

2012

05

Statistické hodnocení biodiverzity

Jiří Jarkovský
Simona Littnerová
Ladislav Dušek

Statistické hodnocení biodiverzity



ISBN 978-80-7204-790-1



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Statistické hodnocení biodiverzity

Jiří Jarkovský, Simona Littnerová, Ladislav Dušek

Únor 2012



Příprava a vydání této publikace byly podporovány projektem ESF č. CZ.1.07/2.2.00/07.0318 „Víceoborová inovace studia Matematické biologie“ a státním rozpočtem České republiky.

Předmluva

Biodiverzita je často diskutovaným pojmem a to zejména v oblasti ekologie a ochrany životního prostředí, kdy je nicméně často používán pouze ve své subjektivní podobě. Zároveň jde o pojem mnohem širší a týkající se nejenom taxonomické rozrůzněnosti živé přírody. Pojem diverzita se prolíná všemi úrovněmi živé hmoty od diverzity biochemických drah v rámci buňky, genetickou diverzitu organismů, složení biologických společenstev až po strukturu ekosystémů. Z matematického hlediska vyrůstá pojem diverzita z informační teorie a zjednodušeně ji můžeme popsat jako míru informace obsažené v hodnoceném systému; řada metod používaných pro hodnocení biodiverzity je v tomto kontextu odvozena a využívána v celé řadě dalších oblastí vědeckého výzkumu.

Cílem předloženého textu je ukázat celou šíři pojmu biodiverzita a zároveň představit objektivní metody jejího statistického hodnocení. Hlavním účelem není podrobný teoretický výklad jednotlivých typů analýz biodiverzity, ale ve stručné a přehledné formě představit principy analýz, objasnit základy jejich využití včetně potenciálně slabých míst a poskytnout návody ke správné interpretaci výsledků; učební text tak slouží zejména jako doplnění přednášek statistického hodnocení biodiverzity a jako referenční text při samostatném studiu problematiky. Studenti v textu naleznou nejenom běžně známé metody indexového hodnocení biodiverzity, ale i méně často používané modely rozdělení druhových abundancí a problematiku využití dat týkajících se biodiverzity ve vícerozměrných statistických analýzách.

Dostupnost nových studijních materiálů, kterých je v současné době stále nedostatek, by měla přispět k zvýšení odbornosti studentů matematické biologie i dalších oborů.

Na tomto místě bychom rádi poděkovali za připomínky recenzentům, jejichž poznámky výrazně zlepšily kvalitu těchto učebních textů.

Příprava a vydání této publikace byly podporovány projektem ESF č. CZ.1.07/2.2.00/07.0318 „Víceoborová inovace studia Matematické biologie“ a státním rozpočtem České republiky.

V Brně 28. 2. 2012

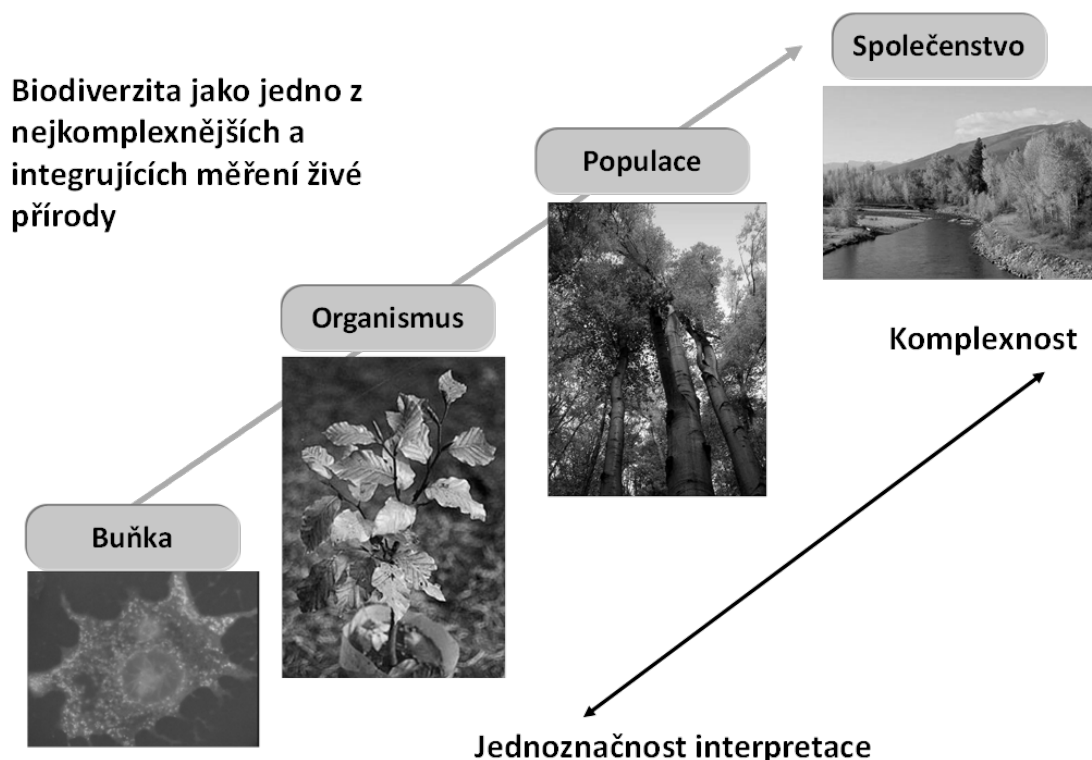
Jiří Jarkovský
Simona Littnerová
Ladislav Dušek

1 Biodiverzita jako pojem

1.1 Definice biodiverzity

1.1.1 Biodiverzita z ekologického hlediska

Biodiverzita jako komplexní ukazatel stavu živé přírody (obr. 1.1) má sama velmi různorodé definice. Za nejjednodušší lze považovat pouhý počet druhů (druhová bohatost) společenstva [1], naopak velmi komplexní je definice, kterou zde uvádíme v původním tvaru i jazyce [2]: “Biodiversity is an attribute of an area and specifically refers to the variety within and among living organisms, assemblages of living organisms, biotic communities, and biotic processes, whether naturally occurring or modified by humans. Biodiversity can be measured in terms of genetic diversity and the identity and number of different types of species, assemblages of species, biotic communities and biotic processes, and the amount (e.g., abundance, biomass, cover, rate) and structure of each. It can be observed and measured at any spatial scale ranging from microsites and habitat patches to the entire biosphere” čili “Biodiverzita je vlastnost dané oblasti vztahovaná zejména k různorodosti živých organismů, jejich seskupení, biologických společenstev a procesů, ať již se vyskytující přirozeně nebo ovlivněných lidskou činností. Diverzita může být měřena ve smyslu genetické diverzity, druhové bohatosti, výskytu skupin druhů, biologických společenstev a procesů a to jak z hlediska kvantity (abundance, biomasa, pokrývnost atd.), tak jejich struktury. Může být měřena na všech prostorových úrovních od mikrohabitatů až po celou biosféru.” Předložený text vychází z této širší definice a snaží se představit metody popisující diverzitu z různých pohledů popsaných v definici.



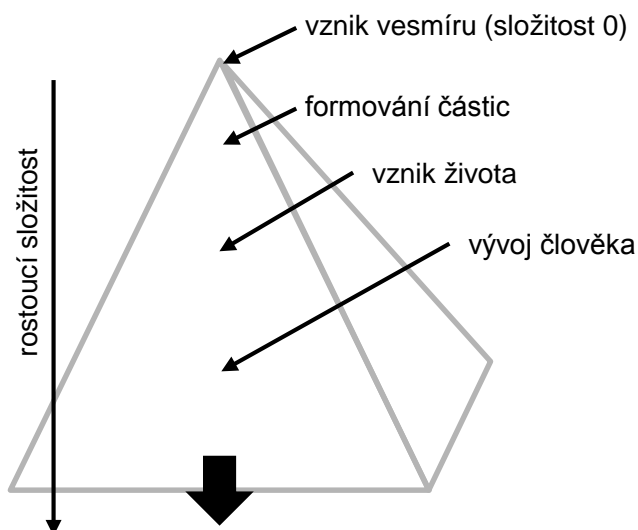
Obr. 1.1 Koncept biodiverzity lze využít na všech úrovních živé přírody

1.1.2 Biodiverzita z matematického hlediska

Z hlediska matematického pozadí výpočtů souvisí analýza diverzity s entropií a informační teorií. Jednotlivé taxony, geny nebo jakékoliv součásti určitého systému představují jednotky informace. Pojem entropie lze chápat jako míru „neuspořádanosti“ zkoumaného systému. Tento pojem zavedl v roce 1865 německý fyzik Rudolf Clausius (1822-1888) v souvislosti s termodynamickými zákony a z hlediska teorie informace jej popsal Claude Elwood Shannon (1916-2001).

Neustálý růst složitosti je podmíněn růstem entropie. Tuto situaci si lze graficky představit jako pyramidu, na jejíž špičce je nula: začátek všeho, nulová entropie, žádný pohyb, žádný čas. Od této chvíle se budují stále složitější struktury a postupujeme směrem dolů ke stavu s vyšší entropií. Pyramida se neustále rozšiřuje, každý následující stav je ovlivněn předchozím stavem. Na určitém místě, relativně blízko špičce, vznikla hmota vesmíru, na jiném daleko níže je vznik prvních živých jednobuněčných organismů a ještě níž člověk, přičemž entropie, tedy složitost systému, stále narůstá (obr. 1.2). Entropie je jedním ze základních a nejdůležitějších pojmů ve fyzice, matematice, teorii pravděpodobnosti a informace a v mnoha dalších oblastech vědy teoretické i aplikované. Setkáváme se s ní všude tam, kde hovoříme o pravděpodobnosti možných stavů daného systému či soustavy.

Chápeme-li entropii jako neurčitost systému, můžeme pojem entropie nahradit pojmem diverzita, a pokud naše zkoumané systémy jsou biologické, lze mluvit o biodiverzitě. Z tohoto hlediska může být biodiverzita chápána jako „neuspořádanost“ biologických systémů.



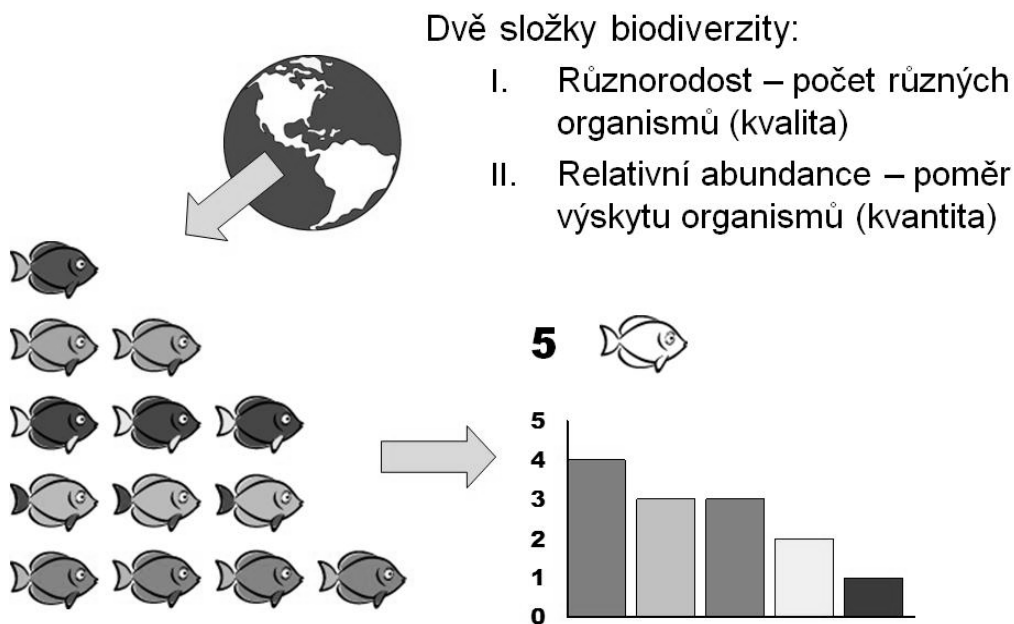
Obr. 1.2 Postupný nárůst entropie od vzniku vesmíru

1.2 Kvalitativní a kvantitativní komponenta biodiverzity

Jak vyplývá z předchozího textu, biodiverzita je velmi široký pojem zahrnující například druhovou diverzitu (počty druhů ve společenstvech), genetickou diverzitu (rozdílnost genomů organismů) nebo ekologickou diverzitu (počet ekosystémů v prostředí).

Pak podle typu hodnocení rozlišujeme biodiverzitu fyziologickou, taxonomickou, strukturní, genetickou, apod. U sledovaných forem se jednak hodnotí jejich přítomnost ve společenstvu („druhová bohatost“) a jednak se vyhodnocuje jejich kvantitativní zastoupení, heterogenita početnosti druhů.

Oba přístupy jsou pro exaktní analýzu biodiverzity nezbytné a hovoříme proto o tzv. duálním konceptu hodnocení biodiverzity [5] (obr. 1.3).



Obr. 1.3 Duální koncept biodiverzity

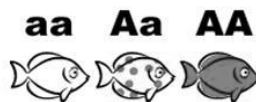
Tradiční chápání biodiverzity je často omezeno pouze na pojem taxonomické biodiverzity, tedy výskytu druhů a jejich abundance (počty jedinců). Tento přístup je nicméně velmi zjednodušující, jak bylo ukázáno v kapitole o matematické podstatě analýz diverzity. Její výpočet se týká všech situací, kdy dokážeme definovat kvalitativní i kvantitativní komponentu nějakého společenstva. V konkrétním případě je nutné vždy zvážit, jaká komponenta biodiverzity je přínosná pro ověření nebo vyvrácení hypotéz prováděné studie.

Kvalitativní komponenta biodiverzity může být definována v podstatě na libovolné úrovni organizace biologického systému, jak je naznačeno na obr. 1.4.

Taxonomická diverzita – výskyt a četnost jedinců druhů nebo jiných taxonomických jednotek



Genetická – výskyt různých kombinací alel v populacích organismů



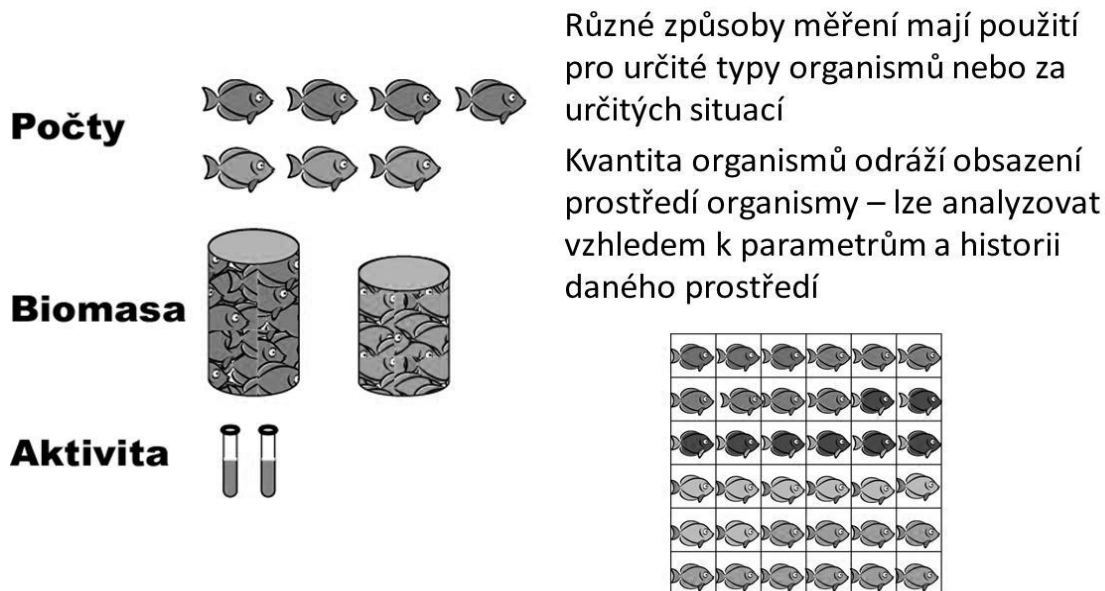
Ekologická/funkční – funkce, kterou organismy vykonávají v rámci společenstva (predátor, parazit, dekompozitor, sesilní mobilní organismy atd.)



Fyziologická/biochemická diverzita – způsoby a biochemické dráhy používané organismy k zpracování substrátu

Obr. 1.4 Příklady kvalitativní komponenty biodiverzity

Kvantitativní komponentou biodiverzity je tradičně abundance, tedy počty jedinců. Tuto kvantitativní komponentu ale získáme pouze v případě jednoznačně ohraničených organismů přibližně stejné velikosti. Pokud nejsou jedinci organismu jednoznačně definováni (např. některé rostliny) nebo jedinci různých taxonů nejsou srovnatelní, je možné využít další kvantitativní míry jako je biomasa, biochemická aktivita atd. (obr. 1.5).



Obr. 1.5 Kvantita organismů může být měřena různými způsoby

Kromě těchto biologických aplikací může být analýza diverzity využívána jako součást výpočtu rozhodovacích stromů (srovnání míry informace ovlivněné daným dělením stromu), hledání optimálních dělení shlukové analýzy (jaký počet shluků nese nejvíce informace) nebo i pro výběr proměnných v predikčních modelech (maximalizace informace sdílené mezi prediktory a predikovanou proměnnou a její současná minimalizace mezi prediktory navzájem).

1.3 α , β , γ diverzita

Z hlediska úrovně geografického detailu můžeme biodiverzitu popisovat ze tří základních pohledů:

1. α diverzita, která se zabývá druhy v rámci společenstev a snaží se popisovat vztahy jejich vzájemného výskytu;
2. β diverzita, která sleduje rozdíly druhového složení ve více společenstvech a
3. γ diverzita zabývající se druhy v rámci regionů (jde tedy o kombinaci α a β diverzity).

V případě α diverzity jde o popis jednotlivých společenstev organismů, v zásadě existují dva možné přístupy:

1. indexy diverzity snažící se postihnout různé aspekty společenstev jednou hodnotou (počet druhů, dominance, vyrovnanost společenstva). Tyto indexy jsou založené buď na empirickém přístupu nebo matematických teoriích, například Shannonově informační teorii.
2. Modely druhové abundance, popisující teoretické rozložení abundancí druhů ve společenstvu, kdy jsou vypočítaná teoretická rozložení porovnávána s reálným společenstvem.

1.4 Biodiverzita jako cílový parametr v hodnocení ekologických rizik

Hodnocení ekologických rizik (EcoRA) je proces, hodnotící pomocí komplexního posuzování pravděpodobnost vlivu stresových faktorů na stav ekosystémů. Hodnocení může být obecně použitelné nebo specifické pro dané místo a je často součástí velice komplexních studií. V současnosti se jedná o nejlepší dostupný nástroj pro predikci potenciálního nebezpečí a negativních dopadů antropogenních aktivit na strukturu a funkci ekosystému.

Některé parametry, používané pro hodnocení účinků, lze získat komplexní analýzou biologických receptorů na úrovni společenstev a ekosystémů. Do této kategorie studií patří i velmi časté analýzy biodiverzity ekosystémů pod vlivem stresu.

Hodnocení tohoto typu jsou na ekosystémové úrovni nezbytnou součástí regionálně specifických přístupů EcoRA, jejich standardizace je však obtížná. Je samozřejmé, že z metodického hlediska má největší význam hledání bioindikátorových systémů, jejichž změna pod vlivem stresových faktorů je jednoznačná a má významné a přímé ekologické interpretace [6].

1.4.1 Biodiverzita jako integrující ukazatel stavu biologických společenstev

Existuje řada parametrů biologických společenstev využitelných jako ukazatele stresu v prostředí. Za systémové a integrující ukazatele však považujeme především indikátory struktury a biodiverzity. Bioindikace na této úrovni je založena na faktu, že výskyt a počty organismů ve společenstvech jsou ovlivněny dvěma základními okolnostmi:

- habitatovou situací na dané lokalitě – souhrn veškerých faktorů včetně stresorů, které na dané lokalitě mohou na společenstvo působit;
- habitatovou preferencí organismu (valencí organismů pro jednotlivé faktory na lokalitě) – výskyt organismu je dán jeho schopností žít na daném místě za daných hodnot environmentálních parametrů, přičemž různé druhy mají různé kombinace optimálních hodnot parametrů prostředí.

Konečný výskyt druhů a jejich početnost jsou tak dány kombinací možností organismů žít na daném místě za daných podmínek. Základní možnosti, jak tento vliv prostředí na výskyt druhů měřit jsou:

- produkce, velikost společenstev – kvantitativní ukazatele stavu společenstva;
- struktura a typ společenstev – kvalitativní ukazatele stavu společenstva;
- biodiverzita společenstev – kvalitativní ukazatele stavu společenstva;
- vývoj a stabilita společenstev – změny společenstva v čase na kvalitativní nebo kvantitativní úrovni.

Každý z těchto ukazatelů se zaměřuje na jiné aspekty existence a vývoje společenstva a pro komplexní popis je vhodné vždy kombinovat několik úhlů pohledu. Při analýze je možné uvažovat buď celkové společenstvo všech organismů (rostliny, živočichové), tento přístup je však velmi náročný na provedení a finanční prostředky a má význam spíše při komplexním vědeckém výzkumu. Druhou možností je využití části společenstva, která nás zajímá z hlediska hodnocení cílových parametrů (např. společenstvo ryb na dané lokalitě pro hodnocení produkce a složení za účelem rybářského využití; společenstvo lišejníků na kmenech stromů pro indikaci atmosférického znečištění; společenstvo rybích parazitů jako model akumulace stresu v ekosystému z pohledu parazitů jako vrcholových predátorů, apod.).

1.4.2 Produkce a velikost společenstev

Produkce společenstev je množství organické hmoty (biomasy) vytvořené za určitou dobu na jednotku plochy nebo objemu systému. Význam tohoto parametru spočívá v hodnocení velikosti a také růstových schopnostech společenstva. Srovnáním výsledků s publikovanými výstupy pro obdobné situace nebo dřívějším výzkumem umožňuje zjistit, zda se společenstvo odchyluje od teoretických nebo dříve naměřených hodnot velikosti/produkce. Příkladem může být rybí společenstvo, kde můžeme jednak sledovat snížení produkce ryb v přítomnosti stresoru nebo změnu početnosti společenstev cizopasníků v důsledku vlivu stresoru na hostitele [12].

V případě srovnávání společenstev velmi odlišných organismů (společenstvo primárních producentů, herbivorů, predátorů atd.) musí být velikost společenstev vyjadřována v termínech biomasy nebo koncentrací produktů metabolismu. Biomasu lze měřit různými způsoby, lze využít jednak přímé měření biomasy (např. metodou sklizně), jednak nepřímá měření založená na znalosti biochemických drah fotosyntézy (koncentrace chlorofylu, úbytky CO₂ a biogenních prvků atd.). Pro zjišťování sekundární produkce na různých trofických úrovních lze jednak měřit biomasu organismů nebo jsou využívány modely průchodu hmoty a energie trofickými úrovněmi společenstva (příklady různých metod např. [13]).

V případě hodnocení společenstev podobných organismů nebo za situace, kdy počet je významnější než biomasa (např. společenstva rybích parazitů, kde má význam celková zátěž hostitele), je velikost společenstva popisována jednotlivými abundancemi, tedy početnostmi jedinců.

1.4.3 Struktura a typ společenstev

Jedná se o kvalitativní hodnocení společenstev. Strukturou společenstev rozumíme zastoupení skupin organismů s různým významem pro společenstvo, různou životní strategií a preferencemi. Strukturu společenstva můžeme sledovat z různých úhlů pohledu:

- trofické úrovně – zastoupení primárních producentů, herbivorů, predátorů, destruentů;
- životní strategie – r a k strategové [14];
- specifita parazitů – je hodnocen podíl parazitů specializovaných na daného hostitele (specialisté) a parazitů bez vyhraněných preferencí (generalisté);
- preference prostředí – podíl organismů s preferencí pro specifické podmínky (např. třídy saprobity);
- potravní, lokomoční a jiné preference – často jsou využívány při hodnocení společenstev makrozoobentosu, jsou hodnoceny například podíly druhů s různým způsobem obživy na lokalitě;
- metabolické charakteristiky – tento přístup je využíván při hodnocení mikrobiálních společenstev, kdy je hodnocena metabolická aktivita různých skupin mikroorganismů.

Velikosti jednotlivých skupin jsou stanovovány na základě měření biomasy, abundance, biochemické aktivity. Protože jednotlivé skupiny tvořící strukturu společenstva mají svůj ekologický význam, je možné z jejich poměrů odvodit podmínky panující na dané lokalitě (a tedy i případné ovlivnění stresorem). Zjištěné výsledky je opět možné porovnat s dosud publikovanými údaji nebo předchozími měřeními. Příkladem mohou být r a k strategové, kdy r strategové s velkým reprodukčním potenciálem a bez mechanismů regulující růst osidlují spíše extrémní nebo stresovaná prostředí než k strategové, kteří jsou typičtější pro stabilní ekosystémy.

V některých případech (např. rostlinná společenstva) lze definovat poměrně striktně daná společenstva spjatá s určitými přírodními podmínkami. Zároveň lze srovnávat odchylky struktury společenstva od teoreticky předpokládaného složení.

1.4.4 Koncept chápání biodiverzity v ekologických studiích

Základem pro hodnocení biodiverzity jsou vstupní data získaná procesem vzorkování. Nejčastěji jimi bývá počet druhů a abundance (početnost). Avšak ty nejsou vhodné pro všechny organismy, proto se využívá i měření biomasy, produkce, pokryvnosti a dalších shrnujících charakteristik, které jsou vhodné zvláště pro rostliny nebo planktonní organismy. Základem pro hodnocení biodiverzity je pak koncept druhové bohatosti (*species richness*), kdy nás zajímá, kolik druhů (forem) a o jaké početnosti (či jiné míře kvantity) nacházíme ve zkoumaném prostředí, z čehož pak počítáme a odvozujeme další charakteristiky.

Jedním z atributů biologického společenstva je druhová diverzita. Pro její měření bylo navrženo mnoho různých způsobů. Avšak před vlastním vyhodnocením je nutno vždy provést řadu dalších kroků.

Prvním předpokladem je, že námi vybraný vzorek je dobře definovaný. Výpočet diverzity vyžaduje přesnou taxonomickou klasifikaci námi zkoumaného vzorku společenstva.

Většina měření diverzity dále předpokládá, že všechny druhy zahrnuté do výpočtu si jsou podobné rozsahy svých nároků. Vypořádat se s tímto omezením není jednoduché.

Posouzení diverzity dále vyžaduje znalost významu jednotlivých druhů pro společenstvo. Jako ukazatel obvykle volíme počet jedinců, biomasu, pokryvnost nebo produkci nějaké chemické látky. Volba částečně závisí na otázce, kterou si klademe. Většinou se používají počty jedinců, ačkoli například při hodnocení planktonu bývá nejlepším parametrem produkce [15].

Otázka s tím spojená je, jak velký vzorek společenstva bychom měli zahrnout do výzkumu. Proto musíme přesně definovat seznam druhů, které se snažíme popsat. Většina autorů si vybírá jistou část společenstva; hodnotí například diverzitu ptáčích nebo stromových společenstev na daném území. Pouze vzácně překračují měření diverzity více trofických úrovní, nebo jsou aplikována na společenstvo jako celek. Podle [16] by se ekologové měli zaměřit na sledování diverzity těch částí celkového společenstva, které jsou ve funkční interakci. Tyto podmnožiny často pokrývají více trofických úrovní a zahrnují druhy vzájemně taxonomicky nepříbuzné. Správný výběr jednotek zahrnutých do studie je velmi důležitý pro korektní výstup a porozumění ekologii společenstva [5].

1.4.5 Vývoj a stabilita společenstev

Analýza vývoje a stability společenstev vnáší do hodnocení biodiverzity časový rozměr. Ten může být retrospektivní, kdy na základě dřívějších dat je možné odhalit periodické změny ve vývoji společenstva (což může zabránit chybné interpretaci dat z jediného odběru), zjistit stabilitu složení společenstva v čase atd. Obecně lze přidat časové měřítko k hodnocení společenstev a zjišťovat tak například změny produkce v průběhu roku, výskyt některých druhů jen v určitém období, změny podílů potravních skupin a jiné. Pro některé typy prostředí jsou definovány sukcesní řady a je možné zjišťovat odchylky od průběhu sukcese předpokládané na základě přírodních podmínek. Další metodickou možností je predikce vývoje společenstva. Aby bylo možné predikovat vývoj společenstva, je třeba pro jeho jednotlivé komponenty zjistit parametry natality, mortality, emigrace a imigrace [17, 18]. Existují i komplexní populační modely vývoje struktury společenstev, například modely společenstev popsané pomocí principu markovovských modelů [19, 20].

1.4.6 Pozice ukazatelů biologických společenstev v celkovém hodnocení vlivu stresoru

Jak již bylo popsáno v předchozích kapitolách, analýza ukazatelů společenstev není příliš specifická, a tudíž není vhodná pro indikaci konkrétních stresorů. Avšak její výhoda tkví především v:

- „paměti“ společenstva, která může uchovávat vliv stresorů i po odeznění jejich přímého působení;
- „zesílení“ odezvy na vliv stresoru prostřednictvím narušení rovnovážných stavů ve společenstvu, kdy i menší narušení této rovnováhy může mít za následek rozsáhlé změny struktury společenstva (i na základě menší chronické zátěže může časem dojít k vyčerpání „pufrační“ schopnosti společenstva a vzniku pozorovatelných změn).

1.4.7 Současný vývoj metod v hodnocení biodiverzity

Hodnocení biodiverzity a ekologické diverzity je významným modelovým cílovým parametrem. Jako vysoce integrující bioindikační parametr představuje biodiverzita typický příklad nutné aplikace biometrických metod, a to především z následujících důvodů: (1) data týkající se biodiverzity jsou většinou velmi heterogenní; (2) běžně dostupné způsoby hodnocení jsou často založeny na výpočtech indexů, jejichž číselné hodnoty a interpretace mohou být jednostranné a zavádějící; (3) biodiverzita a její změny jsou významným faktorem pro hodnocení rizika na ekosystémové úrovni; (4) aplikací širšího spektra metod lze získat spolehlivější a obsáhlejší informace a dosáhnout srovnatelnosti rozdílných souborů dat. K hodnocení lze použít standardní i moderní metody hodnocení:

- indexy diverzity, ukazatele ekvitability,
- predikční techniky a nestandardní techniky odhadu parametrů (rarefakce, bootstrap, jackknife),
- modely druhové abundance,
- grafické techniky posuzující strukturu společenstva,
- vícerozměrné klasifikační a ordinační techniky,
- techniky spojující epidemiologické modely se standardním hodnocením diverzity (při hodnocení vztahů hostitel - parazit).

V hodnocení biodiverzity společenstev lze obecně vyzorovat dva trendy - použití indexů diverzity a popis diverzity společenstva pomocí modelů. Hlavní rozdíl mezi modely druhové abundance a indexy diverzity spočívá v tom, že indexy se snaží sumarizovat v jedné číselné hodnotě diverzitu společenstva, na rozdíl od modelů, které se snaží takové agregaci informace vyhnout a soustředují se na popis celkového tvaru křivky abundancí. Logicky je sice výhodnější využít maximum informace obsažené v datech, avšak v některých případech je praktičtější a výhodnější použití všeobjímajícího indexu diverzity. Tyto techniky jsou blíže popsány v následujících kapitolách.

2 Biodiverzita a biostatistika

2.1 Analýza biodiverzity jako analogie ke klasickým statistickým postupům

Mezi analýzou biodiverzity a statistickou analýzou lze nalézt celou řadu paralel. Přímá analogie leží zejména v následujících oblastech:

- Popisná statistika – indexy diverzity jsou obdobou klasických popisných statistik, jde vlastně o popisné statistiky biodiverzity společenstev.
- Modelová rozložení – prokládání modelů početnosti druhů je obdobou prokládání klasických statistických rozložení.
- Testování hypotéz – indexy diverzity a prokládání modelů početnosti druhů mohou být testovány analogicky k běžným popisným statistikám a prokládání modelových rozložení.
- Vícerozměrná analýza – data o biodiverzitě jsou primárně vícerozměrná a proto je vhodné pro analýzu dat použít vícerozměrné statistické metody; řada metod používaných v hodnocení β diverzity jsou zároveň indexy podobnosti užívané při výpočtu asociačních matic.

2.2 Vzorkování biodiverzity

Pro vzorkování biodiverzity platí obdobná pravidla jako při jakémkoliv jiném vzorkování, při výběru z cílové populace (pojem populace je zde používán ve smyslu statistickém, nikoliv biologickém) musí pro získání kvalitního vzorku platit:

- náhodnost,
- reprezentativnost,
- nezávislost.

