# Incorporating biological and environmental realism into fisheries stock assessment models

**Terrance J. Quinn II**

**School of Fisheries and Ocean Sciences**

**University of Alaska Fairbanks**

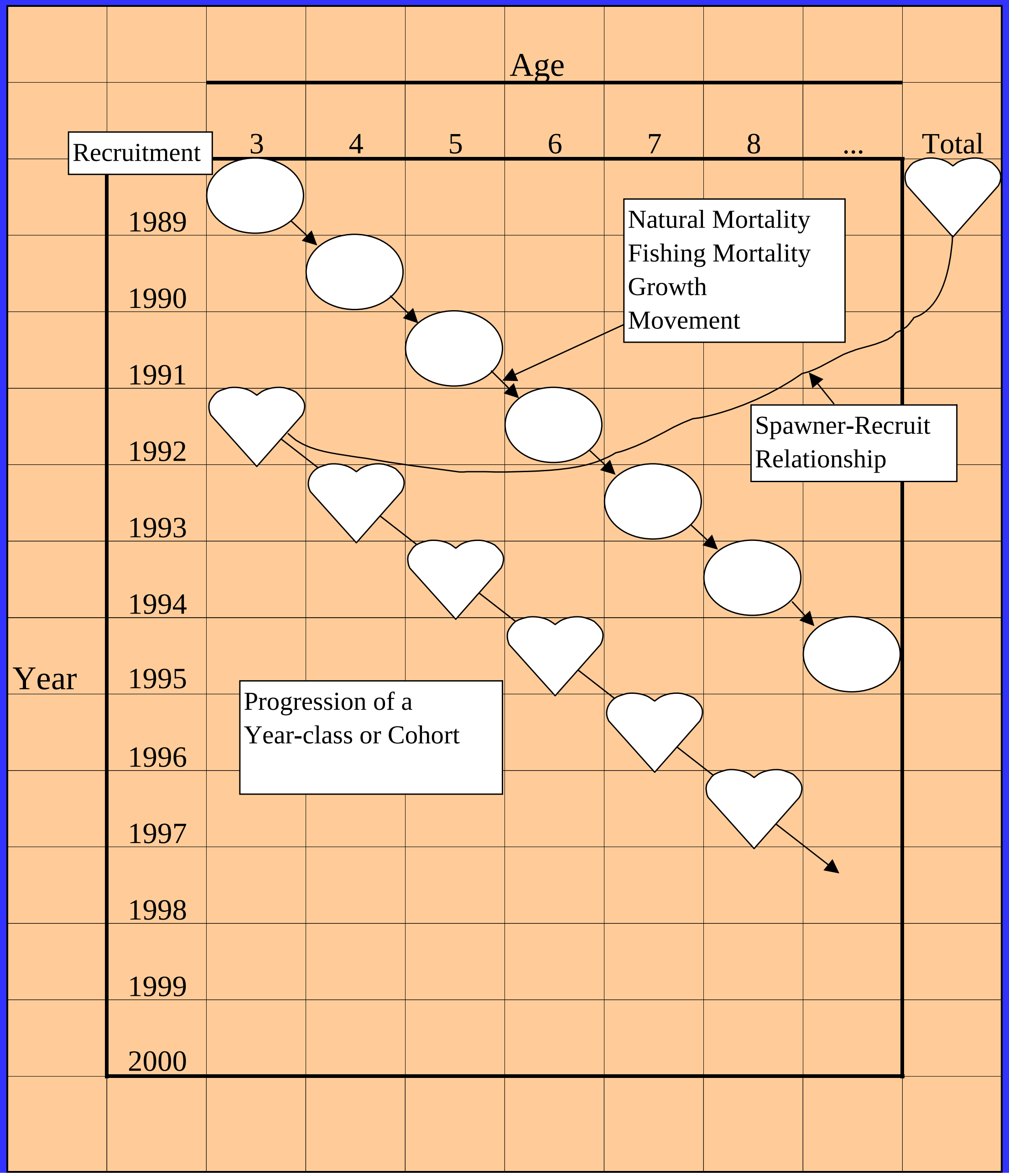
**Juneau Alaska**

**Terry.Quinn@uaf.edu**

# For what purpose?

- Assessment (Accounting, Estimation)
- Forecasting (Prediction, What If Scenarios)
- Cause and Effect (Understanding the Processes, Experiments)

# The Holy Grail: Age-structured Analysis



# Stock Assessment

1. Data Collection
    1. Fishery
    2. Surveys
2. Modeling and analysis
    1. Population dynamics
    2. Uncertainty in measurement and in process
    3. Factors affecting the population (environment)
3. Management recommendations
    1. Biological reference points
    2. Sustainability
    3. Plan of action

# Data from the Fishery

- Harvest data
  - Total catch and kill
    - Should include release and bycatch mortality
  - Composition: length, age, sex
    - Follow year-classes through time
  - Catch-per-unit-effort
    - Index of population change
    - Needs validation as proportional to abundance

# Biological sampling

- Abundance estimation
  - Mark-recapture methods
    - Common approach with recreational fisheries
    - Hundreds of applications
    - Variety of experimental designs, software
  - Line transect methods
  - Removal methods
    - Useful only if significant kill
  - Survey sampling
    - Prevalent with commercial fisheries
    - Simple, stratified, systematic, cluster, adaptive

# Necessary biological information

- Natural mortality $M$ and fishing mortality $F$
- Total mortality $Z = F + M$
- Growth
- Recruitment
- Movement and migration
- Maturity and fecundity (egg production)

# Necessary Modeling

- Connects data and population dynamics
- New abundance = Previous abundance – Fishing Deaths – Natural Deaths + Recruitment + Immigration – Emigration
- Constant and known natural mortality
- Recruitment
  - Related to previous spawning stock
  - Related to previous environmental conditions
  - Related to other species

# Goals of Modeling

- To explain time series of data
- To estimate population parameters
- To determine causes of population change
- To forecast future populations
- To reconcile conflicting information sources
- To specify uncertainty and risk

# What is the objective function?

- The objective function is used in stock assessment models to estimate parameters

- A general equation for the objective function is:

$$O(D) = \sum_{x} \lambda_x G(D_x, P_x)$$

- Here, *G* is some function that relates the data, *D*, to the model predictions, *P*, for some dataset *x*, *λ* is the weighting term.

