**Canadian Journal Of Fisheries And Aquatic Sciences**
November 2012, Volume 69 (11), Pages 1881-1893
http://dx.doi.org/10.1139/f2012-103




# A random effects population dynamics model based on proportions-at-age and removal data for estimating total mortality

Verena M. Trenkel[a, *], Mark V. Bravington[b], Pascal Lorance[a]

[a] Ifremer, rue de l'Île d'Yeu, BP 21105, 44311 Nantes cedex 3, France.
[b] CSIRO, GPO Box 1538, Hobart, Tasmania.

*: Corresponding author : Verena Trenkel, verena.trenkel@ifremer.fr

**Abstract:**

Catch curves are widely used to estimate total mortality for exploited marine populations. The usual population dynamics model assumes constant recruitment across years and constant total mortality. We extend this to include annual recruitment and annual total mortality. Recruitment is treated as an uncorrelated random effect, while total mortality is modelled by a random walk. Data requirements are minimal as only proportions-at-age and total catches are needed. We obtain the effective sample size for aggregated proportion-at-age data based on fitting Dirichlet-multinomial distributions to the raw sampling data. Parameter estimation is carried out by approximate likelihood. We use simulations to study parameter estimability and estimation bias of four model versions, including models treating mortality as fixed effects and misspecified models. All model versions were, in general, estimable, though for certain parameter values or replicate runs they were not. Relative estimation bias of final year total mortalities and depletion rates were lower for the proposed random effects model compared with the fixed effects version for total mortality. The model is demonstrated for the case of blue ling (Molva dypterygia) to the west of the British Isles for the period 1988 to 2011.

**Résumé:**

Les courbes des captures sont largement utilisées pour estimer la mortalité totale des populations marines exploitées. Le modèle de dynamique de population habituel suppose un recrutement et une mortalité totale constants au cours des années. Nous développons ce modèle pour y inclure un recrutement et une mortalité totale variables entre années. Le recrutement est traité comme un effet aléatoire non corrélé tandis que la mortalité totale est traitée comme une marche aléatoire. Les besoins en données sont minimaux car seules les proprotions aux âges et les captures totales sont nécessaires. La taille effective de l'échantillon des données agrégées de proportion aux âges est obtenue en ajustant une distribution Dirichlet-multinomiale aux données brutes d'échantillonnage. L'estimation des paramètres est réalisées par vraisemblance. Des simulations ont été utilisées pour étudier l'estimabilité des paramètres et les biais d'estimation de quatre versions du modèle, dont des modèles traitant la mortalité comme des effets fixes et des modèles avec hypothèses fausses. Toutes les versions du modèle étaient en général estimables sauf pour certaines valeurs de paramètres ou certaines réalisations. Les biais relatif de l'estimation de la mortalité totale de la dernière année et du taux de réduction de la population étaient plus faibles pour le modèle à effet aléatoire proposé que pour les versions à effets fixes pour la mortalité totale. Une application du modèle à la lingue bleue (*Molva dypterygia*) de l'Ouest des Îles Britanniques pour la période de 1988 à 2011 est présentée.

# Introduction

Fisheries stock assessments make use of a range of methods to obtain estimates of the status of exploited stocks which match the diversity of information available. Catch curves and year class curves have been part of the tool box from an early stage (Beverton and Holt 1957; Chapman and Roson 1960; Hilborn and Walters 1992). While catch curves use data from a single year, year class curves follow cohorts in time. Conditional on a few assumptions they allow to estimate total mortality $Z$ of exploited populations based on only numbers or frequencies-at-age, commonly derived from commercial catch data, and the standard population dynamics model describing changes in numbers-at-age $a$, $N_a = N_{a-1}\exp(-Z)$. The limited data requirements are probably responsible for their continuous use for data-poor stocks, but come at the price of strong assumptions. In the original catch curve formulation equilibrium conditions are assumed, i.e. recruitment and total mortality are assumed fixed during the range of ages and years considered and gear selectivity is constant across all considered age classes and years (Chapman and Robson 1960). Consequently numbers at age from a single year are sufficient for estimation.

Early on a range of estimators for $Z$ were developed based on different statistical distributional assumptions (Chapman and Robson 1960). These estimators have been shown to have different degrees of robustness in case of stochastic variability in recruitment, total mortality or age estimation (Dunn et al. 2002). Instead of investigating what happens when assumptions break down, several authors have extended the model to directly allow variations in recruitment, unequal selectivity across age or age varying mortality. For example, Cotter et al. (2007) introduced an age-dependent selectivity term and allowed recruitment to vary between years by using year class curves which are fitted by cohort to numbers- or proportions-at-age. They introduced a polynomial function for $Z$ to allow variation with time and/or age. The data used are catch per unit effort per age (cpue); cpue per age and year are treated as independent. Schnute and Haigh (2007) also allowed varying recruitment. They did this by introducing additional parameters for ages (year classes) for which recruitment was much higher than some average value. In their model the user specifies these age classes, e.g. age 3 and 5. The main parameter of interest remains total mortality $Z$, which is estimated for each annual catch curve (data set). This is somewhat inconsistent as it is assumed that total mortality is constant during the $A$ preceding years corresponding to the $A$ age classes considered. Schnute and Haigh (2007) also included age-specific selectivity. Using only group (aggregated age classes) composition data they noted that not all parameters were estimable. Similarly, Wayte and Klaer (2010) accounted for selectivity changes with age and fitted the catch curve simultaneously to several years of data assuming again constant mortality during that period. Finally, Thorson and Prager (2011) let natural mortality decrease with age in addition to increasing selectivity with age; they found the selectivity aspect was more important in their simulation study compared to natural mortality changes with age. Overall these recent catch curve developments have in common that total mortality is still assumed constant over some time period though other assumptions have been relaxed.

In this manuscript we introduce a new class of models, called multi-year catch curves (MYCC) that allow both recruitment and total mortality to vary in time. MYCC combine the annual view of traditional catch curves and the cohort view of year class models and are formulated using the state space framework which includes both process and observation error. As there are quite a number of parameters to be estimated, we add an additional data source, total catch in numbers and use random effects to achieve parsimonious models. Traditionally catch data has been augmented by effort time series to ensure parameter estimability, e.g. Paloheimo (1958), Deriso et al. (1985), Gudmundsson (1986). However, fishing effort is notoriously difficult to estimate and its relationship with catches or fishing mortality is not necessarily linear, making it a difficult data source. Therefore total catches were used here. Using random effects leads however to the need to estimate the surplus variance in some way, which is rather difficult. We propose to get a handle on this by binding the observation error variance via the effective sample size of the multinomial distribution describing the aggregated observed catch numbers-at-age. A common characteristics of aggregated compositional data is that they are overdispersed with respect to a multinomial distribution. This leads to the notion of effective sample size which corresponds to the sample size for which the variance in a multinomial distribution would be equal to the observed variance, e.g. Pennington et al. (2002). There are several reasons for the overdispersion in aggregated (catch) numbers-at-age: multi-level sampling such as from different hauls, seasons or vessels combined with schooling of similar sized fish and seasonal differences in spatial and depth distributions, and model misspecification in the case of stock assessment models for which the multinomial distribution is assumed to describe the observation process, see review by Maunder (2011),; Hulson et al. (2011). Recently several authors have compared the performance of different methods for estimating the effective sample size using simulations and real data (Candy 2008; Hulson et al. 2011; Maunder 2011). Maunder (2011) concluded that effective samples size was only an issue if it was five times smaller or more than the actual sample. Fitting a Dirichlet distribution produced the least biased estimates of effective sample size, though the difference with the other three tested methods was rather small. Here we used the effective sample size as a means to weigh numbers-at-age when aggregating them across samples, to calculate the effective sample size for the aggregated data set and subsequently to bind the observation error variance in the MYCC for the aggregated numbers-at-age data.

The proposed MYCC is demonstrated for blue ling (*Molva dypterygia*) in the North-east Atlantic which is a deep-water species with a longevity similar to cod. Little data are available, in particular no systematic scientific survey is carried out and commercial fisheries derived data are therefore the main data source for stock assessment and management (Lorance et al. 2010). The majority of catches are taken with bottom trawls primarily by French vessels (Lorance et al. 2010).

The next section introduces the approach used for estimating effective sample sizes of numbers-at-age samples for aggregating them prior to model fitting followed by an introduction to the MYCC model formulation. The salient features of MYCC are restricted data needs, only proportions-at-age and total catch numbers but no abundance indices nor effort data are required, and the use of random effects for total mortality and recruitment. Parameter estimability is then discussed and options for achieving it are studied by simulation. Finally the model is applied to the case of blue ling to obtain annual total mortality estimates.

