

Πανεπιστήμιο Αιγαίου – Τμήμα Επιστημών της Θάλασσας

Παράκτια και μεταβατικά οικοσυστήματα

Διάλεξη 8. Ανάλυση βιοκοινότητας I

Διδάσκων: Αθανάσιος Ευαγγελόπουλος

Γραφείο Α3

(Εργαστήριο Θαλάσσιας Βιοποικιλότητας)

tevagelo@marine.aegean.gr

2017 - 2018

Ανάλυση βιοκοινότητας I

- Μονομεταβλητή vs πολυμεταβλητή ανάλυση οικολογικών δεδομένων
- Συντελεστές ομοιότητας – ανομοιότητας & απόστασης
- Ανάλυση ομαδοποίησης CLUSTER
- Ανάλυση ταξιθέτησης NMDS

Μονομεταβλητή vs πολυμεταβλητή ανάλυση
οικολογικών δεδομένων

Ανάλυση βιοκοινότητας I

Μονομεταβλητή και πολυμεταβλητή ανάλυση οικολογικών δεδομένων

- Τα οικολογικά φαινόμενα παρουσιάζουν πολυπλοκότητα
- Είναι σπάνιο μια μόνο μεταβλητή απόκρισης (π.χ. ένας δείκτης ποικιλότητας) να είναι επαρκής για να περιγράψει ένα οικολογικό σύστημα (= μονομεταβλητή ανάλυση)
- Πολλαπλές μεταβλητές απόκρισης (π.χ. οι αφθονίες πολλών ειδών) συχνά αναλύονται ταυτόχρονα (= πολυμεταβλητή ανάλυση) για να κατανοήσουμε ένα οικολογικό σύστημα
- Είναι συνηθισμένο να εξετάζουμε ταυτόχρονα πολλαπλές ερμηνευτικές μεταβλητές σε μια ανάλυση στην προσπάθεια μας να εξηγήσουμε τη μεταβλητότητα των μεταβλητών απόκρισης

Ανάλυση βιοκοινότητας I

Μονομεταβλητή και πολυμεταβλητή ανάλυση οικολογικών δεδομένων

		Variables			
		R1	E1	E2	E3
Samples	S1	14	A	3.2	12
	S2	10	A	2.1	6
	S3	0	B	0.5	11
	S4	0	B	1.6	22

Τα δεδομένα σε μια μονομεταβλητή ανάλυση περιλαμβάνουν μια μόνο μεταβλητή απόκρισης (R1). Αυτή μπορεί να αναλυθεί σε σχέση με μια ή περισσότερες ερμηνευτικές μεταβλητές (E1-E3).

Ανάλυση βιοκοινότητας I

Μονομεταβλητή και πολυμεταβλητή ανάλυση οικολογικών δεδομένων

		Variables					
		R1... Rn	E1	E2	E3		
Samples	S1	14	18	S1	A	3.2	12
	S2	10	7	S2	A	2.1	6
	S3	0	0	S3	B	0.5	11
	S4	0	1	S4	B	1.6	22

Τα δεδομένα σε μια πολυμεταβλητή ανάλυση περιλαμβάνουν πολλαπλές μεταβλητές απόκρισης (R1...Rn). Οι μέθοδοι πολυμεταβλητής ανάλυσης επιτρέπουν την ταυτόχρονη ανάλυση όλων των μεταβλητών απόκρισης σε σχέση με μια ή περισσότερες ερμηνευτικές μεταβλητές (E1-E3).

Συντελεστές ομοιότητας - ανομοιότητας

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

Εκτίμηση της σχέσης δύο βιοκοινοτήτων με:

Συντελεστές απόστασης
(distance coefficients)

Συντελεστές ανομοιότητας
(dissimilarity coefficients)

Συντελεστές ομοιότητας
(similarity coefficients)

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

Συντελεστές απόστασης:

- Παίρνουν τιμές στο κλειστό-ανοικτό διάστημα $[0, \infty)$
- Η τιμή του συντελεστή μεγαλώνει ανάλογα με την ανομοιότητα

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

Συντελεστές ανομοιότητας:

