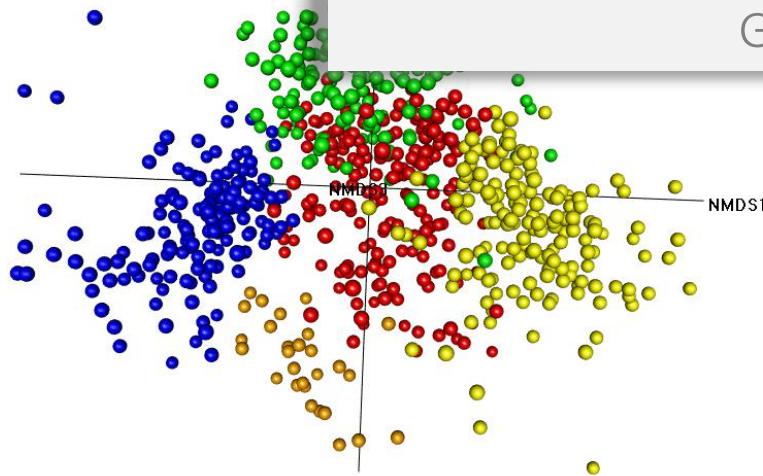


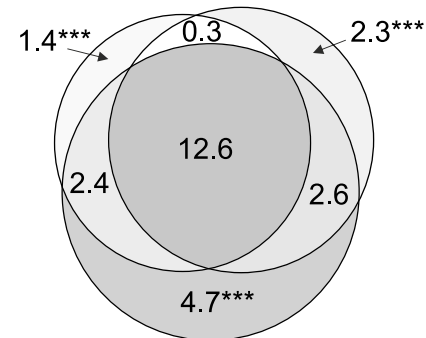
# Metody fyzické geografie 3: Biogeografie & ekologie

Jan Divíšek

Geografický ústav & Ústav botaniky a zoologie



climate: 16.6\*\*\* land-cover: 17.8\*\*\*



natural habitats: 22.3\*\*\*

I TÝ SI ZAPIŠ NOVÝ  
PŘEDMĚT Z 8055  
METODY FYZICKÉ  
GEOGRAFIE 3!



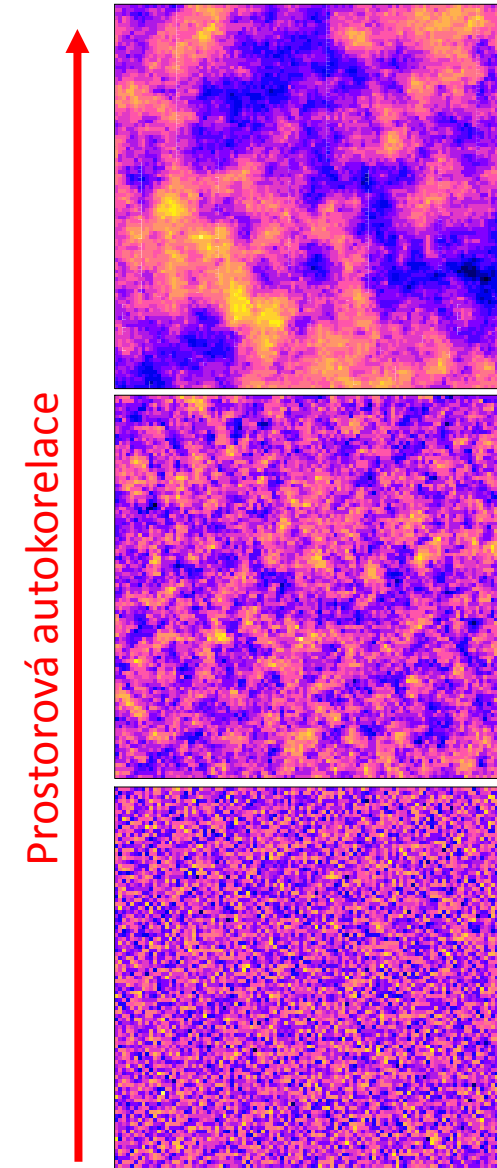
# Metody fyzické geografie 3 – 22. 5. 2017

- Teoretická část
  - Prostorová autokorelace a její měření (Moran's  $I$ , Moran correlogram, Mantel test, Mantel correlogram)
  - Metody prostorové statistiky (GLS, MEM)
- Praktická část
  - zítra

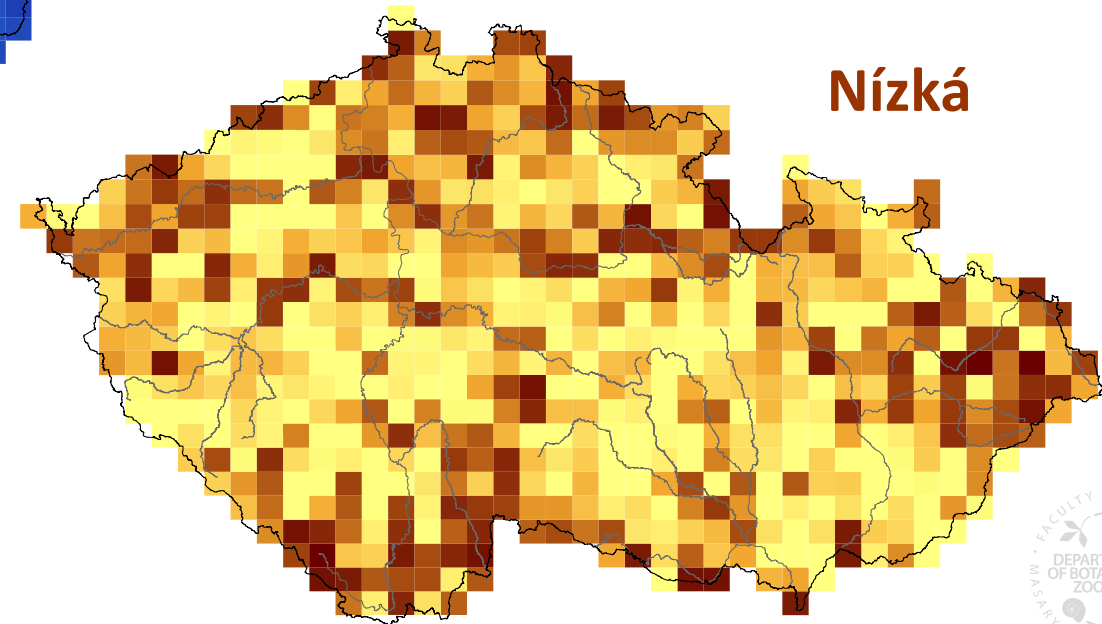
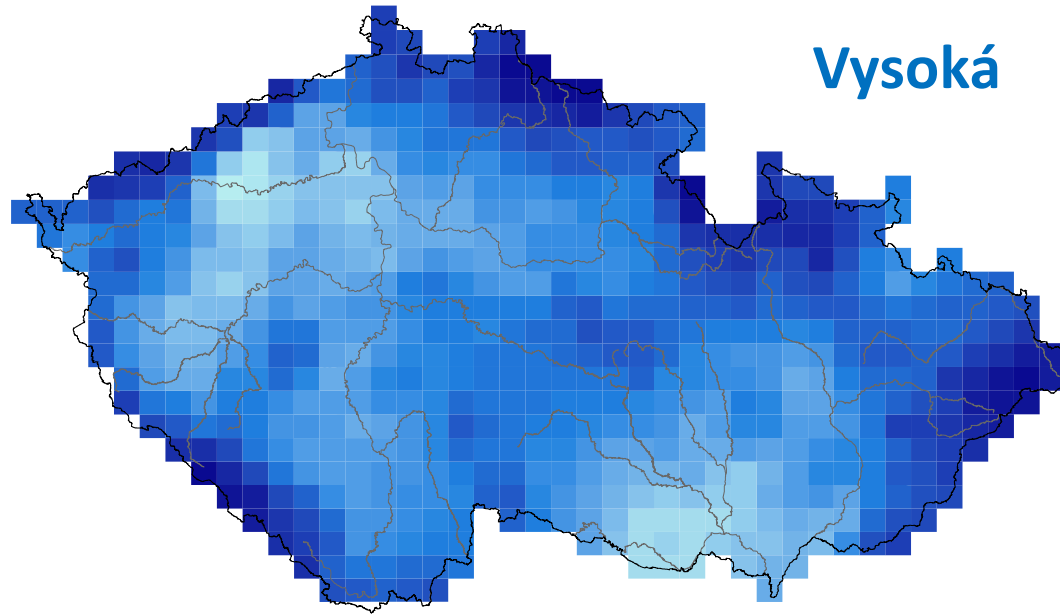
# Prostorová statistika

# Prostorová autokorelace (*Spatial Autocorrelation, SAC*)

- První zákon geografie (Tobler, 1970)
  - „*everything is related to everything else, but near things are more related than distant things*“
- Podobnost dvou míst (lokalit, kvadrátů atp.) koreluje s jejich geografickou vzdáleností
  - bližší lokality jsou si podobnější než vzdálenější → pozitivní autokorelace
  - vzdálenější lokality jsou si podobnější než bližší → negativní autokorelace
- Ekologická data jsou pozitivně autokorelovaná (s výjimkou některých experimentů)
  - Druhové složení → ***distance decay of similarity*** (Nekola & White, 1999)
  - Environmentální podmínky



# Prostorová autokorelace (*Spatial Autocorrelation, SAC*)



# Prostorová autokorelace v biogeografii a ekologii

## Prostorová autokorelace podle Legendre & Legendre (2012)

### 1. Induced spatial dependence

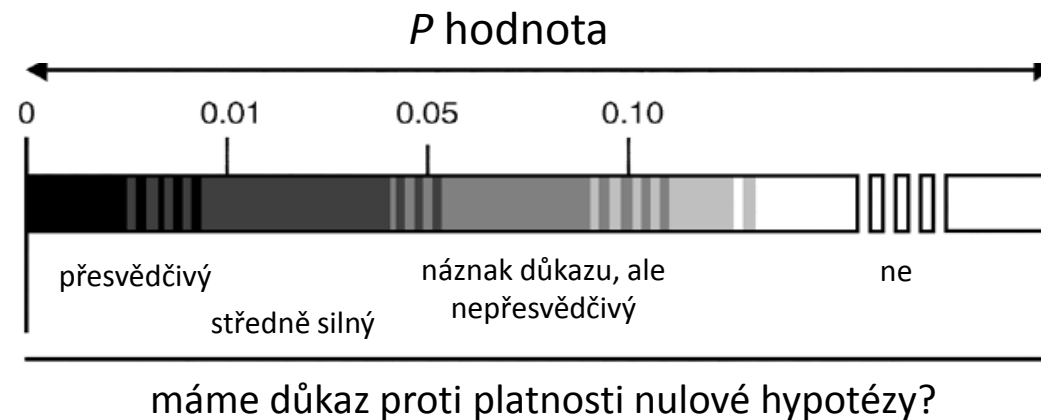
- Prostorová autokorelace v druhových datech v důsledku vlivu autokorelace podmínek prostředí
- Druh reaguje na gradient prostředí a jedinci se šíří se podél něj

### 2. Spatial autocorrelation *sensu stricto*

- Prostorová autokorelace v druhových datech v důsledku procesů ve společenstvu/populaci
- Např. mass effect → z prosperujících populací se jedinci šíří i do míst, kde nemají optimální podmínky

# Důsledky prostorové autokorelace ve statistice

- Testování hypotéz
  - $P$  hodnota jako důkaz, který existuje proti platnosti nulové hypotézy ( $H_0$ ; např. teplota nemá vliv na počet druhů ve vzorcích)



- V případě přítomnosti prostorové autokorelace nejsou jednotlivá pozorování (vzorky) nezávislé
- Každé nové pozorování nepřináší kompletně novou informaci, tj. plný stupeň volnosti
- Počet stupňů volnosti je tak v parametrických testech nadhodnocen → test je příliš liberální, tj.  $H_0$  je zamítána příliš často
  - výsledek vychází signifikantně, i když by neměl
- Pouze pokud závislá i vysvětlující proměnná je prostorově autokorelovaná

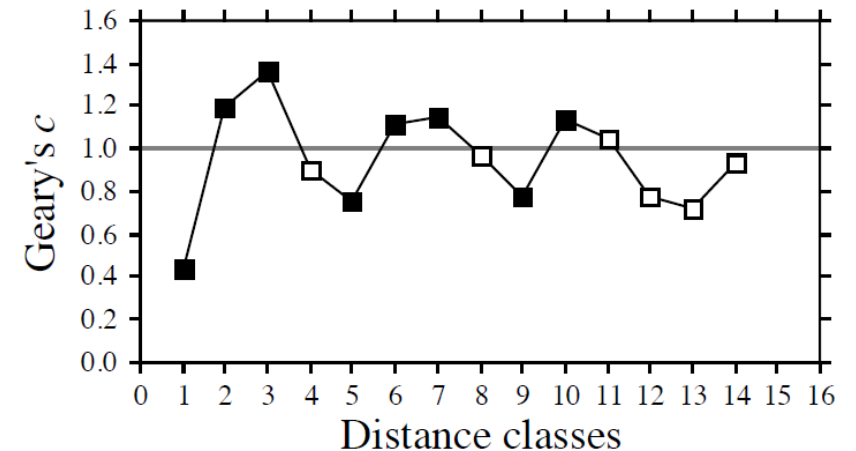
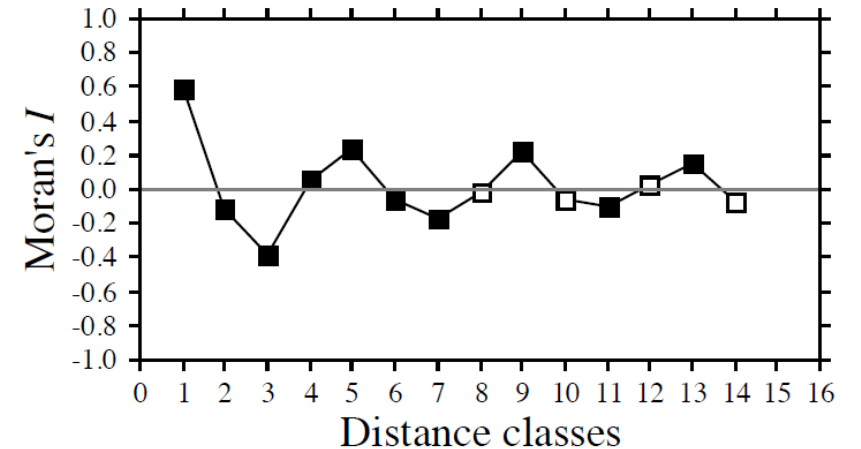
# Měření prostorové autokorelace

## Jednorozměrná data (*univariate data*)

- Globálně (hodnota popisující celkovou autokorelaci proměnné)
  - Moran's  $I$
  - Geary's  $c$
- Lokálně (autokorelace na jednotlivých vzdálenostních třídách)
  - Moran's correlogram
  - Geary's correlogram
  - Semi-variogram

## Vícerozměrná data (*multivariate data*)

- Globálně
  - Mantelova korelace
- Lokálně
  - Mantel correlogram
  - Multivariate variogram





# Global Moran's I

- Index pro měření prostorové autokorelace (Moran, 1950)
- Nabývá hodnot od -1 pro negativní autokorelaci po 1 pro pozitivní autokorelaci

- Očekávaná hodnota Moranova indexu pro nulovou prostorovou autokorelaci ( $H_0$ ) je:

$$E(I) = \frac{-1}{N - 1}$$

- Pro výpočet je nutné stanovit vzdálenosti vzorků, které jsou následně převedeny na „prostorové váhy“ (často inverzní hodnoty ke vzdálenostem)
- V praxi se používají jen vzdálenosti nižší než stanovený threshold

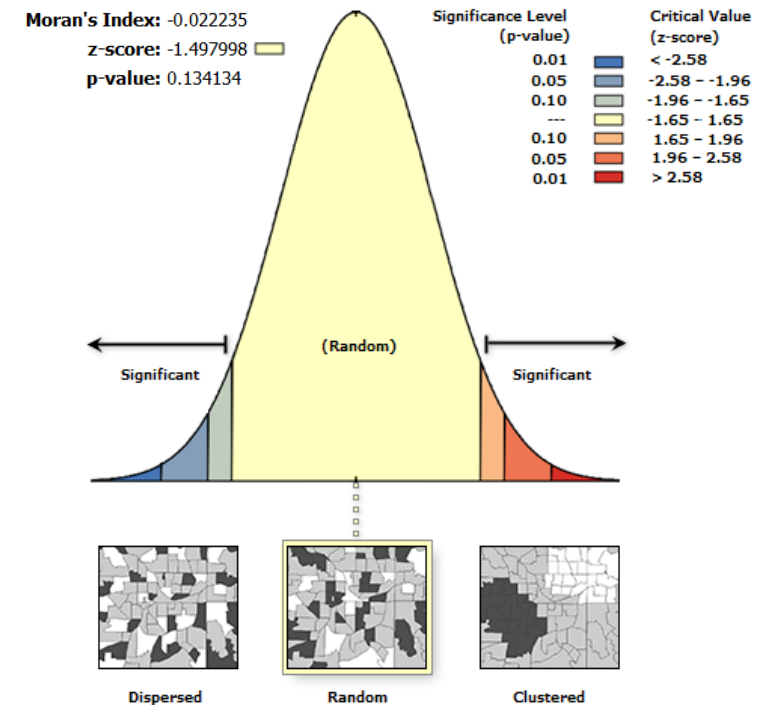
$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

