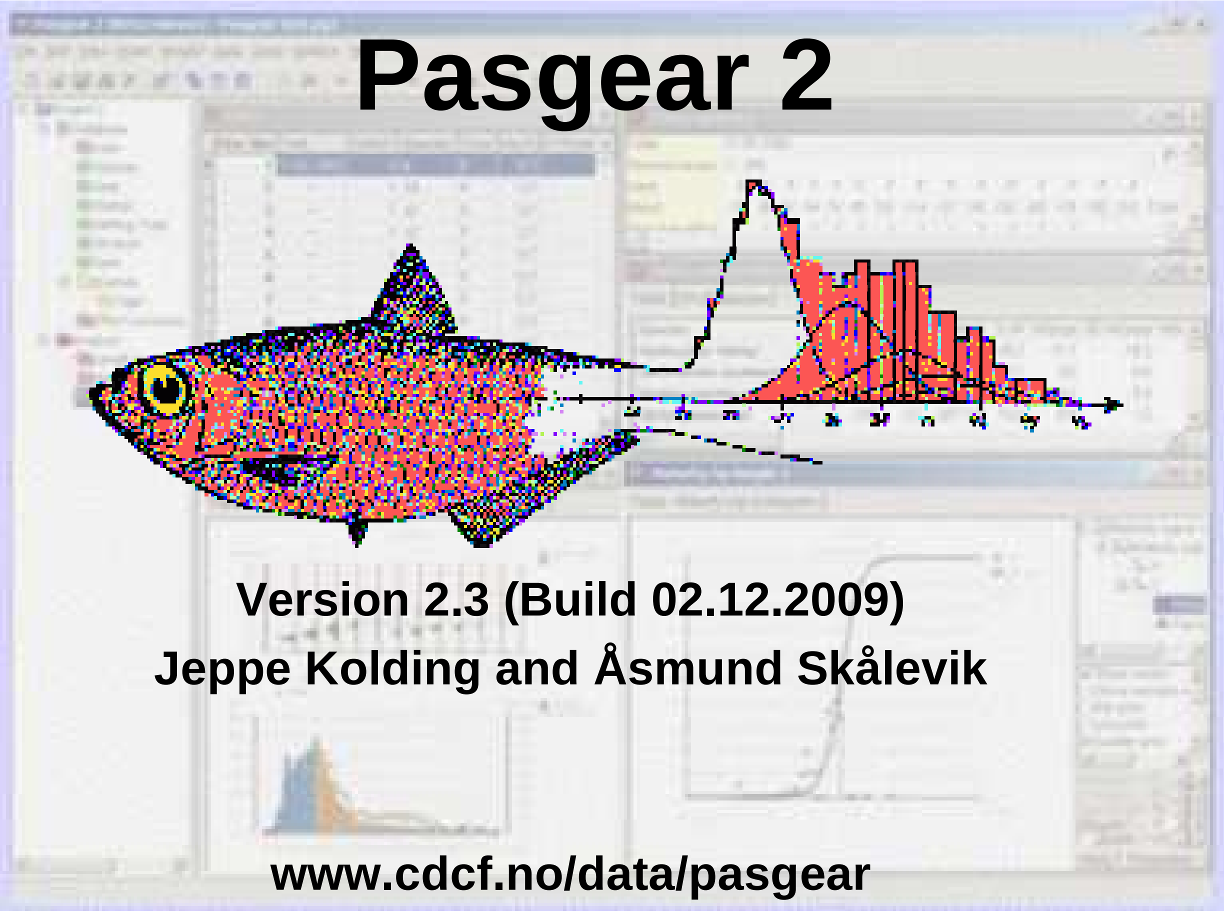
Introduction and Manual to



Pasgear 2

Version 2.3 (Build 02.12.2009)

Jeppe Kolding and Åsmund Skålevik

www.cdcf.no/data/pasgear

Jeppe Kolding

Department of Biology
University of Bergen
N-5020 Bergen
Norway

December 2009

Background and acknowledgements

'PASGEAR' was initially developed in Turbo Pascal for analyzing a series of experimental multi-mesh fishing data from Lake Turkana (Kolding 1989), and was, at that stage, never intended for anything else. In 1991, however, I (Jeppe Kolding) was assigned to computerize, clean and analyze, together with Lawrence Karenge✞, the huge amount of experimental gillnet data collected by Lake Kariba Fisheries Research Institute in Zimbabwe, and the old 'Turkana' programs were dusted of and adapted to the Lake Kariba data set. During this process various modules for facilitating the data entering and cleaning procedures were developed and added, and the idea of developing a more general package was initiated.

Pasgear 1 (DOS) was released in various versions until 2003 by Jeppe Kolding. In February 2003 Åsmund Skålevik was assigned to help redesign and convert the old DOS program into Pasgear II on the Windows Platform. The result, so far, is what you have here.

I am grateful to late Lawrence Karenge for patiently discovering bug after bug and proposing valuable suggestions for improvements. Thanks also to various other users while developing Pasgear 1, particularly Salih El-Thair, Henne Ticheler, and Niklas Mattson, who have discovered more bugs and/or given useful comments and proposals. René Holst, ConStat Denmark, has kindly provided a maximum likelihood optimizer and implemented the gear selectivity functions.

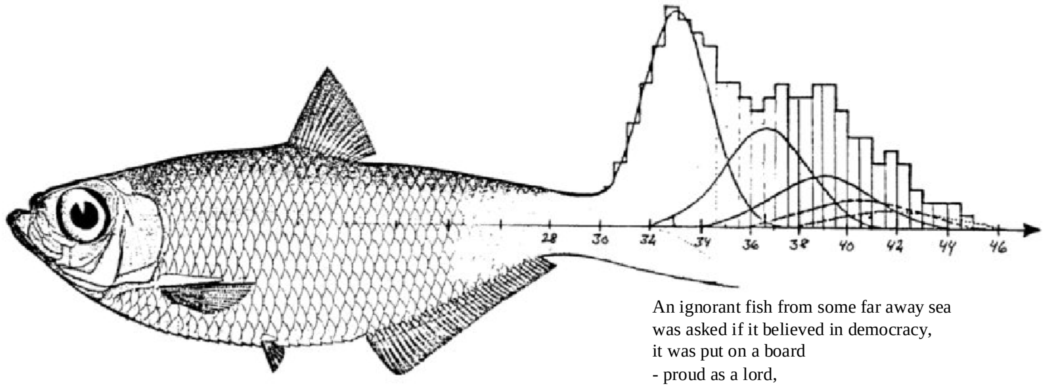
Lastly, a great thanks to Åsmund, who through his impressive programming skills have turned ‘Pasgear’ into a flexible relational database and a tool with almost unlimited possibilities for the user

The development of Pasgear 2 has been kindly been supported by:

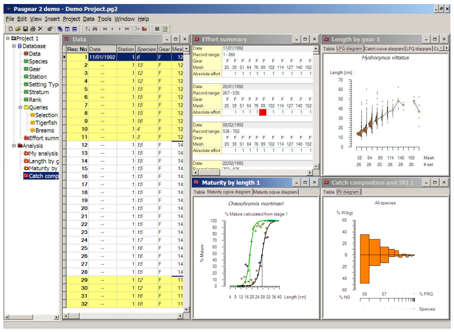
- The Department of Biology, University of Bergen, Norway.
- The European Commission. INCO-DEV programme KNOWFISH
- Ministry of Fisheries and Marine Resources, Namibia
- The Norwegian Research Council
- WorldFish Center
- The Institute of Marine Research (IMR), Bergen, Norway

Table of contents

Overview



An ignorant fish from some far away sea
was asked if it believed in democracy,
it was put on a board
- proud as a lord,
and got lost somewhere in a length-frequency



Rec No	Date	Station	Species	Gear	Mes
1	11/01/1992	1	6	F	12
2	--	1	13	F	12
3	--	1	12	F	12
4	--	1	12	F	12
5	--	1	15	F	12
6	--	1	15	F	12
7	--	1	15	F	12
8	--	1	15	F	12
9	--	1	16	F	12
10	--	1	4	F	12
11	--	1	3	F	12
12	--	1	15	F	14
13	--	1	15	F	14
14	--	1	15	F	14
15	--	1	15	F	14
16	--	1	15	F	14
17	--	1	15	F	14
18	--	1	15	F	14
19	--	1	15	F	14
20	--	1	15	F	14
21	--	1	15	F	14
22	--	1	15	F	14
23	--	1	15	F	14
24	--	1	17	F	14
25	--	1	16	F	14
26	--	1	16	F	14
27	--	1	15	F	14
28	--	1	15	F	14
29	--	1	12	F	11
30	--	1	12	F	11
31	--	1	16	F	11
32	--	1	16	F	11

Date:	11/01/1992										
Record range:	1 - 266										
Gear	F	F	F	F	F	F	F	F	F	F	F
Mesh	25	38	51	64	76	89	102	114	127	140	152
Absolute effort	1	1	1	1	1	1	1	1	1	1	1
Date:	25/01/1992										
Record range:	267 - 535										
Gear	F	F	F	F	F	F	F	F	F	F	F
Mesh	25	38	51	64	76	89	102	114	127	140	152
Absolute effort	1	1	1	1	?	1	1	1	1	1	
Date:	08/02/1992										
Record range:	536 - 702										
Gear	F	F	F	F	F	F	F	F	F	F	F
Mesh	25	38	51	64	76	89	102	114	127	140	152
Absolute effort	1	1	1	1	1	1	1	1	1	1	
Date:	22/02/1992										

PASGEAR 2 is a customised data base package primarily intended for experimental or artisanal fishery data. It is a tool that neatly and quickly lets you store and analyse fishery data from various survey designs.

What Pasgear 2 does

PASGEAR is a customized data base package primarily intended for experimental or artisanal fishery data. The package is developed to facilitate the entering, storage and analysis of the often huge amounts of experimental fishery data, or catch and effort data, accumulated in the research institutes of many countries. It is particularly useful for test-fishing data from multi-mesh gillnets such as 'Lundgren nets' or the 'Drottningholm method' (e.g. Degerman et al. 1989, Fjälling & Fürst 1991.), but can also be used for many other sampling methods such as lines, trawls, or Underwater Visual Census (UVC), Labrosse et al. (2002).

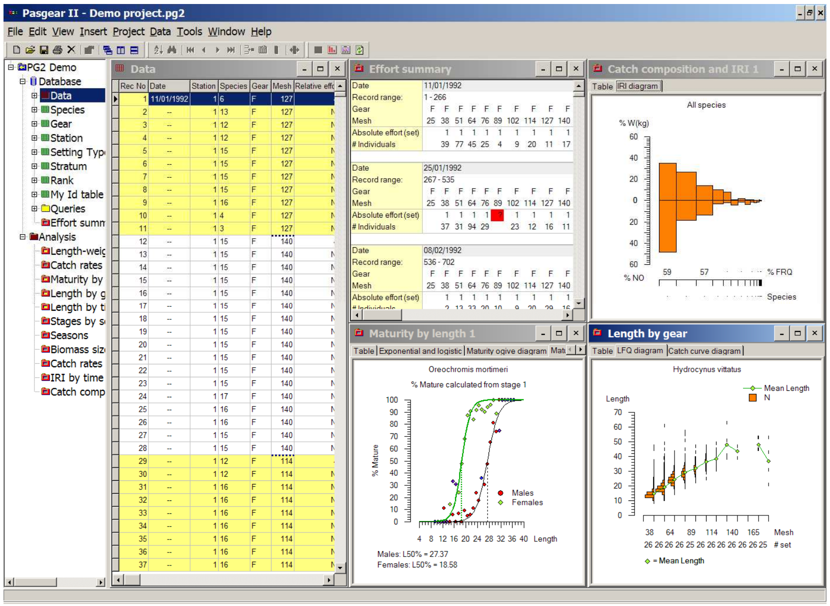
Although PASGEAR has many of the features of an advanced database, it is intended for users with little or no experience in computerized data bases, or with limited access to sophisticated computer equipment. Emphasis has been put on easy data input, editing and manipulation, and on procedures for checking and 'cleaning' data records. With a clean, complete, and accurate data base you are ensured that the calculated results are reliable, at least from a computational point of view. All data can be explored, tabulated, analysed and displayed graphically in the analysis section by query selecting, various groupings, and calculations of a wide range of statistics. A number of predefined extraction, condensing and calculation programmes have been incorporated (see Analysis section) which give an easy overview of large data files and serve the need of the most fundamental data-exploration and most common primary 'analyses' of experimental or statistical fisheries data. Among the several special features included are:

- Automatic estimation of weights from length-weight relationships.
- Standardized (weighed) calculation of CPUE with confidence limits.
- Calculation of different types of confidence limits such as arithmetic, Pennington estimator, and bootstrap.
- Non-linear maximum likelihood estimation of gillnet, hook and trap selectivity probabilities
- Gear selectivity corrected length frequencies and catch curves
- Non-linear least squares estimation of maturity ogives and size at 50% maturity
- Raising of mixed gear catches (CPUE by length groups) for cohort analysis and T&B.

The data-base, or parts of it (any combination of selected records and fields), can be converted into ASCII format for export to spreadsheets, other data bases or statistical packages. Similarly, data can be easily imported from these through the import wizard. All extracts and tables can be copied and pasted in eg. Excel or Word. Length frequencies by time intervals, or mesh sizes, are exportable in FiSAT format for further analysis in ELEFAN (Gayanilo et al. 1989) or FiSAT (Gayanilo et al. 1996). Also total weighted catches by length groups are exportable in FiSAT format as input to length-based cohort analysis (Sparre and Venema 1998).

Pasgear 2 has a powerful graphical interface which lets you plot variables and statistics in almost any combination, as well as fitting a number of different trend curves. The fitting of trends can be done both analytically (linear least square regression method) or by a non-linear iterative numerical solution.

Multiple document interface (MDI)



Rec No	Date	Station	Species	Gear	Mesh
1	11/01/1992	1	6	F	127
2	--	1	13	F	127
3	--	1	12	F	127
4	--	1	12	F	127
5	--	1	15	F	127
6	--	1	15	F	127
7	--	1	15	F	127
8	--	1	15	F	127
9	--	1	16	F	127
10	--	1	4	F	127
11	--	1	3	F	127
12	--	1	15	F	140
13	--	1	15	F	140
14	--	1	15	F	140
15	--	1	15	F	140
16	--	1	15	F	140
17	--	1	15	F	140
18	--	1	15	F	140
19	--	1	15	F	140
20	--	1	15	F	140
21	--	1	15	F	140
22	--	1	15	F	140
23	--	1	15	F	140
24	--	1	17	F	140
25	--	1	16	F	140
26	--	1	16	F	140
27	--	1	15	F	140
28	--	1	15	F	140
29	--	1	12	F	114
30	--	1	12	F	114
31	--	1	16	F	114
32	--	1	16	F	114
33	--	1	16	F	114
34	--	1	16	F	114
35	--	1	16	F	114
36	--	1	16	F	114
37	--	1	16	F	114

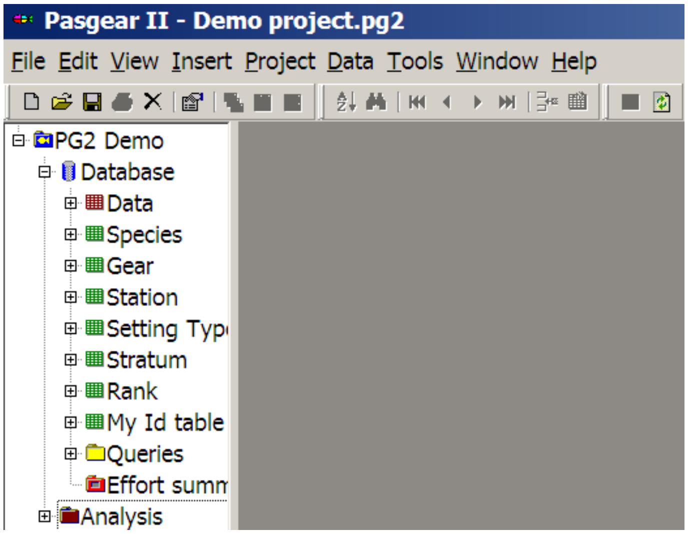
Date	11/01/1992									
Record range:	1 - 266									
Gear	F	F	F	F	F	F	F	F	F	F
Mesh	25	38	51	64	76	89	102	114	127	140
Absolute effort (set)		1	1	1	1	1	1	1	1	1
# Individuals		39	77	45	25	4	9	20	11	17
Date	25/01/1992									
Record range:	267 - 535									
Gear	F	F	F	F	F	F	F	F	F	F
Mesh	25	38	51	64	76	89	102	114	127	140
Absolute effort (set)		1	1	1	1	?	1	1	1	1
# Individuals		37	31	94	29		23	12	16	11
Date	08/02/1992									
Record range:	536 - 702									
Gear	F	F	F	F	F	F	F	F	F	F
Mesh	25	38	51	64	76	89	102	114	127	140
Absolute effort (set)		1	1	1	1	1	1	1	1	1
# Individuals		2	13	33	20	10	9	20	29	16

Pasgear 2 has a multiple document interface which means that you can show any number of tables or analyses simultaneously. Data files, reference tables, queries, and analyses are combined into a ‘Project’. PasGear 2 has a project tree which lets you navigate easily among the various tables, queries and analyses, as well as accessing the properties of these components (objects).

The organization of PASGEAR

Pasgear 2 is both a database and a tool for data exploration and analysis. A “project” is a file (*.pg2) that:

1. Links the database files (tables) to an unlimited set of specified analyses
2. Describes and defines the contents and units of the database and tables
3. Stores the layout of the defined queries, expressions, analyses, charts etc, that connects with the database



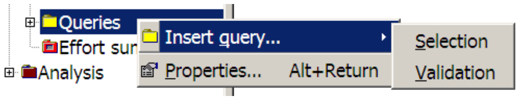
Database

The database is a set of files that stores the data records in binary format (*.bds). These data can be viewed in tables (see Database/Tables).

