Caveats with estimating natural mortality rates in stock assessment models using age aggregated catch data and abundance indices

M. Aldrin [a,*], F.L. Aanes [a], I.F. Tvete [a], S. Aanes [a], S. Subbey [b]

[a] *Norwegian Computing Center, P.O.Box 114 Blindern, N-0314 Oslo, Norway*
[b] *Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway*

ARTICLE INFO

ABSTRACT

We consider the challenge in estimating the natural mortality, M, in a standard statistical fish stock assessment model based on time series of catch- and abundance-at-age data. Though anecdotal evidence and empirical experience lend support to the fact that this parameter may be difficult to estimate, the current literature lacks a theoretical justification. We first discuss the estimatability of a time-invariant M theoretically and present necessary conditions for a constant M to be identifiable. We then investigate the practical usefulness of this by estimating M from simulated data based on models fitted to 19 fish stocks. Using the same data sets, we next explore several model formulations of time varying M, with a pre-specified mean value. Cross validation is used to assess the prediction performance of the candidate models. Our results show that a time-invariant M can be estimated with reasonable precision for a few stocks with long time series and typically high values of the true M. For most stocks, however, the estimation uncertainty of M is very large. For time-varying M, we find that accounting for variability across age and time using a simple model significantly improves the performance compared to a time-invariant M. No significant improvement is obtained by using complex models, such as, those with time dependencies in variability around mean values of M.

1. Introduction

Time series of catch at age and abundance at age indices are critical input for assessing commercially important fish stocks. For this type of data, the assessment models in use for management advice are typically variants of statistical catch at age models, where processes for the fishing mortality rate F, the natural mortality rate M and possibly recruitment are modelled as stochastic processes (e.g. Gudmundsson, 1994; Cook, 2004; Aanes et al., 2007; Nielsen and Berg, 2014). For such models, the natural mortality rate M is considered difficult to estimate precisely when the fishing mortality rate F and the catchability q are to be simultaneously estimated with M, in the absence of additional data, such as capture-recapture data and stomach content data (e.g. Quinn and Deriso, 1999; Bogstad et al., 2000; Cook, 2004; Aanes et al., 2007; Punt et al., 2021). Additional data is often lacking, and M is therefore often fixed at a pre-specified value, as for example in the following stocks managed by ICES (ICES, 2020b,a): Norwegian Spring Spawning Herring ($M = 0.9$ for age two or less, $M = 0.15$ for higher ages), Western Horse Mackerel ($M = 0.15$), Cod in Norwegian Coastal Waters ($M = 0.2$), Beaked Redfish ($M = 0.05$) and Greenland Halibut ($M = 0.1$).

However, Cook (2004) estimated M based on abundance indices for scientific survey data during a period where fishery was closed and found age specific evidence for variation across age, but not over time for a data set on Shetland sandeel. In a simulation experiment based on models fitted to 12 Pacific Coast groundfish stocks, applying the assessment model Stock Synthesis, Lee et al. (2011) concluded that M could be estimated with reasonable precision in most such cases when the model was true. The authors further asserted that unrealistic estimates of M, based on real data, were often due to severe model mis-specification. Both these conclusions were supported by Punt et al. (2021). It should be noted that they had 50 years of data in their simulation study, and that many important fish stocks have considerably less historical data than this. Magnusson and Hilborn (2007) also performed a simulation experiment, but with a less optimistic conclusion. They varied, among others, the exploitation history, and concluded that M was estimatable with some reliability when age composition data were available from before major catches were removedoved but found M to be less estimatable under other fisheries scenarios. Aanes et al. (2007) fitted a model to both real and simulated data for Northeast Arctic cod and found it difficult to estimate the mean value of M. Instead,

they estimated the temporal dynamics in the natural mortality rate around a mean value, using an informative prior for the mean. Other authors have used similar models, but with a fixed, pre-specified value for the mean, and examples of such models are found in Cadigan (2016), Aldrin et al. (2020) and Stock et al. (2021).

To summarize, to our knowledge all studies conclude that the estimation of M is at best difficult without additional data sources or with substantial historical data. The conclusions appear to be based on established practices or empirical findings of model performance, without a theoretical justification.

In this paper, we investigate the estimatability of M further based on time series of catch at age data and abundance at age indices. After introducing a basic assessment model, we investigate necessary conditions for M to be theoretical identifiable, using the framework of Cole and McCrea (2016). Next, we test if it is possible to estimate the level of M in practice, when the model is true, and M is constant over time, by simulating from models fitted to 19 fish stocks. Finally, we compare the prediction performance of models with (i) pre-specified M, (ii) estimated M and (iii) time-varying M around a pre-specified level, by fitting them to real data for the same 19 fish stocks in a cross validation experiment.

2. Models

Here, we present a general stock assessment model consisting of two sub models; a population model for the fish stock and a data model, which links the observed catch at age and abundance at age data to the population model. Ages are indexed by $a = 1, ..., A$, where $a = 1$ is the first age we consider in the model and not necessarily the biological age, and $a = A$ is a plus group. Likewise, years are indexed by $y = 1, ..., Y$, where $y = 1$ and $y = Y$ are the first and last years we consider.

Table 1 gives an overview of relevant variables.

2.1. Data

We consider two types of data, estimates of yearly age specific catches and observed age specific survey index data. Let $\widehat{C}_{a,y}$ denote an estimate of the true, but unknown, number of fish $C_{a,y}$ of age a caught during year y. Let $\widehat{I}_{a,y}(d)$ denote an observed survey index for fish of age a for a survey conducted at the end of day d of year y.

Table 1
Overview of notations. The term "of age a" is dropped from the interpretation text, except for the first line, in order to avoid repetition.

Non-observable, true quantities	
Notation	Interpretation
$N_{a,y}$	True number of fish of age a at the start of year y
$C_{a,y}$	True number of fish caught during year y (i.e. fishing mortality)
$F_{a,y}$	Instantaneous fishing mortality rate during year y corresponding to $C_{a,y}$
$M_{a,y}$	Instantaneous natural mortality rate during year y
$Z_{a,y} = F_{a,y} + M_{a,y}$	Instantaneous total mortality rate during year y
$N_{a,y}(d)$	True number of fish at the end of day d of year y
Q_a	Catchability, assumed to be constant over years
$I_{a,y}(d)$	"True" survey index proportional to $N_{a,y}(d)$, $I_{a,y}(d) = Q_a N_{a,y}(d)$
	$I_{a,y}$ define the expectations from a perfect survey with full coverage

Data, observations or "preliminary" estimates, to be updated when the assessment model is estimated	
Notation	Interpretation
$\widehat{C}_{a,y}$	"Preliminary" estimate of $C_{a,y}$
$\widehat{I}_{a,y}(d)$	Observed survey index, "preliminary" estimate of $I_{a,y}(d)$
$M^*_{a,y}$	Estimate or best guess of $M_{a,y}$

2.2. A general stock assessment model

We divide the stock assessment model into a population model of true, unknown quantities and a data model which defines how the observations are related to the quantities in the population model.

