



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Capture-Recapture Abundance Estimation using a Semi-complete Data Likelihood Approach

Citation for published version:

King, R, McClintock, BT, Kidney, D & Borchers, DL 2016, 'Capture-Recapture Abundance Estimation using a Semi-complete Data Likelihood Approach', *Annals of Applied Statistics*, vol. 10, no. 1, pp. 264-285.
<https://doi.org/10.1214/15-AOAS890>

Digital Object Identifier (DOI):

[10.1214/15-AOAS890](https://doi.org/10.1214/15-AOAS890)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Annals of Applied Statistics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



1 SUPPLEMENT TO “CAPTURE-RECAPTURE ABUNDANCE ESTIMATION
2 USING A SEMI-COMPLETE DATA LIKELIHOOD APPROACH”

3 BY RUTH KING*, BRETT T. MCCLINTOCK†, DARREN KIDNEY‡ AND DAVID BORCHERS‡

4 University of Edinburgh*, NOAA National Marine Mammal Laboratory† and University of St
5 Andrews‡

6 E-MAIL: Ruth.King@ed.ac.uk

7 In this supplement we provide sample JAGS code for model M_h considered in Section 4.1 (Ap-
8 pendix A) and the SECR model considered in Section 4.2 (Appendix B). For each example we
9 provide the model specification component of the JAGS code for the four different model-fitting
10 algorithms: (i) semi-complete data likelihood specifying Jeffreys’ prior on N (SCD1) and alter-
11 native Poisson-Gamma and Beta-Binomial prior specifications; (ii) semi-complete data likelihood
12 specifying the posterior conditional distribution on $N - n$, induced by Jeffreys’ prior on N (SCD2);
13 (iii) super-population complete data likelihood approach of [Royle *et al.* \(2007\)](#) (CD:R) and (iv)
14 super-population complete data likelihood approach of [Durban and Elston \(2005\)](#) (CD:DE).

APPENDIX A: JAGS CODE FOR MODEL M_H

15 In this appendix we provide sample JAGS code for model M_h considered in Section 4.1.

16 **A.1. First semi-complete data likelihood approach - SCD1.** The model component of
17 the JAGS code is provided here for the semi-complete data likelihood approach using the Jeffreys’
18 prior specification for N .

```

19 model{
20   Pi <- 3.14159265359
21   # Priors:
22   alpha ~ dnorm(0.0,0.01)
23   tau ~ dgamma(0.01,0.01)
24   sigma <- 1/sqrt(tau)
25
26   for (i in 1:n) {
27     y[i] ~ dbin(p[i],T)
28     logitp[i] ~ dnorm(alpha,tau)
29     logit(p[i]) <- logitp[i]
30   }
31
32   # Calculate probability of not being observed using Gauss-Hermite quadrature
33   # q = number of quadrature points
34   # weights and nodes correspond to q quadrature points; entered as data
35
36   for(i in 1:q){
37     probi[i] <- 1/sqrt(Pi)*weights[i]*(1/(1+exp(sqrt(2)*sigma*nodes[i]+alpha)))^T
38   }
39   prob<- sum(probi[])
40
41   # Prior for N: Jeffreys’ prior - this is incorporated in the zero trick below
42   # in specifying the likelihood term

```

```

43 # However a prior distribution is needed to be specified on N
44 # Use a discrete Uniform prior so the only influence on the posterior
45 # distribution is the upper limit
46
47 n00 ~ dcat(prior[]) # prior = rep(1/(M+1-n),M+1-n); entered as data
48 n0 <- n00 - 1
49 N <- n + n0
50
51 # Use zero trick for model likelihood
52 # Note loggam(N) instead of loggam(N+1) because of Jeffreys' prior for N
53
54 logzeroprob <- loggam(N) - loggam(n0+1) - loggam(n+1) + n0*log(prob)
55 lambda <- -logzeroprob + 100000
56 dummy ~ dpois(lambda) # dummy = 0; entered as data
57 }

