

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE FAKULTA ŽIVOTNÍHO PROSTŘEDÍ

Katedra aplikované geoinformatiky a územního plánování



**Vliv vstupních dat na identifikaci měřítka,
v němž druhy interagují se svým prostředím**

Diplomová práce

Vedoucí práce: Ing. Vítězslav Moudrý, Ph. D.

Autorka práce: Bc. Kristýna Pokorná

Praha 2017

ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. Kristýna Pokorná

Krajinné inženýrství

Název práce

Vliv vstupních dat na identifikaci měřítka, v němž druhy interagují se svým prostředím

Název anglicky

Species-environment relationships at multiple scales: The influence of input data

Cíle práce

Cílem práce je prozkoumat do jaké míry mohou vstupní data ovlivnit spolehlivost identifikace měřítka ('scale of effect'), v němž druhy interagují se svým prostředím.

Konkrétně je cílem zodpovědět následující otázky:

- Jak ovlivňuje identifikaci měřítka kvalita vstupních dat?
- Liší se výsledky pro různé typy využití území (les, vodní plochy)?
- Liší se výsledky v závislosti na měřítku analýzy (analysis scale)?

Metodika

Identifikace měřítka, v němž druhy interagují se svým prostředím, například s vegetačním pokryvem, je v ekologii běžně řešenou analýzou. Obvykle je druhová diverzita (či abundance) v určitém bodě srovnávána s environmentálními daty (např. plocha lesa) v okolí bodu v bufferech různých velikostí. Měřítka ('scale of effect') je identifikováno jako buffer, ve kterém druhová diverzita a environmentální proměnná nejvíce korelují. Úkolem autorky je simulovat virtuální druh a s využitím různých datových sad (např. CORINE Land Cover, DIBAVOD) posoudit míru ovlivnění analýzy vstupními daty.

Doporučený rozsah práce

40 – 50 stran textu

Klíčová slova

Formulace klíčových slov je úkolem řešitele.

Doporučené zdroje informací

- Crouzeilles, R., & Curran, M. (2016). Which landscape size best predicts the influence of forest cover on restoration success? A global meta-analysis on the scale of effect. *Journal of Applied Ecology*, 53, 440-448.
- Chandler, R., & Hepinstall-Cymerman, J. (2016). Estimating the spatial scales of landscape effects on abundance. *Landscape Ecology*, 1-12.
- Jackson, H. B., & Fahrig, L. (2015). Are ecologists conducting research at the optimal scale?. *Global Ecology and Biogeography*, 24, 52-63.
- Lecours, V., Devillers, R., Schneider, D. C., Lucieer, V. L., Brown, C. J., & Edinger, E. N. (2015). Spatial scale and geographic context in benthic habitat mapping: review and future directions. *Marine Ecology Progress Series*, 535, 259-284.
- Lechner, A. M., Langford, W. T., Jones, S. D., Bekessy, S. A., & Gordon, A. (2012). Investigating species-environment relationships at multiple scales: differentiating between intrinsic scale and the modifiable areal unit problem. *Ecological Complexity*, 11, 91-102.
- Tuanmu, M. N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, 23, 1031-1045.

Předběžný termín obhajoby

2016/17 LS – FŽP

Vedoucí práce

Ing. Vítězslav Moudrý, Ph.D.

Garantující pracoviště

Katedra aplikované geoinformatiky a územního plánování

Elektronicky schváleno dne 5. 4. 2017

doc. Ing. Petra Šimová, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 5. 4. 2017

prof. RNDr. Vladimír Bejček, CSc.

Děkan

V Praze dne 11. 04. 2017

Prohlášení

Prohlašuji, že jsem svou diplomovou práci vypracovala samostatně pod vedením Ing. Vítězslava Moudrého, Ph. D., a že jsem uvedla všechny literární zdroje, z nichž jsem čerpala, i veškeré zdroje dat, s nimiž jsem pracovala.

V Praze 18. 4. 2017

.....

Poděkování

Ráda bych poděkovala vedoucímu diplomové práce Ing. Vítězslavu Moudrému, Ph. D., za jeho zodpovědný přístup a odborné vedení, i za veškeré náměty a připomínky, které mi v průběhu zpracování této diplomové práce poskytoval. Děkuji rovněž za jeho motivaci a morální podporu.

Abstrakt

K identifikaci měřítka, ve kterém druhy interagují se svým prostředím, se v ekologických studiích nejčastěji využívají *multi-scale* analýzy. Obvykle se provádí porovnáváním vztahu mezi ekologickou odpovědí (př. výskyt, abundance, druhová diverzita) a environmentální proměnnou (př. vegetační pokryv) měřených v různých velikých buffrech v okolí zájmových lokalit. Charakteristické měřítko, tedy měřítko, ve kterém druh interaguje se svým prostředím, je identifikováno jako velikost buffru, kde je nejvyšší korelace mezi environmentální a ekologickou proměnnou. Tato práce zkoumá, jak mohou vstupní data ovlivnit spolehlivost identifikace charakteristického měřítka. Vliv několika datových sad z různých zdrojů (CORINE Land Cover, Dibavod, Forest MAP 2000, Global Inland Waters, Globeand30 a Treecover2010) byl testován prostřednictvím lineárně regresního modelu pro 6 oblastí (čtverců 45 km x 45 km) s odlišnou krajinnou konfigurací (strukturou). Simulovaná ekologická odpověď byla odvozena z prostorových dat v závislosti na podílu lesního pokryvu (data Forest MAP) a vodních ploch (data Dibavod). Výsledky jasně prokázaly vliv vstupních dat na identifikaci charakteristického měřítka, celkově došlo ke špatné identifikaci ve více jak 80% případů. Úspěšnost správné identifikace byla vyšší u lesního pokryvu, 28 případů z 90, zatímco u vodních ploch došlo k správné identifikaci jen ve 4 případech z 90. Nespolehlivá identifikace charakteristického měřítka přitom může negativně ovlivnit výsledky ekologických analýz i z nich plynoucí závěry (např. v plánování ochrany).

Klíčová slova: ekologická odpověď, charakteristické měřítko, měřítko, MAUP, prostorová nejistota, vztah druhu a prostředí

Abstract

To identify the scale at which the species interacts with its environment (scale of effect), ecological studies commonly utilise multi-scale analyses. It is usually carried out by comparing the relationship between ecological response (e. g. occurrence, abundance species diversity) and environmental variable (e. g. vegetation cover) measured in multiple buffer sizes surrounding the sites of interest. Scale of effect is identified as the buffer size at which the highest correlation between environmental and ecological variables is. This study examines how the input cover data can influence the reliability of identification of the scale of effect. The effect of several datasets from multiple sources (CORINE Land Cover, Dibavod, Forest MAP 2000, Global Inland Waters, Globeand30 and Treecover2010) was tested by using linear regression model, at 6 sub-regions (squares of 45 km x 45 km) with different landscape configuration (structure). Simulated ecological response was derived from spatial data according to tree cover fraction (Forest MAP) and water surfaces cover fraction (Dibavod). Results demonstrated the influence of input data on identification of scale of effect; there was misidentification in more than 80% of individual experiments. The success of correct identification was higher using forest cover data, in 28 individual experiments out of 80, while using water surfaces data, there was only 4 cases out of 80 of correct identification. Unreliable identification of scale of effect may negatively affects the results and conclusions of ecological analyses (e. g. in conservation management).

Key words: ecological response, scale of effect, scale, MAUP, spatial uncertainty, species-environment relationship

Obsah

1. Úvod	10
2. Literární rešerše.....	12
2.1. Měřítko (scale).....	12
2.1.1. Druhy měřítka	12
2.1.2. Prvky (komponenty) měřítka	13
2.1.3. MAUP	16
2.1.4. Řešení MAUP	18
2.1.5. Scaling.....	19
2.1.6. Scale of effect.....	20
2.1.7. Přístupy detekce scale of effect	20
2.2. Reprezentace dat v GIS – Data reálného světa.....	23
2.2.1. Nejistota dat.....	25
2.2.2. Data DPZ.....	25
2.3. Virtuální druh.....	27
2.4. Regresní model.....	28
2.4.1. Ordinary Least Squares (OLS).....	29
3. Metodika	30
3.1. Vstupní data	30
3.2. Popis experimentu.....	32
4. Výsledky	37
5. Diskuse.....	43
6. Závěr	48
7. Použitá literatura.....	49
8. Seznam příloh.....	55

Seznam zkratk:

CLC	Corine Land Cover
ETRS	Evropský terestrický referenční systém
GIS	geografické informační systémy
MAUP	Modifiable areal unit problem
MMU	minimum mapping unit
MODIS	Moderate Resolution Imaging Spectroradiometer
OLS	Ordinary Least Squares
VÚV TGM, v.v.i.	Výzkumný ústav vodohospodářský T. G. Masaryka, veřejná výzkumná instituce
ZABAGED	Základní báze geografických dat České republiky

1. Úvod

Záležitost měřítka (škálování, ‚*scaling*‘) je základem v ekologii, stejně jako v jiných disciplínách (Wiens, 1989). Hutchinson (1965), který se zabýval prostorovou a časovou variabilitou ekologických procesů, ve své práci *The Ecological Theater and the Evolutionary Play* (1965) uvedl, že pro porozumění událostem je třeba zobrazit je ve vhodném (odpovídajícím) měřítku. Prostorový rozsah, ve kterém druh nejlépe interaguje se svým prostředím (dává nejlepší ekologickou odpověď na krajinnou strukturu) se nazývá ‚*scale of effect*‘ nebo též ‚*characteristic scale*‘ či ‚*appropriate scale*‘ (Jackson et Fahrig, 2012). Obvykle však ‚*scale of effect*‘ není předem známé a jeho přesná identifikace je pro správný popis vztahu druhu a prostředí (‚*species-landscape*‘, ‚*species-environment relationship*‘) nezbytná (Wu et Li, 2006; Jackson et Fahrig, 2015).

Podle Jackson et Fahrig (2015) v řadě výzkumů využívající multiscale přístup není *scale of effect* optimálně identifikované. Měřítko jsou zkoumána buď v příliš úzkém rozsahu (jen v rámci 1 až 2 řádů) nebo je jich testováno málo. Jako nejlépe odpovídající bývá často vyhodnoceno nejmenší nebo naopak největší zkoumané měřítko, což značí, že lepší měřítko se může nacházet mimo zkoumaný rozsah (Jackson et Fahrig, 2015).

Zároveň je nutné podotknout, že hledání *scale of effect* nemusí být vždy úspěšné. Identifikace vhodného měřítka v jednotlivých studiích je omezena rozlišením vstupních dat. Např. ve studii závislosti abundance losa na heterogenitě krajiny, kde se snažili identifikovat vhodné měřítko jednotlivých parametrů ovlivňujících jeho výskyt, může být *scale of effect* jemnější než bylo zjištěno, což je důsledkem limitujícího faktoru kvality vstupních dat, konkrétně vzorkováním hustoty výskytu losa v rozlišení 15km² (Maier et al, 2005).

Běžným přístupem ke zkoumání ekologické odpovědi druhu v různých měřítkách a identifikaci *scale of effect* je hierarchická výstavba víceúrovňových buffrů okolo vzorkovacích míst (Brennan et al., 2002; Boscolo end Metzger, 2009; Lechner et al., 2012a). Nakolik je identifikace *scale of effect* ovlivněna vstupními daty dosud zkoumalo jen několik studií (např. Lechner et al., 2012b). Navíc se studie zaměřily pouze na rozlišení vstupních dat. Nicméně, datové sady se v závislosti na způsobu sběru mohou lišit také ve schopnosti identifikovat jednotlivé krajinné struktury.

Cílem mé práce je zjistit, jak mohou vstupní data ovlivnit spolehlivost identifikace *scale of effect*. Za pomoci simulace virtuálního druhu a jeho interakce s prostředím byla zkoumána míra ovlivnění analýzy různými volně dostupnými datovými sadami

(např. CORINE Land Cover, Dřevost, FOREST Map 2000). Konkrétně ve své práci hledám odpověď na otázky:

- jak ovlivňuje identifikaci měřítka kvalita vstupních dat,
- zda se liší výsledky pro různé typy využití území (les, vodní plochy),
- zda se liší výsledky v závislosti na měřítku analýzy (*analysis scale*).

2. Literární rešerše

2.1. Měřítko (*scale*)

Tento termín je široce používán v mnoha oblastech a jako takový má mnoho někdy protichůdných významů. Například měřítko mapy je kartografický poměr odkazující na vztah mezi vzdáleností nebo plochou zastoupenou v mapě a odpovídající reálnou vzdáleností nebo plochou (Dungan et al., 2002). V krajinné ekologii zas měřítko definuje velikost zrna (*grain*) a rozsah (*extent*; Turner, 1989). Příbuzný termín rozlišení, běžně užívaný v dálkovém průzkumu, je definován jako nejmenší objekt, jenž může být spolehlivě detekován (Dungan et al., 2002).

Zohlednění měřítka je v ekologických studiích zásadní otázkou. Přesto, zatímco v jiných disciplínách už se touto problematikou dávno zabývali, ekologové byli v poznání měřítka pozadu. V oborech jako je geografie, atmosférické procesy a věda o Zemi, fyzické procesy jasně vykazovaly spjatost s lokálními a globálními vzory – tedy odehrávaly se v nějakém měřítku. Naproti tomu ekologové řešili jevy, které brali jako běžné, intuitivně, proto měli tendenci vnímat a studovat tyto jevy v antropocentrickém měřítku, tedy měřítku, jež odpovídá naší zkušenosti. Bývali spíše tradiční a často ve svých výzkumech užívali prvky, které byly použity již v předešlých studiích (Wiens, 1989).

V současné době je určování měřítka ekologických procesů a měřítka závislosti ekologických odpovědí zásadní otázkou krajinné ekologie (Turner, 1989; Wu et Li, 2006).

2.1.1. Druhy měřítka

Lechner et al. (2012b) rozlišují tři hlavní druhy prostorového měřítka: *observation scale*, *analysis scale*, *intrinsic scale* (Obr. č. 1A).

Observation scale (měřítko pozorování), v literatuře také jako *observational scale*, *sampling scale* nebo *measurement scale* (Wu et Li, 2006), popisuje velikost, tvar, rozsah a vzdálenost mezi pozorovanými jednotkami používanými ke vzorkování jevu (Lechner et al., 2012b). *Observation scale* tedy souvisí s jednotkami vzorkování (*sampling units*) a pojmy jako velikost atd. jsou jejich důležitou charakteristikou. Některá ekologická data jsou sbírána na základě přirozených vzorkovacích jednotek, např. pro každý organismus v definované populaci nebo pro část organismu (tím může být třeba list), nebo jednoduše zaznamenaná pozorování na

hnízdo nebo noru. Přirozené jednotky vzorkování můžeme nalézt přímo uvnitř hierarchie, př. hmyz uvnitř semen, zrna v klasu, květy na rostlině, rostliny na louce (Norowi et al. 2000). Ve většině případů však tyto přirozené jednotky nejsou k dispozici a je nutné parametry vzorkovacích jednotek zvolit. Tato volba je často dána použitým nástrojem měření a určitým logistickým omezením na provedení měření. V případě, že nemáme přirozené vzorkovací jednotky, je třeba zvážit 4 hlavní parametry jednotek: 1) velikost (jednotka může být tak malá, že pojme pouze jedince či zlomek, nebo tak velká že vždy pojme hodně), 2) tvar (nejběžnější je čtvercový nebo kruhový tvar, ale může být i nepravidelný), 3) vzájemná vzdálenost (jednotky mohou být přilehlé (přímo sousedící) nebo v náhodném či předem určeném intervalu) a 4) rozsah vzorku (např. vymezená oblast zájmu; Dungan et al., 2002).

Analysis scale (měřítko analýzy) se vztahuje k jednotkám užívaným při analýze dat a obvykle je hrubší než *observation scale* (Lechner et al., 2012a). Při provádění experimentu, prostorová a časová dimenze zkoumaného systému je reprezentována experimentálním měřítkem (*experimental scale*), které je hlavním kritériem při rozhodování mezi micro-, meso- a macro-scale experimentem (Dungan et al., 2002).

Intrinsic scale (skutečné měřítko) je měřítko, ve kterém zkoumaný ekologický jev interaguje nebo vnímá okolní prostředí. Odvíjí se od vztahu organismu a prostředí a měří se nepřímo prostřednictvím násobku *observation scale* a/nebo *analysis scale* (Lechner et al., 2012a). Dalo by se říct, že *intrinsic scale* je pouze pojem zavedený pozorovatelem, jako následek jeho poznání přirozeného měřítka jevu (Allen et Starr, 1982).

Kromě již zmíněných, uvedli Wu et Li (2006) v souvislosti s managementem (řízením) a ochranou životního prostředí další druh měřítka, *policy scale*, které je ovlivněné ekonomickými, politickými a sociálními faktory.

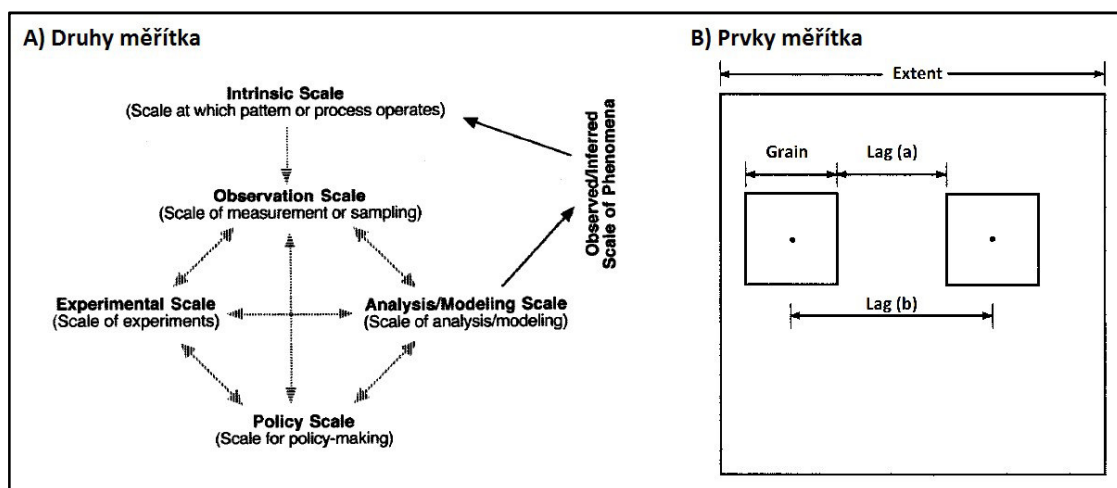
2.1.2. Prvky (komponenty) měřítka

Ve spojitosti s měřítkem se používá několik termínů charakterizujících měřítko (Obr. č. 1B), jako je *rozsah (extent)*, *zrno (grain)*, *rozlišení (resolution)*, *vzorkovací interval (lag)*, *podpora dat (support)*. Existuje spousta definic pro každý z termínů, některé definice se vzájemně překrývají, některé nejsou zcela jednoznačné (Dungan et al., 2002).

