



NOAA Technical Memorandum NMFS-SEFSC-726
doi: 10.25923/6n3b-kd56

VPA-2BOX 4.01

User Guide

BY

Clay E. Porch



U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Southeast Fisheries Science Center
75 Virginia Beach Drive
Miami, Florida 33149

July 2018



NOAA Technical Memorandum NMFS-SEFSC-726
doi: 10.25923/6n3b-kd56

VPA-2BOX VERSION 4.01 USER GUIDE

BY

Clay E. Porch

National Marine Fisheries Service
Southeast Fisheries Science Center
75 Virginia Beach Drive
Miami, Florida 33149

U.S. DEPARTMENT OF COMMERCE
Wilbur L. Ross, Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
Tim Gallaudet, Undersecretary for Oceans and Atmosphere (Acting)

NATIONAL MARINE FISHERIES SERVICE
Christopher Oliver, Assistant Administrator for Fisheries

July 2018

Technical Memoranda are used for documentation and timely communication of preliminary results, interim reports, or special-purpose information. Although the memoranda are not subject to complete formal review, editorial control, or detailed editing, they are expected to reflect sound professional work.

NOTICE

The National Marine Fisheries Service (NMFS) does not approve, recommend or endorse any proprietary product or material mentioned in this publication. No reference shall be made to NMFS or to this publication furnished by NMFS, in any advertising or sales promotion which would imply that NMFS approves, recommends, or endorses any proprietary product or proprietary material mentioned herein which has as its purpose any intent to cause directly or indirectly the advertised product to be used or purchased because of this NMFS publication.

This report should be cited as follows:

Porch, C.E. 2018. VPA-2box Version 4.01 User Guide. NOAA Tech. Memo. NMFS-SEFSC-726. 67 pp. doi: 10.25923/6n3b-kd56.

Sponsored by:

National Marine Fisheries Service, Southeast Fisheries Science Center

Copies may be obtained by writing:

Southeast Fisheries Science Center
National Marine Fisheries Service, NOAA
75 Virginia Beach Drive
Miami, FL 33149

An online version is available at <http://www.sefsc.noaa.gov/publications/>

Acknowledgments

The framework and code for VPA-2BOX evolved largely as a product of the author's work with the Bluefin Tuna working group of the International Commission for the Conservation of Atlantic Tuna (ICCAT). Some of the concepts and algorithms used in the current software were inspired by early versions of code written independently in the 1990s by Joseph Powers, Victor Restrepo and Andre Punt. Since then, the software has benefitted from over two decades of interactions with numerous other scientists from many different working groups, having been applied at one time or the other to most of the species managed under ICCAT, to Albacore by the International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), and to numerous species in the United States. The present version of this manual benefited from helpful reviews by Dan Goethel and Alex Chester. Cover photo by Iñigo Onaindia (AZTI Tecnalia).

Contents

1. GETTING STARTED.....	1
WHAT'S NEW IN VERSION 4.01?.....	1
DISCLAIMER AND CONDITIONS OF USE	1
SYSTEM REQUIREMENTS	2
INSTALLING AND RUNNING VPA-2BOX.....	2
2. THEORY BEHIND VPA-2BOX	4
DETERMINISTIC POPULATION DYNAMICS	5
PARAMETER ESTIMATION.....	7
Phase 1: The Backward Recursion	7
Phase 2: Statistical estimation	10
BIAS AND UNCERTAINTY	21
Retrospective pattern analysis.....	25
MODEL SELECTION.....	26
Hypothesis tests.....	26
Information criteria.....	27
Bayes factors.....	28
Chi-square discrepancy statistic.....	28
NUMERICAL OPTIMIZATION	29
3. INPUT FILES	31
CONTROL FILE : See Appendix 1.....	32
DATA FILE : See Appendix 2	38
TAG RECOVERY DATA FILE : See Appendix 3.....	44
PARAMETER SPECIFICATIONS FILE : See Appendix 4	46
GROWTH.DAT FILE	51
4. OUTPUT FILES.....	52
PARAMETER ESTIMATE FILE.....	52
DIAGNOSTIC AND DERIVED STATISTICS FILE.....	53
DERIVED STATISTICS IN SPREADSHEET FRIENDLY FORMAT.....	55
BOOTSTRAP OUTPUT FILES.....	57
5. LITERATURE CITED	58

6. APPENDICES	60
APPENDIX 1. Sample CONTROL file with line numbers	60
APPENDIX 2. Sample data file with line numbers	62
APPENDIX 3. Sample Tag Data file with line numbers	65
APPENDIX 4. Sample parameter specification file with line numbers.....	66

1. GETTING STARTED

VPA-2BOX is a flexible software tool for analyzing the abundance and mortality of exploited animal populations that is based on the ADAPT framework developed by Parrack (1986) and Gavaris (1988). The primary difference between this package and other versions of ADAPT is the capability of analyzing two different stocks simultaneously, making possible routine quantitative analyses of the effect of sex-specific growth or stock intermixing. In addition, a wide variety of options are provided with respect to the types of data that may be used and the way the parameters are estimated.

WHAT'S NEW IN VERSION 4.01?

This update of versions 3.01 and 4.0 includes the following new features:

- 1) generalization of the tag-attribution model to allow for pop-up satellite tags etc...
- 2) new formats for the tagging data and tagging parameter definitions to make them easier to work with
- 3) use of MFEXP(arg) function to handle large arguments that cause overflows in EXP(arg)
- 4) added switch in the control file to make computing the covariance matrix optional
- 5) option to prorate the plus group out to older ages (makes equilibrium assumption in first year and then tracks each cohort)
- 6) allows stock composition data (fraction of catch comprising of each stock by age and year) to be used to estimate movement rates
- 7) allows indices of mean length or weight

DISCLAIMER AND CONDITIONS OF USE

The manual and executable versions of VPA-2BOX can be obtained free of charge either by contacting the author or by visiting <http://www.iccat.es/>. Source code is available at <https://github.com/ICCAT/software/>. Copies of the executable code and manual for VPA-2BOX may be distributed without restriction. The development of VPA-2BOX is ongoing, and the author appreciates comments regarding it, including any suspected bugs. However, no formal technical support is offered.

There is no warranty of any kind. While considerable effort has been made to ensure the program performs as described in the manual (see Appendix 3), the author and the U.S. government cannot be held responsible for any errors therein.

SYSTEM REQUIREMENTS

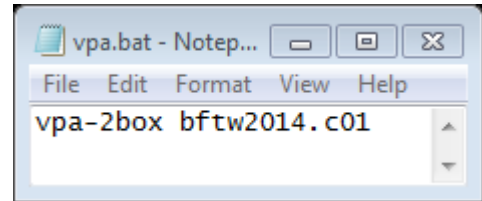
- Compatible with 80486, Pentium and above processors (standard PC's and clones)
- MS-DOS, IBM OS/2, or Microsoft Windows (9x and above) operating systems¹
- 1 MB of available hard disk space (30 MB for some bootstrap applications)

INSTALLING AND RUNNING VPA-2BOX

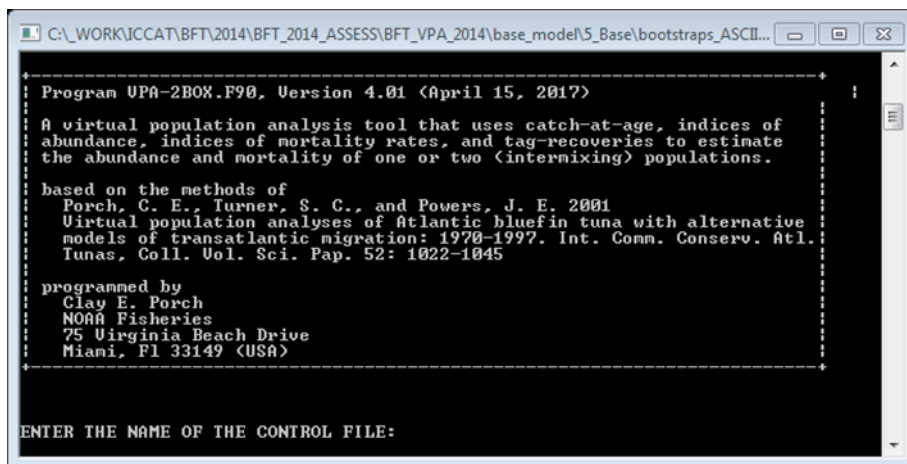
The only file needed for a complete installation is the executable VPA-2BOX.exe. To run the program you can either

- place the executable in a directory of your choosing and add that directory to your PATH statement as described in the documentation for your operating system and restart or
- place a copy of VPA-2BOX.exe in the directory where you plan to do the analyses.

The program may be run from the Windows Run dialog, entering the name VPA-2BOX at a DOS prompt, or double clicking on the VPA-2BOX .EXE icon in Windows Explorer. It may also be run in batch mode from the Windows command line, DOS or other shell environments. For example, a .bat file may be written such as the example `vpa.bat` to the right. The program may be started by double clicking or otherwise calling the batch file. In this example the name of the executable `vpa-2box` is followed by the name of an ASCII text file that contains the specifications needed to make the run (the user may specify the name and extension).



If the name of the control file is not entered in the command line, the program will prompt the user as shown in the screen capture below. There is no graphical user interface; VPA-2BOX runs only as a console-mode program that takes all input from ASCII text files (described in chapter 3). The screen is used only to enter the name of the control file and to display certain error and advisory messages. Hence, the user must be able to create and edit ASCII text files.



¹ Reference to trade names does not imply endorsement by NOAA or the U.S. Government

It is also possible to compile VPA-2BOX on linux and mac platforms. An example of code for a batch file, Makefile.bat, that builds the program for mac platforms and places the executable in a directory called “build” is included in the text box below (courtesy of Matthew Supernaw, Southeast Fisheries Science Center). Note that the compiler, in this case gfortran, is specified on the 12th line by FC = gfortran.

```
## -*- Makefile -*-
## User: matthewsupernaw
## Time: Jul 1, 2016 9:54:54 AM
## Makefile created by Oracle Developer Studio.
##
## This file is generated automatically.
##
# Remove the conflicting .mod extension. Make thinks .mod is a Modula2
# extension but the f90 compiler uses it for a module file.SUFFIXES
.SUFFIXES: .o .f90 .f90~
##### Compiler and tool definitions shared by all build targets #####
FC = gfortran
BASICOPTS = -O
FFLAGS = $(BASICOPTS)

# Define the target directories.
TARGETDIR_vpa-2box=build
all: $(TARGETDIR_vpa-2box)/vpa-2box

## Target: vpa-2box
OBS_vpa-2box = \
    $(TARGETDIR_vpa-2box)/vpa-2box.o
USERLIBS_vpa-2box = $(SYSLIBS_vpa-2box)
DEPLIBS_vpa-2box =
LDLIBS_vpa-2box = $(USERLIBS_vpa-2box)

# Link or archive
$(TARGETDIR_vpa-2box)/vpa-2box: $(TARGETDIR_vpa-2box) $(OBS_vpa-2box) $(DEPLIBS_vpa-2box)
    $(LINK.f) $(FFLAGS_vpa-2box) $(CPPFLAGS_vpa-2box) -o $@ $(OBS_vpa-2box) $(LDLIBS_vpa-2box)

# Compile source files into .o files
$(TARGETDIR_vpa-2box)/vpa-2box.o: $(TARGETDIR_vpa-2box) vpa-2box.f90
    $(COMPILE.f) $(FFLAGS_vpa-2box) $(CPPFLAGS_vpa-2box) -o $@ vpa-2box.f90

##### Clean target deletes all generated files #####
clean:
    rm -f \
        $(TARGETDIR_vpa-2box)/vpa-2box \
        $(TARGETDIR_vpa-2box)/vpa-2box.o
    rm -f *.mod
    rm -f -r $(TARGETDIR_vpa-2box)

# Create the target directory (if needed)
$(TARGETDIR_vpa-2box):
    mkdir -p $(TARGETDIR_vpa-2box)

# Enable dependency checking
.KEEP_STATE:
.KEEP_STATE_FILE::make.state.GNU-x86_64-Linux
```

2. THEORY BEHIND VPA-2BOX

Virtual population analysis (VPA) describes a family of techniques rooted in the seminal papers of Murphy (1965) and Gulland (1965). The method assumes that the catch history of any given year class is known without error, permitting the historical abundance and fishing mortality rates to be computed deterministically from a guess of the abundance or fishing mortality rate on the oldest (terminal) age of the year class. Various ad hoc methods have been proposed for fine tuning the guesses for the terminal parameters, but they invariably involve subjective criteria that may not be applicable to fisheries other than those for which they were developed. Parks (1976) appears to have been the first to estimate the terminal parameters from auxiliary data. He used least-squares to tune the VPA calculations of fishing mortality rates to fishing mortality rates derived from tagging experiments. Subsequently, Gray (1977), Doubleday (1981) and Parrack (1986) developed similar tuning procedures for use with various types of abundance indices.

VPA is not the only method available for analyzing catch-at-age data. A plethora of statistical models now exist that allow for errors in the catch data and a more detailed accounting of the dynamics of the fishery (e.g., Deriso et al., 1985; Fournier et al., 1999, Doonan et al. in press, Methot and Wetzel 2013, Dichmont et al. 2016). Nevertheless, tuned VPA's remain popular, in part because of their long history and relative simplicity, but also because they require fewer estimable parameters and converge faster than their more sophisticated counterparts. Some investigators also find it advantageous that tuned VPA's place no restrictions on the degree to which the selectivity/availability of each age class can vary from year to year

The present program is based on the ADAPT framework, which is essentially an extension of Parrack's (1986) model (see Gavaris, 1988; Conser and Powers, 1990; Powers and Restrepo, 1992). Various implementations of ADAPT have been widely used for domestic fisheries in the United States, South Africa and Canada, as well as in several international arenas, including the International Commission for the Conservation of Atlantic Tuna (ICCAT) and the Northwest Atlantic Fisheries Organization (NAFO).

The most distinctive feature of VPA-2BOX is, of course, its ability to examine two distinct groups simultaneously, making it amenable to sex-specific or stock-specific analyses where the two groups interact at some level. Aldenberg (1975) was the first to apply VPA methods to two intermixing stocks, thus solving the catch equations associated with the box-transfer model proposed by Beverton and Holt (1957). He examined the potential effect of various transfer rates on VPA appraisals of stock status, but did not attempt to estimate the transfer parameters for any stock in particular. Butterworth and Punt (1994) applied a similar box-transfer model to data for Atlantic bluefin tuna and showed that the magnitude of the transfer coefficients affected the model's ability to fit the indices of abundance. Subsequently, several versions of the box-transfer VPA were applied to Atlantic bluefin tuna in an attempt to estimate the transfer rates (NRC, 1994; Porch et al., 1995; Punt and Butterworth, 1995; Porch et al., 1998, 2001).

Note that the following conventions apply in the remainder of this document:

- *Italics* indicate mathematical variables or text to be replaced by the user.
- Courier font indicates text in a file or keystrokes.
- Y represents the number of years in the analysis
- A represents the oldest age in the analysis (a plus-group)
- α represents the youngest age in the analysis

DETERMINISTIC POPULATION DYNAMICS

The program is set up to model two populations simultaneously, either two sexes or two areas. In the case of two areas, the model is structured to accommodate a unique stock associated with each area (management zone). The user can choose between two types of box-transfer models to simulate intermixing between the two stocks: Diffusion and overlap (see Tables 1 and 2 below). The diffusion model assumes a fraction T_{jk} of the population located in management zone j will migrate to management zone k at the beginning of the year and then, if it is mature, contribute to the spawning biomass of the population in zone k (i.e., movement and spawning tendencies are determined based on the area an individual is currently in). When a spawner-recruit relationship is to be generated from the VPA results, then there is also the implicit assumption that immigrating fish will adopt the spawning habits of the native population as well (which may be biologically unrealistic for some stocks, see ICCAT, 1995: pp.108-110). As an alternative, Cooke and Lankester (1996) suggested the so-called ‘overlap’ model, which assumes a fraction \tilde{T}_{sk} of stock s resides in management zone k at any given time, but all mature fish spawn only in their natal area. Thus, fish movement in the overlap model depends on where an individual was spawned (i.e., its natal population) rather than its present location. Both models reduce to the equivalent of two single-stock VPA’s when the transfer fractions are set to zero, otherwise T and \tilde{T} are not directly comparable inasmuch as the diffusion model allows fish born in one zone to accumulate in the other (see Porch and Turner 1999, Porch et al, 2001; Porch, 2002). For example, if the transfer fractions for both stocks were both 10%, the diffusion model (without mortality) would calculate that 10% of the fish born in zone 1 will be living in zone 2 after one year, 18% after two years and 24.4% after three years. In contrast, an overlap model with the same transfer coefficients implies a constant 10% of the fish born in zone 1 will occupy zone 2 and vice versa (in effect all fish are assumed to ‘return’ at the end of the year to their zone of origin and redistribute again in the same proportion at the start of the next year).

Table 1. Standard catch equation assumed in program VPA-2BOX

Catch equation	Description
$C_{k,a,y} = \frac{F_{k,a,y}}{Z_{k,a,y}} N_{k,a,y} (1 - e^{-Z_{kay}})$	Catch of age a in year y from all stocks in management zone k
$Z_{kay} = F_{kay} + M_{kay}$	Total mortality rate in zone k
F_{kay}	Fishing mortality rate in zone k
M_{kay}	Natural mortality rate in zone k
$N_{k,a,y}$	Number in zone k at beginning of the year just <i>after</i> mixing (both stocks combined)

Table 2. Population dynamics equations for the diffusion and overlap frameworks.

Diffusion equations

$$\dot{N}_{k,a+1,y+1} = \begin{cases} R & (a = \alpha - 1) \\ \sum_{j=1}^2 \dot{N}_{j,a,y} T_{j,k,a} e^{-Z_{kay}} & (a < A - 1) \\ \sum_{a=A-1}^A \sum_{j=1}^2 \dot{N}_{j,a,y} T_{j,k,a} e^{-Z_{kay}} & (a = A - 1) \end{cases}$$

Number [of animals] in zone k that are age $a+1$ at the beginning of year y just *before* mixing (note accent over N), where A represents a ‘plus-group’ (age A and older) and $R(E)$ represents the spawner-recruit relationship (discussed later under Maximum Sustainable Yield).

$$N_{k,a,y} = \sum_j T_{j,k,a} \dot{N}_{j,a,y}$$

Number in zone k at beginning of the year just *after* mixing (both stocks combined)

$$T_{j,k,a}$$

Fraction of population in zone j that moves to zone k at the beginning of the year

$$Z_{kay} = F_{kay} + M_{kay}$$

Loss rate coefficients representing fishing (F) and natural mortality (M)

Overlap equations

$$\tilde{N}_{s,a+1,y+1} = \begin{cases} R & (a = \alpha - 1) \\ \tilde{N}_{s,a,y} \sum_{k=1}^2 \tilde{T}_{s,k,a} e^{-Z_{kay}} & (a < A - 1) \\ \sum_{a=A-1}^A \tilde{N}_{s,a,y} \sum_{k=1}^2 \tilde{T}_{s,k,a} e^{-Z_{kay}} & (a = A - 1) \end{cases}$$

Number of stock s that are age $a+1$ at the beginning of year y , where A represents a ‘plus-group’ (age A and older) and $R(E)$ represents the spawner-recruit relationship (described later under Maximum Sustainable Yield)

$$N_{k,a,y} = \sum_s \tilde{T}_{s,k,a} \tilde{N}_{s,a,y}$$

Number in zone k at the beginning of year y just after redistribution (both stocks combined)

$$\tilde{T}_{s,k,a}$$

Fraction of stock s residing in zone k at the beginning of the year

$$Z_{k,a,y} = F_{k,a,y} + M_{k,a,y}$$

Loss rate coefficients representing fishing (F) and natural mortality (M)

As its name suggests, VPA-2BOX considers only two zones, in which case $\tilde{T}_{s,2,a,y} = 1 - \tilde{T}_{s,1,a}$ and $T_{j,2,a} = 1 - T_{j,1,a}$. For convenience of notation the second subscript has henceforth been omitted. In the case of the overlap model, $\tilde{T}_{s,a}$ ($= \tilde{T}_{s,k,a}$) is the fraction of stock s that sojourns away from the natal area and $1 - \tilde{T}_{s,a}$ is the fraction that remains in the natal area. Similarly, for the diffusion model, $T_{j,a}$ ($= T_{j,k,a}$) is the fraction of the population residing in area j that moves to the opposite area and $1 - T_{j,a}$ is the fraction that remains in j .

PARAMETER ESTIMATION

The equations describing the catch and population dynamics (above) include a large number of variables representing the transfer coefficients, natural mortality rates, fishing mortality rates, catches and population abundances. Of these, it is often only the catches that are actually observed. The approach Murphy (1965) and Gulland (1965) took to solve this problem, commonly referred to as Virtual Population Analysis (VPA), was to back-calculate the historical abundance and fishing mortality rate of each cohort from starting guesses of the fishing mortality rate on the oldest observed age classes (F_{Ay} or F_{aY}).

The VPA solution depends on three main assumptions: (1) the catch equation correctly describes the dynamics of the fishery, (2) the observed catches are known with negligible error, and (3) reasonable guesses can be made for M , T , F_{Ay} and F_{aY} . The first assumption is common to all modeling exercises and is discussed further in the section on Model Selection. The second assumption is an important factor in deciding whether the VPA solution is appropriate; if the catches are not well known, then a statistical catch-at-age model (SCAA) may be more suitable².

The third assumption has often been addressed by fixing M , T , F_{Ay} and F_{aY} to several values that seem plausible and then comparing the resulting solutions. The approach taken by ADAPT and related methods differs from this in that at least some of these parameters are estimated internally by use of auxiliary data such as indices of abundance or tag recoveries-- a process sometimes referred to as 'tuning' the VPA. The tuning procedure occurs in two phases. The first phase is the explicit recursion, which determines the historical abundance and mortality rates from the latest estimates for M , T , F_{Ay} and F_{aY} . The second phase uses the abundance and mortality rate estimates from the first step to predict the values of the data. The first and second phase are repeated for different combinations of values for F_{Ay} , F_{aY} , M and T and the best estimates are taken to be those that minimize the discrepancy between the model predictions and the observed values of the data as measured by an appropriate statistical model.

Phase 1: The Backward Recursion

Single stock, cohorts that have not reached the plus-group

A number of the cohorts alive during the last year considered in the analysis will not have reached the plus-group age. In that case the abundance at the beginning of the year for any given cohort can be determined from the fishing mortality rate during that year by rearranging and solving the catch equation

$$N_{a,y} = \frac{Z_{a,y}}{F_{a,y}(1-e^{-Z_{a,y}})} C_{a,y} \quad (2.1)$$

² It is important to keep in mind that VPA does have an important advantage over SCAA in that it allows the selectivity to vary to an unlimited degree from one year to the next (SCAA must constrain changes in selectivity in one way or another). The user must therefore consider the tradeoff between the VPA's ability to accommodate large inter-annual changes in selectivity and the SCAA's ability to deal with imprecise catch-at-age information.

The fishing mortality rate on the next younger age, $F_{a-1,y-1}$, can then be determined by substituting the $N_{a,y}$ calculated above into another equation derived by combining the catch and abundance equations for the next younger age,

$$F_{a-1,y-1} = \frac{Z_{a-1,y-1} C_{a-1,y-1}}{(e^{Z_{a-1,y-1}} - 1)N_{a,y}} \quad (2.2)$$

Note that the value of $F_{a-1,y-1}$ must be solved numerically because it is included in the exponential term (VPA-2BOX uses the bisection routine for this type of univariate search). The two-step cycle above must be initiated with a guess for the value of the abundance or fishing mortality on each cohort during the last year ($N_{a,y}$ or $F_{a,y}$). The resulting estimate of $F_{a-1,y-1}$ is then used to determine $N_{a-1,y-1}$ and $F_{a-2,y-2}$, and so on back to the youngest age class in the analysis.

Single stock, cohorts that reach the plus-group

The plus-group lumps together age A fish from one cohort with older fish from previous cohorts, therefore it is necessary to develop a recursion that allows fish that just turned exactly age A in year y from those that have accumulated from older cohorts. This is done by simultaneously solving for $N_{A,y}$ and $N_{A,y-1}$ by use of (2.1):

$$\begin{aligned} N_{A,y} &= \frac{Z_{A,y}}{F_{A,y}(1 - e^{-Z_{A,y}})} C_{A,y} \\ N_{A,y-1} &= \frac{Z_{A,y-1}}{F_{A,y-1}(1 - e^{-Z_{A,y-1}})} C_{A,y-1} \end{aligned} \quad (2.3)$$

Next, the values for $N_{A,y}$ and $N_{A,y-1}$ are substituted into the equation

$$F_{A-1,y-1} = \frac{Z_{A-1,y-1} C_{A-1,y-1}}{(e^{Z_{A-1,y-1}} - 1)(N_{A,y} - N_{A,y-1}e^{Z_{A,y-1}})} \quad (2.4)$$

which is solved numerically for $F_{A-1,y-1}$. The value of $F_{A-1,y-1}$ is in turn used to initiate the recursive sequence for the younger age classes discussed in the preceding paragraph (equations 2.1 and 2.2).