2.2.1. The population model

We first consider general age and year specific fishing and natural mortalities rates $F_{a,y}$ and $M_{a,y}$, and a general year specific recruitment $N_{1,y}$ for the lowest age. In Sections 2.3, 3 and 4, we will consider specific models for these and other quantities.

We assume that the highest age group A defines a plus group of fish aged A or older. The population model is, for $1 <= y <= Y$

$$N_{a+1,y+1} = exp(-(F_{a,y} + M_{a,y}))N_{a,y} = exp(-Z_{a,y})N_{a,y} \tag{1}$$

for $1 <= a <= A - 2$ and

$$\begin{aligned} N_{A,y+1} &= exp(-(F_{A-1,y} + M_{A-1,y}))N_{A-1,y} + exp(-(F_{A,y} + M_{A,y}))N_{A,y} \\ &= exp(-Z_{A-1,y})N_{A-1,y} + exp(-Z_{A,y})N_{A,y} \end{aligned} \tag{2}$$

for the the plus group.

We assume that the number of fish within an age group decay exponentially during a year, and at day d this number is given by

$$N_{a,y}(d) = exp(-(d/365)Z_{a,y})N_{a,y}. \tag{3}$$

The relationship between the fishing mortality rate and the number of fish caught during a year is uniquely given by Baranov's catch equation (e.g. Quinn and Deriso, 1999)

$$C_{a,y} = (F_{a,y}/Z_{a,y})(1 - exp(-Z_{a,y}))N_{a,y}. \tag{4}$$

Note that even if we use the term "natural mortality rate" for $M_{a,y}$, the model above is well defined also when $M_{a,y}$ includes migration. If the immigration is large enough, then $M_{a,y}$, and even $Z_{a,y}$, may become negative.

2.2.2. The data model

The data model relates observations to the true, unknown, quantities in the population model. We assume that the data are observed for all ages from years 1 to Y. We further assume that the catch at age data $\widehat{C}_{a,y}$ are (mean-)unbiased estimates of the true catch and log-normally distributed, i.e.

$$\widehat{C}_{a,y} = C_{a,y} exp(\varepsilon^C_{a,y}), \tag{5}$$

$$\varepsilon^C_{a,y} \sim N(-1/2\sigma^2_{C,a}, \sigma^2_{C,a}). \tag{6}$$

The corresponding data model for the age specific survey index can be written as

$$\widehat{I}_{a,y}(d) = I_{a,y}(d)exp(\varepsilon^I_{a,y}) = Q_a N_{a,y}(d)exp(\varepsilon^I_{a,y}), \tag{7}$$

$$\varepsilon^I_{a,y} \sim N(-1/2\sigma^2_{I,a}, \sigma^2_{I,a}). \tag{8}$$

Here, $\widehat{I}_{a,y}(d)$ is the noisy observation of an ideal, noise-free survey index $I_{a,y}(d) = Q_a N_{a,y}(d)$, and Q_a an age specific proportionality constant, called catchability.

2.2.3. Estimation

We use the TMB software (Kristensen et al., 2016) to estimate unknown quantities in the model by maximum likelihood. The maximum likelihood estimates are marked with a $\sim$ (e.g. $\widetilde{C}_{a,y}$), to distinguish them from true quantities (e.g. $C_{a,y}$) or observations (e.g $\widehat{C}_{a,y}$).

2.3. Identifiability

In a model, such as an assessment model for a fish stock, it may be that some of the parameters cannot be estimated from observed data. Such parameters are said to be non-identifiable, the remaining parameters being identifiable. A parameter that is non-identifiable with the current observed data may become identifiable if other types of data are added, for instance capture-recapture data. However, even if a parameter is theoretically identifiable, it may be practically non-identifiable because there are too few observations to estimate the parameter with a reasonable precision. A comprehensive overview is given in Cole (2020). Examples investigating parameter identifiability in fisheries models, but with different types of data, are found in Allen et al. (2017), Cole and Morgan (2010), Jiang et al. (2007), Nater et al. (2020) and Polansky et al. (2021).

In this section, we focus on theoretical identifiability, using the methodology described in Cole and McCrea (2016). They developed a method for investigating theoretical parameter identifiability in linear state space models. This method consists of constructing a so-called exhaustive summary vector, and then differentiating it with respect to the parameters. If the resulting matrix is of full rank then all the parameters are identifiable. If not, the structure of the matrix identifies which parameter that cannot be estimated. A linear state space model can be written as

$$x_t = G_t x_{t-1} + \omega_{t-1}, y_t = B_t x_t + \eta_t, \qquad (9)$$

Here, x_t is a vector with the states at time t, which in our case are the stock sizes, and y_t is a vector of observations, in our case the catch and index data. Furthermore, G_t and B_t are matrices and ω_{t-1} and η_t are vectors of random errors. One exhaustive summary for this linear state space model is given in the Supplementary Material in Cole and McCrea (2016):

$$\kappa(\theta) = \begin{bmatrix} E(y_1) \\ \text{vec}(\text{Var}(y_1)) \\ E(y_2) \\ \text{vec}(\text{Var}(y_2)) \\ \vdots \\ E(y_T) \\ \text{vec}(\text{Var}(y_T)) \end{bmatrix} \qquad (10)$$

Here, T is the the final time point considered, $E(y_t)$ is the expected value of y_t and $\text{vec}(\text{Var}(y_t))$ is the vectorised version of the covariance matrix of y_t.

Cole and McCrea (2016) also showed how the method could be used in non-linear state space models, but this is more complicated. We specified the assessment model in Section 2.2 on the original scale, but on the logarithmic scale it is a linear state space model, except for the

age plus group A. Here, we simplify this, and instead assume age A to be a specific age following Eq. (1), and ignore older fish. Furthermore, we choose a specific model for the recruitment, a random walk on the logarithmic scale, given by

$$N_{1,y+1} = N_{1,y} exp(\varepsilon_y^R) \qquad (11)$$

$$\varepsilon_y^R \sim N(0, \sigma_R^2), \qquad (12)$$

but the results presented below holds also for the recruitment model $N_{1,y+1} = \alpha exp(\varepsilon_y^R)$, where α is a parameter to be estimated. We can then use the methodology of Cole and McCrea (2016) to investigate which parameters or unknown quantities can be estimated from the data. More details on the methodology, and the exact formulation of the assessment model in into this framework is given in the Supplementary Material.

