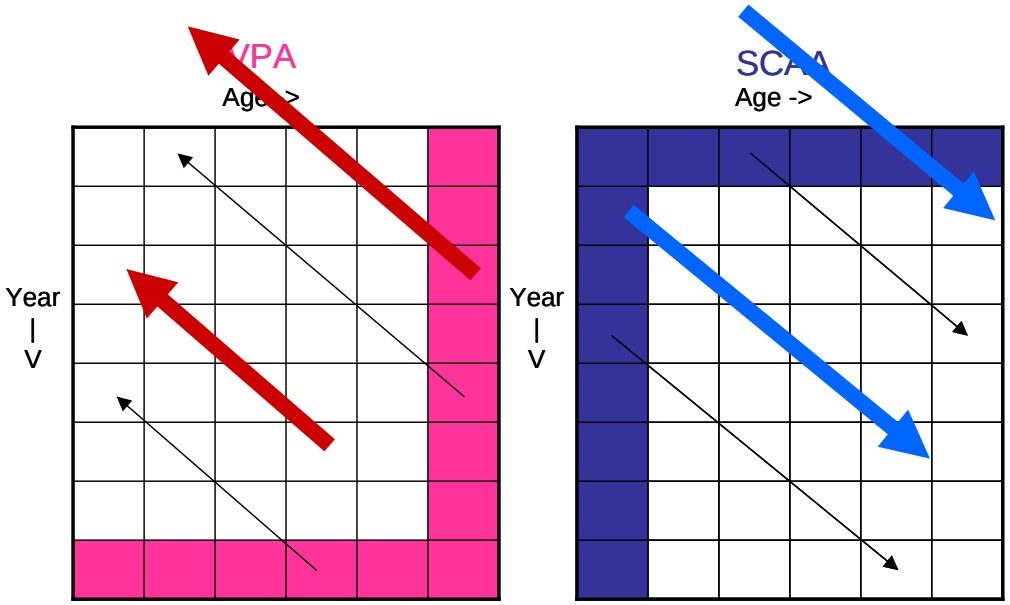
Integrated Analysis

Modified from Richard Methot

Outline

- VPA, Stat. Catch At Age, Integrated Analysis
- IA Concepts
- History of IA
- IA Process
- Data and Priors
- Stock Synthesis Intro