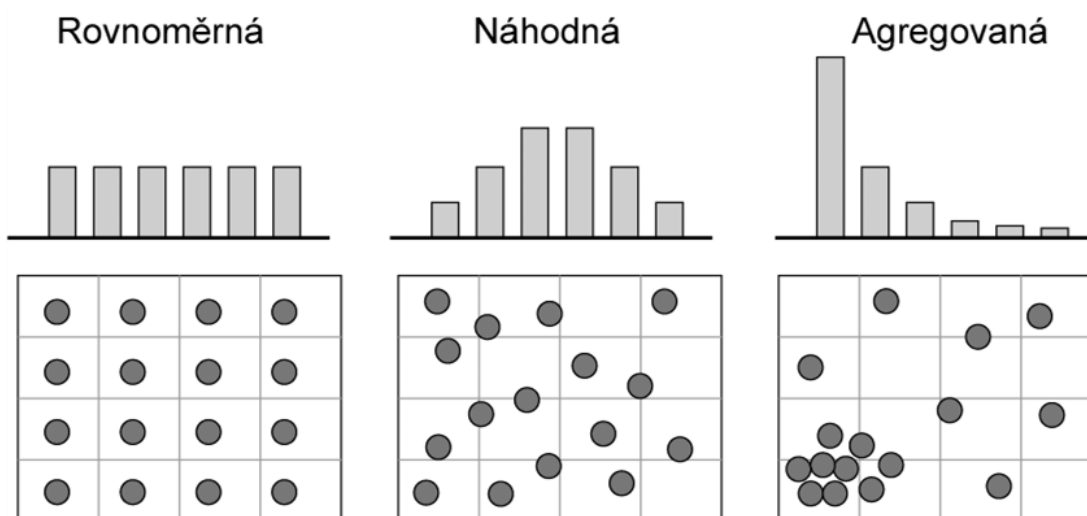
Je samozřejmě pravda, že heterogenita živé přírody klade vysoké nároky na přípravu vzorkovacího plánu, kdy kromě prostorové heterogenity abiotického prostředí je třeba počítat i s různým rozložením organismů v tomto prostředí.

Organismy se v prostředí většinou nevyskytují rovnoměrně rozložené, příčinou je heterogenita prostředí, která ovlivňuje výskyt různých organismů. Způsob, jakým jsou organismy v prostředí rozmístěny, nazýváme jejich frekvenční distribucí. Kromě vzácného rovnoměrného rozložení se můžeme setkat ještě s rozložením náhodným a agregovaným (obr. 2.1). Nejjednodušším postupem pro jejich zjištění je poměr průměru a rozptylu početnosti organismů ve vzorcích prostředí:

- rovnoměrné rozložení má rozptyl menší než průměr,
- náhodné rozložení má rozptyl roven průměru,
- agregované rozložení má rozptyl větší než průměr.

Tyto modely jsou popsány třemi dobře známými rozděleními [24]:

- pozitivní binomické (pro rovnoměrné rozložení jedinců),
- Poissonovo (pro náhodné rozložení jedinců),
- negativní binomické (pro agregovaná rozložení jedinců).



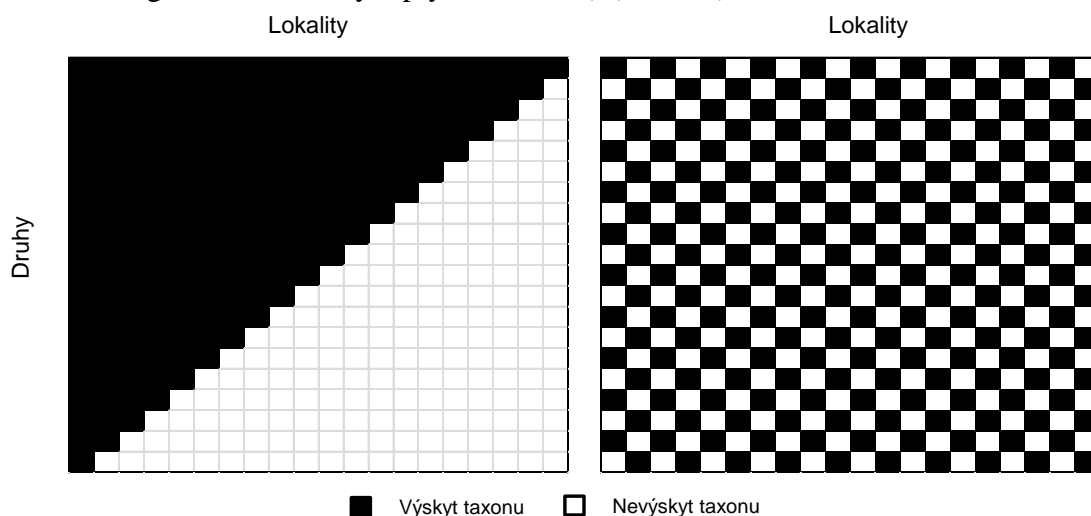
Obr. 2.1 Možné typy frekvenční distribuce organismů v prostředí

Pro měření stupně agregovaného výskytu organismů byl některými autory [25, 26] použit jako index agregace parametr k negativního binomického rozdělení. Při negativním binomickém rozdělení je jeho vztah k průměru a rozptylu následující:

$$k = \frac{\mu^2}{\sigma^2 - \mu} \quad (2.1)$$

Při náhodném rozdělení (Poissonovo), kdy rozptyl se rovná průměru, se hodnota k blíží k nekonečnu, zatímco při agregovaném rozložení získáme nízké hodnoty parametru k . Přestože nelze dosáhnout přesného hodnocení agregace jediným indexem, je parametr k za určitých podmínek použitelný [27].

Další metodou pro určení uspořádanosti systému (agregované nebo náhodné rozložení druhů na lokalitách) je přístup termodynamické uspořádanosti [28], který hodnotí uspořádanost matice „druhy x lokality“ pomocí termodynamické teploty systému. Teplota 0 °C (matice s jasnou strukturou, ekvivalent vody v pevném stavu) odpovídá maximálně uspořádanému (agregovanému) systému, zatímco plně náhodný systém odpovídá teplotě 100 °C (není žádný vztah mezi druhy a lokalitami, analogie molekul vody v plynném stavu) (obr. 2.2).

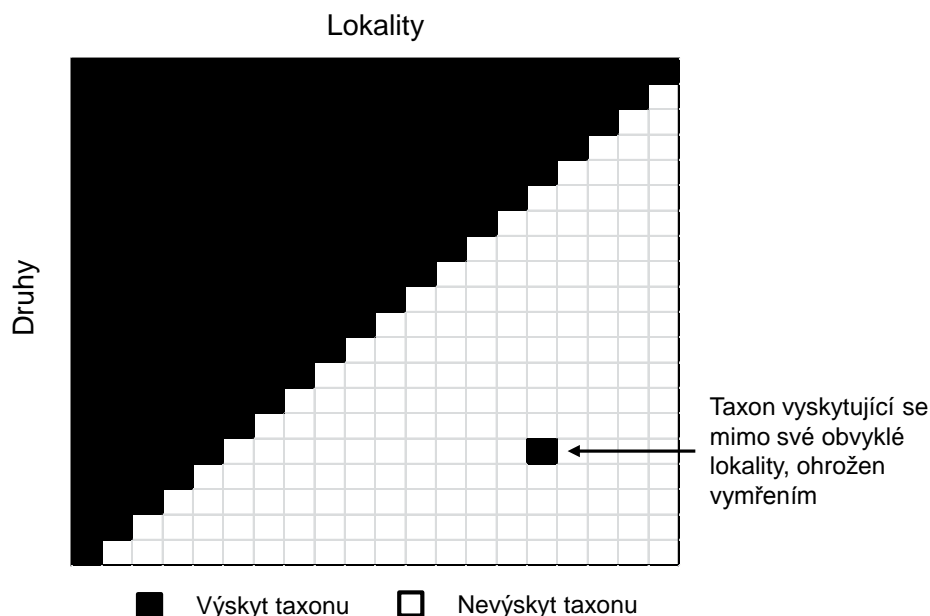


Plně uspořádaný systém (nízká termodynamická teplota)

Chaotický systém (vysoká termodynamická teplota)

Obr. 2.2 Uspořádanost vztahu lokalit a výskytu taxonů

Postup byl původně vyvinut pro účely ochrany přírody. Zde hodnotí, zda výskyt druhů na lokalitách (ve smyslu samostatných „ostrovů“) odpovídá uspořádanému vymírání druhů na druhově méně bohatých lokalitách, nebo zda lokality obsahují druhy, které tomuto uspořádanému systému neodpovídají, a tedy jsou na dané lokalitě ohroženy vymřením (obr. 2.3).



Obr. 2.3 Výskyt taxonu mimo jeho obvyklý vzorec rozšíření jako ukazatel rizika vymření

Autoři doplnili svou práci softwarovým nástrojem pro výpočet termodynamické teploty systému (<http://www.aics-research.com/nestedness/tempcalc.html>), kterou lze následně pomocí Monte-Carlo testu srovnat s náhodným uspořádáním systému (je testována nulová hypotéza, že matice druhů x lokality je náhodně uspořádána).

2.3 Ukládání dat biodiverzity

Při analýze dat o libovolných společenstvech organismů se nevyhneme na první pohled jednoduchému problému jak data uchovávat. Ve skutečnosti jde o klíčový problém. Pokud budeme data zaznamenávat a ukládat nevhodným způsobem, znamená to následně velké problémy pro jednoduché zpracování dat, zbytečnou, často rutinní a zdlouhavou práci a zvýšenou možnost výskytu chyb.

Již v úvodní fázi je nutné definovat, jaké proměnné budeme měřit s ohledem na: (1) maximální výtěžnost informací z dostupného biologického materiálu; (2) zaměření studie; (3) pracnost získání různých typů informací. Je třeba si uvědomit, že informace, které nezaznamenáme, již budeme jen obtížně získávat (např. teplotu vody při odlovu ryb lze získat na základě měření různých institucí, ale nejde samozřejmě o zcela stejná měření), popřípadě je jejich získání zcela nemožné.

Předpokladem pro další zpracování je převedení dat do elektronické podoby, pro jednoduché zpracování na počítači je nejvhodnější forma databázové tabulky. V takové tabulce obsahuje každý sloupec jeden typ informace (např. název druhu, abundance druhu, datum odchyty, lokalita atd.). Řádky tabulky pak představují nejmenší jednotku, pro kterou data získáváme (např. pro tabulku informací o hostitelích je nejmenší informační jednotkou ryba, pro tabulku zaznamenávající lokalizaci parazitů jsou nejmenší jednotkou jednotliví parazité). Příklad datové tabulky je zobrazena na obrázku (obr. 2.4), kde v řádcích jsou jednotlivé lokality a ve sloupcích abundance rybích druhů „navzorkovaných“ na lokalitě. Takto uspořádané tabulky lze snadno zpracovat do

podoby libovolných sumárních tabulek nebo výstupních statistik pomocí například MS Excel, MS Access, Statistica for Windows, SPSS, R apod.

The screenshot shows a Microsoft Excel spreadsheet with the following data:

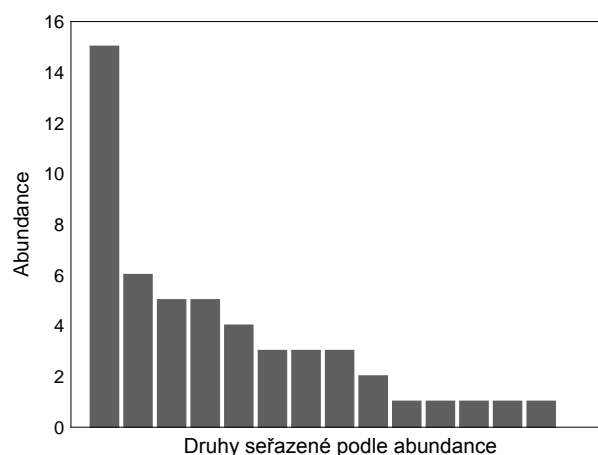
	A	B	C	D	E	F	G
1	Lokalita	Cottus gobio	Salmo trutta fario	Phoxinus phoxinus	Nemacheilus barbatulus	Thymallus thymallus	Telestes soufia agassizi
2	1	0	3	0	0	0	0
3	2	0	5	4	3	0	0
4	3	0	5	5	5	0	0
5	4	0	4	5	5	0	0
6	5	0	2	3	2	0	0
7	6	0	3	4	5	0	0
8	7	0	5	4	5	0	0
9	8	0	0	0	0	0	0
10	9	0	0	1	3	0	0
11	10	0	1	4	4	0	0
12	11	1	3	4	1	1	0
13	12	2	5	4	4	2	0
14	13	2	5	5	2	3	2
15	14	3	5	5	4	4	3
16	15	3	4	4	5	2	4
17	16	2	3	3	5	0	5
18	17	1	2	4	4	1	2
19	18	1	1	3	3	1	1
20	19	0	0	3	5	0	1
21	20	0	0	1	2	0	0
22	21	0	0	1	1	0	0
23	22	0	0	0	1	0	0

Obr. 2.4 Ukázka datové tabulky, kde v řádcích jsou zaznamenány jednotlivé lokality a ve sloupcích druhy ryb vyskytující se na lokalitách

3 Vizualizace biodiverzity

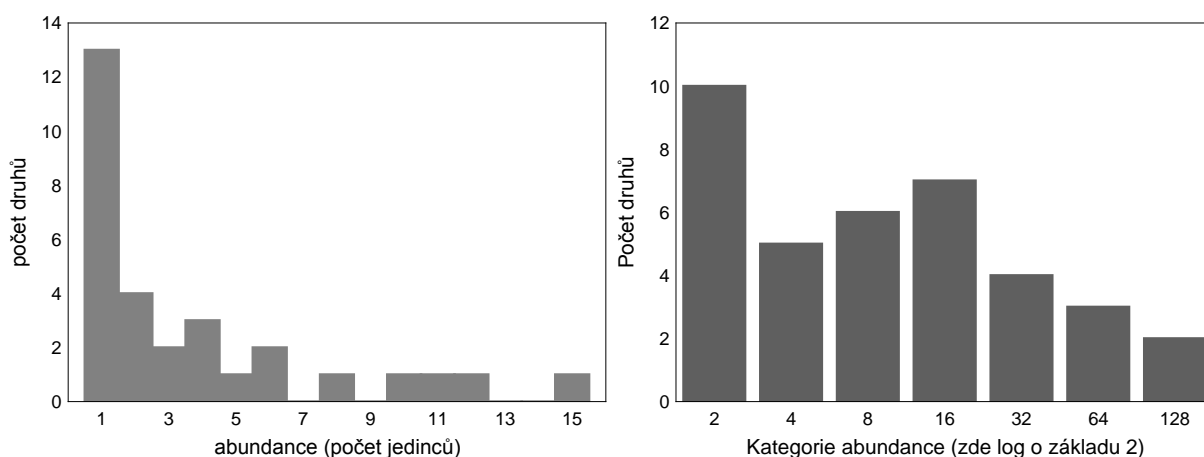
3.1 Složení společenstva a jeho zobrazení z různých pohledů

Nástrojem v hodnocení společenstev je i jejich grafické zobrazení. Je používána řada typů grafů lišících se od sebe úhlem pohledu na společenstvo. Některé z nich jsou zaměřeny spíše na představu o složení společenstva, jiné jsou vhodné pro srovnání více společenstev nebo jsou spojeny se zobrazením některých typů modelů abundance druhů.



Obr. 3.1 Graf pořadí abundancí

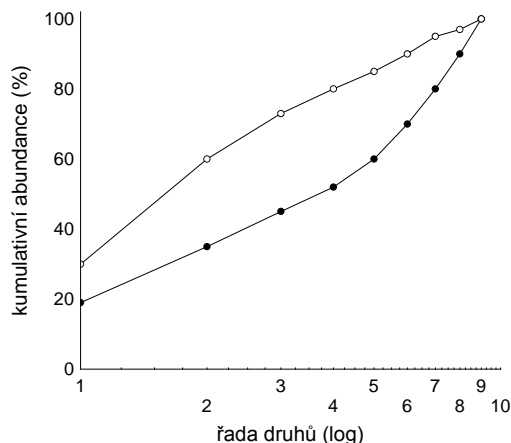
Obrázek 3.1 zobrazuje abundance taxonů seřazených podle jejich abundance, poskytuje přehlednou představu o dominantních i vzácných druzích. Lze jej použít pro srovnání druhového složení a významnosti druhů v různých společenstvech (podle vybraného společenstva je vytvořeno pořadí druhů, podle kterého jsou následně seřazeny druhy v ostatních společenstvech). Osa y může být v absolutních abundancích nebo relativních (podíl počtu jedinců druhů ve společenstvu k celkovému počtu jedinců).



Obr. 3.2 Graf abundance druhů

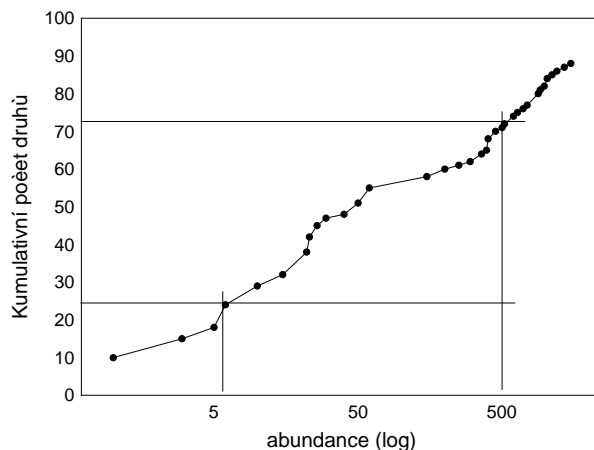
Obrázek 3.2 zobrazuje počty druhů příslušejících dané hodnotě abundance, umožňuje tak posoudit přítomnost vzácných a velmi četných druhů. V tomto typu grafu není zachycena identita jednotlivých druhů, pouze jejich přítomnost.

Nevýhodou grafu je, že v případě velmi početných druhů může mít osa x velký rozsah. Pro přehlednější zobrazení v případě velkého rozsahu hodnot abundancí je možné použít logaritmaci osy x nebo agregovat abundance do tříd abundancí (graf kategorií abundance). Hranice tříd mohou být určeny uživatelsky nebo jsou často používány logaritmické třídy abundance (viz druhý graf). V této formě produkují výsledky některé z matematických modelů rozložení abundancí. Tento graf je vhodným výstupem pro srovnání modelového a reálného rozložení abundancí druhů.



Obr. 3.3 Kumulativní abundance (*K-dominance graf*)

Obrázek 3.3 zobrazuje logaritmovanou řadu druhů seřazených podle abundance proti jejich kumulativní relativní abundanci, díky tomuto relativnímu přístupu je graf vhodný pro optické srovnání průběhu abundancí více společenstev.



Obr. 3.4 Graf kumulativního počtu druhů

Obrázek 3.4 zobrazuje kumulativní počet druhů proti ose jejich logaritmované abundance. Zobrazuje strmost narůstání počtu druhů se stoupající abundancí. Často je používán pouze interkvartilový rozsah druhů jako ukazatel základní části společenstva s vyloučením vzácných a extrémně početných druhů.

3.2 Rozložení a transformace dat při analýze biodiverzity

Jak již bylo zmíněno v části pojednávající o rozmístění organismů v prostředí, abundance taxonů napříč lokalitami nabývají obvykle lognormálního rozdělení. Také poměrně častý je i výskyt odlehklých hodnot, tedy jednotlivých lokalit, na nichž taxon nabývá velmi vysokých abundancí. Tento model je v přírodě velmi častý a platí nejenom pro abundance, ale i pro další míry kvantity organismů nebo i například pro velikost jejich areálu.

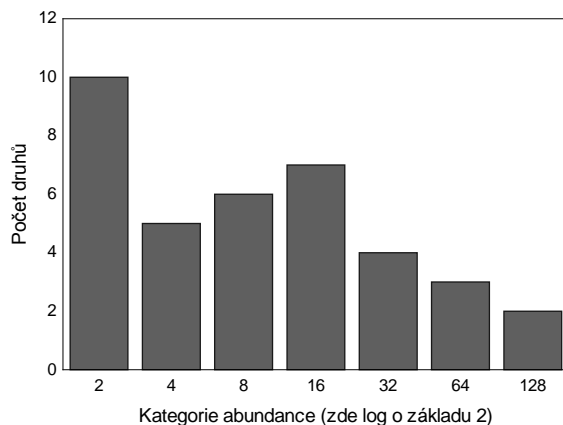
Z tohoto důvodu je třeba věnovat pozornost výběru popisných statistik, kdy kromě běžných parametrických přístupů (průměr, směrodatná odchylka) je vhodné uvádět charakteristiky neparametrické jako je medián nebo percentily.

Další možností je transformace dat pomocí logaritmu o vhodném základu (nejčastěji přirozený nebo dekadický). Alternativou používanou pro vizualizaci dat diverzity je užití logaritmicky škálovaných intervalů v popisných grafech (obr. 3.5).

$$y'_i = \log(b_0 + b_1 y_i), \quad (3.1)$$

kde b_0 a b_1 jsou koeficienty umožňující odstranit problém s nulovými (v případě abundancí nevýskyt taxonu) nebo zápornými hodnotami (v abundancích nenastává) transformované proměnné.

Pokud je kvantita taxonů vyjádřena abundancí, je $b_0 = 1$ a $b_1 = 1$, v případě jiného vyjádření kvantity je třeba výběru koeficientů věnovat větší pozornost, protože $b_0 \sim b_1 y_i$ způsobuje deformaci transformované proměnné.



Obr. 3.5 Popis abundancí pomocí logaritmické škály

Při srovnání diverzity mezi různými společenstvy je třeba zohlednit i rozdílnou celkovou abundanci a data před srovnáním standardizovat. Obdobný postup je také třeba zvolit před vstupem dat do vícerozměrné analýzy. Ke standardizaci je možné přistoupit dvěma způsoby:

- Přepočítání abundancí na procentuální strukturu společenstva – umožňuje vizuální srovnání křivek společenstev nebo může být použit jako vstup do vícerozměrné analýzy. Použijeme jej, pokud požadujeme, aby absolutní velikost procentuálního zastoupení taxonu v jednotlivých společenstvech byla v analýze zohledněna (v analýze se tak projeví zejména taxony s velkým průměrným podílem na společenstvu)
- Standardizace abundance každého taxonu na standardní normální rozdělení – jedná se o vhodnou úpravu dat před vícerozměrnou analýzou, kdy chceme, aby všechny taxony vstoupily do analýzy se stejnou vahou a jejich standardizované hodnoty zohledňovaly rozdíly mezi jednotlivými společenstvy. Standardizovaná data mají nulový průměr a rozptyl roven jedné. Nová hodnota (z_i) se získá odečtením střední hodnoty abundancí daného taxonu od původní hodnoty abundance taxonu ve společenstvu a dělením směrodatnou odchylkou abundancí daného taxonu.

$$z_i = \frac{y_i - \bar{y}}{s_y} \quad (3.2)$$

4 Indexy diverzity a odhady jejich statistické spolehlivosti

4.1 Indexy diverzity jako analogie popisné statistiky

Indexy diverzity mají pro data o abundancích stejný význam jako mají průměr, medián, směrodatná odchylka pro data koncentrací, nadmořských výšek atd. U indexů diverzity je třeba pečlivě zvažovat, jaká je interpretace jednotlivých indexů a zda jejich kombinací nedosáhneme informativnějších výsledků než při použití jednoho indexu nebo naopak, zda použité indexy nejsou navzájem redundantní a tedy zbytečné.

Stejně jako u běžných spojitých dat můžeme použít průměr jako ukazatel střední hodnoty a směrodatnou odchylku jako popis jejich variability. Zde můžeme využít například kombinaci počtu taxonů jako ukazatel kvalitativní stránky diverzity a Shannonovu míru vyrovnanosti jako ukazatel její kvantitativní stránky.

Indexy diverzity zároveň sdílejí i nevýhody běžných popisných statistik, tedy odstranění individuální variability taxonů a prezentace celého souboru (společenstva) jako jedno číslo, postihující pouze jeden aspekt hodnocených dat. Indexy diverzity jsou také bodovými odhady diverzity a mohou být doplněny jak intervalovými odhady, tak testováním statistické významnosti jejich rozdílů mezi společenstvy.

4.2 Indexy diverzity

Indexové hodnocení diverzity se snaží postihnout diverzitu jediným číslem. Indexy diverzity můžeme rozdělit na tři skupiny: (1) indexy založené na počtu druhů; (2) indexy založené na poměru početnosti druhů, které počítají jak s počtem druhů, tak s jejich početností; (3) Q statistika založená na tvaru křivky abundancí kumulativního počtu druhů. Pokud není uvedeno jinak, jsou popisy metod převzaty z práce [23].

4.2.1 Indexy založené na početnosti druhů

Jako nejjednodušší index diverzity je použitelný samotný počet taxonů ve společenstvu, který v sobě nese důležitou informaci o celkovém počtu nalezených druhů.

$$S - \text{počet taxonů} \quad (4.1)$$

Další indexy této skupiny váží podle různých vztahů absolutní počet druhů velikostí vzorku. Indexy této skupiny nejsou obecně příliš vhodné, protože odráží spíše proces vzorkování společenstva než jeho biodiverzitu (obr. 4.1).

Margalefův index [29]

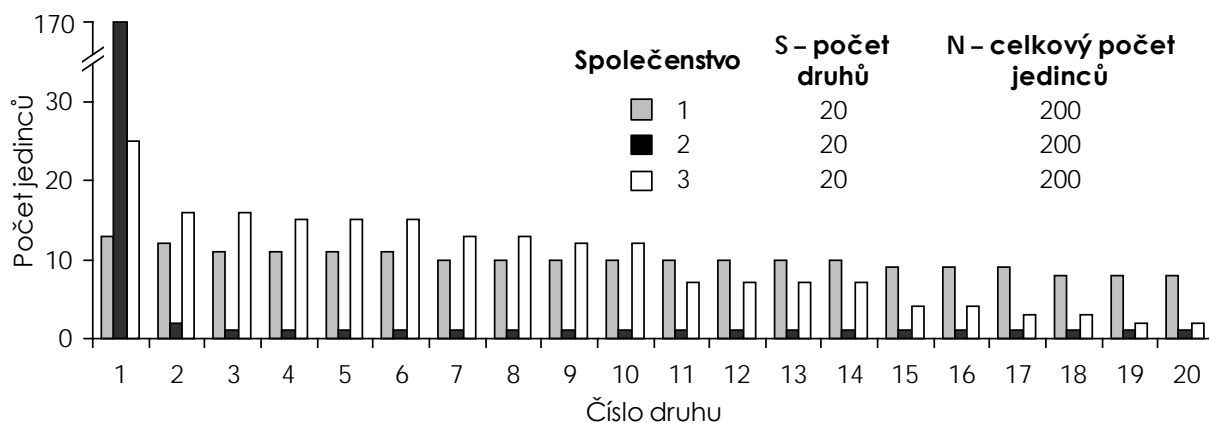
$$D_{Mg} = \frac{(S-1)}{\ln(N)}, \quad (4.2)$$

kde S je počet taxonů a N celkový počet jedinců.

Menhinickův index [30]

$$D_{Mg} = \frac{S}{\sqrt{N}}, \quad (4.3)$$

kde S je počet taxonů a N celkový počet jedinců.



	Počet druhů			Margalefův index			Menhinickův index		
	S			$D_{Mg} = \frac{(S-1)}{\ln N}$			$D_{Mn} = \frac{S}{\sqrt{N}}$		
Společenstvo	1	2	3	1	2	3	1	2	3
Výsledek	20	20	20	3.586	3.586	3.586	1.414	1.414	1.414

Obr. 4.1 Příklad výpočtu indexů založených na početnosti taxonů

4.2.2 Indexy založené na poměru početnosti druhů

Indexy založené na poměrné početnosti druhů počítají s ekvitabilitou, čili rovnoměrností zastoupení druhů, a některé zároveň s druhovou bohatostí. Nemají žádný předpoklad o modelové četnosti druhů a dají se tak považovat za neparametrické indexy. Dále je můžeme rozdělit na indexy vycházející z informační teorie (Shannonův-Weaverův a Brillouinův index) a na indexy dominance.

Shannonův - Weaverův index [31]

Tento index je indexem vycházejícím z informační teorie. Jeho předpokladem je náhodný výběr jedinců z teoreticky neomezeného množství a přítomnost všech druhů společenstva ve vzorku. Obvykle nabývá hodnot od 1,5 až 4,5. Jeho exponenciální hodnota vyjadřuje, kolik stejně početných druhů by vytvořilo Shannonův - Weaverův index o stejné hodnotě. Základní vztah pro výpočet Shannonova - Weaverova indexu je:

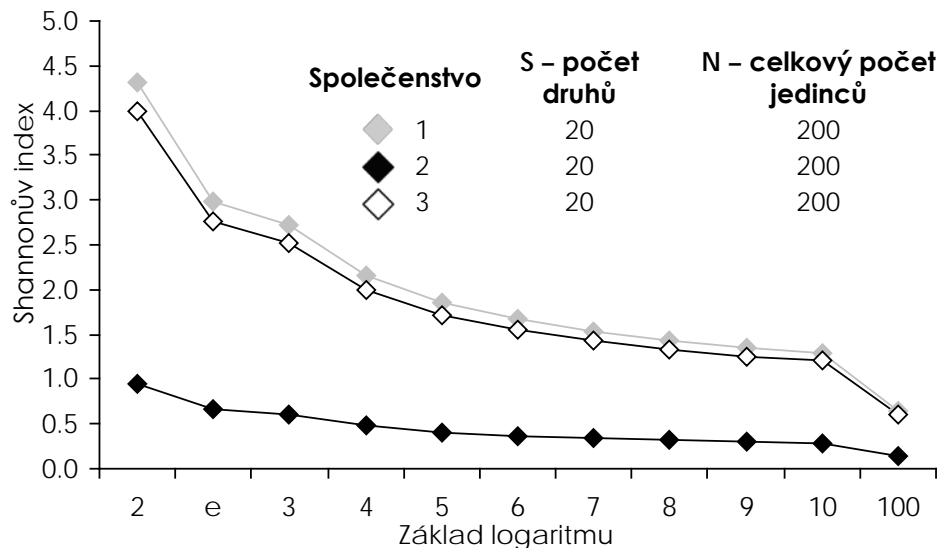
$$H = -\sum_{i=1}^S p_i \ln p_i \quad p_i = \frac{n_i}{N}, \quad (4.4)$$

kde S je celkový počet taxonů, n_i je počet jedinců i -tého druhu a N celkový počet jedinců. Podstatná je také volba základu použitého logaritmu, která ovlivňuje číselný výsledek výpočtu; není tak možné srovnávat hodnoty indexů počítané s různou bází logaritmu (obr. 4.2).

Nicméně, použití n_i/N jako odhadu p_i vytváří vychýlený odhad a proto by měla být používána korigovaná forma Shannonova - Weaverova indexu [32, 33]:

$$H = -\sum_{i=1}^S p_i \ln p_i - \frac{S-1}{N} + \frac{1 - \sum_{i=1}^S p_i^{-1}}{12N^2} + \frac{\sum_{i=1}^S (p_i^{-1} - p_i^{-2})}{12N^3}, \quad (4.5)$$

kde S je celkový počet taxonů, n_i je počet jedinců i -tého druhu a N celkový počet jedinců. Výsledek výpočtu se nicméně významně liší pouze u malých společenstev a tak je ve většině případů akceptovatelné použití základního vzorce (obr. 4.3).



Obr. 4.2 Výsledek výpočtu Shannonova- Weaverova indexu při různé bázi logaritmu

Shannonův-Waeverův index (stejně jako ostatní indexy tohoto typu) je možné vyjádřit také jako vyrovnanost společenstva. Zde je hodnota indexu vztažena na maximální možnou vyrovnanost společenstva a vyjádřena jako podíl z této maximální vyrovnanosti o možném rozsahu od nuly do jedné.