We will consider models with six different restrictions on $M_{a,y}$, but where the Q_a's and the $F_{a,y}$'s are unrestricted, given that both the catch data and the survey index is observed for all years and ages, and that the number of observations ($2 \cdot A \cdot Y$) is at least equal to the total number of parameters, see Table 2.

If $M_{a,y} = M$, i.e. constant over years and ages, then all parameters can be estimated. The same holds for a slightly extended version of this, with $M_{a,y}$ being proportional to a pre-specified $M_{a,y}^*$ as

$$M_{a,y} = \theta \cdot M_{a,y}^*, \qquad (13)$$

where θ is a positive parameter to be estimated.

In a model with more parameters, where the natural mortality rate varies by age ($M_{a,y} = M_a$), all parameters except M_A and the fishing and natural mortality rates for the highest age ($F_{A,y}$ and $M_{A,y} = M_A$) can be estimated. It is intuitive that M_A cannot be estimated since since we have no information on how many fish that survive the upper age A. One way to solve this is to introduce the constraint $M_A = M_{A-1}$.

In a more flexible model, we let the natural mortality rate vary by both age and year by the separable structure $M_{a,y} = M_a^{age} + M_y^{year}$, with constraint $M_0^{year} = 0$, where $M_a^{age}, a = 1, ..., A$ is a set of parameters that depend on age a and $M_y^{year}, y = 1, ..., Y$ is another set of parameters that depend on year y. In this model, the parameters M_A^{age} and M_Y^{year} cannot be estimated. Also the the fishing and natural mortality rates are non-estimable both for the highest age and for the last year. Again, this is quite intuitive, since we have no information on how many fish that survive the upper age and the last year.

If $M_{a,y}$ is unrestricted, then no parameters can be estimated.

In light of these results, we consider a medium flexible model, where $M_{a,y}$ varies around a pre-specified level $M_{a,y}^*$, and are equal for ages $A - 1$ and A (see Table 2 for an exact definition). Then all parameters can be estimated, except for the fishing and natural mortality rates for the last year.

Table 2
Theoretical identifiability of parameters for different restrictions on $M_{a,y}$. The text "All" and "None" means that all or none of the parameters in the corresponding category can be estimated, respectively.

Restriction	$N_{a,1}$	$F_{a,y}$	$M_{a,y}$	Q_a	Variance parameters
$M_{a,y} = M$	All	All	All	All	All
$M_{a,y} = \theta \cdot M_{a,y}^*$	All	All	All	All	All
$M_{a,y} = M_a$	All	All, except $F_{A,y}$, all y	All, except $M_A = M_{A,y}$, all y	All	All
$M_{a,y} = M_a^{age} + M_y^{year}$	All	All, except $F_{A,y}$, all y $F_{a,Y}$, all a	All, except $M_{A,y}$, all y $M_{a,Y}$, all a	All	All
$M_{a,y}$ unrestricted	None	None	None	None	None
$M_{a,y} = M_{a,y}^* + m_{a,y}$ and $(1/(A-1))\sum_{a=1}^{a=A-1} m_{a,y} = 0$ and $m_{A,y} = m_{A-1,y}$	All	All, except $F_{a,Y}$, all a	All, except $M_{a,Y}$, all a	All	All

3. Investigating practical estimatability in the model $M_{a,y} = \theta \cdot M_{a,y}^*$ when the model is true

3.1. Set up for a simulation experiment

In Section 2.3, all parameters are theoretically estimatable when $M_{a,y} = M$ is equal for all years and and ages or when $M_{a,y} = \theta \cdot M_{a,y}^*$, i.e. proportional to the pre-specified values for each age and year. Still, a scarcity of data, such as few observed years, can result in large estimation uncertainty, even if the model is true. This is even more so when parts of the data, for instance the survey indices, are imprecise.

Here, we investigate practical estimability by a simulation experiment. For a given data set, estimability may also be investigated by studying likelihood profiles (Raue et al., 2009), and we demonstrate this method on selected data sets. Other methods for investigating practical estimability on specific data sets include examining the eigen values of the Hessian matrix (Catchpole et al., 2001) and data cloning (Lele et al., 2010).

In the simulation experiment, we first fit an assessment model to 19 fish stock data sets, giving 19 true operating models. These 19 data sets consist of 14 data sets previously analysed in Aldrin et al. (2020), and another five data sets from stockassesment.org (2020) with status "final" on the 25th of May 2020 (Table 3). We use the catch series and one survey index series from each data set and ignore all other data.

The assessment model we fitted to each data set is a specific version of the more general model we introduced in Section 2.2. The natural mortality rate is fixed at pre-specified values $M_{a,y}^*$, i.e. given by Eq. (13) with $\theta = 1$. The values $M_{a,y}^*$ are taken from the respective data sets, and are based on some prior information or best guess of the natural mortality rates. They may vary over ages and years, but are usually the same for all ages and years or they only vary by age.

The fishing mortality rate model is a random walk process on the logarithmic scale, as used in e.g. the SAM model (Nielsen and Berg, 2014):

$$F_{a,y} = F_{a,y-1} \exp\left(\varepsilon_{a,y}^F\right), 1 <= a < A - 1, \tag{14}$$

$$F_{A,y} = F_{A-1,y}, \tag{15}$$

$$\varepsilon_{a,y}^F \sim N(0, \sigma_F^2), \tag{16}$$

with equal standard deviation σ_F for all ages. We further assume that the $\varepsilon_{a,y}^F$'s are independent between years, but positively correlated between ages within the same year, by $\mathrm{corr}(\varepsilon_{a,y}^F, \varepsilon_{a',y}^F) = \rho^{|a-a'|}$, where ρ is a parameter.

The logarithms of the standard deviations $\sigma_{C,a}$ and $\sigma_{I,a}$ and the catchability parameters Q_a are allowed to vary smoothly over ages by quadratic functions of age, using three instead of A parameters to describe each of these (see Supplementary Material for details).

As previously stated, this full-specified assessment model is fitted to each of the 19 data sets and form the basis for our simulation procedure afterwards. In the following simulation procedure we do make an adjustment to the assumptions regarding $F_{a,y}$. The model for $F_{a,y}$ is non-stationary, and can give unrealistic low or high values when simulated over many years. Therefore, for the operating model only, we replace the random walk process on log scale given by Eq. (14) with a stationary AR(1) process on log scale where $log(F_{a,y})$ varies around $log(0.4)$, see Supplementary Material for details. We term the resulting model an operating model, which will serve as the truth in the simulations.

For each of the 19 fitted operating models, one for each stock, we simulate 50 random samples, where the catch and survey data covers the same years and ages as in the real data sets. For each simulated stock, we fit an assessment model almost equal to the operating model, but with $F_{a,y}$ modelled as a random walk on log scale as in Eqs. (14), (15), (16). To restrict θ in Eq. (13) to be positive, we re-parameterise it to the non-restricted parameter $\theta^* = log(\theta)$ when fitting the model. The true value of θ is 1 in the operating model.