VPA vs. SCAA

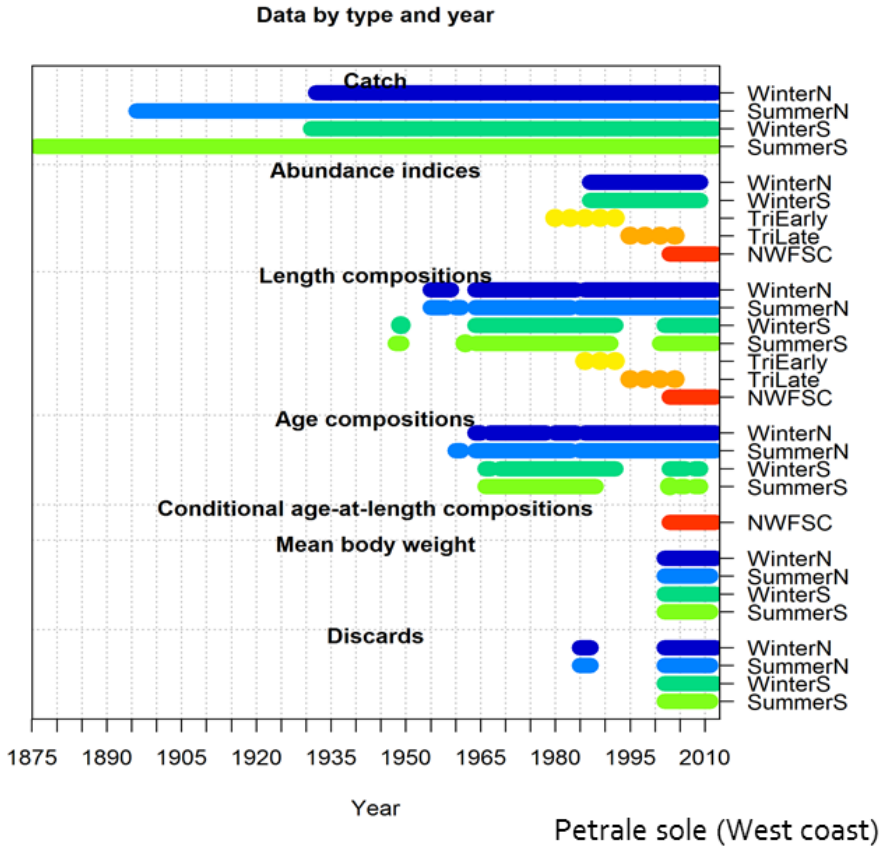


- Calibrated VPA
  - Estimates abundance of the oldest age and current cohorts
  - Calculates abundance back in time
  - Assumes negligible error in the catch at age
  - F-at-age mostly unconstrained
- SCAA
  - Estimates initial abundance at age, recruitments, fishing mortality, selectivity
  - Calculates abundance forward in time
  - Allows error in the catch at age

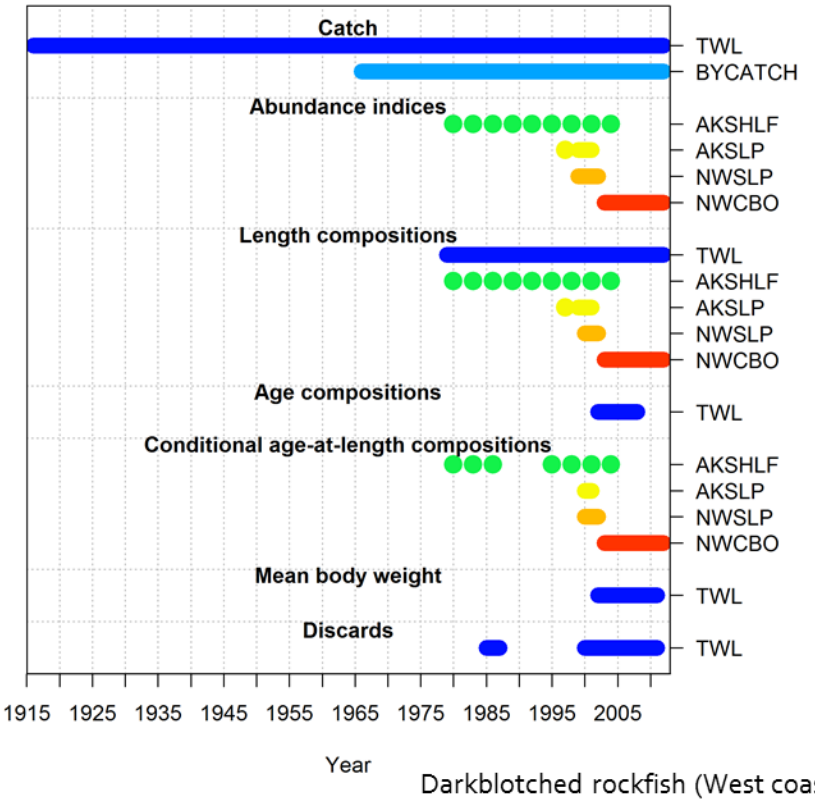
Why Integrated Analysis?

- Long time series of quality catch-at-age and index data are often not available. In response, we may:
  - Truncate time series to shorter period; losing contrast
  - Create catch-at-age from inadequate data sources; losing sense of imprecision
  - Switch to biomass dynamics model with simple parameters linked to population r & K
- Integrate Analysis can:
  - Span data-poor historical periods and current data-rich era
  - Compare its expected values to wide variety of data types
  - Link to population dynamics through spawner-recruitment

Why Integrated Analysis
Data available only for some years



Petrale sole (West coast)



Darkblotched rockfish (West coast)

IA – SCAA Comparison

- SCAA is built around use of fishery catch at age
- IA is a broader and more flexible concept
  - Biological characteristics of catch can be represented by size composition, weight composition, or data-free (biomass dynamics model)
  - Multiple fleets routinely included
  - Predators can be additional sources of mortality
  - Alternative information sources (tag-recapture)
  - Spatial dynamics and movement
  - Less empirical input (such as body wt-at-age)
  - More modeling of processes (growth, size-selectivity, ageing imprecision)

