Chapter 4. ASSESS menu

What you will learn from this chapter

This chapter presents routines for analysis of the various types of data presented in the previous chapter. However, these are presented in a summarized form, and we remind FiSAT users to read the FiSAT reference manual (Gayanilo and Pauly, 1998), and other documents explaining the background of these routines.

The format in module description

The Assess menu is divided into two parts, (i) the routines to estimate parameters (e.g., L_∞, K, C, M, Z, etc.) and (ii) the routines to predict yield or stock-related attributes given certain fishing scenarios. In the following, the modules are presented as given in Chapter 3, with one description added:

Output(s)   *This section defines the output of the model or methodology being described. Examples will be provided when necessary.*

Parameter estimation

The step that should immediately follow data entry and manipulation is the estimation of population parameters. These are required inputs to the second set of analytical routines that deals with predictions.

Direct fit of length-frequency data

The set of routines classified under this heading are those which estimate growth parameters directly from the length composition of the stock, without previously translating the

length scale into an age scale, as an alternative to the MPA approach described below.

ELEFAN I

ELEFAN I is a routine that can be used to identify the (seasonally oscillating) growth curve that "best" fits a set of length-frequency data, using the value of R_n as a criterion. FiSAT II provides three options to the user to identify that "best" growth curve: (1) curve fitting by eye (plotting of the histogram or restructured data may also be accessed from the Support Menu): (2) scan of K-values (Fig. 4.1), and (3) response surface analysis (Fig. 4.2). The plotting and curve fitting by eye is described in detail in the Support Menu below.

Note that in ELEFAN I, the parameter t_o is replaced by the coordinates of a point (any point actually) through which the curve must pass, and whose coordinates consist of SS (a starting sample) and of SL (a starting length).

Required file	Time series of length-frequency data with constant class size.
Input parameters	**Option 1**: Scanning of K-values
	Fixed value of L_∞, C and WP and, as an option, a fixed starting point.
	Option 2: Response surface analysis
	Range of values for two parameters (e.g., L_∞ and K), fixed value for the two others (e.g., C and WP), and, as an option, a fixed starting point.
Functions	In ELEFAN I, data are reconstructed to generate "peaks" and "troughs", and the goodness of fit index (R_n) is defined by $R_n = 10^{ESP/ASP}/10$

where the ASP ("Available Sum of Peaks") is computed by adding the 'best' values of the available 'peaks' and the ESP ("Explained Sum of Peaks") is computed by summing all the peaks and troughs "hit" by a growth curve of the form,

$$L_t = L_\infty (1 - EXP(-K(t - t_o) + S_{ts} + S_{to})$$

where

$S_{ts} = (CK/2\pi) \cdot \sin(2\pi(t\text{-}ts))$,
$S_{to} = (CK/2\pi) \cdot \sin(2\pi(t_O\text{-}ts))$, and
L_t is the length at time t.

Outputs **Option 1**: Scanning of K-values

This plots R_n values for a range of K values (0.10 to 10) on a log-scale.

We recommend use of this plot for <u>all</u> growth analyses, if only to assess how reliable an estimate of K is.

Option 2: Response surface analysis

This outputs an 11 by 11 matrix showing R_n values and in which the 10 best values are highlighted, thus enabling selection of the "best" combination of growth parameters.

User interface ELEFAN I contains three tabs. The first tab is only to identify the length frequency file to use. The tab, "K Scan" (Fig. 4.1), would allow the user to view

the behaviour of the scores given estimates of L_∞.

