Stock-recruitment Part I

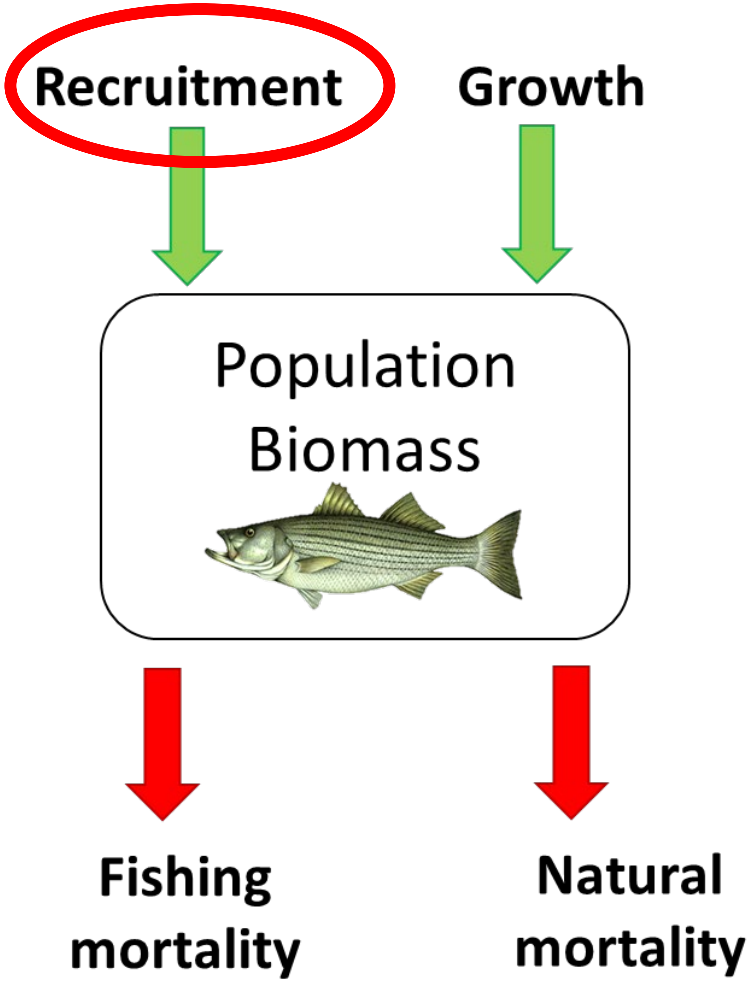
Reading:

Jennings et al. 2001. Marine Fisheries Ecology, Chapter 4 (section 4.2)

Advanced: Quinn and Deriso 1999, Chapter 3

Stock-recruitment

- **Recruitment** = The number of individuals that reach a specified “stage” of their life cycle. Often in reference to a specific age or to a size/age that is vulnerable to fishing



Stock-Recruitment

- Why is this one of the most important topics in fisheries mgmt?
- Answers questions like:
  - How many young are produced by a given amount of adults?
  - How many recruits can we expect in the future for a given stock size?
  - Ability of pop to renew itself is critical for management and conservation
  - How much sustainable harvest can populations generate?

Stock-recruitment

- Interested in finding the relationship between spawners and recruitment

Eggs → larvae → juveniles → spawners

- Side: multiple ways of saying the same thing:
  - Stock-recruitment
  - Stock-recruit
  - Spawner-recruit
  - SR

What is a stock?

What is a stock?

- **Stocks** are arbitrary groups of fish large enough to be essentially self-reproducing, with members of each group having similar life history characteristics
    - Hilborn and Walters 1992

- More generally: **Stock** defines a semi-discrete group of fish with some definable attributes of interest to managers
    - Begg et al. 1999

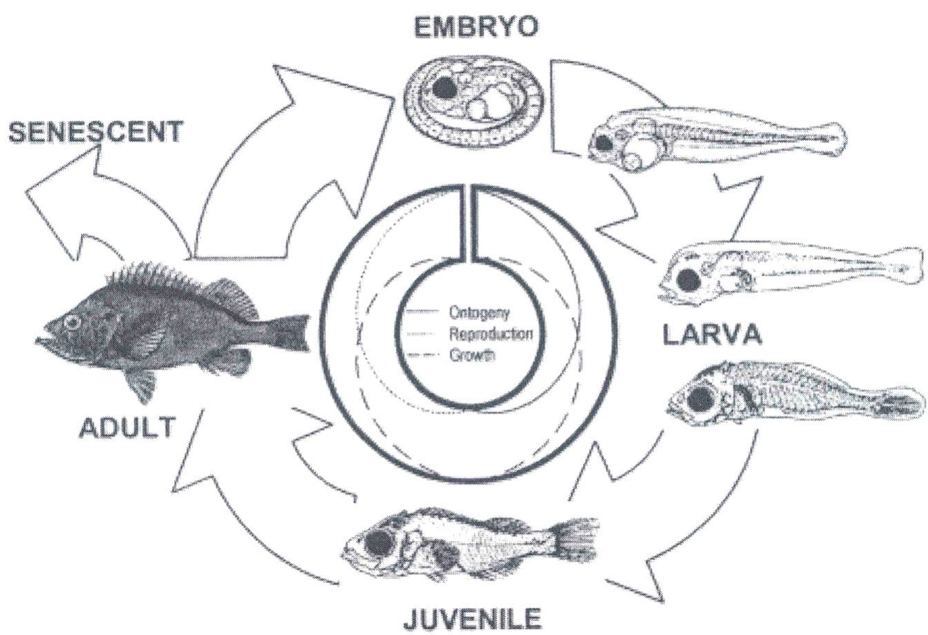
Begg and Waldman 1999

What is a stock?

- ***spawning stock*** - Reproductively active part of the population
- Indices/Proxies include (descending order of quality):
  - Egg (or larval) production in the population
  - Biomass of adult females
  - Biomass of adults
  - Number of adults or adult females
  - CPUE of adults

What is recruitment?

- **Recruitment** = The number of individuals that reach a specified “stage” of their life cycle
  - Typically after egg/larval stage
  - Or referring to fishery
  - Combines reproduction and early life survival
  - Ideally, year class strength set by the selected stage



Where do spawner and recruit data come from?

- Monitoring surveys
  - For eggs/juveniles/adults
- Population models (i.e. stock assessment models)

What does R vs. S look like?

- Draw relationship
- Chat with neighbor
- Why do you think it looks like that?

Shapes of S-R models

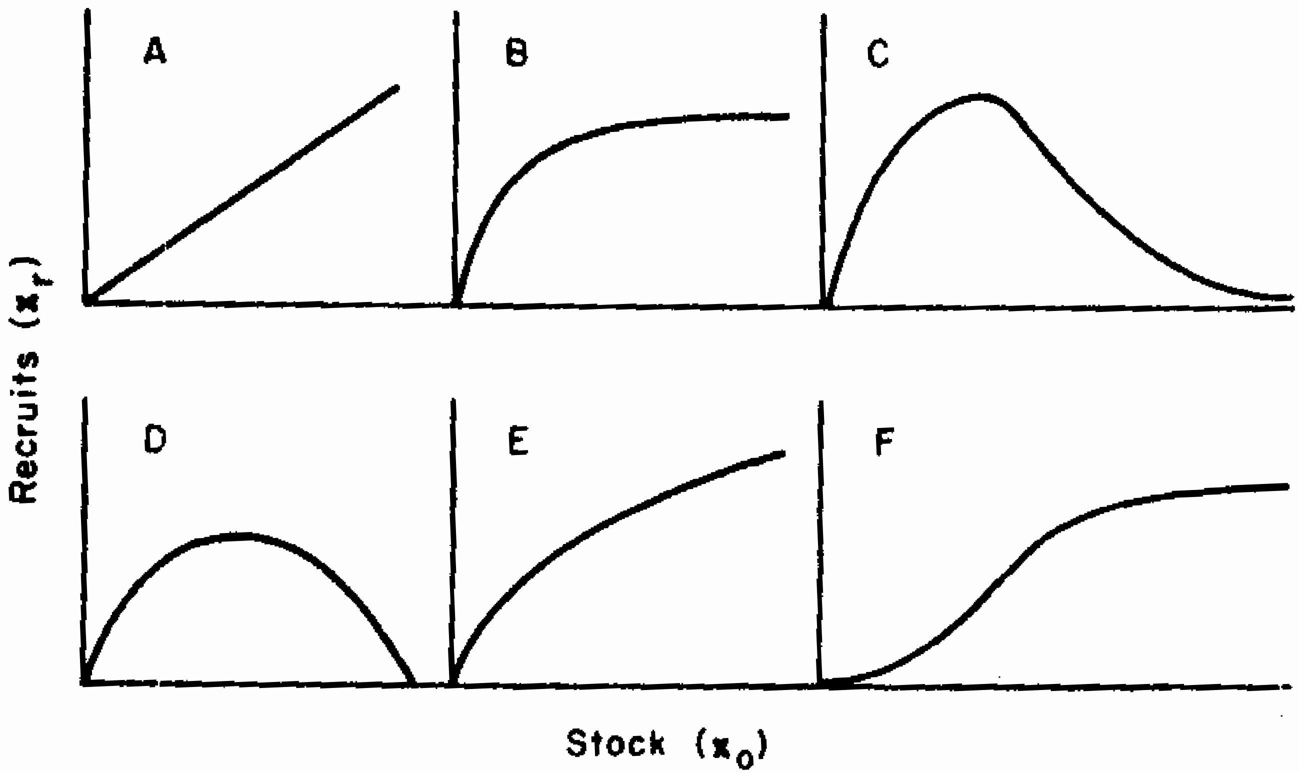
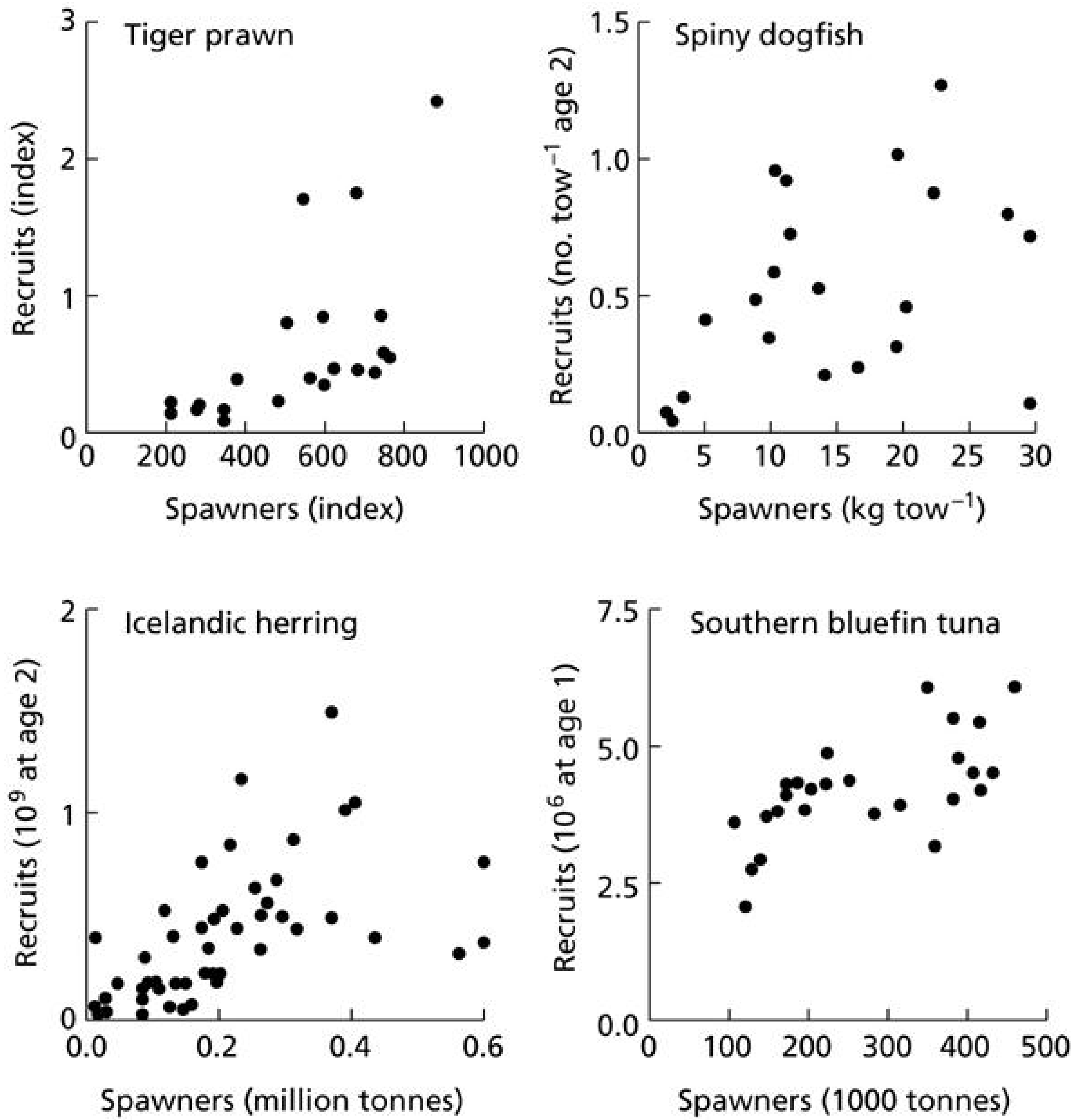


FIGURE 3.1. Nonlinear relationships between newborn individuals (x_0) and survivorship to age r (x_r); see equation (3.8). A, linear; B, Beverton and Holt; C, Ricker; D, Schaefer; E, power function ($0 < \beta < 1$); F, depensation.

Examples (real data)



Jennings, Kaiser, Reynolds
Marine Fisheries Ecology

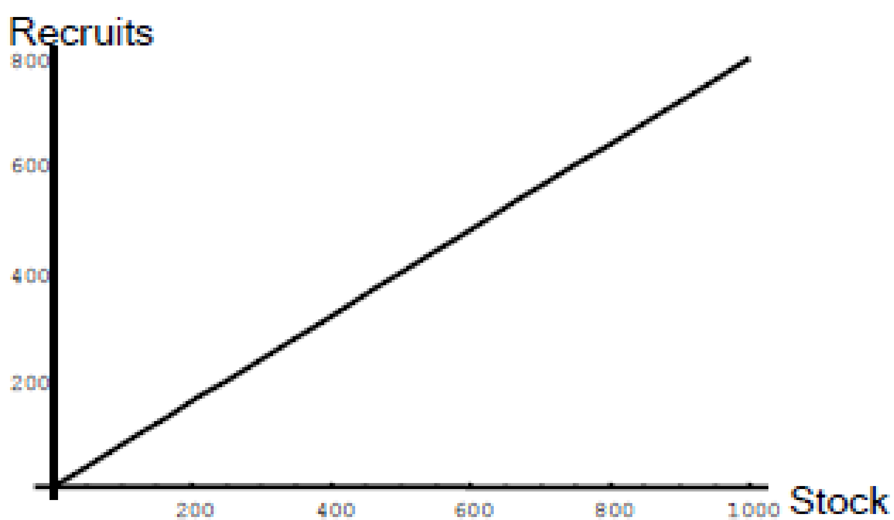
Stock-recruitment models

- 1. Density Independent
- 2. Beverton-Holt
- 3. Ricker
- 4. Shepherd
- 5. Hockey stick
- Others…

Density dependent

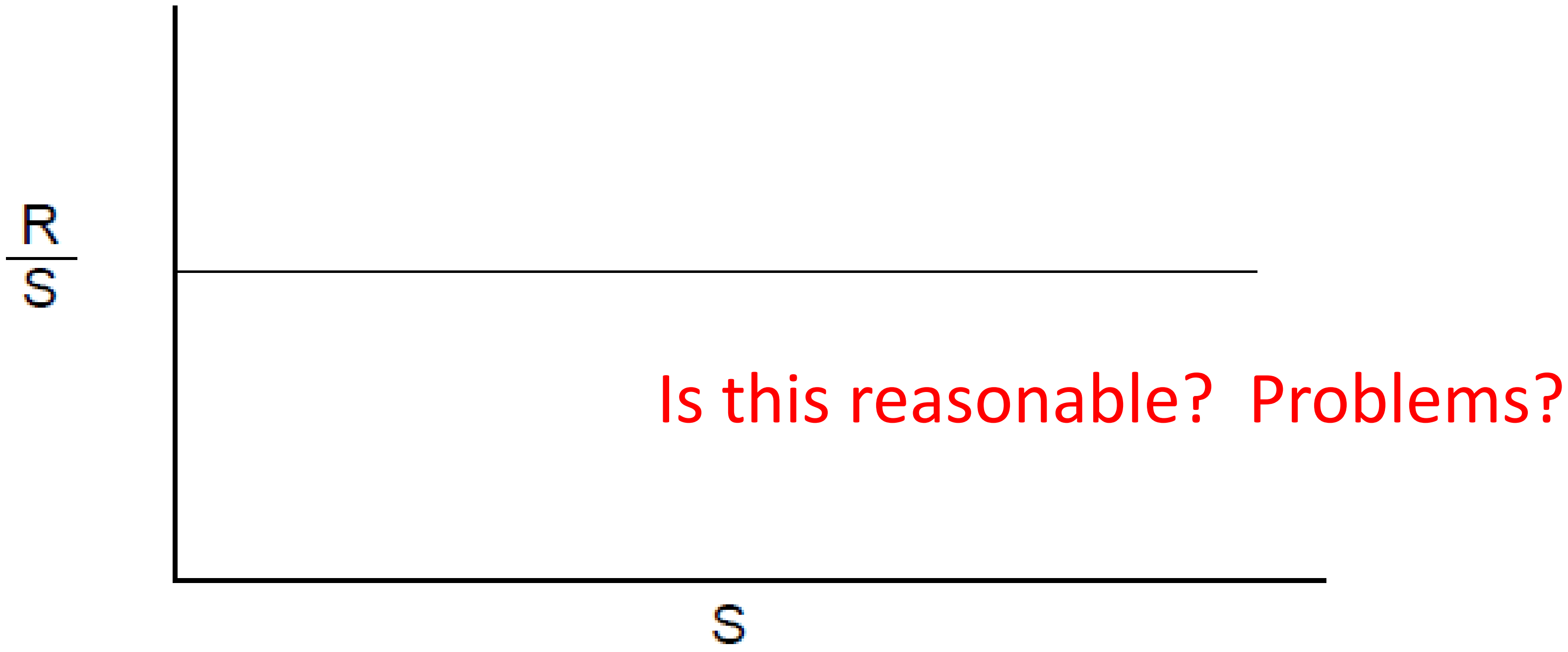
1. Density independent model

- Assumes density independence:
  - Survival (e.g., of eggs, juveniles, adults) does not depend on abundance
  - E.g., no intraspecific competition
- Model: $\boxed{R = a \cdot S}$
  - **a** = productivity or density-independent parameter (combines egg production per spawner and survival to recruit stage)



1. Density independent model

- **R/S = a**;
- R/S can be thought of as an index of survival



Models should be density dependent

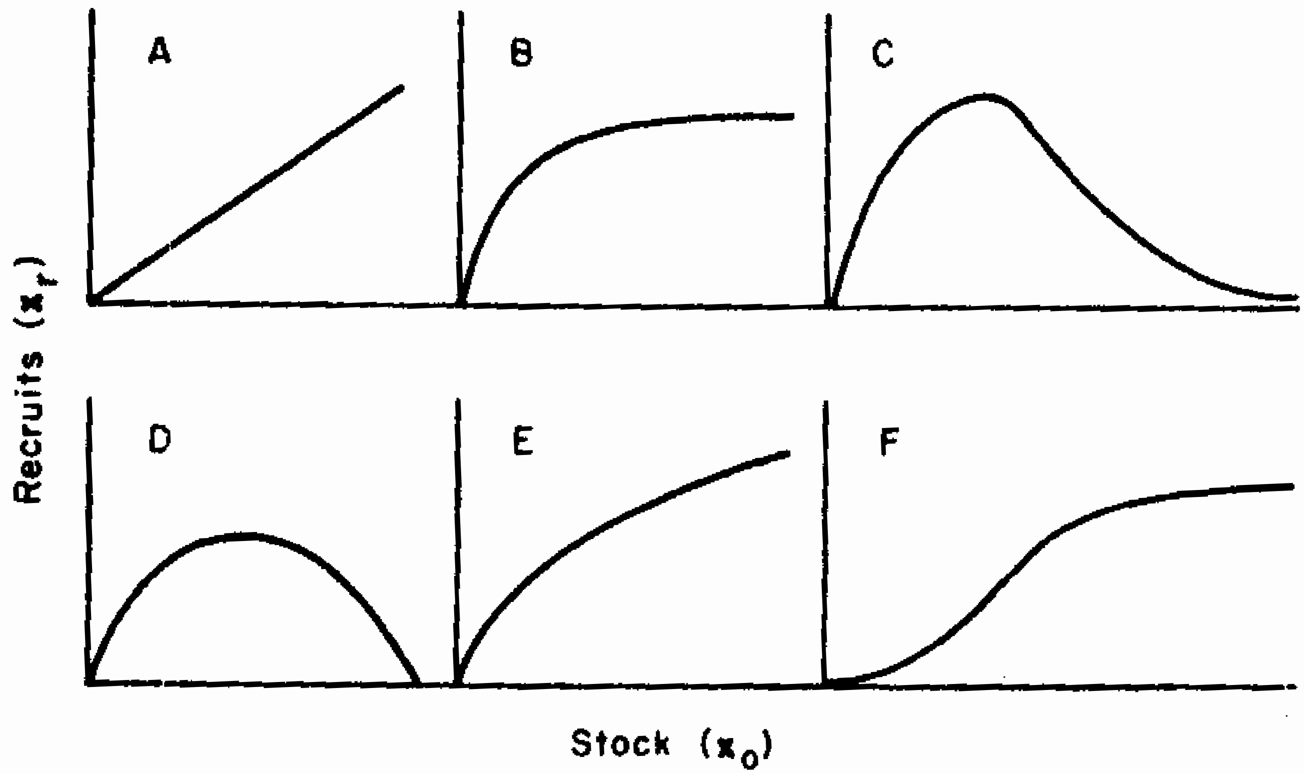
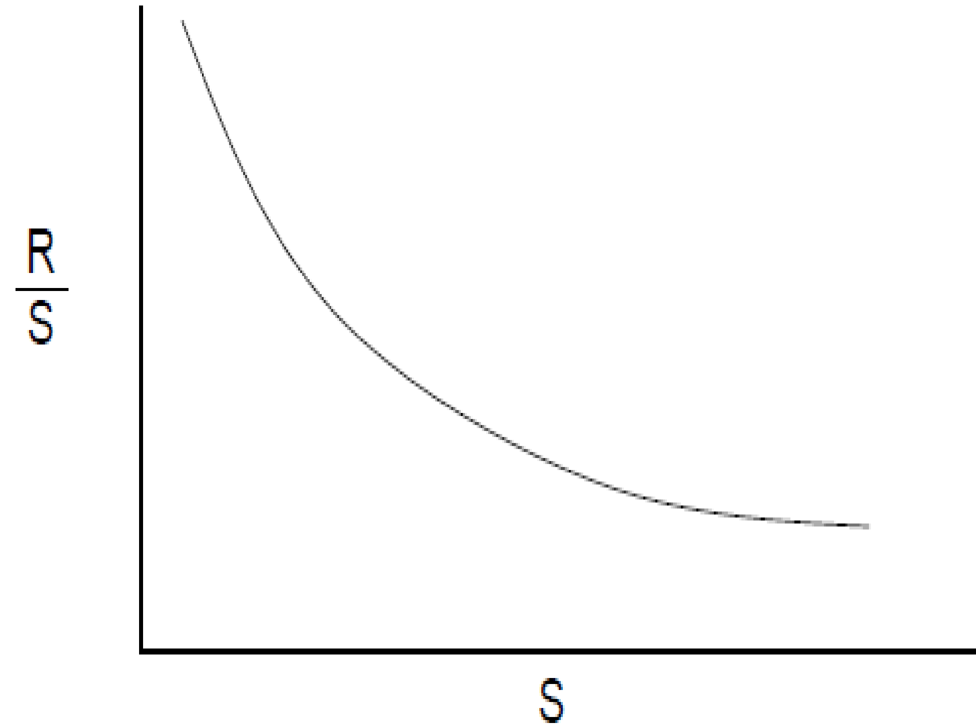


FIGURE 3.1. Nonlinear relationships between newborn individuals (x_0) and survivorship to age r (x_r); see equation (3.8). A, linear; B, Beverton and Holt; C, Ricker; D, Schaefer; E, power function ($0 < \beta < 1$); F, depensation.

Compensatory stock-recruitment

- **Density dependence** can be built in through compensation (ie changes in survival rate)
- **Compensation =**
  - recruits per spawner (R/S) increases as S decreases, *or*
  - R/S decreases, as S increases

- Why?
  - Density independent model leads to exponential growth
  - Compensation assumes survival declines as S increases (e.g., cannibalism, disease, competition for prey, predation)



Compensatory stock-recruitment

- Density-dependent processes thought to act via two mechanisms:
  - **1. acts via the juveniles** (affected by recruit density)
    - E.g., young compete for food or feed on one another
    - → *Leads to Beverton Holt model*
  - **2. acts via the adult stock** (affected by spawner density)
    - Aka “stock dependent”
    - E.g., limited numbers of nests, cannibalism of young by adults, disease transmission from adults to juveniles, density-dependent growth combined with size-specific predation
    - → *leads to Ricker model*

Extra

2. Beverton-Holt Model

- Based on idea of density-dependence acting on the **juveniles**
  - Change in abundance affected by total mortality, Z
  - Assume Z described by linear function of abundance, affected by density independent (a) and density dependent (b) parameters
  - Leads to nonlinear differential equation (see Quinn & Deriso for solution)

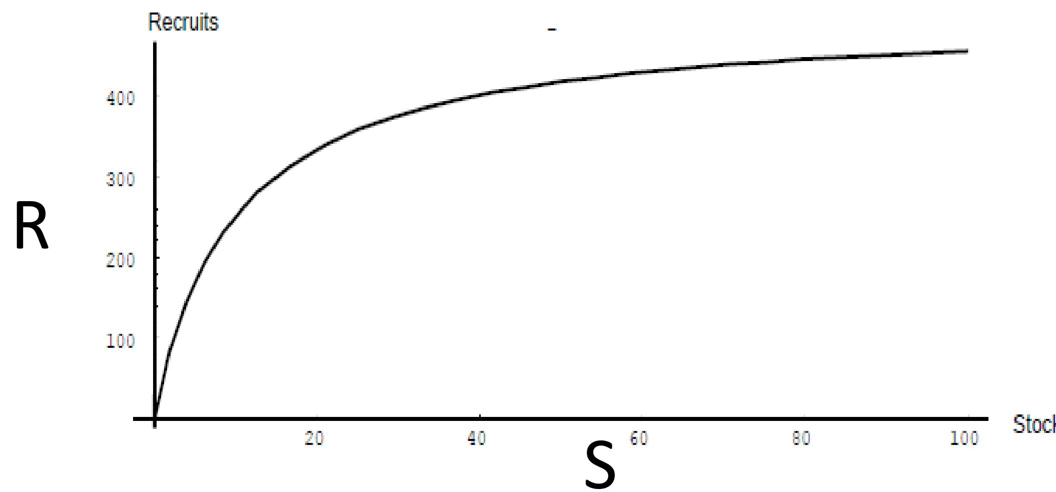
$$\frac{dN}{dt} = -ZN$$

$$Z = a + bN$$

$$\frac{dN}{dt} = -(a + bN)N$$

2. Beverton-Holt Model

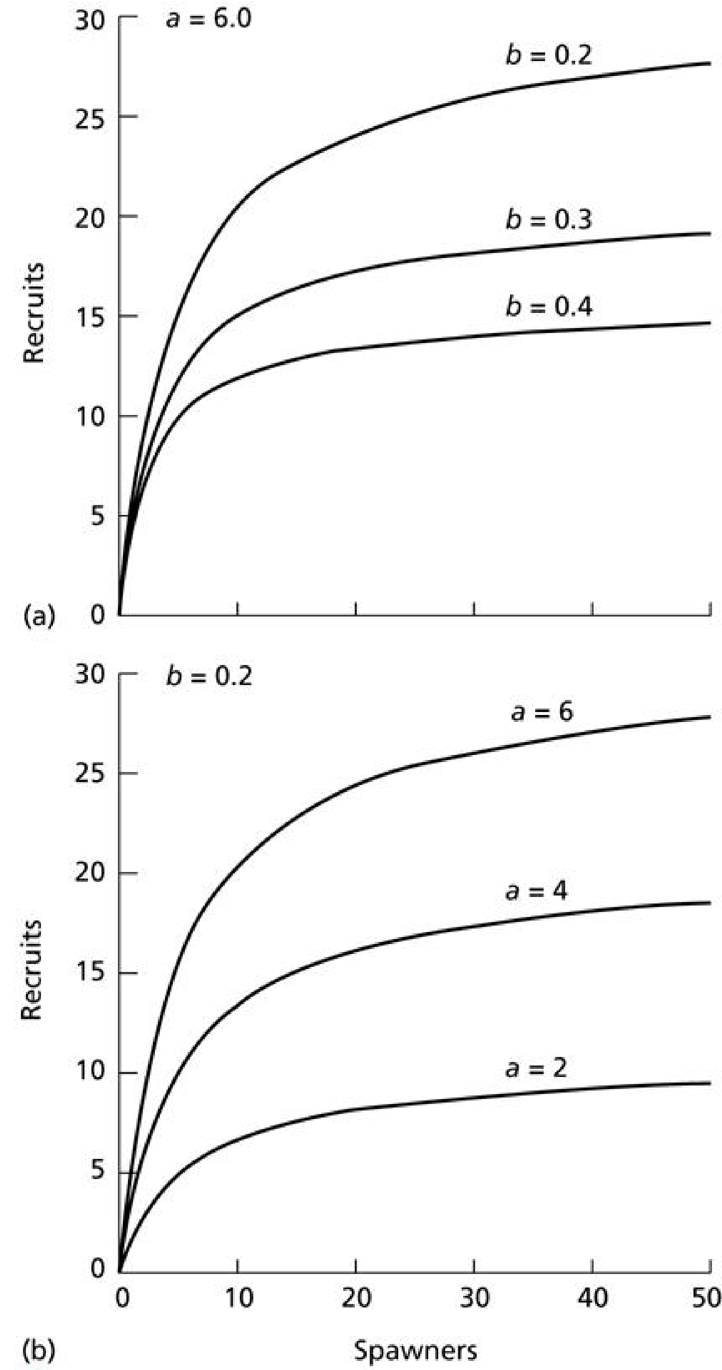
$$R = \frac{aS}{1 + bS}$$



- R = number (or biomass) of recruiting individuals
- S = number (or biomass) of spawners
- a = productivity parameter (number of R per S at low S); slope of curve at the origin
- b = parameter for degree of density dependence (affects rate of approaching asymptote)

- Basic property: R constantly increases toward an asymptote as S increases

2. Beverton-Holt Model



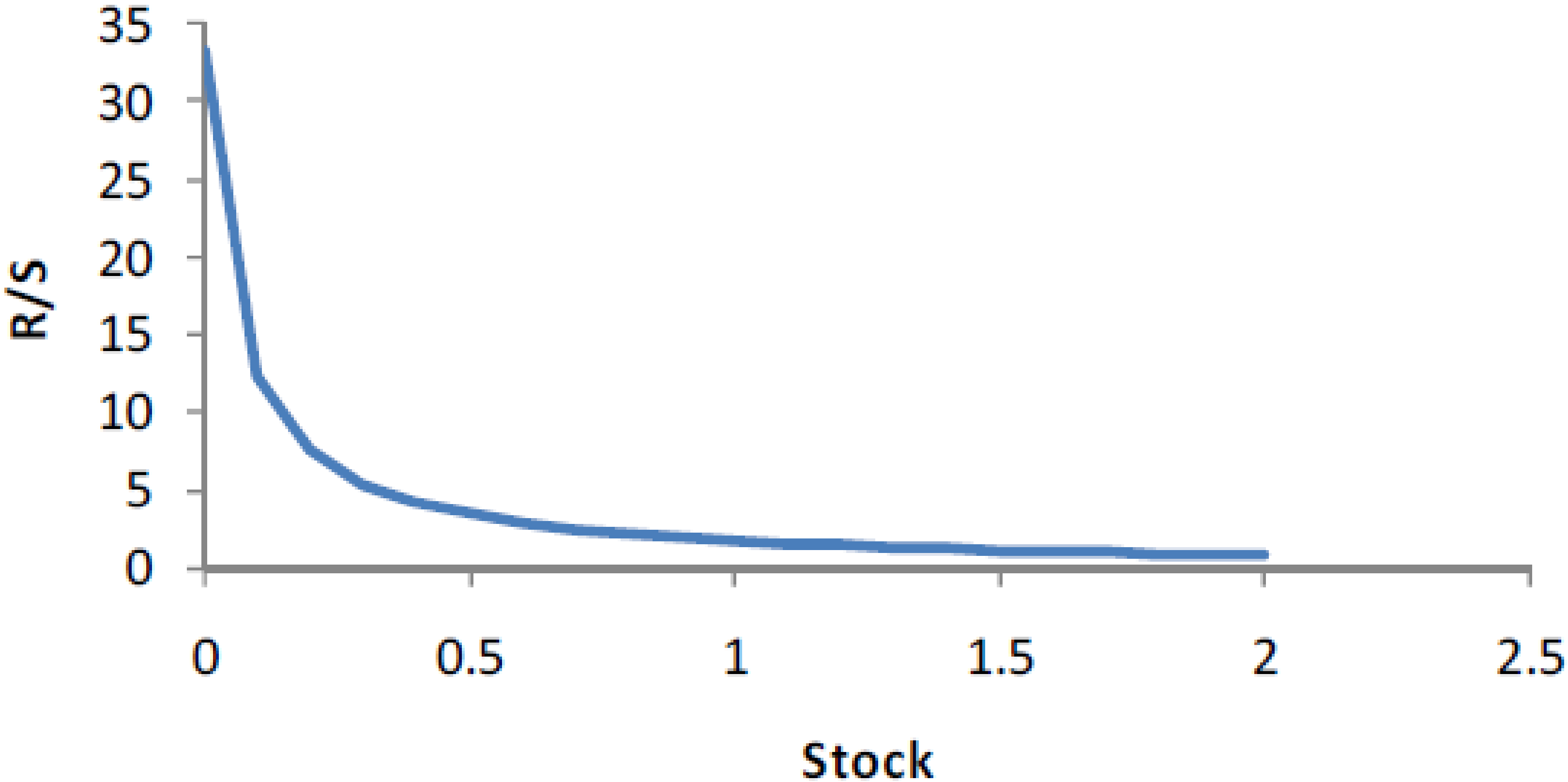
- Effects of changing parameters

$$R = \frac{aS}{1 + bS}$$

- Max R = a/b

2. Beverton-Holt Model

- What does survival index (R/S) look like?



Extra

Beverton Holt parameterizations

- There are different parameterizations for the same model
- Parameter meanings differ
  - a‘ = 1/a (from before)
  - b‘ = b/a (from before)
  - α = maximum # recruits produced (=1/b’)
  - β = spawners needed to produce R= α/2 (=a’/b’)

$$R = \frac{aS}{1 + bS}$$

$$R = \frac{S}{a' + b'S}$$

$$R = \frac{\alpha S}{\beta + S}$$

Extra

Steepness Parameterization of BH

- Commonly used form of the Beverton-Holt Model

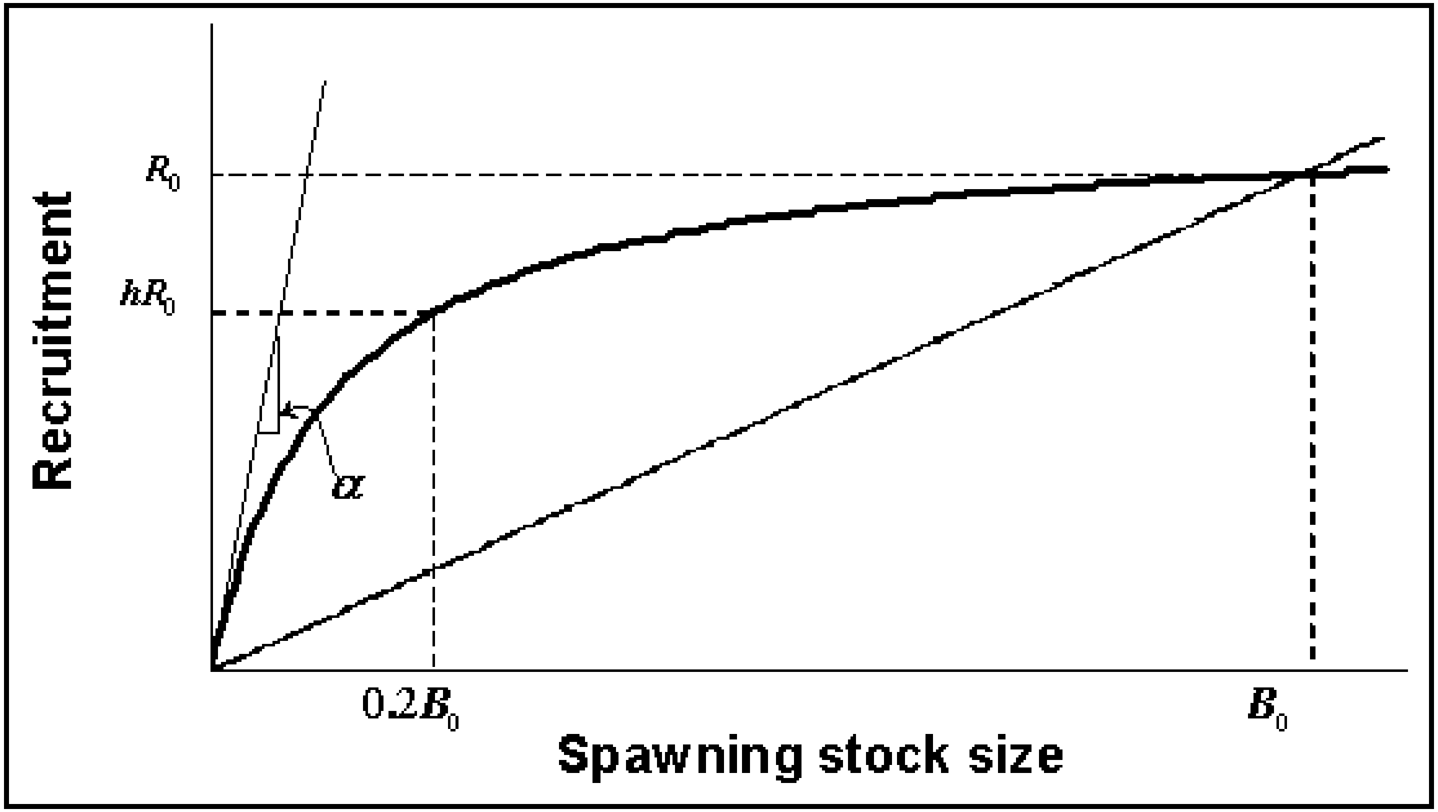
$$\boxed{R = \frac{4hR_{max}S}{S_{max}(1-h)+(5h-1)S}}$$

  - **h=steepness**, measure of how fast asymptote reached
    - Defined as: fraction of max recruitment attained when S is 20% of max (i.e., h = R/Rmax @ 20% Smax)
    - Bounded by: 0.2< h <1
  - R_{max} = Maximum recruitment (ie virgin, unfished R)
  - S_{max} = Maximum observed stock size (virgin, unfished S)

- Why used?
  - Steepness (h) is scale-less, allowing comparison across species
  - Common in marine systems

Extra

Steepness Parameterization of BH



- Note: $R_0 = R_{max}$
- h=steepness; fraction of max recruitment attained when S is 20% of max (i.e., h = R/Rmax @ 20% Smax)

Stock-recruitment Part II

Reading:

Jennings et al. 2001. Marine Fisheries Ecology, Chapter 4 (section 4.2)

Advanced: Quinn and Deriso 1999, Chapter 3

"Stock recruitment funny"

"They're the ideal temp workers! Easy to train, industrious, punctual… And with a 13-day life-span, they conveniently die before collecting a paycheck!"

Recap: Stock-recruitment models

Models

- 1. Density Independent
- 2. Beverton-Holt
- 3. Ricker
- 4. Shepherd
- 5. Hockey stick
- Others…

- What did BH model look like?
- What is compensation?
- How does density dependence play into BH model?

William E. Ricker

Extra

3. Ricker Model

- Based on idea of density-dependence acting on the **adults**
  - Change in juv. abundance (N) affected by total mortality, Z
  - Assume Z described by linear function of *spawner abundance (S)*, affected by density independent (a) and density dependent (b) parameters
  - Leads to nonlinear differential equation (see Quinn and Deriso for solution)

$$\frac{dN}{dt} = -ZN$$

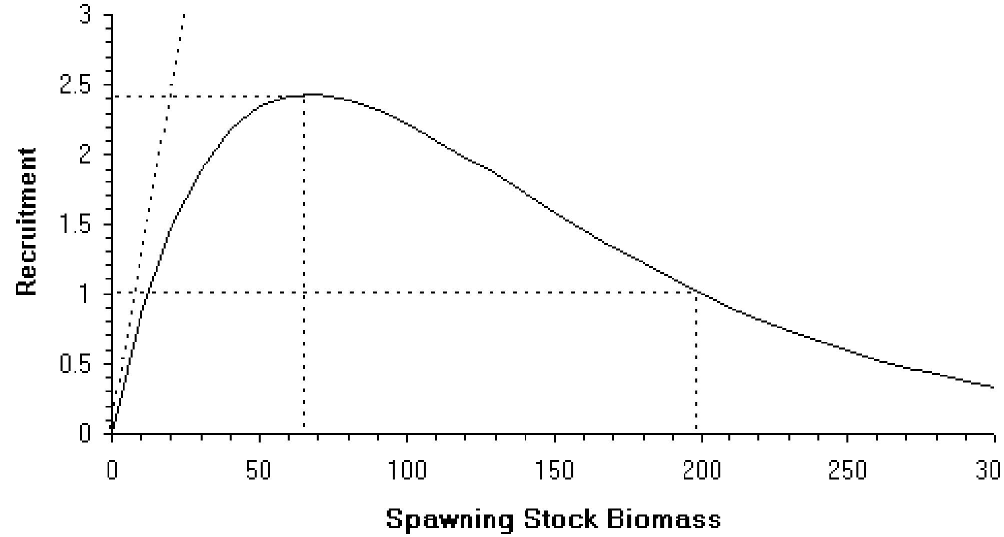
$$Z = a + bS$$

$$\frac{dN}{dt} = -(a + bS)N$$

3. Ricker Model

$$\boxed{R = aSe^{-bS}}$$

R



S

- R = number (or biomass) of recruiting individuals
- S = number (or biomass) of spawners
- a = productivity parameter (number of R per S at low S)
- b = parameter for degree of density dependence

- Basic property: “hump” shaped with declining R at higher S

3. Ricker Model

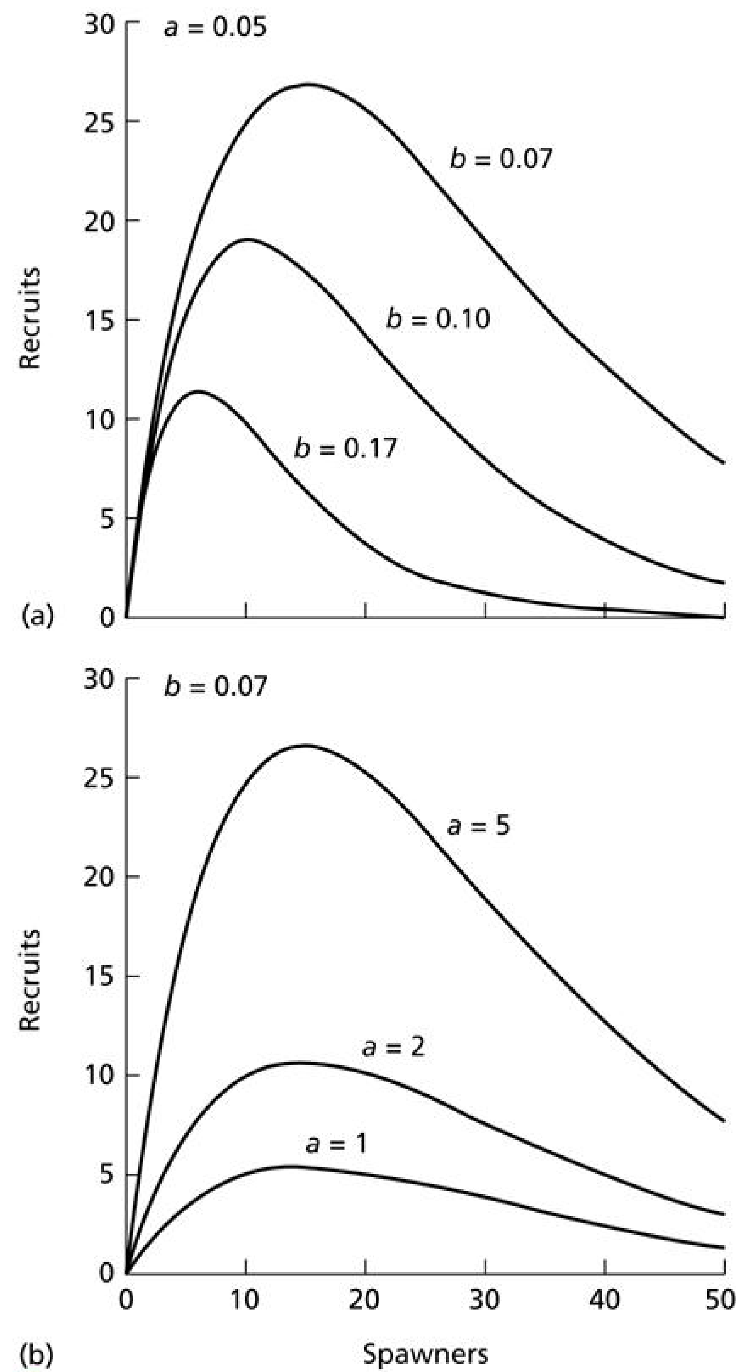
- Effects of changing parameters

$$\boxed{R = aSe^{-bS}}$$

  - Maximum mean R occurs at S=1/b

- Alternative parameterization you might see:

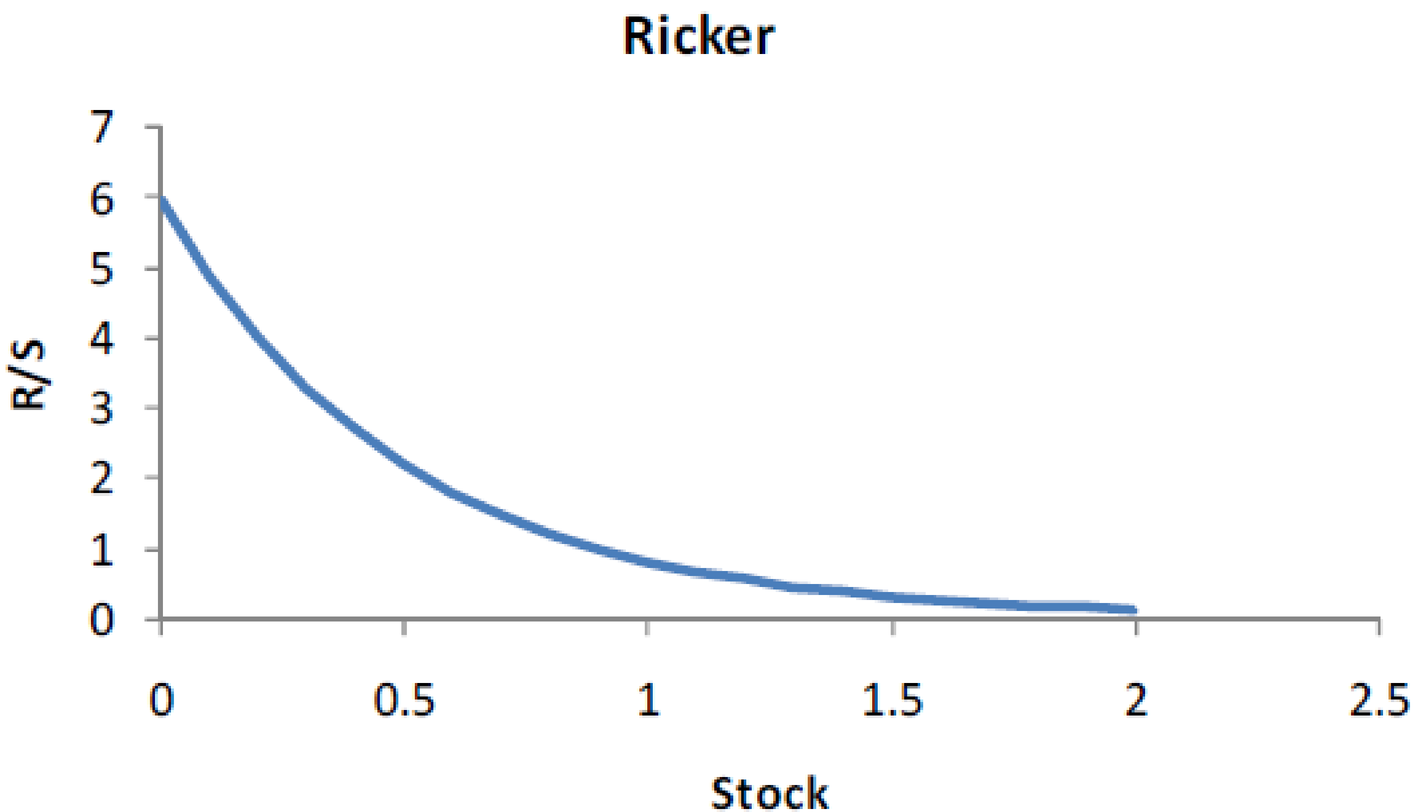
$$R = Se^{a'-bS}$$



Jennings, Kaiser, Reynolds

3. Ricker Model

- What does survival proxy (R/S) look like?
  - Like, BH, has stabilizing effect on population at low S