The recursion for a plus-group (2.4) requires starting guesses for the fishing mortality rate on the oldest age in two years ($F_{k,A,y}$ and $F_{k,A,y-1}$), whereas the recursion without a plus-group requires only $F_{k,A,y}$. This is of little consequence where a succession of cohorts is to be analyzed because starting values must be supplied for all years regardless. However, determining those starting values can be problematic. Program VPA-2BOX adopts the common convention of expressing the fishing mortality rate on the plus-group as a fraction ϕ of the fishing mortality rate on the next younger group, $F_{k,A,y} = \phi_{k,y} F_{k,A-1,y}$. The value of ϕ should be less variable from year to year than the value of F inasmuch as changes in effort are to some degree factored out. In some applications the value of ϕ is assumed to be 1.0 while in others they are estimated.

Two stocks

The recursion for the overlap and diffusion models for two stocks is similar in principle to that above, but requires the linear decomposition of a system of equations. For example, in the case of the overlap model the equation analogous to (2.1) is

$$N_{k,a,y} = \frac{Z_{k,a,y}}{F_{k,a,y}(1-e^{-Z_{k,a,y}})} C_{k,a,y} \quad (2.5)$$

Once $N_{1,a,y}$ and $N_{2,a,y}$ are solved for (using the bisection routine), they are substituted into the system

$$C_{1,a-1,y-1} = \frac{F_{1,a-1,y-1}(1-e^{-Z_{1,a-1,y-1}})}{Z_{1,a-1,y-1}} * \left(\frac{(1-\tilde{T}_{1,a-1})N_{1,a,y}}{(\tilde{T}_{1,a-1}e^{-Z_{2,a-1,y-1}}+(1-\tilde{T}_{1,a-1})e^{-Z_{1,a-1,y-1}})} + \frac{(\tilde{T}_{2,a-1})N_{2,a,y}}{\tilde{T}_{2,a-1}e^{-Z_{1,a-1,y-1}}+(1-\tilde{T}_{2,a-1})e^{-Z_{2,a-1,y-1}}} \right) \quad (2.6)$$

$$C_{2,a-1,y-1} = \frac{F_{2,a-1,y-1}(1-e^{-Z_{2,a-1,y-1}})}{Z_{2,a-1,y-1}} * \left(\frac{(1-\tilde{T}_{2,a-1})N_{2,a,y}}{(\tilde{T}_{2,a-1}e^{-Z_{1,a-1,y-1}}+(1-\tilde{T}_{2,a-1})e^{-Z_{2,a-1,y-1}})} + \frac{(\tilde{T}_{1,a-1})N_{1,a,y}}{\tilde{T}_{1,a-1}e^{-Z_{2,a-1,y-1}}+(1-\tilde{T}_{1,a-1})e^{-Z_{1,a-1,y-1}}} \right)$$

These two equations are solved for $F_{1,a-1,y-1}$ and $F_{2,a-1,y-1}$ given $N_{1,a,y}$ and $N_{2,a,y}$ (and T and M) using Newton's method, then $N_{1,a-1,y-1}$ and $N_{2,a-1,y-1}$ can be computed from the estimates of $F_{1,a-1,y-1}$ and $F_{2,a-1,y-1}$ by use of equation (2.5), thus constituting the backward recursion required by the VPA.

In case of a plus-group, the four catch equations corresponding to ages A and $A-1$ in the two areas can be solved for $N_{1,A-1,y}$, $N_{2,A-1,y}$, $N_{1,A,y}$, $N_{2,A,y}$ in terms of C , F , M and T by linear decomposition. These results in turn are substituted into the system of abundance equations in Table 1, yielding 2 equations in four unknowns ($F_{1,A-1,y}$, $F_{2,A-1,y}$, $F_{1,A,y}$, $F_{2,A,y}$). Program VPA-2BOX adopts the common convention of expressing the fishing mortality rate on the plus group as a fraction ϕ of the fishing mortality rate on the next younger group,

$$F_{k,A,y} = \phi_{k,y} F_{k,A-1,y} \quad , \quad \text{for } k=1,2$$

This reduces the number of unknowns to 2 and allows the system to be solved for $F_{1,A-1,y}$ and $F_{2,A-1,y}$ using Newton's method for 2 equations (the ϕ values either being imposed or estimated).

Phase 2: Statistical estimation

The challenge that remains is to choose appropriate values for M_{kay} , T_{kay} , and the parameters that represent the fishing mortality on the oldest age in the cohort, $\phi_{k,y}$, or in the case of incomplete cohorts represented by the last year of data, F_{kay} . This involves the second phase of the tuning procedure discussed earlier, which is discussed in detail below.

Basic Theory

The basic tenet behind any statistical estimation procedure is that any data element d_j may be described as a scalar function f of a covariate vector X_j (e.g., time), parameter vector Θ , and random noise. One of the most popular estimation approaches historically has been the method of least squares, where the random noise is considered either to be additive ($d_j = f[X_j, \Theta] + \varepsilon$) or multiplicative ($d_j = f[X_j, \Theta] e^\varepsilon$) such that the best estimate of Θ is taken to be that which minimizes, respectively,

$$\sum_j (d_j - f(X_j, \Theta))^2 \quad (2.7)$$

or

$$\sum_j (\log d_j - \log f(X_j, \Theta))^2 \quad (2.8)$$

Another popular approach is the method of maximum likelihood, where the probability of observing the data matrix \mathbf{D} under a proposed model structure is expressed as a probability function:

$$P(D|\Theta, X) = \prod_j P(d_j|\Theta, X_j) \quad (2.9)$$

commonly referred to as the likelihood function. According to classical maximum likelihood theory, the values of Θ that maximize $P(D|\Theta, X)$ will be asymptotically efficient, unbiased estimates of the true values provided the covariates are known without error.

To see how the maximum likelihood method works, consider the special case where the likelihood function P has the form of a normal distribution with constant variance σ^2 . In that case,

$$P(D|\Theta, X) = \prod_j \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(d_j - f(\Theta, X_j))^2}{2\sigma^2}} \quad (2.10)$$

The maximum likelihood estimates (MLE) of Θ are the values that “maximize” the likelihood function P , i.e, the values is most “likely” to produce the observed data D if the model f is correct. Since the maximum of $P(D|\Theta, X)$ occurs at the same values of Θ as the minimum of its negative logarithm, one may equivalently find the values that minimize

$$-\log P(D|\Theta, X) = \sum_j \frac{(d_j - f(\Theta, X_j))^2}{2\sigma^2} - \log \sqrt{2\pi}\sigma \quad (2.11)$$

Notice that (2.11) has a form very similar to the least-squares estimator in (2.7); in fact they are minimized at exactly the same values of Θ inasmuch as σ^2 and 2π are constants. Thus, it can be

said that the additive least-squares estimator is equivalent to the maximum likelihood estimator under a normal distribution with constant variance. A similar equivalency can be shown for the multiplicative least-squares estimator (2.8) and the maximum likelihood estimator under a lognormal distribution with constant coefficient of variation.

As mentioned previously, one of the desirable properties of the maximum likelihood method is that, if an efficient unbiased estimator exists, the maximum likelihood method will produce it-- provided the covariates X are known without error. However, in most stock assessment models some of the covariates will be error-prone and alternative methods of estimation may be required (see Seber and Wild, 1989). One general class of models that deals explicitly with error prone covariates is the state-space model. Typically the observations D (e.g., indices of abundance) are expressed as functions of time-independent parameters Θ and time-dependent covariates X_t that include the unobserved states of the system. A stochastic representation of a state space model is therefore defined by two probability statements: $P(D | \Theta, X)$, quantifying the likelihood of observation errors in the data, and $P(X | \Theta)$, quantifying the likelihood of process errors in the covariates (states).

A statistically rigorous Bayesian treatment of state space models is straightforward (see Schnute, 1994). By Bayes theorem, $P(\Theta, X | D) P(D) = P(D | \Theta, X) P(\Theta, X)$. Inasmuch as $P(D)$ is a constant and $P(\Theta, X)$ may be expressed as $P(X | \Theta) P(\Theta)$,

$$P(\Theta, X | D) \sim P(D | \Theta, X) P(X | \Theta) P(\Theta) \quad (2.12)$$

Here $P(\Theta)$ is the so-called 'Bayes prior', a prior estimate of the probability density for Θ . The classical Bayes moment estimator for a parameter θ_i in the set Θ is:

$$\hat{\theta}_i = \int \theta_i P(D | \Theta, X) P(X | \Theta) P(\Theta) d\theta_i$$

Alternatively, estimates for Θ (and therefore also X) may be obtained by maximizing (2.12), known as the method of highest posterior density (HPD) or penalized likelihood. As for maximum likelihood, the HPD estimates may equivalently be found by minimizing the negative logarithm of (2.12):

$$-\log P(\Theta, X | D) \sim -\log P(D | \Theta, X) -\log P(X | \Theta) -\log P(\Theta) \quad (2.13)$$

Notice that when process errors $P(X | \Theta)$ and priors $P(\Theta)$ are ignored, equation (2.12) reduces to the classical negative log-likelihood expression and the HPD estimates are equivalent to the maximum likelihood (ML) solution. In that case the covariance matrix may be obtained from the inverse of the matrix of second derivatives with respect to Θ .

The use of Bayes priors, penalties and process variance models allow the user to help constrain the model to reflect information other than the observed catch, index or tagging data. They also provide a convenient way of incorporating the uncertainty in expert judgement into the assessment. The primary disadvantage of the use of penalties and priors is that the results tend to be sensitive to the shape of the prior distribution assumed (see Gelman et al., 2000) and can introduce bias. Program VPA-2BOX allows the user to impose priors on all parameters, allows for process variance in several aspects, and incorporates a variety of penalty functions (which act similarly to process errors). Estimation is currently limited to the HPD method.

Estimation from abundance indices

Program VPA-2BOX allows parameters to be estimated by minimizing the negative log-likelihood function of the observed, $I_{i,k,y}$, and predicted, $\hat{I}_{i,k,y}$, values for various indices of abundance, where i is the index series identifier, k is the management zone and y is the year (see Table 3 for the options regarding likelihood functions). This introduces several new variables that need to be accounted for—the index standard error σ , catchability q , and relative vulnerability to the gear ν (which implicitly includes factors such as gear selectivity and the fraction of the population available to be caught). There are many options available in VPA-2BOX pertaining to how these variables may be treated and the reader is referred to the documentation on the input files for a detailed accounting.

One aspect about VPA-2BOX that is important to understand is the way the variance parameter σ^2 is represented. There are two basic components—a fixed value that is determined externally and entered in the data file for each year ($input_{i,k,y}$) and an estimable parameter ($\eta_{i,k}$) that scales the input value. The input value is interpreted as a coefficient of variation (CV) if it is positive and a standard error if it is negative. The scaling parameters may be multiplicative or additive (see line 47 of the CONTROL file). In the multiplicative case

$$\sigma_{i,k,y}^2 = \begin{cases} (input_{i,k,y} \hat{I}_{i,k,y} \eta_{i,k})^2 & \text{input value positive (CV)} \\ (input_{i,k,y} \eta_{i,k})^2 & \text{input value negative (std. error)} \end{cases} \quad (2.14a)$$

In the additive case

$$\sigma_{i,k,y}^2 = \begin{cases} (input_{i,k,y} \hat{I}_{i,k,y})^2 + \eta_{i,k}^2 & \text{input value positive (CV)} \\ (input_{i,k,y})^2 + \eta_{i,k}^2 & \text{input value negative (std. error)} \end{cases} \quad (2.14b)$$

The multiplicative η corresponds to the approaches used by Quinn and Deriso (1999) and Sullivan (1999), where the $input_{i,k,y}$ values are interpreted as expert judgments of the relative variance of each annual observation of each index, presumably reflecting knowledge about how the data were obtained and their representativeness of the overall system (e.g., spatial and temporal coverage). The η 's then simply scale the inputs up or down to reflect the overall variance of the system. The additive η corresponds to the ‘additional variance’ approach of Geromont and Butterworth (2001), where the $input_{i,k,y}$ values are interpreted as the variance of the index associated with observation errors (such as might be obtained from the GLM standardization approach) and the η^2 are interpreted as the additional variance associated with process errors (characterizing the uncertainty in the index as a reflection of the true abundance trends as might occur when the proportion of the stock sampled by the index varies from year to year). A detailed review of the pro's and con's of the additive and multiplicative approaches is beyond the scope of this manual, but some interesting discussions can be found in Legault and Porch (1999) and McAllister et al.(2001).

It is useful to note that the ‘equal weighting’ approach used by many investigators (when presumably they find no basis for weighting one index more than another) can be achieved by setting the $input_{i,k,y}$ values to 1.0 and then estimating (or fixing) a single, multiplicative value of η for all indices. Conversely, the ‘input variance’ approach (used when externally derived values are presumed to be adequate) can be affected by fixing the η 's to a value of 1.0.

Table 3. Negative log-likelihood equations for indices of abundance (index series: i , zone: k , age: a , year: y) corresponding to the probability densities (PDF) that are available in VPA-2BOX.

Equations and variables	PDF	Notes
$\sum_i \sum_k \sum_y 0.5 \left(\frac{\ln(I_{iky}/\hat{I}_{iky})}{\tilde{\sigma}_{iky}} \right)^2 + \ln \tilde{\sigma}_{iky}$	lognormal*	
$\sum_i \sum_k \sum_y 0.5 \left(\frac{I_{iky} - \hat{I}_{iky}}{\sigma_{iky}} \right)^2 + \ln \sigma_{iky}$	normal	For the lognormal distribution $\tilde{\sigma}$ is the standard error of $\log I$ such that,
$\sum_i \sum_k \sum_y \hat{I}_{iky} - I_{iky} \ln(\hat{I}_{iky})$	Poisson	$\tilde{\sigma}_{iky} = \sqrt{\log_e \left\{ \left(\frac{\sigma_{iky}}{\hat{I}_{iky}} \right)^2 + 1 \right\}}$
$\sum_i \sum_k \sum_y 0.5 \frac{(I_{iky} - \hat{I}_{iky})^2}{\sigma_{iky}^2 (\hat{I}_{iky} + 1)}$	Chi-square	and for the gamma distribution:
$\sum_i \sum_k \sum_y \left(\frac{\sqrt{2} I_{iky} - \hat{I}_{iky} }{\sigma_{iky}} \right)^2 + \ln \sigma_{iky}$	Laplace	$\alpha = (\hat{I}_{iky} / \sigma_{iky})^2$ and $\beta = \hat{I}_{iky} / \alpha$
$\sum_i \sum_k \sum_y C \ln \beta - (\alpha - 1) \ln I_{iky} - \frac{I_{iky}}{\beta} - \ln \Gamma(\alpha)$	gamma	
$I_{i,k,y}$		observed value of index
$\hat{I}_{i,k,y} = q_{i,k,y} \sum_a \Delta_{i,k,a,y} v_{i,k,a,y} w_{i,k,a,y} N_{k,a,y}$		predicted value of index
$v_{i,k,a,y}$		relative selectivity/availability at age
$\sigma_{i,k,y}$		standard error of index
$q_{i,k,y}$		scaling (proportionality) coefficient
$w_{i,k,a,y}$		weight
$\Delta_{i,k,a,y}$		adjustment for time of year (see eq. 3.15)

Estimation from mortality rate indices

Parks (1976) estimated the VPA terminal fishing mortality parameters by use of a least-squares fitting to fishing mortality rates obtained from tagging experiments, i.e., by

$$\min \sum_y \sum_a (F_{ay} - {}^\tau F_{ay})^2 \quad (2.15)$$

where the superscript τ denotes the tagging-based values. This procedure assumes that the fishing mortality on the tagged population is the same as on the untagged population, however this may not be true for a number of reasons (see the next section). A less restrictive assumption would be to assume that F is proportional to ${}^\tau F$ by a factor q , i.e.,

$$\min \sum_y \sum_a (qF_{ay} - {}^\tau F_{ay})^2 \quad (2.16)$$

More generally, the objective is to make inferences about the loss rate of a certain target population based on estimates of the loss rate from a subpopulation τ . Suppose that the loss rate experienced by τ has some elements in common with the target population, some elements that are different, and some elements that are proportional. The two mortality rates could then be represented by the formulas $Z = M + F$ and ${}^\tau Z = M + L + qF$. When τ represents a group of tagged fish, for example, these equations could be interpreted to mean that tagged and untagged fish have the same natural loss rate M , tagged fish have an additional loss rate L owing to tag shedding or other factors, and the fishing mortality rate on tagged fish is proportional to that on the untagged population owing to incomplete mixing or a change in catchability.

If the factors q , M and L can be considered relatively constant through time, then M and L may be removed simply by subtracting the observed values of ${}^\tau Z_y$ from the value for one standard year ${}^\tau Z_\psi$

$${}^\tau Z_\psi - {}^\tau Z_y = q(F_\psi - F_y) \quad (2.17)$$

where ψ is the standard year (perhaps the year with the maximum value of ${}^\tau Z$ to keep the deviations positive). Thus, a least squares estimation procedure analogous to (2.15) and (2.16) would be

$$\min \sum_{y \neq \psi} \sum_a (q(F_{\psi a} - F_{ya}) - ({}^\tau Z_\psi - {}^\tau Z_{ya}))^2 \quad (2.18)$$

Note that this is an implicitly less assuming procedure than (2.16) because it does not require the total loss rate estimates for the subpopulation to be partitioned into fishing and non-fishing components. Moreover there is little to gain in using (2.17) even when external estimates of fishing mortality rates are available unless q is known to be 1.0. It should also be noted that the least-squares formulations above have essentially the same form as those for the indices of abundance. It is a trivial matter to derive analogous maximum likelihood formulations (see Table 4).

Two extensions of the above methodology are included as options in the current software. The first allows for a slightly different model, ${}^\tau Z = qZ + b$ (so ${}^\tau Z_\psi - {}^\tau Z_y = q[Z_\psi - Z_y]$). The second option accommodates the possibility that ${}^\tau F$ or ${}^\tau Z$, when they represent the effective mortality rate on multiple age groups, may not represent those age groups in the same proportion as they exist in

target population. This may be useful when it is possible to derive these proportions external to the VPA, but not otherwise as they cannot be estimated from the partial catches as is done for indices of abundance. See the data file section and Table 4 for further details.

Table 4. Models for indices of mortality rate (index series: i , zone: k , age: a , year: y) available in VPA-2BOX. Options 3-6 refer to the third entry on line 56 of the catch data file.

Equations and variables	PDF	Notes
$\sum_i \sum_k \sum_y 0.5 \left(\frac{\ln(D_{iky}/\hat{D}_{iky})}{\tilde{\sigma}_{iky}} \right)^2 + \ln \tilde{\sigma}_{iky}$	lognormal	For the lognormal distribution $\tilde{\sigma}$ is the standard error of $\log I$ such that, $\tilde{\sigma}_{iky} = \sqrt{\log_e \left[\left(\frac{\sigma_{iky}}{\hat{D}_{iky}} \right)^2 + 1 \right]}$
$\sum_i \sum_k \sum_y 0.5 \left(\frac{D_{iky} - \hat{D}_{iky}}{\sigma_{iky}} \right)^2 + \ln \sigma_{iky}$	normal	
$\sum_i \sum_k \sum_y \hat{D}_{iky} - D_{iky} \ln(\hat{D}_{iky})$	Poisson	and for the gamma distribution: $\alpha = (\hat{D}_{iky} / \sigma_{iky})^2$
$\sum_i \sum_k \sum_y 0.5 \frac{(D_{iky} - \hat{D}_{iky})^2}{\sigma_{iky}^2 (\hat{D}_{iky} + 1)}$	Chi-square	and $\beta = \hat{D}_{iky} / \alpha$
$\sum_i \sum_k \sum_y \left(\frac{\sqrt{2} D_{iky} - \hat{D}_{iky} }{\sigma_{iky}} \right)^2 + \ln \sigma_{iky}$	Laplace	
$D_{iky} = \begin{cases} \text{Max}_y \{Z_{iky}\} - Z_{iky} & \text{options 3 or 4} \\ F_{iky} & \text{option 5} \\ Z_{iky} & \text{option 6} \end{cases}$		observed value of standardized mortality rate index
$\hat{D}_{iky} = \begin{cases} q_{iky} \sum_a v_{ikay} (\hat{F}_{kay} - \hat{F}_{kay}) & \text{option 3} \\ q_{iky} \sum_a v_{ikay} (\hat{Z}_{kay} - \hat{Z}_{kay}) & \text{option 4} \\ q_{iky} \sum_a v_{ikay} \hat{F}_{kay} & \text{option 5} \\ q_{iky} \sum_a v_{ikay} \hat{Z}_{kay} & \text{option 6} \end{cases}$		predicted value of standardized mortality rate index
$v_{i,k,a,y}$		relative selectivity/availability at age
$\sigma_{i,k,y}$		standard error of index
$q_{i,k,y}$		scaling (proportionality) coefficient

Estimation from the mean weight of the plus group

Gedamke and Hoenig (2006) demonstrated that trends in total mortality can be estimated from a time series of mean length information for size classes above a certain critical size where the vulnerability at size or age can be assumed constant. In principle a similar approach can be taken to inform mortality rates for the plus-group (where it is implicitly assumed that the mortality rate is constant for all age classes at or above the plus-group age). The average weight of the plus group is a reflection of the relative catch of each age class

$$\bar{W}_{k,A,y} = \frac{\sum_{a=A}^{A+n} C_{k,a,y} W_{k,a,y}}{\sum_{a=A}^{A+n} C_{k,a,y}} \quad (2.19)$$

Inasmuch as the plus-group is usually established because it is not practical to determine the age of fish above a certain threshold A , the $C_{k,a,y}$ will not usually be available for (2.19).

An equivalent predictor for the average weight of the plus group is

$$\bar{W}_{k,A,y} = \frac{\sum_{a=A}^{A+n} N_{k,a,y} W_{k,a,y} F_{k,A,y} (1 - e^{-Z_{k,A,y}}) / Z_{k,A,y}}{\sum_{a=A}^{A+n} N_{k,a,y} F_{k,A,y} (1 - e^{-Z_{k,A,y}}) / Z_{k,A,y}} = \frac{\sum_{a=A}^{A+n} N_{k,a,y} W_{k,a,y}}{\sum_{a=A}^{A+n} N_{k,a,y}} \quad (2.20)$$

where n is the number of age classes older than A that will be tracked in the expanded plus-group (usually chosen to be large enough to where growth is negligible). Note terms involving $F_{k,A,y}$ and $Z_{k,A,y}$ cancel out because they are the same for all ages within the plus group. The application of (2.20) requires calculating the abundance of each of the n age classes contained in the plus-group ($N_{k,a,y}$). The VPA-2BOX program does this using the “expanded plus-group” feature discussed in connection with line 2 of the CONTROL file. Essentially the approach assumes that the members of the plus-group are approximately in equilibrium during the first year such that the sum of the number in each age class a ($a \geq A$) is the same as the number predicted by the VPA for the plus-group as a whole. In the case of a single stock,

$$N_{a,y_0} = N_{A+,y} \frac{e^{-Z_{A,y_0}(a-A)}}{\sum_{j=A}^{A+n} e^{-Z_{A,y_0}(j-A)}} \quad .$$

where the subscript $A+$ in $N_{A+,y}$ is used here to refer to the VPA estimate of the combined abundance of the plus-group, as opposed to the abundance of fish in age class A alone ($N_{A,y}$). Analogous equations are used for the equilibrium condition with two stocks under the diffusion and overlap models as appropriate. Calculations for subsequent years follow the forward recursions in Table 2.

The expected weight of each age class during the year is calculated from a growth curve G that is read from an input file (GROWTH.DAT, see sample file in section 3). If the catch of the plus group is thought to occur concentrated around a certain time of year, the weight may be computed for a specific time of year τ . If the catch is made more or less throughout the year, the average weight may be computed to close approximation by

$$W_{k,a,y} = \frac{\sum_{j=1}^{10} G(t) e^{-Z_{k,A,y}(t)}}{\sum_{j=1}^{10} e^{-Z_{k,A,y}(t)}} \quad , \quad (2.21)$$

where $t = (j-0.5)/10$.

The calculated average weights from (2.20) are compared to the input values in a manner similar to the indices of abundance and the same probability densities are available (the likelihood equations are identical except that $I_{i,k,y}$ is replaced by $\bar{w}_{k,A,y}$).

