Brief introduction to Mark-recapture analysis

Reading: Pine et al. 2003

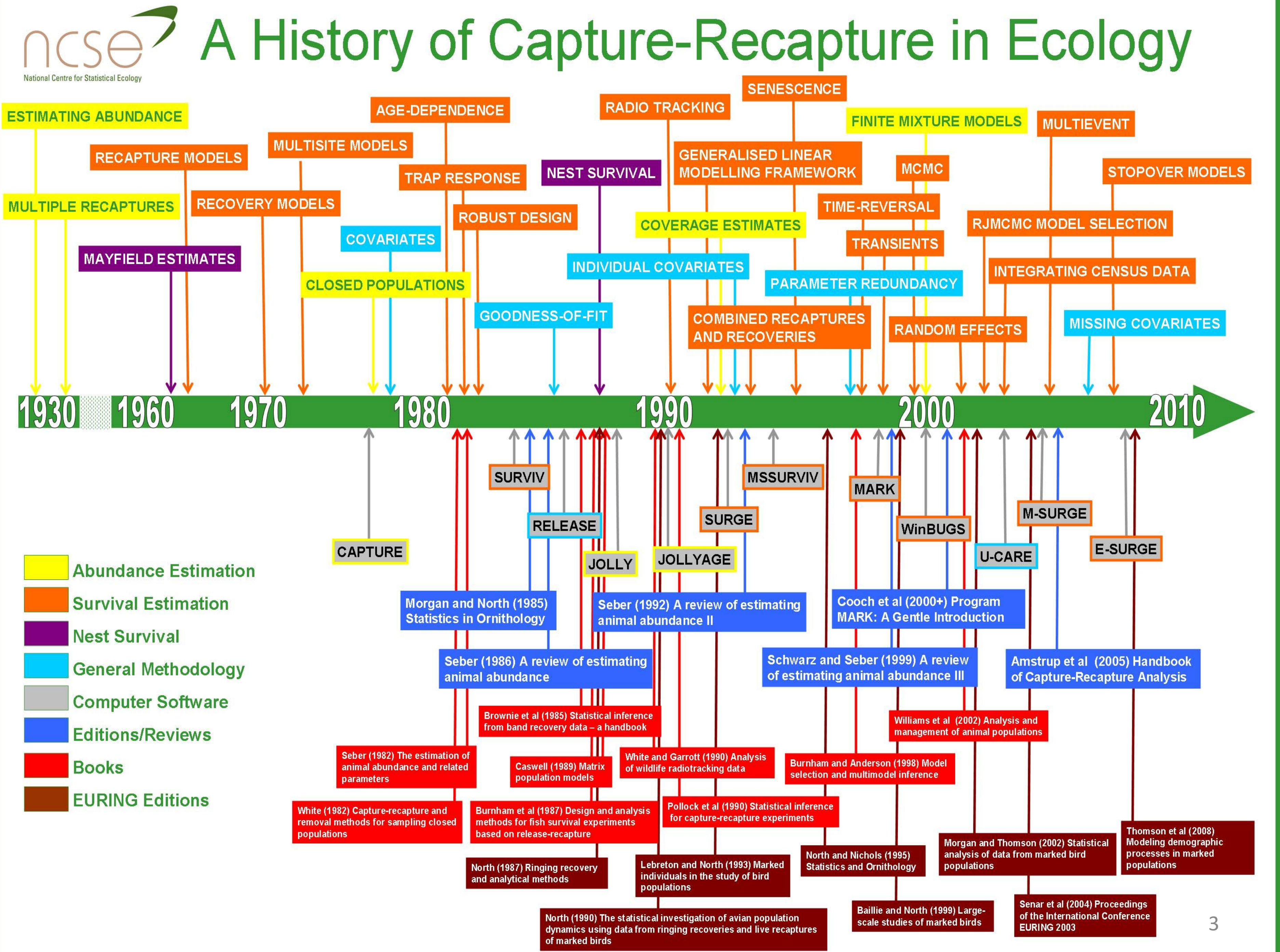
For more: WLDF 578 Advanced Ecology of Wildlife Populations (aka Dan Barton's 'MARK class')

Acknowledgement: Mark Henderson

Outline

- A brief history
- Closed population methods (→ for abundance)
  - Lincoln-Petersen
- Open population methods (→ for survival)
  - Capture-recapture (Cormack-Jolly-Seber) models
  - Recovery (Brownie) models
- Extensions
  - Multi-state models
  - Robust design

A History of Capture-Recapture in Ecology



Mark recapture models

- **Main Purpose is to estimate:**
  1. **Mortality (Z, F, M) or Survival**
  2. **Abundance**
  3. **(or study behavior, movement, habitat use, etc.)**

- Lots of different models & methods!

Table 2. Model, type of mark required (batch or individual), source of fish used in study (research collection or fishery dependent), typical study duration, reporting rate requirement, key parameters, additional information generated, and principal software for estimating population size and mortality components from tagging models discussed in this review.

Model name	Type of mark required	Source of fish	Typical study duration	Reporting rate required?	Key demographic parameters	Additional information generated	Principal software
Lincoln-Peterson	Batch	Research	< 1 month	No	Population size		Calculator, spreadsheet, SPAS
Schnabel	Batch	Research	< 1 month	No	Population size		Calculator, spreadsheet, or CAPTURE
Removal	No mark	Research	< 1 month	No	Population size		CAPTURE or MARK
Closed-CAPTURE models	Unique individual	Research	< 1 month	No	Population size, capture probability		CAPTURE for all closed models or MARK for non-heterogeneity
Jolly-Seber and Cormack-Jolly-Seber	Unique individual	Research	>1 month	No	Population size, apparent survival	Individual growth from recaptures	POPAN, JOLLY, or MARK
Robust	Unique individual	Research	>1 month	No	Population size and growth, apparent survival, temporary emigration	Individual growth from recaptures	CAPTURE and JOLLY together or MARK
Brownie	Unique individual	Fishery	>1 year	No	Survival, total mortality		BROWNIE, MARK
Hoenig/Hearn	Unique individual	Fishery	>1 year	Yes	Survival, fishing and natural mortality		AVOCADO
Telemetry	Unique individual	Research	= 1 year	No	Survival, fishing and natural mortality	Movement, habitat use	SURVIV
Combined telemetry/ tagging	Unique individual	Research/Fishery	> 1 year	No	Survival, fishing and natural mortality	Movement, habitat use	SURVIV

# of tagging events also important

Pine et al. 2003

Many Software Options

Product name	Description	World Wide Web address
MARK	Comprehensive program for most types of capture-recapture analysis including open, closed, and robust design models. Capture probability and survival directly estimated for open, closed, and robust models and population size estimation for closed and robust models.	www.cnr.colostate.edu/~gwhite/mark/mark.htm
CAPTURE	One of the first programs for estimating population size and capture probability in closed populations. Calculates estimates using a variety of models which are able to account for heterogeneity, behavioral response, time variation, in capture probability. Only software that contains heterogeneity models. Can be run as an option within MARK.	www.mbr-pwrc.usgs.gov/software
JOLLY	Program for estimating population size, survival, and capture probability of open populations.	www.mbr-pwrc.usgs.gov/software
SURVIV	Program used to calculate survival rates from user-specified survival functions including tag-return models. Not very user-friendly.	www.mbr-pwrc.usgs.gov/software
POPAN	Program for estimating population size and number of new recruits in open populations.	www.cs.umanitoba.ca/~popan/
SPAS	Program for estimating population size in stratified two sample capture-recapture studies.	www.cs.umanitoba.ca/~popan/

