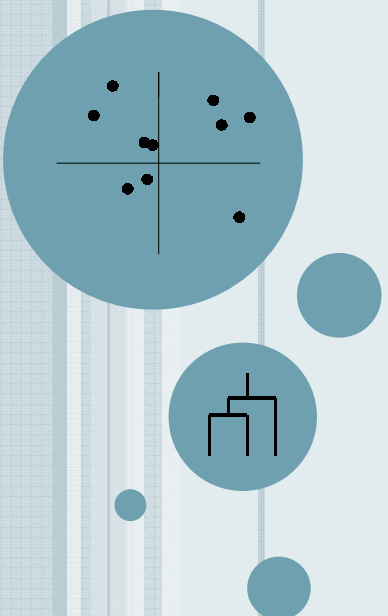


# GRADIENTOVÁ ANALÝZA



# HISTORIE

## WHITTAKER 1956 - PŘÍMÁ GRADIENTOVÁ ANALÝZA

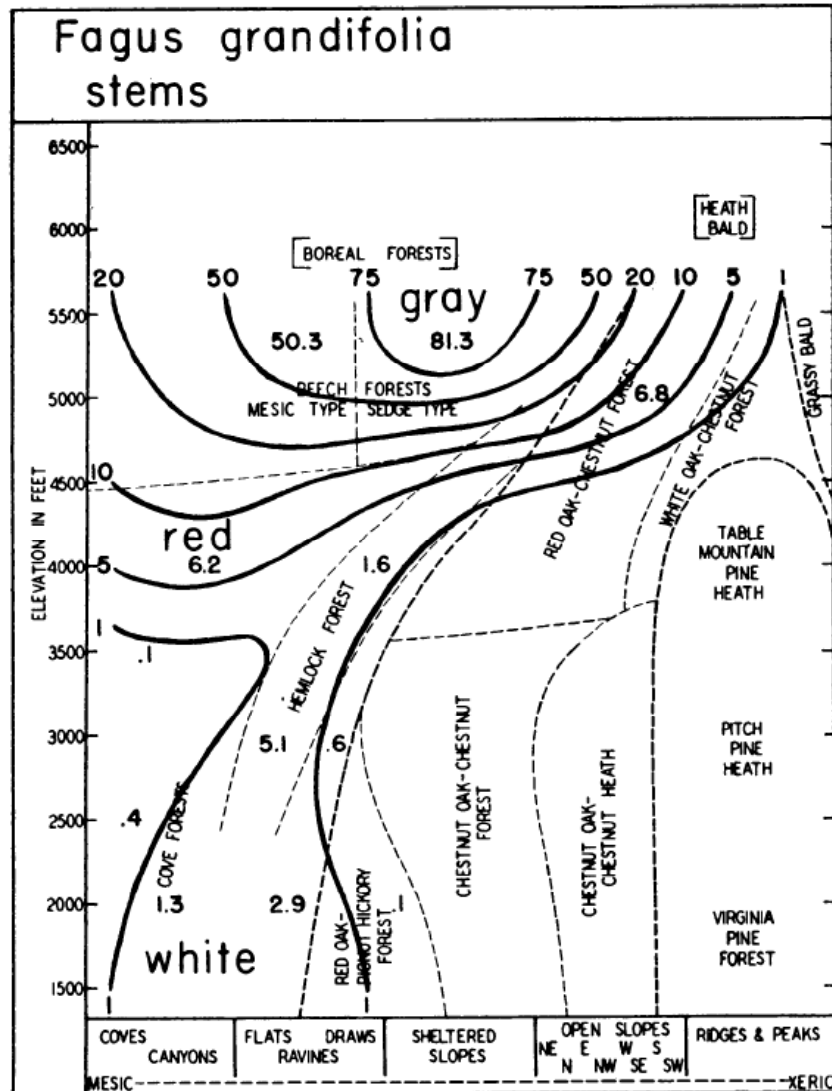


Fig. 1. Mature, forest-covered topography of the Great Smoky Mountains, a view of the southeast slope from Frye Mountain near Bryson City, North Carolina. Reproduced by permission of W. M. Cline Co., Chattanooga, Tenn.

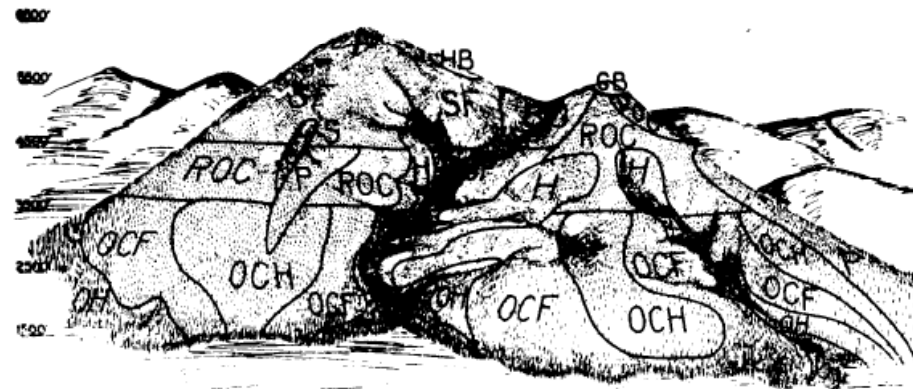


FIG. 21. Topographic disposition of vegetation types.

# HISTORIE

## BRAY & CURTIS 1957 - NEPŘÍMÁ GRADIENTOVÁ ANALÝZA

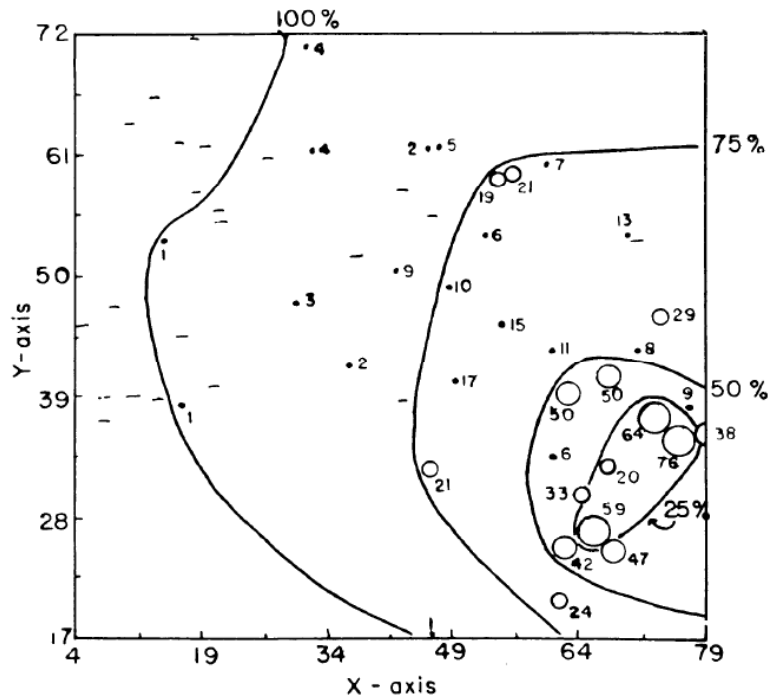


FIG. 3. Demonstration of dominance distribution of *Tilia americana* within the x-y ordination. Each circle or dash represents a stand location. Actual dominance figures in basal area per 100 sq. in. per acre at breast height are given beside each stand location. Values in the upper 25% are represented by the largest circle, values in the 50 to 26% quartile by medium sized circles, values in the 75-51% quartile by small circles, and values in the 100 to 76% quartile by dots. Contour lines are drawn around the 4 quartile lower limits in such a manner as to include all examples of the indicated size class whether or not lesser size class values are present.

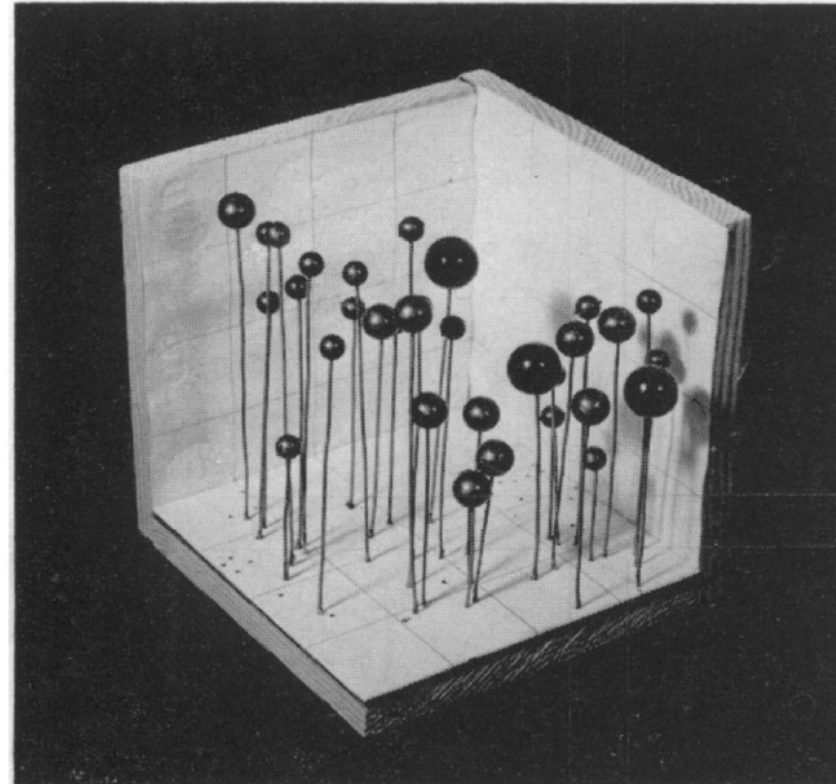
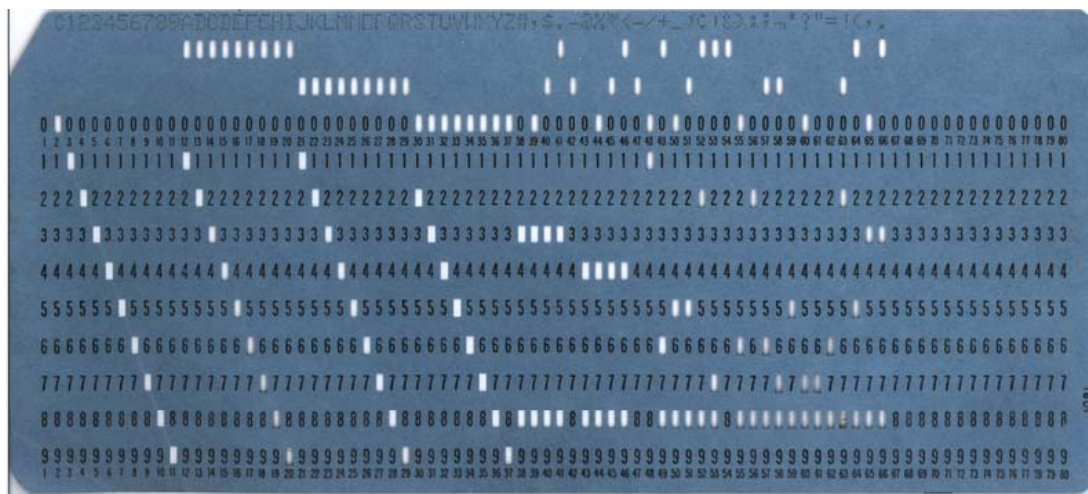


FIG. 7. Three-dimensional model of the dominance behavior of *Quercus borealis* within the ordination. The 3 sizes of spheres indicate the top 3 quartiles of dominance per acre. Stands of the lowest quartile and without the species are represented with holes which appear as dots in the figure. The x axis is on base of model at front from left to right; y axis on base from front to rear; z axis in vertical plane from below to above.

# HISTORIE

*“... P.C.A. cannot be carried out practically without a computer, unless the number of individuals is very small. European phytosociologists generally don't have computers at their disposal. Hence they are in need of approximative methods that can be carried out by hand or with simple desk calculators... “*

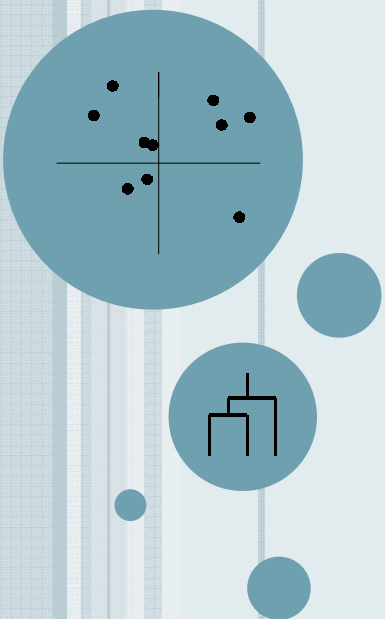
van der Maarel, 1969, *Vegetatio*



# METODY GRADIENTOVÉ ANALÝZY

Data, která máme:		Apriorní znalost vztahů mezi druhy a prostředím?	Použijeme:	Dostaneme:
počet charakteristik prostředí	počet druhů			
1, $n$	1	ne	<b>regrese</b>	závislost druhu na prostředí
žádné	$n$	ano	<b>kalibrace</b>	odhady hodnot charakteristik prostředí
žádné	$n$	ne	<b>nepřímá ordinace</b>	osy variability v druhovém složení
1, $n$	$n$	ne	<b>přímá ordinace</b>	variabilita ve druhovém složení vysvětlená charakteristikami prostředí

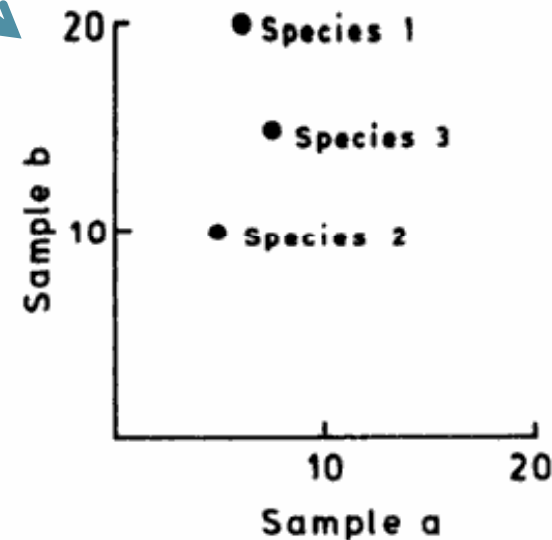
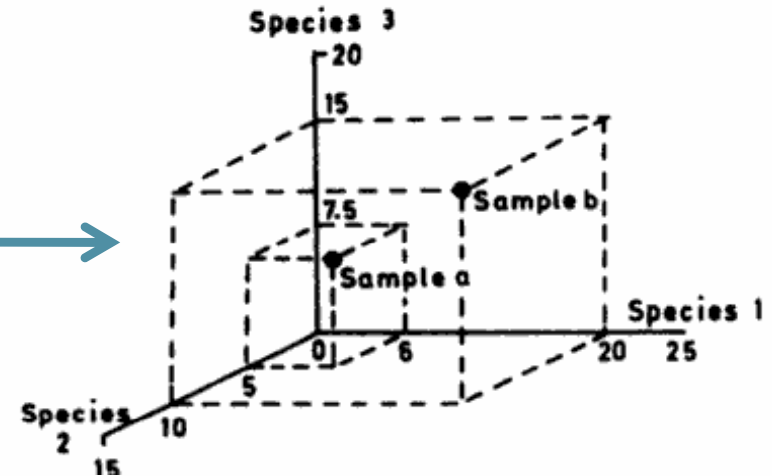
# NEPŘÍMÁ ORDINAČNÍ ANALÝZA



# KONCEPCE MNOHOROZMĚRNÉHO PROSTORU

## Prostor může být definován

- 1) **druhy (*species space*)**
  - druhy jsou osami mnohorozměrného prostoru
  - vzorky jsou body v tomto prostoru
  - zobrazení původní matice druhu-vzorky
- 2) **vzorky (*sample space*)**
  - vzorky jsou osami mnohorozměrného prostoru
  - druhy jsou body v tomto prostoru
  - zobrazení původní matice druhu-vzorky
- 3) **vzdálenostmi mezi druhy (*species dissimilarity space*)**
  - osami jsou vzdálenosti k druhu x
  - druhy jsou body v tomto prostoru
  - zobrazení matice nepodobností mezi druhy
- 4) **vzdálenostmi mezi snímky (*sample dissimilarity space*)**
  - osami jsou vzdálenosti ke vzorku y
  - vzorky jsou body v tomto prostoru
  - zobrazení matice nepodobností mezi vzorky
- 5) **ekologickými gradienty (*ecological space*)**
  - osami jsou ekologické gradienty
  - jako body do něj lze zobrazit druhy i vzorky



