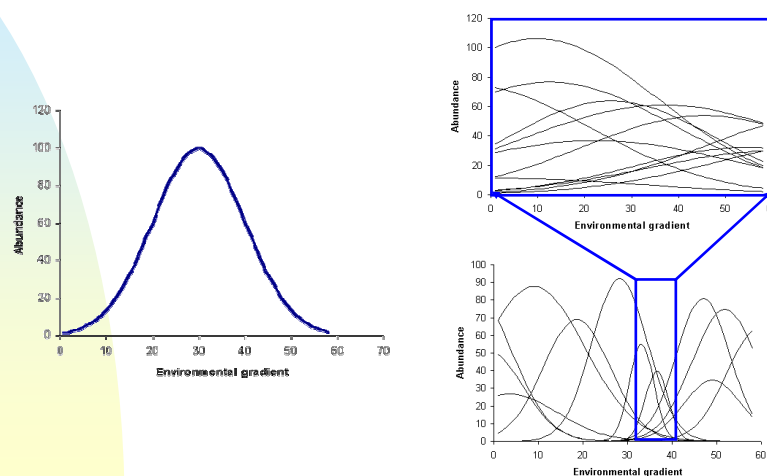


Caratteristiche dei dati ecologici

- I dati sono “sparsi”, cioè hanno molti valori nulli (a volte la maggioranza!)
- La gran parte delle specie presenti è rara.
- I fattori ambientali che influenzano la distribuzione delle specie sono molteplici e combinati fra loro,...
- ...ma quelli veramente importanti sono pochi (bassa dimensionalità intrinseca).
- I dati contengono molto “rumore” sia per eventi stocastici e contingenti, sia per l'errore di osservazione (anche in condizioni ideali le repliche sono diverse!)
- L'informazione è spesso ridondante (la specie A è associata alla specie B, ma questa può essere associata alla specie C, etc.): questo è un problema, ma è anche ciò che rende possibile interpretare i dati ecologici.

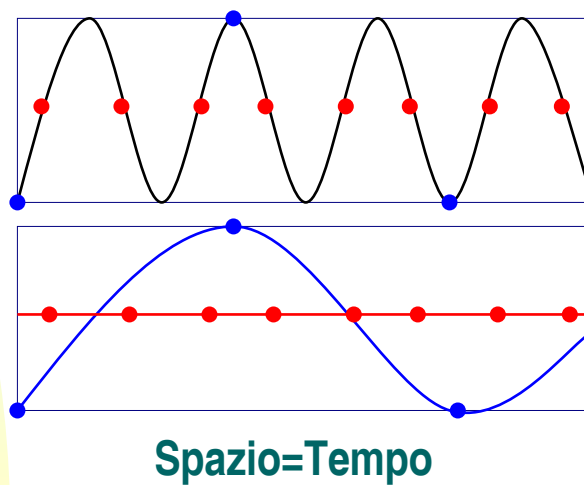
Gradienti ambientali e cenoclini



Piani di campionamento

- Non esistono informazioni preliminari, le scale spazio-temporali non sono note:
 - ◆ **piano randomizzato**
- Esistono informazioni preliminari, le scale spazio-temporali sono note:
 - ◆ **piano regolare**
- Esistono sufficienti dati di riferimento per descrittori accessori (covarianti):
 - ◆ **piano stratificato**

Scale e frequenze di osservazione



I dati

Demo

Sampling stations	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Date	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00
Time	16.05	15.40	15.20	14.50	14.30	14.00	13.35	12.20	11.55	11.25	11.00
Depth (m)	9.9	14.8	19.0	22.8	25.2	25.9	26.3	27.1	28.1	28.0	28.1
Coordinates North	4991467	4991855	4992243	4992628	4993022	4993396	4993792	4994168	4994566	4994951	4995336
Coordinates East	2317820	2318854	2319883	2320915	2321942	2322981	2324016	2325031	2326067	2327096	2328129

taxa	ID	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	
ANPHIPODA	<i>Ampelisca diadema</i>	0	2	0	13	16	3	6	14	15	2	6	
	<i>Ampelisca sarsi</i>	0	10	9	12	13	4	0	0	0	10	0	
	<i>Aora spinicornis</i>	0	0	0	0	0	0	0	0	0	0	0	
	<i>Leptocheirus pectinatus</i>	0	0	0	0	0	0	0	1	0	0	1	
	<i>Apherusa chierighini</i>	0	0	0	0	0	0	0	0	0	0	0	
	<i>Corophium acutum</i>	0	0	0	0	0	0	0	0	0	0	0	
MOLLUSCA	<i>Corophium spp.</i>	0	0	0	0	0	0	0	1	0	1	1	
	<i>Abra nitida</i>	0	3	1	1	2	1	4	4	8	6	1	
	<i>Chlamys varia</i>	0	0	0	0	0	0	0	0	0	0	0	
	<i>Corbula gibba</i>	0	2	88	6	1	24	22	21	11	1	0	
	<i>Phaxas adriaticus</i>	0	1	0	0	0	0	0	2	3	0	1	
	<i>Nassarius reticulatus</i>	0	0	2	0	0	1	0	0	0	0	0	
	<i>Nucula nitidosa</i>	0	0	2	1	1	4	11	0	10	12	1	
	<i>Nuculana pella</i>	0	0	0	0	0	0	0	0	2	0	0	
	<i>Pitar rudis</i>	0	0	3	0	0	2	0	2	2	1	0	
	<i>Psammobia fervensis</i>	0	0	0	0	0	0	0	0	0	1	0	
	<i>Saxicavella jeffreysi</i>	0	0	0	0	0	3	0	0	0	0	0	
	<i>Scrobicularia plana</i>	0	0	0	0	0	0	0	0	0	2	0	
	<i>Smithiella smithi</i>	0	0	0	0	0	0	0	0	0	1	1	
	POLYCHAETA	<i>Phylo cuvieri</i>	0	0	0	0	0	0	0	0	1	1	0
		<i>Aricidea assimilis</i>	0	4	0	0	1	2	1	0	1	1	0
<i>Polydora flava</i>		0	0	0	0	0	0	1	0	6	0	0	
<i>Pronosio caspersi</i>		41	0	1	0	0	0	0	0	0	0	0	
<i>Pronosio malmgreni</i>		3	31	0	0	3	0	9	3	65	2	2	
<i>Pronosio multibranchiata</i>		0	110	0	0	5	1	44	9	52	7	16	
<i>Pseudopolydora antennata</i>		0	0	0	0	0	0	0	0	0	0	2	
<i>Spiophanes kroyeri</i>		0	0	7	0	0	0	0	8	28	13	12	
<i>Magelona allenii</i>		0	0	0	0	0	0	0	0	3	0	0	
<i>Magelona minuta</i>		0	1	0	0	0	0	0	0	5	1	0	
<i>Poecilochaetus fauchaldi</i>		0	0	0	0	0	0	1	3	1	1	1	
...		

Molti dati sono nulli!

