Studijní materiály

Základní studijní materiály

Následují odkazy na dvoje česky psaná skripta, která se přímo týkají probírané látky a mohou sloužit (vedle prezentace z přednášek) jako učební materiál ke zkoušce.

- Herben & Münzbergová (2003) Zpracování gebotanických dat v příkladech. Část 1. Data o druhovém složení
- Lepš & Šmilauer (2001) Mnohorozměrná analýza ekologických dat

Rozšiřující učební materiály k jednotlivým tématům

Kapitoly z různých učebnic, vesměs anglicky psaných, ve kterých najdete podrobnosti k některým probíraným tématům. Jde o rozšiřující materiály, které jdou nad rámec této přednášky a jejich znalost nebudu při zkoušce vyžadovat.

Typy sbíraných dat, příprava dat pro analýzy

- Gotelli & Ellison (2004) A Primer of Ecological Statistics. Chapter 8: Managing and curating data

Ekologická podobnost

- Wildi (2010) Data Analysis in Vegetation Ecology. Chapter 4: Multivariate comparison.

Design ekologických experimentů

- Lepš & Šmilauer (2003) Multivariate Analysis of Ecological Data using CANOCO. Chapter 2: Experimental design
- Gotelli & Ellison (2004) A Primer of Ecological Statistics. Chapter 6: Designing successful field studies. Chapter 7: A bestiary of experimental and sampling design

Odborné články k jednotlivým tématům

Odkazy na odborné články a jiné materiály, které se úzce zaměřují na některá probíraná témata. Jde o články, které mě nějakým způsobem zaujaly, a třeba zaujmou i vás.

Typy sbíraných dat, příprava dat pro analýzy

O'Hara R.B. & Kotze D.J. (2010): Do not log-transform count data. *Methods in Ecology and Evolution*, 1: 118-122.

- Autoři studie varují před bežně zavedenou praxí logaritmické transformace dat představujících počty (jedinců, druhů atd.), a na simulovaných datech ukazují, že zobecněné lineární modely jsou pro analýzu takovýchto dat daleko účinnější.

Wilson J.B. (2012): Species presence/absence sometimes represents a plant community as well as species

abundances do, or better. *Journal of Vegetation Science*, 23: 1013-1023.

- Bastow Wilson argumentuje, že v některých případech vypovídají presenčně/absenční data o vztahu společenstva k faktorům prostředí více než data zahrnující abundance druhů. Týká se to především heterogenních datových souborů. Zajímavý pohled na trochu kontroverzní téma, kolem kterého nepanuje shoda.

Explorační vs. konfirmační analýza

Tukey J.W. (1980): We need both exploratory and confirmatory. *The American Statistician*, 1:23-25.

- John Tukey je považován za zakladetele explorační analýzy (*exploratory data analysis, EDA*), jejíž podstatou je generování hypotéz založených na analýze (exploraci) dat. Ta bývá někdy vnímána jako protiklad analýzy konfirmační (*confirmatory data analysis, CDA*), která je naopak založená na testování předem známých hypotéz pomocí nově sebraných dat. V tomto krátkém článku Tukey shrnuje svůj postoj k této dichotomii a zdůrazňuje, že důležité jsou oba způsoby vědeckého myšlení.

Rosling H. (2010): The Joy of Stats. BBC documentary video.

- Hans Rosling je švédský lékař a statistik, který se proslavil především svou neotřelou vizualizací demografických dat. V tomto dokumentu vytvořeném pro britskou televizi BBC vysvětluje, proč je explorační analýza dat v našem životě důležitá a jak ji prakticky provádět. Inspirující byla už jeho první přednáška v rámci projektu TED (Ideas worth spreading, www.ted.com), kde poprvé ukázal vizualizaci časových trendů pomocí software, kterému říká Gapminder (přednášku najdete zde, ukázku software zde). Hans Rosling je také polykač mečů, což demonstroval na závěr jeho druhého vystoupení v rámci TED (viz 17. minuta).