Zuur et al. (2007)

# ORDINACE

## OPODSTATNĚNÍ

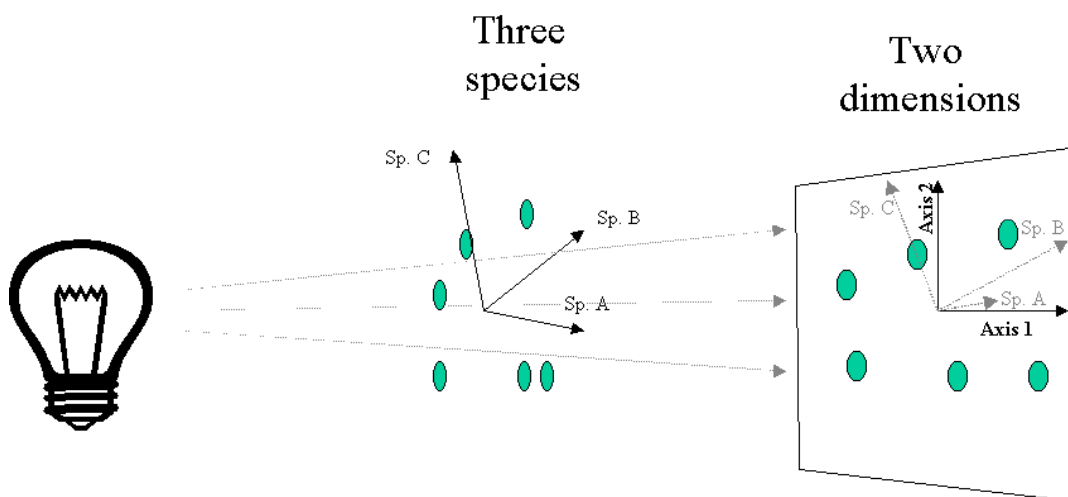
- jeden gradient prostředí většinou ovlivňuje chování (abundanci) několika druhů najednou – základní chování společenstev
  - druhová data jsou **redundantní** – pokud znám chování (abundanci) jednoho druhu, můžu do určité míry odhadnout chování i některých dalších druhů
  - díky této redundanci je možné (a hlavně smysluplné) zredukovat mnohorozměrný prostor, ve kterém jsou druhy/vzorky rozmístěny (prostory 1-4), na několik málo dimenzí *ekologického prostoru* (protor č.5)
- pokud by chování druhů bylo na sobě úplně nezávislé, existovala by celá řada ekvivalentních možností, jak mnohorozměrný prostor zredukovat, a ani jedna by nepřinesla nic nového



# ORDINACE

## RŮZNÉ FORMULACE PROBLÉMU

- 1) hledání skrytých proměnných (ordinačních os)
  - najdi několik proměnných (ordinačních os), které nejlépe vystihují vliv všech druhů
  - ***eigenvalue based methods***
- 2) rozmístění vzorků v ordinačním prostoru
  - najdi takové rozmístění vzorků v redukovaném ordinačním prostoru, aby vzdálenost mezi vzorky co nejvěrněji odrážela jejich nepodobnost vypočtenou z druhového složení jednotlivých vzorků
  - ***distance based methods***



# NEPŘÍMÁ VS. PŘÍMÁ ORDINACE

## *UNCONSTRAINED VS CONSTRAINED ORD.*

### **Nepřímá ordinace**

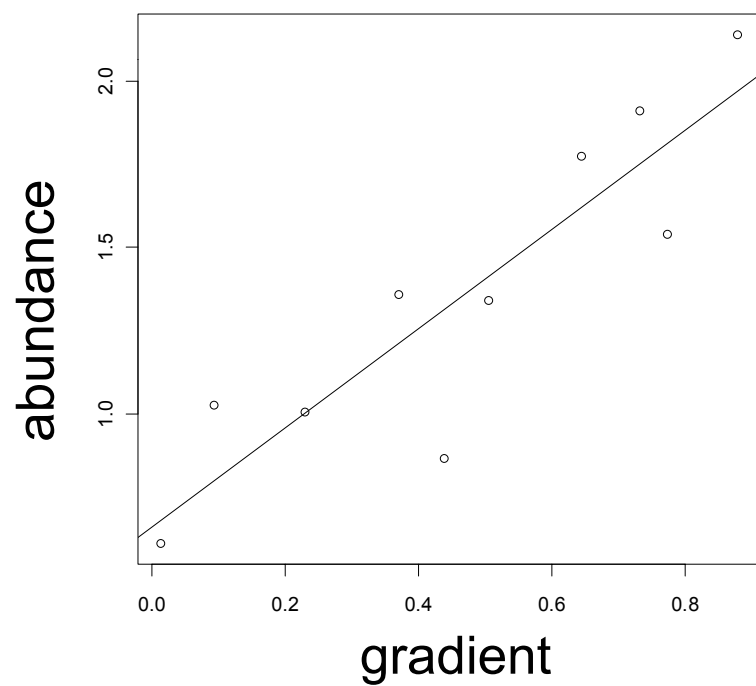
- vychází pouze z matice **vzorky × druhy**
- hledá proměnné (ordinační osy), které nejlépe reprezentují variabilitu v druhových datech
- slouží k popisu mnohorozměrných dat (*pattern description*) a generování hypotéz, ne k testování hypotéz

### **Přímá ordinace**

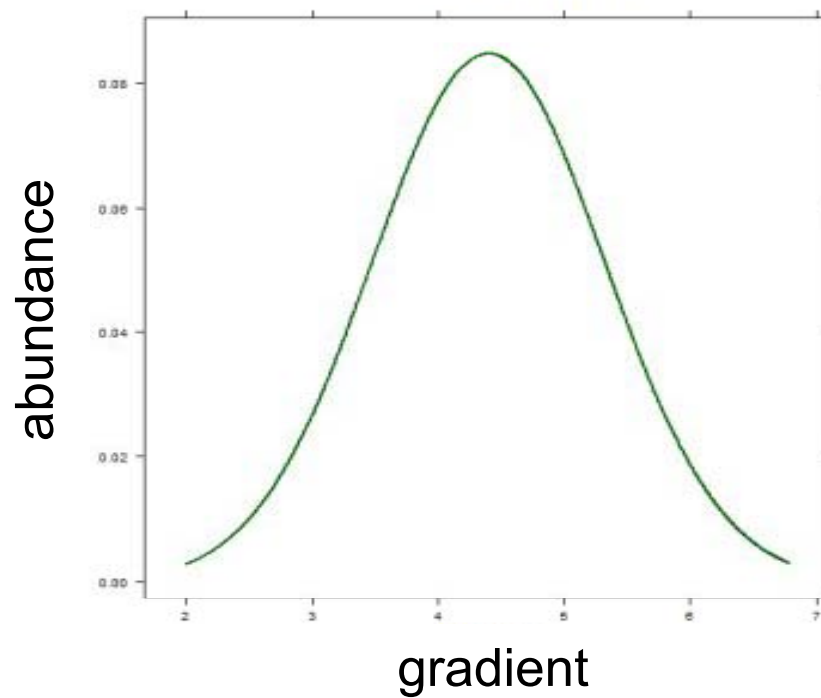
- vychází ze dvou matic: **vzorky × druhy** a **vzorky × proměnné prostředí**
- ordinační osy představují směr největší variability v druhových datech, která může být vysvětlena na základě *a priori* známých proměnných prostředí
- slouží spíše k testování hypotéz o vlivu proměnných prostředí na druhová data, neslouží k popisu dat

# MODELY ODPOVĚDI DRUHŮ NA GRADIENT PROSTŘEDÍ

lineární



unimodální



# ZÁKLADNÍ TYPY ORDINAČNÍCH TECHNIK

## (ZALOŽENÝCH NA VÝPOČTU EIGENVALUES)

	lineární odpověď druhů	unimodální odpověď druhů
nepřímá ordinace ( <i>unconstrained</i> )	<b>PCA</b> ( <i>Principal Component Analysis</i> , analýza hlavních komponent)	<b>CA</b> ( <i>Correspondence Analysis</i> , korespondenční analýza) <b>DCA</b> ( <i>Detrended Correspondence analysis</i> , detrendovaná korespondenční analýza)
přímá ordinace ( <i>constrained</i> )	<b>RDA</b> ( <i>Redundancy Analysis</i> , redundanční analýza)	<b>CCA</b> ( <i>Canonical Correspondence Analysis</i> , kanonická korespondenční analýza)

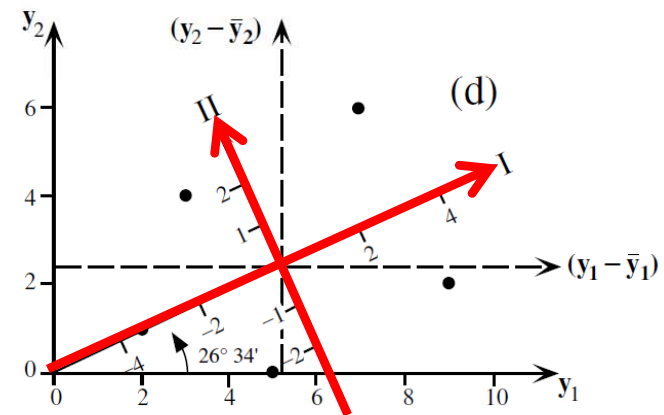
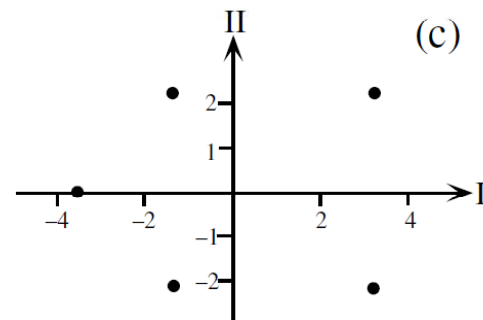
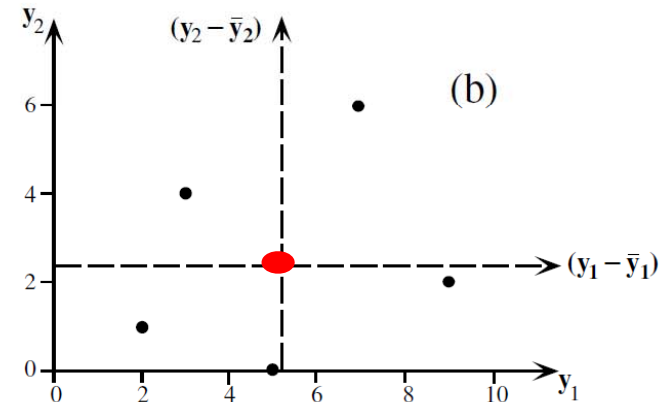
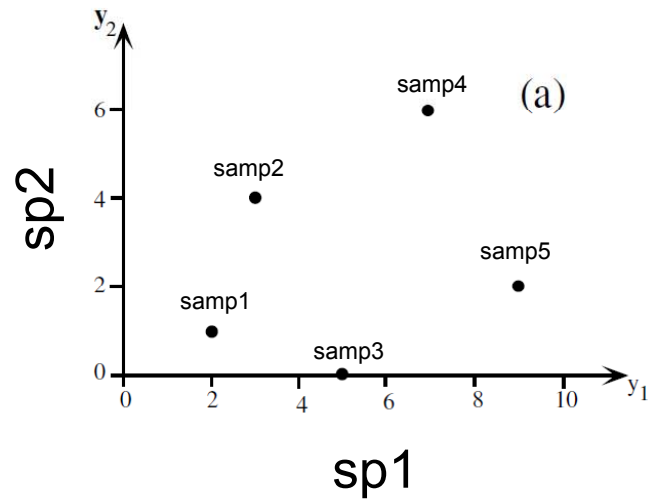
# NEPŘÍMÁ ORDINACE

## PRINCIP

- hledání skryté proměnné (gradientu), který nejlépe reprezentuje chování všech druhů podél tohoto gradientu
- první ordinační osa (*ordination axis*) a skóre vzorků na této ordinační ose (*sample scores*)
- odhad optima (odpovědi) jednotlivých druhů na této ose (*species scores*)
- druhá a vyšší ordinační osy – musejí být lineárně nezávislé na všech nižších ordinačních osách

# NEPŘÍMÁ ORDINACE PRINCIP (PCA)

	sp1	sp2
samp1	2	1
samp2	3	4
samp3	5	0
samp4	7	6
samp5	9	2



Legendre & Legendre (1998)

# NEPŘÍMÁ ORDINACE

## ALGORITMUS (CA)

### 5 výpočetních kroků

1. začni s arbitrárním (náhodným) skóre vzorků ( $x_i$ )
2. vypočti nové skóre pro jednotlivé druhy (*species score*,  $y_i$ ) **jako průměr skóre vzorků  $x_i$  vážený abundancí druhu ve vzorcích**
3. vypočti nové skóre pro jednotlivé vzorky (*sample score*,  $x_i$ ) **jako průměr skóre druhů  $y_i$  vážený abundancí druhů ve vzorku**
4. standardizuj skóre jednotlivých vzorků (natáhni osu)
5. pokud se skóre nemění, zastav, pokud ano, pokračuj krokem 2

# NEPŘÍMÁ ORDINACE

## CA – UNIMODÁLNÍ METODA

**Table 3-5.** Calculation of the first ordination axis by the weighted averaging (WA) method. Further explanation is in the text

	Samp1	Samp2	Samp3	SPWA1	SPWA2	SPWA3	SPWA4
Cirsium	0	0	3	10.000	10.000	10.000	10.000
Glechoma	5	2	1	2.250	1.355	1.312	1.310
Rubus	6	2	0	1.000	0.105	0.062	0.060
Urtica	8	1	0	0.444	0.047	0.028	0.027
<i>Initial value</i>	0	4	10				
SAWA1	1.095	1.389	8.063				
SAWA1resc.	0.000	0.422	10.000				
SAWA2	0.410	0.594	7.839				
SAWA2resc.	0.000	0.248	10.000				
SAWA3	0.376	0.555	7.828				
SAWA3resc.	0.000	0.240	10.000				
SAWA4	0.375	0.553	7.827				
SAWA4resc.	0.000	0.239	10.000				



# NEPŘÍMÁ ORDINACE

## CA – UNIMODÁLNÍ METODA

**Table 3-5.** Calculation of the first ordination axis by the weighted averaging (WA) method. Further explanation is in the text

	Samp1	Samp2	Samp3	SPWA1	SPWA2	SPWA3	SPWA4
Cirsium	0	0	3	10.000	10.000	10.000	10.000
Glechoma	5	2	1	2.250	1.355	1.312	1.310
Rubus	6	2	0	1.000	0.105	0.062	0.060
Urtica	8	1	0	0.444	0.047	0.028	0.027
<i>Initial value</i>	0	4	10				
SAWA1	1.095	1.389	8.063				
SAWA1resc.	0.000	0.422	10.000				
SAWA2	0.410	0.594	7.839				
SAWA2resc.	0.000	0.248	10.000				
SAWA3	0.376	0.555	7.828				
SAWA3resc.	0.000	0.240	10.000				
SAWA4	0.375	0.553	7.827				
SAWA4resc.	0.000	0.239	10.000				

# NEPŘÍMÁ ORDINACE

## CA – UNIMODÁLNÍ METODA

**Table 3-5.** Calculation of the first ordination axis by the weighted averaging (WA) method. Further explanation is in the text

	Samp1	Samp2	Samp3	SPWA1	SPWA2	SPWA3	SPWA4
Cirsium	0	0	3	10.000	10.000	10.000	10.000
Glechoma	5	2	1	2.250	1.355	1.312	1.310
Rubus	6	2	0	1.000	0.105	0.062	0.060
Urtica	8	1	0	0.444	0.047	0.028	0.027
<i>Initial value</i>	0	4	10				
SAWA1	1.095	1.389	8.063				
SAWA1resc.	0.000	0.422	10.000				
SAWA2	0.410	0.594	7.839				
SAWA2resc.	0.000	0.248	10.000				
SAWA3	0.376	0.555	7.828				
SAWA3resc.	0.000	0.240	10.000				
SAWA4	0.375	0.553	7.827				
SAWA4resc.	0.000	0.239	10.000				