Maximální hodnota Shannonova-Weaverova indexu pro dané společenstvo odpovídá logaritmu počtu druhů a ukazuje, jaké hodnoty by index nabyl při shodné početnosti všech druhů společenstva. Hodnota tohoto indexu při maximálně vyrovnaném společenstvu je tedy dána vztahem:

$$H_{max} = -\ln S \quad (4.6)$$

Hodnota Shannonovy-Weaverovy vyrovnanosti vypovídá o poměrné hodnotě diverzity „vyčerpané“ daným společenstvem vzhledem k společenstvu se shodnou početností druhů (obr. 4.4):

$$E = \frac{H}{H_{max}} = \frac{H}{\ln S} \quad (4.7)$$

Pro Shannonův-Weaverův index je definována i jeho variabilita:

$$VarH = \frac{\sum_{i=1}^S p_i (\ln p_i)^2 - \left(\sum_{i=1}^S p_i \ln p_i \right)^2}{N} + \frac{S-1}{2N^2} \quad (4.8)$$

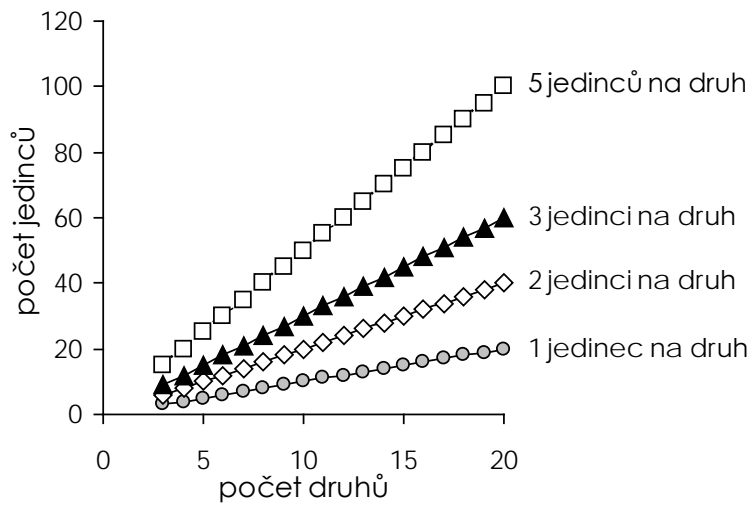
stejně jako i statistický test na bázi t-testu [33]:

$$t = \frac{H_1 - H_2}{\sqrt{VarH_1 + VarH_2}} \quad (4.9)$$

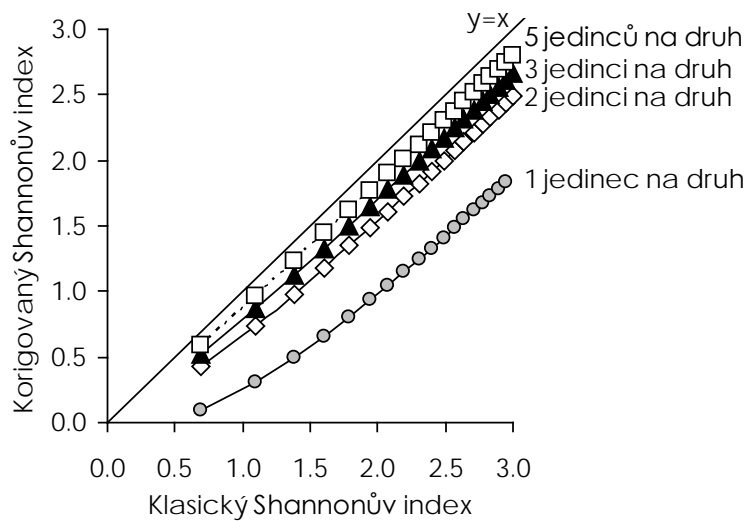
Stupně volnosti testu jsou určeny dle vzorce:

$$df = \frac{(VarH_1 + VarH_2)^2}{\frac{VarH_1^2}{N_1} + \frac{VarH_2^2}{N_2}} \quad (4.10)$$

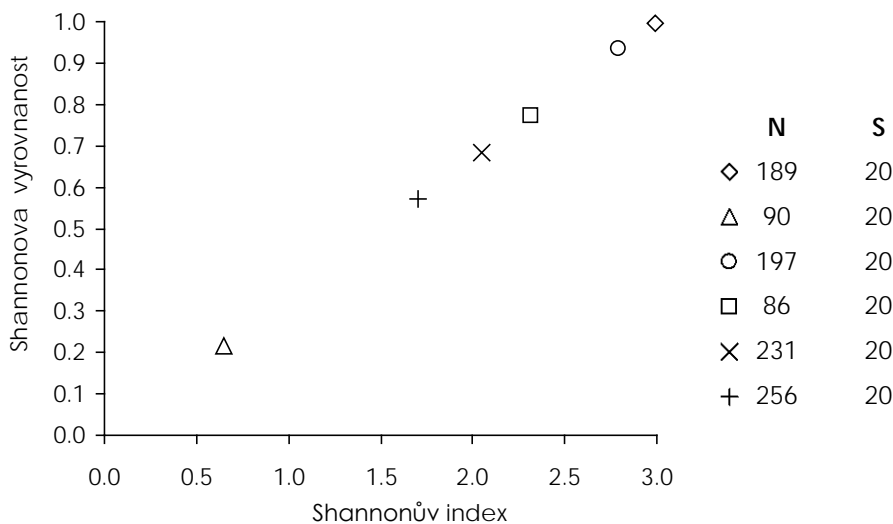
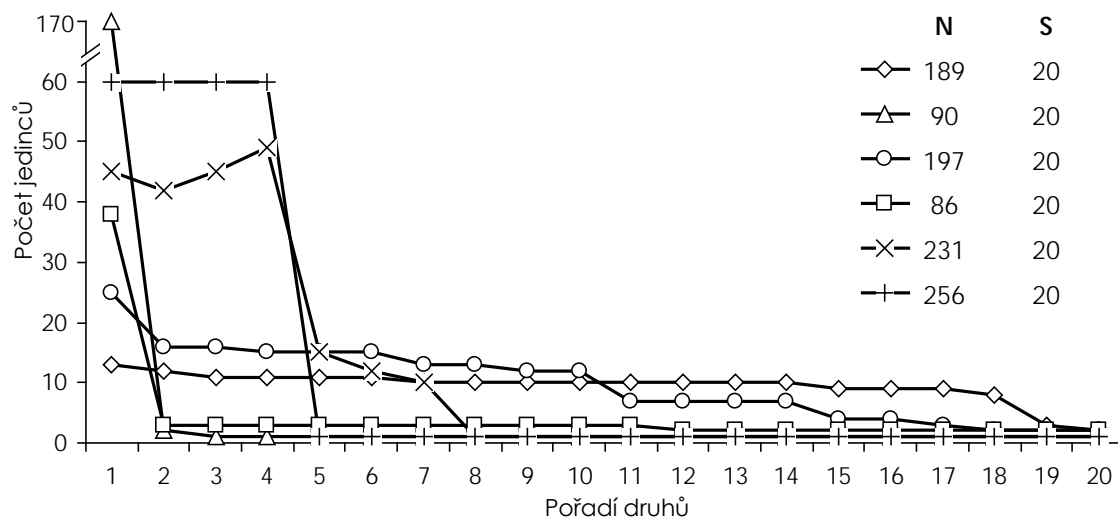
Soubory vyvážených společenstev: každý soubor obsahuje 19 společenstev, která obsahují 2-20 druhů, která mají konstantní počet jedinců (1, 2, 3, 5)



Hodnoty Shannonova indexu ve vztahu s hodnotami korigovaného Shannonova indexu. V závislosti na poměru $S-1/N$ roste velikost korekce Shannonova indexu.



Obr. 4.3 Vliv korekce na hodnotu Shannonova-Weaverova indexu



Obr. 4.4 Vztah Shannonova-Weaverova indexu a jeho vyrovnanosti při různém tvaru křivky společenstva Brillouinův index [31, 34]

V případě, že není možno zajistit náhodnost vzorkování nebo vzorky obsahují všechny členy společenstva, je tento index vhodnější náhradou Shannonova-Weaverova indexu. Ten odhaduje biodiverzitu na základě výběru ze společenstva, zatímco Brillouinův index popisuje biodiverzitu společenstva kompletně navzorkovaného (obr. 4.5). Typickým využitím je studium parazitárních infrakomunit.

$$H_B = \frac{\ln N! - \sum_{i=1}^S \ln n_i}{N}, \quad (4.11)$$

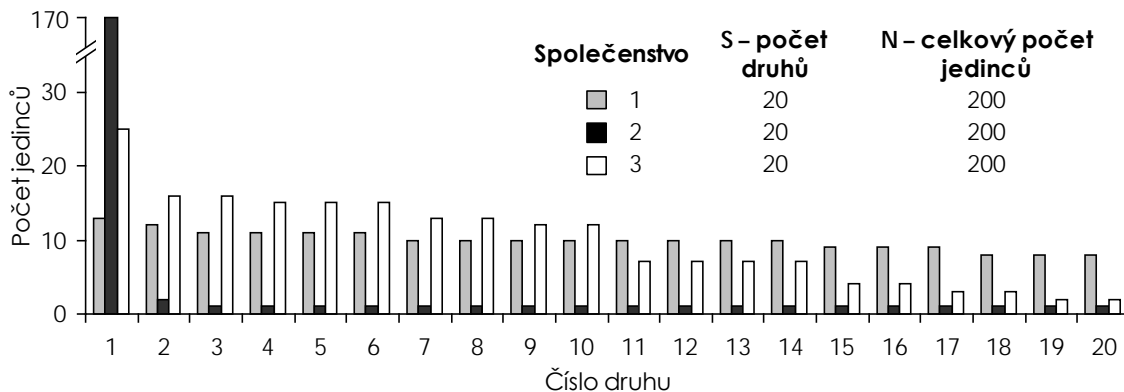
kde S je počet taxonů, n_i počet jedinců i -tého taxonu a N celkový počet jedinců.

I pro Brillouinův index je definována jeho maximální možná hodnota:

$$H_{B-max} = \frac{1}{N} \ln \frac{N!}{\left(\left[\frac{N}{S} \right]! \right)^{s-r} \left(\left(\left[\frac{N}{S} \right] + 1 \right)! \right)^r}, \quad (4.12)$$

kde $\left[\frac{N}{S} \right]$ je celá část $\frac{N}{S}$ a $r = N - S \left[\frac{N}{S} \right]$. Vyrovnanost Brillouinova indexu je následně počítána jako:

$$E_B = \frac{H_B}{H_{B-max}}. \quad (4.13)$$



Shannonův index		Vyrovnanost Shannonova indexu						Brillouinův index				
$H' = -\sum p_i \ln p_i$		$E = \frac{H'}{H_{max}} = \frac{H'}{\ln S}$						$HB = \frac{\ln N! - \sum \ln n_i!}{N}$				
$p_i = \frac{n_i}{N}$												
n_i - je počet jedinců i-tého druhu		H_{max} - maximální hodnota Shannonova indexu						n_i - je počet jedinců i-tého druhu				
	n_i			p_i			$\ln(p_i)$			$\ln(n_i!)$		
	1	2	3	1	2	3	1	2	3	1	2	3
Druh 20	8	1	2	0.040	0.005	0.010	-3.219	-5.298	-4.605	10.605	0.000	0.693
Druh 19	8	1	2	0.040	0.005	0.010	-3.219	-5.298	-4.605	10.605	0.000	0.693
Druh 18	8	1	2	0.040	0.005	0.010	-3.219	-5.298	-4.605	10.605	0.000	0.693
Druh 17	9	1	3	0.045	0.005	0.015	-3.101	-5.298	-4.200	12.802	0.000	1.792
Druh 16	9	1	4	0.045	0.005	0.020	-3.101	-5.298	-3.912	12.802	0.000	3.178
Druh 15	9	1	4	0.045	0.005	0.020	-3.101	-5.298	-3.912	12.802	0.000	3.178
Druh 14	10	1	7	0.050	0.005	0.035	-2.996	-5.298	-3.352	15.104	0.000	8.525
Druh 13	10	1	7	0.050	0.005	0.035	-2.996	-5.298	-3.352	15.104	0.000	8.525
Druh 12	10	1	7	0.050	0.005	0.035	-2.996	-5.298	-3.352	15.104	0.000	8.525
Druh 11	10	1	7	0.050	0.005	0.035	-2.996	-5.298	-3.352	15.104	0.000	8.525
Druh 10	10	1	12	0.050	0.005	0.060	-2.996	-5.298	-2.813	15.104	0.000	19.987
Druh 9	10	1	12	0.050	0.005	0.060	-2.996	-5.298	-2.813	15.104	0.000	19.987
Druh 8	10	1	13	0.050	0.005	0.065	-2.996	-5.298	-2.733	15.104	0.000	22.552
Druh 7	10	1	13	0.050	0.005	0.065	-2.996	-5.298	-2.733	15.104	0.000	22.552
Druh 6	11	1	15	0.055	0.005	0.075	-2.900	-5.298	-2.590	17.502	0.000	27.899
Druh 5	11	1	15	0.055	0.005	0.075	-2.900	-5.298	-2.590	17.502	0.000	27.899
Druh 4	11	1	15	0.055	0.005	0.075	-2.900	-5.298	-2.590	17.502	0.000	27.899
Druh 3	11	1	16	0.055	0.005	0.080	-2.900	-5.298	-2.526	17.502	0.000	30.672
Druh 2	12	2	16	0.060	0.010	0.080	-2.813	-4.605	-2.526	19.987	0.693	30.672
Druh 1	13	170	25	0.065	0.850	0.125	-2.733	-0.163	-2.079	22.552	758.248	58.004

Výsledky								
Shannonův index			Vyrovnanost Shannonova indexu			Brillouinův index		
Společenstvo			Společenstvo			Společenstvo		
1	2	3	1	2	3	1	2	3
2.988	0.661	2.767	0.997	0.221	0.924	2.798	0.521	2.654

Obr. 4.5 Výpočet Shannonova- Weaverova a Brillouinova indexu

Simpsonův index [35]

Jde o nejznámější index skupiny nazývané indexy založené na dominanci. Je silně závislý na nejpočetnějším druhu a méně citlivý ke vzácným druhům. Může nabývat hodnot od nuly do jedné. Hodnota indexu silně záporně koreluje s vyrovnaností, jak Shannonova-Weaverova, tak Brillouinova indexu.

S jeho zvyšující se hodnotou stoupá dominance a klesá vyrovnanost společenstva, proto se často používá jeho převrácená hodnota nebo odpočet od jedné; v případě interpretace publikovaných výsledků je vždy nezbytné ověřit, v jaké formě byl tento index použit. Vztah mezi tímto indexem a počtem druhů, pro vzorky s více než 10 druhů, je silně závislý na rozložení abundancí druhů (na modelech rozdělení abundancie taxonů, tedy species abundance models) ve vzorku [36]. Výpočet základní verze nabývající nejvyšších hodnot při vysoké dominanci a nejnižších při vyrovnaném společenstvu je dán vztahem:

$$D = \sum_{i=1}^S \frac{n_i(n_i - 1)}{N(N - 1)}, \quad (4.14)$$

kde S je počet taxonů, n_i počet jedinců i -tého taxonu a N celkový počet jedinců.

Bergerův-Parkerův index [36, 37]

Bergerův-Parkerův index vyjadřuje poměrnou významnost nejpočetnějšího druhu. Stejně jako u Simpsonova indexu se často používá jeho převrácená hodnota nebo odpočet od jedné. Jeho možný rozsah je od nuly do jedné. Může nabývat podobných hodnot jako Simpsonův index. Je nezávislý na počtu druhů, ale je ovlivněn velikostí vzorku. Vzhledem k tomu, že využívá jiný princip výpočtu než ostatní indexy a nezahrnuje abundancie všech taxonů, je při analýze vhodným doplňkem indexů využívajících abundancie všech taxonů.

$$BP = \frac{N_{max}}{N}, \quad (4.15)$$

kde N_{max} je abundancie nejpočetnějšího taxonu a N celkový počet jedinců ve vzorku.

4.2.3 Q statistika

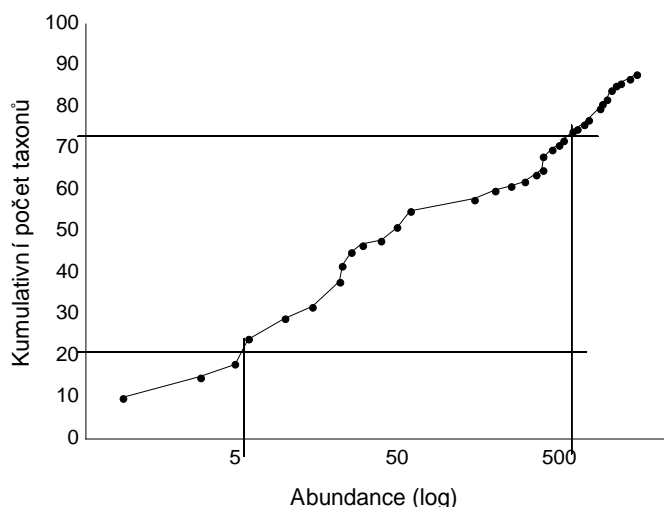
Q statistika je index diverzity založený na měření sklonu křivky abundancí kumulativního počtu druhů (obr. 4.6). Může být ovlivněna malou velikostí vzorku, ale pokud je ve vzorku obsaženo více než 50% druhů, je toto ovlivnění jen malé [38]. V původní podobě [39, 40] je Q statistika měření mezikvartilového úseku křivky početnosti kumulativního počtu druhů, které poskytuje měření diverzity společenstva, při kterém nejsou uvažovány ani velmi četné ani velmi vzácné druhy. V novější formě [41] je v podstatě využito simulace všech možných podspolečenstev daného společenstva k vytvoření rozložení možných Q pro dané společenstvo. Výsledkem tedy není jediná hodnota, ale rozložení hodnot, které lze jednoduše statisticky srovnávat mezi různými společenstvy.

Původní varianta Q statistiky pro mezikvartilový rozsah taxonů společenstva je vypočtena pomocí vztahu:

$$Q = \frac{\frac{1}{2}n_{R_1} + \sum_{r=R_1+1}^{R_2-1} n_r + \frac{1}{2}n_{R_2}}{\ln\left(\frac{R_2}{R_1}\right)}, \quad (4.16)$$

kde $\sum n_r$ je celkový počet taxonů mezi kvartily, S celkový počet taxonů ve vzorku, R_1 a R_2 jsou dolní a horní kvartil, n_{R_1} je počet druhů ve třídě, do níž spadá dolní kvartil, n_{R_2} je počet druhů ve třídě, do níž spadá horní kvartil, R_1 označuje počet jedinců ve třídě, do níž spadá dolní kvartil a R_2 počet jedinců ve třídě, do níž spadá horní kvartil. Kvartily jsou určeny:

$$\sum_1^{R2-1} n_r < \frac{3}{4} S \leq \sum_1^{R2} n_r \quad \text{a} \quad \sum_1^{R1-1} n_r < \frac{1}{4} S \leq \sum_1^{R1} n_r. \quad (4.17)$$



Obr. 4.6 Zobrazení kumulativního počtu druhů proti ose jejich logaritmované abundance, tedy strmost narůstání počtu druhů se stoupající abundancí. Q statistika je počítána ze sklonů spojníc jednotlivých bodů v grafu, popřípadě ze sklonu interkvartilového rozsahu.

Q statistika však může být počítána stejnou metodou také pro celé spektrum taxonů nebo lze její hodnotu odhadnout s pomocí výpočtu:

$$X_i = \frac{S_j - S_{j'}}{\ln\left(\frac{N_j}{N_{j'}}\right)}, \quad (4.18)$$

kde S je kumulativní počet druhů, N počet jedinců ve třídě a r počet tříd. Výpočet je proveden pro všechny páry S_j a $S_{j'}$ a N_j a $N_{j'}$ ($j > j'$, $j = 1, 2, \dots, r$), přičemž počet výpočtů je roven:

$$i = \frac{r(r-1)}{2}. \quad (4.19)$$

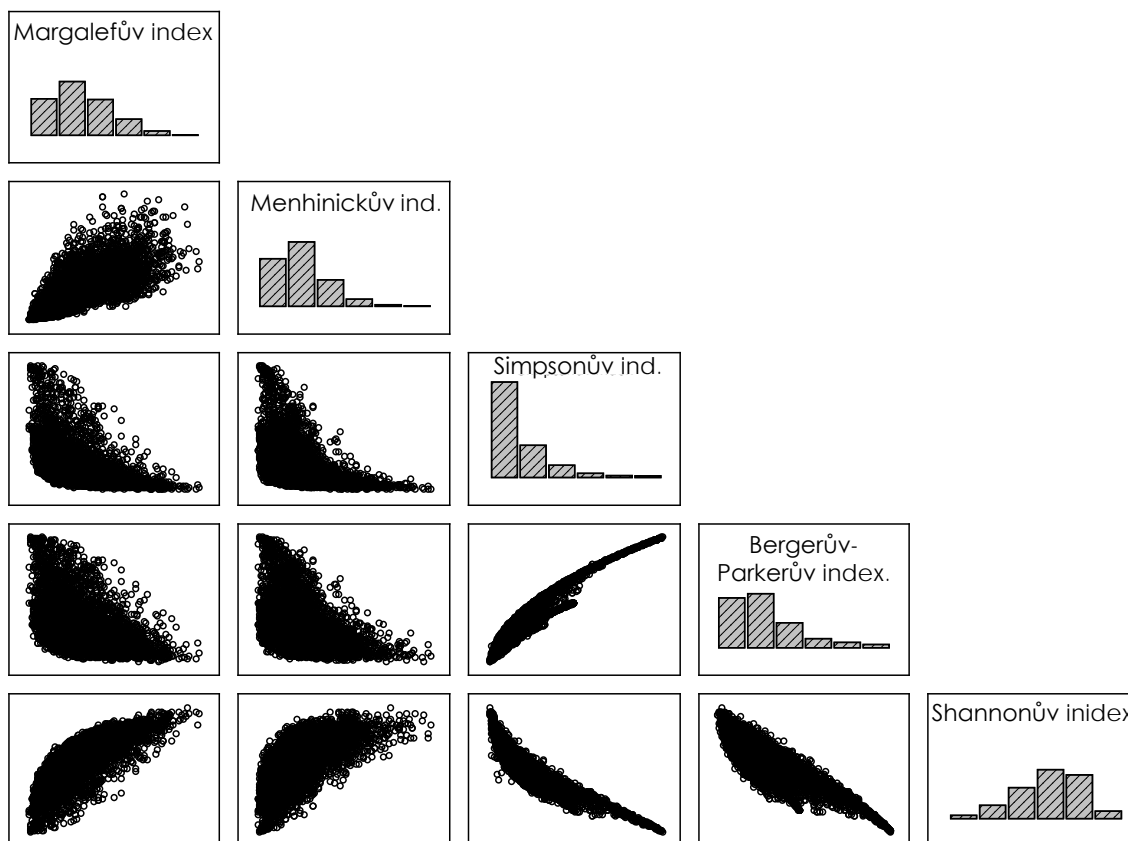
Hodnota Q statistiky se pak spočítá jako medián nebo geometrický průměr vypočítaných hodnot X_i [41]. Vzhledem k tomu, že výsledkem je rozložení hodnot, je možné toto měřítko diverzity jednoduše testovat mezi různými společenstvy, například pomocí neparametrických testů.

4.2.4 Vztahy mezi indexy biodiverzity

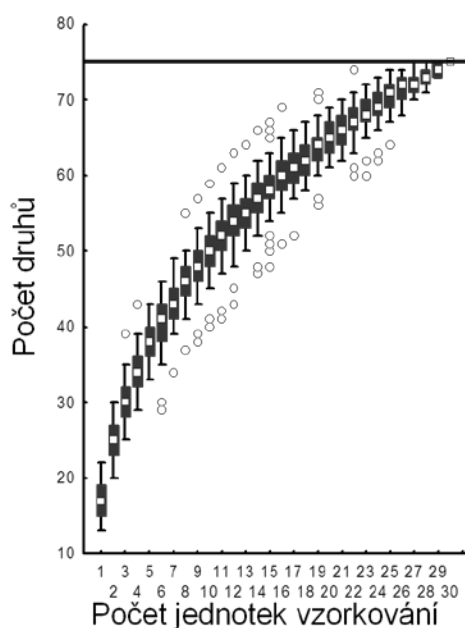
Indexy biodiverzity představují popis jedné vlastnosti živé přírody. Z tohoto důvodu mezi nimi existují různě silné korelační vztahy (obr. 4.7). Při využití indexů biodiverzity v analýze je vhodné vybírat pouze nekorelované indexy, které přinášejí nezávislou informaci o hodnoceném společenstvu.

4.2.5 Velikost vzorku a indexy biodiverzity – rarefakce

Diverzita, stejně jako jakýkoliv jiný výzkum, je závislá na intenzitě a kvalitě vzorkování. Zejména při terénním odběru vzorků se často stává, že společenstva získaná na různých lokalitách si i při stejné metodice sběru neodpovídají svojí početností. Lze tak předpokládat, že naměřená druhová bohatost neodráží pouze rozdíly mezi lokalitami, ale také velikost vzorků společenstev. Toto tvrzení vychází z faktu, že počet druhů nelineárně závisí na počtu jedinců ve vzorku [42] (obr. 4.8).



Obr. 4.7 Vzájemné vztahy vybraných indexů biodiverzity



Obr. 4.8 Vztah mezi velikostí vzorku a počtem detekovaných druhů ve studii monitoringu makrozoobentosu [43]

Vztah mezi velikostí vzorku ve smyslu počtu jedinců a počtem zjištěných druhů lze matematicky popsat. Tento popis lze dále využít pro odhad počtu druhů zjištěných při menší velikosti vzorku než je hodnocené společenstvo. Nejběžnější metodou odhadu je takzvaná rarefakce [44], což je metoda řešící problém srovnání druhové bohatosti ve vzorcích o různé velikosti standardizací obou vzorků na velikost menšího vzorku.

$$E(S_n) = \sum_{i=1}^S \left[1 - \frac{\binom{N-N_i}{n}}{\binom{N}{n}} \right], \quad (4.20)$$

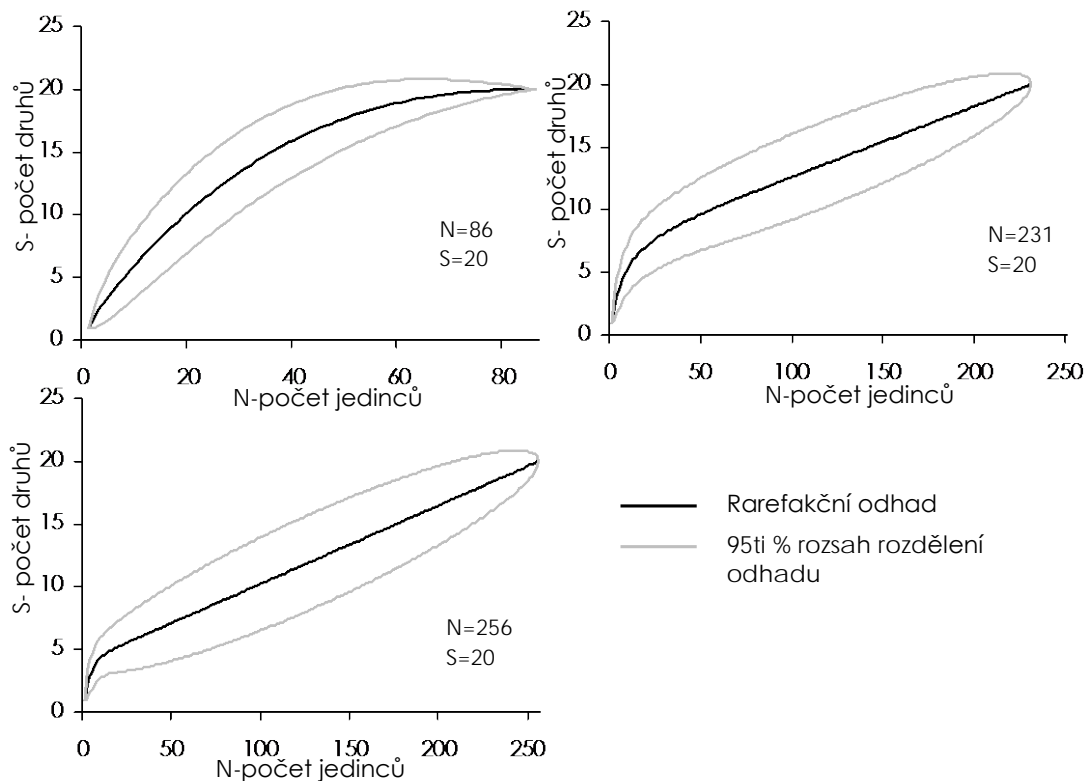
kde $E(S_n)$ je odhadnutý počet druhů ve vzorku o n jedincích, N je celkový počet jedinců, N_i představuje počet jedinců druhu i a n je počet jedinců, pro které je odhad počítán. Variabilita odhadu je dána vztahem:

$$\text{var}(S_n) = \binom{N}{n}^{-1} \left\{ \sum_{i=1}^S \binom{N-N_i}{n} \left[1 - \frac{\binom{N-N_i}{n}}{\binom{N}{n}} \right] + 2 \sum_{i=1}^{S-1} \sum_{j=i+1}^S \left[\binom{N-N_i-N_j}{n} - \frac{\binom{N-N_i}{n} \binom{N-N_j}{n}}{\binom{N}{n}} \right] \right\} \quad (4.21)$$

Rarefakci lze považovat za parametrickou metodu s předpokladem modelu vztahu mezi počtem jedinců a počtem druhů a jako taková má následující předpoklady.

- Společenstva porovnávaná pomocí rarefakce mají mít podobné taxonomické složení.
- Srovnatelné metody vzorkování.
- Rarefakční křivka nemůže být extrapolována za sumu jedinců největšího vzorku.
- Jedinci jsou ve společenstvu náhodně rozmístěni (ve skutečnosti jsou častější shluky jedinců; za těchto podmínek dává rarefakce nadsazené odhady).

Příklad výpočtu rarefakce pro různé počty jedinců a druhů a výsledek jejich odhadu je znázorněn na obrázku 4.9.



Obr. 4.9 Rarefakční odhad pro různé kombinace počtu jedinců a druhů ve společenstvu

4.3 Odhad intervalů spolehlivosti a statistické testování biodiverzitních indexů

Odhad intervalů spolehlivosti pro diverzitní indexy je standardně prováděno pomocí permutačních metod *bootstrap* nebo *jackknife*. Protože obě metody mají obdobný princip, detailněji je v následujícím textu popsán pouze *jackknife* odhad.

Jackknife (název skutečně odkazuje na Jacka rozparovače, kdy v průběhu výpočtu je rozřezáván původní datový soubor) byl na data o biodiverzitě poprvé použit v roce 1977 [45] a obecně je metodika popsána Sokalem a Rohlfem [46]. Metoda nemá žádné předpoklady o rozdělení dat. Namísto toho generuje řadu takzvaných „pseudohodnot“, které mají obvykle normální rozdělení a jejich průměr je nejlepším odhadem dané statistiky; zároveň je možné doplnit i interval spolehlivosti odhadu.

V prvním kroku je odhadnuta diverzita (na základě libovolného indexu diverzity) pro celý vzorek n hodnot (St). Následně je výpočet diverzity n -krát přepočítán, přičemž v každém kroku je ze vzorku odstraněna (odříznuta) jedna hodnota. Každý přepočítání produkuje nový odhad diverzity St_{-i} . „Pseudohodnota“ (ϕ_i) může být následně vypočtena pro každý z n vygenerovaných „podvzorků“.

$$\phi_i = nSt - (n-1)St_{-i} \quad (4.22)$$

Jackknife odhad indexu diverzity je spočítán jako průměr těchto pseudohodnot:

$$\phi = \frac{\sum \phi_i}{n} \quad (4.23)$$

Střední chybu provedeného odhadu vypočteme dle vztahu:

$$S.E.\phi = \sqrt{\frac{\sum (\phi_i - \phi)^2}{n(n-1)}} \quad (4.24)$$

Výpočet intervalu spolehlivosti je počítán dle standardního vztahu

$$\phi \pm t_{0.05}^{n-1} S.E.\phi \quad (4.25)$$

V případě proměnných s omezeným rozsahem (např. 0-1), mezi které spadá řada biodiverzitních indexů, je před výpočtem doporučována transformace [46], například Fisherova z-transformace. Metoda je citlivá na odlehle hodnoty, protože ty jsou zahrnuty do výpočtu původní hodnoty St . Výsledky je samozřejmě třeba hodnotit vzhledem k jejich biologické relevantnosti.

Bootstrapping je obdobnou metodou s vyšší náročností na výpočet, nicméně s předpokladem dosažení lepších výsledků. Základem výpočtu je náhodné opakované vzorkování původního datového souboru, za účelem vytvoření velkého množství kombinací původních pozorování; metodu detailně popisují Sokal a Rohlf [46].

Také testování statistické významnosti rozdílů v hodnotách diverzitních indexů je obvykle prováděno pomocí permutačních metod *bootstrap* nebo *jackknife*. Odhad z *jackknife/bootstrap* procedury a jeho střední chyba odhadu vstupují do t-testu s $n-1$ stupni volnosti [47].

Výjimkou, kdy je k dispozici asymptotický odhad, je Shannonův-Weaverův index, jehož variabilita je dána vztahem [33]:

$$Var(H) = \frac{\sum_{i=1}^S p_i (\ln p_i)^2 - \left(\sum_{i=1}^S p_i \ln p_i \right)^2}{N} - \frac{S-1}{2N^2} \quad (4.26)$$

kde S je celkový počet taxonů, n_i je počet jedinců i -tého druhu a N celkový počet jedinců. Hodnota testové statistiky t je vypočtena:

$$t = \frac{H_1 - H_2}{\sqrt{VarH_1 + VarH_2}} \quad (4.27)$$

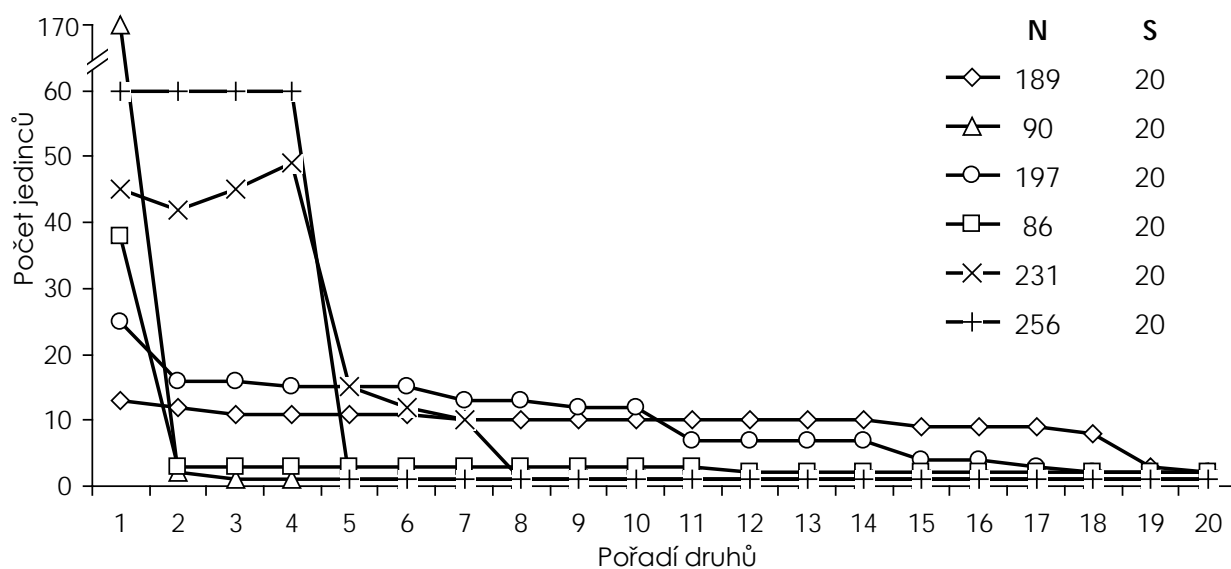
Stupně volnosti testu jsou určeny dle vztahu:

$$df = \frac{(VarH_1 + VarH_2)^2}{\frac{VarH_1^2}{N_1} + \frac{VarH_2^2}{N_2}} \quad (4.28)$$

Příklad výpočtu srovnávajícího společenstva s různou mírou vyrovnanosti je v tabulce 4.1.

