Count data collected using a robust design: models, results and recommendations

Eleni Matechou, Kevin Winner, Rebecca Hutchinson, Jeffrey Pippen & Leslie Ries

ISEC

Seattle, Wednesday, 29/06/16

OUTLINE

▸ Workshop

Motivating data

Modelling

Results- Real data

Results- Simulations

Summary

Future work

MOTIVATION

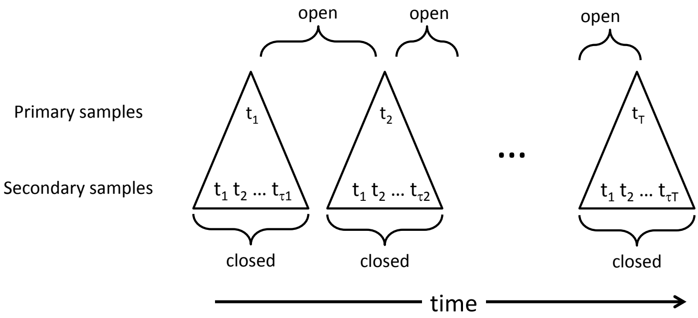
We were motivated by data on butterflies collected at the MPG ranch. ▸ MPG ranch

"Set in the heart of Montana's Bitterroot Valley, MPG Ranch lies on over 9,500 acres of rich undeveloped landscape. Established in 2009, MPG strives to preserve the natural communities that make this area beautiful and focuses on research to restore and protect native diversity."

DATA - ROBUST DESIGN

Sites at the ranch were visited and transects were walked *three times*, instead of once which is the case in the standard design (SD), during each visit.

Result: count data collected under the robust design (RD).[1,2]



[1] Pollock, K. H. (1982). A capture-recapture design robust to unequal probability of capture. *The Journal of Wildlife Management* **46,** 752-757

[2] McCrea, R. S., & Morgan, B. J. T. (2014). Analysis of capture-recapture data. CRC Press.

Questions

We are interested in

- estimating the number of butterflies at each site and
- modelling phenology[3] at each site.

[3] Defined by the Oxford dictionary as "The study of cyclic and seasonal natural phenomena, especially in relation to climate and plant and animal life."

MODEL

We extend the SD Matechou et al. (2014)[4] model to account for the RD.

The model is a Jolly-Seber[5,6] type, motivated by more recent work in stopover capture-recapture models [7,8], but does not require uniquely identifiable individuals.

[4] Matechou, E., Dennis, E. B., Freeman, S. N. , Brereton, T. (2014). Modelling abundance and phenology in (multivoline) butterfly species: a novel mixture model. *JAE* **51,** 766-775

[5] Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration – Stochastic model. *Biometrika* **52,** 225-247.

[6] Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika* **52,** 249-259.

[7] Schwarz, C. J. and Arnason, A. N. (1996). A general methodology for the analysis of open-model capture recapture experiments. *Biometrics* **52,** 860-873.

[8] Pledger S., Efford M., Pollock K., Collazo J. A., Lyons, J. E. (2009). Stopover duration analysis with departure probability dependent on unknown time since arrival. In: DL Thompson, EG Cooch, MJ Conroy (eds) Modeling Demographic Processes in Marked Populations. Springer. pp. 349-363.

MODEL

The model provides estimates of

- N: "super-population" size (site-specific)
- β: entry parameters. They describe the emergence pattern, which is modelled using the pdf of G normal distribution(s), where G is the number of broods of the species. Mean arrival time can be modelled as function of (site) covariates.
- ϕ: survival probability

while taking into account imperfect detection, i.e. p (which can vary according to site and time specific covariates).

SMALL WOOD-NYMPH

$S = 20$ sites were visited on $T = 21$ consecutive weeks and were each walked $K_t = 3$ times for $t = 1, \ldots, T$ resulting in $20 \times 21 \times 3$ counts in total, although 702 counts (i.e. around 50%) are actually missing.

The species is univoltine so we set $G = 1$

The sites are located on different levels of elevation and hence we model the site mean emergence time of the species as a function of site elevation: $\mu_s = \alpha_0 + \alpha_1 x_s$ where x_s is the elevation of site s.

Small wood-nymph

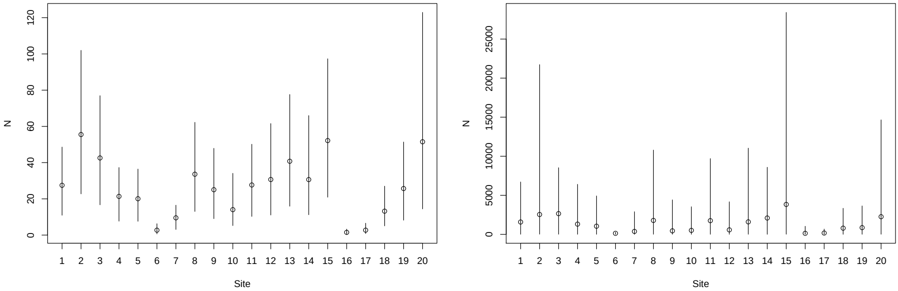
We model detection probability as a function of weather covariates

Model selection results for the small wood-nymph data collected in 2014. The models only differ in the constraints used for detection probability.

