Bayesian integrated population modeling using JAGS

Introduction to matrix projection models

Introduction

- Demographic processes (reproduction, survival, immigration, emigration) drive population dynamics

- Mechanistic understanding of population dynamics requires models that link population changes and demographic processes

- The change of population size from year t to $t+1$ is given by:

$$N_{t+1} - N_t = B_t + I_t - D_t - E_t$$

where:

B_t = Number of births, I_t = number of immigrants, D_t = number of deaths, E_t = number of emigrants

Introduction

- After re-arranging we get:

$$N_{t+1} = N_t + B_t + I_t - D_t - E_t$$

$$N_{t+1} = B_t + I_t + S_t - E_t$$

- Expressed per capita:

$$\lambda_t = \frac{N_{t+1}}{N_t} = \frac{B_t + I_t + S_t - E_t}{N_t} = b_t + i_t + s_t - e_t$$

- The BIDE model assumes that all individuals have the same demographic rates

- Age- or stage structured models relax this assumption and are much more realistic

- Such models are called also *matrix population models, matrix projection models* or *stage-structured population models*
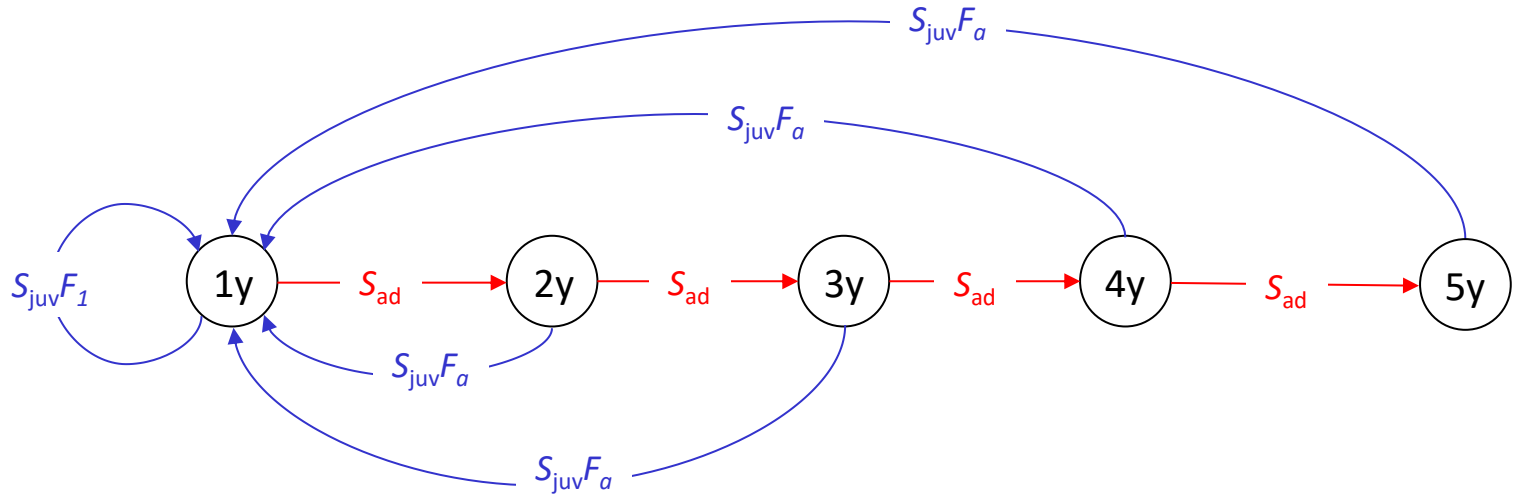
The life cycle graph

- **Key**: the life cycle of the species under study must be known

- Example: woodchat shrike

  - One brood annually
  - Age at first breeding = 1
  - Pre-breeding census

  - Juvenile and adult survival different
  - Productivity of 1y and older individuals different
  - Female-based

S_{juv}: juvenile survival
S_{ad}: adult survival
F_1: productivity of 1y females
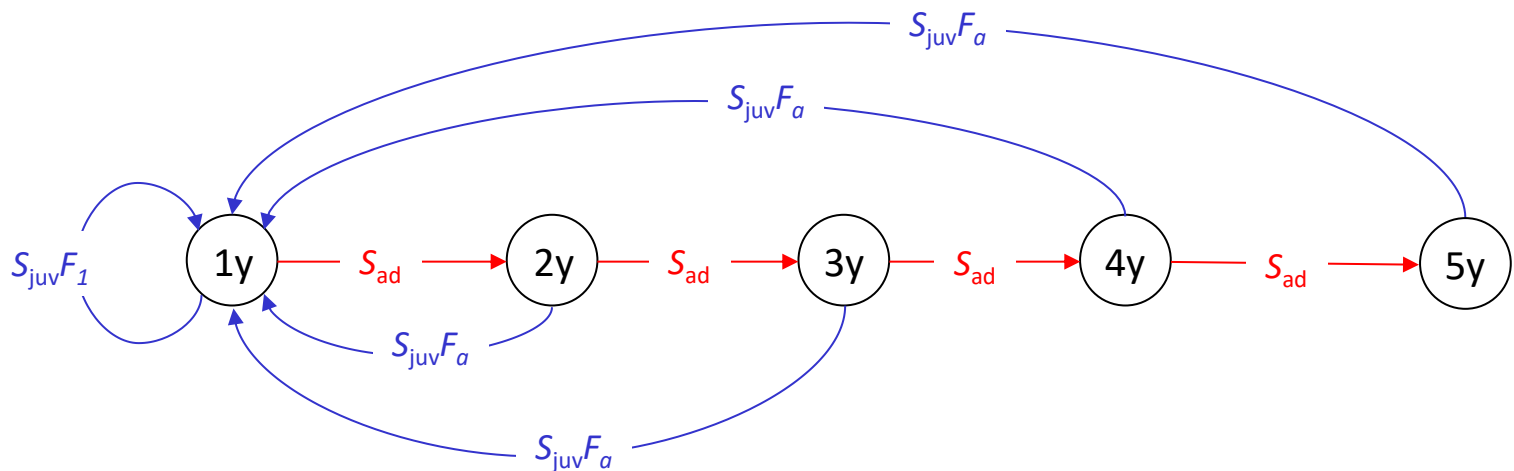F_a: productivity of 1y+ females
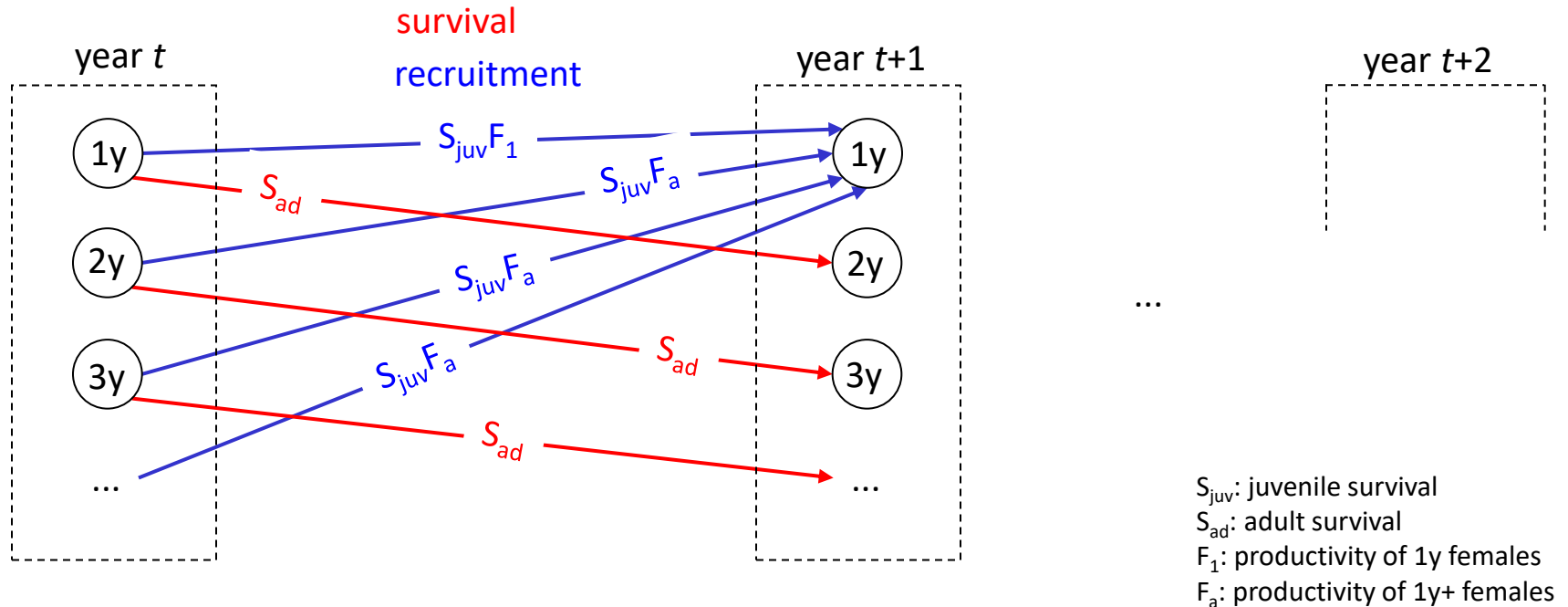
The life cycle graph