- Παίρνουν τιμές στα κλειστά διαστήματα $[0, 1]$ ή $[0, 100]$
- Μέγιστη τιμή συντελεστή = 1 ή 100: Τέλεια ανομοιότητα στη δομή των βιοκοινοτήτων (δείγματα χωρίς κοινά είδη)

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

Συντελεστές ομοιότητας:

- Παίρνουν τιμές στα κλειστά διαστήματα $[0, 1]$ ή $[0, 100]$
- Μέγιστη τιμή συντελεστή = 1 ή 100: Τέλεια ομοιότητα στη δομή των βιοκοινοτήτων (δείγματα ταυτόσημα)

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

Ο συντελεστής Bray–Curtis (Bray & Curtis 1957) συνίσταται για δεδομένα αφθονίας ειδών:

$$B_{jk} = \frac{\sum_{i=1}^p |y_{ij} - y_{ik}|}{\sum_{i=1}^p (y_{ij} + y_{ik})}$$

όπου: y_{ij} και y_{ik} = η αφθονία στα δείγματα j και k αντίστοιχα του είδους i , από ένα σύνολο p ειδών (κοινών και μη κοινών μεταξύ των δειγμάτων)

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

Συντελεστής των Bray-Curtis:

Στο PRIMER, ο συντελεστής των Bray-Curtis χρησιμοποιείται ως ο προεπιλεγμένος συντελεστής ομοιότητας και παίρνει τιμές στο κλειστό διάστημα $[0, 100]$

$\min = 0$ -> απόλυτα ανόμοια δείγματα

$\max = 100$ -> απόλυτα όμοια δείγματα

Ανάλυση βιοκοινότητας I

Συντελεστές ομοιότητας - ανομοιότητας

- Τα άφθονα είδη έχουν μεγάλη βαρύτητα στη τιμή του συντελεστή των Bray-Curtis, ενώ τα σπάνια είδη πολύ μικρή -> είναι συχνά επιθυμητός ο συνολικός μετασχηματισμός των βιοτικών δεδομένων πριν από τον υπολογισμό του συντελεστή
- Επιλογές μετασχηματισμού (με σειρά φθίνουσας επίδρασης στη βαρύτητα των άφθονων ειδών στον υπολογισμό του συντελεστή):
 - Παρουσία/απουσία
 - Λογαριθμικός = τέταρτης ρίζας
 - Τετραγωνικής ρίζας

Ανάλυση ομαδοποίησης CLUSTER

- Η ανάλυση ομαδοποίησης μετατρέπει τα πρωτογενή πολυμεταβλητά δεδομένα σε μικρό αριθμό ομάδων οι οποίες και περιγράφονται αντί των αρχικών πρωτογενών δεδομένων
- Πρόβλημα: Η φύση συνήθως δεν συνίσταται από καλά διακριτές ομάδες οντοτήτων και ως αποτέλεσμα αυτές απαιτούν ιδιαίτερους τύπους ανάλυσης ομαδοποίησης

