An integrated Catch-MSY model for data poor stocks

Steven Martell

September 30, 2015

Abstract

An age-structured Catch-MSY model was developed to address the large number of stocks for which there is a historical time series of catch data, and sparse axillary information on trends in abundance, or trends in mean size, or estimates of density. The model is parameterized in terms of MSY and F_{MSY}, from which values of unfished biomass and the steepness of the stock-recruitment relationship are derived. Parameter distributions are estimated based on non-statistical criterion, and in cases where there are biomass or composition data, a statistical criterion is also incorporated to update the distributions for model parameters. The age-structured model herein requires additional life-history information that is summarized by the single intrinsic rate of increase r in the Schaefer biomass production model. For data poor stocks, this information can be supplemented with life-history invariants information rather than specifying a range of r values based on a subjective and arbitrary choice of resilience for each stock. The catch-MSY method can first be though of as a tool for exploring Tier 5 methods for setting Annual Catch Limits and then has the potential to graduate to Tier 3, or higher, methods if sufficient information exists to statistically estimate stock status and management related parameters (MSY and F_{MSY}).

Introduction

There are a large number of fish and invertebrate species that are listed in fisheries management plans and each of these species requires an Annual Catch Limits (ACL). However, for a vast majority of these species there is insufficient data to conduct routine stock assessments to determine stock status and set appropriate ACLs. The current practice for setting ACLs for many of these data poor species is based solely on: (a) the historical catch information, or (b) use of ratios such as change in mean size or spawning potential ratio (SPR) to infer depletion, or (c) comparative studies on local density in heavily depleted versus near pristine habitats using under water visual census (UWVC).

There are a number of data poor approaches available in the literature that attempt to address data poor stocks. Among these methods that are currently being considered by the Western Pacific Fisheries Management Council (WPFMC) are: (1) the Catch-MSY approach (Martell and Froese, 2012), (2) changes in mean-size to determine the overall mortality rate (Beverton and Holt, 1993), and (3) use of underwater visual transect and tow-board data to estimate fish density in near pristine and exploited reefs. Each of these methods has their own strengths and weaknesses. For example, the catch-MSY approach is not appropriate for determining current stock status or the long-term sustainable fishing mortality rate. But, the method is very useful for approximating the long-term maximum sustainable yield (MSY). The change in mean-size approach and spawning potential ratios (SPR) require information on growth and natural mortality and are reasonable proxies for determining appropriate harvest rates and can give some indication on the current exploitation status if the change in mean size is known. However, their weakness depends on unbiased growth parameters, reasonably accurate age-at-recruitment to the fishery, and relatively stable recruitment (along with a number of other untestable assumptions). Underwater visual surveys provide estimates of population density, but the limitations of this method is spatial expansion and the assumptions about what constitutes appropriate habitat. It is clear that each of these methods alone is insufficient to determine both the current stock status and reasonable estimates of the underlying stock productivity that are necessary for setting model-based Annual Catch Limits (ACLs).

Many age-structured assessment models are parameterized in terms of unfished spawning stock biomass (B_0) and the steepness of the stock recruitment relationship (h). These parameters are then estimated by fitting a model to time series data on relative abundance, which may or may not include additional axillary information on the composition of the catch, tagging data, changes in mean weight, etc. These two parameters are key in determining the MSY-based reference points, where B_0 provides the overall scaling information for MSY, and the steepness of the stock-recruitment relationship provides key information on the optimal harvest rate F_{MSY}. Many authors have noted that there is no simple analytical transformation of B_0 and h into MSY and F_{MSY} when using age or size-structured models. However, the inverse transformation does have a unique analytical solution (Schnute and Richards, 1998). Forrest et al. (2008) develops the Management Oriented Parameterization approach in an age-structured model and provides a comparisons of methods using a single case study and concludes with identical results using either approach.

In this paper, the hypothesis is that the combination of two or more of the above general methods into a single integrative framework will compliment the strengths and weaknesses of each approach, making the integrated framework more informative about stock status and setting appropriate ACLs for data poor species. The simple Catch-MSY model, as suggested by Martell and Froese (2012), is extended into a age-structured model that is parameterized using the maximum sustainable yield (MSY) as the global scaling parameter, and F_{MSY} as the global rate parameter (instead of steepness or intrinsic rate of growth). The advantage over this parameterization is that prior information on historical removals (as is used in the Tier 5 approach) can be used to directly inform the population model rather than having to specify prior distributions for the population carrying capacity and intrinsic rate of growth as suggested by Martell and Froese (2012) and Sabater and Kleiber (2014).

Summary of CIE review

A review of the Augmented Catch-MSY model was conducted by three CIE reviewers in Honolulu, Hawaii from the 30th June to the 3rd July 2014. The following bullet points are my summary of those review comments.

- All reviewers noted a lack of model documentation, and many referred to the code itself to understand the model structure.

  Response: The new Integrated Catch-MSY model is implemented as an R-package with proper package documentation and vignettes to help guide users of proper use and full documentation of the implemented methods and algorithms.

- Caution is warranted in the interpretation of the distribution of MSY estimates; is is not a proper distribution that is grounded in statistical methods (i.e., posterior density or likelihood profile).

  Response: The new Integrated Catch-MSY model is parameterized using MSY and F_{MSY} as leading parameters from which unfished biomass and steepness are derived conditional on assumed age-at-entry (i.e., selectivity). In cases where there are independent estimate of biomass (or trends in biomass) there is the flexibility to use Maximum Likelihood methods, or Importance Sampling to construct a joint posterior density. In cases where there is no data in which to make a formal statistical comparison, the resulting distribution of MSY values is conditioned solely on the catch data and the joint prior density for MSY and F_{MSY}.