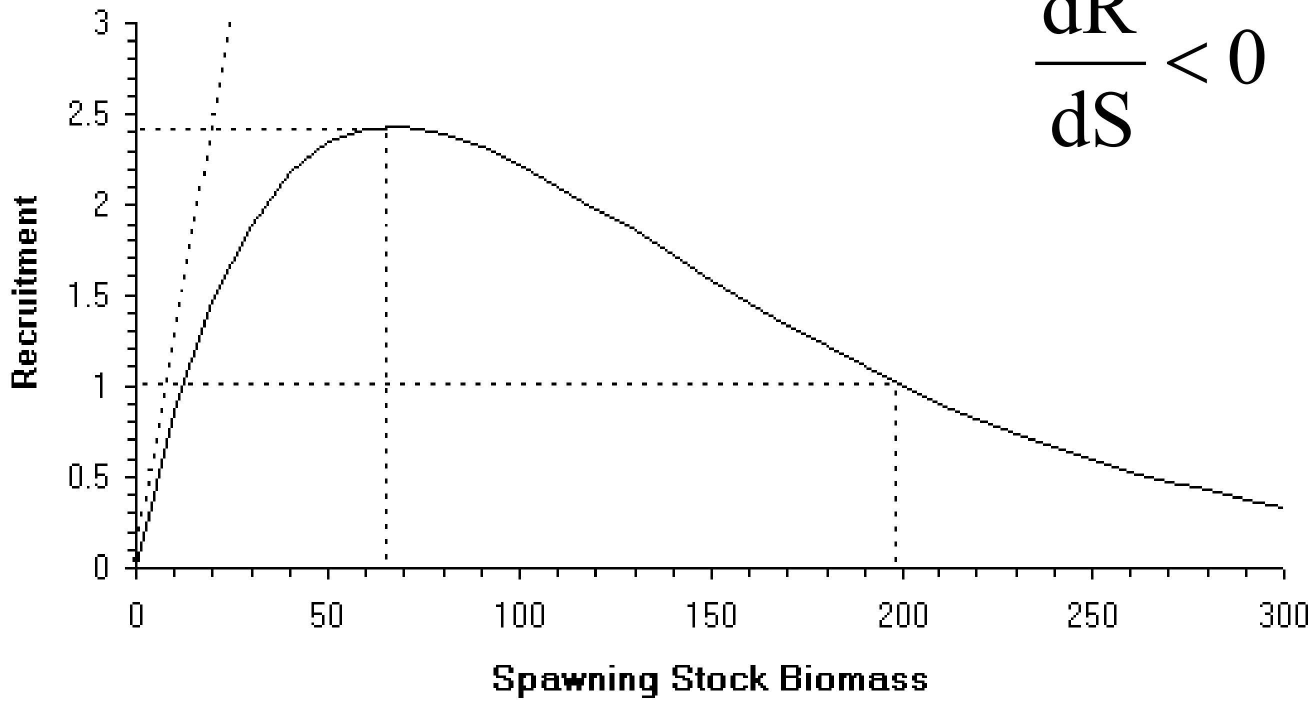


Ricker – overcompensation

- **Overcompensation** – decrease in recruitment with increasing spawning stock

$$\frac{dR}{dS} < 0$$

R



S

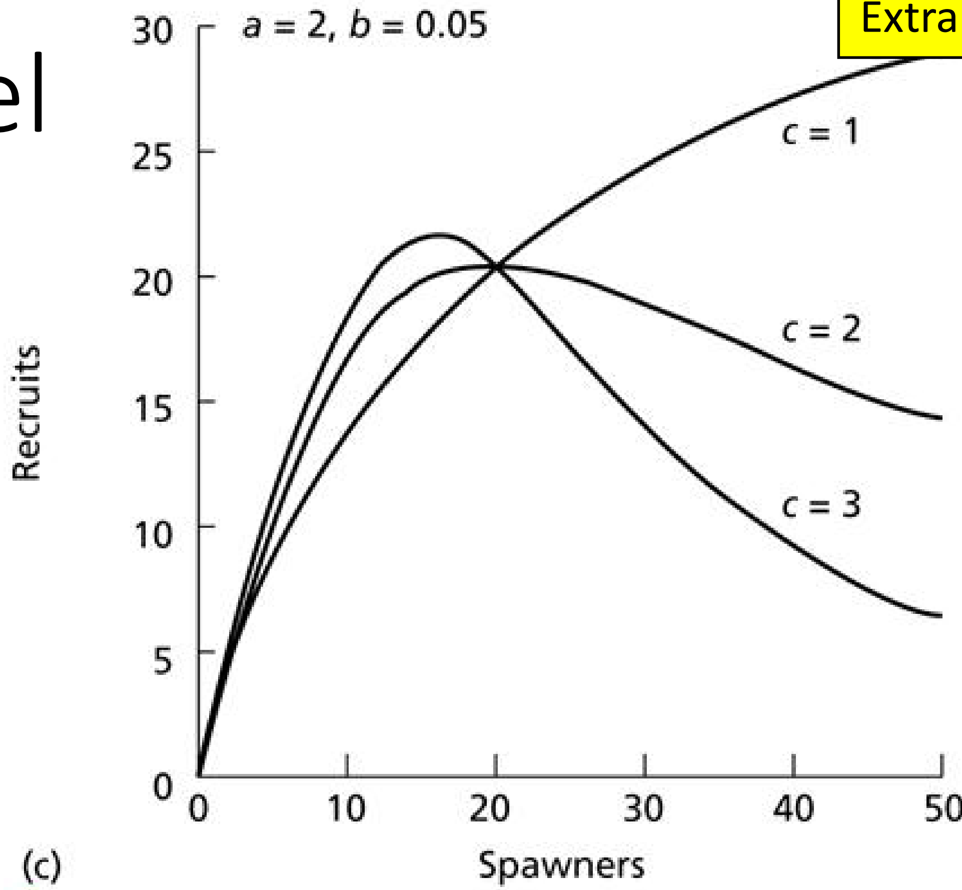
Ricker Overcompensation

- Recruitment decreases at large stock sizes
- **Some possible causes:**
  - 1. Cannibalism of juveniles by adults
  - 2. Disease transmission from adults to juveniles
  - 3. Oxygen limitations due to heavy egg deposit that affects all eggs
  - 4. Spawning site damage by adults
  - 5. Density dependent growth with size-dependent predation

4. Shepherd Model

- Generalizing equation

$$R = \frac{aS}{1+(bS)^c}$$



(c)

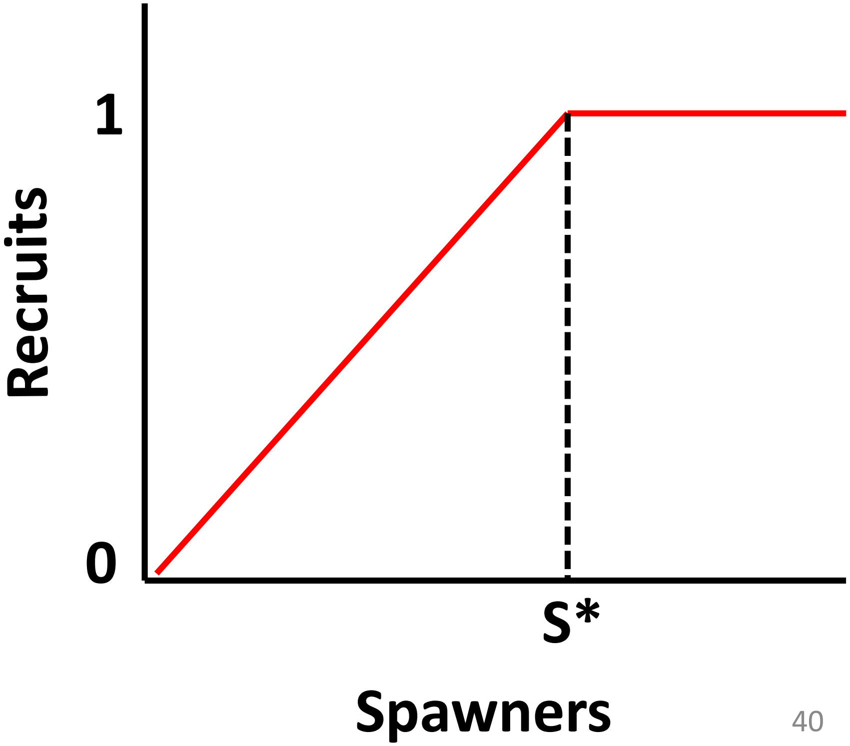
- a = productivity parameter (number of R per S at low S)
- b = parameter for degree of density dependence
- c = shape parameter
    - c<1: density-independent; c=1: Beverton-holt; c>1: Ricker shape
- Basic property: generalizing equation for other model shapes

Extra

5. Hockey stick model

- Segmented (change-point) regression
  - Slope a > 0 at the origin;
  - Slope a = 0 beyond pivotal spawner level, S*

$$R_t = \begin{cases} aS_t & \text{if } \; S_t < S^* \\ aS^* & \text{if } \; S_t \geq S^* \end{cases}$$



Stock-recruitment models

- Most common
  - Ricker
  - Beverton-Holt
- Others
  - Shepherd
  - Deriso-Schnute
  - Cushing
  - “Hockey-stick”
  - Unnormalized gamma density
  - Many others

Fitting Stock-recruitment models

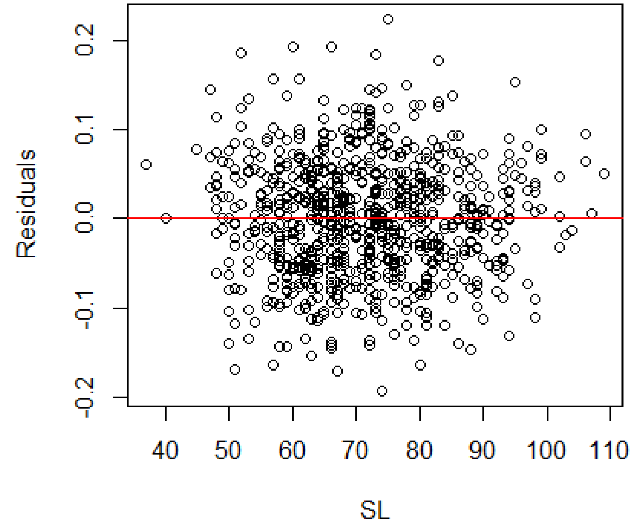
- Recommend using nonlinear regression
- Making the following assumptions:
  - No error in our estimate of S
  - Independent errors*
- But, model equation will depend on whether error is assumed to be additive or multiplicative

*see Quinn and Deriso 1999 (section 3.2) for “measurement error approach” and “autocorrelated errors” if these assumptions are grossly inappropriate

Additive vs. multiplicative error

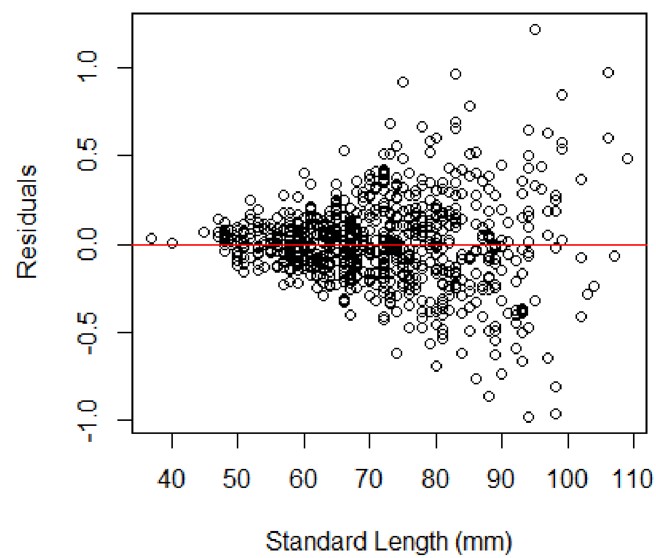
- Additive error
  - for a given model, the residuals will tend to have constant variance
  - Doesn't violate regression assumption of Homogeneity of Variance (HOV)

$$Y = f(X) + \varepsilon$$



- Multiplicative error
  - for a given model, the residuals will increase with higher values of X
  - This violates the HOV assumption for regression→ so log-transform equation to use regression

$$Y = f(X) \cdot e^{\varepsilon}$$



Log transforming to deal with multiplicative error

- Recall our example using W-L allometric model:

$$W = aL^{b}e^{\varepsilon}$$

Log-transform both sides of equation

$$\log(W) = \log(aL^{b}e^{\varepsilon})$$

Algebra

$$\log(W) = \log(a) + b \cdot \log(L) + \varepsilon$$

Now, the error is "additive" in our model, so we can use regression and OLS

Fitting Stock-recruitment models

- Our approach
  - use nonlinear regression (and OLS)
  - The equation we fit with nls() will depend on whether we assume additive or multiplicative error
  - *Multiplicative error is typically more appropriate for SR data*
- If using multiplicative error (& log transformation) → must use bias correction
  - Back transforming (i.e. exponentiating) estimates from log space introduces bias
  - *Bias correction*: multiply the back-transformed predicted values by a correction factor (CF), which depends on the standard error of the estimate (SEE; aka Residual SE)

Quinn and Deriso 1999, section 3.2

Fitting Stock-recruitment models

- Our approach

Equations to fit using nonlinear regression

More common!

	Additive error	**Multiplicative error (log of R=f(S)e$^{\varepsilon}$ → log(R)=log(f(S))+ ε)**
Density Independ.	$\boxed{\mathrm{R} = \mathrm{aS} + \varepsilon}$	$\boxed{\ln(\mathrm{R}) = \ln(aS) + \varepsilon}$
Beverton Holt	$\boxed{\mathrm{R} = \dfrac{\mathrm{aS}}{1 + \mathrm{bS}} + \varepsilon}$	$\boxed{\ln(\mathrm{R}) = \ln(aS/(1 + bS)) + \varepsilon}$
Ricker	$\boxed{\mathrm{R} = aSe^{-bS} + \varepsilon}$	$\boxed{\ln(\mathrm{R}) = \ln(aSe^{-bS}) + \varepsilon}$

Note: "ln" is "natural log", which in R, is written just as "log()"

Quinn and Deriso 1999, section 3.2

Fitting Stock-recruitment models

Example: Ricker Model with Multiplicative Error

$$\mathrm{R} = aSe^{-bS}e^{\varepsilon}$$

Log-transform

$$\ln(\mathrm{R}) = \ln(aSe^{-bS}) + \varepsilon$$

Fit using nonlinear regression, and estimate parameters:

$$\hat{a}, \hat{b}, \hat{\sigma}_{\varepsilon}^{2}$$

Back-transform & bias-correct

$$\hat{\mathrm{R}} = \hat{a}Se^{-\hat{b}S} \cdot e^{(\hat{\sigma}_{\varepsilon}^{2}/2)}$$

$$\mathrm{SEE} = \hat{\sigma}_{\varepsilon}^{2} = \sqrt{\frac{\Sigma(\mathrm{obs.} - \mathrm{pred.})^2}{\mathrm{n - (nos.\ parameters)}}}$$

Note: In R, SEE is the "residual standard error". Value stored in: summary(MyModel)$sigma

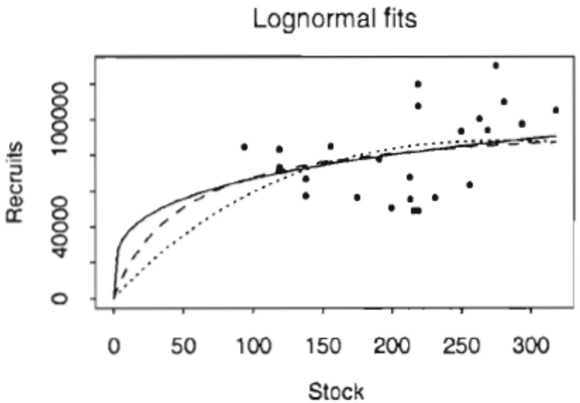
$$CF = e^{(\hat{\sigma}_{\varepsilon}^{2}/2)} = e^{(SEE/2)}$$

Sprugel 1983, Quinn and Deriso 1999

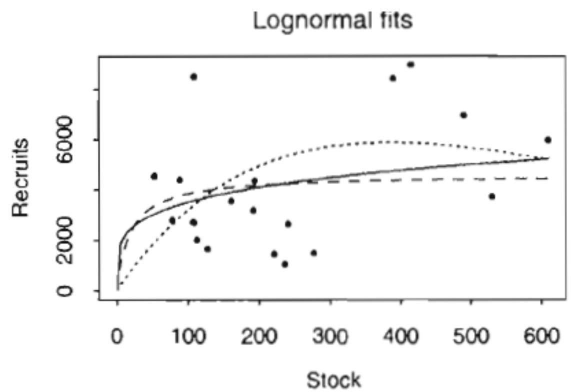
Examples of SR model fits

Thoughts? **Why are fits so poor?**

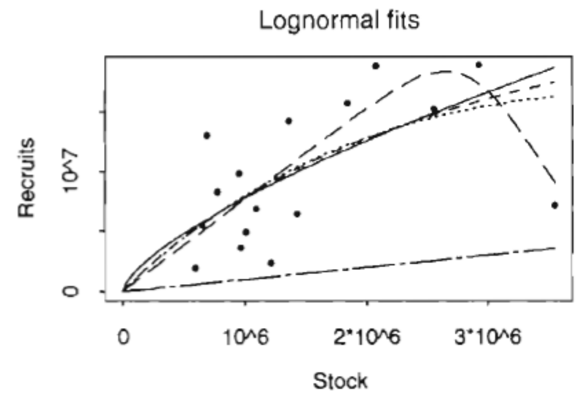
Yellowfin tuna
(Thunnus albacares)

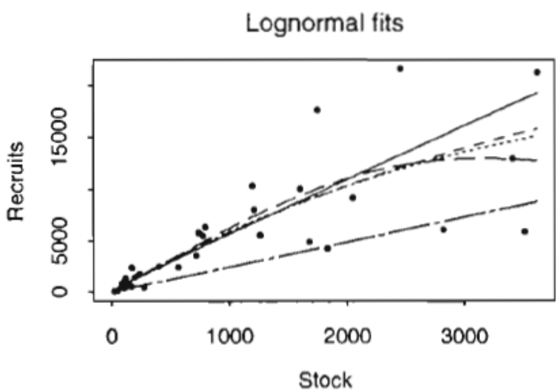


Petrale sole
(Eopsetta jordani)

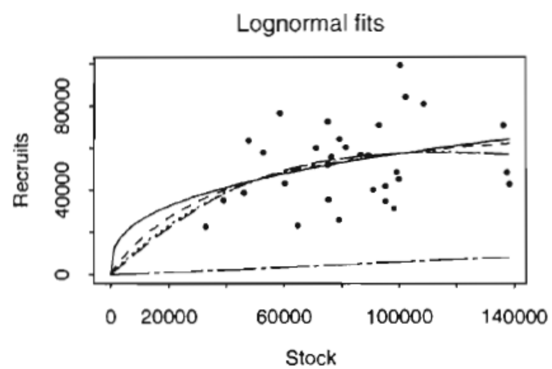


Pink salmon (Fraser river)
(Oncorhynchus gorbuscha)

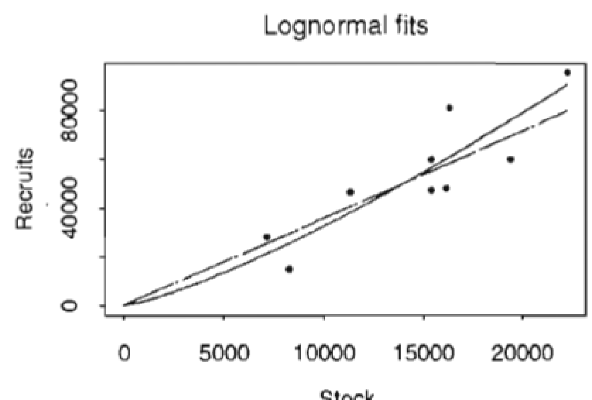




CA Sardine
(Sardinops sagax)



Atlantic cod (NAFO 3Ps)
(Gadus morhua)

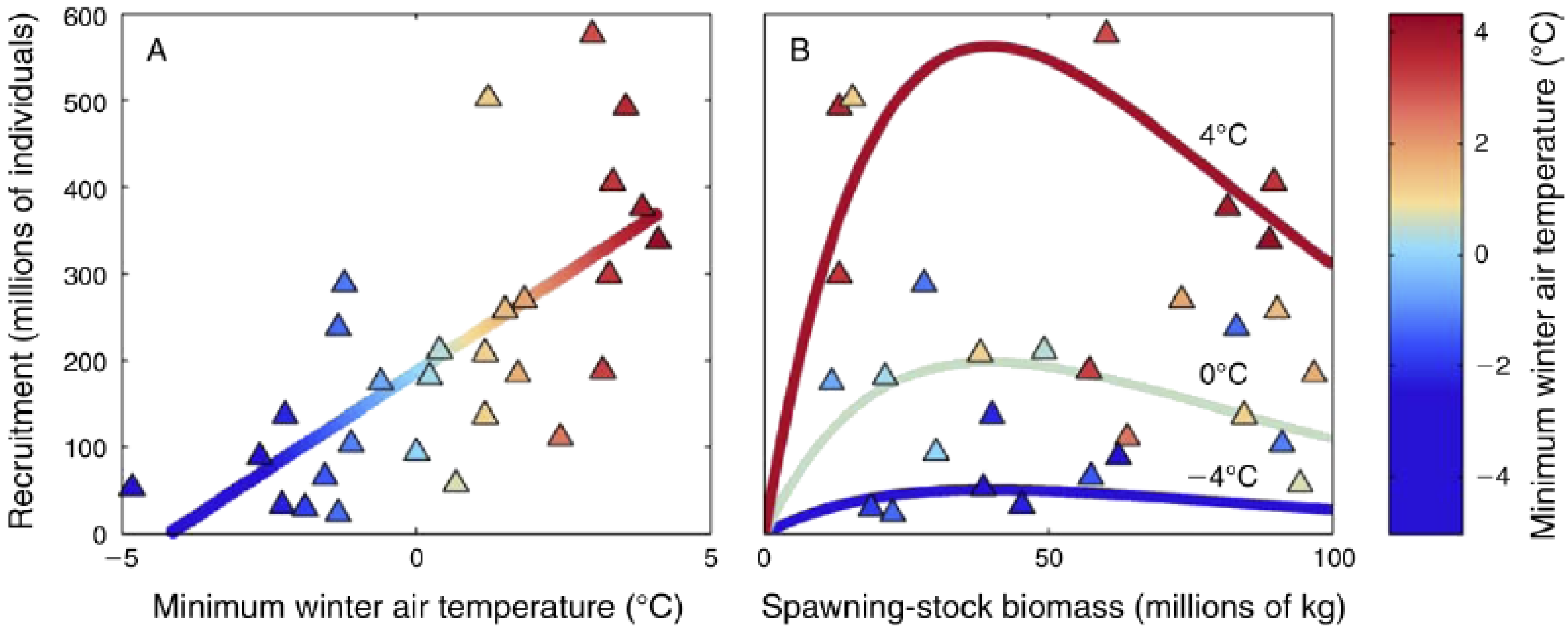


Summer flounder
(Paralichthys dentatus)

Myers et al. 1995

Some modifications to S-R Models

- Possible to build in environmental effects (e.g., temp for Atlantic Croaker)
- Account for error in S estimates
  - see Quinn & Deriso, Section 3.2.3



Hare et al. 2010

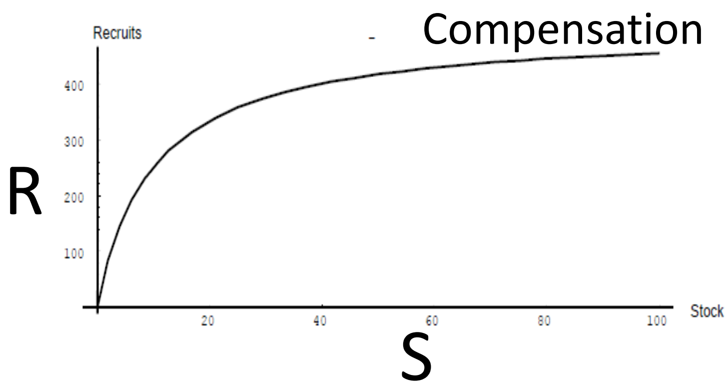
Summary

- Stock recruitment models
  - Relate the production of recruits to adult spawning stock
  - Critical for forecasting, assessing, and managing populations
  - Typically account for some type of density-dependence (DD)
  - Fits can be rather poor → lots of uncertainty
- Know definitions:
  - **Stock, Recruitment, Density dependence, Compensation**
- Stock recruitment models
  - **Beverton-Holt**
  - **Ricker**
  - Shepherd – Generalization of other models
  - "Hockey-stick"
  - Many others (Deriso-Schnute, Cushing, …)
- Fitting models
  - For us: assume multiplicative error (if have HOV problem) → log-transform model → use nonlinear regr. → back-transform & bias correct

Summary of BH and Ricker models

Beverton Holt

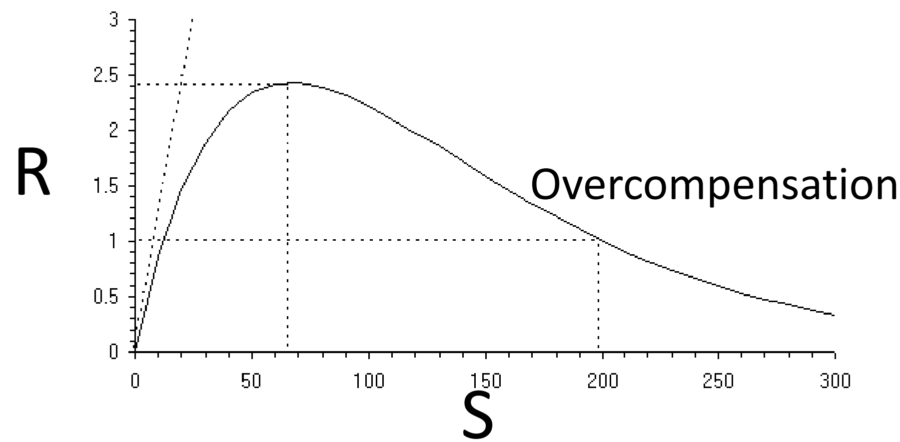
$$\boxed{R = \frac{aS}{1 + bS}}$$



- Density dependence
  - Acts via juvenile stage (*know examples*)
- Parameters
  - a = productivity parameter
  - b = density dependence
- Shape: asymptotic

Ricker

$$\boxed{R = aSe^{-bS}}$$



- Density dependence
  - Acts via adult stage (*know examples*)
- Parameters
  - a = productivity parameter
  - b = density dependence
- Shape: dome

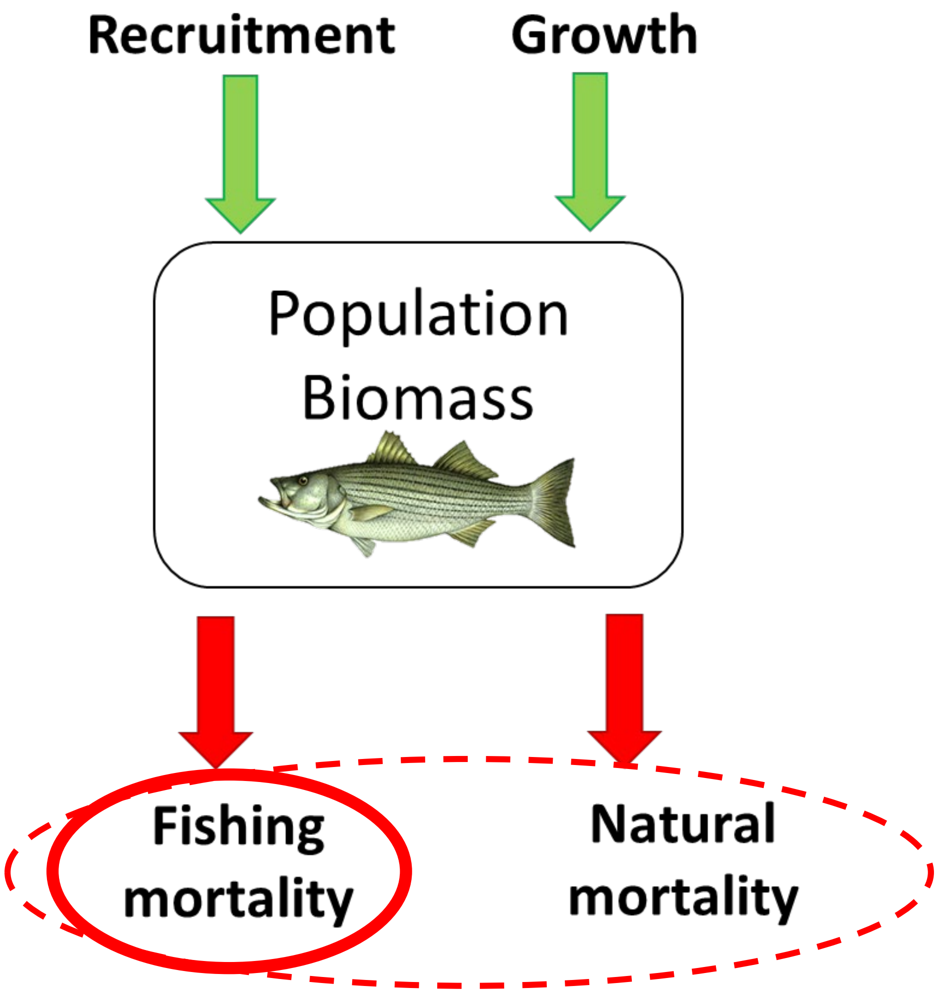
Effects of harvest on populations

Reading:

Jennings et al. 2001 – Chapter 7

Acknowledgment: M. Wilberg

Conceptual population model



Review of exponential mortality

- If following a single cohort, then we can track the mortality losses to the cohort

$$\frac{dN}{dt} = -ZN$$

$$\boxed{N_t = N_0 e^{-Zt}}$$

- Useful for modeling changes in a cohort over time; very common in fisheries science

Derive equations for the following:

S – annual survival (proportion living after one year)

A – the annual mortality rate (proportion dying per year)

Review of exponential mortality

Formulas for converting from instantaneous to annual rates:

- Z = total *instantaneous* mortality rate
- S = *annual* survival rate (proportion surviving in a year)

$$S = e^{-Z} \qquad\qquad Z = -\ln(S)$$

- A = *annual* mortality rate (proportion dying in a year)

$$A = 1 - e^{-Z}$$

Deterministic theory of fishing

- For fished populations, total mortality (Z) is composed of mortality due to harvest and mortality due to natural causes:

$$\boxed{\mathrm{Z} = F + M}$$

- F = instantaneous fishing mortality rate
- M = instantaneous natural mortality rate
  - Includes all non-fishing mortality

Deterministic theory of fishing

- Restate our exponential mortality equations:

$$\frac{dN}{dt} = -(F + M)N$$

$$\boxed{N_t = N_0 e^{-(F+M)t}}$$

Baranov Catch Equation

- How can we calculate how many fish die from fishing (or from natural causes)?
- **Baranov catch equation:**

$$\boxed{C_t = \frac{F}{Z}(1 - e^{-Z})N_t}$$

  - Where C_t = catch in year t
  - What does this mean in words?

Baranov Catch Equation

- How can we calculate how many fish die from fishing (or from natural causes)?
- **Baranov catch equation:**

$$\boxed{C_t = \frac{F}{Z}(1 - e^{-Z})N_t}$$

  - Where C_t = catch in year t
  - What does this mean in words?
    - Catch = (fraction of mortality due to fishing)*(Proportion dying)*abundance
    - Catch = (fraction of mortality due to fishing)*(total number of deaths)

  - What would the equation be for fraction of fish dying from natural causes?

Extra

Baranov Catch Equation

- Other forms:

$$C_t = \frac{F}{Z}(N_t - N_{t+1})$$

$$C_t = \frac{F}{Z}(N_t - N_t e^{-Z})$$

$$C_t = \frac{FA}{Z}N_t$$

$$C = F\bar{N}$$

N(bar) is the average abundance over the course of a year

Converting to annual rates

- Annual rates
  - **A** = proportion of population dying in 1 year
  - ***u*** = ***exploitation rate*** (e.g., proportion of population dying from fishing in 1 year)
  - ***v*** = proportion of population dying from natural causes in one year

$$A = u + v$$

Converting to annual rates

- <u>**Type II fishery** (continuous fishery)</u>
  - fishing and natural mortality continuous throughout the year; more common
  - Calculating *u* requires info on 2 of the following: F,M,Z

$$\boxed{u = \frac{F}{Z}(1 - e^{-Z}) = \frac{FA}{Z}}$$

$$\boxed{v = \frac{M}{Z}(1 - e^{-Z}) = \frac{MA}{Z}}$$

- <u>**Type I fishery** (pulse fishery)</u>
  - Fishing occurs in short pulse (days/weeks), natural mortality elsewhere; less common

$$u = 1 - e^{-F} \qquad v = 1 - e^{-M}$$

Catch in Difference Model

- Catch (C) is the product for the difference equation:

$$\boxed{C_t = u_t N_t}$$

We can incorporate Catch into the logistic model…

  - u_t = exploitation rate at time t
  - N_t = abundance at time t
  - (Assumes catch occurs in one pulse)

Note:
<u>Difference equation</u> – time is viewed as discrete points
<u>Differential equation</u> – time is used as a continuous variable

Review of Logistic Model

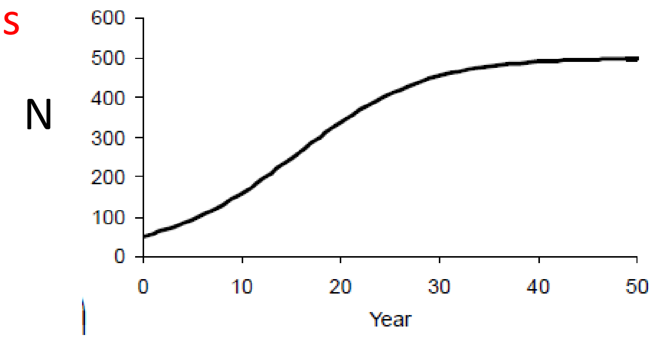
Sidenote: equations work with B or N

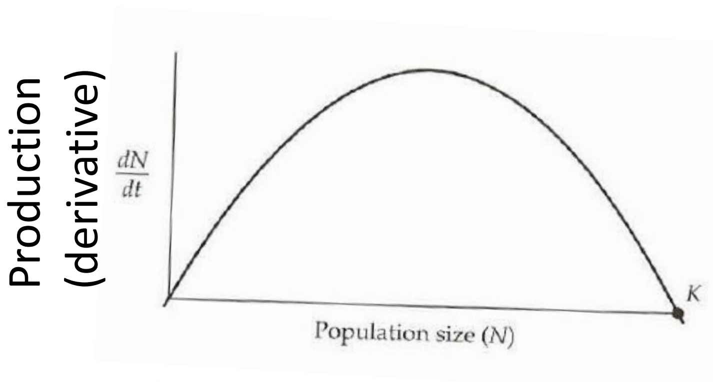
Logistic growth model

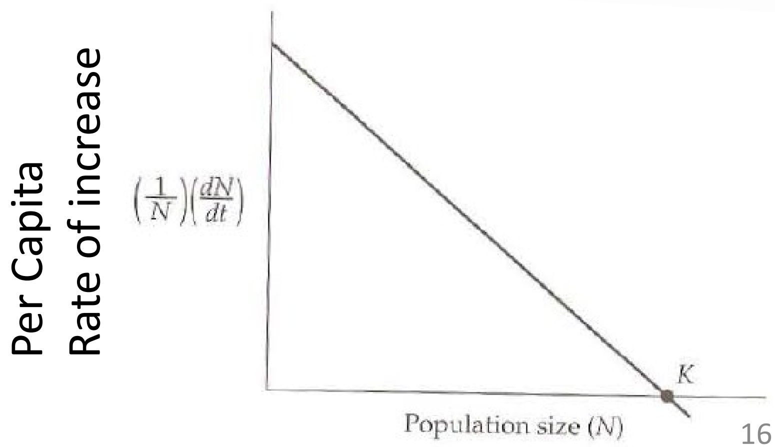
$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

Discrete version
$$N_{t+1} = N_t + rN_t\left(1 - \frac{N_t}{K}\right)$$

- Density-dependent
  - Per-capita growth rate varies with pop. size
- Key parameters:
  - r=intrinsic rate of increase
    K=carrying capacity
- Equilibrium = carrying capacity
- Shape: logistic ("S-shaped")
- For fisheries models:
  - Foundation for production models







Surplus production

$$B_{t+1} = B_t + rB_t\left(1 - \frac{B_t}{K}\right)$$

- Biomass$_{t+1}$ = Biomass$_t$ + **surplus production (SP)**

$$SP = rB_t\left(1 - \frac{B_t}{K}\right)$$

- **<u>Surplus production</u>** = the excess biomass generated above what is needed to maintain the population at its current biomass.

Logistic model **<u>with harvest</u>**

- Logistic growth model adjusted for fishing mortality (F) or exploitation rate (u)

$$\begin{pmatrix}\text{Biomass}\\ \text{next year}\end{pmatrix} = \begin{pmatrix}\text{Biomass}\\ \text{this year}\end{pmatrix} + \begin{pmatrix}\text{Surplus}\\ \text{production}\end{pmatrix} - (\text{Catch})$$

Catch

Discrete version
(difference equation):

$$B_{t+1} = B_t + rB_t\left(1 - \frac{B_t}{K}\right) - C_t$$

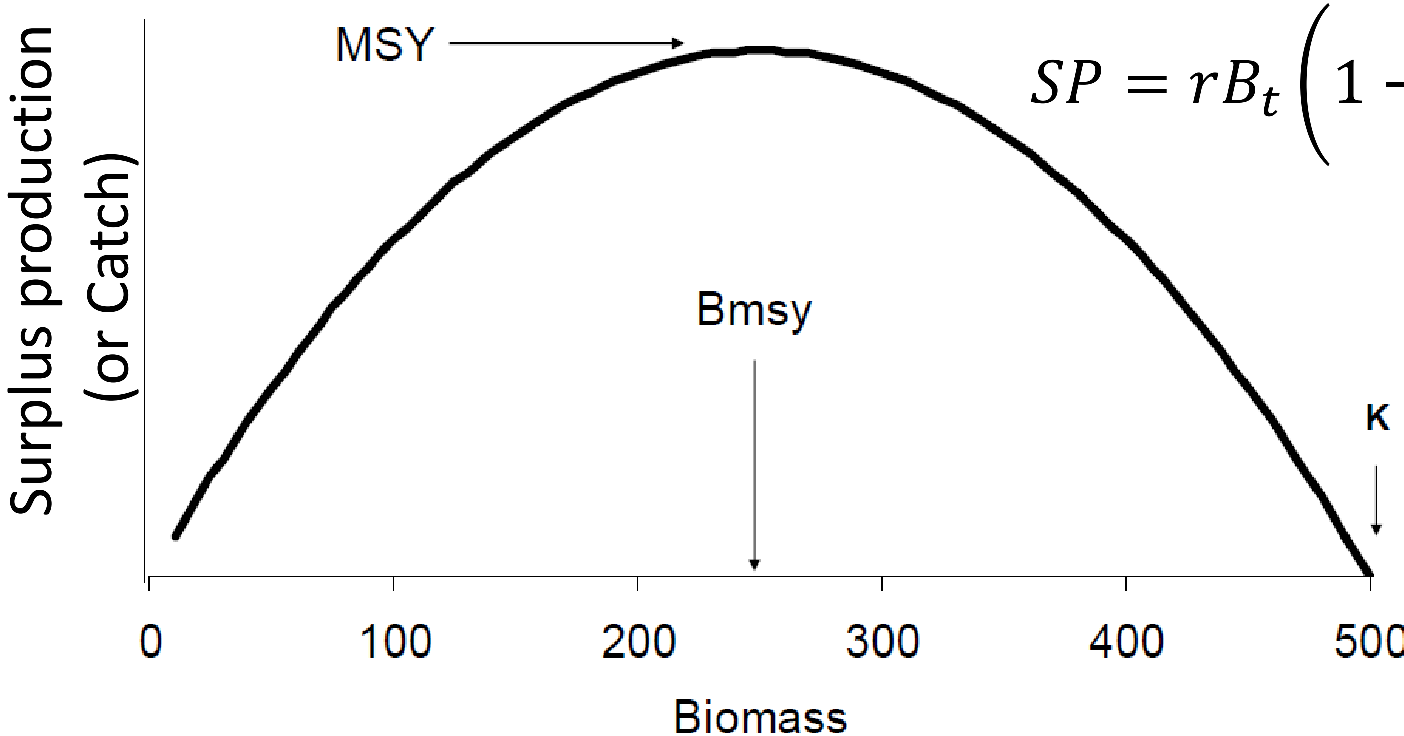
$$B_{t+1} = B_t + rB_t\left(1 - \frac{B_t}{K}\right) - uB_t$$

How can we ensure the Catch doesn't deplete a stock?

Surplus production and MSY

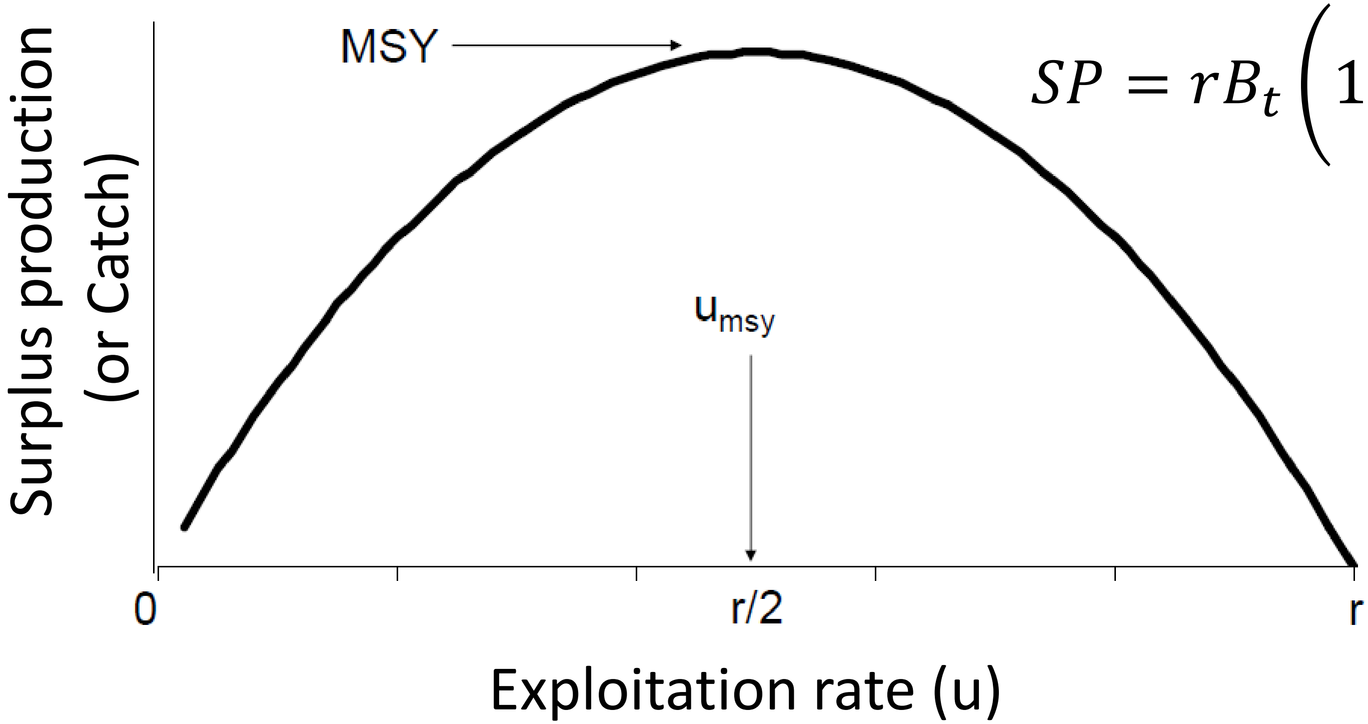
Set Catch to SP

$$SP = rB_t\left(1 - \frac{B_t}{K}\right)$$



- **MSY** = largest catch (yield) that can be supported by the population over an indefinite period
- **B_{msy}** = Biomass at which MSY is generated

Surplus production and MSY



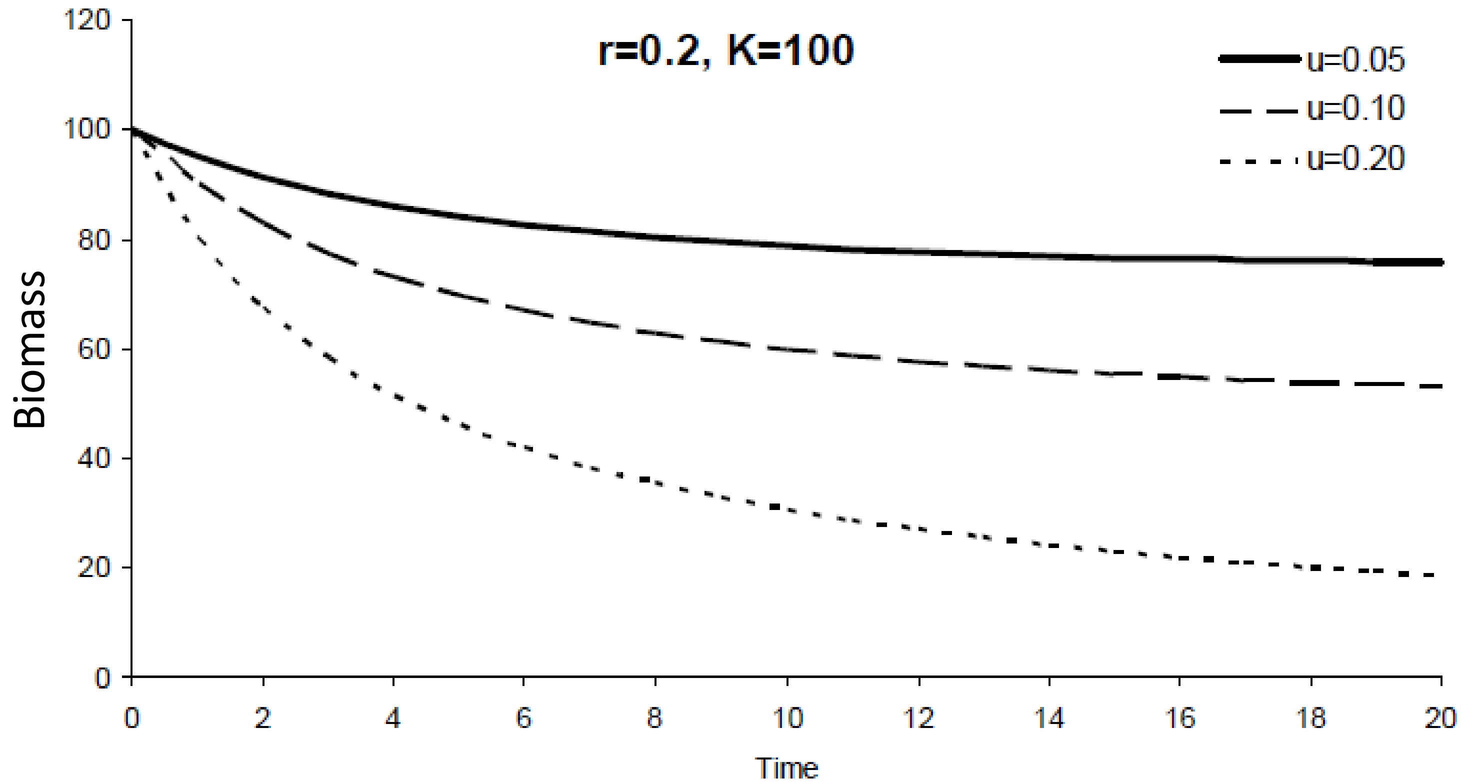
$$SP = rB_t\left(1 - \frac{B_t}{K}\right)$$

How hard should we fish to get MSY and keep the pop. at B_{MSY}?

