# Importance of various data sources in deterministic stock assessment models

Submitted in partial fulfilment
of the requirements of the degree
of Masters of Science
of Rhodes University

**AMANDA ROSALIND NORTHROP**

11th March 2008

**Abstract**

In fisheries, advice for the management of fish populations is based upon management quantities that are estimated by stock assessment models. Fisheries stock assessment is a process in which data collected from a fish population are used to generate a model which enables the effects of fishing on a stock to be quantified. This study determined the effects of various data sources, assumptions, error scenarios and sample sizes on the accuracy with which the age-structured production model and the Schaefer model (assessment models) were able to estimate key management quantities for a fish resource similar to the Cape hakes (*Merluccius capensis* and *M. paradoxus*).

An age-structured production model was used as the operating model to simulate hypothetical fish resource population dynamics for which management quantities could be determined by the assessment models. Different stocks were simulated with various harvest rate histories. These harvest rates produced *Downhill trip* data, where harvest rates increase over time until the resource is close to collapse, and *Good contrast* data, where the harvest rate increases over time until the resource is at less than half of it's exploitable biomass, and then it decreases allowing the resource to rebuild. The accuracy of the assessment models were determined when data were drawn from the operating model with various combinations of error.

The age-structured production model was more accurate at estimating maximum sustainable yield, maximum sustainable yield level and the maximum sustainable yield ratio. The Schaefer model gave more accurate estimates of Depletion and Total Allowable Catch. While the assessment models were able to estimate management quantities using *Downhill trip* data, the estimates improved significantly when the models were tuned with *Good contrast* data. When autocorrelation in the spawner-recruit curve was not accounted for by the deterministic assessment model, inaccuracy in parameter estimates were high. The assessment model management quantities were not greatly affected by multinomial ageing error in the catch-at-age matrices at a sample size of 5000 otoliths. Assessment model estimates were closer to their true values when log-normal error were assumed in the catch-at-age matrix, even when the true

underlying error were multinomial. However, the multinomial had smaller coefficients of variation at all sample sizes, between 1000 and 10000, of otoliths aged.

It was recommended that the assessment model is chosen based on the management quantity of interest. When the underlying error is multinomial, the weighted log-normal likelihood function should be used in the catch-at-age matrix to obtain accurate parameter estimates. However, the multinomial likelihood should be used to minimise the coefficient of variation. Investigation into correcting for autocorrelation in the stock-recruitment relationship should be carried out, as it had a large effect on the accuracy of management quantities.

**Acknowledgements**

Thanks go to Tony Booth for stimulating my interest in fisheries and providing valuable advice and mentorship. Also to Sarah Radloff for her review, excellent suggestions and encouragement. Constructive comments on earlier versions of the manuscript by Rowan Yearsley are greatly appreciated. I am grateful to Nyasha Chigwamba and Oliver King for helping me along on my LaTeX adventure. Earlier versions of my Lyx style file were obtained from Shaun Bangay. Thanks go to my family: my parents, brother, grandmother and aunt and uncle, Peggy and Hans Veene, for their never-ending support and encouragement. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

# Contents

# Chapter 1

# Introduction

## 1.1 Quantitative advice and management

There is sufficient evidence available to show that world's fisheries are under threat (Committee on Fish Stock Assessment Methods, 1998; Phillips et al., 2003; Grafton et al., 2006). Declining catches and catch rates in most of the world's fish stocks indicate widespread over-fishing. Approximately a quarter of the world's fisheries are either over-exploited, depleted or recovering from depletion (17%, 7%, 1%, respectively) (Beddington et al., 2007). Specific examples include the collapse of Atlantic cod (*Gadus morhua)* in Canada (Myers et al., 1997), blue whiting (*Micromesistius poutassou*) in the Northeast Atlantic, and Chilean jack mackerel (*Trachurus murphyi*) in the Southeast Pacific (FAO, 2006).

Amelioration of these situations requires the development of workable management plans and their implementation (Mace et al., 2001; Rose and Cowan, 2003). Management plans require scientific input to provide managers (and politicians) with information on sustainable harvest and effort levels, together with risk assessments of the consequences of alternative harvesting strategies (Francis and Shutton, 1997). The principal input for management decision is the annual determination of the current status of the stock, and what harvest levels would be sustainable over a specified time-frame (Hilborn, 1992). These quantities are estimated using quantitative fisheries stock assessment models.

Fisheries stock assessment is the process of modelling an exploited resource to quantify the effects of harvesting (Mace et al., 2001). Outputs from these stock assessment models are then used as inputs into harvest rules. For example, if the findings show a resource to be below a predefined level, managers of the fishery can reduce annual catches to give the resource the opportunity to recover. Stock assessment models are now statistically and numerically complex (Hilborn, 1992), relying heavily on specific assumptions and require large quantities of input data that can be expensive to collect.

The choice of model to be used for stock assessment is determined by the type and quantity, including the length of the time series, of available data. The fewer the data the simpler the model, and conversely, the more data rich, the more complicated the model. Management-related outputs range from basic (such as annual harvest and effort levels) to detailed (such as length/age at harvesting, harvesting and effort level). The ideal situation is to construct a biologically plausible model that can provide the relevant information to managers and policy makers within a budget commensurate with the value of the resource.

Questions therefore arise as to what type of data to collect, how much data to collect and how much will the data cost? Ideally the data will be sufficient to enable the stock assessment model to estimate management related quantities at some specified level of uncertainty and within some financial constraint.

## 1.2 Stock assessment models

Stock assessment models are created to mimic the population dynamics of fish. These can be categorised by the underlying quantities being estimated and data available. For example, biomass dynamic models only model biomass and only require biomass data. Age-structured and length-structured models model the age/length composition of the resource, respectively. These models are data intensive and require at least estimates of abundance but ideally, include annual age/length composition data.

### 1.2.1 The mathematical and statistical components

There are two components to stock assessment - a mathematical component to simulate the population dynamics of the resource, and a statistical component that 'tunes' the model by comparing the population dynamics component with empirical data. The best model fit is obtained by minimising an appropriate loss function.

The simplest stock assessment model is the biomass dynamic model. Biomass dynamic models do not make any direct assumptions about the age-structure of the stock or the individual growth of fish (Quinn and Deriso, 1999). It models that portion of the resource that is available for harvesting, termed exploitable biomass, as

$$B_{y+1} = B_y + g(B_y) - C_y$$

where $B_{y+1}$ is the exploitable biomass in year $y+1$, $g(B_y)$ is a net growth function, and $C_y$ is the harvest in year $y$ (Haddon, 2001). Pella and Tomlinson (1969) generalised the form of the density-dependent growth function, and defined it as $\frac{dB}{dt}.\frac{1}{B}$, at some productivity parameter $p$, an intrinsic rate of increase $r$ and carrying capacity $EB_0$. The proposed form is expressed as

$$g(B_y) = \frac{dB}{dt}.\frac{1}{B} = \frac{r}{p}B_y\left(1 - \left(\frac{B_y}{EB_0}\right)^p\right) \quad .$$

Their generalisation included Schaefer's (1957) model, which assumes $p = 1$, as

$$g(B_y) = rB_y\left(1 - \frac{B_y}{EB_0}\right)$$

and Fox's (1970) parameterisation that assumes $p \rightarrow 0$ resulting in

$$g(B_y) = rB_y\left(1 - \frac{lnB_y}{lnEB_0}\right) \quad .$$

These forms of the biomass dynamic model are illustrated below in Figure 1.1.

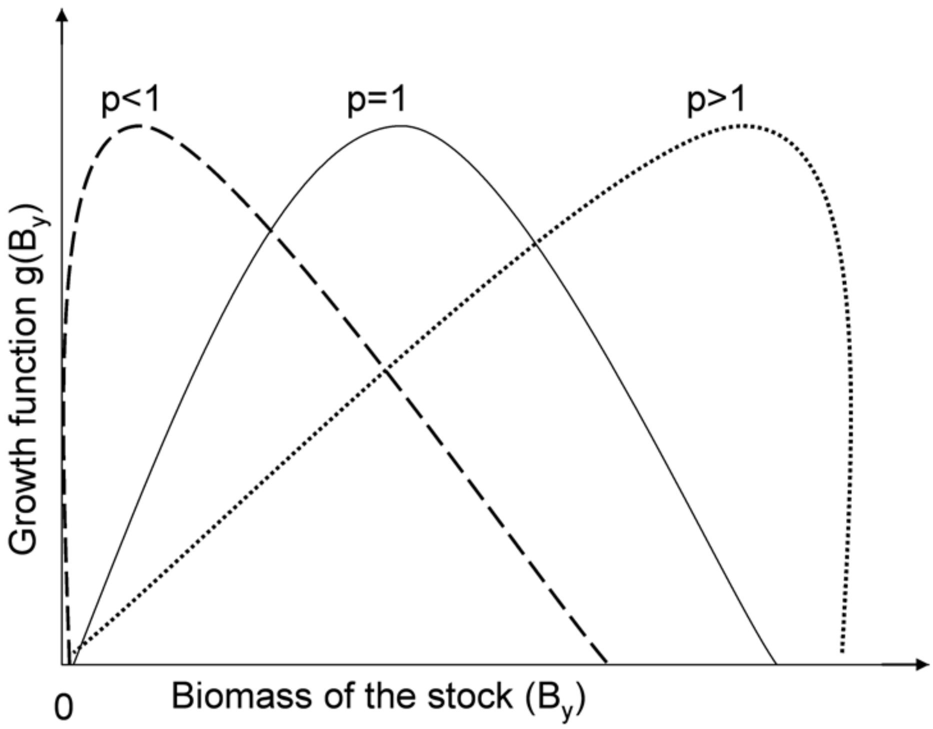


Figure 1.1: Different forms of the Pella and Tomlinson (1969) growth function based on the choice of $p$

The second component of any stock assessment model is statistical, relating the mathematical component to empirical data. This component is also referred to as the 'tuning' component. The population dynamics model is clearly not a perfect model of the system. For example, the stock's resource growth function may fluctuate due to natural variability (Francis and Shutton, 1997). This is known as process error. Process error could follow a log-normal distribution such that

$$g(B_y)e^{\eta_y} \qquad where\ \eta_y \backsim N\left(0, \sigma^2_{process}\right) \quad .$$

The data used to tune the model will also include measurement and sampling error, referred to as observation error (Francis and Shutton, 1997). Therefore, the predicted biomass in year $y$, $\hat{B}_y$, is related to some observed index of abundance $I_y$ as

$$I_y = q\hat{B}_y e^{\varepsilon_y}$$

where $q$ is that proportion of biomass caught by a unit of fishing effort (known as the catchability coefficient) and $\varepsilon_y \backsim N\left(0, \sigma^2_{observation}\right)$.

To include both forms of error, the full form of the Pella and Tomlinson (1969) biomass dynamic model, given a catch series $\boldsymbol{C}$ and an abundance index $\boldsymbol{I}$ is expressed as

$$I_{y+1} = q\left(B_y + \frac{r}{p}B_y\left[1 - \left(\frac{B_y}{EB_0}\right)^p\right]e^{\eta_y} - C_y\right)e^{\varepsilon_y} \quad .$$

The model parameters, $EB_0$, $r$, and $p$, are now estimated using at least one biomass index $\boldsymbol{I}$. Either $\sigma_{process}$ or $\sigma_{observation}$ are estimated, or the ratio between the two. Estimating both components simultaneously is intractable (Punt, 2003). In most instances, only $\sigma_{observation}$ is estimated as it assumes all error lies with the data and the population dynamics is assumed to be perfect (Haddon, 2001) .

The loss function used in the minimisation process has moved from a simple sum-of-squares function to one that is now totally likelihood-based (Quinn and Deriso, 1999). While the two are equivalent in a normal and log-normally distributed single abundance context, the likelihood approach allows for the inclusion of multiple indices if necessary, weighting diverse data in the appropriate manner (Schweder, 1998). Parameters are estimated non-linearly using one of numerous minimisation techniques. These include downhill simplex, Newton-Raphson, Levenberg-Marquardt and Polak-Ribiere approaches. Parameter variability is usually estimated by bootstrapping or from the Hessian matrix derived from the likelihood function (Rademeyer, 2003).

Therefore, by assuming observation error only ($\sigma = \sigma_{observation}$), the population dynamics parameters are estimated by maximising the log-normal likelihood of the form

$$L = \prod_y \frac{1}{I_y \sigma \sqrt{2\pi}} exp\left(\frac{(lnI_y - lnqB_y)^2}{2\sigma^2}\right) \quad .$$

Taking the natural logarithm, negating and dropping terms independent of the model parameters, the function to be minimised becomes

$$-LnL = nln(\sigma) + \frac{1}{2\sigma^2}\sum_y (ln(I_y) - ln(qB_y))^2$$

where $n$ is the number of years.

Maximum likelihood estimates of $\sigma^2$ and $q$ are obtained by taking the first derivative of the likelihood function with respect to the parameter of interest, setting the result to zero and solving for the parameter of interest, such that

$$\hat{q} = exp\left(\frac{1}{n}\sum_y \left[ln(\hat{I}_y) - ln(B_y)\right]\right)$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_y \left(ln(\hat{I}_y) - ln(I_y)\right)^2 \quad .$$

If $i$ indices are used and are assumed to be independent, they are additive on the log-scale such that

$$-LnL = -\sum_i LnL_i \quad .$$

### 1.2.2 Choosing a model

Stock assessment models vary in complexity. Biomass models are the simplest but are criticised because they lack biological realism to obtain mathematically simplistic solutions (Haddon, 2001). Maunder (2001) is of the opinion that if more information is available, then more complicated models should be used. By contrast, Francis and Shutton (1997) take note of the argument that the better a model is at emulating the fishery in question, the more useful one would assume it would be.

These kinds of arguments about the simplicity of models date back to the 13th century to *Occam's razor*. *Occam's razor* can be interpreted in a number of ways. It is often interpreted as if two models give the same result, then the simpler one is better (Fryer, 2003). Balasubramanian (1997) states that *Occam's razor* is the principle that simpler models should be preferred *until the data justify* more complex models. However, Xiao (1997) notes in his study of practical problems in production models: the more assumptions made within a model, the more likely it is that the model will fail. Thus one should move towards age-dependent production models only until the objectives are met. There is trade-off between simplicity and model error.

Model error occurs when the systems within the stock are not known entirely, and thus they are not properly quantified within the model (Francis and Shutton, 1997). As such, the more parameters estimated the higher their individual variability. Examples of this can be the assumptions in the model based upon the observation and process error. There have been a number of studies to evaluate how non-age-structured models predict stock dynamics, particularly if the stock displays distinct age-structure (e.g. Prager et al., 1996), especially if the resource is prone to large recruitment events.

### 1.2.3 Data used in stock assessment

Data availability is the main determinant of the type of model that can be applied to fisheries stock assessment. The more data available, the more complex the model. For example, assume two data sets, one with a single abundance index, and another with many abundance indices each accompanied with its disaggregated age data. The first data series only has an annual harvest rate, while the second series has the harvest estimate disaggregated into age-classes. If these are the only data available, then the first data set can be modelled, from a biomass perspective, only. By contrast, the second data set can be modelled by a fully age-structured model.

Data can be split into two main categories - fisheries dependent and fisheries independent data. Fisheries dependent data are collected from fishers that are targeting the fish for commercial purposes. Fisheries independent data are obtained by scientists. In addition, in age-structured models, data on basic biological parameters such as age-length relationships and maturity are used (Shepherd, 1984). These are usually estimated outside of the model and are assumed to have no error.

The two main sources of dependent fisheries data for most stock assessments are catch-per-unit effort (CPUE) and catch-at-age (CAA) obtained from commercial fishers (Auckland and Ninnes, 2004). The data utilised in most models are usually commercial catches over the years, CPUE, survey abundance indices from research trawls, and CAA information obtained from both commercial fishers and research trawls (Rose and Cowan, 2003).

CPUE is usually standardised using General Linear Modelling that accounts for variables that would affect effort. These include vessel type and size, depth fished, season and gear technology advancements

(Glazer and Butterworth, 2004). The result is to standardise each CPUE observation to *a* standardised vessel, fishing at *a* particular depth and season and using *a* particular gear.

Fisheries independent data are obtained from statistically designed research surveys. There are a number of benefits gained from a modelling perspective from independent data. According to Punt (1991), the advantages of survey trawl data include:

1. The effect of effort is negligible on the fish stock, thus it can be assumed that it has no effect on stock status, and thus on the core model,
2. The surveys are completed such that geographic and seasonal variables are assumed to be controlled,
3. The equipment and vessel used are the same every year. Thus the inter-annual absolute abundance change is due to inherent changes in the fish stock, rather than changes in effort. Technology creep should also not be a factor, and
4. The aim of the survey is solely data collection, thus care is taken by trained scientists to minimise observation error. Surveys usually employ a depth-stratified randomised sampling design.

Survey biomass and CPUE are conceptually similar. Both are assumed to be indices of abundance in relative terms. There are possible confounding variables such as herding effects created by the nets, or the avoidance of nets by fish (Ianelli, 2002). However, these effects most likely occur consistently through all years, thus the assumption of proportionality to the total biomass of the stock is reasonable (Rademeyer, 2003).

To date, studies pay little direct attention to quantify the individual contribution of each data set on a model (Magnusson, 2004). The approach commonly used is to apply *ad hoc* weighting based on consensus between scientists (and possibly industry). Shepherd (1984) rated the various data sets anecdotally, though the conclusions were not tested formally. Schnute and Hilborn (1993) show that some data sets may be completely uninformative and also contradictory. Weighting the data has been suggested by multiplying the data sets likelihood by the inverse of its variance (Haddon, 2001), an approach that has its foundation in likelihood theory. Thus data sets with a large amount of uncertainty are penalised and will have less effect on the overall model. Ludwig and Walters (1985) argued that having more detailed data does not necessarily improve the performance of some models. If a certain data set is the most influential but also has the most error, this may have a large negative effect on the model which also needs to be established.

Shepherd (1984) rated the different data sets with stars which represent utility. Catches were the most important and informative of data sets with three stars. Survey indices were also rated with three stars. Effort data of indicator fleets and length composition received a two star rating. Age composition had little utility with one star. There is a need to establish accuracy of models when different data sources are available for tuning.

## 1.3 Model evaluation and assessment

McCullagh and Nelder (1992, page 8) state that “all models are wrong; some, though, are more useful than others and we seek those.” The state of the real-world, or the ‘truth’, is mostly unknown. Simulations are used to create a hypothetical ‘truth’, such that assessment models can be tested against it. Models are therefore rigorously evaluated and their performance, bias and accuracy assessed using simulations.

Simulations imitate real-world processes over time (Banks et al., 2000) and are used widely in fisheries (Kell et al., 1999). The parameters and their variability for the simulated population models are known and can be sampled. These samples can be inputted in a stock assessment model such that the parameters estimated and the true parameters can be compared under different assessment model assumptions. Examples can be found in Walker (1992), Prager et al. (1996), Polet and Redant (1998) and Punt (2003).

The most commonly used framework used is that of an operating model - assessment model approach. The true underlying dynamics are included in the operating model (Kell et al., 2005). Variation and all parameters of the true population within the operating model are known. Data are then sampled from the operating model in the same form that it is sampled via fishing and fishery independent processes. The sampled data are then used as inputs into a range of different assessment models. The parameters obtained from the assessment model can be compared with the true parameters from the operating model to ascertain the relative error of the assessment models. It is suggested by Kell et al. (2005) that the operating model contain the best known knowledge about the stock. A diagram of the process is shown in Figure 1.2.

It is also possible to arrive at management strategies using the assessment model, and check how the strategies feed-back to, and affect, the true underlying population (Kell et al., 1999).

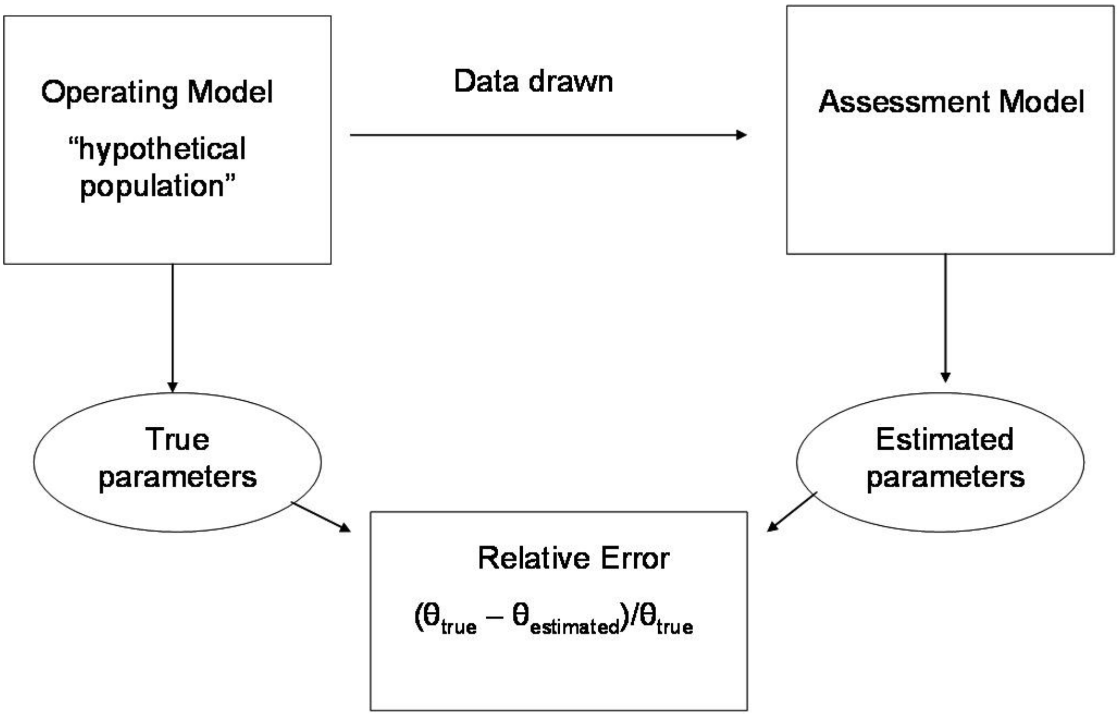


Figure 1.2: Operating model - assessment model simulation approach. A hypothetical population is simulated in the operating model. True parameter values from the operating model are known. Data are drawn from the operating model with error and used as input in the assessment model. Parameters are estimated by the assessment model and are compared with the true parameters from the operating model by a relative error measure.

## 1.4 Thesis structure

This thesis presents an operating model - assessment model simulation study to compare and contrast two commonly applied fisheries stock assessment models.

The aims of this thesis are:

1. To explore whether simple or complicated stock assessment models should be used to assess a hake-like resource.
2. To examine how observation error affects a hake-like stock assessment models. Notably, how does ageing error and various CAA likelihood functions influence the models.
3. To analyse how process error (correlated and uncorrelated) about the stock-recruitment relationship affects the overall performance of the models.
4. To investigate how informative different types of data are.
5. To explore the optimal number of otoliths that are needed to be aged in order to minimise uncertainty in parameter estimates.

Hypothetical fish populations, similar to Cape hake (a two species mixed resource comprising *Merluccius capensis* and *M. paradoxus*), were simulated with random or autocorrelated process error. Two different harvest rates were imitated. Data were drawn from the model with observation error. Various combinations of data sets were tested on two different assessment models. Further investigation into the likelihood function in the ageing error was carried out. Lastly, various sample sizes of CAA data were tested on the assessment model.