$N$  je počet vzorků

$y_i$  a  $y_j$  jsou hodnoty proměnné  $y$  ve vzorku  $i$  a  $j$

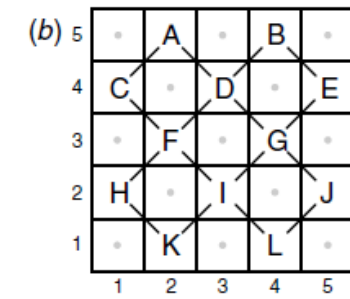
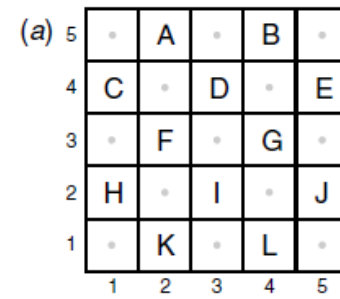
$\bar{y}$  je průměrná hodnota proměnné

$w_{ij}$  je prostorová váha (většinou inverzní vzdálenost) vzorku  $i$  a  $j$



```
moran.test {spdep}  
Moran.I {ape}
```

# Prostorové váhy



Euclidean distance matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	2.00	0										
C	1.41	3.16	0									
D	1.41	1.41	2.00	0								
E	3.16	1.41	4.00	2.00	0							
F	2.00	2.83	1.41	1.41	3.16	0						
G	2.83	2.00	3.16	1.41	1.41	2.00	0					
H	3.16	4.24	2.00	2.83	4.47	1.41	3.16	0				
I	3.16	3.16	2.83	2.00	2.83	1.41	1.41	2.00	0			
J	4.24	3.16	4.47	2.83	2.00	3.16	1.41	4.00	2.00	0		
K	4.00	4.47	3.16	3.16	4.24	2.00	2.83	1.41	1.41	3.16	0	
L	4.47	4.00	4.24	3.16	3.16	2.83	2.00	3.16	1.41	1.41	2.00	0

Connectivity matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	0	0										
C	1	0	0									
D	1	1	0	0								
E	0	1	0	0	0							
F	0	0	1	1	0	0						
G	0	0	0	1	1	0	0					
H	0	0	0	0	0	1	0	0				
I	0	0	0	0	0	1	1	0	0			
J	0	0	0	0	0	0	1	0	0	0		
K	0	0	0	0	0	0	0	1	1	0	0	
L	0	0	0	0	0	0	0	0	1	1	0	0

Distance class matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	2	0										
C	1	3	0									
D	1	1	2	0								
E	3	1	3	2	0							
F	2	2	1	1	3	0						
G	2	2	3	1	1	2	0					
H	3	3	2	2	3	1	3	0				
I	3	3	2	2	2	1	1	2	0			
J	3	3	3	2	2	3	1	3	2	0		
K	3	3	4	4	3	2	2	1	1	3	0	
L	3	3	3	3	3	2	2	3	1	1	2	0

Weighted matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	0	0										
C	0.7	0	0									
D	0.7	0.7	0	0								
E	0	0.7	0	0	0							
F	0	0	0.7	0.7	0	0						
G	0	0	0	0.7	0.7	0	0					
H	0	0	0	0	0	0.7	0	0				
I	0	0	0	0	0	0.7	0.7	0	0			
J	0	0	0	0	0	0	0.7	0	0	0		
K	0	0	0	0	0	0	0	0.7	0.7	0	0	
L	0	0	0	0	0	0	0	0	0.7	0.7	0	0



# Moran's correlogram

- Měří prostorovou autokorelaci na jednotlivých vzdálenostních třídách
  - data jsou rozdělena do tříd ( $d$ ) podle vzdáleností mezi vzorky
  - pro každou třídu je spočítáno Moranovo  $I$
- Detailnější informace o prostorové struktuře dat
- Při testování statistické významnosti
  - Měla by být použita korekce pro vícenásobné testování (např. Bonferoni (konzervativnější) nebo Holm (liberálnější))
  - Nemělo by být prováděno na datech vykazujících trend → detrendování (např. pomocí lineární regrese)

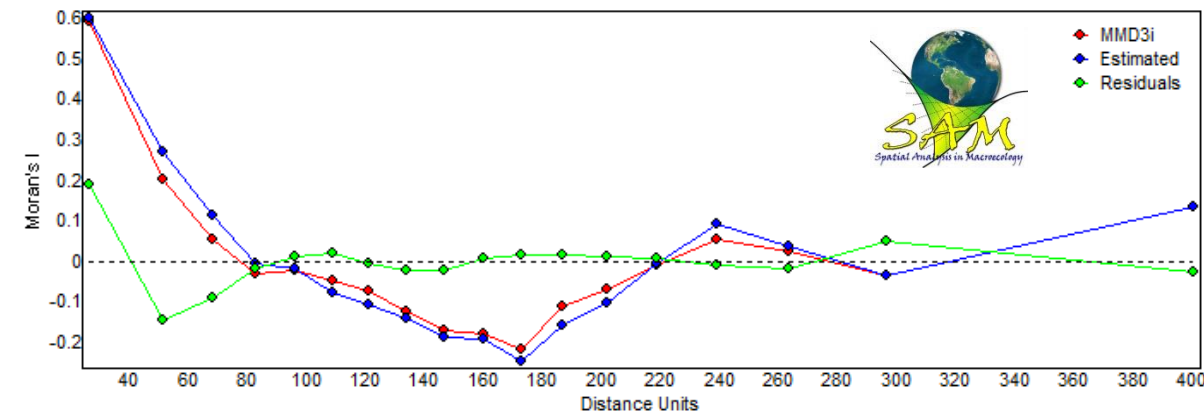
$$I(d) = \frac{\frac{1}{W} \sum_{h=1}^n \sum_{i=1}^n w_{hi} (y_h - \bar{y})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$$

$y_h$  a  $y_j$  jsou hodnoty proměnné  $y$  ve vzorku  $h$  a  $j$

$W$  je počet párů vzorků použitých pro výpočet autokorelace v dané vzdálenostní třídě

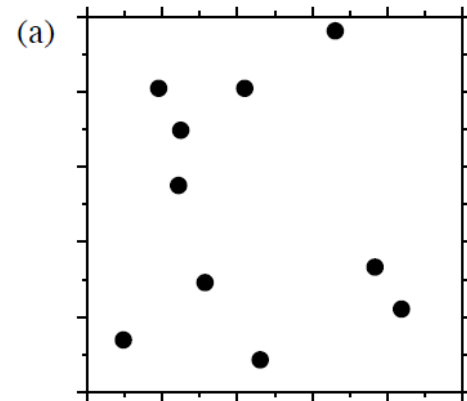
$w_{ij}$  je prostorová váha pro pár vzorků  $h$  a  $i$ ;

$w_{hi} = 1$  pro páry vzorků, které náležejí do stejné vzdálenostní třídy a  
 $w_{hi} = 0$  pro páry vzorků, které nealéžejí do stejné vzdálenostní třídy



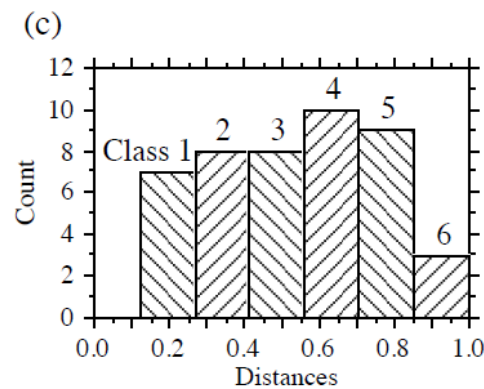
`sp.correlogram {spdep}`

# Výpočet vzdálenostních tříd (lagů)



(b) Distance matrix  $D$  (lower triangular):

0.00	0.52	0.74	0.20	0.31	0.29	0.72	0.72	0.59	0.23
0.52	0.00	0.27	0.41	0.27	0.75	0.52	0.25	0.45	0.53
0.74	0.27	0.00	0.58	0.44	1.00	0.74	0.37	0.70	0.67
0.20	0.41	0.58	0.00	0.15	0.49	0.76	0.65	0.63	0.12
0.31	0.27	0.44	0.15	0.00	0.59	0.68	0.51	0.57	0.26
0.29	0.75	1.00	0.49	0.59	0.00	0.76	0.90	0.64	0.50
0.72	0.52	0.74	0.76	0.68	0.76	0.00	0.40	0.13	0.87
0.72	0.25	0.37	0.65	0.51	0.90	0.40	0.00	0.39	0.77
0.59	0.45	0.70	0.63	0.57	0.64	0.13	0.39	0.00	0.74
0.23	0.53	0.67	0.12	0.26	0.50	0.87	0.77	0.74	0.00



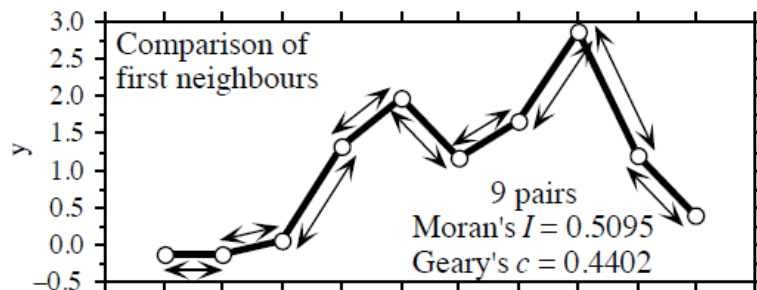
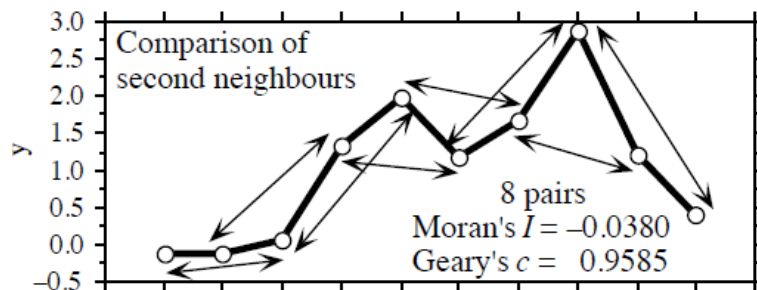
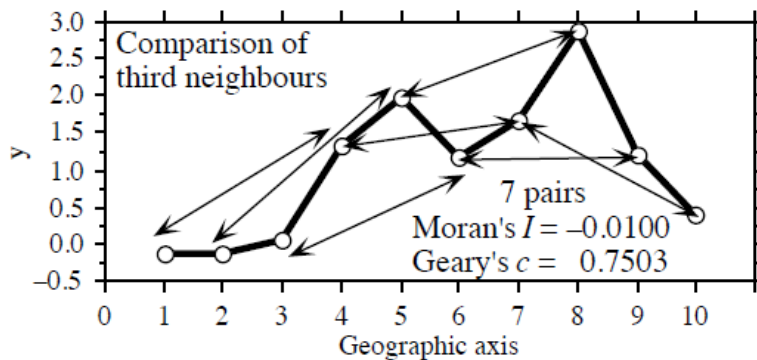
(d) Distance matrix  $D$  (upper triangular):

—									
3	—								
5	2	—							
1	2	4	—						
2	2	3	1	—					
2	5	6	3	4	—				
5	3	5	5	4	5	—			
5	1	2	4	3	6	2	—		
4	3	4	4	4	4	1	2	—	
1	3	4	1	1	3	6	5	5	—

- Sturgesho pravidlo:  $d = 1 + (3.3219 \times \log_{10} n)$   
 $n$  je počet párových vzdáleností

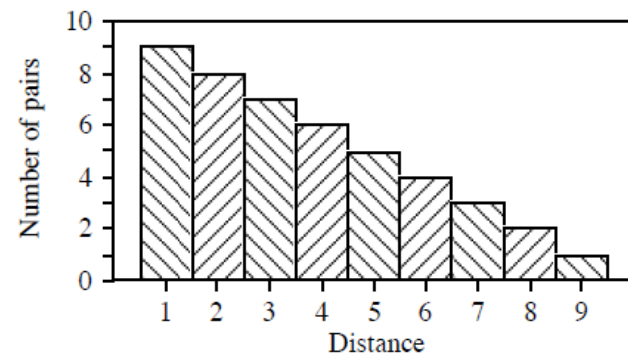
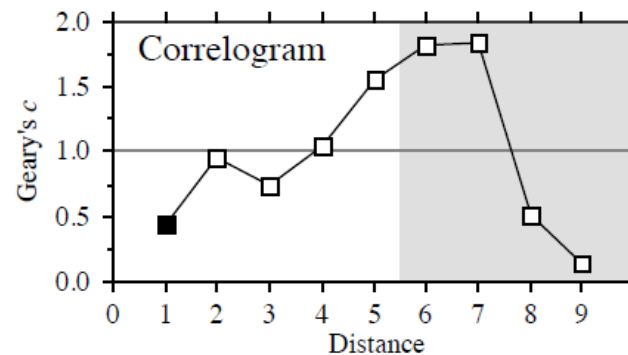
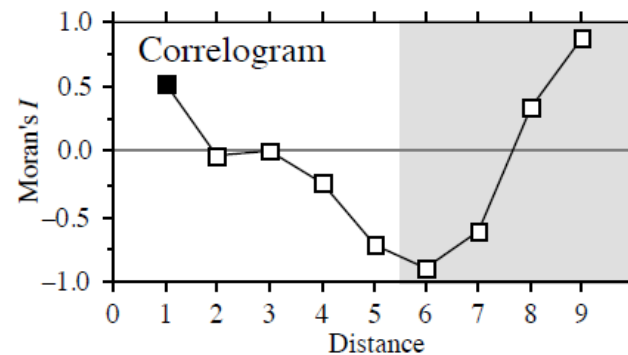
# Výpočet korelogramu

Distance

 $d = 1$  $d = 2$  $d = 3$ 

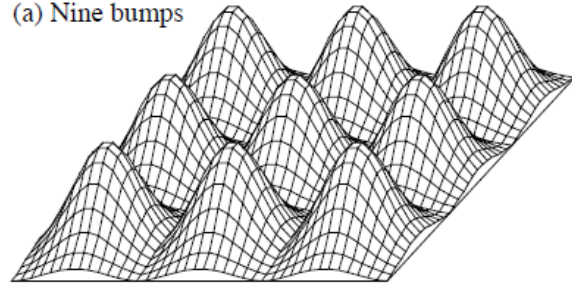
etc.

etc.

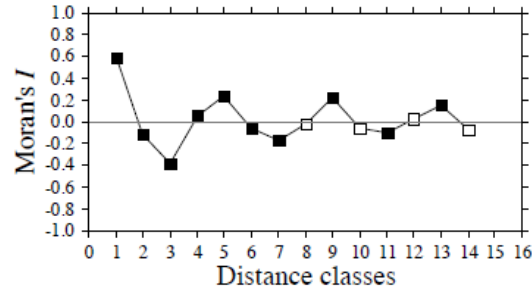


Legendre &amp; Legendre (2012)

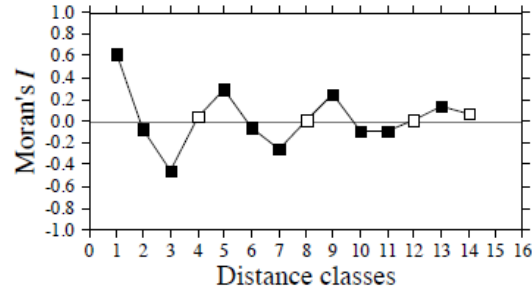
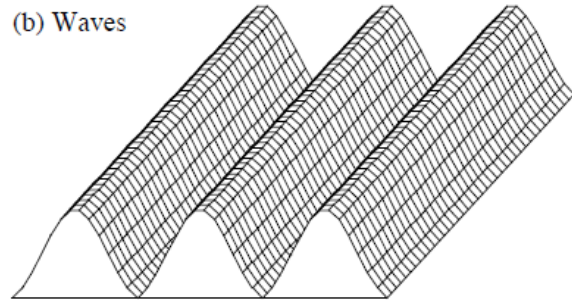
(a) Nine bumps



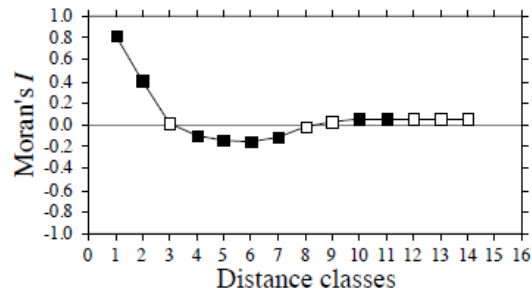
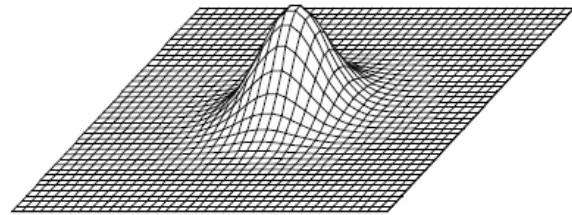
Moran's correlograms