- More simulation and robustness testing, including Management Strategy Evaluation for this data-poor model.

  Response: There was extensive simulation testing of the original Catch-MSY method developed by Martell and Froese (2012) by Rosenberg et al. (2014). In those simulation studies, the authors also examined the performance of the Catch-MSY method and concluded that over a wide range of depletion levels the Catch-MSY method performed relatively well. Moreover, in the original paper, estimates of MSY were compared with other assessments; these were primarily on temperate stocks. No such comparisons have been conducted for the new Integrated Catch-MSY package developed for this project. Furthermore, the effort required to conduct a formal Management Strategy Evaluation (MSE) requires a number of additional components, including specifying objectives in which to develop performance measures and the specification of an operating model to generate alternative states of nature. I agree with the CIE reviewers that this is an important step in developing a management procedure, but not a necessary condition for evaluating the use of a population model to serve as a guide in determining the minimum levels of productivity required to support the historical catch data. Such methods were originally developed by Kimura and Tagart (1982); Kimura et al. (1984).

- Develop a formal statistical setting when including biomass data.

  Response: The original Catch-MSY model was based on a Bernoulli (accept/reject) and not really based on any sort of sampling theory and probabilities. The new Integrated Catch-MSY model maintains this same accept/reject criterion for cases where only catch data exist. There is also the option of including more formal statistical criterion where trends in biomass, or absolute estimates are evaluated assuming log-normal error distributions. There is also a likelihood criterion for changes in mean size. Lastly, there is an option for using a joint prior distribution with a user specified variance covariance matrix for a multivariate normal distribution.

R-package

An R-package (R Core Team, 2015) called catchMSY was developed for this project which contains all of the necessary routines for implementing the Integrated Catch-MSY model. There are two major differences between the tools in the catchMSY package and the original model devleoped by Martell and Froese (2012). First, the underlying population model is an age-structured model. Second, the population model is parameterized using key policy variables: (1) the Maximum Sustainable Yield, and (2) the instantaneous fishing mortality rate that achieves MSY (see Martell et al., 2008, for full details).

The R-package can be installed directly from R using the devtools package. At the R prompt use:

```
> devtools::install_github(smartell/catchMSY)
```

Model Documentation

The R-package contains four major subroutines for the population model, each of which are documented within the package itself. The following documentation will focus on these four major routines: (1) age schedule information, (2) deriving unfished biomass and steepness from MSY and F_{MSY}, (3) the age-structured population model, and lastly (4) the objective function for maximum likelihood estimation and Importance Sampling. The catchMSY package requires the creation of a stock ID object which contains all the necessary life-history information, fishery specs, and time series data on historical catch and if available abundance information and changes in catch composition.

Age-schedule information

Length-at-age is assumed to follow the von Bertallanffy growth function (von Bertalanffy, 1938):

$$
\begin{aligned}
l_a &= l_\infty(1 - \exp(-k * (a - t_0))) && (1)\\
w_a &= a(l_a)^b && (2)\\
m_a &= \left[1 + \exp(-(a - a_{50\%})/\sigma_{a_{50\%}})\right]^{-1} && (3)\\
f_a &= w_a m_a && (4)\\
v_a &= \left[1.0 + \exp\left(-\ln(19)\left\{\frac{a - a_{0.5}}{a_{0.95} - a_{0.5}}\right\}\right)\right]^{-1} && (5)\\
\iota_a &= \begin{cases} 1 & a = 1 \\ \iota_{a-1}\exp(-M_a) & 1 > a \geq A \\ \dfrac{\iota_a}{1 - \exp(-M_a)} & a = A \end{cases} && (6)
\end{aligned}
$$

Given length-at-age (1), weight-at-age is given by the allometric relationship in (2) with input parameters a and b. Maturity-at-age is assumed to follow a logistic curve with the parameters defining the age at 50% maturity ($a_{50\%}$) and the standard deviation in age at 50% maturity ($\sigma_{a_{50\%}}$). Fecundity-at-age is assumed to be proportional to mature weight-at-age (4). Vulnerability to the fishing gear is assumed to be age-specific using an asymptotic logisitc curve parameterized using ages that are 50% and 95% vulnerable (5). Note that a length-based selectivity function may also be desirable, as it is often easier to specify the length rather than the age of fish that are captured in the fishery.

Survivorship to a given age is based on the natural mortality rate M_a and is a recursive function (6). Note that the model assumes the oldest age-class is a plus group and contains individuals ages-A and older.

Deriving unfished biomass and steepness

Deriving B_0 and h from MSY and F_{MSY} is the inverse process of the $(B_0, h) \rightarrow$(MSY,F_{MSY}) transformation. Starting with the catch equation, first take the derivative of the catch equation and set it to $\frac{\partial Y}{\partial F} = 0$. Next solve the equation for F which corresponds to F_{MSY}. Note that the Baranov catch equation is a transcendental equation with respect to F and there is no analytical solution. However, the inverse problem does have an analytical solution for steepness and the following describes the derivation of B_0 and h given estimates of MSY and F_{MSY}.