History of Integrated Analysis

- Fournier & Archibald (1982) provided explicit consideration of errors and use of auxiliary information.
- CAGEAN (Deriso et al 1985) - 10s of parameters
- Stock Synthesis (Methot, 1989) -10s to 100s of parameters; FORTRAN & numerical derivatives
- AD Model Builder (late 1980s) - Computer software to build your own IA, 10s to 1000s of parameters. www.admb-project.org
- MULTIFAN-CL (1998) - 1000s of parameters (age and size, tag recapture, movement)
- ASAP (Legault& Restrepo, 1998). A flexible forward age-structured assessment program.
- Coleraine (Hilborn, Maunder et al, 2000) – comparable to ASAP
- CASAL (Bull et. al 2004; New Zealand) C++ algorithmic stock assessment laboratory); age and size structured, tag recapture, movement
- GADGET (Begley & Howell, 2004) Globally applicable Area-Disaggregated General Ecosystem Toolbox
- Stock Synthesis 2 (Methot, 2005) – ADMB-based; size & age based model with spatial structure, gender and growth-morphs

Bring Model to the Data

- Don’t transform data to meet rigid model structure
- Do add processes to model to develop expected values for diverse, lightly processed data
  - Improves understanding of processes
  - Allows simultaneous use of more types of data
  - Statistical properties of data are preserved and transferred to variance of final model results