The main file is Data which stores all the primary information in records (see the data structure in PASGEAR). Some of this information is stored as codes (so-called foreign keys), e.g. species, gear, station etc in order to save space. Therefore there is a number of additional reference or Id-files that link these codes (by primary key fields) to further information such as e.g. the generic name of the key field code. When the data is tabulated in the Analysis section, all key field codes will be translated into corresponding names from the Id tables.

Note that all entering and editing in the database files (*.bds) is done directly on the disk file (not in memory) so that any changes will be directly and permanently written to the file. All changes on the project file (*.pg2), however, will only be saved on user request (either by pressing save, or press [yes] to save when closing the project.

Queries



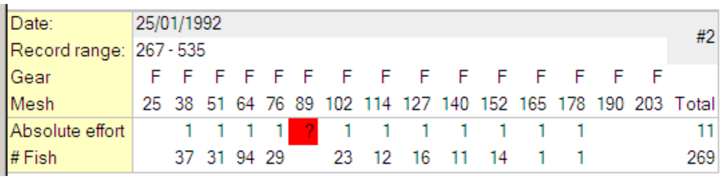
There are two types of queries:

- Selection
- Validation

Selection queries (see Database/Queries) (or selection filters) can be added, defined and applied to any database tables and selection queries on the Data table can also be connected with any analysis item or expression.

Validation queries (see Find and correct records) are used to check whether there are undefined codes in the Data file, i.e. there are codes which have not been defined with key field codes in the corresponding support tables. Validation queries can also be used to find and correct records where the observed length is deviating from the expected length of a fish (calculated from the length-weight relationship) by a defined limit (see Check length/weight and Length-weight relationship).

Effort Summary



Date:	25/01/1992															#2
Record range:	267 - 535															
Gear	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	
Mesh	25	38	51	64	76	89	102	114	127	140	152	165	178	190	203	Total
Absolute effort		1	1	1	1	🟥?	1	1	1	1	1	1	1			11
# Fish		37	31	94	29		23	12	16	11	14	1	1			269

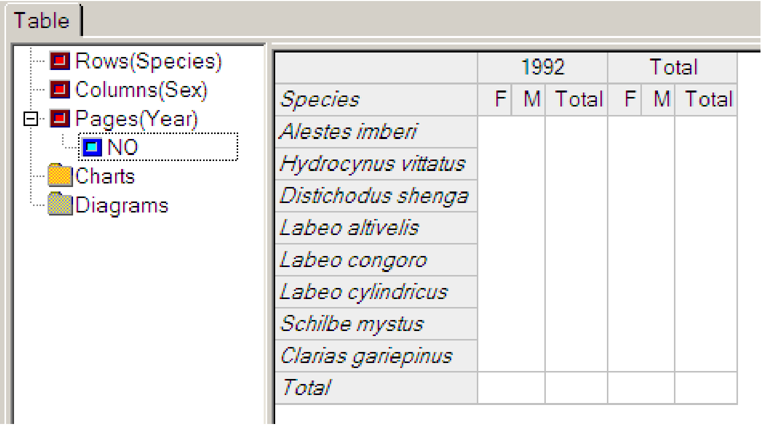
Effort Summary (see Database/Effort summary) is a module that quickly gives you an overview of the data by various date intervals, and specifically lets you check the chronology and the number of primary samples units (see Effort or primary sample unit). It also has an option for inserting ‘empty’ gear settings (see empty settings) into the Data file if these have not been recorded. An empty setting is a gear setting which did not catch any fish. This module should always be run after creating the database and prior to run the analyses in order to secure that the effort and number of primary sample units are correctly counted.

Analysis

The Analysis (see Analysis) part of Pasgear consists of two parts:

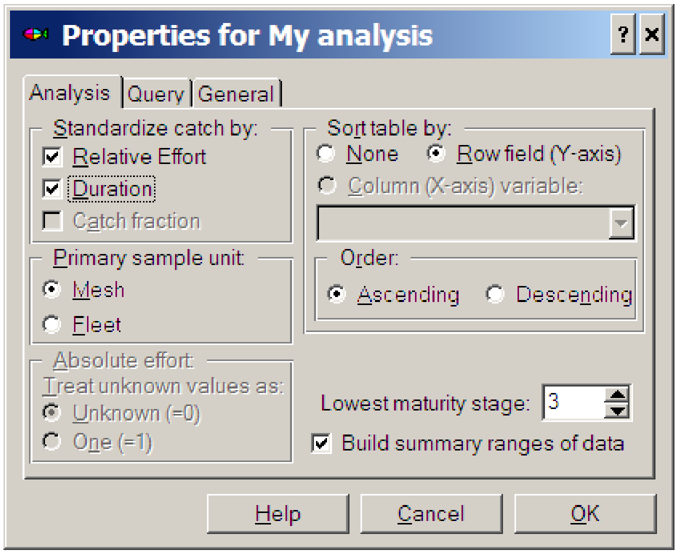
- General analysis
- Predefined analyses

An analysis is generally built as a grouping (condensing) of data fields in one to three dimensions (rows, columns, and pages) with a number of variables/statistics (see Analysis/variables) associated with the groups:



	1992			Total		
Species	F	M	Total	F	M	Total
Alestes imberi						
Hydrocynus vittatus						
Distichodus shenga						
Labeo altivelis						
Labeo congoro						
Labeo cylindricus						
Schilbe mystus						
Clarias gariepinus						
Total						

The outputs off all analyses can be standardized to a given standard effort (see Effort mode in properties for tables) and the result table can be sorted in various ways. Results can also be shown graphically by using the vast range of opportunities in the charts and diagrams objects (see Diagram and chart)



Properties for My analysis

Analysis | Query | General

Standardize catch by:
☑ Relative Effort
☑ Duration
☐ Catch fraction

Primary sample unit:
● Mesh
○ Fleet

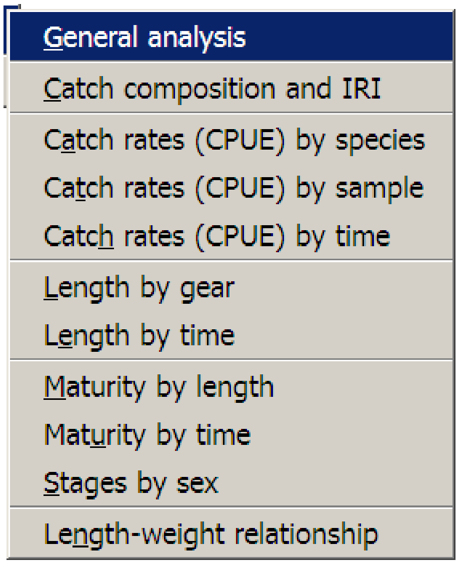
Absolute effort:
Treat unknown values as:
● Unknown (=0)
○ One (=1)

Sort table by:
○ None ● Row field (Y-axis)
○ Column (X-axis) variable:

Order:
● Ascending ○ Descending

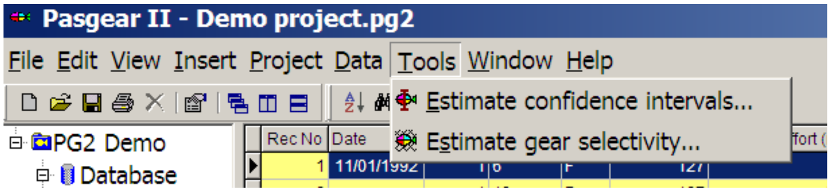
Lowest maturity stage: 3
☑ Build summary ranges of data

Pasgear 2 contains a number of predefined analyses (macros) that can be added (inserted) in a project.



Each of the 10 predefined analyses are special cases of a general analysis, but have been developed to serve the need for the most often used calculations. They all also have predefined diagrams for showing the results. Both the predefined analyses and their diagrams can all be edited or changed nearly unlimited according to the users' wishes. The predefined analyses are all described in more details under Analysis.

Tools



There are two stand alone tools associated with Pasgear 2. These programs can be executed independently from Pasgear 2 and are using their own file system, but can be directly accessed from Pasgear itself. The two tools are:

- Estimation of confidence intervals (conflim.exe)
- Estimation of gear selectivity (gearsel.exe)

The confidence interval tools contains procedures for getting confidence intervals from data that are distributed both normally and highly skewed, such as Pennington's estimator based on the Delta distribution (Pennington, 1983, 1996; Conquest *et al.*, 1996) and bootstrap (Efron & Tibshirani 1986, 1993).

The gear selectivity tool is for indirect estimation of gillnet, hook, and trap selectivity from comparative data of observed catch frequencies across a series of mesh or hook sizes. The general statistical model (SELECT) is described in Millar (1992), and the specific application on gillnets and hooks is described in Millar & Holst (1997) and Millar and Fryer (1999).

The data structure in PASGEAR

PASGEAR is basically a flat data base (one type of records) which can be used to store and analyze fisheries data on various levels, from individual fish with biological measurements to aggregated catch (just numbers and/or weights) and effort data. Each record (Table 1) is consisting of 3 parts:

1. the physical primary sample fields (date, station, gear, gear size or relative effort, duration, set- or sample type, depth etc.),
2. the biological secondary sample fields (species, length, weight, sex, gonads), and
3. the Rank field, which PASGEAR uses for internal data control and validation.

Data can be either punched directly (see Get started/Enter data) or imported from Pasgear 1 (DOS) other sources in text or clipboard format (see Get started/import from Pasgear 1 or import from text)

Field types and codes in a PASGEAR record:

Table 1. The entry fields of one standard (default) record in PASGEAR. The type indicates whether the field is physical, biological or ‘free’. For further details see Database tables/data table/Fields

Entry field	Type	Comments
RECORD NO	NA	Number in file (automatically set)
DATE	Phy	The format depends on the default format chosen for your PC
STATION	Free	Station or location entered as a code (integer) *free field*
SPECIES	Bio	Name of species (or group of species) entered as a code (7 characters, letters or number)
GEAR CODE	Phy	The code of fishing gear used (two letters)
MESH CODE	Phy	The mesh/hook/trap size or code used (integer)
RELATIVE EFFORT	Phy	The relative sample effort or the relative gear effort (e.g. length, area or other unit)
DURATION	Phy	Time duration of effort in hours or minutes
SET TYPE	Free	What kind of setting/sample type or other separator (integer) *free field*
NUMBER	Bio	Number of individuals the record represents (integer)
LENGTH	Bio	Length of the individual (in mm or cm)
WEIGHT	Bio	Weight of the number (defined in the number field) of individuals (in grams or Kg)
SEX	Bio	Male, female or not determined (M, F or X)
GONADS	Bio	Gonad stage (range defined in project properties (integer)
STRATUM	Free	e.g. bottom depth or any other separator (integer) *free field*
RANK		Data validity codes (integer, codes different from 0 has special meanings)

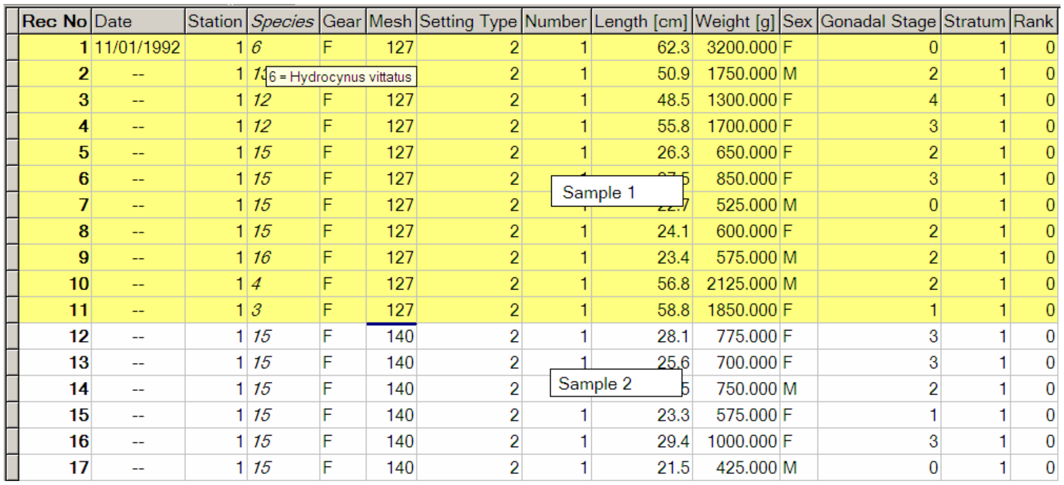
Table 1 shows the standard (default) fields in a Pasgear record. Some of these fields are fixed and you have limited options for editing their properties. Two fields (Date and ‘mesh code’ indicated by yellow) are obligatory for a record to be considered valid. Some default fields can be deleted (indicated by grey background) and others (indicated by red italics) are ‘free’ which means that they can be changed from physical fields (red/blue) to biological fields (green) - or vice versa - according to the users wish. Any field, however, or any visible object in Pasgear, can have its field header renamed.

The user can add any number of additional fields (columns) to the data base, either containing values or calculated from other fields (see Database tables/Fields).

Valid records and missing information

You do not need to have information for all fields in a record for using Pasgear 2. For a record to be valid only **2 physical fields** need to have values: Date and Mesh code (indicated by yellow background in Table 1). Other ‘physical’ fields such as ‘Gear’, ‘Station’, ‘Set type’, and ‘Stratum’ can be empty or renamed and used for other purposes, or – for the two latter – even deleted. For the **biological fields**, only the Species code and number field needs to be specified. Species code = 0 by default means no catch (empty setting or haul). Fields with no information entered will be stored with default values for ‘unknown’ or empty (see Database tables/Fields). You can define the number of fields to be displayed (visually shown) in the data table properties.

2-stage sampling design in one record



Rec No	Date	Station	*Species*	Gear	Mesh	Setting Type	Number	Length [cm]	Weight [g]	Sex	Gonadal Stage	Stratum	Rank
1	11/01/1992	1	*6*	F	127	2	1	62.3	3200.000	F	0	1	0
2	--	1	[illegible]	[illegible]	[illegible]	2	1	50.9	1750.000	M	2	1	0
3	--	1	*12*	F	127	2	1	48.5	1300.000	F	4	1	0
4	--	1	*12*	F	127	2	1	55.8	1700.000	F	3	1	0
5	--	1	*15*	F	127	2	1	26.3	650.000	F	2	1	0
6	--	1	*15*	F	127	2	[illegible]	[illegible]	850.000	F	3	1	0
7	--	1	*15*	F	127	2	[illegible]	[illegible]	525.000	M	0	1	0
8	--	1	*15*	F	127	2	1	24.1	600.000	F	2	1	0
9	--	1	*16*	F	127	2	1	23.4	575.000	M	2	1	0
10	--	1	*4*	F	127	2	1	56.8	2125.000	M	2	1	0
11	--	1	*3*	F	127	2	1	58.8	1850.000	F	1	1	0
12	--	1	*15*	F	140	2	1	28.1	775.000	F	3	1	0
13	--	1	*15*	F	140	2	1	25.6	700.000	F	3	1	0
14	--	1	*15*	F	140	2	[illegible]	[illegible]	750.000	M	2	1	0
15	--	1	*15*	F	140	2	1	23.3	575.000	F	1	1	0
16	--	1	*15*	F	140	2	1	29.4	1000.000	F	3	1	0
17	--	1	*15*	F	140	2	1	21.5	425.000	M	0	1	0

Sampling of biological organisms normally consists of two stages or sample levels:

1. The physical capture or sampling stage where you set the gear.
2. The biological recording stage where you count and measure various parameters in the catch.

The capture or ‘physical’ sampling stage defines the **primary sample unit (PSU)** and is characterised by physical attributes (eg. Date, Station, Stratum, Set Type, Gear, Mesh size etc.) that are unique for each primary sample unit. A primary sample unit in fisheries is typically labelled as a ‘set’ or ‘haul’, and often represent a unit of absolute effort (see Effort definition).

The recording of the biological catch defines the **secondary sample unit (SSU)** and consist of biological attributes (typically Species, Number, Length, Weight, Sex, Gonads). The number of biological records within one primary sample will vary with the size of the catch and the aggregated level of the biological recording (see below under Types of **Catch** data).

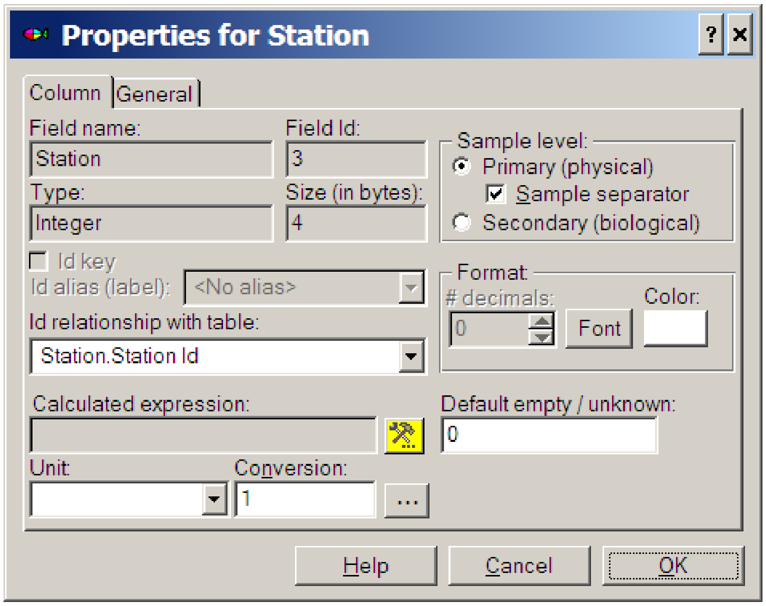
One record in Pasgear consist of both physical (PSU) and biological (SSU) information (indicated by red and green color respectively in Table 1). The minimum of physical information required for a record to be valid is date and mesh (gear), whereas no biological information is required. One record with only physical values but no biological simply means an ‘empty setting’, i.e. the sampling was done but no catch obtained.