Model for p	ν	AIC
1. $\text{logit}(p_{stk}) = \gamma_0$	25	857.5
2. $\text{logit}(p_{stk}) = \gamma_0 + \gamma_1 z_{stk}$	26	860.5
3. $\text{logit}(p_{stk}) = \gamma_0 + \gamma_1 w_{stk}$	26	850.5
4. $\text{logit}(p_{stk}) = \gamma_0 + \gamma_1 z_{stk} + \gamma_2 w_{stk}$	27	849.8

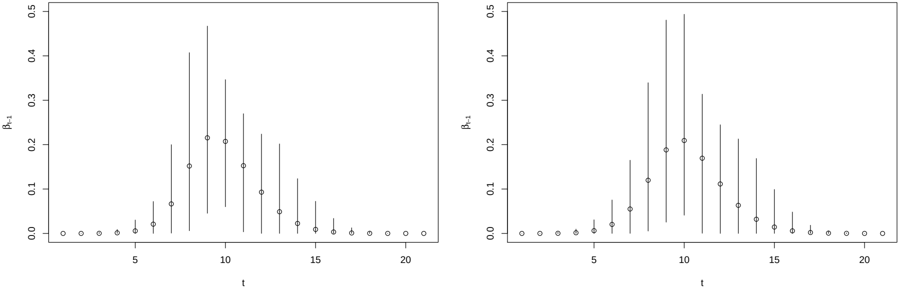
where z is temperature and w is % of cloudless sky.

SMALL WOOD-NYMPH



95% non-parametric bootstrap intervals obtained for $N_1, \ldots, N_{20}$
left: our proposed RD model
right: the Matechou et al. (2014) model keeping one (out of three) counts at random.

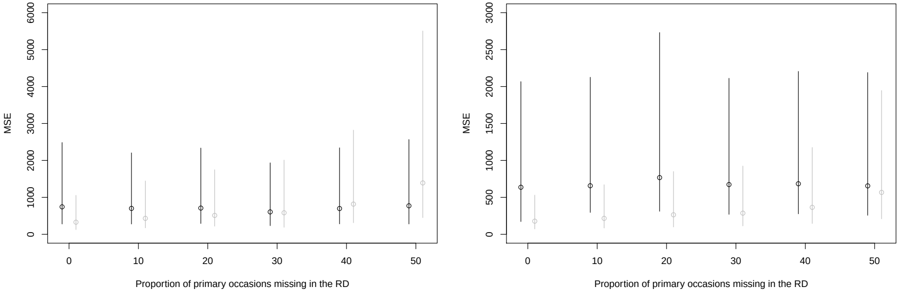
Small wood-nymph



95% non-parametric bootstrap intervals obtained for $\beta_0, \ldots, \beta_{19}$
left: site with lowest elevation
right: site with highest elevation.

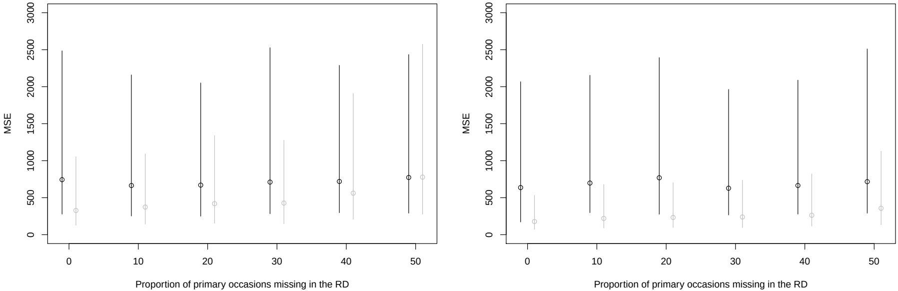
Estimated weekly survival: 0.52 (0.033, 0.696).

CAN THE RD COMPENSATE FOR MISSING PRIMARY OCCASIONS?



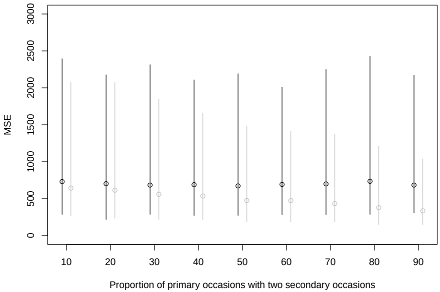
MSE obtained for N for the SD(black) and the RD(gray)
left: p only varies between primary occasions
right: p also varies between secondary occasions.

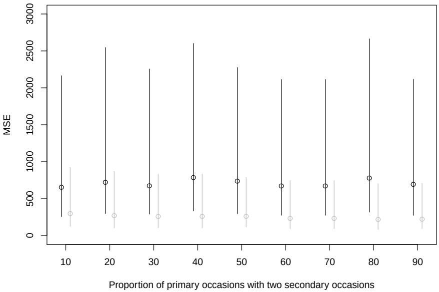
What if the missing primary occasions are mostly at the start and end of the season?



MSE obtained for N for the SD(black) and the RD(gray)
left: p only varies between primary occasions
right: p also varies between secondary occasions.

DO WE NEED TO USE THE RD FOR ALL SITES?





MSE obtained for N for the SD(black) and the RD(gray)
left: p only varies between primary occasions
right: p also varies between secondary occasions.

- The RD can improve accuracy of population size estimates. Other parameters are more robust to the sampling scheme used.
- This improvement can be substantial if detection probability depends on a covariate that varies between both primary and secondary occasions.
- The RD can be used to compensate for missing primary occasions. Example: fewer visits to the site but more counts obtained once there.
- This is especially true, with the RD performing better with fewer counts than the SD, if the covariate for p varies between both primary and secondary occasions and the missing counts are mostly concentrated at the start and end of season.

FUTURE WORK

More simulations! For example, answering questions such as: if one can only go to the site $x\%$ of the weeks, how many counts should they obtain once there to compensate?

Do you know of any schemes where several counts are obtained during each visit eg. double-observer schemes-or could easily be obtained but are not at the moment?

Thank you!