3.2. Results for the simulation experiment

The mean absolute deviance (MAD) of $log(\widetilde{\theta}) - log(1) = log(\widetilde{\theta}))$ for each stock varied from 0.06, a rather precise estimate, to 2.4. Fig. 1 show how MAD($log(\widetilde{\theta})$) tends to decrease by increasing values of the true natural mortality rates (panel a) and by the number of observations in total over years and ages of $\widetilde{C}_{a,y}$ and $\widetilde{I}_{a,y}$ (panel b). The improved precision of $\widetilde{\theta}$ by increasing number of observations is an obvious and expected result. A linear regression of MAD to these two explanatory
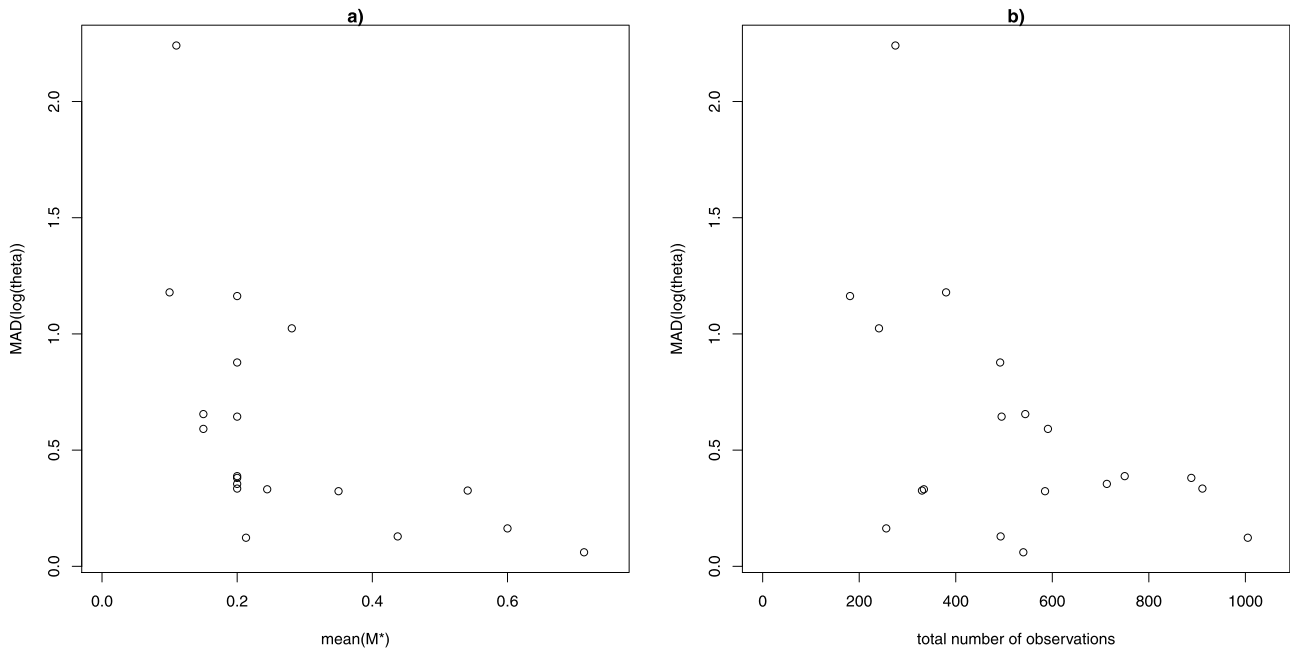
Table 3

Overview of data sets used. Except for the first three data sets, the data sets are downloaded from stockassesment.org (2019) (rows 4–14) and stockassesment.org (2020) (rows 15–19). The minimum and plus ages are the ones used in the models. Together, the catch and survey data cover all years and ages, but usually each data type covers only a subset of the whole year and age ranges. To save computation time, we have ignored catch data before 1961 for North-East Atlantic cod.

Data source	Area	Species	Years catch	Years index	Min. age plus age	Range of M^*a, y	Survey index
(Nielsen and Berg, 2014)	North Sea	Cod	1963–2011	1983–2012	1–7	0.2–1.31	IBTS Q1
(ICES, 2017)	Coast of Norway	Cod	1984–2016	1995–2016	2–10	0.2	Table T26, p. 69
(ICES, 2017)	North-East Atlantic	Cod	1961–2016	1981–2017	3–15	0.2–0.788	Table A3, p. 180
BW_2018	Widely distributed	Blue whiting	1981–2018	2004–2018	1–10	0.2	IBWSS
sole2024_newidx	North Sea	Sole	1984–2017	2004–2017	1–9	0.1	Fisherman
sam-tmb-fsaithe-2017-01	Faroe Plateau	Saithe	1961–2017	1994–2018	3–15	0.2	Spring
sam-tmb-fcod-2017-01	Faroe Plateau	Cod	1959–2017	1996–2017	1–10	0.2	Summer
NSwhiting_2018	North Sea	Whiting	1978–2017	1983–2017	0–8	0.34–2.26	IBTS-Q1
codEastNWWG2018	Iceland/East Greenland	Cod	1973–2017	1982–2017	1–10	0.2–0.5	WH
sam-tmb-fhaddock-2017-01	Faroe Plateau	Haddock	1957–2017	1996–2017	1–10	0.2	Summer
WBSS_mf_004_CB_corrCF	Baltic Sea	Herring	1991–2016	1991–2016	0–8	0.2–0.5	HERAS
PLE2123_WGBFAS2017_Final_run	Baltic Sea	Plaice	1999–2016	1999–2017	1–10	0.10–0.20	IQ IBTS+BITS
WGWIDE2017.V2	Widely distributed	Mackerel	1980–2016	2010–2017	0–12	0.15	Swept-idx
WBcod_2017	Baltic Sea	Cod	1994–2016	2001–2016	0–7	0.2–0.8	SD2224w_4Q
ARU.27.5b6a_WGDEEP_2020_	Faroe Plateau	Greater silversmelt	1995–2019	1998–2019	5–21	0.15	Faroese Summer survey
wit.27.3a47d_2020	North Sea	Whitch	2009–2019	2009–2019	1–10	0.2	Q1
Nea_haddock_2019	North-East Atlantic	Haddock	1950–2018	1991–2017	3–13	0.2	RU-BTr-Q4
HAD7bk_2020_Benchmark_II	Celtic Sea + West of Scotland	Haddock	1993–2018	2003–2018	0–8	0.36–1.09	FRA_IRL_WIBTS_VAST
whg.7b-ce-k_FRA_Tun_longretro	Celtic Sea + West of Scotland	Whiting	1999–2018	2003–2018	0–7	0.36–1.22	IFGS VAST No/Km2

Fig. 1. Scatter plot of MAD($log(\widetilde{\theta})$) vs. the average (over all years and ages) true mortality rates (panel a) and the number of observations (panel b).

variables explained 63% ($R^2 = 0.63$) of the variance of MAD, and both effects were highly significant ($p < 0.001$). None of the other characteristics of the data were significantly related to the precision of $\widetilde{\theta}$.

