Maximum likelihood estimates of mortality rates from catch at age data using survival analysis

Marco Kienzle*and Jason McGilvray†

November 24, 2014

Abstract

Survival analysis was applied to fisheries catch at age data. This novel likelihood approach provide an effective way to estimate natural, fishing mortality and catchability.

1 Introduction

One of the purposes of stock assessment is to estimate mortalities affecting fish stocks. This task is much easier for species that can be aged compared to, for example crustaceans, which aging is not possible. The reason lies in that mortality and longevity are inversely related hence age is a measure, albeit inverse, of mortality. The central mortality model used in fisheries research was proposed by Baranov to describe the variation of the number of fish belonging to a cohort through time [Quinn and Deriso, 1999]. This deterministic exponential model has a statistical counterpart in the form of the exponential probability distribution function which first and second moments quantify the relationship between longevity (age) and mortality rate [Cowan, 1998]. Adopting a statistical view of this problem allowed to develop maximum likelihood estimators [Burnham and Anderson, 2003] of parameters of importance to stock assessment scientists. The branch of statistics focused on survival analysis has created and refined methods to estimate mortality rates [Cox and Oakes, 1984] which are widely applied in the fields of medical research and engineering.

Despite the commonalities between survival analysis for medical and fisheries research, this theory has seldom been applied to animal ecology [Pollock et al., 1989]: to our knowledge, there hasn't been any application to fish age data for the purpose of stock assessment. In this manuscript, we described how to apply survival analysis to create likelihood functions of catch at age for the purpose of estimating natural and fishing mortalities as well as gear selectivity. We started with a simplistic example of constant natural and fishing mortality to introduce fundamental concepts from survival analysis before moving to more sophisticated cases leading to its application to real data from the sea mullet fishery in Queensland (Australia). The proposed methods were tested with simulated data to characterize some of their properties and their capacity to estimate population dynamic parameters of interest. Finally, the application to the mullet fishery case study provided specific estimates of natural mortality, catchability and selectivity.

*Queensland Dept of Agriculture, Fisheries and Forestry, Ecosciences Precinct, Joe Baker St, Dutton Park, Brisbane, QLD 4102, Australia;
University of Queensland, School of Agriculture and Food Sciences, St. Lucia, QLD 4072, Australia

†Queensland Dept of Agriculture, Fisheries and Forestry, Ecosciences Precinct, Joe Baker St, Dutton Park, Brisbane, QLD 4102, Australia

2 Materials and methods

Fish aging is possible thanks to a little bone called the otolith which is present in their ears. An otolith accumulate materials and increases in size throughout the entire lifespan of a fish. Microscopic analyses of sections of an otolith shows a series of marks, similar to tree rings, that can be used to assign each individual to a specific age-group.

Most fisheries institute around the world have a sampling program dedicated to collect a representative sample of fish each year to determine the distribution of age of any species of interest. In most cases, the data are binned into age-groups of width 1 year. For this reason, we split the lifespan of cohorts from their birth ($t \in [0; \infty]$) into n yearly intervals from $a_1 = 0$ to the maximum age of a_{n+1} years. While the theory presented in this document used that particular subdivision of time (t), un-equal ones also applies. In fact, an un-equal subdivision of time was used for the sea mullet case study.

2.1 The likelihood for constant natural and fishing mortalities

The exponential decrease in abundance of individuals belonging to a single cohort due to constant natural (M) and fishing (F) mortalities was described from a survival analysis point of view [Cox and Oakes, 1984] using a constant hazard function of time (t) and parameters θ

$$h(t; \theta) = M + F \tag{1}$$

The probability density function (pdf) describing survival from natural and fishing mortality is

$$f(t; \theta) = (M + F)\ e^{-(M+F)t} = \underbrace{M \times e^{-(M+F)t}}_{=f_1(t;\theta)} + \underbrace{F \times e^{-(M+F)t}}_{=f_2(t;\theta)} \tag{2}$$

