

Chapter 3. FILE menu

What you will learn from this chapter

In this chapter, you will be introduced to the different types of data that can be created and used in FiSAT II, how these data are encoded, edited, imported, exported and options available for manipulating data files.

Types of files

FiSAT II supports several types of data differentiated largely by their filename. Table 3.1. is a summary of the different file-types that can be created in FiSAT II.

Table 3.1. Summary of file-types supported in FiSAT II and their corresponding file extensions (see details of each file type below).

File –Type	File Extension
Grouped length frequencies	*.lfq
Ungrouped length frequencies	*.ulf
Variable grouping of length frequencies	*.vlf
Length-at-age	*.laa
Growth increments	*.gin
Catch-at-age	*.caa
Mean and s.d.	*.msd
Probabilities	*.prb
Value by size	*.val
XY regression file	*.xyf
Species table	*.spc
Fleet table	*.flt
Multispecies/multifleet scenario	*.sce

Note that, although FiSAT supports a number of data types, most of the routines available in FiSAT either require length-frequency data or data derived from length-frequencies. The different data files that can be created and stored using FiSAT are:

Time series of length frequencies

Data files of this form can either be grouped by constant class size (i.e. class interval), grouped by variable class sizes (common for data from commercial sampling of shrimp) or "ungrouped frequencies", i.e. individual measurements requiring subsequent grouping.

Length frequencies grouped by constant class size

Length frequencies of this type are characterized by lower and upper class limits, and a "mid-length", the mean of the limits (Fig. 3.1). FiSAT II does not require inputs for the upper and lower class limits but rather uses the class size (or intervals) and the smallest mid-length, which enables the mid-lengths of the subsequent classes to be computed.

Unlike other data types, this data requires inputs to sample sub-headers (Table 3.2) that include, amongst other things, the sampling date in the format (dd/mm/yyyy). The data file will not be properly saved if the sampling date is not encoded properly. Samples with invalid or missing sampling dates will be automatically deleted. A command button is available in the user interface (Fig. 3.1) to open the sample sub-header. Alternatively, double-clicking on the column or clicking on the column header opens the sample sub-header.

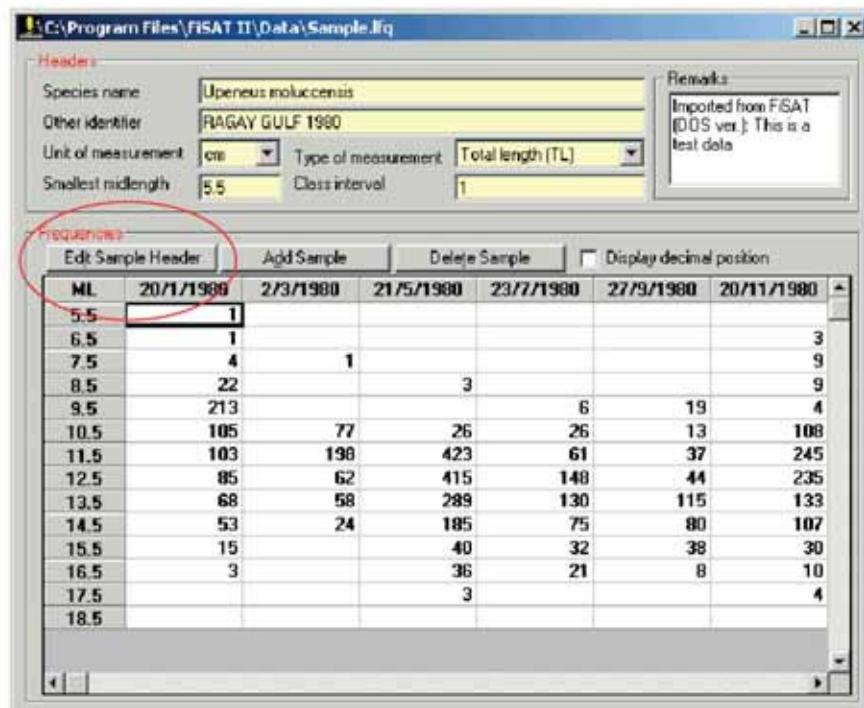


Fig. 3.1. User interface for recording grouped length frequencies. Note the command button to edit the sample sub-header. Alternatively, double-clicking on the column opens the form to encode the parameters for the sub-header.

The required inputs are:

- Date of sampling (DD/MM/YYYY)
- Smallest class mid-length (from 0.25 to 500)
- Class size (from 0.25 to 500)
- Type of measurement
- Frequencies (0 to 999,999) for each class.