Sampling stations	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Date	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00
Time	16.05	15.40	15.20	14.50	14.30	14.00	13.35	12.20	11.55	11.25	11.00
Depth (m)	9.9	14.8	19.0	22.8	25.2	25.9	26.3	27.1	28.1	28.0	28.1
Coordinates North	4991467	4991855	4992243	4992628	4993022	4993396	4993792	4994168	4994566	4994951	4995336
Coordinates East	2317820	2318854	2319883	2320915	2321942	2322981	2324016	2325031	2326067	2327096	2328129

taxa	ID	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	
ANPHIPODA	<i>Ampelisca diadema</i>	0	2	0	13	16	3	6	14	15	2	6	
	<i>Ampelisca sarsi</i>	0	10	9	12	13	4	0	0	0	10	0	
	<i>Aora spinicornis</i>	0	0	0	0	0	0	0	0	0	0	0	
	<i>Leptocheirus pectinatus</i>	0	0	0	0	0	0	0	1	0	0	1	
	<i>Apherusa chierighini</i>	0	0	0	0	0	0	0	0	1	0	0	
	<i>Corophium acutum</i>	0	0	0	0	0	0	0	0	0	1	0	
MOLLUSCA	<i>Corophium spp.</i>	0	0	0	0	0	0	0	1	0	1	1	
	<i>Abra nitida</i>	0	3	1	1	2	1	4	4	8	6	1	
	<i>Chlamys varia</i>	0	0	0	0	0	0	0	0	0	0	0	
	<i>Corbula gibba</i>	0	2	88	6	1	24	22	21	11	1	0	
	<i>Phaxas adriaticus</i>	0	1	0	0	0	0	0	2	3	0	1	
	<i>Nassarius reticulatus</i>	0	0	2	0	0	1	0	0	0	0	0	
	<i>Nucula nitidosa</i>	0	0	2	1	1	4	11	0	10	12	1	
	<i>Nuculana pella</i>	0	0	0	0	0	0	0	0	2	0	0	
	<i>Pitar rudis</i>	0	0	3	0	0	2	0	2	2	1	0	
	<i>Psammobia fervensis</i>	0	0	0	0	0	0	0	0	0	1	0	
	<i>Saxicavella jeffreysi</i>	0	0	0	0	0	3	0	0	0	0	0	
	<i>Scrobicularia plana</i>	0	0	0	0	0	0	0	0	0	2	0	
	<i>Smithiella smithi</i>	0	0	0	0	0	0	0	0	0	1	1	
	POLYCHAETA	<i>Phylo cuvieri</i>	0	0	0	0	0	0	0	0	1	1	0
		<i>Aricidea assimilis</i>	0	4	0	0	1	2	1	0	1	1	0
<i>Polydora flava</i>		0	0	0	0	0	0	1	0	6	0	0	
<i>Pronosio caspersi</i>		41	0	1	0	0	0	0	0	0	0	0	
<i>Pronosio malmgreni</i>		3	31	0	0	3	0	9	3	65	2	2	
<i>Pronosio multibranchiata</i>		0	110	0	0	5	1	44	9	52	7	16	
<i>Pseudopolydora antennata</i>		0	0	0	0	0	0	0	0	0	0	2	
<i>Spiophanes kroyeri</i>		0	0	7	0	0	0	0	8	28	13	12	
<i>Magelona allenii</i>		0	0	0	0	0	0	0	0	3	0	0	
<i>Magelona minuta</i>		0	1	0	0	0	0	0	0	5	1	0	
<i>Poecilochaetus fauchaldi</i>		0	0	0	0	0	0	1	3	1	1	1	
...		

In questo esempio circa 1/3 dei dati è $\neq 0$

Sampling stations	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12		
Date	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00	17/05/00		
Time	16.05	15.40	15.20	14.50	14.30	14.00	13.35	12.20	11.55	11.25	11.00		
Depth (m)	9.9	14.8	19.0	22.8	25.2	25.9	26.3	27.1	28.1	28.0	28.1		
Coordinates North	4991467	4991855	4992243	4992628	4993022	4993396	4993792	4994168	4994566	4994951	4995336		
Coordinates East	2317820	2318854	2319883	2320915	2321942	2322981	2324016	2325031	2326067	2327096	2328129		
taxa	ID												
ANPHIPODA	<i>Ampelisca diadema</i>	AMPDIA	0	2	0	13	16	3	6	14	15	2	6
	<i>Ampelisca sarsi</i>	AMPSAR	0	10	9	12	13	4	0	0	0	10	0
	<i>Aora spinicornis</i>	AORSPI	0	0	0	0	0	0	1	0	0	0	0
	<i>Leptocheirus pectinatus</i>	LEPPPEC	0	0	0	0	0	0	0	1	0	0	1
	<i>Apherusa chiereghini</i>	APHCHI	0	0	0	0	0	0	0	0	1	0	0
	<i>Corophium acutum</i>	CORACU	0	0	0	0	0	0	0	0	0	1	0
	<i>Corophium spp.</i>	CORSPP	0	0	0	0	0	0	0	1	0	1	1
MOLLUSCA	<i>Abra nitida</i>	ABRNIT	0	3	1	1	2	1	4	4	8	6	1
	<i>Chlamys varia</i>	CHLVAR	0	0	0	0	0	0	0	0	0	1	0
	<i>Corbula gibba</i>	CORGBIB	0	2	88	6	1	24	22	21	11	1	0
	<i>Phaxas adriaticus</i>	PHAADR	0	1	0	0	0	0	0	2	3	0	1
	<i>Nassarius reticulatus</i>	NASRET	0	0	2	0	0	1	0	0	0	0	0
	<i>Nucula nitidosa</i>	NUCNIT	0	0	2	1	1	4	11	0	10	12	1
	<i>Nuculana pella</i>	NUCPEL	0	0	0	0	0	0	0	0	2	0	0
	<i>Pitar rudis</i>	PITRUD	0	0	3	0	0	2	0	2	2	1	0
	<i>Psammobia fervensis</i>	PSAFER	0	0	0	0	0	0	0	0	0	1	0
	<i>Saxicavella jeffreysi</i>	SAXJEF	0	0	0	0	0	3	0	0	0	0	0
	<i>Scrobicularia plana</i>	SCRPLA	0	0	0	0	0	0	0	0	0	2	0
	<i>Smithiella smithi</i>	SMISMI	0	0	0	0	0	0	0	0	0	1	1
POLYCHAETA	<i>Phylo cuvieri</i>	PHYCUV	0	0	0	0	0	0	0	0	1	1	0
	<i>Aricidea assimilis</i>	ARIASS	0	4	0	0	1	2	1	0	1	1	0
	<i>Polydora flava</i>	POLFLA	0	0	0	0	0	0	1	0	6	0	0
	<i>Prionospio caspersi</i>	PRICAS	41	0	1	0	0	0	0	0	0	0	0
	<i>Prionospio malmgreni</i>	PRIMAL	3	31	0	0	3	0	9	3	65	2	2
	<i>Prionospio multibranchiata</i>	PRIMUL	0	110	0	0	5	1	44	9	52	7	16
	<i>Pseudopolydora antennata</i>	PSEANT	0	0	0	0	0	0	0	0	0	0	2
	<i>Spiophanes kroeyeri</i>	SPIKRO	0	0	7	0	0	0	0	8	28	13	12
	<i>Magelona allenii</i>	MAGALL	0	0	0	0	0	0	0	0	3	0	0
	<i>Magelona minuta</i>	MAGMIN	0	1	0	0	0	0	0	0	5	1	0
	<i>Poecilochaetus fauchaldi</i>	POEFAU	0	0	0	0	0	0	1	3	1	1	1
...

La cassetta degli attrezzi.

- **Ordinamento** (PCA, MDS, NMDS, CA, DCA, CCA, etc.)
- **Classificazione** (algoritmi gerarchici, k-means, reti neurali, etc.)
- **Analisi spaziale** (correlogrammi, variogrammi, kriging, co-kriging, etc.)
- **Analisi di serie** (periodogrammi, runs tests, cross-correlation, cross-association, etc.)
- **Confronti fra dati multivariati** (MRPP, test di Mantel, INDVAL, etc.)
- **Reti neurali**
- ...