When the records are properly sorted chronologically (see sorting data), this unique system enables Pasgear to automatically keep track of the number of primary sample units. This is a great advantage because Pasgear itself keeps track of the number of sample units (PSU) in calculations (e.g. CPUE see Calculating CPUE) irrespective of the number of records within one sample.

Physical data (the primary sampling unit)

PASGEAR will automatically 'count' a new PSU (setting, fishing operation, or sample) each time one of the fixed physical fields (**Date, Gear or Mesh**) change value or optionally also if any of the other physical fields change value (see Database tables). Thus, each time a physical (PSU) field changes its value, a new primary sample is considered regardless of the number of biological records (SSU) within each PSU.

You can define which optional physical fields should be used to demarcate a PSU by checking these as sample separators under column properties. Each time any one of these fields change its numerical value, compared to the previous record, PASGEAR will consider it a new primary sample and add the number of sample units (n) with +1. In other words when just one of these fields changes its value it simply means that it is now a different fishing operation, i.e. a new sample.



Field name: Station
Field Id: 3
Type: Integer
Size (in bytes): 4
Sample level:
● Primary (physical)
☑ Sample separator
○ Secondary (biological)
☐ Id key
Id alias (label): <No alias>
Id relationship with table: Station.Station Id
Format
decimals: 0
Color:
Calculated expression:
Default empty / unknown: 0
Unit:
Conversion: 1

Biological data (the secondary sampling level)

Types of biological data.

The biological data or catch data (secondary sample unit) can be from 3 different levels of information:

a) The *individual* level where one record is one organism (fish) and all biological fields have a value referring to this single organism. In this case the **NUMBER** field takes the value 1 (see Table 2).

b) The *length-frequency* level where the length is a length-group and the weight is the total weight of individuals in the length group and the **NUMBER** field is the number of individuals in the length group. In this case the number field is normally >1 (see Table 2).

c) The *total "catch"*-level where the length is unknown (value = 0) and the weight is the summed weight of the species caught (in one to several settings or a sample), in this case the **NUMBER** field is the number of fish caught (if not known it is set to 0).

Thus, there are 3 possibilities for entering biological information into a record:

Table 2

	Fields				
Level of information	species	length	Weight	number	sex/gonads
a) Individual fish	x	x or 0	x or 0	1	x or 0
b) Length-frequency	x	x	x or 0	x	0
c) Catch in Number/weight	x	0	x or 0	x or 0	0

where x stands for known information and 0 for unknown (missing) information. If species is 0 and the physical fields (i.e. date and gear fields) are entered then such record will count as an 'empty' setting.

PASGEAR will work on any of these levels, and they can even be mixed into the same file and in the same sample (e.g. when various proportions of a catch is sampled for individual measurements, length-frequencies, and maybe a rest groups of only numbers and weight). PASGEAR will keep track of which information is available in the individual calculation programmes.

Effort definition and sample raising modes.

Effort mode

The most commonly used method of estimating the relative abundance of an exploited fish stock is by using the catch per unit effort (CPUE) as an index of abundance. The nominal fishing effort is expressed in for example the number of fishermen, the number of boat-days, the number of gill-net set, the number of hooks set, the number of hauls or pulls made, etc. However, it is important that the so-called catchability coefficient remains constant. Catchability is defined as the relationship between the catch rate (CPUE) and the true population size (B). So the unit is fish caught per fish available per effort unit and per time unit. Catchability is also called gear efficiency (Hillborn and Walters 1992) or sometimes fishing power. For the gear efficiency to be constant it is important that the relative effort of the nominal fishing effort is also constant.

In PASGEAR 'effort' is considered at 2 different *levels* or types and with two different *kinds* in the primary sampling unit.

The two different levels are:

1. The ***absolute*** effort, i.e. the number of gear settings/hauls/transects etc., or sample units in a data series. Each Primary sample unit (setting or sample) can consist of one to several absolute effort units depending on the defined effort mode on the data table.

2. The ***relative*** effort within one primary sample unit. This can vary between each PSU according to the respective gear size (e.g. area of different mesh panels in a fleet), or the relative sampling effort (e.g. fraction of sampled fish of total catch in a primary sampling unit, or number of settings (effort) for the total catch in a sample).

To illustrate, standardized catch per unit effort (*CPUE*) in PASGEAR is calculated as:

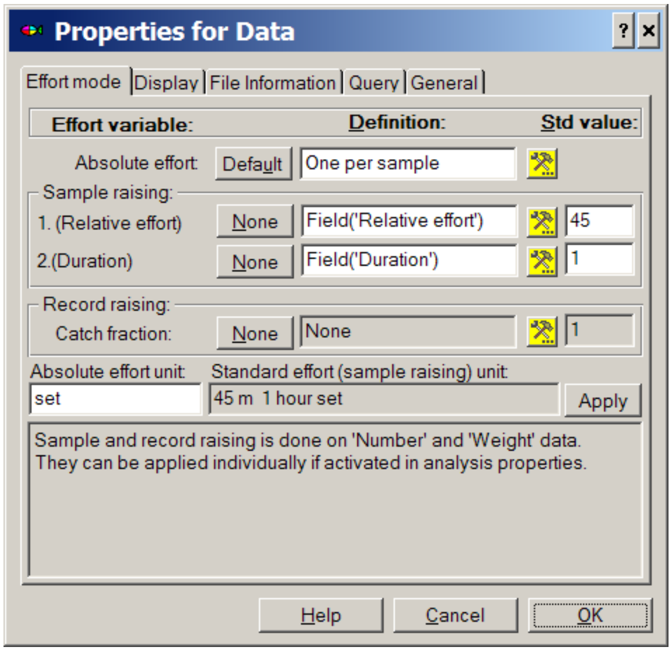
$$CPUE = \frac{1}{y} \sum_{i=1}^{n} W_i \cdot \frac{SU}{U_i} \qquad \text{where}$$

- y is the *absolute* effort and n is number of primary sample units (thus when effort is not a variable then $y = n$),
- W_i is the catch (weight or numbers) of sample unit i
- SU is the standard value of one absolute effort unit (i.e. 100% or *standard effort* e.g. area of a gillnet)
- U_i is the *relative* effort value or ‘weight’ of sample unit i given in e.g. the **RELATIVE EFFORT** field
  (i.e. 50 (percent) if only half the catch was measured, or the actual area of the net used).

PASGEAR can thus easily work on 'sampled' catch and effort, and thus sampled CPUE, even if only a fraction of the catch in one fishing operation (PSU) is actually recorded, or if the recorded catch in a sample is aggregated from several effort units (e.g. nets).

Both absolute (y) and relative (U_i) effort can be stored in PASGEAR, but there are 2 different ‘mode’ or *kind* options for absolute effort within a PSU, depending on whether the number of settings (or absolute effort units) in a primary sample is a variable ($y \neq n$) or not ($y = n$), i.e whether the biological information in a primary sample unit is obtained from one or a variable number of effort units (nets, hauls etc.):

The ‘effort mode’ and standard units are defined on the data table properties:



Effort variable:		Definition:	Std value:
Absolute effort	Default	One per sample	
Sample raising:			
1. (Relative effort)	None	Field('Relative effort')	45
2.(Duration)	None	Field('Duration')	1
Record raising:			
Catch fraction:	None	None	1

Absolute effort unit: set

Standard effort (sample raising) unit: 45 m 1 hour set

Sample and record raising is done on 'Number' and 'Weight' data. They can be applied individually if activated in analysis properties.

The two different kinds of absolute effort mode:

1) The <u>absolute effort</u> is the same as a primary sample unit ($y = n$):

> For most experimental fishing data, the recorded catch is usually separated by each unit of effort (i.e. settings/hauls/mesh size etc.) and each primary sample is from one such effort unit. In this case (which is the 'automatic' **default** effort mode in PASGEAR, see Data properties/effort mode) you do not have to store the number of gears, settings or hauls within a primary sample. Click on 'Default' under absolute effort (definition will then be 1) and PASGEAR will count the effort as the number of primary sampling units.

2) The <u>absolute effort</u> (number of settings or hauls) within a primary sample unit is a variable ($y \neq n$):

> If the catch data consists of pooled data coming from *y* units of effort then the 'automatic' default absolute effort mode is not applicable as '*y*' in the above CPUE equation is a recorded variable that can change with each record or sample (where the number of samples, *n*, are still automatically separated each time any one of the defined physical fields change value). In this case you can use either the **RELATIVE EFFORT** field, or the **DURATION** field, or any other new field, to store the absolute effort variables (and then specify (define) the used location on the property page by giving the column name in the definition field). The number of settings ('*y*' in the above equation) will then be read from the file in the specified column instead of being automatically counted. If you choose the **DURATION** field to store the absolute effort variable then you can still use the **RELATIVE EFFORT** field to give relative weight to the sample as described below under sample raising. Alternatively, if you use the **RELATIVE EFFORT** field to store the absolute effort variable then you can

use the **DURATION** field to standardize (raise) the calculated CPUE by time units, or any other unit (see sample raising).

Sample raising

Sample raising means that the raising of calculated values (number and weight) will take place on each primary sample unit. Thus the whole PSU will be raised according to the ratio of the SU/U_i for the U_i value given in the first record of a new PSU. U_i values in subsequent records within the same PSU are recorded but may be considered redundant. This option should be used if the same proportions of the catch is sampled and recorded within one PSU or the whole sample should be raised according to a given standard value. E.g. the catch was sampled with nets of 45 m length (relative effort = 45) but you want your results standardised to say 100 m net lengths.

The two standard ‘effort’ fields in a record (relative effort, and duration) can be used depending on the setup of the Effort mode. If any one of these is not used for storing absolute effort (i.e. absolute effort is not ‘Default’) then they can also be used for sample raising. However, you can define any new field for storing absolute and relative effort and duration.

The *relative* effort field (e.g. gear length/area) or relative sample effort (fraction of catch recorded of total) is normally given in the **RELATIVE EFFORT** field (Table 1). This field, together with the given **standard** effort unit (defined under ‘effort mode’) is used to give relative 'weight' to the catch within a primary sample unit. If the **standard** effort unit for example is 100 (settings, nets, hauls, hooks, fishermen, pulls, boats, etc...) and the **RELATIVE EFFORT** field for these records is given the value 50, it means that the recorded catch within this setting or sample will be given double weight, or in other words, only half the catch was sampled and recorded. Similarly, if the standard effort unit is set to 100, and the **RELATIVE EFFORT** field is given the value 600 it means that this sample is the combined catch from 6 times the standard effort unit (e.g. 6 settings pooled).

Table 3. Typical relative effort values depending on the survey methodology:

Sampling gear/methodology	**Relative effort**
Gillnets	Area or length of net panels
Lines	Number of hooks
Underwater Visual Census	Area of transect
Trawl	Area swept
Sub-sample	Fraction of total

The *duration* field can be used to store the actual duration of the catch operation, and the corresponding standard value can then be used to raise your catch to standard duration. When both the ‘relative effort’- and ‘duration’ fields are defined then your optional standardization in the analyses can be performed on each level separately or combined (see standardize catch).

Note that the relative effort units, e.g. meter (m) for relative effort expressing net length, and hours or minutes for duration, are defined in the respective column properties/Units. They will appear in the ‘Standard effort (sample raising) unit’ field when pressing [Apply].

If information on the relative effort and/or the duration is not available then click [None] for these under ‘Sample raising’

Record raising

Record raising means that the raising of calculated values will take place at each record. Thus each record will be raised according to the ratio of the SU/U_i for the U_i value given in the record. This option (set in Data table/effort mode) should be used if for example different proportions of the catch is sampled and recorded within one PSU.

Get started

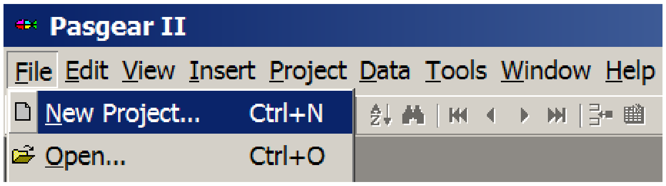
Get started with Pasgear 2

For those not acquainted with Pasgear 1 (DOS), they should first read the Overview to understand the structure of Pasgear 2.

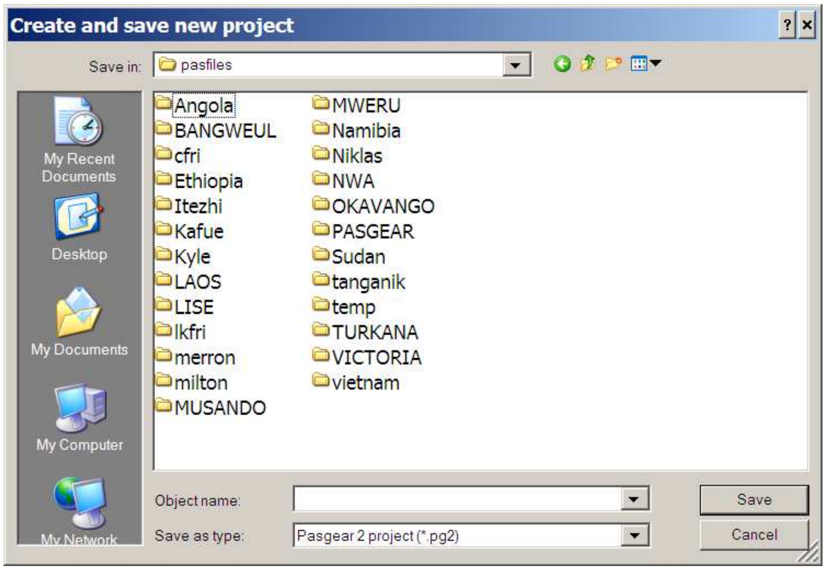
The following points show how to

- Create a new project in Pasgear 2 and migrate from old Pasgear 1 files
- Import data from the clipboard text stream or a text file.
- Punch and edit data directly in Pasgear 2.

Create a new Project



1. Choose **File ►New project** or press Ctrl + N.



Create and save new project

Save in: pasfiles

Save as type: Pasgear 2 project (*.pg2)

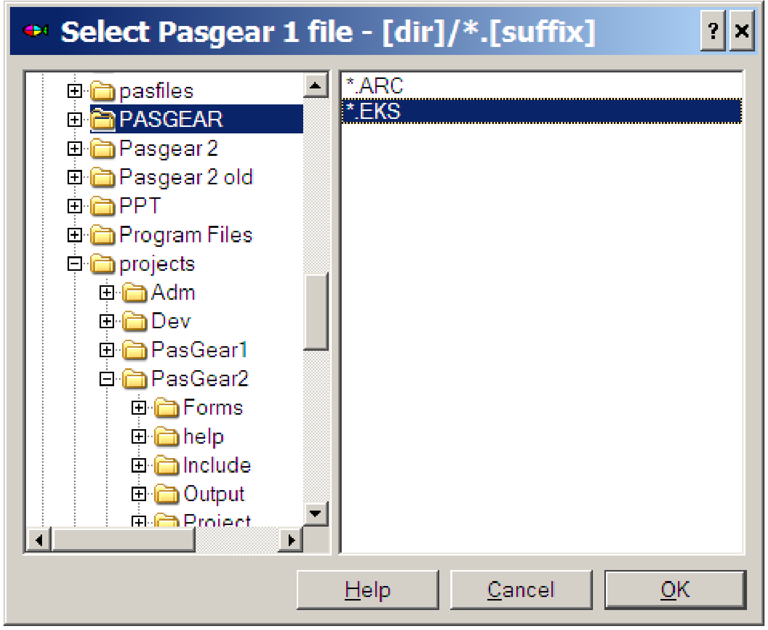
2. Select location, enter Object name and Save.

Depending on the data source there are now 3 options:

- Import data from old Pasgear 1 (DOS) files
- Import data from other sources, e.g. Excel or text files
- Punch data directly into Pasgear 2

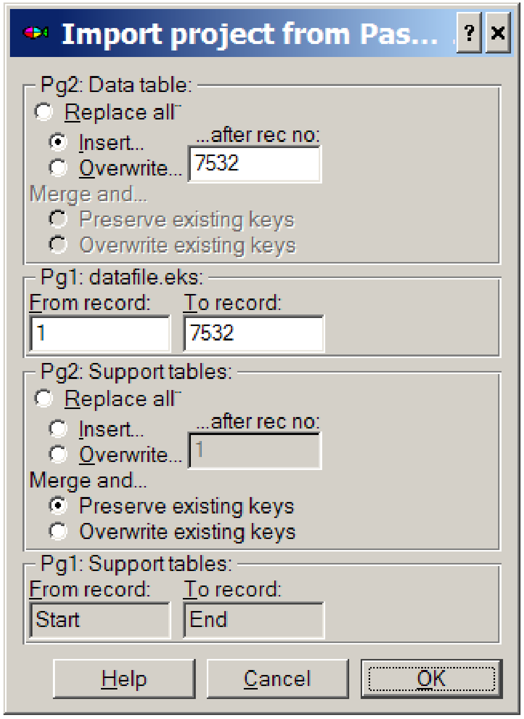
Import from Pasgear 1 (DOS version)

3. Right click the project and choose **Import ► Pasgear 1 (DOS)**.



Select Pasgear 1 file - [dir]/*.[suffix]

4. Choose the location of your Pasgear 1 (DOS) files. Select the *.[extension] and press OK.



Import project from Pas...