# NEPŘÍMÁ ORDINACE

## CA – UNIMODÁLNÍ METODA

**Table 3-5.** Calculation of the first ordination axis by the weighted averaging (WA) method. Further explanation is in the text

	Samp1	Samp2	Samp3	SPWA1	SPWA2	SPWA3	SPWA4
Cirsium	0	0	3	10.000	10.000	10.000	10.000
Glechoma	5	2	1	2.250	1.355	1.312	1.310
Rubus	6	2	0	1.000	0.105	0.062	0.060
Urtica	8	1	0	0.444	0.047	0.028	0.027
<i>Initial value</i>	0	4	10				
SAWA1	1.095	1.389	8.063				
SAWA1resc.	0.000	0.422	10.000				
SAWA2	0.410	0.594	7.839				
SAWA2resc.	0.000	0.248	10.000				
SAWA3	0.376	0.555	7.828				
SAWA3resc.	0.000	0.240	10.000				
SAWA4	0.375	0.553	7.827				
SAWA4resc.	0.000	0.239	10.000				

# NEPŘÍMÁ ORDINACE

## CA – UNIMODÁLNÍ METODA

**Table 3-5.** Calculation of the first ordination axis by the weighted averaging (WA) method. Further explanation is in the text

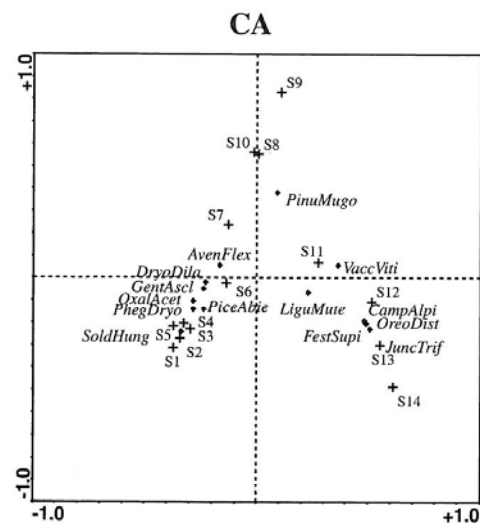
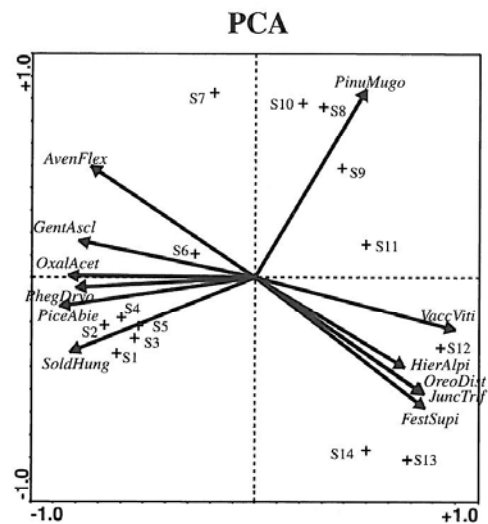
	Samp1	Samp2	Samp3	SPWA1	SPWA2	SPWA3	SPWA4
Cirsium	0	0	3	10.000	10.000	10.000	10.000
Glechoma	5	2	1	2.250	1.355	1.312	1.310
Rubus	6	2	0	1.000	0.105	0.062	0.060
Urtica	8	1	0	0.444	0.047	0.028	0.027
<i>Initial value</i>	0	4	10				
SAWA1	1.095	1.389	8.063				
SAWA1resc.	0.000	0.422	10.000				
SAWA2	0.410	0.594	7.839				
SAWA2resc.	0.000	0.248	10.000				
SAWA3	0.376	0.555	7.828				
SAWA3resc.	0.000	0.240	10.000				
SAWA4	0.375	0.553	7.827				
SAWA4resc.	0.000	0.239	10.000				

# ORDINAČNÍ DIAGRAM

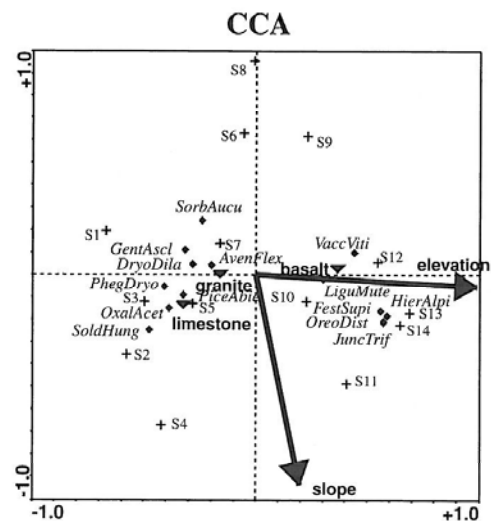
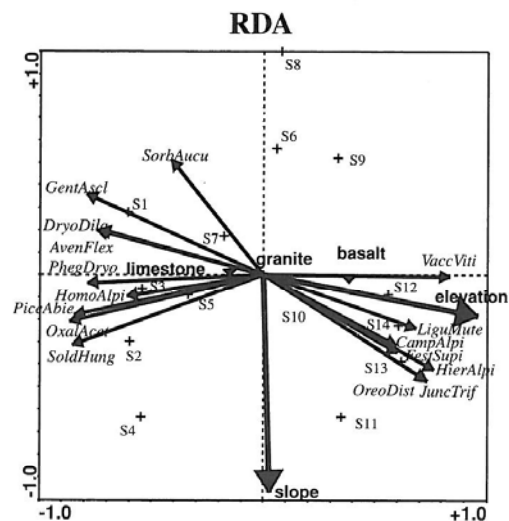
lineární metoda

unimodální metoda

nepřímá  
ordinace



přímá  
ordinace

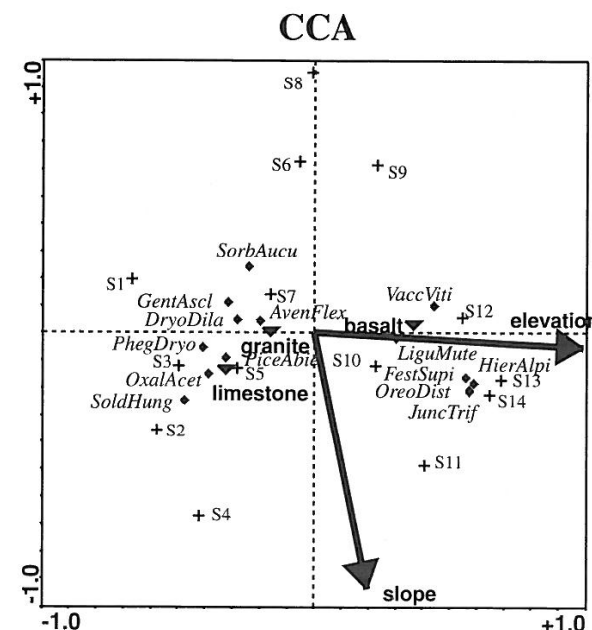
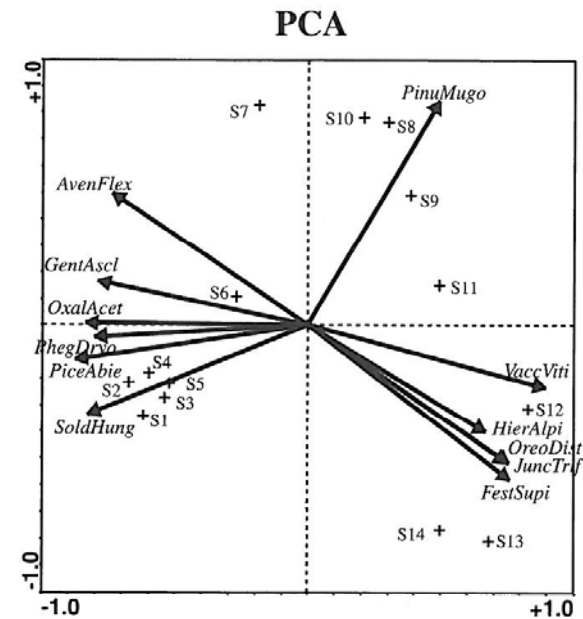


Lepš & Šmilauer (2003) Multivariate analysis of ...

# ORDINAČNÍ DIAGRAMY

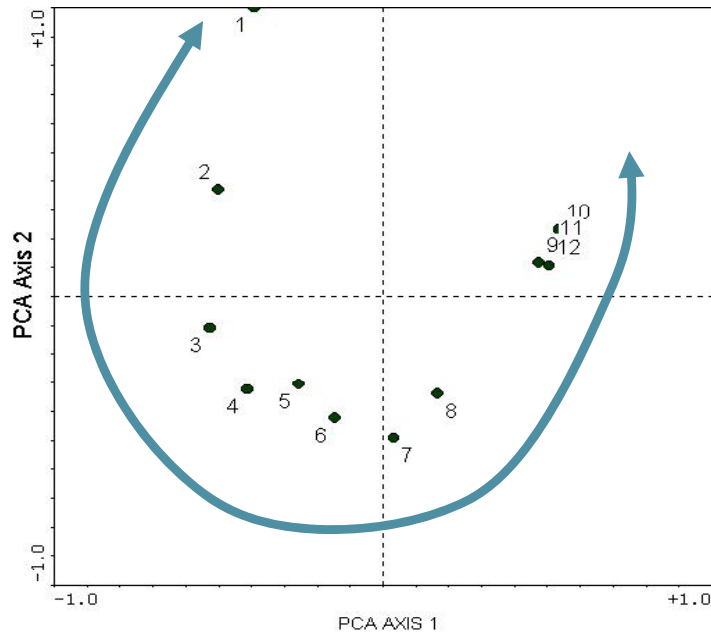
## KONVENCE

- zobrazení vzorků
  - > body
- zobrazení druhů
  - > šipky (lineární metody)
  - > body, centroidy (unimodální metody)
- zobrazení ordinačních os
  - vodorovná bývá osa vyššího řádu (např. první)
  - orientace os je arbitrární
- zobrazení proměnných prostředí
  - šipky (kvantitativní proměnné)
  - centroidy (kategoriální proměnné)
- typ ordinačního diagramu:
  - **scatterplot** - 1 typ dat (vzorky nebo druhy)
  - **biplot** - 2 typy dat (např. vzorky a druhy)
  - **triplot** - 3 typy dat (např. vzorky, druhy a proměnné prostředí)



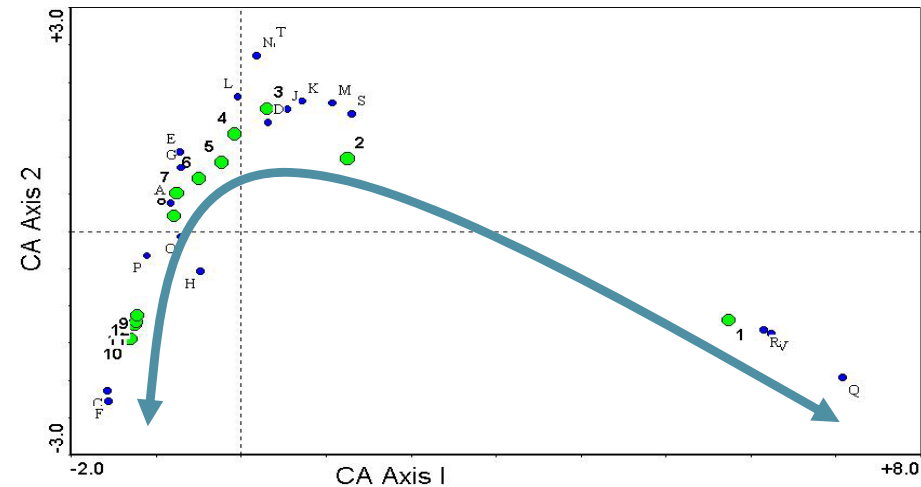
Lepš & Šmilauer (2003) Multivariate analysis of ...

# ARTEFAKTY V ORDINACÍCH



## Podkova (*Horseshoe effect*)

- PCA
- pořadí vzorků podél první osy neodráží jejich skutečnou nepodobnost
- v extrémním případě se mohou okraje přiblížit nebo dokonce překřížit



## Oblouk (*Arch effect*)

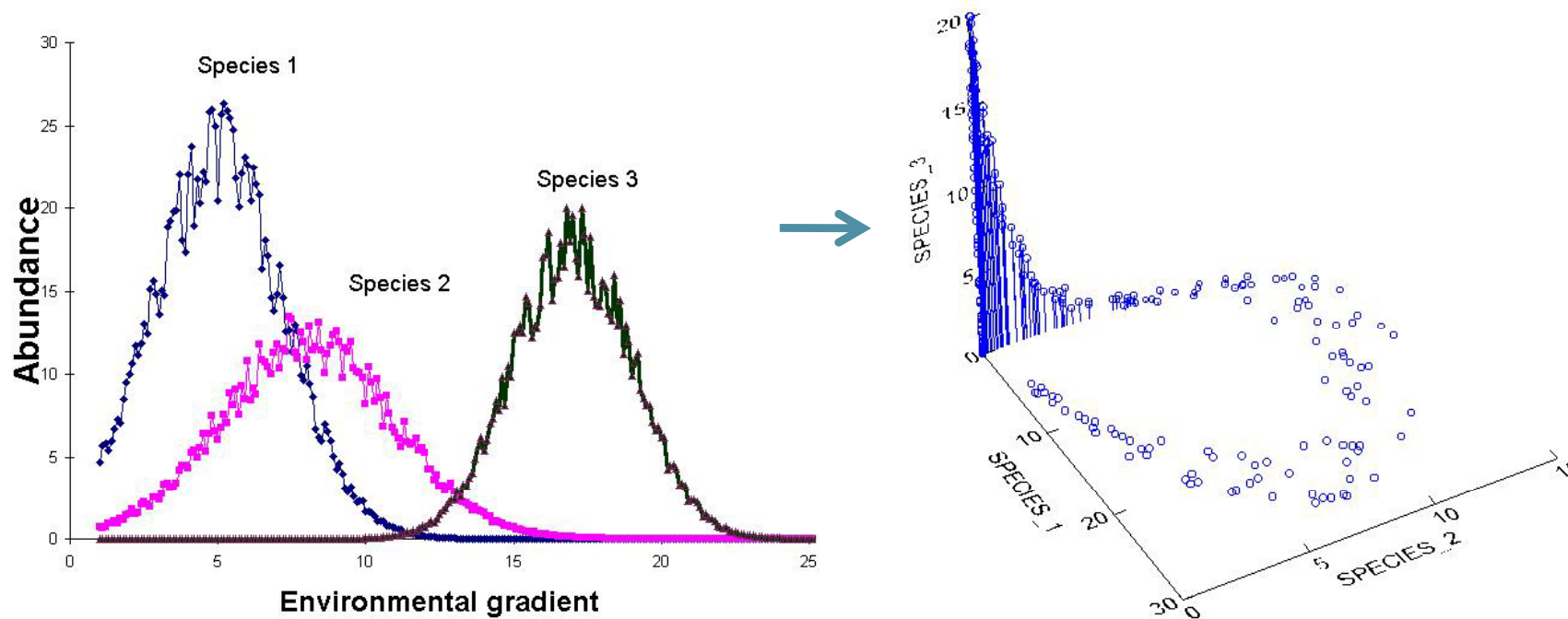
- CA
- pořadí vzorků podél první osy stále odráží jejich nepodobnost
- druhá osa je nelineární kombinací první osy

<http://ordination.okstate.edu>

# ARTEFAKTY V ORDINACÍCH

## Podkova a oblouk (*Horseshoe and arch effect*)

- důsledek algoritmu - každá následující osa musí být *lineárně* nezávislá na předchozích osách, ale neuvažuje se *nelineární* závislost
- důsledek projekce - nelineární vztahy mezi druhy a gradienty prostředí se promítají do lineárního prostoru definovaného Euklidovskými vzdálenostmi