# Materials and Methods

## Aggregation of correlated multinomial samples

Numbers-at-age samples are either obtained directly by random sampling from the target population or by combining length-at-age samples with age-length keys. In the simplest case one sample per quarter is available for each data set. Considering numbers-at-age samples, the sample $\mathbf{y}_i = (y_{i1}, \dots y_{iA})$ is presumed to be drawn from a multinomial distribution with underlying probabilities $\mathbf{p}_i = (p_{i1}, \dots p_{iA})$ for age class $a = 1, \dots A$, and the vector $p_i$ to be drawn from a distribution with the same underlying means $\pi = (\pi_1, \ldots \pi_A)$ for all samples. As a result several numbers-at-age data sets form correlated multinomial samples. If the probabilities for each class come from a Dirichlet distribution, $\mathbf{y}_i$ follows a Dirichlet-multinomial distribution. The Dirichlet-multinomial distribution is a compound multivariate distribution which has probability function

$$P\left(Y_1 = y_1, \ldots Y_A = y_A\right) = \frac{y_+!}{y_1! \ldots y_A!} \frac{\prod_{a=1}^{A} \prod_{r=1}^{y_a} \pi_a(1-\theta)+(r-1)\theta}{\prod_{r=1}^{y_+} 1-\theta+(r-1)\theta} \qquad (1)$$

where $y_+ = \sum_{a=1}^{A} y_a$ is the total sample size and $\theta$ the overdispersion parameter.

To aggregate several numbers-at-age samples, the proportions-at-age of each data set are weighed by the inverse of their variance. This leads to the estimator of the mean proportions-at-age

$$\widehat{\pi}_a = \sum_i w_i \left(\frac{y_{ia}}{y_{i+}}\right) / \sum_i w_i \qquad (2)$$

where $w_i \propto 1/V\left[y_{ia}/y_{i+}\right]$, i.e. the weight is proportional to the inverse of the variance of each sample proportion. From the Dirichlet-multinomial distribution and setting $\theta = 1/(1+\alpha)$ in the notation of Johnson et al. (1997)

$$V\left[y_{ia}\right] = y_{i+}\pi_a(1-\pi_a)(1+(y_{i+}-1)\theta) \quad (3)$$

and thus

$$1/V\left[y_{ia}/y_{i+}\right] = y_{i+}(1+(y_{i+}-1)\theta)^{-1}(\pi_a(1-\pi_a))^{-1} \ (4).$$

Combining (2) and (4) provides the final estimator for the aggregated proportion-at-age

$$\widetilde{\pi}_a = \sum_i y_{i+}\left(1+(y_{i+}-1)\theta\right)^{-1} \frac{y_{ia}}{y_{i+}} / \sum_i y_{i+}\left(1+(y_{i+}-1)\theta\right)^{-1} = \sum_i \widetilde{m}_i \frac{y_{ia}}{y_{i+}} / \widetilde{m} \ (5)$$

with $\widetilde{m}_i = y_{i+}\left(1+(y_{i+}-1)\theta\right)^{-1} \qquad (6)$

and $\widetilde{m} = \sum_i \widetilde{m}_i$ the overall effective sample size.

The aggregated sample for age $a$ class is then simply $\widetilde{y}_a = \widetilde{m}\widetilde{\pi}_a$ with $V[\widetilde{y}_a] = \widetilde{m}\widetilde{\pi}_a(1 - \widetilde{\pi}_a)$. This means that the Dirichlet-multinomial distribution of the raw data $\mathbf{y}$ has the same mean and variance as the pure multinomial likelihood of the aggregated sample vector $\widetilde{\mathbf{y}}$ with sample size $\widetilde{m}$.

Applying the estimator in (5) assumes that the overdispersion parameter $\theta$ is known. In reality it is unknown but can be estimated by maximum likelihood. For the case study the dirmult package in R (Twedebrink 2009) was used to fit the Dirichlet-multinomial distribution and estimate $\theta$ by year.

The above procedure of combining correlated multinomial data samples is generic and can be applied sequentially, e.g. to combine first several numbers-at-age samples from sampling different vessels in a given month and then aggregated samples from different months. The final effective sample size of each stage is then used as input into the next stage. In the case of separate samples of numbers-at-length and age-length keys, samples of each type can be aggregated first before combining them into a single numbers-at-age data set which maintains the appropriate variance structure in the multinomial distribution via the final effective sample size. This procedure is equivalent to the common practice of conditioning inferences on a point estimate of variance. In summary, aggregating numbers-at-age or length-at-age samples prior to model fitting allows to externalize the propagation of sampling variance without confounding potential model misspecification with the treatment of sampling uncertainty.

# Multi-year catch curves (MYCC)

In MYCC population dynamics in numbers are modelled as

$$N_{a,t} = N_{a-1,t-1}e^{-Z_{t-1}} \qquad a_r < a < A_+ \qquad t = 1 \ldots T \quad (7)$$

$$N_{A+,t} = (N_{A_+-1,t-1} + N_{A+,t-1})e^{-Z_{t-1}} \qquad t = 1 \ldots T \quad (8)$$

where $N_{a,t}$ are population numbers at age $a$ in year $t$, $A_+$ is an age plus group and $Z_t$ are annual total mortality rates which are constant across ages. Recruitment at age $a_r$ is assumed to vary randomly over time following a log-normal distribution, similar to other authors, e.g. Deriso et al. (1985)

$$N_{a_r,t} = R_t \qquad R_t \sim logN(\log(\mu_R), \sigma_R) \qquad t = 1 \ldots T \qquad (9)$$

where $\mu_R$ is the mean recruitment on the lognormal scale and $\sigma_R$ the standard deviation on the base normal (log) scale. For ease of interpretation the coefficient of variation ($CV_R$) instead of $\sigma_R$ is used making use of the fact that $var(ln(x)) \approx ln(CV(x)^2 + 1)$. As recruitment is a latent variable, it is treated as an uncorrelated random effect.

Annual total mortality $Z_t$ can either be modelled by $T$ parameters, i.e. treating it as a fixed effect or by a random effect. In many circumstances a reasonable approach is to use a random walk as proposed by Gudmundsson (1986; 1994)

$$Z_t = Z_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma_Z) \qquad t = 1 \ldots T \quad (10)$$

The initial state vector $N_{a,1}$ at the beginning of year $t = 1$ is defined assuming constant historic total mortality $Z0$ and variable recruitment for years $t < 1$

$$N_{a,1} = e^{(a_r-a)Z0}R_{1-a+a_r} \qquad a_r < a < A+ \qquad (11)$$

The initial number of fish in the plus group $N_{A+,1}$ is obtained by an infinite sum over preceding years and assuming constant average recruitment during that period

$$N_{A+,1} = \sum_{t=A}^{\infty} e^{-(t-a_r)Z0}\mu_R = \frac{e^{(a_r-A)Z0}}{1-e^{-Z0}}\mu_R \qquad (12)$$

The observation model has two parts, the first one for aggregated numbers-at-age $\widetilde{y}_{a,t}$ (see previous section), which are assumed to follow a multinomial distribution $\widetilde{y}_{a,t} \sim Multinom(p_{a,t}, \widetilde{m}_t)$ similar to Deriso et al. (1985). The corresponding probability density function is

$$f(\widetilde{y}_{a_r t}, \ldots \widetilde{y}_{A+,t} \mid p_{a_r t}, \ldots p_{A+,t}) = \frac{\widetilde{m}_t!}{\widetilde{y}_{a_r,t}! \ldots \widetilde{y}_{A+,t}!} (p_{a_r,t})^{\widetilde{y}_{a_r,t}} \ldots (p_{A+,t})^{\widetilde{y}_{A+,t}} \qquad a_r \leq a \leq A+ \quad t = 1 \ldots T \qquad (13)$$

where $p_{a,t} = N_{a,t} / \sum N_{a,t}$ are population proportions-at-age and $\widetilde{m}_t = \sum \widetilde{y}_{a,t}$ is the sample size of the aggregated data in year $t$. An important implicit assumption is that selectivity is constant across ages and years.

The second observation model is for the total catch $C_t$ (in numbers) which is assumed to follow a Gamma distribution with parameters $\alpha$ and $\beta$

$$C_t \sim Gamma(\alpha, \beta) \qquad t = 1 \ldots T \qquad (14)$$

$$E\left[C_t\right] = \left(\frac{Z_t - M}{Z_t}\right) (1 - e^{-Z_t}) \sum N_{a,t} \qquad (15)$$

where $E[C_t]$ is the expected catch in numbers. $M$ is natural mortality which is assumed constant over years and age classes. The coefficient of variation ($CV_c$) of the Gamma distribution is related to the $\alpha$ parameter as $CVc = \frac{1}{\sqrt{\alpha}}$ and $\beta = \frac{\alpha}{E[C_t]}$. As CVs are easier to handle, the model is parameterized in terms of $CV_c$. The probability density function for total catch is therefore

$$f(C_t \mid CV_c, E\left[C_t\right]) = \left(\frac{1}{E[C_t] CVc^2}\right)^{1/CVc^2} \frac{1}{\Gamma(1/CVc^2)} C_t^{1/CVc^2 - 1} e^{-1/CVc^2} \quad (16)$$

where $\Gamma()$ is the Gamma function.