Chapter 2 details the methods used in three experiments. The model used to generate all simulations, and which is commonly used in stock assessment, is described. Exact model parameters used in it are then tabulated and presented. The three experiments are described separately. Experiment 1 assessed how accurate the two assessment models were with various data sources, harvest rates, and the presence or absence of observation error in the ageing procedure and process error in the stock-recruitment curve. Experiment 2 compared two different likelihood functions for the CAA matrix while holding data sources constant, though still varying harvest rates, observation error and process error. In Experiment 3 these two likelihoods were used to find the optimal amount of otoliths needed to be aged in order to minimise uncertainty in parameter estimates.

Chapter 3 presents the separate results for the three experiments. Chapter 4 provides a general discussion in which results are synthesised and compared to other studies. Limitations of the study are discussed.

# Chapter 2

# Materials and Methods

The simulation approach used in this study was to construct an operating model that mimics the 'true' population dynamics of the resource and to generate data that realistically simulates fishery dependent and fishery independent data collection processes. Two assessment models were constructed, which, using the same data from the operating model, could be compared and contrasted.

All procedures were carried out within a Monte Carlo framework. Ludwig and Walters (1985) have argued that this framework should always be used when testing a model or estimator, as accuracy cannot be ascertained by using 'real' data. The term simulation 'trial' refers to all simulations for a particular scenario for which many Monte-Carlo simulations are generated. A "single simulation" involves the generation of one particular population within a simulation trial.

There were 100 single simulations generated per trial. The data generated for each single simulation had one long and one short time series, with accompanying age disaggregated data. CPUE was a 23 year-long index, while the survey biomass index was six years long. Methods were similar to that of Magnusson (2004).

There were three experiments. The first experiment considered two harvest rate scenarios, four combinations of observation and process error and two models (a simple model and a more complicated age-structured model). Four combinations of data sources were used to tune the age-structured model. This equated to a total of 40 simulation trials. For the second experiment an additional eight simulation trials were run (two harvest rates x four error scenarios) with a multinomial likelihood such that the log-normal likelihood in the first experiment could be compared with the multinomial. Lastly, 22 simulation trials were run in the third experiment to ascertain an acceptable sample size of otoliths to age to minimise uncertainty of the parameter estimates (two likelihood functions x 11 sample sizes).

## 2.1 Operating Model

The operating model is an age-structured production model (ASPM). It is fully age-structured, is initiated in the year preceding harvesting and includes a stock-recruitment relationship, which is the relationship between adult biomass and zero-year old fish.

### 2.1.1 Resource dynamics

Population abundance, in terms of numbers, follows the recursive equations (Booth and Quinn, 2006)

$$N_{y,a} = \begin{cases} R_y & if\ a = 0 \\ \left(N_{y-1,a-1}e^{-M/2} - C_{y-1,a-1}\right) e^{-M/2} & if\ 0 < a < max \\ \left(N_{y-1,max-1}e^{-M/2} - C_{y-1,max-1}\right) e^{-M/2} & \\ \ + \left(N_{y-1,max}e^{-M/2} - C_{y-1,max}\right) e^{-M/2} & if\ a = max \end{cases}$$

where $N_{y,a}$ is the number of fish at age $a$ at the start of year $y$, $R_y$ is recruitment at the start of year $y$, $M$ is the natural mortality rate, $C_{y,a}$ is the numbers of fish caught at year $a$ in year $y$. After a certain age, it is difficult to determine the age of a fish. Thus, at this threshold level all fish are aggregated into one lumped 'plus-group' (Quinn and Deriso, 1999). This age is the maximum age denoted $max$.

#### Annual recruitment

The number of recruits at the start of year $y$ is considered to be related to the previous year's spawner biomass through the Beverton and Holt (1957) stock-recruitment relationship

$$R_{y+1} = \frac{SB_y}{\alpha + \beta SB_y} e^{\varepsilon_y - \sigma_R^2/2} \tag{2.1}$$

where $SB_y$ is a measure of the spawner biomass stock size and $\alpha$ and $\beta$ determine the shape of the relationship. Random recruitment process error is included into the stock-recruitment relationship with the error term $e^{\varepsilon_y - \sigma_R^2/2}$ which is log-normally distributed such that $\varepsilon_y \sim N(0, \sigma_R^2)$ (Clark, 1993). Recruitment has a right-skewed distribution, thus the assumption of log-normal errors is reasonable (Bradford, 1991). The residuals are adjusted by $exp(\sigma_R^2/2)$, which is the expected value of a log-normally distributed random variate.

Spawner biomass, the biomass of sexually mature fish, is calculated as

$$SB_y = \sum_{a=0}^{max} N_{y,a} W_a \psi_a \quad .$$

The stock recruitment relationship is re-parametrised to contain a single steepness parameter "$h$". This parameter describes the fraction of pristine recruits when spawning biomass is reduced to 20% of carrying capacity (Francis, 1992) such that

$$\alpha = \frac{SB_0(1-h)}{4hR_0}$$

and

$$\beta = \frac{5h-1}{4hR_0} \quad .$$

### 2.1.2 Catches

Annual catches are assumed to occur in the middle of the year and known without error, such that

$$C_y = \mu_y \sum_a W_{a+1/2} N_{y,a} S_a^{com} e^{-M/2} \tag{2.2}$$

where $W_{a+1/2}$ is the mass of a fish at age $a$ in the middle of the year, $\mu_y$ is the harvest rate for year $y$ (effectively the fraction of the resource harvested during the year) and $S_a^{com}$ is selectivity of the commercial fish gear on a fish of age $a$.

### 2.1.3 Initial conditions

The recruitment into the fishery in the first year of exploitation is calculated as

$$R_0 = \frac{K_{sp}}{\sum_{a=0} \tilde{N}_{y,a} W_a \psi_a}$$

where $K_{sp}$ is the carrying capacity of the spawner biomass and $\psi_a$ is the age the fish reaches maturity. Relative numbers at age $a$ are

$$\tilde{N}_a = \begin{cases} 1 & if\ a = 0 \\ \tilde{N}_{a-1} e^{-M} & if\ 1 \leq a < max \\ \frac{\tilde{N}_{a-1} e^{-M}}{1-e^{-M}} & if\ a = max \end{cases} .$$

From this, age-disaggregated numbers of fish at each age for the first year of exploitation $N_{0,a}$ are calculated as

$$N_{0,a} = \begin{cases} R_0 & if\ a = 0 \\ N_{0,a-1} e^{-M} & if\ 1 \leq a < max \\ \frac{N_{o,a-1} e^{-M}}{1-e^{-M}} & if\ a = max \end{cases} .$$

**Gear Selectivity**

Selectivity is a measure of the vulnerability of a fish of a particular size (or age) to a fishing gear. A value of zero implies that the fish is invulnerable to fishing gear, whereas a value of one implies it is completely vulnerable. Selectivity for commercial trawlers $S_a^{com}$, can be calculated in terms of the logistic ogive

$$S_a^{com} = \left(1 + e^{-(a-a_c)/\delta}\right)^{-1}$$

where $a_c$ is age at 50% selectivity and $\delta$ is the inverse rate of selection.

Research surveys are assumed to catch all fish from age-two years and older. Smaller proportions of zero-year ($age0sel$) and one-year olds ($age1sel$) are assumed to be selected, and it is assumed that $age0sel \leq age1sel \leq 1$. Research selectivity at age $a$ is denoted as $S_a^{res}$.

**Maturity**

Maturity $\psi_a$ at a certain age $a$ is one if the fish are sexually mature at age $a_{mat}$, and zero otherwise such that

$$\psi_a = \begin{cases} 0 & if\ a < a_{mat} \\ 1 & if\ a \geq a_{mat} \end{cases} .$$

**Mass-at-age**

The begin-year mass of a fish of age $a$, $W_a$ is calculated using the von Bertalanffy equation coupled with a mass-at-length relationship such that

$$W_a = \phi\left(l_\infty\left(1 - e^{-\kappa(a-t_0)}\right)\right)^{\varphi}$$

where $\phi$ and $\varphi$ govern the relationship between length and weight, $l_\infty$ is the maximum length the fish can theoretically attain, $\kappa$ is the Brody growth coefficient and $t_0$ is the theoretical age the fish would have been at length zero.

Mass-at-age half way through a year $W_{a+1/2}$ is calculated with the inclusion of 0.5 in the exponential term such that

$$W_{a+1/2} = \phi\left(l_\infty\left(1 - e^{-\kappa(a-t_0+0.5)}\right)\right)^{\varphi} \quad .$$

**Natural mortality**

Mortality rate is assumed to be age invariant.

### 2.1.4 Calculation of Maximum Sustainable Yield (MSY)

Yield, as a function of harvest rate $\mu$, $Yield(\mu)$, is calculated as

$$Yield(\mu) = R(\mu)\mu \sum_a \tilde{N}_a S_a^{com} e^{-M/2}$$

where

$$\tilde{N}_a = \begin{cases} 1 & if\ a = 0 \\ \tilde{N}_{a-1}e^{-M/2}(1-\mu S_a^{com}) & if\ 1 \leq a < max \\ \frac{\tilde{N}_{max-1}e^{-M/2}(1-\mu S_{max-1}^{com})}{\left[1-e^{-M/2}(1-\mu S_{max}^{com})\right]} & if\ a = max \end{cases}$$

and the number of recruits as a function of $\mu$ and the stock-recruitment parameters $\alpha$ and $\beta$ is

$$R(\mu) = \frac{SBR(\mu) - \alpha}{SBR(\mu)\beta}$$

and $SBR(\mu)$ is spawner-biomass per recruit defined as

$$SBR(\mu) = \sum_a \tilde{N}_a \psi_a W_a \quad .$$

Maximum sustainable yield (MSY) was calculated by maximising the yield vs harvest rate curve.

The first and second derivatives were calculated numerically to estimate $Yield(\mu)$ as Booth and Quinn (2006)

$$f' = \frac{Yield(\mu + \tau) - Yield(\tau)}{\tau}$$

and

$$f'' = \frac{Yield(\mu + \tau) - 2Yield(\mu) + Yield(\mu - \tau)}{\tau^2}$$

where $\tau = 0.001$. The maximum of the yield curve was estimated by updating $\mu$ iteratively as

$$\mu_i = \mu_{i-1} - (f'/f'')$$

where $i$ is the ith iteration, until a tolerance of $|(\mu_i - \mu_{i+1})/\mu_{i+1}| > 0.0001$ was achieved.

### 2.1.5 Management quantities

Management quantities are derived from stock assessment models, and used by decision-makers to set annual harvest levels (Hilborn, 1992, 2003). Five management quantities were calculated in the operating model. These quantities give the decision-maker an idea of the current stock status, the maximum yield, and what the stock status should be in order to obtain maximum yield. They were

$MSY$ - maximum sustainable yield,

$MSYL$ - the level of exploitable biomass at which $MSY$ is attained,

$MSYR$ - the ratio of $MSY$ to $MSYL$ and corresponds to the harvest rate to achieve $MSY$

$DEPLETION$ - the ratio of the exploitable biomass in the last year modelled to the carrying capacity of the exploitable biomass $EB_{24}/EB_0$. This provides an estimate of the reduction in stock size that has occurred since the start of fishing, and

$TAC$ - the 'total allowable catch' as $TAC = 0.9 \times MSYR \times EB_{24}$.

### 2.1.6 Generation of abundance indices and catch-at-age matrices

Commercial CPUE and CAA were generated under the assumption that commercial catches are made half way through the year. The survey biomass indices and CAA were calculated assuming the surveys occur at the beginning of the year. It was assumed that CPUE and catch data were available since the inception of the fishery (23 years), and that surveys were conducted in the last six years.

An approach similar to Punt (1992), Punt (2003) and Wang et al. (2005) was used. Random log-normal observation error with a mean of zero and standard deviation of 0.2 was added to the CPUE data and survey abundance data at each simulation ($\sigma_{CPUE} = \sigma_{Biomass} = 0.2$), and process error was included in the spawner-recruitment relationship. The final relative biomass values were not altered to a certain value. This is a similar approach to Maunder (2001). Thus all differences in values were due to variation in the indices and the spawner-recruit relationship.

### 2.1.7 Choice of values for the operating model

The operating model simulated a hake-like resource. As such, growth and maturity parameters for the ASPM were taken from a study by Punt and Leslie (1991) on South African hake, and selectivity parameters were estimated in a ASPM stock assessment model using data and methods for hake of the West Coast from Rademeyer (2003). All input values are included in Table 2.3.

Two different harvest rate scenarios were simulated. These were denoted as *Downhill trip* and *Good contrast*. *Downhill trip* data is a hypothetical scenario where the harvest rate of a resource increases over time until the resource is close to collapse. It was calculated such that the harvest rate increases over time, while the CPUE decreases (Hilborn, 1979; Polacheck et al., 1993; Magnusson, 2004). *Good contrast* data is when the harvest rate increases over time until it brings the resource to less than half of it's exploitable biomass, and then it decreases allowing the resource to rebuild (Magnusson, 2004). Harvest rates are summarised in Table 2.1 and illustrated in Figure 2.1. Thus, in general catch trajectories were created such that depletion levels ended at roughly 20% for the *Downhill trip* data, and 40% for the *Good contrast* data. Ludwig and Walters (1985) note that there must be substantial contrast in CPUE data in order to obtain accurate parameter estimates. Hilborn (1979) noted that data collection that is initiated after a stock had been largely depleted resulted in problematic stock assessments.

The carrying capacity of the spawner-stock biomass $K_{sp}$ was assumed to be a constant value of two million tons. This was in order to deplete the stock to 20% for the *Downhill trip* data, and 40% for the *Good contrast* data. The steepness parameter $h$ was chosen as 0.66, such that some level of density dependence existed. Natural mortality $M$ was held at a constant rate of 0.2 per year which is a plausible assumption for hake (Punt, 1992). An arbitrary value for $q_{com}$ was chosen as 0.0001 for the commercial survey. A value of 0.6 was chosen for the research $q_{res}$, as it was assumed that the surveys catch 60% of fish. A total of 5000 fish was assumed to be aged (2500 for commercial catches and 2500 for survey catches).

Any fish over the age of six was treated as 'plus-group' where they were lumped into a '7+' category.

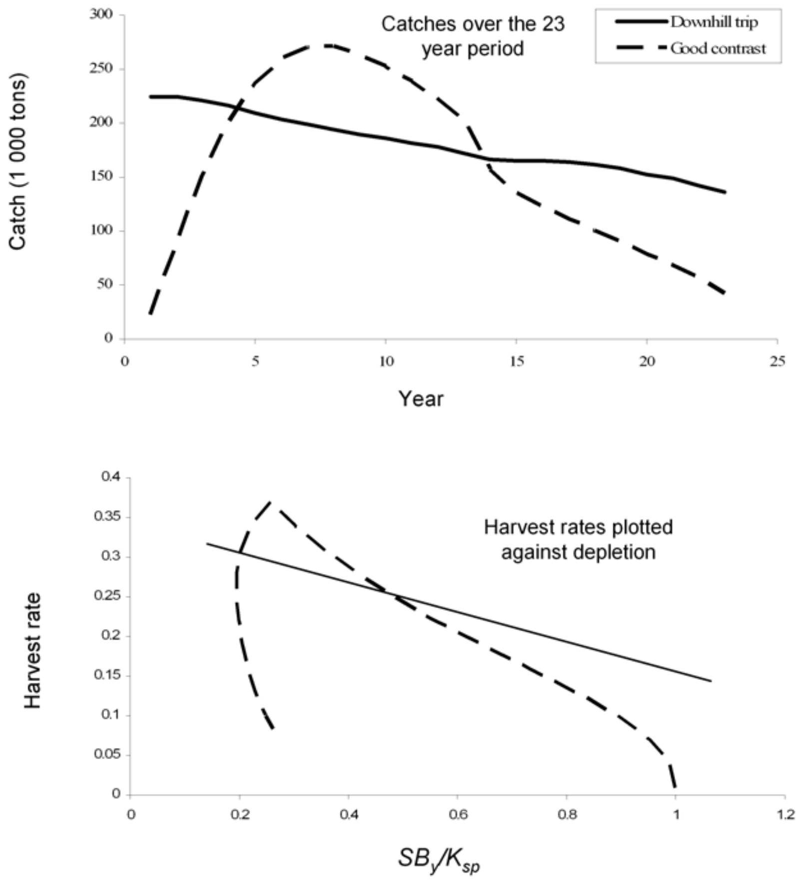


Figure 2.1: Examples of *Downhill trip* and *Good contrast* data for catches over the 23 year period. Harvest rates were also plotted against the depletion level of spawner-stock biomass $SB_y/K_{sp}$.

Table 2.1: Harvest rates for the *Downhill trip* data and *Good contrast* data scenarios

| Year | *Downhill trip* data | *Good contrast* data |
|---|---|---|
| 1 | 0.1 | 0.01 |
| 2 | 0.11 | 0.04 |
| 3 | 0.12 | 0.07 |
| 4 | 0.13 | 0.1 |
| 5 | 0.14 | 0.13 |
| 6 | 0.15 | 0.16 |
| 7 | 0.16 | 0.19 |
| 8 | 0.17 | 0.22 |
| 9 | 0.18 | 0.25 |
| 10 | 0.19 | 0.28 |
| 11 | 0.2 | 0.31 |
| 12 | 0.21 | 0.34 |
| 13 | 0.22 | 0.37 |
| 14 | 0.23 | 0.34 |
| 15 | 0.24 | 0.31 |
| 16 | 0.25 | 0.28 |
| 17 | 0.26 | 0.25 |
| 18 | 0.27 | 0.22 |
| 19 | 0.28 | 0.19 |
| 20 | 0.29 | 0.16 |
| 21 | 0.3 | 0.13 |
| 22 | 0.31 | 0.1 |
| 23 | 0.32 | 0.07 |

### 2.1.8 Error scenarios

Four error scenarios were considered. Scenario 1 had random recruitment process error around the spawner-recruit curve and no ageing error. This was denoted as the "*Baseline*" scenario. Scenario 2 had autocorrelation in the recruit residuals and no ageing error. Scenario 3 had random recruitment and ageing error within the CAA matrices, and Scenario 4 had both autocorrelation in recruitment, and ageing error. Table 2.2 shows the various error scenarios with their corresponding code-names used in figures and tables.

Recruitment variation is largely unknown in South African hake stock assessment (Butterworth and Punt, 1999). A standard deviation of 0.3 (i.e $\sigma_R = 0.3$) was used to generate the error in Equation 2.1.

There are two main factors affecting the amount of recruits. Firstly, the size of the spawning stock, and secondly the environment. For the latter, there appears to be periods where recruits into the fishery are consistently healthy, and equally, periods where the recruits into the fishery are consistently poor (Deriso et al., 1985; Ianelli, 2002). This shows some level of autocorrelation in the spawner-recruitment relationship (Clark, 1993).

For Scenarios 2 and 4 which had autocorrelation, $\varepsilon_y$ in Equation 2.1 was replaced with series $\eta_y$ from the recursive equation

$$\eta_y = \rho\eta_{y-1} + \varepsilon_y\sqrt{1-\rho^2}$$

where $\varepsilon_y \sim N(0, \sigma_R^2)$ distribution. Note if $\rho = 0$ then $\varepsilon_y = \eta_y$. The equation ensures that the overall variance is preserved (Clark, 1993). A value of 0.9 was chosen for $\rho$ based upon Clark's (1993) study.

Multinomial ageing error was included in the CAA matrices in a similar manner to Ianelli (2002) in Scenario 3 and Scenario 4.

Table 2.2: Error scenarios simulated in the operating model. Combinations of process and observation error were simulated.

| Scenario | Name | Random recruitment | Correlated recruitment | No ageing error | Ageing error |
|---|---|---|---|---|---|
| 1 | *Baseline* | x | | x | |
| 2 | *Corr NoAE* | | x | x | |
| 3 | *NoCorr AE* | x | | | x |
| 4 | *Corr AE* | | x | | x |

Table 2.3: Values for input parameters for the operating model. Where possible, values were taken from literature.

| Parameter | Value | Reference |
|---|---|---|
| $K_{sp}$ | 2 000 000 | |
| $h$ | 0.66 | |
| $a_c$ | 2.59 | Rademeyer (2003) |
| $\delta$ | 0.28 | Rademeyer (2003) |
| $age0sel$ | 0.004 | Rademeyer (2003) |
| $age1sel$ | 0.24 | Rademeyer (2003) |
| $M$ | 0.2 | Punt (1992) |
| $a_{mat}$ | 4 | Punt and Leslie (1991) |
| $\phi$ | 0.0055 | Punt and Leslie (1991) |
| $\varphi$ | 3.084 | Punt and Leslie (1991) |
| $\kappa$ | 0.046 | Punt and Leslie (1991) |
| $l_\infty$ | 230.3 | Punt and Leslie (1991) |
| $t_0$ | -0.825 | Punt and Leslie (1991) |
| $\sigma_R$ | 0.3 | |
| $\sigma_{CPUE}$ | 0.2 | |
| $\sigma_{Biomass}$ | 0.2 | |
| $\rho$ | 0.9 | Clark (1993) |
| $q_{com}$ | 0.0001 | |
| $q_{res}$ | 0.6 | |

## 2.2 Assessment Models

Two assessment models were used, that of the Schaefer version of the biomass dynamic model, and the ASPM.

### 2.2.1 Biomass dynamic model

**Population dynamics**

The Schaefer form of the biomass dynamic model

$$g(B_y) = rEB_y\left(1 - \frac{EB_y}{EB_0}\right)$$

was used, where $r$ is the intrinsic rate of growth, $EB_y$ is the exploitable biomass in year $y$ and $EB_0$ is the carrying capacity of the exploitable biomass.

MSY is calculated as

$$MSY = \frac{rEB_0}{4} \quad ,$$

and MSYL as

$$MSYL = \frac{EB_0}{2} \quad .$$

**The likelihood function**

CPUE and survey biomass indices were used to tune the model. The two loss functions are additive, thus the final negative log-likelihood function $-LnL$ is

$$-LnL = -LnL_{CPUE} - LnL_{Biomass} \quad .$$

The negative log-likelihood for the CPUE $(-LnL_{CPUE})$ assuming log-normally distributed error where $\varepsilon_{y,a} \backsim N\left(0, \sigma^2_{CPUE}\right)$ is

$$-LnL_{CPUE} = nln(\sigma_{CPUE}) + \frac{1}{2\sigma^2_{CPUE}} \sum_y (ln(CPUE_y) - ln(q_{com}EB_y))^2 \tag{2.3}$$

where $n$ is the number of years in the series, with the variance estimated by

$$\hat{\sigma}^2_{CPUE} = \frac{1}{n} \sum_y \left( ln(C\hat{PU}E_y) - ln(CPUE_y) \right)^2$$

and $q_{com}$ estimated by

$$\hat{q}_{com} = exp\left( \frac{1}{n} \sum_y \left[ ln(C\hat{PU}E_y) - ln(\hat{EB}_y) \right] \right) \quad .$$

The log-likelihood for the survey biomass $-LnL_{Biomass}$ is

$$-LnL_{Biomass} = nln(\sigma_{Biomass}) + \frac{1}{2\sigma^2_{Biomass}} \sum_y (ln(CPUE_y) - ln(q_{res}EB_y))^2 \tag{2.4}$$

which is calculated in an analogous manner to the CPUE data.

#### Parameters

The estimated parameters were calculated using AD model builder (Otter Research Ltd, 1994). These were the intrinsic rate of growth $r$ and the carrying capacity of the exploitable biomass $EB_0$. Five management quantities were also calculated - MSY, MSYL, MSYR, DEPLETION and TAC.

### 2.2.2 Age-structured production model

#### Population dynamics

The age-structured production model was the same as the operating model. The model is deterministic and has no process error in the stock-recruitment relationship.