Ανάλυση βιοκοινότητας I

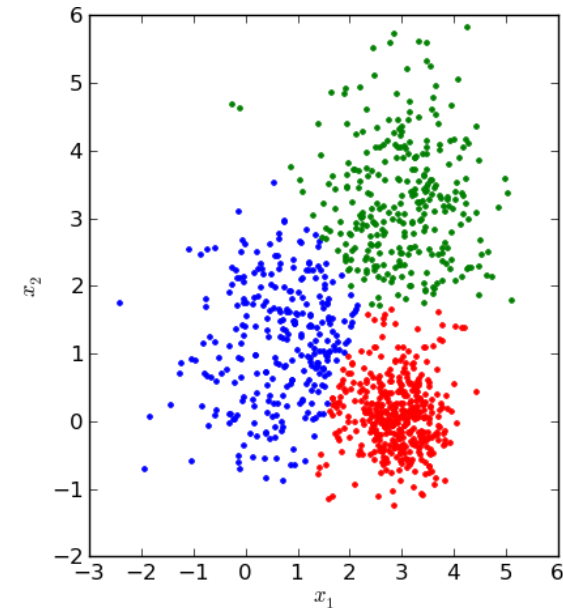
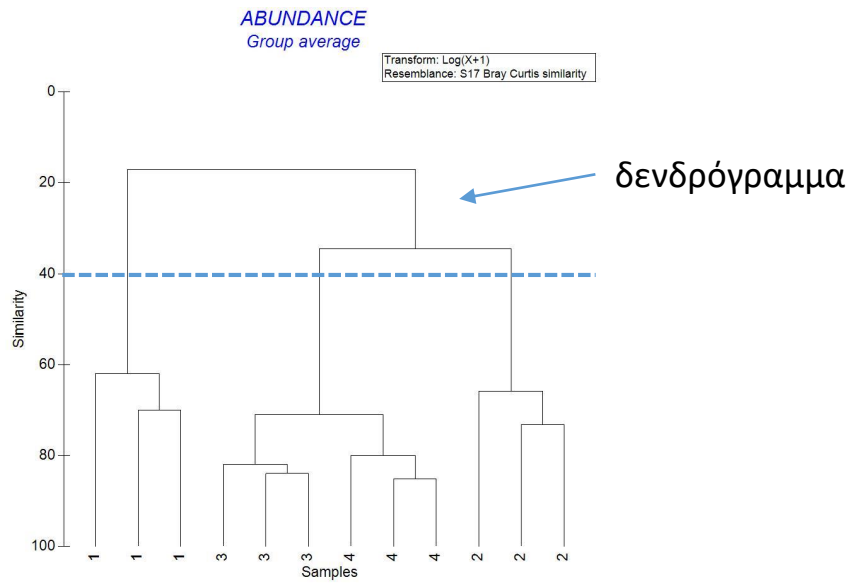
Ανάλυση ομαδοποίησης CLUSTER

- Η ανάλυση ομαδοποίησης μπορεί να είναι ιεραρχική ή μη-ιεραρχική
- Η ανάλυση CLUSTER είναι μια ιεραρχική μέθοδος ομαδοποίησης

Cluster analysis (ιεραρχική μέθοδος ομαδοποίησης)

K-means (μη ιεραρχική μέθοδος ομαδοποίησης)

Τιμές συντελεστή ομοιότητας Bray-Curtis



Σταδιακή ένωση των δειγμάτων σε ομάδες
Διακρίνονται 3 ομάδες δειγμάτων σε επίπεδο ομοιότητας 40%

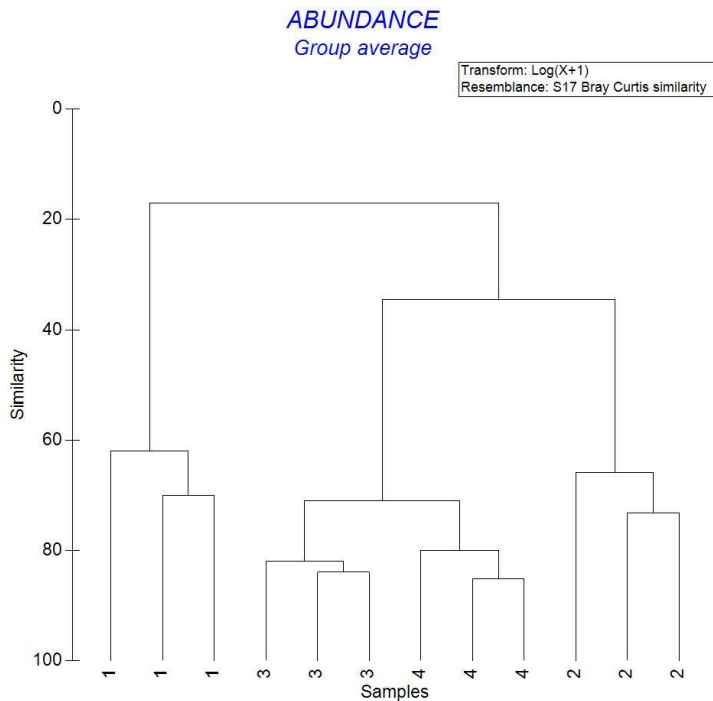
3 ομάδες παρατηρήσεων (χρωματισμένες με
διαφορετικό χρώμα)