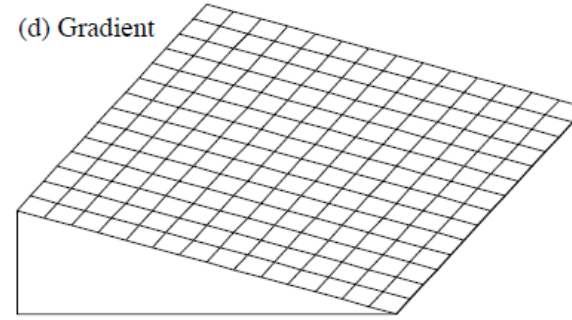
(b) Waves



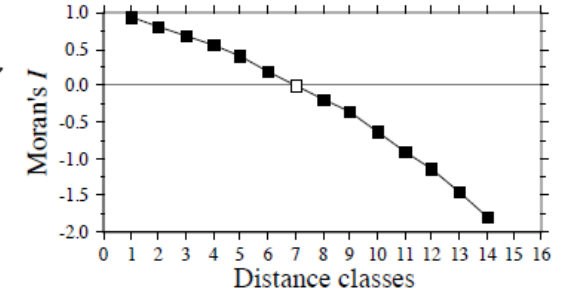
(c) Single bump



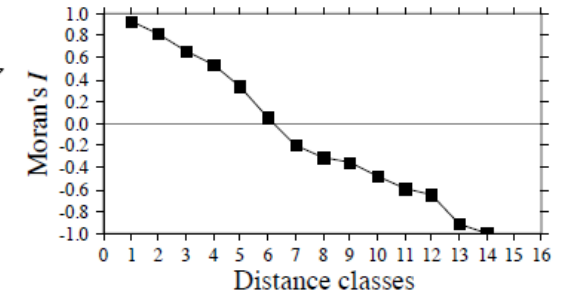
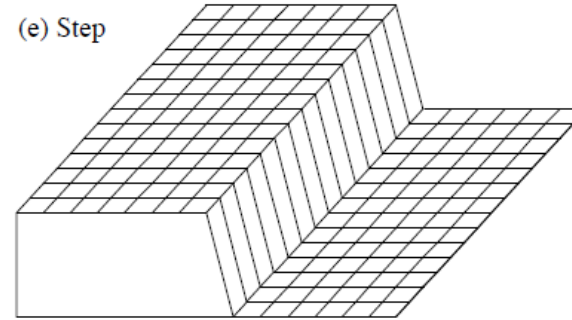
(d) Gradient



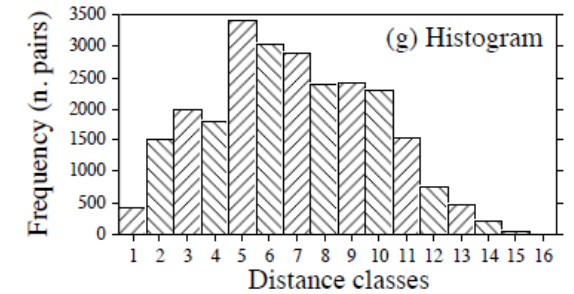
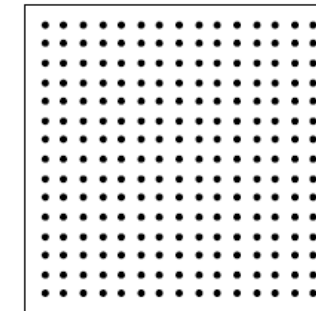
Moran's correlograms



(e) Step



(f) Sampling grid (15 x 15)



# Výstup funkce `sp.correlogram` v R

`sp.correlogram {spdep}`

```
Spatial correlogram for aves$No_species
method: Moran's I
```

		estimate	expectation	variance	standard deviate	Pr(I)	two sided	
1	(628)	3.0684e-01	-1.5949e-03	4.3900e-04	14.7206	< 2.2e-16	***	
2	(628)	1.3306e-01	-1.5949e-03	2.3695e-04	8.7475	< 2.2e-16	***	
3	(628)	5.4154e-02	-1.5949e-03	1.6814e-04	4.2993	0.0001884	***	
4	(628)	1.0936e-02	-1.5949e-03	1.3390e-04	1.0830	1.0000000		
5	(628)	1.6033e-02	-1.5949e-03	1.1401e-04	1.6510	0.6912327		
6	(628)	2.6393e-02	-1.5949e-03	1.0144e-04	2.7789	0.0545481	.	
7	(628)	2.4765e-02	-1.5949e-03	9.3158e-05	2.7311	0.0546199	.	
8	(628)	2.4164e-02	-1.5949e-03	8.8120e-05	2.7440	0.0546199	.	
9	(628)	8.9880e-03	-1.5949e-03	8.5640e-05	1.1436	1.0000000		
10	(628)	3.0362e-03	-1.5949e-03	8.4838e-05	0.5028	1.0000000		
11	(628)	6.4611e-05	-1.5949e-03	8.5581e-05	0.1794	1.0000000		
12	(628)	-6.9681e-03	-1.5949e-03	8.8079e-05	-0.5725	1.0000000		
13	(628)	-1.1453e-02	-1.5949e-03	9.2761e-05	-1.0235	1.0000000		

Vzdálenostní  
třída (lag)

Počet  
pozorování  
zahrnutých v  
lagu

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pozorovaná hodnota  
Moranova I

Očekávaná hodnota  
Moranova I

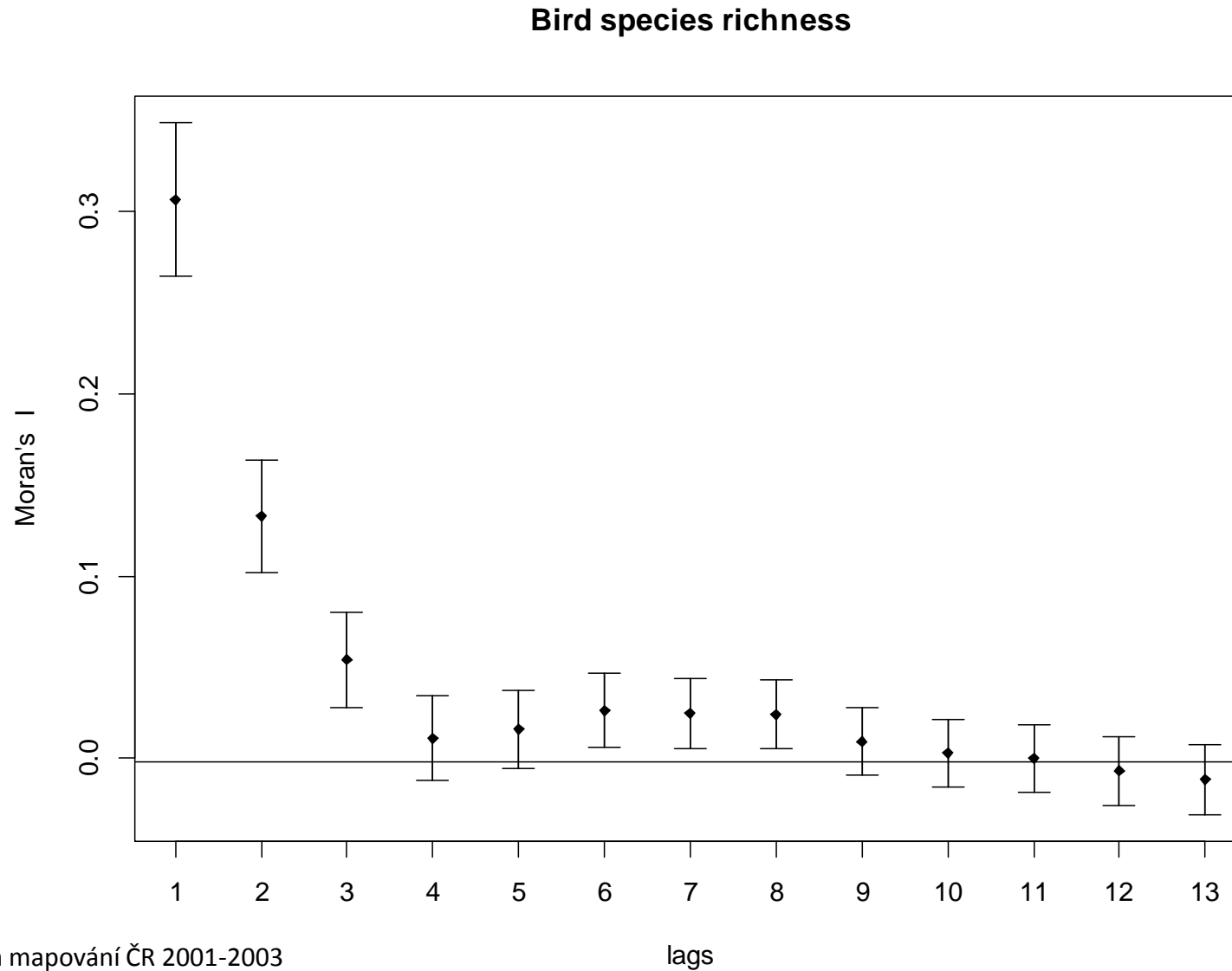
Statistická signifikance založená na randomizaci  
(pukud `randomisation = TRUE`)

Pomocí funkce

```
print(spatial_correlogram,
      p.adj.method = "holm")
je možné provést korekci p-hodnot
```

Diverzita ptáků v kvadrátovém mapování ČR 2001-2003

# Výstup funkce `sp.correlogram` v R

`sp.correlogram {spdep}`

Diverzita ptáků v kvadrátovém mapování ČR 2001-2003



# Semi-variogram

$$\gamma(d) = \frac{1}{2W(d)} \sum_{h=1}^{n-1} \sum_{i=h+1}^n w_{hi} (y_h - y_i)^2$$

- Graf míry **nepodobnosti** vzorků v určité vzdálenostní třídě
- Hojně se využívá v geostatistice → kriging

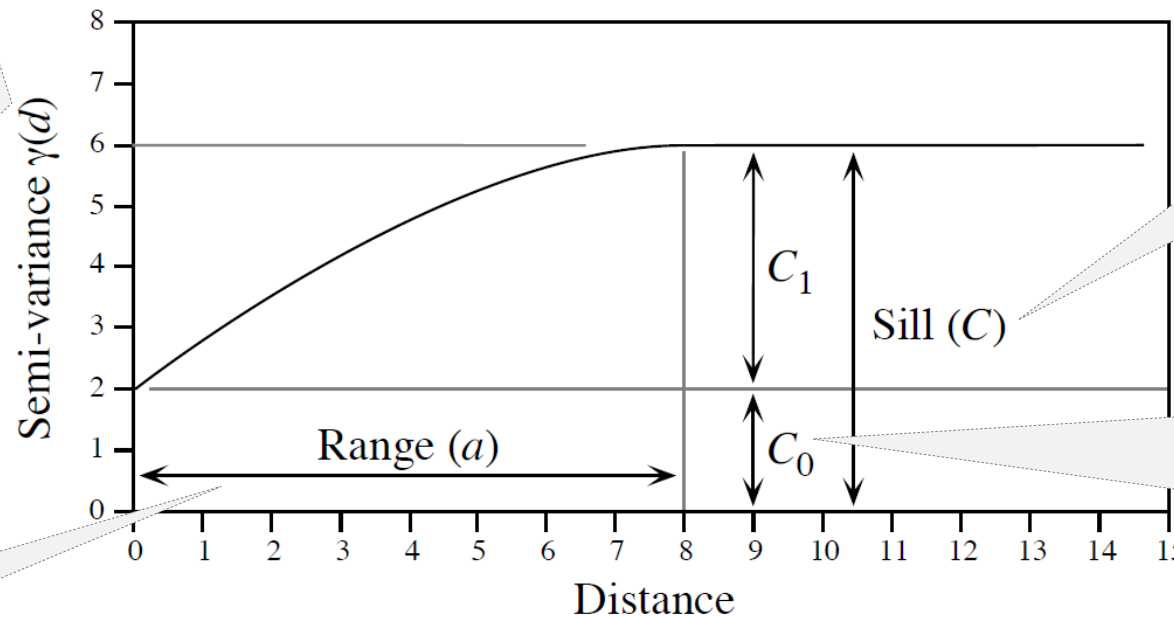
$y_h$  a  $y_j$  jsou hodnoty proměnné  $y$  ve vzorku  $h$  a  $j$

$W$  je počet párů vzorků použitých pro výpočet autokorelace v dané vzdálenostní třídě

$w_{ij}$  je prostorová váha pro pár vzorků  $h$  a  $i$ ;

$w_{hi} = 1$  pro páry vzorků, které náležejí do stejné vzdálenostní třídy a  $w_{hi} = 0$  pro páry vzorků, které neažijí do stejné vzdálenostní třídy

Rozptyl zkoumané proměnné jako funkce vzájemných vzdáleností jednotlivých vzorků. Je-li vzdálenost mezi dvěma body malá a jejich hodnoty jsou podobné, potom je hodnota semi-variance také malá.



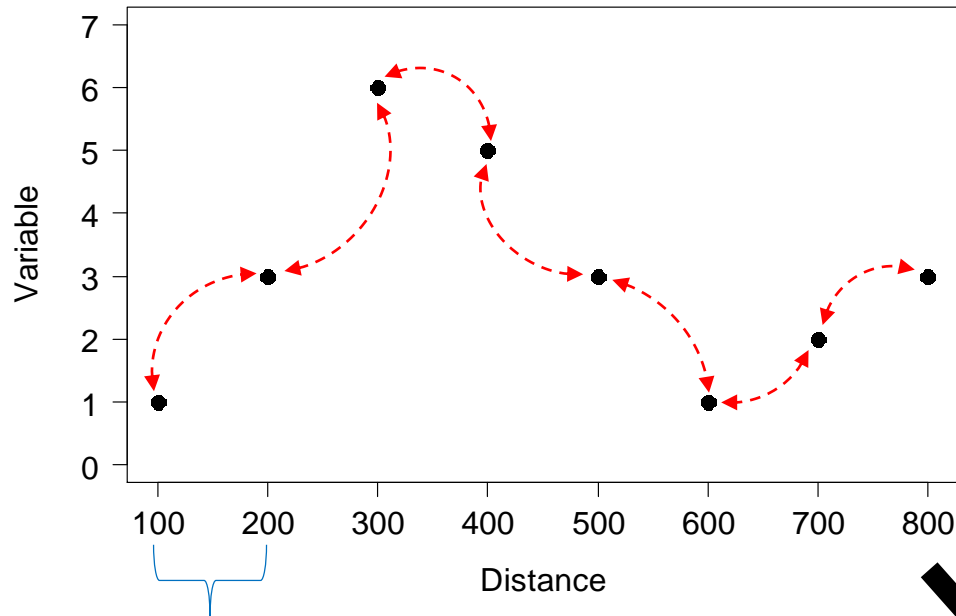
Prahová hodnota, kdy semi-variance je rovna celkovému rozptylu dat a se zvyšující se vzdáleností tedy nemůže dále růst.

Tzv. zbytkový rozptyl (*nugget effect*). Ukazuje na rozptyl menší než je "vzorkovací" vzdálenost, nebo na malou přesnost měření, např. když jsou v datech obsaženy dva vzorky ze stejného místa, pokaždé s jinou hodnotou.

Vzdálenost, do které semi-variance roste.

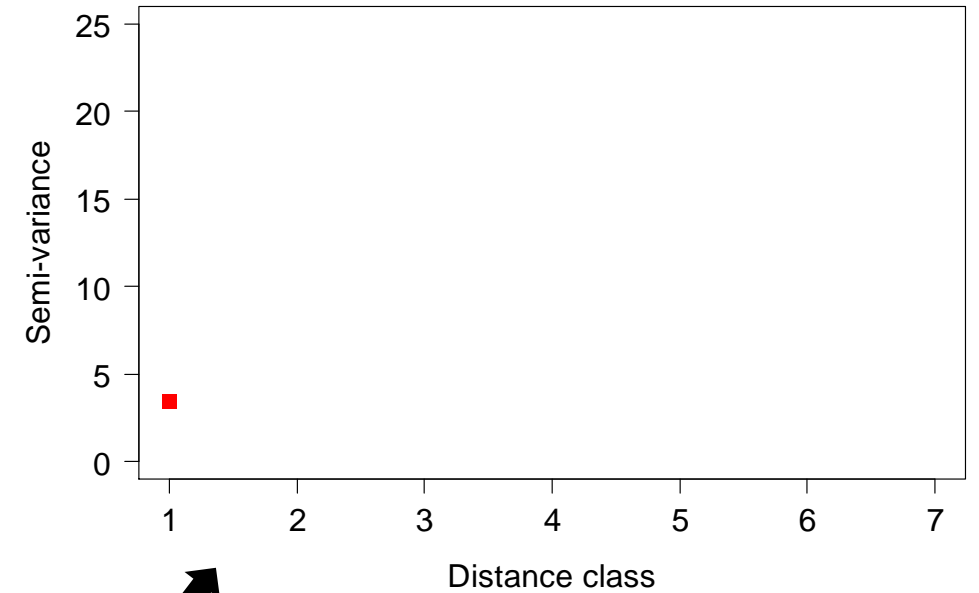
Hranice vzdálenostních tříd nebo vzdálenostní třídy (lagy).