Since age data belonging to individuals dying from natural causes are generally not available to fisheries scientists, we used only the component of the pdf that relates to fishing mortality ($f_2(t; \theta)$). This component of $f(t; \theta)$ integrates over the entire range of t to

$$\begin{aligned}
\int_{t=0}^{t=\infty} f_2(t; \theta)\ dt &= \int_{t=0}^{t=\infty} F \times e^{-(M+F)t}\ dt && (3)\\
&= \int_{t=0}^{t=\infty} f(t; \theta)\ dt - \int_{t=0}^{t=\infty} M \times e^{-(M+F)t}\ dt && (4)\\
&= 1 - \int_{t=0}^{t=\infty} M \times e^{-(M+F)t}\ dt && (5)\\
&= 1 - \frac{M}{M+F} && (6)
\end{aligned}$$

Hence, the pdf of catch at age data was obtained by normalizing $f_2(t; \theta)$

$$\begin{aligned}
g(t; \theta) &= \frac{1}{1 - \frac{M}{M+F}}\ f_2(t; \theta) && (7)\\
&= \frac{M+F}{F}\ F \times e^{-(M+F)t} && (8)\\
&= f(t; \theta) && (9)
\end{aligned}$$

Following Fisher's definition [Edwards, 1992], the likelihood of a sample of fish caught in the fishery (S_i) was written as

$$\mathcal{L} = \prod_{i=1}^{n} \Big(\int_{t=a_i}^{t=a_{i+1}} f(t;\theta)\ dt \Big)^{S_i} \tag{10}$$

$$= \prod_{i=1}^{n} P_i^{S_i} \tag{11}$$

This is often refered to as the likelihood of a multinomial probability $P_i = \int_{t=a_i}^{t=a_{i+1}} f(t;\theta)\ dt$.

The logarithm of the likelihood was

$$\log(\mathcal{L}) = \sum_{i=1}^{n} S_i \log\Big(\int_{t=a_i}^{t=a_{i+1}} f(t;\theta)\ dt \Big) \tag{12}$$

$$= \sum_{i=1}^{n} S_i \log\Big(\int_{t=a_i}^{t=a_{i+1}} (M+F)\ e^{-(M+F)t}\ dt \Big) \tag{13}$$

$$= \sum_{i=1}^{n} S_i \log\big(e^{-(M+F)\times a_i} - e^{-(M+F)\times a_{i+1}} \big) \tag{14}$$

$$\tag{15}$$

This development illustrated an application of survival analysis to estimate mortality rates affecting a cohort of fish by maximum-likelihood using a sample of catch at age. This method was implemented in R [R Core Team, 2013] in the package Survival Analysis for Fisheries Research (SAFR) provided as supplement material. Numerical application were made available using the following commands: **library(SAFR); example(llfunc1)**.

Natural and fishing mortality cannot be disentangled with catch data only but the next section will show that the provision of effort data allowed to estimate both catchability (q) and natural mortality.

2.2 Estimating catchability and natural mortality

In this section, we assumed that a time series of effort (E_i) associated with a sample of catch at age (S_i) was available to the researcher. And the assumption that fishing mortality varied according to fishing effort through constant catchability (q) held: $F(t) = q\ E(t)$. In this situation, the hazard function was written as

$$h(t,\theta) = M + q\ E(t) \tag{16}$$

And the pdf

$$f(t,\theta) = (M + q\ E(t))\ e^{-Mt - q\int_0^t E(t)\ dt} \tag{17}$$

$$= \underbrace{M \times e^{-Mt - q\int_0^t E(t)\ dt}}_{=f_1(t;\theta)} + \underbrace{q\ E(t) \times e^{-Mt - q\int_0^t E(t)\ dt}}_{=f_2(t;\theta)} \tag{18}$$

As in the previous section, we had

$$\int_{t=0}^{t=\infty} f_2(t;\theta)\ dt = 1 - \int_{t=0}^{t=\infty} M \times e^{-Mt - q\int_0^t E(t)\ dt}\ dt \tag{19}$$