# What is *G*?

- In the objective function, *G* is formulated as the likelihood function of our set of parameters given the dataset *x*.

- The function *G* is what connects statistics to our models, or, allows us to quantify uncertainty in our estimates

- For computing purposes, *G* is the negative log-likelihood, and parameters are estimated to minimize *G*

# Examples of *G*: Index data

- $G(D_x,P_x)$ is most often log-normal:

$$G\left(D_x,P_x\right)\cong\frac{1}{\sigma_{D_x}^2}\left(\ln D_x-\ln P_x\right)^2=\lambda_x\left(\ln D_x-\ln P_x\right)^2$$

- Here, the weighting term $\lambda$ is the inverse of the variance of the data, *D.*

- In this case, as the uncertainty in *D* increases the weight, $\lambda$, would decrease.

# Examples of *G*: Compositional data

- Here, a multinomial likelihood can be used, where $G(D_x,P_x)$ is formulated as:

$$G\left(D_x, P_x\right) \simeq n_x \sum_a P_{a,x} \ln D_{a,x} = \lambda_x \sum_a P_{a,x} \ln D_{a,x}$$

- where the *a* subscript denotes ages, and the weighting term $\lambda$ is the sample size *n.*

- In this case, as our sample size *n* increases the weighting term, $\lambda$ increases, or, uncertainty decreases.

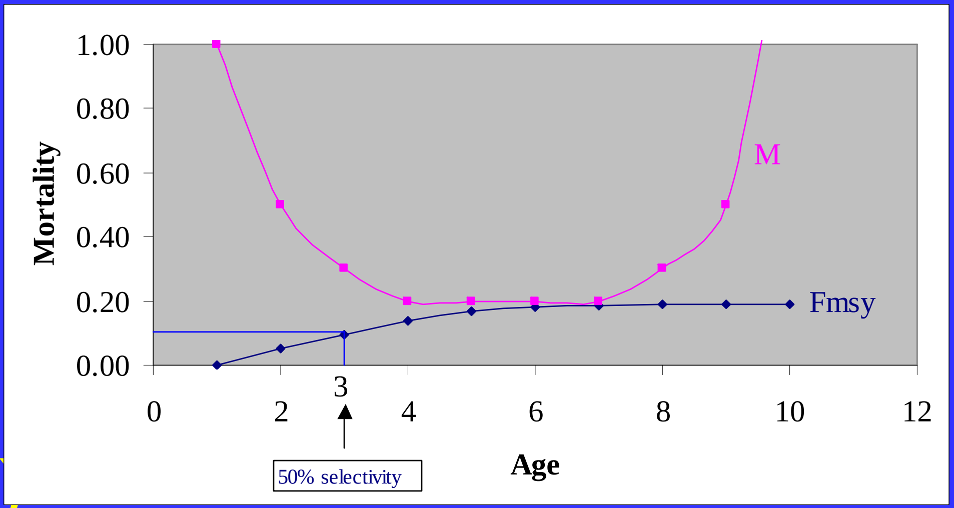
# Software

- Up to hundreds of parameters, thousands of observations
- Excel
- Local products: ADAPT, Stock Synthesis, XSA, etc.
- AD Model Builder (Dave Fournier, automatic differentiation,

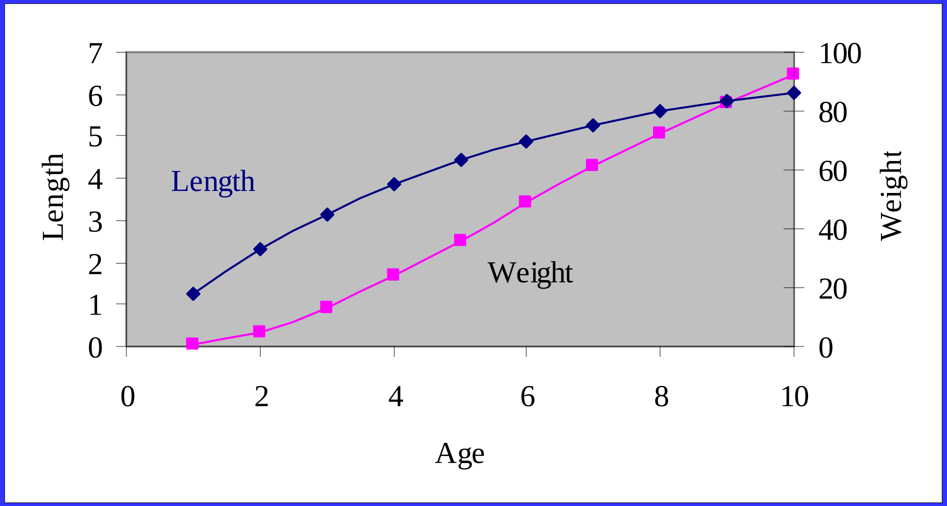
  http://admb-project.org/

# Prototype of Underlying Dynamics

- 10 ages
- M: U-shaped
- F: logistic (50% selectivity at age 3)

