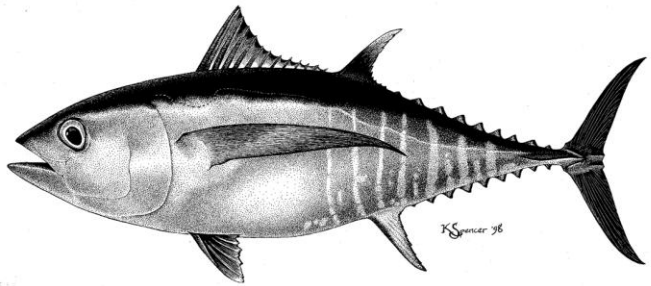
Integrated analysis
- a brief introduction

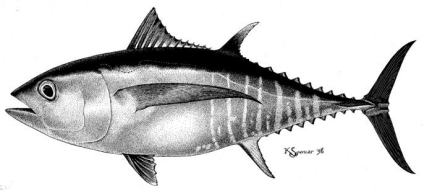


Alexandre Aires-da-Silva and Mark Maunder

IATTC

Outline

- Review of historic method (“two-step”)
- Definition of integrated analysis (“one-step”)
- Example
- Advantages/disadvantages



Traditional “two-step” approach

STEP 1

Fisheries raw data analyzed

Output

Summary statistics

Provide information for parameter estimation

STEP 2

Summary statistics used in population dynamics model

EXAMPLE

STEP 1

Tagging Petersen analysis

Output

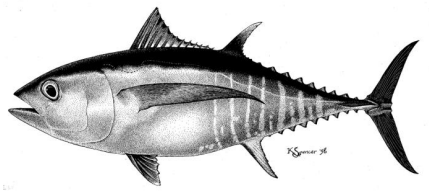
Population size estimate

STEP 2

Population size estimate used in the fitting of a population dynamics model

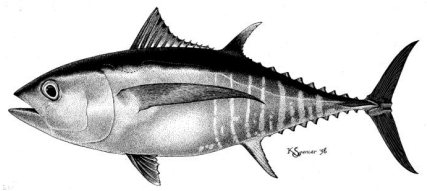
Examples of the “two-step” approach

- Cohort analysis → Stock recruitment analysis
- LF analysis, tagging analysis, age data → YPR analysis
- LF data → Cohort analysis
- Tagging Petersen analysis → Stock assessment model



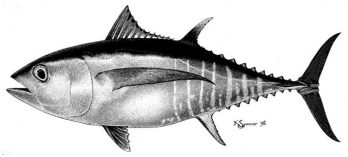
Disadvantages of the “two-step” approach

- Loss of information in the summarization process
- Inconsistencies in assumptions
- Difficulty in determining error structure
- Difficulty in transferring uncertainty
- Reduced diagnostic ability



Integrated analysis - definition

- GOAL: Use all the information in the data and adequately propagate uncertainty
- HOW:
  - Include all data in a single analysis
  - Use data in its rawest form
  - Formulate analyses so that they share parameters for the processes that they have in common
  - Combine objective functions
  - Estimate all parameters simultaneously while optimizing the combined objective function



Model development framework

- Deterministic population dynamics model

$$B_{y+1} = f\left(B_y\right) - C_y$$

- Observation model

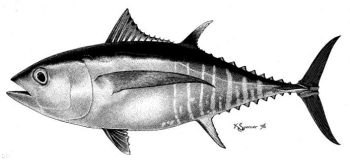
$$CPUE_y = qB_y$$

- Statistical models

$$B_{y+1} = f\left(B_y\right)e^{\nu_y} - C_y \qquad \nu \sim N\left(0, \sigma_\nu^2\right)$$

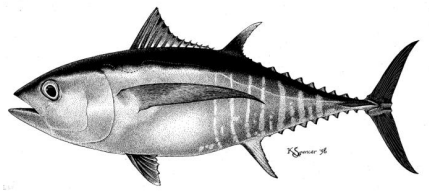
$$CPUE_y = qB_y e^{\varepsilon_y} \qquad \varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$$

- Estimation method
  - Maximum likelihood, Bayesian integration
  - Newton's method, MCMC, SIR



Integrating tagging and population dynamics analyses – simple example

- Tag recapture data – estimate survival and population size parameters
- Forward simulation population model – simulate forward from known population size
- Catch and effort data



Tagging model – Multinomial (simplified)

Not caught Caught

$$L\left(\theta_{\text{tagging model}} \mid \text{tagging data}\right) = \left[1 - \sum p\right]^{N-\sum n} \prod_i p_i^{n_i}$$

$$p_{relage=2,recage=4} = \tau S_{t=1,a=2} S_{t=2,a=3} U_{t=3,a=4}$$

$$N_{a,t} = \frac{n_{a,t}}{U_t}$$

$$S_{a=1,t=1} = S_{a=2,t=1} = S_{a=1,t=2} \ldots ?$$

$$U_{a=1,t=1} = U_{a=1,t=2} = U_{a=2,t=1} \ldots ?$$

Population model

$$N_{t+1,a+1} = N_{t,a} e^{-\left(M_a + F_{t,a}\right)}$$

$$F_{t,a} = qE_t s_a$$

$$s_a = \left[1 + \exp\left(-\ln(19)\frac{a - a_{50}}{a_{50} - a_{95}}\right)\right]^{-1}$$