V geografických informačních systémech (GIS) mají největší význam rozlišení a rozsah. *Rozsah* udává především prostorový rozsah, ale může být definován i v jiných dimenzích, nejčastěji času. Stejně tak rozlišení, neboli stupeň detailu, se primárně týká prostoru a času (Goodchild, 2011). *Rozsah* je celková délka, plocha nebo objem, který existuje nebo je pozorován či analyzován. Dá se použít v případě jevu, pozorování i analýzy. Obdobně se používá i *zrno* (*grain*). *Grain* bývá definováno ve spojitosti s jevem, jako vlastnost struktury mozaiky nebo vlastnost pozorování (Dungan et al., 2002). Je to nejmenší jednotka jevu nebo datové sady, v níž se předpokládá homogenita (Wiens et al., 1989). Nejjednodušší je asi představit si *zrno* jako nějakou oblast pro sběr dat (čtvercový kvadrát nebo dlaždice s definovanou délkou strany). Za *grain* může být brána velikost pixelu v případě rastrových dat, nebo minimum mapping unit (MMU, nejmenší měřitelná jednotka), jak v případě vektorových tak rastrových dat. Zrno analýzy se může od zrna vzorkovací jednotky lišit, třeba v případě, jsou-li vzorkovací jednotky agregované do větších celků. Vzhledem k nedostatečné shodě ve vymezení tohoto pojmu, je dobré, v případě jeho použití, jednoznačně definovat jeho význam (Dungan et al., 2002).

Obecně platí, že velikost zrna by měla být menší než prostorový a časový rozměr zkoumané struktury (zkoumaného vzoru), zatímco extent pro vzorkování (*sampling*) by měl být alespoň tak velký, jako rozsah jevu zkoumaného v rámci studie (Dungan et al., 2002).

Dalším komponentem měřítka je *support*, často používaný geostatický termín. Vztahuje se k nejmenší ploše nebo objemu, pro který je odvozena průměrná hodnota proměnné (Dungan et al., 2002). Protože se jedná o vlastnost proměnné použité k modelování dat, patří *support* k analytické sekci měřítka (Rossi et al., 1992). V mnoha případech mají *grain* a *support* podobný význam a mohou být vzájemně zaměnitelné (Wu et Li, 2006). Nicméně *support* na rozdíl od zrna nezahrnuje pouze velikost n-rozměrného objemu, ale i jeho geometrický tvar a orientaci (Olea, 1990). Prostorová jednotka k vzorkování prostředí byla rovněž považována za *support* dat (Rossi et al., 1992). Intuitivně by člověk řekl, že proměnná je definovaná čistě samotným atributem (abundance, presence/absence, zelenost vegetace atp.), ale rozhodující jsou i prostorové aspekty vzorkovací jednotky. Bráno statisticky, stejný atribut měřený za použití dvou odlišných vzorkovacích jednotek dává dvě odlišné proměnné, neboť mají odlišná rozdělení (rozdílnou distribuci, Dungan et al., 2002).



Obr. č. 1 (A) Druhy měřítka (Wu et Li, 2006 převzato a upraveno z Bierkens et al., 2000), **(B) Prvky měřítka.** Jedná-li se o vzor pravidelné velikosti a tvaru, pak $Lag(a) = Lag(b) - Grain$ (převzato a upraveno z Dungan et al., 2002).

Vzorkovací interval (*lag*) označuje vzdálenost nebo interval mezi dvěma sousedními jednotkami, ať už jednotkami jevu, vzorkování nebo analýzy. V případě, že není interval přesně definován, existuje více možných interpretací. Vzdálenost (interval) mezi jednotkami může být určena např. jako vzdálenost středů jednotek nebo jako vzdálenost mezi nejbližšími okraji těchto jednotek (Dungan et al., 2002).

Bierkens et al. (2000) uvádějí navíc pojem *coverage* (pokrytí) vyjadřující intenzitu vzorkování. Prostorové pokrytí může být reprezentováno jako poměr plochy vzorkování k celkovému rozsahu (Wu et Li, 2006).

Měřítka ekologického jevu lze zkoumat jen volbou měřítek pozorování a analýzy. Proto tato měřítka mohou ovlivnit ekologické závěry. Když změním měřítka pozorování nebo analýzy změnou velikosti, tvaru nebo intervalu vzorkovací jednotky, případně změnou rozsahu, na kterém je vzorkování prováděno, lze očekávat jiné výsledky prostorových statistických analýz. Ve výsledku se změní např. odhady populačního průměru a odchylky, síla a charakter prostorové autokorelace a prostorové anisotropie, velikost plošek (*patch*) a mezer, i vztahy závislé na více proměnných (Dungan et al., 2002). Každá proměnná má své rozdělení nesoucí určité vlastnosti, ty závisí na ploše nebo objemu, nad kterým se měří nebo počítá. Dungan et al. (2002) představili čtyři studie závislosti abundance druhu na krajinném pokryvu ukazující, jak některé změny prostorových charakteristik mohou ovlivnit závěry o ekologických jevech (procesech a strukturách). Přičemž, stejně jako platí působení změn prostorových charakteristik vzorkovacích jednotek jedné proměnné, platí i že se změnou vzorkovací jednotky

dvou nebo více proměnných se změny i kovariance a korelační statistiky (Dungan et al., 2002).

2.1.3. MAUP

Současné vzrůstající využívání dálkového průzkumu a GIS vede k nárůstu prostorových analýz ve velkých měřítkách. Jevy jsou modelovány za pomoci nepřekrývajících se prostorových jednotek, které mohou mít různý tvar a velikost. Zkoumanou oblast lze dělit mnoha různými způsoby (často libovolně). Zvolené měřítko i metoda tvorby těchto prostorových jednotek vede ke komplikacím při statistických analýzách (Openshaw and Taylor, 1979). Např. v biogeografii, většina naší znalosti o jevech velkého měřítka je odvozena ze souhrnných dat dané oblasti, v rámci níž byla data získána z malých ploch (méně než 1 km²), nebo z dat sebraných v konkrétním předem definovaném měřítku, jako jsou třeba snímky dálkového průzkumu (Burke et al., 1991).

Modifiable areal unit problem neboli MAUP představuje citlivost analytických výsledků na definování prostorových jednotek sběru dat a může se vyskytnout v jakémkoliv aspektu prostorově explicitních dat velkého měřítka (Openshaw and Taylor, 1979; Openshaw, 1984b). Změna tvaru a/nebo velikosti prostorových jednotek při analýze dat, může změnit výsledné korelace nebo statistické modely generované z těchto dat (Jelinski et Wu, 1996).

MAUP lze rozlišit na dva základní problémy:

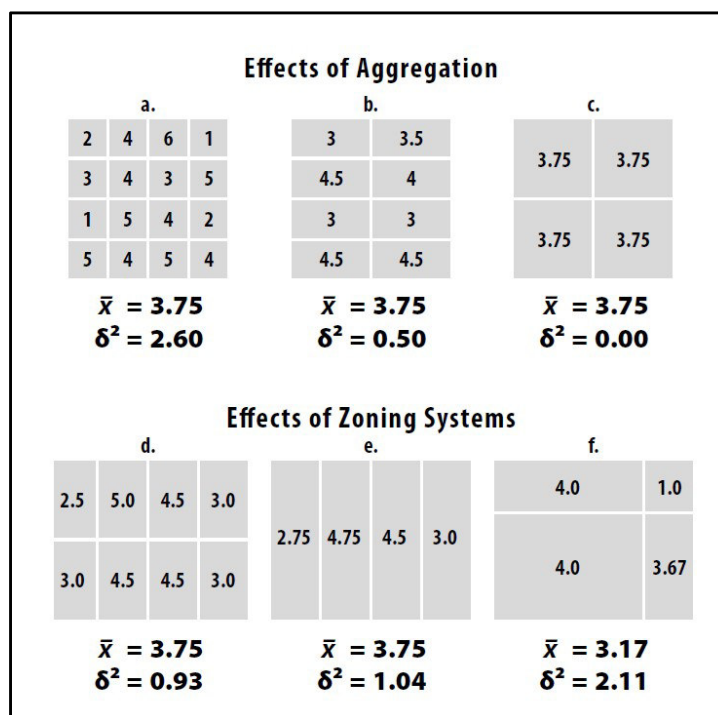
1. Efekt měřítka (scale effect) na změny numerických výsledků se přisuzuje čistě počtu plošných/oblastních jednotek použitých v analýze dané oblasti.
2. Zonační efekt (zonation/zoning effect, nebo též aggregation effect) je přičítán změnám v numerických výsledcích díky procesu, v němž je větší počet menších jednotek uskupen (agregován) do menšího počtu větších prostorových jednotek (Openshaw and Taylor, 1979).

První problém se týká měřítka. Změny ve výsledcích nastanou, když jsou prostorové jednotky progresivně agregovány do menšího počtu větších jednotek pro analýzy. Hodnoty spojené s proměnlivostí dat se sníží, což ovlivní všechny související statistické analýzy. Druhý problém reprezentuje změnu ve výsledcích způsobenou použitím alternativních agregačních schémat ve stejném nebo podobném měřítku (Openshaw, 1984b).

MAUP při získávání dat (při zpracování v GIS) nastává při designu projektu a sběru dat spojeným s terénním mapováním např. vegetačního pokryvu. V případech, kdy jsou vrstvy krajinného pokryvu vytvářeny na základě dat získaných v terénu, je nezbytné zjednodušit a seskupit vegetační typy, jak z časových důvodů, tak kvůli omezeným zdrojům pro sběr dat (Dark et Bram, 2007).

MAUP je rovněž spojen s pojmem ekologická chyba/nepřesnost (*ecological fallacy*). Ta vzniká, když jsou výsledky pocházející ze souhrnných (agregovaných) zonálních dat aplikovány na jednotlivce nebo na konkrétní místa v rámci samotné zóny. Tento problém se objeví, když studie využívající plošné údaje nerozlišuje mezi prostorovými spojitostmi vytvořenými agregací dat a skutečnými spojitostmi, které existují u individuálních dat ještě před prostorovou agregací (Openshaw, 1984b). Jakékoliv statistiky nebo modely, založené na prostorových datových sadách vycházejících z agregovaných dat, mohou být platné v měřítku datové sady, ale jakékoliv vyvození do vyššího nebo nižšího rozlišení dat může vést k chybným výsledkům. Statistické údaje a parametry modelu use liší podle úrovně rozlišení a není možné předpovědět, čemu se hodnoty z menšího rozlišení rovnají ve vyšším rozlišení a naopak (Dark et Bram, 2007).

Studie MAUP probíhají už od 30. let, s největším důrazem koncem 60. a dál v 70. letech. Výsledky studií jsou hodně proměnlivé a poněkud neúplné, proto je těžké udělat ucelený závěr, jak přesně MAUP statistiky ovlivňuje. Nicméně, za tu dobu už byly odhaleny nějaké obecné vzory a pravidla (Dark et Bram, 2007). V univariátních statistických analýzách zatížených MAUP se průměr nemění a odchylka se snižuje s rostoucí agregací (Openshaw, 1984b; Fotheringham and Wong, 1991). V podstatě dojde ke ztrátě informace ve spojitosti s vyhlazovacím efektem (*smoothing effect*) během agregace (Jelinski et Wu, 1996). Zonační efekt má pro průměr a odchylku méně předvídatelné výsledky. Jelinski et Wu (1996) publikovali článek zabývající se MAUP a jeho důsledky pro krajinou ekologii, kde jednoduše demonstrovali účinky měřítka a zonačního efektu (obr č. 2).



Obr. č. 2 Vliv měřítka a zonace. Při agregaci (a-c) se výsledný průměr nemění, odchylka se snižuje. Při seskupení jednotek do zón (d-f) s různou orientací ke světovým stranám, pro d a e nenastane žádná změna v hodnotě průměru, ale změní se odchylka. Při porovnání e a f vidíme, že ačkoliv je počet zón zachován, výsledné hodnoty průměru i odchylky se liší (Jelinski et Wu, 1996).

Uznání faktu, že ekologické procesy a fyzikální charakteristiky v přírodě existují v přirozeném měřítku, bylo jedním z hlavních přínosů pro přírodní vědy. Měřítko studie určuje rozsah vzorů a procesů, jež mohou být během studie detekovány, proto je potřeba určit vhodnou úroveň rozlišení pro zkoumání těchto procesů. Navíc, za účelem propojení a porovnání informací napříč širokým spektrem měřítek, byla zdůrazněna i potřeba příslušných pravidel pro scaling (Dark et Bram, 2007).

2.1.4. Řešení MAUP

Současná znalost problematiky už nabízí několik možností, jak se s MAUP vypořádat. První možností je práce se základními subjekty (*basic entities*) namísto dodatečně stanovených prostorových jednotek. Pojem základní entita se vztahuje k objektu složenému z podobných částí, jež jsou odlišné od samotného objektu. Vezmeme-li např. korunu stromu jako základní entitu, koncepčně je to souhrn listů a větví, z nichž každá jednotlivě náleží do tříd, které mohou být samy o sobě základní entitou. Proto identifikace základních subjektů nabízí nejjasnější východisko z MAUP (Fotheringham, 1989).

Další možností, jak se MAUP vyhnout, je pečlivá výstavba (design) vzorkovacího plánu a analýzy (Dungan et al., 2002). Benz et al. (2004) navrhuje několik strategií pro získávání a zpracování dat, které by lépe propojily teoretickou informaci obsaženou v rastrovém formátu dat dálkového průzkumu s výsledky digitální fotogrametrie a případnou kombinací s vektorovými daty v prostředí GIS. Jsou jimi např. extenze pro zpracování signálu při analýze obrazu pro zachycení významně propojených objektů reálného světa na základě zkoumání hierarchie sítě objektů obrazu, nebo třeba *fuzzy systems* (Townsend et al., 2000; Benz et al. 2004) pro lepší a robustnější modelování závislostí reálného světa spolu s detailní kontrolou kvality výsledného produktu.

Možností, jak přistupovat k MAUP je vícero, většina se však odvíjí od základní myšlenky studování obrazu v různých úrovních měřítka – tzv. *„multiscale approach“*. Při pohledu z blízké vzdálenosti od obrazu máme větší rozlišovací schopnost, můžeme rozpoznat jednotlivé objekty (domy, budovy, cesty, stromy atp.), při zvětšení pohledové vzdálenosti už nerozlišíme jednotlivé objekty, ale spíše různá seskupení, sousedící plochy (sídelní typy atp.). Tyto plochy mohou být význačné (odlišitelné) různou texturou, velikostí a tvarem. Textura ploch je tvořena jejími objekty a strukturami na jemnějším měřítku (domy, cesty, zahrady) a je dána hlavně spektrálními hodnotami, tvarem a také prostorovými vztahy. Z ještě větší vzdálenosti už se jako samostatné objekty budou jevit např. město a okolní zemědělské plochy a/nebo lesy. Tento příklad popisuje přístup 3-úrovňového měřítka (3-scale-level approach), kde rozpoznáme: 1) stromy, budovy a cesty v jemném měřítku; 2) skupiny stromů a skupiny budov tvořící sídelní typy ve středně velkém měřítku; 3) lesy, sídla a otevřenou krajinu v hrubém měřítku. Mezi jednotlivými měřítky vždy existuje hierarchická závislost (Benz et al., 2004).

Pro úspěšnou analýzu obrazu je nezbytná reprezentace jeho obsahu v několika měřítkách zároveň a důkladné prozkoumání hierarchické měřítkové závislosti mezi výslednými objekty (Benz et al., 2004).

2.1.5. Scaling

Obecnou a široce uznávanou definicí termínu *scaling*, hlavně v ekologii a ve vědách o zemi, je převod dat nebo informací mezi/napříč prostorovými a časovými měřítky (Turner et al., 1989a). *Scaling* je kvůli nelinearitě mezi procesy a proměnnými a kvůli heterogenitě vlastností určujících míru dějů poměrně náročný. V praxi může být proveden tzv. *„bottom-up“* nebo *„top-down“*: *upscaling* se skládá z použití informace

z malého měřítka k odvození informace ve větším měřítku, zatímco *downscaling* představuje rozklad informace na složky/součásti z nějakého měřítka do menšího měřítka (Jarvis, 1995).

Pro scaling mohou být použity různé statistické a matematické nástroje, jako např. korelace a extrapolace. Nicméně vzhledem k MAUP, jsou tyto techniky vhodné pouze pro krátkodobé predikce nebo predikce v malém měřítku, případně v rozsahu měřítka, které je pro zkoumaný jev relevantní. Rozšíření za tyto hranice může být nebezpečné vzhledem k nestabilitě dynamiky přechodových zón mezi dvěma oblastmi měřítka (Wiens, 1989). Konkrétně extrapolace měření z konkrétních míst do širšího měřítka může mít velké negativní důsledky (King, 1990), například v krajině ekologii při studování krajiných vzorů/procesů ve více měřítkách (*multiscale* studie; Wu, 2004; Lausch et al., 2013; Schindler et al., 2013).

2.1.6. Scale of effect

Je prokázáno, že většina ekologických vzorů a procesů je závislá na měřítku (Wiens et al., 1989;...) a že existují rozličná charakteristická měřítka, která se je snaží co nejlépe vystihnout (Wu, 1999). Zatímco charakteristická měřítka poskytují koncepční základ a praktické rady pro scaling, kvantitativní popis účinků měřítka může sloužit přímo pro *scaling* vztahů (Wu, 2004).

Základní otázkou při hledání skutečného, charakteristického měřítka - *scale of effect*- je, jaký prostor je ovlivněn daným procesem i opačně jak velký prostor ovlivňuje zkoumaný jev. Scale of effect může být správně určeno, pouze jsou-li správně zvolena měřítka pozorování, analýzy a modelování. Přímý vliv na odhalení scale of effect ve finální analýze má samotná volba vzorkovacích jednotek (*sampling*; Dungan et al., 2002; Wu et Li, 2006). Stejně tak přesnost, se kterou je měřena krajinná proměnná ovlivní přesnost odhadu scale of effect (Miguet et al., 2015).