- $\mathbf{u_{msy}}$ = exploitation rate that generates MSY
- $\mathbf{F_{msy}}$ = instantaneous fishing mortality rate that generates MSY

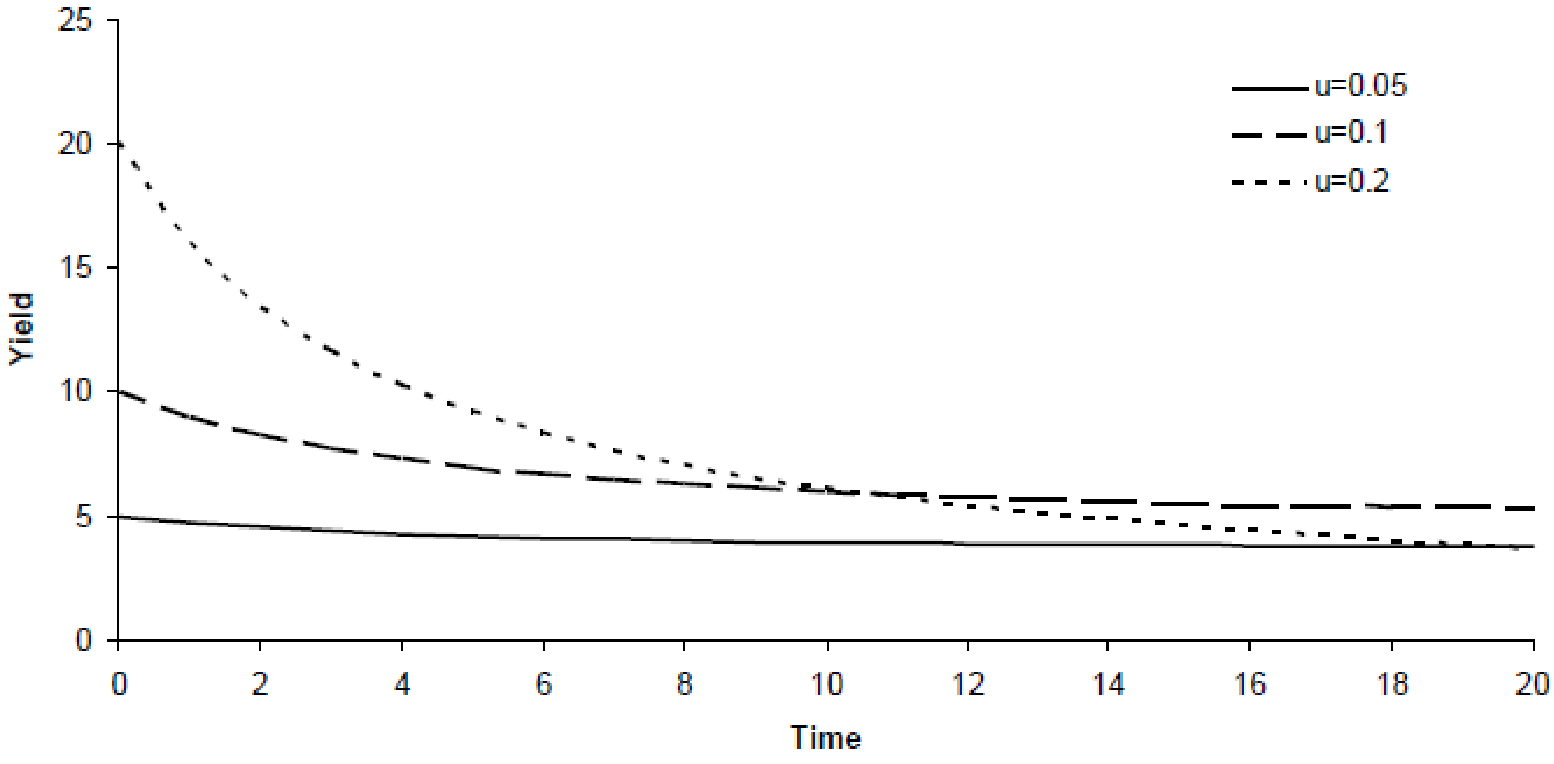
How to calculate MSY?

What happens to pop. through time at different exploitation rates (u)?



- What do you notice about the Biomass?

Yield over time



- What do you notice about the yield over time?
- →MSY is when we maximize this longterm yield

How to calculate MSY

$$B_{t+1} = B_t + rB_t\left(1 - \frac{B_t}{K}\right) - uB_t$$

$$u_{MSY} = \frac{r}{2}$$

$$B_{MSY} = \frac{K}{2}$$

$$\boxed{MSY = u_{MSY}B_{MSY} = \frac{rK}{4}}$$

Management Quantities

To be continued when we get into stock assessments…

Maximum surplus production (MSY)	$rK/4$
Stock size for MSY (B_{MSY})	$K/2$
Rate of exploitation at MSY (u_{MSY})	$r/2$
Effort required to achieve MSY (E_{MSY})	$r/2q$
Maximum rate of exploitation (u_{max})	r
Effort at maximum rate of exploitation (E_{max})	r/q

Summary 1

- Exponential mortality model
  - describes cohort abund. through time
  - Divide total inst. Mortality into: **Z=M+F**

$$\boxed{N_t = N_0 e^{-(F+M)t}}$$

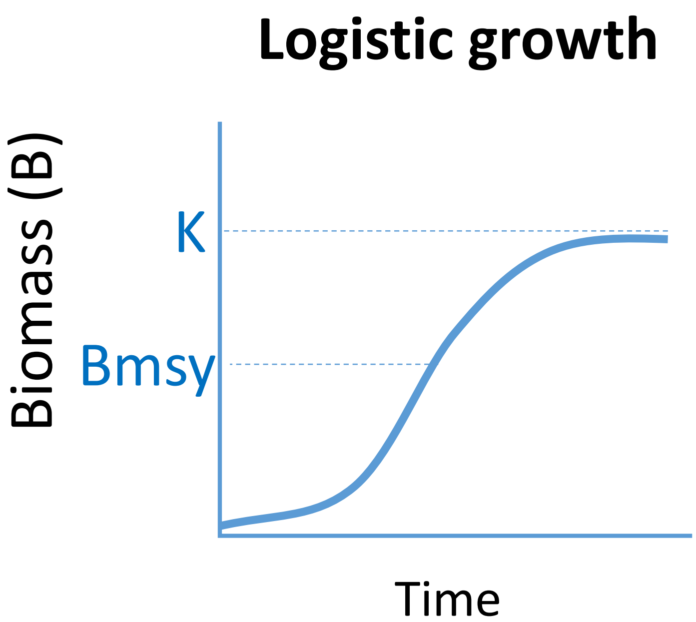
- Definitions
  - A, S, Z, u, v, F, and M
  - Exploitation rate (u)
  - Type I and II fisheries

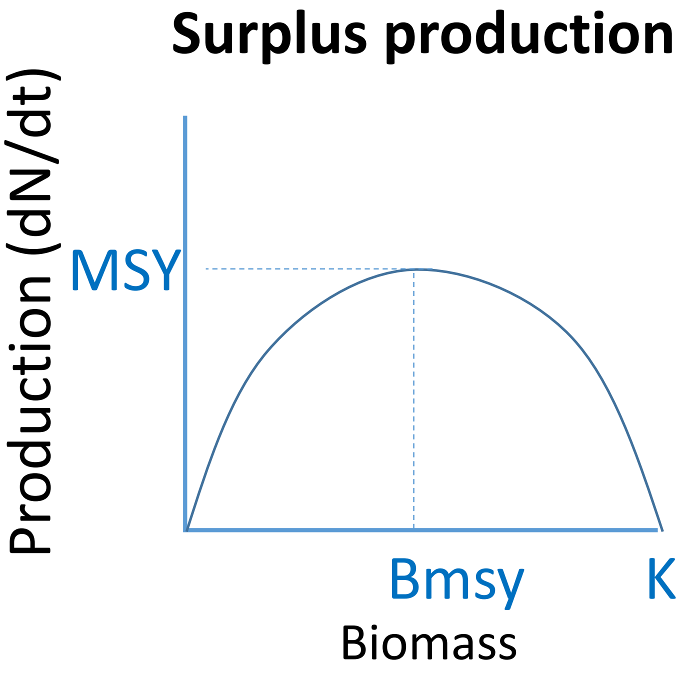
$$\boxed{C_t = \frac{F}{Z}(1 - e^{-Z})N_t}$$

- Baranov Catch Equation (many variants)
  - Catch = (fraction of mortality due to fishing)*(Proportion dying)*abundance
- Know MSY concepts!
  - how MSY related to logistic growth and surplus prod.
  - **MSY, B_{MSY}, u_{MSY}, F_{MSY}**

2 Logistic growth & surplus production

- Key concepts from the simple model:
  - K, r
  - Surplus production
  - MSY, Bmsy
  - u_{MSY} and F_{MSY}





3 Instantaneous & annual mortality

Instantaneous rate	Annual rate (proportion)	Relationship*
Z Total mortality	A Annual mortality	$A = 1-e^{-Z}$ $S = 1–A = e^{-Z}$ $Z = -\ln(S)$
F Fishing mortality	u Exploitation rate	$u = FA/Z$
M Natural mortality	v Prob. death from natural causes	$v = MA/Z$
$Z = F+M$	$A = u+v$	

For u or v, need 2 of the following: F, M, Z

*Note: u and v relationships assume Type II fishery that operates continuously throughout the year

Yield Per Recruit

Reading:

See Jennings et al. 2001, section 7.7

Haddon et al. 2011, section 2.8

Yield Per Recruit (YPR) models

Main questions:

- How hard should we fish to optimize harvest?
- What is the optimum age (or size) at first capture?
  - Lots of small fish or fewer big fish?

YPR is called a "**dynamic pool model**"

What are Dynamic Pool Models?

- Simple age-structured models
- Deterministic models
- Include mortality and growth models
- Widely used to develop reference points used to manage fisheries
- Common Types :
  - **Yield per recruit (YPR) → TODAY**
  - Spawning Stock Biomass per Recruit (SSB/R, S/R)
  - Egg per recruit

YPR – In class exercise

- See hand out
- Explanation of exercise
- Preliminary questions
  - If your goal is to maximize your yield, what fishing strategy do you think will be best: u=80%, 50%, 40%, 30%, 20%, 10%?

- Break up into groups and calculate the values. Then, fill in the summary table on the board.

Example

- *values rounded to 1 decimal place

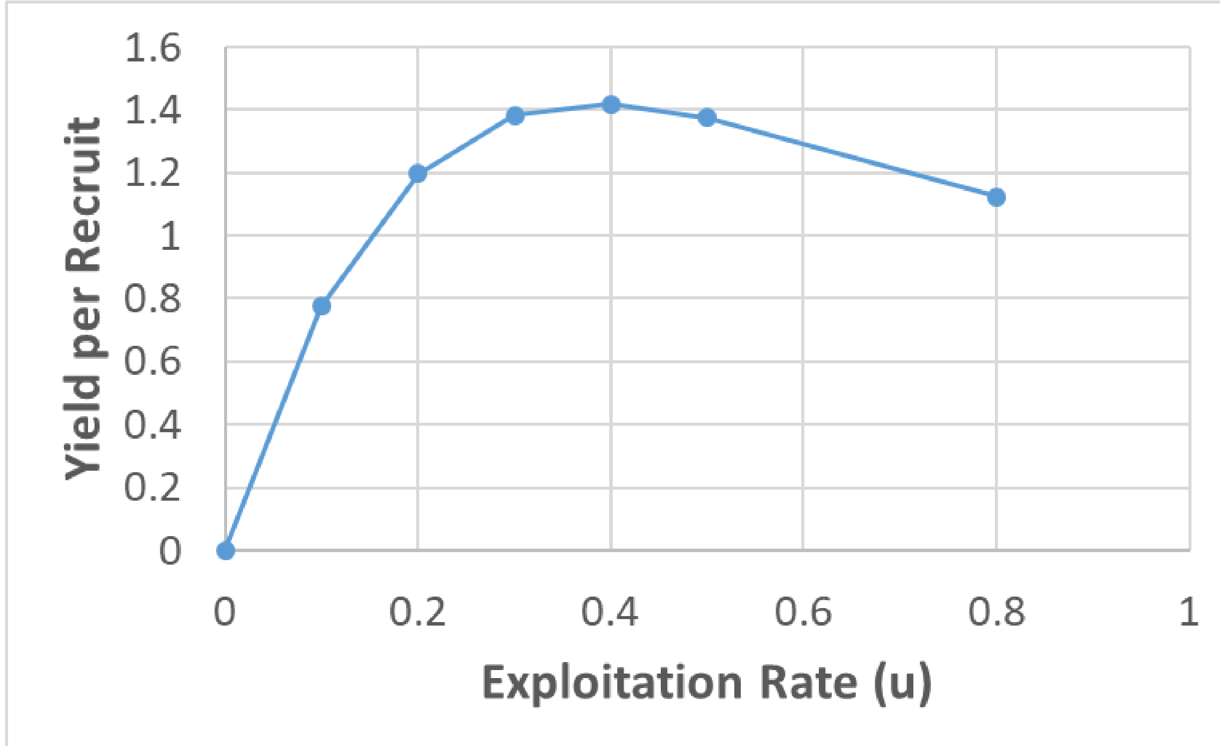
		Exploitation rate of 80% (u=0.8)		
Age (a)	Weight (W, in kg)	Stock Size (N)	Catch (C)	Catch Wt (kg)
1	--	1000	--	--
2	1	200 ($(1-u)*N_{a-1}$)	800 ($u*N_{a-1}$)	800 ($C*W_a$)
3	1.5	40.0	160.0	240.0
4	2	8.0	32.0	64.0
5	2.5	1.6	6.4	16.0
6	3	0.3	1.3	3.8
		TOTAL	1000	1124
			YPR	**1.124**

Write your YPR Results on the board

Fishing Scenario	Total Catch (# of fish)	Total Catch (kg)	Initial Recruits	Yield per recruit
u=0.8			1000	
u=0.5			1000	
u=0.4			1000	
u=0.3			1000	
u=0.2			1000	
u=0.1			1000	

- What do the results mean?
- Did the results match your prediction? Why or why not?
- What were the tradeoffs with the different strategies?
- What would happen if we introduced natural mortality?

YPR Results

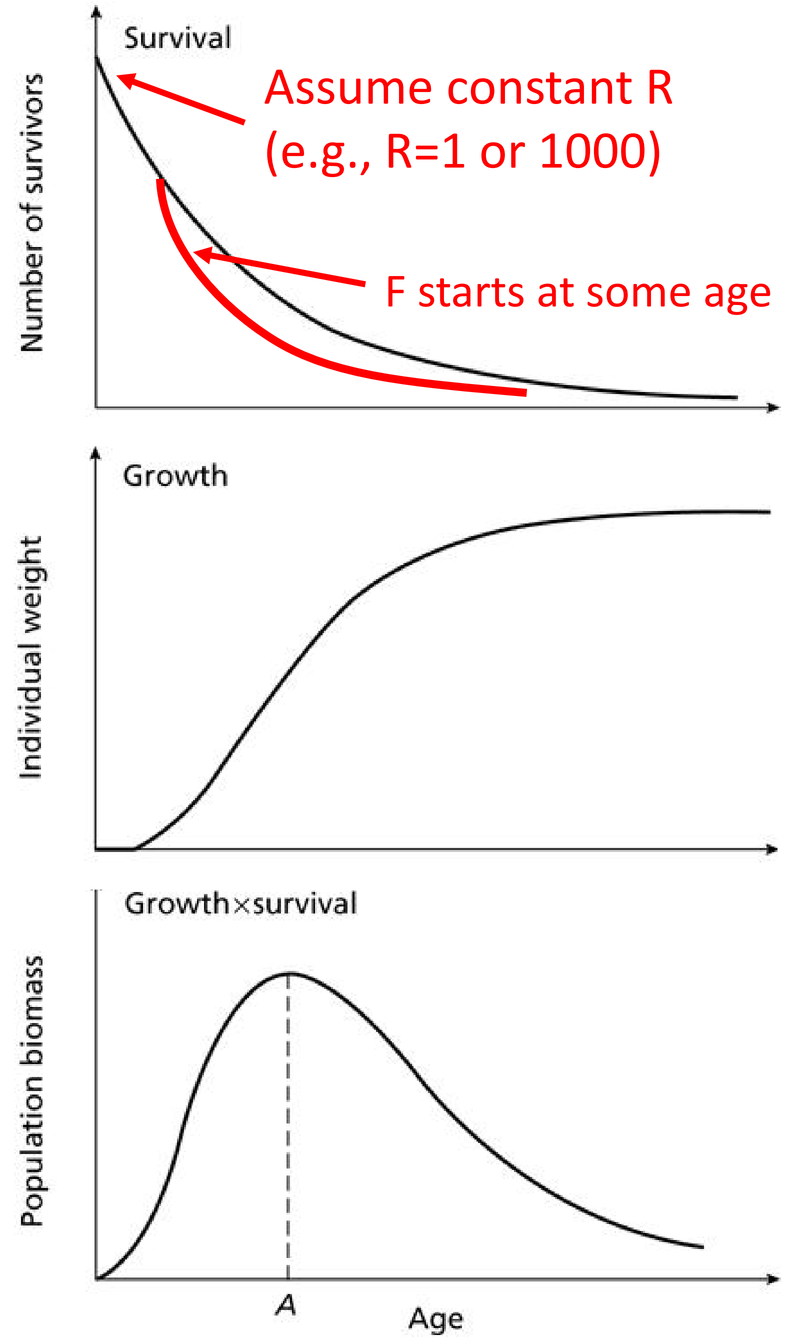


Fishing Scenario	Total Catch (# of fish)	Total Catch (kg)	Initial Recruits	Yield per recruit
u=0.1	410	776	1000	0.776
u=0.2	672	1198	1000	1.198
u=0.3	832	1382	1000	1.382
u=0.4	922	1420	1000	1.420
u=0.5	969	1375	1000	1.375
u=0.8	1000	1124	1000	1.124

Yield Per Recruit (YPR)

Fishing yield affected by a tradeoff between growth and mortality

- Basic Idea
  - Assume constant R
  - Simulate declines in N by age
  - Account for increases in W at age
  - Calculate Biomass at age
  - Estimate yield
    - given F and age at capture
  - Repeat for different fishing scenarios to find the optimum yield



Jennings, Kaiser, Reynolds
Marine Fisheries Ecology

YPR

		Exploitation rate of 80% (u=0.8)		
Age (a)	**Weight (W, in kg)**	**Stock Size (N)**	**Catch (C)**	**Catch Wt (kg)**
1	--	1000	--	--
2	1	200 ($(1-u)*N_{a-1}$)	800 ($u*N_{a-1}$)	800 ($C*W_a$)
3	1.5	40.0	160.0	240.0
4	2	8.0	32.0	64.0
5	2.5	1.6	6.4	16.0
6	3	0.3	1.3	3.8
		TOTAL	1000	1124
			YPR	1.124

How do we express this as an equation?

Yield Per Recruit (YPR)

- Model biomass and yield of a cohort over time, accounting for growth and mortality:

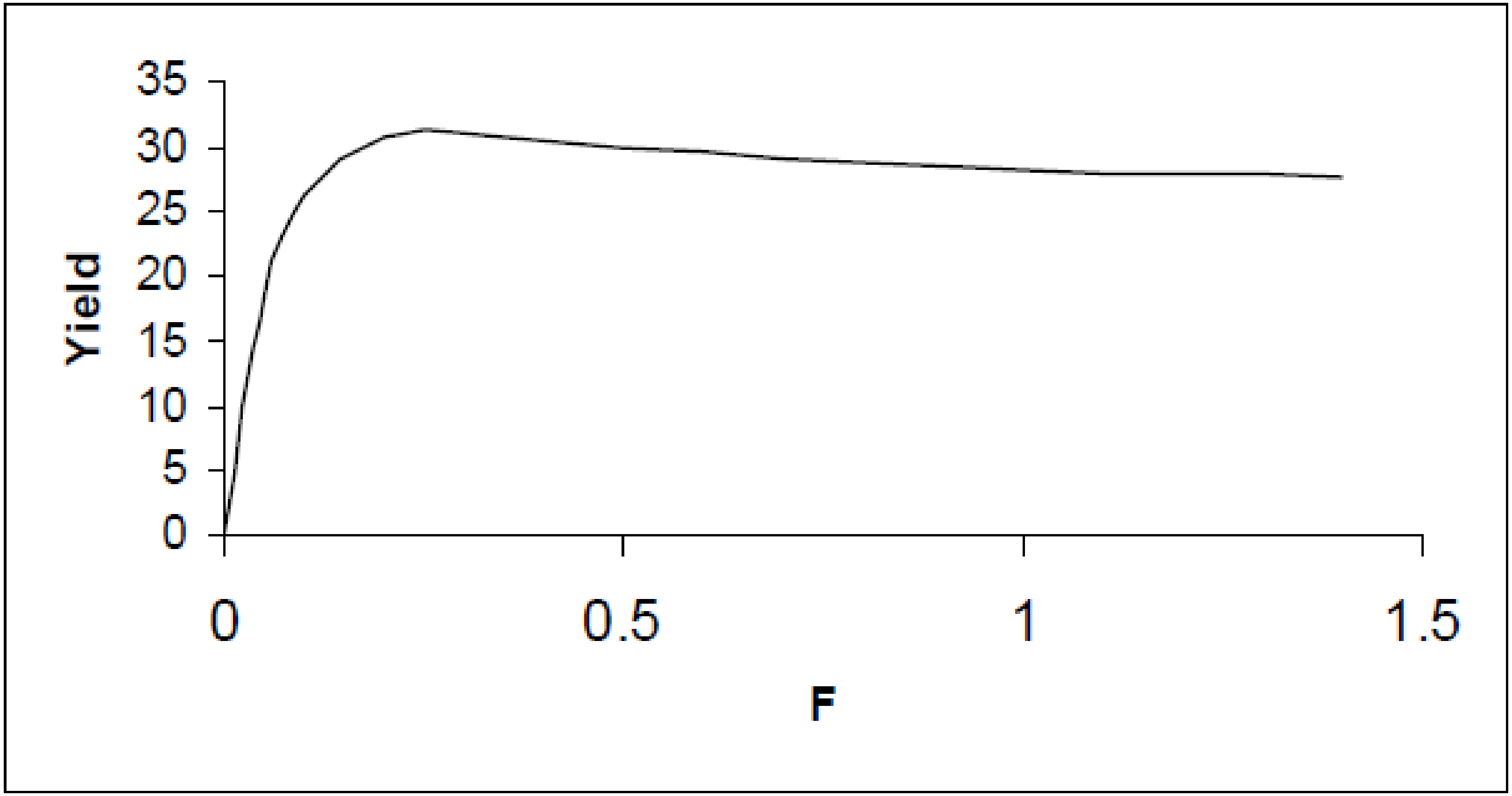
$$Y = \sum_{t=t_c}^{t_{max}} \frac{F_t}{Z_t}(1 - e^{-Z_t})N_t W_t \qquad Y = \sum_{t=t_c}^{t_{max}} u_t N_t W_t$$

  - Y = yield (in biomass) per fish recruited to fishery
  - t_c = age at first capture
  - t_{max} = max age that is captured
  - F_t = instantaneous fishing mortality rate at age t
  - Z_t = instantaneous total mortality rate
  - u_t = exploitation rate at age t (proportion harvested)
  - N_t = abundance at age t
  - W_t = weight at age t

How do we model Nt and Wt?

Yield (per recruit) vs. F

- Yield per recruit curve
  - Model is run for different values of F (e.g. our 6 scenarios)
- What do the different parts of this curve represent?

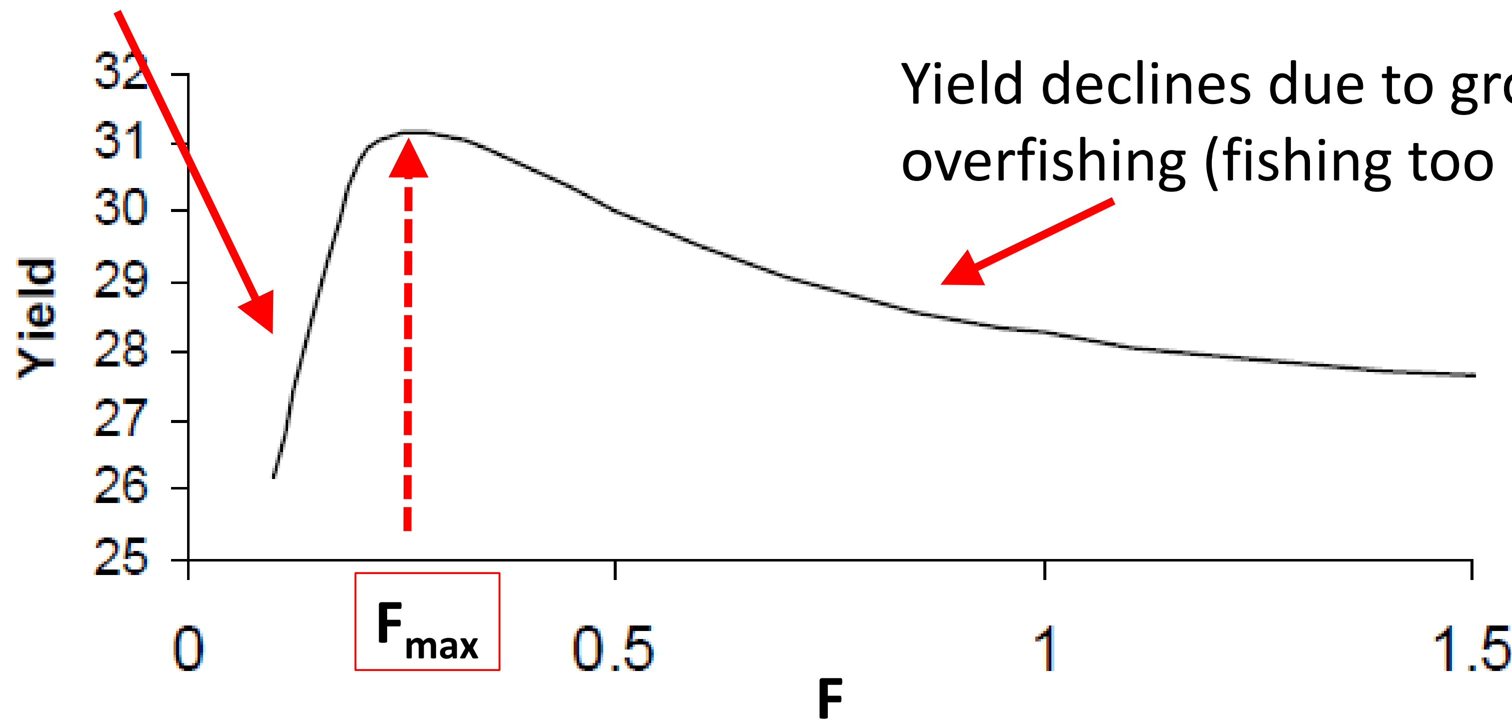


*Not uncommon to have a flat top to these curves

Yield (per recruit) vs. F

Growth overfishing – when fishing rates are too high, preventing the maximum yield being caught (b/c fish are prevented from growing to larger sizes)

Yield declines due to low fishing (not fishing hard enough)



Yield declines due to growth overfishing (fishing too hard)

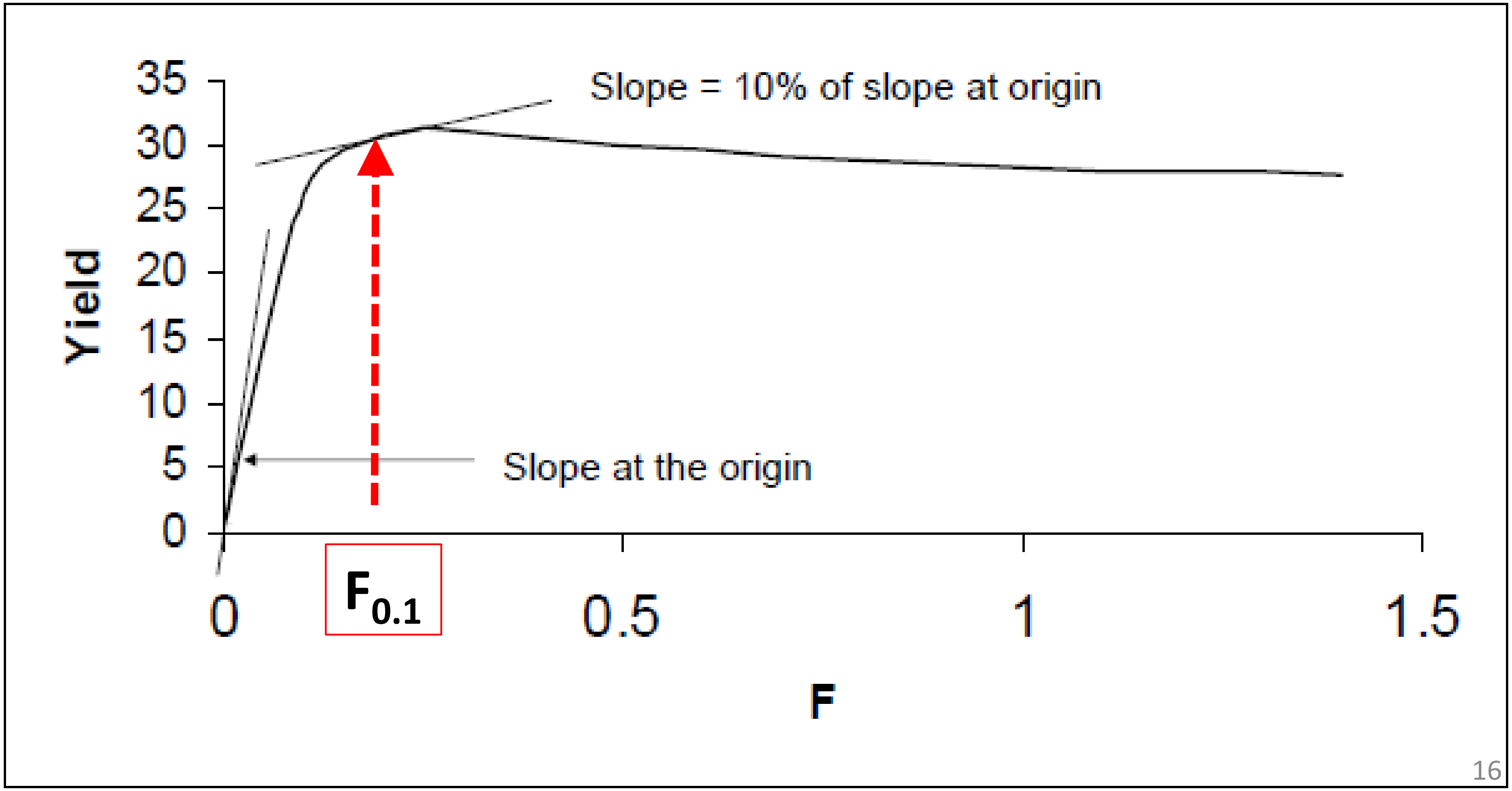
*Change in Y-axis scale for viewing

F_{max} = Fishing mortality rate that maximizes the yield per recruit

Use optimization function to find Fmax

$F_{0.1}$

- **$F_{0.1}$ = F where slope of YPR curve is 10% of slope at the origin**
- Conservative alternative to Fmax

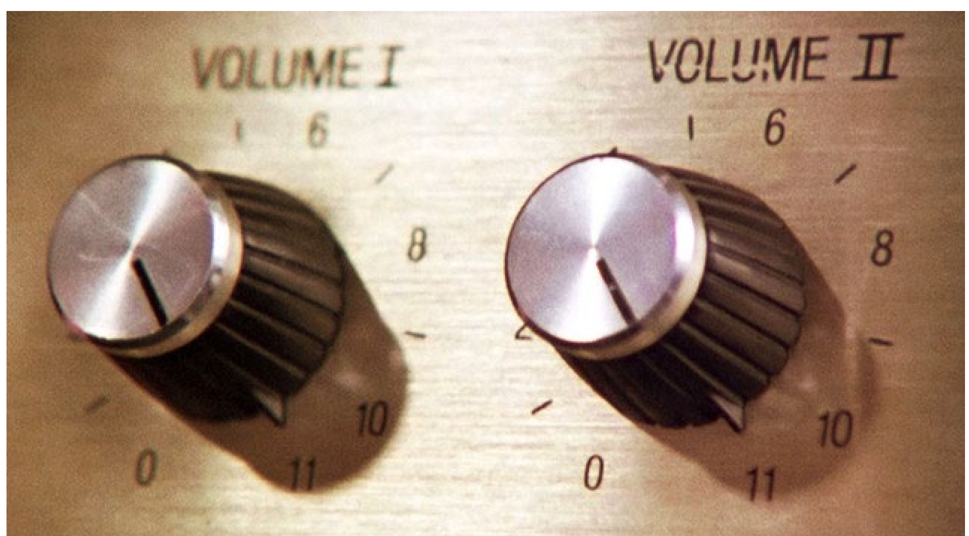


Comments on F_{max} and $F_{0.1}$

- **Biological reference points** – quantitative, biologically-based metric for a stock to inform management
- $F_{0.1}$
  - Ad hoc with no theoretical justification
  - More conservative
  - Greater stock stability with slightly lower yields
- Fmax ≠ Fmsy (don't confuse the two…)
- Usage
  - Still used somewhat in Europe and Canada
  - In US: Spawning Stock Biomass per Recruit favored

Yield per recruit models

- 2 main control variables:
  - Age at first capture, t_c (ie, selectivity pattern)
  - Fishing Mortality, F
- Try different combinations of t_c and F to maximize yield



Age at first capture (t_c)

- Assume constant M, but F only starts at t_c
- But we can change how abruptly F increases by changing the "selectivity"

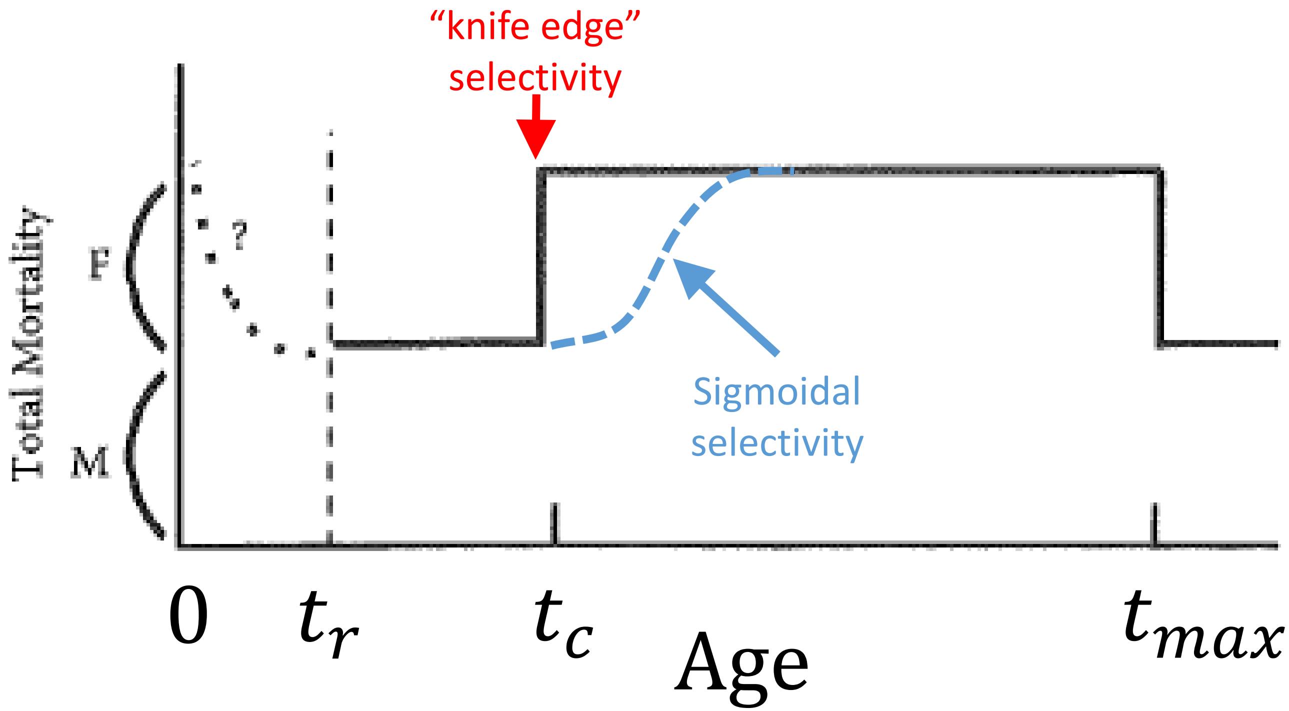
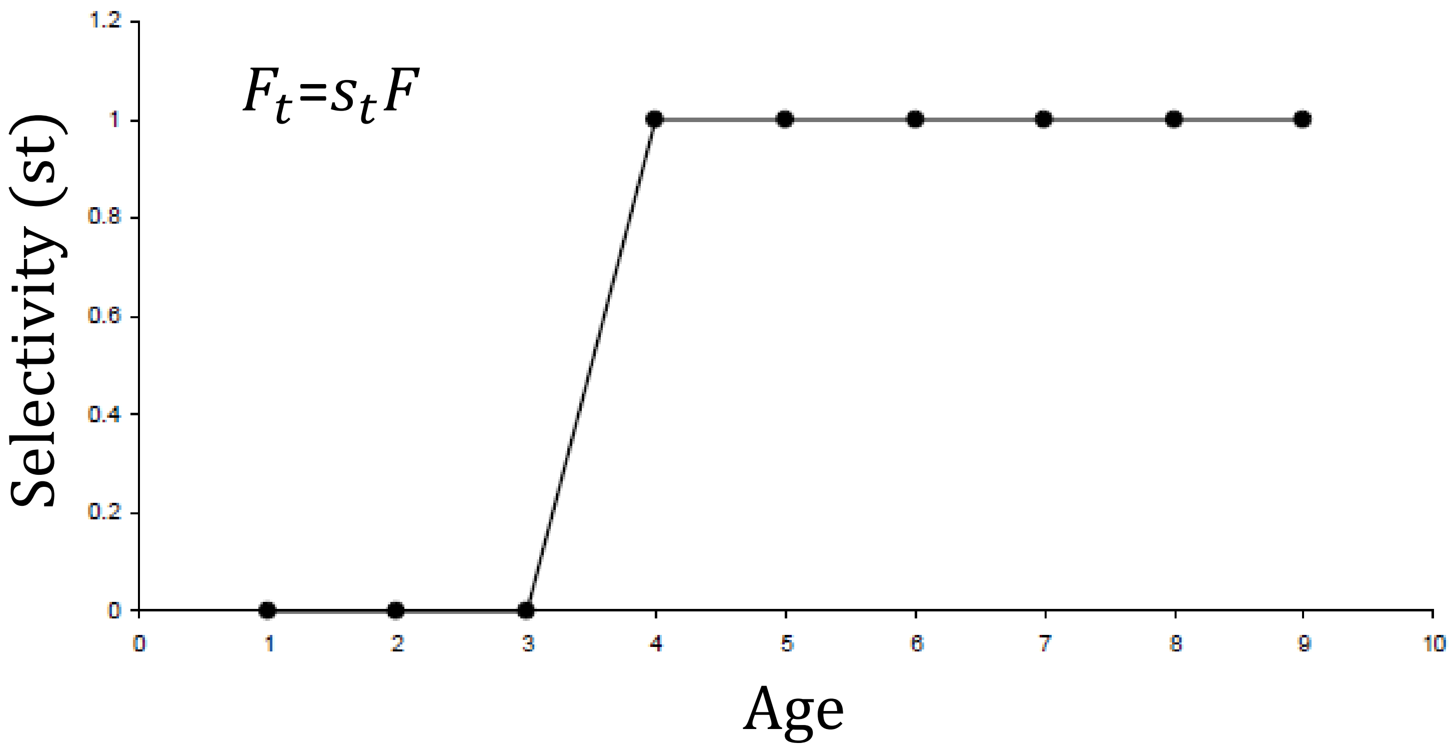


FIGURE 2.14
A diagram of the assumptions made in simple yield-per-recruit analysis. Here t_r is the age at recruitment, t_c is the age at first possible capture, and t_{max} is the age at which fish cease to be vulnerable to fishing. The model does not consider the dynamics of individuals younger than t_r years of age but simply assumes that there are R recruits of this age entering the stock each breeding period. M and F are the constant instantaneous rates of natural and fishing mortality, respectively. Knife-edge selection is shown by the vertical rise in mortality at t_c.

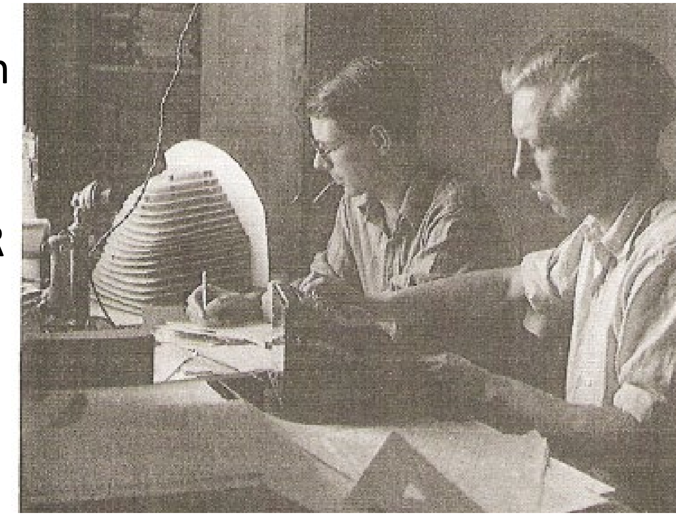
Fishing mortality by age

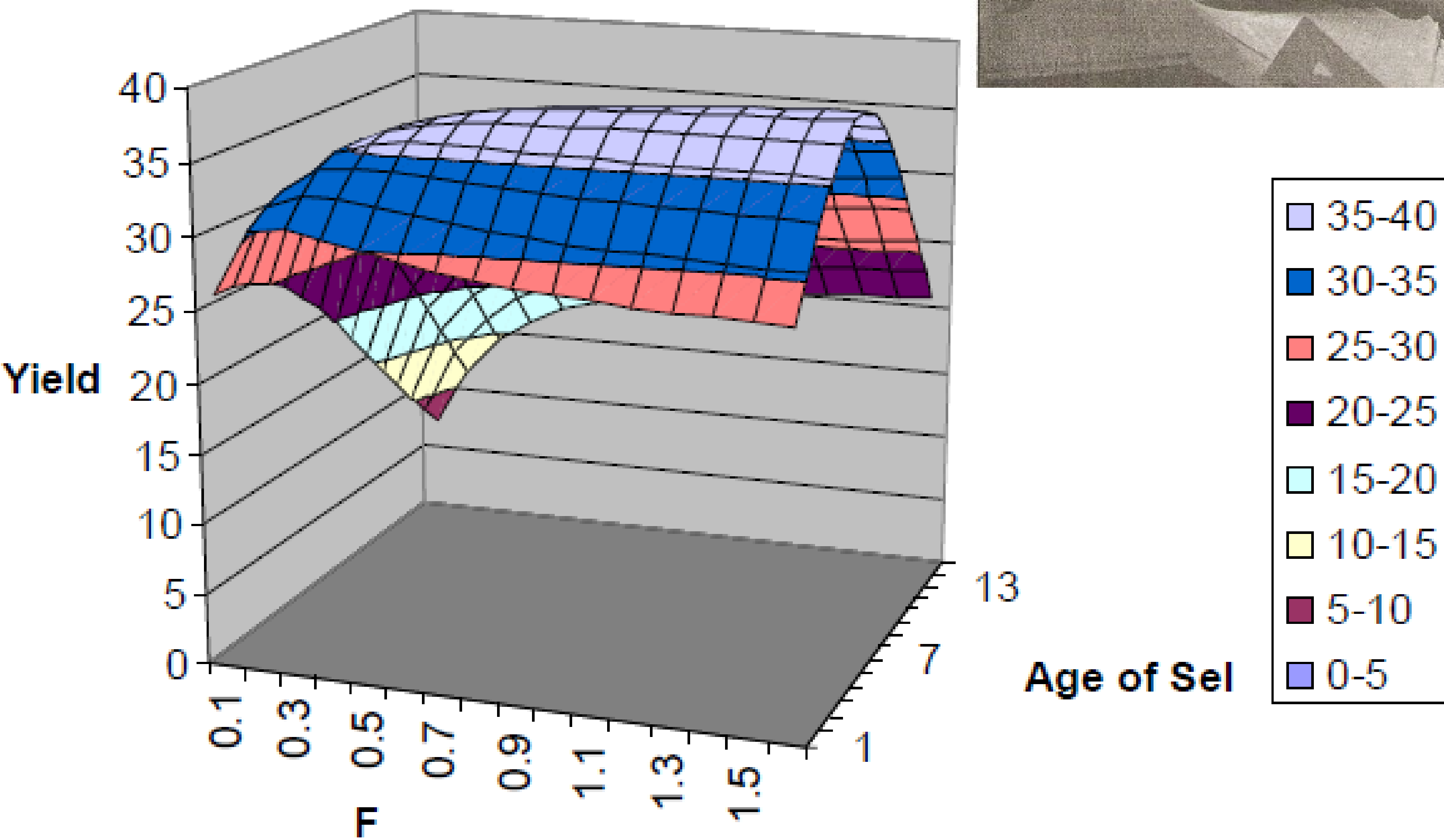
- **Age-specific selectivity** (s_t) = the probability of being vulnerable to fishing at age t (or some size)
  - This is known as "partial recruitment" for the **ypr()** function in R
- Simple model assumes knife-edge selectivity (but can have different shapes):



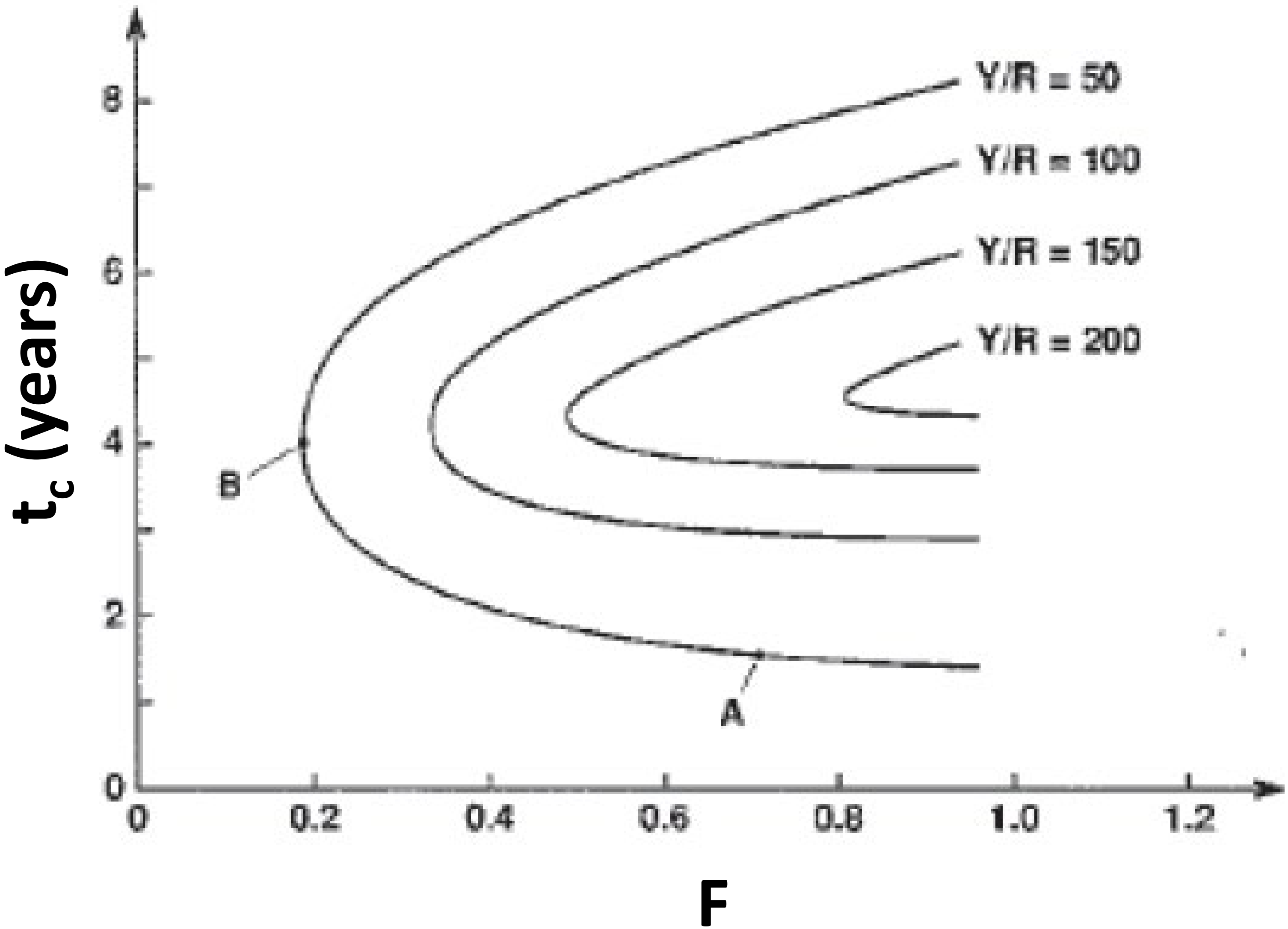
YPR Surface

Side: Beverton & Holt with cardboard cutout for YPR surface





YPR contour plot



Assumptions of YPR Models

- Constant recruitment (ie age-structure is at equilibrium or steady state)
- Constant mortality schedule
- Constant growth schedule

"I'm afarid we need someone with a little more backbone for this position."

