Report for ICCAT-GBYP 04/2013

Marie-Pierre Etienne*[1], Tom Carruthers [2], and Murdoch McAllister[2]

[1]AgroParisTech
[2]UBC

January 15, 2014

Contents

*marie.etienne@agroparistech.fr

1 Introduction

The uncertainty in abundance indices may be important in some fisheries. Furthermore the recruitment model used to represent the population dynamic is a simpler version of a tricky reality. Statistical catch at age model allow to take those sources of uncertainty within an age structure population dynamic model.

The work reported in this document aims at

1. examining the opportunity of using a statistical catch at age for Bluefin tuna stock evaluation,

2. highlighting the structure and the assumptions of the chosen statistical model (`iSCAM` developed by S. Martell [Martell, 2012])

3. running a simulation study for evaluating the quality of the estimation procedure,

4. producing operational codes for running a stock evaluation on Bluefin Tuna east stock with a statistical catch at age model,

5. comparing the results with the results presented in previous stock assessment.

The report is split in five main sections. The first section describes the model proposed in `iSCAM` . The second section describes briefly the data used for this test. Section three presents the practical choices made to run `iSCAM` on this dataset. The detailed simulation study is presented in section 5. The complete results on Bluefin Tuna, eastern stock (BFTE) are developed in the last main section. The discussion sums up the main results and highlights the strengths and the weaknesses of `iSCAM` used for BFTE stock assessment.

All the codes used to produce this work are freely available on `https://github.com/MarieEtienne/ICCAT-BFT`

2 Description of the main aspects of `iSCAM`

This section sums up the principle aspects of `iSCAM` used on BFTE.

2.1 Latent process Level

Population model

iSCAM is an age structured model. Fish population is split into age classes from age *sage* to age *nage*. Let us denote by $\boldsymbol{N}_t = (N_{1,t}, \ldots, N_{A,t})$ the number of individuals in every class age a ($a \in [1, A]$) in year t ($t \in [1; T]$).

$$\text{Population dynamic after year syr (t>1)} \tag{1}$$

$$N_{a,t} = \begin{cases} R_t, \quad a = 1 \\ N_{a-1,t-1} \exp(-Z_{t-1,a-1}), \quad a \in [2; A-1] \\ N_{a-1,t-1} \exp(-Z_{t-1,a-1}) + N_{a,t-1} \exp(-Z_{t-1,a}), \quad a = A, \end{cases} \tag{2}$$

$$\tag{3}$$

where $Z_{a,t}$ stands for the total mortality rate at age a in year y and R_t is the recruitment at year t.

$$\text{Recruitment after year syr (t>1)} \tag{4}$$

$$R_t = \frac{s_0 B_{t-1}}{1 + \beta B_{t-1}} \exp \epsilon_t^R, \tag{5}$$

$$\epsilon_t^R \underset{i.i.d}{\sim} \mathcal{N}\left(-\frac{\tau_R^2}{2}, \tau_R^2\right), \tag{6}$$

$$\tag{7}$$

iSCAM assumes a stock recruitment relationship which links the mature biomass at time t B_t and the recruitment at time $t+1$. The two options for this stock recruitment model are Beverton and Holt (BH) or Ricker (R) model. The presented work used the BH model.

To circumvent estimation issues, iSCAM uses a trick and during the first estimation phases, the recruitment model is specified by the following equation:

$$R_t = \bar{R} e^{\omega_t}, \tag{8}$$

where $\omega_t \overset{i.i.d}{\sim} \mathcal{N}(0, \tau_R^2)$ and $\bar{R}$ is the average recruitment.

This approach in the early phases is close to the VPA from the recruitment point of view, since it doesn't assume any stock recruitment relationship.

During the last phase, the BH relationship is estimated. Turning off this estimation is possible and would allow to base management evaluation strategies on an average annual recruitment.

Special attention is required for the first year of the study. The recruitment parameter R_{init} during this first year has to be estimated. If the stock is considered in unfished condition, then R_{init} is assumed to be equal to R_0, otherwise R_{init} is specifically estimated.

$$\text{Definition of mature biomass} \tag{9}$$

$$B_t = \sum_{a=1}^{A} N_{t,a} f_a, \tag{10}$$

f_a being the fertility at age a.

f_a doesn't depend on year t, fertility at age is considered as fixed over years

$$\text{Definition of the total mortality rate ate age} \tag{11}$$

$$Z_{t,a} = M_a + \sum_{k=1}^{K} F_{k,t} v_{k,t,a}, \quad \text{gear } k, \text{class } a, \text{ at year } t. \tag{12}$$

$$\tag{13}$$

In the BFT data available there is only one time series for catch (catch are not split among different gears), considering that the gear index corresponding to the commercial fisheries equals one, the previous expression may be simplified as :

$$\text{Definition of total mortality rate ate age} \tag{14}$$

$$Z_{t,a} = M_a + F_t v_{1,t,a}, \quad \text{gear } 1, \text{class } a, \text{year } t, \tag{15}$$

$$\tag{16}$$

where F_t is the instantaneous fishing mortality and $v_{1,t,a}$ is the vulnerability for gear 1, in year t for age class a.

The weight at age is derived from an allometric relationship between length and weight, and from the von Bertalanffy growth equation to link age class and length. The fertility at age is assumed to follow a logistic function with parameters $\mu_{mat50\%}$ and $\sigma_{mat50\%}$

$$\text{Life trait specification} \tag{17}$$

$$l_a = l_\infty \left(1 - \exp\left(-k\left(a - t_0\right)\right)\right) \tag{18}$$

$$w_a = a_w \, l_a^{b_w} \tag{19}$$

$$f_a = w_a \frac{1}{1 + \exp\left(\frac{\mu_f - a}{\sigma_f}\right)} \tag{20}$$

$$\tag{21}$$

Vulnerability model

`iSCAM` allows to specify different form for the selectivity/vulnerability. The vulnerability may either be completely specified by the user, or when age composition data are available the vulnerability may be inferred. In the latter case, different form of selectivity curves may be specified: for example the selectivity may be chosen as a logistic function of age, or even, with a more flexible approach, using B-spline. The B-splines functions may even model a variation across years. In this work, only logistic or simple Bspline has been used to circumvent estimation issues and avoid over-parametrization.

2.2 Observation Level

Age composition data

Following [Schnute and Richards, 1995], `iSCAM` uses by default a multivariate logistic function for age composition data. It is assumed that p_{atk} which is the proportion of fish of age a in year t for gear k, is drawn from the following distribution (gear k is omitted for clarity) which is defined thanks to a latent variable X_{at}

$$\begin{aligned}
\epsilon_{at}^{A} &\overset{i.i.d}{\sim} \mathcal{N}(0, \tau_A^2) \\
X_{at} &= \log \mu_{at} + \epsilon_{at}^{A} - \frac{1}{A} \sum_{a=1}^{A} \left(\log \mu_{at} + \epsilon_{at}^{A} \right) \\
e^{X_{at}} &= \frac{\mu_{at} e^{\epsilon_{at}^{A}}}{\left(\prod_{a=1}^{A} \mu_{at} e^{\epsilon_{at}^{A}} \right)^{1/A}} \\
p_{at} &= \frac{e^{X_{at}}}{\sum_{a=1}^{A} e^{X_{at}}} \qquad (22) \\
& \qquad (23)
\end{aligned}$$

The multivariate logistic distribution avoids the drawback of a very high precision when using a classical multinomial distribution.

Abundance indices

The total vulnerable biomass at year t for gear k is defined by

$$V_{k,t} = \sum_{a=1}^{A} N_{t,a} e^{-\lambda_{k,t} Z_{t,a}} v_{k,a} w_a, \qquad (24)$$

where $\lambda_{k,t}$ is the fraction of the mortality to adjust for survey timing, it is specified by the user.

$$\begin{aligned}
I_{k,t} &= q_k V_{k,t} \exp \epsilon_{k,t}^{I} \qquad (25) \\
\epsilon_{k,t}^{I} &\overset{i.i.d}{\sim} \mathcal{N}(0, \tau_I^2)
\end{aligned}$$

Catch

Catch with gear k in year t is denoted by $C_{k,t}$ and defined using Baranov Catch equations by

$$\hat{C}_{k,t} = \sum_{a=1}^{A} \frac{N_{t,a} w_a F_{k,t} v_{k,t,a} (1 - e^{-Z_{t,a}})}{Z_{t,a}},$$

Since all catch are aggregated and treated as one gear, this equation simplifies in

$$\hat{C}_t = \sum_{a=1}^{A} \frac{N_{t,a} w_a F_t v_{1,t,a} (1 - e^{-Z_{t,a}})}{Z_{t,a}}, \tag{26}$$
$$C_t = \hat{C}_t \exp \epsilon_t^C, \quad \epsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, \tau_C^2)$$

2.3 Summary of the symbols in **iSCAM**

Table 1 and 2 recall the main symbols used in this report and give their corresponding name in iSCAM . Further description is available in [Martell et al., 2011]. Some of the names have been changed to unify the notation standards. When the name is shared by this this paper and iSCAM users Guide, nothing is specified.

Parameters	Name in iSCAM	Signification
τ_C	σ_C	Standard deviation in observed catch - This parameter is not estimated but set to a given value
τ_A	τ, mentioned only through its estimation on page 21 of [Martell, 2012]	standard deviation in the multivariate logistic function
$v_{k,a}$		vulnerability for gear k at age a
R_0		Recruitment in unfished conditions
h		Steepness
q_k		Catchability coefficient for gear k
ρ		the proportion of total variation allocated to the variance of the survey index
φ^2	$1/\mathcal{V}^2$	total variance $\varphi^2 = \tau_R^2 + \tau_I^2$
$\tau_I = \varphi * \sqrt{\rho}$	σ	Standard deviation in the survey index
$\tau_R = \varphi * \sqrt{1-\rho}$	τ	Standard deviation in process errors

Table 1: Leading parameters

Other symbols

Name	Name in `iSCAM`	Signification
A		last class age
$\kappa = 4h/(1-h)$		Goodyear compensation ratio for BH
$s_0 = \kappa/\Phi_E$		BH relationship
$\beta = (\kappa - 1)/(R_0\Phi_E)$		BH relationship
ϵ_t^R	δ_t	recruitment errors
$\epsilon_{t,a}^A$	$\eta_{t,a}$	errors in the multivariate logistic
$\epsilon_{k,t}^I$	$\epsilon_{k,t}$	residuals in abundance survey
ϵ_t^C	$\eta_{k,t}$	residuals in catch data
$\bar{R}$		Average recruitment used in first estimation phases as mentioned in equation 8
B_t		Spawning biomass
$N_{a,t}$		Number of individuals of age a in year t
f_a		fecundity at age a
M_a		instantaneous mortality rate at age a
$F_{k,t}, F_t$		Instantaneous fishing mortality for gear k in year t, since there is only one commercial gear considered, index k is mostly omitted
$Z_{a,t}$		Instantaneous total mortality for age a in year t
$X_{a,t}$	not named in `iSCAM`	Quantities involved in the multivariate logistic
$\mu_{a,t}$		unnormalized proportion at age
	$\widehat{p_{a,t}} = \mu_{a,t}/\sum_a \mu_{a,t}$	proportion at age
$p_{a,t}$		Observed proportion at age
$V_{k,t}$		Total vulnerable biomass for gear k in year t
w_a		Weight at age a