But we did not know an analytic solution to the integral since the function $E(t)$ was not specified. Nevertheless, as we knew the value of effort in any given interval ($\int_{t=a_i}^{t=a_{i+1}} E(t)\ dt = E_i = \int_{t=0}^{t=a_{i+1}} E(t)\ dt - \int_{t=0}^{t=a_i} E(t)\ dt, \forall i \in [1;n]$), we could calculate the value of $\int_{t=0}^{t=\infty} f_2(t;\theta)\ dt$ assuming $E(t)$ was constant over each interval i

$$\begin{aligned}
\int_{t=0}^{t=\infty} f_2(t;\theta) &= 1 - \sum_{i=1}^{n} \Big[-\frac{M}{M + q\ E_i} e^{-Mt - q\int_0^t E(t)\ dt}\Big]_{t=a_i}^{t=a_{i+1}} && (20)\\
&= 1 - \sum_{i=1}^{n} \frac{M}{M + q\ E_i}\big(e^{-M\ a_i - q\int_0^{a_i} E(t)\ dt} - e^{-M\ a_{i+1} - q\int_0^{a_{i+1}} E(t)\ dt}\big) && (21)\\
&= \sum_{i=1}^{n} \frac{q\ E_i}{M + q\ E_i}\big(e^{-M\ a_i - q\int_0^{a_i} E(t)\ dt} - e^{-M\ a_{i+1} - q\int_0^{a_{i+1}} E(t)\ dt}\big) && (22)\\
& && (23)
\end{aligned}$$

In practice, $0 \leq \int_{t=0}^{t=\infty} f_2(t;\theta) \leq 1$ and took a specific value depending on the values of M, q and E_i. Naming this constant value K, we could write the pdf of catch at age given that effort data are available as

$$g(t;\theta) = \frac{1}{K}\ f_2(t;\theta) \tag{24}$$

And the log-likelihood:

$$\log(\mathcal{L}) = \sum_{i=1}^{n} S_i\ \log\Big(\int_{t=a_i}^{t=a_{i+1}} g(t;\theta)\ dt\Big) \tag{25}$$

Numerical application of this method were made available using the following commands: **library(SAFR); example(llfunc2)**.

Accounting for age-specific gear selectivity ($s(t)$) effects on fishing mortality ($F(t) = q\ s(t)\ E(t)$) was included in a similar way into the likelihood using constant value for selectivity at age. In practice, it is difficult to estimate n additional selectivity parameters using only the data from a single cohort but processing several cohorts at the same time and assuming separability of fishing mortality rendered estimation of catchability, natural mortality and selectivity possible.

2.3 Estimates from catch at age matrix using fishing mortality separability

This section describes an application of survival analysis to matrices of catch at age, developed for the purpose of estimating catchability (q), selectivity at age ($s(t)$) and constant natural mortality (M). The matrix ($S_{i,j}$) containing a sample of fishes aged to belong to a particular age-group j in year i contains $n + p - 1$ cohorts. These cohorts were indexed by convention using k ($k \in [1, n + p - 1]$) and an increasing number r_k ($1 \leq r_k \leq \min(n,p)$) identifying incrementally each age-group (see appendix p. 15 for more information). Each matrix $S_{i,j}$ has two cohorts with only 1 age-group representing them.

The derivation for a single cohort were the same as those presented in the previous section, here reproduced with indexations relative to a single cohort and accounting for selectivity

$$g_k(t;\theta) = \frac{q\ s(t)\ E(t) \times e^{-Mt - q\int_0^t s(t)\ E(t)\ dt}}{\sum_{l=1}^{r_k} \frac{q\ s_{k,l}\ E_{k,l}}{M + q\ s_{k,l}\ E_{k,l}}\big(e^{-M\ a_{k,l} - q\int_0^{a_{k,l}} s(t)\ E(t)\ dt} - e^{-M\ a_{k,l} - q\int_0^{a_{k,l+1}} s(t)\ E(t)\ dt}\big)} \tag{26}$$

The likelihood function of a catch at age matrix was build using each pdf specific to each cohort ($g_k(t;\theta)$):

$$\mathcal{L} = \prod_{k=1}^{n+p-1} \prod_{l=1}^{r_k} (\int_{t=a_{k,l}}^{t=a_{k,l+1}} g_k(t;\theta)\ dt)^{S_{k,l}} \tag{27}$$

This method was implemented in R [R Core Team, 2013] in the package Survival Analysis for Fisheries Research (SAFR). Numerical application of this method are available using the following commands: **library(SAFR); example(llfunc3); example(llfunc4); example(llfunc5);**.