# Výpočet semi-variogramu v 1D

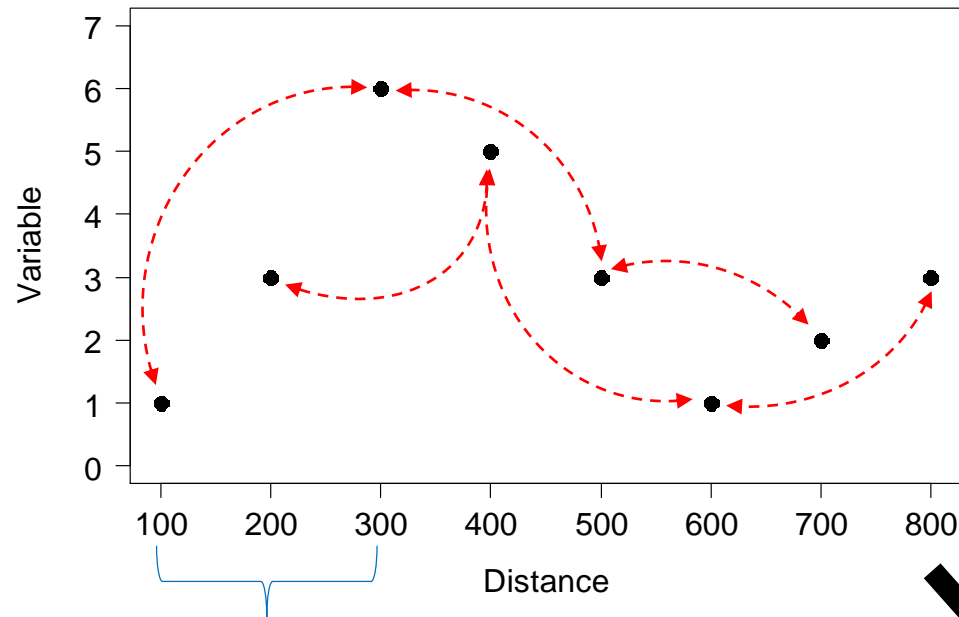


$d_1$  – vzdálenosti  $\leq 100$

$$[(1-3)^2 + (3-6)^2 + (6-5)^2 + (5-3)^2 + (3-1)^2 + (1-2)^2 + (2-3)^2] / 7 = 3.43$$

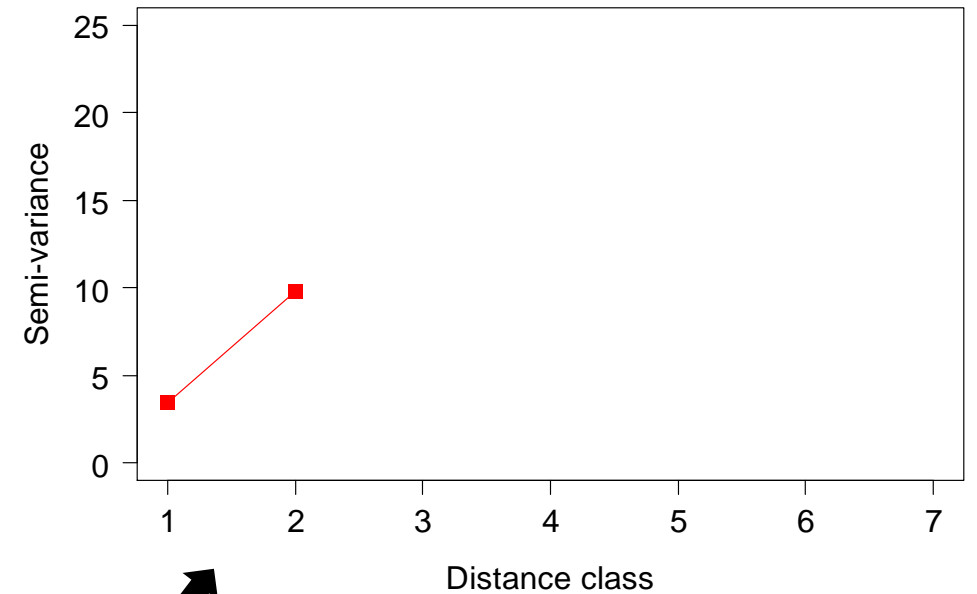


# Výpočet semi-variogramu v 1D

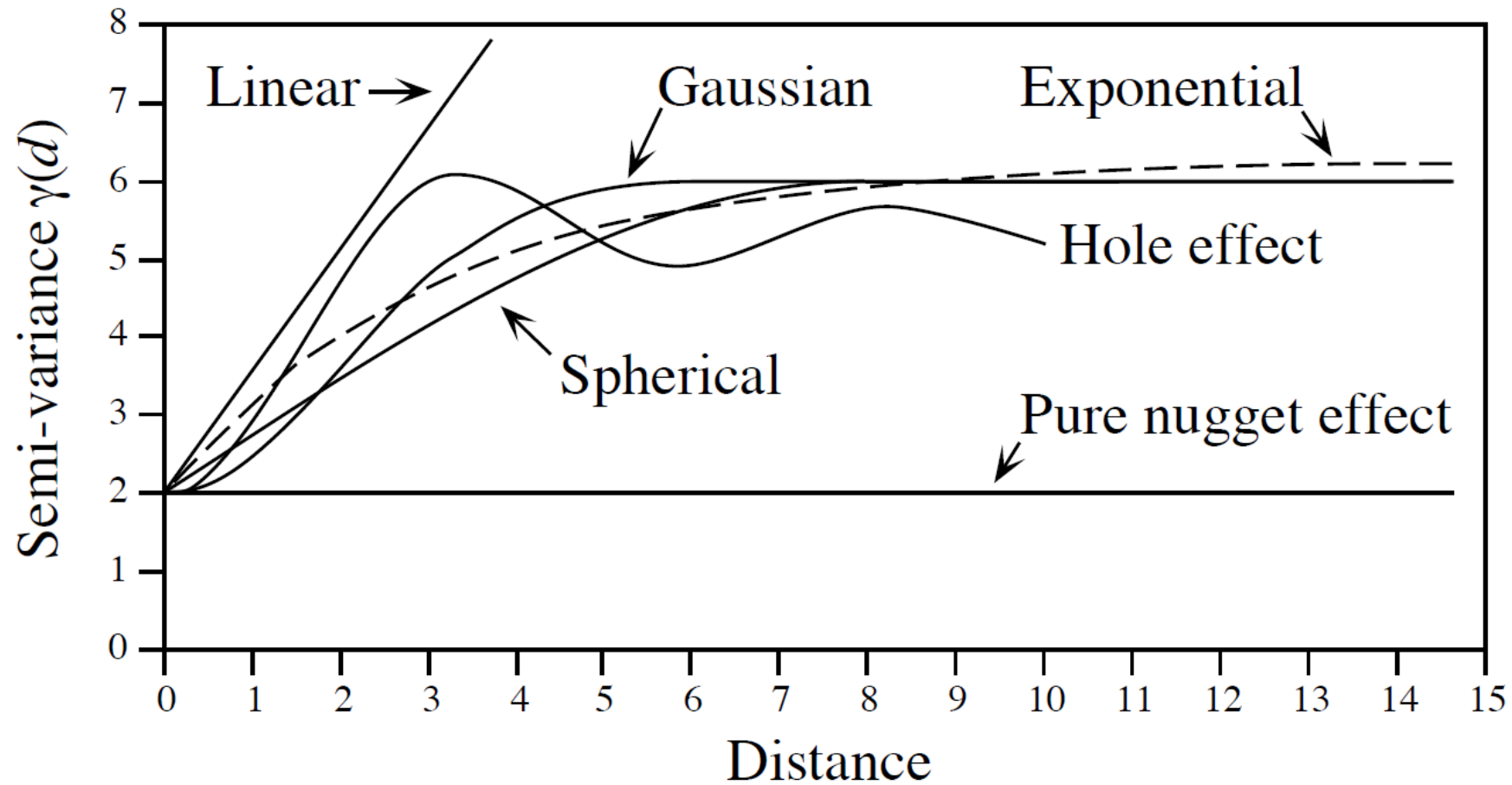


$d_2$  – vzdálenosti  $> 100$  a  $\leq 200$

$$[(1-6)^2 + (3-5)^2 + (6-3)^2 + (5-1)^2 + (3-2)^2 + (1-3)^2] / 6 = 9.83$$

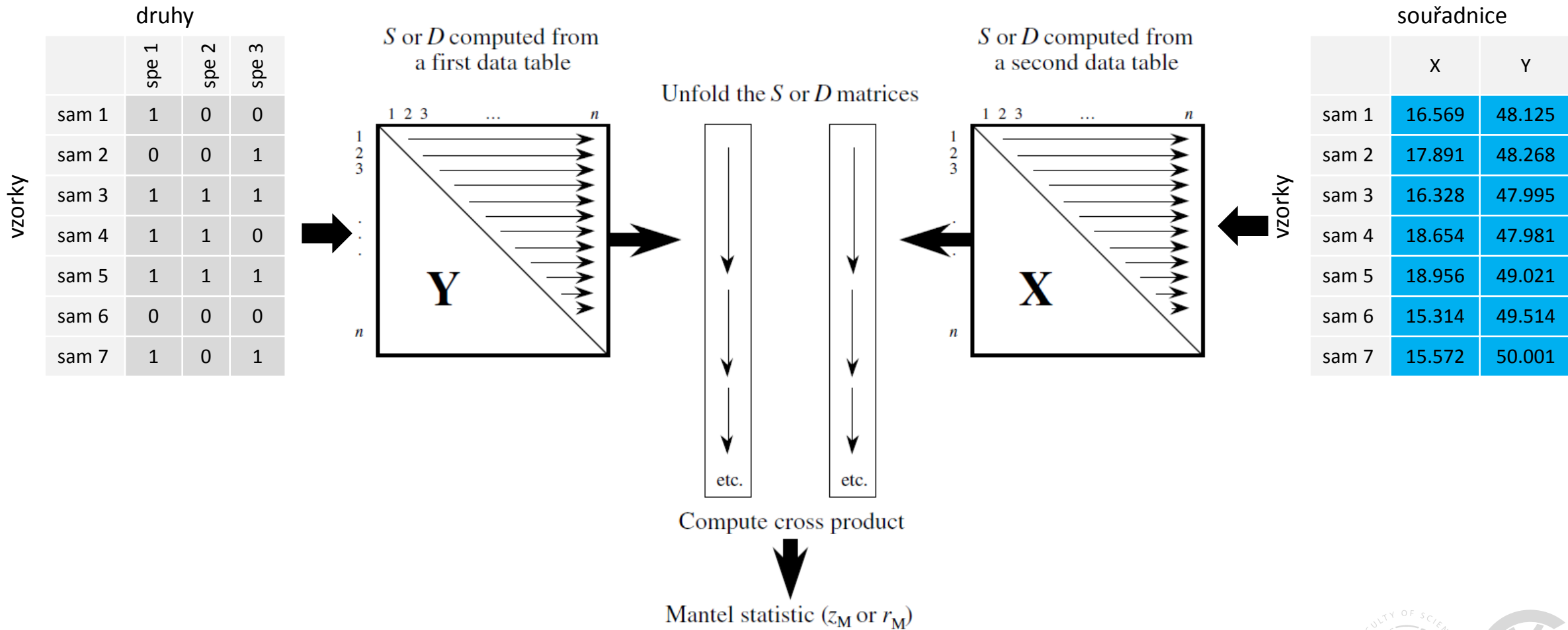


# Nejpoužívanější teoretické semi-variogramy



# Mantelova korelace

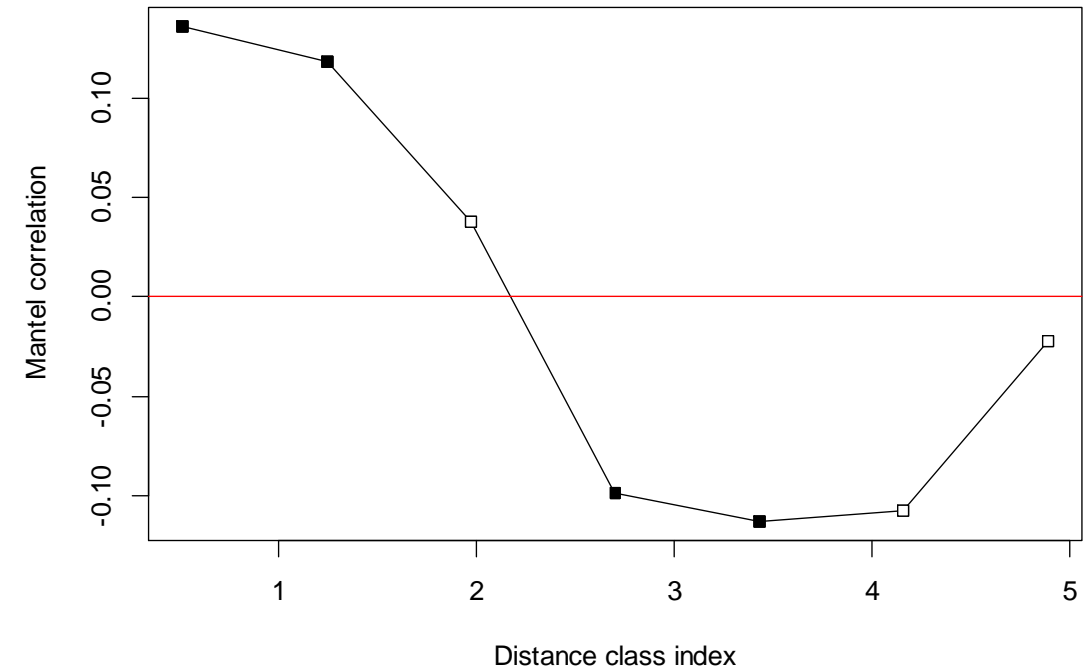
```
mantel {vegan}
```



# Mantelův korelogram

`mantel.correlog {vegan}`

- Měří prostorovou autokorelaci ve vícerozměrných datech na jednotlivých vzdálenostních třídách
  - data jsou rozdělena do tříd ( $d$ ) podle vzdáleností mezi vzorky
  - pro každou třídu je spočítána Mantelova korelace (Pearsonův korelační koeficient)
- Při testování statistické významnosti
  - Měla by být použita korekce pro vícenásobné testování (např. Bonferoni (konzervativnější) nebo Holm (liberálnější))
  - Ve funkci `mantel.correlog` knihovny `vegan` je aplikována progresivní korekce  $P$  hodnot
  - Nemělo by být prováděno na datech vykazujících trend → detrendování (např. pomocí RDA)



# Mantelův korelogram v R

```
mantel.correlog {vegan}
```

Mantel Correlogram Analysis

Call:

```
mantel.correlog(D.eco = aves.hel.det, XY = coord_S42, nperm = 199)
```

Hodnota Mantelovy  
korelace (Pearsonův  
korelační koeficient)

P-hodnota založená na  
permutaci (zde 199)

P-hodnota korigovaná na  
vícenásobné testování (Holm).  
Pokud `progressive = TRUE`,  
potom je provedena progresivní  
korekce P-hodnot. První hodnota  
bez korekce, druhá korigovaná na 2  
testy, třetí na 3 atd.

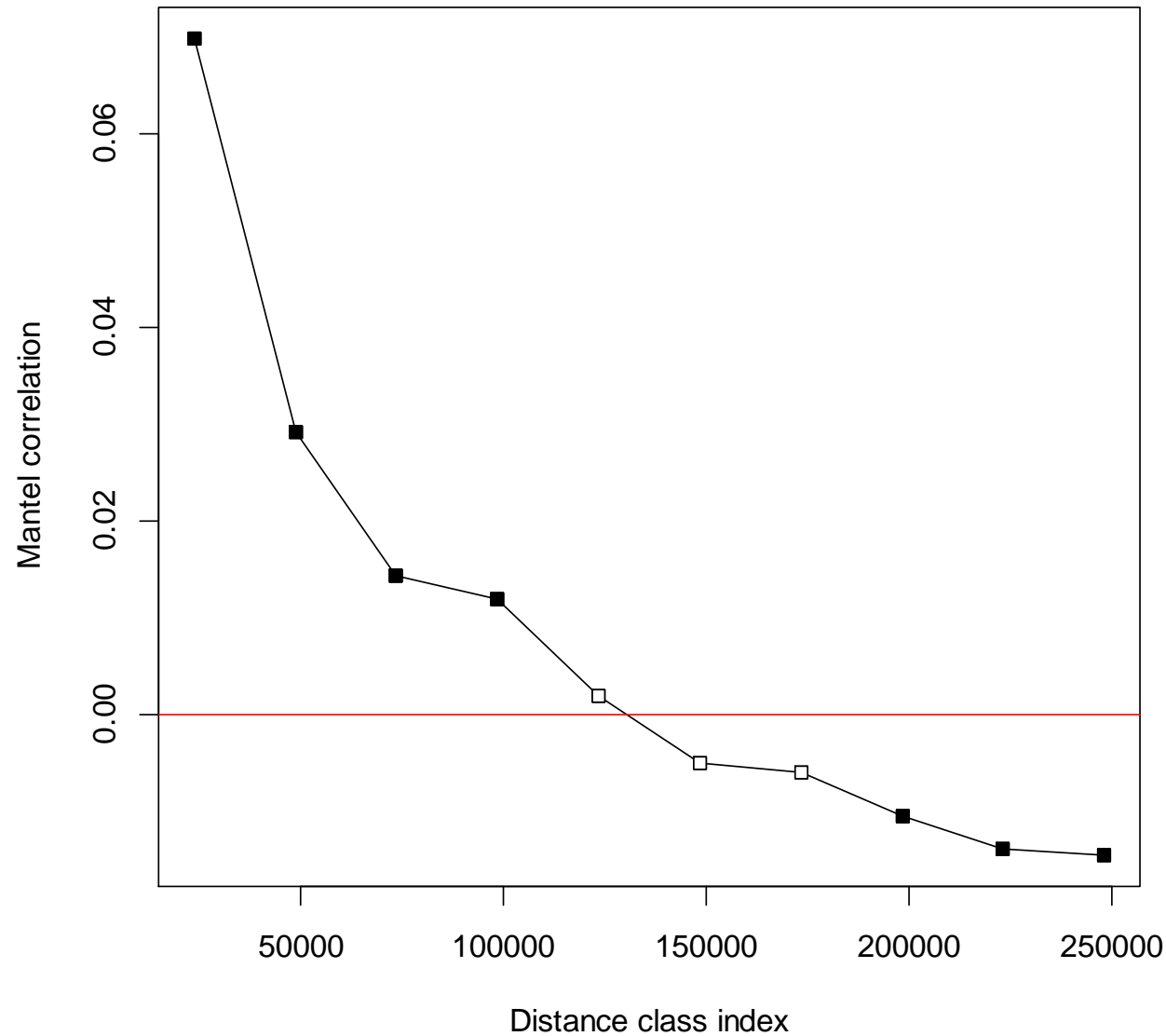
Vzdálenostní  
třída (lag)