The starting point can be set to variable or fixed starting point (see red arrow in Fig. 4.1). If starting point is set to a fixed value, the starting sample and starting length can be selected from a drop-down list.
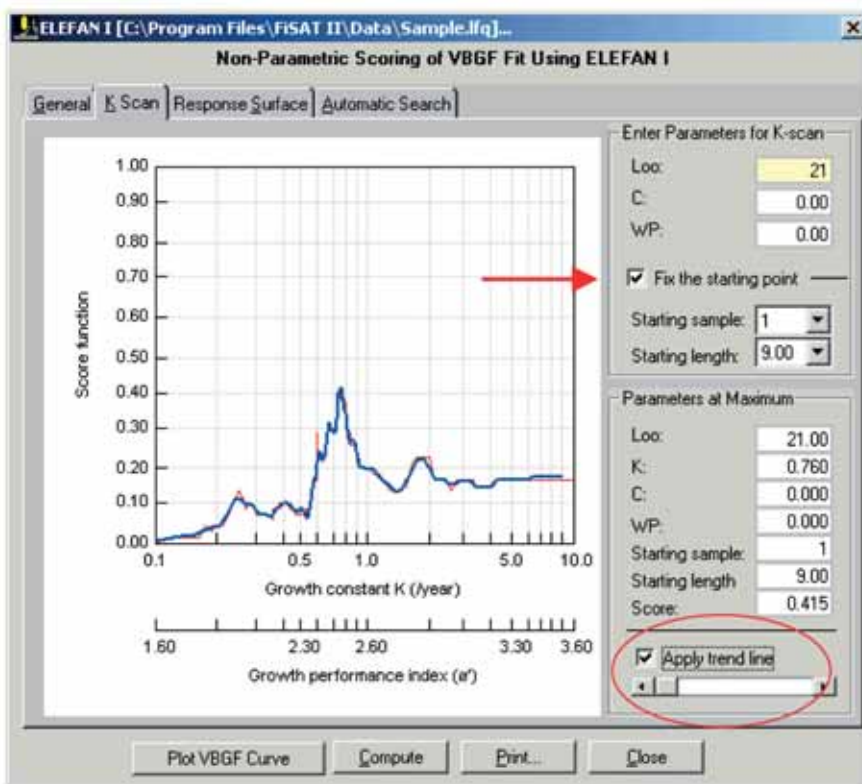

Fig. 4.1. User interface of ELEFAN I, K-scan routine. The red line in the plot is the actual score and the thick blue line is the overlaid trend line.

The behaviour of the scores may be so erratic that assessment of the possible solution is difficult. A trend line (see red circle in Fig. 4.1) may be applied.

The scroll bar is used to change the resolution of the trend line. The third tab is to execute a response surface analysis (Fig. 4.2). The starting point is a requirement of this subroutine. Note that the value of the L_∞ cannot be less than the starting length.
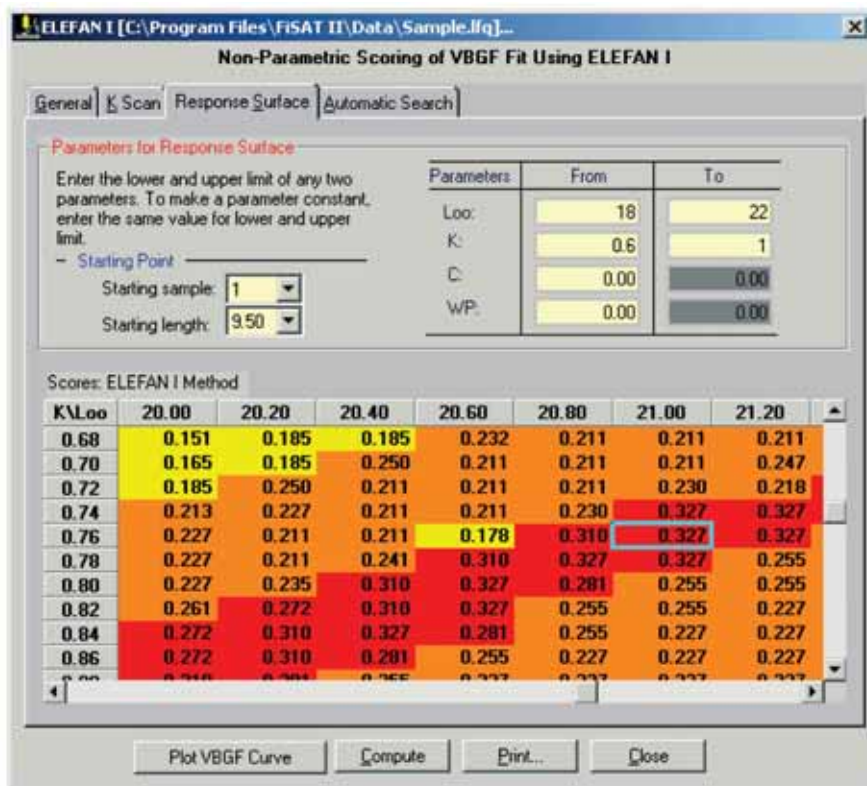

Fig. 4.2. User interface of ELEFAN I, Response surface analysis. Note the change of colour indicating improvement on the score.