Uses of YPR Models

- Develop reference points (F_{max} , $F_{0.1}$) for fisheries management (constant fishing rate harvest strategy)
  - Prevent growth overfishing
  - Evaluates things we can control (F, t_c)
- Determine optimal age at first capture
  - Gear regulations (e.g., mesh sizes, hook sizes, escapement holes, minimum sizes)
  - Season start date

Limitations of YPR Models

In what setting may some of these be more appropriate?

- Constant assumptions
  - Assumes steady state (e.g., constant R & age structure)
  - Constant mortality, selection, size at age schedules
  - This means no density dependence
- Used as a relative measure
  - does not estimate total yield which is of greater interest to fishers. (Total yield requires info on amount of R)
- Ignores effects of harvest on recruitment (no SR model)
  - No indication of whether F_{max} or $F_{0.1}$ is sustainable!
    - E.g., Some could predict t_c for juveniles; some predict infinite Fmax

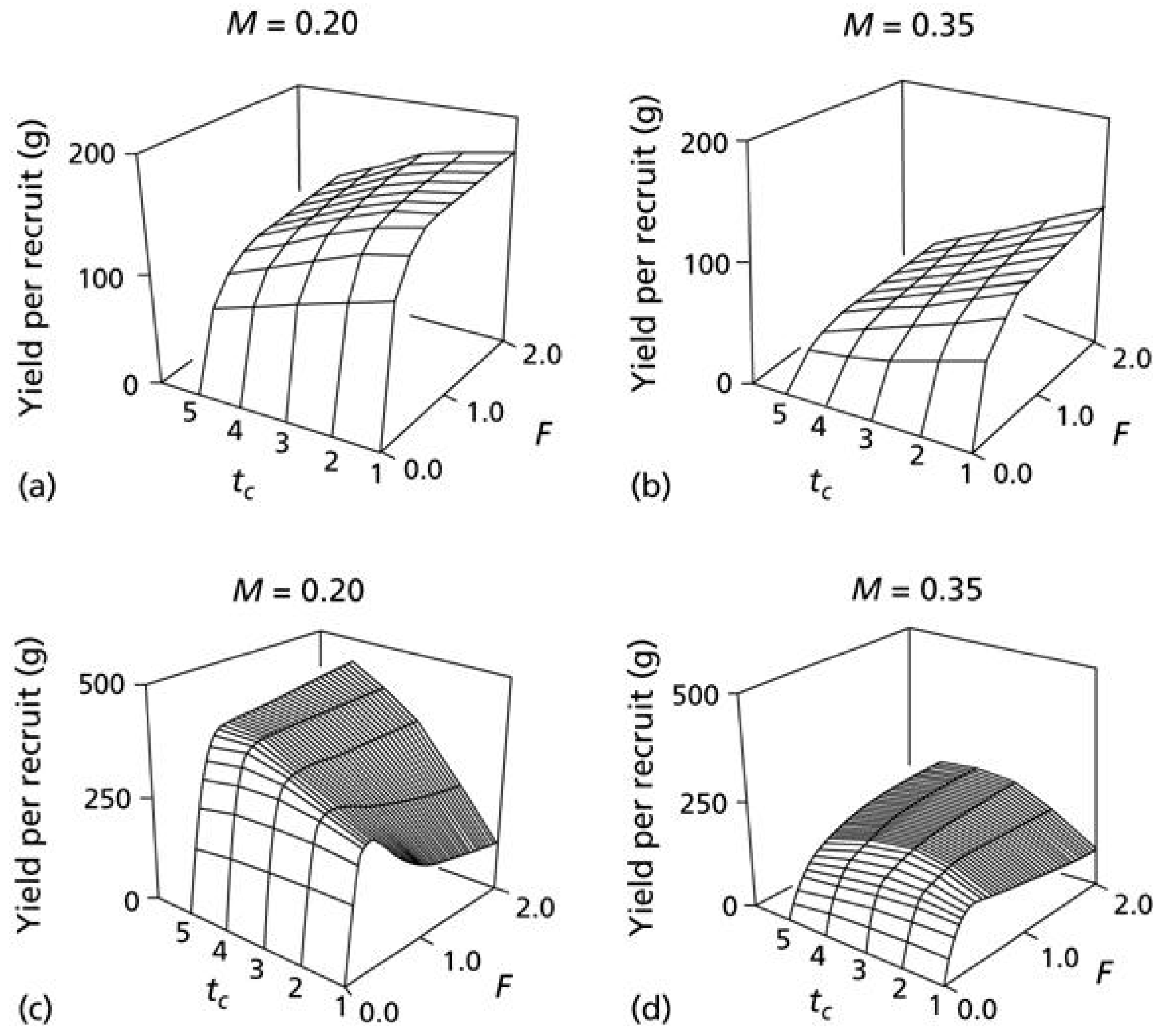
YPR example – Atlantic croaker

- Effect of region and M assumption?

- CB
  - Tc=age-2
  - Z~0.6
  - 1988-91 data
- NC
  - Tc=age-1
  - Z~1.3
  - 1979-81 data

Chesapeake Bay

North Carolina



Jennings, Kaiser, Reynolds

Summary - YPR

$$Y = \sum_{t=t_c}^{t_{max}} \frac{F_t}{Z_t}(1 - e^{-Z_t})N_t W_t$$
$$or\ (u_t N_t W_t)$$

- **Yield per Recruit (YPR)**
  - **YPR** = estimate of the weight that each recruit contributes to fishery harvest over its lifetime (on average)
  - Goal of YPR models: evaluate fishing rates & age at capture (t_c) to maximize YPR
- YPR models account for:
  - age-structure
  - growth and mortality
- YPR models don't account for:
  - Recruitment, changes in growth or mortality schedules, or density dependence
- Provide info and reference points for management
  - Deals with **<u>growth overfishing</u>**
  - $\mathbf{F_{max}}$ and $\mathbf{F_{0.1}}$ (or $\mathbf{u_{max}}$ and $\mathbf{u_{0.1}}$)
  - Provides info on optimal **age at first capture**

know definitions & graphical representation

Spawning per recruit

(Spawning stock biomass per recruit)

Reading:

See Jennings et al. 2001, section 7.7

Haddon et al. 2011, section 2.8

Acknowledgement: M. Wilberg

Announcements

- Still grading take home exams
- Grad students, please review the requirements for your project (see canvas) → will discuss later this week
  - Synopsis Due: March 18

Spawning per recruit

- Developed to address YPR limitations of not accounting for maturation and recruitment
- Goal is to determine the lifetime spawning potential of a recruit and how much it is reduced by fishing
- Spawning: typically SSB or eggs
- Used to calculate reference points for fisheries management.

With Fishing

			Mortality Rate			
Age	Wt (kg)	Maturity at age	Natural (v)	Fishing (u)	Stock Size (#)	SSB (kg)
					1000	
1	0.1	0.00	0.2	0		
2	0.5	0.00	0.2	0.3		
3	1	0.50	0.2	0.3		
4	1.5	1.00	0.2	0.3		
5	2	1.00	0.2	0.3		
					TOTAL SSB	
					SSB/R	

Calculate the following (no calculator should be needed):

- stock sizes based on the mortality rates
- Spawning stock biomass (SSB)
- Total SSB
- SSB per recruit (divide by initial # of recruits)

With Fishing

			Mortality Rate			
Age	Wt (kg)	Maturity at age	Natural (v)	Fishing (u)	Stock Size (#)	SSB (kg)
					1000	
1	0.1	0.00	0.2	0	800	0
2	0.5	0.00	0.2	0.3	400	0
3	1	0.50	0.2	0.3	200	100
4	1.5	1.00	0.2	0.3	100	150
5	2	1.00	0.2	0.3	50	100
					TOTAL SSB	**350**
					SSB/R	**0.35**

Spawning Stock Biomass per Recruit

- Calculation of SSB/R
- Info needed:
  - Maturity at age (Mat_a)
  - Mortality at age (Z_a)
  - Size at age (W_a)

$$SSB/R = \sum_a Mat_a W_a e^{\sum_{j=\min a}^{a-1} -Z_j}$$

- What are the components of this equation?

What is the maximum SSB/R that is possible?

With Fishing

			Mortality Rate			
Age	Wt (kg)	Maturity at age	Natural (v)	Fishing (u)	Stock Size (#)	SSB (kg)
					1000	
1	0.1	0.00	0.2	0	800	0
2	0.5	0.00	0.2	0.3	400	0
3	1	0.50	0.2	0.3	200	100
4	1.5	1.00	0.2	0.3	100	150
5	2	1.00	0.2	0.3	50	100
					TOTAL SSB	**350**
					SSB/R	**0.35**

NO Fishing

			Mortality Rate			
Age	Wt (kg)	Maturity at age	Natural (v)	Fishing (u)	Stock Size (#)	SSB (kg)
					1000	
1	0.1	0.00	0.2	0	800	0
2	0.5	0.00	0.2	0	640	0
3	1	0.50	0.2	0	512	256
4	1.5	1.00	0.2	0	410	614
5	2	1.00	0.2	0	328	655
					TOTAL SSB	**1526**
					SSB/R	**1.53**

Maximum spawning potential (MSP)

Spawning Stock Biomass per Recruit

- **Maximum spawning potential (MSP)**
  - **biomass of reproductively mature fish per recruit *in the absence of fishing***
  - **IE, the max. expected lifetime contribution of a recruit to the spawning stock biomass**
  - **SSB/R with no fishing**

No Fishing (MSP):

$$SSB/R_{F=0} = \sum_{a} Mat_a W_a e^{\sum_{j=\min a}^{a-1} -M_j}$$

- How much are we reducing the lifetime spawning potential by fishing?
  - Calculate ratio of values → "**spawning potential ratio (SPR)**"

With Fishing

			Mortality Rate			
Age	Wt (kg)	Maturity at age	Natural (v)	Fishing (u)	Stock Size (#)	SSB (kg)
					1000	
1	0.1	0.00	0.2	0	800	0
2	0.5	0.00	0.2	0.3	400	0
3	1	0.50	0.2	0.3	200	100
4	1.5	1.00	0.2	0.3	100	150
5	2	1.00	0.2	0.3	50	100
					TOTAL SSB	**350**
					SSB/R	**0.35**

NO Fishing

			Mortality Rate			
Age	Wt (kg)	Maturity at age	Natural (v)	Fishing (u)	Stock Size (#)	SSB (kg)
					1000	
1	0.1	0.00	0.2	0	800	0
2	0.5	0.00	0.2	0	640	0
3	1	0.50	0.2	0	512	256
4	1.5	1.00	0.2	0	410	614
5	2	1.00	0.2	0	328	655
					TOTAL SSB	**1526**
					SSB/R	**1.53**

Maximum spawning potential (MSP)

$$SPR = \frac{SSB/R_{Fishing}}{SSB/R_{NoFishing}} = \frac{0.35}{1.53} = 0.23\ (or\ 23\%)$$

Spawning Stock Biomass per Recruit

- **Spawning potential ratio (SPR)**
  - **the fraction of MSP that is achieved at different fishing mortality rates**
  - **i.e., Ratio of fished SSB/R to unfished SSB/R**

With Fishing:

$$SSB/R = \sum_a Mat_a W_a e^{\sum_{j=\min a}^{a-1} -Z_j}$$

No Fishing (MSP):

$$SSB/R_{F=0} = \sum_a Mat_a W_a e^{\sum_{j=\min a}^{a-1} -M_j}$$

Spawning potential ratio

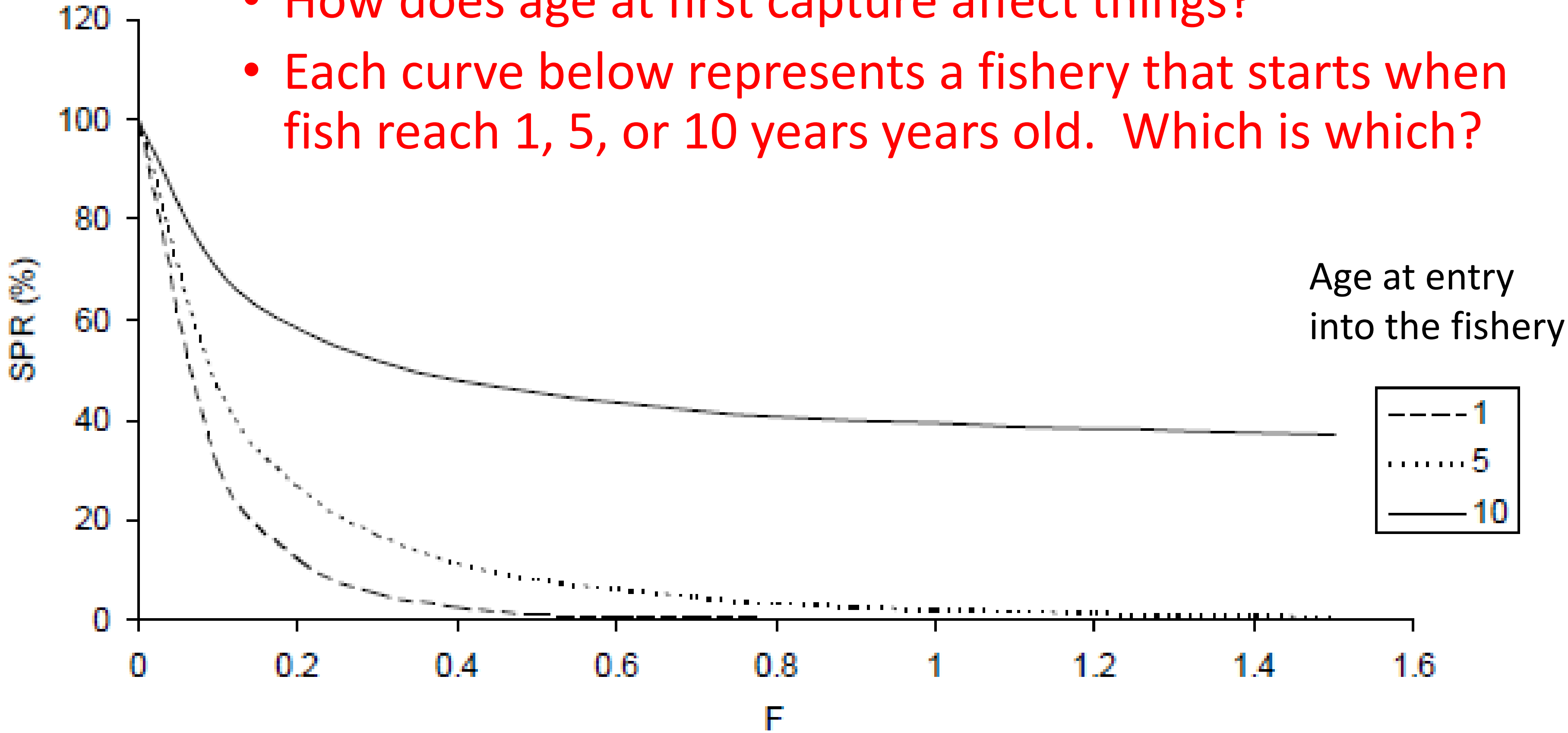
$$\text{SPR} = \frac{SSB/R}{SSB/R_{F=0}} \times 100$$

Terminology/acronyms

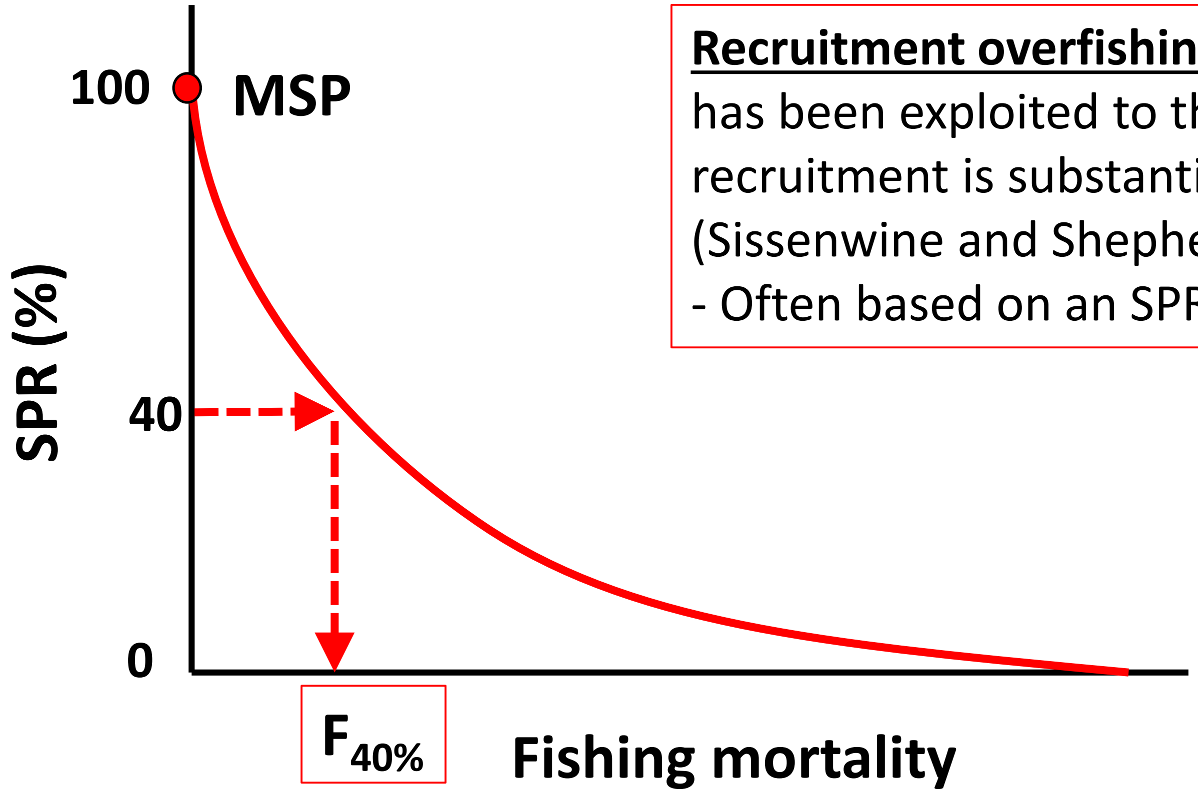
- Spawning stock biomass per recruit
  - Also: spawners per recruit, spawning per recruit
  - Acronyms: SSB/R or S/R or SSBR or SPR (e.g., Gabriel and Mace 1999, Rochet 2000)
  - **Maximum spawning potential (MSP)** – SSB/R with no fishing
- Ratio of SSB/R to unfished SSB/R
  - **Spawning potential ratio** (SPR, or %SPR)
  - %MSP
  - Compensation ratio (CR) (Goodyear 1993)

Spawning potential ratio (SPR)

- Example curves
  - How does age at first capture affect things?
  - Each curve below represents a fishery that starts when fish reach 1, 5, or 10 years years old. Which is which?



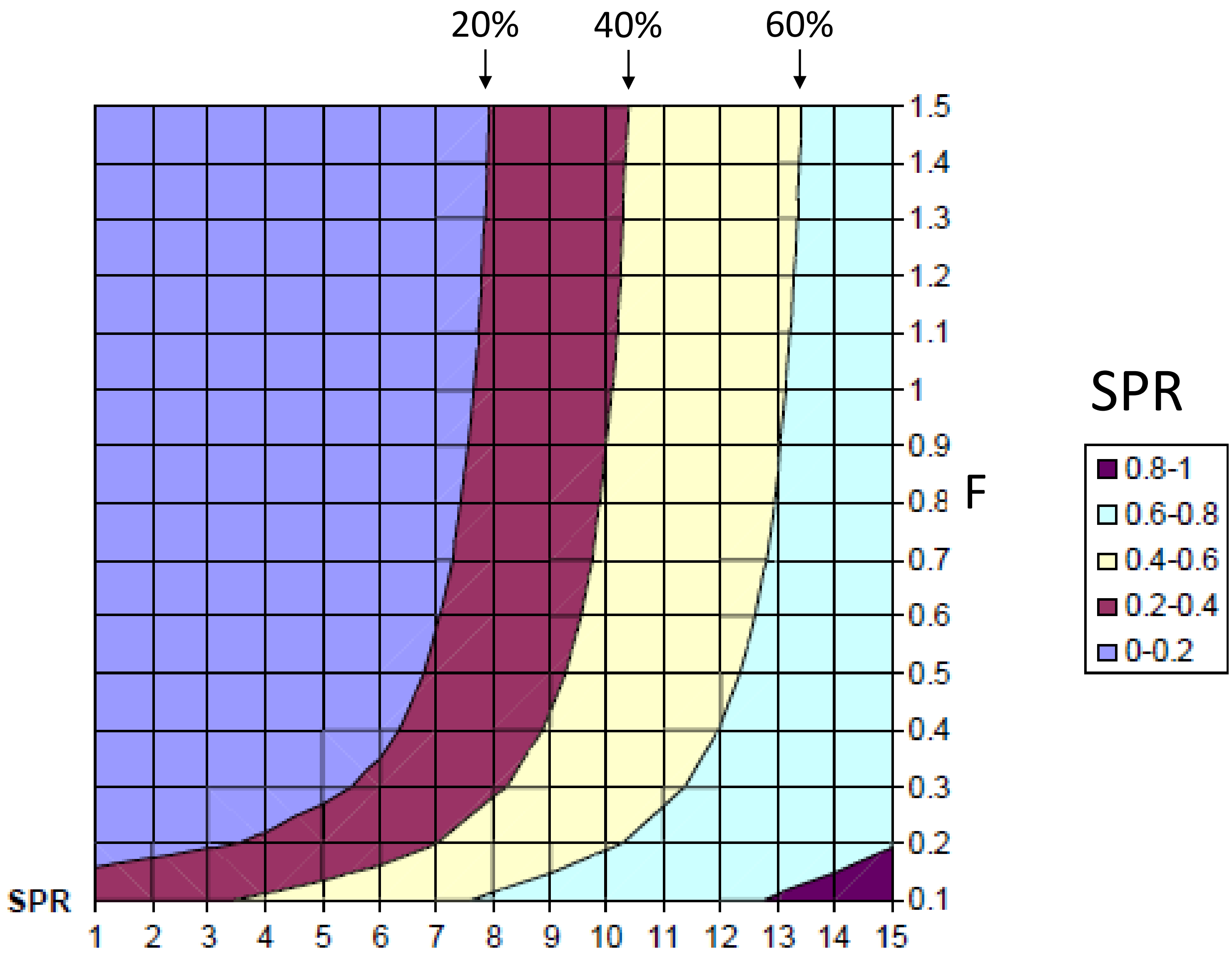
Reference points from SSB/R



Recruitment overfishing – when a population has been exploited to the point at which recruitment is substantially reduced (Sissenwine and Shepherd 1987)
- Often based on an SPR-type reference point

- **$F_{x\%SPR}$** = fishing mortality rate that keeps spawning potential ratio at X% of the maximum spawning potential
  - E.g., $F_{40\%}$ is relatively common.

Spawning Potential Ratio (SPR)



Spawning per recruit assumptions

- Constant life history schedules
  - Mortality at age
  - Growth (weight at age)
  - Maturity at age
  - Fecundity at age (for Eggs-Per-Recruit [EPR] models)

- Assumes spawning occurs over a short period in time (ie not continuous process)

Uses of spawning per recruit

- Generation of biological reference points ($F_{x\%SPR}$)
    - Used as a proxy of F_{msy}
    - 35% SPR suggested as a general reference point, but may be too aggressive for many life histories
    - 40% SPR now commonly used but higher values (50% to 60%) have been suggested for some species.

- Helpful in avoiding recruitment overfishing

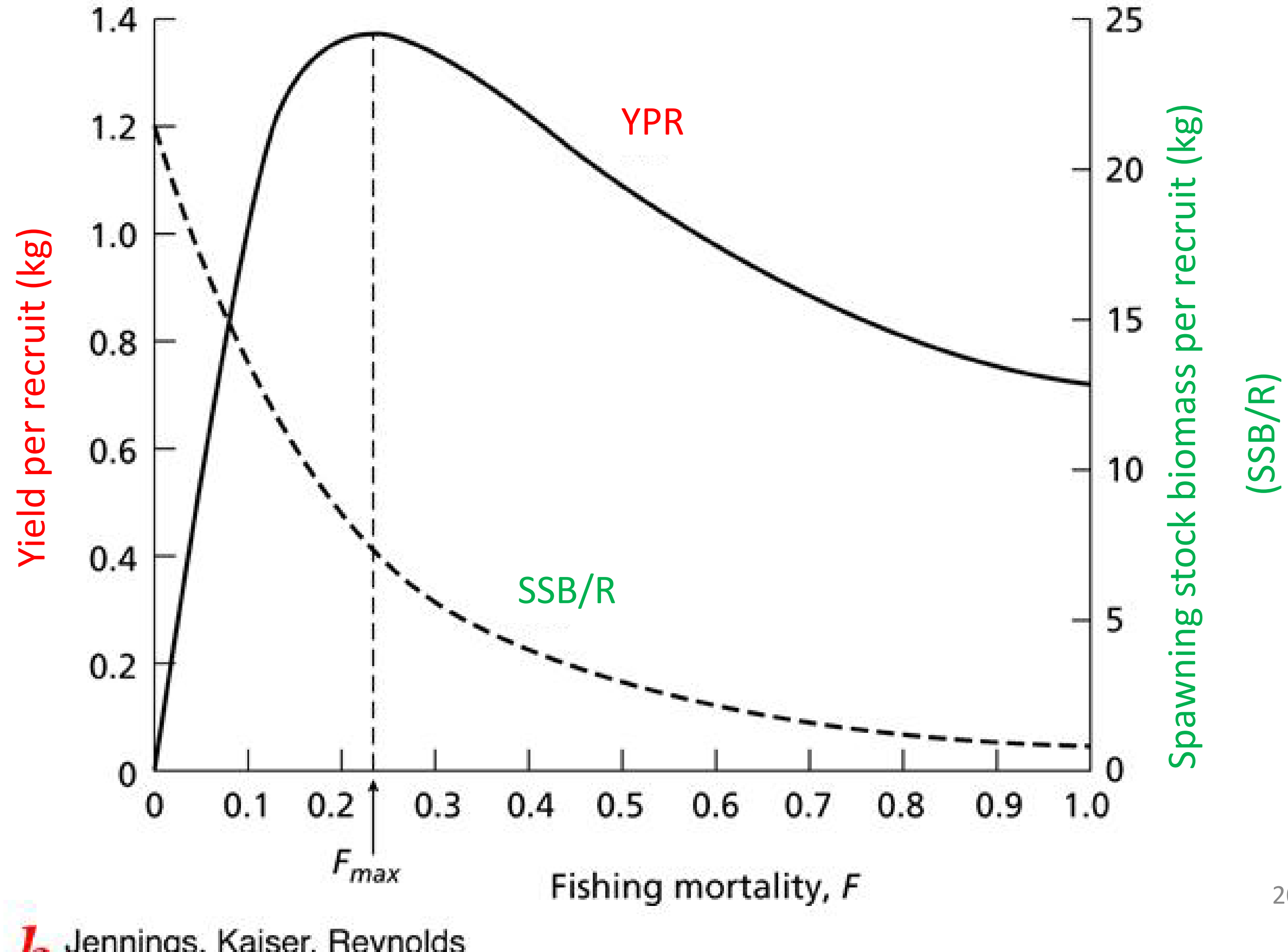
Limitations of SPR Models

- Constant assumptions
- Equilibrium conditions
  - No density dependence
- Does not contain stock-recruitment assumptions
  - But possible to build this in

Examples of challenges/problems

- External drivers of populations
  - Environment, Climate, etc.
- Doesn’t account for density dependence
- Possible changes as density decreases:
  - Growth rate increases (ie increased W at age)
  - Earlier maturity
  - Higher fecundity at age (due to size)
  - Smaller egg sizes (due to younger age at maturity)
  - Decreased natural mortality
- How have life history traits changed over the history of the fishery?
  - Traits/schedules now (which we estimate) may be different than virgin population

Tradeoffs in YPR and SSB/R



Color: SSB/R
Contours: YPR

Example

- Walleye YPR and SSBR in Mille Lacs Lake

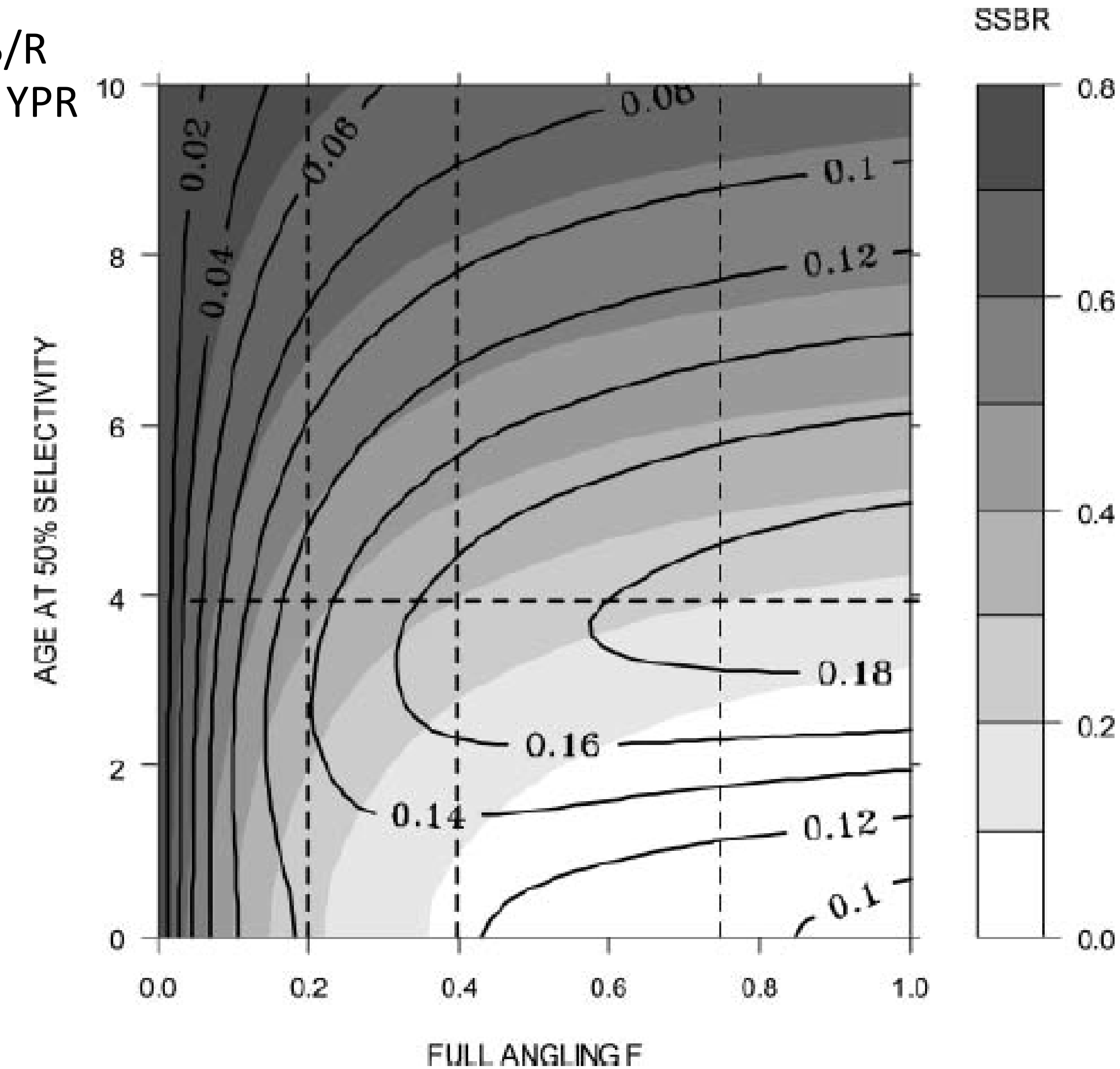
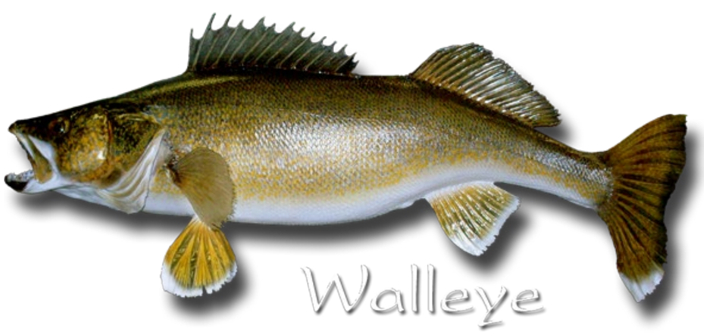


FIGURE 2. Mille Lacs Lake Walleye YPR and SSBR as a function of the age at which 50% selection to the angling fishery occurs, at Full Angling F. The horizontal dashed line represents the age (3.9 years) where 50% selection results in an SSBR = 0.3 at a median $F = 0.4$, as estimated from the SCAA model. The vertical dashed lines represent low $F = 0.2$ (10th percentile of SCAA-estimated F) and high $F = 0.75$ (90th percentile of SCAA-estimated F).



Schmalz et al. 2016

Summary – SSB/R

$$SSB/R = \sum_{a} Mat_a W_a e^{\sum_{j=\min a}^{a} -Z_j}$$

- SSB/R models
  - Examples of "spawning per recruit" models → but they look at different measures of spawning
  - Typically expressed as ratio of fished to unfished values
  - Definitions/concepts to know: **SSB/R**, **SPR**, **MSP**, **recruitment overfishing**
- SSB/R models account for:
  - age-structure
  - growth, mortality, maturity
- Don't account for:
  - Recruitment, changes in growth/mortality/maturity schedules, or density dependence
- Provide info and reference points for management
  - Deals with **recruitment overfishing**
  - $F_{x\%}$ (e.g., $F_{35\%}$ and $F_{40\%}$)
  - Often used as a proxy for F_{MSY} (b/c don't need S-R relationship)
  - Info on best age at first capture
  - These methods can also be modified to calculate eggs per recruit (EPR)

Estimating mortality

Supplemental Readings:

Millar 2015 (CJFAS)

Kenchington 2014 (Fish and Fisheries)

Acknowledgements: M. Wilberg, D. Ward, D. Secor

Announcements

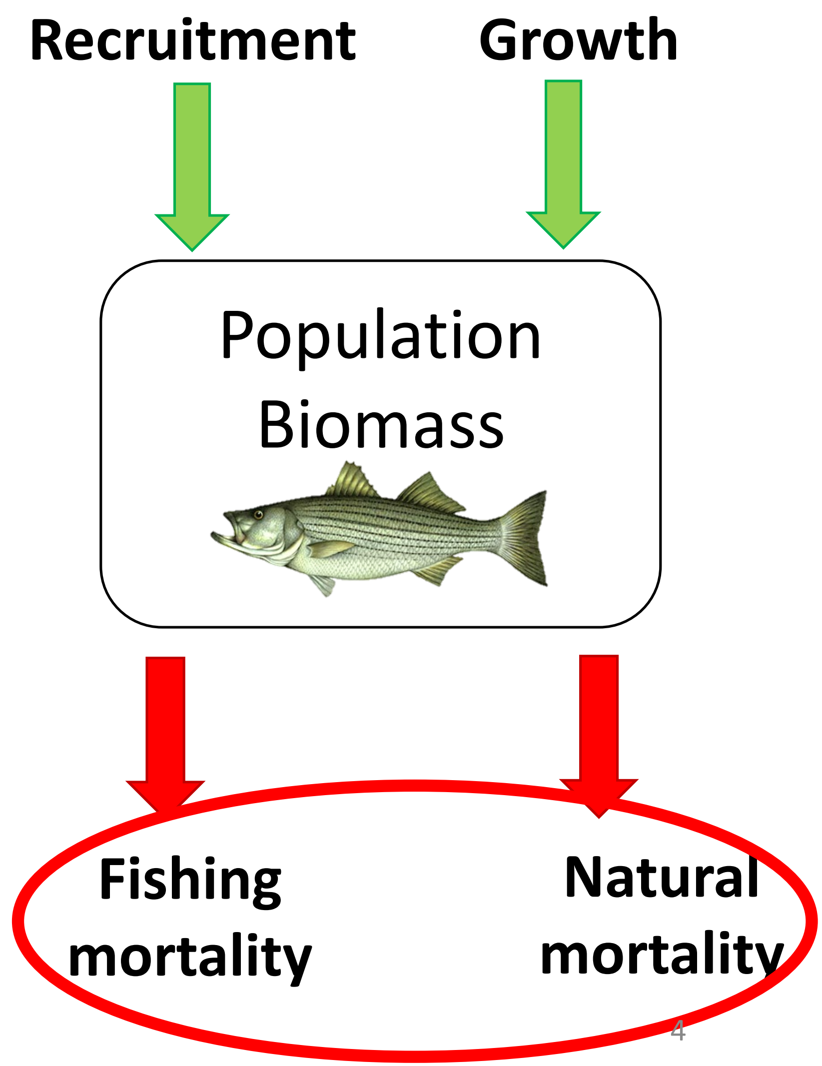
- **Grad students, please review the requirements for your project (see canvas)**
  - **will discuss on Friday**
  - **Synopsis Due: March 18**

Mortality

<u>Sources of mortality</u>

- Exploitation
- Predation
- Disease
- Starvation
- Senescence

- Grouped into fishing and natural mortality



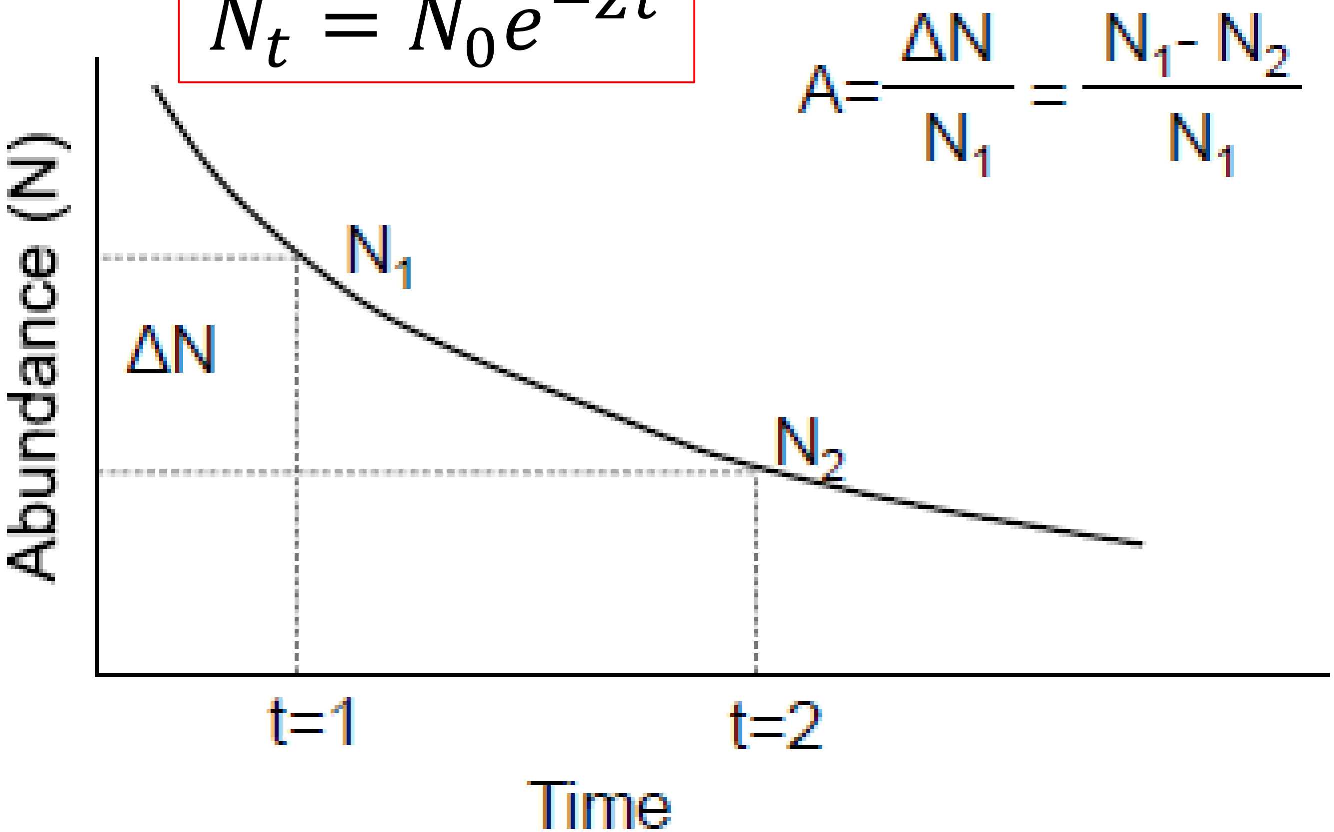
Russell 1931

"MY PARENTS DIED. THEIR PARENTS DIED. THEIR PARENTS DIED... IT RUNS IN THE FAMILY."

Exponential mortality model

- Assume exponential decline in abundance over time (ie, constant proportional change)

$$\boxed{N_t = N_0 e^{-Zt}}$$



How to estimate total instantaneous mortality Z?

How to estimate Z?

1. Catch curve and related methods Will focus on this
2. Length-based estimators
3. Mark-recapture
4. Population models (or two population size estimates)

How to estimate Z?