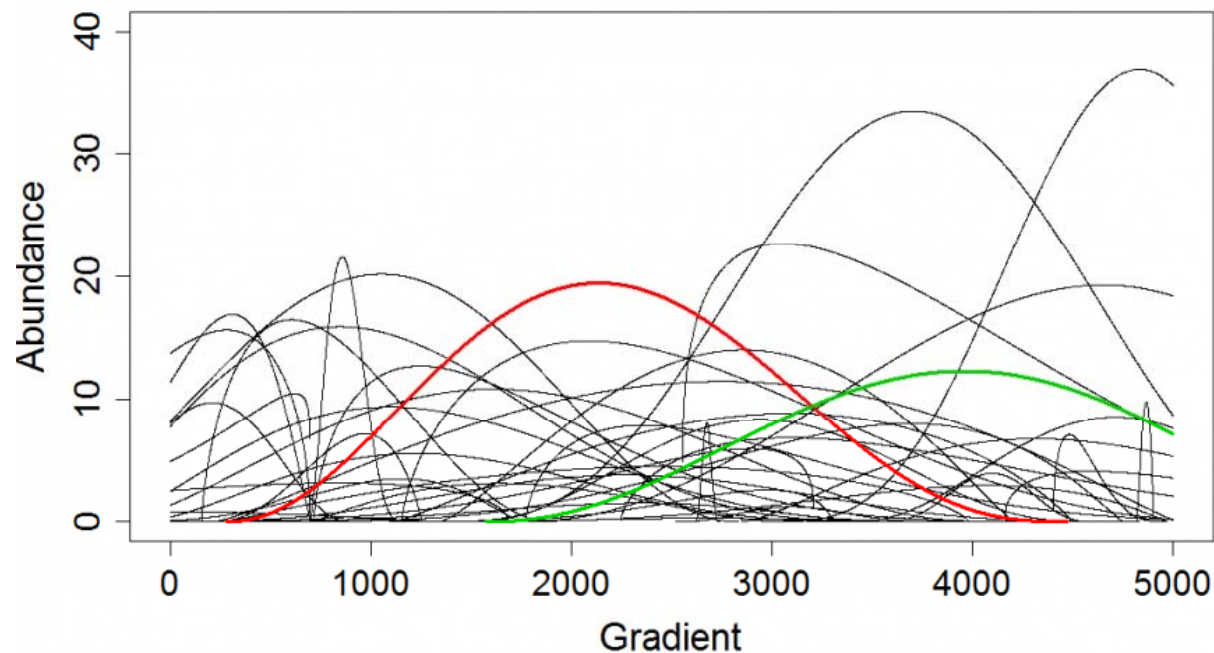
<http://ordination.okstate.edu>



# SIMULOVANÁ DATA

## POUZE JEDEN EKOLOGICKÝ GRADIENT

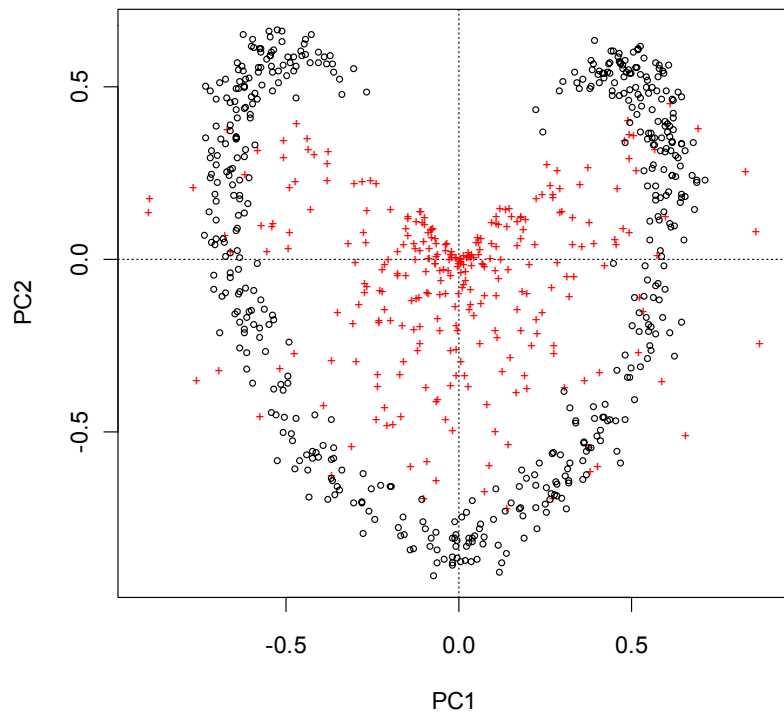
- simulovaný gradient dlouhý 5000 jednotek
- 300 druhů s unimodální odpovědí, různými šířkami nik
- 500 vzorků náhodně rozmístěných podél gradientu



# SIMULOVANÁ DATA

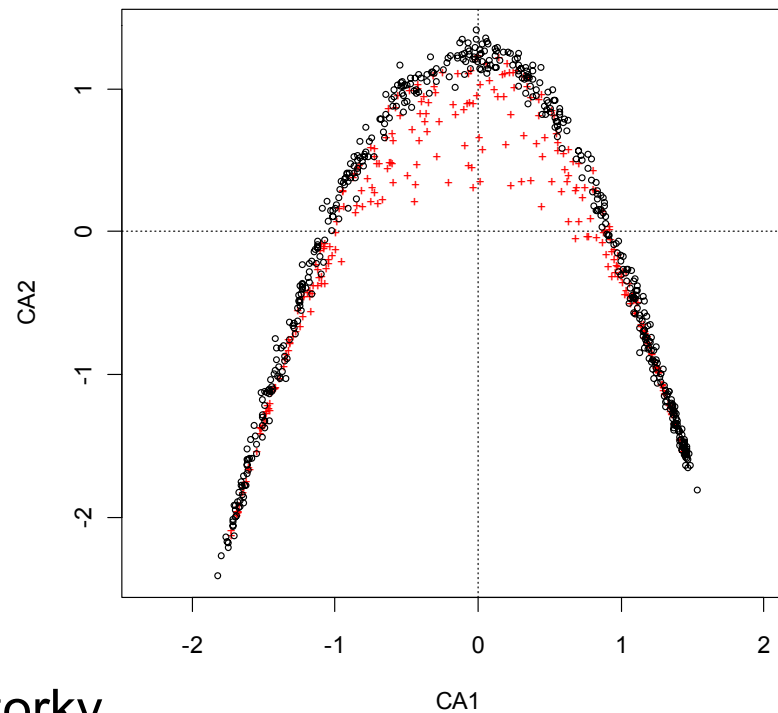
## ARTEFAKTY

PCA - podkova



o vzorky  
+ druhy

CA - oblouk



# ARTEFAKTY V ORDINACÍCH

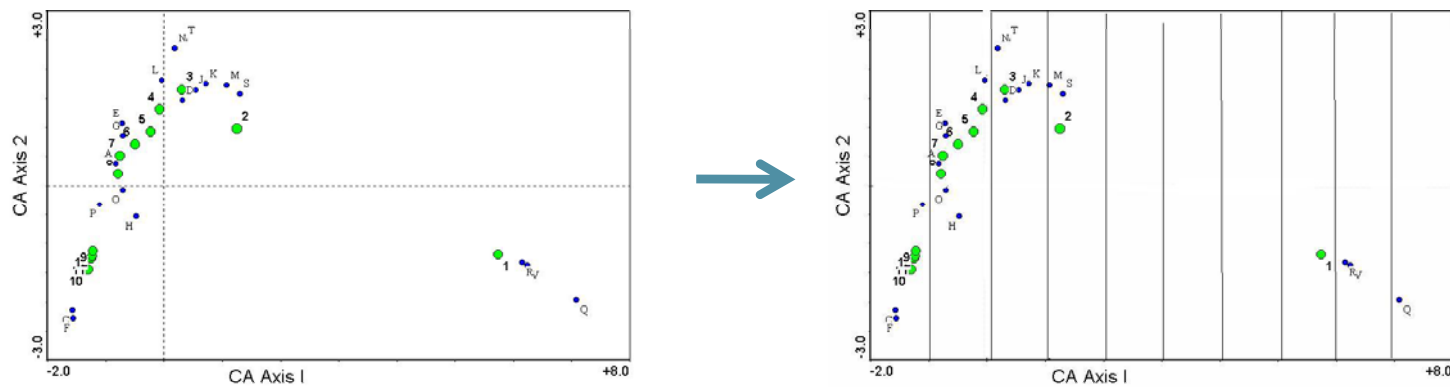
## MOŽNOSTI ŘEŠENÍ

- odstranění trendu z ordinačních os (***detrending***)
  - detrendovaná korespondenční analýza, *Detrended Correspondence Analysis* (DCA, Hill & Gauch 1980)
  - *detrending by segments* (nejčastější)
  - *detrending by polynomials* (pokud v analýze používám kovariáty)
- použití takových ordinačních technik, které umožňují ordinaci vzorků v prostoru pomocí jiných metrik než je Euklidovská distance (PCA) nebo chi-kvadrát distance (CA)
  - analýza hlavních koordinát, *Principal Coordinate Analysis* (PCoA)
  - nemetrické mnohorozměrné škálování, *Non-metric Multidimensional Scaling* (NMDS)

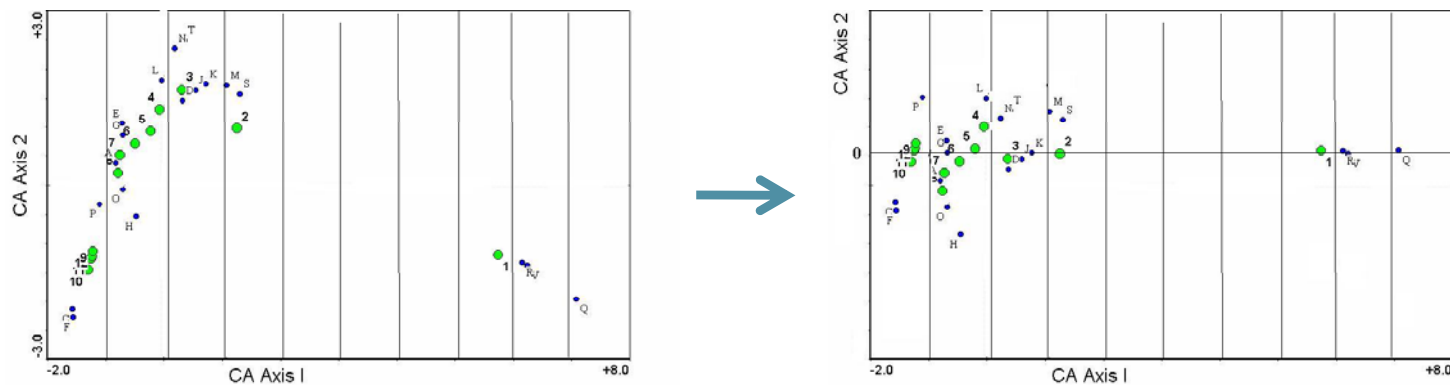
# DETRENDED CORRESPONDENCE ANALYSIS

## PROCES ODSTRANĚNÍ TRENDU

Krok 1 – rozdělení první osy na několik segmentů



Krok 2 – vycentrování druhé osy každého segmentu kolem nuly



<http://ordination.okstate.edu>

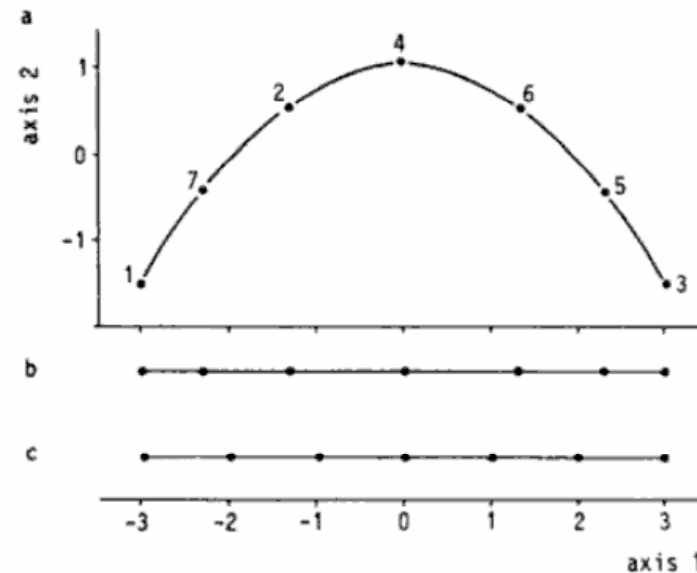
# DETRENDED CORRESPONDENCE ANALYSIS

## PROCES ODSTRANĚNÍ TRENDU

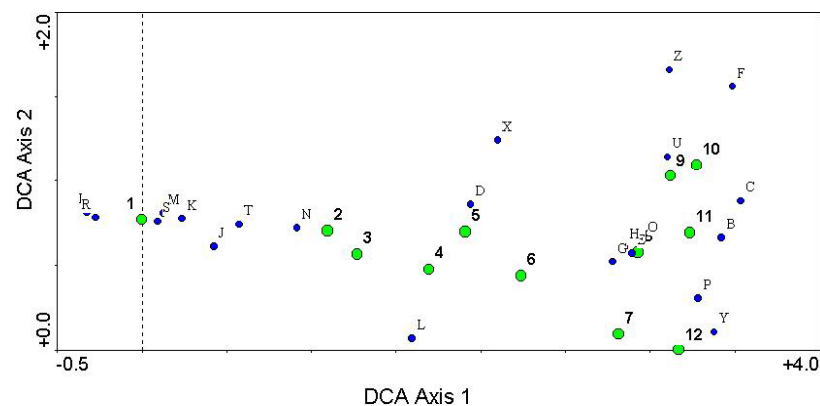
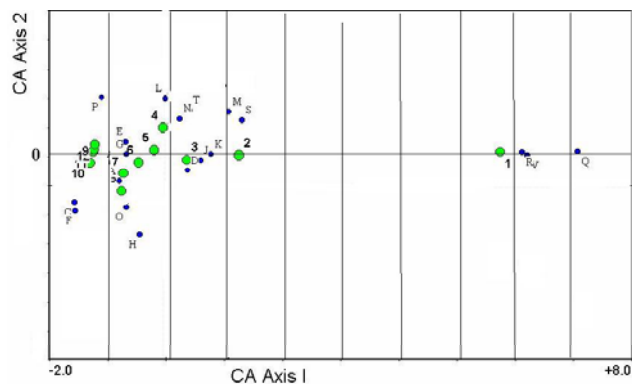
Krok 3 – nelineární přeškálování první osy, které odstraňuje nahloučení vzorků na koncích gradientů

-> výsledný ordinační diagram má osy naškáované v jednotkách **směrodatné odchylky (SD)**

-> platí, že druhové složení se změní o polovinu na gradientu o délce 1-1,4 SD (*half-change in species composition*)



ter Braak (1987)