Remarks	It is imperative that users of ELEFAN I read detailed accounts of this method.
Readings	Pauly (1982)
	Pauly and David (1981)
	Pauly and Morgan (1987)

Shepherd's method

Conceptually, this approach is similar to ELEFAN I in that it is designed to maximize a non-parametric scoring function. Two options for identifying optional values of L_∞ and K are available: (1) response surface analysis and (2) scan of K-values, both of which are very similar, in display and operations, to those of the ELEFAN I routine (see above).

Required file	Time series of length-frequency data with constant class size.
Input parameters	**Option 1**: Response surface analysis
	Range of values for L_∞ and K.
	Option 2: K-scanning
	Fixed value of L_∞.
Functions	The score (S) for Shepherd's method is defined by:

$$S = (s_A^2 + s_B^2)^{\frac{1}{2}}$$

where s_A and s_B are the goodness-of-fit scores (s_{tz}) obtained with the origin of the VBGF in calendar time (t_z) set to 0 and 0.25, respectively. s_{tz} is defined by:

$$s_{tz} = \sum_i T_i \cdot \sqrt{N_i}$$

where

N_i = frequency for length group i,
$T_i = D \cdot \cos 2\pi (t-t_i)$,
$D = (\sin\pi (\Delta t)/\pi(\Delta t))$,
$t = \Delta t/2$,
$\Delta t = t_{max} - t_{min}$,
$t_i = t_z - (1/K)\cdot\ln(1-(L_i/L_\infty))$, and
$t_z = (1/2\pi) \cdot \tan^{-1}(s_B/s_A)$.

Outputs

Option 1: This outputs an 11 by 11 matrix of S values (with S_{max} standardized to 1) and in which the 10 best scores are highlighted, thus enabling selection of the "best" combination of L_∞ and K;

Option 2: This generates a plot of S values (with S_{max} standardized to 1) for a range of K values (0.1 to 10 year^{-1}) on a log scale, thus enabling the identification of the best value of K for a given value of L_∞.

User interface

The user interface of this routine is very similar to that in ELEFAN I. It also contains three tabs where the first tab is used only to identify the file to be analysed. The second tab (Fig. 4.3) is to examine the S values given a fixed L_∞ and a fixed range of K values.
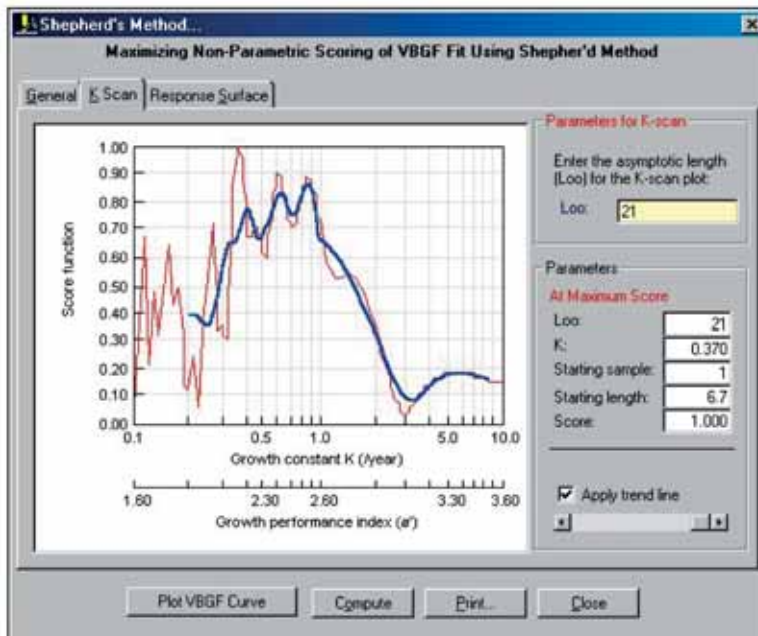
Fig. 4.3. User interface of Shepherd's method. The trend line indicates a K value in the same range as that generated by ELEFAN I in Fig. 4.1.