# Parameter estimation

All model parameters $\Omega = (\mu_R, CV_{_R}, log(\sigma_Z), Z0, M)$ are estimated using maximum likelihood based on the observation vector $\mathbf{d} = (Y_{a_R,1}, \ldots, Y_{A+,T}, m_1, \ldots, m_T, C_1, \ldots C_T)$ which has conditional density $f_\Omega(\mathbf{d} \mid \mathbf{u}, \mathbf{v})$ where $\mathbf{u} = (R_1, ..., R_T)$ is the vector of the latent random recruitment variable (eq. 9) with marginal density $h_\Omega(\mathbf{u})$ and $\mathbf{v} = (Z_1, ..., Z_{T-1})$ is the total mortality random effects variable (eq. 6) with marginal density $g_\Omega(\mathbf{v})$.

The marginal likelihood function is obtained by integrating out $\mathbf{u}$ and $\boldsymbol{v}$ from the joint density

$$\mathcal{L}(\Omega) = \int\int f_\Omega(\mathbf{d} \mid \mathbf{u}, \mathbf{v})\ h_\Omega(\mathbf{u}) g_\Omega(\mathbf{v}) d(\mathbf{u}) d(\mathbf{v}) \qquad (17)$$

The double integral in eq. 17 is evaluated using the Laplace approximation as implemented in the random effects module of AD Model builder (Fournier et al. 2012) described in Skaug and Fournier (2006). AD Model builder automatically calculates standard deviations of estimates based on the observed Fisher Information matrix.

# Estimability of MYCC parameters

Depending on the data set not all MYCC parameters might be estimable which manifests itself by certain parameter estimates lying on the boundary imposed during the estimation process or the non-convergence of the estimation procedure. The main options for ensuring parameter estimability are to fix certain parameters, i.e. treat them as constants, reparameterize the model or a combination of both.

Depending on the application, for certain parameters it might be easy to identify suitable values to treat them as constant. One set of parameters easy to fix might be the total catch observation error $CV_c$ and historical total mortality $Z0$ (eq. 12), which could be set equal to natural mortality $M$ if the data starts from the beginning of the fishery and some reasonable estimate of $M$ were available. If an estimate of $M$ is available it can be assumed constant in eq. 15.

# Simulation studies

Two simulation studies were carried out to (i) explore the estimability of model parameters for model variants and (ii) evaluate the robustness to model misspecification. The following two models were compared in both studies:

RE-Z model: $Z$ random effect (4 parameters)

RE-Z & M model: $Z$ as random effect plus $M$ estimated (5 parameters)

In addition, in simulations study 1 the performance of the random walk formulation for total mortality was compared with the more traditional fixed effect approach in which each $Z_t$ is a separate independent parameter using the two models

FE-Z model: $Z$ fixed effect (2 + # years-1 parameters)

FE-Z & M model: $Z$ fixed effect plus $M$ estimated (3 + # years-1 parameters)

# Design

## Simulation study 1

To evaluate MYCC parameter estimability, the true population and the observation data (numbers-at-age and total catches) were simulated using the MYCC model and a full factorial design for four parameters, keeping natural mortality $M$ and mean recruitment $\mu_R$ constant (see Table 1). For each model parameter combination, two values for the number of ages $(A_+ - a_r + 1)$, number of years $T$ and aggregated sample size $m_t$, i.e. not simulating the sample aggregation process, were used. This lead to 32 distinct combinations (using the same for the two aggregated sample sizes). In each case 50 replicate data sets were created and the four model variants were fitted. Model performance was compared using the percentage of estimable replicates, i.e. no parameter estimates lying on the boundaries (see Table 1) and overall convergence, and relative estimation error (obs-true)/true of total mortality in the final year $Z_T$ and of the population depletion rate $N_T/N_1$. To investigate the impact of parameter values on parameter estimability and relative errors, regression tree analyses were carried out. Interquartile ranges across replicates were also calculated for the two relative estimation errors.

**Simulation study 2**

To evaluate the effects of model misspecification, three scenarios were compared. In the "Base" scenario the simulation model is the same as estimation model. In the "Rdec" scenario, mean recruitment $\mu_R$ decreases linearly over time, and finally in the "Sel" scenario selectivity for the numbers-at-age data is not constant but increasing for first two ages classes. The true populations were simulated using the parameters estimated for blue ling below (setting $a_r = 1$; $A_+ = 11$; $T = 24$), but without any missing data. Sample size for numbers-at-age was set to $m_t = 300$ for all years. For scenario "R-dec", $\mu_R$ was linearly reduced to 20% of the starting value at the end of the time period. For scenario "Sel", selectivity for numbers-at-age was set to 0.5 for the first age and 0.8 for the second, i.e. $y^{Sel}_{1,t} = 0.5y_{1,t}$ and $y^{Sel}_{2,t} = 0.8y_{2,t}$. For each scenario 50 replicate data sets were created and the two MYCC model variants were fitted. Model performance was compared using the percentage of estimable replicates, i.e. no parameter estimates lying on the boundaries (see Table 1) and overall convergence, and relative estimation errors for annual total mortality $Z_t$ and total abundance $N_t$.

# Simulation results

## Simulation study 1

The two values used for the sample size $m_t$ in the observation model (eq. 13) did not lead to different results so results are only shown for $m_t = 400$. The proportion of simulation runs leading to all parameters being estimable ranged from 24 to 100% depending on the parameter set (Figure 1, Table 2). Estimating natural mortality reduced parameter estimability for both the random effects (RE-Z) and fixed effects (FE-Z) model versions. On average fixed effects model had a slightly higher parameter estimability in the case of FE-Z model runs but lower when $M$ was also estimated in the FE-Z & M model (Table 2). In a few cases parameter estimates ended up on the boundary. In the RE-Z models the parameters to hit the lower boundary were $Z0$ and $M$. In the FE-Z models it was $Z0$, $M$ and $CV_R$ (Table 2).

The regression tree analysis showed that the most important parameter in terms of parameter estimability was the variance of the $Z$ random walk, here $log(\sigma_Z)$, for which cases with small value (-3) lead to more replicates with all parameters being estimable compared to those simulated with a large value (-1); the second influential parameter was the length of the time series $T$ (Figure 2). The value of $Z0$ played a role when estimating natural mortality in the RE-Z model with smaller $Z0$ values leading to more replicate runs with all model parameters being estimable. For the FE-Z & M model it was the length of the time series $T$ rather than $Z0$ that explained parameter estimability (higher proportion of estimability for longer times series).

Final year estimates for total mortality $Z_T$ and the depletion rate were estimated without bias in general for all parameter sets and model variants which can be seen from the fact that the interquartile range across the 50 replicates all included 0 (Figure 3). However, for certain parameter combinations large negative bias in $Z_T$ and positive bias in the depletion rate were observed in particular for the two fixed effect model variants. Using Spearmans rank correlation test (e.g. Conover 1971) a negative correlation between the relative bias of the two quantities was found($p<0.0001$). Inspection of the relative errors confirmed that random effect models lead to smaller bias in $Z_T$; relative bias was also smaller for simulation sets with smaller interannual variability in $Z$. No systematic parameter value effect was found for the relative bias of the population depletion rate.

## Simulation study 2

The percentage of estimable model runs for the two scenarios with model misspecification did no differ much from the base runs where the simulation model and the estimation models were identical (Table 2, final columns). In terms of relative estimation errors, halving mean recruitment over time (Rdec scenario) or assuming an increasing selectivity with age (Sel scenario) both lead to overestimation of total mortality, with increasing errors over time (Figure 4, left column) while the estimation error in total abundance was negative (Figure 4, right column). Overall relative estimation errors were larger for the Rdec scenario compared to the Sel scenario while estimates for the Base scenario were unbiased on average for total abundance (Figure 34b) and becoming slightly positively biased for total mortality at the end of the time period (Figure 4a).