Pg2: Data table:
○ Replace all
● Insert... ...after rec no: 7532
○ Overwrite...
Merge and...
○ Preserve existing keys
○ Overwrite existing keys

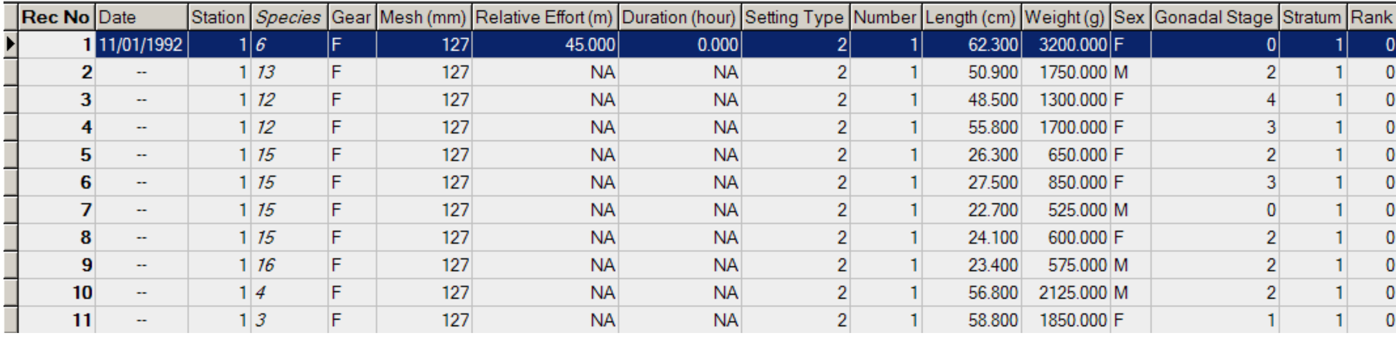
Pg1: datafile.eks:
From record: 1
To record: 7532

Pg2: Support tables:
○ Replace all
○ Insert... ...after rec no: 1
○ Overwrite...
Merge and...
● Preserve existing keys
○ Overwrite existing keys

Pg1: Support tables:
From record: Start
To record: End

5. Choose the import range (From record and To record), or accept defaults (whole range) and press OK.

6. Double click 'Data' to view imported data in the Data table.



Rec No	Date	Station	*Species*	Gear	Mesh (mm)	Relative Effort (m)	Duration (hour)	Setting Type	Number	Length (cm)	Weight (g)	Sex	Gonadal Stage	Stratum	Rank
1	11/01/1992	1	*6*	F	127	45.000	0.000	2	1	62.300	3200.000	F	0	1	0
2	--	1	*13*	F	127	NA	NA	2	1	50.900	1750.000	M	2	1	0
3	--	1	*12*	F	127	NA	NA	2	1	48.500	1300.000	F	4	1	0
4	--	1	*12*	F	127	NA	NA	2	1	55.800	1700.000	F	3	1	0
5	--	1	*15*	F	127	NA	NA	2	1	26.300	650.000	F	2	1	0
6	--	1	*15*	F	127	NA	NA	2	1	27.500	850.000	F	3	1	0
7	--	1	*15*	F	127	NA	NA	2	1	22.700	525.000	M	0	1	0
8	--	1	*15*	F	127	NA	NA	2	1	24.100	600.000	F	2	1	0
9	--	1	*16*	F	127	NA	NA	2	1	23.400	575.000	M	2	1	0
10	--	1	*4*	F	127	NA	NA	2	1	56.800	2125.000	M	2	1	0
11	--	1	*3*	F	127	NA	NA	2	1	58.800	1850.000	F	1	1	0

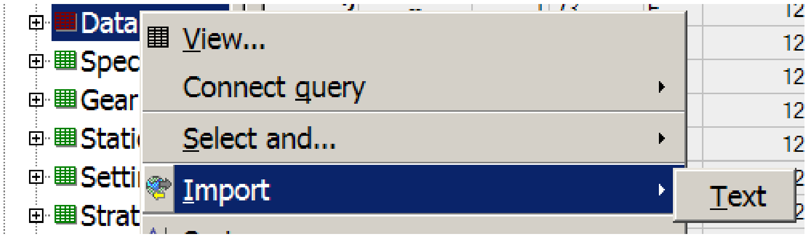
7. Open view Id tables (species, Station, Setting type etc.)

8. Open Project properties (see Project properties) and set all definitions and default settings for the project. This is equivalent to the old CONSTANT.* file in Pasgear 1 (DOS)

9. Open Data properties (see Properties for tables) and set the Effort mode.

10. You are ready to build and run analyses (see Analysis)

Import from Clipboard or text files

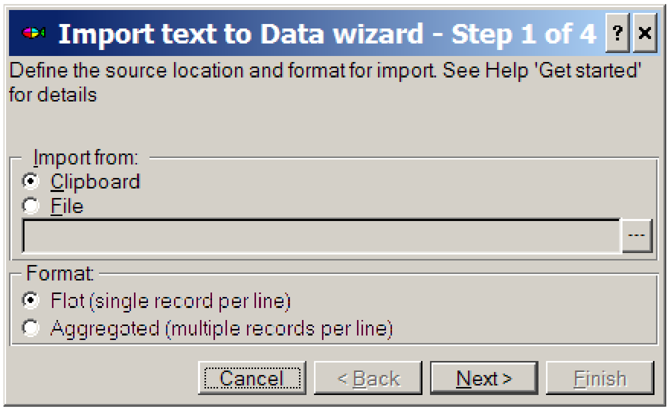


Importing data from external sources (e.g. Excel) must be done on the Data table and/or each Id Table separately.

3. Right click on the table, choose **Import ► Text**.

This will bring you into the Import Wizard with 4 steps:

Step 1.


Import text to Data wizard - Step 1 of 4

Define the source location and format for import. See Help 'Get started' for details

Import from:
● Clipboard
○ File

Format
● Flat (single record per line)
○ Aggregated (multiple records per line)

Define the source origin:

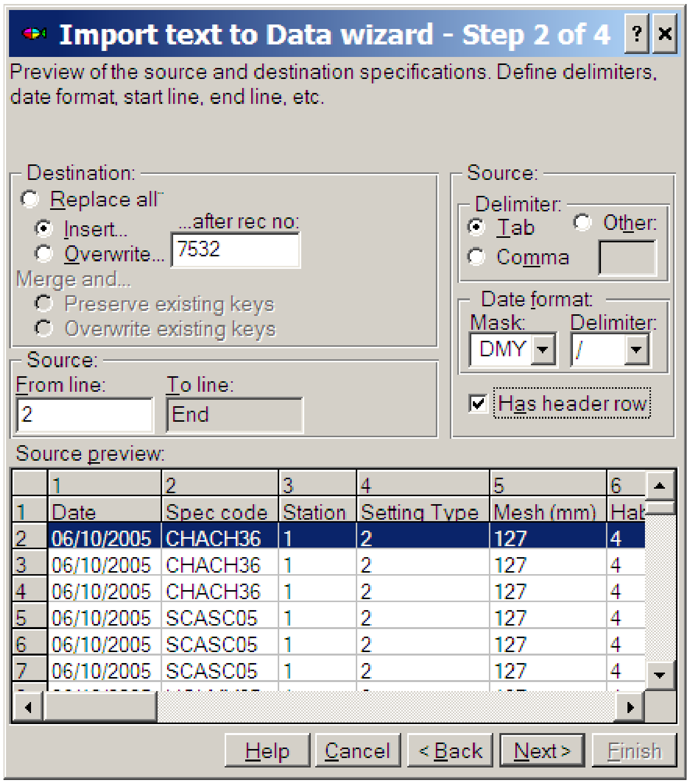
- Clipboard or
- File

If file (must be a text file) then open the file by pressing ...
Define the format: Flat or aggregated.

- A flat file is one record with various fields per line.
- An aggregated file is a matrix with rows and columns, typically a length frequency file with length intervals in rows and e.g. dates or mesh sizes in columns.

Press **Next** for next step

Step 2.



Import text to Data wizard - Step 2 of 4

Preview of the source and destination specifications. Define delimiters, date format, start line, end line, etc.

Destination:
- ○ Replace all
- ● Insert... ...after rec no: 7532
- ○ Overwrite...

Merge and...
- ○ Preserve existing keys
- ○ Overwrite existing keys

Source:
- From line: 2
- To line: End

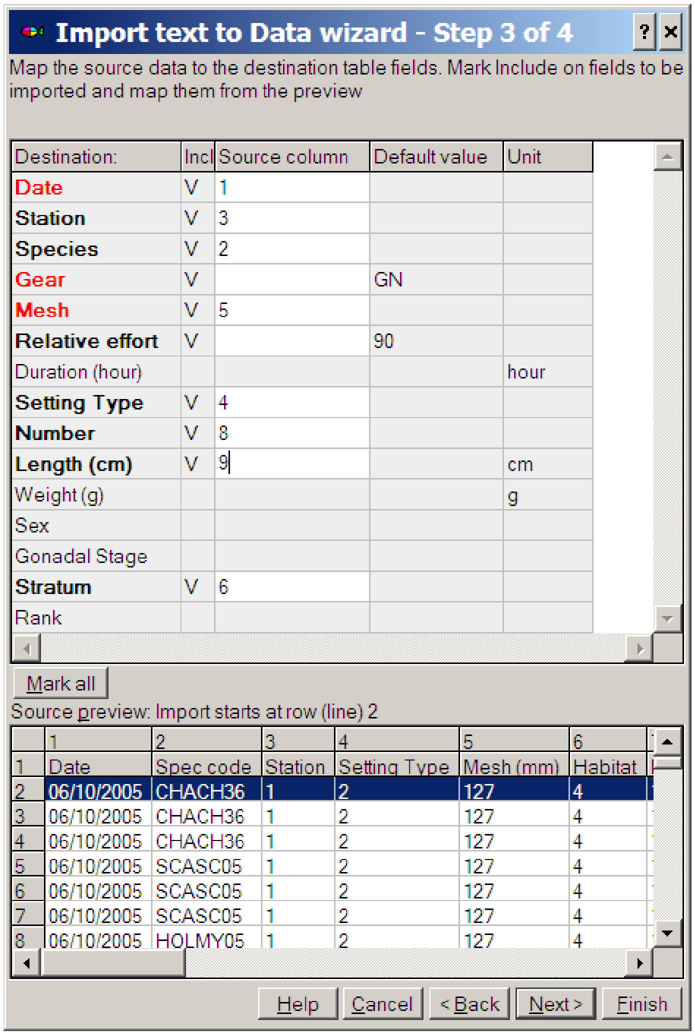
Source:
- Delimiter: ● Tab ○ Comma ○ Other:
- Date format: Mask: DMY Delimiter: /
- ☑ Has header row

Source preview:

	1	2	3	4	5	6
1	Date	Spec code	Station	Setting Type	Mesh (mm)	Ha
2	06/10/2005	CHACH36	1	2	127	4
3	06/10/2005	CHACH36	1	2	127	4
4	06/10/2005	CHACH36	1	2	127	4
5	06/10/2005	SCASC05	1	2	127	4
6	06/10/2005	SCASC05	1	2	127	4
7	06/10/2005	SCASC05	1	2	127	4

Here you have various options on the Destination (i.e. the Pasgear file you import to) on whether to replace, insert, or overwrite records. You must also define the format of the external source data (date format, field delimiter), as well on which lines the import should start. If there is a header line for the columns to be imported then check [ ] Has header row, and the import will automatically start at line2. The source preview gives you the layout on how the data are read from the source.

Step 3



Map the source data to the destination table fields. Mark Include on fields to be imported and map them from the preview

Destination:	Incl	Source column	Default value	Unit
Date	V	1		
Station	V	3		
Species	V	2		
Gear	V		GN	
Mesh	V	5		
Relative effort	V		90	
Duration (hour)				hour
Setting Type	V	4		
Number	V	8		
Length (cm)	V	9		cm
Weight (g)				g
Sex				
Gonadal Stage				
Stratum	V	6		
Rank				

Source preview: Import starts at row (line) 2

	1	2	3	4	5	6
1	Date	Spec code	Station	Setting Type	Mesh (mm)	Habitat
2	06/10/2005	CHACH36	1	2	127	4
3	06/10/2005	CHACH36	1	2	127	4
4	06/10/2005	CHACH36	1	2	127	4
5	06/10/2005	SCASC05	1	2	127	4
6	06/10/2005	SCASC05	1	2	127	4
7	06/10/2005	SCASC05	1	2	127	4
8	06/10/2005	HOLMY05	1	2	127	4

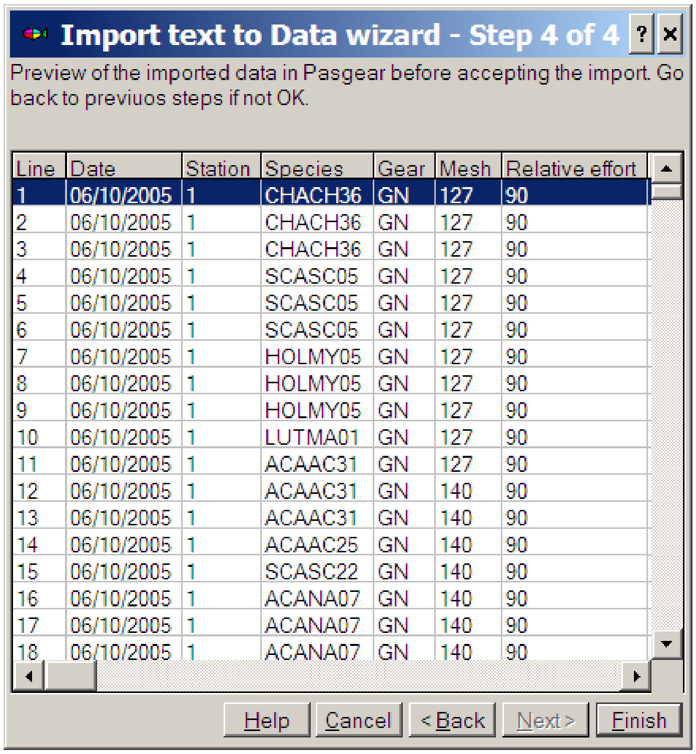
Pasgear 2 will import any number of fields in any (specified) order from the clipboard a text file where each line is a record (if not an aggregated matrix) as long as they correspond to the Pasgear formats.

First mark [V] the fields you want to import, if all then use [Mark all].

Next, The 'column'-number (1..n) given in the preview of a particular variable/field in the source-file should be indicated to the right of the corresponding Pasgear 2 field. This means that the source file can contain both more or less fields, and in any different order, than the Pasgear 2 fields, and that fields in the source file which are not to be imported are just 'skipped'. You can also enter a default value for fields where you don't have information in the source file, such as e.g. the gear code or any other field with constant values. It is only important that the values in the source file are within the range and format accepted by Pasgear 2 (see Table 1). The only exception is the date field which will accept a range of optional input values in Step 2.

Lastly, you should give the units (where applicable) to the imported fields. If some field already have defined units (see Database tables/column properties) then these will appear in combo boxes. Note that for a Pasgear record to be recognised as valid, then 3 fields are obligatory (Date, gear and mesh – indicated in red). If you do not have import values for any one of these fields then give a constant default value.

Step 4.


Import text to Data wizard - Step 4 of 4

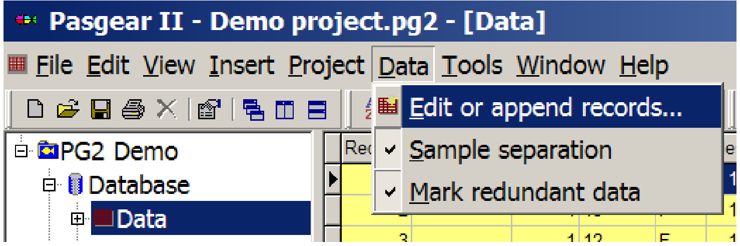
Preview of the imported data in Pasgear before accepting the import. Go back to previuos steps if not OK.