Package "RMark" is an R interface that uses program MARK in the background

Pine et al. 2003

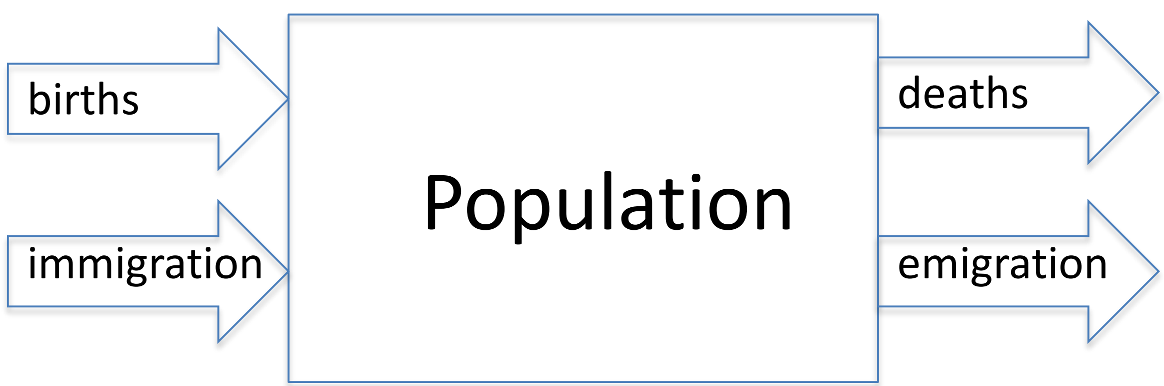
Open vs closed population

- Closed population*
  - No changes in population size (births, deaths, immigration, or emigration)

***Note this is slightly different definition than what we used before. Before, "closed" was defined by not having migration (but allowed births and deaths)**

Open vs closed population

- Closed population
  - No changes in population size (births, deaths, immigration, or emigration)
- Open population



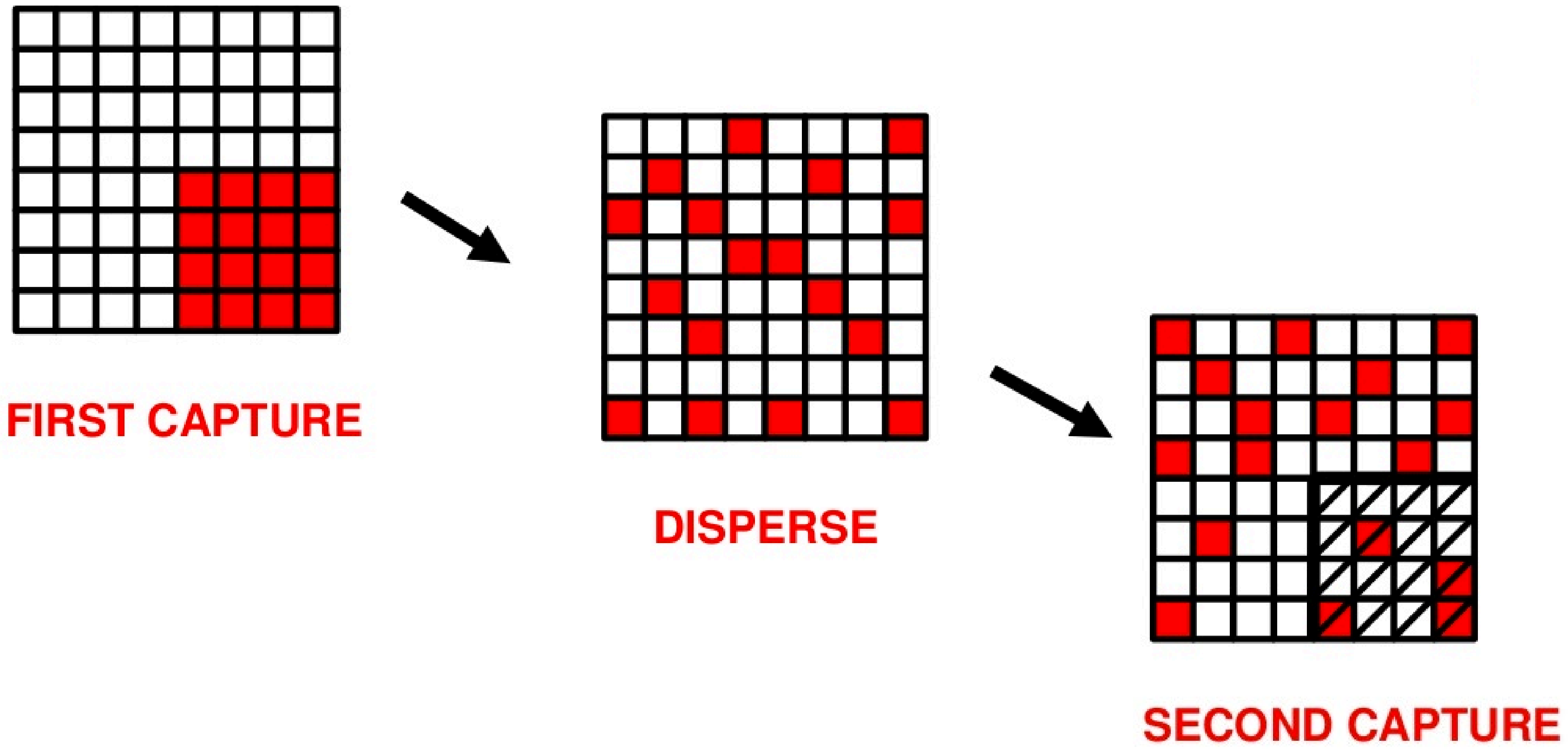
Outline

- A brief history
- Closed population methods
  - Lincoln-Petersen
- Open population methods
  - Capture-recapture (Cormack-Jolly-Seber) models
  - Recovery (Brownie) models
- Extensions
  - Multi-state models
  - Robust design

Closed Populations

- No additions
  - Births or immigration
- No deletions
  - Death or emigration
- **Primarily used to estimate abundance**

Closed Populations



Lincoln-Petersen Estimator

$$\underset{N}{\frac{\overset{n_1}{\text{\# captured in 1st sample}}}{\text{Population abundance}}} = \underset{n_2}{\frac{\overset{m_2}{\text{\# marked in 2nd sample}}}{\text{\# captured in 2nd sample}}}$$

$$\frac{n_1}{N} = \frac{m_2}{\mathrm{n}_2}$$

Solve for N →

$$N = \frac{n_1 n_2}{m_2}$$

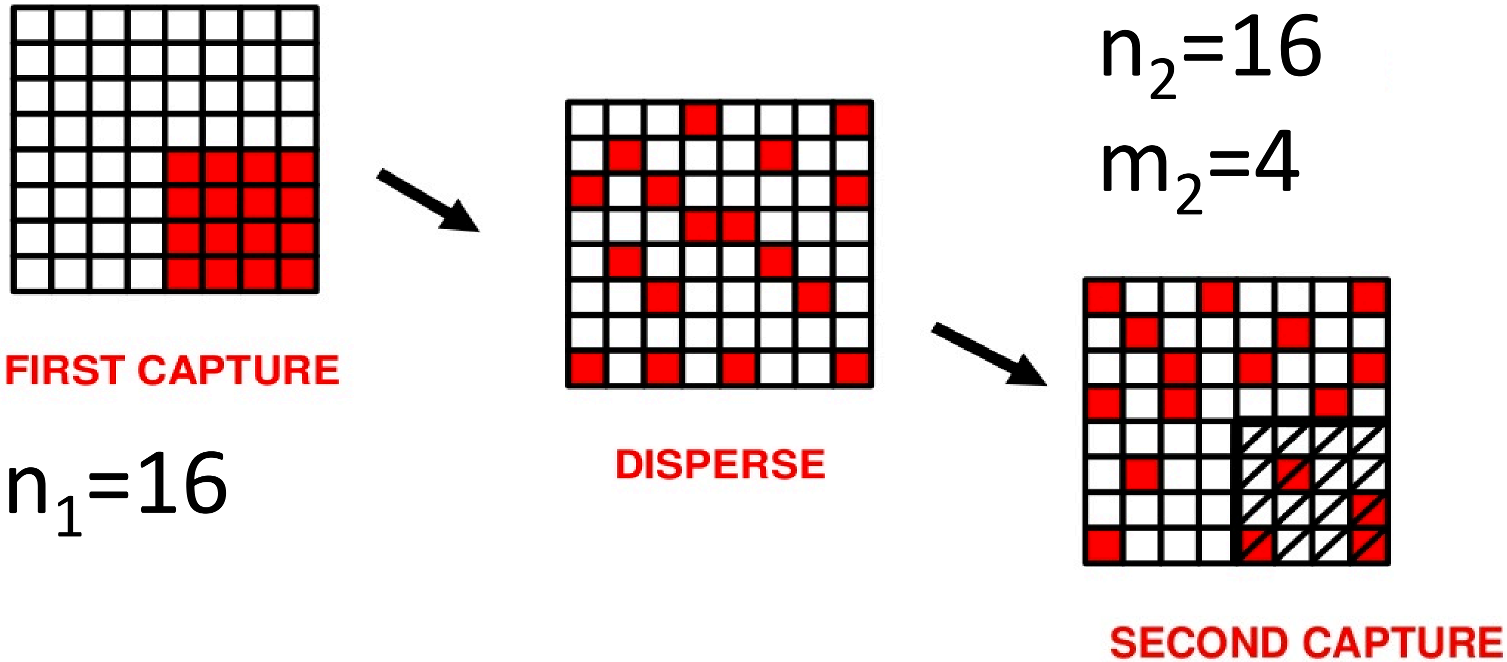
Lincoln-Petersen Correction

- Chapman correction
  - Reduces bias
  - Defined even if m_2=0

$$N_{chap} = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1$$

https://derekogle.com/NCNRS349/modules/MarkRecap/BKG

Lincoln-Petersen Example



n_1=16

n_2=16

m_2=4

$$N = \frac{n_1 n_2}{m_2} = \frac{16 * 16}{4} = 64$$

As an exercise, calculate N assuming m_2 was 2 or 5.