Estimation from stock composition

The proportion of the catch in a management zone that is made up of the stock that originates in that zone can provide important information on the relative abundance of two overlapping stocks. It is not informative for the diffusion model used here because the identity of a fish is assumed to change with its location (i.e., the equations would need to be changed to track where the animal originated from). The expectation is straightforward

$$\hat{p}_{k,a,y} = \frac{\tilde{N}_{k,a,y}(1-\tilde{T}_{k,a,y})}{N_{k,a,y}} , \quad (2.22)$$

where $\tilde{N}_{k,a,y}(1-\tilde{T}_{k,a,y})$ is the abundance of stock k in zone k and $N_{k,a,y}$ is the combined abundance of both stocks in zone k . If the stock identity of each fish can be determined with little error, the total number of animals in a sample of $v_{a,y}$ fish that come from stock k ($v_{k,a,y}$) is binomial distributed with probability

$$p_{k,a,y} = \frac{v_{k,a,y}}{v_{a,y}} , \quad (2.23)$$

and variance

$$VAR(v_{k,a,y}) = v_{k,a,y} \left(1 - \frac{v_{k,a,y}}{v_{a,y}}\right) . \quad (2.24)$$

In practice, however, the identity of each fish is seldom known perfectly so that there is some uncertainty about $v_{k,a,y}$ beyond that associated with sample size. This suggests that some form of over-dispersed binomial distribution is more appropriate. Program VPA-2BOX takes advantage of the fact that the binomial distribution tends to the normal distribution for large $v_{a,y}$, in which case an appropriate likelihood expression is

$$\sum_j \frac{(v_{k,a,y} - v_{a,y} \hat{p}_{k,a,y})^2}{2\sigma_{v_{k,a,y}}^2} - \log \sqrt{2\pi} \sigma_{v_{k,a,y}} \quad (2.25)$$

where

$$\sigma_{v_{k,a,y}}^2 = \begin{cases} v_{k,a,y} \left(1 - \frac{v_{k,a,y}}{v_{a,y}}\right) + \eta_{k,a}^2 & \text{additive} \\ v_{k,a,y} \left(1 - \frac{v_{k,a,y}}{v_{a,y}}\right) \eta_{k,a} & \text{multiplicative} \end{cases} \quad (2.26)$$

Here η is a variance scaling parameter that can be estimated analogous to that discussed in connection with equations (2.14a) and (2.14b).

Estimation from tag releases and recoveries

Program VPA-2BOX models the temporal and spatial distribution of tag recoveries R from a group (cohort i) of tagged animals with a common age, stock of origin and release venue (date and zone) as described in Table 5. The likelihood forms shown in Tables 3 and 4 are available for comparing the observed and expected number of recoveries in each strata, but the multinomial likelihood shown in Table 5 is the most common choice in the literature (see Schweigert and Schwarz, 1993; Anganuzzi et al., 1994). The tag attrition model includes terms for the reporting rate (ρ), immediate loss of tags owing to misapplication or tagging-induced mortality (d), and chronic loss of tags owing to shedding or fouling (λ). The first two terms always occur together as the product $(1-d)\rho$ and cannot be distinguished using tag recovery data alone; therefore only one of the terms may be estimated and the other must be fixed. The chronic tag loss rate λ is similarly confounded with the natural mortality rate M , but independent estimation becomes possible (in principle at least) when abundance indices are available in addition to the tagging data. The model in Table 5 also allows for the possibility that the effective fishing mortality on the tagged population may differ from that on the untagged population by use of pre-mixing adjustment factors γ_{ty} . The values of γ_{ty} can be difficult to distinguish precisely from F when only tag recovery data are available (Hoenig et al., 1998), but the task is made easier when catch and abundance data are included because they too contain information on F .

An important adjustment to consider has to do with the fact that the VPA accounting is by year, so the value of M and F are essentially averages over the course of the year. Tag releases, on the other hand, tend to occur at various times within a year, sometimes before and sometimes after the periods of most intense fishing. Consider, for example, a group of fish that were tagged and released at the end of the 7th month after the major fishery was over. In that case the tagged fish would be exposed to five months of tag loss and natural mortality, but none would be recaptured. If no adjustments are made to account for this, the model would erroneously interpret the lack of recaptures for that year as no fishing for the year, when in fact the fishing pressure could have been very high. The fishing fraction parameter ϕ in Table 5 adjusts the fishing mortality rate parameter by the fraction of the total fishing pressure exerted during the year that was experienced by tags. This requires information on the seasonality of the fisheries beyond that which is normally required by VPA (see discussion on line 34 of the control file). Note that this problem only occurs for the year when the fish were first tagged. In subsequent years the tagged population will face the full cycle of fishing such that $\phi=1$.

Finally, in some cases with two stocks, the origin of the tagged fish may not have been determined. In the case of the diffusion model this is unimportant because the two stocks are defined by management zone (rather than by origin)--fish born in zone 1 and tagged in the zone 2 are considered to be members of the zone 2 stock and vice versa. In the case of the overlap model, where the two stocks are defined by origin, fish that have not been identified by stock of origin are assumed to be tagged in proportion to their estimated relative abundance in each management zone, i.e.,

$$R_{s,i} = R_i \frac{\bar{T}_{s,\kappa,\alpha,\psi} \bar{N}_{s,\alpha,\psi}}{\sum_s \bar{T}_{s,\kappa,\alpha,\psi} \bar{N}_{s,\alpha,\psi}} \quad (2.27)$$

where R_i is the number of reported recaptures from the i th cohort (tagged in zone k), $R_{s,i}$ is the calculated number of reported recaptures from the i th cohort that belong to stock s , and the \bar{N} and

\tilde{T} values are the estimates from the VPA.

Table 5. Model for tag recoveries. The subscript i denotes a unique group (cohort) of tag releases distinguished by the zone (κ), year (ψ) and age (α) of release. The remaining subscripts are as defined previously (stock: s , zone: k , year: y , age: $a = \alpha + y - \psi$).

Equations and variables	Description
$-\log P(\tilde{r}, \Theta) =$ $-\frac{1}{\omega_i^2} \left[\left(R_i - \sum_{k,y} (r_{i,k,y} + r_{i,k,0}) \right) \ln \left(1 - \frac{\sum_{k,y} (\hat{r}_{i,k,y} + \hat{r}_{i,k,0})}{R_i} \right) \right.$ $\left. + \sum_{k,y} \left(r_{i,k,y} \ln \frac{\hat{r}_{i,k,y}}{R_i} + r_{i,k,0} \ln \frac{\hat{r}_{i,k,0}}{R_i} \right) \right]$	Multinomial negative log-likelihood for tag recoveries from all cohorts
R_i	number of tag releases for cohort i
$r_{i,k,y}$	number of recaptures reported by the fishery of cohort i in zone k during year y
$r_{i,k,0}$	number of survivors that are detected at the end of the experiment (e.g., from satellite signals)
$\hat{r}_{i,k,y} = \rho_{i,k} n_{i,k,y} \frac{f_{i,k,a,y}}{Z_{i,k,a,y}} (1 - e^{-Z_{i,k,a,y}})$	expected number of recaptures reported by the fishery of cohort i in zone k during year y
$\hat{r}_{i,k,0} = \rho_{i,0} \tilde{n}_{i,k,y} e^{-Z_{i,k,a,y}}$	expected number of survivors that are detected at the end of the experiment (e.g., from satellite signals)
ω_i	weighting factor used to discount or emphasize tag data
$Z_{i,k,a,y} = f_{i,k,a,y} + (M_{k,a,y} + \lambda_{iy})(1 - \tau_{iy})$	total loss rate of cohort i
$f_{i,k,a,y} = \phi_{i,y} \gamma_{i,y} F_{k,a,y}$	fishing mortality rate of cohort i
$\gamma_{i,y}$	pre-mixing adjustment for cohort i
$\phi_{i,y}$	fishing fraction in first year of release (= 1 for $y > \psi$)
$\lambda_{i,y}$	chronic tag loss rate of cohort i
d_i	immediate tag loss of cohort i
τ_{iy}	release date of cohort i (= 0 for $y > \psi$)
$\rho_{i,k}$	fraction of recaptured tags reported by fisheries in zone k
$\rho_{i,0}$	fraction of surviving tags that are detected at the end of the experiment (e.g., from satellite signals)

Table 5 Continued on next page

Table 5 continued. Model for tag recoveries. The subscript i denotes a unique group (cohort) of tag releases distinguished by the zone (κ), year (ψ) and age (α) of release. The remaining subscripts are as defined previously (stock: s , zone: k , year: y , age: $a = \alpha + y - \psi$).

Overlap tag attrition model

$$n_{i,k,y} = \begin{cases} R_i(1 - d_i) & y = \psi, k = \kappa \\ \sum_s \tilde{T}_{s,k,a,y} \tilde{n}_{i,s,y} & y > \psi \end{cases} \quad \begin{array}{l} \text{number of survivors with tags from cohort } i \text{ in zone } k \text{ at} \\ \text{start of year } y \end{array}$$

$$\tilde{n}_{i,s,y+1} = \begin{cases} R_{i,s}(1 - d_i)e^{-Z_{i,\kappa,a,y}(1-\tau_{i,y})} & y = \psi \\ \tilde{n}_{i,s,y} \sum_k \tilde{T}_{s,k,a,y} e^{-Z_{i,k,a,y}} & y > \psi \end{cases} \quad \begin{array}{l} \text{number of survivors with tags from cohort } i \text{ and stock } s \text{ at} \\ \text{start of year } (R_{i,s} \text{ is the number of releases by stock}) \end{array}$$

Diffusion tag attrition model

$$n_{i,k,y+1} = \begin{cases} R_i(1 - d_i)e^{-Z_{i,\kappa,a,y}(1-\tau_{i,y})} & y = \psi, k = \kappa \\ \sum_j T_{j,k,a,y} n_{i,j,y} e^{-Z_{i,j,a,y}} & y > \psi \end{cases} \quad \begin{array}{l} \text{number of survivors with tags from} \\ \text{cohort } i \text{ in zone } k \text{ at start of year } y \end{array}$$

BIAS AND UNCERTAINTY

In any statistical analysis there is a tradeoff between precision and accuracy. Generally, adding parameters to a model should improve its ability to reflect the dynamics of the population (making it more accurate). However, if the data are limited, adding too many parameters will lead to a situation where the estimates are highly imprecise, i.e., many combinations of parameter values can explain the data equally well (give the same posterior/likelihood values). In this regard, one of the challenges to the investigator is to determine whether the additional parameters improve the model's ability to fit the data in a statistically significant way. If not, the new model is generally rejected in favor of the simpler model (see the section on model selection below). Subsequent inferences are usually made under the implicit assumption that the selected model is correct and the parameter estimates obtained by the nonlinear optimization are unbiased.

Maximum likelihood estimators, where the objective function being minimized is the negative log-likelihood, are asymptotically unbiased under fairly general conditions (Hoel et al., 1980). In principle, asymptotically efficient estimates of the covariance of the parameters may be obtained from the inverse of the Hessian matrix (matrix of second derivatives), allowing one to quantify the precision with which the parameters were estimated. However, in most stock assessments the number of data points available is seldom much larger than the number of parameters, so the asymptotic theory may not apply and a substantial bias can be incurred in both the point estimates and their corresponding Hessian-based covariances. Moreover, the asymptotic covariance matrix is not strictly given by the inverse of the expected Hessian matrix when penalties, priors or process errors are imposed (Seber and Wild, 1989).

The best way to ascertain the bias and variance of any given estimator is to apply that estimator to a large number of random samples from a population with known distributional characteristics. In practice of course, one does not know the distribution of the population in question and, in many cases, does not have the capacity to resample that population multiple times. Therefore, one must turn to various approximations, such as Box's (1971) linear approach, which involves computing both first and second derivatives (see discussion in Gavaris, 1993). I am unaware of an analogous technique for determining the bias of the corresponding estimators for the covariance of the estimates. Moreover, it is unclear how the Box (1971) approach would apply when the objective function is not simply the negative log-likelihood, but also includes process error, priors and other penalties.

An alternative way of determining the bias and covariance of the point estimates is by use of a technique known as bootstrapping. Punt and Butterworth (1993) found that, on average, the estimates of variance from the inverse of the Hessian matrix were about the same as the estimates from bootstrapping, but it is not clear that this will always be the case for the reasons discussed above. In practice, the bootstrap estimator may often perform better than the inverse of the Hessian matrix because the latter is based only on the shape of the negative log-likelihood surface in the vicinity of the minimum, which of course presupposes that the true minimum has in fact been found and is based on a single configuration of the data. The bootstrap estimator, on the other hand, implicitly takes into account the possibility that the minimum may not always be found and the sensitivity of the solution to perturbations in the data.

Bootstrap procedures

The concept behind bootstrapping is that the distribution of values from any given random sample is the best guide to the distribution of values in the population the sample came from. Therefore, the next best thing to resampling a population is resampling the sample. Regression analyses may be bootstrapped either by resampling the regression residuals or by resampling the data (Manly, 1998, p 167), but the latter is seldom an option for VPA tuning approaches because there is only one input value for each index for any given time unit (year). Accordingly, most of the literature on bootstrapping VPA's involves resampling the residuals.

There are two principal ways of bootstrapping from regression residuals: parametrically by sampling from a known distribution with a prescribed variance and non-parametrically by sampling directly from the residuals of the model fit (random draws with replacement). In either case, the sampled residuals are added to the model expectation (not the observations themselves) to generate a pseudo-data set. The model is then refit to each of B number of pseudo data sets and the results used to construct the sampling distributions of the estimators in question.

VPA-2BOX can presently conduct parametric and non-parametric bootstrap analyses of the indices of abundance and indices of mortality. In principle, one should choose the parametric approach only if the distribution of the statistic being bootstrapped is known (e.g., lognormal with a variance of 0.2). Punt and Butterworth (1993) found that the non-parametric and parametric bootstrap procedures gave similar results in their examples, but Freedman and Peters (1984) and others have found that the non-parametric approach generally gives less-biased results.

The parametric bootstrap entails the following steps:

- (1) Fit the model to the actual data;
- (2) Create new "bootstrap" indices, ${}^b I_{i,k,y}$, by drawing at random with replacement from normal or lognormal distributions with mean and standard error equal to the model expectations of I (or $\ln I$) and σ (i denotes a particular series for zone k and y denotes a particular year):

$${}^b I_{i,k,y} = \begin{cases} \hat{I}_{i,k,y} + {}^b \eta_{i,k,y} & \text{additive} \\ \hat{I}_{i,k,y} e^{b \eta_{i,k,y}} & \text{multiplicative} \end{cases}$$

$${}^b \eta_{i,k,y} \sim N\{\mu_{i,k,y}, \sigma_{i,k,y}\} \quad (2.28)$$

- (3) Fit the model to the bootstrap data set created in step 2;
- (4) Repeat steps 2 and 3 until the required number of replications is obtained.

A potential problem with this approach occurs when the standard error term σ is not known and is instead replaced by a value estimated in the model, $\hat{\sigma}$. In many cases the estimated value $\hat{\sigma}$ may be smaller than the true value, causing the bootstrap residuals π to be smaller than they should be. As a result, the bootstrap could indicate that the model parameters are well-estimated when in fact they are poorly estimated and $\hat{\sigma}$ just happens to be too low. For this reason the parametric approach is not recommended unless one is confident in the values being used for σ .

The non-parametric bootstrap is similar in principle, but steps 1 and 2 become:

- (1) Fit the model to the actual data and store the residuals—

$$\varepsilon_{i,k,y} = \begin{cases} I_{i,k,y} - \hat{I}_{i,k,y} & \text{additive} \\ \ln I_{i,k,y} - \ln \hat{I}_{i,k,y} & \text{multiplicative} \end{cases} \quad (2.29a)$$

- (2) Create new “bootstrap” indices of abundance, ${}^b I_{k,y}$, by drawing at random with replacement from the index-specific sets of residuals–

$${}^b I_{i,k,y} = \begin{cases} \hat{I}_{i,k,y} + {}^b \eta_{i,k,y} & \text{additive} \\ \hat{I}_{i,k,y} e^{b \eta_{i,k,y}} & \text{multiplicative} \end{cases}$$

$${}^b \eta_{i,k,y} \in \{\varepsilon_{i,k,y}\} \quad (2.29b)$$

As it stands, the approaches above tend to perform poorly when the number of estimated parameters is large relative to the number of data points. In such cases the model may be able to provide a near perfect fit to the data with many different combinations of parameter values. The small residuals in turn will lead to bootstrap data sets that are essentially identical to the original data set, ultimately leading to artificially small bootstrap estimates of the parameter variances. The reason for this is easy to see in the context of a regression involving a single index of abundance. In that case the variance of the bootstrap residuals (being random draws from the finite population η_y) is

$$VAR({}^b \eta_{i,k,y}) = \frac{n_{\Theta} - n_D}{n_D} \sigma_{i,k}^2 \quad (2.30)$$

where n_{Θ} is the number of estimated parameters and n_D is the number of observations. Hence, the variance of the residuals is too small by a factor of $1 - n_{\Theta}/n_D$.

Tukey (1987) and Stine (1990) recommend ‘fattening’ the residuals from linear regressions by dividing each by a factor $(1 - n_{\Theta}/n_D)^{1/2}$. However, this may not be the correct adjustment in more complicated regressions, such as when more than one index of abundance is being used. Efron and Tibshirani (1993) suggest that the random variability in the estimates of variance is more important than the bias caused by factors such as $(1 - n_{\Theta}/n_D)^{1/2}$ and that the problem is not too worrisome unless n_{Θ}/n_D is greater than about 0.25. Nevertheless, in the case of multiple indices it is reasonable to take an intermediate position and at least adjust for the number of estimated parameters unique to each index (optional in VPA-2BOX). Thus, the residuals in step 2 of the non-parametric procedure are inflated as

$${}^b \eta_{i,k,y} \in \left\{ \frac{\varepsilon_{i,k,y}}{\sqrt{1 - n_{\Theta,i,k}/n_{D,i,k}}} \right\} \quad (2.31)$$

where $n_{D,i,k}$ and $n_{\Theta,i,k}$ respectively indicate the number of observations and estimated parameters unique to index i (namely the catchability coefficients $q_{i,k}$ and variances $\sigma_{i,k}$). In point of fact the above adjustment is still too low owing to the bias caused by not adjusting for the parameters shared by each index, but it is not clear how such further adjustments should be accomplished. Therefore, a warning message is output by program VPA-2BOX whenever $n_{\Theta}/n_D > 0.25$.

Finally, it may sometimes happen that the variance of an index differs from one year to the next owing to differences in sample size or other factors and that the investigator may wish to deal with this heterogeneity explicitly. With the parametric approach this is straightforward and no further modifications are required, but the non-parametric approach will lead to homoscedastic bootstrap samples. In such cases residuals η may be transformed to standard normal deviates by

dividing by the standard errors. Thus, step 2 of the nonparametric procedure are replaced by the parametric approach below:

$${}^b I_{i,k,y} = \begin{cases} \hat{I}_{i,k,y} + {}^b \eta_{i,k,y} & \text{additive} \\ \hat{I}_{i,k,y} e^{b \eta_{i,k,y}} & \text{multiplicative} \end{cases}$$

$${}^b \eta_{i,k,y} \in \left\{ \frac{\varepsilon_{i,k,y}}{\sigma_{i,k,y}} \right\} \quad (2.32)$$

Note: This standardization procedure is an *ad hoc* approach employed to account for heterogeneous variances among years; it would not be necessary if multiple observations were available for each year, in which case one could simply resample the data.

Bootstrap estimates of bias and variance

The expected value of an estimator applied to bootstrap pseudo-data sets is, by definition, the point estimate associated with original model fit. Therefore, if the estimator is unbiased, the mean of the estimates obtained from each bootstrap data set should be the same as the original point estimate. Otherwise, the difference between the mean and the point estimate is taken to be an indicator of the bias, i.e., the difference between the point estimate and the true value. Program VPA-2BOX computes the bias and variance of any given point estimate $\hat{\phi}$ via the estimators suggested by Efron (1990):

$$\text{bias}_B(\hat{\phi}) = \frac{\sum {}^b \phi}{B} - \phi(\bar{\mathbf{I}})$$

$$V_B(\hat{\phi}) = \frac{\sum_{b=1}^B ({}^b \phi - \frac{\sum {}^b \phi}{B})^2}{B-1} \quad (2.33)$$

where $\phi(\bar{\mathbf{I}})$ denotes the estimate obtained by fitting the model to the average of the bootstrap data sets (i.e., $\bar{I}_{ky} = \sum {}^b I_{ky}/B$). Note that some investigators prefer to use the median of the bootstraps to calculate the bias rather than the mean because it is less sensitive to rare, large aberrations:

$$\text{bias}_B(\hat{\phi}) = \text{median}\{{}^b \phi\} - \hat{\phi} \quad (2.34)$$

This has not yet been automated in VPA-2BOX, but can be computed from the binary output files.

Frequently one may wish to correct the statistic $\hat{\phi}$ in order to make it less biased:

$$\phi_{corrected} = \hat{\phi} - \text{bias}_B(\hat{\phi}) \quad (2.35)$$

Note that $\phi_{corrected}$ is not equivalent to the median or mean of the bootstraps, a fairly common

mistake (in fact, if $\hat{\phi}$ is biased then, by definition, the bootstrap mean and median are even more so). Efron and Tibshirani (1993) point out that bias-correcting is a dangerous affair; even if $\phi_{corrected}$ is less biased than $\hat{\phi}$, its standard error may be much greater. For this reason they suggest that the bias statistic is most useful as a diagnostic tool and that it is safer to use $\hat{\phi}$ than $\phi_{corrected}$ if the bias is small relative to the estimated standard error. On the other hand, a large bias relative to the standard error may indicate that the statistic $\hat{\phi}$ is too poorly determined to be useful.

How many bootstraps?

Porch (1999) found that the estimates of bias and coefficients of variation converged to stable values as the number of bootstrap replicates increased to 200 or more, but even as few as 50 replicates gave reasonably good estimates. Therefore, there seems to be little to gain from using more than 200 replicates and as few as 50 may be adequate when computing time is limited. This conclusion seems to be typical of many bootstrap problems (Efron, 1987; Punt and Butterworth, 1993; Smith and Gavaris, 1993), although it is possible that very large data sets may require more replicates.

Cautions

It must be emphasized that the bootstrapping procedure only addresses errors that arise specifically because a nonlinear model is being applied to a relatively small data set. Other biases may occur when the error structure of the dependent variables is mis-specified or because the independent variables (e.g., the predicted indices of abundance) are treated as though they were error-free when in fact they do have error. In case of the latter, for example, the inverse of the Hessian matrix will usually lead to underestimates of variance (page 496 in Seber and Wild, 1989). Still other uncertainties may arise because the catches (and even some of the model parameters) are treated as though they were perfectly known. In the case of catches, VPA-2BOX does include an option to generate pseudo-catch-at-age data from the observed values assuming a particular distribution type and variance (or CV) analogous to the parametric bootstrap discussed above. This approach is useful for characterizing the additional uncertainty attributable to the catches, but is biased because it is centered on the data (see Poole et al., 1999) and should not be used for bias-correction.

Retrospective pattern analysis

One means of detecting some potential biases not discernable from the bootstrap is retrospective pattern analysis. The idea here is to determine if there is a systematic trend in the model estimates when the same model is applied to successively shorter data sets (retrospective “peals”). For example, one might compare the fishing mortality estimates when the model was applied to the data from 1975 to 1999, then again to the data from 1975 to 1998, then again to the data from 1975 to 1997, and so on. A systematic trend in the F estimates could suggest a number of

possible biases, including mis-specification of some of the fixed parameters (often the natural mortality rate), consistent under or over-reporting of catches, indices of abundance that do not represent abundance, and ageing errors.

Various *ad hoc* procedures have been developed to try to adjust the estimates to account for potential biases indicated by the retrospective pattern (e.g., ICCAT 1995, page 48), however, the statistical properties of these adjustments are unclear. Certainly no adjustment should be attempted if the standard errors of the estimates are larger than the indicated retrospective bias because the bias adjustment itself is likely to be poorly estimated. Moreover, retrospective patterns usually dissipate as more data are added owing to the convergent properties of the VPA, so it is unclear how the estimates farther back in time should be adjusted. For these reasons I recommend retrospective analyses primarily as a diagnostic tool to help identify failings in the model or data.

MODEL SELECTION

Model selection is the process of identifying the model that provides the ‘best’ description of a given set of data from a suite of competing models. There are two important caveats that must be addressed at the outset. First, none of the methods discussed below are useful for comparing models applied to different data sets—they all assume the same data are used. Second, one should determine the most theoretically plausible model structures *a priori* and then limit the comparisons to these. The strategy should never be to try every model under the sun and then select the one that provides the best fit, the danger being that the selection process will identify a model that explains much of the variation in the data but has little connection to reality. Draper (1995), for example, has shown that traditional methods of model selection can lead to models with apparently strong predictive power even for randomly generated data. This problem is analogous to what happens in step-wise regression procedures with a very large number of potential covariates: the probability of finding some combination of covariates that accounts for the variation in the data may be great even when none of the covariates are related to the data (Manly, 1998).