Misure di distanza

- Distanza euclidea

$$D_{jk} = \sqrt{\sum_{i=1}^p (x_{ij} - x_{ik})^2}$$

- Distanza di Manhattan

$$D_{jk} = \sum_{i=1}^p |x_{ij} - x_{ik}|$$

- Distanza di Canberra

$$D_{ij} = \sum_{i=1}^p \frac{|x_{ij} - x_{ik}|}{(x_{ij} + x_{ik})}$$

- Metrica di Minkowski

$$D_{jk} = \sqrt[r]{\sum_{i=1}^p |x_{ij} - x_{ik}|^r}$$

Misure di similarità

Per i coefficienti binari
(dati di presenza/assenza)

		Osservazione <i>j</i>	
		1	0
Osservazione <i>k</i>	1	<i>a</i>	<i>b</i>
	0	<i>c</i>	<i>d</i>

$p = a + b + c + d$

- Numero di taxa presenti in entrambi i campioni
- Numero di taxa presenti solo nel campione *k*
- Numero di taxa presenti solo nel campione *j*
- Numero di taxa **assenti** in entrambi i campioni

Misure di similarità

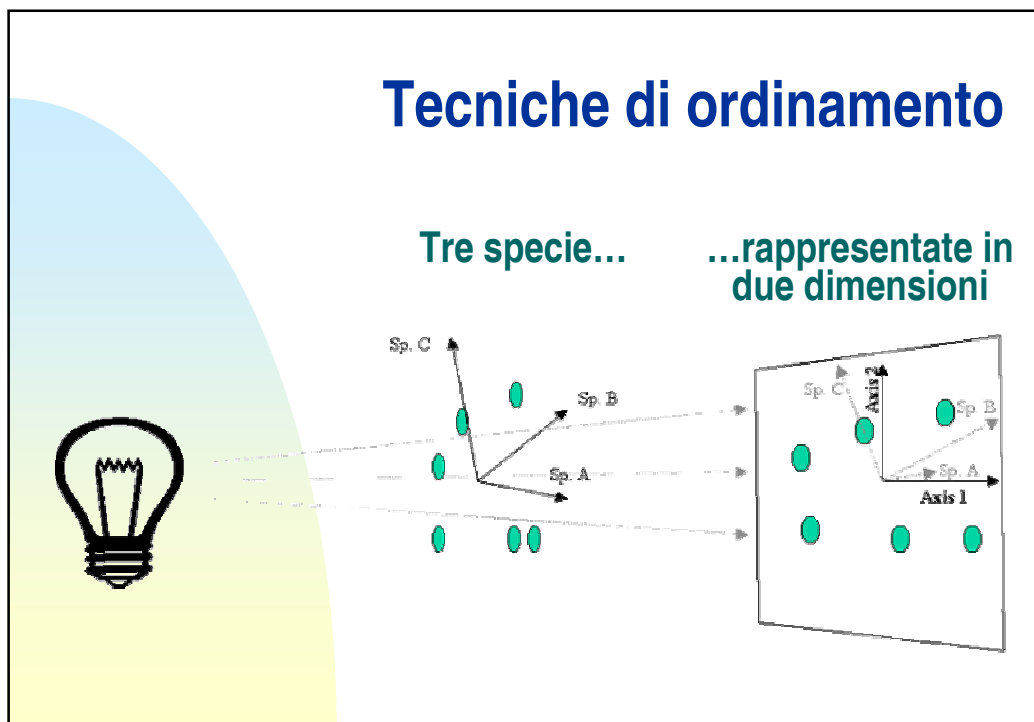
Dati binari (qualitativi)	■ Indice di concordanza	$S_{jk} = \frac{a+d}{a+b+c+d}$	simmetrico
	■ Indice di Jaccard	$S_{jk} = \frac{a}{a+b+c}$	asimmetrico
	■ Indice di Sørensen	$S_{jk} = \frac{2a}{2a+b+c}$	asimmetrico
Dati quantitativi	■ Coeff. di Bray-Curtis	$S_{ij} = \frac{\sum_{i=1}^s x_{ij} - x_{ik} }{\sum_{i=1}^s (x_{ij} + x_{ik})}$	

Similarità e dissimilarità

Dissimilarità = 1 - Similarità

Un coefficiente di dissimilarità è di tipo metrico se:

1. $D_{ij}=0$ se $j=k$ 2. $D_{jk}>0$ se $j \neq k$ 3. $D_{jk}=D_{kj}$ 4. $D_{jk}+D_{kh} \geq D_{jh}$ (assioma della disuguaglianza triangolare)	}	Semi-metrica
---	---	---------------------



Perchè l'ordinamento?

"Ordination primarily endeavors to represent sample and species relationships as faithfully as possible in a low-dimensional space."

Gauch (1982)

Perchè...

- E' impossibile visualizzare efficacemente insiemi di dati multidimensionali complessi
- Un'analisi multivariata è più economica e più efficiente di n analisi univariate
- Gli assi dello spazio ridotto di solito rappresentano gradienti ambientali interpretabili
- Se si effettuano anche test statistici, si evitano i problemi legati alle comparazioni multiple
- Concentrando l'attenzione solo su alcuni assi si evita di considerare il "rumore"

E inoltre...

- Fino a non molto tempo fa l'obiettivo dei metodi di ordinamento era di tipo prettamente descrittivo (più un'arte che una scienza, quindi!).
- Con la CCA, i test statistici d'ipotesi sono stati accoppiati alle tecniche di ordinamento, superando l'approccio descrittivo (cfr. ter Braak 1985)

Analisi indiretta di gradiente

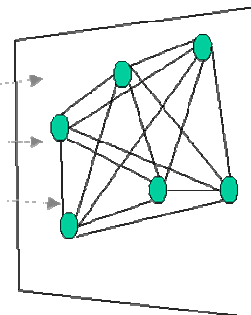
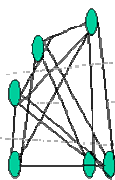
- Metodi basati su distanze
 - ◆ Ordinamento polare (Bray-Curtis)
 - ◆ Analisi delle Coordinate Principali (PCoA)
 - ◆ Multidimensional Scaling Nonmetrico (NMDS)
- Metodi basati su autovalori/autovettori
 - ◆ Modello lineare
 - ☞ Analisi delle Componenti Principali (PCA)
 - ◆ Modello unimodale
 - ☞ Analisi delle Corrispondenze (CA)
 - ☞ Analisi delle Corrispondenze Detrendizzata (DCA)

Le distanze nello spazio originale sono approssimate al meglio dalle distanze (euclidee) nello spazio ridotto (cioè nell'ordinamento)

PCoA

2 dimensioni (assi)

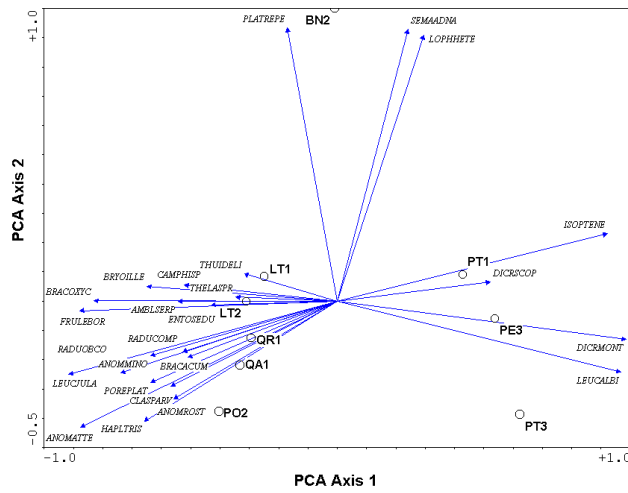
n dimensioni (taxa)



Rototraslazione rigida degli assi originali.

Si assumono relazioni lineari fra le variabili.

