MINISTRY OF AGRICULTURE, FISHERIES AND FOOD
DIRECTORATE OF FISHERIES RESEARCH

INFORMATION TECHNOLOGY SERIES
NUMBER 1

Virtual Population Analysis: version 3.1 (Windows/DOS) user guide

C.D. Darby and S. Flatman

LOWESTOFT
1994

Contents

Info. Tech. Ser., MAFF Direct. Fish. Res., Lowestoft, (1): 85pp.
c Crown copyright, 1994

Requests for reproduction of material contained in this publication should be addressed to the authors.

CONTENTS

Page

Example data and results files included on the diskette

Executable programs:	VPADOS.EXE	DOS	disk 1
	VPAWIN.EXE	Windows	disk 2
	RETVPA.EXE	Windows	disk 2
An example icon file:	ASSESS.ICO	Windows	disk 2

In the \DATA directory of disk 1

1	Stock index file	- COD7AIND.DAT
2	Landings file	- COD7ALA.DAT
3	Catch numbers file	- COD7ACN.DAT
4	Catch weights file	- COD7ACW.DAT
5	Stock weights file	- COD7ASW.DAT
6	Natural mortality file	- COD7ANM.DAT
7	Maturity ogive file	- COD7AMO.DAT
8	Proportion of F before spawning	- COD7APF.DAT
9	Proportion of M before spawning	- COD7APM.DAT
10	F on the oldest group by year	- COD7AFO.DAT
11	F at age in last year	- COD7AFN.DAT
12	Fleet catch and effort file	- COD7ATUN.DAT

1. INTRODUCTION

Contents

The methods programmed within the Lowestoft Virtual Population Analysis (VPA) assessment suite have been designed for use in the analysis of fisheries population data. Each method estimates fishing mortality and numbers at age in a stock using data on international catches-at-age and estimates (or assumed values) of natural mortality. Some of the methods make use of fleet-disaggregated catch-at-age data, to calibrate or 'tune' the fishing mortality and stock number estimates to observed trends in effort or in abundance indices. Given the appropriate weight and maturity-at-age data sets, the program will calculate stock and spawning biomasses at the start of the year or at spawning time.

This user guide has evolved with the suite. It provides summaries of the rationale underlying the techniques and of the algorithms used. Where appropriate, guidance notes are included which provide an aid to navigation through the options available within each procedure, or describe data-related problems that may occur during a run. For detailed information on individual techniques the user is advised to read the references listed within the relevant sections.

This version of the Lowestoft VPA suite has been incorporated within the International Council for the Exploration of the Sea (ICES) standard assessment software package.

Acknowledgements : Thanks are due to J.G.Shepherd, J.G.Pope, B.Mesnil, J.W.Horwood and R.M.Cook for suggestions, advice and useful source code.

2. PROGRAM AND USER GUIDE HISTORY

Contents

1983: Shepherd and Stevens (1983) release a Separable VPA program and user guide for the HP1000 computer.

1984: Code incorporated within the ICES assessment suite.

1985: Flatman (1985) describes the development of the assessment suite on the Lowestoft HP1000 computer.

1986: Incorporation of the *ad hoc* tuning procedures.

1987: Version 2.0. Initial release of the DEC VAX implementation. The transfer of the program to the DEC VAX 8200 mini-computer provided the opportunity to correct known bugs and make some enhancements. The user guide for this version was not published.

1988: Version 2.1. A major modification to the *ad hoc* tuning module accepting 20 fleets and 40 years of data. Each fleet could have a different first year of data. Extra summary statistics were produced in the report file. The number of available analysis options was reduced.

1992: Version 3.0. Code modified to run on IBM-compatible PCs within DOS (e.g. MS DOS[1] and DR DOS[2]) and Windows[1]. Four major changes were made to version 2.1 such that :
- (1) The program only handles single sex (or unsexed) data.
- (2) Within *ad hoc* tuning, fleet-derived final year F estimates can only be combined using inverse variance weights; time is the only explanatory variate for the regressions.
- (3) The assessment methods now include Extended Survivors Analysis (XSA).
- (4) The program produces comma separated output files, compatible with other applications.

1993: Version 3.1, with a new user guide. Two new features have been introduced :
- (1) *Ad hoc* tuning has been modified to handle missing data more effectively.
- (2) Additional diagnostics have been added to the XSA tuning output.

---

[1] *MS DOS and Windows are trademarks of the Microsoft Corporation.*

[2] *DR DOS is a trademark of Digital Research Inc.*

3. INSTALLATION OF THE VPA 3.1 SUITE

Contents

3.1 Hardware requirements

Two versions of VPA 3.1 are included on the software disk:

VPADOS: designed to operate within a DOS operating system environment (e.g. MS DOS, DR DOS). It requires 510k of conventional memory and can therefore be run within the 640k of memory available with older PCs.

VPAWIN: compiled to utilise the greater memory capacity available on PCs running Microsoft Windows versions 3.0 and later. Windows runs on the 80286/386/486 family of processors and requires at least 2 megabytes of extended memory. VPAWIN provides some of the features associated with the Windows graphics user interface, particularly the ability to access tuning diagnostic output or other applications during a VPA run.

A maths co-processor is not essential for running the programs, but considerably reduces computation times if available.

3.2 Installation within DOS and Windows

The following instructions will assume that the user wishes to run the program from the hard disk of their PC. The procedures are straightforward but, if problems arise, consult the appropriate DOS or Windows manuals for the specific PC.

The programs are installed by copying either VPADOS.EXE or VPAWIN.EXE and ASSESS.ICO, from the distribution disks into the required hard disk working directory. If required, the example data files in the \DATA directory may be copied into a hard disk directory or accessed directly on the floppy disk.

DOS Installation :

After copying onto the hard disk, VPADOS is ready for use. If the user requires access to the program from within any directory, the installation directory should be added to the environment path.

Windows Installation :

The instructions for setting up VPAWIN under Windows will assume that the user is familiar with accessing the menu bar options using the <alt> key procedures or a mouse; they are described in detail in the Windows manuals. The instructions describe installation of VPAWIN within the user's Windows Applications group.

Without a mouse

Open or select the Windows Applications group or the user's preferred choice, using the <alt> key procedures. From the Program Manager menu bar select **F**ile and Ne**w**. Windows will open a New Program Object window. The Type box should indicate a window item. Use <Tab> to move to the OK button and press *Return*. Windows will open a Program Item Properties window. Using <Tab> to move between boxes, type a name for the program in the Description box, the *drive:\path*\VPAWIN.EXE in the Command line box, and the path of the directory containing the data files in the Working directory box. Using the Change Icon button enter *drive:\path*\ASSESS.ICO in the file name box. Tab to the OK buttons and exit the boxes by pressing *Return*. The Program-item icon will be added to the group.

With a mouse

The above procedure can be used by 'clicking' on the menu items and buttons. Alternatively, start File Manager and open the directory containing VPAWIN.EXE. Open or select the Windows Applications group. Arrange the two windows so that they are adjacent. Within File Manager, place the mouse pointer on the VPAWIN.EXE file name. Press the mouse button and, keeping it depressed, drag the pointer to the Windows Applications window. Drop the new application in the window by releasing the mouse button. Windows will create a new Program-item icon for the VPAWIN program.

To customise the application interface, use the mouse to highlight the new VPAWIN Program-item icon (click on it once). Select the **F**ile option from the Program Manager menu and then Proper**t**ies. Enter a name for the application in the Description box and a directory name in the Working directory box. Select the Change Icon button. In the box for the file name, type *drive:\path*\ASSESS.ICO or any preferred icon file name. Press *Return* or click the OK button. Customisation is completed by using the Properties window OK button.

Note:
If a series of Program-item icons is set up, each customised to start the program in a separate directory, they can be used to provide immediate access to stock data directories. Once created, a Program-item icon can be copied by pointing with the mouse to the icon, holding down the <Ctrl> key and the mouse button and dragging a copy of the icon to the required destination. The Proper**t**ies window of the new Program-item icon must then be edited to enter a new start up directory.

4. DATA REQUIREMENTS

Contents

The following section describes the data files required to run the program. Example data files are provided in the \DATA directory on the distribution disks; these will allow the user to gain familiarity with the data structures and to perform trial runs. The files are similar to those used for the 1992 ICES Division VIIa cod assessments carried out by the Northern Shelf Demersal Working Group (Anon., 1993(a)).

4.1 Program constraints

The program imposes the following constraints on the data sets:

	VPADOS	VPAWIN
The maximum number of years	30	40
Ages must be in the range	0 - 20	0 - 25
The maximum number of fleets	10	20

In addition:

With one exception[1] , all data must be annual.
All data are single sex or unsexed.[2]
References to years must use 4 digits.
The start of the fishing year is January 1st.
Landings are in tonnes.
Numbers of fish are in thousands.
Fish weights at age are in kilograms.

4.2 The stock index file

The program first opens and reads a user-defined stock index file. This is a control file for the stock, and contains the main title, some control parameters, and the names of the data files to be read.

Note*: The file positions are fundamental to the correct operation of the program. It is necessary to place the dummy file name * in the relevant position(s) if one (or more) of the files is not available at the time that the index file is created.**

[1] The constraint that all data must be annual does not hold for the fleet catch data file. Data within this file may represent shorter time periods e.g. seasonal catches, survey data etc.
[2] In order to maintain compatibility with data sets used for previous versions, the sex identifier has been retained in the data file structure. It is always set to 1.

The layout of the stock index file must use the following format:

Index file contents	*Index N^o* [1]
Stock title (80 characters maximum)	
Sex option (value 1 only)	
'Landings (tonnes)' file name	*1*
'Catch-at-age (numbers)' file name	*2*
'Catch weight-at-age' file name	*3*
'Stock weight-at-age' file name	*4*
'Natural mortality' file name	*5*
'Proportion mature-at-age' file name	*6*
'Proportion of F before spawning' file name	*7*
'Proportion of M before spawning' file name	*8*
'F on oldest age group by year' file name (optional)	*9*
'F at age in last year' file name (optional)	*10*
'Fleet tuning data' file name (optional)	*11*

Notes:
(i) It is suggested that 'Stock title' should include species, ICES Division, assessment working group and last amendment date.

(ii) Where estimates of discards are available, 'landings' refers to the sum of the landed catch and discards.

(iii) Catch weight-at-age is the mean weight of a fish at each age, in the catch. Similarly, stock weight-at-age is the mean weight of a fish at each age, in the population (preferably at spawning time - stock weights are primarily used for the calculation of spawning stock biomass).

(iv) File names within the index file should be prefixed by the *drive:\path* for the data directory or disk and should not exceed 8 characters with a 3 character file extension. COD7AIND.DAT in the \DATA directory of the program disk is an example of a stock index file.

The minimum requirements for a run are the catch-at-age and natural mortality files. These will allow a Separable VPA to be performed, and a VPA or Cohort analysis if terminal fishing mortalities are also supplied (interactively or as files 9 and 10). A fleet tuning file (effort and catch data) is required for *ad hoc* tuning or Extended Survivors Analysis (XSA). The remaining files are required for the calculation of stock and spawning stock biomasses.

[1] For reference only - not included in the index file.

4.3 The stock data files (1 - 10)

The general format for each data file is as follows:

Title (80 characters maximum)
Sex (**use only 1**) Index Nº
First year Last year
First age Last age
Data format identifier (see below)
Data .... ⇨ (by age)
"
" (by year)
⇩

It is suggested that in addition to the reference information entered in the stock index file title, a data file's title should include a file contents description. The sex identifier has been retained in order to maintain compatibility with older data sets. It is always set to 1. The Index Nº is a reference for file identification (the appropriate values are listed in Section 4.2) The first age, last age and first and last years must be consistent throughout files 1-10. The last age can be a true age or a plus group. Data values should be space or comma separated.

Reading of each data file is controlled by the data format identifier (DFI). A table defining the DFI options available is given in Appendix 1. A variety of data structure examples are provided in the data files on the program disk. They include:

DFI	Data structure	Example file name
1	2 Dimensional array	COD7ASW.DAT
2	Row array	COD7AMO.DAT
3	Scalar value	COD7ANM.DAT
5	Column array	COD7ALA.DAT

Notes:
(i) Insufficient data in a data file will cause the program to stop, with an appropriate error message displayed on the screen. Too much data, i.e. extra values or an incorrect DFI, will not stop the program since excess values in a record or excess records in the file will be ignored. It is therefore important to specify the correct DFI. A visual check is provided by the program: whilst a data file is being read, a description of the format appears on the screen, together with the file title, Index Nº and contents description.

(ii) Note (i) describes how the program ignores excess data on a line. This can be used to advantage. When handling data sets with a long time series, identification of the location of data for specific years can be tedious. By creating data files with a year label appended to the end of the row, where it will be ignored, the process is simplified. *The first and last ages must, however be correct.*

4.4 The fleet tuning data file (11)

This file is used to supply the fleet data needed for the *ad hoc* and XSA tuning modules. It has a different structure to the stock data files. The data are presented by fleet and are constructed from effort and catch-at-age data. Data values should be space or comma separated. An example of a two fleet data file is provided in COD7ATUN.DAT.

For use within *ad hoc* tuning **all** fleet catch data must have the same final data year (the year for which F estimates are to be made). Missing data (see note (ii) below) for individual ages in the final year are permitted. For XSA tuning it is preferable, but not necessary, for all of the fleets' tuning data to be consistent with the final year of the overall catch-at-age data.

The data file **must** have the following structure:

			Line No [1]
	File title.		*1*
	Number of fleets (+100).		*2*
	Fleet name.		*3*
	First year Last year	(for the fleet)	*4*
Either	Sex code (**1**) Effort code (**1**)	(for *ad hoc* tuning)	*5*
or	Sex code (**1**) Effort code (**1**) Alpha Beta	(for XSA and *ad hoc*)	*5*
	First age Last age	(for the fleet)	*6*
	Effort value Catch numbers at age.......⇨	(for the fleet)	*7*
	"		
	" (by year)		
	⇩ ⇩		

Lines 3 to 7 are repeated for each fleet present in the data file.

Notes:

File title	- May be up to 80 characters.
N° of fleets	- **Must** be 100 plus the number of fleets in the file. Maximum : VPADOS = 10 fleets, entered as 110 VPAWIN = 20 fleets, entered as 120
Fleet name	- May be up to 20 characters.
Sex code	- value **1** only (retained for compatibility with old data sets).
Effort code	- value **1** only (*ditto*).
Alpha and beta	- the start and end of the fishing period for the fleet. They are given as fractions of a year, so that annual data would be entered as alpha = 0, beta = 1.0, second quarter effort as alpha = 0.25, beta = 0.5.
Last age	**- The last age is the oldest true age**; plus group data are not used in the tuning process.

(i) The same tuning data files can be used for XSA and *ad hoc* tuning; alpha and beta values will be ignored by the *ad hoc* routines.

(ii) **Version 3.1** handles zero catches as missing data within both XSA and *ad hoc* tuning. They are weighted out of the analysis. It is assumed that populations do not have zero recruitment and that fishing cannot remove all individuals from a year class. Zero values are considered to be the result of incomplete sampling. If the user considers that the year class strength is near zero (migration effects etc.) replace zero with a suitably low value (but keep an eye on the residuals). *Catch numbers are in thousands so that 1 fish should be entered as 0.001.*

[1] For reference only - not included in the data file.

4.5 Suggested data file naming conventions

Standard formats for data file names avoid the need for separate lists linking file names with particular stocks, and aid directory 'housekeeping'. e.g.

MAFF (or ICES) 3 character species code: e.g. Haddock - HAD, Whiting - WHG

ICES Division: as two character division codes (4A, 4B etc.)
(Total sub-region: use nZ e.g. 4Z - North Sea.)

File contents:

LA	- landings (tonnes - round fresh weight)
CN	- catch-at-age (thousands)
CW	- catch weight-at-age (kg)
SW	- stock weight-at-age (kg)
NM	- natural mortality
MO	- maturity-at-age ogive
PF	- proportion of F before spawning
PM	- proportion of M before spawning
FO	- F on oldest age in each year
FN	- F at age in last year
TUN	- Fleet catch and effort data

An example would be SOL7ECN.DAT which contains sole catch numbers from ICES division VIIe.

4.6 Output file names

If the .csv file name extension is used for all tuning report and output table files, the files can be read directly into spreadsheets e.g. Excel, Lotus, SuperCalc.

5. RUNNING THE VPA SUITE

Contents

5.1 Starting the program

To operate the DOS version, type VPADOS within the VPA working directory, or, if the directory is within the PC environment path, within any working directory.

There are several methods which can be used to run VPAWIN:

(i) After installing the program as an application within a Windows group. Double click on the VPA icon. This method is recommended for quick/easy access to the application.
(ii) Within File Manager, select the working directory from the list of hard disk directories. Highlight the VPAWIN.EXE file name using the cursor or mouse. Press *Return*, or double click on the file name.
(iii) Within Program Manager, select **F**ile from the menu bar. Select the **R**un option, and type the path for the working directory followed by VPAWIN.

VPAWIN is handled as a normal Windows application and provides concurrent access to other applications e.g. file managers, text editors and spreadsheets. The main restriction to usage is that only one VPA window can be opened or held iconised at any time.

5.2 Interactive responses

The program requires information to be supplied by the user. Generally the required responses are either:

(i) a data file name
(ii) a Yes/No answer
(iii) a number, to select a menu option or supply a value.

Data file names and Yes/No responses may be given in upper or lower case. **Y**es/**N**o answers require the initial character only (plus *Return* (or *Enter*)). If the data file is not in either : (a) the start-up directory, defined during the installation procedure, or (b) the directory containing the VPAWIN file, then file names should include the directory path e.g. *drive:\path \filename.ext*. Press the *Return* (*Enter*) key after each answer.

Wherever possible the program offers defaults (enclosed within < >) in the prompt for input. They are set at commonly used or recommended choices. Thus, if in doubt, the user may generally accept the default offered as a sensible choice. Default responses are taken by pressing the *Return* (*Enter*) key only.

After completion of a run the DOS version returns to the DOS prompt. The Windows version presents the user with a termination box. Pressing *Return* (*Enter*) will keep the window open, tabbing to Yes closes the window. This feature allows the user to scroll back through the screen display to check input options (a procedure which can be used at any suitable pause during the interactive session). If the window has been kept open, it is closed by typing *<alt>*, *f*, *x*, or double clicking with the mouse in the top left hand corner.

5.3 Run time errors

Two types of error message are produced by the executable program - features and bugs.

Features are errors messages that are written within the program window. They describe data file format or interactive input errors. They have been designed to help the user identify and solve the specific problem.

Bugs appear within a new window (VPAWIN only). They are run time errors that occur during the execution of the run (e.g. division by 0) and are generated by debugging code incorporated within the compiled program.

During its development, the program code has been designed to cope with as many of these errors as possible. However, all of the potential combinations of data structure cannot realistically be tested. Future developments of the program are anticipated and the authors are therefore interested in all occurrences of bugs. If found, a copy of the data set creating the problem, and a description of the assessment being attempted (with all selected parameter options and values) can be sent to the authors for examination.

6. PROGRAM OVERVIEW

Contents

The following sections describe the execution of the program, review the assessment methods, and discuss the interpretation of results. All examples in the text refer to the data set provided on the program disks, or to the results of runs performed on them. The example data files will allow the user to become familiar with all of the methods within the assessment suite.

On starting the program, the first screen presents the package and describes the use of default options. At the foot of the screen is a prompt for input of the stock index file name. The *drive:\path* will be required if the index file is not in the directory containing the executable program.

6.1 Stock data input

The first part of the program reads the data files which define the stock being analysed. After the user has given the name of the stock index file, the program offers three options:

1) An option to read the first eight data files in the index file - the default option.

2) A fast option to read only the catch-at-age and natural mortality files for a quick run. This run will not produce spawning stock and biomass information.

3) User prompts for input of each of the 8 data file names. The program offers the appropriate name from the index file list as a default. This option is used if data files, other than those listed in the index file, are required for the run.

Any inconsistent age or year range parameters in files 1-8 will cause the program to stop and display a screen table showing the location of incorrect values. Index files 9 - 11 are opened and read only if required, later in the run.

Note:
For new users, getting the program to read the data files is the most frequent source of frustration. If the file formats described earlier, have been adhered to <u>strictly</u>, the program <u>will</u> run. If not check that:

(i) The data file names entered in the index file are correct and that they are preceded by the path name giving the drive and directory in which the files are located (the majority of input errors are produced by incorrect file specification).

