European
Commission

JRC TECHNICAL REPORTS

a4a assessment model simulation testing

Jardim E.
Millar C.
Ferretti M.
Mosqueira I.
Osio C.
Scott F.

2013

Report EUR 26156 EN

Joint
Research
Centre

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information
Millar Colin
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 051, 21027 Ispra (VA), Italy
E-mail: colin.millar@jrc.ec.europa.eu
Tel.: +39 0332 78 5208
Fax: +39 0332 78 9658

http://www.jrc.ec.europa.eu/

JRC 84620

EUR 26156 EN

ISBN 978-92-79-33134-3

ISSN 1831-9424

doi:10.2788/22675
Luxembourg: Publications Office of the European Union, 2013

Printed in Italy

a4a assessment model simulation testing

Jardim, E., Millar, C., Ferretti, M., Mosqueira, I., Osio, C. and Scott, F.

August 29, 2013

Contents

1 Introduction

Simulation testing provides information about the performance of the model in specific circumstances, which is a major source of information for analysts. Informing about regions of the parameters space that may require more attention in order to overcome problems that surfaced during testing phase. Simulation testing *Per se* does not cover the full range of situations that may occur, these are a lot wider than it's possible to simulate. The simulation environment does allow a controlled experiment to be carried out which gives information about the capacity of the model to reconstruct the underlying reality under specific conditions. Furthermore, these simulation studies allows testing the adjustment of models without much human intervention (automatic adjustments), a major objective of the framework developed here.

To run the simulation study the following algorithm was applied:

1. Use information about life history traits of several fish stocks to build coherent population dynamics under no-exploitation scenario.

2. Simulate for each stock a 50 year exploitation history based on the common development-over exploitation-recovery pattern.

3. Trim the simulations in four 15 year periods, generating a range of exploitation patterns that are commonly observed in global fisheries: (i) "development" - from non-exploited fisheries up to over-exploitation, (ii) "development plus over-exploitation" - five years of increasing fishing mortality and 10 years of over-exploitation, (iii) "over-exploitation" - 15 years of stable fishing mortality at over-exploitation levels, and (iv) "recovery" - 5 years of over-exploitation and 10 years decreasing fishing mortality down to the maximum sustainable yield (MSY) fishing mortality.

4. Add observation error in abundance indices considering several assumptions about catchability and independent lognormal errors with several levels of variability.

5. Add observation error in catch in numbers at age in the form of independent lognormal errors with several levels of variability.

6. Fit a random assessment model to each simulation by selecting the sub-models from a set of three distinct F models, five distinct Q models and two distinct R models.

7. Compute performance statistics.

The approach taken regarding the population dynamics does not try to simulate exactly specific species, it aims to simulate stocks and exploitation histories that are consistent with population dynamics theory and loosely based on the biology of the species, so that a large range of life histories traits and commercial exploitation patterns are considered.

For this simulation exercise, life history parameters were extracted from FishBase [1] for all marine species available. Over-exploitation was defined by fishing levels of 80% of the crash fishing mortality [2] or five times the level of $F_{0.1}$ [3].

The results were integrated in an on-line application, using a database and a visualization tool that allows the users to search and download subsets of the data, as well as visualizing directly the performance for particular scenarios. Instead of presenting a massive amount of analysis in this paper the information can be easily analysed on-line or extracted. The information can be used to evaluate the performance of the model, but it can also be used as a starting point to carry out stock assessments, learn how to use the framework, etc. Ultimately it shows that it is possible to run a massive number of stock assessments and make the results available and manageable for those interested in meta analysis.

Note that all the analysis are carried out in R[1], using the FLR[2] libraries and the a4a package for the stock assessment model. The document is written in LaTeX using the R package knitr[3], which embeds the code in the document. The major advantage is allowing full replicability of the analysis and readers to check the code.

This document is supplemental material to the paper Colin, et.al (2013) and shows the scenarios and code used for the simulation study 1.

