

N-Mixture Models

Data Integration

Casey Setash, Colorado Parks and Wildlife

Casey.setash@state.co.us

Motivation: N-Mixture Models

- Royle 2004
- Another example of a hierarchical model
- One of the ways to estimate abundance with imperfect detection
 - Mark-Recapture
 - Removal sampling
 - Distance sampling
 - Double observer
 - N-mixture models

Basic Model Description

- Use: estimate abundance when mark-recapture is not feasible
 - Originally built for sparse data (few sites, few occasions, small detection probability, some zeros)
- Data inputs: repeated counts at the same sites
- Assumptions:
 1. Poisson and binomial distributions are true descriptions of state/observation processes
 1. Variation in counts within a site is due only to observer error
 2. All between-site variation stems from underlying abundance process
 2. Closed population with respect to births, deaths, and migration between occasions
 3. Abundance at each site is random and independent of abundance at all other sites
 4. Constant detection probability p
 5. No double counting

Basic Model Structure

- $y_{i,t}$ = number of individuals observed at sites $i=1\dots R$ during $t=1\dots T$ sampling occasions

$$y_{i,t} \sim \text{binomial}(N_i, p)$$

- N_i = the size of the population at site i
- $\sum_{i=1}^R N_i$ = total population size

Basic Model Structure

Observation Process: $y_{i,t} \sim \text{binomial}(N_i, p)$

State Process: $N_i \sim \text{Poisson}(\lambda)$

Each site has expected abundance λ

Basic Model Structure

Observation Process: $y_{i,t} \sim \text{binomial}(N_i, p)$

State Process: $N_i \sim \text{Poisson}(\lambda)$

Priors: $\log(\lambda) \sim \text{normal}(0, 1000)$

$\text{logit}(p) \sim \text{normal}(0, 0.5)$

Data Structure Example

Site	N	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
1	6	2	3	3	5	3
2	2	0	2	1	1	1
3	4	0	0	3	3	2
4	1	0	1	0	0	1

JAGS Code

```
model{  
  # Priors  
  lambda ~ dgamma(0.01, 0.01)  
  p ~ dbeta(1, 1)  
  
  # Likelihood  
  for(j in 1:J){  
    ## state model  
    N[j] ~ dpois(lambda)  
  
    ## observation model  
    for(k in 1:K){  
      y[j, k] ~ dbinom(p, N[j])  
    } # end k loop  
  } # end j loop  
  
} # end model
```


Model Structure Extensions

Observation Process: $y_{i,t} \sim \text{binomial}(N_i, p)$

State Process: $N_i \sim \text{Poisson}(\lambda_i)$

Covariates: $\log(\lambda_i) = \alpha_0 + \alpha_1 x_1 + \epsilon_i$

Random effect: $\epsilon_i \sim \text{normal}(0, \tau)$

$\text{logit}(p_{i,t}) = \beta_0 + \beta_1 x_1 \dots$

Model Structure Extensions – Dail-Madsen

Observation Process: $y_{i,t} \sim \text{binomial}(N_{i,t}, p)$

State Process: $N_{i,1} \sim \text{Poisson}(\lambda)$

$$N_{i,t} = S_{i,t} + G_{i,t}$$

$$S_{i,t} \sim \text{binomial}(N_{i,t-1}, \phi)$$

$$G_{i,t} \sim \text{Poisson}(\gamma N_{i,t-1})$$

Model Structure Extensions

- Other distributions
 - Negative binomial
 - Zero-inflated Poisson
 - Multinomial
- Integrated – known fate or distance
- Multiple age/sex “states”
- Spatial movement of animals

Coding Options

- Maximum Likelihood:
 - Unmarked (Fiske and Chandler 2011) – pcount function
- Bayesian Estimation:
 - Specify yourself using JAGS, Nimble, INLA, or STAN
 - ubms package (Kellner et al. 2022)

Simple Example

- Basile et al. (2021) Ecological Indicators
- Abundance, species richness and diversity of forest bird assemblages
 - The relative importance of habitat structures and landscape context
- How does bird abundance respond to landscape context in forests of similar structure?

