

การวิเคราะห์ข้อมูลด้วยสถิติหลายตัวแปรทางชีววิทยาเบื้องต้น (Multivariate Analysis of Biological data)

ผศ.ดร. บุญเสฐียร บุญสูง

ภาควิชาสัตววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

งานวิจัยทางชีววิทยาส่วนใหญ่มักจะมีตัวแปรหลายตัว ซึ่งตัวแปรเหล่านั้นมีความสัมพันธ์กันไม่มากก็น้อย เทคนิคการหาความสัมพันธ์ของตัวแปรหลายตัว (multivariate analysis) นั้นมีหลายเทคนิค ในที่นี้จะขอกกล่าวถึงเทคนิคที่นิยมใช้กันมากในทางชีววิทยา คือ

เทคนิคการจัดกลุ่ม (cluster analysis)

เป็นเทคนิคที่ใช้ในการแบ่งหรือจัดกลุ่ม Case (rows) หมายถึง สถานี พื้นที่ ฯลฯ ออกเป็นกลุ่มย่อยโดยให้ Case (rows) ที่มีความคล้ายกันอยู่ในกลุ่มเดียวกัน โดยใช้ลักษณะหรือตัวแปรที่คาดว่าจะทำให้ Case (rows) ต่างกัน ในทางนิเวศวิทยาส่วนมากใช้ species composition ในแต่ละ Case (rows) ซึ่งอาจเป็นข้อมูลความชุกชุม หรือข้อมูลพบ/ไม่พบชนิดของสิ่งมีชีวิต

เทคนิคการจัดอันดับ (ordination analysis)

เป็นเทคนิคการจับกลุ่มหรือรวมตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่มหรือ Factor เดียวกัน ตัวแปรที่อยู่ใน Factor เดียวกันจะมีความสัมพันธ์กันมาก โดยความสัมพันธ์นั้นอาจจะเป็นในทิศทางบวก (ไปในทางเดียวกัน) หรือทิศทางลบ (ไปในทางตรงข้ามกัน) ก็ได้ ส่วนตัวแปรที่อยู่คนละ Factor จะไม่มีความสัมพันธ์กัน หรือมีความสัมพันธ์กันน้อยมาก

การวิเคราะห์ข้อมูลด้วยสถิติหลายตัวแปรทางชีววิทยาเบื้องต้นโดยใช้โปรแกรมสำเร็จรูป PC-ORD for Windows

1. โปรแกรม PC-ORD for Windows

PC-ORD เป็นโปรแกรม Windows สำหรับวิเคราะห์ข้อมูลหลายตัวแปรทางนิเวศวิทยา โปรแกรมนี้เป็นเครื่องมือในการทดสอบ non-parametric และนำเสนอในรูปแบบกราฟ มีการสุ่มเพื่อวิเคราะห์ข้อมูลทางนิเวศวิทยา นอกจากนี้ยังสามารถใช้แปลงข้อมูล และจัดการแฟ้มข้อมูลได้ โปรแกรม PC-ORD มีเทคนิคการจัดอันดับ และเทคนิคการจัดกลุ่มหลายวิธี เช่น CCA, DCA, MRPP, perMANOVA, two-way clustering, TWINSpan, Beals smoothing, diversity indices, species lists, Mantel test, various rotation methods, 3D ordination graphics, indicator species analysis, Bray-Curtis ordination, city-block distance measures, species-area curves, tree data summaries, publication-quality dendrograms และ autopilot NMS โปรแกรมนี้สามารถวิเคราะห์ข้อมูลที่มีจำนวนมากได้ถึง 32,000 rows/columns

ขั้นตอนหลักในการวิเคราะห์ข้อมูลด้วยสถิติหลายตัวแปรทางชีววิทยาเบื้องต้นโดยใช้โปรแกรมสำเร็จรูป PC-ORD มี 4 ขั้นตอน ดังนี้

- 1) การเตรียมข้อมูล (Prepare Your Data)
- 2) เปิดโปรแกรม PC-ORD และแปลงข้อมูล (Open PC-ORD and your data files)
- 3) วิเคราะห์ข้อมูล (Run an analysis)
- 4) การแสดงผลกราฟ (Explore your results graphically)

2. การเตรียมข้อมูล (Prepare Your Data)

ข้อมูลที่จะนำไปวิเคราะห์ด้วยสถิติหลายตัวแปรด้วยโปรแกรม นั้น จะต้องเตรียมข้อมูลใน Excel ก่อน และต้องจัดเก็บแฟ้มข้อมูลเป็นแบบ Lotus 1-2-3 (*.wk1) โดยเตรียมข้อมูลใน Excel (ดังภาพ) ดังนี้

row 1 ให้ใส่จำนวนสถานี/พื้นที่ หรือตัวแปรที่เราต้องการจัดกลุ่มหรือวิเคราะห์องค์ประกอบ

row 2 ให้ใส่จำนวนชนิดของนก หรือตัวแปรที่เราต้องนำมาใช้วิเคราะห์หรือทดสอบ

row 3 ให้ใส่ Q สำหรับข้อมูลที่เป็น quantitative (abundance) หรือ C สำหรับข้อมูลที่เป็น categorical (binary, 0/1, presence/absence)

row 4 ให้ใส่อักษรย่อชนิดของนก (species) (ไม่เกิน 8 ตัวอักษร)

row 5 (column 1) ให้ใส่อักษรย่อของสถานีหรือพื้นที่ (samples) (ไม่เกิน 8 ตัวอักษร)

ตัวอย่าง ข้อมูลสำรวจชนิดของนก 10 ชนิดในบริเวณพื้นที่ที่แตกต่างกัน 5 แบบใหญ่ 10 แบบย่อย ซึ่งข้อมูลอาจเป็นข้อมูลความชุกชุมของนก (abundance data) หรืออาจเป็นข้อมูลพบ/ไม่พบ (presence/absence data) ถ้าพบนกชนิดนั้นให้กรอกเป็น 1 ถ้าไม่พบนกชนิดนั้นให้กรอกเป็น 0

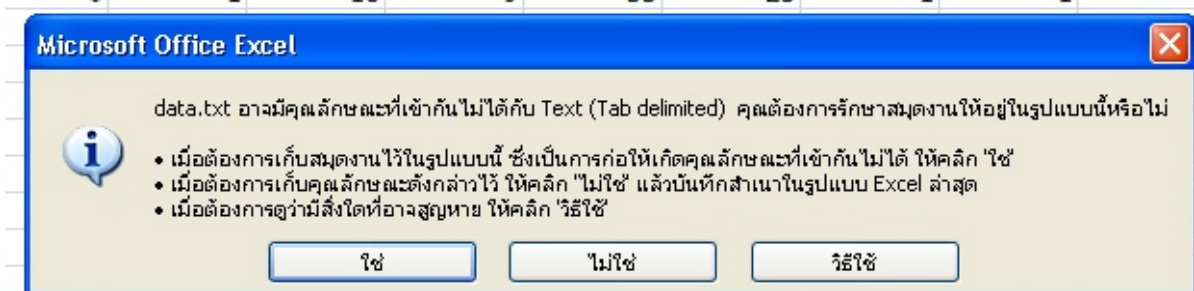
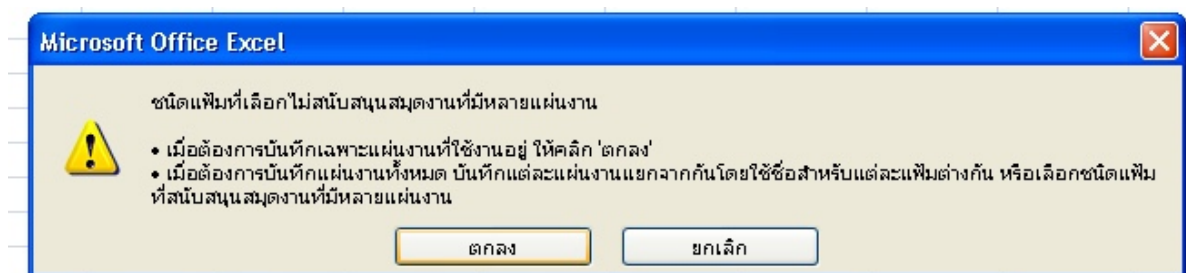
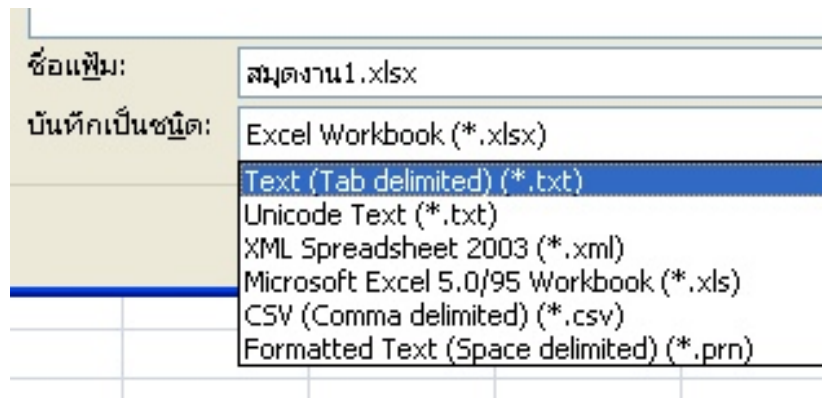
ข้อมูล abundance

	A	B	C	D	E	F	G	H	I	J	K
1	10 sites										
2	10 species										
3		Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
4		spA	spB	spC	spD	spE	spF	spG	spH	spI	spJ
5	Forest1	25	22	42	10	8	2	0	5	2	2
6	Forest2	18	19	38	9	14	3	0	7	2	3
7	City1	0	0	1	18	9	33	25	1	1	1
8	City2	1	0	0	24	12	35	37	0	1	1
9	Planta1	0	2	11	46	37	0	0	0	0	0
10	Planta2	1	2	10	41	47	2	0	0	0	0
11	Wetland1	1	3	1	23	0	0	0	7	2	0
12	Wetland2	2	2	2	30	0	0	0	6	5	0
13	Wetland3	3	1	0	25	5	2	12	2	2	1
14	Seashore	1	1	0	11	9	0	2	38	24	20
15											

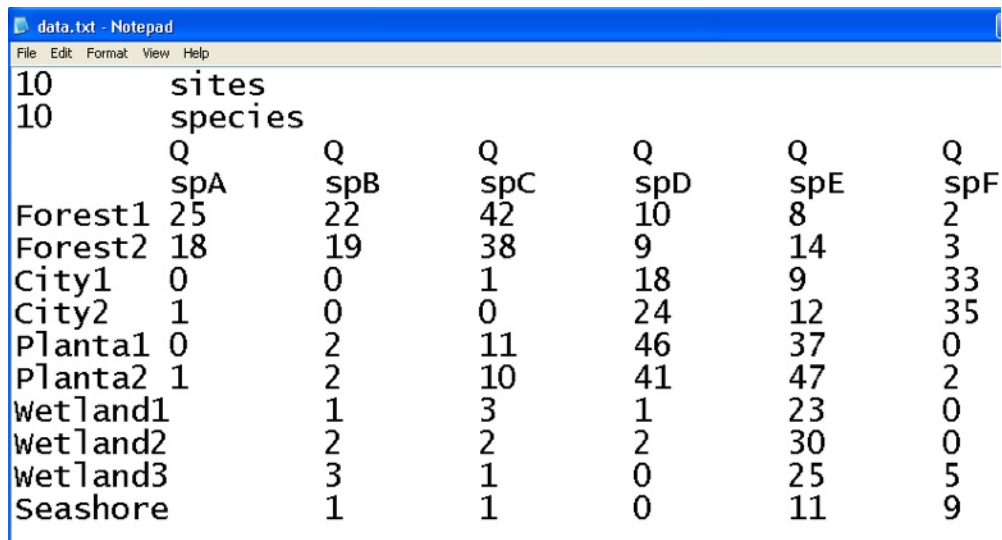
ข้อมูล presence/absence (binary, 0/1)

	A	B	C	D	E	F	G	H	I	J	K
1	10 sites										
2	10 species										
3		C	C	C	C	C	C	C	C	C	C
4		spA	spB	spC	spD	spE	spF	spG	spH	spI	spJ
5	Forest1	1	1	1	1	1	1	0	1	1	1
6	Forest2	1	1	1	1	1	1	0	1	1	1
7	City1	0	0	1	1	1	1	1	1	1	1
8	City2	1	0	0	1	1	1	1	0	1	1
9	Planta1	0	1	1	1	1	0	0	0	0	0
10	Planta2	1	1	1	1	1	1	0	0	0	0
11	Wetland1	1	1	1	1	0	0	0	1	1	0
12	Wetland2	1	1	1	1	0	0	0	1	1	0
13	Wetland3	1	1	0	1	1	1	1	1	1	1
14	Seashore	1	1	0	1	1	0	1	1	1	1