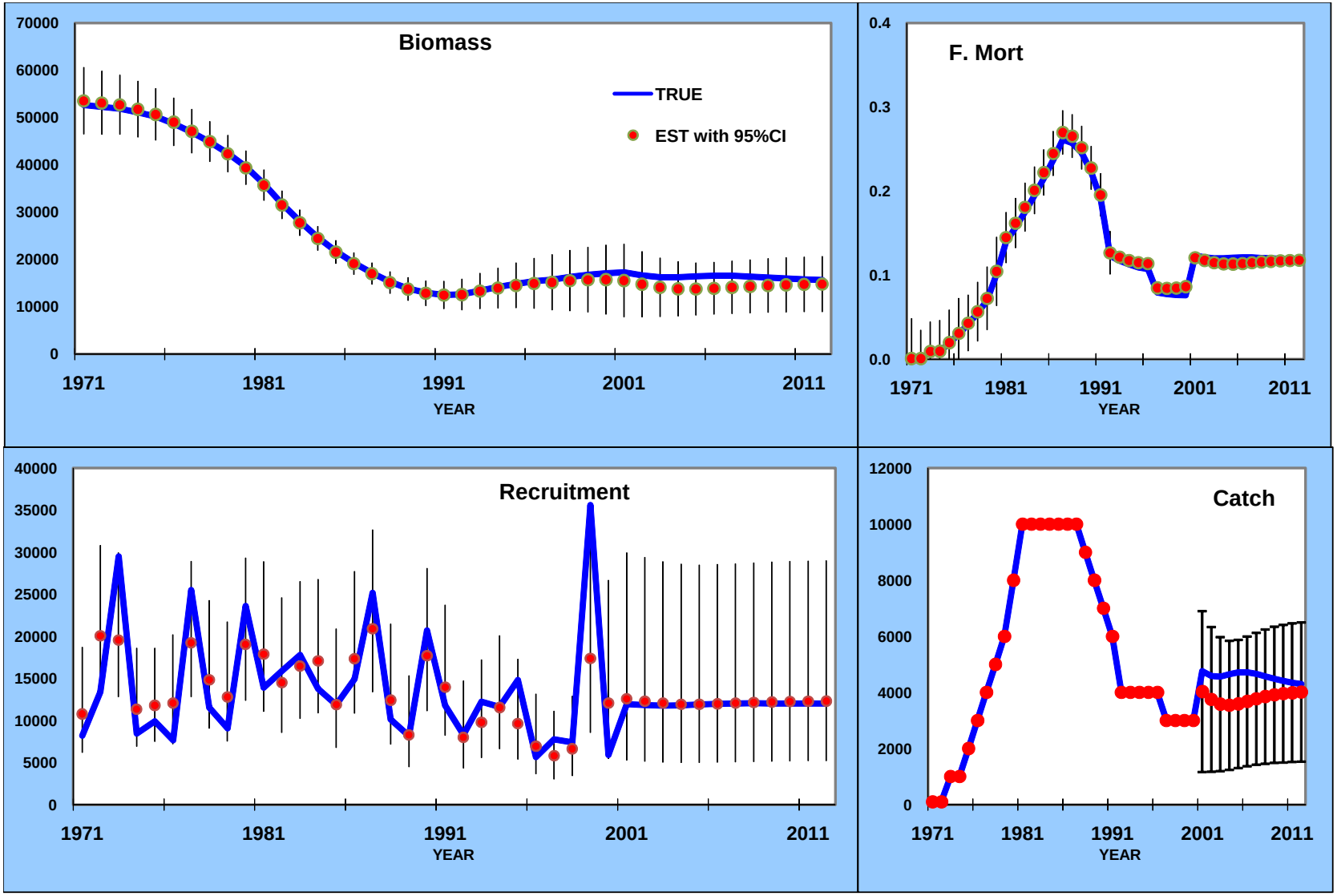
Estimation, Benchmarks, Forecasts

- Typically, we use a sequence of separate analyses:
  1. Estimate population abundance time series
  2. Calculate benchmark quantities: target and limit F rates, sometimes based on first fitting spawner-recruitment curve.
  3. Forecast future abundance and catch using the target F
- Integrated Analysis can:
  - Bring all steps into one analytical package
  - Parameter variance from population estimation gets propagated to quantities in forecast
  - Example output: probability that stock abundance will dip below the overfished threshold 5 years into the future, and the standard error of this probability