WARNING! Although it is possible to copy contents from other Microsoft Windows software through the Microsoft clipboard (i.e. Copy->Paste procedure), such as Microsoft Excel, care should be taken to check for the input to sampling dates.

Table 3.2. Sub-header for size (length or weight) frequency data grouped by constant class interval.

Field	Type [†]	Remarks
Sample No.	N	The values for this field are automatically provided by FiSAT and cannot be edited.
Day of sampling	N	The upper limit is dependent on the month and year. A default DAY (equal to 15) is provided for convenience.
Month of sampling	N	The default value is the previous month + 1.
Year of sampling	N	Year of sampling, from 1900 to 2100; the default for this field is the year of the previous sample.
Sex*	C	Valid entries to this field are: female, male or undefined/both. Only the first letter of the choice needs to be entered; FiSAT will complete the entry when <ENTER> is pressed.
Depth (m)	C	This field should indicate the (mean) depth at which the sample was obtained.
Area	C	This indicates the area from which a sample originates.
Latitude	C	The input to this field, indicating the location of the sample will not be validated. However, it is recommended that a standard format be used, e.g., XDD:MM:SS, where X refers either to N (north) or S (south), DD refers to degrees of latitude, MM to minutes and SS to seconds (example: S35:30:00)
Longitude	C	As above. Thus, YDDD:MM:SS, where Y is either E (east) or W (west), DDD refers to degrees of longitude, MM to minutes and SS to seconds (example: W12:10:30)
Gear type	C	This field identifies the gear type used for sampling (example: <i>Trawl</i>).
Gear characteristics	C	This field should describe specific characteristics of the gear used (example: 2.5 cm mesh)
Vessel	C	This field should contain the vessel name or type.

[†]C - character field; N - numeric field; Ch - limited choice field

*The fields from Sex to Remarks are optional fields

Table 3.2 (Cont). Sub-header for size (length or weight) frequency data grouped by constant class interval.

Field	Type [†]	Remarks
Time of day	C	This field refers, on a 24h basis, to the time the collection started or ended in hours and minutes. (example: 14:30).
Bottom type	C	This field indicates the bottom structure at the sampling site (example: <i>sandy</i>).
Type of catch	C	This field may be used to describe the data at hand, e.g., commercial or research data.
Effort	N	This field should contain a measure of the effort expended to obtain the sample in question (e.g., no. of hauls, no. of hours, no. of boats, etc.). This field will be used when the data are to be re-expressed as catch per unit of effort (C/f); thus it is important that the unit used is standardized.
Wt. of the sample	N	This field should indicate the weight of the sample, as taken out of the total catch. This implies that the unit used in measuring weight should be consistent for all samples.
Total wt. of species	N	This field should indicate the weight of the species as a subset of the sample collected. It is important to note that the value provided should be less than the WT. OF THE SAMPLE and that the unit used in measuring weights should be the same throughout.
Total wt. of catch	N	This field is for a value representing the total multi-species catch. As in the two previous fields, the unit used in measuring weight should be the same.
a	N	The coefficient "a" of the length-weight relationship ($W=aL^b$). The default value is the one used in the previous sample.
b	N	The coefficient "b" of the length-weight relationship complementing the previous field. The default value is the one used in the previous sample.
Remarks	C	Other information not included in any of the above fields.

Ungrouped length frequencies

This form of length frequencies is characterized by only one column of inputs (Fig. 3.2), the length of individual specimens. As with the previous data form, data of this form cannot be used for analysis in FiSAT II unless regrouped with constant class size. This is provided as an option in FiSAT II when saving this type of file.

The screenshot shows the FiSAT II data entry interface for a file named 'Sample.ulf'. The window is divided into several sections:

- Headers:** Contains fields for Species name ('Upeneus moluccensis'), Other identifier ('Ragay Gulf 1980, Philippines'), Sampling date ('15 Julho 1981'), Unit of measurement ('cm'), and Type of measurement ('Total length (TL)').
- Frequencies:** A table with two columns: 'Obs' (Observation number) and 'Length (cm)'. The data is as follows:

Obs	Length (cm)
1	11.00
2	9.00
3	9.00
4	12.00
5	12.00
6	12.00
7	16.00
8	7.00
9	9.00
10	16.00
11	14.00
12	14.00
- Remarks:** A text area containing the note 'This is only a sample data.'