We have a closer look at two data sets, the Northeast Atlantic Cod

(NEAC) and the Baltic Sea Cod (BSC). The average true mortality rate is similar for the two data sets (0.21 for NEAC and 0.28 for BSC), but NEAC has 1005 observations and BSC only 241. We therefore expect that θ and other quantities will be more precisely estimated for NEAC than for BSC.
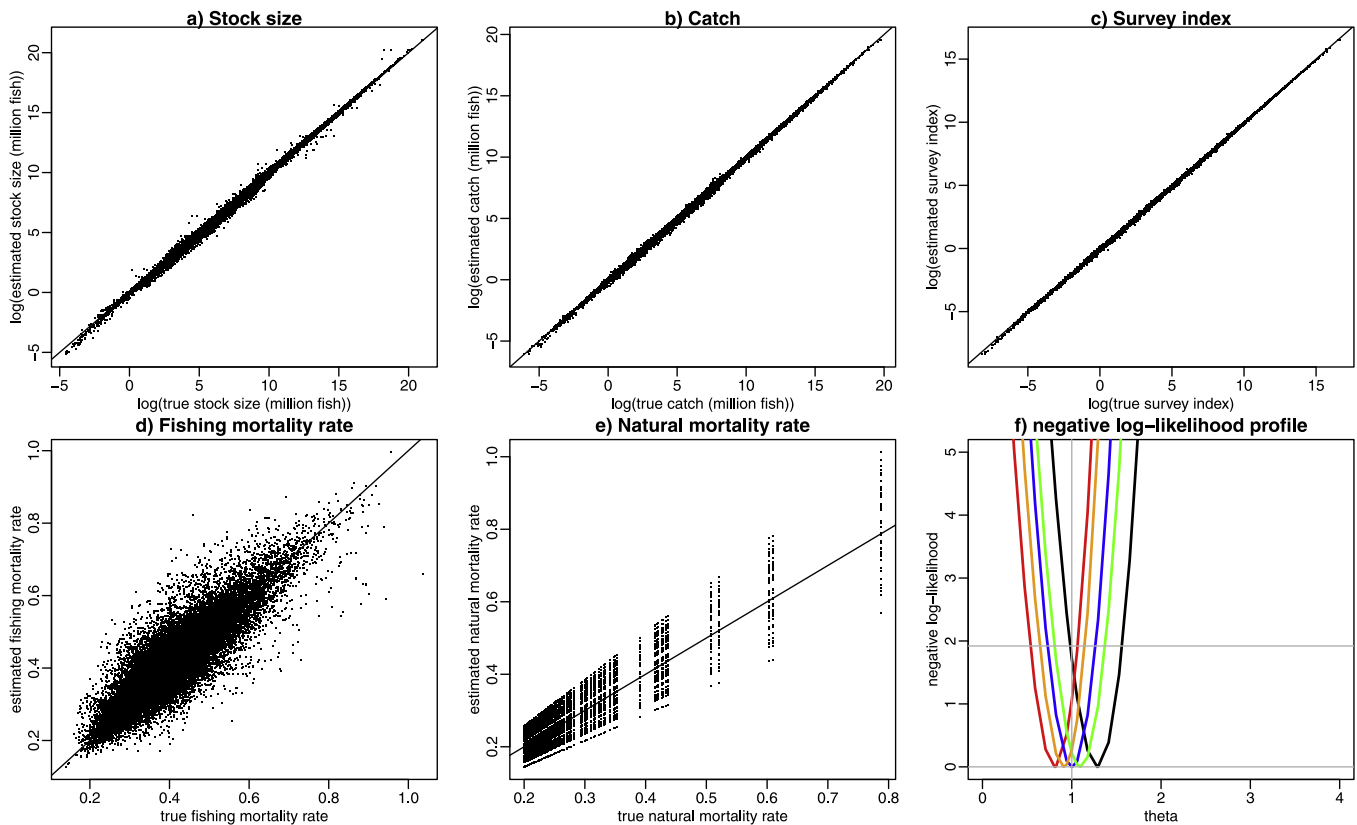