(ii) If the program stops during the reading of the data files, examine the name of the data file that the program is attempting to read, it will be printed at the bottom of the screen (an error number is also printed, it is only used for debugging the program code). Examine the entry for the named file in the stock index file. If the file name has been entered correctly, ensure that the data file's header section conforms to that shown in Section 4.3. The specified format controls the number of values to be read from the data file, too few data points will stop the reading process.

6.2 Year and age range selection

The user is asked to select the range of years and ages over which the assessment is to be performed. The selected ranges can be a subset of the years and ages defined within the input data files (the default values). The only condition imposed by the program is that the youngest age in the data files is the youngest age for the run.

6.3 Plus groups and the oldest true age

After selecting the oldest age for the analysis, the user must inform the program whether the oldest age defined in the data files was created as a plus group. If the oldest age selected by the user is younger than the oldest age defined in the data files, a new plus group is automatically created by summing the catch data of the selected age with the data for older ages. The new plus group catch, stock weights and other data attributes are recalculated as catch-number-weighted means. The age preceding the plus group age becomes the oldest true age for the analysis.

All VPA and Cohort analyses are initiated from the oldest true age. Stock numbers for the plus group are estimated independently using the plus group catch number and the F on the oldest true age in the same year.

If the user wishes to perform a run without a plus group, the full age range defined within the header section of the data files **must** be used. The data files should be edited to specify required age range. Data for ages outside this range will be ignored. During interactive input, select the default values offered for the age range (the data file values) and answer 'No' to the question asking whether the oldest age in the data files is a plus group.

Selection of the working plus group age is usually based on the quality of the catch-at-age and fleet catch and effort data for the oldest ages in the data structure. If the data for the oldest ages are known to be of poor quality, or there are significant numbers of missing values, then the ages will contribute little to the assessment; they may even prevent the methods from achieving a converged solution. It would therefore be appropriate to include them in the working plus group.

The initial selection for the plus group age can be modified after an inspection of the output produced by each assessment method. The results contain diagnostic statistics that can be used to detect problems in the data for the older ages. Once the appropriate plus group age has been selected for a stock, it is usually held constant in subsequent assessments, unless there have been significant changes in the quality of the data sets (in either direction).

6.4 Definition of averages for the main output tables

The next menu enables user-defined averages to be set for (a) the fishing mortality results table, and (b) the stock number results table. Averages are defined over a series of years or ages and, with some restrictions, can be unweighted or weighted. Three selection options and a help screen are offered:

1) Full default settings.
2) Selection of the year (column) means for F table only (the remainder are set to the default values).
3) Selection of all the means and ranges interactively.
4) Help.

The default values are calculated from the year and age ranges selected at the start of a run; the settings are described in Appendix 2. During a run the values can be examined within option 4 (Help).

(a) The fishing mortality table: up to three averages may be defined for the year (column) means of the fishing mortality table. Four calculation methods are available:

- (i) Arithmetic mean weighted by catch number per recruit (FBARC).
- (ii) Arithmetic mean weighted by average population number per recruit (FBARP).
- (iii) Arithmetic mean unweighted (FBAR).
- (iv) Exploitation pattern weighting (FBARS).

FBARC and FBARP are described in Shepherd (1983), FBARS in Appendix 3.

The first user-defined average will be used as the reference F. This is generally used by ICES Working Groups to describe the overall level of fishing mortality in each year. The choice is restricted to either (i) or (iii). The default is (iii), the method required by the ICES Advisory Committee on Fisheries Management (ACFM).

The reference F is used to normalise the fishing mortality-at-age values when producing the relative F table. Relative F's allow the examination of year to year variation in the exploitation pattern.

If option 3 has been selected, up to two averages can be defined for the age groups (rows) in the F table. They are also calculated for the relative F table. The averages are restricted to unweighted arithmetic means with the user defining the range of years over which they are calculated.

(b) The stock number table: averages are only user-defined if option 3 has been selected. Up to two averages are available for the age groups (rows) in the table. They are either arithmetic or geometric unweighted means. Defaults are listed in Appendix 2.

6.5 Selection of the assessment method

The Central Menu, which is the next screen displayed, provides the user with a choice of assessment methods. User-defined VPA and Cohort analysis, Separable VPA and *ad hoc* tuning (options 1, 2 and 3) are methods for determining terminal fishing mortalities to be used in a final VPA or Cohort analysis. Extended Survivors Analysis (option 4) is an alternative assessment method to VPA. Option (9) provides a menu for output of the main data and results tables. This section is cyclical. Once a selected option has been completed the menu is re-displayed and the user prompted to select another option. In order to produce output tables of fishing mortalities, population numbers and biomasses, option 9 **must** be chosen after each assessment. If not, the next assessment will overwrite the results of the previous run, or the user will exit the program without a complete output.

7. USER-DEFINED VPA AND COHORT ANALYSIS

Contents

7.1 Description of the method

The methods are 'user-defined' in the sense that the user must supply values for the terminal F's of a VPA or Cohort analysis. Appendix 5 summarises the calculation algorithms.

7.2 Using the method

Selection of option 1 at the central menu results in the presentation of two sub-menus to control the supply of terminal F values.

The first sub-menu offers the user four options for definition of the F values on the oldest true age in each year:

(1) Supplied from a data file (the default file name is file N^{o} 9 from the index file).
(2) Interactively entered by the user.
(3) Left unchanged from the previous run.
(4) A user-defined average of F on one or more younger ages (backwards extension, see *ad hoc* tuning, Section 9.2.2).

The second sub-menu offers the user three options to supply F values for each age in the most recent year. The values may be supplied using the methods (1) - (3) described above.

WARNING: There is an automatic zero catch adjustment in the program. If, for a given cohort, the catch from the oldest age group is zero, the program will move on to the next youngest age group in the cohort. It will continue searching the cohort until a non-zero catch is found and use the user-defined terminal F at that age. This may not be an appropriate terminal F for the age. In this instance, either amend the relevant value in the terminal F file (9 or 10), or adjust the value interactively during a re-run.

Finally the user must select either a full VPA or a Cohort analysis.

7.3 Results tables

This method produces no output directly. The results of the VPA should be printed or written to a file by selecting option 9 from the central menu.

The data files COD7AFO.DAT and COD7AFN.DAT contain the terminal F values for the oldest true age and last year produced by the 1992 VIIa cod assessment (Anon., 1993(a)). They can be used as examples for a user-defined VPA or Cohort analysis.

8. SEPARABLE VPA

Contents

8.1 Description of the method

The development of the Separable VPA procedure has been described by Pope (1977,1979) and Pope and Shepherd (1982). It is implemented in this program with a method for weighting the residuals described by Stevens (1984). Summaries of the algorithms and weighting modifications are presented in Appendix 6.

Separable VPA determines values of fishing mortality from a matrix of catch-at-age data, on the assumption that the exploitation pattern is constant. The procedure is used when tuning data are not available or when tuning data are noisy and may have unreasonably large errors in the estimates for the final year (see also XSA). In addition, the method provides a useful filter for examining the total international catch-at-age data before tuning; high individual residuals may indicate data anomalies.

Pope (1979) demonstrated that the information contained within the data matrix analysed by a Separable VPA is insufficient for the definition of a unique solution. If a unique solution is to be achieved, the number of estimated parameters must be reduced. In addition to the natural mortality value(s) entered within the data files, the user has to specify a 'reference age for unit selection', against which the selection values for other ages will be scaled; and values for:

(a) The fishing mortality on the reference age in the last year (terminal F).

(b) The terminal selection value, i.e. that for the oldest independent age in the data range (used for all years). Selection-at-age (S(a)) is the fishing mortality-at-age relative to that on the reference age.

The method calculates all the other fishing mortalities subject to these constraints. Two algorithm constraints should also be noted :

1) The maximum number of iterations allowed is currently 150 and there is no option for the user to change this value.

2) Any zero catch-at-age values in the data set will be replaced with 10% of the minimum non-zero catch number. This prevents problems with the log catch ratios in the separable analysis.

Note :
"Separable VPA does not use any information which is not available to a traditional VPA, but, given certain starting assumptions (i.e. a constant exploitation pattern, M, terminal F, and terminal S), merely automates the procedure of generating an 'internally consistent' VPA. It is less sensitive to errors than traditional VPA. Its virtue is that it automatically and objectively does what a skilled analyst might do by hand, whilst clearly identifying the assumptions made, and forcing one to think about the existence (or not) of evidence for or against the many possible interpretations of the catch-at-age data." Shepherd and Stevens (1983)

8.2 Using the method

The method is selected by choosing option 2 from the central menu.

The opening questions relate to the definition of the weights to be applied to the log catch ratio matrix used to estimate the fishing mortalities (Appendix 6). The next section requires the user to supply the reference age for unit selection, the number of terminal F's and their values, and the number of terminal S's and their values. Each terminal F/terminal S combination will produce a unique separable analysis. A maximum of 3 terminal F's and 3 terminal S's are allowed. The following is adapted from Anon., (1991(a)) and Shepherd and Stevens (1983).

8.2.1 *Year ratio weights*

The main assumption of Separable VPA is that the exploitation pattern is constant. The use of a long series of data may be disadvantageous as the exploitation pattern may have changed through time. Year ratio weights are used to specify a range of years over which the exploitation pattern is considered to have been relatively constant. Five to ten of the most recent years of reliable data are sufficient, and the program offers the last six years as a default. If other weighting options are required, such as the down-weighting of specific years, weights can be entered by hand.

8.2.2 *Age ratio weights*

For each age group, the age ratio weights may be entered manually or calculated by the program (see Appendix 6). They can be used to down-weight the effect of poorly sampled age groups upon the estimated fishing pattern. Consequently, automatic weights will normally be lowest on the younger and older age groups. Very low weights associated with low selection values on the younger age groups may suggest that including them in the analysis serves no useful purpose. Similarly very low weights on older age groups may suggest that lowering the plus group age could be worthwhile. The user can 'remove' age ratios (if so desired), by specifying manual weighting when prompted, and giving a value close to zero for those age groups.

8.2.3 *Reference age for unit selection*

This is the age at which the value of selection is to be taken as unity. The results of the analysis are not affected by the choice of reference age, in any essential way, but confusing side effects can be avoided by selection of an appropriate value.

The reference age should not be chosen too low (in the partially recruited range) because this leads to the majority of selection values becoming greater than one, and possible interactions with inappropriate choices of terminal selection to produce spuriously domed selection patterns. It should not be too high, near the maximum age, since this makes the procedure liable to crash. The ideal choice is the first age at which the selection pattern may be regarded as fully recruited and subsequently flat. When there are high F values in the middle of the age range (Figure 1a), it is advisable to choose the reference age so as to normalise either on the maximum values or on the flat part of the range so that some intermediate values exceed 1.0 (see Table 1).

Note:
For an initial value try the age group contributing the greatest catch (in number).

Table 1. Two, equally valid, solutions for selection-at-age resulting from differing choices of reference age and terminal S. The results relate to the F pattern illustrated in Figure 1(a).

Reference age	terminal F	terminal S	Age										
			0	1	2	3	4	5	6	7	8	9	10
			Selection-at-age										
4	0.85	0.47	0.12	0.41	0.71	0.94	1.0	0.88	0.71	0.53	0.47	0.47	0.47
8	0.40	1.00	0.25	0.88	1.50	2.00	2.13	1.88	1.50	1.13	1.00	1.00	1.00

8.2.4 *Terminal F*

Terminal F is the (user-defined) F value assigned to the reference age in the last year. Initially, the user is advised to try only one value of terminal F. After which, with (hopefully) a clearer understanding of any patterns apparent within the data and/or results, the user can try widely spaced terminal F values and gradually reduce the range in successive runs. The solutions are equally good interpretations of the data (as judged by the final sum of squares), each statistically valid. The choice as to which is the appropriate interpretation can only be made using additional information e.g. trends in effort over time, groundfish survey data, assumptions about exploitation patterns, etc. An appropriate example is the Separable VPA assessment carried out for the Western mackerel (Anon., 1993(b)). Spawning stock biomasses (SSB) generated by a Separable VPA are 'tuned' to estimates of SSB derived from triennial egg surveys.

8.2.5 *Terminal S*

Terminal S is the (user-defined) selection value assigned to the oldest true age, for all years. For the initial run, try one value of terminal S, after which determination of the appropriate value is a process of refinement. Trial runs are carried out with a range of values and the resulting exploitation patterns examined. The terminal S value should be chosen with regard to the previous parameter selections. By definition S on the reference age is 1.0. Using the same value for S on the oldest age, without thought, can lead to:

(i) an increasing trend in F with age for the older ages if one has selected a reference F at an intermediate, high F value (S=1.0 and 0.6 in Figure 1(a)).

(ii) a spuriously domed exploitation pattern if one has selected a reference F for a partially recruited age group (S=1.0 in Figure 1(b)).

The values of natural mortality-at-age and of selection-at-age are confounded within the separable model (see equation 7 in Appendix 6). Therefore, the user-defined pattern of natural mortality-at-age can influence the shape of the selection-at-age pattern derived from the analysis. If natural mortality varies with age, the influence of the variation on the selection pattern must be taken into consideration.

The final choice is made on the basis of the user's perception of the most likely shape of the selection-at-age curve. In the absence of any prior information, and if natural mortality is considered to be constant for the oldest ages, it may be prudent to choose a terminal selection value that produces a level exploitation pattern for the oldest ages.

Notes:
(i) If available, even noisy fleet data can be used within *ad hoc* tuning or XSA to produce a terminal F value for running a Separable VPA. The terminal F value entered for the run should be an arithmetic average of recent F's on the reference age.

(ii) Within the Separable VPA methodology, the final year F on the oldest true age is a product of the user-defined terminal F and terminal S values. After selection of the terminal F value, if the exploitation pattern derived from a tuned assessment is to be recreated, the terminal S value must be adjusted accordingly.

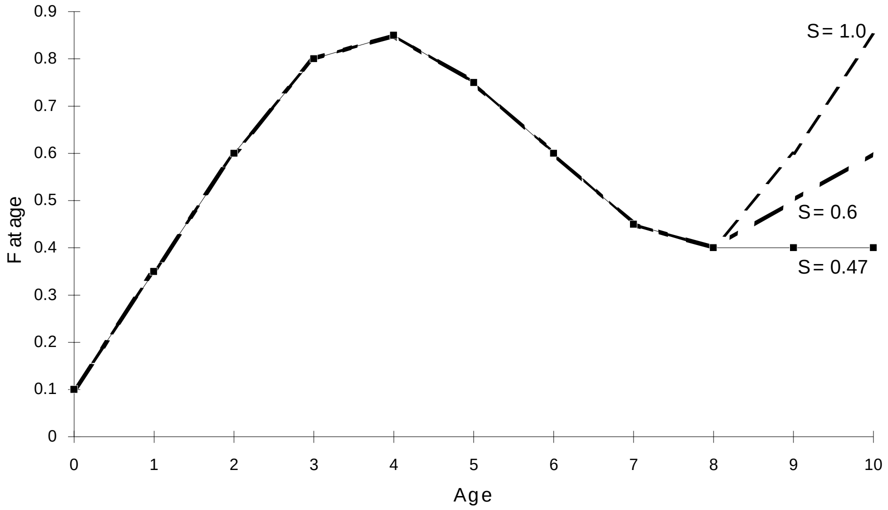


Figure 1(a). An illustration of the effects on estimated F at age of an inappropriate selection for the value of S on the oldest age. Reference age = 4, terminal F = 0.85

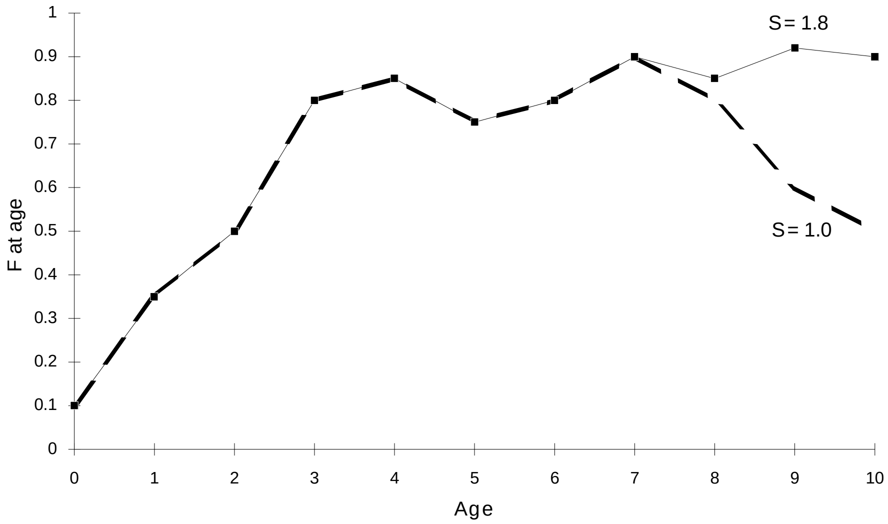


Figure 1(b). An illustration of the effects on estimated F at age of an inappropriate selection for the value of S on the oldest age. Reference age = 2, terminal F = 0.5

8.2.6 S*eparable F and population numbers*

The user is asked whether output of the separably generated F and population numbers tables are required. The separably generated F values are obtained by multiplying the overall fully recruited fishing mortality for the year ($Fo(y)$) by the selection-at-age ($S(a)$) (Appendix 6). The separably generated population numbers are obtained by calculating the recruitment (initial population abundance) for each cohort that would, using the separable F values, give the best fit to the catch-at-age data over the whole cohort (Appendix 6). The catches implied by the smooth separable F values and population abundance estimates will not match the catch data exactly.

8.2.7 *A Separable VPA initialised VPA/Cohort analysis of the catch data*

If only 1 terminal F and 1 terminal S have been chosen, the user may choose to run a final VPA or Cohort analysis (the recommended procedure), in order to generate fishing mortality and population estimates. The terminal F values for the oldest cohort age, which initialise the calculations, are derived from the Separable VPA terminal population abundance (estimated at the start of the year) and the total catch at that age. The population numbers derived from the VPA/Cohort analysis will differ from the Separable VPA population numbers, as the former are derived from the catch-at-age data and not the weighted estimate of recruitment generated by the separable model.

8.3 Analysis procedure

The separable method always produces output which is independent of option 9 from the central menu. An example of the independent output from a run on the test data set is illustrated in Appendix 11. The numbers at the beginning of each of the following paragraphs refer to the circled reference numbers added to Appendix 11.

File or printed output is available, consisting of:

The title, time and date of the run (1), the year and age range of the data and the terminal F and terminal S value for this run.

The number of iterations taken to reach the solution (2), and the initial and the final sum of squared unweighted residuals (SSQ). This provides a measure of the fit to the separable model and should be reduced in the final solution. The final value can be used to derive the root mean square residual (≡ standard error) of the fit to the log catch ratios, an approximation for the coefficient of variation (see Appendix 4) implied if all the lack of fit were due to uniform random variation in the catch-at-age data.

$$\text{Model RMSE} \cong \text{catch-at-age data CV} \approx \sqrt{\frac{\text{Final SSQ}}{2((a-1)(y-1)-2)}}$$

where a is the number of ages and y the number of years of catch-at-age data. The variance of the fit to the log catch ratios is 2x that of the fit to the catch-at-age data. Often the lack of fit is not due to uniform variation and a few residuals contribute a significant proportion.

The matrix of residuals showing the difference between the observed log catch ratio and the estimated log catch ratio (3). Row and column totals of weighted residuals are given, as is the grand total, which the algorithm is attempting to minimise. The row and column totals should be near zero. If they are not the analysis is a poor fit. Row and column weights are printed at the edges of the table.

Often the SSQ value is the result of a few high residuals which indicate poor data for that year and age; these may occur with poorly sampled age groups. The automatic weighting should cope with this adequately, but occasionally it may be necessary to either (i) exclude the age groups by removing younger ages from the analysis or incorporating the older ages in the plus group, or (ii) down-weight specific years manually.

Pattern in the residuals may indicate systematic lack of fit to the model (i.e. a changing selection pattern). Look for year effects running down the columns, age effects across the rows and year class effects which follow the cohort diagonals. If the selection pattern has changed a chequered flag effect can result with positive residuals in diagonally opposed quadrants and negative residuals in the other two.
The fully exploited fishing mortality Fo(y) for each year (4), referred to the reference age. F for each cell in the age/year matrix is obtained from the product of the overall fully exploited fishing mortality for the year (Fo(y)) and the selection-at-age value for the particular age (S(a)).