The derivation of B_0 and h given an initial guess for MSY ($\dot{C}$) and F_{MSY} ($\dot{f}$) starts with calculating the age-specific mortality rates and survivorship under fished conditions ($\hat{\imath}$).

$$z_a = M_a + \dot{f} v_a \tag{7}$$

$$s_a = \exp(-z_a) \tag{8}$$

$$o_a = (1 - s_a) \tag{9}$$

$$q_a = v_a o_a / z_a \tag{10}$$

$$\hat{\imath}_a = \begin{cases} 1 & a = 1 \\ \hat{\imath}_{a-1} \exp(-z_a) & 1 > a \geq A \\ \dfrac{\hat{\imath}_a}{o_a} & a = A \end{cases} \tag{11}$$

$$\phi_q = \sum_a \hat{\imath}_a w_a q_a \tag{12}$$

Note that expressions (7)-(10) are intermediate vectors that simplify the algebra in the following equations. These vectors are the age-specific total mortality rate (7), age-specific discrete survival rate (8), age-specific discrete mortality rate (9), and the age-specific fraction that are vulnerable to harvest (10). The survivorship to a given age under fished conditions is denoted by (11), and the per recruit yield for each unit of fishing mortality is given by (12). Thus, the yield per recruit at MSY is given by YPR $= \dot{f}\phi_q = F_{\mathrm{MSY}}\phi_q$.

Given YPR, the equilibrium yield at $F_{\mathrm{MSY}}$is given by $\dot{C} = \dot{f}\dot{R}\phi_q$. So the next step is to derive the equilibrium recruitment at MSY. This is done by solving the previous equation for $R_{\mathrm{MSY}} = \dot{C}/(\dot{f}\phi_q)$. Note that the '$\dot{}$' notation denotes variables at MSY; example, $\dot{R}$ denotes the recruitment at MSY levels of spawning biomass. It is also worth mentioning again, that $\dot{f}$ and $\dot{C}$ are initial guesses at what the true underlying MSY values are, and these are akin to the $r - k$ pairs in the original Catch-MSY method outlined in Martell and Froese (2012).

The implied steepness of the stock recruitment relationship is determined by $\dot{f}$, along with the age-schedule information that determines the relative fecundity required to sustain the yield per recruit. The details of this derivation are fully documented in Martell et al. (2008) which result in the following expression for recruitment compensation (κ) for the Beverton-Holt stock recruitment model:

$$\kappa = \phi_e/\phi_f - \frac{\dot{f}\phi_q \dfrac{\phi_e}{\phi_f^2}\dfrac{\partial \phi_f}{\partial \dot{f}}}{\phi_q + \dot{f}\dfrac{\partial \phi_q}{\partial \dot{f}}} \tag{13}$$

$$h = \frac{\kappa}{\kappa + 4} \tag{14}$$

$$\kappa = \frac{4h}{1 - h} \tag{15}$$

The relationship between recruitment compensation κ and steepness as defined by Mace and Doonan (1988) is given by (14), and the reverse transformation by (15). In (13), ϕ_e and ϕ_f represent the spawning biomass per recruit in unfished and fished conditions, respectively.

The rate of change in spawning biomass per recruit with respect to changes in $\dot{f}$ is denoted by $\dfrac{\partial \phi_f}{\partial \dot{f}}$ and the rate of change in the yield per recruit with respect to $\dot{f}$ is given by $\dfrac{\partial \phi_q}{\partial \dot{f}}$. The following partial derivatives are used to calculate the change in spawning stock biomass with respect to $\dot{f}$ and the change in yield per recruit with respect to $\dot{f}s$:

$$\frac{\partial \hat{\iota}_a}{\partial \dot{f}} = \begin{cases} 0, & a = 1 \\ \dfrac{\partial \hat{\iota}_{a-1}}{\partial \dot{f}} - \iota_{a-1} s_{a-1} v_{a-1} & 1 < a \leq A \\ \dfrac{\partial \hat{\iota}_{a-1}}{\partial \dot{f}} - \dfrac{\iota_{a-1} s_{a-1} v_{a-1} s_a}{o_a^2} & a = A \end{cases} \tag{16}$$

$$\frac{\partial \phi_f}{\partial \dot{f}} = \sum_a f_a \frac{\partial \hat{\iota}_a}{\partial \dot{f}} \tag{17}$$

$$\frac{\partial \phi_q}{\partial \dot{f}} = \sum_a w_a q_a \frac{\partial \hat{\iota}_a}{\partial \dot{f}} + \hat{\iota_a} \frac{w_a v_a^2}{z_a} \left(s_a - \frac{o_a}{z_a} \right) \tag{18}$$

Recruitment is assumed to follow a Beverton-Holt stock recruitment model:

$$R = \frac{s_o B}{1 + \beta B} \tag{19}$$

where B is the spawning stock biomass, s_o is the maximum juvenile survival rate from egg to age-1, s_o/β is the asymptote of the function at infinite B, and R is defined as the number of age-1 recruits. This function can be re-parametrized using the spawning biomass per recruit incidence functions (ϕ_e, ϕ_f) and the recruitment compensation ratio κ as

$$\dot{R} = R_0 \frac{\kappa - \phi_e/\phi_f}{\kappa - 1} \tag{20}$$

and solving for the unfished recruitment R_0

$$R_0 = \dot{R} \frac{\kappa - 1}{\kappa - \phi_e/\phi_f} \tag{21}$$

$$B_0 = R_0 \phi_e \tag{22}$$

Note that (20) implies a constraint where the recruitment compensation ratio must be greater than 1.0, otherwise the function results in negative recruitment or is undefined if $\kappa = 1.0$. In addition, there is a limit where $\phi_e/\phi_f = \kappa$ that corresponds to the maximum fishing mortality rate that would lead to extinction (or $\mathrm{F}_{\mathrm{MAX}}$), which simply implies that $F_{\mathrm{MSY}} < \mathrm{F}_{\mathrm{MAX}}$. To determine the unfished spawning stock biomass, first calculate the unfished equilibrium recruitment using (21), then use (22) to compute B_0.