#### The likelihood function

A maximum of four data sources were used to tune the ASPM model. Using all data sources, the negative log-likelihood function is

$$-LnL = -LnL_{CPUE} - LnL_{Biomass} - LnL_{CAA}^{com} - LnL_{CAA}^{res} \quad .$$

Both $-LnL_{CPUE}$ and $-LnL_{Biomass}$ are the same as the biomass dynamic model (Equations 2.3 and 2.4).

In the modelling process it is often assumed that CAA matrices are not affected by ageing error (Patterson et al., 2001). If error is assumed, it is frequently presumed to be either log-normally or multinomially distributed. If most of the error in the CAA matrix is from ageing inaccuracy multinomial error is used (Fournier and Archibald, 1982; Deriso et al., 1985). If most of the error is assumed to be from changes in effort and selectivity then the log-normal distribution is used (Rademeyer, 2003).

CAA data usually have large sample sizes and thus have a great influence on an age-structured production model (Maunder and Langley, 2004). Thus the likelihood section of the CAA data can be weighted down by its variation, or weighting is done *ad hoc*.

The negative log-likelihood for the commercial CAA $-LnL_{CAA}^{com}$ assuming log-normally distributed error where $\epsilon_{y,a} \backsim N\left(0, \frac{\sigma^2}{\hat{p}_{y,a}^{com}}\right)$ from Rademeyer (2003) is

$$-LnL_{CAA} = \sum_{y}\sum_{a}\left[ln\left(\sigma_{com}/\sqrt{\hat{p}_{y,a}^{com}}\right) + \hat{p}_{y,a}^{com}\left(lnp_{y,a} - ln\hat{p}_{y,a}^{com}\right)^2 / \left(2\sigma_{com}^2\right)\right] \tag{2.5}$$

with standard deviation estimated by $\hat{\sigma}^{com} = \sqrt{\sum_y \sum_a \hat{p}_{y,a}^{com}\left(ln\hat{p}_{y,a}^{com} - lnp_{y,a}^{com}\right)^2 / \sum_y \sum_a 1}$

where

$\hat{p}_{y,a}^{com} = \hat{C}_{y,a} / \sum_a \hat{C}_{y,a}$

and $\hat{C}_{y,a}^{com} = N_{y,a}e^{-M/2}S_{y,a}^{com}\mu_y$ for the commercial CAA where $\mu_y$ is obtained from Equation 2.2.

Research CAA for surveys are estimated in an analogous manner as for commercial CAA, except the estimated proportions are calculated as:

$$\hat{p}_{y,a}^{res} = \frac{S_a^{res}N_{y,a}}{\sum_{a'=0} S_a N_{y,a'}} \quad .$$

Alternatively, it can be assumed that CAA for both commercial or survey trawls is multinomially distributed, such that the log-likelihood is expressed as

$$-LnL^{i}_{CAA} = \sum_{y}\sum_{a} N_{y,a} ln(\hat{p}^{i}_{y,a}) \tag{2.6}$$

where $i$ is $com$ for the commercial trawls or $res$ for the survey trawls.

#### Parameters

Mortality $M$, maturity $\psi$ and begin year mass-at-age $W_a$ were kept the same as the operating model. The ASPM estimated a maximum of six parameters based upon the data sets used to tune it. These were (Table 2.3) the carrying capacity of the spawning biomass $K_{sp}$, the steepness parameter $h$, the commercial selectivity parameters $a_c$ and $\delta$, and the selectivity for age-zero and age-one fish for the survey trawls.

The estimated parameters were calculated using AD model builder (Otter Research Ltd, 1994). Management quantities calculated were MSY, MSYL, MSYR, DEPLETION and TAC.

## 2.3 Experiment 1 - Different data sources

There were five main "estimators" that used combinations of data that were investigated. Estimator here is defined as either different models or combinations of data sources. These were the ASPM with all data sources, with both commercial and survey indices, with only commercial data sources and with only survey data sources, and a Schaefer model that used both indices. Estimators are summarised in Table 2.4 with their corresponding code-names.

Thus the combination of the two harvest rates, four error scenarios, and the five estimators yielded a total of 40 trials, corresponding to 4000 simulations.

## 2.4 Experiment 2 - Likelihood functions

For the aim of investigating how the multinomial and weighted log-normal CAA likelihoods influence the model, these two likelihood functions (Equations 2.6 and 2.5) for the CAA data were compared. The ASPM with all data sources was used in the analyses. For the log-normal likelihood, proportions obtained from the operating model CAA matrices were used in the assessment model as the true CAA proportions.

Eight hundred additional simulations (two harvest rates x four error scenarios x 100 simulations) were run with all data sources in the ASPM model, with the multinomial likelihood. Actual numbers of fish aged were used in the CAA matrix (which was assumed to be 5000 a year) rather than proportions. This was carried out using the error scenarios and harvest rate scenarios described above in Experiment 1 such that the *ASPM - all* data scenario with the log-normal likelihood and the *ASPM - all* data scenario with the multinomial likelihood could be directly compared.

Table 2.4: The five "estimators" tested using data drawn from the operating model. An 'x' shows if a particular data source was used in the analysis for either the age-structured production model (ASPM) or the Schaefer model.

| Estimator | Name | CPUE | Survey biomass | Commercial CAA | Survey CAA |
|---|---|---|---|---|---|
| ASPM - all data | *ASPM - all* | x | x | x | x |
| ASPM - only biomass indices | *ASPM - indices* | x | x | | |
| ASPM - commercial data | *ASPM - comm* | x | | x | |
| ASPM - all research survey data | *ASPM - survey* | | x | | x |
| Schaefer model | *Schaefer* | x | x | | |

## 2.5 Experiment 3 - Optimal sample sizes

Age-determination is crucial for age-structured production models (Reeves, 2003). Error in the ageing process may accordingly have a profound effect on the model. Assuming ageing error is random and unbiased, the more fish aged, the more precision in the CAA matrix (Richards et al., 1992) . However, resources are used in the ageing of fish, so there should be an optimal amount of fish aged such that the trade off between precision and cost of ageing are apposite (Smith, 1989). There needs to be a consistent minimum number aged though, as Miller and Skalski (2006) found that fish move in shoals or "clusters" that have similar ages. Thus when these shoals are harvested and aged, they may not be fully representative of the age-distribution within the population.

Simulations were conducted to ascertain how many otoliths would result in obtaining an optimum level of precision in parameter estimates. The same operating/assessment models were used as above with only the ASPM with all data sources (*ASPM - all*). The ageing error scenario was used (*NoCorr AE*) with *Downhill trip* and *Good contrast* data. 11 sample sizes were tested: 400, 1000, 2000, 3000,....,10000. However, these sample sizes were simulated only for the last six years when surveys had been conducted. Previous to that, half the number of otoliths aged was simulated for the commercial ships. This was carried out with the weighted log-normal likelihood function initially, and then repeated with the multinomial likelihood. Thus there were 44 trials (11 sample sizes x two harvest rates x two likelihood functions).

## 2.6 Assessing model performance

There are a "bewildering" suite of measures one can use to measure model performance (Francis and Shutton, 1997, page 1706). For example, the root mean square error can be used as a measure of overall performance as well as bias, though it is extremely sensitive to outliers (Punt, 2003). The use of $R^2$ was also investigated, though it has been found to give smaller confidence interval estimates than is the true case (Polacheck et al., 1993). Punt (2003) recommends the use of the Median Absolute Relative Error (MARE) estimate, which is less sensitive to outliers and more robust.

Relative error for the estimated management quantities was measured in Experiments 1 and 2 by

$$RelativeError = \frac{\theta_{true} - \theta_{est}}{\theta_{true}} \tag{2.7}$$

where $\theta_{true}$ is the true value of a single simulation and $\theta_{est}$ is the estimated value. Ideally, one would want the relative error to be as close to zero as possible. If the relative error is negative, it would mean the estimated parameter is larger than the true parameter, and vice-verse for a positive relative error.

For each trial, 100 relative error values would be calculated. Bias was determined by visually assessing the plots of the relative error distributions. By the nature of the relative error statistic, if values are less than zero, or negatively biased, then the assessment model is *overestimating* the parameter. If the value is greater than zero, or positively biased, then the assessment model is *underestimating* the parameter.

The absolute value of each one relative error was recorded, and finally the median of these (the MARE) were used as a measure of overall performance, shown in percentage in graphs and figures.

A single initial simulation was first carried out to double-check that the MARE for all parameters were zero when no error is added at all.

The different sample sizes in Experiment 3 did not seem to have an effect on the MAREs. However, they did have an effect on another performance statistic, the coefficient of variation (CV). This was a similar finding to Coggins and Quinn (1998) who found that the relative error was not affected, but the variation was. Therefore the CV was used as a performance measure in this particular experiment. The CV value for each trial was the average of the 100 Hessian-derived CV values obtained for each individual simulation. Standard deviations were obtained from the variance-covariance matrix which were computed by AD Model Builder. Ideally, variation measures obtained from the profile likelihoods should rather be used, and are more accurate than this Hessian-based, quadratic measure (Maunder, 2001). However, due to the long computational time involved in obtaining profile likelihoods, the Hessian-based computations were used.

Only one performance measure was used per experiment. This was because using too many performance statistics can lead to "information pollution" (Punt, 1992, page 946) where conclusions cannot be reached due to the sheer volume of information.

## 2.7 Failures and extreme outliers

The number of failures in each trial was recorded for Experiments 1 and 2, as well as extreme outliers for the estimated parameters.

A failure is most likely caused by problems inverting a non-positive definite Hessian matrix (the matrix of second derivatives) when fitting the model to the data. This indicates the function minimiser may not have obtained a unique minimum for the objective function (Maunder, 2001). It may also imply "deficiencies" in the data (Hilborn, 1979) and possible problems with parameters too close to pre-defined minima and maxima.

Extreme outliers are defined as having a MARE greater than 100% (Punt, 2003).

## 2.8 Interactions

To gain an understanding of the interaction of the various kinds of estimators, likelihoods, error and harvesting rates, General Linear Models (GLM) were run for each of the five management reference points for both Experiment 1 and 2. This is a novel approach in the comparison of performance measures in stock assessment. The arcsin of the square-root of the absolute relative errors transformation was used in order to stabilise the variance. Levene's test was used to check for equality of variance, and Scheffé's post-hoc test was used to find the individual significant differences. Significance was assessed at a 5% level.

The response variable was the absolute relative error. For Experiment 1, the factors were the types of estimators $\alpha^*$(five levels), errors scenarios $\beta^*$ (four levels), and *Downhill trip* and *Good contrast* harvest rates $\gamma^*$ (two levels). The equation for the GLM was thus

$$Y_{ijkl} = \mu^* + \alpha_i^* + \beta_j^* + \gamma_k^* + \delta_{ij}^* + \tau_{ik}^* + \omega_{jk}^* + \varsigma_{ijk}^* + \varepsilon_{ijkl}^*$$

for $i = 1, ...., 5$ $j = 1, ..., 4$, $k = 1, 2$, $\varepsilon_{ijkl}^* \sim N(0, \sigma^2)$ and $\mu^*$ is the overall mean, $\delta_{ij}^*$, $\tau_{ik}^*$, $\omega_{jk}^*$ and $\varsigma_{ijk}^*$ are the corresponding interactions.

For Experiment 2, the factors were the two different likelihood functions $\alpha'$ (two levels), errors scenarios $\beta^*$ (four levels), and *Downhill trip* and *Good contrast* harvest rates $\gamma^*$ (two levels). The equation for the GLM was thus

$$Y_{qjkl} = \mu' + \alpha_q' + \beta_j^* + \gamma_k^* + \delta_{qj}' + \tau_{qk}' + \omega_{jk}' + \varsigma_{qjk}' + \varepsilon_{qjkl}'$$

for $q = 1, 2$, $j = 1, ..., 4$, and $k = 1, 2$, $\varepsilon_{qjkl}' \sim N(0, \sigma^2)$, $\mu'$ is the overall mean, $\delta_{qj}'$, $\tau_{qk}'$, $\omega_{jk}'$ and $\varsigma_{qjk}'$ are the corresponding interactions.

## 2.9 Choice of 100 simulations

The minimum number of simulations used is generally 100 simulations per scenario (Bradford, 1991; Francis and Shutton, 1997; Ianelli, 2002; Punt, 2003; Reeves, 2003). This number was chosen in this study due to reasonable computational time. To investigate if Monte Carlo variability had been accounted for with 100 simulations rather than a larger number, Punt (1991) conducted a study where 2000 simulations were generated and the corresponding parameters calculated. This was compared to parameter estimates calculated from 100 simulations. Punt (1991) found that although there were certain statistics that were slightly imprecise, in general the Monte Carlo variability did not have a sufficient effect on conclusions of his study to warrant criticism.

A similar comparison procedure to Punt (1991) was conducted in this study, using the *Baseline* error scenario and *Good contrast* harvest rate with all data included in the analysis. The trial with 2000 simulations took approximately eight hours of computer time. The results of the comparison are summarised in Table 2.5. The largest relative difference was 3%, with the variability reduced with increased simulations. The difference was not considered significant and 100 simulations were used for the rest of the study. The comparison equation is

$$\frac{MARE(\theta_{2000}) - MARE(\theta_{100})}{MARE(\theta_{100})} \times 100$$

where $MARE(\theta_{100})$ and $MARE(\theta_{2000})$ are the median absolute relative error of a parameter for 100 and 2000 simulations, respectively.

Table 2.5: Comparison between 100 simulations and 2000 simulations for management reference points. The second column shows the difference between MAREs for 100 and 2000 simulations of the MARE of 100 simulations.

| Management parameter | Comparison |
|---|---|
| MSY | 0% |
| MSYL | -2% |
| MSYR | -2% |
| DEPLETION | -3% |
| TAC | -3% |

# Chapter 3

# Results

## 3.1 Generated data sets

Examples of commercial CPUE and CAA series generated for the simulation period are presented in Figures 3.1 and 3.2.

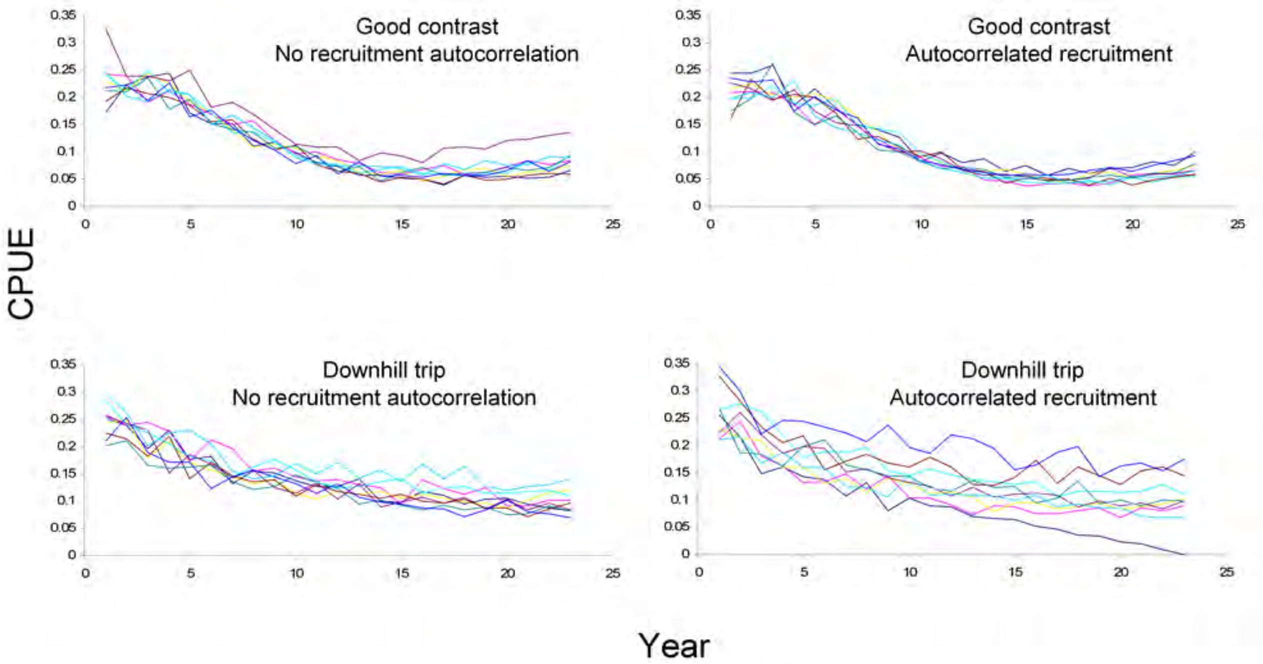


Figure 3.1: Ten randomly selected single simulations over the 23 year simulation period given two harvest rates and two process error combinations. All simulations include observation error.

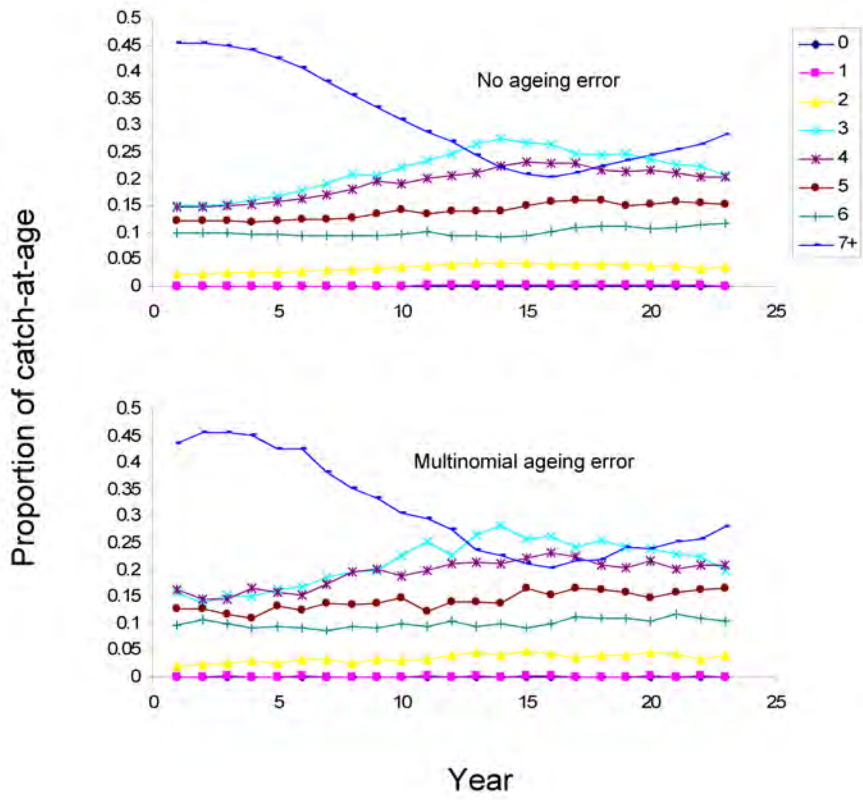


Figure 3.2: Proportions of catch-at-age for *Good contrast* data with and without ageing error, drawn from a multinomial distribution with a sample size of 5000 otoliths.

## 3.2 Experiment 1: Comparison of data sources

### 3.2.1 Bias

Figures 3.3 and 3.4 illustrate the relative error distributions for *Downhill trip* data and *Good Contrast* data, respectively.

Relative error variability tended to be largest and more biased for the *Downhill trip* data. One exception was that *Downhill trip* data were generally less variable (though still more biased) than *Good contrast* data when the model estimated TAC. As expected, the *Baseline* error scenario was generally the least variable due to the presence of observation error in the abundance indices only. Relative errors were slightly more variable than the *Baseline* case with the addition of ageing error (the *NoCorr AE* scenario). Scenarios with autocorrelation in the recruitment residuals (*Corr NoAE* and *Corr AE*) often had substantially higher variation than the *Baseline* scenario. Relative error variability was generally high when only survey data were used *(ASPM - survey).*

In terms of management quantities, general results showed that MSYL was often the best estimated management quantity, with, in general, the least variation and bias. TAC was, on the whole, the most variable and exhibited stronger bias given the *Downhill trip* data, underestimating TAC. The Schaefer produced DEPLETION and TAC estimates that were considerably less biased and variable than the ASPM model.

MSY was consistently underestimated in all ASPM scenarios, and consistently overestimated in the Schaefer model. This suggests the ASPM model outcomes would be more likely to result in lower catches than what is possible based on the true population. The Schaefer model would, by contrast, be too optimistic, showing unsustainably high catches.

MSYL, the size of the stock needed to obtain MSY, seems to be the least biased parameter in all ASPM trials. The Schaefer model shows that for all error scenarios, MSYL was consistently overestimated.

MSYR was consistently underestimated for both assessment models. The ASPM trials are relatively right-skewed, while the Schaefer model trials are more left-skewed.

The relative distributions for DEPLETION were highly variable when only survey data were used in the analysis. It was, however, relatively unbiased. DEPLETION was underestimated by ASPM *Good contrast* data, suggesting a healthier stock than what is known. The Schaefer model, by contrast, was too conservative - predicting a resource that is overly depleted.

For the *Downhill trip* data for every trial, TAC was constantly underestimated. Estimated TACs for *Good contrast* data were less biased but more variable.

### 3.2.2 Evaluation of Median Absolute Relative Errors

Typically, all management quantities had lower MAREs with *Good contrast* data, and when there was no autocorrelation within the stock-recruitment relationship. The Schaefer model parameter estimates, in particular, were often affected by autocorrelated recruitment. The most suitable estimator was generally the ASPM with all data included in the analysis, the best harvest rate scenario was *Good contrast* data, and the best error scenario was the *Baseline* case. The models performed badly when there was autocorrelated recruitment within the stock-recruitment relationship.

For all 40 trials, the MAREs for MSY for all ASPM data sources ranged between 13 and 27%, slightly lower than those of the Schaefer model which were clumped at 20% and 32% (Figure 3.5).

