2. Fishing mortality calculation

Several options are available for undertaking the calculation of fishing mortality and, within each, there are the special cases for either missing catch or effort for a given fishing incident. The options are controlled by age_flags(92):

- 0 = catch errors method (default); is effort-based, i.e. F = q * Effort * Selectivity
- 1 = un-implemented
- 2 = catch calculated by Baranov catch equation
- 3 = catch calculated by linearized catch equation, i.e. catch-conditioned as in SS3 – Pope's approximation

Other influential flags referred to during the catch estimation are:

- age_flags(34) – activate estimation of effort_dev_coff
- age_flags(94) – specifies the starting population strategy, where = 2 uses average Z
- age_flags(115) – specifies the catch-conditioned method as in SS3
- age_flags(116) – specifies the maximum allowable fishing mortality level
- age_flags(143) – is assigned within control phases (set_ctrl.cpp) that enables the activation of estimating effort_dev_coffs
- age_flags(157) – undocumented, may be an option not yet implemented or a special case, if = 1, see rshort1.cpp where global_vars are set, and short2.cpp where `xcatch_equations_calc`() is called.
- age_flags(177) – undocumented, may be an option not yet implemented or a special case; in routine catch_equations_calc2() it implements a special case of initial population calculation; and totpop_coff is assigned to a different set_value() routine with different penalty.

The key point in the code where the different methods are employed is in `do_everything_calc()` in short2.cpp, where there are various calls to routines for the fishing mortality calculation and the switch-case block exists for those options where age_flags(92) != 0.

For the case age_flags(92) == 0, the routine is called: `catch_equations_calc`()

For all cases > 1, the routine is called: `new_catch_equations_calc_implicit_experiment_loop()`

Catch-errors method

The documentation of the catch-errors method is well documented in the User Guide in the Technical Appendix A.1.6 to A.1.8.

Catch-conditioned method

Essentially the catch-conditioned method is that employed in most contemporary fish population models, such that the fishing mortalities are solved for conditional upon the observed catches. The formulation used in this development assumed that the Baranov catch equation held and the

Newton-Raphson (N-R) procedure was used to solve for the fishing mortality levels among the various fisheries for which there were incidents within a time period.

There is a fundamental difference in the dynamic model calculations when using the catch-conditioned method compared to the catch-errors method. With the latter, all the estimated parameters for calculating fishing mortality (catchability, effort deviates etc.) are available at the beginning of the model calculation period **for all time periods**; and therefore, the numbers-at-age calculation is a linear sequence starting from the first model year. In contrast, for the catch-conditioned method, in the first model time period none of the fishing mortalities are available, but rather, they are calculated sequentially for each fishing incident with each time period, starting in the first year. The numbers-at-age calculation therefore includes both the dynamic age-matrix and the fishing mortality calculations in each time period.

Given this fundamental difference, there is no requirement for a total catch likelihood within the catch-conditioned method. Hence, in **newl2.cpp** in the routine `fit_totals()` there is the condition with respect to age_flags(92) when calling the total catch likelihood routine: `total_catch_or_weight_fit()`.

The standard form of the Baranov catch equation is employed, with the supplied observed catches for each fishery and the component incidents where available. The fishing mortality is found that satisfies: the predicted catch is equal to the observed catch, i.e. without error. The solution is found by the Newton-Raphson procedure with matrix parameterisation, and that is solved for via various penalty functions that are included within the auto-differentiation minimisation of the integrated model.

Flag settings

The flag settings that are new and specific to the catch-conditioned method are listed, along with some existing flags that are influential to the calculations.

Newton-Raphson:
parest_flags(373) = 1 - activates the catch-conditioned option with Baranov equation
age_flags(92) = 2 - specifies the catch-conditioned option with Baranov equation
age_flags(92) = 3 - specifies the catch-conditioned option with Pope's Approximation
age_flags(115) = 1 - specifies the catch-conditioned option with Pope's Approximation
age_flags(116) = specifies the maximum allowable fishing mortality level, Zmax_fish
parest_flags(382) = n - weight for Zmax_fish penalty, default=100 to achieve Zmax_flag=0
age_flags(189) = the fraction of Zmax_fish above which the penalty is calculated; fraction= age_flags(189)/100
age_flags(157) = not certain of its function, may be a special case that calls routine xcatch_equations_calc()