The third tab (Fig. 4.4) is the response surface analysis given a range of L_∞ and K values. The main difference with ELEFAN I is that, in this approach, the starting point and seasonality parameters (C and WP) are not identified.
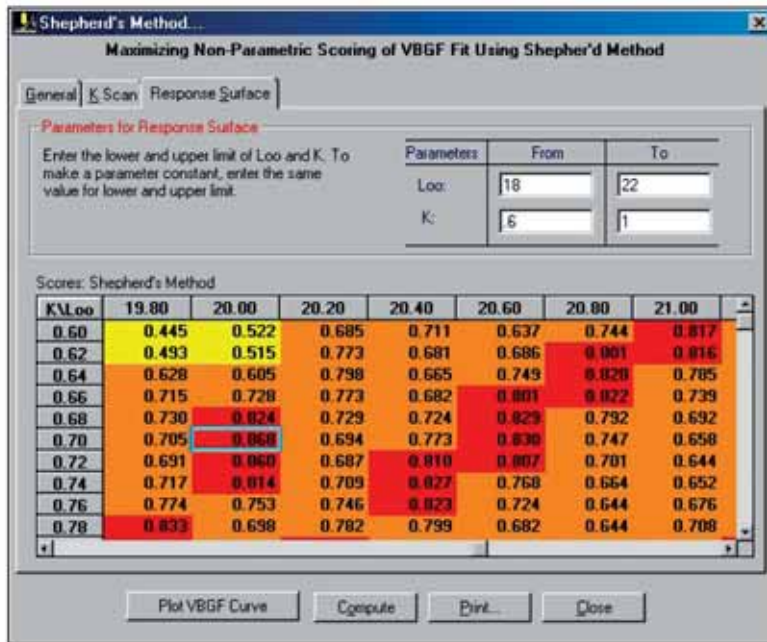
Fig. 4.4. Results of the response surface analysis where red (i.e. darker coloured cells) indicates improvement on the score.

Remarks

As for the ELEFAN I routine, this method is best applied when L_∞ has been estimated using another method (e.g., the Powell-Wetherall plot, see below). Note however, that this method differs from ELEFAN I in that it cannot deal with seasonal growth oscillations. Also, note that in this implementation, the score function is standardized to 1, i.e. S_{max} is made equal to unity. Further, the parameter "t_z" (which is *not* similar to t_o), is here replaced by a "starting point" (as for ELEFAN I, see above), and hence the ELEFAN I output routine can be used to display growth curves whose parameters have been estimated using Shepherd's method.

Readings

Shepherd (1987) and Isaac (1990)

Powell-Wetherall plot

This method allows estimation of L_∞ and Z/K from a sample representing a steady-state population, as can be approximated by pooling a time series of length-frequency data.

Required file   Length-frequency   data   file   with constant class size

Input parameter   Graphical   identification   of   smallest length fully recruited by the gear (L', or cut-off length):

Functions $(\overline{L}\text{-}L') = a + b \cdot L'$

where

$$\overline{L} = \left(\frac{L_\infty + L'}{1 + (Z/K)} \right)$$

from which

L_∞ = -a/b,
Z/K = - (1+b)/b.

Outputs L_∞ and Z/K estimates and regression parameters.