Simple Example

- Black Forest, SW Germany
- Point counts repeated up to 3 times per year in 126 one-ha plots
 - Species divided up by feeding/nesting guilds
- 1. Assemblage process: $w_k \sim \text{Bernoulli}(\Omega)$
- 2. Single species abundance: $N_{ik}/w_k \sim \text{Poisson}(w_k \lambda_k)$
- 3. Detection process: $y_{ijk}/N_{ik} \sim \text{Binomial}(N_{ik}, p_{ijk})$

$\log(\lambda_k) = \text{model coefficients} * \text{forest covariates}$

N-Mixture Models

- Pros

- Another option to estimate abundance when data are limited
 - Count data are “cheap”
- Don’t need auxiliary info (distance, double observer, etc.)
- Easy to describe and implement the model
- Widely applicable to count datasets

- Cons

- Trade-off between temporal replication and spatial replication
- Parameters may not be identifiable
- Sensitive to model assumption violations

Identifiability

- Definition reminder – the ability to estimate every parameter in a model individually rather than as the product of a smaller set of parameters
- Example:
 - One sampling occasion ($T=1$)
 - Observed counts are distributed as $\text{Poisson}(p\lambda)$
 - p and λ are not identifiable
 - Even when $T=2$ or 3 , the likelihood surface that you are maximizing may be too flat – end up with infinite MLEs of expected abundance
 - More occasions and including covariates on p and λ can help

Controversy

- Barker et al. (2017)
 - Need auxiliary information on p to estimate it separately
 - Even if you assume p is constant, data can be generated by processes other than the N-mixture model
 - Essentially estimating relative abundance
 - Might as well just use Poisson regression

Controversy

- Kery (2017)
 - Tested Barker et al's doubts about identifiability with 137 bird datasets
 - No issues for Poisson and ZIP N-mixture models
 - Negative binomial models had issues for 25% of datasets
 - Small sample sizes caused issues (<267 sites)
 - Covariates did not affect identifiability

Controversy

- Link et al. (2018)
 - N-mixture models are *very* sensitive to violation of assumptions
 - Tested double counting, unmodeled variation in λ , and unmodeled variation in p
 - Small violations have been shown to result in large biases in abundance estimates
 - Violations are hard to detect using goodness-of-fit tests

N-Mixture Models – Final Thoughts

“We stress from the outset that, in principle, N-mixture models work: if the model is correct (an exact depiction of the data generating process), if the parameter values are not extreme, and if we have adequate data, the N-mixture models can be estimated with high precision.”

“Our conclusion is that for estimating absolute abundance, there is no substitute for mark–recapture analysis: N-mixture modeling relies too heavily on questionable and poorly verifiable model assumptions.”