2.4 Testing methods by simulation

Methods to estimate mortality and selectivity from a matrix containing a sample of number at age were tested with simulated datasets to characterise their performance. Variable number of cohorts $(n + p - 1)$; sample size and amount of white noise were simulated. The simulated datasets were created by generating an age-structure population using random recruitment for each cohort, natural mortality constant through years and age was randomly generated between ... and ... in each simulated dataset, random catchability and random fishing effort in each year (Tab. 1). A catch at age matrix was calculated using a logistic gear selectivity with 2 parameters.

$$s_{a_i} = \frac{1}{1 + exp(\alpha - \beta \times a_i)} \tag{28}$$

Several sampling strategies were implemented to assess how it affected mortality estimates. The problem with sampling fish cohorts is that the surveyor never has in front of him/her an entire cohort to randomly chose from. Instead s/he has access to "slices" of it each year in the form a catch. The magnitude of catch changes from year to year as a factor of fishing effort (and many other factors including variations in selectivity, behavioural changes, etc...) distorting the representation of cohorts. The challenge is to design a sampling strategy which distort as little as possible the sampling of each cohort. To study this phenomenom, we implemented first a sampling strategy that consist in sampling randomly from the entire simulated catch at age data (SS1), a strategy not applicable to real life but that has the merit to allow to assess the behaviour the maximum likelihood estimator and provide a benchmark for other sampling strategies. Second, we implemented a sampling strategy that collected a fixed number of sample (N) each year (SS2). Finally we implemented a third strategy (SS3) which weighted number at age in the sample $(S_{i,j})$ by estimated total catch at age $(\hat{C}_{i,j})$:

$$\hat{C}_{i,j} = P_{i,j} \times (C_j \otimes I(j)) \tag{29}$$

$$S^*_{i,j} = S_{i,j} \times \frac{\hat{C}_{i,j}}{C_j} \tag{30}$$

A sample of N individuals per year was drawn at random from this matrix of catch at age, creating a matrix of sampled number at age data, with dimensions $n \times p$ containing $n \times N$ data, that were processed with survival analysis methods described in previous sections.

2.5 A case study: Queensland's sea mullet fishery

The straddling sea mullet (*Mugil cephalus*) population stretches along the east coast of Australia, with most landings occurring between 19^oS (approx. Townsville) and 37^oS (roughly the border between New South Wales and Victoria). Following recommendations from Bell et al. [2005], the scientific survey design was modified from 2007 to include both estuarine and ocean habitats in order to provide representative demographic statistics for Queensland component of this fishery. Number at age obtained by otolithometry

(Tab. 2) were analyzed to estimate natural mortality, catchability and gear selectivity.

Sea mullet spawn in oceanic waters adjacent to ocean beaches from May to August each year. By convention, the birth date was assumed to be on July 1^{st} each year. Otolith ring formation occurs during spring and early summer (September to December). Biologists have come to the conclusion that the first identifiable ring is formed 14 to 18 months after birth, all subsequent rings forming at a yearly schedule. So each fish in the sample was assigned to an age group based on ring counts and translucent material at the margin of otoliths. Age group 0–1 comprised fish up to 18 months old ($a_1 = 18$ months) while all subsequent age groups spanned 12 months ($a_2 = 30$ months, $a_3 = 42$ months, etc ...).

Sensitivity of survival analysis estimates to these data, a matrix containing 7 years and 16 age-groups, were performed by truncating the dataset in 2 ways to assess the robustness of the method to varying number of years and age-groups. The first truncation removed the last and last-two years of data to evaluate the sensitivity of parameters estimates to addition/omission of data in order to anticipate possible effects of future addition of newly available data. The second truncation removed older age-groups from 10–11 to 15–16 to evaluate the importance of few old fish on natural mortality estimates as one could think *a priori* that these longer-lived individuals provided a lot of information on mortality.

3 Results

3.1 Method tests using simulated data

Natural mortality estimates were biased using a fixed number of sample per years while they were not when numbers at age in the samples were weighted by total catch (Fig 1). This correction to the sample data reduced considerably the uncertainty on natural mortality and remove almost completely bias: a small of amount of bias was still noticeable at the extremity of the range of natural mortality (0.1–1.0) tested. Increasing number of samples reduced uncertainty associated with natural mortality estimates.

Estimates of catchability were much more consistent across the range of value (1–10 10^{-4}) tested between methods (Fig. 2). The bias of the un-weighted approach was often similar to that of the weighted one. But uncertainty associated with the former approach were much larger than the latter. They appeared to be fairly in-sensitive to sample size.