User interface   In this routine, as with other routines in FiSAT II, options are provided to select and deselect samples to be included in the analysis. To deselect a sample, uncheck the sample represented by the sampling date. Three options are provided to do a temporary data transformation (see Chapter 3 for details). Use the radio buttons (Fig. 4.5) to select the option. In addition to the

available temporary data transformation, temporal weighting may also be applied to the selected samples by checking the appropriate box (see Chapter 3 for details of this option).
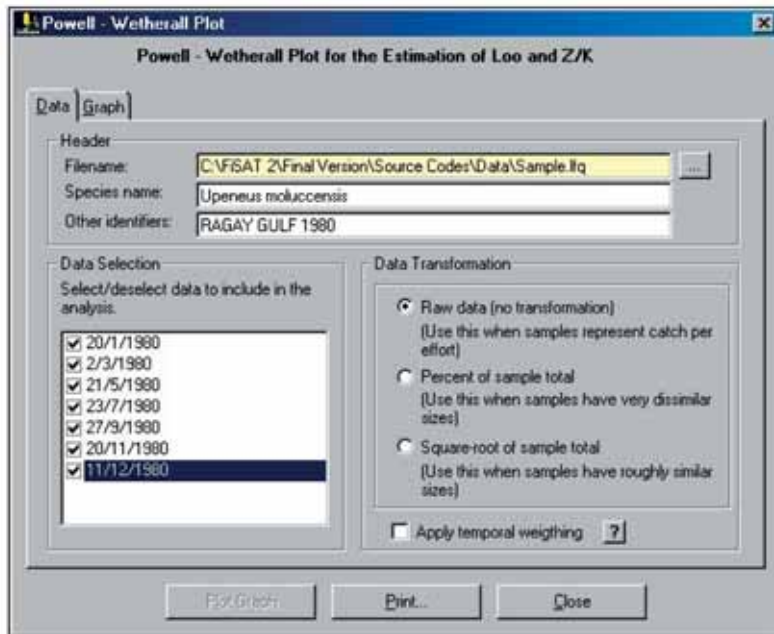

Fig. 4.5. File identification tab and options to temporarily transform data before data analysis. These options are available in other routines of FiSAT II that sums a time series of length frequencies to approximate a steady-state condition.

Once the initial inputs are satisfied, the data points may be plotted (Fig. 4.6). In FiSAT II, the identification of the cut-off length (L') is facilitated by an adjacent graph (pseudo-catch curve; Fig. 4.7), showing one's L/F data, and from which L' is often identifiable. The cut-off length may be identified using a mouse pointing device directly from the Powell-Wetherall plot or from the pseudo-catch curve to estimate L_∞ and Z/K or from the pseudo-catch curve.

A command button, "Open pseudo-catch curve" is available in Fig. 4.6, if the pseudo-catch curve (to help identify the cut-off length) was previously closed.
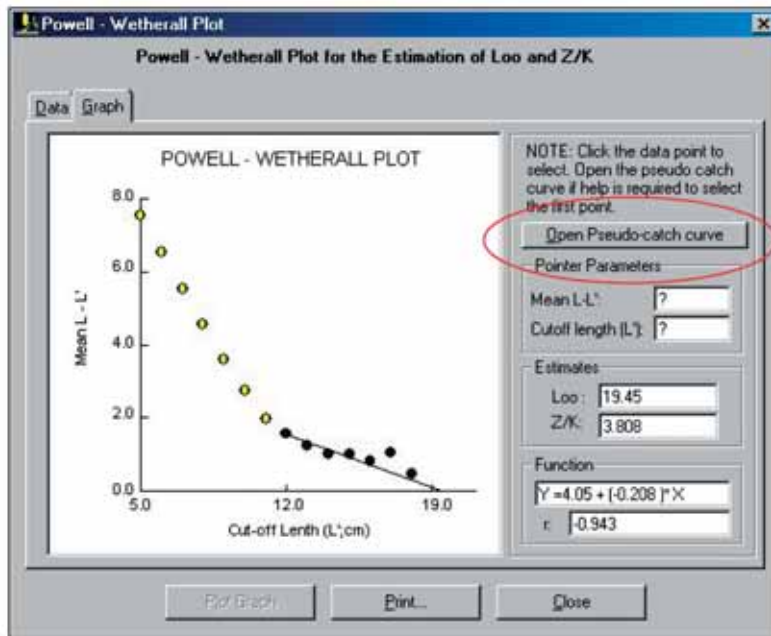