1. **Catch curve and related methods**
    - Basic catch curve
        - Cohort or year-specific method
    - Related methods:
        - Chapman and Robson 1960
        - Maceina and Bettoli 1998
        - Mixed-effects Poisson log-linear model (Millar 2015)

1. Catch curve method

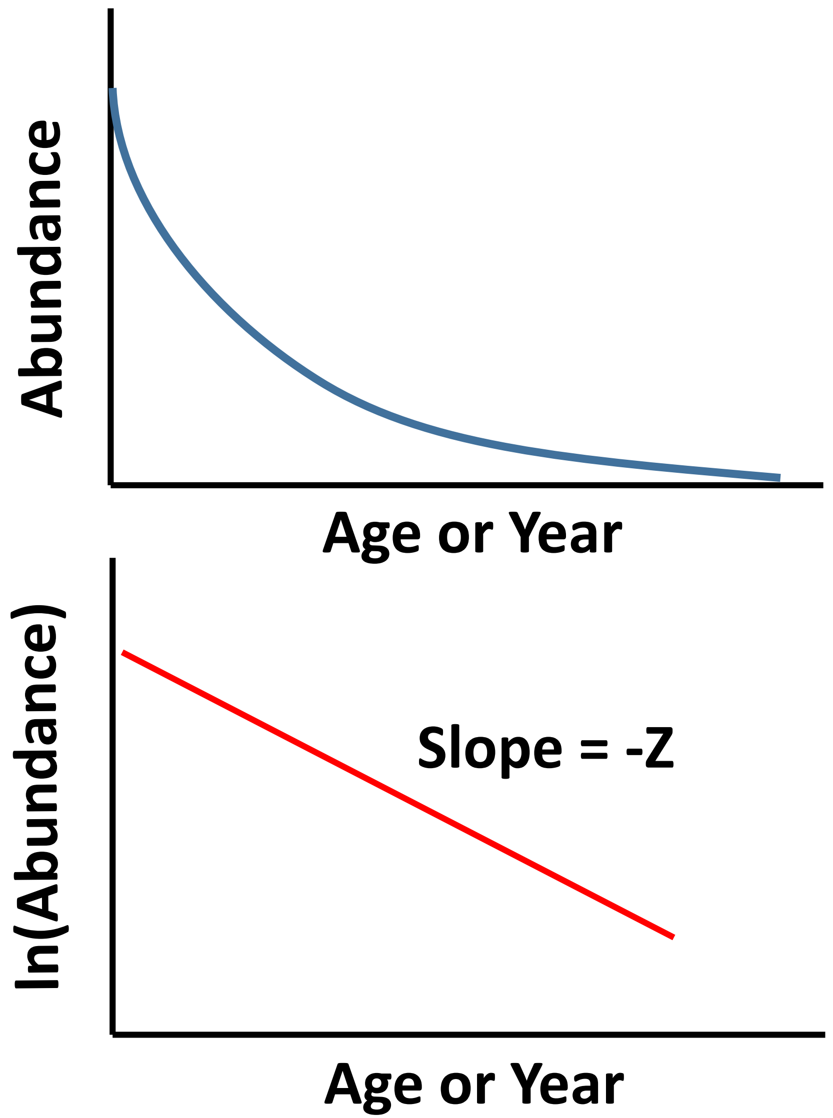
- General idea
  - Look at changes in relative abundance
  - Requires catch-at-age data
  - Linearize exponential mortality model

$$\boxed{N_t = N_0 e^{-Zt} e^{\varepsilon_t}}$$

Log transform

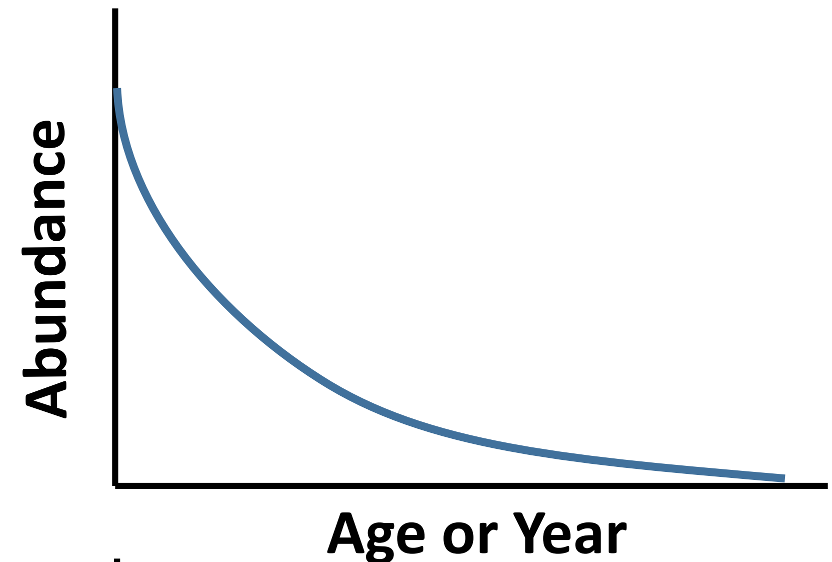
$$\boxed{\ln(N_t) = \ln(N_0) - Zt + \varepsilon_t}$$

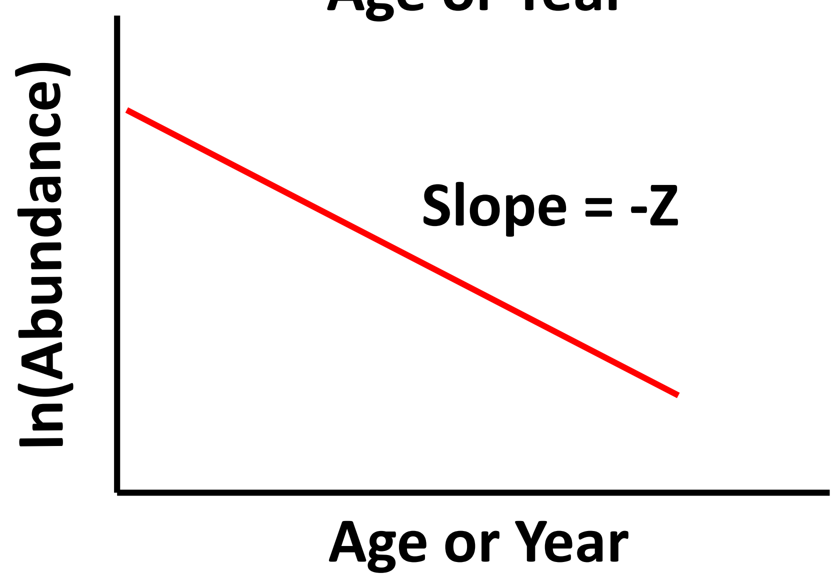
Note inclusion of error term



1. Catch curve method

- 2 options:
  - <u>Cohort-specific</u>: Track cohort abundance through time
  - <u>Year-specific</u>: Look at age distribution snapshot in 1 year





Cohort vs. year specific

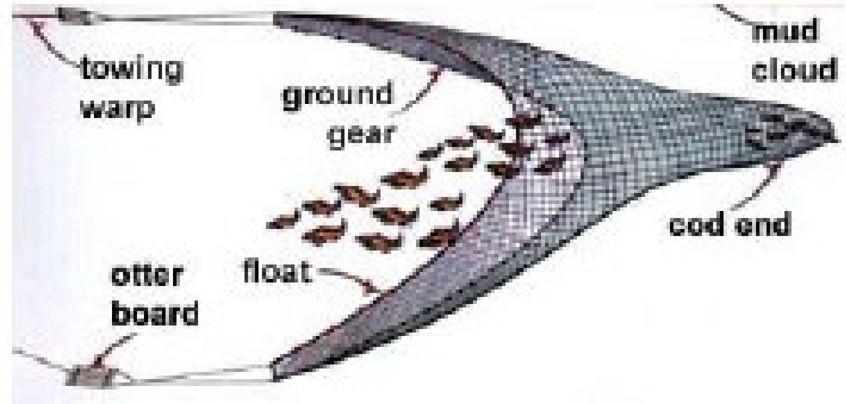
Year	Ages 3	4	5	6	7	8	9
1978	13	129	646	954	99	19	4
1979	19	169	416	1031	243	47	18
1980	40	354	606	479	152	18	7
1981	32	606	1424	644	157	23	17
1982	0	226	1178	1156	116	16	5
1983	2	165	593	982	428	22	11
1984	53	209	560	410	30	1	4
1985	0	105	674	446	16	2	2
1986	46	422	838	726	70	4	4
1987	3	310	1224	1068	65	0	0
1988	14	354	1264	1172	69	0	6
1989	6	429	1222	1067	192	0	0

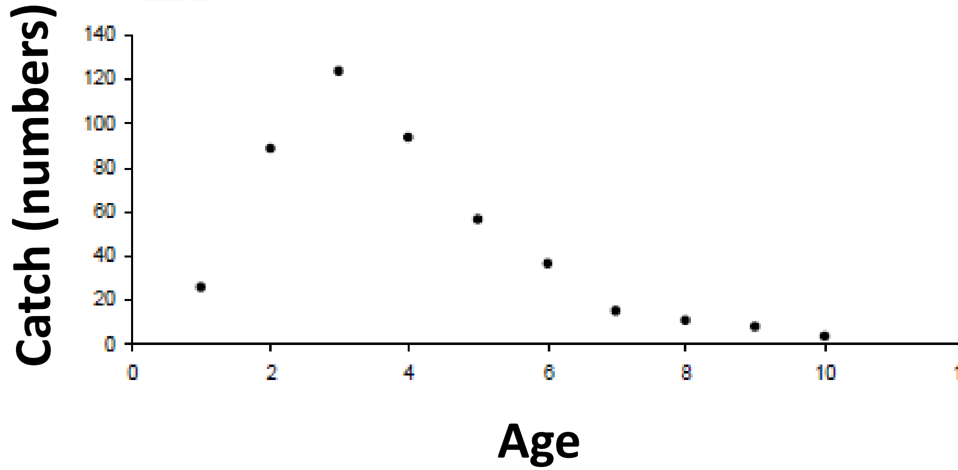
Year

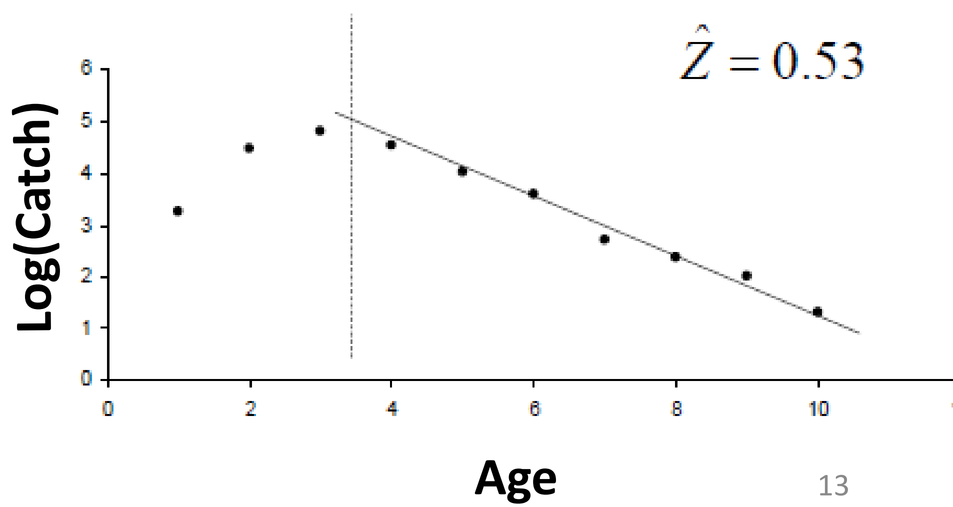
Cohort

Approach

1. Log transform abundance index data
    - Or use catch-at-age from fisheries data
2. Plot log(catch) vs. age
3. Select the first "fully selected age"
    - Typically the peak or next value
4. Fit regression using selected ages
    - Ie, use first fully selected age and older







Examples

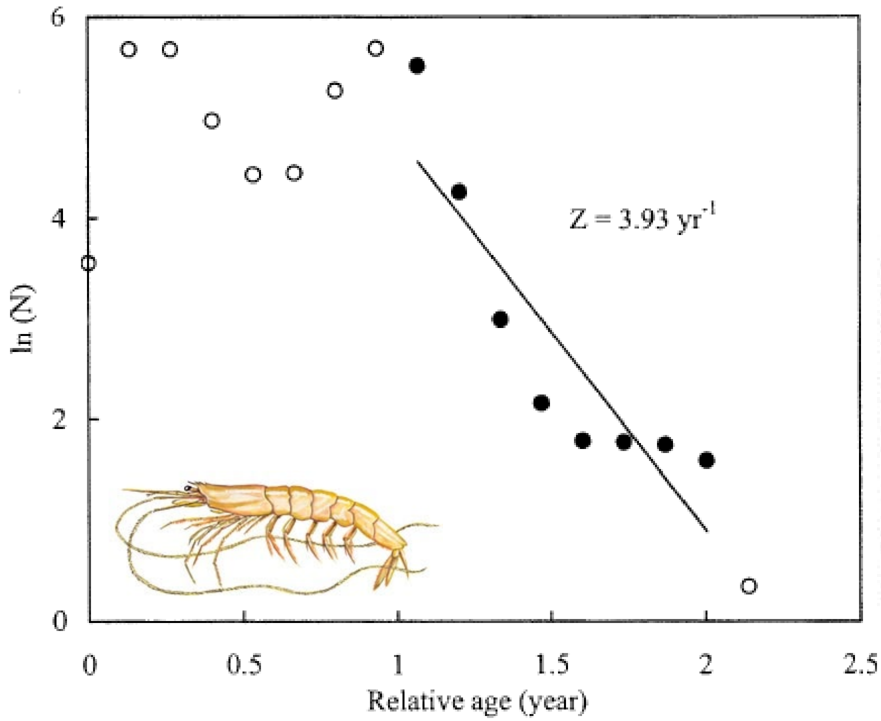
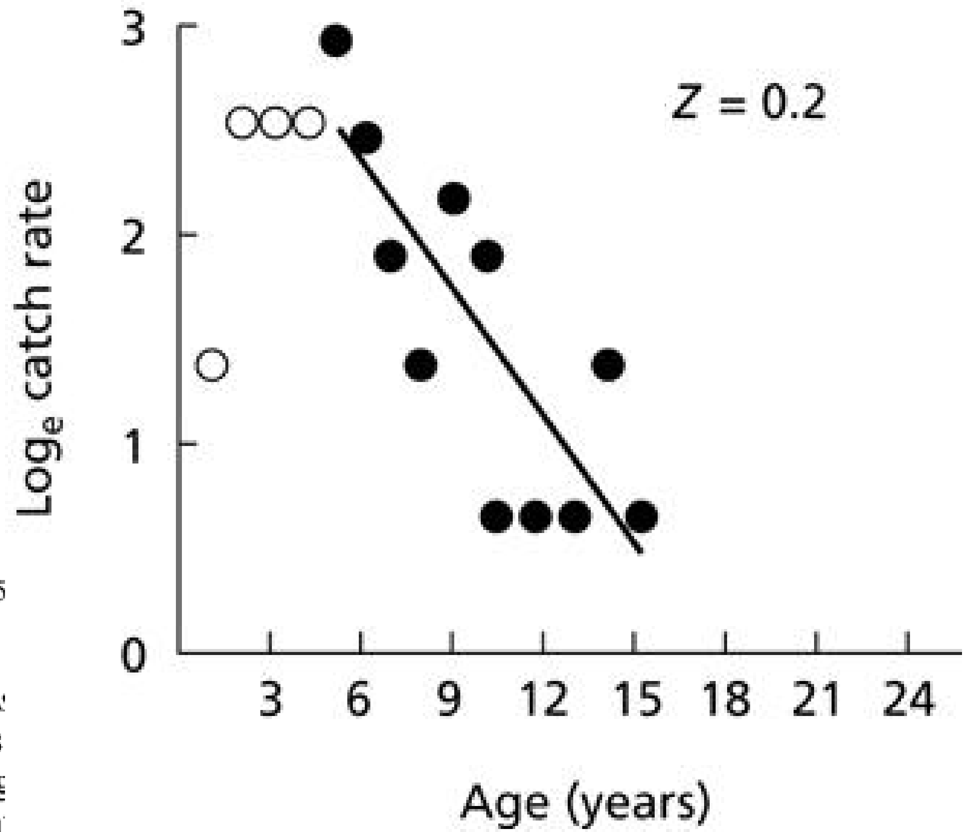


Fig. 9. Length-converted catch curve of *Acetes chinensis* based on length-frequency data during the study periods The darkened circles represent the points used in estimating Z through regression analysis. The open circles represent points either not fully recruited or nearing L_∞, hence discarded from the calculation.

Oh and Jeong J Crust Bio 2003

Surgeonfish (*Acanthurus nigrofuscus*)

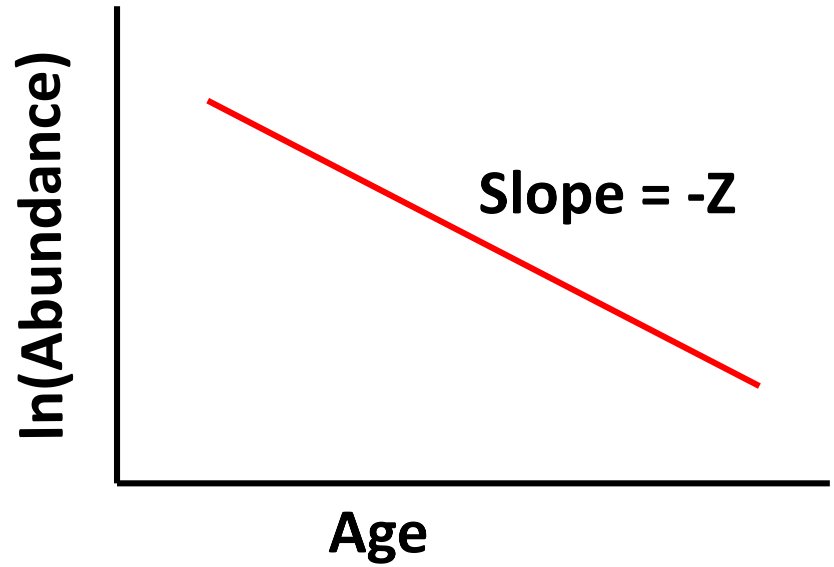


Hart and Russ 1996
(in Jennings et al. 2001)

1. Catch curve method

<u>Assumptions</u>

- cohort-specific
  - **No errors in ages estimates**
  - **Constant mortality** in all years and age classes
  - **Constant catchability**, or vulnerability to gear, over ages and years
- Year-specific
  - **Constant recruitment**
  - (rest are same as above)



Year	Ages 3	4	5	6	7	8	9
1978	13	129	646	954	99	19	4
1979	19	169	416	1031	243	47	18
1980	40	354	606	479	152	18	7
1981	32	606	1424	644	157	23	17
1982	0	226	1178	1156	116	16	5
1983	2	165	593	982	428	22	11
1984	53	209	560	410	30	1	4
1985	0	105	674	446	16	2	2
1986	46	422	838	726	70	4	4
1987	3	310	1224	1068	65	0	0
1988	14	354	1264	1172	69	0	6
1989	6	429	1222	1067	192	0	0

If recruitment is not constant, what type of pattern would be most problematic?

Advantages/disadvantages of each approach

- Cohort-specific
  - PRO: No need to assume constant recruitment
  - PRO: Allows estimation of mortality rates for individual cohorts
  - CON: many years of data needed
  - CON: It is in the past
- Year-specific
  - PRO: Only one year of data is needed
  - PRO: Represents what is occurring in recent time
  - CON: Need to assume constant recruitment

Other catch-curve-like methods

- Chapman and Robson 1960
- Weighted regression (Maceina and Bettoli 1998)
- Mixed-effects Poisson log-linear model (Millar 2015)
- Modifications
  - E.g., logistic selectivity & age-specific M (Thorson and Prager 2011)

Chapman & Robson (CR) method

Extra

- Uses mean age, age at full selection, and the number of samples (above the age of full selection)

$$\hat{S} = \frac{\bar{a} - a_r}{a_r - \bar{a} + \frac{n-1}{n}} \qquad \hat{Z} = -\log(\hat{S})$$

  - S is survival
  - a_r is the age at full selection
  - $\bar{a}$ is the mean age of the sample
  - n is the sample size

How is the numerator reflective of survival?

- Notes:
  - assumes duration of life follows a geometric distribution
  - Variance estimate should be corrected for bias from overdispersion

Chapman and Robson 1960; Millar 2015

Weighted linear regression
(Maceina and Bettoli 1998)

Extra

- Two-step process
- 1) do the normal regression estimate (i.e., catch curve)
- 2) Use the estimated $\log(C_a)$ values from the first regression to do a weighted regression
  - Essentially, this is an *ad hoc* method to give higher weight to ages with higher abundances

Extra

Mixed effects Poisson Model

- Uses GLMM (generalized, linear, mixed-effects model) and maximum likelihood
- Builds in recruitment variability using a random intercept
- Assumes a Poisson distribution with log link

Millar 2015

Recommendations for catch-curve-like methods

- For year-specific approach (ie “cross-sectional”)
  - Basic catch curve should **<u>not</u>** be used
  - Smith et al. 2012 (TAFS) recommended the Chapman Robson method corrected for overdispersion
    - Use: age of max catch + 1 as the lower age limit
  - Millar 2015 recommended Poisson GLMM

2. Length-based estimators for Z

Extra

- Use length as a proxy for age
- When is this useful?
- Resources:
  - Reviews:
    - Hoenig et al. 1983;
    - Shepherd and Breen 1992
  - Examples:
    - Beverton and Holt 1957; Ehrhardt and Ault 1992; Gedamke and Hoenig 2006; Then et al. 2015

2. Length-based estimators for Z

Extra

- Example: Beverton and Holt (1957)
  - Assumptions: Von Bert growth; knife edge selectivity at L_c; constant mortality

$$Z = K \frac{L_{inf} - \bar{L}}{\bar{L} - L_c}$$

  - K & L_{inf} from von Bertalanffy model;
  - Lbar is mean length in sample
  - L_c is length at first capture

How to estimate Z?

1. Catch curve method
2. Length-based estimators
3. Mark-recapture
   - Estimate disappearance of marked individuals over time
   - Need multiple detection events
   - Assume survival of marked individuals = survival of unmarked individuals
   - *Future lecture…*
4. Population models
   - Survival = N_{t+1}/N_t (if no recruitment)
   - *Future lectures on stock assessment models…*

Estimating instantaneous natural mortality (M)

Estimating natural mortality (M)

1. Catch curve analysis
2. Length-based estimators
3. Mark recapture methods
4. Life-history methods (empirical methods)
5. Population models (Multispecies VPA)
6. Pope's derivation

Will focus on this

Estimating natural mortality

1. Catch curve analysis

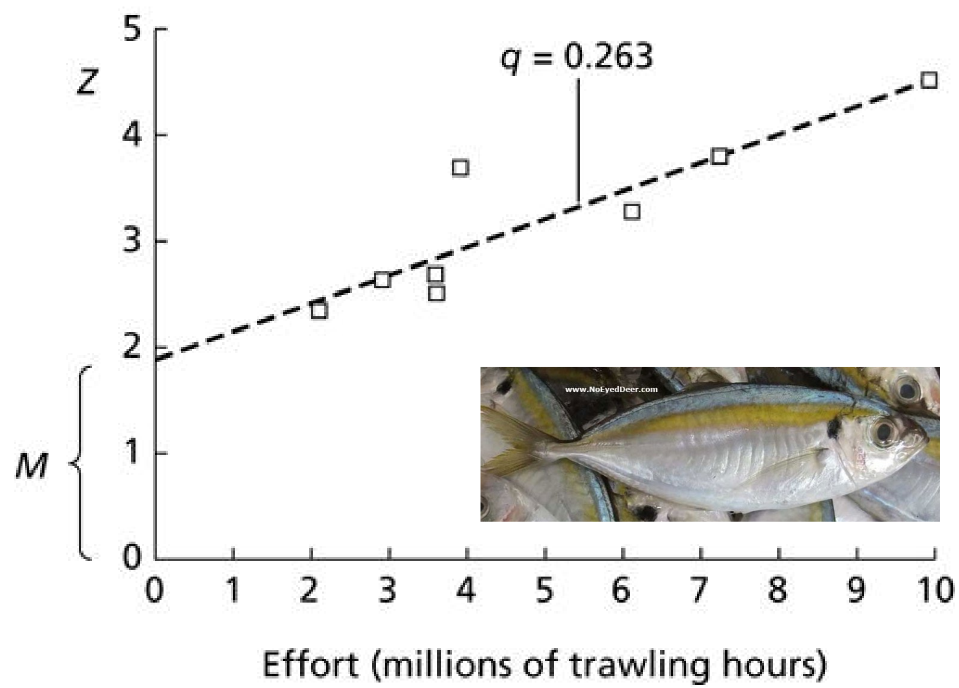
A. Catch curve on population with no fishing
   - Restricted applicability given our interest in fishing

B. Use relationship of Z to fishing effort (E)

$$Z = M + F \qquad F = qE$$

$$Z = M + qE$$

- M=natural mortality (y intercept)
- q=catchability coefficient
- E=fishing effort
- Requires wide range of fishing effort



Smooth-tailed trevally (*Selaroides leptolepis*)
(Pauly 1982, in Jennings et al. 2001)

Estimating natural mortality

2. Length based estimators

- Good for data-limited fisheries
- Use estimates of Z (as described before) and F to get M
- See Hoenig et al. 1983; Shepherd and Breen 1992

Estimating natural mortality

3. **Mark-recapture**
   - Estimate disappearance of marked individuals over time
   - Need multiple detection events
   - Assume survival of marked individuals = survival of unmarked individuals
   - Partition total mortality into fishing and natural

   - *Future lecture…*

Estimating natural mortality

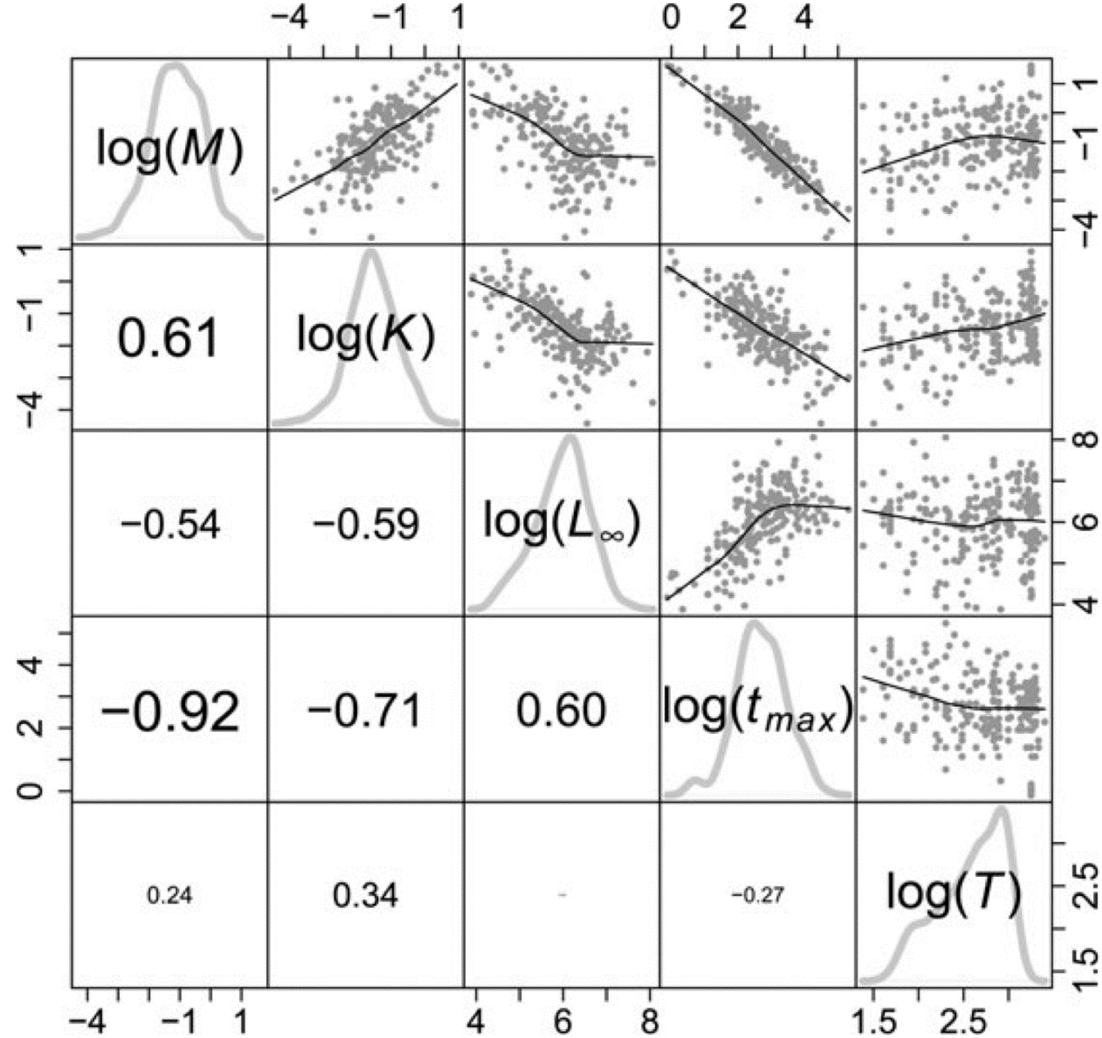
4. **Use empirical estimators**

What life history parameters do you think would be related to M?

Draw how you think M is related to:

- K (from von Bertalanffy model),
- Linf (from von Bert model; asymptotic max length)
- Max Age (t_{max}),
- mean temperature (T)

Figure 1. Scatterplot of pairs of log-transformed variables in the upper half of the panel, with LOWESS smooths. (>200 M estimates)



Then et al. 2015. ICES Journal of Marine Science, Volume 72, Issue 1, 20 August 2014, Pages 82–92, https://doi.org/10.1093/icesjms/fsu136

Estimating natural mortality

4. Use empirical estimators

- Often used for information limited fisheries
- M related to life history parameters (e.g., K, Linf, mean Temp)
  - www.Fishbase.org is a great source of values!
- At least 30 equations proposed!
- 2 of the more robust models:
  - **Pauly 1980**:

$$\ln(M) = -0.0152 - 0.279\ln(Linf) + 0.6543\ln(K) + 0.4634\ln(T)$$

  - **Jensen 1996**

$$M = 1.5K$$

Review: Kenchington 2014 (Fish and Fisheries)

Sidenote: Size/Age-specific M

- Many empirical relationships developed.
- Decent examples:
  - Lorenzen 1996

$$M_w = 3.00w^{-0.288}$$

  - Gislason et al. 2010

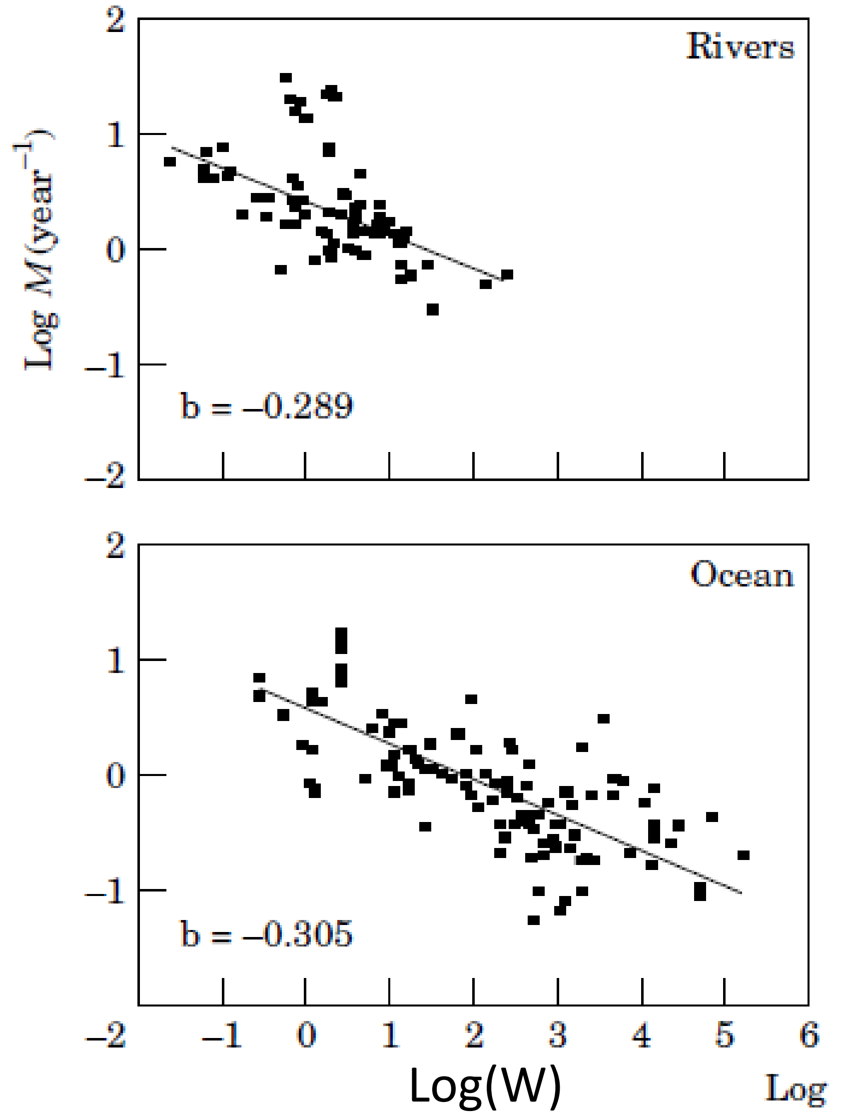
$$M_l = 1.73l^{-1.61}L_\infty^{1.44}K$$

  - Charnov et al. 2012

$$M_l = K\left(\frac{l}{L_\infty}\right)^{-1.5}$$

Kenchington 2014 (Fish and Fisheries)

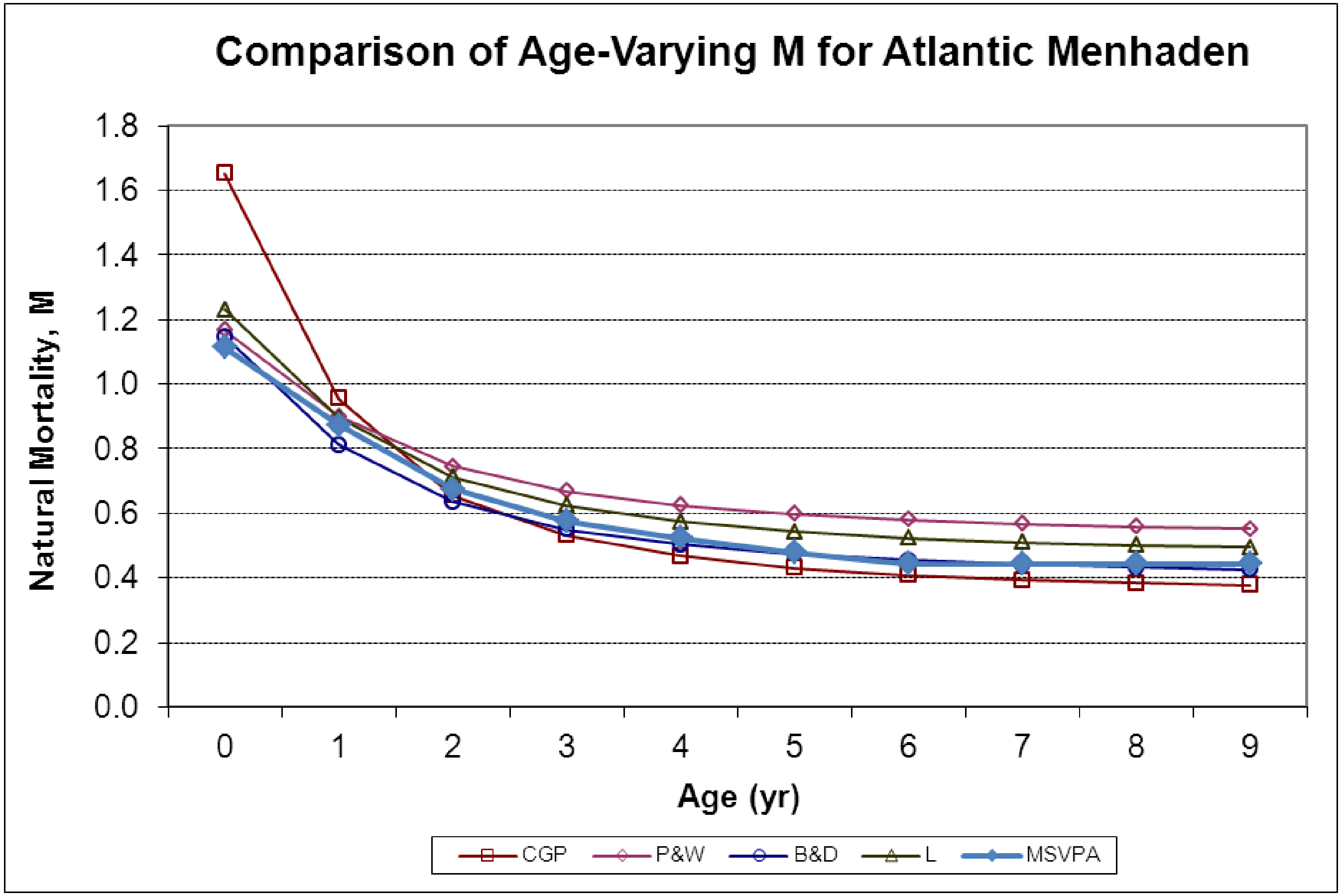
Lorenzen 1996



M_w = M at weight, *w*
M_l = M at length, *l*
Linf, K = Von Bert parms

Sidenote: Size/Age-specific M

- Example
  - More realistic biologically, but hard to estimate accurately
  - Can scale to other mortality estimates (e.g., tagging-based)



Estimating natural mortality

5. Population Models

- Multispecies models (e.g., MS VPA)
  - Multispecies models account for consumption by predator species
  - Use this info to estimate predation mortality
  - *Future lecture(s)*
- Estimate M within stock assessment (or ecosystem) models

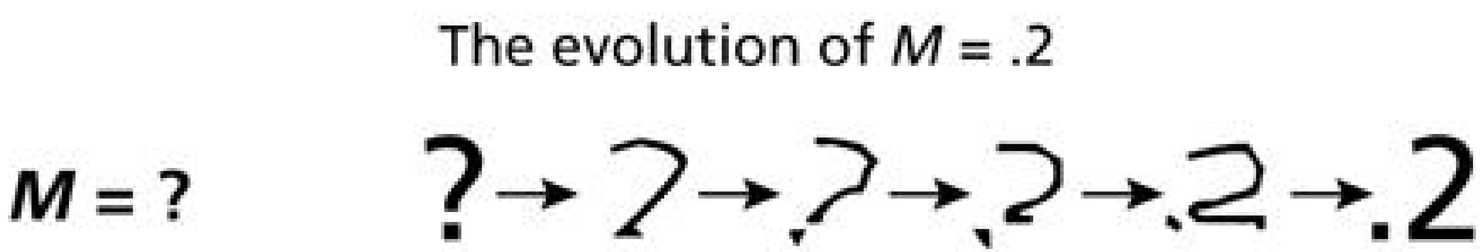
- Comments
  - Data intensive
  - Can be controversial given uncertainties

Kenchington 2014 (Fish and Fisheries)

Estimating natural mortality

6. Pope's Derivation

- Many models just assume M=0.2 if no other info available
- Not ideal, but common
- Note: this is NOT an "estimate" of M… just an assumption



Summary

- Mortality is critical component for population dynamics and management
  - → challenging to estimate precisely
- General approaches for estimating Z or M:
  - (Z or M) **Catch curve** and related methods
    - Know general catch curve process, diff btw cohort & year- specific
    - Chapman/Robson Method or Poisson GLMM recommended (for Z)
    - (M) Regress Z on Effort → intercept=M
  - (Z) Length-based estimators
  - (M) **Life-history methods (meta-analysis)** – for M only
    - Size/length specific estimates possible
  - (Z or M) Mark-recapture (*future lectures*)
  - (Z or M) Population models (*future lectures*)
  - (M) "Pope's derivation" → many just assume M=0.2 in the absence of other information. Not ideal. Also would <u>not</u> be considered a "<u>estimate</u> of M".

Replacement Lines

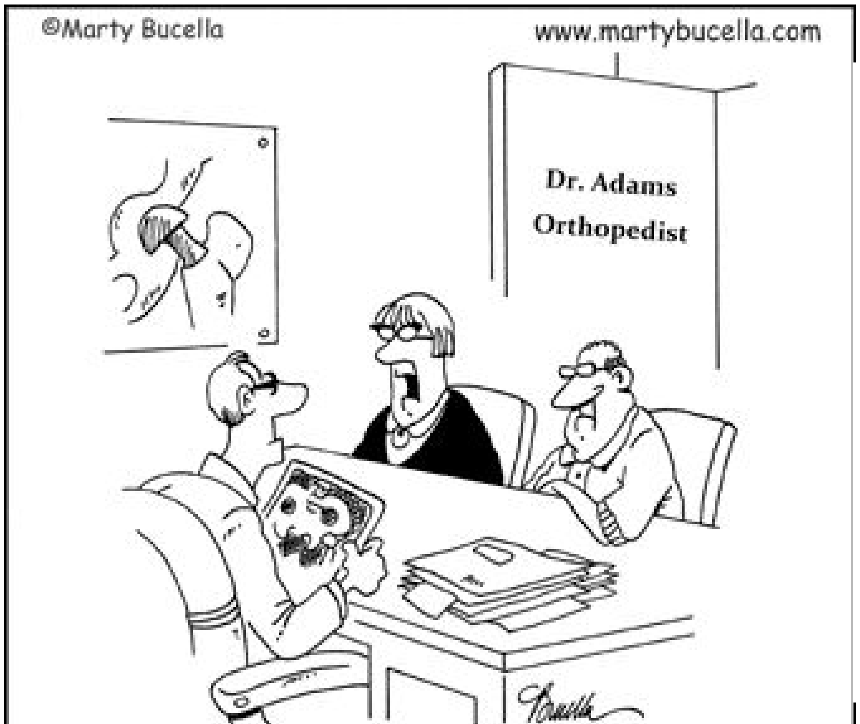
Reading:

Jennings et al. 2001, section 7.8

Supplemental: Mace and Sissenwine 1993

Supplemental: Gabriel et al. 1989

Acknowledgement: M. Wilberg



"Hip replacement? He was never hip to begin with."

"Tom, we're letting you go, but we'd like you to stay on and train your replacement so they know what not to do."

Background

- Dynamic pool models (YPR, SSB/R, EPR) did not account for stock-recruit relationships

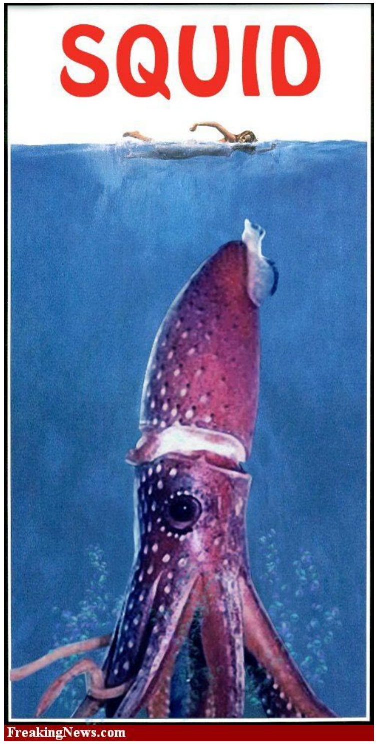
- Why is this potentially problematic?

Background

- **Replacement line** = Line on a stock-recruit plot indicating the amount of recruits per spawner needed to replenish the population on average
  - Semelparous species – can use stock recruit models alone
  - Iteroparous species – combine SSB/R models with stock-recruit models

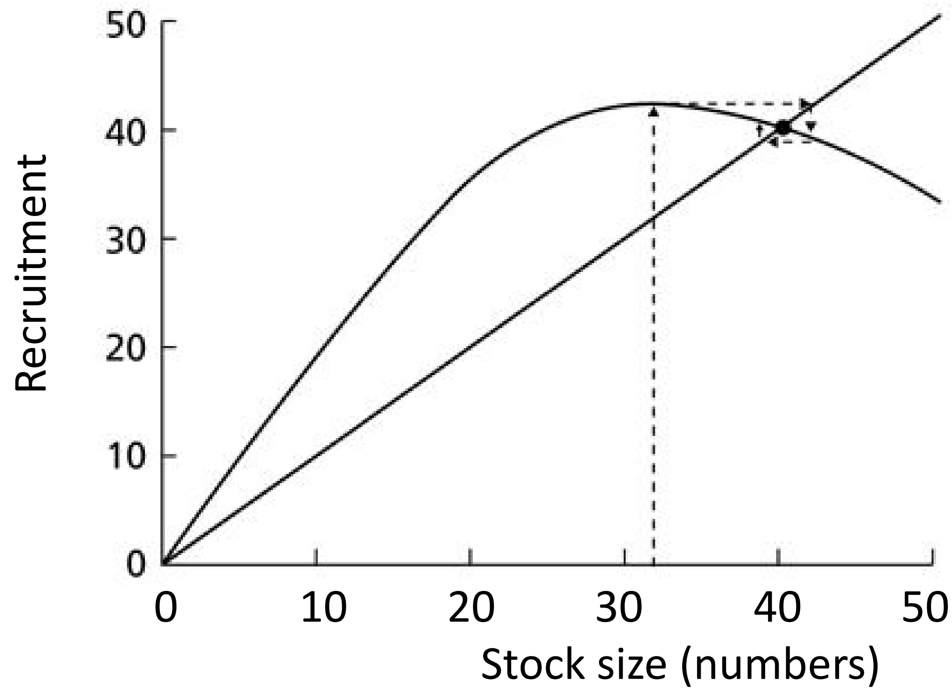
Example of replacement lines

- **Example**: consider a semelparous species with a life span of 1 year (e.g., squid)
- Draw Ricker Curve with replacement line. Discuss:
  - Meaning
  - Equilibrium
  - Effect of fishing on replacement line



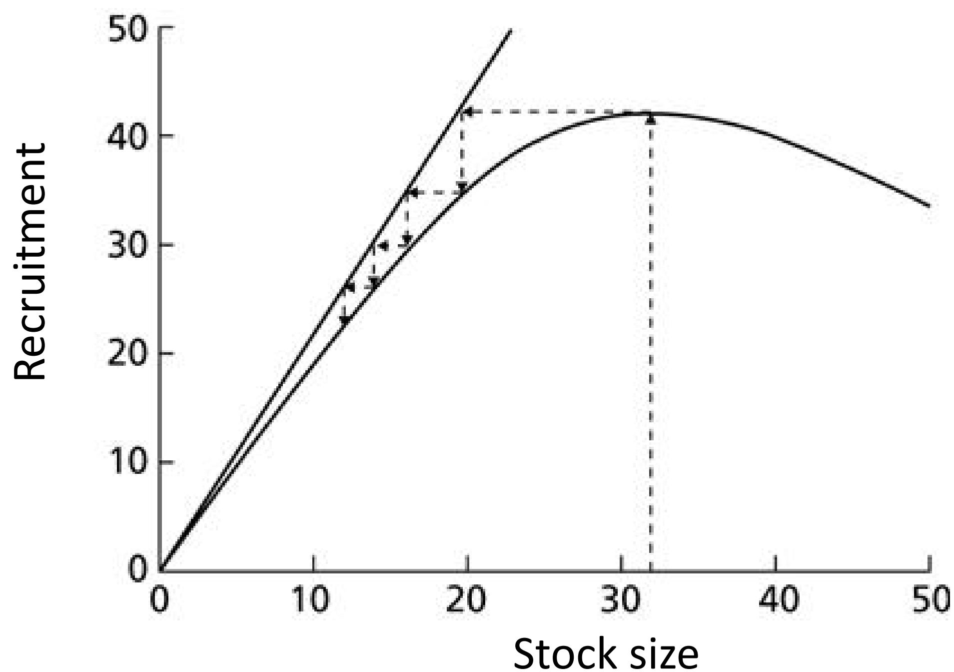
Example

- Ricker SR model
- 1:1 replacement line
    - 1:1 b/c dealing with annual species; assume all mortality occurs prior to recruitment
- If SR curve is above replacement line, we get a stable equilibrium (at intersection)



Example

- Steeper replacement line to accommodate fishing
- If SR curve is below the replacement line, population crashes



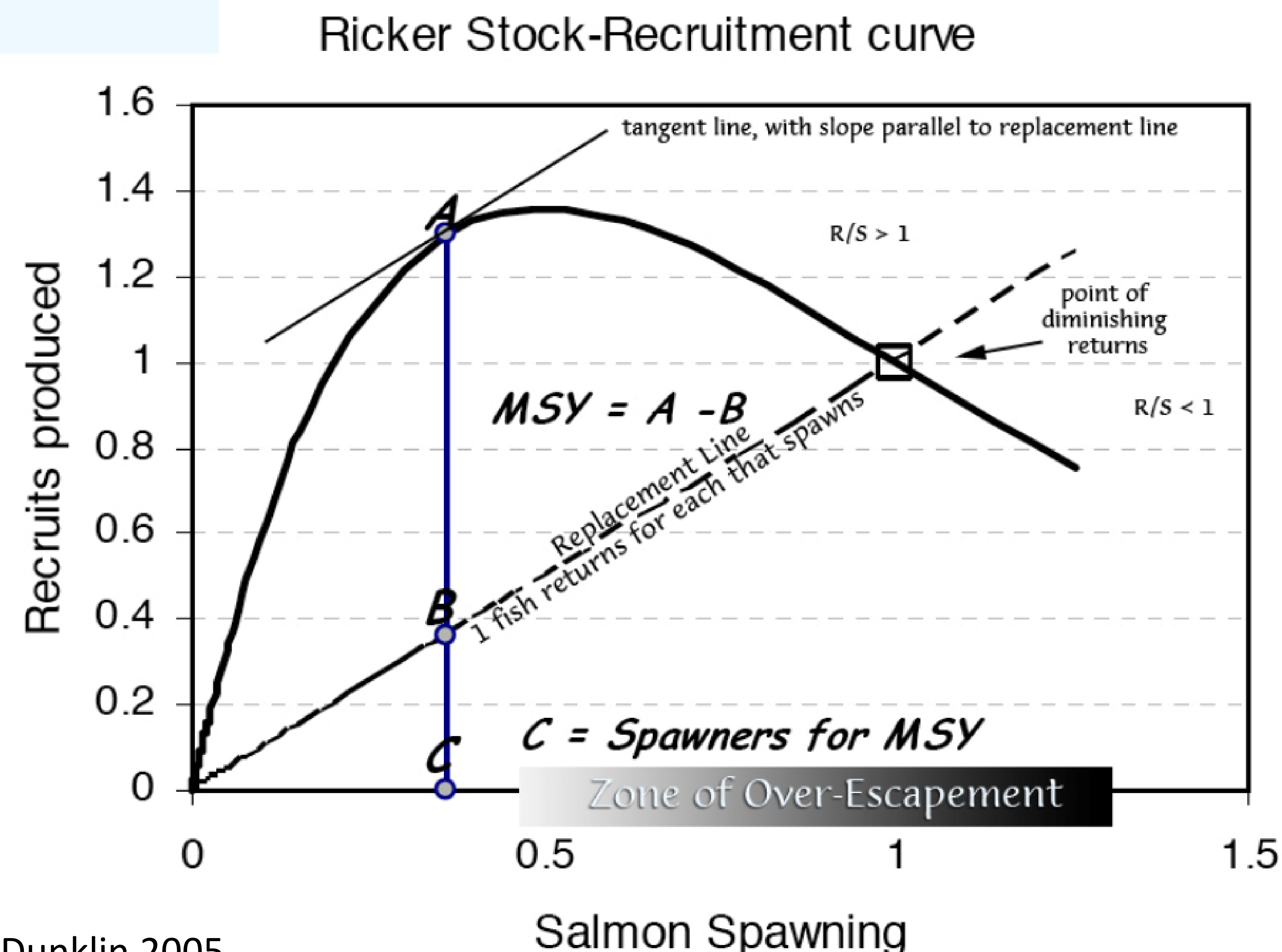
Replacement lines and MSY **for semelparous species**

- Area above the replacement line equates to surplus production
- Use this to calculate MSY from SR models
- Qualifications:
    - Semelparous species only
    - Recruitment is measured after all non-fishing mortality
    - Fixed timing of life cycle (ie non-overlapping generations)
    - Recruitment needs to be adjusted for years at large
- Best example: Pacific salmon

→ Draw a Ricker S-R Model with a replacement line and use that to identify the following values: S_{MSY}, MSY

Hilborn and Walters 1992 - Chapter 7

MSY from replacement lines in Stock-Recruitment Models



Escapement – the number of fish that are not caught in the fishery and allowed to survive and reproduce

Dunklin 2005

Extra

Ricker Model – MSY calculations

- Approximations to estimate MSY reference points from SR parameters
- Side: can use numerical methods to get exact values

$$R = \alpha S e^{-\beta S}$$

$$S_{MSY} \sim \frac{\log_e \alpha}{\beta}(0.5 - 0.07\log\alpha)$$

$$MSY \sim \alpha S_{MSY} e^{-\beta S_{MSY}} - S_{MSY}$$

$$u_{MSY} \sim 0.5\log_e \alpha - 0.07(\log\alpha)^2$$

Hilborn and Walters 1992 (Chapter 7)

