

Motivation/Introduction

Capture-Recapture data is widely used by statisticians and biologists to understand the demographics of animal populations. For example, they often want to know how survival rates are affected by time and environmental change for the purpose of conservation.

The Cormack-Jolly-Seber Model

The Cormack-Jolly-Seber (CJS) model for open capture-recapture data allows estimation of;

- ϕ_i , the **apparent** survival probability between occasions i and $i + 1$,
- p_i , the probability of **recapture** in occasion i

To turn capture histories into a likelihood for inference, an M-array of sufficient statistics is constructed.

Number Released	Number Recaptured				Never Recaptured
R_1	$m_{1,2}$	$m_{1,3}$	$m_{1,4}$	$m_{1,5}$	$m_{1,0}$
R_2		$m_{2,3}$	$m_{2,4}$	$m_{2,5}$	$m_{2,0}$
R_3			$m_{3,4}$	$m_{3,5}$	$m_{3,0}$
R_4				$m_{4,5}$	$m_{4,0}$

Table 1: M-array for 5 capture occasions

We have;

- R_i animals captured in occasion i , marked and then released,
- $m_{i,j}$ animals released in occasion i and **first recaptured** in occasion j ,
- $m_{i,0}$ animals released in occasion i and never recaptured

Each row in the m-array is modelled with a multinomial distribution so that;

$$\mathbf{m}_i \sim \text{Multi}(R_i; \mathbf{q}_i),$$

where \mathbf{q}_i is a vector of probabilities that an animal belongs to each $m_{i,j}$. For example;

- $m_{1,4}$ requires survival from occasions 1 to 4, being uncaptured in occasions 2 to 3 and capture in occasion 4. Hence, $q_{1,4} = \phi_1(1 - p_2)\phi_2(1 - p_3)\phi_3p_4$.

A likelihood function can then be formed from these distributions.

Random Effects and Covariates

We can extend the CJS model by including environmental covariates \mathbf{x}_i^T (e.g. minimum winter temperature) from the time of our study and a random effects term ϵ_i . We thus model;

$$\log\left(\frac{\phi_i}{1 - \phi_i}\right) = \mathbf{x}_i^T \beta + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma^2)$, for each i . Including these terms lets us directly model the natural variation in apparent survival over time as well as the effect of environmental change.

To estimate our parameters we now use MCMC as finding the MLEs would require integrating over the random effects.