Ekologická podobnost

Bahns A.J., Pickett K.M. & Crandall Ch.C. (2011): Social ecology of similarity: Big schools, small schools and social relationships. *Group Processes & Intergroup Relations*

- Studie, která se netýká přímo ekologické podobnosti, ale podobnosti mezi přáteli. Vědci se zaměřili na studenty na různě velkých univerzitách a zjišťovali, jaký je rozdíl mezi přáteli, pokud se setkají na velké nebo na malé univerzitě. Zjistili, že na velkých univerzitách si studenti hledají přátele, kteří jsou jim více podobní v názorech, zatímco na malých univerzitách jsou si přátelé spíše nepodobní. Zároveň ale platí, že přátelství podobných na velkých univerzitách bývá povrchnější a krátkodobější než přátelství rozdílných na malých univerzitách. Na velkých univerzitách je velký “species pool” přátel, je jednodušší najít si někoho jiného, zatímco na malých univerzitách je výběr omezený a člověk spíš investuje do přátelství s někým, kdo mu není až tak podobný. Více podrobností například v tomto čtenářském výcucu (*readers's digest*): http://www.scientificamerican.com/article.cfm?id=a-friend-like-me.

Ordinační analýza

Økland H. (1996): Are ordination and constrained ordination alternative or complementary strategies in general ecological studies? *Journal of Vegetation Science*, 7:289-292.

- Autor zmiňuje základní rozdíly, jak chápat a používat nepřímou a přímou ordinační analýzu (o nepřímé mluví jako o *ordination*, přímou označuje jako *constrained ordination*).

Wartenberg D., Ferson S. & Rohlf F.J. (1987): Putting things in order: a critique of detrended correspondence analysis. *Americal Naturalists*, 129: 434-448.

- Autoři kritizují DCA, která podle nich ve skutečnosti neřeší problém s obloukem coby artefaktem v

ordinačním diagramu, a doporučují DCA vůbec nepoužívat.

Peet R.K., Knox R.G., Case J.S & Allen R. B. (1988): Putting things in order: the advantages of detrended correspondence analysis. *American Naturalists*, 131: 924-934.

- Shrnutí výhod a nevýhod detrendované korespondenční analýzy, jedné z v té době nejpoužívanějších metod, které "dávají věci do pořádku". Reakce na předchozí článek Wartenberg et al. (1987).

Palmer M.W. (1993): Putting things in even better order: the advantages of canonical correspondence analysis. *Ecology*, 8: 2215-2230.

- Mike Palmer za pomoci simulovaných dat ukazuje, že CCA je ještě lepší metodou než DCA. To bylo v době, kdy se CCA běžně používala pro deskriptivní účely stejně jako nepřímé ordinace (i když CCA je ordinací přímou).

Ukázky článků, které používají různé typy ordinačních metod

Nobis M. & Wohlgemuth T. (2004): Trend words in ecological core journals over the last 25 years (1978-2002). *Oikos*, 106: 411-420.

- Článek nezpracovává ekologická data, ale data o trendech v názvech článků v ekologických časopisech. Autoři vytvořili databázi názvů článků v hlavních ekologických časopisech za 25 let a pomocí ordinačních analýz (PCA, CA) a klastrové analýzy ukazují na vývoj v trendech používání různých ekologických termínů.

Chytrý M., Grulich V., Tichý L. & Kouřil M. (1999): Phytogeographical boundary between the Pannonicum and Hercynicum: a multivariate analysis of landscape in the Podyjí/Thayatal National Park, Czech Republic, Austria. *Preslia*, 71:1-19.

- Data pocházejí z floristického mapování ve čtvercové síti, která pokrývá celý NP Podyjí/Thayatal. Pomocí ordinační metody DCA autoři hledali hlavní směry ve variabilitě druhových dat v krajině a doložili existenci hranice dvou fytogeografických okresů.

Vojta J. (2007): Relative importance of historical and natural factors influencing vegetation of secondary forests in abandoned villages. *Preslia*, 79: 223-244.

- Studie srovnává druhové složení lesů, které vznikly na místech opuštěných vesnic ve vojenském výcvikovém prostoru Doupovské hory, s druhovým složením lesů přirozených, a sledoval, které druhy jsou indikátory sekundárních lesů. Pro srovnání sekundárních a původních lesů autor použil přímou ordinační analýzu RDA.