Initial equilibrium population:
parest_flags(393) - activates the estimation of and penalty for catch-conditioned parameters: kludged_equilib_coffs
parest_flags(375) = 1000 - penalty wt on kludged initial survival relationship : actual survival

parest_flags(374) = 7 - spline degree for kludged initial survival relationship
parest_flags(376) = n - no. of additional spline nodes to be added to the degrees for kludged initial survival relationship
parest_flags(379) = n – weight on the normal penalty of the higher level (>4) spline degrees for the kludged_equilib_coffs; default=10.
parest_flags(394) – a development/debug function. If set = 1 avoids using the kludged_equilibrium_coffs parameters and uses fixed initial survival = 0.8 for establishing the initial conditions.
age_flags(94) – 1 specifies N_{init} as a function of M; 2 – specifies N_{init} in exploited equilibrium relative to average total mortality over a defined period
age_flags(95) – specifies initial periods used for calculating average total mortality for deriving N_{init} in exploited equilibrium

Fml_effort_rltnshp:
parest_flags(377) = 1 – activates estimation of fm_level:effort fit relationship parameters and implements penalty; and generates reports
parest_flags(396) = 0 implements the fm_level:effort fit relationship regression using the independent parameters; or, = 1 implements the fm_level:effort fit relationship regression using the least-squares solution for the parameters
parest_flags(378) = activates G-S design matrix, adds regression penalty to integrated likelihood, and selects options for the fm_level:effort fit relationship: 1 - activates robust option for "messy" effort data; 2 – activates robust-normal mixture option; 3 – activates the student-t option
parest_flags(383) = n = weight for normal penalty on implicit_fm_level_regression_pars; default = 10.
fish_pars(31) - parameter for the robust-normal mixture option in fitting the fm_levels:effort relationship that seeks to get a dimensionless, well-scaled sigma because it is not normally distributed
fish_pars(32) – degrees of freedom parameters for the students-t option in fitting the fm_levels:effort relationship
fish_flags(81) 1 – activates estimation of fish_pars(31)
fish_flags(93) 1 – activates estimation of fish_pars(32)
parest_flags(362) – setting for an assumed fixed variance for robust-normal mixture option, where fish_flags(81) **must** be set = 0. Fixed variance = 1.0/parest_flags(362)
fish_flags(73) = 5 - degrees of polynomial for fm_level:effort fit relationship; a value of 0 estimates only the mean catchability (i.e. constant)
fish_flags(29) = n - takes into account grouping of catchability for the polynomial for fm_level:effort fit relationship and its associated parameters (fish_pars(31))
fish_flags(27) 1 - takes into account (seasonal catchability) for the polynomial for fm_level:effort fit relationship

Survey fishery CPUE likelihood:
fish_flags(92) = 1 specifies the fisheries having a "survey index" of relative abundance in respect of the catch and standardized effort; fitted in a specific likelihood term with penalty weight = ff92/1000.
fish_flags(i,66) = 1 activates the option for index-specific variance is to be included in likelihood calculation
fish_flags(i,99) = specifies the grouping to be used for fisheries assumed to have stationary catchability.

parest_flags(381) = 1 - activate debug_par

There is a sanity check for the survey fisheries grouping. So if one assumes stationary q for survey fisheries, then the data_fish_flags(1,fi) in the *.frq should all match with the settings in fish_flags(fi,99). This is to ensure that mixed frequency/biomass CPUE indices do not occur in the same group.

This checks:

data_fish_flags(1,fi) – for the survey fisheries are all identical (either all 0 or all 1)

fish_flags(fi,99) – are correctly sequentially numbered as grouping flags should be: e.g. 1 – 31, 32, 32, 32, … 32 (for grouped survey fisheries 32 to *n* fisheries)

Parameters

Fewer independent variables are required for the catch-conditioned method by comparison with the catch-errors method (by up to 73%), and they concern the regression of the estimated fishing mortality levels (fm_levels) and the observed effort, and the initial equilibrium population conditions. Besides fish_pars(31, 32), these parameters are new sections of the .par file version 1061 and higher.

fish_pars(31) - parameters for the robust option in fitting the fm_levels:effort relationship
fish_pars(32) - parameters (dof) for the student-t option in fitting the fm_levels:effort relationship

implicit_fm_level_regression_pars(fi,npars) - fm_levels:effort relationship regression parameters with npars(fi) being the degrees of the polynomial regression for each fishery, fi. Section of the .par file: # fsh.implicit_fm_level_regression_pars

kludged_equilib_coffs - spline nodes for the kludged equilibrium survival for the initial population. A matrix: regions-seasons. Section of the .par file: # kludged_equilib_coffs

kludged_equilib_level_coffs - the region-specific kludged levels. Section of the .par file: # kludged_equilib_level_coffs