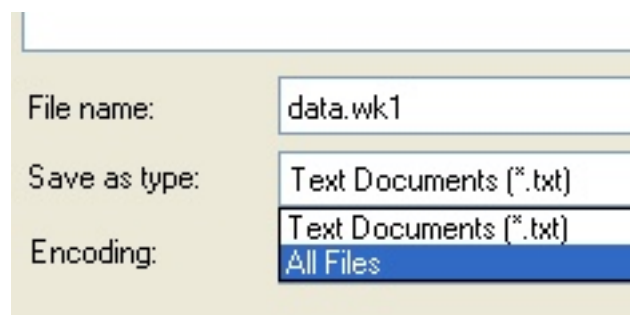
เนื่องจาก Microsoft Office Excel 2007 ไม่มีในส่วนการจัดเก็บแฟ้มเป็นแบบ *.wk1 จึงต้องบันทึกแฟ้มเป็น *.text ก่อน ซึ่งแฟ้มที่จัดเก็บนี้ไม่ได้เป็นแบบปกติ ดังนั้นเมื่อ Excel ถามให้ตอบตกลงและใช่ ดังภาพ



จากนั้นให้เปิดโปรแกรม Notepad เพื่อเปิดแฟ้ม *.text แล้วไปที่ File คำสั่งย่อย Save as ให้เลือก Save as type เป็น All Files และให้พิมพ์ชื่อไฟล์ใหม่แล้วตามด้วย .wk1 ดังภาพ

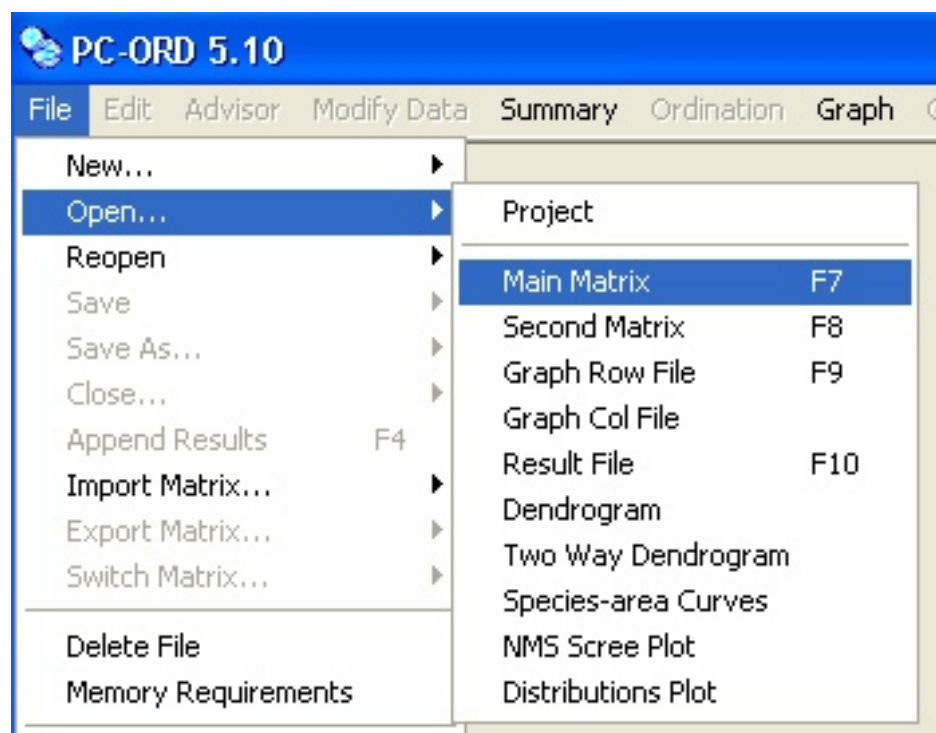


sites	Q	Q	Q	Q	Q	Q
species	spA	spB	spC	spD	spE	spF
Forest1	25	22	42	10	8	2
Forest2	18	19	38	9	14	3
City1	0	0	1	18	9	33
City2	1	0	0	24	12	35
Planta1	0	2	11	46	37	0
Planta2	1	2	10	41	47	2
Wetland1		1	3	1	23	0
Wetland2		2	2	2	30	0
Wetland3		3	1	0	25	5
Seashore		1	1	0	11	9

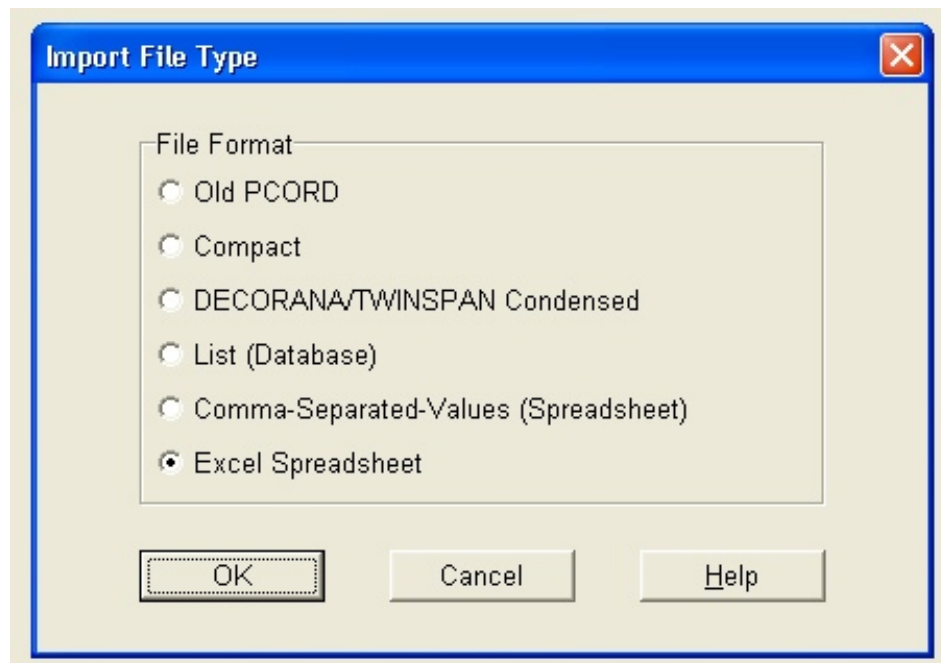
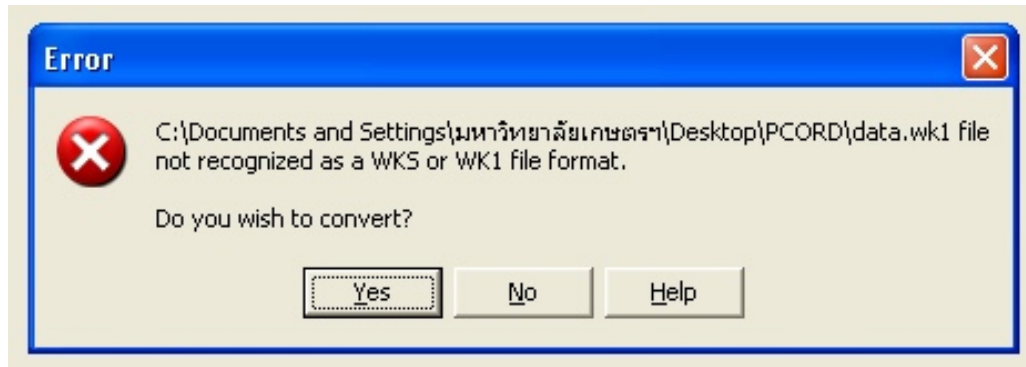


2. เปิดโปรแกรม PC-ORD และแปลงข้อมูล (Open PC-ORD and your data files)

เมื่อเปิดโปรแกรม PC-ORD for Windows ขึ้นมา ไปที่คำสั่ง File > Open คำสั่งย่อย Main Matrix

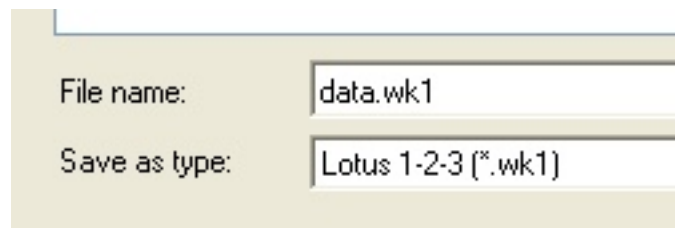
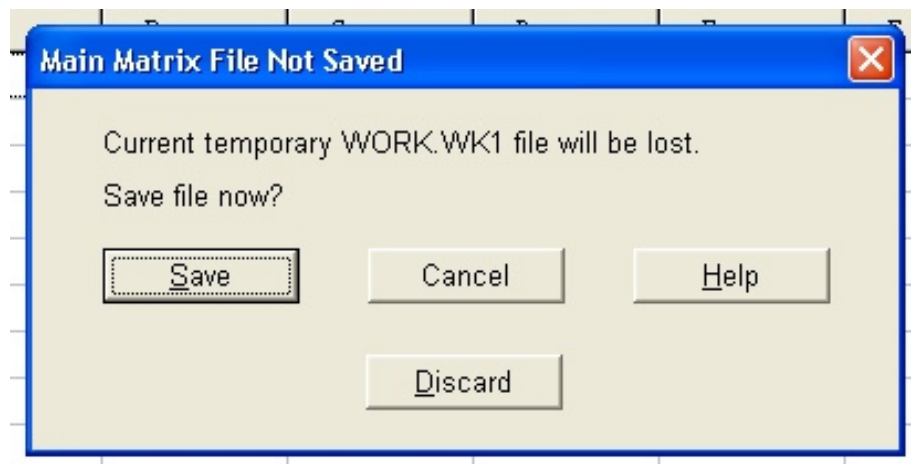
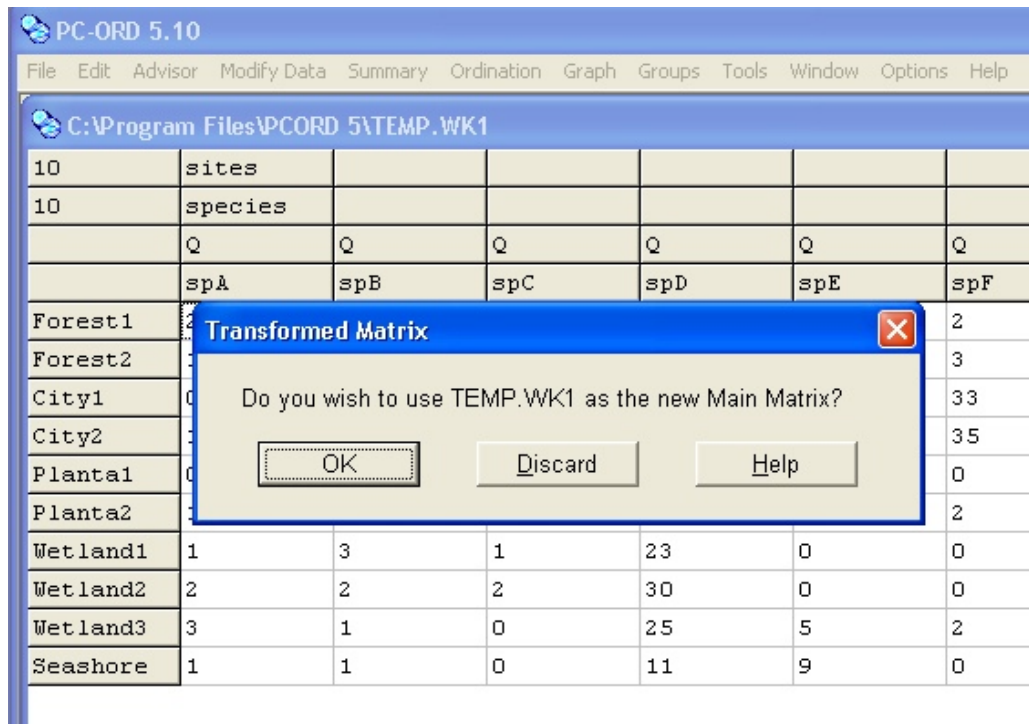


ทำการเปลี่ยนไฟล์ที่เปิดให้เป็น wk1 file format อีกครั้ง โดยเลือก Yes ในการแปลงไฟล์ และเลือก file format เป็น Excel Spreadsheet แล้วกด **OK**



เมื่อเปิดไฟล์ได้แล้วโปรแกรมจะทำการบันทึกไฟล์อัตโนมัติเป็น TEMP.WK1 กด OK เพื่อใช้ไฟล์นี้เป็น Main Matrix จากนั้นบันทึกไฟล์และตั้งชื่อใหม่ ดังภาพ