Tabulka 4.1 Příklad testování statistické významnosti rozdílů Shannona-Weaverova indexu mezi společenstvy s různou mírou vyrovnanosti

	Vyrovnané společen.	◇	○	□	×	+	△
Vyrovnané společen.	2.996	0.082	<0.001*	<0.001*	<0.001*	<0.001*	<0.001*
◇	0.0477	2.948	0.002*	<0.001*	<0.001*	<0.001*	<0.001*
○	0.2019	0.154	2.794	0.007	<0.001*	<0.001*	<0.001*
□	0.6770	0.629	0.475	2.319	0.222	0.001	<0.001*
×	0.9462	0.898	0.744	0.269	2.050	<0.001*	<0.001*
+	1.2890	1.241	1.087	0.612	0.343	1.707	<0.001*
△	2.3512	2.303	-2.149	-1.674	-1.405	-1.062	0.645



4.4 Biotické indexy

Za speciální skupinu indexů diverzity lze považovat takzvané biotické indexy, které do výpočtu indexu kromě přítomnosti a abundance taxonů přidávají i jejich ekologické charakteristiky (*species traits*). Výsledný index pak lze popsat jako souhrnné hodnocení společenstva vycházející z výskytu a abundance taxonů vážené jejich ekologickými charakteristikami. Tento přístup je zároveň výhodou i nevýhodou, kdy na jedné straně umožňuje přímou ekologickou interpretaci výsledků, ale na straně druhé je závislý na dostupnosti detailních datových podkladů o jednotlivých taxonech.

Za nejznámější index tohoto typu lze považovat **saprobní index** [48], který je určen k hodnocení znečištění vod na základě druhové skladby společenstva podle:

$$S_i = \frac{\sum_{i=1}^S s_i h_i i_i}{\sum_{i=1}^S h_i i_i}, \quad (4.29)$$

kde s_i je individuální saprobní index i -tého druhu, h_i je jeho početnost a i_i jeho individuální indikační váha.

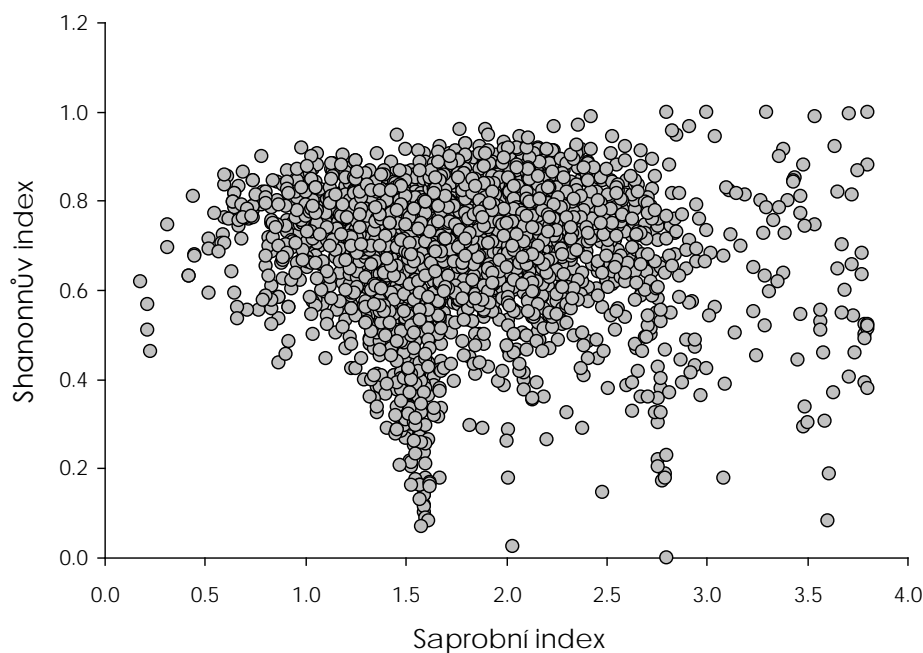
Mírou spolehlivosti saprobního indexu je jeho disperze [49, 50], což je hodnota udávající spolehlivost saprobního indexu založená na skutečnosti, zda se ve vzorku vyskytují druhy s obdobnými (spolehlivá hodnota indexu) nebo rozdílnými saprobiologickými vlastnostmi:

$$D = \sqrt{\frac{\sum_{i=1}^S (s_i - S_i)^2 h_i i_i}{N - \sum_{i=1}^S h_i i_i}}, \quad (4.30)$$

kde s_i je individuální saprobní index i -tého druhu, h_i je jeho početnost a i_i jeho individuální indikační váha. Při hodnotě disperze převyšující 0,2 je saprobní index považován za nespolehlivý.

Na obdobném principu existuje celá řada indexů. Jejich hlavním rozdílem je taxonomická úroveň a skupina, na níž je vlastnost (váha) taxonu číselně vyjádřena, dále aspekt společenstva hodnocený danou vlastností (saprobita, kyselost, citlivost k toxickým látkám, potravní strategie apod.) a výpočet na bázi přítomnosti nebo abundance taxonů.

Biotické indexy přináší do analýzy biodiverzity další rozměr. Zatímco „čisté“ biodiverzitní indexy popisují míru uspořádanosti a informační bohatost systému, biotické indexy odráží funkční stav společenstva (obr. 4.10).



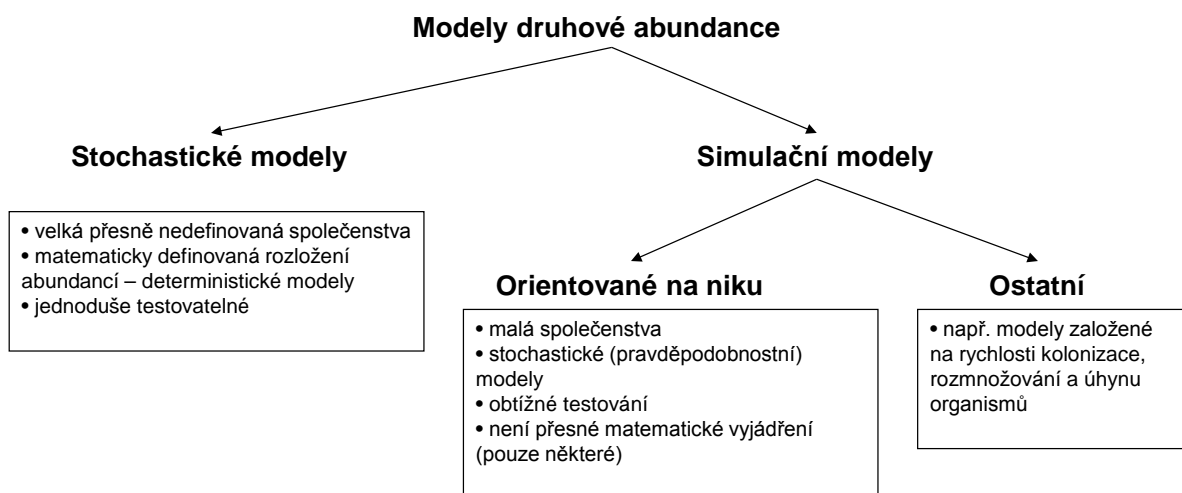
Obr. 4.10 Vztah mezi hodnotou Shannonova-Weaverova a saprobního indexu v datech biomonitoringu makrozoobentosu

5 Modely druhové abundance a stochastické modely

5.1 Modely druhové abundance jako analogie prokládání statistických rozložení

Modely rozložení abundance taxonů (početnost, biomasa, aktivita) představují snahu o kategorizaci společenstev podle typizovaných průběhů křivek abundancí. Tyto typizované průběhy by měly mít i určitou biologickou interpretaci. Cílem zhodnocení křivky průběhu druhů tedy není pouze její přiřazení k určitému modelu, ale především biologická interpretace. Tedy teoreticky by mělo platit, že pokud můžeme přiřadit společenstvo k modelu, máme představu, jakou cestou společenstvo vzniklo, jaké vlivy se mohly podílet na jeho utváření a současném stavu apod.

V tvorbě modelů se odráží dva hlavní přístupy: (1) přiřazení průběhu křivky abundancí k běžným matematickým funkcím, jejichž průběh může být interpretován i z hlediska reálného rozložení abundancí a (2) myšlenková konstrukce chování druhů ve společenstvu je převáděna do podoby matematického modelu, který je testován oproti reálným datům. Vlastní model je často definován ne striktně matematicky, ale pouze pravděpodobnostně a jeho průběh je vytvářen počítačovou simulací (základní rozdělení modelů na obrázku 5.1).



Obr. 5.1 Funkční rozdělení modelů druhové abundance

Modely druhové abundance tedy můžeme rozdělit na simulačně založené (orientované a neorientované na niku) a na modely stochasticky založené [22]. Hranice těchto modelů ovšem není ostře vedena, protože mnohé z těchto modelů jsou zároveň simulačními i stochastickými modely nebo byly původně prezentovány jako zástupci jiné skupiny modelů. Přehled modelů a jejich rozdělení do skupin obsahuje tabulka 5.1. Stochastické modely jsou obecně zaměřeny na velká společenstva, kdy se uplatňují obecné biologické a matematické zákonitosti. Simulační modely jsou naproti tomu zaměřeny na malá homogenní společenstva a vliv biologických jevů a vlastností taxonů na lokální úrovni.

Tabulka 5.1 Přehled modelů rozložení abundancí podle [22], anglické názvosloví zachováno pro snazší orientaci vzhledem v literatuře

Typ modelu	Model	Autor
Stochastické modely	Logaritmická řada (<i>Logaritmick series</i>)	[51]
	Log normální (<i>Log normal</i>)	[52, 53]
	Negativní binomické (<i>Negative binomial</i>)	[54, 55]
	Zipfův-Mandelbrotův model (<i>Zipf-Mandelbrot</i>)	[56, 57]
Simulační na niku orientované modely	Geometrická řada (<i>Geometric series</i>)	[58]
	Částicová nika (<i>Particulate niche</i>)	[59]
	Překryv nik (<i>Overlapping niche</i>)	[59]
	Zlomená hůlka (<i>Broken stick</i>)	[59]
	MacArthurova frakcionace (<i>MacArthur fraction</i>)	[21]
	Předpoklad dominance (<i>Dominance pre-emption</i>)	[21]
	Náhodná frakcionace (<i>Random fraction</i>)	[21]
	Sugiharův model postupného dělení (<i>Sugihara's sequential breakage</i>)	[60]
	Odmítnutí dominance (<i>Dominance decay</i>)	[21]
Náhodné roztržení (<i>Random assortment</i>)	[21]	
Složený model (<i>Composite model</i>)	[21]	
Jiné simulační modely	Dynamický model (<i>Dynamic model</i>)	[61, 62]
	Neutrální model (<i>Neutral model</i>)	[63]

Na proces výpočtu parametrů modelu pro daná reálná data navazuje testování shody modelu s těmito daty. Nejjednodušším postupem je optická kontrola průběhu modelu a reálných dat v některém typu grafu, tento postup je však silně subjektivní. Častou možností je použití χ^2 testu dobré shody pro otestování počtu druhů v kategoriích abundance, kdy očekávané hodnoty získáme z průběhu modelu a pozorované hodnoty na základě reálných dat. Nevýhodou je jeho použití pouze pro velká společenstva (vliv síly testu), která ovšem většinou nepředstavují homogenní společenstva. Jedná se tedy o test s obdobným použitím a omezením jako mají matematické modely. Test není vhodný pro malá společenstva. Dalším testem vhodným pro deterministické modely je Kolmogorův-Smirnovův test, který je však používán méně často. Jeho výhodou oproti χ^2 test je použitelnost i pro velmi malá společenstva. Dalším způsobem je využití vícerozměrné vzdálenosti reálného a modelového společenstva, jako je například Hellingerova vzdálenost D^2 [22]. Vzdálenost je definovaná jako

$$D_{ab}^2 = \sum \left[\sqrt{P_a(i)} - \sqrt{P_b(i)} \right]^2, \quad (5.1)$$

kde $P_a(i)$ a $P_b(i)$ jsou abundance třídy i ve vzorku a (pozorování) a b (teorie). Tento způsob byl použit u geometrické řady, log-normálního rozložení, a modelu zlomené hůlky. Pro rutinní hodnocení je ovšem tento přístup špatně použitelný vzhledem k obtížnému testování významnosti.

Pro testování stochastických modelů vztahujících se na niku není možné použít χ^2 ani Kolmogorův-Smirnovův test, protože se porovnává teoretický odhad vzniklý z velké množiny opakování s pozorováním, která představuje jedno reálné společenstvo. Proto byl [21] uveden test, který by tento problém měl vyřešit. Z naměřených dat je spočítána průměrná abundance x_i pro třídu $i = 1$ (třída s nejvyšší abundancí) až $i = S$ (třída s nejnižší abundancí), kde S je počet druhů. Pro naměřená data jsou nutná opakovaná pozorování, tedy jedna reálná lokalita zde představuje více provedených měření, které by ovšem měly mít obdobné druhové složení. Problém postupu je v počtu druhů S . Pokud se během opakovaného pozorování mění počet druhů, musí být

S natolik velké, aby pokrylo nejméně 95% druhů sledovaného společenstva v případě, že počet druhů ve vzorku vzroste. Pomocí počítačové simulace je vytvořen velký počet společenstev (N minimálně 999 [64]) obsahujících S druhů, které představují rozložení nasimulovaných modelových abundancí jednotlivých druhů a jsou základem pro další výpočet. Z nasimulovaného rozložení je vypočten průměr X_i a odchylka σ_i pro třídy $i = 1$ až S a vypočteny teoretické intervaly spolehlivosti

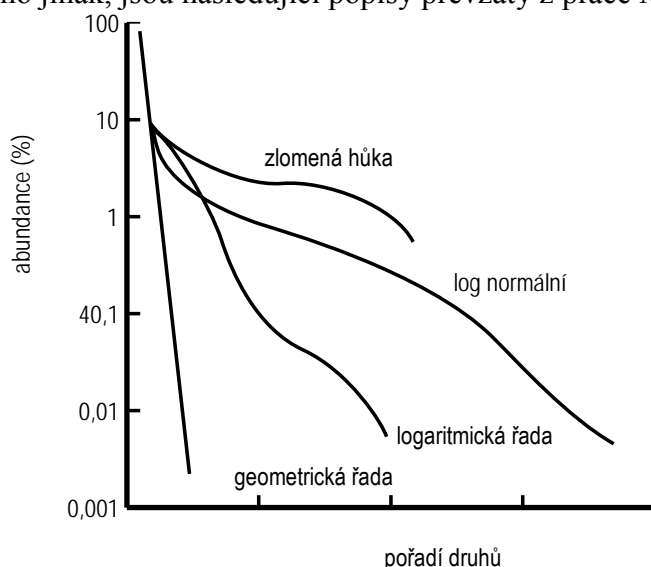
$$IS(X_i): X_i \pm r \frac{\sigma_i}{\sqrt{n}}, \quad (5.2)$$

kde $r = 1,96$ pro 95% a $1,65$ pro 90% interval spolehlivosti jsou odvozeny ze standardního normálního rozložení. Teoretické hodnoty jsou následně porovnány s pozorovaným průměrem x_i . Pokud všechny hodnoty x_i spadají do intervalů spolehlivosti nasimulované průměrné abundance, dá se říci, že pozorovaný tvar křivky druhové abundance je ve shodě s očekáváním modelu.

Složitější a výpočetně náročnou metodou testování je metoda Monte Carlo [65]. Tato metoda by měla řešit nedostatky předchozího testu: citlivost na počet druhů v testovaném společenstvu a problémy s heterogenitou dat a modelu [64]. Principem je vygenerování teoretického rozložení průměrů a odchylek pro každou třídu daného modelu.

5.2 Stochastické modely

Stochastické modely jsou odvozeny z obecných matematických rozložení, jsou matematicky definovány a existuje pro ně postup výpočtu parametrů rozložení. Vzhledem k jejich matematickým vlastnostem a možnostem testování jsou primárně zaměřeny na velká společenstva a obecné zákonitosti (zákon velkých čísel, výskyt vzácných jevů apod.). Obecný průběh čtyř základních stochastických modelů zobrazuje obrázek 5.2. Modely jsou detailně popsány v následujících kapitolách. I když jsou s těmito modely spjaty různé biologické interpretace, jejich platnost není všeobecně přijímána. Například velmi často popisovaný přechod rozložení abundancí od log-normálního k logaritmickému pod vlivem znečištění není vzhledem k nejednoznačnému proložení těchto typů modelů interpretovatelný z hlediska trendů ve změně struktury společenstva [22, 23]. Pokud není uvedeno jinak, jsou následující popisy převzaty z práce Magurran [23].



Obr. 5.2 Průběh vybraných stochastických modelů

5.2.1 Geometrická řada [58]

Tento model je shodný s na niku orientovaným modelem geometrické řady. Jedná se o jeden z modelů, který je popsán jak matematicky, tak modelem dělení niky.

Předpokladem tohoto modelu je využití nadpoloviční poměrné části celkové niky (k) prvním druhem, následné využití stejného poměru ze zbylé niky dalším druhem atd. Hodnota k je shodná pro všechny druhy a její odchylky jsou dány pouze chybou vzorkování. Každý druh tak obsazuje větší část niky než všechny následující druhy dohromady. Při vykreslení do grafu pořadí druhů proti jejich abundanci můžeme tímto grafem proložit lineární regresi. Model byl ověřen v případě malých společenstev. Zde však může jít pouze o výsledek malého počtu datových bodů (druhy v grafu pořadí/abundance), kde se projevují pouze malé odchylky od lineární regrese než v případě velkých společenstev s mnoha druhy.

Pokud uvažujeme, že proporce celkové niky obsazené jednotlivými druhy odpovídá jejich poměru abundancí, můžeme vypočítat očekávané abundance jednotlivých druhů společenstva pomocí vztahu [36, 58]:

$$n_i = NC_k k(1-k)^{i-1}, \quad (5.3)$$

kde n_i je počet jedinců i -tého druhu, N celkový počet jedinců a

$$C_k = [1 - (1-k)^S]^{-1} \quad (5.4)$$

je konstanta zabezpečující, že součet získaných počtů jednotlivých druhů dá dohromady celkový počet jedinců. Nejdůležitějším úsekem výpočtu je získání hodnoty k odpovídající poměrné části zbývající niky zabrané jednotlivými druhy. k je získáno iteračním řešením rovnice

$$\frac{n_{min}}{N} = \frac{\frac{k}{1-k}(1-k)^S}{1 - (1-k)^S}, \quad (5.5)$$

kde n_{min} je počet jedinců nejméně početného taxonu a S počet taxonů ve společenstvu.

Hodnoty abundance, vypočítané na základě k pro jednotlivé druhy, mohou být porovnány s abundancí druhů reálného společenstva například pomocí χ^2 testu dobré shody nebo jinými metodami.

5.2.2 Logaritmická řada (série)

Logaritmické řady se stejně jako geometrické řady hodí pro popis situací, kdy ekologii společenstva dominuje jeden nebo málo faktorů. Avšak pokles abundance mezi jednotlivými druhy (tedy dominance) je pomalejší než v případě geometrické řady. Od jejich rozložení je odvozen index diverzity α , což je vlastně koeficient příslušné funkce. Výpočet probíhá podle vztahu:

$$S(n) = \alpha \frac{x^n}{n}, \quad (5.6)$$

kde $S(n)$ je počet druhů v kategorii s n jedinci a x je iteračně odhadováno ze vztahu

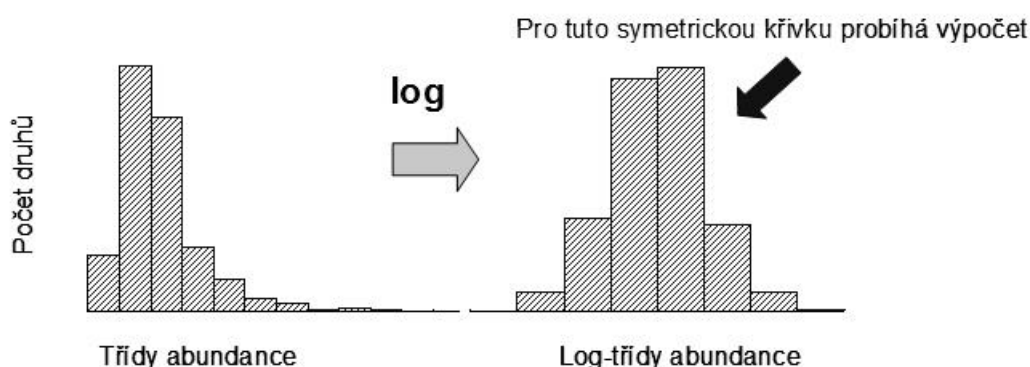
$$\frac{S}{N} = \frac{(1-x)[- \ln(1-x)]}{x}, \quad (5.7)$$

kde S je celkový počet druhů, N je celkový počet jedinců a α index diverzity (může sloužit též jako index diverzity) a jeho variabilita je:

$$\alpha = \frac{N(1-x)}{x} \text{ a } \text{Var}(\alpha) = \frac{\alpha}{- \ln(1-x)}. \quad (5.8)$$

5.2.3 Log-normální rozložení

Mnoho studovaných společenstev organismů je log-normálně rozloženo [60], tento jev je v biologii poměrně obecný. Příkladem může být rozložení početnosti bílých krvinek v populaci. Předpokládá se, že toto rozložení se vyskytuje u velkých, pestrých a rozvinutých společenstev. Jedním z vysvětlení je, že jde o důsledek statistických vlastností velkých čísel a centrálního limitního teorému [36]. Centrální limitní teorém říká, že při působení velkého množství faktorů na proměnnou způsobí náhodné variace těchto faktorů normální rozložení této proměnné; tento vliv se zesiluje při vzrůstajícím počtu faktorů. V našem případě jsou proměnnou abundancí taxonů (logaritmicke standardizované) a faktory jsou všechny procesy působící v prostředí na společenstvo. Rozložení bylo poprvé použito Prestonem [52], který použil logaritmus o základu dvě pro standardizaci abundancí (vznikly tzv. oktávy – třídy abundancí představující zdvojnásobení počtu jedinců (obr. 5.3)).



Obr. 5.3 Vztah log-normálního a normálního rozložení abundancí

Rozložení je nejčastěji popisováno ve formě

$$S(R) = S_0 \exp(-a^2 R^2), \quad (5.9)$$

kde $a = (2\sigma^2)^{0.5}$ je inverzní šířka rozložení, $S(R)$ je počet druhů v R -té třídě v levé i pravé části symetrické křivky, S_0 se označuje počet druhů v nejpočetnější třídě. Empirické studie ukázaly, že a je obvykle blízké 0,2 [66].

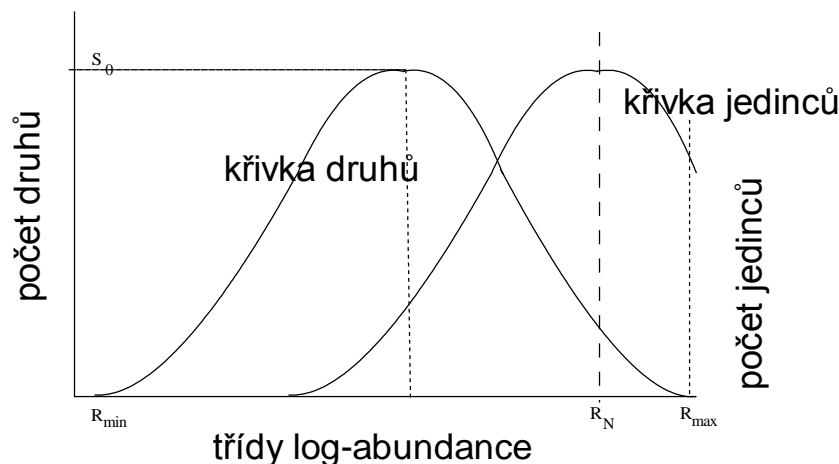
S log normálním rozložením je spjat také index diverzity γ , který je mírou vztahu mezi vrcholem křivky jedinců (modální oktáva křivky jedinců) a horní hranicí křivky druhů (oktáva s nejpočetnějšími druhy). Výpočet probíhá podle vztahu

$$\gamma = \frac{R_N}{R_{max}} = \frac{\ln 2}{2a(\ln S_0)^{0.5}}, \quad (5.10)$$

kde R_N je označena nejpočetnější třída křivky jedinců a R_{max} třída křivky druhů obsahující nejpočetnější druhy (obr. 5.4). Jde vlastně o odhad počtu druhů v oktávě, kde křivka jedinců dosáhne vrcholu.

V mnoha případech leží vrchol křivky jedinců (R_N) ve stejném bodě jako horní hranice křivky druhů (R_{max}), potom $\gamma \sim 1$ a hovoříme o takzvaném kanonickém log-normálním rozložení [53] (kanonická hypotéza), kdy je směrodatná odchylka omezena jen na úzké rozpětí hodnot ($a \sim 0.2$). Protože podobná situace nastává i u neekologických dat, je možné, že log-normální rozložení nemá žádné biologické pozadí, ale jde pouze o matematický artefakt [36]. Na druhou stranu [60] existuje řada příkladů společenstev, které lze vysvětlit kanonickou hypotézou. Sugihara [60] také podává biologické vysvětlení kanonického log-normálního rozložení prostřednictvím modelu dělení niky, kde dělení niky přirovnává k rozbíjení kamene; velikost jeho úlomků bude také log-normálně rozložena, tato hypotéza tedy propojuje čistě stochastický model rozložení abundancí s na niku orientovanými modely. I když se uvádí [36], že korelace s empirickými daty

nemusí být zárukou správnosti, jde o vhodnou pracovní hypotézu. Diskuse nad biologickým nebo matematickým původem log-normálního rozložení abundancí ovšem neustávají. Podle Uglanda [67] budou druhově bohatá společenstva (s více než 50 druhy) nejspíše vždy kanonická, neboť $\gamma \sim 1$ je jejich matematickou vlastností. Také v případě $a \sim 0,2$ se zdá, že jde o matematický artefakt velkého počtu druhů [36, 67].



Obr. 5.4 Vztah křivky druhů a křivky jedinců u log-normálního rozložení abundancí (index diverzity γ)

S výpočtem log-normálního rozložení abundancí se váže ještě další problém, a to nezachycení velmi vzácných druhů (levá strana křivky). Výsledkem je takzvaná zkrácená log-normal křivka, její výpočet popsal Pielou [31].

Alternativní metodou pro proložení log-normálního rozložení daty abundancí je metoda popsaná [68, 69], takzvaná Poissonova log-normální metoda, která předpokládá, že spojitá log-normální křivka je reprezentována řadou nespojitých tříd abundancí, které se chovají jako složené Poissonovské proměnné.

Společenstva popsaná log-normálním rozložením lze často úspěšně proložit také rozložením logaritmické řady nebo i jinými modely. Také by mohlo k jejich vzniku dojít díky špatné identifikaci druhů nebo chybou vzorkování. Podle některých autorů by log-normální rozložení nemělo být používáno pro data abundancí [62, 70]. I přesto je zde mnoho příkladů vzorkování, které lze úspěšně popsat log-normálním rozložením [60, 71]. Zdá se, že log-normální rozložení zůstane i nadále významným nástrojem při studiu diverzity.

5.2.4 Zlomená hůlka (*broken stick*)

Tento model je shodný s modelem na niku orientovaným. Jedná se o jeden z modelů, který je popsán jak matematicky, tak modelem dělení niky.

MacArthur [59] definoval tento model jako případ, kdy celková nika je náhodně a současně rozdělena do úseků zabraných jednotlivými druhy. Model může být popsán jako skupina druhů s podobnými kompetičními schopnostmi, které současně obsazují celkovou niku a navzájem se vytlačují a tak určují své hranice niky. Model je použitelný pro malá společenstva blízce příbuzných druhů, která dosáhla vyrovnaného stavu.

Model zlomené hůlky odráží mnohem více vyrovnaný stav než ostatní rozložení. Je to vlastně vyjádření rovnoměrného rozložení druhů. Není od něj odvozen žádný index diverzity. A protože je charakterizován jen jediným parametrem, počtem druhů, je silně ovlivněn velikostí vzorku. Výpočet probíhá pomocí vztahu

$$S(n) = \frac{S(S-1)}{N} \left(1 - \frac{n}{N}\right)^{S-2}, \quad (5.11)$$

kde $S(n)$ je počet druhů v kategorii s n jedinci, S je celkový počet druhů a N celkový počet jedinců.

5.2.5 Zipfovy-Mandelbrotovy modely

Nespokojenost s existujícími modely heterogenity společenstva přiměla řadu ekologů k hledání dalším možných modelů pro lepší popis těchto vztahů.