Age-structured population model

Assuming the model starts at unfished conditions, the initial numbers-at-age are initialized using the survivorship under unfished conditions (23).

Initial states

$$N_{t,a} = R_0 \iota_a, \qquad t = 1, \forall a \tag{23}$$

$$s_o = \kappa/\phi_e \tag{24}$$

$$\beta = (\kappa - 1)/B_0 \tag{25}$$

The stock recruitment parameters (24,25) are initialized using the recruitment compensation ratio (κ) and the unfished spawning biomass (B_0).

The spawning biomass in any given year is the sum of products between the numbers-at-age and the fecundity-at-age (26). The total age-specific mortality rate is given by (27), and the instantaneous fishing mortality rate conditioned on the observed catch (see next sub-section).

$$\text{Dynamic states (t > 1)}$$

$$B_t = \sum_a N_{t,a} f_a \tag{26}$$

$$z_{t,a} = M_a + F_t v_a \tag{27}$$

$$N_{t,a} = \begin{cases} \dfrac{s_o B_{t-1}}{1 + \beta B_{t-1}} & a = 1 \\ N_{t-1,a-1} \exp(-z_{t-1,a-1}) & 1 < a < A \\ N_{t-1,a-1} + N_{t-1,a} \exp(-z_{t-1,a}) & a = A \end{cases} \tag{28}$$

$$C_t = \sum_a \frac{N_{t,a} w_a v_a F_t (1 - \exp(-z_{t,a}))}{z_{t,a}} \tag{29}$$

Lastly, the age-structured model assumes that both natural mortality and fishing mortality occur simultaneously. The catch equation (29) is based on the fraction of total age-specific mortality that is associated with fishing mortality (F_t).

Conditioning the model on catch

To condition the age-structured model on the observed catch (which is assumed to be in units of weight), an iterative approach is used to determine the instantaneous fishing mortality rate F_t, because (29) does not have a closed form solution for F_t.

An initial value for F_t is based on an approximation to the to the analytical solution using:

$$F_t = \frac{C_t}{\sum_a N_{t,a} \exp(-0.5 M_a) w_a v_a} \tag{30}$$

Note that (30) is also known as Popes' approximation in the fisheries literature and is a fairly accurate approximation upto values of $F_t = 0.3$. To obtain the exact solution, Newtons' root finding method (31) is used to iteratively solve for F_t given the observed catch $\hat{C}_t$:

$$F_t^{(i+1)} = F_t^{(i)} - \frac{(C_t - \hat{C}_t)}{\frac{\partial C_t}{\partial F_t}} \tag{31}$$

$$\frac{\partial C_t}{\partial F_t} = \sum_a \left[\frac{v_a w_a o_{t,a} N_{t,a}}{z_{t,a}} - \frac{F_t v_a^2 w_a o_{t,a} N_{t,a}}{(z_{t,a})^2} + \frac{F_t v_a^2 w_a N_{t,a} \exp(-z_{t,a})}{z_{t,a}} \right] \tag{32}$$

$$\text{where} \quad o_{t,a} = (1 - \exp(-z_{t,a}))$$

The derivative of the catch equation with respect to the instantaneous fishing mortality rate is given by (32). This algorithm converges in roughly 3-7 iterations depending on the initial value of F_t relative to the converged value.

Observation Models

The catch-MSY package currently accommodates 4 different types of data in which to make any sort of statistical inference about the likelihood of the observed data given the model parameters. These data are: (1) relative abundance data, (2) absolute abundance data, (3) change in mean length data, (4) mean weight of the catch.

Relative abundance data

In the case of fitting the model to trends in abundance, it is assumed that observation errors in abundance trends are lognormal. The predicted abundance index is assumed to be proportional to population biomass:

$$\hat{I}_t = qB_t e^{\epsilon_t} \tag{33}$$

$$\ln(q) = \ln(\hat{I}_t) - \ln(B_t) - \frac{1}{n}\sum_{t\in\hat{I}_t} \ln(\hat{I}_t) - \ln(B_t) \tag{34}$$

$$\epsilon_t = \ln(\hat{I}_t) - \ln(q) - \ln(B_t) \tag{35}$$

$$\ell_I = n[0.5\ln(2\pi) + \ln(\sigma_{\epsilon_t})] + \sum_{t\in I_t} \frac{\epsilon_t^2}{2\sigma_{\epsilon_t}^2} \tag{36}$$

The slope of the relationship between survey estimates of abundance and biomass is given by the parameter q, and the conditional maximum likelihood estimate of q is given by (34). The vector of residuals is defined by (35). Note that the number of survey observations must greater than or equal to at least 2 observations, and the model assumes the same q for each observation. The negative loglikelihood for trend data is given by (36).

Absolute abundance data

The previous biomass augmented catch-MSY approach (Sabater and Kleiber, 2014) specified a range of acceptable biomass estimates and would only accept parameter combinations that resulted in biomass trajectories that would fall within a biomass interval in the year in which the survey was conducted. In this application, a more formal statistical approach is adopted that allows for non-linear parameter estimation and using gradient methods to obtain maximum likelihood estimates of model parameters, a variance covariance matrix, and to conduct Importance Sampling for constructing joint posterior distributions.