- Another, and more usual, way to show the life cycle graph, does not distinguish explicitly between different years
- Thus, transitions from one node to the next jointly show changes of age classes and years
- Maximal age: 5 years

From a life cycle graph to the population model

Population sizes of the age classes:

$$N_{1\,y,t+1} = N_{1\,y,t}S_{juv}F_1 + N_{2\,y,t}S_{juv}F_a + N_{3\,y,t}S_{juv}F_a + N_{4\,y,t}S_{juv}F_a + N_{5\,y,t}S_{juv}F_a$$

$$N_{2\,y,t+1} = N_{1\,y,t}S_{ad}$$

$$N_{3\,y,t+1} = N_{2\,y,t}S_{ad}$$

$$N_{4\,y,t+1} = N_{3\,y,t}S_{ad}$$

$$N_{5\,y,t+1} = N_{4\,y,t}S_{ad}$$

From a life cycle graph to the population model

Population sizes of the age classes:

$$N_{1y,t+1} = N_{1y,t}S_{juv}F_1 + N_{2y,t}S_{juv}F_a + N_{3y,t}S_{juv}F_a + N_{4y,t}S_{juv}F_a + N_{5y,t}S_{juv}F_a$$

$$N_{2y,t+1} = N_{1y,t}S_{ad}$$

$$N_{3y,t+1} = N_{2y,t}S_{ad}$$
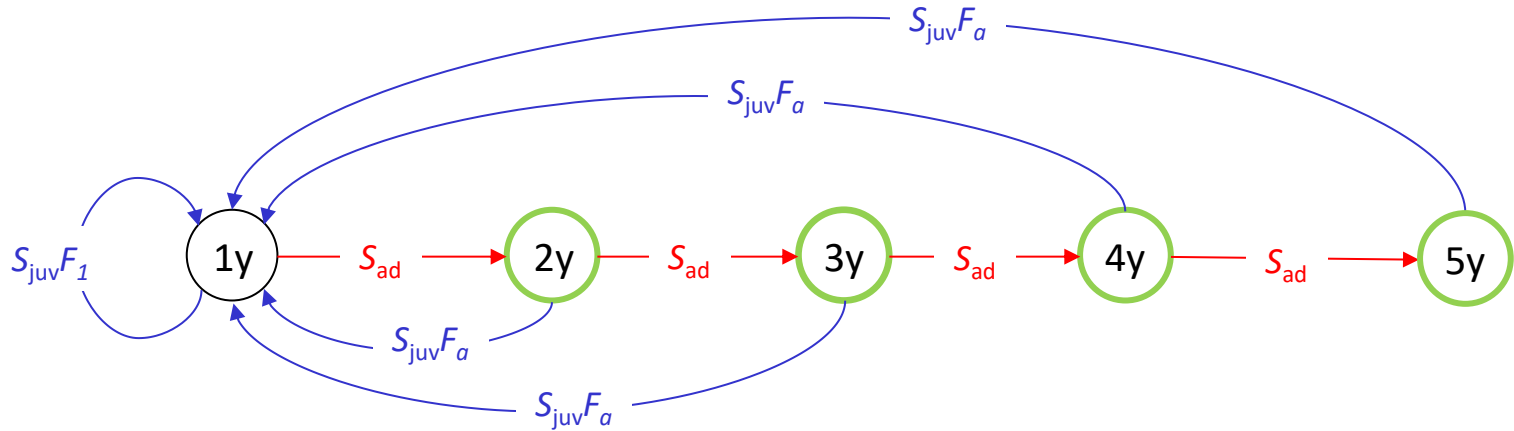
$$N_{4y,t+1} = N_{3y,t}S_{ad}$$

$$N_{5y,t+1} = N_{4y,t}S_{ad}$$

The same written with a matrix and vectors

$$
\begin{bmatrix} N_{1y} \\ N_{2y} \\ N_{3y} \\ N_{4y} \\ N_{4y} \end{bmatrix}_{t+1}
=
\begin{bmatrix}
S_{juv}F_1 & S_{juv}F_a & S_{juv}F_a & S_{juv}F_a & S_{juv}F_a \\
S_{ad} & 0 & 0 & 0 & 0 \\
0 & S_{ad} & 0 & 0 & 0 \\
0 & 0 & S_{ad} & 0 & 0 \\
0 & 0 & 0 & S_{ad} & 0
\end{bmatrix}
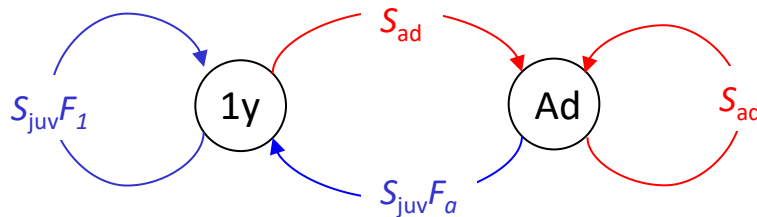\begin{bmatrix} N_{1y} \\ N_{2y} \\ N_{3y} \\ N_{4y} \\ N_{5y} \end{bmatrix}_{t}
$$

Leslie matrix

From a life cycle graph to the population model

The age classes 2y, 3y, and 4y are identical from a demographic point of view (arrows that leave the node are identical). Therefore, these nodes can be collapsed.
We can include node 5y as well if we relax the assumption of the maximal age. The model can be simplified as:

From a life cycle graph to the population model

$$N_{1y,t+1} = N_{1y,t}S_{juv}F_1 + N_{2y,t}S_{juv}F_a + N_{3y,t}S_{juv}F_a + N_{4y,t}S_{juv}F_a + N_{5y,t}S_{juv}F_a$$

$$N_{2y,t+1} = N_{1y,t}S_{ad}$$

$$N_{3y,t+1} = N_{2y,t}S_{ad}$$

$$N_{4y,t+1} = N_{3y,t}S_{ad}$$

$$N_{5y,t+1} = N_{4y,t}S_{ad}$$

The simplified Leslie matrix (often called Lefkovich matrix):

$$
\begin{bmatrix} N_{1y} \\ N_{ad} \end{bmatrix}_{t+1}
=
\begin{bmatrix} S_{juv}F_1 & S_{juv}F_a \\ S_{ad} & S_{ad} \end{bmatrix}
\begin{bmatrix} N_{1y} \\ N_{ad} \end{bmatrix}_{t}
$$

or:

$$N_{1y,t+1} = N_{1y,t}S_{juv}F_1 + N_{ad,t}S_{juv}F_a$$

$$N_{ad,t+1} = N_{1y,t}S_{ad} + N_{ad,t}S_{ad}$$

Growth of an stage-structured population

S_{juv}: 0.3 $N_{1y,1} = 1$
S_{ad}: 0.55 $N_{ad,1} = 9$
F_1: 1.3
F_a: 1.8

$$\begin{bmatrix} N_{1y} \\ N_{ad} \end{bmatrix}_2 = \begin{bmatrix} 0.39 & 0.54 \\ 0.55 & 0.55 \end{bmatrix} \begin{bmatrix} 1 \\ 9 \end{bmatrix}_1$$
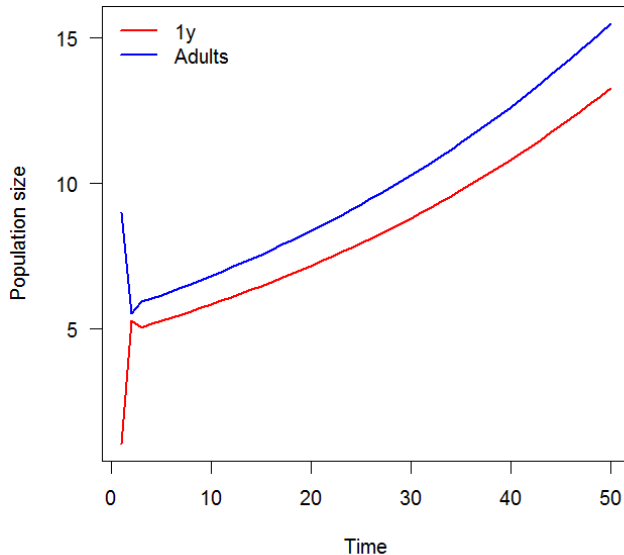
Mean population growth rate:

$$\lambda = \sqrt[40]{\frac{N_{50}}{N_{10}}}$$

$\lambda_{juv} = \lambda_{ad} = 1.020818$

Age distribution:

$$\frac{N_{1y,10}}{N_{1y,10} + N_{ad,10}} = 0.46 \qquad \frac{N_{ad,10}}{N_{1y,10} + N_{ad,10}} = 0.54$$

Once a stable age ratio is reached, all age classes of the population are growing exponentially at the same rate. This is the **asymptotic population growth rate**.