# Application to blue ling

## Data

Three data sets were available for blue ling; all three come from commercial fishing operations. The first data set consists of annual international landings in weight for the area to the north and west of the British Isles (ICES subareas VI and VII, ICES division Vb) for the years 1966 to 2010. The second data set are numbers per 1-cm length group (length-frequency data set) per quarter from harbour sampling of French landings (1984-2010, no data in 1986 and 87). The third data set are proportions of ages-at-size (so called age-length keys) per quarter for the years 1991, 1992, 1993, 1994, 2009, and 2010, and on an annual basis for 1988 and 1995; again for samples from French landings only. Though blue ling exhibit sexual dimorphism with females growing larger, no sex information was available so both sexes had to be treated together.

Total annual landings in numbers were calculated by dividing landings in weight by the mean individual weight and multiplying by the proportion in weight of individuals aged 9 and older, corresponding to the age range considered here. Mean individual weight was calculated from the length-frequency data set by first transforming length into weight (in gram) using the relationship $W = 0.00191 * L^{3.14882}$ (Dorel 1986) and then averaging across individuals. The annual proportion of individuals older than 9 years in the landings was estimated from the length-frequency data set assuming a mean size for age 9 of 84 cm, which in turn was derived by combining length-frequency data with the age-length keys.

To obtain proportions-at-age, quarterly age-length keys were first multiplied by quarterly numbers per length class from the length-frequency data set and scaled to the sample size of the age length key; 5 cm length classes were used for this calculation. Then separate Dirichlet distributions were fitted to each annual set of quarterly numbers per age group and their effective sample size was estimated. Quarterly data sets were aggregated to an annual proportions-at-age data set as weighted average of quarterly values with effective sample size as weighing factor (see methods section). The annual aggregated sample size is the sum of quarterly effective sample sizes.

The RE-Z and RE-Z & M models were then fitted for the period 1988 to 2011 using the prepared annual aggregated proportions-at-age with the estimated aggregated sample sizes and total landings in numbers. Catch uncertainty was set to $CV_C = 0.02$, thus assuming transformed landings were reliable estimates of catch numbers. For natural mortality in the RE-Z model values of $M = (0.16, 0.17, 0.18)$ were tested. The upper value of 0.18 was obtained using Pauly's empirical formula (Pauly 1980) with growth parameters $K = 0.152$ and $L\infty = 125$ estimated for both sexes combined by Ehrich and Reinsch (1985). Residuals were examined to investigate model fit and the number of positive eigenvalues of the Hessian matrix at the maximum likelihood was checked to determine parameter identifiability. As the FE-Z models did not provide reliable estimates, no results are presented.

# Results

## Initial analyses

International blue ling landings reached their all time high in the late 1970 and decreased thereafter (Figure 5a). Mean individual weight in landings decreased from 1984 to the late 1990s, and more or less stabilized thereafter (Figure 5b). Years with low mean weight, e.g. 1998 and 2007, probably indicate strong recruitment. The proportion of individuals

>9 years in the landings followed the same time trend as mean weight (Figure 5b).

Proportions-at-age per quarter varied over time, with a higher proportion of older individuals earlier in the time series (Figure 6). For the analysis the model and data were restricted to the fully recruited age classes assumed to be from age 9 onwards based on visual inspection of figure 6. Further, ages >19 years were grouped into a 19+ group.

Annual effective sample size of numbers-at-age data sets obtained by fitting Dirichlet-multinomial distributions and aggregating data across quarters ranged from 130 to 458 which corresponds to 21 to 60% of the raw data (Table 3).

## Model results

All four model parameters, $\mu_R, CV_R$, $Z0$ and $log(\sigma_Z)$ were estimable for the RE-Z model but $M$ was not estimable in the RE-Z & M model; the value of $M$ was driven to the lower boundary (close to zero) during the estimation process. No convergence was achieved in the run using $M = 0.16$. When comparing the run with $M = 0.18$ to that with 0.17 a slightly larger likelihood was achieved for the later case. Therefore only results from the RE-Z fit with $M = 0.17$ are presented. The precision of estimated model parameters ranged from a coefficient of variation of 0.07 to 0.2 (Table 4).

Inspection of the posterior modes of the estimated random effects for total mortality (Figure 7a) and recruitment (Figure 7b) revealed no major deviations from assumptions. There was little evidence of autocorrelation in the residuals of total landings or any other pattern (Figure 7c); predicted landings were linearly related to observed total landings (Figure 7d). Residuals for proportions-at-age showed no year effect but a slight age effect with younger ages having more positive residuals and older ones more negative (Figure 7e). Predicted and observed proportions-at-age showed good agreement with slightly increasing differences as proportions increased (Figure 7f).

The estimated total mortalities started from 0.36 in the late 1980s, reached a peak of 0.56 around 2002 and decreased since then to $Z_{2010} = 0.26$ in 2010 (Figure 8). Total stock abundance for age 9+ decreased during the first half of the period coinciding with high fishing mortalities (assuming constant $M$) and increased slightly since about 2004. Recruit estimates (age 9) were highly variable over time and significantly autocorrelated ($R(1) = 0.51$). The uncertainty of estimated total mortality was highest during 1996-2005, which corresponds to the period with no proportion-at-age data available but only total landings, while estimates of $Z$ for 2006-2008 were somewhat more precise despite also a lack of proportions-at-age data.

## Interpretation for blue ling

All parameters except natural mortality were estimable for blue ling despite the large data gap in the available proportions-at-age data. Natural mortality was set to $M = 0.17$ for the final estimates which is smaller than the average estimated fishing mortality of 0.22 for the period 1988 to 2010 obtained when subtracting $M$ from the estimated total mortality values. However it is much larger than the estimated fishing mortality for 2010 ($F_{2010} = 0.093$) for ages 9 years and older.

Aggregated sample sizes for the blue ling number-per-age data sets were less than half the nominal sample size in most years; the large number of age classes might have contributed to this. Given these values and judging from the results obtained by Maunder (2011), effective sample size might not have been a big issue for blue ling. Also, little differences were found in the simulation study when values of 50 and 400 were compared.

Blue ling appear in commercial landings from about age 6, though their importance increases for up to age 9 (Figure 5). The main factor for this is probably a lack of availability rather than trawl selectivity. Young blue ling (≤6 years) have been reported in very small numbers only in surveys using small mesh trawls in the study area to the west of the British Isles (Bridger 1978; Ehrich 1983; Gordon and Hunter 1994). Further, these small individuals may belong to the closely related *Molva macrophthalma*, which was not considered as a separated species in the past (Whitehead et al. 1986). The only known nursery area for blue ling is located in Icelandic waters with probably some juveniles also occurring in Faroese waters, where blue ling below 30 cm have been caught in small numbers (Magnússon et al. 1997; Magnussen 2007). Juveniles blue ling are not known to occur to the west of the British Isles (F. Neat, personal communication, Marine Scotland-Science, Aberdeen, United Kingdom, 2011). As a consequence estimated recruitment also includes movement to the area the fisheries is operating in. Given that few young individuals are caught by the blue ling fishery, discards are minor or inexistent (ICES 2011) and consequently landings correspond to catches. Taking these elements together means that assuming constant catchability and selectivity of blue ling from age 9 by the commercial fishery and non age-specific total mortality beyond age 9 was reasonable. If however selectivity increased with age, the results of the second simulation study suggested that total mortality could be overestimated and total abundance underestimated. In terms of management this would mean that mortality and abundance estimates should be conservative.

The time series for total mortality estimates for blue ling aged 9 years to 19+ was found to be dome shaped with values around 0.6 in the early 2000s. Thus the management measures that were implemented from 2003 seemed to have been effective in that estimated total mortality started to decline and total abundance to increase from about 2005. Using haul by haul landings and effort data Lorance et al. (2011) derived an abundance index for blue ling in the same area which was stable from 2000-2007 and then increased generally in agreement with the current estimates for a much longer period. Total mortality estimates have been used in harvest control rules for managing relatively data-poor stocks (Wayte and Klaer 2010). The next step would now be to test harvest control rules for blue ling based on $Z_t$ estimates.

# Discussion

The proposed MYCC model is a statistical catch-at-age model similar to those that have been in use in stock assessments for decades, e.g.Paloheimo (1958), Doubleday (1976), and Deriso et al. (1985). It adopts a time series approach as pioneered by Gudmundsson (1994). Contrary to many traditional approaches, parameter estimation is by maximum likelihood and both process and observation errors are implemented in a state-space approach. A parsimonious formulation is achieved by using random effects, an approach which is increasingly being used in fisheries stock assessment models e.g. Fryer (2002), Trenkel (2008), and Nielsen (2009). Random effects have the advantage of being able to handle missing years of data and offer an appropriate way for dealing with latent variables such as recruitment and mortality. However, they come with the challenge of having to estimate a number of different variances. We mastered this challenge by setting the sample size of proportions-at-age data to appropriate values obtained externally by fitting Dirichlet-multinomial distributions before aggregating across samples and fixing (somewhat arbitrarily) the coefficient of variation for total catches. Depending on the application it might also be possible to use knowledge and common sense to fix one of the random effects variances, i.e. for recruitment or total mortality.