The exploitation pattern S(a) for each age (5), referred to unity on the reference age and set to the user-defined value on the oldest age.

Options (1) - (5) will be repeated for each terminal F/terminal S combination.

If the user has chosen to use only one terminal F and terminal S and also to output the separably generated F and population numbers tables:

The separably generated F tables (6). These are the F values generated by multiplying the overall fully exploited fishing mortality for the year with the selection value for each age.

The Separable VPA populations-at-age table (7). Obtained by calculating the recruitment (i.e. initial population for each cohort) values that would, using the separable F values, give the best fit to the catch-at-age data over the whole cohort.

After a run with only one value for terminal F and terminal S, the user can choose whether to run a VPA or Cohort analysis. The terminal F starting values for the run are calculated using the raw catch data (including errors) and the 'smooth', Separable VPA generated, terminal population abundances (estimated at the start of the year). The F and population numbers tables generated by the VPA or Cohort analysis (tables 8 and 10 from option 9 of the main menu), are produced by an exact fit to the raw catch data. They will exhibit differences from the 'smoothed' Separable VPA tables ((6) and (7)). The differences are reproduced in (8), the F residuals (F_{sep} - F_{vpa}).

The VPA analysis should be taken as the final solution (as recommended by ACFM (Anon., 1991(a))).

8.4 Results tables

If the user has run a VPA or Cohort analysis, option 9 should be selected from the central menu to output the results of the VPA and biomass calculations.

9. AD HOC VPA TUNING

Contents

9.1 Description of the method

The *ad hoc* tuning algorithms are a family of VPA tuning procedures which derive estimates of fishing mortality at age in the final year from an analysis of the logarithms of fleet catchabilities. They are based on the assumption that catchability (q) is separable by fleet and by age within a fleet. The methods have been reviewed and tested by Pope and Shepherd (1985). Many of them were originally presented at ICES Working Group meetings, and further enhancements have subsequently been produced at such meetings. The algorithms have no formal statistical basis (hence the generalised description, *ad hoc*), but are simple to program and, at the time of their development, provided faster solutions than alternative methods using least squares or maximum-likelihood.

An iterative algorithm (described in Appendix 7) is used to derive estimates of fleet catchability-at-age in the final year. Fleet catchabilities and effort in the final year are used to calculate partial F-at-age: the fraction of overall F-at-age contributed by each fleet. Fleet partial F's are then 'raised' by the ratio of the total catch-at-age and the fleet catch-at-age to give (fleet-based) estimates of total F-at-age. Final year F's for each new VPA iteration are derived from a weighted average of the fleet-based estimates.

In this program, two options for prediction of the final year catchabilities are available. The Laurec-Shepherd procedure assumes constant catchability-at-age with respect to time for each fleet. At each age, a fleet's catchability in the final year is the geometric mean of the tuning time series values. The Hybrid procedure allows the user to apply a constant catchability model to some of the fleets (preferably as many as possible), and let the catchability of the remaining fleets vary exponentially with time. For each age, a linear regression of log catchability against time is used to fit the relationship and derive estimates of catchability in the final year.

The procedures do not provide estimates of terminal F for cohorts that reach the maximum age used for the assessment in years prior to the final year (F on the oldest age). These must be entered by the user or derived as a ratio (proportion) of the mean fishing mortality over some preceding ages in the cohort's terminal year.

Note :
In previous versions of the *ad hoc* tuning algorithm (VPA 3.0 and earlier), zero values in the fleet or total international catch-at-age data have been treated as small catches and replaced with a low value, relative to the catches in the data set being examined (0.2 x the smallest catch-at-age value for the fleet or total catch). This removed the problem of zero values within the calculations for log catchability, but introduced a potential under-estimation bias. The XSA procedure makes the assumption that populations do not have zero recruitment and that fishing mortality at realistic levels cannot remove all individuals from a year class. Zero values are considered to be the result of incomplete sampling and are weighted out from the analysis, removing the potential under-estimation bias. In version 3.1 the *ad hoc* tuning algorithms have been updated to follow the XSA methodology.

9.2 Using the method

The method is selected using option 3 at the central menu.
In the first section of the program, the user supplies the name of the fleet-disaggregated data file, a name for the tuning summary report file and the tuning data year range.

9.2.1 *Time series weights*

The user defines the time series weights to be applied to the tuning data. This option allows the down-weighting or exclusion of older data in the tuning range. The defaults are set at the recommended options - a tricubic time series taper over 20 years - giving most weight to recent data.

The user supplies the weights interactively or selects one of the weighting functions described in Appendix 9.

9.2.2 *F on the oldest age (backwards extension)*

Terminal F for cohorts that reach the maximum assessment age, in years prior to the final year (F on the oldest age), cannot be determined directly from the tuning process. The values must be entered by the user interactively, taken from a data file (the default being the file name given in the index file (File 9)), or calculated during each iteration. The calculations assume that, <u>during the time period used for the assessment</u>, the exploitation pattern on the oldest ages was constant. The oldest age F can therefore be derived as a ratio (proportion) of the average of the fishing mortality on *n* younger ages in the same year. If the exploitation pattern on the oldest ages has evolved with time, the values entered for the ratio and age range should represent a compromise between recent and historical estimates.

The user enters the oldest age F values, a file name, or the values for the ratio and number of ages.

9.2.3 *Missing tuning data*

The program will prompt the user to input terminal F values for ages without final year data. This occurs when:

(i) the user-defined assessment age range is greater than the range of ages for which tuning data are available;

(ii) for any age, either all of the final year tuning data, or the total catch value, are missing (zero).

The user enters a final year F value for each age with missing tuning data.

Note:
For both instances shrinkage to the mean F is recommended. If shrinkage to the mean F has been selected (see Section 9.2.6 and Appendix 10), during each iteration, any user-defined terminal F value will be replaced by an estimate derived from a mean of recent fishing mortality values for that age.

9.2.4 *The fleet catchability time series model*

The user then selects the method to be used for tuning.

1. Modified Hybrid: fits a time trend in log catchability for each age of one or more fleets. To ensure stability of the model at least one fleet must have constant catchability. This is actually a mixed Hybrid + Laurec-Shepherd method, and must not be used for single fleet data sets.

2. Laurec-Shepherd: uses the mean log catchability at each age for all fleets.

9.2.5 *The minimum number of data points required for an estimate from a fleet*

The program requires a threshold to be set for the minimum number of non-zero catchability values that are used for either; the calculation of the mean in Laurec-Shepherd tuning, or the regression analysis used within the Hybrid method. The values include the catchability estimate for the final year derived from the previous iteration. If, for any age, the fleet data set contains fewer values than the threshold, the fleet data will not be used in the overall weighted mean for the age.

The recommended (default) value for the minimum number of data points for a regression is 5. This should prevent the assessment from being dominated by estimates of raised F with low standard errors, associated with small numbers of data points.

9.2.6 *Shrinkage to the mean F*

For each age, the overall estimate of fishing mortality in the final year is a weighted geometric mean of the raised fleet F's (Appendix 7). Raised F's are weighted by the inverse of the fleet's final year log catchability prediction variance.

The overall estimate of F in the final year can be shrunk to a mean of recent F values for the age. A description of the rationale behind shrinkage is given in Appendix 10. The shrinkage mean F used within *ad hoc* tuning is an arithmetic mean of the fishing mortality values calculated for the 5 years preceding the final year. The F values are taken from the previous iteration's VPA.

The user must supply a value for the standard error of the logarithm of the mean F, it is used to weight the means, for all ages, during shrinkage. A value of 0.5 (a 50% coefficient of variation, see Appendix 4) is suggested as a starting point (Anon., 1993(c)). Retrospective analysis procedures for the selection of an appropriate value are described in Section 12.

9.2.7 *The tuning convergence criteria*

Termination of the tuning process occurs when:

$$\Sigma_a \left| \left(F_{a,ty,i} - F_{a,ty,(i-1)} \right) \right| < 0.0001$$

where a indicates age, ty the final data year and i the iterations. This is the sum (across age groups) of the absolute residuals (between iterations) of terminal F. If the routine has not converged after 10 iterations the program displays the residual total on the screen and asks whether the user wishes to continue for a further 10 iterations. This continues until convergence, or the user stops the tuning.

Note:
With some data sets the program may not reach a converged solution before generating extremely low (zero) values of F. This usually requires a high number of iterations (>30). If this occurs the program may fail when calculating subsequent outputs. It is recommended that when using *ad hoc* tuning, the user monitors the residuals displayed after each set of iterations and does not progress beyond 30 iterations before stopping the tuning run and examining the diagnostics file. [*Use the advantages offered by the Windows environment to examine the tuning file without leaving the VPA program*]. If convergence has not occurred, the F-at-age values for the final year, calculated during the final two iterations, are recorded and can be compared. They can be used to identify the ages which are not converging.

9.2.8 *A VPA or Cohort analysis*

Finally the user must select between a VPA or Cohort analysis initialised by the tuned terminal F values.

9.3 Analysis procedure

The user is encouraged to adopt an experimental approach to the selection of the values entered for each assessment parameter. A progressive series of investigative runs should be used to examine the effect of variation in the selected values on the final predictions.

It is suggested that for an initial run the following options are adopted:

(i) Run an *ad hoc* assessment over the full age range using the default time series weights (20 years, power 3) this will focus the analysis on the most recent fleet data.
(ii) Use the defaults to calculate F for the oldest age as the mean of the preceding 3 ages (a ratio of 1.0)
(iii) Use the default tuning procedure, Laurec-Shepherd (constant catchability). <u>Do not</u> shrink to the mean.
(iv) Note the number of iterations taken for the procedure to converge. Terminate the run after 30 iterations if convergence has not been reached, or earlier if diverging.

Examine the tuning output file. Appendix 12 presents an example run, annotated for the following description.

The date, time and tuning file used for the run (1).

A log of the selected assessment options (2).

The convergence results (3). Appendix 12 presents the results for a converged run.
If convergence is not achieved, the final year F values from the last two iterations are printed (see the XSA tuning results in Appendix 13 for an example of the latter). The F values will indicate the ages that are varying between iterations and the degree of variation.

Examine the fishing mortality values resulting from the run (4). Check for extreme values, especially those at the older ages which might better be incorporated into the plus group.

Note:
VPA versions 2.1 and earlier produced this table during the penultimate iteration of the run; it would then be checked against the table of F's produced by option 9 of the main menu. Differences between the two tables would indicate that the procedure had not converged. Following the introduction of the convergence test to version 3.0, this table is now produced during the final iteration and is equivalent to that produced by option 9. The degree of convergence is examined in (3).

Examine the log catchability residuals for each age for all fleets (5). 99.99 indicates a missing (zero) total catch or fleet catch value. The values can indicate changes in the stock - fleet interactions. Look for year effects running down the columns, age effects across the rows and year class effects which follow the cohort diagonals. Recent and sudden changes in catchability may require removal of the fleet from the assessment.

For each age, plots of the residuals against time can be used to reveal trends in log catchability. One way to achieve this is to give the tuning output file a comma separated file name extension (.csv) and import it into a spreadsheet package.

Note:
If only one fleet data set is available and the Laurec-Shepherd constant log catchability model is used without shrinkage to the mean, the residuals in the final year will all be 0.0; the terminal F values are generated using the fleet's average catchability for the age. If shrinkage to the mean is selected or the assessment is tuned with more than one fleet, F in the final year is a weighted mean. The estimate of catchability derived for each age will differ from the fleet's mean and the final year residuals will not be zero.

The significance of trends in catchability, examined in (5), can be tested using the diagnostics presented in the summary statistics (6). As a quick check, look at the slope of the log catchabilities for each age (8), for each fleet separately. Slopes which have the same sign, and exceed twice their standard error consistently for most of the important age groups, are considered significant and indicate that the Hybrid model could be appropriate. If the Hybrid model is used, constant catchability should be maintained for as many fleets as possible unless there is strong evidence against it. Remember that these are log catchabilities and that a trend with time indicates an exponential trend in catchability.

Note:
In the diagnostic statistics produced by Laurec-Shepherd tuning, a previous version of the VPA suite (2.1) did not calculate slopes for log catchabilities. A slope of zero indicated that the Laurec-Shepherd (constant catchability) algorithm had been used. Later versions calculate the slope and intercept of the log catchabilities for both the Laurec-Shepherd and the Hybrid methods. This allows investigation of possible trends with time under both prior assumptions about catchability (the slope from a Laurec-Shepherd run will not be the same as that from a Hybrid analysis).

Examine the mean log catchability (pred. log q) and its standard error for each age and fleet (7). The standard error of the log catchability is an indicator of the quality of the data (a fractional coefficient of variation, see Appendix 4). Values greater than 0.5 indicate problems with that age for the fleet. High standard errors for the older ages of all fleets indicate that the assessment should probably be re-run with the problem ages incorporated into a younger plus group.

When combining fleet-derived estimates of terminal F at each age, weighting by the inverse of the prediction variance of the log catchability will reduce the influence of poor fleet data. However, if for any fleet, the standard errors of the majority of the important ages are poor, the user may wish to remove the fleet from the analysis altogether.

The estimate of the partial F contributed by the fleet (9) and the raised F (10) are printed. Raised F's are the individual fleet predictions of overall F: the level that would have been recorded if the fleet had taken the whole of the international catch for that age. The values can be used to identify incompatible predictions from the individual fleet data sets.

For each age, the overall weighted mean terminal F is printed (11) along with its internal (SIGMA(int)) and external (SIGMA(ext)) log standard errors. Also given is the overall standard error (SIGMA(overall) (12)); it is the larger of the internal and external values.

The internal standard error for an age is calculated from the (prediction) standard errors of the fleet's final year log catchabilities; it corresponds to the within samples variance. The external standard error is calculated from the scatter of the logarithms of the raised F values; it corresponds to the between samples variance (Topping, 1978). The formulae are given in Appendix 7. If shrinkage to the mean has been selected, the internal and external standard errors include the F shrinkage value.

SIGMA(overall) is a good approximation to the fractional coefficient of variation of the mean F and should be used as a measure of the accuracy of the prediction. If it is large (greater than 0.3) for important age groups, then the assessment should be treated with caution.

If the values of the internal and external standard errors differ significantly, there is a discrepancy between the fleet's estimates for overall F (the raised F's (10)). The variance ratio (13), (external s.e.)^2/(internal s.e.)^2, may be tested as an F statistic with *n* and *n* -1 degrees of freedom, where *n* is the number of fleets contributing a raised F estimate. Values exceeding 3 imply conflicting signals from the fleets. Too small a value implies an unexpected correspondence of the tuning fleets in relation to the inherent noise.

After examining the results from the preliminary run(s), the user can progressively refine the assessment by making adjustments to the selected options:

In the initial run it was assumed that the exploitation pattern for the older ages was flat. A ratio value of 1 was used to determine F on the oldest age. This assumption needs to be verified. After selection of a new plus group age, the age preceding this becomes the oldest (true) age for the next assessment. New values for the age range and the ratio used to estimate F on the oldest age can be estimated from the F-at-age table ((4) in Appendix 12). Using the results from the assessment on the complete age range, choose a block of F values for which the calculations have some convergence. The recommended procedure is to use a block of six - ten years with a final year that predates the terminal year by at least three years. Select a range of ages preceding the new oldest age that, if possible, do not include pre-recruits. The selected range will be the number of ages to be used in calculating F on the oldest age in the next assessment. For each year in the block of data, calculate a mean F over the new age range. The ratios of the F values for the new oldest age, to the means calculated for the younger ages, provide estimates for the ratio to be used in the next assessment. If the ratios exhibit strong variation, increase or reduce the number of ages in the average in an attempt to stabilise the means. If there are no extreme values, use the average of the ratio estimates.

For stocks with only a few ages contributing to the fishery, the number of ages available for the selected range may be small. In cases such as this try to avoid using pre-recruit ages for which catchability may be related to year class abundance. Extended Survivors Analysis (XSA) uses an alternative method for estimating F on the oldest age and is more appropriate for such stocks.

Once the age range, time series weights and analysis model have been selected, the shrinkage option can be used to introduce a degree of rigidity to the terminal F estimates and reduce potential biases in the results (see Appendix 10).

9.4 Results tables

In order to output the VPA or Cohort analysis results, generated using the terminal F values obtained by tuning, it is necessary to select option 9 from the central menu.

10. EXTENDED SURVIVORS ANALYSIS (XSA) Contents

Ad hoc methods for tuning single species VPA's to fleet catch per unit effort (CPUE) data are sensitive to observation errors in the final year because they make the assumption that the data for that year are exact. In addition, the methods fail to utilise all of the year class strength information contained within the catches taken from a cohort by the tuning fleets. Extended Survivors Analysis (XSA), (Shepherd, 1992), an extension of Survivors Analysis (Doubleday, 1981), is an alternative approach which overcomes these deficiencies. In general, the algorithms used within the *ad hoc* tuning procedures, exploit the relationship between fishing effort and fishing mortality. XSA focuses on the relationship between catch per unit effort and population abundance, allowing the use of a more complicated model for the relationship between CPUE and year class strength at the youngest ages. A summary of the XSA algorithm is presented in Appendix 8.

10.1 Description of the method

The XSA algorithm performs:

1) A Cohort analysis of the total catch-at-age data to produce estimates of population abundance-at-age, and total fishing mortalities.

2) Adjustment of the CPUE values for the period of fishing defined using the alpha and beta parameters in the fleet tuning file, into CPUE values that would have been recorded if the fleet had fished only at the beginning of the year. The adjusted values are directly comparable with the population abundances at the beginning of the year.

3) Calculation of fleet-based estimates of population abundance-at-age from the adjusted CPUE values and fleet catchabilities. Fleet catchabilities-at-age are assumed to be constant with respect to time (for ages considered to be 'recruited'), or dependent on year class abundance (for ages to be treated as 'recruits') within the model described by Shepherd (1993) and implemented in the program RCT3 (previously RCRTINX2) described by Darby and Shepherd (in prep.).

4) Calculation of a least squares estimate (weighted mean) of the terminal population (survivors at the end of the final assessment year) for each cohort in the tuning range using the fleet-derived estimates of population abundance-at-age. These terminal populations are used to initiate the Cohort analysis in the next iteration.

The process iterates until the convergence criteria described for *ad hoc* tuning are achieved. Various options are available for catchability analysis, time series weighting and shrinkage of the weighted estimates.

10.2 Using the method

The method is selected using option 4 of the central menu. The user supplies the name of the fleet tuning file, the name of an output file for the XSA tuning diagnostics, and the range of years to be read from the tuning data file. Subsequent selections determine the characteristics of the tuning run.

10.2.1 *The ages at which catchability is independent of year class strength*

The first option relates to the treatment of the ages considered to be recruits. It allows for the possibility that the catchability of the younger age classes is **dependent** on year class strength. A power law model is used to describe the relationship between catchability and abundance (Shepherd, 1993 and Appendix 8). A least squares linear regression is used to determine the relationship between log CPUE and log VPA population abundance for the recruiting ages of each fleet. The CPUE values are then re-used as indices within the fitted relationship to derive fleet-based estimates of population abundance in each year.

The user enters the first age at which catchability is considered to be **independent** of year class strength: the value applies to all fleets. If the youngest age, for which tuning data is available, is chosen, all ages will be treated as having catchability independent of year class strength.

10.2.2 *The age above which catchability is independent of age*

In common with the *ad hoc* tuning procedures, the catchability for the oldest assessment age is under-determined (Appendix 8). In order to overcome this uncertainty, XSA makes the assumption that catchability is independent of age (constant) above a specific, user-defined age. The procedure uses the catchability value derived for the selected age to calculate estimates of population abundance for all older ages. A default constraint is imposed to ensure that, at the very least, the catchability of the oldest true age is fixed to that of the preceding age.

The user enters the age above which catchability is to be independent of age (constant): the value is used for all fleets.

10.2.3 *The XSA default settings*

The program presents the user with a screen illustrating the default settings for the calculations used within the XSA algorithm. They are commonly used ICES Assessment Working Group options. If the options are suitable, pressing *Return* (or *Enter)* or answering no (N/n) will use the displayed settings. Answering yes (Y/y) will allow the user to select values for the following options:

10.2.4 *Time series weighting*

The user defines the weights to be applied to the tuning data time series. This option allows the down-weighting or exclusion of earlier data in the tuning range (see Appendix 9). The defaults are set at the recommended values - a tricubic time series taper over 20 years. This is a compromise between allowing recent data to receive the strongest weight during the estimation of the final year F and population abundances, whilst at the same time, allowing historical data to influence values derived for earlier years.