In the case of fitting the model to an estimate of absolute abundance in a given year, or multiple years if such data are available, it is also assumed that measurement errors are lognormal. The likelihood is exactly the same as (36); however, the residuals do not include the latent variable q as it is assumed that the observation is an estimate of absolute abundance.

$$\varepsilon_t = \ln(\hat{B}_t) - \ln(B_t) \tag{37}$$

$$\ell_{\hat{B}} = n[0.5\ln(2\pi) + \ln(\sigma_{\varepsilon_t})] + \sum_{t \in \hat{B}_t} \frac{\varepsilon_t^2}{2\sigma_{\varepsilon_t}^2} \tag{38}$$

where ε_t is the log residual difference between the observed abundance index ($\hat{B}_t$) and the predicted abundance index (B_t).

Changes in mean length

For cases in which there are observations on the changes in mean length from samples collected at two distinct time periods, the catch-MSY model must convert numbers-at-age, to numbers-at-length. This is accomplished using the inverse of a age-length key (if empirical length-age data are available), or using a growth model to predict mean length-at-age and the standard deviation in length-at-age. To derive an age-length key from the growth model, the catch-MSY package assumes that length-at-age has a normal distribution. The probability that a fish of a given age a is in the length-interval l is given by:

$$P(l|a) = \int_{l-\Delta_l}^{l+\Delta_l} \frac{1}{\sqrt{2\pi}\sigma_a} \exp\left(-\frac{(l-l_a)^2}{2\sigma_a^2}\right) dl \tag{39}$$

where Δ_l is half the interval width, l_a is the mean length-at-age and σ_a is the standard deviation in length-at-age. Ideally, length-age data would be available to estimate the mean length-at-age and the standard deviation in length-at-age using one of many alternative growth models. Absent these data, a reasonable approximation for the standard deviation in length-at-age is to assume a CV$\approx$ 0.1, where $\sigma_a = CV \cdot l_a$ and use the von Bertalanffy growth model to approximate the mean length-at-age.

The predicted mean length of the catch, or mean length of an underwater visual census, involves 2 distinct probabilities: (1) the probability of sampling an individual of a given length l, and (2) the probability of an individual of length l existing in the population. The former refers to what is widely known in the fisheries literature as selectivity, and the latter refers to the relative abundance of different length/age classes in the population. It's generally safe to assume that very small individuals are less conspicuous than large individuals and one could use a logistic function to represent the probability of sampling. It may also be appropriate to consider other alternative functions (i.e., dome-shaped functions) that would represent the idea of larger fishing being more difficult to detect if the sampling frame is restricted relative to the distribution of size/age classes. For example, if the size-composition information is obtained from a depth restricted underwater visual census (e.g. Richards et al., 2011), but the species is also known to have ontogenic movement to deeper waters (common in the Serranidae family). In this case a dome-shaped function might be more appropriate, but an asymptotic function would result in a more conservative estimate of MSY-based reference points.

Given the length-age key in (39) the predicted length-vector is given by the following joint probability distribution:

$$\vec{l_t} = (\vec{p} \cdot \vec{v}) \cdot P(l|a) \tag{40}$$

where

$$p_{t,a} = \frac{N_{t,a}}{\sum_a N_{t,a}} \tag{41}$$

$$\bar{L}_t = \sum_l \left[\frac{(l_l)_t}{\sum_l (\vec{p} \cdot \vec{v}) \cdot P(l|a)} \right] \tag{42}$$

$$\eta_t = \hat{L}_t - \bar{L}_t \tag{43}$$

$$\ell_{\hat{L}} = n[0.5 \ln(2\pi) + \ln(\sigma_\eta)] + \sum_{t \in \hat{L}_t} \frac{\eta_t^2}{2\sigma_\eta^2} \tag{44}$$

where p_a is the proportion of individuals of age a in the age-structured population dynamics mode, v_a is the probability of sampling an individual of age a (i.e., selectivity in eq. 5), and $P(l|a)$ is the age-length key. The residual difference between the observed mean length $\hat{L}$ and the predicted mean length $\bar{L}$ is given by (43), and the negative log-likelihood by (44).

Changes in mean weight

The predicted average weight of the catch in year t is given by (45), which is the sum of products between thee proportions-at-age, the weight-at-age (w_a), and the selectivity-at-age (v_a).

$$\bar{w}_t = \sum_a p_{t,a} w_a v_a \tag{45}$$

$$\nu_t = \hat{w}_t - \bar{w}_t \tag{46}$$

$$\ell_{\hat{w}} = n[0.5 \ln(2\pi) + \ln(\sigma_\nu)] + \sum_{t \in \hat{w}_t} \frac{\nu_t^2}{2\sigma_\nu^2} \tag{47}$$

The residual difference between the observed ($\hat{w}_t$) and predicted mean weight ($\bar{w}_t$) is given by (46), and the negative loglikelihood by (47).

Objective Function

The objective function is made up of two components: (1) statistical components that describe the goodness of fit to observed data, and (2) non-statistical components that describe the prior beliefs about underlying model parameters or predicted state variables that are derived from the same model parameters. The statistical components are listed in Table 1. In the catch-MSY method developed by (Martell and Froese, 2012), there were no statistical components in the methods, and the model outputs were strictly based on prior beliefs about resilience, depletion in the terminal year, and the catch history which is used to condition the model. The new integrated catch-MSY model maintains this "data free" model fitting based only on prior beliefs. But there is also the option to incorporate biomass and composition data.

