

# MODELI S PROSTORSKO OMEJITVIJO ZA OCENJEVANJE GOSTOT SUPERPOPULACIJ

## SPATIALLY EXPLICIT MODELING OF SUPERPOPULATION DENSITY

ROMAN LUŠTRIK

15. april 2013

### Kazalo

<b>1 Znanstveno področje na katerega se nanaša predlagana tema doktorske disertacije</b>	<b>1</b>
<b>2 Prikaz dosedanjih raziskovanj opredeljenega raziskovalnega problema</b>	<b>1</b>
<b>3 Raziskovalne hipoteze</b>	<b>3</b>
<b>4 Metode raziskovanja</b>	<b>3</b>
<b>5 Pričakovani rezultati in prispevek disertacije k razvoju znanosti</b>	<b>4</b>

### **1 Znanstveno področje na katerega se nanaša predlagana tema doktorske disertacije**

Naloga je interdisciplinarna in se nanaša na področje statistike in biologije.

### **2 Prikaz dosedanjih raziskovanj opredeljenega raziskovalnega problema**

Populacija je opredeljena kot skupina osebkov, ki pripadajo isti vrsti in naseljujejo skupni prostor v določenem času (Krebs, 2001). Populacije so umeščene v prostor, nanje pa vplivajo abiotski in biotski dejavniki. Za preučevanje in upravljanje populacij je eden ključnih parametrov njihova velikost. Velikost populacije pogosto ocenjujemo s pomočjo modelov ulova in ponovnega ulova (označimo kot  $\hat{N}$ ). Pri tej metodi osebek označimo in izpustimo, nato pa spremljamo njegovo ulovljivost (verjetnost, da bomo osebek zaznali) skozi (ne)odlov v odlovnih intervalih. (Williams *et al.*, 2002).

Za ocenjevanje velikosti populacij pogosto uporabljamo modele za zaprte populacije (angl. *closed population capture-recapture models*, Cooch & White (2012)), saj so praktični za uporabo. Da bodo cenilke parametrov za te modele nepristranske, morajo med vzorčenjem veljati predpostavke o a) prostorski in demografsko zaprti populaciji (angl. *closure*), b) enako verjetni ulovljivosti za vse osebkke in c), da so oznake osebkov zanesljive. Za naše potrebe bomo predpostavili, da je populacija zaprta. Da je vzorčena populacija v času vzorčenja kar se da zaprta, se v praksi vzorči v relativno kratkem časovnem obdobju na relativno velikem območju. Zaradi zanesljivosti določitve osebkov z molekularnimi metodami lahko predpostavimo zanesljivost oznak (Beja-Pereira *et al.*, 2009). Podrobneje bomo preučili predpostavko o enako verjetni ulovljivosti.

V praksi so vzorčenja navadno omejena finančno, časovno ali fizično na določeno območje (npr. mejo rezervata ali državno mejo), zaradi česar je populacija lahko relativno odprta. Poleg živali, ki so stalno v območju vzorčenja naletimo še na živali, ki del časa preživijo zunaj območja vzorčenja in tvorijo t.i. superpopulacijo. Ker območje vzorčenja ni dovolj veliko, da bi lahko zaobjeli celotno superpopulacijo, nekateri osebkki prehajajo rob območja, zato je kršena predpostavka o zaprtosti populacije. Drugače rečeno, vzorčeno območje in domači okoliš osebkov se lahko le delno prekrivata. Zaradi prehajanja roba je verjetnost ulovljivosti spremenjena, saj osebek ni vedno na voljo za ulov. Kendall (1999) je pokazal, da je za populacije, ki prehajajo rob vzorčenega območja naključno,  $E(\hat{p}_i) = \tau_i p_i$ , kjer je  $\tau_i$  verjetnost, da vzorčimo osebek  $i$  iz superpopulacije v danem odlovnem intervalu in  $p_i$  ulovljivost. Ocena gostote populacije bo nepristranska, če bodo imeli vsi osebkki v vzorčenem območju  $\tau = 1$ . Če imajo nekateri osebkki  $\tau < 1$ , bo ocena pristranska. Pristranskost parametrov zaradi prehajanja meje vzorčenega območja, kjer je  $\tau < 1$ , ima za posledico t.i. *učinek roba*.