หากมี Error เกิดขึ้นในช่วงที่เปิดไฟล์ ให้กลับไปตรวจสอบการกรอกข้อมูลใน Excel Spreadsheet อีกครั้งว่ามีความผิดพลาดหรือไม่



เนื่องจากข้อมูลเป็น abundance data จำเป็นต้องแปลงข้อมูล (transformation) ก่อนนำไปวิเคราะห์ ถ้าข้อมูลเป็น presence/absence data สามารถนำไปวิเคราะห์ข้อมูลได้เลย การแปลงข้อมูลทำได้โดยไปที่คำสั่ง Modify Data > Relativizations > Main Matrix > General Relativization เมื่อขึ้นหน้าต่าง Main Matrix General Relativization ใส่คำสั่ง p=1 แล้วกด OK แล้วทำการบันทึกไฟล์ใหม่ดังภาพ

PC-ORD 5.10

File Edit Advisor Modify Data Summary Ordination Graph Groups Tools Window Options Help

Main - data.wk1

	Q	Q	Q	Q	Q	Q	Q
	spA	spB	spC	spD	spE	spF	spG
Forest1	25	22	42	10	8	2	0
Forest2	18	19	38	9	14	3	0
City1	0	0	1	18	9	33	25
City2	1	0	0	24	12	35	37
Planta1	0	2	11	46	37	0	0
Planta2	1	2	10	41	47	2	0
Wetland1	1	3	1	23	0	0	0
Wetland2	2	2	2	30	0	0	0
Wetland3	3	1	0	25	5	2	12
Seashore	1	1	0	11	9	0	2

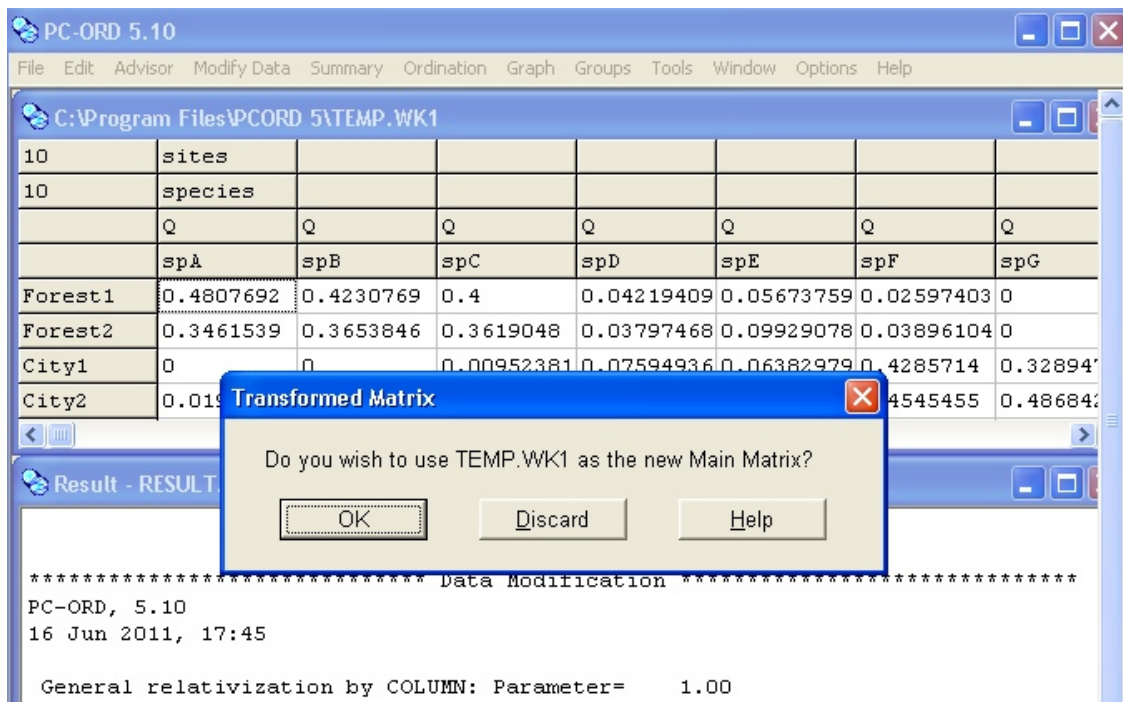
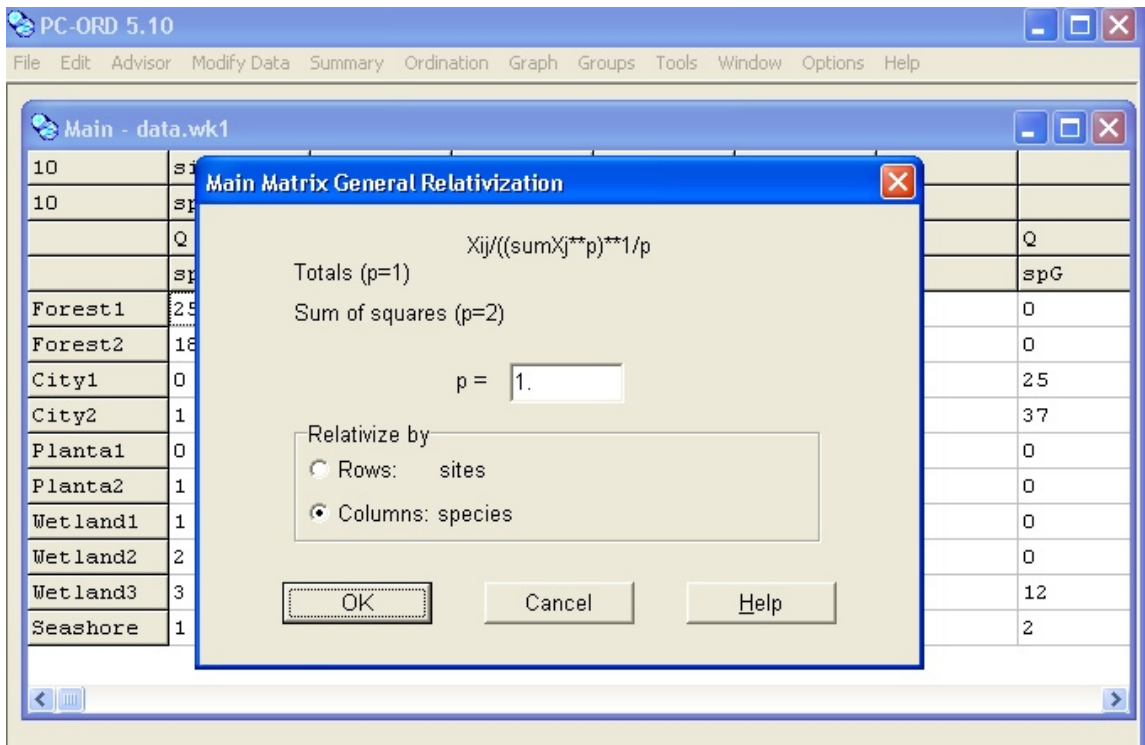
PC-ORD 5.10

File Edit Advisor **Modify Data** Summary Ordination Graph Groups Tools Window Options Help

Main - data

- Transformations...
- Relativizations...
 - Main Matrix
 - General Relativization
 - Relativization By Maximum
 - Adjust To Mean
 - Adjust To Standard Deviate
 - Binary W/Respect To Mean
 - Rank Adjustment
 - Binary W/Respect To Median
 - Variates By Ubiquity Row/Col
 - Info Function Of Ubiquity
 - Second Matrix
- Beals Smoothing
- Multiply Matrices
- Append Matrix
- Transpose Matrix
- Delete Rows...
- Delete Columns...
- Multiply Or Add A Constant
- Random Sample
- Shuffle Data

	Q	Q	Q	Q	Q	Q	Q
	spA	spB	spC	spD	spE	spF	spG
Forest1	25	22	42	10	8	2	0
Forest2	18	19	38	9	14	3	0
City1	0	0	1	18	9	33	25
City2	1	0	0	24	12	35	37
Planta1	0	2	11	46	37	0	0
Planta2	1	2	10	41	47	2	0
Wetland1	1	3	1	23	0	0	0
Wetland2	2	2	2	30	0	0	0
Wetland3	3	1	0	25	5	2	12
Seashore	1	1	0	11	9	0	2



The screenshot shows the PC-ORD 5.10 software interface. The 'File' menu is open, showing options like 'New...', 'Open...', 'Save As...', 'Close...', 'Append Results', 'Import Matrix...', 'Export Matrix...', 'Switch Matrix...', 'Delete File', 'Memory Requirements', 'Print...', 'Print Setup...', and 'Font...'. The 'Save As...' option is selected, and a sub-menu is visible with 'Project' selected. The main window displays a data table with columns 'Q' and 'spD' and 'spE'. The table contains numerical data for various cases.

	Q	Q	spD	spE
			0.0421940	0.056737
	9048		0.0379746	0.099290
	5238		0.0759493	0.063829
			0.1012658	0.085106
	0.0384615	0.1047619	0.1940928	0.262411
	0.0384615	0.0952381	0.1729958	0.333333
	0.0576923	0.0095238	0.0970464	0

3. วิเคราะห์ข้อมูล (Run an analysis)

3.1 การจัดกลุ่ม (cluster analysis)

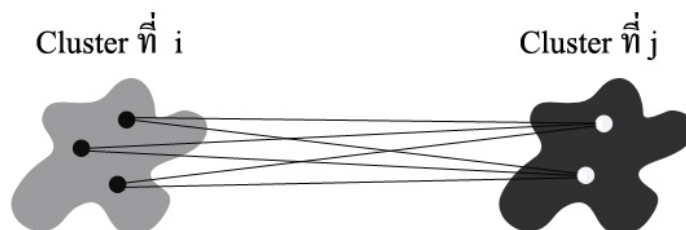
ดังที่กล่าวไปแล้วข้างต้นว่า หลักของเทคนิคการจัดกลุ่มจะใช้ในการแบ่งกลุ่ม Case (rows) ออกเป็น กลุ่มย่อยโดยให้ Case (rows) ที่มีความคล้ายกันอยู่ในกลุ่มเดียวกัน ซึ่งวัดด้วยระยะห่าง (distance หรือ dissimilarity measure) หรือความคล้ายกัน (similarity measure) เทคนิคการจัดกลุ่มแบ่งเป็นหลายเทคนิคย่อย เทคนิคที่นิยมใช้กันมากคือ Hierarchical Cluster Analysis ซึ่งเป็นเทคนิคที่ใช้แบ่งกลุ่ม Case (rows) หรือแบ่งกลุ่มตัวแปรโดยมีเงื่อนไขดังนี้

- จำนวน Case (rows) ควรต่ำกว่า 200 และจำนวนตัวแปรต้องไม่มากเกินไป
- ไม่จำเป็นต้องทราบจำนวนกลุ่มมาก่อน
- ไม่จำเป็นต้องทราบว่าตัวแปรใดหรือ Case (rows) ใดอยู่กลุ่มใดก่อน

กรณีที่วัดความคล้ายด้วยระยะห่าง ถ้าระยะห่างระหว่าง Case (rows) คู่ใดต่ำ แสดงว่า Case (rows) นั้นอยู่ใกล้กันหรือมีความคล้ายกันควรจะจัดอยู่ในกลุ่มหรือ cluster เดียวกัน สำหรับวิธีคำนวณ distance measure นั้นมีหลายวิธีขึ้นอยู่กับชนิดของข้อมูล เช่น ถ้าเป็นข้อมูล binary (0/1) นิยมวัดด้วย Jaccard distance เป็นต้น

ส่วนหลักเกณฑ์ในการรวมกลุ่ม (group linkage method) นั้นมีหลายวิธี วิธีที่นิยมใช้ในโปรแกรม PC-ORD คือ Group Average ซึ่งวิธีนี้จะคำนวณหาระยะห่างเฉลี่ยของทุกคู่ของ Case (rows) โดยที่ Case (rows) หนึ่งอยู่ใน Cluster ที่ i ส่วนอีก Case (rows) หนึ่งอยู่ใน Cluster ที่ j , $i \neq j$

ถ้า Cluster ที่ i มีระยะห่างเฉลี่ยจาก Cluster ที่ j สั้นกว่าระยะห่างจาก Cluster อื่น จะนำ Cluster ที่ i และ j มารวมเป็น Cluster เดียวกัน



การวิเคราะห์การจัดกลุ่มในโปรแกรม PC-ORD ทำได้โดยไปที่คำสั่ง Groups > Cluster Analysis เมื่อขึ้นหน้าต่าง Cluster Setup ในส่วน Distance Measure ให้เลือกเป็น Sørensen (Bray-Curtis) เป็นวิธีที่นิยมใช้กันมากในทางนิเวศวิทยาสำหรับข้อมูล abundance ส่วนข้อมูล binary ให้เลือก Jaccard และ Group Linkage Method ให้เลือก Group Average แล้วเลือกลักษณะของ Dendrogram ดังภาพ

The screenshot shows the PC-ORD 5.10 interface. The main window displays a data matrix with columns for sites, species, and various variables (Q, spA, spB, spC, spF, spG). A 'Cluster Analysis' menu is open, listing options like Two-way Cluster Analysis, MRPP, Blocked MRPP (MRBP), TWINSpan, Indicator Species Analysis, Mantel Test, and PerMANOVA.