Statistical criterion

For the statistical component of the objective function, the sum of the negative log-likelihoods is used to inform the search gradient, or importance weight:

$$\ell(\hat{I}, \hat{B}, \hat{L}, \hat{w}|\Theta) = \ell_{\hat{I}} + \ell_{\hat{B}} + \ell_{\hat{L}} + \ell_{\hat{w}} \tag{48}$$

For the non-statistical component of the objective function, the catch-MSY method makes use of a number of criterion to accept or reject a particular combination of $\Theta = (m, \dot{f}, \dot{C})$. These criterion are detailed in the next section.

Table 1: List of statistical objective function components in the catch-MSY package.

Type of Data	Symbol	Error distribution	Required	Implemented
Catch data	$\hat{C}$	NONE	YES	✓
Relative abundance	$\hat{I}$	lognormal	NO	✓
Absolute abundance	$\hat{B}$	lognormal	NO	✓
Change in mean length	$\hat{L}$	normal	NO	
Mean weight in catch	$\hat{W}$	normal	NO	

Non-statistical criterion

A key feature that makes the catch-MSY approach attractive for data poor assessments is that it is extremely simple (Rosenberg et al., 2014). Martell and Froese (2012) warn that the method, in the absence of axillary data on relative abundance is only informative about about the lower bounds of stock productivity. That is, any parameter combination that leads to population extinction prior to the terminus of the time series is very unlikely. What the authors also demonstrate, is that additional constraints (or bounds) on the current status of the stock (depletion level relative to the unfished state) are very informative about the overall scale of MSY provided that the catch time series does induce depletion. Hilborn and Walters (1992) demonstrate that despite the strong negative correlations in scale (k) and productivity parameter (r) that arises in stocks with a 'one-way trip' dataset, the assessment model is very informative about MSY because the confounding is reduced in the MSY=$\frac{rk}{4}$. The catch-MSY method exploits this particular tautology to provide reasonable bounds for MSY.

There are five non-statistical components that are considered in the integrated catch-MSY model: (1) the parameter bounds, (2) extinction prior to the terminal year, (3) infinite biomass (4) upper bound on fishing mortality rates, and (5) the range of spawning depletion. The exit codes for each model run are summarized in Table 2.

Setting the upper and lower bounds for the three leading model parameters (M, F_{MSY}, and MSY) will influence the overall distribution for MSY. For example if there are no constraints on the upper bound of the depletion model, then the right-hand tail of the resulting MSY distribution will correspond to the upper bound specified for MSY. In other words, there is no information in the catch data alone that will inform the overall scale of the stock. However, the catch data do provide information on the lower bound of the MSY distribution, otherwise the population would go extinct in the model.

Another non-statistical criterion is parameter combinations that lead to population extinction in prior to the terminal year. Under the assumption of deterministic production (i.e., no process errors in the form of recruitment deviations), parameter combinations that result in the stock being fished to extinction would have to be rejected because there are continued removals in the fishery.

Similarly to extinction is infinite biomass. It is also possible that randomly draws from a range of possible parameter combinations can result in an infinite population biomass due to numerical precision of the computers, or taking the logarithm of 0. In this case the parameter combination is not accepted.

The forth constraint is an acceptable range of depletion levels (default values are 0.0 – 1.0), where 0.0 would indicate that the fishery just removed that last individual from the population, and 1.0 implies a pristine or virgin stock. In cases where there is no abundance information, or changes in size-composition over time, it is often desirable to restrict the upper bound of this range in order to provide an upper bound for MSY (n.b., the previous comment about parameter bounds). If there are data on trends in abundance, or absolute abundance, than it may not be necessary to restrict the depletion range as the catch data and abundance data together are likely to help inform estimates of depletion Walters et al. (2006).

The fifth constraint pertains to estimates of fishing mortality rates. It is possible that the observed catch is greater than the exploitable biomass (i.e., the numerator in eq. 30 is $>$ the denominator). This is entirely possible as the fishing mortality rate is represented as

Table 2: Default limits for accepting a random parameter combination and model exit codes for each sample.

Constraint	**Default Limit**	**User specified**	**Exit Code**
Pass	-	-	0
Biomass extinction	0	-	1
Infinite biomass	∞	-	2
Depletion lower bound	0.0	✓	3
Depletion upper bound	1.0	✓	4
Maximum fishing rate	5.0	-	5

an instantaneous rate and the annual discrete rate is approximately defined by the following relationship: $U_t = 1 - \exp(-F_t)$. For example, and instantaneous fishing mortality rate of $F_t = 1.2$ is approximately equal to an annual exploitation rate of 0.70. In particular cases where the biomass trajectory of the stock is low, then the corresponding estimates of F are very large. The user must specify an upper range for the maximum F that is acceptable, and the criterion is to accept parameter combinations that result in $F's$ that fall below the maximum. The default upper limit for max F is 5.0, which corresponds to an annual exploitation rate of 0.99.

Non-linear search

One of the research recommendations by the CIE review on the methods developed by Sabater and Kleiber (2014) was to develop a more formal statistical approach to the catch-MSY method. In consideration of this recommendation, the new integrated catch-MSY package also incorporates a non-linear search routine to obtain the MLE estimates for the vector Θ. The Hessian matrix is also used to construct approximate estimates of the standard errors for each of the model parameters and the variance covariance matrix. This is extremely useful for constructing the joint posterior distribution using Importance Sampling (see next section).

Importance Sampling

For cases in which there is a statistical component to the objective function, the catch-MSY package also has the option to sample the joint posterior density using Importance Sampling. This was another recommendation by the CIE review.