Properties of the Leslie matrix

Leslie/Lefkovitch matrix:

• Basic papers by Leslie (1945, Biometrika) and Lefkovitch (1965, Biometrics)

• Caswell (2001) *Matrix population models*. Sinauer

• Very flexibel tool

The Leslie matrix has very nice properties. It allows the calculation of life history summaries from the population without knowing their size:

• Asympotic population growth rate

• Stable age/stage distributions

• Age-specific reproductive values

• Net reproductive rate R_0

• Generation time

• Growth rate sensitivities and elasticities

Asymptotic population growth rate

- Either projection of population size, or
- Dominant eigenvalue of the Leslie matrix
- The dominant eigenvalue needs the solution of the characteristic equation (with respect to λ): det($\mathbf{A}$-$\lambda\mathbf{I}$) = 0
- Calculated in R as:

```
Sjuv <- 0.3
Sad <- 0.55
F1 <- 1.3
Fa <- 1.8
A <- matrix(c(Sjuv*F1, Sjuv*Fa,
              Sad,      Sad), ncol = 2, byrow = TRUE)

max(Re(eigen(A)$values))        # A is the Leslie matrix

[1] 1.020818
```

Stable stage distribution

- Proportion of individuals belonging to each age class
- Calculated as the right eigenvector of the Leslie matrix corresponding to the dominant eigenvalue
- In R:

```
u <- which.max(Re(eigen(A)$values ))
revec <- Re(eigen(A)$vectors[,u])
stand.revec <- matrix(revec/sum(revec)) # standardised eigenvector
```

Woodchat shrike example: the standardized stable age distribution:

1y: 0.46

Adults: 0.54

Growth rate sensitivities and elasticities

- Sensitivity and elasticity determine, how the growth rate changes if a matrix entry (or a single vital rate) changes by a small amount
- Prospective perturbation
- Sensitivity and elasticity are related: elasticity = proportional sensitivity
- Sensitivity is the slope of a demographic rate against growth rate

Sensitivity
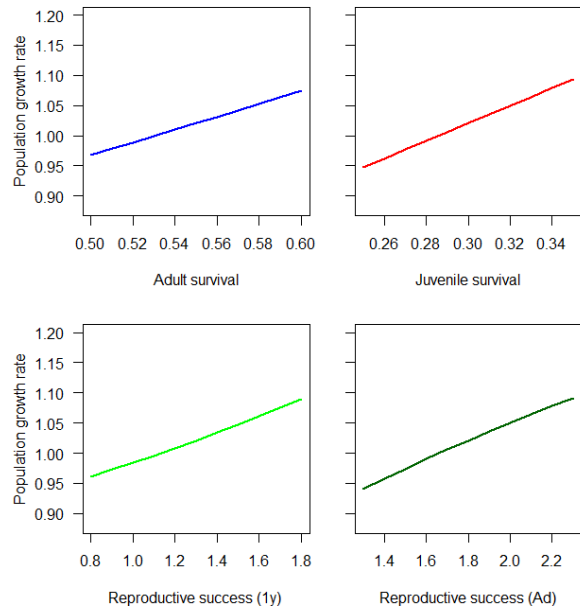
Elasticity
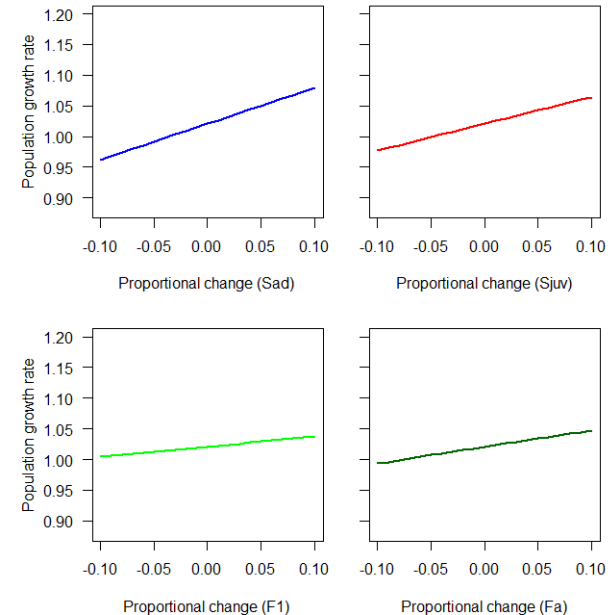
Relaxing assumptions

- So far:
  - Constant demographic rates
  - No density-dependence   Exponentially growing population
  - No demographic stochasticity
  - No parameter uncertainty

- Relaxing these assumptions usually requires simulations to calculate features of population dynamics

- In the following we want to include:
  - **Environmental stochasticity** (time-dependent demographic rates)
  - **Density-dependence** (demographic rates depend on population size)
  - **Demographic stochasticity** (statistical distributions are used to project the population over time)
  - **Parameter uncertainty**

Including environmental stochasticity

Calculation of the stochastic population growth rate

$$\mathbf{N}_{t+1} = \mathbf{A}_t \mathbf{N}_t$$

$$r_s = \frac{1}{T-u} \sum_{i=u}^{T} \log\left(\frac{N_{i+1}}{N_i}\right)$$

note that $r_s = \log(\lambda_s)$

T should be large (>10'000)

Including environmental stochasticity

Calculation of the stochastic population growth rate

$$\mathbf{N}_{t+1} = \mathbf{A}_t \mathbf{N}_t$$

$$r_s = \frac{1}{T-u} \sum_{i=u}^{T} \log\left(\frac{N_{i+1}}{N_i}\right)$$

note that $r_s = \log(\lambda_s)$

T should be large (>10'000)

Stochastic growth rate sensitivities

$$\mathbf{N}_{t+1}^* = \mathbf{A}_t^* \mathbf{N}_t^* \qquad r_s^* = \frac{1}{T-u} \sum_{i=u}^{T} \log\left(\frac{N_{i+1}^*}{N_i^*}\right)$$

where $\mathbf{A}^*$ differs from $\mathbf{A}$, such that a demographic parameter θ_t is always $\theta_t + \Delta$, with Δ being small (e.g. 0.001)

$$S(\theta) = \frac{r_s^* - r_s}{\Delta}$$

Including environmental stochasticity

- We need a transition matrix for many, many years. Thus, we need demographic rates that are measured for many, many years. How can we do that?

- There are basically two ways :

  - 1. If demographic rates for many year are avaliable, draw at random with replacement from these rates

  - 2. If the mean (m) and the temporal variation (σ^2) of the rates is known, draw a value from a normal distribution with N(m, σ^2).

- The first option has the advantage that correlations among demographic rates can easily be maintained. Yet, it has the disadvantage that values from many years must be available, such that the temporal variability is well represented by these values

- The second option is the most approriate in many practical situations. Correlations among demographic rates can also be included, but estimates of temporal covariances among demographic rates are needed
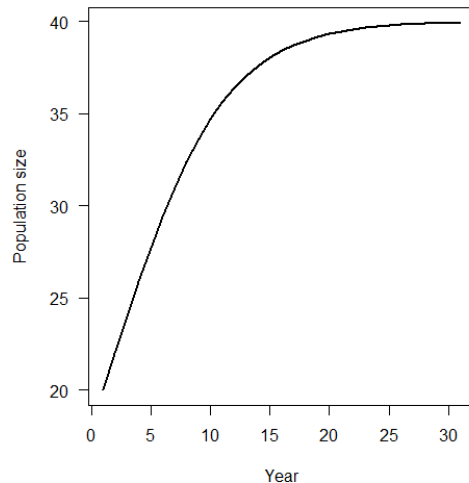
Including density-dependence

- One or several demographic rates (θ) change as a function of population size, e.g.:

$$\theta_t = \alpha + \beta N_t$$

- For some parameters scale transformations are needed (e.g. logit for probabilities) to ensure that realized values are within a reasonable range

