขาวซีดในขณะที่ใบแก่ยังเขียวสด
แมงกานีส	ช่วยในการสังเคราะห์และการทำงานของเอนไซม์บางชนิด	ใบอ่อนมีสีเหลืองในขณะที่เส้นใบยังเขียว ต่อมาใบที่มีอาการดังกล่าวจะเหี่ยวแล้วร่วงหล่น
โมลิบดินัม	ช่วยให้พืชใช้ไนเตรตให้เป็นประโยชน์ในการสังเคราะห์โปรตีน	พืชมีอาการคล้ายขาดไนโตรเจน ใบมีลักษณะโค้งคล้ายถ้วย ปรากฏจุดเหลืองๆ ตามแผ่นใบ

ตารางที่ 2 (ต่อ)

ธาตุ	หน้าที่สำคัญ	อาการขาดธาตุ
สังกะสี	ช่วยในการสังเคราะห์ออกซิน (ฮอร์โมนพืชชนิดหนึ่ง) คลอโรฟิลล์ และแป้ง	ใบอ่อนมีสีเหลืองซีดและปรากฏสีขาวๆ ประปรายตามแผ่นใบ โดยเส้นใบยังเขียวรากสั้นไม่เจริญตามปกติ

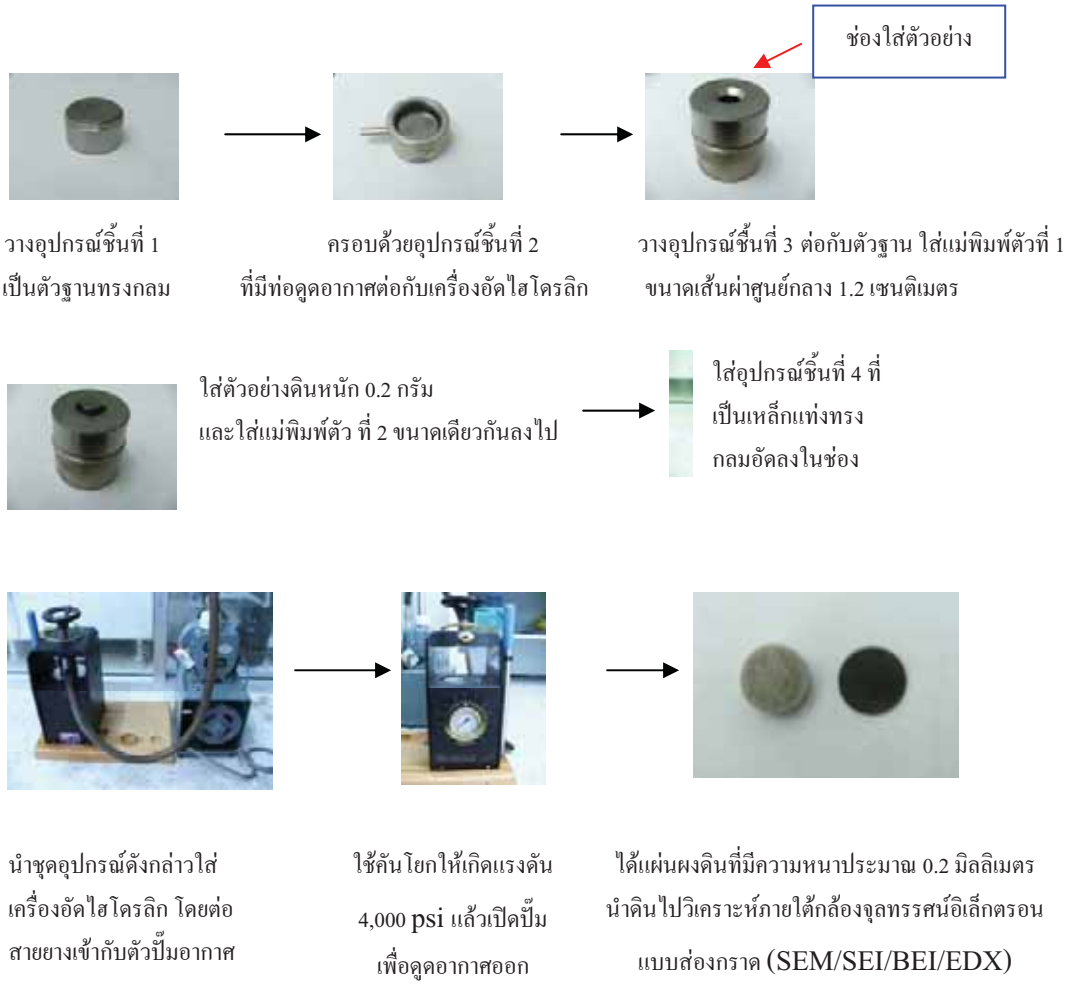
2. แหล่งที่มาของตัวอย่างดิน

2.1 ข้อมูลที่มาของตัวอย่างดิน

ข้อมูลที่ใช้ในงานวิจัยนี้ได้รับความอนุเคราะห์จาก ผศ.ดร.มยุวา อารีกิจเสรี โครงการศึกษาผลของปริมาณกรดฮิวมิคและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดิน และโลหะที่เป็นพิษกรณีศึกษาภาคตะวันตกของประเทศไทย โดยนำข้อมูลบางส่วนมาใช้ในงานวิจัยคือ ตัวอย่างดิน 58 ตัวอย่าง จาก 3 จังหวัด ได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม นำมาผึ่งให้แห้งและแยกขนาดของดินให้มีขนาดที่ 212 และ 65 ไมครอน ก่อนนำไปทำการวิเคราะห์ต่อไปโดยการศึกษาองค์ประกอบทางเคมีของดินจะทำการวัดโดยการใช้เทคนิคการถ่ายภาพด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope, SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์ที่วัดได้จากตัวอย่างดินด้วยหัวตรวจวัด Energy Dispersive X-ray Spectrometer (EDX) ช่วยให้สามารถอนุมานได้ว่าตัวอย่างของดินที่นำมาทำการศึกษานั้นประกอบด้วยองค์ประกอบทางเคมีของแร่ธาตุอาหารใด ในการศึกษาองค์ประกอบทางเคมีของดิน จะมีองค์ประกอบหลักคือ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (P) แคลเซียม (Ca) เหล็ก (Fe) ไทเทเนียม (Ti) และโซเดียม (Na) ซึ่งในการวิจัยนี้จะทำการศึกษาที่ขนาดตัวอย่าง ขนาด 65 ไมครอน (เป็นขนาดที่ดูดซับโลหะมากที่สุด)

2.2 วิเคราะห์ตัวอย่างดินโดยเทคนิคนาโนเทคโนโลยี

ตัวอย่างดินทั้ง 58 ตัวอย่าง อย่างละ 3 ซ้ำ ได้ถูกนำไปศึกษาโครงสร้างทางสัณฐานวิทยาด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning electron microscope) ของบริษัท JOEL รุ่น JSM6400LV โดยใช้ Secondary electron imaging (SEI) และ Back-scattered electron imaging (BEI) เพื่อวิเคราะห์หาธาตุต่าง ๆ ตลอดจนลักษณะการจัดเรียงตัวของมันในธรรมชาติ ขั้นตอนการเตรียมตัวอย่างดังแผนภาพต่อไปนี้

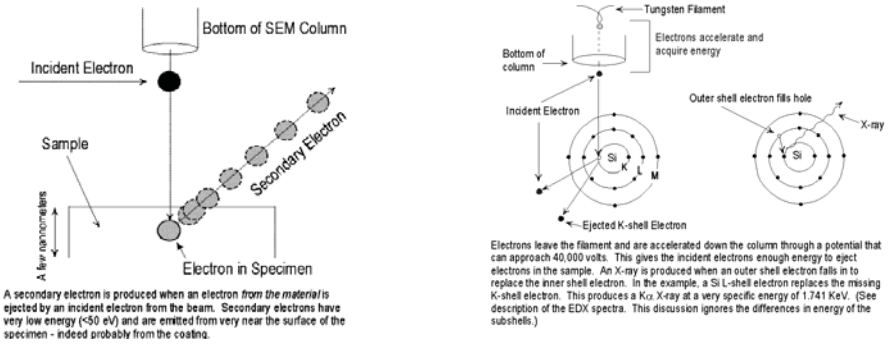


ภาพที่ 2 ภาพขั้นตอนการอัดผงดิน

ที่มา: นัทธีรา และคณะ, “รายงานผลของปริมาณกรดฮิวมิกและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดิน และโลหะที่เป็นพิษกรณีศึกษาภาคตะวันตกของประเทศไทย,” สถาบันวิจัยและพัฒนา มหาวิทยาลัยศิลปากร, 2551. (อัดสำเนา)

ในการศึกษานี้แบ่งการวิเคราะห์เป็น 2 แบบคือ 1) การวิเคราะห์แบบเฉพาะจุด (Point analysis) โดยการวิเคราะห์ธาตุที่จุดเพียงจุดเดียวตลอดการวิเคราะห์ มีรัศมีในการวิเคราะห์ประมาณ 1 ไมโครเมตร และมีความลึกประมาณ 1 ไมโครเมตรจากผิวตัวอย่าง ใช้เวลาสำหรับการวิเคราะห์จุดละ 120 วินาที 2) การวิเคราะห์แบบรวม (Region analysis) โดยการฉายกราดลำอิเล็กตรอนลงบนผิวตัวอย่างในพื้นที่ที่ได้กำหนดไว้มีความลึกประมาณ 1 ไมโครเมตรจากผิวตัวอย่าง โดยใช้เวลาสำหรับการวิเคราะห์พื้นที่ละ 120 วินาที

ภาพแสดงการสะท้อนของอิเล็กตรอนทุติยภูมิ และอะตอมของธาตุในตัวอย่างปลดปล่อยพลังงานแสดงในภาพที่ 3



ก) ภาพสะท้อนของอิเล็กตรอนทุติยภูมิ ข) อะตอมของธาตุตัวอย่างปลดปล่อยพลังงาน

ภาพที่ 3 ภาพสะท้อนของอิเล็กตรอนทุติยภูมิ และอะตอมของธาตุตัวอย่างที่ปลดปล่อยพลังงาน จากคู่มือการใช้งาน JOEL JSM6400LV

ที่มา: นัทธีรา และคณะ, “รายงานผลของปริมาณกรดฮิวมิกและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดิน และโลหะที่เป็นพิษกรณีศึกษาภาคตะวันตกของประเทศไทย,” สถาบันวิจัยและพัฒนา มหาวิทยาลัยศิลปากร, 2551. (อัดสำเนา)

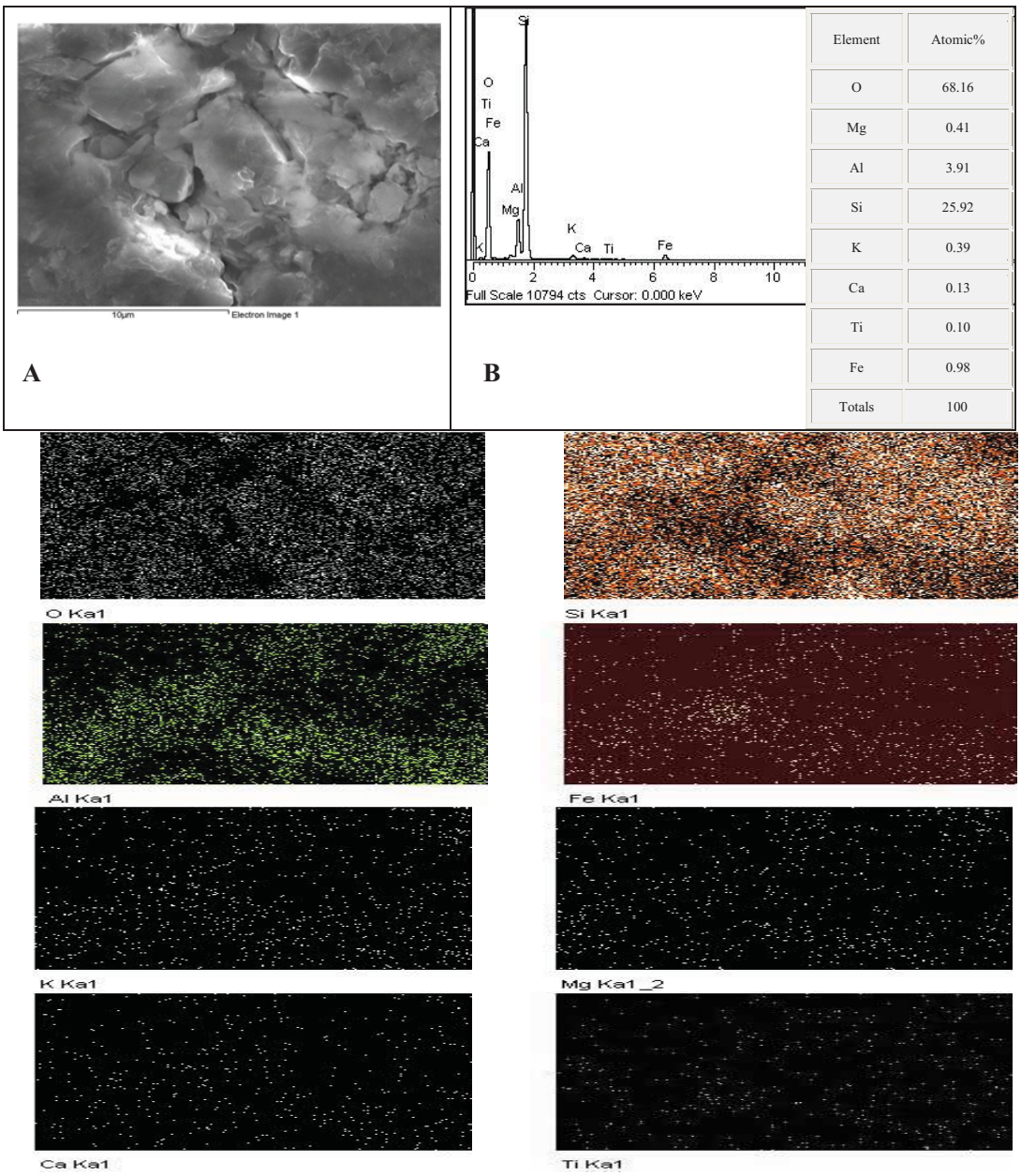
2.3 การศึกษาสัณฐานวิทยาของดินและธาตุที่พบในดิน (Morphological study and Elemental analysis)

จากการศึกษาด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราดที่ใช้หัวตรวจวัดอิเล็กตรอนทุติยภูมิ (SEI) พบว่าโครงสร้างดินที่ศึกษาเป็นดินที่มีรูปร่างกลม (spherical shape) (ตามภาพที่ 4 A) เป็นโครงสร้างที่มีก้อนดินขนาดเล็ก รูปกลม จัดเรียงกันอยู่หลวมๆ ผสมกับก้อนกลมทึบ (granular) ซึ่งโครงสร้างแบบนี้แสดงว่าจะไม่เชื่อมยึดกันอย่างแน่น หมายถึงช่องว่างภายในที่มีปริมาณมาก ทำให้ดินมีลักษณะร่วนซุยดี เกิดการระบายน้ำและอากาศได้ดี

ผลจากการศึกษาด้วยหัวตรวจวัดรังสีเอ็กซ์เฉพาะธาตุ (EDX) โดยวิเคราะห์ 3 ซ้ำเชิงคุณภาพ (Qualitative analysis) ในแต่ละตัวอย่าง (ซึ่ง 1 บริเวณเก็บ 3 จุด) พบว่าดินทั้งหมดที่ทำการศึกษาประกอบด้วยธาตุดังต่อไปนี้ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมินัม(Al) โพแทสเซียม (K) ซิลิคอน (Si) แคลเซียม (Ca) ไทเทเนียม (Ti) และเหล็ก (Fe) และการวิเคราะห์ทางด้านปริมาณ (Quantitative analysis) พบว่าธาตุแต่ละชนิดมีเปอร์เซ็นต์จากมากไปน้อยดังนี้ O >

Si > Al > Fe > K > Mg > Ca > Ti (ภาพที่ 4 B) และเมื่อศึกษาแผนที่และตำแหน่งของธาตุที่อยู่ในผงดิน (Mapping and location of elements) ที่ทำการศึกษาด้วยโปรเกรม Smiling program (บริษัท JOEL JSM6400LV) พบว่าธาตุแบ่งได้เป็น 3 กลุ่ม คือกลุ่มที่ 1 ประกอบด้วย Si และ O มีปริมาณสูงมากที่สุดและมีการกระจายตัวในลักษณะกลุ่มก้อน (ภาพที่ 4 C และ D) กลุ่มที่ 2 ประกอบด้วย Al และ Fe นั้นมีปริมาณน้อยรองลงมามีการกระจายตัวในลักษณะกลุ่มก้อนเช่นกัน (ภาพที่ 4 E และ F) ในขณะที่กลุ่มที่ 3 ประกอบด้วย K, Mg, Ca และ Ti เป็นกลุ่มที่มีปริมาณธาตุน้อยที่สุดและมีการกระจายตัวในลักษณะเดี่ยวๆ กระจายตัวไปทั่วๆ (ภาพที่ 4 G และ J) จากการศึกษานี้ทำให้ทราบว่า mineral composition ในดินที่ทำศึกษาคือ SiO_2, Al_2O_3 และ FeO_3 เป็นส่วนใหญ่

ผลจาก SEM/EDX นี้แสดงให้ทราบว่าการกระจายตัวของธาตุ (distribution of elements) มีความแตกต่างกัน สรุปที่ได้ว่าธาตุที่มีเปอร์เซ็นต์สูง (High percentage of elements) คือ O, Si, Al และ Fe และอยู่เป็นกลุ่มก้อนในขณะที่ธาตุกลุ่มที่มีเปอร์เซ็นต์ต่ำคือ Mg, K, Ca และ Ti มีการกระจายตัวทั่วๆ และบ่งบอก mineral composition ของธาตุในที่ต่างๆ

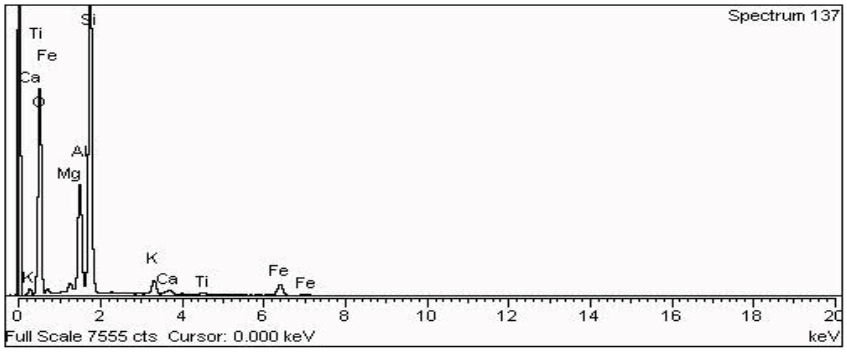


ภาพที่ 4 (A) รูปจาก SEM ใช้หัวตรวจวัด EDX ทำการศึกษาด้วยโปรแกรม Smiling programs (กำลังขยาย 4000 เท่า) ลักษณะของ X-rays character emitted จาก oxygen (O), magnesium (Mg), aluminum (Al), potassium (K), silicon (Si), calcium (Ca), titanium (Ti) และ Iron (Fe) รูป C-J แสดงแผนที่และตำแหน่งของธาตุที่อยู่ในผงดินแสดงตำแหน่งของ O, Si, Al, Fe, K, Mg, Ca และ Ti

ที่มา: นัทธีรา และคณะ, “รายงานผลของปริมาณกรดฮิวมิกและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดิน และโลหะที่เป็นพิษกรณีศึกษาภาคตะวันตกของประเทศไทย,” สถาบันวิจัยและพัฒนา มหาวิทยาลัยศิลปากร, 2551. (อัดสำเนา)

Sample: sample 2 65*1
Type: Default
ID:

Project Notes:



Spectrum processing :
No peaks omitted

Processing option : All elements analyzed (Normalised)
Number of iterations = 4

Standard :
O SiO2 1-Jun-1999 12:00 AM
Mg MgO 1-Jun-1999 12:00 AM
Al Al2O3 1-Jun-1999 12:00 AM
Si SiO2 1-Jun-1999 12:00 AM
K MAD-10 Feldspar 1-Jun-1999 12:00 AM
Ca Wollastonite 1-Jun-1999 12:00 AM
Ti Ti 1-Jun-1999 12:00 AM
Fe Fe 1-Jun-1999 12:00 AM

Element	Weight%	Atomic%
O K	56.44	70.61
Mg K	0.73	0.60
Al K	9.47	7.02
Si K	26.65	18.99
K K	1.59	0.81
Ca K	0.66	0.33
Ti K	0.44	0.19
Fe K	4.01	1.44
Totals	100.00	

ภาพที่ 5 ตัวอย่างของผลที่ได้จากการวิเคราะห์ด้วย SEM/EDX
ที่มา: นัทธีรา และคณะ, “รายงานผลของปริมาณกรดฮิวมิกและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดิน และโลหะที่เป็นพิษกรณีศึกษาภาคตะวันตกของประเทศไทย,” สถาบันวิจัยและพัฒนา มหาวิทยาลัยศิลปากร, 2551. (อัดสำเนา)

2.4 ลักษณะที่ตั้งและสภาพแวดล้อมของแหล่งที่มาของดิน (แผนที่ภูมิศาสตร์ของกรมแผนที่ทหาร พิมพ์ครั้งที่ 1 RTSD ลำดับชุด L 7017 ระวาง 5138 IV มาตราส่วน 1:50,000)

1. จังหวัดนครปฐม มีเนื่องที่ 2,168.327 ตารางกิโลเมตร แบ่งการปกครองออกเป็น 7 อำเภอ คือ อำเภอเมืองนครปฐม อำเภอกำแพงแสน อำเภอนครชัยศรี อำเภอบางเลน อำเภอสามพราน อำเภอดอนตูม และอำเภอพุทธมณฑล

ทิศเหนือ ติดต่อกับจังหวัดสุพรรณบุรี

ทิศใต้ ติดต่อกับจังหวัดสมุทรสาคร

ทิศตะวันออก ติดต่อกับจังหวัดนนทบุรี และกรุงเทพฯ

ทิศตะวันตก ติดต่อกับจังหวัดราชบุรี

ลักษณะภูมิประเทศ จังหวัดนครปฐม มีพื้นที่เป็นที่ราบลุ่ม ไม่มีภูเขา มีที่ดอนเฉพาะทางทิศตะวันตกเฉียงเหนือของอำเภอเมือง และอำเภอกำแพงแสน เท่านั้น ส่วนที่ราบลุ่มบริเวณลุ่มน้ำท่าจีน (แม่น้ำนครชัยศรี) ได้แก่ท้องที่อำเภอนครชัยศรี อำเภอสามพราน และอำเภอบางเลน เป็นที่อุดมสมบูรณ์ มีการประกอบการเกษตรกรรมและเลี้ยงสัตว์

2. จังหวัดสมุทรสาคร มีเนื้อที่ 851 ตารางกิโลเมตร แบ่งการปกครองเป็น 3 อำเภอ คือ อำเภอเมืองสมุทรสาคร อำเภอกระทุ่มแบน และอำเภอบ้านแพ้ว

ทิศเหนือ ติดต่อจังหวัดนครปฐม

ทิศใต้ ติดต่ออ่าวไทย

ทิศตะวันออก ติดต่อกรุงเทพฯ และสมุทรปราการ

ทิศตะวันตก ติดต่อจังหวัดสมุทรสงคราม และราชบุรี

ลักษณะภูมิประเทศ จังหวัดสมุทรสาคร มีพื้นที่เป็นที่ราบลุ่ม ตั้งอยู่บนปากน้ำท่าจีน ห่างจากทะเลเพียง 2 กิโลเมตร ประชากรส่วนใหญ่มีอาชีพเกษตรกรรม และการประมง เป็นจังหวัดที่อุดมสมบูรณ์ด้วยทรัพยากรสัตว์น้ำ โดยเฉพาะอย่างยิ่งสัตว์น้ำทะเล

3. จังหวัดสมุทรสงคราม มีพื้นที่ 416 ตารางกิโลเมตร แบ่งการปกครองออกเป็น 3 อำเภอ คือ อำเภอเมืองสมุทรสงคราม อำเภออัมพวา และอำเภอบางคนที

ทิศเหนือ ติดต่อกับจังหวัดราชบุรี

ทิศใต้ ติดต่อกับจังหวัดเพชรบุรี และอ่าวไทย

ทิศตะวันตก ติดต่อกับจังหวัดเพชรบุรี และราชบุรี

ทิศตะวันออก ติดต่อกับจังหวัดสมุทรสาคร

ลักษณะภูมิประเทศ จังหวัดสมุทรสงครามมีลักษณะภูมิประเทศคล้ายกับจังหวัดสมุทรสาคร มีพื้นที่เป็นที่ราบลุ่ม มีแม่น้ำแม่กลองไหลผ่าน ติดทะเล ประชากรส่วนใหญ่มีอาชีพเกษตรกรรม และการประมง เป็นจังหวัดที่อุดมสมบูรณ์ด้วยทรัพยากรสัตว์น้ำ

3. การวิเคราะห์องค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุด (Principal Component Analysis based on Maximum likelihood Estimates)

การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) เป็นการอธิบายโครงสร้างของความแปรปรวน และความแปรปรวนร่วมระหว่างตัวแปรต่างๆ โดยใช้ผลบวกเชิงเส้นของตัวแปรเหล่านั้น โดยมีวัตถุประสงค์เพื่อลดข้อมูล (data reduction) และเพื่อให้ตีความข้อมูลง่ายขึ้น โดยสูญเสียสารสนเทศที่มีอยู่ในข้อมูลน้อยที่สุด สมมติมีตัวแปรที่สนใจศึกษา p ตัวแปร ความแปรผันรวม (total variability) คือความผันแปรที่เกิดจากตัวแปร p ตัว พบว่าบ่อยครั้งที่ความแปรผันรวมนี้จริงๆ แล้วเกิดจากองค์ประกอบที่หลักๆ เพียงไม่กี่องค์ประกอบ สมมติเกิดจาก k องค์ประกอบ ดังนั้นสารสนเทศที่มีอยู่ใน k องค์ประกอบนี้เกือบสมบูรณ์เหมือนกับที่มาจากตัวแปรเดิม p ตัว จึงอาจใช้องค์ประกอบหลัก k องค์ประกอบนี้เกือบสมบูรณ์เหมือนกับที่มาจากตัวแปรเดิม p ตัว จึงอาจใช้องค์ประกอบหลัก k องค์ประกอบนี้แทนตัวแปร p ตัวเดิม และข้อมูลชุดเดิมที่ประกอบด้วย n ค่าสังเกตของ p ตัวแปร จึงลดลงเหลือเป็น n ค่าสังเกตของ k องค์ประกอบหลักแทน (ปราณี นิลกรณ์ 2547)

การวิเคราะห์องค์ประกอบหลักช่วยให้มองเห็นความสัมพันธ์ที่เดิมอาจไม่เคยเห็น ดังนั้นอาจช่วยในการตีความที่ปกติอาจตีความไม่ได้

3.1 องค์ประกอบหลักของประชากร (Population principal components)