Line	Date	Station	Species	Gear	Mesh	Relative effort
1	06/10/2005	1	CHACH36	GN	127	90
2	06/10/2005	1	CHACH36	GN	127	90
3	06/10/2005	1	CHACH36	GN	127	90
4	06/10/2005	1	SCASC05	GN	127	90
5	06/10/2005	1	SCASC05	GN	127	90
6	06/10/2005	1	SCASC05	GN	127	90
7	06/10/2005	1	HOLMY05	GN	127	90
8	06/10/2005	1	HOLMY05	GN	127	90
9	06/10/2005	1	HOLMY05	GN	127	90
10	06/10/2005	1	LUTMA01	GN	127	90
11	06/10/2005	1	ACAAC31	GN	127	90
12	06/10/2005	1	ACAAC31	GN	140	90
13	06/10/2005	1	ACAAC31	GN	140	90
14	06/10/2005	1	ACAAC25	GN	140	90
15	06/10/2005	1	SCASC22	GN	140	90
16	06/10/2005	1	ACANA07	GN	140	90
17	06/10/2005	1	ACANA07	GN	140	90
18	06/10/2005	1	ACANA07	GN	140	90

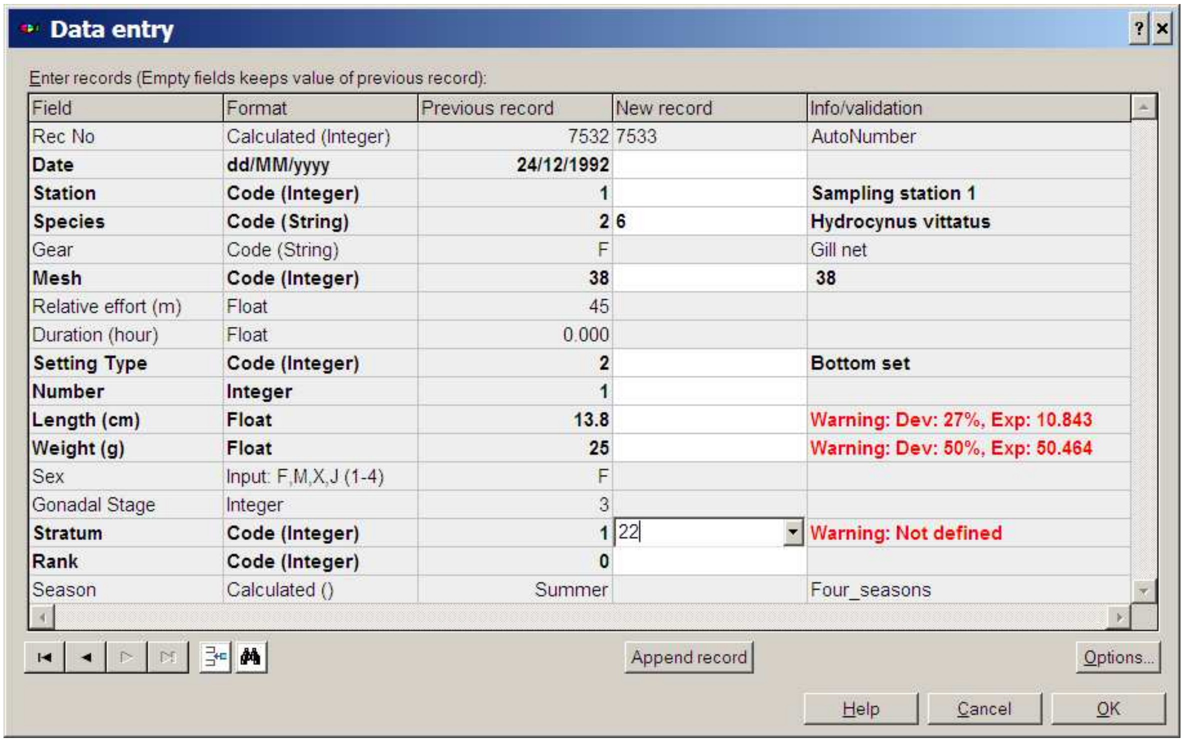
The last step in the Import wizard gives you a preview of how the data will be imported into Pasgear 2. If it is OK then press Finish, otherwise go back to previous steps and adjust your settings.

Enter data directly into Pasgear 2

Choose **Edit or Append records** or press ▦ in the main menu



The data entry dialogue:



Enter records (Empty fields keeps value of previous record):

Field	Format	Previous record	New record	Info/validation
Rec No	Calculated (Integer)	7532	7533	AutoNumber
Date	**dd/MM/yyyy**	**24/12/1992**		
Station	**Code (Integer)**	**1**		**Sampling station 1**
Species	**Code (String)**	**2**	**6**	**Hydrocynus vittatus**
Gear	Code (String)	F		Gill net
Mesh	**Code (Integer)**	**38**		**38**
Relative effort (m)	Float	45		
Duration (hour)	Float	0.000		
Setting Type	**Code (Integer)**	**2**		**Bottom set**
Number	**Integer**	**1**		
Length (cm)	**Float**	**13.8**		**Warning: Dev: 27%, Exp: 10.843**
Weight (g)	**Float**	**25**		**Warning: Dev: 50%, Exp: 50.464**
Sex	Input: F,M,X,J (1-4)	F		
Gonadal Stage	Integer	3		
Stratum	**Code (Integer)**	**1**	22	**Warning: Not defined**
Rank	**Code (Integer)**	**0**		
Season	Calculated ()	Summer		Four_seasons

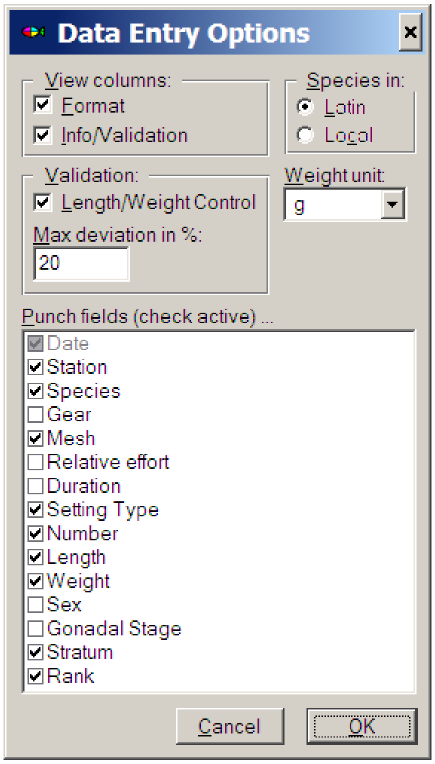
Each record in the DATA table (= each individual or frequency of same organism) contains and array of codes or Id fields (foreign codes) and field data values. Once the values have been entered in a field (i.e. after the first record data entry) they will remain unchanged in the subsequent new records by just pressing [Enter] on the field. Only fields that are actually entered with new Id’s or values have their contents changed. This feature makes it easy and fast to enter records as most fields do not change very often.

Data entry is fastest using the numeric keyboard, therefore the Sex fields (F,M, J or X = unknown) can be entered by corresponding numbers 1,2,4 or 3.

You can also set up the punching module in [Options] to jump over fields that are not changed often (they will simply inherit their value from the previous record) or connect the fields with the Id-Tables to give warnings when unknown codes and/or unrealistic combinations of length and weights are entered.
Once a record has been entered you append it to the data base by pressing [Append record], or just press [Enter] when you reach the bottom of the field array, and you are ready for the next record entry.

Data entry options

In addition there is a series of **Options** for Data entry, which can be accessed by pressing [Options..]



View columns:
☑ Format
☑ Info/Validation

Species in:
● Latin
○ Local

Validation:
☑ Length/Weight Control
Max deviation in %: 20

Weight unit: g

Punch fields (check active) ...
☑ Date
☑ Station
☑ Species
☐ Gear
☑ Mesh
☐ Relative effort
☐ Duration
☑ Setting Type
☑ Number
☑ Length
☑ Weight
☐ Sex
☐ Gonadal Stage
☑ Stratum
☑ Rank

- **View Columns.** Check these if you want these columns to appear in the data entry dialogue
  If the Id Tables (see database tables) have been created and entered, and you have the **Info/Validation** view checked on, you can optionally enter data while the **names** corresponding to the codes of species, stations, setting types etc. are simultaneously displayed. This will help you to remember and check the codes while entering the data. Ids that are not defined in the corresponding Id tables will be given a Warning: Not defined.

- **Species in Latin or Local**. This is for display of corresponding species names to the Id codes in either local language or in Latin.

- **Length-weight control**. Similarly, if the Id Table for Species has been created and the species-specific length-weight coefficients (see Analysis/Length-weight relationship) have been entered in the table, you can optionally enter data while the corresponding length and weight data are simultaneously controlled for mutual consistency (see also Queries/Length-weight validation). This will help you to avoid punching in obvious mistakes and/or disclose major errors in the field observations. Default level of warning is more than 20% deviance between observed and expected length.

- **Weight Unit.** Specify the unit of the weight field (all data will by default be converted and stored in grams)

- **Punch fields**. This for shortening the data entries by excluding fields that are not changed often. Fields that are not checked will be jumped when pressing [Enter] = next field in the data entry dialogue.

Edit previously punched records

You can use the data entry dialogue to go back and check/edit previously punched data. Use
⏮ ◀ ▷ ⏭ or the ‘Goto’ button 🔍 for navigating around in the data table.
With ⇥ you can insert new records anywhere in the file.

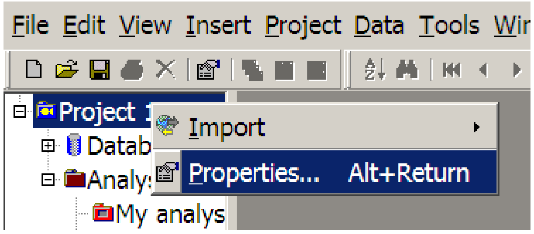
Project

Project properties (definitions and layout)

A Pasgear “project” is a set of data- and Id-files in binary format (*.bds) and a *.pg2 file in XML (eXtended Markup Language) format. The ‘bds’ files contain the data and the ‘pg2’ file holds the project properties, and keeps all the settings, queries, expressions, analyses, layouts, and graphics. When data are stored in binary format they cannot be red without using the programme.

Note that all editing or changes on the bds files is **always** done directly on the disk and **cannot be ‘undone’**. It is therefore highly recommended to have regular backups of the bds files (or the whole project by using Save as.. in the File menu). The pg2 file, however, will only be saved to disk on prompt. So any changes in running Pasgear 2, or changing properties, will only be kept when responding [Yes] to saving on exit, or pressing the save button 💾.

For accessing the project properties then *Right click the Project and choose properties from the popup menu.*

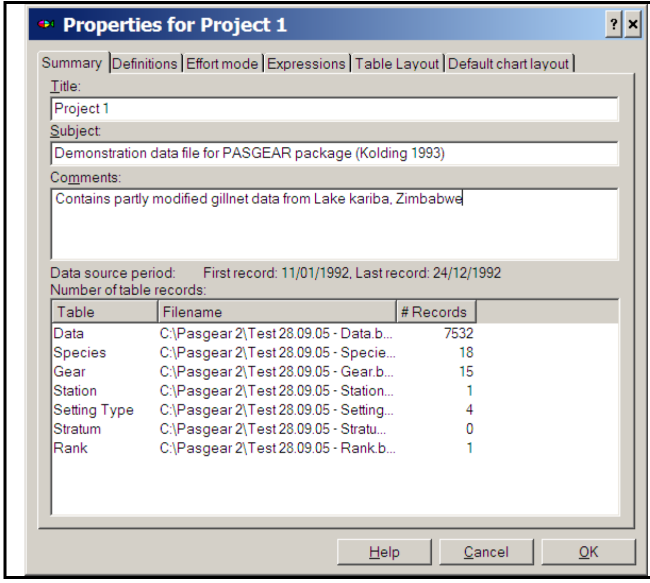


The project properties consist of a series of Tabs (each explained more fully below) that set up the configuration and properties of a Project:

- Summary

- Definitions
- Expressions
- Table layout
- Default chart layout

Summary



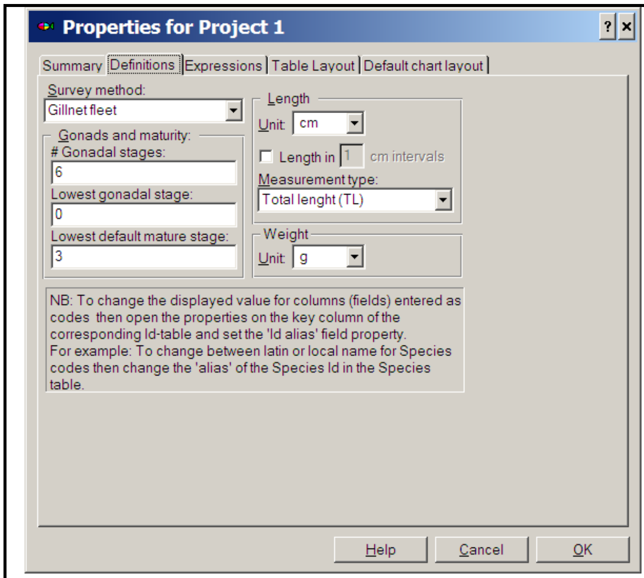
Title:
Project 1
Subject:
Demonstration data file for PASGEAR package (Kolding 1993)
Comments:
Contains partly modified gillnet data from Lake kariba, Zimbabwe

Data source period: First record: 11/01/1992, Last record: 24/12/1992
Number of table records:

Table	Filename	# Records
Data	C:\Pasgear 2\Test 28.09.05 - Data.b...	7532
Species	C:\Pasgear 2\Test 28.09.05 - Specie...	18
Gear	C:\Pasgear 2\Test 28.09.05 - Gear.b...	15
Station	C:\Pasgear 2\Test 28.09.05 - Station...	1
Setting Type	C:\Pasgear 2\Test 28.09.05 - Setting...	4
Stratum	C:\Pasgear 2\Test 28.09.05 - Stratu...	0
Rank	C:\Pasgear 2\Test 28.09.05 - Rank.b...	1

- Title: Project title
- Subject: Project subject
- Comments: Project comments
- Data source period: The first and last records date of data table.
- Tables: List of connected data and Id tables, their names, files and # records.

Definitions



Survey method:
Gillnet fleet

Gonads and maturity:
Gonadal stages:
6
Lowest gonadal stage:
0
Lowest default mature stage:
3

Length
Unit: cm
☐ Length in 1 cm intervals
Measurement type:
Total lenght (TL)

Weight
Unit: g

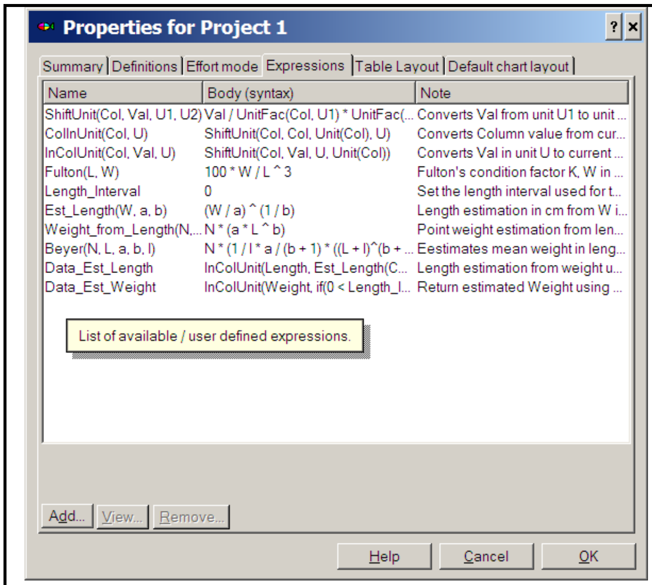
NB: To change the displayed value for columns (fields) entered as codes then open the properties on the key column of the corresponding Id-table and set the 'Id alias' field property.
For example: To change between latin or local name for Species codes then change the 'alias' of the Species Id in the Species table.

- Survey method: Choose sampling method or define your own.
- Gonads and maturity: (Depends on this data field exist).
    - Gonadal stages: # of defined stages allowed in data field 'gonads'.
    - Lowest gonadal stage: lowest stage allowed in data field 'gonads'.
    - Lowest default mature stage: Lowest mature stage for mature individuals. Used by the analysis variable 'Mature percentage'. This property can be overridden at the analysis. (See Analysis properties).
- Length: Defines the length unit used in groupings and display. This is a short cut to the unit

definitions in the length column (see column properties).

- Check length in [ ] cm intervals if the data consist of length frequencies instead of individually measured organisms (see Types of biological data under Overview).
- Weight: Defines the weight unit used in groupings and display. This is a short cut to the unit definitions in the Weight column (see column properties).
- To change a coded column alias, e.g. species name in Latin or local. (see column properties).

Expressions

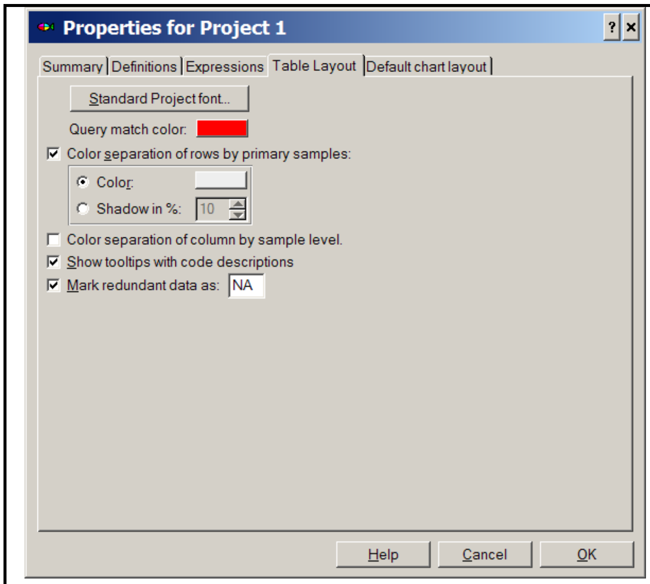


Name	Body (syntax)	Note
ShiftUnit(Col, Val, U1, U2)	Val / UnitFac(Col, U1) * UnitFac(...	Converts Val from unit U1 to unit ...
ColInUnit(Col, U)	ShiftUnit(Col, Col, Unit(Col), U)	Converts Column value from cur...
InColUnit(Col, Val, U)	ShiftUnit(Col, Val, U, Unit(Col))	Converts Val in unit U to current ...
Fulton(L, W)	100 * W / L ^ 3	Fulton's condition factor K, W in ...
Length_Interval	0	Set the length interval used for t...
Est_Length(W, a, b)	(W / a) ^ (1 / b)	Length estimation in cm from W i...
Weight_from_Length(N,...	N * (a * L ^ b)	Point weight estimation from len...
Beyer(N, L, a, b, l)	N * (1 / l * a / (b + 1) * ((L + l)^(b + ...	Eestimates mean weight in leng...
Data_Est_Length	InColUnit(Length, Est_Length(C...	Length estimation from weight u...
Data_Est_Weight	InColUnit(Weight, if(0 < Length_I...	Return estimated Weight using ...

List of available / user defined expressions.

