Report of the
Workshop on Advanced Fish Stock Assessment Techniques (WKAFAT)

ICES Headquarters
3-10 March 2004

International Council for the Exploration of the Sea

Conseil International pour l'Exploration de la Mer

Contents

1 Introduction

1.1 Participants

Pablo Abaunza	Spain
Jesper Boje	Denmark
José Castro	Spain
Wim Demaré	Belgium
Mark Dickey-Collas	Netherlands
Christian von Dorrien	Germany
Georg Engelhard	United Kingdom
Jens Flöter	Germany
Zinaida Frolova	Russia
Sergey Golovanov	Russia
Mikko Heino	Norway
Einar Hjörleifsson (Co-Chair)	Iceland
Jan Horbowy	Poland
Carsten Hvingel	Greenland
Lawrence Kell (Co-Chair)	United Kingdom
Ciaran Kelly	Ireland
Georgs Kornilovs	Latvia
Bjarne Lyberth	Greenland
Dawn Maddock Parsons	Canada
Rasmus Nielsen	Denmark
Rick Officer	Ireland
Lise Helen Ofstad	Faroe Islands
Maris Pliksh	Latvia
Luis Ridao	Faroe Islands
Alexey Russkikh	Russia
Are Salthaug	Norway
Michael Schirripa	USA
Alexandra Silva	Portugal
Dankert Skagen (Co-Chair)	Norway
Marie Storr-Paulsen	Greenland
D. Valentinsson	Sweden
Yvonne Walther	Sweden

1.2 Terms of reference

A Workshop on Advanced Fish Stock Assessment Techniques [WKAFAT] (Co-Chairs: D. Skagen, Norway, E. Hjorleifsson, Iceland, and L. Kell, UK) will be held at ICES Headquarters from 3–10 March 2004 to:

    a)   teach a course covering stock assessment methodology, including evaluation of data consistency, estimation of the state of a stock, projection of stock status, uncertainty evaluation and risk assessment;

    b)   present the open computing environment for fishery science and management currently under development within the Working Group on Methods of Fish Stock Assessments.

    WKAFAT will report by 31 March 2004 for the attention of the Resource Management and the Living Resources Committees, as well as ACFM.

2 Workshop summary

2.1 Background

This is the third consecutive course in fish stock assessment techniques, the first one being held in 2002. The previous ones concentrated on understanding and using standard ICES software. When the future of these courses was discussed in the Resource Management Committee at the ASC in 2003, it was concluded that there is a need to extend and broaden courses in assessment methodology to take on recent developments in assessment methods, emphasise the importance of understanding and scrutinise of the input data, as well as the simulation methods that are developing rapidly at present. It was suggested that this would require a series of courses at different levels, both elementary courses in standard methods as an introduction to the field for newcomers, and more advanced courses to update more experienced analysts on recent developments. The course for 2004 was designed with a dual scope, firstly to consider model design, assumptions and critical evaluation of assessment data in general, and secondly to demonstrate an integrated framework (FLR) based upon R that allows assessments and evaluation of uncertainty to be conducted within a common framework.

The first aspect reflects a concern that the current practise of standardisation may lead to a somewhat uncritical application of analysis tools, without considering the adequacy of the method in sufficient depth. Therefore, the course was designed to give people with some experience in assessment work a general overview of how assessment models are designed, ways to analyse signals in the data, what kind of assumptions can be relevant and other methodological aspects of fish stock assessments. The purpose was to enable analysts to perform analysis of the data adapted to the properties specific for the stock and the data available, as well as to stock specific management procedures.

In addition in recent years, a number of management agencies have funded studies to investigate longer-term management strategies (e.g. MATES/MATACS Kell *et al*, submitted) and there is now an urgent requirement to build upon this accumulated expertise and provide generic tools and a computational frameworks within which stock assessments, statistical analyses, predictions and management evaluations can be undertaken in an integrated manner. Recognising this the WGMG (2004) stated that there is an urgent need for a common software framework that would allow the building of simulation models and to undertake stock assessment and to move away from tools such as XSA to a more flexible environment in which model assumptions can be tested. Therefore, parts of the course was an introduction to an integrated framework that allows assessments and evaluation of uncertainty to be conducted within a common framework.