Jedním z takto získaných modelů je i Zipfův-Mandelbrotův model [57], který má podobně jako Shannonův-Weaverův index základ v lingvistické a informační teorii. Zipfův zákon popisuje pořadí slov podle frekvence jejich výskytu v textu. G. K. Zipf studoval otázky statistického rozložení slov v textu. Zjistil, že když vynese logaritmy frekvence nejčastěji se vyskytujících slov proti logaritmu jejich pořadí, tvoří body na počátku grafu zcela přijatelnou přímku. Konec představovaný řídkce se vyskytujících slovy potom začne od přímky odpadat. Zipf použil pořadovou statistiku:

$$N(x) = \frac{a}{x^b}, \quad (5.12)$$

kde a a b jsou konstanty a x je nezávislá proměnná. Kvůli lepšímu dosazení dat navrhl Mandelbrot jeho zobecnění

$$N(x) = \frac{a}{(c+x)^b}, \quad (5.13)$$

kde a , b a c jsou konstanty. Zipfovo-Mandelbrotovo rozložení, které vzniklo v kontextu statistické mechaniky, může být obecně zapsáno takto:

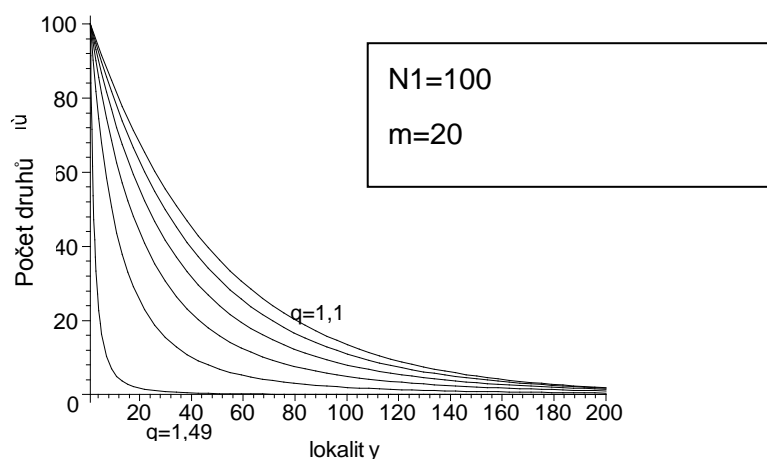
$$N(x) = N_0 [1 - (1-q)\lambda x]^{1/(1-q)}, \quad (5.14)$$

kde N_0 , λ a q jsou reálné parametry. Tato funkce při $\lim q \rightarrow 1$ přechází v normální exponenciální rozložení

$$N(x) = N_0 \exp(-\lambda x). \quad (5.15)$$

Navíc tato funkce vyhovuje rovnici normálního rozložení:

$$\frac{d}{dx} \left(\frac{N(x)}{N_0} \right) = -\lambda \left(\frac{N(x)}{N_0} \right)^q. \quad (5.16)$$



Obr. 5.5 Simulace vlivu q na tvar Zipfova-Mandelbrotova rozložení

Příkladem využití může být modelování počtu vzácných druhů na lokalitách. Nejprve, aby mohly být odstraněny N_0 a λ , je třeba vyjádřit počet lokalit se vzácnými taxony, s výskytem pou-

ze jednoho jedince a průměrný počet taxonů na lokalitu. Mimoto je také dobré nahradit diskrétní průměr spojitým:

$$m = \frac{\int_0^{\infty} xN(x)}{\int_0^{\infty} N(x)} = \frac{1}{\lambda(3-2q)}. \quad (5.17)$$

Parametr q může být považován za míru normality rozložení, s rostoucím q se zvětšuje zakřivení paraboly (obr. 5.5). Po předchozích úpravách může být rovnice 5.14 pro rozložení taxonů přepsána jako:

$$N(x) = N(1) \frac{\left[1 - \frac{(1-q)}{(3-2q)m}\right]^{\frac{1}{1-q}}}{\left[1 - \frac{(1-q)}{(3-2q)m}x\right]^{\frac{1}{1-q}}}. \quad (5.18)$$

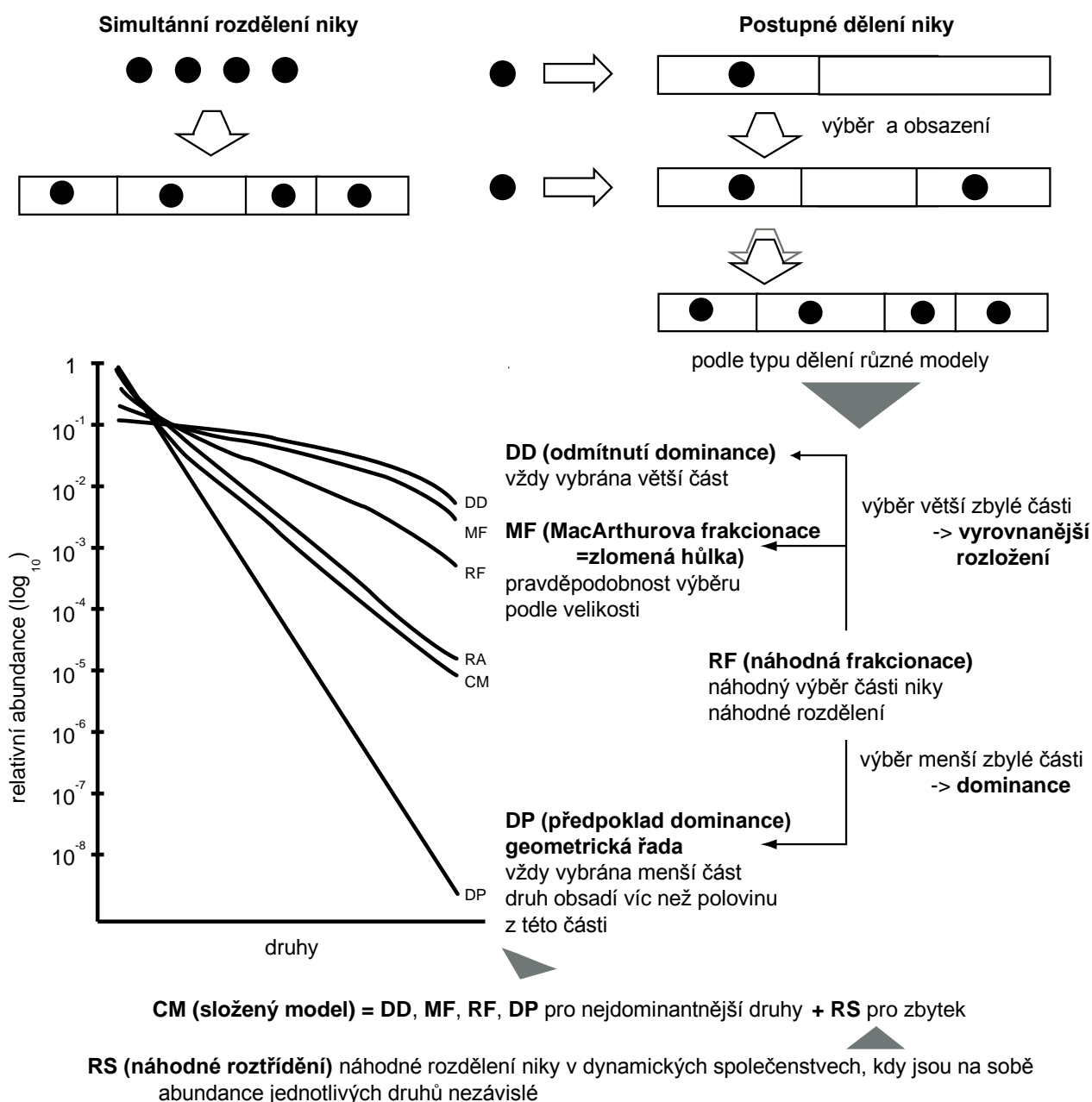
V ekologickém podání je obecně známa interpretace Zipfových-Mandelbrotových modelů odrážející sukcesní proces, ve kterém pozdní kolonizátoři mají více specifické požadavky a jsou proto vzácnější než druhy, jež obsadily prostor dříve [72]. Model předpokládá pevně dané pořadí kolonizátorů se stejnými druhy přítomnými ve stejném bodě sukcese v podobných habitatech, ovšem v reálných podmínkách tento předpoklad není vždy dodržen [73]. Přestože model není přesný ve společenstvech s velkým množstvím vzácných druhů, byl model úspěšně aplikován na řadu společenstev [72-74].

5.3 Simulační, na niku orientované modely

Tyto modely představují jednu ze skupin simulačních modelů. Představují spojnicí mezi koncepcí niky a výslednými abundancemi druhů. Proto je jejich vývoj spjat s poznáním biologických procesů podílejících se na rozložení abundancí druhů; jsou odvozeny ze způsobu obsazování nik organismy.

Základní podkladem při jejich definici je představa dostupných nik na lokalitě jako objektu, který si jednotlivé druhy mezi sebou lámou, přičemž velikost částí připadajících na druhy se v realitě odráží jako jejich biomasa nebo abundance. V zásadě jsou možné dva způsoby dělení a to simultánní, kdy si druhy dostupnou niku rozdělí současně a postupný, kdy v každém kroku vstupuje do systému nový druh a zabírá část z dostupné niky. Protože je problematické simultánní postup převést z myšlenkového konceptu do podoby návodu k výpočtu, je většina zde představených modelů ve formě modelů postupného dělení (obr. 5.6).

Protože tento typ modelů představuje možné biologické scénáře a ne stochastické modely, je zde možné, že některé z modelů mohou vést ke shodnému rozložení abundancí odlišným postupem. Obrázek 5.6 ukazuje graf základního průběhu různých modelů (z nichž některé poskytují obdobné výsledky) a jejich seřazení podle výsledného společenstva (spíše vyrovnaného nebo spíše s dominantními druhy). S výjimkou geometrické řady a modelu zlomené hůlky (MacArthurův frakční model) nemají tyto modely stochastické vyjádření. Detailní přehled modelů a jejich vlastností je obsažen v následujících kapitolách.



Obr. 5.6 Základní koncepce a vlastnosti na niku orientovaných modelů [22, 75]

5.3.1 Geometrická řada [58]

Tento model je shodný s matematickým modelem geometrické řady, jde o jeden z modelů, který je popsán jak matematicky, tak modelem dělení niky.

Geometrické řady se využívá u druhově chudých společenstev nebo u společenstev v raném stadiu sukcese. Tento model je založen na předpokladu, že nejdominantnější druh obsadí určitou poměrnou část zdrojů, druhý stejnou poměrnou část zbytku atd.

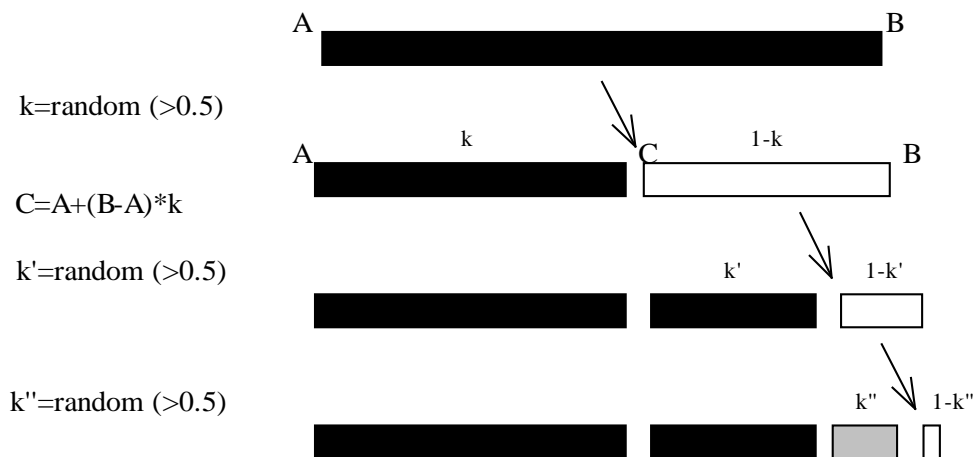
První druh obsadí část k ($k \in (0,5;1,0)$) celkové niky, druhý druh část k' zbytku, třetí druh část k'' toho, co zůstane po umístění prvního a druhého druhu, atd. Hodnota k je shodná pro všechny druhy a její odchylky jsou dány pouze chybou vzorkování. Každý druh tak obsazuje větší část niky než všechny následující druhy dohromady (obr. 5.7).



Obr. 5.7 Schéma frakcionace niky u modelu geometrické řady

5.3.2 Předpoklad dominance (*dominance preemption*) [21]

V tomto modelu si druhy postupně okupují dominantní část ($>0,5$) ze zbývající části niky. Tento přístup zajišťuje dominanci druhu nad všemi ostatními následujícími druhy. V Tokeshiho [21] originální verzi se předpokládá, že tato vyčerpaná část niky sahá náhodně od 0,5 do 1 a hodnot abundancí vypočítaných z geometrické řady dosahuje při $k = 0,75$ (obr. 5.8).

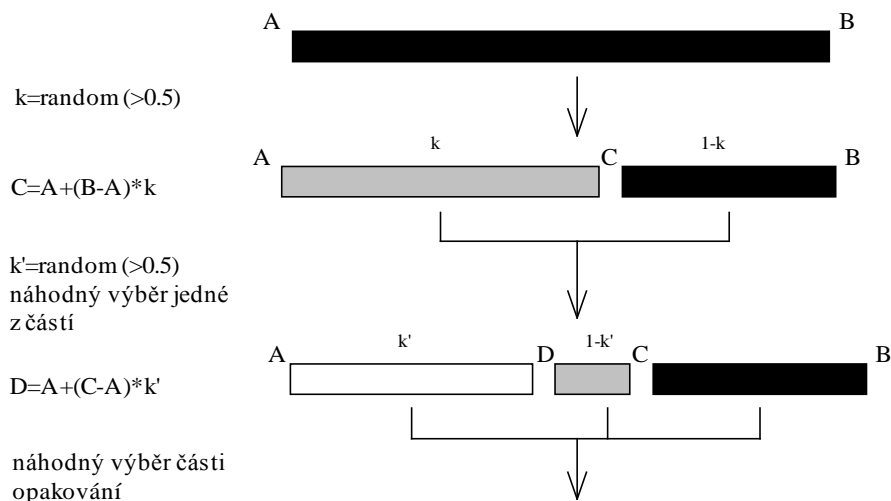


Obr. 5.8 Schéma frakcionace niky u modelu s předpokladem dominance (*Dominance preemption*)

5.3.3 Náhodná frakcionace (*random fraction*) [21]

Jde opět o model, který postupně rozkládá celkovou niku na části. Náhodnost procesu probíhá na dvou úrovních, nejprve je celková nika rozdělena na dvě části, z nich je dále náhodně jedna vybrána a náhodně rozdělena. V dalším kroku je některá z výsledných částí opět náhodně (uniformně) vybrána a rozdělena atd. Interpretace spočívá v příchodu nového druhu, který vstoupí do niky již přítomného druhu a vezme si z ní určitou část [21].

Nejprve je celková nika náhodně rozdělena na dvě části (rovnoměrné rozložení dělicích bodů k z intervalu $(0,5;1)$). Jedna z částí je náhodně vybrána pro další dělení a opět náhodně rozdělena. Vzniklé tři části jsou opět podrobeny náhodnému výběru a z nich vybraná část náhodně rozdělena. Každá z částí má stejnou pravděpodobnost být vybrána. Tento postup představuje situaci, kdy si nově přichodící druh náhodně vybere niku jiného druhu a obsadí její náhodně velkou část. Speciálními případy jsou předpoklad dominance, pro dělení se vždy vybere ta nejmenší část; a odmítnutí dominance, pro dělení se vybírá největší část (obr. 5.9).



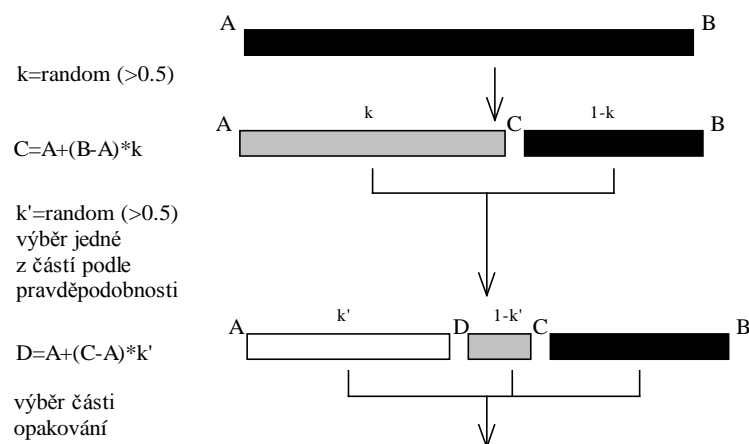
Obr. 5.9 Schéma frakcionace niky u modelu s náhodnou frakcionací (Random fraction)

5.3.4 Sugiharův model postupného dělení (Sugihara sequential breakage model) [60]

Základem modelu je frakcionace celkové niky analogicky k náhodnému frakcionačnímu modelu. Rozdíl je v distribuci náhodnosti. V případě tohoto modelu jde o rozložení $p = 0,5$ až 1 , s vrcholem v $0,75$, zatímco v případě náhodné frakcionace jde o uniformní náhodné rozložení v rozmezí $p = 0,5$ až 1 .

5.3.5 Zlomená hůlka (broken stick) [59] alias MacArthurova frakcionace [21]

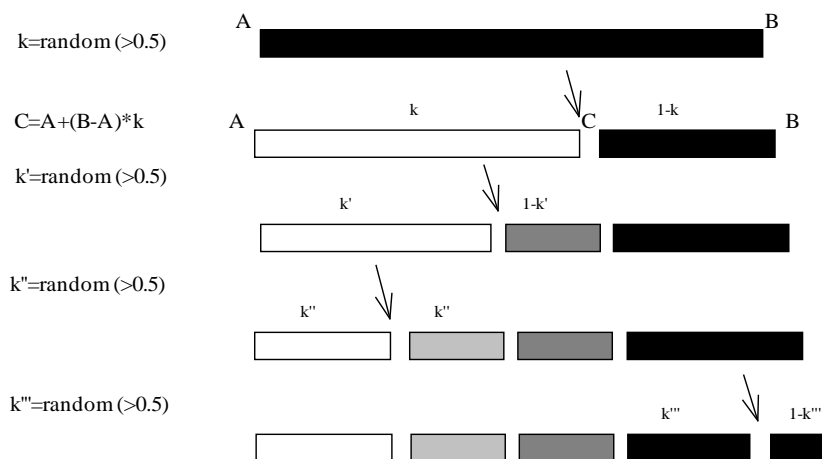
Kvůli nemožnosti srovnání s ostatními modely postupného dělení niky byl původní model zlomené hůlky popsán jako model postupného dělení (obr. 5.10). V modelu je celková nika náhodně rozdělena. Vzniklé niky jsou pro další dělení vybrány s pravděpodobností, která je ve vztahu s jejich velikostí; celková nika se tedy dělí obdobně jako v případě modelu náhodné frakcionace. Jediný rozdíl spočívá při výběru části pro další dělení. U náhodné frakcionace byl tento výběr čistě náhodný. MacArthurova frakcionace předpokládá pravděpodobnostní výběr (větší část má větší pravděpodobnost být vybrána pro dělení – je zde lineární vztah velikosti části a pravděpodobnosti jejího výběru, tedy nově přichází druh spíše obsazuje niku více početných druhů).



Obr. 5.10 Schéma frakcionace niky u modelu zlomená hůlka (broken stick) alias MacArthurova frakcionace

5.3.6 Odmítnutí dominance (*dominance decay*) [21]

Tento model je inverzí modelu předpokladu dominance (*dominance preemption*), kdy je největší nika objektem následujícího dělení. Invazivní druh tak obsazuje niku nejpočetnějšího druhu existujícího společenstva. Tento proces vyúsťuje ve společenstvo více vyrovnané než v případě modelu zlomené hůlky (*broken stick model*). Dominance je zde popírána a výsledkem tohoto dělení jsou víceméně vyrovnané abundance jednotlivých druhů (obr. 5.11).



Obr. 5.11 Schéma frakcionace niky u modelu odmítnutí dominance (*dominance decay*)

K podobným výsledkům vede i model částicové niky (*particulate niche*) [59]. Tento model předpokládá existenci pevně daného počtu jednotek abundance (částice niky) založené na náhodném přiřazení těchto jednotek mezi druhy, které mají shodnou pravděpodobnost jejich získání. Tento model představuje společenstvo vyrovnaných druhů ještě více než v případě modelu zlomené hůlky a tedy obdobně jako model odmítnutí dominance. Předpokládá se shodný růst populace, a pokud je nika obsazena nepřítomnost kompetice. Teoreticky vede k Poissonovu rozložení abundance druhů.

5.3.7 Náhodné roztržení (*random assortment*) [21]

V tomto případě se abundance různých druhů liší navzájem zcela nezávisle. Jde o systém, kde není dostatek času pro kompetici na základě zdrojů a nika není zcela vyplněna.

5.3.8 Složený model (*composite model*) [21]

Složený model představuje kombinaci více různých modelů. Při jeho tvorbě předpokládáme, že společenstvo je vytvářeno na základě několika různých kritérií. Pro několik nejpočetnějších druhů (dva až tři druhy) použijeme některý ze základních modelů dělení niky. Pro zbylé vzácné druhy je využit model náhodného roztržení. Zde vycházíme z předpokladu, že u vzácných druhů můžeme jen obtížně stanovit jejich přesné počty z důvodu malého vzorku, a tedy jejich abundance můžeme pokládat za náhodné.

5.4 Hodnocení kompetice, agregace a šíře niky

Zajímavým měřítkem vlastností druhu souvisejícím s niku orientovanými modely je „agregovanost“ jedinců stejného druhu. Zde dochází ke kombinaci vztahů jedinců různých druhů na obsazené lokalitě a velikostí prostoru zaujímaného druhem vzhledem k celkové nize.

5.4.1 Vnitrodruhová agregace

Vhodné měřítko pro tento druh problému vytvořil Ives [76, 77]. Jde vlastně o měřítko poměrného nárůstu kompetitorů stejného druhu vzhledem k náhodnému rozložení. Výpočet probíhá podle vztahu

$$J_1 = \frac{\sum_{i=1}^p \frac{n_{1j}(n_{1i} - 1)}{m_1}}{m_1}, \quad (5.19)$$

kde n_i a m_i jsou počet a rozptyl počtu jedinců druhu 1 v části niky i (celková nika je rozdělena na p částí). Pokud je hodnota $J = 0$, jde o náhodné rozložení, hodnota $J = 0,5$ znamená 50% nárůst jedinců stejného druhu v segmentech niky oproti jejich náhodnému rozložení.

5.4.2 Mezidruhová agregace

Pro měření mezidruhové kompetice vyvinul [76, 77] obdobné měřítko vyjadřující poměrný nárůst kompetitorů jiného druhu oproti náhodnému rozložení.

$$C_{1,2} = \frac{\sum_{i=1}^p \frac{n_{1i}n_{2i}}{m_1 p} - m_2}{m_2} = \frac{Cov_{12}}{m_1 m_2}, \quad (5.20)$$

kde Cov je kovariance mezi dvěma druhy. Pokud je $C > 0$, jsou druhy spolu pozitivně asociovány; pokud je $C < 0$, jde o negativní asociaci.

Snížení kompetice díky vnitrodruhové agregaci (relativní síla vnitrodruhové agregace oproti mezidruhové agregaci) pak může být spočítána jako [77, 78]

$$A_{12} = \frac{(J_1 + 1)(J_2 + 1)}{(C_{12} + 1)^2}. \quad (5.21)$$

Pokud je A_{12} větší než 1, je vnitrodruhová agregace silnější než mezidruhová.

5.4.3 Šíře a překryv niky

Pro výpočet šíře niky druhu lze využít Levinovu šíři niky [79]:

$$B = \frac{1}{\sum (p_j^2)}, \quad (5.22)$$

kde p_j je podíl jedinců daného druhu nalezených v segmentu j niky.

Pro zjištění překryvu nik dvou druhů lze využít takzvané Renkonenovo číslo [80]:

$$R = 1 - \frac{\sum |p_{ia} - p_{ja}|}{2}, \quad (5.23)$$

kde p_{ia} je podíl jedinců druhu i v sektoru a , stejně tak pro druh j .

6 Možnosti frakcionace biologických společenstev a následná analýza biodiverzity získaných složek

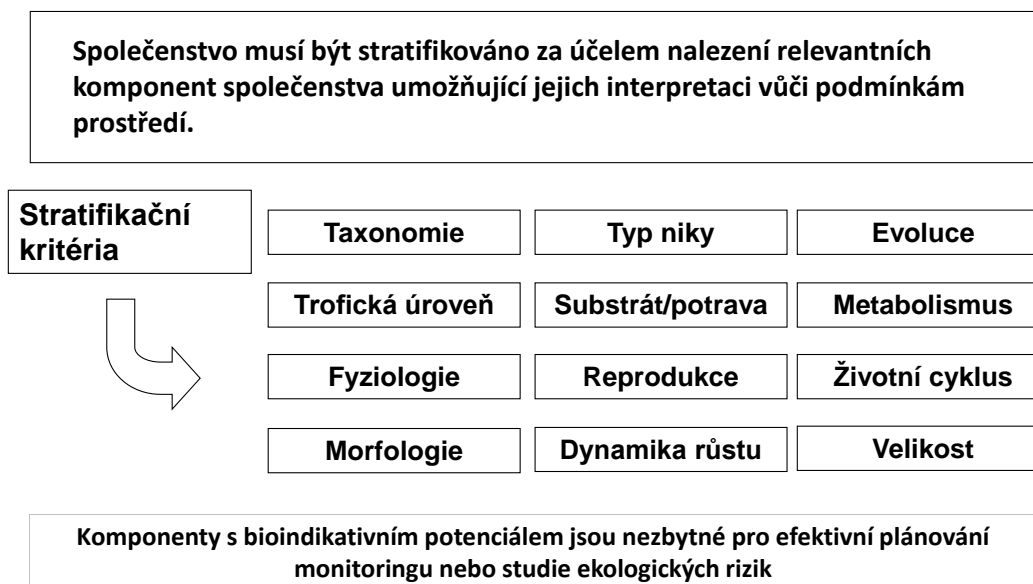
6.1 Rozdílný vliv prostředí na různé složky společenstva

Biologická společenstva jsou důsledkem působení podmínek prostředí na živé organismy. Zároveň ale platí, že různé organismy mohou na podmínky prostředí reagovat různým způsobem. Jejich protichůdné reakce mohou ztížit nebo i znemožnit interpretaci hodnocení biodiverzity vzhledem k podmínkám prostředí. Zvláště významné to je pro využití biodiverzity v bioindikaci a hodnocení ekologických rizik.

Možným řešením je frakcionace společenstva dle různých charakteristik taxonů za účelem nalezení skupin taxonů s homogenní reakcí na prostředí. Do určité míry je tento postup analogický k výpočtům biotických indexů, kdy vlastnosti taxonů jsou do analýzy zapracovány nikoliv prostřednictvím vah, ale stratifikací společenstva podle těchto vlastností.

6.2 Frakcionace společenstva – taxonomická, funkční aj.

K frakcionaci společenstva je možné přistoupit řadou různých způsobů (obr. 6.1), přičemž expertní znalost vlastností daných taxonů je nezbytná. Zcela jinou stratifikaci budou mít například parazitární společenstva, jinou společenstva lišejníků a opět jinou společenstva makrozoobentosu.

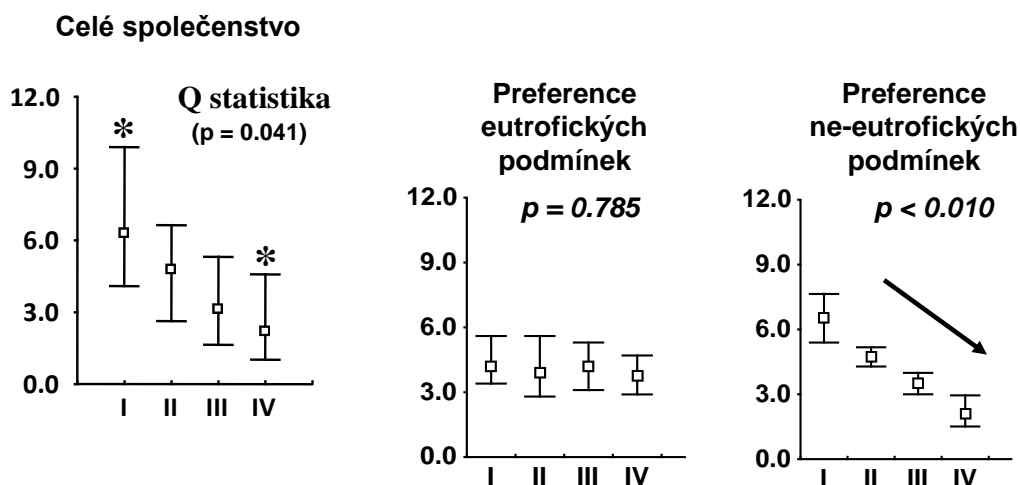


Obr. 6.1 Možnosti frakcionace společenstev

6.3 Identifikace složek společenstva s konzistentní reakcí na faktory prostředí

K identifikaci vztahu dílčích komponent společenstva s podmínkami prostředí je možné využít standardního statistického testování, jak je vidět na příkladu na obrázku 6.2.

Společenstva lišejníků ve čtyřech oblastech s různou mírou imisní zátěže byla stratifikována dle morfologie, substrátu a preference niky. Hodnocení společenstva prostřednictvím Q indexu diverzity identifikovalo preferenci niky (eutrofické nebo neeutrofické podmínky) jako hlavní faktor odlišující lokality s různou mírou imisní zátěže. Druhy lišejníků s preferencí neeutrofických podmínek citlivě odlišily všechny čtyři skupiny různě zatížených lokalit.



Vzrůstající imisní zátěž I < II < III < IV

Obr. 6.2 Příklad identifikace komponenty společenstva s bioindikačním potenciálem

7 Aplikovatelnost parametrických a neparametrických statistických technik při hodnocení biodiversity

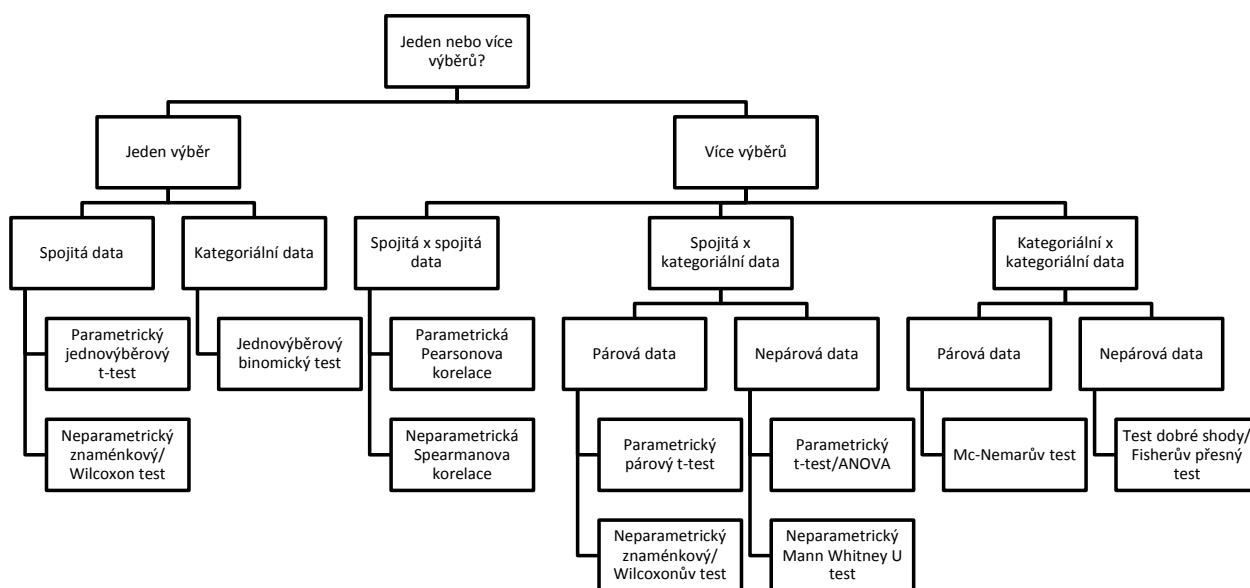
Vlastní prezenze a abundance taxonů, stejně jako výsledky výpočtů biodiverzitních indexů a zařazení společenstev dle modelů profilů abundance mohou vstupovat do libovolných statistických výpočtů. Vstupní a výstupní data z analýzy biodiversity jsou zde v pozici popisných charakteristik společenstev a lokalit, obdobně jako například nadmořská výška nebo koncentrace chemických látek ve vodě či půdě.

S biodiverzitními charakteristikami lze pro skupiny společenstev (lokalit) počítat (obr. 7.1):

- souhrnné popisné statistiky,
- srovnání dvou nebo více skupin společenstev,
- korelace diverzitních charakteristik navzájem nebo s podmínkami prostředí,
- modelování diverzity v závislosti na podmínkách prostředí,
- vícerozměrná analýza prezencí/abundancí taxonů nebo biodiverzitních charakteristik.

Klasické statistické testy nejsou součástí těchto předkládaných skript, na rozdíl od vybraných vícerozměrných analýz, o kterých pojednává samostatná kapitola.

Stejně jako v jakékoliv jiné statistické analýze i zde platí, že „Statistika využívá matematické modely reality k zobecnění výsledků experimentů a vzorkování. Statistika funguje korektně, pouze pokud jsou splněny předpoklady jejích metod a modelů“ a s biodiverzitními charakteristikami je třeba pracovat jako s kterýmikoliv jinými daty.



Obr. 7.1 Rozhodovací proces výběru statistických testů

8 Metody hodnocení β biodiverzity

8.1 β biodiverzita a vícerozměrná analýza biodiverzity

β biodiverzita sleduje rozdíly druhového složení mezi více společenstvy a měří tak vlastně podobnost společenstev. Vzhledem k tomu, že společenstvo je popsáno vícerozměrně přítomností nebo abundancí jeho taxonů, jde o vícerozměrnou podobnost těchto společenstev. A na rozdíl od indexů a modelů druhové abundance α biodiverzity zohledňuje současně všechny taxony tvořící společenstvo při zachování jejich individuality, která není „zprůměrována“ do formy indexového hodnocení jedním číslem nebo přiřazením k modelu rozdělení abundancí.

Indexy β biodiverzity jsou na pozadí řady vícerozměrných statistických metod, kdy v obvyklém případě vystupují jednotlivé taxony v roli proměnných, které definují podobnost společenstev ať již v čase nebo v prostoru.

Výsledkem vícerozměrných analýz mohou být seskupení druhově podobných si lokalit (společenstev), které následně mají ekologickou interpretaci i z hlediska abiotických podmínek, které mohou determinovat výskyt organismů.

Při výběru vhodné metodiky analýzy β biodiverzity je třeba zohlednit i základní ekologické charakteristiky vztahu mezi organismy a prostředím a to zejména zákon ekologického optima a minima, které jednoznačně determinují výběr vícerozměrné statistické metodiky.

8.2 Problém dvojité nepřítomnosti (*double-zero problem*)

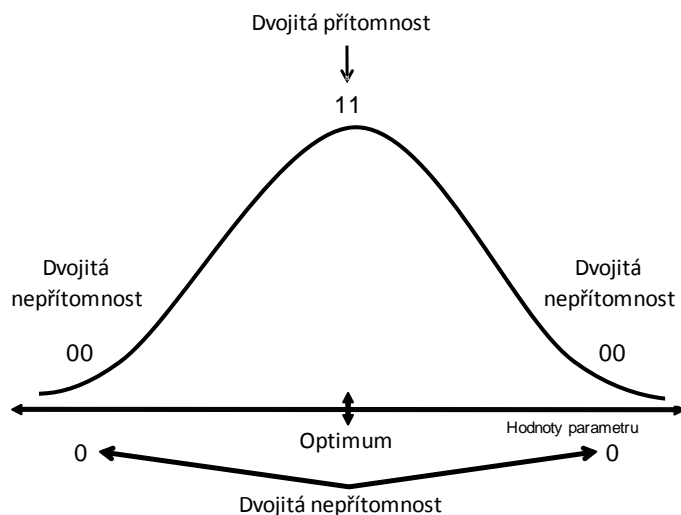
Dvojitá nepřítomnost je častým problémem β biodiverzity a vícerozměrné analýzy v ekologii. Podobnost společenstev hodnotíme podle druhů tvořících jejich strukturu. Druhy se vyskytují na místech, kde mají vhodné podmínky pro život. Závislost výskytu organismů na prostředí je většinou unimodální (má pouze jedno maximum). Rozložení organismů je závislé na optimálních hodnotách gradientů životních podmínek, avšak dokáže také tolerovat určité rozpětí těchto podmínek. Na začátku a na konci valenční křivky se objevují body zlomu, kde počet jedinců druhů začíná růst nebo klesat (obr. 8.1). Na lokalitách, které nemají tyto optimální hodnoty může být daný druh vzácný nebo se zde nemusí vůbec vyskytovat. Když se druh vyskytuje na dvou lokalitách, poukazuje na určitou podobnost mezi těmito dvěma lokalitami (mají podobné podmínky – optimum \pm rozpětí). Avšak když se druh nevyskytuje ani na jedné lokalitě, mohou mít lokality úplně rozdílné podmínky (jedna nebo druhá strana optima) (obr. 8.1). Dvojitá přítomnost je za této situace lepším ukazatelem podobnosti lokalit než dvojitá nepřítomnost [81].

