**A Simple User Guide for CMSY**

N. Demirel, G. Coro, R. Froese and D. Miller

October 2015

This is an accompanying document for Froese, R., Demirel, N., Coro, G., Kleisner, K., Winker, H. (2015), Estimating Fisheries Reference Points from Catch and Resilience (submitted to Fish and Fisheries in 2015). A short description of the method and application to some ICCAT stocks is given in Froese (2015), which is available as part of the download (see below).

This document provides a simple step-by-step guide for researchers who want to apply the CMSY methods to their own data.

The required R-code and some example input files can be downloaded at

<http://data.d4science.org/uri-resolver/id?fileName=CMSY_-_Windows_OS_-_Package.zip&smp-id=56150c97e4b02e1b6570e0fe&contentType=application%2Fzip>

# Installation instructions

1) Install a recent version of R in your computer. CMSY was tested under R version 3.1.1, available from <http://www.r-project.org/>, but newer versions should also work.

2) We suggest using RStudio as an R development environment. RStudio is a free software that is available for several Operating Systems (Windows, OS, Linux, ...) and can be downloaded at <http://www.rstudio.com/products/rstudio/download/>

3) Install JAGS 3.4 for your Operating System. A version for MS Windows is already included in our package downloadable from the link above. For other Operating Systems, refer to the following web site: http://sourceforge.net/projects/mcmc-jags/files/JAGS/3.x/

4) In order to run the code, several R-packages are required. In the R Console execute the following commands (i.e. cut & paste into the Console window, then hit Enter and wait):

install.packages("R2jags")

install.packages("coda")

install.packages("lattice")

(pay attention to possible required sub-packages, as alerted in warning/error messages form the R installer, and install missing packages if needed)

5) Two different data files are required by CMSY, which should be placed in the same directory as the script. The names of these files are specified in the code, in values of variables named “catch\_file” and “id\_file”. The default file names are *Catch.csv* and *ID.csv*. Examples are already present in our package and their structure is specified in next section.

6) Make sure that the source file and the downloaded R script are in the same directory.

7) Open the downloaded CMSY R script (CMSY\_44\_WKLIFEV\_7.R) in RStudio. Use the tab “Session” and select “Set Working Directory” -> “To Source File Location”, so that the code will find the data.

8) If you want to use your own input files, just change the file names in the *catch\_file* and *id\_file* variables in the “Required settings, File names” section of the code. If you create your own input files, make sure you use the same headers (case sensitive) as in the provided example files. Make sure you are using comma-delimited (.csv) files (look at the data in a text editor to check). You can also add your data to the provided example files (recommended).

9) The R-code analyzes can either analyze all stocks in the data files (default) or a single stock can be specified in the “Select stock to be analyzed” section of the code, according to the stock identification specified at line 33. To specify the stock to analyze just put the unique fish stock name or identifier of the stock there (e.g. stocks <- “cod-2532”). To make the code run on all the stocks in the files, just comment out the whole “Select stock to be analyzed” section (put # in front of each line).

10) In RStudio, click on “Source” (or press Ctrl+Shift+S) to execute the code.

11) When the analysis is complete, results can be found in the console window as well as in the CMSY graphs window (which can be saved manually) and in an output .txt file (having the “Out\_” prefix), whose name depends on the date of the execution. An output file in .csv format containing the results is also created in the same directory, which can be opened in Excel for example. The titles of the columns are in the same order as the results in the console.

# Structure of the input files

**Structure of *Catch.csv***

For each stock, the following information must be specified (in the corresponding columns):

**stock**: a unique fish stock name or identifier (e.g. “cod-2532”), repeated for each year

**yr**: the reporting year of the catch (e.g. 2004). One row for each year.

**ct**: catch value, in tonnes (e.g. 12345). One row for each year, interpolating missing values.

**TB**: the value of the Biomass (in tonnes, e.g. 34567), or the value of the CPUE or stock size index, or NA if there is no information. Note that shorter time series with total biomass or CPUE are acceptable, but they must not contain gaps. Fill any gaps by interpolation.

**Structure of *ID.csv***

For each stock, the following information must be specified (in the corresponding columns):

**Region**: a string indicating the catch area, e.g. “Northeast Atlantic”

**stock**: a unique fish stock name or identifier (corresponding to the one in the ‘stock’ column in the catch\_file)

**Name**: a common name of the species, e.g. “Makala”

**EnglishName**: a common English name of the species, e.g. “Greater forkbeard”

**ScientificName**: the scientific name of the species, e.g. “Phycis blennoides”

**Source**: the source where the catch statistics were taken from, e.g. <http://www.ices.dk/sites/pub/Publication%20Reports/Advice/2014/2014/gfb-comb.pdf>

**MinOfYear**: the start year of the catch report

**MaxOfYear**: the last year of the catch report

**StartYear**: the start year of the Biomass/CPUE time series. This value can also be used to specify a subset of the year data to use for the CMSY.

**EndYear**: the end year of the Biomass/CPUE time series. This value can also be used to specify a subset of the year data to use for the CMSY.

**Resilience**: prior estimate of resilience, corresponding to intrinsic growth rate ranges (see next section). Allowed values are “High”, “Medium”, “Low”, “Very Low”.

**r\_low/r\_high**: an optional pair of parameters to specify the range of intrinsic growth rate for the species. Set these to NA to use the range specified in Resilience. If values are given, the Resilience column will be ignored.