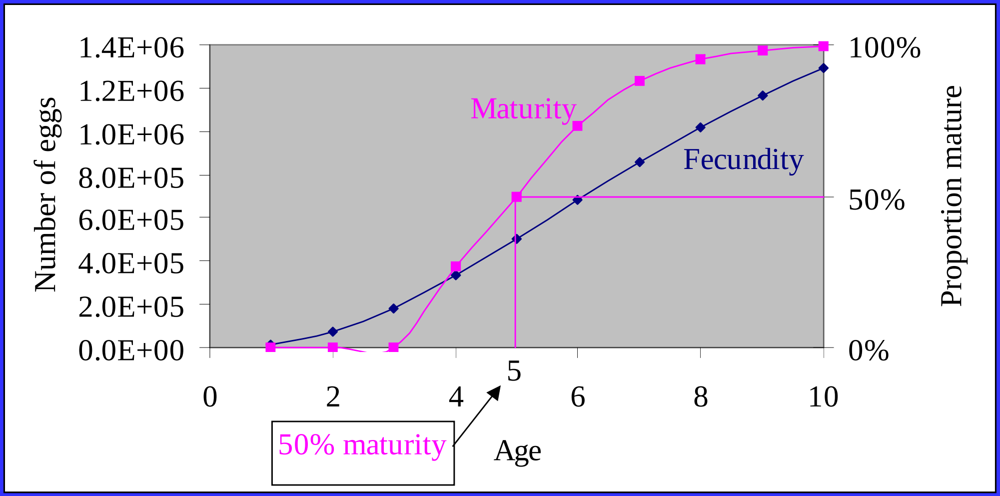



- L: LVB
- W: isometric



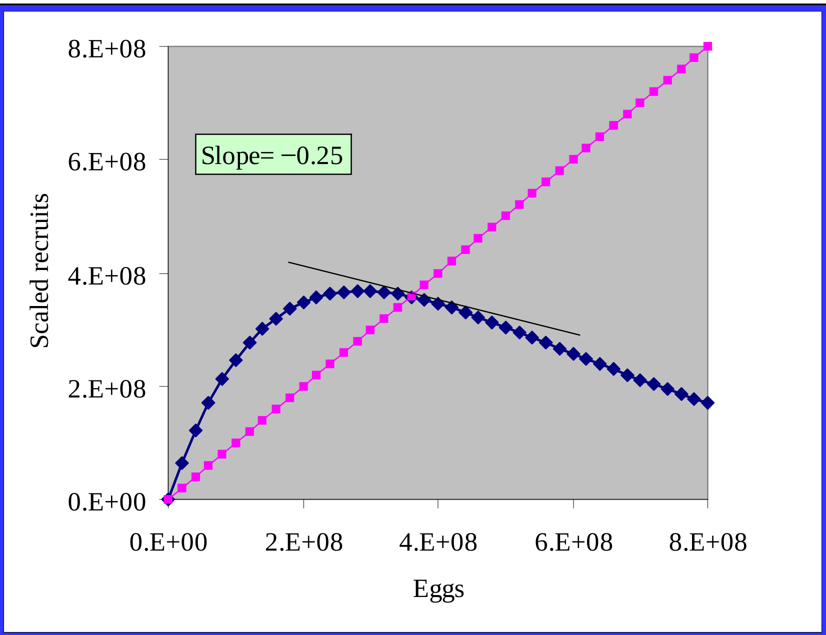
# Prototype (continued)