```

We also provide the associated JAGS code for two alternative prior specifications on N . In particular we consider (i) $N|\mu \sim Po(\mu)$ with $\mu \sim \Gamma(a, b)$, which is equivalent to $N \sim Negative - Binomial\left(a, \frac{b}{b+1}\right)$ and (ii) $N|\psi \sim Binomial(M, \psi)$ where $\psi \sim Beta(a, b)$, equivalent to $N \sim Beta - Binomial(M, a, b)$. To use these priors, replace the line defining `logzeroprob` with:

```

62 # For Negative-Binomial(a,b/(b+1)) prior on N:
63 # a and b read in as data - for approximate Jeffreys' prior a = b = 0.000001
64
65 p1 <- b/(1+b)
66 logzeroprob <- - loggam(n0+1) - loggam(n+1) + n0*log(prob) +
67               loggam(N+a+1) - loggam(a+1) + a*log(p1) + N*log(1-p1)
68
69 # For Beta-Binomial(M,a,b) prior on N:
70 # a and b read in as data - for approximate Jeffreys' prior a = 0.001, b = 1
71
72 logzeroprob <- - loggam(n0+1) - loggam(n+1) + n0*log(prob) + loggam(M+1) -
73               loggam(M-N+1) + loggam(N+a) + loggam(M-N+b) - loggam(M+a+b) +
74               loggam(a+b) - loggam(a) - loggam(b)

```

Note: using the Negative-Binomial model specification achieved a better ESS and ESS/s than the equivalent Poisson-Gamma hierarchical prior model specification (not provided) where the mean Poisson parameter is treated as an auxiliary variable and imputed within the algorithm. Similarly for the Beta-Binomial model specification, using the explicit probability mass function achieved a better ESS and ESS/s than the equivalent hierarchical prior model specification (not provided) where the Binomial probability parameter is imputed within the MCMC algorithm.

A.2. Second semi-complete data likelihood approach - SCD2. The model component of the JAGS code is provided here for the semi-complete data likelihood approach, specifying the posterior conditional distribution of $N - n$ to be of Negative-Binomial form.

```

84 model{
85   Pi <- 3.14159265359
86   # Priors:
87   alpha ~ dnorm(0.0,0.01)
88   tau ~ dgamma(0.01,0.01)
89   sigma <- 1/sqrt(tau)

```

```

90   for (i in 1:n) {
91     y[i] ~ dbin(p[i],T)
92     logitp[i] ~ dnorm(alpha,tau)
93     logit(p[i]) <- logitp[i]
94   }
95
96   # Posterior conditional distribution for N-n (and hence N):
97
98   n0 ~ dnegbin(pstar,n)
99   N <- n + n0
100
101   # Calculate probability of not being observed using Gauss-Hermite quadrature
102   # q = number of quadrature points
103   # weights and nodes correspond to q quadrature points; entered as data
104   for(i in 1:q){
105     probi[i] <- 1/sqrt(Pi)*weights[i]*(1/(1+exp(sqrt(2)*sigma*nodes[i]+alpha)))^T
106   }
107   pstar <- 1-sum(probi[])
108
109   # Use zero trick for initial 1/(pstar)^n
110
111   loglikterm <- -n*log(pstar)
112   lambda <- -loglikterm + 100000
113   dummy ~ dpois(lambda) # dummy = 0; entered as data
114 }

```

115 **A.3. Super-population complete data likelihood approach - CD:R.** The model com-
116 ponent of the JAGS code for the super-population complete data likelihood approach of [Royle et](#)
117 [al. \(2007\)](#).

```

118 model{
119   # Priors:
120   psi ~ dbeta(0.001,1)
121   alpha ~ dnorm(0.0,0.01)
122   tau ~ dgamma(0.01,0.01)
123   sigma <- 1/sqrt(tau)
124
125   # Complete data likelihood:
126   for(i in 1:M){
127     y[i] ~ dbin(pi[i],T)
128     pi[i] <- z[i]*p[i]
129     z[i] ~ dbern(psi)
130     logit(p[i]) <- logitp[i]
131     logitp[i] ~ dnorm(alpha,tau)
132   }
133
134   # Calculate N:
135   N <- sum(z[1:M])
136 }