2.2 The workshop

Course material was principally a collection of power point slides and an integrated package of data-generating and analysis tools implemented in Excel worksheets. The data sets used were synthetic and thus with known properties. The aim of the lectures was to give a theoretical overview of the generic properties of various methods used to generate a historical stock and exploitation estimates. Emphasis was put on classifying the various assumptions as well as strength and weaknesses of different methods. In the practical part of the course emphasis was on exploring the input data (the "observables") and get a hands on understanding of the family models that assume error both in the tuning data and in the catch at age matrix. The prototype to convey the understanding was a separable model as the base. Main emphasis was on historical assessment. It was the intent of the instructors that the main message that participants would take home from this part of the course should be:

- The input data ("observables") contain the principal information about the stock and exploitation. Thus analysis of input data should be considered an integral part of an assessment and should help in determining what type of assessment model is appropriate.
- Understand the principal assumptions that are made in different type of models.
- Not to infer more from the data than they actually can tell.

The integrated Excel tools were the backbone of the practical training in the course. The tools consisted of: 1) a flexible population simulator were various types of stock and exploitation development could be generated and observable could be generated with user controlled degree of error and amount of unaccounted mortality. 2) A statistical catch at age model with varying degree of flexibility in the assumptions made (separable model, power function, non-constant catchability, demonstration of the principal concept of random walk and bootstrap). 3) A standard set of tools displaying principal results as well as diagnostics, and 4) A tool that provided a quick graphical comparison of the known truth (from the simulator) with the estimates from the model. Since the tools were integrated, studying effect of different exploitation development, observation errors as well as the effect of unaccounted mortality could be made relatively easily. The standard output of results and diagnostics provided the group a uniform platform as a basis for discussions.

Three synthetic data sets of "observables" were provided to the group, all based on the same population trajectory. The "observables" were a single catch-at-age matrix, with a plus group, and two sets of age based tuning fleet indices.

Data set 1. Fixed exploitation pattern throughout the history.

Data set 2. Increasing exploitation of younger age groups with time

Data set 3. Same as data set 2, but with doubled noise in the observables.

In addition to the use of these data sets to gain understanding of the principles of stock assessment and to describe the principal building blocks of stock assessment software they where used to study the effect violation in model assumptions (e.g. constant separability, plus group) and to introduce the concept of random walk, bootstrap and influence diagnostics.

An introduction FLR, a framework in R for fisheries modelling, was given. The framework attempts to provide a flexible set of tools in which it is possible to explore alternative hypothesises about stock and fishery dynamics and in particular to consider uncertainty in process, measurement error, estimation, model and implementation error. In this approach complex models, based upon a range of alternative hypotheses, are used to represent alternative states of nature and their probabilities. The frameworks makes it possible to implement and evaluate alternative estimators of stock status, biological reference points, data collection regimes and harvest control rules.

Daily overview of the topics that were covered are given in the in text table below:

Day	Plenary	Practical exercises
1	Basic theory, analysis of observables, introduction to population simulation	Analysis of input data (catch curves, log-ratio). Handout of simulator.
2	VPA, separable models, objective functions	Building statistical catch at age from scratch in Excel
3	Sources of error, model misspecification, diagnostics	Violation of model assumptions
4	More diagnostics, random walk, uncertainty, bootstrap, influence diagnostics.	Practical exercise in random walk, demonstration of bootstrap, influence diagnostics.
5	Day off	Day off
6	More on uncertainty, predictions, overview of available assessment software	Practical exercise in bootstrap, demonstration of NOOA toolbox software.
7	Some basics of R	Practical exercise in R
8	An introduction to the FLR Framework	Practical exercises in FLR

3   Conclusions

Given that the participants had a very diverse practical background in stock assessment there was a consensus that the course was a fair compromise. Those with relatively little practical experiences appreciated the step wise approach, while the more experienced said that the course may have raised more questions than it had answered. One of the participant with the longest practical experience in stock assessment said that the emphasis of analysis input data was valuable and that going back over the basics was always helpful. The introduction to perturbation analysis and principle of bootstrap was thought useful.

It was considered that the use of simulated observables with known error generated from known population and known fisheries to explore the sensitivity of different model assumptions was an extremely good educational approach. The more experience participants thought that the tools made one more critical in the use of assessment models in general. In addition it was thought that the tool would be helpful to address various hypothesis that could explain patterns observed in real data sets. During the course it was learned that the error structure in the simulated data were however somewhat simplistic (random noise only) and could be improved by adding more realistic noise (year effects, autocorrelation). This is necessary to demonstrate how some assessment errors are caused by the noise in the data.