MSYL MAREs were generally higher in the ASPM trials when there was autocorrelated recruitment. For all ASPM trials, the MAREs ranged between 5% and 23%. For the Schaefer model, the MAREs were higher ranging from 15% to 38%. MSYL was reasonably well estimated when all data were used in the model and when only the commercial data series were used. The estimators gave the most inaccurate management quantity estimates when there was no ageing data. Even the trials with only the six years of survey biomass and CAA data performed better in general than when there was no CAA data. A clear

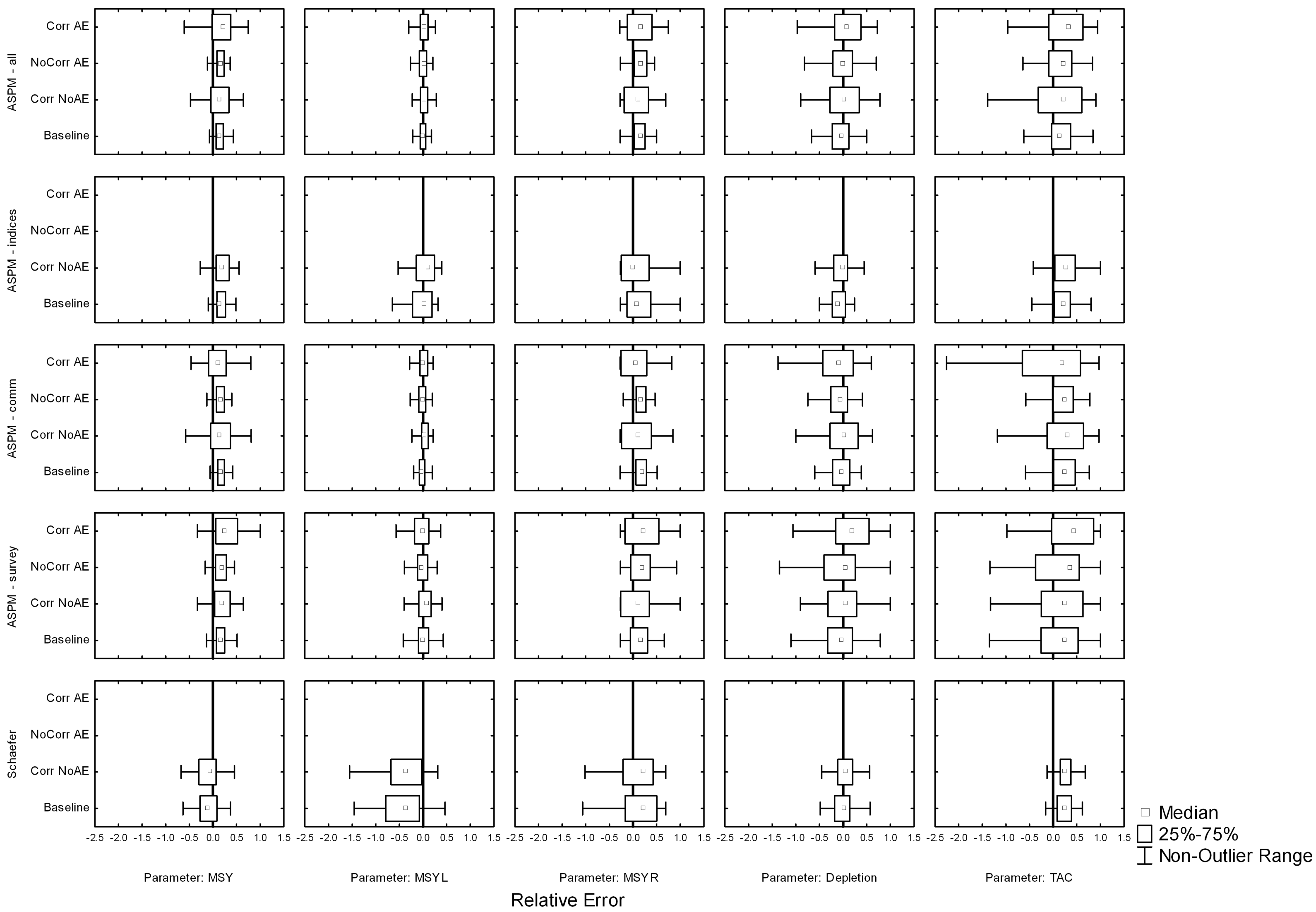


Figure 3.3: Relative error distributions for five management quantities with *Downhill trip* data for Experiment 1. The error scenarios included *Baseline*, which had no autocorrelation in recruitment in the stock-recruitment relationship and no ageing error, *Corr NoAE* which had autocorrelation in the stock-recruitment relationship and no ageing error, *NoCorr AE* which had no autocorrelation in the stock-recruitment relationship and multinomial ageing error in the CAA data and *Corr AE* which had both correlation in the stock-recruitment relationship and multinomial ageing error. *ASPM - all* used all data (commercial and survey biomass indices and corresponding CAA data) in the Age-structured production model (ASPM), *ASPM - indices* used biomass indices from commercial and survey trawls in the ASPM, *ASPM - comm* which used only commercial data in the ASPM, *ASPM - survey* which used only the survey data and lastly, the Schaefer model which used both commercial and survey data.

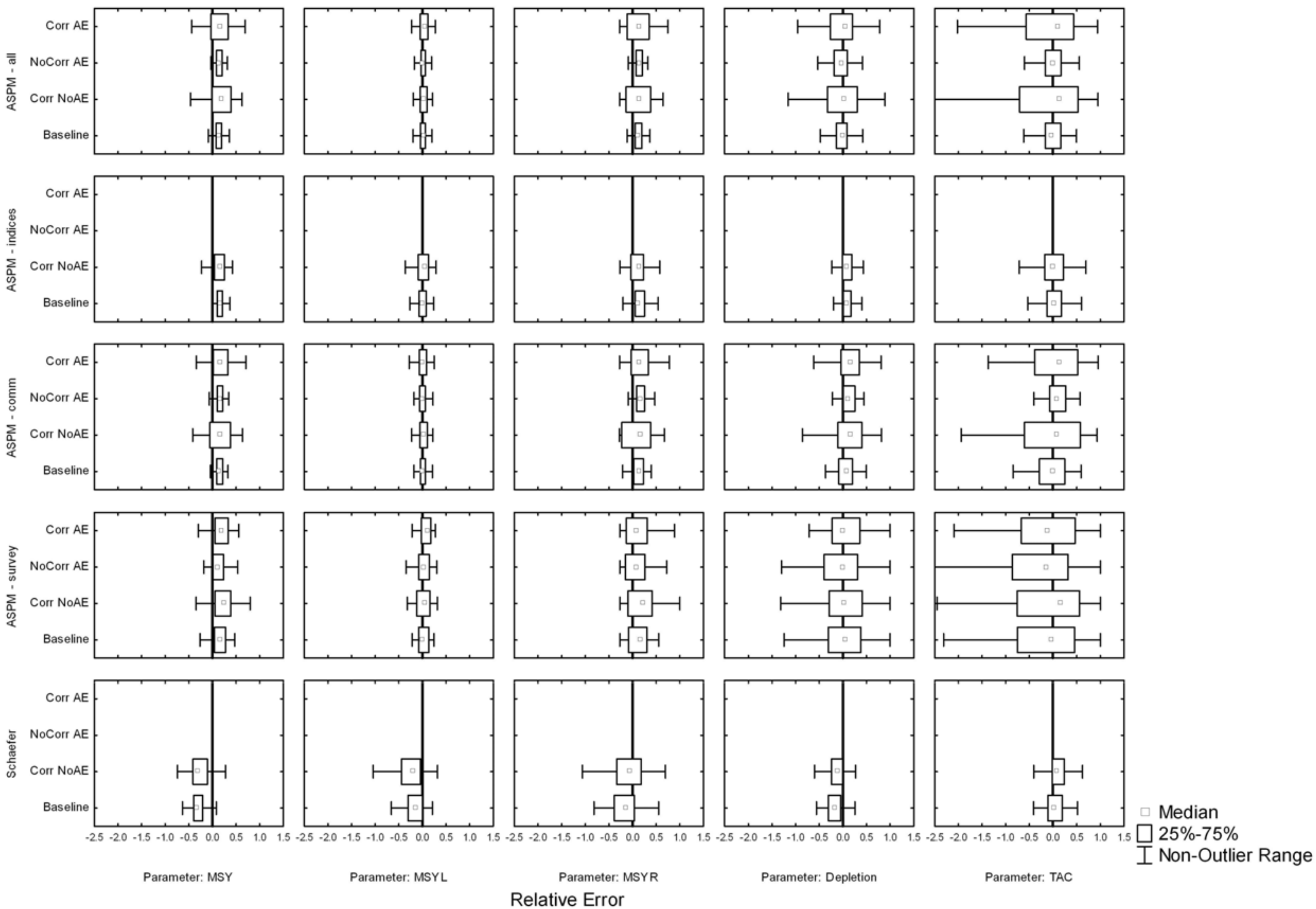


Figure 3.4: Relative error distributions for trials with *Good contrast* data for Experiment 1. See Figure 3.3 for an explanation of scenarios.

trend in the data were noticeable in Figure 3.6 where *Downhill trip* data had a negative effect on the model when there was no ageing data for tuning the model.

For MSYR, the ASPM trials generally had MAREs that ranged between 13% and 27%. The Schaefer model had MAREs between 24% and 40%. The *Good contrast* data generally obtained lower MAREs than the *Downhill trip* data, particularly for the Schaefer model (Figure 3.7).

DEPLETION was better estimated by the Schaefer model. While it ranged between 13% and 20%, the ASPM ranged from 11% to a maximum of 37% (Figure 3.8).

TAC was the most poorly estimated parameter with MAREs ranging from 16%-67% for the ASPM trials, and between 15%-27% for the Schaefer model. The Schaefer model outperformed the ASPM model for TAC even when all data sources were available to the ASPM (Figure 3.9).

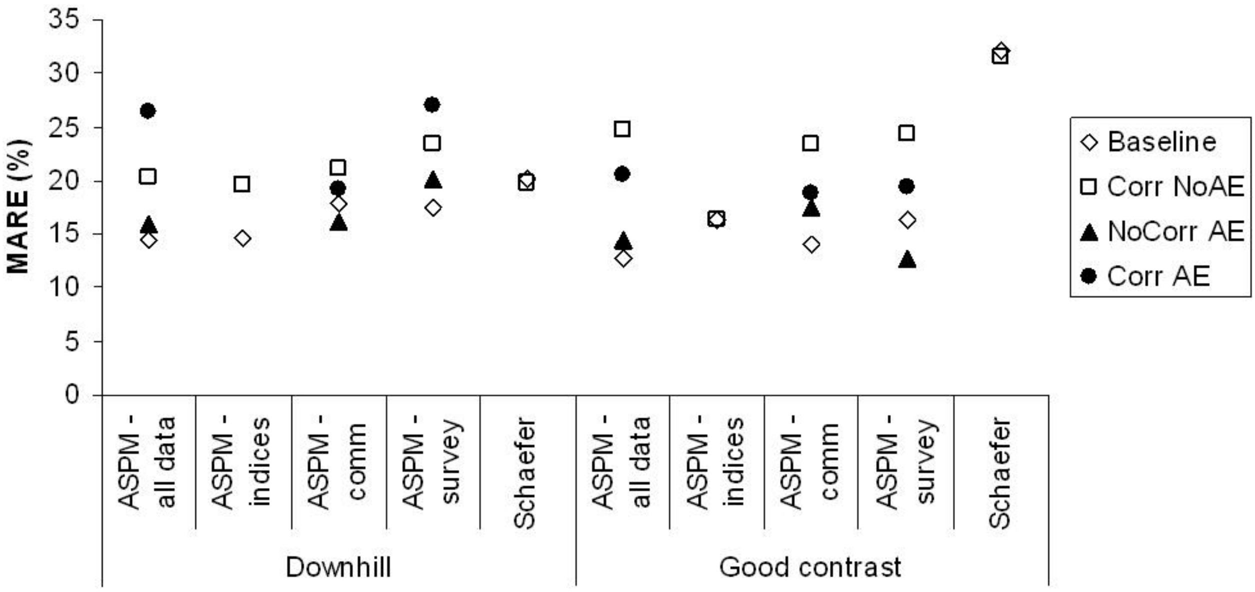


Figure 3.5: MSY Median Absolute Relative Error (MARE) values for Experiment 1. Harvest scenarios had *Downhill trip* data and *Good contrast* data . The error scenarios included *Baseline*, which had no autocorrelation in the stock-recruitment relationship and no ageing error, *Corr NoAE* which had autocorrelation in the stock-recruitment relationship and no ageing error, *NoCorr AE* which had no autocorrelation in the stock-recruitment relationship and multinomial ageing error in the catch-at-age (CAA) matrix and *Corr AE* which had correlation in the stock-recruitment relationship and multinomial ageing error. *ASPM - all* used all data (commercial and survey biomass indices and corresponding CAA data) in the age-structured production model (ASPM), *ASPM - indices* which used biomass indices from commercial and survey trawls in the ASPM, *ASPM - comm* which used only commercial data in the ASPM, *ASPM - survey* which used only the survey data, and lastly the Schaefer model which used both commercial and survey biomass indices.

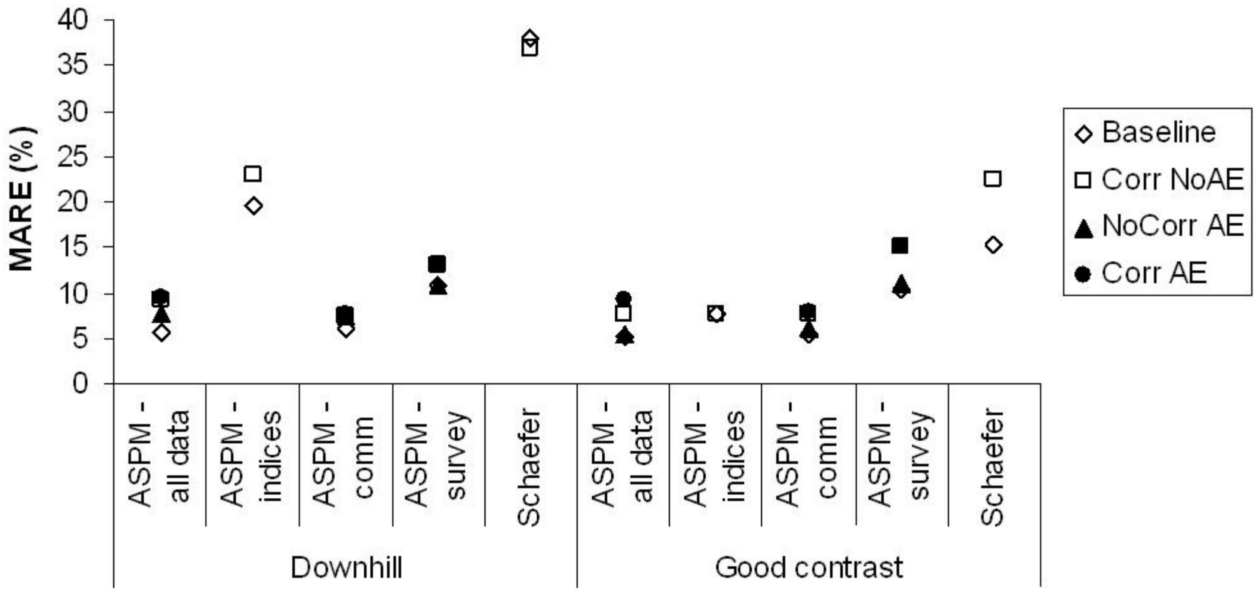


Figure 3.6: MSYL Median Absolute Relative Error (MARE) values for Experiment 1. See Figure 3.5 for an explanation of scenarios.

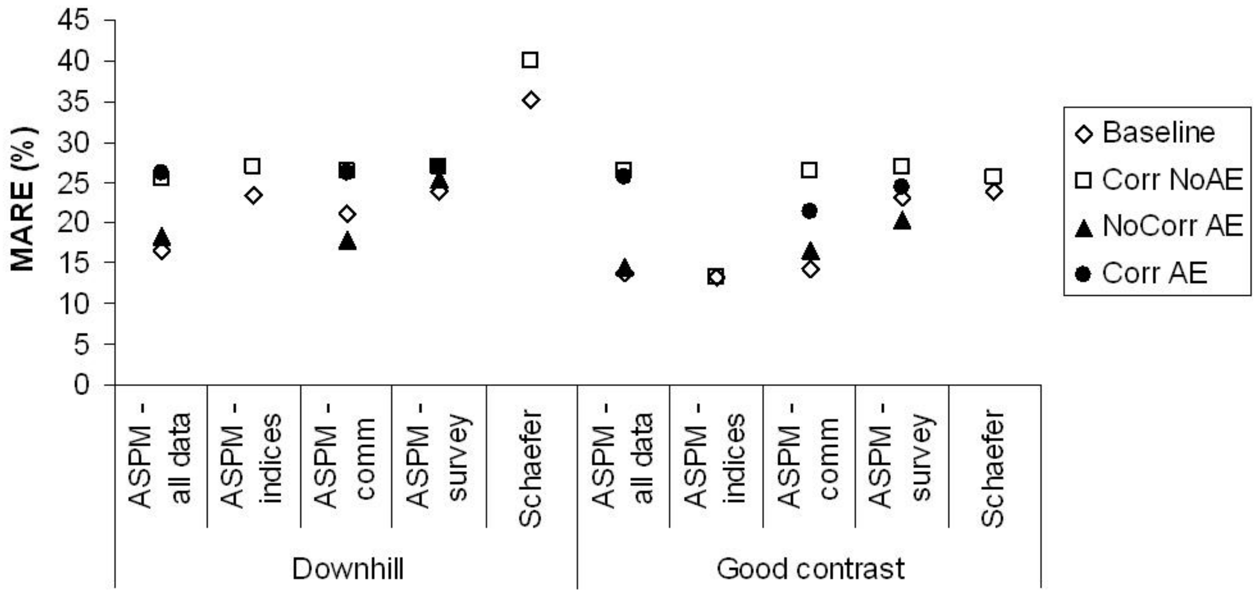


Figure 3.7: MSYR Median Absolute Relative Error (MARE) values for Experiment 1. See Figure 3.5 for an explanation of scenarios.

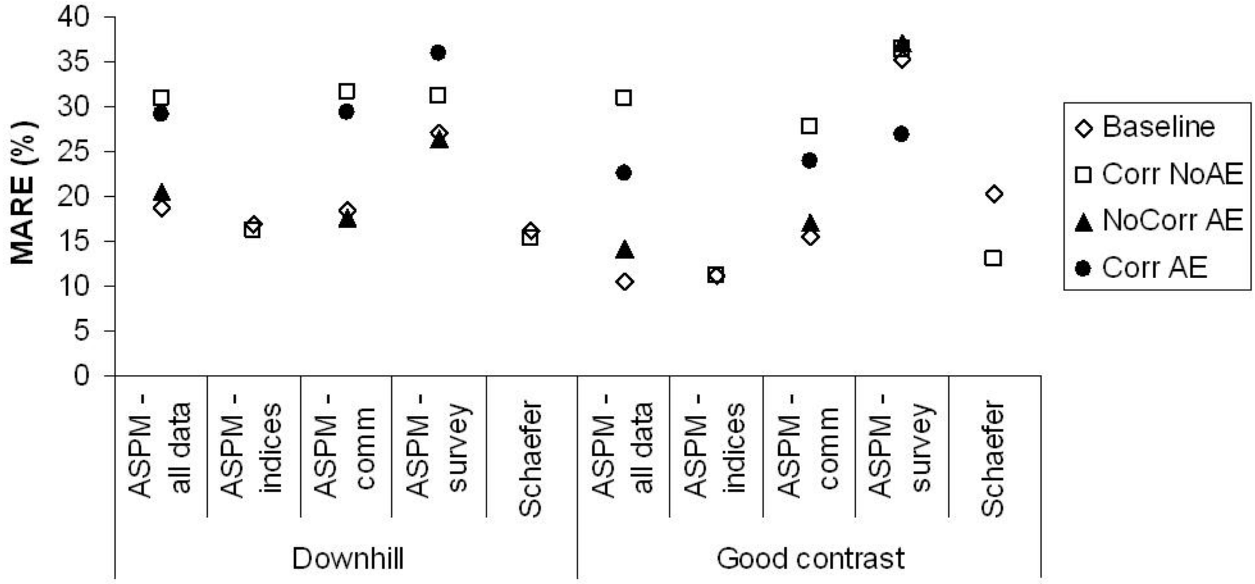


Figure 3.8: DEPLETION Median Absolute Relative Error (MARE) values for Experiment 1. See Figure 3.5 for an explanation of scenarios.

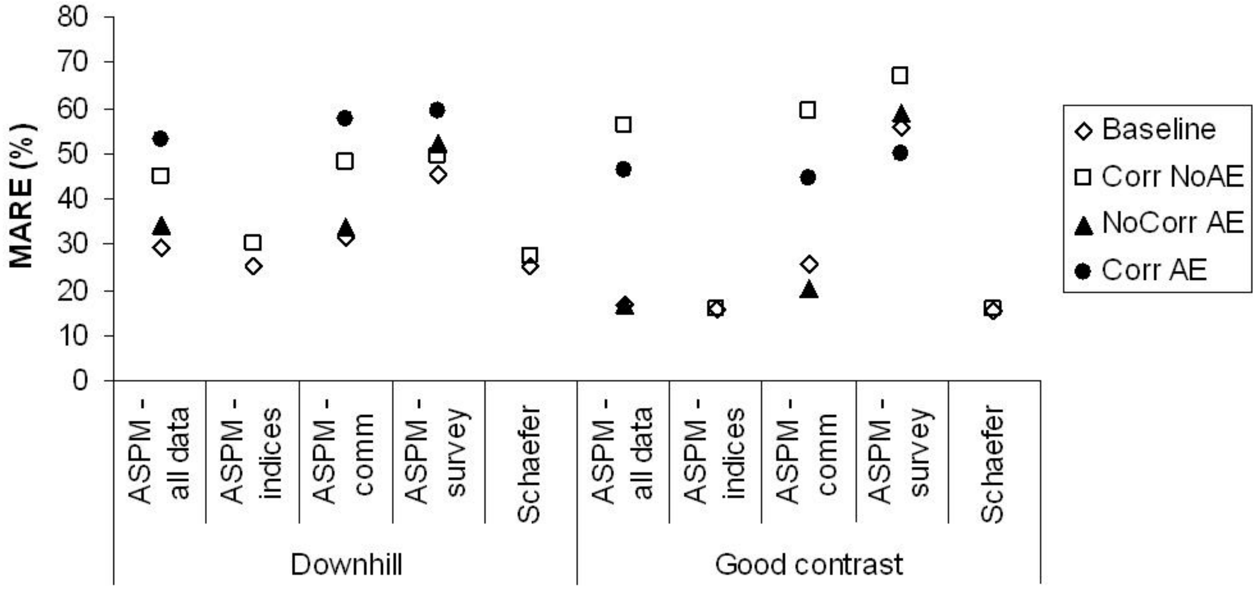


Figure 3.9: TAC Median Absolute Relative Error (MARE) values for Experiment 1. See Figure 3.5 for an explanation of scenarios.

### 3.2.3 Interactions

In general, the GLMs showed there were significant differences between the management quantity absolute error values when the two different harvest rates were used. On occasion, no significant differences were found between the *Baseline* error scenario and the error scenario with ageing error *NoCorr AE.* Error scenarios with autocorrelation in the stock-recruitment relationship, were frequently significantly different from the other two error scenarios. The ASPM estimators were generally not significantly different to each other, except when the survey data were used. The Schaefer model was generally significantly different from the ASPM models, given the different error scenarios.

For the MSY GLM, significant differences were noted between estimator effects (F-statistic=20.3; df=(4,3566); p<0.001), error effects (F-statistic=36.7; df=(3, 3566); p<0.001), interaction between the estimator and error (F-statistic=2.6; df=(12,3566); p=0.001) and the interaction between the harvest rate and estimator (F-statistic=18.7; df=(4,3566); p<0.001). Harvest rates had an insignificant effect on MSY absolute values (F-statistic=0.2; df=(1,3566); p=0.683). The ASPM with all data sources had significantly lower absolute error values with the Schaefer model having the highest (Scheffé p<0.001). The Schaefer model's absolute error values were significantly higher when there were *Good contrast* data compared to most ASPM trials with *Good contrast* data (Scheffé p<0.001). The error scenario that yielded the highest absolute error values was when there was autcorrelation present in the stock-recruitment relationship in conjunction with ageing error (Scheffé p<0.001).

The GLM for MSYL showed a clear pattern of the use of a *Downhill trip* harvest rate data effect as it yielded higher absolute errors compared to that of the *Good contrast* data. There were significant differences for all predictors and interactions except for the interaction between harvest rate and type of error. [Harvest rate effect (F-statistic=107.6; df=(1,3498); p<0.001). Estimator effect (F-statistic=246.4; df=(4,3498); p<0.001). Error effect (F-statistic=13.3; df=(3,3498); p<0.001). Harvest and estimator interaction (F-statistic=20.6; df=(4,3498); p<0.001). Estimator and error interaction (F-statistic=2.2; df=(12,3498); p=0.0083). Harvest rate, estimator type and error interaction (F-statistic=1.9; df=(12,3498); p=0.036)]. It was clear that the Schaefer model performed poorly, whereas the ASPM, with all data sources included, performed the best for every simulation trial (Scheffé p<0.001).