Učinek roba je že dolgo proučevan problem (Efford, 2004), ki pa do danes še nima zadovoljive rešitve. Prvič se v literaturi z njim ukvarja Dice (1938, 1941), ki predlaga, da se vzorčeno območje poveča za polmer domačega okoliša. Dosedanje raziskave s področja metod ulova in ponovnega ulova se osredotočajo predvsem na vzorčenje s pomočjo pasti na mreži ali situ (pregled teh metod najdemo v Williams *et al.* (2002), poglavje 14). Dela, ki se ukvarjajo s popravljanjem učinka roba, navadno uporabljajo vzorčenje na situ (Parmenter *et al.*, 2003).

Prvi prispevek, ki predpostavlja heterogeno ulovljivost posameznega osebkka, so objavili Otis *et al.* (1978), vendar so že v tem delu predvideni zapleti, ker je potrebno ocenjevati veliko število parametrov. Število parametrov lahko zmanjšamo tako, da predpostavimo nekaj "tipov" osebkov. Na primer, take, ki so stalno prisotni v vzorčenem območju in take, ki občasno prehajajo rob vzorčenega območja. Miller *et al.* (2005) predstavijo model, ki po naravi ni prostorski, izrabi pa dejstvo, da je bilo v enem odlovnem intervalu za posamezen osebek zabeleženih več vzorcev. S heterogenostjo ulovljivosti se ukvarjajo še Burnham & Overton (1978), Chao (1988) in Pledger (2005).

Z razvojem genetike in GPS tehnologije je osebek mogoče "odloviti" tudi s pomočjo neinvazivnih genetskih vzorcev, za kar so razvili modele, ki vključujejo prostorsko komponento (angl. *spatially explicit models*). Tako dobimo (grobo) predstavo o gibanju živali (lahko samo z enim odlovnim intervalom). Ker modeli izrabljajo prostorsko informacijo je mogoče oceniti velikost območja, kjer se nahaja superpopulacija, kar je ključen podatek za zanesljivo računanje popula-

cijske gostote (Gardner *et al.* , 2009; Royle *et al.* , 2011; Efford, 2011; Borchers, 2012; Royle *et al.* , 2013). Od teh sta metodi avtorjev Royle *et al.* (2011) in Efford (2011) primerni za podatke, ki niso vezani na odlov s pomočjo pasti, kar bo pomembno za naše delo.

### 3 Raziskovalne hipoteze

- Pričakujemo, da bomo z upoštevanjem prostorske komponente (kot individualne kovariate, ki do neke mere opiše gibanje osebkov) izboljšali oceni velikosti in gostote populacije. Z upoštevanjem dodatne informacije o gibanju bomo izboljšali oceno ulovljivosti, saj bomo vključili eno od pomembnih komponent, ki vplivajo na zaznavnost ( $\tau$ ) in ulovljivost ( $p$ ) osebkov. Gostoto bomo lahko ocenili z boljšim poznavanjem velikosti prispevnega območja superpopulacije, pri čemer nam bo pomagalo razumevanje gibanja osebkov.
- Pričakujemo, da bo razlika med ocenjeno in dejansko gostoto odvisna od števila (oz. gostote) simuliranih živali. Pri nizki gostoti bo učinek roba zanemarljiv, s povečevanjem gostote se bo odstopanje ocenjene gostote od prave vrednosti povečevalo. Ker cenilka predpostavlja enako ulovljivost ( $\hat{N} = \frac{1}{\hat{p}_i}$ ), bo zaradi večjega števila napačno ocenjenih ulovljivosti ocena velikosti populacije pristranska.