PCA

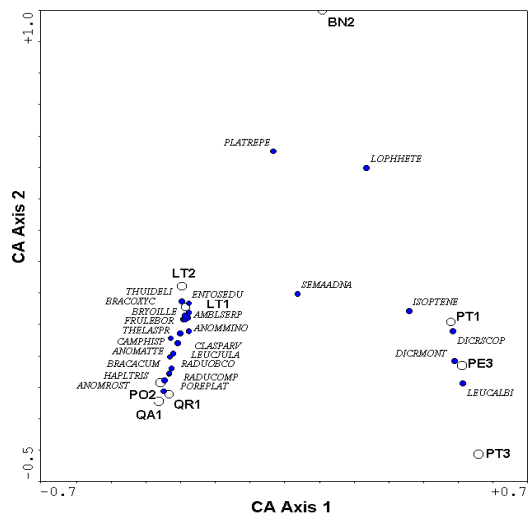


Metrica di χ^2

Si assumono risposte unimodali dei taxa.

Ordinamento simultaneo di campioni e taxa (oggetti e variabili).

CA



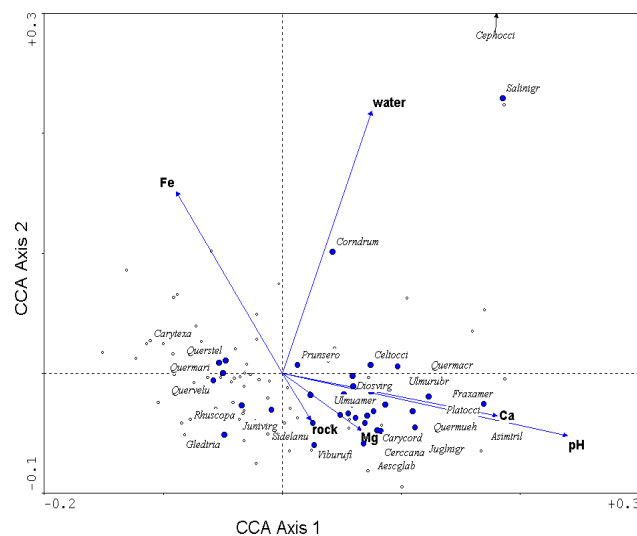
Analisi diretta di gradiente

- Modello lineare
 - ◆ **Analisi di Ridondanza (RDA)**
- Modello unimodale
 - ◆ **Analisi Canonica delle Corrispondenze (CCA)**
 - ◆ **Analisi Canonica delle Corrispondenze Detrendizzata (DCCA)**

E' concettualmente affine alla CA, ma l'ordinamento di campioni e taxa è vincolato a combinazioni lineari di variabili ambientali.

Rappresentazione simultanea di campioni, taxa e gradienti ambientali.

CCA



Clustering (classificazione)

■ **Obiettivi:**

- ◆ Formare gruppi omogenei di entità (osservazioni, campioni, siti, specie, etc.)
- ◆ Identificare discontinuità (nello spazio, nel tempo)

■ **Algoritmi:**

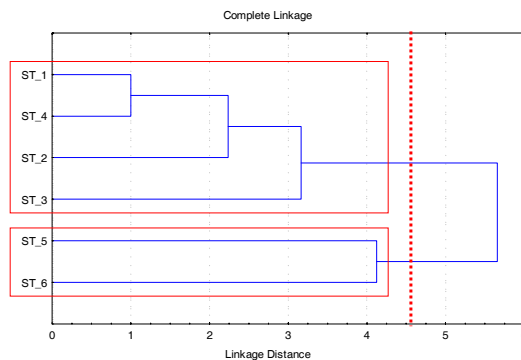
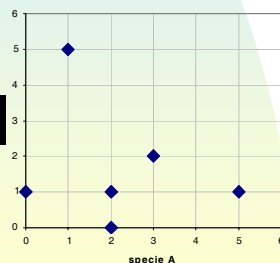
- ◆ Gerarchici
 - ◆ Non gerarchici
- } •Vincolati
} •Non vincolati

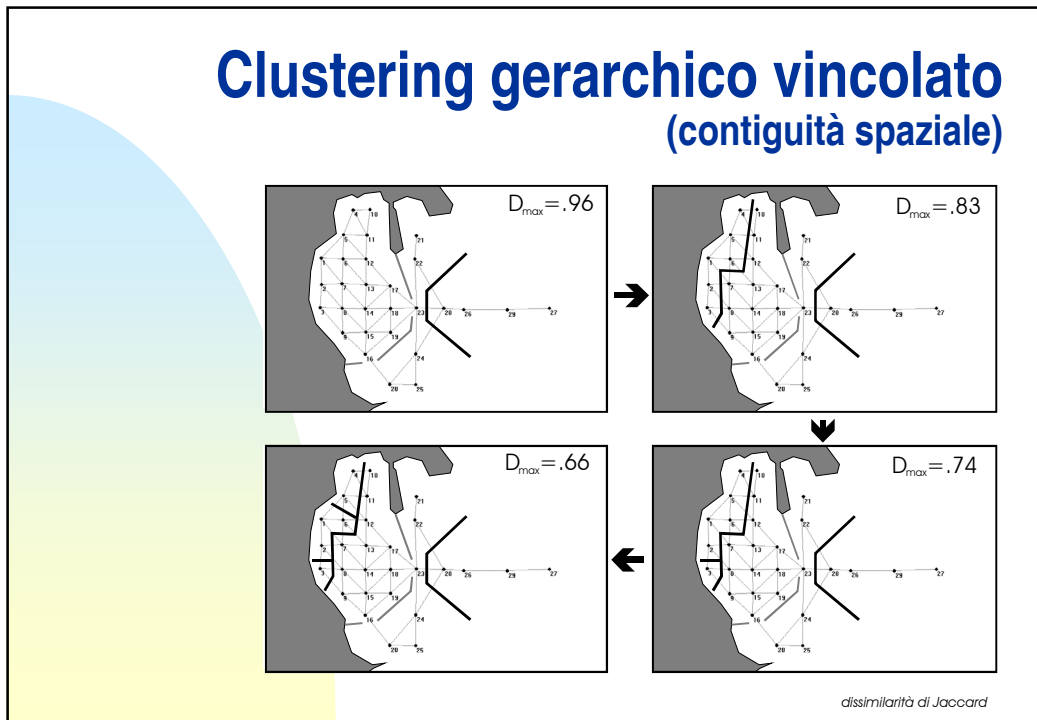
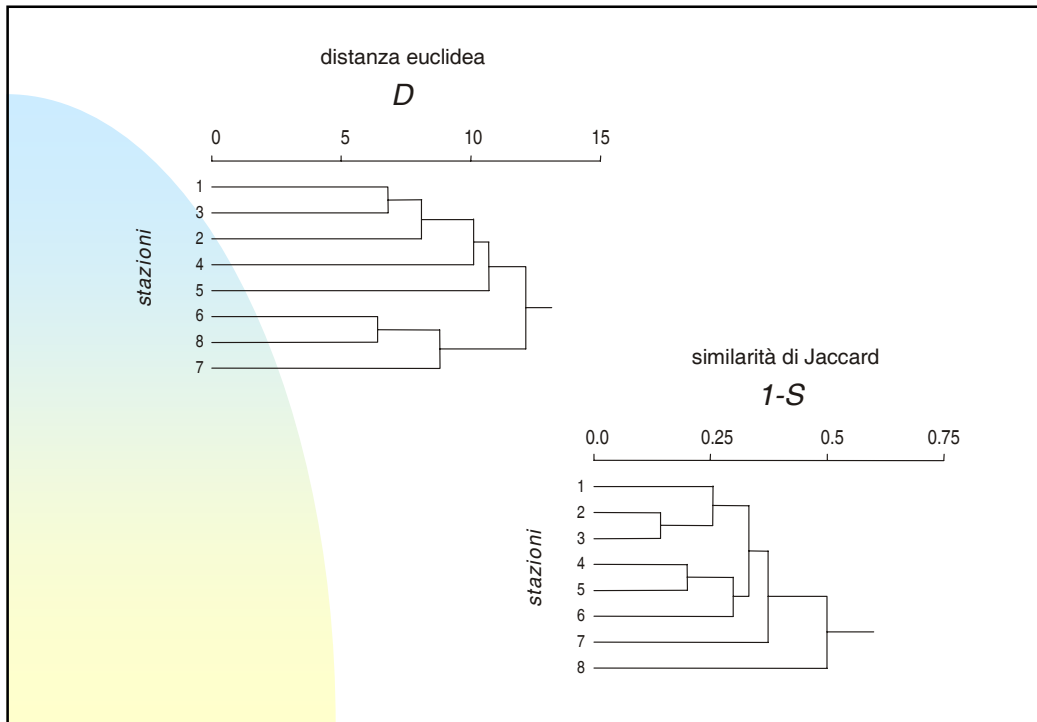
Clustering gerarchico