Může se stát, že se jednotlivé druhy měřítka pro ten samý ekologický jev liší. K řešení tohoto problému nesouladu měřítek nebo ke vztažení jednoho typu měřítka k jinému slouží převod měřítka (*scale transfer*) nebo *scaling* (Bierkens et al., 2000).

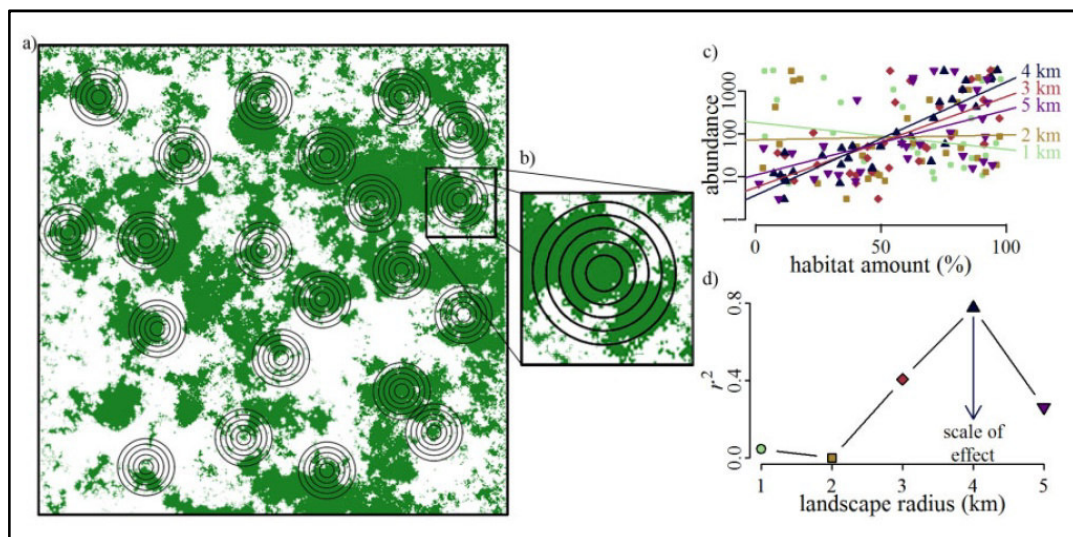
2.1.7. Přístupy detekce scale of effect

Roku 1989 uvedl Wiens jako běžný způsob, jak se vypořádat s měřítkem, porovnání vzorů mezi několika libovolně zvolenými body měřítkové škály.

Vzhledem k tomu, že scale of effect není předem známé, existuje k jeho zjištění řada metod. Mezi ně patří např. ‚spectral analysis‘, ‚fractal analysis‘, ‚wavelet analysis‘, ‚scale variance‘, ‚geostatistics‘, ‚multipleobject-specific analysis‘ (př. Turner et al., 1991; Wu et al., 2000; Hay et al., 2001; Dale et al., 2002; Hall et al., 2004). Obvykle je nezbytné mít množství datových sad v různých měřítkách (Wu, 1999). Účinné odhalení měřítka pak vyžaduje, aby měřítko analýzy odpovídalo scale of effect daného jevu (Dungan et al., 2002; Legendre et al., 2002).

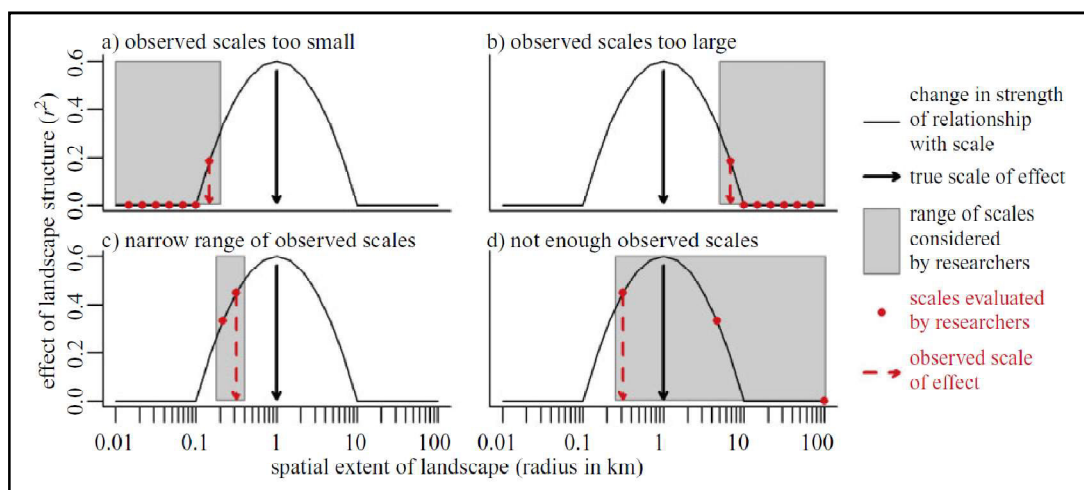
Asi nejjednodušší způsob identifikace scale of effect je pomocí *multiscale study design* (obr. č. 3; Brennan et al., 2002; Jackson et Fahrig, 2015), kde se hodnotí síla vztahu mezi krajinnou strukturou a ekologickou odpovědí. Spočívá v měření ekologické odpovědi druhu (př. abundance) v několika zájmových bodech a měření krajinné struktury (př. podíl habitatu, zvoleného krajinného pokryvu) v oblastech různého prostorového rozsahu v okolí těchto zájmových bodů (Jackson et Fahrig, 2015). Poté je vystavěn pro každý prostorový rozsah statistický model vztahující krajinnou proměnnou k biologické odpovědi (obvykle používá lineární regresi; Miguet et al., 2015). Scale of effect je pak rozsah, ve kterém má krajinná struktura při predikci ekologické odpovědi druhu nejvyšší korelační koeficient, případně nejstrmější regresní křivku nebo nejnížší AIC (*Akaike information criterion*) nebo BIC (*Bayesian information criterion*; Jackson et Fahrig, 2015). Tato metoda nebere v potaz směr vztahu mezi krajinnou proměnnou a biologickou odpovědí, například neuvažuje, jestli je kladný nebo záporný, protože to nijak neovlivňuje měřítko, v němž je vztah nejsilnější (Miguet et al., 2015)

Měření krajinných proměnných se provádí většinou v oblastech ve tvaru kruhu, přestože druh může interagovat se svým okolím i na území jiného tvaru než je kruh. Vědci ale nemají předem dostatek informací o daném vztahu, aby mohli predikovat jiný složitější tvar (Miguet et al., 2015)



Obr. č. 3 Design hypotetické multiscale studie s vybranými místními body. (a) Ekologická odpověď (př. abundance) v zájmových místech rozprostřených v oblasti zájmu. (b) Krajinná struktura je měřena v několika bufferech různého rozsahu koncentrických se středem v zájmovém bodě. V tomto případě je měřeno procentuální zastoupení habitatu (zeleně). (c) Vztah mezi ekologickou odpovědí a krajinnou strukturou je hodnocen pro každý prostorový rozsah - každý buffer. (d) Scale of effect je prostorový rozsah, kde krajinná struktura nejlépe predikuje odpověď druhu. V tomto případě je scale of effect 4 km, kde je sklon křivky mezi abundancí a krajinnou strukturou nejstrmější (graf c) a kde je r^2 nejvyšší (graf d). Jakmile jsou prostorová měřítka (buffery) umístěna, krajinné metriky sousedících měřítek jsou korelované. Proto pokud krajinná proměnná ovlivňuje ekologickou odpověď, síla vlivu (graf r^2) roste až k dosažení scale of effect (4 km) a poté se opět snižuje. Nutno poznamenat, že pokud by podíl habitu byl měřen pouze v 1 km, byl by nesprávně stanoven malý negativní vztah mezi abundancí a mírou (Jackson et Fahrig, 2015).

V mnoha multiscale studiích dochází ke špatnému určení skutečného scale of effect. Nejčastěji proto, že evaluovaná měřítka jsou: a) příliš malá, b) příliš velká, c) příliš malého rozsahu, a/nebo d) příliš malého počtu (obr. č. 4; Jackson et Fahrig, 2015). Brennan et al. (2002) doporučují zakládat výběr evaluovaných měřítek pro zjištění scale of effect na typických rysech zkoumaného druhu. Ideálně by studie měla vycházet z očekávaného měřítka využití prostoru a zkoumat krajinné proměnné v měřítkách v rozsahu hodnot od hodnoty menší než je domovský okrsek druhu, po hodnotu 4-9 krát větší než je průměrná disperzní vzdálenost (Jackson et Fahrig, 2012). Obvykle však měřítka nejsou vybírána explicitně na základě druhových rysů, ale jejich volba je nebiologického charakteru. Většina studií tak dosud nezahrnuje relevantní měřítka využití prostoru, přitom zahrnutí většího rozsahu nebo větší hustoty (většího počtu) evaluovaných měřítek zlepšuje výsledky studie (Jackson et Fahrig, 2015).



Obr. č. 4 Možnosti, kdy je zjištěné scale of effect odlišné od skutečného scale of effect. Jsou-li evaluovaná krajinná měřítka (a) příliš malá, (b) příliš velká, (c) v příliš úzkém rozsahu, nebo (d) je evaluovaných měřítek malý počet, může se zjištěné scale of effect (přerušovaná červená šipka) lišit od skutečného scale of effect (plná černá šipka; Jackson et Fahrig, 2015).

Teoretické studie zastávají názor, že některé druhové rysy skutečně ovlivňují scale of effect (Jackson et Fahrig, 2012; Miguet et al., 2015). Ricci et al. (2013) potvrdili ve své simulaci očekávaný pozitivní vztah mezi disperzní vzdáleností a scale of effect. U těch druhů, u kterých byla do rozsahu evaluovaných měřítek zahrnuta měřítka využití prostoru daným druhem, bylo zjištěné scale of effect výrazně větší než velikost domovského okrsku nebo velikost teritoria, ale nijak významně se nelišilo od průměrné natální disperzní vzdálenosti nebo od očekávaných hodnot (Jackson et Fahrig, 2012). Předpokládá se, že velikost těla a trofická úroveň (herbivor < karnivor) mohou rovněž pozitivně souviset se scale of effect, z důvodu jejich korelace s velikostí teritoria a disperzní vzdáleností (Jenkins et al., 2007; Hendriks et al., 2009). Několik empirických studií ukázalo pozitivní vztah mezi velikostí těla a scale of effect (Holland et al., 2005; Thornton et Fletcher, 2014), ale výsledky studie Jackson et Fahrig (2015) přinesly jen slabé důkazy o tomto vztahu. Empirická opora pro spojitosti mezi druhovými rysy a scale of effect není jednotná a je do značné míry neprůkazná (Jackson et Fahrig, 2015).

2.2. Reprezentace dat v GIS – Data reálného světa

Reálný geografický svět je komplexním souborem různých jevů a pravidla integrovaná do našeho systému a použitá pro geografickou reprezentaci jevů je potřeba zjednodušit. Obecná teorie poskytující jednodušší soubor/sadu stavebních

bloků pro geografickou reprezentaci může poskytnout lepší podporu pro vědecký výzkum a management povrchu Země, včetně jeho popisu, reprezentace, analýzy, vizualizace a simulace (Goodchild et al., 2007).

Existují dva obecně uznávané způsoby jak pohlížet na jevy rozprostírající se na zemském povrchu. Koncept diskrétních (jednotlivých, oddělených) objektů (*discrete objects*) bere povrch Země jako desku stolu - vyjma míst, kde se nacházejí diskrétní, počítatelné objekty, je povrch prázdný. Tyto objekty se mohou překrývat. V mnoha případech si zachovávají svou integritu (celistvost), jak v čase, tak při přemístění. Tento koncept je vhodný hlavně pro objekty, jako jsou biologické organismy, vozidla nebo budovy. V praxi jsou diskrétní objekty reprezentovány jako body, linie, oblasti (plochy, území) nebo objemy a to v závislosti na jejich velikosti a smyslu/účelu reprezentace (Goodchild, 2011).

Jiné jevy, které nemají tak jasně definované okraje, jimiž jsou hlavně přírodní krajinné prvky jako hory, jezera a řeky, je lepší brát jako kontinuální pole (*continuous-fields*). V tomto konceptu jde o spojení místa s konkrétní hodnotou nebo třídou, takže každé místo v časoprostoru má právě jednu hodnotu každé proměnné. Kompletní časoprostor o čtyřech dimenzích bývá často zjednodušován, ignorováním času, v případě statických jevů, nebo neuvažováním třetí prostorové dimenze, případně obojím. Tento koncept se hodí například pro topografii (místopis), kde je mapován celý povrch, každé místo (x, y) má svou hodnotu z , stejně tak se hodí pro mapování půdních typů, kde každé místo (x, y) je spojeno s nějakou třídou. Hodně fyzikálních jevů (např. teplota vzduchu, vlhkost půdy, krajinný pokryv) je pojata spíše jako *continuous-fields* než jako soubor diskrétních objektů (Goodchild, 2011).

Oba koncepty mohou být reprezentovány rastrovými nebo vektorovými daty, a ačkoliv jsou v zásadě nezávislé na měřítku, v praxi jejich digitální reprezentace do jisté míry měřítko vyjadřuje. V případě rastru je rozlišení dáno explicitně velikostí buněk rastru. V případě vektoru je to složitější. Data jsou zachycena nepravidelně rozmístěnými body, a i když existuje určité opodstatnění pro využití vzdálenosti mezi body jako míry rozlišení, není nikde přesně řečeno, zda by měla být brána minimální, průměrná nebo maximální hodnota této vzdálenosti. Data představující atributy oblastí jsou v geometrické formě obvykle reprezentovány jako polygony, míra rozlišení pak vyvstává v několika možných formách: v ochotě zastoupení hranic jako nekonečně tenkých, v hustotě vzorkování hranic, v kolísání uvnitř oblasti, které je nahrazeno předpokladem homogenity, ve velikosti oblastí. To pak vytváří mylný dojem, že vektorová data mají nekonečně jemné rozlišení. Podobně je

tomu i u linií, které jsou v digitální formě reprezentovány jako *polyline*. Představují sadu bodů spojených přímými spojnicemi, mýlka o dokonalém rozlišení plyne z hustoty vzorkování linií a jejich nekonečně malé šířce. Omezená možnost stanovení míry rozlišení vektorových dat je silným argumentem pro použití rastrových dat v důsledných vědeckých výzkumech (Goodchild, 2011).

2.2.1. Nejistota dat

Nejistota vzniká, protože geografické informace reprezentující prostředí jsou nedokonalé, vždy jsou do jisté míry nepřesné, nejasné, neurčité nebo nějakým způsobem zavádějící. Nepřesnost je výsledkem neúplnosti informací, nedostatečného postihu detailu. Člověk nemůže nikdy pozorovat a zaznamenat úplný detail, a to ani za pomoci měřících přístrojů. Informace získaná pomocí technologií nemůže nikdy plně popsat ani fyzikální vlastnosti prostředí (Veregin, 1999).

Přesnost byla definována ve statistickém smyslu jako ‚odchylka pozorování od skutečné hodnoty‘. Takovéto odkazy na skutečnost jsou vysoce problematické, proto se přesnost v praxi běžně hodnotí porovnáním s nezávislým datovým zdrojem vyšší přesnosti. Tato definice však představuje nepříjemné točení se v kruhu, obvykle je možné provést subjektivní hodnocení relativní přesnosti různých datových sad. Realističtější definicí přesnosti je ‚odchylka pozorované hodnoty od hodnoty vykazované z nezávisle odvozené, spolehlivější datové sady‘. Dalším důležitým a často ignorovaným zdrojem nepřesnosti je zjednodušující předpoklad, že geografické prostředí je statické. Většina geografických informací není explicitně temporální, tudíž se stává nepřesnou v důsledku dynamického vývoje životního prostředí v průběhu času (Duckham et Sharp, 2005).

2.2.2. Data DPZ

Dálkový průzkum Země je v současnosti jedním z nejčastějších zdrojů dat. Snímky dálkového průzkumu mají široké uplatnění v mnoha vědních oborech. V oblasti ekologie jejich využití jednoznačně zvýšilo naše porozumění prostorovým složkám přírodního prostředí, je však důležité uvědomit si omezení těchto dat (Hay et al., 2001).