– Link et al. 2018, Ecology

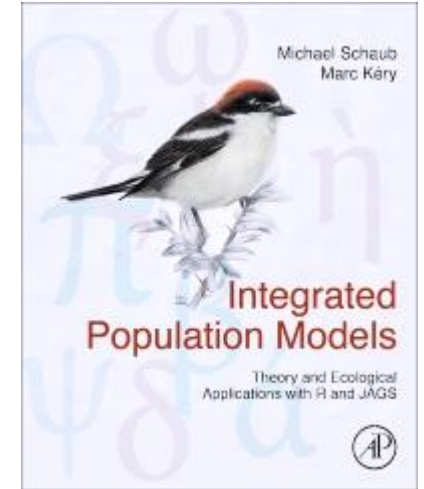
SWITCHING GEARS

Data Integration

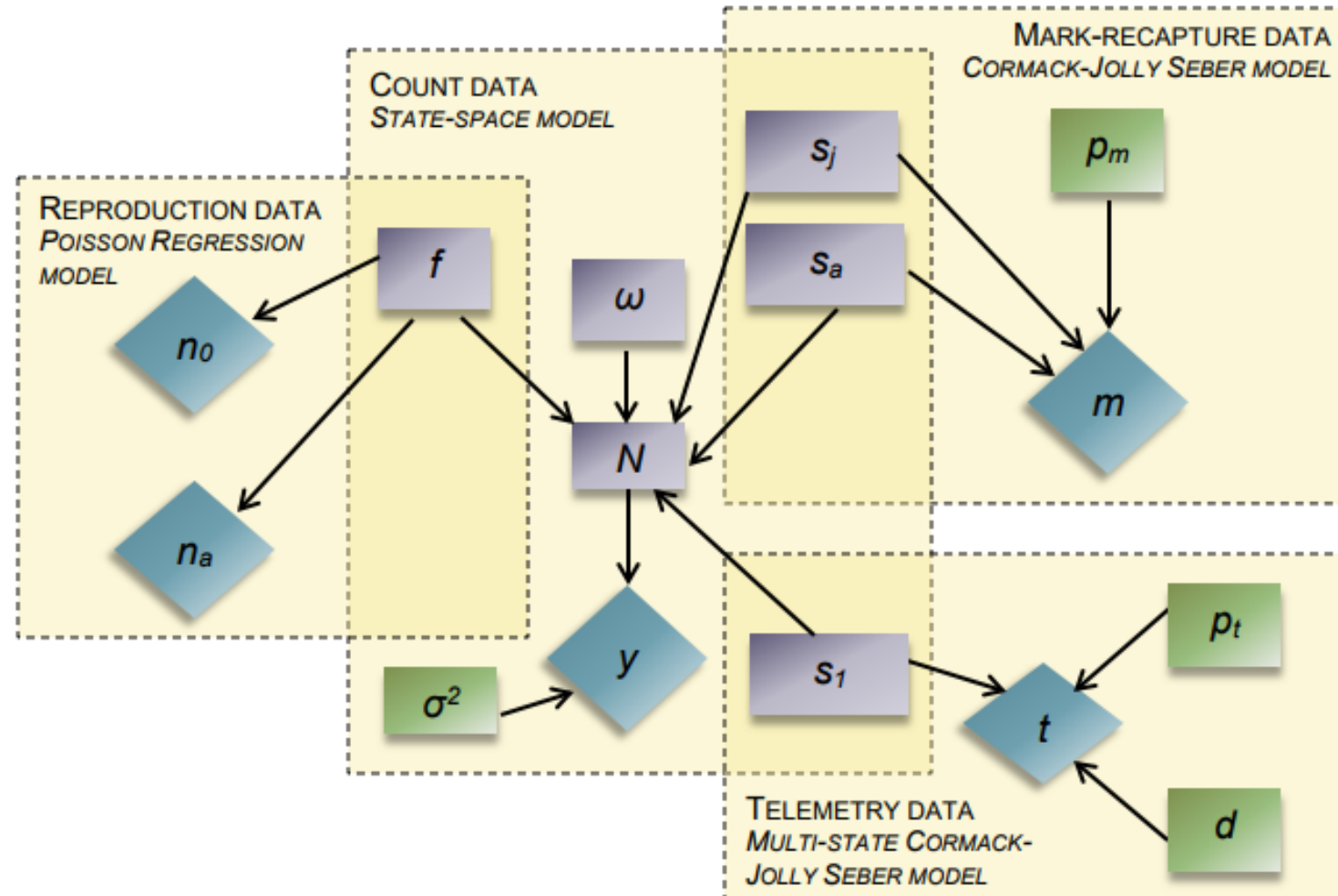
- Moving on!
- Other names: integrated models, data fusion, assimilation, combination
- Commonly heard terms:
 - Integrated population model (IPM)
 - Integrated species distribution model (iSDM or IDM)
 - Integrated community model (ICM)
 - Integrated dynamic N-mixture model
- Super trendy topic in the past several decades

What is it?

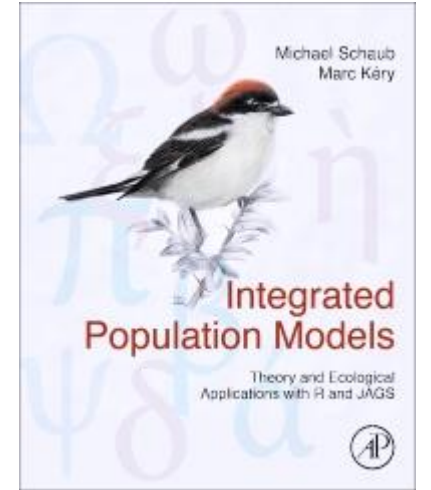
- A type of hierarchical model
- Multiple disparate data sets that each contain information about a shared process with one or more shared parameters
- Combine the processes into a single model
 - Identify shared parameters
 - Write a model that links them based on the different data sets
- “...datastream-specific observation submodels that inform a common ecological model” – Gilbert et al. 2021