As shown by the first simulation study all MYCC model parameters were estimable in principle though natural mortality $M$ was the most difficult to estimate. The overall failure rate was increased from 11 to 35% when $M$ was tried to be estimated. This failure rate is comparable to what has been found for other age-structured models (Magnusson and Hilborn 2007) and indicates that $M$ might have to be fixed in practical applications; indeed this was the case here for blue ling.

The most important factor affecting parameter estimability was the value of the standard deviation of the total mortality random walk ($log(\sigma_Z)$) determining the interannual variability in total mortality if natural mortality was assumed known and historic total mortality $Z0$ when $M$ was also estimated. Small values of $Z0$, implying a larger ratio of $M/Z0$ lead to higher estimability of $M$. Time series length was the second (third when $M$ was estimated) most important factor. The model versions with fixed effects for total mortality (but keeping a random effect for recruitment) were generally equally estimable, though the dependence on the particular parameter value set was somewhat stronger. For longer time series ($T$=20) the interannual variability of simulated $Z_t$ values was the most important factor, with smaller values increasing parameter estimability. In terms of relative estimation errors of model outputs of interest, RE-Z models had smaller relative errors compared to FE-Z models. Hence there seems to be an advantage in using the random walk formulation for total mortality. However, the suitable model for total mortality, random walk or fixed effect, will depend on the particular application. A fisheries closure might preclude the use of the random walk approach unless an explanatory variable is introduced which models the step change in mean fishing mortality. If ignored a step change in a variable that is modelled by a random walk will lead to estimates exhibiting a time delay. This time delay was found by Mesnil et al. (2009) for a biomass model with a random effect for biomass growth when survey catchability was increased step wise. In terms of parameters, $log(\sigma_Z)$ impacted total mortality estimates in the final year $Z_T$ but not population depletion rate estimates. Relative bias of depletion rate estimates was generally higher than for $Z_T$ and both very negatively correlated. No other factor was found to be important, neither the length of the times series, nor the number of age classes or the sample size for numbers at age (here effective sample size); these factors have been found to impact bias of fishing mortality estimates in a simulation study using the Stock Synthesis model (Yin and Sampson 2004). The objective of the first simulation study was to study parameter estimability. Testing the robustness of the method in the face of model misspecification or biased observations was the objective of the second simulation study. It showed that decreasing recruitment or increasing selectivity will indeed lead to positively biased total mortality and negatively biased total abundance estimates. The degree of bias will of course depend on the misspecification scenario. The second simulation study only provides a first evaluation. Full exploration of the misspecification issue would require setting up simulation studies where the data are simulated with an operating and observation model which are not identical to the estimation model. This robustness testing and comparison is common practice in fisheries science (e.g. Mesnil et al. 2009) and a logical next step for evaluating the proposed method. However, as there are a wide range of possibilities, case specific simulation tests need to be set up.

Contrary to many routinely used assessment methods such as virtual population analysis (VPA) derived methods, total catches are not assumed to be known with certainty in MYCC. Further, no survey abundance index is required nor any other cpue tuning series which avoids the need for notoriously difficult to estimate effort time series and assumptions about the relationship between cpue and abundance which for many species has been found to be non-linear (Harley et al. 2001). Trends in fishing effort have been found to lead to biased stock abundance and mortality estimates (Dickey-Collas et al. 2010). Of course, if available survey data should always be used. The proposed MYCC fills the gap of methods applicable in situations with no survey data. This is commonly the case for deep-water species exploited on the continental shelf in European waters. As illustrated with the blue ling example missing data are easily handled.

The price to pay for limited data requirements are assumptions regarding the representativeness of the available proportions-at-age information, recruitment dynamics and total mortality being constant across the considered age classes. The appropriateness of these assumptions might depend on the particular stock. Selectivity and catchability varying strongly with age would clearly invalidate the first assumption as would large unknown discards of certain age classes. The approach adopted here for the blue ling case study was to only model the population from an age where individuals can be assumed fully recruited to the trawl fishery and hence proportions-at-age in the landings would be representative of the population.

In conclusion, we expect MYCC to fill the gap in the stock assessment toolbox for cases with no fisheries independent survey or fishing effort data but reliable information on proportions-at-age and total catches. The use of random effects make MYCC suitable for missing data situations. To take account of particular life history traits or particular fishing histories (closures, etc.) case specific model variants could be developed that make use of auxilliary information indicating total mortality or recruitment changes.

# Acknowledgments

The study was carried out with financial support from the Commission of the European Communities under the DEEPFISHMAN project (grant agreement no. 227390). The work was initiated when VT was visiting Tasmania with financial support from CSIRO.

# References


Beverton, J.H, and Holt, S.J. 1957. On the dynamics of exploited fish populations. Chapman and Hall, New York.

Bridger, J.P. 1978. New deep-water trawling grounds to the west of britain. Laboratory Leaflet, Ministry of Agriculture Fisheries and Food (MAFF), Directorate of Fisheries Research, Lowestoft, United Kingdom.

Candy, S.G. 2008. Estimation of effective sample size for catch-at-age and catch-at-length data using simulated data fromthe dirichlet-multinomial distribution. CCAMLR Science **15**: 115-138.

Chapman, D.G., and Robson, D.S. 1960. The analysis of a catch curve. Biometrics **16**: 354-541.

Conover, W. J. 1971. Practical nonparametric statistics. 2nd edition. John Wiley & Sons, New York.

Cotter, A.J.R., Mesnil, B., and Piet, G.J., 2007. Estimating stock parameters from trawl cpue-at-age series using year-class curves. ICES J. Mar. Sci. **64**: 234-247.

Deriso, R.B., Quinn II, T.J., and Neal, P.R. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. **42**: 815-824.

Dickey-Collas, M., Pastoors, M.A., and van Keeken, A. 2010. Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of north sea plaice. ICES J. Mar. Sci. **66**: 1641-1649.

Dorel, D. 1986. Poissons de l'Atlantique Nord-Est. Relations taille-poids. DRV-86-001/RH/NANTES. Ifremer, La Rochelle, France, 183 pp.

Doubleday, W.G. 1976. A least squares approach to analysing 552 catch at age data. ICNAF Res. Bull. **12**: 69-81.

Dunn, A., Francis, R.I.C.C., and Doonan, I.. 2002. Comparison of Chapman-Robson and regression estimators of Z from catch-curve data when non-sampling stochastic error is present. Fish. Res. 59: 149-159.

Ehrich, S. 1983. On the occurrence of some fish species at the slopes of the Rockall Trough. Archiv für Fischereiwissenschaft **33**: 105-150.

Ehrich, S., and Reinsch, H.H. 1985. Investigations on the blue ling stock (*Molva dypterygia dypt.*) in the waters west of the british isles. Archiv für Fischereiwissenschaft **36**: 97-113.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A. and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods. Softw. **27** :233-249.

Fryer, R. 2002. TSA: is it the way? Technical report, Appendix D in Report of Working 567 Group on Methods of Fish Stock Assessment, Dec. 2001. ICES CM 2002/D:01, 86-93 pp.

Gordon, J.D.M., and Hunter, J.E. 1994. Study of deep-water fish stocks to the west of Scotland. Scottish Association for Marine Science, Oban (Scotland), volume 1: 141-570.

Gudmundsson, G. 1986. Statistical considerations in the analysis of catch-at-age observations. J. Cons. int. Explora. Mer **43**: 83-90.

Gudmundsson, G. 1994. Time series analysis of catch-at-age observations. Appl. Statist. 43: 117-126.

Harley, S.J., Myers, R.A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. **58**: 1760-1772.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York.

Hulson, P.-J.F., Henselman, D.H., and Quinn II, T.J. 2011. Effects of process and observation errors on effective sample size of fishery and survey age and length composition using variance ratio and likelihood methods. ICES J. Mar. Sci. **68**: 1548-1557.

ICES 2011. Report of the working group on biology and assessment of deep-sea fisheries resources (WGDEEP). ICES CM 2011/ACOM:17.

Johnson, N.L., Kotz, S., and Balafishnan, N. 1997. Discrete multivariate distributions. John Wiley & Sons, New York, United States of America.

Lorance, P., Agnarsson, S., Damalas, D., des Clers, S., Figueiredo, I., Gil, J., and Trenkel, V.M. 2011. Using qualitative and quantitative stakeholder knowledge: examples from European deep-water fisheries. ICES J. Mar. Sci. **68**: 1815–1824.