Klíčové omezení pramenící ze spojitosti s MAUP je problematika prostorového měřítka (*scale effect*). Data dálkového průzkumu jsou většinou prezentována v rastrovém formátu (Benz et al., 2004). Jsou získávána v předem definovaném

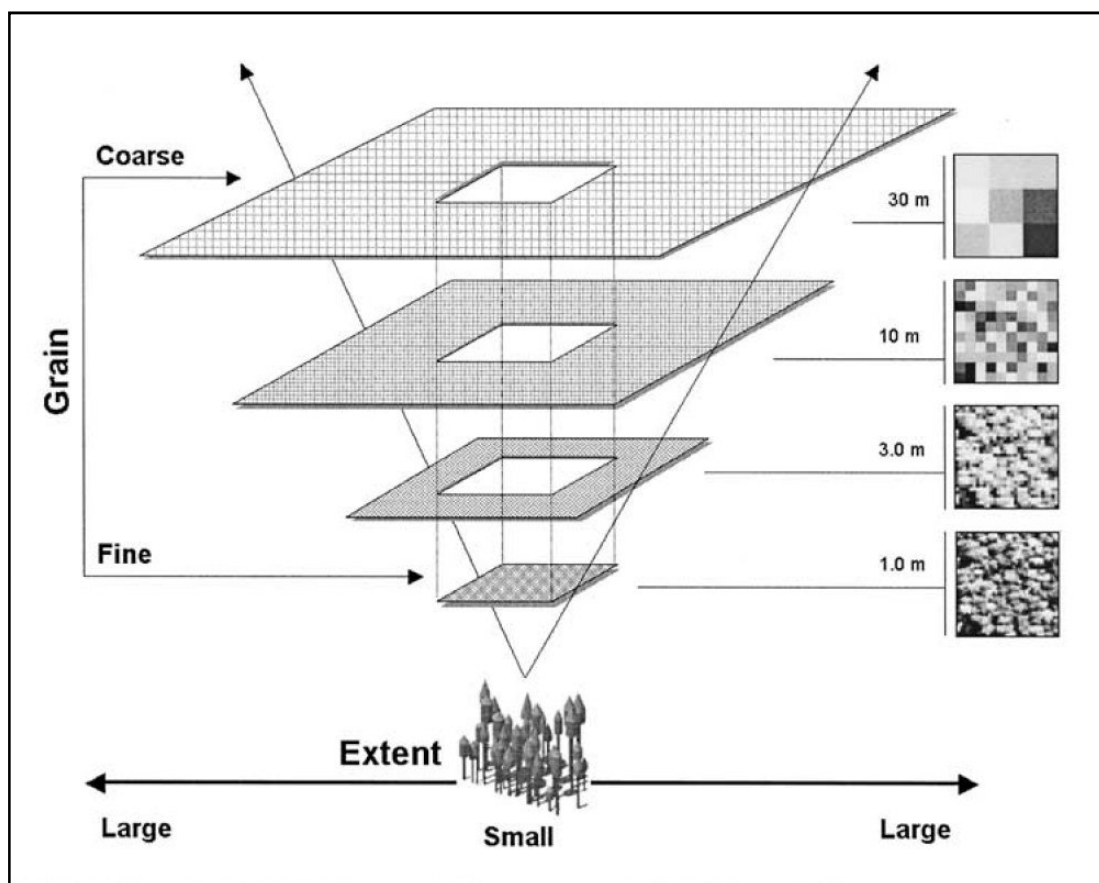
měřítka bez ohledu na konkrétní jevy, které jsou posléze z těchto dat zjišťovány (Hay et al., 2001), nevnímají informace v kontextu (Benz et al., 2004). Měřítka analýzy u satelitních snímků je předem stanoveno rozlišením dat (Jelinski et Wu, 1996). Snímky dálkového průzkumu v podstatě představují libovolný vzorkovací grid (mřížku) promítnutý na povrch Země (obr. č. 5), vzorkovací jednotky jsou tudíž dány technickými parametry satelitu namísto ekologické významnosti. A právě kvůli agregaci reálných dat na základě velikosti gridu a ne na základě přirozených přírodních procesů jsou data dálkového průzkumu vysoce náchylná k MAUP. Navíc jsou získané snímky pro potřeby dalšího výzkumu často převzorkovány (*re-sampled*), při čemž jsou hodnoty sousedních buněk zprůměrovány nebo jinak matematicky kombinovány za účelem vyhlazení nebo získání tzv. filtrů (Lillisand et al., 2004). Toto převzorkování odpovídá agregačnímu efektu, který může mít za následek nesprávnost produkovaných dat nebo v případě extrahování dat pro statistické analýzy i chybné výsledky analýz (Dark et Bram, 2007).

Benz et al. (2004) shrnuli několik typů přirozených neurčitostí a nejistot, jež ovlivňují získávání informací z dat dálkového průzkumu Země. Nejistota začíná při měření senzorem s omezenou přesností, pokračuje degradací informace metodami zpracování signálu pro kompresi a filtrování dat, dále nejednoznačnou extrakcí prvků a klasifikací a končí nepřesnými koncepty tříd krajinného pokryvu (*landcover*) a tříd využití krajiny (*landuse*).

Senzorová měření mají i při pečlivé kalibraci přístroje limitované radiometrické a geometrické rozlišení, což ovlivňuje schopnost rozlišit jednotlivé kategorie krajinného pokryvu, a tím vede k omezené možnosti jednoznačné klasifikace (Benz et al., 2003).

Proces generování obrazu představuje převedení (konverzi) sensorových měření na obrazová data. Data musí být navíc komprimována pro snížení nároků na jejich uchování a přenos. Ve většině případů tyto kroky zanechávají určité pozůstatky a nejasnosti (rozpory), které vedou k datovému šumu, a tudíž další nejistotě ve výsledných obrazových datech (Benz et al., 2004).

Pro krajinný pokryv a využití krajiny existují obvykle jen rozostřené koncepty. Neexistuje žádný přesný práh mezi hustě a řídce osídlenou oblastí, ani mezi nízkou a vysokou vegetací, proto kdykoliv jsou nějaké prahy definovány, jsou většinou neuspokojivou idealizací reálného světa a vedou k problémům při klasifikaci dat. Je jasné, že po použití těchto prahů pro vymezení tříd nebude výsledná klasifikace optimální (Benz et al., 2004).



Obr. č. 5 Rozlišení dat dálkového průzkumu Země (Hay et al., 2001).

2.3. Virtuální druh

Simulace virtuálního druhu je široce užívanou technikou zejména v modelování druhové distribuce, slouží k validaci predikce druhové distribuce a dalších biogeografických hypotéz (Leroy et al., 2015; Miller, 2014). Jejich výhodou oproti reálným datům je znalost a ovladatelnost základního mechanismu vztahu druhu k prostředí i generování distribuce druhu (Thibaud et al., 2014). Protože vlastnosti druhu mají na scale of effect nezanedbatelný vliv (Miguet et al., 2015), byl virtuální druh použit i při tomto experimentu, analýza tak není ovlivněna vlastními prostorovými nároky druhu a detekcí skutečného scale of effect.

V podstatě se jedná o předem definovanou ekologickou odpověď (výskyt, abundanci) jako závislé proměnné v simulovaném vztahu druhu a prostředí. Simulovaný vztah druhu a prostředí je odvozen od prostorových dat - environmentální proměnné, kterou je např. podíl určitého typu krajinného pokryvu (Lechner et al., 2012a) nebo jiná environmentální proměnná (Meynard et Kaplan, 2013; Garzon-Lopez et al., 2015).

Dobrým nástrojem pro simulaci virtuálního druhu je open-source package pro prostředí R, který navržený tak, aby poskytoval kompletní rámec pro simulaci virtuální druhové distribuce, jež umožňuje generovat virtuální druhy prostřednictvím různorodých vztahů druhu a prostředí (různorodých tvarů odpovědných křivek), a zároveň je robustní, uživatelsky přívětivý, a hlavně umožňuje simulaci virtuálního druhu s důrazem na vyšší ekologický realismus (tj. simulovaný vztah druhu a prostředí je blízký reálnému; Leroy et al., 2015).

Využití simulovaných dat v našem experimentu umožní rozlišit mezi skutečnými účinky simulovaného vztahu druhu a prostředí a vlastními účinky prostorového rozlišení podkladových dat (Lechner et al., 2012a).

2.4. Regresní model

Regresní analýza je jednou z nejpoužívanějších statistických metod. Používá se k hodnocení vztahů mezi dvěma nebo více atributy (vlastnostmi). Identifikace a kvantifikace vztahu umožňuje lépe porozumět, co se na určitém místě děje, predikovat, kde něco pravděpodobně nastane, nebo zkoumat důvody, proč se věci dějí v daných místech (Mitchel, 2015). Základem je regresní rovnice, matematický předpis aplikovaný na vysvětlující proměnné (*explanatory variables*), za účelem co nejlepší předpovědi závislé proměnné (*dependent variable*). Závislá proměnná je v rovnici označována y , vysvětlující nebo též nezávislé proměnné zastupují x . Každá nezávislá proměnná je spjata s regresním koeficientem β vyjadřujícím sílu a směr (pozitivní/negativní) vztahu dané proměnné k závislé proměnné. Celý zápis vypadá následovně: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$. Rezidua představují nevysvětlitelnou část závislé proměnné a jsou v rovnici reprezentována jako náhodná chyba symbolem ε . K vystavění a kalibraci regresního modelu jsou použity známé hodnoty závislé proměnné. Použitím známých hodnot závislých a vysvětlujících proměnných, nástroj pro vytváření regresních modelů vytvoří rovnici, která se snaží predikovat tyto známé hodnoty závislé proměnné y co možná nej přesněji. Nicméně predikované hodnoty jen zřídka kdy přesně odpovídají pozorovaným hodnotám, a právě rozdíly mezi nimi se nazývají rezidua (Regression analysis basics—ArcGIS Pro | ArcGIS Desktop (online) [cit. 10.3.2017]).

Lineární vztahy jsou buď pozitivní, nebo negativní. Pozitivní nastává, když se jedna veličina zvyšuje s růstem jiné veličiny, nebo když jedna veličina klesá s poklesem jiné. Negativní vztah je takový, kdy jedna z veličin roste a jiná klesá. Může nastat i případ, že mezi proměnnými není žádný vztah. Korelační analýzy testují sílu vztahu

mezi dvěma proměnnými, zatímco regresní analýzy se snaží ukázat, do jaké míry jedna nebo více proměnných potenciálně podporuje pozitivní nebo negativní změnu jiné proměnné (Regression analysis basics—ArcGIS Pro | ArcGIS Desktop (online) [cit. 10.3.2017]).

2.4.1. Ordinary Least Squares (OLS)

Nejznámější regresní metodou je *Ordinary Least Squares* (OLS). Poskytuje globální model proměnné nebo procesu, které se snažíme pochopit nebo predikovat. K reprezentaci procesu používá jednu regresní rovnici (Mitchel, 2015).

Nejdůležitějšími výstup OLS pro tento experiment jsou bodové grafy popisující vztah mezi závislou a vysvětlující proměnnou, přičemž každý graf je doplněn histogramem ukazujícím rozdělení proměnných. Dalšími nezbytnými výstupy jsou výsledné koeficienty analýz. Mezi základní patří vícenásobné R^2 (*Multiple R-Squared*) a upravené R^2 (*Adjusted R-Squared*). Oba slouží k hodnocení výkonnosti/přesnosti modelu. Mohou dosahovat hodnot od 0 do 1 včetně. Upravené R^2 je vždy o něco nižší než vícenásobné R^2 , protože se v něm odráží i složitost/komplexnost modelu (počet proměnných). Proto se považuje za přesnější při měření přesnosti modelu (Regression analysis basics—ArcGIS Pro | ArcGIS Desktop (online) [cit. 10.3.2017]). V našem případě, kdy do modelu vstupovala vždy jen jedna vysvětlující a jedna závislá proměnná, byl rozdíl výše zmíněných R^2 minimální (zanedbatelný).

3. Metodika

V následujících podkapitolách jsou uvedeny použité datové sady a jejich základní charakteristiky. Dále jsou popsány jednotlivé kroky výstavby multi-scale studie a způsob vyhodnocení míry ovlivnění identifikace scale of effect vstupními daty při analýze vztahu mezi environmentální proměnnou a ekologickou odpovědí.

3.1. Vstupní data

Vstupní data tvoří několik datových sad vybraných na základě dostupnosti, charakteristik (typ reprezentace prostorových dat – rastr nebo vektor, přesnost atp.) a časového určení. Všechny datové sady jsou volně dostupné na internetu, jejich přehled je v následující tabulce (tab. č. 1)

Název	Zdroj dat	Rozlišení	Rok	Zájmová třída	Zkratka
CORINE Land Cover	http://eea.europa.eu/	100 m	2000	les, voda	L1, V1
CORINE Land Cover	http://eea.europa.eu/	250m	2000	les, voda	L2, V2
CORINE Land Cover	http://eea.europa.eu/	vektor	2000	les, voda	LV, VV
Dibavod	http://www.dibavod.cz/	vektor	2004	voda	Di
Forest MAP 2000	http://forest.jrc.ec.europa.eu/	25 m	2000	les	FM
Global Inland Water	http://www.landcover.org/	30 m	2000	voda	GW
GlobeLand30	https://landcover.usgs.gov/	30 m	2010	les, voda	GL, GV
Treecover2010	https://landcover.usgs.gov/	30 m	2012	les	TC

Tab. č. 1 Přehled vstupních datových sad

CORINE Land Cover (Coordinate Information on the Environment) – Datová sada vyvinutá na základě multispektrálních satelitních snímků Landsat-7 ETM s geometrickou přesností ≤ 25 m, MMU 25 ha. Dosažená tematická přesnost CLC je ≥ 85 % (CORINE Land Cover — Copernicus Land Monitoring Service (online) [cit. 29.8.2016]). Rrastr bylo nutné reklasifikovat na binární rastr zájmového krajinného pokryvu (vodní plochy – „water bodies“ dle Level 1, les – „forests“ dle Level 2); v případě vektoru bylo nutno exportovat prvky zájmového pokryvu do nových vrstev.

Dibavod (Digitální báze vodohospodářských dat) – Databáze vektorových dat, primárně vytvořená z ZABAGED©, funguje jako její tematická vodohospodářská nadstavba. Obsahuje 75 objektů v 10 účelově členěných skupinách. Je průběžně aktualizovaný a spravovaný VÚV TGM, v.v.i. (VÚV T.G.Masaryka - Oddělení GIS - O projektu DIBAVOD (online) [cit. 29.8.2016]).

Forest MAP 2000 - Datová sada je odvozena z 415 Landsat ETM+ satelitních snímků a CORINE Land Cover 2000 (Pekkarinen et al., 2009). Obsahuje kategorie forest, non-forest a clouds/snow. Sada je poskytována jako jednotlivé vzájemně navazující vrstvy. Vrstvy byly nejprve složeny v jeden rastr pomocí *Mosaic to new raster*, který byl poté reklasifikován na binární.

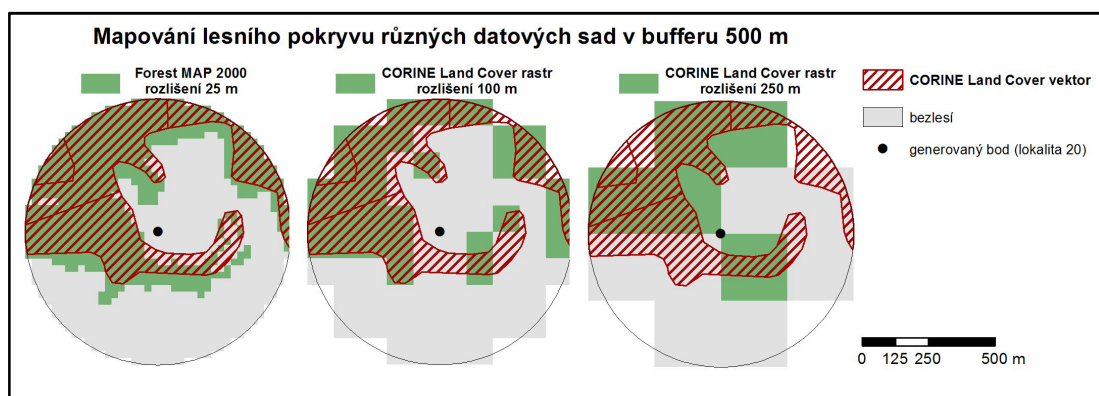
Global Inland Water – Globální dataset vnitrozemských vodních ploch vytvořený pomocí automatického topograficko-spektrálního klasifikačního algoritmu, s využitím snímků Landsat na bázi odrazivosti povrchu, multispektrálních vodních a vegetačních indexů, terénních metrik a stávajících satelitních dat s hrubým rozlišením MODIS (Feng et al., 2016). Dataset je poskytován po částech (různý extent skenování), jednotlivé části se v některých místech překrývají a mohou se lišit v záznamu. Rastr nabývá pěti hodnot (1 - land, 2 - water, 3 - snow/ice, 200 – cloudshadow, 201 - cloud). Části byly nejprve reklasifikovány na binární rast zájmového pokryvu (vodní plochy) a poté složeny v jeden rastr pomocí funkce *Mosaic to new raster*. Při odlišnosti hodnot v překrývajících se částech, pokud alespoň v jedné vrstvě byla detekována voda, byla zaznamenána i ve výsledném rastru.

GlobeLand30 – Produkt globálního mapování krajinného pokryvu vzniklý v rámci *National High Technology Research and Development Program of China*. Dataset zahrnuje 10 tříd krajinného pokryvu (cultivated land, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surfaces, bareland, permanent snow and ice). Podkladem byly multispektrální snímky Landsat TM a ETM+ a multispektrální snímky Chinese Environmental Disaster Alleviation Satellite. Vybírány byly prvořadě bezoblačné snímky z roku 2010 ± 1 rok, v případě, že v oblasti chyběly vhodné snímky, zvětšil se časový rámec (Glc 30 Tianditu (online) [cit. 12.12.2016]). Data jsou poskytována jako jednotlivé, částečně se překrývající vrstvy. Vrstvy byly reklasifikovány na binární rasty zájmového pokryvu (les, vodní plochy) a potom funkcí *Mosaic to new raster* složeny v celistvé vrstvy zájmového pokryvu.

Treecover2010 – Data globálního mapování lesního pokryvu, kde každý pixel udává celočíselný odhad procentuálního maxima lesního pokryvu (1-100) odvozeného ze snímků Landsat 7 ETM+ pořízených každoročně během vegetačního období od roku 2000 do 2012 včetně (Hansen et al., 2013). Mezery a rozpory v datech z jednotlivých let byly nahrazeny hodnotami víceletých mediánů. Data jsou poskytována jako jednotlivé vrstvy. Nejprve byla provedena reklasifikace vrstev na binární rastry zájmového pokryvu, 50% a víc označeno jako les, a potom funkcí *Mosaic to new raster* byly jednotlivé vrstvy složeny v jednu.

Práce s daty probíhala v souřadnicovém systému ETRS 1989 LAEA. Průzkum v lokálním měřítku byl prováděn v rozsahu České republiky.

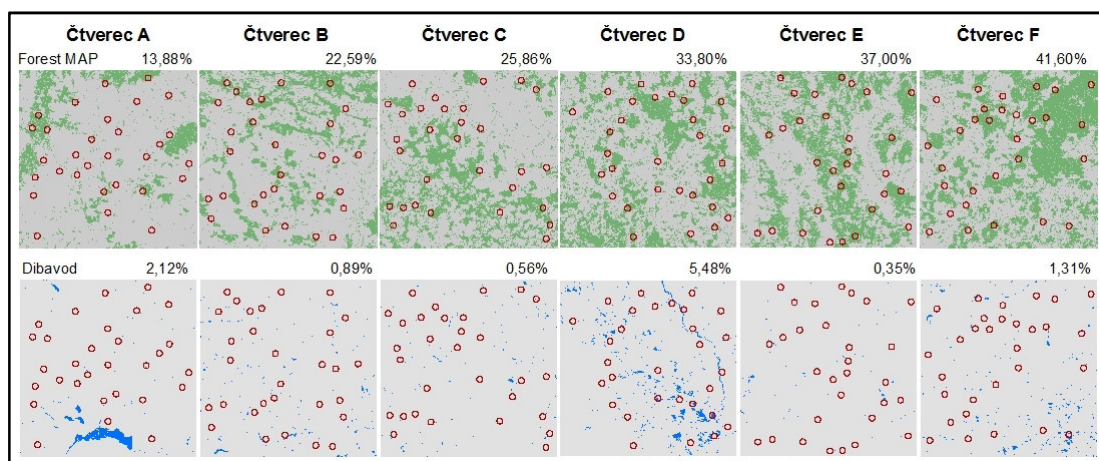
Přestože vstupní data charakterizují tu samou krajinu, v důsledku jiné metody jejich získávání a zpracování, nebo kvůli jinému období pořizování, mohou jednotlivé datové sady charakterizovat krajinu odlišně (obr. č. 6). Proces získávání a zpracování dat je určující pro výsledné rozlišení dat, mapování v jiném časovém období zas postihuje vývoj a proměnu krajiny v čase.