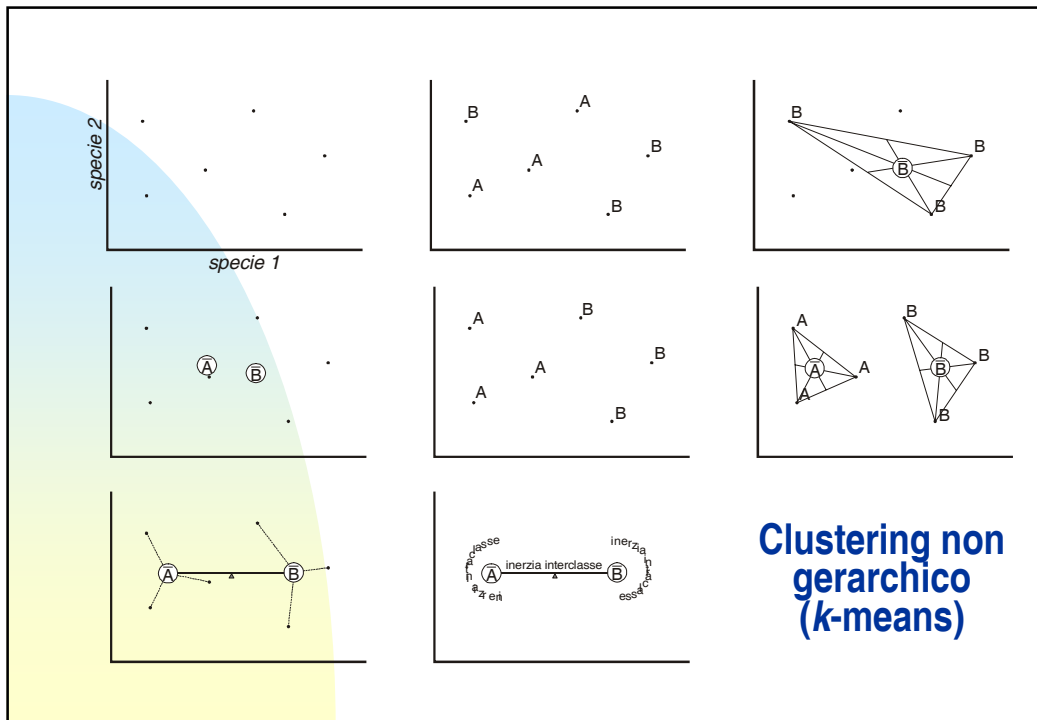
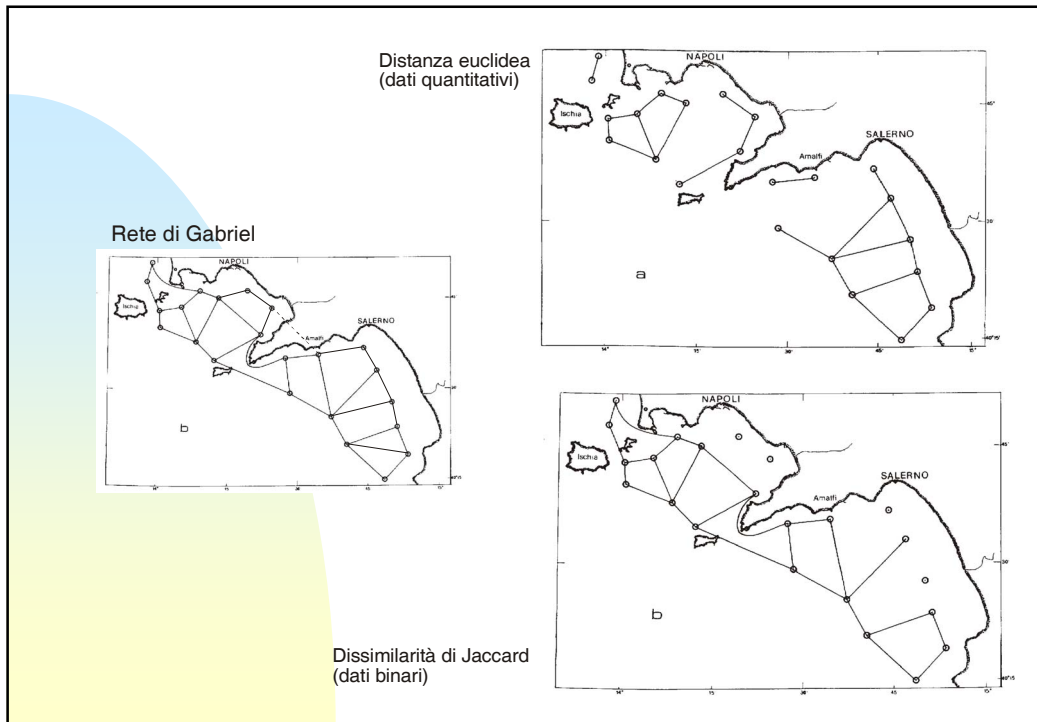
siti	specie	
	A	B
1	2	1
2	3	2
3	5	1
4	2	0
5	0	1
6	1	5

$$D_{jk} = \sqrt{\sum_{i=1}^p (x_{ij} - x_{ik})^2}$$

siti	siti					
	1	2	3	4	5	6
1	0.000	1.414	3.000	1.000	2.000	4.123
2	1.414	0.000	2.236	2.236	3.162	3.606
3	3.000	2.236	0.000	3.162	5.000	5.657
4	1.000	2.236	3.162	0.000	2.236	5.099
5	2.000	3.162	5.000	2.236	0.000	4.123
6	4.123	3.606	5.657	5.099	4.123	0.000







Self Organizing Maps

	dove	hen	duck	goose	owl	hawk	eagle	fox	dog	wolf	cat	tiger	lion	horse	zebra	cow
<i>is</i>	<i>small</i>	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0
	<i>medium</i>	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
	<i>big</i>	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
<i>has</i>	<i>2 legs</i>	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	<i>4 legs</i>	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	<i>hair</i>	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	<i>hooves</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	<i>mane</i>	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0
<i>feathers</i>	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
<i>likes to</i>	<i>hunt</i>	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0
	<i>run</i>	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0
	<i>fly</i>	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0
	<i>swim</i>	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0

horse			cow		duck											
		zebra													goose	
														hen	dove	
		lion		tiger												
															owl	
wolf																hawk
				fox										cat		
		dog														eagle

Test basati su permutazioni (MCP)

- Sono un caso speciale dei test di randomizzazione, che utilizzano serie di numeri casuali formulare delle inferenze statistiche.
- La potenza di calcolo dei moderni PC ha reso possibile la loro applicazione diffusa.
- Questi metodi non richiedono che siano soddisfatte particolari assunzioni circa la distribuzione dei dati.
- Quindi, questi metodi sono molto più adatti dei tradizionali test statistici (es. *t*-tests, ANOVA, etc.) in applicazioni ecologiche.

