

Posudek na diplomovou práci:

Tereza Pašková: Časování výletové aktivity netopýrů z úkrytů lokalizovaných v urbánním prostředí

Diplomová práce je přiměřená v rozsahu i formě, navíc s dobrou dokumentací použitých dat. Titul dobře odpovídá obsahu. Téma behaviorálních adaptací k urbánnímu prostředí, jmenovitě časování aktivity, je velmi aktuální.

K práci nemám zásadní výhrady týkající se celkové koncepce, zvoleného řešení, ani způsobu prezentace. Oceňuji úsilí nutné k získání vlastních dat, stejně jako invenci při vymýšlení netradičních metod vyhodnocení. Ke konkrétním postupům, výsledkům a jejich interpretaci mám však řadu dotazů, snad pramenících z nepochopení a z nejednoznačného popisu. Následující připomínky necht' jsou brány jako náměty k diskusi spíš než jako kritika odvedené práce.

Načasování výletu

Předesílám, že použité statistické metody považuji za zcela adekvátní účelu DP. Nicméně v práci postrádám výslovné vyjádření v tom smyslu, že veškeré analýzy jsou univariátní, tj. nezohledňují možné vlivy kovariát. Autorka si je tohoto problému vědoma, jak vidno např. z analýzy reprodukčních fází rozdělené po letech (Obr. 10). Dalším obecným problémem je pseudoreplikace, zmíněná a řešená v případě analýzy typu úkrytu (s. 27). Týká se to ovšem v různé míře všech analýz, přičemž jednoduché řešení v podobě agregace dat nelze vždy použít.

Problém mám s interpretací vztahu mezi časem výletu a intenzitou světla. Vztah mezi intenzitou světla a denní dobou je čistě deterministický. Rozptyl naměřených dat (Obr. 6) je výsledkem nezohlednění vlivu dalších proměnných, především zastínění a oblačnosti v čase a místě měření. Vysoce významný výsledek testu lineární regrese ($R^2 > 0,6$) tak říká pouze to, že později večer je větší tma.

Jisté rezervy vidím také v popisu použitých metod a způsobu prezentace výsledků, které jsou místy méně přehledné až matoucí. Zčásti je to důsledek neobvyklých výrazů a neobratných formulací. Např. čím se liší „všeobecná významnost jednotlivých proměnných“ a „celková statistická významnost“? Patrně je tím někde myšlen rozdíl mezi testem kategorického prediktoru a testem pro jeho jednotlivé hladiny. Asi proto lze nalézt částečné duplicity mezi testy ukázanými v tabulkách 3 a 4. Obojí lze přitom prezentovat přehledněji a informativněji ve společné tabulce – pro každý prediktor ukázat regresní koeficient(y)±SE (= směr a velikost efektu) a příslušný test.

Nesouhlasím s tvrzením, že „P-hodnoty jsou všeobecné zhodnocení významnosti proměnných“ (Tabulka 3). P-hodnota vyjadřuje pouze jeden z aspektů statistické významnosti proměnné. Proto je celkem zbytečné, až zavádějící, vynášet P-hodnoty do grafů (Obr. 4 a 17).

Nepochopil jsem rozdíl mezi proměnnou „změna teploty“ (3 hladiny, proto $df = 2$, Tabulka 3) a „příznak změny teploty“ ($df = 1$, Tabulka 4), což se týká i dalších proměnných tohoto typu.

Proč má rok v tabulce 4 pouze 1df (data jsou ze 3 let)?

Ne každý test založený na statistice χ^2 (např. testy prediktorů v logistické regresi) je testem dobré shody (ten je v práci použit pouze pro analýzu klastrovitosti).

Wilcoxonův test nelze použít pro post-hoc párová porovnání po Kruskal-Wallisově anově.

V analýze typu úkrytů není jasné, zda uváděné rozdíly v načasování výletu (s. 27) jsou rozdílem mezi mediány mediánů (Obr. 8), nebo rozdílem mezi průměry mediánů.

Vliv proměnných na pozitivní výletovou aktivitu

Obecné připomínky ke způsobu prezentace platí i zde. Především si ale na základě dostupného popisu nejsem jistý, zda je skutečně analyzována pravděpodobnost výletu (např. Obr. 15), což bylo zamýšleným cílem, nebo spíš obsazenost dutiny. Netopýři vylétli ve 112 případech z celkem 247 pozorování. Je jisté, že daná dutina byla v době pozorování opravdu obsazená, tj. že netopýři vylétnout mohli, ale nevylétli?

Chybové úsečky v obr. 15 neudávají rozsah naměřených hodnot (za předpokladu, že jde o binární logistickou regresi pracující s daty 0/1), ale patrně jde o konfidenční interval pro predikovanou pravděpodobnost.