- Maturity: logistic (50% mature at age 5)
- Fecundity: isometric

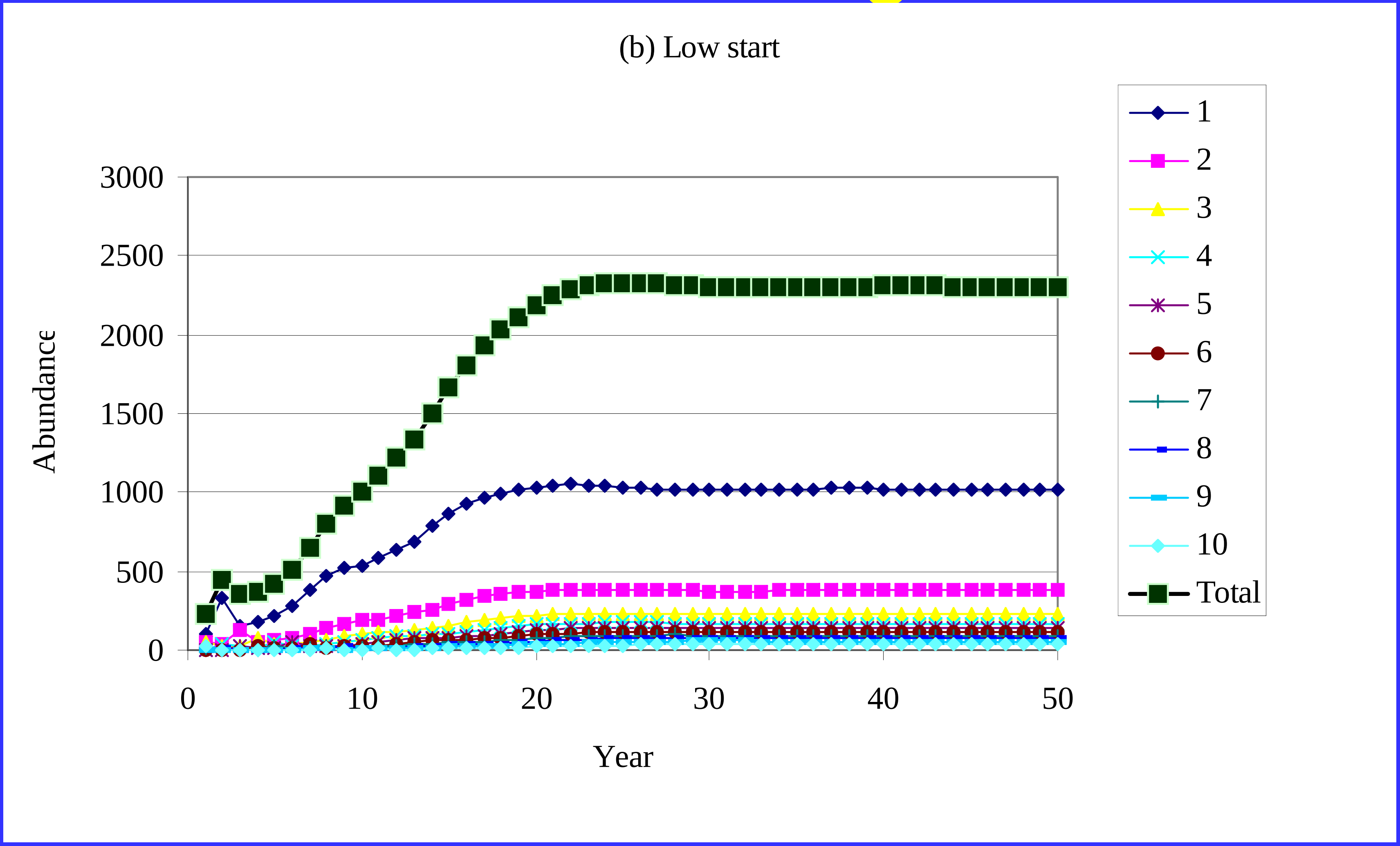



- Spawner-recruit relationship: Ricker

$$R = \alpha S \exp(-\beta S)$$

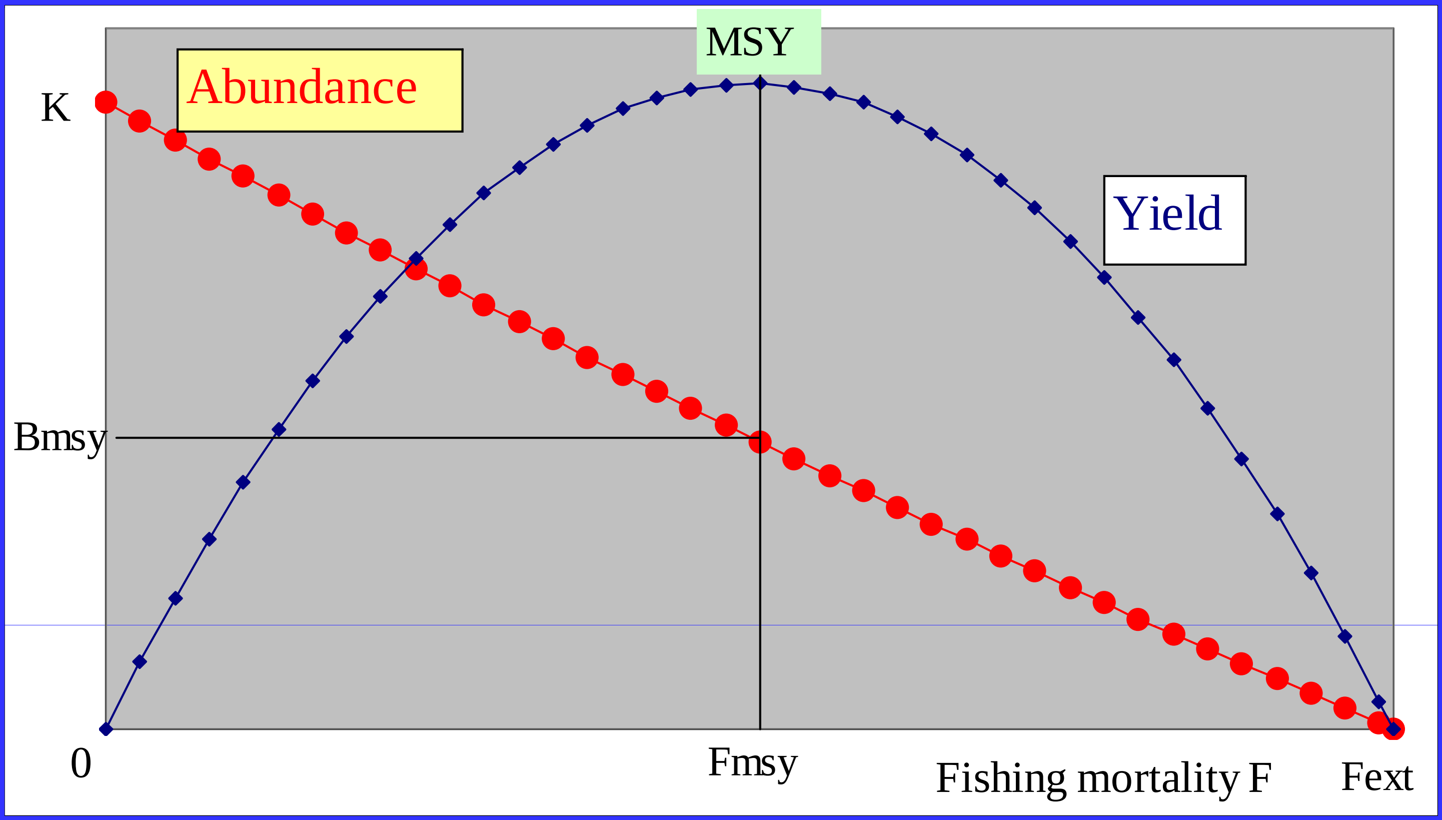


# No fishing



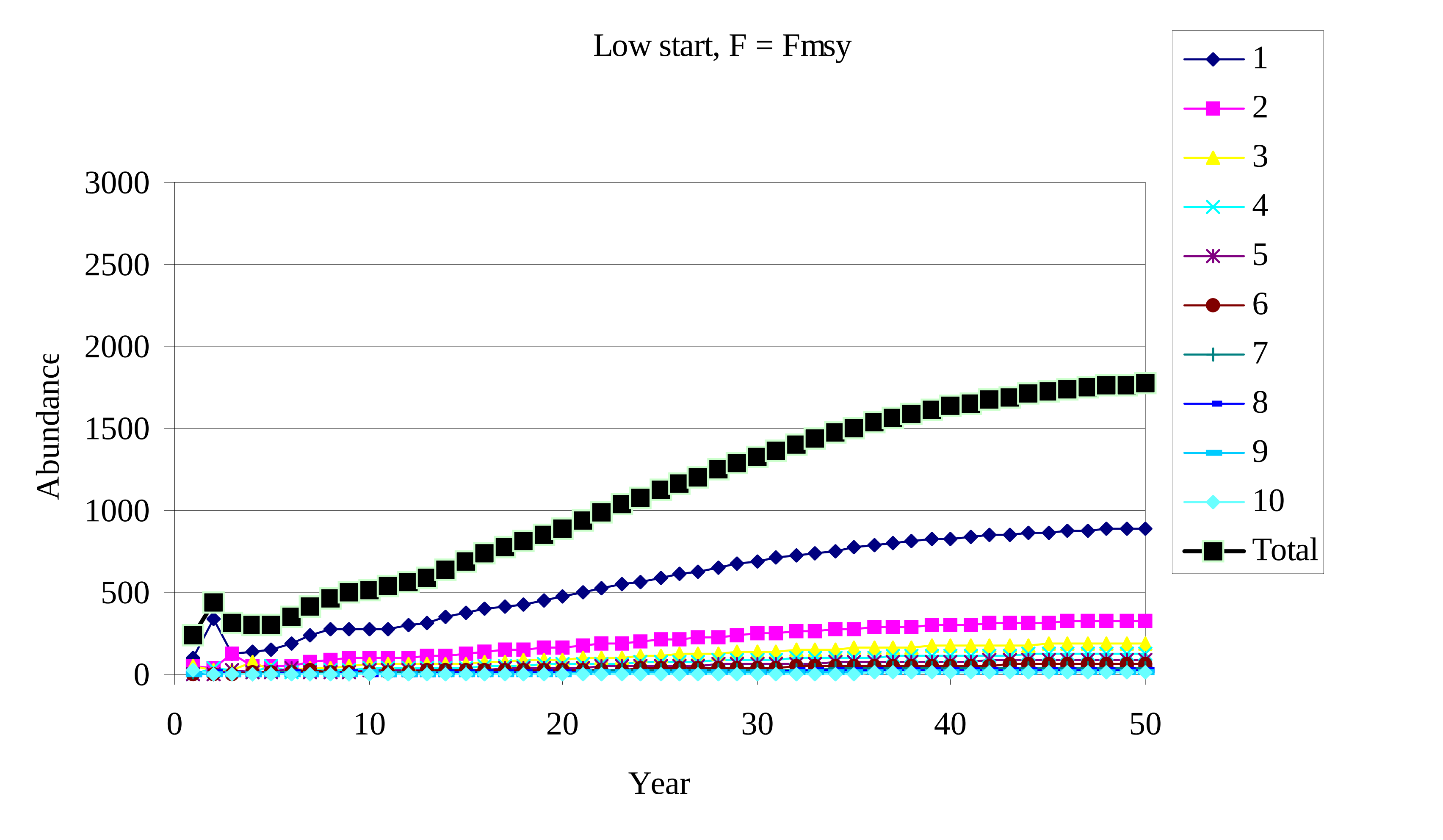