$$N_{chap} = \frac{(n_1 + 1)(n_2 + 1)}{m_2 + 1} - 1 = \frac{17 * 17}{5} - 1 = 56.8$$

Assumptions

- Population is closed to additions or deletions
  - i.e., no births, deaths, migration
- All individuals have the same capture probability

- Individuals do not lose marks

Outline

- A brief history
- Closed population methods
  - Lincoln-Petersen
- Open population methods
  - Capture-recapture (Cormack-Jolly-Seber) models
  - Recovery (Brownie) models
- Extensions
  - Multi-state models
  - Robust design

Capture Recapture Models

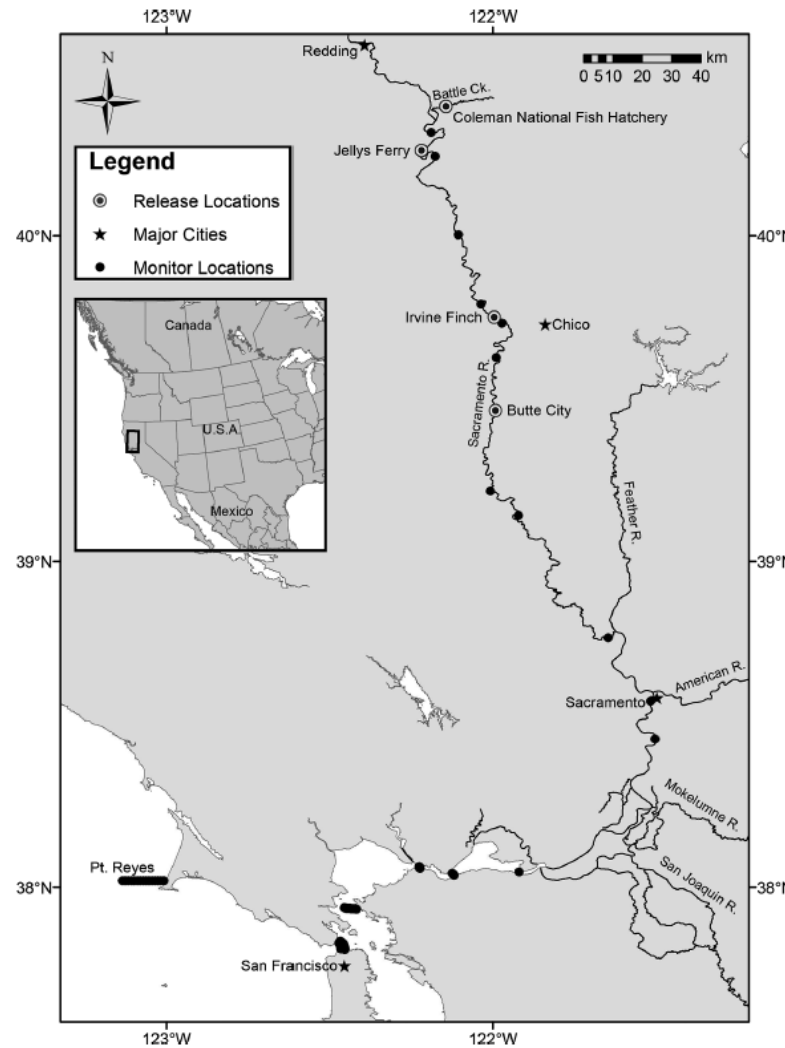
- Individuals are captured and then released
  - Can be recaptured in multiple sampling occasions

Cormack Jolly Seber (CJS) Model

- Individuals are captured, tagged, and released
  - Can be recaptured on multiple sampling occasions
  - Model accounts for imperfect re-capture (i.e, imperfect detection)
- Only uses marked individuals
  - **Primarily used to estimate survival**
  - **Cannot estimate abundance or recruitment**
- Requires at least 3 samples
- Developed independently in 1960s by three researchers (Cormack, Jolly, and Seber)

CJS Example

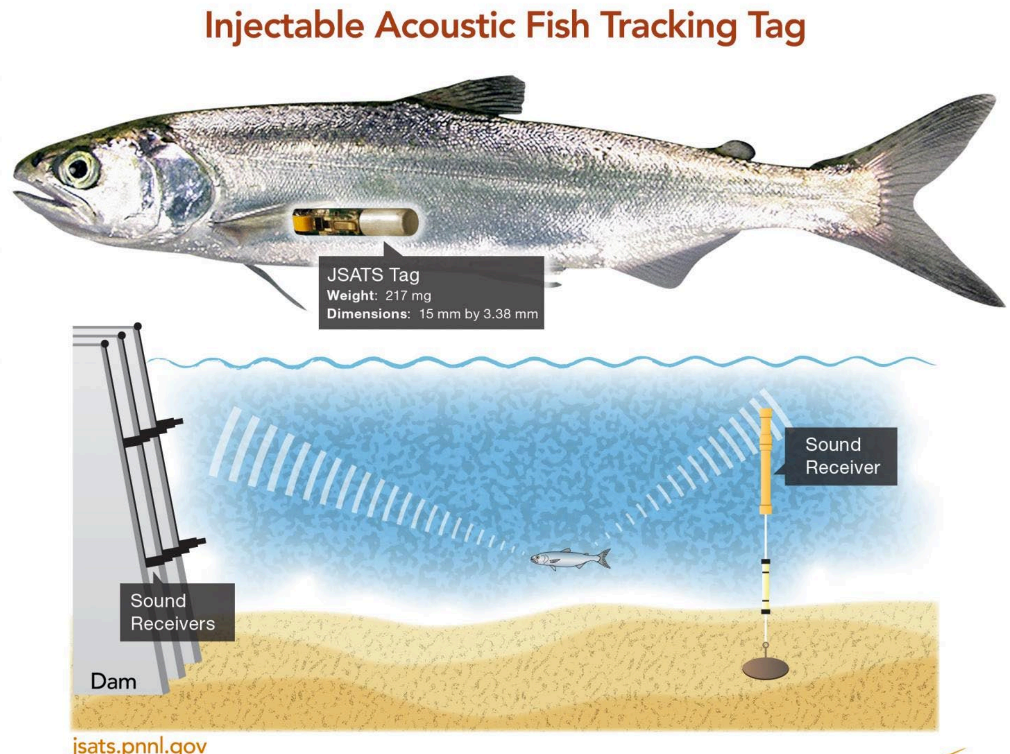
- Implanted acoustic tags in 1350 late-fall Chinook smolts
- Tracked movements with acoustic receivers
- Applied CJS model to estimate survival throughout Sac River, Delta, and SF estuary