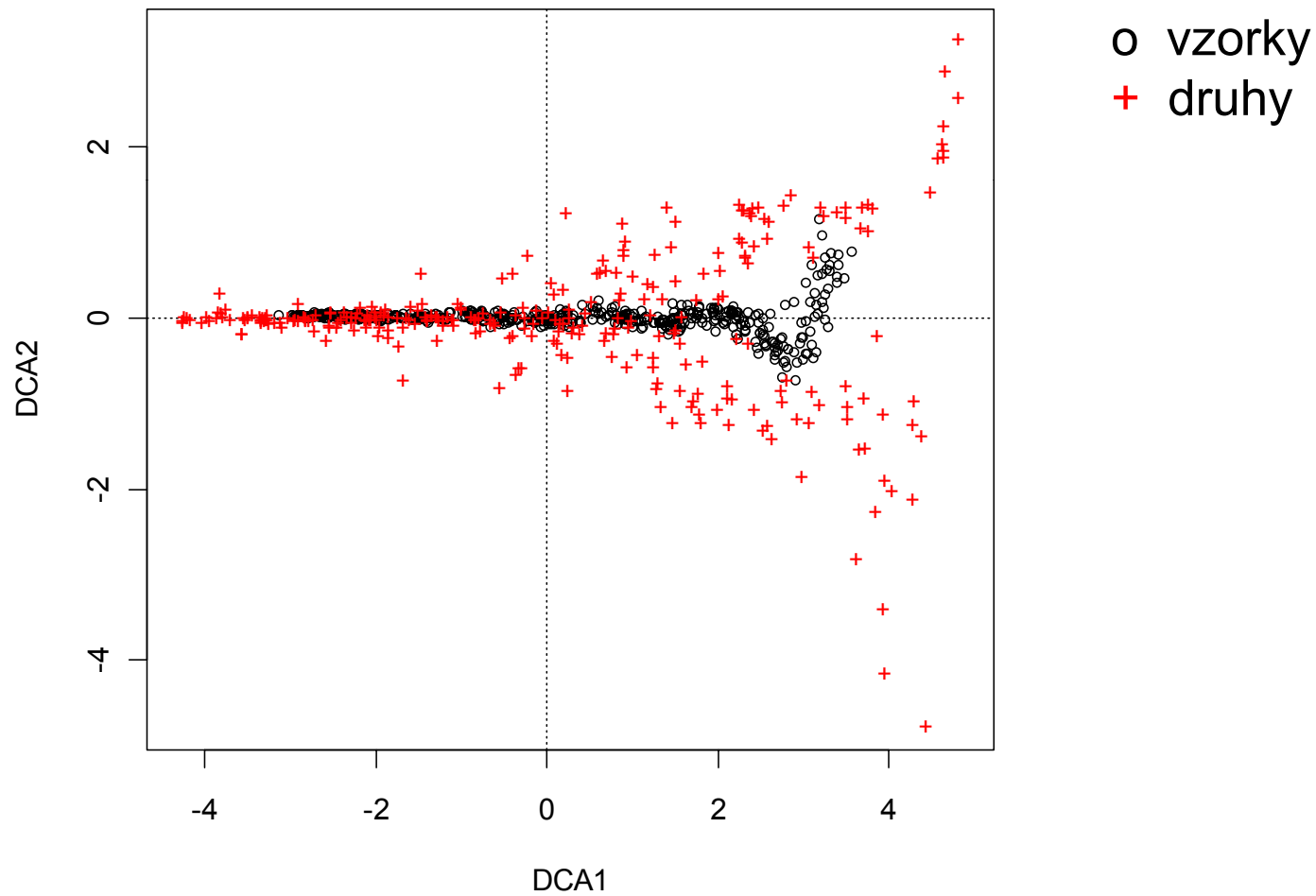
# DETRENDED CORRESPONDENCE ANALYSIS

## VÝHODY A NEVÝHODY

- ☹️ neelegantní metoda, která je někdy přirovnávána k použití kladiva na data (hlavně část týkající se rozdělení osy na segmenty a jejich centrování)
- ☹️ výsledek je silně ovlivněn arbitrárním rozhodnutím o počtu segmentů (doporučuje se vyzkoušet více možností)
- ☹️ pokud jsou v datech dva nebo více hlavních gradientů (ordinačních os), DCA si s nimi neporadí (*detrending* do určité míry poškodí druhou a vyšší ordinační osy)
- 😊 i kladivo, pokud je v rukou odborníka, může být použito efektivně - metoda často dává ekologicky dobře interpretovatelné výsledky
- 😊 osy DCA jsou v jednotkách SD, které umožňují zjistit, jak dlouhý gradient naše data pokrývají

# SIMULOVANÁ DATA (JEDEN EKOLOGICKÝ GRADIENT)

## DCA



# VÝBĚR ORDINAČNÍ METODY NA ZÁKLADĚ DCA

## LINEÁRNÍ NEBO UNIMODÁLNÍ?

- lineární metody vyžadují homogenní data, unimodální jsou vhodná i pro data heterogenní
- kuchařka *alá* Lepš & Šmilauer (2003) - zjištění délky gradientu (heterogeneity dat) pomocí metody DCA, *detrending by segments*
- pokud je délka 1. osy DCA
  - menší než 3 SD** – použijí lineární techniku
  - větší než 4 SD** – použijí unimodální techniku
  - v rozmezí 3-4 SD** – obě techniky pracují rozumně
- alternativní doporučení (Jari Oksanen) – pokud se nemůžu rozhodnout mezi lineární a unimodální metodou, použijí unimodální – ta se chová rozumně i na lineárních datech



# PCoA (*PRINCIPAL COORDINATE ANALYSIS*)

## ORDINACE ZALOŽENÁ NA DISTANCÍCH

- syn. MDS – *Metric Dimensional Scaling*
- alternativní metoda nepřímé ordinace
- vstupní data – matice nepodobností mezi vzorky
- výpočet matice nepodobností – jakýkoliv index nepodobnosti
  - pokud zvolím Euklidovskou vzdálenost -> identické s PCA
  - pokud zvolím Chi-kvadrát vzdálenost -> obdoba CA
- v CANOCO se počítá programem PrCoord



# NMDS (*NON-METRIC MULTIDIMENSIONAL SCALING*)

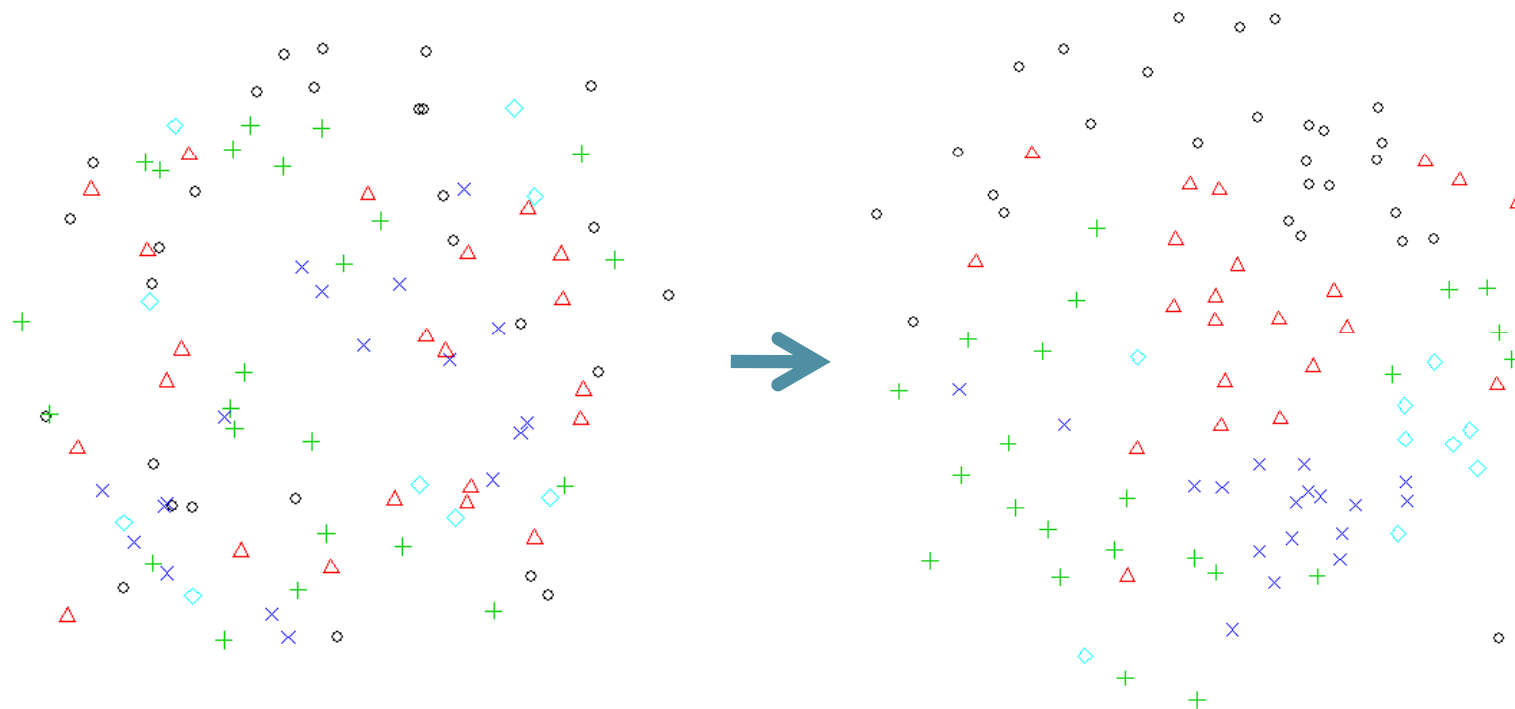
## ORDINACE ZALOŽENÁ NA DISTANCÍCH

- další alternativa nepřímých ordinací, nemetrická varianta PCoA
- vstupní data – matice nepodobností mezi vzorky
- výpočet matice nepodobností – jakýkoliv index nepodobnosti
- výsledek je značně ovlivněn výběrem indexu nepodobnosti
- iterativní algoritmus, který nemusí pokaždé dojít ke stejnému výsledku (lokální optima)
- na začátku je nutno určit počet dimenzí, se kterými bude metoda pracovat (obvykle  $k = 2$  nebo  $3$ )
- při větším množství dat VELMI časově náročná
- v CANOCO se počítá programem WinKyst, který je ke stažení zde: <http://www.canodraw.com/winkyst.htm>



# NMDS

## *NON-METRIC MULTIDIMENSIONAL SCALING*



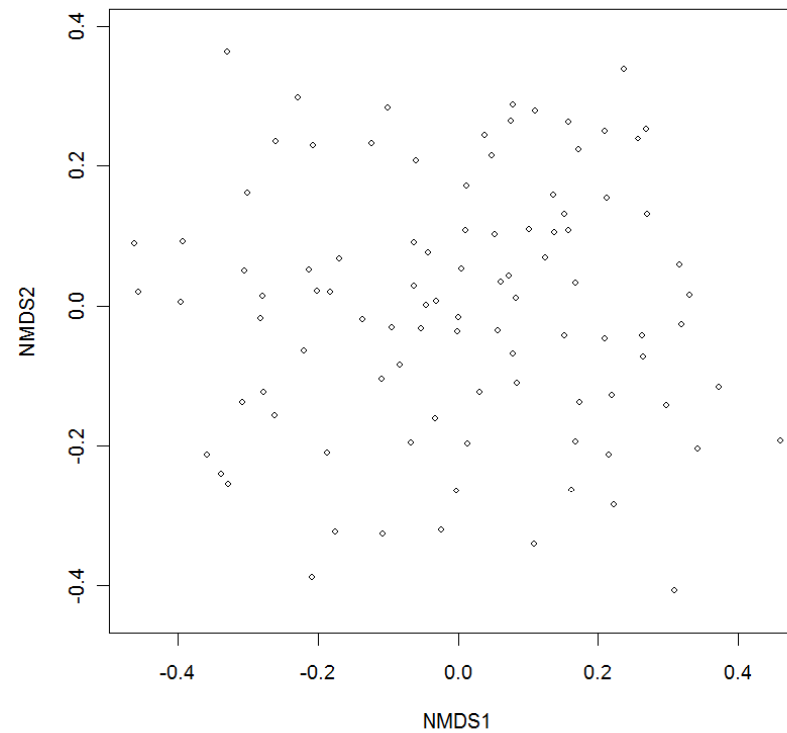
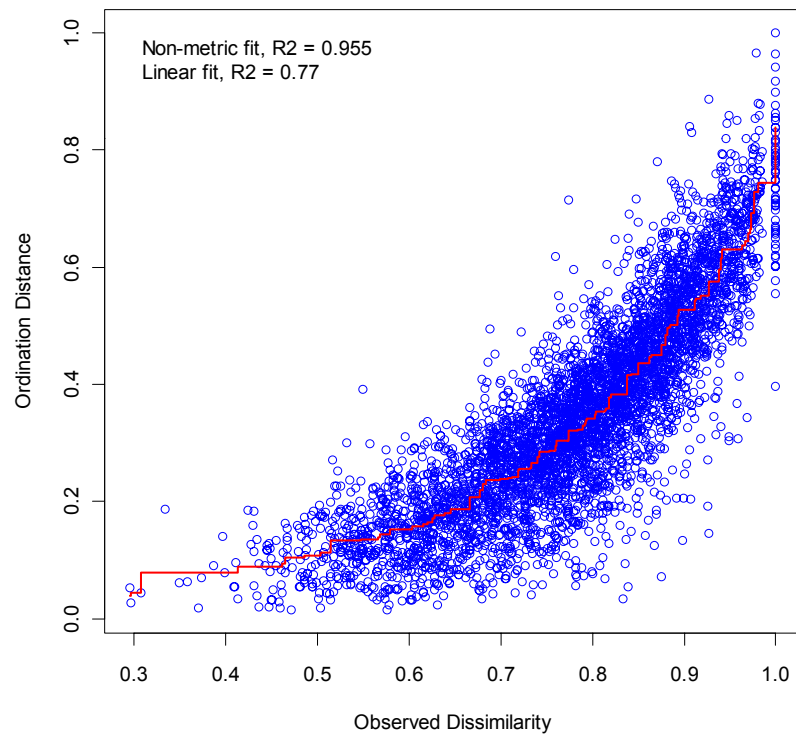
náhodné rozmístění vzorků v  
prostoru

rozmístění vzorků v prostoru  
respektuje jejich nepodobnost

# NMDS

## NON-METRIC MULTIDIMENSIONAL SCALING

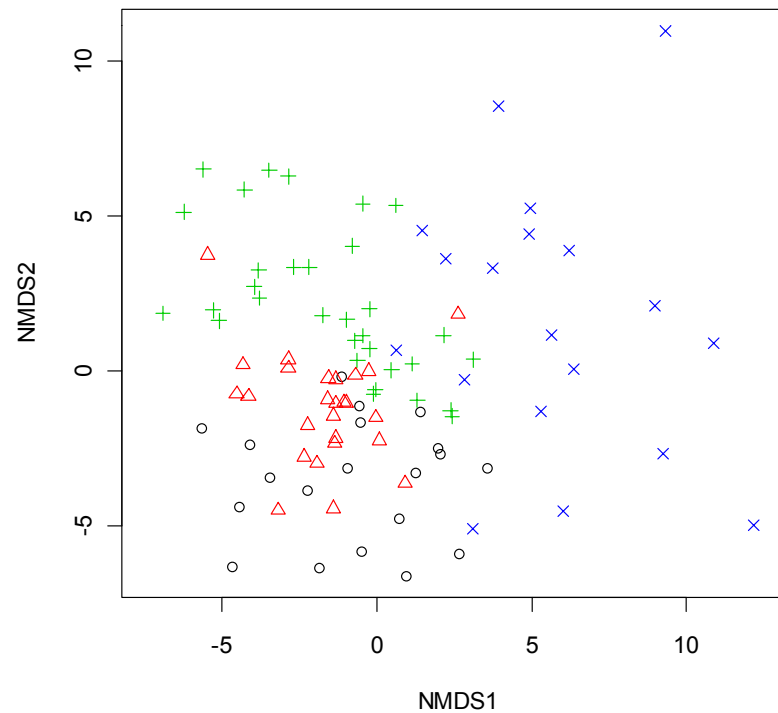
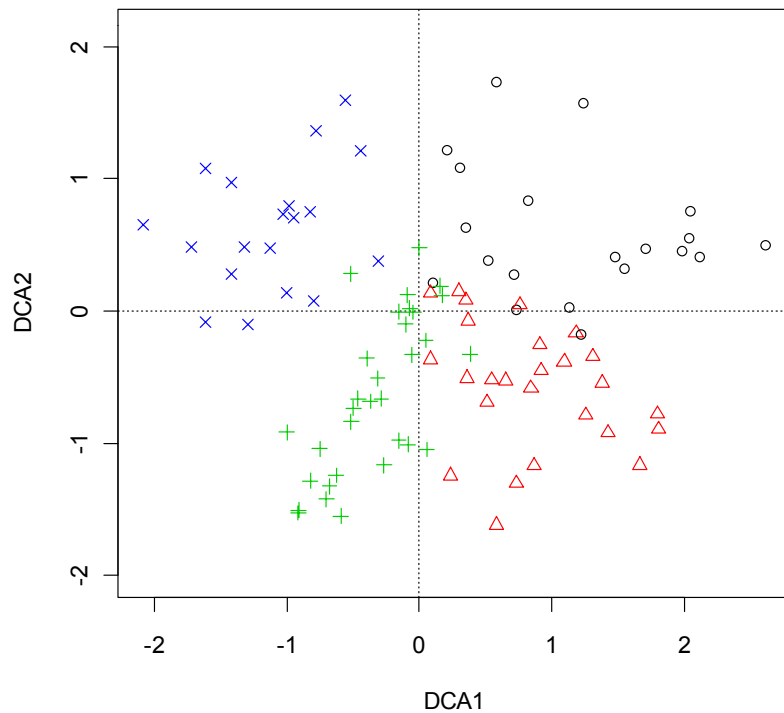
stress = 7.47