Fig. 4.6. The Powell-Wetherall's plot showing the regression line with the Y-intercept as the estimate of L_∞.
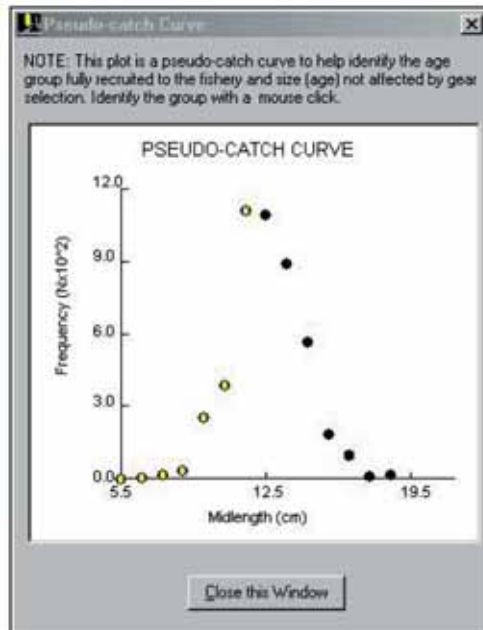
Fig. 4.7. Pseudo-catch curve to help identify cut-off length in Powell-Wetherall's plots. Selecting the cut-off length can also be done through this plot.

Remarks   This approach should be used only when sufficient samples are available.

Readings   Pauly (1986)
Powell (1979)
Wetherall (1986)

Analysis of length-at-age data

This routine allows non-linear estimation of growth parameters from length-at-age data.

Required file   Length-at-age data, consisting of at least four age-length data pairs.

Input parameters **Option 1: Non-seasonal growth**

The minimum and maximum constraints and starting values for the growth parameters L_∞, K, and t_0.

Option 2: Seasonal growth

The minimum and maximum constraints, and starting values for the growth parameters L_∞, K, t_0, C and WP (where WP = $t_s + 0.5$).

Functions Find the global minimum of the following functions using Marquardt's algorithm:

Option 1

$$SSE = \sum_i \left\{ L_i - L_\infty \left[1 - EXP(-K(t_i - t_0)) \right] \right\}^2$$

Option 2

$$SSE = \sum_i \left\{ L_i - L_\infty \left[1 - EXP(-K(t_i - t_0) + S_{ts} + S_{t0}) \right] \right\}^2$$

where

$S_{ts} = (C \cdot K/2\pi) \cdot \sin(2\pi(t - t_s))$,
$S_{to} = (C \cdot K/2\pi) \cdot \sin(2\pi(t_0 - t_s))$.

Outputs L_∞, K, t_0, and their standard errors for **Option 1** and in addition, C and WP for **Option 2**. For either option, the coefficient of determination (r^2 and adjusted r^2), phi prime (ø'), and its standard error.

User interface   The user interface of this routine contains only two tabs (Fig. 4.8 and Fig. 4.9). The first tab (Fig. 4.8) provides the option to enter the range of values identified by the minimum and maximum constraints, and the starting estimates for each parameter. After the iteration process, the results are also displayed in this tab.
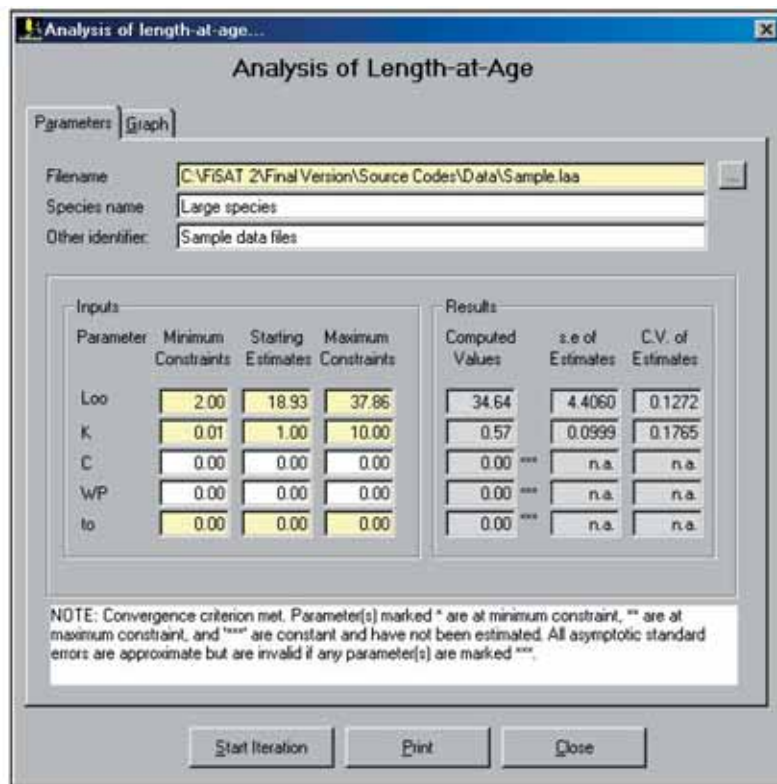