3.2 Mortality estimates for sea mullet

Sea mullet catchability was estimated to be equal to 7.068 $\pm$ 3.674 per boat-day (Tab. 3). Sensitivity analysis suggested this estimation was robust to data truncations, with catchability estimated to vary between [6.906; 7.142] with an average of 7.041 boat-day^{-1}.

Natural mortality for sea mullet was estimated to 0.328 $\pm$ 0.187 year^{-1} using the entire dataset (comprising 2013 and 16 age-groups, Tab. 4). The sensitivity analysis to data truncation showed the method is robust to addition/delition of data with estimates varying between [0.292; 0.384] with an average 0.325 year^{-1}.

4 Discussion

This application of survival analysis to fisheries catch at age data provides a novel likelihood approach to estimate natural, fishing mortalities and gear selectivity. Testing this method with simulated data showed that it provided un-biased estimates of natural mortality and catchability over a wide range of simulated values. This likelihood method should find its place naturally in integrated stock assessment [Maunder and Punt, 2013] as it provide an efficient way of dealing with catch at age data.

The simulations used a logistic gear-selectivity to generate and fit the data although we would have preferred to generate data from one of the many possible gear-selectivity functions or even using non-parametric procedures. Simulations showed that gear selectivity was most difficult to estimate. The sea mullet case study was in fact not fitted with a logistic curve but selectivity was estimated through a tedious process to search each proportion retained at age that best fitted the data as measured by the likelihood. This tedious process could not be automatised into the simulation-testing framework to provide automatic identification of gear-selectivity. Therefore it was left out for a future work. Criticisms that this somewhat simplified the problem would be correct. But the current article was designed as an introduction to these methods in fisheries research not one that solves all problems at once. The purpose of writing the present manuscript was to provide likelihood methods allowing to identify the correct gear-selectivity.

Parasite data suggested that the bulk of sea mullet caught in Queensland fishery is based on local fish populations and not migrating from NSW [Lester et al., 2009]. While genetic analyses could not identify differences in single nucleotide polymorphism between samples from south QLD and NSW [Krück et al., 2013].

The sensitivity analysis to data truncation showed a weak trend in increasing uncertainty associated with natural mortality as more older age-groups containing few observations were added to the analysis suggesting that lack of information with those create large uncertainties on gear selectivity estimates which translate into greater estimates of natural mortality uncertainty.

This survival analysis approach to fisheries catch at age data can be further expanded to provide an estimator of recruitment.

References

P.A. Bell, M.F. O'Neill, G.M. Leigh, A.J. Courtney, and S.L. Peel. Stock assessment of the Queensland-New South Wales sea mullet fishery (*Mugil cephalus*). Technical Report QI05033, Queensland Government, 2005.

K. P. Burnham and D. Anderson. *Model Selection and Multi-Model Inference*. Springer-Verlag, second edition, 2003.

G. Cowan. *Statistical Data Analysis*. Oxford Science Publications, 1998.

D.R. Cox and D. Oakes. *Analysis of survival data*. Chapman and Hall, 1984.

A.W.F. Edwards. *Likelihood*. Johns Hopkins University Press, 1992.

N.C. Krück, D.I. Innes, and J.R. Ovenden. New SNPs for population genetic analysis reveal possible cryptic speciation of eastern australian sea mullet (*Mugil cephalus*). *Molecular Ecology Resources*, 13(4):715–725, 2013.

R.J.G. Lester, S.E. Rawlinson, and L.C. Weaver. Movement of sea mullet mugil cephalus as indicated by a parasite. *Fisheries Research*, 96(23):129 – 132, 2009.

Mark N. Maunder and André E. Punt. A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142:61–74, 2013.

K.H. Pollock, S.R. Winterstein, and M.J. Conroy. Estimation and analysis of survival distributions for radio-tagged animals. *Biometrics*, 45(1):pp. 99–109, 1989.

T. J. Quinn and R. B. Deriso. *Quantitative Fish Dynamics*. Oxford University Press, 1999.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL `http://www.R-project.org/`.