Reparameterized tagging model

$$S_{a,t} = \mathrm{e}^{-\left(M + qE_t s_a\right)}$$

$$U_{a,t} = (1 - e^{-(M + F_{a,t})}) \frac{F_{a,t}}{M + F_{,t}}$$

$$F_{a,t} = qE_t s_a$$

$$p_{rel=2,rec=4} = \tau \, \mathrm{e}^{-\left(M + F_{a=2,t=1}\right)} \, \mathrm{e}^{-\left(M + F_{a=3,t=2}\right)} (1 - e^{-(M + F_{a=4,t=3})}) \frac{F_{a=4,t=3}}{M + F_{a=4,t=3}}$$

$S_{t=1,a=2}$ $S_{t=2,a=3}$ $U_{t=3,a=4}$

Adding information for catch and effort data

$$\hat{C}_{t,a} = \frac{F_{t,a}}{Z_{t,a}}\left[1-\exp\left(-Z_{t,a}\right)\right]N_{t,a}$$

$$Z_{t,a} = M_a + F_{t,a}$$

$$F_{t,a} = qE_t s_a$$

$$L\left(\theta_{\text{population model}} \mid \text{catch, effort}\right) = \prod_{t,a}\frac{1}{\sqrt{2\pi\sigma}}\exp\left[-\frac{\left(C_{t,a}^{obs}-\hat{C}_{t,a}\right)^2}{2\sigma^2}\right]$$

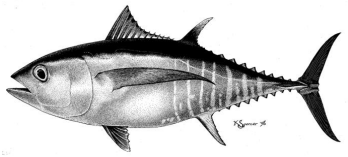
Integrated Model

$$L(\theta \mid \text{data}) = L(\theta \mid \text{taging data})\, L(\theta \mid \text{catch and effort})$$

$$\theta = \theta_{\text{tagging model}} \cup \theta_{\text{population model}}$$

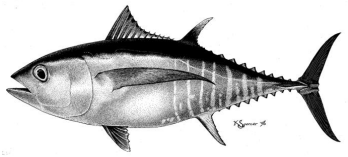
Advantages of integrated analysis

- Includes all the information
- Consistent assumptions
- Error structure based on sampling
- Uncertainty propagated automatically
- All residuals available in one analysis
- Flexibility



There are problems...

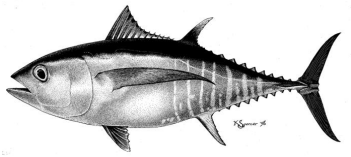
- Computational time constraints
- Convergence
- Parameter confounding
- Model misspecification
- Weighting factors for each data set



IATTC workshops

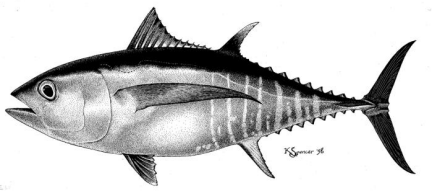
- Oct 2005: IATTC workshop on “Stock Assessment Methods”
- Oct 2007: IATTC workshop on “Using Tagging Data for Fisheries Stock Assessment and Management”
- Oct 2008: IATTC workshop on “Spatial Analysis for Stock Assessment"

For reports and additional information visit:
http://www.iattc.org/MeetingsENG.htm

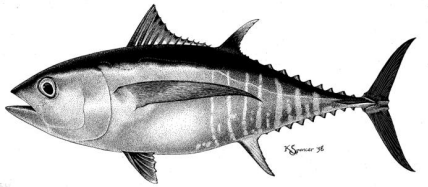


Data - Comparisons

Data	MULTIFAN-CL	Stock Synthesis II	CASAL
Catch-at-age	X	X	X
Catch-at-length	X	X	X
Abundance index (CPUE)		X	X
Catch-at-weight	X	X	
Age-length		X	X
Average weight		X	
Discard (fit)		X	
Tagging	X	X	X

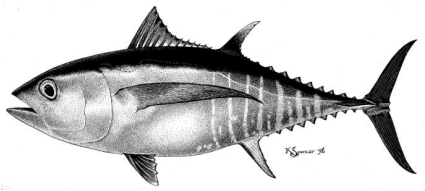


General - comparisons



	MULTIFAN-CL	Stock Synthesis II	CASAL
Bayesian	No	MCMC	MCMC
Bootstrapping	No	Automatic	Automatic
Review	Publication review, comparisons with A-SCALA and MF-CL (no spatial or tagging)	Modeling workshop with independent review. Panel intensive reviews, comparisons with other models, simulation tests, comparisons with other models at UW	Comparison with Coleraine in several assessments, comparison with existing Hoki and Paua models, comparisons with other models at UW, applications reviewed by independent experts
Assessments	WCPO YFT, BET, ALB, SKJ, BUM, SWO, Blue shark, Lobster, Atlantic BET and ALB	15 west coast and Alaska groundfish assessments, SEPO swordfish, EOP BET, PBF, WCPO ALB	From 10 to 20 stocks in NZ, fin fish and shellfish
Aprox. max. params in application	3000*	200	200
Aprox. time required for EPO YFT moel	4 hours	40 min	Not evaluated

Model structure - comparisons



Structure	MULTIFAN-CL	Stock Synthesis II	CASAL
Spatial population dynamics	Yes	Yes	Yes
Modeling of discards	No	Yes	No
Sex-structured	Under development	Optional	Optional
Selectivity	Functional forms, nonparametric with smoothless penalties, and cubic splines	Functional forms and nonparametric	Functional forms and nonparametric with smoothless penalties
Selectivity basis	Age or length	Age, length, and sex	Age, length and partition
S-R relationship	B-H	B-H, Ricker	B-H, Ricker
M	Full age-structure with smoothness	2 breakpoints, full age-structure	Full age-structure with smoothness
Aging error	No	Yes	Yes
Variable lenght bin size	No	Yes	Yes