

# Quantitative methods for population dynamics

## Introduction

Olivier Gimenez, Sarah Cubaynes, Aurélien Besnard  
Coline Canonne, Thierry Chambert, Thibaut Couturier, Valentin Lauret

# Montpellier and its charms





# CORONAVIRUS DISEASE (COVID-19)



zoom

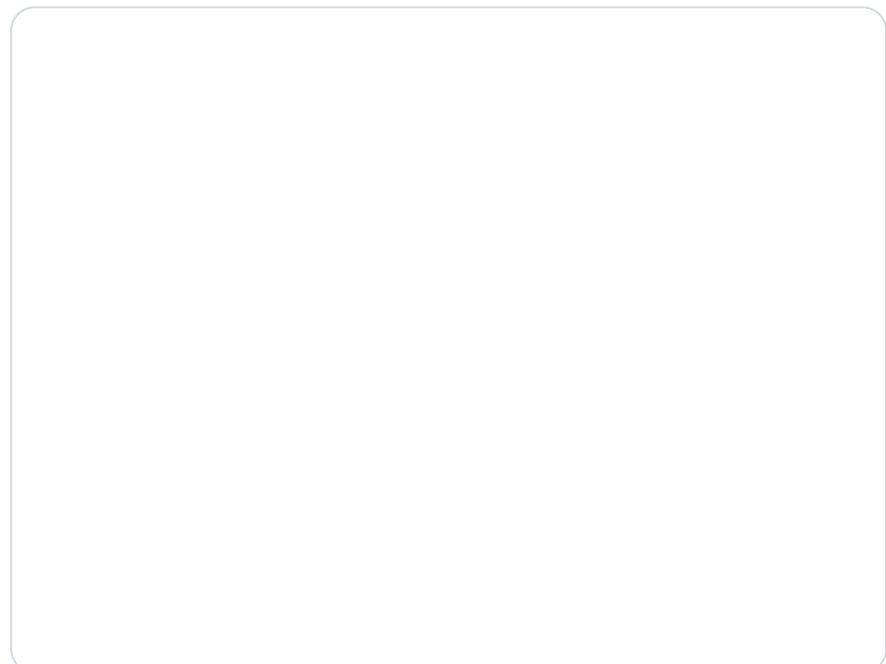
# How it all started

Olivier Gimenez 🙋

@oaggimenez



📢 Together with @SarahCubaynes and @abesnardEPHE we will be giving a 2-day introductory workshop March 21-22 on quantitative methods for population dynamics in R #rstats Join us 😊 It's online and free of charge 😊 You just need to register forms.gle/HkSYYm8MDkT96... Please RT 😊



2:42 PM · Oct 11, 2021



# Where we're at

Questions Réponses 2542 Paramètres



## Introductory workshop on quantitative methods for population dynamics in R. March 21-22, 2022.

Quantifying population dynamics is crucial for the conservation and management of animal and plant species. This workshop deals with the analysis and modelling of population dynamics. In this introductory workshop, we will cover population projection matrix models, the estimation of demographic parameters (e.g. survival, dispersal) using capture-recapture models and the estimation of population density and abundance using capture-recapture, N-mixture and distance sampling models. The format of the workshop will be a combination of lectures and practical sessions in R with packages popbio, RMark and unmarked.

The workshop is online, free of charge, and will be video recorded.

Olivier Gimenez, Sarah Cubaynes, Aurélien Besnard.

# Quantitative methods for population dynamics

# Population dynamics

# Processes







# Credits and inspiration

- Past workshops on capture-recapture models with Roger Pradel, Rémi Choquet and Jean-Dominique Lebreton.
- Past workshop on Population dynamics back in 2007.
- Lectures for Master students with Olivier Duriez.

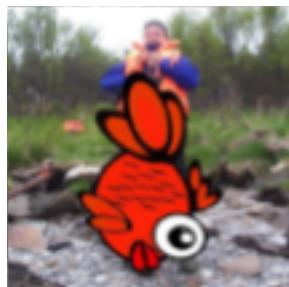
# The team - lecturers



Sarah Cubaynes



Aurélien Besnard



Olivier Gimenez

# The team - helpers



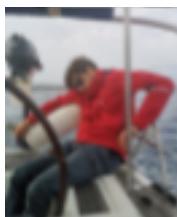
Coline Canonne



Thibaut Couturier



Thierry Chambert



Valentin Lauret

# On our plate

- Estimating abundance and density of populations
  - Conventional methods and the detectability issue
  - Capture-recapture, distance sampling and N-mixture
- Estimation of demographic parameters
  - Survival estimation from capture-recapture data
  - Using covariates to explain variation
  - Inference about dispersal
- Population projection matrix models
  - Count-based projection models
  - Age- and stage-structured models
  - Sensitivity analyses

# Philosophy of teaching

- Lots of attendees, with huge heterogeneity in knowledge of statistical analyses, methods in population dynamics, and R.
- It is our hope that everyone will find something to take home.
- We've packed a lot of things in two days.
- We do not expect you to digest everything.
- All material (including videos) on website <https://oliviergimenez.github.io/popdyn-workshop/>.
- Feel free to play around with material while we walk through it, and afterwards.
- The workshop is organized in modules, each module is a combination of lectures and live demos.

# The way we will interact with each other

- We'll be streaming lectures and live coding demos on Youtube.
- Everything is video recorded.
- The recordings will be made available on the website.
- Questions and answers via the chat on Youtube. You need to have a YouTube account to use the chat.
- We'll do our best to answer your questions, but with > 2500 attendees, we'll probably miss some of them. Apologies in advance.

# Estimating abundance

Aurélien Besnard for the team

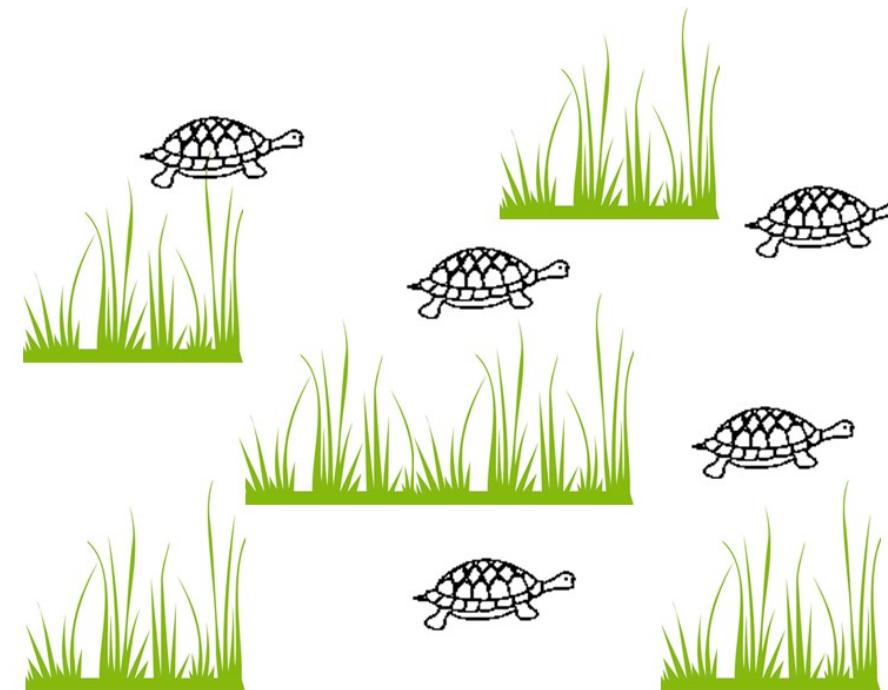
last updated: 2022-03-18

# Population dynamics

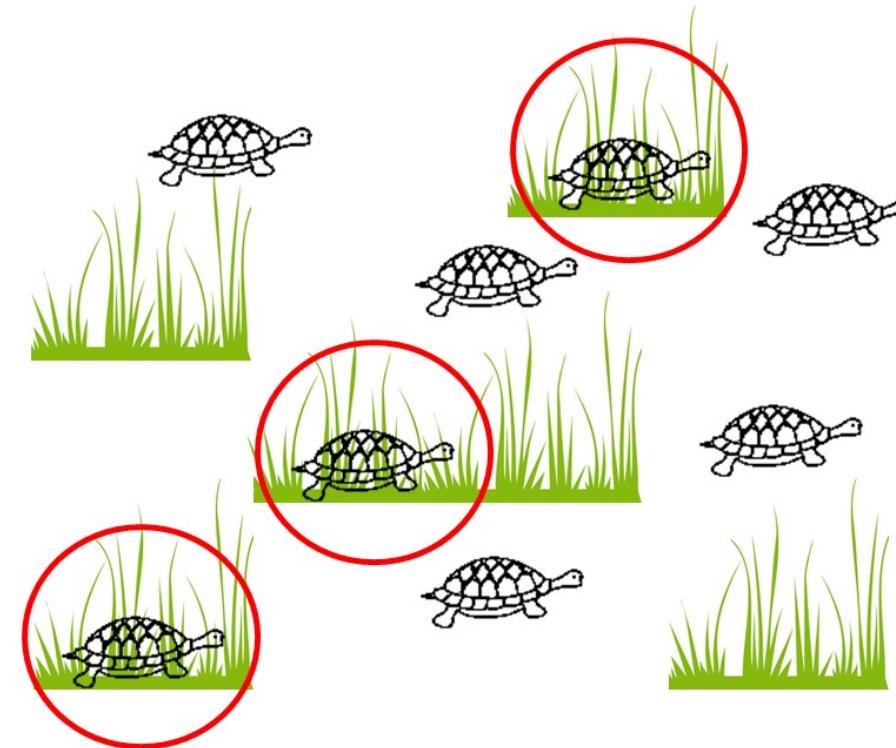


# Background - Detectability issues

Counting animals or plants in the field



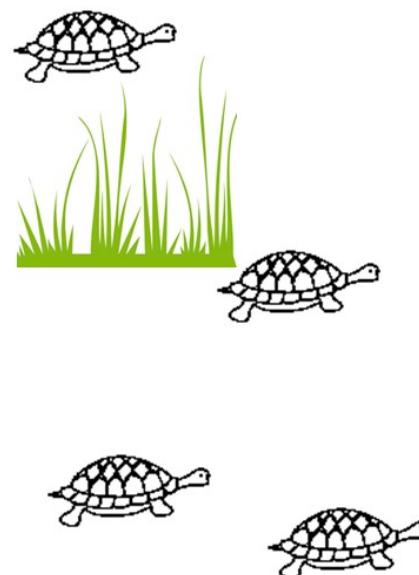
# Background - Detectability issues



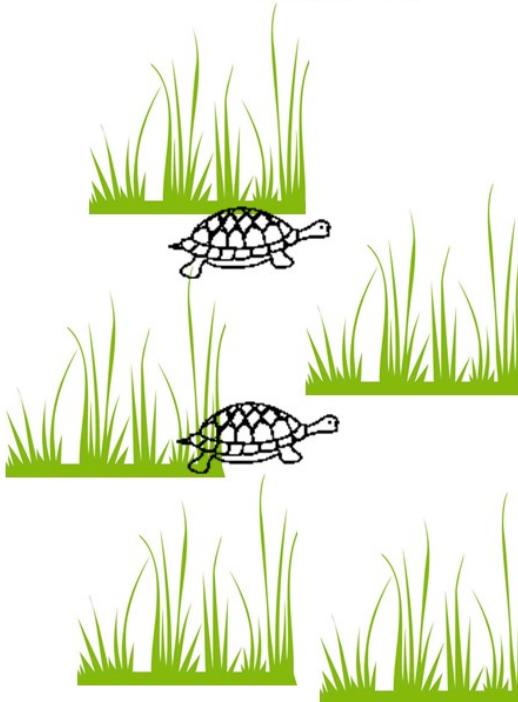
Detectability is rarely (never) exhaustive

# Background - Detectability issues

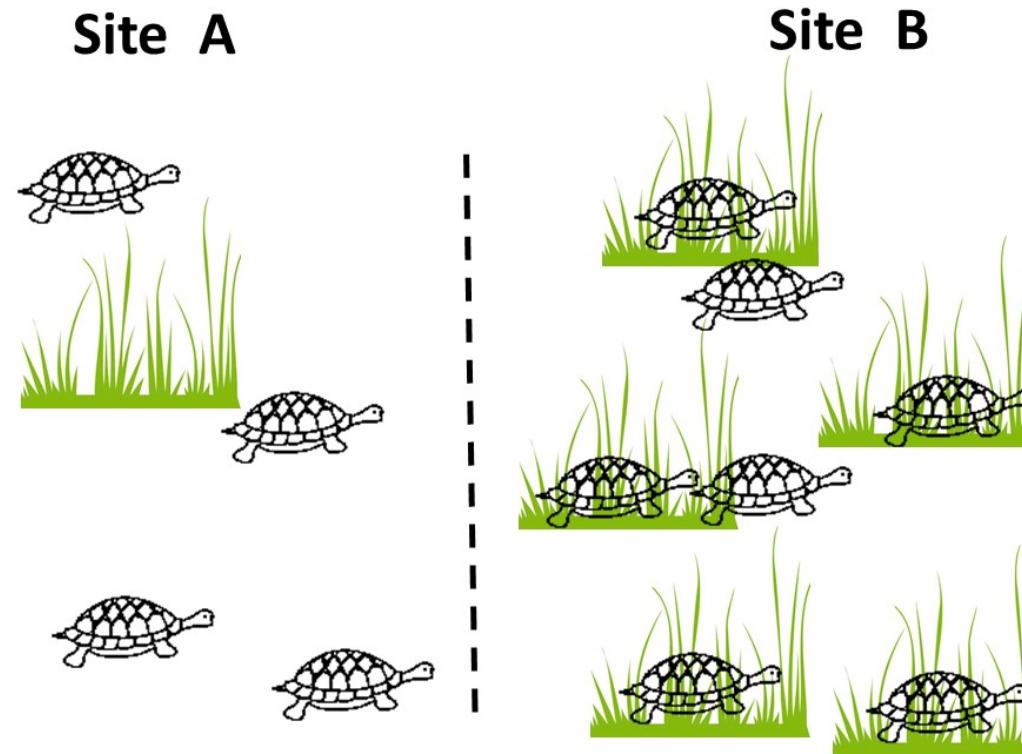
**Site A**



**Site B**

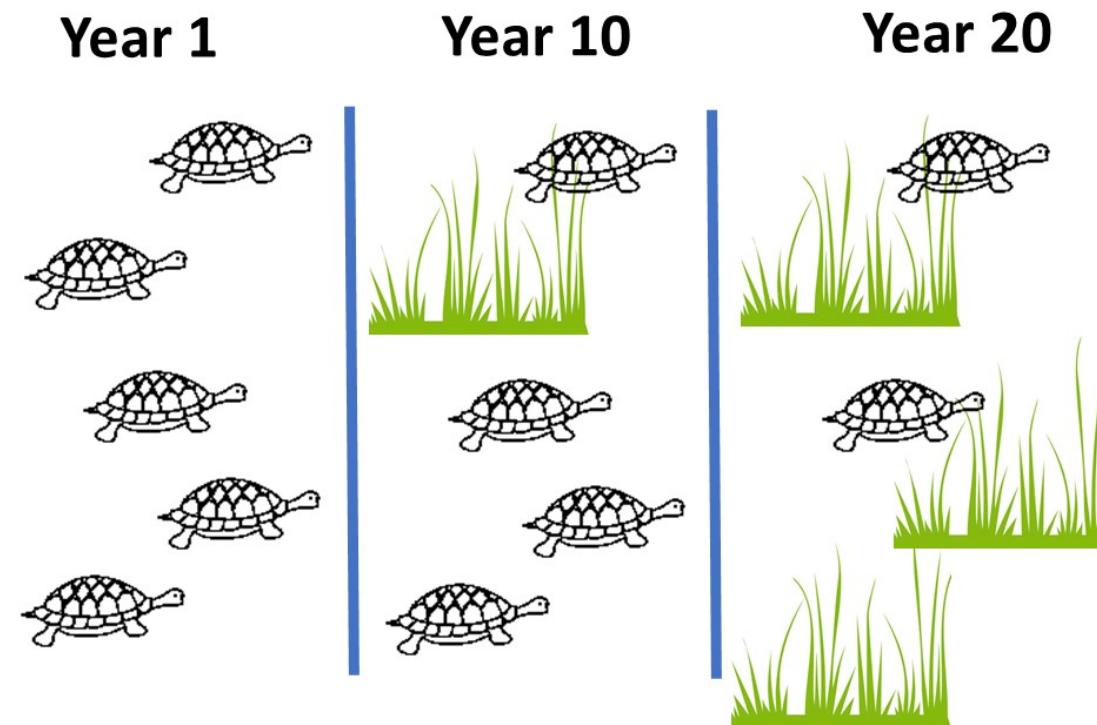


# Background - Detectability issues

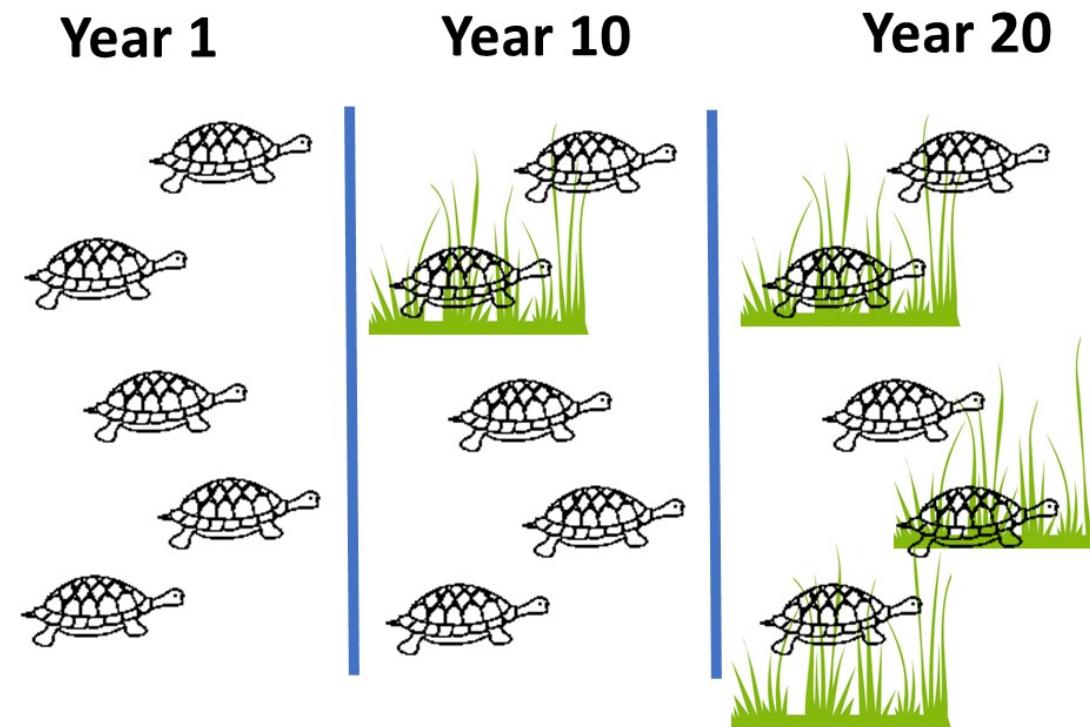


Detectability (usually) varies over space

# Background - Detectability issues



# Background - Detectability issues



Detectability (usually) varies over time

# Background - Detectability issues

- If  $N$  is the abundance of a particular species at a particular study area
- And  $C$  the number of individuals counted in the field in this area
- We have  $E(C) = pN$
- Where  $p$  is the probability of counting an individual in the study area

# Background - Detectability issues

- Eventually population abundance is given by:

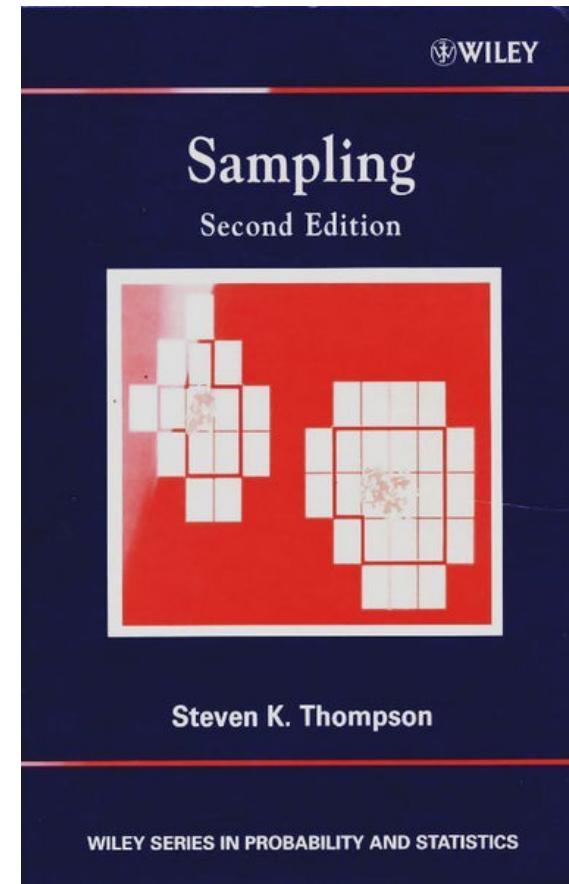
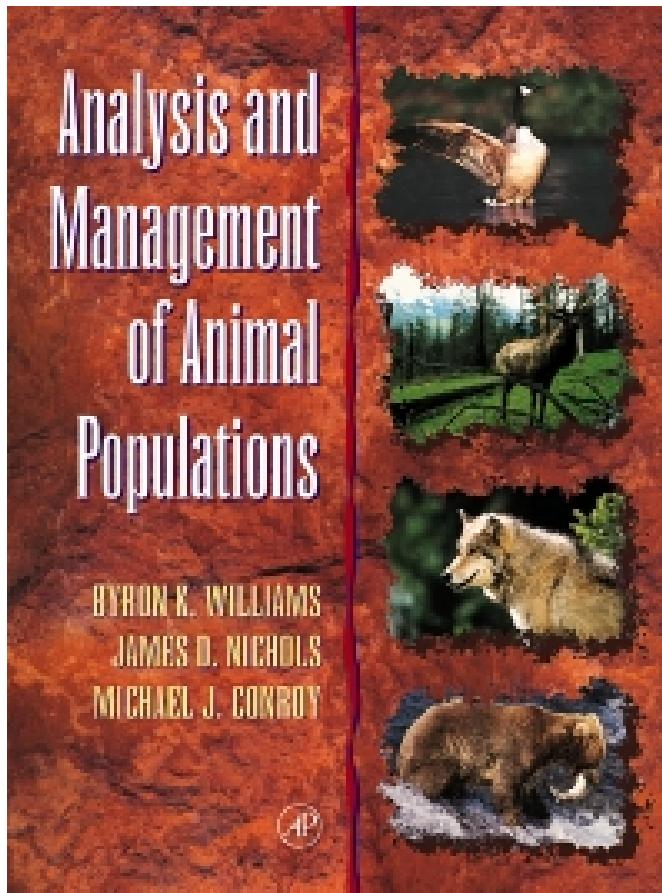
$$N = \frac{C}{\alpha \times p}$$

- $\alpha$  being the fraction of the study area that is sampled
- $p$  needs to be estimated to provide unbiased estimates
- And allows inference regarding abundance variations over time and space

# Sampling design

- How to select sampling units in the study area to obtain unbiased estimates of abundance is a complex task
- Need to define the statistical population of sample units and rely on some random selection process of the units to survey
- Several sampling designs exist: Random, systematic, stratified, spatially balanced, etc
- Not covered in this course

# Some useful reading on sampling design



# Issue 1: Estimating abundance when $p < 1$

- Several methods exist, depending on whether population is "closed" or "open"
- Closed populations: Capture-recapture, Distance sampling, N-mixture
- Open populations: Capture-recapture, open N-mixture
- In this course:
  - Capture-recapture for closed populations
  - Distance sampling
  - N-mixture for closed populations

# What does "closed" populations mean?

- Demographic closure:
  - no birth
  - no death
- Geographic closure:
  - no immigration
  - no emigration
- Between first and last field sessions
- Implies to work over short time intervals
- 'Short' depending on species life history traits

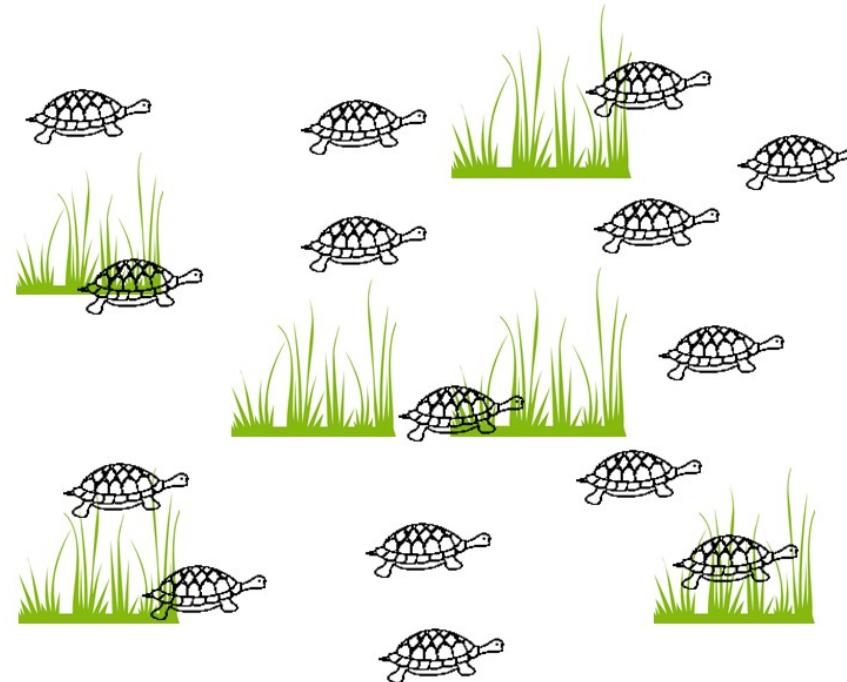
# What does "closed" populations mean?



# Capture-recapture for closed populations

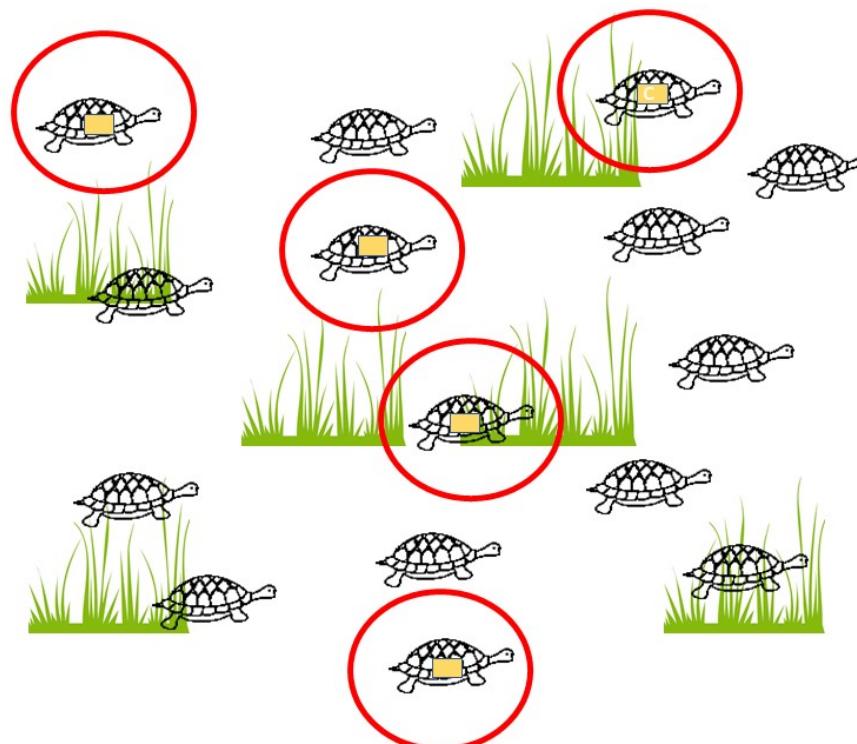
# Capture-recapture methods for closed populations

Sample in a closed population



# Capture-recapture methods for closed populations

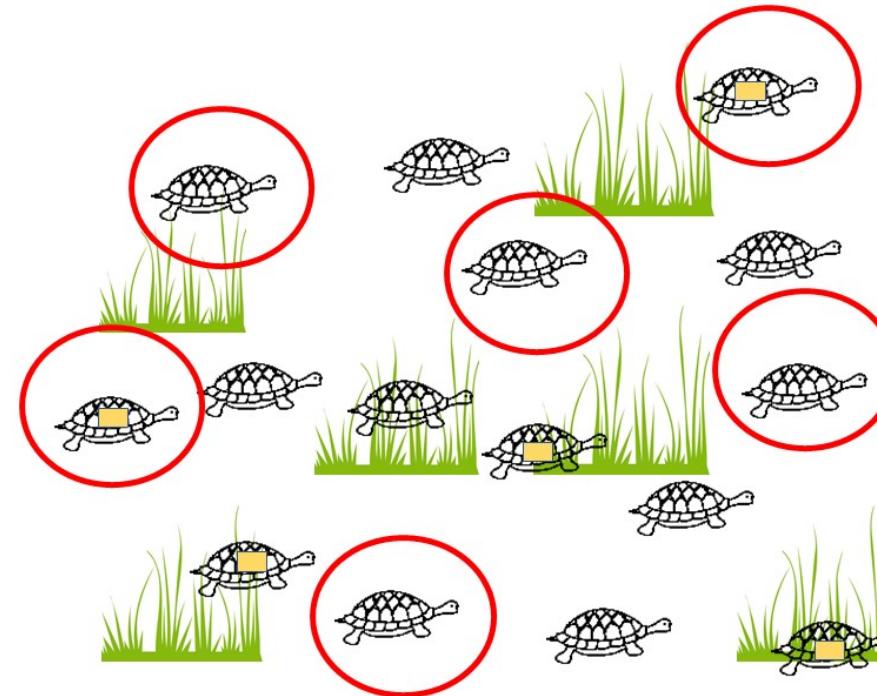
Capture, mark and release some individuals (noted  $n_1$ )



$p$  is typically below 1, not all present individuals are captured (here  $p \approx 1/3$ )

# Capture-recapture methods for closed populations

Make a second visit, capture a second independent sample of individuals



Among newly captured individuals, noted  $n_2$ , some are marked, noted  $m$ , others aren't

# Capture-recapture methods for closed populations

- The proportion of marked individuals after the first sample is  $\frac{n_1}{N}$
- The proportion of marked individuals in the second sample is  $\frac{m}{n_2}$
- Assuming sampling sessions are independent (individuals have equal detection probability whether they have already been captured or not)
- Then the proportion of marked individuals in the second sample,  $m$ , is the proportion of marked individuals in the entire population:

$$\hat{N} = \frac{n_1 \times n_2}{m}$$

# Capture-recapture methods for closed populations

Another way of writing the Lincoln-Petersen index

$$n_{10} = N \times p_1 \times (1 - p_2)$$

$$n_{01} = N \times (1 - p_1) \times p_2$$

$$n_{11} = N \times p_1 \times p_2$$

with  $p_1$  and  $p_2$  the capture probability at first and second session respectively

3 equations, 3 unknown values  $N, p_1, p_2, \dots$ , no difficulties

# Capture-recapture methods for closed populations

Rearranging the equations also leads to:

$$N = \frac{n_{10} + n_{01} + n_{11}}{1 - (1 - p_1) \times (1 - p_2)}$$

or

$$N = \frac{n_{10} + n_{01} + n_{11}}{p^*}$$

with  $p^*$  the probability to be captured at least once and  $n_{10} + n_{01} + n_{11}$  the total number of different individuals captured

*Remember*

$$N = \frac{C}{p}$$

# The Lincoln-Petersen index

Petersen (fishes 1894)



Carl George Johannes Petersen

Lincoln (birds 1930)

# Lincoln-Petersen index requires individual identification

- Usually implies some marking
- Group marking can also works

# Lincoln-Petersen index requires individual identification

Physical alteration of individuals



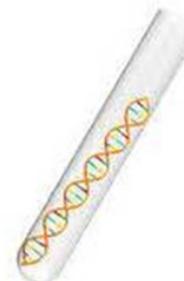
# Lincoln-Petersen index requires individual identification

Add some tags



# Lincoln-Petersen index requires individual identification

Natural marks (including DNA)

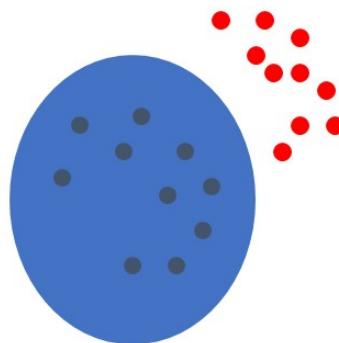


k6293533 fotosearch.com ©

# Lincoln-Petersen assumptions 1

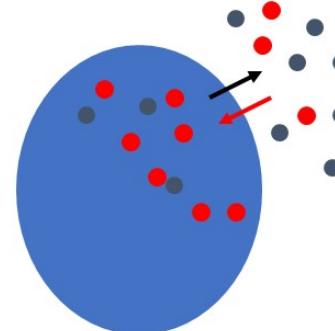
Population is closed demographically and geographically

Session 1



All individuals in the study area are marked  
 $N = n_1 = 10$

Session 2



All individuals in the study area are marked  
 $N = n_2 = 10$   
 $M = 3/10$

Population is estimated to 30 individuals (supra-population ?)