### 4 Metode raziskovanja

Denimo, da imamo za vsak ujet osebek v odlovnem intervalu vsaj en podatek o lokaciji odlova. Več podatkov kot imamo, bolj lahko opišemo velikost in gibanje osebkov v domačem okolišu. Informacijo o lokaciji domačega okoliša in njegove velikosti lahko izkoristimo za generiranje nove statistike. Ker imamo navadno premalo točk, da bi lahko zadovoljivo opisali gibanje osebkov znotraj domačega okoliša, imamo dve možnosti.

Prva je, da na podlagi parametričnega modela predpostavimo verjetnost pojavljanja osebkov na določeni točki znotraj domačega okoliša. V ta namen lahko, med drugimi, uporabimo pol-normalno ali inverzno-normalno porazdelitev (bivariatna normalna porazdelitev). Odstotek prekrivanja domačega okoliša (glede na funkcijo, ki opisuje verjetnost pojavljanja v domačem okolišu) je predmet preučevanja (individualna kovariata).

Druga možnost je neparametrična. Na podlagi razdalj parov točk osebkov, ki imajo vsaj dve vzorčni točki, izračunamo empirično porazdelitveno funkcijo razdalj. To funkcijo uporabimo za računanje verjetnosti nahajanja osebkov v dani točki domačega okoliša. Vzorčne točke za izračun razdalj lahko dobimo s pomočjo drugih metod, npr. GPS-GSM telemetrijo (po potrebi točke redčimo za zmanjševanje avtokorelacije (Royle *et al.* , 2013)).

Odstotek prekrivanja domačega okoliša osebkov z vzorčenim območjem pove nekaj o prispevku h gostoti osebkov v območju vzorčenja in o potencialni ulovljivosti. To lahko uporabimo kot individualno kovariato v katerem izmed modelov ulova in ponovnega ulova, ki temelji na generaliziranih linearnih modelih (npr. Hugginsov model, Schwartz-Arnasonova modifikacija Jolly-Seberjevega modela ipd).

Izračunali bomo različne informacije, ki opisujejo gibanje v domačem okolišu (uporaba različnih predpostavk - krivulj) in preverili njihovo ustreznost kot pokazatelje modifikatorje ulovljivosti tako, da jih bomo vključili v modele, ki primarno niso prostorski. Za primerjavo bomo analizirali podatke še z modeli, ki vključujejo prostorsko komponentno oz. predpostavijo nekaj različnih skupin, ki imajo enako ulovljivost.

S simulacijami bomo preverili, ali predlagane izračunane kovariate izboljšajo ocene velikosti superpopulacije. Na večjem območju bomo generirali osebkke in jih vzorčili znotraj manjšega območja. Vsi osebki, ki kadarkoli pridejo v stik z vzorčenim območjem pripadajo superpopulaciji. V odlovnih intervalih bomo vzorčili po eno točko. Ker bomo kontrolirali verjetnost ulovljivosti in poznali število osebkov, ki jih je v danem trenutku mogoče vzorčiti, bomo lahko preverili zanesljivost naše ocene.

Podatke odlovne zgodovine bomo analizirali s pomočjo programa MARK (oz. R paketa `RMark`). Preverili bomo, če bo model, kjer vključimo individualno kovariato, boljši od tistega, kjer individualne kovariate ne vključimo (in predpostavljamo homogeno ulovljivost). Modele bomo primerjali s pomočjo popravljenega Akaikevega informacijske kriterija (AICc) (Akaike, 1974; Burnham & Anderson, 2002). Drugi pokazatelj bo odstopanje ocenjene gostote od simulirane. S simulacijami bomo poskušali zaobjeti čim večji parametrski prostor. Simulacije bodo sprogramirane v programskem jeziku R (R Core Team, 2012), in bodo po potrebi poganjane na visoko zmogljivem računalniškem sestavu.

## 5 Pričakovani rezultati in prispevek disertacije k razvoju znanosti

Prispevali bomo k boljšemu poznavanju problema učinka roba na ocenjevanje velikosti populacij in njihovih gostot. Ker je učinek roba že dolgo proučevan pojav, ki še ni bil zadovoljivo rešen, pričakujemo, da bomo lahko rezultate objavili v vplivnejših znanstvenih revijah s področja biologije.