Horsák M., Hájek M., Tichý L. & Juřičková L. (2005): Plant indicator values as a tool for land mollusc autecology assesment. *Acta Oecologica*, 32: 161-171.

- Mohou Ellenbergovy indikační hodnoty pro rostliny vysvětlit druhovou variabilitu ve složení malakofauny? Autoři se snaží na tuto otázku odpovědět pomocí dat z ploch, na kterých byla současně sbírána vegetační a malakologická data. Vegetační data byla použita pro výpočet průměrných Ellenbergových indikačních hodnot pro několik ekologických faktorů a tyto faktory byly použity pro zjištění ekologických optim vybraných druhů měkkýšů.

Numerická klasifikace

Roleček J., Tichý L., Zelený D. & Chytrý M. (2009): Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. *Journal of Vegetation Science*, 20:596-602.

- Článek, který detailně popisuje, jak pracuje modifikovaný TWINSPAN. Mimo jiné je zde rozebráno, jak se metoda chová při použití různých indexů betadiverzity, kterými se hodnotí heterogenita jednotlivých

skhluků - na základě této heterogenity se metoda rozhoduje, které shluky bude dělit jako první. Pokud plánujete modifikovaný TWINSPAN použít, měli byste si tento článek přečíst.

P hodnoty a testování signifikance

Murtaugh P.A. (2014) In defence of P values. *Ecology* 95:611-617 (a další články diskutující téma P hodnot)

- Aktuální diskuze na stránkách časopisu *Ecology* o problematice používání P hodnot v analýze a interpretacích ekologických studií. Článek, který napsal Paul Murtaugh, rozhodně stojí za přečtení, protože shrnuje běžná nedorozumění, kterých se uživatelé P hodnot (míry signifikance) často dopouštějí, a která vedou k nedocenění významu vlastních P hodnot. Murtaugh zdůrazňuje, že P hodnoty by měly být vnímány jako důveryhodnost důkazu, který máme proti platnosti nulové hypotézy, a ne jako kritérium pro binární rozhodnutí *signifikantní* vs *nesignifikantní*.

"Fourth-corner problem" a RLQ analýza

Dray S., Choler P., Dolédec S., Peres-Neto P.R., Thuiller W., Pavoine S. & ter Braak C.J.F. (2014): Combining the fourth-corner and the RLQ methods for assessing trait responses to environmental variation. *Ecology*, 95: 14-21.

- Krátký článek ilustrující použití dvou metod pro analýzu druhových vlastností společně s proměnnými prostředí, na příkladu vegetačních dat z vysohoroských sněhových výležisek a funkčních vlastnostech rostlinných druhů ve vztahu k délce sněhové pokrývky a dalším faktorům prostředí. Autoři ukazují jak prakticky spočítat fourth-corner a RLQ analýzu, jak výsledky interpretovat a jak je zkombinovat. Jde o asi nejnovější studii demonstrující a diskutující použití těchto dvou metod.

Ellenbergovy indikační hodnoty

Zelený D. (2012): Poznámky k používání průměrných Ellenbergových indikačních hodnot při analýze vegetačních dat. *Zprávy České Botanické Společnosti*, 47: 159-178.

- K průměrným Ellenbergovým indikačním hodnotám je při analýzách třeba přistupovat jinak než k měřeným faktorům prostředí, protože tyto hodnoty jsou přímo odvozeny z druhového složení vegetace. Problém se objeví v okamžiku, kdy průměrné EIH analyzujeme společně s vegetačními daty, ze kterých jsou odvozeny - příkladem může být pasivní proložení EIH do ordinačního diagramu, nebo ANOVA Ellenbergových hodnot mezi vegetačními typy. Článek shrnuje možnosti, jak tento problém ošetřit (více na webových stránkách).

Alfa, beta a gama diversita

Jurasinski G., Retzer V. & Beierkuhnlein C. (2009): Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. *Oecologia* 1:15-26.

- Poznámky k modernímu použití termínů alfa, beta a gama diverzita. Podobných článků se poslední dobou objevilo několik, tento je z nich nejkratší...