# Lincoln-Petersen assumptions 2

Population is closed demographically and geographically

**All individuals are equally likely to be caught in each sample**

$$n_1 = N \times p_1, n_2 = N \times p_2, m = N \times p_1 \times p_2$$

# Lincoln-Petersen assumptions 3

Population is closed demographically and geographically

All individuals are equally likely to be caught in each sample

**Marks are not lost (or unobserved)**

$m$  is under-estimated and then  $\hat{N}$  over-estimated

# Lincoln-Petersen and then ?

Population is closed demographically and geographically

**All individuals are equally likely to be caught in each sample**

Marks are not lost (or unobserved)

Several sources of variation in capture:

- time -  $t$  (weather, sampling effort)
- behavior -  $b$  (trap-effect)
- individual heterogeneity -  $h$  (age, sex, social status)

... and combination of these effects

To test these effects, we need more than 2 sessions

# More than two sessions - capture histories

Each line = one individual / each column = one field session

History	Comment
1011	seen first session, not seen second session, seen third and fourth session
0110	not seen first session, seen second and third session, not seen fourth session
1010	
...	
...	
1101	

# More than two sessions - assumptions

Population is closed demographically and geographically

Marks are not lost and individuals are identified without error

All individuals are equally likely to be caught in each sample

# More than two sessions - capture histories

The  $M_0$  model - constant detection probability (among individuals, among sessions)

Generalisation of Lincoln-Petersen index to more than two sessions

# Maximum likelihood

- The number of individuals with history 111 is provided by  $n_{111} = N \times p \times p \times p$
- We can do that for all histories, remember the equations we wrote for the Lincoln-Petersen index
- Solving the equation system becomes more and more challenging as the number of field sessions increases
- Then using an algorithm called "maximum likelihood" we can search for the values of  $p$  and  $N$  that lead to the closest match between the number of observed histories ( $n_{111}, n_{110}$ , etc) and their expected number given  $\hat{p}$  and  $\hat{N}$

# More than two sessions

- The  $M_0$  model is very constrained, possible to relax some assumptions
- $M_t$  = Variation over time of detection probabilities
- $M_b$  = Capture and recapture can be different (trap-dependence)

# Model Mt

- Capture probabilities vary with session (weather conditions, field effort, observer experience, etc)
- With  $K$  sessions,  $K + 1$  parameters to estimate:  $N, p_1, p_2, \dots, p_K$

# Model Mt

History	M0	Mt	Mb
111		$p_1 p_2 p_3$	
110		$p_1 p_2 (1-p_3)$	
101		$p_1 (1-p_2) p_3$	
100		$p_1 (1-p_2) (1-p_3)$	
011		$(1-p_1) p_2 p_3$	
010		$(1-p_1) p_2 (1-p_3)$	
001		$(1-p_1) (1-p_2) p_3$	
000		$(1-p_1) (1-p_2) (1-p_3)$	

# Model Mb

- Probability of first capture  $p$  differs from probability of recapture  $c$
- Individuals can become trap-shy with  $c < p$  or trap-happy with  $c > p$
- 3 parameters to be estimated:  $N, p, c$

# Model Mb

History	M0	Mt	Mb
111			pcc
110			$pc(1-c)$
101			$p(1-c)c$
100			$p(1-c)(1-c)$
011			$(1-p)pc$
010			$(1-p)p(1-c)$
001			$(1-p)(1-p)p$
000			$(1-p)(1-p)(1-p)$

# All models

History	M0	Mt	Mb
111	ppp	p1p2p3	pcc
110	pp(1-p)	p1p2(1-p3)	pc(1-c)
101	p(1-p)p	p1(1-p2)p3	p(1-c)c
100	p(1-p)(1-p)	p1(1-p2)(1-p3)	p(1-c)(1-c)
011	(1-p)pp	(1-p1)p2p3	(1-p)pc
010	(1-p)p(1-p)	(1-p1)p2(1-p3)	(1-p)p(1-c)
001	(1-p)(1-p)p	(1-p1)(1-p2)p3	(1-p)(1-p)p
000	(1-p)(1-p)(1-p)	(1-p1)(1-p2)(1-p3)	(1-p)(1-p)(1-p)

# Model with heterogeneity

- Capture can depend on several unknown variables (age, sexe, social status, etc)
- Each individual has a unique capture probability
- $N + 1$  parameters to estimate:  $N, p_i$  with  $i = 1, 2, \dots, N$
- This model has too many parameters, we can use Pledger's mixture models

# Pledger's mixture model

- Assume individuals belong to two groups with different capture probabilities
- A proportion  $\pi$  belongs to the first group with capture probability  $p_H$
- And  $(1 - \pi)$  to the second group with  $p_L$
- Then  $p = \pi \times p_H + (1 - \pi) \times p_L$
- We can just use this equation in  $M_0$  and get four parameters:  $N$ ,  $\pi$ ,  $p_H$  and  $p_L$

# More complex models

- Combining different effects:  $M_{bt}$ ,  $M_{bh}$ ,  $M_{th}$
- Use some individuals covariates: Age, sex, body conditions, etc

# Model selection

- We can fit several models, each returning a different  $\hat{N}$
- How to select for the one that is the best given the data ?
- The most complex models are always the one that explain best the data (increase explained variance)
- Yet sometimes the complexification yields to only very small improvement of explained variance
- Idea: **penalize models with too many parameters**

# Akaike information criterion (AIC)

$$AIC = \text{Deviance} + 2K$$

- with *Deviance*, a measure of the unexplained variance of the model
- $K$  the number of parameters in the model
- AIC makes the balance between *quality of fit* and *complexity* of a model
- Best model is the one with lowest AIC value

Note that two models are difficult to distinguish if  $\Delta\text{AIC} < 2$

# Some thoughts on closure assumption

- Some tests of closure exist (Otis et al 1978, Stanley and Burnham 1999), available in R (package "secr")
- Yet very sensitive to other sources of variations such as capture heterogeneity
- Low performance when  $p$  is low
- Some authors now rely more on ecological justification than statistical tests

# Cited references

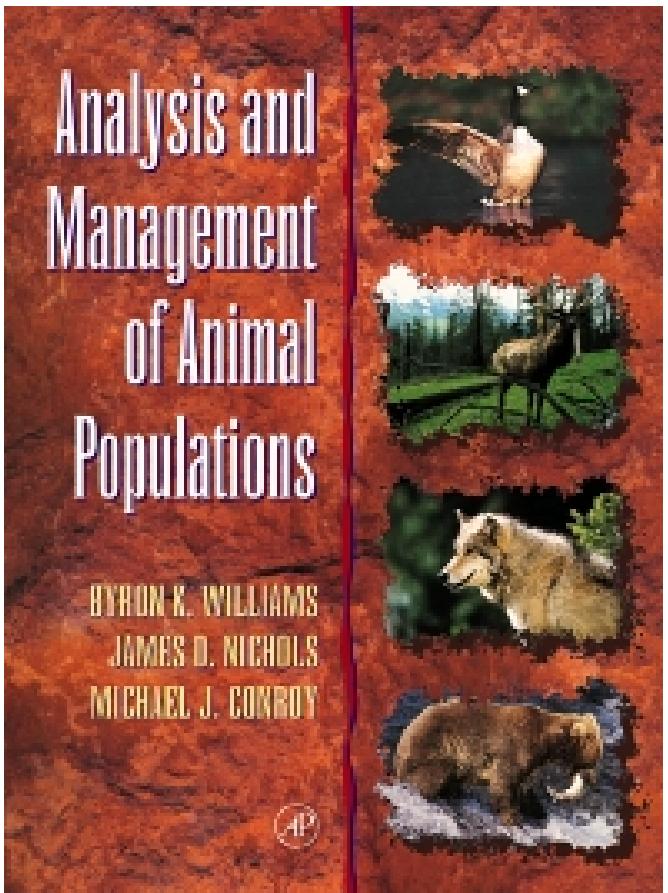
Huggins R. M. (1989). On the statistical analysis of capture-recapture experiments. *Biometrika* 76:133-140.

Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs* 62: 1–135.

Pledger S. (2000). Unified Maximum Likelihood Estimates for Closed Capture–Recapture Models Using Mixtures. *Biometrics* 56: 434-442.

Stanley, T. R. and Burnham, K. P. (1999). A closure test for time-specific capture–recapture data. *Environmental and Ecological Statistics* 6:197–209.

# Further reading



<http://www.phidot.org/software/mark/docs/book/>

A screenshot of the 'program MARK' website. The header reads 'program MARK' and 'A Gentle Introduction'. Below the header is a navigation bar with links for Home, About, Download, Support Forum, and Training. The main content area has a section titled 'Overview' which includes a thumbnail image of a book cover for 'Program MARK - a gentle introduction' and some descriptive text about the program. To the right of the overview is a sidebar with sections for 'Book chapters &amp; data files' (with a dropdown menu) and 'Thoughts for the day...' (with a quote from the book).

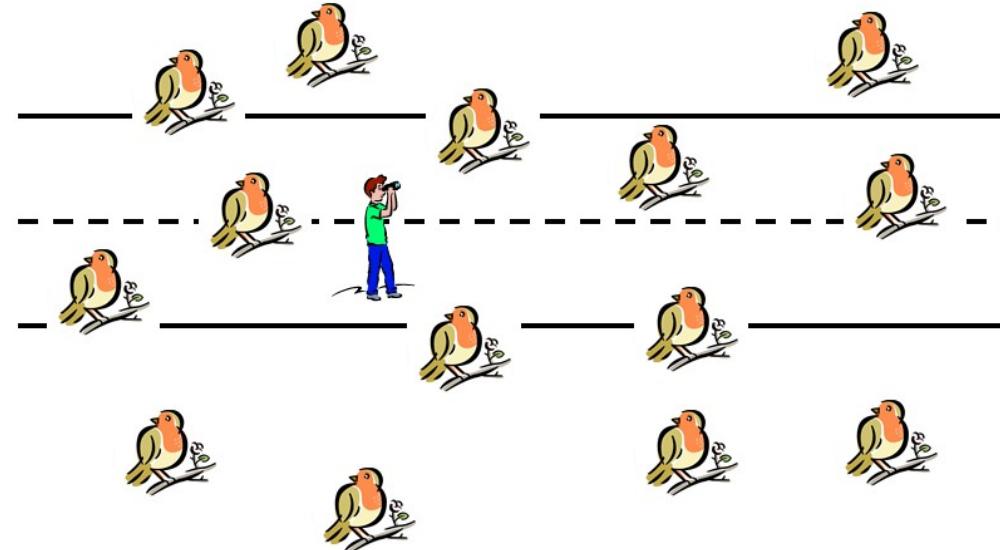
# Live demo in R



# Distance sampling

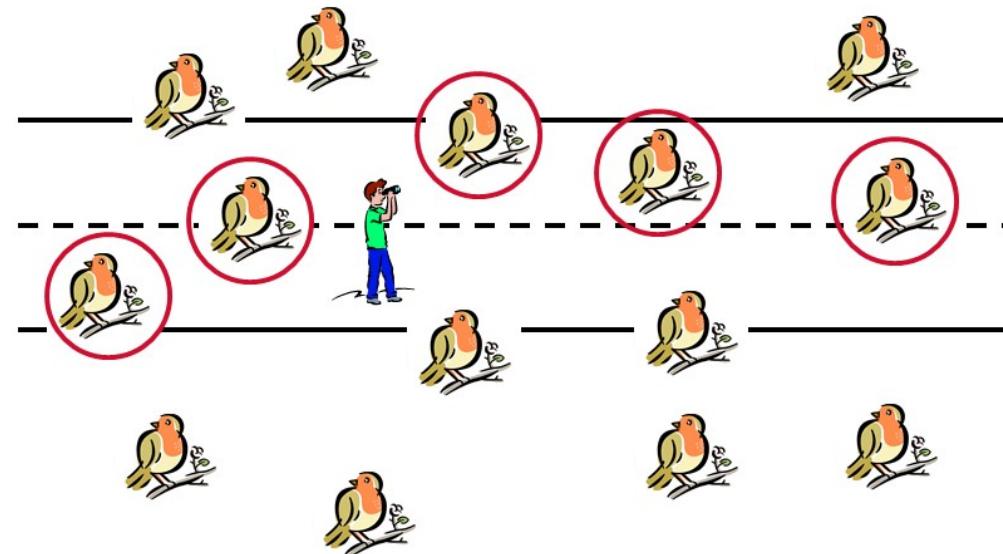
# Count individuals along strip transects

Traditional methods = count animals (or plants) along transects of fixed width



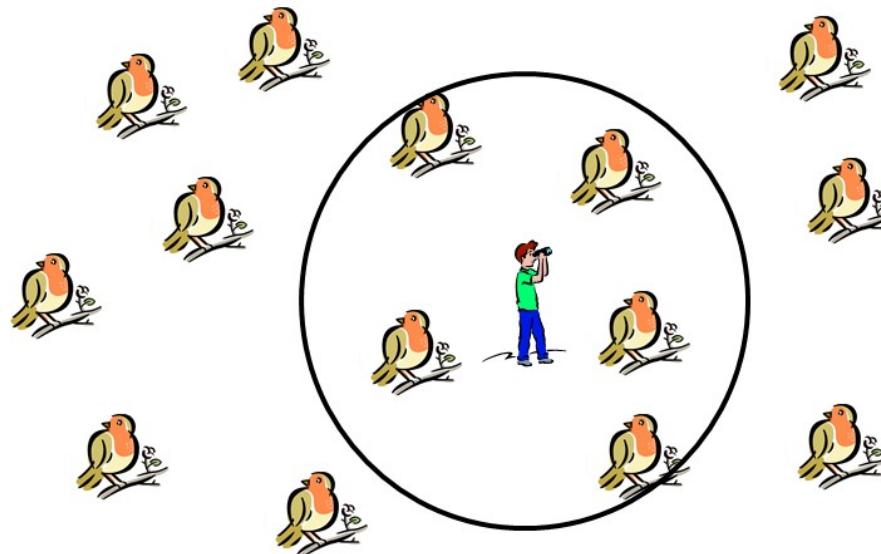
# Count individuals along strip transects

Assume all individuals in the strip are detected and counted



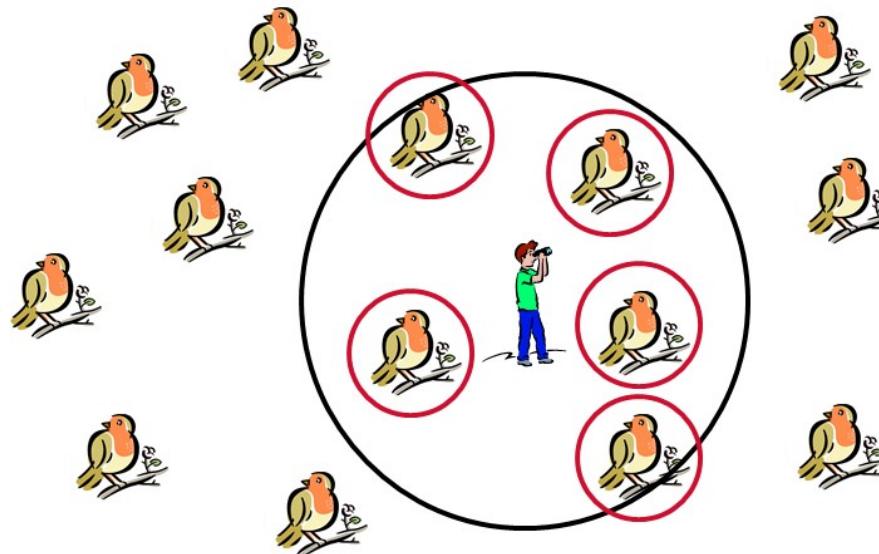
# Count individuals inside a circular area

The survey area can also be a circle = circular plot



# Count individuals inside a circular area

Assume all individuals in the circle are detected and counted



# Estimating density and abundance using counts (1)

- Using raw counts in the sample strips/circles
- Density is easily estimated using:

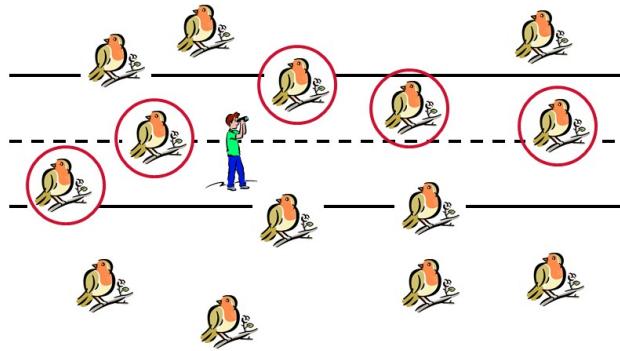
$$\hat{D} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n a_i}$$

- With  $y_i$  the count at each sample unit, and  $a_i$  the surface of each sample unit
- That can be translated into abundance  $\hat{N}$  with:

$$\hat{N} = \frac{\sum_{i=1}^n y_i}{\alpha}$$

- Where  $\alpha$  is the proportion of the study area eventually sampled

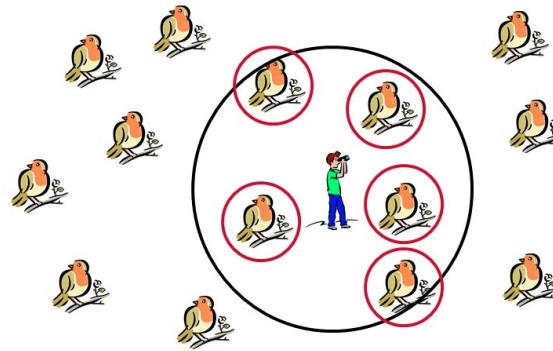
# Estimating density and abundance using counts (2)



$L$  = transect length,  $w$  = transect width,  $y$  = number of birds detected

$$\hat{D} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n L_i \times w_i}$$

# Estimating density and abundance using counts (3)



$w$  = radius,  $y$  = number of birds detected

$$\hat{D} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n \pi \times w_i^2}$$

# Assumptions

- All individuals in the sample area are detected and counted
- Yet detection probability is (always) below 1 (see lecture on detection issue)
- It can vary between sample units and even in sample units (depending on vegetation cover for instance)

# Distance sampling - background (1)

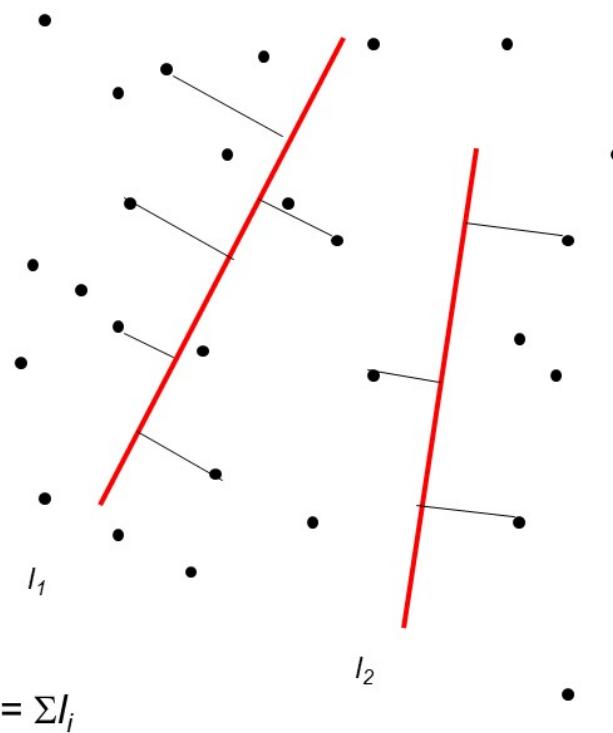
- Distance sampling rational was formulated by Steve Buckland, early 1980
- The number of detections tends to decrease with distance to observer...
- ...some information on detection probability is encapsulated in distances
- Can we estimate detection probability by recording the distances of detected individuals rather than their presence only ?

# Distance sampling - background (2)

- Observations along a transect
- Record perpendicular distances to transects, assume not all individuals are detected

# Distance sampling - background (3)

Observations along several transects



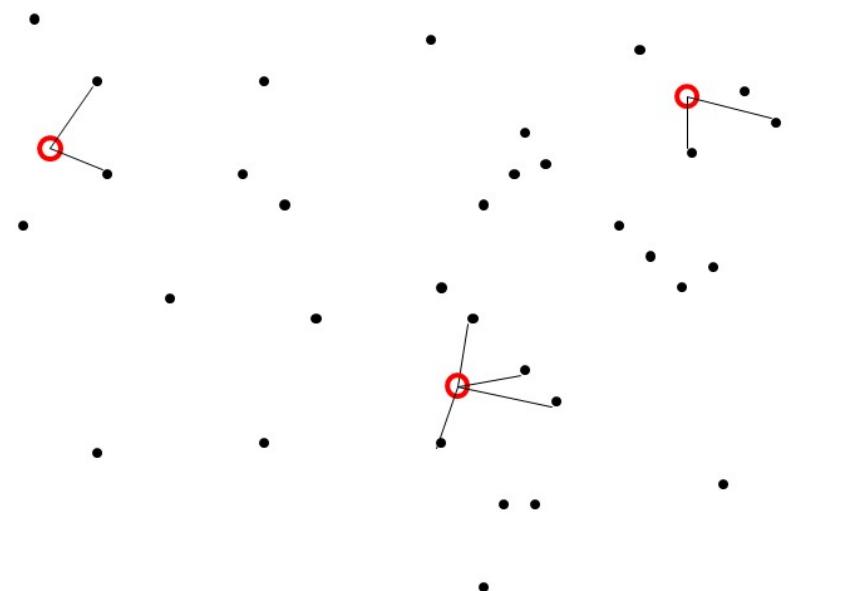
# Distance sampling - background (4)

Observations from a fixed location

Record radial distances to the point, assume not all individuals are detected

# Distance sampling - background (5)

Observations from several fixed locations



# Estimating density when detection < 1

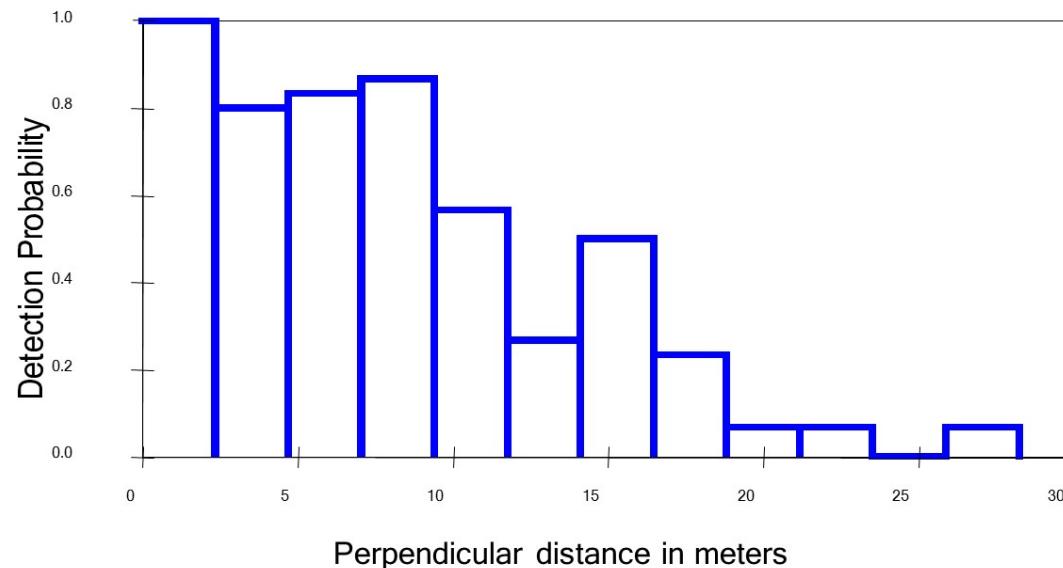
- Remember first lecture on detection issue, but also Capture-recapture:  $N = \frac{C}{p}$

$$\hat{D} = \frac{\sum_{i=1}^n y_i}{a \times \hat{p}}$$

- With  $p$  the mean detection probability on the entire sampled area of surface  $a$
- How to estimate  $p$  using distance data ?

# Recorded data

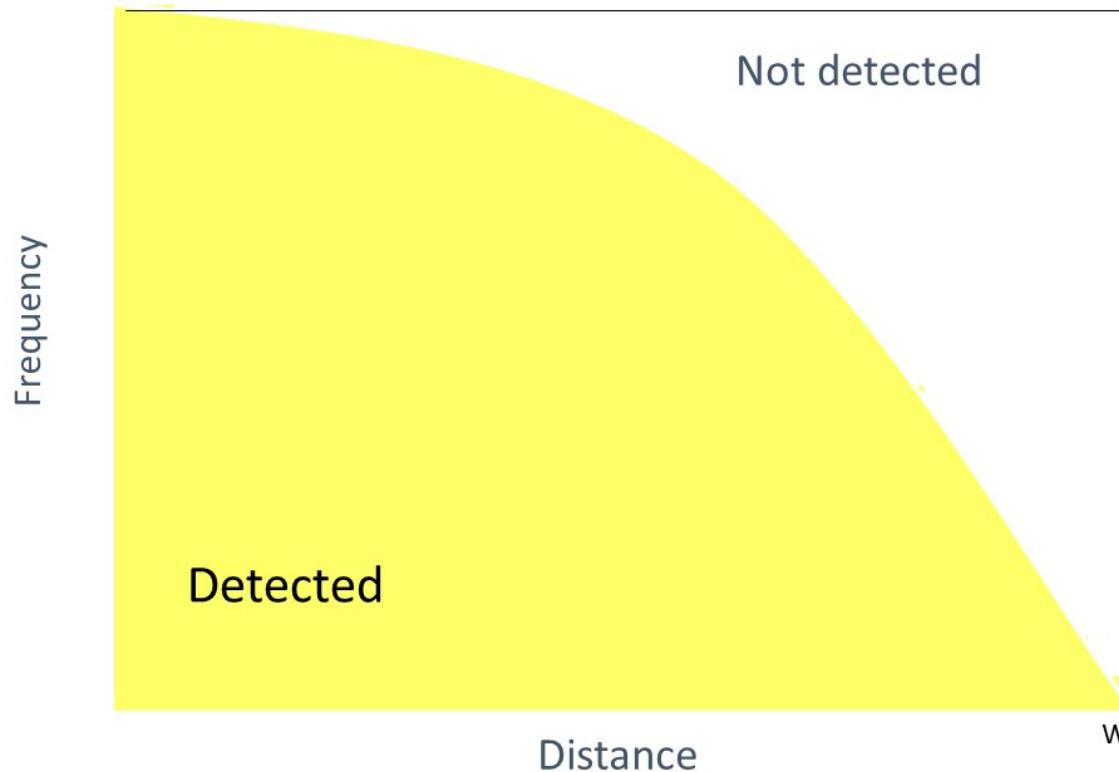
- Distance sampling recorded data typically look like:



- If  $p$  was equal to one, the histogram would be flat
- The decrease results from a decreasing detection prob with distance to observer

# Estimating detection with distance data (1)

Some individuals are detected, some are not



# Estimating detection with distance data (2)

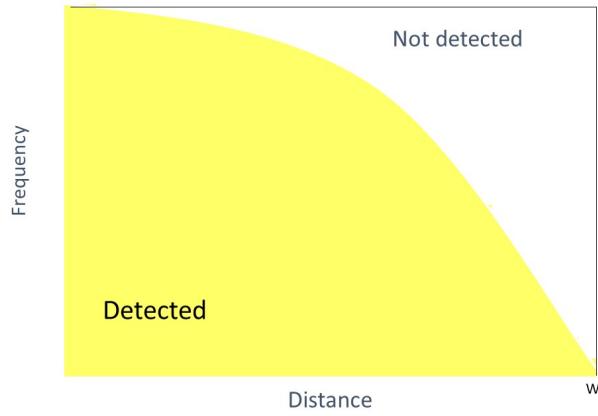
For instance, if 100 individuals are observed, 70% of the individuals are detected (below the curve) and 30% not detected

$$\hat{N} = \frac{100}{0.70} = 143$$

# How to extract $p$ from these data ?

- We need to estimate  $p$
- To do that, we need to fit a detection curve to the raw data

# How to extract $p$ from these data ?



$$\hat{p} = \frac{\int_0^w \hat{g}(y) dy}{w}$$

with  $w$  the distance at which data are truncated and  $\hat{g}(y)$  the detection curve

it is the area under the curve  
area of the rectangle

# Modelling the detection function (1)

- $g(y)$ , detection probability given the distance, needs to be estimated
- $g(y)$  is not known and may vary a lot due to observer experience and environment
- Distance sampling strategy is to go for a few models for  $g(y)$  with good properties:
  - model robustness
  - shape criterion
  - efficiency

# Modelling the detection function (2)

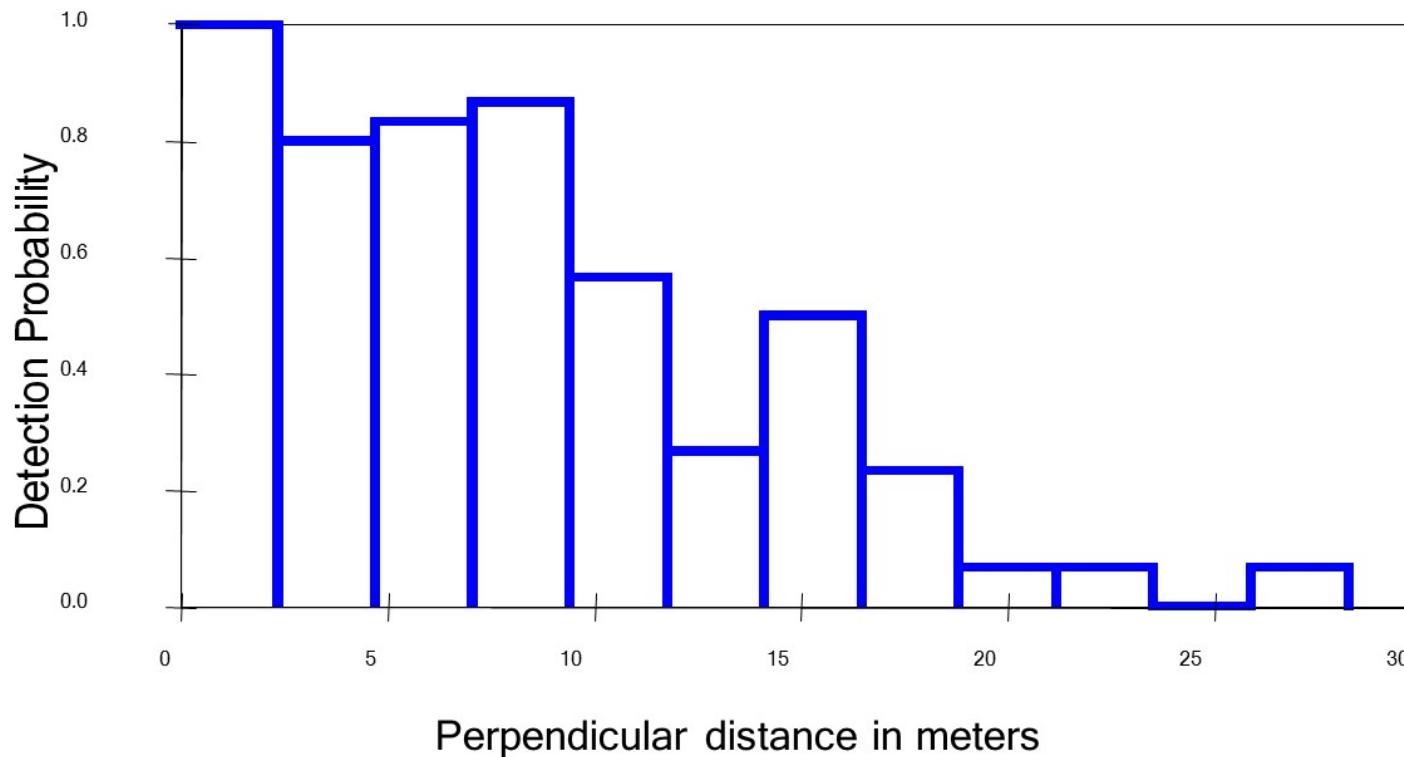
- Model robustness
  - General function that can take a variety of shapes
  - Robust w.r.t. pooling
- Shape criterion
  - A “shoulder” near the transect line (detection approx. 1 around the line / point)
- Efficiency
  - Provides precise estimates – use Maximum Likelihood

# Modelling the detection function (3)

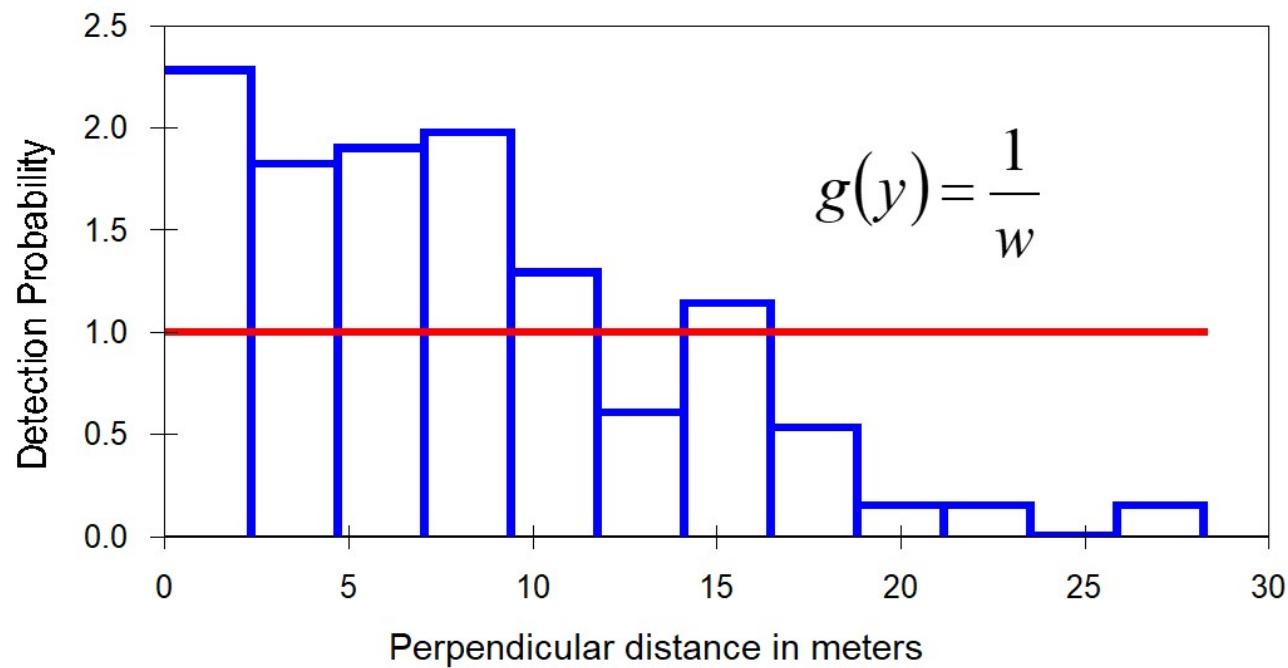
Two components for  $g(y)$ :