- Project population size over time $\mathbf{N}_{t+1} = \mathbf{A}_{\mathbf{N}_t} \mathbf{N}_t$
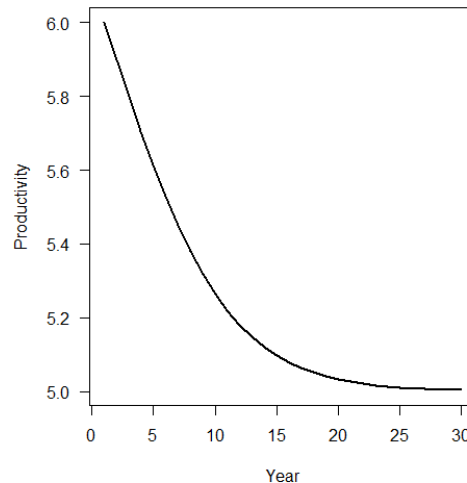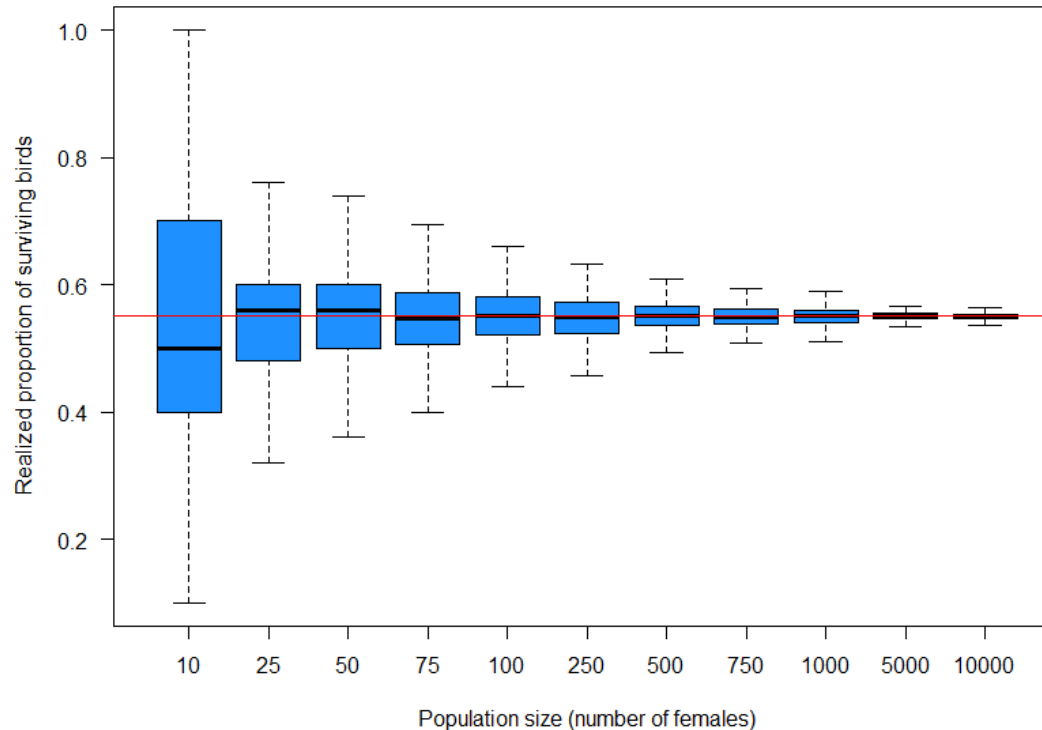
Including demographic stochasticity

- Demographic stochasticity originates from random variation in the realized number of births and deaths caused by the discrete nature of individuals

- Experience demographic stochasticity in the survival process:

    - Mean survival is 0.55

    - Start with a varying number of individuals, use Binomial model (coin flip) to determine how many survive

    - Calculate realized survival as the ratio of the number of survivors and the starting number of individuals

Including demographic stochasticity

- Demographic stochasticity originates from random variation in the realized number of births and deaths caused by the discrete nature of individuals

Including demographic stochasticity

- Deterministic population equations are replaced by stochastic ones:

$$N_{1y,t+1} = N_{1y,t}S_{juv}F_1 + N_{ad,t}S_{juv}F_a$$

$$N_{ad,t+1} = N_{1y,t}S_{ad} + N_{ad,t}S_{ad}$$

$$N_{1y,t+1} \sim Poisson\left(N_{1y,t}S_{juv}F_1 + N_{ad,t}S_{juv}F_a\right)$$

$$N_{ad,t+1} \sim Binomial\left(\left(N_{1y,t} + N_{ad,t}\right), S_{ad}\right)$$
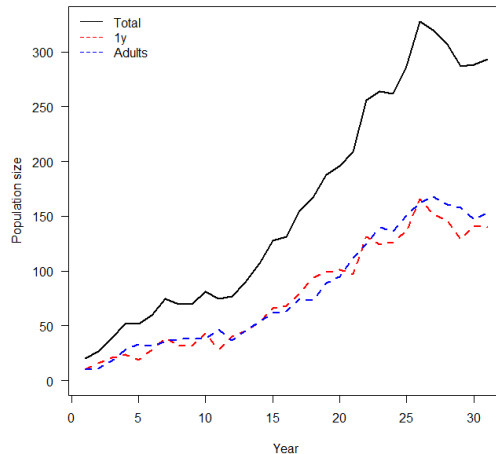
- Perform many simulations to get population growth rate

- Result depends on population size of the first year

Parameter uncertainty

- Usually are demographic rates estimated → subject to estimation uncertainty

- Simulations (usually bootstrap or Monte Carlo simulation) needed

- Pseudo-code to calculate the population growth rate:

  – Pick each demographic rate at random from a statistical distribution using known means and SDs

  – Construct a Leslie matrix using these generated demographic values

  – Calculate asymptotic population growth rate (dominant eigen value) and safe it

  – Repeat these steps many times

  – Summarize mean and SD of the growth rates

Next steps

- Show briefly how a matrix projection model is analysed in R

- Show how a matrix projection model is analysed with JAGS

  - Without uncertainty, without stochasticity

  - With uncertainty, without stochasticity

- Inclusion of demographic data in a matrix projection model

- Steps towards an integrated population model

- (You will) develop an IPM with demographic and environmental stochasticity

Exercises

- **Exercise 1**: Fit an IPM that includes demographic stochasticity

- **Exercise 2**: Fit an IPM in which adult survival is variable over time and juvenile survival shows a linear temporal trend

- **Exercise 3**: Fit an IPM that uses the Poisson distribution for the observation model of the state-space model

(always use the woodchat shrike data, and take the model `m3.jags` as a starting model and adapt)