Michel et al. 2015

Capture Recapture Example

- Juvenile Salmon Acoustic Telemetry System



Pacific Northwest National Laboratory

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

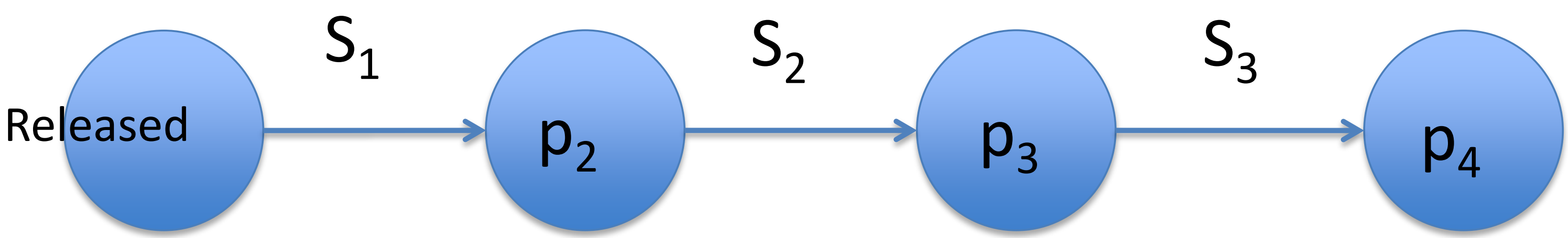
Parameters

- **Capture probability** (p_i)
  - Probability that a marked fish is captured in period (i)
- **Apparent survival** (S_i) (often use "Phi" Φ_i)
  - Probability that an animal alive in time (i) survives until (i+1) *and* does not permanently emigrate
  - Cannot distinguish between death and permanent emigration

Sidenote: Φ is also used in math for the [Golden ratio (1.618…)](), but this is not what we are using

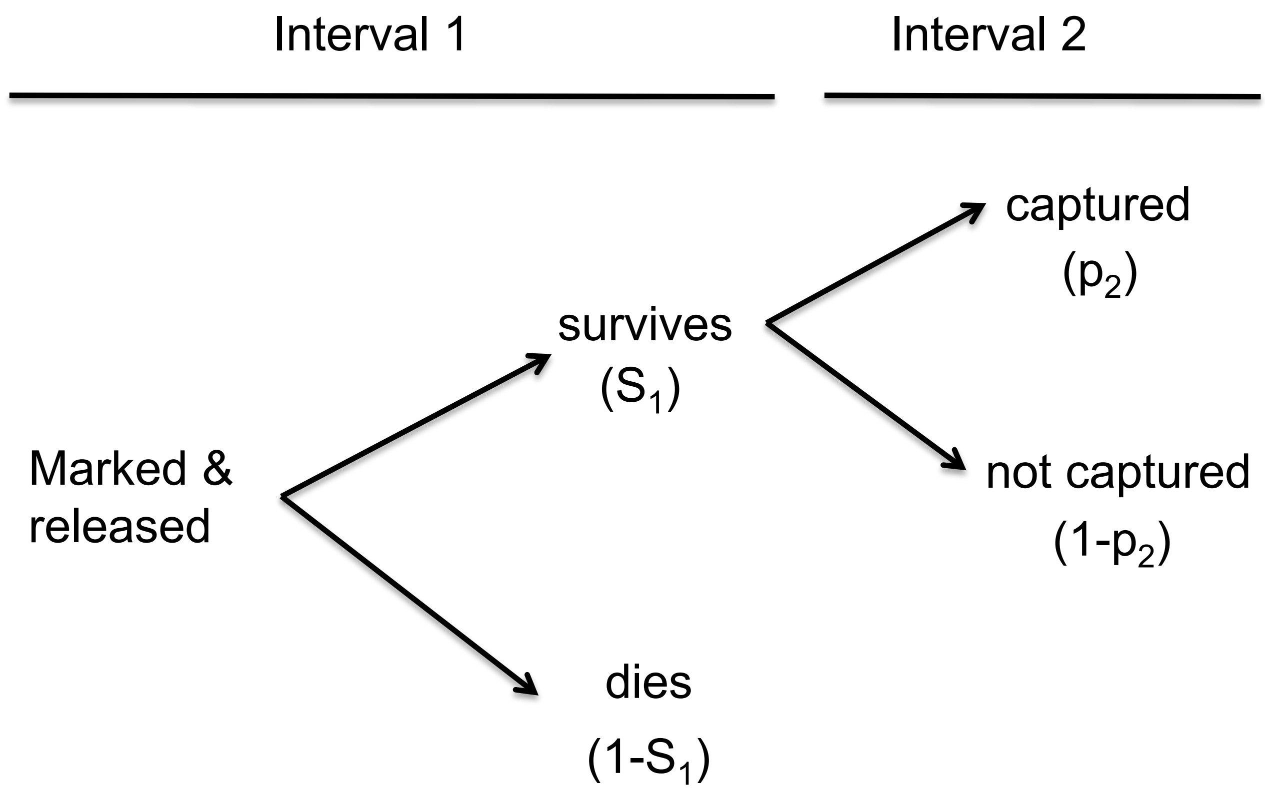
CJS data

For each fish, we may (or may not) detect it on each sampling occasion



$$\text{E.g.:}\quad 1 \xrightarrow{\quad S_1 \quad} \underset{1-p_2}{0} \xrightarrow{\quad S_2 \quad} \underset{p_3}{1} \xrightarrow{\quad S_3 \quad} \underset{p_4}{1}$$

CJS Model



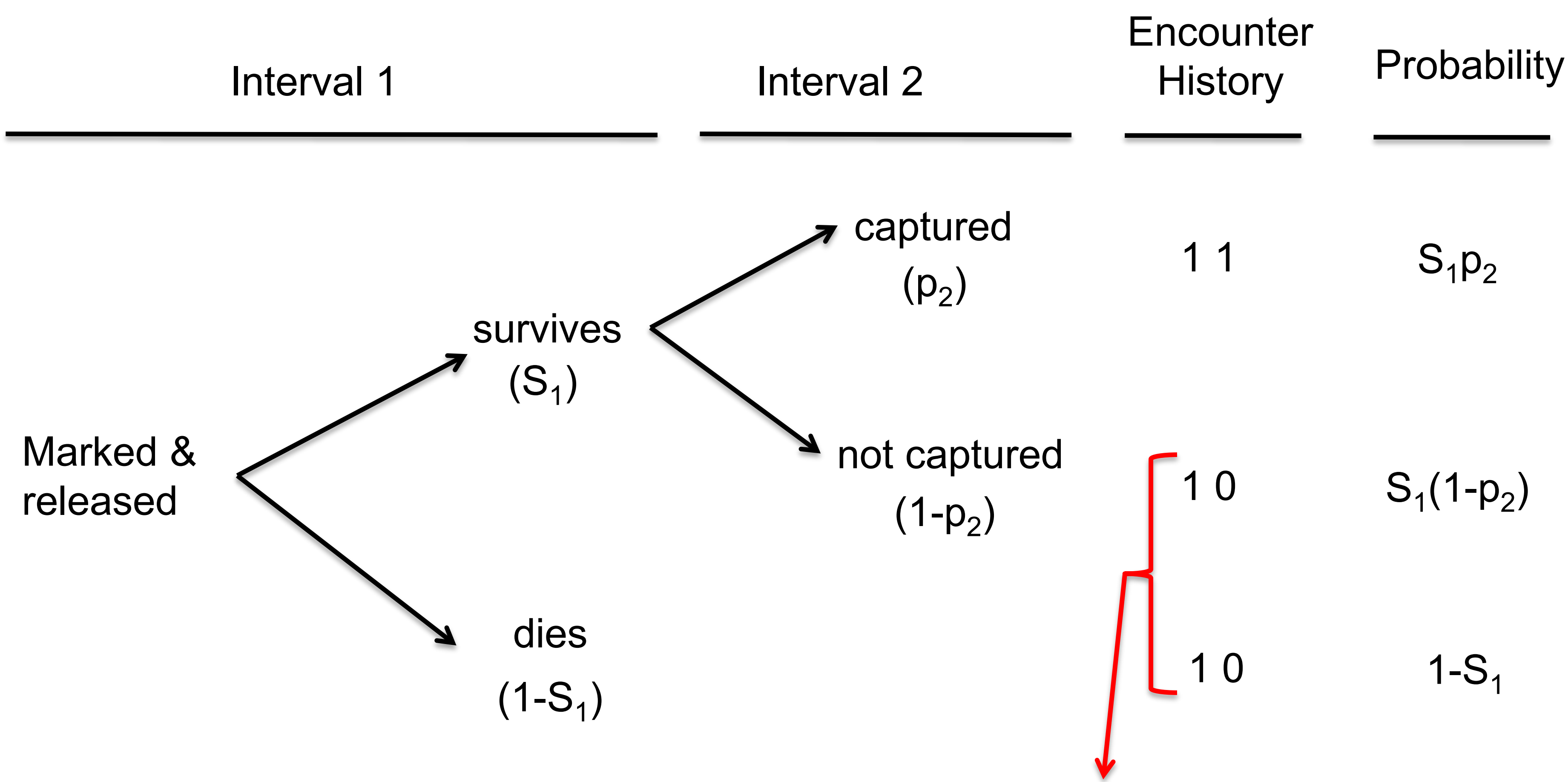
Modified from Williams et al 2002

CJS Model

Interval 1		Interval 2	Encounter History	Probability
Marked & released	survives (S_1)	captured (p_2)	1 1	S_1p_2
		not captured ($1-p_2$)	1 0	$S_1(1-p_2)$
	dies ($1-S_1$)		1 0	$1-S_1$