- A ‘key’ function: Uniform, half-normal, hazard-rate
- A ‘series’ expansion: Cosine, simple polynomial, hermite polynomial

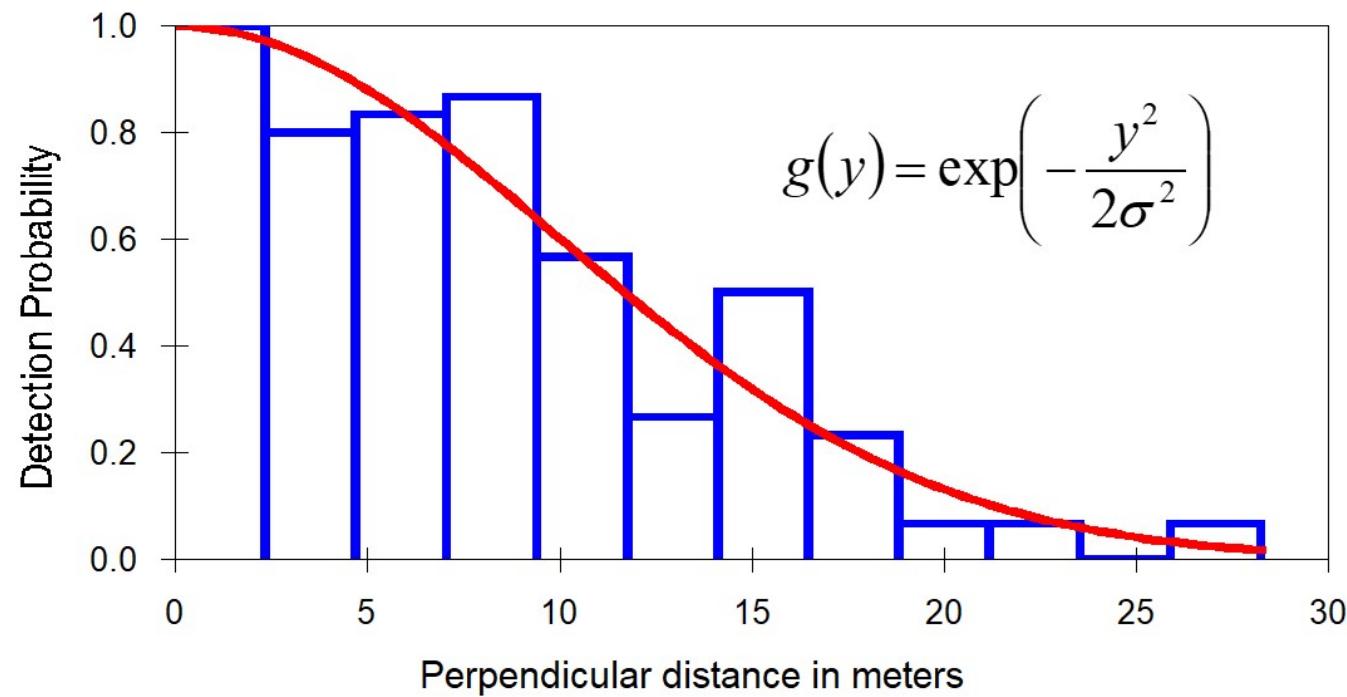
# Fit a detection curve to the raw data



# Uniform key

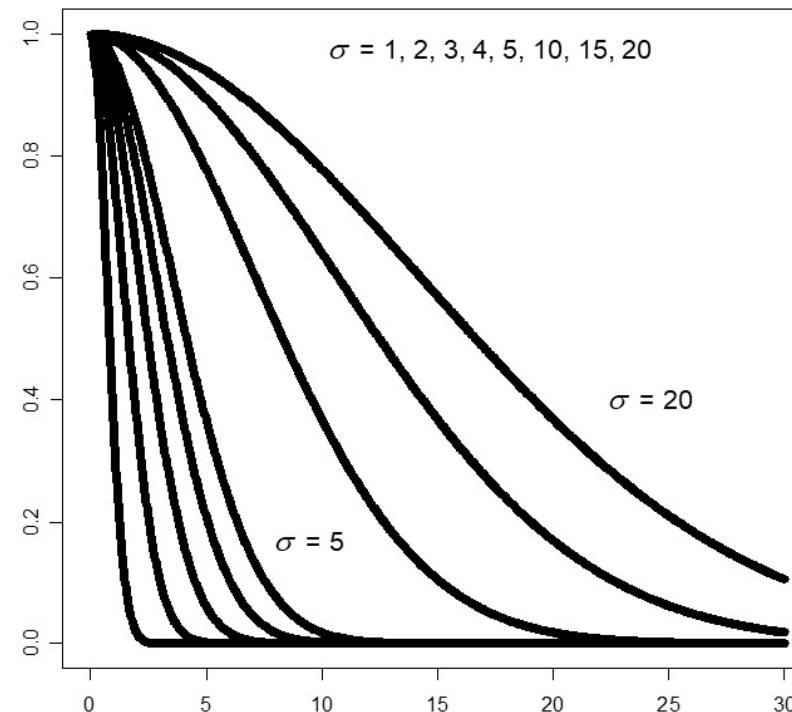


# Half-normal key

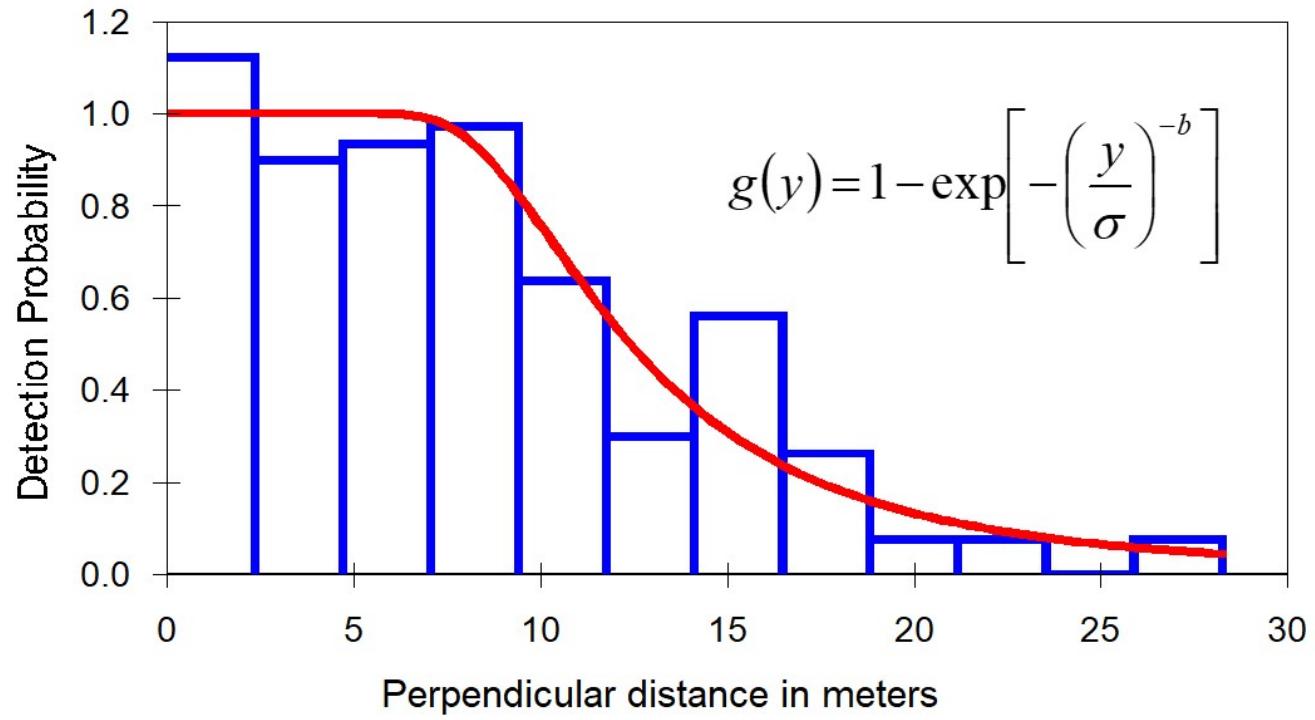


# Half-normal key

Changing  $\sigma$  allows changing tail length

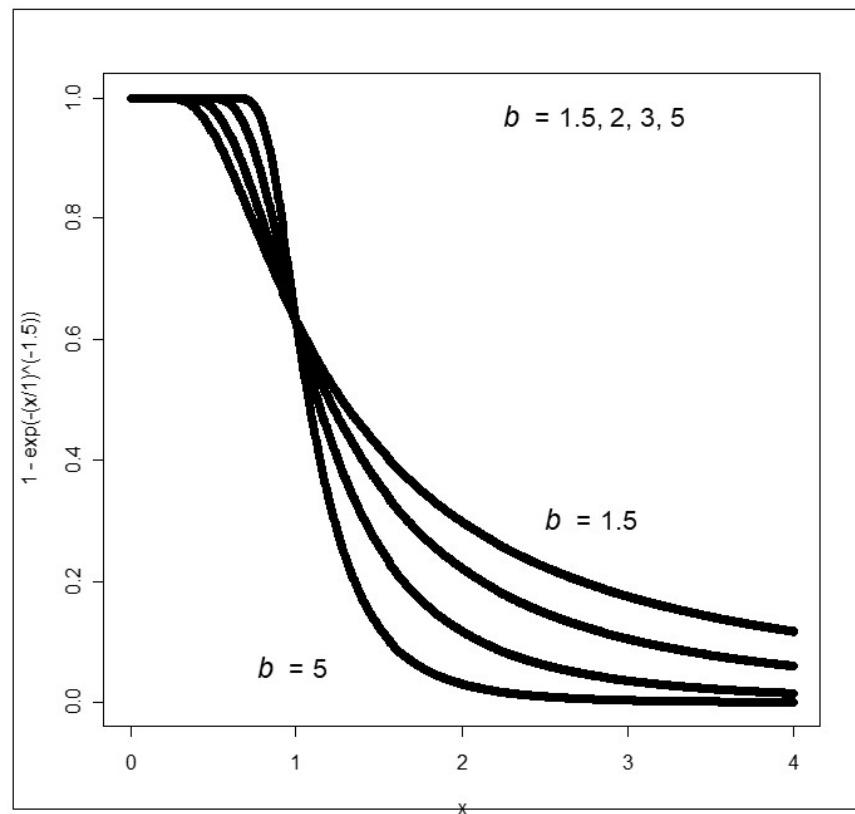


# Hazard-rate key



# Hazard-rate key

Changing  $b$  allow changing shoulder and tail

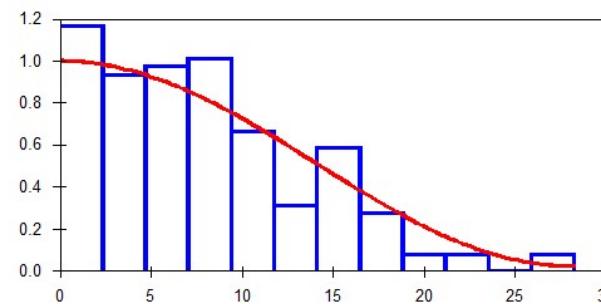
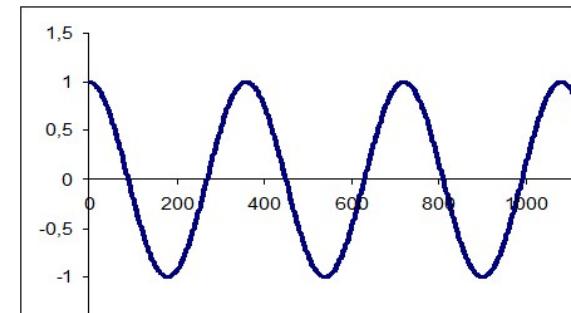
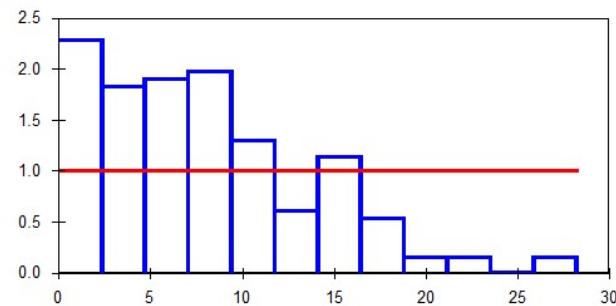


# Series expansion

- Cosine  $\sum a_j \cos\left(\frac{j\pi y}{w}\right)$
- Simple polynomial  $\sum a_j \left(\frac{y}{w}\right)^2 j$
- Hermite polynomial  $\sum a_j H_2 j\left(\frac{y}{\sigma}\right)$

# Series expansion

Uniform key function + single cosine adjustment term

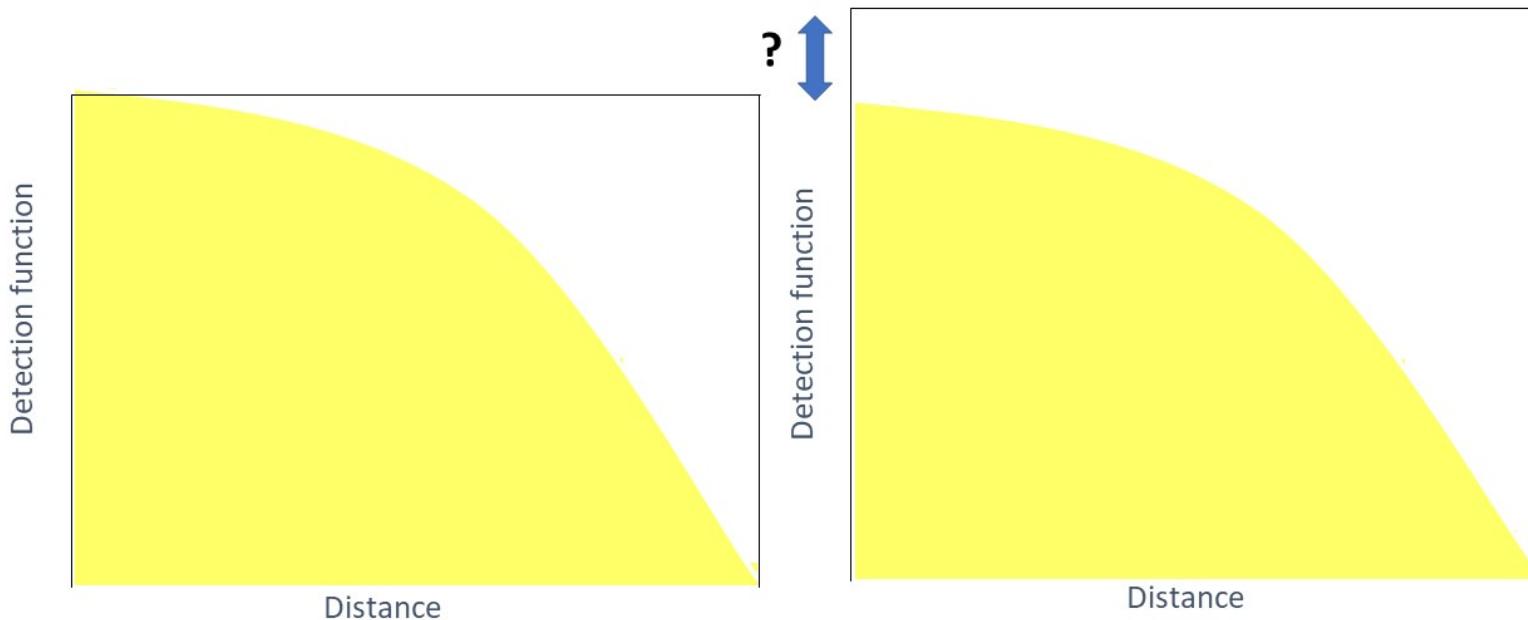


# Three critical assumptions

- Individuals on the line or point are all detected  $g(0) = 1$
- Individuals are detected at their initial location
- Distances are recorded accurately

# Individuals on the line or point are all detected

If  $g(0)$  is below 1 and unknown, density is underestimated by an unknown factor



# Individuals on the line or point are all detected

- Possible to estimate  $g(0)$  using two or more independent observers or radiotracked individuals
- Ensure  $g(0) = 1$  using cameras, trained pointing dogs, but be aware of potential bias
- Use more recent modelling approach, e.g. distance sampling with temporary emigration (requires repeated surveys in the same sampling unit)

# Individuals are detected at their initial location

- Assumption 2: Animals do not move before detection (snapshot)
- Effect of violation:
  - Random movements induce positive bias; provided object movement is slow relative to movement of the observer, bias is small
  - Responsive movements can cause large bias; positive if there is attraction to the observer, negative if there is avoidance
- Recommendations:
  - In point transects, wait once on site before proceeding
  - In line transects, look well ahead

# Measures are recorded accurately

- Assumption 3: Distances are measured accurately
- Effect of violation:
  - The estimator is fairly robust to random errors in measurement
  - It is sensitive to extreme outliers and to rounding distances
- Recommendations:
  - If exact measurements are difficult, use intervals (group data)
  - If outliers, use truncation (5-10% of the largest observations)
  - If heaping (convenient ‘rounding’), use intervals and choose cutpoints such that “heaps” are at the midpoint of an interval
  - Accurate measurement is most effective solution, so use appropriate tools (tape measures; laser range finders; compass for angles)

# Goodness-of-fit tests

- We need to verify that the fit of the model makes sense
- Several diagnostics and tests have been developed (QQplot, Chi-square tests, etc), depends on whether data are continuous or in classes of distance
- The general principle is to verify that there is no large discrepancy between the prediction from the fitted model and the raw data

# Line or point transects ?

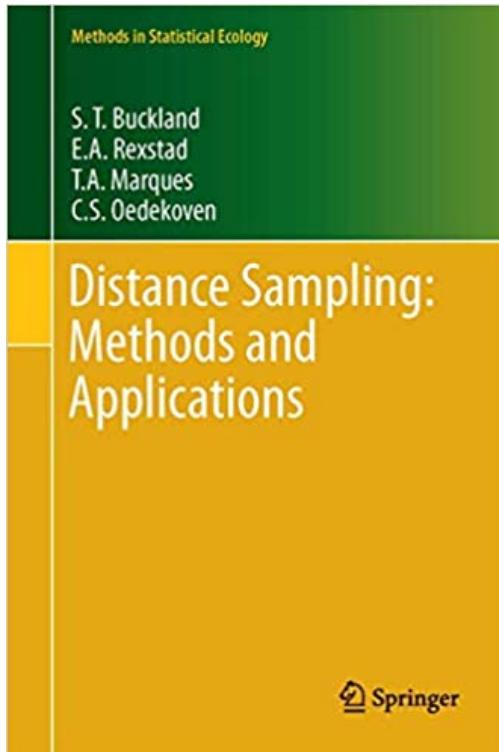
Point transects are:

- Relevant for populations distributed in patches
- Convenient to use when the area is difficult to access / survey
- Naturally suited for stratification

Line transects are:

- Efficient for sparsely distributed populations
- Effective in low densities

# Further readings



# Live demo in R

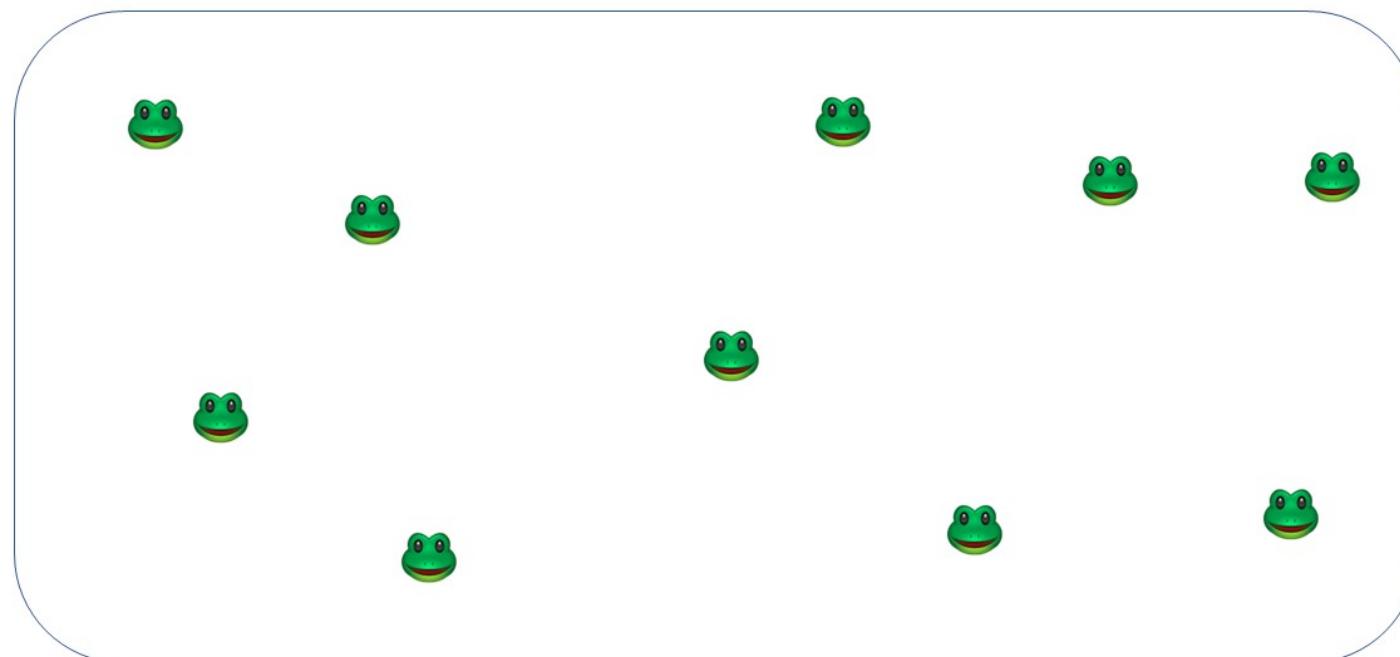


# N-mixture

# Count on a fixed sample area

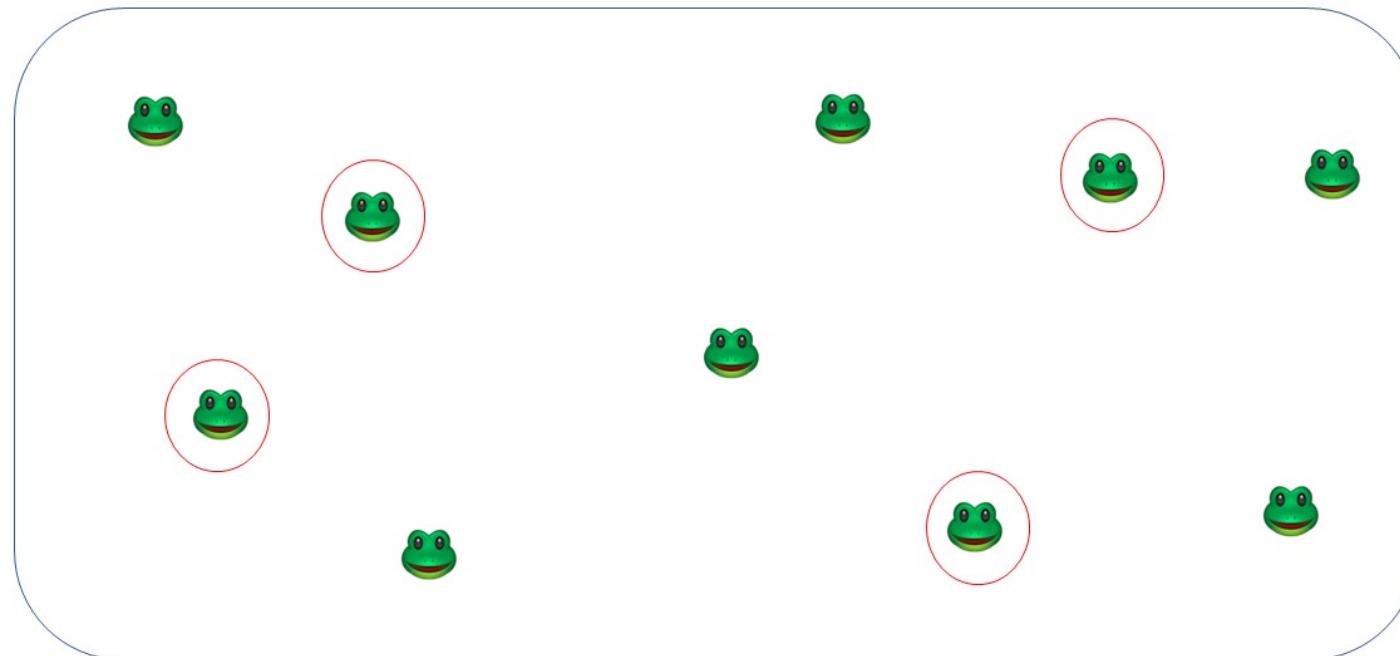
- Traditional methods = count animals (or plants) along transects of fixed width or on quadrat, or on circular plots
- These counts are affected by detection issues

# Repeated counts at the same sample unit



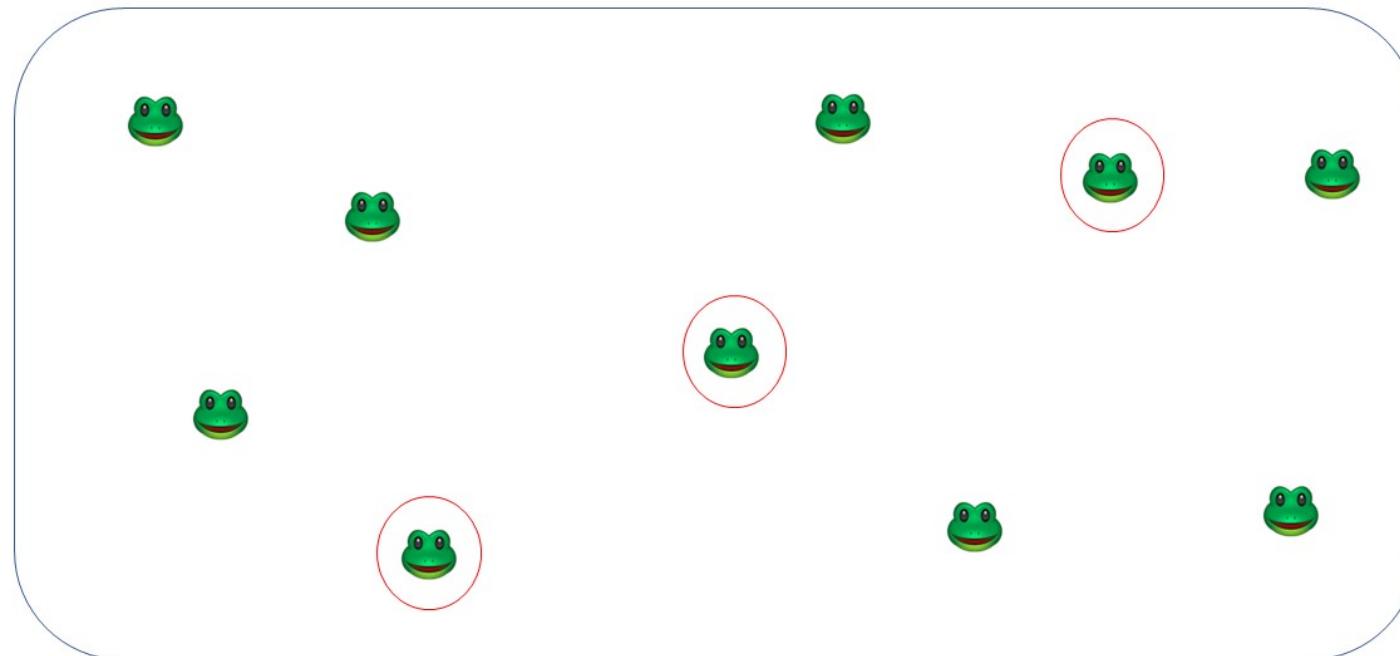
True abundance  $N = 10$

# Repeated counts at the same sample unit



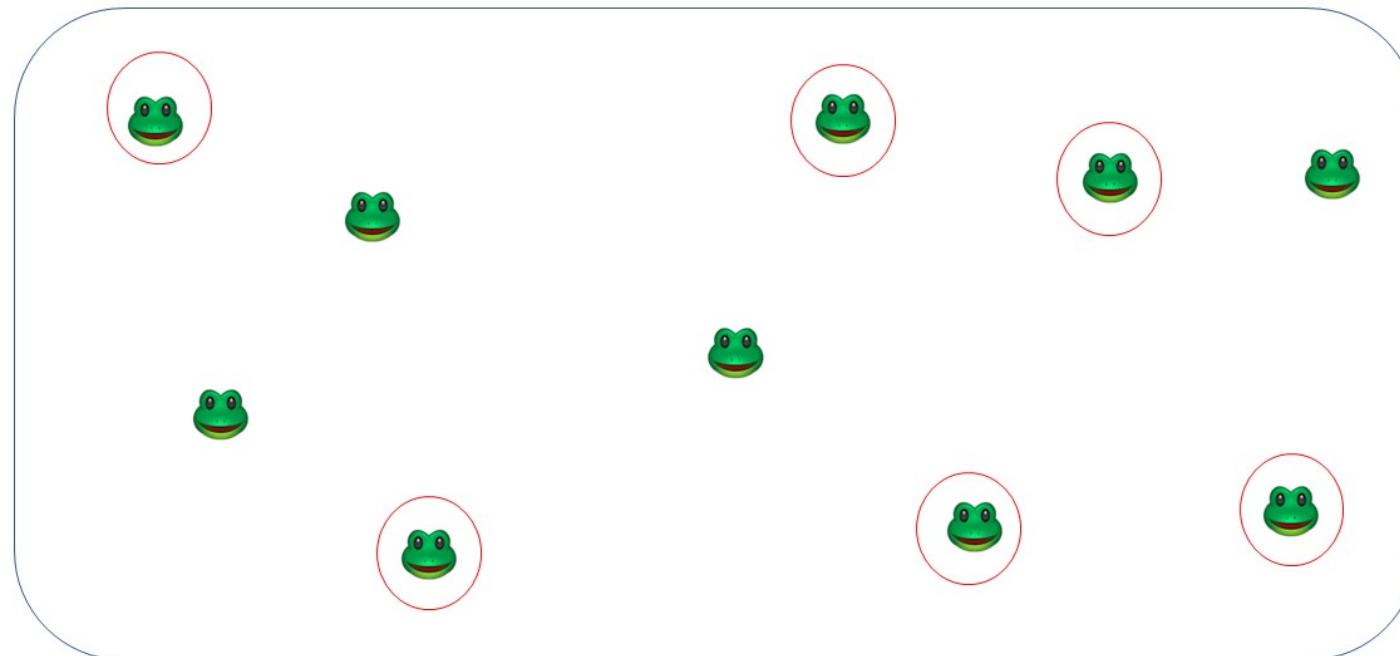
True abundance  $N = 10$ , count at first visit is 4

# Repeated counts at the same sample unit



True abundance  $N = 10$ , count at second visit is 3

# Repeated counts at the same sample unit



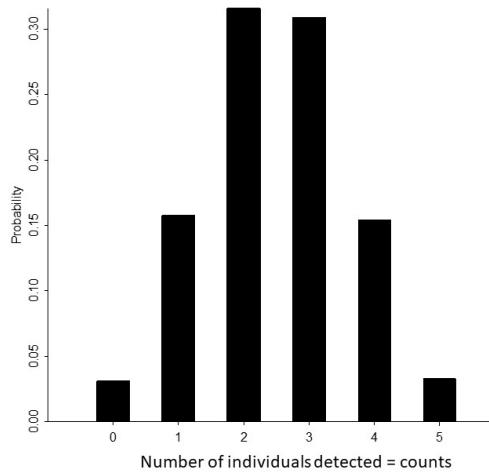
True abundance  $N = 10$ , count at third visit is 6

# Repeated counts with imperfect detection

- When counts are repeated on the same sampling unit, some variations are observed
- This is due to random process related to imperfect detection
- Example: sex-ratio in humans is close to 0.5
  - Probability to have a son is therefore 0.5
  - Expected number of sons in four siblings is 2
  - Yet some parents have four sons and zero daughters

# Repeated counts with imperfect detection

- Example: true abundance is 5, detection probability is 0.5
- Repeated counts on the same unit with imperfect detection leads to variations



- Repeated counts on the same units encapsulate information on detection and abundance

# N-mixture models (Royle 2004)

- N-mixture models take advantage of these variations between counts at the same sample unit
- Requires some spatial replications (several sampling units) and temporal replications (repeated counts at the same units) - same design as *site occupancy models*
- Some strong assumptions:
  - Abundance at a sampling unit does not change between successive counts ("closed population")
  - Individuals have the same detection probability (no heterogeneity)
  - Abundances at the sampling units are Poisson distributed
  - Sampling sessions are independent (no trap-dependence)

# N-mixture models (Royle 2004)

Typical dataset used for N-mixture models

# N-mixture models formulation

- N-mixture can be formulated as hierarchical models
- A state process describing the system

$$N_i \sim \text{Poisson}(\lambda)$$

with  $\lambda$  the mean abundance at the sites

- An observation process corresponding to detection issues

$$y_{it} \sim \text{Binomial}(N_i, p)$$

- Covariates can be added on abundance (e.g. vegetation cover, elevation) and on detection probability (e.g. vegetation cover, date, observer experience)

# N-mixture models - to use with caution

- The method has been highly controversial for years
- Couturier et al. 2013 - Identifiability problems when  $p$  is low or highly variable
- Barker et al. 2017 - Without auxillary information about  $p$ , count data cannot distinguish between N-mixture model or other possible models of  $N$
- Dennis 2015, Kéry 2017 - Some parameters not identifiable especially with negative binomial distribution instead of Poisson
- Link et al. 2018 - Estimates sensitive to violation of double counting and constant  $\lambda/p$ 
  - GOF unable to detect this
- Conclusion: check parameter  $K$  (see live demo), perform GOF, do not rely on negative binomial distribution

# N-mixture models - further reading

- Open N-mixture - Dail and Madsen 2011 / Hostetler and Chandler 2015
- Generalized Distance sampling - Accomodates for temporary emigration - Chandler (2011)
- Community N-mixture models - Yamaura et al. (2012)

# Cited references (1)

Barker, R.J., Schofield, M.R., Link, W.A. and Sauer, J.R. (2018). On the reliability of N-mixture models for count data. *Biometrics* 74: 369-377.

