HYBRID – A MODELLING FRAMEWORK TO SIDESTEP STRUCTURAL UNCERTAINTY IN MODELS

Divya Varkey, Jonathan Babyn
Paul Regular, Rajeev Kumar

INTRODUCTION: A SMALL NICHE

Dr. Maunder's questionnaire highlights the vast use of varied approaches in fisheries stock assessment modelling.

Within the vastness, HYBRID represents a small niche for data types:
- Survey indices for abundance at age
- Catch-at-age data

HYBRID is a modelling framework to explore multiple models
- allows the user to explore and compare different model structures.
  - **what-if** we modelled fisheries selectivity differently?
  - **what-if** natural mortality is changing over time?

Method of multiple working hypotheses (Chambelain 1890)
→ Hilborn and Mangel's Ecological Detective "confrontation between more than one model arbitrated by data underlies science"

HYBRID STRUCTURE

TMB – R package

Built as a generic modelling framework with different options for

- F structure
- M structure
- Fitting catch-at-age
- Fitting surveys: Missing data points

Flexdashboard for model comparison (Dr. Paul Regular)

THE STATE EQUATION

State equation follows the parameterization in the State-space Assessment Model (SAM) (Nielsen and Berg, 2014)

Recruitment: Only Random walk

$$log\ N_{1,y} = log\ N_{1,y-1} + \eta_{1,y};\ where, \eta_{1,y} \sim N(0, \sigma R)$$

$$log\ N_{a,y} = log\ N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y};\ where, 2 \leq a < A; \eta_{2:A,y} \sim N(0, \sigma P)$$

$$log\ N_{A,y} = log\begin{pmatrix} N_{A,y-1} * \exp(-F_{A,y-1} - M_{A,y-1}) + \\ N_{A-1,y-1} * \exp(-F_{A-1,y-1} - M_{A-1,y-1}) \end{pmatrix} + \eta_{A,y};\ where, A = plus\ group$$

PARAMETERIZATION OF F

F as parameters and Fit to catch-at-age

5 Options for time-varying fisheries selectivity in the model

- Different levels of flexibility in the connection between ages and years
- Key questions:
  - Does the age pattern change over time – was there change in gear composition in the fishery?
  - How much dependence in F between years?

Option 1: Non parametric (not time varying)

- For each age (fixed age pattern)
- Random walk over ages (Cadigan 2010)
- Time-blocks can be implemented

$$\log(s_a) = \log(s_{a-1}) + \omega_a;$$
$$where, \omega_a \sim N(0, \sigma_s)$$

$$F_{a,y} = s_a * f_y$$

PARAMETERIZATION OF F

Option 2: Parametric

Little flexibility in pattern over age.

- Logistic (flat-topped)

$$s_a = \frac{1}{1 + \exp\left(-b_1(a - a_{50})\right)}$$

- Double logistic (domed)

$$s_a = \frac{1}{1 + \exp(-b1(a - a1_{50}))} \cdot \frac{1}{1 + \exp(b2(a - a2_{50}))}$$

- Time blocks implemented
  (Radomski et al. 2005)

$$\log(a_{50\,y}) = \log(a_{50}) + sdev_y;$$
$$where, sdev_y \sim N(0, \sigma_{sel2})$$