The GLM for MSYR showed significant differences between the harvest rate effects (F-statistic=87.0; df=(1,3550); p<0.001), the estimator effects (F-statistic=27.6; df=(4,3550); p<0.001), the error effects (F-statistic=36.4; df=(3,3550); p<0.001), the interaction of the harvest and the estimator (F-statistic=8.9; df=(4,3550); p<0.001) and the interaction of the estimator and the error (F-statistic=2.7; df=(12,3550); p<0.001). When only survey data were used there were no significant difference between the two harvest rates (Scheffé p=0.440). The ASPM with all data sources, with only CPUE, and with all commercial data all performed equally well under the *Baseline* error scenarios. The Schaefer model and the ASPM obtained higher absolute values when only abundance indices were included in the analysis. They were particularly affected by the *Downhill trip* harvest data (Scheffé p<0.001) .

For DEPLETION there were a great number of significant differences. Harvest rates were found to be significant (F-statistic=4.7; df=(1,3551); p=0.031), and all their interactions involving the harvest rates [Harvest rate and estimator (F-statistic=2.9; df=(4,3551); p=0.021), Harvest rate and error (F-statistic=3.1; df=(3,3551); p=0.027). The estimator that only used the six years of survey information with the *Good contrast* data performed poorly for every error scenario (Scheffé p<0.001). The Schaefer model absolute errors were significantly different to that of the ASPM with all data sources, with indices and with commercial data (Scheffé p<0.02).

Lastly, for TAC, there were significances found between all predictors and their interactions [Harvest rate effect (F-statistic=77.8; df=(1,3400); p<0.001). Estimator effect (F-statistic=112.0; df=(4,3400); p<0.001). Error effect (F-statistic=54.4; df=(3,3400); p<0.001). Harvest rate and estimator interaction (F-statistic=2.6; df=(4,3400); p=0.0333). Harvest rate and error interaction (F-statistic=5.5; df=(3,3400); p<0.001). Estimator and error interaction (F-statistic=6.3; df=(12,3440); p<0.001). Harvest, estimator and error interaction (F-statistic=2.0; df=(12,3400); p=0.022).] The *Downhill trip* data had significantly higher absolute errors than the *Good contrast* data (Scheffé p<0.001). The Schaefer model estimated this quantity the best, whereas the ASPM with only survey data estimated it the worst (Scheffé p<0.001) as shown by MAREs.

### 3.2.4 Failures and extreme outliers

There were a total of 353 failures - equating to nine percent of all simulations. Of these, the most failures occurred when the data had ageing error (Figure 3.10).

When focusing on different estimator trials, when only the survey data were used in the ASPM, most of the failures occurred. The ASPM with only biomass indices had the least failures. The *Downhill trip* harvesting scenarios had the same number of failures as *Good contrast* data scenarios.

There were 1094 extreme outliers equating to 2.5% of the 43000 absolute error values for all estimated parameters and quantities (not just the five derived management quantities). The ASPM with only survey data were responsible for 47% of the outliers, followed by ASPM with all data sources which were responsible for 28% of extreme outliers. (Figure 3.11). Most of these extreme outliers in the two estimators were due to poor estimation of age-zero selectivity for survey trawls as well the catchability coefficient for the survey trawls. This is presumably because there were only six years of data to estimate these parameters, and therefore insufficient information in the data to estimate them.

There was no substantive difference between the error scenarios, with the four each contributing roughly 25% of the total number of extreme outliers.

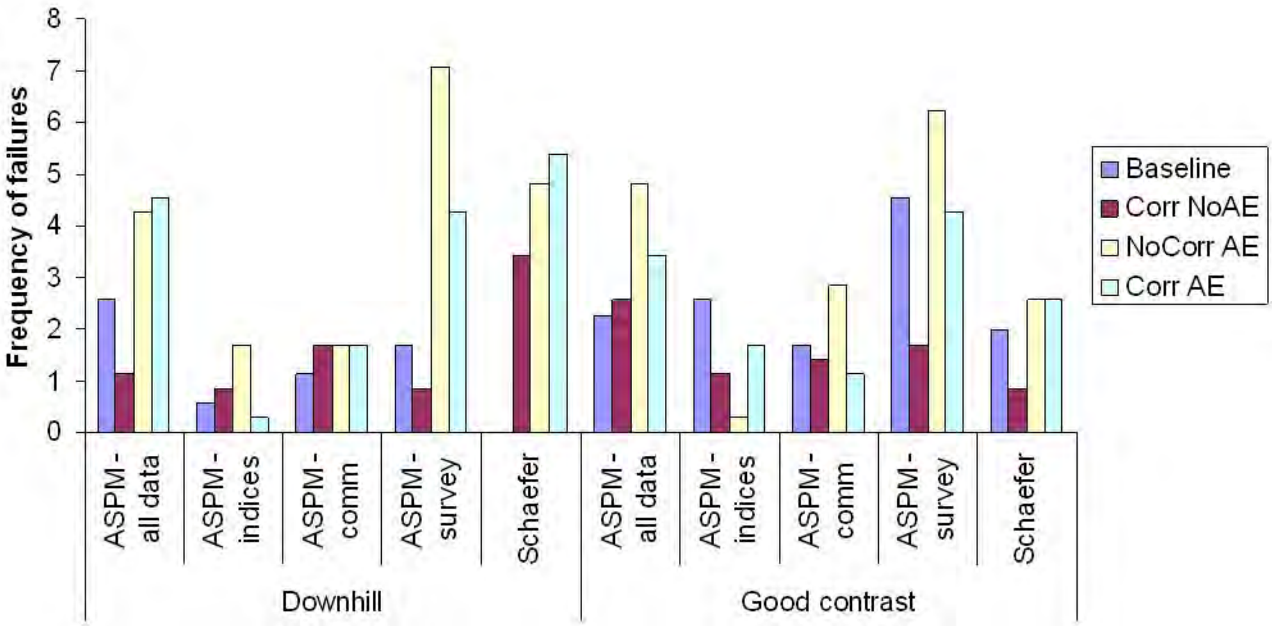


Figure 3.10: Frequency of failures for various estimators trials for Experiment 1 as a percentage of all simulations. Harvest scenarios had *Downhill trip* data and *Good contrast* data . The error scenarios included *Baseline*, which had no autocorrelation in the stock-recruitment relationship and no ageing error, *Corr NoAE* which had autocorrelation in the stock-recruitment relationship and no ageing error, *NoCorr AE* which had no autocorrelation in the stock-recruitment relationship and multinomial ageing error in the catch-at-age (CAA) matrix and *Corr AE* which had correlation in the stock-recruitment relationship and multinomial ageing error. *ASPM - all* used all data (commercial and survey biomass indices and corresponding CAA data) in the age-structured production model (ASPM), *ASPM - indices* which used biomass indices from commercial and survey trawls in the ASPM, *ASPM - comm* which used only commercial data in the ASPM, *ASPM - survey* which used only the survey data, and lastly the Schaefer model which used both commercial and survey biomass indices.

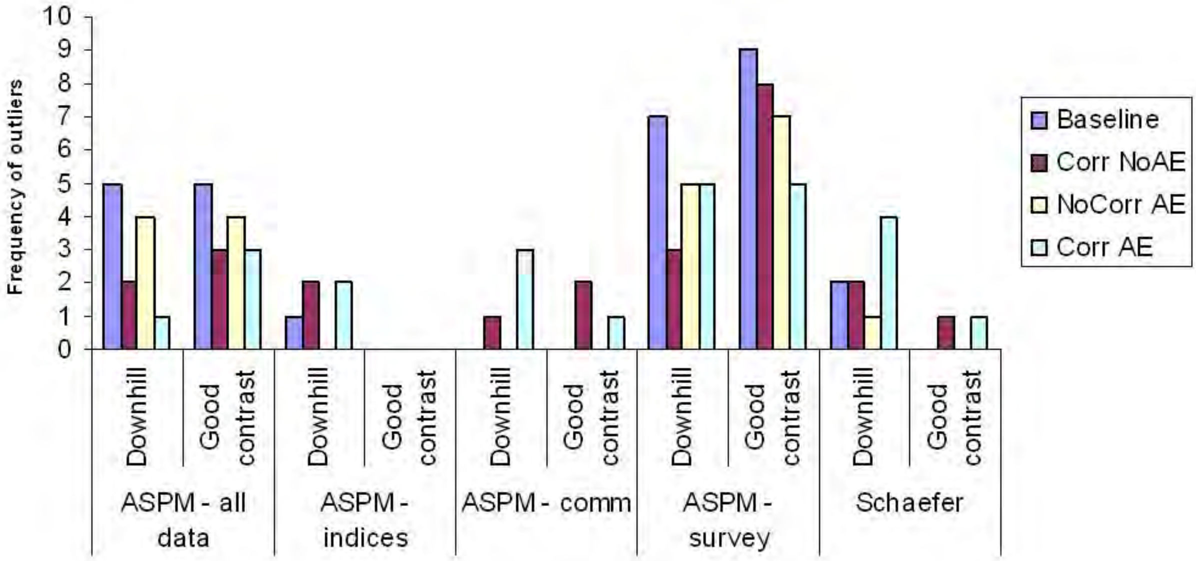


Figure 3.11: Frequency of outliers for Experiment 1 as a percentage of the total. See Figure 3.10 for an explanation of scenarios.

## 3.3 Experiment 2: Comparison of two likelihood functions for catch-at-age

### 3.3.1 Bias

Box-and-whisker plots for the relative error estimates for the different log-likelihood functions are illustrated in Figures 3.12 and 3.13. Overall, the bias was similar for the two likelihood functions considered in the experiment. Use of the multinomial likelihood gave more variable relative error estimates for DEPLETION and TAC than that of the log-normal when there was no autocorrelation in the stock-recruitment relationship.

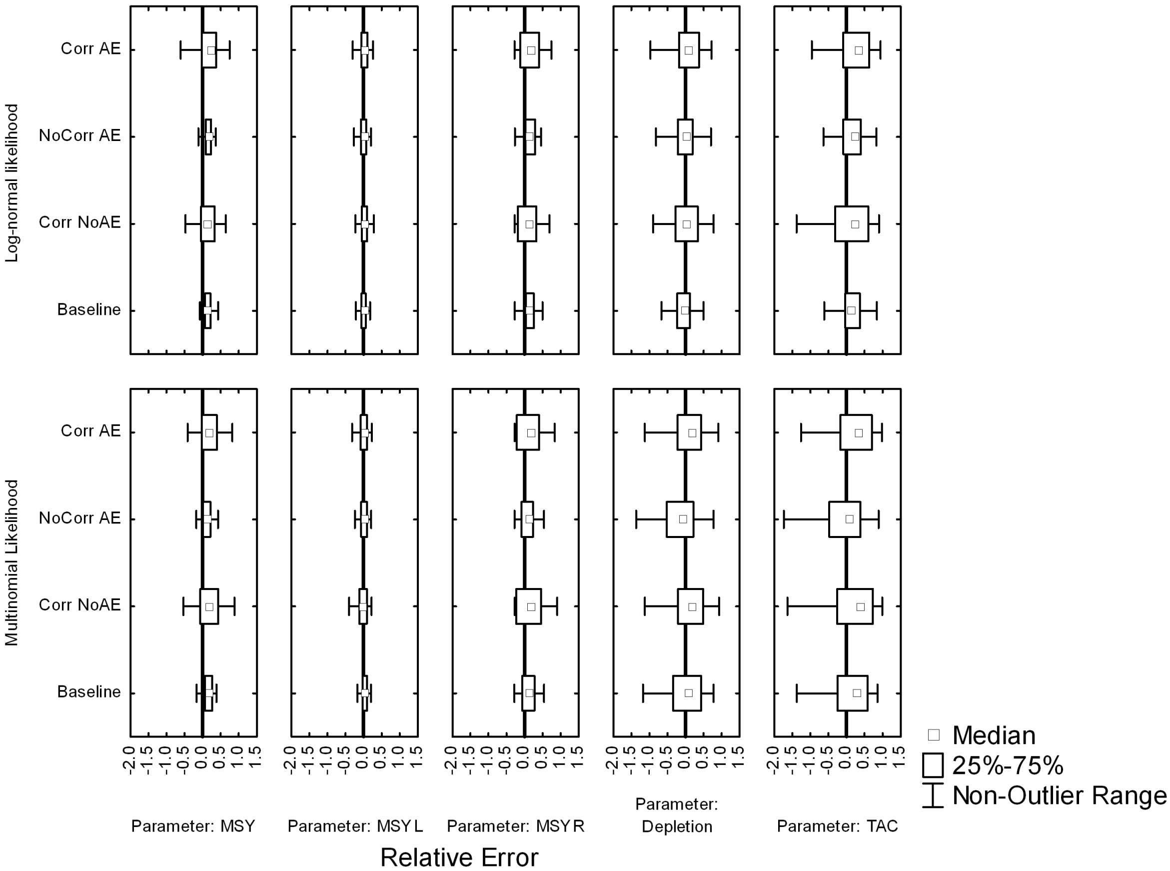


Figure 3.12: Relative error distributions for five management quantities with *Downhill trip* data for Experiment 2. The error scenarios included *Baseline*, which had no autocorrelation in recruitment in the stock-recruitment relationship and no ageing error, *Corr NoAE* which had autocorrelation in the stock-recruitment relationship and no ageing error, *NoCorr AE* which had no autocorrelation in the stock-recruitment relationship and multinomial ageing error in the CAA matrix and *Corr AE* which had both correlation in the stock-recruitment relationship and multinomial ageing error. Two likelihood functions for the CAA matrix were investigated, the log-normal and the multinomial.

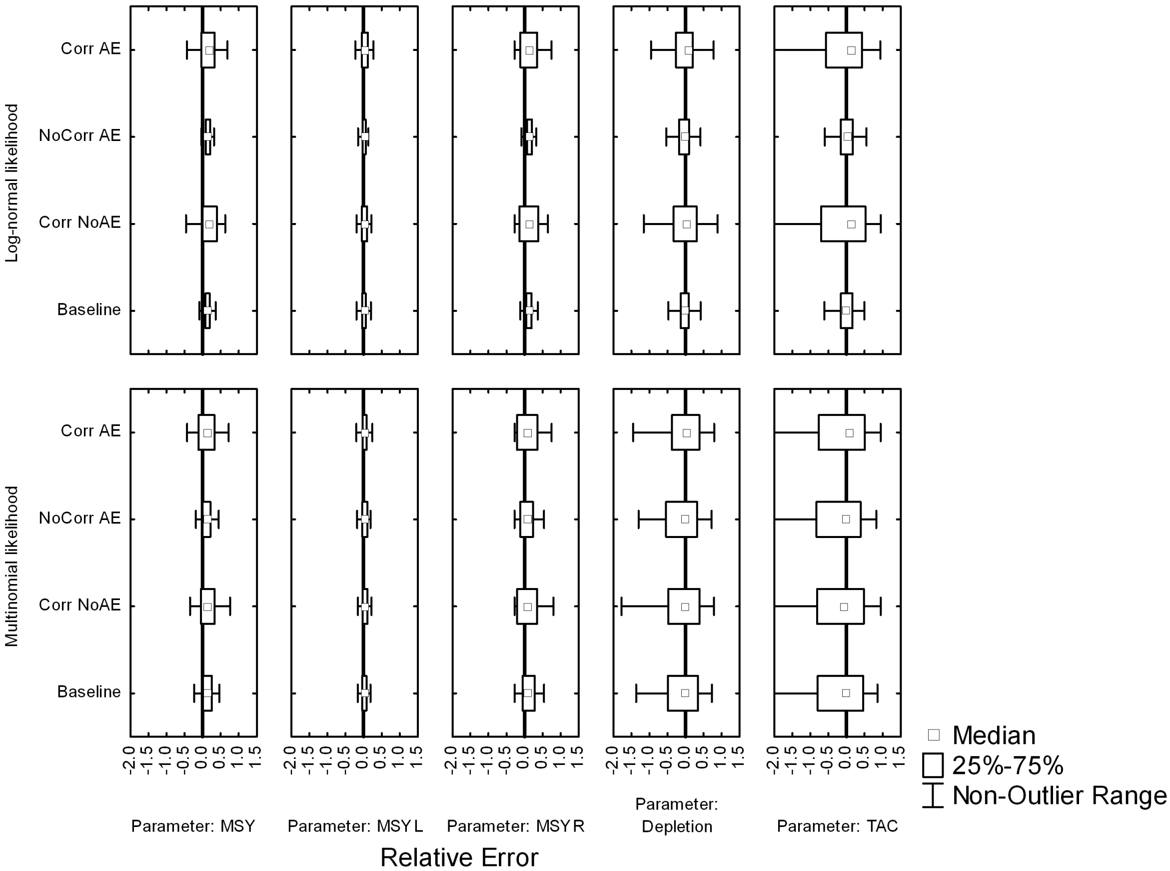


Figure 3.13: Relative error distributions for trials with *Good contrast* data for Experiment 2. See Figure 3.12 for an explanation of scenarios.

### 3.3.2 Evaluation of Median Absolute Relative Errors

In general, the MAREs were similar when log-normal and multinomial likelihoods were utilised in the catch-at-age matrix . Notably, the MAREs were larger for the *Downhill trip* data.

The MAREs for MSY, MSYL and MSYR were roughly equivalent for both the log-normal and the multinomial likelihoods, with MSYL being the best estimated parameter with values ranging from 5 to 10% (Figures 3.14 - 3.18).

When the log-normal likelihood was used, MAREs for DEPLETION and TAC were consistently lower than the multinomial by approximately 10%. This may show that the log-likelihood function is preferable when estimating these two management quantities at large sample sizes such as the one used in this study (5000 otoliths).

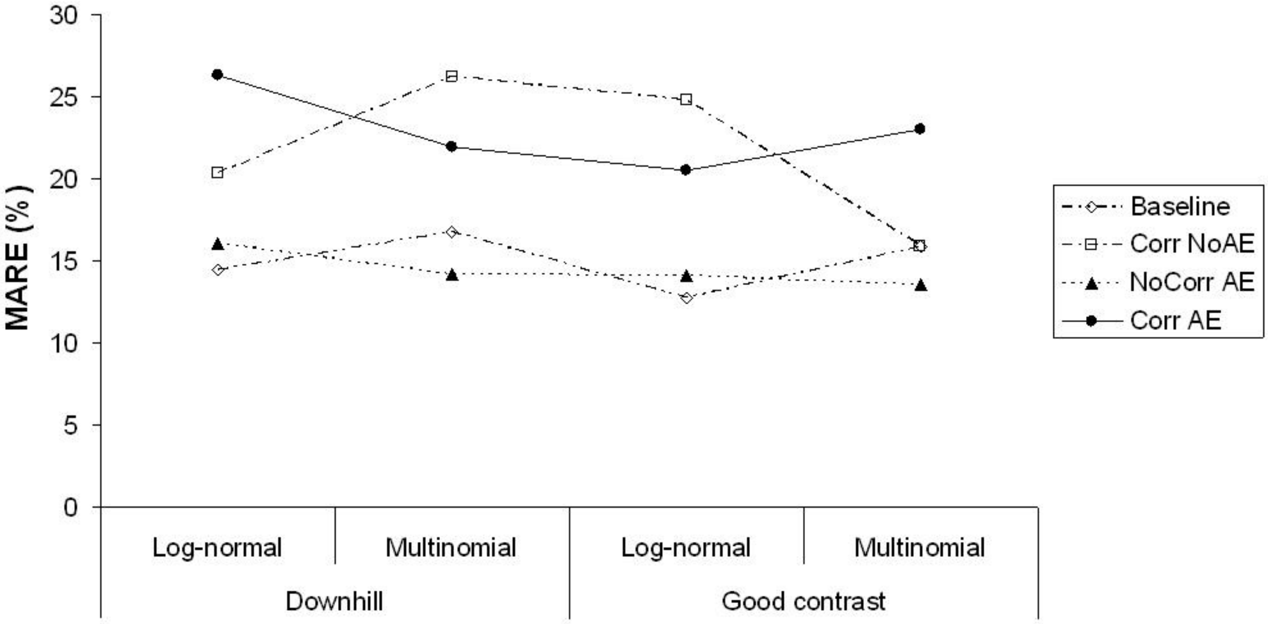


Figure 3.14: MSY MAREs for Experiment 2. There was random (uncorrelated and autocorrelated) recruitment variation in the stock-recruitment relationship and multinomial ageing error in the catch-at-age data. Two likelihoods were used in the assessment model CAA matrix, the log-normal likelihood and the multinomial likelihood. The error scenarios included *Baseline*, which had no autocorrelation in recruitment in the stock-recruitment relationship and no ageing error, *Corr NoAE* which had autocorrelation in the stock-recruitment relationship and no ageing error, *NoCorr AE* which had no autocorrelation in the stock-recruitment relationship and multinomial ageing error in the CAA matrix and *Corr AE* which had both correlation in the stock-recruitment relationship and multinomial ageing error.

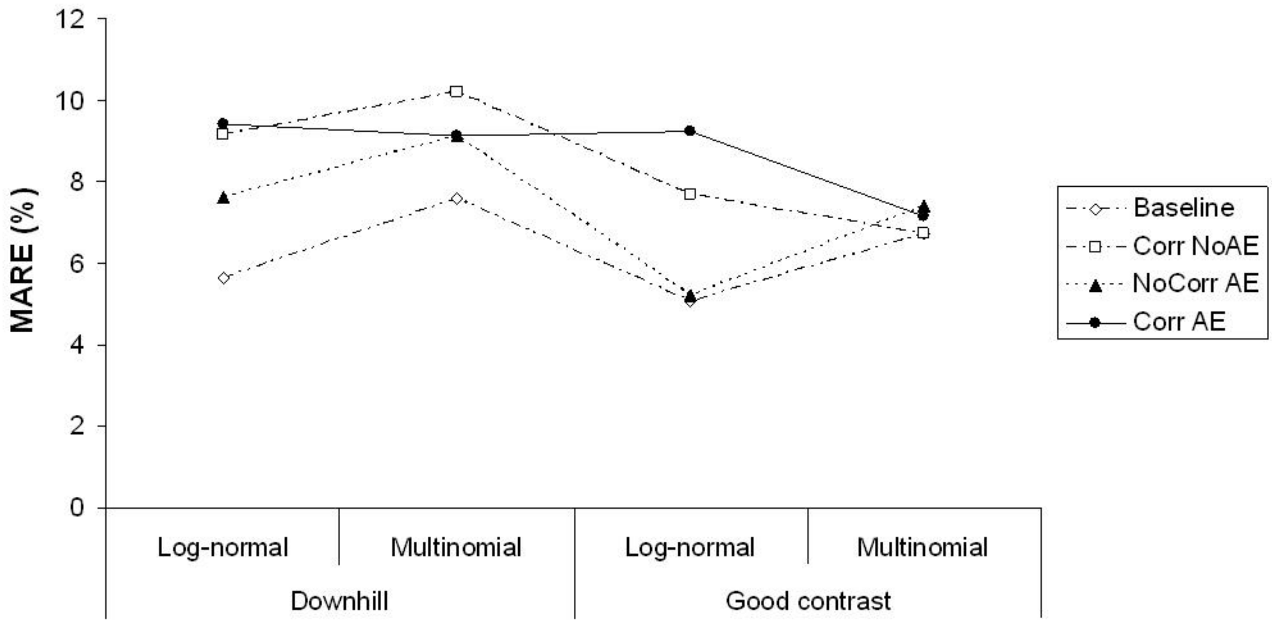


Figure 3.15: MSYL MAREs for Experiment 2. See Figure 3.14 for an explanation of scenarios.