No matter whether the population starts low or high, it equilibrates to its carrying capacity (2300).

# When fishing occurs



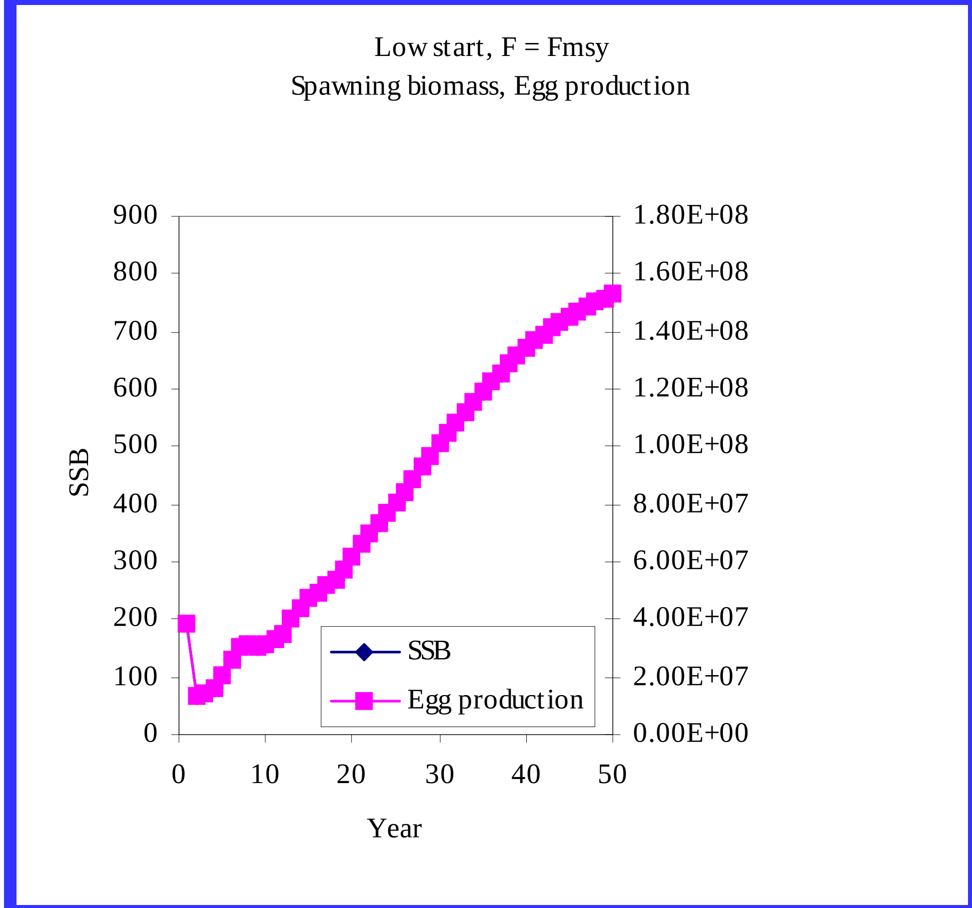

- Continuum of sustainable yields and populations
- Extremes: *B*=*K* at *F*=0 and *B*=0 at *F*=*F*ext
- Optimal: *B*=*B*msy at *F*=*F*msy