Ανάλυση βιοκοινότητας I

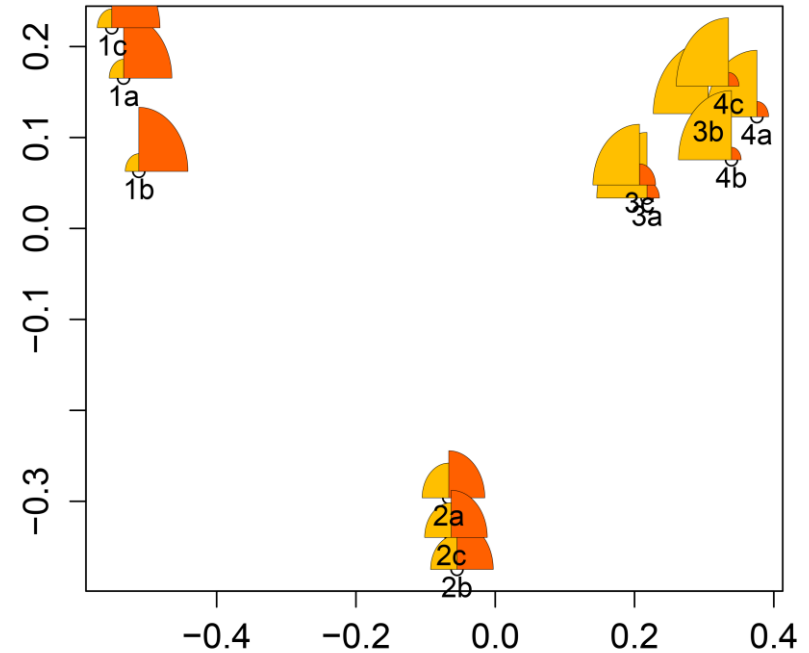
Ανάλυση ομαδοποίησης CLUSTER

- Η ανάλυση ομαδοποίησης μπορεί να είναι αυστηρή ή ασαφής
- Η ανάλυση CLUSTER είναι μια αυστηρή μέθοδος ομαδοποίησης

cluster analysis (αυστηρή μέθοδος ομαδοποίησης)



FCM classification (ασαφής μέθοδος ομαδοποίησης)



Ένα αντικείμενο μπορεί να ανήκει σε μια μόνο ομάδα

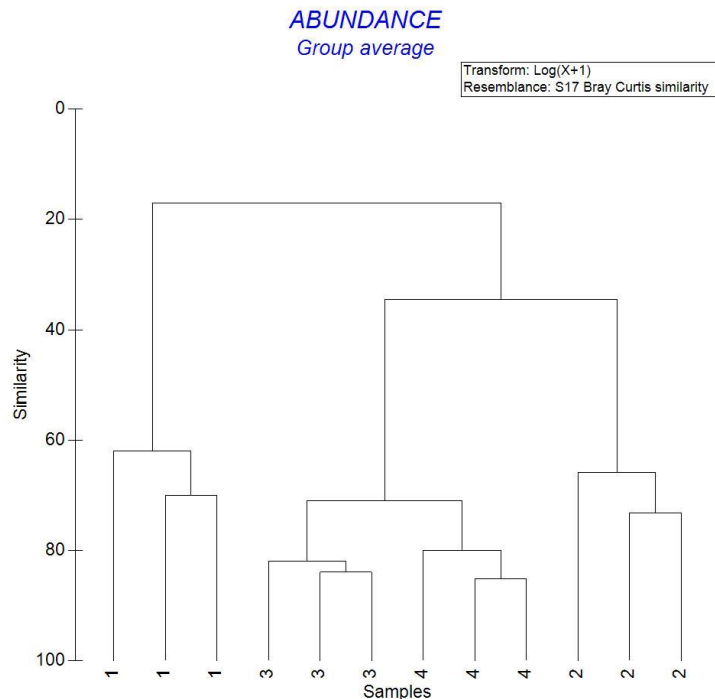
Ένα αντικείμενο μπορεί να ανήκει σε πάνω από μια ομάδες
Οι τιμές συμμετοχής κάθε αντικειμένου στις ομάδες υπολογίζονται βάσει μιας συνάρτησης συμμετοχής μ