Table 2: List of symbols used in `iSCAM`

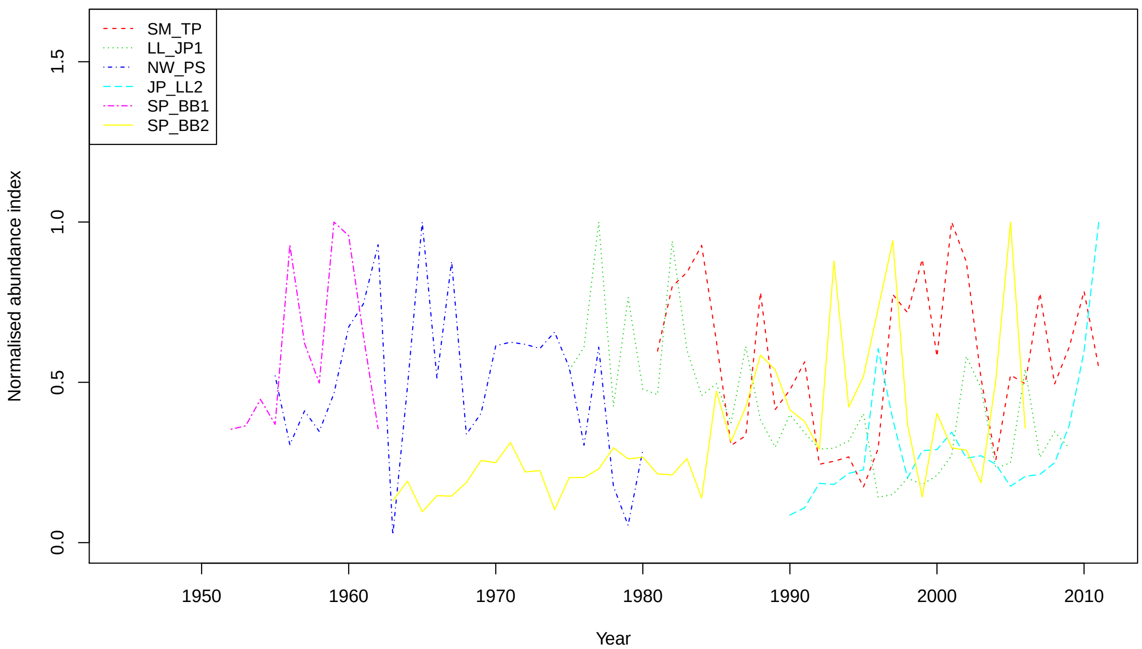


Figure 1: Scaled abundance indices

3 Description of Bluefin Tuna data

As requested, the presented work focus on BFTE data. There are seven abundance indices and the total commercial catch available. This section describes the available data and the way they have been included in the model.

3.1 Abundance indices

Seven abundance indices are available. the oldest one starts in year 1950, the last indices considered stop in 2011. Three indices are assumed to be proportional to the total number of tuna (`SM_TP`, `LL_JP1`, `LL_JP2`), the four others are assumed to be proportional to the total biomass (expressed as weight).

We focus only on the six first abundance indices and ignore the index `SP_BB3`. This last one holds only for a short period of time and lead to very poor estimation which produces numerical instability.

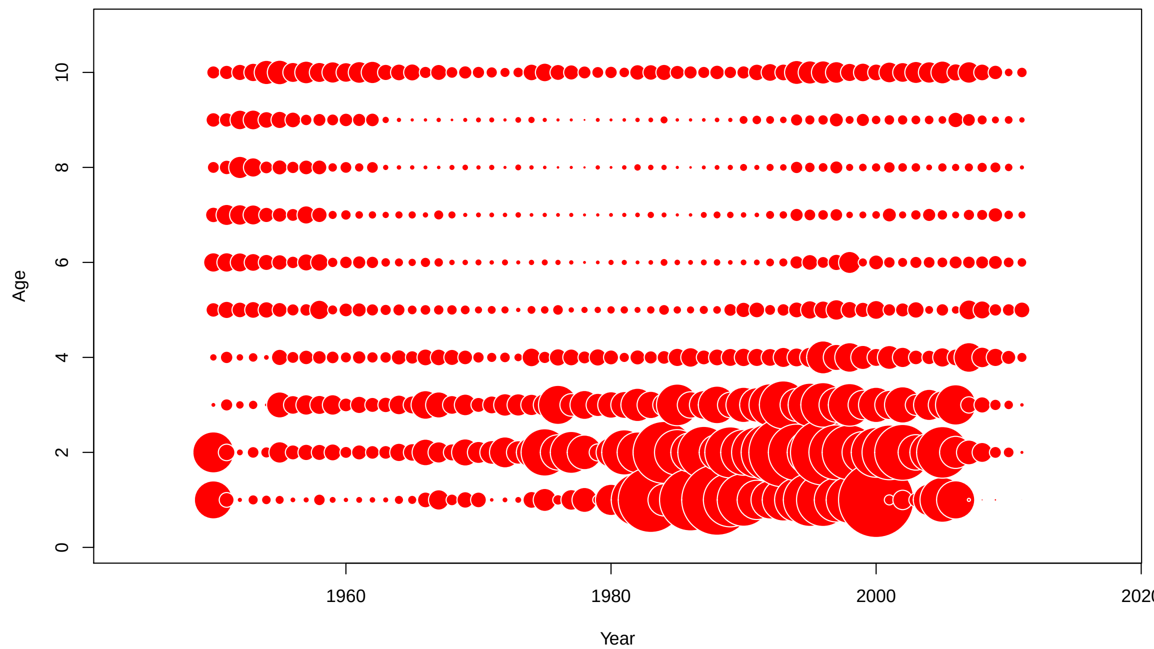


Figure 2: Catch split by age over the whole period. Area of red circles is proportional to the catch

3.2 Available catch

The total catch in numbers are given from 1950 to 2011. Figure 2 illustrates the evolution of the catch across years.

It would appear on figure 2, that there are possibly numerous different periods with very different vulnerability at age, 1950-1955, 1956-1975, 1976-1982 or so, 1983-1990, 1991-1995, and 1996-present. The selectivity should be allowed to vary over time. In order to avoid over-parametrization and to keep the model simple in this first approach, the selectivity has been chosen constant over years.

3.3 Data on Selectivity

Catch at age data are available for the commercial fisheries and for six of the seven abundance indices. It is specified that `NW_PS` is proportional to last age class (more than 10 years old tuna).

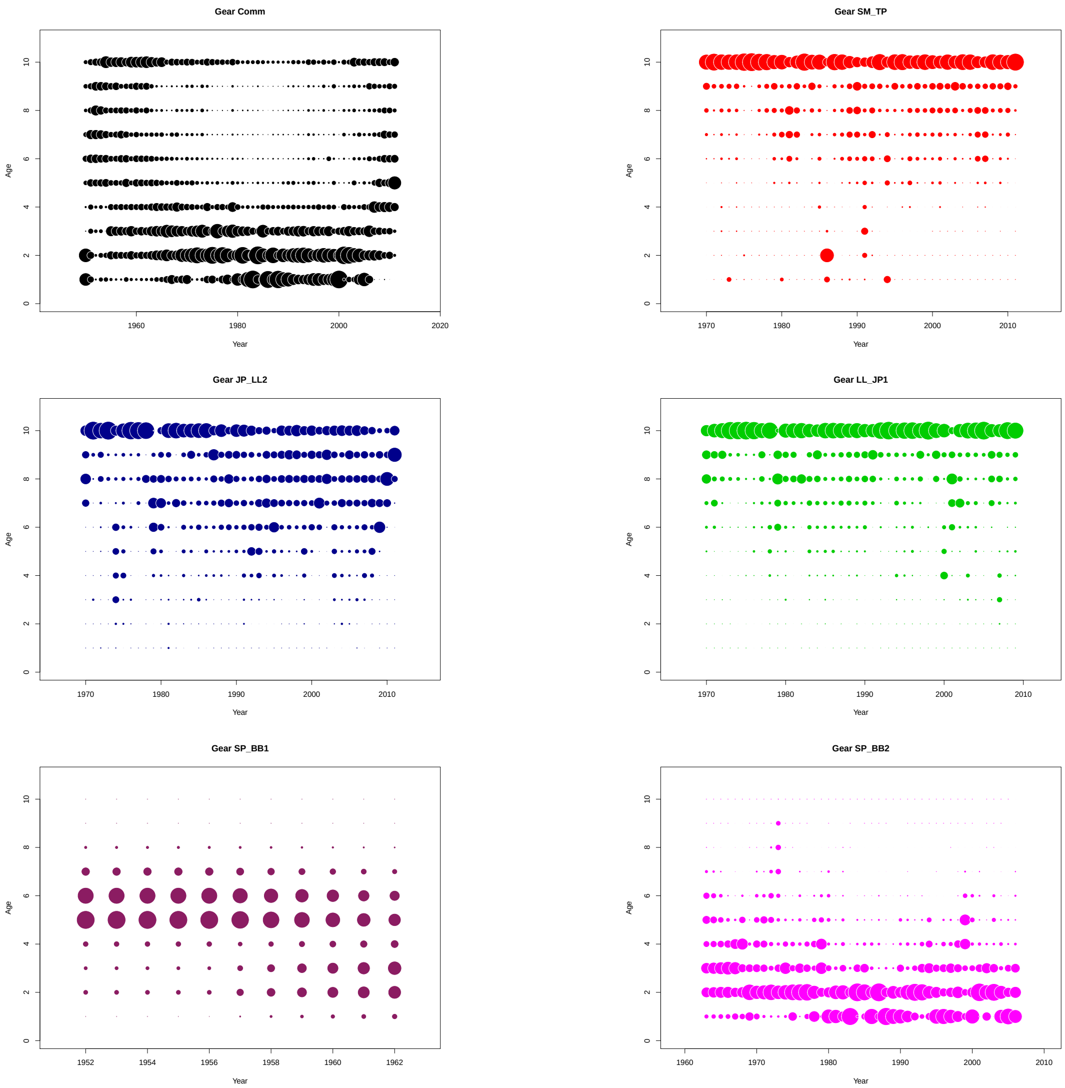


Figure 3: Catch split by age and gear over the whole period. Area of the circles is proportional to the Catch

To investigate the change in selectivity, we can look at composition of the catch per gear and per year

4 Practical aspects of running `iSCAM` on Bluefin Tuna data

The first section aims at clarifying the choices made to run `iSCAM` on BFT data. The second section shows the estimation of the leading parameters in order to justify the simulation framework develop in section 5. The detailed complete results are given in section 6.

4.1 Tunning up `iSCAM` - choices

Preliminary remarks The first age class is 1 year, recruitment at year t depends on mature biomass at year $t-1$, according to [SCRSGroup, 2012]. The Beverton-Holt model is used for recruitment.

To account for the total commercial catch (only one time series) and for the six abundance indices considered, seven gears have been declared. Gear one corresponds to commercial fisheries, gears 2 to 7 correspond to the six abundance indices, each gear having its one vulnerability specification. The fishing mortality is driven by the catch for gear 1 and therefore $F_{k,t} = 0$ for $k \geq 2$. For clarity, index k will be omitted in the sequel since there is no ambiguity.

Mortality rate In [SCRSGroup, 2012], the mortality rate is defined by

	1	2	3	4	5	6	7	8	9	10
Mortality	0.49	0.24	0.24	0.24	0.24	0.20	0.17	0.15	0.12	0.10

But `iSCAM` doesn't allow to use instantaneous mortality at age whereas it is possible to use a time varying mortality, using a random walk centered on a given value. Because of the well known difficulty for estimating the natural mortality rate, it has been decided to use a fixed instantaneous mortality rate. This parameter won't be estimated and is fixed to 0.23.

Selectivity As mentioned in [Martell, 2012], `iSCAM` allows to specify different form for the selectivity/vulnerability. When age composition data are available, choice has been made to model the selectivity using Bspline curves or logistic function. The selectivity has been chosen to stay constant over years. The specific choice between B-splines and logistic curves has been made to assure the convergence of the optimization algorithm and the obtention of a definite positive Hessian matrix.

Gear 4 corresponds to the Norwegian Purse seine abundance index which is documented to focus only on last age class. Since it is not possible to specify 0 vulnerability for some age (due to a log transformation), the selectivity curve has been set to a logistic function with fixed parameters, the parameters being chosen to mimic the desired behavior.