ในเชิงพีชคณิต องค์ประกอบหลักเป็นผลบวกเชิงเส้นของตัวแปรสุ่ม $X_1, X_2, ..., X_p$ จำนวน p ตัว โดย

p แทนจำนวนตัวแปรที่ศึกษา

n แทนจำนวนค่าสังเกต

x_{ij} คือค่าสังเกตของตัวแปรที่ i บนหน่วยสังเกตที่ j

เมื่อ $i = 1, 2, ..., p, j = 1, 2, ..., n$

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ . & . & . & . \\ . & . & . & . \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{bmatrix}_{p\times n}$$

และ μ แทนเวกเตอร์ค่าเฉลี่ย

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ \mu_p \end{bmatrix}_{p\times 1}$$

ในเชิงเรขาคณิต ผลบวกเชิงเส้นเหล่านี้แทนการเลือกระบบพิกัดใหม่โดยการหมุนระบบพิกัดเดิมที่มี $X_1, X_2, ..., X_p$ เป็นแกนของพิกัด แกนพิกัดจะแทนทิศทางที่มีความแปรปรวนสูงสุดและทำให้โครงสร้างของความแปรปรวนง่ายขึ้นและประกอบด้วยตัวแปรน้อยลง องค์ประกอบหลักขึ้นอยู่กับเมทริกซ์ความแปรปรวนร่วม Σ ของ $X_1, X_2, ..., X_p$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ . & . & . & . \\ . & . & . & . \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}_{p\times p}$$

เมื่อ σ_{ii} แทนความแปรปรวนของ X_i

σ_{ij} แทนความแปรปรวนร่วมระหว่าง X_i กับ X_j เมื่อ $i, j = 1,2,...,p$ โดย $i \neq j$

หรือเมทริกซ์สหสัมพันธ์ ρ เท่านั้น

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2p} \\ . & . & . & . \\ . & . & . & . \\ \rho_{p1} & \rho_{p2} & \dots & \rho_{pp} \end{bmatrix}_{p\times p}$$

เมื่อ ρ_{ij} แทนสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X_i กับ X_j

ไม่จำเป็นต้องใช้ข้อสมมติว่าตัวแปรสุ่มมีการแจกแจงร่วมแบบปกติ แต่องค์ประกอบหลักที่ได้จากการแจกแจงแบบปกติของหลายตัวแปรมีประโยชน์ในการตีความในรูปของทรงรีความหนาแน่นคงที่ (constant density ellipsoids) ดังนั้น ถ้าประชากรมีการแจกแจงแบบปกติอาจตีความเกี่ยวกับองค์ประกอบหลักได้มากกว่า

ให้เวกเตอร์สุ่ม $X' = [X_1, X_2, ..., X_p]$ มีเมทริกซ์ความแปรปรวนร่วมเป็น Σ ที่มีค่าไอเกน $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_p \geq 0$ และเวกเตอร์ไอเกน $e' = [e_1, e_2, ..., e_p]$ พิจารณาผลบวกเชิงเส้น

$$Z_1 = e_1'X = e_{11}X_1 + e_{21}X_2 + ... + e_{p1}X_p$$
$$Z_2 = e_2'X = e_{12}X_1 + e_{22}X_2 + ... + e_{p2}X_p$$
$$\vdots$$
$$Z_p = e_p'X = e_{1p}X_1 + e_{2p}X_2 + ... + e_{pp}X_p$$

ดังนั้น $\text{var}(Z_i) = e_i'\Sigma e_i = \lambda_i, i = 1,2,..., p$

$Cov(Z_i, Z_k) = e_i'\Sigma e_k = 0, i, k = 1,2,..., p$ โดย $i \neq k$

องค์ประกอบหลัก คือ ผลบวกเชิงเส้น $Z_1, Z_2, ..., Z_p$ ที่ไม่มีความสัมพันธ์กันและมีความแปรปรวนของ Z_p สูงสุดเท่าที่จะสูงได้ โดยองค์ประกอบหลักขององค์ประกอบแรกมีผลบวกเชิงเส้นที่มีความแปรปรวนสูงสุด องค์ประกอบที่ 2 มีผลบวกเชิงเส้นที่มีความแปรปรวนรองลงมาจากองค์ประกอบแรก และตั้งฉากกับผลบวกเชิงเส้นขององค์ประกอบแรก องค์ประกอบที่ 3 มีผลบวกเชิงเส้นที่มีความแปรปรวนรองลงมาจากองค์ประกอบที่ 2 และตั้งฉากกับผลบวกเชิงเส้นขององค์ประกอบแรกและองค์ประกอบที่ 2 ตามลำดับ

3.2 การอธิบายความแปรปรวนของตัวอย่างด้วยองค์ประกอบหลัก

สมมติ ข้อมูล $X_1, X_2, ..., X_n$ แทนการสุ่ม n ครั้งอย่างเป็นอิสระกันจากประชากร p มิติที่มีเวกเตอร์ค่าเฉลี่ย μ และเมตริกซ์ความแปรปรวนร่วม Σ ข้อมูลเหล่านี้ ให้เวกเตอร์ค่าเฉลี่ย $\overline{X}$

$$\overline{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ . \\ . \\ \bar{x}_p \end{bmatrix}_{p\times 1}$$

และเมทริกซ์ความแปรปรวนร่วมของตัวอย่าง S

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ . & . & . & . \\ . & . & . & . \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix}_{p\times p}$$

เมื่อ s_{ii} แทนความแปรปรวนของ X_i

s_{ij} แทนความแปรปรวนร่วมระหว่าง X_i กับ X_j เมื่อ $i, j = 1,2,...,p$ โดย $i \neq j$

และเมทริกซ์สหสัมพันธ์ของตัวอย่าง R

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ . & . & . & . \\ . & . & . & . \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix}_{p\times p}$$

เมื่อ r_{ij} แทนสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X_i กับ X_j

จุดประสงค์คือ ต้องการสร้างผลบวกเชิงเส้นของคุณลักษณะที่วัดมาโดยผลบวกเชิงเส้นที่สร้างนี้ต้องไม่สัมพันธ์กันและสามารถใช้อธิบายความแปรปรวนส่วนใหญ่ของตัวอย่างได้ ผลบวก

เชิงเส้นที่ไม่สัมพันธ์กันและมีความแปรปรวนสูงสุดนี้ เรียก องค์ประกอบหลักจากตัวอย่าง (Sample principal component)

ผลบวกเชิงเส้น n ผลบวก

$$l'_1 x_j = l_{11}x_{ij} + l_{21}x_{2j} + ... + l_{p1}x_{pj}, j = 1,2...,n$$

มีค่าเฉลี่ยจากตัวอย่างเป็น $l'_1\bar{x}$ และความแปรปรวนจากตัวอย่างเป็น $l'_1 S l_1$ และความแปรปรวนร่วมระหว่าง $(l'_1 x, l'_2 x)$ เป็น $l'_1 S l_2$

นิยามองค์ประกอบหลักจากตัวอย่างเป็นผลบวกเชิงเส้นที่มีความแปรปรวนสูงสุด เช่นเดียวกับกรณีองค์ประกอบหลักของประชากร โดยมีข้อจำกัดว่าเวกเตอร์สัมประสิทธิ์ l_i ต้องมีความยาวเป็น 1 หรือ $l'_i l_i = 1$ จะได้ว่า

องค์ประกอบหลักจากตัวอย่างองค์ประกอบแรก คือผลบวกเชิงเส้นของ $l'_1 x_j$ ที่ทำให้ความแปรปรวนจากตัวอย่างของ $l'_1 x_j$ สูงสุดโดยมีข้อจำกัดว่า $l'_1 l_1 = 1$

องค์ประกอบหลักจากตัวอย่างองค์ประกอบที่สอง คือผลบวกเชิงเส้นของ $l'_2 x_j$ ที่ทำให้ความแปรปรวนจากตัวอย่างของ $l'_2 x_j$ สูงสุดโดยมีข้อจำกัดว่า $l'_2 l_2 = 1$ และ $l'_1 x_j$ และ $l'_2 x_j$ ไม่มีความสัมพันธ์กัน

ขั้นที่ i องค์ประกอบหลักจากตัวอย่างองค์ประกอบที่ i คือผลบวกเชิงเส้น $l'_i x_j$ ที่ทำให้ความแปรปรวนจากตัวอย่างของ $l'_i x_j$ สูงสุดโดยมีข้อจำกัดว่า $l'_i l_i = 1$ และ $l'_i x_j$ และ $l'_k x_j$ ไม่มีความสัมพันธ์กันสำหรับทุก $k < i$

องค์ประกอบหลักองค์ประกอบแรกทำให้ $l'_1 S l_1$ มีค่าสูงสุดหรือทำให้ $\frac{l'_1 S l_1}{l'_1 l_1}$ มีค่าสูงสุดเช่นเดียวกับในกรณีของประชากรได้ว่า

ค่าความแปรปรวนสูงสุดที่ต้องการคือค่าไอเกนที่มีค่ามากที่สุดของ S ซึ่งคือ $\hat{\lambda}_1$ ซึ่งได้เมื่อ l_1 คือเวกเตอร์ไอเกน l_1 ของ S

การเลือกองค์ประกอบหลักตัวต่อๆ ไปคือ เลือก l_i ที่ทำให้ $\frac{l'_i S l_i}{l'_i l_i}$ สูงสุดโดยมีข้อจำกัดว่า $l'_i S l_k = 0$ หรือ l_i ต้องตั้งฉากกับ l_k เมื่อ k < i ผลที่ได้จะคล้ายกับกรณีประชากร

ถ้า $S = \{S_{ik}\}$ เป็นเมทริกซ์ความแปรปรวนร่วมจากตัวอย่างขนาด $p \times p$ ที่มีค่าไอเกนและเวกเตอร์ไอเกนแต่ละคู่ คือ $(\hat{\lambda}_1, l_1)(\hat{\lambda}_2, l_2)...(\hat{\lambda}_p\, l_p)$ แล้วองค์ประกอบหลักที่ i คือ

$$\hat{Z}_i = l_i'x = l_1x_1 + l_2x_2 + ... + l_px_p, i = 1,2,....., p$$

โดยที่ $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_p \geq 0$ และ X_{ij} เป็นค่าสังเกตของตัวแปร $X_1, X_2, ..., X_p$

นอกจากนั้น

ความแปรปรวนของตัวอย่างของ $\hat{Z}_k = \hat{\lambda}_k, k = 1,2...., p$

ความแปรปรวนร่วมของตัวอย่างระหว่าง $\hat{Z}_i$ กับ $\hat{Z}_k = 0, i \neq k$

ความแปรปรวนของตัวอย่าง (total sample variance) $= \sum_{i=1}^{p} S_{ii} = \hat{\lambda}_1 + \hat{\lambda}_2 + ... + \hat{\lambda}_p$

และ $$r_{\hat{Z}_i, x_k} = \frac{l_{ki}\sqrt{\hat{\lambda}_i}}{\sqrt{S_{kk}}}, i, k = 1,2,..., p$$

3.3 การพลอตกราฟขององค์ประกอบหลัก

การพลอตกราฟขององค์ประกอบหลักอาจช่วยให้มองเห็นค่าสังเกตที่น่าสงสัย และช่วยให้ตรวจสอบข้อสมมติเกี่ยวกับการแจกแจงแบบปกติได้ เนื่องจากองค์ประกอบหลักเป็นผลบวกเชิงเส้นของตัวแปรเดิม ถ้าตัวแปรเดิมมีการแจกแจงแบบปกติ องค์ประกอบหลักจะมีการแจกแจงแบบปกติด้วย (ปราณี นิลกรณ์ 2547) โดยสรุปได้ดังนี้

1. ในการตรวจสอบว่าข้อมูลมีการแจกแจงแบบปกติหรือไม่ ให้พลอตแผนภาพการกระจายขององค์ประกอบหลักคู่แรกๆ และ Q – Q plot ของค่าตัวอย่างขององค์ประกอบหลักแต่ละองค์ประกอบ

2. พลอตแผนภาพการกระจายและ Q – Q plot สำหรับองค์ประกอบหลักตัวหลังๆ เพื่อดูว่าค่าสังเกตใดผิดปกติบ้าง

3. พลอตกราฟ 2 มิติ ของค่าองค์ประกอบที่ 1 กับองค์ประกอบที่ 2 เพื่อเป็นการแยกประเภทหรือกลุ่มของตัวอย่าง (Baxter 1995)

4. พลอตกราฟ 2 มิติ ของค่าองค์ประกอบที่ 1 กับองค์ประกอบที่ 2 เพื่อเป็นการดูค่านอกกลุ่ม (Qutlier) (Banning 2000)

4. การวิเคราะห์องค์ประกอบหลักโดยใช้ตัวประมาณค่าแบบแกร่งของเมทริกซ์ความแปรปรวนร่วม (Robust Principal Component Analysis based on covariance matrix)

Marden (1999) สมมติให้ X แทนเวกเตอร์เชิงสุ่มมิติ $p \times 1$ โดยมีการแจกแจง F_x ซึ่งจุดประสงค์ของการวิเคราะห์องค์ประกอบหลัก คือ การหาผลบวกเชิงเส้นของเวกเตอร์เชิงสุ่ม X (เวกเตอร์ของข้อมูล) ที่ให้ข้อมูลสารสนเทศมากที่สุด โดยข้อมูลสารสนเทศที่ได้นั้นจะถูกวัดจาก

ความสามารถในการอธิบายความผันแปรของข้อมูลนั่นเอง ผลบวกเชิงเส้นเหล่านี้คือค่าเวกเตอร์ไอเกนของเมทริกซ์ความแปรปรวนร่วม (สมมติว่าเราสามารถหาเวกเตอร์ไอเกนได้) ดังนั้นเราสามารถเขียนเมทริกซ์ความแปรปรวนร่วม ดังนี้

$$\Sigma \equiv Cov(X) = \Gamma \Lambda \Gamma' \tag{2}$$

โดยที่ Γ คือ เมทริกซ์ ของเวกเตอร์ไอเกนมิติ p × p , $\Gamma = [\gamma_1, \gamma_2, ..., \gamma_p]$

Λ คือเมทริกซ์ทแยงมุมมิติ $p \times p$ ที่มีค่าไอเกน λ_i เป็นสมาชิกในแนวทแยงมุมลงมา ดังนั้นสมาชิกของเมทริกซ์นี้ตัวที่ ii คือ $\Lambda_{ii} = \lambda_i$ และโดยที่มี $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_p$ (เวกเตอร์ไอเกนจะต้องไม่เป็นเวกเตอร์ที่มีลักษณะเฉพาะตัวเด็ดขาดถ้ามีการผสมที่ซับซ้อนของเวกเตอร์ไอเกน)

ดังนั้น Z_1 คือ องค์ประกอบหลักที่ 1 ซึ่งเป็นผลบวกเชิงเส้นเวกเตอร์เชิงสุ่ม X ที่ให้ข้อมูลสารสนเทศมากที่สุด และ

Z_2 คือ องค์ประกอบหลักที่ 2 ซึ่งเป็นผลบวกเชิงเส้นเวกเตอร์เชิงสุ่ม X ที่ให้ข้อมูลสารสนเทศมากที่สุดรองลงมา (ดูได้จาก Mardia et al., 1979)

เมื่อเราพิจารณาตามรูปแบบดังกล่าว การแจกแจก F_W ของเวกเตอร์เชิงสุ่ม W มิติ $p \times 1$ จะมีพิกัดสมมาตรของศูนย์ถ้า

$$gW \overset{D}{=} W \tag{3}$$

เมื่อ g ใดๆ เป็นสมาชิกของ G $(g \in G)$ โดยที่กลุ่มของเมทริกซ์ทแยงมุมมิติ $p \times p$ จะมีสมาชิกเป็น $g_{ii} \in \{-1, +1\}$ ซึ่งสมการที่ (3) หมายความว่า gW และ W จะมีการแจกแจงแบบเดียวกัน เราสมมติให้เวกเตอร์เชิงสุ่ม X สามารถหาได้บนเมทริกซ์เชิงตั้งฉาก Z และให้เวกเตอร์เชิงสุ่ม b มีมิติ $p \times 1$ ดังนั้นสามารถเขียนเวกเตอร์เชิงสุ่ม X ใหม่ดังนี้

$$X = \Gamma W + b \tag{4}$$

สำหรับเวกเตอร์เชิงสุ่ม W บางตัวที่มีพิกัดสมมาตรที่ศูนย์ เราจะสมมติว่า cov(W) สามารถหาได้และปราศจากการสูญเสียลักษณะทั่วไปโดยที่

$$\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_p \text{ เมื่อ } \lambda_i = \text{Var}(W_j) \tag{5}$$

ดังนั้นจากสมการที่ (3) ที่เราได้คอลัมน์ของเมตริกซ์ Z จะเป็นเวกเตอร์ไอเกนของเมตริกซ์ Σ

เราจะให้สมการที่ (4) เป็นสมการของตัวแปรพหุปกติและมีการแจกแจงลักษณะทรงรีที่สมมาตร

จากนี้ไปเราจะสมมติว่าสมการที่ (4) ของเรามีจุดประสงค์ที่จะประมาณค่าของเมทริกซ์ Γ ซึ่งตั้งอยู่บนเวกเตอร์เชิงสุ่มตัวอย่างที่มีการแจกแจงแบบเดียวกันและเป็นอิสระกันซึ่งคือเวกเตอร์ $X_1, X_2, ..., X_n$ ซึ่งมีการแจกแจง F_x การประมาณค่าเมทริกซ์ความแปรปรวนร่วมแบบปกติ (regular estimate) คือ การประมาณค่าบนเมทริกซ์ความแปรปรวนร่วมของตัวอย่าง ดังนี้

$$\hat{\Sigma}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)(X_i - \overline{X}_n)' \qquad (6)$$

เมื่อ $\overline{X}_n$ คือ เวกเตอร์ค่าเฉลี่ยตัวอย่าง

สำหรับตัวประมาณค่าแบบอื่นๆ ที่จะเลือกใช้คือ ตัวแปรพหุเรขาคณิตที่มีเครื่องหมายและตำแหน่งสามารถดูจาก Koltehinskii (1997), Chaudhuri (1996) และ Mottonen et al. (1997)

4.1 สถิติเครื่องหมาย (Sign statistics)

สำหรับ $w \in R^p$ ให้นิยามของ w ซึ่งเป็นตัวแปรพหุเครื่องหมายโดย

$$S(w) = \frac{w}{\|w\|} \qquad \text{ถ้า} \qquad w \neq 0$$

และ $$S(w) = 0 \qquad \text{ถ้า} \qquad w = 0 \qquad (7)$$

$S(w)$ คือเวกเตอร์หนึ่งหน่วยที่มีทิศทางเดียวกับ w สำหรับเวกเตอร์เชิงสุ่ม X ที่มีค่าตัวที่กำหนดให้ในสมการที่ (4) พวกเราต้องการที่จะหาค่าของ $S(x-b)$ เนื่องจาก b เป็นสัญลักษณ์ที่เราไม่ทราบค่า ดังนั้นเราจะทำการประมาณค่า b ด้วย $\hat{b}_n$ จากตัวอย่างโดยที่ $\hat{b}_n$ คือค่า spatial median (ค่ามัธยฐาน)ของตัวอย่าง

$$\sum_{i=1}^{n}\|x_i - a\| \qquad (8)$$

ดูได้จากบทความของ Small (1990) ที่นำเสนอวิธีการนี้พร้อมทั้งตัวแปรพหุแบบมัธยฐานแบบอื่นๆ ในที่นี้เราจะสามารถประมาณค่าเมทริกซ์ความแปรปรวนร่วมของ $S(x-b)^n$ ซึ่งแทนด้วย Σ_s ด้วยตัวประมาณ ดังนี้

$$\hat{\Sigma}_{Sn} = \frac{1}{n}\sum_{i=1}^{n} S(X_i - \hat{b}_n)S(X_i - \hat{b}_n)' \qquad (9)$$

4.2 สถิติตำแหน่ง (Rank statistics)

ทฤษฎีตำแหน่งของ w จะมีความเกี่ยวข้องกับการแจกแจง F_X คือ

$$R(w, F_X) = \int \frac{w-x}{\|w-x\|} F_X(dx) \qquad (10)$$

ตำแหน่งของ w จะมีความเกี่ยวข้องกับตัวอย่าง นั่นคือทฤษฎีของตำแหน่งจะมีความสัมพันธ์กับการแจกแจง empirical ของฟังก์ชัน $\hat{F}_n$ โดยมีค่าดังนี้

$$R(w,\hat{F}_n)=\frac{1}{n}\sum_{i=1}^{n}\frac{w-x_i}{\|w-x_i\|} \tag{11}$$

เราจะประมาณค่าเมทริกซ์ความแปรปรวนร่วม Σ_r ของ $R(X,\hat{F}_x)$ โดยใช้ตัวอย่าง ซึ่งจะได้ตัวประมาณดังนี้

$$\hat{\Sigma}_{Rn}=\frac{1}{n=1}\sum_{i=1}^{n}R(X_i,\hat{F}_n)R(X_i,\hat{F}_n)' \tag{12}$$

ข้อเสนอแนะ ถ้าสมการที่ 3 ครอบคลุมแล้ว

$$\sum_I=\Gamma\Lambda_I\Gamma' \tag{13}$$

เมื่อ Λ_I คือ เมทริกซ์ทแยงมุมที่มีสมาชิกเป็นค่าไอเกนของเมทริกซ์ความแปรปรวนร่วม Σ_r โดยที่ I เท่ากับ R หรือ S

ผลเด่นชัดที่ตามคือข้อสังเกตว่าเมทริกซ์เชิงตั้งฉาก Γ ในแนวทแยงของเมทริกซ์ความแปรปรวนร่วมของ X และ เครื่องหมายของ $X-b$ และตำแหน่งของ X จะเหมือนกัน ดังนั้นเราสามารถประมาณค่า Γ ด้วยเวกเตอร์ไอเกนของ $\hat{\Sigma}_n,\hat{\Sigma}_{Sn}$ หรือ $\hat{\Sigma}_{Rn}$ อย่างไรก็ตามแม้ว่าตัวประมาณค่าทั้งสามจะประมาณจากเซตของเวกเตอร์ไอเกนชุดเดียวกัน มันก็จะมีค่าน้อยที่สุดเมื่อเวกเตอร์ไอเกนมีลักษณะเฉพาะตัว (unique) แต่ไม่มีอะไรสามารถการันตีได้ว่า เวกเตอร์เหล่านี้จะมีลำดับเดียวกันในการเขียนแยกเชิงสเปกตรัม ดังนั้น สมาชิกในแนวทแยงของเมทริกซ์ Λ_S หรือ Λ_R ก็จะไม่ลดลง

5. วรรณกรรมที่เกี่ยวข้อง

Baxter (1991) ได้ศึกษาการประยุกต์การวิเคราะห์องค์ประกอบหลักบนองค์ประกอบทางเคมีของแก้วซึ่งค่าตอบสนองที่ถูกวัดขึ้นมาจากตัวแปรที่นำมาศึกษาเป็นค่าเปอร์เซ็นซึ่งเป็นข้อมูลในลักษณะ compositive data มีลักษณะไม่เหมาะสมกับการวิเคราะห์โดยตรงบนข้อมูลดิบ เนื่องจากว่าเงื่อนไขของทุกผลรวมของทุกตัวแปรมีค่าเป็น 100% จึงได้ใช้การแปลงข้อมูลซึ่งถูกเสนอโดย Aitchison (1986) ดังสมการ(1) ก่อนการนำไปวิเคราะห์ PCA ภายใต้เมทริกซ์ความแปรปรวนร่วม และประสบปัญหาว่าสมการการแปลงข้อมูลของ Aitchison ไม่ได้นิยามเมื่อค่า x_{ij} เท่ากับศูนย์ จึงกำหนดค่าศูนย์ให้มีค่าน้อยกว่าค่าที่น้อยที่สุดเล็กน้อย เขาสรุปว่าการวิเคราะห์ PCA โดยใช้ข้อมูลดิบหรือข้อมูลที่แปลงแล้วให้สารสนเทศที่เป็นประโยชน์ แต่การใช้ PCA บนข้อมูลที่แปลงแล้วควรใช้ด้วยความระมัดระวังเพราะจะมีความไวต่อค่าที่เล็ก ผลที่ได้อาจถูกกำหนดโดยตัวแปรที่มีค่าน้อย

Baxter (1995) ได้ศึกษาการวิเคราะห์องค์ประกอบหลักเกี่ยวกับการแปลงเชิงเส้นของตัวแปรที่มีความสัมพันธ์กันหลายตัวที่นำมาใช้บ่อย โดยอาศัยการนำค่าจากการวิเคราะห์องค์ประกอบหลักที่ 1 , 2 หรือ 3 เพื่อแสดงโครงสร้างข้อมูล โดยทั่วไปตัวแปรถูกแปลงให้เป็นค่ามาตรฐานที่มีค่าเฉลี่ย 0 และมีค่าความแปรปรวนเป็น 1 และโดยทั่วไปเป็นสิ่งจำเป็นถ้าตัวแปรต่าง ๆ ถูกวัดมาด้วยหน่วยที่ต่างกัน อย่างไรก็ตามถ้าตัวแปรต่าง ๆ มีหน่วยการวัดเหมือนกันทางเลือกที่จะทำให้เป็นค่ามาตรฐานนี้ขึ้นอยู่กับเทคนิคของแต่ละคน อีกทางเลือกหนึ่งก่อนที่จะทำการวิเคราะห์องค์ประกอบหลักคือ อาจมีหรือไม่มีการแปลงข้อมูลก็ได้ เช่นข้อมูลที่อยู่ในมาตรวัดอัตราส่วน (ratio scale) อาจใช้การแปลงโดยใช้ log หลังจากมีการแปลงอาจทำให้เป็นค่ามาตรฐานหรือไม่ก็ได้ Baxter จึงตรวจสอบการแปลงข้อมูลและทำให้เป็นค่ามาตรฐานแบบต่าง ๆ โดยในเรื่องของการวิเคราะห์องค์ประกอบหลักทางเคมีของโบราณวัตถุปัญหาที่พบบ่อย ๆ ปัญหาหนึ่งคือออกไซด์และธาตุต่าง ๆ อาจมีค่าต่ำกว่าค่าต่ำสุดที่ใช้ตรวจจึงบันทึกค่าเป็นศูนย์ ฉะนั้นการแปลงโดยใช้ log จึงเกิดปัญหา นอกจากนี้การแจกแจงของตัวแปรบางตัวในส่วนที่ศึกษาอาจผิดปกติไปมากมายและไม่เหมาะกับการวิเคราะห์องค์ประกอบหลัก ดังนั้น Baxter จึงได้เสนอให้ใช้การแปลงข้อมูลแบบอันดับ (rank transformation) โดยเปรียบเทียบการแปลงข้อมูลหลายแบบบนชุดข้อมูลองค์ประกอบเคมีของแก้ว