A broad syllabus specifying the main outline of the course should have been made available well in advance. This would have helped potential participants in deciding whether the course was applicable to their needs as well as help in determining the preparatory work needed.

Although there was an understanding that a course like this could not cover all topics some practical exercises on VPA type modelling approach would have been useful since it is currently the backbone of most stocks assessed by ICES. And although the forward based approach covered the principles of short term projections, practical exercises on uncertainties and risks would have been welcome.

4   Recommendations

The number of interested people exceeded the number that could be accommodated, and the group should definitely not be bigger than it was this year. Furthermore, the participants had a more diverse background than expected, ranging from experienced assessment scientists to people with hardly any experience who basically needed an elementary course. It may be suggested to arrange a more elementary course again. This clearly illustrates the need for further courses in assessment methodology.

ICES is heavily dependent on the assessment skills of Working Group participants, and both the conveners and the participants felt strongly that ICES needs to develop a clear strategy to ensure the necessary competence by those who do the assessment work. The diversity in background amongst the participants indicates a need for training at several levels. One may suggest:

A basic, introductory course, covering a basic understanding of data, the basic equations and statistical assumptions, outline of common methods, and some training in using selected assessment tools. This should bring the participants in a position where they can contribute constructively in a working group context.

An advanced course, along the lines of the present one, but spending less time on elementary items, This should be intended for people with good control over the material covered in the basic course, and with some experience from practical assessment work. A course at this level may qualify for some authorisation by ICES as a "Master of assessment" , recognising that the quality of the assessor is as important as the quality of the software.

Specialist courses, bringing in top expertise to cover special items in depth. Examples may be special methods (length-age based methods, Collie-Sissenwine analysis, etc.), statistical aspects (frequentist and Bayesian, weighing of alternative hypothesis, parameterisation), optimisation methods or new approaches in advice (mixed fisheries approach, decision tables). This should be directed mostly at people who are in the position to develop assessment methods, be it within or outside the ICES Working Group framework.

A course, which deals with assessment methodology within the context of management under uncertainty. This course should cover methods for the design and testing of management procedures (i.e. the combination of a particular stock assessment method with particular harvest control rules and their implementation) that consider uncertainty in process, measurement error, estimation, model and implementation error. The participants should be people who at least have the level of experience gained from course 1 above. Due to the advanced nature of the material the number of participants should be limited to 10-15.

It is concluded that ICES should formulate a long term strategy for the continuing education of its members and it is proposed that this should be dealt with in a special study group. The above points may serve as a guideline to the group. The group should include tutors from ICES courses of the recent past to ensure that the practical experience gained will be propagated into the long term strategy. Although the development of a clear long term goal is important there is a need to address the short term needs in a timely fashion. In that regard ICES may consider to offer two courses for next year, covering any two of the four groups of topics described above.

5   Annex

5.1   Further elaborations on the catch equation.

The more common way of defining mortality are as instantaneous relative disappearance rates:

Within a year class cohort, let $N(t)$ be the number at time t, and define

$$Z(t) = 1/N(t)*dN(t)/dt \qquad (A1.1)$$

This differential equation has the solution

$$N(t) = N(t_0)*\exp\left(-\int_{t_0}^{t} z(\tau)d\tau\right) \qquad (A1.2)$$

By assuming that Z is constant over the year, from t to $t+1$,

$$N(t+1) = N(t)*\exp(-Z) \qquad (A1.3)$$

Thus, when deriving Z as $Z = log(N(t)/N(t+1))$, it is tacitly assumed that Z is constant over that period, or in other words, Z is the mean total mortality over the period.

The mean value of N in the time period $t \in [0,1]$, when Z is constant, becomes

$$\overline{N} = \int_{0}^{1} N(\tau)d\tau = N(0)*(1-exp(-Z))/Z = (N(0)-N(1))/Z \qquad (A1.4)$$

Then, the loss becomes

$$\Delta N = N(0)-N(1) = \overline{N}*Z \qquad (A1.5)$$

If we decompose ΔN into additive components, i.e. numbers dead from different causes, we may define the mortality from each cause as corresponding additive components of Z. In particular, the number dead from the fishing is the catch

$$C = F * \overline{N} = F/Z *(N(0)-N(1)) = F/Z*N(0)*(1-exp(-Z)) \qquad (A1.6)$$

where F/Z is the fraction of the total mortality that is due to the removal that is accounted as catch. Equation A1.6 is often referred to as Baranovs catch equation.