Fig. 3.2. User interface in encoding ungrouped length frequencies. This data type may also be saved as grouped length frequencies.

The required inputs are:

- Date of sampling (DD/MM/YYYY);
- Measured lengths (0.25 to 999)

When saving to a form with a constant class size, the required inputs are:

- Class interval to use (0.25 to 500);
- Smallest class mid-length to use (0.25 to 500);
- Filename to store the new file (if the filename already exists, the data will be appended to the existing file).

Data grouped by variable class sizes

This form of length frequency data is characterized by explicit lower and upper limits for each class (Fig. 3.3). The mid-length cannot be used in this case. Files in this form can be stored, but they cannot be used within FiSAT II unless regrouped with constant class size.

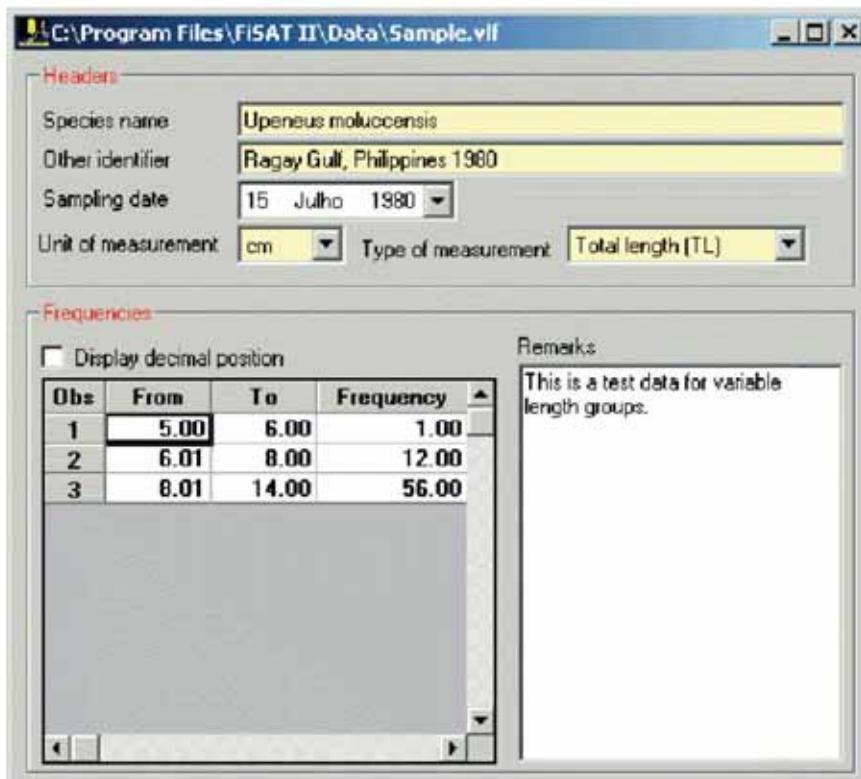


Fig. 3.3. User interface to encode length frequencies with variable length classes.

The required inputs are:

- Date of sampling (DD/MM/YYYY);
- All lower and upper class limits (0.25 to 999);
- Frequencies (0 to 999,999) for each class.

When saving to a form with a constant class size, the required inputs are:

- Class size (0.25 to 500);
- Smallest mid-length to use (0.25 to 500);
- Filename to store the new file (if the filename already exists, the data will be appended to the existing file).

Growth increments

Growth increments are usually referred to as "mark-recapture" data (because mark-recapture studies are a major source of growth increment data). Pairs of mean lengths that are linked in modal progression analysis may be saved in FiSAT II and analysed as growth increments.

This data type is characterized by four columns of entries (Fig. 3.4); the initial date of observation (relative to marking), the initial size, the date of recapture or date of second reading of the same specimen and the size at recapture or second reading. FiSAT II will accept lengths at recapture implying zero or negative growth, but will reject a date pertaining to a second reading that is earlier than the initial date or date of first reading.

Headers

Species name	Large fish		
Other identifier	New major exercise (trawler)		
Unit of length measurement	cm	Type of measurement	Others/Unknown

Frequencies

(NOTE: Yellow column refers to tagging data and unshaded, i.e. white, refers to data at recapture.)