Lorance, P., Pawlowski, L., and Trenkel, V.M. 2010. Deriving blue ling abundance indices from industry haul by haul data. ICES J. Mar. Sci. **67**: 1650-1658.

Magnussen, E. 2007. Interpopulation comparison of growth patterns of 14 fish species on Faroe bank: are all fishes on the bank fast-growing? J. Fish. Biol. **71**: 453-475.

Magnusson, A., and Hilborn, R. 2007. What makes fisheries data informative? Fish Fish. 8: 337-358.

Magnússon, J.V., Bergstad, O.A., Hareide, N.-R., Magnússon, J., and Reinert, J. 1997. Ling, blue ling and tusk of the northeast Atlantic (TemaNord 1997:535). Nordic Council of Ministers, Copenhagen, page 63 pp.

Maunder, M.N. 2011. Review and evaluation of likelihood functions for composition data in stock-assessment models: Estimating the effective sample size. Fish. Res. **109**: 311-319.

Mesnil, B., Cotter, J., Fryer, R.J., Needle, C.L., and Trenkel, V.M. 2009. A review of fishery-independent assessment models, and initial evaluation based on simulated data. Aquat. Liv. Res. **22**: 207-216.

Nielsen, A. 2009. State-space assessment models. Report of the ICES working group on methods of fish stock qssessment, October 2009. ICES CM 2009/RMC: 12.

Paloheimo, J.E. 1958. A method of estimating natural and fishing mortalities. J. Fish. Res. Board Can. **15**: 749-758.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 stocks. J. Cons. Int. Explor. Mer. **39**: 175-192.

Pennington, M., Burmeister, L.-M., and Hjellvik, V. 2001. Assessing the precision of frequency distributions estimated from trawl-survey samples. Fish. Bull. **100**: 74-80.

Schnute, J.T., and Haigh, R., 2007. Compositional analysis of catch curve data, with an application to *Sebastes maliger*. ICES J. Mar. Sci. **64:** 218-233.

Skaug, H.J., and Fournier, D.A. 2006. Automatic approximation of the marginal likelihood in non-gaussian hierarchical models. Comp. Stat. & Data Anal. **51**: 699-709.

Thorson, J.T., and Prager, M.H. 2011. Better catch curves: Incorporating age-specific natural mortality and logistic selectivity. Trans. Americ. Fish. Soc. **140**: 356-366.

Trenkel, V.M. 2008. A two-stage biomass random effects model for stock assessment without catches: What can be estimated using only biomass survey indices? Can. J. Fish. Aquat. Sci. **65**: 1024-1035.

Twedebrink, T. 2009. Dirmult: estimation in Dirichlet-Multinomial distribution, R Package Version 0.1.2. http://cran.r-project.org/web/packages/dirmult/dirmult.pdf.

Wayte, S.E., and Klaer, N.L. 2010. An effective harvest strategy using improved catch-curves. Fish. Res. **106**: 310-320.

Whitehead, P.J.P., Bauchot, M.L., Hureau, J.-C., Nielsen, J., and Tortonese, E. 1986. Fishes of the north-eastern Atlantic and the Mediterranean, Vol. 2. UNESCO, Paris.

Yin, Y., and Sampson, D.B. 2004. Bias and precision of estimates from an age-structured stock assessment program in relation to stock and data characteristics. N. Am. J. Fish. Manag. **24**: 865-879.

# Tables

Table 1. Model parameters, value used for simulation study and boundary values for estimation (min; max). Parameter set numbers with given values are provided in last column. CV coefficient of variation·

| Parameter | Description | Eq. | Value | Bounds | Parameter sets |
|---|---|---|---|---|---|
| T | number of years | 7-15 | 10 | | 1-4, 9-12, 17-20, 25-28 |
| | | | 2 | | 5-8, 13-16, 21-24, 29-32 |
| A | number of ages | 7-15 | 3 | | 1,2, 5, 6, 9, 10, 13, 14, 17, 18, 21, 22,25, 26, 29, 30 |
| | | | 10 | | 3, 4, 7, 8, 11, 12, 15, 16, 19, 20, 23, 24, 27,28, 31, 32 |
| $\mu_R$ | mean recruitment | 9 | 150 | [1,$10^9$] | 1-32 |
| $CV_R$ | CV recruitment | 9 | 0.4 | [0.01,3] | 1-16 |
| | | | 0.8 | | 17-32 |
| $log(\sigma_Z)$ | log(std. dev.) total mortality | 10 | -3 | [-4,50] | 1-8, 17-24 |
| | | | -1 | | 9-16, 25-32 |
| Z0 | historic total mortality | 8 | 0.4 | [0.01,2] | 1, 3, 5, 7, 9, 11, 13, 15,17, 19, 21, 23, 25, 27, 29, 31 |
| | | | 0.8 | | 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24,26, 28, 30, 32 |
| $CV_c$ | CV total catch | 15 | 0.02 | fixed | 1-32 |
| M | natural mortality | 15 | 0.2 | [0.001,2] | 1-32 |

Table 2. Percentage of simulation runs in which the parameter hit the boundary and overall percentage of simulations runs in which all parameters were estimable (not on bounds and convergence). c not estimated.

| Model | Parameter on boundary (%) | | | | | | Estimable runs (%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Simulation study 1 | | | | | | | | Simulation study 2 | | |
| | $\mu_R$ | $CV_R$ | $Z0$ | $log(\sigma_Z)$ | $Z_t$ | $M$ | $m_t$=50 | $m_t$=400 | Base | Rdec | Sel |
| RE-Z | 0 | 0 | 0 | 0 | - | c | 89 | 89 | 48 | 41 | 44 |
| RE-Z & M | 0 | 0 | 0.1 | 0 | - | 22.1 | 65 | 65 | 50 | 26 | 30 |
| FE-Z | 0 | 2.5 | 0.3 | - | 0.2 | c | 97 | 97 | - | - | - |
| FE-Z & M | 0 | 1.8 | 18.3 | - | 0 | 22.2 | 59 | 59 | - | - | - |

Table 3. Raw and aggregated sample size $\widetilde{y}_+$ for annual numbers-at-age data for blue ling. Raw sample size corresponds to sum of individuals in quarterly age-length keys.

| Year | Raw $y_+$ | Aggregated $\widetilde{y}_+$ | Ratio |
|---|---|---|---|
| 1988 | 295 | 155 | 0.52 |
| 1991 | 283 | 124 | 0.44 |
| 1992 | 1310 | 458 | 0.35 |
| 1993 | 918 | 352 | 0.38 |
| 1994 | 633 | 377 | 0.6 |
| 1995 | 643 | 226 | 0.35 |
| 2009 | 558 | 214 | 0.38 |
| 2010 | 615 | 130 | 0.21 |

Table 4. Estimated model parameters and their precision (SD standard deviation; CV coefficient of variation) for blue ling. $M = 0.17$.

| Parameter | Estimate | SD | CV |
|---|---|---|---|
| $\mu_R$ | 4085400 | 525100 | 0.129 |
| $CV_R$ | 0.451 | 0.073 | 0.162 |
| $log(\sigma_Z)$ | -2.984 | 0.221 | 0.074 |
| $F0 = Z0 - M$ | 0.094 | 0.019 | 0.202 |

# Figure legends

Figure 1. Simulation study 1: proportion of replicate runs in which all parameters were estimable by model parameter set (see Table 1). (a) RE-Z model, (b) RE-Z & M model, (c) FE-Z model, and (d) FE-Z & M model.

Figure 2. Simulation study 1: regression trees for estimable parameters as a function model parameter values (see Table 1) for four model variants. The dependent variable is 1 if all parameters were estimable and 0 otherwise. The inequalities at each branching level indicate the parameter values for each branch. For example, for the RE-Z model, the top inequality logsd_Z>= -2 means that in the branches on the left hand side the value for $log(\sigma_Z)$ is bigger than $-2$. CV_R = $CV_R$.

Figure 3. Simulation study 1: interquartile range of relative estimation bias across replicate runs (simulation study 1) for the estimates of (a) final year total mortality $Z_T$ and (b) depletion rate, $N_T/N_1$ for four model variants and 32 parameter sets (see Table 1).

Figure 4. Simulation study 2: boxplots of relative errors in total mortality $Z_t$ and total abundance $N_t$ estimates for model misspecification scenarios. Base : no model misspecification, Rdec: recruitment linearly decreasing with time, Sel: selectivity of observations increasing with age. Whiskers extend to extreme values, boxes stretch from 25 to 75 percentiles.

Figure 5. (a) Blue ling international landings, (b) mean individual weight (symbols) and proportion of individuals larger than 84 cm years (line) in French landings from the

west of the British Isles.