- Give the expressions library of the project. An expression is any function, query, or constant that can be called and used in Queries, Analyses, Calculated fields, or for replacement of values in the data base.
- An expression consists of a **name** (with or without ingoing parameters), a **body**, which is the syntax that can be compiled, and a **note** describing what the expression does.
- Some of the expressions are predefined and cannot be edited or removed. But the user can define any number of expressions and add to the library by pressing [Add] on the property form, or using [Add to library] on the expression builder. User defined expressions can be edited or removed.

Table layout



Query match color: 🟥

☑ Color separation of rows by primary samples:

● Color:

○ Shadow in %: 10

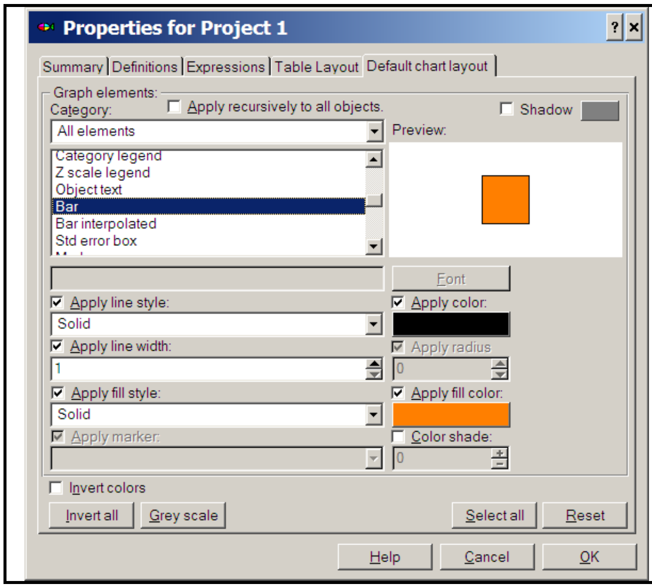
☐ Color separation of column by sample level.

☑ Show tooltips with code descriptions

☑ Mark redundant data as: NA

- The Table layout gives options for how the tables should be displayed
- The **Standard font**. Sets the default font used on all tables and new objects.
- **Query match color.** Sets the text color used for records that matches the connected query.
- **Color separation of rows by primary samples**. If checked the primary sample units (PSU) will be separated visually in the data table by a color or a shade depending on the choice. If you have colors on the columns (see column properties) you should use shadow**.**
- **Color separation of columns by sample level**. This option will automatically color the columns depending on their defined sample level: PSU (red), physical (blue) or biological (green). (see column properties).
- **Show tool tips with code description**. This option enables you to directly see the lookup alias of codes from the related Id-Table.
- **Mark redundant data as NA**. If checked then illegal records **or** data not used in analyses will be marked as NA (default) in table, or user defined layout.

Default chart layout



Category: All elements

☑ Apply line style: Solid

☑ Apply color:

☑ Apply line width: 1

☑ Apply radius: 0

☑ Apply fill style: Solid

☑ Apply fill color:

Pasgear has an inbuilt default for the layout of all graph elements (objects). This, however, can be modified by the user. All newly created charts and diagrams will follow the default setup. Graph elements can be categorized into

- All elements
- Font elements
- Fill elements
- Line elements

Each element can be individually designed (size, color, fill style etc.)

Expressions
    Expression builder
        Formal and actual parameters
Expression Library

Expressions

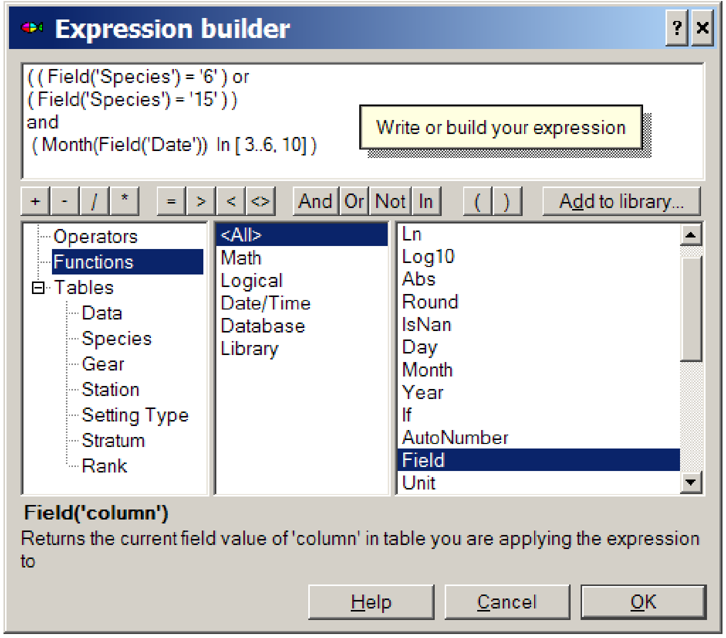
Pasgear 2 has an expression compiler that enables you to solve a large range of mathematical or logical expressions (see Project properties/expressions). Expressions are used as queries, functions (mathematical calculations), or data base operations such as e.g. lookup codes in Id-tables or replacement of values. By using expressions on added calculated columns in the data table, or in the queries, the user has virtually unlimited possibilities for selecting, grouping and analyzing the data (see example in Database tables/Calculated fields). Expressions are compiled in Pasgear, and therefore needs to be constructed with the correct syntax for the compiler to deal with it. For complicated sequential expression the use of parentheses is therefore important to divide the operations into smaller parts.

Expressions can be any combinations of the following:
1. A number (float) – i.e. a simple floating value like 2, 2.1 or 2e-2
2. A variable - a quoted expression for an object representing a value. 'var1'
3. A binary operation - an arithmetic or logical operation with 2 operands (+, -, *, /, DIV, MOD, ^, = , <, etc..)
4. A unary operation - an arithmetic or logical operation with 1 operands (LN, LOG10, EXP, ABS, etc..)
5. A function - a freely defined function with an optional parameter list like FuncName(Arg1, Arg2..ArgN).

6. Another expression – identified by a name and stored in the Expression library.

Expression builder

By using the expression builder you can see the available operators and functions and their syntax, as well as the available tables and fields to select values (variables) from:



```
( ( Field('Species') = '6' ) or
( Field('Species') = '15' ) )
and
 ( Month(Field('Date'))  In [ 3..6, 10] )
```

Field('column')

Returns the current field value of 'column' in table you are applying the expression to

The expression builder has an edit field where you can directly type in the expression. Alternatively you can add operators or parentheses by using the speed buttons

Below the speed buttons you can select among the main expressions groups (Operators, Functions, or Tables) in the left pane. The middle pane will then give you the group categories:

- **Operators**: All or Arithmetic, Comparison, Logical, and set
- **Functions**: All or Mathematical, Logical, Date/time, Database, and Library
- **Tables**: The fields (columns) of the selected table.

If you have chosen operators or functions you can then select and insert (double click) the chosen operator or function in the expression pane.

Below the pane the syntax is shown, and - if necessary - a help text of the chosen operator or function is given.

Formal and actual parameters

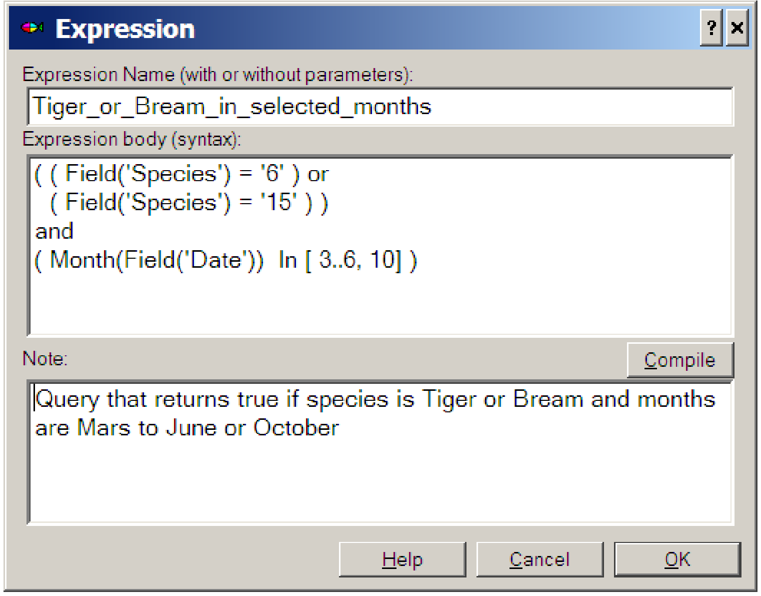
Some functions have formal parameters that must be specified to actual parameters. The formal parameters are given in parenthesis after the expression (function) name: <name>(formal param1, formal param2,…).

For example if you want to look up a value from a specific table field you use the function **Lookup(table, column)**, which has two formal parameters: table and column. To make this function operational you must the insert the actual table name (choose group table, choose table and double click) instead of formal parameter ‘table’, and the actual column name in chosen table instead of formal parameter ‘column’ (choose a column in second pane and double click).

Expression Library

You can add any new expressions (queries, functions, calculated fields, etc.) to the Project Expression library by pressing Add to library...

This will open a new dialogue where you must give the expression a header (name), specify if there are formal parameters, and optionally give an explanatory note (annotation) to you expression:



Tiger_or_Bream_in_selected_months

```
( ( Field('Species') = '6' ) or
  ( Field('Species') = '15' ) )
and
( Month(Field('Date'))  In [ 3..6, 10] )
```

Query that returns true if species is Tiger or Bream and months are Mars to June or October

When pressing OK the expression will be saved to the library and you can recall it anytime you need it by just referring to the name. For example, if you want to use it inside other expressions. Expressions added to the library will be displayed and can be retrieved in the expression builder under Functions/Library.

To delete or edit a user defined expression you must use the [Edit] or [Remove] option in Project properties/Expressions. If you change or edit the syntax you can check if it compiles without errors by pressing [Compile].

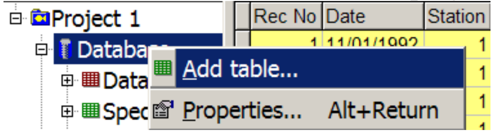
Database

Database tables

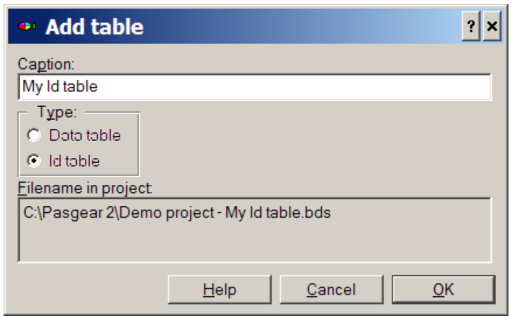
A standard (default) Pasgear database has one table for data and a series of 6 Id tables (or reference tables for translating codes (foreign keys) in the data table to corresponding names and additional information) for species, gear, station, set-type, stratum and rank. The unique Id field (or primary key) in the Id tables corresponds to an Id-field (foreign key) in the data table. (i.e. Species-Id, Gear-Id etc.).

Add tables

You can add more tables to the project by right-click on Database and choose Add table:



This will open the Add table dialog:



Caption:
My Id table

Type:
○ Data table
● Id table

Filename in project
C:\Pasgear 2\Demo project - My Id table.bds

Here you give the name (caption) of the table. The file name will automatically be a combination of the project name + <table caption>.bds, and located on the same path as the project.

You can choose between two table types:

- Data table
- Id table

NB! Although you can add a data table, these cannot yet be accessed by the analyses so this option is for future use only.

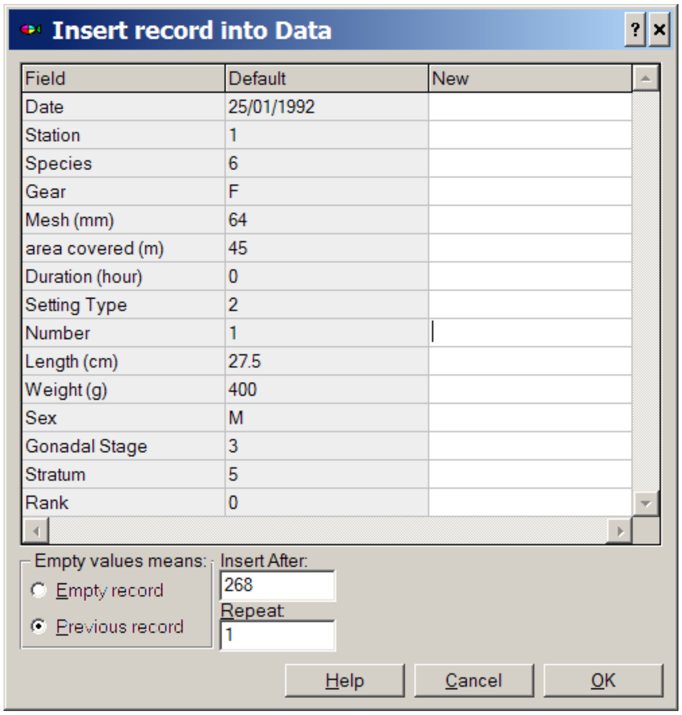
The Id-tables (or reference tables) are for creating relationships for translating codes in the data table to corresponding names and additional information. Thus if you want more code fields in the data tables and associate these with a new Id-Table you can create one and set the relationship on the data field property.

Delete tables

You can delete some of the standard Pasgear Id-tables (Station, Setting type, and Stratum) and any of the tables you have added yourself by highlight the table and choose delete.

Insert records

You can insert new records anywhere in a table by right click and choose **Insert record** or simply press [Ins]. The insert record dialog consist of 2 columns, where the first (**Default**) contains the field values of the last record you want to insert after. This means you only have to fill out the fields in the **New** column where information has changed. All the others will inherit their values form the Default column. The number of records to be inserted is by default 1, but you can insert any number of records by changing the value in the **Repeat** field. If you have many records to insert, it is easier to simply append them at the end using Enter data, and then sort them in to their right place in the table.



Field	Default	New
Date	25/01/1992	
Station	1	
Species	6	
Gear	F	
Mesh (mm)	64	
area covered (m)	45	
Duration (hour)	0	
Setting Type	2	
Number	1	
Length (cm)	27.5	
Weight (g)	400	
Sex	M	
Gonadal Stage	3	
Stratum	5	
Rank	0	

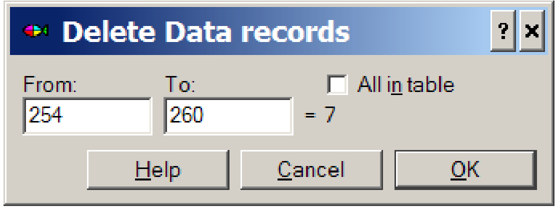
Empty values means:
○ Empty record
● Previous record

Insert After: 268
Repeat: 1

Delete records

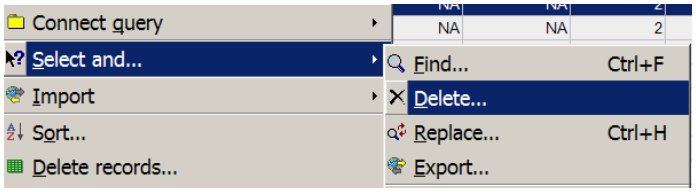
Choose the starting record, the right click on the table and choose **Delete records**. In the **From** and **To** fields you can then define the number of records to be deleted. You can use arrow keys, PgUp, PgDown, Home or End when specifying the number of records to be deleted.

Note that once you press [OK], the records will be deleted permanently; there are no Undo or Ctrl Z options in Pasgear when working on the tables. You should therefore always make sure you have a backup of the tables and project.

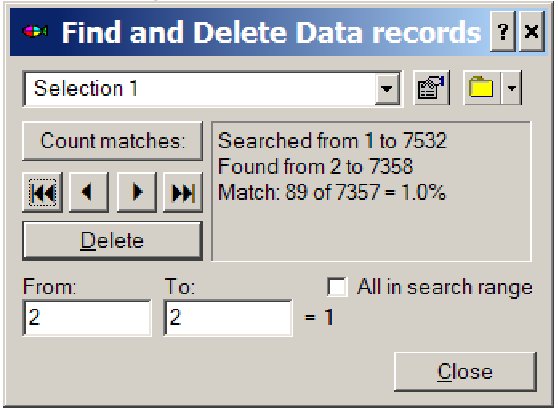


From: 254
To: 260 = 7
☐ All in table

You can also select and delete records using a Query by right click on the Table and choose



By going to the first record matching the Query filter (press ⏮) you can then delete one by one, a range by increasing the **To** field, or all found by checking [V] the box ‘**All in search range**’.



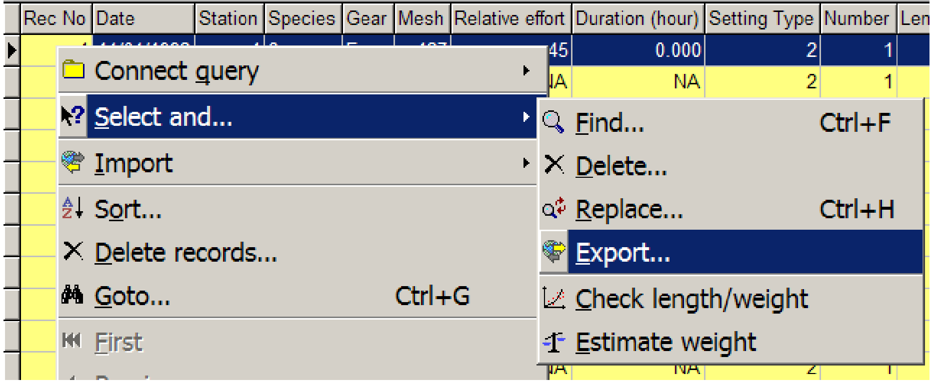
Selection 1

Searched from 1 to 7532
Found from 2 to 7358
Match: 89 of 7357 = 1.0%

From: 2
To: 2 = 1
☐ All in search range

Export data

Data can be selected and exported from any table by right click on table ► **Select** and.. ► **Export**..