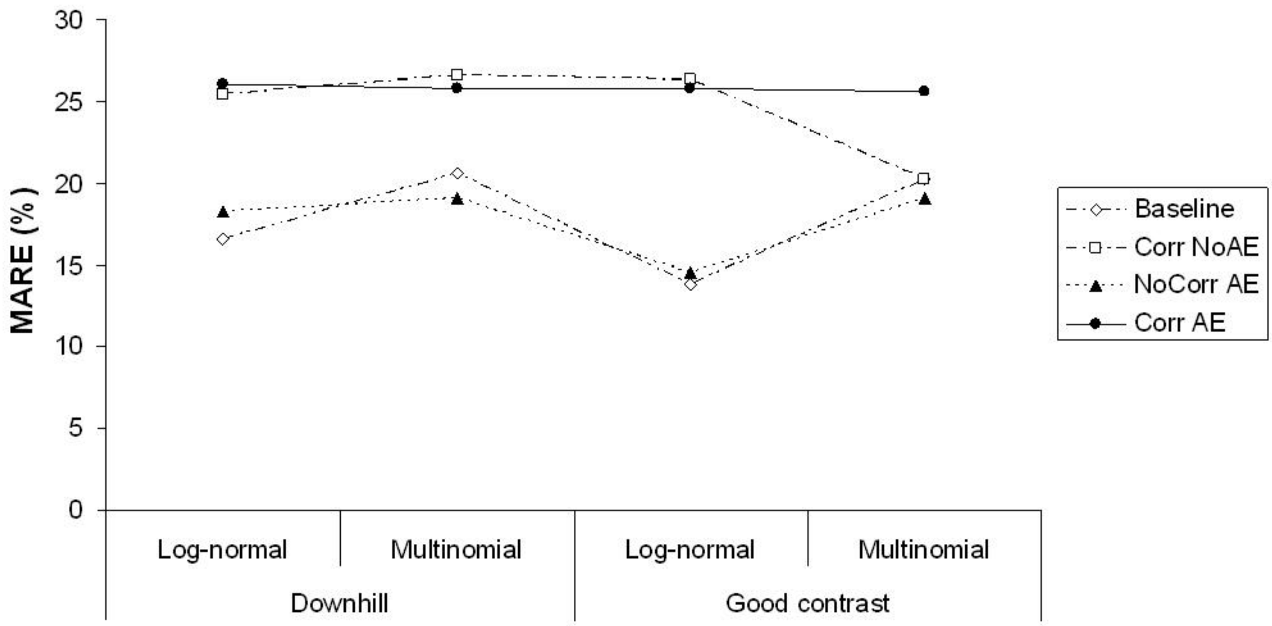


Figure 3.16: MSYR MAREs for Experiment 2. See Figure 3.14 for an explanation of scenarios

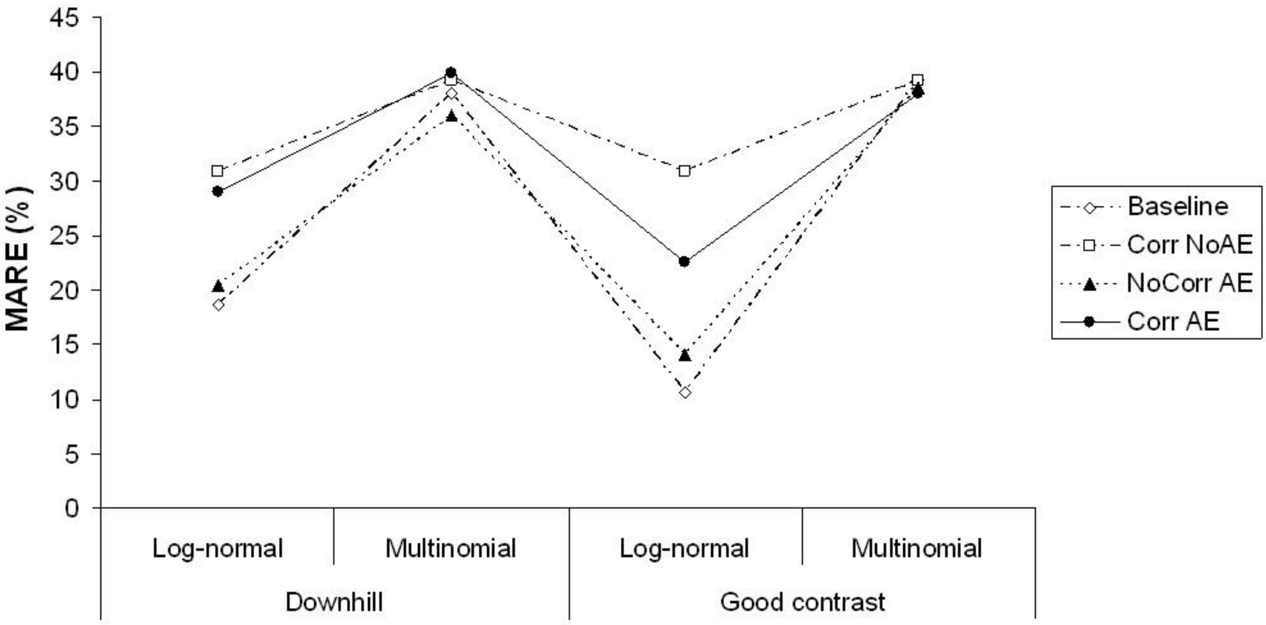


Figure 3.17: DEPLETION MAREs for Experiment 2.See Figure 3.14 for an explanation of scenarios

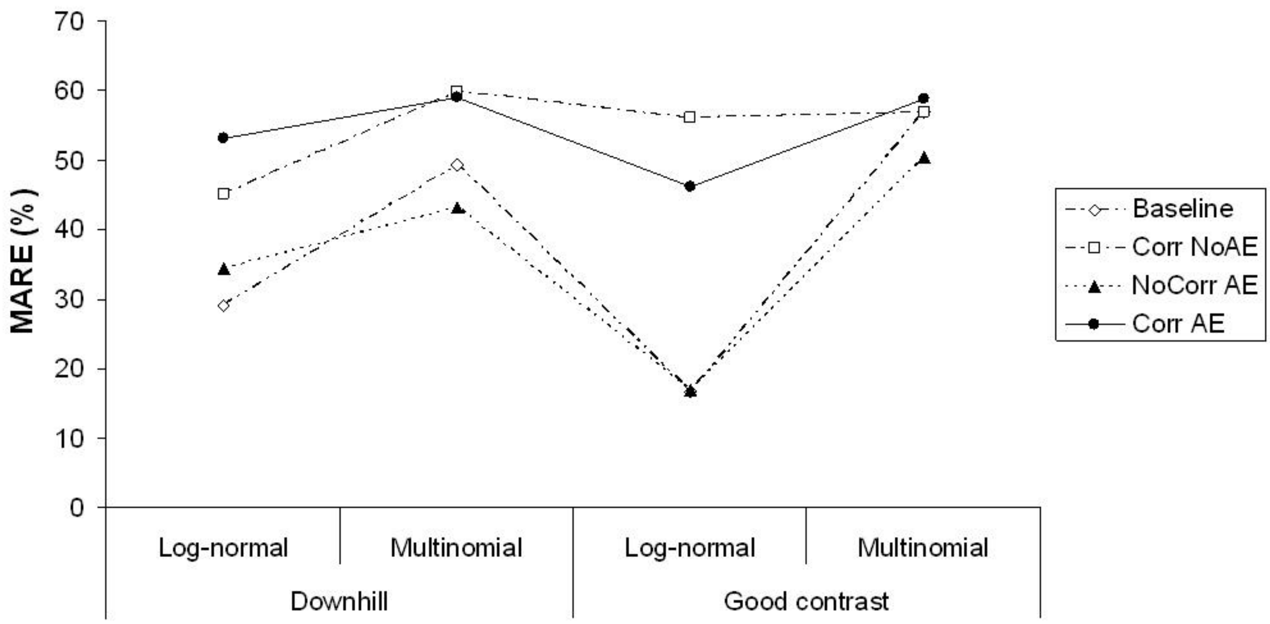


Figure 3.18: TAC MAREs for Experiment 2. See Figure 3.14 for an explanation of scenarios

### 3.3.3 Interactions

GLMs were applied to each management quantity separately. For MSY, significant differences were found between harvest rate effects (F-statistic=7.3; df=(1,1448); p=0.007) and error effects (F-statistic=41.7; df=(3,1448); p<0.001). Autocorrelated recruitment had the most effect on increasing the absolute values of MSY (Scheffé p<0.001). Both likelihoods yielded higher absolute values for *Downhill trip* data (Scheffé p=0.007). This was the same case for MSYL.

For MSYR, there were significant differences between effects of harvest rates (F-statistic=13.9; df=(1,1448); p<0.001), likelihood functions (F-statistic=9.5; df=(1,1448); p=0.002), error scenarios (F-statistic=43.4; df=(3,1448); p<0.001) and the interaction of the three factors (F-statistic=2.9, df=(3,1448), p<0.035). When the log-normal likelihood was used, the model predicted MSYR significantly better than when the multinomial likelihood was used (Scheffé p=0.003). Both likelihoods yielded significantly lower absolute values for *Good contrast* data (Scheffé p<0.001).

For DEPLETION, there were significant differences between the likelihoods (F-statistic=89.1; df=(1,1381); p<0.001), the error scenarios (F-statistic=11.0; df=(3,1381); p<0.001), and the interaction of these two (F-statistic=5.7; df=(3,1381); p<0.001). The log-normal likelihood yielded significantly lower absolute values than the multinomial likelihood (Scheffé p<0.001).

Lastly, the TAC only had significant differences between the likelihoods (F-statistic 64.6; df=(1,1291); p<0.001) and the error scenario (F-statistic=27.0; df=(3,1291); p<0.001). It appears the multinomial likelihood always had higher absolute values (Scheffé p<0.001).

### 3.3.4 Failures and extreme outliers

Nine percent of the simulations failed, two thirds of which were due to the log-normal likelihood. Most of them occurred when there were ageing errors. Results can be seen in Figure 3.19.

There were only 544 extreme outliers, which equated to 2.2% of all absolute error estimates for all parameters, with 56% of them were from the log-normal likelihood and 44% from the multinomial likelihood (Figure 3.20). More in-depth analysis revealed age-zero selectivity for research trawls had the highest probability of being an extreme outlier when the log-normal distribution was used. TAC had the highest probability of being an extreme outlier when the multinomial likelihood function was used.

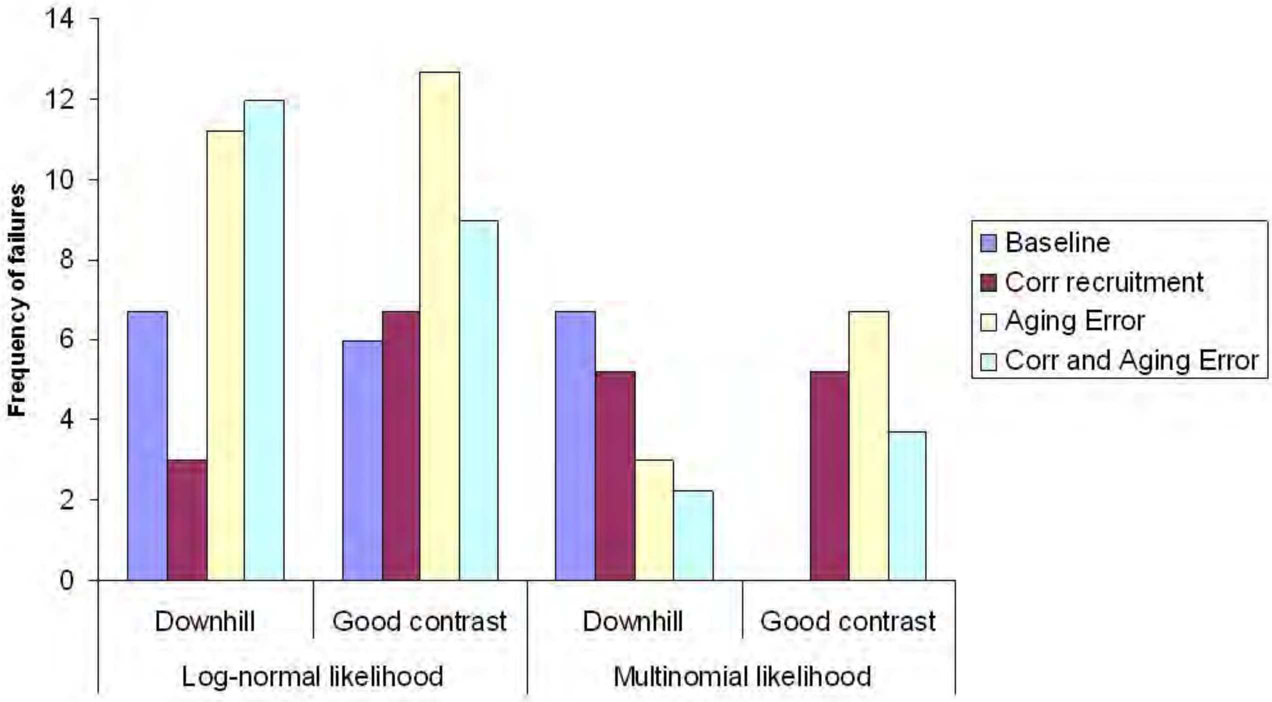


Figure 3.19: Frequency of failures for when the log-normal likelihood and multinomial likelihood were used in Experiment 2 as a percentage of the total amount. Harvest scenarios were *Downhill trip* and *Good contrast* data. The error scenarios included *Baseline*, which had no autocorrelation in recruitment in the stock-recruitment relationship and no ageing error, *Corr NoAE* which had autocorrelation in the stock-recruitment relationship and no ageing error, *NoCorr AE* which had no autocorrelation in the stock-recruitment relationship and multinomial ageing error in the CAA matrix and *Corr AE* which had both correlation in the stock-recruitment relationship and multinomial ageing error.

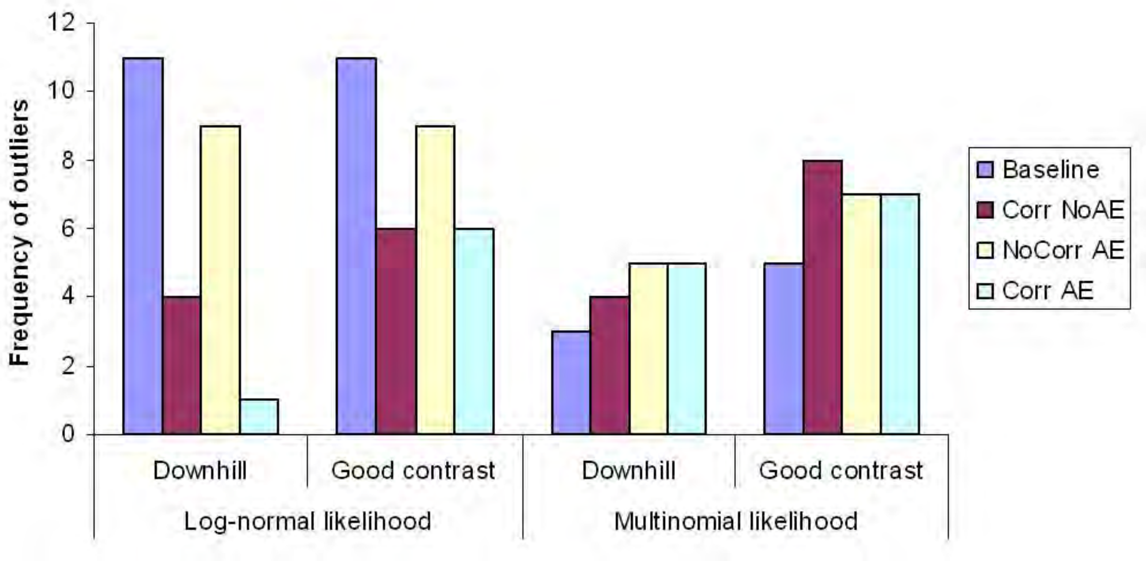


Figure 3.20: Frequency of outliers for when the log-normal likelihood and multinomial likelihood were used in Experiment 2 as a percentage of the total amount. See Figure 3.19 for an explanation of scenarios.

## 3.4 Experiment 3 - Otolith sample size

It appears that coefficients of variation (CV) of the management quantities when the log-likelihood function is used is relatively constant after 2000 otoliths are aged. Use of the multinomial likelihood shows a logarithmic relationship between the CV and the number of otoliths aged.

When the log-normal likelihood was used in the CAA data, coefficients of variation for the five management quantities dropped substantially at a sample size of 2000, and remained at this relatively constant level for higher sample sizes (Figure 3.21).

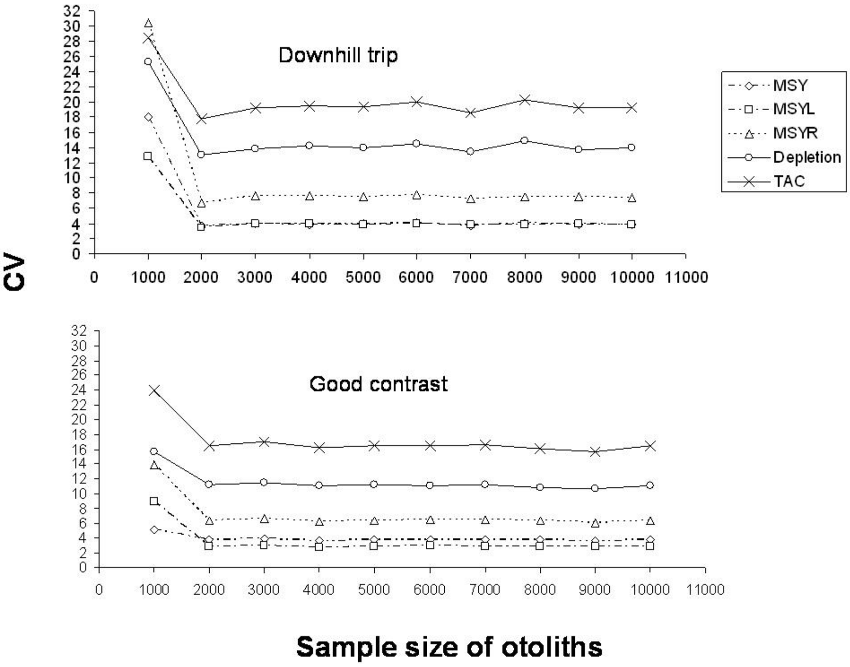


Figure 3.21: Coefficients of variation (CV) obtained from ageing various sample sizes of otoliths for the log-normal likelihood for the two harvest rates. Five management quantities were calculated.

The CVs for the management quantities using a multinomial log-likelihood followed distinctive negative logarithmic shapes (Figure 3.22). Equations are in Table 3.1 with $x$ as the number of otoliths aged which interpolates the CV trend for each management parameter. The percentage returns of an extra 1000 otoliths aged diminish logarithmically, suggesting further increases in input lead to proportionately smaller variance. Percentage decrease of CV on an extra 1000 otoliths aged were calculated as

$$Percentage\ decrease = \frac{CV(x-1000) - CV(x)}{CV(x-1000)} \quad .$$

The calculated percentage decrease of an extra 1000 otoliths aged are presented in Tables 3.2 and 3.3 for *Downhill trip* and *Good contrast* data respectively. For example, if a value of -20% is obtained at a sample size of $x$ otoliths, it means a 20% decrease in CV will result if a thousand more otoliths are aged to reach $x$.

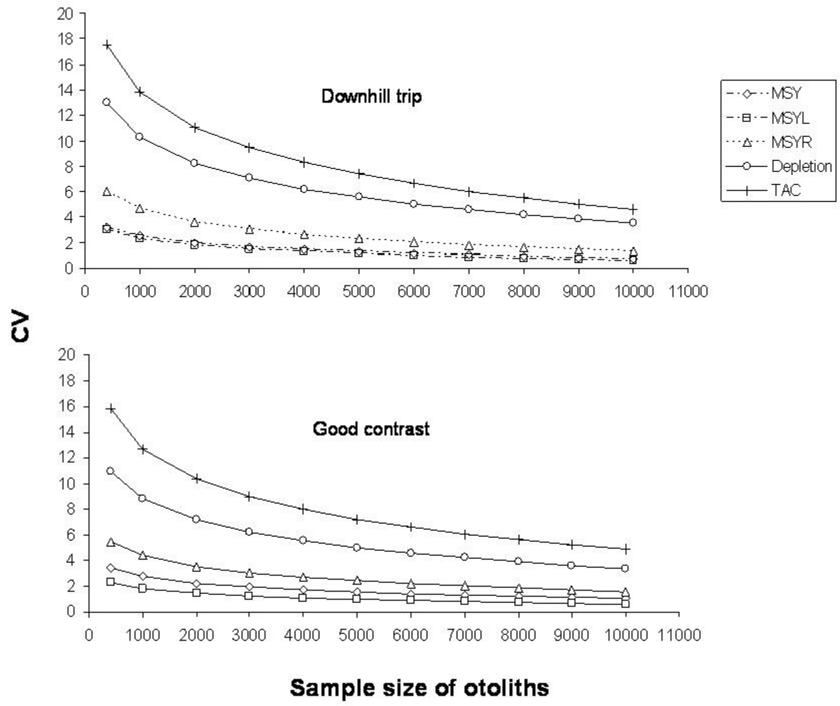


Figure 3.22: Coefficient of variation (CV) obtained from ageing various sample sizes of otoliths for using the multinomial likelihood to tune the catch-at-age matrix for two harvest rates. Five management quantities were calculated.