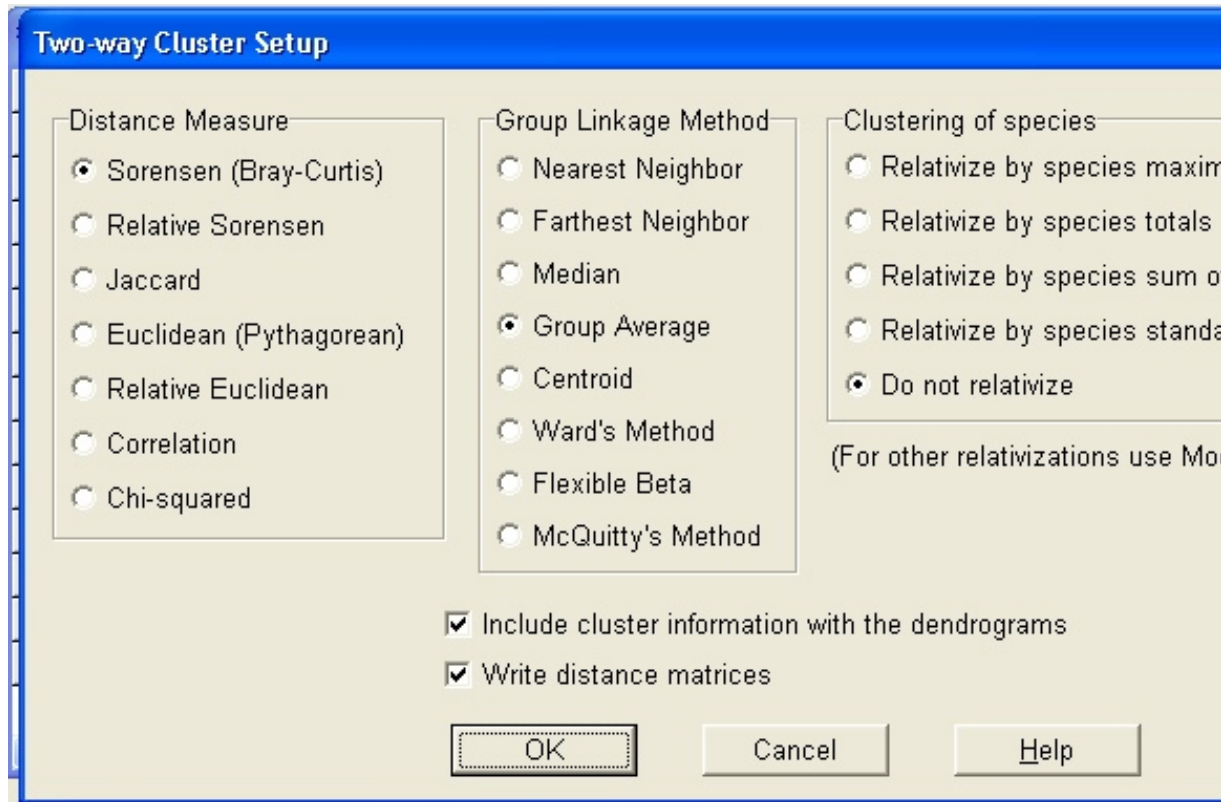
	Q	Q	Q				
	spA	spB	spC		spF	spG	
Forest1	0.4807692	0.4230769	0.4		0.0259740	0	
Forest2	0.3461539	0.3653846	0.3619048	0.0379746	0.0992907	0.0389610	0
City1	0	0	0.0095238	0.0759493	0.0638297	0.4285714	0.32894
City2	0.0192307	0	0	0.1012658	0.0851063	0.4545455	0.48684
Planta1	0	0.0384615	0.1047619	0.1940928	0.2624114	0	0
Planta2	0.0192307	0.0384615	0.0952381	0.1729958	0.3333333	0.0259740	0
Wetland1	0.0192307	0.0576923	0.0095238	0.0970464	0	0	0

The 'Cluster Setup' dialog box is shown with the following settings:

- Distance Measure:** Sørensen (Bray-Curtis) (selected), Relative Sørensen, Jaccard, Euclidean (Pythagorean), Relative Euclidean, Correlation, Chi-squared.
- Group Linkage Method:** Nearest Neighbor, Farthest Neighbor, Median, Group Average (selected), Centroid, Ward's Method, Flexible Beta, McQuitty's Method.
- Dendrogram Width:** Narrow Width (selected), Wide Width.
- Dendrogram Spacing:** Single-spaced, Double-spaced (selected).
- Include cluster information with the dendrogram
- Log transform dendrogram scale
- Write distance matrix
- Add group membership variable to second matrix
- Write crude dendrogram to output file (high quality dendrogram always available)
- Main matrix is already a distance matrix
- Distances in main matrix are already squared

Buttons: OK, Cancel, Help

สำหรับโปรแกรม PC-ORD version 5.10 นั้นสามารถวิเคราะห์การจัดกลุ่มแบบสองทาง (Two-way Cluster Analysis) ได้ โดยจัดกลุ่ม rows และ columns ไว้ในเมทริกเดียว (single matrix) ดังนี้
 ไปที่คำสั่ง Groups > Cluster Analysis เมื่อขึ้นหน้าต่าง Two-way Cluster Setup ในส่วน Distance Measure ให้เลือกเป็น Sorensen (Bray-Curtis) สำหรับข้อมูล abundance ส่วนข้อมูล binary ให้เลือก Jaccard และ Group Linkage Method ให้เลือก Group Average ดังภาพ



3.2 การจัดอันดับ (ordination analysis)

การวิเคราะห์การจัดอันดับในโปรแกรม PC-ORD มีหลายวิธี เช่น Bray-Curtis, CCA, DCA, NMS, NMS Scores, PCA, RA และ Weighted Averaging (ดังภาพ) ในที่นี้จะขอกล่าวถึงวิธีที่นิยมใช้กันมากในทางนิเวศวิทยา คือ Principal Components Analysis (PCA), Non-metric Multidimensional Scaling (NMS) และ Canonical Correspondence Analysis (CCA)

Principal Components Analysis (PCA) เป็นวิธีการสกัดปัจจัยของการวิเคราะห์องค์ประกอบ (factor analysis) ที่ได้รับความนิยมมาก ซึ่งมีวัตถุประสงค์ที่จะนำรายละเอียดของตัวแปรที่มีจำนวนตัวแปรหลายๆ มาไว้ในปัจจัยที่มีเพียงไม่กี่ปัจจัย โดยพิจารณาจากรายละเอียดทั้งหมดจากแต่ละตัวแปร วิธีนี้อาศัยหลักความสัมพันธ์เชิงเส้นระหว่างตัวแปรที่ใช้เป็นข้อมูลองค์ประกอบหลักตัวแปรคือการผสมเชิงเส้นตรง (linear combination) การวิเคราะห์ PCA ทำได้โดย ไปที่คำสั่ง Ordination > PCA เมื่อขึ้นหน้าต่าง Principal Components Analysis เลือก Correlation สำหรับ Cross-product Matrix และเลือก distance-based biplot สำหรับ Calculate Scores For Species แล้วกด OK

PC-ORD 5.10

File Edit Advisor Modify Data Summary **Ordination** Graph Groups Tools Window Options Help

Main - data.wk1

	Q	Q		Q	Q	Q
	spA	spB		spE	spF	spG
Forest1	0.4807692	0.4230		0.421940	0.0567375	0.0259740
Forest2	0.3461539	0.3653846	0.3619048	0.0379746	0.0992907	0.0389610
City1	0	0	0.0095238	0.0759493	0.0638297	0.4285714
City2	0.0192307	0	0	0.1012658	0.0851063	0.4545455
Planta1	0	0.0384615	0.1047619	0.1940928	0.2624114	0
Planta2	0.0192307	0.0384615	0.0952381	0.1729958	0.3333333	0.0259740
Wetland1	0.0192307	0.0576923	0.0095238	0.0970464	0	0
Wetland2	0.0384615	0.0384615	0.0190476	0.1265823	0	0
Wetland3	0.0576923	0.0192307	0	0.1054852	0.0354609	0.0259740
Seashore	0.0192307	0.0192307	0	0.0464135	0.0638297	0

Bray-Curtis
CCA
DCA (DECORANA)
NMS
NMS Scores
PCA
RA
Weighted Averaging

PC-ORD 5.10

File Edit Advisor Modify Data Summary Ordination **Graph** Groups Tools Window Options Help

Main - data.wk1

	Q	Q		Q	Q	Q
	spA	spB		spE	spF	spG
Forest1	0.4807692	0.4230		0.421940	0.0567375	0.0259740
Forest2	0.3461539	0.3653846	0.3619048	0.0379746	0.0992907	0.0389610
City1	0	0	0.0095238	0.0759493	0.0638297	0.4285714
City2	0.0192307	0	0	0.1012658	0.0851063	0.4545455
Planta1	0	0.0384615	0.1047619	0.1940928	0.2624114	0
Planta2	0.0192307	0.0384615	0.0952381	0.1729958	0.3333333	0.0259740
Wetland1	0.0192307	0.0576923	0.0095238	0.0970464	0	0
Wetland2	0.0384615	0.0384615	0.0190476	0.1265823	0	0
Wetland3	0.0576923	0.0192307	0	0.1054852	0.0354609	0.0259740
Seashore	0.0192307	0.0192307	0	0.0464135	0.0638297	0

Principal Components Analysis

Cross-products Matrix

- Correlation
- Variance/Covariance (centered)
- Non-centered

Calculate Scores For species

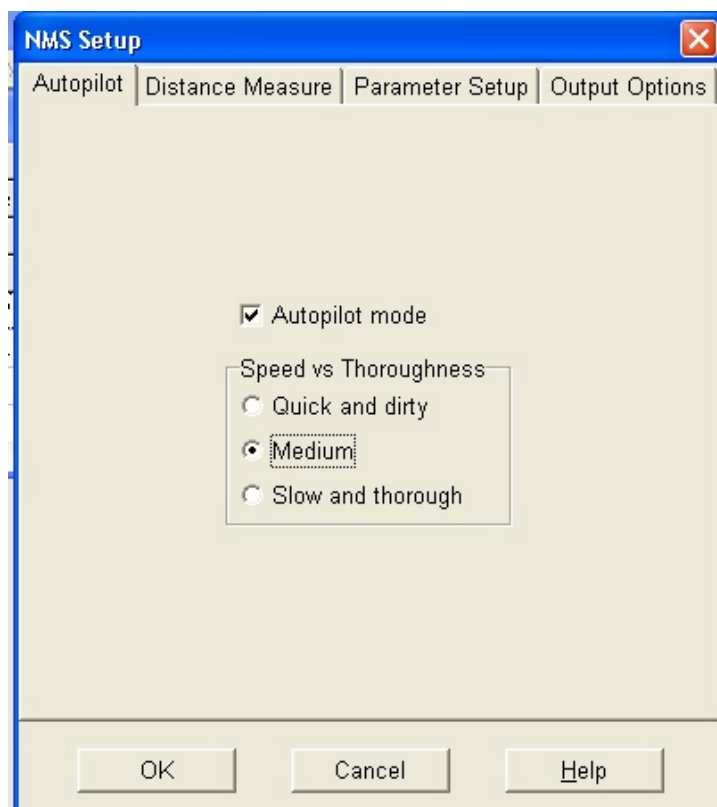
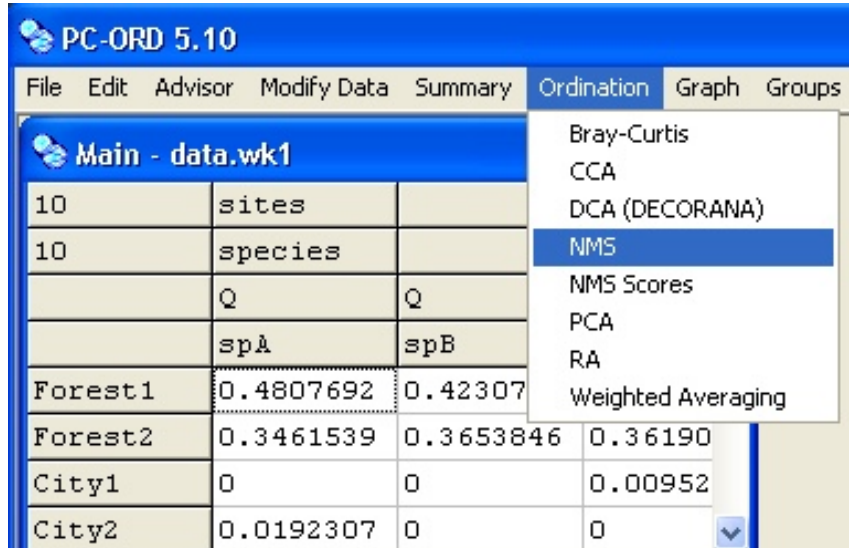
- none
- distance-based biplot
- by weighted averaging