[1] R Core Team ...
[2] Kell et. al
[3] REF

2 Webscrap fishbase

```
## Error:  there is no package called 'FLAdvice'
## Error:  no existing definition for function 'sv'
```

```
library(FLCore)
library(FLAdvice)
library(rfishbase)
library(longtable)
library(XML)
fish.data <- loadCache()

# ============================================================
#' extract information in rfishbase
# ============================================================
ids <- unlist(lapply(fish.data, "[", "id"))
order <- unlist(lapply(fish.data, "[", "Order"))
family <- unlist(lapply(fish.data, "[", "Family"))
spp <- unlist(lapply(fish.data, "[", "ScientificName"))
maxage <- getSize(fish.data, "age")
maxlen <- getSize(fish.data, "length")
maxwgt <- getSize(fish.data, "weight")
mar <- vector("logical", length = length(ids))
mar[grep("marine", unlist(lapply(fish.data, "[", "habitat")))] <- TRUE

# ============================================================
#' init results dataframe with data from rfishbase
# ============================================================
lhPars <- data.frame(id = ids, order = order, family = family, species = spp,
    marine = mar, a = NA, b = NA, k = NA, linf = NA, t0 = NA, a50 = NA, l50 = NA,
    lmax = maxlen, amax = maxage, wmax = maxwgt, stringsAsFactors = FALSE)

# ============================================================
#' scrap
# ============================================================
for (i in ids) {
    cat(i, ";")
    # --------------------------------------------------------
    #' growth
    # --------------------------------------------------------
    addr <- paste("http://www.fishbase.org/PopDyn/PopGrowthList.php?ID=", i,
        sep = "")
    tab <- try(readHTMLTable(addr))
    if (is(tab, "try-error"))
        tab$dataTable <- NULL
    if (!is.null(tab$dataTable)) {
        # linf
        v <- as.character(tab$dataTable[, "Loo(cm)"])
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "linf"] <- median(as.numeric(v), na.rm = T)
        # k
        v <- as.character(tab$dataTable[, "K(1/y)"])
```

```
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "k"] <- median(as.numeric(v), na.rm = T)
        # t0
        v <- as.character(tab$dataTable[, "to(years)"])
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "t0"] <- median(as.numeric(v), na.rm = T)
    }
    # ----------------------------------------------------------
    #' maturity
    # ----------------------------------------------------------
    addr <- paste("http://www.fishbase.org/Reproduction/MaturityList.php?ID=",
        i, sep = "")
    tab <- try(readHTMLTable(addr))
    if (is(tab, "try-error"))
        tab$dataTable <- NULL
    if (!is.null(tab$dataTable)) {
        # l50
        v <- as.character(tab$dataTable[, "Lm(cm)"])
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "l50"] <- median(as.numeric(v), na.rm = T)
        # a50
        v <- as.character(tab$dataTable[, 8])
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "a50"] <- median(as.numeric(v), na.rm = T)
    }
    # ----------------------------------------------------------
    #' l~w
    # ----------------------------------------------------------
    addr <- paste("http://www.fishbase.org/PopDyn/LWRelationshipList.php?ID=",
        i, sep = "")
    tab <- try(readHTMLTable(addr))
    if (is(tab, "try-error"))
        tab$dataTable <- NULL else names(tab)[3] <- "dataTable"
    if (!is.null(tab$dataTable)) {
        # a
        v <- as.character(tab$dataTable[, 2])
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "a"] <- median(as.numeric(v), na.rm = T)
```

```
        # b
        v <- as.character(tab$dataTable[, 3])
        for (j in 1:length(v)) {
            vj <- v[j]
            vj <- utf8ToInt(vj)
            v[j] <- intToUtf8(vj[vj >= 46 & vj <= 57])
        }
        lhPars[lhPars$id == i, "b"] <- median(as.numeric(v), na.rm = T)
    }
}
```

```
# ============================================================
#' cleaning & replacing missing values
# ============================================================
lhMar <- subset(lhPars, marine == TRUE)
lhMar[lhMar == 0] <- NA
# ------------------------------------------------------------
#' remove sea horses
# ------------------------------------------------------------
lhMar01 <- subset(lhMar, !(is.na(a) | is.na(b) | is.na(k) | is.na(linf)) & order !=
    "Syngnathiformes")
# ------------------------------------------------------------
#' id those that were observations
# ------------------------------------------------------------
lhMar01$l50obs <- !is.na(lhMar01$l50)
lhMar01$a50obs <- !is.na(lhMar01$a50)
lhMar01$amaxobs <- !is.na(lhMar01$amax)
lhMar01$t0obs <- !is.na(lhMar01$t0)
lhMar01 <- transform(lhMar01, obs = (l50obs + a50obs + amaxobs + t0obs) > 0)
```

```
# ============================================================
#' estimate missing values for l50, a50 and amax
# ============================================================
rlm01 <- rlm(log(l50) ~ log(linf), data = subset(lhMar01, linf < 300))
lhMar01[is.na(lhMar01$l50), "l50"] <- exp(predict(rlm01, newdata = data.frame(linf = lhMar01[is.na(lhMa
    "linf"])) + rnorm(sum(is.na(lhMar01$l50)), 0, sd(residuals(rlm01))))

rlm02 <- rlm(log(t0) ~ log(k), data = subset(lhMar01, linf < 300))
lhMar01[is.na(lhMar01$t0), "t0"] <- exp(predict(rlm02, newdata = data.frame(k = lhMar01[is.na(lhMar01$t
    "k"])) + rnorm(sum(is.na(lhMar01$t0)), 0, sd(residuals(rlm02))))

rlm03 <- rlm(log(amax) ~ log(linf), data = lhMar01)
lhMar01[is.na(lhMar01$amax), "amax"] <- exp(predict(rlm03, newdata = data.frame(linf = lhMar01[is.na(lh
    "linf"])) + rnorm(sum(is.na(lhMar01$amax)), 0, sd(residuals(rlm03))))

vv <- -log(1 - (lhMar01$l50/lhMar01$linf))/lhMar01$k + lhMar01$t0
lhMar01[is.na(lhMar01$a50), "a50"] <- vv[is.na(lhMar01$a50)]
```