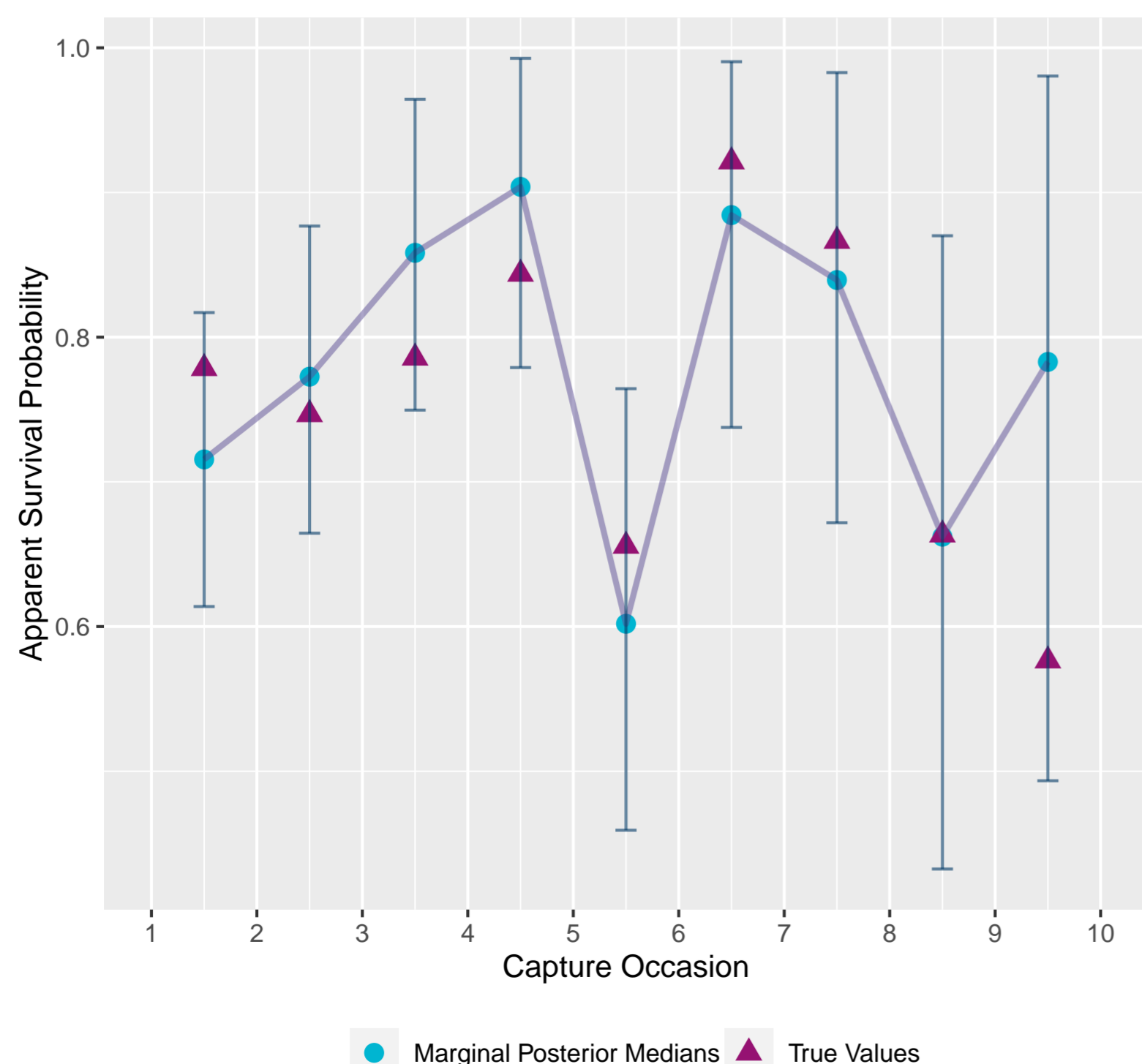


Figure 1: True values and 95% credible intervals for ϕ_i from random effects MCMC

Random Effects Example

We simulated 10 capture occasions under parameters $N = 300, p = 0.5$ (for all occasions), a mean apparent survival rate $\bar{\phi} = 0.8$ ($\beta_0 = 1.38$ on the logit scale) and random variation ϵ_i with $\sigma^2 = 0.5$.

After a 5000 sample burn-in, 95,000 posterior draws were made using WinBUGS with a thinning rate of 5 and uninformative priors for all parameters.

- Figure 1 shows the sample median and 95% credible intervals for ϕ_i with their true values,
- The MCMC sample medians for the other parameters were $\hat{\beta}_0 = 1.402$ (0.68, 2.6), $\hat{\sigma}^2 = 0.867$ (0.068, 5.94), and $\hat{p} = 0.467$ (0.42, 0.51).

Multi-Species Modelling

It may be preferable and ecologically reasonable to simultaneously analyse more than one species in the same spatial region. For multiple species, we can model apparent survival rates using the model of J. J. Lahoz-Monfort et al. (2011);

$$\log\left(\frac{\phi_{i,s}}{1 - \phi_{i,s}}\right) = f_s(\mathbf{x}_{i,s}^T) + \delta_i + \epsilon_{i,s}$$

Where;

- $\delta_i \sim N(0, \sigma_\delta^2)$ is a random effect for **all species** in the region,
- $\epsilon_{i,s} \sim N(0, \sigma_s^2)$ is a random effect exclusive to species s ,
- $f_s(\mathbf{x}_{i,s}^T)$ describes the covariate response for species s .

If variation in yearly survival is intrinsic to the environment, we would expect synchrony in species survival which is accounted for by δ_i . To quantify species synchrony we can use the **Intra-Class Coefficient** given by;

$$\text{ICC}_s = \frac{\hat{\sigma}_\delta^2}{\hat{\sigma}_\delta^2 + \hat{\sigma}_s^2}$$

for each species. The ICC describes the synchrony of species s by the proportion of its total variance present in all species.

We can determine if the environmental covariates synchronise the species by fitting the random effects model with and without covariates. We denote;

- $\hat{\sigma}_s^2(\text{res})$, the residual variance from the covariate model,
- $\hat{\sigma}_s^2(\text{total})$, the total variance from the intercept only model.

Hence, for each species (and for $\hat{\sigma}_\delta^2$) we define;

$$C_s = 1 - \frac{\hat{\sigma}_s^2(\text{res})}{\hat{\sigma}_s^2(\text{total})},$$

If the covariates explain the synchrony, we expect $\hat{\sigma}_s^2(\text{res})$ to be small relative to $\hat{\sigma}_s^2(\text{total})$ and thus C_s to be close to 1.

Further Applications and Limitations

This multiple random effects approach can be extended for more complex scenarios.

- Ghislain et al. (2024) implemented up to 4 random effects to account for variance in species, migration strategy, sampling location and individuals.
- J. Lahoz-Monfort et al. (2017) incorporated multiple random effects models into a single Integrated Population Model (IPM).

Overall, variance decomposition with random effects provides good models and insight into the scale of multi-species synchrony. However, in this form results should only serve as a guideline on potential causes of synchrony in ecosystems.

References

- Ghislain, M., Bonnet, T., Godeau, U., Dehorter, O., Gimenez, O., & Henry, P.-Y. (2024). Synchrony in adult survival is remarkably strong among common temperate songbirds across France. *Ecology*, 105(6), e4305.
- Lahoz-Monfort, J., Harris, M., Wanless, S., Freeman, S., & Morgan, B. (2017, 04). Bringing it all together: Multi-species integrated population modelling of a breeding community. *Journal of Agricultural, Biological and Environmental Statistics*, 22.
- Lahoz-Monfort, J. J., Morgan, B. J. T., Harris, M. P., Wanless, S., & Freeman, S. N. (2011). A capture-recapture model for exploring multi-species synchrony in survival. *Methods in Ecology and Evolution*, 2(1), 116-124.