Obs	Day	Month	Year	Length	Day	Month	Year	Remarks
1	15	1	1985	14.90	15	2	1985	Imported from FiSAT (DOS ver.). Data computed from length frequencies; Generated from linking of means
2	15	2	1985	16.08	15	3	1985	
3	15	3	1985	17.08	15	4	1985	
4	15	4	1985	18.34	15	5	1985	
5	15	5	1985	18.70	15	6	1985	
6	15	6	1985	20.13	15	7	1985	
7	15	1	1985	0.04	15	2	1985	
8	15	2	1985	10.29	15	3	1985	
9	15	3	1985	11.37	15	4	1985	
10	15	4	1985	12.32	15	5	1985	
11	15	5	1985	12.92	15	6	1985	
12	15	6	1985	14.58	15	7	1985	

Fig. 3.4. User interface in encoding growth increments. The columns with a dark yellow background refer to tagging data and those with a white background are recaptured data.

The required inputs are:

- Initial date or date of first reading (DD/MM/YYYY);
- Corresponding length (0.25 to 999);
- Final date or date of second reading (DD/MM/YYYY);
- Corresponding length (0.25 to 999).

Length-at-age

The age of a fish may be obtained by directly reading otoliths, scales or other hard parts. Pairs of observations, (age, length) collected for individual fish, form length-at-age data. Series of mean lengths, as identified by modal class progression analysis, can also be treated as length-at-age data, although they actually represent lengths at *relative* age (precluding the direct estimation of t_0).

This type of data is characterized by two column entries; the ages (or relative ages if derived from modal progression analysis) and the corresponding length measurements (Fig. 3.5). The ages can either be in days, months or years and may be any values from 0 to 999. The length values can only be from 0.25 to 999, and the units must either be in cm, mm or inches. FiSAT II provides a third column entry for a weighting factor (WF) for each observation. These weighting factors (default of 1.0; acceptable range 0.01-100) can be used to emphasize or de-emphasize observations during data analysis.

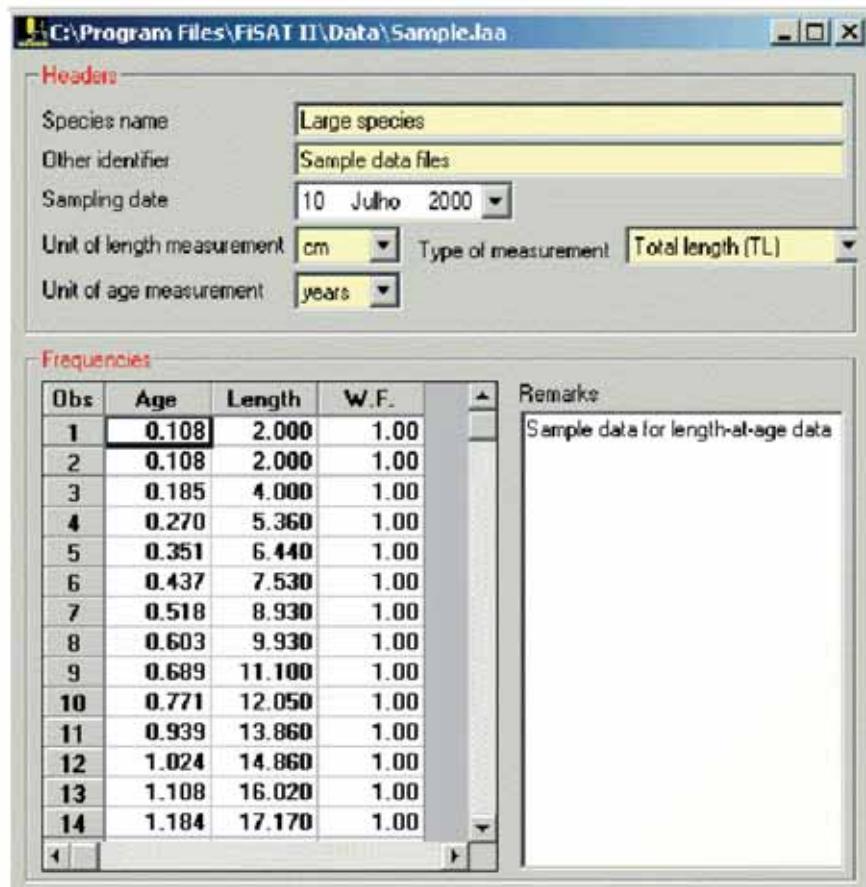


Fig. 3.5. User interface in encoding length-at-age data in FiSAT II.

The required inputs are:

- Ages (0.01 to 999);
- Lengths (0.25 to 999);
- Weighting factors (default is 1.0; 0.01 to 100)

Catch in numbers by age

This data type (also referred to as catch-at-age, Fig. 3.6) can be used only for one routine, i.e. age-structured VPA. The required inputs are similar to length-at-age data. The required input consists of two columns: (i) the ages

(starting with the youngest group) and (ii) the catches in numbers.