The user selects one of the weighting functions described in Appendix 9.

10.2.5 S*election of the recruitment regression model*

If the catchability of recruiting ages is to be treated as dependent on year class strength, the type of the regression to be used in the analysis must be selected. The options are C - calibration and P - predictive regression. The recommended option is calibration regression (Rosenberg *et al.*,1992; Shepherd, 1993).

A threshold value must be entered for the minimum number of points required for a regression analysis. For each recruitment age, a fleet's catchability is considered to be constant with respect to abundance, if the number of non-zero data points is lower than the threshold value.

Note:
(i) Unlike RCT3, XSA uses all the data points in the regression, calculating a population abundance estimate from each index value. An analysis with a threshold value of 5 non-zero data points will utilise all of the tuning data from a fleet with 5 years catch data.

(ii) It has been noted that, with some data sets, calibration regression fails to establish a relationship between CPUE and population abundance (indicated by a low r-square value). The procedure fits regression lines with severe slopes, which generate values too large for the PC to handle, when used for the prediction of population abundance estimates. In order to cope with this situation, the extreme values are weighted out from the analysis. This is achieved by setting the prediction s.e. for the problem age, in the offending fleet, to a high value (99.99). A warning on the screen indicates the fleet and age for which this is occurring.

10.2.6 *Shrinkage to the mean*

The next two questions are associated with the two forms of shrinkage available for the analysis. The rationale behind shrinkage is described in Appendix 10. The equations used for shrinking the XSA terminal population estimates are given in Appendix 8.

The first form of shrinkage used in XSA is shrinkage to the population mean, originally introduced in the recruitment analysis programs RCRTINX2 and RCT3. It is applied to the recruit ages <u>only</u>. Terminal population estimates (calculated at the end of a year) for the recruiting age *a* are shrunk to the time series weighted geometric mean of the population abundance estimates for age *a*+1 (calculated by the preceding VPA iteration, at the beginning of a year). The weight given to the shrinkage mean is the inverse of the variance of the weighted geometric mean.
Rosenberg *et al.* (1992) have used simulation analysis to show that when estimating year class strength, prediction accuracy can be improved by the use of <u>calibration</u> regression with shrinkage to the population mean. The default settings supply this combination. If <u>predictive</u> regression is used, shrinkage to the population mean is equivalent to a double shrinkage and should be avoided.

The second form of shrinkage is shrinkage to the mean F, introduced after the Working Group on Methods of Fish Stock Assessment (the Methods Working Group) , examined the biases in assessment predictions of terminal F and survivors (Anon., 1991(b), 1993(c), see Section 12 and Appendix 10). For each <u>age</u>, the program calculates a mean of the F values for the *j* years in the assessment range that precede the final year. The mean F and the total catch in the final year are used to calculate an estimate of the survivors at the end of the year. It is this estimate towards which the fleet-derived predictions of terminal population for the final year are shrunk (Appendix 8). The survivors in <u>years</u> prior to the final year are shrunk to a terminal population derived from an average of the F values of *k* ages that precede the oldest true age, and the total catch value for the oldest true age.

The number of years and ages used for the shrinkage mean (*j* and *k*) should be customised to the dimensions of the data set being studied. The number of years should be less than or equivalent to the longest time period over which the exploitation pattern is considered to have been constant. For stocks with only a few ages contributing to the catches, particular attention should be given to the number of ages in the shrinkage range. Pre-recruit ages should not be included.

The program requires a weight to be applied to the means used in F shrinkage. It is used for all ages and years. Retrospective analysis procedures for the selection of an appropriate value are described in Section 12.

The user enters a weight for the F shrinkage mean. A value of 0.5 is suggested as a starting point (Anon., 1993(c)).

***Note*:**
For VPA's extending, historically, past the years for which tuning data are available, it is strongly recommended that XSA is used with F shrinkage.

At initialisation, the program uses a seed value for the terminal population of each cohort (usually giving terminal $F \approx 0.65$ and equal for all the cohorts). New terminal population estimates are produced during successive iterations until convergence is achieved. If the shrinkage to the mean option is not selected, the terminal populations of cohorts without tuning data will not be modified. If shrinkage is used, the terminal populations for the oldest age are calculated from F values for younger ages, a procedure equivalent to the fixed exploitation pattern used within *ad hoc* tuning. Terminal populations at the end of the final year are derived from recent F values and, for the ages treated as recruits, the population abundance means of the next age. Cohorts without tuning data are therefore initialised by survivor estimates derived from all of the information available.

10.2.7 *A threshold value for the minimum standard error*

A minimum value can be set for the log catchability standard errors used as weights in the estimation of survivors. This option is designed to reduce the possible influence, on the weighted estimates of terminal population, of fleets with a low variance resulting (by chance) from the use of a few data points. Fleet standard errors that fall below it, are replaced by the threshold value.

The user defines the level at which the threshold is set; the suggested default is 0.3.

Note: An examination of an XSA diagnostics file produced with a low shrinkage weight (high CV, e.g. 1.0) and without the s.e. threshold option, will give an indication of the magnitude of the raw fleet catchability standard errors.

10.2.8 *Prior fleet weighting*

The final option will only be available if more than one tuning fleet is used for an assessment run. During calculation of the weighted terminal population means, it allows down-weighting (or removal) of all estimates derived from a particular fleet. Weights should be in the range 0.0 to 1.0 (corresponding to complete removal through to no change).

10.2.9 *The user-selected tuning results*

After completion of the user input, the program runs until the convergence criteria described for *ad hoc* tuning are achieved. If not reached after 30 iterations, the program will request clearance for additional iterations, repeated in sets of 10.

After convergence, or user termination of the iterations, the program writes the XSA diagnostics file. Output of additional tuning information can be requested at this stage:

1) The estimates of terminal population abundance and their weights, used for the overall mean, are printed within a 'long' format output. The 'short' format gives the summary data only.

2) The CPUE data, adjusted to the beginning of the year, can be printed for additional analysis. The transformation uses the alpha and beta values defined in the tuning data file and the total mortalities from the final iteration.

The program then returns to the main menu to allow output of the population statistics tables.

10.3 Analysis procedure

The user is encouraged to experiment with the XSA package to become familiar with the data set being assessed. Several runs will be required for selection of the options to be used in the final XSA assessment. Annual repetitions of the procedures described below should not be necessary: once determined, the XSA input parameters selected for the assessment of a particular stock may be held constant from year to year. However, significant alterations to the tuning fleet structure, or to the age range of the assessment, may require repetition of the full procedure.

For an exploratory run use all available ages and select the following options :

(i) Set the age at which catchability is to be <u>independent</u> of year class strength to the oldest true age (plus group age - 1). This treats all ages except the oldest as 'recruits' (i.e. allows for the possibility of catchability varying with abundance).

(ii) Set the age at which catchability is to be independent of age at the age selected in (i).

Answer 'No' to changing the default settings. This will run an assessment with:

(iii) A tricubic weighting function with a range of 20 years.

(iv) All ages, except the oldest, treated as catchability proportional to population abundance. A calibration regression with a requirement for a minimum of 5 data points is used to fit the catchability model.

(v) Shrinkage of the terminal population estimates of age *a* to the mean population abundance of age *a*+1, for all ages except the oldest.

(vi) Shrinkage to the mean F, using a standard error of 0.5, with means taken over 5 years and 5 ages.

(vii) Log catchability standard errors lower than 0.3 replaced by a standard error of 0.3.

(viii) No prior fleet weighting.

Select the 'short' format output and do not print the corrected CPUE data, as these are not required at this stage.

Examine the XSA diagnostics file. An annotated example is given in Appendix 13, it is not an exploratory run.

The user should check:

The stock title, run time and data file name (1).

The listing of the options selected when defining the method (2).

Either, the number of iterations performed to reach convergence, or if convergence was not achieved (as in the example), the final year F values for the last two iterations (3).

The time series weights (4).

Selection of the age at which catchability is independent of year class strength:

The diagnostics from the initial run can be used to select the age at which the catchabilities of all important fleets are independent of year class strength.

For each fleet, examine the regression statistics (13) for the ages with catchability dependent on year class strength, especially the slope, the R square and the overall regression standard error (Reg. s.e.). The slopes should be tested to see whether they are significantly different from 1.0: if not then catchability is constant with respect to population abundance. The t-value given in the table is derived from : t = (slope - 1.0)/ s.e slope. It can be tested against the t statistic for the required confidence level, obtained from Student's t table with *n*-2 degrees of freedom - *n* is the number of data points used for the regression (No Pts). Find the first age at which catchability is independent of year class strength for all important fleets.

If all of the first age regression slopes appear to be independent of year class strength, the age of independence should be chosen with reference to the information sources from which the tuning data are derived. If the tuning data for the first age do not include survey data sets and the fleet data for this age are known to be well sampled, then the use of constant catchability (with respect to year class strength and time) for all ages is appropriate. If the tuning data for the first ages include survey data then it is recommended that, at the very least, the first age is treated as recruits. This approach utilises the RCT3 model for the estimation of the terminal population at the youngest age in the final year (including shrinkage to the population mean). Terminal population estimates are therefore obtained using a combined approach (Anon., 1993(c)).

Note:
If requested, XSA will print the final iteration's transformed CPUE values after a run (see (29)). Plotting the log of the CPUE values against the log of the VPA population abundance estimates given in (6), allows an examination of the distribution of the data points about the fitted regression relationships. The graph can be used to examine whether one or two extreme values are dominating the relationships. This practice has also proved useful when examining the fleet CPUE data for ages at which calibration regression generates extreme values that are subsequently weighted out from the tuning process.

Selection of the age at which catchability is independent of age (definition of the catchability plateau):

Re-run the assessment with the new value for the age at which the catchability is considered to be independent of year class strength. Set the age at which catchability is independent of age at the penultimate true age. This run is used as a standard against which further runs are tested.

Catchability on the oldest age is poorly determined by the model (Appendix 8) and, to overcome this difficulty, the catchability values for the oldest ages are taken to be equivalent to that of a younger but **fully recruited**, age. In order to introduce the greatest possible degree of stability to the assessment, it is necessary to set the age at which catchability is independent of age as low as possible in the fully recruited age range, without affecting the fit of the model at the older ages. The selection of the appropriate age is a process of refinement.

Examine the log catchability values for the ages with constant catchability with respect to time (11) and their standard errors (12). This can be done by plotting the means for all fleets against age on the same graph. If, for the oldest ages, catchability does not exhibit large variation from age to age and there are no trends with respect to age, the youngest fully recruited age at which catchability appears to be independent of age is the preferred choice. For the selected age, examine the log catchability standard errors for each fleet. An alternative selection may be required if all of the important fleets' log catchabilities, at the selected age, are poorly estimated by the model (s.e.'s >0.5). If the log catchability standard errors are acceptable, a series of runs with a stepwise reduction in the age above which catchability is fixed, from the oldest true age-1 to the selected age, should be carried out and the log catchabilities and their standard errors compared with the standard run. Noticeable differences between runs should indicate when to stop.

Notes:
(i) One reason for choosing the penultimate age for the initial run is that if a trend in catchability with age exists, it is possible to force an inappropriate plateau by selecting too young an age.

(ii) Large variations in catchability for all of the oldest ages in the assessment make it difficult to choose an appropriate age for fixing catchability.

In either of these situations it is recommended that the assessment is carried out with catchability for the oldest age determined from the penultimate age. This removes the constraints on the older ages and allows the model to determine the majority of their catchability values independently. In addition, F shrinkage must be used, otherwise the model is badly under-determined and noisy. Due to the increased freedom within the model, the run may require more iterations to achieve a solution.

Run the assessment with the new catchability analysis parameters:

If there are only a few ages contributing to the assessment, examine the regression statistics for the youngest ages. Changes to the catchability model applied to the older ages may have influenced the fit at the youngest ages.

The fishing mortality-at-age table (5). If F on the oldest age is substantially higher than F at the preceding ages this may be an indication of an inappropriate selection for the age at which catchability has been set independent of age (fixed). It is possible that the catchability-at-age distribution is dome-shaped and catchability has been fixed at a high value which is inappropriate for the subsequent ages. This results in artificially high F's on the oldest age. The log catchability residuals for each fleet can be examined (see below) as a check.

Population numbers-at-age (6).

Survivor estimates for the end of the final year (the terminal populations) (7).

The taper weighted geometric mean of the final VPA populations (8). The terminal populations for the recruiting ages are shrunk to these means. For example, the survivors for age 1, estimated at the end of the year, are shrunk to the mean of the abundances estimated for age 2 (at the beginning of the year). The weight applied to the population shrinkage mean is the reciprocal of the square of the standard error listed in (9).

For each fleet in the data series:

The log catchability residuals (10). 99.99 indicates a missing (zero) total catch or fleet catch value. The values can indicate changes in the stock - fleet interactions (changes in catchability). Look for year effects running down the columns (e.g. 1985,1987,1990 in Appendix 13), age effects across the rows and year class effects which follow the cohort diagonals (e.g. the 1987 year class in Appendix 13). Recent and sudden changes in catchability may require removal of the fleet from the assessment.

The log catchability residuals for ages above the age at which catchability is constant with respect to age are derived from :

$$\text{residual} = \text{Ln}(q_{est(y,a,f)}) - \overline{\text{Ln}(q_{est(af,f)})}$$

where $\overline{\text{Ln}(q_{est(af,f)})}$ is the mean log catchability for the age at which catchability is fixed (af), and $q_{est(y,a,f)}$ are the values derived from the final iteration $(CPUE_{(y,a,f)} / N_{(y,a)})$. It is often seen that, if the age at which catchability is held constant is inappropriate, the catchability residuals for the subsequent ages generate blocks of all positive or negative values.

For the ages with constant catchability with respect to time, the log catchability mean (11) and its standard error (12). Examine the standard errors of log catchability for each age and fleet. The standard error of the log catchability is an indicator of the quality of the data (a fractional coefficient of variation of the fleet catchability for the age, Appendix 8). Values greater than 0.5 indicate problems with that age in the fleet data. High standard errors for the older ages indicate that the assessment should probably be re-run with the problem ages incorporated in a younger plus group.

When combining estimates of terminal population derived from the fleet catches taken at each age, weighting by the inverse of the log catchability variance will reduce the influence of poor fleet data. However, if the standard errors of the majority of the important ages for a fleet are poor, the user may wish to remove the fleet from the analysis altogether.

The user should then review some of the parameter selections used to define the characteristics of the assessment run:

The number of ages and years used in F shrinkage at the margins of the assessment. Are the F values for the selected age and year ranges stable? Would an increase or reduction in the selected ranges add greater stability to the assessment and/or reduce the influence of changes in exploitation pattern and extreme years? Does the range of ages selected include pre-recruits?

Does the standard error applied as a weight in the F shrinkage introduce or remove bias in the estimates of final year F and populations-at-age generated from the assessment (see Section 12)?

Are the catchability values at a few ages or from one fleet dominating the assessment? Does an increase in the cut-off threshold for the minimum standard error reduce the dominance and produce a different result when estimates are derived from all fleets equally?

The final assessment run:

After completing the definition of the assessment parameters, the user should make a final run printing the full diagnostic output. This will allow a full analysis of the final year survivor estimates. Short format outputs will only print the fleet and overall summary tables.

10.4 XSA diagnostic statistics new to version 3.1

Additional diagnostic statistics have been incorporated in the XSA output: they are illustrated in Appendix 13.

The XSA algorithm fits the catchability proportional to year class abundance regression to all ages, regardless of whether the results are used within the analysis (14). This allows an examination of the regression slopes and standard errors for ages fitted with the catchability independent of year class strength model.

For each final year terminal population, the program prints the year class, the age of the cohort in the final year and the model used to derive catchability-at-age (15).

If the user has selected the long format diagnostics output:

For each fleet and each age in the cohort's history, the estimate of the terminal population at the end of the final assessment year (16) and its raw weight (17) are printed. The raw weights are defined in Appendix 8. They are used with the individual estimates of survivors to calculate the fleet-based and overall weighted means. Zero values ((16a) and (17a)) indicate that the fleet has no data for the age.

If the short diagnostics output was selected the individual fleet estimates at age will be omitted and only the following statistics will be tabulated:

A fleet-based weighted mean of the cohort's survivors (18). This is derived from the estimates obtained from the fleet catches at each age in the cohort's history (the raw weights, printed in the long format output, can be used to identify the specific contribution of each estimate).

The internal standard error of the terminal population estimate obtained from a fleet (19). It is derived by combining the standard errors associated with each estimate in the weighted mean and corresponds to the within samples variance of the fleet-based terminal population estimate (Appendix 8).

The external standard error of the estimate of survivors obtained from each fleet (20). This is the standard error of the terminal population estimates derived at each age (Appendix 8); it corresponds to the between samples variance.

If the values of the internal and external standard errors differ significantly, this indicates a discrepancy between the individual estimates generated by the fleet catches. The variance ratio (21), $(\text{external s.e.})^2/(\text{internal s.e.})^2$, may be tested as an F statistic with n and n-1 degrees of freedom. n is the number of estimates of terminal population abundance contributing to the mean, i.e. the number of years in which the fleet removed catches from the cohort. Values exceeding 3 imply that the independent estimates obtained at each age are providing conflicting signals. Too small a value implies an unexpected correspondence of the tuning fleets in relation to the inherent noise.

The scaled weights (22) are a measure of the proportional contribution of the fleet's estimates (for all ages) to the overall survivors estimate for the cohort. The weights are not actually used in the derivation of the overall mean, which is a weighted mean (using the raw weights (17)) of all the disaggregated (by fleet and age) estimates, including the population and F shrinkage means (if used). The scaled weight is given so that contributions from each fleet can be compared.

The terminal F that would be generated by using the estimate of survivors derived from the fleet to initiate the VPA (23). This value is equivalent to the fleet's raised F generated by the *ad hoc* tuning procedures. Discrepancies in the signals provided by the fleet data sets can be detected by comparing the F values or the survivor estimates.

If the age is a recruiting age in the assessment and shrinkage to the population mean has been selected, then the estimate of survivors used in the population shrinkage is printed with its standard error, scaled weight and F (24). The F shrinkage terminal population, the s.e. supplied by the user, scaled weight and F, are also given (25).

The overall geometric weighted estimate of survivors at the end of the final year (26). This value is derived by combining all of the estimates of terminal population abundance; the estimates at each age from all fleets and the shrinkage estimates (Appendix 8). The raw weights used for the overall weighted mean are listed in (17).

The internal standard error and external standard error of the overall mean (see Appendix 8), and the variance ratio (27). If the variance ratio exceeds 3, conflicting signals are being given by the disaggregated (by fleet and age) estimates of terminal population. The F test carried out for the individual fleet estimates can be repeated for the overall mean. In this case n is the summation, across fleets, of the number of years in which a fleet removed catches from the cohort. The individual estimates of terminal population (16) and the fleet variance ratios (21) can be used to identify the fleets and/or ages that are causing problems.

The overall terminal F value for the cohort (28). This is calculated using the overall weighted mean terminal population and the catch in the final year.

After the diagnostics for each age:

An optional output of each fleet's corrected CPUE data (29). The data are transformed to the beginning of the year using the total fishing mortality values from the final iteration and the alpha and beta values entered in the diagnostics file. The data can be used to examine the distribution of data points about the fitted catchability regressions, as described previously (13).

10.5 Results tables

It is necessary to select option 9 from the central menu, in order to output the full time series of XSA-generated fishing mortalities, population numbers and biomasses.

11. SELECTION OF OUTPUT TABLES

Contents

Selection of option 9 from the central menu will display a menu showing the 17 different output tables available.

<u>Menu of Tables</u>

Table	1	Catch numbers at age
Table	2	Catch weights at age (kg)
Table	3	Stock weights at age (kg)
Table	4	Natural mortality (M) at age
Table	5	Proportion mature at age
Table	6	Proportion of M before Spawning
Table	7	Proportion of F before Spawning
Table	8	Fishing mortality (F) at age
Table	9	Relative F at age
Table	10	Stock number at age (start of year)
Table	11	Spawning stock number at age (spawning time)
Table	12	Stock biomass at age (start of year)
Table	13	Spawning stock biomass at age (spawning time)
Table	14	Stock biomass at age (start of year) with SOP correction
Table	15	Spawning stock biomass (spawning time) with SOP correction
Table	16	Summary (without SOP correction)
Table	17	Summary (with SOP correction)
Code	18	Will produce data tables 1,2,3,4,5,6,7
Code	19	Will produce result tables 8 to 17 inclusive

Tables 1 to 7 may be selected at any time and print the data entered into the program from the stock data files. Using code 18 will produce all 7 tables.