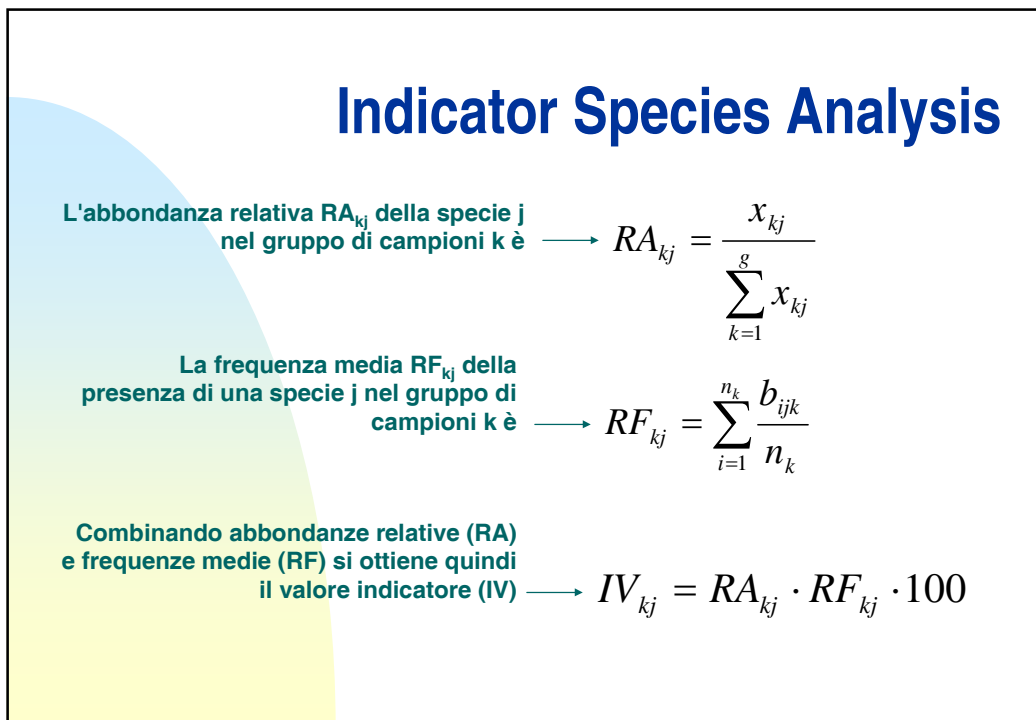
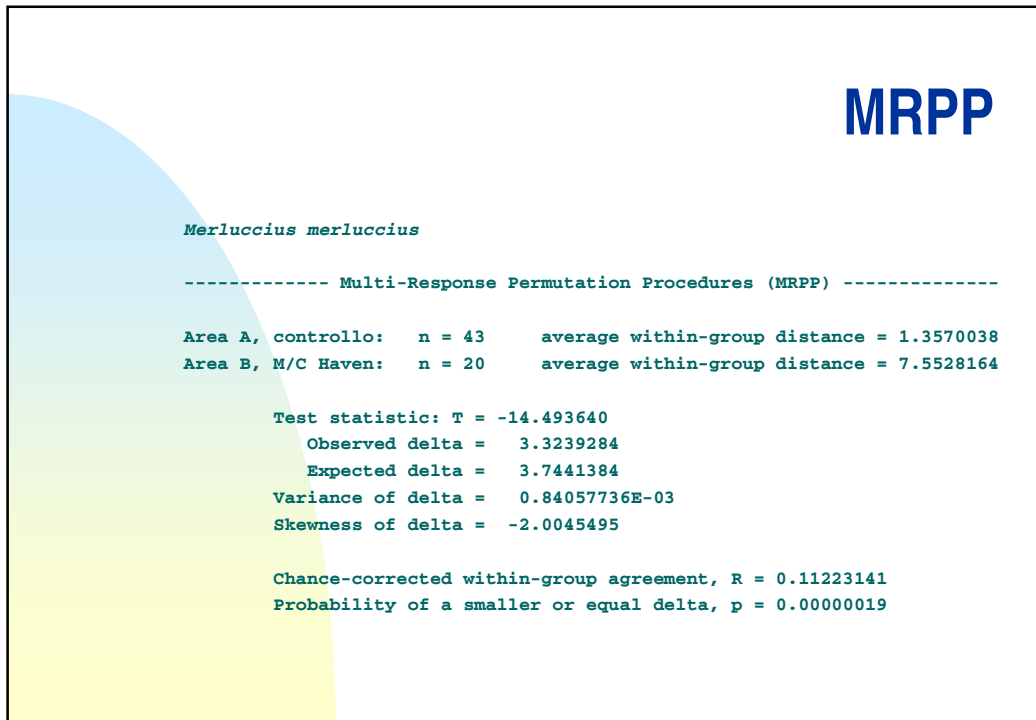
Test basati su permutazioni (MCP)

- Si definisce una statistica il cui valore sia proporzionale all'intensità del processo o della relazione studiati
- Si definisce un'ipotesi nulla H_0
- Si crea un set di dati basati sul "rimiscolamento" di quelli realmente osservati (la modalità di "rimiscolamento" viene definita in funzione dell'ipotesi nulla)
- Si ricalcola la statistica di riferimento e si compara il valore con quello osservato
- Si ripetono gli ultimi due punti molte volte (es. 1000 volte)
- Se la statistica osservata è maggiore del limite ottenuto nel 95% dei casi basati su "rimiscolamento", si **rigetta H_0**

MRPP

Se Δ è la distanza media intragruppo osservata fra i gruppi definiti *a priori*, pesata per la dimensione dei diversi gruppi, allora:

$$T = \frac{\Delta_{osservato} - \Delta_{atteso}}{\sqrt{\sigma_{\Delta}^2}} \qquad R = 1 - \frac{\Delta_{osservato}}{\Delta_{atteso}}$$