Marden (1999) ได้เสนอการวิเคราะห์องค์ประกอบหลักอย่างแกร่ง (Robust Principal Component Analysis: RPCA) แทนการวิเคราะห์องค์ประกอบหลักแบบปกติ (PCA) ในกรณีที่ตัวอย่างของข้อมูลที่ทำการศึกษามีค่านอกกลุ่มโดยที่ค่านอกกลุ่มไม่มีความผิดปกติอย่างรุนแรง โดยทำการแปลงข้อมูลตัวแปรแบ่งออกเป็น 2 วิธี คือ ตัวแปรพหุแบบเครื่องหมาย (multivariate signs) และตัวแปรพหุแบบอันดับ (multivariate rank)ต่อจากนั้นก็จะหาค่าของเวกเตอร์ไอเกนของเมทริกซ์ความแปรปรวนร่วมของตัวอย่างของข้อมูลตัวแปรที่ได้ทำการแปลงดังกล่าวแล้ว หลังจากนั้นเขาก็ทำการประมาณค่าเมทริกซ์ความแปรปรวนร่วมด้วยตัวประมาณค่า 3 ตัวคือตัวประมาณค่าแบบเครื่องหมาย (the sign estimator, $\hat{\sum}_{Sn}$) ตัวประมาณค่าแบบอันดับ (the rank estimator, $\hat{\sum}_{Rn}$) และตัวประมาณค่าแบบวงรีต่ำสุด (the minimum-volume-ellipsoid estimator: $\hat{\sum}_{MVE}$) เปรียบเทียบกับตัวประมาณแบบปกติ (the regular estimator, $\hat{\sum}_{n}$) ที่ได้จากวิธีวิเคราะห์องค์ประกอบหลักแบบปกติ (PCA)โดยข้อมูลที่นำมาทำการศึกษาได้มาจากการจำลองข้อมูล (Simulations) 1,000 ครั้งให้มีการแจงแจกแบบปกติสองตัวแปรที่ขนาดตัวอย่างต่างๆ การแจกแจงแบบโคชีสองตัวแปร การแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียลสองตัวแปร การแจกแจงแบบยูนิฟอร์มสองตัวแปรและจากการจำลองข้อมูลที่มีข้อมูลผิดปกติปนอยู่ด้วย (Contaminate) ที่ระดับต่างๆ จากการศึกษาค่า MSE efficiencies และค่า Relative efficiencies ของแต่ละตัวประมาณค่าเมทริกซ์ความแปรปรวนร่วม

พบว่าโดยทั่วไปแล้วตัวประมาณค่าแบบเครื่องหมายและตัวประมาณแบบอันดับจะเป็นตัวประมาณค่าที่ดีที่สุด ยกเว้นในกรณีที่ข้อมูลผิดปกติในระดับสูงตัวประมาณค่าแบบวงรีต่ำสุดจะเป็นตัวประมาณค่าที่ดีที่สุด ส่วนตัวประมาณค่าแบบปกติจะให้ผลใช้ได้ในกรณีที่ข้อมูลมีการแจงแจกแบบปกติสองตัวแปรที่ขนาดตัวอย่างต่างๆ การแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียลสองตัวแปร และการแจกแจงแบบยูนิฟอร์มสองตัวแปร และนอกจากนั้นเขาได้ประยุกต์กับข้อมูลจริงของตัวอย่างของรถยนต์ 111 ตัวอย่างซึ่งประกอบด้วยตัวแปร p = 11 ตัวโดยตัวแปรเหล่านี้ได้มาจากการวัดค่าต่างๆ ของรถยนต์ จากการศึกษาพบว่ามีอย่างน้อย 3 ตัวประมาณแรกคือ $\hat{\sum}_{Sn}$, $\hat{\sum}_{Rn}$ และ $\hat{\sum}_{n}$ ของการประมาณค่าเมทริกซ์ความแปรปรวนร่วมซึ่งมีกลุ่มของค่านอกกลุ่มแบ่งออกเป็น 2 กลุ่ม โดยที่กลุ่มแรกทีค่านอกกลุ่มของตัวอย่างรถยนต์ 5 ค่า ซึ่งจะอยู่ไกลไปทางด้านซ้ายของข้อมูลทั้งหมดและกลุ่มที่สองมีค่านอกกลุ่มมากกว่า 10 ค่า ซึ่งค่านอกกลุ่มทั้งสองนี้กลุ่มแรกเป็นการรวมกลุ่มของตัวอย่างของรถยนต์สปอร์ต 2 ที่นั่ง ซึ่งจะอยู่ไปทางซ้ายของข้อมูลทั้งหมดและส่วนกลุ่มที่ 2 เป็นค่าของตัวอย่างรถยนต์มินิแวน (เมื่อพิจารณาตามความเป็นจริงแล้วค่านอกกลุ่มทั้ง 15 ค่านี้ เป็นค่าของรถยนต์ที่ไม่มีห้องเก็บของอย่างเป็นทางการและมีค่าของข้อมูลบางตัวที่มีค่าเป็น -2 หรือ -3) และเมื่อทำการประมาณวิเคราะห์องค์ประกอบหลักในกรณีที่มีข้อมูลสมบูรณ์ไม่ตัดค่านอกกลุ่มออกและกรณีที่ข้อมูลตัดค่านอกกลุ่มออก พบว่าค่าประมาณองค์ประกอบหลักที่1และองค์ประกอบหลักที่ 2 ที่ได้จากการประมาณค่าเมทริกซ์ความแปรปรวนร่วมด้วยตัวประมาณแบบปกติค่าประมาณองค์ประกอบหลักที่1-2ในทั้งสองกรณีคือกรณีที่มีข้อมูลสมบูรณ์ไม่ตัดค่านอกกลุ่มออก และกรณีที่ข้อมูลตัดค่านอกกลุ่มออกค่าประมาณที่ได้จะไม่คงที่ เมื่อเปรียบเทียบกับค่าประมาณองค์ประกอบหลักที่การประมาณค่าเมทริกซ์ความแปรปรวนร่วมอีก 3 ตัวที่ค่อนข้างจะคงที่ แสดงให้เห็นว่าการตัดค่านอกกลุ่มออกมีผลต่อการวิเคราะห์องค์ประกอบหลักที่ได้การประมาณค่าเมทริกซ์ความแปรปรวนร่วมด้วยตัวประมาณแบบปกติ(สอดคล้องกับความจริงที่ว่าเมทริกซ์ความแปรปรวนร่วมจะมีความไวต่อค่านอกกลุ่ม) โดยเฉพาะอย่างยิ่งค่าประมาณองค์ประกอบหลักที่ 1 ที่ได้จากตัวประมาณแบบปกติค่าในกรณีที่มีข้อมูลสมบูรณ์ไม่ตัดค่านอกกลุ่มออก และกรณีที่ข้อมูลตัดค่านอกกลุ่มออกจะแตกต่างกันอย่างเห็นชัดเจนนั่นคือค่าประมาณองค์ประกอบหลักที่ 1 ในกรณีที่มีข้อมูลสมบูรณ์ไม่ตัดค่านอกกลุ่มออก จะมีค่าต่ำกว่ากรณีที่ข้อมูลตัดค่านอกกลุ่มออก

ภานุพงษ์ พนมวัน (2547) ได้ศึกษาข้อมูลทางโบราณคดีที่มีลักษณะหลายตัวแปร มีวัตถุประสงค์เพื่อจัดกลุ่มของลูกปัดแก้วจำนวน 206 ลูก จากการขุดค้นแหล่งโบราณคดีพรหมทินใต้ อำเภอโคกสำโรง จังหวัดลพบุรี ประกอบด้วยตัวแปร 15 ตัวคือ อลูมินัม (Al) เหล็ก (Fe) ฟอสฟอรัส (P) แคลเซียม (Ca) ไทเทเนียม (Ti) แมกนีเซียม (Mg) ซิลิคอน (Si) คลอรีน (Cl) ตะกั่ว (Pb) โพแทสเซียม (K) โซเดียม (Na) ทองแดง (Cu) ดีบุก (Sn) แมงกานีส (Mn) และปรอท (Hg) หน่วยที่

วัดเป็นเปอร์เซ็นต์ของปริมาณธาตุในลูกปัดแก้วแต่ละลูกซึ่งเป็นข้อมูลในลักษณะ compositive data การวิเคราะห์ใช้การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) โดยการพลอตกราฟ 2 มิติจากค่าที่ได้จากองค์ประกอบที่ 1 ที่ 2 และที่ 3 และการวิเคราะห์กลุ่ม (Cluster Analysis) ในข้อมูลที่แตกต่างกัน 5 ลักษณะคือ ข้อมูลดิบ ข้อมูลดิบที่เป็นค่ามาตรฐาน ข้อมูลที่มีการแปลงแบบอันดับ ข้อมูลดิบที่มีการแปลงโดย log และข้อมูลที่มีการแปลงโดย log ที่เป็นค่ามาตรฐาน ผลการวิเคราะห์พบ 2 ลักษณะที่ให้สารสนเทศที่ดี คือ ข้อมูลดิบจากการวิเคราะห์องค์ประกอบหลักและการวิเคราะห์กลุ่มด้วยวิธี Complete Linkage สามารถแยกได้ 4 กลุ่มในเรื่องของสีลูกปัดแก้ว และข้อมูลดิบที่แปลงเป็นค่ามาตรฐานจากการวิเคราะห์องค์ประกอบหลักสามารถแยกได้ 3 กลุ่มในเรื่องความโปร่งใสลูกปัดแก้วและการวิเคราะห์กลุ่มด้วยวิธีเดียวกันกับข้อมูลดิบสามารถแยกได้ 4 กลุ่มในเรื่องของสีลูกปัดแก้วเช่นเดียวกันกับข้อมูลดิบ สำหรับข้อมูลอีก 3 ลักษณะที่เหลือคือ ข้อมูลที่มีการแปลงแบบอันดับ ข้อมูลดิบที่มีการแปลงโดย log และข้อมูลที่มีการแปลงโดย log ที่เป็นค่ามาตรฐานไม่สามารถจัดกลุ่มของลูกปัดแก้วได้อย่างชัดเจน

กมลชนก พานิชการ (2550) ได้ศึกษาการจำแนกกลุ่มของดินตามองค์ประกอบทางเคมีจากตัวอย่างดิน 58 ตัวอย่างที่ถูกเก็บมาจาก 3 จังหวัดได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม แยกขนาดของดินให้มีขนาดเล็กกว่า 212 และ 65 ไมครอน ประกอบด้วยตัวแปร 10 ตัวคือ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) ไทเทเนียม (Ti) คาร์บอน(C) และโซเดียม (Na) หน่วยที่วัดเป็นเปอร์เซ็นต์ของปริมาณธาตุในดินแต่ละตัวอย่างซึ่งเป็นข้อมูลในลักษณะ compositive data โดยใช้การวิเคราะห์องค์ประกอบหลักและการวิเคราะห์กลุ่มในข้อมูล 2 ลักษณะคือข้อมูลดิบ และข้อมูลที่แปลงด้วยวิธีของ Aitchison (1986)โดยในส่วนของการวิเคราะห์นั้นจะไม่พิจารณาตัวแปรออกซิเจนเนื่องจากเป็นตัวแปรที่เป็นธาตุหลักในดินและไม่พิจารณาตัวแปรโซเดียมเนื่องจากมีค่าเป็นศูนย์โดยส่วนใหญ่และจะพิจารณาเฉพาะชนิดของดินหลัก 3 ชนิดที่พบคือ ดินเหนียว ดินร่วนเหนียว และดินร่วนรวมเป็น 54 ตัวอย่าง ผลการวิเคราะห์พบว่าการวิเคราะห์องค์ประกอบหลักบนข้อมูลดิบ สำหรับดินขนาด 65 ไมครอน ผลที่ได้จากกราฟ 3 มิติของค่าองค์ประกอบหลักที่ 1,3 และ 4 สามารถจำแนกตัวอย่างของดินตามชนิดของดินทั้ง 3 ที่นำมาพิจารณาได้ค่อนข้างชัดเจนแสดงว่าองค์ประกอบทางเคมีมีส่วนในการอธิบายลักษณะการแบ่งกลุ่มของดิน สำหรับดินขนาด 212 ไมครอนพบว่าให้ผลการวิเคราะห์ลักษณะเดียวกับการวิเคราะห์บนดินขนาด 65 ไมครอน โดยผลที่ได้จากกราฟ 3 มิติของค่าองค์ประกอบหลักที่ 1,3 และ 4 สามารถจำแนกตัวอย่างของดินตามชนิดของดินทั้ง 3 ที่นำมาพิจารณาแต่ไม่ชัดเจนเหมือนกับการจำแนกบนดินขนาด 65 ไมครอน และการวิเคราะห์กลุ่มบนข้อมูลดิบ สำหรับดินขนาด 65 ไมครอนพบว่าสามารถจำแนกกลุ่มของตัวอย่างดินได้ 3 กลุ่มที่ระยะ

ประมาณ 5 แต่การแบ่งกลุ่มไม่สอดคล้องกับชนิดของดิน 3 ชนิดที่นำมาพิจารณา สำหรับดินขนาด 212 ไมครอนสามารถจำแนกกลุ่มของตัวอย่างดินได้ 4 กลุ่มที่ระยะประมาณ 3 แต่การแบ่งกลุ่มไม่สอดคล้องกับชนิดของดินเช่นเดียวกัน ดังนั้นจึงสามารถกล่าวได้ว่าการวิเคราะห์ข้อมูลดิบโดยการวิเคราะห์องค์ประกอบหลักสามารถแบ่งกลุ่มของข้อมูลซึ่งสอดคล้องกับชนิดของดินได้ดีกว่าการวิเคราะห์กลุ่ม สำหรับการวิเคราะห์ข้อมูลที่แปลงด้วยวิธีของ Aitchison (1986) จะทำการวิเคราะห์เฉพาะดินขนาด 65 ไมครอน (เนื่องจากผลการวิเคราะห์องค์ประกอบหลักและการวิเคราะห์กลุ่มบนข้อมูลดิบมีลักษณะคล้ายกันทั้งสองขนาดของดิน) พบว่าผลจากการวิเคราะห์องค์ประกอบหลักโดยการพลอตกราฟ 3 มิติของค่าองค์ประกอบหลักที่ 1,3 และ 4 สามารถแบ่งกลุ่มของตัวอย่างของดินได้แต่ไม่สอดคล้องกับชนิดของดินที่นำมาพิจารณาและในทำนองเดียวกันผลจากการวิเคราะห์กลุ่มก็สามารถแบ่งกลุ่มของตัวอย่างของดินได้แต่ไม่สอดคล้องกับชนิดของดินที่นำมาพิจารณาเช่นเดียวกัน จึงอาจกล่าวได้ว่าการแปลงข้อมูลช่วยให้สามารถมองเห็นการจำแนกกลุ่มของตัวอย่างดินแต่ไม่ได้ให้สารสนเทศที่เป็นประโยชน์มากขึ้นในการจำแนกกลุ่มที่สอดคล้องกับชนิดของดิน จากผลที่ได้ผู้วิจัยแนะนำให้มีการพิจารณาปัจจัยทางสิ่งแวดล้อมอื่น ๆ เช่น ลักษณะทางกายภาพของดิน (สี ค่า pH) ปริมาณโลหะหนัก หรือปริมาณจุลธาตุอาหารในดิน

บทที่ 3
วิธีดำเนินการวิจัย

การดำเนินการวิจัยเพื่อการจำแนกกลุ่มดินตามองค์ประกอบทางเคมี ภายใต้วิธีวิเคราะห์องค์ประกอบหลักบนเมทริกซ์ความแปรปรวนร่วม มีขั้นตอนดังนี้

1. ดำเนินการติดต่อขอข้อมูลโดยข้อมูลได้รับการอนุเคราะห์จาก ผศ.ดร.มยุวา อารีกิจเสรี โครงการศึกษาผลของปริมาณกรดฮิวมิกและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดินและโลหะที่เป็นพิษกรณีศึกษาภูมิภาคตะวันตกของประเทศไทย

2. ตรวจสอบข้อมูลเบื้องต้น

3. วิเคราะห์ข้อมูล

4. ศึกษาและตีความจากวิธีการวิเคราะห์ข้อมูลทั้ง 2 เทคนิค

โดยแต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

1. ข้อมูลที่นำมาศึกษาเป็นองค์ประกอบทางเคมีของดินซึ่งถูกวัดโดยการใช้เทคนิคทางกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope, SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์เรย์ที่วัดได้จากตัวอย่างวัตถุด้วยหัวตรวจวัด Energy Dispersive X-ray Micro-analysis (EDX) จากตัวอย่างดิน 58 ตัวอย่างที่ถูกเก็บมาจาก 3 จังหวัด ได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม นำมาผึ่งให้แห้งและแยกขนาดของดินให้มีขนาดที่ 212 ขนาดและ 65 ไมครอน ก่อนนำไปวิเคราะห์ด้วยเทคนิคทางกล้องจุลทรรศน์โดยใช้หัวตรวจวัด Secondary Electron Image Detector กับ Energy Dispersive X-ray Detector (SEM/EDX) ซึ่งในการวิจัยนี้จะทำการศึกษาตัวอย่างของดินขนาดที่ 65 ไมครอน ซึ่งเป็นขนาดที่ดูดซึมโลหะมากที่สุด โดยตัวแปรที่ถูกนำมาศึกษาได้แก่ ออกซิเจน (O) แมกนีเซียม(Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) ไทเทเนียม (Ti) คาร์บอน (C) และโซเดียม (Na) โดยค่าตอบสนองที่ถูกวัดขึ้นมาจากตัวแปรที่ถูกนำมาศึกษาเป็นค่าเปอร์เซ็นต์ซึ่งเป็นข้อมูลในลักษณะ compositive date

2. การตรวจสอบข้อมูลเบื้องต้นของตัวแปรที่ศึกษา เพื่อดูลักษณะการกระจายของข้อมูลและตรวจสอบค่าผิดปกติ (Outlier) จะทำการตรวจสอบข้อมูลเบื้องต้นของทั้ง 10 ตัวแปร โดยการตรวจสอบด้วย Box plot และพิจารณาจากค่าต่ำสุด สูงสุด ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (SD) ของทั้ง 10 ตัวที่นำมาทำการศึกษา

3. วิธีวิเคราะห์ข้อมูลในที่นี้จะทำการศึกษาด้วยเทคนิค 2 เทคนิค คือ

3.1 เทคนิคการวิเคราะห์องค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุด(Principal Component Analysis based on Maximum likelihood Estimates) โดยการพลอตค่าองค์ประกอบหลัก (Principal Component Score) 2 องค์ประกอบแรก โดยข้อมูลที่นำมาวิเคราะห์มี 2 ลักษณะ คือ

3.1.1 ข้อมูลดิบ (raw data)

3.1.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison (1986)

3.2 เทคนิคการวิเคราะห์องค์ประกอบหลักแบบแกร่งบนเมตริกซ์ความแปรปรวนร่วม(Robust Principal Component Analysis based on covariance matrix) แบ่งออกเป็นตัวประมาณค่า 2 แบบ คือ

3.2.1 ประมาณค่าเมทริกซ์ความแปรปรวนร่วมโดยตัวประมาณค่าแบบเครื่องหมาย ($\hat{\sum}_{Sn}$)ในสมการที่ (9) บนข้อมูลดังนี้

3.2.1.1 ข้อมูลดิบ (raw data)

3.2.1.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison (1986)

3.2.2 ประมาณค่าเมทริกซ์ความแปรปรวนร่วมโดยตัวประมาณค่าแบบอันดับ ($\hat{\sum}_{Rn}$)ในสมการที่ (12) บนข้อมูลดังนี้

3.2.2.1 ข้อมูลดิบ (raw data)

3.2.2.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison (1986)

แล้วจึงนำค่าประมาณเทริกซ์ความแปรปรวนร่วมที่ได้ไปทำการวิเคราะห์องค์ประกอบหลักต่อไป และทำการพลอตค่าองค์ประกอบหลัก (Principal Component Score) 2 องค์ประกอบแรก

4. ศึกษาและตีความจากการจำแนกกลุ่มที่ได้จากการวิเคราะห์องค์ประกอบทางเคมี ของดินจากวิธีวิเคราะห์องค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุดและแบบแกร่ง ทำโดยคำนวณค่าสถิติต่างๆ ที่เกี่ยวข้อง และพลอตกราฟแบบต่างๆ เพื่อตรวจสอบข้อมูลและกลุ่มของข้อมูล และการประมวลผลเพื่อการวิเคราะห์ข้อมูลใช้โปรแกรม SAS version 9.1 และ SPSS version 15.5

บทที่ 4
ผลการวิจัย

การศึกษาเรื่องการประยุกต์การวิเคราะห์องค์ประกอบหลักโดยใช้การประมาณแบบภาวะน่าจะเป็นสูงสุดและแบบแกร่งของเมทริกซ์ความแปรปรวนร่วมในการจำแนกกลุ่มดินตามองค์ประกอบทางเคมี: กรณีศึกษาภูมิภาคตะวันตกของประเทศไทย โดยใช้วิธีวิเคราะห์องค์ประกอบหลักภายใต้ตัวแปรที่เป็นองค์ประกอบหลักทางเคมีของดินขนาด 65 ไมครอน มีผลการวิจัยดังนี้

1. การตรวจสอบข้อมูลเบื้องต้นและสหสัมพันธ์ขององค์ประกอบทางเคมีของดิน

2. การวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด โดยใช้ข้อมูล 2 ลักษณะคือ

2.1 ข้อมูลดิบ (raw data)

2.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison

3. การวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมาย โดยใช้ข้อมูล 2 ลักษณะคือ

3.1 ข้อมูลดิบ (raw data)

3.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison

4. การวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ โดยใช้ข้อมูล 2 ลักษณะคือ

4.1 ข้อมูลดิบ (raw data)

4.2 ข้อมูลที่ทำการแปลงด้วยวิธีของ Aitchison

โดยมีรายละเอียดในแต่ละหัวข้อดังนี้

1. การตรวจสอบข้อมูลเบื้องต้นและสหสัมพันธ์ขององค์ประกอบทางเคมีของดิน

องค์ประกอบทางเคมีของดินที่ถูกเก็บมาจาก 3 จังหวัด ได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม โดยทำการเก็บตัวอย่างดิน 58 ตัวอย่าง เมื่อถูกนำมาวิเคราะห์องค์ประกอบทางเคมีโดยใช้เทคนิคทางกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope,

SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์เรย์ที่วัดได้จากตัวอย่างวัตถุด้วยหัวตรวจวัด Energy Dispersive X-ray Micro-analysis (EDX) ซึ่งในการวิจัยนี้ตัวแปรที่ถูกนำมาศึกษาประกอบด้วย 10 ตัวแปร คือ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) คาร์บอน (C) ไทเทเนียม (Ti) และโซเดียม (Na) หน่วยที่วัดข้อมูลของทุกตัวแปรวัดเป็นเปอร์เซ็นต์ของปริมาณธาตุในดิน โดยแสดงค่าสถิติเบื้องต้นไว้ในตารางที่ 3

ตารางที่ 3 ค่าสถิติเบื้องต้นขององค์ประกอบทางเคมีของดิน

ตัวแปร	ค่าต่ำสุด	ค่าสูงสุด	Mean	SD
ออกซิเจน (O)	50.39	56.91	54.46	1.56
แมกนีเซียม (Mg)	0.29	1.17	0.79	0.16
อลูมิเนียม (Al)	7.24	10.41	8.67	0.73
ซิลิคอน (Si)	22.71	31.75	26.82	2.36
โพแทสเซียม (K)	1.27	2.41	1.76	0.30
แคลเซียม (Ca)	0	1.33	0.58	0.24
เหล็ก (Fe)	2.78	5.08	3.84	0.56
คาร์บอน (C)	0	18.02	3.38	3.49
ไทเทเนียม (Ti)	0	0.53	0.37	0.10
โซเดียม (Na)	0	0.38	0.04	0.09

จากตารางที่ 3 ตัวแปรออกซิเจน (O) เป็นส่วนประกอบหลักของดินที่มีค่าเฉลี่ยเท่ากับ 54.46 (SD = 1.56) มีค่าต่ำสุดเท่ากับ 50.39 และค่าสูงสุดเท่ากับ 56.9 ตัวแปรซิลิคอน (Si) เป็นส่วนประกอบรองลงมาของดินที่มีค่าเฉลี่ยเท่ากับ 26.82 (SD = 2.36) มีค่าต่ำสุดเท่ากับ 22.71 และค่าสูงสุดเท่ากับ 31.75 ตัวแปรอลูมิเนียม (Al) เป็นส่วนประกอบรองลงมาของดินตามลำดับที่มีค่าเฉลี่ยเท่ากับ 8.67 (SD = 0.73) มีค่าต่ำสุดเท่ากับ 7.24 และค่าสูงสุดเท่ากับ 10.41 จนถึงตัวแปรโซเดียม (Na) เป็นส่วนประกอบที่มีค่าเฉลี่ยน้อยที่สุดของดินเท่ากับ 0.045 (SD = 0.09) มีค่าต่ำสุดเท่ากับ 0 และค่าสูงสุดเท่ากับ 0.38

ตารางที่ 4 ค่าสถิติเบื้องต้นขององค์ประกอบทางเคมีของดินโดยแยกตามชนิดของดิน

ตัวแปร	ดินเหนียว		ดินร่วนเหนียว		ดินร่วน		ดินร่วนปนทรายแป้ง		ดินร่วนปนทราย	
	(26 ตัวอย่าง)		(12 ตัวอย่าง)		(16 ตัวอย่าง)		(1 ตัวอย่าง)		(3 ตัวอย่าง)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ออกซิเจน (O)	54.21	1.60	54.84	1.75	54.48	1.53	**55.66**	-*	54.67	0.87
แมกนีเซียม (Mg)	0.75	0.17	0.79	0.11	**0.88**	0.15	0.87	-*	0.61	0.03
อลูมินัม (Al)	**8.97**	0.86	8.36	0.42	8.56	0.54	8.25	-*	8.13	0.84
ซิลิกอน (Si)	26.10	2.63	26.34	2.26	**26.52**	2.20	25.89	-*	24.67	1.85
โพแทสเซียม (K)	1.63	0.19	1.51	0.18	2.07	0.19	2.12	-*	**2.20**	0.19
แคลเซียม (Ca)	0.58	0.20	**0.77**	0.28	0.45	0.22	0.45	-*	0.61	0.10
เหล็ก (Fe)	**3.94**	0.54	3.92	0.59	3.77	0.58	3.63	-*	3.19	0.30
คาร์บอน (C)	**3.84**	4.19	2.84	2.72	2.69	2.89	2.69	-*	5.51	2.91
ไทเทเนียม (Ti)	0.35	0.13	0.37	0.08	**0.40**	0.07	0.35	-*	0.37	0.09
โซเดียม (Na)	0.01	0.03	0.05	0.09	**0.10**	0.13	**0.10**	-*	0.03	0.05

หมายเหตุ * มีจำนวนตัวอย่าง 1 หน่วย

จากตารางที่ 4 เมื่อพิจารณาองค์ประกอบทางเคมีของดินทั้ง 5 ชนิดที่นำมาพิจารณา พบว่าออกซิเจน (O) เป็นส่วนประกอบหลักที่มากที่สุด ซิลิกอน (Si) เป็นส่วนประกอบรองลงมา และโซเดียม (Na) เป็นส่วนประกอบที่น้อยที่สุด และเมื่อพิจารณาตามแต่ละธาตุที่เป็นองค์ประกอบทางเคมีของดินพบว่าออกซิเจน (O) และไทเทเนียม (Ti) พบมากที่สุดในดินร่วนปนทรายแป้ง แมกนีเซียม (Mg) ซิลิกอน (Si) และไทเทเนียม (Ti) พบมากที่สุดในดินร่วน อลูมินัม (Al) เหล็ก (Fe) และคาร์บอน (C) พบมากที่สุดในดินเหนียว โพแทสเซียม (K) พบมากที่สุดในดินร่วนปนทราย และแคลเซียม (Ca) พบมากที่สุดในดินร่วนเหนียว

การตรวจสอบการแจกแจงปกติของแต่ละตัวแปร พบว่ามีตัวแปรหลายตัวแปรที่ไม่ได้มีการแจกแจงแบบปกติ ดังแสดงในตารางที่ 5