	class.index	n.dist	Mantel.cor	Pr (Mantel)	Pr (corrected)
D.cl.1	2.3596e+04	1.7014e+04	6.9862e-02	0.005	0.005 **
D.cl.2	4.8550e+04	2.8498e+04	2.9236e-02	0.005	0.010 **
D.cl.3	7.3504e+04	3.6594e+04	1.4367e-02	0.005	0.015 *
D.cl.4	9.8458e+04	4.0234e+04	1.2032e-02	0.010	0.020 *
D.cl.5	1.2341e+05	4.5454e+04	1.9082e-03	0.355	0.355
D.cl.6	1.4837e+05	4.3338e+04	-5.0490e-03	0.165	0.330
D.cl.7	1.7332e+05	3.9334e+04	-5.9919e-03	0.145	0.435
D.cl.8	1.9827e+05	3.4810e+04	-1.0481e-02	0.005	0.040 *
D.cl.9	2.2323e+05	2.9192e+04	-1.3889e-02	0.005	0.045 *
D.cl.10	2.4818e+05	2.2874e+04	-1.4450e-02	0.005	0.050 *
D.cl.11	2.7314e+05	1.7826e+04	NA	NA	NA
D.cl.12	2.9809e+05	1.3912e+04	NA	NA	NA
D.cl.13	3.2304e+05	1.0184e+04	NA	NA	NA
D.cl.14	3.4800e+05	6.8500e+03	NA	NA	NA
D.cl.15	3.7295e+05	4.1820e+03	NA	NA	NA
D.cl.16	3.9791e+05	2.2360e+03	NA	NA	NA
D.cl.17	4.2286e+05	9.4000e+02	NA	NA	NA
D.cl.18	4.4781e+05	2.4200e+02	NA	NA	NA
D.cl.19	4.7277e+05	4.0000e+01	NA	NA	NA

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Mantelův korelogram v R

```
mantel.correlog {vegan}
```





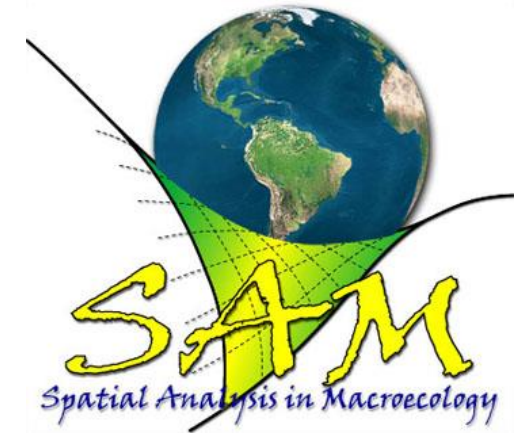
# Modelování prostorové struktury dat

- Mnoho metod sloužících k popisu prostorové struktury v datech
  - Kvantifikace prostorové autokorelace
    - porovnávání vlivu environmentálních proměnných a geografické vzdálenosti lokalit
  - Zohlednění prostorové informace ve statistických modelech
    - korektní testování hypotéz v případě autokorelovaných dat
- Korelační analýza s korekcí stupňů volnosti
- Metody prostorové regresní analýzy
  - Generalized Least Squares (GLS)
  - Simultaneous Autoregression (SAR)
  - Conditional Autoregression (CAR)
  - Geographically Weighted Regression (GWR)
- Metody založené na generování prostorových proměnných (*spatial filters*)
  - Moran's eigenvector maps (MEM)

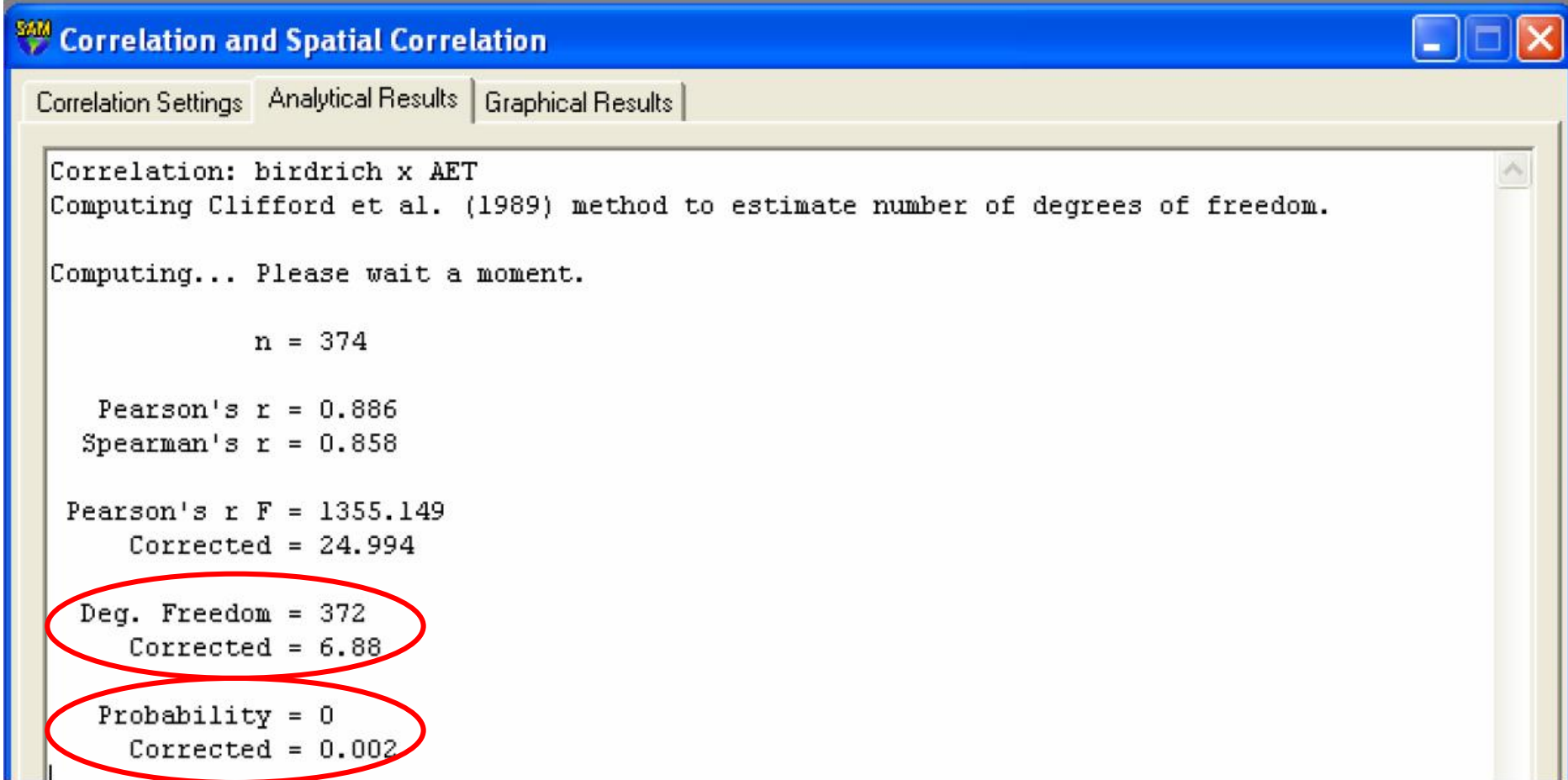
# Korelační analýza s korekcí stupňů volnosti

- Testy statistické významnosti korelačního koeficientu jsou provedeny s redukovaným počtem stupňů volnosti (*degrees of freedom*)
- Clifford et al. (1989)
  - Relativně rychlý výpočet
  - Pracuje dobře na velkých datech s pozitivní autokorelací
- Dutilleul (1993)
  - Pomalejší výpočet
  - Pracuje dobře pro jakoukoliv velikost dat a typ prostorové struktury

```
modified.ttest {SpatialPack}
```



# Korelační analýza v S.A.M.



The screenshot shows the SAM software window titled "Correlation and Spatial Correlation". It has three tabs: "Correlation Settings", "Analytical Results", and "Graphical Results". The "Analytical Results" tab is active, displaying the following text:

```
Correlation: birdrich x AET
Computing Clifford et al. (1989) method to estimate number of degrees of freedom.

Computing... Please wait a moment.

      n = 374

Pearson's r = 0.886
Spearman's r = 0.858

Pearson's r F = 1355.149
      Corrected = 24.994

Deg. Freedom = 372
      Corrected = 6.88

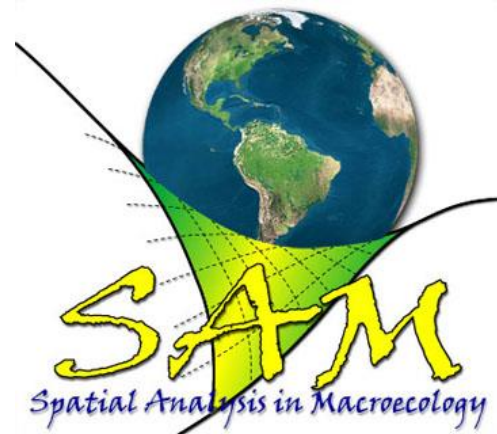
Probability = 0
      Corrected = 0.002
```

Two pairs of lines are circled in red: "Deg. Freedom = 372" and "Corrected = 6.88", and "Probability = 0" and "Corrected = 0.002".

# Generalized Least Squares (GLS)

- Metoda prostorové regrese někdy nazývaná také jako „kriging regression“
- Použití
  - Pro data, která obsahují prostorovou/časovou strukturu
  - Pokud jsou rezidua z lineárního regresního modelu autokorelovaná (*correlated errors*)
- Reziduální autokorelace z lineárního regresního modelu je zde modelována pomocí semi-variogramu
- S jeho pomocí se pak vypočítají (odhadnou) regresní koeficienty

```
gls {nlme}
```



# GLS v S.A.M.

**GLS Settings** | Modeling Residual Spatial Component

Semi-Variance

Distance Units

Empirical

Estimated

OLS Residuals Semi-Variance:

D.Class	Up.Limit	Emp.Semi-Var.	Est.Semi-Var.
1	3	2108.932	2063.994
2	4.123	2608.54	2270.417
3	5.385	2779.428	2496.991
4	6.325	3002.705	2660.953
5	7.211	2965.359	2811.31
6	8.062	3089.185	2951.093
7	9.055	3112.802	3107.817
8	10	2979.392	3249.784

Spatial Modeling

Available Models:

- Spherical Model
- Exponential Model
- Gaussian Model
- Hole Effect Model
- Powered Exponential Model
- Matérn Model
- Truncated Linear Model
- Pure Nugget Model

Note: Not all parameter combinations yield an invertible covariance matrix

Parameters

$C_0$ : 1500      a: 21

$C_1$ : 2650      b:

Model Diagnostics:

n: 17       $r^2$ : 0.829      AIC: 205.434

	Coeff	P
Intercept:	760.395	0.027
Slope:	0.759	0.02

Compute GLS

Nastavení parametrů semi-variogramu

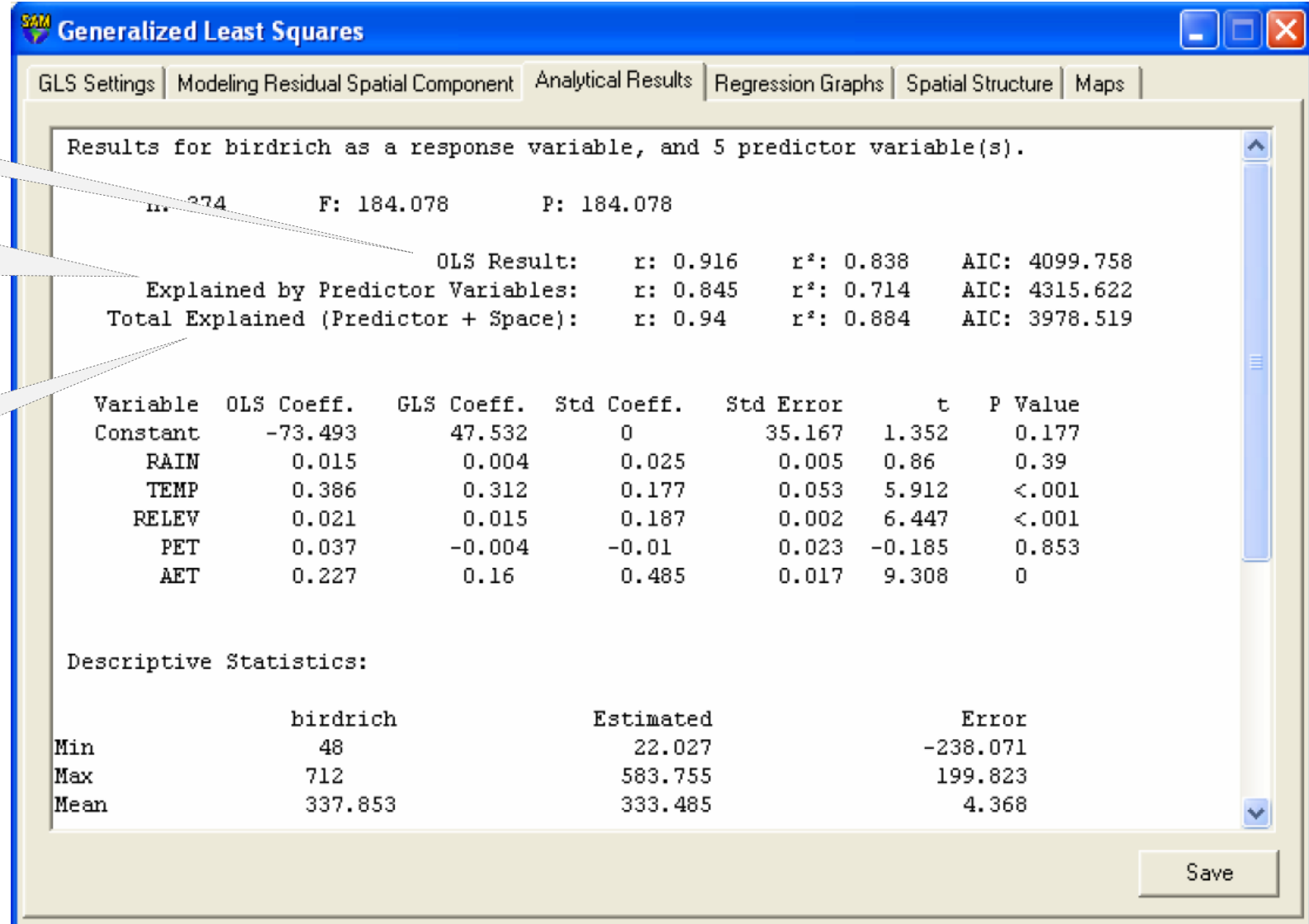
„Fit“ prostorového modelu

# GLS v S.A.M.

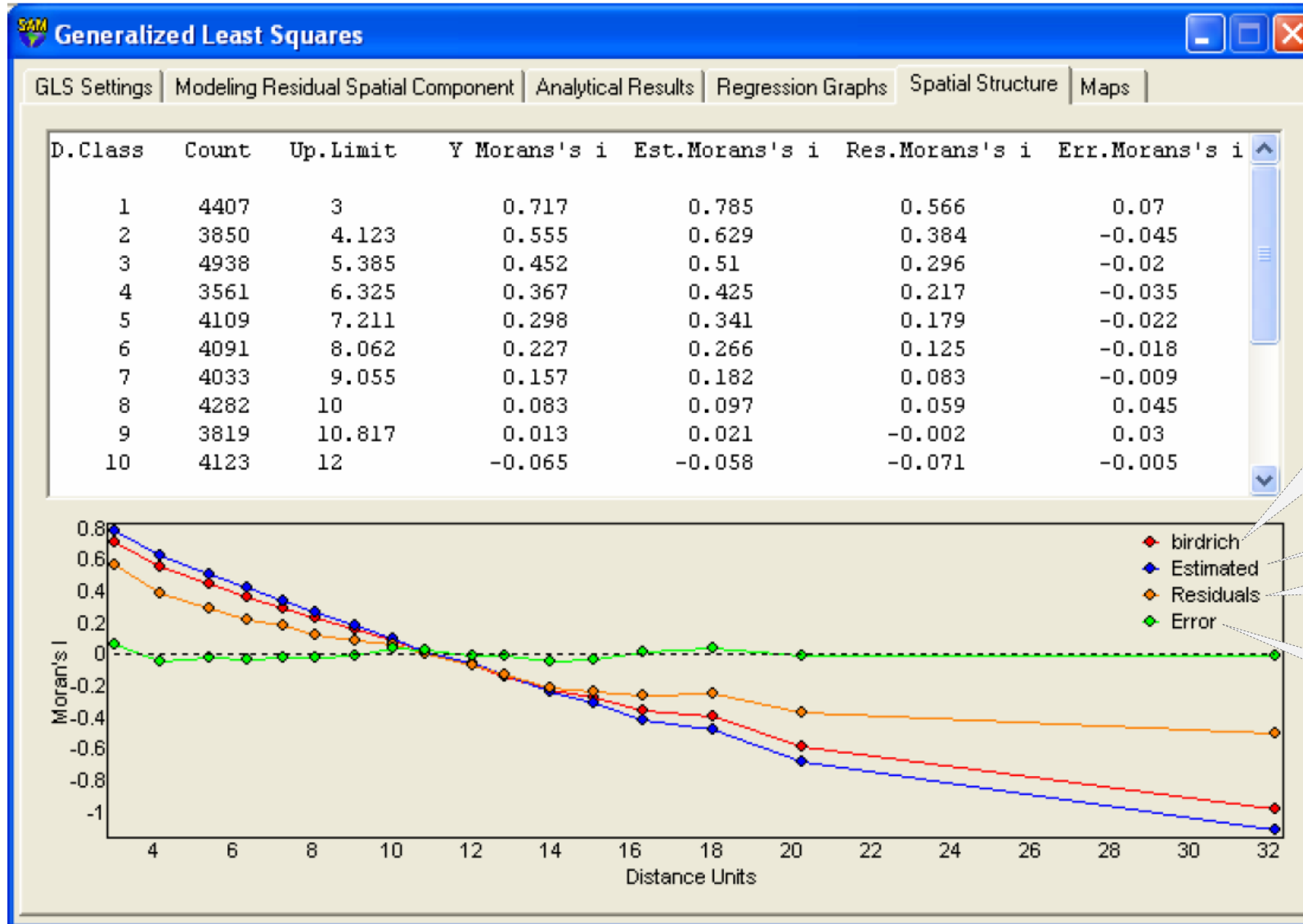
$R^2$  klasické lineární regrese

pseudo- $R^2$  vypočítaný pomocí korelace pozorovaných a predikovaných hodnot

Celkový  $R^2$  modelu, tj. včetně prostorové struktury



# GLS v S.A.M.



Prostorová struktura pozorovaných hodnot

Prostorová struktura predikovaných hodnot

Prostorová struktura v datech, která nebyla vysvětlena vysvětlujícími proměnnými

Residua – bez prostorové struktury



# GLS v R

```
gls {nlme}
```

```
gls(No_species ~ TMAX + Wetland, data=aves, corr=corExp(form=~X + Y))
```

Vysvětlovaná (závislá)  
proměnná

Vysvětlující proměnné

Korelační struktura definovaná typem  
teoretického semi-variogramu.

Další možnosti:

```
corGaus = Gaussian model  
corSpher = Spherical model  
corLin = Linear model
```

Geografické souřadnice  
(metrické – zde S42)



# GLS v R

```
Generalized least squares fit by REML
Model: No_species ~ TMAX + Wetland
Data: aves
      AIC      BIC    logLik
5466.41 5488.598 -2728.205
```

```
gls {nlme}
```

Akaikeho informační kritérium. Čím menší, tím lepší model

```
Correlation Structure: Exponential spatial correlation
Formula: ~X + Y
Parameter estimate(s):
  range
9006.713
```

Typ použitého teoretického semi-variogramu

Koeficienty pro vysvětlující proměnné

```
Coefficients:
              Value Std.Error  t-value p-value
(Intercept) 92.64952  1.5370879  60.27600    0
TMAX         5.32310  1.2006515   4.43351    0
Wetland      7.39775  0.9046259   8.17769    0
```

Statistická významnost. Obdobné jako v lineární regresi

Korelační matice

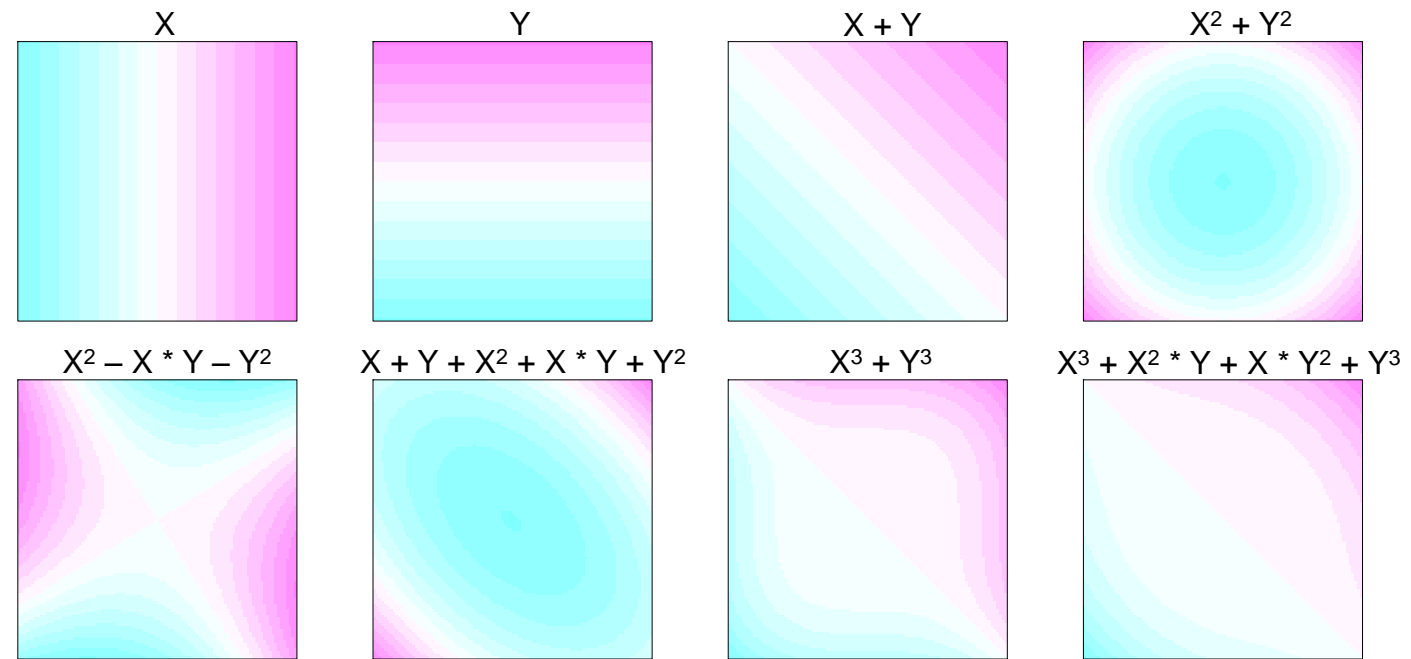
```
Correlation:
      (Intr) TMAX
TMAX   0.042
Wetland -0.006 -0.036
```

```
Standardized residuals:
      Min      Q1      Med      Q3      Max
-4.4829575 -0.5066319  0.1213412  0.7284539  2.7023428
```

```
Residual standard error: 20.18827
Degrees of freedom: 628 total; 625 residual
```

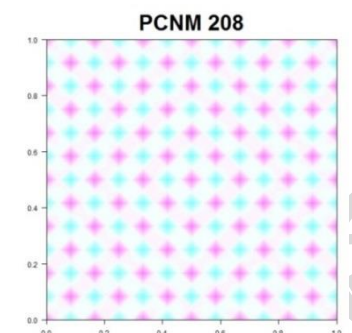
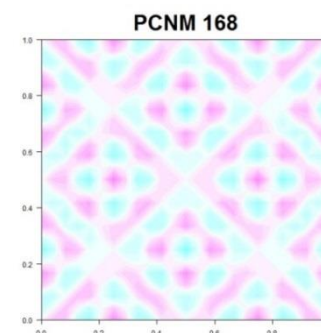
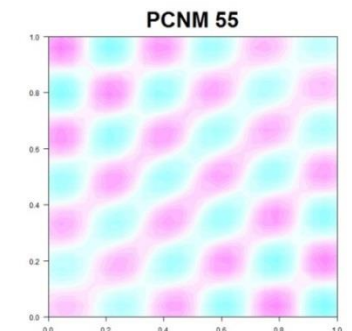
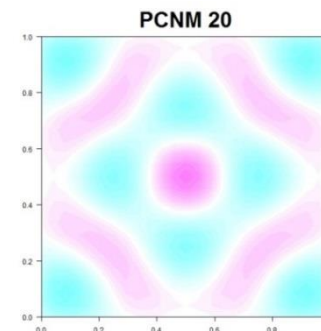
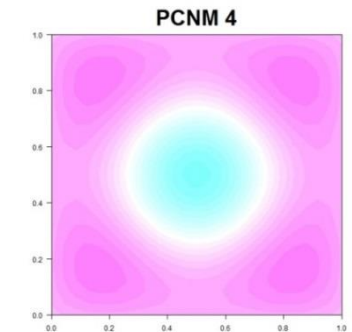
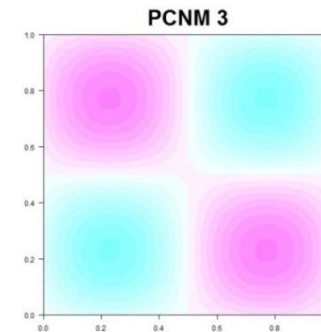
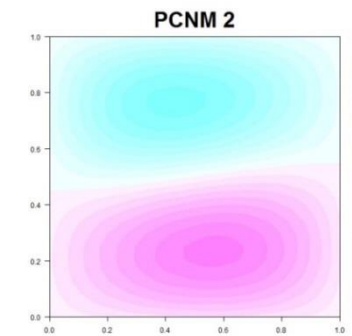
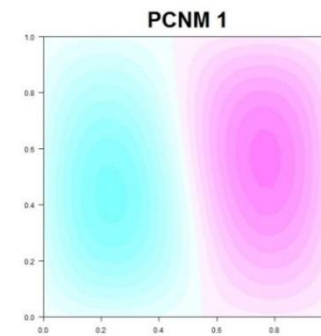
# Metody založené na „prostorových proměnných“

- Nejprve jsou vygenerovány tzv. prostorové proměnné (*spatial filters*)
  - Geografické souřadnice
  - Trend surface analysis
  - Moran's eigenvector maps (MEM)
- Prostorové proměnné mohou být následně použity jako vysvětlující proměnné v regresních nebo ordinačních metodách

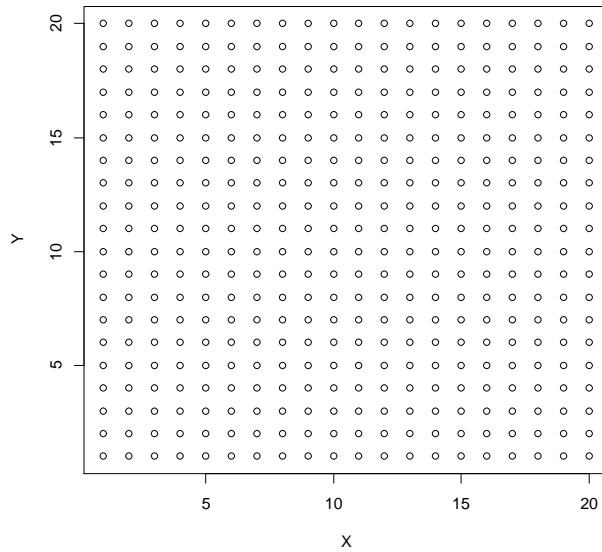


# Moran's eigenvector maps

- Soubor metod prostorového modelování (Borcard & Legendre 2002, Dray et al. 2006)
- Dříve Principal Coordinates of Neighbour Matrices (PCNM)
- Analýzou hlavních koordinát (PCoA) je vygenerována sada proměnných (PCoA os), jimiž lze modelovat prostorovou strukturu v datech
  - vysvětlující proměnné v regresních nebo vícerozměrných analýzách
- Počet os je roven počtu vzorků -1
- Každá osa má určitou hodnotu prostorové autokorelace, přičemž první osa má autokorelaci nejvyšší a poslední nejnižší



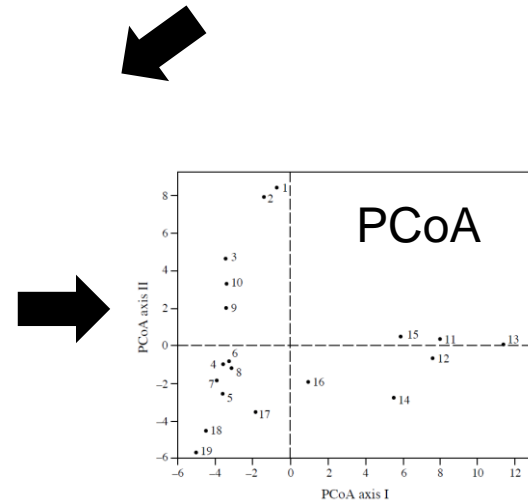
# Principal Coordinates of Neighbour Matrices (PCNM)



	X	Y
Samp1	1	1
Samp2	2	1
Samp3	3	1
Samp4	4	1
Samp5	5	1
Samp6	6	1
Samp7	7	1
Samp8	8	1

	Samp1	Samp2	Samp3	Samp4	Samp5	Samp6
Samp1	0	1	2	3	4	5
Samp2	1	0	1	2	3	4
Samp3	2	1	0	1	2	3
Samp4	3	2	1	0	1	2
Samp5	4	3	2	1	0	1
Samp6	5	4	3	2	1	0

	Samp1	Samp2	Samp3	Samp4	Samp5	Samp6
Samp1	0	1	4	4	4	4
Samp2	1	0	1	4	4	4
Samp3	4	1	0	1	4	4
Samp4	4	4	1	0	1	4
Samp5	4	4	4	1	0	1
Samp6	4	4	4	4	1	0



(pouze osy s pozitivními eigenvalue)

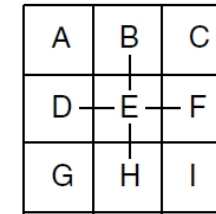
	PCNM1	PCNM2	PCNM3
Samp1	0.006031	-0.03556	0.050104
Samp2	0.010544	-0.06935	0.095756
Samp3	0.012178	-0.09971	0.1329
Samp4	0.00987	-0.12524	0.158235
Samp5	0.002973	-0.14487	0.16951
Samp6	-0.00867	-0.15794	0.165723
Samp7	-0.0247	-0.16423	0.147211
Samp8	-0.04424	-0.16397	0.115619

Vzdálenosti nad zvoleným thresholdem nahrazeny hodnotou  $4 \times \text{threshold}$ .

# Jaký „threshold“ zvolit?

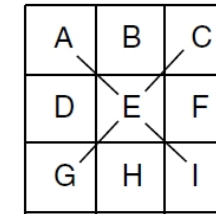
- Vzorke rozmístěné v pravidelné síti
  - Rook connection
  - Bishop connection
  - Queen connection
  - Kritérium vzdálenosti
- Vzorke rozmístěné nepravidelně (všechny vzorky by měly být propojené)
  - Delaunay triangulation
  - Gabriel criterion
  - Relative neighbourhood
  - Maximum distance
  - Minimum spanning tree (MST)

Rook



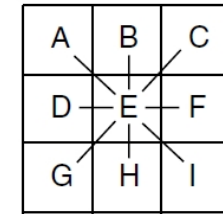
	A	B	C	D	E	F	G	H	I
A	-								
B	0	-							
C	0	0	-						
D	0	0	0	-					
E	0	1	0	1	-				
F	0	0	0	1	1	-			
G	0	0	0	0	0	0	-		
H	0	0	0	1	1	0	0	-	
I	0	0	0	0	0	0	0	0	-