Ανάλυση βιοκοινότητας I

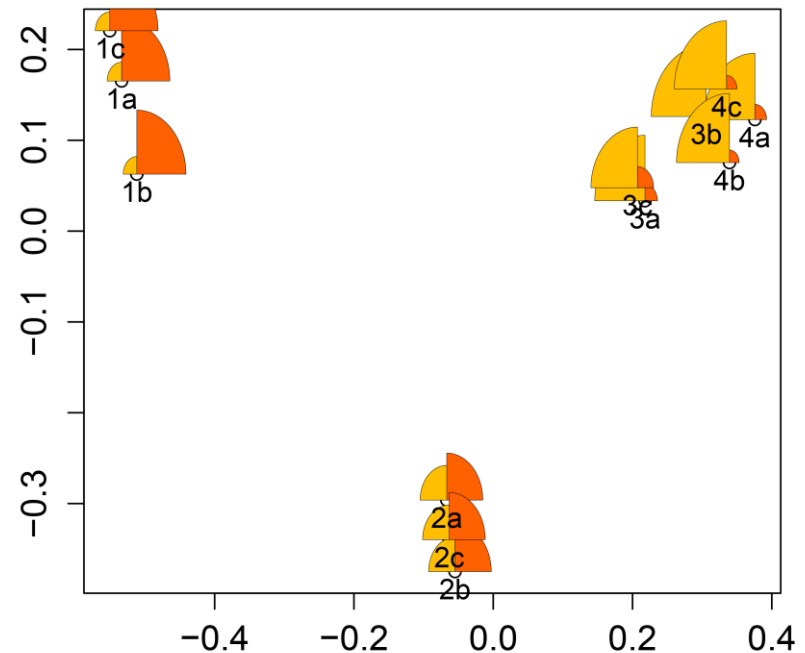
Ανάλυση ομαδοποίησης CLUSTER

- Στην ασαφή μέθοδο ομαδοποίησης (FCM) μπορούμε να διακρίνουμε πως η δομή της βιοποικιλότητας στο σταθμό 2 είναι ενδιάμεση εκείνης των σταθμών 1 και 3&4, ενώ στην αυστηρή μέθοδο ομαδοποίησης (CLUSTER) αυτό δεν είναι δυνατό

cluster analysis (αυστηρή μέθοδος ομαδοποίησης)



FCM classification (ασαφής μέθοδος ομαδοποίησης)



Ανάλυση βιοκοινότητας I

Ανάλυση ομαδοποίησης CLUSTER

- Η γειτνίαση δύο δειγμάτων στο δενδρόγραμμα της ανάλυσης ομαδοποίησης δεν είναι ένδειξη ομοιότητας αν οι κλάδοι ενώνονται σε χαμηλή τιμή ομοιότητας
- Το λογισμικό συνήθως διευθετεί τις θέσεις των κλάδων του δενδρογράμματος με τυχαίο τρόπο ή βάσει απλών κριτηρίων
- Η ανάλυση ομαδοποίησης CLUSTER δεν αποτελεί στατιστικό έλεγχο

Ανάλυση ταξιθέτησης NMDS

Ανάλυση βιοκοινότητας I

Ανάλυση ταξιθέτησης NMDS

- Η ανάλυση ταξιθέτησης αναπαριστά οντότητες (π.χ. δείγματα και/ή είδη) ως σημεία κατά μήκος ενός ή περισσότερων αξόνων αναφοράς
- Βασικός στόχος μιας ανάλυσης ταξιθέτησης είναι η απλοποίηση και οπτικοποίηση των πολυμεταβλητών δεδομένων, ώστε να μπορούμε να διακρίνουμε σημαντικά πρότυπα ή διαβαθμίσεις που υπάρχουν σε αυτά