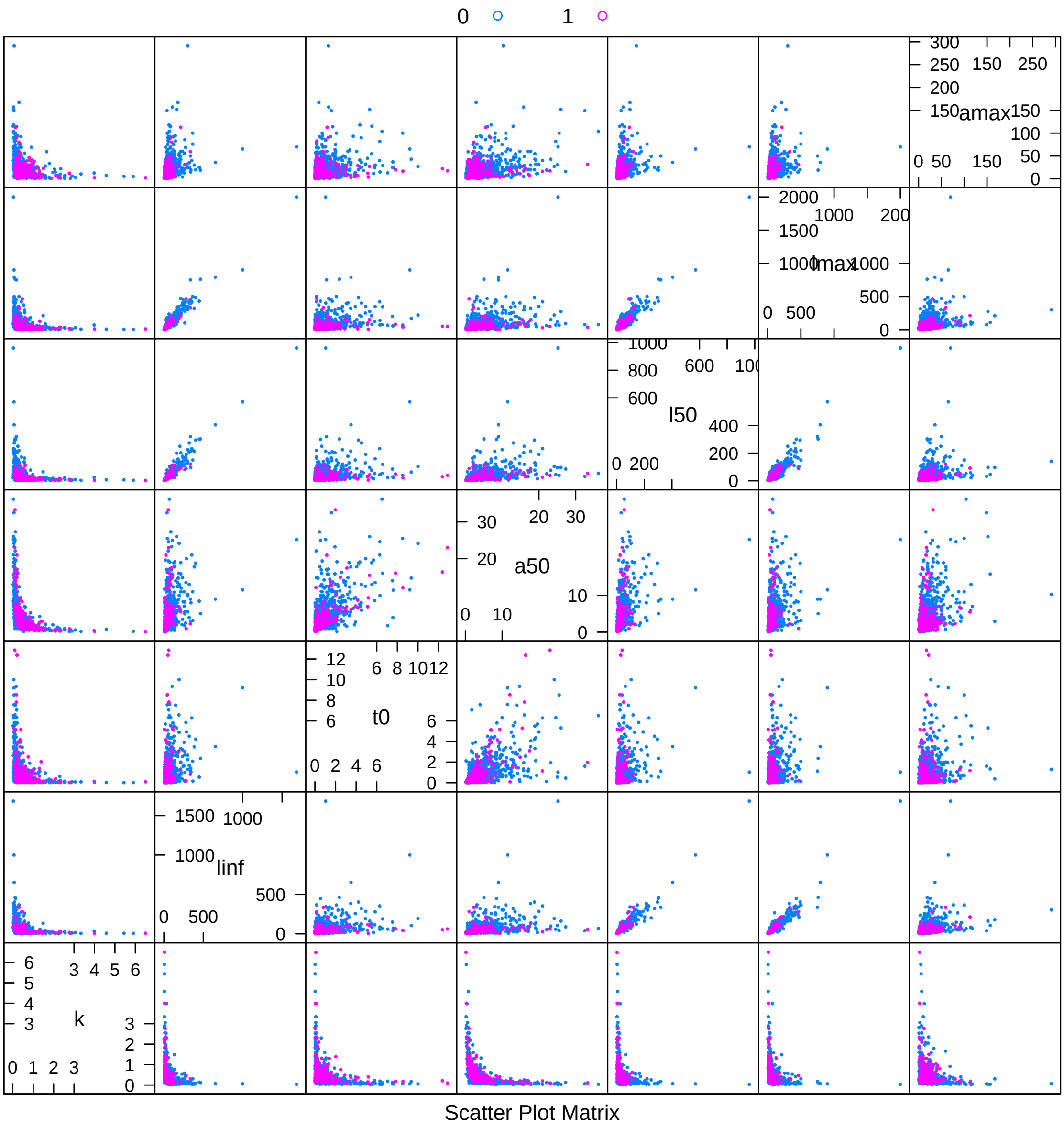

Figure 1. Scrapped values and linear model estimates to replace missing values

```
# ------------------------------------------------------------
#' variables from fishbase
# ------------------------------------------------------------
names(lhMar01)

## [1] "id"       "order"    "family"   "species" "marine"  "a"       "b"
## [8] "k"        "linf"     "t0"       "a50"     "l50"     "lmax"    "amax"
## [15] "wmax"     "l50obs"   "a50obs"   "amaxobs" "t0obs"   "obs"

# ------------------------------------------------------------
#' number of species
# ------------------------------------------------------------
nrow(lhMar01)

## [1] 1053
```

3 Scenarios

For each species the following scenarios were simulated.

```
# ------------------------------------------------------------
#' Operating models
# ------------------------------------------------------------
#' stock recruitment models
srMod <- c("bevholt", "ricker")
#' steepness of stock recruitment relationship
s <- c(0.8, 0.6)
#' age of 50% selectivity (maximum of double normal model)
a1 <- c(0.7, 1)  # relative to a50
#' variance of the right half of the double normal
sr <- c(1, 100)
#' variance of the left half of the double normal
sl <- c(1, 100)

# ------------------------------------------------------------
#' Observation error
# ------------------------------------------------------------
#' coefficient of variation for catch at age lognormal errors
oe.ccv <- c(0.1, 0.3)
#' coefficient of variation for index at age lognormal errors
oe.icv <- c(0.2, 0.5)
#' linear increase by year for index catchability
oe.iq <- c(1, 1.05)  # technical creep, q = 0.01
```