Fig. 4.8. The "Parameters" tab of the user interface when analysing length-at-age data. This tab identifies the file and range of values to search for the growth parameters. The results, including the statistics to measure the fit, are also displayed after iteration.

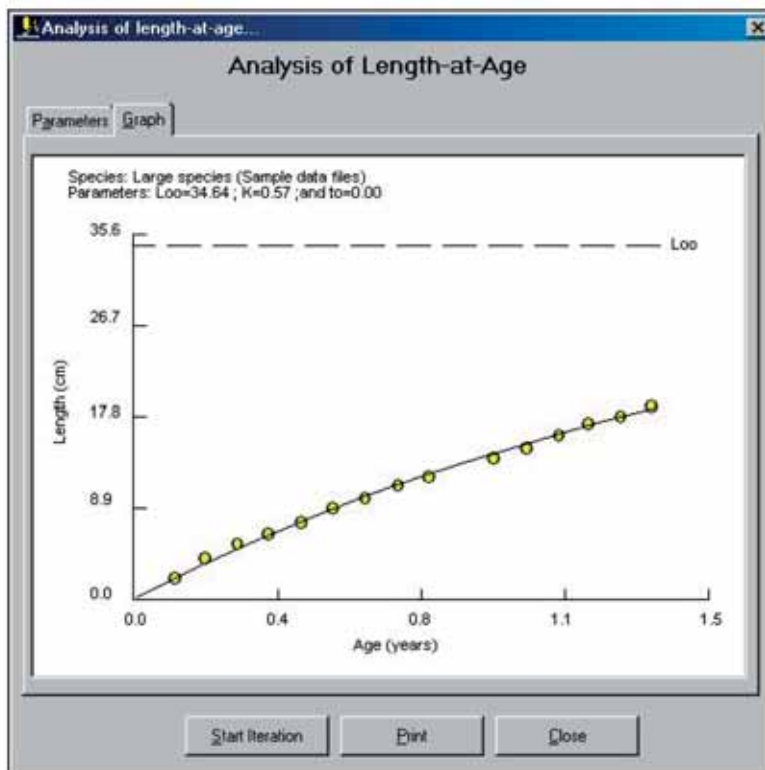The second tab (Fig 4.9) is a plot of the results, i.e. the data points overlaid by the computed growth curve.

Fig. 4.9. Plot of the growth curve based on the computed growth parameters.

Remarks   For rapid convergence, the routine should be initiated with a *wide* range of constraints. To fix inputs at certain set values, enter minimum constraint = maximum constraint.

Note that up to 200 iterations may be performed until the routine converges or stops.

The method for estimating the standard error of ø' stems from J. Hoenig (pers. comm.).

Readings   Kvalseth (1985)
Pauly *et al*. (1992)
Saila, Recksiek and Prager (1988)
Soriano and Jarre (1988)

Modal progression analysis

Modal progression analysis (MPA) refers to a methodology that infers growth from the apparent shift of the modes or means in a time series of length-frequency samples. MPA involves three stages: (i) decomposition of composite distributions into their components to identify means, (ii) subjective identification and "linking" of the means perceived to belong to the same cohorts and (iii) using the growth increments and size-at-age (relative) data resulting from the linking to estimate growth parameters.

In FiSAT II, two methods are provided to decompose composite length-frequency distributions:(i) Bhattacharya's method and (ii) NORMSEP.

Bhattacharya's method

Required file   Length or weight frequencies with constant class size.