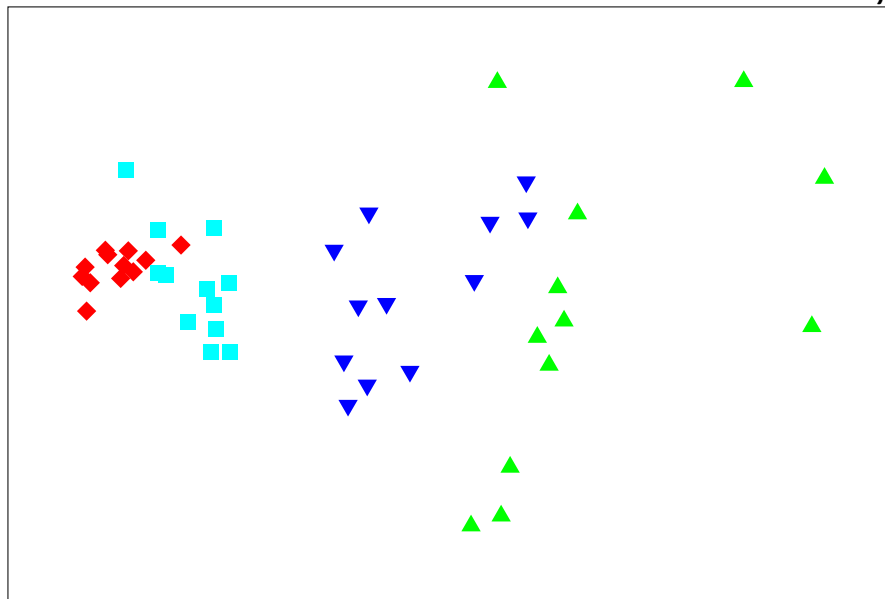
Ανάλυση βιοκοινότητας I

Ανάλυση ταξινόμησης NMDS

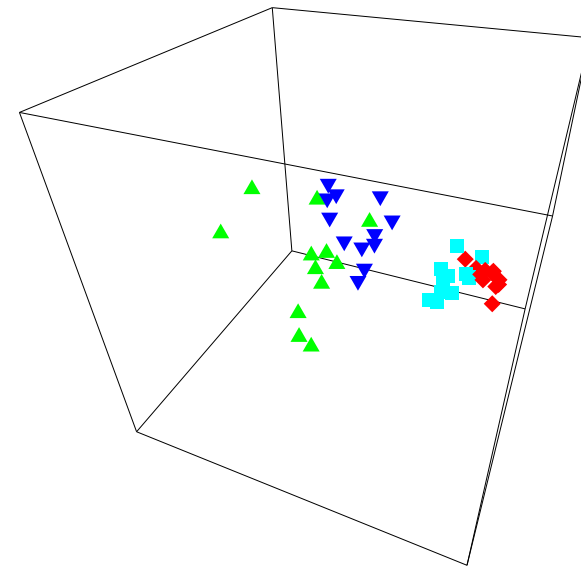
Μη-μετρική Πολυδιάστατη Διευθέτηση (NMDS):

Μικρότερη απόσταση
δειγμάτων στο γράφημα =
μεγαλύτερη ομοιότητα

Οι άξονες δεν αναπαριστούν τις πιο σημαντικές διαβαθμίσεις (όπως συμβαίνει σε άλλους τύπους ανάλυσης ταξινόμησης, π.χ. PCA) στα δεδομένα και η σειρά τους (1ος άξονας, 2ος άξονας κλπ.) είναι αυθαίρετη