Beverton Holt Model – MSY calculations

Extra

- Analytic solutions for BH model: $R = \dfrac{\alpha S}{\beta + S}$

$$S_{MSY} = \alpha\sqrt{\frac{\beta}{\alpha}} - \beta$$

$$MSY = \frac{\alpha S_{MSY}}{\beta + S_{MSY}} - S_{MSY}$$

$$u_{MSY} = 1 - \sqrt{\frac{\beta}{\alpha}}$$

Replacement lines **for iteroparous species**

- Challenges determining the replacement line (e.g., iteroparous species):
  - Depends on maturation schedule, longevity, mortality rates
- Need to combine SR model with YPR and SSB/R models
- No easy analytical solution for MSY reference points

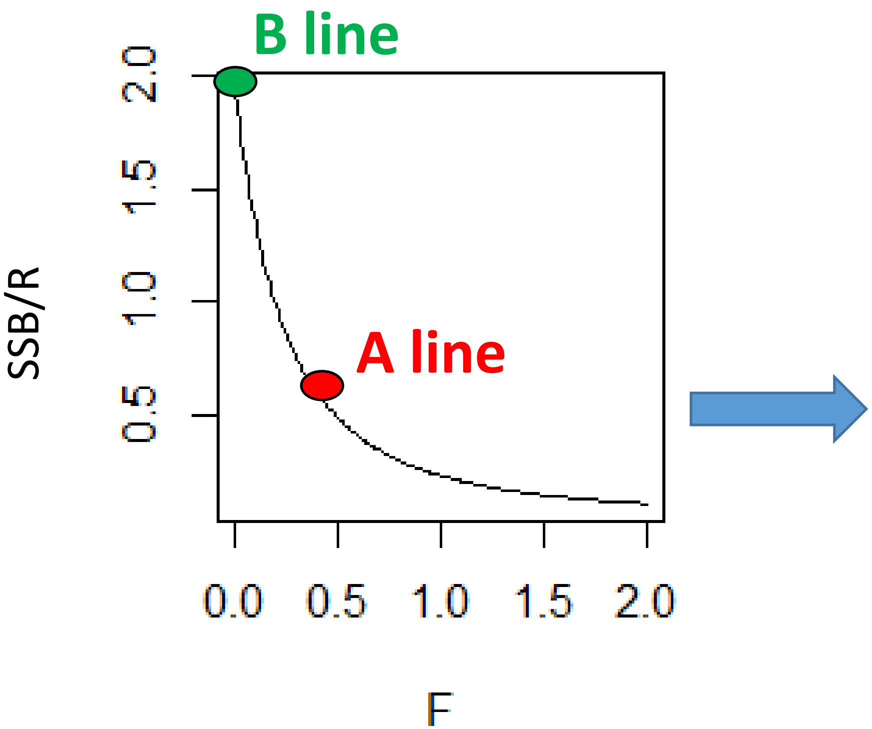
Relating SSB/R and R/SSB

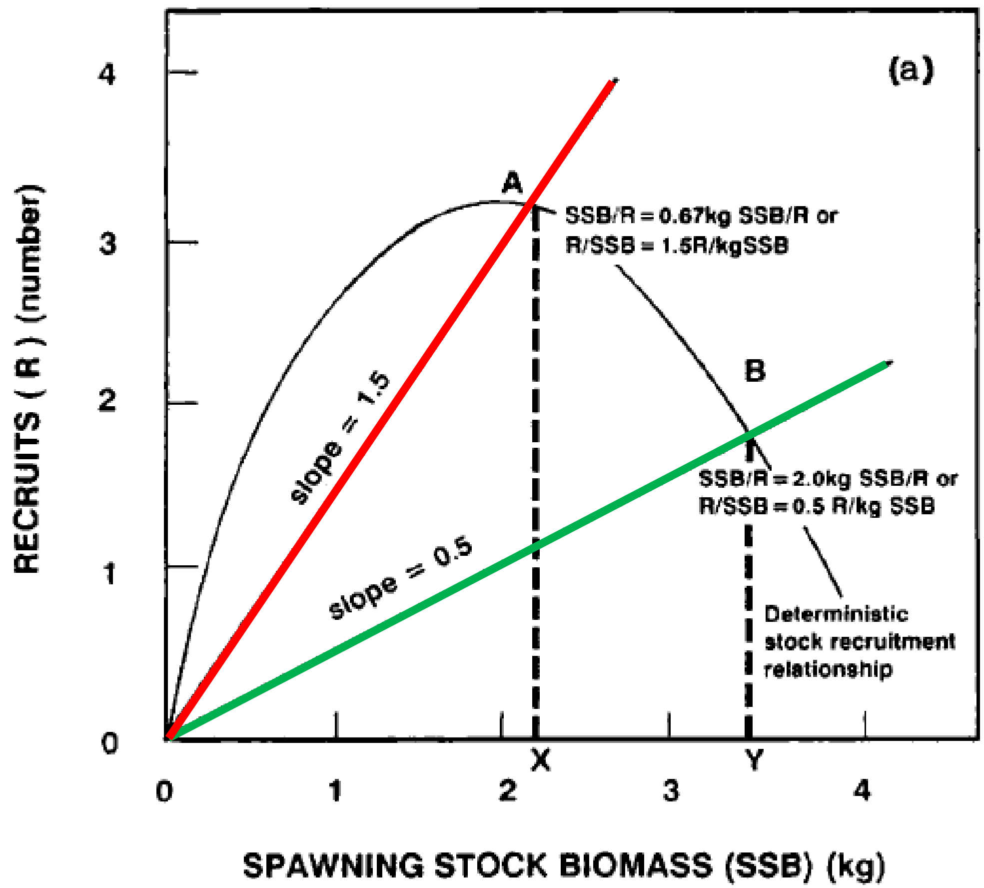
- Select desired **SSB/R** for a given level of fishing
- The inverse (**R/SSB**) is the survival ratio for our stock-recruit model
- **Replacement line** = recruits per spawner needed to replenish the population on average
  - **OR**: line in a stock-recruit plot with a slope equal to the observed average *survival ratio* (R/SSB)

Mace and Sissenwine 1993

Relating SSB/R and R/SSB

- Slope on S-R curve depends on the fishing mortality rate
- Slope on S-R = inverse of SSB/R

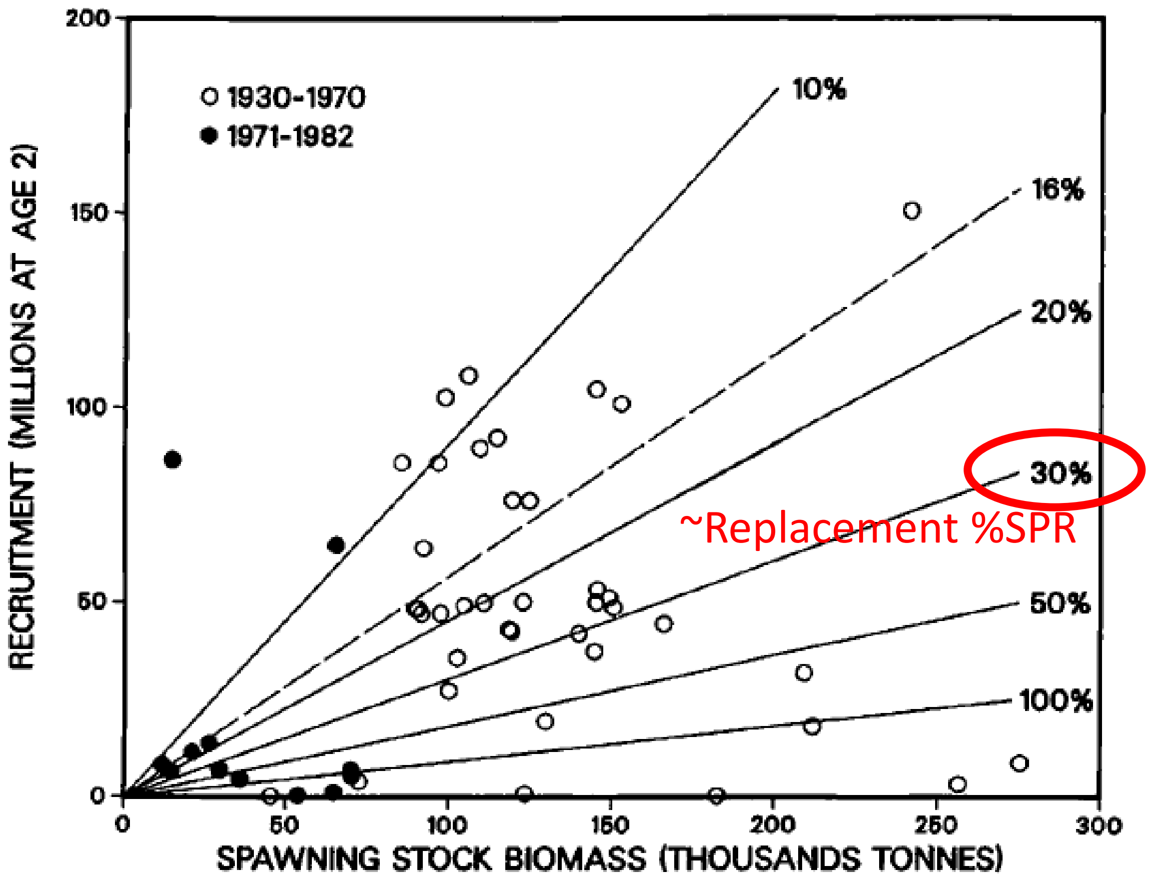




Gabriel et al. 1989

Example with real data: haddock

- Replacement lines labeled as % of maximum spawning potential
- Points above a line = R/SSB sufficient to maintain or increase pop (on average)
- One meta-analysis recommend 30%SPR as guide



Gabriel et al. 1989, Mace and Sissenwine 1993

Benefits of replacement lines

- Explicitly account for the effect of fishing on future recruitment (ie accounts for stock-recruit relationship)
- Addresses recruitment overfishing
- Can be used to determine biological reference points
  - E.g., F_{MSY}, F_{rep}, replacement %SPR
  - (Mace and Sissenwine 1993)
- Estimates of MSY are possible:
  - Easier for semelparous species like Pacific salmon
  - For iteroparous species, see Shepherd 1982

Limitations of replacement lines

- Data limitations
  - S-R models may not fit observed data well
  - narrow range of S
  - imprecise estimates
- Don’t account for changes in survival rates (R/S)
  - environmental effects
  - changes through time
- Only accounts for density dependence in production

Summary – Replacement lines

- **Replacement lines** represent the recruits per spawner needed to replenish a population on average
- Deals with **recruitment overfishing** (know def. from before)
- For semelparous species like Pacific salmon, they can be used with SR model to estimate MSY reference points
  - [know conceptually how to get MSY from stock recruitment model & replacement line]
- For most species, replacement lines can be drawn using SSB/R models in conjunction with Stock-Recruit models
  - Higher fishing → higher slope of replacement line
- Biological reference points for management can be derived from these concepts

Maximum Likelihood Estimation

Readings:

Haddon 2011 (Section 3.4)

Acknowledgements: M. Wilberg, R. Latour

Announcements

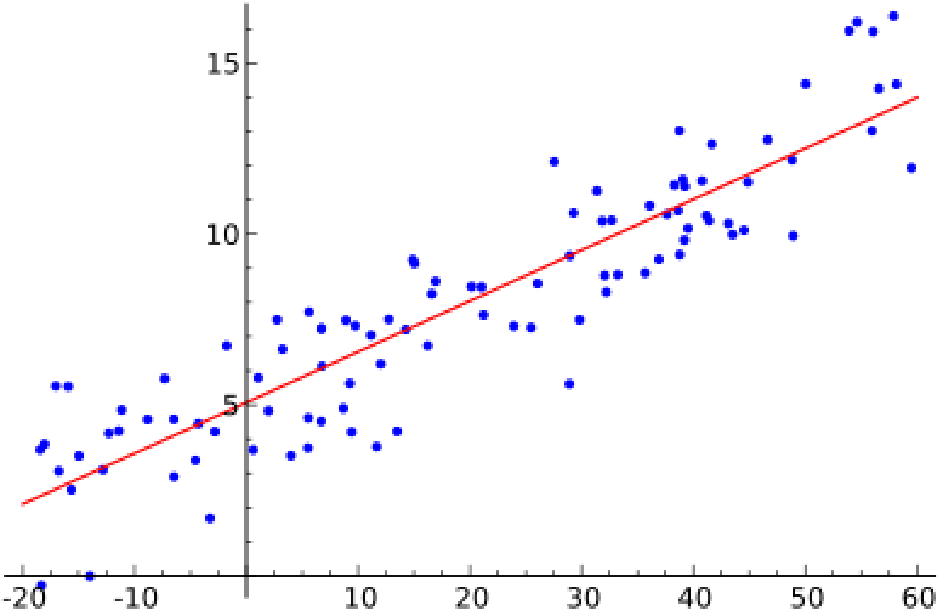
- 558 project synopses → Due Wednesday, 3/18 (11:59pm)

How do we fit models to data?

- Least squares (see refresher below)
- Maximum likelihood
- Bayesian methods (not for this class)

$$residual = \varepsilon_i = Y_i - \widehat{Y}_i$$

$$RSS = \sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2$$



Maximum Likelihood

- Alternative way of fitting models and getting estimates…
- Need a quick review of probability

- Simple intro/explanation:
  - https://towardsdatascience.com/probability-concepts-explained-maximum-likelihood-estimation-c7b4342fdbb1

Probability

Probability – Frequency View

- Probability is long-run relative frequency (from a random process)
- Same as relative frequency in the population

- Examples
  - Dice toss p(1) = p(2) = …=p(6) = 1/6
  - Coin flip p(Head) = p(Tail) = .5

Probability

- Some characteristics
  - Between 0 and 1
  - Probabilities must be non-negative.
  - The sum of probabilities over all possible mutually exclusive outcomes must equal one (e.g., dice)
  - If two events, A and B, are mutually exclusive, the probability of observing either of the events is the sum of their individual probabilities.

$$P(A \cup B) = P(A) + P(B)$$

"Union"
P(A or B)

Mutually exclusive: related in such a way that each thing makes the other thing impossible; not able to be true at the same time or to exist together

E.g., what is P(1 or 2) when rolling a die?

Independence

- **Joint probability** is the probability that two (or more) different events will occur
- Statistical **independence** means that knowledge of one event provides no information about the probability that another event will occur

$$P(A,B) = P(A)P(B)$$

intersection

P(A,B) is the same as P(A and B) or $\mathrm{P}(A \cap B)$

E.g., what is P(1 and 2) when rolling two die?

Independence

- We usually assume that observations are in some way independent of one another when we fit models

$$P(y_1, y_2, \ldots, y_n) = P(y_1)P(y_2) \ldots P(y_n) = \prod_{i=1}^{n} P(y_i)$$

- E.g., in linear regression, we assume that the errors are independent

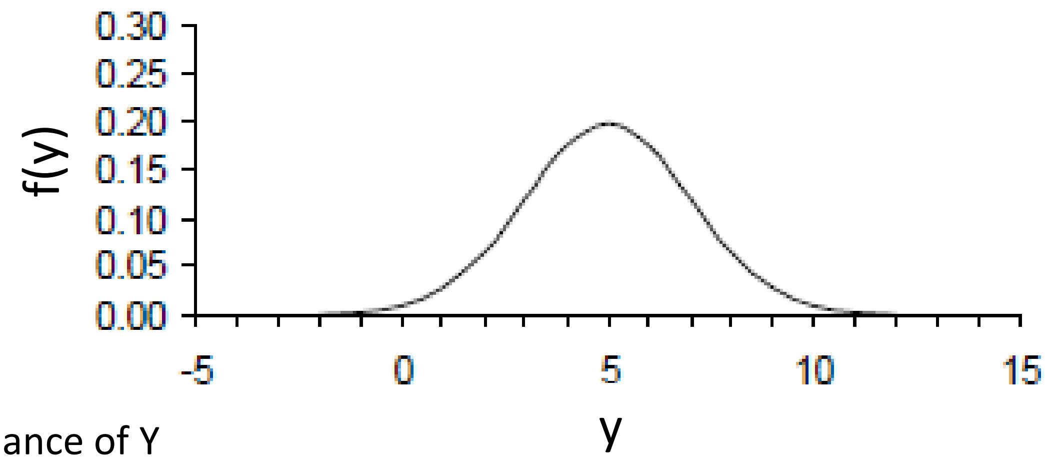
pdf for the normal distribution

- **Probability density function (pdf)** describes the probability of an event occurring for a *continuous* distribution.
  - The total area under the curve = 1
- The normal distribution is one of the most common pdfs:

$$f(y|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{y-\mu}{\sigma}\right)^2}$$

$$E(Y) = \mu$$

$$Var(Y) = \sigma^2$$



E(Y) – "expected value of Y"; Var(Y) – variance of Y

Maximum Likelihood

Maximum Likelihood Estimation (MLE)

- ML techniques are a powerful and flexible method for parameter estimation
- Alternative to least squares methods
- **This technique involves finding the parameters that maximize the probability of generating the observed data** (this is how we define “best fit” with ML)
- Likelihood ≠ probability

ML - Developed by R. A. Fischer (published this while a junior in college!)

Maximum Likelihood Parameter Estimation

- ML parameter estimates maximize the probability of observing the data
- Reverses the role of parameters and data (compared to probability)
- **Treat the data as fixed and find parameters that maximize the probability of observing those data**

$$L(parameters \mid data) = P(data \mid parameters)$$

Likelihood (L) is conditioned on the data

Likelihood for continuous distributions

- For a continuous distribution, the likelihood is calculated as:

Probability density function

$$L(\theta \mid Y_i) = \prod_{i=1}^{n} \mathrm{pdf}(Y_i \mid \theta)$$

Parameters

Product

Data

- “the likelihood of the parameter(s) θ (theta) is the product of the pdf values for each of the n observations Y_i given the parameter(s) θ”

Haddon 2011, p. 79

Example 1 – Normal distribution

- To get probability of whole data set (given parameters):
  - Multiply prob. of each data point together b/c independent
  - $P(y_{all}) = P(y_1) x P(y_2) x \ldots x P(y_n)$
- Use normal PDF for getting probabilities.
- Parameters for normal are μ and σ.

Data

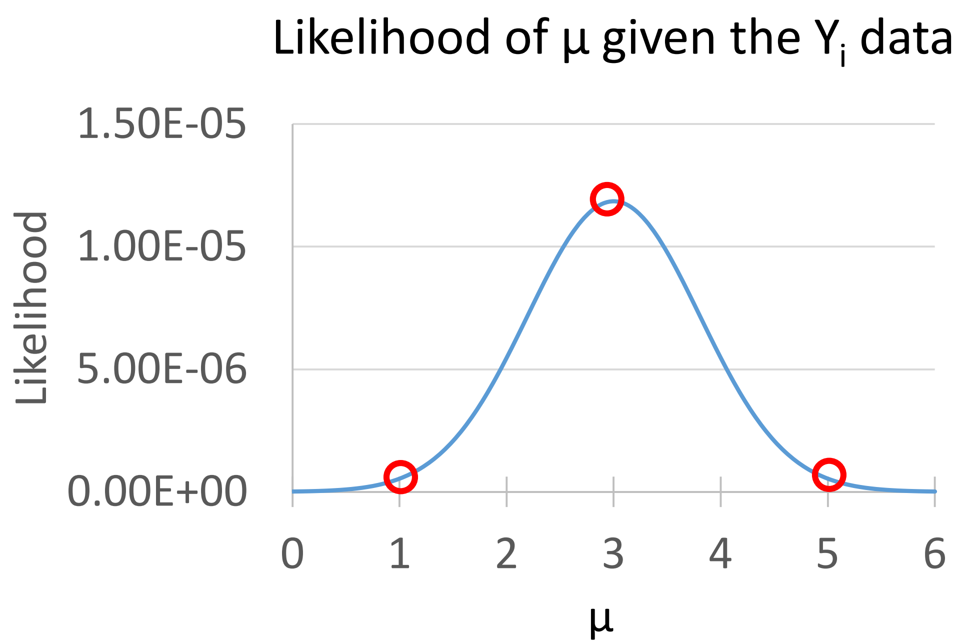
$$L(\mu, \sigma \mid y_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{(y_i - \mu)^2}{\sigma}\right)}$$

Parameters

Maximum Likelihood Example 1

- You take a sample (n=4) from a population and you want to estimate the population average (μ) using ML.
- Samples of Y_i = 0, 2, 4, 6
- Calculate L for different values of μ (here, using sample SD as σ=2.6). Find the value of μ that would “maximize the likelihood”.

	$P(Y_i \mid \mu)$		
Y_i	**If μ=1**	**If μ=3**	**If μ=5**
0	0.127	0.027	0.001
2	0.127	0.127	0.027
4	0.027	0.127	0.127
6	0.001	0.027	0.127
L	5.35E-07	1.19E-05	5.35E-07
log(L)	-14.44	-11.34	-14.44
-log(L)	14.44	11.34	14.44



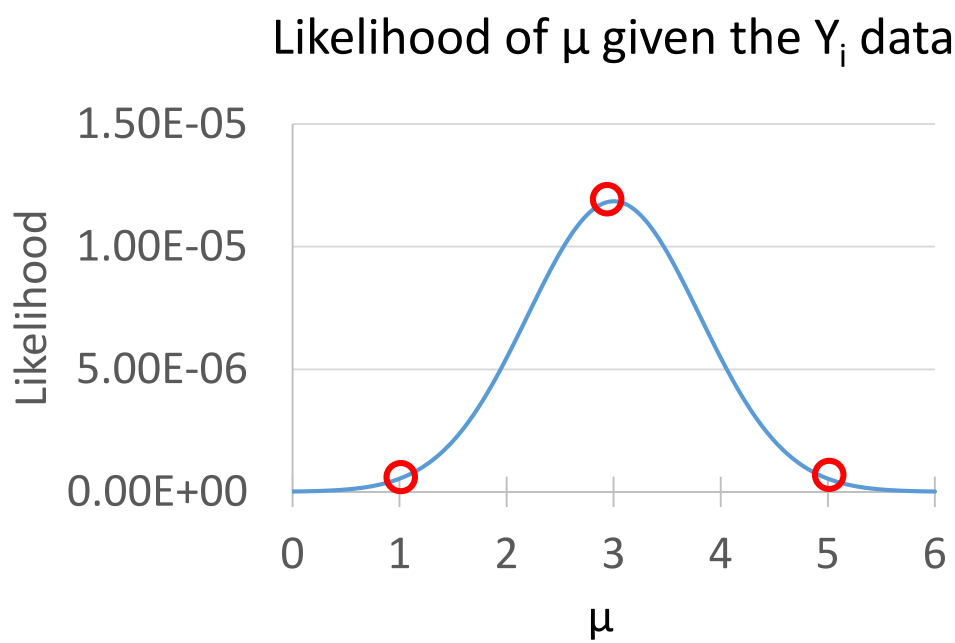
$$L(\mu,\sigma \mid y_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{(y_i-\mu)^2}{\sigma}\right)}$$ = $P(Y_1)xP(Y_2)xP(Y_3)xP(Y_4)$

Maximum Likelihood Example 1

- You take a sample (n=4) from a population and you want to estimate the population average (μ) using ML.
- Samples of Y_i = 0, 2, 4, 6
- Calculate L for different values of μ (here, using sample SD as σ=2.6). Find the value of μ that would “maximize the likelihood”.

Our ML estimate of μ is 3.

- This is the value that maximizes the likelihood on the graph
- note this matches our sample mean ($\bar{Y}$ = (0+2+4+6)/4 =3), which is an unbiased estimator of μ.



$$L(\mu,\sigma \mid y_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{(y_i-\mu)^2}{\sigma}\right)}$$ = $P(Y_1)xP(Y_2)xP(Y_3)xP(Y_4)$

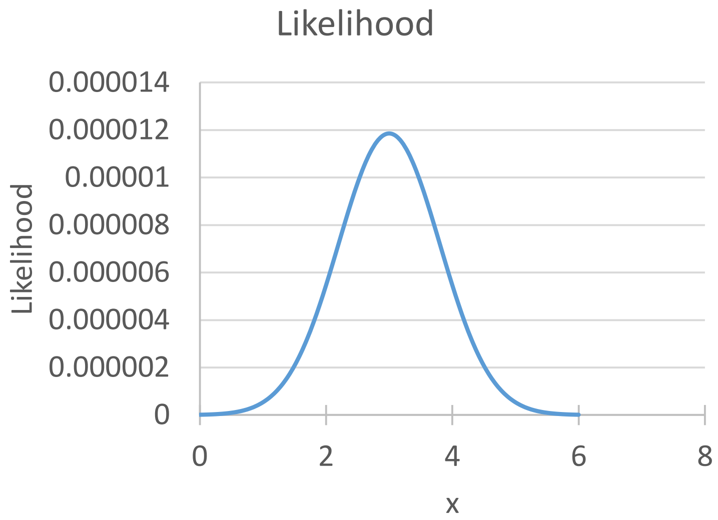
Properties of Likelihoods

- Likelihoods do not need to sum to one
  - They are NOT probabilities
- Likelihoods are a *relative* (not absolute) measure of model fit
- Calculating products is difficult, so logs are often used

Likelihood Nomenclature

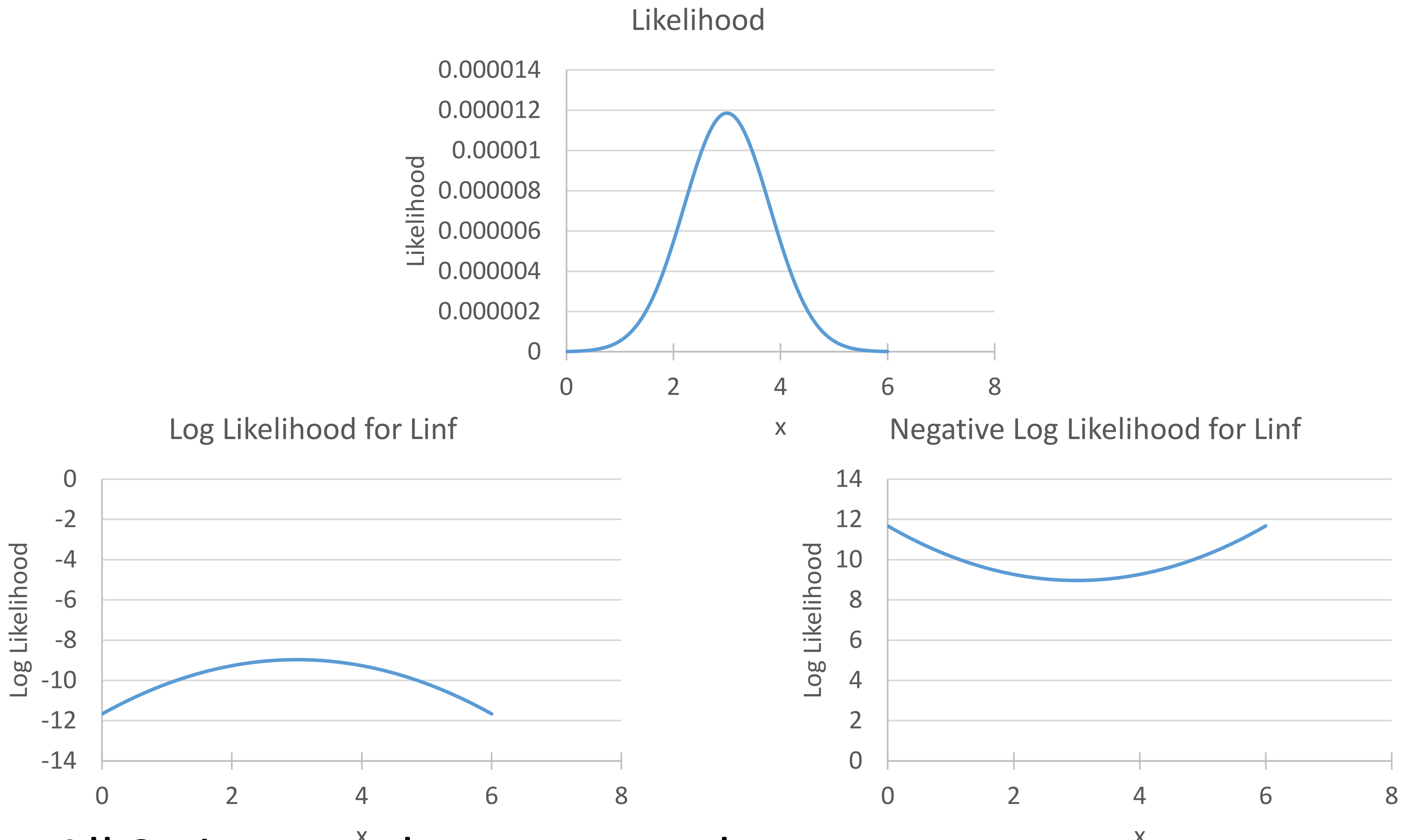
- Terms used to identify likelihood methods can be confusing:
- **Likelihood** $L(\theta|data)$ [where θ are parameters]
  - Remember: goal is to maximize likelihood
- **Log likelihood** $\log(L(\theta|data))$ or LL
  - Used because easier to deal with sums than with products.
- **Negative log likelihood** $-\log(L(\theta\ |data))$ or –LL or NLL
  - Used for historical reasons (e.g., folks used to minimizing Sums of Squares), and software more common previously for minimizing functions

ML Example 1, n=4



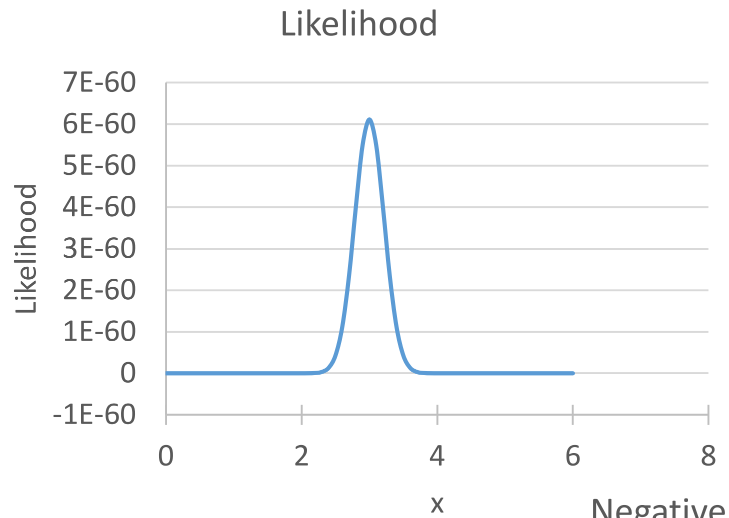
- What would LL and –LL look like?

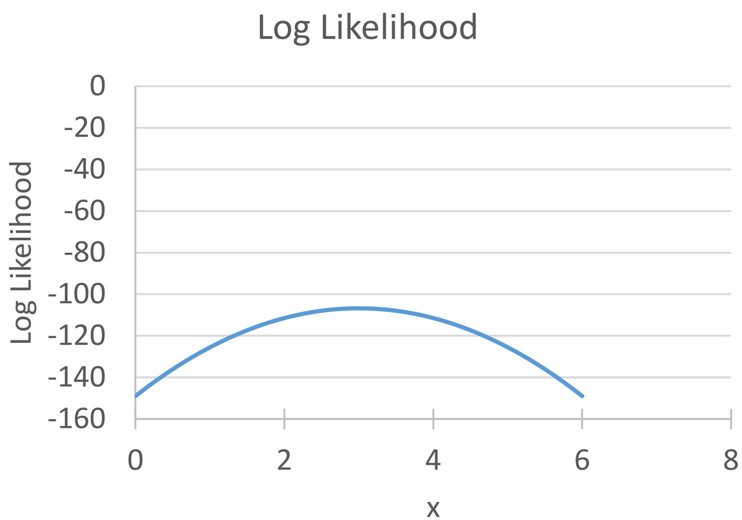
ML Example 1, n=4

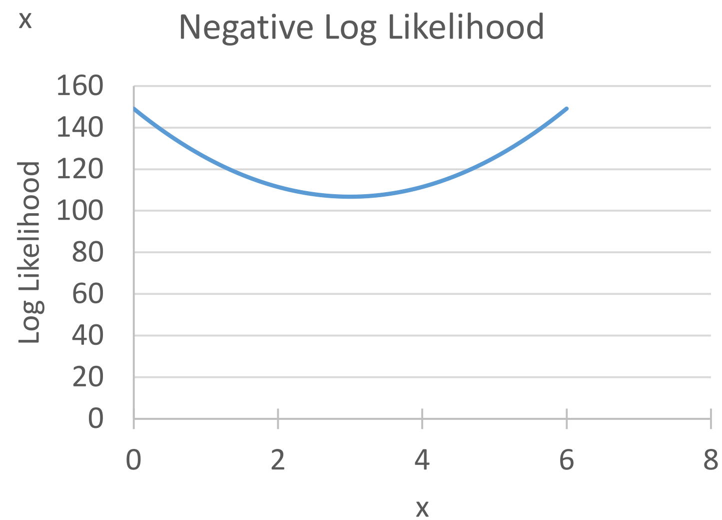


- All 3 give you the same result.
- Profile shapes give you some information of the relative likelihood of different results

ML Example 2, n=48







- As sample size increases, the amount of information increases, so you have a narrower range of values near the “best estimate” → **greater confidence in value**

-LL eqns. – Normal distribution

- Negative log likelihood (negLL, -LL) [to be minimized!]
  - Done to facilitate calculations

$$\boxed{\text{negLL} = \sum_{i=1}^{n}\left(\frac{1}{2}\log(2\pi) + \log(\sigma) + \frac{1}{2}\left(\frac{y_i - \mu}{\sigma}\right)^2\right)}$$

- Negative log likelihood with constants removed:

$$\boxed{\text{negLL} = \sum_{i=1}^{n}\left(\log(\sigma) + \frac{1}{2}\left(\frac{y_i - \mu}{\sigma}\right)^2\right)}$$

μ would be whatever your model is!
e.g., $\mu = \beta_0+\beta_1X$

Properties of MLEs

- **Invariant to transformations**
- Asymptotically efficient (**lowest possible variance**)
- Asymptotically **normally distributed**
  - Useful for getting CI intervals
- Asymptotically **unbiased** (expected value of the estimated parameter equals the true value)
- **ML estimates will be the same as least squares estimates if errors are normal, additive, and with constant variance**

"Asymptotic" - deals with the conditions at infinitely large sample sizes.

Potential Challenges of ML Methods

- Programming may be required because usually problem specific
- Likelihood equations need to be worked out for a given problem
- Numerical techniques are often required to find MLE
- MLEs may be biased for small samples and asymptotic benefits may not apply to small samples

Summary 1: Maximum likelihood

- Goal: find the parameter values that make the observed data most likely
  - *Maximize* likelihood (L), *maximize* log-likelihood (LL), or *minimize* negative log-likelihood (-LL) → give same result
- Likelihood is different than probability
  - Probability: Knowing parameters → Prediction of data
  - Likelihood: Observation of data → Estimation of parameters
- ML is an alternative to least squares for fitting models
  - ML and Least Squares give same results if errors are normal and additive with constant variance
- Writing a likelihood function relies on the equation for your distribution
  - See Haddon 2011 for examples, or lecture for eqns.

Model fitting with nonlinear optimization

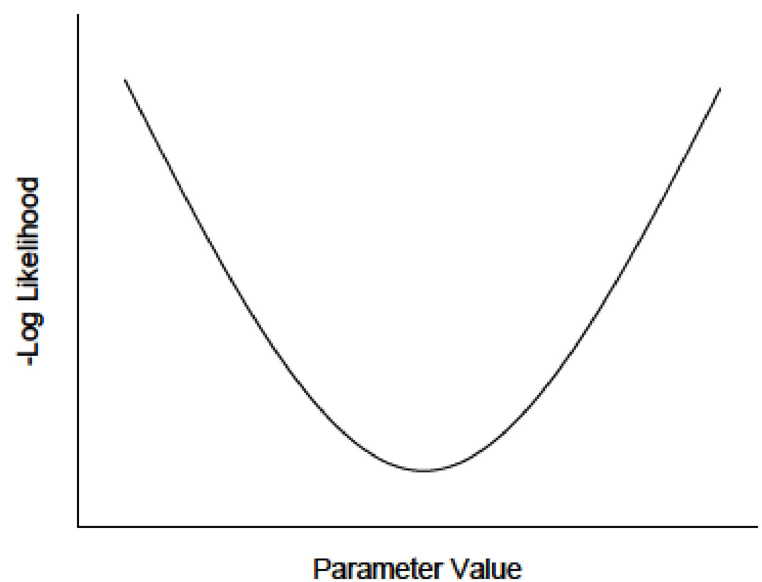
Readings:

Haddon 2011 (Section 3.4)

Acknowledgements: M. Wilberg, R. Latour

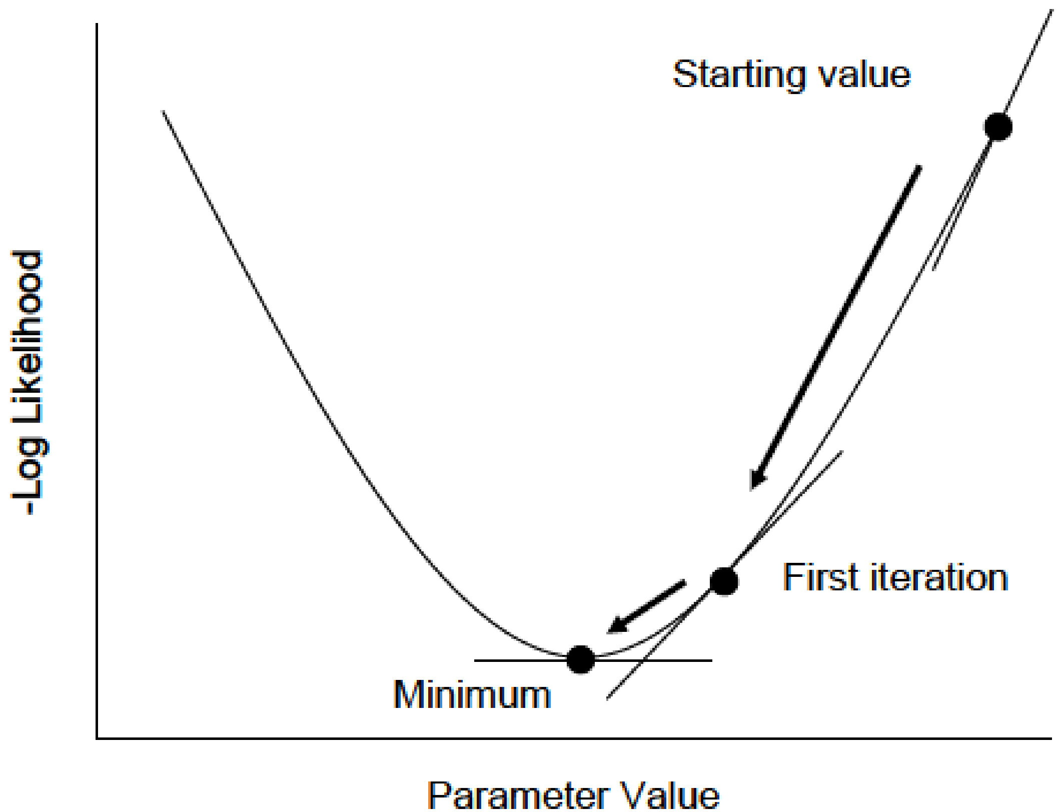
How to find minimum of -LL? Nonlinear Optimization

- **Nonlinear optimization** is the general term for trying to find the maximum or minimum of a function (e.g., likelihood fxn)
- Numerical solution (vs. Analytical solution) – approximate solution found through iterative searching

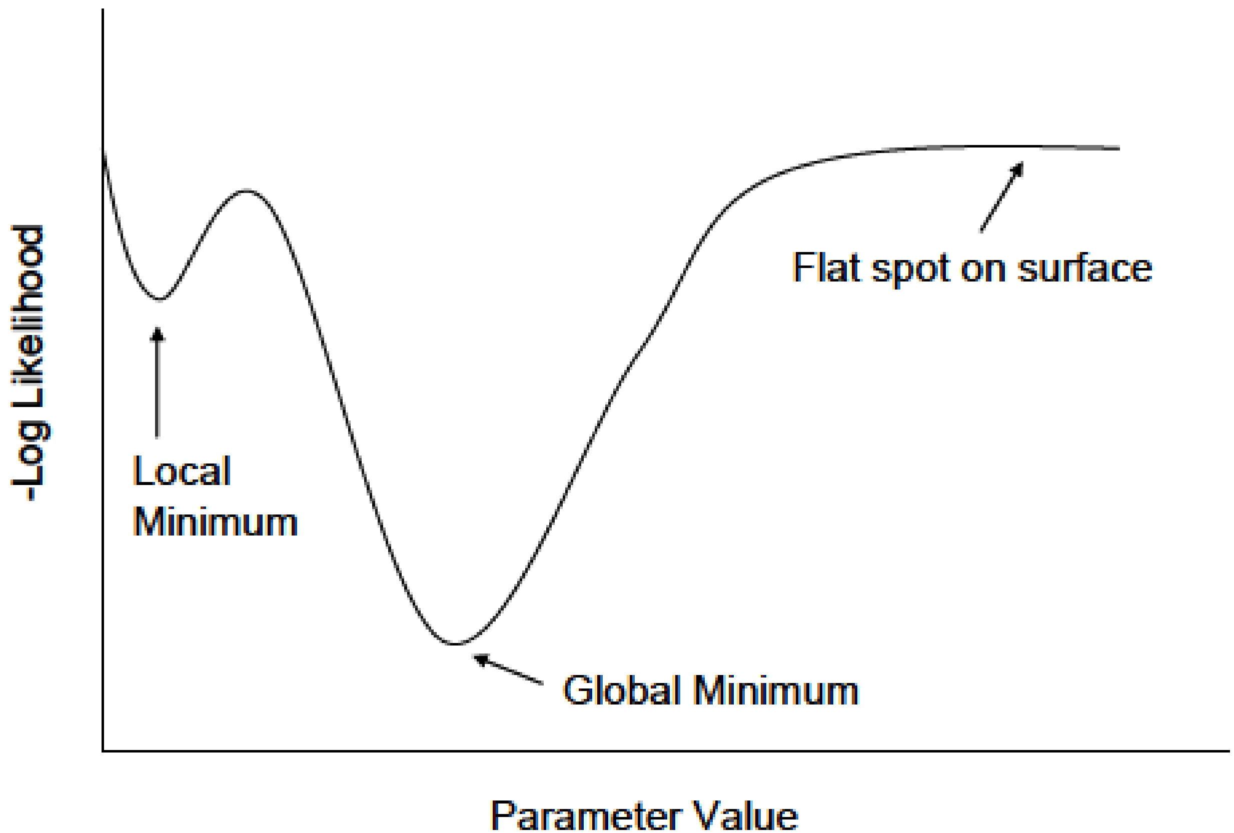


Minimization

- R has functions to do this: e.g., optim()



Potential Nonlinear Optimization Problems



optim() in R

- optim() uses an algorithm that relies on derivatives to find the optimal solution
- Some common outcomes:
  - it finds the “best” solution
  - it reaches its limit of iterations (can change *maxit*)
  - Misc. errors
- Check that convergence criteria were met:
  - **$convergence = 0 → SUCCESS!**
  - $convergence = 1 → max iterations reached

Potential Nonlinear Optimization Problems

1. Improper starting values
2. Model specification or coding error
3. Negative log likelihood function may be undefined for some parameter values
4. Parameterization problems (scaling, correlated parameters)
5. Uninformative data

1. Starting values

- Parameters need to be provided "good" starting values for all nonlinear optimization routines
  - The starting value should be: of the same sign and similar magnitude the expected solution
- Obtain starting values by eye
- Use several sets of starting values
- Can increase max number of iterations (using *maxit*) to allow longer search
  - optim(…, control=list(maxit=10000))

2. Model Specification and Coding Errors

- R cannot detect mistakes in your model
- Make sure that model predictions make sense
  - Do the numbers make sense?
  - Graph predictions over the data
- Try specific sets of parameter values for which you know the correct answer
- Fit model to simulated data

3. negLL function may be undefined for some parameter values

- Common example:
  - Sigma (σ) can’t be negative – can’t have a negative SD of residuals
  - Warning

    ```
    Warning messages:
    1: In log(sigma) : NaNs produced
    ```

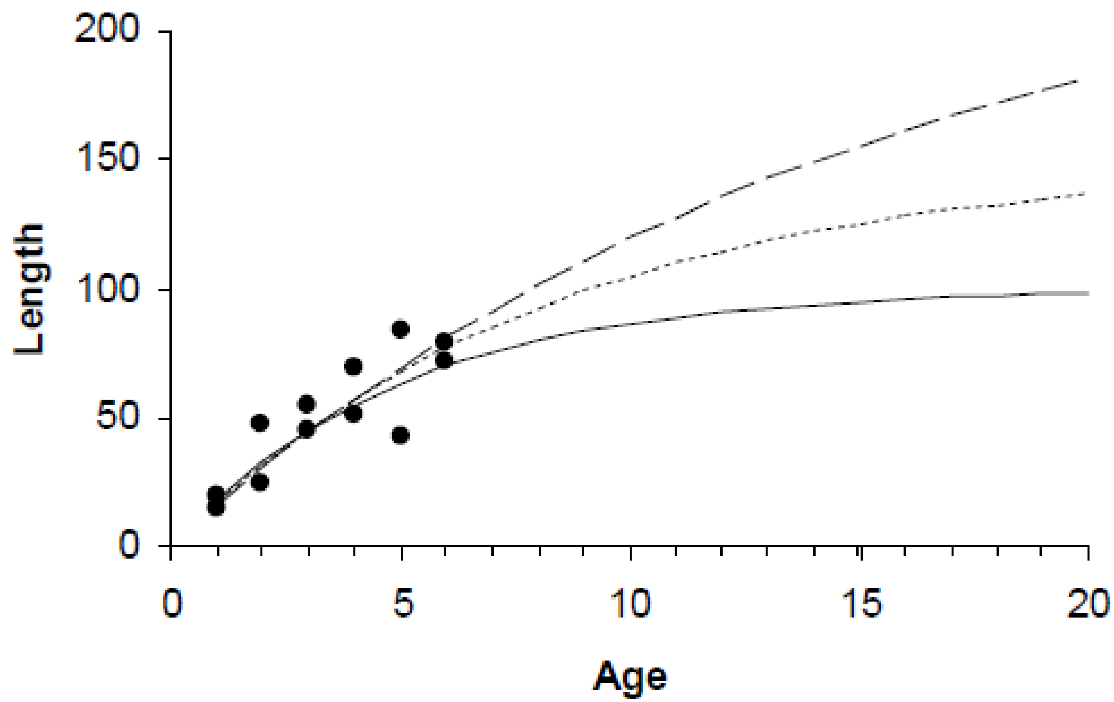
  - Possible solutions:
    - Ignore
    - Use a different method that allows constraints (e.g., "L-BFGS-B“)
    - Estimate log(σ) which can be negative, then backtransform

4. Parameterization Problems

- Many models can be written in several forms, each called a parameterization
- Parameters should have low correlations with one another
- Parameters should be of similar magnitude (e.g., avoid 0.000000529 and 3.54)
    - → estimating parameters in log space can help (e.g., ln(5.3E-07) = -14.45

5. Uninformative data

- The data may have little or no information about one or several parameters
- Only able to statistically estimate n-1 parameters



Summary 2 - Nonlinear Optimization

- **Nonlinear optimization** - general term for trying to find the maximum or minimum of a function (e.g., likelihood function).
  - Numerical solution used when can't Analytical solution)

Potential problem with nonlinear optimization:

1. Improper starting values
2. Model specification or coding error
3. Negative log likelihood function may be undefined for some parameter values
4. Parameterization problems (scaling, correlated parameters)
5. Uninformative data

Occupancy Modeling

Reading:

MacKenzie et al. 2002

Announcements

- Monday, April 1 – No Class
- Wed, April 3 - In-class exam 2
- Tues, April 9 – Take-home exam 2 due (will have 10-14 days to complete.

Occupancy

- Occupancy (ψ)
  - Measuring the presence or absence of a species in a location
  - (Alternative to abundance)
- Uses for occupancy data/studies? Examples?