```

137 **A.4. Super-population complete data likelihood approach - CD:DE.** The model com-
 138 ponent of the JAGS code for the super-population complete data likelihood approach of [Durban](#)
 139 [and Elston \(2005\)](#).

```

140 model{
141   # Priors:
142   alpha ~ dnorm(0.0,0.01)
143   tau ~ dgamma(0.01,0.01)
144   sigma <- 1/sqrt(tau)
145
146   # Prior for N: (Jeffrey's prior over {n,n+1,...,M} following Link 2013).
147
148   n00 ~ dcat(prior[]) # prior = rep(1/(M+1-n),M+1-n); entered as data
149   n0 <- n00 - 1
150   N <- n+n0
151
152   # Use zero trick for factorial term
153   # Note loggam(N) instead of loggam(N+1) because of Jeffrey's prior for N
154
155   logzeroprob <- loggam(N) - loggam(n0+1) - loggam(n+1)
156   lambda <- -logzeroprob + 1000
157   dummy ~ dpois(lambda) # dummy = 0; entered as data
158
159   # Complete data likelihood:
160
161   for (i in 1:M){
162     y[i] ~ dbin(pi[i],T)
163     pi[i] <- z[i]*p[i]
164     z[i] <- step(N-i)
165     logit(p[i]) <- z[i]*logitp1[i] + (1-z[i])*logitp2[i]
166
167     logitp1[i] ~ dnorm(alpha,tau)
168     logitp2[i] ~ dnorm(alphaprior,tauprior)
169
170     # alpha prior and tauprior are pseudo-prior parameters entered as data
171
172   }
173 }
```

APPENDIX B: JAGS CODE FOR SECR MODEL

174 In this appendix we provide sample JAGS code for the SECR models considered in Section 4.2.

175 **B.1. First semi-complete data likelihood approach - SCD1.** The model component of
 176 the JAGS code is provided here for the semi-complete data likelihood approach using the Jeffreys'
 177 prior specification for N .

```

178 model{
179   # Priors:
180   sigma ~ dunif(0,10)
181   tau <- 1/(sigma*sigma)
182   for(i in 1:n){
```

```

183     X[i] ~ dunif(xlim[1], xlim[2])
184     Y[i] ~ dunif(ylim[1], ylim[2])
185 }
186
187 # pdot = probability of being detected at least once (given location)
188 # Calculate esa numerically using the integration grid
189
190 for(i in 1:G){ # G = number of points on integration grid
191   for(s in 1:S){
192     for(k in 1:K){
193       one_minus_detprob[i,s,k] <- 1 - exp(-dist2[i,k]*tau/2)
194     }
195   }
196   pdot.temp[i] <- 1 - prod(one_minus_detprob[i,,])
197   pdot[i] <- max(pdot.temp[i], 1.0E-10)
198 }
199 esa <- sum(pdot[])*a # a = size of grid square in numerical integration
200 pstar <- esa / A
201
202 # Prior for N: Jeffreys' prior - this is incorporated in the zero trick below
203 # in specifying the likelihood term
204 # However a prior distribution is needed to be specified on N
205 # Use a discrete Uniform prior so the only influence on the posterior
206 # distribution is the upper limit
207
208 n00 ~ dcat(prior[]) # prior = rep(1/(M+1-n),M+1-n); entered as data
209 n0 <- n00 - 1
210 N <- n + n0
211
212 # Zero trick for likelihood component for unobserved individuals
213 logzeroprob <- loggam(N) - loggam(n0+1) - loggam(n+1) + n0*log(1-pstar)
214 lambda <- -logzeroprob + 1000
215 dummy ~ dpois(lambda) # dummy = 0; entered as data
216
217 # Model for capture histories of observed individuals:
218 for(i in 1:n){
219   for(k in 1:K){
220     for(s in 1:S){
221       capthist[i,s,k] ~ dbern(detprob[i,s,k])
222       detprob[i,s,k] <- exp(-r2[i,k] * tau/2 )
223     }
224     r2[i,k] <- pow(X[i] - traps[k,1], 2) + pow(Y[i] - traps[k,2], 2)
225   }
226 }
227 }