- Random variation within time blocks

$$F_{a,y} = s_a * f_y$$

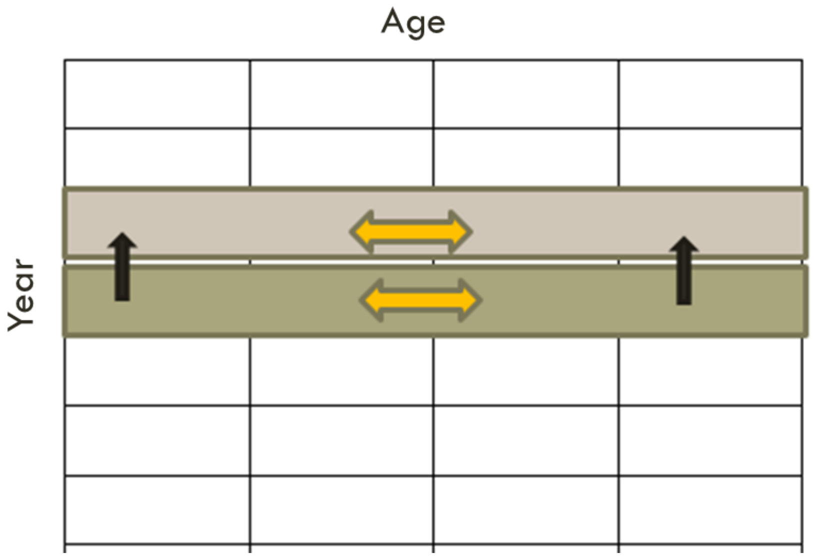
PARAMETERIZATION OF F

Option 3: MVN Random Walk (Nielsen and Berg, 2014)

- Flexibility in age and year patterns
- Multivariate Normal (MVN) random walk over years
- Autoregressive (AR) process for the correlation between ages
- similar age groups develop similar trends in the fishing mortality

$$\log(F_{1:A,y}) = \log(F_{1:A,y-1}) + e_{1:A,y};$$
$$where, e_{1:A,y} \sim MVN_{1:A}(0, \Sigma)$$

$$\Sigma_{a,\bar{a}} = \rho^{|a-\bar{a}|}\sigma_a^2$$



PARAMETERIZATION OF F

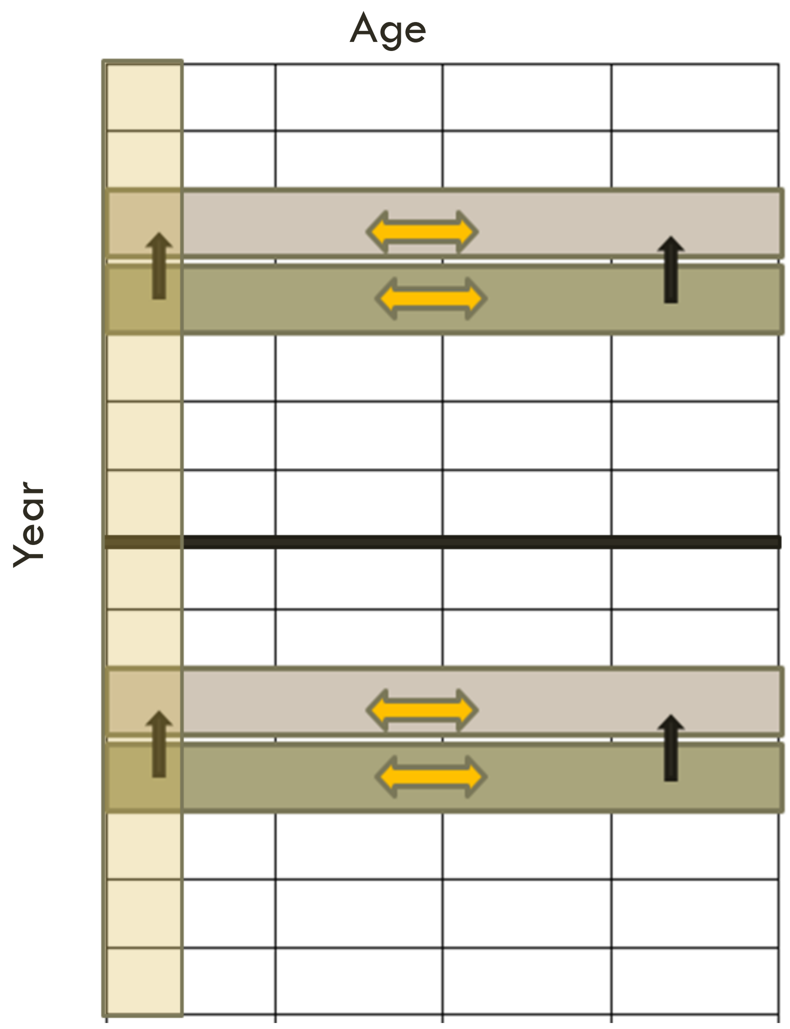
Option 4: Similar to option 3

To account for fisheries management changes:

- Restart the MVN random walk at the beginning of the fishing moratorium

F on young ages may not correlate with F on older ages

- De-correlate the standard deviation for the young ages
- Choice for which ages to de-correlate in the covariance matrix



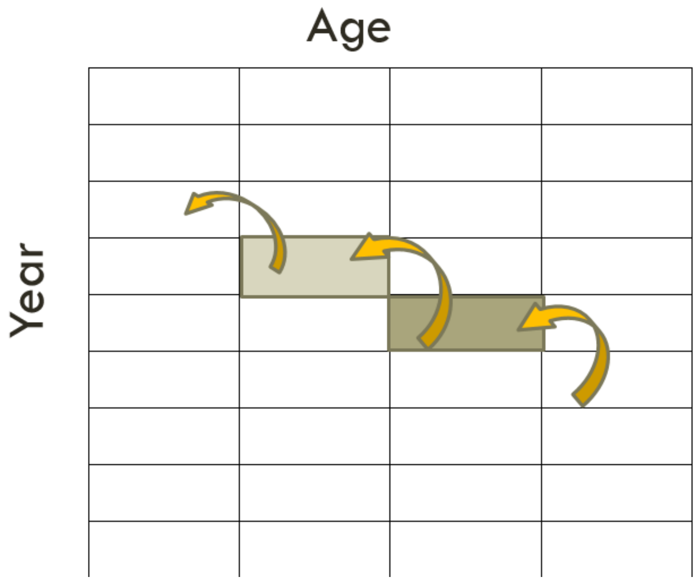
PARAMETERIZATION OF F

Option 5: Correlated separable AR1 pattern in year and age

F in a given age and year is the product of a mean F and correlated age-year deviations (Cadigan 2016)

$$log\ F_{a,y} = \mu\ logF_{a,y} + \Delta_{a,y}$$

$$Corr[\Delta_{a,y}, \Delta_{a-m,y-n}] = \varphi_{Fa}^{|m|}\varphi_{Fy}^{|n|}$$



Stronger connection between years compared to Options 3 and 4

Perhaps ideal for fisheries that target strong cohorts moving through the fishery.

FITTING CATCH-AT-AGE

2 Options based on Reliability of Catch Numbers-at-age data

Option 1: Fairly reliable time series

- Fit to Catch Numbers at age

Option 2: Reliability of time series varies over time (Cadigan 2016)

- Fit to Catch-Proportions at age
- Magnitude of the catch fit using landings
- Censored likelihood for landings to account for different levels of reliability of the catch magnitude over time.

FITTING CATCH-AT-AGE: OPTION 2 CONTD..

$$P_{a,y} = \frac{C_{a,y}}{\sum_1^A C_{a,y}}$$

$$\pi_{a,y} = Prob(age = a | age \geq a) \frac{P_{a,y}}{\sum_a^A P_{a,y}}$$

$$X_{a,y} = \log\left(\frac{\pi_{a,y}}{1-\pi_{a,y}}\right); where\ a = 1{:}A-1$$

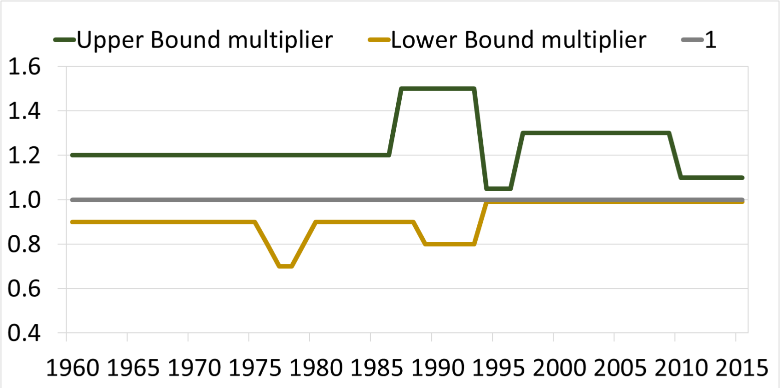
Proportions at age using continuation ratio logits

$$Xo_{a,y} = X_{a,y} + \epsilon_{a,y}; where\ \epsilon_{1:A-1,y} \sim MVN(0, \Sigma)$$