Integrated Population Model



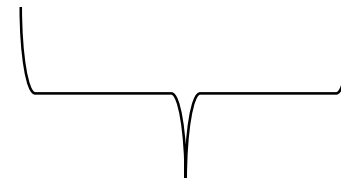
Integrated Population Model



- Shared latent state: most commonly abundance
- **Step 1:** Set up a population model that links abundance and vital rates

Most simple: Balance Equation $N_{t+1} = N_t (1 + B_t + I_t - D_t - E_t)$

$$N_{t+1} = N_t (1 + B_t - D_t)$$



$$\lambda_t = \exp(r_t)$$

Integrated Population Model

- Structured version of balance equations
 - Matrix projection models (Caswell 2001)
 - Integral projection models

$$n_{t+1} = \mathbf{A}_t n_t$$

$$\mathbf{A}_t = \begin{bmatrix} f_1 & f_2 \\ s_1 & s_2 \end{bmatrix}$$

State-Space Models

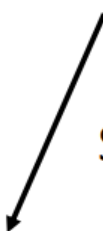
- Hierarchical model
- Can decouple observation error from process variation of interest (e.g. environmental stochasticity)
- Key component of an IPM
- Two sets of equations:
 - State process – describes the development of the true state of a system over time
 - Observation process – relates the true (latent) state to the observations (counts)

State-Space Likelihood in IPM

1. State process

$$\begin{bmatrix} n_{J,t+1} \\ n_{A,t+1} \end{bmatrix} = \begin{bmatrix} 0 & F_{A,t} \\ S_{J,t} & S_{A,t} \end{bmatrix} \begin{bmatrix} n_{J,t} \\ n_{A,t} \end{bmatrix}$$

2. Observation process



Sum

$$\log(y_t) \mid \log(N_t) \sim \text{Normal}(\log(N_t), \sigma_y^2)$$

Y_t population count at time t

σ_y^2 residual variation (observation error)

Integrated Population Model

S: survival
F: productivity
N_{juv}: pop size of juveniles
N_{ad}: pop size of adults
 σ_y^2 : census/residual error

y: counts

- **Step 2:** Define likelihoods for different datasets
- Counts: state-space model

State process: $N_{juv,t+1} \sim \text{Poisson}(N_{juv,t}S_{juv,t}F_{juv,t} + N_{ad,t}S_{juv,t}F_{ad,t})$

$$N_{ad,t+1} \sim \text{binomial}(N_{juv,t} + N_{ad,t}S_{ad,t})$$

Observation Process: $y_t \sim \text{normal}((N_{juv,t} + N_{ad,t}), \sigma_y^2)$

Integrated Population Model

S: survival

P: recapture probability

m: m-array

- **Step 2:** Define likelihoods for different datasets
- Capture-Recapture Data: CJS or multinomial

$$L_{CJS}(S_{juv}, S_{ad}, P | m)$$

Integrated Population Model

F: productivity

J: number of nestlings

- **Step 2:** Define likelihoods for different datasets
- Productivity Data: Poisson regression

$$J \sim \text{Poisson}(F)$$

$$\log(F) = \beta_0 + \beta_1 x_1 \dots$$

Integrated Population Model

- **Step 3:** Put them all together into a joint likelihood

$$L_{IPM}(N, S_{juv}, S_{ad}, F, P, \sigma_y^2 | y, m, J) = L_1(N_1) \times L_o(N, \sigma_y^2 | y) \times L_s(N, S_{juv}, S_{ad}, F) \times L_{CJS}(S_{juv}, S_{ad}, P | m) \times L_p(F | J)$$

IPM Example

- Woodchat Shrike
- Monitored for 34 years in Germany
- Nested in extensive orchards, ate large insects
- Long-term decline in Germany, France, and Switzerland
- Searched nesting territories
- Located nests, counted eggs, hatchlings, and fledglings
- Marked nestlings and adults with color bands



IPM Example

- Life cycle graph
- Pre-breeding census, female-based model
- Local recruits, surviving adults, immigrants