S tímto problémem také souvisí problém vzácných druhů. Společenstva se většinou skládají z malého počtu dominantních druhů a velkého počtu druhů vzácných. Pak datová matice obsahuje velký počet nul. To má za následek, že míra podobnosti společenstev je založena na mnoha párech nul, které nemají zcela jasný biologický význam [81]. Problém dvojité nepřítomnosti je jedním z nejpodstatnějších faktorů, které ovlivňují vícerozměrné analýzy biologických společenstev a musí s ním být v analýze vždy počítáno.

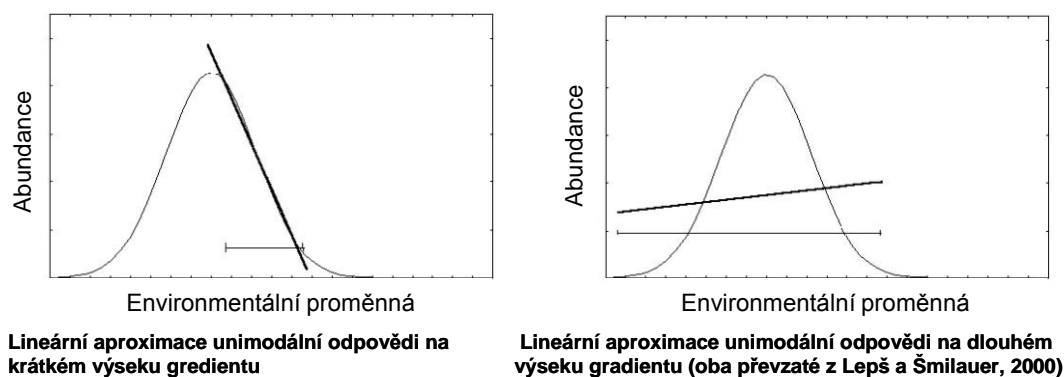
V případech, kdy data obsahují celý gradient s unimodální odpovědí organismu, je třeba tuto skutečnost (výskyt dvojí nepřítomnosti) zohlednit při výpočtu. Naopak, v případě existence lineární odpovědi druhu na prostředí (data obsahují pouze část gradientu, v němž se taxon vzhledem k podmínkám prostředí chová lineárně), je problém dvojí nepřítomnosti možné zanedbat (obr. 8.2).

Lineární model – model odpovědi druhu na gradienty prostředí, lineární odpověď je nejjednodušším odhadem na krátkém gradientu (tj. výrazně užším, než je toleranční rozpětí organismů). Můžeme zde použít lineární aproximace kterékoliv funkce, problém dvojité nepřítomnosti se neprojevuje.

Unimodální model – odpověď předpokládá, že druh má na gradientu prostředí své optimum, používá se na dlouhém gradientu (výrazně širší než je toleranční rozpětí organismů, např. hodnocené lokality jsou hodně odlišné). Zde je aproximace lineární funkcí nevhodná a je třeba při výpočtu počítat s problémem dvojité nepřítomnosti.



Obr. 8.1 Problém dvojité nepřítomnosti. Dvojitá nepřítomnost není stejná jako dvojitá přítomnost.



Obr. 8.2 Modely odpovědi taxonu na gradienty prostředí (dle [82])

8.3 Koeficienty podobnosti pro data o biodiverzitě (β biodiverzita)

Indexy β biodiverzity můžeme zároveň brát jako koeficienty podobností mezi společenstvy, které sledují rozdíly v druhovém složení více společenstev.

K analýze podobností společenstev můžeme přistoupit jak prostřednictvím výskytu taxonů, tak i s využitím jejich abundancí, jde o základní dělení podobnostních koeficientů.

8.3.1 Binární koeficienty podobnosti

Binární metriky počítají pouze s přítomností nebo absencí taxonů. Dále je můžeme rozdělit na symetrické a asymetrické. Symetrické metriky nerozlišují mezi případem 1–1 a 0–0, obě události mají stejnou váhu. Asymetrické koeficienty rozlišují mezi 1–1 a 0–0, přičemž existuje mnoho postupů, jak spočítat míru asociace mezi druhy založené na binárních datech. Tyto koeficienty jsou založené na 2×2 frekvenční tabulce (obr. 8.3).

		Společenstvo 1		
		1	0	
Společenstvo 2	1	a	b	a + b
	0	c	d	c + d
		a + c	b + d	

a, b, c, d – počet případů, kdy souhlasí binární charakteristika společenstev

$a + b + c + d = n$

Obr. 8.3 Výpočet binárních metrik podobnosti na základě 2×2 frekvenční tabulky

V případě lineární odpovědi organismů na prostředí je možné využít symetrické binární koeficienty.

Jednoduchý srovnávací koeficient [83], kdy metodou pro výpočet podobnosti mezi dvěma objekty je podíl počtu deskriptorů, které kódují objekt stejně a celkového počtu deskriptorů.

$$S(x_1, x_2) = \frac{a + d}{p}, \quad (8.1)$$

kde x_1 a x_2 jsou srovnávaná společenstva, a je počet současných výskytů, d počet současných nevýskytů a p počet všech taxonů ve srovnávaných společenstvech. Při použití tohoto koeficientu předpokládáme, že není rozdíl mezi nastáním 0 a 1 u deskriptorů.

Rogersův a Tanimotův koeficient [84] dává větší váhu rozdílům než podobnostem:

$$S(x_1, x_2) = \frac{a + d}{a + 2b + 2c + d}, \quad (8.2)$$

kde x_1 a x_2 jsou srovnávaná společenstva, a je počet současných výskytů, d počet současných nevýskytů a b, c počet neshodujících se taxonů ve srovnávaných společenstvech. Obdobné indexy, které stejně jako 8.2 snižují vliv dvojité nuly, byly navrženy Sneathem a Sokalem [85]:

$$S(x_1, x_2) = \frac{2a + 2d}{2a + b + c + 2d}, \quad (8.3)$$

tento koeficient dává dvakrát větší váhu shodným deskriptorům než rozdílným.

$$S(x_1, x_2) = \frac{a + d}{b + c} \quad (8.4)$$

porovnává shody a rozdíly prostým podílem v měřítku jdoucím od 0 do nekonečna.

$$S(x_1, x_2) = \frac{1}{4} \left[\frac{a}{a+b} + \frac{a}{a+c} + \frac{d}{b+d} + \frac{d}{c+d} \right] \quad (8.5)$$

porovnává shodné deskriptory se součty okrajů tabulky.

$$S(x_1, x_2) = \frac{a}{\sqrt{(a+b)(a+c)}} \frac{d}{\sqrt{(b+d)(c+d)}} \quad (8.6)$$

je vytvořen z geometrických průměrů členů vztahujících se k a a d .

Pro eliminaci vlivu problému dvojité nuly jsou využity asymetrické binární koeficienty. Nejznámější z nich je Jaccardův koeficient [86-88]:

$$S(x_1, x_2) = \frac{a}{a+b+c}, \quad (8.7)$$

kde x_1 a x_2 jsou srovnávaná společenstva, a je počet současných výskytů a b , c je počet neshodujících se taxonů ve srovnávaných společenstvech.

Sørensenův koeficient [89] jako varianta předchozího koeficientu dává dvojnásobnou váhu dvojitým prezencím

$$S(x_1, x_2) = \frac{2a}{2a+b+c}, \quad (8.8)$$

protože se může zdát, že přítomnost druhů je více informativní než jejich absence, která může být způsobena různými faktory a nemusí nutně odrážet rozdílnost prostředí. Prezence druhu na obou lokalitách je silným ukazatelem jejich podobnosti. (8.7) je monotónní k (8.8), proto podobnost pro dvě dvojice objektů vypočítaná podle (8.7) bude podobná stejnému výpočtu (8.8). Oba koeficienty se liší pouze v měřítku. Jiná varianta tohoto koeficientu dává duplicitním prezencím trojnásobnou váhu:

$$S(x_1, x_2) = \frac{3a}{3a+b+c}. \quad (8.9)$$

Následující koeficient [85] byl navržen jako doplněk Rogersova a Tanimotova koeficientu a dává dvojnásobnou váhu rozdílům ve jmenovateli.

$$S(x_1, x_2) = \frac{a+d}{a+2b+2c} \quad (8.10)$$

Russelův a Raův koeficient [90] umožňuje porovnání počtu duplicitních prezencí (v čitateli) proti celkovému počtu druhů, nalezených na všech lokalitách, zahrnujícím druhy, které chybějí (d) na obou uvažovaných lokalitách.

$$S(x_1, x_2) = \frac{a}{p} \quad (8.11)$$

Kulczynski [91] nabízí koeficient porovnávající duplicitní prezence s diferencemi:

$$S(x_1, x_2) = \frac{a}{b+c}. \quad (8.12)$$

Mezi svými koeficienty pro presence/absence dat zmiňují Sneath a Sokal [85] tuto verzi koeficientu:

$$S(x_1, x_2) = \frac{1}{2} \left[\frac{a}{a+b} + \frac{a}{a+c} \right], \quad (8.13)$$

kde jsou duplicitní prezence srovnávány se součty okrajů tabulky $(a+b)$ a $(a+c)$.

Ochia [92] použil obdobně jako míru podobnosti geometrický průměr poměrů a k počtu druhů na každé lokalitě, tj. se součty okrajů tabulky $(a+b)$ a $(a+c)$:

$$S(x_1, x_2) = \sqrt{\frac{a}{(a+b)(a+c)} \frac{a}{(a+b)(a+c)}} = \frac{a}{\sqrt{(a+b)(a+c)}}. \quad (8.14)$$

Faith [93] popsal koeficient, kde je neshoda (přítomnost na jedné a absence na druhé lokalitě) vážena proti duplicitní prezenci. Hodnota klesá s růstem počtem dvojí nepřítomnosti.

$$S(x_1, x_2) = \frac{a + d / 2}{p} \quad (8.15)$$

V biogeografii je pro asociaci taxonů na bázi společného výskytu na lokalitách často používán Baroniův-Urbaniův a Buserův index podobnosti [94], který lze zapsat jako:

$$S_B(x_1, x_2) = \frac{\sqrt{c \times d} + c}{\sqrt{c \times d} + a + b - c}, \quad (8.16)$$

kde a je počet lokalit, ve kterých je přítomný druh A, b je počet lokalit, ve kterých je přítomný druh B, c je počet lokalit, ve kterých jsou přítomny oba druhy A i B a d je počet lokalit, ve kterých není přítomen ani jeden z porovnávaných druhů A a B. Na rozdíl od Jaccardova indexu, který nebere v úvahu společné absence obou porovnávaných druhů (d), se u tohoto indexu uvádí [94], že společná absence druhů na lokalitách je důležitá, neboť může mít ekologickou či historickou příčinu a proto by neměla být opomíjena. Baroniův-Urbaniův a Buserův index dává větší významnost společné přítomnosti tím, že společné prezenze jsou násobeny společnými absencemi. Tímto se eliminuje chybný výsledek, že dva druhy jsou svým rozšířením považovány za podobné pouze kvůli společným absencím [94].

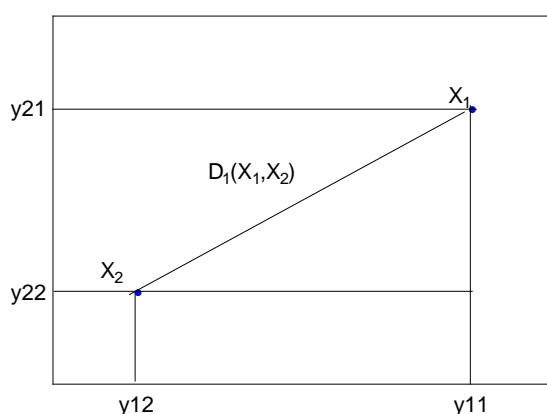
8.3.2 Kvantitativní koeficienty podobnosti a vzdálenosti společenstev

V případě použití abundancí taxonů je opět nutné rozhodnout, zda bude použit některý koeficient pro data na krátkém environmentálním gradientu, tedy bez problému dvojité nuly, nebo koeficient řešící tento problém na dlouhém environmentálním gradientu.

V případě nepřítomnosti problému dvojité nuly je možné využít základní euklidovskou metriku, která počítá vzdálenost objektů obdobně, jako Pythagorova věta počítá přeponu pravoúhlého trojúhelníku (obr. 8.4). Metoda je citlivá na rozdílný rozsah hodnot vstupujících proměnných (vhodným řešením může být standardizace) a problém dvojité nuly. Nemá horní hranici hodnot a standardně je vyjádřena ve formě vzdálenosti.

$$D(x_1, x_2) = \sqrt{\sum_{j=1}^p (y_{1j} - y_{2j})^2}, \quad (8.17)$$

kde x_1 a x_2 jsou srovnávaná společenstva, p je počet taxonů ve společenstvu a y_{1j} a y_{2j} jejich abundance ve společenstvech.



Obr. 8.4 Princip výpočtu euklidovské vzdálenosti

Její variantou je manhattanská vzdálenost, kdy jde o součet rozdílů jednotlivých parametrů popisujících objekty.

$$D(x_1, x_2) = \sum_{j=1}^p |y_{1j} - y_{2j}| \quad (8.18)$$

Modifikací euklidovské vzdálenosti je takzvaná tětivová vzdálenost (chord distance) [95]. Odstraňuje problém dvojité nuly a vliv rozdílného počtu jedinců druhů ve vzorcích při výpočtu euklidovské vzdálenosti. Její maximální hodnota je 1 a minimum 0. Při výpočtu počítá pouze s poměry druhů v rámci jednotlivých vzorků. Jde vlastně o euklidovskou vzdálenost počítanou pro vektory vzorků standardizované na délku 1, nebo je možný přímý výpočet už zahrnující standardizaci.

$$D(x_1, x_2) = \sqrt{2 \left(1 - \frac{\sum_{j=1}^p y_{1j} y_{2j}}{\sqrt{\sum_{j=1}^p y_{1j}^2 \sum_{j=1}^p y_{2j}^2}} \right)} \quad (8.19)$$

Vnitřní část výpočtu je vlastně cosinus úhlu svíraného vektory, zápis vzorce je možný i v této formě.

$$D = \sqrt{2(1 - \cos \theta)} \quad (8.20)$$

Whittakerův asociační index [96] je dobře použitelný pro data abundancí, každý druh je nejprve transformován ve svůj podíl ve společenstvu, následný výpočet je opět jako u manhattanské vzdálenosti.

$$D(x_1, x_2) = \frac{1}{2} \sum_{j=1}^p \left| \frac{y_{1j}}{\sum_{j=1}^p y_{1j}} - \frac{y_{2j}}{\sum_{j=1}^p y_{2j}} \right| \quad (8.21)$$

Jeho hodnota je 0 v případě identických proporcí druhů. Stejný výsledek lze získat i jako součet nejmenších podílů v rámci obou vzorků.

$$D(x_1, x_2) = \left[1 - \min \left(\frac{y_j}{\sum_{j=1}^p y_j} \right) \right] \quad (8.22)$$

Canberrská vzdálenost [97] je opět varianta manhattanské vzdálenosti (před výpočtem musí být odstraněn problém dvojité přítomnosti). Stejný rozdíl mezi početnými druhy ovlivňuje vzdálenost méně než mezi druhy vzácnějšími.

$$D(x_1, x_2) = \sum_{j=1}^p \left[\frac{|y_{1j} - y_{2j}|}{(y_{1j} + y_{2j})} \right] \quad (8.23)$$

χ^2 metrika [98] náleží do skupiny metrik založených na χ^2 pro výpočet vzdáleností odběrů založených na abundancích druhů nebo jiných frekvenčních datech (nejsou přípustné žádné záporné hodnoty). Data původní matice abundancí/frekvencí Y jsou nejprve přepočítána do matice poměrných frekvencí (součty frekvencí v řádcích (odběry) jsou rovny 1). Jako dodatečné charakteristiky uplatňované při výpočtu jsou spočteny součty řádků y_{i+} a sloupců y_{+j} celé matice $n(i)$ odběrů \times $p(j)$ druhů.

$$Y = \begin{bmatrix} y_{ij} \\ \vdots \\ y_{ij} \end{bmatrix} \begin{bmatrix} y_{i+} \\ \vdots \\ y_{i+} \end{bmatrix} \rightarrow \begin{bmatrix} y_{ij} / y_{i+} \\ \vdots \\ y_{ij} / y_{i+} \end{bmatrix} \quad (8.24)$$

Výpočet odstraňuje problém dvojité nepřítomnosti. Nejjednodušším výpočtem je obdoba euklidovské vzdálenosti.

$$D(x_1, x_2) = \sqrt{\sum_{j=1}^p \left(\frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2}, \quad (8.25)$$

která je dále vážena součty jednotlivých druhů

$$D(x_1, x_2) = \sqrt{\sum_{j=1}^p \frac{1}{y_{+j}} \left(\frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2}. \quad (8.26)$$

Index je možné využít i pro měření vzdálenosti mezi druhy na základě jejich rozložení v odběrech.

χ^2 relativní metrika [99] je podobná χ^2 metrice, ale vážení je prováděno relativní četností řádku v matici místo jeho absolutního součtu, při výpočtu se užívá parametr y_{++} (celkový součet matice). Je využívána také při výpočtu vztahů řádků a sloupců kontingenční tabulky.

$$D(x_1, x_2) = \sqrt{\sum_{j=1}^p \frac{1}{y_{+j}} \left(\frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2} = \sqrt{y_{++}} \sqrt{\sum_{j=1}^p \frac{1}{y_{+j}} \left(\frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2} \quad (8.27)$$

V případě analýz na dlouhém environmentálním gradientu jsou v ekologii dlouhodobě etablovány následující koeficienty podobnosti:

Brayův–Curtisův koeficient [100] - tento koeficient porovnává dvě společenstva z hlediska minimální abundance u každého druhu.

$$C_{BC}(x_1, x_2) = \frac{\sum_{j=1}^p |y_{1j} - y_{2j}|}{\sum_{j=1}^p (y_{1j} + y_{2j})} = 1 - \frac{2W}{A + B}, \quad (8.28)$$

kde W je součet minimálních abundancí druhu, A a B je součet abundancí jednotlivých společenstev.

Často je používána také kvantitativní varianta Sørensenova koeficientu [101]

$$C_N = \frac{2jN}{(aN + bN)}, \quad (8.29)$$

kde aN a bN jsou celkové počty jedinců ve společenstvu na lokalitách A a B a jN je suma abundancí druhů, pokud se druh vyskytuje v obou společenstvech.

Dalším významným koeficientem je Morisitův – Hornův index [102, 103]

$$C_{MH} = \frac{2 \sum (a_i b_i)}{(da + db) * aN * bN} \quad da = \frac{\sum a_i^2}{aN^2}, \quad (8.30)$$

kde aN je počet jedinců druhu ve společenstvu A a a_i počet jedinců druhu i ve společenstvu A (obdobně platí pro společenstvo B).

8.4 Vícerozměrné analýzy s přímou vazbou na analýzu biodiverzity

8.4.1 Shluková analýza

Přímo jsou pro analýzu biodiverzity využitelné shlukovací algoritmy pracující s libovolnou asociační maticí (hierarchické aglomerativní shlukování) nebo postavené na metrice řešící problém dvojí nepřítomnosti (hierarchické divizivní shlukování TWINSpan). Ostatní metody shlukování, jako je například algoritmus k -průměrů, které využívají euklidovské vzdálenosti, jsou použitelné pouze nepřímou, a to po převodu asociační matice do euklidovského prostoru.

8.4.2 Analýza hlavních koordinát – PCoA (klasické, metrické škálování MDS)

Vstupem do PCoA je asociační matice podobnosti, popřípadě nepodobnosti mezi vzorky, která se vypočítá z původní matice. Metoda je použitelná na druhová data, která obsahují velký počet dvojí nepřítomnosti (druhová přítomnost/nepřítomnost nebo abundance jednotlivých druhů). A to díky tomu, že matice vzdálenosti může být počítána na metrikách, které umí problém

dvojí nepřítomnosti odstranit. V ordinačním diagramu jsou podobná nebo blízká pozorování vyznačena blízko sebe a pozorování odlišná jsou od sebe dále [115].

Na rozdíl od běžných ordinačních analýz nespécializovaných na analýzu biodiverzity, kdy hlavní komponenty jsou lineární kombinací originálních nebo standardizovaných proměnných, jsou koordináty v PCoA funkcemi originálních proměnných zprostředkovanými přes funkci podobnosti nebo vzdálenosti.

8.4.3 Korespondenční analýza – CA, Detrendovaná korespondenční analýza – DCA

Korespondenční analýza, také reciproké vážení [105], je v podstatě analýzou kontingenčních tabulek. Zkoumá vztah mezi dvěma proměnnými. Kontingenční tabulka je frekvenční tabulka, kde stav jedné proměnné v řádcích je porovnáván se stavem druhé proměnné ve sloupcích. Data v každé buňce jsou frekvence, počet objektů, který kombinuje obě proměnné.

CA se používá na data dimenzionálně homogenní (všechny proměnné mají stejnou dimenzi) a obsahují pouze pozitivní hodnoty nebo nuly. Pro stanovení vztahu mezi sloupci a řádky se používá koeficient χ^2 (chí-kvadrát). V ekologii se tato metoda používá nejvíce na analýzu druhových dat (přítomnost/nepřítomnost nebo abundance). Je založena na unimodálním modelu a dá se použít na dlouhém gradientu [81].

Jedná se o problém vlastní analýzy, kde vlastní hodnoty asociační matice udávají stupně korespondence (indexy) mezi objekty a parametry. Pomocí těchto „indexů“ jsme schopni určit, která sloupcová kategorie má větší nebo menší váhu. Vysoká vlastní hodnota ukazuje na dlouhý a důležitý gradient [105]. Může být také použita pro jednoduché vážené průměrování. Hodnotí rozdíl daného rozložení objektů proti rozložení objektů, které by nastalo, kdyby objekty byly vzhledem ke všem parametrům umístěny náhodně.

Výpočet korespondenční analýzy zahrnuje tři kroky: (1) Kontingenční tabulka je transformovaná na tabulku příspěvku pomocí Pearsonovy χ^2 statistiky, (2) pomocí metodiky rozkladu na singulární hodnoty se vypočítají vlastní hodnoty a vlastní vektory asociační matice, (3) následně jsou pomocí těchto maticových operací vykresleny ordinační diagramy. Kromě toho, že CA můžeme použít jako ordinační metodu, lze pomocí CA zkoumat blízkost vztahu mezi řádky (nebo sloupci) kontingenčních tabulek.

Grafické znázornění vztahů, které obdržíme z CA, je založeno na myšlence, jak prezentovat všechny sloupce a řádky a interpretovat relativní pozice bodů, jako jsou váhy patřící k danému sloupci a řádku. Systém takto získaných indexů nám bude poskytovat souřadnice každého sloupce a řádku, které následně vykreslíme do grafu. Z něj pak můžeme odvodit, které sloupcové kategorie jsou důležité vůči řádkovým kategoriím a naopak.

Problémem CA je takzvaný podkovovitý efekt (*arch efekt*) vycházející z vazby taxonů na prostředí. Taxony mívají podél gradientu unimodální rozložení, efekt gradientů na vzdálenosti vztahu mezi objekty bývá počítán na druhových datech a je nevyhnutelně nelineární. Ordinační metody usilují o překlad této nelinearity do euklidovského dvourozměrného prostoru. Tím vzniká tento artefakt výpočtu. Tvar této křivky záleží na použité vzdálenosti mezi objekty. Uvedený problém můžeme odstranit pomocí zbavení se trendu CA (detrendováním), to znamená použít její detrendovanou formu (DCA). Detrendování je operace prováděná na ordinační ose CA, jejíž pomocí se eliminuje podkovovitý efekt na gradientu tak, že vyjde najevo lineární uspořádání lokalit [81].

8.4.4 Mnohorozměrné nemetrické škálování – NMDS

Jedná se o neparametrickou metodu, která stejně jako PCoA dokáže pracovat s libovolnou asociační maticí, tedy i maticí indexů β diverzity. NMDS se velmi liší od ostatních ordinačních technik. Ve většině ostatních ordinačních metod se počítá mnoho dimenzí, ale jen dvě bývají

zobrazeny. V MDS se na počátku zadá počet dimenzí a následuje výpočet těchto dimenzí tak, aby počet zadaných dimenzí co nejlépe vysvětloval datovou matici. Výpočet běží iterativně a zastaví se na nejlepším možném výsledku nebo po daném počtu iterací. Nepracuje s vlastními čísly, ani vlastními vektory asociační matice.

Metoda používá pouze hodnot (dimenze) nebo informaci z asociační matice podobnosti nebo nepodobnosti. Hodnota matice nám určuje počet lineárně nezávislých řádků matice. Objevují se případy, kde přesné zachování vzdáleností není tak důležité, prioritou místo toho je reprezentovat objekty v menším, specifikovaném počtu dimenzí (často dvě nebo tři, počet nezávislých řádků v matici) [81]. Záměrem NMDS je převod libovolné matice podobnosti do euklidovského prostoru. Běžné NMDS techniky se pokouší lokalizovat proměnné v redukovaném prostoru tak, že vzdálenost v ordinačním prostoru má stejnou hodnotu pořadí jako vzdálenosti v asociační matici. Pak algoritmus používá pouze hodnotu pořadí vzdáleností a ne jejich velikost.

Výhoda této metody je, že nepředpokládá linearitu v datech a je méně citlivá na podkovovitý efekt. Není ovlivněna skupinami proměnných, nepravidelně rozloženými pozorováními podél skrytých gradientů, odlehlými hodnotami ani mírným šumem v datech.

NMDS má také určitá omezení. Za prvé, nemůžeme se spolehnout na získání správné konfigurace, dokonce ani na nasimulovaných datech. Část tohoto problému náleží k prvnímu kroku analýzy, ve které jsou data přepočítávána na asociační matici. Tato matice neobsahuje explicitní informaci o proměnné (pouze podobnosti proměnných), a proto jsou ochuzeny ve srovnání s originální datovou maticí. Za druhé, výběr počáteční konfigurace používaný v NMDS je jeden z největších problémů této metody. Když počáteční konfigurace není vhodně stanovena, je možné, že algoritmus bude umístěn v lokálním minimu, čímž se získávají falešná ordinační skóre. Protože volba počáteční konfigurace je libovolná, nezískáme unikátní řešení. Za třetí, metoda předpokládá, že počet dimenzí datové matice je známý. V praxi se vědec musí spoléhat na své znalosti ohledně optimálního počtu dimenzí. NMDS je citlivé vůči nesprávnému stanovení dimenzionality. Ke stanovení dimenzionality můžeme použít jinou ordinační metodu nebo také Shepardův diagram.

Pro každou iteraci MNDS počítá odchylku od počáteční konfigurace pomocí euklidovské vzdálenosti. Na této vzdálenosti se provádí regrese podle původní datové matice a jsou předpovídány ordinační vzdálenosti pro každý vzorek. Čtverec rozdílu odchylky počáteční konfigurace od předpovídané vzdálenosti se podle regrese používá ke stanovení dobrého „dosednutí“ regrese. Toto „dosednutí“ se nazývá stres a může být počítáno několika způsoby, nejčastěji se používá vztah pro Kruskalův stres (*Kruskal's stress formula*):

$$stres = \sqrt{\frac{\sum_{h,i} (d_{hi} - \hat{d}_{hi})^2}{\sum_{h,i} \hat{d}_{hi}^2}}, \quad (8.31)$$

kde d_{hi} je vzdálenost od počáteční konfigurace a \hat{d}_{hi} je předpovídaná vzdálenost podle regrese.

Díky tomu, že NMDS může počítat nad libovolnou maticí vzdáleností, lze tímto odstranit problém dvojité nuly, což je největší výhoda NMDS.

8.4.5 Kanonická analýza

Všechny zatím zmíněné techniky vícerozměrných metod byly nepřímé (*unconstrained ordination*). Počítají nad proměnnými, kde není definovaná závislá a nezávislá proměnná, a snaží se získat model z těchto proměnných bez ohledu na jiné proměnné, které nejsou v této sadě proměnných obsaženy. Nepopisují vztah mezi tímto modelem a kauzálními faktory, ani neinformují, jak zacházet se vzniklým modelem. Pouze nám popisují hlavní model, inherentní pro daná data. Nepřímé ordinační techniky bohužel neřeknou, co nalezený model v datech znamená a čím je způsoben. Osy v nepřímých ordinacích jsou nejdůležitější gradienty v datech, přesto bývá

jejich interpretace značně složitá. Interpretace ekologického významu os musí být spojena s jiným ekologickým faktorem v izolované sekundární analýze. Často je lepší rovnou umístit jednu část proměnných na osu, která je kombinací druhé části proměnných, zvláště když první proměnné jsou druhová data (abundance) a druhé jsou environmentální proměnné. Tomuto se říká přímá ordinace (*constrained ordination*) [105].

Při tomto druhu vícerozměrné analýzy rozdělujeme datovou matici do dvou skupin: první jsou nejčastěji druhová data, označovaná také závislé proměnné nebo také matice Y, a druhou skupinou bývají environmentální proměnné, které označujeme jako nezávislé, vysvětlující nebo matici X.

8.4.6 Kanonická korespondenční analýza – CCA

Kanonická korespondenční analýza je přímá ordinační metoda. V podstatě jde o hybrid mezi ordinací a mnohorozměrnou regresí. Výpočet CCA vyžaduje přidání vícenásobné regrese metodou nejmenších čtverců [117], druhů na environmentální proměnné funkcí váženého průměrování. CCA extrahuje gradient z dat tak, že může být počítán pomocí měřené vysvětlující veličiny. To se ideálně děje, když proměnné první matice (závislé, druhová data) mají Gaussovo rozdělení s ohledem na kombinace v druhé matici (nezávislé, faktory prostředí). CCA je dvojitá ordinace vzorků a druhů. V CCA ordinačním diagramu, kterému se říká triplet, rozložení druhů a vzorků dohromady představuje dominantní ekologický vztah. Vzdálenost mezi nimi nám říká, jak můžeme tento vztah vysvětlit pomocí vysvětlované proměnné. Největší výhodou této metody oproti nepřímým metodám je, že osy jsou lineární kombinací nezávislých proměnných tak, že vysvětlují nejvíce variability v závislých proměnných. To nám okamžitě umožňuje pochopit význam osy a způsob, jak jsou nezávislé proměnné spojeny s distribucí závislých proměnných.

Vysvětlující proměnné jsou v tripletu vyznačeny jako linie. Jejich směr v redukovaném prostoru poukazuje na směr největší změny strukturní proměnné. Délka linie je rovna stupni změny váženého průměrování podél osy a míry, jak moc se mění distribuce druhů podél vysvětlující proměnné. Obecně, délka linie ukazuje důležitost environmentální proměnné. Jejich směr vztahující se k ose říká, co daná osa reprezentuje. Kosinus úhlu mezi linií a osou je korelační koeficient mezi proměnnou a osou. Podobně úhly mezi liniemi indikují korelační koeficient mezi environmentálními proměnnými. „Druhové“ body a linie společně odrážejí rozdělení druhů na každé environmentální proměnné. Linie a osy společně odrážejí kompozici hlavního ekologického gradientu v datech.

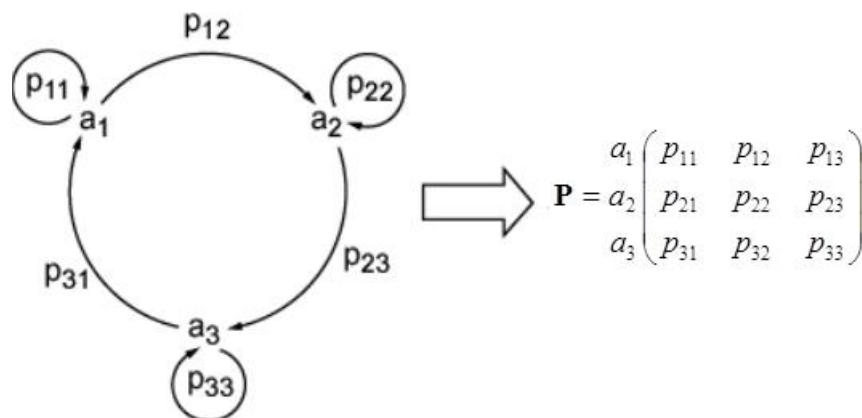
Pozice druhů vztahující se k liniím indikuje environmentální podmínky každého vzorku. Potom lokalizace „druhových“ bodů související s liniemi ukazuje na charakteristiku optima každého druhu. Abundance nebo pravděpodobnost výskytu druhů se snižuje se vzdáleností od druhového bodu. Kromě toho, protože každý vzorek leží na centroidu výskytu druhu na lokalitě, diagram ukazuje, které druhy se více vyskytují na kterých místech.