Tables 8 to 17 can be produced after running one of the four options from the central menu. The results from a pass through one of those methods should be printed before re-selecting any of those options or before selecting option 0 to stop the program. Using code 19 will produce all 10 results tables.

Selection of summary Table 16 and/or 17 will also produce Table 8 - Fishing Mortality and Table 10 - Stock Number.

The file RESULT.CSV contains a complete set of all the tables.

Average fishing mortalities and stock numbers defined by the user (Section 6.4) will appear on the corresponding tables. Within the tables, numbers are usually printed in thousands, biomass weights are in tonnes. <u>The tables are automatically scaled by the program</u>. The magnitude of the scaling is printed at the top of the numbers tables (e.g. 10∗∗-3 : thousands, 10∗∗-6 : millions), and at the base of the summary tables.

SOP correction refers to the sums of products correction factor:

$$\text{SOPCOFAC} = \text{Landings} / \Sigma_a \text{ (total catch numbers at age } a \text{ x catch weight-at-age } a\text{)}$$

It is assumed that the landings weights are exact, and if this factor is not 1.0, there are errors in either the total international catch numbers-at-age or the values of catch weight-at-age used for the assessment. Biomass and spawning stock biomasses can be corrected for the errors. The without SOP correction options are available for assessments for which the SOP corrections have been applied during the creation of the catch-at-age or weight-at-age data sets.

12. RETROSPECTIVE ANALYSES

Contents

12.1 Introduction

Retrospective studies have established that patterns of consistent under- or over-estimation bias in estimates of F and population numbers-at-age can be produced by the application of assessment methodologies to fish stock data (Sinclair *et al*., 1990, Anon., 1991(b)). Such biases may cause problems in the advice given to managers and therefore need to be examined and if possible removed from the assessment and subsequent predictions.

Sinclair *et al*. (1990) and Anon,. (1991(b)) have shown that the biases in F and N estimates appear to be stock specific, and data induced. They are not attributable to a particular tuning methodology. Sinclair *et al*. (1990) concluded that the retrospective patterns found in the estimates for the stocks of the Northwest Atlantic could result from patterns of misreporting, trends in catchability, or mis-specification of natural mortality. Each will affect the data in a particular way and therefore influence the outcome of the tuning procedures.

Anon., (1991(b)) established that the degree of bias could usually be reduced by the introduction of shrinkage to the mean F (Appendix 10) to the assessment packages. Subsequent work by the Methods Working Group has examined the influence of the degree of shrinkage imposed on the assessment (Anon., 1993(c)). It recommended that retrospective analyses are used regularly to screen stock assessments.

12.2 Description of the method

For each stock and analysis procedure a series of assessments are performed with the terminal year decreased by one year at each run, simulating the results of assessments in progressively earlier years. All input parameters to the analysis are held constant, e.g. number of tuning data years, time series weights, reference ages, shrinkage standard errors. The values estimated by the most recent assessment, derived from all available data, are assumed to be the 'truth' and compared with the estimates from the runs which pre-date it. The accuracy of an assessment methodology, for the particular parameter selections used in the analysis, is determined by its ability to consistently predict the 'truth'. Bias is the degree to which the method consistently under- or over-estimates the 'truth'. The analysis procedure usually involves the creation of retrospective time series plots for particular assessment predictions (e.g. F or population numbers-at-age, SSB) followed by a statistical or subjective analysis of the accuracy and bias of the method.

Retrospective series can be used to investigate the influence of particular assessment parameters (e.g. shrinkage to the mean F) on the accuracy and bias of the terminal year estimates. Retrospective runs are performed with a range of values for the selected parameter (all other parameter values are held constant), and the value producing the 'best' retrospective pattern is chosen as the optimum value for the assessment of the particular stock. In order to simplify the analysis, it is assumed that there are no interactions in the effects on the assessment predictions.

'Best retrospective pattern' has been interpreted in two ways. In the investigative analyses carried out by the Methods Working Group (Anon., 1991(b), 1993(c)) the stock age range was divided into three categories, recruits, pre-recruits and fully recruited ages. Time series plots were created using average F values and average population numbers for each of the three categories and a series of statistics computed in order to compare the accuracy and bias of assessment predictions. Assessment Working Groups are usually under restrictions of time and have tended to examine the variation in predictions for average F over a series of fully recruited ages (Fbar) and/or the predicted total spawning stock biomasses (Anon., 1993(a)). Subjective comparisons of the retrospective time series plots are made by eye. If XSA is used with recruitment survey information included within the tuning fleet data sets (the recommended procedure, Anon., (1993(c))), the user should also examine the accuracy and bias in the estimates of recruitment to the stock.

When carrying out retrospective analyses the selection of tuning fleets to be used in the assessments is important. Fleets with short time series should be avoided. As the program steps back through the data range they may drop out when there are insufficient years of fleet data for the specified analysis. In addition, short series with artificially low standard errors may erroneously dominate the assessment. The use of short time series can introduce sudden changes in the retrospective patterns and should be avoided. If required, the short series can be reintroduced for restricted retrospective analyses after the full runs.

12.3 A program to produce retrospective runs

A program designed specifically for the purpose of performing retrospective *ad hoc* tuning and XSA runs (RETVPA.EXE (Windows only)) is included on the program disks.

HEALTH WARNING : **The program is at an early stage of development. It has not been tested to the same degree as the VPA suite and may have bugs. At the present stage of development it is recommended that when running the program the user does not perform repeat runs for each method, but re-starts the program for each retrospective series.**

The program is started in a similar manner to VPAWIN or VPADOS. It requires the same format for the index and data files. After selection of the assessment procedure, the program will not require input for the range of years to be used in tuning, it will read all the tuning data.

A directory must be entered for the tuning files generated by each assessment in the retrospective series. This is followed by input of a three character prefix for the tuning file name. The program will create each tuning file by adding RT<yr>.CSV to the end of the prefix, <yr> represents the terminal year for the assessment being performed.

In addition to the tuning diagnostics files, the program generates output tables with one of two user selected formats. They are:

1) F and population numbers-at-age tables for each assessment in the retrospective series (Tables 8 and 10 from the main menu of the output from the main suite) and the stock time series summary, Table 16. The program will create each output file by adding RO<yr>.CSV to the end of the user-defined prefix.

2) A single file containing the F and population numbers tables from each run in the format defined for the SAS program used to generate the figures and summary tables presented in Anon., (1991(b)) and Anon., (1993(c)).

All of the inputs used to select the tuning procedure parameters are those described for the main suite programs. After selection of the method and parameters, the program asks an additional series of questions in order to define the characteristics of the retrospective run:

The first defines the number of years for which the retrospective analysis is to be run.
Enter the finishing year for the run; the earliest terminal year. Acceptable values lie between the penultimate year in the data file and the earliest year in the complete tuning range that allows 5 years of data for each fleet. Short fleet tuning series should be removed if a retrospective series

with sufficient comparison years for an acceptable analysis is to be achieved (8 years will give 4 assessments in a retrospective series).

The next question refers to the use of the tuning data time series. The user can select between (1) a tuning range window, e.g. 10 years of fleet data, which is moved backwards with the terminal year for each new assessment, or (2) the use of the full data range in the tuning file and the removal of the most recent years data as the program steps back for each new terminal year.

Time series weights, if used, are moved back with the assessment terminal year.

The program then begins the retrospective analysis of the data sets, printing the terminal year for the current assessment to the screen.

If the assessment has not converged after the required numbers of iterations (described earlier for each of the methods) the program will request clearance for further iterations.

When converged or the current assessment is terminated by the user, the program will write the output data to the file defined earlier. It will then proceed with the next assessment in the series.

12.4 An example

The 1992 Northern Shelf Working Group (Anon., 1993(a)) used retrospective analysis when selecting the most appropriate method for assessment of the VIIa cod stock (the example data set). The method was used to examine the bias in estimates derived from two assessment methods, Laurec-Shepherd tuning and XSA, with and without shrinkage. All other parameters were held constant. Figure 2 illustrates the results of this exercise. Only the Fbar values for ages 2-5 are illustrated.

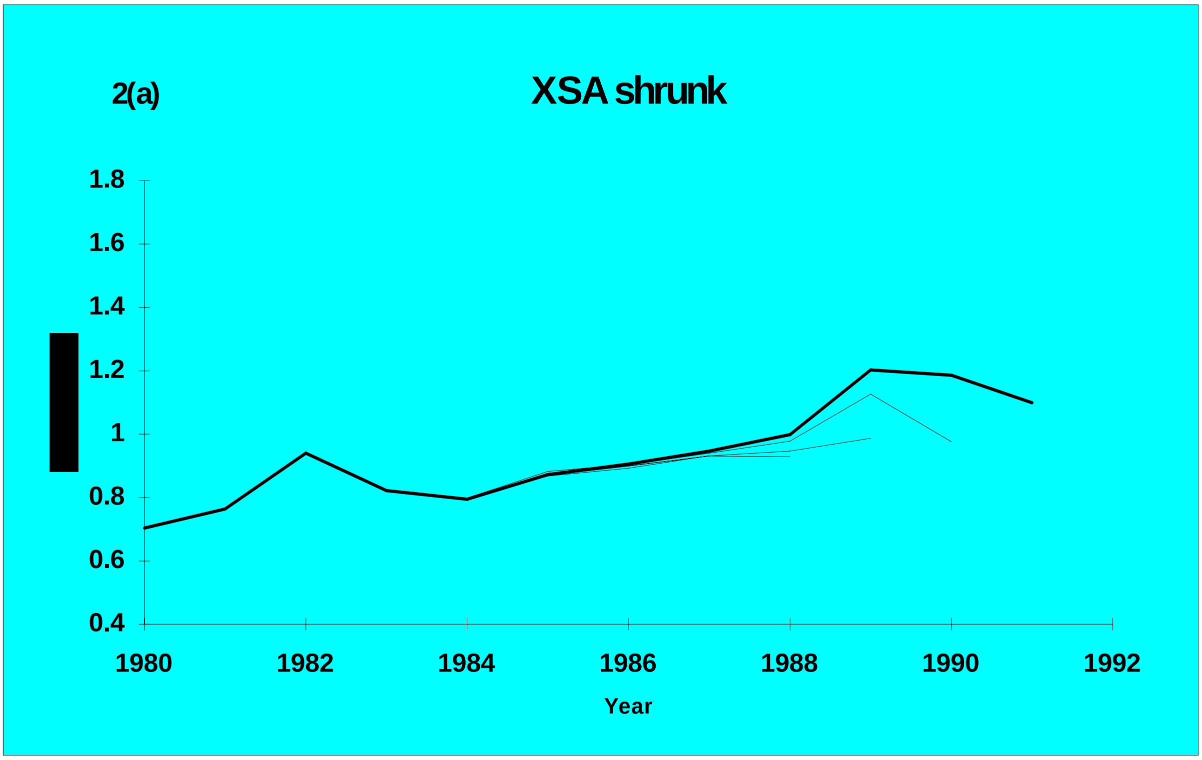
The plots show the Fbar values for the years 1980 to 1991. Figure 2(a) illustrates the variation between XSA assessment predictions with shrinkage, Figure 2(b) a comparative plot without shrinkage. Figures 2(c) and 2(d) illustrate the plots for Laurec-Shepherd tuning.

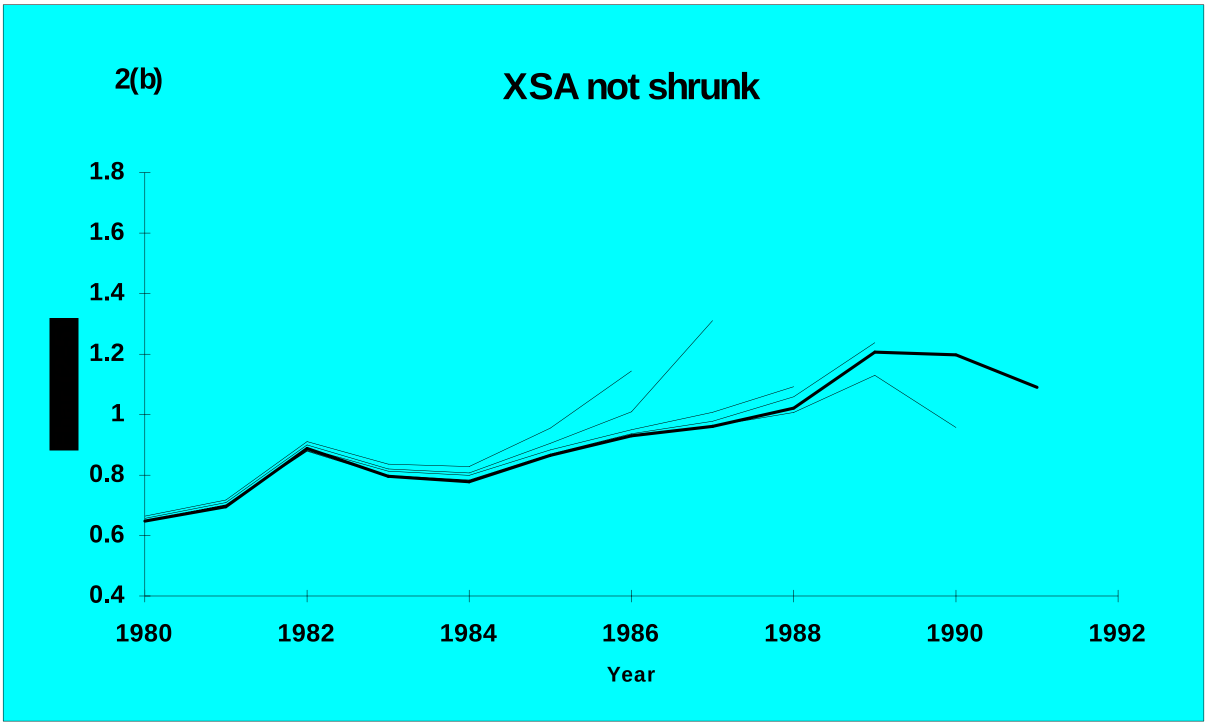
The objective of the analysis is to compare the variation between the 'truth', the final assessment in the series (bold lines), and the values estimated for each terminal year by earlier assessments.

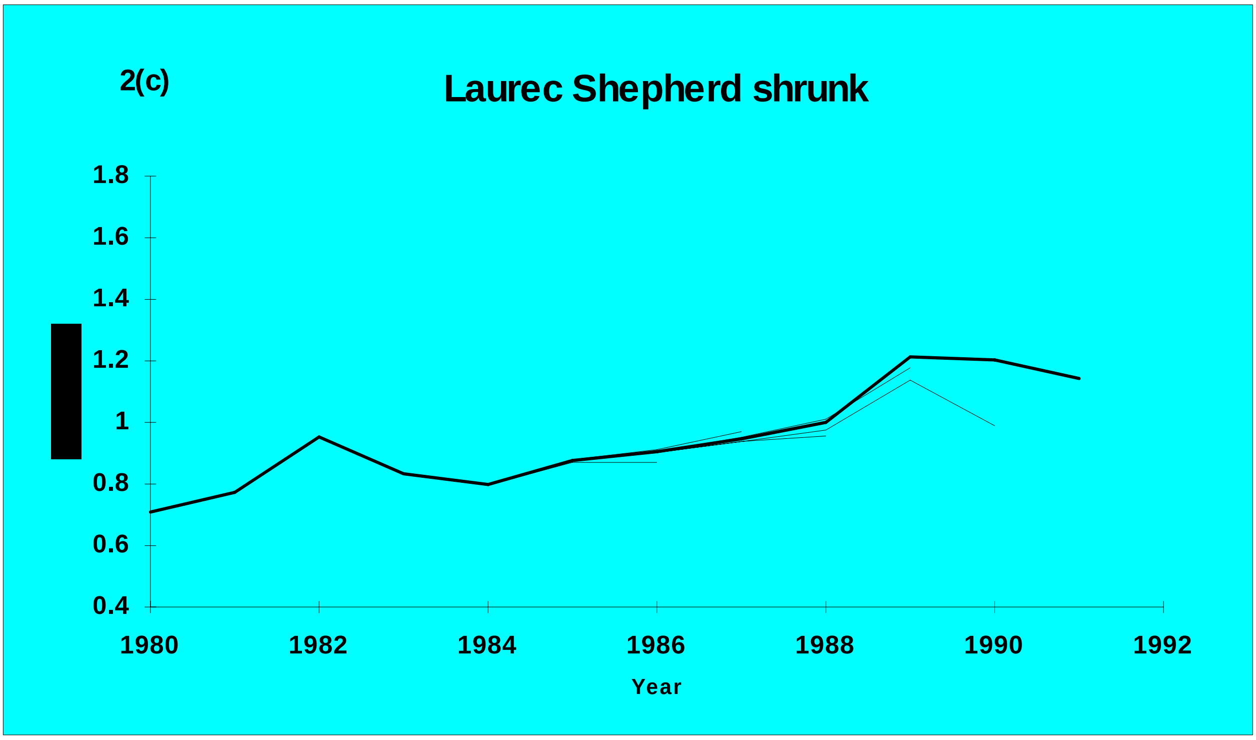
Figure 2 illustrates that, for this stock, shrinkage to the mean improves the prediction accuracy of the final year Fbar values, for both XSA and Laurec-Shepherd tuning. Apart from the penultimate year, both un-shrunk methods have a degree of overestimation bias in the estimated terminal Fbar values. The bias is reduced but reversed by shrinkage to the mean, suggesting that the weight applied to the mean during the shrinkage was too great. It would therefore be concluded that shrinkage to the mean:

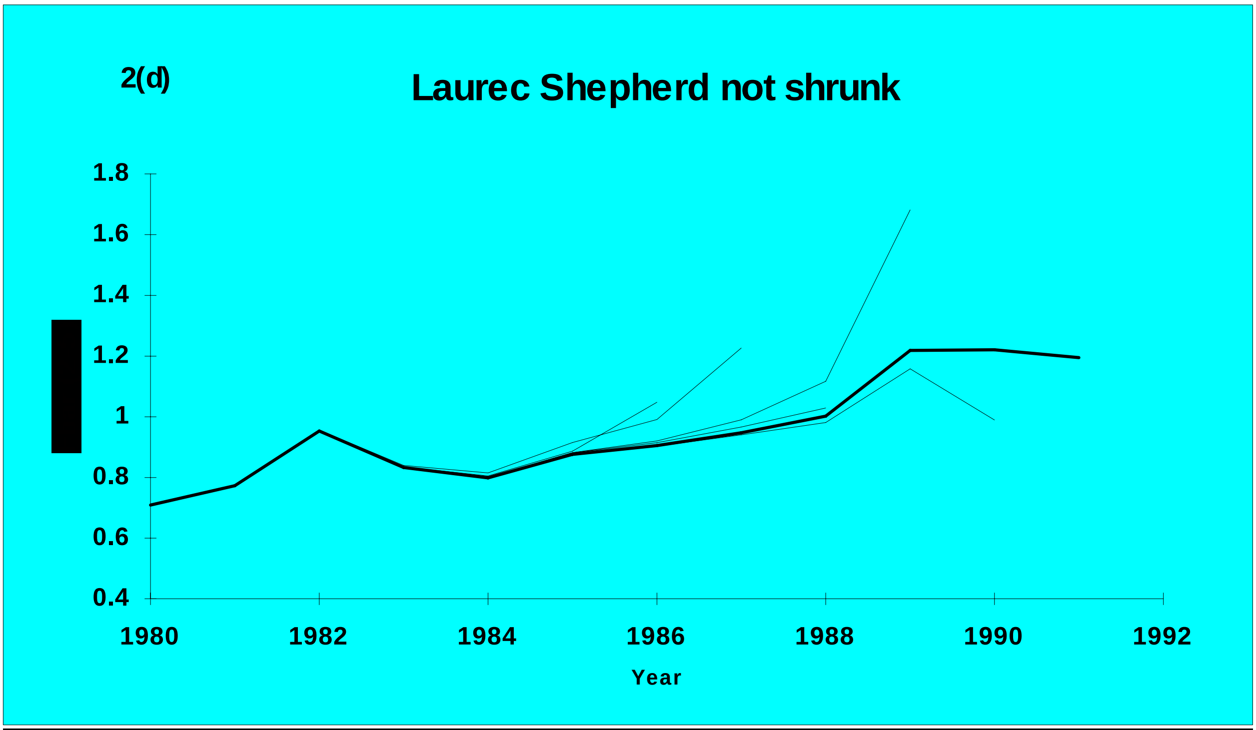
- (i) reduces assessment bias for this stock;
- (ii) improves the accuracy of the Fbar predictions;
- (iii) at the level of shrinkage selected for the retrospective runs, Laurec-Shepherd tuning appears to give the most consistent results.