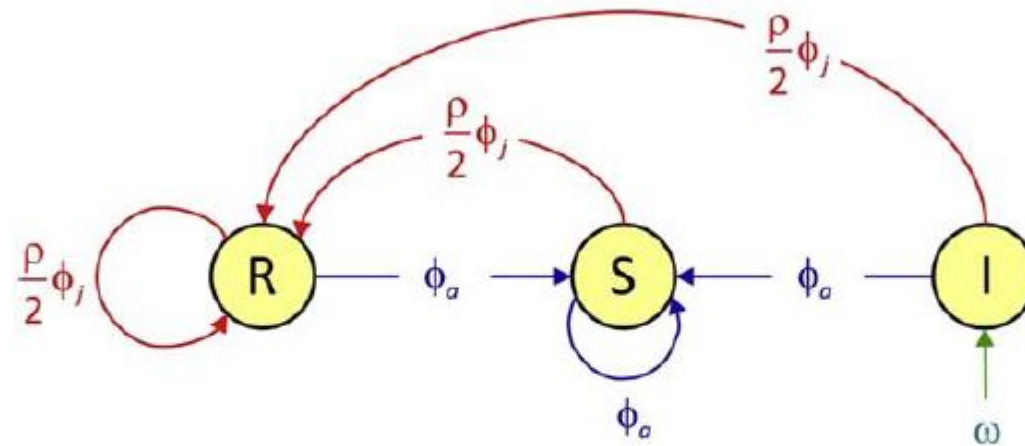


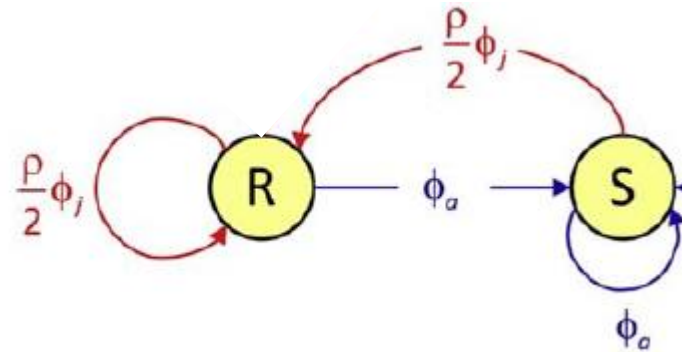
FIGURE 11.1

Life-cycle graph for the woodchat shrike model. The three population stages are local recruits (R), surviving adults (S) and immigrants (I), and the demographic processes are productivity (governed by productivity per female and year, ρ), apparent survival (governed by ϕ_j for juveniles and ϕ_a for adults), and immigration (governed by ω). Transitions involving local recruitment are shown as red arrows, pure survival transitions in blue, and immigration in green.



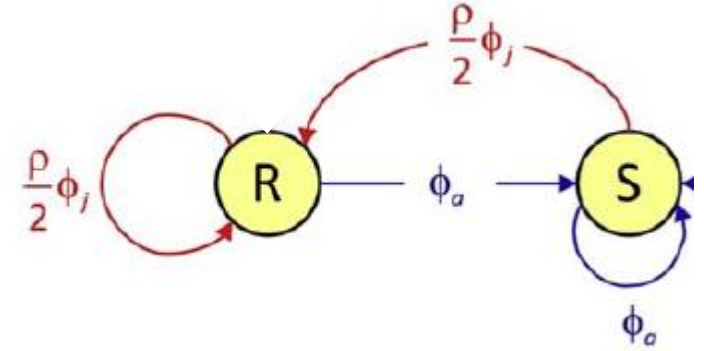
IPM Example

- Life cycle graph
- Pre-breeding census, female-based model
- Local recruits, surviving adults



IPM Example

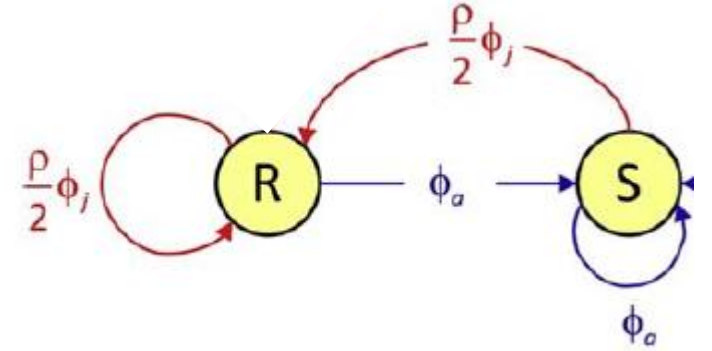
- Population model



$$\begin{bmatrix} N_{1,t+1} \\ N_{ad,t+1} \end{bmatrix} = \begin{bmatrix} \frac{f_{1,t}}{2} \phi_j & \frac{f_{a,t}}{2} \phi_j \\ \phi_a & \phi_a \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{ad,t} \end{bmatrix}$$