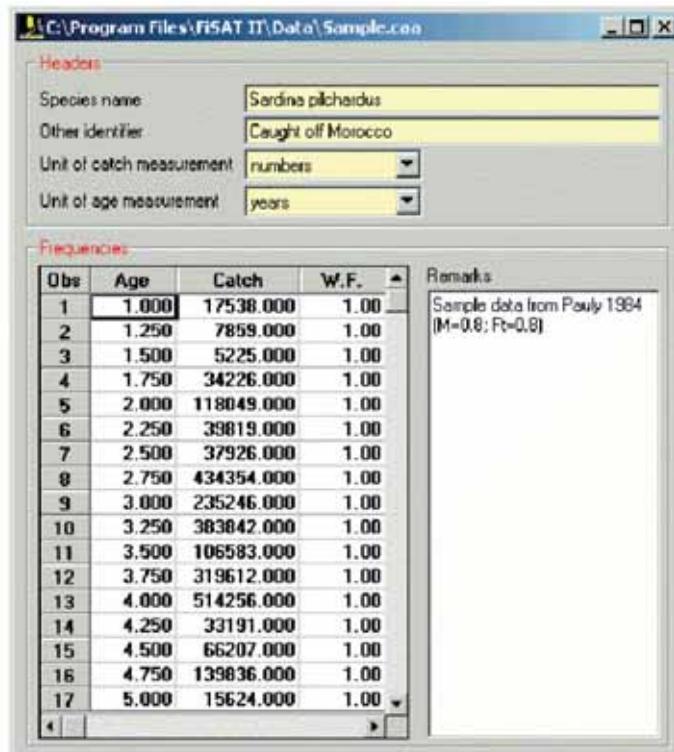


Fig. 3.6. User interface to encode catch-at-age data in FiSAT II.

The required inputs are:

- Age (0.01 to 999);
- Catches, in numbers (1 to 999,999).

Series of mean lengths and standard deviations

In most cases, the data in this file type will be products of the decomposition of normal distributions using, e.g., Bhattacharya's method or Hasselblad's NORMSEP as implemented in FiSAT II, which automatically generate such files (see next chapter). However, creation of such

files from external sources for subsequent analysis by FiSAT II is possible.

The file is characterized by three columns of entries; (i) the date of sampling, (ii) the mean length, and (iii) the standard deviation (Fig. 3.7).

Obs	Day	Month	Year	Mean	s.d.
1	15	1	1985	8.84	1.09000
2	15	1	1985	14.90	1.82000
3	15	1	1985	20.95	2.66000
4	15	2	1985	2.00	0.58000
5	15	2	1985	10.29	1.36000
6	15	2	1985	16.08	1.84000
7	15	3	1985	4.00	0.48000
8	15	3	1985	11.37	1.54000
9	15	3	1985	17.09	1.90000
10	15	4	1985	5.37	0.68000
11	15	4	1985	12.32	1.65000
12	15	4	1985	18.34	1.81000
13	15	5	1985	6.46	0.78000
14	15	5	1985	12.92	1.31000
15	15	5	1985	18.70	2.29000
16	15	6	1985	7.54	0.93000
17	15	6	1985	14.58	1.99000
18	15	6	1985	20.13	2.29000

Fig. 3.7. User interface in encoding means and standard deviation file.

The required inputs are:

- Date of observation (DD/MM/YYYY);
- Mean length (0.25 to 999);
- Standard deviation.

Probabilities of capture

One common cause of bias in length-frequency data is the selectivity of the gear(s) used to obtain the samples. This bias can be in part overcome by correcting size frequency samples with the probabilities of capture stored in a "probabilities of capture file", and generated either through

a selection experiment, or a length-converted catch curve (see next chapter).

This type of data file (Fig. 3.8) is characterized by two columns of entries; (i) the class mid-length (0.25 to 999) and (ii) the probability of capture which may range from 0 to a maximum of 1. Entries less than 0.001 will be registered as 0. Note that a file in which all values are equal to 0 or 1 will not be stored.