ตารางที่ 5 ค่าสถิติสำหรับการทดสอบการแจกแจงปกติ

	Kolmogorov-Smirnov		
	ค่าสถิติ	**องศาอิสระ**	**p-value**
O	0.139	58	0.007
Mg	0.078	58	0.200
Al	0.113	58	0.065
Si	0.119	58	0.041
K	0.093	58	0.200
Ca	0.060	58	0.200
Fe	0.116	58	0.051
C	0.196	58	0.000
Ti	0.189	58	0.000
Na	0.433	58	0.000

จากตารางที่ 5 แสดงค่าสถิติและค่า p-value สำหรับการทดสอบการแจกแจงปกติด้วยการทดสอบของ Kolmogorov Smirnov พบว่าที่ระดับนัยสำคัญ 0.05 ตัวแปรแมกนีเซียม (Mg) อลูมินัม (Al) โพแทสเซียม (K) แคลเซียม (Ca) และเหล็ก (Fe) มีการแจกแจงแบบปกติ ส่วนตัวแปรออกซิเจน (O) ซิลิคอน (Si) คาร์บอน (C) ไทเทเนียม (Ti) และโซเดียม (Na) ไม่ได้มีการแจกแจงปกติ

การตรวจสอบความสัมพันธ์ของตัวแปร พบว่ามีตัวแปรหลายคู่ที่มีความสัมพันธ์เชิงเส้น ดังแสดงในตารางที่ 6

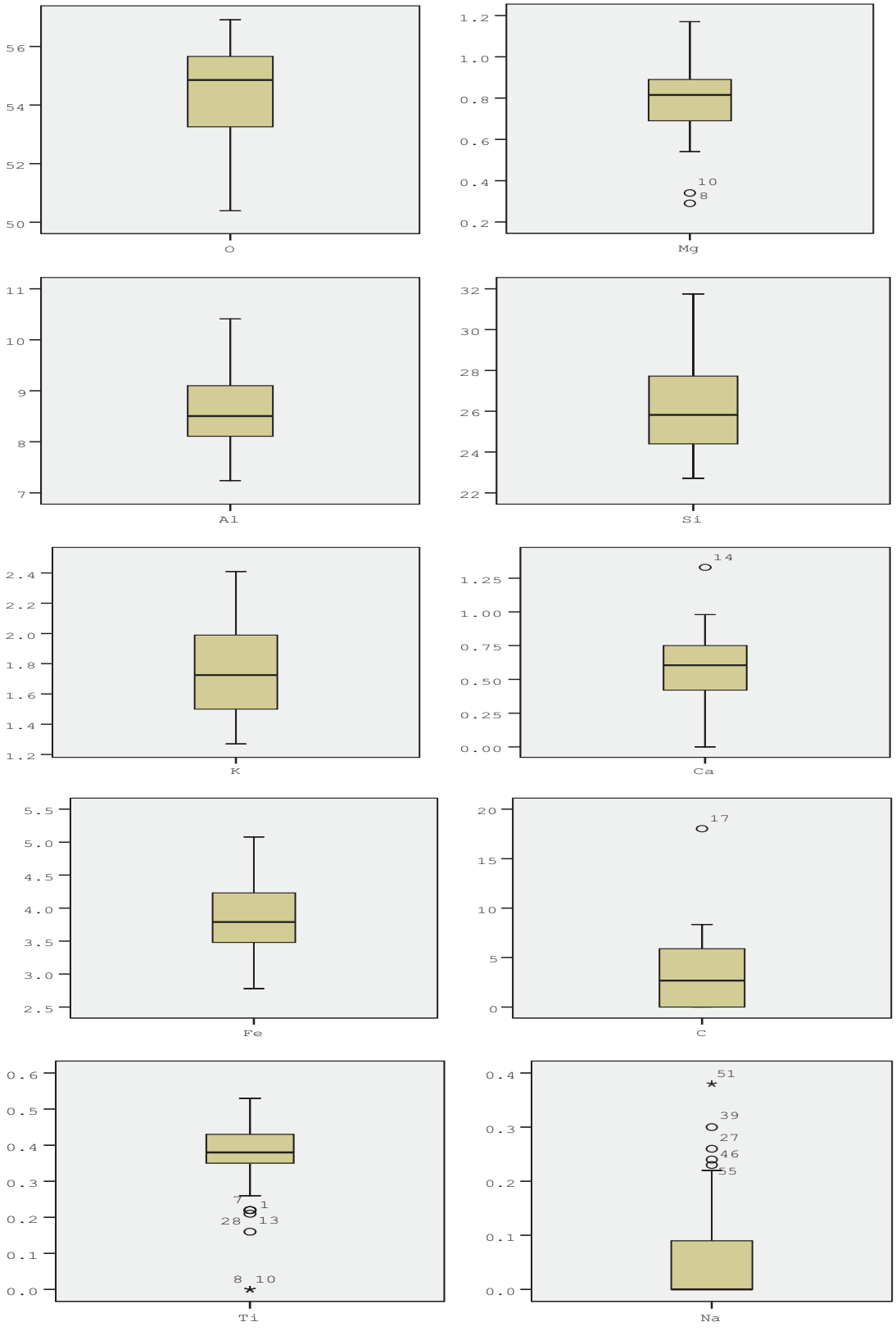
ตารางที่ 6 เมตริกซ์สหสัมพันธ์ของเพียร์สันขององค์ประกอบทางเคมีของดิน

	O	Mg	Al	Si	K	Ca	Fe	C	Ti	Na
O	1									
Mg	-0.016	1								
Al	**-0.459**	-0.028	1							
Si	**-0.563**	0.213	**0.744**	1						
K	-0.282	0.257	0.239	**0.400**	1					
Ca	0.061	-0.148	-0.066	-0.017	**-0.450**	1				
Fe	**-0.678**	0.307	**0.619**	**0.811**	0.213	-0.010	1			
C	0.193	**-0.369**	**-0.470**	**-0.749**	-0.297	-0.112	**-0.636**	1		
Ti	0.184	**0.556**	-0.122	-0.085	0.074	0.147	0.027	-0.058	1	
Na	0.124	**0.445**	-0.068	0.098	0.284	-0.301	0.107	-0.200	0.239	1

หมายเหตุ ค่าตัวเลขที่แสดงตัวหนามีระดับนัยสำคัญที่ $\alpha = 0.05$

จากตารางที่ 6 พบว่าที่ระดับนัยสำคัญ 0.05 ตัวแปรออกซิเจน (O) มีความสัมพันธ์ทางลบกับอลูมินัม (Al) ซิลิกอน (Si) และเหล็ก (Fe) โดยออกซิเจน (O) มีความสัมพันธ์ทางลบกับเหล็ก (Fe) มากที่สุด (r = -0.678) ตัวแปรแมกนีเซียม (Mg) มีความสัมพันธ์ทางบวกกับโซเดียม (Na) และไทเทเนียม (Ti) โดยแมกนีเซียม (Mg) มีความสัมพันธ์ทางบวกกับไทเทเนียม (Ti) มากที่สุด (r = 0.556) และนอกจากนี้แมกนีเซียม (Mg) มีความสัมพันธ์ทางลบคาร์บอน (C) (r = -0.369) ตัวแปรอลูมินัม (Al) มีความสัมพันธ์ทางบวกกับซิลิคอน (Si) และเหล็ก (Fe) โดยอลูมินัม (Al) มีความสัมพันธ์ทางบวกกับซิลิคอนมากที่สุด (r = 0.744) และนอกจากนี้อลูมินัม (Al) ยังมีความสัมพันธ์ทางลบกับออกซิเจน (O) และคาร์บอน (C) โดยอลูมินัม (Al) มีความสัมพันธ์ทางลบกับคาร์บอน (C) มากที่สุด ตัวแปรซิลิคอน (Si) มีความสัมพันธ์ทางบวกกับอลูมินัม (Al) โพแทสเซียม (K) และเหล็ก (Fe) โดยซิลิคอน (Si) มีความสัมพันธ์ทางบวกกับเหล็ก (Fe) มากที่สุด (r = 0.811) ในขณะที่มีความสัมพันธ์ทางลบกับออกซิเจน (O) และคาร์บอน (C) โดยมีความสัมพันธ์ทางลบกับคาร์บอน (C) มากที่สุด (r = -0.749) และตัวแปรโพแทสเซียม (K) มีความสัมพันธ์ทางบวกกับซิลิคอน (Si) ตามที่ได้กล่าวมาแล้ว ตัวแปรโพแทสเซียม (K) ยังมีความสัมพันธ์ทางลบกับแคลเซียม (Ca) (r = -0.450)

การตรวจสอบค่านอกกลุ่มเบื้องต้น พบว่าตัวแปรหลายตัวมีค่านอกกลุ่มปะปนอยู่ ดังแสดงในภาพที่ 6 โดยสัญลักษณ์ o แทนค่านอกกลุ่มปานกลาง (mild outliers) และสัญลักษณ์ * แทนค่านอกกลุ่มสุดปลาย (extreme outliers)



ภาพที่ 6 Boxplot ของตัวแปรที่เป็นองค์ประกอบทางเคมีของดิน

จากภาพที่ 6 แสดงให้เห็นว่ามีจำนวน 5 ตัวแปรที่มีค่านอกกลุ่มปะปนอยู่ด้วยคือตัวแปรแมกนีเซียม (Mg) แคลเซียม (Ca) คาร์บอน (C) ไทเทเนียม (Ti) และ โซเดียม (Na) โดยตัวแปรแมกนีเซียม (Mg) มีค่านอกกลุ่ม 2 ค่าคือค่าจากตัวอย่างที่ 8 และ 10 ตัวแปรแคลเซียม (Ca) และคาร์บอน (C) มีค่านอกกลุ่ม 1 ค่าคือค่าจากตัวอย่างที่ 14 และ 17 ตามลำดับ ตัวแปรไทเทเนียม (Ti) มีค่านอกกลุ่ม 6 ค่าคือค่าจากตัวอย่างที่ 1, 7, 8, 10, 13 และ 28 ตัวแปรโซเดียม (Na) มีค่านอกกลุ่ม 5 ค่าคือค่าจากตัวอย่างที่ 27, 46, 39, 51 และ 55

จากผลการวิเคราะห์ข้อมูลด้วย Boxplot พบว่าข้อมูลองค์ประกอบทางเคมีของดินมีค่านอกกลุ่ม (outlier) ในบางตัวแปร แต่ผู้วิจัยต้องการทดสอบโดยใช้การวิเคราะห์องค์ประกอบหลักแบบแกร่ง (Robust Principal Component Analysis: RPCA) ซึ่งน่าจะเหมาะกับข้อมูลที่มีค่านอกกลุ่มที่ไม่ผิดปกติอย่างรุนแรง ดังนั้นในงานวิจัยนี้จึงไม่มีการตัดค่านอกกลุ่มในการวิเคราะห์ขั้นต่อไป

2. การวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด

ในส่วนของการวิเคราะห์องค์ประกอบหลักจะพิจารณาข้อมูลองค์ประกอบทางเคมีของดินจำนวน 8 ตัวแปรได้แก่ แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) คาร์บอน (C) และไทเทเนียม (Ti) แต่จะไม่พิจารณาตัวแปรออกซิเจน (O) เนื่องจากเป็นตัวแปรที่เป็นฐานหลักที่พบในดินมีปริมาณในช่วง 50%-57% และไม่พิจารณาตัวแปรโซเดียม (Na) เนื่องจากมีค่าเป็นศูนย์ส่วนใหญ่ พบเฉพาะดินที่อยู่ใกล้ทะเล โดยทำการวิเคราะห์องค์ประกอบหลักจากตัวอย่างจำนวน 58 ตัวอย่างซึ่งถูกจำแนกเป็นชนิดของดิน 5 ชนิดตามเนื้อดินดังแสดงในตารางที่ 7

ตารางที่ 7 การจำแนกตัวอย่างของดินจำนวน 58 ตัวอย่างตามชนิดของดิน

ชนิดของดิน (soiltype)	**จำนวนตัวอย่าง**
ดินเหนียว (Clay)	26
ดินร่วนเหนียว (Clay loam)	12
ดินร่วน (Medium loam)	16
ดินร่วนปนทรายแป้ง (Silty loam)	1
ดินร่วนปนทราย (Sand loam)	3

โดยวิเคราะห์องค์ประกอบหลักข้อมูล 2 ลักษณะดังนี้

2.1 ข้อมูลดิบ (raw data)

ผลการวิเคราะห์ข้อมูลดิบโดยการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดแสดงในตารางที่ 8-9 โดยตารางที่ 8 แสดงค่าความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ และตารางที่ 9 แสดงค่าสัมประสิทธิ์ขององค์ประกอบหลัก 5 องค์ประกอบแรกจาก 8 องค์ประกอบหลักที่สามารถคำนวณได้

ตารางที่ 8 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ

	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	0.0256							
Al	-0.0032	0.5386						
Si	0.0806	1.2885	5.5738					
K	0.0125	0.0532	0.2869	0.0923				
Ca	-0.0057	-0.0118	-0.0097	-0.0332	0.0588			
Fe	0.0277	0.2562	1.0798	0.0364	-0.0014	0.3182		
C	-0.2059	-1.2013	-6.1644	-0.3147	-0.0951	-1.2512	12.1553	
Ti	0.0090	-0.0091	-0.0207	0.0023	0.0036	0.0015	-0.0204	0.0103

จากตารางที่ 8 ค่าความแปรปรวนของคาร์บอน (C) มีค่าสูงสุดเท่ากับ 12.1553 รองลงมาเป็นซิลิคอน (Si) เท่ากับ 5.5738 และน้อยที่สุดเป็นไทเทเนียม (Ti) เท่ากับ 0.0103 ความแปรปรวนร่วมระหว่างซิลิคอน (Si) กับคาร์บอน (C) มีค่าสูงสุดเท่ากับ -6.1644 รองลงมาเป็นอลูมินัม (Al) กับซิลิคอน (Si) เท่ากับ 1.2885 และน้อยสุดเป็นไทเทเนียม (Ti) กับเหล็ก (Fe) เท่ากับ 0.0015

ตารางที่ 9 ค่าสัมประสิทธิ์ขององค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ

	PC1	PC2	PC3	PC4	PC5
Mg	-0.013	-0.021	-0.085	0.032	0.213
Al	-0.109	0.272	**0.946**	-0.084	0.051
Si	**-0.513**	**0.787**	-0.283	-0.084	-0.163
K	-0.026	0.040	-0.120	**-0.555**	**0.538**
Ca	-0.005	-0.031	0.009	**0.289**	**-0.553**
Fe	-0.103	0.146	-0.011	**0.770**	**0.572**
C	**0.845**	**0.532**	-0.056	0.017	-0.005
Ti	0.000	-0.014	-0.013	0.031	0.062
Eigenvalue	16.223	2.067	0.216	0.119	0.092
Cumulative	0.864	0.974	0.986	0.992	0.997

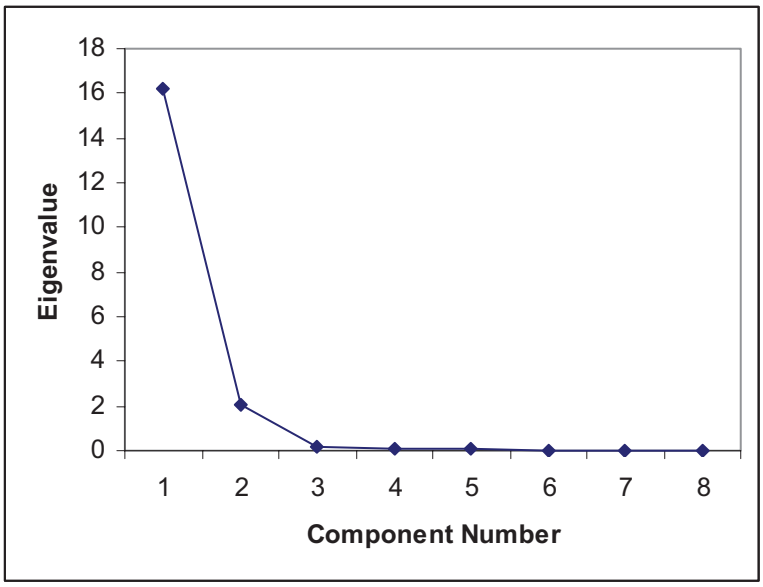
องค์ประกอบหลักแรกมีค่าน้ำหนักสูงที่ตัวแปรซิลิคอน (Si) และคาร์บอน (C) ดังนั้น องค์ประกอบแรกจึงเป็นตัวแทนของทั้งสองตัวแปรดังกล่าว โดยถ้าค่าขององค์ประกอบแรกมีค่าสูง แสดงว่าเป็นกลุ่มดินที่มีค่าคาร์บอน (C) สูงเนื่องจากน้ำหนักมีค่าเป็นบวก ส่วนค่าองประกอบแรกที่มีค่าต่ำแสดงว่าเป็นกลุ่มดินที่มีค่าซิลิคอน (Si) สูงเนื่องจากน้ำหนักมีค่าเป็นลบ ถ้าเป็นกลุ่มดินที่มีค่าซิลิคอน (Si) และคาร์บอน (C) พอกันจะทำให้ได้ค่าองค์ประกอบหลักแรกกลางๆ

องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) เช่นเดียวกับในองค์ประกอบแรกแต่อธิบายไปในทิศทางเดียวกัน องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรอลูมินัม (Al) เนื่องจากมีค่าน้ำหนักสูงที่ตัวแปรอลูมินัม (Al) เพียงตัวเดียว องค์ประกอบที่สี่และห้าเป็นตัวแทนของตัวแปรเหล็ก (Fe) โพแทสเซียม (K) และแคลเซียม (Ca)

ตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลย เนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล

องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 86.4% องค์ประกอบหลักที่ 2 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 11% องค์ประกอบหลักที่ 3 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดความจากแตกต่างของดิน 1.2% สำหรับองค์ประกอบหลักที่ 4, 5, 6, 7 และ 8 เป็นตัวอธิบายความแตกต่างใน

เรื่องของตัวแปรต่าง ๆ ออกไป ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 99.7%



ภาพที่ 7 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ

การทดสอบการเท่ากันของค่าไอเกนทำการทดสอบโดยใช้สถิติทดสอบ u (Rencher 1996 : 436) ดังนี้

$$u = \left(n - \frac{2p+11}{6} \right) \left(k \ln \bar{\hat{\lambda}} - \sum_{i=p-k+1}^{p} \ln \hat{\lambda}_i \right) \qquad (14)$$

เมื่อ n คือจำนวนหน่วยตัวอย่าง

p คือ จำนวนตัวแปร

k คือ จำนวนค่าไอเกนขององค์ประกอบหลักสุดท้ายจากตัวอย่าง

$\bar{\hat{\lambda}}$ คือ ค่าเฉลี่ยผลรวมของค่าไอเกนขององค์ประกอบหลักสุดท้ายจากตัวอย่าง

$\hat{\lambda}_i$ คือ ค่าไอเกนขององค์ประกอบหลักที่ i ของตัวอย่าง

โดยมีสมมติฐานในการทดสอบคือ

$$H_{0k} : \lambda_{k-p+1} = \lambda_{k-p+2} = ... = \lambda_p \quad (15)$$

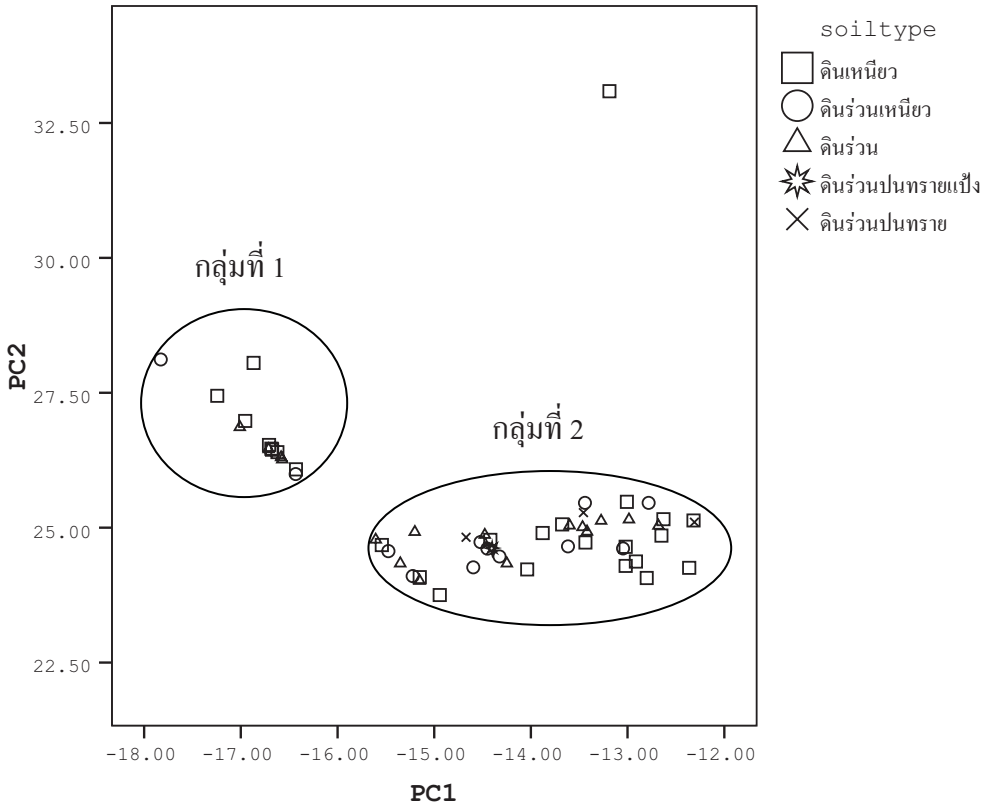
เมื่อ $\lambda_1, \lambda_2, ..., \lambda_p$ คือค่าไอเกนขององค์ประกอบหลักที่ i จากประชากร และปฏิเสธ H_0 ถ้า $u \geq \chi^2_{\alpha,\nu}$ เมื่อองศาอิสระ $\nu = \frac{1}{2}(k-1)(k+2)$

จากการทดสอบพบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน ดังแสดงในตารางที่ 10

ตารางที่ 10 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ

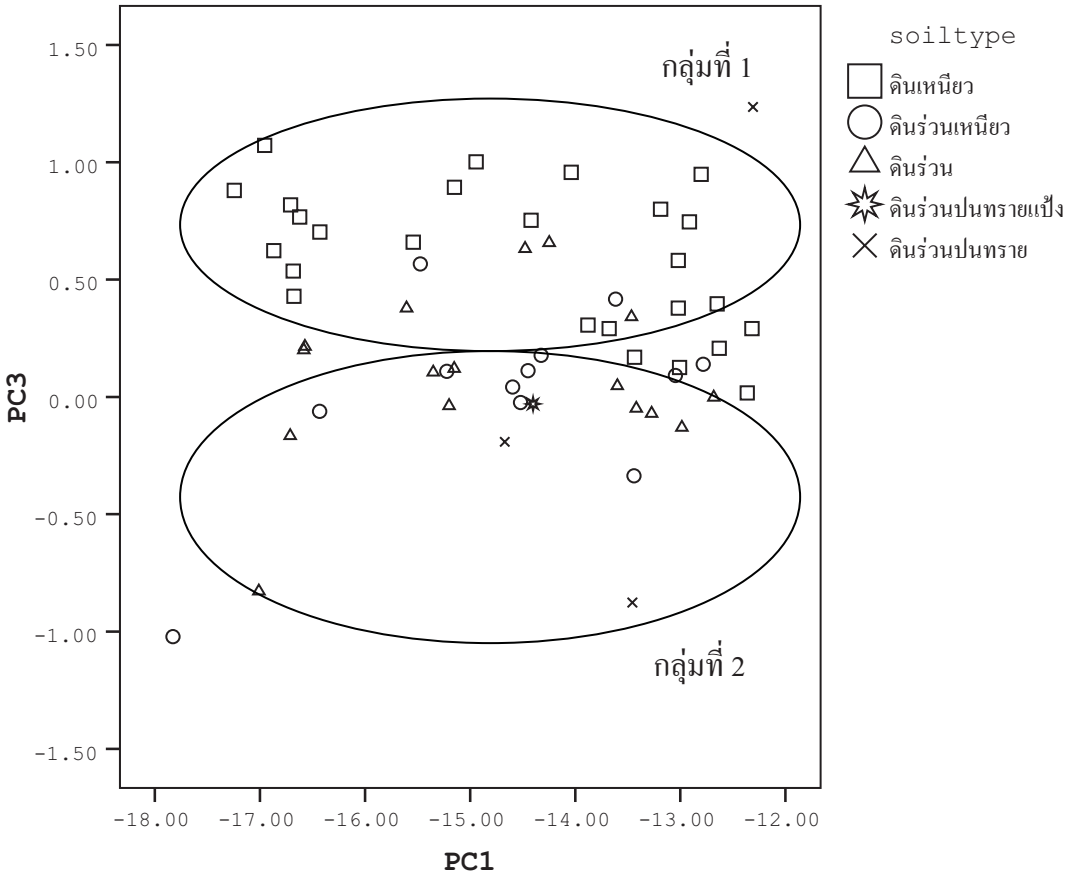
Eigenvalue	k	องศาอิสระ	u	$\chi^2_{0.05}$
16.223	8	35	1182.605	49.766
2.067	7	27	588.462	40.113
0.216	6	20	197.156	31.410
0.119	5	14	140.426	23.685
0.092	4	9	104.850	16.919
0.032	3	5	45.138	11.071
0.021	2	2	29.176	5.991
0.004	1			

แม้ว่าผลจากการทดสอบการเท่ากันของค่าไอเกนจะแสดงให้เห็นว่าค่าไอเกนแตกต่างกันอย่างมีนัยสำคัญและมีผลต่อการใช้จำนวนองค์ประกอบหลักที่มากขึ้นซึ่งจะทำให้สามารถอธิบายความแปรปรวนของข้อมูลได้มากขึ้น แต่เนื่องจากการวิเคราะห์องค์ประกอบหลักนั้น มีจุดมุ่งหมายที่จะลดจำนวนข้อมูลตัวแปรของตัวแปรเดิมและแทนที่ด้วยตัวแปรใหม่ที่ได้จากการวิเคราะห์องค์ประกอบหลัก โดยยังคงรักษาสารสนเทศของข้อมูลเดิม โดยปกติการวิเคราะห์องค์ประกอบหลักบนเมตริกซ์ความแปรปรวนร่วมจะใช้จำนวนองค์ประกอบหลักที่สามารถอธิบายความแปรปรวนของข้อมูลได้ประมาณ 80% (Rencher 1996 : 436) และเมื่อพิจารณาร่วมกับภาพที่ 7 ซึ่งแสดง Scree plot ของค่าไอเกนองค์ประกอบหลักจากข้อมูลดิบ ดังนั้นในกรณีนี้อาจใช้องค์ประกอบหลัก 3 แรกก็เพียงพอที่จะอธิบายความแปรปรวนของข้อมูลได้ถึง 98.6% โดยยังคงรักษาสารสนเทศของข้อมูลเดิมไว้ได้



ภาพที่ 8 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ

จากภาพที่ 8 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะเห็นการจำแนกกลุ่มของดินได้ 2 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน และจะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 อยู่กลางๆ มีส่วนประกอบของซิลิคอน (Si) สูงและคาร์บอน (C) ปนอยู่บ้าง ในขณะที่กลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าสูงกับค่าองค์ประกอบหลักที่ 2 มีค่าต่ำมีส่วนประกอบของคาร์บอน (C) สูงและซิลิคอน (Si) ปนอยู่เล็กน้อย