**stb\_low/stb\_hi**: the estimated relative biomass (B/k) range at the beginning of the catch report period (see next section).

**intyr**: the year in the catch time series range, to be taken as reference for intermediate biomass level. Set it to NA to have it estimated from the year with maximum or minimum catch

**intbio\_low/intbio\_high**: the estimated intermediate relative biomass range (see next section). Set it to NA to make to have it estimated from maximum or minimum catch, according to some simple rules (note: these may not be appropriate for your stock).

**endbio\_low/ endbio\_hi**: the estimated relative biomass (B/k) range at the end of the catch time series (see next section).

**Btype**: the type of information in the TB column of the catch\_file. Allowed values are “observed” (when total biomass is reported), “CPUE” or “None”

**FutureCrash**: indication about possible crash of the stock at the end of the catch time series. Allowed values are “No”, if the stock is deemed well above Blim and “Possible” when the stock is close to or below Blim.

**Comment**: a comment on the stock, to be reported in the output statistics.

Remember that the files must be saved in “csv” (comma delimited) format. Double-check that indeed a comma (and not a semi-colon) is used as delimiter.

# Suggestion for parameters setting

Table 1 reports a set of questions that can help to set the CMSY input parameters. Please note that answers can also be derived other stock assessment tools, such as length frequency analysis or catch per unit of effort.

**Table 1**. Example of questions to be put to experts to establish priors for CMSY analysis.

|  |  |
| --- | --- |
| **Prior** | **Question to experts** |
| Start year for catch time series | From what year onward are catch data deemed reliable? |
| Relative start and end biomass  *B/B0* | What was the most likely stock status at the beginning and end of that time series: light, full, or overfishing? |
| Relative intermediate biomass  *B/B0* | Is there an intermediate year where biomass is considered to have been particular low or high, e.g., exploitation changed from light to full, or where an extraordinary large year class entered the fishery? |
| Resilience prior | What is your best guess for the range of values including natural mortality of adults (*M*)? Considering the relationship *M* ≈ *r/2* |
| Resilience prior | What is your best guess for the range of values including maximum sustainable fishing mortality (*Fmsy*)? Considering the relationship *Fmsy* ≈ *r/2* |
| FutureCrash: Possible / No | If current catches continue, is it likely that the stock will be outside of safe biological limits within the next 3 years? E.g. *B/B0* < 0.2? |

Table 2 suggests ranges for relative biomass to be used as input parameters, depending on the depletion status of the stock. Alternatively, you can get preliminary estimates of *r* from the following empirical relations:

*r* ≈ 2 *M* ≈ 2 *Fmsy* ≈ 3 *K* ≈ 3/*tgen* ≈ 9/*tmax*

where *r* is the intrinsic rate of population increase, *M* is the rate of natural mortality, *Fmsy* is the maximum sustainable fishing mortality, *K* is the somatic growth rate, *tgen* is generation time, and *tmax* is maximum age. If point estimates are very close to each other, assume a range of uncertainty of +/- 50%.

**Table 2**. Prior relative biomass (b/k) ranges forCMSY.

|  |  |  |
| --- | --- | --- |
| **Point in time series** | **Strong depletion** | **Low depletion** |
| Beginning | 0.1 – 0.5 | 0.5 – 0.9 |
| Intermediate | 0.01 – 0.4 | 0.3 – 0.9 |
| End | 0.01 – 0.4 | 0.4 – 0.8 |

Table 3 reports the *r* ranges automatically associated by CMSY to the resilience parameter values.

**Table 3**. Prior ranges for parameter *r*, based on classification of resilience.

|  |  |
| --- | --- |
| **Resilience** | **prior *r* range** |
| High | 0.6 – 1.5 |
| Medium | 0.2 – 0.8 |
| Low | 0.05 – 0.5 |
| Very low | 0.015 – 0.1 |

When setting an intermediate biomass, it often improves the CMSY analysis if the end of a period with low biomass is indicated by setting the intermediate year to the last year with low biomass, and indicating a respective relative range, e.g. as 0.01 – 0.4. Similarly, indicate a period of large biomass by setting the intermediate year to the last year with high biomass and indicate a respective range, e.g. as 0.4 – 0.8. In general, the width of relative biomass windows should not be less than 0.4, as in the previous examples. Setting a range of 0 to 1 is also possible, and would indicate no information at all about stock status, which is, however, unlikely. If a stock is fished it must be smaller than 1. If it is delivering decent catches, it must be larger than 0.01. See Table 1 for guidance on how to get priors from interviews with fishers or experts (or yourself).

Note that if Biomass/CPUE values are provided, part of the CMSY analysis is also an analysis with a Bayesian state-space implementation of a Schaefer model (BSM). These results are shown in red in the graphical output. You can change the minimum number of years with total biomass or CPUE required for BSM analysis (variable ‘nab’ in the “General settings for the analysis” section, see row 46 of the code), but it should not be much less than the recommended 10 years. You can also change the uncertainty associated with catch data (variable ‘dataUncert’ in the “General settings for the analysis” section of the code), in row 38, but it should not be much higher than 0.2, because without reliable catch data, CMSY makes no sense. The default uncertainty is 0.1 (i.e. 10%).

If you have any questions, contact Rainer Froese ([rfroese@geomar.de](mailto:rfroese@geomar.de)) or Gianpaolo Coro (gianpaolo.coro@isti.cnr.it) or Nazli Demirel (ndemirel@geomar.de).

Enjoy using CMSY.