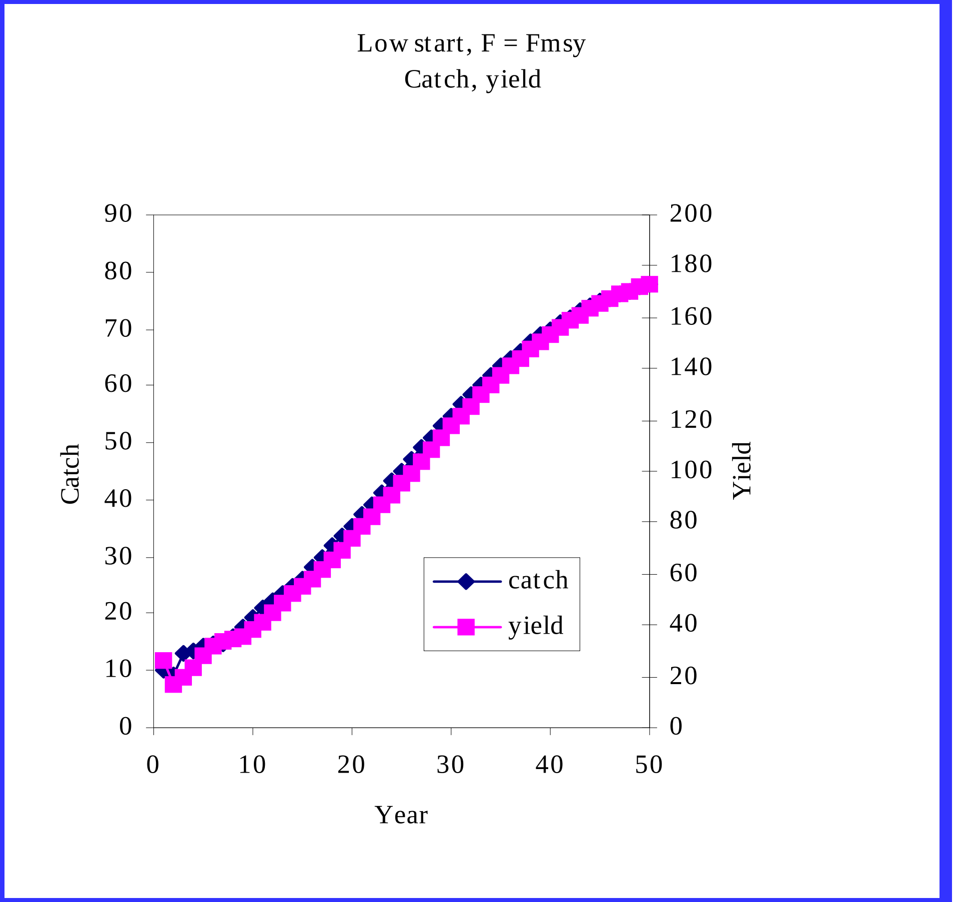
# Trajectory when F=Fmsy




Population equilibrates at the Bmsy level (1800).

# Reproduction and catch
# Low start, F=Fmsy



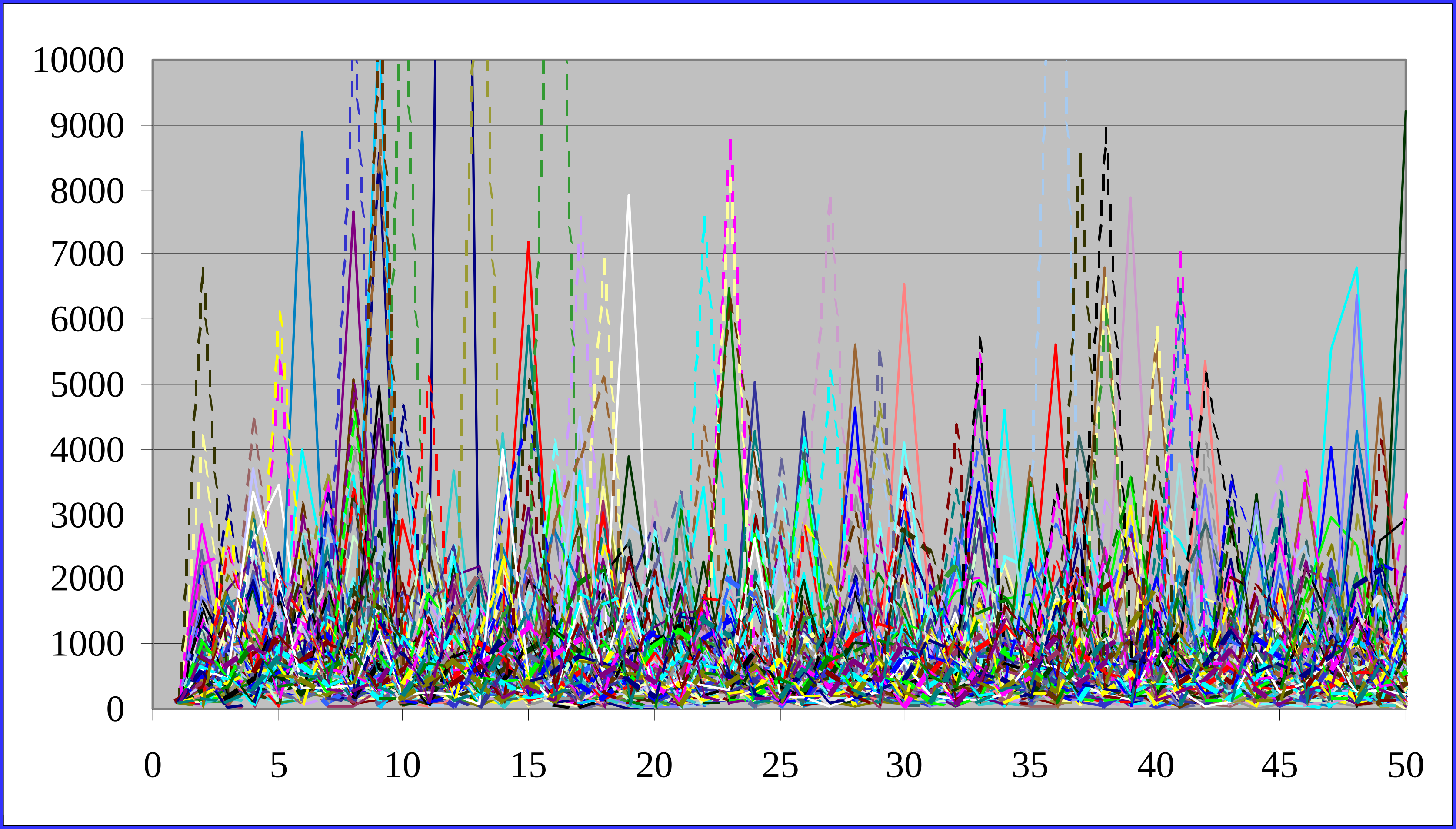
# Challenge 1: Stochasticity

- Ricker spawner-recruit relationship
- Need stochastic effects for temporal change, environment
- Lognormal variability, E(*R*)= deterministic