Hypothesis tests

Hypothesis tests are often used to select the best approximating model. If the models are fitted by maximum likelihood, for example, the standard likelihood ratio test may be used to determine if the addition of n number of parameters to a previous model significantly improved the fit to the data. The test statistic,

$$-2\log \left\{ \frac{P(\mathbf{D} | \Theta_2, \mathbf{X}_2)}{P(\mathbf{D} | \Theta_1, \mathbf{X}_1)} \right\}, \quad (2.36)$$

is chi-square distributed with n degrees of freedom (Hoel et al., 1981). Thus, the augmented model (model 2) would be judged significantly better than its simpler predecessor (model 1) if the test statistic was greater than the value from the chi-square distribution with n degrees of freedom and the prescribed probability level (say 5%).

The primary drawback to likelihood ratio tests is that they are made with the understanding

that one model is a reduced version of the other. Accordingly, models with different distributional assumptions about the data (e.g., lognormal versus normal distributed indices) cannot be compared with likelihood ratio tests. For this reason it is often preferable to work with information criteria, which do not require the competing hypotheses to be nested.

Information criteria

Perhaps the most familiar information criterion is Akaike's (1973) AIC metric,

$$AIC = -2 \log P(D|\Theta, \mathbf{X}) + 2n_{\Theta} \quad (2.37)$$

and the small sample bias-adjusted version AIC_c (Hurvich and Tsai, 1995),

$$AIC_c = -2 \log P(D|\Theta, \mathbf{X}) + 2n_{\Theta} \left(1 + \frac{n_{\Theta} + 1}{n_D - n_{\Theta} - 1} \right) \quad (2.38)$$

The variables n_{Θ} and n_D represent the number of estimated parameters and number of data points, respectively. The philosophy behind the AIC is that reality is high-dimensional, possibly requiring infinitely many parameters to describe it. Model selection is therefore seen as a process of identifying the best approximating model. The dimension of this 'best' approximation would be expected to be low when the data are sparse and increase with the quantity and quality of the available data. The 'best' model is taken to be that with the lowest AIC (AIC_c) value.

Another metric that is sometimes used is the Bayes Information Criterion BIC,

$$BIC = -2 \log P(D|\Theta, \mathbf{X}) + n_{\Theta} \log(n_D) \quad (2.39)$$

The philosophy underlying the BIC criterion is that reality is low dimensional and model selection is seen as the process of identifying the true model (the usual Bayesian approach). As with AIC, the 'best' model is taken to be that with the lowest BIC value. An interesting implication of the BIC is that it should be proportionately harder to add parameters as more data become available, which seems somewhat counterintuitive.

Burnham et al. (1994) and Buckland et al. (1997) favor the AIC/ AIC_c approach over BIC, arguing that reality is rarely low-dimensional. VPA-2BOX includes both metrics, but I do not generally recommend the BIC because the systems one is likely to analyze with VPA are likely to be much more complex than any model that can be supported by the data. I generally use the AIC_c metric rather than AIC because of the relatively small ratio of data to parameters in most VPA applications.

It is important to note that the negative log-likelihood expression $-\log P(D | \Theta, \mathbf{X})$ used in criteria (2.37 -2.39) must include all constant terms if comparisons are to be made between models with different distributional assumptions (e.g., normal versus lognormal distributed indices of abundance). An alternative form of the criteria uses the model *deviance* in place of $-2 \log P(D | \Theta, \mathbf{X})$, however it is unclear how this form could be applied when the variance terms are estimated. (Model deviance is twice the difference between the negative loglikelihood function associated with the model fit and the negative loglikelihood that would be obtained if the model were able to

fit the data perfectly.)

Bayes factors

A related point to keep in mind is that the AIC, BIC, likelihood ratio tests, and related criteria are based on likelihood theory and should not be applied when Bayesian priors and penalty functions are incorporated unless the priors and penalties are relatively uninformative (and so have little influence on the likelihood function) or can in some sense be interpreted as data. Otherwise, a discrete set of Bayesian models should be compared by contrasting their posterior densities,

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D)}, \quad (2.40)$$

where H_i denotes the i 'th of several alternative models to be tested, $P(D|H_i)$ is the marginal density of observing the data if hypothesis H_i were true, $P(H_i)$ is the prior representing expert opinions on probability that H_i is true relative to the competing hypotheses, and $P(D)$ is the unknown probability density of the data. If the same data are used for all models, $P(D)$ is a constant that can be ignored. In that case the ratios of $P(H_i|D)$ for two competing models (H_1 and H_2) may be compared:

$$\frac{P(H_2|D)}{P(H_1|D)} = \frac{P(D|H_2)P(H_2)}{P(D|H_1)P(H_1)}. \quad (2.41)$$

Model H_2 would then be interpreted as 'better' than H_1 if the ratio were greater than 1.

Unfortunately, the computation of the ratio $P(D|H_2)/P(D|H_1)$, known as the Bayes factor, has a number of practical limitations. One of the most important of these limitations is that they cannot reasonably be applied to noninformative priors (See Gelman et al. 2000, pp. 175-177). Another is that their computation requires performing the multidimensional integration

$$P(D | H_i) = \int_{\Theta} P(D | \Theta, X_i) P(X_i | \Theta) P(\Theta) d\Theta, \quad (2.33)$$

which may demand considerable computing time. A number of alternatives to the Bayes factor have been suggested, including the controversial 'posterior Bayes factor' suggested by Aitkin (1991) and appropriated (in one of its limiting forms) for stock assessments by Fournier et al., (1998). Each such alternative has its own peculiarities and, at this point, none have been programmed into VPA-2BOX.

Chi-square discrepancy statistic

One useful check of model performance that can be applied in both Bayesian and non-Bayesian constructs is the Chi-square discrepancy statistic (Gelman et al., 1995; BFT, 2001):

$$\chi^2 \text{ discrepancy} = \sum_j \frac{(d_j - f[\hat{\Theta}, \mathbf{X}_j])^2}{V(d_j)}, \quad (2.34)$$

where $f[\hat{\Theta}, \mathbf{X}_j]$ is the model expectation of d_j given the estimated parameter vector and $V(d_j)$ is the variance. This statistic is approximately Chi-square distributed with $n_D - n_\Theta$ degrees of freedom, so the probability p of observing a value of χ^2 that exceed the calculated value may be read from a standard Chi-square table. The idea here is to assess whether the observed discrepancies between the data and model predictions could have arisen by chance under the model's own assumptions. Major failures of the model should lead to either very high p -values (say, over 0.99, in which case the model probably has too many parameters) or very low p -values (say, less than 0.01, in which case the model is inconsistent with the data). The chi-square statistic can also be used to identify changes in the model, other than error structure, that augment its performance and bring the p -values to a reasonable range. It should not be used to compare different assumptions about error structure because it would, by its very nature, always favor weighted least-squares estimators. Note that VPA-2BOX computes this statistic for the indices of abundance/mortality only; it does not compute it for tag-recapture data.

NUMERICAL OPTIMIZATION

The section on parameter estimation discussed how the ability of a model to fit the data could be expressed as a single objective function Ψ that depends on a number of independent variables (parameters). The goal was to find the values of those variables where Ψ takes on its lowest possible value, i.e., the minimum. In certain special cases, such as when the model is linear in its parameters, the minimum can be found exactly by setting the derivatives of Ψ with respect to the parameters equal to zero and solving the resulting system of equations analytically. In the case of nonlinear models, however, it is seldom possible to find a solution analytically and numerical procedures must be employed.

Numerical optimization methods essentially amount to trying a large number of different combinations of parameter values and finding the combination with the optimal (in this case lowest) value of Ψ . Ideally, one would compare the values of Ψ obtained with every possible combination of feasible parameter values (or increments of values in the case of continuous variables). Unfortunately, the number of function evaluations (parameter combinations) required for this strategy increases geometrically with the number of values (n) evaluated and parameters considered (p): evaluations = n^p . Thus, the use of even a very coarse grid with only ten values for each parameter would require ten million function evaluations for seven parameters (a fairly modest number). At this writing, even the fastest PC processors require several hours to compute a typical VPA-type objective functions ten-million times. Grid searches with any more than 7 parameters or ten values can therefore be seen to be completely impractical. For this reason a number of algorithms have been developed to reduce the number of parameter combinations that must be evaluated in an intelligent way.

There are no foolproof numerical methods that are universally applicable to multivariate problems; any given algorithm is prone to be 'fooled' in certain situations. Perhaps the most common failure involves mistaking a local minimum in the solution surface for the overall (global)

minimum. It is beyond the scope of this manual to discuss the pros and cons of the innumerable number of search algorithms that have been developed. Interested readers are referred to any standard text on the subject (e.g., Press et al., 1992).

A two-step approach is used in VPA-2BOX. The first step involves the OSSRS algorithm of Sheela (1979), which was designed to rapidly converge on the general vicinity of the global minimum. Once in that vicinity, however, the OSRSS algorithm converges very slowly. At that point VPA-2BOX automatically switches to the Nelder-Mead simplex routine AMOEBA of Press et al. (1992). This routine has the advantage of not requiring analytical or numerical derivatives, as do quasi-Newton and related algorithms.

The AMOEBA algorithm is restarted multiple times to avoid being fooled by local minima. New initial vertices are selected for each such restart using the formula

$$\theta_j = \theta_{0j} e^{\sigma\eta\delta} \quad (i,j = 1, 2, \dots, p) \quad (2.35)$$

where θ_{0j} is the value of the j^{th} parameter at the presumed minimum, θ_{ij} is the value of the j^{th} coordinate (parameter) in the i^{th} vertex of the initial simplex, η is a standard normal variate, σ is a user-prescribed standard deviation, and δ is equal to one if i equals j and zero otherwise. Subsequent 'restarts' continue until a prescribed number of consecutive sets of parameter estimates differ by less than one percent.

Again, no algorithm is foolproof and the combination of OSSRS and AMOEBA used here is no exception. While this combination has fared well in extensive comparisons with several other VPA programs that use different search algorithms (usually finding the same minimum and often a lower one), it is always good practice to make multiple runs with different starting guesses to be certain that the global minimum has been obtained.

3. INPUT FILES

Program VPA-2BOX always requires three ASCII text files: 1) a control file that designates the type of model to be run, 2) a parameter file that specifies how the parameters should be estimated, 3) a data file containing the catches, weights and indices of abundance. Two other text files may also be required if mean weight or tag-recovery information is to be used. A template for each of these files appears in an appendix. The explanation for each line item in the appropriate appendix is then given under the corresponding heading below.

Each of the files is read in free-format fashion, that is, the input data need not be in any particular column. However, if there are multiple entries on one line, they must be entered in the proper sequence. For example, catch-at-age information must be entered in the sequence {year, catch-at-youngest-age, catch-at-youngest-age+1, ..., catch-at-oldest-age}. Comment lines may be inserted anywhere in the file provided they are preceded by a # symbol in the first column. Text placed after a # symbol will not be read. Comments may also appear at the end of the line after all required inputs. For example, the following comments are acceptable:

```
# the catches of ages 1 and 2 in 1995 are unknown (1996 catches are used)
1995 13456 23454 68906 80001 25501
1996 13456 23454 78906 90001 24565
```

or

```
1995 13456 23454 68906 80001 25501 the catches of ages 1 and 2 in 1995
1996 13456 23454 78906 90001 24565 are unknown (1996 catches are used)
```

In contrast, the comment forms below are unacceptable (the first is improperly indented and the second has comments in the middle of the data stream):

```
  # the catches of ages 1 and 2 in 1995 are unknown (1996 catches are used)
1995 13456 23454 68906 80001 25501
1996 13456 23454 78906 90001 24565
```

or

```
1995 13456 23454 the catches of ages 1 and 2 in 1995 68906 80001 25501
1996 13456 23454 are unknown (1996 catches are used) 78906 90001 24565
```

CONTROL FILE : See Appendix 1

This is the only file VPA-2BOX will prompt the user for. It contains the file names of all the other input files. Among other things, it specifies the type of model being used (one stock or two, overlap versus diffusion), imposes certain constraints on the model (e.g., a stock recruitment penalty), controls the performance of the search algorithm, and conducts bootstrap or retrospective analyses. What follows is a line by line explanation of the entries in Appendix 1.

Line Explanation

Comments

1 - 13 Comments preceded by the # symbol in the first column.

Names (must be placed within single quotes)

14 The title of the run, which may be up to 50 characters long.

15 The name of the file with the catch, weight and index data (50 characters).

16 The name of the file with the parameter specifications (50 characters).

17 The name of the output file with the results (50 characters).

18 The name of the output file with the parameter estimates (50 characters).

19 The name of the output file with the results in a spreadsheet friendly format (50 char.)

20 The name of the file with the tag-recovery data (50 characters). The name of this file is immaterial if the tagging data switch on line 34 is set to 0, but some kind of name must still be entered

Model type

24 Number of zones or stocks being considered (either 1 or 2)

25 The class of model being used (overlap or diffusion, see Table 2). If only 1 stock is considered then the overlap and diffusion models are exactly the same.

Tagging data controls

34 Fourteen specifications must be entered here. The first is the tagging data switch: a value less than or equal to 0 tells the program to ignore the tag-recovery information and the other 13 entries can be left out. Otherwise, a positive value tells the program which likelihood function to use. A value of 3 invokes the multinomial likelihood listed in Table 5, but other likelihood functions can be invoked (1=lognormal, 2=normal, 3=Poisson). Note that if the lognormal or normal options are invoked, the variance functions are obtained from values given in the tag recovery file (and appendix 3). The second entry on this line is interpreted as a multiplicative weighting factor ω that controls the influence that the tagging data has on the estimation procedure. The objective function minimized in program VPA-2BOX is

$$\Psi(\Theta, X) = L(\vec{I}) + \omega L(\vec{r}) + \text{priors and other terms} \quad (3.1)$$

where $L(\vec{I})$ and $L(\vec{r})$ are the likelihood functions for the indices and tag recoveries, respectively. Thus, a value of ω greater than 1.0 magnifies the influence of the tagging, whereas a value less than 1.0 decreases its influence.

The next 12 entries give the relative amount of fishing E_j expected to occur in each month j . These are used to calculate the fraction of the annual fishing pressure ϕ_i each cohort of tags was exposed to during the calendar year they were released (see Table 5):

$$\phi_i = \frac{(\delta_i - \tau_i)E_{\delta_i} + \sum_{j > \delta_i}^{12} E_j}{\sum_{j=1}^{12} E_j} \quad (3.2)$$

where τ_i is the time of release in terms of elapsed months and d_i is the month the release occurred (an integer value from 1 to 12). The monthly effort indices E can be in any units so long as they are consistent. Note, however, that the above formula implicitly assumes that the E values are fairly constant from year to year. If they are not, the release dates should be adjusted outside program VPA-2BOX and the monthly values for E all set to the same value (say 1) so that VPA-2BOX does not modify them.

Search algorithm controls

- 38 The seed for the random number generator, which can be any negative integer. You would want to change this if you were running separate bootstrap analyses (perhaps on different machines) and planned later to combine the results into one grand bootstrap analysis. If you do not change the random number seed the bootstrap runs will be identical whether you run them at different times or on different machines.
- 39 The maximum number of Amoeba simplex restarts (see Numerical Optimization section) can be fixed to prevent the algorithm from running too long. Most applications will converge within 10 restarts, however some two-stock analyses with tagging data can take 50 or more. Usually, this many restarts does not lead to a noticeable improvement in the model fit and is the result of one or two parameters being poorly determined. Therefore one may wish to cap the number of restarts at the point where noticeable improvements cease, particularly when doing bootstrap analyses.
- 40 The number of consecutive restarts over which the parameter estimates must not vary by more than one percent (see discussion of equation 2.35). Usually three is enough, but one may wish to use four or five with especially difficult problems.
- 41 This is the standard deviation parameter used in equation (2.35). It controls the extent to which the restart vertices span the parameter space. If the value is large, the initial simplex at each restart will be large, which will lead to rapid convergence initially, but slower convergence when the algorithm gets closer to the minimum. Conversely, small values tend to enhance the performance in the vicinity of the minimum, but the initial rate of convergence will be slow. Experience has shown that, for VPA analyses, the best tradeoff usually occurs with values of 0.4 - 0.5. However it is always useful to try a range of values to see which works best for any particular application

Index weighting controls

- 45 This option allows one to divide each index of abundance by its arithmetic mean (to do so enter any positive nonzero value). This scales all of the indices to an order of 1. It is important to do this if you are assuming the indices are normally distributed and wish to weight all of the indices the same (choice -1 at line 46 below). Otherwise, the objective function will be dominated by the indices with the largest values. If you are assuming the indices are lognormally distributed or are allowing each series to be weighted by different variances, then no scaling is necessary.
- 46 This option allows one to substitute a default value for the index variance inputs discussed in connection with line 63 of the data file. The available options are:

<u>Option</u>	<u>Result</u>
0	no action, program uses the INPUT _{iky} values specified in the data file
+ value	any positive nonzero input is read as a default CV, and $\sigma_{i,k,y}^2 = (value * \hat{I}_{i,k,y})^2$
- value	any negative input is read as a default standard error, and $\sigma_{i,k,y}^2 = (value)^2$
999	a value with the integer part equal to 999 causes the variance for each index to be estimated by the concentrated maximum likelihood formulae

$$\sigma_{iky}^2 = \sqrt{\frac{\sum_y (I_{iky} - \hat{I}_{iky})^2}{\sum_y 1.0}} \quad \text{(Normal distribution)} \quad (3.3)$$

$$\tilde{\sigma}_{iky}^2 = \sqrt{\frac{\sum_y (\ln I_{iky} - \ln \hat{I}_{iky})^2}{\sum_y 1.0}} \quad \text{(Lognormal distribution)} \quad (3.4)$$

This option is included for comparison with previous ADAPT programs and should only be used if the indices of abundance are being modeled as normal or lognormal distributed variates. The concentrated likelihood approach is effectively equivalent to the method of iterative re-weighting used by some other ADAPT programs such as STAATS (J. E. Powers, National Oceanic and Atmospheric Administration, Southeast Fisheries Science Center, Miami Laboratory, USA). Note that VPA-2BOX also allows the variances to be represented by parameters estimated in the search (see discussion on line 47 below). If the

solution is well-determined, the maximum likelihood and concentrated maximum likelihood estimates should be the same.

- 47 This option tells the program whether the variance scaling parameters (v_{ik}) specified in the parameter file (see line 77 of Appendix 4) should be multiplied or added to the input values entered in the data file or line 46 above. It is important to remember that the input values may be either CV's or negative standard errors. The scaling parameters therefore modify those values as follows:

<u>Option</u>	<u>Result</u>
0	$\sigma_{i,k,y}^2 = (\text{INPUT}_{iky} * \hat{I}_{i,k,y} * v_{ik})^2$ if input value is a positive CV $\sigma_{i,k,y}^2 = (\text{INPUT}_{iky} * v_{ik})^2$ if input value is a negative std. error
1	$\sigma_{i,k,y}^2 = (\text{INPUT}_{iky} * \hat{I}_{i,k,y})^2 + v_{ik}^2$ if input value is a positive CV $\sigma_{i,k,y}^2 = (\text{INPUT}_{iky})^2 + v_{ik}^2$ if input value is a negative std. error

As is true of any parameter listed in the parameter file, the variance parameters may be fixed or estimated in various ways. However, the program will send an error message if you enter 999 at line 46 and then try to estimate the variance parameters entered in the parameter file since one would be trying to estimate the same parameter in two different ways simultaneously.

One very important item to remember is that the multiplicative factors are made after the inputs are converted into the appropriate variances. Therefore, when a lognormal distribution is specified the v_i^2 values are added or multiplied to the inputs after they are converted to logscale variances $\tilde{\sigma}_{iky}^2$, i.e.,

<u>Option</u>	<u>Result</u>
0	$\tilde{\sigma}_{iky}^2 = \ln(\text{INPUT}_{iky}^2 + 1) * v_{ik}^2$ if input value is a positive CV $\tilde{\sigma}_{iky}^2 = \ln((\text{INPUT}_{iky}/I_{iky})^2 + 1) * v_{ik}^2$ if input value is a negative std. error
1	$\tilde{\sigma}_{iky}^2 = \ln(\text{INPUT}_{iky}^2 + 1) + v_{ik}^2$ if input value is a positive CV $\tilde{\sigma}_{iky}^2 = \ln((\text{INPUT}_{iky}/I_{iky})^2 + 1) + v_{ik}^2$ if input value is a negative std. error

Note that in this case the input values I_{iky} are used rather than the model predictions $\hat{I}_{i,k,y}$. Constant variance scenarios equivalent to the concentrated likelihood method (line 46) may be imposed by use of the additive structure with INPUT=0 or by use of the multiplicative structure with INPUT=1.

Constraints on vulnerability

- 56 The vulnerability (partial recruitment) for a subset of age groups can be constrained to change slowly over the last several years of the time series. The number of years included in the penalty (n) is entered first, followed by the log-scale standard error σ_v (which controls the severity of the penalty) and the youngest age (α) and oldest age (A). If n is less than 2, no penalty is imposed and the program does not attempt to read σ_v , α and A , so they do not need to be entered.

The penalty term that is added to the objective function is a correlated random walk:

$$\sum_{y=Y-n+2}^Y \sum_{a=\alpha}^A \left\{ 0.5 \left(\frac{\ln v_{k,a,y} - \ln v_{k,a,y-1}}{\sigma_v} \right)^2 + \ln \sigma_v \right\}$$

$$v_{k,a,y} = \frac{F_{k,a,y}}{\max_a F_{k,a,y}}, \quad (3.5)$$

where Y is the last year in the time series and v is the relative vulnerability. Note that the maximization in the vulnerability computation is over all ages, not just those from α to A . Also, the penalty is on the relative vulnerability, not the fishing mortality rates themselves (which are more likely to vary from year to year owing to changes in effort).

The vulnerability penalty can be very useful in situations where there are few data covering several age groups during the recent time period. Usually, the VPA solution can be stabilized by linking the last two or three years with a σ_v between 0.4 and 0.8, although one should be careful not to impose values of σ_v too much smaller than the standard errors of the indices abundance or the penalty will dominate the objective function (Walter and Porch 2012). Smaller values on the order of 0.2 may be appropriate if it is known apriori that the fishery has changed very little in the affected years, otherwise if the penalty is too strong it may force rather strange solutions in order to achieve an exact match to the catch-at-age. Of course the value of σ_v should be large (> 1.0) if there is evidence to suggest that the vulnerability has changed substantially during the last n years of the time series. Alternatively, one could set $n = 0$ to eliminate the vulnerability penalty altogether and then fix the partial recruitments to some predetermined values (see discussion of line 41 in the parameter file).

Constraints on recruitment

62 The recruitments R (abundance of youngest age group) can be linked as a correlated random walk. The number of years included in the penalty (n) is entered first, followed by the log-scale standard error σ_R (which controls the severity of the penalty):

$$\sum_{y=Y-n+2}^Y \left\{ 0.5 \left(\frac{\ln R_{k,y} - \ln R_{k,y-1}}{\sigma_R} \right)^2 + \ln \sigma_R \right\} \quad (3.6)$$

As with the vulnerability penalty above, this recruitment penalty is most useful where there is little information to determine the last several recruitments (which is almost always true). A strong penalty can force nearly constant recruitment over the most recent two or three years, but earlier recruitment estimates are seldom affected, even when n is set to include the entire time period, owing to the well known convergent properties of VPA. However, bizarre solutions can arise if the penalty is too strong ($\sigma_R \ll 0.1$) because the penalty will try to force essentially constant recruitment, but the VPA must still match the catches.

63 The recruitments R of the two separate stocks may be linked by the penalty term

$$\sum_{y=Y-n+2}^Y \left\{ 0.5 \left(\frac{\ln rR_{2,y} - \ln R_{1,y-1}}{\sigma_r} \right)^2 + \ln \sigma_r \right\} \quad (3.7)$$

Here n refers to the number of years to link (including the most recent year Y), σ_r is the standard error and r is the recruitment ratio (number of recruits from stock 1 divided by the number from stock 2). The entries for σ_r and r are not read if n is set to 0.

The main purpose for this penalty is to accommodate situations where the growth and fishing pressure on a stock may differ by sex, but the sex ratio is fixed at the time of recruitment. Restrepo and Porch (2000) applied this penalty to Atlantic swordfish and found that it had a strong effect on the recruitments during the most recent years, but little effect farther back in time (again, owing to the well known convergent properties of VPA). As with the other penalties, setting σ_R much lower than 0.1 can lead to aberrant solutions and poor fits to the data.