# POROVNÁNÍ METOD DCA A NMDS

DCA

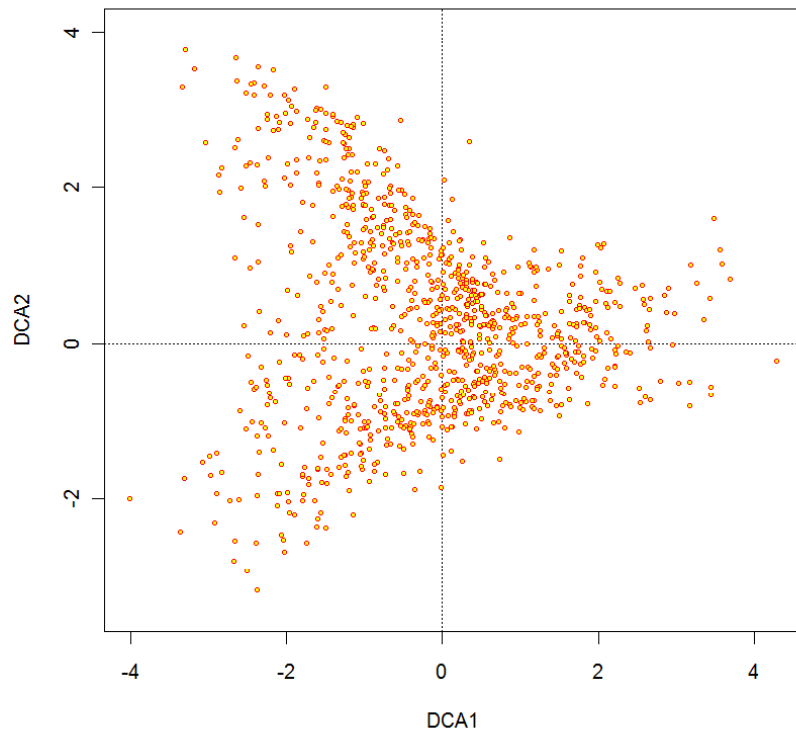
NMDS



data z údolí Vltavy, klasifikace metodou **TWINSPAN** (Zelený & Chytrý 2007)

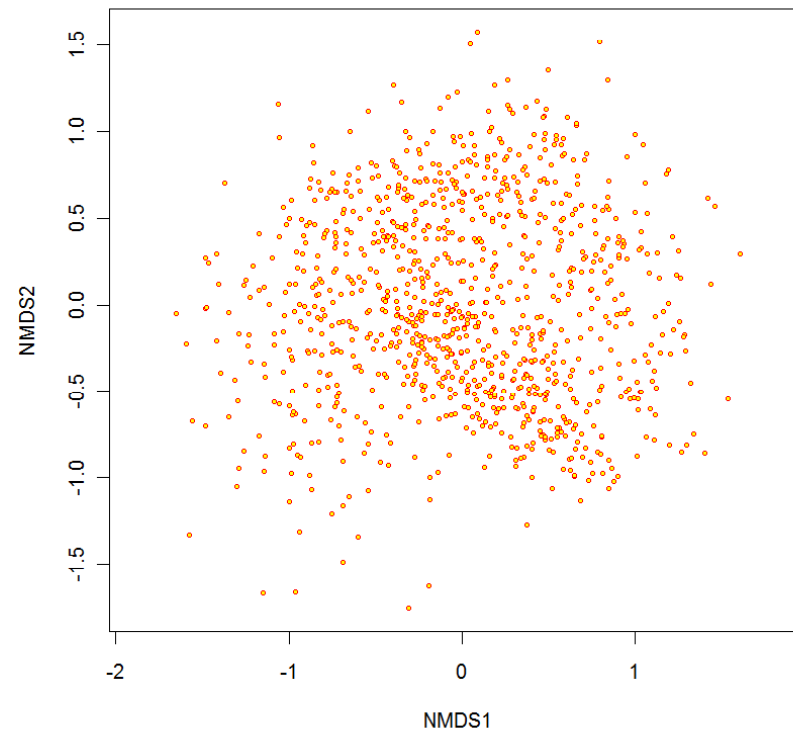
# POROVNÁNÍ METOD DCA A NMDS

DCA



při větším počtu vzorků tvoří trojúhelník  
nebo pěticípou hvězdu (artefakt)

NMDS



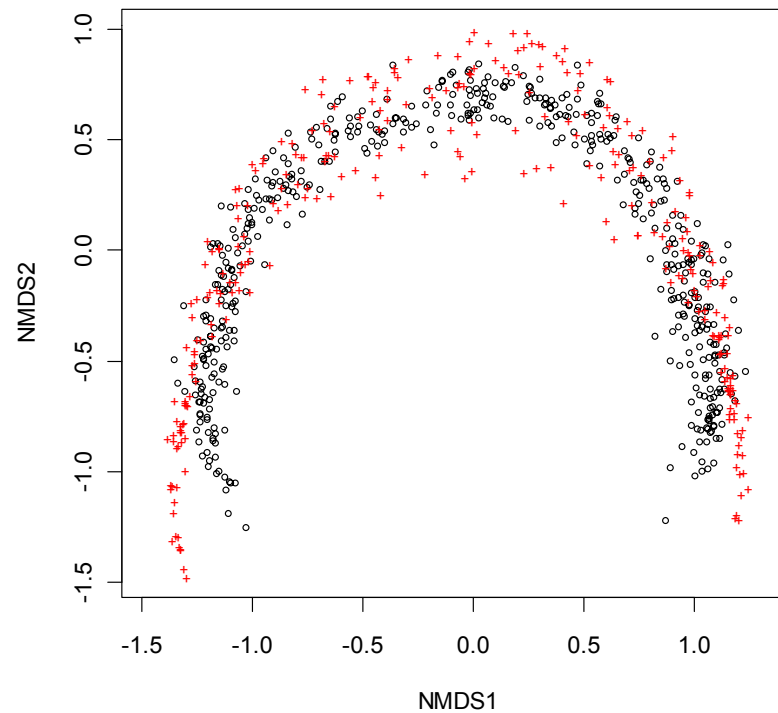
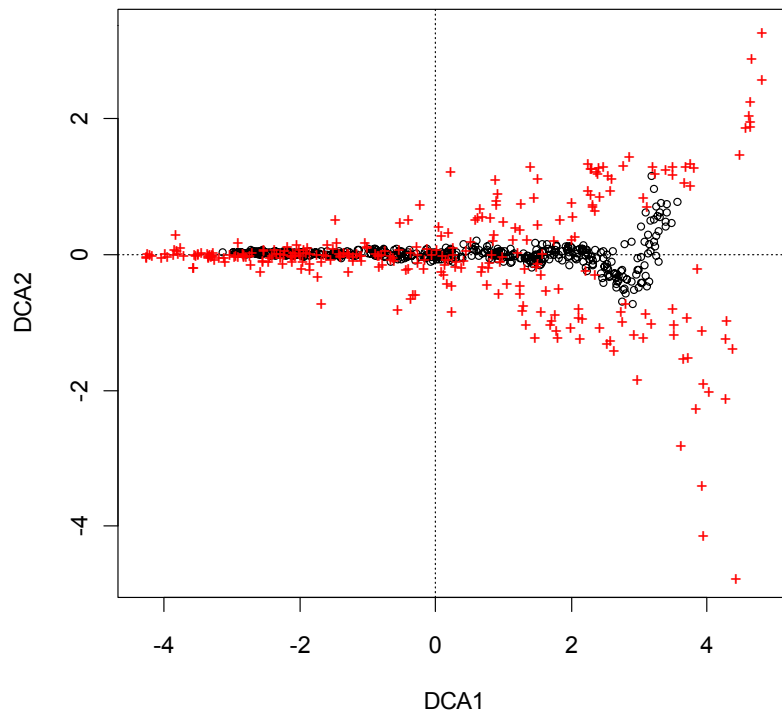
má tendenci jakákoliv data zobrazit  
jako kouli

# POROVNÁNÍ METOD DCA A NMDS

## SIMULOVANÁ DATA (JEDEN GRADIENT)

DCA

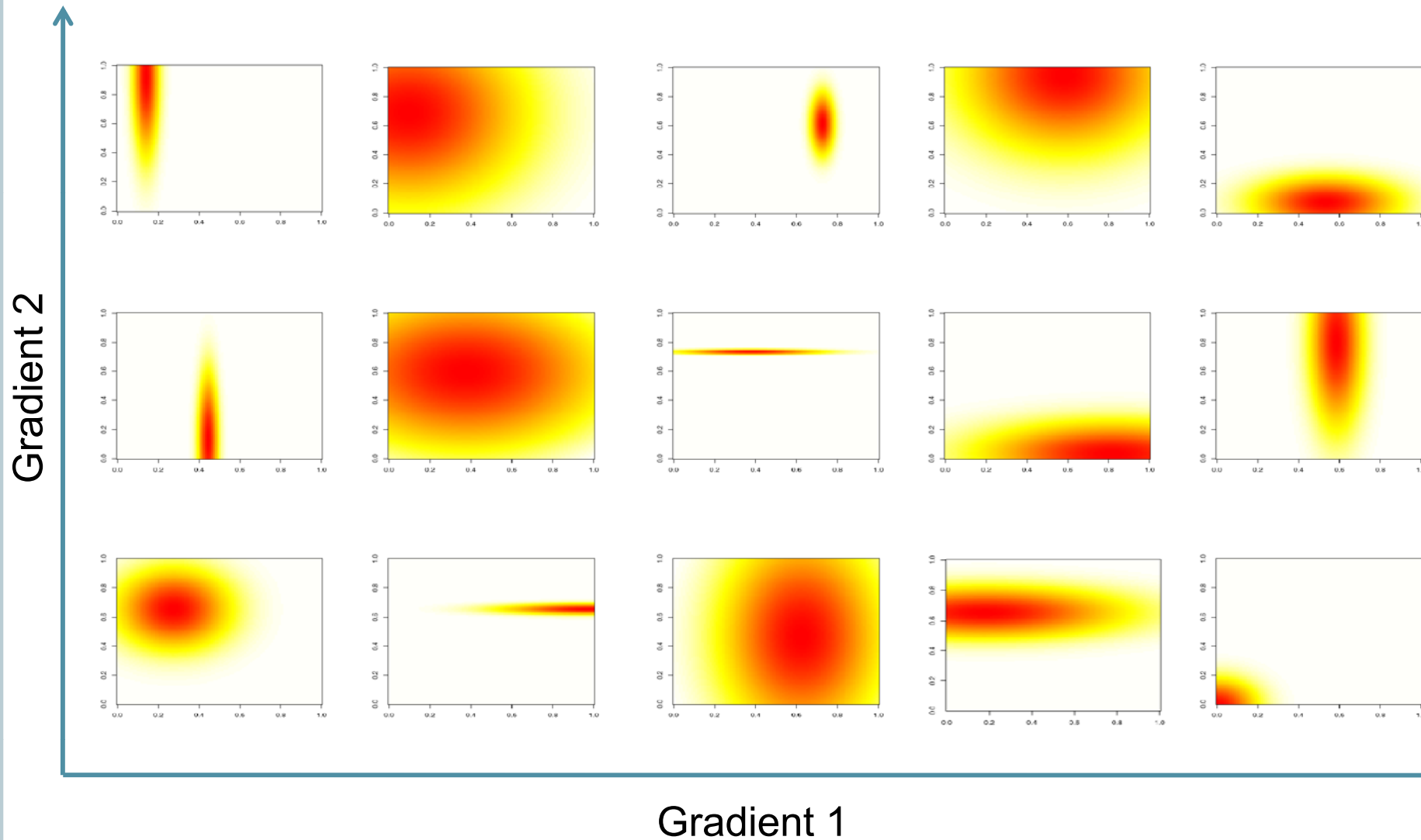
NMDS



o vzorky  
+ druhy

# SIMULOVANÁ DATA

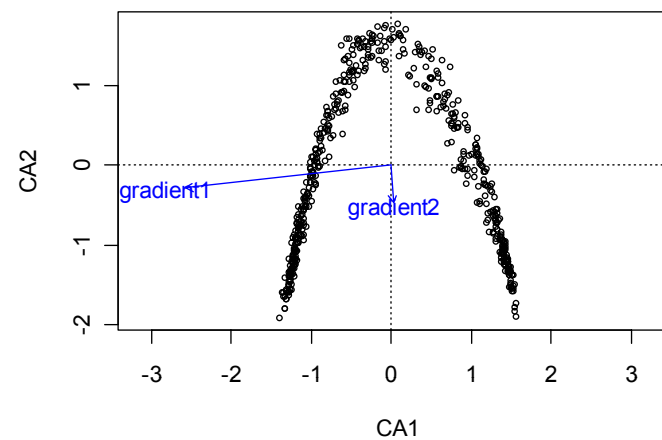
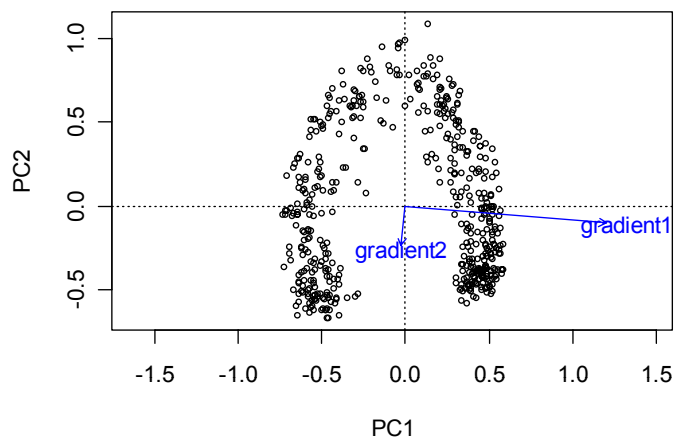
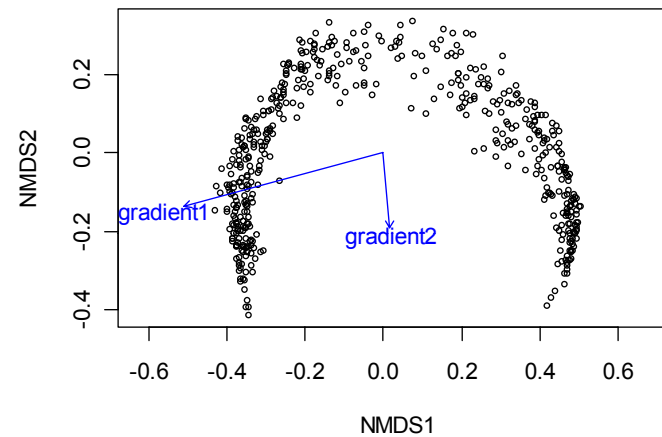
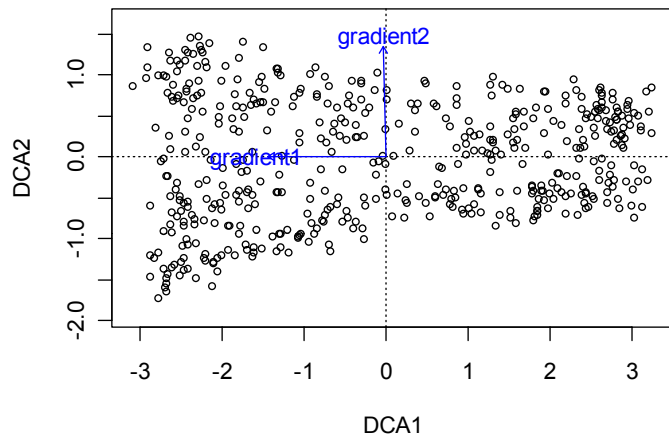
## DVA RŮZNĚ DLOUHÉ GRADIENTY