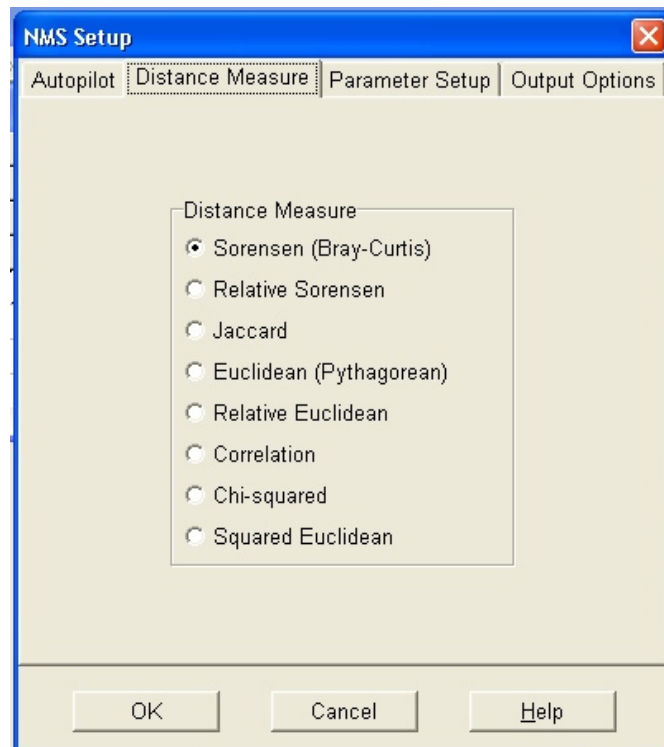
List the Cross-products Matrix

Randomization test

OK Cancel Help

Non-metric Multidimensional Scaling (NMS) หรือ MDS เป็นเทคนิคที่ใช้แสดงตำแหน่งข้อมูลที่ใช้เป็น dissimilarity data หรือข้อมูลที่แสดงความแตกต่าง/ความคล้าย การวิเคราะห์ PCA ทำได้โดย ไปที่คำสั่ง Ordination > NMS เมื่อขึ้นหน้าต่าง NMS Setup หน้าต่างย่อย Autopilot เลือก Autopilot mode และ Medium สำหรับ Speed vs Thoroughness หน้าต่างย่อย Distance Measure เลือก Sorensen (Bray-Curtis) แล้วกด OK





Canonical Correspondence Analysis (CCA) เป็นเทคนิคการหาความสัมพันธ์ระหว่างเมทริกของ species data (samples-by-species) และเมทริกของข้อมูลสิ่งแวดล้อม (environmental data) การวิเคราะห์ CCA ทำได้โดย การเตรียมข้อมูลสิ่งแวดล้อม ซึ่งเตรียมเหมือนกับ Main matrix เพื่อเป็น Second matrix (ตั้งภาพ) ทำการเปิดไฟล์ species data เป็น Main matrix และเปิด environmental data เป็น Second matrix จากนั้นไปที่คำสั่ง Ordination > CCA เมื่อขึ้นหน้าต่าง CCA Setup เลือก Centering and normalizing, Optimize rows: sites และ sites scores are derived from species แล้วกด OK

	A	B	C	D	E	F	G
1	10 sites						
2	6 variable						
3	Q	Q	Q	Q	Q	Q	Q
4	Canopy	Tdensity	GroundC	Rainfall	Arthrop	Mollusc	
5	Forest1	89	22	2	77	76	10
6	Forest2	93	19	3	79	58	5
7	City1	7	2	13	23	7	9
8	City2	5	3	17	45	8	11
9	Planta1	68	30	7	58	23	1
10	Planta2	76	30	8	74	34	2
11	Wetland1	1	1	34	48	11	22
12	Wetland2	0	2	43	50	17	31
13	Wetland3	1	3	50	39	18	28
14	Seashore	0	0	0	37	22	56
15							

PC-ORD 5.10

File Edit Advisor Modify Data Summary Ordination Graph Groups Tools Window Options Help

New... Open... Reopen Save Save As... Close... Append Results F4 Import Matrix... Export Matrix... Switch Matrix... Delete File Memory Requirements Print... Print Setup... Font... Exit

Project
Main Matrix F7
Second Matrix F8
Graph Row File F9
Graph Col File
Result File F10
Dendrogram
Two Way Dendrogram
Species-area Curves
NMS Scree Plot
Distributions Plot

	Q	Q	Q	Q
	spD	spE	spF	spG
	0.0421940	0.0567375	0.0259740	0
	0.0379746	0.0992907	0.0389610	0
	0.0759493	0.0638297	0.4285714	0.328947
	0.1012658	0.0851063	0.4545455	0.486842
	0.0384615	0.1047619	0.1940928	0.2624114
	0.0384615	0.0952381	0.1729958	0.3333333
	0.0576923	0.0095238	0.0970464	0
	0.0384615	0.0190476	0.1265823	0
	0.0192307	0	0.1054852	0.0354609
Seashore	0.0192307	0.0192307	0	0.0464135

PC-ORD 5.10

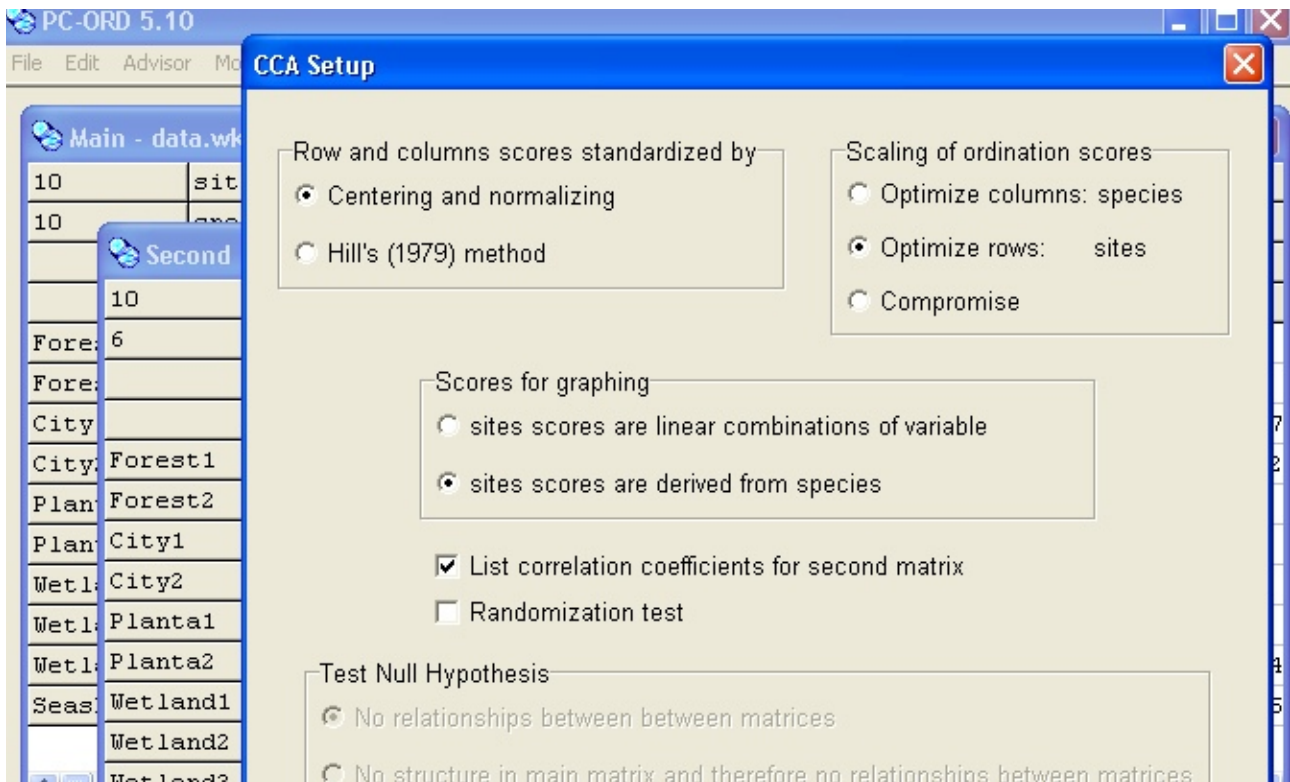
File Edit Advisor Modify Data Summary Ordination Graph Groups Tools Window Options Help

Main - data.wk1
10 sites
10 species

Second - variable.wk1
10 sites
Fore: 6 variable
Fore: Q Q Q Q
City Canopy Tdensity GroundC Rainfall
City: Forest1 0.2617647 0.1964286 0.0112994 0.145283
Plan: Forest2 0.2735294 0.1696429 0.0169491 0.1490566
Plan: City1 0.0205882 0.0178571 0.0734463 0.0433962
Wetl: City2 0.0147058 0.0267857 0.0960452 0.0849056
Wetl: Planta1 0.2 0.2678571 0.0395480 0.109434
Wetl: Planta2 0.2235294 0.2678571 0.0451977 0.1396226
Seas: Wetland1 0.0029411 0.0089285 0.1920904 0.0905660
Wetland2 0 0.0178571 0.2429378 0.0943396