Fig. 2. Estimated vs. true values and negative log-likelihood profiles for a simulation experiment based on the Northeast Atlantic Cod stock, with true $M_{a,y} = M^*_{a,y}$. (a) stock size, (b) catch, (c) survey index, (d) fishing mortality rate, (e) natural mortality rate, (f) negative log-likelihood profiles for five simulations. The solid black lines in panels a)-e) indicate the 1-1 relationships. Each negative log-likelihood profile in panel f) is subtracted by the minimum value of the corresponding negative log-likelihood, so all curves have minimum value 0. The vertical grey line in panel e) indicates the true value $\theta = 1$ and the horizontal grey line is set at 1.92.
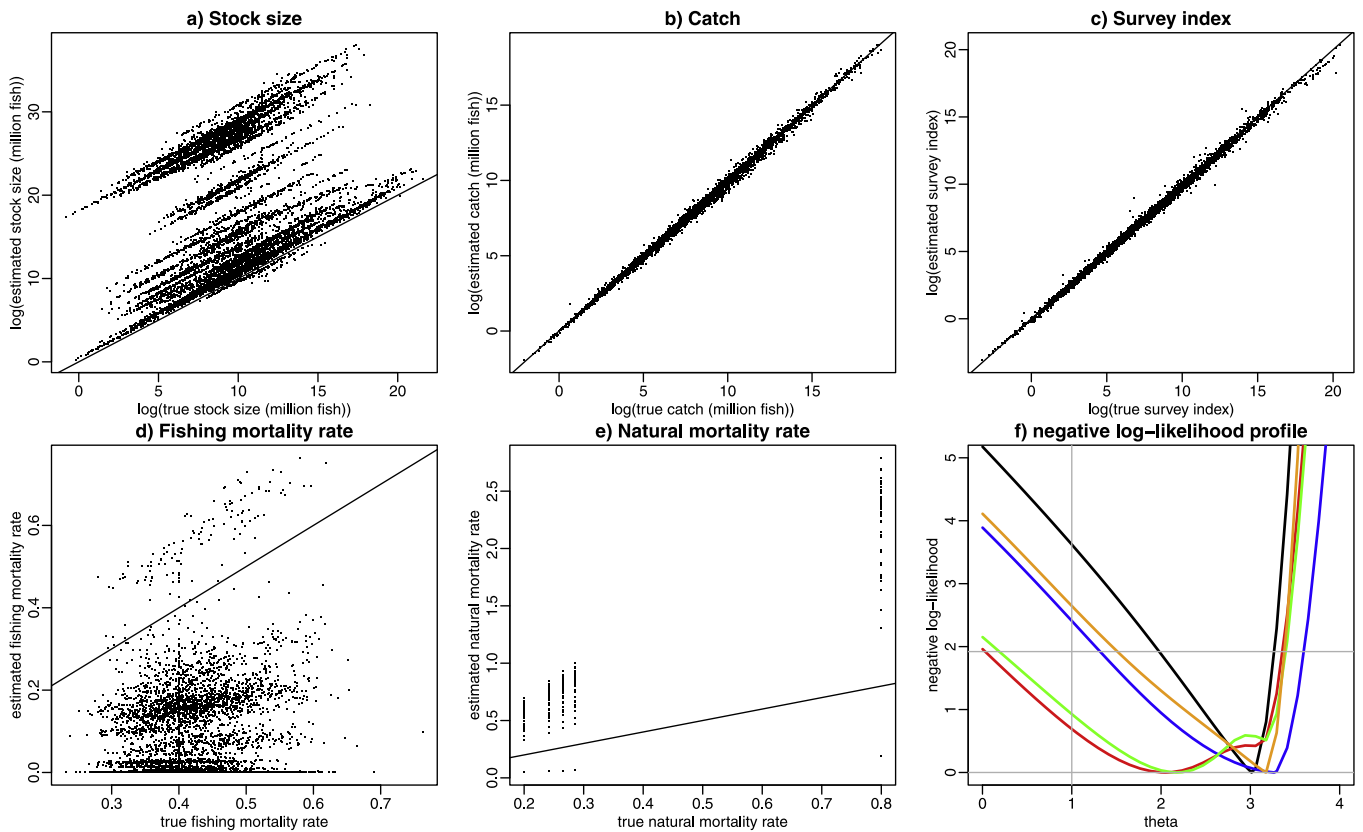
Fig. 3. Estimated vs. true values and negative log-likelihood profiles for a simulation experiment based on the Baltic Sea Cod stock, with true $M_{a,y} = M_{a,y}^*$. See Fig. 2 for explanation of the various panels.

It turns out that for NEAC, MAD($log(\widetilde{\theta})$) becomes 0.12, and for BSC it becomes 1.02. As a consequence, also other quantities are more precisely estimated for NEAC than for BSC (panels a)-e) in Figs. 2 and 3).

The negative log-likelihood profile for θ is a function of θ, where all the parameters are optimised for the given value of θ. For a given data set, here given by each simulation, one can use the negative log-likelihood profile to validate whether θ can be estimated with sufficient precision or not. Panels f) in Figs. 2 and 3 show the negative log-likelihood profiles for θ for five different simulations for NEAC and BSC. The maximum likelihood estimates of θ are given at the bottom of each curve. The difference in -2 log-likelihood between a model where θ is estimated and a model where θ is fixed to 1 is asymptotically chi square distributed with 1 degree of freedom. Therefore, if the negative log-likelihood curve is above 1.92 (the horizontal grey line) for $\theta = 1$ (the vertical grey line), the corresponding estimate of θ is significantly different from 1 at 5% level. For NEAC (Fig. 2) each of the five curves are quite narrow and below 1.92 for $\theta = 1$, i.e. none of the fives estimates are significantly different from the true value. On the other hand, for BSC (Fig. 3) the curves are much wider, all estimates of θ are larger than 1 and three out of five estimates are significantly different from the true value.

Based on the theoretical results in Section 2.3 and the simulation study, we conclude that it is in principle possible to estimate a fixed $M_{a,y}$ common for all years and ages, or to estimate θ in the model $M_{a,y} = \theta \cdot M_{a,y}^*$, based on catch at age and abundance at age data only, but to get a reasonable precision, one needs many observations. For many ICES stocks today, there are probably not enough observations to estimate M precisely, but this may change in the future.

4. Empirical study based on 19 fish stocks

4.1. Empirical study for six models for $M_{a,y}$

We perform an empirical study for selected sub models for $M_{a,y}$, to investigate the practical consequences of the theoretical findings in Section 2.3 and the simulation experiment in Section 3. We use data from the same 19 fish stocks as in Section 3. The model is the same as the one used in Section 2.3, except that we specify six different sub models for the natural mortality rate $M_{a,y}$:

- Model M0: $M_{a,y} = M_{a,y}^*$, where $M_{a,y}^*$ is fixed and based on external information or "best guess", separate for each fish stock.
- Model M1: $M_{a,y} = \theta M_{a,y}^*$, where θ is a positive parameter to be estimated. This is the same model as were used in the simulation experiment in Section 3, defined by Eq. (13).
- Model M2: $M_{a,y} = M_{a,y}^* + \varepsilon_{a,y}^M$, where $\varepsilon_{a,y}^M \sim N(0, \sigma_{M,a}^2)$, but where $\varepsilon_{a,y}^M$ is truncated at $\pm M_{a,y}^*$ to ensure that the natural mortality rate is non-negative. An alternative would be to use a multiplicative model with lognormal errors, but we have previously experienced that this can give severe convergence problems (Aldrin et al., 2020).
- Model M3: As model M2, but where $\varepsilon_{a,y}^M$ follow an AR(1) process in time, i.e. $\varepsilon_{a,y}^M = \rho \varepsilon_{a,y-1}^M + \omega_{a,y}$, where $\omega_{a,y} \sim N(0, \sigma_{o,a}^2)$ and $\omega_{a,y}$ is correlated over ages in the same way as $\varepsilon_{a,y}^F$.
- Model M4: As model M2, but with unrestricted $M_{a,y}$, i.e. $M_{a,y}$ may be negative, and thus account for possible immigration.
- Model M5: As model M4, but where $\varepsilon_{a,y}^M$ is correlated in time and between ages as in model M3.