Chandler, R.B., Royle, J.A. and King, D.I. (2011). Inference about density and temporary emigration in unmarked populations. *Ecology* 92: 1429-1435.

Couturier, T., Cheylan, M., Bertolero, A., Astruc, G. and Besnard, A. (2013). Estimating abundance and population trends when detection is low and highly variable: A comparison of three methods for the Hermann's tortoise. *Journal of Wildlife Management* 77: 454-462.

Dail, D. and Madsen, L. (2011). Models for Estimating Abundance from Repeated Counts of an Open Metapopulation. *Biometrics* 67: 577-587.

Dennis, E.B., Morgan, B.J. and Ridout, M.S. (2015). Computational aspects of N-mixture models. *Biometrics* 71: 237-246.

# Cited references (2)

Hostetler, J.A. and Chandler, R.B. (2015). Improved state-space models for inference about spatial and temporal variation in abundance from count data. *Ecology* 96: 1713-1723.

Kéry, M. (2018). Identifiability in N-mixture models: a large-scale screening test with bird data. *Ecology* 99: 281-288.

Link, W.A., Schofield, M.R., Barker, R.J. and Sauer, J.R. (2018). On the robustness of N-mixture models. *Ecology* 99: 1547-1551.

Royle, J.A. (2004). N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. *Biometrics* 60: 108-115.

Yamaura, Y., Royle, J.A., Shimada, N. et al. (2012). Biodiversity of man-made open habitats in an underused country: a class of multispecies abundance models for count data. *Biodivers Conserv* 21, 1365-1380.

# Live demo in R



# Estimation of demographic parameters

Olivier Gimenez for the team





# On our plate

- Estimating survival with capture-recapture models
- Including discrete and continuous covariates
- Estimating transitions between sites/states

# On our plate

- **Estimating survival with capture-recapture models**
- Including discrete and continuous covariates
- Estimating transitions between sites/states

# Capture, mark and recapture



Emmanuelle Cam & Jean-Yves Monnat



Kelly Powell

## Artificial marks

# Capture, mark and recapture



Natural marks

Capture-recapture data

## 3 capture occasions, 2 time intervals

1	0	0
1	1	1
1	1	1
1	1	1
1	0	0
1	1	0
1	0	0
1	1	1
1	0	0

# Parameters





# Definition of parameters

- Parameter  $\varphi_i$  is the probability that a marked animal in the population at sampling occasion  $i$  survives until occasion  $i + 1$  and remains in the population (does not permanently emigrate).
- Parameter  $p_i$  is the probability that a marked animal in the population is detected (captured or observed) at sampling occasion  $i$ .
- While detection operates at a sampling occasion, survival operates over a time interval.
- How to estimate these parameters from the data we have?

# Encounter histories















# Probabilities

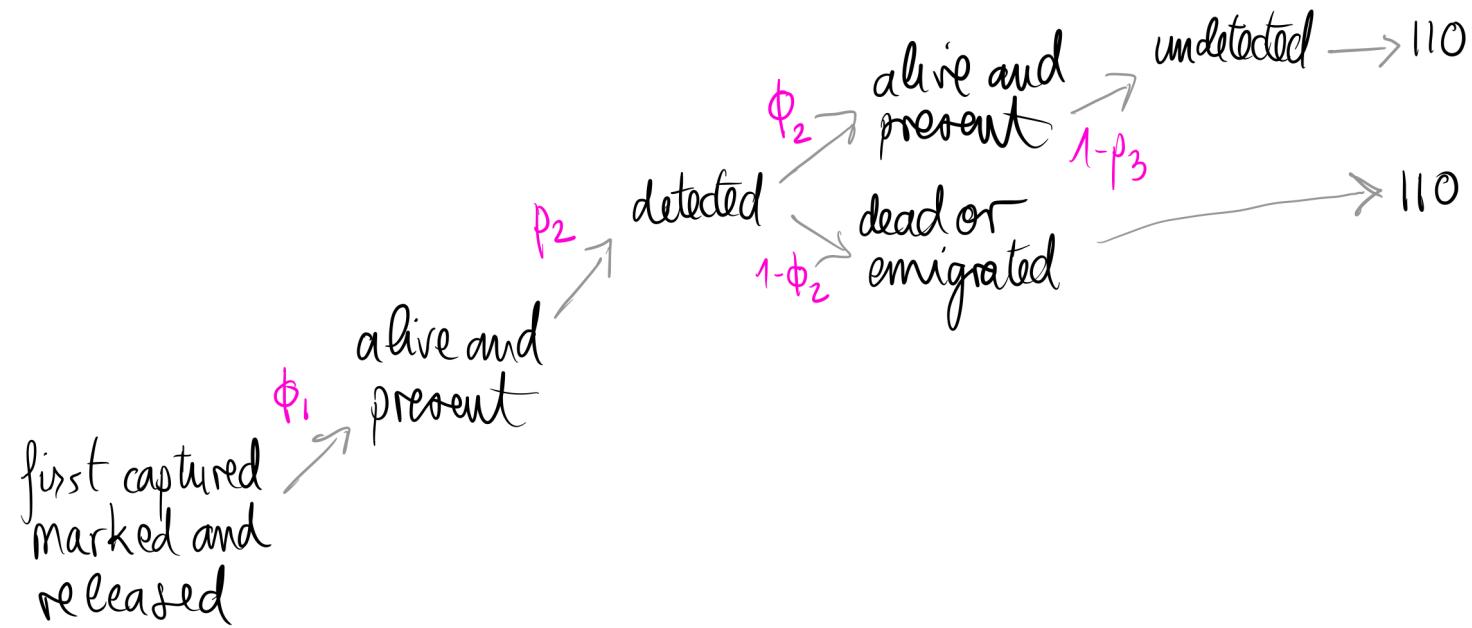




$t_1$

$t_2$

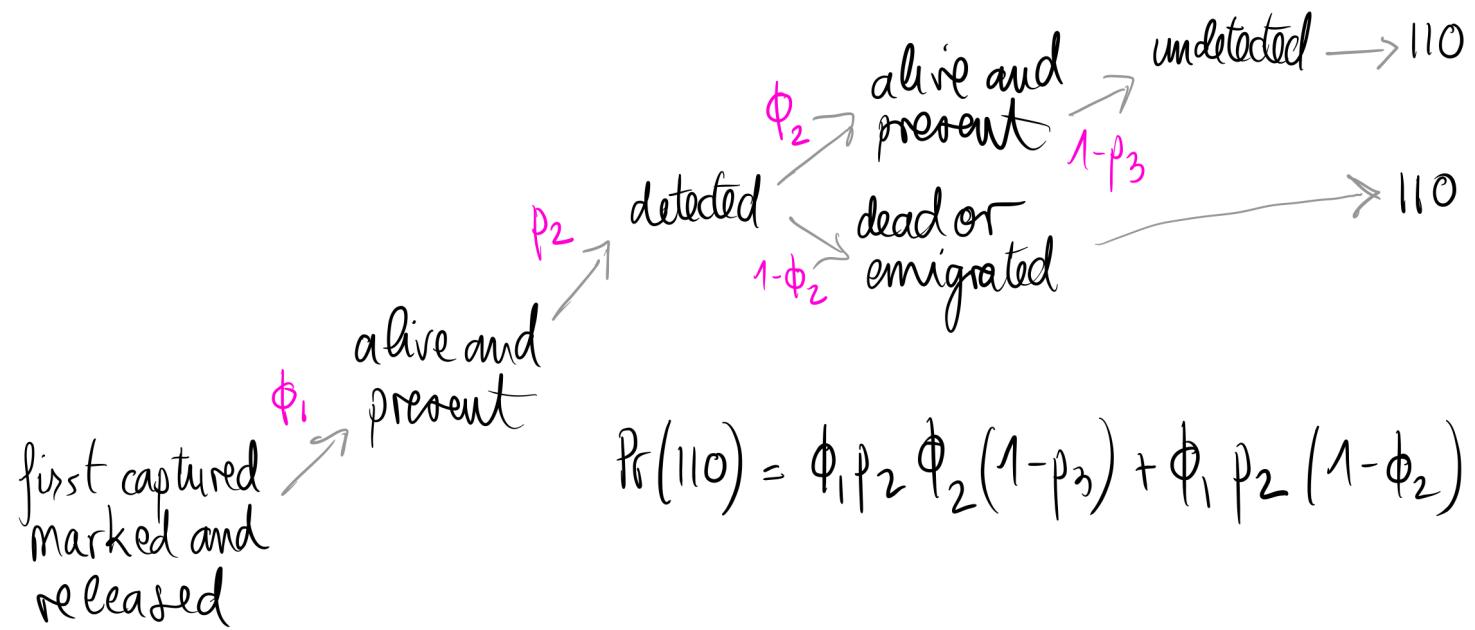
$t_3$

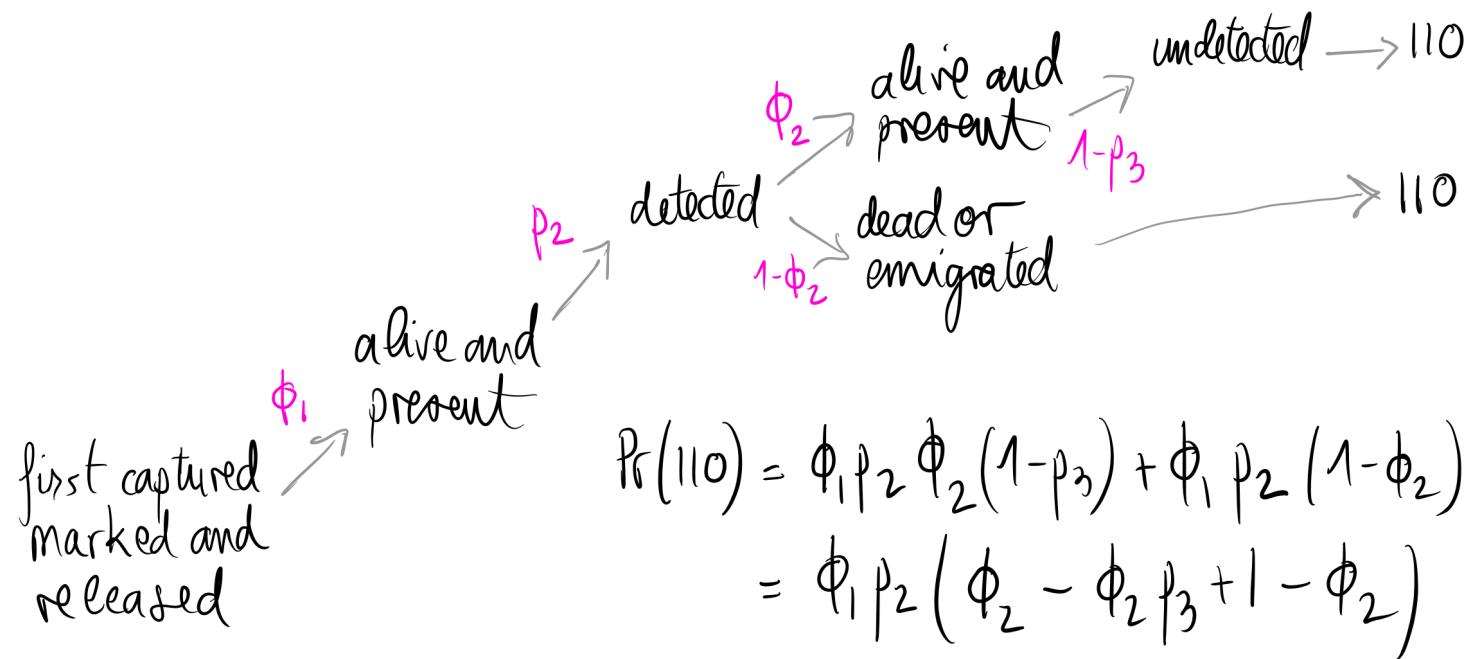


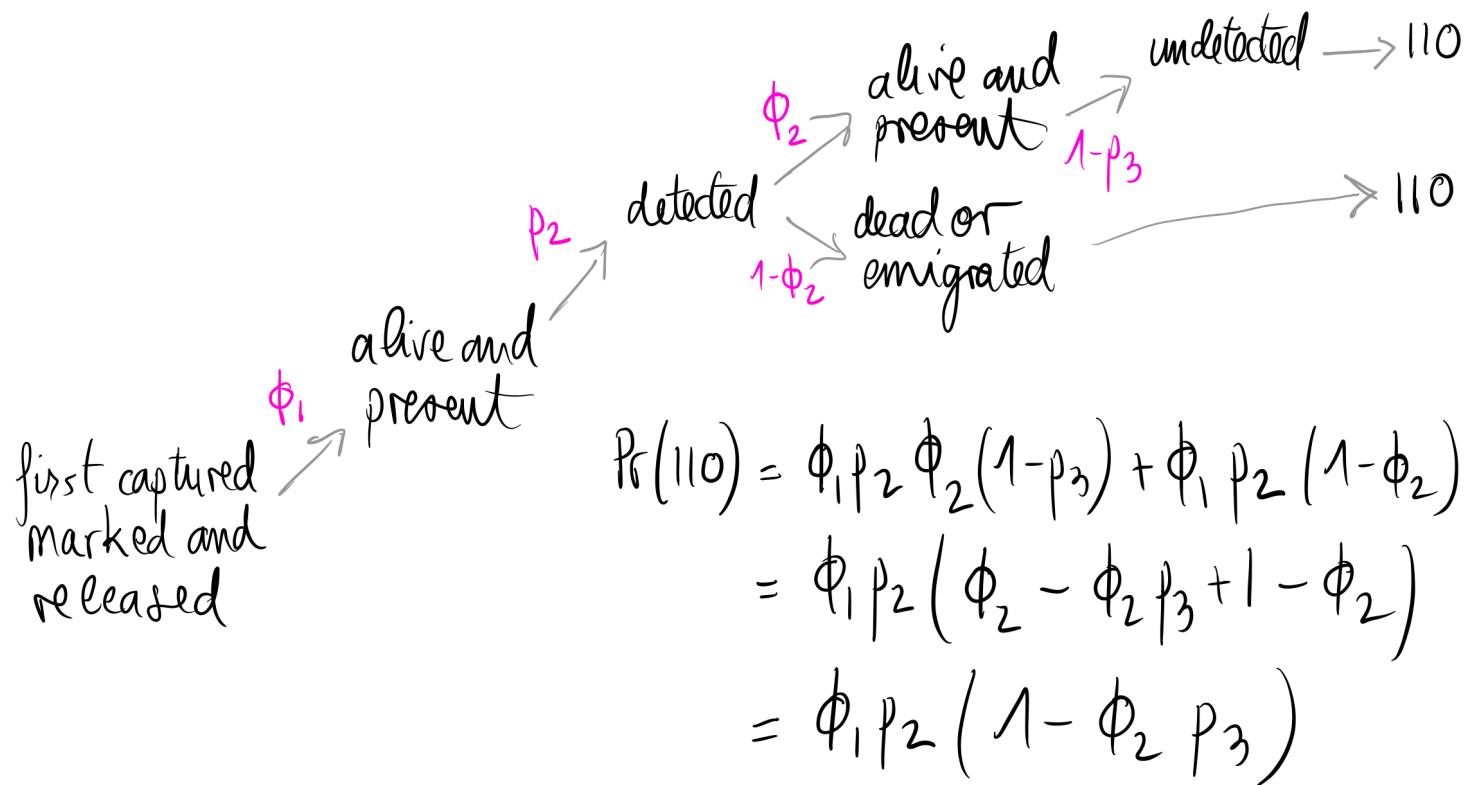
$t_1$

$t_2$

$t_3$



$t_1$  $t_2$  $t_3$ 

$t_1$  $t_2$  $t_3$ 

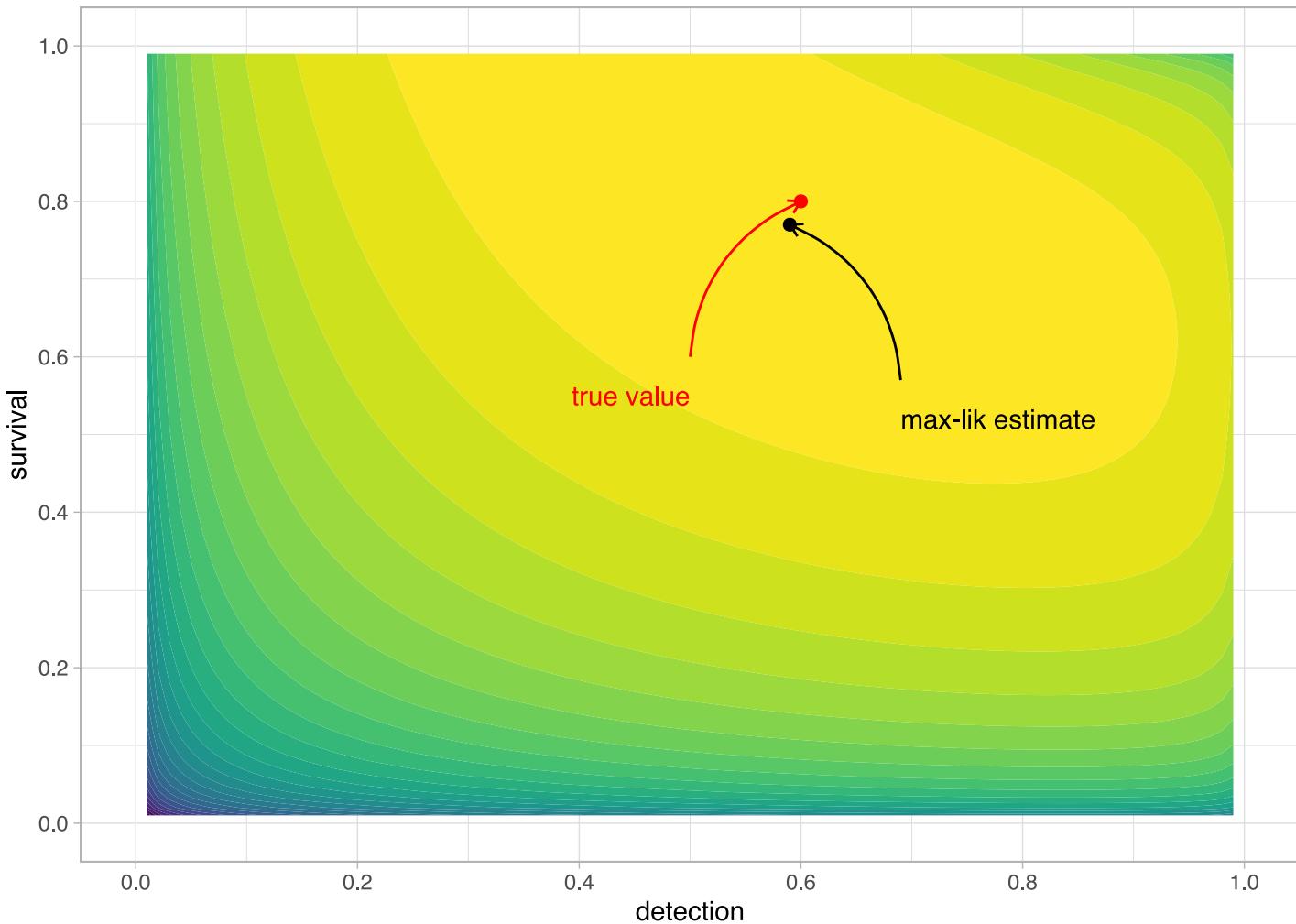
## Probabilities of all possible encounter histories with 3 occasions

- $\text{Pr}(111) = \varphi_1 p_2 \varphi_2 p_3$
- $\text{Pr}(110) = \varphi_1 p_2 (1 - \varphi_2 p_3)$
- $\text{Pr}(101) = \varphi_1 (1 - p_2) \varphi_2 p_3$
- $\text{Pr}(100) = 1 - \varphi_1 p_2 - \varphi_1 (1 - p_2) \varphi_2 p_3$

# Model likelihood

# Likelihood

- When survival and detection probabilities vary over time, we call it the Cormack-Jolly-Seber model.
- Its likelihood is  $L(\text{data} \mid \text{parameters}) = \Pr(\text{data} \mid \varphi_1, \varphi_2, p_2, p_3)$ .
- You get maximum likelihood estimates by maximising the likelihood function (!)  $L(\text{parameters} \mid \text{data})$ .
- It is  $L = \text{constant} \times \Pr(111)^{n_{111}} \times \Pr(110)^{n_{110}} \times \Pr(101)^{n_{101}} \times \Pr(100)^{n_{100}}$ .
- We search for  $\varphi_1, \varphi_2, p_2, p_3$  that maximize the likelihood function.
- Let's inspect visually the likelihood function in the simple case where all parameters are constant over time, that is  $\varphi_1 = \varphi_2 = \varphi$  and  $p_2 = p_3 = p$ .



A bit of history

# History of the Cormack-Jolly-Seber (CJS) model

S.T. Buckland (2016). A Conversation with Richard M. Cormack. Statistical Science 31: 142-150.

**Buckland:** George Jolly was a colleague of yours in the 1960s. Could you describe your interactions with him?

**Cormack:** George was in the ARC Unit of Statistics in the same corridor as I was. His main job was designing and conducting agricultural surveys in Scotland. There wasn't a practice of giving seminars in the department to talk to colleagues about what one was doing, and David Finney's appointees had been chosen to cover all the varied areas of statistics rather than build a research group in a particular area. So despite the fact that I met George every day at coffee, and, indeed, we caused David a lot of angst as, on many mornings, we played kriegspiel (a version of chess where you don't see the other person's board and a referee judges—very good for developing inference), we never mentioned work and mark-recapture. I don't remember George noticing my Biometrika paper in 1964 (Cormack, 1964), or indeed the practical paper in British Birds in 1963 (Dunnet, Anderson and Cormack, 1963). It was completely unknown to the two of us that we were working in the same area.

**Buckland:** What interactions did you have with George Seber?

**Cormack:** Before the 1965 papers (Jolly, 1965; Seber, 1965), George Seber and I had no contact whatsoever. After the papers, yes, we did. We got into deep communication after the first papers, and he was all for sending me drafts of everything he did. He produced stuff at a colossal rate and his encyclopaedic knowledge was unbelievable. I'm not sure he ever actually worked closely with biologists, but, when he was writing his book, he asked if I would comment on the draft chapters on the bits I knew about. But you have to realise that communication between opposite corners of the world took time. At one point, I received a plaintive handwritten letter saying "The University has cut down on postage and I'm not allowed to post the draft chapter airmail and you will have to wait for it to come by surface mail from New Zealand." By the time it arrived, I already had another airmail letter from him saying, "I'm sorry you haven't been able to comment on the chapter—I've had to submit it!" To some extent, the opposite is true now: response is too quick.

## MODELING SURVIVAL AND TESTING BIOLOGICAL HYPOTHESES USING MARKED ANIMALS: A UNIFIED APPROACH WITH CASE STUDIES<sup>1</sup>

JEAN-DOMINIQUE LEBRETON

*CEFE/CNRS, BP 5051, 34033 Montpellier Cedex, France*

KENNETH P. BURNHAM

*Colorado Cooperative Fish and Wildlife Research Unit, U.S. Fish and Wildlife Service,  
201 Wagar Building, Fort Collins, Colorado 80523 USA*

JEAN CLOBERT

*Laboratoire d'Ecologie, Ecole Normale Supérieure, 46 rue d'Ulm 75231, Paris Cedex 05 France*

DAVID R. ANDERSON

*Colorado Cooperative Fish and Wildlife Research Unit, U.S. Fish and Wildlife Service,  
201 Wagar Building, Fort Collins, Colorado 80523 USA*

Let's practice!

# The famous Dipper example



White-throated Dipper (*Cinclus cinclus*)



Gilbert Marzolin

294 dippers captured and recaptured between 1981 and 1987 with known sex and wing length

year_1981	year_1982	year_1983	year_1984	year_1985	year_1986	year_1987	sex
1	1	1	1	1	1	1	0 M
1	1	1	1	1	0	0	0 F
1	1	1	1	0	0	0	0 M
1	1	1	1	0	0	0	0 F
1	1	0	1	1	1	0	0 F
1	1	0	0	0	0	0	0 M
1	1	0	0	0	0	0	0 M
1	1	0	0	0	0	0	0 M

# Live demo #1



# On our plate

- Estimating survival with capture-recapture models
- **Including discrete and continuous covariates**
- Estimating transitions between sites/states

# Can we explain individual heterogeneity?

- Discrete covariate like, e.g., sex
- Continuous covariate like, e.g., mass or size

# Sex and wing length in Dipper

year_1981	year_1982	year_1983	year_1984	year_1985	year_1986	year_1987	sex
1	1	1	1	1	1	1	0 M
1	1	1	1	1	0	0	0 F
1	1	1	1	0	0	0	0 M
1	1	1	1	0	0	0	0 F
1	1	0	1	1	1	0	0 F
1	1	0	0	0	0	0	0 M
1	1	0	0	0	0	0	0 M
1	1	0	0	0	0	0	0 M

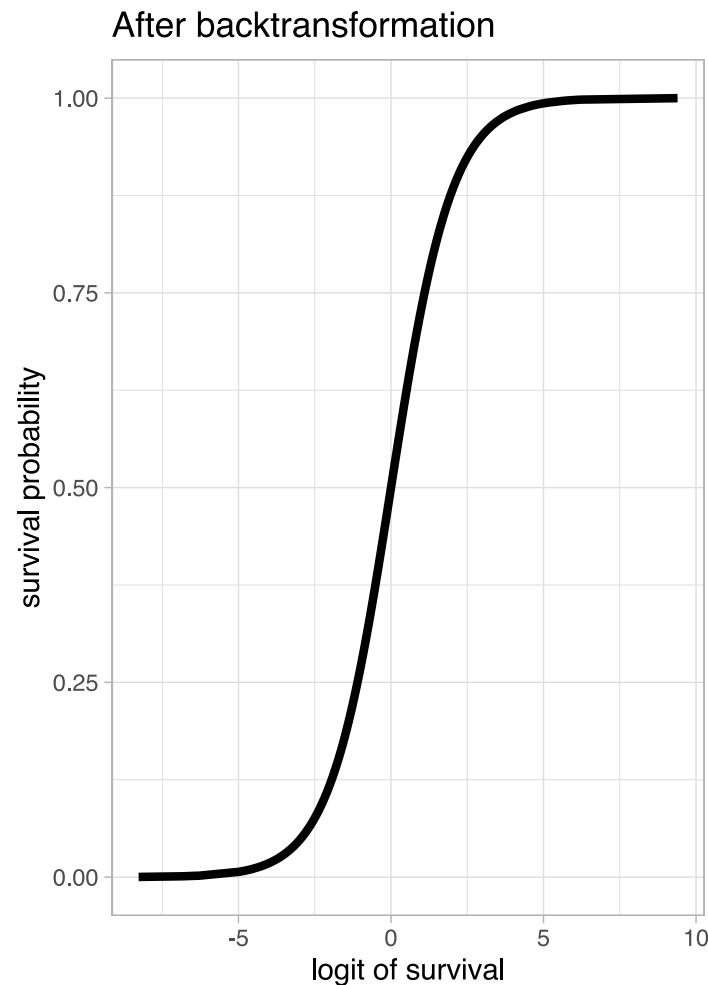
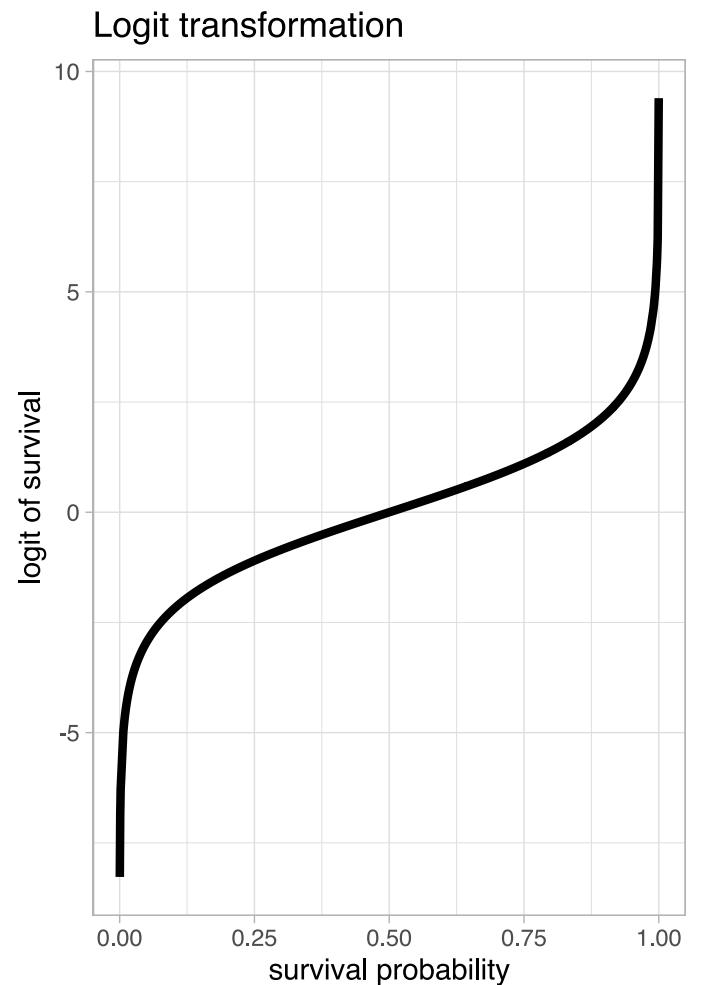
# Sex effect

- Let's use a covariate sex that takes value 0 if female, and 1 if male
- And write  $\text{logit}(\varphi_i) = \beta_1 + \beta_2 \text{ sex}_i$  for bird i
- What is this logit thing?

# Logit link function

- When modeling covariates, you could write  $\varphi_i = \beta_1 + \beta_2 x_i$
- However, the risk is that values of the  $\beta$ 's make survival  $\varphi_i$  negative or greater than 1, which is not possible for a probability
- The idea is to estimate parameters on a safe scale, then goes back to the (0,1) probability scale
- We use  $\text{logit}(y) = \log\left(\frac{y}{1-y}\right)$
- And the back-transformation is  $\text{logit}^{-1}(x) = \frac{\exp(x)}{1 + \exp(x)} = \frac{1}{1 + \exp(-x)}$

# Logit link function



# Sex effect

- Let's use a covariate sex that takes value 0 if female, and 1 if male
- And write  $\text{logit}(\varphi_i) = \beta_1 + \beta_2 \text{ sex}_i$  for bird i
- Then female survival is

$$\text{logit}(\varphi_i) = \beta_1$$

- And male survival is

$$\text{logit}(\varphi_i) = \beta_1 + \beta_2$$

- Very much alike an ANOVA with one factor (two levels)

# Wing length effect

- Include continuous individual covariate, say  $x_i$ .
- $\text{logit}(\varphi_i) = \beta_1 + \beta_2 x_i$ .
- Let's investigate the effect of sex and wing length on dipper survival.

# Live demo #2



Can we explain time variation?

# Embrace heterogeneity

- Include temporal covariates, say  $x_t$ .
- $\text{logit}(\varphi_t) = \beta_1 + \beta_2 x_t$ .
- Let's investigate the effect of an extreme climatic event on dipper survival.

# Live demo #3



# What if covariates vary with individual and time?

- Think of age for example.
- In capture-recapture, age is the time elapsed since first capture. This is true age when animals are marked at birth.
- Now, think of body size across life.
- Problem is we cannot record size when animal is non-detected.
- Discretize in small, medium and large and treat as a state – more later.
- Assume a model for covariate and fill in missing values (imputation).

# Live demo #4



# Capture-recapture models rely on assumptions

- Design
  - No mark lost
  - Identity of individuals recorded without error (no false positives)
  - Captured individuals are a random sample
- Model
  - Homogeneity of survival and recapture probabilities
  - Independence between individuals (overdispersion)
- Test validity of assumptions
  - These assumptions should be valid, whatever inferential framework
  - Use goodness-of-fit tests – Pradel et al. (2005)
  - R implementation with package R2ucare ([paper](#) and [vignette](#))

# What does survival actually mean in capture-recapture ?

- Survival refers to the study area.
- Mortality and permanent emigration are confounded.
- Therefore we estimate apparent survival, not true survival.
- Apparent survival probability = true survival  $\times$  study area fidelity.
- Consequently, apparent survival < true survival unless study area fidelity = 1.
- Use caution when interpreting survival. If possible, combine with ring-recovery data, or go spatial to get closer to true survival.