IPM Example

- Population model



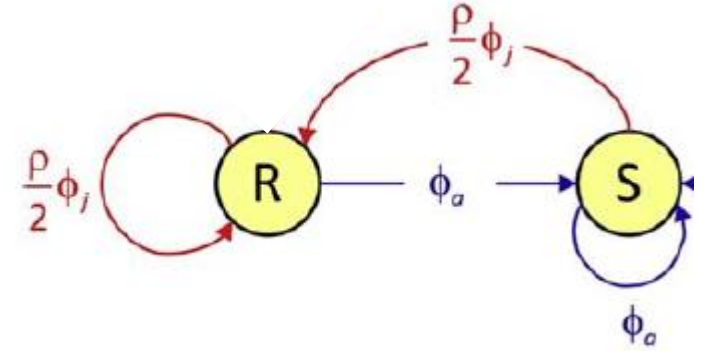
$$\begin{bmatrix} N_{1,t+1} \\ N_{ad,t+1} \end{bmatrix} = \begin{bmatrix} \frac{f_{1,t}}{2} \phi_j & \frac{f_{a,t}}{2} \phi_j \\ \phi_a & \phi_a \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{ad,t} \end{bmatrix}$$

$$N_{1,t+1} = N_{1,t} \frac{f_{1,t}}{2} \phi_j + N_{ad,t} \frac{f_{a,t}}{2} \phi_j$$

$$N_{ad,t+1} = (N_{1,t} + N_{ad,t}) \phi_a$$

IPM Example

- Population model



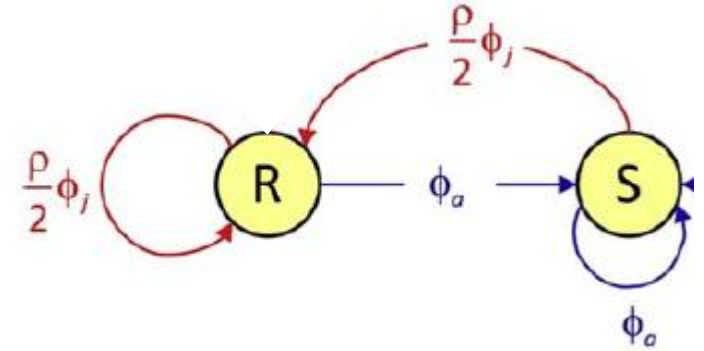
$$\begin{bmatrix} N_{1,t+1} \\ N_{ad,t+1} \end{bmatrix} = \begin{bmatrix} \frac{f_{1,t}}{2} \phi_j & \frac{f_{a,t}}{2} \phi_j \\ \phi_a & \phi_a \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{ad,t} \end{bmatrix}$$

$$N_{1,t+1} \sim \text{Poisson}\left(N_{1,t} \frac{f_{1,t}}{2} \phi_j + N_{ad,t} \frac{f_{a,t}}{2} \phi_j \right)$$

$$N_{ad,t+1} \sim \text{binomial}(\phi_a, N_{1,t} + N_{ad,t})$$

IPM Example

- Population model



$$\begin{bmatrix} N_{1,t+1} \\ N_{ad,t+1} \end{bmatrix} = \begin{bmatrix} \frac{f_{1,t}}{2} \phi_j & \frac{f_{a,t}}{2} \phi_j \\ \phi_a & \phi_a \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{ad,t} \end{bmatrix}$$

```
# Process model over time: our model of population dynamics
for (t in 1:(n.occasions-1)){
  N[1,t+1] ~ dpois(N[1,t] * f[1,t] / 2 * sj[t] + N[2,t] * f[2,t] / 2 * sj[t])
  N[2,t+1] ~ dbin(sa[t], N[1,t] + N[2,t])
}
```

IPM Example

- Observation model
- C = counts (data)

$$C_t \sim \text{normal}(N_{1,t} + N_{ad,t}, \sigma^2)$$

IPM Example

- Poisson model
- J = number of fledglings

$$J_t \sim \text{Poisson}(f_{i,t})$$

```
# Productivity data (Poisson regression model)
for (i in 1:length(J)){
  J[i] ~ dpois(f[age[i],year[i]])
}
```

- CJS model

[illegible]