Gotelli N.J & Colwell R.K. (2011) Estimating species richness. In: A.E. Magurran & B.J. Mc Gill, *Biological diversity: frontiers in measurement and assessment*. Oxford University Press, Oxford, UK.

- Tento učebnicový text mimo jiné zahrnuje i detailní vysvětlení konceptu akumulačních a rarefakčních křivek a jejich aplikaci.

Spellerberg I.F. & Fedor P.J. (2003): A tribute to Claude Shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon-Wiener' Index. *Global Ecology & Biogeography* 12:

177-179.

- Co je správně, Shannon-Wiener nebo Shannon-Wiever? Claude Shannon (1916-2001) publikoval svůj index H' v roce v roce 1948 v časopise *Bell System Technical Journal*. V tomto článku Shannon píše, že „filozofie a teorie informační teorie vděčí mnohé matematikovi Norbertu Wienerovi", a informační teorie je zde zmíněna, protože na ní je Shannonův index postaven. Shannon zde také cituje několik Wienerových publikací. Původní Shannonův článek byl později (1949) téměř doslovně otištěn v knize, která se jmenuje *The mathematical theory of communication*, vyšla v roce 1949 a jejími autory jsou Shannon a matematik Warren Weaver. Kniha sestává ze dvou kapitol, první napsané Shannonem (právě přetisk onoho článku), a druhé napsané Warrenem Weaverem s názvem *Recent contributions to the mathematical theory of communication* (tato kapitola nebyla před tím publikována, ale kondenzovaná verze se objevila v časopise *Scientific American* v červenci 1949). Tato kniha se stala nejčastější citací pro Shannonův index, a odsud pochází to nedorozumění. Z důvodů, že Shannon navazuje na práci Norberta Wienera v informační teorii, je Shannonův index přezdíván Shannon-Wienerův. Ale protože je často citována ona kniha, kde je druhým autorem Weaver, který má podobně znějící jméno, začalo se mu omylem říkat i Shannon-Weaverův, i když Warren Weaver vlastně na vyvinutí Shannonova indexu nemá zásluhu. Průzkum vědecké literatury ukazuje, že tento index je nejčastěji nazýván jako Shannonův (1105 krát), Shannon-Wienerův (214 krát) a Shannon-Wieverův (165 krát), a občas se objeví překlep[1] v podobě Shannon-Weinerův (namísto Shannon-Wienerův).

Design ekologických experimentů

Hurlbert S.H. (1984): Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54: 187-211.

- Článek, ve kterém Hurlbert zavedl pojem *pseudoreplikace*, a ve kterém čtivou formou vysvětluje rozdíly mezi jednotlivými designy ekologických experimentů a problémy s analýzou jejich výsledků. Mimochodem jeden z nejcitovanějších ekologických studií osmdesátých let (na Web of Science má téměř 5.000 citací).

Lájer K. (2007): Statistical tests as inappropriate tools for data analysis performed on non-random samples of plant communities. *Folia Geobotanica* 42: 115–122.

- Protože vegetační data nejsou sbírána náhodně, nejsou tedy náhodným výběrem a nelze na ně aplikovat klasické statistické metody - tak by se dal shrnout obsah tohoto článku, který svého času vzbudil bouřlivou diskuzi na téma jak tedy data sbírat, abychom je mohli statisticky zpracovávat. Lájerův článek i reakce na něj najdete ve zvláštním čísle *Folia Geobotanica*, které vzniklo právě s ohledem na onu bouřlivou diskuzi.

Michalcová D., Lvončík S., Chytrý M. & Hájek O. (2011): Bias in vegetation databases? A comparison of stratified-random and preferential sampling. *Journal of Vegetation Science*, 22: 281–291.

- Studie porovnává vlastnosti vegetačních dat sbíraných preferenčně a stratifikovaně náhodně; preferenčně sbíraná data mají větší beta diverzitu a vyšší zastoupení ohrožených druhů, pouze částečně pak platí že mají obecně vyšší druhovou bohatost a nižší zastoupení invazních druhů a neofytů.

[1] např. v knize *A Primer of Ecological Statistics* od N. Gotelli a A. Ellison