# On our plate

- Estimating survival with capture-recapture models
- Including discrete and continuous covariates
- **Estimating transitions between sites/states**

On the move: Transition estimation

Two sites A and B











A little bit of history

## THE ESTIMATION OF POPULATION SIZE, MIGRATION RATES AND SURVIVAL IN A STRATIFIED POPULATION

A. Neil ARNASON

Computer Science Department, University of Manitoba,  
Winnipeg, Canada

### INTRODUCTION

CHAPMAN and JUNGE (1956, hereafter referred to as C & J) developed estimates of stratum size and migration rates for a population divided into  $n \geq 2$  areas (strata) when animals were free to migrate from area to area. The method was based on data from sampling and marking observations on two occasions. The method was extended by DARROCH (1961) to allow sampling in different numbers of strata at the two sampling times, and to show how to treat some special problems that arise when using the method. These problems arise when a particular data matrix (which must be inverted) is singular or ill-conditioned. The same problems could occur with the estimates which will be given in this paper.

In order to account for deaths or losses from the areas due to permanent emigration out of the areas being sampled, it is necessary to sample on at least three occasions. I developed estimates for the three sample experiment on two areas

BIOMETRICS 49, 177-193  
March 1993

## Estimating Migration Rates Using Tag-Recovery Data

Carl J. Schwarz

Department of Statistics, University of Manitoba,  
Winnipeg, Manitoba R3T 2N2, Canada

Jake F. Schweigert

Biological Sciences Branch, Pacific Biological Station,  
Department of Fisheries and Oceans, Nanaimo, British Columbia V9R 5K6, Canada

and

A. Neil Arnason

Department of Computer Science, University of Manitoba,  
Winnipeg, Manitoba R3T 2N2, Canada

### SUMMARY

Tag-recovery data are used to estimate migration rates among a set of strata. The model formulation is a simple matrix extension of the formulation of a tag-recovery experiment discussed by Brownie et al. (1985, *Statistical Inference from Band-Recovery Data—A Handbook*, 2nd edition, Washington, D.C.: U.S. Department of the Interior). Estimation is more difficult because of the convolution of parameters between release and recovery and this convolution may cause estimates of the survival/

Sites may be states.

## ESTIMATING TRANSITION PROBABILITIES FOR STAGE-BASED POPULATION PROJECTION MATRICES USING CAPTURE–RECAPTURE DATA<sup>1</sup>

JAMES D. NICHOLS AND JOHN R. SAUER

*United States Fish and Wildlife Service, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

KENNETH H. POLLOCK

*Institute of Statistics, North Carolina State University, Box 8203, Raleigh, North Carolina 27695-8203 USA*

JAY B. HESTBECK<sup>2</sup>

*United States Fish and Wildlife Service, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

**Abstract.** In stage-based demography, animals are often categorized into size (or mass) classes, and size-based probabilities of surviving and changing mass classes must be estimated before demographic analyses can be conducted. In this paper, we develop two procedures for the estimation of mass transition probabilities from capture–recapture data. The first approach uses a multistate capture–recapture model that is parameterized directly with the transition probabilities of interest. Maximum likelihood estimates are then obtained numerically using program SURVIV. The second approach involves a modification of Pollock's robust design. Estimation proceeds by conditioning on animals caught in a particular class at time  $i$ , and then using closed models to estimate the number of these that are alive in other classes at  $i + 1$ . Both methods are illustrated by application to meadow vole, *Microtus pennsylvanicus*, capture–recapture data. The two methods produced reasonable estimates that were similar. Advantages of these two approaches include the directness of estimation, the absence of need for restrictive assumptions about the independence of survival and growth, the testability of assumptions, and the testability of related hypotheses of ecological interest (e.g., the hypothesis of temporal variation in transition probabilities).

**Key words:** capture–recapture models; *Microtus pennsylvanicus*; multistate models; parameter estimation; Pollock's robust design; stage-based population projection matrices; stage transition prob-

## ESTIMATING BREEDING PROPORTIONS AND TESTING HYPOTHESES ABOUT COSTS OF REPRODUCTION WITH CAPTURE–RECAPTURE DATA<sup>1</sup>

JAMES D. NICHOLS AND JAMES E. HINES

*National Biological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

KENNETH H. POLLOCK

*Institute of Statistics, North Carolina State University, Box 8203, Raleigh, North Carolina 27695-8203 USA*

ROBERT L. HINZ AND WILLIAM A. LINK

*National Biological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

**Abstract.** The proportion of animals in a population that breeds is an important determinant of population growth rate. Usual estimates of this quantity from field sampling data assume that the probability of appearing in the capture or count statistic is the same for animals that do and do not breed. A similar assumption is required by most existing methods used to test ecologically interesting hypotheses about reproductive costs using field sampling data. However, in many field sampling situations breeding and nonbreeding animals are likely to exhibit different probabilities of being seen or caught. In this paper, we propose the use of multistate capture–recapture models for these estimation and testing problems. This methodology permits a formal test of the hypothesis of equal capture/sighting probabilities for breeding and nonbreeding individuals. Two estimators of breeding proportion (and associated standard errors) are presented, one for the case of equal capture probabilities and one for the case of unequal capture probabilities. The multistate modeling framework also yields formal tests of hypotheses about reproductive costs to future reproduction or survival or both fitness components. The general methodology is illustrated using capture–recapture data on female meadow voles, *Microtus pennsylvanicus*. Resulting estimates of the proportion of reproductively active females showed strong seasonal variation, as expected, with low breeding proportions in midwinter. We found no evidence of reproductive costs extracted in subsequent survival or reproduction. We believe that this methodological framework has wide application to problems in animal ecology concerning breeding proportions and phenotypic reproductive costs.

**Key words:** capture/sighting probability; *Microtus pennsylvanicus*; multistate capture–recapture

# Examples of multistate models

- *Epidemiological or disease states:* sick/healthy, uninfected/infected/recovered.
- *Morphological states:* small/medium/big, light/medium/heavy.
- *Breeding states:* e.g. breeder/non-breeder, failed breeder, first-time breeder, breeder with 1 or 2 offspring.
- *Developmental or life-history states:* e.g. juvenile/subadult/adult.
- *Social states:* e.g. solitary/group-living, subordinate/dominant.
- *Death states:* e.g. alive, dead from harvest, dead from natural causes.

**States = individual, time-specific categorical covariates.**

Let's practice!

# Wintering site fidelity in Canada Geese



3 sites Carolinas, Chesapeake, Mid-Atlantic, with 21277 banded geese,  
data kindly provided by Jay Hestbeck

	year_1984	year_1985	year_1986	year_1987	year_1988	year_1989
	0	2	2	0	0	0
	0	0	0	0	0	2
	0	0	0	1	0	0
	0	0	2	0	0	0
	0	3	0	0	3	2
	0	0	0	2	0	0
	2	2	0	2	3	2
	0	0	0	0	2	2

# Live demo #5



# On our plate

- Estimating survival with capture-recapture models
- Including discrete and continuous covariates
- Estimating transitions between sites/states

# Further reading

- Lebreton, Burnham, Clobert, and Anderson (1992). *Modeling Survival and Testing Biological Hypotheses Using Marked Animals: A Unified Approach with Case Studies.* Ecological Monographs 62: 67-118.
- Lebreton, Nichols, Barker, Pradel and Spendelow (2009). *Modeling Individual Animal Histories with Multistate Capture–Recapture Models.* Advances in Ecological Research 41:87-173.
- Cooch and White (2021). *Program MARK - a 'gentle introduction'.*
- Williams, Nichols and Conroy (2002). *Analysis and Management of Animal Populations.*
- Pradel, Gimenez and Lebreton (2005). *Principles and interest of GOF tests for multistate capture-recapture models.* Animal Biodiversity and Conservation 28: 189–204.

# Population Projection Models (PPMs)

Sarah Cubaynes for the Team

last updated: 2022-03-18

# What we've learned so far

Obj. 1: Assess current and past trends in population abundance

- see class 1

Obj. 2: Estimate demographic parameters and identify causes of variation

- see class 2

Obj. 3: Evaluate population viability to inform decision about management actions

- some hints about this now!

# Population Viability Analysis (PVA)

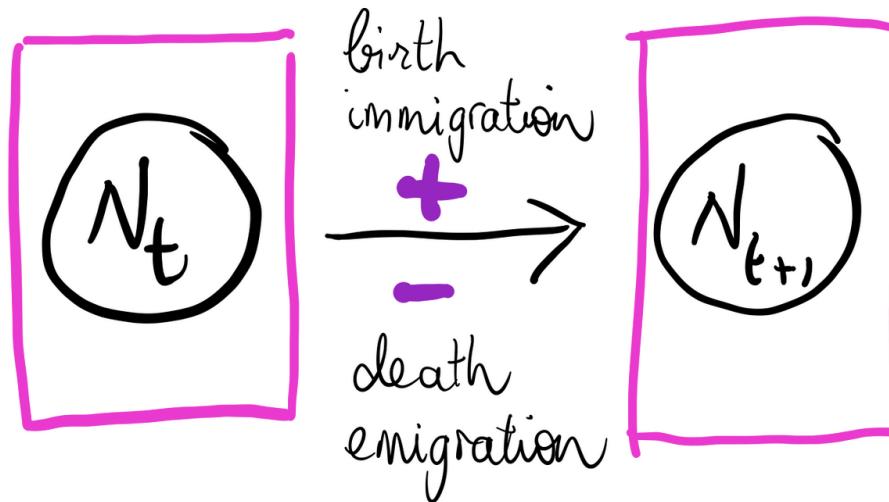
Morris et al. (1999). *A practical handbook for population viability analysis*. The Nature Conservancy.

- Use of **quantitative methods** to predict the **likely future status** of a population or collection of populations of conservation concern
- **Tentative assessments** based upon what we now know **rather than as iron-clad predictions** of population fate

# Why is PVA useful ?

- Quantify rate of population change over time
- Estimate extinction risks (used by IUCN)
- Identify key parameters for population management
- Evaluate and compare relative impact of population management actions

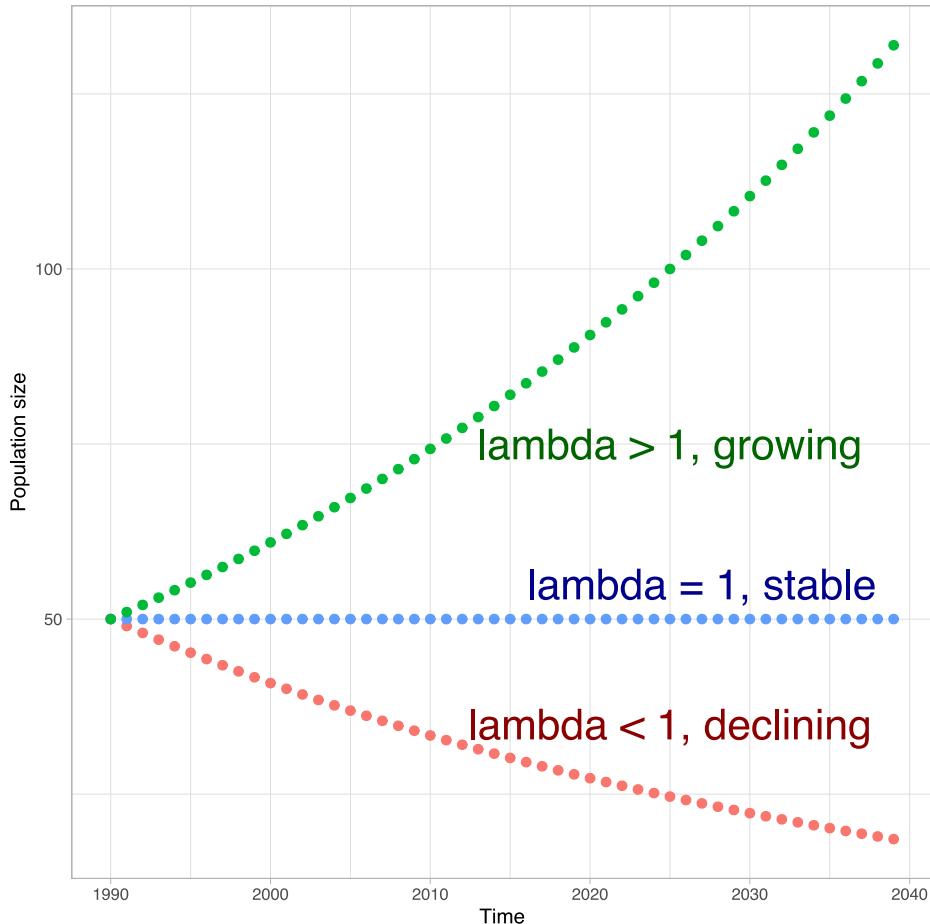
# Can we assess population viability from counts ?



Population size and  
density

# 1- Count-based PVA

# Finite population growth rate $\lambda$



- Exponential growth or decay at a constant rate of change
- $\lambda = \frac{N_{t+1}}{N_t}$  gives the **proportional change** in population size
- After  $t$  time steps, population size will be  $N_t = N_0 \cdot \lambda^t$
- $\lambda$  is log-normally distributed

# Growth rate versus intrinsic rate of increase

- Population growth rate  $\lambda \sim \text{Lognormal}(\mu, \sigma^2)$
- Easier to work with the **intrinsic rate of increase**

$$r = \log(\lambda) = \log\left(\frac{N_{t+1}}{N_t}\right) = \log(N_{t+1}) - \log(N_t)$$

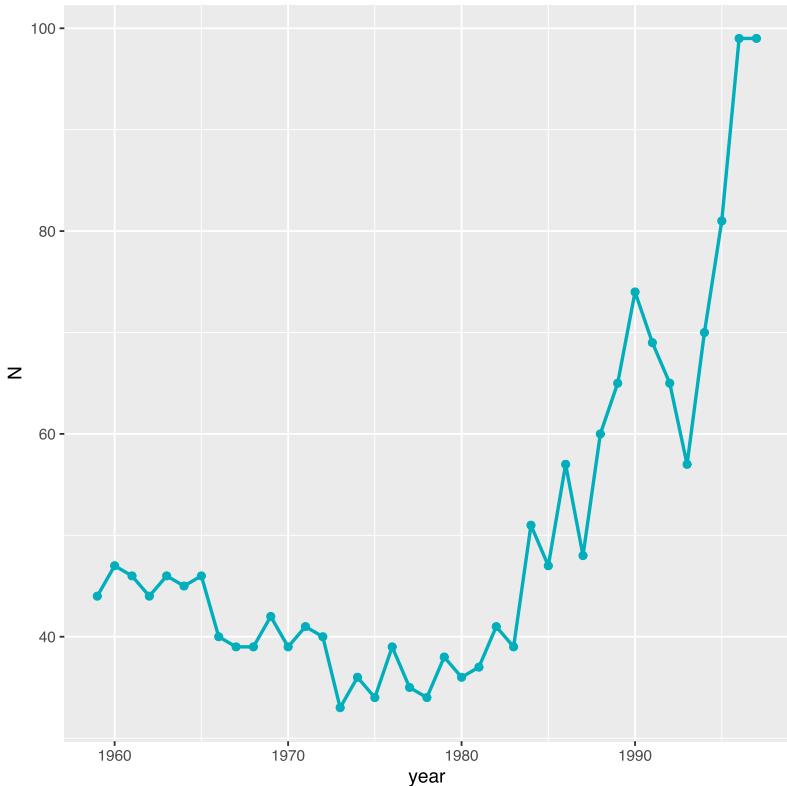
- Which is normally distributed  $r \sim N(\mu, \sigma^2)$
- $\mu$  is the **mean rate of increase**
- $\sigma^2$  is the **environmental variance**

# Environmental variance

- If the mean rate of increase  $\mu < 0$ , extinction will certainly occur.
- A population can still decline or go extinct even if the mean rate of increase  $\mu > 0$ , because of environmental variance  $\sigma^2$
- Variable environments increase extinction risks

# Example of the Yellowstone grizzly bear population

From Morris & Doak (2002). *Quantitative conservation biology: Theory and practice of population viability analysis*. Massachusetts, USA.



# Step 1: Calculate $\mu$ and $\sigma^2$ from the data

```
# rate of increase over years  
logN <- log(N[-1]/N[-length(N)]) # log(Nt+1) - log(Nt)  
#mean rate of increase  
mu <- mean(logN)  
#environmental variance  
sigma2 <- var(logN)
```

```
mu      sigma2  
1 0.02134027 0.01305092
```

- $\mu > 0$  so on average the population is growing
- $\sigma^2 = 0.013$  reflects low inter-annual variance in the rate of increase

# or using linear regression for unequal time intervals

```
x <- sqrt(grizzly$year[-1] - grizzly$year[-length(grizzly$year)]  
y <- logN / x  
mod <- lm(y ~ 0 + x) # forcing a intercept of zero  
mod
```

Call:

```
lm(formula = y ~ 0 + x)
```

Coefficients:

x
0.02134

```
mu <- coef(mod) # slope = mean intrinsic rate of increase
```

# or using linear regression for unequal time intervals

```
# get an estimate for sigma2  
anova(mod)
```

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x	1	0.01731	0.017306	1.326	0.2569
Residuals	37	0.48288	0.013051		

```
sigma2 <- anova(mod)[["Mean Sq"]][2] # environmental variance
```

# Get confidence intervals for $\mu$ and $\sigma_2$

```
## Confidence interval for mu  
confint(mod, 1)
```

```
2.5 %      97.5 %  
x -0.01620969 0.05889023
```

- Confidence interval of mean rate of increase encompasses 0, therefore we cannot rule out a potential risk of decline!

```
## Confidence interval for sigma 2  
df1 <- length(logN) - 1  
df1 * sigma2 / qchisq(c(.975, .025), df = df1)
```

```
[1] 0.008674359 0.021844393
```

# Back-transform to get finite population growth rate $\bar{\lambda}$

```
lambda <- exp(mu)
lambda # average growth rate
```

x  
1.02157

- Here  $\bar{\lambda} > 1$ , so the grizzly population is growing on average
- It does not rule out the possibility of a decline owing to the chance occurrence of a sequence of bad years (remember confidence interval)

# Step 2: Project the population

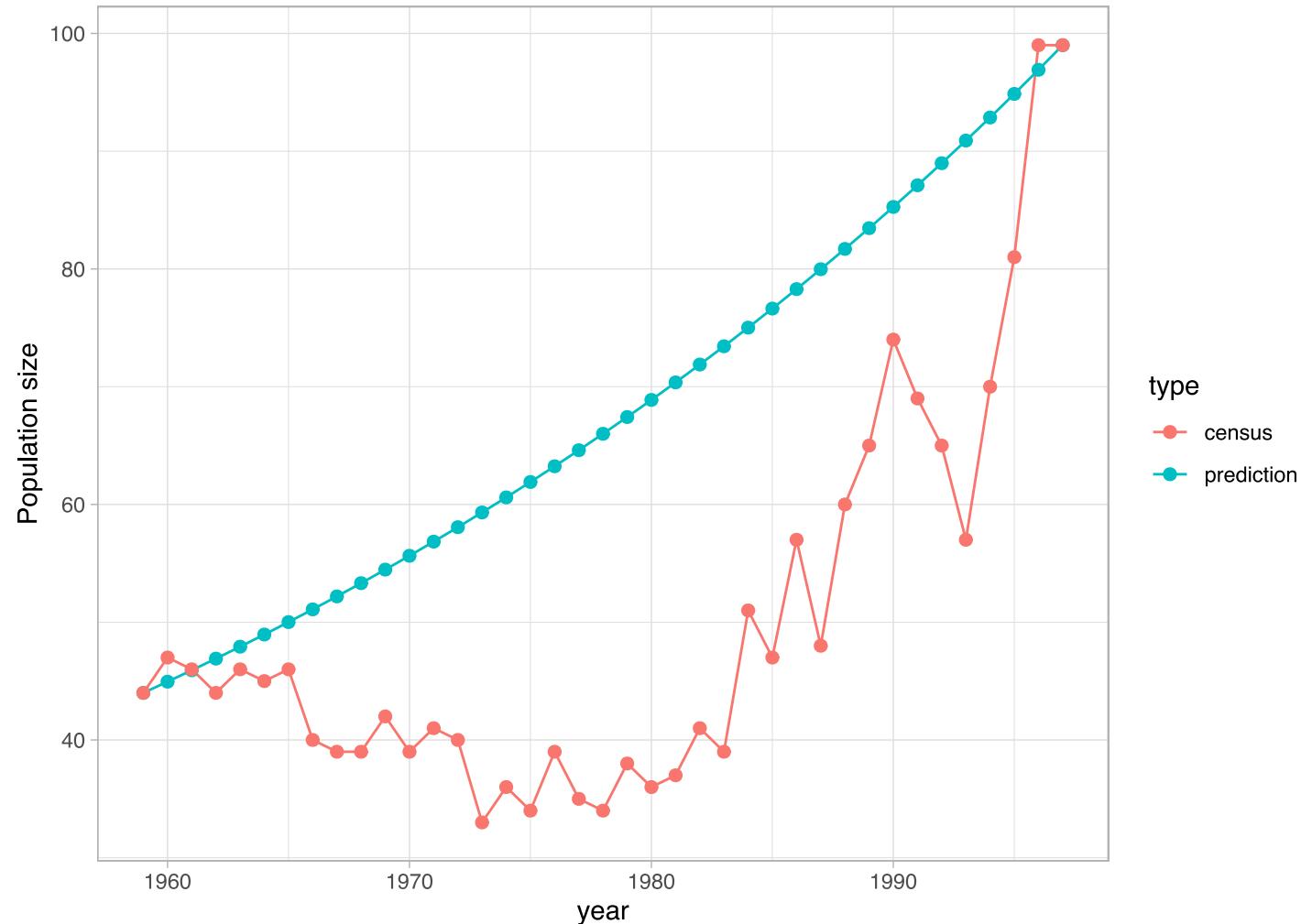
- Expected population size using mean rate of increase and ignoring environmental variance (not recommended)

$$N_t = N_0 \cdot \mu^t$$

$$\ln(N_t) = \ln(N_0) + \mu \cdot t$$

- Initial population  $N_0 = 44$
- Time steps to project over:  $t = 38$  years (nb. years - 1)
- Mean rate of increase  $\mu = 0.021$

# Step 2: Project the population



# Step 2: Project the population

Let's account for observed variation in growth rate

- First, set the initial population, number of time steps, and number of repetitions:

```
n0 <- grizzly$N[1] # initial pop.
T <- 50 # time iterations to project over
runs <- 500 # number of repetitions (pop. trajectories)
stoch.pop <- matrix(NA,T,runs) # to store results
stoch.pop[1,] <- n0 # initiate
```

# Step 2: Project the population

Let's account for observed variation in growth rate

- Then set a quasi-extinction threshold

```
Ne <- 30 # threshold for minimum viable pop.
```

- 1 female or a minimum viable population (genetic drift, demographic stochasticity)
- Can also be the lowest level of abundance at which it remains feasible to attempt intervention to prevent further decline.

# Now run the projections

```
# let's project the population
for (i in 1:runs){
  for (t in 2:T){
    # Draw r from normal using estimates of mu and sigma2
    r <- rnorm(n = 1, mean = mu, sd = sqrt(sigma2))
    # back-transform to get lambda and get pop. size
    lambda <- exp(r)
    #project one time step from the current pop size
    stoch.pop[t,i] <- stoch.pop[(t-1),i] * lambda
    # leave the loop if pop <= threshold
    if(stoch.pop[t,i] <= Ne){
      stoch.pop[t,i] <- 0
      i < i+1}
  }
}
```

# Now run the projections

```
# let's project the population
for (i in 1:runs){ # loop over repetitions
  for (t in 2:T){ # loop over years
    # Draw r from normal using estimates of mu and sigma2
    r <- rnorm(n = 1, mean = mu, sd = sqrt(sigma2))
    # back-transform to get lambda and get pop. size
    lambda <- exp(r)
    #project one time step from the current pop size
    stoch.pop[t,i] <- stoch.pop[(t-1),i] * lambda
    # leave the loop if pop <= threshold
    if(stoch.pop[t,i] <= Ne){
      stoch.pop[t,i] <- 0
      i < i+1}
  }
}
```

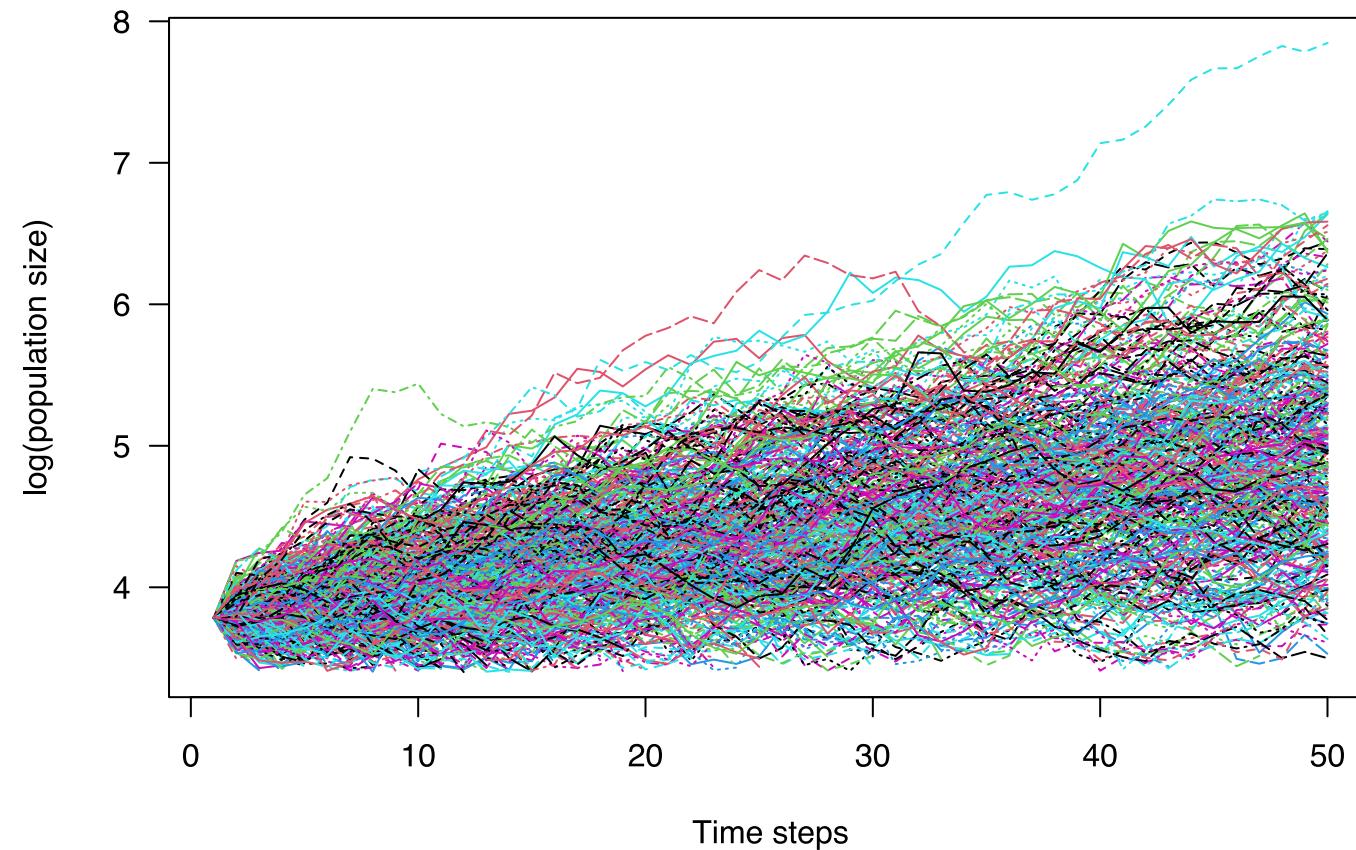
# Now run the projections

```
# let's project the population
for (i in 1:runs){ # loop over repetitions
  for (t in 2:T){ # loop over years
    # Draw r from normal using estimates of mu and sigma2
    r <- rnorm(n = 1, mean = mu, sd = sqrt(sigma2))
    # back-transform to get lambda and get pop. size
    lambda <- exp(r)
    #project one time step from the current pop size
    stoch.pop[t,i] <- stoch.pop[(t-1),i] * lambda
    # leave the loop if pop <= threshold
    if(stoch.pop[t,i] <= Ne){
      stoch.pop[t,i] <- 0
      i < i+1}
  }
}
```

# Now run the projections

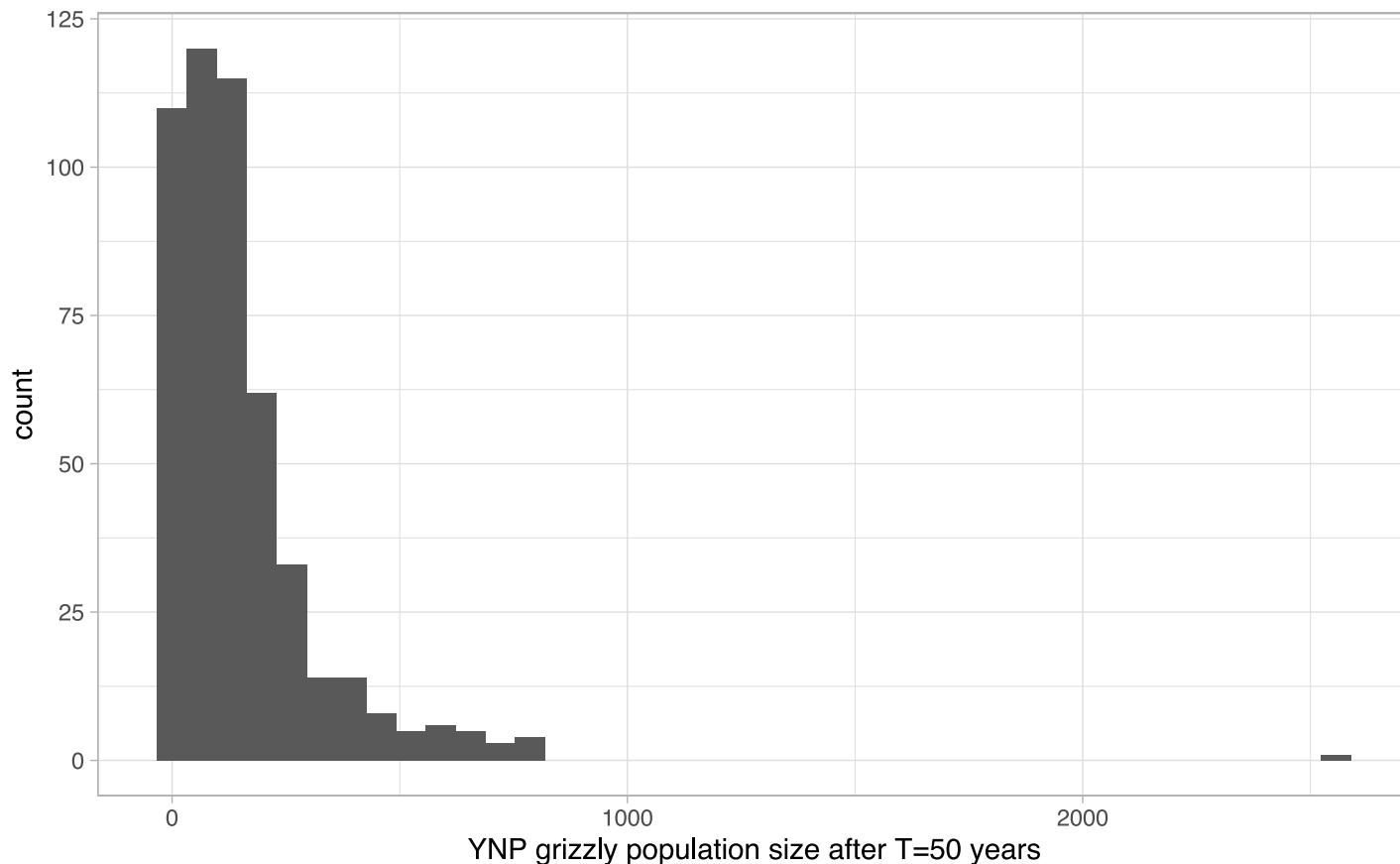
```
# let's project the population
for (i in 1:runs){ # loop over repetitions
  for (t in 2:T){ # loop over years
    # Draw r from normal using estimates of mu and sigma2
    r <- rnorm(n = 1, mean = mu, sd = sqrt(sigma2))
    # back-transform to get lambda and get pop. size
    lambda <- exp(r)
    #project one time step from the current pop size
    stoch.pop[t,i] <- stoch.pop[(t-1),i] * lambda
    # leave the loop if pop <= threshold
    if(stoch.pop[t,i] <= Ne){
      stoch.pop[t,i] <- 0
      i < i+1}
  }
}
```