Obr. č. 6 Příklad reprezentace lesního pokryvu v okolí lokality (generovaného bodu) v buffru o poloměru 500 m.

3.2. Popis experimentu

Pro experiment bylo na území ČR náhodně vybráno 6 čtverců (sub-regionů) A až F, které charakterizují krajinu jiným poměrem zastoupení zájmové třídy krajinného pokryvu (voda, les) a jiným krajinným vzorem (obr. č. 7). Větší počet čtverců pojme širší rozpětí ekologicky významných krajinných prvků (Lechner et al., 2012a). Velikost čtverce 45 km x 45 km byla zvolena s ohledem na jejich dostatečný počet v rámci ČR a s ohledem na volbu velikosti měřítek (buffrů) reprezentujících scale of effect. V rámci každého čtverce bylo náhodně vygenerováno 30 bodů (běžně užívaný počet; Pearman et al., 2002; Taki et al., 2007; Louzada et al., 2009; Ricci et al., 2009; Lechner et al., 2012a), kolem nichž jsou následně tvořeny buffry. Body jsou generovány v dostatečné vzájemné vzdálenosti (větší než dvojnásobek poloměru největšího buffru použitého v analýze), aby se snížil vliv prostorové autokorelace a nedošlo k porušení předpokladu nezávislosti při lineární regresi (Lechner et al., 2012a).



Obr. č. 7 Čtverce charakterizující krajinu. V horní řadě data Forest MAP, mapování lesního pokryvu (zeleně) s procentuálním podílem lesa v každém čtverci. V dolní řadě data Dibavod, mapování vodních ploch (modře) s procentuálním zastoupením vodních ploch v každém čtverci. V každém čtverci 30 náhodně vygenerovaných bodů (hnědé kružnice, rádius 750 m).

Samotný experiment se skládá z několika kroků. Do analýzy vstupovalo 6 čtverců, 3 možné velikosti skutečného (správného) scale of effect, 5 datových sad mapujících vodní plochy a 5 datových sad mapujících les, což dává $6 \times 3 \times 5 = 90$ jednotlivých experimentů pro vodu a stejně tak pro les. Zjednodušené schéma experimentu ukazující posloupnost některých kroků je zobrazeno v příloze 1.

- 1) Prvním krokem každého experimentu je definování ekologické odpovědi a správné velikosti scale of effect.

Pro reprezentaci skutečného krajinného pokryvu byly vybrány 2 datové sady: Dibavod (reprezentující vodní plochy) a Forest MAP (reprezentující les), ty slouží pro simulaci ekologické odpovědi. Kolem vygenerovaných bodů byly v rámci všech čtverců vytvořeny buffry velikosti 250 m, 500 m a 750 m, které představují skutečné scale of effect. Pro každou lokalitu a každou velikost buffru je spočítán procentuální podíl zájmové třídy krajinného pokryvu z dat Dibavod a Forest MAP. Na začátku každého experimentu je jedna z velikostí těchto buffrů prohlášena za skutečné scale of effect pro ten daný experiment.

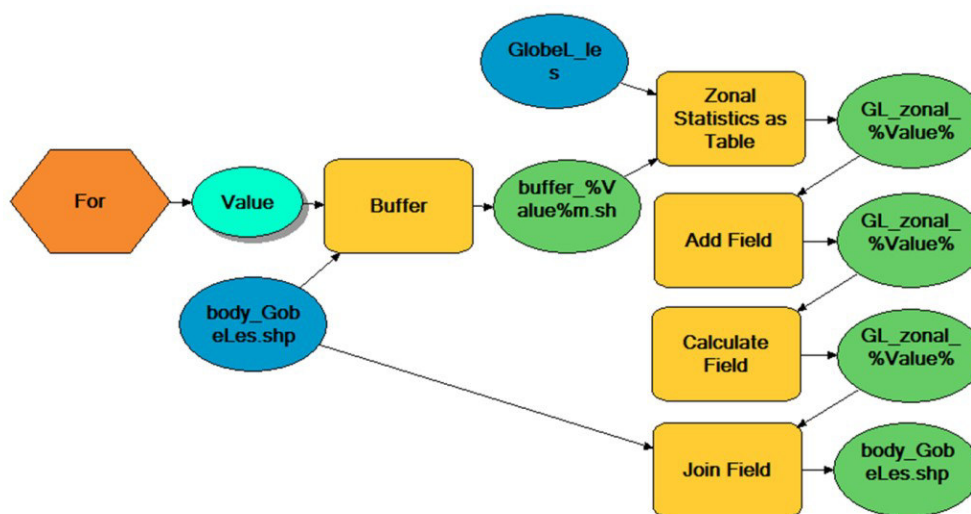
Ekologickou odpovědí je abundance (y) závislá na podílu zastoupení zájmového krajinného pokryvu (x) v jedné ze tří možných velikostí buffru (té, která byla prohlášena na počátku jako skutečné scale of effect). Zvolena byla nejjednodušší možná lineární závislost ($y = x$), která se běžně užívá ve studiích zabývajících se vztahem druhu a prostředí (Lechner et al., 2012a). Samotná výchozí abundance je tudíž rovna procentuálnímu zastoupení zájmového

krajinného pokryvu v buffru scale of effect, a představuje vysvětlující proměnnou (*explanatory variable*) v lineárně regresním modelu.

Pro každou kombinaci jednoho čtverce (A až F), jedné velikosti scale of effect (250 m, 500 m nebo 750 m) a jedné výchozí datové sady (Dibavod nebo Forest MAP) tedy existuje správná (skutečná) korelační křivka RCC (*recover correlation curve*), jež popisuje korelaci mezi ekologickou odpovědí a podílem zájmového krajinného pokryvu.

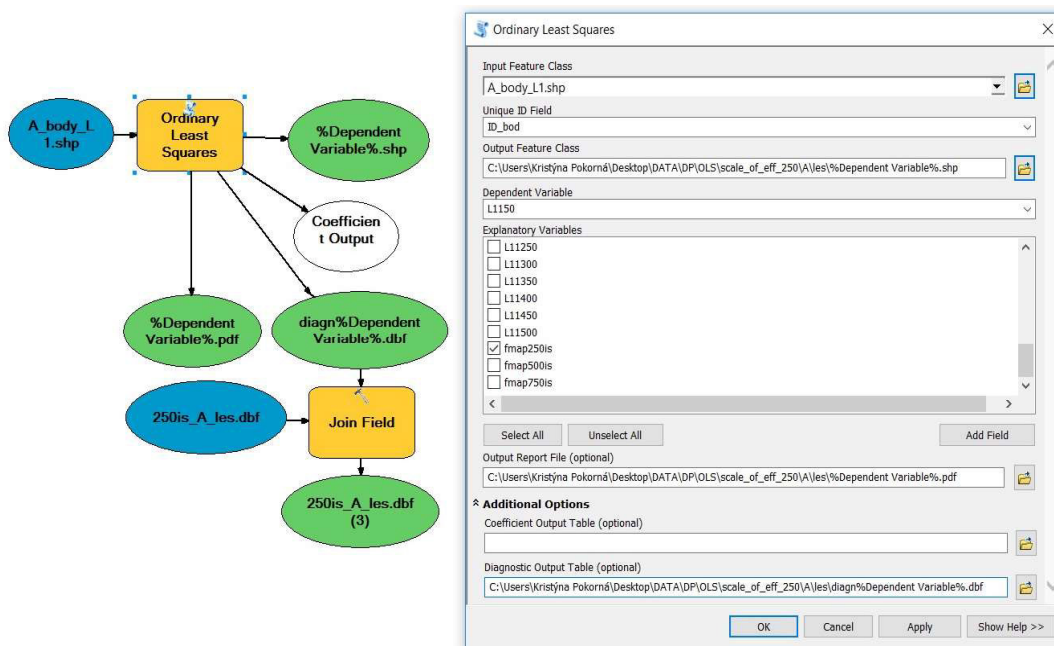
- 2) V dalším kroku jsou pro každý čtverec při použití ostatních datových sad (tab. č. 1) odvozeny zdánlivé RCC následujícím způsobem.

Hodnoty procentuálního zastoupení zájmové třídy krajinného pokryvu ve 30-ti lokalitách v rámci jednoho čtverce (pro výchozí datovou sadu a skutečnou velikost scale of effect) zjištěné v prvním kroku představují vysvětlující proměnnou. Pro tytéž lokality je znovu měřen procentuální podíl zájmové třídy krajinného pokryvu (obr. č. 8), tentokrát však za použití ostatních datových sad a pro větší škálu velikostí bufferů (rádius 150 m až 1 500 m po 50 m). Naměřené hodnoty představují závislou proměnnou (*dependent variable*) v lineárně regresním modelu.



Obr. č. 8 Model výpočtu procentuálního zastoupení zájmového krajinného pokryvu v buffrech pro rastrové datové sady. Pomocí iterátoru „For“ se kolem generovaných bodů postupně tvoří buffry od 150 m do 1500 m vždy po 50 m, Nástroj „Zonal Statistics as Table“ následně v těchto buffrech spočítá celkovou plochu zájmového krajinného pokryvu (pro každou lokalitu zvlášť). Poté je přidán nový atribut, v němž je pomocí „Calculate field“ spočítán procentuální podíl pokryvu vůči celkové ploše buffru. Tento atribut je nakonec přidán do atributové tabulky původní vrstvy bodů a celý cyklus se opakuje pro další velikost buffru.

Samotná analýza lineární regrese je provedena rovněž v ArcGIS a to pomocí nástroje *Ordinary Least Squares* (OLS; obr. č. 9), kde jsou naměřené hodnoty pro každou datovou sadu a každou velikost buffru porovnány s ekologickou odpovědí měřenou v kroku 1. Prakticky to znamená porovnání hodnot procentuálního zastoupení zájmové třídy krajinného pokryvu naměřených v tomto kroku (pro všechny datové sady a větší škálu buffrů) s hodnotami procentuálního zastoupení krajinného pokryvu zjištěnými v prvním kroku (pro Dibavod a Forest MAP ve správných scale of effect).



Obr. č. 9 Model OLS. Vstupní vrstvou jsou body ve čtverci A, atributová tabulka vrstvy obsahuje procentuální zastoupení lesa z dat CORINE Land cover raster 100 m pro jednotlivé buffry 150 m až 1500 m (dependent variable) a procentuální zastoupení lesa z dat FMAP pro buffry scale of effect 250 m, 500 m, 750 m (explanatory variable).

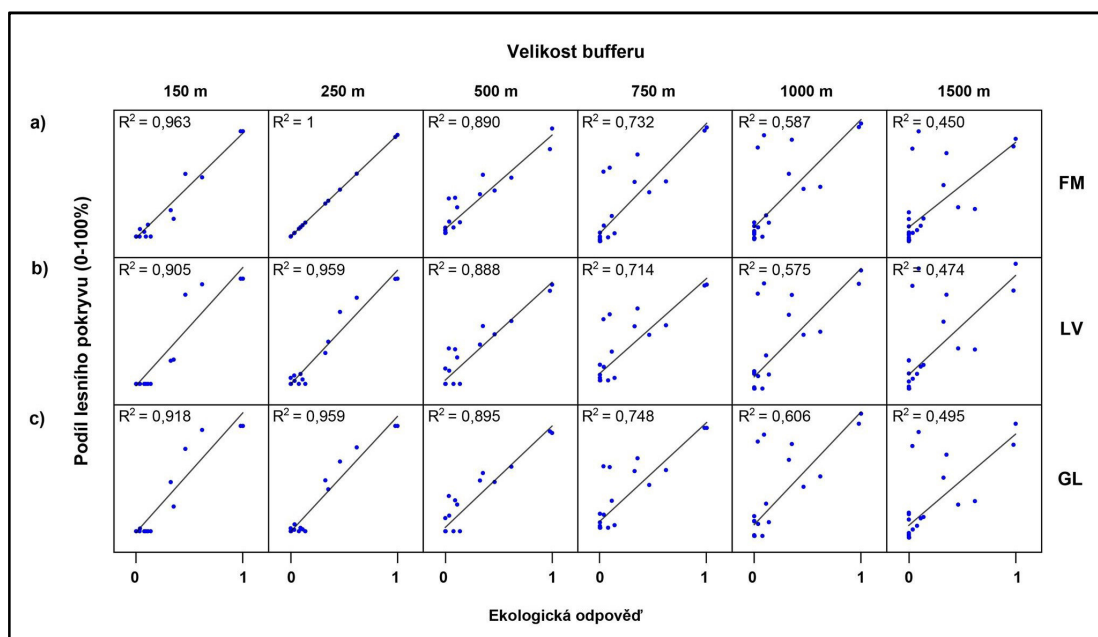
- 3) Výsledný koeficient R^2 značí korelaci mezi proměnnými. Může nabývat hodnot 0 až 1 včetně, kde 1 vyjadřuje nejvyšší možnou korelaci. Velikost buffru, v němž je odvozena správná ekologická odpověď pro data Dibavod a Forest MAP, představuje skutečnou velikost scale of effect, v tomto buffru jsou použita totožná data pro závislou i vysvětlující proměnnou, a proto R^2 nabývá hodnoty 1. Se změnou velikosti buffru a/nebo datové sady se R^2 sníží. Velikost buffru vykazující nejvyšší hodnotu R^2 na zdánlivých RCC je označena jako zdánlivé scale of effect.
- 4) Vliv vstupních dat na scale of effect zjistíme vzájemným porovnáním správných RCC generovaných při použití datových sad Dibavod nebo Forest MAP se

zdánlivými RCC generovanými za použití ostatních datových sad. Maxima (vrcholy) a minima RCC odpovídají ekologické odpovědi na velikost buffru, stejně jako v jiných ekologických studiích (Pearman, 2002; Holland et al., 2004). Má-li zdánlivá RCC víc vrcholů, může být identifikováno vícerozměrné scale of effect, přestože skutečná RCC byla odvozena pouze z jednoho scale of effect. V případech, kdy maxima dosahuje více velikostí buffru (tj. v případech, kdy je tvar křivky plochý a nemá žádný vrchol), je scale of effect identifikováno jako průměr maximálních dosažených hodnot. Rozdíly mezi RCC pro různé datové sady, stejně jako rozdíly scale of effect ze skutečné RCC a zdánlivé RCC, ukazují na vliv vstupních dat (přítomnost MAUP) a tím i nespolehlivost při identifikaci scale of effect.

4. Výsledky

Síla vztahu vegetačního pokryvu a ekologické odpovědi se mění se změnou datové sady a velikosti buffru, a to jak pro lesní pokryv, tak pro vodní plochy. V případě lesa, jako zájmové třídy krajinného pokryvu, proběhla analýza OLS bez problémů, v každém čtverci existovaly pro všechny datové sady lesního pokryvu (L1, L2, LV, FM, GL, TC) a pro všechny velikosti buffru nějaké nenulové záznamy procentuálního podílu lesních ploch. Naproti tomu v případě vodních ploch, ve 4 čtvercích z 6 (pro všechny lokality v rámci čtverce), byl pro některé datové sady vodních ploch v některých velikostech buffru zaznamenán nulový podíl vodních ploch, což způsobilo nemožnost provedení analýzy OLS. Zdánlivé RCC a identifikace zdánlivého scale of effect tak bylo omezeno pouze na oblast buffrů, kde byly nenulové záznamy procentuálního podílu vodních ploch.

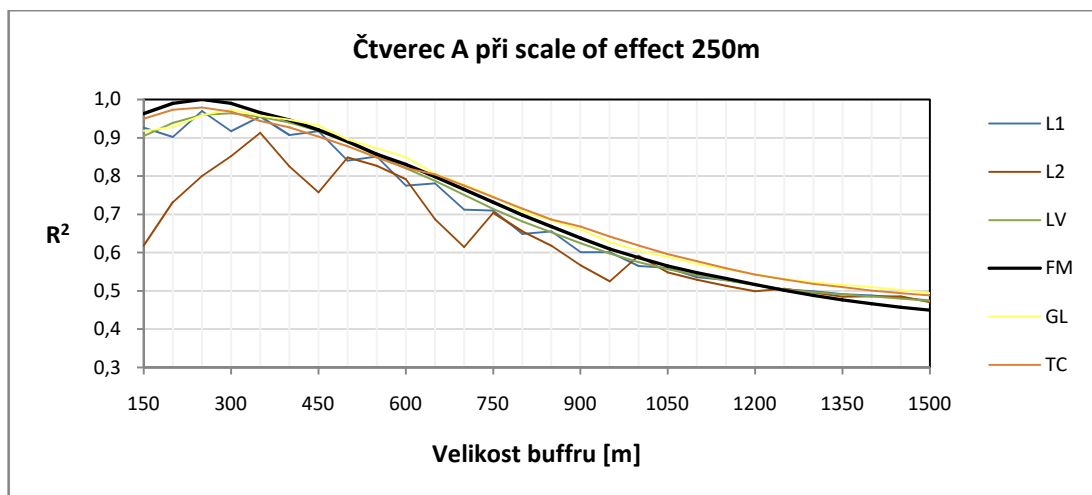
Obr. č. 10 ukazuje jako příklad tři sady bodových grafů (v několika vybraných velikostech buffru) popisujících vztah mezi procentuálním podílem lesního pokryvu a ekologickou odpovědí, které jsou výstupem analýzy OLS. U každého grafu je i naměřená hodnota R^2 . Vrchní řada grafů (obr. č. 10a) slouží pro tvorbu správné RCC. Grafy byly generovány z dat Forest MAP pomocí lineární regrese analyzující vztah mezi podílem lesního pokryvu ve správném scale of effect (250 m) a podílem lesního pokryvu ve všech velikostech buffru. Ekologická odpověď byla odvozena z buffru 250 m, proto je R^2 v buffru velikosti 250 m rovno 1 a na grafu je vidět dokonalý lineární vztah. Se změnou velikosti buffru se změní korelace mezi podílem lesního pokryvu a síla ekologické odpovědi klesne, hodnota R^2 se sníží. Další dvě řady grafů (obr. č. 10b, 10c) ukazují výsledek analýzy OLS jiných datových sad lesního pokryvu při zachování stejného scale of effect. V případě vektorových dat CORINE Land Cover (tab. č. 2) bylo nejvyšší dosažené R^2 (0,964) zaznamenáno v buffru 300 m, což vede k nesprávné identifikaci scale of effect. Stejně tomu bylo i u dat GlobeLand30 (tab. č. 2), kdy nejvyšší R^2 (0,974) bylo rovněž při buffru 300 m.