IPM Example

- CJS model

$$\begin{bmatrix} S_{j,tp_t} & S_{a,tp_t} & \dots & \dots & \dots & \dots & \dots & (1 - P(\text{othercells})) \\ 0 & S_{j,tp_t} & S_{a,tp_t} & \dots & \dots & \dots & \dots & (1 - P(\text{othercells})) \\ 0 & 0 & S_{j,tp_t} & S_{a,tp_t} & \dots & \dots & \dots & (1 - P(\text{othercells})) \\ 0 & 0 & 0 & S_{j,tp_t} & S_{a,tp_t} & \dots & \dots & (1 - P(\text{othercells})) \\ 0 & 0 & 0 & 0 & S_{j,tp_t} & S_{a,tp_t} & \dots & (1 - P(\text{othercells})) \\ 0 & 0 & 0 & 0 & 0 & S_{j,tp_t} & S_{a,tp_t} & (1 - P(\text{othercells})) \\ 0 & 0 & 0 & 0 & 0 & 0 & S_{j,tp_t} & (1 - P(\text{othercells})) \end{bmatrix}$$

IPM Example

- CJS model

$$\left(\begin{array}{cccccccccccccccccccccccccccc} 8 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 \end{array} \right) \sim$$

$$\text{multinomial}(\begin{array}{cccccccc} S_{j,tp_t} & S_{a,tp_t} & \dots & \dots & \dots & \dots & \dots & (1 - P(\text{othercells}) \end{array}, 51)$$

IPM Example

- CJS model

```
104 # Capture-recapture data (CJS model with multinomial likelihood)
105 # Define the multinomial likelihood
106 for (t in 1:(n.occasions-1)){
107   marr.j[t,1:n.occasions] ~ dmulti(pr.j[t,], rel.j[t])
108   marr.a[t,1:n.occasions] ~ dmulti(pr.a[t,], rel.a[t])
109 }
110 # Define the cell probabilities of the m-arrays
111 for (t in 1:(n.occasions-1)){
112   # Main diagonal
113   q[t] <- 1 - p[t] # Probability of non-recapture
114   pr.j[t,t] <- sj[t] * p[t]
115   pr.a[t,t] <- sa[t] * p[t]
116   # Above main diagonal
117   for (j in (t+1):(n.occasions-1)){
118     pr.j[t,j] <- sj[t] * prod(sa[(t+1):j]) * prod(q[t:(j-1)]) * p[j]
119     pr.a[t,j] <- prod(sa[t:j]) * prod(q[t:(j-1)]) * p[j]
120   } #j
121   # Below main diagonal
122   for (j in 1:(t-1)){
123     pr.j[t,j] <- 0
124     pr.a[t,j] <- 0
125   } #j
126 } #t
127 # Last column: probability of non-recapture
128 for (t in 1:(n.occasions-1)){
129   pr.j[t,n.occasions] <- 1-sum(pr.j[t,1:(n.occasions-1)])
130   pr.a[t,n.occasions] <- 1-sum(pr.a[t,1:(n.occasions-1)])
131 }
```

IPM Example

- Priors

```
# Priors and linear models
for (t in 1:(n.occasions-1)){
  logit.sj[t] ~ dnorm(mu.sj, tau.sj)
  sj[t] <- ilogit(logit.sj[t])           # Back-transformation from logit scale
  logit.sa[t] ~ dnorm(mu.sa, tau.sa)
  sa[t] <- ilogit(logit.sa[t])          # Back-transformation from logit scale
  p[t] <- mean.p
}

for (t in 1:n.occasions){
  log.f[1,t] ~ dnorm(mu.f[1], tau.f[1])
  f[1,t] <- exp(log.f[1,t])             # Back-transformation from log scale
  log.f[2,t] ~ dnorm(mu.f[2], tau.f[2])
  f[2,t] <- exp(log.f[2,t])             # Back-transformation from log scale
}

mean.sj ~ dunif(0, 1)
mu.sj <- logit(mean.sj)                 # Logit transformation
mean.sa ~ dunif(0, 1)
mu.sa <- logit(mean.sa)                 # Logit transformation
sigma.sj ~ dunif(0, 3)
tau.sj <- pow(sigma.sj, -2)
sigma.sa ~ dunif(0, 3)
tau.sa <- pow(sigma.sa, -2)

for (j in 1:2){
  mean.f[j] ~ dunif(0, 10)
  mu.f[j] <- log(mean.f[j])             # Log transformation
  sigma.f[j] ~ dunif(0, 3)
  tau.f[j] <- pow(sigma.f[j], -2)
}

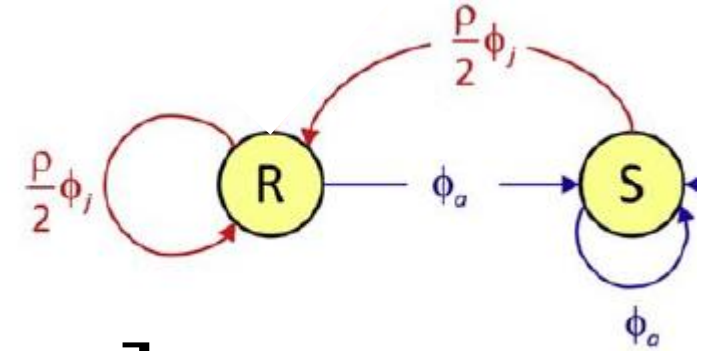
mean.p ~ dunif(0, 1)

sigma ~ dunif(0.5, 100)
tau <- pow(sigma, -2)
```