Bray-Curtis
CCA
DCA (DECORANA)
NMS
NMS Scores
PCA
RA
Weighted Averaging

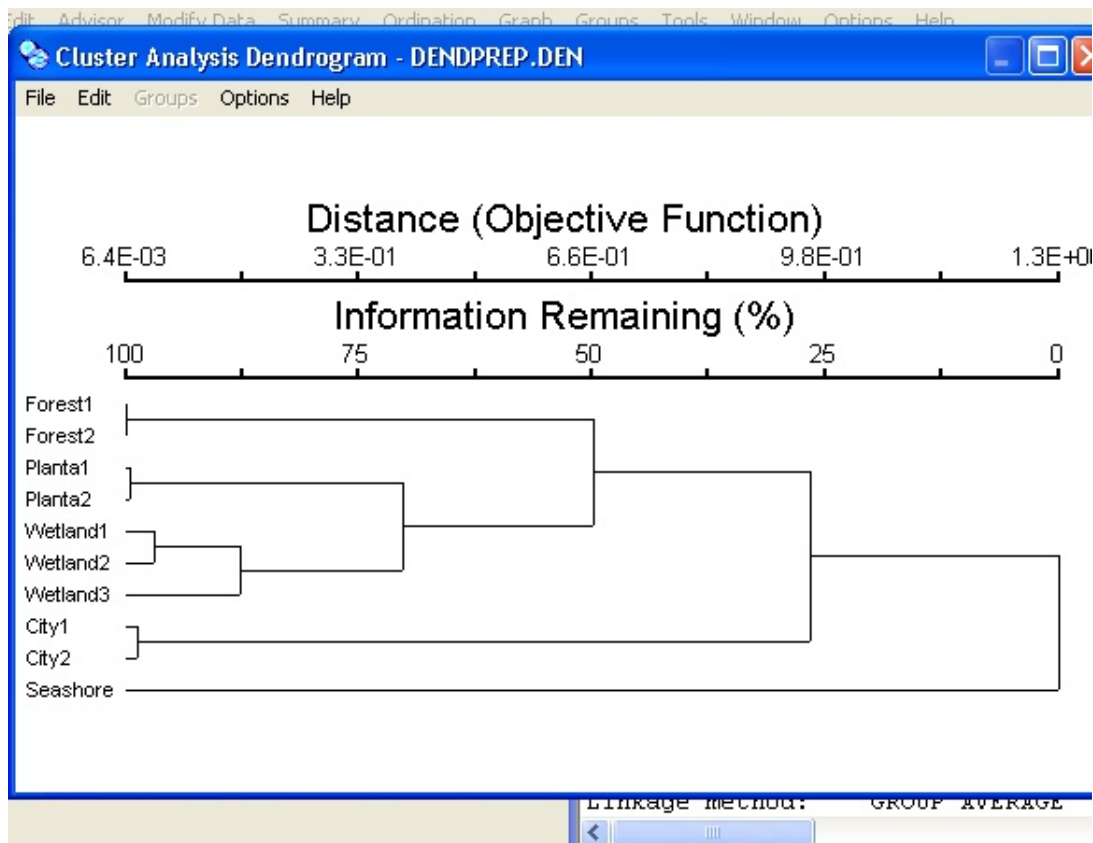
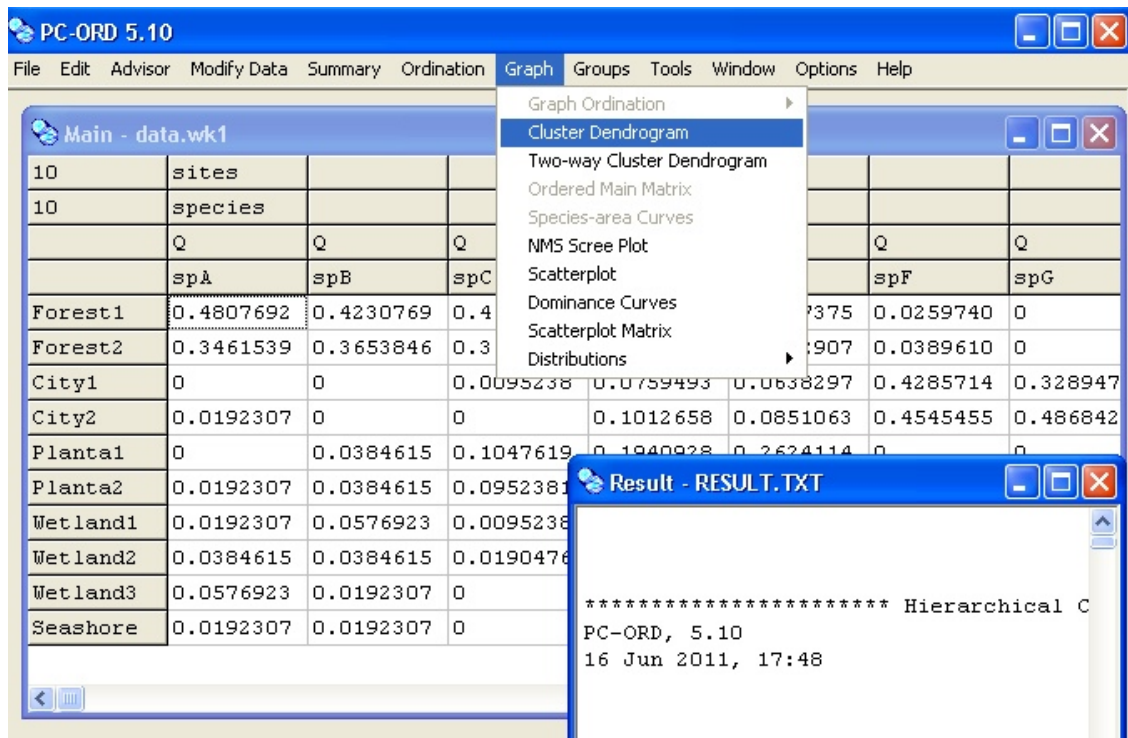
	Q	Q	Q	Q
	spF	spG		
	0.0259740	0		
	0.0389610	0		
	0.4285714	0.328947		
	0.4545455	0.486842		
	0	0		
	0.0259740	0		
	0	0		
	0.0259740	0.157894		
	0	0.026315		

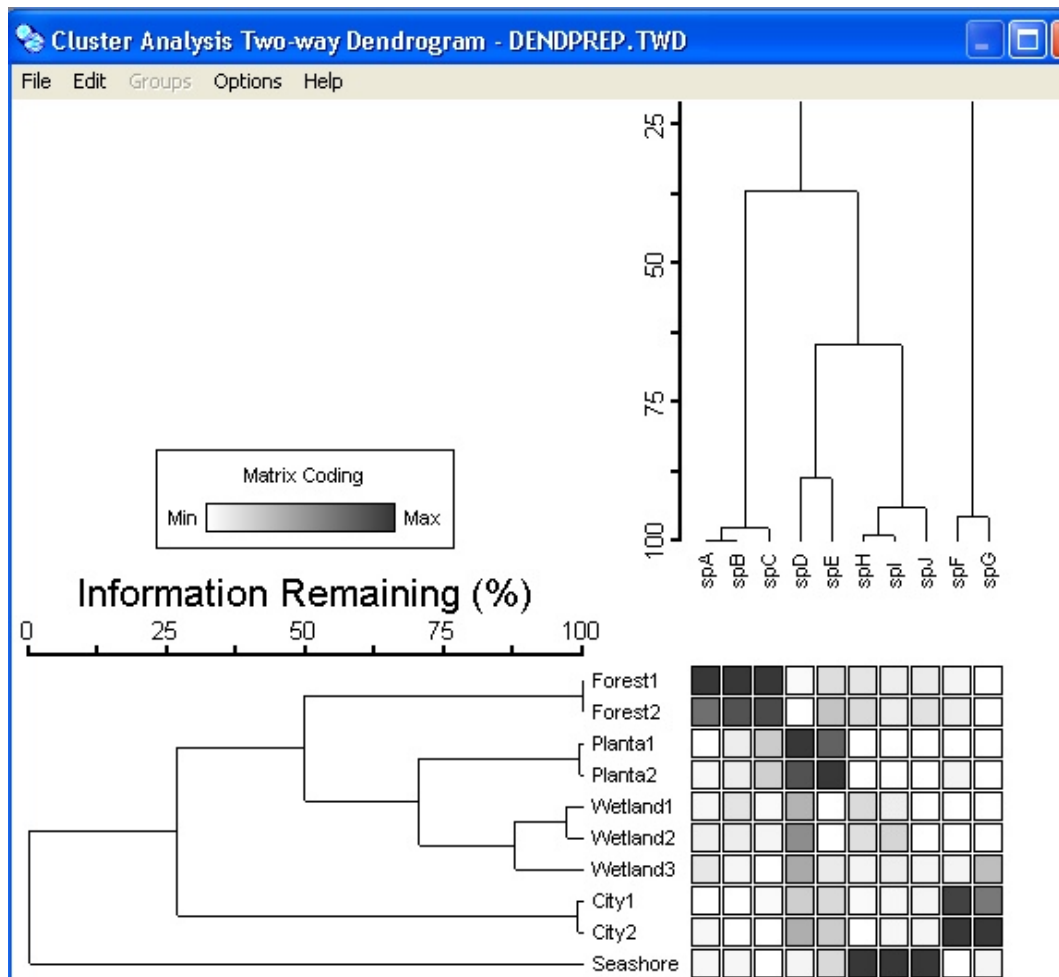
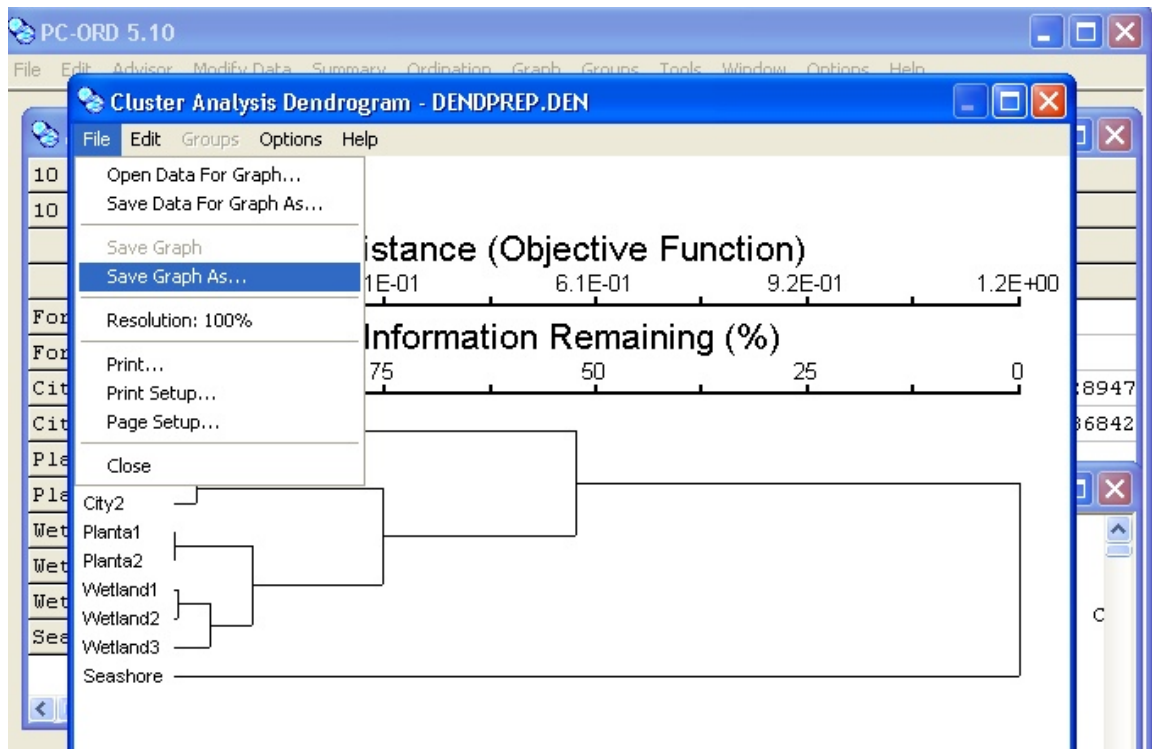


4. การแสดงผลกราฟ (Explore your results graphically)

4.1 การจัดกลุ่ม (cluster analysis)

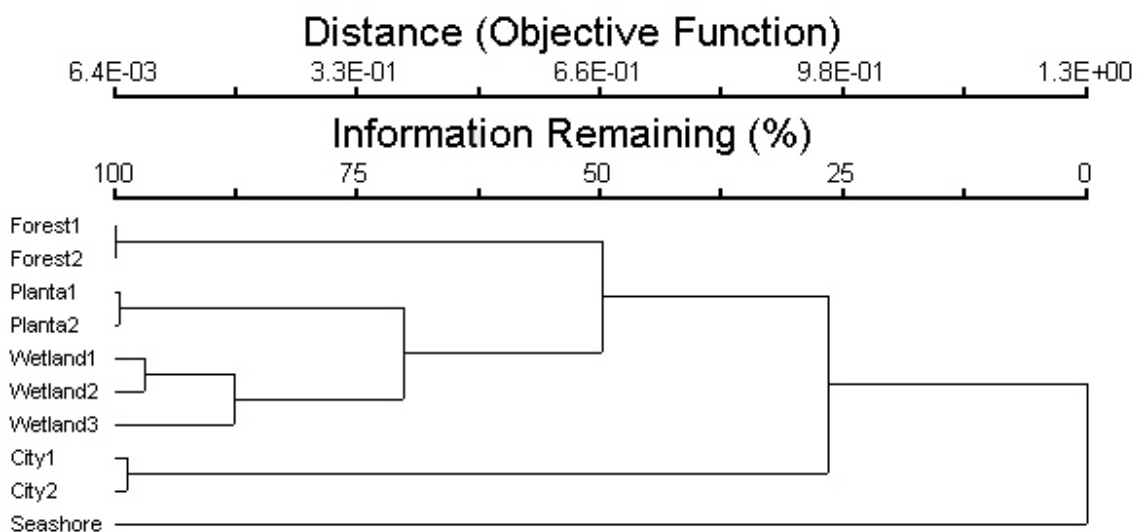
การแสดงผลกราฟหรือเดนโดรแกรม (dendrogram) ของการจัดกลุ่มนั้นทำได้โดย ไปที่คำสั่ง Graph > Cluster Dendrogram หรือ Two-way Cluster Dendrogram สำหรับการวิเคราะห์การจัดกลุ่มแบบ Two-way Cluster Analysis ซึ่งจะขึ้นหน้าต่าง Cluster Analysis Dendrogram หากต้องการบันทึก dendrogram ไปที่ File > Save Graph As ดังภาพ