2D

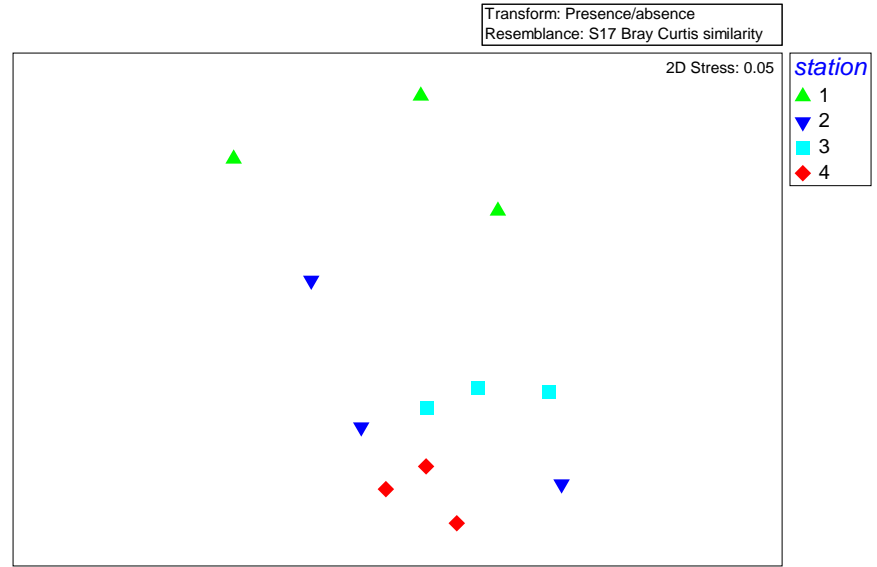
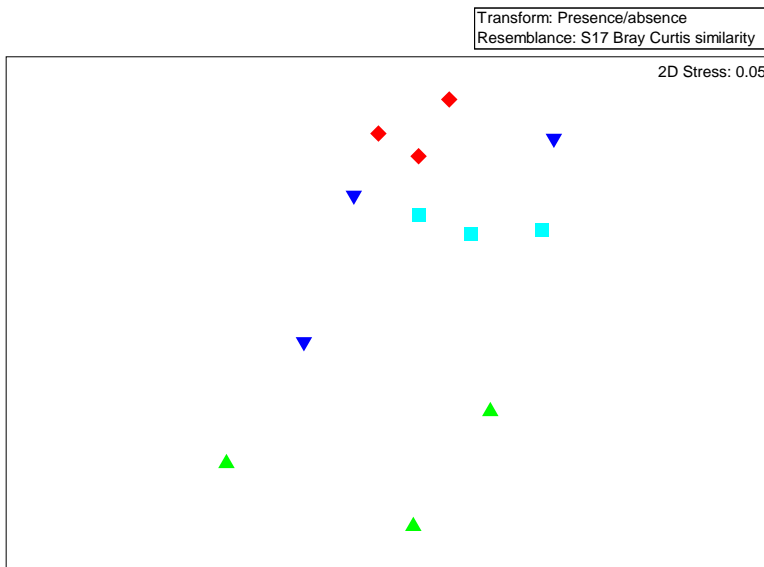


3D

Ανάλυση βιοκοινότητας I

Ανάλυση ταξίθετης NMDS

Επειδή οι άξονες της NMDS δεν αναπαριστούν διαβαθμίσεις, μπορούν να περιστραφούν ή να αντιστραφούν ή να αντιμετατεθούν όπως επιθυμούμε για την καλύτερη οπτικοποίηση των αποτελεσμάτων:



αντιστροφή του άξονα γ

Πλεονεκτήματα:

- Μπορεί να βασιστεί σε οποιονδήποτε συντελεστή ομοιότητας (αντίθετα, η PCA για παράδειγμα μπορεί να βασιστεί μόνο στην ευκλείδιο απόσταση)
- Αναπαριστά σε ικανοποιητικό βαθμό τις ομοιότητες των αντικειμένων που ταξιθετούμε

Μειονεκτήματα:

- Οι άξονες δεν αναπαριστούν διαβαθμίσεις
- Δεν ταξιθετούνται ταυτόχρονα δείγματα και είδη (όπως π.χ. στην CCA)
- Δεν υπάρχει μια και μοναδική λύση (τα αποτελέσματα είναι δυνατό να διαφέρουν λίγο κάθε φορά που τρέχουμε την ανάλυση)
- Ο αλγόριθμος δεν βρίσκει πάντοτε τη βέλτιστη λύση (για το λόγο αυτό το λογισμικό τρέχει τον αλγόριθμο πολλές φορές)

"Στρες" μιας λύσης NMDS (2D), δηλ. αξιοπιστία των αποτελεσμάτων:

<0.05 = άριστη ταξιθέτηση

<0.1 = καλή ταξιθέτηση

<0.2 = δυνητικά χρήσιμη

<0.3 = μη χρήσιμη

Ανάλυση βιοκοινότητας I

Βιβλιογραφία

McCune, B. & Grace, J.B., 2002. Analysis of ecological communities. MjM Software Design.

Gotelli, M.J. & Ellison, A.M., 2004. A primer of ecological statistics. Sinauer Associates

Krebs, C.J., 1999. Ecological Methodology. Addison-Wesley.

Καρανδεινός, Μ.Γ., 2007. Ποσοτικές οικολογικές μέθοδοι. Πανεπιστημιακές Εκδόσεις Κρήτης.