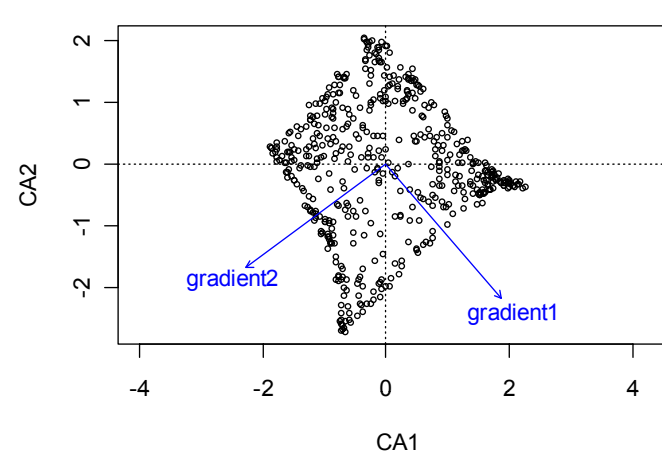
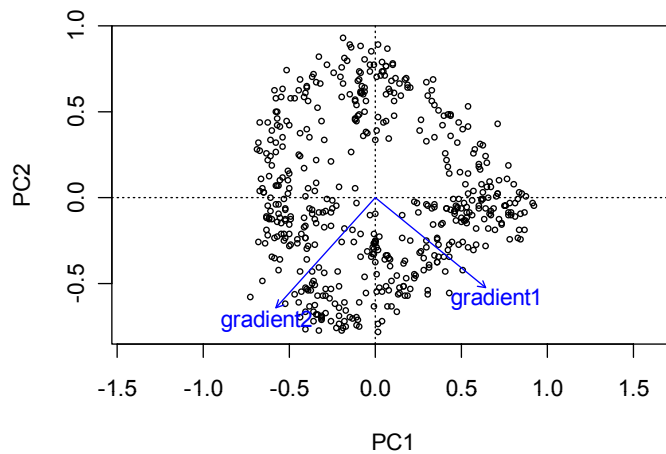
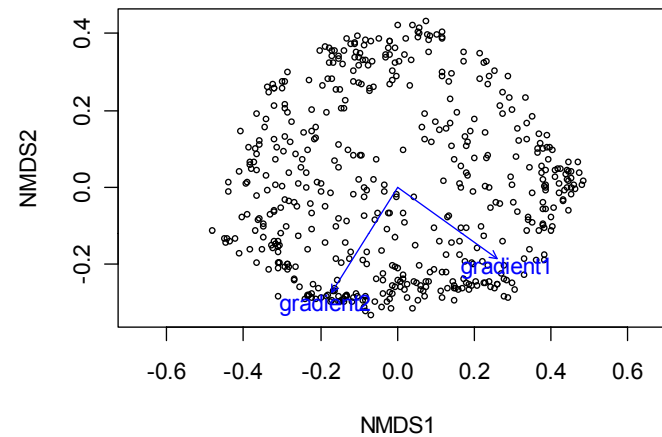
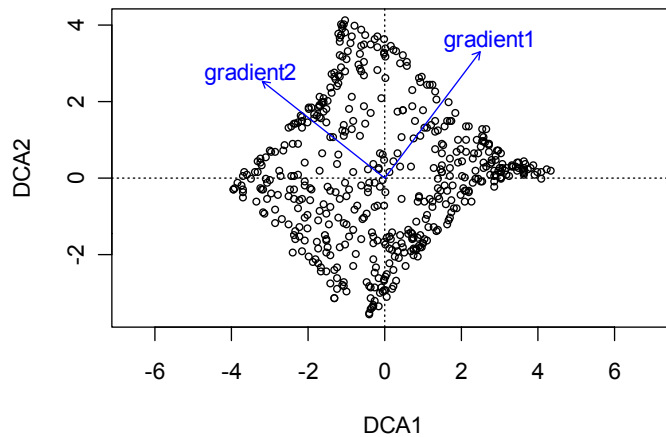
# SIMULOVANÁ DATA

## DVA RŮZNĚ DLOUHÉ GRADIENTY



# SIMULOVANÁ DATA

## DVA STEJNĚ DLOUHÉ GRADIENTY

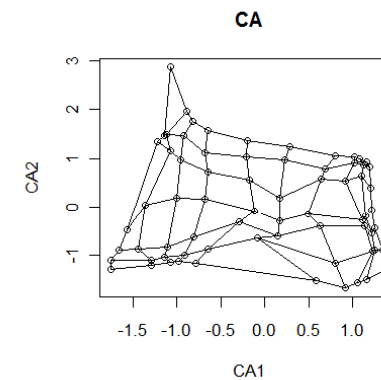
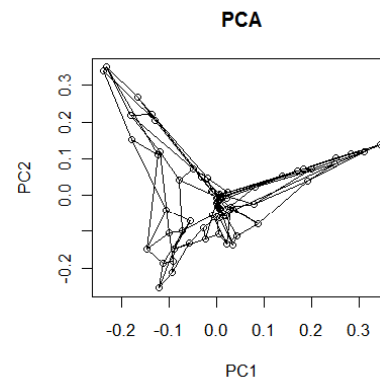
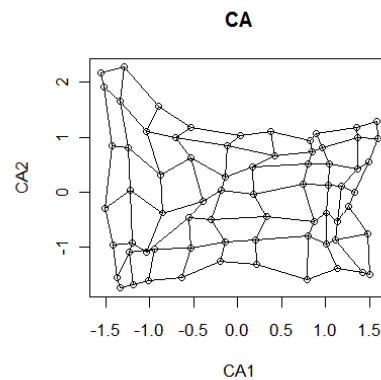
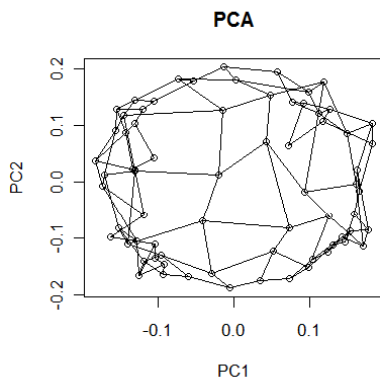
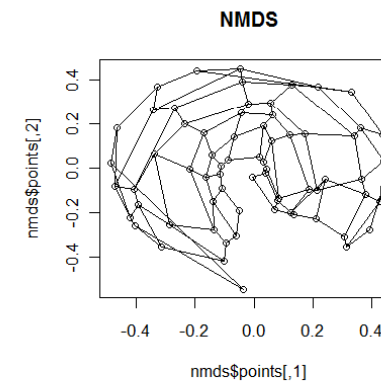
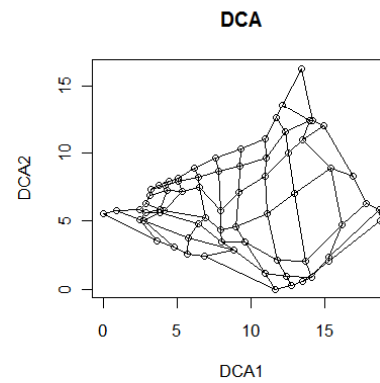
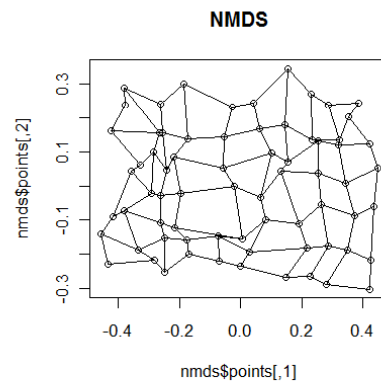
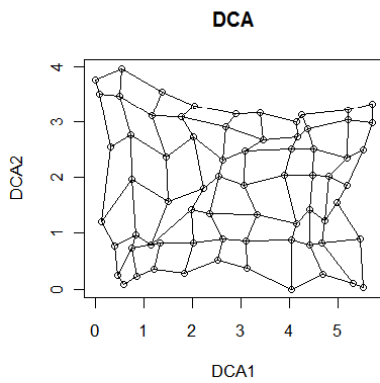


# SIMULOVANÁ DATA

## DVA RŮZNĚ DLOUHÉ GRADIENTY

krátké gradienty

dlouhé gradienty



# POROVNÁNÍ METOD ZALOŽENÝCH NA VÝPOČTU EIGENVALUES A DISTANCÍ

## ***Eigenvalue-based ordination methods***

- DCA, PCA a CA a jejich omezené (constrained) varianty DCCA, RDA a CCA
- vstupní data = matice vzorky x druhy, ze kterých se extrahují hlavní ordinační osy (*eigenvectors*)
- interpretace zaměřena na směry variability v datech, vysvětlené jednotlivými ordinačními osami

## ***Distance-based ordination methods***

- NMDS a PCoA
- vstupní data = matice nepodobností
- interpretace zaměřena se na vzdálenosti mezi vzorky v redukovaném ordinačním prostoru

# POUŽITÍ PROMĚNNÝCH PROSTŘEDÍ V ORDINACI

## DVA ALTERNATIVNÍ POSTUPY

### 1. nepřímá ordinace + korelace

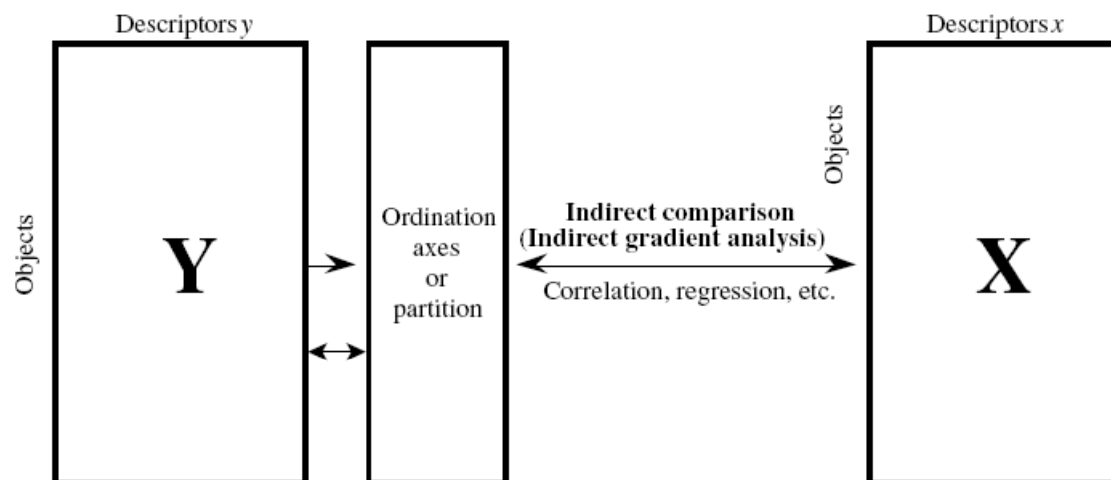
- získám skóre vzorků na hlavních ordinačních osách
- skóre vzorků koreluje s jednotlivými proměnnými prostředí
- + určitě zachytím hlavní gradienty v druhovém složení
- nemusím zachytit tu část variability v druhovém složení, která je vztažená k jednotlivým proměnným prostředí

### 2. přímá ordinace

- proměnné prostředí vstupují přímo jako vysvětlující proměnné do ordinace
- skóre vzorků na osách je ovlivněno vztahem k těmto proměnným prostředí
- + určitě zachytím variabilitu v datech, která se vztahuje k jednotlivým proměnným prostředí
- nemusím zachytit část variability v druhových datech, která není vysvětlena žádnou proměnnou prostředí

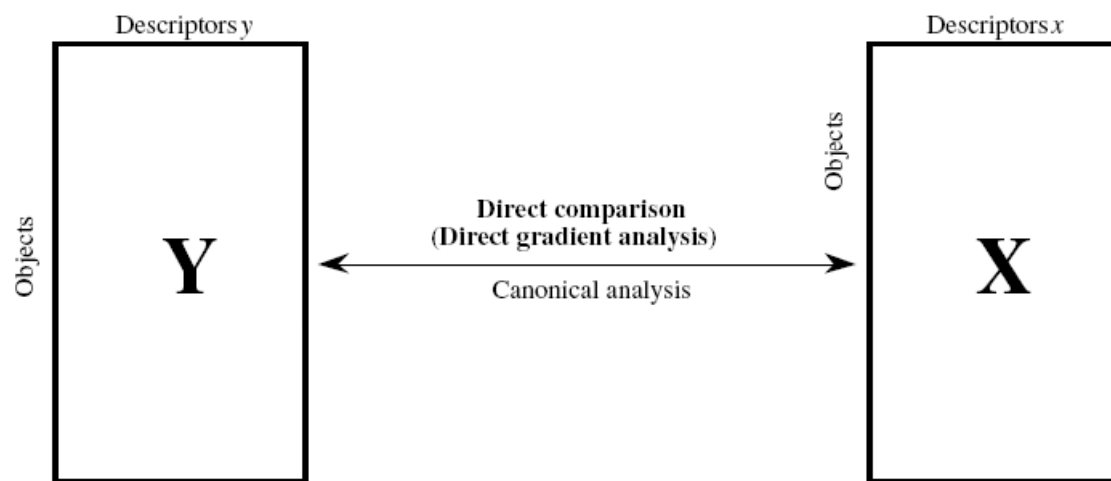
# POUŽITÍ PROMĚNNÝCH PROSTŘEDÍ V ORDINACI

## DVA ALTERNATIVNÍ POSTUPY



**Y** – samples × species matrix

**X** – samples × environmental factors matrix



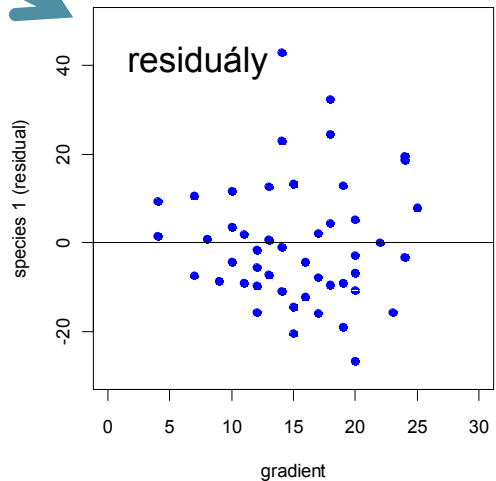
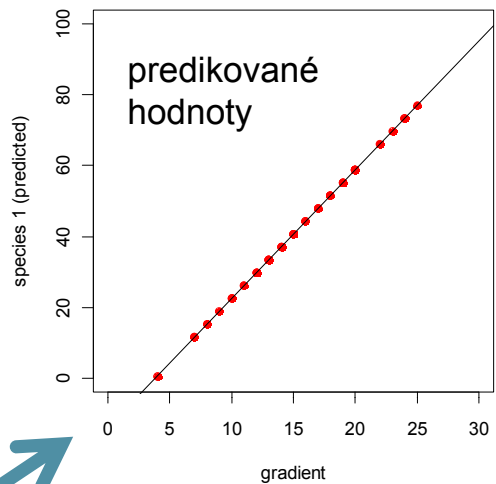
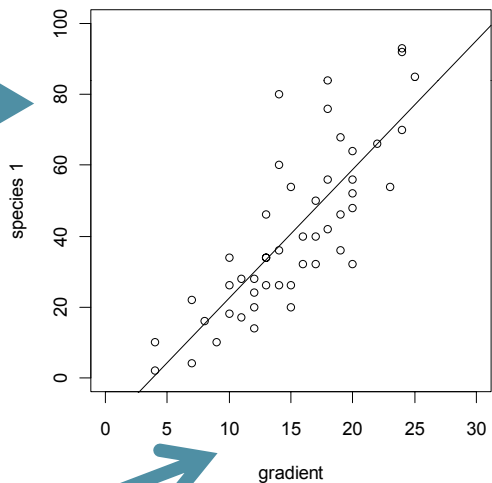
oba přístupy jsou relevantní a navzájem se doplňují!

# PŘÍMÁ ORDINAČNÍ ANALÝZA

původní druhová matice


	spe 1	spe 2	spe 3
sam 1			
sam 2			
sam 3			
sam 4			
sam 5			
sam 6			
sam 7			

regrese abundance druhu na proměnné prostředí




	gradient
sam 1	
sam 2	
sam 3	
sam 4	
sam 5	
sam 6	
sam 7	

ordinační osy s omezením



	spe 1	spe 2	spe 3
sam 1			
sam 2			
sam 3			
sam 4			
sam 5			
sam 6			
sam 7			



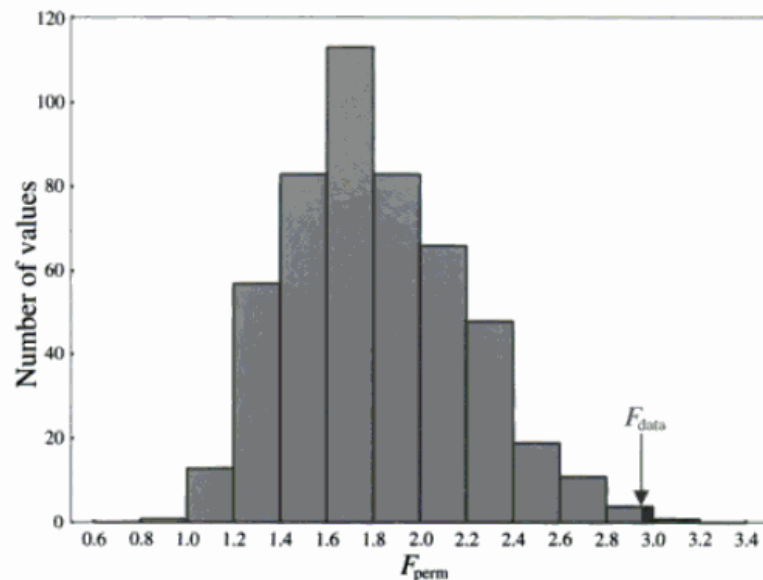
	spe 1	spe 2	spe 3
sam 1			
sam 2			
sam 3			
sam 4			
sam 5			
sam 6			
sam 7			

ordinační osy bez omezení

# PŘÍMÁ ORDINAČNÍ ANALÝZA

## MONTE-CARLO PERMUTAČNÍ TEST

- testuje nulovou hypotézu, že druhové složení je nezávislé na vysvětlující proměnné
- test první kanonické osy – vliv jen jedné kvantitativní proměnné
- test všech kanonických os – vliv všech proměnných, nebo vliv jedné nominální proměnné s více kategoriemi (počet os = počet kategorií – 1)



$$P = \frac{n_x + 1}{N + 1}$$

$n_x$  – počet permutací s

$F_{perm} \geq F_{data}$

$N$  – celkový počet permutací

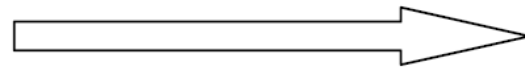


# PŘÍMÁ ORDINAČNÍ ANALÝZA

## MONTE-CARLO PERMUTAČNÍ TEST

ošetřeno	plocha	pokr.
1	1	8
1	2	7
1	3	7
1	7	6
1	8	5
1	9	6
0	4	5
0	5	5
0	6	6
0	10	1
0	11	2
0	12	2
	F =	0.413

RANDOMIZACE: Každé ploše se přiřadí náhodně jeden vegetační zápis (v našem případě jedna hodnota pokrývnosti). Tím se zruší jakákoliv závislost mezi proměnnou prostředí (tj. ošetřením) a hodnotami pokrývnosti.