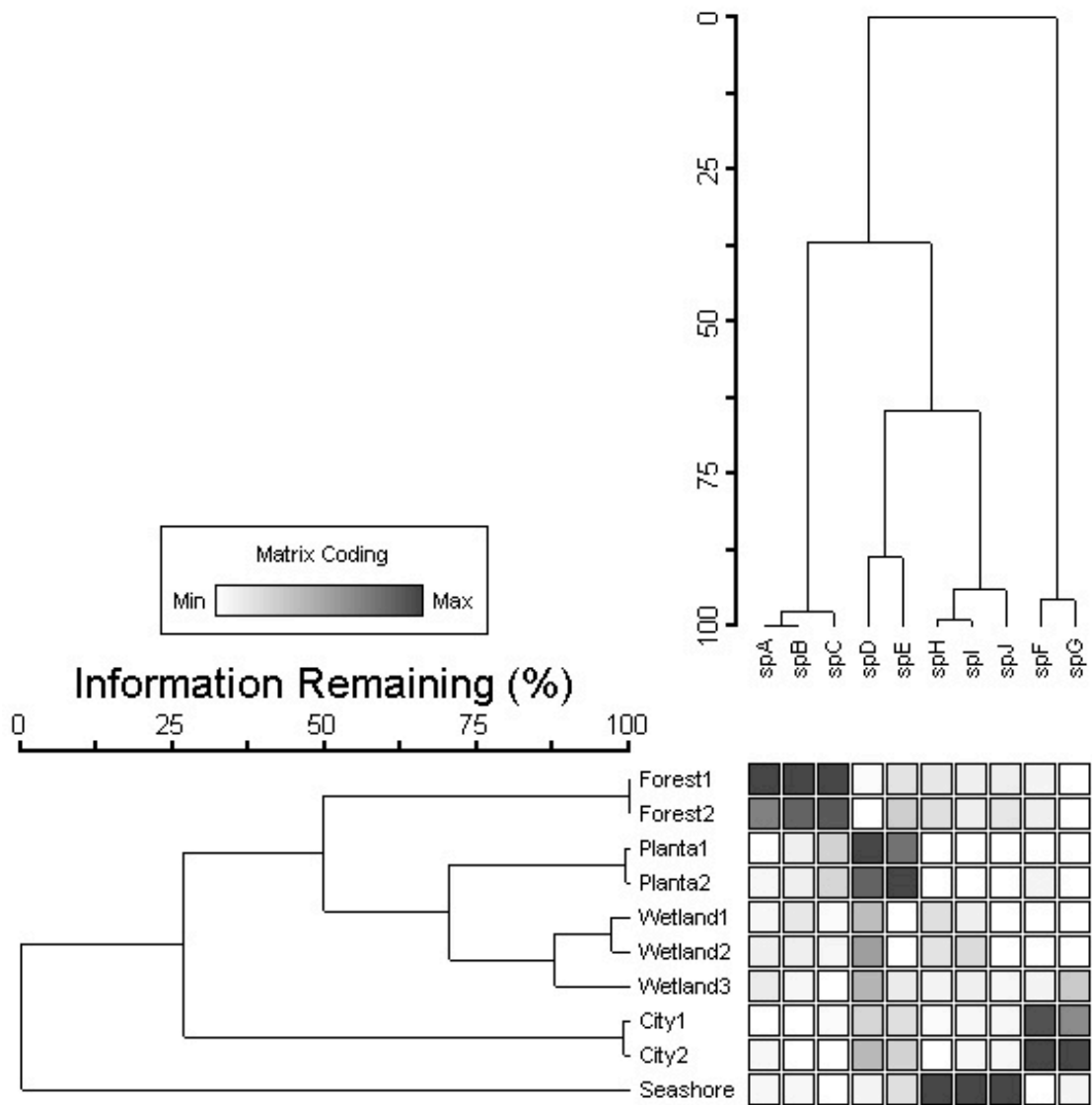
จาก Cluster Analysis Dendrogram เมื่อพิจารณาความห่างหรือความคล้ายกันของ 10 samples พบว่าแบ่งออกเป็น 4 กลุ่มที่ความห่างหรือความคล้าย 50% ดังนี้

- ☑ กลุ่มที่ 1 ได้แก่ Forest1 และ Forest2 มีความห่าง 0% (หรือคล้ายกัน 90%)
- ☑ กลุ่มที่ 2 แบ่งเป็น 2 กลุ่มย่อย ได้แก่ Planta1, Planta2 และ Wetland1, Wetland2, Wetland3 โดยกลุ่มที่ 2 มีความห่างจากกลุ่มที่ 1 ที่ 50%
- ☑ กลุ่มที่ 3 ได้แก่ City1 และ City2 โดยกลุ่มที่ 3 มีความห่างจากกลุ่มที่ 1 และ 2 ที่ 75% (หรือคล้ายกัน 25%)
- ☑ กลุ่มที่ 4 ได้แก่ Seashore โดยกลุ่มที่ 4 มีความห่างจากกลุ่มอื่นที่ 100% (หรือคล้ายกัน 0%)



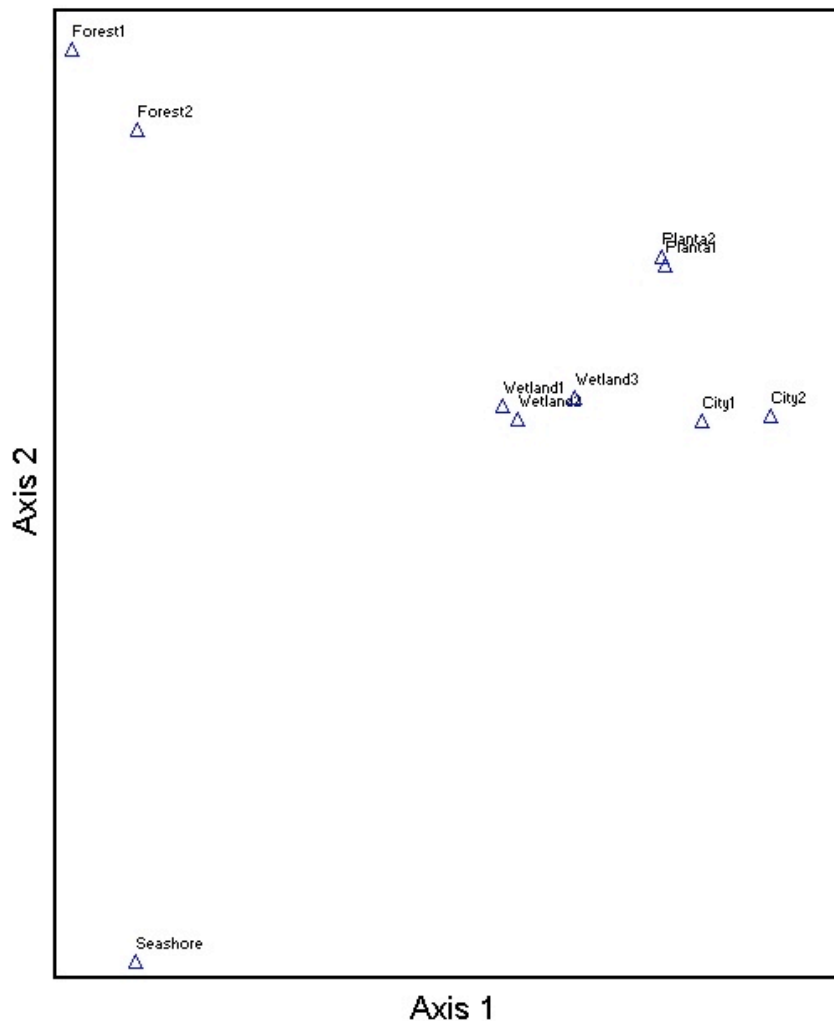
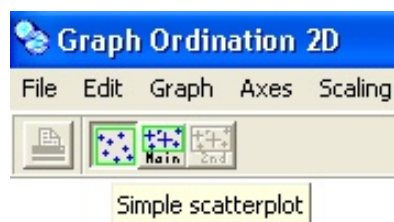
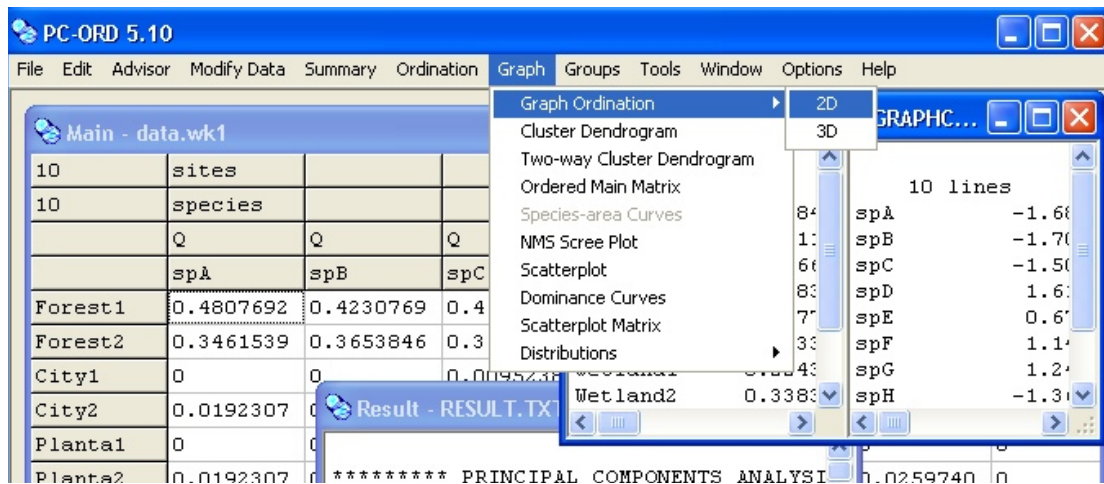
จาก Cluster Analysis Two-way Dendrogram ซึ่งแสดงเดนโดแกรมทั้ง rows (sites) และ columns (species) เมื่อพิจารณาความห่างหรือความคล้ายกันของ 10 sites พบว่ามีการจัดกลุ่มเหมือนกับ Cluster Analysis Dendrogram ทุกประการ ส่วนความห่างหรือความคล้ายกันของ 10 species พบว่าแบ่งออกเป็น 4 กลุ่มโดยที่ กลุ่มที่ 1 (Forest1 และ Forest2) มีความสัมพันธ์กับนกชนิด spA, spB, และ spC ซึ่งความเข้มของสีแสดงให้เห็นถึงความชุกชุมของนกชนิดนั้นในพื้นที่นั้น

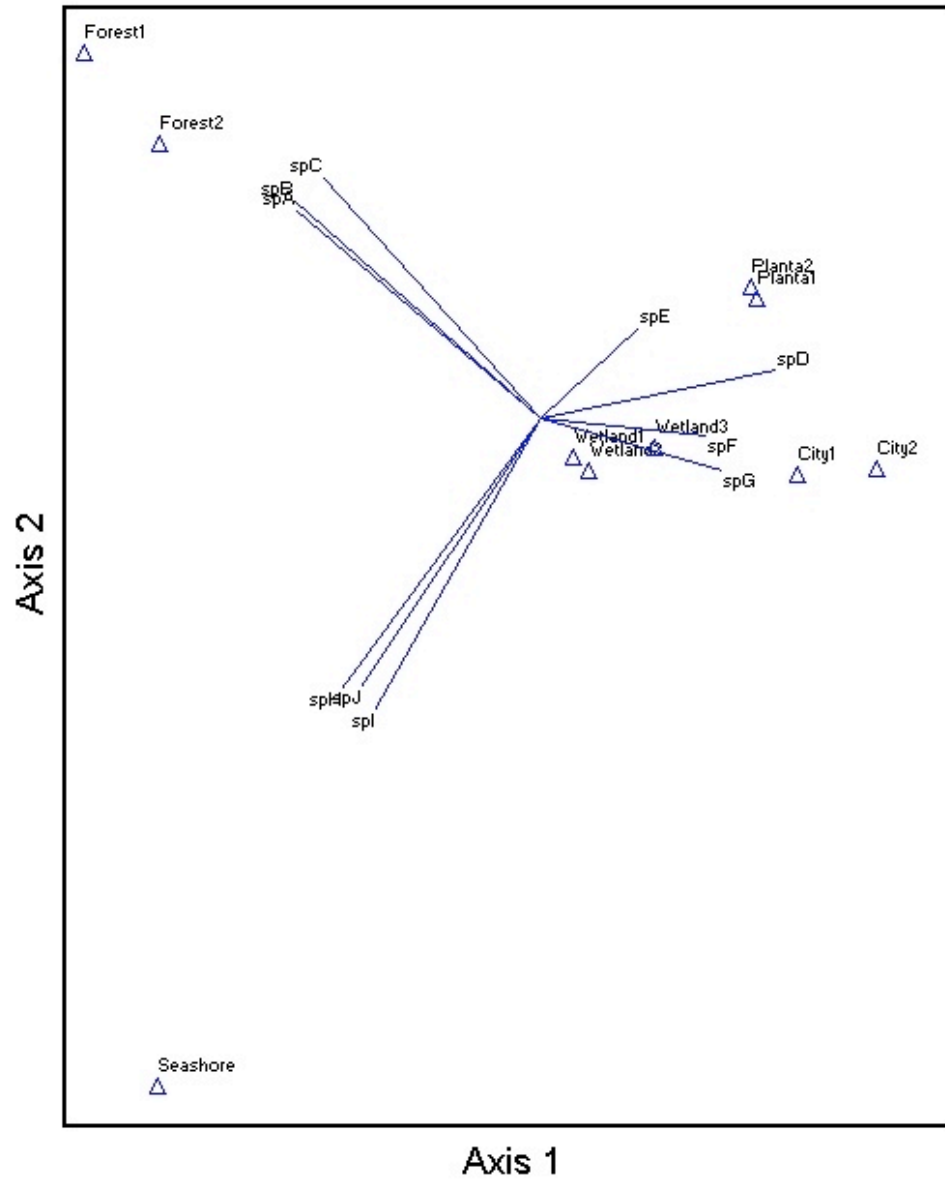
กลุ่มที่ 2 (Planta1, Planta2 และ Wetland1, Wetland2, Wetland3) มีความสัมพันธ์กับนกชนิด spD และ spE กลุ่มที่ 3 (City1 และ City2) มีความสัมพันธ์กับนกชนิด spF และ spG และกลุ่มที่ 4 (Seashore) มีความสัมพันธ์กับนกชนิด spH, spI และ spJ