Constraint on spawner-recruit relationship

72 This constraint penalizes departures from the Beverton and Holt (1957) spawner-recruit relationship. It takes the form of a first-order autoregressive error structure:

$$\sum_k \frac{1}{2\sigma_{BH,k}^2} \left[(1-\rho_k^2) \varepsilon_{k,\psi}^2 + \sum_{y=\psi}^{\Psi-1} (\varepsilon_{k,y+1} - \rho \varepsilon_{k,y})^2 \right] + n_k \ln \sigma_{BH,k} - \frac{\ln(1-\rho_k^2)}{2} \quad (3.8)$$

where the form of the ε 's depends on the distribution specified by the first entry. The penalty is turned off by entering a value of zero. It is invoked by entering either 1, 2, -1, or -2; followed by the first and last years corresponding to the range of recruitments one wishes to use (ψ and Ψ):

$$\varepsilon_{k,y} = \begin{cases} \ln(R_{k,y}) - \ln\left(\frac{a_k SSF_{k,y-\alpha}}{b_k + SSF_{k,y-\alpha}}\right) & \text{lognormal (entry = 1 or -1)} \\ R_{k,y} - \frac{a_k SSF_{k,y-\alpha}}{b_k + SSF_{k,y-\alpha}} & \text{normal (entry = 2 or -2)} \end{cases} \quad (3.9)$$

Here n_k is the number of spawner-recruitment pairs, α is the youngest age in the VPA, and SSF is the average spawning stock fecundity computed from user-supplied fecundity-at-age data as described for line 147 in the data file section. The variables a and b are parameters of the Beverton and Holt spawner-recruit relationship and ρ is the correlation coefficient, all three of which must be specified in the parameter file. The term σ_{BH} , which refers to the standard error of the random component of the recruitment deviations, may either be specified in the parameter file (if the first entry is 1 or 2) or else estimated by the concentrated likelihood approach (if the entry is -1 or -2):

$$\hat{\sigma}_{BH,k}^2 = \frac{1}{n_k} \left[(1-\rho_k^2) \varepsilon_{k,\psi}^2 + \sum_{y=\psi}^{\Psi-1} (\varepsilon_{k,y+1} - \rho \varepsilon_{k,y})^2 \right] \quad (3.10)$$

As for the previous two stock-recruitment penalties, most of the effect is on the recruitment estimates for the most recent years. Strong effects on the early part of the time series are usually not achieved without a very strong penalty (under about 0.1), in which case there is a danger of bizarre solutions because of the conflicting need to match the catch-at-age history exactly.

Parameter estimation options

78 The parameters for each age group a on the last year Y may be represented by the fishing mortality rate in that year $F_{a,Y}$ (entry = 1) or the abundance at the beginning of the next year after mixing $N_{a+1,Y+1}$ (entry =2). Normally the first option is to be preferred because it is easier to guess the magnitude of F than of N , but the latter is useful to facilitate comparisons with older ADAPT programs.

79 This option allows one to estimate the catchability parameters q_{ki} by the concentrated likelihood approach (enter 0) rather than include them in the search list specified by the parameter file (enter 1). The concentrated likelihood approach is computationally more efficient and generally is the method of choice if one is willing to assume catchability does not change with time. Otherwise, one should specify the q 's in the parameter file. The concentrated likelihood approach is available for the lognormal, normal and Poisson distributions as shown in equation (3.11). If the chi-square, Laplace or gamma distributions are selected for the indices (see Table 4), then one should specify the q 's in the parameter file (otherwise the program will substitute the concentration formula for the normal distribution).

$$q_{ik} = \left\{ \begin{array}{l} \exp \left[\frac{\sum_y \frac{\ln I_{iky} - \ln \left(\sum_a v_{ikay} w_{ikay} \tilde{N}_{kay} \right)}{\tilde{\sigma}_{iky}^2}}{\sum_y \frac{1}{\tilde{\sigma}_{iky}^2}} \right] \\ \sum_y \frac{I_{iky} \sum_a v_{ikay} w_{ikay} \tilde{N}_{kay}}{\sigma_{iky}^2} \\ \sum_y \frac{\left(\sum_a v_{ikay} w_{ikay} \tilde{N}_{kay} \right)^2}{\sigma_{iky}^2} \\ \sum_y I_{iky} \\ \sum_y \sum_a v_{ikay} w_{ikay} \tilde{N}_{kay} \end{array} \right. \begin{array}{l} \text{lognormal} \\ \\ \\ \text{normal} \\ \\ \text{Poisson} \end{array} \quad (3.11)$$

Bootstrap analyses

87 This is the number of bootstraps you wish to conduct. A positive integer indicates they are to be nonparametric bootstraps and a negative integer indicates they are to be parametric bootstraps (e.g., entering -500 would tell the program to do 500 parametric bootstraps). An input value of 0 means no bootstraps are to be conducted. If you chose to conduct a nonparametric bootstrap, then you must enter another integer on the same line that specifies whether or not to inflate the bootstrap residuals by use of the Stine correction factor discussed in regards to equation 2.31 (1 = yes, do it; 0 = no, do not do it).

Retrospective analyses

91 This entry specifies the number of years to go back for annual retrospective analyses (up to 20). The program automatically removes one year from the data and relevant parameters and then performs the VPA on the reduced data set. This procedure is repeated in annual steps until the data are reduced by the specified number of year (no retrospective analyses will be conducted if a 0 is entered). The results from each successive retrospective run (x) are written to files with the same format as the output files from the base run: MINUSx.R gives the diagnostic output, MINUSx.EST gives the parameter estimates, and MINUSx.SPD gives the spreadsheet friendly output (see the explanation of the output files in chapter 4).

One potential pitfall of automating this process occurs when when some of the data do not go as far back in time as the specified retrospective analyses. For example, suppose there is an index of abundance that only goes back 7 years and you are attempting to estimate both a catchability q and standard error σ for that index. The fifth retrospective run will have to estimate those two parameters from only two index points and any subsequent run will not be able to estimate them at all. A similar problem would be encountered when estimating tagging parameters from too few data. The user should be mindful that the program will attempt to run retrospective VPA's regardless of whether they make sense or not. The nonsensical results should be evident from the output files and the user should consider setting up those retrospective analyses by hand.

Note that you cannot run a retrospective analysis and a bootstrap analysis with the same call to the program. You will need to initiate two different runs with the same random number seed (line 38 above), one with the control file pointing to retrospectives and another with it pointing to bootstraps

92 The @ symbol is read as the end of the file, but is not necessary here.

DATA FILE : See Appendix 2

This file contains all the information pertaining to the data used in the VPA with the exception of the tag-recoveries. One feature that distinguishes the format of this file from the control and parameter files is the use of negative entries to demarcate the end of a data input stream.

Remember, the line numbers given here are only for referencing the example file in Appendix 2; they will change depending on the placement of comments and the length of the data series. What is important is to enter the data and specifications in the correct order, not the specific line number.

Line Explanation

Comments

1 - 6 Comments preceded by the # symbol in the first column.

General specifications

7 The first and last year in the data. It is not necessary to use four digit years, but of course you must be consistent in whatever accounting you use.

8 The youngest and oldest age classes in the data followed by two values relating to the “plus-group”: the plus-group age A and the “expanded plus-group” age Ω .

As discussed above, program VPA-2BOX makes calculations using a plus-group, where the parameters are assumed to be constant for all ages at or older than the plus-group age A . The choice of the plus-group age is usually based on the inability to determine the age of animals older than a certain critical limit or because those older fish are too rare to be well represented in the catch. An implication of using a plus-group is that fish can live to an indeterminate age. This generally is of little consequence unless the natural mortality rate is very low, in which case it is possible that the model could predict a finite numbers of animals with impossibly long lifespans. In such cases, the user may wish to consider using an older plus-group with an artificially high natural mortality rate.

If the plus-group specified on line 8 is less than the oldest age, the program reads the data from the youngest age to the oldest age and automatically accumulates the catch of age classes greater than or equal to the plus-group. If there is weight information given in the data file, the average weight of the plus-group is computed from the weights of the older age classes as leveraged by the corresponding catches (see discussion of lines 141 and 147). The oldest age is then reset to the value of the plus-group and all subsequent calculations use the plus-group age as the maximum age. (This makes it easier to examine the effect of changing the age of the plus-group).

The “expanded plus-group” feature is needed when the catch-at-age data are censored into a plus-group that is smaller than the oldest age class included in an index of abundance or mortality (for example, if the intended plus-group is age 10, but the oldest fish in the index are age 13), or when information on the mean-size of the plus-group is available. Essentially the approach assumes that the members of the plus-group are approximately in equilibrium during the first year such that the sum of the number in each age class a ($a \geq A$) is the same as the number predicted by the VPA for the plus-group as a whole. The abundance of the cohorts in the expanded plus-group are then tracked through time using the recursions in Table 2 as appropriate. Note that the effect of the equilibrium assumption is mitigated with time as the affected cohorts disappear from the fishery and will usually be negligible after $n = \Omega - A$ years.

In the case of a single stock the equilibrium expansion for the first year is

$$N_{A+t,y_0} = \begin{cases} \frac{N_{A+,y_0} e^{-tZ_{A,y_0}}}{\sum_{j=A}^{A+n-1} e^{-(j-A)Z_{A,y_0} + e^{-nZ_{A,y_0}}/(1-e^{-Z_{A,y_0}})}} & \text{for } t = 0, 1, \dots, n-1 \\ \frac{N_{A+,y_0} e^{-nZ_{A,y_0}}/(1-e^{-Z_{A,y_0}})}{\sum_{j=A}^{A+n-1} e^{-(j-A)Z_{A,y_0} + e^{-nZ_{A,y_0}}/(1-e^{-Z_{A,y_0}})}} & \text{for } t = n \end{cases}$$

where y_0 refers to the first year in the VPA and the term $N_{A+,y}$ (with subscript $A+$) refers to the VPA estimate of the combined abundance of the plus-group, as opposed to the abundance of fish in age class A alone ($N_{A,y}$).

In the case of two stocks, it is necessary to account for the relative magnitude of each stock, so the equilibrium population structure is found using the recursions in Table 6 (below). Note that the formula applied when $t = n$ is only approximate if T or \tilde{T} are nonzero, so the value of Ω should be set large enough to

ensure N_{A+n,y_0} is small (thus avoiding potential bias).

Table 6. Equilibrium diffusion and overlap equations expanded plus-groups.

Diffusion equilibrium recursion	
$u_{k,a,y_0} = \begin{cases} \hat{N}_{k,A+,y_0} & (a = A) \\ \sum_{j=1}^2 u_{j,a-1,y_0} T_{j,k,A} e^{-Z_{k,A,y_0}} & (A < a \leq \Omega) \\ \frac{u_{k,\Omega,y_0}}{(1 - e^{-Z_{k,A,y_0}})} & (a = \Omega) \end{cases}$	<p>Step 1: Relative number [of animals] in zone k that are age a at the beginning of year y just <i>before</i> mixing, where Ω represents the oldest age in the ‘expanded plus-group’</p>
$\hat{N}_{k,a,y_0} = \hat{N}_{k,A+,y_0} \frac{u_{k,a,y_0}}{\sum_{a=A}^{\Omega} u_{k,a,y_0}}$	<p>Step 2: Rescale to match VPA estimate of plus-group</p>
Overlap equilibrium recursion	
$u_{s,a,y_0} = \begin{cases} \tilde{N}_{s,A+,y_0} & (a = A) \\ u_{s,a-1,y_0} \sum_{k=1}^2 \tilde{T}_{s,k,A} e^{-Z_{k,A,y_0}} & (A < a \leq \Omega) \\ \frac{u_{s,\Omega,y_0}}{(1 - e^{-Z_{k,A,y_0}})} & (a = \Omega) \end{cases}$	<p>Step 1: Relative number of stock s that are age a at the beginning of year y. Here $A+$ represents the plus-group (age A and older) and Ω represents the oldest age in the ‘expanded plus-group’</p>
$\tilde{N}_{s,a,y_0} = \tilde{N}_{s,A+,y_0} \frac{u_{s,a,y_0}}{\sum_{a=A}^{\Omega} u_{s,a,y_0}}$	<p>Step 2: Rescale to match VPA estimate of plus-group</p>

Input for first stock/zone

- 9-11 These comments remind the user to begin inputting all the data that pertains specifically to the first zone or stock in the analysis. If only one zone/stock is specified in the control file, then the information will only be read for that one zone and the user can disregard the comment on line 174.
- 12 Number of indices of abundance (or mortality) to be read. This should include any indices that you list later in the data file, even if you do not intend to use them in the analysis.
- 13 Spawning date (t_k), in months elapsed. A value of 0 indicates the beginning of the year and a value of 12 indicates the end of the year (but before mixing occurs, if applicable).
- 14 Fecundity modifier (p) for each age class. This is used to determine the spawning stock fecundity SSF as follows:

$$SSF_{k,y} = \begin{cases} \sum_a p_{k,a} f_{k,a,y} N_{k,a,y} e^{-Z_{kay} t_k} & \text{if } t_k \geq 0 \\ \sum_a p_{k,a} f_{k,a,y} N_{k,a,y} \frac{(1 - e^{-Z_{kay}})}{Z_{kay}} & \text{if } t_k < 0 \end{cases}$$

(3.12)

where the fecundity information (f) is specified at the bottom of the file (line 146). The fecundity modifier can be used, for example, as an index of spawning fraction when computing spawning stock biomass from weight information.

17 Title of the stock/zone to which the following catch data apply (must be 50 characters or less) followed by the probability density (pdf) the catch data and a measure of its standard error σ_C (a positive value is interpreted as a CV and a negative value as a standard error). The pdf and σ_C specifications are only used if a bootstrap analysis is specified in the control file. If the pdf is set to 0 the catches are held constant, otherwise a parametric bootstrap of the catches is carried out according to the specified distribution (where the expectation is set equal to the observed catch at age). The options available for the pdf are summarized in Table 7. Note that the specifications here pertain only to the catch; the bootstrap specifications for the indices of abundance appearing elsewhere.

Catch data input

22 - 46 Here is where the catch at age vectors are input. Each line must include the year followed by the catch observations for every age class from youngest to oldest (recall that if the age of the plus-group is less than the oldest age indicated on line 8, then the program will read from youngest to oldest and automatically combine the catches into a plus-group).

Years with no catch need not be input; the missing years will automatically be assigned catch-at-age values of 1.0. Also, catch-at-age entries that are less than or equal to zero are replaced by a value of 1.0. Zero catches are not acceptable because the explicit VPA recursion would be undefined, so a small value of 1 fish is used instead. If vectors for the same year are entered on more than one line, the catches corresponding to the last line with that year will be used.

47 A negative value (e.g., -1) must be entered after the catch-at-age matrix to tell the program to move on to the index specifications.

Table 6. Probability density functions available for parametric bootstraps of catch and index data. In case of catch data the expectation μ is set equal to the observed catch-at-age, whereas with the index data μ is set equal to the model prediction of the index value. The methods of generating the variables are described in Law and Kelton (1982).

Option	Mathematical representation	How x is generated
1: lognormal	$\frac{1}{\sqrt{2\pi\tilde{\sigma}^2}x} e^{-0.5(\log_e [x] - \mu)^2 / \tilde{\sigma}^2}$	polar method to get normal variate y , then take $x = e^y$
2: normal	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-0.5(x - \mu)^2 / \sigma^2}$	polar method
4: Poisson	$\frac{\mu^x e^{-\mu}}{x!}$	Law and Kelton (1982)
7: uniform	$\begin{cases} \frac{1}{\sqrt{12}\sigma} & \text{for } \mu(1 - \frac{\sqrt{3}\sigma}{\mu}) \leq x \leq \mu(1 + \frac{\sqrt{3}\sigma}{\mu}) \\ 0 & \text{otherwise} \end{cases}$	3 linear congruential generators (see Press et al., 1995)
8: triangular (centered at μ)	$\begin{cases} \frac{x - \mu + \sqrt{6}\sigma}{6\sigma^2} & \text{if } -\sqrt{6}\sigma \leq x - \mu \leq 0 \\ \frac{\mu + \sqrt{6}\sigma - x}{6\sigma^2} & \text{if } 0 \leq x - \mu \leq \sqrt{6}\sigma \\ 0 & \text{otherwise} \end{cases}$	Law and Kelton (1982)

Index specifications

56-57 There are 8 entries on each line that tell the program how to interpret the index time series that follow. There should be one line for each index. The **first entry** identifies the index and must be an integer between 1 and the number entered on line 12. The **second entry** identifies the error structure one wishes to employ (see Table 6) and must be one of the following choices: 0 = do not use this index, 1 = lognormal, 2 = normal, 4 = Poisson, 5 = robust Chi-square, 6 = Laplace double exponential, 7 = gamma, 12 = standard normal. Please note, however, that only the lognormal, normal and Poisson forms can be used for parametric bootstraps. The last choice, 12 = standard normal, uses the same formula as choice 2 (normal), but the input data are interpreted as standard normal deviates created by subtracting the series mean from each observation and then dividing by the series standard deviation. This option has been included to accommodate situations where the only available indices had been standardized in this manner. However I recommend against the use of this standardization approach if at all possible inasmuch as the difference between a fairly flat trend with low variance and a steep trend with high variance is obscured.

The **third entry** is an integer that specifies the units of the index as follows:

<u>Option</u>	<u>Prescription</u>
1	index of abundance in numbers
2	index of biomass (abundance in terms of weight)
3	index of relative fishing mortality rate, $\text{Max}\{F_y\} - F_y$
4	index of relative total mortality rate, $\text{Max}\{Z_y\} - Z_y$
5	index of absolute fishing mortality rate
6	index of absolute total mortality rate
7	proportion of population in area k that originated in the opposite area (calculated in numbers)
8	proportion of population in area k that originated in the opposite area (calculated in biomass)
9	index of average weight

The first two options above relate to indices of abundance and correspond to Table 3, where for choice (1) $w_{kiay} = 1.0$ and for choice (2) $w_{kiay} =$ the weight-at-age specified later in the data file. Options 3, 4, 5 and 6 relate to indices of mortality and correspond to Table 4. Options 7 and 8 refer to stock composition data (see equations 2.22-2.27) and may only be invoked when two stocks are being analyzed. Option 9 refers to the use of average weight data, specifically for estimating the mortality rate on the plus-group, and requires use of the expanded plus group discussed above in connection with line 8 of the data file (see also equations 2.19-2.22).

The **fourth entry** is an integer that specifies the method that is to be used to determine the index-specific vulnerability vector v_{ikay} . There are three basic choices here:

<u>Option</u>	<u>Prescription</u>
1	Vulnerability vector read as input (line 111)
2 - 3	$v_{ikay} = \frac{C_{ikay} F_{kay} / C_{kay}}{\text{MAX}_a \{C_{ikay} F_{kay} / C_{kay}\}} \quad (3.13)$
4	$v_{ika} = \frac{\sum_y C_{ikay} F_{kay} / C_{kay}}{\text{MAX}_a \left\{ \sum_y C_{ikay} F_{kay} / C_{kay} \right\}} \quad (3.14)$

When options (3) or (4) are selected, the values read from the file (see line 111) are assumed to be the fishery-specific (partial) catches C_{ikay} . When option (2) is selected, the values from line 111 are assumed to be the fractions C_{ikay}/C_{kay} (but otherwise the computations are identical to option 3). The value of F_{kay} is the fishing mortality rate estimated by the VPA at the current iteration.

Option 1 should be used whenever the index represents a single age class, in which case one would normally assume the vulnerability is constant at 1.0 (the default value if no vulnerabilities are input). Option

1 is the only option available for indices of mortality. The choice between options 2 or 3 (Powers and Restrepo, 1992) and option 4 (Geromont and Butterworth, 1999) depends upon how stable one believes the vulnerability-at-age vector is likely to be. The former allows the vector to change from year to year whereas the latter assumes it is constant over all of the years covered by the index. One may be tempted to go with the seemingly more flexible approach of option 2-3, however it turns out that option 4 nearly always produces a better fit to the indices of abundance (and a lower value of the objective function). There reasons for this are two inherent inconsistencies in the method of options 2-3. The first inconsistency arises because the partial catches are usually observed with some error and yet must be matched exactly; large random errors in the partial catches will tend to skew the estimated vulnerabilities too far towards younger ages in some years and too far towards older ages in other years, making it difficult for the model to fit the index data. The method of option 4 will tend to average these random errors away. The second inconsistency arises with the usual practice of holding the catchability coefficient q constant. In point of fact, one would expect a change in S to be accompanied by an effective change in q since some of the same factors affect both; one would, for example, generally expect very different q values to accompany the relative vulnerability vectors $v=\{1,1,1,1,1\}$ and $v=\{0,0,0,0,1\}$. However, when option 4 is used and q is at the same time held constant, this likely scenario is ruled out, creating a potentially serious model mis-specification problem. In general, I do not recommend the use of options 2 and 3 unless q is allowed to vary from year to year (see the discussion on specifications for the q parameters in the parameter file).

The **fifth entry** specifies when the index observations were made in terms of months elapsed ($indextime_{i,k}$). A negative integer (e.g., -1) indicates the observations represent the average abundance during the year. The value of $indextime_{i,k}$ is used to determine the value of the adjustment factor Δ :

$$\Delta_{i,k,a,y} = \begin{cases} \frac{(1-e^{-Z_{k,a,y}})}{Z_{k,a,y}} & \text{if } indextime_{i,k} < 0 \\ e^{-Z_{k,a,y}(\frac{indextime_{i,k}}{12})} & \text{if } indextime_{i,k} \geq 0 \end{cases} \quad (3.15)$$

The **sixth** and **seventh entries** give the youngest and oldest age represented by the index. The oldest age should not exceed the age of the plus-group. The **eighth entry** is the title of the index you want to appear in the output (must be 50 characters or less).

58 A negative value (e.g., -1) must be entered after the index specifications to tell the program to move on from the specifications to reading the index data.

Index data

63-104 Here the actual values of the indices of abundance or mortality are read. The **first entry** is the integer identifying the index and the **second entry** is the year. The **third entry** is the value of the index, which is ignored if it is negative unless the standard normal option is selected. The **fourth entry** is interpreted as a coefficient of variation if is positive and a standard error if it is negative. It corresponds to the value labeled $INPUT_{iky}$ in the discussion on line 47 of the control file. Note that when units of the index are specified as stock composition data (options 7 or 8), values for the third entry greater than 1 are interpreted as the number (option 7) or aggregate weight (option 8) of fish recovered in the specified zone (rather than as a proportion). Similarly, when units of the index are specified as stock composition data, but entered as proportions, values for the fourth entry exceeding 2 are interpreted as sample sizes.

105 A negative value (e.g., -1) must be entered after the index specifications to tell the program to move on to reading the vulnerability data.

Index vulnerability information

110-134 Here the values of the vulnerabilities or partial catches for the indices are read. The **first entry** is the integer identifying the index and the **second entry** is the year. After these must come the vulnerability-at-age vector for that year, starting with the youngest age and continuing to the oldest age (as per line 8). If no lines are input for a given index or year a value of 1.0 is assumed for all ages (thus you do not need to list an index where the vulnerabilities are the same for all ages in all years).

If the plus-group age is less than the oldest age and the inputs are fixed vulnerabilities, the vulnerability given for the plus-group age will be used. Otherwise, the partial catches or fractional catches of the age classes

older than the plus-group age will be combined into the plus-group. Whether the values are read as vulnerabilites, partial catches or catch fractions is specified by the fourth entry on line 56.

135 A negative value (e.g., -1) must be entered after the index specifications to tell the program to move on to reading the index weight data.

Index weight information

140 Here the weight-at-age vectors for the indices are read. This information is only used if the third entry on line 56 is set to 2 (for biomass units). The **first entry** is the integer identifying the index and the **second entry** is the year. After these must come the weight-at-age vector for that year, starting with the youngest age and continuing to the oldest age (as per line 8). If the plus-group age is less than the oldest age, the weight of the plus-group is computed from the partial catches if they are available and the total catch otherwise, i.e., as $\Sigma C_{ikay}w_{ikay}/\Sigma C_{ikay}$ or $\Sigma C_{kay}w_{ikay}/\Sigma C_{kay}$.

141 A negative value (e.g., -1) must be entered after the index specifications to tell the program to move on to reading the fecundity data.

Fecundity information

146 Here the fecundity-at-age data are entered (the variable *FEC* discussed in connection with line 14 and equation 3.12). This can be any measure of the relative production of the stock, such as weight-at-age for computation of spawning stock biomass or the per capita fecundity (e.g., average gonad weight, average number of eggs per female, etc.). The **first entry** is the integer identifying the index and the **second entry** is the year. After these must come the fecundity-at-age vector for that year, starting with the youngest age and continuing to the oldest age (as per line 8). If the plus-group age is less than the oldest age, the fecundity of the plus-group is computed from the total catch as $\Sigma C_{kay}FEC_{kay}/\Sigma C_{kay}$.