Likewise, it follows from equation (A1.2) or from equation (A1.3) that mortalities are additive over time. Thus, if the total mortality in year 1 is Z_1 and in year 2 is Z_2, then after these two years what remains is

$$N(2) = N(1)*\exp(-Z_2) = N(0)*\exp(-Z_1) * \exp(-Z_2) = N(0)* \exp(-Z_1-Z_2) \qquad (A1.7)$$

It follows from the catch equation (A1.6) that C can be derived when N and F and M (i.e. Z) are known. It is also often needed to find $N(1)$ (and implicitly F) from this equation when $N(0)$ and C are known. This problem has a unique solution, but it cannot be solved explicitly. One alternative is to solve it numerically. The other alternative is to use an approximation.

Popes approximation

This approximation to the catch equation is obtained by assuming that all the catch is taken instantaneously in the middle of the year. Before that, the N is reduced by M for ½ year. Then the catch is subtracted, and finally the remainder is reduced by M for another ½ year. This is illustrated in Figure A1.1. The final result becomes

$$N(1) = N(0)*\exp(-M) - C*\exp(-M/2) \tag{A1.8}$$

N(0)

N(0)*exp(-M/2)

N(0)*exp(-M/2) - C

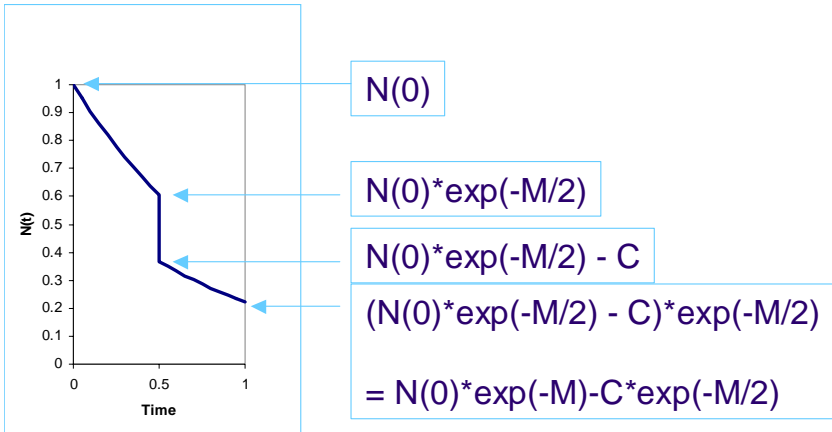(N(0)*exp(-M/2) - C)*exp(-M/2)

= N(0)*exp(-M)-C*exp(-M/2)

Figure A1.1

This approximation is quite close, and it is often used e.g. in spreadsheets, to allow analytic expressions. It is also useful for understanding how the stock is actually estimated from the catches – as a sum of catches expanded to account for the natural mortality, + a term representing what still remains in the stock. See annex 3 for more detail about this.

5.2 Mortality signal in catch data

The signal about mortality is how a year class shows in the catches over time. At any time, the mortality is the catch vs. the standing stock. However, also the standing stock is subject to mortality, and it is assumed that within a cohort, the only change in abundance is the loss due to mortality.

To simplify the discussion, we use the following notation convention:

Referring to any age a and year y:

N0: N(a,y)   C0: C(a,y)   F0: F(a,y)   Z0): Z(a,y)
N1 = N(a+1,y+1)   C1 = C(a+1,y+1)   F1 = F(a+1,y+1)   Z1 = Z(a+1,y+1)

and similar for later years, ages

Now,

$$C0 = N0*F0*(1-\exp(-Z0))/Z0 \tag{A2.1a}$$
$$C1 = N0*\exp(-z0)*(1-\exp(-Z1))/Z1 \tag{A2.1b}$$

In the ratio, the *N0* cancels, making the ratio dependent just on mortalities.

The awkward thing is the expression *(1-exp(-Z)/Z*. One may express this as *exp(-ξZ)*, which allows a simple expression for the ratio *C0/C1*, if ξ is known.