Table 3.1: Equations describing the coefficient of variation (CV) values for the different management quantities when $x$ is the number of otoliths aged.

| | *Downhill trip* harvest rate | | *Good contrast* harvest rate | |
|---|---|---|---|---|
| Parameter | Equation | $R^2$ | Equation | $R^2$ |
| MSY | -$0.7695ln(x) + 7.8434$ | 0.97 | $-0.7393ln(x) + 7.8429$ | 0.99 |
| MSYL | $-0.7491ln(x) + 7.4941$ | 0.96 | $-0.5326ln(x) + 5.4915$ | 0.99 |
| MSYR | $-1.4551ln(x) + 14.701$ | 0.97 | $-1.2224ln(x) + 12.833$ | 0.99 |
| DEPLETION | $-2.9331ln(x) + 30.539$ | 0.99 | $-2.3466ln(x) + 24.997$ | 0.99 |
| TAC | $-4.0155ln(x) + 41.583$ | 0.99 | $-3.3955ln(x) + 36.141$ | 0.99 |

Table 3.2: Percentage decrease in the coefficient of variation for five management quantities for an extra 1000 otoliths aged for *Downhill trip* data when the multinomial distribution was used for the catch-at-age data.

| Sample size of otoliths | MSY % decrease | MSYL % decrease | MSYR % decrease | DEPLETION % decrease | TAC % decrease |
|---|---|---|---|---|---|
| 2000 | -21.1 | -22.4 | -21.7 | -19.8 | -20.1 |
| 3000 | -15.6 | -16.9 | -16.2 | -14.4 | -14.7 |
| 4000 | -13.2 | -14.4 | -13.7 | -12.0 | -12.2 |
| 5000 | -11.8 | -13.0 | -12.3 | -10.5 | -10.8 |
| 6000 | -10.9 | -12.3 | -11.5 | -9.6 | -9.9 |
| 7000 | -10.3 | -11.8 | -11.0 | -9.0 | -9.3 |
| 8000 | -10.0 | -11.6 | -10.7 | -8.6 | -8.9 |
| 9000 | -9.8 | -11.6 | -10.5 | -8.3 | -8.6 |
| 10000 | -9.7 | -11.7 | -10.5 | -8.1 | -8.4 |

Table 3.3: Percentage decrease in the coefficient of variation for five management quantities for an extra 1000 otoliths aged for *Good contrast* data when the multinomial distribution was used for the catch-at-age data.

| Sample size of otoliths | MSY % decrease | MSYL % decrease | MSYR % decrease | DEPLETION % decrease | TAC % decrease |
|---|---|---|---|---|---|
| 2000 | -18.7 | -20.4 | -19.3 | -18.5 | -18.6 |
| 3000 | -13.5 | -15.0 | -14.0 | -13.3 | -13.3 |
| 4000 | -11.1 | -12.5 | -11.5 | -10.9 | -10.9 |
| 5000 | -9.6 | -11.1 | -10.1 | -9.5 | -9.5 |
| 6000 | -8.7 | -10.2 | -9.2 | -8.5 | -8.6 |
| 7000 | -8.1 | -9.6 | -8.6 | -7.9 | -7.9 |
| 8000 | -7.6 | -9.2 | -8.1 | -7.4 | -7.5 |
| 9000 | -7.3 | -8.9 | -7.8 | -7.1 | -7.1 |
| 10000 | -7.0 | -8.7 | -7.6 | -6.8 | -6.8 |

# Chapter 4

# Discussion

The use of simulations, particularly the use of an operating model - assessment model approach, has been used extensively in fisheries modelling as it allows scientists to test and quantify model precision and bias under different conditions. The result is a better understanding of which assessment models, and assumptions, yield the most accurate results for a particular set of conditions. For example, simulations can be used to select the most accurate management quantities for a particular set of data. Also, simulations can be used to determine which data sources are most informative (Magnusson, 2004), thus limited resources can be used to the best effect. In the context of the South African hake resource, simulations have been used extensively to quantify the risk of harvesting the stock unsustainably given different harvesting strategies (Rademeyer, 2003). An example is Punt (1992) who found in order to minimise risk the best strategy was to use maximum allowable catch. Previous studies have investigated the effects of various data sources, harvest rates, assumptions and error scenarios in isolation. Combinations of process and observation error were investigated on Namibian hake by Punt (2003). Punt (1991) investigated various management procedures for a large number of scenarios. However, few investigations have quantified the effect of various combinations of these data sources, error scenarios and assumptions on model accuracy that were investigated in this study .

The aims of this study were to ascertain the effect of different combinations of data sources, ageing error in the CAA matrix and process error in the stock-recruitment relationship. The study also compared the Schaefer model with the ASPM, the use of two different likelihood functions for the catch-at-age data, and determined the optimal amount of otoliths that needed to be aged for input data.

The ASPM generally derived the management quantities better than the Schaefer model for the *Baseline* error scenarios and when there was ageing error. Bearing in mind that the management quantities were generated with an ASPM model, the Schaefer model performed well, thus the model should not be discarded due to it's simplicity. The ASPM derived MSY, MSYL and MSYR more accurately than the Schaefer model. The Schaefer model typically overestimated these management quantities. Therefore, it is recommended that the Schaefer model is not used to estimate these quantities, as it may result in management decisions that would suggest unsustainably high catches. The management quantities from the ASPM, on the other hand, were generally less variable, more accurate and were conservative. Therefore the ASPM model should be used to estimate them. By contrast, the two management quantities, DEPLETION and TAC, were generally better estimated by the Schaefer model. DEPLETION and TAC alone could possibly be used as a basis for management strategies. Thus the Schaefer model, which derives these management quantities accurately and requires less expensive data than that of the ASPM, could be used as a basis for management strategies when resources are limited. However, when an entire suite of management quantities are needed, a combination of the Schaefer and the ASPM models will give the best results.

The Schaefer model tended to obtain higher absolute error values with *Downhill trip* data. Prager et al. (1996) used an extension of the Schaefer model, in which selectivity varied over time, and found that *Downhill trip* data did not allow the objective function to find a unique minimum. In the current study, the Schaefer model was able to estimate values when given *Downhill trip* data, albeit poorly. Therefore it appears the Schaefer model may be used with *Downhill trip* data under certain conditions, though it is not as accurate as the ASPM model. Hilborn (1979), Ludwig and Walters (1985) and Polacheck et al. (1993) showed there must be sufficient contrast in the abundance index in order to model a resource accurately. Therefore, in order to estimate productivity the resource should be fished to below MSYL and then be allowed to rebuild. Thus models tuned with *Good contrast* data will estimate management quantities better than those tuned with *Downhill trip* data. Likewise, this study found the *Downhill trip* harvest data did not, in general, estimate management quantities as well as models tuned with *Good contrast* data. However, the trials with *Downhill trip* data did perform better than were prescribed by Hilborn (1979), Ludwig and Walters (1985) and Polacheck et al. (1993), and did not result in substantially more failures than that of the *Good contrast* data. This finding was similar to that of Magnusson and Hilborn (2007).

When all data sources were used in the ASPM model, estimation of the management quantities were generally the best. When only the six years of survey data were used, the model performed poorly. Thus it should not be the only source of data used in the ASPM model. However, Walters and Maguire (1996) surmised that stock size is overestimated when only CPUE is used due to increasing catchability over time, and that more investment should be placed in survey trawls regardless of the model used. There is a trade-off between model accuracy and increased cost associated with the collection of data from more sources. This will be particular to different situations. The point at which the cost of collecting more data is no longer justified by a gain in model accuracy needs to be determined in future studies.

The study found there was a high probability of obtaining extreme outliers when estimating age-zero selectivity for research trawls. Thus it appears that this parameter should possibly not be estimated, and rather given a constant value.

Autocorrelation in the stock-recruitment relationship is one of the main uncertainties surrounding stock assessment (Clark, 1993; Francis and Shutton, 1997). Models also appear to be very sensitive to it (Butterworth and Punt, 1999). Ianelli (2002) found that CVs are considerably higher when autocorrelation is not accounted for in the ASPM. This study found a similar result using MAREs as a performance measure. Punt (1997) showed autocorrelation led to overly optimistic representations of the system with an age-structured Virtual Population Analysis model. Similar to Punt's (1997) results, autocorrelation in the stock-recruitment relationship caused the ASPM model to give an optimistic representation of the stock. Thus when there is autocorrelation in the stock-recruitment relationship there is a danger of the stock being overfished. A possible solution is applying time-series techniques with lags to correct for the autocorrelation. In the current study, autocorrelation also increased the variance of relative error distributions quite substantially. Clark (1993) obtained similar results in his study on the effects of autocorrelation on groundfish, although his measure was CV and not relative error.

Multinomial ageing error had a small increasing effect on the MAREs when log-normal errors were assumed in the assessment model. However, the addition of multinomial ageing error was not significantly different to when there was no ageing error. This may have been from the relatively large (5000) sample size of otoliths which could have been fully representative of the population structure. There have been various studies on the effects of ageing error, though very few concentrate on how it affects management quantities. Reeves (2003) investigated how differing bias in age-reading across nations can effect scientific advice. He found that noise-only error, such as what is investigated in this study, consistently underestimates spawner stock biomass, though as found in this study, the differences were not large.

Use of the log-normal likelihood in the catch-at-age data generally showed lower levels of MARE for DEPLETION and TAC, compared to the use of multinomial likelihood, as well as smaller variation in their relative error distributions. For the other management quantities the log-normal likelihood produced similar results to the multinomial. The ageing error that was added to the data drawn in the operating model was multinomial. Thus the log-normal likelihood would appear to be robust, even when the true underlying error is not log-normal. It is thus recommended that log-normal errors should be assumed in the catch-at-age data to gain accuracy. This finding is supported by Kimura (1990) who tested the effects of ageing error on age-structured separable sequential population analysis, and Deriso et al. (1985) in his investigation of assumptions that stabilised estimates for the Pacific halibut (*Hippoglossus stenolepis*).

Any increases in sample size after ageing a sample size of 2000 otoliths when using the log-normal likelihood function in the CAA data did not have an effect on the CVs of management quantities. In comparison, use of the multinomial likelihood yielded CVs for management quantities that decrease logarithmically with sample size between the range of 1000 and 10000. Therefore in the current study, when using the multinomial likelihood, the percentage decrease in CV of ageing an extra thousand otoliths otoliths are diminished due to the logarithmic shape of the CV to sample size relationship. Furthermore, use of the multinomial likelihood resulted in lower CVs than the log-normal likelihood function for both harvest rates scenarios and at any sample size between 1000 and 10000. Thus it is recommended that if precision in the data is to be maximised, the multinomial likelihood should be used. This is similar to the study by Crone and Sampson (1998) who found the multinomial distribution was best when describing error in samples of landed data for five groundfish species in Oregon. There have been many simulation studies involving sample size (Magnusson, 2004). For example, Sullivan et al. (1993) found the optimal sample size for otoliths when using an age-structured production model for Blue Grenadier (*Macruronus novaezelandiae*) was between 1000-1500. These kinds of information are used to improve stock assessment models. Richards et al. (1992) stated that in order to obtain more precision, the sample size for ageing information must be increased. While this study suggested the same conclusion when the multinomial likelihood is used, the conclusion was only valid up to a certain point (i.e. between 1000 and 2000 otoliths) when using the log-normal likelihood.

It is recommended that stock assessment models should be chosen not only based on the available data, but also on the management quantity of interest. The log-normal likelihood function should be used to obtain accuracy in the estimates, while to minimise variation in the management quantity estimates, the multinomial likelihood should be used. However, if TAC is the quantity of interest, the multinomial likelihood should not be used as there is a higher probability of obtaining extreme outliers. The current model used to assess hake in South Africa is the Fox-form of the biomass dynamic model (Auckland and Ninnes, 2004). Reasons for using this model are the decreased computation time and simplicity compared to the ASPM. If management continues to use this model to estimate management quantities, then based on the results of this study, DEPLETION and TAC (as defined in this study) should be derived using this model. More investigation into correcting for autocorrelation in the stock-recruitment relationship should be carried out, as it has a large effect on the accuracy of management quantities.

## 4.1 Limitations

The "true" underlying parameters were given to the ASPM when it could not estimate these with the data sources it had. This means caution is needed when making conclusions by directly comparing how the different data sources affect the MAREs of the ASPM. This is because more information was given to the stock assessment model than what would normally be known to it when "real" data is used, which is unrealistic (Wang et al., 2005). Thus sensitivity analysis should be carried out to test model reaction to varying input parameters, different from that of the operating model.

The GLM was used to investigate interactions in the data. Transformation of the absolute errors stabilised the variances of the distributions substantially, though the Levene's test would sporadically show unequal absolute error variances. Thus the assumption of homogeneity of variances was sometimes violated. However, Quinn and Keough (2002) state that with large sample sizes, such as the ones used in this study, the tests of homogeneity of variances are frequently rejected but the ANOVA F-test is still reliable. Therefore because of the large sample sizes used in the current study it was assumed that the GLM would still provide accurate analysis.

# References


Auckland, R. and C. Ninnes. 2004. An assessment of the state of commercial fisheries catch data in the BCLME region. Technical report, BCLME. Project LMR/CF/03/02.

Balasubramanian, V. 1997. Statistical inference, Occam's razor, and statistical mechanics on the space of probability distributions. Neural Computations, **9**:349–368.

Banks, J., J. S. Carson, B. L. Nelson, and D. M. Nicol. 2000. Discrete-event system simulation. Prentice Hall, New Jersey.

Beddington, J. R., D. J. Agnew, and C. W. Clark. 2007. Current problems in the management of marine fisheries. Science, **316**:1713–1716.

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

Booth, A. J. and T. J. Quinn. 2006. Maximum likelihood and Bayesian approaches to stock assessment when data are questionable. Fisheries Research, **80**:169–181.

Bradford, M. J. 1991. Effects of ageing errors on recruitment time series estimated from sequential population analysis. Canadian Journal of Fisheries and Aquatic Sciences, **48**:555–558.

Butterworth, D. S. and A. E. Punt. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, **56**:985–998.

Clark, W. G. 1993. The effect of recruitment variability on the choice of a target level of spawning biomass per recruit. In G. Kruse, D. M. Eggers, R. Marasco, C. Pautze, and T. J. Quinn II, editors, Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations, pp 233–246. Alaska Sea Grant College Program, University of Alaska, Fairbanks.

Coggins, L. G. and T. J. Quinn. 1998. A simulation study of the effects of aging error and sample size on sustained yield estimates. In F. Funk, T. J. Quinn II, J. Heifetz, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C. I. Zhang, editors, Proceedings of the International Symposium on Fishery Stock Assessment Models, pp 955–975. Alaska Sea Grant College Program, University of Alaska, Fairbanks.

Committee on Fish Stock Assessment Methods. 1998. Improving fish stock assessments. National Academy Press, Washington.

Crone, P. R. and D. B. Sampson. 1998. Evaluation of assumed error structure in stock assessment models that use sample estimates of age composition. In F. Funk, T. J. Quinn II, J. Heifetz, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C. I. Zhang, editors, Proceedings of the International Symposium on Fishery Stock Assessment Models, pp 355–370. Alaska Sea Grant College Program, University of Alaska, Fairbanks.

Deriso, R. B., T. J. Quinn II, and P. R. Neal. 1985. Catch-age analysis with auxiliary information. Canadian Journal of Fisheries and Aquatic Sciences, **42**:815–824.

FAO. 2006. The state of world fisheries and aquaculture. Technical report, FAO Corporate Document Repository, Rome.

Fournier, D. and C. P. Archibald. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences, **39**:1195–1207.

Fox, W. W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Transactions of the American Fisheries Society, **99**:80–88.

Francis, R. I. C. C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Canadian Journal of Fisheries and Aquatic Science, **49**:922–930.

Francis, R. I. C. C. and R. Shutton. 1997. "Risk" in fisheries management: a review. Canadian Journal of Fisheries and Aquatic Sciences, **54**:1699–1715.

Fryer, D. 2003. Microeconomics and society: towards a synthesis of orthodox and heterodox explanations. Ph.D. thesis, Rhodes University, Grahamstown.

Glazer, J. P. and D. S. Butterworth. 2004. The general linear models applied to standardize the catch per unit effort data of the Cape hake stocks, *Merluccius capensis* and *M. paradoxus*, caught offshore off the coast of South Africa. Technical report, BENEFIT/NRF/BCLME.

Grafton, R. Q., J. Kirkley, T. Kompas, and D. Squires. 2006. Economics and fisheries management. Ashgate Publishing Company, Aldershot.

Haddon, M. 2001. Modelling and quantitative methods in fisheries. Chapman and Hall/CRC, Boca Raton.

Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. Journal of the Fisheries Research Board of Canada, **36**:1477–1489.

Hilborn, R. 1992. Current and future trends in fisheries stock assessment and management. South African Journal of Marine Science, **12**:975–988.

Hilborn, R. 2003. The state of the art in stock assessment: where we are and where we are going. Scientia Marina, **67**:15–20.

Ianelli, J. N. 2002. Simulation analyses testing the robustness of productivity determinations from west coast Pacific ocean perch stock assessment data. North American Journal of Fisheries Management, **22**:301–310.

Kell, L. T., C. M. O'Brien, M. T. Smith, T. K. Stokes, and B. D. Rackham. 1999. An evaluation of management procedures for implementing a precautionary approach in the ICES context for North Sea plaice (*Pleuronected platessa* L.). ICES Journal of Marine Science, **56**:834–845.

Kell, L. T., G. M. Pilling, G. P. Kirkwood, M. Pastoors, B. Mesnil, K. Korsbrekke, P. Abaunza, R. Aps, A. Biseau, P. Kunzlik, C. Needle, B. A. Roel, and C. Ulrich-Rescan. 2005. An evaluation of the implicit management procedure used for some ICES groundfish stock. ICES Journal of Marine Science, **62**:750–759.

Kimura, D. K. 1990. Approaches to age-structured separable sequential population analysis. Canadian Journal of Fisheries and Aquatic Sciences, **47**:2364–2374.

Ludwig, D. and C. J. Walters. 1985. Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences, **42**:1066–1072.

Mace, P. M., N. W. Bartoo, A. B. Hollowed, P. Kleiber, R. D. Methot, S. A. Murawski, J. E. Powers, and G. P. Scott. 2001. Marine fisheries stock assessment improvement plan. Technical report, National Oceanic and Atmospheric Administration. NOAA Technical Memorandum NMFS-F/SPO-56.

Magnusson, A. 2004. Information about stock abundance: data, models, and assumptions. Ph.D dissertation proposal.

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

Maunder, M. N. 2001. A general framework for integrating the standardization of catch-per-unit-effort into stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences, **58**:795–803.

Maunder, M. N. and A. D. Langley. 2004. Integrating the standardization of catch-per-unit-of-effort into stock assessment models: testing a population dynamics model and using multiple data types. Fisheries Research, **70**:389–395.

McCullagh, P. and J. A. Nelder. 1992. Generalized linear models. Chapman and Hall, London.

Miller, T. J. and J. R. Skalski. 2006. Integrating design- and model-based inference to estimate length and age composition in North Pacific longline catches. Canadian Journal of Fisheries and Aquatic Sciences, **63**:1092–1114.

Myers, R. A., J. A. Hutchings, and N. J. Barrowman. 1997. Why do fish stocks collapse? The example of cod in Atlantic Canada. Ecological Applications, **7**:91–106.

Otter Research Ltd. 1994. An introduction to AD model builder. Otter Research Ltd: Nanaimo, Canada.

Patterson, K., R. Cook, C. Darby, S. Gavaris, L. Kell, P. Lewy, B. Mesnil, A. Punt, V. Restrepo, D. W. Skagen, and G. Stefansson. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries, **2**:125–157.

Pella, J. J. and P. K. Tomlinson. 1969. A generalized stock-production model. Bulletin of the Inter-American Tropical Tuna Commission, **13**:421–458.

Phillips, B., T. Ward, and C. Chaffee. 2003. Eco-labelling in fisheries. Blackwell, Oxford.

Polacheck, T., R. Hilborn, and A. E. Punt. 1993. Fitting surplus production models: comparing methods and measuring uncertainty. Canadian Journal of Fisheries and Aquatic Sciences, **50**:2597–2607.

Polet, H. and F. Redant. 1998. Effect of population structure, sampling strategy and sample size on the estimates of selection parameters for shrimp *(Crangon crangon)* trawls. Fisheries Research, **40**:213–225.

Prager, M. H., C. P. Goodyear, and G. P. Scott. 1996. Application of a surplus production model to a swordfish-like simulated stock with time-changing gear selectivity. Transactions of the American Fisheries Society, **125**:729–740.

Punt, A. E. 1991. Management procedures for Cape hake and baleen whale resources. Ph.D. thesis, University of Cape Town.

Punt, A. E. 1992. Selecting management methodologies for marine resources, with an illustration for southern African hake. South African Journal of Marine Science, **12**:943–958.

Punt, A. E. 1997. The performance of VPA-based management. Fisheries Research, **29**:217–243.

Punt, A. E. 2003. Extending production models to include process error in the population dynamics. Canadian Journal of Fisheries and Aquatic Sciences, **60**:1217–1228.

Punt, A. E. and R. W. Leslie. 1991. Estimates of some biological parameters for the Cape hakes off the South African west coast. South African Journal of Marine Science, **10**:271–284.

Quinn, G. P. and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge.

Quinn, T. J. and R. B. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, Oxford.

Rademeyer, R. A. 2003. Assessment of and management procedures for the hake stocks off Southern Africa. Master's thesis, University of Cape Town.

Reeves, S. A. 2003. A simulation study of the implications of age-reading errors for stock assessment and management advice. ICES Journal of Marine Science, **60**:314–328.

Richards, L. J., J. T. Schnute, A. R. Kronlund, and R. J. Beamish. 1992. Statistical models for the analysis of ageing error. Canadian Journal of Fisheries and Aquatic Sciences, **49**:1801–1815.

Rose, K. A. and J. H. Cowan. 2003. Data, models and decision in US marine fisheries management: lessons for ecologists. Annual Review of Ecology, Evolution, and Systematics, **34**:127–51.

Schaefer, M. B. 1957. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-American Tropical Tuna Commission, **1**:247–285.

Schnute, J. T. and R. Hilborn. 1993. Analysis of contradictory data sources in fish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, **50**:1916–1137.

Schweder, T. 1998. Fisherian or Bayesian methods of integrating diverse statistical information? Fisheries Research, **37**:61–75.

Shepherd, J. G. 1984. The availability and information content of fisheries data. In R. M. May, editor, Exploitation of marine communities, pp 95–109. Dahlem workshop 1984, Springer-Verlag, Berlin.

Smith, P. J. 1989. Is two-phase sampling really better for estimating age-composition? American Statistical Association, **84**:916–921.

Sullivan, K., D. C. Smith, and A. Gason. 1993. Getting the right information for age-structured models. In D. A. Hancock, editor, Population Dynamics for Fisheries Management, pp 91–98. Australian Society For Fish Biology, Perth.

Walker, T. I. 1992. Fishery simulation model for sharks applied to the gummy shark, *Mustelus antarcticus* Gunther, from southern Australian waters. Australian Journal of Marine and Freshwater Resources, **43**:195–212.

Walters, C. and J. Maguire. 1996. Lessons for stock assessment from the northern cod collapse. Reviews in Fish Biology and Fisheries, **6**:125–137.

Wang, S., C. Sun, A. E. Punt, and S. Yeh. 2005. Evaluation of a sex-specific age-structured assessment model for the swordfish, *Xiphias gladius*, in the North Pacific Ocean. Fisheries Research, **73**:79–97.

Xiao, Y. 1997. Subtleties in, and practical problems with, the use of production models in fish stock assessment. Fisheries Research, **33**:17–36.