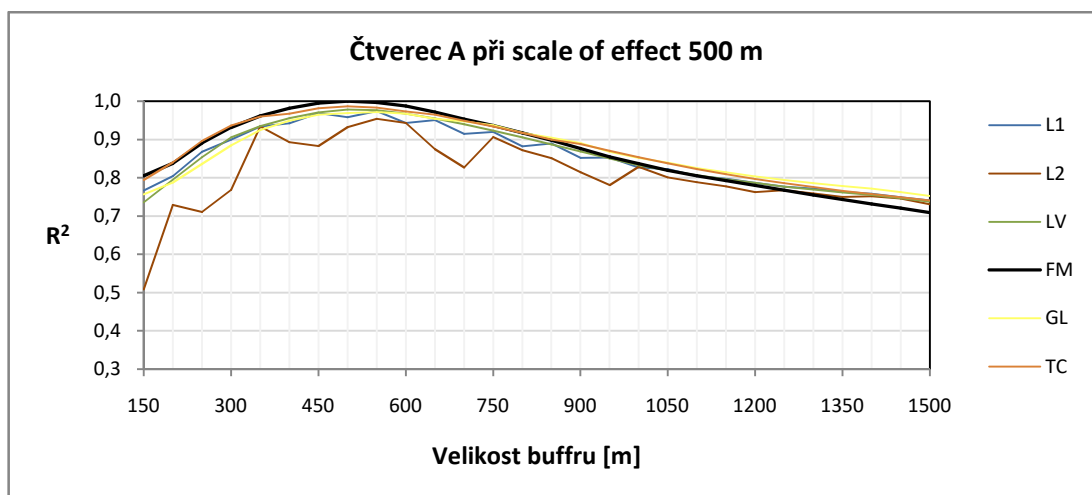
Obr. č. 10 Bodové grafy vztahu ekologické odpovědi a procentuálního podílu zámrového krajinného pokryvu (lesa) z analýzy OLS pro čtverec A. Body v grafu jsou měření v jednotlivých lokalitách. Body jsou proloženy nejlépe odpovídající přímkou lineární regrese. U každého grafu je uvedeno výsledné R^2 . a) data Forest MAP 2000, výchozí datová sada, b) data CORINE Land Cover vektor, c) data GlobeLand30.

Hodnoty R^2 v různých velikostech buffrů pro každou datovou sadu jsou použity ke generování RCC (obr. č. 11), jež ukazuje závislost R^2 na velikosti buffru. Křivka FM odkazuje k první sadě bodových grafů (obr. č. 10a), křivka LV vychází z prostřední řady bodových grafů (obr. č. 10b) a GL z poslední řady grafů (obr. č. 10c). Na grafu RCC (obr. č. 11) je vidět, že jediná datová sada, jejíž křivka dosáhla maximální možné hodnoty R^2 , tedy hodnoty 1, je Forest MAP (FM). Datová sada Forest MAP byla výchozí datovou sadou pro odvození ekologické odpovědi, reprezentuje skutečnou krajinu. Křivka FM je správnou RCC a velikost buffru, ve kterém křivka dosahuje svého maxima ($R^2 = 1$) odpovídá správnému scale of effect.

Pro srovnání, na obr. č. 12 je další graf RCC pro čtverec A, tentokrát při správném (výchozím) scale of effect 500 m. Z něho je patrné, že při menších velikostech buffru (150 až 250 m), jsou dosažené hodnoty R^2 celkově menší oproti hodnotám R^2 při scale of effect 250 m (obr. č. 11), což vypovídá o menší korelaci mezi podílem lesního pokryvu a ekologickou odpovědí. Vrcholy RCC křivek se při správném scale of effect 500 m posunou do rozmezí velikosti buffru 450 m až 650 m, v tomto rozmezí je ekologická odpověď nejsilnější. Při větším správném scale of effect (500 m) jsou ve větších buffrech hodnoty R^2 celkově vyšší (obr. č. 12), což značí silnější lineární závislost právě v této oblasti.



Obr. č. 11 Graf RCC ve čtverci A při skutečném scale of effect 250 m. Skutečná RCC odvozená z dat Forest MAP (FM) a zdánlivé RCC odvozené z ostatních datových sad (L1 – CORINE Land Cover rastr 100 m, L2 – CORINE Land Cover rastr 250 m, LV – CORINE Land Cover vektor, GL – GlobeLand30, TC – Treecover2010). Scale of effect je identifikováno jako velikost buffru, při které křivka dosáhne nejvyšší hodnoty R^2 .



Obr. č. 12 Graf RCC ve čtverci A při skutečném scale of effect 500 m. Skutečná RCC odvozená z dat Forest MAP (FM) a zdánlivé RCC odvozené z ostatních datových sad (L1 – CORINE Land Cover rastr 100 m, L2 – CORINE Land Cover rastr 250 m, LV – CORINE Land Cover vektor, GL – GlobeLand30, TC – Treecover2010). Scale of effect je identifikováno jako velikost buffru, při které křivka dosáhne nejvyšší hodnoty R^2 .

Tab. č. 2 ukazuje, kdy se skutečné a identifikované (zdánlivé) scale of effect liší a kdy se shodují. U dat lesního pokryvu došlo ke špatné identifikaci scale of effect v 62 případech z 90, tedy v téměř 70% případů. Největší shoda byla zaznamenána u vektorových dat CORINE Land Cover, z čehož vyplývá, že tyto data nejlépe odpovídají datové sadě Forest MAP, na kterých byla generována ekologická odpověď. Naproti tomu nejhůř dopadl rastr CORINE Land Cover s rozlišením 250 m,

tedy datová sada s nejnižším rozlišením v rámci datových sad použitých v analýze. Nejčtenější shoda (8 případů z 15) byla ve čtverci A. Oproti tomu ve čtverci E nenastala shoda ani jednou. Kompletní přehled dosažených R^2 a dalších koeficientů, včetně grafů RCC pro data lesního pokryvu jsou uvedeny v příloze 2.

Skutečné scale of effect	Čtverec	L1		L2		LV		GL		TC	
		Zdánlivé scale of effect	R^2	Zdánlivé scale of effect	R^2	Zdánlivé scale of effect	R^2	Zdánlivé scale of effect	R^2	Zdánlivé scale of effect	R^2
250	A	250	0,969	350	0,913	300	0,964	300	0,974	250	0,979
	B	350	0,956	400	0,891	300	0,975	300	0,949	250	0,968
	C	200	0,893	300	0,797	250	0,915	250	0,911	250	0,979
	D	300	0,955	400	0,914	300	0,954	200	0,935	300	0,983
	E	350	0,899	350	0,780	300	0,903	300	0,899	200	0,924
	F	150	0,921	350	0,798	150	0,929	250	0,888	250	0,960
500	A	550	0,974	550	0,954	500	0,978	550	0,972	500	0,986
	B	500	0,981	600	0,943	500	0,983	500	0,972	450	0,986
	C	600	0,904	650	0,834	550	0,908	550	0,911	500	0,985
	D	500	0,953	400	0,902	500	0,952	500	0,948	450	0,989
	E	450	0,927	600	0,846	400	0,919	450	0,921	450	0,944
	F	450	0,911	550	0,837	500	0,914	600	0,946	500	0,976
750	A	750	0,988	750	0,970	750	0,989	750	0,989	800	0,991
	B	700	0,983	900	0,956	750	0,984	800	0,979	700	0,989
	C	800	0,934	850	0,896	750	0,940	800	0,933	800	0,983
	D	700	0,941	650	0,896	750	0,942	800	0,938	700	0,988
	E	550	0,902	650	0,896	650	0,900	600	0,915	700	0,957
	F	650	0,933	750	0,886	700	0,936	750	0,945	650	0,985
Počet správně identifikovaných scale of effect		4		2		9		6		7	
Počet špatně identifikovaných scale of effect		14		16		9		12		11	

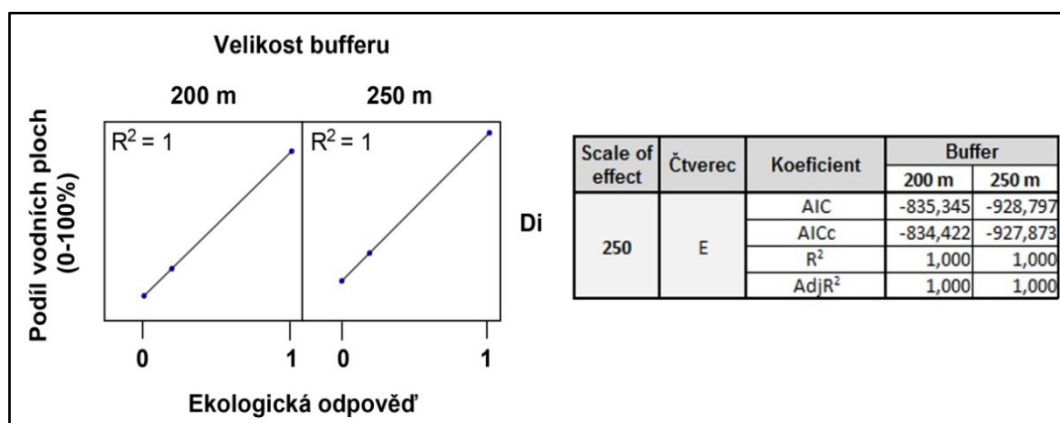
Tab. č. 2 Maximální dosažená R^2 a identifikovaná (zdánlivá) scale of effect pro datové sady lesního pokryvu v jednotlivých čtvercích při různých skutečných scale of effect. Správně identifikované scale of effect je zvýrazněno zeleně.

V případě dat vodních ploch byly mezi jednotlivými datovými sadami výraznější rozdíly v záznamech procentuálního zastoupení krajinného pokryvu a často se také stalo, že v jednotlivých lokalitách daného čtverce byl do určité velikosti buffru nulový podíl vodních ploch. To vedlo k velkému kolísání RCC křivek a méně spolehlivé identifikaci scale of effect. Příklad záznamů procentuálního zastoupení vodních ploch jednotlivých datových sad ve čtverci A je v tab. č. 3. Při velikosti buffru 250 m mají nenulové záznamy pouze data Dibavod. V největším buffru (1500 m) pak polovina datových sad obsahuje pouze jeden nenulový záznam. Analýza OLS může proběhnout pouze v případě, že existuje alespoň jeden nenulový záznam. V opačném případě jsou analýza OLS, grafy RCC i oblasti pro identifikaci scale of effect omezeny jen na rozsah buffrů s alespoň jedním nenulovým záznam. Je tedy předem jasné, že pokud zdánlivá RCC začíná až při velikosti buffru větší než je velikost buffru správného scale of effect, nemohou se zdánlivé a správné scale of effect shodovat.

ID_bod	Buffer 250 m						Buffer 500 m						Buffer 750 m						Buffer 1500 m					
	Di	V1	V2	VV	GV	GW	Di	V1	V2	VV	GV	GW	Di	V1	V2	VV	GV	GW	Di	V1	V2	VV	GV	GW
31	0,160	0	0	0	0	0	0,784	0	0	0	0	0	0,348	0	0	0	0	0	0,107	0	0	0	0	0
32	0	0	0	0	0	0	0,068	0	0	0	0	0	0,337	0	0	0	0	0	0,121	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0,056	0	0	0	0	0	0,014	0	0	0	0	0
35	1,858	0	0	0	0	0	0,939	0	0	0	0	0	0,739	0	0	0	0	0,408	0,248	0	0	0	0	0,102
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0,014	0	0	0	0	0	0,211	0	0	0	0	0	0,094	0	0	0	0	0	0,023	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,055	0	0	0	0	0
39	0	0	0	0	0	0	2,943	2,597	0	5,715	0	2,288	7,197	9,143	10,345	9,802	1,580	11,825	2,576	3,693	3,571	3,561	0,471	4,532
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0,379	0	0	0	0	0	0,236	0	0	0	0	0,356
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,194	0	0	0	0	0,305
45	0	0	0	0	0	0	0	0	0	0	0	0	0,053	0	0	0	0	0	0,159	0	0	0	0	0
46	0	0	0	0	0	0	0,210	0	0	0	0	0	0,093	0	0	0	0	0	0,023	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,479	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,011	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,012	0	0	0	0	0
52	0	0	0	0	0	0	0,021	0	0	0	0	0	0,220	0	0	0	0	0	1,219	0	0	0	0	2,012
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,017	0	0	0	0	0
54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	0	0,081	0	0	0	0	0	0,068	0	0	0	0	0	0,234	0	0	0	0	0,255
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0,030	0	0	0	0	0	0,279	0	0	0	0	0	0,124	0	0	0	0	0	0,031	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,127	0
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,019	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,055	0	0	0	0	0

Tab. č. 3 Procentuální zastoupení vodních ploch v jednotlivých lokalitách ve čtverci A pro jednotlivé datové sady a 4 různé velikosti buffru.

V následující tabulce (tab. č. 4) jsou uvedeny maximální dosažená R^2 a identifikovaná (zdánlivá) scale of effect pro datové sady vodních ploch. Pokud je nejvyššího R^2 dosaženo ve více velikostech buffru, je identifikováno více zdánlivých scale of effect (v tab. č. 4 označeno hvězdičkou). Např. ve čtverci A u všech tří sad CORINE Land Cover (V1, V2, VV) bylo R^2 shodné pro všechny analyzované velikosti buffru, a to při všech třech správných (výchozích) scale of effect. K identifikaci více scale of effect došlo v jednom případě (čtverec E při správném scale of effect 250 m) i u výchozí datové sady Dibavod (obr. č. 13, tab. č. 4), kde kromě správně identifikovaného buffru velikosti 250 m byl za scale of effect označen i buffer 200 m.



Obr. č. 13 Vícečetné zdánlivé scale of effect datové sady Dibavod v analýze vodních ploch ve čtverci E. V levé části bodové grafy vztahu ekologické odpovědi a procentuálního podílu zájmového krajinného pokryvu (vodních ploch) z analýzy OLS spolu s výsledným R^2 . V pravé části tabulka s dalšími koeficienty: AIC - Akaike's Information Criterion, AICc - Corrected Akaike's Information Criterion, R^2 - R-Squared, Adj R^2 - Adjusted R-Squared.

Podrobně jsou výsledky analýz pro datové sady vodních ploch spolu s grafy RCC v příloze 3.

Většina RCC je polynomickou funkcí, což má u dat vodních ploch za následek výskyt vícečetných vrcholů značící více scale of effect, oproti jednomu správnému (výchozímu) scale of effect.

Skutečné scale of effect	Čtverec	V1		V2		VV		GV		GW		Di	
		Zdánlivé scale of effect	R2	Zdánlivé scale of effect	R2	Zdánlivé scale of effect	R2	Zdánlivé scale of effect	R2	Zdánlivé scale of effect	R2	Zdánlivé scale of effect	R2
250	A	*	0,001	*	0,001	*	0,001	*	0,001	*	0,001	250	1
	B	*	0,172	1050	0,120	*	0,172	350	0,637	300	0,849	250	1
	C	1500	0,002	*	0,002	1500	0,002	*	0,990	200	0,992	250	1
	D	300	0,943	150, 200	0,784	1500	0,945	300	0,944	300	0,987	250	1
	E	1500	0,004	1500	0,004	1500	0,004	1500	0,004	1500	0,004	200, 250	1
	F	350	0,993	400	0,989	1300	0,993	300	0,982	300	0,982	250	1
500	A	*	0,845	*	0,845	*	0,845	*	0,845	750	0,862	500	1
	B	*	0,722	1050	0,547	*	0,722	550	0,899	550	0,911	500	1
	C	1500	0,003	*	0,002	1500	0,002	650	0,970	450	0,989	500	1
	D	550	0,980	600	0,823	550	0,986	600	0,988	550	0,996	500	1
	E	1400	0,872	1500	0,782	1400	0,891	1400	0,889	450	0,946	500	1
	F	300	0,994	450, 700	0,994	200	0,981	500	0,981	450	0,999	500	1
750	A	*	0,984	*	0,984	*	0,984	*	0,984	750	0,989	750	1
	B	*	0,797	1050	0,609	*	0,797	600	0,902	700	0,937	750	1
	C	*	0,027	*	0,027	*	0,027	850	0,925	750	0,868	750	1
	D	650	0,989	600	0,903	550	0,948	800	0,991	800	0,996	750	1
	E	1500	0,978	1500	0,929	1450	0,988	1400	0,985	650	0,994	750	1
	F	700	0,995	800	0,995	150	0,994	800	0,997	700	0,999	750	1

Tab. č. 4 Maximální dosažená R^2 a identifikovaná (zdánlivá) scale of effect pro datové sady vodních ploch v jednotlivých čtvcích při různých skutečných scale of effect. Správně identifikované scale of effect je zvýrazněno zeleně. Pole označené hvězdičkou značí, že v daném případě byla identifikována více než 2 scale of effect.

Účinek jednotlivých datových sad, znázorněný změnou hodnot statistických koeficientů a průběhem křivek RCC, se liší v jednotlivých čtvcích a závisí na velikosti buffru použitého k odvození RCC. Při použití jiné datové sady než byla výchozí sada pro odvození ekologické odpovědi, došlo vždy k oslabení ekologické odpovědi, a tedy snížení hodnoty koeficientu R^2 oproti R^2 ve správném scale of effect. U všech datových sad byla prokázána špatná identifikace scale of effect a to při použití tří různých velikostí správného scale of effect (250 m, 500 m a 750 m).