The expression *(1-exp(-Z))/Z* is the integral $\int_0^1 \exp(-Z\tau)d\tau$. According to the mean value theorem of calculus, for a continuous function *f(ξ)* on the interval [a,b]:

$$\exists \xi \in [a,b]: \quad \frac{1}{b-a}\int_a^b f(\tau)d\tau = f(\xi) \tag{A2.2}$$

This ensures that there is a $\xi \in [0,1]$ so that

$$(1-\exp(-Z))/Z= \exp(-\xi Z). \tag{A2.3}$$

The ξ will depend on Z, but not very much, and for all practical purposes it will be a number slightly below 0.5, typically about 0.45. Inserting this is eqns.. A2.1a,b and taking the ratio gives, with a fair precision:

$$\frac{C0}{C1} \approx \frac{N0*\exp(-\xi Z0)*F0}{N0*\exp(-Z0)*\exp(-\xi Z1)*F1} \tag{A2.4}$$

Taking the log and rearranging gives:

log (C0/C1) ≈ (1-ξ)Z0+ξZ1+log(F0/F1) (A2.5)

Thus, the log catch ratio is approximately the average Z over the two years, adjusted for the change in *F*.

The curves can be quite useful to control if the fishery is separable, i.e. we can write *F(a,y)= S(a)*f(y)*, where the selection S(a) is invariant over time. Then,

log (F0/F1) = log (S0/S1) + log (f0/f1)

In this case, the term *log(F0/F1)* in eqn. *A2.5* will be the same for all ages.

The term *(1-ξ)Z0+ξZ1* can be written *S0f0-ξ(S0f0+S1f1),* i.e. the change due to a change in *f* will depend on the selection at the two ages. Thus, this term will not be quite independent on age, unless the selection is flat. This implies that if the selection is constant i.e. the fishery is 'separable', curves of *log (C0/C1)*over time should look almost parallel, but the distance between curves will vary when *f* varies. This is shown in Figure A2.1. The curves to the left are for a population with sinusoidal changes in *f*, the ones to the right are for a step increase in *f*. In both, there is no noise in the catch data. The years and ages are the first years and ages in the pair. Notice that a step increase in *f* first gives a drop, and then a rise in the curves – the increased catch due to the increased *f* first shows in the numerator, i.e. the most recent year in the pair.
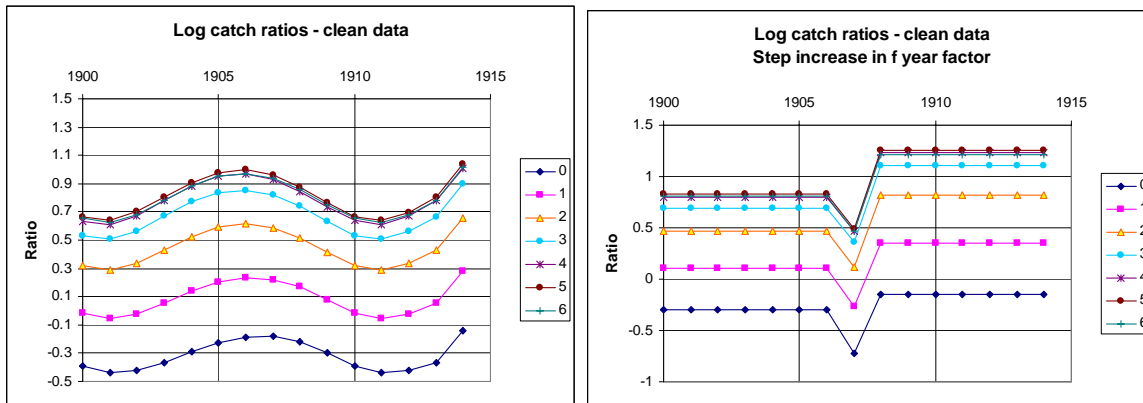