This will open the export dialogue where you choose the export format:

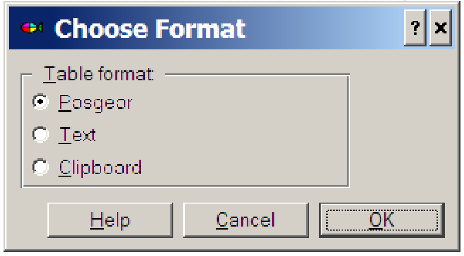
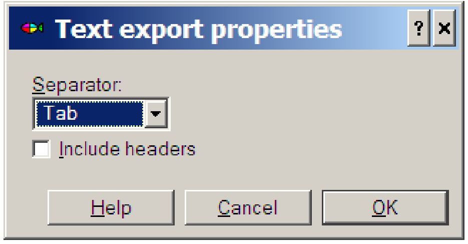


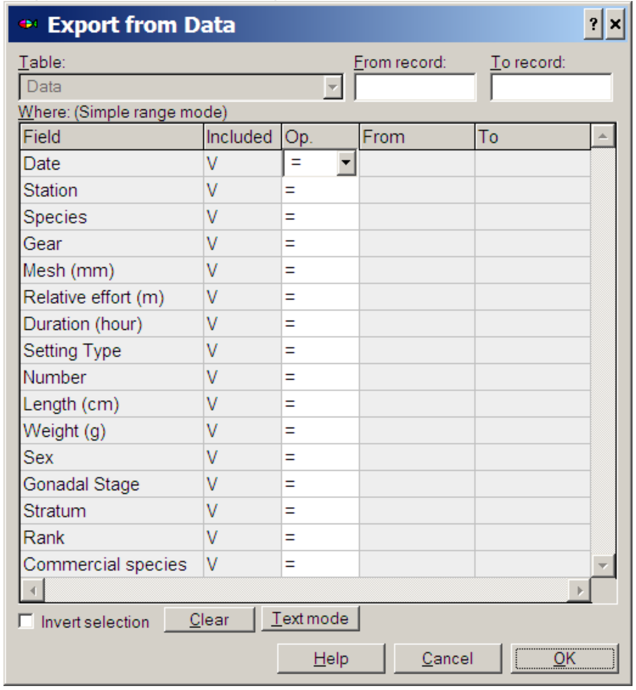
Table format
● Posgeor
○ Text
○ Clipboard

If you choose ‘text’ or ‘clipboard’ then you must define the field separator and whether you want to include the column headers when exporting (check Include headers on).



Separator: Tab
☐ Include headers

Finally specify which record range to export (in From record – To record) and the Fields – by checking [V] Included on/off + optionally the field ranges you want to export.
If you don't make any specifications or restrictions the whole data table will be exported.



Field	Included	Op.	From	To
Date	V	=		
Station	V	=		
Species	V	=		
Gear	V	=		
Mesh (mm)	V	=		
Relative effort (m)	V	=		
Duration (hour)	V	=		
Setting Type	V	=		
Number	V	=		
Length (cm)	V	=		
Weight (g)	V	=		
Sex	V	=		
Gonadal Stage	V	=		
Stratum	V	=		
Rank	V	=		
Commercial species	V	=		

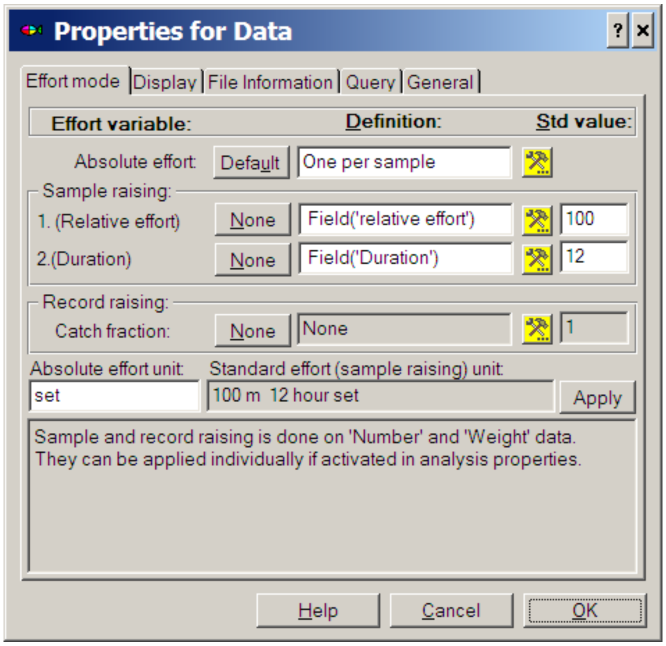
Properties for tables

Right click table in project tree view and choose properties or highlight table and press

Each table has a property frame where you can define the display, connected query, and the table caption (under the general tab). In addition to can see the file information of the table (path, size and last modified).

Effort mode

If the table is a **Data table** it has one more frame for effort mode:



Effort variable: | Definition: | Std value:

Absolute effort: Default | One per sample

Sample raising:
1. (Relative effort) None | Field('relative effort') | 100
2.(Duration) None | Field('Duration') | 12

Record raising:
Catch fraction: None | None | 1

Absolute effort unit: set
Standard effort (sample raising) unit: 100 m 12 hour set | Apply

Sample and record raising is done on 'Number' and 'Weight' data.
They can be applied individually if activated in analysis properties.

Here you define:

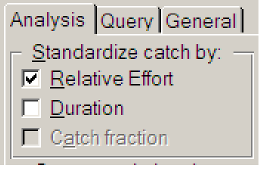
1) The absolute effort (value and unit), (see also Overview/ Effort definition and sample raising modes). Default (1) means that the absolute effort is the same as a primary sample unit (one sample = one absolute effort unit). If this is not the case you must specify in the definition field the column (or expression) where the absolute effort values are given.

2) The relative effort and the duration effort used for sample raising to a given standard value. If not applicable then press None and definition and standard value will both be set to "None" (code =1). Otherwise you must specify in the definition field the column (or expression) where the relative effort and duration values are given.

3) The record raising (if applicable see Effort definition and sample raising modes) and the column (or expression) where these values are given

4) The relative effort units, e.g. m for relative effort expressing net length, and hours or minutes for duration, are defined (if applicable) in the respective column properties/Units. They will appear in the combined definition for 'Standard effort (or sample raising) unit' field when pressing Apply.

By changing the standard values of the relative efforts, all calculations will be adjusted accordingly if 'Standardize catch by' is checked in the Analysis properties:

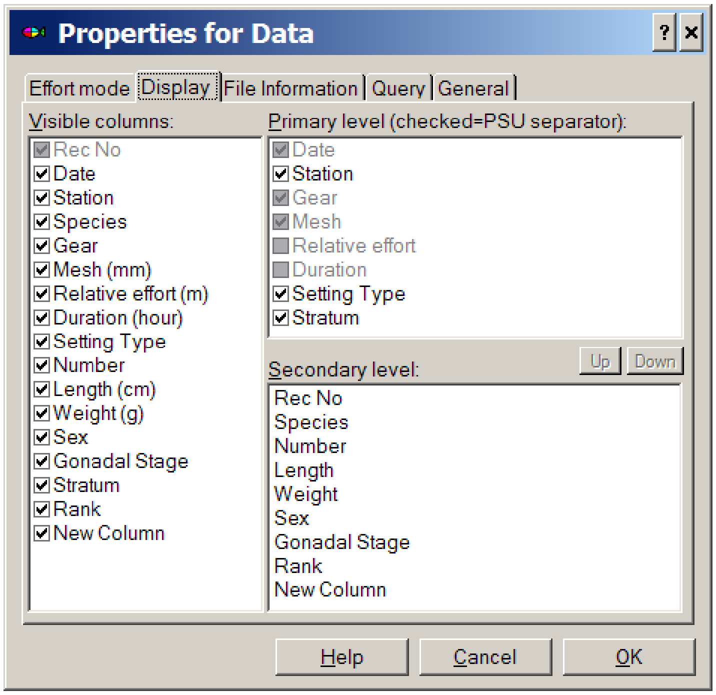


Standardize catch by:
☑ Relative Effort
☐ Duration
☐ Catch fraction

If any of the raising options are set to None in the data table properties (definition and standard values = 1), then these will be disabled in the Analysis properties, as e.g. Catch fraction in the example above.

Display

Under the Display tab you can define the number of visible columns (= checked), and for the data table you can change some of the field columns to being physical (with or without PSU separation) or biological (see 2-stage sampling design in one record).



Visible columns:

☑ Rec No
☑ Date
☑ Station
☑ Species
☑ Gear
☑ Mesh (mm)
☑ Relative effort (m)
☑ Duration (hour)
☑ Setting Type
☑ Number
☑ Length (cm)
☑ Weight (g)
☑ Sex
☑ Gonadal Stage
☑ Stratum
☑ Rank
☑ New Column

Primary level (checked=PSU separator):

☑ Date
☑ Station
☑ Gear
☑ Mesh
☐ Relative effort
☐ Duration
☑ Setting Type
☑ Stratum

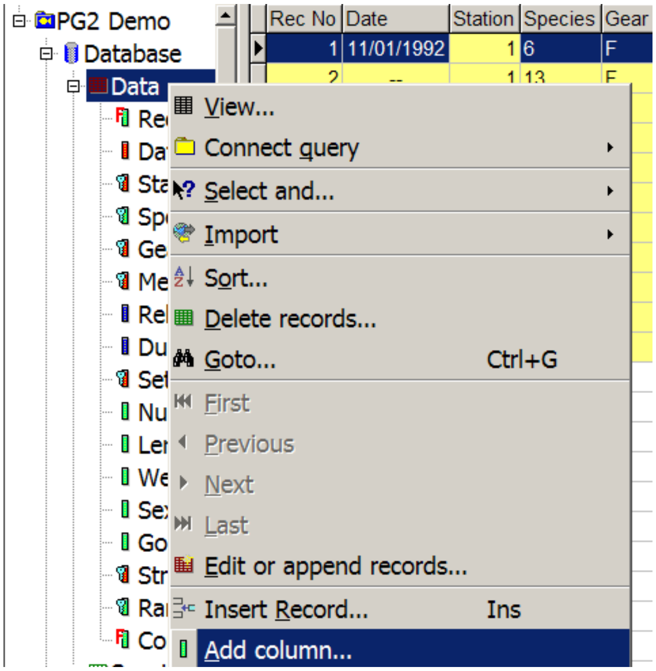
Secondary level:

Rec No
Species
Number
Length
Weight
Sex
Gonadal Stage
Rank
New Column

In the project tree view you can see some of the properties of the fields by their icons. Fields that are used as sample separators (PSU fields) are red, other physical fields are blue, while biological fields are green. In addition you can see if the field is a code field with an Id-relationship (reference) to one of the Id-tables (marked by a key🔑), or if it is a calculated field (marked by a red **F** for function).

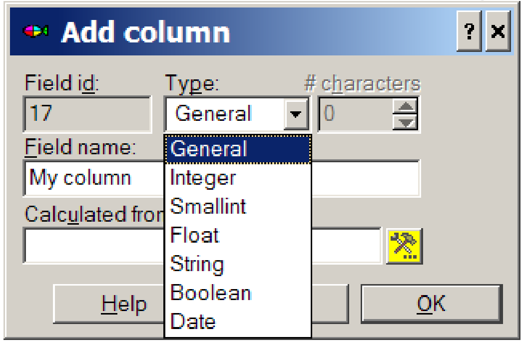
Add new fields (columns)

You can add any number of new fields (columns) to a table by right click on the table and choose ‘add column’ in the pop-menu:



Rec No	Date	Station	Species	Gear
1	11/01/1992	1	6	F
2	--	1	13	F

This will open the Add column dialog:



Field id: 17

Type: General

characters 0

Field name: My column

Here you can give the column a name and specify the type you want. If you choose string you must also give the maximum number of characters that the filed can take. The type ‘General’ is for calculated fields.

Calculated fields

You can choose the column to be a calculated expression. In that case the type will automatically be set to ‘General’. Calculated fields will take no physical space in the database. You must then specify the expression, optionally by using the expression builder . Calculated fields are very useful for making new groupings on the data. If, for example, you want to group the species into commercial and non-commercial you can do this by making a calculated field and give it the following expression:

If (Field('Species') in [‘code1’, ‘code2’, ‘code3’, etc], 'commercial', 'non-commercial')

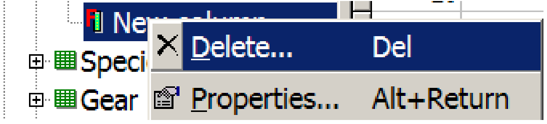
Or you can nest several ‘If’ expressions inside each other to make several groups

If (Field('Species') = 'code1', 'name1', If(Field('Species') = 'code2', 'name2', 'Other'))

You can then group your analyses on this calculated field.

Delete columns

Of the standard (default) Pasgear columns, you can only delete Stratum, Gonadal stages, Sex, Setting type and Duration. In addition you can delete any of the new columns you have added to a table by highlighting the column and press delete:

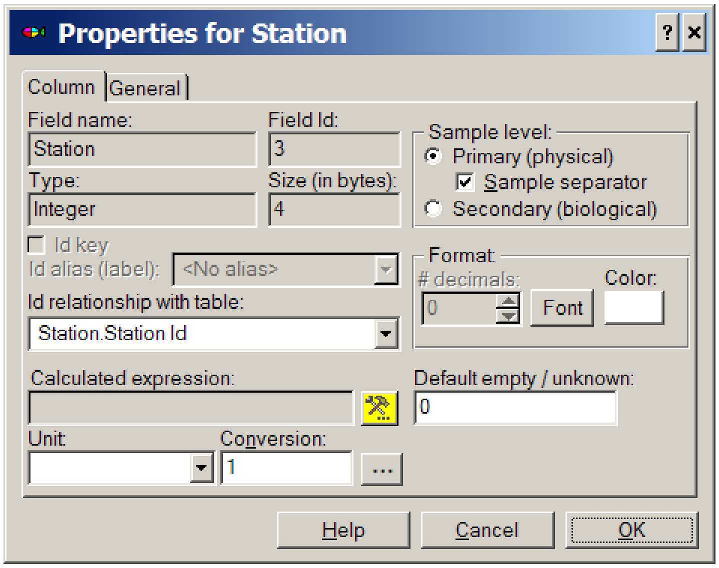


Field (column) properties

Right click field in project tree view and choose properties or highlight field and press

Depending on the type of the field (column), and table, you have various options:

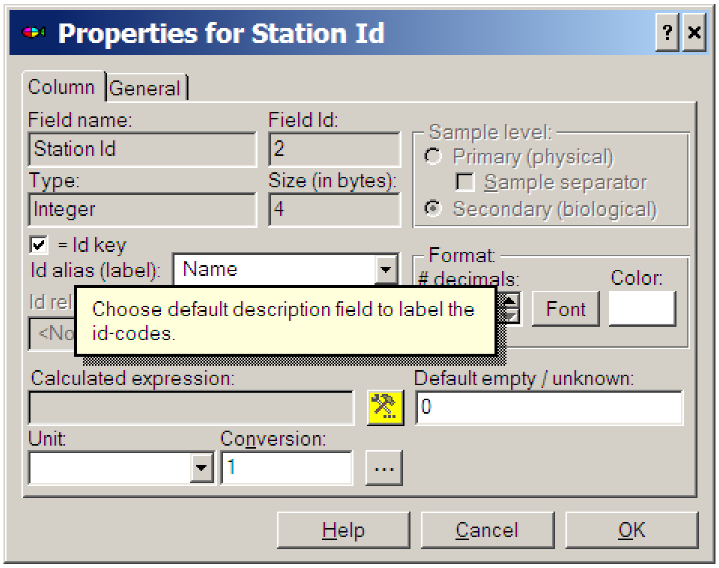
- If the table is an Id-table (reference table) you can define the field as a primary key field by checking ☑ Id key
- If the field is a primary key field you can then choose the default Id ‘alias’ (Label) that the key shall translate into.
- If the field is a code (foreign key) in a data table you can set the Id relationship with an Id-table and the corresponding primary key. (see example below on how the <station> field in the data table is related to the <Station Id> field in the ‘Station’ table, and that the default alias is the <name> field)
- If the field is a calculated expression (only applicable to new columns, see add fields), you must specify the expression, optionally by using the expression builder
- If the field has a metric unit (stored unit has conversion = 1) you can choose or define other units, e.g. stored length unit = mm, then cm has conversion 0.1. You can also add more units to the list by pressing …
- You can set the default empty or unknown value for the field
- Sample level: For some of the standard Pasgear fields, and all new fields you can define the sample level as physical (with or without being a sample separator) or biological (see 2-stage sampling design in one record)
- Finally you can specify the field format (# decimals, font and background color)


Properties for Station

Column | General

Field name: Station
Field Id: 3
Type: Integer
Size (in bytes): 4
Sample level:
● Primary (physical)
☑ Sample separator
○ Secondary (biological)
☐ Id key
Id alias (label): <No alias>
Id relationship with table: Station.Station Id
Format:
decimals: 0
Color:
Calculated expression:
Default empty / unknown: 0
Unit:
Conversion: 1

This means that field <Station> in the data table has Id relationship with the <Station Id> field in the ‘Station’ Id table:


Properties for Station Id

Column | General

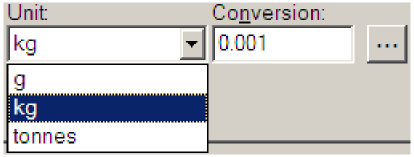
Field name: Station Id
Field Id: 2
Type: Integer
Size (in bytes): 4
Sample level:
○ Primary (physical)
☐ Sample separator
● Secondary (biological)
☑ = Id key
Id alias (label): Name
Id rel
<No
Choose default description field to label the id-codes.
Format:
decimals:
Color:
Calculated expression:
Default empty / unknown: 0
Unit:
Conversion: 1

The primary key field <Station Id> in the ‘Station’ Id table will be translated into the content of the <Name> field when running analyses.