Landings using censored bounds

- Where LB and UB are lower and upper bounds
- Fairly flat likelihood inside bounds depending on σ_L

$$l(L_{obs1,\ldots}L_{obsY}|\theta) = \sum_{y=1}^{Y} log\left\{\Phi_N\left[\frac{\log\left({}^{UB_y}/_{L_y}\right)}{\sigma_L}\right] - \Phi_N\left[\frac{\log\left({}^{LB_y}/_{L_y}\right)}{\sigma_L}\right]\right\}$$



- For more detail, please see Cadigan 2016 and Bousquet et al. 2010 for more details

PARAMETERIZATION OF M

Option 1: Invariant over age and year

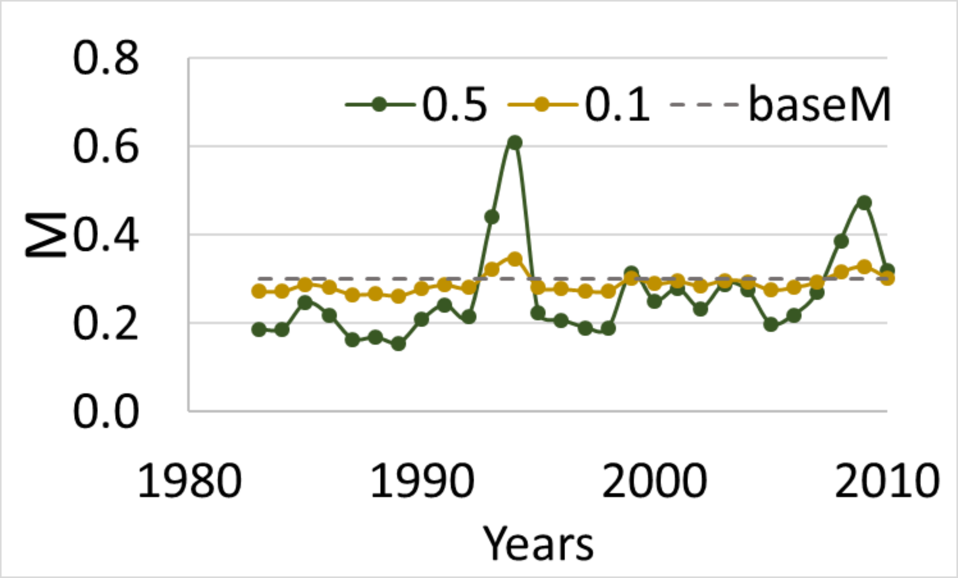
Time varying options:

Option 2: Size specific (Miller and Hyun 2017)

$$\log M_{a,y} = b_0 + b_1 * \log W_{a,y} ;$$
$$where\ b_1 = -0.305\ (Lorenzen\ 1996)$$

Option 3: Mortality follows trend in an index

- Scales above or below a base level M
- Equation structure from Kumar et al. 2013
- Estimates parameter Mscale for effect of index



$$M_{a,y} = \mathrm{baseM} * \exp\left(Mscale_a * Normalized\ Index_y\right)$$

Option 4: Mortality follows trend in an index - *Additive effect (not implemented)*

LIKELIHOOD FOR FITTING TO SURVEY INDICES OF ABUNDANCE

The model fits to log indices from the survey

$$\log I_{a,y,s} = \log q_{a,s} + \log N_{a,y} - sf_{y,s} * Z_{a,y} + e_{a,y,s}; where\, e_{a,y,s} \sim N(0, \sigma_{ag,s})$$

Choice to use censored likelihood for missing values, or ignore the missing values

- When censoring is applied, the log-likelihood will be very small if the predicted index is lower than the bound (Cadigan 2016).

$$l(I_{a,y,s} = 0|\theta) = log\left\{\Phi_N\left[log[0.004/E(I_{a,y,s})]\right]/\sigma_{ag,s}\right\}$$