Figure A2.1

5.3   Relative year class strength.

By using the approximation in A2.3, we can write:

logC0 = logN0 – ξZ0 + logF0

logC1 = logN0 – Z0 – ξZ1 + logF1

logC2 = logN0 – Z0 –Z1 –ξZ2 + logF2

logCm = logN0 – Z0 – Z1 – . . – Zm-1 – ξZm + logFm

Taking the average on both sides gives, after some rearranging:

$$\text{logN0} = \frac{1}{m+1}\sum_i \log Ci - \frac{1}{m+1}\sum_i \log Fi + \frac{1}{m+1}\sum_i Zi*(m-i+\xi) \qquad (A2.6)$$

i.e. the recruitment *N0* is the geometric mean of the catches from the cohort, scaled with terms that only depend on the mortality. Thus, if the mortality can be assumed to have been relatively stable, a geometric mean of the catches at some ages from a cohort can give a reasonable indication of the relative year class strength. Figure A2.2 below gives an example. This is from the same data set as above, where the mortality varied cyclically over time. This creates some difference between the curves, but weak and strong year classes are picked up quite well. One may also notice , however, that the information to evaluate the most recent year classes is only the catches at the youngest ages. A common problem in practical assessment appears when there are large catches at the youngest ages in the most recent year. The common sense argument, that this can be either because of targeting of young fish or because of a strong year class, is confirmed by the theoretical considerations here.
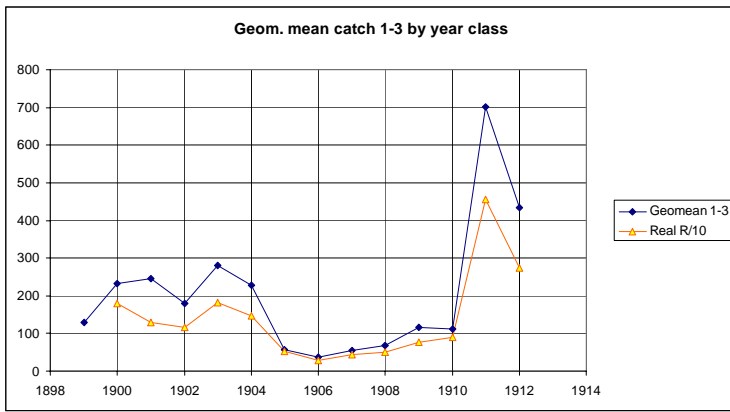
Figure A2.2

5.4 Partition of the year class using Popes approximation

One may argue that the numbers at some age in a year class calculated in a VPA, is the number that is needed to account for the subsequent catches plus the numbers still remaining in the stock, taking into account that there is also loss due to natural mortality. Using Popes approximation, this can be demonstrated very clearly in algebraic terms. Let N0 be the number in the cohort at some age, N1 the number the year after, N2 the number in the year after that, and so forth. Define the catches the same way, and assume for simplicity that Ms are the same everywhere. Then, we can express N0 as follows:

$$N0 = N1*exp(M) + C0*(exp(M/2)) \qquad (A3.1)$$

And

$$N1 = N2*exp(M) + C1*(exp(M/2)) \qquad (A3.2)$$

Gives

$$N0 = \{N2*exp(M) + C1*exp(M/2)\}*exp(M) + C0*(exp(M/2)) =$$
$$N2*exp(2M) + C1*exp(3M/2) + C0*(exp(M/2)) \qquad (A3.3)$$

Repeating this exercise gives:

$$N0 = Np*exp(p*M) + \sum_{i=0,p-1}Ci*exp((2i+1)*M/2) \qquad (A3.4)$$

Thus, *N0* can be written as a sum of one term with *Np*, the number still remaining in the stock at time *p*, and one term which is a weighted sum of the catches after time 0, where the weighting is cumulated natural mortality. The figure A3.1below shows this partition.
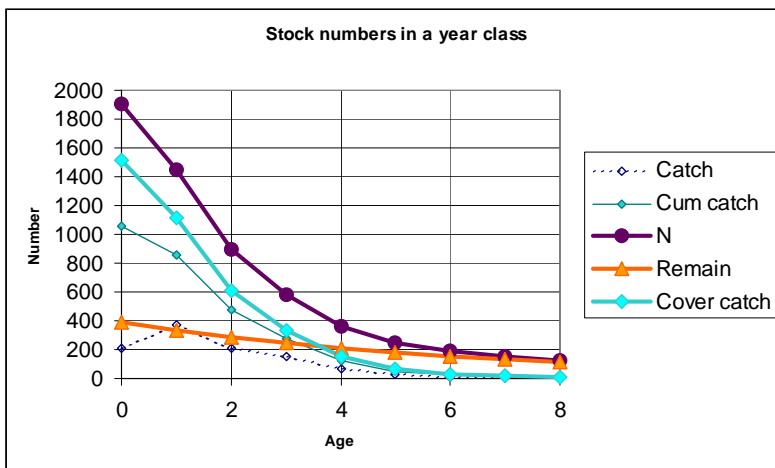

Figure A3.1

5.5   Optimisation

This section is intended for people with some mathematical insight, who want some idea about how optimisation routines work.