Parameter samples are drawn from a multivariate normal distribution where the user must specify the mean and variance-covariance matrix. These values could be obtained from the non-linear search routine, and the variance-covariance matrix could be inflated to ensure sufficient sampling occurs at the margin of the distribution. It's also possible for the user to directly specify the parameter correlation matrix and standard deviations of the prior distributions for each of the parameters. This can then be used to construct the appropriate variance-covariance matrix for use in Importance Sampling. For the importance function, a multivariate normal distribution is used along with the statistical components in (48) to determine the probability of obtaining the data for a given parameter vector.

Discussion

The catch-MSY approach (Martell and Froese, 2012) in its simplest form is an alternative to a Tier 5 approach to specifying annual catch limits (ACLs). The Tier 5 approach uses the average catch (over some arbitrary time period) to set the ABC and OFLs. There is certainly merit in using this method, or other similar algorithmic approaches (e.g., Kimura and Tagart, 1982; Dick and MacCall, 2011; MacCall, 2009) to set catch levels when only catch information exists. But these methods are not meant to provide estimates of stock productivity or stock status in the absence of abundance or composition data.

A great deal of algebra and calculus is involved in the derivation of B_0 and steepness parameters given the management related variables (MSY and F_{MSY}) and some assumptions about fisheries selectivity. The objective of adopting this management oriented approach is two fold: (1) to create a simpler framework where the catch data alone can be used to specify reasonable priors for the age-structured model, and (2) to allow for the potential of using life-history invariants (e.g., Beddington and Kirkwood, 2005) to directly specify priors for $F_{\mathrm{MSY}}$and natural mortality.

Under the first objective, using a prior range, or prior density, for M, F_{MSY}, and MSY, implies a prior density for the derived quantities B_0 and steepness. The catch data are largely informative about the global population scaling and to some degree estimates of MSY scale with B_0 (Hilborn and Walters, 1992). The challenge with the conventional parametrization, when the only available data is a catch time series, is to come up with reasonable bounds for B_0. This was noted in the CIE review for specifying the analogous scaling parameter (carrying capacity, K). This alternative parameterization allows the catch data themselves to be used to specify a reasonable range for MSY directly. In fact, this could be encoded where, for example, the 5th and 95th percentiles of the catch distribution be used to specify the range of values for MSY.

The second objective is to allow for the use of life-history invariants to specifying life-history variables. There is a large literature built upon meta-analytic work that bears the relationship between 3 common life-history invariants (or sometimes called Beverton and Holt life history invariants Jensen, 1996). These three relations are: $C_1 = MA_m$, $C_2 = M/K$, and $C_3 = L_m/L_\infty$. Where C_1, C_2 and C_3 are constants, M is the instantaneous natural mortality rate, A_m is the age-at-maturity, K is the von Bertalanffy growth coefficient, L_m is the length at maturity and L_∞ is the asymptotic length. Jensen (1996) estimated the value of the three constants via fecundity maximization that optimizes the trade-off between survival and fecundity. The values obtained using this method fall within the range of numerous empirical studies. Given these constants, and well know life-history invariants (Charnov, 1993), nearly all of the information to approximate age-at-maturity, or length-at-maturity, natural mortality, can be readily obtained using an estimate of the von Bertalanffy growth coefficient and estimates of the asymptotic length (see Walters and Martell, 2002, for full details). At first it may seem a bit daunting coming up with all the necessary information to set up an age-structured model for a data poor species, but using relationships such as the life-history invariants, is really not all that difficult nor that different than a subjective choice on resilience (as in Martell and Froese, 2012). If fact, it is more grounded in the meta-analytic work that has been done in this field and is much simpler than the methods developed by McAllister et al. (2001).

One of the primary deficiencies of the catch-MSY method that Sabater and Kleiber (2014) attempted to address is to augment the method to address estimates of stock productivity and stock status such that many of these stocks could be elevated from a catch only (Tier 5) method for setting ACLs to a model-based approach (or Tier 3). This was accomplished by incorporating a biomass estimate into the non-statistical criterion for accepting $r - k$ pairs. In addition to specifying an acceptable range of depletion values, the selected parameter combinations would also have to fit through the approximate 95% confidence interval of absolute biomass. In theory, this should vastly improve the estimate of MSY, and where it could potentially be misleading is if the estimate of biomass is biased. For example, if the biomass-interval estimate is biased by 50%, then model estimates of biomass would be biased by approximately the same amount, and it would be logical to conclude that MSY is also biased low. One way to address this unmeasurable bias is to attempt to integrate other additional sources of information into the same framework.

Another way to tackle the overall scaling problem in these data poor models is to incorporate additional information about the total mortality rate Z. Absent process errors in the form of recruitment variation, changes in the mean size of fish sampled in the population over time can be informative about the relative changes in Z provided there are unbiased estimates of growth parameters and selectivity schedules. Constant recruitment is one of the many underlying assumptions in the simple catch-curve methods that are used to infer changes in Z (Hilborn and Walters, 1992; Walters and Martell, 2004). One of the minor differences in using a full age-structure population model with a stock-recruitment relationship is that the estimates of Z are consistent with the stock-depletion and the resulting recruitment (i.e., recruitment is a deterministic function of spawning biomass, and is not not constant over the entire range of depletion levels). Another advantage of incorporating even an single year of mean size data is that the model can be fitted to this observation based on the predicted mean size. This provides additional information that can help inform selectivity (or at a minimum age-at-first capture).