Modified from Williams et al 2002

CJS Model



$$Pr(10) = (1-S_1)+S_1(1-p_2)$$

Modified from Williams et al 2002

CJS Likelihood

- Capture history, *CH* (conditional on release)

Interval 1	Interval 2	Interval 3	Interval 4
1	0	1	0

$$\boxed{\Pr(CH) = S_1(1-p_2)}\ \boxed{S_2 p_3}\ \boxed{((1-S_3)+S_3(1-p_4))}$$

$$L(S,p) = \prod_{i=1}^{h} (\Pr(CH_h))^n$$

Where h = unique capture histories
n = individuals with capture history i

Example

Ind	Sample2	Sample3	Sample4	Probability	Likelihood
1	1	1	1	$S_1 p_2 S_2 p_3 S_3 p_4$	0.106
2	1	0	1		
3	0	1	1		
4	1	1	0		
5	1	0	0		

Initial parameter values: $S_1=0.5$, $S_2=0.6$, $S_3=0.7$; $p_2=0.9$, $p_3=0.8$, $p_4=0.7$

Example

Ind	Sample2	Sample3	Sample4	Probability	Likelihood
1	1	1	1	$S_1p_2S_2p_3S_3p_4$	0.106
2	1	0	1	$S_1p_2S_2(1\text{-}p_3)S_3p_4$	0.027
3	0	1	1		
4	1	1	0		
5	1	0	0		

Initial parameter values: S_1=0.5, S_2=0.6, S_3=0.7; p_2=0.9, p_3=0.8, p_4=0.7

Example

Ind	Sample2	Sample3	Sample4	Probability	Likelihood
1	1	1	1	$S_1p_2S_2p_3S_3p_4$	0.106
2	1	0	1	$S_1p_2S_2(1-p_3)S_3p_4$	0.027
3	0	1	1	$S_1(1-p_2)S_2p_2S_3p_4$	0.012
4	1	1	0		
5	1	0	0		

Initial parameter values: S_1=0.5, S_2=0.6, S_3=0.7; p_2=0.9, p_3=0.8, p_4=0.7

Example

Ind	Sample2	Sample3	Sample4	Probability	Likelihood
1	1	1	1	$S_1p_2S_2p_3S_3p_4$	0.106
2	1	0	1	$S_1p_2S_2(1\text{-}p_3)S_3p_4$	0.027
3	0	1	1	$S_1(1\text{-}p_2)S_2p_2S_3p_4$	0.012
4	1	1	0	$S_1p_2S_2p_3[1\text{-}S_3+S_3(1\text{-}p_4)]$	0.065
5	1	0	0		

Initial parameter values: S_1=0.5, S_2=0.6, S_3=0.7; p_2=0.9, p_3=0.8, p_4=0.7

Example

Ind	Sample2	Sample3	Sample4	Probability	Likelihood
1	1	1	1	$S_1p_2S_2p_3S_3p_4$	0.106
2	1	0	1	$S_1p_2S_2(1\text{-}p_3)S_3p_4$	0.027
3	0	1	1	$S_1(1\text{-}p_2)S_2p_2S_3p_4$	0.012
4	1	1	0	$S_1p_2S_2p_3[1\text{-}S_3+S_3(1\text{-}p_4)]$	0.065
5	1	0	0	$S_1p_2\,\{1\text{-}S_2+S_2(1\text{-}p_3)[1\text{-}S_3+S_3(1\text{-}p_4)]\}$	0.208

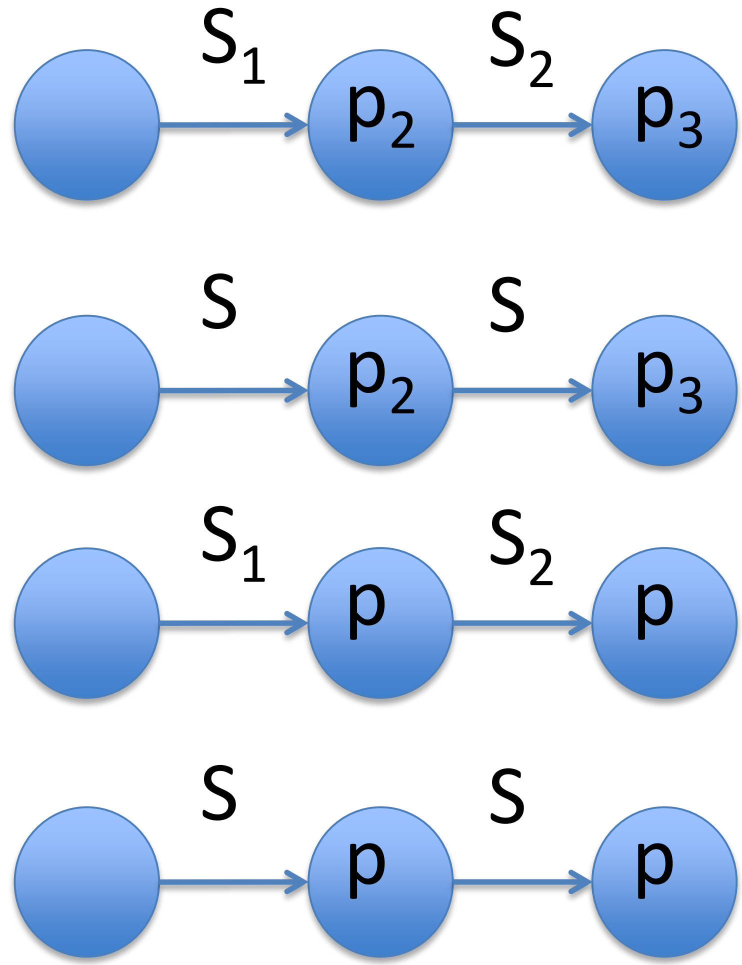
Initial parameter values: S_1=0.5, S_2=0.6, S_3=0.7; p_2=0.9, p_3=0.8, p_4=0.7

Model assumptions

- Survival is equal for marked and unmarked animals
- Tagging does not influence capture probability
- Sampling is ‘instantaneous’
- Tags are not lost or overlooked
- Fate of each fish is independent

Examples of Potential Models

- Time dependent
- Constant Survival
- Constant Cap Prob
- Time independent



Note the different subscripts on the parameters;
different subscripts denote different values

Model Selection

- Akaike's Information Criterion (AIC)

$$\mathrm{AIC} = -2\log(\hat{\mathrm{L}}) + 2\mathrm{p}$$

  - $\hat{L}$ = likelihood value for a model evaluated at the parameter estimates
  - p = number of parameters (including the estimated error term, σ^2)
  - Lower values are better

- Other criteria also exist (QAIC, BIC, etc.)