4.2 การจัดอันดับ (ordination analysis)

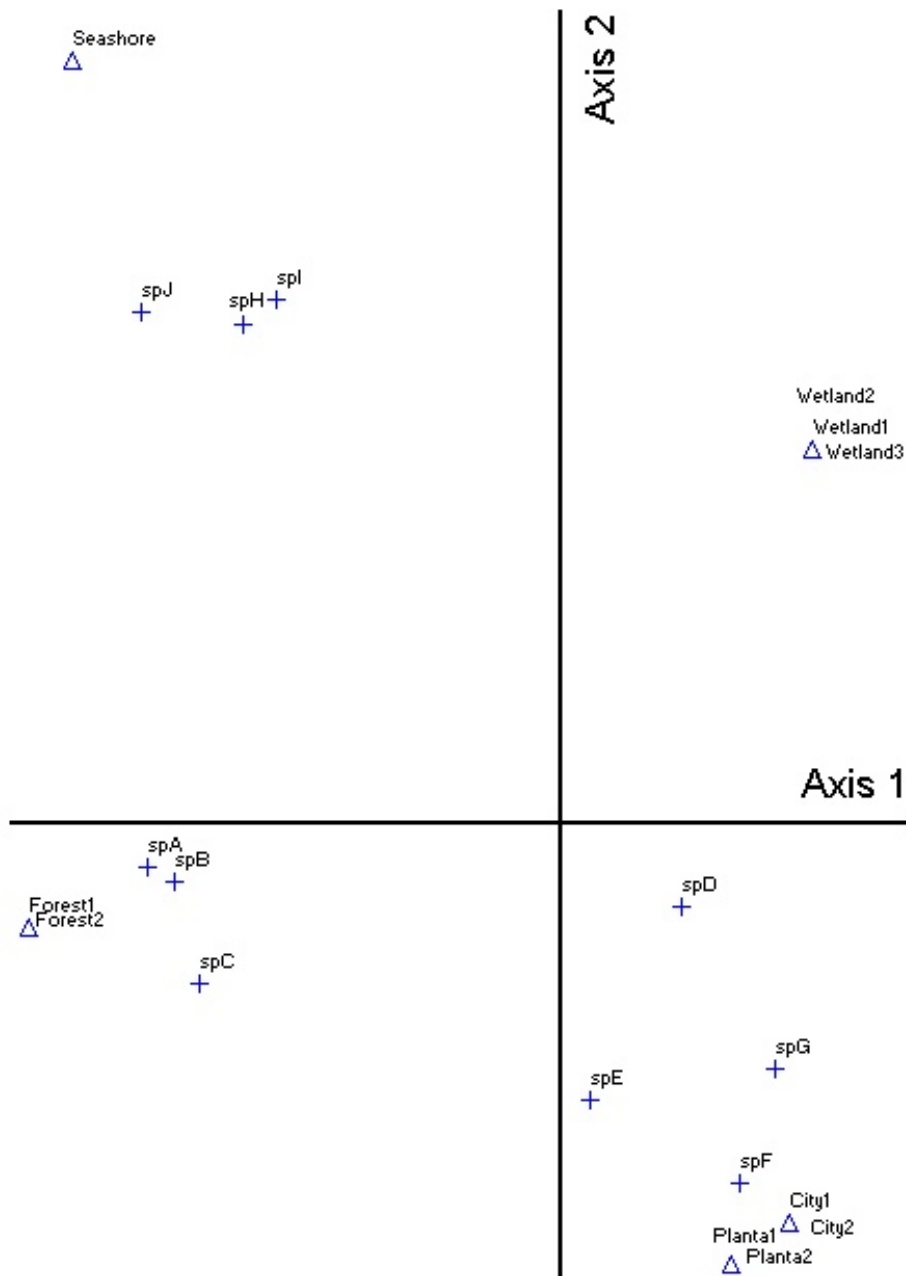
การแสดงผลกราฟของการจัดอันดับนั้นทำได้โดย ไปที่คำสั่ง Graph > Graph Ordination > 2D ซึ่งจะขึ้นหน้าต่าง Graph Ordination 2D และกด Simple scatterplot เพื่อแสดงกราฟ หากต้องการบันทึกกราฟ ไปที่ File > Save Graph As ดังภาพ หากต้องการแก้ไขกราฟ





กราฟ PCA

สำหรับกราฟ NMS ไปที่คำสั่ง Graph > Graph Ordination > 2D ซึ่งจะขึ้นหน้าต่าง Graph Ordination 2D และกด Simple scatterplot เพื่อแสดงกราฟ หากต้องการบันทึกกราฟ ไปที่ File > Save Graph As ดังภาพ

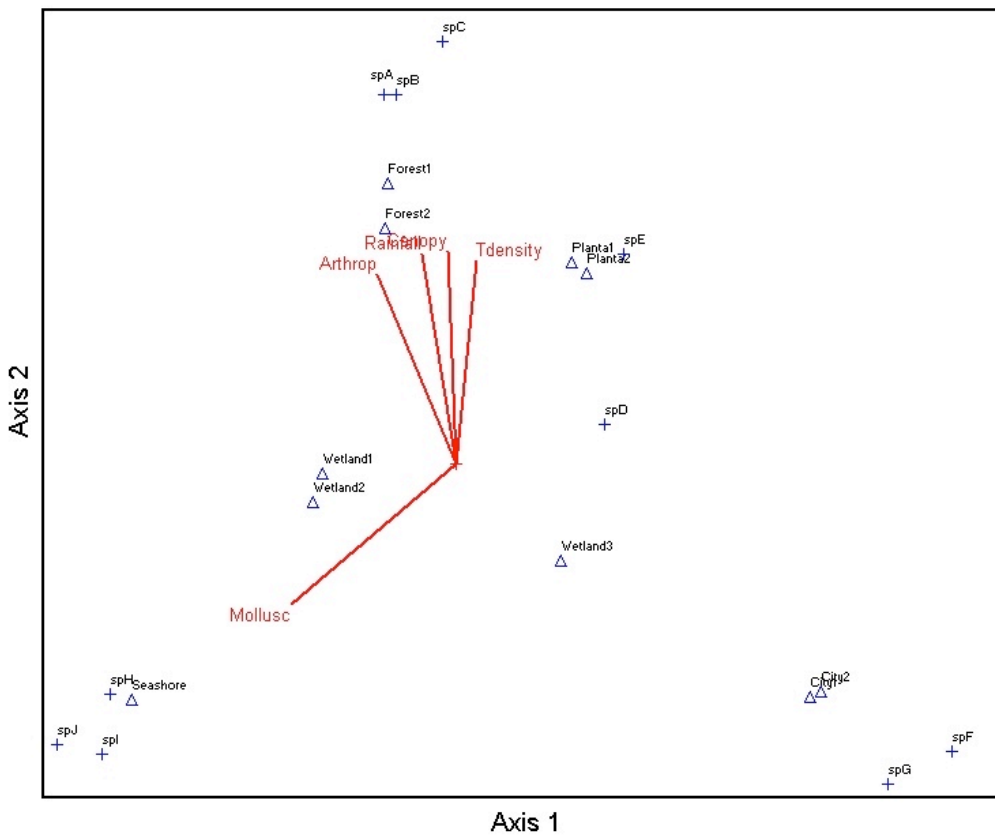
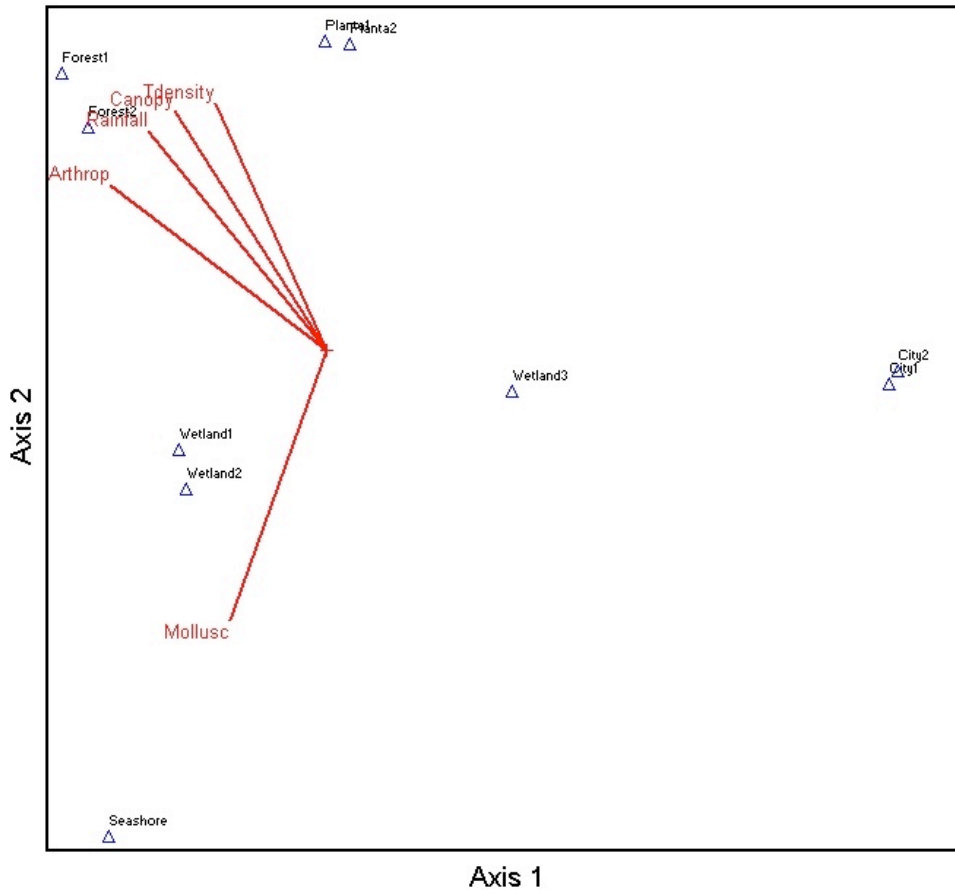


กราฟ NMS

สำหรับกราฟ CCA ไปที่คำสั่ง Graph > Graph Ordination > 2D ซึ่งจะขึ้นหน้าต่าง Graph Ordination 2D และกด Simple scatterplot และ Joint plot (biplot) เพื่อแสดงกราฟ หากต้องการบันทึกกราฟ ไปที่ File > Save Graph As ดังภาพ

หากต้องการหมุนแกนหรือหมุนกราฟ สามารถทำได้โดย ไปที่ Rotate > By Angle Continuous เลือก 5° for the increment และคลิก Next เพื่อหมุนแกนไปเรื่อยๆ ดังภาพ





หากต้องการแก้ไขกราฟ สามารถไปคลิก Preferences และหน้าต่างย่อย Format > Styles > Fonts ได้ดังภาพ



Preferences

Format | **Styles** | Fonts

Graph title
 Axis titles
 Tick marks
 Tick labels
 Label every other tick
 Overlay tick marks
 Hide toolbar
 Transparent legends
 Show graph hints
 Quadrant labeling

Background Color
 White
 Gray
 Black
 Navy

Print in black and white
 Save in black and white
 Copy in black and white

Plot Points
 sites
 species

Label Points
 sites
 species

sites Labels From
 Graph Row File
 Second Matrix
 Group Number

Joint plot labels

Groups
 Symbol code categories
 Color code categories
 Symbols and Colors
 Show legend
 Show legend frame
 Color code vectors

Quantitative Overlays
 Show side scatterplots
 Show regression line
 Show envelope line
 Flexibility = 3
 Standard deviation = 2.0

Plexus For species Scores
 Weak association lines
 Strong association lines
 Weak association = 0.3
 Strong association = 0.5

Small Expected Values
 Exclude associations with an expected value less than 5
 Apply Yates' correction to expected values less than 5
 Always apply Yates' correction

OK Cancel Page Defaults All Defaults Help