To estimate parameters, we find the optimum of an objective function Φ that depends – through the models for the population and 'observables' – on the parameters. The parameters constitute a vector $\boldsymbol{\theta}$. Thus, we formally consider a functional $\Phi(\boldsymbol{\theta})$ on the space Ω of vectors $\boldsymbol{\theta}$. The task is to find the vector $\boldsymbol{\theta} \in \Omega$ which corresponds to the minimum of $\Phi(\boldsymbol{\theta})$.

This is a large field of research in its own right, and we can only scratch the surface here. We shall, however, outline some approaches that are in common use in assessment methods.

Fixed point iteration

The idea here is to use the old parameter set to generate a better parameter set. That is, to find a mapping $\mathbf{g}$: $\boldsymbol{\theta}_{k+1} = \mathbf{g}(\boldsymbol{\theta}_k)$, so that $\Phi\,(\mathbf{g}(\boldsymbol{\theta}_k)) < \Phi(\boldsymbol{\theta}_k)$. According to Banahcs fixed point theorem, the process will converge if the mapping is a contraction mapping, i.e. if the condition: $|g(x)-g(y)| < K|x-y|$ for some $K < 1$, is satisfied.

In practical assessment, the function $\mathbf{g}$ is a two step procedure. The algorithm in XSA can serve as an example:

Start with guessing a set of terminal stock numbers (survivors)

Calculate the N-matrix starting with the survivors and using the catches

Calculate survey catchabilities $q(a) = 1/r(a)$ by

$\log r(a) = 1/\text{number of years} * \Sigma_{\text{years}} \log N(a,y)/\log U(a,y)$

Calculate revised survivors numbers:

$\text{Survivors}_{\text{new}}(a) = 1/\text{number of years} * \Sigma_i \; \log U(a\text{-}i,y\text{-}i)*r(a\text{-}i)\text{-}\text{CumZ}(a\text{-}i,y\text{-}i)$

where Cum Z is the cumulated mortality in the year class.

Goto 2, unless convergence criterium is satisfied.

Thus, the function $\mathbf{g}$ consists of steps 2-5, generating a new set of survivors numbers from the previous ones.

Fixed point iterations are robust provided the contraction mapping condition is satisfied. They are also quite fast, because they do not require calculation of the gradient. They were typically preferred in the early days of fish stock assessment, when computers were less efficient.

Non-linear optimisation

These methods have become very wide-spread in many fields of science, and also in assessment methods. Essentially, these are also fixed point iterations, where the mapping $\mathbf{g}$ is designed to move the multi-dimensional vector $\boldsymbol{\theta}$ as efficiently as possible towards the optimum. The idea is that the gradient shows the direction towards the minimum.

In order to have a minimum of $\Phi(\boldsymbol{\theta})$ at $\boldsymbol{\theta}^*$, two conditions must be satisfied:

The gradient $\nabla \Phi(\theta) = 0$
The Hessian $\nabla^2\Phi(\boldsymbol{\theta})$ is positive definite.

The gradient is the vector $\nabla \Phi(\boldsymbol{\theta}) = \left(\partial \Phi/\partial \theta_i \right)$ and the Hessian is the matrix of mixed partial derivatives: $\mathbf{H} = \left(\partial^2\Phi/\partial \theta_i \, \partial \theta_j \right)$. That the matrix is positive definite means that for any vector $\mathbf{d} \in \Omega$, $\mathbf{d}^T\mathbf{H}\mathbf{d} \geq 0$.

To get a smooth start, consider the one-dimentional case $y=f(x)$. The minimum is where $dy/dx = 0$. If dy/dx is positive at the current x, the function will increase as you increase x, so you need to decrease x, and *vice versa* for negative dy/dx. The question is how much you shall decrease x.