Adding covariates

- Use Logistic regression and Logit Link
  - Logit transforms values (e.g., Φ) that are btw 0 and 1 to make it go from –inf to +inf

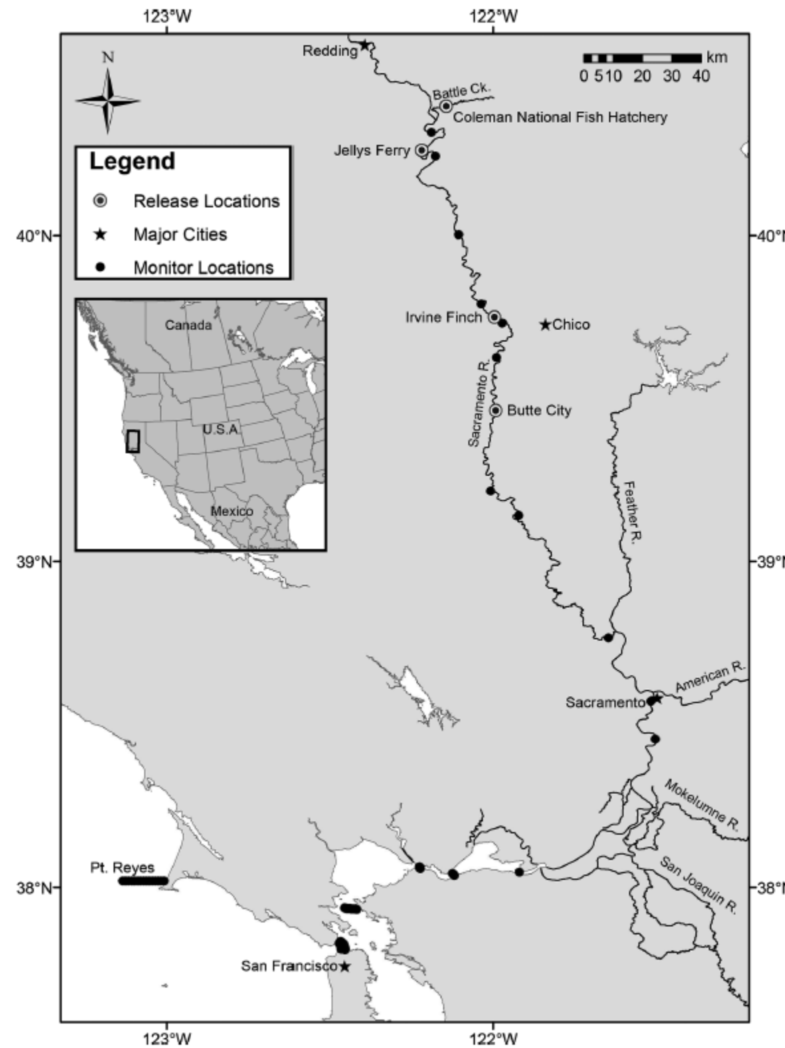
$$\log\left(\frac{\Phi_i}{1-\Phi_i}\right) = \beta_0 + \beta_1 X_{1i} \ldots$$

Alternative formulation (in probability form)

$$\Phi_i = \frac{e^{\beta_0+\beta_1 X_{1i}\cdots}}{1+e^{\beta_0+\beta_1 X_{1i}\cdots}}$$

CJS Example

- Implanted acoustic tags in 1350 late-fall Chinook smolts
- Tracked movements with acoustic receivers
- Applied CJS model to estimate survival throughout Sac River, Delta, and SF estuary



Michel et al. 2015

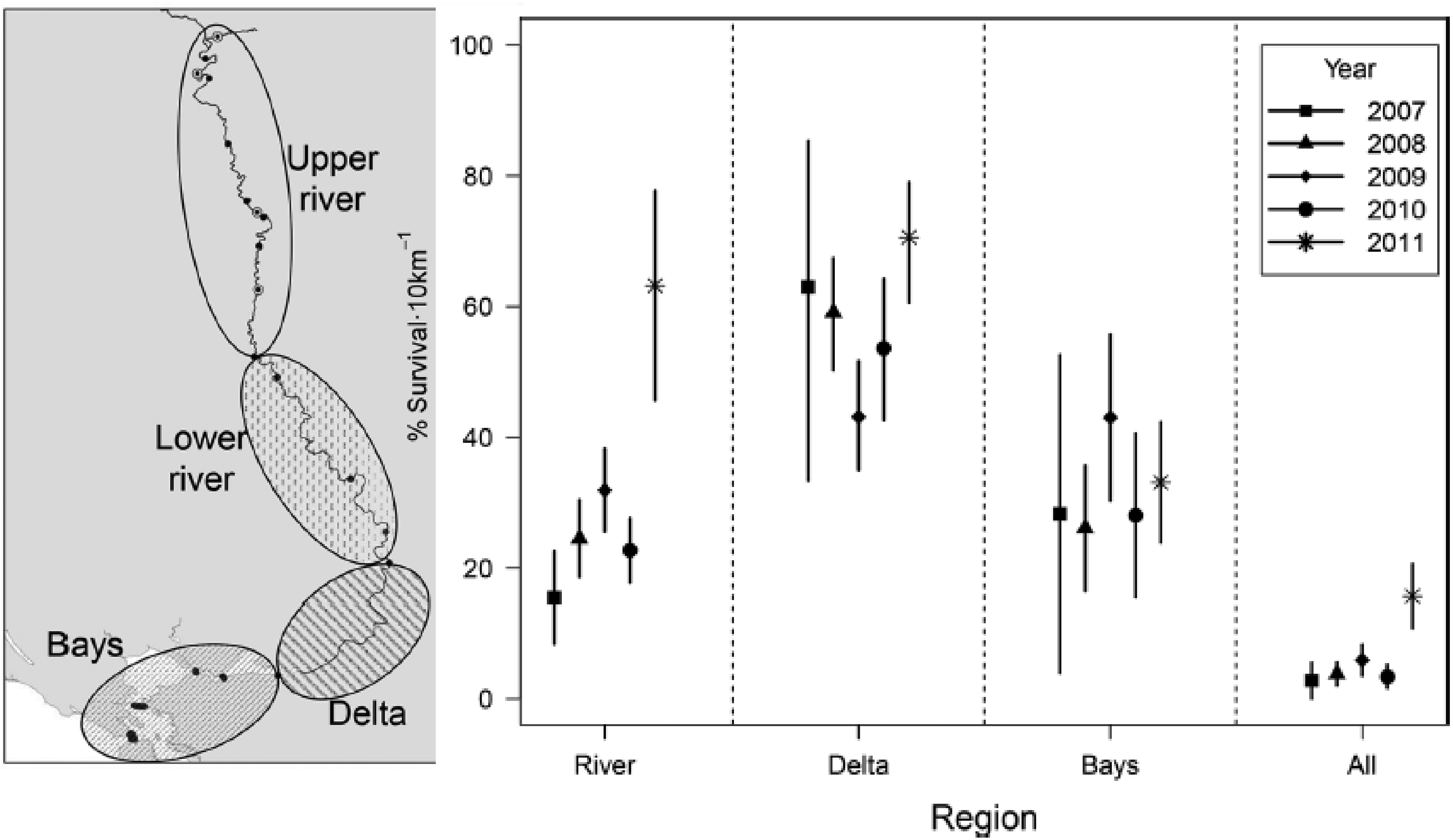
CJS Example

Table 3. Survival models for different spatial and temporal factors, as well as individual covariates, ordered from lowest to highest $QAIC_c$, omitting 2011 data.

Survival (φ) treatment	$\Delta QAIC_c$	No. of parameters
(River survival × year) × reach	0.0	126
(Delta survival × year) × reach	25.3	93
Base model (reach)	26.6	90
Reach + length	26.6	91
Reach × year	27.9	144
Reach × length	40.0	108
(Bays survival × year) × reach	49.0	105
Reach × mass	50.0	108
Reach × release	53.8	126
Reach × year × release	270.8	288
Null model (constant survival)	308.4	73

Michel et al. 2015

CJS Example



Michel et al. 2015

Outline

- A brief history
- Closed population methods
  - Lincoln-Petersen
- Open population methods
  - Capture-recapture (Cormack-Jolly-Seber) models
  - Recovery (Brownie) models
- Extensions
  - Multi-state models
  - Robust design