9 Využití markovských řetězců v analýze biodiverzity

Složitějším přístupem, který do hodnocení společenstev zavádí i pojem času, jsou markovské řetězce. Markovské řetězce popisují systémy, které jsou definovány svými stavy, které nastávají v různém čase (a_1, a_2, \dots, a_n), přičemž mezi stavy (např. typ společenstva na lokalitě) existuje pravděpodobnost přechodu (obr. 9.1). Může být popsán pomocí přechodové matice:

$$\mathbf{P} = \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}, \quad (9.1)$$

kde $p_{m,n}$ je pravděpodobnost přechodu ze stavu a_m do stavu a_n v jednom časovém kroku systému. Pro přechodovou matici platí, že součet v jejích řádcích je vždy jedna. Na základě znalosti stavu systému na počátku $\mathbf{a}^{(0)} = [a_1, a_2, a_3]$ a přechodové matice \mathbf{P} můžeme zjistit stav systému v čase n kroků pomocí jednoduchého vzorce $\mathbf{a}^{(n)} = \mathbf{a}^{(0)}\mathbf{P}$. V případě dodržení určitých podmínek pro matici \mathbf{P} (regulární matice) je pro matici možné spočítat takzvaný pevný bod \mathbf{w} , pro který platí $\mathbf{w} = \mathbf{w}\mathbf{P}$, interpretace tohoto pevného bodu je, že jde o stabilní poměr stavů $[a_1, a_2, a_3]$, ke kterému systém (za předpokladu, že matice \mathbf{P} je regulární, což lze snadno ověřit) směřuje za dlouhý počet kroků. Pevný bod tedy představuje poměr stavů systému (např. ve významu typů společenstev), který se v dlouhodobém horizontu stabilizuje.



Obr. 9.1 Systém popsatelný markovským řetězcem

Výše popsaný markovský řetězec představuje nejjednodušší případ, kdy se přechodová matice \mathbf{P} nemění s časem, jde o takzvaný homogenní markovský řetězec. Markovský řetězec může obsahovat také takzvané absorbující stavy, to jsou stavy, z nichž se už systém nemůže dostat jinam (úmrtí jedince, zánik lokality apod.). Kromě homogenních existují ještě nehomogenní markovské řetězce, které používají při výpočtu několik přechodových matic, typickým příkladem jsou cyklické jevy, jako je střídání ročních období. Při d různých přechodech, kde $d \in \mathbb{N}$, pak pro přechodovou matici \mathbf{P} platí $\mathbf{P} = (n.d + s) = \mathbf{P}(s)$, kde s je číslo přechodu. Kromě výhod definice stabilního stavu systému a predikce jeho stavu mají markovské řetězce také nevýhody. Pro analýzu společenstev je nejkritičtější potřeba velkých souborů dat, které by navíc měly být získávány po dlouhé časové období. Markovské řetězce jsou používány pro modelování řady biologických jevů, mimo jiné i pro modelování stavu společenstev [118, 119]. Nehomogenní markovské řetězce byly použity s dobrými výsledky například pro studium změn společenstev zoobentosu v jednotlivých ročních obdobích [123]. V parazitologii jsou markovské řetězce využívány často pro modelování dynamiky infekcí a v epidemiologii [120-122].

10 Software pro analýzu biodiverzity

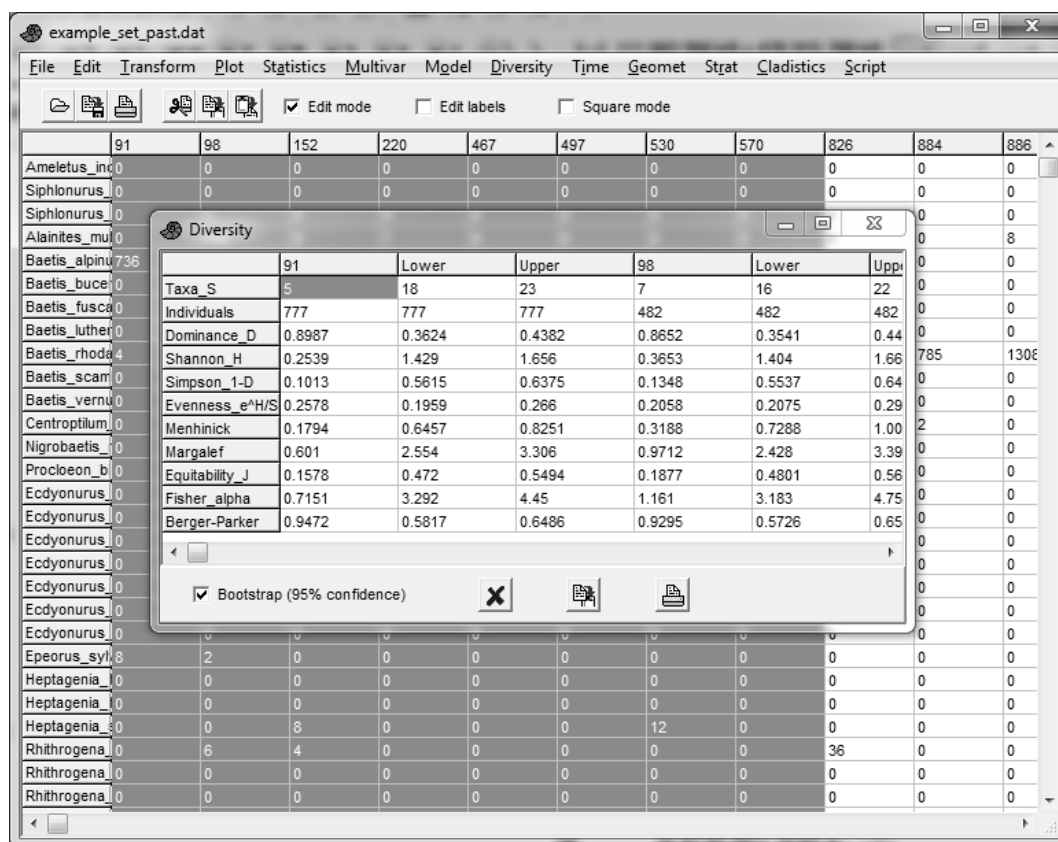
10.1 Standardní komerční statistické softwary

Biodiverzitu lze analyzovat celou řadou statistických softwarů, od standardních komerčních statistických balíčků až po nejrůznější volně šiřitelné nástroje.

Běžné komerční software (SPSS – www.spss.com, Statistica – www.statsoft.com, SAS www.sas.com, Stata – www.stata.com, S-PLUS - spotfire.tibco.com, CANOCO) umožňují standardní statistické výpočty (popisná statistika, testy, modelování), ale s výjimkou vícerozměrné analýzy dat (asymetrické koeficienty podobnosti, korespondenční analýza, nemetrické mnoho-rozměrné škálování) nedisponují specializovanými biodiverzními výpočty (indexy diverzity, modely profilů abundance). Pro biodiverztní výpočty jsou pak vhodnější specializované nástroje představené v následujících kapitolách.

10.2 PAST

Software PAST (Paleontological STatistics) stažený volně z folk.uio.no/ohammer/past/ je univerzálním statistickým nástrojem vyvinutým původně pro analýzu paleontologických dat (obr. 10.1).



Obr. 10.1 Software PAST

Z hlediska analýzy biodiverzity umožňuje:

- výpočet indexů diverzity včetně bootstrapových odhadů jejich intervalů spolehlivosti,
- testování indexů mezi společenstvy pomocí asymptotického i bootstrapového výpočtu,
- rarefakce,
- matematické modely profilů společenstev,
- širokou škálu vícerozměrných statistických analýz společenstev.

10.3 Power niche 1.0

Software Power niche je volně stažitelný z adresy

- www.entu.cas.cz/png/PowerNiche/NicheHlp.html
- jako makro pro MS Excel a obsahuje Tokeshiho modely profilu abundancí.

10.4 EstimateS

Software EstimateS 8 je dostupný ke stažení po registraci na viceroy.eeb.uconn.edu/EstimateSPages/AboutEstimateS.htm#MajorFeatures. Obsahuje následující biodiverzitní výpočty:

- výpočet indexů diverzity včetně bootstrapových a jackknife odhadů jejich intervalů spolehlivosti,
- rarefakce,
- indexy podobnosti společenstev.

10.5 MVSP

MVSP je komerční software (www.kovcomp.com/mvsp/index.html) umožňující:

- výpočet indexů diverzity,
- širokou škálu vícerozměrných analýz zaměřených na data společenstev včetně běžně nedostupných asociačních koeficientů (Gower, Canberra, Baroni-Urbani a jiné).

10.6 PC-ORD

PC-ORD je komerční software (home.centurytel.net/~mjm/pcordwin.htm), jehož hlavní těžiště leží ve vícerozměrné analýze ekologických dat, doplňující funkcí je výpočet indexů diverzity.

10.7 R a jeho knihovny specializované na analýzu biodiverzity

Software R je univerzálním, volně šiřitelným nástrojem pro analýzu dat (www.r-project.org). Pro software R je k dispozici celá řada knihoven (*packages*) pro analýzu biodiverzity:

Vegan

Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens and Helene Wagner (2011). *vegan: Community Ecology Package*. R package version 1.17-10. <http://CRAN.R-project.org/package=vegan>

- poskytuje nástroje pro analýzu diverzity a složení společenstva
- poskytuje nástroje pro vícerozměrnou analýzu dat.

Labdsv

- ordinační analýza především pro data o rostlinných společenstvech.

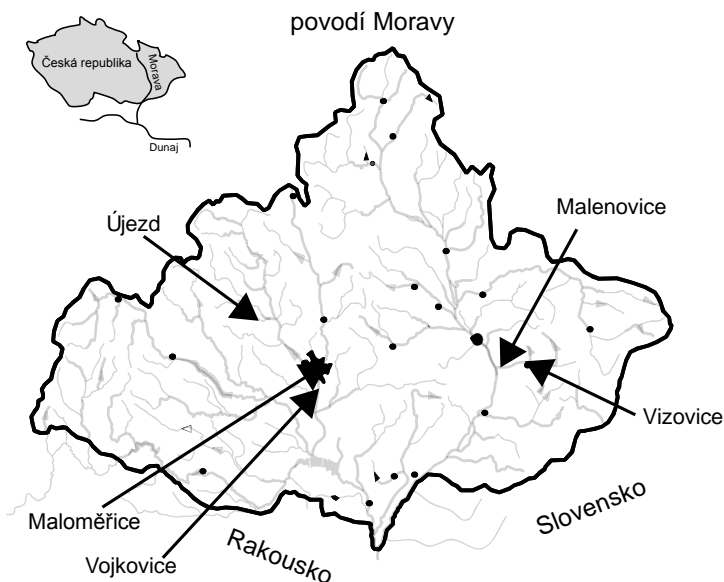
BiodiversityR

Kindt, R. & Coe, R. (2005) *Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies*. World Agroforestry Centre (ICRAF), Nairobi. ISBN 92-9059-179-X:

- poskytuje GUI (Graphical User interface),
- statistické funkce využívající knihovnu “vegan“,
- analýza biodiverzity a ekologických společenstev zahrnující indexy diverzity, Renyiovy profily, zobecněné lineární modely pro analýzu druhové abundance a presence-absence, matici vzdáleností, Mantel test, shlukovou analýzu, přímou a nepřímou ordinační analýzu.

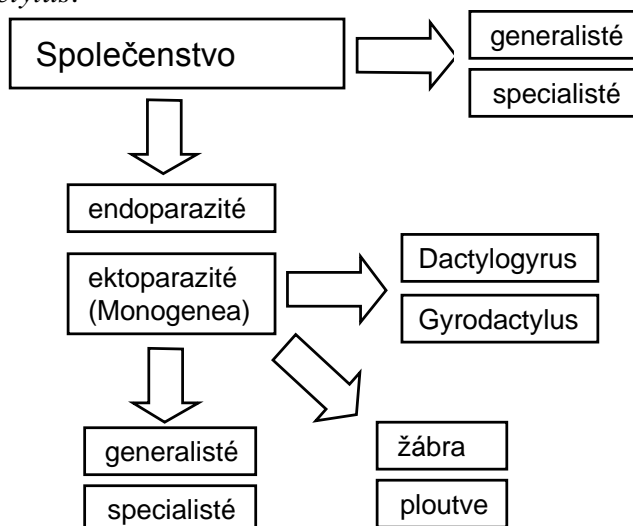
11 Případová studie: Parazitární společenstva

Parazitární společenstvo jelce tlouště (*Leuciscus cephalus*) bylo sledováno na celkem pěti lokalitách lišících se úrovní organického znečištění a teplotou vody (obr. 11.1).¹



Obr. 11.1 Odběrové lokality

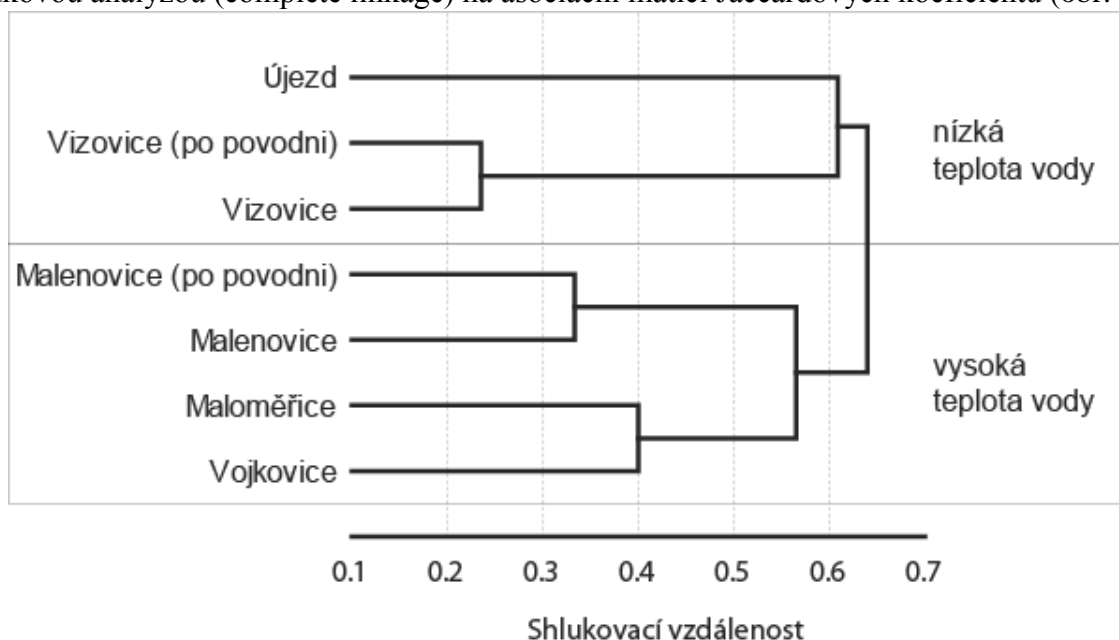
Pro minimalizaci vlivu sezóny a velikosti hostitele byly odchyty prováděny pouze v letních měsících a zaměřeny na ryby přibližně stejné velikosti. Na dvou lokalitách došlo v průběhu zkoumání ke katastrofickým povodním, které se tak staly dalším studovaným faktorem. Protože různé složky parazitárního společenstva mohou vykazovat rozdílnou reakci na různé stresující faktory (obr. 11.2), byly kromě celého společenstva studovány i jeho funkční nebo taxonomické podsložky – endoparazité/ektoparazité; generalisté/specialisté (celé společenstvo a Monogenea); *Dactylogyrus*/*Gyrodactylus*.



Obr. 11.2 Frakcionace společenstva

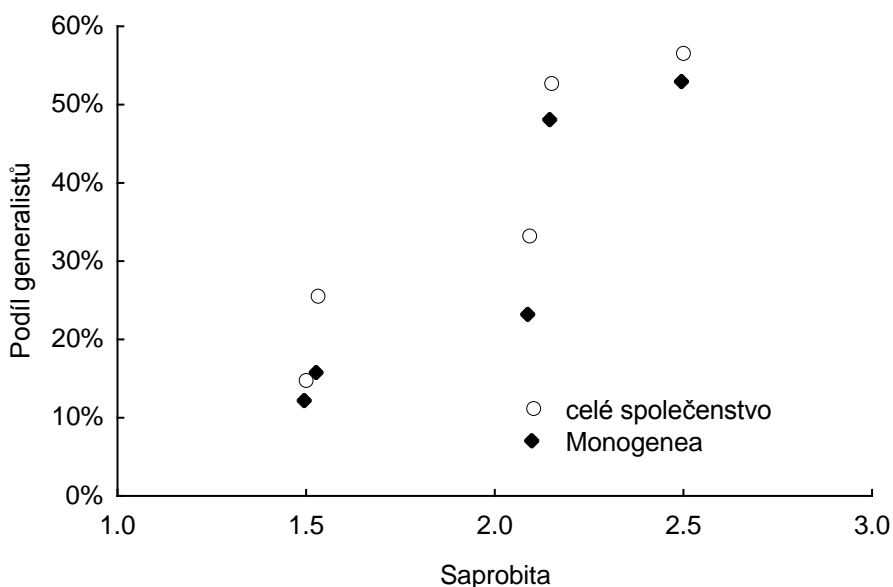
¹ Případová studie vychází z nepublikované práce Jarkovský J., Gelnar M. „Parasite communities of *Leuciscus cephalus* under different environmental conditions“.

Použitými metodami pro srovnání společenstev v různých podmínkách byly indexy diverzity (stochastická Q statistika a Bergerův-Parkerův index), druhy společenstev byly popsány standardními epidemiologickými charakteristikami. Druhové složení společenstev bylo srovnáno shlukovou analýzou (complete linkage) na asoiační matici Jaccardových koeficientů (obr. 11.3).



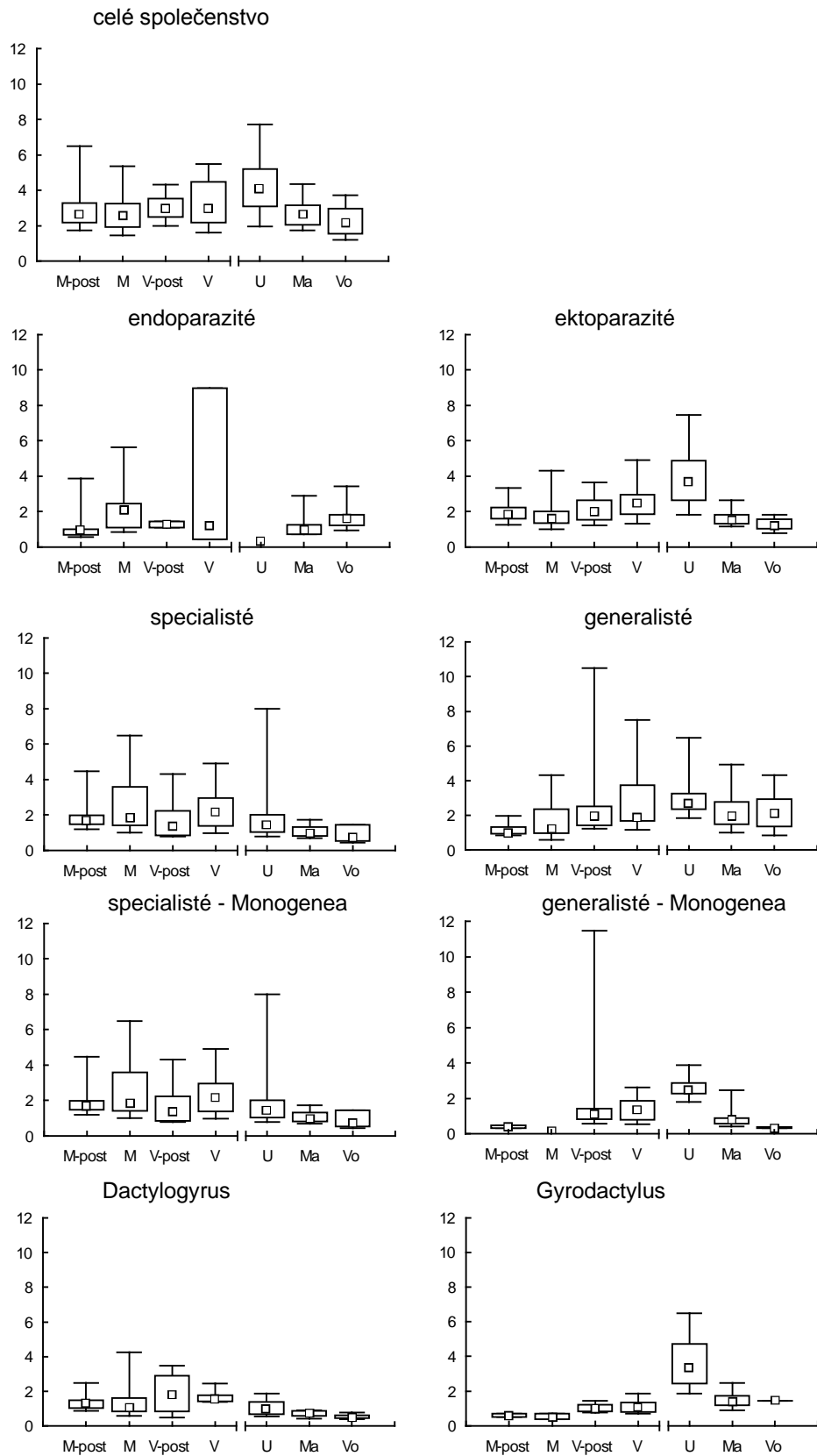
Obr. 11.3 Shluková analýza společenstev

Celkový vzorek představoval 187 ryb s celkovým počtem 4994 parazitů patřících do 28 druhů, největší podíl jak jedinců, tak druhů představovala Monogenea, s výjimkou jediné lokality představovala Monogenea také všechny dominantní druhy.



Obr. 11.4 Vztah mezi organickým zatížením lokalit a podílem generalistů ve společenstvu

Druhové složení společenstev se zdá být určováno zejména lokálními podmínkami lokalit a teplotou vody, při srovnání byly jako nejpodobnější zjištěny před a povodňové vzorky, další úroveň podobností druhového složení byla spjata s teplotou vody, kdy lokality s vyšší a nižší teplotou vody vytvořily dva oddělené shluky. Při zhodnocení vlivu působících faktorů byl zjištěn nižší účinek povodní na diverzitu společenstva, nejmarkantnějším důsledkem zde byl pokles dominance ve společenstvu. Tento pokles dominance byl zjištěn také v případě vyšší úrovně organického znečištění a zdá se, že jde o obecný trend v přítomnosti stresujícího faktoru.



Obr. 11.5 Vztah mezi diverzitou společenstev (Q statistika) a lokalitami s různými environmentálními podmínkami

Vliv hladiny organického znečištění byl již zřetelnější, zejména při rozdělení společenstva na podskupiny. Nejvýraznější byl vzestup diverzity společenstva ektoparazitů (Monogeneí) v podmínkách nižšího znečištění a opačný trend u společenstva endoparazitů, který může souviset s vyšší abundancí jejich mezipřenositelů v podmínkách vyššího organického znečištění. V případě rozdělení společenstva podle specializace byl u generalistů zjištěn nárůst podílu jejich jedinců v podmínkách vyššího organického znečištění, který může souviset s jejich lepším uplatněním ve společenstvu v případě stresu hostitele (obr. 11.4). Pokud bylo společenstvo Monogeneí rozděle-
no na rody *Dactylogyrus* a *Gyrodactylus*, byl u rodu *Dactylogyrus* zjištěn trend poklesu diverzity v podmínkách organicky znečištěných lokalit, který však nebyl statisticky významný, naproti tomu rod *Gyrodactylus* se zdá spíše inertní vzhledem k organickému znečištění; jeho diverzita byla ovšem silně ovlivňována teplotními podmínkami lokalit, kdy vysoká diverzita společenstva *Gyrodactylů* byla spjata s nižší teplotou vody na lokalitě (obr. 11.5).

Závěrem lze říci, že v analyzovaných datech byl nalezen vliv stresujících faktorů na společenstva cizopasníků, tento vliv byl výraznější v případě organického znečištění a to zejména po rozdělení společenstva na jeho logické funkční celky. Nejobecnějším ukazatelem vlivu stresujícího faktoru pak byl pokles dominance ve společenstvu, týkající se jak organického znečištění, tak vlivu povodní.

12 Případová studie: Lišejníky a znečištění ovzduší

Předmětem vyhodnocení jsou soubory dat získané z různých oblastí České republiky v rámci dílčích studií o využití lišejníků k indikaci imisní zátěže v lesních porostech.² Všechny oblasti byly hodnoceny pomocí stejné metodiky, což umožňuje jejich společné vyhodnocení. Cílem analýzy je:

- zhodnocení biodiverzity lišejníkových společenstev pod vlivem imisní zátěže,
- využití parametrů biodiverzity k bioindikačním účelům.

Soubor dat reprezentuje situaci v pohraničních horách ČR, zahrnuje 5 oblastí, které se od sebe výrazně liší v celkové imisní zátěži a postihují prakticky celou škálu od velmi vysoké zátěže až po nízkou v následujícím pořadí:

- Jizerské hory (JH)
- Beskydy (BE)
- Orlické hory (OH)
- Kralický Sněžník (KS)
- Novohradské hory (NH).

V každé oblasti bylo vybráno cca 15 lokalit, na každé lokalitě byla popsána lišejníková vegetace na cca 10 až 20 stromech, které vyhovovaly standardizaci (tabulka 12.1).

Tabulka 12.1 Struktura datového souboru

	Oblast	celková imisní zátěž	pořadí	počet lokalit	počet stromů
JH	Jizerské hory	velmi vysoká	1	10	141
BE	Beskydy	vysoká	2 – 3	15	183
OH	Orlické hory	vysoká	2 – 3	12	239
KS	Kralický Sněžník	střední	4	10	122
NH	Novohradské hory	nízká	5	15	189

Na každém stromě byla popsána přítomnost 10 bioindikačních lišejníkových druhů s různou citlivostí k emisím. Vzhledem ke specifickým společenstvům lišejníků jsou u každého druhu sledovány abundance a vitalita pomocí kvantitativní stupnice, jako ukazatel kvantitativní stránky společenstva je potom využit jejich součin, takzvaná indikační kapacita, která popisuje celkový stav druhu na stromě a jeho perspektivu (tabulka 12.2).

Tabulka 12.2 Popis kvantitativní stránky společenstva lišejníků

Abundance	Vitalita	Indikační kapacita
5 - velmi hojně se vyskytující druh	1,0 - stélky normálně vyvinuté, bez známek poškození a zakrnění	Stupnice od 5,0 – ideální stav do 0,2 druh prakticky již ustoupil – zůstává 1 odumřelá stélka).
4 - hojně se vyskytující druh	0,8 - stélky zakrnělé, případně mírně poškozené	Druh, který se na stromě (lokalitě) nevyskytuje, má kapacitu $k = 0$
3 - středně se vyskytující	0,6 - stélky s výraznými stopami poškození a odumírání	
2 - málo se vyskytující, pouze několik stélek	0,4 - stélky z velké části odumřelé	
1 - jedna nebo dvě stélky	0,2 - stélky zcela odumřelé	

² Data případové studie byla pro výukové účely poskytnuta RNDr. Petrem Andělem, CSc.

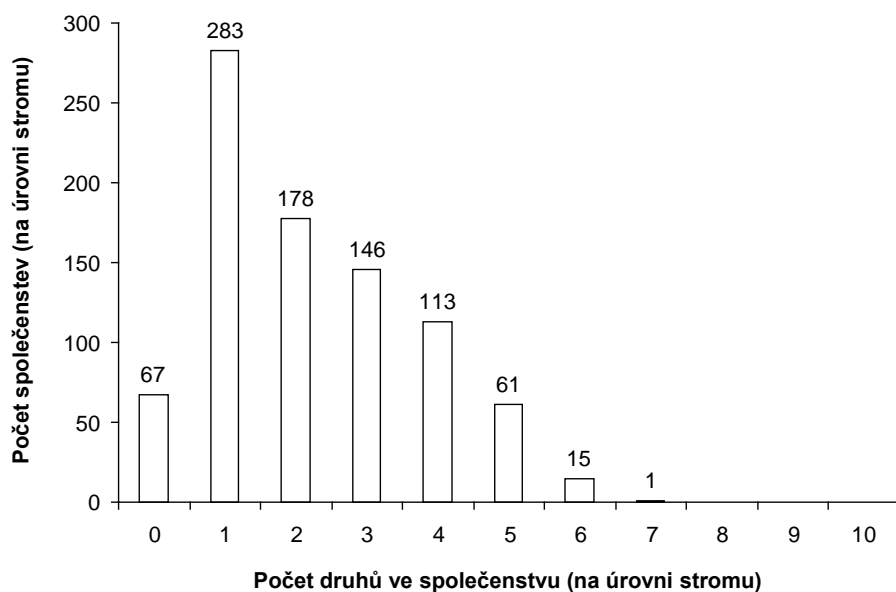
První krok analýzy pracuje se základní úrovní sběru dat, tedy s jednotlivými stromy (obr. 12.1).

	1 oblast	2 NewVar	3 lokalita	4 strom	5 hphf	6 pamf	7 psaf	8 pfuf	9 pglf	10 cchf	11 cpif	12 phyf	13 bryf	14 usnf
52	BE	BE_4	4	3	3	0	0	0	1	0	0	0	0	0
53	BE	BE_4	4	4	2	0	0	0	0	0	0	0	0	0
54	BE	BE_4	4	5	3	0	0	0	0	0	0	0	0	0
55	BE	BE_4	4	6	5	0	0	0	0	0	0	0	0	0
56	BE	BE_4	4	7	4	0	0	0	0	0	0	0	0	0
57	BE	BE_5	5	1	3	0	0	0	0	0	0	0	0	0
58	BE	BE_5	5	2	3	0	0	0	0	0	0	0	0	0
59	BE	BE_5	5	3	3	0	0	0	0	0	0	0	0	0
60	RF	RF_5	5	4	4	n	n	n	n	n	n	n	n	n

Obr. 12.1 Datový soubor společenstev na úrovni jednotlivých stromů

Jak ale ukazuje obrázek 12.2, většina společenstev na úrovni stromů obsahuje jeden až tři druhy, což poskytuje pouze malý prostor pro aplikaci biodiverzitních indexů, které mezi jednotlivými lokalitami nebo oblastmi vykazují minimální rozdíly. Rozmístění druhů na jednotlivých stromech navíc nemusí souviset s celkovou imisní zátěží lokalit, ale spíše s jinými faktory jako je vzájemná kompetice druhů a jiné lokální vlivy; biodiverzita na této úrovni není pravděpodobně dobrým ukazatelem imisního zatížení lokalit.

Z výše uvedených důvodů byla před další analýzou provedena agregace dat na úroveň jednotlivých lokalit, kdy výsledná indikační kapacita druhu je sumou indikačních kapacit na jednotlivých stromech na lokalitě. Výsledný soubor má již pro srovnání různých oblastí daleko variabilnější data odrážející podmínky na celé lokalitě, potažmo oblasti (obr. 12.3).



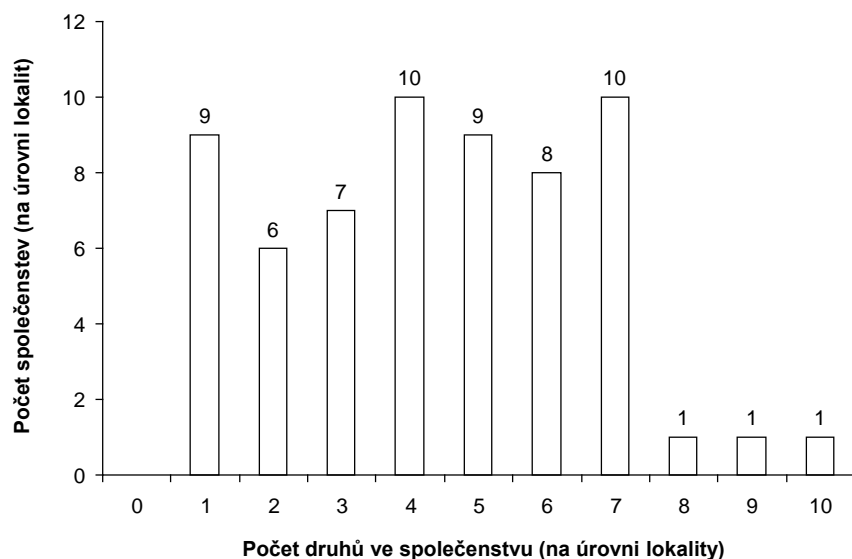
Obr. 12.2 Počet druhů ve společenstvu na úrovni jednotlivých stromů

Pro společenstva na úrovni lokalit bylo spočteno indexové hodnocení biodiverzity, byl použit Shannonův-Weaverův index, permutační varianta Q statistiky a Shannonova-Weaverova vyrovnanost společenstva. Mezi jednotlivými oblastmi byly zjištěny statisticky významné rozdíly v biodiverzitě společenstev. Největší kontrasty mezi oblastmi byly zjištěny pomocí Q statistiky, která je tak i vhodným indexem pro aplikaci v bioindikaci imisní zátěže oblastí (obr. 12.4).