Penalties

Various penalties for estimating the catch-conditioned parameters are added to the overall integrated model total likelihood. These are listed briefly with the code location provided, but a detailed explanation is given in the programmer's guide below:

- Actual **initial survival** (taken from the dynamic model calculation) versus the **kludged equilibrium survival** (as estimated from the parameters: kludged_equilib_coffs, kludged_equilib_coffs_levels): kludged_equil_surv.cpp:26
- **kludged_selmean_square** – in respect of the kludged_equilib_coffs parameters themselves
- penalty on the higher level (>4) spline degrees for the **kludged_equilib_coffs**, with parest_flags(379) specifying the penalty weight
- A maximum fishing mortality penalty **Zmax_fish**: nrcatch3.cpp:269
- Various penalties that depend upon the option used for the form of the **fm_level:effort** fit relationship specified by the switch-case options provided by parest_flags(378):
    - 0: Least-squares option (**DO NOT USE THIS**)
    - 1: Robust option using robust_normal_cauchy

- 2: Less robust option using robust_normal_mixture
    - 3: students-t option
- Normal penalty on **implicit_fm_level_regression_pars**, with parest_flags(383) specifying the penalty weight. Note callpen.cpp that `int lb=icol+2;` so as not to impact on the lower order coefficients.
- **Maximum fm_level** for other_mort calculation: nrcatch4.cpp:165; but fpen does not appear to integrated?

Initial conditions

Since the average initial total mortality that is used to calculate the initial exploited equilibrium population, i.e. the initial survival rate, is not known at the beginning of year 1, a method was developed to allow for the estimation of the survival rate to be used for calculating the equilibrium population within the overall integrated model fit. The standard approach is to specify a period at the beginning of the model from which the average total mortality will be used for calculating the initial exploited population at equilibrium. A spline formulation was used for the age-specific estimated survival rates with the nodes being the independent variables. The initial population at equilibrium is then calculated with these rates, and the subsequent population calculation is done for all time periods following. A penalty function ensures the survival rate estimates are closely consistent with that of the actual total mortalities for the specified period used for the equilibrium calculation. Provision is made for the case where the spline degree can be increased from that of a previous solution.

Regression between estimated fishing mortality levels and observed effort

The N-R solution for the fishing mortality levels applies in the cases where an observed catch is available. The exceptions are the cases where observed catch is unavailable. An orthogonal-polynomial regression model using a Gram-Schmidt (G-S) design matrix is formulated, and a robustified fit using a mixture distribution, to estimate fishery-specific relationships between the: **fishing mortality levels** (estimated from the N-R) for those incidents where catch is available, and the **observed effort**. This estimated relationship is denoted here as the **fm_level_rltnshp**. The degrees for the estimated polynomial for each fishery are specified, and may include seasonality in the catchabilities, that define the dimensions of the G-S, and the parameters are estimated within the integrated model fit. Then for those incidents having missing catch, the fishing mortality level prediction from the relationship using the observed effort, is applied to predict the catch. The calculation is also extended to cope with the case of grouped fisheries. The method entailed the following algorithm:

- Obtain the fishery and grouping specifications and construct the fm_level_rltnshp parameters
- Build the G-S design matrix for each fishery/grouping – store
- Enter the fitting procedure
- Obtain the first model iteration values for the fm_level_rltnshp parameters
- For incidents with observed catches, use the N-R to calculate the fishing mortality levels

- Using the fm_level_rltnshp parameters, the G-S design matrix, and observed effort, predict the fishing mortality levels for the incidents with missing catch
- Calculate the regression distribution function for the fm_level_rltnshp, and add to the integrated model fit
- Enter next model iteration
- Repeat iterations until convergence

In the implementation of the method within MULTIFAN-CL a diagnostic report for all incidents having missing catches is produced containing the predictions obtained from the fm_level_rltnshp. A penalty on the maximum age-specific fishing mortality (`Zmax_fish`) is applied that prevents the N-R estimates exceeding levels above which the population cannot sustain, and therefore causing minimisation instability. The weight of this penalty is assigned by a flag setting (see above).

The fm_level_rltnshp plays a fundamental role in estimating fishing mortality for fisheries with standardised indices of relative abundance. Presently the sigma (maximum likelihood estimate of the standard deviation) for the regression is calculated and reported for each fishery. In theory, the fm_level_rltnshp should be self-scaling, but how to achieve the correct relativity between the normal and the standardised fisheries remains to be developed. Some of the fishery groups may assume seasonality in catchability, while others that lack effort in all but the final year will estimate only the mean level.

