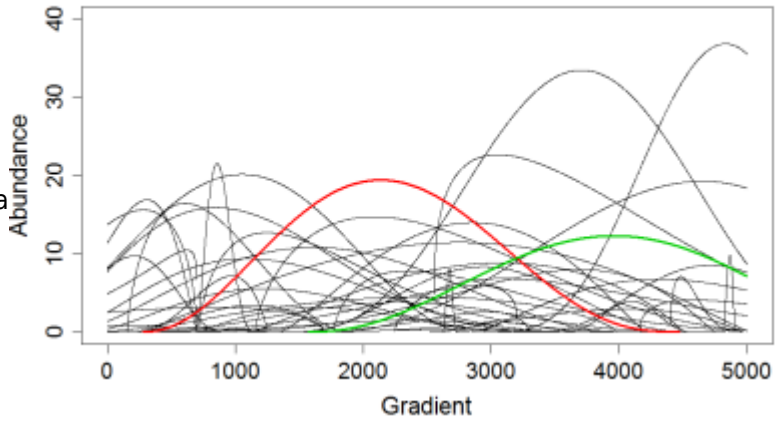
Simulovaná ekologická data

Zdroj dat

Zelený D. (nepubl.), upravený skript založený na simulačním modelu, který napsal Fridley et al. (2007) (viz Appendix S2 tohoto článku)

Popis datového souboru



Základem simulace těchto ekologických dat je model společenstva, který plně odpovídá teorii ekologické niky. Podél jednoho, případně dvou virtuálních gradientů jsou náhodně rozmístěny unimodální odpovědní křivky druhů (které odpovídají pravděpodobnosti výskytu druhu v dané části gradientu a jsou modelovány pomocí tzv. Beta funkce, viz obrázek), přičemž každý druh je definován svým optimem na daném gradientu, šířkou niky a několika dalšími parametry odpovědní křivky. V dalším kroku jsou v rámci gradientu náhodně vygenerovány lokality pro jednotlivá pozorování, a v rámci každého pozorování jsou 'sbírány' jednotlivé druhy následovně: nejdříve se vygeneruje náhodný počet jedinců, který má být pro dané pozorování vybrán, a následně je každý jedinec náhodně přiřazen k některému druhu, přičemž pravděpodobnost přiřazení jedince k druhu je vážena pravděpodobností výskytu jednotlivých druhů v dané části gradientu. Jeden druh může tedy být v daném pozorování zastoupen několika jedinci, pokud v dané pozici gradientu má narozdíl od ostatních vysokou pravděpodobnost výskytu. V případě dvou gradientů je pravděpodobnost výskytu druhu pro danou kombinaci hodnot jednotlivých gradientů definována jako násobek pravděpodobností pro každý gradient zvlášť. Podrobnosti viz přiložené skripty.

Parametry jednotlivých souborů

- simul1 - 1 gradient (délka 5000 jednotek), 500 vzorků, 300 druhů[1)]
- simul2 - 2 nestejně dlouhé gradienty (5000 a 2000 jednotek), 500 vzorků, 300 druhů
- simul3 - 2 stejně dlouhé gradienty (5000 jednotek), 500 vzorků, 300 druhů
- simul.short - 2 nestejně dlouhé gradienty, které jsou relativně krátké (1100 a 800 jednotek), 70 vzorků, 300 druhů, vzorky rovnoměrně rozmístěny podél gradientu v odstupech 100 jednotek
- simul.long - 2 nestejně dlouhé gradienty, které jsou relativně dlouhé (5500 a 4000 jednotek), 70 vzorků, 300 druhů, vzorky rovnoměrně rozmístěny podél gradientu v odstupech 100 jednotek

Pozn.: jde o parametry simulovaného modelu - počty druhů ve výsledné matici nemusejí odpovídat (a také neodpovídají) počtu simulovaných druhů, protože některé druhy s nízkou abundancí nebyly do výsledné matice vybrány.

Data ke stažení

Soubory s příponou -spe obsahují presenčně absenční matici vzorky-druhy (vzorky jsou v řádcích, druhy ve sloupcích), soubory s příponou -env obsahují hodnotu simulovaného gradientu pro patřičný vzorek (analogie měřené proměnné), soubory s příponou -specvalues obsahují údaje o pozici optima druhu podél gradientu (*optimum*) a šířce niky (*niche.width*, v arbitrárních jednotkách gradientu). Data simulovaná podél dvou gradientů mají pro každý druh dva údaje o optimu a šířce niky, každý vztahující se k jednomu gradientu.

- simul1-spe.csv
- simul1-env.csv
- simul1-specvalues.csv
- simul2-spe.csv
- simul2-env.csv
- simul2-specvalues.csv
- simul3-spe.csv
- simul3-env.csv
- simul3-specvalues.csv
- simul.short-spe.csv
- simul.short-env.csv
- simul.long-spe.csv
- simul.long-env.csv

Skript pro přímé načtení dat do eRka

```
simul1.spe <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul1-spe.txt', row.names = 1)
simul1.env <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul1-env.txt', row.names = 1)
simul1.specvalues <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul1-specvalues.txt',
row.names = 1)

simul2.spe <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul2-spe.txt', row.names = 1)
simul2.env <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul2-env.txt', row.names = 1)
simul2.specvalues <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul2-specvalues.txt',
row.names = 1)

simul3.spe <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul3-spe.txt', row.names = 1)
simul3.env <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul3-env.txt', row.names = 1)
simul3.specvalues <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul3-specvalues.txt',
row.names = 1)

simul.short.spe <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul.short-spe.txt', row.names
= 1)
simul.short.env <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul.short-env.txt', row.names
```

```
= 1)

simul.long.spe <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul.long-spe.txt', row.names
= 1)
simul.long.env <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/simul.long-env.txt', row.names
= 1)
```

Skripty pro vytvoření simulovaných dat

- simulovaná data podél jednoho gradientu
- simulovaná data podél dvou gradientů

Literatura

Fridley, J.D., Vandermast, D.B., Kuppinger, D.M., Manthey, M., and Peet, R.K. (2007) Co-occurrence-based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. *Journal of Ecology* 95: 707-722 pdf Appendix S2

1)
Počet druhů ve výsledném souboru je mírně nižší - 300 druhů bylo použito pro simulované společenstvo, ale ne všechny druhy se při náhodném výběru dostaly do výsledného souboru dat.

From:
https://anadat-r.davidzeleny.net/ - **Analysis of community ecology data in R**

Permanent link:
https://anadat-r.davidzeleny.net/doku.php/cs:data:simul

Last update: **2017/10/11 20:36**