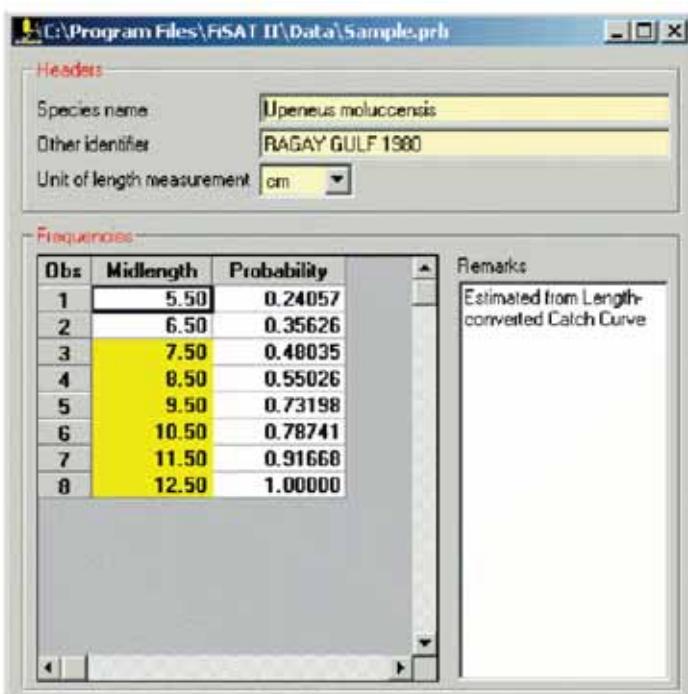


Fig. 3.8. User interface in encoding probabilities in FiSAT II.

The required inputs are:

- Smallest class mid-length (0.25 to 500);
- Class size (0.25 to 500);
- Probability of capture (default is 1.0; 0 to 1.0)

Table of values

The commercial (e.g., market price) or biological value (e.g., in terms of fecundity) of fish often varies with size. This can be accommodated, for various assessments by using size-specific values as multipliers. In FISAT II, the values are used as a multiplier in the Thompson and Bell Yield-Stock Prediction routine.

This data type (Fig. 3.9) is characterized by two columns of data: (i) Class midlength, and (ii) value for each group.

Obs	Midlength	Value
1	5.50	1.00
2	6.50	1.00
3	7.50	1.00
4	8.50	1.00
5	9.50	1.00
6	10.50	1.00
7	11.50	1.00
8	12.50	1.00
9	13.50	1.00
10	14.50	1.00
11	15.50	2.00
12	16.50	2.00
13	17.50	2.00
14	18.50	2.00

Fig. 3.9. User interface for encoding table of values.

The required inputs are:

- Date of sampling (DD/MM/YYYY);
- Lower and upper class limits (0.25 to 999);
- Value (0 to 999) for each class.

Two-column regression files

Linear regression analysis is an important statistical tool for fish stock assessment. One common application is to derive the relationship between the (log-transformed) lengths and weights of fish.

This data type (Fig. 3.10) is characterized by two columns of entries, the X-values and the Y-values, both of which may range from 0 to 999. In FiSAT II, only the regression analysis routine can use this type of data. It is also the only routine whose headers are used only for labelling graphs.

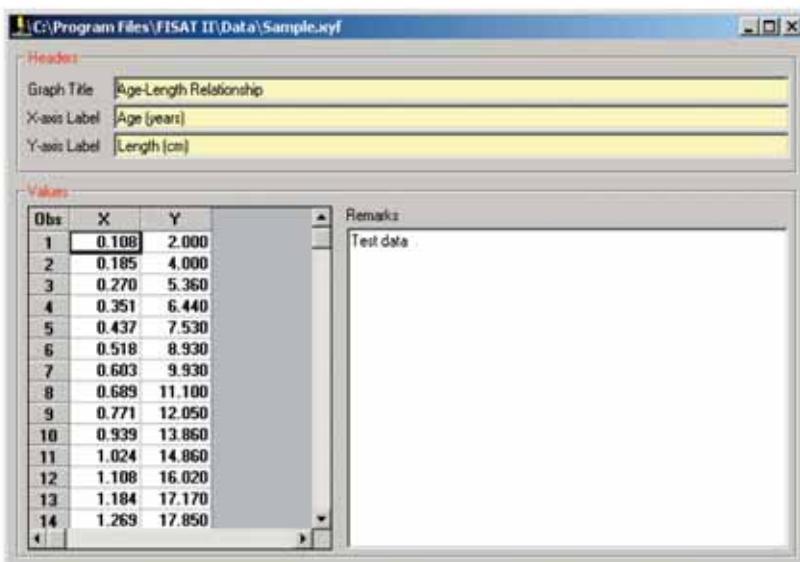


Fig. 3.10. User interface for encoding two-column data for regression analysis.