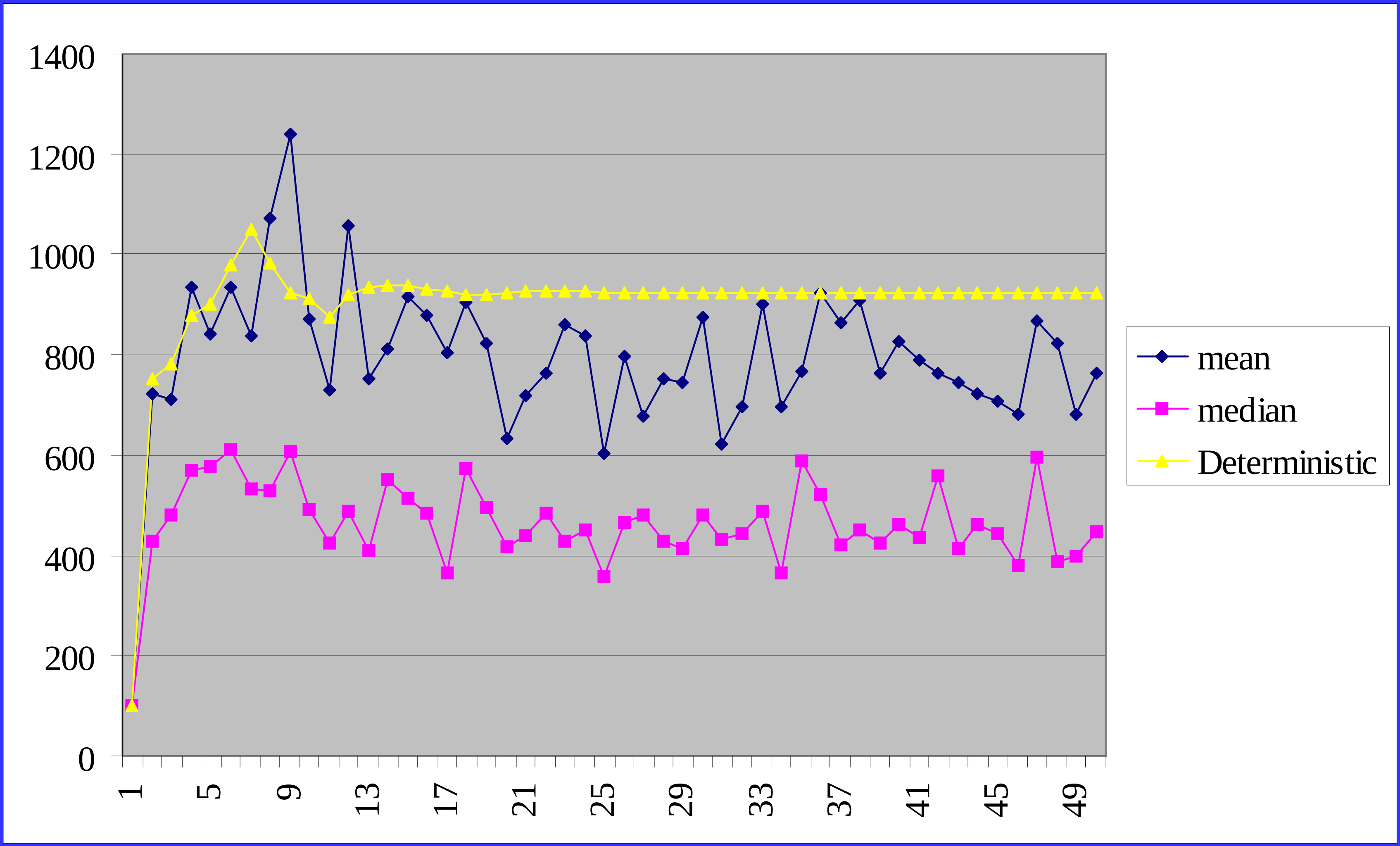
$$R = \alpha S \exp(-\beta S)\exp(\varepsilon - \tfrac{1}{2}\sigma^2), \quad \varepsilon \sim N(0,\sigma^2)$$

- CV = 1 (fairly high for illustration)
- 100 replications
- Compare mean and median parameters with deterministic ones.

# Recruitment replications



# Mean and median recruitment



# Stochastic conclusions

- Stochastic effects are large on all population parameters.
- These effects occur at all life stages.
- The effect is downward: Yield, population abundance, and egg production are lower than the deterministic case.
  - Solution: More conservative action is necessary if stochasticity is present.
- Density dependence is poorly estimated.
  - Solution: Bayesian hierarchical models, meta-analyses

# Challenge 2: Varying natural mortality

- U-shaped distribution not well determined
- A function of predators and disease
  - Solution 1. Covariates (disease prevalence, predator abundance)
  - Solution 2. Multi-species models (more realistic but more uncertain, requires consumption data)

Cause and effect requires study of early life history (expensive, complex)

- **Deconstruct Z into:**
  - **Fishing mortality *F***
  - **Predation mortality *P***
  - **Residual natural mortality *M***

$$N_{i,a+1,t+1} = N_{i,a,t} e^{(-M - F - P_1 - P_2 .... P_n)}$$

**The Multispecies Model is simply an extension of the single species model, in which *Z = F + M + P!***

# Modeling predation

$$P_{i,a,t} = \frac{1}{N_{i,a,t} W_{i,a}} \sum_j \sum_b I_{j,b} N_{j,b,t} \frac{\phi_{i,a,j,b,t}}{\phi_{j,b,t}}$$

i = prey species
j = predator species
a = prey age
b = predator age

Annual Ingestion

Predator abundance

Total ingestion by predator j * Proportion of the ingested food that is prey i, age a

Σ Total amount of prey i consumed by predator j

$$P_{i,a,t} = \frac{\text{Total amount of prey i consumed by all predators}}{\text{Biomass of prey i}}$$