We compare the performance of these sub models by investigating their prediction performance based on cross validation, and by studying
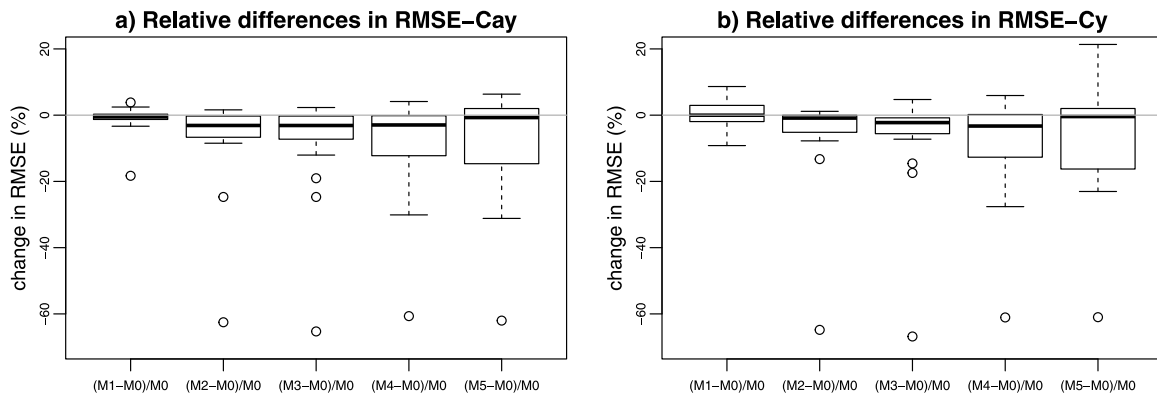
Fig. 4. Box plots with percentage changes in cross-validated prediction errors for catch by changing from model M0 to models M1-M5, for age-and-year specific errors $RMSE^{Cay}$ (panel a) and for year specific errors $RMSE^{Cy}$ (panel b).

how their estimated stock sizes differ.

4.2. Cross validation

We perform cross validation for each of the 19 data sets. We omit data for one year at a time, giving a training data set. We estimate the model on this training data and predict the catches and indexes for the year omitted. Data for each year are left out once, except for the first year, which are always included in the training data. We calculate two root-mean-square error measures of prediction performance for catch, one per age and year ($RMSE^{Cay}$), and another for total catch in a year ($RMSE^{Cy}$):

$$\text{RMSE}^{\text{Cay}} = \sqrt{\left(1 \Big/ n^{\text{Cay}}\right) \sum_y \sum_a \left(\widehat{C}_{a,y} - \widetilde{C}_{a,y,(-y)}\right)^2}, \tag{17}$$

$$\text{RMSE}^{\text{Cy}} = \sqrt{\left(1 \Big/ n^{\text{Cy}}\right) \sum_y \left(\sum_a \widehat{C}_{a,y} - \sum_a \widetilde{C}_{a,y,(-y)}\right)^2}, \tag{18}$$

where n is the number of elements in each sum, and the subscript $(-y)$ means that catch and index data for year y was omitted for estimation. We also compute two corresponding measures for the indices.

4.3. Results for the empirical study

Fig. 4 shows the relative changes in the cross validated root-mean-square errors $RMSE^{Cay}$ and $RMSE^{Cy}$ for catch, when changing from
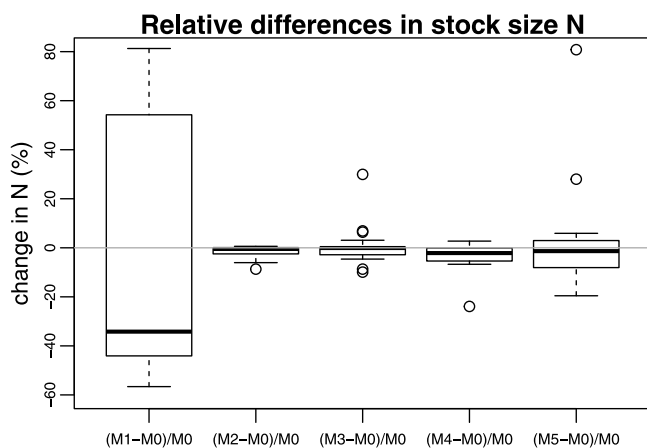

Fig. 5. Experiment 1: Percentage change in estimated average total stock size by changing from model M0 to models M1-M5.

model model M0 ($M_{a,y} = M_{a,y}^*$) to each of the others. Estimating the level of the natural mortality rate by multiplying $M_{a,y}^*$ by a factor (model M1), does not systematically improve the predictions. On the other hand, to allow for independent random variations around a fixed expected value equal to $M_{a,y}^*$ (model M2 for non-negative $M_{a,y}$ and model M4 for unrestricted $M_{a,y}$) tend to improve the predictions, and in one case RMSE is reduced by 60% (cod by Iceland and East Greenland). Extending the model for the random variations by including correlations between years and between ages (models M3 and M5) may improve the predictions further in some cases but make them more imprecise in other cases. The results for predicting the survey indices are similar (Fig. 1 in Supplementary Material).

The estimated total stock size can be changed significantly in both directions using model M1 instead of M0 (Fig. 5). This is consistent with the results from the simulation experiment in Section 3, where the estimated natural mortality rate could be very low or very high. For instance, an overestimated natural mortality can appear together with an underestimated fishing mortality rate and an overestimated stock size (e.g. panels a), d) and e) in Fig. 3). However, both the natural mortality rates and the stock sizes are in reality unknown, so we can not from this alone claim that the estimates from model M1 are wrong. But estimates of the natural mortality rates that are very far from the pre-guessed values $M_{a,y}^*$ may be considered implausible. Using models M2-M5 tend to give slightly lower estimates of stock size than model M0, but the results are more variable for the two most complex models M3 and M5.

5. Conclusions

We present a general stock assessment model that is to be estimated on catch and survey index data only. We investigate whether one should try to estimate the year and age specific natural mortality rates $M_{a,y}$ or fix them at pre-specified values $M_{a,y}^*$.

It is theoretically possible to estimate a common natural mortality rate for all years and ages using only catch-at-age data and abundance indices at age. However, a common $M_{a,y} = M$ for all years and ages is not a realistic assumption for many fish stocks, and may fail to detect changes in stock sizes if its contribution to the total mortality varies, i.e. $M_{a,y} \neq M$. It is also theoretically possible to estimate the natural mortality rate in models where the pre-specified $M_{a,y}^*$ is scaled by an estimated factor. However, for both models a large number of observations is required to achieve a reliable estimate of M or θ. We think that for many of the fish stocks presently available one should avoid to estimate the level of the natural mortality rate. Instead, without other data-sources, we believe it is more useful to treat the pre-specified $M_{a,y}^*$'s as fixed, *expected,* values of the natural mortality rates, and allow independent, random variations around these. This gives improved predictions, which indicates that such models also will give more precise

estimates of the unobservable quantities we are interested in, such as abundance at age or spawning stock biomass. We believe it is often not worthwhile to build more complex models for these random variations in the natural mortality rates, unless the data contains a large amount of observations or additional data is available for use. However, as the times goes by, more data will be collected, so the value of more complex models will increase in the future.