Figure 6. Quarterly proportions-at-age for blue ling obtained from combing length-frequency market samples and age-length keys. The vertical line indicates age 9 above which fishing selectivity is assumed to be constant.

Figure 7. Model diagnostics for blue ling RE-Z model. (a) qq-plot for $Z$ random effect, (b) qq-plot for $R$ random effect, (c) raw residuals for total landings, (d) observed vs. predicted total catches, (e) raw residuals for proportions-at-age (grey positive, white negative), (f), observed vs. predicted proportions-at-age. Fixed

natural mortality $M = 0.17$.

Figure 8. Blue ling estimates for RE-Z model. (a) total mortality, (b) total population abundance ($\geq$ age 9) and (c) recruits (age 9). Grey areas are 95% confidence bands. Fixed natural mortality $M = 0.17$.

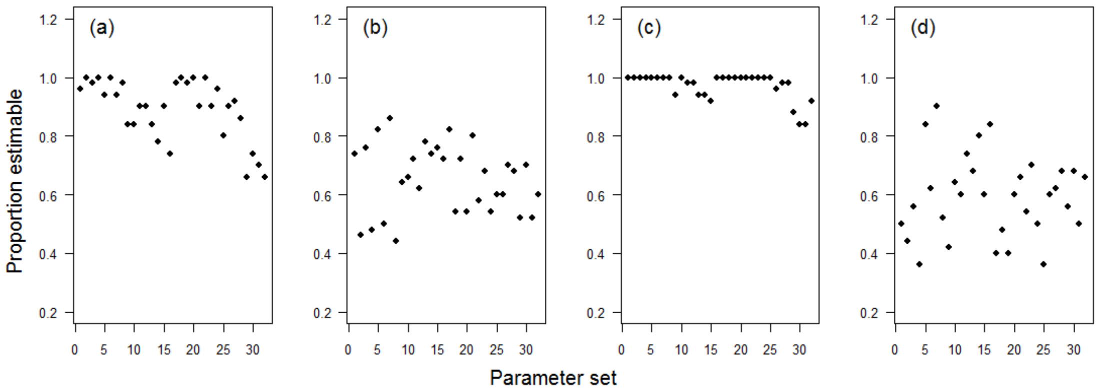

(a)
(b)
(c)
(d)
Proportion estimable
Parameter set
0.2
0.4
0.6
0.8
1.0
1.2
0
5
10
15
20
25
30

Figure 1

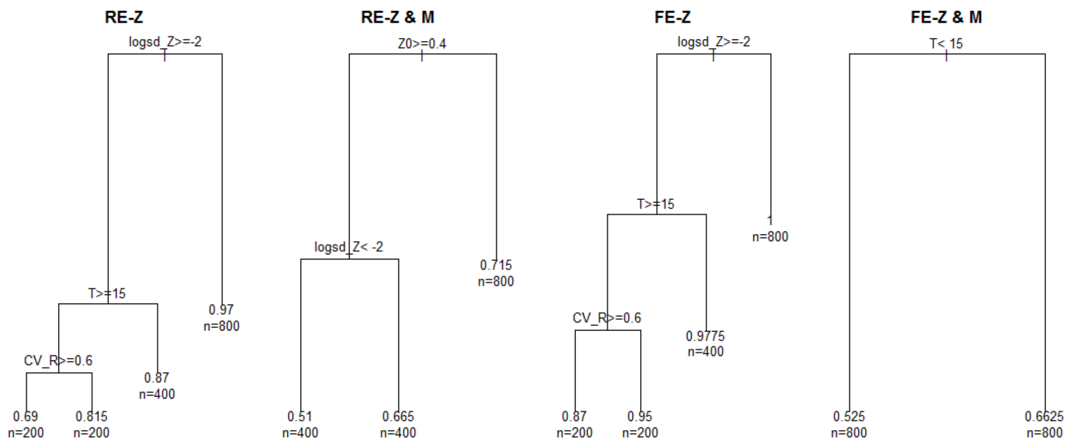

RE-Z
logsd_Z>=-2
T>=15
0.97
n=800
CV_R>=0.6
0.87
n=400
0.69
n=200
0.815
n=200
RE-Z & M
Z0>=0.4
logsd_Z< -2
0.715
n=800
0.51
n=400
0.665
n=400
FE-Z
logsd_Z>=-2
T>=15
1
n=800
CV_R>=0.6
0.9775
n=400
0.87
n=200
0.95
n=200
FE-Z & M
T< 15
0.525
n=800
0.6625
n=800

Figure 2

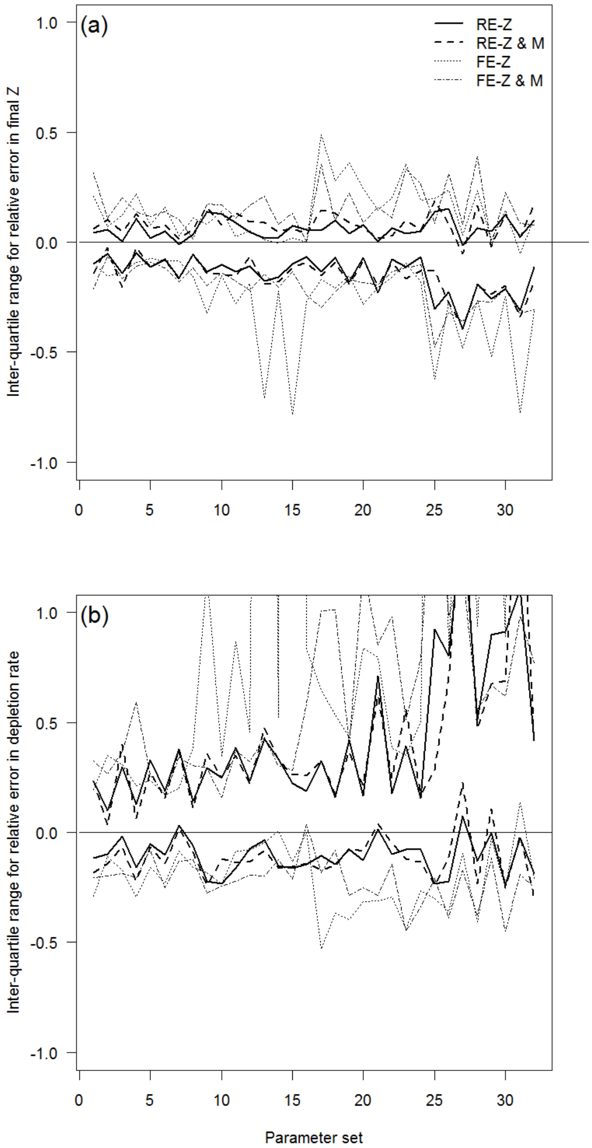

(a)
RE-Z
RE-Z & M
FE-Z
FE-Z & M
Inter-quartile range for relative error in final Z
(b)
Inter-quartile range for relative error in depletion rate
Parameter set

Figure 3

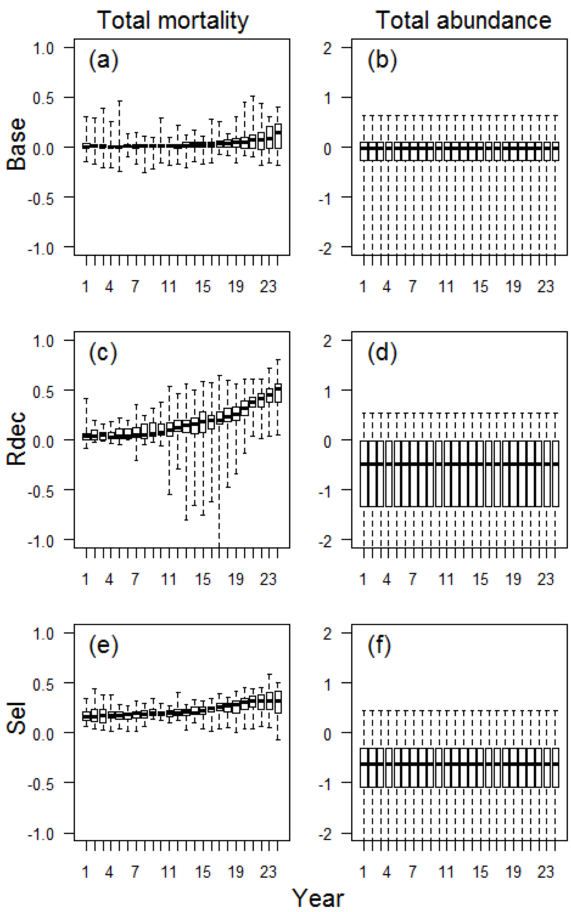

Total mortality
Total abundance
Base
Rdec
Sel
(a)
(b)
(c)
(d)
(e)
(f)
1.0
0.5
0.0
-0.5
-1.0
2
1
0
-1
-2
1 4 7 11 15 19 23
Year