Bishop



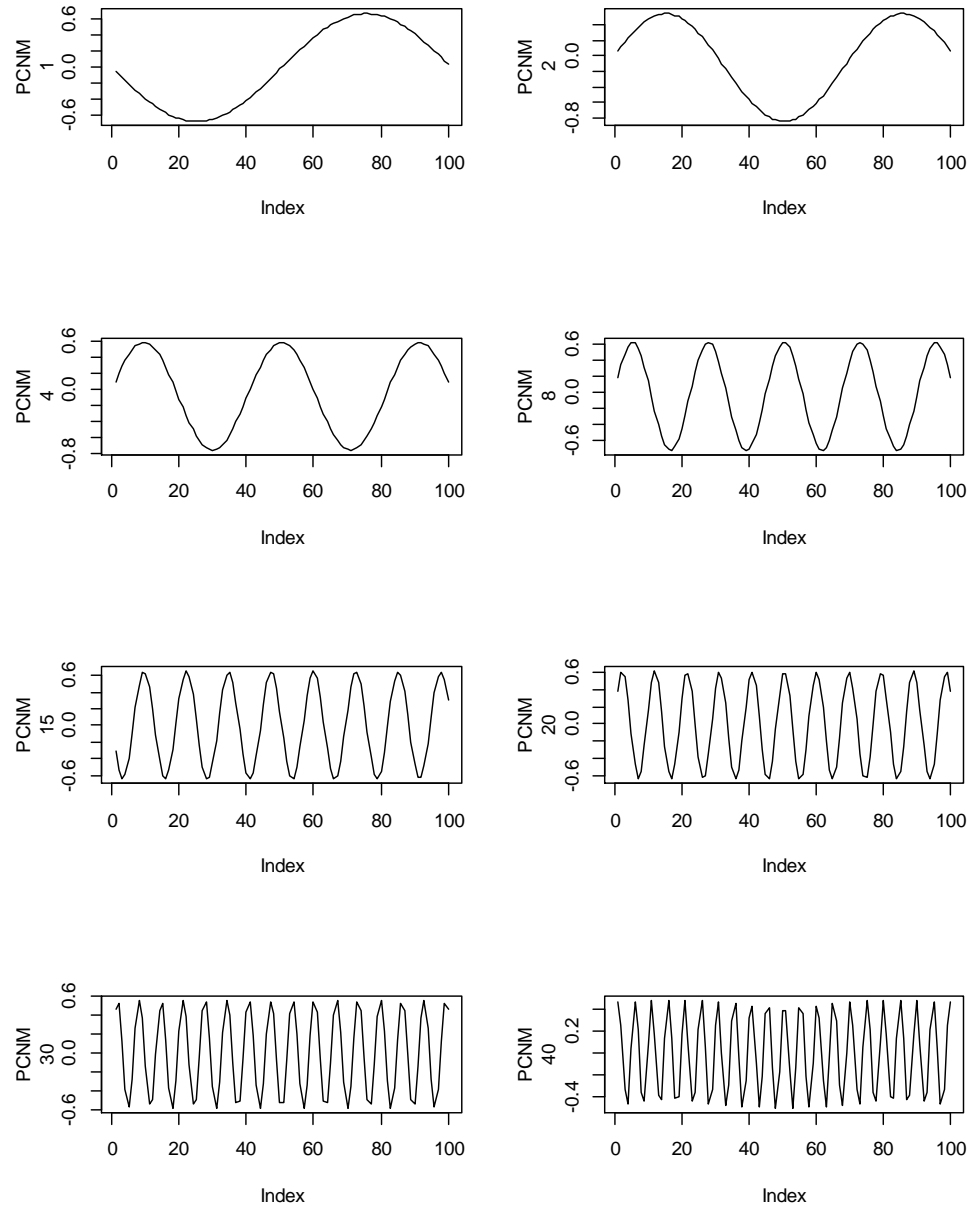
	A	B	C	D	E	F	G	H	I
A	-								
B	0	-							
C	0	0	-						
D	0	0	0	-					
E	1	0	1	0	-				
F	0	0	0	0	0	-			
G	0	0	0	0	1	0	-		
H	0	0	0	0	0	0	0	-	
I	0	0	0	0	1	0	0	0	-

Queen

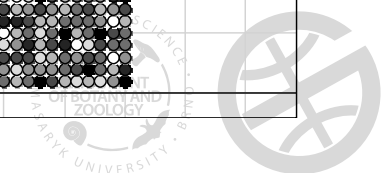
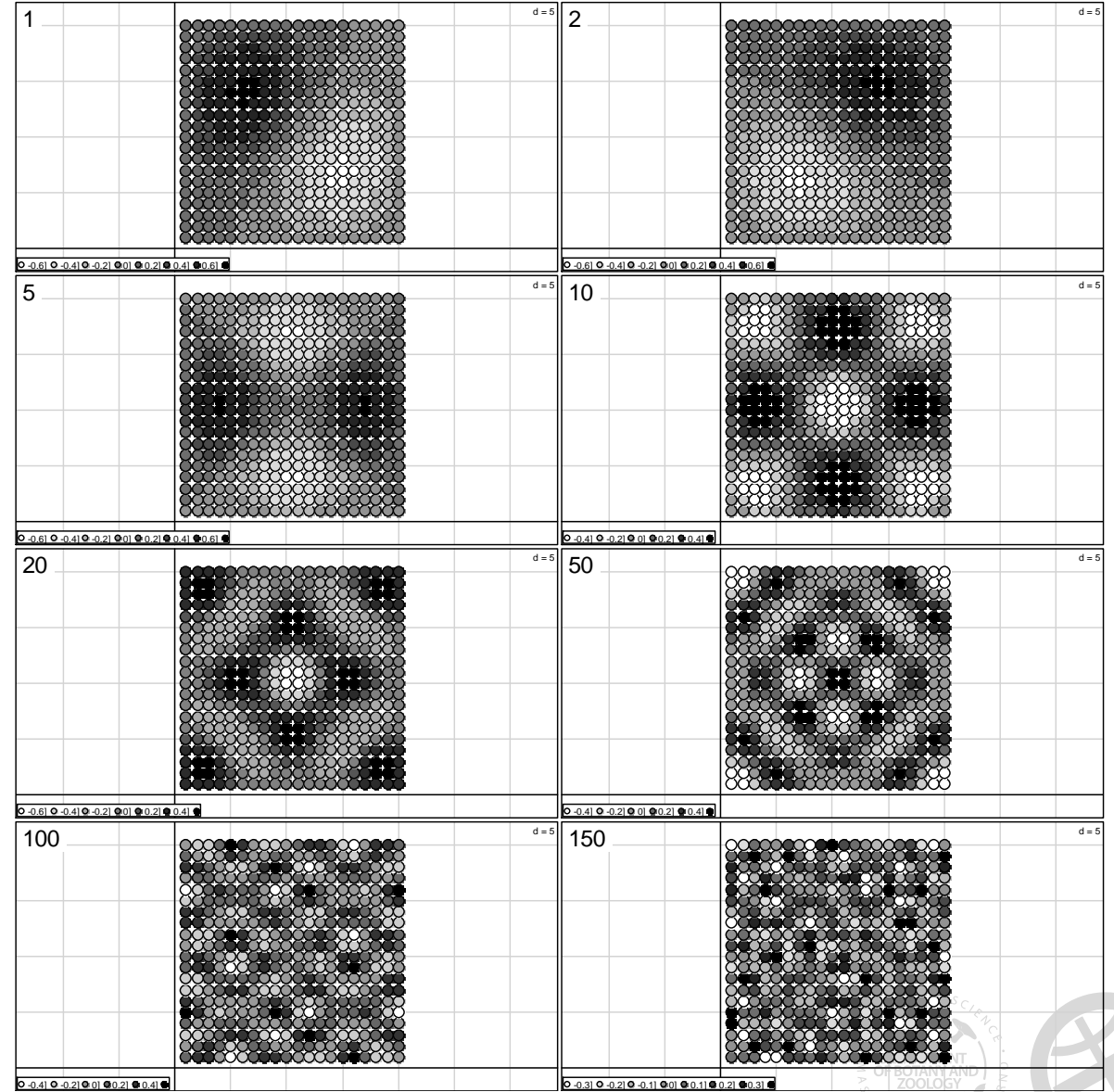


	A	B	C	D	E	F	G	H	I
A	-								
B	0	-							
C	0	0	-						
D	0	0	0	-					
E	1	1	1	1	-				
F	0	0	0	0	1	-			
G	0	0	0	0	1	0	-		
H	0	0	0	0	1	0	0	-	
I	0	0	0	0	1	0	0	0	-