# Step 3: Examine the results



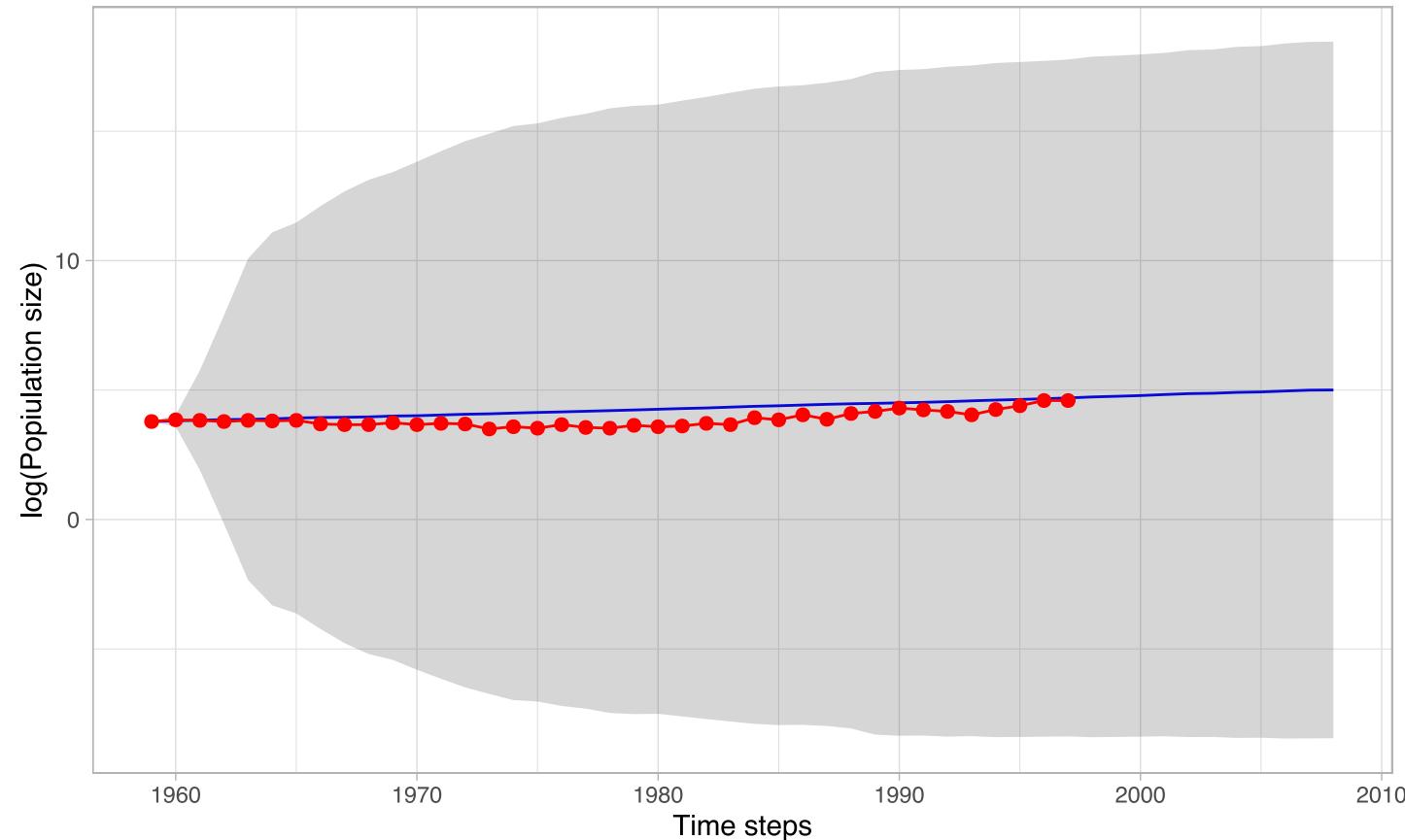
# Step 3: Examine the results

Plot population size at the last time step:



# Step 3: Examine the results

Plot prediction together with observed counts:



# Step 4: Quantify extinction risks

- The average population growth rate doesn't do a good job at predicting what most population realizations will do
- What are the chances that the population will go extinct at various times?
  - Extinction risk
  - Time to extinction

# Step 4: Quantify extinction risks

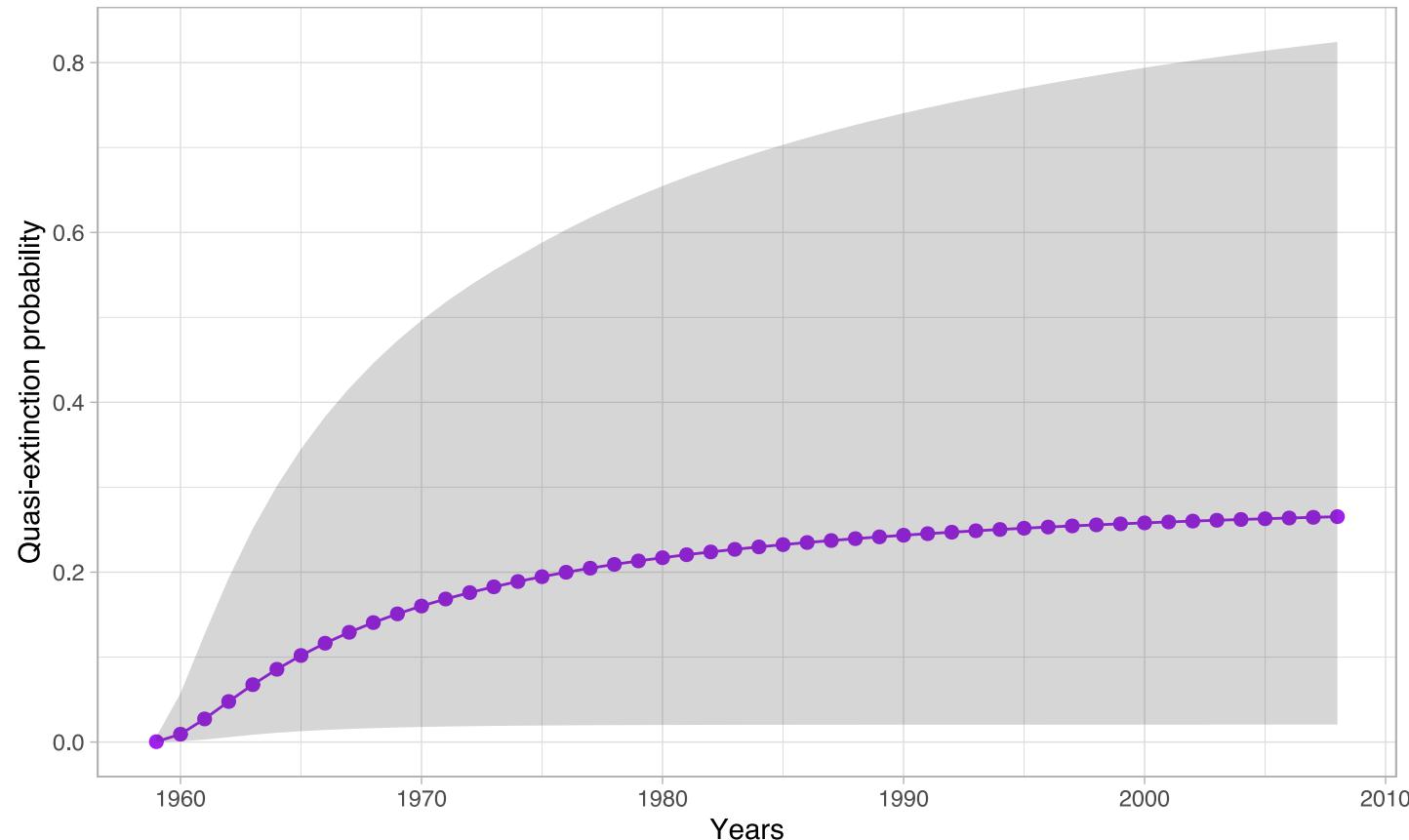
- Ultimate extinction probability = percentage of trajectories (over the 500 runs) reaching the extinction threshold at some point (over T years)

```
Pr.ext <- sum(lastN <= Ne) / runs # prob. to reach the extinction threshold  
Pr.ext
```

```
[1] 0.22
```

# Step 4: Quantify extinction risks

- Cumulative extinction risk over the years



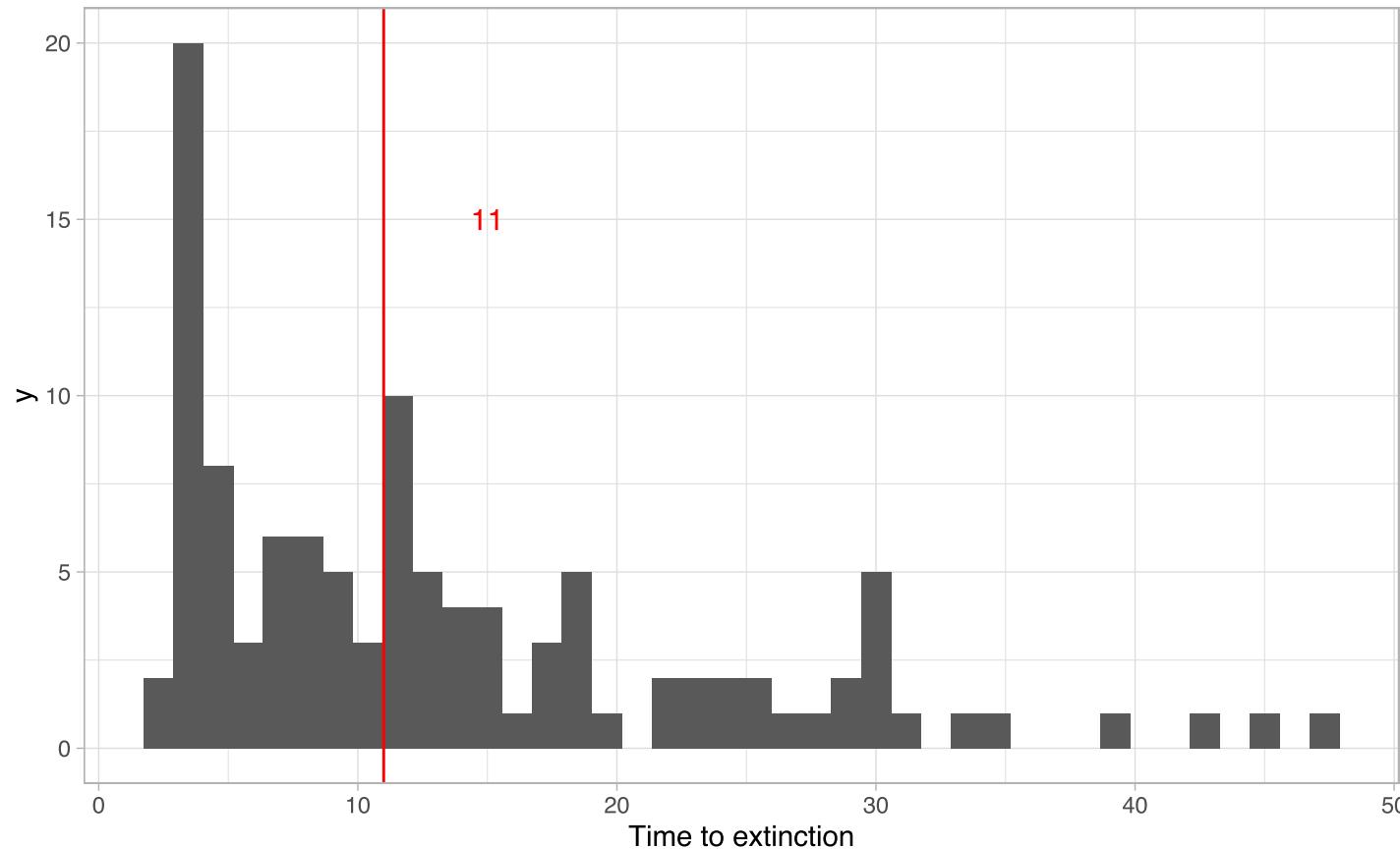
# Step 5: Quantify time to extinction

- Mean time to extinction is an overestimation (because of few pop. growing fast)
- Median time to extinction is a better measure

```
# Time to reach extinction for extinct pop.
maxt <- NULL      # empty vector to store results
for (i in 1:runs){    # loop over repetitions
  N <- stoch.pop[,i]
  # max time N > threshold
  maxt[i] <- max(which(N>0)) }
# time at extinction for pseudo-extinct populations
time.ext <- maxt[maxt < T]
median(time.ext)
```

[1] 11

# Step 5: Quantify time to extinction



# Step 6: Perturb and run the model

Interesting to evaluate the sensitivity of the results to changes in:

- Initial population size
- Extinction threshold
- Amount of environmental variance
- Number of time steps

# Live demo on grizzly bears

From Morris & Doak (2002). *Quantitative conservation biology: Theory and practice of population viability analysis*. Massachusetts, USA.



# Count-based extinction analyses are based on strong hypotheses

- Exhaustive counts (no sampling error)
- No density-dependence (exponential growth)
- Only source of variation is environmental stochasticity (no demographic stochasticity, no trends in mean or variance over time, uncorrelated environment among successive years)
- Moderate environmental variability (no catastrophe, no bonanzas)

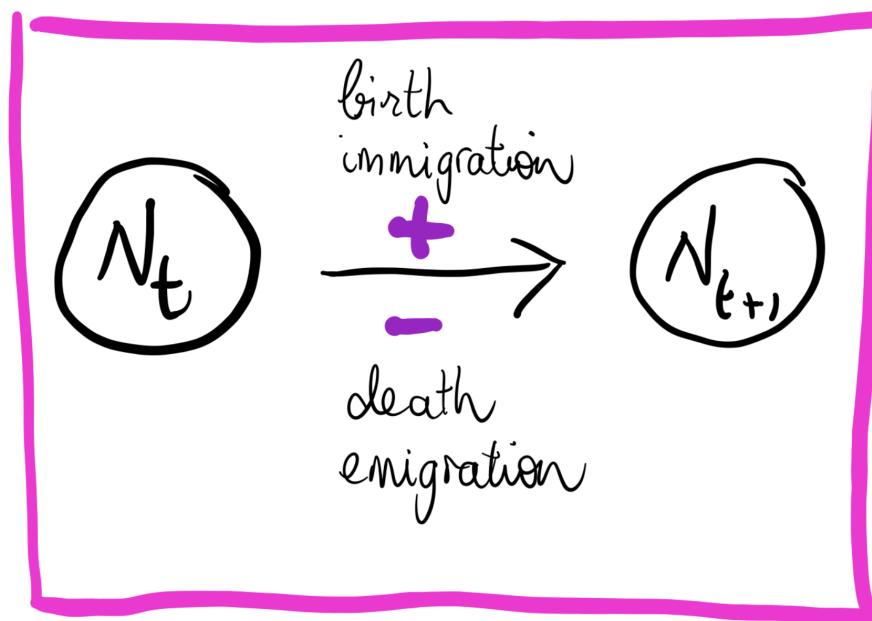
# Advantages

- Simplicity (data requirement at least 10 censuses and calculation)
- Work relatively well when assumptions are met
- Assess model quality by hindcasting

# Limitations

- $\lambda$  is only a summary of the population dynamics
- No info about the mechanisms governing the dynamics
- No hints about which management action might be most efficient?
- Is it better to act on survival? Fecundity? That of adults? Of juveniles?

# Can we assess population viability from demographic parameters ?



Population projection models

## 2- PVA using Matrix Projection Models (MPMs)

# Let's assume no migrations for now

- Survival and fecundity rates are enough to fully describe the population dynamics

$$N_{t+1} = N_t * F + N_t * S$$

- F = fecundity
- S = survival (= 1- mortality)

# Demographic parameters (see class 2)

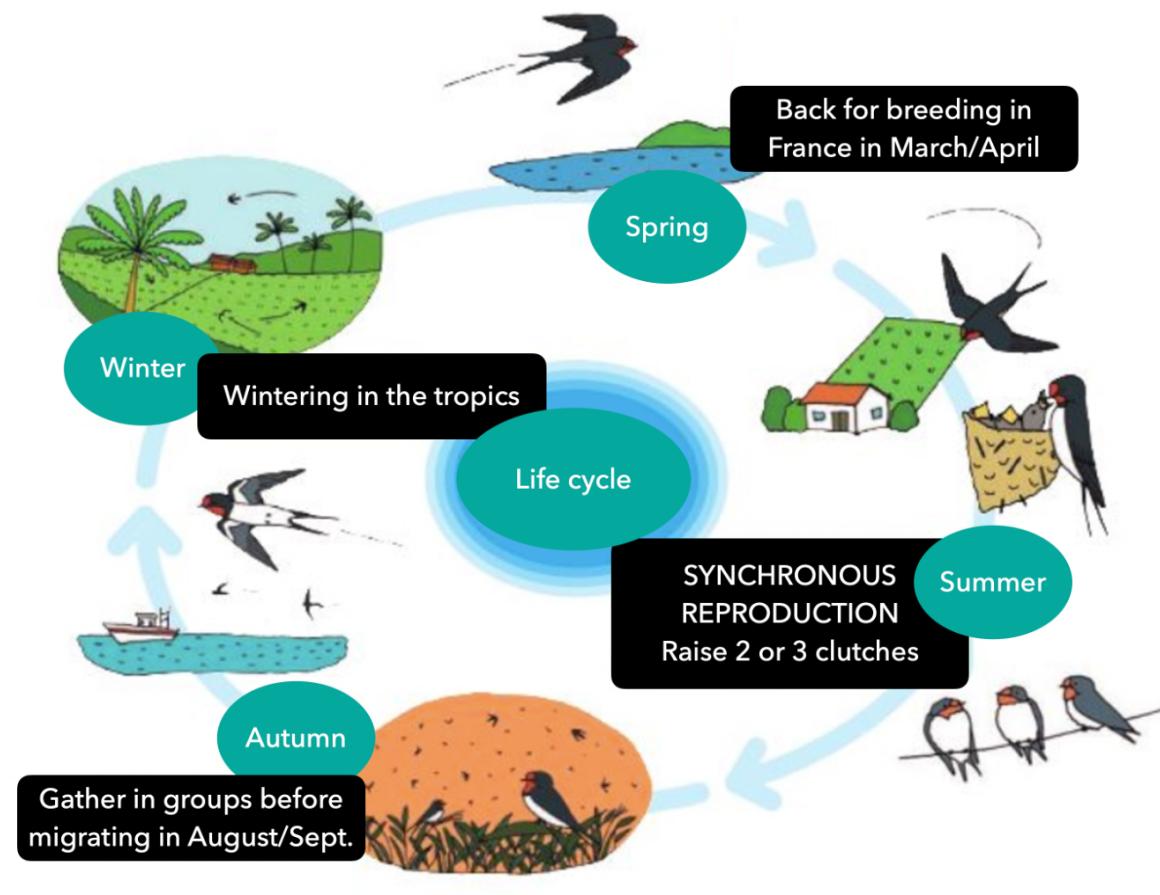
- Demographic parameters are often **heterogeneous across the life cycle**: survival and fecundity vary with age and/or stage in many species.



# Matrix Population Models (MPMs)

- Incorporate vital rates that are heterogeneous across the life cycle
- Project the population based on the matrix summarizing the age- or stage- dependent demographic parameters

# The Barn swallow (*Hirundo rustica*) example



# Demographic parameters

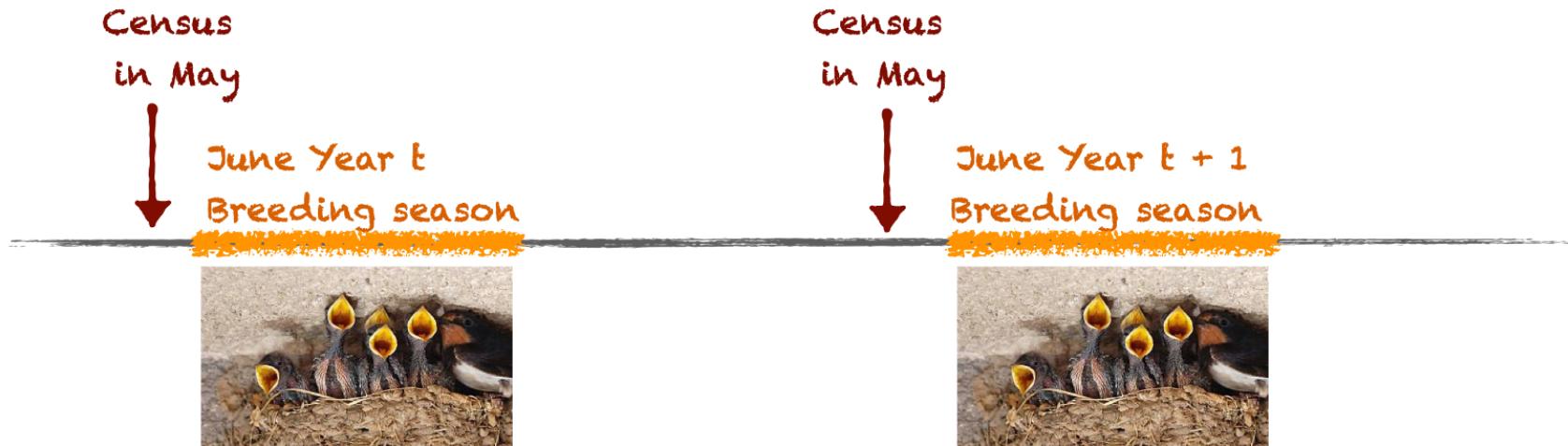
Survival rates:

- $S_0$  chick survival
- $S_1$  juvenile survival (1 yo)
- $S_2$  adult survival (2+ yo)

Fecundity:

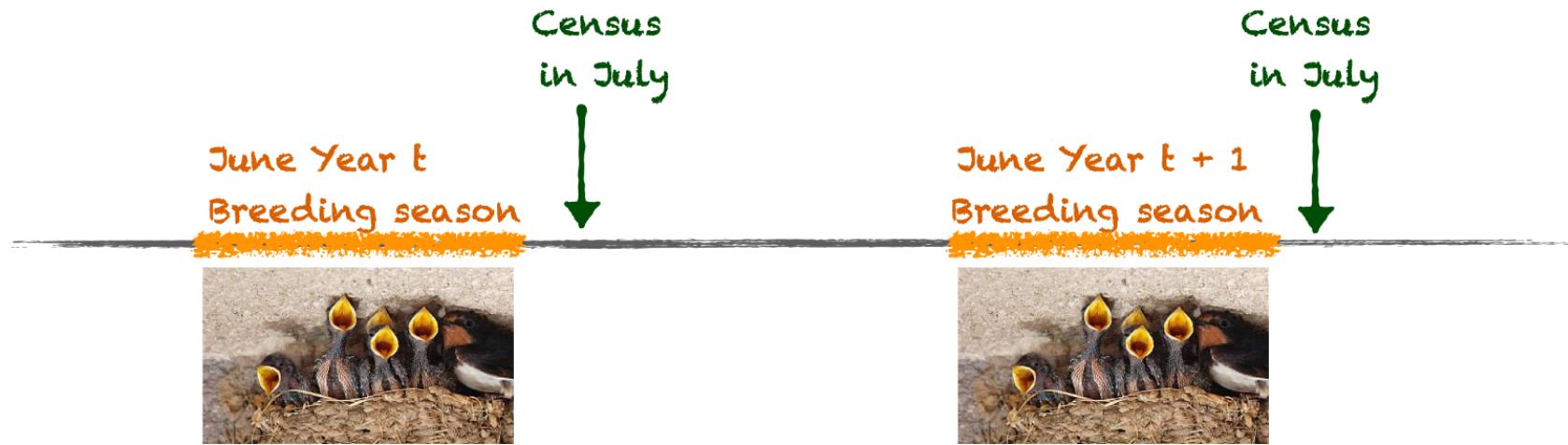
- $F_1$  number of females produced by a juvenile female
- $F_2$  number of females produced by an adult female

# Timing of data collection ?



- Pre-breeding census

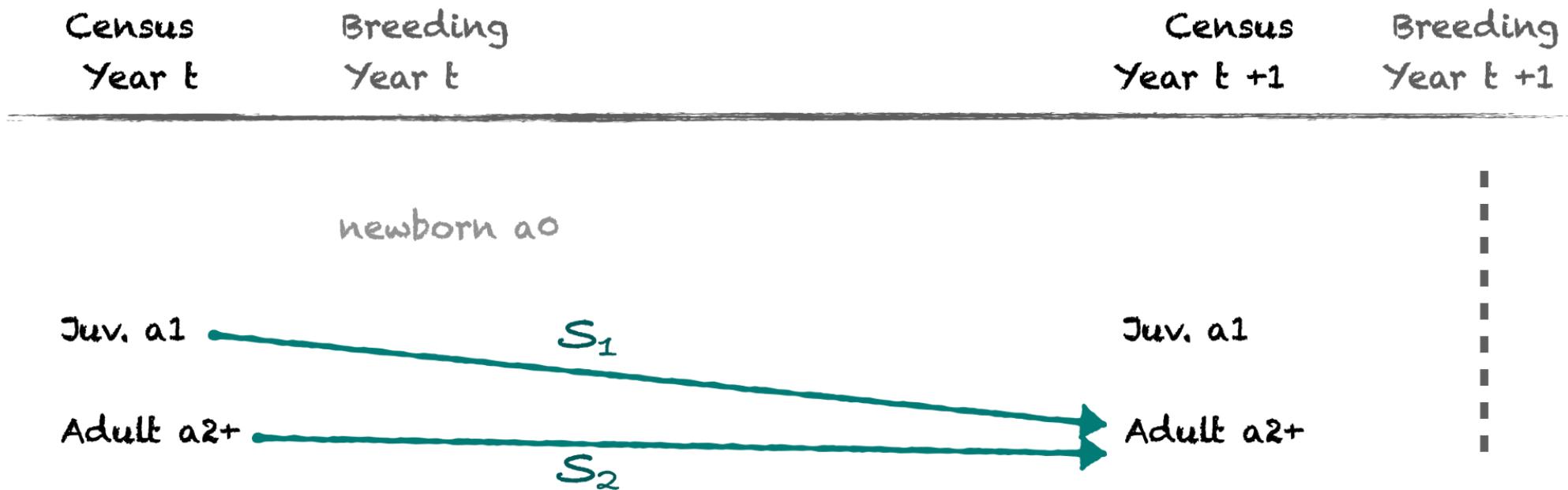
# Timing of data collection ?



- Post-breeding census

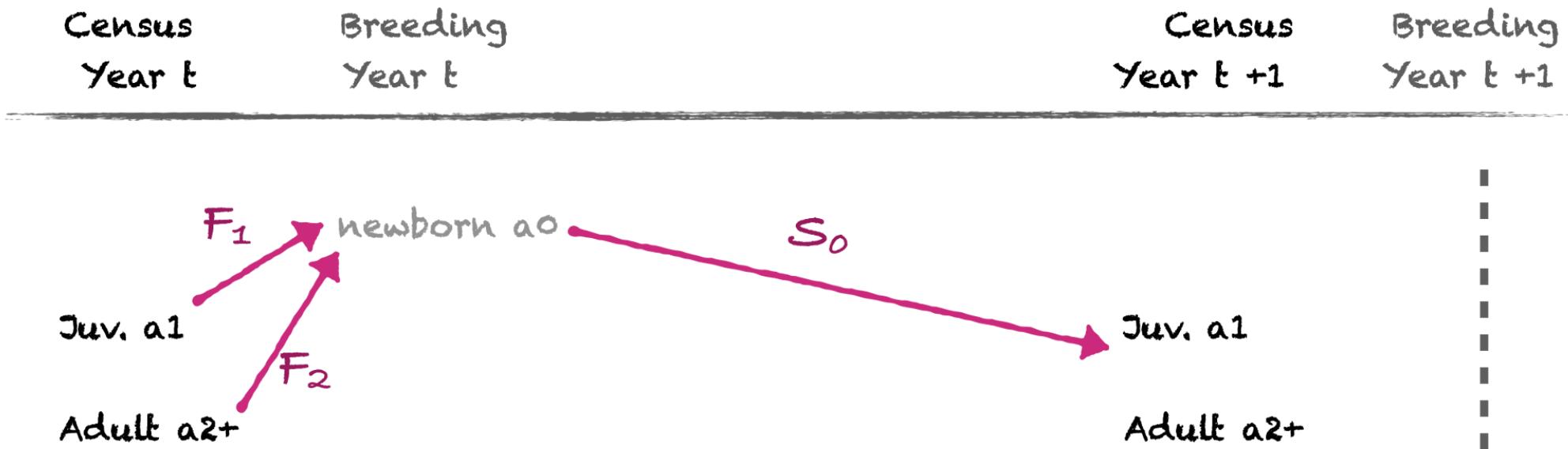
# Step 1: write the agenda of events

- Let's first consider a pre-breeding census



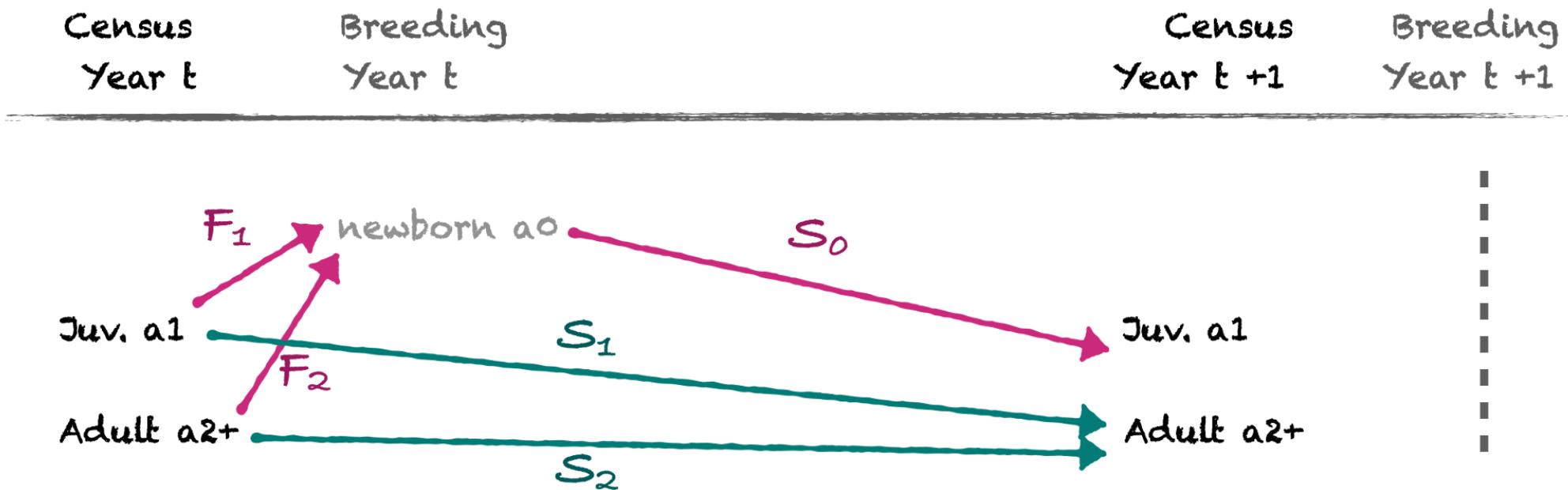
# Step 1: write the agenda of events

- Let's first consider a pre-breeding census



# Step 1: write the agenda of events

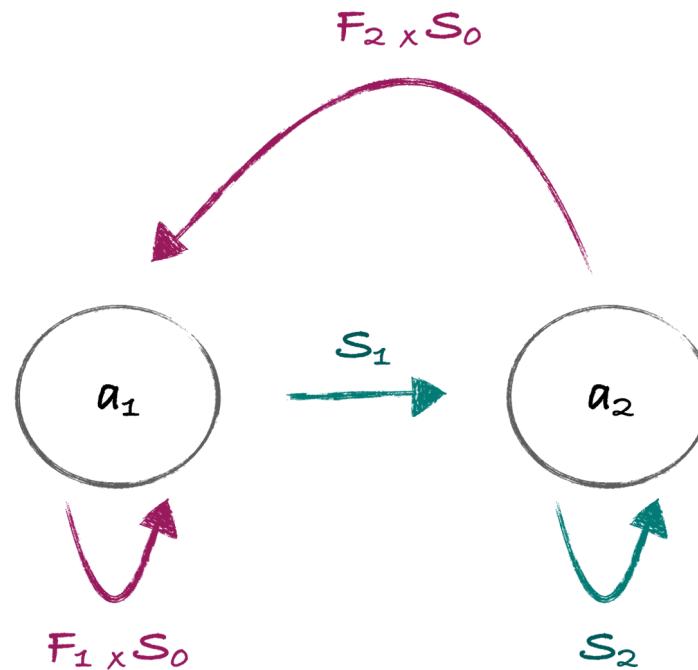
- Let's first consider a pre-breeding census



- Note that newborn are not observed directly!