ภาพที่ 9 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลดิบ

จากภาพที่ 9 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะเห็นการจำแนกกลุ่มของดินได้ 2 กลุ่ม กลุ่มที่ 1 เป็นการรวมกลุ่มของดิน 3 ชนิดคือดินเหนียว ดินร่วนเหนียวและดินร่วน ซึ่งชนิดของดินในกลุ่มนี้ส่วนใหญ่เป็นดินเหนียว มีค่าองค์ประกอบหลักที่ 1 มีค่าอยู่กลาง ๆ กับค่าองค์ประกอบหลักที่ 3 ที่มีค่าสูง ที่มีส่วนประกอบที่มีอลูมินัม (Al) สูง ซิลิคอน (Si) และคาร์บอน (C) ปนอยู่ปานกลาง กลุ่มที่ 2 เป็นการรวมกลุ่มของดินอีก 4 ชนิดคือดินร่วนเหนียว ดินร่วน ดินร่วนปนทรายแป้งและดินร่วนปนทรายและมีตัวอย่างของดินที่เป็นดินร่วนเหนียวปนอยู่ด้วยเพียงเล็กน้อย ซึ่งทำให้ไม่สามารถที่จะจำแนกชนิดของดินตามธาตุที่เป็นองค์ประกอบของดินในแต่ละชนิดได้อย่างชัดเจน แต่อาจกล่าวได้ว่าดินทั้ง 4 ชนิดนี้เป็นกลุ่มชนิดของดินที่มีอะลูมินัม (Al) ต่ำกว่าดินเหนียว โดยมีค่า

องค์ประกอบหลักที่ 1 และ 3 อยู่กลาง ๆ ซึ่งจะเห็นได้ว่าตัวอย่างของดินในกลุ่มที่ 2 นี้มีค่าองค์ประกอบหลักที่ 3 ต่ำกว่าตัวอย่างของดินในกลุ่มที่ 1

2.2 ข้อมูลแปลงด้วยวิธีของ Aitchison

ผลการวิเคราะห์ข้อมูลแปลงด้วยวิธีของ Aitchison โดยการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดแสดงในตารางที่ 11-12 โดยตารางที่ 11 แสดงค่าความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง และตารางที่ 12 แสดงค่าสัมประสิทธิ์ขององค์ประกอบหลัก 5 องค์ประกอบแรกจาก 8 องค์ประกอบหลักที่สามารถคำนวณได้

ตารางที่ 11 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	0.0567							
Al	-0.0021	0.0069						
Si	0.0021	0.0055	0.0078					
K	0.0062	0.0038	0.0063	0.0292				
Ca	-0.0201	-0.0027	-0.0043	-0.0403	0.2419			
Fe	0.0069	0.0074	0.0101	0.0056	-0.0084	0.0212		
C	-0.0519	-0.0403	-0.0592	-0.0456	-0.0087	-0.0791	0.5953	
Ti	0.0271	-0.0006	-0.0012	0.0048	0.0046	0.0041	-0.0174	0.0622

จากตารางที่ 11 ค่าความแปรปรวนของคาร์บอน (C) มีค่าสูงสุดเท่ากับ 0.5953 รองลงมาเป็นแคลเซียม (Ca) เท่ากับ 0.2419 และน้อยที่สุดเป็นอลูมินัม (Al) เท่ากับ 0.0069 ความแปรปรวนร่วมระหว่างเหล็ก (Fe) กับคาร์บอน (C) มีค่าสูงสุดเท่ากับ -0.0791 รองลงมาเป็นซิลิคอน (Si) กับคาร์บอน (C) เท่ากับ -0.0592 และน้อยสุดเป็นไทเทเนียม (Ti) กับอลูมินัม (Al) เท่ากับ -0.0006

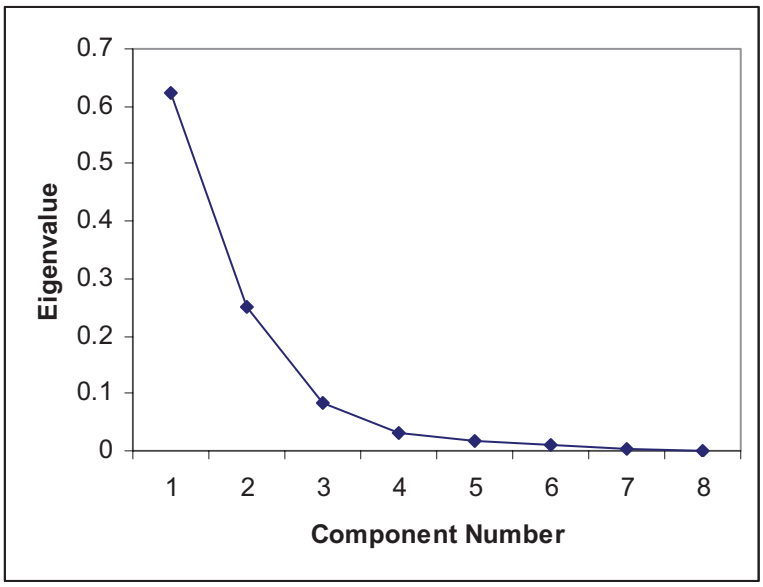
ตารางที่ 12 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

	PC1	PC2	PC3	PC4	PC5
Mg	-0.095	-0.089	**0.618**	**-0.735**	0.158
Al	-0.066	-0.003	-0.070	0.133	-0.067
Si	-0.098	-0.007	-0.062	0.049	0.012
K	-0.081	-0.168	0.023	0.280	**0.908**
Ca	0.040	**0.979**	0.054	-0.024	0.169
Fe	-0.133	-0.020	-0.002	0.075	-0.292
C	**0.975**	-0.066	0.077	-0.001	0.035
Ti	-0.036	0.012	**0.774**	**0.596**	-0.174
Eigenvalue	0.625	0.250	0.083	0.031	0.018
Cumulative	0.611	0.857	0.938	0.968	0.986

องค์ประกอบหลักแรกมีค่าน้ำหนักสูงที่ตัวแปรคาร์บอน (C) ดังนั้นองค์ประกอบแรกจึงเป็นตัวแทนของตัวแปรคาร์บอน (C) องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรแคลเซียม (Ca) องค์ประกอบหลักที่สามและสี่เป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) เนื่องจากมีค่าน้ำหนักสูงที่ตัวแปรที่ตัวแปรทั้งสอง โดยองค์ประกอบหลักที่สามอธิบายน้ำหนักของตัวแปรทั้งสองไปในทิศทางเดียวกัน ส่วนองค์ประกอบที่สี่อธิบายน้ำหนักของตัวแปรทั้งสองไปในทิศทางตรงกันข้าม องค์ประกอบหลักที่ห้าเป็นตัวแทนของตัวแปรโพแทสเซียม (K)

ตัวแปรอลูมินัม (Al) ซิลิคอน (Si) และเหล็ก (Fe) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลย เนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล

องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 61.1% องค์ประกอบหลักที่ 2 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 24.5% องค์ประกอบหลักที่ 3 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดความจากแตกต่างของดิน 8.1% องค์ประกอบหลักที่ 4 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 3.1% สำหรับองค์ประกอบหลักที่ 5, 6, 7 และ 8 เป็นตัวอธิบายความแตกต่างในเรื่องของตัวแปรต่าง ๆ ออกไป ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 93.8% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6%



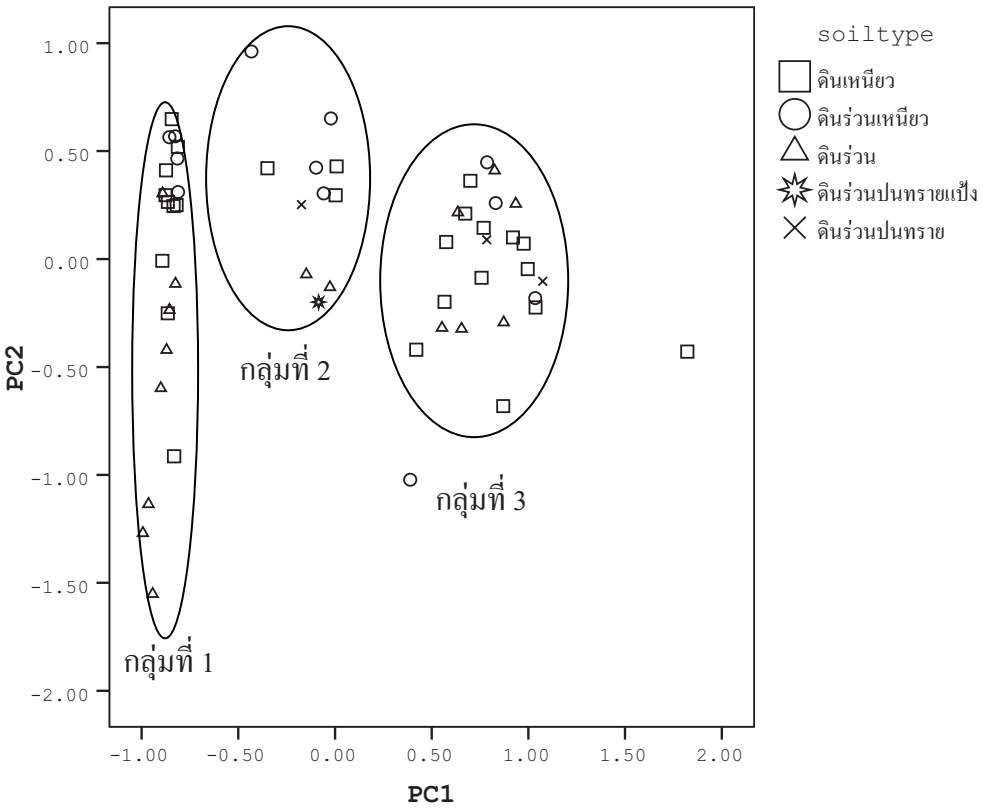
ภาพที่ 10 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

การทดสอบการเท่ากันของค่าไอเกนทำการทดสอบโดยใช้สถิติทดสอบในสมการที่ (14) ภายใต้สมมติฐานในสมการที่ (15) พบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน ดังแสดงในตารางที่ 13

ตารางที่ 13 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

Eigenvalue	k	องศาอิสระ	u	$\chi^2_{0.05}$
0.625	8	35	683.690	49.766
0.250	7	27	464.617	40.113
0.083	6	20	273.746	31.410
0.031	5	14	165.204	23.685
0.018	4	9	117.407	16.919
0.011	3	5	75.831	11.071
0.003	2	2	19.475	5.991
0.001	1			

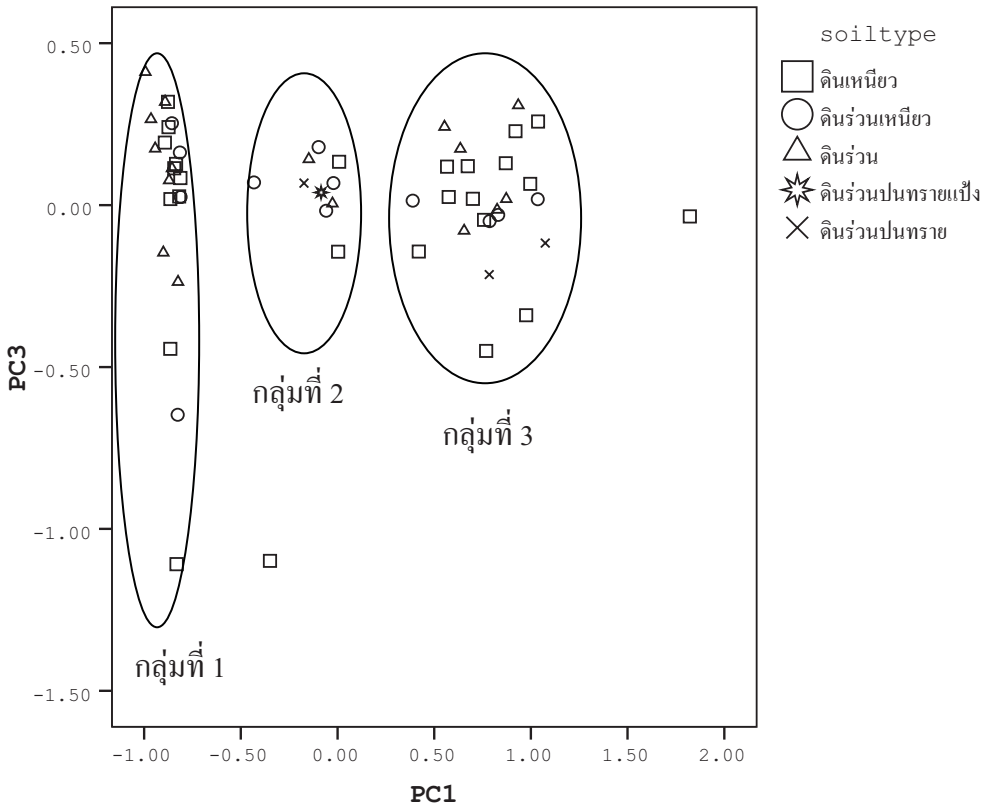
ผลจากการทดสอบการเท่ากันของค่าไอเกนจากข้อมูลแปลงแสดงให้เห็นว่าค่าไอเกนแตกต่างกันอย่างมีนัยสำคัญซึ่งให้ผลเช่นเดียวกันกับผลการทดสอบการเท่ากันของค่าไอเกนจากข้อมูลดิบ แต่เมื่อพิจาณาถึงวัตถุประสงค์ของการวิเคราะห์องค์ประกอบหลักร่วมกับการพิจารณาภาพที่ 10 ซึ่งแสดง Scree plot ของค่าไอเกนองค์ประกอบหลักของดินจากข้อมูลแปลง ดังนั้นในกรณีนี้อาจใช้องค์ประกอบหลัก 3 แรกก็เพียงพอที่จะอธิบายความแปรปรวนของข้อมูลได้ถึง 93.8%



ภาพที่ 11 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

จากภาพที่ 11 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะเห็นการจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน และจะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 มีค่าอยู่ในช่วงปานกลางมีส่วนประกอบของแคลเซียม

(Ca) ปนอยู่ปานกลางและคาร์บอน (C) ปนอยู่บ้าง และกลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 มีค่าสูงมีส่วนประกอบของแคลเซียม (Ca) สูงและคาร์บอน (C) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 กับค่าองค์ประกอบหลักที่ 2 ที่มีค่าอยู่กลางๆ มีส่วนประกอบของคาร์บอน (C) และแคลเซียม (Ca) สูง



ภาพที่ 12 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

จากภาพที่ 12 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะเห็นการจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดินเช่นเดียวกัน ซึ่งผลที่ได้นั้นใกล้เคียงกับผลที่ได้จากภาพที่ 11 และจะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้าง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำแต่มากกว่ากลุ่มที่ 1 กับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและ

คาร์บอน (C) ปนอยู่บ้างแต่มากกว่ากลุ่มที่ 1 และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าอยู่กลางๆ กับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของคาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) สูง

3. การวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมาย

การวิเคราะห์องค์ประกอบหลักจะพิจารณาข้อมูลองค์ประกอบทางเคมีของดินจำนวน 8 ตัวแปรได้แก่ แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) คาร์บอน (C) และไทเทเนียม (Ti) จากตัวอย่างจำนวน 58 ตัวอย่าง ซึ่งจากการตรวจสอบเบื้องต้นพบว่ามีจำนวน 4 ตัวแปรที่มีค่านอกกลุ่มปะปนอยู่ด้วยคือตัวแปรแมกนีเซียม (Mg) แคลเซียม (Ca) คาร์บอน (C) และไทเทเนียม (Ti) ดังนั้นในการวิเคราะห์ข้อมูลจะทำการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมาย

3.1 ข้อมูลดิบ

ผลการวิเคราะห์ข้อมูลดิบโดยการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายแสดงในตารางที่ 14-15 โดยตารางที่ 14 แสดงค่าความแปรปรวน และความแปรปรวนร่วมด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ และตารางที่ 15 แสดงค่าสัมประสิทธิ์ขององค์ประกอบหลัก 5 องค์ประกอบแรกจาก 8 องค์ประกอบหลักที่สามารถคำนวณได้

ตารางที่ 14 ความแปรปรวน และความแปรปรวนร่วมของธาตุ ๆ ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ

	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	0.0074							
Al	-0.0065	0.0685						
Si	0.0023	0.0469	0.2928					
K	0.0034	0.0074	0.0166	0.0445				
Ca	-0.0034	-0.0065	-0.0006	-0.0204	0.0232			
Fe	0.0044	0.0107	0.0446	-0.0095	0.0055	0.0377		
C	-0.0129	-0.0706	-0.3147	-0.0158	-0.0121	-0.0573	0.5249	
Ti	-0.0002	0.0007	0.0006	0.0006	0.0009	0.0006	-0.0038	0.0009

จากตารางที่ 14 ค่าความแปรปรวนของคาร์บอน (C) มีค่าสูงสุดเท่ากับ 0.5249 รองลงมาเป็นซิลิกอน (Si) เท่ากับ 0.2419 และน้อยที่สุดเป็นไทเทเนียม (Ti) เท่ากับ -0.0002 ความแปรปรวนร่วมระหว่างเหล็ก (Fe) กับคาร์บอน (C) มีค่าสูงสุดเท่ากับ -0.0573 รองลงมาเป็นอลูมินัม(Al) กับซิลิคอน (Si) เท่ากับ 0.0469 และน้อยสุดเป็นอลูมินัม (Al) กับไทเทเนียม (Ti) เท่ากับ -0.0002

ตารางที่ 15 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ

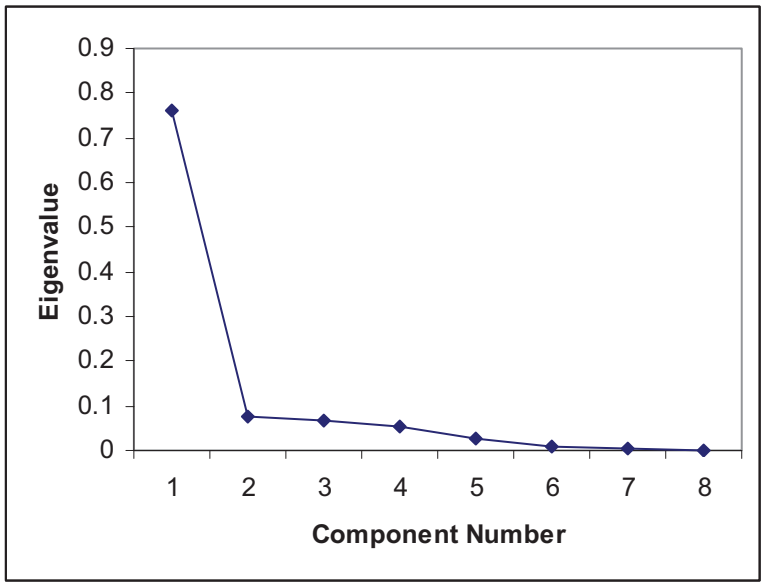
	PC1	PC2	PC3	PC4	PC5
Mg	0.015	-0.058	-0.029	-0.161	0.307
Al	0.122	0.108	**0.692**	**0.673**	-0.090
Si	**0.564**	**0.736**	-0.328	0.077	-0.135
K	0.031	0.303	**0.493**	**-0.545**	0.246
Ca	0.011	-0.256	-0.325	0.229	-0.277
Fe	0.100	-0.028	-0.212	0.367	**0.861**
C	**-0.809**	**0.533**	-0.137	0.180	0.010
Ti	0.005	-0.021	0.011	-0.000	0.008
Eigenvalue	0.763	0.077	0.066	0.054	0.026
Cumulative	0.763	0.840	0.906	0.960	0.986

องค์ประกอบหลักแรกมีค่าน้ำหนักสูงที่ตัวแปรซิลิคอน (Si) และคาร์บอน (C) ดังนั้นองค์ประกอบแรกจึงเป็นตัวแทนของทั้งสองตัวแปรดังกล่าว โดยถ้าค่าขององค์ประกอบแรกมีค่าสูงแสดงว่าเป็นกลุ่มดินที่มีค่าซิลิคอน (Si) สูงเนื่องจากน้ำหนักมีค่าเป็นบวก ส่วนค่าองประกอบแรกที่มีค่าต่ำแสดงว่าเป็นกลุ่มดินที่มีค่าคาร์บอน (C) สูงเนื่องจากน้ำหนักมีค่าเป็นลบ ถ้าเป็นกลุ่มดินที่มีค่าซิลิคอน (Si) และคาร์บอน (C) พอกันจะทำให้ได้ค่าองค์ประกอบหลักแรกกลางๆ

องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) เช่นเดียวกับในองค์ประกอบแรกแต่อธิบายไปในทิศทางเดียวกัน องค์ประกอบหลักที่สามและสี่เป็นตัวแทนของตัวแปรอลูมินัม (Al) และโพแทสเซียม (K) เนื่องจากมีค่าน้ำหนักสูงที่ตัวแปรทั้งสอง โดยองค์ประกอบหลักที่สามอธิบายน้ำหนักของตัวแปรทั้งสองไปในทิศทางเดียวกัน ส่วนองค์ประกอบที่สี่อธิบายน้ำหนักของตัวแปรทั้งสองไปในทิศทางตรงกันข้าม องค์ประกอบที่ห้าเป็นตัวแทนของตัวแปรเหล็ก (Fe)

ตัวแปรแมกนีเซียม (Mg) แคลเซียม (Ca) และไทเทเนียม (Ti) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล

องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 76.3% องค์ประกอบหลักที่ 2 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 7.7% องค์ประกอบหลักที่ 3 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 6.6% องค์ประกอบหลักที่ 4 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 5.4% สำหรับองค์ประกอบหลักที่ 5, 6, 7 และ 8 เป็นตัวอธิบายความแตกต่างในเรื่องของตัวแปรต่าง ๆ ออกไป ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 90.6% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6%



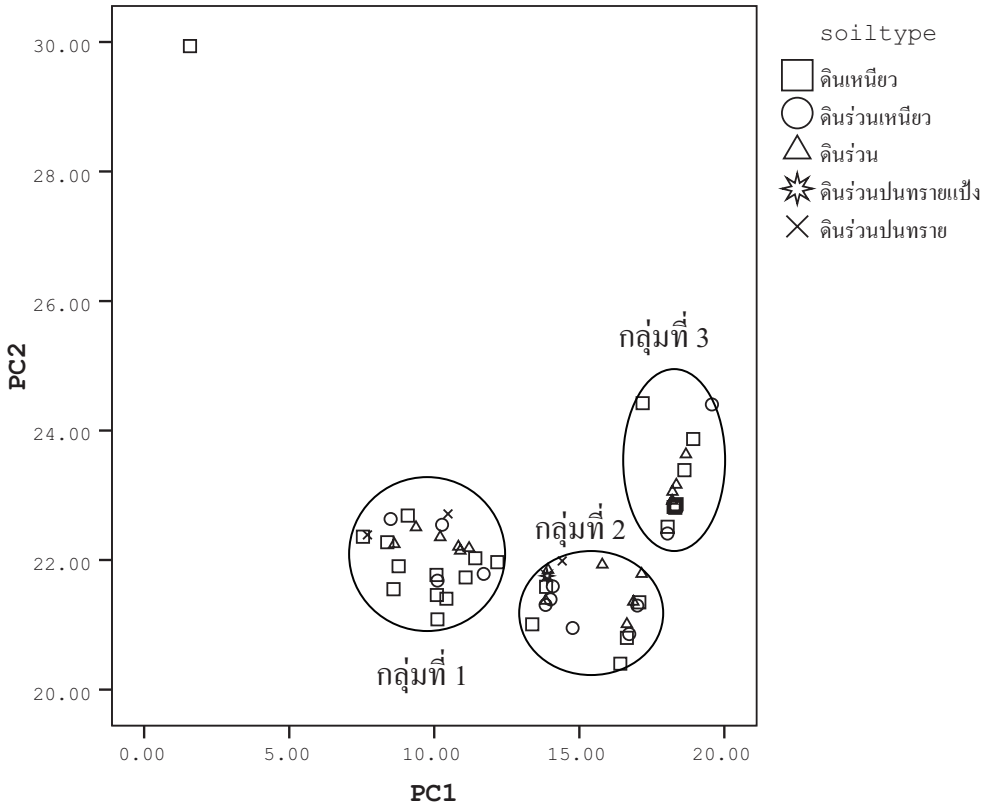
ภาพที่ 13 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ

การทดสอบการเท่ากันของค่าไอเกนทำการทดสอบโดยใช้สถิติทดสอบในสมการที่ (14) ภายใต้สมมติฐานในสมการที่ (15) พบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน ดังแสดงในตารางที่ 16

ตารางที่ 16 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ

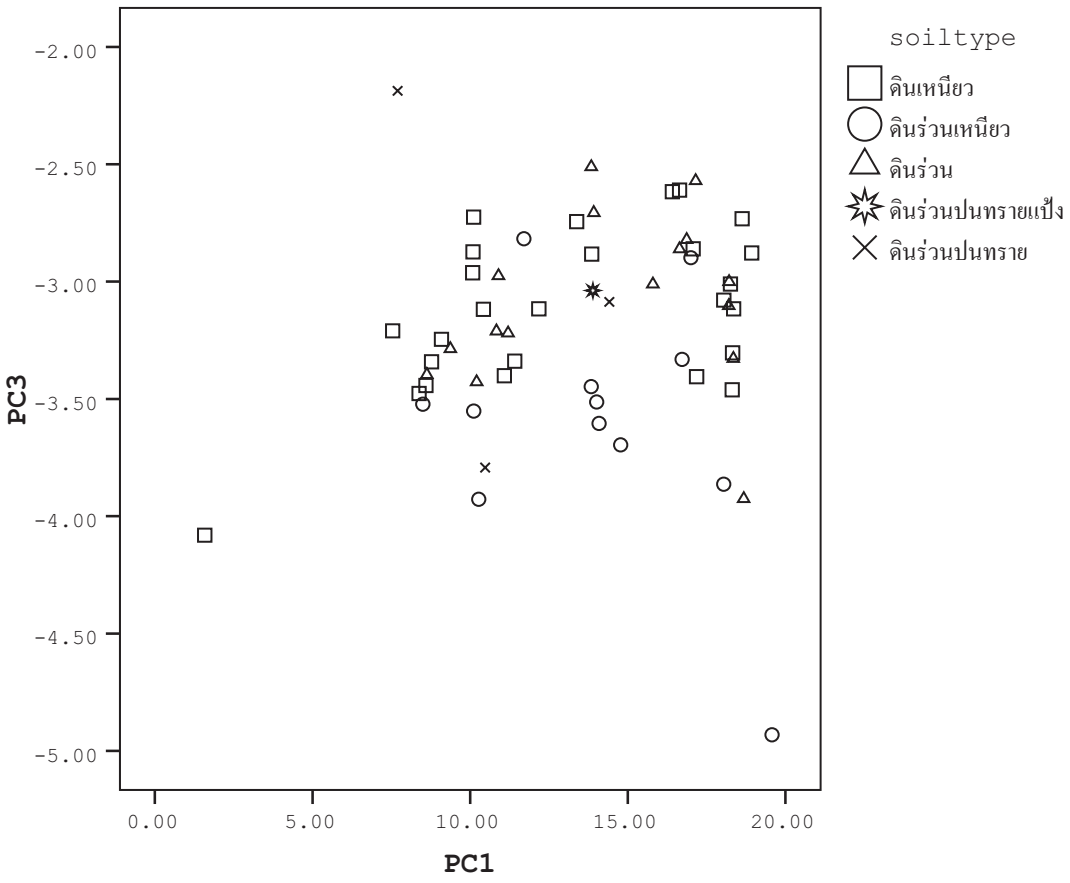
Eigenvalue	k	df	u	$\chi^2_{0.05}$
0.763	8	35	700.596	49.766
0.077	7	27	308.058	40.113
0.066	6	20	275.270	31.410
0.054	5	14	230.834	23.685
0.026	4	9	151.990	16.919
0.010	3	5	80.692	11.071
0.003	2	2	24.508	5.991
0.001	1			