# Challenge 3a: Multiple datasets

- Data weighting issues (back to objective function

$$\lambda_i = \sigma_1^2 / \sigma_i^2 \text{ [ratio of variances, dataset 1 to dataset } i]$$

$$\max \ln L = \sum_i -\frac{n_i}{2}\left[\ln(2\pi\hat{\sigma}_1^2 / \lambda_i) + 1\right], \text{ in which}$$

$$\hat{\sigma}_1^2 = \sum \lambda_i RSS_i / \sum n_i \text{ [weighted residual sum of squares]}$$

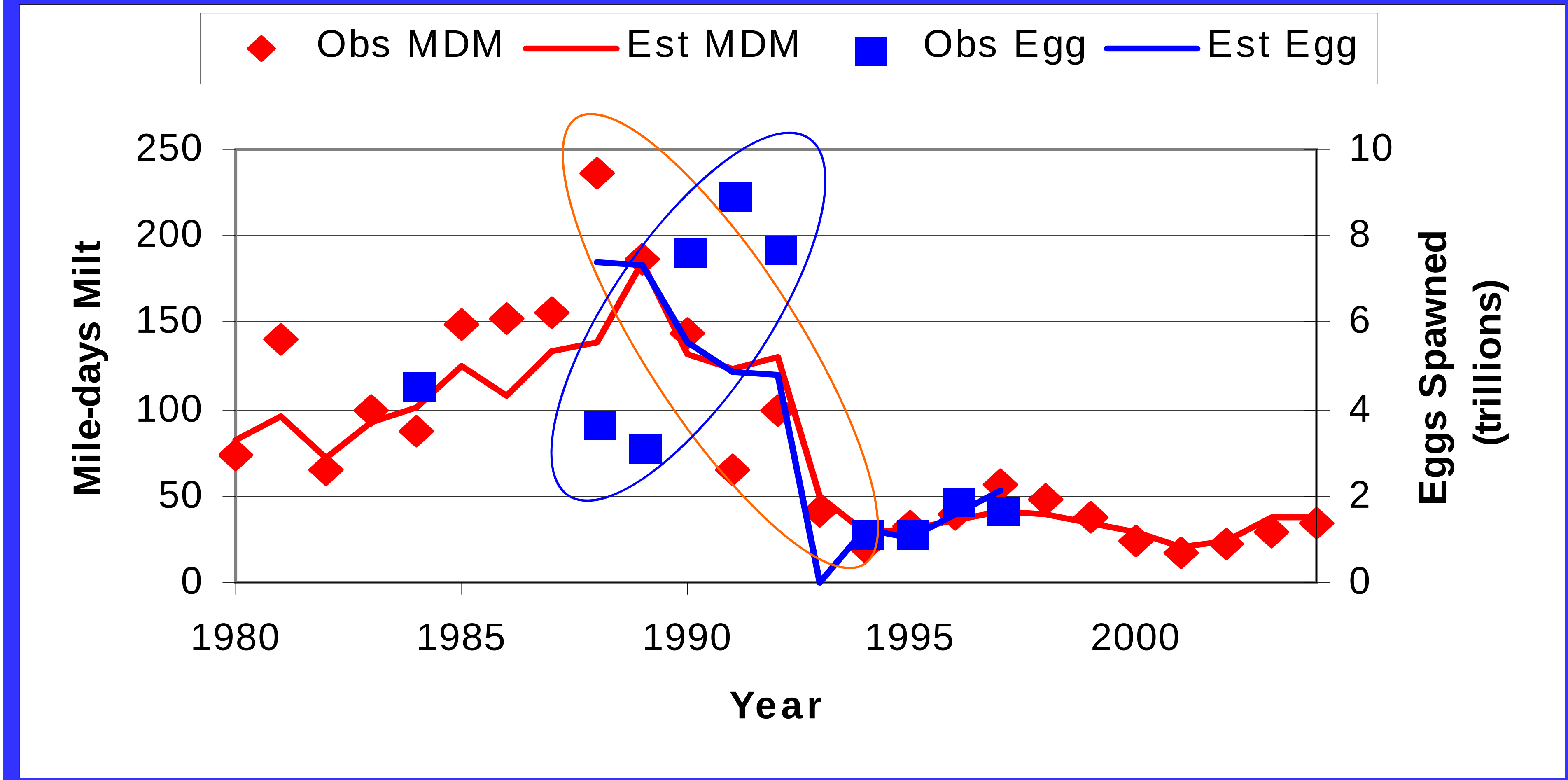
$$\hat{\sigma}_i^2 = \hat{\sigma}_1^2 / \lambda_i.$$

- What to do about weightings $\{\lambda_i\}$?
  - Pre-specify and do sensitivity study
  - Estimate them: iterative reweighting
  - Theory is not definitive.

# Challenge 3a: Multiple datasets

- Data conflicts: Can affect interpretation of population dynamics
- Case study: Prince William Sound herring
  - Data since 1980
  - Exxon Valdex oil spill: March, 1989
  - Age-structured model, multiple datasets
  - Conflict between mile-days of milt and egg production
  - No *a priori* reason to reject either dataset

# Conflict between reproductive datasets




•Greater belief in Mile-days of Milt: Decline in egg production and spawning biomass began in 1989.

•Greater belief in Egg Survey: Egg production and spawning biomass collapsed in 1993.

# Challenge 3b: Conflicts

- Indirect conflicts with other datasets: spawning and catch age composition, disease prevalence
- At least it is better to expose conflicts and state uncertainty than to ignore it or hide it.

# Challenge 4: Parameter inflation for biological realism

- For each year of new data, any number of parameters can chang $t \rightarrow \infty, p \rightarrow \infty$ )
- Examples: natural mortality, gear selectivity, survey catchability, maturity
- There is little theory for highly-parameterized models
  – Solution: AICc, BIC, DIC for parsimony

# Summary

- Both biological and statistical issues are critical in fishery modeling
- Lots of data; lots of parameters, yet we still feel uncertain
- Innovative solutions have and will occur.
- Many interesting theoretical issues need attention.