Indicator Species Analysis

■ *Merluccius merluccius* - Indicator Values

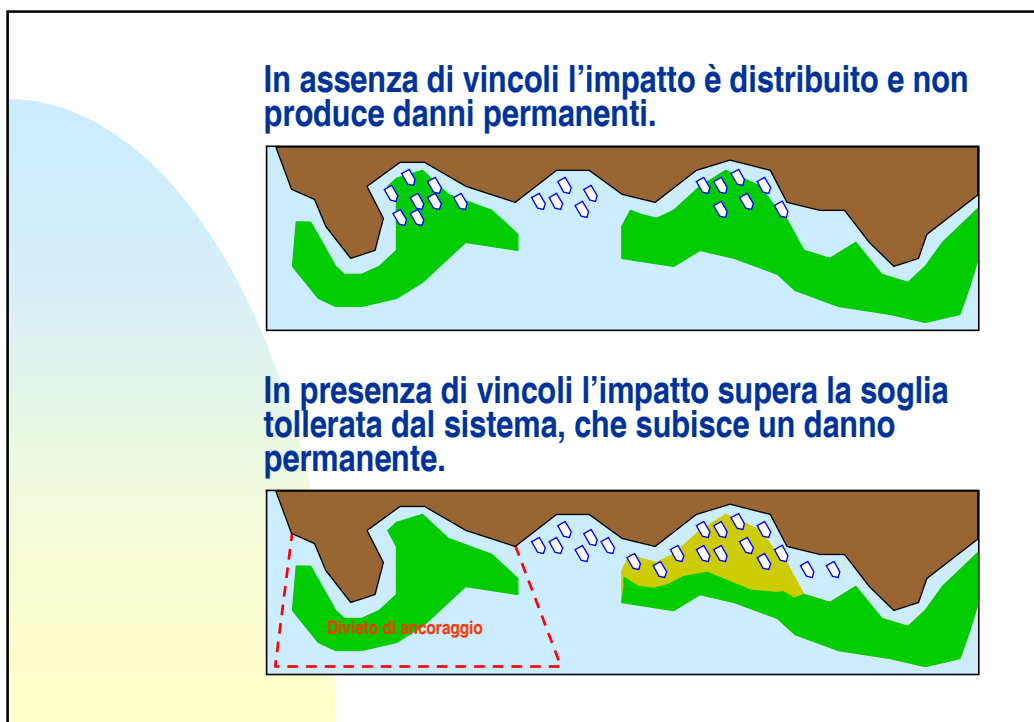
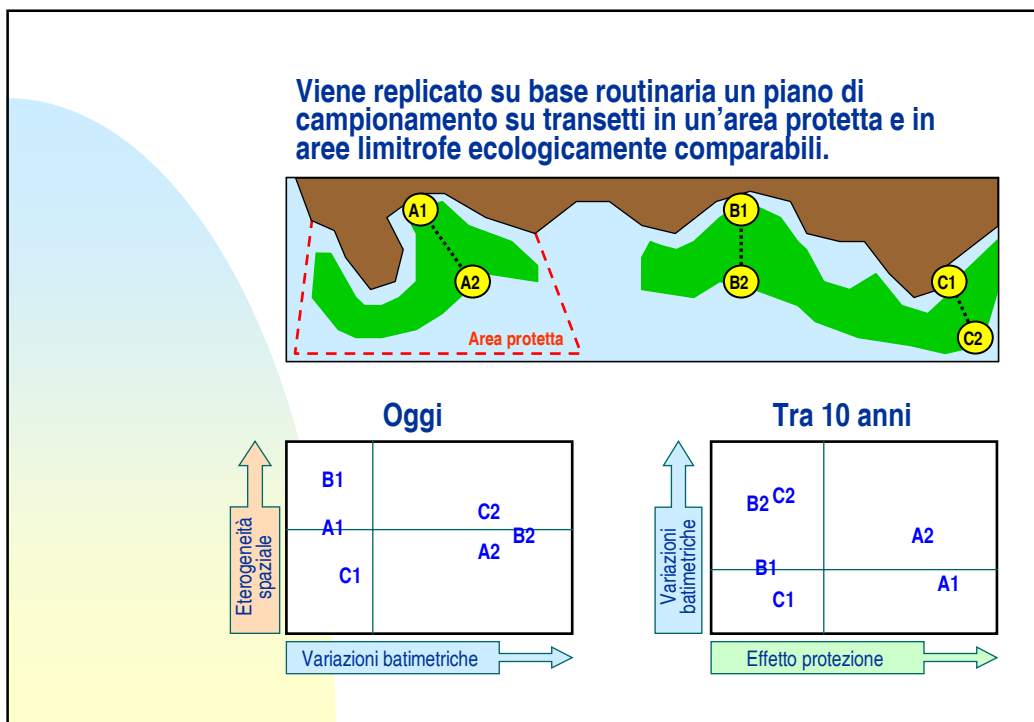
■ Group: A B

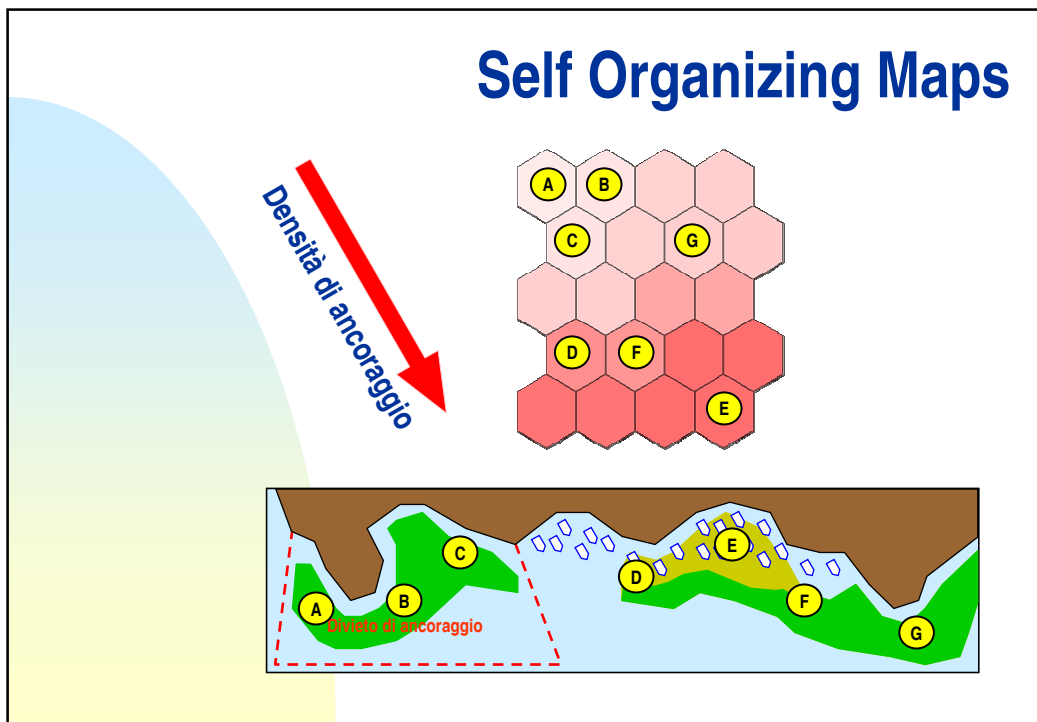
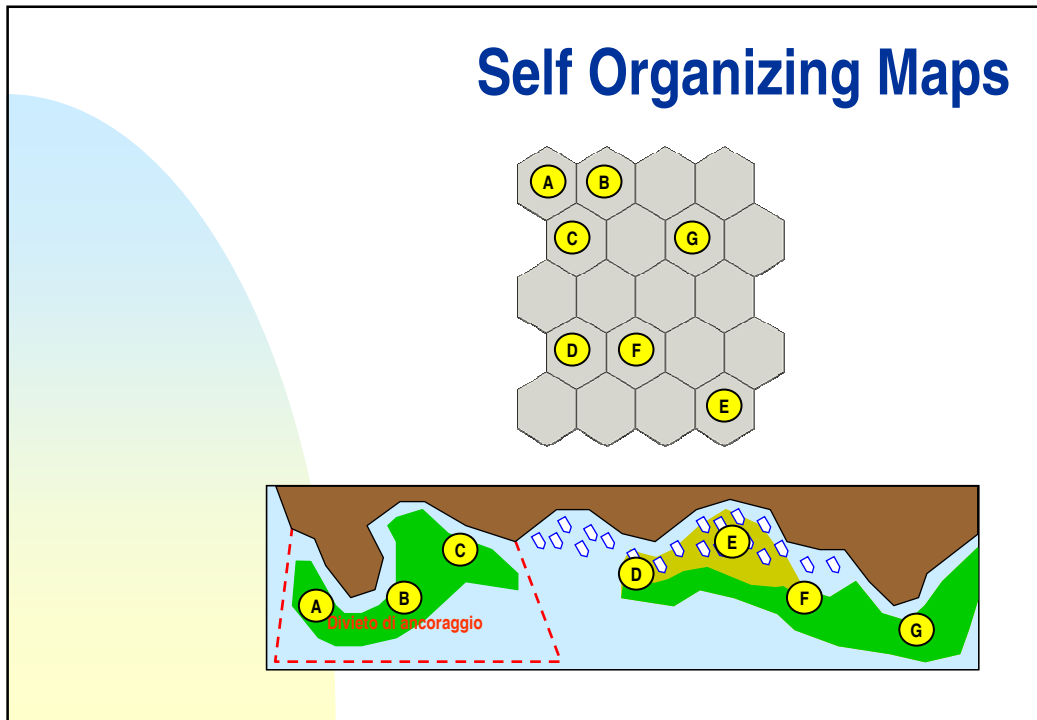
■ Number of items: 43 20