Analýza klastrovitosti

Za největší výzvu považuji analýzu klastrovitosti počtu výletujících netopýrů. V současných analýzách tuším konceptuální problém. Předně, v čem by měla spočívat výhodnost přístupu, který není založen na časových intervalech (s. 18)? Biologicky relevantní je určitě nejen samotná klastrovitost, ale i distribuce klastrů v čase. Pokud už ale byla zvolená metoda použita, nedomnívám se, že je vhodné stanovit jako referenční úroveň nulovou klastrovitost = uniformní distribuci počtu výletujících jedinců (s. 18 a 19). Osobně bych tuto referenční klastrovitost definoval např. jako výsledek nezávislého a náhodného chování (vylétnutí) N jedinců během určitého časového intervalu, při zohlednění relevantních omezení (např. kapacity výletového otvoru, apod.). Výsledkem takového chování by určitě nebyla uniformní distribuce, ale nějaká „náhodná“ klastrovitost. Biologicky relevantní otázkou potom je, jak moc a jakým směrem se pozorovaná míra klastrovitosti liší od té náhodně vygenerované.

Ptám se, zda není pozitivní vztah mezi celkovým počtem výletujících netopýrů a mírou klastrovitosti (Obr. 18) nevyhnutelný? Očekával bych, že s počtem jedinců nutně roste šance, že někteří vylétí krátce po sobě, tedy vytvoří klastr. Smysluplnější otázka potom je, jak se liší sklon závislosti v pozorovaných datech a datech generovaných jako „náhodné“ výlety daného počtu jedinců.

Analýzu frekvence výletů jsem z popisu metod nepochopil a mám důvodné podezření, že není zcela v pořádku. Frekvence byla počítána jako „počet výletů za časový interval od posledního výletu“ (s. 19). První netopýr podle definice vylétí vždy v čase 0, další potom řekněme v čase 10 a další 2 jedinci např. v čase 100. Co je potom vyneseno v grafech (Obr. 23 a příloha 8.2)? Ať už je to cokoli, křivka lokální regrese v ukázaných grafech je zjevně počítána pouze z datových bodů odpovídajících (patrně) pozorovaným časům výletu. Obávám se, že tímto postupem jsou vynechána negativní pozorování (nulové frekvence v době když netopýři nelétali), což nutně ovlivní tvar závislosti na čase. Nehledě na výše uvedené jsou stávající grafy špatně už z toho důvodu, že predikované frekvence z principu nemohou nabývat negativních hodnot (viz příloha 8.2).

Domnívám se, že mnohem intuitivnější by bylo zobrazení distribuce počtu výletujících netopýrů v čase (jako obr. 3, ale se zahrnutím všech jedinců, nikoli jen prvního). Počty jedinců v jednotlivých časových intervalech stejné délky lze potom interpretovat jako frekvence výletů. Pokud netopýři vylétají ve více vlnách, projevilo by se to jako píky histogramu. Předpokládám, že při daném počtu jedinců by z tradičního histogramu nebylo nic moc vidět, což lze ale řešit grafem kernelovy hustoty.

Formální připomínky

Nelze spoléhat na automatické generátory citací; seznam použité literatury vždy vyžaduje kontrolu a doplnění. Jinak se může stát, že jako zdroj je uveden časopis „October“ (místo Journal of Zoology), jindy zdroj zcela chybí, aniž by bylo zřejmé, o co se vůbec jedná. U periodik je nedůsledně uváděno číslo v rámci svazku a rozsah stran. Opakovaně v textu citovaný zdroj „Bats in the Anthropocene 2016“ v seznamu chybí. Také odkazy v textu nejsou formátovány jednotně: v úvodu jsou odkazy v závorce řazeny podle autora, zatímco v diskuzi jsou řazeny chronologicky, jak bývá obvyklé; zkratka et al. není používána důsledně; neskloňovaná jména autorů působí v textu (mimo závorku) krkolomně.

Odstavec by neměl začínat formulací „Tato průměrná hodnota...“ (s. 47), neboť čtenář neví, které z mnoha hodnot probíraných v předchozím odstavci se to týká.

Pozor na neurčitá vyjádření, např. v prvním odstavci kap. 3.1: „Většinou“, „během některých večerů“, „do 10 případů ze všech pozorování“ ($n = ?$).

V odborném textu neradno zaměňovat termíny hypotéza a teorie.

Ne všechny popisky grafů a tabulek jsou samovysvětlující, zejména s ohledem na použité metody.

Podle označení (Tabulka 1–6) nejde rozlišit tabulky v textu od tabulek v příloze.

Tabulka 2 (v textu) nemá obvyklý formát ostatních tabulek.

Tabulka 5 = Tabulka 6 (obojí v příloze).

Závěr:

Nehledě k výše uvedeným připomínkám považuji DP za zdařilou, výsledky jistě poslouží jako cenný podklad pro další bádání. Autorka musela vynaložit značné úsilí na sběr dat a zvládnutí řady metod jejich zpracování a nesporně tak prokázala schopnost samostatné tvůrčí práce. Diplomovou práci doporučuji k obhajobě.