171 A negative value (e.g., -1) must be entered after the index specifications to tell the program to move on to the next zone/stock.

Input for second stock/zone

172-175 These comments remind the user to begin inputting all the data that pertains specifically to the second zone or stock in the analysis. If only one zone/stock is specified in the control file, then any information entered from here on will not be read. Otherwise, the information in lines 12 to 172 will need to be repeated for the second stock (the number of lines will of course change, but the order and formats must be identical to those used for the information on lines 12 to 172).

TAG RECOVERY DATA FILE : See Appendix 3

This file contains all the information pertaining to the tag-recovery data. The example in Appendix 3 has been abridged to appear as though there were only two years of releases and 3 years of recoveries; the reader is reminded that the line numbers given here are only for referencing the example file in Appendix 3; they will change depending on the placement of comments and the length of the data series. What is important is to enter the data and specifications in the correct order, not the specific line number.

Line Explanation

Comments

1 - 25 Comments preceded by the # symbol in the first column.

Release information

27-28 Here you define the tag cohort, starting with an id number i (which will be used to link with the release information) followed by release zone, stock of origin (enter 0 if the stock of origin is unknown), year released, time of year released (τ_{iy} , in elapsed months, averaged over all animals in the cohort), year and month when the experiment ends (normally the last year of the VPA or when recapture data are no longer being collected), age class of the fish when they were tagged, and the number of released animals with these same release attributes. The average time of the year of the releases (τ_{iy}) is used in the calculation to determine the fraction of the fishing year that this group of tags was exposed to (see discussion of line 34 in the control file and equation 3.2).

The information defining a tag cohort must then be followed (on the same line) by the weighting factor ω_i and a series of indices that identify the additional parameters that must be accounted for when tagging data are used (see Table 5). The weighting factor is used to discount tag-recovery information that is deemed somehow less reliable than that of other cohorts (see Table 5). One might wish to do this, for example, if the number of releases that survived the tagging process for one cohort was more uncertain than for other cohorts. If the PDF chosen in the control file is multinomial (option 3), the weighting factor is interpreted as an inverse weight analogous to a standard error (i.e., the multinomial likelihood is divided by the value squared). Otherwise, it is interpreted as a coefficient of variation if is positive and as a standard error if it is negative. If you do not wish to include any special weighting, just set all the weighting factors equal to -1.0, otherwise values with absolute values greater than 1.0 will decrease the weight attributed to the tagging data for that cohort (one can also weight the entire tagging data set using the default weighting factor discussed on line 34 of the example control file).

The entries following the weighting factor are integers that map a series of parameters specified in the parameter file (below) to the appropriate tag cohort. The maps are unique for each basic type of parameter (see lines 93-110 of the example parameter specification file): fraction of animals that survive shortly after release (d), chronic tag loss (λ), reporting rate (ρ), and the “pre-mixing adjustments” accounting for incomplete mixing of the tagged population with the untagged population (γ). Thus, the value for the fraction of animals in tag cohort i that survive shortly after release (d_i) is set to the value in vector element $dparameter(dmap_i)$, and chronic tag loss for tag cohort i (λ_i) is set to the value in vector element $\lambdaparameter(\lambda map_i)$.

The three reporting rates are set to the value in vector element $pparameter(\rho map_{i,j})$ corresponding to ($j=1$) reporting at the end of the experiment ($\rho_{i,0}$) and ($j=2,3$) reporting by fisheries in each zone ($\rho_{i,k}$). Thus, it is possible to make all three reporting rates the same (i.e., represent all three reporting rates by the same parameter) by setting $\rho map_{i,1} = \rho map_{i,2} = \rho map_{i,3}$. Note that for single-stock analyses one should set $\rho map_{i,2} = \rho map_{i,3}$ or fix $\rho map_{i,3}$ to an arbitrary value (don't estimate). Similarly, when conventional tags are used, where there is no means of recovering tags at the end of the experiment (other than the fishery), the value of the $\rho_{i,0}$ should be fixed to 0 by fixing $pparameter(\rho map_{i,1})$ at a value of zero.

The two “pre-mixing adjustments” accounting for incomplete mixing in the first year and second years at large ($\gamma_{i,y}$) are set to the value in vector element $\gammaparameter(\gamma map_{i,j})$ corresponding to ($j=1$) mixing in year 1 ($\gamma_{i,1}$) and ($j=2$) mixing in year 2 ($\gamma_{i,2}$). It is possible to make the values for $\gamma_{i,1}$ and $\gamma_{i,2}$ the same by setting $\gamma map_{i,1} = \gamma map_{i,2}$.

In the example, cohorts 1 and 2 are represented by exactly the same parameters, meaning one is assuming that the values for those parameters are the same for the two cohorts. The values for $\rho_{i,0}$, $\rho_{i,1}$, and $\rho_{i,2}$ are associated with different parameters since $\rho map_{i,1} = 1$, $\rho map_{i,2} = 2$ and $\rho map_{i,3} = 3$. Similarly, the values

for $\gamma_{i,1}$ and $\gamma_{i,2}$ are allowed to differ by associating them with different parameters by setting $\gamma_{map_{i,1}} = 1$ and $\gamma_{map_{i,2}} = 2$. It is also possible to have different parameters for some or all of the cohorts specified in the tag data file. For example, in the example below, the values for d_i , $\rho_{i,0}$, $\rho_{i,1}$, and $\rho_{i,2}$ are associated with different parameters for each cohort (so there will need to be six parameter specifications in the corresponding section of the parameter file). The values for $\gamma_{i,1}$ and $\gamma_{i,2}$ in the same example are associated with the same parameter within a cohort (i.e., the level of incomplete mixing in the first and second year is assumed identical), but allowed to vary among cohorts.

```

*****
...* indices linking tag cohorts with parameter estimates *
*****
...dmapi (initial survival of tags)
...| λmapi (chronic tag loss)
...| | ρmapi (reporting rates) | γmapi (partial-mixing factors)
...| | -----
...| | by remote by fisheries | for | for
...| | sensing area 1 area 2 | year 1 | year 2
...| | j=1 j=2 j=3 | j=1 | j=2
...| | | | | | |
-----
...1 1 1 2 3 1 1
...1 2 4 5 6 2 2

```

- 36 A negative value (e.g., -1) must be entered after the release information to tell the program to move on to reading the recovery data. A value of -1 tells the program to read the format described below. A value of -2 tells the program to read a similar, less compact format for compatibility with older versions (cohort, area recovered, number recaptured by year, starting with the first year in the VPA).

Recovery data

- 45-48 Here you must indicate the tag cohort id, zone recovered, number recovered at the end of the experiment by means other than fishing ($r_{i,k,0}$), and the number of tags recovered from the cohort in that zone during each year of the tagging experiment ($r_{i,k,y}$). In the case of the recaptures, the number of entries should equal the number of years from the date of release to the end of the experiment.

- 49 A negative value (e.g., -1) can be entered after the recovery information to tell the program to stop reading data.

PARAMETER SPECIFICATIONS FILE : See Appendix 4

This file determines how the parameters of the model will be estimated. The format for all parameter specifications is as follows:

lower bound	best estimate (prior expectation)	upper bound	method of estimation	log-scale standard error of prior
0.1	1.2	2.0	1	0.1 {anything after the fifth entry is not read}

There must be one line of specifications for each parameter, however several consecutive parameters with exactly the same prescription may be represented by a single line if it is preceded by a \$ symbol in the first column followed by an integer value that indicates the number of parameters the line represents. For example, writing the line

\$ 3 0.1 1.2 2.0 1 0.1

is equivalent to writing

0.1	1.2	2.0	1	0.1
0.1	1.2	2.0	1	0.1
0.1	1.2	2.0	1	0.1

The lower bound (l) refers to the lowest reasonable value of the parameter you will accept, the best starting estimate (b) refers to the value of the parameter you think is most likely, and the upper bound (u) refers to the highest reasonable value of the parameter you will accept. Solutions with parameter values (θ) that are outside the upper and lower bounds incur a penalty,

$$penalty = \begin{cases} 10+1000\left(\frac{\theta-u}{b}\right)^2 & \text{if } \theta > u \\ 10+1000\left(\frac{\theta-l}{b}\right)^2 & \text{if } \theta < l \end{cases} \quad (3.16)$$

which is added to the objective function. This usually helps the search algorithm find a solution faster by forcing it to concentrate on parameter values within the feasible range. A judicious choice for the starting value b will also usually improve the search algorithm's performance.

The 'method of estimation' indicator tells the program how to handle each parameter θ_j :

Method	Parameter structure	Description
0	$\theta_j = b_j$	fixed constant at best starting guess
1	$\theta_j = \theta_j$	estimated as 'frequentist' parameter (no Bayes prior)
2 (0.3)	$\theta_j = b_j e^{\varepsilon_j}$	estimated with Bayes prior having median b_j
3 (0.1)	$\theta_j = \theta_{j-1} e^{\varepsilon_j}$	estimated with correlated process error (random walk)
4 (0.2)	$\theta_j = \theta_{ref} e^{\varepsilon_j}$	estimated with uncorrelated process error
-0.1	$\theta_j = \theta_{est}$	set to value of the closest previous estimated parameter
-n	$\theta_j = \theta_n$	set to value of n th parameter in the list (whether estimated or not)

where $\varepsilon_k \rightarrow \text{Normal}(0, \tilde{\sigma}_k)$

Note: the Bayesian method designations in parentheses--0.1, 0.2 and 0.3--were used in previous versions of VPA-2BOX and are still accommodated (i.e., the old parameter files can still be used). The variable θ_{ref} refers to the closest previous parameter determined by any method other than method 4 (an uncorrelated random deviation from the preceding uncorrelated random deviation becomes effectively a random walk). The variable θ_{est} refers to the closest previous parameter that was actually estimated (methods 1 to 4). The variable $\tilde{\sigma}_k$ is the log-scale standard error and is specified by the last entry on the parameter specification line (which is read but not used with options 0, 1, -n, and -0.1).

The choice of methods 1 to 4 causes the program to estimate the parameter by searching for the value that minimizes the objective function. Method 1 takes the usual frequentist approach of minimizing the negative log-likelihood function that measures the discrepancy between the data and the model predictions (subject to the constraints discussed in the description of the control file). Methods 2, 3 and 4 however, expand the objective function to include terms that penalize discrepancies between the parameter estimates and their preconceived values:

Method	Term added to objective function
2 (0.3)	$0.5 \left(\frac{\ln \theta_k - \ln b_k}{\tilde{\sigma}_k} \right)^2 + \ln \tilde{\sigma}_k$
3 (0.1)	$0.5 \left(\frac{\ln \theta_k - \ln \theta_{k-1}}{\tilde{\sigma}_k} \right)^2 + \ln \tilde{\sigma}_k$
4 (0.2)	$0.5 \left(\frac{\ln \theta_k - \ln \theta_{ref}}{\tilde{\sigma}_k} \right)^2 + \ln \tilde{\sigma}_k$

Method 2 invokes a lognormal Bayesian prior for the parameter centered on the input best guess b and having a log-scale standard error of $\tilde{\sigma}_k = \sqrt{\ln(CV^2 + 1)}$. This construct is useful when there is some external information on a parameter, but additional insight may be gained from the data used in the VPA. For example, the natural mortality rate is notoriously difficult to estimate and is usually fixed to some predetermined constant. An alternative is to impose a prior centered at that predetermined value with a variance term that reflects its uncertainty.

Methods 3 and 4 may be interpreted as autocorrelated process errors with correlation coefficients of 1 and 0, respectively. From this perspective the parameter θ_k is viewed as a state variable (see Chapter 2, Parameter Estimation) with expectation θ_{k-1} or θ_{ref} (if θ_{ref} is estimated). The two methods are identical when $\theta_{ref} = \theta_{k-1}$, however they have very different implications when a long series of parameters are linked. Generally, one should chose method 3 (known as a random walk; see Porch, 1999) if there is likely to be a consistent trend and method 4 otherwise. For example, the catchability coefficient q of a catch per unit effort series might be expected to increase through time owing to technological improvements made by the fishing fleet, in which case one might prefer the random walk. On the other hand, the catchability coefficient for a research survey might be expected to be devoid of any trends yet still vary owing to fluctuations in the spatial distribution of the stock relative to the locations where the survey was conducted, in which case method 4 might be preferred.

The process error approach tends to the frequentist approach as the process variance becomes either very small ($\tilde{\sigma}_k \rightarrow 0$) or very large ($\tilde{\sigma}_k \rightarrow \infty$). In case of the former, it is as if only θ_{ref} were being estimated, whereas in the case of the latter it is as if one were trying to estimate every θ_k , as a free parameter (generally impossible because of insufficient data). Thus, moderate values of $\tilde{\sigma}_k$ (say between 0.1 and 1.0) can be thought of as a middle ground adding flexibility to the model while still preserving the estimability of the parameters. The best choice for $\tilde{\sigma}_k$ depends on how many other parameters are being estimated, what data are available, and how variable the state variables are likely to be. The bootstrap and other diagnostic tools may give some insight as to when the value of $\tilde{\sigma}_k$ is too large, but the choice is really more of an art than a science.

As mentioned previously, all of the parameters in this file are represented in the same fashion (with one exception to be discussed). However, the order of parameter input is very important. The following discussion is based on the file in Appendix 4 and the reader is reminded that the line numbers given here are only for referencing

that example; they will change depending on the placement of comments and the number of parameters. What is important is to enter the specifications in the correct order, not the specific line number. It is also important to recognize that a single parameter can be used to represent several categories (age class, year, or index) simply by estimating the value for one category and then using the -0.1 or -n method prescriptions to set subsequent parameters equal to that estimated value. Also, one should take care not to try to estimate parameters where there is no data to do so, as might happen if you try to estimate the variance scaling factor for an index that you have in the data file but told the program not to use (by setting the pdf in the data file to 0).

Line Explanation

Comments

1 - 40 Comments preceded by the # symbol in the first column.

Terminal-year parameters

41-47 The parameters for each age group a on the last year Y may be represented by the fishing mortality rate in that year $F_{a,Y}$ or the abundance at the beginning of the next year after mixing $N_{a+1,Y+1}$, depending on the prescription on line 78 of the control file. Each age must be represented by one specification line with the exception of the oldest age, which is represented by the F-ratio parameter discussed in connection with line 51 below. Thus, the youngest age is represented by the first parameter (first specification line) in the file. The format used to specify the terminal-year parameters is the same as for the other parameters below, with one exception. If the method indicator is set to 0 and the value of the best estimate is less than 9, then the best estimate is interpreted as the vulnerability on that age a relative to a reference age j (which is entered in place of the standard deviation) such that $F_{a,Y} = b_a F_{j,Y}$.

If the option to estimate $N_{a+1,Y+1}$ (rather than $F_{a,Y}$) is chosen, then the specifications for $N_{a+1,Y+1}$ are entered in the same position as for $F_{a,Y}$ and the reference age j must still refer to the age during the terminal year (not the age +1). Note that when the overlap model is specified, the program will attempt to solve for $F_{1,a,Y}$ and $F_{2,a,Y}$ simultaneously from $N_{1,a+1,Y+1}$ and $N_{2,a+1,Y+1}$; therefore the same age groups must be used for both stocks (an error message will be generated if you attempt otherwise).

The terminal parameters for the first three or four age groups tend to be rather poorly determined because they are affected by relatively few data points, but their estimability can be improved by implementing the vulnerability or recruitment constraints discussed in the control file. Where those constraints do not seem plausible, one may fix the relative vulnerabilities to some predetermined values and link them to a reference age that can reasonably be estimated.

In the example file (Appendix 4), line 41 tells the program that $F_{1,Y} = 0.2 F_{2,Y}$ and line 42 tells the program to estimate $F_{2,Y}$ with a starting value of 0.4886. Lines 43 to 46 likewise tell the program to estimate the F's on ages 3 to 6, but line 47 tells the program to link age 7 to age 6 such that $F_{7,Y} = 0.9 F_{6,Y}$.

F-ratio parameters

51 This is the ratio of the fishing mortality on the plus-group to the next younger age, $\phi_{ky} = F_{k,A,y}/F_{k,A-1,y}$. There must be one F-ratio specification for each year. Typically the values are fixed to 1.0 unless there is some compelling biological reason to suggest otherwise. VPA-2BOX allows separate values of ϕ_{ky} to be estimated as free parameters for every year, but this is seldom practical owing to the scarcity of data. Generally one will have to limit the number of parameters by assuming ϕ_{ky} is constant during several blocks of years or limit the amount ϕ_{ky} can vary from year to year by use of the process error formulations with moderate levels of $\tilde{\sigma}_k$.

Natural mortality parameters

55-57 One natural mortality parameter M_{ka} must be specified for each age class (inter-annual variations are not accommodated). In the example file there are two M parameters being estimated, one for age 1 and another for age 2 and older (both with starting values of 0.3). In general, the natural mortality rate will not be well-estimated unless the indices of abundance cover a time without fishing or the tag-recovery data come from a well-executed experiment. In most cases it will be necessary to impose informative Bayesian priors on the M_{ka} 's or else fix them to some externally derived constants.

Mixing parameters

61 One mixing parameter T_{ka} must be specified for each age class (inter-annual variations are not

accommodated). These parameter are very unlikely to be well-estimated without tag-recovery data, in which case one must either fix them or impose tight Bayesian priors. If a single stock is being modeled these parameters are read, but ignored.

Stock recruit parameters

65-69 Five parameters are specified for the autocorrelated stock recruit relationship discussed in connection with line 72 of the control file (equation 3.8). The first two parameters represent a_k and b_k of the Beverton and Holt (1957) spawner-recruit curve:

$$R_{k,y} = \frac{a_k SSF_{k,y-\alpha}}{b_k + SSF_{k,y-\alpha}} \quad (3.17)$$

The third specification refers to a parameter that will be used to introduce a three parameter spawner-recruit curve in a later version, but is not now being used and should have the estimation method set to 0. The fourth and fifth specifications are the correlation coefficient ρ_k and standard error σ_k ($\tilde{\sigma}_k$ in the lognormal case) of the process error, respectively. Obviously one cannot estimate any of these parameters unless the stock-recruit penalty on line 72 of the control file is turned on.

Variance scaling parameters

77-78 These are the scaling parameters v_{ik} discussed in connection with equation 2.15 and line 47 of the control file. One v parameter must be specified for each index listed in the data file. Obviously the values of the v_{ik} parameters cannot be estimated if the corresponding index is not being used, so be sure to check that the estimation method for these parameters is either set to zero or equated to a previous parameter as in the example here.

Catchability parameters

83-88 These represent the catchability coefficients for the indices of abundance and proportionality coefficients for the indices of mortality (q_{iky}). This section must not appear if the option on line 79 of the control file specifies for the q values to be estimated by the concentrated likelihood method. Otherwise, there must be one specification line for every year regardless of the number of years for which there are data. If you intend to estimate only a single value of q for all years, then you can estimate the value for the first year and set the values for the remaining years equal to that value as in lines 83-84. However, if you intend to allow q to vary from year to year, it generally does not make sense to try to estimate q for the years where there is no data and you should fix the q values for the years without data to some arbitrary constant (as done in line 88, since the second index in the corresponding data file does not include 1999) or to the value estimated in a preceding year (the program won't predict the value of the index for years with no data). If you do estimate a q parameter for a year with no data, the search algorithm will take unnecessarily long because there is no basis for estimation. One exception to this rule might be when the index is missing a year or two in the midst of the time series. In that case it would still be appropriate to try to estimate the q values for the missing year using one of the process error approaches (provided the imposed $\tilde{\sigma}_k$ is not too large),

Tag-recovery parameters

89-110 The rest of the specifications in this example file pertain to the tag recovery parameters in Table 5 and lines 26-27 in the tag recovery file above. This section must not appear if tag-recovery information is not being used (or be placed after the @ symbol discussed below at line 111).

93 This line(s) specifies the vector of parameters $d\text{vector}(d\text{map}_i)$ representing the immediate loss of tags owing to misapplication or tagging-induced mortality (d_i of Table 5). One parameter must be specified for each map index value in the tag data file ($d\text{map}_i$).

97 This line(s) specifies the vector of parameters $\lambda\text{vector}(\lambda\text{map}_i)$ representing the chronic loss rate of tags owing to shedding or fouling (λ_i of Table 5). One parameter must be specified for each map index value in the tag data (λmap_i). If indices of abundance are available this parameter is estimable, otherwise it is confounded with the natural mortality rate M and one may prefer to set it equal to values determined from double tagging experiments.

101 These lines specify the vector of parameters $\rho\text{vector}(\rho\text{map}_i)$ representing the reporting rates (ρ_{ik} of Table 5). One parameter must be specified for each map index value in the tag data file (ρmap_i). It is important to recognize that this term always occurs together with d (see line 92) as the product $(1-d)\rho$ in the tag-recovery

equations, therefore the two sets of parameters may not be estimated independently for every cohort no matter how good the tag-recovery data are.

109 These lines specify the vector of parameters γ vector(γ map_{*i*}) representing incomplete mixing of tags with the untagged populations (γ_i of Table 5). Unique values can be specified for the first and second years of the tagging experiment by setting differ values for (γ map_{*i*}). Hoenig et al. (1998) points out that the values of γ can be difficult to distinguish precisely from F when only tag recovery data are available, but this should not be too serious of a problem here because F is largely determined from the catch and index data.

@ The end of file indicator

111 The appearance of an @ symbol in the first column tells the program to stop reading the parameter file at that point. It is superfluous if it comes at the end of the file, but is useful if you want to store alternative parameter specifications in the same file. For example, if the concentrated likelihood method were to be used in runs that did not use the tagging data, then an @ could be placed in the first column of line 79 to avoid having delete or comment out all the q and tagging parameters.

GROWTH.DAT FILE

This file contains the growth parameters needed to predict mean size at age for estimation from the mean weight of the plus group (see description of equations 2.16-2.21). It is not needed if no indices of mean weight are to be included. The format for the file is given below:

```

# Zone curve applies to
# | Index number (see index specifications in data file)
# | | Growth curve type (1=von Bertalanffy or Chapman Richards, 2=Gompertz)
# | | | Linfinity
# | | | | k
# | | | | | t0
# | | | | | | Chapman-Richards skew parameter (m=1 for von Bert)
# | | | | | | weight parameter a of w=a1^b
# | | | | | | weight parameter b of w=a1^b
# | | | | | | offset (months elapsed)
# | | | | | |
# | | | | | |
1 | 5 | 1 | 315.0 | 0.089 | -1.13 | 1 | 0.0000152 | 3.05305 | 4
2 | 2 | 1 | 315.0 | 0.089 | -1.13 | 1 | 0.0000152 | 3.05305 | 4
    
```

The first entry in the file specifies the zone that the curve applies to and the second indicates one of two possible **curve types**:

- 1 Chapman-Richards equation,

$$f_a = c \left[L_\infty \left(1 - m e^{-K(a + \frac{t}{12} - \frac{offset}{12} - t_0)} \right)^{1/m} \right]^b \tag{3.18}$$

- 2 Gompertz equation,

$$f_a = W_\infty e^{-e^{-K(a + \frac{t}{12} - \frac{offset}{12} - t_0)}} \tag{3.19}$$

The third through fifth entries correspond to the growth parameters L_∞ (or W_∞), K , and t_0 . The sixth entry is the skew parameter m (note that the Chapman-Richards option reduces to the von Bertalanffy curve when $m = 1$). The seventh and eight entries correspond to parameters c and b of the weight-length power-function used for option 1 (equation 3.18). The entries for m , c and b are superfluous when the Gompertz curve is selected, but dummy values must still be inserted. The eighth and final entry is an offset parameter meant to convert the integer age class (a) used in the accounting of the assessment projection model to the age currency used for the growth curve (usually the actual age of the animal). The conversion is necessary because the definition of the year used in an assessment model is often chosen to coincide with the way the data are collected (e.g., a calendar year, January through December) rather than the birth date of the animal. In such cases, the usual way to keep track of individual year-classes (fish born in a particular year) is to assign them to an integer age class based on the real age of the animal at the beginning and end of the accounting year. For example, consider the case where the animals are born (peak spawning occurs) on May 1st, which is 4 months into the calendar year. On January 1st of its first year of life the year-class is 8 months old and on the 31st of the following December it is 20 months old. Accordingly, any fish caught during the calendar year that were deemed by the age determination process to be between the ages of 8 months to 20 months would be classified as belonging to age class 1, fish aged between 20 months to 32 months would be classified as belonging to age class 2 and so on. More generally, for a given cohort born x months into the year, the age range corresponding to age class a is $a-x/12$ to $a+1-x/12$. If the growth curve is calculated based on the true age of the fish, then x is the offset required in line 57. In some cases the offset x will be equal to the spawning date, but care should be taken to understand how animals are being assigned to age classes to make sure the correct offset is used (it is the amount that must be subtracted from the integer age class used in the model in order to arrive at the actual age at the beginning of the accounting year). For example, suppose the convention had been to assign animals between the ages 8 months to 20 months to age class 0. In that case the true age of an animal belonging to age class a at the beginning of the calendar year is $a - 4/12 + 1 = a + 8/12$. Accordingly, the offset is -8, because 8 months must be added to (-8 months must be subtracted from) the model age.