Gear	Gear Name	Selectivity shape
1	Catch	Constant (over years) cubic B-spline with 5 nodes
2	SM_TP	Logistic shape
3	LL_JP1	Logistic shape
4	NW_PS	fixed logistic $\mu_{50} = 9.9$, $sd = 0.1$
5	JP_LL2	Logistic shape
6	SP_BB1	Constant (over years) cubic B-spline with 5 nodes
7	SP_BB2	Constant (over years) cubic B-spline with 5 nodes
8	SP_BB3	Logistic shape

Table 3: Choice for the selectivity curve for each gear

Abundance indices The total vulnerable biomass at year t for gear k is defined by

$$V_{k,t} = \sum_{a=1}^{A} N_{t,a} e^{-\lambda_{k,t} Z_{t,a}} v_{k,a} w_a, \tag{27}$$

where $\lambda_{k,t}$ is the fraction of the mortality to adjust for survey timing, it is specified by the user. They have been specified according to the input data file

Respectively, the total vulnerable number of fish at year t for gear k is defined by

$$V_{k,t}^{N} = \sum_{a=1}^{A} N_{t,a} e^{-\lambda_{k,t} Z_{t,a}} v_{k,a}, \tag{28}$$

The value of $\lambda_{k,t}$ are taken constant over years. Five of them are specified as average indices along the year in the stock assessment, the last one is assumed to occur one month after the start of the year. To mimic those choices, the timing parameters λ are equal to 0.5 for five of the six indices, the last one is set to 1/12. Those choices are sum up in table 4

Gear	Gear Name	Number/Biomass	λ_k: Timing for the survey
2	SM_TP	Number	0.5
3	LL_JP1	Number	0.5
4	NW_PS	Biomass	0.5
5	JP_LL2	Number	1/12
6	SP_BB1	Biomass	0.5
7	SP_BB2	Biomass	0.5

Table 4: Choice for the type and the timing of abundance indices

Recruitment The estimation of the recruitment parameters in `iSCAM` is split into two steps. In early phases, the recruitment is assumed to be stock independent using formula 8. In the last phases, a stock recruitment relationship is used as specified in equation 7.

Life trait history The parameters regarding the life trait history are not estimated and are treated as fixed when running `iSCAM` . Actually, there is some possibility to use weight at age data to estimate the weight at age relationship, but trying to use this option leads to inconsistent results. The values of the concerned parameters have been fixed according to [SCRSGroup, 2012] and are summarized in table 5.

Parameters	Name in `iSCAM`	Set Value	Signification
l_∞		319	vonB parameters
k		0.093	vonB parameters
t_0		-0.97	vonB parameters
a_w	a	$1.95e-05$	Weight at age allometric parameter
b_w	b	3.009	Weight at age allometric parameter
μ_f	$\dot{a}$	4	age for 50% maturity
σ_f	$\dot{\gamma}$	0.8	Standard deviation at 50% maturity
Φ_E		201	incidence function

Table 5: The parameters reported in this table are considered as fixed to the specified value

Initial values for estimated parameters

The leading parameters are given in table 1. `iSCAM` requires to provide initial values to the parameters to start the optimization algorithm. Since the role of the recruitment differs between first phases and the last ones, it is also required to give initial values for R_{init} and $\overline{R}$. The proportion of the total variance allocated to the error in the observation process is not estimated and is set to $\rho = 0.4$.

4.2 Bayesian approach - Prior specification

A Bayesian approach requires to specify adequate prior distributions on leading parameters.

Prior on R_0, R_{init} and $\bar{R}$ The prior on R_0 is build with the idea of a range between 10^5 et 10^6 Tons for the unfished total biomass. Φ_E being the contribution of a recruit to the total biomass during its life, $B_0 = R_0\Phi_E$. Φ_E depends on the weight at age and the survivorship, both being fixed. Therefore R_0 may be derived and the corresponding

Parameters	Initial Value	interval	Phase	Signification
$\log R_0$	13	$[-5,30]$	1	Logarithm of recruitment (numbers) in unfished conditions
h	0.85	$[0.2,0.99]$	3	Steepness
$\log \bar{m}$	-1.47	$[-5,0]$	-1	Logarithm of the average natural mortality rate
$\bar{R}$	12.5	$[-5,20]$	1	Logarithm of the average recruitment
R_{init}	12.5	$[-5,20]$	1	Logarithm of the initial recruitment
ρ	0.4	$[0.001,0.999]$	-1	proportion of the total variance for the observation process
$1/\varphi$	0.8	$[0.001,12]$	3	the root of the precision of the total error

Table 6: Initial values and range for main parameters. The estimation phase is also given.

value for the logarithm of R_0 is in the interval $[12.9, 15.2$. Relaxing this assumption, the prior has been specified as in equation 29.

$$\log(R_0) \sim \mathcal{U}[12,16] \tag{29}$$

The same priors have been chosen for R_{init} and $\bar{R}$.

Prior for steepness h A beta distribution has been chosen as a prior for the steepness. The parameters of the beta distribution has been chosen so that the mode of the steepness is 0.9 and the coefficient of variation is 10%. This consideration leads to the specification in 30.

$$h \sim \beta\,(14, 2.44) \tag{30}$$

Prior for the total variance φ^2 being the total variance, `iSCAM` allows to specify a prior distribution on $1/\varphi^2$ using classically a gamma distribution. The parameters of this gamma distribution has been chosen so that to define a vague prior.

$$1/\varphi \sim \Gamma\,(0.1, 0.1) \tag{31}$$

4.3 First run

To illustrate the results obtained on BFT East Stock, we focus on one datafile:
`Inputs/bfte/2012/vpa/inflated/high/bfte2012.d1`

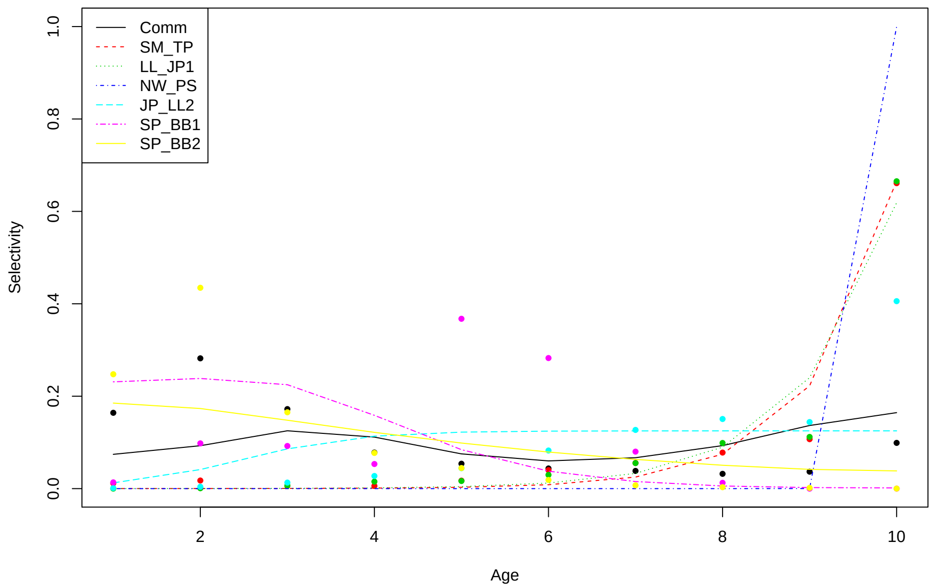


Figure 4: Estimated selectivity. The additional points represents the average empirical proportion at ages

Figure 4 seems to present a good coherence between the empirical proportion at ages and the estimated vulnerability. Unfortunately figure 5 highlights the large variability around the average behavior.

This first run leads to the Kobe plot presented in figure 6.

Selectivity The vulnerability matrix is estimated to

$$V = (v_{k,a}) = \begin{pmatrix} 0.063 & 0.085 & 0.134 & 0.132 & 0.092 & 0.074 & 0.079 & 0.097 & 0.117 & 0.127 \\ 0.000 & 0.000 & 0.001 & 0.002 & 0.006 & 0.015 & 0.041 & 0.108 & 0.264 & 0.563 \\ 0.000 & 0.000 & 0.001 & 0.002 & 0.006 & 0.018 & 0.053 & 0.137 & 0.297 & 0.486 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.000 \\ 0.010 & 0.038 & 0.086 & 0.115 & 0.123 & 0.125 & 0.125 & 0.126 & 0.126 & 0.126 \\ 0.205 & 0.221 & 0.229 & 0.175 & 0.098 & 0.045 & 0.018 & 0.006 & 0.002 & 0.001 \\ 0.224 & 0.204 & 0.162 & 0.122 & 0.090 & 0.066 & 0.048 & 0.035 & 0.026 & 0.023 \end{pmatrix} \tag{R1}$$

Parameters	Initial value	Reported value
$\log R_0$	13	14.75
h	0.85	0.934
$log(\bar{R})$	12.5	14.38
R_{init}	12.5	15.22
ρ	0.4	0.4
$1/\varphi$	1.25	1.09
$log(B_msy)$		18.81
(f_msy)		0.14
$log(B_{2011})$		19.72
(f_{2011})		0.037

Table 7: Initial values and range for main parameters.

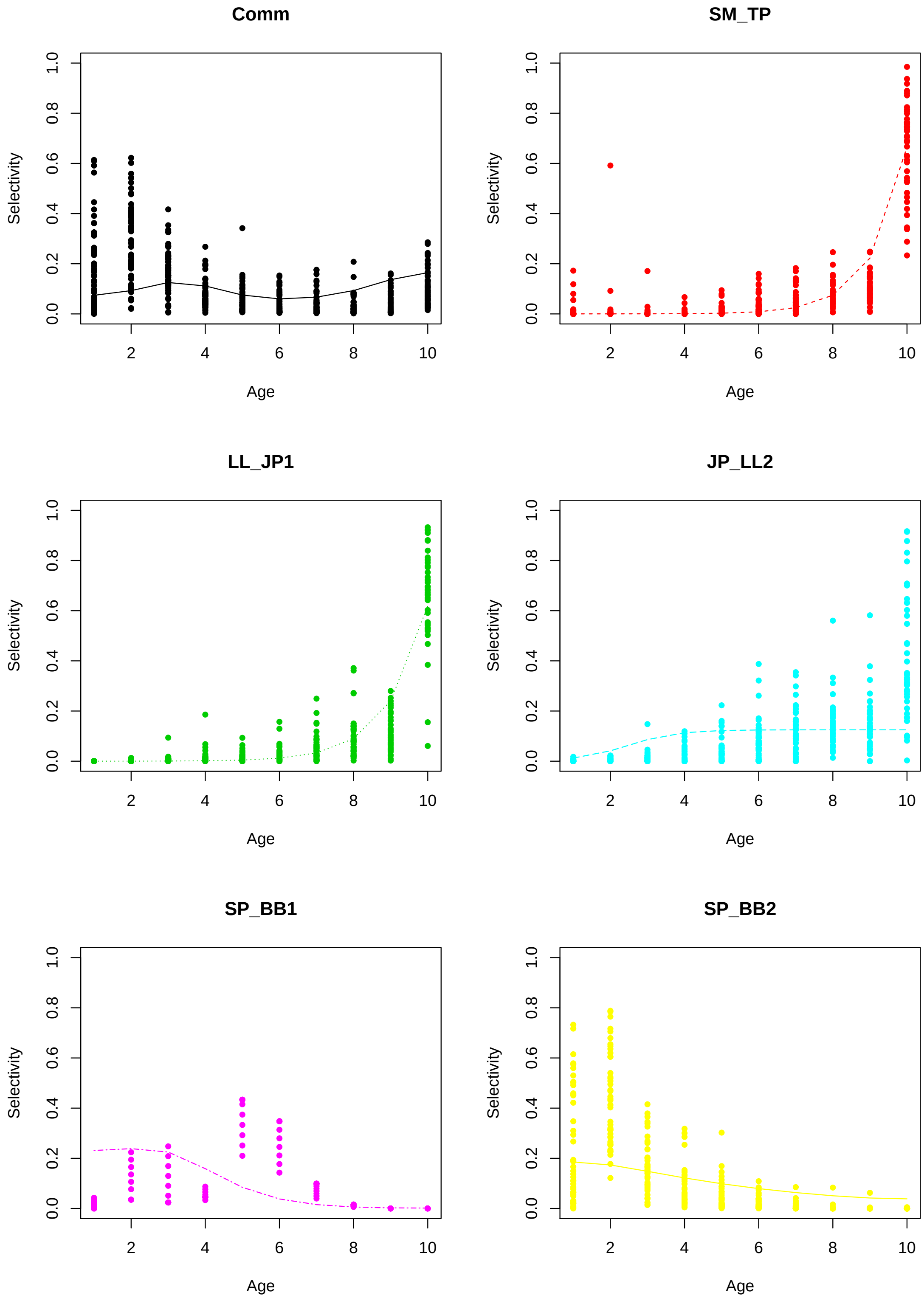


Figure 5: Estimated selectivity and the empirical proportion at ages.