จากการทดสอบการเท่ากันของค่าไอเกนให้ผลของการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด โดยแสดงให้เห็นว่าค่าไอเกนแตกต่างกันอย่างมีนัยสำคัญ แต่เมื่อพิจารณาภาพที่ 13 ซึ่งแสดง Scree plot ของค่าไอเกนองค์ประกอบหลักด้วยค่าตัวประมาณแบบเครื่องหมายจากข้อมูลดิบ ดังนั้นในกรณีนี้อาจใช้องค์ประกอบหลัก 3 แรกก็เพียงพอที่จะอธิบายความแปรปรวนของข้อมูลได้ถึง 90.6%



ภาพที่ 14 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ

จากภาพที่ 14 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะเห็นการจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน และจะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าปานกลางกับค่าองค์ประกอบหลักที่ 2 มีค่าต่ำมีส่วนประกอบของซิลิคอน (Si) และคาร์บอน (C) ปนอยู่พอกัน กลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าสูงกับค่าองค์ประกอบหลักที่ 2 มีค่าต่ำมีส่วนประกอบของคาร์บอน (C) สูงและซิลิคอน (Si) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าสูงกับค่าองค์ประกอบหลักที่ 2 มีค่าค่อนข้างปานกลางมีส่วนประกอบของซิลิคอน (Si) และคาร์บอน (C) สูง



ภาพที่ 15 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบเครื่องหมาย จากข้อมูลดิบ

จากภาพที่ 15 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะไม่เห็นการจำแนกกลุ่มของดิน

3.2 ข้อมูลแปลงด้วยวิธีของ Aitchison

ผลการวิเคราะห์ข้อมูลแปลงด้วยวิธีของ Aitchison โดยการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายแสดงในตารางที่ 17-18 โดยตารางที่ 17 แสดงค่าความแปรปรวนและความแปรปรวนร่วมด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง และตารางที่ 18 แสดงค่าสัมประสิทธิ์ขององค์ประกอบหลัก 5 องค์ประกอบแรกจาก 8 องค์ประกอบหลักที่สามารถคำนวณได้

ตารางที่ 17 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง

	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	0.0614							
Al	-0.0002	0.0083						
Si	0.0030	0.0049	0.0073					
K	0.0074	0.0042	0.0050	0.0609				
Ca	-0.0136	-0.0001	0.0005	-0.0567	0.2178			
Fe	0.0104	0.0070	0.0085	-0.0027	0.0116	0.0225		
C	-0.0551	-0.0435	-0.0561	-0.0392	-0.0358	-0.0690	0.5669	
Ti	0.0126	0.0027	0.0012	0.0094	0.0164	0.0087	-0.0339	0.0550

จากตารางที่ 17 ค่าความแปรปรวนของคาร์บอน (C) มีค่าสูงสุดเท่ากับ 0.5669 รองลงมาเป็นแคลเซียม (Ca) เท่ากับ 0.2419 และน้อยที่สุดเป็นซิลิคอน (Si) เท่ากับ 0.0073 ความแปรปรวนร่วมระหว่างเหล็ก (Fe) กับคาร์บอน (C) มีค่าสูงสุดเท่ากับ -0.0690 รองลงมาเป็นโพแทสเซียม (K) กับแคลเซียม (Ca) เท่ากับ -0.0567 และน้อยสุดเป็นแมกนีเซียม (Mg) กับอลูมินัม (Al) เท่ากับ -0.0002

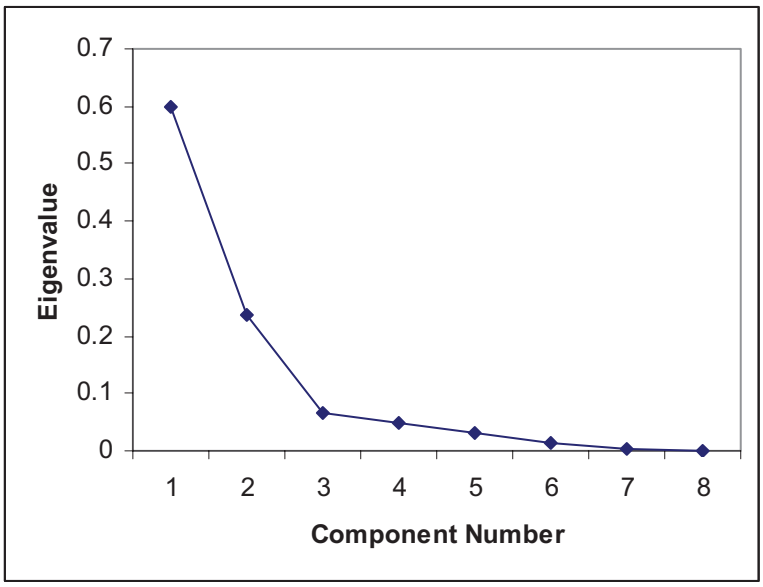
ตารางที่ 18 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง

	PC1	PC2	PC3	PC4	PC5
Mg	-0.103	-0.098	**0.649**	**-0.661**	0.328
Al	-0.075	-0.015	-0.057	0.070	-0.079
Si	-0.096	-0.018	-0.063	0.017	-0.013
K	-0.065	-0.316	0.193	**0.575**	**0.687**
Ca	-0.085	**0.940**	0.080	0.105	0.299
Fe	-0.123	0.033	0.068	-0.105	-0.253
C	**0.971**	0.055	0.137	0.004	0.031
Ti	-0.069	0.052	**0.710**	**0.453**	**-0.509**
Eigenvalue	0.599	0.237	0.065	0.050	0.031
Cumulative	0.599	0.836	0.901	0.951	0.983

องค์ประกอบหลักแรกมีค่าน้ำหนักสูงที่ตัวแปรคาร์บอน (C) ดังนั้นองค์ประกอบแรกจึงเป็นตัวแทนของตัวแปรคาร์บอน (C) องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรแคลเซียม (Ca) องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) เนื่องจากมีค่าน้ำหนักสูงที่ตัวแปรที่ตัวแปรทั้งสอง องค์ประกอบหลักที่สี่เป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) โพแทสเซียม (K) และไทเทเนียม (Ti) องค์ประกอบหลักที่ห้าเป็นตัวแทนของตัวแปรโพแทศเซียม (K) และไทเทเนียม (Ti)

ตัวแปรอลูมินัม (Al) ซิลิคอน (Si) และเหล็ก (Fe) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลย เนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล

องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 59.9% องค์ประกอบหลักที่ 2 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 23.7% องค์ประกอบหลักที่ 3 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 6.5% องค์ประกอบหลักที่ 4 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 5% สำหรับองค์ประกอบหลักที่ 5, 6, 7 และ 8 เป็นตัวอธิบายความแตกต่างในเรื่องของตัวแปรต่าง ๆ ออกไป ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 90.1% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.3%



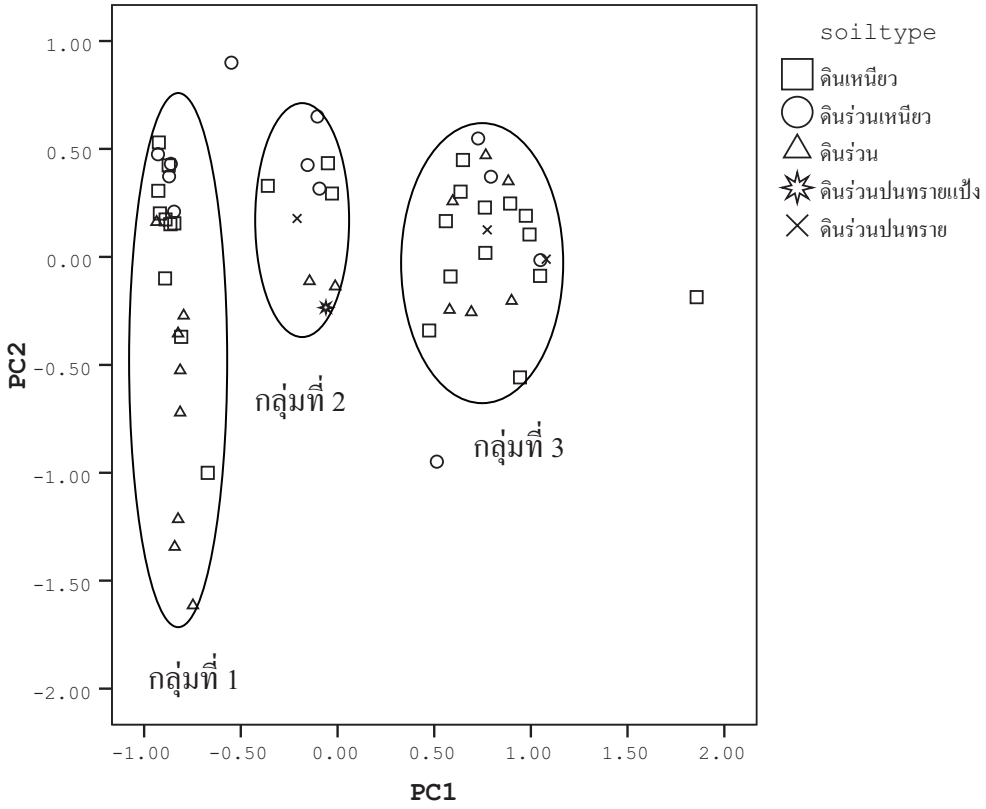
ภาพที่ 16 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง

การทดสอบการเท่ากันของค่าไอเกนทำการทดสอบโดยใช้สถิติทดสอบในสมการที่ (14) ภายใต้สมมติฐานในสมการที่ (15) พบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน ดังแสดงในตารางที่ 19

ตารางที่ 19 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง

Eigenvalue	k	df	u	$\chi^2_{0.05}$
0.599	8	35	604.481	49.766
0.237	7	27	396.506	40.113
0.065	6	20	234.602	31.410
0.050	5	14	194.881	23.685
0.031	4	9	140.503	16.919
0.012	3	5	71.043	11.071
0.004	2	2	21.452	5.991
0.001	1			

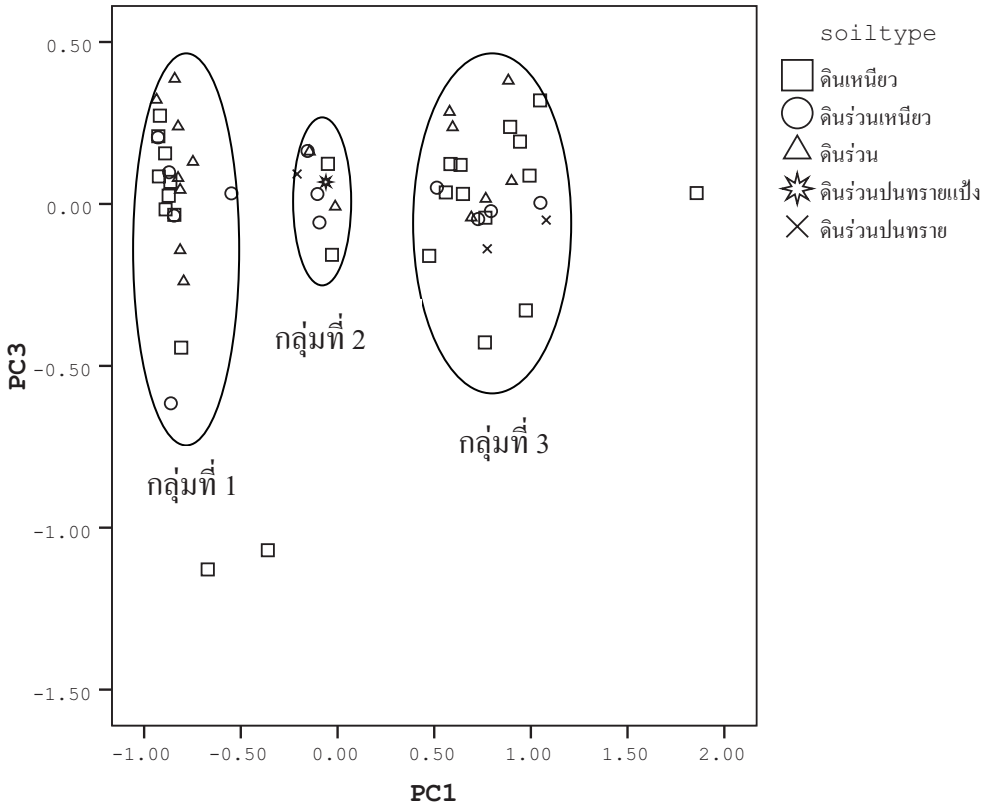
ผลจากการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลงจะเห็นว่าให้ผลของการทดสอบเช่นเดียวกับการทดสอบค่าไอเกนจากข้อมูลดิบอีกทั้งยังให้ผลการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด โดยแสดงให้เห็นว่าค่าไอเกนแตกต่างกันอย่างมีนัยสำคัญ แต่เมื่อพิจาณาร่วมกับภาพที่ 16 ซึ่งแสดง Scree plot ของค่าไอเกนองค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายของดินจากข้อมูลแปลง ดังนั้นในกรณีนี้อาจใช้องค์ประกอบหลัก 3 แรกก็เพียงพอที่จะอธิบายความแปรปรวนของข้อมูลได้ถึง 90.1%



ภาพที่ 17 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง

จากภาพที่ 17 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 และจากรูปที่ 18 แสดงการพลอตองค์ประกอบหลักที่ 1 และ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จะเห็นว่าผลของการจำแนกกลุ่มของดินตามชนิดของดินของทั้งสองภาพให้ผล

ที่คล้ายกันนั่นคือสามารถจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน และจากภาพที่ 17 จะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 มีค่าอยู่ในช่วงปานกลางมีส่วนประกอบของแคลเซียม (Ca) ปนอยู่ปานกลางและคาร์บอน (C) ปนอยู่บ้าง และกลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 มีค่าสูงมีส่วนประกอบของแคลเซียม (Ca) สูงและคาร์บอน (C) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าอยู่กลางๆ กับค่าองค์ประกอบหลักที่ 2 มีค่าสูงมีส่วนประกอบของคาร์บอน (C) และแคลเซียม (Ca) สูงและจากภาพที่ 18 จะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้าง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำแต่มากกว่ากลุ่มที่ 1 กับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้างแต่มากกว่ากลุ่มที่ 1 และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าอยู่กลางๆ กับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของคาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) สูง อีกทั้งผลที่ได้จากการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 และค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลแปลงนี้ยังให้ผลการวิเคราะห์คล้ายกันกับการพลอตค่าองค์ประกอบหลักดังกล่าวด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลงในภาพที่ 11 และ 12 อีกด้วย



ภาพที่ 18 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบเครื่องหมาย จากข้อมูลแปลง

4. การวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ

การวิเคราะห์องค์ประกอบหลักจะพิจารณาข้อมูลองค์ประกอบทางเคมีของดินจำนวน 8 ตัวแปรได้แก่ แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) คาร์บอน (C) และไทเทเนียม (Ti) จากตัวอย่างจำนวน 58 ตัวอย่าง ซึ่งจากการตรวจสอบเบื้องต้นพบว่ามีจำนวน 4 ตัวแปรที่มีค่านอกกลุ่มปะปนอยู่ด้วยคือตัวแปรแมกนีเซียม (Mg) แคลเซียม (Ca) คาร์บอน (C) และไทเทเนียม (Ti) ดังนั้นในการวิเคราะห์ข้อมูลจะทำการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ

4.1 ข้อมูลดิบ (raw data)

ผลการวิเคราะห์ข้อมูลดิบโดยการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับแสดงในตารางที่ 20-21 โดยตารางที่ 20 แสดงค่าความแปรปรวน และความแปรปรวนร่วมด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ และตารางที่ 21 แสดงค่าสัมประสิทธิ์ขององค์ประกอบหลัก 5 องค์ประกอบแรกจาก 8 องค์ประกอบหลักที่สามารถคำนวณได้

ตารางที่ 20 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ

	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	0.00208							
Al	-0.00096	0.02298						
Si	0.00112	0.02519	0.13120					
K	0.00108	-0.00007	0.00915	0.00983				
Ca	-0.00088	-0.00046	-0.00189	-0.00368	0.00634			
Fe	0.00155	0.00624	0.02336	-0.00080	-0.00033	0.01605		
C	-0.00405	-0.02917	-0.11492	-0.00521	-0.00175	-0.02039	0.20343	
Ti	0.00057	-0.00063	-0.00078	0.00025	0.00045	0.00068	-0.00124	0.00084

จากตารางที่ 20 ค่าความแปรปรวนของคาร์บอน (C) มีค่าสูงสุดเท่ากับ 0.20343 รองลงมาเป็นซิลิคอน (Si) เท่ากับ 0.13120 และน้อยที่สุดเป็นไทเทเนียม (Ti) เท่ากับ 0.00084 ความแปรปรวนร่วมระหว่างซิลิคอน (Si) กับคาร์บอน (C) มีค่าสูงสุดเท่ากับ -0.11492 รองลงมาเป็นอลูมินัม (Al) กับคาร์บอน (C) เท่ากับ -0.02917 และน้อยสุดเป็นโพแทสเซียม (K) กับไทเทเนียม (Ti) เท่ากับ 0.00025

ตารางที่ 21 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ

	PC1	PC2	PC3	PC4	PC5
Mg	0.032	-0.027	-0.078	0.061	0.241
Al	0.140	0.094	**0.897**	-0.387	0.030
Si	**0.587**	**0.759**	-0.172	-0.012	-0.207
K	0.033	0.097	-0.292	**-0.524**	**0.522**
Ca	0.000	-0.068	0.073	0.286	**-0.520**
Fe	0.109	0.173	0.254	**0.698**	**0.591**
C	**-0.789**	**0.609**	0.053	-0.002	-0.041
Ti	0.002	-0.027	-0.023	0.063	0.070
Eigenvalue	0.297	0.050	0.018	0.013	0.010
Cumulative	0.755	0.881	0.928	0.96	0.986

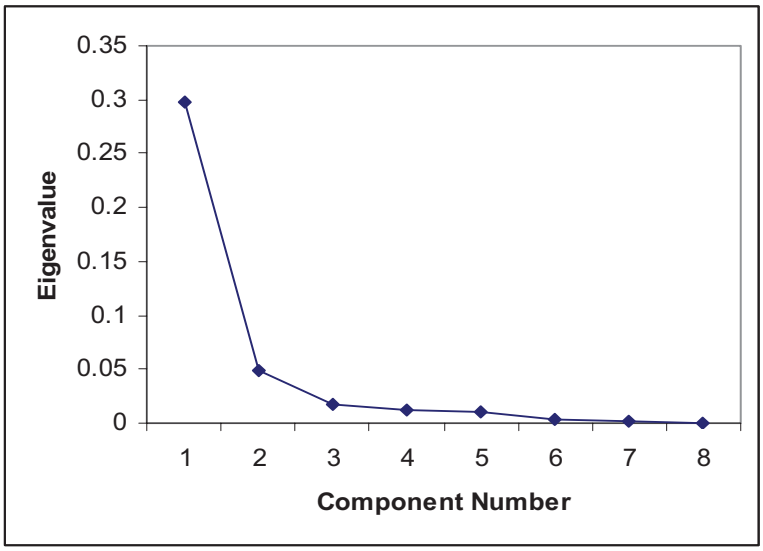
องค์ประกอบหลักแรกมีค่าน้ำหนักสูงที่ตัวแปรซิลิคอน (Si) และคาร์บอน (C) ดังนั้น องค์ประกอบแรกจึงเป็นตัวแทนของทั้งสองตัวแปรดังกล่าว โดยถ้าค่าขององค์ประกอบแรกมีค่าสูง แสดงว่าเป็นกลุ่มดินที่มีค่าซิลิคอน (Si) สูงเนื่องจากน้ำหนักมีค่าเป็นบวก ส่วนค่าองประกอบแรกที่มีค่าต่ำแสดงว่าเป็นกลุ่มดินที่มีค่าคาร์บอน (C) สูงเนื่องจากน้ำหนักมีค่าเป็นลบ ถ้าเป็นกลุ่มดินที่มีค่าซิลิคอน (Si) และคาร์บอน (C) พอกันจะทำให้ได้ค่าองค์ประกอบหลักแรกกลางๆ

องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) เช่นเดียวกับในองค์ประกอบแรกแต่อธิบายไปในทิศทางเดียวกัน องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรอลูมินัม (Al) เนื่องจากมีค่าน้ำหนักสูงที่ตัวแปรอลูมินัม (Al) เพียงตัวเดียว องค์ประกอบหลักที่สี่เป็นตัวแทนของตัวแปรโพแทสเซียม (K) และเหล็ก (Fe) องค์ประกอบที่ห้าเป็นตัวแทนของตัวแปรโพแทสเซียม (K) เหล็ก (Fe) และแคลเซียม (Ca)

ตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลย เนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล

องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 75.5% องค์ประกอบหลักที่ 2 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 12.5% องค์ประกอบหลักที่ 3 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 4.6% องค์ประกอบหลักที่ 4 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 3.3% สำหรับองค์ประกอบหลักที่ 5, 6, 7 และ 8 เป็นตัวอธิบายความแตกต่างในเรื่องของตัวแปรต่าง ๆ ออกไป ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 92.8% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6%

การทดสอบการเท่ากันของค่าไอเกนทำการทดสอบโดยใช้สถิติทดสอบในสมการที่ (14) ภายใต้สมมติฐานในสมการที่ (15) พบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน ดังแสดงในตารางที่ 22



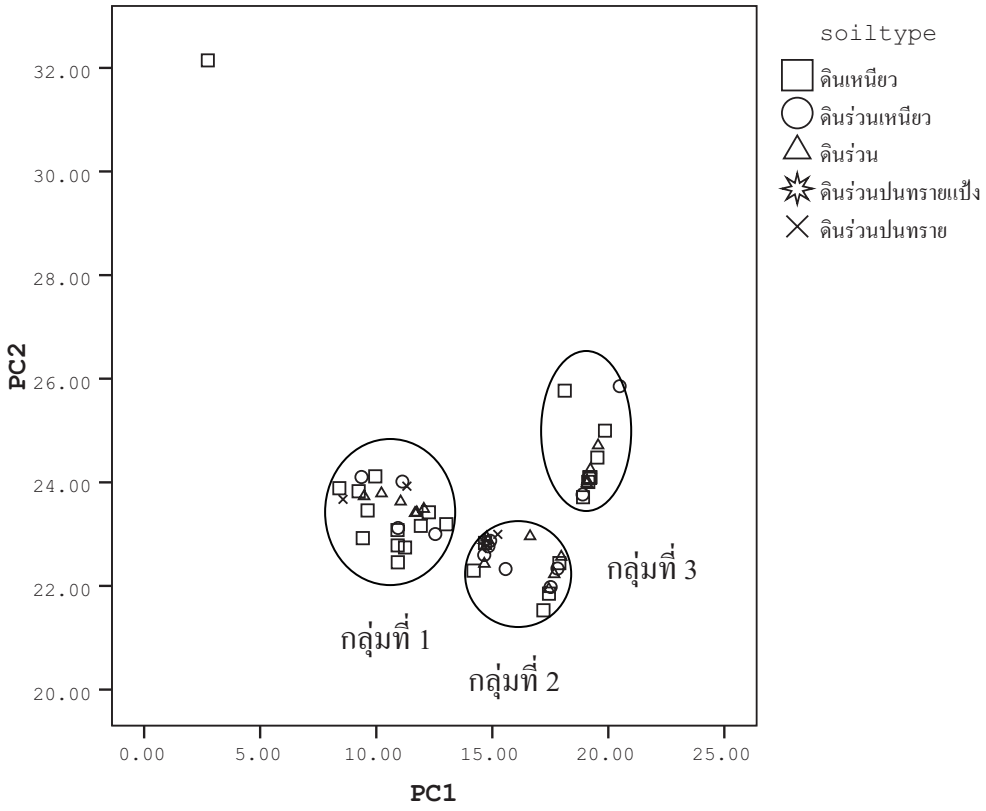
ภาพที่ 19 Scree plot ขององค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ

ตารางที่ 22 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ

Eigenvalue	k	df	u	$\chi^2_{0.05}$
0.297	8	35	687.150	49.766
0.050	7	27	306.156	40.113
0.018	6	20	192.296	31.410
0.013	5	14	154.035	23.685
0.010	4	9	116.467	16.919
0.004	3	5	50.445	11.071
0.001	2	2	15.953	5.991
0.000	1			

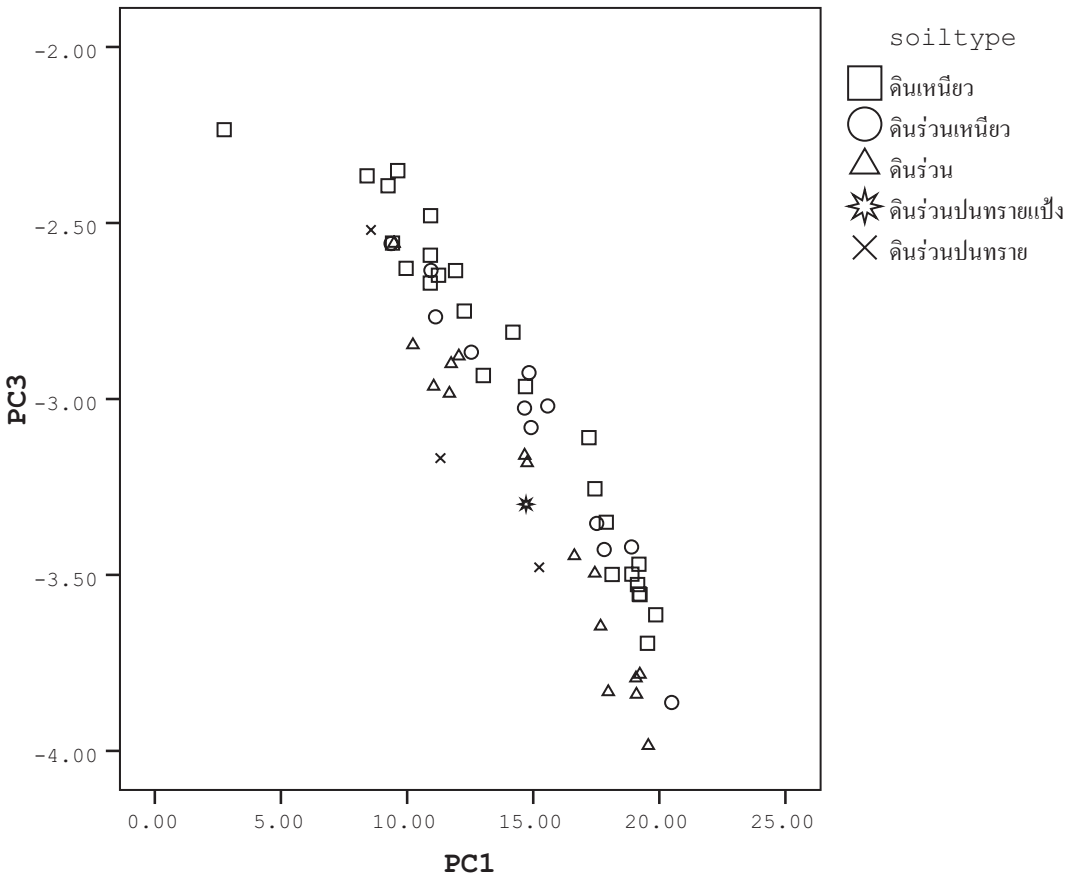
ผลจากการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบอันดับจากข้อมูลดิบจะเห็นว่าให้ผลของการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดและค่าประมาณแบบเครื่องหมาย โดยแสดงให้เห็นว่าค่าไอเกนแตกต่างกันอย่างมีนัยสำคัญ เมื่อพิจารณาร่วมกับภาพที่ 19 ซึ่งแสดง Scree plot ของค่าไอเกนองค์ประกอบหลัก