# Appendix A

# Acronyms and definitions

**ASPM** Age Structured Production Model

**BDM** Biomass Dynamic Model

**CAA** Catch-at-age

**CPUE** Catch Per Unit Effort

**CV** Coefficient of variation

**GLM** General Linear Model

**ICSEAF** International Commission for South East Atlantic Fisheries

**Large recruitment events** A large number of fish from a specific spawning events survive, such that there is a pulse in a certain age-group biomass

**MARE** Median Absolute Relative Error

**MSY** Maximum Sustainable Yield

**Recruits** The amount of zero-year old fish born into the fishery in a particular year

**TAC** Total Allowable Catch

# Appendix B

# AD model builder code

## B.1 Age-structured production model

```
// ASPM model for hake off the West Coast
// Amanda Northrop 2007
DATA_SECTION
// taken from .dat file
init_int minage;
init_int maxage;
init_int comm_minage;
init_int comm_maxage;
init_int survey_minage;
init_int survey_maxage;
init_int yearbegin;
init_int yearend;
init_int yearbegincomm;
init_int yearendcomm;
init_int yearbeginsurvey;
init_int yearendsurvey;
init_vector Catch(yearbegin,yearend);
init_vector CPUE(yearbegincomm,yearendcomm);
```

```
init_vector survey(yearbeginsurvey,yearendsurvey);
init_matrix
comm_CAA(yearbegincomm,yearendcomm,comm_minage,comm_maxage);
init_matrix survey_CAA(yearbeginsurvey,yearendsurvey,survey_minage,survey_maxage);
init_vector weight(minage,maxage);
init_vector weightplushalf(minage,maxage);
init_vector mortality(minage,maxage);
init_vector maturity(minage,maxage);
init_vector comms(minage,maxage);
init_vector surveys(minage,maxage);
init_int phaseK;
init_int phaseh;
init_int phaseM;
init_int phasealphamort;
init_int phasebetamort;
init_int phaseacf;
init_int phasedelta;
init_int phaseage0sel;
init_int phaseage1sel;
init_number maxmu;
init_number parameters;
init_vector useindex(1,4);
//________________________________________________________
PARAMETER_SECTION
// created bounds for two estimable parameters
init_bounded_number Ksp(500,10000,phaseK);
init_bounded_number h(0.2,1,phaseh);
init_bounded_number acf(0.01,4,phaseacf);
init_bounded_number delta(0.01,5,phasedelta);
init_bounded_number age0sel(0.0001,1,phaseage0sel);
init_bounded_number age1sel(0.0001,1,phaseage1sel);
```

```
//estimable parameters
vector commsel(minage,maxage);
vector surveysel(minage,maxage);
// start-off parameters for model
vector Ntilda(minage,maxage);
number SBPR;
number EBPR;
number alpha;
number beta;
number Rzero;
vector Bsp(yearbegin,yearend+1);
vector Bexplcomm(yearbegin,yearend+1);
vector Bexplsurvey(yearbegin,yearend+1);
vector mu(yearbegin,yearend+1);
matrix numbersmatrix(yearbegin,yearend+1,minage,maxage);
matrix catchmatrix(yearbegin,yearend+1,minage,maxage);
number ncomm;
number nsurvey;
// CPUE index parameters
vector Bexplc(yearbegincomm,yearendcomm);
vector qcommvector(yearbegincomm,yearendcomm);
vector qBcomm(yearbegincomm,yearendcomm);
vector residualscomm(yearbegincomm,yearendcomm);
number qcomm;
number sigmacomm;
// survey index parameters
vector Bexpls(yearbeginsurvey,yearendsurvey);
vector qsurveyvector(yearbeginsurvey,yearendsurvey);
vector qBsurvey(yearbeginsurvey,yearendsurvey);
vector residualssurvey(yearbeginsurvey,yearendsurvey);
number qsurvey;
```

```
number sigmasurvey;
// commercial CAA
matrix catchm(yearbegincomm,yearendcomm,comm_minage,comm_maxage);
matrix predpropcomm(yearbegincomm,yearendcomm,comm_minage,comm_maxage);
vector sumagecomm(yearbegincomm,yearendcomm);
matrix sigmacommCAAmatrix(yearbegincomm,yearendcomm,comm_minage,comm_maxage);
number sigmacommmatrixno;
number sigmacommCAA;
matrix LogLcommCAA(yearbegincomm,yearendcomm,comm_minage,comm_maxage);
// survey CAA
matrix predcatchessurvey(yearbeginsurvey,yearendsurvey,survey_minage,survey_maxage);
matrix predpropsurvey(yearbeginsurvey,yearendsurvey,survey_minage,survey_maxage);
vector sumagesurvey(yearbeginsurvey,yearendsurvey);
matrix sigmasurveyCAAmatrix(yearbeginsurvey,yearendsurvey,survey_minage,survey_maxage);
number sigmasurveymatrixno;
number sigmasurveyCAA;
matrix LogLsurveyCAA(yearbeginsurvey,yearendsurvey,survey_minage,survey_maxage);
//MSY calculations
vector nos(minage,maxage);
number Fmax;
number MSY;
number MSYL;
number MSYR;
number TAC;
// declarations needed to obtain standard deviations of reference points
sdreport_number MSYrep;
sdreport_number MSYLrep;
sdreport_number MSYRrep;
sdreport_number depletion;
sdreport_number qcommrep;
sdreport_number qsurveyrep;
```

```
sdreport_number bendplus1;
sdreport_number bbeginning;
sdreport_number TACrep;
// values for objective functions
number penalty;
vector LogL(1,4);
objective_function_value f;
//___________________________________________________
PROCEDURE_SECTION
calculate selectivity parameters
calculate_estimable_parameters();
calculate_kick_off();
calculate_numbers();
LogL = 0;
if (useindex(1) == 1)
calculate_comm(); //use commercial indices to obtain likelihood
if (useindex(2) == 1)
calculate_survey(); // ditto for survey biomass indices
if (useindex(3) == 1)
calculate_comm_CAA(); // commercial CAA information
if (useindex(4) == 1)
calculate_survey_CAA(); // survey CAA information
f += sum(LogL);
// obtain likelihood profiles for some of the parameters
if (last_phase())
{
calculate_likelihood_profiles();
//calculate reference points
Fmax = calculate_muMSY();
MSY = yield(Fmax);
MSYL = exploitablebiomass(Fmax);
```

```
MSYR = MSY/MSYL;
TAC = 0.9*MSYR*Bexplcomm(yearend+1);
}
if (sd_phase())
{
MSYrep = MSY;
MSYLrep = MSYL;
MSYRrep = MSYR;
depletion = Bexplcomm(yearend+1)/Bexplcomm(yearbegin);
qcommrep = qcomm;
qsurveyrep = qsurvey;
bendplus1 = Bexplcomm(yearend+1);
bbeginning = Bexplcomm(yearbegin);
TACrep = TAC;
}
//______________________________________________________________
FUNCTION calculate_estimable_parameters
//used to calculate selectivity parameters for both commercial and survey trawls
//commercial selectivity
if (phaseacf>0)
{
for (int age = minage; age<=maxage;age++)
commsel(age) = 1/(1+exp(-(age-acf)/delta));
}
else
{
for (int age = minage; age<=maxage; age++)
commsel(age) = comms(age);
}
// survey selectivity
if (phaseage0sel>0)
```

```
{
for (int age = minage; age<=maxage;age++)
{
if(age==0)surveysel(age) = age0sel;
if(age==1)surveysel(age) = age1sel;
if(age>1)surveysel(age) = 1;
}
}
else
{
for (int age = minage; age<=maxage; age++)
surveysel(age) = surveys(age);
}
//____________________________________________________
FUNCTION calculate_kick_off
// uses input paramters to find the spawning biomass of an individual fish
over its lifetime
// it also finds the reparameterised values of the beverton-holt relationship
by using h
//find the number of years we have information for
ncomm = yearendcomm-yearbegincomm+1;
nsurvey = yearendsurvey-yearbeginsurvey+1;
SBPR = 0;
Ntilda(minage) = 1;
for (int y = minage+1; y <= maxage-1; y++)
{
Ntilda(y) = Ntilda(y-1)*exp(-mortality(y-1));
}
Ntilda(maxage)=
(Ntilda(maxage-1)*exp(-mortality(maxage-1)))/(1-exp(-mortality(maxage)));
for (int y = minage; y<= maxage; y++)
```

```
SBPR += Ntilda(y)*maturity(y)*weight(y);

Rzero = Ksp/SBPR;

alpha = SBPR*(1-h)/(4*h);

beta = (5*h-1)/(4*Rzero*h);

//______________________________________________________________

FUNCTION calculate_numbers

// this is the crux of the model

// the numbers at age are calculated, as well as catch-at-age, spawning and exploitable biomass over the years

// and also the fraction of exploitable fish caught by commercial fishers in a year

penalty=0;

for(int year=yearbegin; year <= yearend+1; year++)

{

Bsp(year) = 0;

Bexplcomm(year) = 0;

Bexplsurvey(year) = 0;

}

for (int year = yearbegin; year <=yearend+1; year++)

{

if(year == yearbegin)

{

for (int age = minage; age <= maxage; age++) //values calculated for the first year only

{

if(age==minage)numbersmatrix(year,age) = Rzero;

if((age>minage)&&(age<maxage))numbersmatrix(year,age)

= numbersmatrix(year,age-1)*exp(-mortality(age-1));

if(age==maxage)numbersmatrix(year,maxage)

= (numbersmatrix(year,age-1)*exp(-mortality(age-1)))/(1-exp(-mortality(age)));

Bsp(year) += numbersmatrix(year,age)*weight(age)*maturity(age);

Bsp(year) = posfun(Bsp(year),1,penalty);

Bexplcomm(year)
```

```
+= numbersmatrix(year,age)*weightplushalf(age)*commsel(age)*exp(-mortality(age)/2);
Bexplcomm(year) = posfun(Bexplcomm(year),1,penalty);
Bexplsurvey(year) += numbersmatrix(year,age)*weight(age)*surveysel(age);
Bexplsurvey(year) = posfun(Bexplsurvey(year),1,penalty);
}
}
if(year>=(yearbegin+1)) //values for the rest of the years
{
for (int age = minage; age <= maxage; age++)
{
if(age==minage)numbersmatrix(year,age) = Bsp(year-1)/(alpha+beta*Bsp(year-1));
if((age>minage)&&(age<maxage))numbersmatrix(year,age)
= (numbersmatrix(year-1,age-1)*exp(-mortality(age-1)/2)-catchmatrix(year-1,age-1))*
exp(-mortality(age-1)/2);
if(age==maxage)numbersmatrix(year,age)
= (numbersmatrix(year-1,age-1)*exp(-mortality(age-1)/2)-catchmatrix(year-1,age-1))*
exp(-mortality(age-1)/2)+(numbersmatrix(year-1,age)*exp(-mortality(age)/2) -catchmatrix(year-1,age))
*exp(-mortality(age)/2);
Bsp(year) += numbersmatrix(year,age)*weight(age)*maturity(age);
Bsp(year) = posfun(Bsp(year),1,penalty);
Bexplcomm(year)
+= numbersmatrix(year,age)*weightplushalf(age)*commsel(age)*exp(-mortality(age)/2);
Bexplcomm(year) = posfun(Bexplcomm(year),1,penalty);
Bexplsurvey(year) += numbersmatrix(year,age)*weight(age)*surveysel(age);
Bexplsurvey(year) = posfun(Bexplsurvey(year),1,penalty);
}
}
mu(year) = Catch(year)/Bexplcomm(year);
if(mu(year)>maxmu)mu(year)=maxmu;
for (int age = minage; age <= maxage; age++)
{
```

```
catchmatrix(year,age) = numbersmatrix(year,age)*commsel(age)*mu(year)*exp(-mortality(age)/2);
}
}
//________________________________________________________________
FUNCTION calculate_comm
// finds the log-likelihood function for the commercial CPUE series
for (int i=yearbegincomm;i<=yearendcomm;i++)
{
Bexplc(i) = Bexplcomm(i); // get into the right format
}
qcommvector = log(CPUE)-log(Bexplc);
qcomm = exp(sum(qcommvector)/(ncomm)); // catchability coefficient
qBcomm = qcomm*Bexplc; // predicted CPUE
residualscomm = log(CPUE)-log(qBcomm); // residuals
sigmacomm = sqrt(norm2(residualscomm)/ncomm); // sigma of residuals
LogL(1) = ncomm*log(sigmacomm)+ncomm/2; // log-likelihood
FUNCTION calculate_survey
// finds the log-likelihood function for the survey series
for(int i=yearbeginsurvey;i<=yearendsurvey;i++)
{
Bexpls(i) = Bexplsurvey(i);
}
qsurveyvector = log(survey)-log(Bexpls);
qsurvey = exp(sum(qsurveyvector)/(nsurvey));
qBsurvey = qsurvey*Bexpls;
residualssurvey = log(survey)-log(qBsurvey);
sigmasurvey = sqrt(norm2(residualssurvey)/nsurvey);
LogL(2) = nsurvey*log(sigmasurvey)+nsurvey/2;
//__________________________________________________________
FUNCTION calculate_comm_CAA
// finds the log-likelihood for commercial CAA
```

```
sigmacommmatrixno = 0;
// the predicted catches are summed for age 2
for (int year = yearbegincomm; year <= yearendcomm; year++)
{
for (int age = comm_minage; age <= comm_maxage; age++)
{
if(age==comm_minage)catchm(year,age)
= catchmatrix(year,(age-2))+catchmatrix(year,(age-1))+catchmatrix(year,age);
if(age>comm_minage)catchm(year,age) = catchmatrix(year,age);
}
}
sumagecomm = rowsum(catchm); // gets the total catches in numbers for the year
for (int year = yearbegincomm; year <= yearendcomm; year++)
{
for (int age = comm_minage; age <= comm_maxage; age++)
{
//calculates the proportions of catches from the model
predpropcomm(year,age) = catchm(year,age)/sumagecomm(year);
sigmacommCAAmatrix(year,age)
= predpropcomm(year,age)*pow((log(comm_CAA(year,age))-log(predpropcomm(year,age))),2);
sigmacommmatrixno += sigmacommCAAmatrix(year,age);
}
}
// sigma of the data
sigmacommCAA = sqrt(sigmacommmatrixno/(ncomm*(comm_maxage-comm_minage+1)));
for (int year = yearbegincomm; year <= yearendcomm; year++)
{
for (int age = comm_minage; age <= comm_maxage; age++)
{
LogLcommCAA(year,age) =
log(sigmacommCAA/sqrt(predpropcomm(year,age)))+sigmacommCAAmatrix(year,age) /(2*square(sigmacomm
```

```
//LogLcommCAA(year,age)
= ncomm*(comm_maxage-comm_minage)*comm_CAA(year,age)*log(predpropcomm(year,age));
LogL(3) += LogLcommCAA(year,age); //log-likelihood of this sample
if (predpropcomm(year,age) < 0) cout << "negative props in predpropcomm(year,age)" << endl;
}
}
//____________________________________________________________________________
FUNCTION calculate_survey_CAA
// finds log-likelihood of survey CAA data
sigmasurveymatrixno = 0;
// calculates the predicted catches based on the model
for (int year = yearbeginsurvey; year <= yearendsurvey; year++)
{
for (int age = survey_minage; age <=survey_maxage; age++)
{
predcatchessurvey(year,age) = numbersmatrix(year,age)*surveysel(age);
}
}
sumagesurvey = rowsum(predcatchessurvey); //gets the total number of catches in a particular year
for (int year = yearbeginsurvey; year <= yearendsurvey; year++)
{
for (int age = survey_minage; age <= survey_maxage; age++)
{
predpropsurvey(year,age) = predcatchessurvey(year,age)/sumagesurvey(year);
sigmasurveyCAAmatrix(year,age)
= predpropsurvey(year,age)*square(log(survey_CAA(year,age))-log(predpropsurvey(year,age)));
sigmasurveymatrixno += sigmasurveyCAAmatrix(year,age);
}
}
sigmasurveyCAA = sqrt(sigmasurveymatrixno/(nsurvey*(survey_maxage-survey_minage+1)));
//sigma of the survey
```

```
for (int year = yearbeginsurvey; year <= yearendsurvey; year++)
{
for (int age = survey_minage; age <= survey_maxage; age++)
{
LogLsurveyCAA(year,age)
= log(sigmasurveyCAA/sqrt(predpropsurvey(year,age)))+sigmasurveyCAAmatrix(year,age)/
(2*square(sigmasurveyCAA));
//LogLsurveyCAA(year,age)
= nsurvey*(survey_maxage-survey_minage)*survey_CAA(year,age)*log(predpropsurvey(year,age));
LogL(4) += LogLsurveyCAA(year,age);
if (predpropsurvey(year,age) < 0) cout << "negative props in predpropsurvey(year,age)" << endl;
}
}
// _______________ Per recruit modeling ____________
FUNCTION dvariable yield_per_recruit(dvariable F)
dvariable out;
out = 0;
for (int y = minage; y <= maxage; y++)
{
if(y==minage)nos(y) = 1;
if((y>minage)&&(y<maxage))nos(y) = nos(y-1)*exp(-mortality(y-1))*(1-F*commsel(y));
if(y==maxage)nos(y)=
(nos(maxage-1)*exp(-mortality(maxage-1))*(1-F*commsel(y-1)))/(1-exp(-mortality(maxage))*
(1-F*commsel(y)));
out += F*nos(y)*weightplushalf(y)*commsel(y)*exp(-mortality(y)/2);
}
return out;
FUNCTION dvariable spawnerbiomass_per_recruit(dvariable F)
dvariable out;
out = 0;
for (int y = minage; y <= maxage; y++)
```

```
{
if(y==minage) nos(y) = 1;
if((y>minage)&&(y<maxage)) nos(y) = nos(y-1)*exp(-mortality(y-1))*(1-F*commsel(y-1));
if(y==maxage) nos(y) = nos(y-1)*exp(-mortality(y-1))*(1-F*commsel(y-1))/
(1-exp(-mortality(maxage))*
(1-F*commsel(y)));
out += nos(y)*weight(y)*maturity(y);
}
return out;
FUNCTION dvariable recruit(dvariable F)
dvariable out;
out = (spawnerbiomass_per_recruit(F)-alpha)/(spawnerbiomass_per_recruit(F)*beta);
return out;
FUNCTION dvariable yield(dvariable F)
dvariable out;
out = recruit(F)*yield_per_recruit(F);
return out;
FUNCTION dvariable calculate_muMSY(void)
dvariable F, Fold, fdash, fdashdash;
Fold = 0;
F = 0.00001;
while ((fabs((F-Fold)/F) > 0.0000001) && (F < maxmu))
{
fdash = (yield(F+0.00001) - yield(F))/0.00001; //get first derivative
//second derivative
fdashdash = (yield(F+0.00001)-2*yield(F)+yield(F-0.00001))/(0.00001*0.00001);
Fold = F;
F -= fdash/fdashdash; // update F
}
return F;
FUNCTION dvariable exploitablebiomass(dvariable F)
```

```
dvariable out;
out = 0;
for (int y = minage; y <= maxage; y++)
{
if(y==minage)nos(y) = 1;
if((y>minage)&&(y<maxage))nos(y) = nos(y-1)*exp(-mortality(y-1))*(1-F*commsel(y));
if(y==maxage)nos(y) = (nos(maxage-1)*exp(-mortality(maxage-1))*(1-F*commsel(y-1)))
/(1-exp(-mortality(maxage))*(1-F*commsel(y)));
out +=nos(y)*weightplushalf(y)*commsel(y)*exp(-mortality(y/2));
}
return out*recruit(F);
//spawnerbpr +=
//exploitablebrp += nos(y)*weightplushalf(y)*commsel(y)*
exp(-mortality(y)/2); //*(1-commsel(y)*FRFM/2);
//___________________likelihood profiles__________________
FUNCTION calculate_likelihood_profiles
//likeA = Ksp;
//likeB = h;
//likeC = M;
//likeD = alphamort;
//likeE = betamort;
//likeF = acf;
//likeG = delta;
//likeH = age0sel;
//likeI = age1sel;
//_____________________________________________________
TOP_OF_MAIN_SECTION
arrmblsize=30000000;
gradient_structure::set_CMPDIF_BUFFER_SIZE(6000000);
gradient_structure::set_GRADSTACK_BUFFER_SIZE(1000000);
//___________________________________________________
```

```
REPORT_SECTION
report << "Carrying capacity" << endl;
report << Ksp << endl;
report << "h" << endl;
report << h << endl;
report << "LogL " << endl;
report << LogL << endl;
report << "parameters " << endl;
report << parameters << endl;
report << "Bsp " << endl;
report << Bsp << endl;
report << "Bexplcomm " << endl;
report << Bexplcomm << endl;
report << "Bexplsurvey "<< endl;
report << Bexplsurvey << endl;
report << "qsurvey " << endl;
report << qsurvey << endl;
report << "qcomm " << endl;
report << qcomm << endl;
report << "MSY " << endl;
report << MSY << endl;
report << "MSYL " << endl;
report << MSYL << endl;
report << "MSYR " << endl;
report << MSYR << endl;
report << "f " << endl;
report << f << endl;
report << "mu(yearend)" << endl;
report << mu(yearend) << endl;
report << "Fmax" << endl;
report << Fmax << endl;
report << "Depletion" << endl;
report << Bexplcomm(yearend+1)/Bexplcomm(yearbegin) << endl;
report << "TAC " << endl;
report << TAC << endl;
```

## B.2 Schaefer model

```
//Schaefer model for hake stocks
//Amanda Northrop 2007
DATA_SECTION
init_int yearbegin;
init_int yearend;
init_int yearbeginsurvey;
init_int yearendsurvey;
init_vector Catch(yearbegin,yearend);
init_vector CPUE(yearbegin,yearend);
init_vector survey(yearbeginsurvey,yearendsurvey);
init_int phaseK;
init_int phaser;
INITIALIZATION_SECTION
K 3000;
r 0.3;
PARAMETER_SECTION
init_bounded_number K(500,10000,phaseK);
init_bounded_number r(0.10,1,phaser);
vector biomass(yearbegin,yearend+1);
vector biomasscomm(yearbegin,yearend);
vector biomasssurvey(yearbeginsurvey,yearendsurvey);
vector qcommvector(yearbegin,yearend);
vector qBcomm(yearbegin,yearend);
vector residualscomm(yearbegin,yearend);
number qcomm;
number ncomm;
vector qsurveyvector(yearbeginsurvey,yearendsurvey);
vector qBsurvey(yearbeginsurvey,yearendsurvey);
vector residualssurvey(yearbeginsurvey,yearendsurvey);
```

```
number qsurvey;
number nsurvey;
vector LogL(1,2);
number sigmacomm;
number sigmasurvey;
number MSY;
number MSYL;
number MSYR;
number depletion;
number TAC;
objective_function_value f;
PROCEDURE_SECTION
ncomm = yearend-yearbegin+1;
nsurvey = yearendsurvey-yearbeginsurvey+1;
//calculate_biomass
biomass(yearbegin)=K;
for (int y = yearbegin+1; y <= yearend+1; y++)
{
biomass(y) = biomass(y-1) +
r*biomass(y-1)*(1-biomass(y-1)/K)-Catch(y-1);
if (biomass(y) <= 1) biomass (y)= 1;
}
calculate_comm();
calculate_survey();
// calculate_loglikelihood;
f = sum(LogL);
if (last_phase())
{
//calculate_likelihood_profiles();
//calculate reference points
MSY = r*biomass(yearbegin)/4;
```

```
MSYL = biomass(yearbegin)/2;
MSYR = MSY/MSYL;
TAC = 0.9*MSYR*biomass(yearend);
}
FUNCTION calculate_comm
for (int i=yearbegin;i<=yearend; i++)
{
biomasscomm(i) = biomass(i);
qcommvector(i) = log(CPUE(i))-log(biomasscomm(i));
}
qcomm = exp(sum(qcommvector)/(ncomm));
qBcomm = qcomm*biomasscomm;
residualscomm = log(CPUE)-log(qBcomm);
sigmacomm = sqrt(norm2(residualscomm)/ncomm);
LogL(1) = ncomm*log(sigmacomm)+ncomm/2;
FUNCTION calculate_survey
for (int i = yearbeginsurvey; i<=yearendsurvey;i++)
{
biomasssurvey(i) = biomass(i);
qsurveyvector(i) = log(survey(i))-log(biomasssurvey(i));
}
qsurvey = exp(sum(qsurveyvector)/(nsurvey));
qBsurvey = qsurvey*biomasssurvey;
residualssurvey = log(survey)-log(qBsurvey);
sigmasurvey = sqrt(norm2(residualssurvey)/nsurvey);
LogL(2) = nsurvey*log(sigmasurvey)+nsurvey/2;
REPORT_SECTION
report << "biomassend "<< endl;
report << biomass(yearend+1) << endl;
report << "qsurvey " << endl;
report << qsurvey << endl;
```

```
report << "qcomm " << endl;
report << qcomm << endl;
report << "MSY " << endl;
report << MSY << endl;
report << "MSYL " << endl;
report << MSYL << endl;
report << "MSYR " << endl;
report << MSYR << endl;
report << "Depletion" << endl;
report << biomass(yearend+1)/biomass(yearbegin) << endl;
report << "TAC " << endl;
report << TAC << endl;
```