In practice additional runs would be made with a range of shrinkage weights for each assessment method, to select the most appropriate value. The effects of shrinkage on the population numbers-at-age estimates, especially recruitment, would also be investigated.


2(a)
XSA shrunk
[redacted]
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
1980
1982
1984
1986
1988
1990
1992
Year


2(b)
XSA not shrunk
[redacted]
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
1980
1982
1984
1986
1988
1990
1992
Year


2(c)
Laurec Shepherd shrunk
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
1980
1982
1984
1986
1988
1990
1992
Year


2(d)
Laurec Shepherd not shrunk
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
1980
1982
1984
1986
1988
1990
1992
Year

13. A SUGGESTED PROCEDURE FOR A FULL STOCK ASSESSMENT.

Contents

If fleet catch and effort data are not available:

The choice of assessment procedure is limited to a user-defined VPA using fixed terminal F values, or Separable VPA (the latter being less susceptible to error). The choice as to which solution to take as the final run may be guided by fishery independent information such as SSB estimates from egg surveys or biomass data from acoustic surveys.

If fleet catch and effort data are available:

A Separable VPA, a process which is independent of the fleet CPUE data, can provide a useful test for catch data outliers. Check the residuals for patterns that may indicate changes in exploitation pattern and age/year effects.

Carry out a Laurec-Shepherd *ad hoc* tuning run, without shrinkage, for each of the fleet data sets independently. Plot and examine the catchability residuals. This process is used to 'screen' the fleet data sets, comparisons with the residuals from the Separable VPA run can indicate anomalies in the tuning data. Look for trends in catchability in the most important ages of the dominant (in terms of landings and low standard errors) fleets. Recent and sudden changes in catchability may require removal of the fleet from the assessment.

If there is strong evidence for a trend in catchability, with respect to time, for some of the important ages of the main fleets, the user has two options:

(i) Accept the trend, if it is considered to be a real effect. In this case use the Hybrid model for the final assessment.

(ii) Reject the trend and either, try to remove it by revising the fleet catch and effort data, or remove the fleet from the analysis. Use XSA for the final assessment.

Use the procedural menus described in the appropriate sections, and retrospective analyses (Section 12) for guidance in the selection of the assessment parameter values.

14. REFERENCES

Contents

Aitchinson, J. and J. A. C. Brown, 1957. The lognormal distribution. Cambridge University Press, Cambridge. 176 pp

Anon., 1991(a). Guidance to Assessment Working Groups from ACFM and the ICES Secretariat. International Council for the Exploration of the Sea., 14 pp (mimeo)

Anon., 1991(b). The Working Group on Methods of Fish Stock Assessment. ICES CM,1991/Assess:25, 147 pp (mimeo)

Anon., 1993(a). Report of the Working Group on the Assessment of Northern Shelf Demersal stocks. ICES CM.1993/Assess:2, 449 pp (mimeo)

Anon., 1993(b). Report of the Working Group on the Assessment of Mackerel, Horse Mackerel Sardine and Anchovy. ICES CM.1993/Assess:19, 274 pp (mimeo)

Anon., 1993(c). Report of the Working Group on Methods of Fish Stock Assessment. ICES CM,1993/Assess:12, 86 pp (mimeo)

Armstrong, D. W., 1985. Catchability analysis. ICES C.M. 1985/Assess:9, Appendix 2., 9 pp (mimeo)

Copas, J. B., 1983. Regression, prediction and shrinkage (with discussion). J. R. Statist. Soc. B 45 (3): 311-354

Darby, C. D., and J. G. Shepherd, Combination of recruit indices by weighted averages using RCT3: A users guide. (In prep.)

Doubleday, W. G., 1981. A method for estimating the abundance of survivors of an exploited fish population using commercial catch-at-age and research vessel abundance indices. Can. Spec. Pub. Fish. Aquat. Sci. 58: 164-178.

Flatman, S., 1985. Virtual Population Analysis: Users guide. Part 1: Traditional VPA and Cohort analysis. Int. Rep., MAFF Direct. Fish. Res., Lowestoft, (13), 25 pp.

Gray, D. F., 1977. An iterative derivation of fishing and natural mortality from catch at age data. ICES C.M. 1977/F:33, 13pp (mimeo)

Pope, J. G., 1972. An Investigation of the Accuracy of Virtual Population Analysis Using Cohort Analysis. ICNAF Research Bulletin No. 9. 10 pp

Pope, J. G., 1977. Estimation of fishing mortality, its precision and implications for the management of fisheries. *In* Fisheries mathematics, pp. 63-76. Ed. J.H.Steele. Academic Press, London, New York. 198 pp

Pope, J. G., 1979. Population dynamics and management: current status and future trends. Invest. Pesq., 43: 199-221.

Pope, J. G., and J. G. Shepherd, 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. Int. Explor. Mer, 40: 176-184.

Pope, J. G., and J. G. Shepherd, 1985. A comparison of the performance of various methods for tuning VPA's using effort data. J. Cons. Int. Explor. Mer., 42:129-51.

Rosenberg, A. A., G. P. Kirkwood, R. M. Cook, R. A. Myers, 1992. Combining information from commercial catches and research surveys to estimate recruitment: a comparison of methods. ICES, Mar. Sci. 49: 379-387.

Seber, G. A. F., 1982. The Estimation of Animal Abundance and Related Parameters. London, Griffin, 506pp.

Shepherd, J. G., 1983. Two measures of overall fishing mortality. J. Cons. Int. Explor. Mer, 41: 76-80.

Shepherd, J. G., 1992. Extended survivors' analysis: an improved method for the analysis of catch-at-age data and catch-per-unit-effort data. Working paper No.11 ICES Multispecies Assessment Working Group, June 1992, Copenhagen, Denmark. 22pp (mimeo)

ISSN: 1054-3139

Shepherd, J. G. 1999. Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices ICES Journal of Marine Science Vol. 56, No. 5, October 1999 pp. 584-591

Shepherd, J. G., 1993. Prediction of yearclass strength by calibration regression analysis of multiple recruit index series. ICES, Mar. Sci., In prep.

Shepherd, J. G., and S. M. Stevens, 1983. Separable VPA: User's guide. Int. Rep., MAFF Direct. Fish. Res., Lowestoft, (8): 13pp

Sinclair, A., D. Gascon, R. O'Boyle, D. Rivard and S. Gavaris, 1990. Consistency of some Northwest Atlantic groundfish stock assessments. NAFO SCR Doc. 90/96., 26pp (mimeo)

Stevens, S. M., 1984. A method for weighting residuals in Separable VPA. ICES CM 1984/D:4, 4pp (mimeo)

Topping, J. 1978. Errors of observation and their treatment. Chapman and Hall Ltd. London. 119pp.

APPENDICES

Appendix notation list

a An index denoting the age of the parameter.

y An index for the year in which the parameter is recorded or estimated.

f An index for fleets.

t An index denoting age a in year y; t+1 denotes age a+1 in year y+1.

e^x the exponential of x

Parameters indexed with t (e.g. N_t) are a general reference to any cohort. Those indexed with a and y (e.g. N(y,a)) refer to specific ages and years.

N The number of individuals at the beginning of a time period.

C The catch of individuals removed from the stock during a specific time period.

F The instantaneous rate of fishing mortality.

M The instantaneous rate of natural mortality. Usually assumed constant with respect to age and time.

Z The instantaneous rate of total mortality. Z = F + M.

U Catch per unit effort.

E(y,f) Effort exerted by a fleet in a particular year.

q Catchability.

Fo(y) The overall fully exploited instantaneous fishing mortality in year y.

S(a) The selection-at-age. A parameter relating the instantaneous fishing mortality at age a (F(y,a)) to the fully exploited instantaneous fishing mortality estimated for the year Fo(y). (F(y,a) = Fo(y) S(a))

Appendix 1. The Data Format Identifier (DFI)

Contents

This parameter informs the program of the data format in a file. Valid parameters and their respective data formats are:

DFI		Expected data format
1	i.e. an array of ages → by year ↓.	data data data .... data .... data ....
2	i.e. a row vector for data by age.	data data data ....
3	i.e. a scalar value.	data
4	No data expected.	
5	i.e. a column vector for data by year.	data data data :

DFI 4 is for the special case where catch weight-at-age data are also used as stock weights at age. Spaces or commas are standard separators.

The table below defines the valid use of each DFI within the data files.

Data file#	DFI				
	1	2	3	4	5
	Array	Row	Scalar	None	Column
1 (Landings)					√
2 (Catch numbers)	√				
3 (Catch weights)	√	√			
4 (Stock weights)	√	√		√	
5 (Natural mortality)	√	√	√		√
6 (Maturity proportion)	√	√			
7 (Prop. of F before sp.)	√	√	√		√
8 (Prop. of M before sp.)	√	√	√		√
9 (F on oldest age)					√
10 (F at age)		√			

e.g. DFI 3 may only be used in data files #5, #7 and #8.

Failure to observe these requirements could result in either a program error message or a runtime error message being displayed on the terminal. In both cases the program will stop.

Appendix 2. The defaults for the F and stock number output table means Contents

Firstage - the first assessment age; defined in the input data files.
Lastage - the last assessment age; defined during a run.
Firstyear - the first assessment year; defined during a run.
Lastyear - the last assessment year; defined during a run.

(a) The fishing mortality-at-age table.

Column means - 1 unweighted arithmetic mean (Fbar). If the number of assessment ages selected by the user is greater than 5, the age range is (firstage +2) to (lastage -2), otherwise the average is calculated over all ages.

Fbar is generally used by working groups to describe the overall level of fishing mortality in each year.

Row means - 1 unweighted arithmetic mean. Year range: (lastyear -2) .. lastyear.

This average provides the exploitation pattern used in catch prediction programs. Working groups usually scale the values to provide the same mean F (usually Fbar) as that in the most recent year.

(b) The stock numbers-at-age table.

Row means - 1 unweighted geometric mean. Year range: firstyear .. (lastyear -2).
1 unweighted arithmetic mean. Year range: firstyear .. (lastyear -2).

Averages are calculated over the range firstyear .. (lastyear -2) in order to utilise the years for which convergence of the VPA equations has produced well-conditioned values.

Appendix 3. Estimation of fishing mortality levels consistent with specified exploitation patterns (J.G.Shepherd. 10/2/83)

Contents

Given a specified exploitation pattern (relative F at age) S_j (such as that determined by Separable VPA) one may wish to estimate the overall level of fishing mortality which, combined with that exploitation pattern, best represents any vector F_j of fishing mortality at age. This problem commonly occurs when one has determined values of fishing mortality at age for a recent year, which are somewhat 'rough' because they include the effects of sampling errors etc., and wishes to compare these with another set based on a smoothed exploitation pattern. When the exploitation pattern is unknown, or known to have changed, the catch-weighted measure $\overline{F}_c$ (Shepherd, 1983) may be used. When the pattern can be specified, however, and no systematic changes are believed to have occurred, one can avoid the possible bias of that procedure, in the following way.

Denote the overall fishing mortality by $\overline{F}_s$. If the exploitation pattern S_j applied exactly, one would have fishing mortality-at-age given by

$$F_j^{1} = \overline{F}_s S_j \qquad (1)$$

However the actual vector of fishing mortality-at-age is F_j, which is different. Choose $\overline{F}_s$ in such a way that the sum of squares of deviations between F_j^1 and F_j,

$$T = \sum_j \left(F_j - F_j^{1}\right)^2 = \sum_j \left(F_j - \overline{F}_s S_j\right)^2 \qquad (2)$$

is minimised. This occurs when $dT / d\overline{F}_s = 0$, i.e. when

$$\sum_j S_j \, (F_j - \overline{F}_s S_j)^2 = 0 \qquad (3)$$

and therefore when

$$\overline{F}_s = \sum_j F_j S_j \Big/ \sum_j S_j^2 \qquad (4)$$

This is the required definition of $\overline{F}_s$. It is important to note that $\overline{F}_s$ is conditional on the specified selection pattern (hence the notation adopted for it). If any of the S_j (or even the normalisation adopted for them) is changed, the value of $\overline{F}_s$ will change also. This is obvious from equation (4), which also shows that if all the S_j are multiplied by the same factor, the value of $\overline{F}_s$ will be divided by that same factor, as would be expected.

Appendix 4. The standard error of the logarithm of a variable

Contents

The majority of commonly used tuning methodologies (e.g. the *ad hoc* procedures and XSA) use calibration procedures that are based on relationships between log-transformed quantities. During the interpretation of the output from the tuning procedures, or when estimating values for the weight to be applied to a shrinkage mean, it should be remembered that the standard error of a log transformed quantity is approximately equivalent to the fractional coefficient of variation of the untransformed values. If x is a normally distributed value, $N(\mu_x,\sigma_x^2)$, and $Y = f(x)$, then the expected value of y is given by $E[y] = f(\mu_x)$, and the variance of y can be estimated from :

$$V[y] \cong V[x].\left(\frac{df(\mu_x)}{dx}\right)^2 \qquad \text{Seber (1982)}$$

when y = Log (x)

$$V[y] \cong \sigma_x^2.\,(1/\mu_x)^2$$

and therefore

$$\text{s.e.}[\text{Log}(x)] \cong \sigma_x / \mu_x = \text{c.v.}[x]$$

This approximation only holds for C.V. below 0.7 (Aitchinson and Brown. 1957)

Appendix 5. VPA/Cohort analysis algorithms

Contents

A5.1 Cohort analysis

This option uses Pope's approximation (Pope, 1972) to estimate the cohort population abundances :

$$N_t = N_{(t+1)}\, e^{M} + C_t\, e^{M/2} \qquad (1)$$

and for the corresponding total and fishing mortalities:

$$Z_t = \mathrm{Ln}\,\frac{N_t}{N_{(t+1)}} \qquad (2)$$

$$F_t = Z_t - M_t \qquad (3)$$

A5.2 Virtual Population Analysis (VPA)

The VPA catch equation is

$$C_t = \frac{F_t}{Z_t} N_t \left(1 - e^{-Z_t}\right) \qquad (4)$$

This can be modified to give the catch in year t in terms of the population abundance at the beginning of the following year ($N_{(t+1)}$)

$$C_t = \frac{F_t}{Z_t} N_{(t+1)}\, e^{Z_t} \left(1 - e^{-Z_t}\right) \qquad (5)$$

From which the VPA equation is derived

$$\frac{C_t}{N_{(t+1)}} = \frac{F_t}{Z_t}\, e^{Z_t} \left(1 - e^{-Z_t}\right) \qquad (6)$$

Given C_t and N_{t+1}, the problem is to find F_t that solves the equation. This can be conveniently done by writing the equation as

$$Y = e^{-Z_t} \frac{C_t}{N_{(t+1)}} - \frac{F_t}{Z_t} \left(1 - e^{-Z_t}\right) \qquad (7)$$

and finding the value of F_t for which Y is zero.

First estimates of F_t are obtained from Cohort analysis and refined using an approximation to the Newton-Raphson method which finds an iterative solution for a variable *x* such that *f*(*x*) = 0. This gives fresh estimates of F_t as

$$F_{t\,(new)} = F_{t\,(old)} - \frac{Y}{dY / dF_{t(old)}} \qquad (8)$$

Pope (1972) has shown that within the range M < 0.3 and F < 1.2

$$\frac{Z\,(1-e^{-F})}{F\,(1-e^{-Z})} \approx e^{M/2} \qquad (9)$$

Re-arranging and substituting (9) into (7) gives

$$Y = e^{-Z_t}\frac{C_t}{N_{(t+1)}} - (1-e^{-F_t})\,e^{-M/2} \qquad (10)$$

therefore

$$\frac{dY}{dF_t} = -e^{-Z_t}\frac{C_t}{N_{(t+1)}} - e^{-F_t}\,e^{-M/2}$$

$$= -e^{-F_t}\left[\frac{e^{-M}C_t}{N_{(t+1)}} + e^{-M/2}\right]$$

simplifying using $e^{3M/2}$ gives

$$\frac{dY}{dF_t} = -e^{-F_t}\,e^{-3M/2}\left[\frac{e^{M/2}C_t + e^{M}N_{(t+1)}}{N_{(t+1)}}\right]$$

and

$$\frac{dY}{dF_t} = -e^{-F_t}\,e^{-3M/2}\left[\frac{N_t}{N_{(t+1)}}\right] \qquad (11)$$

Where N_t is the Cohort approximation for N_t. Iteration using equations (7), (8) and (11) continues until | Y | < 0.0001. This gives estimates of F_t which satisfy the VPA equation with a precision of $\pm$ 0.0001 $e^{M/2}$.

Appendix 6. Separable VPA. Contents

A 6.1 The Separable VPA equations

The development of the Separable VPA methodology is described in Pope (1977, 1979) and Pope and Shepherd (1982). The algorithm used in this program is that developed by Pope and Shepherd (1982) and implemented in a program by Shepherd and Stevens (1983).

The model assumes that fishing mortality-at-age in any year is the product of a year effect, the overall fully exploited fishing mortality in a year (Fo(y)), and an age dependent exploitation pattern effect (S(a)). The fishing mortality-at-age acting in a year is given by

$$F(y,a) = Fo(y)\, S(a) \qquad (1)$$

The catch-at-age data are used to derive a matrix of log catch ratios $D_{(y,a)}$ where

$$D_{(y,a)} = \mathrm{Ln}\,(C_{(y+1,a+1)} / C_{(y,a)}) \qquad (2)$$

This removes the year class effects by normalising the matrix with respect to recruitment.

Given $\hat{F}o(y)$ and $\hat{S}(a)$, the ratios can be modelled in terms of F and M alone.

$$\hat{D}_{(y,a)} = \mathrm{Ln}\left[\frac{F_{(y+1,a+1)} Z_{(y,a)} (1-e^{-Z_{(y+1,a+1)}}) e^{-Z_{(y,a)}}}{F_{(y,a)}\ Z_{(y+1,a+1)} (1-e^{-Z_{(y,a)}})}\right] \qquad (3)$$

Using the approximation given by Gray (1977)

$$\frac{F\,(1-e^{-Z})}{Z} \approx F\, e^{(-Z/2.25)} \qquad (4)$$

equation (3) can be re-written as

$$\hat{D}_{(y,a)} = \mathrm{Ln}\left[\frac{F_{(y+1,a+1)} e^{-Z_{(y+1,a+1)}/2.25} e^{-Z_{(y,a)}}}{F_{(y,a)}\ e^{-Z_{(y,a)}/2.25}}\right] \qquad (5)$$

and simplified to

$$\hat{D}_{(y,a)} = \mathrm{Ln}\left[\frac{F_{(y+1,a+1)}}{F_{(y,a)}}\right] - 0.4444\, F_{(y+1,a+1)} - 0.5556\, F_{(y,a)} - M \qquad (6)$$

The log catch ratios can therefore be described in terms of just the year effects ($\hat{F}o(y)$), age effects ($\hat{S}(a)$) and their interactions.

$$\hat{D}_{(y,a)} = \mathrm{Ln}\left[\frac{\hat{F}o_{(y+1)}}{\hat{F}o_{(y)}}\right] + \mathrm{Ln}\left[\frac{\hat{S}_{(a+1)}}{\hat{S}_{(a)}}\right] - 0.4444\, \hat{F}o_{(y+1)} \hat{S}_{(a+1)} - 0.5556\, \hat{F}o_{(y)} \hat{S}_{(a)} - M \qquad (7)$$

Pope (1977,1979) and Pope and Shepherd (1982) show the similarity between equation (7) and a two way ANOVA.