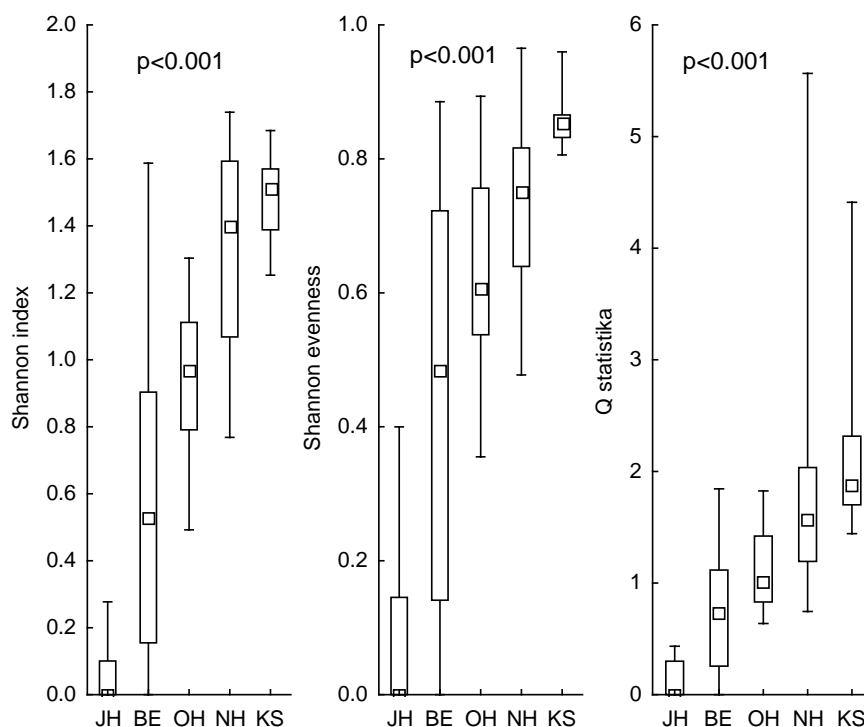
Jako doplněk k biodiverzitním indexům byly v analýze použity Tokeshiho modely druhové abundance. Jednotlivé Tokeshiho modely byly simulovány pro společenstvo o deseti druzích, každý model byl simulován pro 1000 společenstev. Reálná společenstva jednotlivých oblastí

byla vizualizována jako grafy pořadí indikační kapacity a srovnána s 5-95ti percentilovým rozsahem simulovaných Tokeshiho modelů (obr. 12.5). Grafy ukazují jednak výrazné rozdíly v profilu společenstev oblastí, jednak různou míru jejich shody s jednotlivými Tokeshiho modely. Výsledky analýzy lze shrnout:

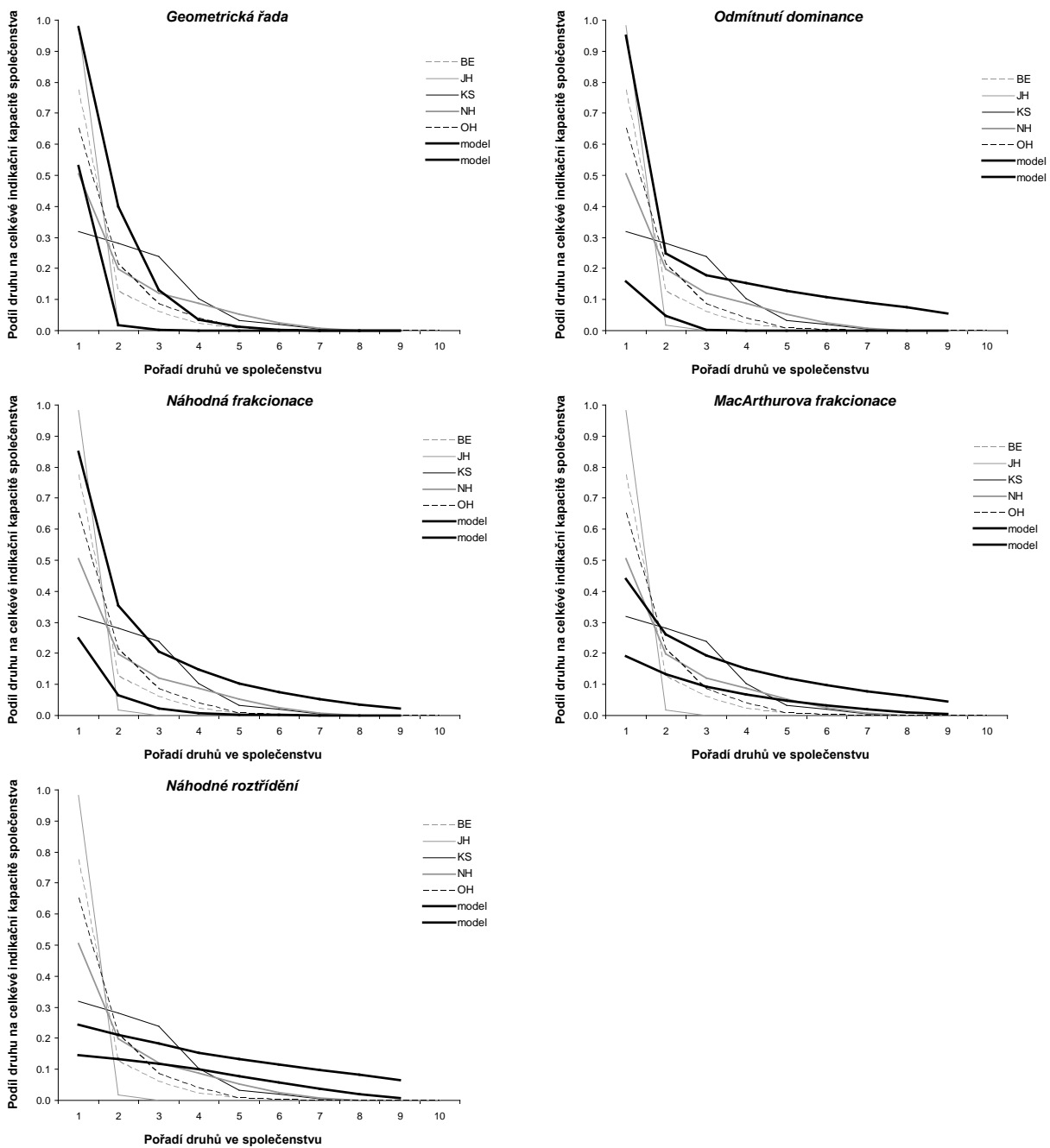
- pro analýzu biodiverzity je třeba zvolit vhodnou úroveň agregace společenstev umožňující zodpovědět položené hypotézy,
- biodiverzitní indexy citlivě odlišily oblasti s různou imisní zátěží a jsou využitelné v biondikaci takto zatížených oblastí,
- pro vizualizaci výsledků lze s úspěchem využít profilů společenstev a jeho srovnání s modely druhových abundancí (poznámka: zde je namísto abundance využita indikační kapacita druhu).



Obr. 12.3 Počet druhů ve společenstvu po agregaci na úrovni lokality



Obr. 12.4 Biodiverzitní indexy v jednotlivých oblastech (popsány mediánem, kvartily a 5-95ti percentilovým rozsahem); statistická významnost hodnocena Kruskalovým-Wallisovým testem.



Obr. 12.5 Srovnání profilů společenstev lišejníků v jednotlivých oblastech s Tokeshiho modely druhové abundance

Seznam doporučené literatury

- [1] Fiedler, R.A., Jain, S.K. Conservation Biology: The Theory and Practice of Nature Conservation, Preservation and Management. Chapman and Hall, New York. (1992)
- [2] DeLong, D. Defining biodiversity. Wildlife Society Bulletin 24: 738-749. (1996)
- [3] Rényi, A. On measures of entropy and information. In On measures of entropy and information. University of California Press. (1961)
- [4] Hill, M. Diversity and Evenness - Unifying Notation And Its Consequences. Ecology 54: 427-432. (1973)
- [5] Krebs, C.J. Ecological Methodology. Harper Collins Publishers: New York. (1989)
- [6] Carignan, V., Villard, M. Selecting indicator species to monitor ecological integrity: A review. Environmental Monitoring and Assessment 78: 45-61. (2002)
- [7] Nicholls, R., Hoozemans, F., Marchand, M. Increasing flood risk and wetland losses due to global sea-level rise: regional and global analyses. Global Environmental Change-Human and Policy Dimensions 9: S69-S87. (1999)
- [8] Boughton, D., Smith, E., O'Neill, R. Regional vulnerability: A conceptual framework. Ecosystem Health 5: 312-322. (1999)
- [9] Detenbeck, N., Batterman, S., Brady, V., Brazner, J., Snarski, V., Taylor, D., Thompson, J., Arthur, J. A test of watershed classification systems for ecological risk assessment. Environmental Toxicology and Chemistry 19: 1174-1181. (2000)
- [10] Williams, L., Kapustka, L. Ecosystem vulnerability: A complex interface with technical components. Environmental Toxicology and Chemistry 19: 1055-1058. (2000)
- [11] Braak, C.J.F. Unimodal models to relate species to environment. DLO-Agricultural Mathematics Group: Wageningen. (1996)
- [12] Poulin R. Toxic Pollution and Parasitism in Fresh-Water Fish. Parasitology Today 8: 58-61. (1992)
- [13] Southwood, T.R.E., Henderson, P.A. Ecological methods. Blackwell Science Ltd.: Oxford, UK. (2000)
- [14] Boyce, M. Restitution of R-Selection and K-Selection as a Model of Density-Dependent Natural-Selection. Annual Review of Ecology and Systematics 15: 427-447. (1984)
- [15] Dickman, M. Some indices of diversity. Ecology 49: 1191-1193. (1968)
- [16] Colwell, R.K. Toward a unified approach to the study of species diversity. In Toward a unified approach to the study of species diversity. International Cooperative Publishing House: Fairland, Md. (1979)
- [17] Odum, E.P. Základy ekologie. Academia. (1977)
- [18] Wilson, E.O., Bossert, W.H. A Primer of Population Biology. Sinauer Ass. (1971)
- [19] Williamson, M.H. The Analysis of Biological Populations. Arnold. (1972)
- [20] Usher, M.B. Developments in the Leslie Matrix Model. Jeffers, J.N.R. (ed). (1972)
- [21] Tokeshi, M. Niche Apportionment or Random Assortment - Species Abundance Patterns Revisited. Journal of Animal Ecology 59: 1129-1146. (1990)
- [22] Tokeshi, M. Species Abundance Patterns and Community Structure. Advances in Ecological Research 24: 111-186. (1993)
- [23] Magurran, A.E. Measuring biological diversity. Blackwell Publishing: Malden, Mass. (2004)
- [24] Cox, F.E.G.. Modern Parasitology. Blackwell Scientific Publication. (1993)
- [25] Bundy, D., Cooper, E., Thompson, D., Didier, J., Simmons, I. Epidemiology and Population-Dynamics of Ascaris-Lumbricoides and Trichuris-Trichiura Infection in the Same Community. Transactions of the Royal Society of Tropical Medicine and Hygiene 81: 987-993. (1987)

- [26] Haswellelkins, M., Elkins, D., Anderson, R. The Influence of Individual, Social Group and Household Factors on the Distribution of *Ascaris-Lumbricoides* within a Community and Implications for Control Strategies. *Parasitology* 98: 125-134. (1989)
- [27] Scott, M.E., Smith G. Parasitic and Infectious Diseases (Epidemiology and Ecology). Academic Press, Inc. (1994)
- [28] Atmar, W., Patterson, B. The Measure of Order and Disorder in the Distribution of Species in Fragmented Habitat. *Oecologia* 96: 373-382. (1993)
- [29] Clifford, H.T., Stephenson, W. An Introduction to Numerical Classification. Academic Press: London. (1975)
- [30] Whittaker, R.H. Evolution of Species Diversity in Land Communities. Hecht. M.K., Steere, W.C., Wallace, B (eds). Plenum: New York. (1977)
- [31] Pielou, E.C. Ecological Diversity. Wiley: New York. (1975)
- [32] Bowman, K., Hutcheson, K., Odum, E.P., Shenton, L.R. Comments on the Distribution of Indices of Diversity. Pennsylvania State University Press: University Park, PA. (1971)
- [33] Hutcheson, K. A Test for Comparing Diversities Based on Shannon Formula. *Journal of Theoretical Biology* 29: 151-160. (1970)
- [34] Pielou, E.C. An Introduction to Mathematical Ecology. Wiley: New York. (1969)
- [35] Simpson, E. Measurement of Diversity. *Nature* 163: 688-688. (1949)
- [36] May, R.M. Patterns of Species Abundance and Diversity. Harvard University Press: Cambridge, MA. (1975)
- [37] Berger, W., Parker, F. Diversity of Planktonic Foraminifera in Deep-Sea Sediments. *Science* 168: 1345-1355. (1970)
- [38] Kempton, R., Wedderburn, R. Comparison of 3 Measures of Species-Diversity. *Biometrics* 34: 25-37. (1978)
- [39] Kempton, R., Taylor, L. Models and Statistics for Species-Diversity. *Nature* 262: 818-820. (1976)
- [40] Kempton, R., Taylor, L. Q-Statistic and Diversity of Floras. *Nature* 275: 252-253. (1978)
- [41] Dusek, L., Gelnar, M., Sebelova, S. Biodiversity of parasites in a freshwater environment with respect to pollution: metazoan parasites of chub (*Leuciscus cephalus* L.) as a model for statistical evaluation. *International Journal for Parasitology* 28: 1555-1571. (1998)
- [42] Colwell, R., Coddington, J. Estimating Terrestrial Biodiversity Through Extrapolation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 345: 101-118. (1994)
- [43] Jarkovsky, J., Brabec, K. Multihabitatový vzorek makrozoobentosu z pohledu variability jeho podjednotek. *Zoologické dny* (2004)
- [44] Simberloff, D. Properties of Rarefaction Diversity Measurement. *American Naturalist* 106: 414-424. (1972)
- [45] Zahl, S. Jackknifing An Index of Diversity. *Ecology* 58: 907-913. (1977)
- [46] Sokal, R.R., Rohlf, F.J. Biometry. Freeman: New York. (1995)
- [47] Adams, J.E., McCune, E.D. Application of the generalized jackknife to Shannon's measure of information used as an index of diversity. International Co-operative Publishing House: Fairland, Maryland. (1979)
- [48] Marvan, P. Primetchnija k primeneniju statističeskich metodovpo opredeleniju saprobnosti. (1969).
- [49] Friedrich, G. A Revision Of The Saprobic System. *Zeitschrift Fur Wasser Und Abwasser Forschung-Journal For Water and Wastewater Research* 23: 141-152. (1990)
- [50] Rosenberg, D.M., Resh, V.H. Freshwater biomonitoring and benthic macroinvertebrates. Chapman & Hall: New York and London. (1993)
- [51] Fisher, R., Corbet, A., Williams, C. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12: 42-58. (1943)

- [52] Preston, F. The Commonness, and Rarity, of Species. *Ecology* 29: 254-283. (1948)
- [53] Preston, F. Canonical Distribution of Commonness and Rarity. *Ecology* 43: 185-195. (1962)
- [54] Anscombe, F. Sampling Theory of the Negative Binomial and Logarithmic Series Distributions. *Biometrika* 37: 358-382. (1950)
- [55] Bliss, C., Fisher, R. Fitting the Negative Binomial Distribution to Biological Data - Note on the Efficient Fitting of the Negative Binomial. *Biometrics* 9: 176-200. (1953)
- [56] Zipf, G.K. *Human Behaviour and the Principle of Least Effort*. Hafner: New York. (1965)
- [57] Mandelbrot, B.B. *The Fractal Geometry of Nature*. Freeman. (1982)
- [58] Motomura, I. A statistical treatment of associations (in Japanese and cited in May, 1975). *Japanese Journal of Zoology* 44: 379-383. (1932)
- [59] MacArthur, R. On the Relative Abundance of Bird Species. *Proceedings of the National Academy of Sciences of the United States of America* 43: 293-295. (1957)
- [60] Sugihara, G. Minimal Community Structure - An Explanation of Species Abundance Patterns. *American Naturalist* 116: 770-787. (1980)
- [61] Hughes, R.A. Model of the Structure and Dynamics of Benthic Marine Invertebrate Communities. *Marine Ecology-Progress Series* 15: 1-11. (1984)
- [62] Hughes, R. Theories and Models of Species Abundance. *American Naturalist* 128: 879-899. (1986)
- [63] Caswell, H. Community Structure - Neutral Model Analysis. *Ecological Monographs* 46: 327-354. (1976)
- [64] Bersier, L., Sugihara, G. Species abundance patterns: the problem of testing stochastic models. *Journal of Animal Ecology* 66: 769-774. (1997)
- [65] Manly, B. On The Statistical-Analysis of Niche Overlap Data. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* 68: 1420-1422. (1990)
- [66] Whittaker, R.H. Evolution and measurement of species diversity. *Taxon* 21: 213-251. (1972)
- [67] Ugland, K., Gray, J. Lognormal Distributions and the Concept Of Community Equilibrium. *Oikos* 39: 171-178. (1982)
- [68] Bulmer, M. Fitting Poisson Lognormal Distribution to Species-Abundance Data. *Biometrics* 30: 101-110. (1974)
- [69] Kempton, R, Taylor, L. Log-Series and Log-Normal Parameters as Diversity Discriminants for Lepidoptera. *Journal of Animal Ecology* 43: 381-399. (1974)
- [70] Lambshead, J., Platt, H.M. *Structural patterns of marine benthic assemblages and their relationship with empirical statistical models*. Cambridge University Press, Cambridge. (1985)
- [71] Bates, T.L.R., Williams, Hutchinson A variety of diversities. Blackwell. (1978)
- [72] Frontier, S. Diversity and Structure in Aquatic Ecosystems. *Oceanography and Marine Biology* 23: 253-312. (1985)
- [73] Gray, N. Ecology of Nematophagous Fungi - Effect of the Soil Nutrients N, P and K, and 7 Major Metals on Distribution. *Plant and Soil* 108: 286-290. (1988)
- [74] Reichelt, R., Bradbury, R. Spatial Patterns in Coral-Reef Benthos - Multiscale Analysis of Sites From 3 Oceans. *Marine Ecology-Progress Series* 17: 251-257. (1984)
- [75] Tokeshi, M. *Species Coexistence: Ecological and Evolutionary Perspectives*. Blackwell Science Ltd.: Oxford. (1999)
- [76] Ives, A. Aggregation and the Coexistence of Competitors. *Annales Zoologici Fennici* 25: 75-88. (1988)
- [77] Ives, A. Aggregation and Coexistence in a Carrion Fly Community. *Ecological Monographs* 61: 75-94. (1991)

- [78] Sevenster, J. Aggregation and coexistence.1. Theory and analysis. *Journal of Animal Ecology* 65: 297-307. (1996)
- [79] Levins, R. *Evolution in Changing Environments: Some Theoretical Exploration*. Princeton University Press: Princeton, New Jersey. (1968)
- [80] Geets, A., Coene, H., Ollevier, F. Ectoparasites of the whitespotted rabbitfish, *Siganus sutor* (Valenciennes, 1835) off the Kenyan Coast: Distribution within the host population and site selection on the gills. *Parasitology* 115: 69-79. (1997)
- [81] Legendre, L., Legendre, P. *Numerical ecology*. Elsevier Science BV. (1998)
- [82] Lepš, J., Šmilauer, P. *Mnohorozměrná analýza ekologických dat*. Biologická fakulta Jihočeské univerzity v Českých Budějovicích: České Budějovice. (2000)
- [83] Sokal, R.R., Michener, C.D. A Statistical Method for Evaluating Systematic Relationships. *The University of Kansas Scientific Bulletin* 38: 1409-1438. (1958)
- [84] Rogers, D., Tanimoto, T. Computer Program for Classifying Plants. *Science* 132: 1115-1118. (1960)
- [85] Sneath, P., Sokal, R. Numerical Taxonomy. *Nature* 193: 855-865. (1962)
- [86] Jaccard, P. Contribution au problème de l'immigration post-glaciaire de la flore alpine. *Bull. Soc. Vaudoise Sci. Nat.* 36: 87-130. (1900)
- [87] Jaccard, P. Etude comparative de la distribution florale dans une portion des Alpes et du Jura. *Bull. Soc. Vaudoise Sci. Nat.* 37: 547-579. (1901)
- [88] Jaccard, P. Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaudoise Sci. Nat.* 44: 223-270. (1908)
- [89] Sørensen, T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analysis of the vegetation on Danish commons. *Bial. Skr.* 5. (1948)
- [90] Russel, P.F., Rao, T.R. On habitat and association of species of anopheline larvae in south-eastern Madras. *J. Malar. Inst. India* 3: 153-178. (1940)
- [91] Kulczynski, S. Die Pijanzenassoziationen der Pieninen. *Bull. Int. Acad. Pol. Sci. Lett. Cl. Sci. Math. Nat. Ser. B, Suppl. II* 57-203. (1928)
- [92] Ochia, A. Zoogeographic studies on the soleoid fishes found in Japan and its neighbouring regions. *Bull. Jpn. Soc. Sci. Fish.* 22.
- [93] Faith, D. Asymmetric Binary Similarity Measures. *Oecologia* 57: 287-290. (1983)
- [94] Baroni-Urbani, C., Buser, M. Similarity of Binary Data. *Systematic Zoology* 25: 251-259. (1976)
- [95] Orloci, L. Geometric Models in Ecology I. Theory and Application of Some Ordination Methods. *Journal of Ecology* 54: 193-203. (1966)
- [96] Whittaker, R. A Study Of Summer Foliage Insect Communities in The Great Smoky Mountains. *Ecological Monographs* 22: 1-44. (1952)
- [97] Lance, G., Williams, W. Computer Programs for Hierarchical Polythetic Classification (Similarity Analyses). *Computer Journal* 9: 60-70. (1966)
- [98] Roux, M., Reyssac, J. Essaid 'applicationa u phytoplanton marin de m6thodes statistiques utilisees en phytosociologie terrestre. *Annales de l'Institut Oceanographique* 51. (1975)
- [99] Lebart, L., Fénelon, J.P. *Statistique et informatique appliquées*. Dunod: Paris. (1971)
- [100] Bray, R.J. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr.* 27: 325-349. (1957)
- [101] Watson, L., Williams, W.T., Lance, G.N. Angiosperm taxonomy: a comparative study of some novel numerical techniques. *J. Linn. Soc. Lond. Bot.* 59. (1966)
- [102] Morisita, M. Measuring of the dispersion and analysis of distribution patterns. *Memoires of the Faculty of Science, Kyushu University, Series E. Biology* 2. (1959)
- [103] Horn, H. Measurement of Overlap in Comparative Ecological Studies. *American Naturalist* 100: 419-429. (1966)

- [104] Meloun, M., Militký, J. Kompendium statistického zpracování dat: metody a řešené úlohy včetně CD. Academia: Praha. (2002)
- [105] McGarigal, K., Cushman, S., Strafgord, S. Multivariate Statistics for Wildlife and Ecology Research. Springer: New York. (2000)
- [106] Jonsson, B., Jonsell, M. Exploring potential biodiversity indicators in boreal forests. *Biodiversity and Conservation* 8: 1417-1433. (1999)
- [107] Bergeron, Y. Species and stand dynamics in the mixed woods of Quebec's southern boreal forest. *Ecology* 81: 1500-1516. (2000)
- [108] Trzcinski, M., Fahrig, L., Merriam, G. Independent effects of forest cover and fragmentation on the distribution of forest breeding birds. *Ecological Applications* 9: 586-593. (1999)
- [109] Vega, M., Pardo, R., Barrado, E., Deban, L. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research* 32: 3581-3592. (1998)
- [110] Hair, J.E., Anderson, R.E., Tatham, R.L. Multivariate Data Analysis. Macmillan Publishers: New York. (1987)
- [111] Tabachnick, B.G., Fidell, L.S. Using multivariate statistics. Harper & Row. (1989)
- [112] Reid, S., Mandrak, N., Carl, L., Wilson, C. Influence of dams and habitat condition on the distribution of redbreast sunfish (*Moxostoma*) species in the Grand River watershed, Ontario. *Environmental Biology of Fishes* 81: 111-125. (2008)
- [113] Gabriel, K. Biplot Graphic Display of Matrices With Application to Principal Component Analysis. *Biometrika* 58: 453-463. (1971)
- [114] Zaniewski, A., Lehmann, A., Overton, J. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling* 157: 261-280. (2002)
- [115] Ludwig, J.A., Reynolds, J.F. Statistical ecology: a primer on methods and computing. John Wiley & Sons: New York. (1988)
- [116] Duque, A., Duivenvoorden, J., Cavelier, J., Sanchez, M., Polania, C., Leon, A. Ferns and Melastomataceae as indicators of vascular plant composition in rain forests of Colombian Amazonia. *Plant Ecology* 178: 1-13. (2005)
- [117] McCune, B. Influence of noisy environmental data on canonical correspondence analysis. *Ecology* 78: 2617-2623. (1997)
- [118] Hill, M., Witman, J., Caswell, H. Spatio-temporal variation in Markov chain models of subtidal community succession. *Ecology Letters* 5: 665-675. (2002)
- [119] Billheimer, D., Guttorp, P., Fagan, W. Statistical interpretation of species composition. *Journal of the American Statistical Association* 96: 1205-1214. (2001)
- [120] Diggle, P., Moyeed, R., Rowlingson, B., Thomson, M. Childhood malaria in the Gambia: a case-study in model-based geostatistics. *Journal of the Royal Statistical Society Series C-Applied Statistics* 51: 493-506. (2002)
- [121] McElhany, P., Real, L., Power, A. Vector Preference and Disease Dynamics - A Study of Barley Yellow Dwarf Virus. *Ecology* 76: 444-457. (1995)
- [122] Roepstorff, A., Nilsson, O., O'Callaghan, C., Oksanen, A., Gjerde, B., Richter, S., Ortenberg, E., Christensson, D., Nansen, P., Eriksen, L., Medley, G. Intestinal parasites in swine in the Nordic countries: multilevel modelling of *Ascaris suum* infections in relation to production factors. *Parasitology* 119: 521-534. (1999)
- [123] Patoucheas, D.P., Stamou, G. Non homogeneous Markovian models in ecological modelling: A study of zoobenthos dynamics in Thermaikos Gulf, Greece, *Ecol. Model.* 66, 197-215. (1993)

Obsah

Předmluva.....	2
1 Biodiverzita jako pojem	3
1.1 Definice biodiverzity	3
1.1.1 Biodiverzita z ekologického hlediska.....	3
1.1.2 Biodiverzita z matematického hlediska.....	4
1.2 Kvalitativní a kvantitativní komponenta biodiverzity	4
1.3 α , β , γ diverzita	6
1.4 Biodiverzita jako cílový parametr v hodnocení ekologických rizik	7
1.4.1 Biodiverzita jako integrující ukazatel stavu biologických společenstev	7
1.4.2 Produkce a velikost společenstev	8
1.4.3 Struktura a typ společenstev	8
1.4.4 Koncept chápání biodiverzity v ekologických studiích.....	9
1.4.5 Vývoj a stabilita společenstev	9
1.4.6 Pozice ukazatelů biologických společenstev v celkovém hodnocení vlivu stresoru.....	10
1.4.7 Současný vývoj metod v hodnocení biodiverzity.....	10
2 Biodiverzita a biostatistika	11
2.1 Analýza biodiverzity jako analogie ke klasickým statistickým postupům	11
2.2 Vzorkování biodiverzity	11
2.3 Ukládání dat biodiverzity	13
3 Vizualizace biodiverzity.....	15
3.1 Složení společenstva a jeho zobrazení z různých pohledů	15
3.2 Rozložení a transformace dat při analýze biodiverzity.....	17
4 Indexy diverzity a odhady jejich statistické spolehlivosti	18
4.1 Indexy diverzity jako analogie popisné statistiky.....	18
4.2 Indexy diverzity	18
4.2.1 Indexy založené na početnosti druhů	18
4.2.2 Indexy založené na poměru početnosti druhů	19
4.2.3 Q statistika.....	24
4.2.4 Vztahy mezi indexy biodiverzity.....	25
4.2.5 Velikost vzorku a indexy biodiverzity – rarefakce.....	25
4.3 Odhad intervalů spolehlivosti a statistické testování biodiverzitních indexů.....	28
4.4 Biotické indexy	30
5 Modely druhové abundance a stochastické modely	31
5.1 Modely druhové abundance jako analogie prokládání statistických rozložení.....	31
5.2 Stochastické modely	33
5.2.1 Geometrická řada [58].....	34
5.2.2 Logaritmičká řada (série)	34
5.2.3 Log-normální rozložení.....	35
5.2.4 Zlomená hůlka (broken stick).....	36
5.2.5 Zipfovy-Mandelbrotovy modely	37
5.3 Simulační, na niku orientované modely	38
5.3.1 Geometrická řada [58].....	39
5.3.2 Předpoklad dominance (dominance preemption) [21]	40
5.3.3 Náhodná frakcionace (random fraction) [21]	40
5.3.4 Sugiharův model postupného dělení (Sugihara sequential breakage model) [60]	41
5.3.5 Zlomená hůlka (broken stick) [59] alias MacArthurova frakcionace [21].....	41
5.3.6 Odmítnutí dominance (dominance decay) [21]	42
5.3.7 Náhodné roztržení (random assortment) [21]	42
5.3.8 Složený model (composite model) [21]	42
5.4 Hodnocení kompetice, agregace a šíře niky	42
5.4.1 Vnitrodruhová agregace	43
5.4.2 Mezidruhová agregace	43
5.4.3 Šíře a překryv niky	43
6 Možnosti frakcionace biologických společenstev a následná analýza biodiverzity získaných složek..	44
6.1 Rozdílný vliv prostředí na různé složky společenstva.....	44

6.2	Frakcionace společenstva – taxonomická, funkční aj.....	44
6.3	Identifikace složek společenstva s konzistentní reakcí na faktory prostředí	45
7	Aplikovatelnost parametrických a neparametrických statistických technik při hodnocení biodiverzity.....	46
8	Metody hodnocení β biodiverzity	47
8.1	β biodiverzita a vícerozměrná analýza biodiverzity.....	47
8.2	Problém dvojité nepřítomnosti (double-zero problem)	47
8.3	Koeficienty podobnosti pro data o biodiverzitě (β biodiverzita).....	49
8.3.1	Binární koeficienty podobnosti	49
8.3.2	Kvantitativní koeficienty podobnosti a vzdálenosti společenstev.....	51
8.4	Vícerozměrné analýzy s přímou vazbou na analýzu biodiverzity	53
8.4.1	Shluková analýza	53
8.4.2	Analýza hlavních koordinát – PCoA (klasické, metrické škálování MDS)	53
8.4.3	Korespondenční analýza – CA, Detrendovaná korespondenční analýza – DCA.....	54
8.4.4	Mnohorozměrné nemetrické škálování – NMDS.....	54
8.4.5	Kanonická analýza	55
8.4.6	Kanonická korespondenční analýza – CCA.....	56
9	Využití markovských řetězců v analýze biodiverzity	57
10	Software pro analýzu biodiverzity	58
10.1	Standardní komerční statistické softwary.....	58
10.2	PAST	58
10.3	Power niche 1.0	59
10.4	EstimateS.....	59
10.5	MVSP	59
10.6	PC-ORD	59
10.7	R a jeho knihovny specializované na analýzu biodiverzity.....	59
11	Případová studie: Parazitární společenstva	61
12	Případová studie: Lišejníky a znečištění ovzduší.....	65
	Seznam doporučené literatury	69
	Summary	76

Summary

The publication *Statistical Evaluation of Biodiversity* was funded as a part of the ESF project no. CZ.1.07/2.2.00/07.0318 entitled „MULTIDISCIPLINARY INNOVATION OF STUDY IN COMPUTATIONAL BIOLOGY“, which was investigated at the Faculty of Science, Masaryk University. This project aimed to improve study courses that form a core of the Computational Biology study programme at the Masaryk University.

Biodiversity is commonly used term especially in the context of ecology and environmental protection; nevertheless, its meaning is much broader and applicable to any level of living organisms from cell metabolic pathways, genetics and biological communities the whole ecosystems. From the mathematical point of view the methods for biodiversity analysis are derived from the information theory and they are also used in many other fields of science.

Therefore, our target readers are the students of Computational Biology, to whom we want to provide a comprehensive overview of the essentials of biodiversity data analysis in the context of real-life biological data. The ambition of this textbook is to show complete range of biodiversity term meaning and the methods for its objective statistical evaluation; not only commonly used methods but also more advanced and recent methods.

First chapter of this publication serves as an introduction to biodiversity analysis and for definition of terms. The second and third chapter shows links between biodiversity analysis and biostatistics and data visualisation. Chapter 4 is focused on diversity indices, chapter 5 on species abundance curves. Chapter 6 is aimed on stratification of biological communities data as a tool for finding contrasts in biodiversity analysis. Chapters 7 and 8 show applicability of standard biostatistical methods for data concerning biodiversity especially aimed on multivariate data analysis. Chapter 9 is devoted to Markov chain models in biodiversity, chapter 10 provides a list of available software for biodiversity data analysis. Chapters 11 and 12 are practical case studies of biodiversity data analysis.

The main goal is to provide students not with detailed theoretical description of methods but show their main principles and computations together with their weak points and provide guide for their interpretation. The textbook is thus supplementary to lectures and we hope it will also serve students as a reference text for their own data assessment in bachelor and master theses. We hope that availability of these new study materials will enhance the knowledge of students of computational biology and other study branches.

Statistické hodnocení biodiverzity

RNDr. Jiří Jarkovský, Ph.D., RNDr. Simona Littnerová, Doc. RNDr. Ladislav Dušek, Ph.D.

Recenzenti: doc. RNDr. Eva Bulánková, Ph.D., doc. RNDr. Milan Gelnar, CSc.

Jazykové korekce: Ing. Marie Juranová

Obálka: Radim Šustr, DiS.

Vydalo: AKADEMICKÉ NAKLADATELSTVÍ CERM, s.r.o. Brno,

Purkyňova 95a, 612 00 Brno

www.cerm.cz

Tisk: FINAL TISK s.r.o. Olomučany

Náklad: 200 ks

Vydání: první

Vyšlo v roce 2012

ISBN 978-80-7204-790-1