Tam, kde došlo ke správné identifikaci scale of effect, byla vstupními daty ovlivněna pouze síla vztahu mezi ekologickou odpovědí a podílem krajinného pokryvu (zaznamenáno poklesem R^2).

5. Diskuse

Lechner et al. (2012a) prokázali, že prostorová nejistota dat pramenící z MAUP může mít za následek špatné (zavádějící) výsledky při ekologických analýzách. Při svém výzkumu zjistili, že pozorovaná ekologická odpověď opakovaně vykazovala MAUP a že zmíněná ekologická odpověď nebyla pouze vlastností velikosti buffru, ale i prostorového rozlišení dat. Zjistili tak vysokou míru variability v působení (účincích) různých dat na přesnost identifikace scale of effect mezi různými čtverci (s jinou krajinnou strukturou) a pro různé kombinace velikostí pixelu a velikostí buffru použitých k odvození RCC. Vliv dat na identifikaci scale of effect byla potvrzena i při tomto experimentu, avšak zatímco ve studii Lechner et al. (2012a) byla použita pouze jedna datová sada DPZ, která byla za účelem získání několika alternativních datových sad reprezentujících tentýž krajinný pokryv, ale s jiným rozlišením, opakovaně upravena pomocí agregace dat a vyhlazovacího filtru (*smoothingfilter*), při této studii byly použity datové sady z různých zdrojů. Rozdíly datových sad tudíž vznikly už při snímání povrchu, kdy se uplatňuje hlavně rozlišovací schopnost senzoru, a následně pak při zpracování a klasifikaci dat. Navíc, oproti studii Lechner et al. (2012a) byla v analýze použita i vektorová data krajinného pokryvu (Dibavod, CORINE Land Cover). Každá použitá prostorová datová sada vstupující do modelu, ať už vektorová nebo rastrová, představuje riziko zavedení vlivu MAUP. Kromě velikosti pixelu a vyhlazovacího efektu, který testovali Lechner et al. (2012a), může být MAUP výsledkem i dalších měřítkově závislých faktorů, jako MMU, tematické rozlišení a další (Buyantuyev et Wu, 2007; Lechner et al., 2009).

Určování scale of effect je doporučeno provádět z dat s nejvyšší přesností (Miguet et al., 2015). V případě rastru je otázka přesnosti jednoduchá, většinou se jedná o data s nejvyšším rozlišením, u vektorových dat je hodnocení/posouzení přesnosti složitější. Závislost ekologické odpovědi (abundance) na podílu celkového habitatu lesního pokryvu byla generována z dat Forest MAP 2000 s rozlišením 25 m. Celkově nejpočetnější správné identifikace scale of effect bylo dosaženo při použití vektorových dat CORINE Land Cover, což vyvolává otázku, která ze zmíněných dvou datových sad je přesnější, a jak by se změnila výsledky analýzy, pokud by výchozí datovou sadou ke generování ekologické odpovědi byla právě vektorová data CORINE Land Cover.

Vodní plochy na území ČR představují drobné prvky, k jejich reprezentaci je zapotřebí detailnější mapování. Nejpřesnější databází vodohospodářských dat na

našem území je Dibavod. Žádná volně dostupná rastrová data, která byla použita v analýze, neměla dostatečně vysoké rozlišení, aby postihla drobné vodní povrchy (Shanlong et al., 2011), proto shoda záznamů jednotlivých datových sad v podílu vodních ploch v jednotlivých lokalitách byla nízká. Tudíž, ani nejvyšší dostupné rozlišení (30 m) GlobeLand30 a Global Inland Water nepřineslo pozitivní výsledky při identifikaci scale of effect. Klasifikace jednotlivých tříd krajinného pokryvu při rozlišení 100 m nebo dokonce 250 m CORINE Land Cover je otázkou desítek metrů, proto v jejich analýze vodních ploch nikdy nedošlo ke správné identifikaci scale of effect. Naproti tomu lesní pokryv tvoří i větší plochy, které je možné v rastru CORINE Land Cover 100 m a 250 m zaznamenat, identifikace správného scale of effect je tak v tomto případě úspěšnější. Nízká četnost správné identifikace scale of effect při použití rastrových dat CORINE Land Cover, ve srovnání s ostatními přesnějšími datovými sadami, vede k závěru, že pro ekologické analýzy jsou tato data méně vhodná a bylo by lepší použít data přesnější. Účinky MAUP při použití prostorových dat např. z DPZ totiž mohou způsobit, že výsledky některých prostorových statistických analýz budou bezvýznamné (Jelinski et Wu, 1996; Wu et al., 1997; Nelson, 2001).

Ekologická odpověď v experimentu závisela na jedné proměnné, a to celkovém množství biotopu. Ekologické procesy však ovlivňují i jiné krajinné metriky, např. krajinný vzor, konfigurace krajiny (Borgella et Gavin, 2005; Jackson et Fahrig, 2012; Martensen et al., 2012; Šímová et Gdulová, 2012). Bylo prokázáno, že charakteristiky jako výskyt, abundance nebo druhová bohatost jsou spojeny např. s konektivitou a velikostí fragmentů (Martensen et al., 2008; Boscolo et Metzger, 2009; Martensen et al., 2012). Při řešení otázky, jaká velikost krajiny (scale of effect) je nejvhodnější k predikci vlivu lesního pokryvu na úspěšnost obnovy (*restoration successs*), byly do měření vegetační struktury zahrnuty kromě pokryvnosti a hustoty vegetace i parametry jako výška vegetace, biomasa a opad (Crouzeilles et Curran, 2016).

Ekologické analýzy uvažující scale of effect mají i svou nevýhodu. Připomeňme, že scale of effect je správné (nejvhodnější) měřítko, při kterém je ekologická odpověď (např. abundance) v zájmové lokalitě nejlépe předpovězena krajinnou strukturou (např. množstvím biotopu). Tento přístup zaručuje, že krajinné metriky jsou měřeny v měřítku odpovídajícím zájmovému druhu, čímž se maximalizuje pravděpodobnost nalezení vztahu druhu k prostředí, pokud nějaký existuje. Nevýhodou však je, že scale of effect je detekováno až po vzorkování, kdy už je pozdě na využití této informace k designu (návrhu) studie. Proto je potřeba provést vzorkování zájmových

lokalit daleko od sebe, aby se oblasti zatím neznámé velikosti vzájemně nepřekrývaly, tj. byly prostorově nezávislé. Pokud by šlo scale of effect určit předem, bylo by možné zvýšit počet míst ke vzorkování a zároveň by se ušetřil čas a náklady vynaložené na cestování mezi oblastmi (Jackson et Fahrig, 2012).

Virtuálnímu druhu může být nadefinována různorodá funkční odpověď na prostředí (Meynard et Kaplan, 2013). Komplexní ekologické vztahy mohou být ne-lineární a statisticky těžko odvoditelné kvůli faktorům jako je např. prostorová autokorelace (Legendre et al., 2002; Wintle et al., 2005). V tomto experimentu byl pro popis vztahu mezi druhem a prostředím použit jednoduchý lineární model, běžně užívaný jinými autory (Pearman, 2002; Holland et al., 2004; Taki et al., 2007). Důvodem byla jednak možnost provedení analýzy vztahu proměnných v prostředí ArcGIS (pomocí OLS) a jednak relativně následná interpretace výstupů.

V ekologii je uměle stanovená odpověď druhu na environmentální proměnné (virtuální druh) atraktivní zejména na poli modelování druhové distribuce, protože umožňuje kontrolovat prevalenci druhu, tvar křivky reakce (odpovědi) druhu na gradienty prostředí a další charakteristiky (Meynard et Quinn, 2007; Elith et Graham; 2009), což dovoluje provádět ‚virtuální experimenty‘ ve velkých prostorových měřítkách (Zurell et al., 2010). Z této práce však vyplývá, že výsledky ekologických analýz a experimentů se odvíjí nejen od samotného druhu a jeho vztahu k prostředí, ale i od vstupních dat, s nimiž jsou informace o daném druhu v souvislosti. Vliv odlišných typů funkčních odpovědí na výkonnost (přesnost) distribučního modelu zkoumali např. Meynard et Quinn (2007), Elith et Graham (2009), Meynard et Kaplan (2012).

Některé výsledné RCC dosahují vrcholu (maxima) hned na svém počátku, při nejmenší zvolené velikosti buffru, nebo naopak na konci při největším analyzovaném buffru. To ukazuje na možnost, že skutečné scale of effect se nachází mimo rozsah zkoumaných velikostí buffru (Jackson et Fahrig, 2015). Pro zjištění skutečného scale of effect běžnou výstavbou víceúrovňových buffrů je zapotřebí zvolit dostatečný rozsah i počet. Velikost scale of effect (případně vícenásobné scale of effect) je vysoce variabilní v závislosti na daném druhu. Při designu experimentu v případě reálného druhu mohou s volbou rozpětí pro detekci správného scale of effect pomoci druhové rysy, jako např. velikost těla, disperze, velikost domovského okrsku (Jackson et Fahrig, 2012; Stevens et al., 2014). Při důkladné studii vztahu mezi druhovými rysy (disperzní vzdáleností, schopností reprodukce a pohybovým chováním) a scale of effect při závislosti abundance na

množství biotopu bylo zjištěno, že disperzní vzdálenost nejvíce pozitivně ovlivňuje scale of effect, jinak řečeno je jeho nejsilnějším prediktorem. Dalším důležitým prediktorem bylo pohybové chování, nejméně důležitá byla schopnost (míra) reprodukce. Naproti tomu velikost populace měla silně negativní spojitost se scale of effect, tzn. že růst populace způsobil zmenšení scale of effect (Jackson et Fahrig, 2012). Schopnost reprodukce, jež negativně koreluje s velikostí těla (Fagan et al., 2010), zas může objasnit některá kolísání ve vztahu velikosti těla a scale of effect (Jackson et Fahrig, 2012). Asi největší problém je určit, které druhové rysy jsou důležité. Studie dokazuje, že velikost scale of effect reálného druhu je ovlivněna působením více druhových rysů současně (Jackson et Fahrig, 2012).

Mezi další proměnné (kromě již zmíněných - struktury krajiny a druhových rysů), které mohou ovlivnit scale of effect, patří mezidruhové interakce a kvalita biotopu. Obojí může určitým způsobem ovlivnit růst populace a pohybovou aktivitu a tím změnit velikost scale of effect (Jackson et Fahrig, 2012). Použití více vysvětlujících proměnných pro vyjádření složitějších vztahů reálných dat potenciálně zesiluje vliv MAUP na výsledky analýzy (Lechner et al., 2012a). Výběru relevantních proměnných při identifikaci scale of effect jsme se v experimentu vyhnuly právě použitím virtuálního druhu. Detekce skutečného scale of effect navíc nebyla cílem této práce a menší velikost buffru pro měření podílu zastoupení zájmových tříd krajinného pokryvu za účelem dohledání skutečného scale of effect by byla bezpředmětná, vzhledem k rozlišení vstupních datových sad 100 m a 250 m u CORINE Land Cover.

Za scale of effect jsme stejně jako většina studií uvažovali kruhovou oblast (Jackson et Fahrig, 2012; Lechner et al., 2012a; Miguet et al., 2015) v lokálním měřítku, scale of effect by však mělo platit i pro posouzení populace ve velkých oblastech nepravidelného tvaru (Jackson et Fahrig, 2012).

Při porovnání hodnot odvozených korelačních koeficientů s výsledky studie Lechner et al. (2012a), jejich simulační model vykazoval v celkovém průměru vyšší hodnoty a menší rozsah hodnot korelačních koeficientů. Maximální rozsah r pro libovolný čtverec testovaný v jejich analýze byl přibližně 0,57–1,0, což ukazuje na silnější vztah vyplývající z absence jiných neměřených faktorů, které běžně ovlivňují vztah druhu a prostředí (Lechner et al., 2012a). V našem experimentu byl u lesního pokryvu rozsah hodnot $R^2 \sim 0,291-0,989$, u vodních ploch $\sim 0,000-0,999$, v podstatě tedy téměř celý potenciálně možný rozsah R^2 . Další studie zkoumající vztah mezi ekologickým atributem a vegetačním pokryvem měly oproti Lechner et al. (2012a)

rovněž větší rozsah a celkově menší hodnoty r . Nutno však podotknout, že oproti simulačnímu modelu Lechner et al. (2012a) se jednalo o reálné studie. Např. u Pearman (2002) byly hodnoty $r^2 \sim 0,35-0,7$; Holland et al. (2004) dosáhli Pearsonova korelačního koeficientu $r \sim 0,0-0,3$ a Taki et al. (2007) dosáhli $r^2 0,045-0,165$. Bližší vzájemné srovnání je však obtížné vzhledem k různorodosti metod korelačních měření použitých v jednotlivých studiích a vzhledem k různé škále použitých buffrů.

Mohli bychom rovněž spekulovat o významnosti přesné identifikace scale of effect z hlediska dosažených odchylek. U lesního pokryvu, kde máme k dispozici kompletní data, se zdánlivé a skutečné scale of effect lišily ve většině případů (cca 89%) maximálně o 100 m, v cca 61% případů jen o 50 m, což při naší volbě velikostí buffrů znamená odchylku pouze o jednu velikost buffru. Otázkou tedy je, na kolik je taková odchylka velikosti měřítka interakce druhu a prostředí relevantní, pokud budeme modelovat reálný druh s velkou schopností disperze či s velkým domovským okrskem (např. ptáci)

6. Závěr

Provedeným experimentem bylo demonstrováno, že vstupní a mohou výrazně ovlivnit výsledky ekologických analýz. Vliv vstupních dat na identifikaci scale of effect byl prokázán mnohonásobnou analýzou v 6-ti čtvercích (A až F) s různou krajinnou strukturou, za použití několika vybraných, volně dostupných datových sad dvou základních tříd krajinného pokryvu (lesa, vodních ploch). Použitím více čtverců bylo dokázáno, že vliv MAUP v datech krajinného pokryvu se neomezuje pouze na určitý krajinný vzor. Nicméně pro zobecnění, jakým způsobem struktura krajiny ovlivňuje účinky MAUP, je zapotřebí komplexnějších studií pracujících s větší škálou krajinné konfigurace, při současné kontrole všech faktorů, které mají na daný vztah vliv.

Míra vlivu vstupních dat na přesnost identifikace scale of effect vykazovala při různých kombinacích datových sad a velikostí buffru vysokou variabilitu. Úspěšnost správné identifikace scale of effect byla výrazně vyšší v analýze dat lesního pokryvu oproti analýze vodních ploch.

Studie ukázala, že ekologické výzkumy vyžadují citlivostní analýzu měřítka, neboť vztah mezi scale of effect, velikostí buffru a prostorovým rozlišením se může podstatně měnit. Při výzkumech, které používají pevně stanovené měřítko, ať už pozorování nebo analýzy (např. jednu velikost pixelu nebo buffru), může dojít k přehlédnutí důležitého měřítkově závislého vztahu nebo MAUP (Lechner et al., 2012a). Testování citlivosti analýzy vůči měřítku však slouží jen k odhalení existence prostorových nejistot, nepředstavuje účinné řešení tohoto problému. Ideální by bylo rozšířit výzkum a vypracovat účinné metody pro stanovení příčin (zdrojů) měřítkových závislostí a vyvinout metody, jež by snížily účinky MAUP.

Virtuální druh umožní zjednodušení reálných přírodních dějů a spolu s volbou výchozí datové sady, jež reprezentuje skutečný krajinný pokryv, dokáže dokonale popsat, jak daný druh interaguje se svým prostředím. Poskytne tak dostatek prostoru pro zkoumání určitého konkrétního problému, bez vlivu neznámých prostorových nejistot pramenících z reálných dat, u kterých nikdy nelze určit/determinovat tyto vztahy s naprostou jistotou.

7. Použitá literatura

- Allen, T. F., et Starr, T. B. (1982). Hierarchy perspectives for ecological complexity. In Wu, J., et Li, H. (2006). Concepts of scale and scaling. *Scaling and uncertainty analysis in ecology* (pp. 3-15). Springer Netherlands.
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., et Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of photogrammetry and remote sensing*, 58(3), 239-258.
- Bierkens, M., Finke, P., et De Willigen, P. (2000). Upscaling and downscaling methods for environmental research. In Wu, J., et Li, H. (2006). Concepts of scale and scaling. *Scaling and uncertainty analysis in ecology* (pp. 3-15). Springer Netherlands.
- Borgella, R., et Gavin, T. A. (2005). Avian community dynamics in a fragmented tropical landscape. *Ecological Applications*, 15(3), 1062-1073.
- Boscolo, D., et Metzger, J. P. (2009). Is bird incidence in Atlantic forest fragments influenced by landscape patterns at multiple scales?. *Landscape Ecology*, 24(7), 907-918.
- Brennan, J. M., Bender, D. J., Contreras, T. A., et Fahrig, L. (2002). Focal patch landscape studies for wildlife management: optimizing sampling effort across scales. *Integrating landscape ecology into natural resource management*, 68-91.
- Burke, I. C., Kittel, T. G. F., Laurenroth, W. K., Snook, P., Yonker, C. M., et Parton, W. J. (1991). Regional analysis of biomass, primary productivity, and net CO₂ flux. *Vegetation*, 80, 71-89.
- Buyantuyev, A., et Wu, J. (2007). Effects of thematic resolution on landscape pattern analysis. *Landscape Ecology*, 22(1), 7-13.
- CORINE Land Cover — Copernicus Land Monitoring Service (online), dostupné z <<http://land.copernicus.eu/pan-european/corine-land-cover>>.
- Crouzeilles, R., et Curran, M. (2016). Which landscape size best predicts the influence of forest cover on restoration success? A global meta-analysis on the scale of effect. *Journal of Applied Ecology*.
- Dale, M. R., Dixon, P., Fortin, M. J., Legendre, P., Myers, D. E., et Rosenberg, M. S. (2002). Conceptual and mathematical relationships among methods for spatial analysis. *Ecography*, 25(5), 558-577.
- Dark, S. J., et Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, 31(5), 471-479.
- Duckham, M., et Sharp, J. (2005). Uncertainty and geographic information: Computational and critical convergence. *Re-presenting GIS*. Wiley: Chichester, 113-124.
- Dungan, J. L., Perry, J. N., Dale, M. R. T., Legendre, P., Citron-Pousty, S., Fortin, M. J., Jakomulska, A., Miriti, M., et Rosenberg, M. (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25(5), 626-640.
- Elith, J., et Graham, C. H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32(1), 66-77.