Scenarios with flat selectivity and a1=0.7 were removed once a1 has no effect. Also cases that end up with missing values in the age of 50% mature fish were removed. Finally the number of scenarios were 224 that when merged with the species and for each combination five exploitation histories were simulated, giving raise to approximately 1.15 million simulations.

4 Model fit

Afterwards a set of models were built, considering the most common options used in stock assessment (see below) which generate a total of 30 combinations. For each simulation one of these combinations was allocated randomly to be used in the model fit.

Table 1. Sub-models used for fitting.

submodel	code	formula
fishery	fm1	~factor(age) + factor(year)
fishery	fm2	~bs(age, 4) + bs(year, 10)
fishery	fm3	~te(age, year, bs = c("tp", "tp"), k = c(4, 15))
catchability	qm0	~1
catchability	qm1	~age
catchability	qm2	~factor(age)
catchability	qm3	~bs(age, 4)
catchability	qm4	~bs(age, 4) + bs(year, 15)
recruitment	rm1	~factor(year)
recruitment	rm2	~bs(year, 15)

5 Results

The results are made available on a dedicated web page on the a4a website. The page accesses a database with all the results and provides visualization and exporting tools.

5.1 Export file data structure

The data is split in two tables, "stats" and "summs". The first has summary stats of the tests. Table "stats" is linked one to many with table "summs", which has summaries of the stock assessment fits. Both tables have the field "scnid", which is the scenario identification field. The scenario identification kwy is build by concatenating other fields, which are identified by the string "scnid" in bold. This field can be used to link both tables.