Z hodnot pokrývnosti spočtu množství variability vysvětlené proměnnými prostředí (F-statistika). Pro hodnoty:

výchozí                      randomizované

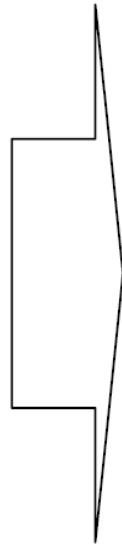
ran.1	ran.2	ran.3	ran.4
7	2	5	7
6	5	8	2
5	6	1	5
1	8	6	7
6	1	7	6
6	5	6	2
8	5	6	5
5	6	7	1
5	7	2	6
2	6	2	5
7	2	5	6
2	7	5	8
0.175	0.005	0.005	0.076

A tak dále. Až po dosažení kýženého počtu randomizací.

# PŘÍMÁ ORDINAČNÍ ANALÝZA

## MONTE-CARLO PERMUTAČNÍ TEST

zásah	pl.	blok	pokr.
<u>1</u>	<u>1</u>	<u>1</u>	<u>8</u>
<u>1</u>	<u>2</u>	<u>1</u>	<u>7</u>
<u>1</u>	<u>3</u>	<u>1</u>	<u>7</u>
<u>0</u>	<u>4</u>	<u>1</u>	<u>5</u>
<u>0</u>	<u>5</u>	<u>1</u>	<u>5</u>
<u>0</u>	<u>6</u>	<u>1</u>	<u>6</u>
1	7	2	6
1	8	2	5
1	9	2	6
0	10	2	1
0	11	2	2
0	12	2	2



ran.1	ran.2	ran.3	ran.4
<u>7</u>	2	<u>5</u>	<u>7</u>
6	<u>5</u>	<u>8</u>	2
<u>5</u>	6	1	<u>5</u>
1	<u>8</u>	6	<u>7</u>
<u>6</u>	1	<u>7</u>	6
6	<u>5</u>	6	2
<u>8</u>	5	<u>6</u>	<u>5</u>
<u>5</u>	<u>6</u>	<u>7</u>	1
5	<u>7</u>	2	<u>6</u>
2	6	2	5
<u>7</u>	2	<u>5</u>	6
2	<u>7</u>	5	<u>8</u>

zásah	pl.	blok	pokr.
<u>1</u>	<u>1</u>	<u>1</u>	<u>8</u>
<u>1</u>	<u>2</u>	<u>1</u>	<u>7</u>
<u>1</u>	<u>3</u>	<u>1</u>	<u>7</u>
<u>0</u>	<u>4</u>	<u>1</u>	<u>5</u>
<u>0</u>	<u>5</u>	<u>1</u>	<u>5</u>
<u>0</u>	<u>6</u>	<u>1</u>	<u>6</u>
1	7	2	6
1	8	2	5
1	9	2	6
0	10	2	1
0	11	2	2
0	12	2	2



ran.1	ran.2	ran.3	ran.4
<u>7</u>	<u>5</u>	<u>6</u>	<u>7</u>
<u>5</u>	<u>7</u>	<u>7</u>	<u>5</u>
<u>7</u>	<u>6</u>	<u>7</u>	<u>6</u>
<u>5</u>	<u>5</u>	<u>5</u>	<u>8</u>
<u>8</u>	<u>7</u>	<u>8</u>	<u>5</u>
<u>6</u>	<u>8</u>	<u>5</u>	<u>7</u>
5	6	2	6
6	2	6	2
2	1	2	6
2	5	1	2
6	2	6	1
1	6	5	5

randomizace ploch bez omezení (*unrestricted randomization*)

randomizace ploch v blocích (*randomization within blocks defined by covariables*)

# JAK ČÍST VÝSLEDKY ORDINAČNÍCH METOD?

- procento variability vysvětlené hlavními osami
  - CANOCO: *cumulative percentage variance of species data*
  - vypočte se také jako *eigenvalue / total variance*
  - ukazuje, jak úspěšný byl celý proces ordinace
  - čím více jsou jednotlivé druhy korelované, tím více variability bude vysvětleno několika málo hlavními osami
  - **má smysl srovnávat** vysvětlenou variabilitu hlavních os různými ordinačními technikami na stejných datech
  - **nemá smysl srovnávat** vysvětlenou variabilitu hlavních os stejnými ordinačními technikami na různých datech (*eigenvalues* jsou závislé na počtu hráčů ve hře – druhů, vzorků)
- skóry (souřadnice) závisle proměnných (druhů) na osách
  - u lineárních technik **skóre = regresní koeficient**, v ordinačních diagramech zobrazeny jako **šipky**
  - u unimodálních technik **skóre = optimum druhu**, v ordinačních diagramech zobrazeny jako **body**

# JAK ČÍST VÝSLEDKY ORDINAČNÍCH METOD?

- skóry vzorků (snímků) na osách
  - v ordinačních diagramech vzorky zobrazeny jako body (lineární i unimodální techniky)
  - vzdálenost mezi body v ordinačním prostoru odpovídá nepodobnosti mezi vzorky (ne ale nepodobnosti celého floristického složení, ale jenom té části, která je vyjádřena zobrazenými ordinačními osami)
- skóry nezávislých (vysvětlujících proměnných) \*
  - regresní koeficienty, důležitá jsou jejich znaménka
- test signifikance (Monte-Carlo permutační test) \*
  - ukazuje na statistickou významnost použitých vysvětlujících proměnných

# JEDNOTLIVÉ PROMĚNNÉ

## TERMINOLOGIE

- vysvětlované / závislé proměnné
  - CANOCO: druhy (*species*)
- vysvětlující / nezávislé proměnné, prediktory \*
  - CANOCO: proměnné prostředí (*environmental variables*)
  - měřené nebo odhadované proměnné
- vzorky, objekty, případy (*cases*)
  - CANOCO: snímky (*samples*)
- kovariáty, nezajímavé vysvětlující / nezávislé proměnné \*
  - CANOCO: kovariáty (*covariables*)
  - proměnné, jejichž vliv nás nezajímá a chceme ho z analýzy odstranit

\* jen přímé ordinační techniky

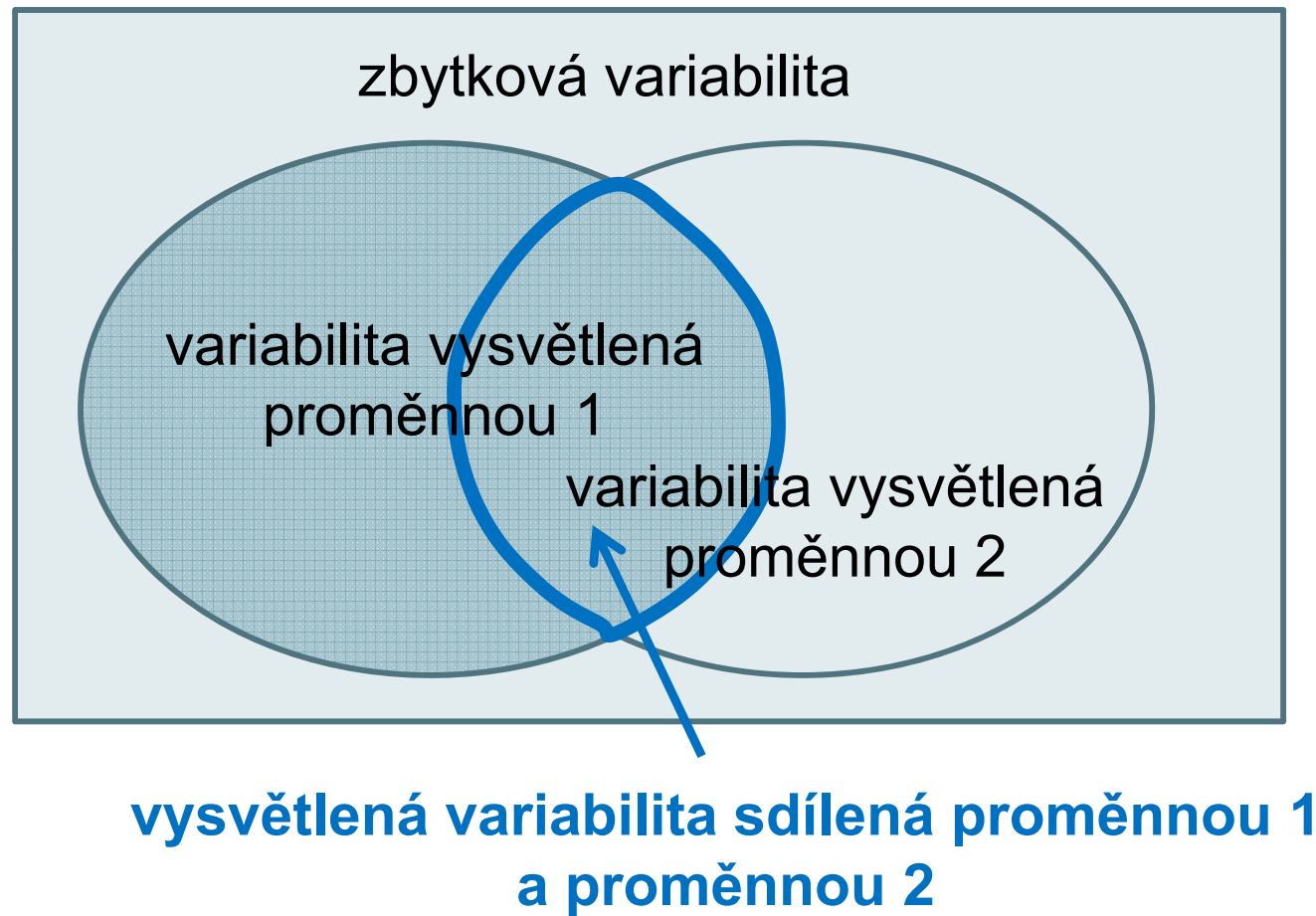
# PARCIÁLNÍ ORDINACE

## *PARTIAL ORDINATION*

- odstraňuje část variability vysvětlené proměnnými, které jsou pro nás nezajímavé (například vliv umístění ploch do bloků)
- následně se přímou nebo nepřímou ordinací analyzuje zbytková variabilita
- „nezajímavé“ proměnné se definují jako **kovariáty**
- pokud následuje přímá ordinace – ordinační osy představují čistý vliv ostatních vysvětlujících proměnných bez vlivu kovariát
- pokud následuje nepřímá ordinace – ordinační osy zachycují zbytkovou variabilitu v druhových datech po odstranění vlivu kovariát

# ROZKLAD VARIANCE

## VARIANCE PARTITIONING



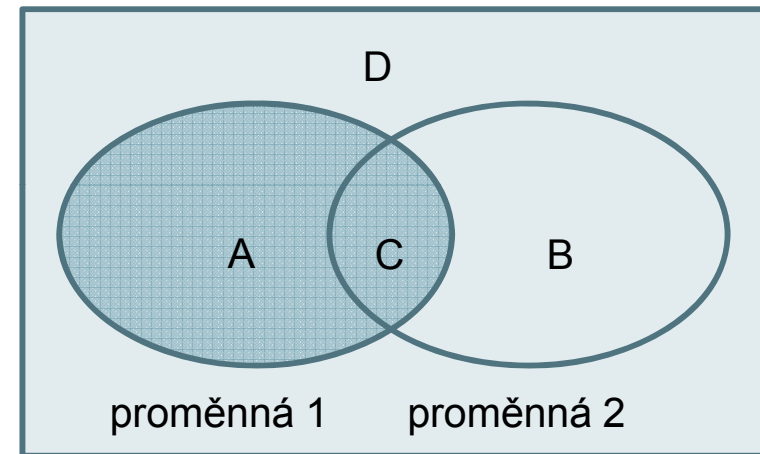
Borcard et al. 1992, *Ecology* 73: 1045–1055

# ROZKLAD VARIANCE

## VARIANCE PARTITIONING

Postup výpočtu:

vysvětlující proměnná	kovariáta	vysvětlená variabilita
1 a 2	není	A + B + C
1	2	A
2	1	B
1	není	A + C



sdílená variabilita  $C = (A + C) - A$

nevysvětlená variabilita  $D = Total\ inertia - (A + B + C)$



# ROZKLAD VARIANCE

## VARIANCE PARTITIONING

ALE:

- ordinační metody jsou založené na modelu (lineární nebo unimodální) odpovědi druhu na gradient prostředí, který je velkým zjednodušením skutečnosti
- variance nevysvětlená modelem (složka D) ve skutečnosti obsahuje variabilitu, která by mohla být vysvětlena některou z proměnných, pokud by se data chovala podle teoretického modelu
- varianci nevysvětlenou modelem tedy nelze interpretovat jako zbytkovou variabilitu, která je dána jednak šumem v datech, jednak faktem, že ne všechny proměnné prostředí byly měřeny
- *Total inertia* proto není měřítkem celkové variability v druhových datech, ale variability, kterou je možné zachytit pomocí zvoleného modelu (lineárního nebo unimodálního)
- variabilita vysvětlená danou proměnnou prostředí a vypočtená jako *eigenvalue / total inertia* je proto podhodnocená
- **doporučení:** nepočítejte procento variability, kterou proměnná vysvětlí, jako poměr *eigenvalue / total inertia*; místo toho se zaměřte na relativní množství variability, kterou daná proměnná vysvětlí z celkové variability vysvětlené všemi proměnnými prostředí

Økland (1999) J. Veg. Sci. 10: 131-136

# POSTUPNÝ VÝBĚR VYSVĚTLUJÍCÍCH PROMĚNNÝCH

## *FORWARD SELECTION*

- ze souboru vysvětlujících proměnných umožňuje vybrat jen ty, které mají průkazný vliv
- v každém kroku testuje zvlášť vliv jednotlivých proměnných (Monte-Carlo permutační test)
- vybere tu proměnnou, která vysvětlí nejvíce variability a zároveň je signifikantní; tuto proměnnou pak do modelu zahrne jako kovariátu
- v dalším kroku znovu testuje vliv jednotlivých proměnných na druhová data (s odstraněním vlivu kovariát) a opakuje předchozí kroky
- testy signifikance jsou zatíženy mnohonásobným porovnáním, a jsou proto poměrně liberální (počet signifikantních proměnných je často nerealisticky vysoký a vyžaduje např. Bonferroniho korekci)

# PROČ ORDINACE?

- není možné zobrazit zároveň víc jak 3 dimenze – přitom ekologická data mají často dimenzí stovky
- redukovaný ordinační prostor reprezentuje hlavní a interpretovatelné ekologické gradienty a redukuje šum (*ordination = noise reduction technique*)
- přímá ordinační metoda v případě statistického testování netrpí problémem mnohonásobného porovnání
- umožňuje určit relativní význam jednotlivých gradientů
- některé techniky (DCA) vyjadřují míru betadiversity v datech
- grafický výstup (ordinační diagram) často nabízí přehledný a snadno interpretovatelný náhled do vztahů mezi druhy (případně vzorky) a proměnnými prostředí