Recovery Model

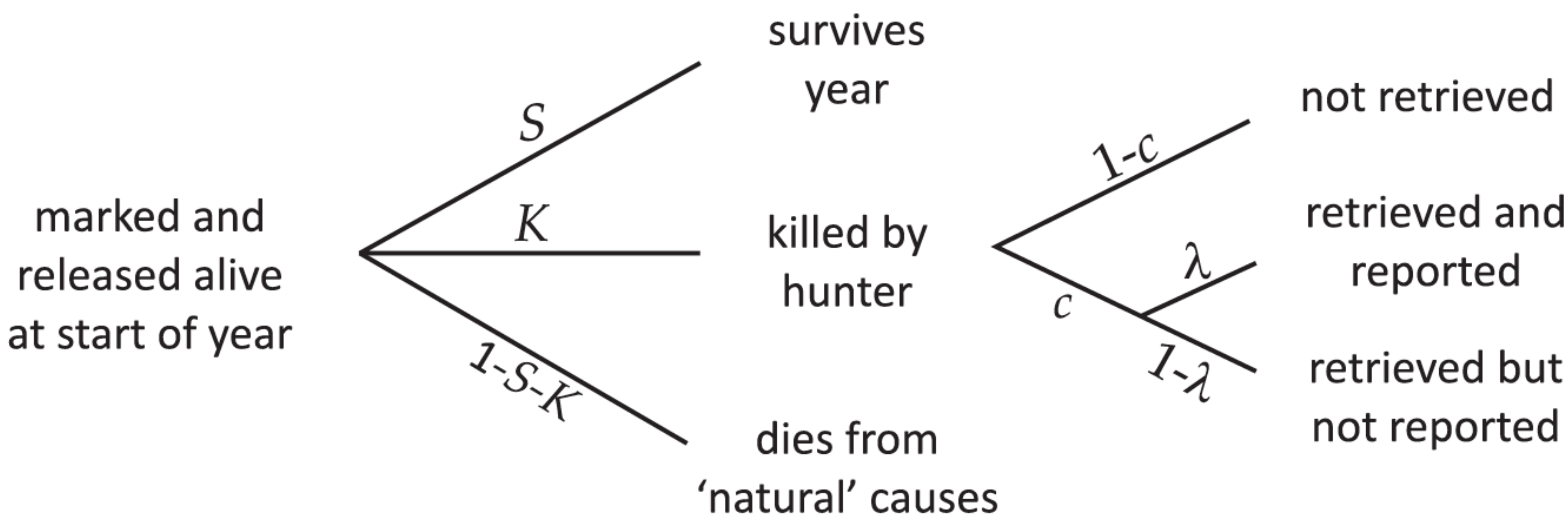
- Models where a tag is ‘recovered’ and not returned to the population
  - Harvest

www.skysguideservice.com

Recovery Model

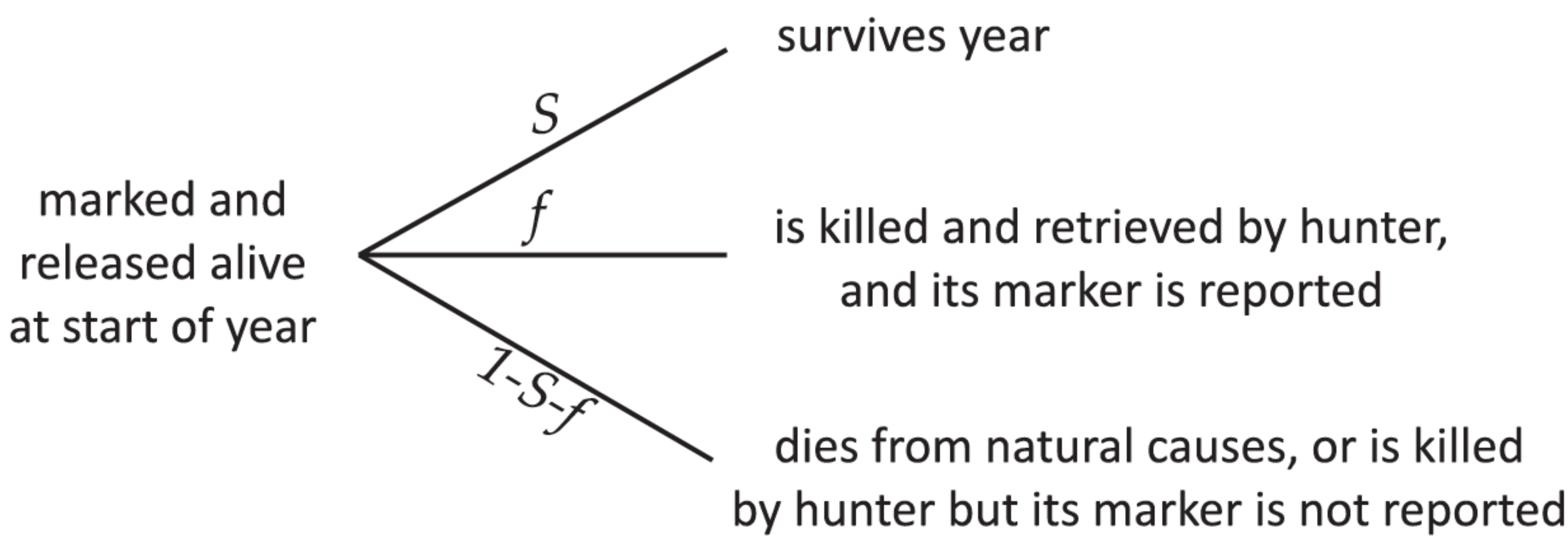
- Models where a tag is 'recovered' and not returned to the population
  - Harvest
- Also requires at least 3 sampling occasions
- Often referred to as 'Brownie models'
  - Brownie et al. 1985

Recovery Model



Cooch and White chapter 8

Recovery Model



Brownie et al 1985

Recovery Model

		year recovered			
year marked	*number marked*	1	2	3	$l = 4$
1	N_1	$N_1 f_1$	$N_1 S_1 f_2$	$N_1 S_1 S_2 f_3$	$N_1 S_1 S_2 S_3 f_4$
2	N_2		$N_2 f_2$	$N_2 S_2 f_3$	$N_2 S_2 S_3 f_4$
3	N_3			$N_3 f_3$	$N_3 S_3 f_4$
$k = 4$	N_4				$N_4 f_4$

N_t = Number of individuals marked at time t
f_t = proportion harvested and reported
S_t = proportion surviving (or harvested and not reported)

Recovery matrix

TABLE 16.8 Recoveries of Adult Male Mallards Banded during January/February in Illinois[a]

		Recovered during hunting season										
Year	**Number banded**	**1963**	**1964**	**1965**	**1966**	**1967**	**1968**	**1969**	**1970**	**1971**	**1972**	**1973**
1963	2583	91	89	24	18	16	11	8	7	7	2	6
1964	3075		141	45	52	50	17	30	21	16	7	3
1965	1195			27	31	21	8	19	7	9	4	3
1966	3418				156	92	44	50	49	34	23	5
1967	3100					113	68	57	65	41	23	10
1968	2400						63	52	59	44	30	12
1969	2601							91	80	58	37	25
1970	4433								222	169	95	46

[a]From Brownie *et al.* (1985).

Parameter Estimation

- Method of Moments

		year recovered			
year marked	*number marked*	1	2	3	$l = 4$
1	N_1	$N_1 f_1$	$\boxed{N_1 S_1 f_2}$	$N_1 S_1 S_2 f_3$	$N_1 S_1 S_2 S_3 f_4$
2	N_2		$\boxed{N_2 f_2}$	$N_2 S_2 f_3$	$N_2 S_2 S_3 f_4$
3	N_3			$N_3 f_3$	$N_3 S_3 f_4$
$k = 4$	N_4				$N_4 f_4$

TABLE 16.8 Recoveries of Adult Male Mallards

		Re			
Year	**Number banded**	**1963**	**1964**	**1965**	**1966**
1963	2583	91	89	24	18
1964	3075		141	45	52
1965	1195			27	31
1966	3418				156
1967	3100				

Parameter Estimation

- Method of Moments

year marked	*number marked*	*year recovered* 1	2	3	$l = 4$
1	N_1	$N_1 f_1$	$N_1 S_1 f_2$	$N_1 S_1 S_2 f_3$	$N_1 S_1 S_2 S_3 f_4$
2	N_2		$N_2 f_2$	$N_2 S_2 f_3$	$N_2 S_2 S_3 f_4$
3	N_3			$N_3 f_3$	$N_3 S_3 f_4$
$k = 4$	N_4				$N_4 f_4$

$$\frac{r_{12}}{r_{22}} = \frac{N_1 S_1 \cancel{f_2}}{N_2 \cancel{f_2}}$$

$$S_1 = \frac{r_{12} N_2}{r_{22} N_1}$$

TABLE 16.8 Recoveries of Adult Male Mallards

Year	Number banded	Re… 1963	1964	1965	1966
1963	2583	91	89	24	18
1964	3075		141	45	52
1965	1195			27	31
1966	3418				156
1967	3100				

$$S_1 = \frac{89 \cdot 3075}{141 \cdot 2583}$$

$$S_1 = 0.75$$

Parameter Estimation

- Method of Moments
- Maximum Likelihood
  - Uses all available data
  - Reduces bias
  - Better (and more common) than the method of moments approach