Pope (1979) demonstrated that the information contained within the catch-at-age matrix is insufficient for a unique definition of the values of Fo(y) and S(a), given a realistic level of sampling error. If a unique solution is to be achieved, the number of estimated parameters must be reduced. In the current approach the user defines M, the fishing mortality acting on a reference age in the terminal year and the selection value for the oldest independent age in the data set.

A6.2 The Separable VPA algorithm

If R(y,a) is the residual between the estimated log catch ratio and log catch ratio estimated from the data set,

$$\mathrm{R(y,a)} \;=\; \mathrm{D(y,a)} \;-\; \hat{\mathrm{D}}\mathrm{(y,a)} \qquad (8)$$

Then the goal of the fitting algorithm is the derivation of estimates for Fo(y) and S(a) that minimise the row and column residual totals i.e.

$$\mathrm{R(y,.)} = \sum\nolimits_{i=Firstage}^{i=Lastage-1} \mathrm{R}(\mathrm{y},i) = 0 \;\text{ for each year in the data matrix}$$

and $$\mathrm{R(.,a)} = \sum\nolimits_{i=Firstyear}^{i=Lastyear-1} \mathrm{R}(i,\mathrm{a}) = 0 \text{ for each age in the matrix.}$$

The row and column totals are derived by combining equations (2), (7) and (8). Pope and Shepherd (1982) show that if the age range for the catch-at-age array is 1.. a, and y is the numbers of years, the <u>column</u> totals are given by

$$\mathrm{R(y,.)} = (\mathrm{a}-1)\,\mathrm{Ln}\left[\frac{\hat{\mathrm{F}}\mathrm{o}_{(\mathrm{y}+1)}}{\hat{\mathrm{F}}\mathrm{o}_{(\mathrm{y})}}\right] + \mathrm{Ln}\left[\frac{\hat{\mathrm{S}}_{(\mathrm{a})}}{\hat{\mathrm{S}}_{(1)}}\right] - 0.4444\hat{\mathrm{F}}\mathrm{o}_{(\mathrm{y}+1)}\sum_{\mathrm{i}=1}^{\mathrm{a}-1}\hat{\mathrm{S}}_{(\mathrm{i}+1)}$$
$$-\,0.5556\,\hat{\mathrm{F}}\mathrm{o}_{(\mathrm{y})}\sum_{\mathrm{i}=1}^{\mathrm{a}-1}\hat{\mathrm{S}}_{(\mathrm{i})} - (\mathrm{a}-1)\mathrm{M} - \sum_{\mathrm{i}=1}^{\mathrm{a}-1}\mathrm{Ln}\left[\mathrm{C}_{(\mathrm{y}+1,\,\mathrm{i}+1)}/\mathrm{C}_{(\mathrm{y},\,\mathrm{i})}\right] \qquad (9)$$

The algorithm used for estimating values of $\hat{\mathrm{F}}\mathrm{o(y)}$ that achieve zero column residual totals is based on the approximation that, in equation (9), the dependence on all variables except $\hat{\mathrm{F}}\mathrm{o(y)}$ can be ignored. Trials with dependence on other variables ($\hat{\mathrm{F}}\mathrm{o(y+1)}$) produced no improvement in the iterative procedure (Pope and Shepherd 1982). This allows the Newton-Raphson approximation to be used to find an iterative solution for the value of $\hat{\mathrm{F}}\mathrm{o(y)}$ that results in $\sum_{i=Firstage}^{i=Lastage-1}\mathrm{R}(\mathrm{y},i) = 0$. Each new estimate of $\hat{\mathrm{F}}\mathrm{o(y)}$, $\hat{\mathrm{F}}\mathrm{o(y)}_{[\mathrm{new}]}$ is given by

$$\hat{\mathrm{F}}\mathrm{o(y)}_{[\mathrm{new}]} \;=\; \hat{\mathrm{F}}\mathrm{o(y)}_{[\mathrm{old}]} \;-\; \mathrm{R(y,.)}\Big/\frac{d\mathrm{R(y,.)}}{d\hat{\mathrm{F}}\mathrm{o(y)}} \qquad (10)$$

where

$$\frac{d\,\mathrm{R(y,.)}}{d\hat{\mathrm{F}}\mathrm{o(y)}} = -\frac{(\mathrm{a}-1)}{\hat{\mathrm{F}}\mathrm{o(y)}} - \frac{(\mathrm{a}-1)}{2} \qquad (11)$$

and therefore

$$\hat{F}o(y)_{[new]} = \hat{F}o(y)_{[old]}\left[1+\frac{R(y,.)/(a-1)}{1+\hat{F}o(y)/2}\right] \qquad (12)$$

The algorithm developed by Pope and Shepherd (1982) ignored the correction term 1+Fo(y)/2 and replaced (a-1) by a to introduce an element of 'under-relaxation'. Subsequent testing has revealed that a more robust approach to the solution is achieved by increasing the degree of under-relaxation and using 2a. Also, since Fo(y) must be positive, the algorithm uses the approximation

$$\hat{F}o(y)[new] = \hat{F}o(y)[old]\; e^{\,(R(y,.)/2a)}$$

Similar derivations can be applied to obtain iterative estimates for the row totals, giving

$$\hat{S}(a)[new] = \hat{S}(a)[old]\; e^{\,(R(.,a)/2y)}$$

A6.3 Population numbers

After estimation of the $\hat{F}o(y)$ and $\hat{S}(a)$ values the program estimates the recruitment to each cohort in the time series. The appropriate value is that which minimises the sum of squares of the log catch residuals over the whole cohort, conditional on the estimated values of $\hat{F}o(y)$ and $\hat{S}(a)$.

$$SS = \sum_{k=kmin}^{kmax}\left[Ln\, C_{(i+k-kmin,\,k)} - LnN_{(i,\,kmin)}H_{(i+k-kmin,\,k)}\right]^2$$

where

$$H_{(i+k-kmin,\,k)} = \frac{F_{(i+k-kmin,\,k)}}{Z_{(i+k-kmin,\,k)}}\left(1-e^{-Z_{(i+k-kmin,\,k)}}\right) e^{-\sum_{j=kmin}^{k-1} Z_{(i+j-kmin,k)}}$$

i is the year and k is the kth age in the cohort, kmin the first, and kmax the oldest cohort age. SS is at a minimum when

$$\frac{dSS}{dN(i,\,kmin)} = 0$$

i.e. when

$$Ln\, N_{(i,\,kmin)} = \frac{1}{kmax+1-kmin}\sum_{k=kmin}^{kmax}\left[LnC_{(i+k-kmin,\,k)} - Ln\, H_{(i+k-kmin,\,k)}\right]$$

A6.4 Year and age weights

A modification to the algorithm implemented by Shepherd and Stevens (1983) was introduced to enable the use of specific year and age weights (Stevens 1984).

Year weights allow the period during which selection-at-age is considered to have been constant to be a subset of the assessment time series. The year weights are user-defined ($0.001 \le Y(y) \le 1.0$).

Age weights permit the down-weighting of residuals for ages which exhibit a poor fit within the model. The age weights (A(a)) are either user-defined or calculated by the program as the reciprocal of the root mean square error about the log catch ratio model for each of the (number of ages-1) matrix rows.

$$A_{raw}(a) = \sqrt{\frac{y-2}{\sum_{i=1}^{y-1} (\hat{D}(i,a) - D(i,a))^2 - \frac{[\sum_{i=1}^{y-1} (\hat{D}(i,a) - D(i,a))]^2}{y-1}}}$$

The weights are normalised to the largest reciprocal.

$$A(a) = A_{raw}(a)/A_{maximum}$$

Weights are assigned during the calculation of row and column totals such that the algorithm seeks

$$\sum_{i=Firstage}^{i=Lastage-1} A(i)\ R(y,i) = 0 \quad \text{for each year in the data matrix}$$

and $$\sum_{i=Firstyear}^{i=Lastyear-1} Y(i)\ R(i,a) = 0 \quad \text{for each age in the matrix.}$$

Weighted estimates of recruitment are derived from

$$\mathrm{Ln}\ N_{(i,\ kmin)} = \frac{1}{W} \frac{1}{kmax+1-kmin} \sum_{k=kmin}^{kmax} A_{(k)} Y_{(k)} [\mathrm{LnC}_{(i+k-kmin,\ k)} - \mathrm{Ln}\ H_{(i+k-kmin,\ k)}]$$

$$W = \sum_{k=kmin}^{kmax} A_{(k)} Y(i+k-kmin)$$

Appendix 7. The *ad hoc* tuning algorithms.

Contents

In this program, final year F-at-age values can be derived by fitting one of two models for fleet catchability-at-age. The Laurec-Shepherd procedure assumes constant fleet catchability-at-age with respect to time and uses the mean for the predicted value (a geometric mean of q). The Hybrid procedure allows the user to apply Laurec-Shepherd tuning to some of the fleets (preferably as many as possible), and to let the catchability for the remaining fleets vary exponentially with time.

A7.1 Partial F-at-age

At each iteration a VPA is performed on the catch-at-age data. The assumption is then made that, at each age, an estimate of the fishing mortality contributed by a fleet (the fleet's partial F-at-age - $Fp_{(y,a,f)}$), can be obtained from the VPA estimate of total international fishing mortality-at-age $F_{(y,a)}$ and the ratio of the fleet catch-at-age in the year to that of the total international catch ($Ct_{(y,a)}$) :

$$Fp_{(y,a,f)} = F_{(y,a)} \frac{C_{(y,a,f)}}{Ct_{(y,a)}} \qquad (1)$$

It follows that the fleet's catchability-at-age in year y is :

$$q_{(y,a,f)} = \frac{Fp_{(y,a,f)}}{E_{(y,f)}} \qquad (2)$$

Fleet catchabilities-at-age are assumed to have log normally distributed errors.

A7.2 Estimation of fleet catchability in the final year

The fitting procedure used to derive fleet catchability in the final year is a weighted least squares predictive regression of the log catchabilities against time. In VPA version 3.1 the regression procedure treats zero values as missing data.

The Laurec-Shepherd algorithm assumes that a fleet's catchability-at-age is constant with respect to time and log normally distributed. For each age, the time series weighted geometric mean (taken from the regression equations) is used for the fleet's catchability in the final year. By definition, the Laurec-Shepherd procedure makes the assumption that there is no relationship between log catchability and time. The regression statistics, including the slope and its standard error, are printed in the tuning output to permit tests for departures from this assumption.

The Hybrid method assumes that catchability-at-age <u>for all ages</u> of particular fleets varies exponentially with time. The remaining fleets are fitted with the constant catchability model. For each age in the selected fleets, the regression fits a linear log catchability trend with time and uses the slope and intercept to estimate a value for the final year.

For both methods the standard errors of the predicted fleet log catchabilities-at-age are given by

$$\sigma_{(ty,a,f)} = \sqrt{\frac{\sum_{i=1}^{i=ty} \omega_{(i,a,f)} [\mathrm{Ln}(q_{(i,a,f)}) - \overline{\mathrm{Ln}(q_{(a,f)})}]^2}{n_{(a,f)} - 2}} \sqrt{1 + \frac{1}{\sum_{i=1}^{i=ty} \omega_{(i,a,f)}} + \frac{(ty - \bar{i})^2}{\sum_{i=1}^{i=ty} \omega_{(i,a,f)} (i - \bar{i})^2}} \qquad (3)$$

Where i is a year index, ty the index value for the final year, $n_{(a,f)}$ is the number of non-zero total annual catches of fish aged a removed by fleet f during the tuning time series, and the $\omega_{(y,a,f)}$ values are weights.

$$\omega_{(y,a,f)} = \theta_{(y,a,f)}\ \tau_{(y)}$$

where $\tau(y)$ are the time series weights defined in Appendix 9, and $\theta_{(y,a,f)}$ are the missing value weights defined as :

$\theta_{(y,a,f)} = 1.0$ - If a catch from fleet f <u>and</u> a total international catch value were both <u>recorded</u> at age a in year y.

$\theta_{(y,a,f)} = 0.0$ - If either a catch from fleet f <u>or</u> the total international catch value were zero <u>(missing)</u> at age a in year y.

The first term on the r.h.s of equation 3 is the standard error about the prediction regression, the second is the error introduced by the distance of the terminal year from the mean year. For the Laurec-Shepherd constant catchability estimates the second term reduces to

$$\sqrt{1 + \frac{1}{\sum \omega_{(y,a,f)}}}$$

A7.3 Fleet based estimates of overall F

When the final year (ty) fleet catchabilities-at-age have been estimated, final year partial F values (F_p) for each fleet are calculated as

$$Fp_{(ty,a,f)} = q_{(ty,a,f)}\ E_{(ty,f)} \qquad (4)$$

and final year estimates of total international F ('raised F's') calculated using the fleet catch-at-age and the total international catch-at-age

$$\hat{F}_{(ty,a,f)} = Fp_{(ty,a,f)} \frac{Ct_{(ty,a)}}{C_{(ty,a,f)}} \qquad (5)$$

The terminal F for the next iteration is derived as a weighted geometric mean of the fleet-based estimates. The weights are the inverse of the variances of the final year log raised F's. Fleet effort for the final year is assumed to be exact so the variance of a log raised F estimate for each age is equal to that of the log catchability ($\sigma_{(ty,a,f)}{}^2$).

$$\mathrm{Ln}(F_{(ty,a)}) = \frac{\sum_{f=1}^{f=\text{No Fleets}} \frac{1}{\sigma_{(ty,a,f)}{}^2} \mathrm{Ln}(\hat{F}_{(ty,a,f)})}{\sum_{f=1}^{f=\text{No Fleets}} \frac{1}{\sigma_{(ty,a,f)}{}^2}} \qquad (6)$$

A7.4 Shrinkage to the mean

If the user has selected the shrinkage to the mean option, the mean F for the age over the 5 years preceding the terminal year ($\overline{F}_{((ty-6,\,ty-1),\,a)}$) is added to equation (6) with the user-defined weight ($\sigma_{(shrse)}$).

$$\mathrm{Ln}(F_{(ty,a)}) = \frac{\sum_{f=1}^{f=\text{No Fleets}} \left[\frac{1}{\sigma_{(ty,a,f)}{}^2} \mathrm{Ln}(\hat{F}_{(ty,a,f)})\right] + \frac{1}{\sigma_{(shrse)}{}^2} \mathrm{Ln}(\overline{F}_{((ty-6,\,ty-1),\,a)})}{\sum_{f=1}^{f=\text{No Fleets}} \left[\frac{1}{\sigma_{(ty,a,f)}{}^2}\right] + \frac{1}{\sigma_{(shrse)}{}^2}} \qquad (7)$$

Once the F-at-age values in the final year have been estimated, the program performs a VPA for each of the cohorts (except the oldest) terminating in that year. The oldest age F is either user-defined or derived as a proportion of an average F of ages preceding it. The program then steps backwards through the time series performing a VPA on each cohort using either the user-defined terminal F, or backwards extension (Section 9.2.2).

A7.5 Standard errors for the overall F estimates

Assuming that the shrinkage option has been selected, the internal standard error ($\sigma_{\mathrm{int(ty,\,a)}}$) of terminal F for age a is derived from

$$\sigma_{\mathrm{int(ty,\,a)}} = \sqrt{\sum_{f=1}^{f=\mathrm{No\ Fleets}} \left[\frac{1}{\sigma_{(\mathrm{ty,\,a,\,f})}{}^{2}} \right] + \frac{1}{\sigma_{(\mathrm{shrse})}{}^{2}}} \qquad (8)$$

and the external standard error ($\sigma_{\mathrm{ext(ty,\,a)}}$) from

$$\sigma_{\mathrm{ext(ty,\,a)}} = \sqrt{\frac{1}{(n(a)-1)}} \sqrt{\frac{\left[\dfrac{\sum_{f=1}^{f=\mathrm{No\ Fleets}} \left[\mathrm{Ln}(\hat{\mathrm{F}}_{(\mathrm{ty,\,a,\,f})}) - \mathrm{Ln}(\mathrm{F}_{(\mathrm{ty,\,a})}) \right]^{2}}{\sigma(\mathrm{ty,a,f})^{2}} \right] + \dfrac{\left[\mathrm{Ln}(\overline{\mathrm{F}}_{((\mathrm{ty\,-\,6,\,ty\,-\,1),\,a})}) - \mathrm{Ln}(\mathrm{F}_{(\mathrm{ty,\,a})}) \right]^{2}}{\sigma_{(\mathrm{shrse})}{}^{2}}}{\sum_{f=1}^{f=\mathrm{No\ Fleets}} \left[\dfrac{1}{\sigma(\mathrm{ty,a,f})^{2}} \right] + \dfrac{1}{\sigma_{(\mathrm{shrse})}{}^{2}}}} \qquad (9)$$

where F(ty,a) is the overall mean F derived from equation (7), and *n*(a) is the number of fleets contributing a raised F estimate.

Appendix 8. The XSA algorithm

Contents

The *ad hoc* tuning algorithms use the relationship between fleet partial F and fleet effort to derive estimates of fleet catchability-at-age and subsequently overall F in the final year. XSA utilises the relationship between catch per unit effort (CPUE) and population abundance. Shepherd (1992) describes the least squares derivation of the XSA equations.

A8.1 Transformation of fleet CPUE

CPUE (U) indices from each fleet are assumed to be related to population abundance by the constant catchability model

$$U_{(y,a,f)} = q_{(a,f)} A_{(y,a,f)} N_{(y,a)} \qquad (1)$$

A is an averaging factor relating the population abundance during the time at which the catch was taken to the population abundance at the beginning of the year ($N_{(y,a)}$). A is given by :

$$A_{(y,a,f)} = \frac{\left(e^{-\alpha_{(f)} Z_{(y,a)}} - e^{-\beta_{(f)} Z_{(y,a)}}\right)}{\left(\beta_{(f)} - \alpha_{(f)}\right) Z_{(y,a)}} \qquad (2)$$

where α is the start of the period of fishing and β the end for each fleet (both expressed as fractions of a year). In practice XSA adjusts the CPUE to the beginning of the year.

$$U'_{(y,a,f)} = \frac{U_{(y,a,f)}}{A_{(y,a,f)}} \qquad (3)$$

Estimates of population abundance-at-age at the beginning of a year are derived for each fleet by

$$N_{est(y,a,f)} = \frac{U'_{(y,a,f)}}{q_{est(a,f)}} \qquad (4)$$

A8.2 The underlying Cohort analysis model

The method assumes that the population abundances derived from the VPA (N_{vpa}) are exact and that the values of N_{est} are estimates of these values. Shepherd (1992) shows that by a rearrangement of Pope's Cohort analysis equation, the VPA population abundance for age a in year y is given by

$$N_{vpa\,(y,a)} = ECM_{(y,a)}\, P_{t(Ty,Ta)} + P_{c\,(y,a)} \qquad (5)$$

$P_{t(Ty,Ta)}$ is the terminal population (survivors at the <u>end of the year</u>) at the oldest cohort age. ECM is the exponential cumulative natural mortality, the sum of the natural mortality values for all subsequent ages within the cohort.

$$ECM_{(y,a)} = e^{\left(M_{(y,a)} + M_{(y+1,a+1)} + M_{(y+2,a+2)} \ldots\ldots + M_{(Ty,Ta)}\right)}$$

Ty is the terminal year for the cohort, either the year in which it reaches the oldest age, or the final year of the assessment. Ta is the terminal age for the cohort, either the oldest age in the assessment or the cohort's age in the final year (Fa will be used to refer to the youngest cohort age). Similar definitions apply to ECF and ECZ, used later.

In the following text $\sum_{i=a}^{i=Ta}$ is used to describe the summation of all cohort ages from age a to Ta i.e.

$$\sum_{i=a}^{i=Ta} X(y,i) = X(y,a) + X(y+1,a+1) + X(y+2,a+2)\ldots\ldots\ldots X(Ty,Ta)$$

the increment of y has been omitted for clarity. Thus, in equation (5) Pc is given by

$$Pc_{(y,a)} = \sum_{i=a}^{i=Ta} [\ ECM_{(y,i)}\ C_{(y,i)}\ e^{-0.5M}\ ] \tag{6}$$

$Pc_{(y,a)}$ therefore describes the contribution of the raised accumulated catches to the population abundance in year y at age a. Since M and the catches are constant between iterations, Pc is constant.

A8.3 Fleet catchability-at-age

Two methods for the calculation of fleet-based estimates of population abundance-at-age ($Nest_{(y,a,f)}$) are available. They differ in the model applied to estimate catchability-at-age. Fleet catchabilities are assumed to be either constant with respect to time (the fully recruited ages of all fleets), or proportional to year class abundance (used for the recruiting ages of all fleets).