Fishing impact analysis - initial conditions

For the fishery impact analysis, there is the added complication that the initial equilibrium survival rates estimated for the normal fishing run, will now be invalid because catches for particular fisheries are set to zero. For the impact analysis scenario, the initial survival rate must be recalculated as it now excludes the fishing mortalities for these fisheries with zero catches. An approach was developed that calculates the initial population using a nominal survival rate for the first iteration, and it is then solved for iteratively using the actual fishing mortalities of the dynamic model within the impact analysis, i.e. replacing the survival rates with the actual values at each iteration. Preliminary tests indicated this takes only several iterations to converge.

Note that the fisheries having non-zero fishing mortalities will have `fm_level` different to that for the normal fishing model evaluation, since the population size will be larger; and so for the observed catches, the associated fishing mortalities will be lower. This will be accounted for during the iteration procedure as the new `fm_level` values from the N-R are calculated.

Attention was also given to the formulation of the equilibrium population under conditions of multiple recruitments, regions, and movement within a calculation time period. A prototype formulation was developed that may improve on that currently being implemented in MULTIFAN-CL.

3. Survey fishery relative abundance index likelihood

Survey fisheries include indices within the catch-effort data that reflect trends in population relative abundance, i.e. CPUE observations. These fisheries include size composition data for estimating selectivity patterns that can be used for calculating fishery-specific vulnerable biomasses. Using these components, the indices may be fitted using a suitable likelihood formulation.

The survey fishery data are input as a "normal" fishery, but are controlled in how the data is treated in the model based on **fish_flags(92)**. Two configurations of a survey fishery are managed in respect of their catch and effort data.

An **extraction survey fishery** has catch data are as for a normal fishery, and therefore having removals from the population resulting from substantial fishing mortalities which are estimated via the Newton-Raphson procedure. For the fishing mortality:observed effort regression, time-invariant catchability is assumed in the relationship for this fishery. This is because this regression does also function as the method for fittting observations of relative abundance. The effort data for this fishery are standardised, such that the CPUE are indices of relative abundance, and given assumed constant catchability for the fishery in this respect, the indices are fitted using the relationship with a suitable penalty weight.

A **non-extraction survey fishery** has catch data for a "true" survey fishery in that they are negligible, e.g. 1 fish per fishing incident. As such, the fishery is excluded from the Newton-Raphson procedure for estimation of fishing mortalities. The reciprocal of the CPUE index denotes the effort data. Since the catch is arbitrary and low, it would be desirable to not require the model to fit the catch and only fit the CPUE index in its own likelihood function.

In both configurations, the selectivity patterns would be estimated via the observed size composition data.

How each fishery is modelled is controlled by fish_flags(92). The functionality of this fish_flag would be active when using the catch conditioned method for fishing mortality estimation (see section 2). **The CPUE likelihood term is only activated by this fish_flag != 0 and when the non-extraction survey fishery configuration is employed with data for the survey fisheries. Note: this flag also specifies the relative weight of the survey index likelihood: pen = ff92/10.**

Flags

fish_flags(99,fi) – specify grouping of the survey fisheries such that fisheries in a group are assumed to share the same stationary catchability. Consequently, the absolute index values among the fisheries impact upon the model predictions. This enables the assumption that the indices among fisheries may reflect relative abundance in the population. These flags operate as per normal grouping flags; i.e., fisheries that don't share stationary catchabilities should have no grouping (each fishery will have a unique group number); while survey fisheries that do share stationary catchabilities have the same group number. For example:

- num_fisheries = 20
- Fisheries 1 to 17 are normal non-survey fisheries: fish_flags(99, fi = 1 to 17) = 1 to 17
- Fisheries 18 to 20 are survey fisheries that share stationary catchability: fish_flags(99, fi = 18 to 20) = 18, 18, 18

fish_flags(92,fi) - A non-zero value activates the survey fishery CPUE index likelihood calculation for fishery fi; it also specifies the penalty weight of the survey fishery CPUE index likelihood such that pen = fish_flags(92)/10.0. For the above example:

- Fisheries 1 to 17 are normal non-survey fisheries: fish_flags(92, fi = 1 to 17) = 0
- Fisheries 18 to 20 are survey fisheries that share stationary catchability: fish_flags(92, fi = 18 to 20) = pen*10.0

fish_flags(66,fi) - A value = 1 activates the survey fishery CPUE index likelihood calculation that includes observation-specific estimates of relative precision for fishery fi. These are input as is done for the catch-errors method in column 7 (time-dependent CPUE penalty) of the fisheries data section of the *.frq input file. In the case of the survey fishery CPUE index likelihood calculation the values in this column are the estimates of relative precision for the individual indices in each time period, e.g., coefficients of variation or standard errors.