224

Figures

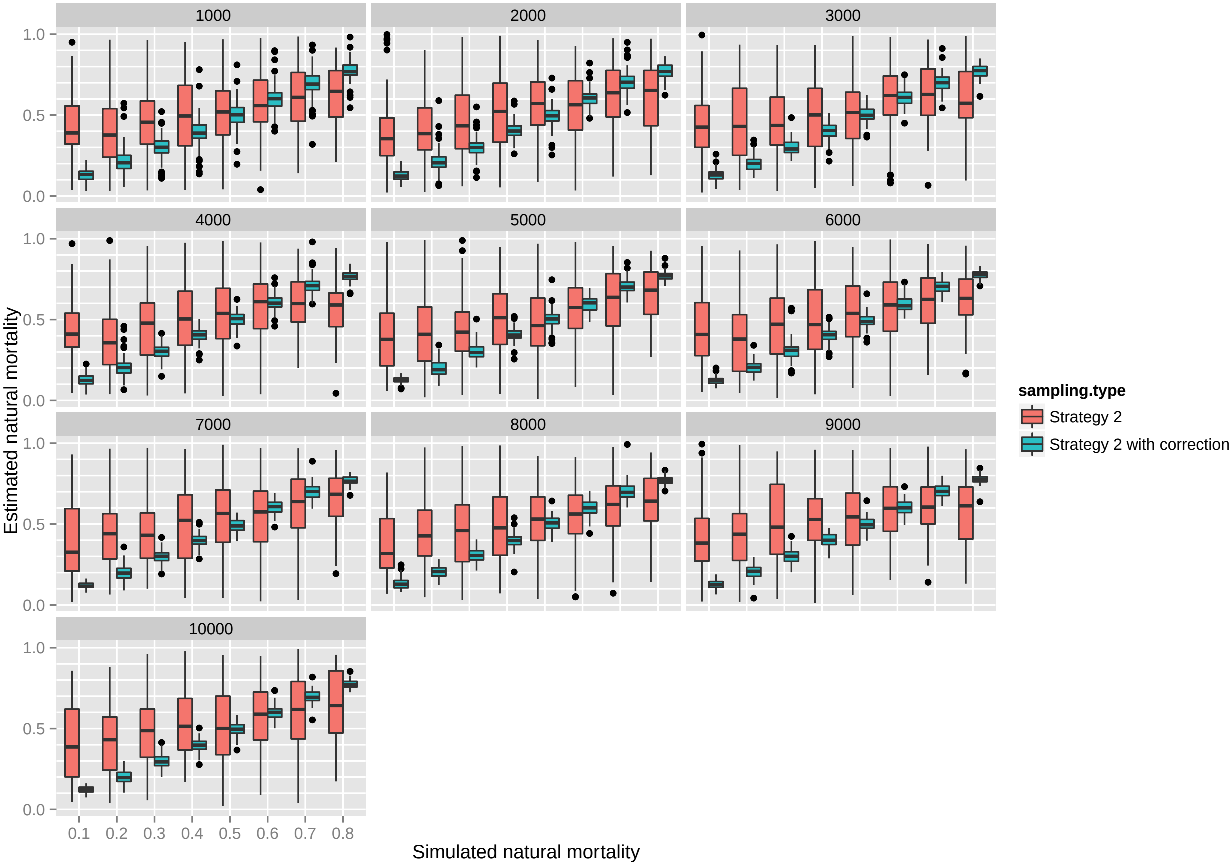

Figure 1: Comparison between simulated natural mortality (x-axis) and estimated using (a) a fixed number of sample each year or (b) data weighted by catch. Each panel correspond to an increasing number of samples per year varying from 1000 to 10.000.