Benefits of IPMs

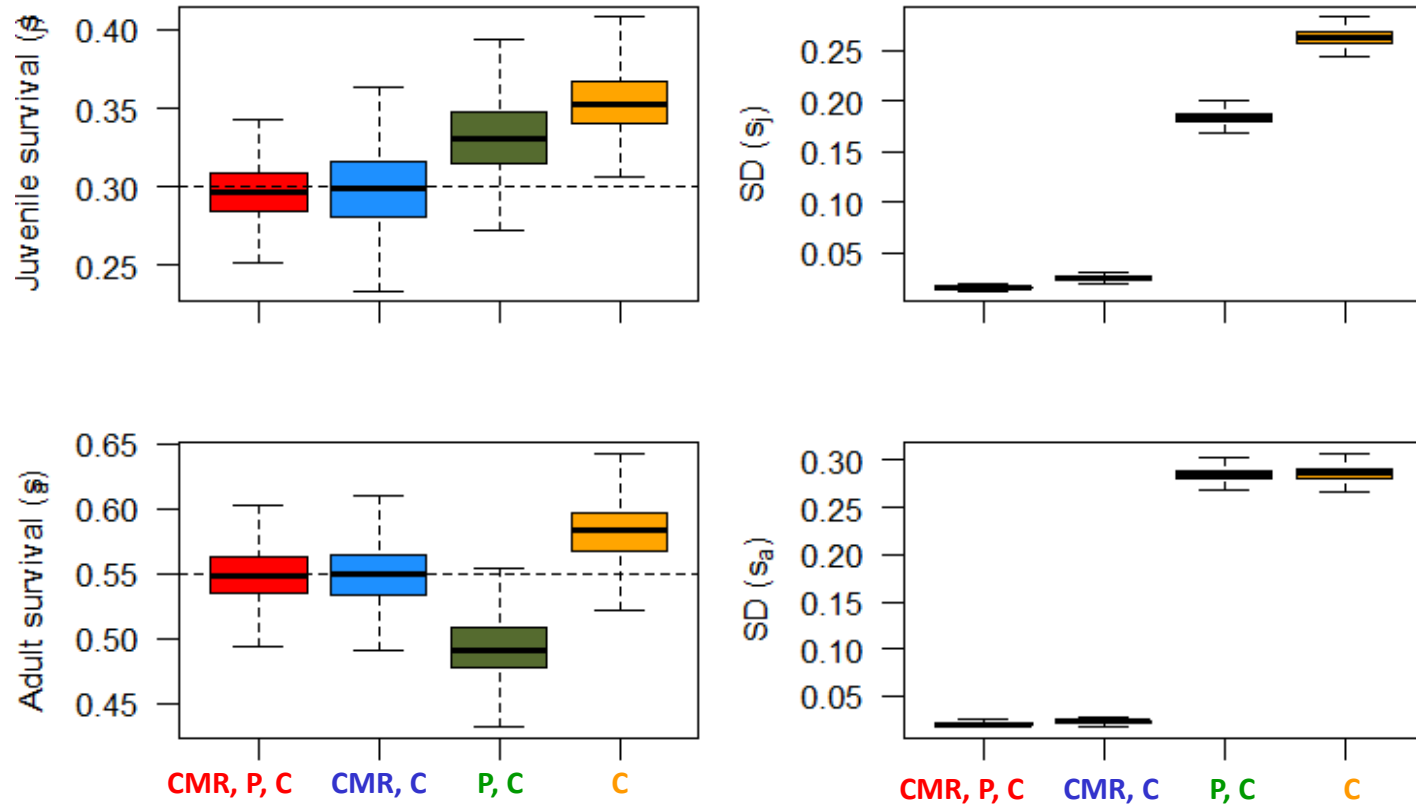- More efficient use of data: increased precision of parameters
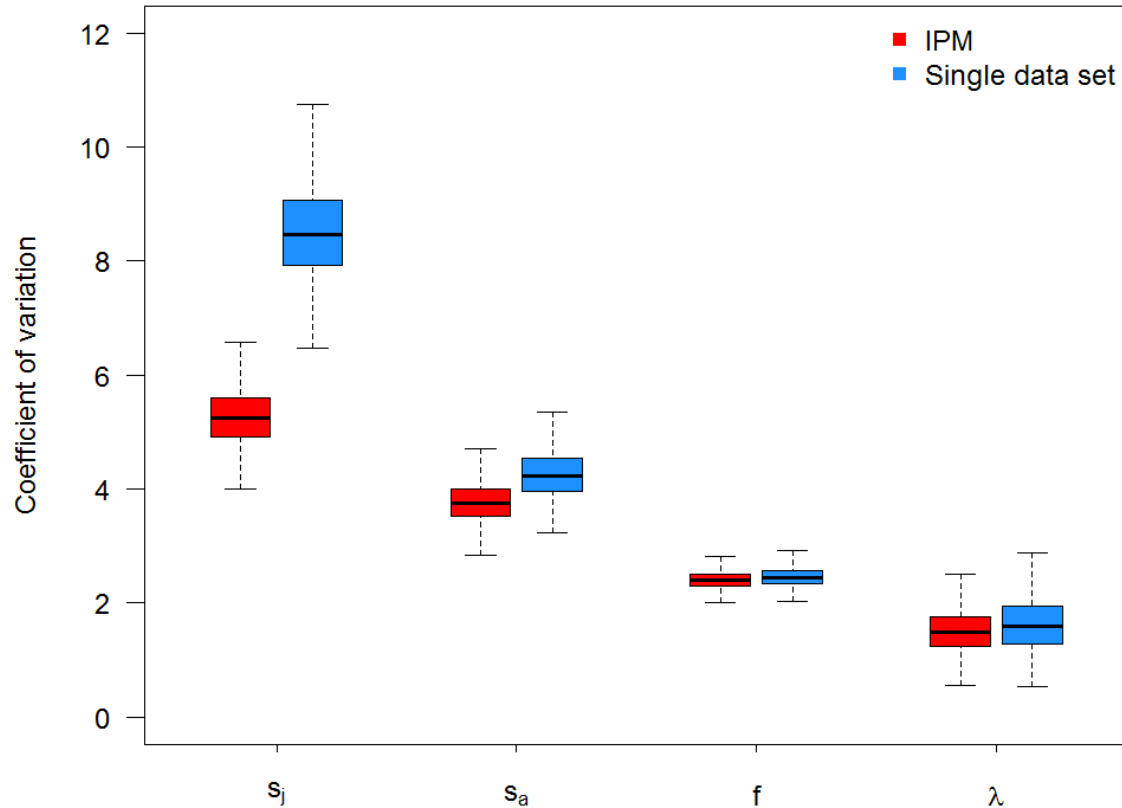
Benefits of IPMs

- More efficient use of data: increased precision of parameters
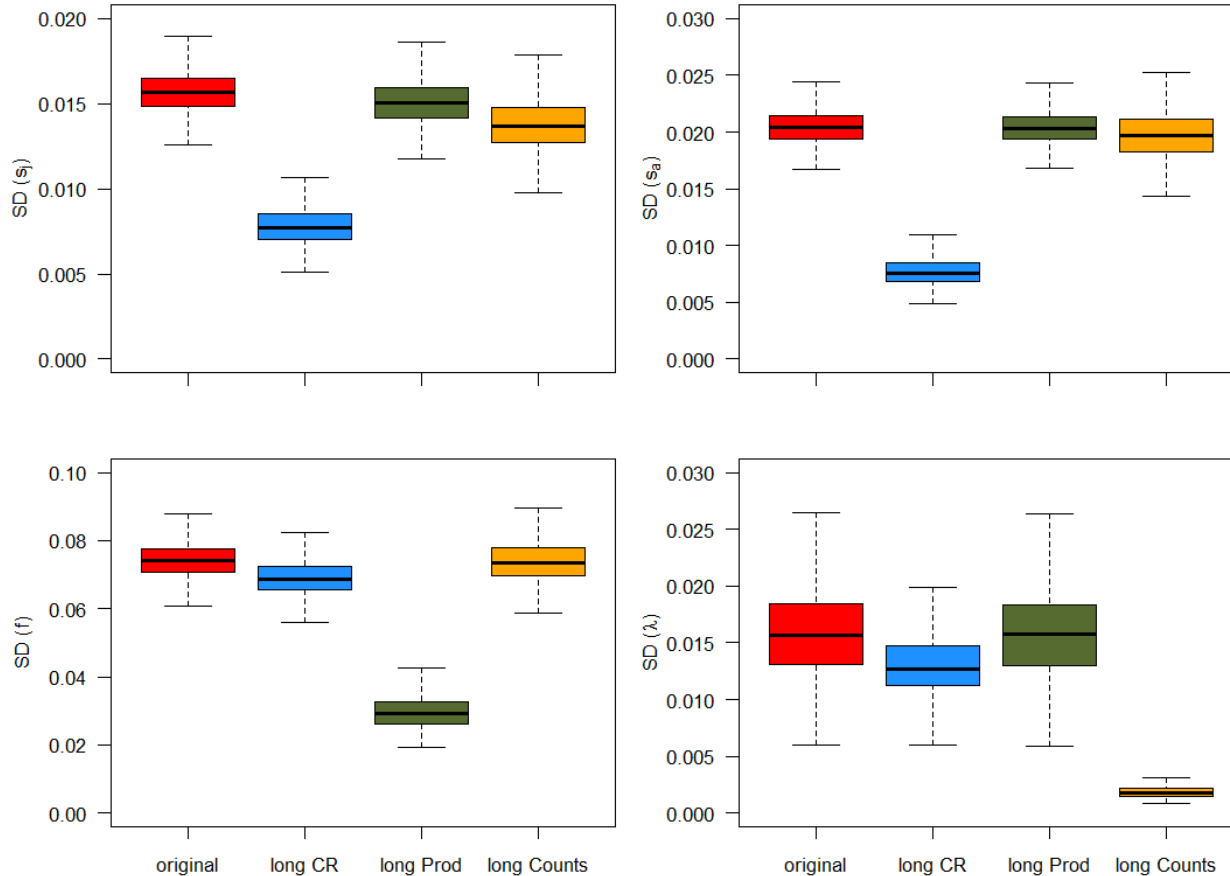
Benefits of IPMs

- More efficient use of data: increased precision of parameters

- Ability to estimate a demographic process for which we have no explicit data

Benefits of IPMs

- More efficient use of data: increased precision of parameters

- Ability to estimate a demographic process for which we have no explicit data

- **Exercise 4**: Fit an IPM to the shrike data. Assume that only capture-recapture data of the adults [CH.A.marray], counts and productivity data are available.