Occupancy

- Occupancy (ψ)
  - Measuring the presence or absence of a species in a location
  - Given as a probability or frequency
  - (Alternative to abundance)
- Uses for occupancy data/studies? Examples?
  - Studies of species distributions
    - What factors affect presence?
    - Habitat modeling
  - Metapopulation dynamics
    - where site (or patch) occupancy is related to site (or patch) characteristics
    - Can look at extinction and colonization probabilities

Occupancy

- How might someone estimate occupancy (ψ) if we had data from a survey at *n* sites?

$$\psi = \frac{n_{occupied}}{n_{total}}$$

Problem with this?

Problem?

- Was species *present and not detected* or was it absent?
- The measure of occupancy is **confounded** with the detectability of the species
  - **Confounded**: an inability to separate multiple factors potentially contributing to an observed pattern.
- Using presence/absence data alone will UNDERESTIMATE or OVERESTIMATE true occupancy.

Problem?

- Was species *present and not detected* or was it absent?
- The measure of occupancy is **confounded** with the detectability of the species
  - **Confounded**: an inability to separate multiple factors potentially contributing to an observed pattern.
- Using presence/absence data alone will UNDERESTIMATE or OVERESTIMATE true occupancy.

- **Detectability**
  - refers to the reality that it is very common for animals and even entire species to be missed and go undetected.
  - Expressed as a probability of detection

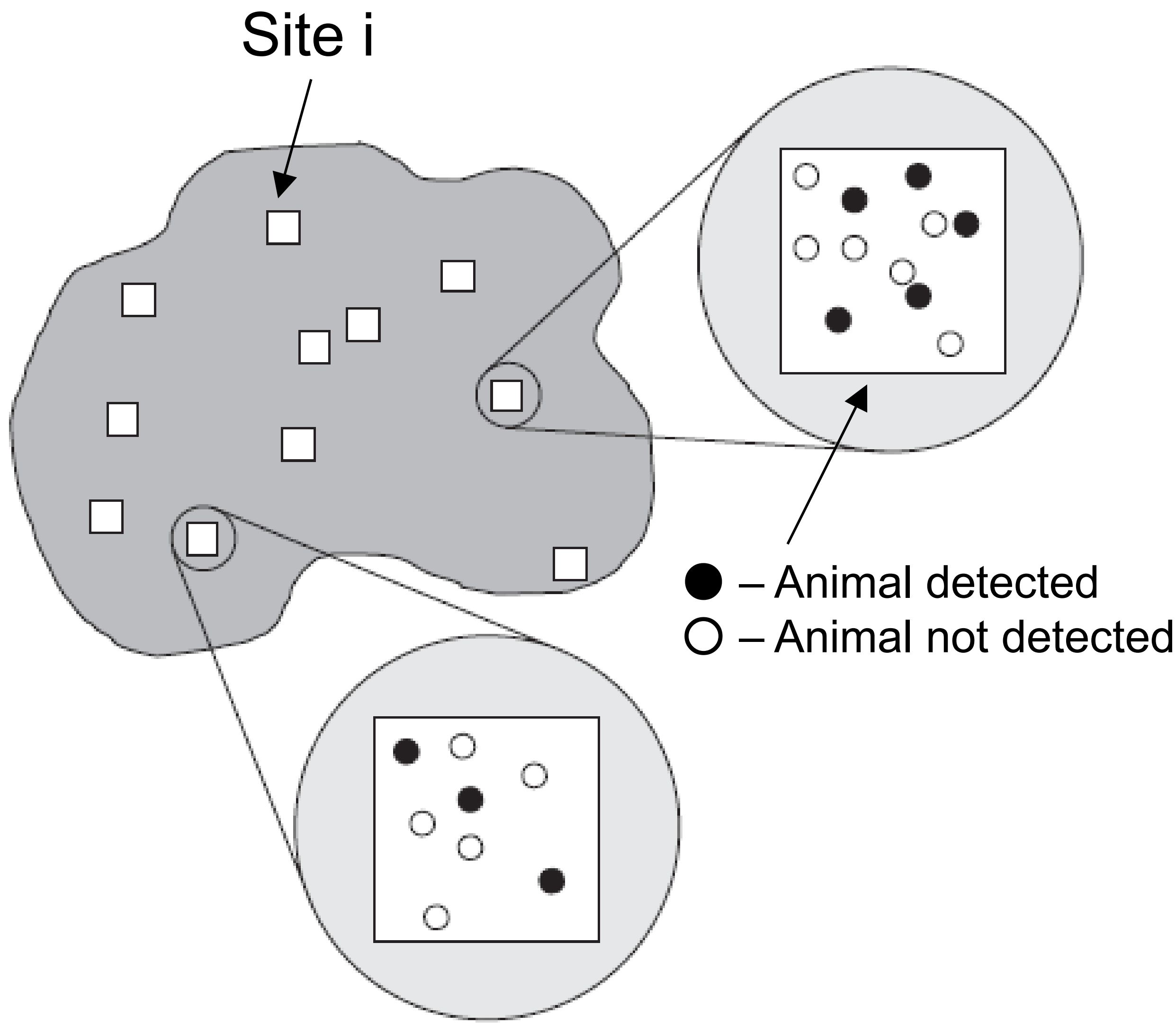


FIGURE 1.1 Illustration of the two critical aspects of sampling animal populations, spatial variation and detectability. The shaded region indicates the area or population of interest, with the small squares representing the locations selected for sampling. Within each sampling location, animals will be detected (filled circles) or undetected (hollow circles) during a survey or count.

Occupancy models

- **Models that deal with the problem of imperfect detectability**
- **Need info from repeated observations at each site to estimate detectability!**
  - Can be multiple visits, multiple observers, single observer & multiple passes, etc.
- Detectability can vary with site characteristics (e.g., habitat variables) or survey characteristics (e.g., weather)
  - Occupancy relates only to site characteristics

Effect of imperfect detection

Consensus models
(assume perfect detection)

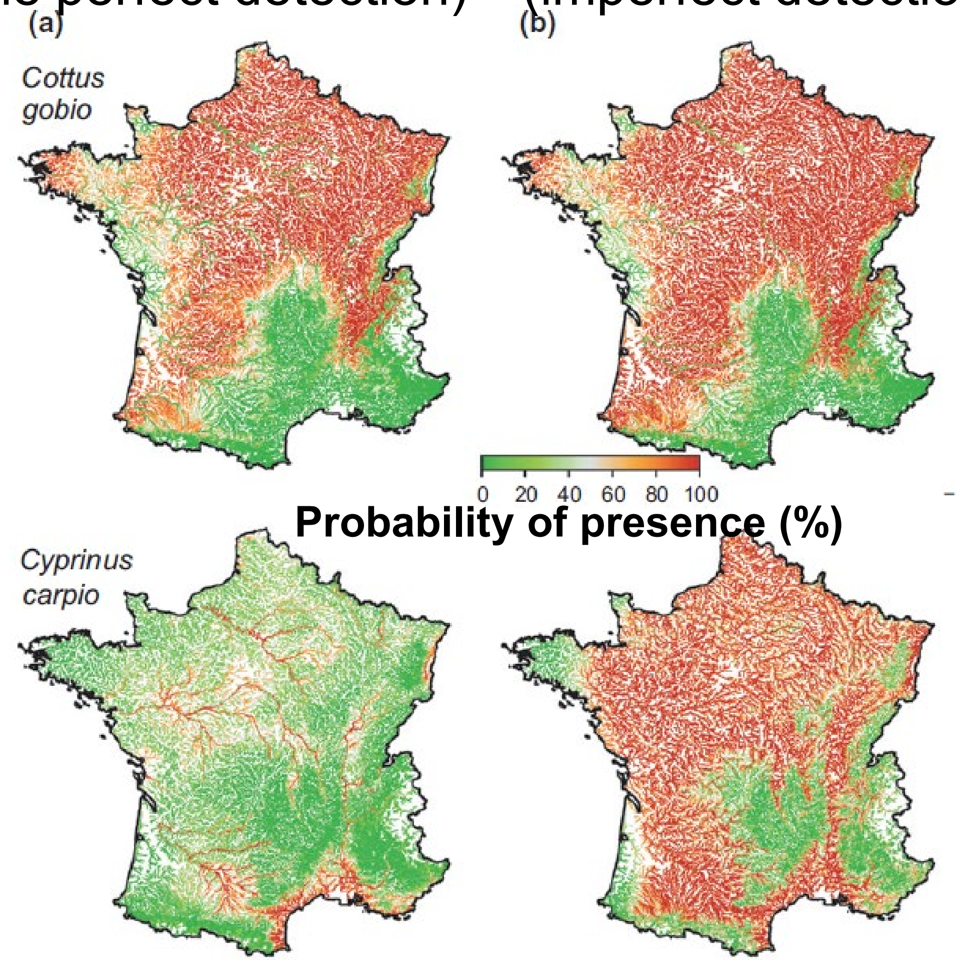
Occupancy models
(imperfect detection)

European bullhead

High detectability

Common carp

Low detectability



What patterns do you see?

Comte and Grenouillet 2013

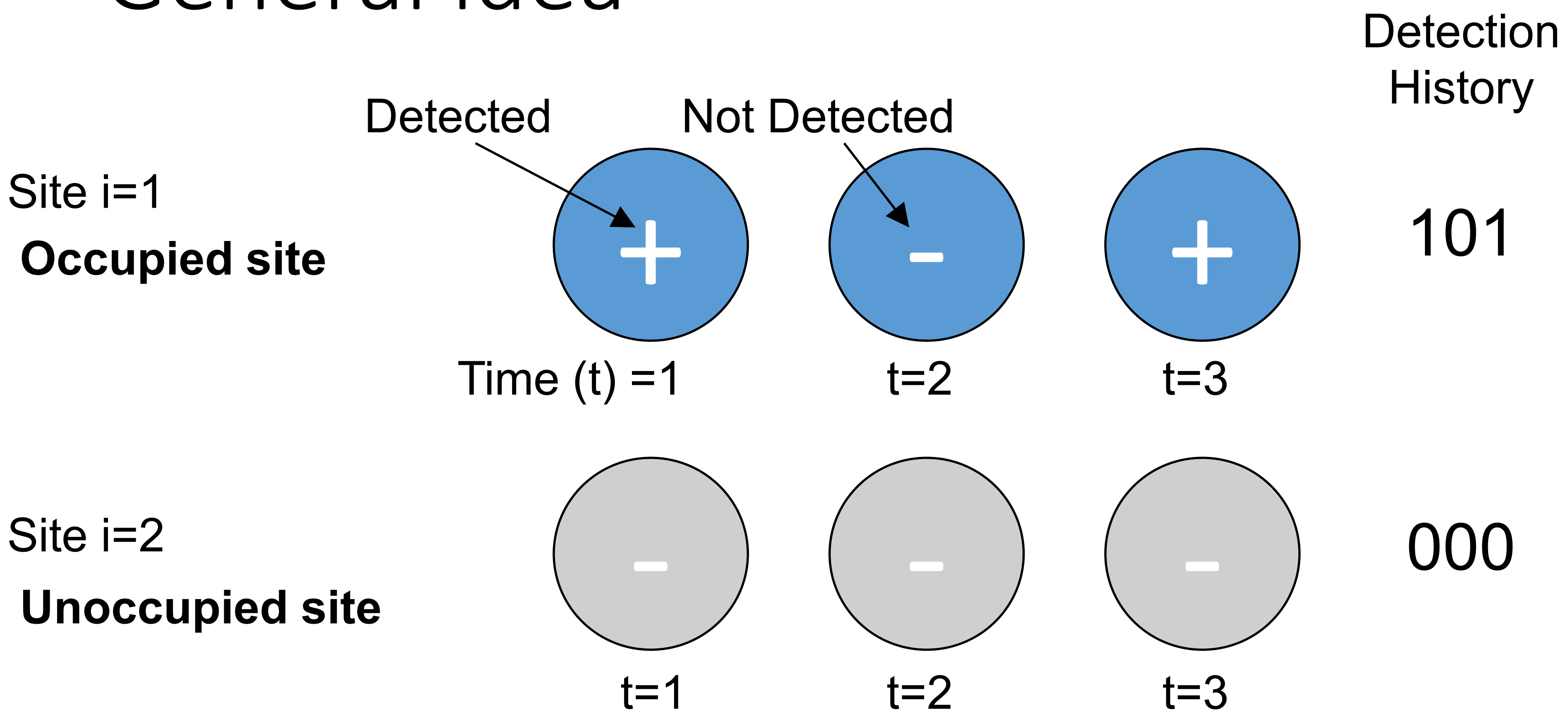
Next installment of "Google This"!

- "occupancy modeling funny"

Data needs

- Detection history
  - record of pres/abs data for each site (multiple occasions)
- Other variables of interest
  - Environmental variables pertaining to site
  - Variables pertaining to detection during repeat sampling events or surveys (e.g., weather)

General idea



Main parameters

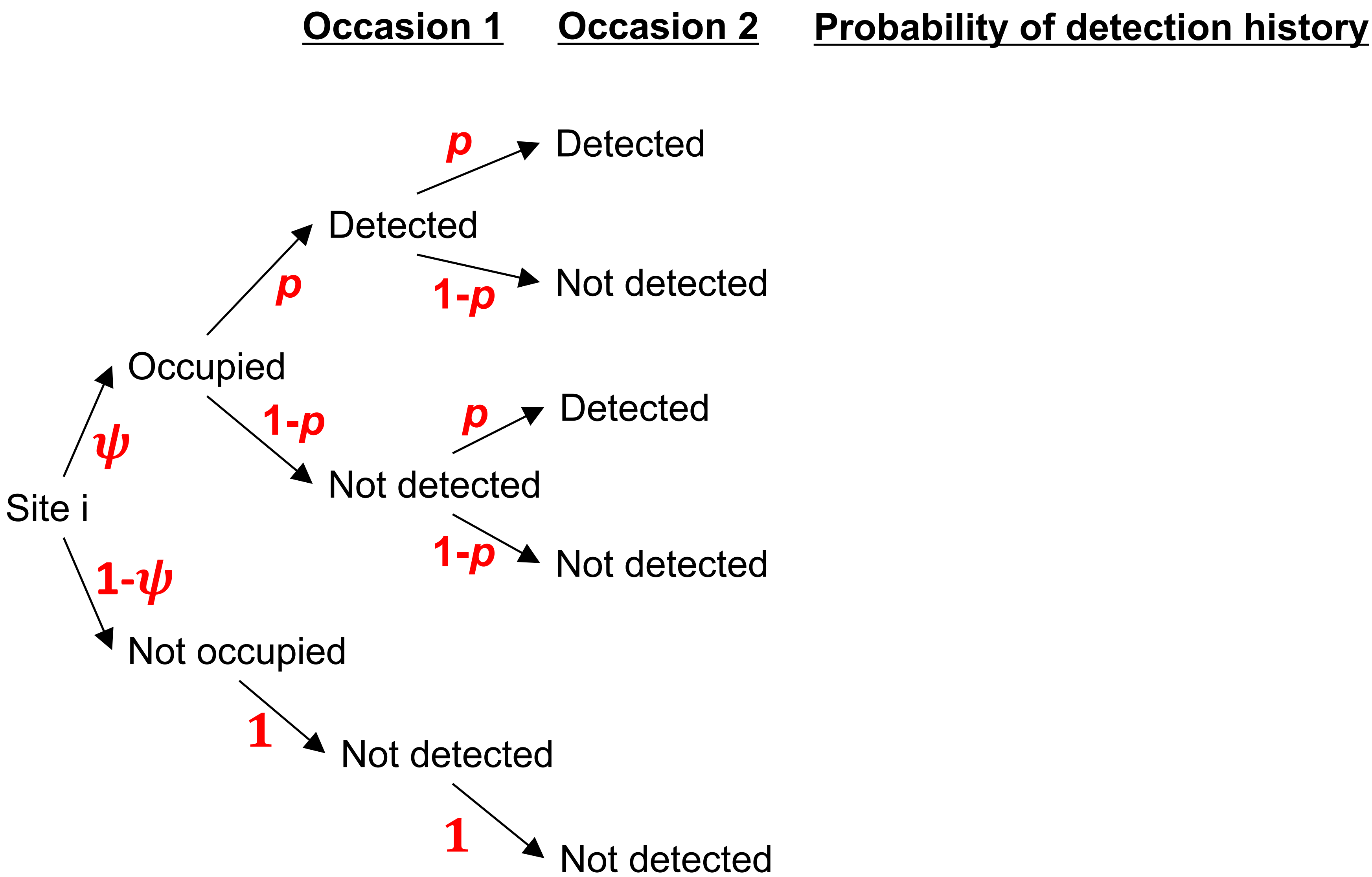
- **ψ – probability of occupancy** = probability of the species being present at site i
- **p – conditional detection probability** = probability of detecting the species, given that it was present
  - This takes advantage of the multiples detections (e.g., over time t)

Single seas. occupancy (2 sampling occasions)

	Occasion 1	Occasion 2	Probability of detection history
Site i			

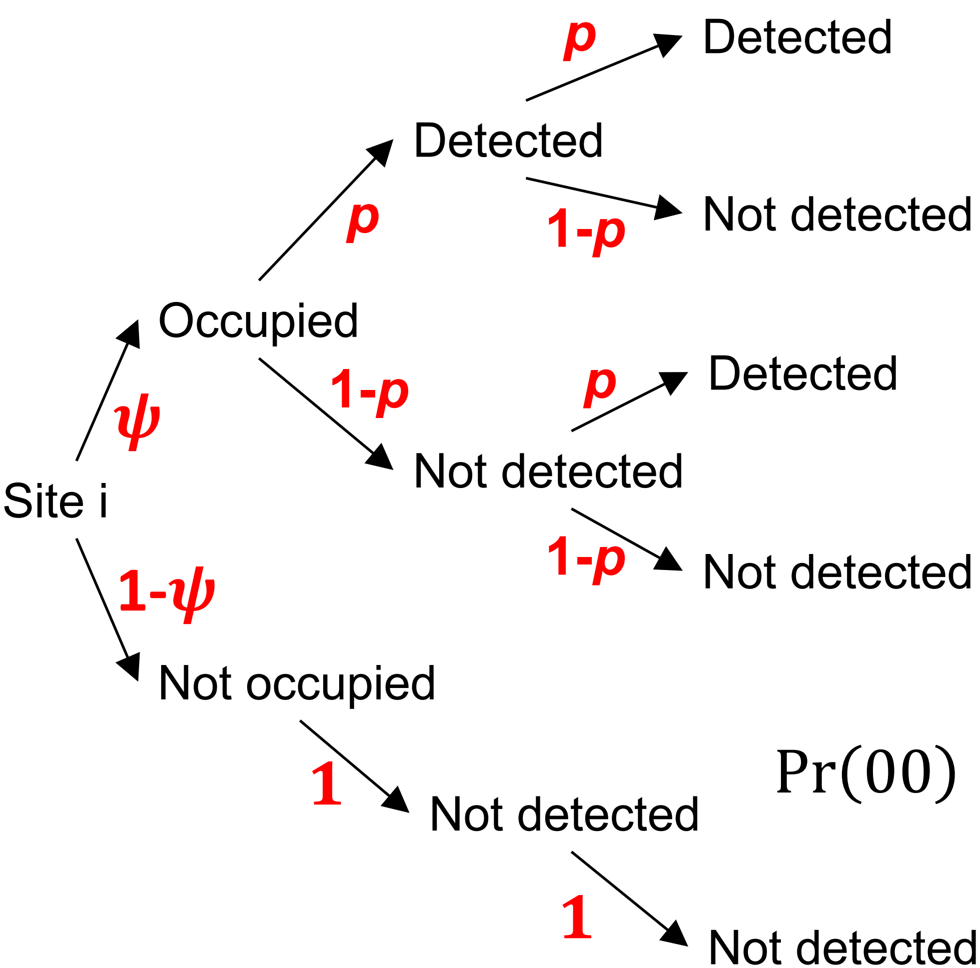
Modified from D. Barton

Single seas. occupancy (2 sampling occasions)



Modified from D. Barton

Single seas. occupancy (2 sampling occasions)



Probability of detection history

$$\Pr(11) = \psi[p_1 p_2]$$

$$\Pr(10) = \psi[p_1(1 - p_2)]$$

$$\Pr(01) = \psi[(1 - p_1)p_2]$$

$$\cancel{\Pr(00)} = \psi[(1 - p_1)(1 - p_2)]$$

"false absence"

$$\Pr(00) = \psi[(1 - p_1)(1 - p_2)] + (1 - \psi)$$

$$\cancel{\Pr(00)} = (1 - \psi)$$

"true absence"

Modified from D. Barton

Draw hypothetical example…

Suppose you sample 5 sites to detect the endangered tidewater goby,
And you have 2 survey occasions per site:

Likelihood and log likelihood

Probability of detection history

n_{01} = number of sites with detection history of 01

$$\Pr(11) = \psi[p_1 p_2]$$

$$\Pr(10) = \psi[p_1(1-p_2)]$$

$$\Pr(01) = \psi[(1-p_1)p_2]$$

$$\Pr(00) = \psi[(1-p_1)(1-p_2)] + (1-\psi)$$

- **Likelihood** of the dataset, given N total sites ($N=n_{11}+n_{10}+n_{01}+n_{00}$), is:

$$L(\psi, p|N) = \Pr(11)^{n_{11}} \Pr(10)^{n_{10}} \Pr(01)^{n_{01}} \Pr(00)^{n_{00}}$$

- And the **log likelihood** is:

$$\boldsymbol{\log L(\psi, p|N)} = n_{11}\log(\psi[p_1 p_2]) + n_{10}\log(\psi[p_1(1-p_2)]) + n_{01}\log(\psi[(1-p_1)p_2]) + n_{00}\log[\psi[(1-p_1)(1-p_2)] + (1-\psi)]$$

Modified from D. Barton

How to add covariate effects?

- Model ψ or p as a function of covariates (i.e., logistic regression)

Generic parameter (representing either ψ or p)

$$\theta = \frac{e^{\boldsymbol{X\beta}}}{1+e^{\boldsymbol{X\beta}}} = \frac{e^{\beta_0+\beta_1 X_1 \dots}}{1+e^{\beta_0+\beta_1 X_1 \dots}}$$

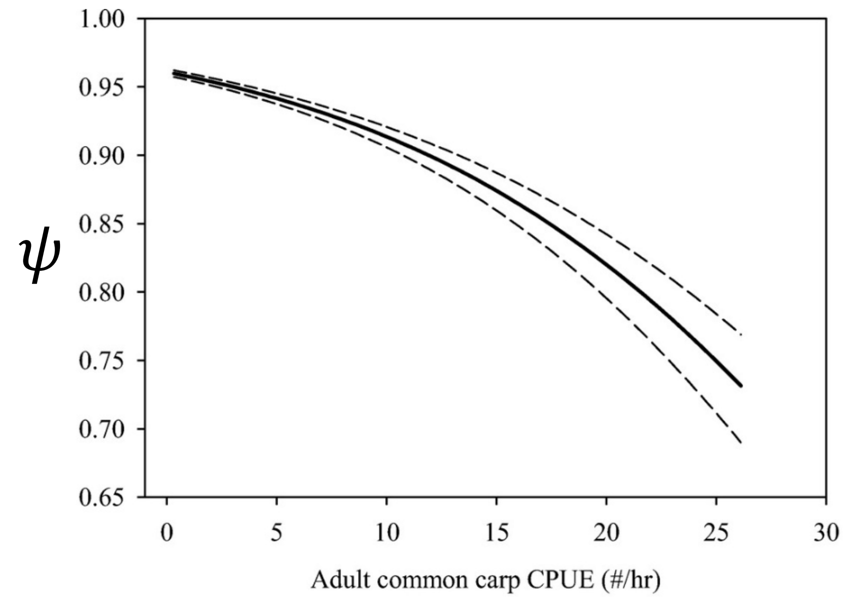
- Examples of covariates?

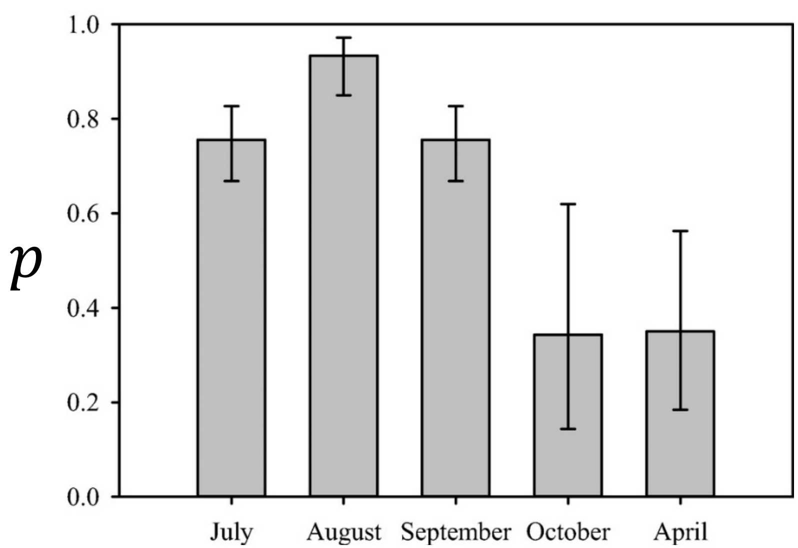
- Covariates can be for sites or for sampling occasions (aka surveys)

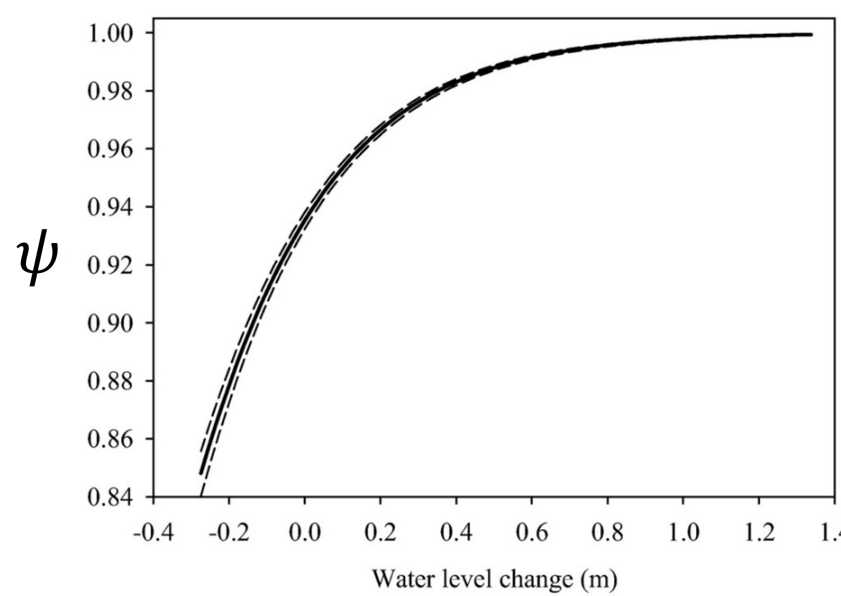
MacKenzie et al. 2002

Visualizing covariate effects

- Plot predicted psi and p as a function of your covariates (typically with 95% CI)
  - Can be continuous or categorical variables
- E.g.: Study on age-0 common carp occupancy in South Dakota







Weber and Brown 2019 (CJFAS)

How to add covariate effects?

- To get the average occupancy across sites (when modeled with covariates):
    - Need to average occupancy across covariate levels:

$$\overline{\hat{\psi}} = \frac{\sum_{i=1}^{N} \hat{\psi}_i}{N}$$

MacKenzie et al. 2002

Missing observations

- The model can handle missing observations
  - E.g., due to logistical issues or sampling design

$$\Pr(10\_11) = \psi p_1 (1 - p_2) p_4 p_5$$

- If you have some sites with only 1 sampling occasion:
  - Won't inform detection probability estimate,
  - but can apply estimates of p (and perhaps covariate effects) to those sites

Standard error of parameters

- Recommend bootstrapping
  - Resample N sites (with replacement) from the N sites
  - Refit model many times
  - Look at distribution of values
- Asymptotic standard error not recommended
  - ASE based on second order partial derivative of model likelihood

MacKenzie et al. 2002

Assumptions

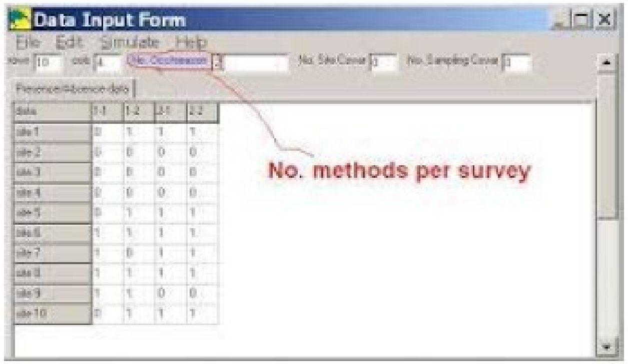
- Occupancy state is “closed” (no change in occupancy)
- Sites are independent
- No unexplained heterogeneity in occupancy
- No unexplained heterogeneity in detectability

How to meet assumptions?

- Occupancy state is "closed" (no change in occupancy)
  - Use life history or existing data to inform sampling
- Sites are independent
  - Select sites using probability based sampling scheme (e.g., simple-random; stratified-random)
- No unexplained heterogeneity in occupancy
  - Collect info about important variables
- No unexplained heterogeneity in detectability
  - Collect info about important variables

Occupancy model software

Program PRESENCE



Program MARK

R package Unmarked

Example

Spring peeper

American toad

- Covariates
  - Temperature
  - Habitat type
- 29 Sites
  - visited 2-66 times; mean ~9

MacKenzie et al. 2002

Example

What do the results of this table mean?
(Recall "Lab07 – AIC background.pdf")

- Can compare evidence for different hypotheses (ie models)!

TABLE 1. Relative difference in AIC (ΔAIC), AIC model weights (w_i), overall estimate of the fraction of sites occupied by each species ($\bar{\hat{\psi}}$), and associated standard error ($\mathrm{SE}(\bar{\hat{\psi}})$).

Model, by species	ΔAIC	w_i	$\bar{\hat{\psi}}$	$\mathrm{SE}(\bar{\hat{\psi}})$
American toad				
ψ(Habitat) p(Temperature)	0.00	0.36	0.50	0.13
$\psi(\cdot)$ p(Temperature)	0.42	0.24	0.49	0.14
ψ(Habitat) $p(\cdot)$	0.49	0.22	0.49	0.12
$\psi(\cdot)$ $p(\cdot)$	0.70	0.18	0.49	0.13
Spring peeper				
ψ(Habitat) p(Temperature)	0.00	0.85	0.84	0.07
$\psi(\cdot)$ p(Temperature)	1.72	0.15	0.85	0.07
ψ(Habitat) $p(\cdot)$	40.49	0.00	0.84	0.07
$\psi(\cdot)$ $p(\cdot)$	42.18	0.00	0.85	0.07

MacKenzie et al. 2002

Simulation results

Take-homes

- Better performance with greater sampling occasions
- Lower precision with lower sampling occasions
- Poorer performance at lower p with T=2

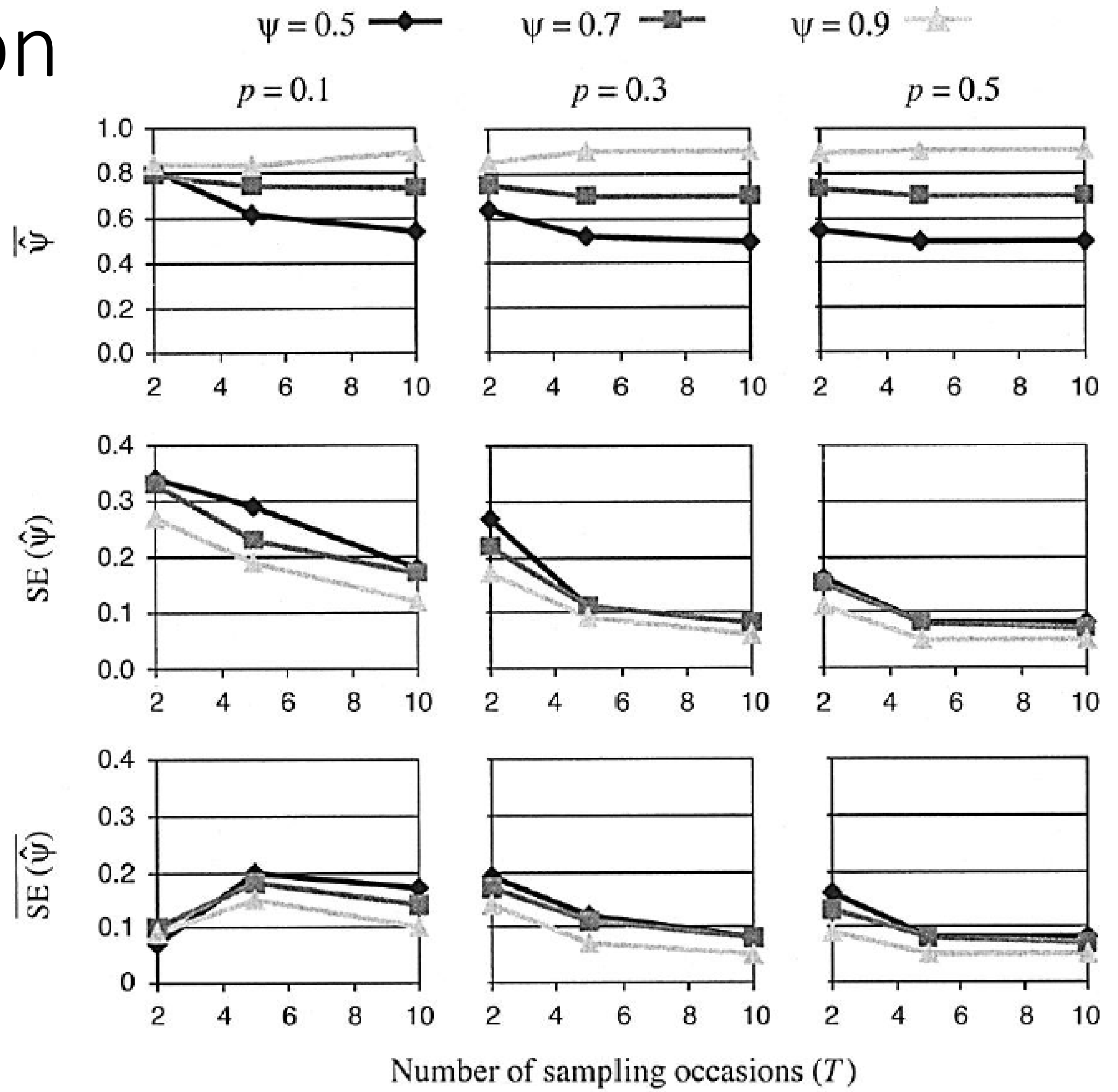


FIG. 1. Results of the 500 simulated sets of data for $N = 40$, with no missing values. Indicated are the average value of $\hat{\psi}$, $\overline{\hat{\psi}}$; the replication-based estimate of the true standard error of $\hat{\psi}$, $\mathrm{SE}(\hat{\psi})$; and the average estimate of the standard error obtained from 200 nonparametric bootstrap samples, $\overline{\mathrm{SE}(\hat{\psi})}$, for various levels of T, p, and ψ.

Considerations

- Recommended number of site visits?
  - More site visits = greater precision
  - If visiting only twice: best if occupancy is >0.7 and detection probability > 0.3
- Be skeptical of values if:
  - ψ close to 1, and p is low (<0.15)

MacKenzie et al. 2002

Extensions

- Having detection probability depend on abundance
  - Royle and Nichols 2003
- Multiple seasons
  - MacKenzie et al. 2003
- Multiple species
  - MacKenzie et al. 2004
- Multi-scale occupancy using multiple detection methods
  - Nichols et al. 2008
- …

MacKenzie et al. 2006 (book)

Ecology, 83(8), 2002, pp. 2248–2255
© 2002 by the Ecological Society of America

ESTIMATING SITE OCCUPANCY RATES WHEN DETECTION PROBABILITIES ARE LESS THAN ONE

DARRYL I. MACKENZIE,[1,5] JAMES D. NICHOLS,[2] GIDEON B. LACHMAN,[2,6] SAM DROEGE,[2] J. ANDREW ROYLE,[3] AND CATHERINE A. LANGTIMM[4]

Ecology, 84(8), 2003, pp. 2200–2207
© 2003 by the Ecological Society of America

ESTIMATING SITE OCCUPANCY, COLONIZATION, AND LOCAL EXTINCTION WHEN A SPECIES IS DETECTED IMPERFECTLY

DARRYL I. MACKENZIE,[1,5] JAMES D. NICHOLS,[2] JAMES E. HINES,[2] MELINDA G. KNUTSON,[3] AND ALAN B. FRANKLIN[4]

Journal of Animal Ecology 2004 **73**, 546–555

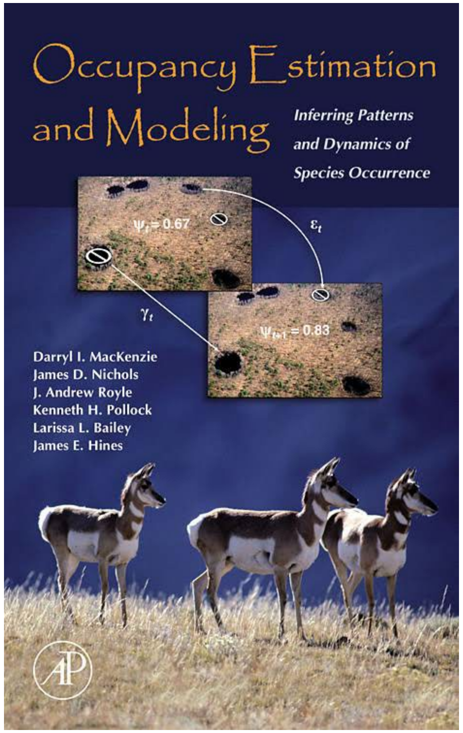
Investigating species co-occurrence patterns when species are detected imperfectly

DARRYL I. MACKENZIE*, LARISSA L. BAILEY† and JAMES. D. NICHOLS‡

Journal of Applied Ecology 2008, **45**, 1321–1329 doi: 10.1111/j.1365-2664.2008.01509.x

Multi-scale occupancy estimation and modelling using multiple detection methods

James D. Nichols[1]*, Larissa L. Bailey[1], Allan F. O'Connell Jr.[2], Neil W. Talancy[3], Evan H. Campbell Grant[1], Andrew T. Gilbert[4], Elizabeth M. Annand[5], Thomas P. Husband[3] and James E. Hines[1]



- MacKenzie et al. 2006 (book)
- Donovan, T. M. and J. Hines. 2007. Exercises in occupancy modeling and estimation. www.uvm.edu/rsenr/vtcfwru/spreadsheets/occupancy/occupancy.htm

E.g., Multi-season occupancy model

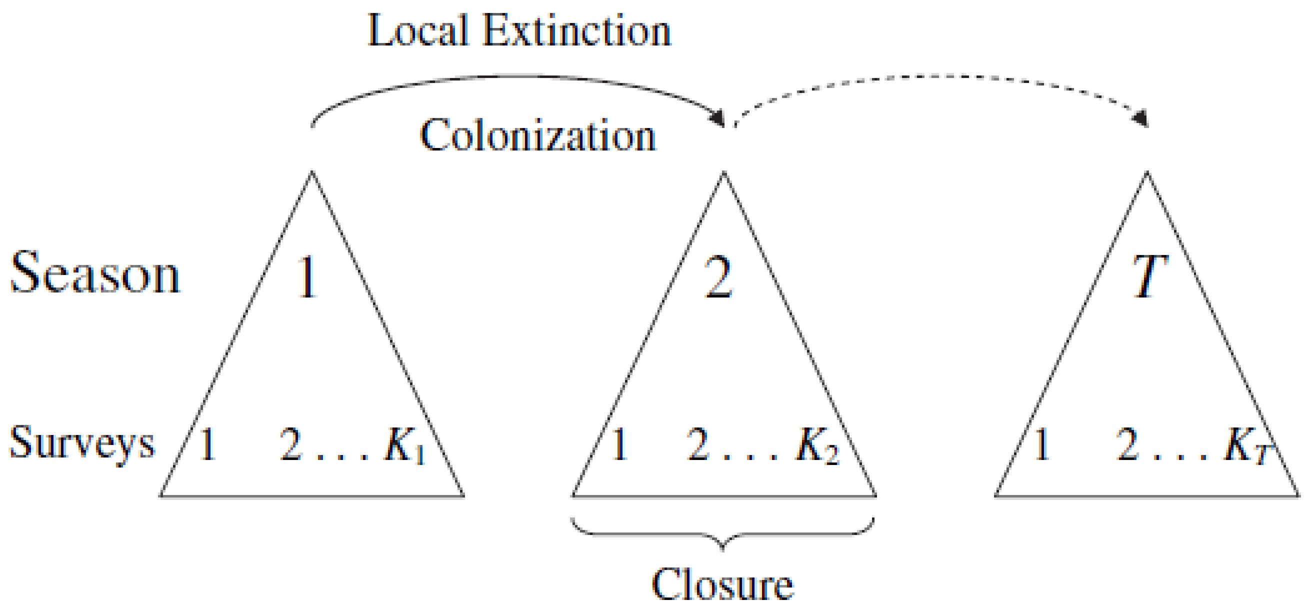


FIGURE 7.1 Graphical representation of the sampling situation for a multi-season occupancy study. Each triangle represents a season (t), with multiple (K_t) surveys within seasons. Sites are closed to changes in occupancy within seasons, but changes may occur between seasons through the processes of colonization and local extinction.

MacKenzie et al. 2006 (book)

Method is akin to "Pollock's robust design" for mark-recapture studies

E.g., Multi-season occupancy model

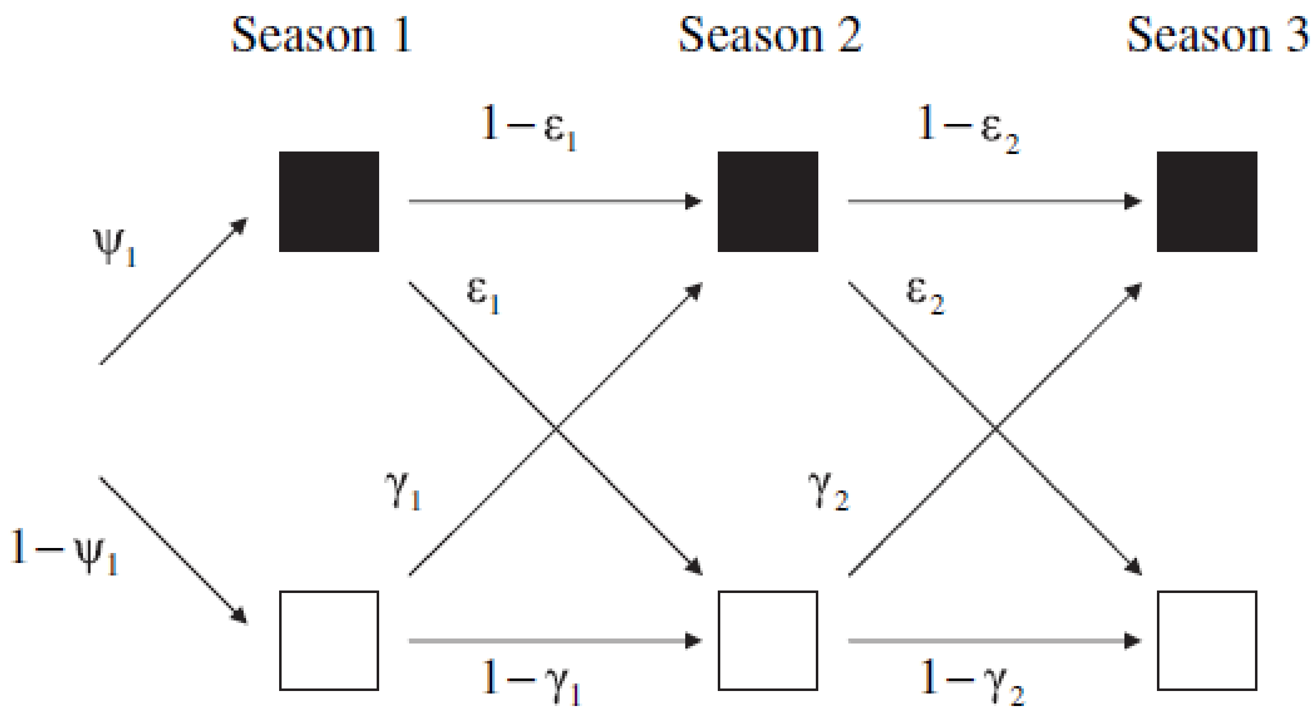


FIGURE 7.2 Representation of how the occupancy state of a site may change between seasons in terms of the processes of occupancy (ψ), colonization (γ), and local extinction (ε). Filled boxes indicate that the site is occupied (species present) in that season, while empty boxes indicate that the site is unoccupied (species absent).