5.1.1 Table "stats" fields

- scenario parameters
  - order, text, taxon order,
  - family, text, taxon family,
  - species, text, taxon species, **scnid**
  - marine, bolean, marine species or not,
  - a, numeric, parameter of the weight-length relationship,
  - b, numeric, parameter of the weight-length relationship,
  - k, numeric, parameter of the von Bertalanffy growth model,
  - linf, numeric, parameter of the von Bertalanffy growth model,
  - t0, numeric, parameter of the von Bertalanffy growth model,
  - a50, numeric, age of 50% maturity,
  - l50, numeric, length of 50% maturity,
  - amax, numeric, maximum age,
  - lmax, numeric, maximum length,
  - wmax, numeric, maximum weight,
  - l50obs, bolean, was l50 observed or not,
  - a50obs, bolean, was a50 observed or not,
  - amaxobs, bolean, was amax observed or not,
  - t0obs, bolean, was t0 observed or not,
  - obs, ignore,,
  - srMod, text, stock-recruitment model, **scnid**
  - s, numeric, steepness parameter of the stock-recruitment curve, **scnid**
  - v, numeric, virgin biomass parameter of the stock-recruitment curve, **scnid**
  - a1, numeric, age of 50% selectivity parameter of the selectivity double normal model, **scnid**
  - sr, numeric, right variance parameter of the selectivity double normal model, **scnid**
  - sl, numeric, left variance parameter of the selectivity double normal model, **scnid**
  - oe.icv, numeric, abundance index coefficient of variation parameter of the observation error model, **scnid**
  - oe.iq, numeric, abundance index catchability increase parameter of the observation error model, **scnid**
  - oe.ccv, numeric, catch at age coefficient of variation parameter of the observation error model, **scnid**

  - qmodel, text, catchability model, **scnid**
  - rmodel, text, recruitment model, **scnid**
  - fmodel, text, fishing mortality model, **scnid**
  - fmsy, numeric, fmsy fishing mortality reference point,
  - f0.1, numeric, f0.1 fishing mortality reference point,
  - m, numeric, natural mortality average over age range,
  - expl, numeric, exploitation history pattern, **scnid**

- fitting statistics
  - nopar, numeric, number of model parameters,
  - nlogl, numeric, negative log likelihood of the fit,
  - maxgrad, numeric, maximum gradient of the negative log likelihood surface,
  - npar, ignore,,
  - logDetHess, ignore,,

- comparison statistics
  - ssbrbias, numeric, SSB relative bias,
  - ssbmse, numeric, SSB mean square error,
  - fbarrbias, numeric, fishing mortality relative bias,
  - fbarmse, numeric, fishing mortality mean square error,
  - recrbias, numeric, recruitment relative bias,
  - recmse, numeric, recruitment mean square error,
  - catrbias, numeric, catch relative bias,
  - catmse, numeric, catch mean square error,
  - qrbias, numeric, catchability relative bias,
  - qmse, numeric, catchability mean square error,

- scnid, text, scenario id,
- dynid, text, population dynamics id,

5.1.2 Table "stats" fields

- comparison time series
  - y, numeric, year,
  - stat, text, statistic with values "S" (spawning stock biomass) "F" (fishing mortality) "R" (recruitment) or "C" (catch),
  - src, text, source of information with values "obs" (observed/simulated) or "hat" (estimated)
  - val, numeric, value

- scnid, text, scenario id,
- dynid, text, population dynamics id,

References

[1] Froese R, Pauly D, Editors (2013). FishBase. World Wide Web electronic publication. URL `www.fishbase.org`. Version (02/2013).

[2] Quinn TJ, Deriso RB (1999) Quantitative fish dynamics. Oxford University Press, USA.

[3] Gulland JA, Boerema LK (1973) Scientific advice on catch levels. Fishery Bulletin 71(2): 325-335.

European Commission
EUR 26156 – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: a4a assessment model simulation testing

Author(s): Jardim, E., Millar, C., Ferretti, M., Mosqueira, I., Osio, C. and Scott, F.

Luxembourg: Publications Office of the European Union

2013 – 10 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online), ISSN 1018-5593 (print)
ISBN 978-92-79-33134-3

doi:10.2788/22675

Abstract

The a4a initiative aims to provide timely and cost effective advice for the circa. 250 fish stocks that, through the EU Data Collection Framework, will have at least 10 years of data by the year 2020. Current processes for assessing the state of and managing fish stocks are intensive processes, each stock requiring the attention of one or more stock assessment scientist to produce preliminary catch advice, which is subsequently reviewed by one or two committees before the final catch advice is published. Ingrained in the development of these processes has been the development of more and more complex stock assessment models which typically require highly skilled personnel to set up and run.

The a4a initiative seeks to overcome these issues by developing a flexible, robust and easy to use stock assessment model, thus making stock assessment accessible to a wide range of scientists that do not have the high skilled quantitative background required to run very complex models. Forthcoming research will describe how to overcome the burden of producing catch advice for such a large number of stocks. This technical report presents assessment model simulation testing undertaken under the a4a Initiative.

LB-NA-26156-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle. Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

ISBN 978-92-79-33134-3

Publications Office