4. OUTPUT FILES

Program VPA-2BOX always produces four output files containing the: 1) parameter estimates, 2) diagnostic statistics and derived quantities, 3) derived quantities in a spreadsheet friendly format and 4) log of performance statistics. If a retrospective analysis is conducted, then each of the first three files is reproduced for every retrospective run as explained for line 90 of the control file. If a bootstrap analysis is conducted, then a summary file is produced along with a set of binary files with the output from each individual bootstrap run.

PARAMETER ESTIMATE FILE

This format of this file is very much like that of the parameter specification file and can be used as the parameter specification file for subsequent runs simply by changing its name to the pointer in line 16 of the control file³. The only differences are that the estimated value is substituted for the ‘best guess’, the parameters are enumerated, a coefficient of variation (CV) is computed and an exponential format is used. Also, the flag BOUND appears when the parameter is close to the constraints.

lower bound	estimated value	upper bound	method of estimation	log-scale std. error of prior	parameter ID	estimate ID	CV (%)	Flag
0.0000E+00	0.6000E+00	0.3000E+01	0.0	0.2000E+01	1			
0.1000E-01	0.7337E+00	0.5000E+01	1.0	0.1000E+00	2	1	21.	
0.1000E-01	0.4999E+01	0.5000E+01	1.0	0.1000E+00	2	1		1.

BOUND

The column labeled ‘parameter ID’ identifies the numerical order of the adjacent parameter specification. The column labeled ‘estimate ID’ is similar, but refers to the numerical order of the parameters that are estimated. Both identifiers are useful when trying to reference certain error messages produced by VPA-2BOX. The parameter ID is also useful when one wishes to link one parameter to another parameter specified earlier in the file (estimation method *-n* discussed above).

The coefficient of variation (CV) is the standard error of the estimator divided by the value of the estimate, here expressed as a percentage. Typically, values of less than 10 or 20% are interpreted as indicating the parameter estimates are fairly precise, whereas values of more than 50% indicate the parameters are poorly estimated. However, it is important to realize that the standard errors are themselves estimated by inverting the Hessian matrix, which assumes the likelihood⁴ surface is quadratic near the minimum. In practice, this is only approximately true and the

³Note, however, that the new estimates will be used as the central tendencies for any Bayes priors that are imposed (estimation method 2). If you wish to preserve the old priors, you will have to manually change the values in the second column to the ‘best’ values specified in the original parameter file..

⁴The asymptotic covariance matrix is not generally given by the inverse of the Hessian matrix when informative Bayes priors or process errors are incorporated since the posterior is no longer a true likelihood. It may, however, be given to

estimates of standard error tend to be less precise than the estimates of the parameters themselves. In extreme cases a true minimum may not have been found or the surface may be so far from quadratic that the Hessian is not be positive semi-definite and estimates of the CV cannot be produced. Furthermore, the CV's will tend to be biased if the variance parameters for the indices of abundance/mortality are incorrectly specified. Hence, it is recommended that the CV's derived from the Hessian matrix not be trusted unless the variance terms for the indices are estimated (either in the search or by the concentrated likelihood method) and the Hessian appears well-behaved (see discussion on the log file below).

DIAGNOSTIC AND DERIVED STATISTICS FILE

This file contains most of the output statistics one might expect from a VPA program in a format that is easy to read. The first set of information presented is a synopsis of the model's performance:

```

Total objective function = 10602.56 value of the negative log posterior function at its minimum
Number of parameters = 73 number of parameters that are estimated
Number of data points = 8691
AIC {Deviance+2P } = 2531.44
AICc {Deviance+2P(...)} = 2532.69
BIC {Deviance+Plog(nd)} = 3047.55
Chi-square discrepancy = 869.30
Log-likelihood (deviance)= -10600.77 ( 2385.44) the values in parentheses
    effort data = -54.67 ( 326.16) are the deviances
    tagging data = -10546.10 ( 2059.27)
Log-posteriors = -1.79
    catchability = 0.00
    f-ratio = -1.79
    natural mortality = 0.00
    mixing coeff. = 0.00
    initial tag survival = 0.00
    tag shedding rate = 0.00
    tag reporting rate = 0.00
    tag non-mixing factor = 0.00
Constraints = 0.00
    terminal F = 0.00
    stock-rec./sex ratio = 0.00
Out of bounds penalty = 0.00

```

Explanations of these output statistics are provided in previous sections of this manual (see index). Note that the components of the objective function pertaining to the data, priors and process errors and constraints are the positive logarithms of the respective likelihood and posterior function— not the negative logarithms. Thus, a better fit is indicated by larger (less negative) values. The quantity being minimized (first entry above), therefore, is the negative of the sum of the components listed below it.

The remaining components are fairly self explanatory. The first three tables give the estimated fishing mortality rates, estimated abundances and observed catches for each year and age group.

TABLE 1.1 FISHING MORTALITY RATE FOR EAST

	1	2	3	4	5
70	0.219	0.231	0.129	0.099	0.037
71	0.011	0.281	0.267	0.112	0.098

good approximation if the variances of any penalties, priors or process errors are much larger than those associated with the data observation errors (the posterior effectively reducing to a likelihood).

72	0.171	0.235	0.433	0.140	0.094
73	0.145	0.178	0.200	0.111	0.042
74	0.127	0.234	0.273	0.237	0.168

TABLE 2.1 ABUNDANCE AT THE BEGINNING OF THE YEAR [BY AREA] FOR EAST

```
=====
```

	1	2	3	4	5
70	839440.	407538.	256252.	207707.	305586.
71	1244733.	413005.	247664.	172667.	150988.
72	933262.	754513.	238717.	145232.	123160.
73	1313402.	481718.	456478.	118423.	100363.
74	1931654.	696225.	308705.	285930.	83930.
75		1042334.	421733.	179783.	177808.

TABLE 3.1 CATCH OF EAST

```
=====
```

	1	2	3	4	5
6	7	8	9	10	
70	131834.	73211.	26994.	17842.	9793.
71	10384.	88198.	50413.	16589.	12777.
72	116980.	137100.	73231.	17152.	9923.
73	140707.	68048.	71707.	11121.	3710.
74	182955.	125976.	64125.	54045.	11608.

A fourth table is included that gives the abundance by stock (rather than area) when the overlap model is specified for analyzing two stocks simultaneously.

The next set of tables gives the spawning stock biomass, which is computed from the weights and other specifications given in the catch data file:

TABLE 4.1 SPAWNING STOCK FECUNDITY AND RECRUITMENT OF EAST

```
=====
```

year	spawning biomass	recruits from VPA
70	161011.	839440.
71	183815.	1244733.
72	191700.	933262.
73	188799.	1313402.
74	194959.	1931654.

An additional column containing the expected value of recruitment from the spawner-recruit relationship is included if that constraint is employed (see discussion of line 72 in the control file).

The next set of Tables summarizes the fits to the indices of abundance and indices of mortality.

TABLE 5 FITS TO INDEX DATA EAST

```
=====
```

5.1
Lognormal dist.

average numbers
Ages 8 - 10
log-likelihood = 1.46
deviance = 64.17
Chi-sq. discrepancy= 75.96

Year	Observed	Predicted	Residuals (Obs-pred)	Standard Deviation	Q Catchabil.	Untransformed Observed	Untransformed Predicted	Chi-square Discrepancy
70	-0.456	-0.295	-0.161	0.226	0.136E-05	0.634	0.745	0.420
71	-0.177	-0.398	0.221	0.227	0.136E-05	0.838	0.672	1.156
72	0.232	-0.509	0.741	0.232	0.136E-05	1.261	0.601	21.786
73	-0.498	-0.518	0.020	0.230	0.136E-05	0.608	0.596	0.007
74	-0.708	-0.553	-0.155	0.246	0.136E-05	0.493	0.575	0.331

Selectivities by age

Year	8	9	10
70	0.810	1.000	0.791
71	0.810	1.000	0.791
72	0.810	1.000	0.791
73	0.810	1.000	0.791
74	0.810	1.000	0.791

Note that the columns labeled 'untransformed' refer to the value of the indices on an arithmetic scale. They should be the same as the original input values unless the user specified in the control file that they should be divided by the series mean. The columns labeled 'observed', 'predicted', 'residuals', and 'standard deviation' contain the observed values and model predictions on a log scale when the lognormal distribution is specified in the data file, but otherwise are presented on an arithmetic scale (in which case the 'untransformed' columns do not appear). The chi-square discrepancy is computed for each index point for ease of identifying possible outliers (they should sum to the total given at the top of the file, but in this case the entire table has not been presented).

Finally, the last tables in this output file give the fits to the tag-recovery data:

TABLE 6.1 FITS TO TAGS RELEASED IN AREA OF EAST

Multinomial dist.
log-likelihood = -1770.26
deviance = 367.20

Year	Release Age	Recapture Area	Type	Recaptures by year following release			
				1	2	3	4
70	1	1	obsd	3.00	2.00	0.00	0.00
			pred	0.60	0.36	0.03	0.01
71	1	1	obsd	23.00	1.00	3.00	2.00
			pred	7.28	2.18	1.13	0.13

Note that chi-square discrepancy statistics are not computed for tag-recapture data.

DERIVED STATISTICS IN SPREADSHEET FRIENDLY FORMAT

This file contains the fishing mortality rate estimates, abundance estimates, observed catches, input weight at age for spawners, and the observed and predicted values of the indices. There are 100 spaces (for up to 100 years) devoted to each matrix, enabling the user to set up a spreadsheet file that need not be readjusted for every particular application. The index matrix has two columns devoted to each index for each area; the first contains the observed values and the second contains the predicted values (-9's are written as place holders where data are missing).

LOG OF PERFORMANCE STATISTICS

This file, called VPA-2BOX.LOG, records the value of the objective function after each restart of the objective function. Then it displays the results of a first derivative test to see if the proposed minimum point is in fact a local extremum (the hope of course is that it is also a minimum rather than a saddle point, in which case the Hessian matrix must also be positive semi-definite).

The first derivative test results are presented in a table that looks like the following:

FIRST DERIVATIVE TEST

```
=====
          -h          central          +h
          -----          -----          -----
1 : -0.1003D-01  0.1703D-01  0.1904D-01
2 : -0.1001D-01  0.1621D-01  0.1821D-01
3 :  0.8303D-02  0.8309D-02  0.9315D-02      FAIL: Backward step :
```

where in this case there are three estimated parameters. If a local extremum has been found, then all of the first derivatives should be zero at that point. Owing to the large number of options available, VPA-2BOX does not compute analytical derivatives, but instead employs finite difference approximations. In that case, a good indication of whether a true minimum has been reached is if the central difference approximation is close to zero while the backwards difference (-h) is slightly negative and the forward difference (+h) is slightly positive. Several flags are written to alert the user when these criteria are not met (as for the third parameter in the example above). Such failures may occur for a number of reasons, including (1) one or more parameters are estimated near the boundary constraints, (2) the simplex search has not found a true minimum and (3) surface of objective function is not approximately quadratic near the minimum (either very flat or very jagged). The first possibility can easily be checked by inspection of the parameter estimate file discussed above. The second possibility can be addressed by restarting the algorithm with several different initial values for the parameters; if the same minimum is found each time then it is likely that the lowest point has been found. One should also check to see if the search was terminated prematurely because the number of restarts exceeded the limit indicated in the control file. The third possibility suggests either the data are too noisy, conflicting or sparse to provide useful parameter estimates; the user should consider reducing the number of estimable parameters.

It may happen that the first derivative test above indicates that a minimum had been found, but the Hessian is not positive semi-definite and therefore not invertible. This may occur for the same reasons outline above, and the remedies are the same. When the Hessian matrix is invertible, the resulting covariance matrix is printed as well as the corresponding matrix of correlation coefficients. As mentioned previously, the covariance matrix will be biased if the variances of the index data are incorrectly specified or the objective function is not a true likelihood. Moreover, the numerical derivatives in the Hessian matrix are sometimes very sensitive to the step-size used owing to the jagged terrain of the solution surface in the vicinity of the lowest point found by the simplex algorithm (which may not be a well-behaved minima). The correlation coefficients derived from the inverse-Hessian covariances,

$$\hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\hat{\sigma}_i \hat{\sigma}_j},$$

appear to be more stable than the covariance estimates themselves inasmuch as some of the biases cancel out. Accordingly, they are probably a more reliable indicator of the quality of the solution than the magnitudes of the covariances (correlations of 0.1 or 0.2 are considered very low, whereas correlations above 0.9 are excessively high and may indicate that some of the parameters involved should somehow be combined or eliminated altogether). In summary, I do not recommend placing much confidence in the inverse-Hessian covariance estimates or the coefficients of variation derived from them. I have found the correlation coefficients to be somewhat useful during the model development phase, but recommend the bootstrap procedure for characterizing the uncertainty in any model used to generate management advice.

BOOTSTRAP OUTPUT FILES

The file BOOTSTRP.OUT records the bias and standard error of the fishing mortality rate and abundance estimates, which are computed according to the methods discussed in chapter two. The bias and standard errors are also computed for the parameter estimates with the following format

TABLE 1E. PARAMETER ESTIMATES FOR Northern Albacore

```
=====
TERMINAL AGE STRUCTURE OF POPULATION ABUNDANCE
=====
```

Age	MLE	Average of bootstraps	Bias	Std. Error	% CV
2	0.125E+08	0.134E+08	0.657E+06	0.491E+07	36.6
3	0.167E+07	0.183E+07	0.118E+06	0.834E+06	45.6
4	0.784E+06	0.882E+06	0.353E+05	0.267E+06	30.3
5	0.625E+06	0.610E+06	-0.258E+05	0.155E+06	25.4
6	0.454E+06	0.433E+06	0.664E+04	0.124E+06	28.7
7	0.145E+06	0.147E+06	0.151E+05	0.700E+05	47.5
8	0.458E+06	0.492E+06	0.111E+05	0.779E+05	15.8

Here the column labeled “MLE” stands for the best (maximum likelihood or highest posterior density) estimates, the “average of bootstraps” is the mean value of the parameter estimates from all of the bootstrap runs, and the bias is computed from the “average of bootstraps” and a run using the averages of the bootstrap data as described in chapter two.

The results from each individual bootstrap run are stored in a set of 4-bit binary files (.bin extension) with the following formats:

Abundance, fishing mortality, catch, natural mortality, and transfer rate coefficients

```
record length = 4*(Y-y0+1)*(A-firststage+1)*NBOX
NAA.bin: ((N(k,a,y),y=y0,Y),a=firststage,A),k=1,NBOX)
FAA.bin: ((F(k,a,y),y=y0,Y),a=firststage,A),k=1,NBOX)
CAA.bin: ((C(k,a,y),y=y0,Y),a=firststage,A),k=1,NBOX)
MAA.bin: ((M(k,a,y),y=y0,Y),a=firststage,A),k=1,NBOX)
TAA.bin: ((T(k,a,y),y=y0,Y),a=firststage,A),k=1,NBOX)
```

Observed indices of abundance and mortality

```
record length = 4*(Y-y0+1)*(number of indices)
IND.bin ((I(i,1,y),y=y0,Y),i=1,Nind(1)),((I(i,2,y),y=y0,Y),i=1,Nind(2))
```

Terminal vulnerabilities (during last year)

```
record length = 4*(A-firststage+1)*NBOX
TERM.bin: ((Term(k,a),a=firststage,A),k=1,NBOX)
```

Stock recruitment parameters

```
record length = 4*(5+Y-y0+1)*NBOX
SR.bin: ((SR(k,Y),Y=1,5+Y),k=1,NBOX)
```

where y_0 and Y are the first and last years in the model, A is the age of the plus-group, $NBOX$ is the number of zones (1 or 2), and $Nind$ is the number of indices representing each zone.

These files are useful if one wishes to use other software to compute confidence limits and alternative forms of bias correction. They may also be read into programs designed to project the VPA results into the future (such as the companion program to VPA-2BOX, called PRO-2BOX, Porch 2017). Note that an ASC-II file called BAD.OUT is also produced that identifies bootstrap runs that may not have converged or gave otherwise unreasonable results (which the careful investigator may wish to exclude from any further calculations).

5. LITERATURE CITED

- Anganuzzi, A., Hilborn, R. and Skalski, J. R. 1994. Estimation of size selectivity and movement rates from mark-recovery data. *Can. J. Fish. Aquat. Sci.* 51: 734-742
- Baglin, R. E., Farber, M. I., Lenarz, W. H. and Mason, J. M. 1980. Shedding rates of plastic and metal dart tags from Atlantic bluefin tuna, *Thunnus thynnus*. *Fishery Bulletin (US)* 78: 179-185.
- Block, B. A., Dewar, H., Blackwell, S. B., Williams, T., Prince, E. D., Farwell, C., Boustany, A., and Seitz, A. In press. Archival tagging of Atlantic bluefin tuna. ICCAT Coll. Vol. Sci. Pap. (SCRS/99/103)
- Brown, B., and M. L. Parrack, 1985. Status of the Atlantic bluefin tuna resource. Pages 279-289 in Stroud, R. H., ed., *World Angling Resources and Challenges*. Int. Game Fish Assoc. 390 pp.
- Brownie, C., Anderson, D. R., Burnham, K. P., and Robson, D. S. 1985. *Statistical inference from band recovery data: a handbook*. 2nd ed. U.S. Fish and Wildl. Serv. Resour. Publ. No. 156.
- Buckland, S. T., Burnham, K. P., and Augustin, N. H. 1997. Model selection: An integral part of inference. *Biometrics* 53: 603-618
- Butterworth, D. S., and A. E. Punt. 1994. The robustness of estimates of stock status for the western north Atlantic bluefin tuna population to violations of the assumptions underlying the associated assessment models. ICCAT Coll. Vol. Sci. Pap. 42(1): 192-210.
- Cooke, J. G., and K. Lankester. 1996. Examination of alternative stock distribution models for the interpretation of bluefin tuna (*Thunnus thynnus*) tag-recovery data. ICCAT Coll. Vol. Sci. Pap. 45(2): 135-138.
- Dichmont, C.M., Deng, R., Punt, A. E., Brodziak, J., Chang, Y., Cope, J.M., Ianelli, J.M., Legault, C.M., Methot, R. D., Porch, C.E., Prager, M.H., Shertzer, K.W. 2016. A review of stock assessment packages in the United States. *Fisheries Research* 183: 447–460.
- Doonan, I. L., K. Dunn, A. Rasmussen, S. Marsh, C. Mormede, S (In Press). *Casal2: New Zealand's integrated population modelling tool*. Fisheries Research.
- Hilborn, R. 1990. Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. *Can. J. Fish. Aquat. Sci.* 47: 635-643.
- Hoening, J. M., Barrowman, N. J., Pollock, K. H., Brooks, E. N., Hearn, W. S., and Polacheck, T. 1998. Models for tagging data that allow for incomplete mixing of newly tagged animals. *Can. J. Fish. Aquat. Sci.* 55: 1477-1483.
- Hurvich, C. M., and Tsai, C. 1995. Model selection for extended quasi-likelihood models in small samples. *Biometrics* 51: 1077-1084.
- ICCAT. 1994. Report for the biennial period 1992-1993. Part II (1993). 395 pp.
- ICCAT. 1995. Report for the biennial period 1994-1995. Part I (1994). 395 pp.
- Lutcavage, M. E., Brill, R. W., Skomal, G. B., Chase, B. C., and Howey, P. W. 1999. Results of pop-up satellite tagging on spawning size class fish in the Gulf of Maine: Do North Atlantic bluefin tuna spawn in the mid-Atlantic? *Can. J. Fish. Aquat. Sci.* 56: 173-177.
- Mather, F. J. 1980. A preliminary note on migratory tendencies and distributional patterns of Atlantic bluefin tuna on recently acquired and cumulative tagging results. ICCAT Coll. Vol. Sci. Pap. 11: 478-490.
- Mather, F. J., Rothschild, B. J., Paulik, G. J. and Lenarz, W. H. 1974. Analysis of migrations and mortality of bluefin tuna, *Thunnus thynnus*, tagged in the northwestern Atlantic Ocean. *Fishery Bulletin (US)* 72: 900-911.
- Mather, F. J., Mason, J. M., and Jones, A. C. 1995. Historical Document: Life history and fisheries of Atlantic bluefin tuna. 1995. NOAA Tech. Memo. NMFS-SEFSC-370. 165 p.
- Methot, R. D. and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142:86–99
- NRC. 1994. Report of the National Research Council review of Atlantic bluefin tuna. National Academy Press. Washington, D.C. 148 pp.
- Porch, C. E. 1995. A two-area VPA with discrete mixing: Can we discriminate between mixing rates given the present condition of the data? ICCAT Coll. Vol. Sci. Pap. 44(1): 198-208.
- Porch, C.E. 2017. Pro-2box Version 3.0 User Guide. NOAA Tech. Memo. NMFS-SEFSC-708 49 pp.
- Porch, C., P. Kleiber, S. Turner, J. Sibert, R. Bailey and J. Cort. 1998. The efficacy of VPA models in the presence of

- complicated movement patterns. Proceedings of the 25 Anniversary ICCAT Tuna Symposium. ICCAT Coll. Vol. Sci. Pap. 50(2):591-622.
- Porch, C. E., Restrepo, V. R., Turner, S. C., and G. P. Scott. 1995. Virtual population analyses of Atlantic bluefin tuna incorporating movement and tagging data. ICCAT Coll. Vol. Sci. Pap. 44(1): 183-190.
- Porch, C. E. and S. C. Turner. 1998. Virtual population analyses of Atlantic bluefin tuna with alternative models of trans-Atlantic migration. ICCAT Coll. Vol. Sci. Pap. 49:291-305.
- Powers J.E.; Conser R.J and Parrack M.L. 1983. An assessment of Atlantic bluefin tuna resources. ICCAT Coll. Vol. Sci. Pap. 18(2): 563-611
- Powers, J. E., and J. Cramer. 1996. An exploration of the nature of Atlantic bluefin tuna mixing. ICCAT Coll. Vol. Sci. Pap. 45(2): 173-181.
- Punt, A. E., and D. S. Butterworth. 1995. Use of tagging data within a VPA formalism to estimate migration rates of bluefin tuna across the north Atlantic. ICCAT Coll. Vol. Sci. Pap. 44(1): 166-182.
- Restrepo, V. R. and Porch, C. E. 2000. Options for conducting swordfish sex-specific assessments. ICCAT Collect. Vol. Sci. Pap. 51:1375-1386.
- Schweigert, J. F., and Schwarz, C. J. 1993. Estimating migration rates for Pacific herring (*Clupea pallasii*) using tag-recovery data. Can. J. fish. Aquat. Sci. 50: 1530-1540.
- Sibert, R., and D. A. Fournier. 1994. Evaluation of advection-diffusion equations for estimation of movement patterns from tag recapture data. FAO Fish. Tech. Pap. 336/1:108-121.
- SCRS. 2001. Report of the ICCAT SCRS West Atlantic Bluefin Tuna Stock Assessment Session (Madrid, Spain, September 18 - 22, 2000). ICCAT Col. Vol. Sci. Pap. 52: 831-958.
- Turner, S. C. 1986. An analysis of recaptures of tagged bluefin with respect to the mixing assumption. ICCAT Col. Vol. Sci. Pap. 24: 196-202.
- Turner, S. C., and J. E. Powers. 1995. Review of information related to Atlantic bluefin tuna east-west movement. ICCAT Col. Vol. Sci. Pap. 44(1): 191-197.
- Walter, JF, C.E. Porch. 2012. Three Strategies for Modeling Terminal-Year Fishing Mortality Rates in Virtual Population Analyses of Western Atlantic Bluefin Tuna: Retrospective Patterns and Consequences for Projections. Aquatic Living Resources 25(04): 333 – 340.