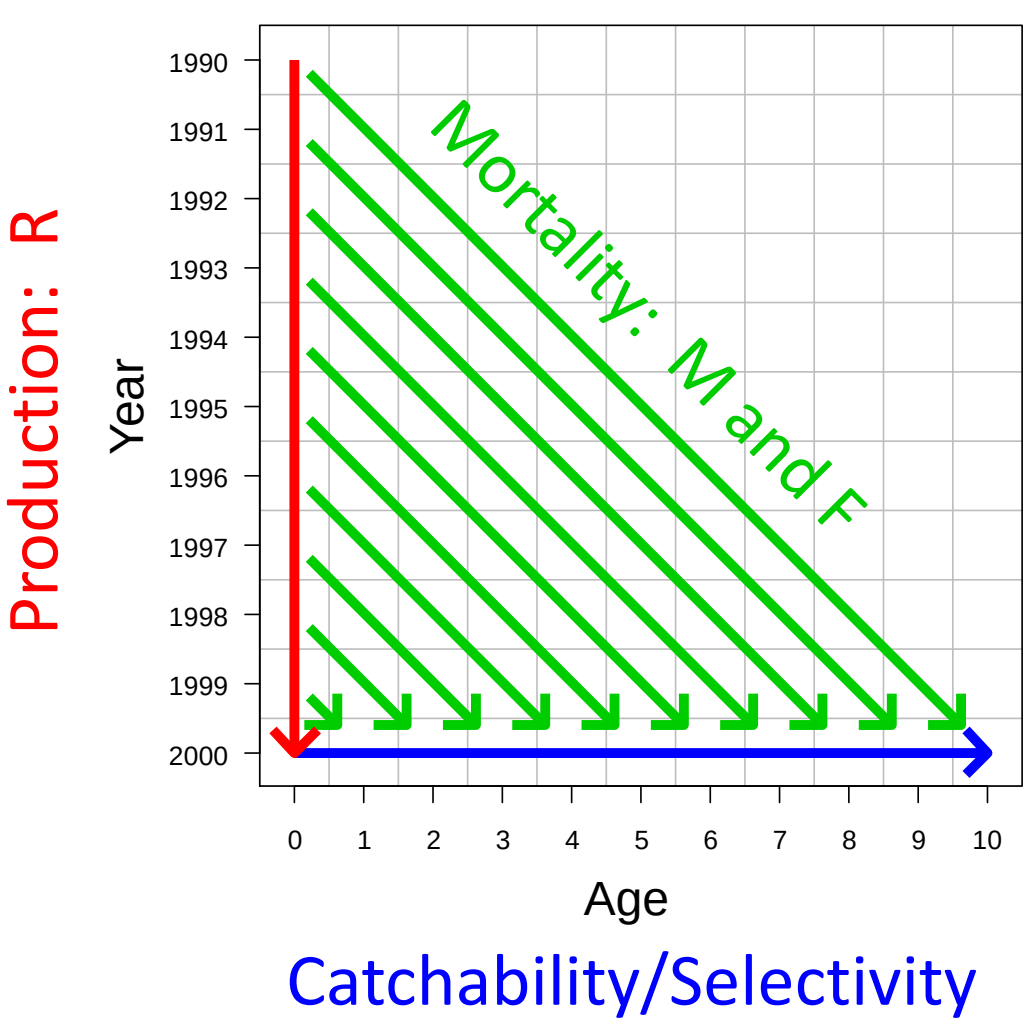
Integrated Analysis Sub-Models

- Population Model
  - Recruitment, mortality, growth
  - Age and/or size structured
- Observation Model
  - Derive expected values for data
- Likelihood-based Statistical Model
  - Quantify goodness-of-fit
- Algorithm to search for parameter set that maximizes the likelihood
  - Auto-Differentiation Model Builder (ADMB)
- Cast results in terms of management quantities
- Propagate uncertainty onto confidence interval for management quantities