(always use the woodchat shrike data, and take the model `m3.jags` as a starting model and adapt)

Benefits of IPMs

- Ability to estimate a demographic process for which we have no explicit data: how far can we go?
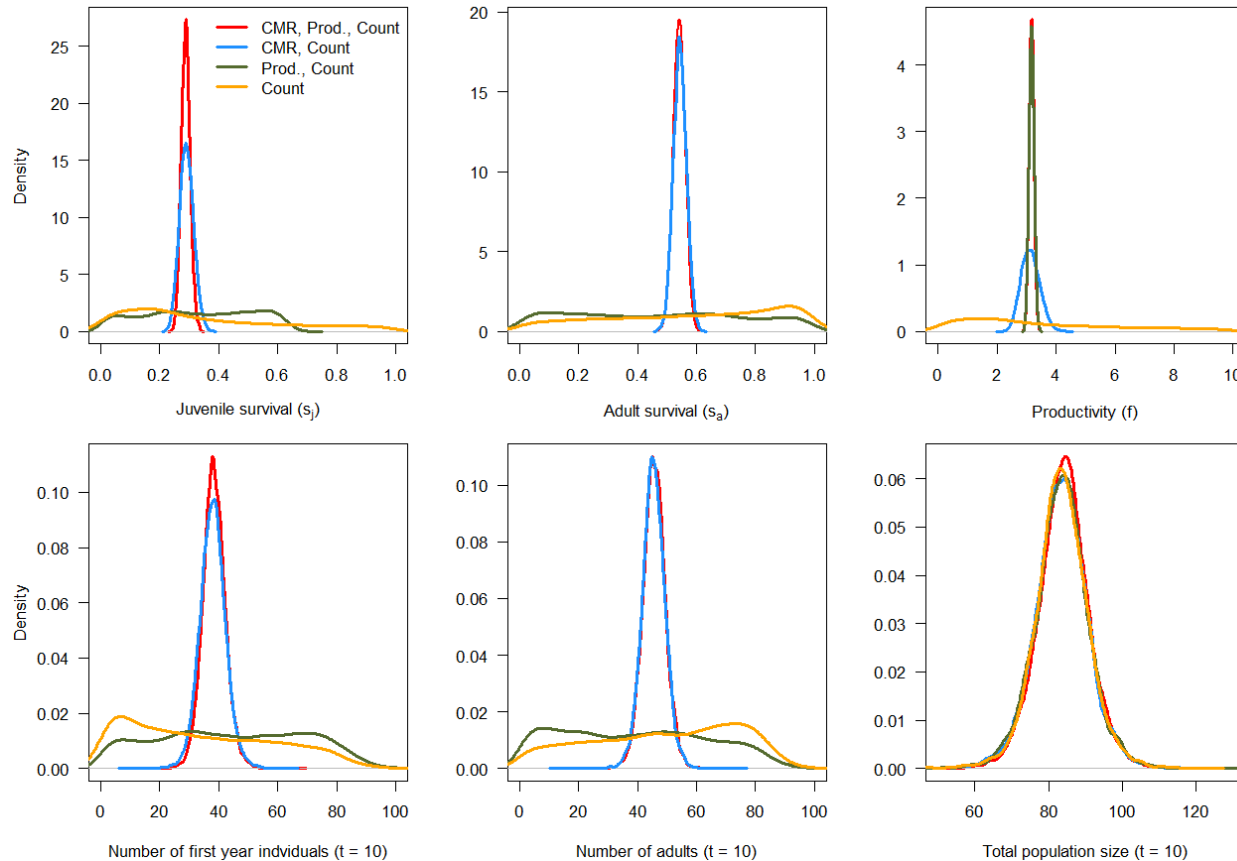
Benefits of IPMs

- Ability to estimate a demographic process for which we have no explicit data: how far can we go?

Benefits of IPMs

- Ability to estimate a demographic process for which we have no explicit data: how far can we go?
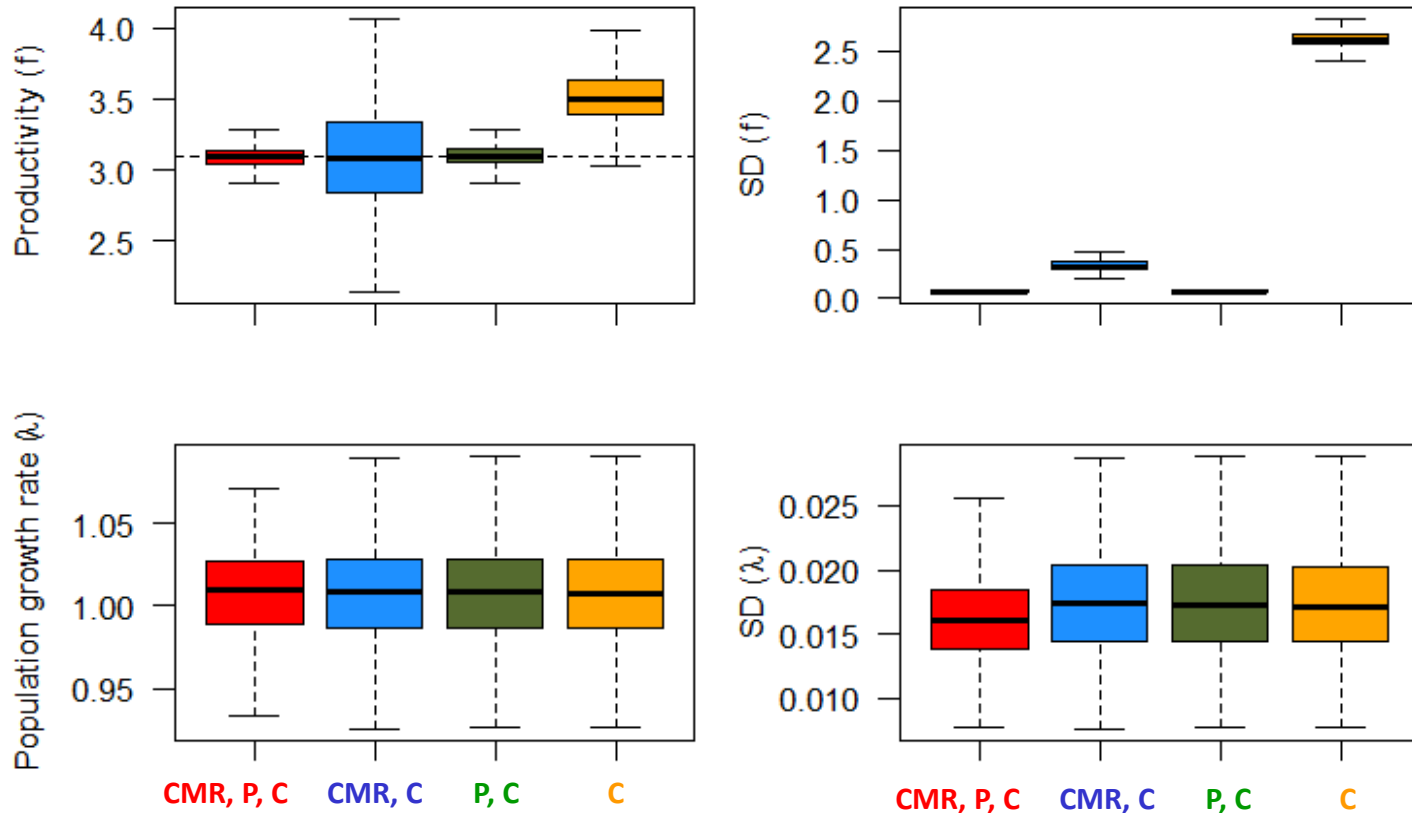
Benefits of IPMs

- Ability to estimate a demographic process for which we have no explicit data: how far can we go?

- Be careful: some parameters may not be identifiable!

- How to spot non-identifiability?