### 1D

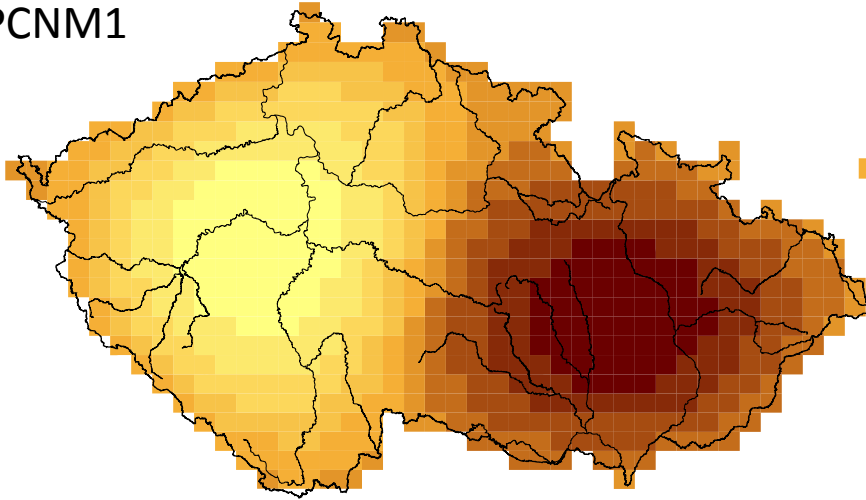


### 2D

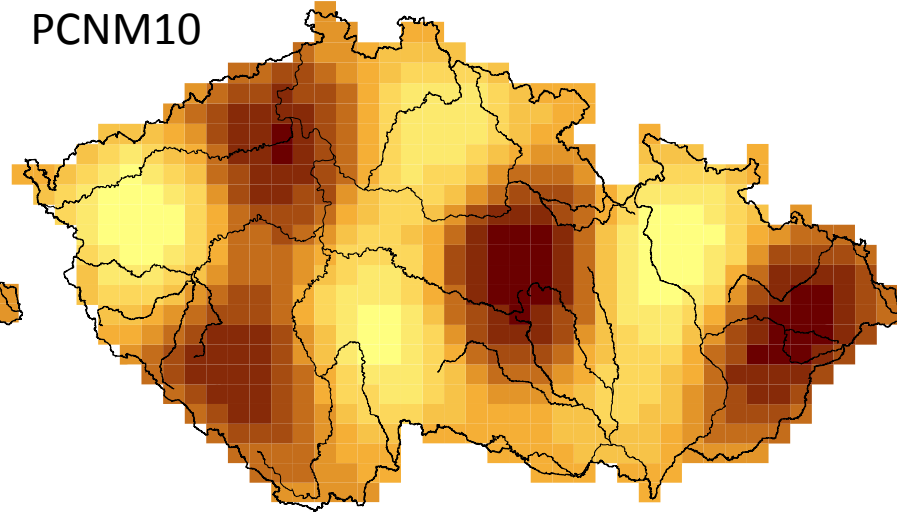


# PCNM osy pro mapovací kvadráty ČR

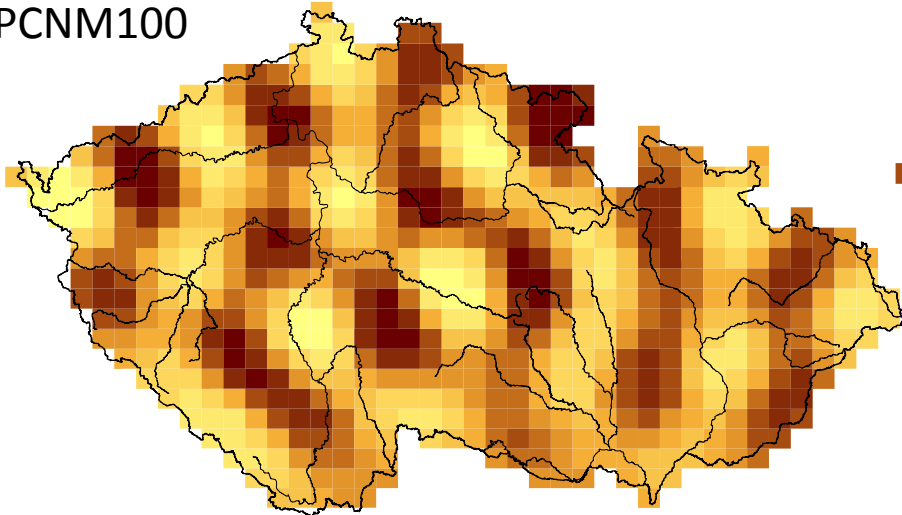
PCNM1



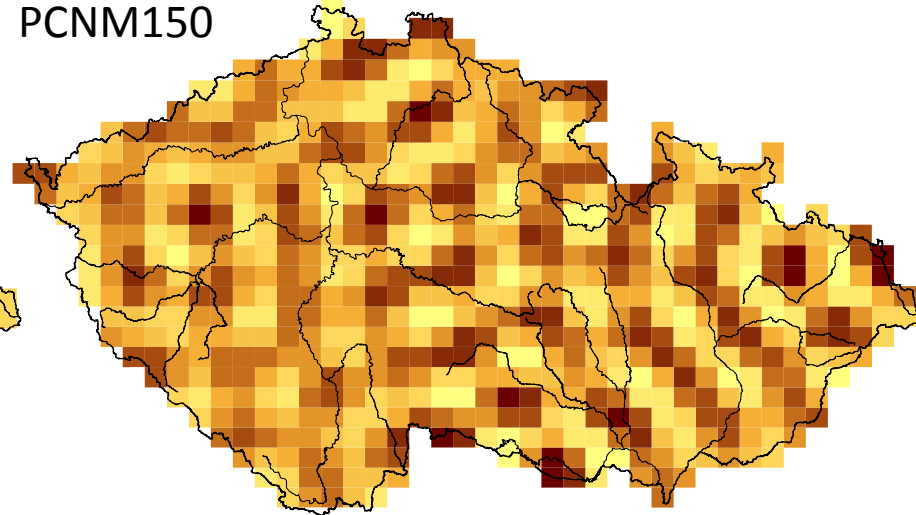
PCNM10



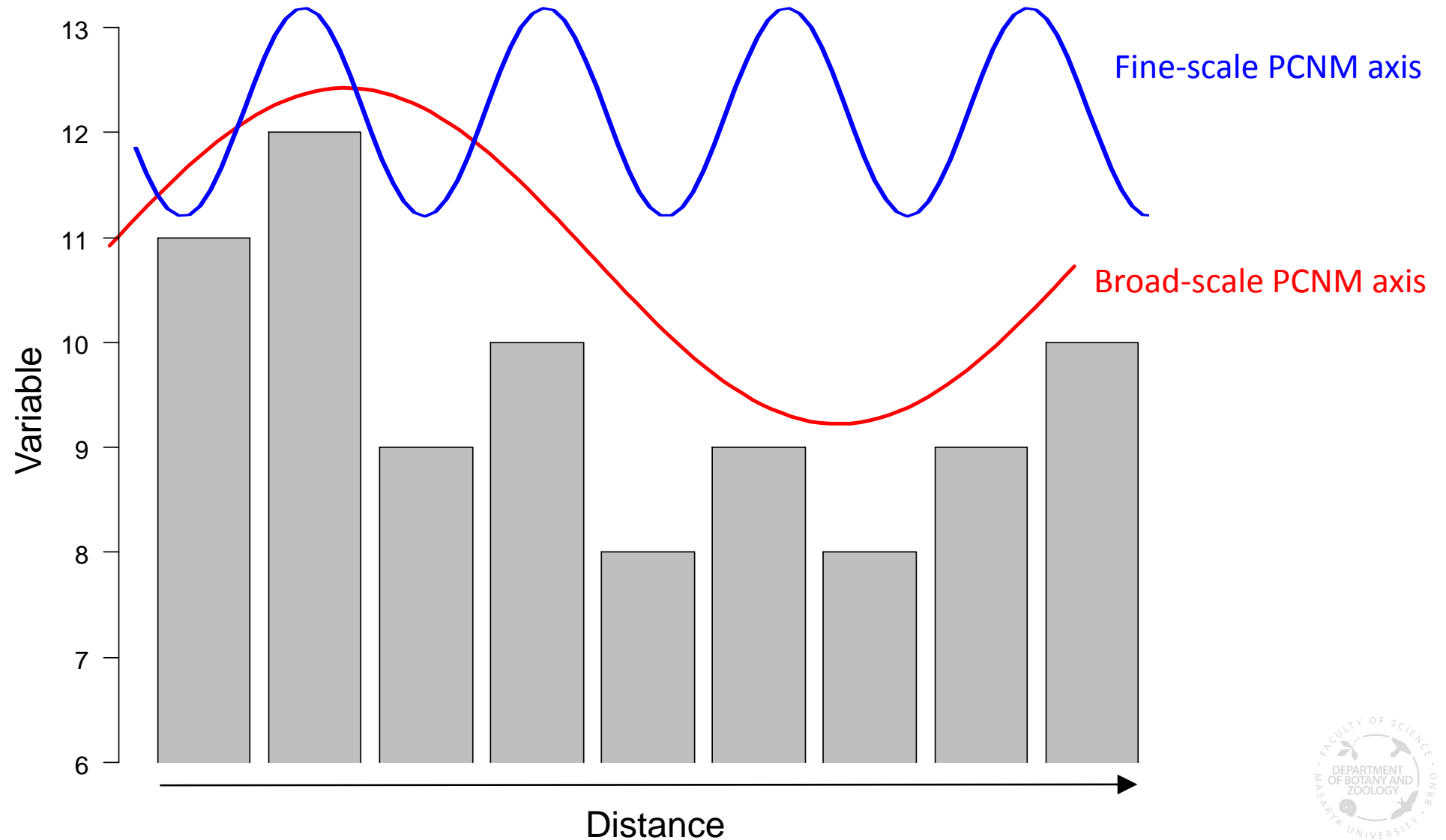
PCNM100



PCNM150

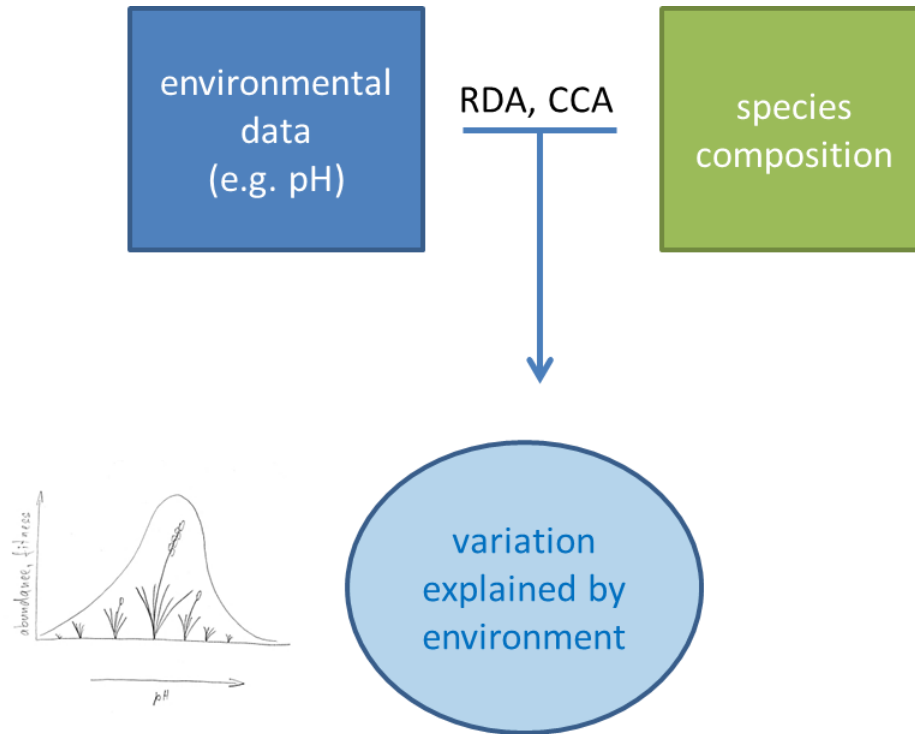


# Broad-scale & fine-scale variability

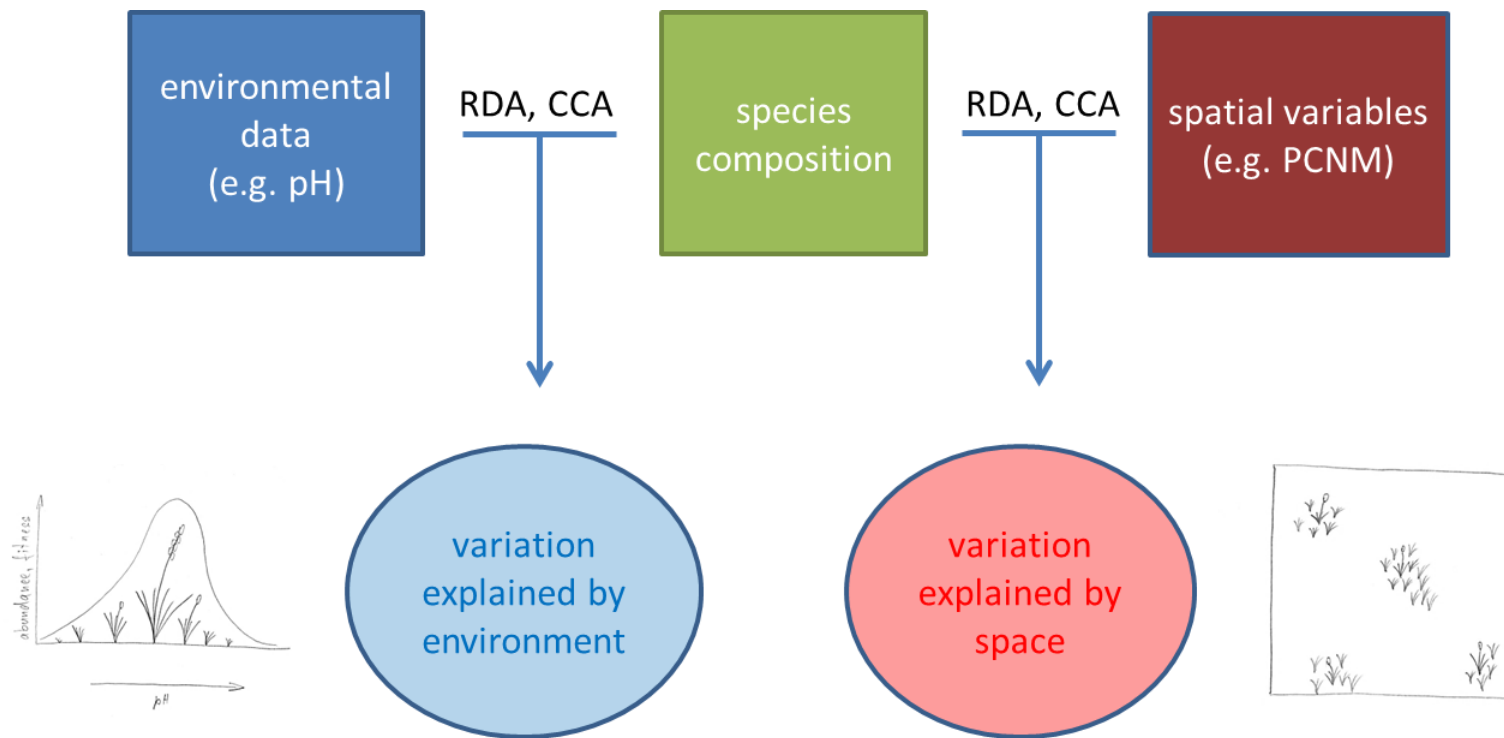




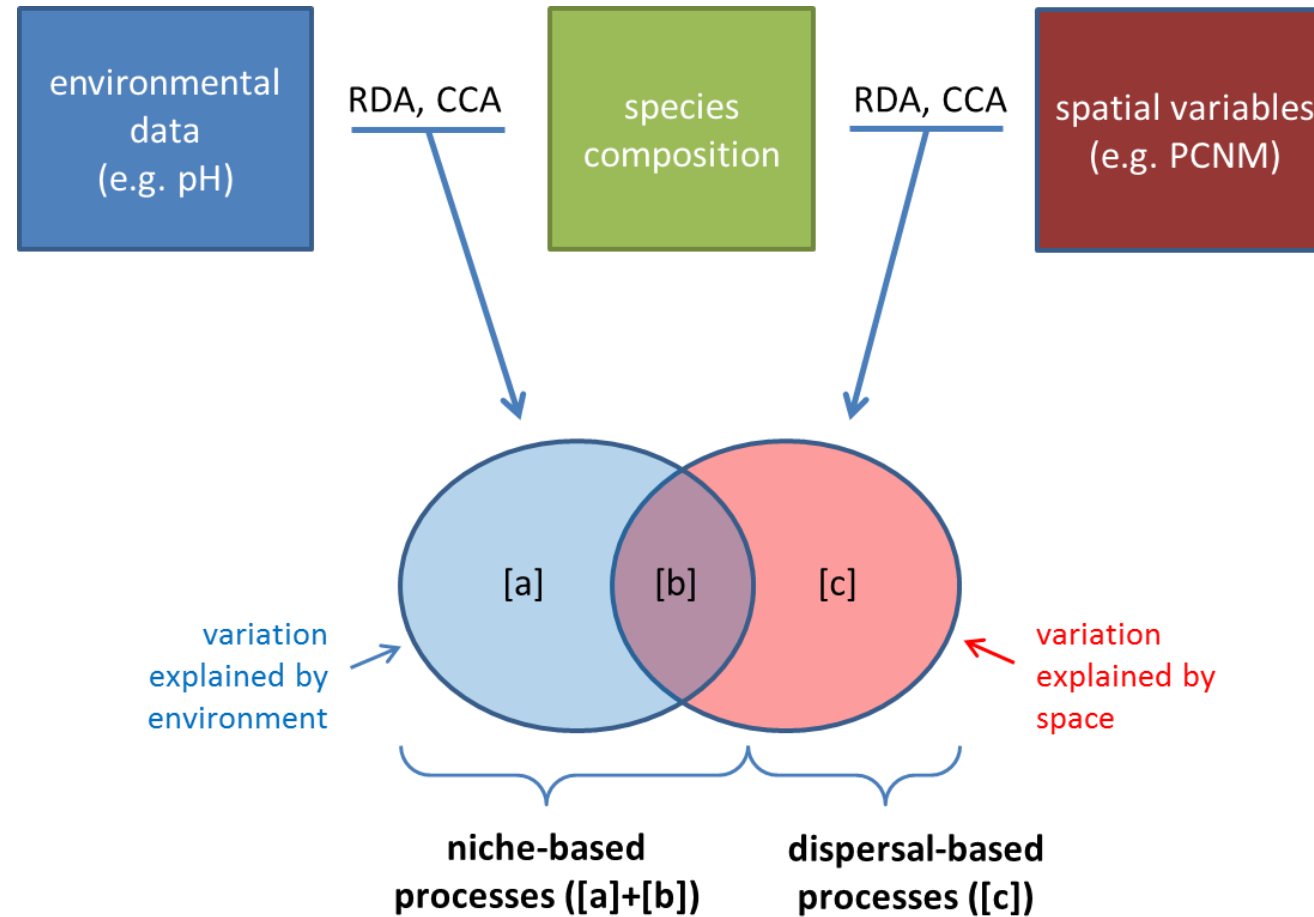
# Použití MEM analýzy



# Použití MEM analýzy



# Použití MEM analýzy



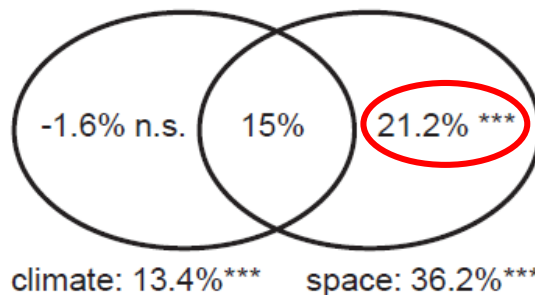


# Dispersal limitation is stronger in communities of microorganisms than macroorganisms across Central European cities

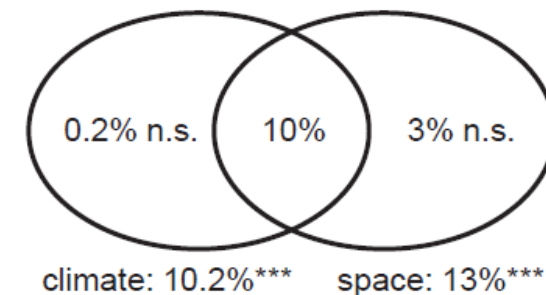
Milan Chytrý<sup>1</sup>, Zdeňka Lososová<sup>1,2\*</sup>, Michal Horský<sup>1</sup>, Bohuslav Uher<sup>1</sup>, Tomáš Čejka<sup>3</sup>, Jiří Danihelka<sup>1,4</sup>, Karel Fajmon<sup>1</sup>, Ondřej Hájek<sup>1</sup>, Lucie Juříčková<sup>5</sup>, Kateřina Kintrová<sup>1</sup>, Deana Láníková<sup>1,4</sup>, Zdenka Otýpková<sup>1</sup>, Vladimír Řehořek<sup>1</sup> and Lubomír Tichý<sup>1</sup>



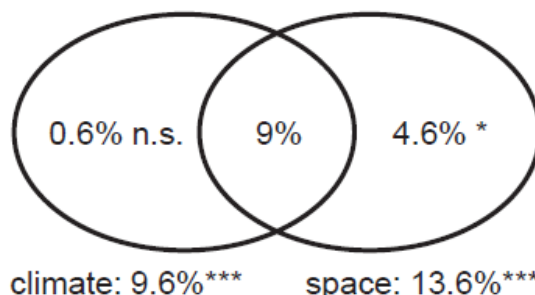
(a) cyanobacteria/algae: 34.6%



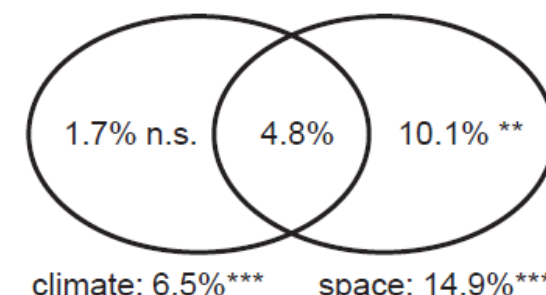
(d) herbs: 13.2%



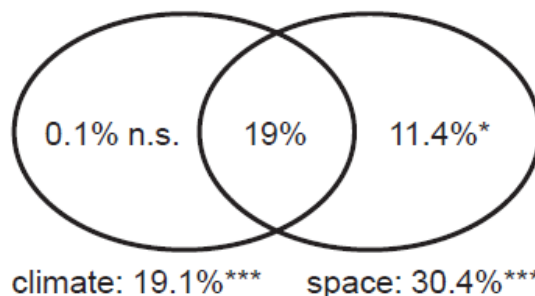
(b) plants: 14.2%



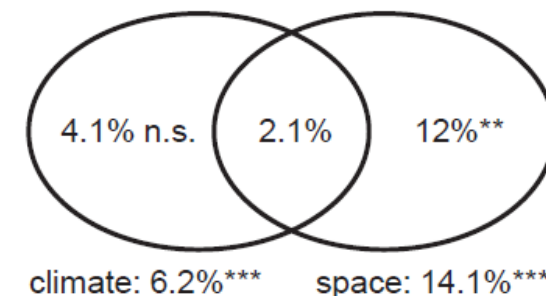
(e) animal-dispersed trees and shrubs: 16.6%



(c) snails: 30.5%



(f) wind-dispersed trees and shrubs: 18.2%



# Výhody a nevýhody MEM analýzy

## Výhody

1. Velmi efektivní a flexibilní
  - dokáže modelovat jakékoliv prostorové struktury
2. Použití v regresních i ordinačních metodách

## Nevýhody

1. Má tendenci zachytávat „šum“ v datech (Gilbert & Bennett, 2010)
  - variabilita vysvětlená prostorovými proměnnými je často nadhodnocená → nutné pečlivě vybírat jednotlivé eigenvektory
2. Při variation partitioning „podhodnocuje“ čistý vliv prostředí (Smith & Lundholm, 2010)
  - Interpretace by měla být založena na porovnávání frakcí [a] + [b] s frakcí [c] (nikoliv [a] vs [c])
3. Variabilita zachycená prostorovými proměnnými může být ve skutečnosti způsobena neměřenou proměnnou (Dray et al., 2012)



# Výběr MEM eigenvektorů

## (A) Select MEM eigenvectors to describe all spatial variation in species data

Variation in composition of animal assemblages (e.g. mammals)



=

Variation explained by selected MEM eigenvectors (spatially structured variation)



+

Residual variation which has not spatial structure



## (B) Check for spatial autocorrelation in residuals of the environmental model

Variation in composition of animal assemblages (e.g. mammals)



=

Variation explained by environmental variables (environmental model)



+

Residual variation with significant spatial autocorrelation



*Hatched part denotes spatially structured variation explained by environment*

*Hatched part denotes spatially structured residual variation*

## (C) Generate the set of spatial filters

Spatial filters representing spatial variation in species data independent on environment (Independent Spatial Variation)



=

Selected MEM eigenvectors representing spatially structured variation in species data



Environmental variables representing environmental variation in species data



-

## (D) Test environmental model

Variation in composition of animal assemblages (e.g. mammals)



=



+



+



Environmental model corrected for residual spatial autocorrelation (ISV used as covariables)

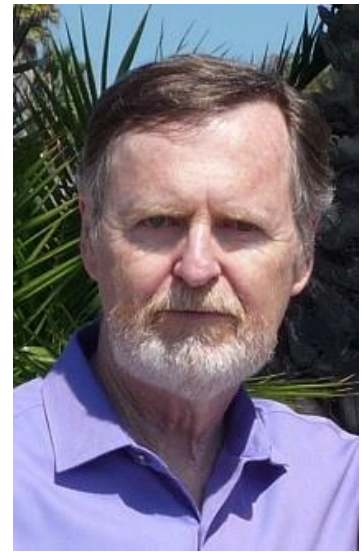




GUEST  
EDITORIAL

## Eight (and a half) deadly sins of spatial analysis

Bradford A. Hawkins



1. Spatial autocorrelation generates bias
2. Spatial regression is best
3. The world is stationary
4. Partial regression coefficients mean something
5. Regression coefficients identify effects
6. Species richness generates bias
7. The Earth is round ( $P < 0.05$ )
8. Spatial processes explain spatial patterns
- 8.5 Spatial autocorrelation causes red shifts in regression models



# Literatura

- Legendre, P. & Legendre, L. (2012): Numerical ecology. Third Edition. Elsevier, Amsterdam.
- Borcard, D., Gillet, F. & Legendre, P. (2011): Numerical ecology with R. Springer, New York.
- Fortin, M.-J. & Dale, M.R.T. (2005): Spatial analysis: a guide for ecologists. Cambridge University Press. New York.
- Borcard, D. & Legendre, P. (2002): All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* 153: 51-68.
- Dray, S., Legendre, P. & Peres-Neto, P.R. (2006): Spatial modelling: a comprehensive Framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* 196: 483-493.
- Clifford, P., Richardson, S. & Hémon, D. (1989): Assessing the significance of the correlation between two spatial processes. *Biometrics* 45: 123-134.
- Dutilleul, P. (1993): Modifying the t test for assessing the correlation between two spatial processes. *Biometrics* 49: 305-314.
- Gilbert, B. & Bennett, J.R. (2010): Partitioning variation in ecological communities: do the numbers add up? *Journal of Applied Ecology* 47: 1071-1082.
- Smith, T.W. & Lundholm, J.T. (2010): Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography* 33: 648-655.
- Dray, S., Pelissier, R., Couteron, P., Fortin, M.-J., Legendre, P., Peres-Neto, P.R., Bellier, E., Bivand, R., Blanchet, F.G., De Caceres, M., Dufour, A.-B., Heegaard, E., Jombart, T., Munoz, F., Oksanen, J., Thioulouse, J. & Wagner, H.H., 2012. Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs* 82: 257-275.