Integrated Analysis Output

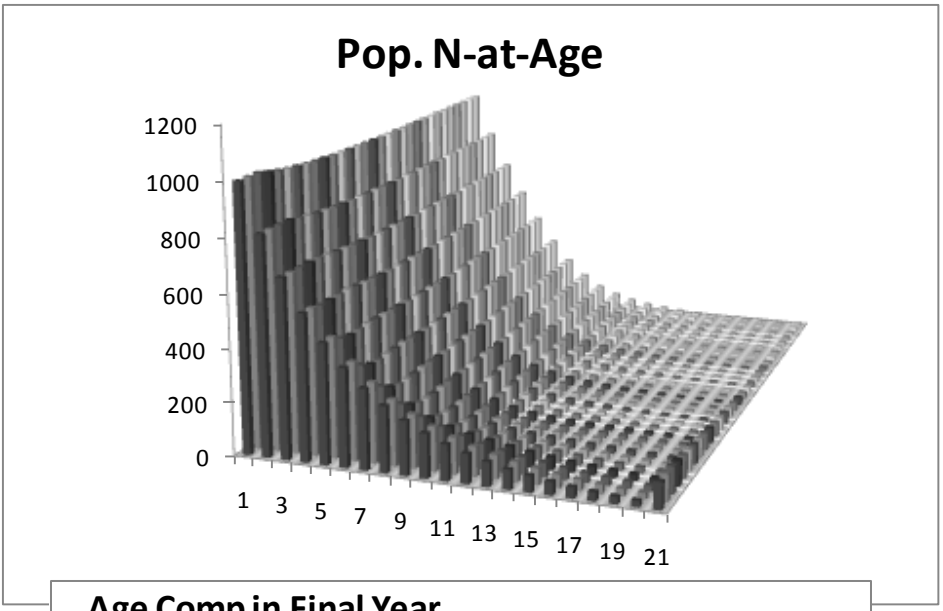


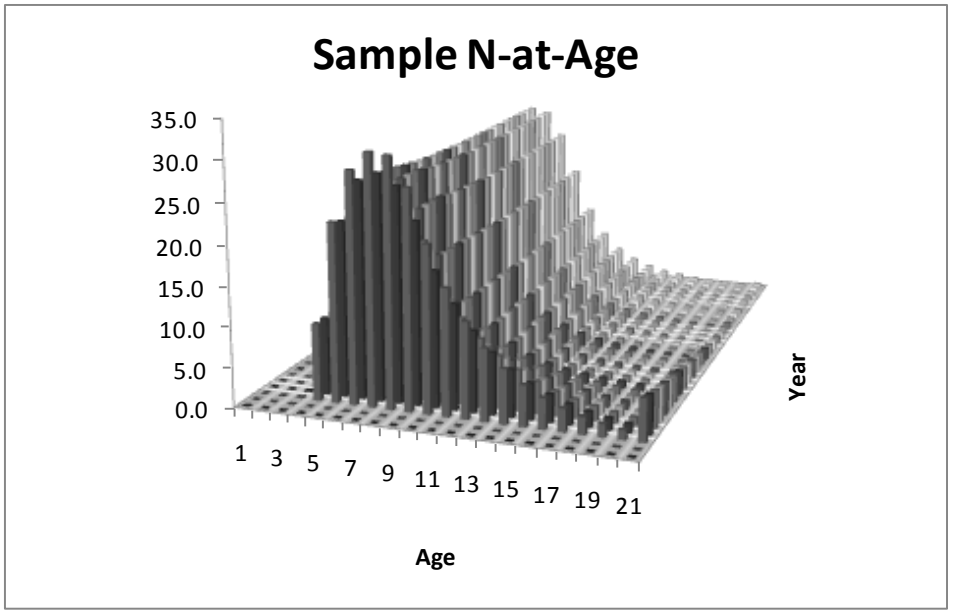
Fundamental Processes

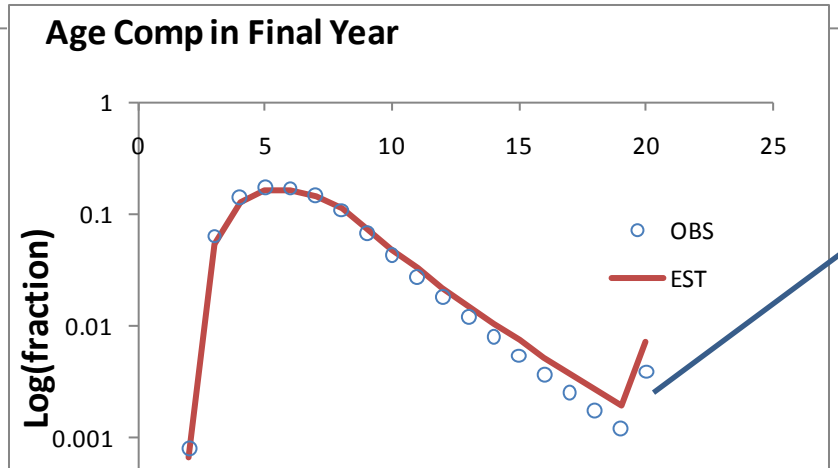


- Population
  - Production: R
  - Mortality: M and F
  - Visibility: Q and Selectivity
  - Location and Movement
- Individual
  - Growth
  - Maturation and Fecundity