The required inputs are:

- X-axis and Y-axis labels;
- X-values (0 to 999);
- Y-values (0 to 999).

Species table

The Thompson and Bell Yield-Stock Prediction routine requires population parameters and importantly, an array of fishing mortality (F). In FiSAT II, these data are results when executing length-structured VPA. To simplify the procedures in creating a scenario for the Thompson and Bell Yield-Stock Prediction routine, these results are saved as a Species Table (Fig. 3.11 and Fig. 3.12). Note, however, that these data may also be manually encoded and/or edited by the user.

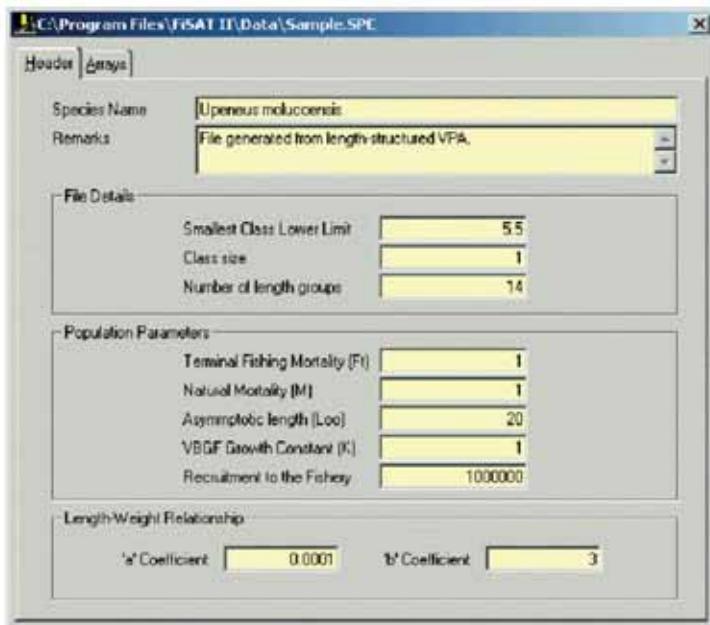
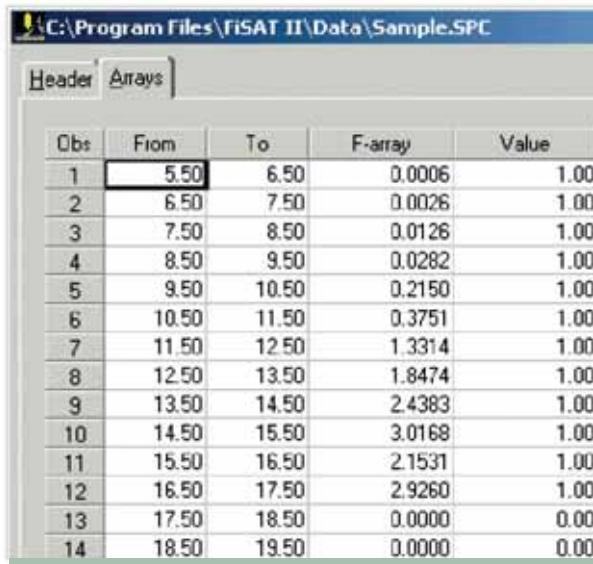


Fig. 3.11. User interface for encoding population parameters when creating/editing Species Tables.



Obs	From	To	F-array	Value
1	5.50	6.50	0.0006	1.00
2	6.50	7.50	0.0026	1.00
3	7.50	8.50	0.0126	1.00
4	8.50	9.50	0.0282	1.00
5	9.50	10.50	0.2150	1.00
6	10.50	11.50	0.3751	1.00
7	11.50	12.50	1.3314	1.00
8	12.50	13.50	1.8474	1.00
9	13.50	14.50	2.4383	1.00
10	14.50	15.50	3.0168	1.00
11	15.50	16.50	2.1531	1.00
12	16.50	17.50	2.9260	1.00
13	17.50	18.50	0.0000	0.00
14	18.50	19.50	0.0000	0.00

Fig. 3.12. Portion of the user interface (second tab as shown in Fig. 3.11) to record F-array and table of values.

The required inputs are:

- Population parameters (class details, F_t , M , L_∞ , K , number of recruits, and coefficients of the length-weight relationship [a , b]);
- F-array;
- Table of values.