    - Not always easy

    - Compare prior and posterior

    - Simulation

    - Inspect correlation between parameters

- Parameters may be biased if lack of fit in the other "parts" of the model is present

Benefits of IPMs

- More efficient use of data: increased precision of parameters

- Ability to estimate a demographic process for which we have no explicit data

- Allows insights into the population structure

Another approach for developing an IPM

Three basic steps:

1. Set up a population model that links demographic rates with population size

2. Write the likelihood for the single data sets

3. Develop the joint likelihood and analyse it

1. Set up a population model

Pre-breeding census
Female-based
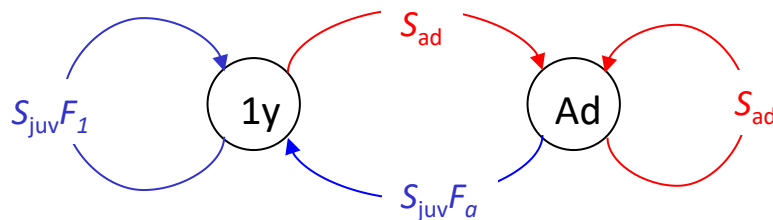
Population sizes:

$$N_{1y,t+1} = N_{1y,t}S_{juv}F_1 + N_{ad,t}S_{juv}F_a$$
$$N_{ad,t+1} = N_{1y,t}S_{ad} + N_{ad,t}S_{ad}$$

To include demographic and environmental stochasticity, we write

$$N_{1y,t+1} \sim \text{Poisson}\left(N_{1y,t}S_{juv,t}F_{1,t} + N_{ad,t}S_{juv,t}F_{a,t}\right)$$

└───── generates an integer number between 0 and ∞

$$N_{ad,t+1} \sim \text{Binomial}\left(N_{1y,t} + N_{ad,t}, S_{ad,t}\right)$$

generates an integer number between 0 and $N_{1y,t} + N_{ad,t}$

Parameters
S: survival probability
F: productivity
N_{1y}: population size of 1y
N_{ad}: population size of adults

2. Likelihood for the different data sets

A. Counts: state-space model

State process equations:

$$N_{1y,t+1} \sim \text{Poisson}\left(N_{1y,t} S_{1,t} F_{1,t} + N_{ad,t} S_{1,t} F_{a,t}\right)$$

$$N_{ad,t+1} \sim \text{Binomial}\left(N_{1y,t} + N_{ad,t}, S_{ad,t}\right)$$

Observation process equation:

$$y_t \sim \text{Normal}\left(\left(N_{1y,t} + N_{ad,t}\right), \sigma_y^2\right)$$

Likelihood:

$$L_{SS}\left(\mathbf{N}, \mathbf{S_{juv}}, \mathbf{S_{ad}}, \mathbf{F}, \sigma_y^2 \middle| \mathbf{y}\right) = L_1\left(\mathbf{N_1}\right) \times L_O\left(\mathbf{N}, \sigma_y^2 \middle| \mathbf{y}\right) \times L_S\left(\mathbf{N}, \mathbf{S_{juv}}, \mathbf{S_{ad}}, \mathbf{F}\right)$$

Parameters
S: survival probability
F: productivity
N_{1y}: population size of 1y
N_{ad}: population size of adults
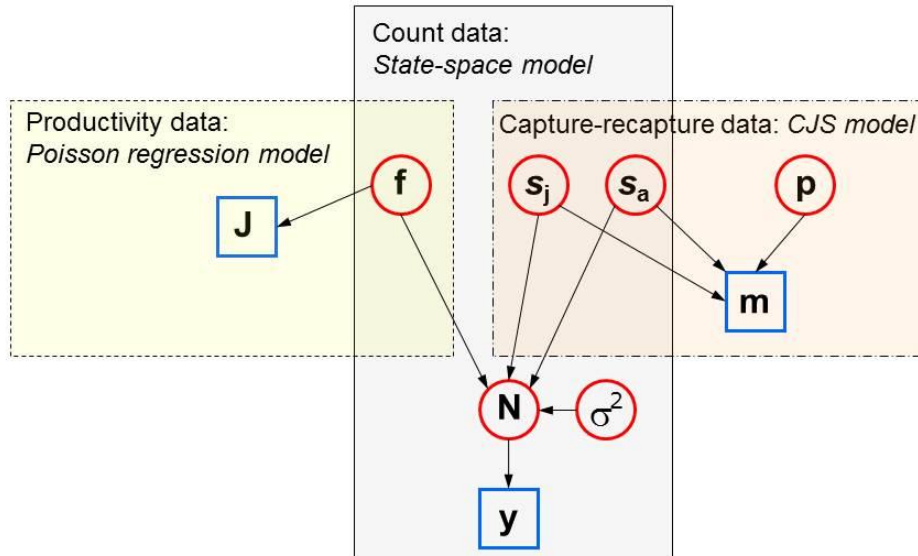σ_y^2: census/residual error
Data
y: counts

B. Capture-recapture data

- As introduced earlier
- Better use multinomial likelihood

Likelihood:

$$L_{CJS}\left(\mathbf{S_{juv}}, \mathbf{S_{ad}}, \mathbf{p} \mid \mathbf{m}\right)$$

C. Data on reproductive output

- Poisson regression model

$$\mathbf{J} \sim Pois(\mathbf{F})$$

Likelihood:

$$\mathbf{L_P}\left(\mathbf{F} \mid \mathbf{J}\right)$$

Parameters
S: survival probability
p: recapture probability
F: productivity
Data
m: m-array
J: # nestlings

3. Joint likelihood

- Assume independence among data sets

- Joint likelihood

$$L_{IPM}\left(\mathbf{N},\mathbf{S_{juv}},\mathbf{S_{ad}},\mathbf{F},\mathbf{p},\sigma_y^2\middle|\mathbf{y},\mathbf{m},\mathbf{J}\right)=L_1\left(\mathbf{N_1}\right)\times L_O\left(\mathbf{N},\sigma_y^2\middle|\mathbf{y}\right)\times L_S\left(\mathbf{N},\mathbf{S_{juv}},\mathbf{S_{ad}},\mathbf{F}\right)\times$$

$$L_{CJS}\left(\mathbf{S_{juv}},\mathbf{S_{ad}},\mathbf{p}\middle|\mathbf{m}\right)\times L_p\left(\mathbf{F}\middle|\mathbf{J}\right)$$

- Analyse the joint likelihood

Parameters
S: survival probability
p: recapture probability
F: productivity
N_{1y}: population size of 1y
N_{ad}: population size of adults
σ_y^2: census/residual error

Data
y: counts
m: m-array
J: # nestlings

3. Joint likelihood

Graphical relationship between data and parameters:
(Directed acyclic graph (DAG) without priors)

Parameters
S: survival probability
p: recapture probability
F: productivity
N_{1y}: population size of 1y
N_{ad}: population size of adults
σ^2_y: census/residual error
Data
y: counts
m: m-array
J: # nestlings

Some comments on IPMs

Where is the joint likelihood in the BUGS code?

- Is defined implicitly by a series of simpler, local relationships
- Practically, the same name of parameters in the different likelihoods ensures the integration

Independence assumption:

- Required because of the joint likelihood
- Too precise results expected, if not fulfilled
- Simulation study to see effects of the violation (e.g. Abadi et al. 2010)
- Different sampling protocols for the different data sets needed (not necessarily different individuals)
- Key for modeling dependent data sets: model explicitly the fate of all marked and unmarked individuals in the population (Lee at al. 2015, Béliveau et al. in prep.)
- Common demography assumption is more likely to be violated, if data sets stem from different populations

Some comments on IPMs

What happens if the common demography assumption is not fulfilled?

$S_j = 0.3$	$S_j = 0.3$	$S_j = 0.3$	$S_j = 0.3$	$S_j = 0.3$
$S_a = 0.55$	$S_a = 0.55$	$S_a = 0.55$	$S_a = 0.64$	$S_a = 0.55$
$F = 3.1$	$F = 3.1$	$F = 3.1$	$F = 3.1$	$F = 3.7$
$\lambda = 1.015$	$\lambda = 1.015$	$\lambda = 1.015$	$\lambda = 1.105$	$\lambda = 1.105$
↓	↓	↓	↓	↓
CMR	**P**	**Count**	**Count**	**Count**

IPMs:

CMR, P, Count   CMR, Count

CMR, P, Count   CMR, Count

CMR, P, Count   CMR, Count

Some comments on IPMs

What happens if the common demography assumption is not fulfilled?
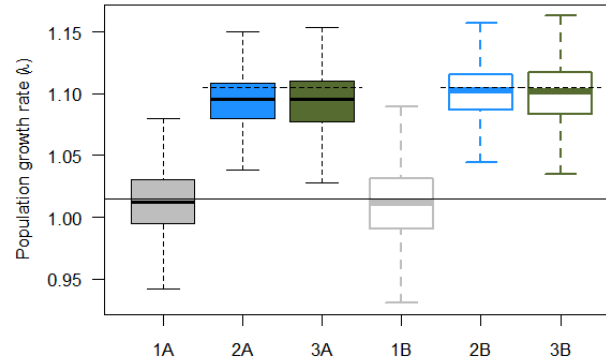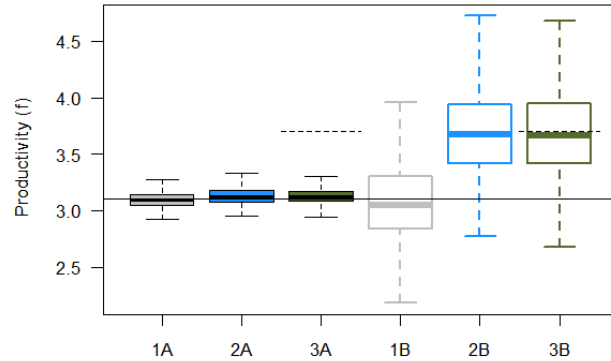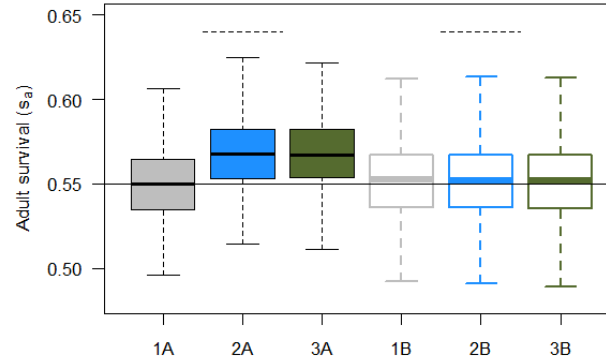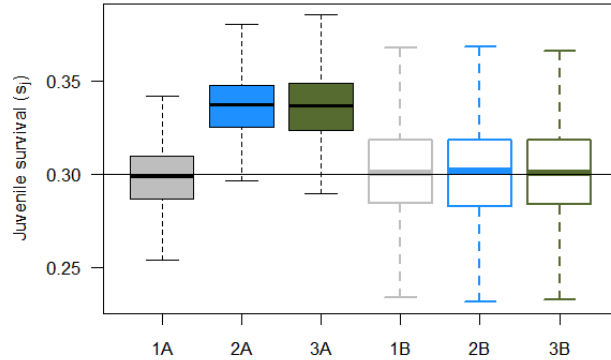
1A: CMR, P, Count
2A: CMR, P, Count
3A: CMR, P, Count

1B: CMR, Count
2B: CMR, Count
3B: CMR, Count

Some comments on IPMs

How accurate is population size estimated with an IPM?

- Depends on the type of error that we have while counting the individuals

- Are counts correct on average (both types of error [double counting, missing individuals] occur) ?

- Only imperfect detection ?

- Small simulation study to visualize effects

Some comments on IPMs

How accurate is population size estimated with an IPM?

- Simulate 3 independent populations

- Sample from population 1 CMR data, from population 2 productivity data, from population 3 count data

- Different ways to get count data:

  - 1. Normal distribution

  - 2. Binomial distribution, with constant detection probability

  - 3. Binomial distribution, with temporally variable detection probability

  - 4. Binomial distribution, with linearly increasing detection probability

- Analyse data sets with IPMs that has the standard observation model: $C_t \sim N(N_t, \sigma)$

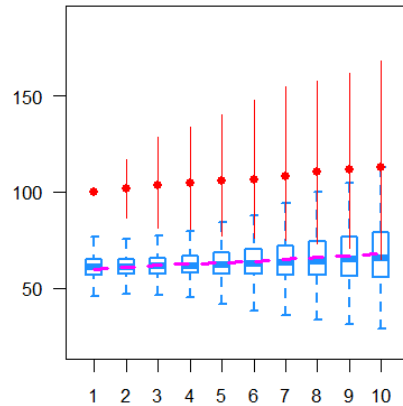- Also use an IPM without productivity data

Some comments on IPMs
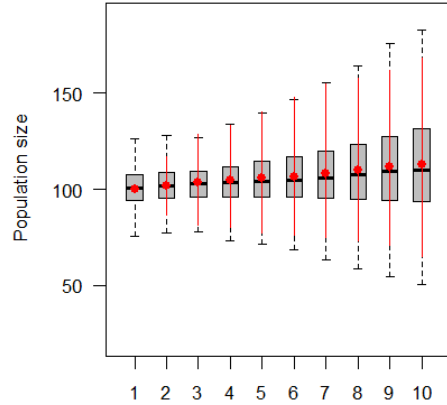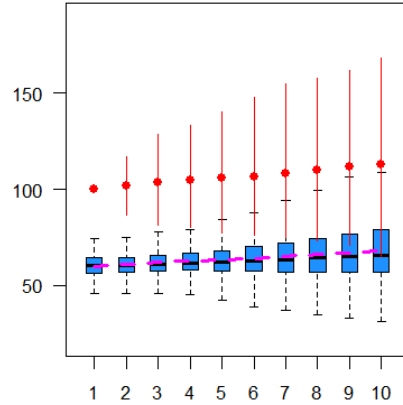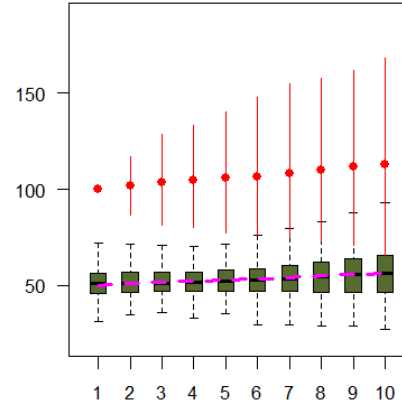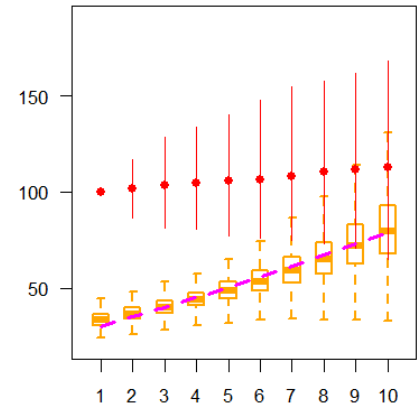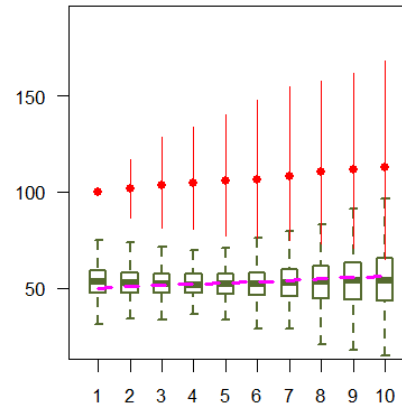
Some comments on IPMs
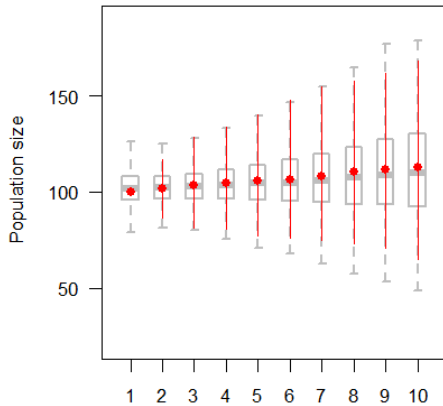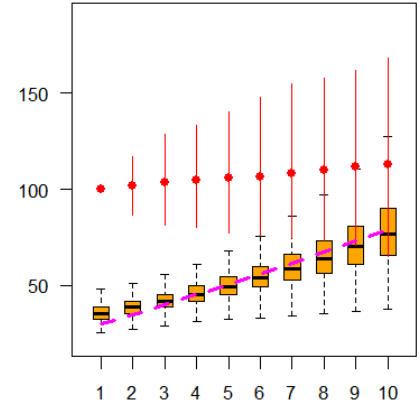
Some comments on IPMs

Goodness-of-fit test:

- Difficult (but important), area of current research (one recent paper: Besbeas & Morgan 2014, Methods in Ecology & Evolution)

- Assess the fit of each submodel, by e.g. posterior predictive checks

- Difficulty is the judgement of an «overall» fit

- Pragmatic approach: assess the GOF of the single data sets with classical tests

- Parameters which are estimated «without» explicit data are expected to be particularly biased if the single data models do not fit

Some comments on IPMs

Mode of analysis

- Bayesian
  - is the most flexible
  - due to the flexible and transparent BUGS language, Bayesian is easier for a typical ecologist
- Frequentist
  - Advantages:
    - faster computation
    - model selection (AIC)
  - Disadvantages:
    - Normal distributions for the state-process required
    - Kalman Filters needed, seems more complicated (at least to me)

Exercises

- **Exercise 5**: Fit an IPM to the woodchat shrike data. Assume now, that the shrikes start to breed only when they are 2 years old.

- **Exercise 6**: Fit an IPM to the woodchat shrike data. Still use the pre-breeding census model, but change the model in such a way that we monitor in addition the number of fledglings.

(always use the woodchat shrike data, and take the model `m3.jags` as a start and adapt)