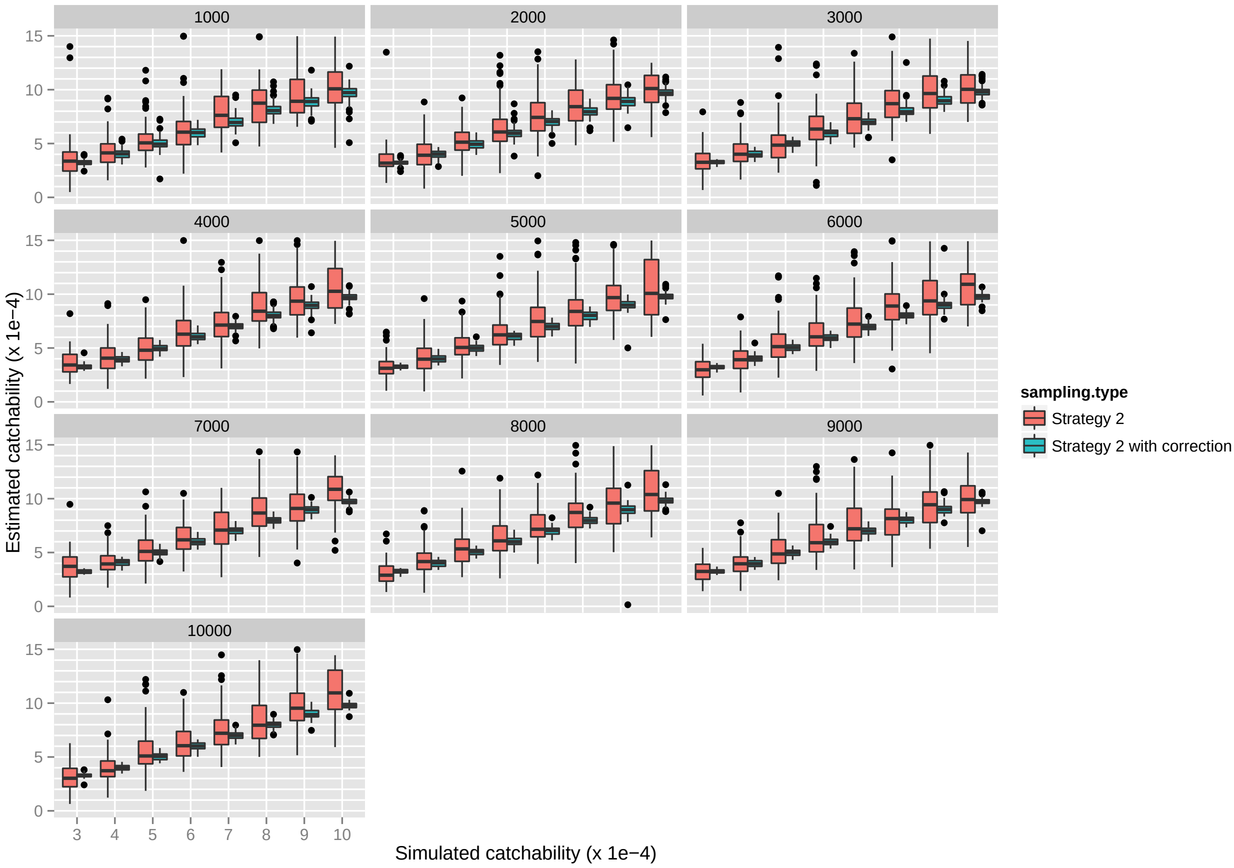

Figure 2: Comparison between simulated catchability (x-axis) and estimated using (a) a fixed number of sample each year or (b) data weighted by catch. Each panel correspond to an increasing number of samples per year varying from 1000 to 10.000.

Tables

Variable type	Distribution	Parameters
recruitment	uniform	min=1e6, max=1e7
natural mortality	uniform	min=0.1, max=0.8
catchability	uniform	min=3e-4, max=1e-3
fishing effort	uniform	min=1e3, max=5e3
gear selectivity α	uniform	min=8, max=12
gear selectivity β	uniform	min=1, max=3

Table 1: Distribution and range of value taken by different type of random variable in simulations.

	0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12	12–13	13–14	14–15	15–16	Catch	Effort
2007	11	180	517	561	118	105	45	24	11	3							1350	7400
2008		42	468	618	409	100	57	21	10	8	2	2		2			1795	7875
2009	1	110	280	679	251	151	29	17	6	1	3	2	1	1			1815	6529
2010	2	239	541	250	200	97	50	11	9	2							1757	6109
2011	6	244	598	500	115	71	35	10	2	4		1	1		1	1	1542	6412
2012	1	99	633	563	298	57	32	15	11								1649	6993
2013		89	405	955	532	183	25	24	5					1			1993	6667

Table 2: Distribution of yearly samples (in rows) of sea mullet into age-groups of width 1 year (in columns); catch in tonnes and effort in boat-days.