Assumptions

(Same as CJS model):

- Survival is equal for marked and unmarked animals
- Tagging does not influence capture probability
- Sampling is ‘instantaneous’
- Tags are not lost or overlooked
- Fate of each fish is independent

Extra

Mortality

- The parameter f_i (prob of harvest and reporting) is a function of the probability of harvest (u_i) and the tag-reporting rate (λ_i)

$$f_i = u_i \lambda \longrightarrow u_i = \frac{f_i}{\lambda}$$

Extra

Mortality

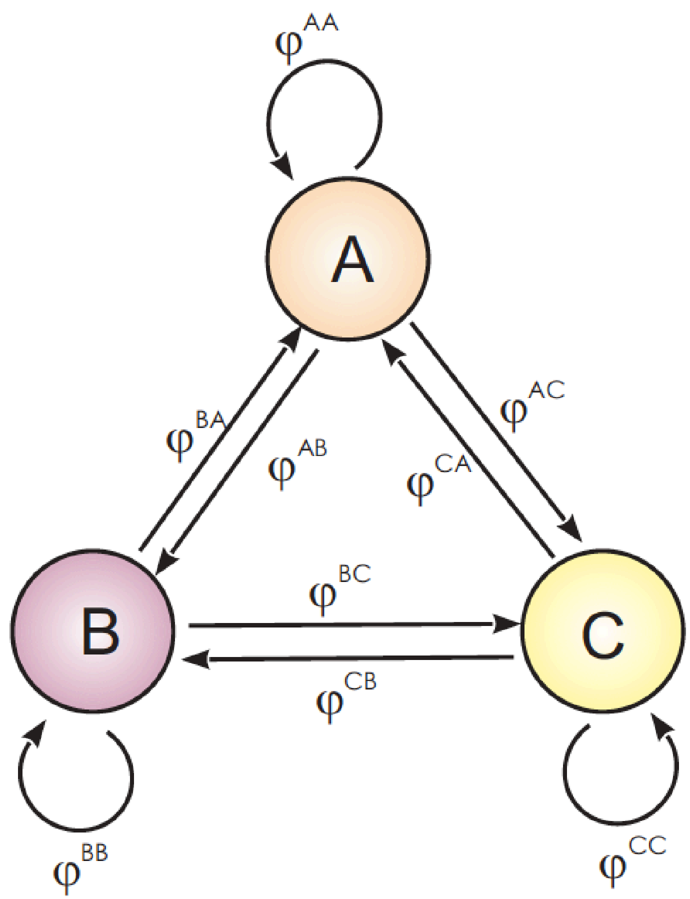
- The parameter f_i (prob of harvest and reporting) is a function of the probability of harvest (u_i) and the tag-reporting rate (λ_i)
- Methods to estimate reporting rate
  – High reward tags (Pollock et al. 2001)
  – Planted tags
  – Port sampling
  – Two samples per year (Hearn et al 1998)

Outline

- A brief history
- Closed population methods
  - Lincoln-Petersen
- Open population methods
  - Capture-recapture (Cormack-Jolly-Seber) models
  - Recovery (Brownie) models
- Extensions
  - Multi-state models
  - Robust design

Multistate Models

- Animals can move between different discrete ‘states’

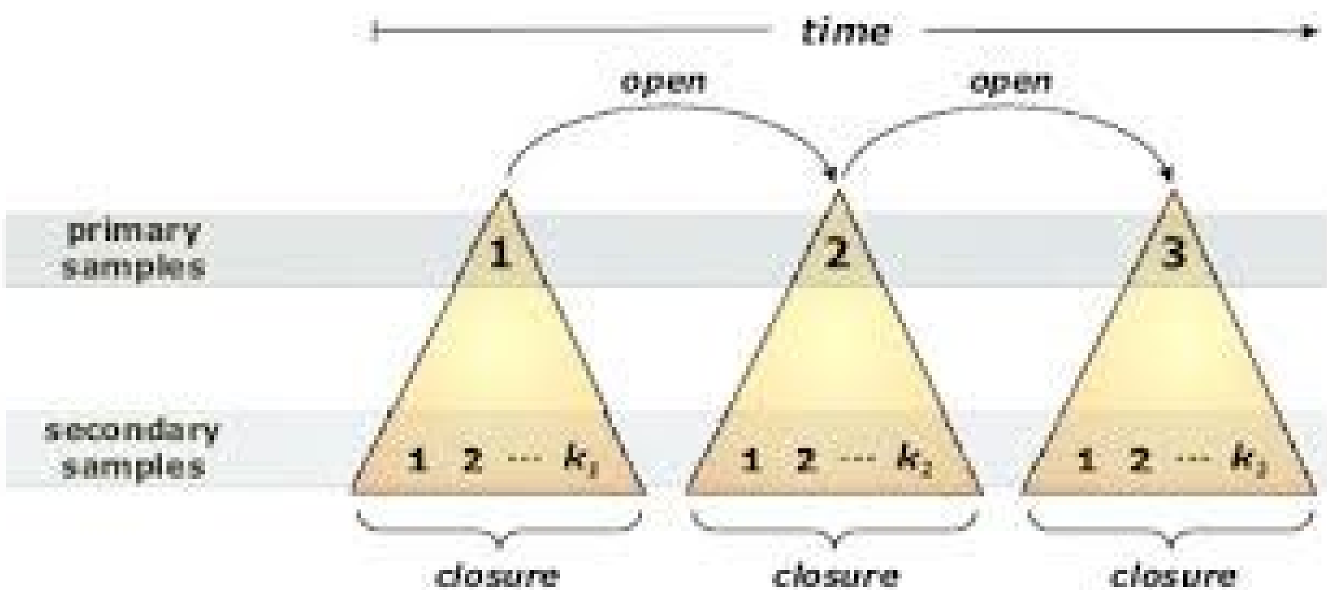


Multistate Models

- Animals can move between different discrete ‘states’
- Examples of states
  - Spatial (spawning grounds)
  - Ontogenetic (immature vs mature)
  - Related to harvest (sublegal vs legal)
- Advantages
  - Can separate the probability of survival and moving between states
- Disadvantages
  - Very data hungry!

Robust Design

- Combine intense sample over short period (closed population) with longer term sampling (open population)



Robust Design

- Combine intense sample over short period (closed population) with longer term sampling (open population)
- Advantages
  - Robust estimates of abundance and recruitment
  - More precise estimates
  - Estimate temporary emigration

Useful References

- Amstrup S.C., McDonald T.L., and Manly B.F.J. (2005) Handbook of Capture-Recapture Analysis. Princeton University Press, Princeton, NJ.
- Brownie C., Anderson D.R., Burnham K.P., and Robson DS (1985) Statistical inference from band recovery data: a handbook. United States Fish and Wildlife Service Resource Publication Number 156
- **Cooch E.G. and White G.C. (2014) Program MARK: A gentle introduction**
  - **http://www.phidot.org/software/mark/docs/book/**
- Lebreton J-D, Burnham K.P., Clobert J., and Anderson D.R. (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62: 67-118.
- **Pine, William E., Kenneth H. Pollock , Joseph E. Hightower , Thomas J. Kwak & James A. Rice. (2003) A Review of Tagging Methods for Estimating Fish Population Size and Components of Mortality. Fisheries 28:10-23.**
- Williams B.K., Nichols J.D., and Conroy, M.J. (2001) Analysis and Management of Animal Populations. Academic Press, San Diego, CA.

Summary - mark-recapture models

- General
  - Purpose: estimate mortality (Z,F,M) or abundance
  - Numerous methods with diff. assumptions/goals!
  - Methods for **open** vs. **closed** populations
- Closed population methods
  - **Lincoln-Petersen** – Estimate N; simple; know basics & assumptions
- Open population methods
  - **Cormack-Jolly-Seber (CJS) models** – Estimate survival & capture probability; know basic idea; be able to write out capture history probabilities (for simple example)
  - **Recovery (Brownie) models** – tags are harvested by fishery; **tag reporting rate** is important; ML method better than method of moments
- Extensions
  - **Multi-state models**
  - **Robust design**