- Feng, M., Sexton, J. O., Channan, S., et Townshend, J. R. (2016). A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. *International Journal of Digital Earth*, 9(2), 113-133.
- Fotheringham, A. S. (1989). Scale-independent spatial analysis. *Accuracy of spatial databases*, 221-228. In Hay, G. J., Marceau, D. J., Dube, P., et Bouchard, A. (2001). A multiscale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology*, 16(6), 471-490.
- Garzon-Lopez, C. X., Bastin, L., Foody, G. M., et Rocchini, D. (2016). A virtual species set for robust and reproducible species distribution modelling tests. *Data in brief*, 7, 476-479.
- Glc 30 Tianditu (online), dostupné z <<http://glc30.tianditu.com/Enbackground.html>>.
- Goodchild, M. F. (2011). Scale in GIS: An overview. *Geomorphology*, 130(1), 5-9.
- Goodchild, M. F., Yuan, M., et Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International journal of geographical information science*, 21(3), 239-260.
- Haklay, M., et Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, 7(4), 12-18.
- Hall, O., Hay, G. J., Bouchard, A., et Marceau, D. J. (2004). Detecting dominant landscape objects through multiple scales: An integration of object-specific methods and watershed segmentation. *Landscape Ecology*, 19(1), 59-76.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *science*, 342(6160), 850-853.
- Hay, G. J., Marceau, D. J., Dube, P., et Bouchard, A. (2001). A multiscale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology*, 16(6), 471-490.
- Hendriks, A. J., Willers, B. J., Lenders, H. J., et Leuven, R. S. (2009). Towards a coherent allometric framework for individual home ranges, key population patches and geographic ranges. *Ecography*, 32(6), 929-942.
- Holland, J. D., Bert, D. G., et Fahrig, L. (2004). Determining the spatial scale of species' response to habitat. *BioScience*, 54(3), 227-233.
- Holland, J. D., Fahrig, L., et Cappuccino, N. (2005). Fecundity determines the extinction threshold in a Canadian assemblage of longhorned beetles (Coleoptera: Cerambycidae). *Journal of Insect Conservation*, 9(2), 109-119.
- Hutchinson, G. E. (1965). *The ecological theater and the evolutionary play*. In Wiens, J. A. (1989). Spatial scaling in ecology. *Functional ecology*, 3(4), 385-397.
- Jackson, H. B., et Fahrig, L. (2012). What size is a biologically relevant landscape?. *Landscape ecology*, 27(7), 929-941.
- Jackson, H. B., et Fahrig, L. (2015). Are ecologists conducting research at the optimal scale?. *Global Ecology and Biogeography*, 24(1), 52-63.
- Jarvis, P. G. (1995). Scaling processes and problems. *Plant, Cell et Environment*, 18(10), 1079-1089.

Jelinski, D. E., et Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape ecology*, 11(3), 129-140.

Jenkins, D. G., Brescacin, C. R., Duxbury, C. V., Elliott, J. A., Evans, J. A., Grablow, K. R., Hillegass, M., Lyon, B. N., Metzger, G. A., Olandese, M. L., Pepe, D., Silvers, G. A., Suresch, H. N., Thompson, T. N., Trexler, Ch. M., Williams, G. E., Williams, N. C., et Williams, S. E. (2007). Does size matter for dispersal distance?. *Global Ecology and Biogeography*, 16(4), 415-425.

King, A. W. (1991). 19. Translating Models Across Scales in the Landscape. *Quantitative methods in landscape ecology: the analysis and interpretation of landscape heterogeneity*, 82, 479.

Lausch, A., Pause, M., Merbach, I., Zacharias, S., Doktor, D., Volk, M., et Seppelt, R. (2013). A new multiscale approach for monitoring vegetation using remote sensing-based indicators in laboratory, field, and landscape. *Environmental monitoring and assessment*, 185(2), 1215-1235.

Legendre, P., Dale, M. R., Fortin, M. J., Gurevitch, J., Hohn, M., et Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, 25(5), 601-615.

Lechner, A. M., Langford, W. T., Jones, S. D., Bekessy, S. A., et Gordon, A. (2012a). Investigating species–environment relationships at multiple scales: Differentiating between intrinsic scale and the modifiable areal unit problem. *Ecological Complexity*, 11, 91-102.

Lechner, A. M., Langford, W. T., Bekessy, S. A., et Jones, S. D. (2012b). Are landscape ecologists addressing uncertainty in their remote sensing data? *Landscape ecology*, 27(9), 1249-1261.

Lechner, A. M., Stein, A., Jones, S. D., et Ferwerda, J. G. (2009). Remote sensing of small and linear features: quantifying the effects of patch size and length, grid position and detectability on land cover mapping. *Remote Sensing of Environment*, 113(10), 2194-2204.

Leroy, B., Meynard, C. N., Bellard, C., et Courchamp, F. (2015). Virtualspecies, an R package to generate virtual species distributions. *Ecography*.

Lillesand, T., Kiefer, R. W., et Chipman, J. (2014). *Remote sensing and image interpretation*. In Dark, S. J., et Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, 31(5), 471-479.

Louzada, J., Lima, A. P., Matavelli, R., Zambaldi, L., et Barlow, J. (2010). Community structure of dung beetles in Amazonian savannas: role of fire disturbance, vegetation and landscape structure. *Landscape Ecology*, 25(4), 631-641.

Maier, J. A., Ver Hoef, J. M., McGuire, A. D., Bowyer, R. T., Saperstein, L., et Maier, H. A. (2005). Distribution and density of moose in relation to landscape characteristics: effects of scale. *Canadian Journal of Forest Research*, 35(9), 2233-2243.

Marceau, D. J., Howarth, P. J., et Gratton, D. J. (1994). Remote sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem. *Remote Sensing of environment*, 49(2), 93-104.

Martensen, A. C., Pimentel, R. G., et Metzger, J. P. (2008). Relative effects of fragment size and connectivity on bird community in the Atlantic Rain Forest: implications for conservation. *Biological Conservation*, 141(9), 2184-2192.

- Martensen, A. C., Ribeiro, M. C., Banks-Leite, C., Prado, P. I., et Metzger, J. P. (2012). Associations of forest cover, fragment area, and connectivity with neotropical understory bird species richness and abundance. *Conservation Biology*, 26(6), 1100-1111.
- Meynard, C. N., et Kaplan, D. M. (2012). The effect of a gradual response to the environment on species distribution modeling performance. *Ecography*, 35(6), 499-509.
- Meynard, C. N., et Kaplan, D. M. (2013). Using virtual species to study species distributions and model performance. *Journal of Biogeography*, 40(1), 1-8.
- Meynard, C. N., et Quinn, J. F. (2007). Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, 34(8), 1455-1469.
- Miguet, P., Jackson, H. B., Jackson, N. D., Martin, A. E., et Fahrig, L. (2015). What determines the spatial extent of landscape effects on species?. *Landscape Ecology*, 1-18.
- Mitchel, A. (2005). The ESRI Guide to GIS analysis, Volume 2: Spatial measurements and statistics. *ESRI Guide to GIS analysis*.
- Nelson, R. (2001). *Geologic analysis of naturally fractured reservoirs*. Gulf Professional Publishing.
- Norowi, H. M., Perry, J. N., Powell, W., et Rennolls, K. (2000). The effect of spatial scale on interactions between two weevils and their parasitoid. *Ecological Entomology*, 25(2), 188-196.
- Olea, R. A. 1990. Geostatistical glossary and multilingual dictionary. In Dungan, J. L., Perry, J. N., Dale, M. R. T., Legendre, P., Citron-Pousty, S., Fortin, M. J., ... et Rosenberg, M. (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25(5), 626-640.
- Openshaw, S. (1984a). Concepts and Techniques in Modern Geography Number 38. *The Modifiable Areal Unit Problem Geo Books Norwich*. In Dungan, J. L., Perry, J. N., Dale, M. R. T., Legendre, P., Citron-Pousty, S., Fortin, M. J., ... et Rosenberg, M. (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25(5), 626-640
- Openshaw, S. (1984b). Ecological fallacies and the analysis of areal census data. *Environment and Planning A*, 16(1), 17-31. In Dark, S. J., et Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, 31(5), 471-479.
- Openshaw, S., et Taylor, P. J. (1979). A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *Statistical applications in the spatial sciences*, 21, 127-144.
- Pearman, C. P., Adams, C. S., Cox, S. G., Griffin, P. F., Smith, D. A., et Hughes, I. G. (2002). Polarization spectroscopy of a closed atomic transition: applications to laser frequency locking. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 35(24), 5141.
- Pekkarinen, A., Reithmaier, L., et Strobl, P. (2009). Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(2), 171-183.

Regression analysis basics—ArcGIS Pro | ArcGIS Desktop (online), dostupné z <<http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/regression-analysis-basics.htm>>.

Ricci, B., Franck, P., Toubon, J. F., Bouvier, J. C., Sauphanor, B., et Lavigne, C. (2009). The influence of landscape on insect pest dynamics: a case study in southeastern France. *Landscape Ecology*, 24(3), 337-349.

Ricci, B., Franck, P., Valantin-Morison, M., Bohan, D. A., et Lavigne, C. (2013). Do species population parameters and landscape characteristics affect the relationship between local population abundance and surrounding habitat amount?. *Ecological Complexity*, 15, 62-70.

Rossi, R. E., Mulla, D. J., et Franz, E. H. Geostatistical tools for modeling and interpreting ecological spatial dependence¹. *Ecological Monographs*, 62(2).

Shanlong, L., Bingfang, W., et Fapeng, L. (2011). Wetland pattern change in Hai Basin. *Journal of Remote Sensing*, 15(2), 349-371.

Schindler, S., von Wehrden, H., Poirazidis, K., Wrbka, T., et Kati, V. (2013). Multiscale performance of landscape metrics as indicators of species richness of plants, insects and vertebrates. *Ecological Indicators*, 31, 41-48.

Stevens, V. M., Whitmee, S., Galliard, L., Clobert, J., Böhning-Gaese, K., Bonte, D., Brändle, M., Dehling, D. M., Hof, Ch., Trochet, A., et Baguette, M. (2014). A comparative analysis of dispersal syndromes in terrestrial and semi-terrestrial animals. *Ecology letters*, 17(8), 1039-1052.

Šímová, P., et Gdulová, K. (2012). Landscape indices behavior: a review of scale effects. *Applied Geography*, 34, 385-394.

Taki, H., Kevan, P. G., et Ascher, J. S. (2007). Landscape effects of forest loss in a pollination system. *Landscape Ecology*, 22(10), 1575-1587.

Thibaud, E., Petitpierre, B., Broennimann, O., Davison, A. C., et Guisan, A. (2014). Measuring the relative effect of factors affecting species distribution model predictions. *Methods in Ecology and Evolution*, 5(9), 947-955.

Thornton, D. H., et Fletcher, R. J. (2014). Body size and spatial scales in avian response to landscapes: a meta-analysis. *Ecography*, 37(5), 454-463.

Townsend, P. A. (2000). A quantitative fuzzy approach to assess mapped vegetation classifications for ecological applications. *Remote Sensing of Environment*, 72(3), 253-267.

Turner, M. G. (1989). Landscape ecology: the effect of pattern on process. *Annual review of ecology and systematics*. In Lechner, A. M., Langford, W. T., Jones, S. D., Bekessy, S. A., et Gordon, A. (2012). Investigating species–environment relationships at multiple scales: differentiating between intrinsic scale and the modifiable areal unit problem. *Ecological Complexity*, 11, 91-102.

Turner, M. G., Dale, V. H., et Gardner, R. H. (1989a). Predicting across scales: theory development and testing. *Landscape ecology*, 3(3-4), 245-252.

Turner, M. G., O'Neill, R. V., Gardner, R. H., et Milne, B. T. (1989b). Effects of changing spatial scale on the analysis of landscape pattern. *Landscape ecology*, 3(3-4), 153-162.

- Turner, S. J., O'NEILL, R. V., Conley, W., Conley, M. R., et Humphries, H. C. (1991). Patternandscale: statistics for landscape ecology. *Ecological studies*, 82, 17-49.
- Veregin, H. (1999). Data quality parameters. *Geographical information systems*, 1, 177-189.
- VÚV T.G.Masaryka - Oddělení GIS - O projektu DIBAVOD (online), dostupné z <<http://www.dibavod.cz>>.
- Wiens, J. A. (1989). Spatial scaling in ecology. *Functional ecology*, 3(4), 385-397.
- Wintle, B. A., Elith, J., et Potts, J. M. (2005). Fauna habitat modelling and mapping: a review and case study in the Lower Hunter Central Coast region of NSW. *Austral Ecology*, 30(7), 719-738.
- Wu, J. (1999). Hierarchy and scaling: extrapolating information along a scaling ladder. *Canadian journal of remote sensing*, 25(4), 367-380.
- Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape ecology*, 19(2), 125-138.
- Wu, J., Gao, W., & Tueller, P. T. (1997). Effects of changing spatial scale on the results of statistical analysis with landscape data: A case study. *Geographic Information Sciences*, 3(1-2), 30-41.
- Wu, J., Jelinski, D. E., Luck, M., et Tueller, P. T. (2000). Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics. *Geographic Information Sciences*, 6(1), 6-19.
- Wu, J., et Li, H. (2006). Concepts of scale and scaling. In *Scaling and uncertainty analysis in ecology* (pp. 3-15). Springer Netherlands.
- Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., ... et Grimm, V. (2010). The virtual ecologist approach: simulating data and observers. *Oikos*, 119(4), 622-635.

8. Seznam příloh

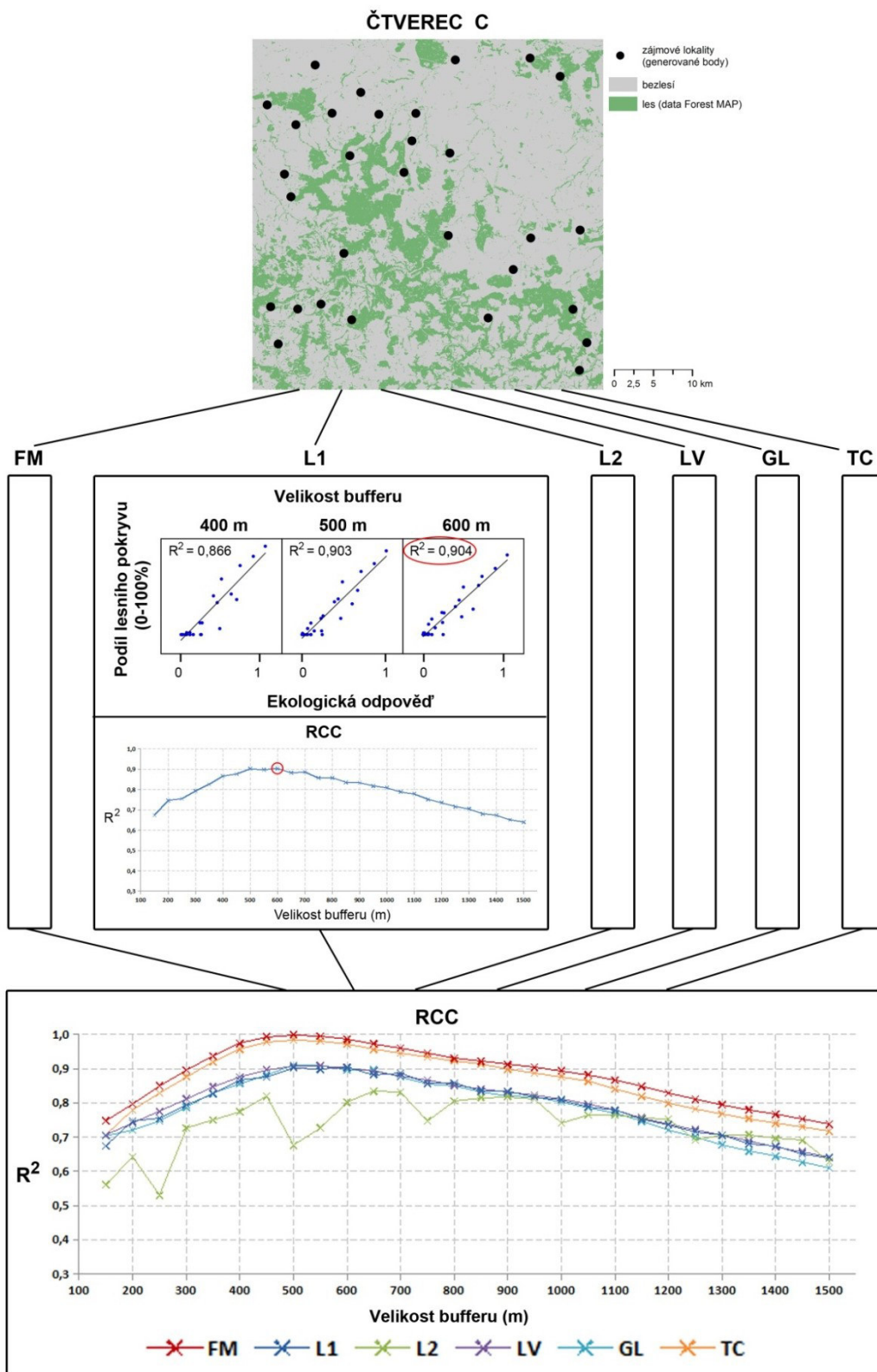
Obrazové přílohy:

Příloha 1 Schéma základních kroků experimentu (čtverec C při skutečném scale of effect 500 m).

Elektronické přílohy:

Příloha 2 Výsledné koeficienty analýzy OLS a grafy RCC pro data lesního pokryvu.

Příloha 3 Výsledné koeficienty analýzy OLS a grafy RCC pro data vodních ploch.



Příloha 1 Schéma základních kroků experimentu (čtverec C při skutečném scale of effect 500 m). Ekologická odpověď odvozená ve 30-ti lokalitách každého čtverce z dat Forest MAP (FM) je pomocí analýzy OLS porovnávána s ostatními datovými sadami v buffrech 150 m až 1500 m. Z výsledných R^2 pro každou datovou sadu je generována křivka RCC a identifikováno zdánlivé scale of effect (nejvyšší dosažené R^2 dat CORINE Land Cover rastr 100 m (L1) je zakroužkováno červeně). Jednotlivé křivky RCC jsou pak vzájemně porovnány.