6. APPENDICES

APPENDIX 1. Sample CONTROL file with line numbers

First column of file

```
1 #####
2 ###          CONTROL FILE FOR PROGRAM VPA-2BOX, Version 4.01      ###
3 #####
4 # INSTRUCTIONS: the control options are entered in the order specified.
5 #           Additional comment lines may be inserted anywhere in this
6 #           file provided they are preceded by a # symbol in the FIRST
7 #           column, otherwise the line is perceived as free-format data.
8 #####
9 #
10 #####
11 # TITLES AND FILE NAMES (MUST BE PLACED WITHIN SINGLE QUOTES)
12 #####
13 #|-----must be 50 characters or fewer-----|
14 'NORTHERN ALBACORE 1975-97'          TITLE OF RUN
15 'ALB00.d01'                          DATA FILE NAME (INPUT)
16 'ALB00.p01'                          PARAMETER SPECIFICATION FILE (INPUT)
17 'ALB00.r01'                          RESULTS FILE NAME (OUTPUT)
18 'ALB00.e01'                          PARAMETER ESTIMATE FILE NAME (OUTPUT)
19 'ALB00.SPD'                          SPREADSHEET FRIENDLY RESULTS (OUTPUT)
20 'none for ALB '                      TAGGING DATA FILE (INPUT)
21 #####
22 # MODEL TYPE OPTIONS
23 #####
24 1          NUMBER OF ZONES (1 OR 2)
25 1          MODEL_TYPE (1=DIFFUSION, 2=OVERLAP)
26 #####
27 # TAGGING DATA SWITCH
28 #####
29 # tagging data switch (0=do not use tagging data, 1=use tagging data)
30 # | weighting factor for modifying importance of tagging data in objective function
31 # | Relative amount of fishing in each month (not necessary tag data switch = 0)
32 # |
33 # |
34 0 | 1.0 | 0 0 0 1 1 1 1 1 1 0 0 0          TAGGING MODEL CONTROLS
35 #####
36 # SEARCH ALGORITHM CONTROLS
37 #####
38 -911      RANDOM NUMBER SEED
39 20       MAXIMUM NUMBER OF AMOEBIA SIMPLEX SEARCH RESTARTS
40 3        NUMBER OF CONSECUTIVE RESTARTS THAT MUST VARY BY LESS THAN 1% TO STOP SEARCH
41 0.4     PDEV (standard deviation controlling vertices for Initial simplex of each restart)
42 #####
43 # INDEX WEIGHTING CONTROLS
44 #####
45 0       SCALE (DIVIDE INDEX VALUES BY THEIR MEAN)- ANY VALUE > 0 = YES
46 1.0    INDEX WEIGHTING:(0)INPUT CV's, (+)DEFAULT CV, (-)DEFAULT STD. DEV., (999)MLE
47 0      (0) MULTIPLICATIVE VARIANCE SCALING FACTOR or (1) ADDITIVE VARIANCE SCALING FACTOR
48 #####
49 # CONSTRAINT ON VULNERABILITY (PARTIAL RECRUITMENT)
50 #####
51 # apply this penalty to the last N years (SET N = 0 TO IGNORE)
52 # | standard deviation controlling the severity of the penalty
53 # | first age affected
54 # | last age affected
55 # |
56 0 | .4 | 1 | 9          LINKS THE VULNERABILITIES IN THE LAST N YEARS
57 #####
58 # CONSTRAINTS ON RECRUITMENT
59 #####
60 # apply this penalty to the last N years (SET N = 0 TO IGNORE)
61 # | standard deviation controlling the severity of the penalty
62 0 | .1          LINKS THE RECRUITMENTS IN THE LAST N YEARS
63 0 | .1 | 1          LINKS THE RECRUITMENTS OF THE TWO STOCKS
64 # |
65 # | ratio of stock (sex) 1 to stock (sex) 2 {a value of 1 means a 1:1 ratio}
66 #####
67 # CONSTRAINT ON SPAWNER-RECRUIT RELATIONSHIP
```

```
68 #####
69 # PDF of spawner-recruit penalty: 0=none, 1=lognormal, 2=normal (-)=estimate sigma by MLE
70 # | first and last years to use in fitting (in terms of recruits)
71 # | |
72 # 0 1975 1996 PENALIZES DEPARTURES FROM BEVERTON AND HOLT STOCK-RECRUIT CURVE
73 # (note: check the parameter file to make sure you are estimating the S/R
74 # parameters when pdf not 0, or not estimating them when pdf=0))
75 #####
76 # PARAMETER ESTIMATION OPTIONS
77 #####
78 # 1 OPTION TO USE (1) F'S OR (2) N'S AS TERMINAL YEAR PARAMETERS
79 # -1 ESTIMATE Q IN (+) SEARCH or (<0) by concentrated MLE's
80 #####
81 # BOOTSTRAP ANALYSES
82 #####
83 # Number of bootstraps to run (negative value = do a parametric bootstrap)
84 # | Use Stine correction to inflate bootstrap residuals (0=NO)
85 # | | File Output Toggle (-1 output as ASCII file, +1 output as Binary file)
86 # | | |
87 # 0 1 -1
88 #####
89 # RETROSPECTIVE ANALYSES (CANNOT DO RETROSPECTIVE ANALYSES AND BOOTSTRAPS AT SAME TIME)
90 #####
91 # 4 NUMBER OF YEARS TO GO BACK FOR RETROSPECTIVE ANALYSES
92 @EOF@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
```

APPENDIX 2. Sample data file with line numbers

First column of file

```

1 |
2 | #####
3 | # DATA FILE FOR PROGRAM VPA-2BOX, Version 4.01
4 | # The data and specifications are entered in the order indicated
5 | # by the existing comments. Additional comments must be preceded by a # symbol
6 | # in the first column, otherwise the line is perceived as free format input.
7 | #####
8 | 1975 1999 FIRST AND LAST YEAR
9 | 1 8 8 8 FIRST AGE, LAST AGE, PLUSGROUP AGE, EXPANDED PLUS GROUP AGE
10 | #####
11 | # BEGIN INPUT FOR Zone/STOCK 1
12 | #####
13 | 2 NUMBER OF INDICES in this file (whether used or not)
14 | 6 SPAWNING SEASON (elapsed months, 0 is beginning of year)
15 | 0 0 0 0 0.5 1 1 1 FECUNDITY MODIFIER (SPAWNING FRACTION) AT AGE
16 | # 50 character title within single quotes ' ' ---->] PDF OF CATCH
17 | # | SIGMA CATCH
18 | 'Northern Albacore' 0 1
19 | #####
20 | # NOW ENTER THE CATCH-AT-AGE DATA. ROW=YEAR, COLUMN=AGE
21 | #####
22 | #YEAR 1 2 3 4 5 6 7 8+ <--AGE
23 | 1975 315733 1066318 1237487 299015 280100 186106 212809 68034
24 | 1976 931707 2228301 935988 823222 448893 313768 151055 68898
25 | 1977 428823 2433768 1397655 321475 454251 277723 108428 72114
26 | 1978 2487904 2179912 1182421 418111 219033 213060 84188 31706
27 | 1979 883487 3295114 1767666 179078 150122 108815 51052 121773
28 | 1980 1661960 1372292 1533717 280723 85452 49805 31750 67929
29 | 1981 1127638 1524756 1019549 302997 74073 51145 42296 99892
30 | 1982 239468 1844670 1610921 284073 77559 70137 37154 173027
31 | 1983 890767 1577368 1628346 619212 207258 156937 80563 106758
32 | 1984 432954 1204226 972542 275997 209046 124572 123533 225416
33 | 1985 1052931 1354882 996685 235554 231200 137424 40219 212844
34 | 1986 821188 1559763 1172319 343754 409680 166179 78931 146896
35 | 1987 378762 2372142 1358300 146963 50677 45487 27354 70074
36 | 1988 1725728 2070614 1039478 140129 33163 15686 14680 21340
37 | 1989 1113168 1826467 1385188 106313 41240 23941 9374 10010
38 | 1990 1101655 2609762 790747 212087 97017 74879 60175 63898
39 | 1991 1197076 2190165 588433 96016 105347 34942 32053 51821
40 | 1992 1367443 1895016 833741 161041 47378 69838 56897 22928
41 | 1993 988442 2102203 1031346 208371 83237 52766 79599 178022
42 | 1994 708148 2863041 696340 160490 36962 27918 40370 163956
43 | 1995 1327523 2302110 992772 108378 118960 122804 78394 93902
44 | 1996 1343761 2438753 342995 96818 55599 39962 46529 77277
45 | 1997 1769533 1666480 763456 96324 49881 53179 39700 27793
46 | 1998 1847969 1933866 546087 83660 31479 32148 32259 36206
47 | 1999 1852736 1612315 1087974 248920 68592 52359 43632 50509
48 | -1 end of catch data
49 | #####
50 | # NOW ENTER IN THE ABUNDANCE INDEX SPECIFICATIONS
51 | #####
52 | # INDEX PDF (0= do not use,1=lognormal, 2=normal)
53 | # | UNITS (1 = numbers, 2 = biomass)
54 | # | VULNERABILITY (1=fixed, 2=frac.catches, 3=part. catches, 4=Butt. & Gero.
55 | # | TIMING (-1=average, +integer = number of months elapsed)
56 | # | FIRST AGE LAST AGE TITLE (IN SINGLE QUOTES)
57 | 1 1 1 1 -1 2 2 'SP. TROLL 2'
58 | 2 1 1 4 -1 2 8 'TAIWAN LL'
59 | -1 end index specifications
60 | #####
61 | # NOW ENTER IN THE INDICES OF ABUNDANCE
62 | #####
63 | #INDEX YEAR INDEX VALUE INDEX STD ERROR
64 | 1 1981 40.50319669 -1 SP Troll 2
65 | 1 1982 53.95664758 -1
66 | 1 1983 38.30118752 -1
67 | 1 1984 31.99824001 -1
68 | 1 1985 24.86051748 -1
69 | 1 1986 37.62169586 -1
70 | 1 1987 42.23633635 -1
71 | 1 1988 41.50778091 -1
72 | 1 1989 22.77312019 -1
73 | 1 1990 38.9929516 -1
74 | 1 1991 57.01309062 -1
75 | 1 1992 49.02747727 -1

```

75	1	1993	42.29127929	-1
76	1	1994	66.17355432	-1
77	1	1995	51.94995761	-1
78	1	1996	61.3982235	-1
79	1	1997	38.33567411	-1
80	1	1998	34.43869895	-1
81	1	1999	28.71963116	-1
82	2	1975	14.3106	-1
83	2	1976	16.0662	-1
84	2	1977	12.3064	-1
85	2	1978	14.0685	-1
86	2	1979	16.532	-1
87	2	1980	17.3829	-1
88	2	1981	16.3676	-1
89	2	1982	20.168	-1
90	2	1983	19.0735	-1
91	2	1984	15.3925	-1
92	2	1985	13.0577	-1
93	2	1986	11.33	-1
94	2	1987	10.342	-1
95	2	1988	17.2733	-1
96	2	1989	14.2442	-1
97	2	1990	11.6289	-1
98	2	1991	9.6744	-1
99	2	1992	11.0432	-1
100	2	1993	14.2994	-1
101	2	1995	17.2424	-1
102	2	1996	6.6744	-1
103	2	1997	7.9734	-1
104	2	1998	14.3053	-1

Chinese-Taipe LL

105 -1 end index data

106 #=====
 107 # NOW ENTER IN THE VULNERABILITIES OR PARTIAL CATCHES FOR THE INDICES OF ABUNDANCE
 108 #=====

#INDEX YEAR	AGE1	AGE 2	AGE 3	AGE 4	AGE 5	AGE 6	AGE 7	AGE 8		
109	2	1975	1	8004	117604	136701	80679	45508	53501	28105
110	2	1976	1	20625	110919	267439	206916	127664	62566	29343
111	2	1977	1	25506	130719	122264	227435	142092	60194	39084
112	2	1978	1	2358	58387	111893	122751	125723	50297	15838
113	2	1979	1	32263	71613	48715	68205	51369	30048	62593
114	2	1980	1	14023	119100	111482	56875	34060	21647	48676
115	2	1981	1	29024	62736	96319	38412	26434	27092	69689
116	2	1982	1	67783	151933	119530	58476	53898	25025	115500
117	2	1983	1	49625	162733	239222	118373	100967	52898	82696
118	2	1984	1	56542	128831	167504	151424	79661	62894	136840
119	2	1985	1	31390	106341	135704	176272	102264	29379	165136
120	2	1986	1	32052	92679	207010	348902	142385	65886	124171
121	2	1987	1	85954	104652	60162	27649	27904	19321	61656
122	2	1988	1	4083	71760	42457	6332	2291	3959	13382
123	2	1989	1	13	26034	19085	21453	11262	1522	349
124	2	1990	1	8	8887	23971	38597	11777	1697	4862
125	2	1991	1	14	94566	27532	92448	23152	7031	12421
126	2	1992	1	289	17639	22351	22613	31948	17784	6396
127	2	1993	1	19154	84338	65883	69497	24628	33747	56260
128	2	1994	1	25322	39961	26040	10477	10969	22885	121014
129	2	1995	1	18269	88299	49967	45340	33248	12944	7432
130	2	1996	1	22114	48631	34637	24666	11195	7296	48784
131	2	1997	1	2994	99172	33814	16296	20193	14837	10667
132	2	1998	1	316	89790	33267	11842	14431	15955	17393
133	2	1999	1	530	169504	62954	24445	24473	25948	33850

134 -1 end index vulnerabilities

135 #=====
 136 # NOW ENTER IN THE WEIGHTS AT AGE FOR THE INDICES OF ABUNDANCE (row=year, col=age)
 137 #=====

#Index year	age 1	age 2	age 3	age 4	age 5	age 6	age 7	age 8		
138	1	1975	2.59	5.88	10.55	15.19	18.75	22.92	29.64	38.99

139 -1

140 #=====
 141 # NOW ENTER IN THE FECUNDITY AT AGE FOR THE SPAWNING STOCK BIOMASS (row=year, col=age)
 142 #=====

#year	age 1	age 2	age 3	age 4	age 5	age 6	age 7	age 8	
143	1975	2.59	5.88	10.55	15.19	18.75	22.92	29.64	38.99
144	1976	2.52	5.94	10.27	15	19.17	23.57	28.9	39.01
145	1977	2.45	5.84	10.64	15.01	19.01	23.88	29.18	38.48
146	1978	2.69	6.2	10.55	15.04	19.59	23.54	28.78	39.14
147	1979	2.59	5.99	10.92	15.45	18.79	22.51	28.33	37.22
148	1980	2.65	6.03	10.57	15.81	19.06	22.93	26.91	37.07
149	1981	2.74	6.08	10.49	15.48	19.77	23.29	26.94	36.78
150	1982	2.49	5.88	10.67	15.75	18.61	21.58	26	36.37
151	1983	2.78	6.1	10.51	15.32	19.13	22.05	27.47	37.87
152	1984	2.84	6.01	10.64	15.18	19.32	21.58	26.02	37.14

156	1985	2.72	6.23	10.67	15.04	19.31	21.81	26.9	36.8
157	1986	2.64	6.1	10.66	15.21	19.06	22.83	28.87	37.45
158	1987	3.1	6.02	10.94	15.98	19.13	23.08	28.02	36.47
159	1988	2.81	6.28	10.88	15.28	19.6	23.18	28.08	37.33
160	1989	2.64	6.07	11.11	16.37	19	22.16	27.69	38.13
161	1990	2.65	6.07	10.7	15.39	19.54	22.37	27.46	38.1
162	1991	2.67	6.12	10.98	16.14	17.63	22.96	26.87	37.11
163	1992	2.68	6.08	10.7	15.56	18.64	22.22	28.49	38.84
164	1993	2.6	6.02	10.66	15.31	17.82	22.06	27.23	36.19
165	1994	2.56	6.04	10.86	15.66	19.27	22.01	27.25	36.83
166	1995	2.59	6.15	10.93	15.16	18.27	22.34	27.30	37.61
167	1996	2.62	6.37	11.09	15.62	18.87	21.98	27.60	36.84
168	1997	2.65	6.20	11.07	15.69	18.27	22.95	29.10	37.97
169	1998	2.65	6.02	10.89	15.97	18.36	22.28	27.75	37.89
170	1999	2.65	6.12	10.67	15.78	18.92	22.98	28.10	37.51
171		-1							
172	#####								
173	# BEGIN INPUT FOR Zone/STOCK 2								
174	#####								

APPENDIX 3. Sample Tag Data file with line numbers

```

=====
First column of file
|
1 #-----
2 # Release information
3 #-----
4 # tag cohort id number
5 #|   area released
6 #|   |   stock affiliation (enter 0 if the stock of origin is unknown)
7 #|   |   |   year released
8 #|   |   |   |   month released (time elapsed)
9 #|   |   |   |   |   year experiment ends (last year when data are collected)
10 #|   |   |   |   |   |   month experiment ends (last month when data are collected)
11 #|   |   |   |   |   |   |   age at release
12 #|   |   |   |   |   |   |   |   number released (enter negative number to temporarily exclude tags from analysis)
13 #|   |   |   |   |   |   |   |   |   weight given to tagging data in objective function (acts like a standard deviation)
14 #|   |   |   |   |   |   |   |   |   |
15 #|   |   |   |   |   |   |   |   |   |   *****
16 #|   |   |   |   |   |   |   |   |   |   * indices linking tag cohorts with parameter estimates *
17 #|   |   |   |   |   |   |   |   |   |   *****
18 #|   |   |   |   |   |   |   |   |   |   dmapi (initial survival of tags)
19 #|   |   |   |   |   |   |   |   |   |   λmapi (chronic tag loss)
20 #|   |   |   |   |   |   |   |   |   |   ρmapij (reporting rates)   γmapij (partial-mixing factors)
21 #|   |   |   |   |   |   |   |   |   |   -----
22 #|   |   |   |   |   |   |   |   |   |   by remote by fisheries   for   for
23 #|   |   |   |   |   |   |   |   |   |   sensing   area 1   area 2   year 1   year 2
24 #|   |   |   |   |   |   |   |   |   |   j=1     j=2     j=3     j=1     j=2
25 #|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
26 #-----
27 1 2 2 1998 8.6 2000 8.9 10 50 1 1 1 1 2 3 1 2
28 2 2 2 1998 8.6 2000 8.9 10 -50 1 1 1 1 2 3 1 2
29 #-----
30 # Recapture data
31 #-----
32 # Format of data (enter -1 to read data using the standard format for all tag types that reads:
33 #|   cohort, area recovered, number recovered at end of experiment, number recaptured by year starting with year of release
34 #|   -2 to read the data using the old format for conventional tags that reads:
35 #|   cohort, area recovered, number recaptured by year starting with first year in VPA
36 -1 entry must be negative!
37 #-----
38 # tag cohort id number
39 #|   area recovered
40 #|   |   number recovered in area at end of experiment (e.g., by survival)
41 #|   |   |   number recovered from fisheries in area during each year of study (beginning with the year the study began and ending with
42 #|   |   |   |   either the year the study ended or else the last year of the analysis)
43 #|   |   |   |   |   1   2   3   (1998, 1999, 2000)
44 #|   |   |   |   |   |   |   |
45 1 1 7.472967228 0.896763349 2.755045041 1.063809403
46 1 2 14.49862687 1.834338914 4.13903679 1.85780102
47 2 1 8 1 3 1
48 2 2 15 5 5 3
49 -1 End of Recapture Data

```

APPENDIX 4. Sample parameter specification file with line numbers

```

=====
First column of file
|
1 #####
2 # PARAMETER FILE FOR PROGRAM VPA-2BOX, Version 4.01
3 #
4 # The specifications are entered in the order indicated
5 # by the existing comments. Additional comments must be preceded by a # symbol
6 # in the first column, otherwise the line is perceived as free format input.
7 #
8 # Each parameter in the model must have its own specification line unless a $
9 # symbol is placed in the first column followed by an integer value (n), which
10 # tells the program that the next n parameters abide by the same specifications.
11 #
12 # The format of each specification line is as follows
13 #
14 # column 1
15 # | number of parameter to which these specifications apply
16 # | lower bound
17 # | best estimate (prior expectation)
18 # | upper bound
19 # | method of estimation
20 # | standard deviation of prior
21 # |
22 # $ 5 0 1.2 2.0 1 0.1
23 #
24 # The methods of estimation include:
25 # 0 set equal to the value given for the best estimate (a fixed constant)
26 # 1 estimate in the usual frequentist (non-Bayesian) sense
27 # 2(0.3) estimate as a random deviation from the previous parameter
28 # 3(0.1) estimate as a random deviation from the previous constant or type 1 parameter
29 # 4(0.2) estimate as random deviation from the best estimate.
30 # -0.1 set equal to the value of the closest previous estimated parameter
31 # -n set equal to the value of the nth parameter in the list (estimated or not)
32 #####
33 #-----
34 # TERMINAL F PARAMETERS: (lower bound, best estimate, upper bound, indicator, reference age)
35 # Note 1: the method indicator for the terminal F parameters is unique in that if it is
36 # zero but the best estimate is set to a value < 9, then the 'best estimate'
37 # is taken to be the vulnerability relative to the reference age in the last
38 # (fifth) column. Otherwise these parameters are treated the same as the
39 # others below and the fifth column is the standard deviation of the prior.
40 # Note 2: the last age is represented by an F-ratio parameter (below), so the number
41 # of entries here should be 1 fewer than the number of ages
42 #-----
43 # $ 1 0 .2 3 0 2 first age (here age 1)
44 # $ 1 0 .4886 2 1 .1
45 # $ 1 0 .5118 2 1 .1
46 # $ 1 0 .1416 2 1 .1
47 # $ 1 0 .2439 2 1 .1
48 # $ 1 0 .1738 2 1 .1
49 # $ 1 0 .9 2 0 6 next to last age
50 #-----
51 # F-RATIO PARAMETERS F{oldest}/F{oldest-1} one parameter (set of specifications) for each year
52 #-----
53 # $ 25 0.1 1.0 5.0 0 .2 fixed
54 #-----
55 # NATURAL MORTALITY PARAMETERS: one parameter (set of specifications) for each age
56 #-----
57 # $ 1 0 0.30 1.0 1 .1
58 # $ 1 0 0.30 1.0 1 .1
59 # $ 6 0 0.30 1.0 -0.1 .1
60 #-----
61 # MIXING PARAMETERS: one parameter (set of specifications) for each age
62 #-----
63 # $ 8 0 0.0 1.0 0 .1
64 #-----
65 # STOCK-RECRUITMENT PARAMETERS: five parameters so 5 sets of specifications
66 #-----
67 # 0 0.11d+08 1.D20 0 0.4 maximum recruitment
68 # 0 0.85d+04 1.D20 0 0.0 spawning biomass scaling parameter

```

```

67 0 0.000 0.9 0 0.0 extra parameter (not used yet)
68 0 0 1 0 0 autocorrelation parameter
69 0 0.4 2 0 0 variance of random component (discounting the autocorrelation)
70 #=====
71 # VARIANCE SCALING PARAMETER (lower bound, best estimate, upper bound, indicator, std. dev.)
72 # this parameter scales the input variance up or down as desired
73 # In principal, if you estimate this you should obtain more accurate estimates of the
74 # magnitude of the parameter variances-- all other things being equal.
75 # (1 parameter so 1 set of specifications)
76 #-----
77 $ 1 0 0.5 1.0 1 .1
78 $ 1 0 0.5 1.0 -.1 .1
79 #=====
80 # CATCHABILITY PARAMETERS (do not try to estimate q's for indices you are not using!)#
81 #-----
82 # index 1
83 $ 1 0. .0 1.0 1 .2 estimate
84 $ 24 0. .00001 2.0 -.1 .2 set equal to above estimate
85 # index 2
86 $ 1 0. .00001 1.0 1 .2 1975 estimate
87 $ 23 0. .00001 1.0 3 .2 1976-98 random walk
88 $ 1 0. .0 1.0 0 .2 1999 ignore (don't estimate)
89 #
90 #=====
91 # dvector: FRACTION SURVIVING THE INITIAL TAGGING PROCESS
92 #-----
93 0.0000D+00 1.0D+00 0.1000D+01 0.0 0.1000D+00
94 #=====
95 # λvector: CHRONIC TAG LOSS (SHEDDING) RATE
96 #-----
97 0.0000D+00 0.2 0.1000D+01 1.0 0.1000D+00
98 #=====
99 # ρvector: TAG REPORTING RATE
100 #-----
101 0.0000D+00 1.000D+0 0.1000D+01 0 0.1000D+00 satellite
102 0.1000D+00 1.00000 0.1000D+01 0.0 0.1000D+00 east <1989
103 0.1000D+00 1.00000 0.1000D+01 0.0 0.1000D+00 west <1989
104 0.1000D+00 1.00000 0.1000D+01 0.0 0.1000D+00 east 1990-
105 0.1000D+00 1.00000 0.1000D+01 0.0 0.1000D+00 west 1990-
106 #=====
107 # γvector: PRE-MIXING ADJUSTMENT
108 #-----
109 0.0000D+00 1.00000D+00 0.5000D+01 1.0 0.1000D+00
110 0.0000D+00 1.00000D+00 0.5000D+01 1.0 0.1000D+00
111 @ end of file

```