Population Scenario



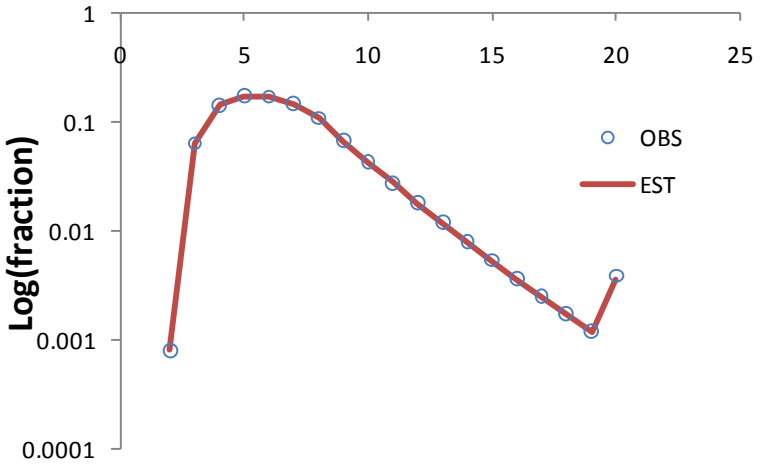




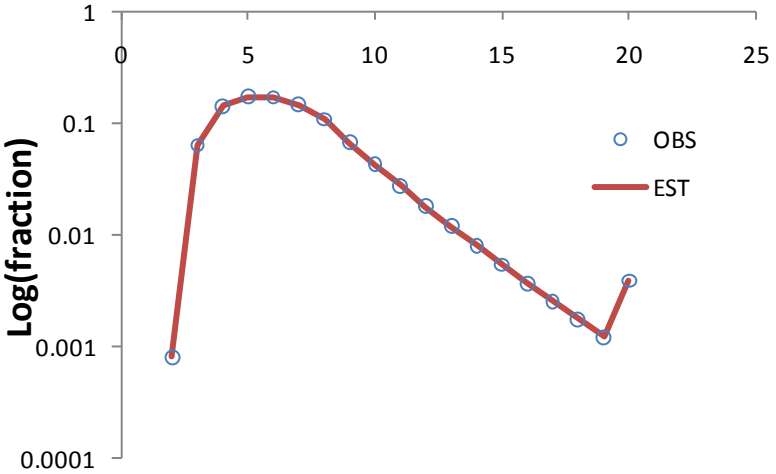
What process' parameter is causing this deviation?

Equally Likely Solutions

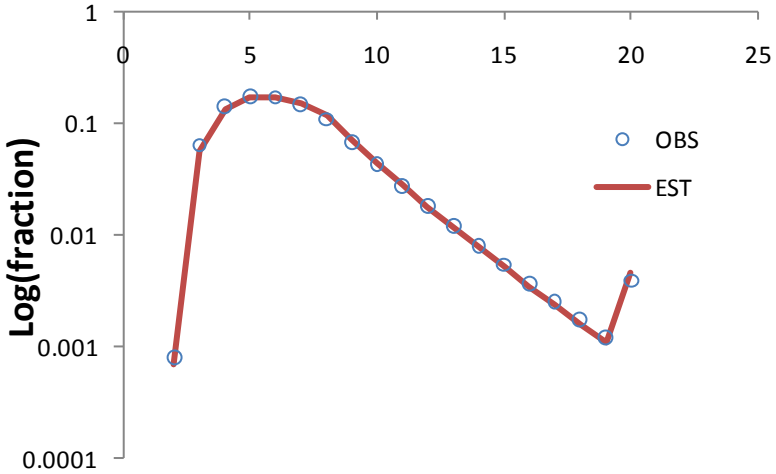
Change M from .2 to .24



change recruit trend from -.03 to 0.0



Change dome-selex from -.1 to -.2



- Productivity, mortality and selectivity are confounded
- Attempt to estimate all 3 parameters with one datum would produce parameter correlation near 1.0
- Unique solutions require more data with sufficient contrast along relevant dimensions

IA: No Magic Bullet

- Allows many kinds of data, but data does not assure contrast
- Allows many processes to be investigated, but cannot magically remove confounding
- Fixing parameter values for some processes (M) will tighten confidence intervals by excluding some alternative explanations for the data
- Result probably will have more variance than result from a simpler model – that's good
- A fishery interacting with its ecosystem is complex process; our models should not overly simplify this process just because the data are lacking

From Data to Results

- Result will be a complex function of fit to all included data;
  - Type, contrast and precision of data determine its influence
  - Examine residuals and root mean squared error of fit to data
  - Parsimoniously, add enough process to remove pattern to residuals
  - Judicious re-weighting of inputs to match rmse of output

Data, Penalties, Priors

- Penalties and Priors are information about parameters in a model
  - Example: maximum age used to create prior on M
- Data are information about a derived quantity
  - Expected value for this quantity is derived from model parameters and structure
  - Example: Age composition of catch from a fleet
- In IA, the expected value for maximum observed age could be derived as a function of M, then observed maximum age could be included as model data
- Concept of Data and Priors blur; it's all information

Relation to Bayesian Analysis

- **Bayesian Analysis (BA) requires prior *pdf* for all parameters and integrates across this *pdf* and the posterior *pdf* to create a *pdf* of results**
- **Information (meta-data) typically used to create priors for a Bayesian analysis can be included as likelihood penalties in IA for some, but not necessarily all, parameters**
  - **Maximum likelihood (MLE) result is like mode of BA**
  - **MCMC with IA's parameter penalties produces *pdf* like that of BA**
  - **The normal *pdf* based on the uncertainty of the MLE is comparable to the BA *pdf* from MCMC, at least in many applications. More on this tomorrow.**

Stock Synthesis

- An implementation of IA
- Age and size structured, with geographic areas
- High diversity of data types, including tags
- Fully integrates population estimation, benchmark calculation, and forecasting