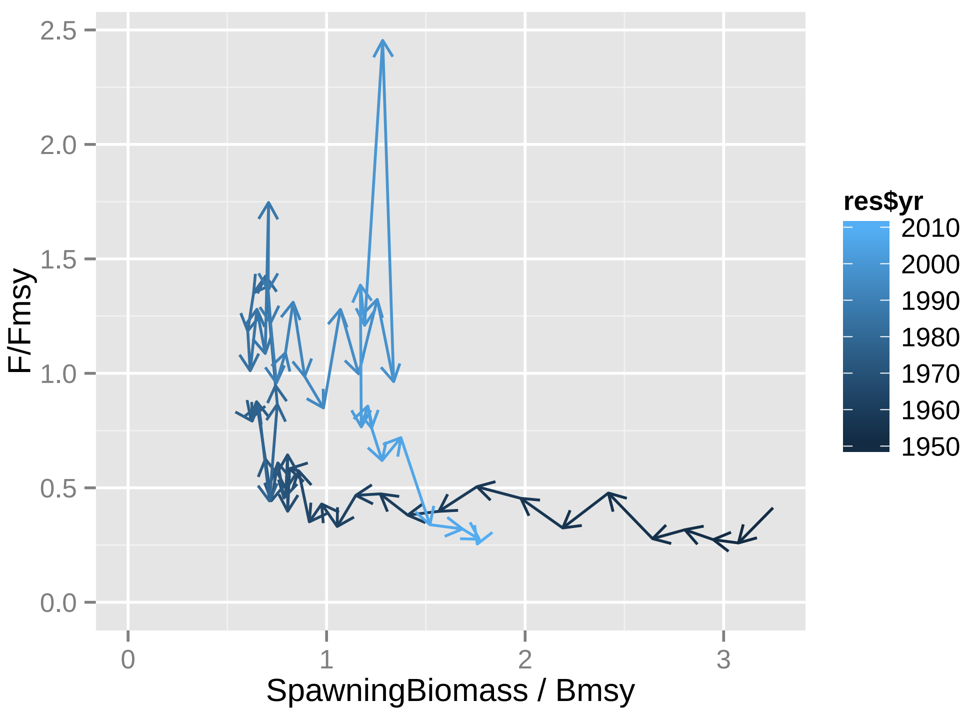


Figure 6: Evolution of the fishing effort and the stock status.

5 Simulations study

Simulations have been run to check the ability of `iSCAM` at estimating the parameters. The values of the parameters are chosen to be realistic in regard of the parameter estimates on BFTE data.

5.1 Simulation model

`iSCAM` proposes a simulation model but in order to fully control the simulation framework (especially the fishing mortality), a `R` code has been developed to simulate a stock dynamic according to the model specified in `iSCAM` . In `iSCAM` some realistic fishing mortality rate is derived from the true catch and then used to simulate new catch observation. To avoid this back and forth procedure, a simpler version is developed: the fishing mortality is taken constant over year, the value being chosen to avoid stock extinction.

5.2 Simulation framework

The purpose of this section is to investigate the possibility for `iSCAM` to reestimate the actual parameters. The value of the parameters have been chosen to be similar to the estimates obtained for the first run. But the strength of the information contained in the data is obviously a key point, therefore in order to investigate the amount of noise `iSCAM` is able to handle, nine different scenario have been studied from a very precise one up to a very poorly informative situation..

The total variance observed on the BFTE data is estimated to 1.16 and the proportion of the process error is set to 0.4 which corresponds to 0.68 for the standard deviation of the observation errors and 0.83 for the process errors. This situation is worse than scenario 9 in terms of information in the data. This situation is investigated in scenario 10.

The remaining parameters used for this simulation study are specified in 9.

For every scenario 600 datasets have been drawn and the parameters have been reestimated. When `iSCAM` produces a non definite positive Hessian matrix, the simulation is dropped, but this case never happens within our simulation framework.

5.3 Results on leading parameters

5.4 Comparison of the results for the different scenarios

The results on the different simulation studies for the nine first scenarios considered prove that the model and the estimation procedure produce reasonably good results when the total variance is low. As expected, increasing the total variance increases the variability of the estimates but more surprisingly also produces bias estimates.

When the total variance increases, the recruitment in unfished conditions tends to be overestimated. Conversely the initial recruitment is underestimated. The steepness also

Scenario	Process error	Obs. error	τ_R	τ_I	Total Precision	Portion of variance ρ
1	low	low	0.1	0.1	50	0.5
2	low	med	0.1	0.2	20	0.8
3	low	high	0.1	0.5	3.84	0.96
4	med	low	0.2	0.1	20	0.2
5	med	med	0.2	0.2	12.5	0.5
6	med	high	0.2	0.5	3.45	0.86
7	high	low	0.5	0.1	3.84	0.04
8	high	med	0.5	0.2	3.45	0.14
9	high	high	0.5	0.5	2	0.5
10	BFTE	BFTE	0.83	0.68	0.86	0.4

Table 8: Summary of the considered scenarios in term of variance errors for the simulation

tends to be overestimated when the data are too noisy. The consequence of such behavior is to over estimate the resilience of the stock and to produce higher values for B_{MSY}.

The figures are not presented here, but there is also a systematic slight positive bias in the estimation of the total variance and therefore τ_I and τ_R tend to be slightly overestimated.

Focus on Scenario 10 - BFTE like parameters value

This section focus on the results obtained under scenario 10, chosen to be comparable to the BFTE datasets.

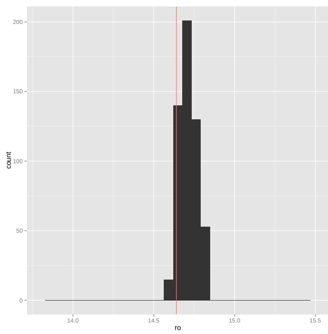
Table 10 sums up the reestimated values through the obtained quantiles.The obtained distribution show heavy tails but removing only the 2 most extremal percents of the estimations produce better estimates.

The figures 9 illustrates the bias in the reestimation. There is an important bias on the initial recruitment and on the steepness. The total variance is also overestimated.

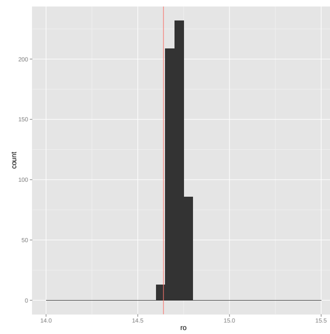
5.5 Effect of parameter τ_A

The proportion at age are assumed to follow a multivariate logistic distribution which has been shown [Schnute and Richards, 1995] to be more flexible than the classical multinomial distribution. The τ_A parameter of this distribution is not really intuitive. It is important to have some clues on how the expected proportion at age are closed to the empirical proportions at age according to this parameter. A very simple simulation has been used to investigate this relationship. Using notation presents in 23, figure 10 presents the largest absolute difference between the expected proportion μ_a and the observed proportion for different value of τ_A. Because μ_a are probabilities, the difference is bounded by 1.

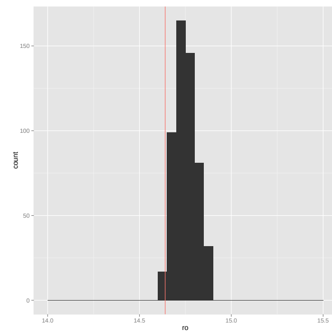
Only very small values of τ_a produces precise realization of the multivariate logistic distribution.



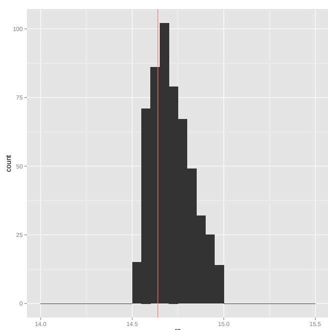
(a) Scenario 1



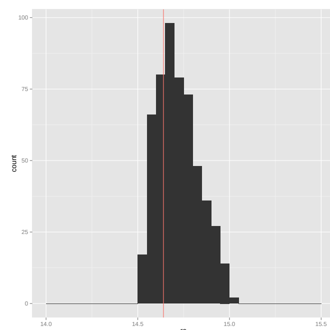
(b) Scenario 2



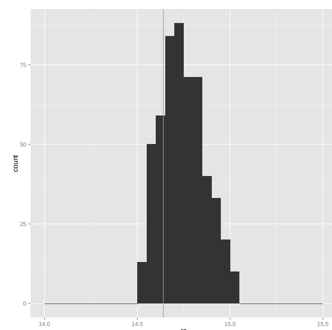
(c) Scenario 3



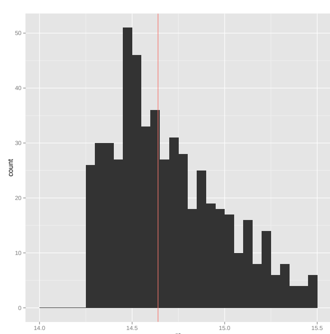
(d) Scenario 4



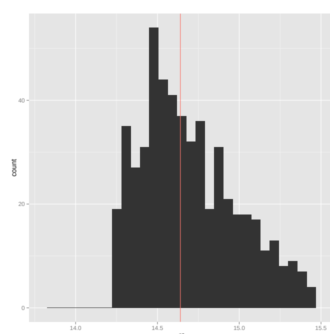
(e) Scenario 5



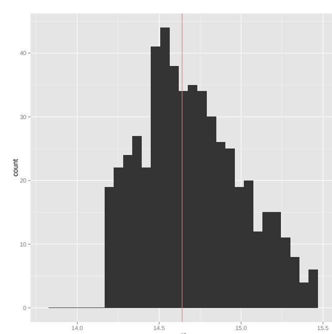
(f) Scenario 6



(g) Scenario 7

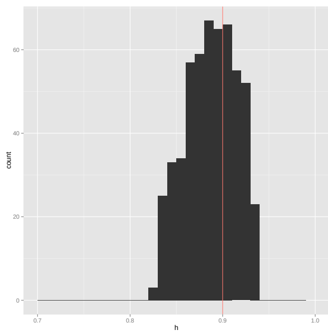


(h) Scenario 8

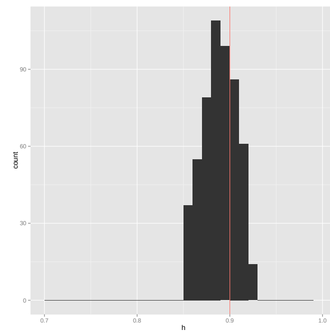


(i) Scenario 9

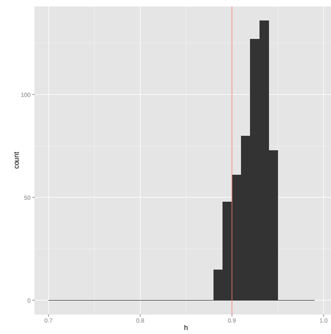
Figure 7: Reestimation for R_0 for the 9 considered scenarios. The red line represents the value used for the simulation.