ด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ ดังนั้นในกรณีนี้อาจใช้องค์ประกอบหลัก 3 องค์ประกอบแรกก็เพียงพอที่จะอธิบายความแปรปรวนของข้อมูลได้ถึง 92.8%



ภาพที่ 20 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ

จากภาพที่ 20 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะเห็นการจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน และจะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิกอน (Si) และคาร์บอน (C) ปานกลาง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิกอน (Si) และคาร์บอน (C) สูง และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิกอน (Si) ปานกลางและคาร์บอน (C) สูง เมื่อพิจารณาจากองค์ประกอบหลักที่ 1 ซึ่งผลการวิเคราะห์ดังกล่าวมีลักษณะใกล้เคียงกับผลการวิเคราะห์ที่ได้จากการพลอตกราฟองค์ประกอบที่ 1 และ 2 ด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบในภาพที่ 14



ภาพที่ 21 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบอันดับจากข้อมูลดิบ

จากภาพที่ 21 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 7 จากภาพจะไม่เห็นการจำแนกกลุ่มของดิน

4.1 ข้อมูลแปลงด้วยวิธีของ Aitchison

ผลการวิเคราะห์ข้อมูลแปลงด้วยวิธีของ Aitchison โดยการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับแสดงในตารางที่ 23-43 โดยตารางที่ 23 แสดงค่าความแปรปรวน และความแปรปรวนร่วมด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง และตารางที่ 24 แสดงค่าสัมประสิทธิ์ขององค์ประกอบหลัก 5 องค์ประกอบแรกจาก 8 องค์ประกอบหลักที่สามารถคำนวณได้

ตารางที่ 23 ความแปรปรวน และความแปรปรวนร่วมของธาตุต่าง ๆ ด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง

	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	0.02882							
Al	-0.00053	0.00415						
Si	0.00082	0.00217	0.00351					
K	0.00275	0.00138	0.00276	0.02438				
Ca	-0.00893	-0.00076	-0.00076	-0.01846	0.10921			
Fe	0.00529	0.00348	0.00418	0.00005	-0.00079	0.01328		
C	-0.01751	-0.01563	-0.02125	-0.01441	-0.00613	-0.02632	0.21246	
Ti	0.01018	0.00039	-0.00027	0.00365	0.00575	0.00527	-0.00803	0.03766

จากตารางที่ 23 ค่าความแปรปรวนของคาร์บอน (C) มีค่าสูงสุดเท่ากับ 0.21246 รองลงมาเป็นแคลเซียม (Ca) เท่ากับ 0.10921 และน้อยที่สุดเป็นซิลิคอน (Si) เท่ากับ 0.00351 ความแปรปรวนร่วมระหว่างเหล็ก (Fe) กับคาร์บอน (C) มีค่าสูงสุดเท่ากับ -0.02632 รองลงมาเป็นซิลิคอน (Si) กับคาร์บอน (C) เท่ากับ -0.02917 และน้อยสุดเป็นซิลิคอน (Si) กับไทเทเนียม (Ti) เท่ากับ -0.00027

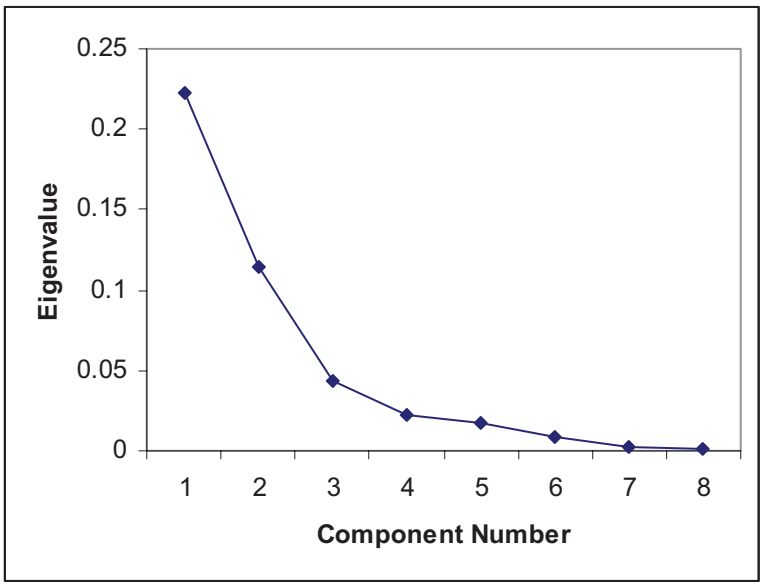
ตารางที่ 24 ค่าสัมประสิทธิ์ขององค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง

	PC1	PC2	PC3	PC4	PC5
Mg	-0.094	-0.105	**0.486**	**-0.682**	**0.516**
Al	-0.073	-0.010	-0.031	0.059	-0.091
Si	-0.099	-0.015	-0.037	0.043	0.002
K	-0.072	-0.203	0.106	**0.631**	**0.699**
Ca	-0.034	**0.972**	0.033	0.041	0.215
Fe	-0.129	-0.014	0.124	-0.133	-0.232
C	**0.975**	0.008	0.113	-0.008	0.015
Ti	-0.054	0.048	**0.849**	0.334	-0.367
Eigenvalue	0.223	0.114	0.044	0.023	0.018
Cumulative	0.514	0.778	0.879	0.931	0.971

องค์ประกอบหลักแรกมีค่าน้ำหนักสูงที่ตัวแปรคาร์บอน (C) ดังนั้นองค์ประกอบแรกจึงเป็นตัวแทนของตัวแปรคาร์บอน (C) องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรแคลเซียม (Ca) องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) องค์ประกอบหลักที่สี่และห้าเป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และโพแทสเซียม (K) เนื่องจากมีค่าน้ำหนักสูงที่ตัวแปรที่ตัวแปรทั้งสอง

ตัวแปรอลูมินัม (Al) ซิลิคอน (Si) และเหล็ก (Fe) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลย เนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล

องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 51.4% องค์ประกอบหลักที่ 2 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 26.4% องค์ประกอบหลักที่ 3 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 10.1% องค์ประกอบหลักที่ 4 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 5.2% สำหรับองค์ประกอบหลักที่ 5, 6, 7 และ 8 เป็นตัวอธิบายความแตกต่างในเรื่องของตัวแปรต่าง ๆ ออกไป ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 87.9% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 97.1%



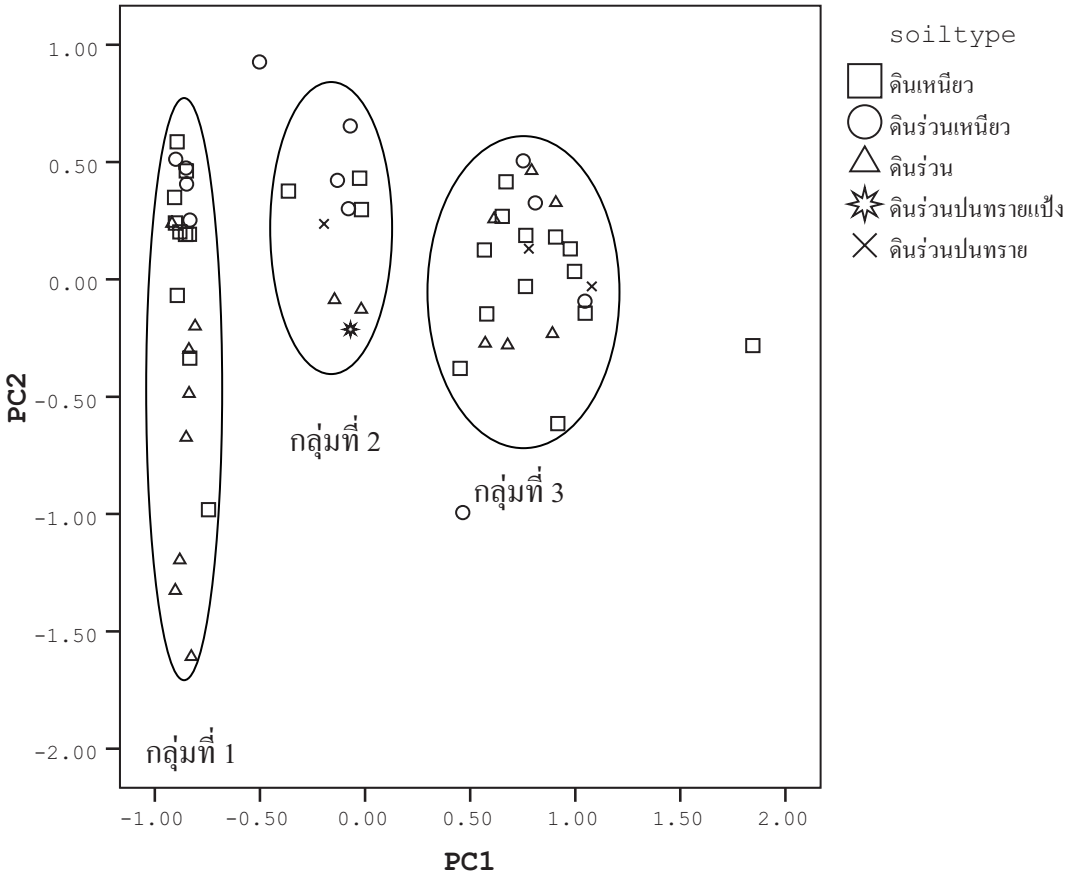
ภาพที่ 22 Scree plot ของค่าไอเกนองค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง

การทดสอบการเท่ากันของค่าไอเกนทำการทดสอบโดยใช้สถิติทดสอบในสมการที่ (14) ภายใต้สมมติฐานในสมการที่ (15) พบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน ดังแสดงในตารางที่ 25

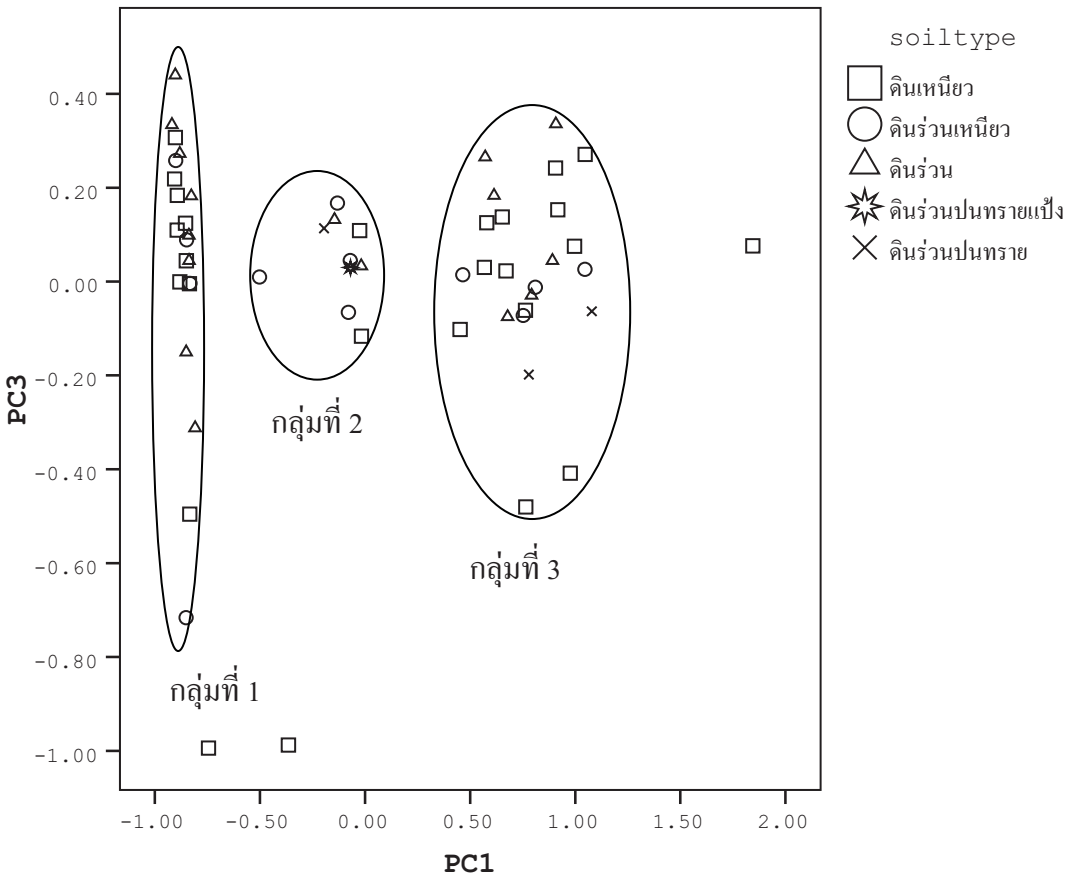
ตารางที่ 25 ค่าสถิติของการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง

Eigenvalue	k	df	u	$\chi^2_{0.05}$
0.223	8	35	489.311	49.766
0.114	7	27	344.928	40.113
0.044	6	20	214.979	31.410
0.023	5	14	155.289	23.685
0.018	4	9	123.709	16.919
0.009	3	5	74.088	11.071
0.002	2	2	14.836	5.991
0.001	1			

ผลจากการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง จะเห็นได้ว่าให้ผลของการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนจากข้อมูลดิบซึ่งสามารถที่จะกล่าวได้ว่าการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด ค่าประมาณแบบเครื่องหมาย และค่าประมาณแบบอันดับให้ผลการทดสอบเช่นเดียวกัน โดยแสดงให้เห็นว่าค่าไอเกนแตกต่างกันอย่างมีนัยสำคัญ สำหรับในกรณีนี้เมื่อพิจารณาจุดมุ่งหมายของการวิเคราะห์องค์ประกอบหลักร่วมกับภาพที่ 22 ซึ่งแสดง Scree plot ของค่าไอเกนองค์ประกอบหลักด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง ดังนั้นในกรณีนี้อาจใช้องค์ประกอบหลัก 3 แรกก็เพียงพอที่จะอธิบายความแปรปรวนของข้อมูลได้ถึง 87.9%



ภาพที่ 23 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง



ภาพที่ 24 กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบอันดับจากข้อมูลแปลง

จากภาพที่ 23 แสดงการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 และจากภาพที่ 24 แสดงการพลอตองค์ประกอบหลักที่ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณาดังแสดงในตารางที่ 6 จะเห็นว่าผลของการจำแนกกลุ่มของดินตามชนิดของดินของทั้งสองรูปให้ผลที่คล้ายกันนั่นคือสามารถจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน และจากภาพที่ 23 จะเห็นว่ากลุ่มที่ 1 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 มีค่าอยู่ในช่วงปานกลางมีส่วนประกอบของแคลเซียม (Ca) ปนอยู่ปานกลางและคาร์บอน (C) ปนอยู่บ้าง และกลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 2 มีค่าสูงมีส่วนประกอบของแคลเซียม (Ca) สูงและคาร์บอน (C) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าอยู่กลางๆ กับค่าองค์ประกอบหลักที่ 2 มีค่าสูงมีส่วนประกอบของคาร์บอน (C) และแคลเซียม (Ca) สูงและจากภาพที่ 24 จะเห็นว่ากลุ่มที่ 1 เป็น

กลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำกับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้าง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าต่ำแต่มากกว่ากลุ่มที่ 1 กับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้างแต่มากกว่ากลุ่มที่ 1 และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีค่าองค์ประกอบหลักที่ 1 มีค่าอยู่กลางๆ กับค่าองค์ประกอบหลักที่ 3 มีค่าสูงมีส่วนประกอบของคาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) สูง แต่ประเด็นที่น่าสนใจคือผลที่ได้จากการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 และค่าองค์ประกอบหลักที่ 1 กับ 3 ด้วยค่าประมาณแบบอันดับจากข้อมูลแปลงนี้ยังให้ผลการวิเคราะห์คล้ายกันกับการพลอตค่าองค์ประกอบหลักดังกล่าวด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด และค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง ดังนั้นจึงอาจกล่าวได้ว่าการแปลงข้อมูลด้วยวิธีของ Aitchison มีผลให้การจำแนกกลุ่มของดินโดยวิเคราะห์การพลอตกราฟขององค์ประกอบหลักดังกล่าว ให้ผลการวิเคราะห์ที่คล้ายคลึงกัน

บทที่ 5
สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยเพื่อศึกษาการประยุกต์การวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด (Principal Component Analysis based on Maximum likelihood Estimates) และแบบแกร่งของเมตริกซ์ความแปรปรวนร่วม (Robust Principal Component Analysis based on covariance matrix) ในการจำแนกกลุ่มดินตามองค์ประกอบทางเคมี กรณีศึกษาภูมิภาคตะวันตกของประเทศไทย สามารถสรุปผลได้ดังนี้

สรุปผลการวิจัย

การตรวจสอบข้อมูลเบื้องต้นของตัวอย่างของดินที่ถูกเก็บมาจาก 3 จังหวัด ได้แก่ นครปฐม สมุทรสาคร และสมุทรสงคราม จำนวน 58 ตัวอย่าง ซึ่งถูกวัดโดยการใช้เทคนิคทางกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด (Scanning Electron Microscope, SEM) ร่วมกับการวิเคราะห์รังสีเอ็กซ์เรย์ที่วัดได้จากตัวอย่างวัตถุด้วยหัวตรวจวัด Energy Dispersive X-ray Micro-analysis (EDX) จากตัวอย่างดิน ซึ่งในการวิจัยนี้ทำการศึกษาขนาดของดินที่ขนาดตัวอย่าง ขนาด 65 ไมครอน (เป็นขนาดที่ดูดซับโลหะมากที่สุด) ตัวแปรที่พิจารณาคือ ออกซิเจน (O) แมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) ไทเทเนียม (Ti) คาร์บอน (C) และโซเดียม (Na) ได้ผลการสรุปดังนี้

1. ปริมาณองค์ประกอบทางเคมีที่พบในตัวอย่างของดินคือ O > Si > Al > Fe > C > K > Mg > Ca > Ti > Na

2. การทดสอบการแจกแจงปกติพบว่าที่ระดับนัยสำคัญ 0.05 ตัวแปรแมกนีเซียม (Mg) อลูมินัม (Al) โพแทสเซียม (K) แคลเซียม (Ca) และเหล็ก (Fe) มีการแจกแจงแบบปกติ ส่วนตัวแปรออกซิเจน (O) ซิลิคอน (Si) คาร์บอน (C) ไทเทเนียม (Ti) และโซเดียม (Na) ไม่ได้มีการแจกแจงปกติ

3. การตรวจสอบค่านอกกลุ่มโดยใช้ Boxplot พบว่ามีจำนวน 5 ตัวแปรที่มีค่านอกกลุ่มปะปนอยู่ด้วยคือตัวแปร แมกนีเซียม (Mg) แคลเซียม (Ca) คาร์บอน (C) ไทเทเนียม (Ti) และโซเดียม (Na)

สำหรับตัวแปรที่ใช้ในวิเคราะห์องค์ประกอบหลักผู้วิจัยใช้ตัวแปร 8 ตัวแปรคือแมกนีเซียม (Mg) อลูมิเนียม (Al) ซิลิคอน (Si) โพแทสเซียม (K) แคลเซียม (Ca) เหล็ก (Fe) คาร์บอน (C) และไทเทเนียม (Ti) โดยผู้วิจัยได้ตัดตัวแปร 2 ตัวแปรออกคือตัวแปรออกซิเจน (O) เนื่องจากเป็นตัวแปรที่เป็นส่วนประกอบหลักที่พบในดินมีปริมาณในช่วง 50%-57% และตัวแปรโซเดียม (Na) เนื่องจากมีค่าเป็นศูนย์ส่วนใหญ่ พบเฉพาะดินที่อยู่ใกล้ทะเล

ผลการวิเคราะห์ข้อมูลใช้วิธีวิเคราะห์องค์ประกอบหลัก 3 วิธีคือการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด การวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายและการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ โดยการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 และ 1 กับ 3 ภายใต้ข้อมูล 2 ลักษณะคือ ข้อมูลดิบ (raw data) และข้อมูลแปลงด้วยวิธีของ Aitchison สามารถสรุปผลการวิจัยได้ดังนี้

1. ผลการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด ประกอบด้วย

1.1 ข้อมูลดิบ (raw data) ผลการวิเคราะห์องค์ประกอบหลักพบว่า องค์ประกอบหลักแรกเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 86.4% องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) เช่นเดียวกันสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 11% องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรอลูมินัม (Al) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 1.2% สำหรับองค์ประกอบที่สี่และห้าเป็นตัวอธิบายตัวแปรเหล็ก (Fe) โพแทสเซียม (K) และแคลเซียม (Ca) และสำหรับตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 99.7% จากการทดสอบการเท่ากันของค่าไอเกนพบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน และจากการกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดคือ ดินเหนียว ดินร่วนเหนียว ดินร่วน ดินร่วนปนทรายแป้ง และดินร่วนปนทราย จะเห็นการจำแนกกลุ่มของดินได้ 2 กลุ่มแต่ไม่สอดคล้องกับชนิดของดินโดยกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิคอน (Si) สูงและคาร์บอน (C) ปนอยู่บ้าง ในขณะที่กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) สูงและซิลิคอน (Si) ปนอยู่เล็กน้อย และจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 จะเห็นการจำแนกกลุ่มของดินได้ 2 กลุ่มแต่ไม่ชัดเจนนัก กลุ่มที่ 1 เป็นการรวมกลุ่มของดิน 3 ชนิดคือดินเหนียว ดินร่วนเหนียว และดินร่วน ซึ่ง

ชนิดของดินส่วนใหญ่ในกลุ่มนี้เป็นดินเหนียว ที่มีส่วนประกอบที่มีอลูมินัม(Al) สูง กลุ่มที่ 2 เป็นการรวมกลุ่มของดินอีก 4 ชนิดคือดินร่วนเหนียว ดินร่วน ดินร่วนปนทรายแป้ง และดินร่วนปนทรายและมีตัวอย่างของดินที่เป็นดินเหนียวร่วนอยู่ด้วยเพียงเล็กน้อยเป็นกลุ่มชนิดของดินที่มีอะลูมินัม (Al) ต่ำกว่าดินเหนียว

1.2 ข้อมูลแปลงด้วยวิธีของ Aitchison ผลการวิเคราะห์องค์ประกอบหลักพบว่าองค์ประกอบหลักแรกเป็นตัวแทนของตัวแปรคาร์บอน (C) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 61.1% องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรแคลเซียม (Ca) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 24.5% องค์ประกอบหลักที่สามและสี่เป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) องค์ประกอบหลักที่สามสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 8.1% และองค์ประกอบหลักที่ 4 สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดความแตกต่างของดิน 3.1% สำหรับองค์ประกอบหลักที่ห้าเป็นตัวแทนของตัวแปรโพแทสเซียม (K) และสำหรับตัวแปรอลูมินัม (Al) ซิลิคอน (Si) และเหล็ก (Fe) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 93.8% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6% จากการทดสอบการเท่ากันของค่าไอเกนให้ผลเช่นเดียวกันกับข้อมูลดิบนั่นคือที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน และจากการกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 และ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจาณา จะเห็นว่าผลของการจำแนกกลุ่มของดินตามชนิดของดินของทั้งสองรูปให้ผลที่คล้ายกันนั่นคือสามารถจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน โดยกลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของแคลเซียม (Ca) ปนอยู่ปานกลางและคาร์บอน (C) ปนอยู่บ้าง และกลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของแคลเซียม (Ca) สูงและคาร์บอน (C) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) และแคลเซียม (Ca) สูง ในขณะที่กลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้าง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้างแต่มากกว่ากลุ่มที่ 1 และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) สูง

2. ผลการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมาย ประกอบด้วย

2.1 ข้อมูลดิบ (raw data) ผลการวิเคราะห์องค์ประกอบหลักพบว่า องค์ประกอบหลักแรกเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 76.3% องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) เช่นเดียวกันสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 7.7% องค์ประกอบหลักที่สามและสี่เป็นตัวแทนของตัวแปรอลูมินัม (Al) และโพแทสเซียม (K) องค์ประกอบหลักที่สามสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 6.6% และองค์ประกอบหลักที่สี่สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 5.4% สำหรับองค์ประกอบที่ห้าเป็นตัวแทนของตัวแปรเหล็ก (Fe) และสำหรับตัวแปรแมกนีเซียม (Mg) แคลเซียม (Ca) และไทเทเนียม (Ti) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 90.6% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6% จากการทดสอบการเท่ากันของค่าไอเกนให้ผลของการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดนั่นคือที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน และจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา จากรูปจะเห็นการจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน โดยกลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิคอน (Si) และคาร์บอน (C) ปนอยู่พอกัน กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) สูงและซิลิคอน (Si) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิคอน (Si) และคาร์บอน (C) สูง และจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 จากภาพจะไม่เห็นจำแนกกลุ่มของดินได้

2.2 ข้อมูลแปลงด้วยวิธีของ Aitchison ผลการวิเคราะห์องค์ประกอบหลักพบว่าองค์ประกอบหลักแรกเป็นตัวแทนของตัวแปรคาร์บอน (C) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 59.9% องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรแคลเซียม (Ca) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 23.7% องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 6.5% องค์ประกอบหลักที่สี่เป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) โพแทสเซียม (K) และไทเทเนียม (Ti) อธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 5% สำหรับองค์ประกอบหลักที่ห้าเป็น

ตัวแทนของตัวแปรโพแทศเซียม (K) และไทเทเนียม (Ti) และสำหรับตัวแปรอลูมินัม (Al) ซิลิคอน (Si) และเหล็ก (Fe) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 90.1% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.3% จากการทดสอบการเท่ากันของค่าไอเกนของข้อมูลแปลงให้ผลเช่นเดียวกับข้อมูลดิบอีกทั้งยังให้ผลการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด นั่นคือที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกันและจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 และ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา จะเห็นว่าการจำแนกกลุ่มของดินตามชนิดของดินของทั้งสองรูปให้ผลที่คล้ายกันนั่นคือสามารถจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน โดยกลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของแคลเซียม (Ca) ปนอยู่ปานกลางและคาร์บอน (C) ปนอยู่บ้าง และกลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของแคลเซียม (Ca) สูงและคาร์บอน (C) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) และแคลเซียม (Ca) สูง ในขณะที่กลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้าง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้างแต่มากกว่ากลุ่มที่ 1 และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) สูง อีกทั้งผลที่ได้คล้ายกันกับกราฟ 2 มิติของค่าองค์ประกอบหลักดังกล่าวด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดจากข้อมูลแปลง

3. ผลการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ ประกอบด้วย

3.1 ข้อมูลดิบ (raw data) ผลการวิเคราะห์องค์ประกอบหลักพบว่า องค์ประกอบหลักแรกเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 75.5% องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรซิลิคอน (Si) และคาร์บอน (C) เช่นเดียวกันสามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 12.5% องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรอลูมินัม (Al) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 4.6% องค์ประกอบหลักที่สี่เป็นตัวแทนของตัวแปรโพแทสเซียม (K) และเหล็ก (Fe) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 3.3% สำหรับองค์ประกอบที่ห้าเป็นตัวแทนของตัวแปรโพแทสเซียม (K) เหล็ก (Fe) และแคลเซียม (Ca) และสำหรับตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) ไม่ถูก

อธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูล ดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 92.8% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 98.6% จากการทดสอบการเท่ากันของค่าไอเกนบนข้อมูลดิบจะเห็นว่าให้ผลของการทดสอบเช่นเดียวกับการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดและค่าประมาณแบบเครื่องหมาย นั่นคือที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน และจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา จะเห็นการจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน โดยกลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิคอน (Si) และคาร์บอน (C) ปนอยู่พอกัน กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) สูงและซิลิคอน (Si) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของซิลิคอน (Si) และคาร์บอน (C) สูง ซึ่งผลการวิเคราะห์ดังกล่าวมีลักษณะใกล้เคียงกับผลการวิเคราะห์ที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักดังกล่าวด้วยค่าประมาณแบบเครื่องหมายจากข้อมูลดิบ และจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา จะไม่เห็นการจำแนกกลุ่มของดิน

3.2 ข้อมูลแปลงด้วยวิธีของ Aitchison ผลการวิเคราะห์องค์ประกอบหลักพบว่าองค์ประกอบหลักแรกเป็นตัวแทนของตัวแปรคาร์บอน (C) ความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดินได้ 51.4% องค์ประกอบหลักที่สองเป็นตัวแทนของตัวแปรแคลเซียม (Ca) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างได้ 26.4% องค์ประกอบหลักที่สามเป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti) สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 10.1% องค์ประกอบหลักที่สี่และห้าเป็นตัวแทนของตัวแปรแมกนีเซียม (Mg) และโพแทสเซียม (K) องค์ประกอบหลักที่สี่สามารถอธิบายความแปรปรวนในข้อมูลที่เกิดจากความแตกต่างของดิน 5.2% สำหรับตัวแปรอลูมินัม (Al) ซิลิคอน (Si) และเหล็ก (Fe) ไม่ถูกอธิบายด้วยองค์ประกอบใดเลยเนื่องจากเป็นตัวแปรที่มีค่าความแปรปรวนน้อย ไม่มีผลต่อการอธิบายความผันแปรของข้อมูลดังนั้น 3 องค์ประกอบหลักแรกสามารถอธิบายความแปรปรวนในข้อมูลได้ 87.9% และ 5 องค์ประกอบหลักสามารถอธิบายความแปรปรวนในข้อมูลได้ 97.1% จากการทดสอบการเท่ากันของค่าไอเกนบนข้อมูลแปลงจะเห็นได้ว่าให้ผลเช่นเดียวกับข้อมูลดิบ ซึ่งสามารถที่จะกล่าวได้ว่าการทดสอบการเท่ากันของค่าไอเกนด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด ค่าประมาณแบบเครื่องหมาย และค่าประมาณแบบอันดับให้ผลการทดสอบเช่นเดียวกัน นั่นคือที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 มีความแตกต่างกัน และจากกราฟ 2 มิติ

ของค่าองค์ประกอบหลักที่ 1 กับ 2 และ 1 กับ 3 โดยจำแนกตามชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา จะเห็นว่าการจำแนกกลุ่มของดินตามชนิดของดินของทั้งสองรูปให้ผลที่คล้ายกันนั่นคือสามารถจำแนกกลุ่มของดินได้ 3 กลุ่มแต่ไม่สอดคล้องกับชนิดของดิน โดยกลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของแคลเซียม (Ca) ปนอยู่ปานกลางและคาร์บอน (C) ปนอยู่บ้าง และกลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของแคลเซียม (Ca) สูงและคาร์บอน (C) ปนอยู่เล็กน้อย และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) และแคลเซียม (Ca) สูง ในขณะที่กลุ่มของดินทั้ง 3 กลุ่มที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 คือกลุ่มที่ 1 เป็นกลุ่มของดินที่มีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้าง กลุ่มที่ 2 เป็นกลุ่มของดินที่มีส่วนประกอบของแมกนีเซียม (Mg) กับไทเทเนียม (Ti) สูงและคาร์บอน (C) ปนอยู่บ้างแต่มากกว่ากลุ่มที่ 1 และกลุ่มที่ 3 เป็นกลุ่มของดินที่มีส่วนประกอบของคาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) สูง แต่ประเด็นที่น่าสนใจคือผลที่ได้จากกราฟ 2 มิติของค่าองค์ประกอบหลักดังกล่าวด้วยค่าประมาณแบบอันดับจากข้อมูลแปลงนี้ยังให้ผลการวิเคราะห์คล้ายกับกราฟ 2 มิติของค่าองค์ประกอบหลักดังกล่าวที่ได้จากค่าประมาณแบบภาวะน่าจะเป็นสูงสุด และค่าประมาณแบบเครื่องหมายจากข้อมูลแปลง ดังนั้นจึงอาจกล่าวได้ว่าการแปลงข้อมูลด้วยวิธีของ Aitchison มีให้ผลการจำแนกกลุ่มของดินที่คล้ายคลึงกัน

ตารางที่ 26 ตัวแปรที่มีค่าสัมประสิทธิ์สูงในแต่ละองค์ประกอบหลักที่ 1-3 โดยจำแนกตามตัวประมาณค่าเมทริกซ์ความแปรปรวนร่วมบนข้อมูลดิบ และข้อมูลแปลงด้วยวิธีของ Aitchison

ตัวประมาณค่าเมทริกซ์ความแปรปรวนร่วม	ข้อมูลดิบ			ข้อมูลแปลงด้วยวิธีของ Aitchison		
	PC1	PC2	PC3	PC1	PC2	PC3
ตัวประมาณค่าแบบภาวะน่าจะเป็นสูงสุด	Si,C	Si,C	Al	C	Ca	Mg,Ti
ความแปรปรวนรวม (%)	86.4	97.4	98.6	61.1	85.7	93.8
ตัวประมาณค่าแบบเครื่องหมาย	Si,C	Si,C	Al	C	Ca	Mg,Ti
ความแปรปรวนรวม (%)	76.3	84	90.6	59.9	83.6	90.1
ตัวประมาณค่าแบบอันดับ	Si,C	Si,C	Al	C	Ca	Mg,Ti
ความแปรปรวนรวม (%)	75.5	88.1	92.8	51.4	77.8	87.9

อภิปรายผล

ผลการวิเคราะห์ข้อมูลใช้วิธีวิเคราะห์องค์ประกอบหลัก 3 วิธีคือการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุด การวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายและการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับ โดยการพลอตค่าองค์ประกอบหลักที่ 1 กับ 2 และ 1 กับ 3 ภายใต้ข้อมูล 2 ลักษณะคือ ข้อมูลดิบ (raw data) และข้อมูลแปลงด้วยวิธีของ Aitchison มีประเด็นที่น่าสนใจดังนี้

1. องค์ประกอบหลักที่ 1, 2 และ 3 ที่ได้จากการวิเคราะห์องค์ประกอบหลักทั้ง 3 วิธีบนข้อมูลดิบเป็นตัวแทนของตัวแปรตัวเดียวกันคือ องค์ประกอบหลักที่ 1 กับ 2 เป็นตัวแทนของตัวแปรซิลิคอน (Si) กับคาร์บอน (C) และองค์ประกอบหลักที่ 3 เป็นตัวแปรอลูมินัม (Al) ในขณะที่องค์ประกอบหลักที่ได้จากการวิเคราะห์องค์ประกอบหลักทั้ง 3 วิธีบนข้อมูลแปลง มีองค์ประกอบหลักที่ 1 เป็นตัวแทนของตัวแปรตัวเดียวกันคือตัวแปรคาร์บอน (C) เพียงตัวเดียว องค์ประกอบหลัก

ที่ 2 เป็นตัวแทนของตัวแปรตัวเดียวกันคือตัวแปรแคลเซียม (Ca) และองค์ประกอบหลักที่ 3 เป็นตัวแทนของตัวแปรตัวเดียวกันคือตัวแปรแมกนีเซียม (Mg) และไทเทเนียม (Ti)

2. เมื่อทำการเปรียบเทียบผลจากการวิเคราะห์ความสามารถของการอธิบายความแปรปรวนในข้อมูลของผลรวม 3 องค์ประกอบแรกของข้อมูล 2 ลักษณะคือข้อมูลดิบ และข้อมูลแปลงด้วยวิธีของ Aitchison พบว่า

2.1 ผลจากวิเคราะห์ความสามารถของการอธิบายความแปรปรวนของข้อมูลของผลรวม 3 องค์ประกอบแรกจากวิเคราะห์องค์ประกอบหลักทั้ง 3 วิธีบนข้อมูลดิบพบว่า ผลรวมของ 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบภาวะน่าจะเป็นสูงสุดสามารถอธิบายความแปรปรวนในข้อมูลได้มากที่สุดคือ 98.6% รองลงมาคือผลรวมของ 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบอันดับสามารถอธิบายความแปรปรวนในข้อมูลได้ 92.8% และผลรวมของ 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบเครื่องหมายสามารถอธิบายความแปรปรวนในข้อมูลได้น้อยที่สุดคือ 90.6%

2.2 ผลจากการวิเคราะห์ความสามารถของการอธิบายความแปรปรวนในข้อมูลของผลรวม 3 องค์ประกอบแรกจากวิเคราะห์องค์ประกอบหลักทั้ง 3 วิธีบนข้อมูลแปลงพบว่า ผลรวมของ 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบภาวะน่าจะเป็นสูงสุดสามารถอธิบายความแปรปรวนในข้อมูลได้มากที่สุดคือ 93.8% รองลงมาคือผลรวมของ 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบเครื่องหมายสามารถอธิบายความแปรปรวนในข้อมูลได้ 90.1% และผลรวมของ 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบอันดับสามารถอธิบายความแปรปรวนในข้อมูลได้น้อยที่สุดคือ 87.9%

3. เมื่อทำการเปรียบเทียบผลจากการวิเคราะห์ความสามารถของการอธิบายความแปรปรวนของข้อมูลของผลรวม 3 องค์ประกอบแรกของทั้ง 3 วิธีพบว่า ผลรวม 3 องค์ประกอบแรกที่ได้จากค่าประมาณแบบภาวะน่าจะเป็นสูงสุดสามารถอธิบายความแปรปรวนของข้อมูลได้มากที่สุดทั้งบนข้อมูลดิบและข้อมูลแปลง

4. ผลการทดสอบการเท่ากันของค่าไอเกนที่ได้จากการวิเคราะห์องค์ประกอบหลักทั้ง 3 วิธีทั้งบนข้อมูลดิบและข้อมูลแปลง พบว่าที่ระดับนัยสำคัญ 0.05 ค่าไอเกนที่ 1-8 ของทุกกรณีมีความแตกต่างกัน

5. ผลจากการวิเคราะห์กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 ของการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดบนข้อมูลดิบ สามารถจำแนกกลุ่มได้ค่อนข้างชัดเจน 2 กลุ่ม แต่ลักษณะการแบ่งกลุ่มไม่สอดคล้องกับชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา และจะเห็นได้ว่าตัวแปรที่มีความสำคัญในการแบ่งกลุ่มของดินที่ได้ทั้ง 2 กลุ่มคือ ซิลิคอน

(Si) และคาร์บอน (C) โดยทั้งสองกลุ่มจะมีความแตกต่างกันในปริมาณส่วนประกอบของตัวแปรดังกล่าว ในขณะที่ผลที่ได้จากการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายและค่าประมาณแบบอันดับบนข้อมูลดิบ ให้ผลการวิเคราะห์ที่คล้ายกันคือ สามารถทำให้เห็นการแบ่งกลุ่มของตัวอย่าง ค่อนข้างชัดเจน 3 กลุ่ม แต่ลักษณะการแบ่งกลุ่มไม่สอดคล้องกับชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา และจะเห็นได้ว่าตัวแปรที่มีความสำคัญในการแบ่งกลุ่มของดินที่ได้ทั้ง 3 กลุ่มคือ ซิลิคอน (Si) และคาร์บอน (C) โดยแต่ละกลุ่มจะมีความแตกต่างกันในปริมาณส่วนประกอบของตัวแปรดังกล่าวเช่นกัน

6. ผลจากการวิเคราะห์กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 ของการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบภาวะน่าจะเป็นสูงสุดบนข้อมูลดิบ สามารถจำแนกกลุ่มของดินได้ 2 กลุ่ม แต่ไม่ชัดเจนนัก โดยผลของการวิเคราะห์ทำให้อาจกล่าวได้ว่ากลุ่มของดินเหนียวจะมีอลูมินัม (Al) สูงกว่าดินอีก 4 ชนิดที่เหลือ ในขณะที่ผลที่ได้จากการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบเครื่องหมายและผลที่ได้จากการวิเคราะห์องค์ประกอบหลักด้วยค่าประมาณแบบอันดับบนข้อมูลดิบ จะให้ผลการวิเคราะห์ที่คล้ายกันคือจะไม่สามารถจำแนกกลุ่มของดินได้

7. ผลจากการวิเคราะห์กราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 และ 1 กับ 3 ของการวิเคราะห์องค์ประกอบหลักทั้ง 3 วิธีบนข้อมูลแปลงพบว่าทั้ง 3 วิธีให้ผลการวิเคราะห์ที่คล้ายกันคือ สามารถทำให้เห็นการแบ่งกลุ่มของตัวอย่าง ค่อนข้างชัดเจน 3 กลุ่ม แต่ลักษณะการแบ่งกลุ่มไม่สอดคล้องกับชนิดของดินทั้ง 5 ชนิดที่นำมาพิจารณา แต่มีประเด็นที่น่าสนใจคือตัวแปรที่มีความสำคัญในการแบ่งกลุ่มของดินที่ได้ทั้ง 3 กลุ่มจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 2 คือ คาร์บอน (C) และแคลเซียม (Ca) โดยแต่ละกลุ่มจะมีความแตกต่างกันในปริมาณส่วนประกอบของตัวแปรดังกล่าว ในขณะที่ตัวแปรที่มีความสำคัญในการแบ่งกลุ่มของดินที่ได้ทั้ง 3 กลุ่มจากกราฟ 2 มิติของค่าองค์ประกอบหลักที่ 1 กับ 3 คือ คาร์บอน (C) แมกนีเซียม (Mg) และไทเทเนียม (Ti) โดยแต่ละกลุ่มจะมีความแตกต่างกันในปริมาณส่วนประกอบของตัวแปรดังกล่าว นอกจากนี้อาจกล่าวได้ว่าการแปลงข้อมูลด้วยวิธีของ Aitchison สามารถทำให้เห็นการจำแนกกลุ่มอย่างชัดเจน

ข้อเสนอแนะจากงานวิจัย

1. การแปลงข้อมูลโดยใช้สมการที่ (1) สามารถทำให้เห็นการแบ่งกลุ่มของตัวอย่างดินค่อนข้างชัดเจน 3 กลุ่มทั้งในการวิเคราะห์องค์ประกอบหลักแบบภาวะน่าจะเป็นสูงสุดและการวิเคราะห์องค์ประกอบหลักแบบแกร่งบนเมตริกซ์ความแปรปรวนร่วม แต่ลักษณะการแบ่งกลุ่มไม่

สอดคล้องกับชนิดของดิน 5 ชนิดที่นำมาพิจารณา ดังนั้นจึงอาจพิจารณาปัจจัยทางสิ่งแวดล้อมอื่น เช่น ลักษณะทางกายภาพของดิน (สี ค่า pH) ปริมาณโลหะหนัก หรือปริมาณจุลธาตุอาหารในดิน

2. งานวิจัยในครั้งต่อไปอาจพิจารณาการวิเคราะห์องค์ประกอบหลัก (PCA) บนเมทริกซ์สหสัมพันธ์ (Correlation Matrix)

บรรณานุกรม

ภาษาไทย

กมลชนก พานิชการ. “การวิเคราะห์ตัวแปรพหุในการจำแนกกลุ่มดินตามองค์ประกอบทางเคมี.” งานวิจัยฉบับสมบูรณ์ทุนอุดหนุนการวิจัย ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร, 2550. (อัดสำเนา)

นัทธีรา และคณะ. “รายงานผลของปริมาณกรดฮิวมิกและกรดฟัลวิคของดินในพื้นที่เกษตรกรรมต่อระดับจุลธาตุอาหารในดิน และโลหะที่เป็นพิษกรณีศึกษาภาคตะวันตกของประเทศไทย.” สถาบันวิจัยและพัฒนามหาวิทยาลัยศิลปากร, 2551. (อัดสำเนา)

ปราณี นิลกรณ์. “องค์ประกอบหลัก.” เอกสารประกอบการสอนวิชา 515 514 ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร, 2547. (อัดสำเนา)

ภานุพงษ์ พนมวัน . “การวิเคราะห์องค์ประกอบหลักและการวิเคราะห์กลุ่มสำหรับองค์ประกอบทางเคมีของลูกปัดแก้วที่ได้จากการขุดค้นแหล่งโบราณคดีพรหมทินใต้ ลพบุรี.” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติประยุกต์ มหาวิทยาลัยศิลปากร, 2547.

มหาวิทยาลัยเกษตรศาสตร์. คณะเกษตร. ภาควิชาปฐพีวิทยา. ปฐพีวิทยาเบื้องต้น. พิมพ์ครั้งที่ 9. กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์, 2544.

ภาษาอังกฤษ

Aitchison, J. “Principal Component Analysis of Composition Data.” Biometrika 70(1983) : 57-65.

Baxter, M.J. “An empirical study of principal component and correspondence analysis of glass compositions.” Archaeometry 33(1991): 29-41.

Baxter, M.J. “Statistical analysis of chemical compositional data and the comparison of analysis.” Archaeometry 34(1992): 267-277.

Baxter, M.J. “Standardization and Transformation in Principal Component Analysis with Applications to Archeology.” Applied Statistics 44(1995): 513-527.

Chaudhuri, P. “Ona geometric notion of quantile for multivariate data.” J Amer. Statist. Assoc. 91(1996): 862-872.

Koltchinski, V.I. “M-estimation, convexity and quantiles.” Ann. Statist. 25(1997): 435-477.

Marden, I.J. “Some robust estimates of principal components.” Statistics and Probability Letters 43(1999): 349-359.

Möttönen, J., Oja, H. and Tienari, J. “On the efficiency of multivariate spatial sign and rank test.” Ann. Statist. 25(1997): 542-552.

Rencher, Alvin C. Method of Multivariate Analysis. 2nd ed. New York: John Wiley & Sons Inc., 1996.

Small, C.G. “A survey of multidimensional median.” Int. Statist. Rev. 58(1990): 263-277.

ภาคผนวก
ข้อมูลที่ใช้ในงานวิจัย

ตารางที่ 27 ข้อมูลองค์ประกอบทางเคมีของดินขนาด 65 ไมครอนจำนวน 58 ตัวอย่าง

location	soiltype *	O	Mg	Al	Si	K	Ca	Fe	C	Ti	Na
1	1	52.19	0.88	10.41	29.86	1.83	0.4	4.23	0	0.21	0
2	1	55.7	0.76	9.21	27.39	1.71	0.64	4.12	0	0.47	0
3	1	56.15	0.79	9.23	26.67	1.62	0.64	3.9	0	0.4	0
4	1	55.79	0.68	8.11	23.75	1.53	0.66	3.43	5.42	0.43	0
5	1	55.24	0.59	8.96	25.77	1.63	0.69	3.88	2.77	0.37	0
6	1	54.93	0.69	8.56	23.24	1.48	0.77	3.48	5.52	0.37	0
7	1	55.41	0.63	8.44	23.58	1.4	0.61	3.19	5.72	0.22	0
8	1	50.39	0.29	10.05	29.99	1.97	0.75	4.68	1.87	0	0
9	1	56.37	0.82	9.21	26.24	1.55	0.84	4.06	0	0.41	0.09
10	1	52.05	0.34	10.32	30.3	1.99	0	4.99	0	0	0
11	1	55.08	0.85	8.97	25.03	1.53	0.81	3.9	2.85	0.39	0
12	1	54.63	0.75	8.33	24.85	1.5	0.58	3.78	5.03	0.36	0
13	2	50.76	0.82	8.74	31.75	1.67	0.91	5.08	0	0.16	0
14	2	55.42	0.83	8.23	26.08	1.41	1.33	3.98	1.76	0.37	0.1
15	1	52.16	0.79	10	29.43	1.79	0.98	4.41	0	0.44	0

ตารางที่ 27 (ต่อ)

location	soiltype *	O	Mg	Al	Si	K	Ca	Fe	C	Ti	Na
16	3	55.01	0.68	7.92	24.52	1.71	0.83	2.78	6.12	0.35	0
17	1	55.36	0.54	9.92	25.79	1.65	0.37	3.32	18.02	0.4	0
18	3	55.42	0.66	8.92	25.95	1.86	0.46	3.52	2.77	0.42	0
19	1	52.35	0.96	9.71	29.41	1.75	0.78	4.57	0	0.45	0
20	2	56.3	0.76	9.09	27.36	1.65	0.68	3.75	0	0.41	0
21	2	55.85	0.71	8.32	25.76	1.42	0.98	3.6	2.65	0.41	0
22	2	54.44	0.63	7.7	24.79	1.44	0.69	3.5	6.27	0.37	0.09
23	2	52.26	0.82	8.94	29.02	1.65	0.89	4.92	0	0.51	0
24	2	55.76	0.7	7.87	23.96	1.33	0.83	3.33	5.89	0.33	0
25	2	56.17	0.81	8.22	26.14	1.34	0.69	3.7	2.6	0.34	0
26	2	54.76	0.84	8.31	25.85	1.53	0.8	4.29	2.59	0.42	0
27	2	56.76	0.96	8.49	26.98	1.49	0.8	3.84	0	0.4	0.26
28	1	56.32	0.73	7.55	22.97	1.27	0.57	2.86	6.96	0.22	0
29	1	52.96	0.94	9.74	28.93	1.74	0.69	4.47	0	0.51	0
30	1	54.54	0.79	8.08	23.25	1.39	0.6	3.8	7.14	0.43	0

ตารางที่ 27 (ต่อ)

location	soiltype *	O	Mg	Al	Si	K	Ca	Fe	C	Ti	Na
31	2	54.4	0.67	7.97	23.77	1.3	0.44	3.15	7.74	0.38	0.17
32	1	52.72	0.95	9.91	29.24	1.79	0.51	4.61	0	0.44	0
33	1	54.9	0.59	8.39	25.16	1.59	0.34	3.53	4.28	0.37	0
34	1	54.51	0.83	8.08	24.4	1.47	0.44	4.03	5.12	0.38	0.09
35	1	54.58	0.92	8.08	23.93	1.89	0.29	3.84	7.25	0.35	0.09
36	1	52.71	0.87	9.57	29.39	1.63	0.66	4.8	0	0.37	0
37	1	55.72	0.74	8.37	23.83	1.45	0.49	3.17	5.9	0.33	0
38	1	53.26	0.76	7.94	23.35	1.42	0.52	3.68	7.66	0.36	0
39	3	55.47	0.9	8.83	25.44	2.08	0.51	3.65	2.53	0.39	0.3
40	3	54.16	0.83	8.38	24.53	2.02	0.71	3.46	5.44	0.39	0
41	3	52.45	0.9	9.4	29.27	2.37	0.73	4.36	0	0.52	0
42	3	53.63	0.72	7.79	24.01	1.95	0.42	3.2	6.83	0.36	0
43	3	56.29	0.9	9.07	27.72	2.18	0.47	3.11	0	0.26	0
44	3	55.49	0.74	7.96	24.6	1.86	0.4	3.19	5.46	0.32	0
45	3	53.08	0.88	9.39	29.33	2.29	0.29	4.23	0	0.31	0

ตารางที่ 27 (ต่อ)

location	soiltype *	O	Mg	Al	Si	K	Ca	Fe	C	Ti	Na
46	3	56.47	0.85	8.63	27.22	2.13	0.41	3.62	0	0.43	0.23
47	1	53.53	0.93	7.97	22.82	1.75	0.46	3.73	8.35	0.39	0
48	3	54.44	0.89	8.12	24.74	1.93	0.41	3.83	5.21	0.42	0
49	3	53.72	0.84	7.79	23.4	1.93	0.75	3.73	7.4	0.44	0
50	4	55.66	0.87	8.25	25.89	2.12	0.45	3.63	2.69	0.35	0.1
51	3	51.83	1.17	9.1	29.5	2.35	0.15	4.89	0	0.53	0.38
52	3	52.42	1.12	8.6	30.25	2.23	0.17	4.61	0	0.45	0.15
53	3	56.91	0.95	8.52	26.81	2	0.34	3.87	0	0.38	0.22
54	2	55.22	0.98	8.43	24.59	1.9	0.2	3.85	4.53	0.31	0
55	3	54.81	1.05	8.53	27.02	2.19	0.11	4.26	1.25	0.43	0.24
56	5	55.29	0.63	8.22	26.38	2.41	0.7	3.53	2.38	0.46	0
57	5	55.05	0.63	7.24	24.93	2.15	0.62	3.06	6.02	0.29	0
58	5	53.67	0.58	8.92	22.71	2.05	0.51	2.98	8.14	0.35	0.09

หมายเหตุ * หมายเลข 1 แทนดินเหนียว หมายเลข 2 แทนดินร่วนเหนียว หมายเลข 3 แทนดินร่วน หมายเลข 4 แทนดินร่วนปนทรายแป้ง และหมายเลข 5 แทนดินร่วนปนทราย

ประวัติผู้วิจัย

ชื่อ-สกุล นายวัชรพงษ์ ว่องนิยมเกษตร

ที่อยู่ 9/4 หมู่ 8 ตำบลปากแพรก อำเภอเมืองกาญจนบุรี
จังหวัดกาญจนบุรี 71000

ที่ทำงาน โรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ วิทยาเขต
กำแพงแสน ศูนย์วิจัยและพัฒนาการศึกษา ตำบลกำแพงแสน
อำเภอกำแพงแสน จังหวัดนครปฐม โทรศัพท์ (034) 351396

ประวัติการศึกษา

พ.ศ. 2545	สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาคณิตศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
พ.ศ. 2547	ศึกษาต่อระดับปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ประวัติการทำงาน

พ.ศ. 2545-2547	อาจารย์โรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน ศูนย์วิจัยและพัฒนาการศึกษา
พ.ศ. 2549- ปัจจุบัน	อาจารย์โรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน ศูนย์วิจัยและพัฒนาการศึกษา