MacKenzie et al. 2006 (book)

Summary – Single season occupancy models

- General, flexible approach that models **occupancy** (ψ) while accounting for imperfect **detection probability** (p)
- Data needs:
  - detection history at sites (need to have repeat sampling at each site!);
  - covariates (if any) – covariates can be for the site OR for the sampling occasion
- Strengths
  - More accurate than assuming p=1
  - Evaluate covariate effects; assess competing hypotheses
  - Flexible, statistical approach
- Drawbacks:
  - Requires more sampling at each site
  - Modeling pres/abs and not abundance
- Extensions
  - Have p depend on abundance; Multiple seasons; multiple species;…
- Know the general branching diagram and be able to derive an equation for 1) the probability for a given detection history, and 2) the likelihood for a simple example

Brief introduction to Mark-recapture analysis

Reading: Pine et al. 2003

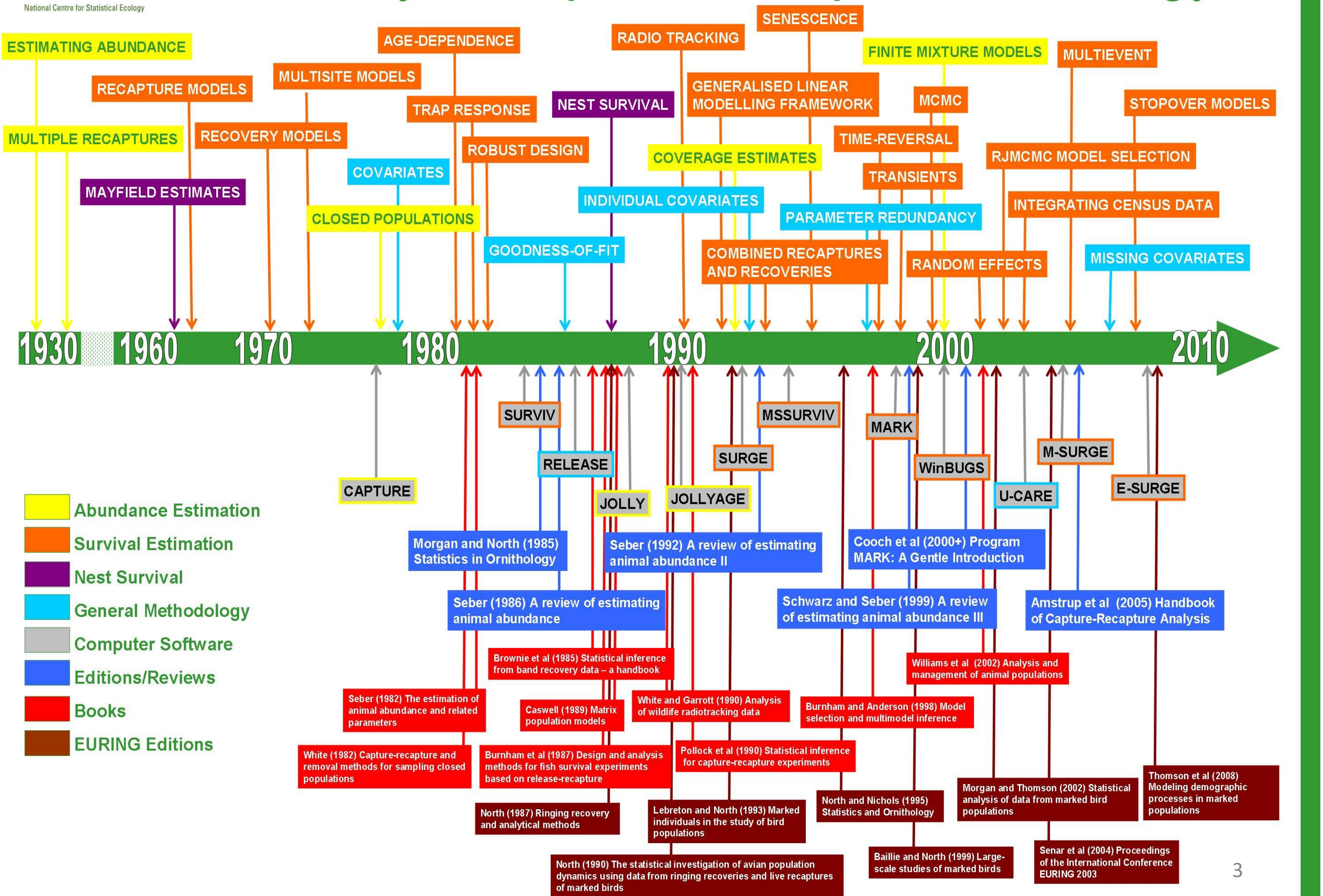
For more: WLDF 578 Advanced Ecology of Wildlife Populations (aka Dan Barton’s 'MARK class')

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

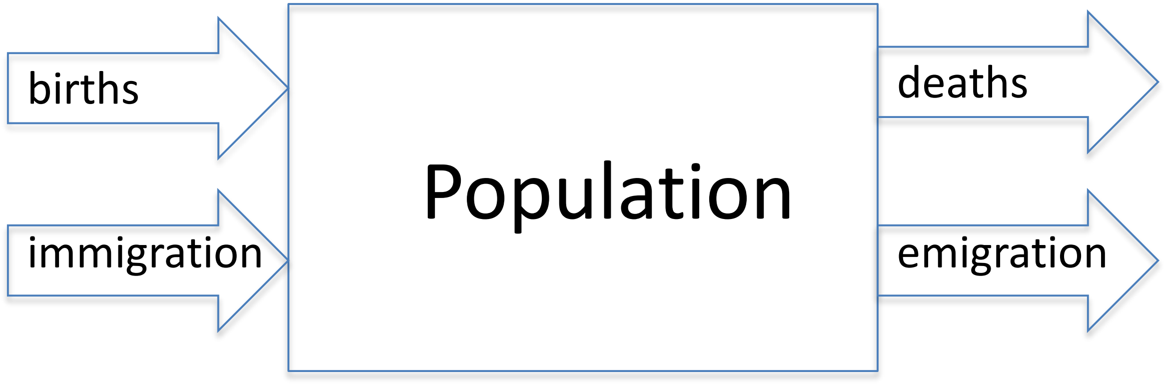
Open vs closed population

- 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



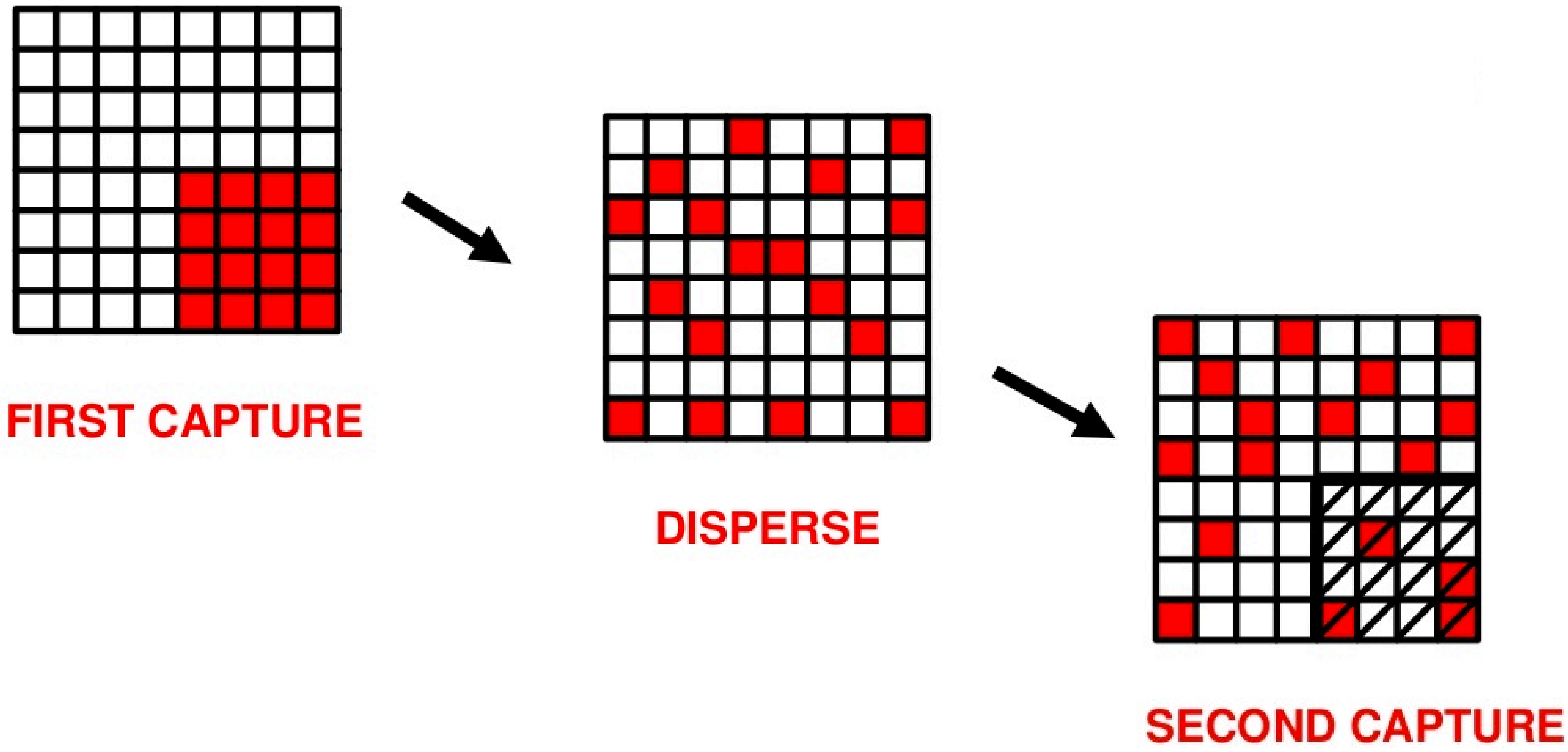
Outline

- 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

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

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

Solve for N →

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

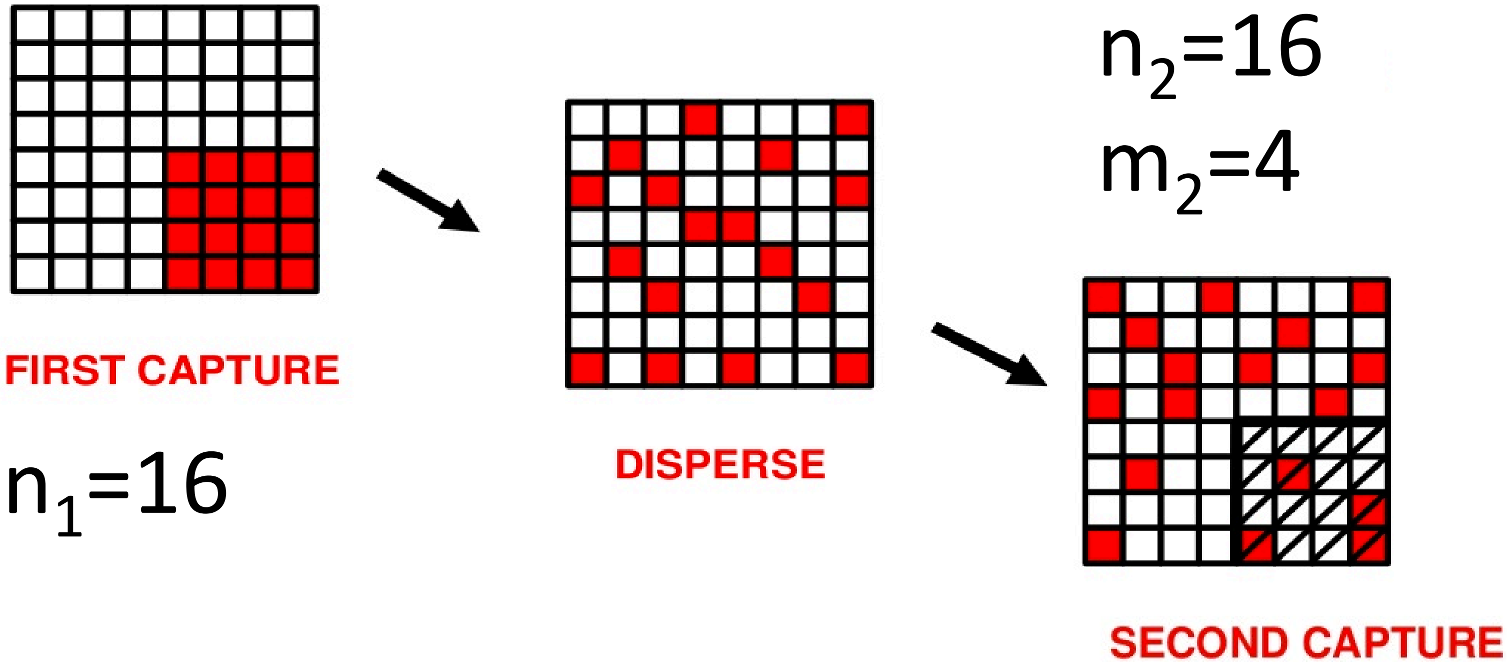
Lincoln-Petersen Correction

- 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

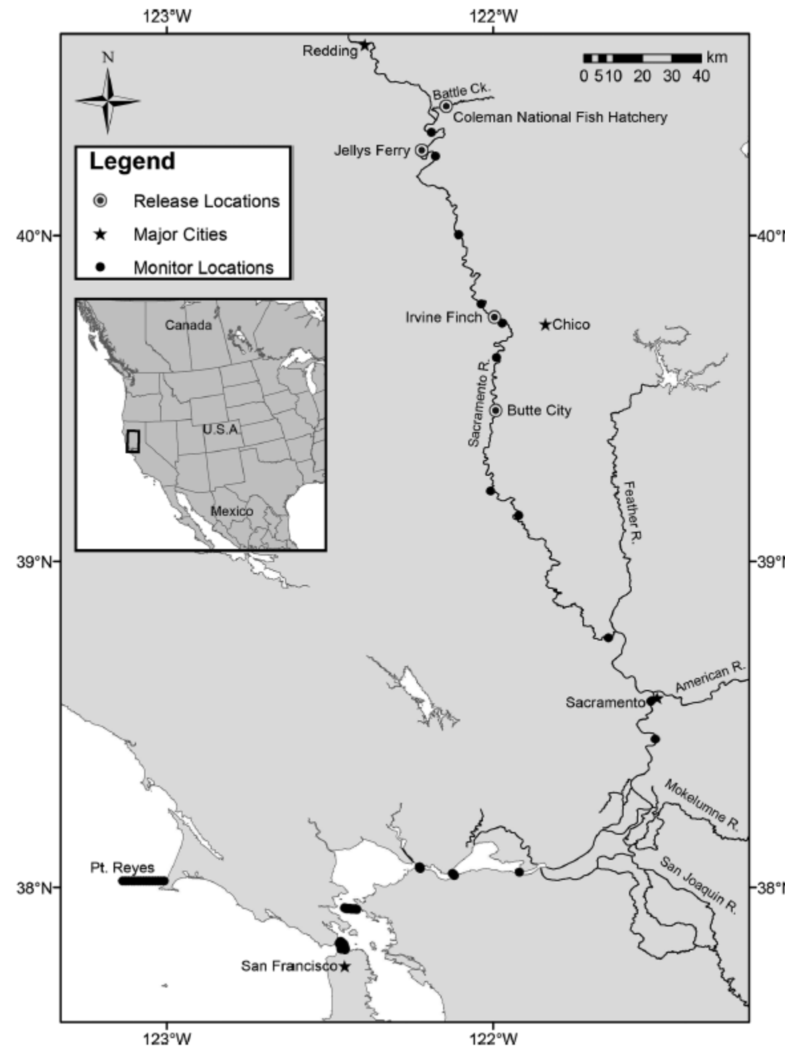
- Individuals are captured and then released
  - Can be recaptured in multiple sampling occasions

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

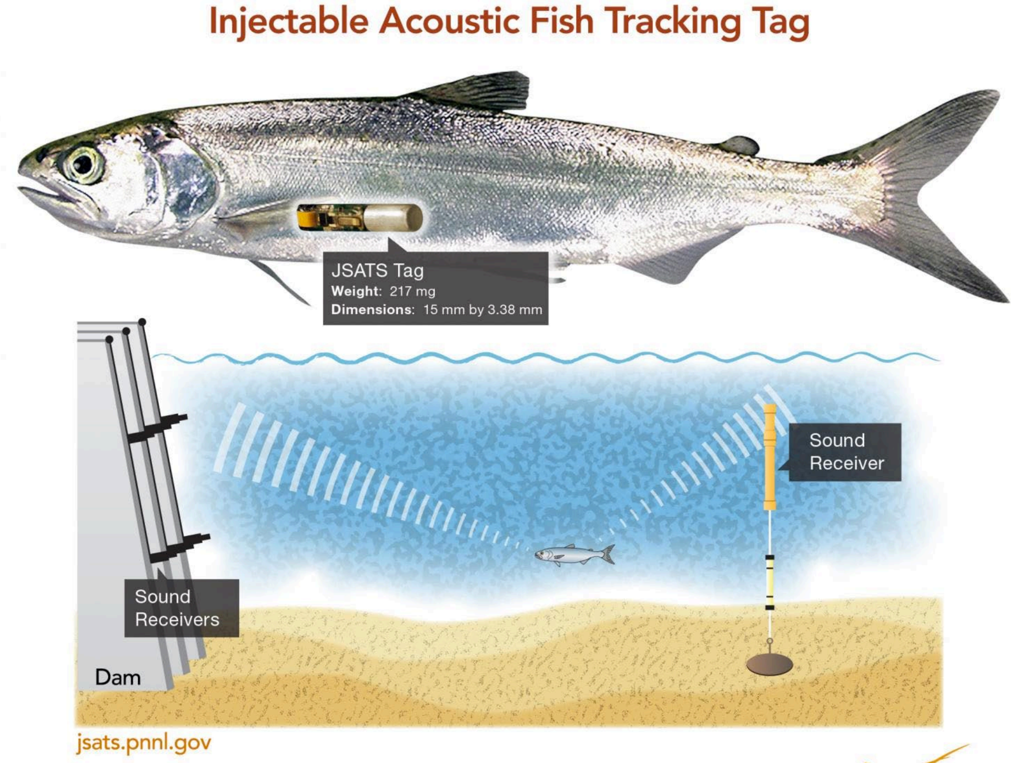
- Implanted acoustic tags in 1350 late-fall Chinook smolts
- Tracked movements with acoustic receivers
- Applied CJS model to estimate survival throughout Sac River, Delta, and SF estuary



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

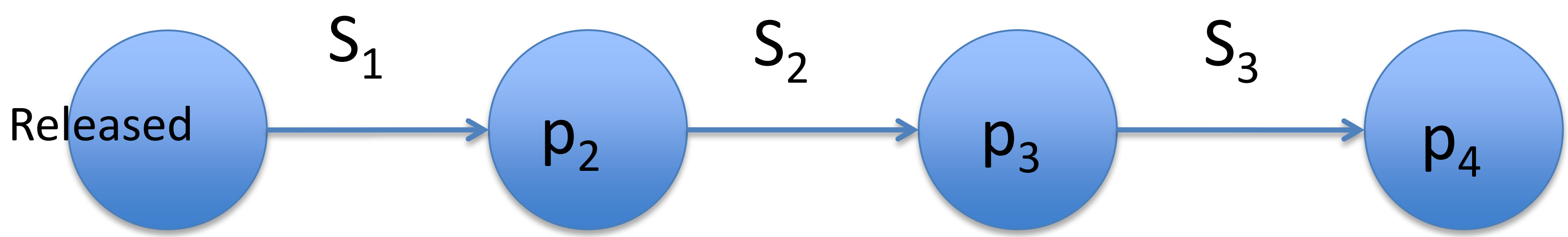
Parameters

- **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

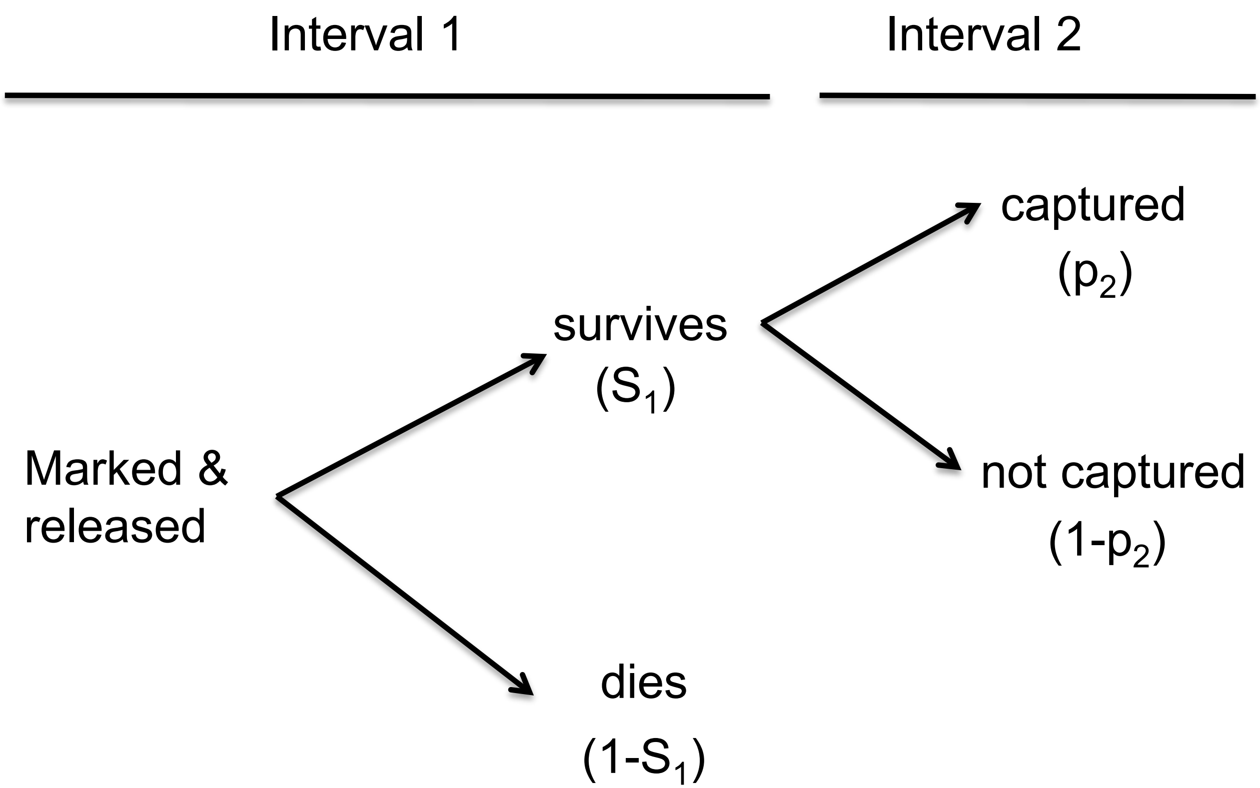
CJS data

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



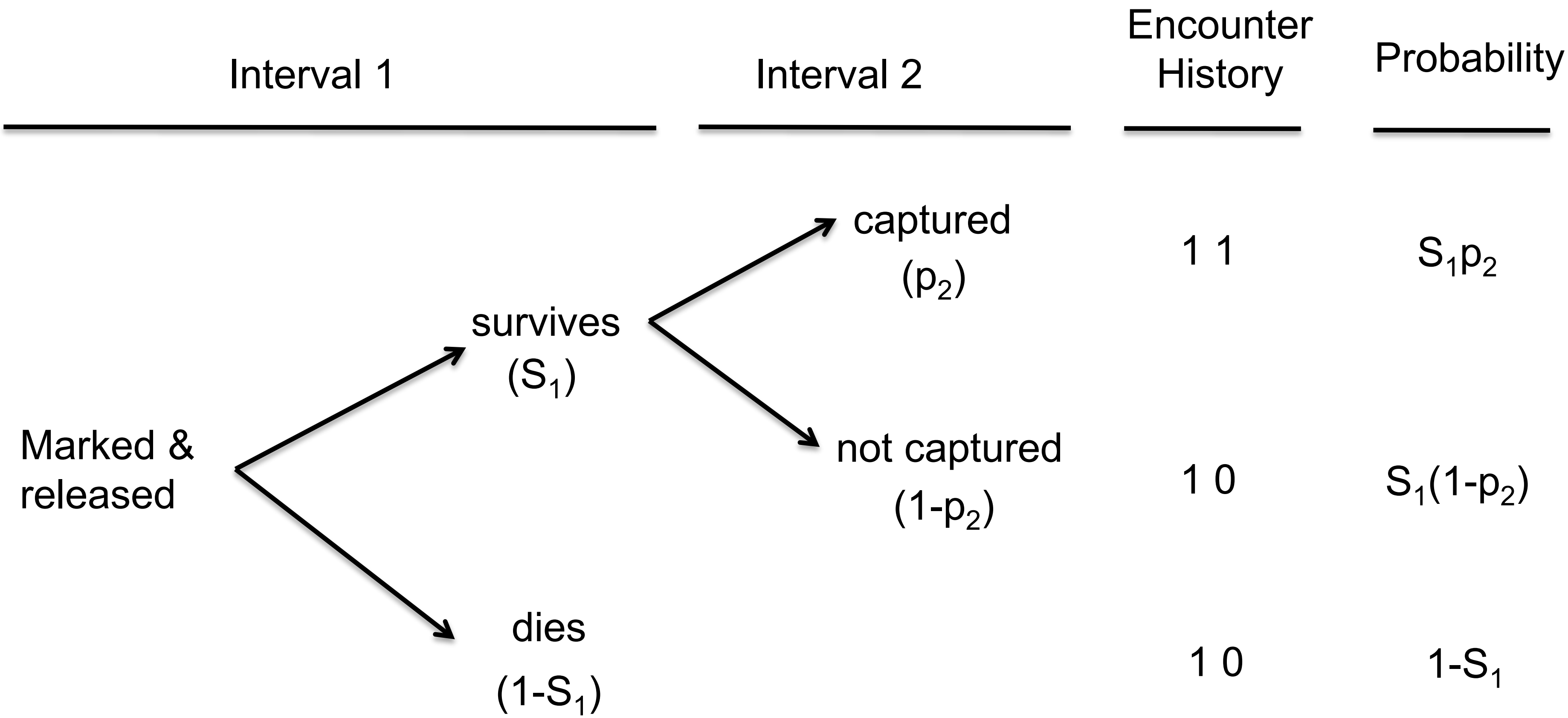
$$\begin{array}{lccccccc} \text{E.g.:} & 1 & \xrightarrow{S_1} & 0 & \xrightarrow{S_2} & 1 & \xrightarrow{S_3} & 1 \\ & & & 1-p_2 & & p_3 & & p_4 \end{array}$$

CJS Model



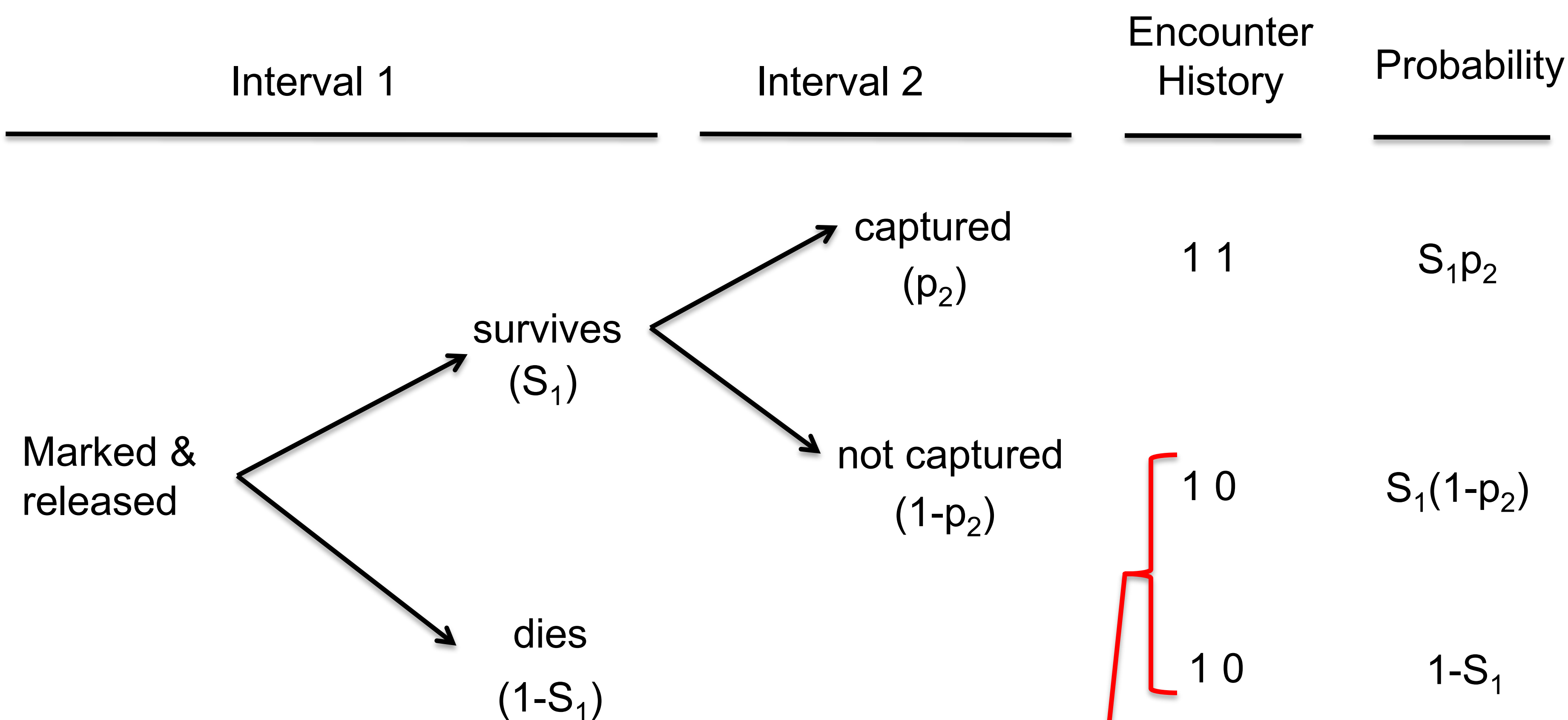
Modified from Williams et al 2002

CJS Model



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	$S_1p_2S_2p_3S_3p_4$	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	$S_1p_2S_2(1\text{-}p_3)S_3p_4$	0.027
3	0	1	1	$S_1(1\text{-}p_2)S_2p_2S_3p_4$	0.012
4	1	1	0		
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	$S_1p_2S_2p_3[1-S_3+S_3(1-p_4)]$	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

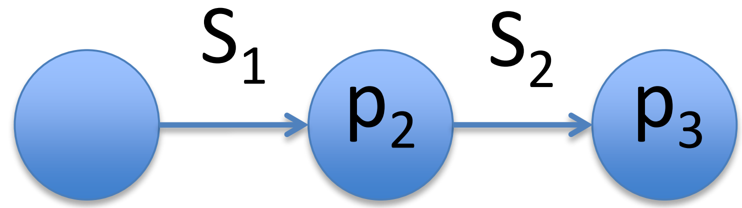
Initial parameter values: S_1=0.5, S_2=0.6, S_3=0.7; p_2=0.9, p_3=0.8, p_4=0.7

Model assumptions

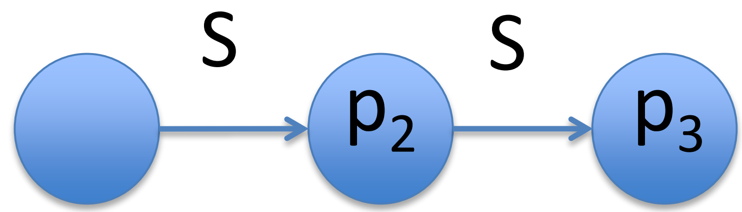
- Survival is equal for marked and unmarked animals
- Tagging does not influence capture probability
- Sampling is ‘instantaneous’
- Tags are not lost or overlooked
- Fate of each fish is independent

Examples of Potential Models

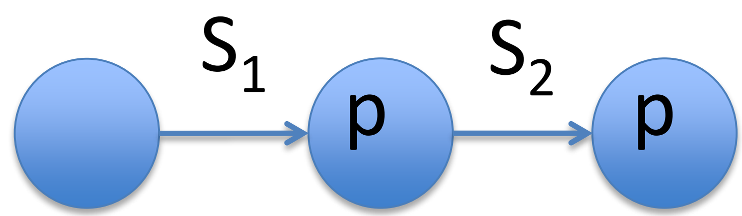
- Time dependent



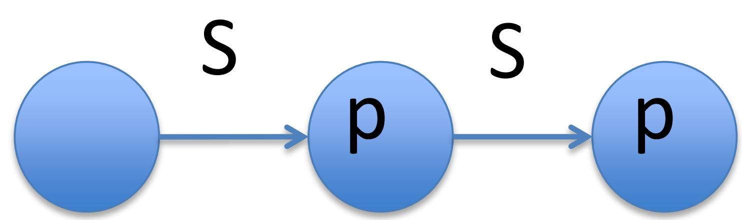
- Constant Survival



- Constant Cap Prob



- Time independent



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

Model Selection

- Akaike's Information Criterion (AIC)

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

  - $\hat{L}$ = 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

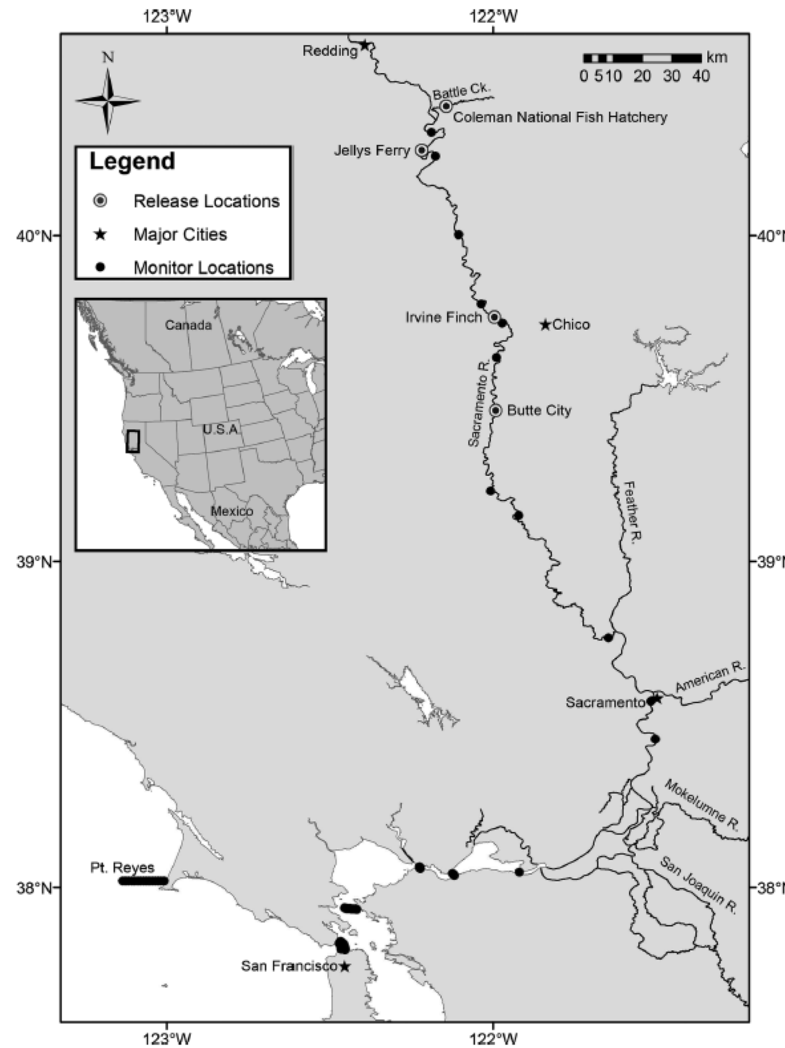
$$\log\left(\frac{\Phi_i}{1-\Phi_i}\right) = \beta_0 + \beta_1 X_{1i} \ldots$$

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

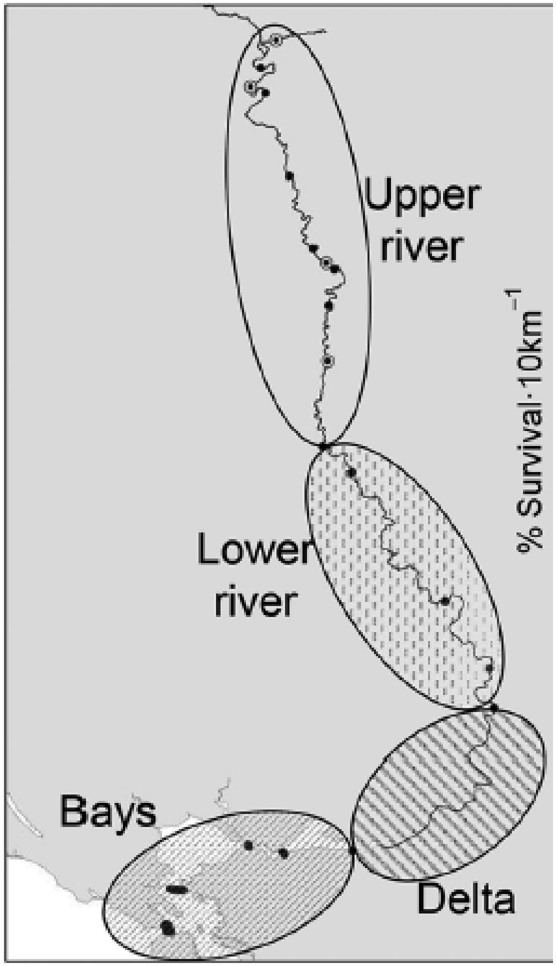
CJS Example

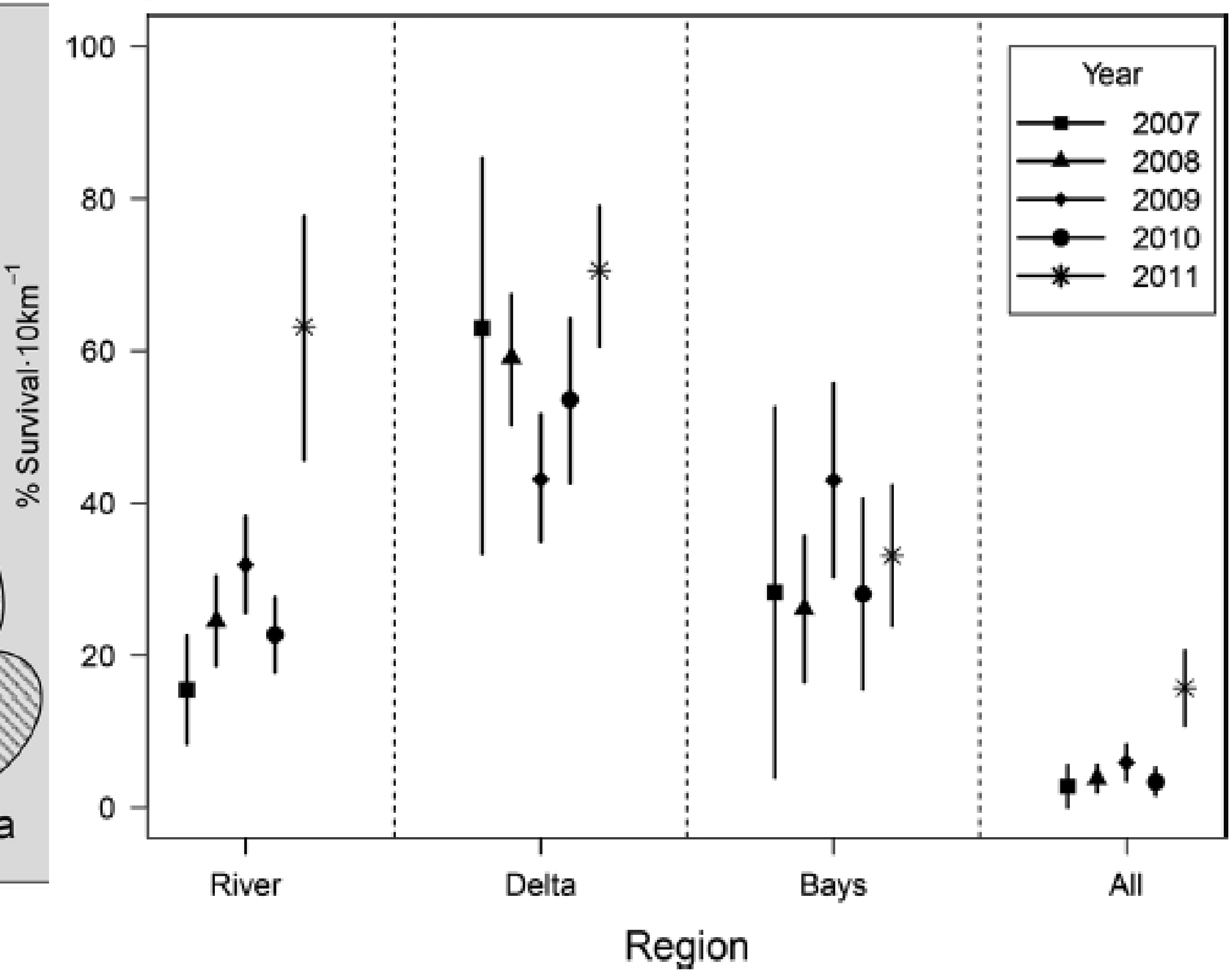
Table 3. Survival models for different spatial and temporal factors, as well as individual covariates, ordered from lowest to highest $QAIC_c$, omitting 2011 data.

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

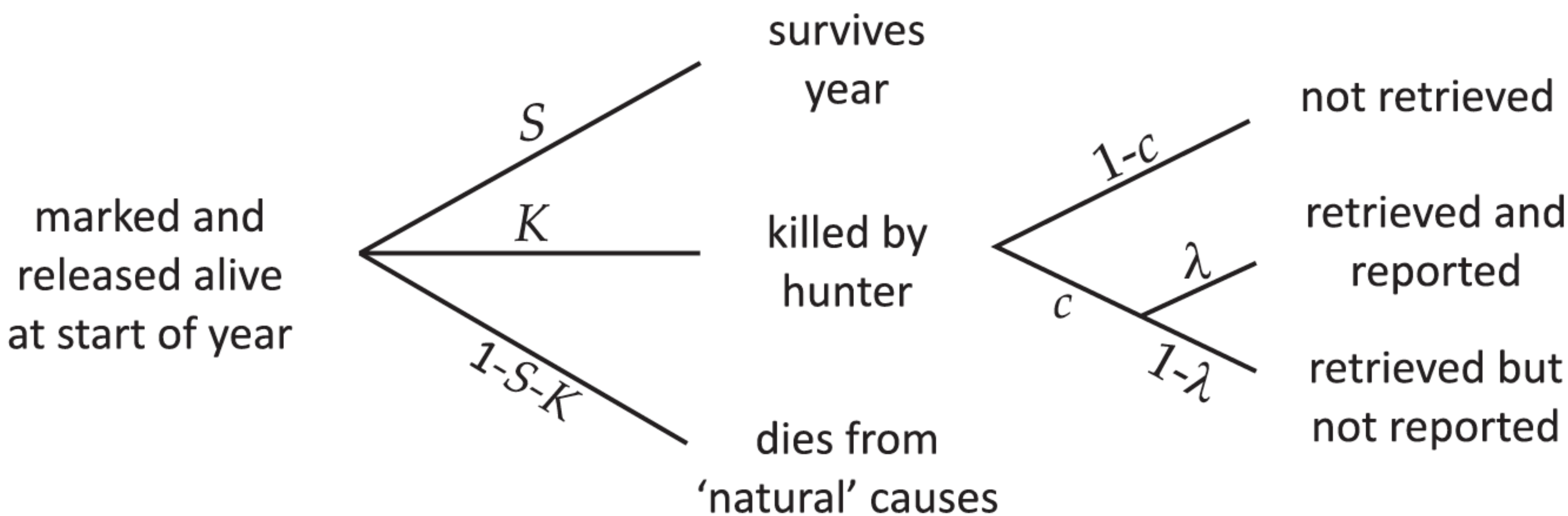
- Models where a tag is ‘recovered’ and not returned to the population
  - Harvest

www.skysguideservice.com

Recovery Model

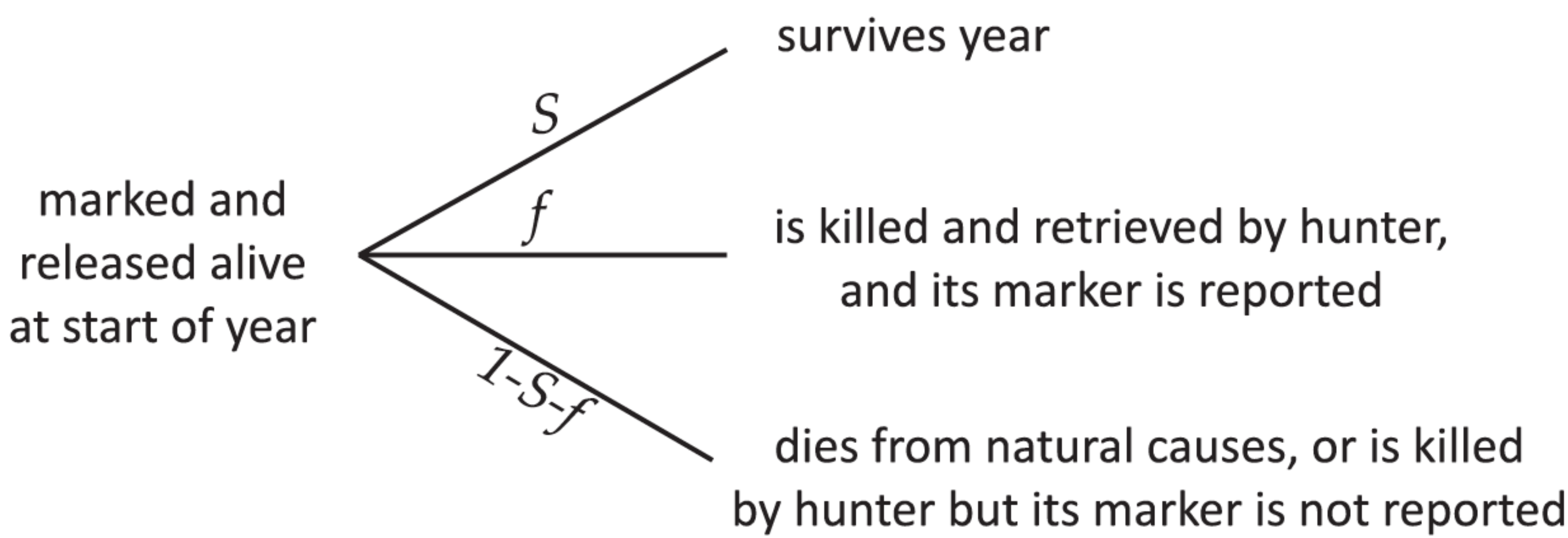
- Models where a tag is ‘recovered’ and not returned to the population
  - Harvest
- Also requires at least 3 sampling occasions
- Often referred to as ‘Brownie models’
  - Brownie et al. 1985

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	$N_1 f_1$	$N_1 S_1 f_2$	$N_1 S_1 S_2 f_3$	$N_1 S_1 S_2 S_3 f_4$
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				$N_4 f_4$

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 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				$N_4 f_4$

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

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

TABLE 16.8 Recoveries of Adult Male Mallards

		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				

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

$$S_1 = 0.75$$

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

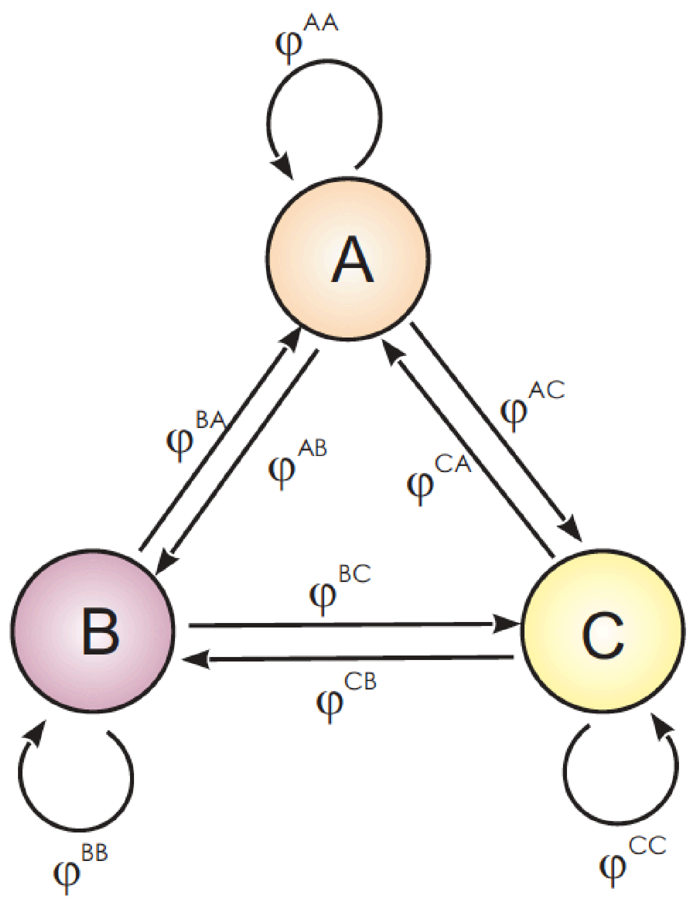
- The parameter f_i (prob of harvest and reporting) is a function of the probability of harvest (u_i) and the tag-reporting rate (λ_i)
- Methods to estimate reporting rate
  – High reward tags (Pollock et al. 2001)
  – Planted tags
  – Port sampling
  – Two samples per year (Hearn et al 1998)

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’

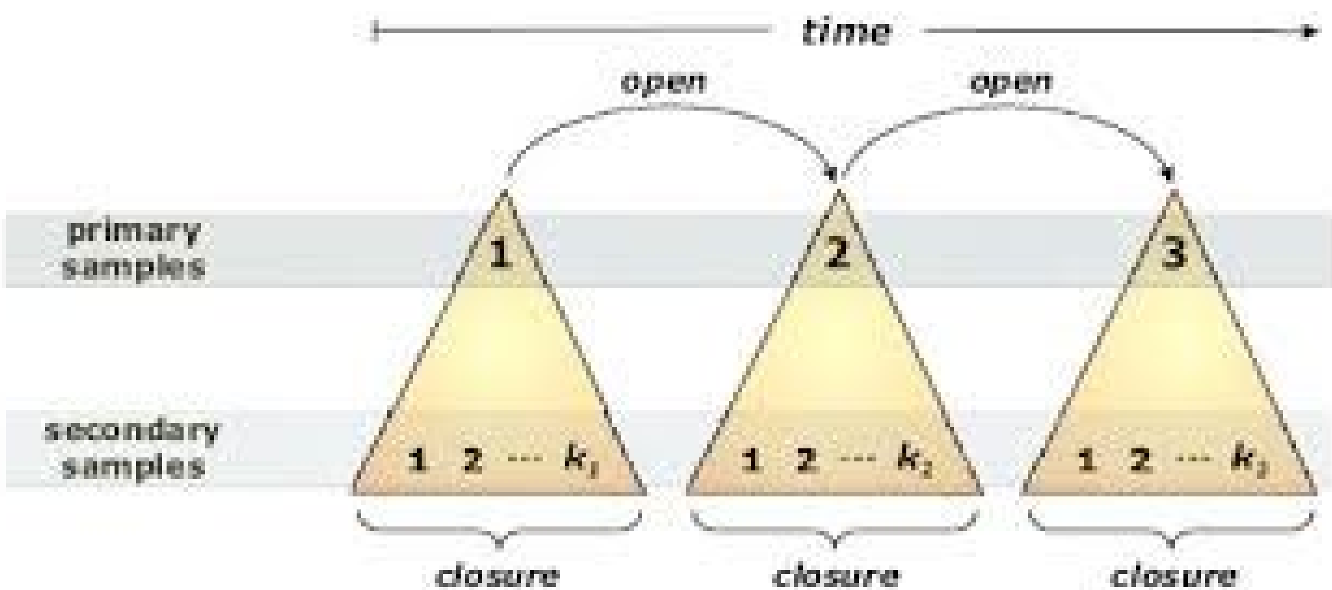


Multistate Models

- 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**