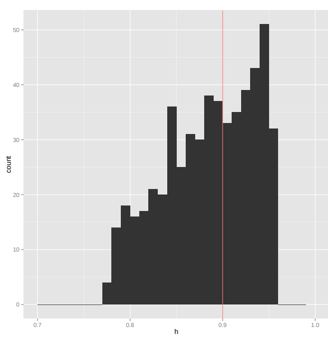
(a) Scenario 1



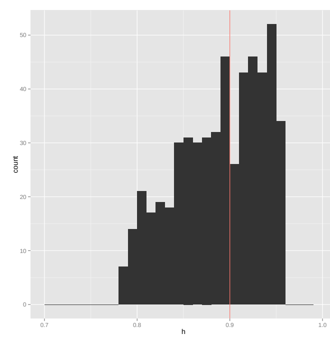
(b) Scenario 2



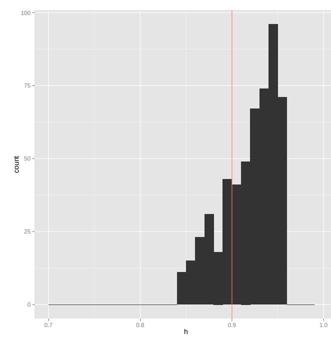
(c) Scenario 3



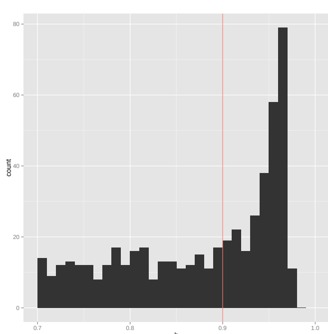
(d) Scenario 4



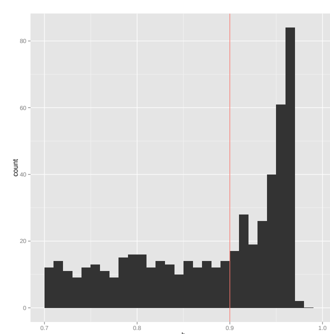
(e) Scenario 5



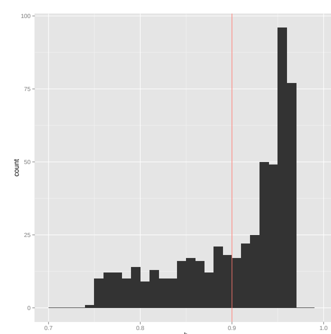
(f) Scenario 6



(g) Scenario 7

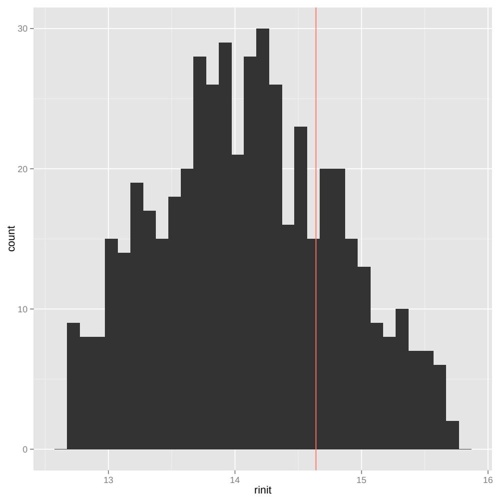


(h) Scenario 8

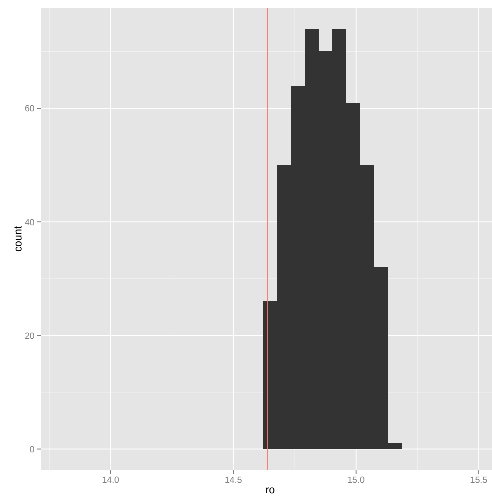


(i) Scenario 9

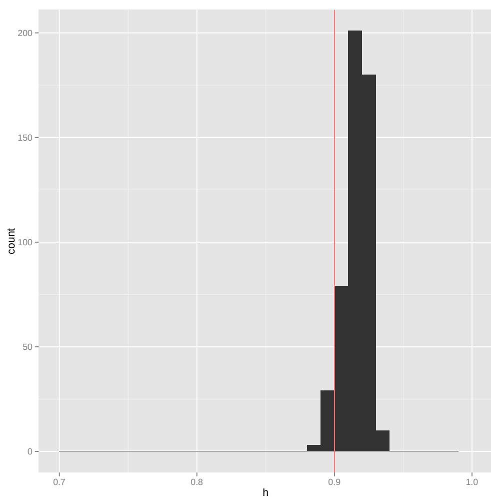
Figure 8: Reestimation for the steepness h for the 9 considered scenarios. The red line represents the value used for the simulation.



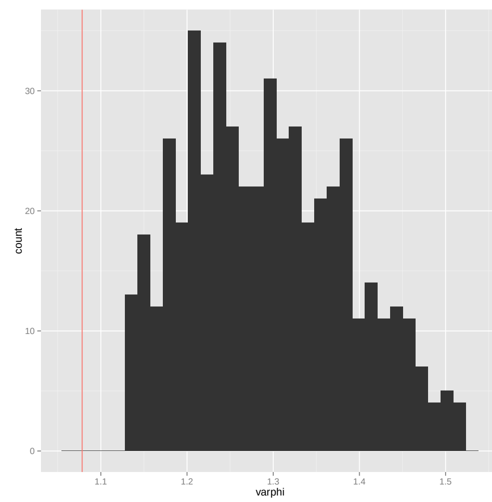
(a) Histogram of the reestimation of the initial recruitment $log(R_{init})$



(b) Histogram of the reestimation of unfished recruitment $log(R_0)$



(c) Histogram of the reestimation of the steepness h



(d) Histogram of the reestimation of the total variance φ

Figure 9: Summary of the reestimated values. The red lines is the true value used for the simulation. The 5 most extremal percents have been dropped.

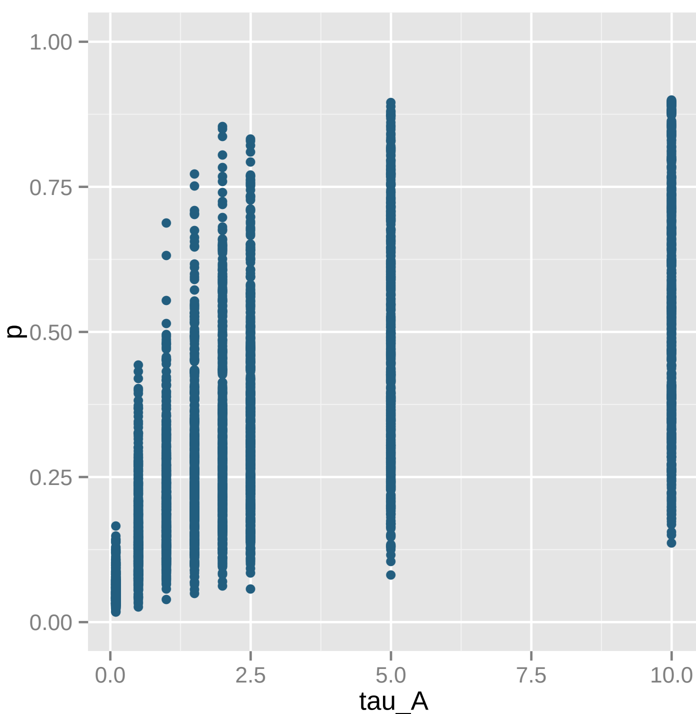


Figure 10: Largest absolute difference between expected proportion at age p_a and the realization of a multivariate logistic distribution $MV(p_a, \tau_A)$ centered on p_a with standard deviation τ_A.

Name	Value	Notes
	Parameters be estimated	
R_0	$exp(14.64)$	
h	0.9	
τ_C	0.1	This parameter is fixed in `iSCAM` not estimated
τ_A	0.1	
$v_{k,a}$	see R1	the selectivity table is taken according to the results obtained on BFT
q_k	10^{-5}	Same value for all gear
Other quantities of interest (fixed parameters or derived from parameters)		
f	0.2	the instant fishing mortality, chosen to be constant over year
Φ_E	317	
M_a	$(0.490, 0.24, 0.24, 0.24, 0.240,$ $0.20, 0.175, 0.15, 0.125, 0.10)$	
m_{50}	4	
σ_{50}	0.8	
s_0	0.113	fixed according to R_0 and h using $s_0 = \kappa/\Phi_E$ and $\kappa = 4h/(1-h)$
β	$4.84\,10^{-08}$	fixed using $\beta = (\kappa - 1)/(R_0 * \phi_E)$

Table 9: Leading parameters

5.6 Conclusion on simulation results

The estimation procedure proposed in `iSCAM` behaves reasonably well for small standard deviation in the abundance indices but produces highly biased estimates (and not only inaccurate estimates) when the quality of the abundance indices decrease. Especially it tends to overestimate the steepness and underestimate the initial recruitment, leading to optimistic predictions on the resilience of the stock.

Unfortunately, `iSCAM` estimates a high level of uncertainty on the BFTE abundance indices, which mean that the estimates and the prediction have to be carefully considered since the abundance indices seem to be very inaccurate.

6 Complete results on Bluefin tuna data - East stock

6.1 Maximum likelihood approach

This section reports the Maximum likelihood estimates obtained using `iSCAM` . Table 11 presents the value and if available reports the values presented in [SCRSGroup, 2012]. The presented table assumes that the initial recruitment may be different form the recruitment in unfished condition.

In [SCRSGroup, 2012] the logarithm of the initial recruitment is reported to vary be-

	$log(R_0)$	$\log(R_{init})$	φ	τ_I	τ_R	$\log MSY$	$Fmsy$	$\log Bmsy$
0%	14.43	9.09	0.94	0.59	0.73	16.89	0.08	18.45
1%	14.54	12.13	1.01	0.64	0.78	17.00	0.13	18.56
2.5%	14.59	12.50	1.03	0.65	0.80	17.05	0.13	18.60
5%	14.63	12.67	1.05	0.67	0.82	17.09	0.14	18.64
50%	14.84	14.04	1.20	0.76	0.93	17.30	0.15	18.86
95%	15.09	15.98	1.39	0.88	1.07	17.54	0.15	19.13
97.5%	15.19	17.73	1.45	0.92	1.13	17.64	0.15	19.21
99%	15.28	18.68	1.51	0.96	1.17	17.70	0.15	19.32
100%	29.99	19.99	1.67	1.05	1.29	32.08	0.16	34.31

Table 10: Empirical quantiles obtained by reestimation on simulated data for the main parameters.MSY and B_{msy} are expressed in Kg, R_0 and R_{init} are expressed in numbers.

tween 14.5 and 15.6 depending on the scenarios considered. The logarithm of the minimum recruitment is evaluated to 13.8. The spawning biomass is reported to range from 150 000 tons to 300 000 tons and the current biomass is evaluated to 0.96 of the maximal spawning biomass.