Credit author statement

Magne Aldrin: Writing – Original Draft; Conceptualization; Methodology; Software; Visualization.

Fredrik Lohne Aanes: Writing – Original Draft; Methodology; Software.

Ingunn Fride Tvete: Writing – Review & Editing; Data Curation.
Sondre Aanes: Writing – Review & Editing; Conceptualization.
Samuel Subbey: Writing – Review & Editing; Conceptualization.

Conflict of interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.fishres.2021.106071.

References

Aanes, S., Engen, S., Sæther, B.E., Aanes, R., 2007. Estimation of the parameters of fish stockdynamics from catch-at-age data and indices of abundance: can natural and fishing mortality be separated? Can. J. Fish. Aquat. Sci. 64, 1130–1142. https://doi.org/10.1139/f07-074.

Aldrin, M., Tvete, I., Aanes, S., Subbey, S., 2020. The specification of the data model part in the sam model matters. Fish. Res. 229 https://doi.org/10.1016/j.fishres.2020.105585.

Allen, S., Satterthwaite, W., Hankin, D., Cole, D., Mohr, M., 2017. Temporally varying natural mortality: sensitivity of a virtual population analysis and an exploration of alternatives. Fish. Res. 185, 185–197. https://doi.org/10.1016/j.fishres.2016.09.002.

Bogstad, B., Haug, T., Mehl, S., 2000. Who eats whom in the barents sea? NAMMCO Scientific Publications 2, 98–119. https://doi.org/10.7557/32975.

Cadigan, N., 2016. A state-space stock assessment model for northern cod, including under-reporting catches and variable natural mortality rates. Can. J. Fish. Aquat. Sci. 73, 296–308. https://doi.org/10.1139/cjfas-2015-0047.

Catchpole, E., Kgosi, P., Morgan, B., 2001. On the near-singularity of models for animal recovery data. Biometrics 57, 720–726. https://doi.org/10.1111/j.0006-341X.2001.00720.x.

Cole, D., 2020. Parameter Redundancy and Identifiability. Chapman and Hall/CRC.

Cole, D., McCrea, R., 2016. Parameter redundancy in discrete state-space and integrated models. Biom. J. 58, 1071–1090. https://doi.org/10.1002/Fbimj.201400239.

Cole, D., Morgan, B., 2010. A note on determining parameter redundancy in age-dependent tag return models for estimating fishing mortality, natural mortality and selectivity. J. Agri. Biol. Envir. Stat. 15, 431–434. https://doi.org/10.1007/s13253-010-0026-6.

Cook, R.M., 2004. Estimation of the age-specific rate of natural mortalityfor shetland sandeels. ICES J. Mar. Sci. 61, 159–164. https://doi.org/10.1016/j.icesjms.2003.11.005.

Gudmundsson, G., 1994. Time series analysis of catch-at-age observations. J. Roy. Stat. Soc, C - App. 43, 117–126. https://doi.org/10.2307/2986116.

ICES, 2017. Report of the Working Group on Widely Distributed Stocks (WGWIDE), 30 August - 5 September 2017. Technical Report ICES CM 2017/ACOM:23. ICES. Copenhagen, Denmark. ICES, Copenhagen, Denmark, 994 pp. http://www.ices.dk/sites/pub/Publication Reports/Expert Group Report/acom/2017/WGWIDE/01 Report of the Working group on Widely Distributed Stocks (WGWIDE).pdf.

ICES, 2020a, 2020a. Arctic Fisheries Working Group (AFWG). Technical Report ICES Scientific Reports. 2:52. ICES. Copenhagen, Denmark. https://doi.org/10.17895/ices.pub.6050, 577 pp.

ICES, 2020b, 2020b. Working Group on Widely Distributed Stocks (WGWIDE). Technical Report ICES Scientific Reports. 2:82. ICES. Copenhagen, Denmark. https://doi.org/10.17895/ices.pub.7475, 1019 pp.

Jiang, H., Pollock, K., Brownie, C., Hightower, J., Hoenig, J., Hearn, W., 2007. Age-dependent tag return models for estimating fishing mortality, natural mortality and selectivity. J. Agri. Biol. Envir. Stat. 12, 177–194. https://doi.org/10.1198/108571107X197382.

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., Bell, B., 2016. Tmb: automatic differentiation and Laplace approximation. J. Stat. Softw. 70, 1–21. https://doi.org/10.18637/jss.v070.i05.

Lee, H.H., Maunder, M., Piner, K., Methot, R., 2011. Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. Fish. Res. 109, 89–94. https://doi.org/10.1016/j.fishres.2011.01.021.

Lele, S., Nadeem, K., Schumland, B., 2010. Estimability and likelihood inference for generalized linear mixed models using data cloning. J. Am. Stat. Assoc. 105, 1617–1625. https://doi.org/10.1198/jasa.2010.tm09757.

Magnusson, A., Hilborn, R., 2007. What makes fisheries data informative? Fish Fish. 8, 337–358. https://doi.org/10.1111/j.1467-2979.2007.00258.x.

Nater, C., Vindenes, Y., Aass, P., Cole, D., Langangen, O., Moe, S., Rustadbakken, A., Turek, D., Vollestad, L., Ergon, T., 2020. Size- and stage-dependence in cause-specific mortality of migratory brown trout. J. Anim. Ecol. 89, 2122–2133. https://doi.org/10.1111/1365-2656.13269.

Nielsen, M., Berg, C., 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158, 96–101. https://doi.org/10.1016/j.fishres.2014.01.014.

Polansky, L., Newman, K., Mitchell, L., 2021. Improving inference for nonlinear state-space models of animal population dynamics given biased sequential life stage data. Biometrics 77, 352–361. https://doi.org/10.1111/biom.13267.

Punt, A., Castillo-Jordán, C., Hamel, O., Cope, J., Maunder, M., Ianelli, J., 2021. Consequences of error in natural mortality and its estimation in stock assessment models. Fish. Res. 233, 105759. https://doi.org/10.1016/j.fishres.2020.105759.

Quinn, T., Deriso, R., 1999. Quantitative Fish Dynamics. Oxford University Press, Oxford, pp. 355–362.

Raue, A., Kreutz1, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25, 1923–1929. https://doi.org/10.1093/bioinformatics/btp358.

Stock, B., Xu, H., Miller, T., Thorson, J., Nye, J., 2021. Implementing two-dimensional autocorrelation in either survival or natural mortality improves a state-space assessment model for southern new england-mid atlantic yellowtail flounder. Fish. Res. 237, 105873. https://doi.org/10.1016/j.fishres.2021.105873.

stockassesment.org, 2019. https://www.stockassessment.org/login.php. Accessed 20. October 2019.

stockassesment.org, 2020. https://www.stockassessment.org/login.php. Accessed 25. May 2020.