```

228 Alternative prior specifications for N can be easily incorporated. Example code for the Negative-
 229 Binomial and Beta-Binomial priors follow analogously to those provided in Section A.1.

230 **B.2. Second semi-complete data likelihood approach - SCD2.** The model component
 231 of the JAGS code is provided here for the semi-complete data likelihood approach, specifying the

232 posterior conditional distribution of $N - n$ to be of Negative-Binomial form.

```

233 model{
234   # Priors:
235   sigma ~ dunif(0,10)
236   tau <- 1/(sigma*sigma)
237   for(i in 1:n){
238     X[i] ~ dunif(xlim[1], xlim[2])
239     Y[i] ~ dunif(ylim[1], ylim[2])
240   }
241
242   # Posterior conditional distribution for N-n (and hence N):
243
244   n0 ~ dnegbin(pstar,n)
245   N <- n + n0
246
247   # pdot = probability of being detected at least once (given location)
248   # calculate esa numerically using the integration grid
249
250   for(i in 1:G){ # G = number of points on integration grid
251     for(s in 1:S){
252       for(k in 1:K){
253         one_minus_detprob[i,s,k] <- 1 - exp(-dist2[i,k] * tau/2)
254       }
255     }
256     pdot.temp[i] <- 1 - prod(one_minus_detprob[i,,])
257     pdot[i] <- max(pdot.temp[i], 1.0E-10)
258   }
259   esa <- sum(pdot[])*a # a = size of grid square in numerical integration
260   pstar <- esa / A
261
262   # Zero trick for initial 1/pstar^n
263
264   loglikterm <- -n * log(pstar)
265   lambda <- -loglikterm + 1000
266   dummy ~ dpois(lambda) # dummy = 0; entered as data
267
268   # Model for capture histories of observed individuals:
269
270   for(i in 1:n){
271     for(k in 1:K){
272       for(s in 1:S){
273         capthist[i,s,k] ~ dbern(detprob[i,s,k])
274         detprob[i,s,k] <- exp(-r2[i,k] * tau/2 )
275       }
276       r2[i,k] <- pow(X[i] - traps[k,1], 2) + pow(Y[i] - traps[k,2], 2)
277     }
278   }
279 }

```

280 **B.3. Super-population complete data likelihood approach - CD:R.** The model com-
 281 ponent of the JAGS code for the super-population complete data likelihood approach.

```

282 model {
283   # Priors:
284   psi ~ dbeta(0.001,1)
285   sigma ~ dunif(0,10)
286   tau <- 1/(sigma*sigma)
287   for(i in 1:M){
288     z[i] ~ dbern(psi)
289     X[i] ~ dunif(xlim[1], xlim[2])
290     Y[i] ~ dunif(ylim[1], ylim[2])
291   }
292
293   # Complete data likelihood component:
294
295   for(i in 1:M){
296     for(k in 1:K){
297       for(s in 1:S){
298         capthist[i,s,k] ~ dbern(detprob[i,s,k])
299         detprob[i,s,k] <- z[i] * exp(-r2[i,k] * tau/2)
300       }
301       r2[i,k] <- pow(X[i] - traps[k,1], 2) + pow(Y[i] - traps[k,2], 2)
302     }
303   }
304
305   # Calculate N:
306   N <- sum(z[])
307 }