As already mentioned, `iSCAM` seems to estimate similar recruitment but to produce different spawning biomass. This may arise from the difference in the natural mortality rate. The current `iSCAM` version doesn't allow to specify mortality rate at age and therefore the mortality parameter has been chosen constant over ages and fixed to an arbitrary value of 0.23. This may explain the difference observed between the two approaches. This development of `iSCAM` is in progress and should be shortly available.

Figure 12 exhibits important changes in the recruitments with very high recruitment levels in the 1990's and a decrease during the last years. The very last years have been dropped due to well known difficulties to estimate the last recruitments.

Abundance indices

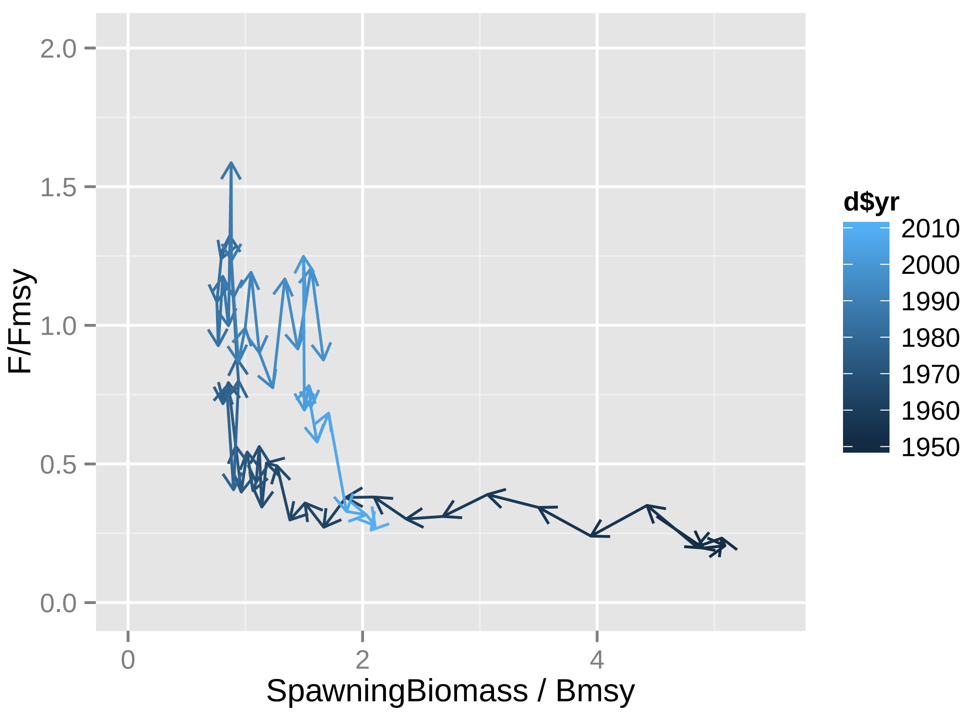
To diagnose the fit to the data, one may compare the expected abundance indices derived from the results to the observed abundance indices. As the value τ_I is high it is not surprising to denote a lack of fit between the observed abundance indices and their expected values, as presented on figure 13. This behavior may be improved by giving different weight to each index, according to its quality. This option is available in `iSCAM` .

Observed Catch

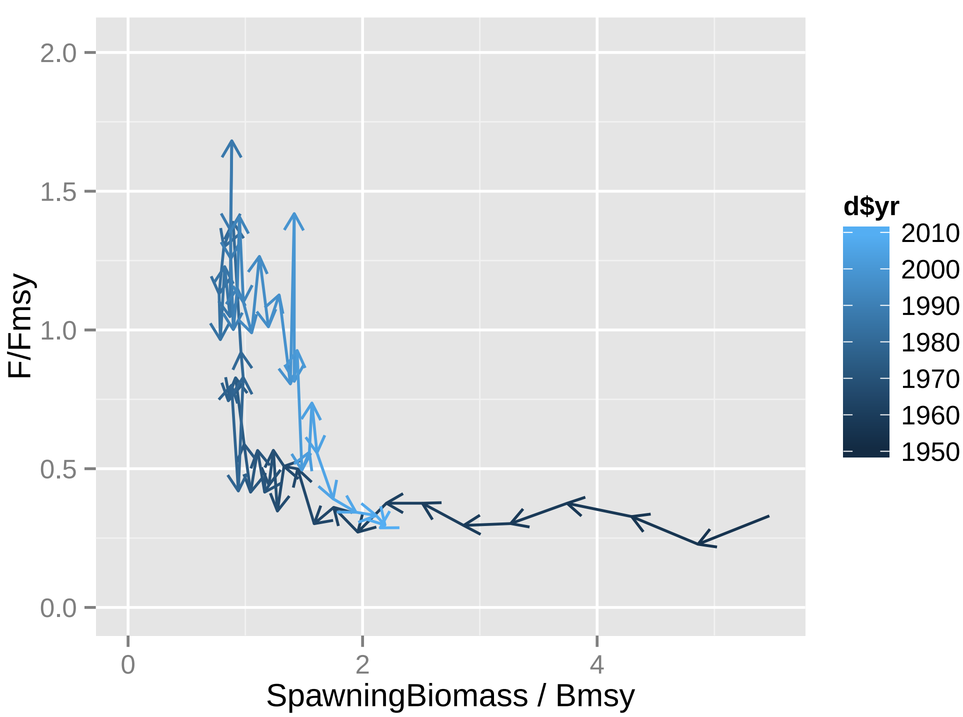
Another diagnostic would be to compare the observed catch and the expected catch. There is a very small differences between the two series.

Parameters	Inflated - $R_{init} \neq R_0$		Reported $R_{init} \neq R_0$		ICCAT 2012 report
$\log R_0$	14.75	$[14.57, 15.00]$	14.65	$[14.48, 14.88]$	
$\log(R_{init})$	15.22	$[14.04, 16.20]$	15.31	$[14.05, 16.37]$	$[14.5, 15.6]$
$\log R_min$	$log(R_{2007}) = 13.0$		$log(R_{2007}) = 12.9$		13.8
$log(\bar{R})$	14.38	$[13.97, 14.81]$	14.28	$[13.88, 14.71]$	
h	0.934	$[0.82, 0.97]$	0.937	$[0.82, 0.97]$	
ρ	0.4		0.4		
$1/\varphi$	1.09	$[1.06, 1.52]$	1.09	$[1.01, 1.5]$	
$log(B_msy)$	18.81	$[18.66, 19.11]$	18.68	$[18.53, 18.99]$	
(f_msy)	0.144	$[0.12, 0.15]$	0.146	$[0.12, 0.15]$	
$log(B_{2011})$	19.72	$[19.24, 19.89]$	19.62	$[19.15, 19.84]$	
SSB_{2011}/SSB_{max}	0.4		0.33		
(f_{2011})	0.037	$[0.022, 0.06]$	0.037	$[0.024, 0.07]$	
	Inflated - $R_{init} = R_0$		Reported $R_{init} = R_0$		
$\log R_0$	14.89	$[14.76, 15.17]$	14.86	$[14.69, 15.09]$	
$\log(R_{init})$	NA		NA		
$\log R_{min}$	13.15		13.05		
$log(\bar{R})$	14.35	$[13.98, 14.96]$	14.24	$[13.87, 14.83]$	
h	0.92	$[0.75, 0.95]$	0.91	$[0.78, 0.96]$	
ρ	0.4		0.4		
$1/\varphi$	1.08	$[0.97, 1.48]$	1.08	$[0.94, 1.47]$	
$log(B_msy)$	18.98	$[18.85, 19.35]$	18.94	$[18.77, 19.23]$	
(f_msy)	0.14	$[0.11, 0.15]$	0.14	$[0.11, 0.15]$	
$log(B_{2011})$	19.72	$[19.29, 20.06]$	19.63	$[19.21, 19.84]$	
SSB_{2011}/SSB_{max}	0.54		0.51		
(f_{2011})	0.036	$[0.019, 0.058]$	0.039	$[0.025, 0.059]$	

Table 11: Estimated values for the main parameters. The 90% credibility interval specified are obtained within the Bayesian framework

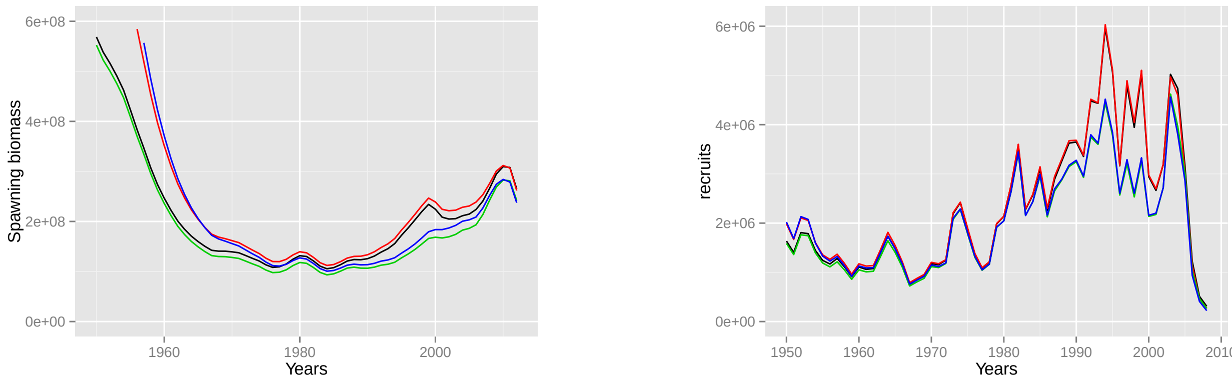


(a) Inflated Catch



(b) Reported Catch

Figure 11: Kobe plot for the two scenarios



(a) Spawning biomass evolution

(b) Recruitment evolution

Figure 12: The evolution of the spawning biomass and of the number of recruits. The black line (resp the red line) corresponds to the inflated catch assuming $R_{init} = R_0$ (resp $R_{init} \neq R_0$). The green line (resp the blue line) corresponds to the reported catch assuming $R_{init} = R_0$ (resp $R_{init} \neq R_0$).

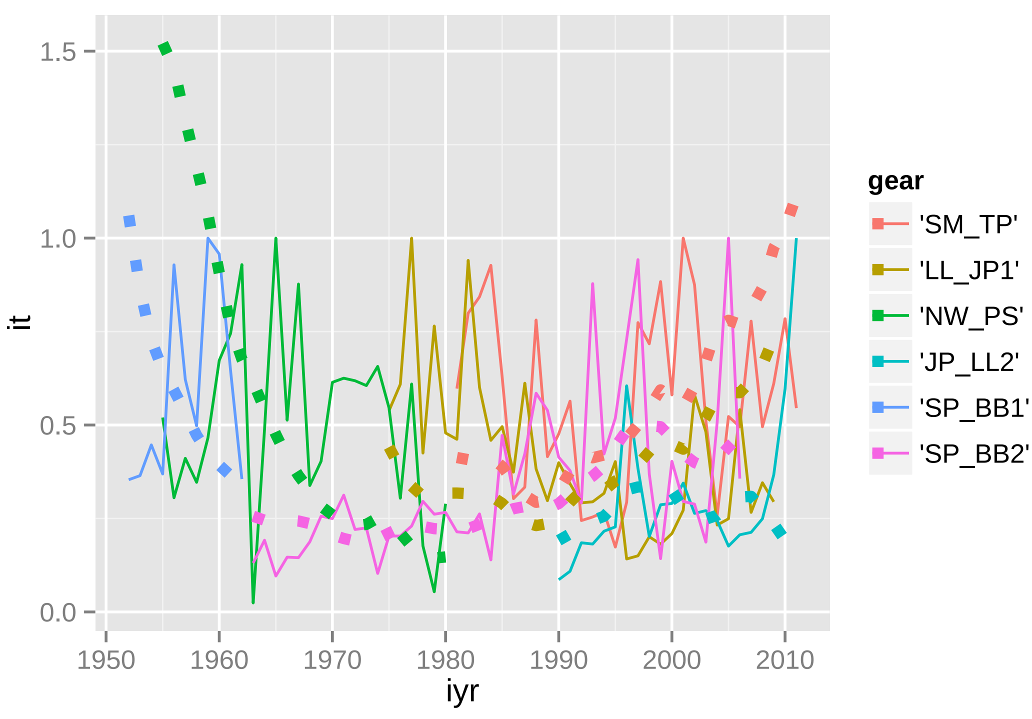


Figure 13: Scaled abundance indices. The observed value are given by the plain lines, the predicted values by the dotted ones. The values have been scaled to stay between 0 and 1.