IPM Example – Practical notes

- Where is the joint likelihood in the JAGS code?
 - It's defined implicitly by a series of simpler, local relationships
 - Practically, the same name of parameters in the different likelihoods insures the integration
- Assumptions
 - The assumptions that each submodel requires also apply to the IPM
 - Additional assumptions due to integration
 - Independence of datasets
 - Common demography – datasets stem from the same population

IPM Example



$$\begin{bmatrix} N_{1,t+1} \\ N_{ad,t+1} \end{bmatrix} = \begin{bmatrix} \frac{f_{1,t}}{2} \phi_j & \frac{f_{a,t}}{2} \phi_j \\ \phi_a & \phi_a \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{ad,t} \end{bmatrix}$$

```
# Process model over time: our model of population dynamics
for (t in 1:(n.occasions-1)){
  N[1,t+1] ~ dpois(N[1,t] * f[1,t] / 2 * sj[t] + N[2,t] * f[2,t] / 2 * sj[t])
  N[2,t+1] ~ dbin(sa[t], N[1,t] + N[2,t])
}

# Observation model
for (t in 1:n.occasions){
  C[t] ~ dnorm(N[1,t] + N[2,t], tau)
}
```


IPM Example – Practical notes

- Where is the joint likelihood in the JAGS code?
 - It's defined implicitly by a series of simpler, local relationships
 - Practically, the same name of parameters in the different likelihoods insures the integration
- Assumptions
 - The assumptions that each submodel requires also apply to the IPM
 - Additional assumptions due to integration
 - Independence of datasets
 - Common demography – datasets stem from the same population

IPM Example – Practical notes

- How accurate is the pop size estimated with an IPM?
 - Depends on the type of error that we have while counting individuals
 - Are counts correct on average (both types of error – double counting and under counting – occur)?
 - Only imperfect detection?

IPM Example – Practical notes

- Goodness-of-fit:
 - Can be extremely challenging (but important) – area of current research
 - Assess the fit of each submodel (e.g., posterior predictive checks)
 - Derived/latent quantities expected to be the most biased if the submodels do not fit well

IPM Example – Practical notes

- Modes of analysis
- Bayesian
 - Most flexible and often easiest for ecologists
- Frequentist
 - Faster computation
 - Easier model selection (AIC)
 - State process requires normal distribution – not always representative of data generating process
 - Kalman Filters needed – can be more complicated

Integrated Species Distribution Model

- “...datastream-specific observation submodels that inform a common ecological model” – Gilbert et al. 2021
- Spatial point process – expected abundance determines whether a species occurs in an area and true abundance within that area
 - Hunter harvest data and camera trap data

Data Integration - Advantages

- Often results in improved accuracy and precision
- Makes use of multiple datasets
- Can reconcile bias among datasets
- Can make additional parameters estimable
 - Time series of counts + capture-recapture = productivity
- Propagation of uncertainty from vital rate estimates to other outcomes (abundance, population growth rates)

Data Integration - Words of Warning

- “Moreover, the advent of data integration makes rigorous data even more valuable as a reference for identifying and adjusting for deficiencies in opportunistic data in the context of data mining.” – Latif et al. 2024