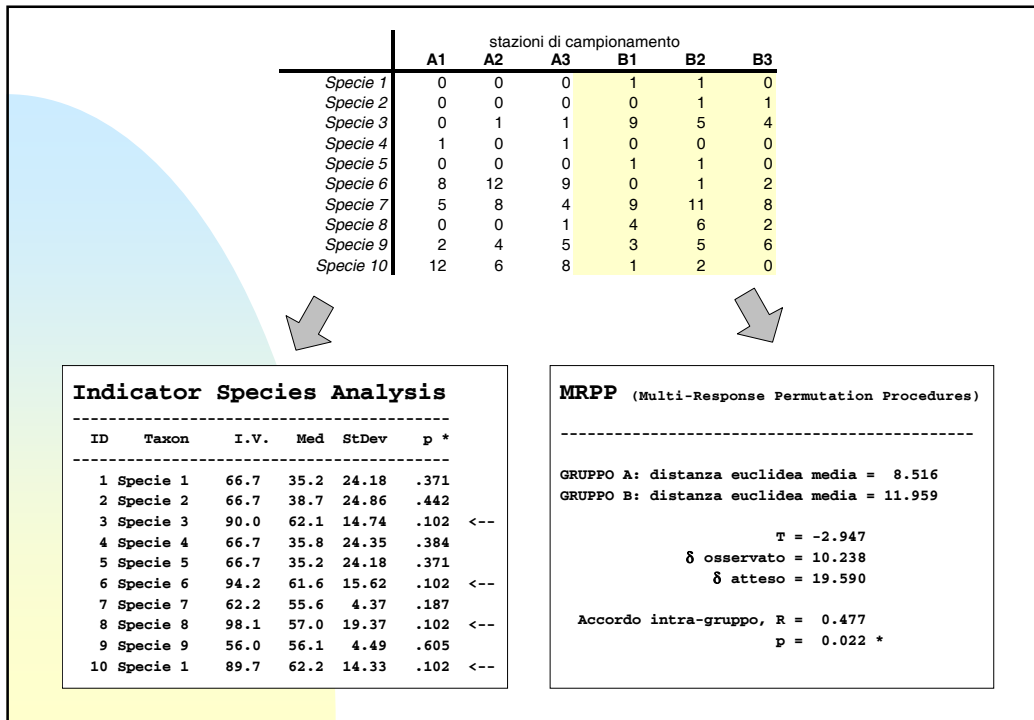
n	ID	Avg	Max	INDVAL	p	taxon	
1	EUFASI	24	44	4	44	0.017	Eufasiacei
2	THYSAN	12	25	0	25	0.004	Thysanopoda aequalis
3	RESPES	13	26	26	1	0.046	Resti pesci
4	RESCRO	10	20	0	20	0.009	Resti crostacei
5	MISIDA	2	5	5	0	0.595	Misidacei nc
6	DECAPO	2	3	2	3	0.999	Decapodi nc
7	CEPALO	2	5	5	0	0.554	Cefalopodi
8	CHLORO	3	5	0	5	0.299	Chlorotocus crassicornis
9	CRANGO	1	2	2	0	0.999	Crangon sp
10	SARDIN	3	5	0	5	0.299	Sardina pilchardus
11	ROCINE	3	5	0	5	0.307	Rocinela sp
12	POLICH	1	2	2	0	0.999	Policheti

Problema: valutare gli effetti della protezione sulla struttura delle comunità.

- La struttura delle comunità varia in funzione delle risposte delle specie a gradienti ambientali complessi.
- L'impatto antropico altera o modula questi gradienti.
- La protezione dell'ambiente riduce l'impatto antropico o lo trasferisce in altre aree.



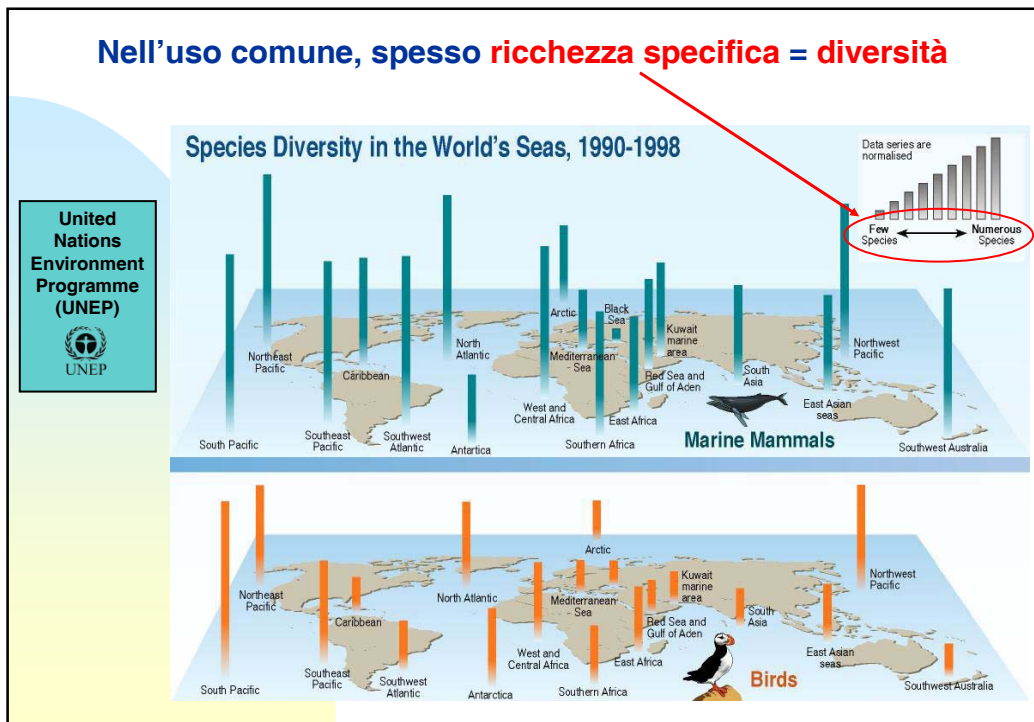




La ricchezza specifica



- La ricchezza specifica è il numero di specie che compongono una comunità.
- Il termine è stato coniato da McIntosh (1967).
- Rappresenta la più elementare misura delle diversità.
- **Attenzione!** Dipende dalla dimensione del campione.



La diversità

- In senso strettamente ecologico, la diversità di una comunità deve esprimere la complessità della sua struttura.
- La diversità è massima quando la probabilità che due individui estratti a caso appartengano alla stessa specie è minima.
- Ovvero, la diversità è massima quando tutte le specie hanno abbondanze uguali.

- Proporzione di una determinata specie nella comunità:

$$p_i = \frac{n_i}{\sum_{i=1}^s n_i}$$
- Probabilità di estrazione casuale di due individui della i-ma specie:

$$P = p_i \cdot p_i = p_i^2$$
- Probabilità di estrazione casuale di due individui di una qualsiasi specie:

$$P = \sum_{i=1}^s p_i^2$$
- Diversità $\propto P^{-1}$

diversità

alta

bassa

Indici di diversità

- Margalef:
$$D = \frac{s-1}{\ln(N)}$$
- Simpson:
$$D = \frac{1}{\sum_{i=1}^s p_i^2}$$
- Shannon-Wiener (o Shannon-Weaver):
$$H = -\sum_{i=1}^s p_i \log_2 p_i$$

Evenness

- I valori degli indici di diversità non sono sempre comparabili fra loro e dipendono dai limiti entro i quali essi possono effettivamente variare.
- La evenness è una misura di diversità normalizzata su una scala prefissata (es. da 0 a 1) e consente di effettuare tali confronti.
- La evenness può essere definita come:

$$J = \frac{H}{H_{\max}} \quad \text{oppure} \quad J = \frac{H - H_{\min}}{H_{\max} - H_{\min}}$$