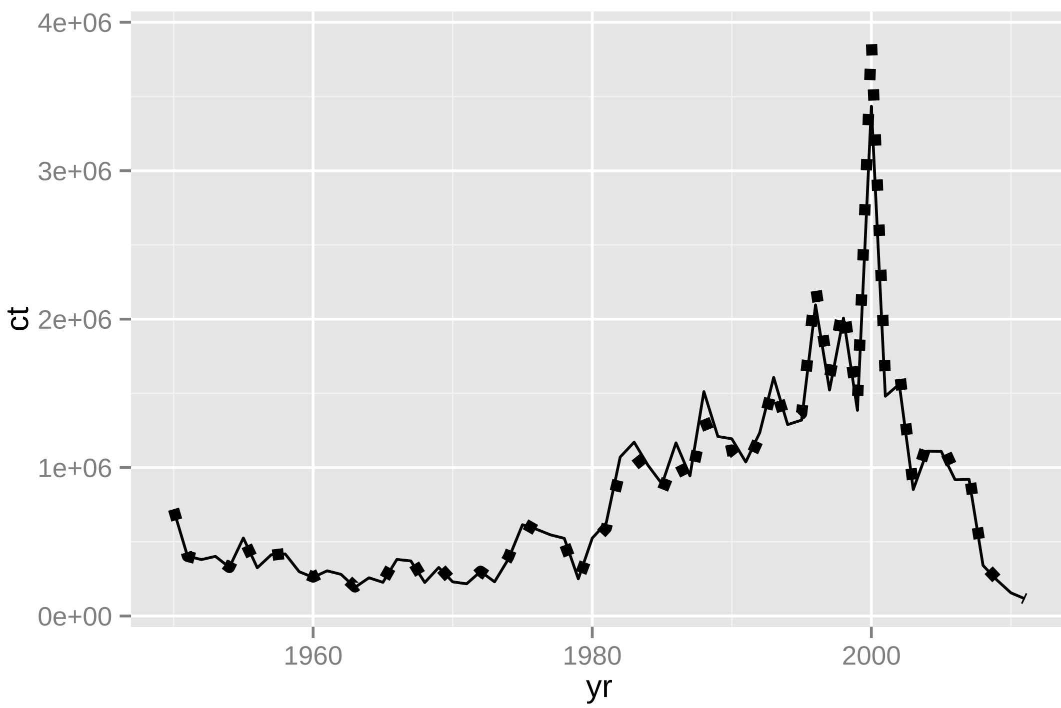


Figure 14: Catch series in numbers. The observed value are given by the plain lines, the predicted values by the dotted ones. The values have been scaled to stay between 0 and 1.

6.2 Retrospective analysis

Retrospective analyses have been run on the two dataset corresponding to inflated catch or reported catch, dropping up to 8 years. Figure 15 presents the evolution of the total biomass. this figure is consistent with the last stock assessment. When estimating the stock level using data until 2003, the trend is a constant decrease over the last years. The last years (between 2003 and 2010) explain the recovery of the stocks. The current trend seems to prove an increase of the spawning biomass, whatever scenario is used.

6.3 Bayesian analysis

A Bayesian approach is possible with `iSCAM` . 50000 iterations of a Metropolis Hastings algorithm have been performed. A burn in period of 25000 iterations is used, one iteration every twenty being saved.

Convergence The convergence of the algorithm has been visually checked using figure 16.

Stock Status The posterior density for the leading parameters are illustrated in figures 17 to 19, the prior have been defined in equation 29 to 31.

The posterior density on R_{init} when R_{init} is assumed to be equal to R_0 are not presented, since R_{init} is not actually defined in this case. (`iSCAM` produces value of R_{init} sampled from the prior distribution). The posterior density of R_{init} is quite broad, the 90% credibility interval correspond to a number of recruits from 1 million to 10 millions. While the corresponding 90% interval for R_0 gives a range of million and up to 3.2 millions of recruits at the unfished state. Every scenario exhibits a very high level of recruitment between 1990 and 2005. This exceptional level has also been reported in [SCRSGroup, 2012].

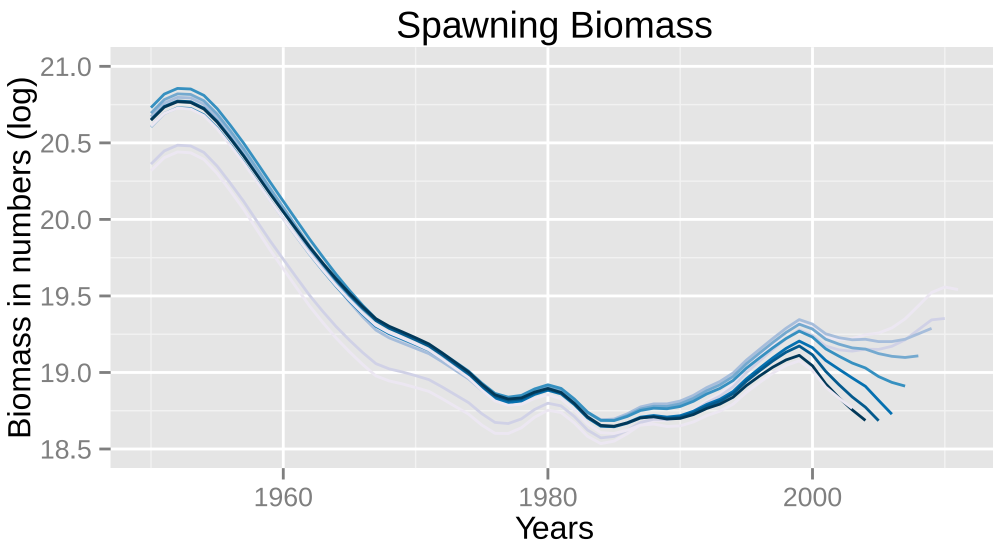
Whatever scenario is used, the posterior distribution on the steepness is vague and the posterior is not really different of the prior distribution. The difficulty of estimating the steepness is well known. This parameter is also a key parameter in the dynamic of the stock.

One of the quantity of main interest is the current stock status. Figure 20 presents the uncertainty for the state of the stock in 2011. For both scenario (reported or inflated catch).

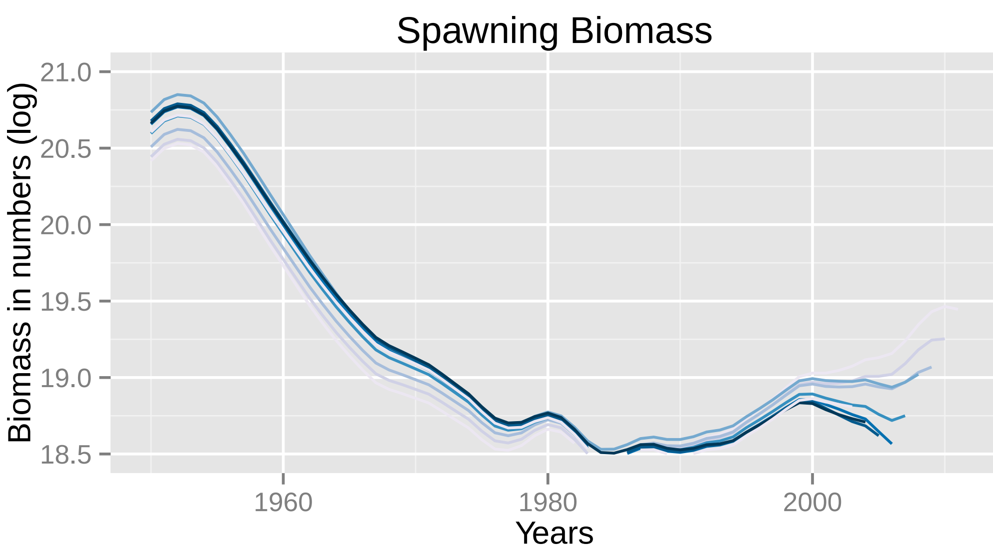
7 Discussion

The results obtained with `iSCAM` are consistent with the results described in [SCRSGroup, 2012] and obtained using a virtual population analysis. Furthermore, as `iSCAM` used the stock recruitment relationship in the last phase of estimation, the recruitment may be almost freely estimated as in a VPA approach.

Some of the parameters have to be fixed and can't be estimated with the data available for this study. `iSCAM` may be more finely tuned using more expertise on the studied

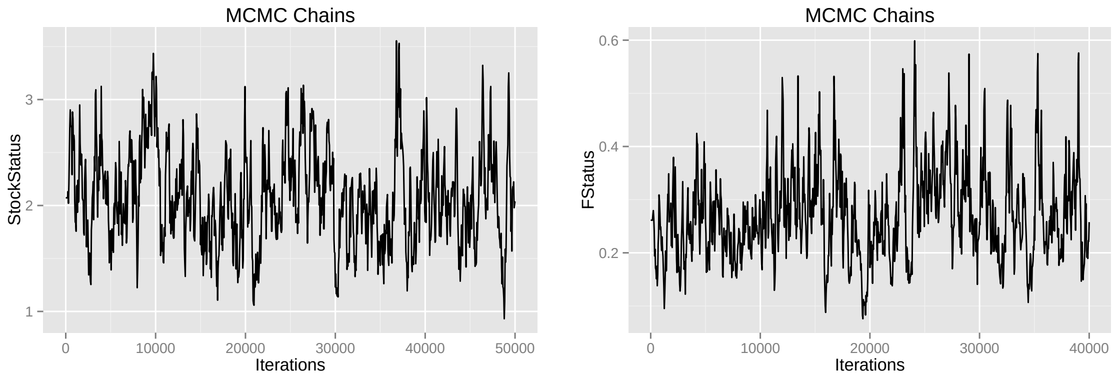


(a) Inflated Catch



(b) Reported Catch

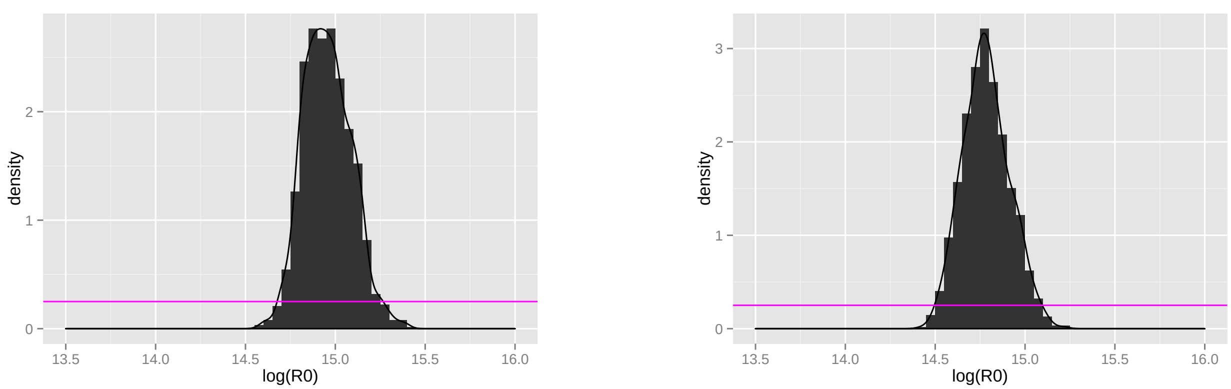
Figure 15: Retrospective analysis, dropping 1 to 8 years