S pomočjo pravilnejših ocen gostot bodo raziskovalci, upravljalci parkov in lovišč dobili boljši vpogled v populacijo in s tem imeli možnost izboljšanja upravljanja. Rezultati bodo imeli praktičen pomen v ekologiji in varstveni genetiki.

## Literatura

- Akaike, H. 1974. A new look at the statistical model identification. *Automatic control, IEEE transactions*, **19**, 716–723.
- Beja-Pereira, A., Oliveira, R., Alves, P. C., Schwartz, M. K., & Luikart, G. 2009. Advancing ecological understanding through technological transformations in noninvasive genetics. *Molecular ecology resources*, **9**, 1279–1301.
- Borchers, D. 2012. A non-technical overview of spatially explicit capture-recapture. *Journal of ornithology*, **152**, S435–S444.
- Burnham, K. P., & Anderson, D. R. 2002. *Model selection and multimodel inference: A practical information-theoretic approach (2nd ed.)*. Springer-Verlag New York, Inc.
- Burnham, K. P., & Overton, W. S. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika*, **65**(3), 625–633.
- Chao, A. 1988. Estimating animal abundance with capture frequency data. *Journal of wildlife management*, **52**(2), 295–300.
- Cooch, E., & White, G. 2012. *Program MARK: A gentle introduction (11th edition)*. 11 edn.
- Dice, L. R. 1938. Some census methods for mammals. *The journal of wildlife management*, **2**(3), 119–130.
- Dice, L. R. 1941. Methods for estimating populations of mammals. *The journal of wildlife management*, **5**(4), 398–407.
- Efford, M. 2004. Density estimation in live-trapping studies. *Oikos*, **106**, 598–610.
- Efford, M. G. 2011. Estimation of population density by spatially explicit capture-recapture analysis of data from area searches. *Ecology*, **92**(12), 2202–2207.
- Gardner, B., Royle, J. A., & Wegan, M. T. 2009. Hierarchical models for estimating density from DNA mark-recapture studies. *Ecology*, **90**(4), 1106–1115.
- Kendall, W. L. 1999. Robustness of closed capture-recapture methods to violations of the closure assumption. *Ecology*, **80**(8), 2517–2525.
- Krebs, C. J. 2001. *Ecology: the experimental analysis of distribution and abundance (5th ed.)*. Benjamin Cummings (California, USA).
- Miller, C. R., Joyce, P., & Waits, L. 2005. A new method for estimating the size of small populations from genetic mark-recapture data. *Molecular ecology*, **14**, 1991–2005.
- Otis, D. L., Burnham, K. P., White, G.C., & Anderson, D. R. 1978. Statistical inference from capture data on closed animal populations. *Wildlife monographs*, **62**, 3–135.

- Parmenter, R. R., Yates, T. L., Anderson, D. R., Burnham, K. P., Dunnun, J. L., Frankin, A. B., Friggens, M. T., Lubow, B. C., Miller, M., Olson, G. S., Parmenter, C. A., Pollard, J., Rextad, E., Shenk, T. M., Stanley, T. R., & White, G. C. 2003. Small-mammal density estimation: A field comparison of grid-based vs. web-based density estimators. *Ecological monographs*, **73**(1), 1–26.
- Pledger, S. 2005. The performance of mixture models in heterogeneous closed population capture-recapture. *Biometrics*, **61**, 868–876.
- R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Royle, A. J., Chandler, R. B., Sun, C. C., & Fuller, A. K. 2013. Integrating resource selection information with spatial capture-recapture. *Methods in ecology and evolution*, 1–42.
- Royle, J. A., Kéry, M., & Guélat, J. 2011. Spatial capture-recapture models for search-encounter data. *Methods in ecology and evolution*, 1–10.
- Williams, B. K., Nichols, J. D., & Conroy, M. J. 2002. *Analysis and Management of Animal Populations: Modeling, Estimation and Decision Making*. Academic press.