Fleet table

In simulating a multi-fleet scenario when executing the Thompson and Bell Yield Stock routine, a fleet indicator file is required to differentiate the catches recorded by various fleets. In FiSAT II, a Fleet Table can be created for the purpose (Fig. 3.13).

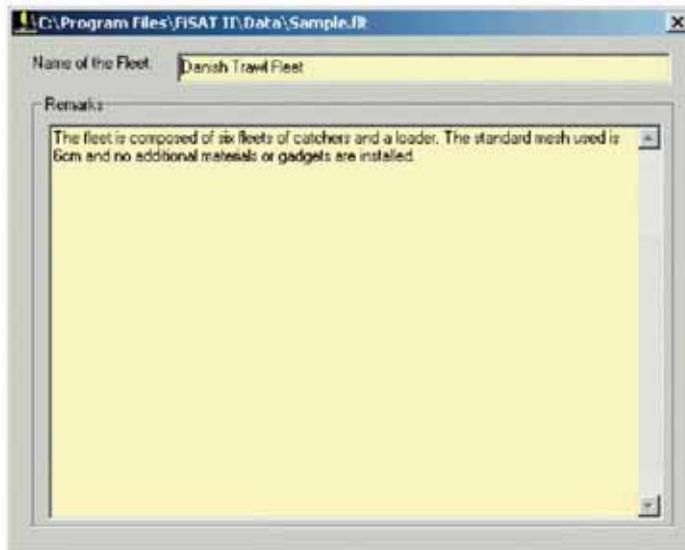


Fig. 3.13. User interface to encode the descriptive parameters defining a fleet.

The required inputs are:

- Fleet name;
- Fleet description.

Multispecies/multifleet scenario

The file that sets the relationship between the exploited species and fishing fleet is the scenario file (Fig. 3.14 and Fig. 3.15). To compute the contribution of each fleet to the species' F-array, catches by fleet are required. To allow the simulation on the change of mesh sizes, selection parameters of each fleet to a specific species also have to be encoded.

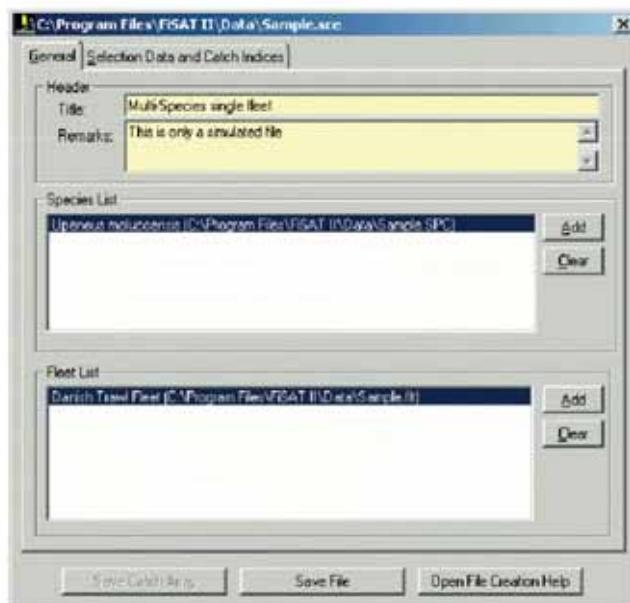


Fig. 3.14. Header records of the Multi-species/multi-fleet scenario file identifying the list of species and fleets to be included in the simulation.

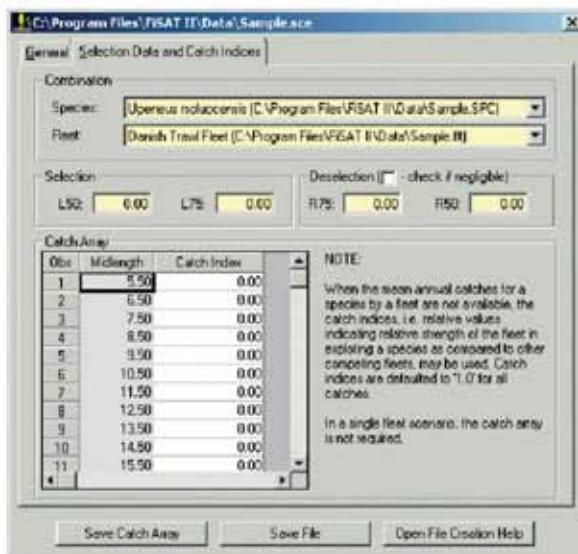


Fig. 3.15. Tab to encode the catches and selection parameters for a species by fleet.