Figure 4

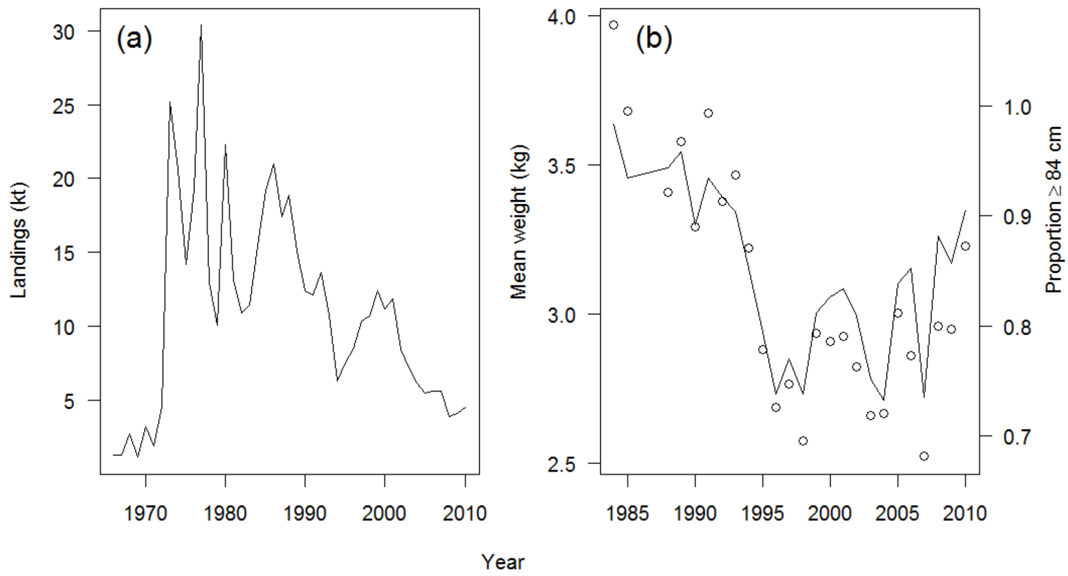

(a)
Landings (kt)
30
25
20
15
10
5
1970
1980
1990
2000
2010
(b)
Mean weight (kg)
4.0
3.5
3.0
2.5
Proportion ≥ 84 cm
1.0
0.9
0.8
0.7
1985
1990
1995
2000
2005
2010
Year

Figure 5

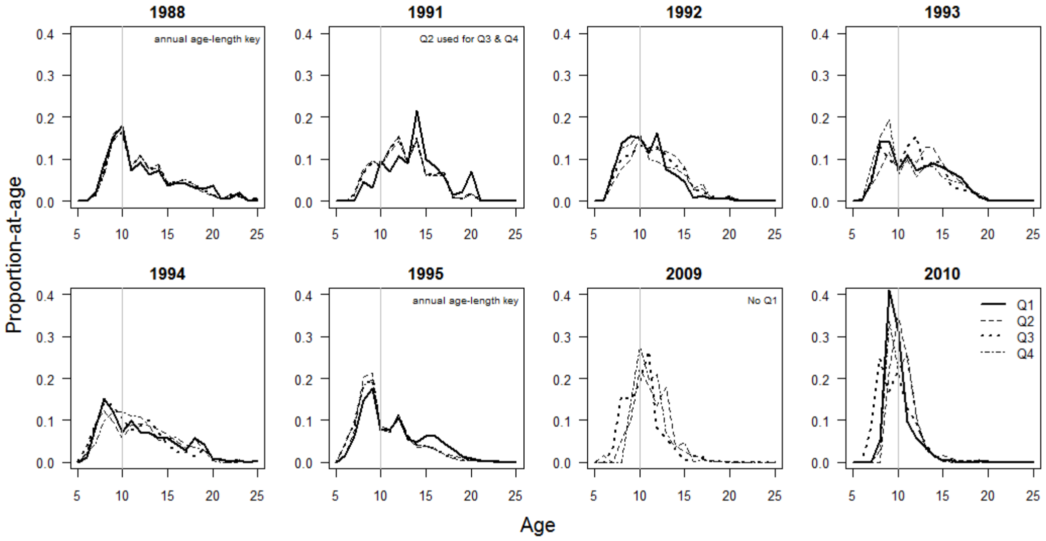
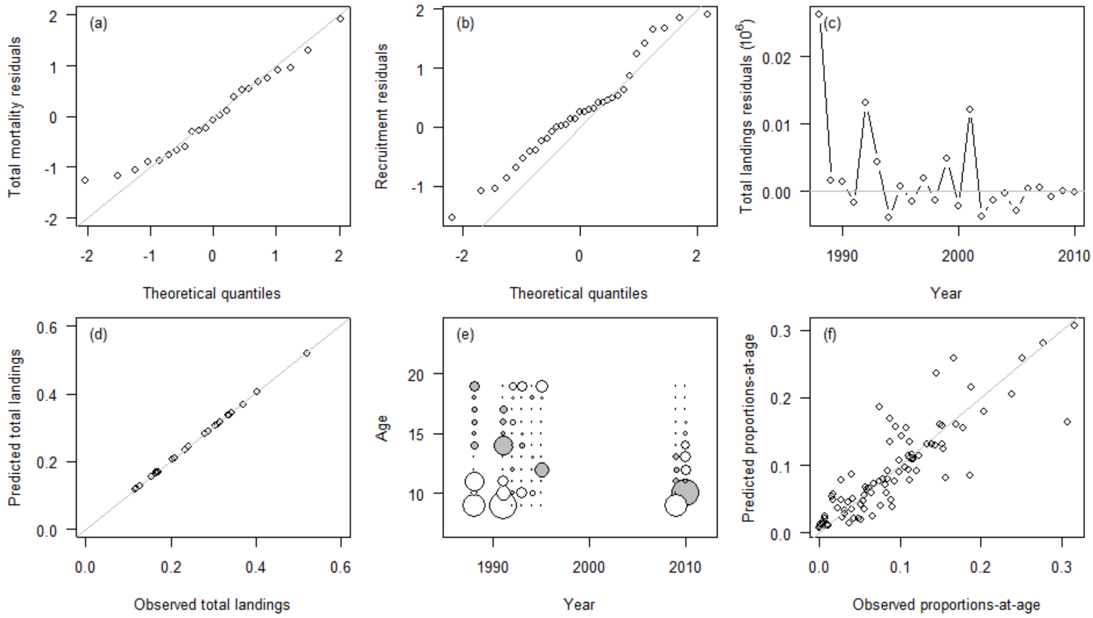
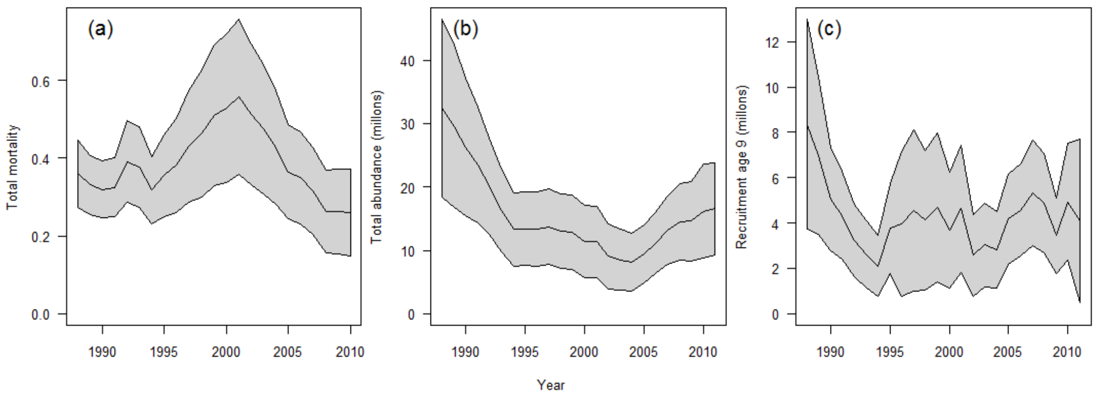

1988
annual age-length key
1991
Q2 used for Q3 & Q4
1992
1993
1994
1995
annual age-length key
2009
No Q1
2010
Q1
Q2
Q3
Q4
Proportion-at-age
Age

Figure 6




711

712 Figure 7



(a)
(b)
(c)
Total mortality
Total abundance (millons)
Recruitment age 9 (millons)
Year


Figure 8