The relationship between fleet CPUE and year class abundance for the recruiting ages is derived from the power-law model described by Shepherd (1993).

$$U'_{(y,a,f)} = \varphi\, Nvpa_{(y,a)}{}^{\gamma} \tag{7}$$

φ and γ are assumed to be constant with respect to time. They are usually estimated using a weighted linear calibration regression.

$$\text{Ln } Nvpa_{(y,a)} = \frac{1}{\gamma}\ \text{Ln } U'_{(y,a,f)} - \frac{\varphi}{\gamma} \tag{8}$$

Fleet-based estimates of population numbers-at-age ($Nest_{(y,a,f)}$) are obtained from the fitted relationship during each iteration.

$$\text{Ln } Nest_{(y,a,f)} = \frac{\text{Ln } U'_{(y,a,f)} - \varphi}{\gamma} \tag{9}$$

Standard errors for the recruitment age $Nest_{(y,a,f)}$ values are derived from the equation for the standard error of a value predicted by linear regression :

$$\sigma_{(y,a,f)} = \sqrt{\frac{\sum_{i=1}^{i=ty} \omega_{(i,a,f)} (Nest_{(i,a,f)} - \overline{Nest}_{(a,f)})^2}{n_{(a,f)} - 2}} \sqrt{1 + \frac{1}{\sum_{i=1}^{i=ty} \omega_{(i,a,f)}} + \frac{(U'_{(y,a,f)} - \overline{U'}_{(a,f)})^2}{\sum_{i=1}^{i=ty} \omega_{(i,a,f)} (U'_{(i,a,f)} - \overline{U'}_{(a,f)})^2}} \tag{10}$$

Where i is a year index, ty the index value for the final year, $n_{(a,f)}$ is the number of non-zero data points recorded for fleet f at age a, and the $\omega_{(y,a,f)}$ values are the product of the missing value and time series weights, defined earlier in Appendix 7. The first term on the r.h.s is the standard error about the regression, the second is the error introduced by the distance of the index value ($U'_{(y,a,f)}$) from the mean.

For each fleet age at which catchability is considered to be constant the program derives estimates of the log reciprocal catchability (1/q) as a weighted geometric mean of the time series values.

$$\mathrm{Ln}\left[\frac{1}{q_{est(a,f)}}\right] = \frac{\sum_y \omega_{(y,a,f)} \left[\mathrm{Ln}(N_{vpa(y,a)}) - \mathrm{Ln}(U'_{(y,a,f)})\right]}{\sum_y \omega_{(y,a,f)}} \tag{11}$$

Fleet-based estimates of population numbers-at-age ($N_{est(y,a,f)}$) are obtained from equation (4).

For the ages with constant catchability, the standard errors of the estimated population abundances are assumed to be log normal and equal to those of the fleet's reciprocal catchability-at-age.

$$\sigma_{(y,a,f)} = \sqrt{\frac{\sum_y \omega_{(i,a,f)} \left[\mathrm{Ln}\left[\frac{N_{vpa(y,a)}}{U'_{(y,a,f)}}\right] - \mathrm{Ln}\left[\frac{1}{q_{est(a,f)}}\right]\right]^2}{n-1}} \sqrt{1 + \frac{1}{\sum_{i=1}^{i=ty} \omega_{(i,a,f)}}} \tag{12}$$

As catchability is assumed to be constant w.r.t time $\sigma_{(y,a,f)}$ is constant for all y.

A8.4 Catchability for the oldest assessment age

Estimates of the catchability for the oldest age in an assessment, tuned by the *ad hoc* or XSA procedures, are directly dependent on the terminal population or F values used to initialise the underlying VPA. Catchability at the oldest age is therefore under-determined and cannot be utilised without additional information. Within the *ad hoc* tuning procedures the additional information is obtained by making the assumption that the exploitation pattern on the oldest ages was constant during the assessment time series. F on the oldest age in the final year is estimated as a proportion of an average of the F for preceding ages in the same year. XSA uses an alternative approach by making the assumption that fleet catchability is constant (independent of age) above a certain age. The age (constant for all fleets) is user-defined. For each fleet, the catchability value estimated at the specified age, is used to derive population abundance estimates for all subsequent ages in the fleet data set.

A8.5 The estimation of terminal populations (survivors)

For each cohort, the terminal populations which initialise the next iteration, are derived from :

$$\mathrm{Ln}\ P_{t(Ty,Ta)} = \frac{\sum_f \sum_{i=Fa}^{i=Ta} [\, w'_{(y,i,f)}\ (\mathrm{Ln}\ N_{est(y,i,f)} - \mathrm{Ln}\ ECZ_{(y,i)})]}{\sum_f \sum_{i=Fa}^{i=Ta} [w'_{(y,i,f)}]} \tag{13}$$

where $$w'_{(y,i,f)} = \frac{\theta_{(y,i,f)}\ \tau_{(y)}\ \phi_{(f)}}{\sigma^2_{(y,i,f)}\ ECF_{(y,i)}}$$

$\tau(y)$ are the time series weights described in Appendix 9, $\phi(f)$ are the user defined fleet weights, $ECF_{(y,i)}$ is the exponential cumulative fishing mortality summed from the oldest cohort age back to age i, $\sigma(y,i,f)$ are the standard errors of the $N_{est(y,i,f)}$ values, and $\theta_{(y,a,f)}$ are the missing value weights defined as:

$\theta_{(y,a,f)} = 1.0$ - if a catch from fleet f and a total international catch value were both recorded at age a in year y;

$\theta_{(y,a,f)} = 0.0$ - if either a catch from fleet f or the total international catch value were zero (missing) at age a in year y.

Equation (13) is a weighted geometric mean of the population abundance estimates-at-age reduced by total mortality to the end of the final year. The division by ECF progressively reduces the influence of estimates from the younger ages in the cohort.

A8.6 Shrinkage to the mean

Two forms of shrinkage are available, shrinkage to the mean population and shrinkage to the mean F (applied to all ages). They are both implemented by modifying equation (13).

$$\mathrm{Ln}\,P_{t(Ty,Ta)} = \frac{\sum_f \sum_{i=Fa}^{i=Ta} [\, w'_{(y,i,f)}\,(\mathrm{Ln}\,N_{est(y,i,f)} - \mathrm{Ln}\,ECZ_{(y,i)})] + \frac{\mathrm{Ln}\,\overline{P}t_{Fshr(y,a)}}{\sigma^2_{Fshr}} \left[+ \frac{\mathrm{Ln}\,\overline{P}t_{Pshr(a)}}{\sigma^2_{Pshr(a)}} \right]}{\sum_f \sum_{i=Fa}^{i=Ta} [w'_{(y,i,f)}] + \frac{1}{\sigma^2_{Fshr}} \left[+ \frac{1}{\sigma^2_{Pshr(a)}} \right]} \tag{14}$$

Where $\overline{P}t_{Pshr(a)}$ is the population shrinkage mean and the inverse of its variance ($\sigma^2_{Pshr(a)}$) the weight applied to it. Similarly, $\overline{P}t_{Fshr(j,a)}$ is the F shrinkage mean and the inverse of its variance (σ^2_{Fshr}), the user-defined weight applied to it. The square brackets indicate that only the recruit ages are shrunk to the population means. All ages are shrunk to the mean F's.

Shrinkage to the population mean (P shrinkage)

Shrinkage to the population mean, initially used within RCT3 (Shepherd 1994), can be applied to the terminal population values estimated for the recruit ages. The terminal population estimate for recruit age *a*, at the end of the final year, is assumed to lie within the log normal distribution of the VPA population abundances at age *a*+1 (estimated at the start of a year). The population shrinkage mean, towards which the age *a* terminal population estimate is shrunk ($\overline{P}t_{Pshr(a)}$), is therefore a time series weighted geometric mean of the VPA population abundances of age *a*+1:

$$\mathrm{Ln}\,\overline{P}t_{Pshr(a)} = \frac{\sum_{i=1}^{i=Ty} \tau(i)\,\mathrm{Ln}\,N_{vpa(i,a+1)}}{\sum_{i=1}^{i=Ty} \tau(i)} \tag{15}$$

The weight applied to the mean is the inverse of the variance of the taper weighted log values ($1/\sigma^2_{Pshr(a)}$). If *n* is the number of years in the assessment time series :

$$\sigma_{Pshr(a)} = \sqrt{\frac{\sum_{i=1}^{i=Ty} \tau(i)\left[\mathrm{Ln}\,N_{vpa(i,a+1)} - \mathrm{Ln}\,\overline{P}t_{Pshr(a)}\right]^2}{n-1}} \tag{16}$$

Shrinkage to the mean F (F shrinkage)

Shrinkage to the mean F shrinks the terminal population estimates of all cohorts in the assessment; including those that have previously been shrunk to the population mean. For each age, the program calculates a mean of the F values for *j* years that precede the final assessment year (ty).

$$\overline{F}\mathrm{shr}_{(ty,a)} = \frac{\sum_{i=ty-1}^{i=ty-(j+1)} F_{(i,a)}}{j} \tag{17}$$

The mean F and the total international catch in the final year are used to calculate an estimate for the survivors at the <u>end</u> of the final year.

$$\overline{P}t_{\,F\mathrm{shr}(ty,a)} = \frac{C_{(ty,a)}\,(\overline{F}\mathrm{shr}_{(ty,a)} + M_{(ty,a)})}{\overline{F}\mathrm{shr}_{(ty,a)}\,[e^{(\overline{F}\mathrm{shr}(ty,a)+M(ty,a))} - 1]} \tag{18}$$

The survivors in years prior to the final year are shrunk to a terminal population derived from an average of the F values of the *k* ages that precede the oldest true age (ta).

$$\overline{F}\mathrm{shr}_{(y,ta)} = \frac{\sum_{i=ta-1}^{i=ta-(k+1)} F_{(y,i)}}{k} \tag{19}$$

$$\overline{P}t_{\,F\mathrm{shr}(y,ta)} = \frac{C_{(y,ta)}\,(\overline{F}\mathrm{shr}_{(y,ta)} + M_{(y,ta)})}{\overline{F}\mathrm{shr}_{(y,ta)}\,[e^{(\overline{F}\mathrm{shr}(y,ta)+M(y,ta))} - 1]} \tag{20}$$

The shrinkage CV, applied to all F shrinkage means, is the user-defined value and is equal for all years and ages.

A8.7 The standard errors of the terminal population estimates

The fleet based terminal population estimates

The internal standard error of a fleet-based estimate of the cohort's terminal population is derived from the standard errors of the terminal population estimates obtained at each age in the cohort's history.

$$\sigma_{\mathrm{int}(Ty,\,Ta,\,f)}{}^2 = \frac{\sum_{i=Fa}^{i=Ta}\left(\frac{\theta_{(y,i,f)}\,\tau_{(y)}\,\phi_{(f)}}{\sigma^2_{(y,i,f)}\,\mathrm{ECF}_{(y,i)}{}^2}\right)}{\left(\sum_{i=Fa}^{i=Ta}\frac{\theta_{(y,i,f)}\,\tau_{(y)}\,\phi_{(f)}}{\sigma^2_{(y,i,f)}\,\mathrm{ECF}_{(y,i)}}\right)^2} \tag{21}$$

The external standard error of a fleet-based estimate of the cohort's terminal population is the standard error of the estimates obtained at each cohort age ($Pt_{est(y,a,f)}$).

$$\sigma_{ext(Ty,\,Ta,\,f)} = \sqrt{\frac{1}{n(f)-1}} \sqrt{\frac{\sum_{i=Fa}^{i=Ta} w'_{(y,\,i,\,f)} \left[Pt_{est(y,\,i,\,f)} - \overline{Pt_{(Ty,Ta,f)}} \right]^2}{\sum_{i=Fa}^{i=Ta} w'_{(y,\,i,\,f)}}} \qquad (22)$$

where $Pt_{est(y,\,a,\,f)} = \text{Ln } N_{est(y,a,f)} - \text{Ln } ECZ_{(y,\,a)}$

and $w'_{(y,i,f)} = \dfrac{\theta_{(y,\,i,\,f)}\ \tau_{(y)}\ \phi_{(f)}}{\sigma^2_{(y,\,i,\,f)}\ ECF_{(y,\,i)}}$

$n(f)$ is the number of years in which catches were removed from the cohort by fleet f, and $\overline{Pt_{(Ty,Ta,f)}}$ the fleet's weighted mean estimate of the cohort's terminal population.

The overall mean terminal population estimates

Assuming that both of the shrinkage options have been selected during the estimation of a terminal population at a recruiting age, the internal standard error of the overall terminal population estimate derived from all ages of all fleets is derived from

$$\sigma_{int(Ty,\,Ta,\,f)}{}^2 = \frac{\sum_f \sum_{i=Fa}^{i=Ta} \left(\frac{\theta_{(y,\,i,\,f)}\ \tau_{(y)}\ \phi_{(f)}}{\sigma^2_{(y,\,i,\,f)}\ ECF_{(y,\,i)}{}^2} \right) + \frac{1}{\sigma_{(Fshr)}{}^2} + \frac{1}{\sigma_{(Pshr)}{}^2}}{\left(\sum_f \sum_{i=Fa}^{i=Ta} \frac{\theta_{(y,\,i,\,f)}\ \tau_{(y)}\ \phi_{(f)}}{\sigma^2_{(y,\,i,\,f)}\ ECF_{(y,\,i)}} + \frac{1}{\sigma_{(Fshr)}{}^2} + \frac{1}{\sigma_{(Pshr)}{}^2} \right)^2} \qquad (23)$$

The external standard error of the overall terminal population estimate is given by

$$\sigma_{ext(Ty,\,Ta)} = \sqrt{\frac{1}{\sum_f n_{(f)} + 1}} \sqrt{\frac{\sum_f \sum_{i=Fa}^{i=Ta} w'_{(y,\,i,\,f)} \left[Pt_{est(y,\,i,\,f)} - \overline{Pt_{(Ty,Ta)}} \right]^2 + \frac{1}{\sigma^2_{Fshr}} \left[Pt_{Fshr} - \overline{Pt_{(Ty,Ta)}} \right]^2 + \frac{1}{\sigma^2_{Pshr}} \left[Pt_{Pshr} - \overline{Pt_{(Ty,Ta)}} \right]^2}{\sum_f \sum_{i=Fa}^{i=Ta} w'_{(y,\,i,\,f)} + \frac{1}{\sigma^2_{Fshr}} + \frac{1}{\sigma^2_{Pshr}}}} \qquad (24)$$

$\overline{Pt_{(Ty,Ta)}}$ is the overall weighted mean estimate of the cohort's terminal population.

Appendix 9. Time series weighting

Contents

The *ad hoc* tuning and XSA modules present the user with a range of weighting options that can be applied to the time period selected for the assessment, either :

(i) No weighting of the regressions.
(ii) Weights supplied by the user for each year (*ad hoc* tuning only).
(iii) A weighting function which assigns each year a weight whose value is related to its distance in time from the final year of the assessment.

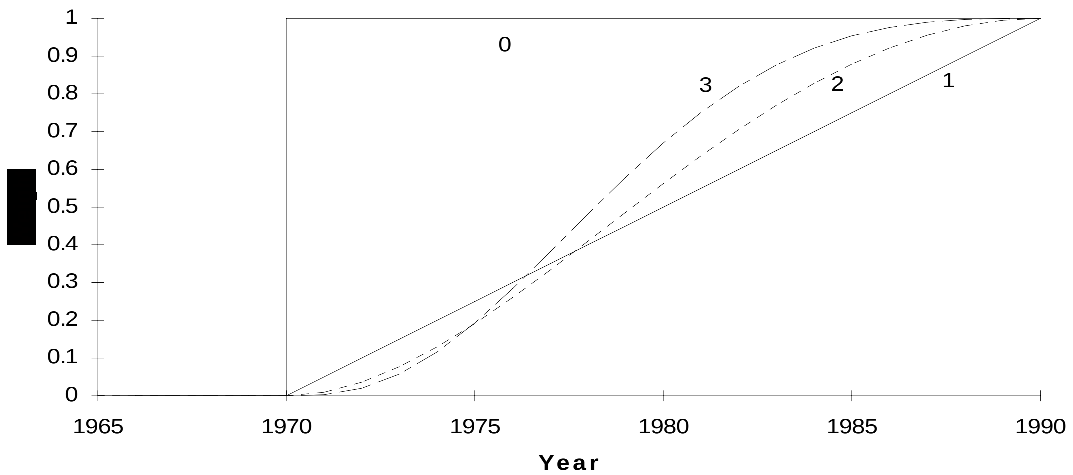
The weighting function, a modification to that described by Armstrong (1985), is :

$$W_{(ty-d)} = \left[1 - \left(\frac{R - d}{R}\right)^{p}\right]^{p} \qquad d = 0 .. R$$

where $W_{(ty-d)}$ = Weight in year ty - d
ty = Terminal year
R = Taper range
p = the function power: 0, 1, 2 & 3 giving uniform, linear, biquadratic or tricubic tapers

The weighting functions are illustrated in Figure 3. The *ad hoc* tuning and XSA default times series weights are set at a tricubic taper with a range of 20 years. This is a compromise between allowing recent data to receive the strongest weight during the estimation of the final year F and population abundances, whilst at the same time, allowing historical data to influence values derived for earlier years. After each run the generated values are printed in the tuning report file.

Figure 3. An illustration of the weights applied to a data set, by the weighting function with a series of powers (0-3). Taper range = 20 years.



Time series weights are mainly used for counteracting departures from the assumptions underlying an assessment methodology. Selection of the weighting pattern should be made with regard to the user's perception of the evolution of the interactions between the fleets and the exploited stock. In Separable VPA the weights are used to reduce the influence of changes in the exploitation pattern. Within XSA and Laurec-Shepherd tuning their main use is in maintaining the assumption of constant catchability during the time period used in the assessment. The influence of trends or sudden changes in fleet catchability (detected in the residuals and diagnostic statistics), can often be removed by the selection of the appropriate taper range.

A secondary effect of the use of time series weights is the concentration of the analysis on the most recent years for which data quality may be better. The Separable VPA weighting procedure allows removal of individual years.

Notes: The uniform weighting (power = 0) is not the same as no weighting. It applies to the taper range only; this may differ from the tuning range.

Appendix 10. Shrinkage to the mean.

Contents

The accuracy of predictions made by fitting an explanatory model to a data set can sometimes be improved by the use of the information contained within a mean of previous observations of the estimated parameter. The incorporation of the mean, with an appropriate weight, into the model reduces the prediction variance at the cost of a bias towards the mean. This process has been called 'shrinkage to the mean' (Anon., 1991(b), 1993(c)) due to similarities with shrinkage in multiple regression (Copas 1983). In terms of an assessment, the procedure can be described as making the assumption that, if a time series is being used to predict a current value of a particular parameter, e.g. F-at-age, and no major changes are known to have taken place; then as an initial starting value for the estimate, a mean of recent values of the parameter *"is not entirely stupid"* (*J.G.Shepherd pers.comm.*).

Anon.,(1993(c)) examines the use of shrinkage with a linear regression model and follows with an examination of its application to Laurec-Shepherd tuning. The study established that the 'best' estimate of a predicted value (terminal F), for a range of assessment data error scenarios, can be derived from a weighted average of the mean F and the predictions from the linear model. 'Best estimate' is defined as the smallest mean square error for the predicted value. The strength of the bias towards the mean (strength of the shrinkage) should be dependent on the quality of the data to which the model is fitted. Estimates derived from poor quality data with a low signal to noise ratio should have a greater contribution from the mean (heavier shrinkage). If the data quality is good, a strong signal to noise ratio, the variance of the prediction from the fitted model is reduced and the mean should be given less weight (lighter shrinkage). In practice this is achieved by using the inverse of the variance of the estimates as weights. In both *ad hoc* tuning and XSA the weights are entered as a fractional coefficient of variation for the shrinkage means. They are therefore functionally equivalent to the standard errors of the log catchabilities, with which they are combined in the overall weighted means.

Two forms of shrinkage are employed in the assessment suite. Shrinkage to the population mean (XSA), first used in RCT3 (Shepherd 1994), and shrinkage to a mean of the most recent F values (XSA and *ad hoc* tuning). The formulations are described in detail in Appendices 7 and 8.

If predictive regression is used (not recommended), shrinkage to the population mean is equivalent to a double shrinkage and should be avoided.