```

308 **B.4. Super-population complete data likelihood approach - CD:DE.** The model com-
 309 ponent of the JAGS code for the super-population complete data likelihood approach.

```

310 model{
311
312   # Priors:
313   psi ~ dbeta(0.001,1)
314   sigma ~ dunif(0,10)
315   tau <- 1/(sigma*sigma)
316
317   # Data augmentation part - using Durban and Elston approach:
318
319   for(i in 1:M){
320
321     # Define the first N individuals to be in population of interest
322
323     z[i] <- step(N-i) # z = 1 if i \le N; z = 0 if i > N.
324
325     # Prior for home range centre for an individual in the population
326
327     X1[i] ~ dunif(xlim[1], xlim[2])

```



```

328   Y1[i] ~ dunif(ylim[1], ylim[2])
329
330   # Set pseudo-prior for home range centre for an individual in the population
331   # Independent Beta priors for (x,y) location scaled to be in specified region
332
333   Xtemp ~ dbeta(xprior[1],xprior[2]) # xprior - pseudo-prior parameters entered as data
334   Ytemp ~ dbeta(yprior[1],yprior[2]) # yprior - pseudo-prior parameters entered as data
335
336   X2[i] <- xlim[1] + Xtemp*xlim[2]
337   Y2[i] <- ylim[1] + Ytemp*ylim[2]
338
339   X[i] <- z[i] * X1[i] + (1 - z[i]) * X2[i]
340   Y[i] <- z[i] * Y1[i] + (1 - z[i]) * Y2[i]
341
342 }
343
344 # Prior for N: (Jeffrey's prior over {n,n+1,...,M} following Link 2013).
345
346 n00 ~ dcat(prior[]) # prior = rep(1/(M+1-n),M+1-n); entered as data
347 n0 <- n00 - 1
348 N <- n+n0
349
350 # Use zero trick for factorial term
351 # Note loggam(N) instead of loggam(N+1) because of Jeffrey's prior for N
352
353 logLik <- loggam(N) - loggam(n0 + 1) - loggam(n + 1)
354 phi <- -logLik + 100000
355 dummy ~ dpois(phi) # dummy = 0; entered as data
356
357 for(i in 1:M){
358   for(k in 1:K){
359     for(s in 1:S){
360       capthist[i,s,k] ~ dbern(detprob[i,s,k])
361       detprob[i,s,k] <- z[i] * exp(-r2[i,k] * tau/2)
362     }
363     r2[i,k] <- pow(X[i] - traps[k,1], 2) + pow(Y[i] - traps[k,2], 2)
364   }
365 }
366 }
367

```

REFERENCES

- 368 Durban, J. W. and Elston, D. A. (2005) Mark-recapture with occasion and individual effects: abundance estimation
369 through Bayesian model selection in a fixed dimensional parameter space. *Journal of Agricultural, Biological, and*
370 *Environmental Statistics* **10**, 291-305
- 371 Royle J. A., Dorazio R. M. and Link W. A. (2007) Analysis of multinomial models with unknown index using data
372 augmentaton. *Journal of Computational and Graphical Statistics*, **16**, 67-85

373 ADDRESS OF KING

SCHOOL OF MATHEMATICS, UNIVERSITY OF EDINBURGH,
JAMES CLERK MAXWELL BUILDING,
THE KING'S BUILDINGS, PETER GUTHRIE TAIT ROAD,
EDINBURGH, UK. EH9 3FD .

ADDRESS OF McCLINTOCK

NATIONAL MARINE MAMMAL LABORATORY,
ALASKA FISHERIES SCIENCE CENTER,
NOAA NATIONAL MARINE FISHERIES SERVICE,
7600 SAND POINT WAY NE, SEATTLE,
WASHINGTON 98115 USA

374

ADDRESS OF KIDNEY AND BORCHERS

CENTRE FOR RESEARCH INTO ECOLOGICAL AND ENVIRONMENTAL MODELLING
AND SCHOOL OF MATHEMATICS AND STATISTICS,
THE OBSERVATORY, BUCHANAN GARDENS,
UNIVERSITY OF ST ANDREWS, ST ANDREWS, UK. KY16 9LZ.