Sanity checks on the input flags are done:

- Among fish_flags(92,fi) and fish_flags(99,fi) to ensure vector shapes match; check for holes in grouping relative to penalty weightings.
- Among data_fish_flags(1,fi) and fish_flags(99,fi) to ensure all fisheries in the group have the same index type (fish numbers or biomass).

21.1 Likelihood formulation

A general form for the negative log-normal likelihood for the survey index for a specific fishery, *f*, is:

$$0.5\sum_i \log(\lambda_i\sigma^2) + 0.5\sum_i \frac{(P_i - O_i)^2}{\lambda_i\sigma^2}$$

Where O_i is the *i*th observed index, *n* is the number of observations (i.e. time periods), λ_i is the relative precision of the i^{th} observation, and B^{mid}_{if} is the stock abundance vulnerable to the survey fishery in the middle of the i^{th} time period. Taking the derivative with respect to σ and setting it equal to zero:

$$n/\sigma - \sum_i \frac{(P_i - O_i)^2}{\lambda_i\sigma^3} = 0$$

$$\widehat{\sigma}^2 = \sum_i \frac{(P_i - O_i)^2}{n\lambda_i}$$

So, the concentrated likelihood is:

$$0.5 \sum_i \log(\lambda_i) + 0.5n \log\left(\sum_i \frac{(P_i - O_i)^2}{n\lambda_i}\right)$$
$$+ 0.5 \sum_i \frac{(P_i - O_i)^2}{\lambda_i \, \widehat{\sigma}^2}$$

Conversion from a catch-errors to a catch-conditioned model – doitall flag setting changes

Flag conversions as applied for developing the SKJ2019 catch-conditioned model based upon the existing catch-errors example:

```
Phase 1:
#-------------------------------------------------------------------------------
# Catch-conditioned flags
1 382 10
1 381 0 #- de-activate debug_par
1 378 1 #- activates robust option for fm_level:effort fit relationship
-999 73 5  # activate estimation and specify degrees of polynomial for fm_level:effort fit relationship
-3 73 0  # estimate only the mean for the kludged effort fisheries
-6 73 0
-9 73 0
…
-31 73 0
1 375 1000 #- penalty wt on kludged initial survival relationship : actual survival
2 171 0 #- de-activate SRR multiplier for no-fishing option
1 374 8 #- spline degree for kludged initial survival relationship
1 373 1 #- activates the catch-conditioned option with Baranov equation
1 393 1 - activates the estimation of catch-conditioned params
2 92 2 #- specifies the catch-conditioned option with Baranov equation
1 377 1 #- activate fm_level:effort fit relationship; and generates reports
#
# Settings for catch-cond
# - grouping for fm_levels:effort regression
-1 29 1
-2 29 2
-3 29 3
…
-31 29 31
#
-999 27 0  # check if active for catch-errors method - is used in polynomial for fm_levels relationship
-999 81 1  # activate estimation of fish_pars(31) kludged_equil_coffs
#

AND

# De-activate catch-errors method flags
-999 1 0
-999 4 0
-999 10 0
-999 23 0  # turn off random walk
-999 15 0
#
```

```
  -999 13 0
 -999 66 0     # turn off for all

Phase 2:
#-----------------------------------------------------------------------------
    2 113 0       # de-activate rec_init_diff
# De-activate catch-errors method flags
    -999 1 0      # de-activate catch-errors method
    -999 4 0      # de-activate catch-errors method
#    2 35 12       # Effort deviate boundary
#    2 34 1        # Turn on eff. Devs – these get turned on in Phase 1 control phases
#  -999 4 4        # Estimate effort deviates (used for specific fisheries)
  2 144 0        # de-activate catch likelihood penalty
#    -999 45 100000   # Catch likelihood penalty

#  PHASE 4
# ---------
# Estimate seasonal catchability for all fisheries - needed for CC design matrix
 -999 27 1
 1 382 100     # Increase the penalty on Zmax_fish

#  PHASE 6
# ---------
# De-activate catch error flags
# JPPL catchability devs fsh(,15) don't matter when we actually have an index
#   -1 10 0    -1 15 0    -1 23  0
#   -2 10 1    -2 15 0    -2 23 23
…
```