(a) Stock status at the last year

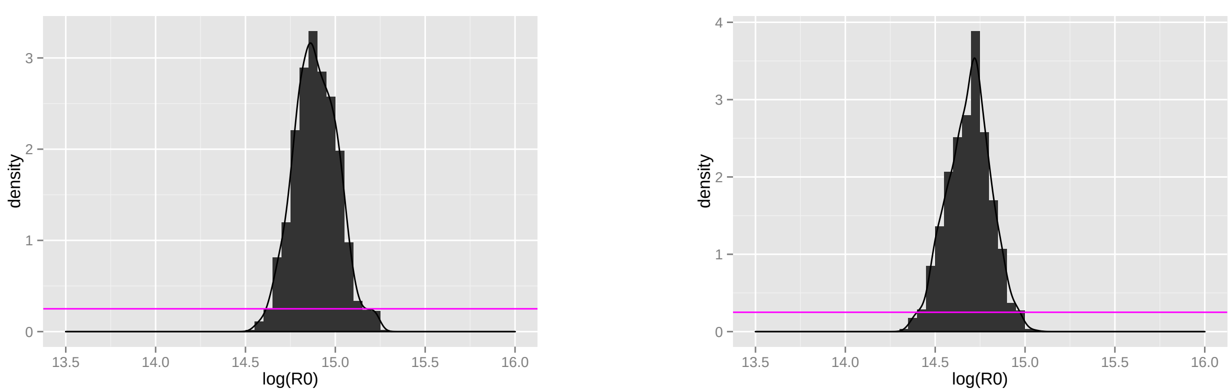
(b) Fishing status at the last year

Figure 16: Graph for MCMC chains



(a) Inflated Catch - $R_{init} = R_0$

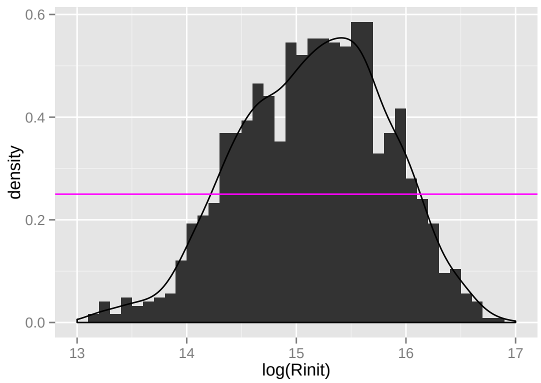
(b) Inflated Catch - $R_{init} \neq R_0$



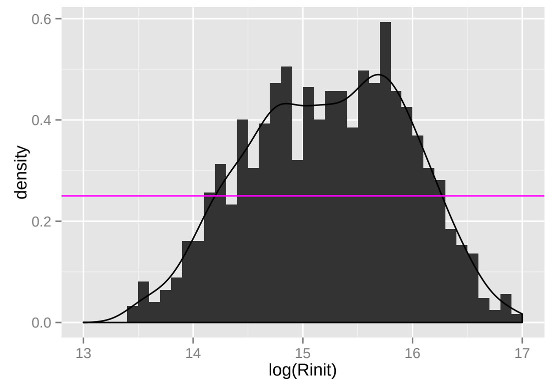
(c) Reported Catch - $R_{init} = R_0$

(d) Reported Catch - $R_{init} \neq R_0$

Figure 17: Posterior distribution for the recruitment in unfished conditions R_0. The purple line represents the prior distribution.

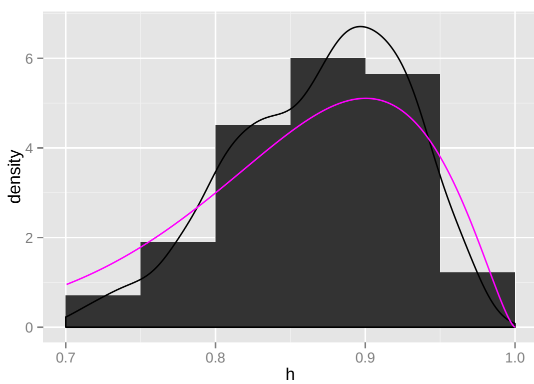


(a) Inflated Catch - $R_{init} \neq R_0$

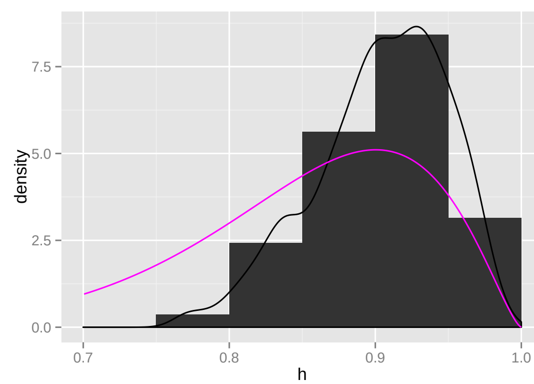


(b) Reported Catch - $R_{init} \neq R_0$

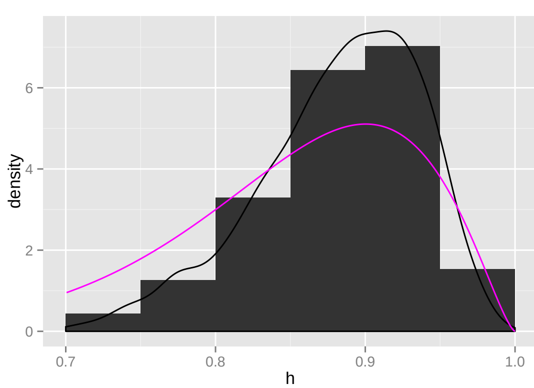
Figure 18: Posterior distribution for the initial recruitment R_{init}. The purple line represents the prior distribution.



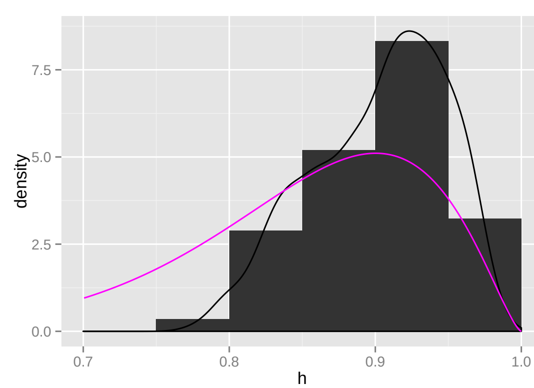
(a) Inflated Catch - $R_{init} = R_0$



(b) Inflated Catch - $R_{init} \neq R_0$

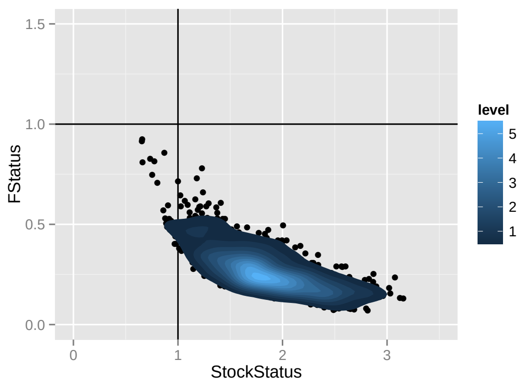


(c) Reported Catch - $R_{init} = R_0$

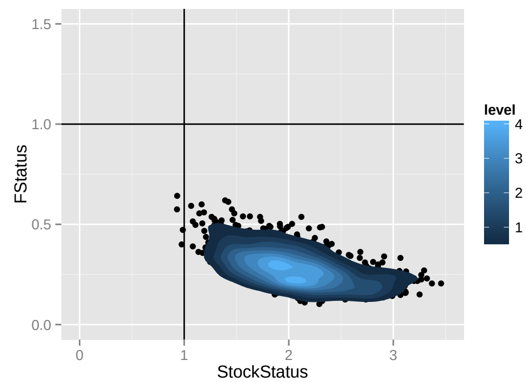


(d) Reported Catch - $R_{init} \neq R_0$

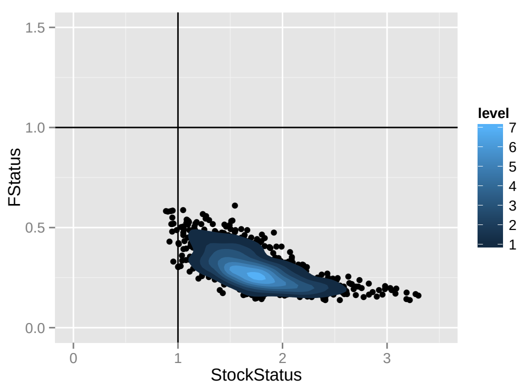
Figure 19: Posterior distribution for the steepness, h. The purple line represents the prior distribution.



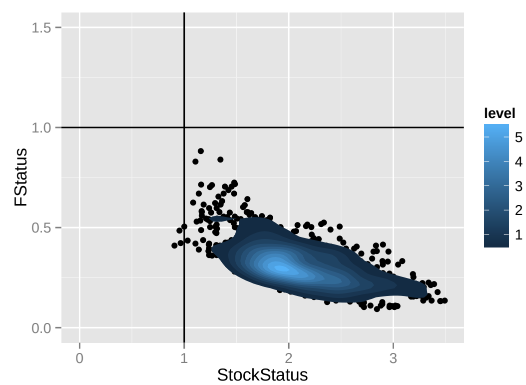
(a) Inflated Catch - $R_{init} = R_0$



(b) Inflated Catch - $R_{init} \neq R_0$



(c) Reported Catch - $R_{init} = R_0$



(d) Reported Catch - $R_{init} \neq R_0$

Figure 20: Posterior Kobe plot, assuming $R_{init} = R_0$, or not, considering both inflated or reported catch.

stock; this expertise may be included as a prior for example but also uses to propose more adequate choices. By instance, the catch at age data show strong pattern, this may correspond to a known change in the selectivity and should be included in the model.

On the other hand, the mortality rate is a key parameter and has been chosen constant throughout ages, but recent development of `iSCAM` should allow to use a mortality rate varying with age and produce results closer from the previous stock assessment. `iSCAM` also proposes some prediction of the evolution of the stock according to fixed TAC policy. The code using this possibility of `iSCAM` is also available.

Running `iSCAM` on BFTE will require more precise tunning of the parameters. The improvement obtained with a fine tunning will not circumvent the issue of the amount of noise (estimated by `iSCAM`) in the data. This high level of noise in the data may produce biased estimates on the key parameters and in any case this variability in the recruitment process and in the observed abundance indices make the stock assessment difficult. However this variability can not be omitted and has to be considered. Therefore, using a Bayesian framework, would allow to include prior information and to produce predictive interval which account for both uncertainty described by the model and its variances parameters and uncertainty on those parameters itself. Taking all sources of uncertainty into consideration would help to develop decision analysis process under uncertainty and should be strongly recommended.

References

[Martell, 2012] Martell, S. (2012). iscam users guide & documentation version 1.01. Technical report, University ofBritish Columbia.

[Martell et al., 2011] Martell, S., Schweigert, J., Haist, V., and Cleary, J. (2011). Moving towards the sustainable fisheries framework for pacific herring: data, models, and alternative assumptions; stock assessment and management advice for the british columbia pacific herring stocks: 2011 assessment and 2012 forecasts. Technical report, CSAP Research Document- 2011/136.

[Schnute and Richards, 1995] Schnute, J. T. and Richards, L. J. (1995). The influence of error on population estimates from catch-age models. 52(10):2063–2077.

[SCRSGroup, 2012] SCRSGroup (2012). Report of the 2012 atlantic bluefin tuna stock assesmnent meeting. Doc. No. SCI-033 / 2012.