	10	11	12	13	14	15	16
2010	6.9 ± 11.2	6.9 ± 14.6	6.9 ± 22.1	6.9 ± 48	6.9 ± 20	6.9 ± 13.8	6.9 ± 11
2011	7.1 ± 1.3	7.1 ± 1.8	7.1 ± 4.3	7.1 ± 1.9	7.1 ± 4	7.1 ± 6.2	7.1 ± 8.2
2012	7.1 ± 2.2	7.1 ± 1.4	7.1 ± 2.4	7.1 ± 2.5	7.1 ± 3.2	7.1 ± 3.7	7.1 ± 2.2
2013	7 ± 1.1	7 ± 1.7	7 ± 3.6	7 ± 4.1	7.1 ± 0.9	7.1 ± 5.4	7.1 ± 3.7

Table 3: Sensitivity of catchability estimates (in boat-day^{-1}) to data truncations. Rows indicate the most recent year of data and columns the maximum age-group included in the analysis.

	10	11	12	13	14	15	16
2010	0.29 ± 0.38	0.29 ± 0.48	0.29 ± 0.79	0.29 ± 1.75	0.29 ± 0.82	0.29 ± 0.56	0.29 ± 0.45
2011	0.32 ± 0.1	0.33 ± 0.1	0.34 ± 0.15	0.32 ± 0.15	0.32 ± 0.32	0.32 ± 0.48	0.32 ± 0.61
2012	0.35 ± 0.09	0.37 ± 0.05	0.35 ± 0.11	0.35 ± 0.11	0.35 ± 0.14	0.35 ± 0.16	0.37 ± 0.05
2013	0.31 ± 0.09	0.32 ± 0.14	0.32 ± 0.31	0.32 ± 0.35	0.38 ± 0.05	0.33 ± 0.29	0.33 ± 0.19

Table 4: Sensitivity of natural mortality estimates (in year^{-1}) to data truncations. Rows indicate the most recent year of data and columns the maximum age-group included in the analysis.

Appendices

Definitions of some mathematical symbols

This appendice contains definitions of some of the mathematical symbols used in previous sections

- $S_{i,j}$: a matrix of dimensions $n \times p$ ($i \in [1, n]$ and $j \in [1, p]$) containing a number of fishes that were aged and found to belong to specific age-groups j in a particular year i. This matrix contains data belonging to $n + p - 1$ cohorts, which by convention were labeled using k varying from 1 on the top-right corner of the matrix to $n + p - 1$ on the bottom-left (Tab. 5).

	1		...			p
1	...	...	...	3	2	1
	...	...	...	...	3	2
⋮	...	...	...	...	4	3
	...	...	k	...	...	...
	...	...	...	...	...	...
n	$n+p-1$	...	...	...	...	...

Table 5: Convention used to associate each element of the catch at age matrix ($C_{i,j}$) with particular cohort referred to as with the number given in this table.

  The number of data in $S_{i,j}$ belonging to each cohort (r_k) varies from 1 to $\min(n, p)$ and was determined as follow:

$$r_k = \begin{cases} i - j + p \text{ if } k < \min(n,p) \\ \min(n,p) \text{ if } \min(n,p) \leq k < \max(n,p) \\ j - i + n \text{ if } k \geq \max(n,p) \end{cases} \tag{31}$$

  Each element of the $S_{i,j}$ matrix is uniquely identified using indices i and j ($1 \leq i \leq n$ and $1 \leq j \leq p$) or indices k and l ($1 \leq k \leq n + p - 1$ and $1 \leq l \leq r_k$), so for example

$$\sum_{i,j} S_{i,j} = \sum_{k,l} S_{k,l} \tag{32}$$

- $P_{i,j}$: a matrix of dimensions $n \times p$ ($i \in [1, n]$ and $j \in [1, p]$) containing the proportion at age in the sample ($S_{i,j}$). Rows of this matrix sum to 1.

$$P_{i,j} = \frac{S_{i,j}}{\sum_i S_{i,j}} \tag{33}$$

$$S^*_{i,j} = \frac{S_{i,j}}{\sum_i S_{i,j}} \tag{34}$$

- $F_{i,j}$ a matrix of fishing mortality with dimension $n \times p$ ($i \in [1, n]$ and $j \in [1, p]$). This matrix was constructed as the outer product of year specific fishing mortalities ($q\ E_i$) and selectivity at age (s_j):

$$F_{i,j} = q\ E_i \otimes s_j \tag{35}$$