Input parameters   Visual identification of frequencies perceived to belong to one age group, using the graphs provided for this.

Functions $\ln(N_{i+1}) - \ln(N_i) = a_j + b_j \cdot L_i$

where

N_i and N_{i+1} are successive frequencies of the same component of a group of fish in a sample (i.e. representing age group j) and where L_i is the upper class limit of N_i.

From this, the mean of the normal distribution is

$$L_j = -a_j/b_j$$

while its standard deviation (σ) is

$$\sigma_j = (-\Delta\overline{L}/b_j)^{1/2}$$

where $\Delta\overline{L}$ is the constant class size.

Also, a separation index (SI) is computed.

$$SI = \Delta\overline{L}_j/\Delta\sigma_j$$

where $\Delta\overline{L}_j$ is the difference between two successive means, and $\Delta\sigma_j$ is the difference between their estimated standard deviations.

The separation of length-frequency samples into their components is an iterative process in that every identified component is subtracted from the remainder of the sample using the Gaussian function,

$$N_{2i+1} = N_{i+} - \left\{\left(\frac{1}{\sigma_j\sqrt{2\pi}}\right) \cdot \text{EXP}\left(-\left(\frac{(L_i - L_j)^2}{2\sigma_i^2}\right)\right)\right\}$$

where N_{i+} refers to the previous set of frequencies and N_{2i+} is the new set of frequencies, less the component identified so far. Also, the user can backtrack and redo any step of the analysis without losing intermediate results.

Outputs Mean lengths, population sizes (in numbers), standard deviations and separation index (SI) for the age groups identified.

These results are automatically saved to disk as a "mean and standard deviation" file, using the same filename as the input.

User interface The module contains three tabs (Fig. 4.10). As with previous routines, the first tab is used only to identify the file to analyse. The "Plots" tab shows the log-plots of the slopes between successive frequencies and the resulting distributions.

The process starts with the selection of the sample to analyse (see red circle in Fig. 4.10) and identification of the first and last data points that identify a group using the mouse.
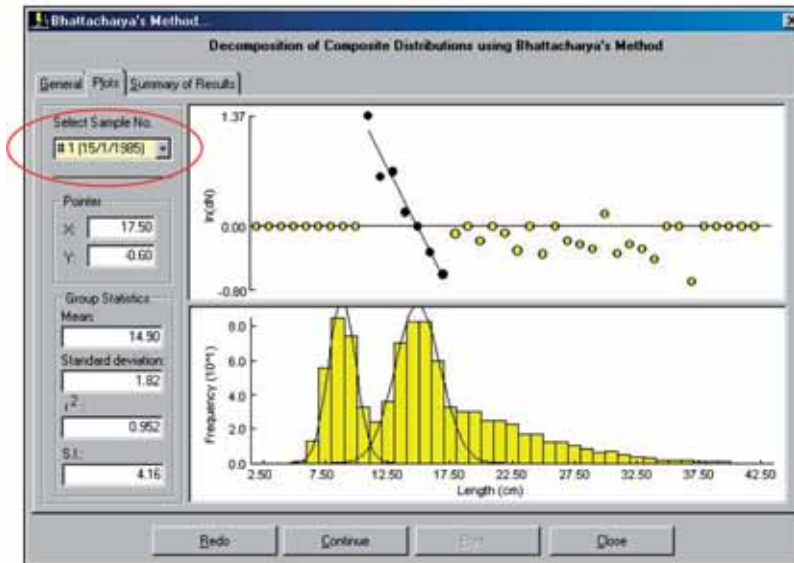
Fig. 4.10. User interface for Bhattacharya's method in FiSAT II.

Clicking the "Redo" command button will re-initialize the display and the user will be provided with the option to reselect data points. Clicking the "Continue" command button indicates acceptance of the results and the group will be eliminated from the composite distribution to emphasize the next possible group (if any). This process is continued until all groups are identified. The result can be saved (Fig. 4.11) by clicking the "Save MSD" command button that replaces the "Continue" command button after group identification. The results will be saved as a "mean and standard deviation file" that can later be linked to form growth increments or saved as length-at-age (relative) data.