SURVEY CATCHABILITY

Estimated parameters feed into a matrix of catchability by age and survey

Between survey series: allows for in model adjustments between surveys

- Such as correction factor for change in vessel/gear over time in surveys
- For example: Catchability of Survey S3 depends is related to catchability of Survey S1

Parameters for q

	qS1	qS2
age1	0.2	0.1
age2	0.3	0.15
age3	0.32	0.4
age4	0.35	0.5
age5	0.36	0.5
age6	0.4	0.5
age7	0.4	0.5

Survey catchability in model

	qS1	qS2	qS3<->qS1
age1	0.2	0.1	0.3
age2	0.3	0.15	0.45
age3	0.32	0.4	0.52
age4	0.35	0.5	0.55
age5	0.36	0.5	0.56
age6	0.4	0.5	0.6
age7	0.4	0.5	0.6
age8	0.4	0.5	0.6
age9	0.4	0.5	0.6
age10	0.4	0.5	0.6

CORRELATED YEAR-EFFECTS

All ages are not necessarily affected equally

Applied in our model as survey year-effects –

- Explore changes in catchability over time
- Value is unclear – when it cannot be propagated forward

$$YE_{1,y,s} = N\left(0, \left(\frac{\sigma_{ye_s}^2}{\left(1 - {\varphi_{ye}}_s^{\,2}\right)}\right)^{1/2}\right)$$

$$YE_{2:A,y,s} = N\left({\varphi_{ye}}_s * YE_{a-1,y,s}, {\sigma_{ye}}_s\right)$$

SUMMARY OF FEATURES

F	F-Case 1: Non-parametric age effect
	F-Case 2: Logistic
	F-Case 3: MVN Random walk (SAM) Style
	F-Case 4: Year and Age split in the MVN Random walk
	F-Case 5: Age and Year correlated structure
M	M-Case 1: Fixed (M=0.3)
	M-Case 2: Size dependent
	M-Case 3: Scaled index
R	Random Walk
Catch	Limited uncertainty: Fitting to Catch Numbers at age
	Uncertainty about magnitude of catch over the time series: Fitting catch proportions and magnitude separately, censored fitting of landings, Correlated likelihood
Survey	Ignore zeroes
	Censored likelihood for missing data points
	Survey year effects
others	Retrospective patterns, Projections

MODEL COMPARISON IN FLEXDASHBOARD

Compare models based on

1. Residuals for survey-fits by age
2. Residuals for catch-at-age fits
3. Process error comparisons
4. Model outputs (F-at-age, Recruitment, Biomass etc..)

QUESTIONS/ CLARIFICATIONS/ FEEDBACK?

REFERENCES

Cadigan, N.G. (2010). Trends in Northwest Atlantic Fisheries Organization (NAFO) Subdivision 3Ps Cod (*Gadus morhua*) stock size based on a separable total mortality model and the Fisheries and Oceans Canada Research Vessel survey index. DFO Can. Sci. Advis. Sec. Res. Doc. 2010/015.

Cadigan, N. G. (2016). A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 73(2), 296-308.

Chamberlain, T. C. (1890). The method of multiple working hypothesis. Science 366:92-96. Reprinted in Journal of Geology 5: 837-48. Reprinted in Science 148: 754-59 (1965). Reprinted in Ecological Detective (1997).

Hilborn, R., & Mangel, M. (1997). The ecological detective: confronting models with data. Princeton University Press.

Kumar, R., Martell, S. J., Pitcher, T. J., & Varkey, D. A. (2013). Temperature-driven decline of a cisco population in Mille Lacs Lake, Minnesota. North American journal of fisheries management, 33(4), 669-681.

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of fish biology, 49(4), 627-642.

Miller, T. J., & Hyun, S. Y. (2017). Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. Canadian Journal of Fisheries and Aquatic Sciences, 75(5), 691-703.

Nielsen, A., & Berg, C. W. (2014). Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research, 158, 96-101.

Radomski, P., Bence, J. R., & Quinn Ii, T. J. (2005). Comparison of virtual population analysis and statistical kill-at-age analysis for a recreational, kill-dominated fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, *62*(2), 436-452.