The catch-MSY packaged developed herein is not limited to just biomass and composition information as describe in the methods section. It is possible to expand the package options to include the wide-variety of data that are now available in many of the more sophisticated stock assessment programs. For example, the catch-MSY package already predicts a size-frequency distribution to estimate the means size of the catch. There is no reason not to directly fit the model to size-composition directly as it may provide better information to resolve selectivity.

Lastly, one of the major assumptions herein, and in most stock assessment models, is the notion of a stationary production relationship over the time period in question. For example, a vast majority of the assessments assume many of the life-history parameters are time-invariant. Obviously changes in size-at-age associated with factors other than fishing would have profound impacts on the estimates of total mortality rate. In fact, these dynamic variables that are assumed to be static are now creeping into the more sophisticated, data & information rich assessments (e.g., Methot and Wetzel, 2013). Changes in the stock-recruitment relationship have been the subject of much research in ecosystem models (Christensen and Walters, 2004)

The original catch-MSY method was never intended to be used as a stock assessment framework for estimating stock status. It was developed under the basic idea that catch information could inform the parameter domain which defined the lower bounds of MSY. It was the addition of prior information about the current stock status (depletion level) that was necessary to put any sort of upper bound on MSY. The work done herein, and for the catch-MSY package, has maintained the same simple elegance of the original work where only 2-3 unknown parameters are required for a much more sophisticated population dynamics model that is capable of integrated a wider array of information on size-composition, biomass, trends in biomass, etc. and can also address growth overfishing and the potential to manage the effect through size-limits or other incentives.

References

Beddington, J. and Kirkwood, G. (2005). The estimation of potential yield and stock status using life–history parameters. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1453):163.

Beverton, R. J. and Holt, S. J. (1993). *On the dynamics of exploited fish populations*, volume 11. Springer.

Charnov, E. L. (1993). *Life history invariants: some explorations of symmetry in evolutionary ecology*, volume 6. Oxford University Press, USA.

Christensen, V. and Walters, C. J. (2004). Ecopath with ecosim: methods, capabilities and limitations. *Ecological Modelling*, 172:109–139.

Dick, E. and MacCall, A. D. (2011). Depletion-based stock reduction analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. *Fisheries Research*, 110(2):331–341.

Forrest, R. E., Martell, S. J. D., Melnychuk, M. C., and Walters, C. J. (2008). An age-structured model with leading management parameters, incorporatig age-specific selectivity and maturity. *Can. J. Fish. Aquat. Sci.*, 65:286–296.

Hilborn, R. and Walters, C. J. (1992). *Quantitative fisheries stock assessment: choice, dynamics and uncertainty*, volume 2. Chapman and Hall (New York).

Jensen, A. (1996). Beverton and holt life history invariants result from optimal trade-off of reproduction and survival. *Can. J. Fish. Aquat. Sci.*, 53:820–822.

Kimura, D., Balsiger, J., and Ito, D. (1984). Generalized stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 41(9):1325–1333.

Kimura, D. and Tagart, J. (1982). Stock reduction analysis, another solution to the catch equations. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(11):1467–1472.

MacCall, A. (2009). Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. *ICES Journal of Marine Science: Journal du Conseil*, 66(10):2267.

Mace, P. M. and Doonan, I. (1988). *A generalised bioeconomic simulation model for fish population dynamics*. MAFFish, NZ Ministry of Agriculture and Fisheries.

Martell, S. and Froese, R. (2012). A simple method for estimating msy from catch and resilience. *Fish and Fisheries*.

Martell, S. J. D., Pine, W. E., and Walters, C. J. (2008). Parameterizing age-structured models from a fisheries management perspective. *Can. J. Fish. Aquat. Sci.*, 65:1586–1600.

McAllister, M. K., Pikitch, E. K., and Babcock, E. A. (2001). Using demographic methods to construct bayesian priors for the intrinsic rate of increase in the schaefer model and implications for stock rebuilding. *Can. J. Fish. Aquat. Sci.*, 58:1871–1890.

Methot, R. D. and Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Richards, B. L., Williams, I. D., Nadon, M. O., and Zgliczynski, B. J. (2011). A towed-diver survey method for mesoscale fishery-independent assessment of large-bodied reef fishes. *Bulletin of Marine Science*, 87(1):55–74.

Rosenberg, A. A., Fogarty, M. J., Cooper, A. B., Dickey-Collas, M., Fulton, E. A., Gutiérrez, N., Hyde, K., Kleisner, K., Kristiansen, T., Longo, C., et al. (2014). *Developing new approaches to global stock status assessment and fishery production potential of the seas*. Food and Agriculture Organization of the United Nations.

Sabater, M. and Kleiber, P. (2014). Improving specifications of acceptable biological catches of data-poor reef fish stocks using a biomass-augmented catch-msy approach. 60p. Technical report, Western Pacific Regional Fishery Management Council, Honolulu, HI 96813 USA.

Schnute, J. T. and Richards, L. J. (1998). Analytical models for fishery reference points. *Can. J. Fish. Aquat. Sci.*, 55:515–528.

von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws ii). *Human biology*, 10(2):181–213.

Walters, C. and Martell, S. (2002). Stock assessment needs for sustainable fisheries management. *Bulletin of Marine Science*, 70(2):629–638.

Walters, C. and Martell, S. (2004). *Fisheries Ecology and Management*. Princeton University Press.

Walters, C., Martell, S., and Korman, J. (2006). A stochastic approach to stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(10):212–223.