# Step 2: Translate into life cycle graph or 'Caswell representation'

- A trick is to go 'up the arrows'



# Step 3: Translate into equations

- Link  $N_{(t+1)}$  to  $N_{(t)}$  via survival and fertility rates
- A trick is to read the parameters going up the arrows

$$N_{(1,t+1)} = F_1 \cdot S_0 \cdot N_{(1,t)} + F_2 \cdot S_0 \cdot N_{(2,t)}$$

$$N_{(2,t+1)} = S_1 \cdot N_{(1,t)} + S_2 \cdot N_{(2,t)}$$

# Step 4: Arrange in a matrix format

- Called the **transition matrix**, or the **projection matrix**

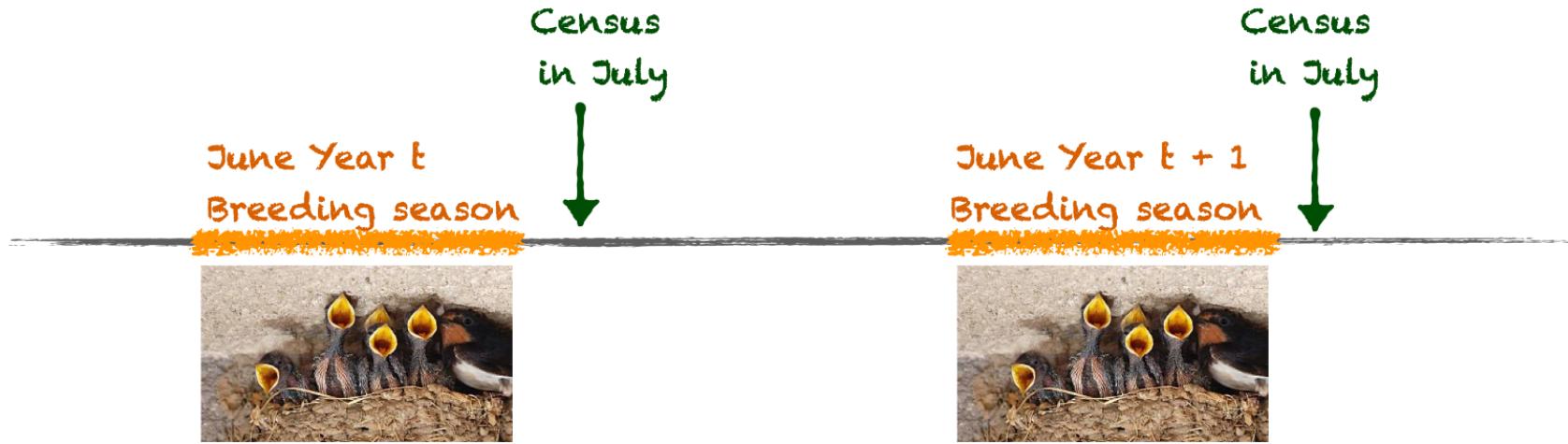
$$N_{(t+1)} = \begin{bmatrix} F_1 \cdot S_0 & F_2 \cdot S_0 \\ S_1 & S_2 \end{bmatrix} \cdot N_{(t)}$$

# Step 4: Arrange in a matrix format

- With  $S_0 = 0.2$ ,  $S_1 = 0.5$  and  $S_2 = 0.65$
- $F_1 = 3/2$  and  $F_2 = 6/2$

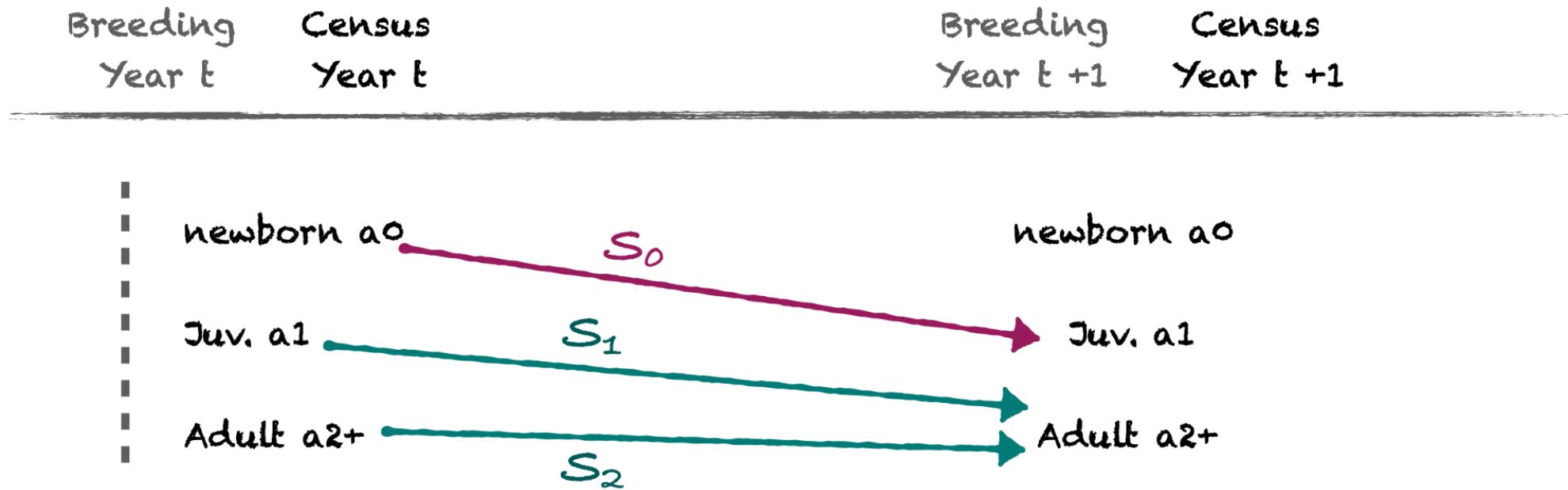
$$N_{(t+1)} = \begin{bmatrix} 0.3 & 0.6 \\ 0.5 & 0.65 \end{bmatrix} \cdot N_{(t)}$$

# What is the difference with a post-breeding census ?



- Post-breeding census

# Step 1: write the agenda of events

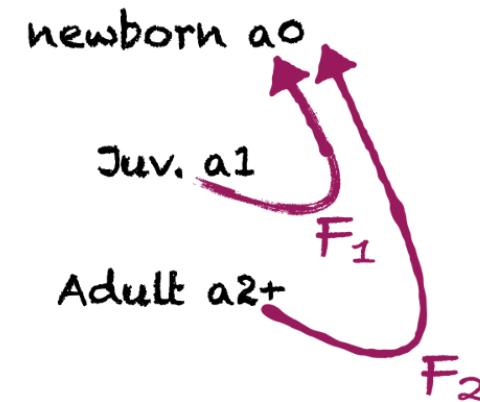


# Step 1: write the agenda of events

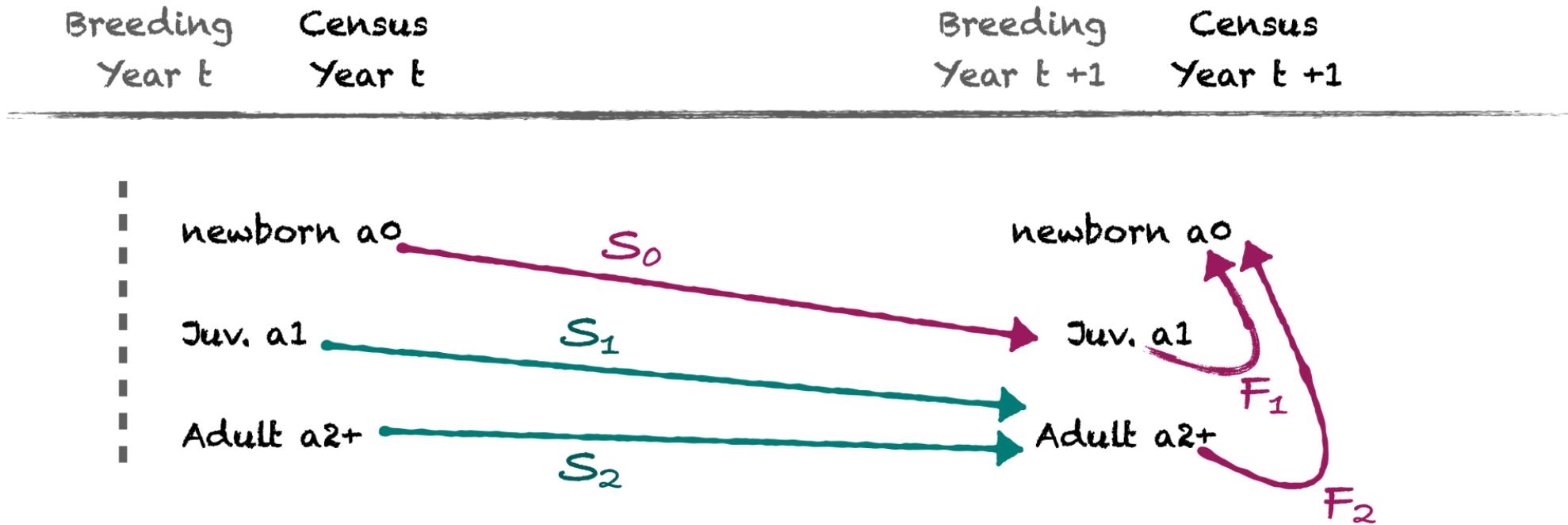
Breeding      Census  
Year t          Year t

Breeding      Census  
Year t +1      Year t +1

newborn a0  
Juv. a1  
Adult a2+

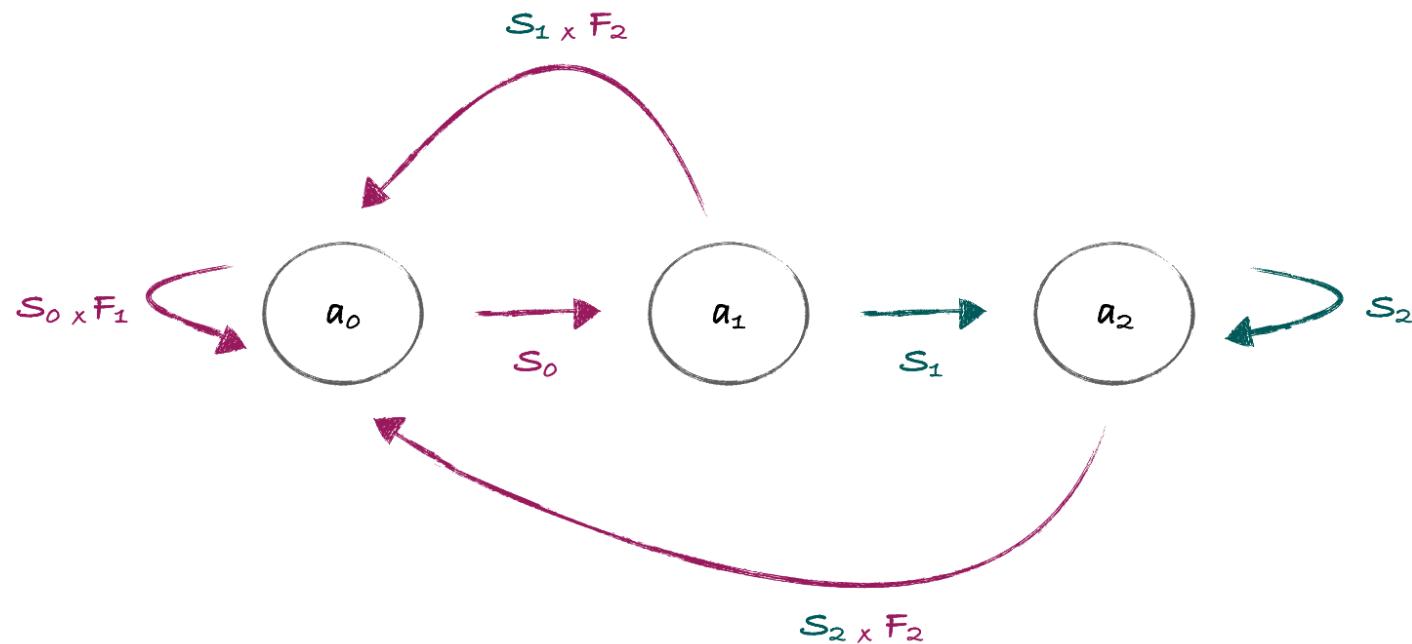


# Step 1: write the agenda of events



- Note that newborn are now observed!

# Step 2: Translate into life cycle graph or 'Caswell representation'



## Step 3: Translate into equations

$$N_{(0,t+1)} = S_0 \cdot F_1 \cdot N_{(0,t)} + S_1 \cdot F_2 \cdot N_{(1,t)} + S_2 \cdot F_2 \cdot N_{(2,t)}$$

$$N_{(1,t+1)} = S_0 \cdot N_{(0,t)}$$

$$N_{(2,t+1)} = S_1 \cdot N_{(1,t)} + S_2 \cdot N_{(2,t)}$$

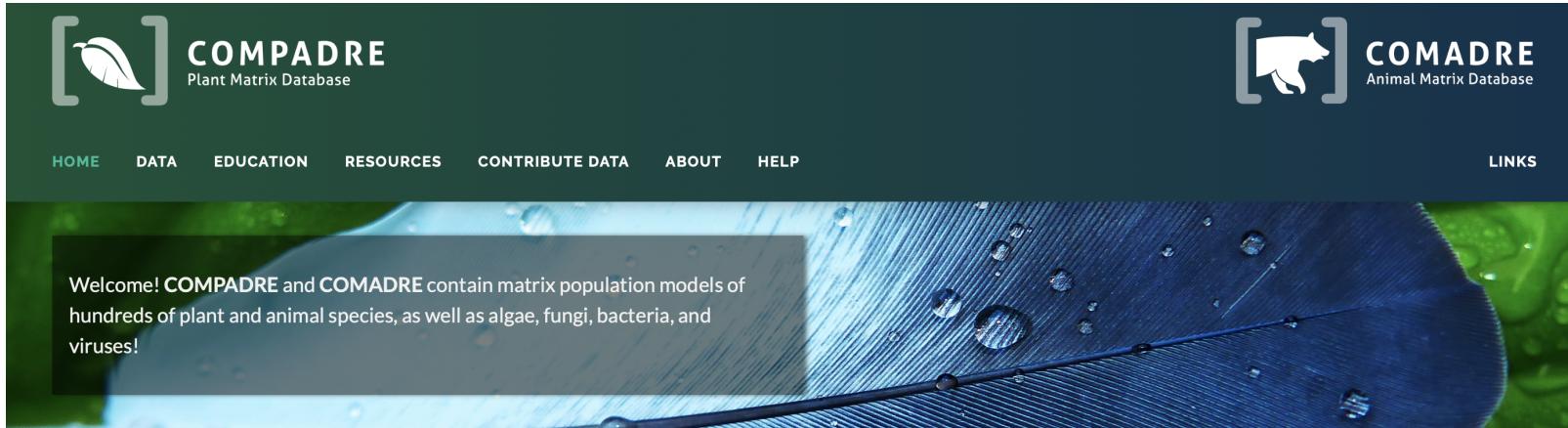
## Step 4: Arrange in matrix format

$$N_{(t+1)} = \begin{bmatrix} S_0 \cdot F_1 & S_1 \cdot F_2 & S_2 \cdot F_2 \\ S_0 & 0 & 0 \\ 0 & S_1 & S_2 \end{bmatrix} \cdot N_{(t)}$$

# Why is the matrix format interesting ?

- Easier to read than multiple equations
- Intrinsic numeric features (back to it later)
- Work the same way for complex life cycles

# Examples



Taxonomic Species

755

Studies

638

Matrix Population Models

8687

Taxonomic Species

415

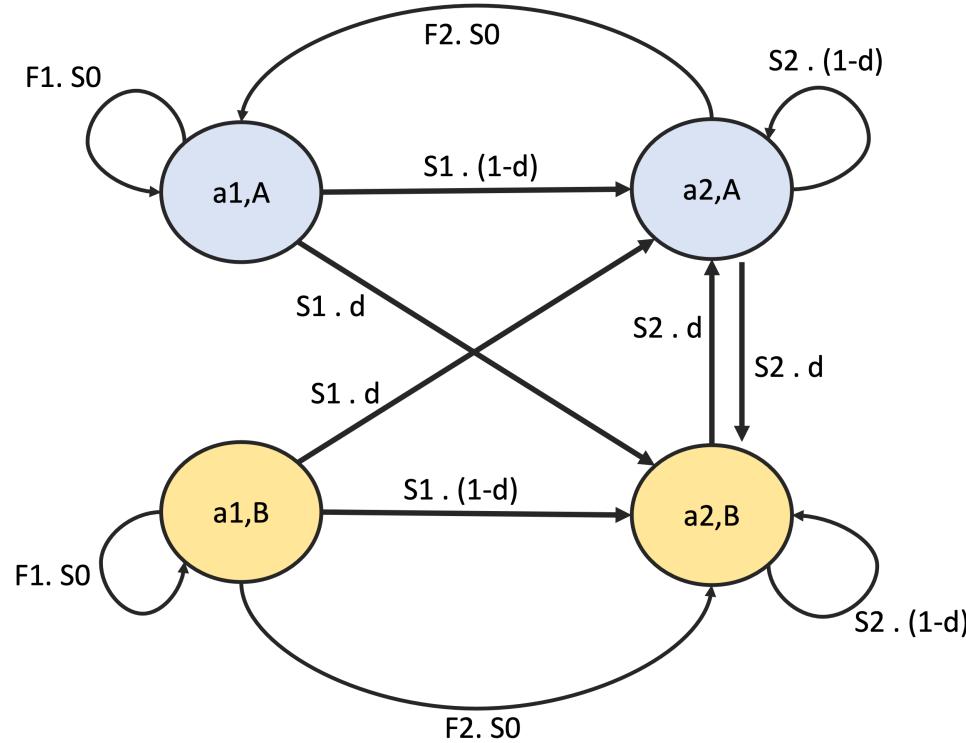
Studies

395

Matrix Population Models

3317

# Several sites: Barn swallow example

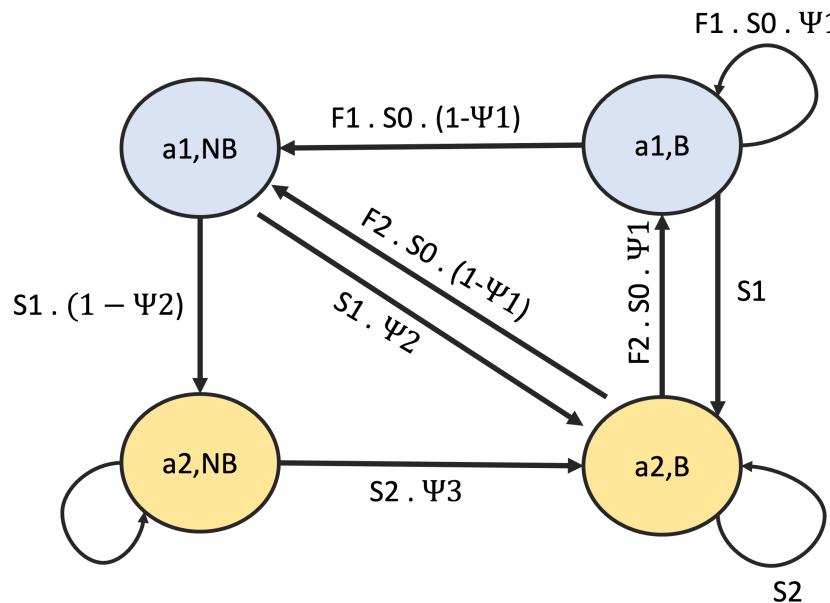


$d$  dispersal probability



$$\begin{bmatrix} F_1 \cdot S_0 & F_2 \cdot S_0 & 0 & 0 \\ S_1 \cdot (1-d) & S_2 \cdot (1-d) & S_1 \cdot d & S_2 \cdot d \\ 0 & 0 & F_1 \cdot S_0 & F_2 \cdot S_0 \\ S_1 \cdot d & S_2 \cdot d & S_1 \cdot (1-d) & S_2 \cdot (1-d) \end{bmatrix}$$

# Variable age at first reproduction: Slender-billed gull example

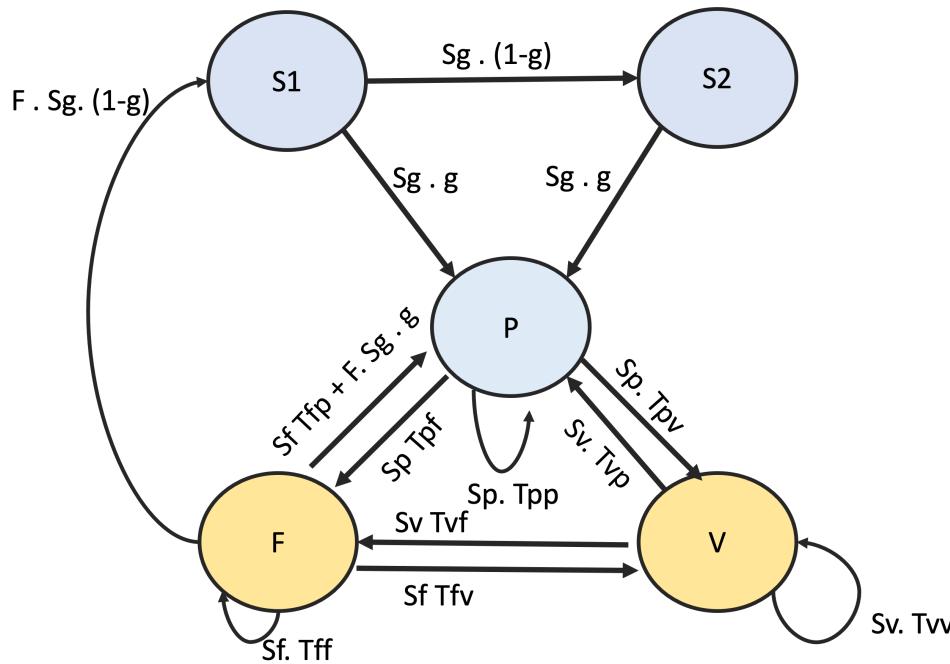


$\Psi_a$  age-specific recruitment probability



$$\begin{bmatrix} F_1 \cdot S_0 \cdot \Psi_1 & F_2 \cdot S_0 \cdot \Psi_1 & 0 & 0 \\ S_1 & S_2 & S_1 \cdot \Psi_2 & S_2 \cdot \Psi_3 \\ F_1 \cdot S_0 \cdot (1-\Psi_1) & F_2 \cdot S_0 \cdot (1-\Psi_1) & 0 & S_1 \cdot (1-\Psi_2) \\ 0 & 0 & S_1 \cdot (1-\Psi_2) & S_2 \cdot (1-\Psi_3) \end{bmatrix}$$

# Transition among stages: Peony example



$$\begin{bmatrix} 0 & 0 & 0 & 0 & F.Sg.(1-g) \\ Sg.(1-g) & 0 & 0 & 0 & 0 \\ Sg.g & Sg.g & Sp.Tpp & Sv.Tvp & Sf.Tfp+F.Sg.g \\ 0 & 0 & Sp.Tpv & Sv.Tvv & Sf.Tfv \\ 0 & 0 & Sp.Tpf & Sv.Tvf & Sf.Tff \end{bmatrix}$$

S seed, P small seedling, V tall vegetative plant, F flowering plant

g germination probability

F #female seeds produced by one female plant (=n/2),

$T_{xy}$  = transition from stage x to stage y

# The projection matrix

- How many age classes / stages in the life cycle?
- Age at first reproduction?
- Maximum age at death fixed or not?
- Pre- or post-breeding census?

Age structured

Leslie matrix

$$\begin{pmatrix} f_1 & f_2 & f_3 & \dots & f_{\max-1} & f_{\max} \\ s_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & s_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & s_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & s_{\max-1} & 0 \end{pmatrix}$$

Stage structured

Lefkovitch matrix

$$\begin{pmatrix} f_1 & f_2 & f_3 & \dots & f_{\max-1} & f_{\max} \\ g_{12} & g_{22} & g_{32} & \dots & \dots & \dots \\ g_{13} & g_{23} & g_{33} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & g_{\max\max} \end{pmatrix}$$

$G$  : transition among stages

# 1- Deterministic MPMs

# Step 5: Project the population

- Write the transition matrix in R:

```
A.swallow <- matrix(c(0.3, 0.6, 0.5, 0.65), # pre-breeding Leslie  
                      nrow = 2,  
                      byrow = TRUE)
```

A.swallow

---

```
      [,1] [,2]  
[1, ] 0.3 0.60  
[2, ] 0.5 0.65
```

# Step 5: Project the population

- Start from an initial population  $n_0$  at time  $t = 0$ :

```
n0 <- c(50,30) # vector with initial population  
n0
```

```
[1] 50 30
```

# Step 5: Project the population

- Project to the next time step:

```
n1 <- A.swallow %*% n0 # matrix product  
n1
```

```
[ , 1 ]  
[1, ] 33.0  
[2, ] 44.5
```

# Step 5: Project the population

- Project the population over 10 years:

```
require(matrixcalc)
t <- 10
# matrix product
n10 <- matrix.power(A.swallow, t) %*% n0
n10
```

```
[ , 1]
[1, ] 53.82434
[2, ] 67.28043
```

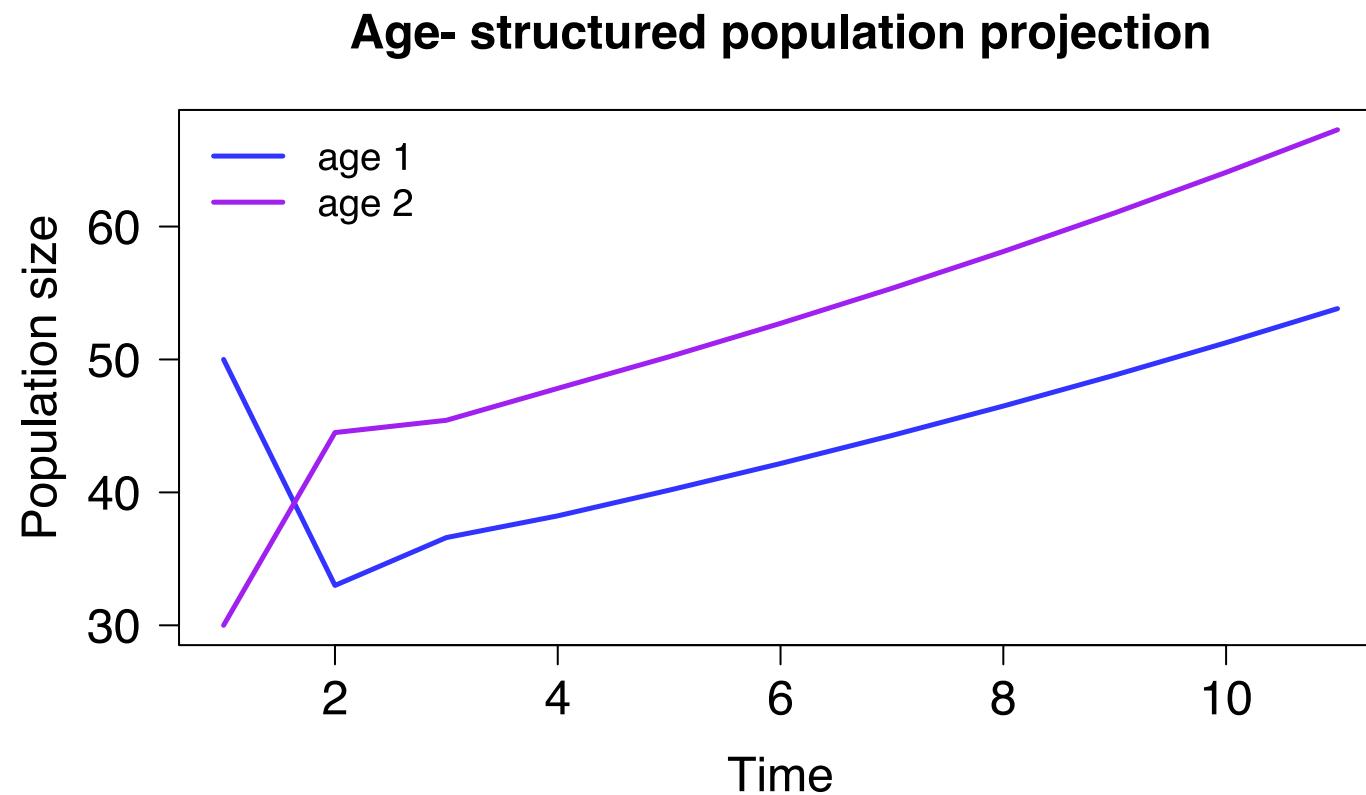
# Step 5: Project the population

- Or using built-in functions from package `popbio`:

```
library(popbio) # load package
t <- 11
# project the population
results <- pop.projection(A.swallow,n0,iterations = t)
results
```

# Step 6: Examine the results

Let's plot the projection



# Convergence to a stable distribution

Transient dynamics:

- Depends on the initial population
- Damping ratio measures how fast the population converges toward equilibrium

Stationary phase:

- Independent of initial conditions
- Depends on the transition matrix only
- Constant growth rate = stable or **asymptotic population growth rate** ( $\lambda$ )
- Constant proportion of individuals per age/stage = **stable age/stage structure**

## Step 6: Examine the results

```
results <- pop.projection(A.swallow, n0, iterations = t)  
names(results)
```

```
[1] "lambda"          "stable.stage"    "stage.vectors" "pop.sizes"  
[5] "pop.changes"
```

## Step 6: Examine the results

First element contains the average growth rate in stationary phase = **asymptotic growth rate**

```
results$lambda
```

```
[1] 1.05
```

## Step 6: Examine the results

Second element contains the **stable age/stage structure** = % of each age/stage in the population in stationary phase

```
results$stable.stage
```

```
[1] 0.4444444 0.5555556
```

## Step 6: Examine the results

Third element contains pop. sizes per age/stage class at each time step

```
results$stage.vectors
```

	0	1	2	3	4	5	6	7	
[1, ]	50	33.0	36.600	38.23500	40.16625	42.17261	44.28144	46.49549	48.820
[2, ]	30	44.5	45.425	47.82625	50.20456	52.71609	55.35177	58.11937	61.025
	9	10							
[1, ]	51.26128	53.82434							
[2, ]	64.07660	67.28043							

## Step 6: Examine the results

Fourth element contains total pop. size at each time step

```
results$pop.sizes
```

```
[1] 80.00000 77.50000 82.02500 86.06125 90.37081 94.88870 99.63  
[8] 104.61486 109.84560 115.33788 121.10477
```

## Step 6: Examine the results

Fifth element is  $\lambda_t$  the rate of change at each time step

```
results$pop.changes
```

```
[1] 0.968750 1.058387 1.049208 1.050076 1.049993 1.050001 1.050000 1.0  
[9] 1.050000 1.050000
```

# Quantities in stationary phase

Calculated directly from the transition matrix:

```
lambda(A.swallow) #stable population growth rate
```

```
[1] 1.05
```

# Quantities in stationary phase

Calculated directly from the transition matrix:

```
stable.stage(A.swallow) # stable age/stage structure
```

```
[1] 0.4444444 0.5555556
```

# Quantities in stationary phase

Calculated directly from the transition matrix:

```
# relative contribution of each age/stage to the next generation  
reproductive.value(A.swallow)
```

```
[1] 1.0 1.5
```

# Quantities in stationary phase

Calculated directly from the transition matrix:

```
generation.time(A.swallow) # average time between generations
```

```
[1] 4.150924
```

```
# average age of mothers at birth of their daughters
```

# Step 7: Sensitivity analysis

Let's perturb the model: What happens if female adult survival is reduced by 50%?

```
A.swallow.modified <- matrix(c(0.3, 0.6, 0.5, 0.65/2), # pre-breeding survival rates
                                nrow = 2,
                                byrow = TRUE)
```

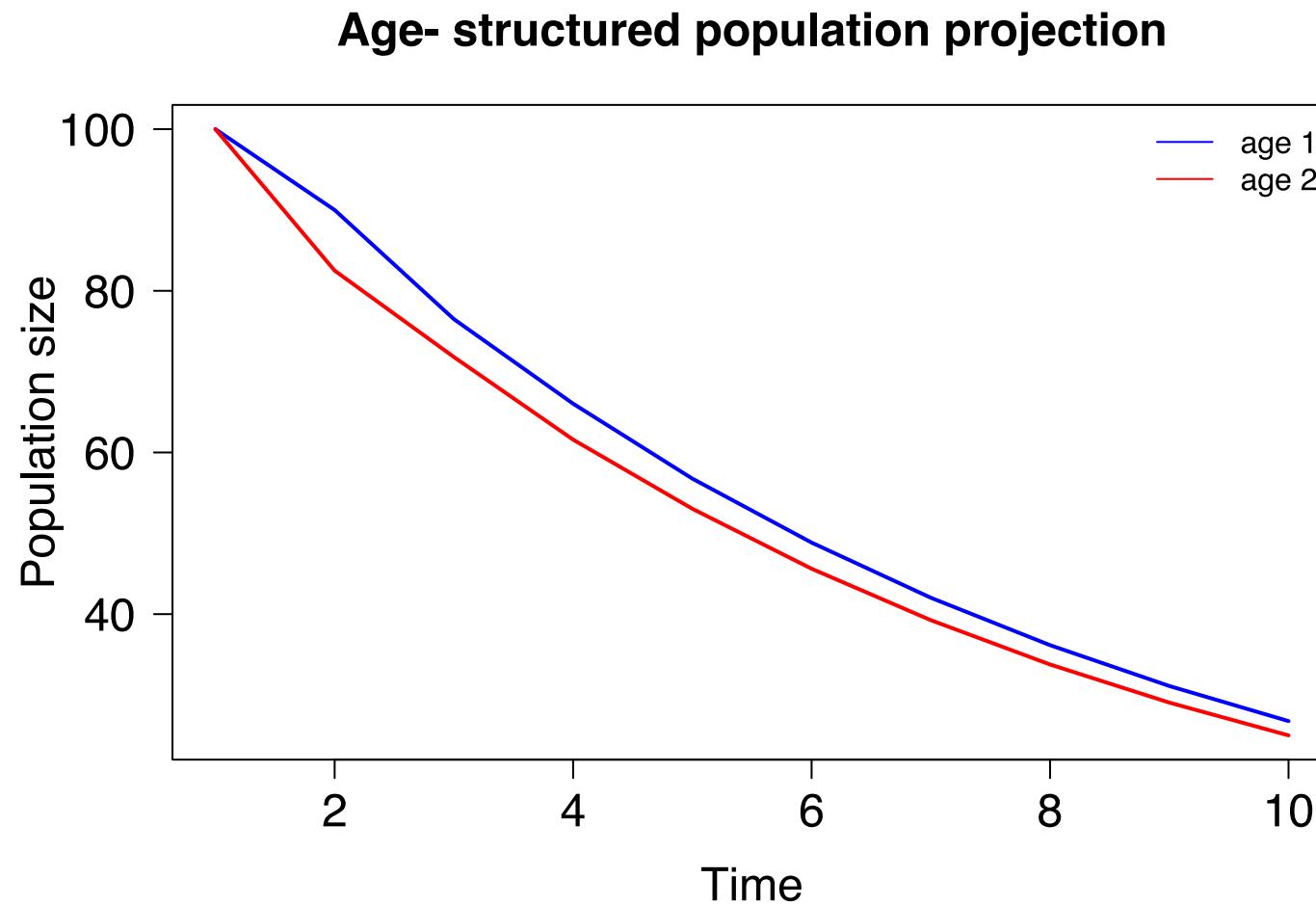
A.swallow.modified

```
      [,1]  [,2]
[1, ]  0.3  0.600
[2, ]  0.5  0.325
```

# Step 7: Sensitivity analysis

```
n0 <- c(100,100)
t <- 10
results.modified <- pop.projection(A.swallow.modified,
                                      n0,
                                      iterations = t)
```

# Step 7: Sensitivity analysis



# Step 7: Sensitivity analysis

```
lambda(A.swallow.modified)
```

```
[1] 0.8603652
```

- The population is now declining
- Adult survival is important to population growth

# Step 7: Sensitivity analysis

- What happens if juvenile survival is reduced ?
- What happens if fecundity or chick survival is reduced ?
- Which demographic parameter contributes most to population dynamics ?