One way is the method of steepest decent. The trick is to construct a function of a parameter α as $g(\alpha) = y(x-\alpha * dy/dx)$, and look for the value α^* of α that gives this function a minimum. That gives you a new value of x: $x^*=x - \alpha^* * dy/dx$. Then you repeat the exercise until convergence. In one dimension, this is not so helpful, but when x is multidimensional, it reduces the problem to a search in one dimension. The multi-dimensional analogue is to search for $\alpha > 0$ so that

$g(t) = \Phi\{\boldsymbol{\theta} - \alpha * \nabla \Phi(\boldsymbol{\theta})\}$

as a minimum. Set $\boldsymbol{\theta}_{new} = \boldsymbol{\theta} - \alpha * \nabla \Phi(\boldsymbol{\theta})$, and repeat until convergence.

Another way - Newton's method - is to assume that $\Phi(\boldsymbol{\theta})$, at least locally, can be approximated by a second order polynomial, that is, that the function $\Phi(\boldsymbol{\theta})$ is represented by the first terms in its Taylor series.

In one dimension, if we want to find the minimum of a function y(x), and we start at x_k, let

$f(x) \approx q(x) = f(x_k) + f'(x_k)(x-x_k) + \frac{1}{2}F''(x_k)(x-x_k)^2$
We want x_{k+1} where q(x) has its minimum, i.e. where q'(x_{k+1})=0
Taking the derivative of q(x) in x_{k+1}, we get $f'(x_k) + f''(x_k)(x_{k+1}-x_k) = 0$

or

$x_{k+1} = x_k - f'(x_k)/f''(x_k)$

Generalising this to a multi-dimensional vector $\boldsymbol{\theta}$, the gradient takes over for the first derivative and the Hessian for the second derivative:

$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{H}^{-1} \nabla \Phi$

so here instead of searching for a minimum along the gradient, we search along the gradient modified by the Hessian. An obvious modification is to search for a more precise minimum in this direction, by searching for an α that minimises

$\Phi(\boldsymbol{\theta}_{k+1})$ with $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \alpha \mathbf{H}^{-1}(\boldsymbol{\theta}_k) \nabla \Phi(\boldsymbol{\theta}_k)$

The next step is to avoid calculating the Hessian, which easily becomes a big task. The idea here is to use the information in the change of the gradient as $\boldsymbol{\theta}$ changes in the iteration.

There are numerous algorithms for Newtonian optimisation, which mostly differ in the way the Hessian is updated. They all start with a proposal for the Hessian, which may be just the identity matrix, and end with a fair approximation to the actual Hessian at the optimum.

Both these approaches lead to a search in one dimension. Several algorithms exist for searching in one dimension, some more efficient than others.

Non-linear optimisation is a wide field, which reflects that there is no straight and standard method applicable to all cases. Many methods have elements of both steepest decent and Newtonian optimisation. Crucial problems are that the algorithm may go astray if the starting values are too far from the optimum, that it may not always find the optimum, that it may get stuck in local optima and that it is very computer intensive. The most computer-demanding part is computation of the gradient. Modern algorithms like automated differentiation can reduce the time to compute the gradient dramatically in selected cases. The commercially available toolbox AD model builder is based on automated differentiation. The Solver add-on facility in EXCEL is a non-linear optimisation routine, but the details of the algorithm are not very transparent to the user. The NAG library has several variants, and is quite well documented. ICA uses a (now obsolete) NAG routine for finding the minimum sum of squares.

Searching routines

This is the most simple-minded way of finding a minimum. Take one parameter, try to increase it a bit and to decrease it a bit, and choose the best value. Then proceed to the next parameter. The first problem is the sequence in which the parameters are evaluated – if parameters are confounded, the first will determine the remainder. One option is to change the sequence randomly each time it starts over again, and this apparently works quite well. The next problem is how large increments one shall use. Clearly, one should start with relatively large increments, and reduce them as the solution comes closer. One may also have to increase intervals again.

Searching routines take time, because the objective function, and accordingly, the population, has to be re-calculated a large number of times. However, their simplicity and robustness may make them preferable if computing time is not critical. AMCI uses a routine of this kind as the primary tool for optimisation, and the experience is that it causes far less problems than the alternative quasi-Newton routine.

5.6   Lecture slides

Lecture slides can be found on the ICES website (www.ices.dk).