# Step 7: Sensitivity analysis

- Measuring the **impact of a change in a specific demographic parameter on population dynamics**
- **Sensitivity** measures **absolute change** (e.g. -0.1 in parameter)

$$\frac{\delta \lambda}{\delta \theta}$$

- **Elasticity** measures **relative change** (e.g. -0.1% change in parameter)

$$\frac{\delta \lambda}{\delta \theta} \cdot \frac{\theta}{\lambda}$$

- Both are useful, elasticity is better to compare parameters that are on different scales (change in survival versus fertility)

# Step 7: Sensitivity analysis

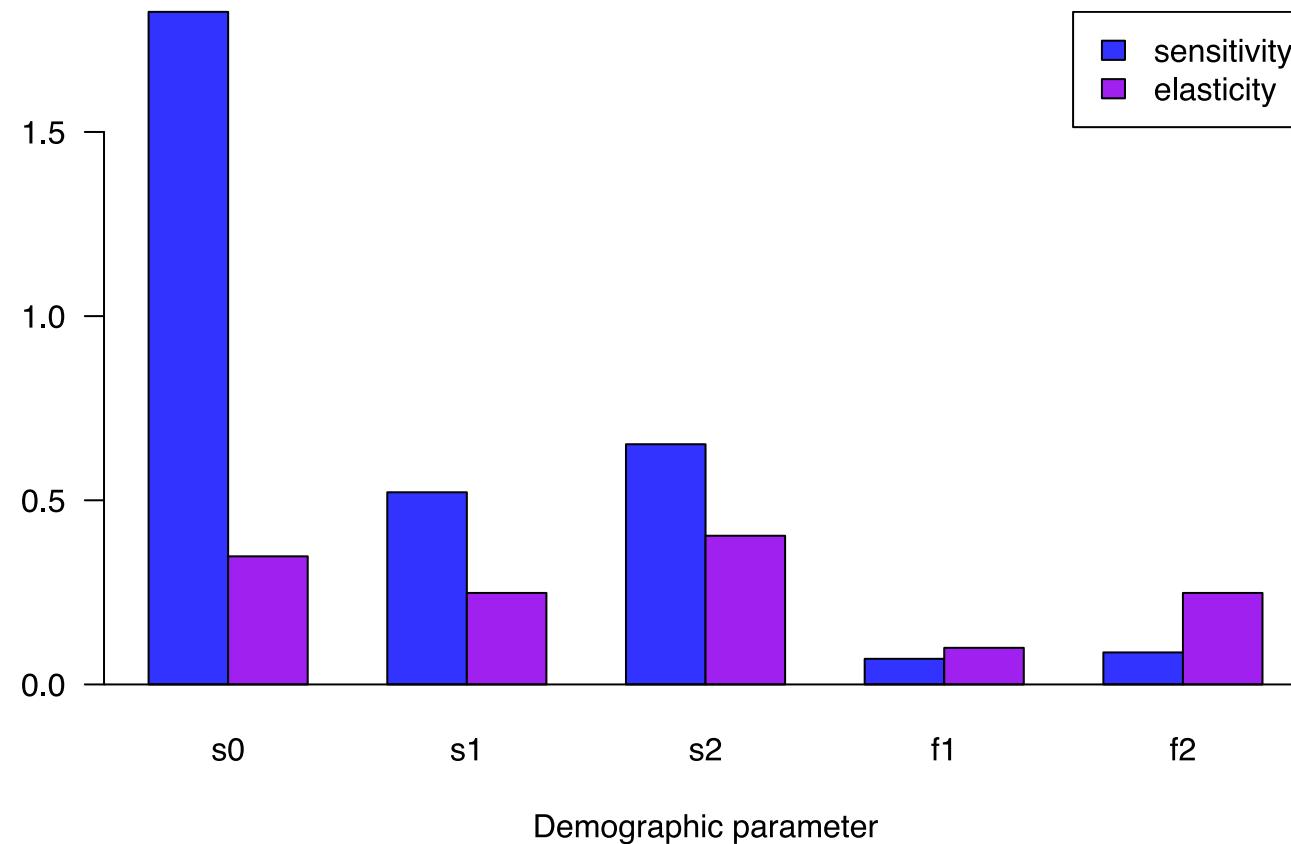
```
swallow.param <- list(s0 = 0.20,  
                      s1 = 0.5,  
                      s2 = 0.65,  
                      f1 = 3/2,  
                      f2 = 6/2)  
swallow.equation <- expression( s0 * f1, s0 * f2, s1, s2)  
VS <- vitalsens(swallow.equation, swallow.param)
```

# Step 7: Sensitivity analysis

```
VS <- vitalsens(swallow.equation, swallow.param)  
VS
```

	estimate	sensitivity	elasticity
s0	0.20	1.82608696	0.34782609
s1	0.50	0.52173913	0.24844720
s2	0.65	0.65217391	0.40372671
f1	1.50	0.06956522	0.09937888
f2	3.00	0.08695652	0.24844720

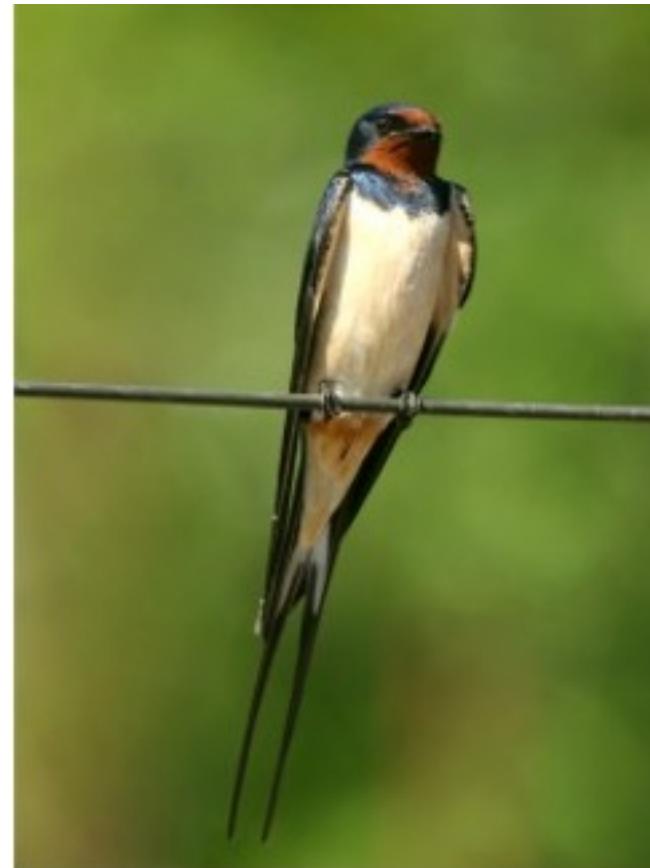
# Step 7: Sensitivity analysis



# Implications for management

- Identify key parameters for population management
  - Best strategy here is to reduce adult mortality
- Evaluate the impact of relative management actions
  - Actions focused on fecundity (e.g. nest protection) will have little impact

# Live demo on the barn swallow



# Assumptions and limitations of deterministic MPMs

- One sex model: Females drive the demography, males are not limiting
- Synchronous breeding: discrete time
- No density-dependence (exponential growth or decay)
- Demographic parameters are constant in time
- No environmental stochasticity
- No demographic stochasticity

# Interest of deterministic models

- Species living in stable environments (such as protected areas)
- Exponential growth (recolonization, recovering from over-exploitation)
- Sensitivity analyses provide useful information to identify key parameters for population management

## 2- Stochastic MPMs

# Environmental stochasticity

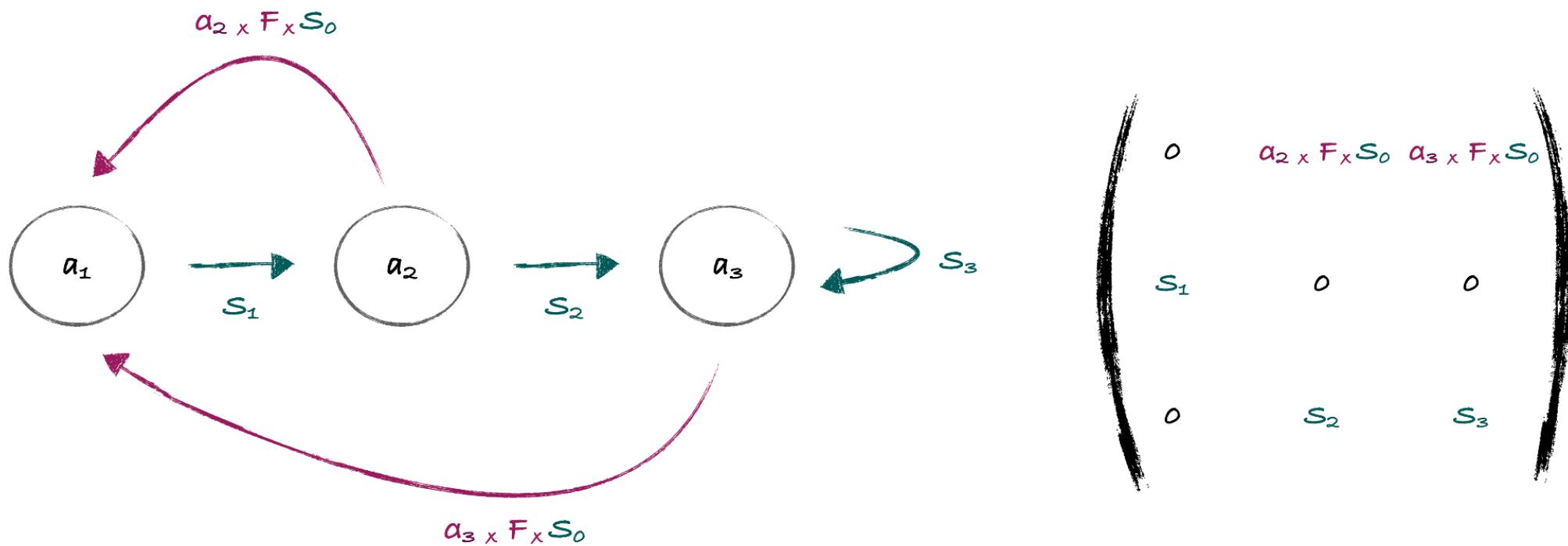
- Unpredictable fluctuations in environmental conditions
- Demographic rates vary over the years
- Stochastic fluctuations in survival and reproductive rates reduce long-run population growth rate
- The chance of occurrence of consecutive periods of unfavourable environmental conditions can drive small populations to extinctions
- **MPMs in stochastic environments** project the population using a series of annual stochastic transition matrices (instead of one transition matrix)

# The crested newt example

High inter-annual variation in demographic parameters:

- Mean adult survival = 0.52 varied between 0.22 et 0.74 in 8 years of monitoring
- Mean adult fecundity = 3.07 juvenile females / adult female varied from 0.31 to 5.40 in 10 years
- Bad years with dry pond in spring (about 1 out of 3 years) induces quasi-complete failure of reproduction

# Crested newt (*Triturus cristatus*) life cycle and transition matrix



$a_1$  immatures,  $a_2$  subadults,  $a_3$  adults

# Several options exist to include annual variability on demographic parameters

- Random annual variation around mean values
- Catastrophic events

Several options exist to include annual variability on demographic parameters

- Random annual variation around mean values
- Catastrophic events

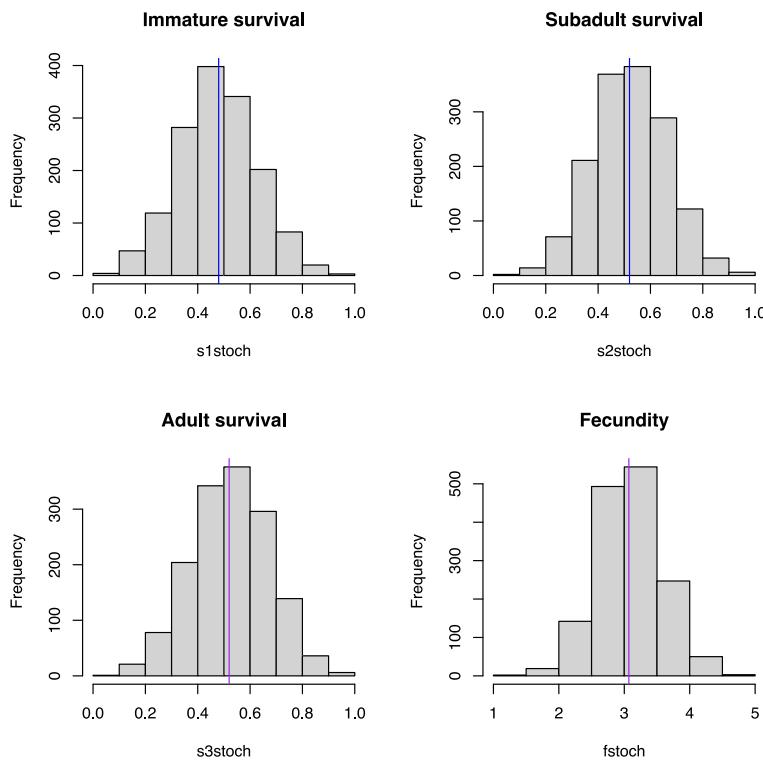
# Step 1: Start from the deterministic transition matrix

```
[ ,1]    [ ,2]    [ ,3]  
[1, ] 0.00 1.03152 1.2894  
[2, ] 0.48 0.00000 0.0000  
[3, ] 0.00 0.52000 0.5200
```

## Step 2: Draw annual values for each demographic parameters from a normal distribution

```
s0stoch <- rnorm(n=1500,mean=s0,sd=0.15)
s1stoch <- rnorm(n=1500,mean=s1,sd=0.15)
s2stoch <- rnorm(n=1500,mean=s2,sd=0.15)
s3stoch <- rnorm(n=1500,mean=s3,sd=0.15)
fstoch <- rnorm(n=1500,mean=f,sd=0.5)
```

# Step 2: Draw annual values for each demographic parameters from a normal distribution



# Step 3: Pile up the stochastic transition matrices

```
# Create a list of stochastic transition matrices
A.newtSE <- list()
# fill in by sampling from distrib. of demo. param.
for(i in 1:ns){
  A.newtSE[[i]] <- matrix( c( 0,
    sample(alpha2,1) * sample(fstoch,1) * sample(s0stoch,1),
    sample(alpha3,1) * sample(fstoch,1) * sample(s0stoch,1),
    sample(s1stoch,1), 0, 0,
    0, sample(s2stoch,1), sample(s3stoch,1) ) ,
    nrow = 3,
    ncol = 3)
}
```

## Step 4: Project the population using the stochastic transition matrices

```
T <- 30
runSE <- stoch.projection(A.newtSE,
                           n0 = c(50, 50, 50),
                           tmax = T,
                           nreps = 1000,
                           verbose = FALSE)
runSE
```

## Step 4: Project the population using the stochastic transition matrices

```
runSE <- stoch.projection(A.newtSE,
    # initial population
    n0 = c(50, 50, 50),
    tmax = T,
    nreps = 1000,
    verbose = FALSE)
runSE
```

## Step 4: Project the population using the stochastic transition matrices

```
runSE <- stoch.projection(A.newtSE,
    n0 = c(50, 50, 50),
    # number of time steps to project over
    tmax = T,
    nreps = 1000,
    verbose = FALSE)
runSE
```

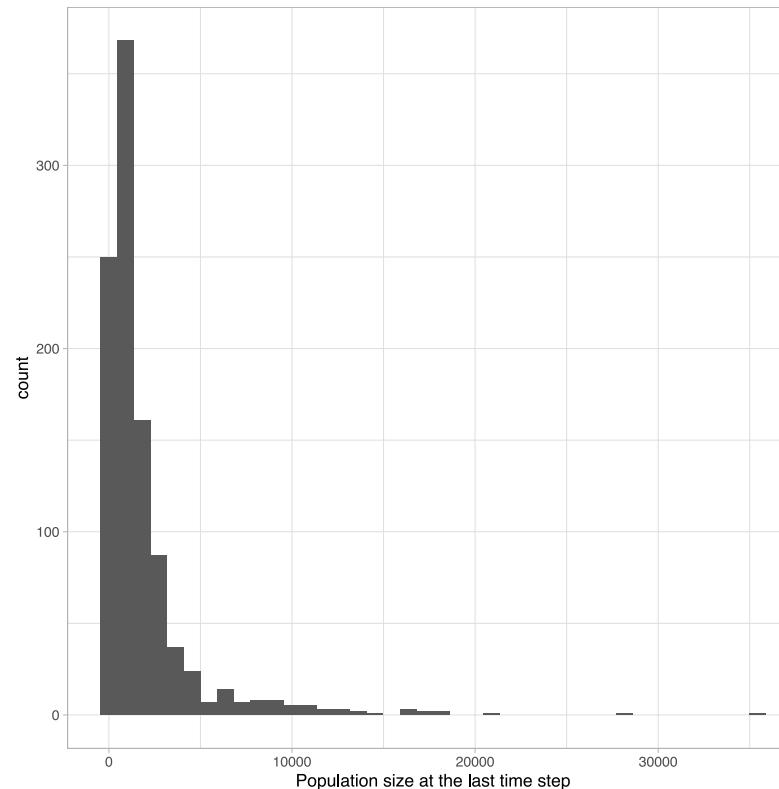
## Step 4: Project the population using the stochastic transition matrices

```
runSE <- stoch.projection(A.newtSE,
                           n0 = c(50, 50, 50),
                           tmax = T,
                           # number of repetitions
                           nreps = 1000,
                           verbose = FALSE)
head(runSE)
```

	[,1]	[,2]	[,3]
[1, ]	59.83276	88.17996	91.30614
[2, ]	92.12902	196.82366	274.23069
[3, ]	616.31127	1282.68469	1301.33291
[4, ]	71.46646	97.31048	223.07366

# Step 5: Examine the results

Distribution of population sizes after  $T = 20$  years?



# Step 5: Examine the results

Long-run stochastic growth rate  $\lambda_s$ :

```
lambdastoch <- stoch.growth.rate(A.newtSE,
                                    maxt = 5000,
                                    verbose = FALSE)
names(lambdastoch)
```

```
[1] "approx" "sim"    "sim.CI"
```

# Step 5: Examine the results

Long-run stochastic growth rate  $\lambda_s$  on a log-scale:

```
lambdastoch$approx # by Tuljapukar's approximation
```

```
[1] 0.0365452
```

```
lambdastoch$sim # by simulation
```

```
[1] 0.0335274
```

```
lambdastoch$sim.CI # with confidence interval
```

```
[1] 0.02813619 0.03891860
```

## Step 5: Examine the results

Long-run stochastic growth rate  $\lambda_s$ :

```
exp(lambdastoch$approx) # exponentiate to get stochastic growth
```

```
[1] 1.037221
```

# Step 5: Examine the results

Probability of extinction:

```
proba.ext <- stoch.quasi.ext(A.newtSE,
  # initial population size
  n0 = c(50, 50, 50),
  Nx = 30,
  nreps = 1000,
  tmax=50,
  maxruns = 10,
  verbose = FALSE)
```

# Step 5: Examine the results

Probability of extinction:

```
proba.ext <- stoch.quasi.ext(A.newtSE,
  # initial population size
  n0 = c(50, 50, 50),
  Nx = 30,
  nreps = 1000,
  tmax = 50,
  maxruns = 10,
  verbose = FALSE)
```

# Step 5: Examine the results

Probability of extinction:

```
proba.ext <- stoch.quasi.ext(A.newtSE,
  n = c(50, 50, 50),
  # quasi-extinction threshold
  Nx = 30,
  nreps = 1000,
  tmax = 50,
  maxruns = 10,
  verbose = FALSE)
```

# Step 5: Examine the results

Probability of extinction:

```
proba.ext <- stoch.quasi.ext(A.newtSE,
  n = c(50, 50, 50),
  Nx = 30,
  # number of runs
  nreps = 1000,
  tmax = 50,
  maxruns = 10,
  verbose = FALSE)
```

# Step 5: Examine the results

Probability of extinction:

```
proba.ext <- stoch.quasi.ext(A.newtSE,
  n = c(50, 50, 50),
  Nx = 30,
  nreps = 1000,
  # number of time steps
  tmax = 50,
  maxruns = 10,
  verbose = FALSE)
```

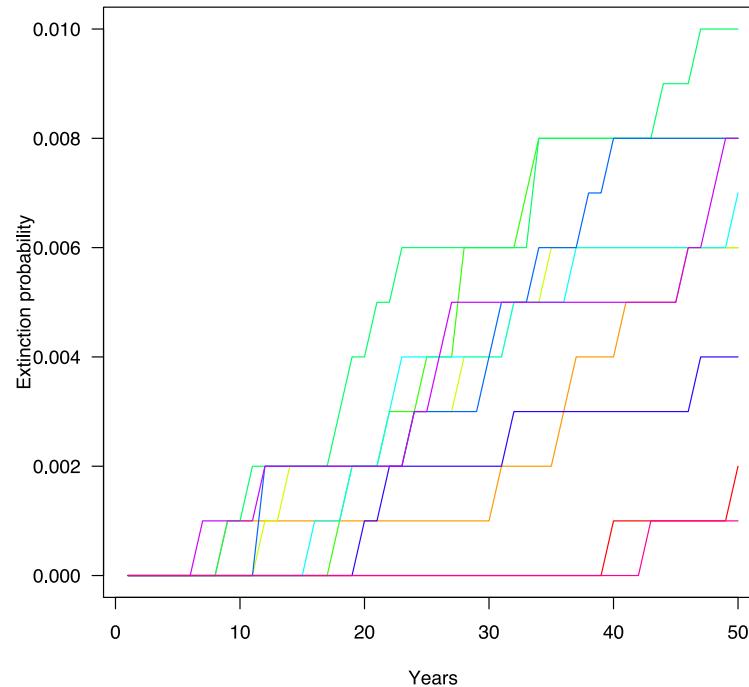
# Step 5: Examine the results

Probability of extinction:

```
proba.ext <- stoch.quasi.ext(A.newtSE,
  n = c(50, 50, 50),
  Nx = 30,
  nreps = 1000,
  tmax = 50,
  # repeat to get robust estimates
  maxruns = 10,
  verbose = FALSE)
```

# Step 5: Examine the results

Probability of extinction:



# Step 5: Examine the results

Probability of extinction:

```
proba.ext.mean <- apply(proba.ext, 1, mean)  
proba.ext.mean[20] # in 20 years
```

[1] 0.0016

# Several options to generate stochastic transition matrices

Include annual variability on demographic parameters:

- Random annual variation around mean values
- **Catastrophic events**

# Step 1: Define transition matrix in good vs bad years

```
fgood <- 3.07 # fecundity in normal years  
fbad <- 0 # fecundity in years with Spring dryness of the pond
```

## Step 1: Define transition matrix in good vs bad years

# Step 1: Define transition matrix in good vs bad years

```
$good
```

```
      [,1]     [,2]     [,3]
[1, ] 0.00 1.03152 1.2894
[2, ] 0.48 0.00000 0.0000
[3, ] 0.00 0.52000 0.5200
```

```
$bad
```

```
      [,1] [,2] [,3]
[1, ] 0.00 0.00 0.00
[2, ] 0.48 0.00 0.00
[3, ] 0.00 0.52 0.52
```

## Step 2: Define frequency of catastrophic events

```
# Spring dyness of the pond occurs every 3 years in average  
freqbad <- 1/3
```

# Step 3: Project the population

Project using the 'good' or 'bad' transition matrix with probability defined above:

```
stochCATA <- stoch.projection(A.newtCATA,
                                prob = c( (1-freqbad), (freqbad)),
                                n0 = c(50, 50, 50),
                                tmax = 100,
                                nreps = 1000,
                                verbose = FALSE)

head(stochCATA)
```

	[,1]	[,2]	[,3]
[1, ]	0.0006161299	0.0000000000	0.0002779838
[2, ]	0.0020843707	0.0000000000	0.0009356701
[3, ]	0.0029163790	0.0000000000	0.0011761417
[4, ]	0.0000000000	0.0000000000	0.0218283887

# Step 3: Project the population

Project using the 'good' or 'bad' transition matrix with probability defined above:

```
stochCATA <- stoch.projection(A.newtCATA,
                                # frequency of castastrophic events
                                prob = c( (1-freqbad), (freqbad) ),
                                n0 = c( 50, 50, 50 ),
                                tmax = 100,
                                nreps = 1000,
                                verbose = FALSE)
```

# Step 3: Project the population

Project using the 'good' or 'bad' transition matrix with probability defined above:

# Step 3: Project the population

Project using the 'good' or 'bad' transition matrix with probability defined above:

```
stochCATA <- stoch.projection(A.newtCATA,
                                prob = c( (1-freqbad), (freqbad)),
                                n0 = c(50, 50, 50),
                                # number of time steps
                                tmax = 100,
                                nreps = 1000,
                                verbose = FALSE)
```

# Step 3: Project the population

Project using the 'good' or 'bad' transition matrix with probability defined above:

```
stochCATA <- stoch.projection(A.newtCATA,
                                prob = c( (1-freqbad), (freqbad)),
                                n0 = c(50, 50, 50),
                                tmax = 100,
                                # number of replicates
                                nreps = 1000,
                                verbose = FALSE)
```

# Step 4: Examine the results

Long-run stochastic growth rate:

```
lambdaCATA <- stoch.growth.rate(A.newtCATA,  
                                 prob = c( (1-freqbad), (freqbad)  
                                 maxt = 5000,  
                                 verbose = FALSE)  
  
exp(lambdaCATA$approx)
```

```
[1] 0.8871908
```

# Step 4: Examine the results

Probability of extinction:

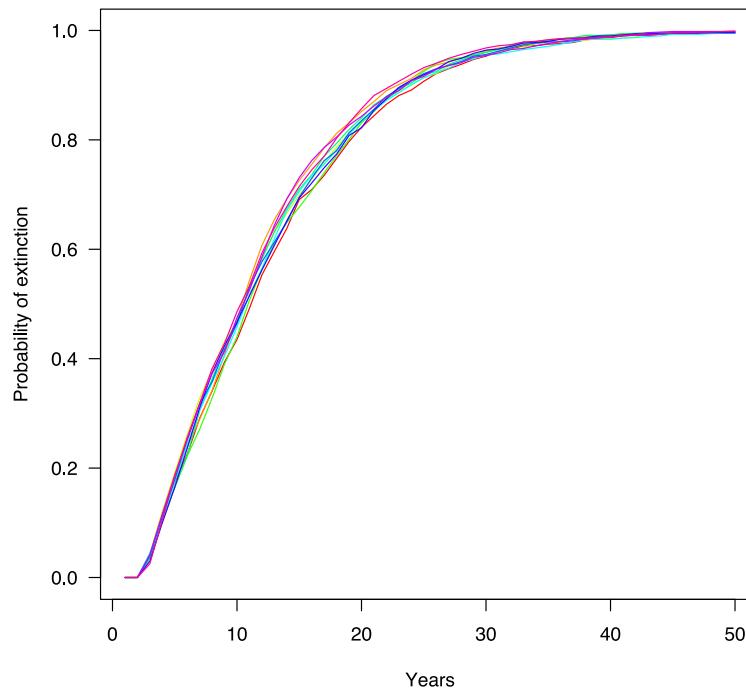
```
proba.extCATA <- stoch.quasi.ext(A.newtCATA,  
                                    prob = c( (1-freqbad), (freqbad),  
                                    n = c(50, 15, 50), # initial pop  
                                    Nx = 30, # Quasi-extinction threshold  
                                    nreps = 1000, # nb. of replicates  
                                    tmax = 50, # nb. of time steps  
                                    maxruns = 10, # nb of repetitions  
                                    verbose = TRUE)
```

Calculating extinction probability for run 1

Calculating extinction probability for run 2

# Step 4: Examine the results

Probability of extinction:



# Step 4: Examine the results

Probability of extinction:

```
proba.ext.mean <- apply(proba.extCATA, 1, mean)  
proba.ext.mean[20] # in 20 years
```

[1] 0.8371

# Step 5: Interesting to evaluate the sensitivity of the results to changes in:

- Amount of environmental variance
- Frequency of catastrophic events
- Initial population size
- Extinction threshold
- Number of time steps

# Live demo on crested newts



# Assumptions and limitations

- One sex model: Females drive the demography, males are not limiting
- Synchronous breeding: discrete time
- No density-dependence (exponential growth or decay)
- No demographic stochasticity
- No trend or change of mean in environmental variance

# Assumptions and limitations

- Keep a critical eyes on results: efficiency of PVAs debated in the litterature
- We are making assumptions about demographic parameters in the future
- Short-term studies = underestimation of variance in demographic rates while extinction risks increase with increased temporal variance in pop size

# In which cases constant MPMs are useful despite its limitations?

- Good knowledge of the species biology, life cycle and estimate of demographic parameters are critical
- Sensitivity analyses remain a powerful tool to identify key demographic parameters and evaluate management actions.
- Especially for big populations (little impact of demographic stochasticity)
- When survival and fertility are best structured by age or stage

# Other (more complicated) models exist

- MPMs including demographic stochasticity for small populations
- Density-dependent MPMs
- Two-sex models for species with skewed reproductive success
- Integral projection models for size-structured demography
- Continuous-time models
- Multi-species models...

# Useful references

Morris, William, et al. (1999). *A practical handbook for population viability analysis*. The Nature Conservancy.

Caswell, H. (2000). *Matrix population models* (Vol. 1). Sunderland, MA: Sinauer.

Brook, Barry W., et al. (2000). Predictive accuracy of population viability analysis in conservation biology. *Nature* 404.6776: 385-387.

Beissinger, Steven R., and Dale R. McCullough, eds. (2002). *Population viability analysis*. University of Chicago Press.

Reed, J. Michael, et al. (2002). Emerging issues in population viability analysis. *Conservation Biology* 16.1: 7-19.

# Quantitative methods for population dynamics

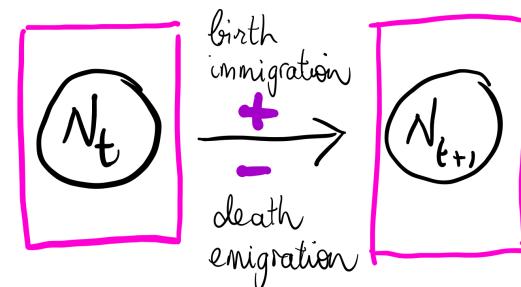
## Conclusions

Olivier Gimenez, Sarah Cubaynes, Aurélien Besnard  
Coline Canonne, Thierry Chambert, Thibaut Couturier, Valentin Lauret

# Population dynamics

# What we covered (1/3)

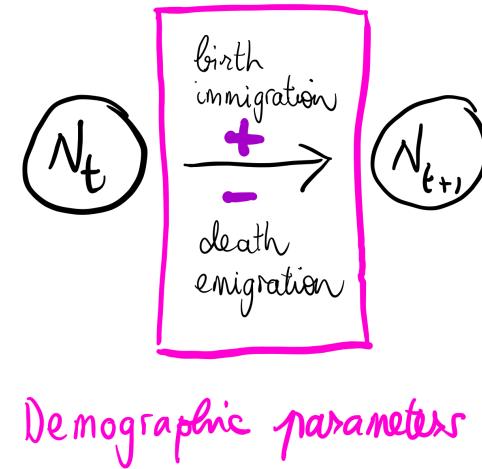
- Abundance and density
  - Counting methods
  - Detectability issue
  - Capture-recapture, distance sampling and N-mixture



Population size and density

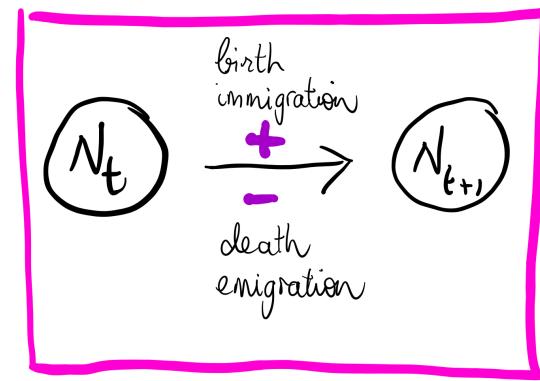
# What we covered (2/3)

- Demographic parameters
  - Capture-recapture data
  - Survival
  - Covariates
  - Transitions



# What we covered (3/3)

- Population projection models
  - Count-based PVA
  - Age- and stage-structured models
  - Sensitivity analyses



Population projection models

# Take-home messages and recommendations

# The art of modelling

- Make your ecological question explicit.
- Think of your model.
- Start simple. Make sure everything runs smoothly.
- Add complexity one step at a time.

# Till next time

- Website will be updated with
  - video recordings
  - your feedbacks
- Please, fill in feedback form at <https://forms.gle/XoXvz58RQq6kAMT38>. It will help us to improve materials and delivery for next time.
- Certificate of attendance on request.
- A book is on its way. More in 2023 hopefully.