Thus on the data table you can set Id relationships of the values which are in codes to a set of primary keys (list of all codes used) in an Id-Table. On the Id-Table you give the field that contains the information the codes should be translated into. For example if you want the species codes to be translated in ‘Local name’ instead of ‘Latin name’ you change the Id alias from the field with caption ‘Latin name’ to the field with caption ‘Local name’ in the species table on the ‘Species Id’.

Defining and adding units

Numeric fields can have automatic units (e.g. length and weight) where you can toggle between different units by choosing from the ‘Unit’ combobox:

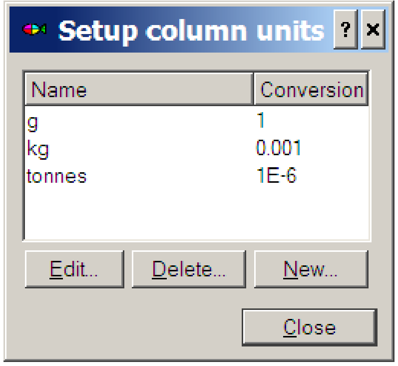


Unit: kg

Conversion: 0.001

The ‘Conversion’ field will then give you the multiplication factor between the displayed unit and the value as stored in the database. Data are preferably stored at the lowest resolution possible, so that for weight, for example, the stored unit is grams, so that the conversion factor between grams and kg (the optional displayed unit) will be 0.001.

You can add any units of your own by pressing ... , which will open the Unit editor:

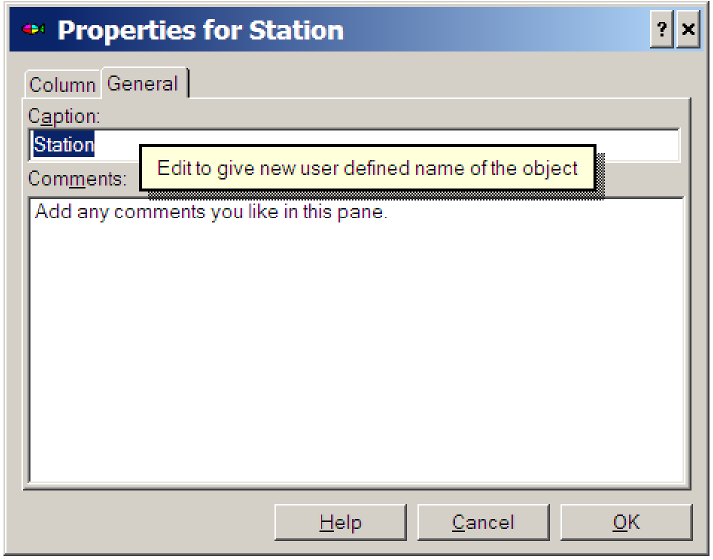


Name	Conversion
g	1
kg	0.001
tonnes	1E-6

Press [New] to define a new unit and corresponding conversion factor

Renaming fields (and any othe object)

Almost any object in a Pasgear Project can be renamed according to your preferences. If the property page has a tab called ‘General’ (usually located as the last tab on any property page), then by opening this you can rename the object your are addressing. Thus if you want to rename any of the fields, tables, queries, analyses, charts etc. then open property page and change the name in the ‘Caption’ edit field. Similarly you can add any personal comments, memos, etc. to any of the objects in the ‘Comments’ pane.



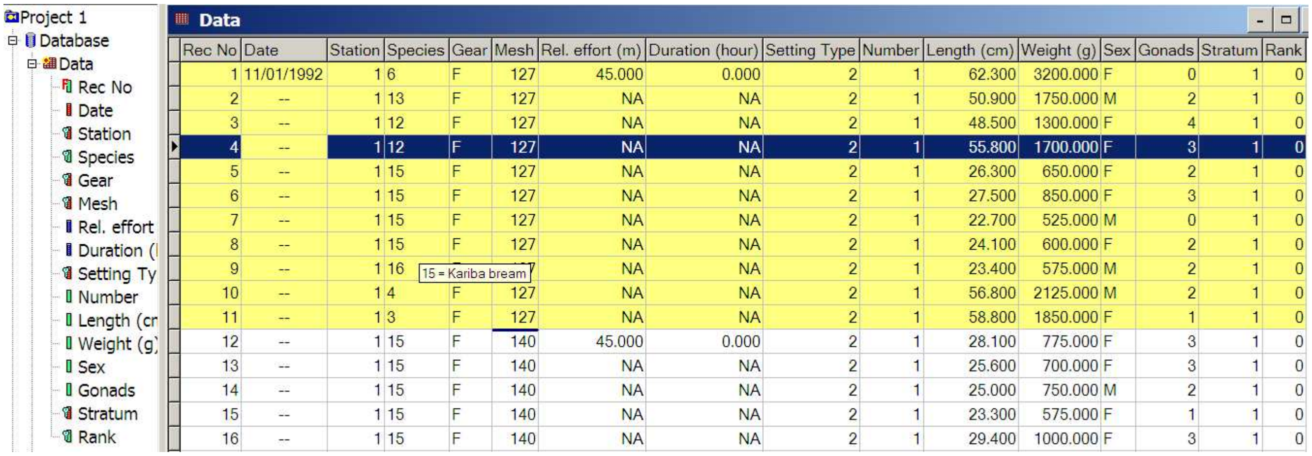
Caption: Station

Edit to give new user defined name of the object

Comments:

Add any comments you like in this pane.

Data table



Rec No	Date	Station	Species	Gear	Mesh	Rel. effort (m)	Duration (hour)	Setting Type	Number	Length (cm)	Weight (g)	Sex	Gonads	Stratum	Rank
1	11/01/1992	1	6	F	127	45.000	0.000	2	1	62.300	3200.000	F	0	1	0
2	--	1	13	F	127	NA	NA	2	1	50.900	1750.000	M	2	1	0
3	--	1	12	F	127	NA	NA	2	1	48.500	1300.000	F	4	1	0
4	--	1	12	F	127	NA	NA	2	1	55.800	1700.000	F	3	1	0
5	--	1	15	F	127	NA	NA	2	1	26.300	650.000	F	2	1	0
6	--	1	15	F	127	NA	NA	2	1	27.500	850.000	F	3	1	0
7	--	1	15	F	127	NA	NA	2	1	22.700	525.000	M	0	1	0
8	--	1	15	F	127	NA	NA	2	1	24.100	600.000	F	2	1	0
9	--	1	16	[illegible]	[illegible]	NA	NA	2	1	23.400	575.000	M	2	1	0
10	--	1	4	F	127	NA	NA	2	1	56.800	2125.000	M	2	1	0
11	--	1	3	F	127	NA	NA	2	1	58.800	1850.000	F	1	1	0
12	--	1	15	F	140	45.000	0.000	2	1	28.100	775.000	F	3	1	0
13	--	1	15	F	140	NA	NA	2	1	25.600	700.000	F	3	1	0
14	--	1	15	F	140	NA	NA	2	1	25.000	750.000	M	2	1	0
15	--	1	15	F	140	NA	NA	2	1	23.300	575.000	F	1	1	0
16	--	1	15	F	140	NA	NA	2	1	29.400	1000.000	F	3	1	0

Fields (columns):

In a standard Pasgear data base there are the following fields by default (see also table 1):

- Record number (automatically calculated)
- Date: Uses local short date format (In MS Windows, choose Control Panel ▶ Regional options ▶ Date). NB. A date is mandatory for the record to be accepted as valid (see 2-stage sampling design in one record).
- Station: Numeric station Id. Corresponds to Station table ▶ Id
- Species: String Id. (7 chars) Corresponds to Species table ▶ Id. Unknown default = '0'. NB: If Species Id is known, the record is considered as a valid secondary sample record (SSU), or valid biological record (see 2-stage sampling design in one record).
- Gear: String Id. (2 chars) Corresponds to Gear table ▶ Gear Id
- Mesh: Numeric Id: Corresponds to Gear table ▶ Mesh Id. NB: If 'Mesh' is known (<> 0), the record is considered as a valid primary sample record (PSU), or physical record (see 2-stage sampling design in one record).
- Relative Effort field. Relative effort type is defined in the Data property page ▶ effort mode. (see also Effort definition and sample raising modes or Table 3)
- Duration field. Duration type is defined in the Data property page ▶ effort mode. (see also Effort definition and sample raising modes) .
- Setting type: Numeric Id: Corresponds to 'Id' field in Setting type table.
- Number: Number of individuals per record. Unknown = 0. (see Biological data (the secondary sampling level))
- Length: length of individuals in record. Unknown = 0
- Weight: Weight of individuals in record. Unknown = 0
- Sex: The values Males (M), females (F), unknown (X), juvenile (J) is allowed here. Default = unknown = X.
- Gonadal stage: The gonadal stage (numeric value) of the individuals in the record. The range of valid gonadal stages is defined in the Project Properties.
- Stratum: Numeric Id: Corresponds to 'Id' field in Stratum table.
- Rank: Numeric Id: Corresponds to 'Id' field in Rank table.

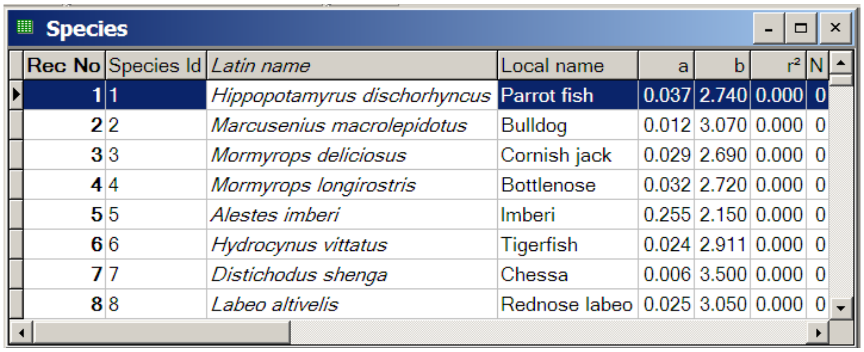
Id Tables or Reference Tables

The Id tables are used to store information on the Id's used in the data table, such as names etc. This works a bit like a relational database and the purpose is to save space in the often very large data tables.

All Id tables can automatically be filled out with Id's from the Data table by right click and choose 'Build from table..' on the pop-up menu and choose the table to build from. This will add all the codes (foreign keys) found in the chosen table.

Standard Pasgear Id-tables are the following:

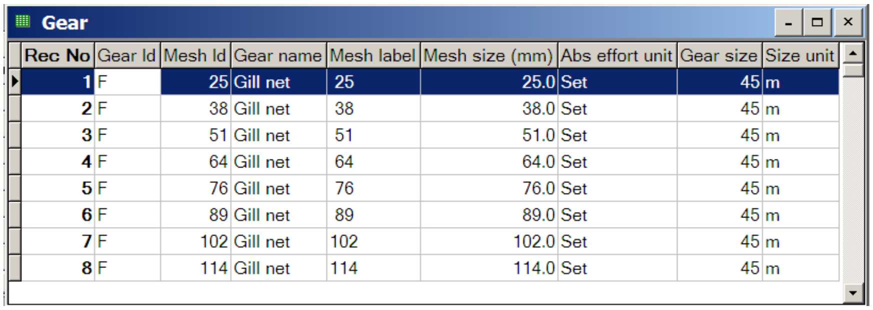
Species table



Rec No	Species Id	*Latin name*	Local name	a	b	r²	N
1	1	*Hippopotamyrus dischorhyncus*	Parrot fish	0.037	2.740	0.000	0
2	2	*Marcusenius macrolepidotus*	Bulldog	0.012	3.070	0.000	0
3	3	*Mormyrops deliciosus*	Cornish jack	0.029	2.690	0.000	0
4	4	*Mormyrops longirostris*	Bottlenose	0.032	2.720	0.000	0
5	5	*Alestes imberi*	Imberi	0.255	2.150	0.000	0
6	6	*Hydrocynus vittatus*	Tigerfish	0.024	2.911	0.000	0
7	7	*Distichodus shenga*	Chessa	0.006	3.500	0.000	0
8	8	*Labeo altivelis*	Rednose labeo	0.025	3.050	0.000	0

The **Species Id** (7 character text string) should correspond with the ones you have used in the species field in the DATA table. Here you can now enter the Latin name well as the common or local name. The **a** and **b** fields are the length-weight coefficients for the particular species, and the $\mathbf{r}^2$ and **N** fields are statistics describing their estimates from the data (if applicable). The length-weight coefficients can be entered manually (values for most species can be found in FishBase http://www.fishbase.org/search.php) or automatically when calculated from your own data (see Analysis/Length-weight relationship).

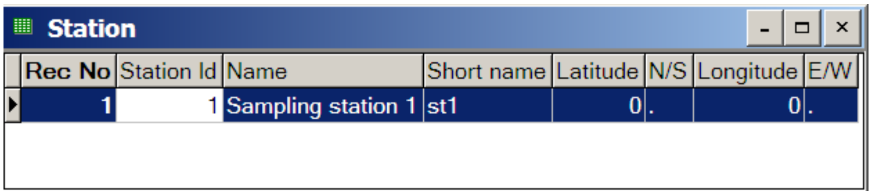
Gear table



Rec No	Gear Id	Mesh Id	Gear name	Mesh label	Mesh size (mm)	Abs effort unit	Gear size	Size unit
1	F	25	Gill net	25	25.0	Set	45	m
2	F	38	Gill net	38	38.0	Set	45	m
3	F	51	Gill net	51	51.0	Set	45	m
4	F	64	Gill net	64	64.0	Set	45	m
5	F	76	Gill net	76	76.0	Set	45	m
6	F	89	Gill net	89	89.0	Set	45	m
7	F	102	Gill net	102	102.0	Set	45	m
8	F	114	Gill net	114	114.0	Set	45	m

The **Gear Id** (2 characters) and **Mesh Id** (integer) should correspond with the ones you have used in the DATA table. Under **Gear name** you can enter the name of the gear (e.g. gill net) and the **Mesh label** (= output format) you want to be used in the Analysis Tables. The **Mesh Id** and the **Mesh label** can be identical, and usually are when operating with different mesh sizes. However, the order of **Mesh label**s given in the gear file will also determine the order they are shown in some of the Analysis tables calculated by PASGEAR, and the order they are presented in the gear selectivity tool. It is therefore recommended to arrange and enter (or sort) various mesh or hook sizes in ascending order.
The **Mesh size (unit)** field is for calculating the selectivity coefficients in the Ger selectivity tool and should always be numerical. The **Abs effort unit** field is used to define the fishing unit (Set, haul, etc.). The field **Gear size** field is used if for example the different mesh panels, long lines or traps have different dimensions. They may be identical to the **Relative effort** field in the DATA table, but this is not critical for most calculations of standardized CPUE. For calculations of standardized CPUE, only the **Relative effort** field in the data base is used, except when **Gear table** is chosen in Data properties/**Effort mode**/Sample raising/relative effort where the value in the **Gear size** field will be used. The **Size unit** gives the unit of the **Gear size**.

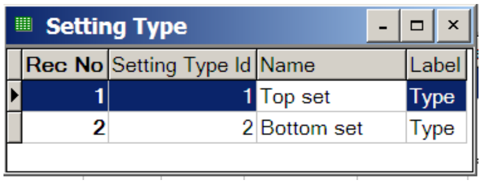
Station table



Rec No	Station Id	Name	Short name	Latitude	N/S	Longitude	E/W
1	1	Sampling station 1	st1	0	.	0	.

The **Station Id** (integer) should correspond with the ones you have used in the DATA table. Here you can now enter the **name** of the station, and a **short** print name for graphs, and the geographic position of the station.

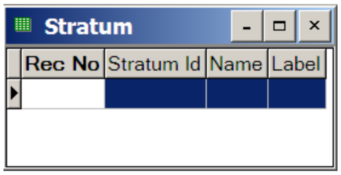
Setting type table



Rec No	Setting Type Id	Name	Label
1	1	Top set	Type
2	2	Bottom set	Type

The **Setting type Id** (integer) should correspond with the ones you have used in the DATA table. Here you can now enter the name of the particular **Type** of setting, and the **label** (e.g. Sector, Strata, Type, etc.). A type of setting can be any of your choice. For example 'Bottom set', 'Top set', 'Perpendicular to the shore' or whatever your particular samplings programme wants to stratify by or differentiate between.

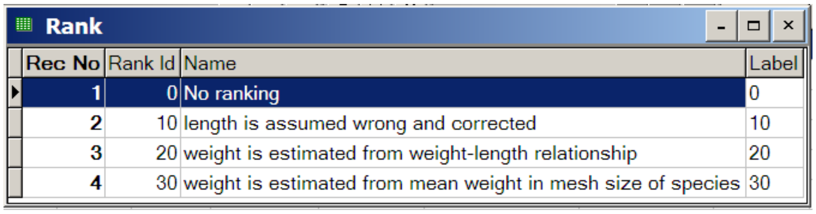
Stratum table



Rec No	Stratum Id	Name	Label

This table works like the Setting type table. The **Stratum Id** (integer) should correspond with the ones you have used in the DATA table etc.

Rank table

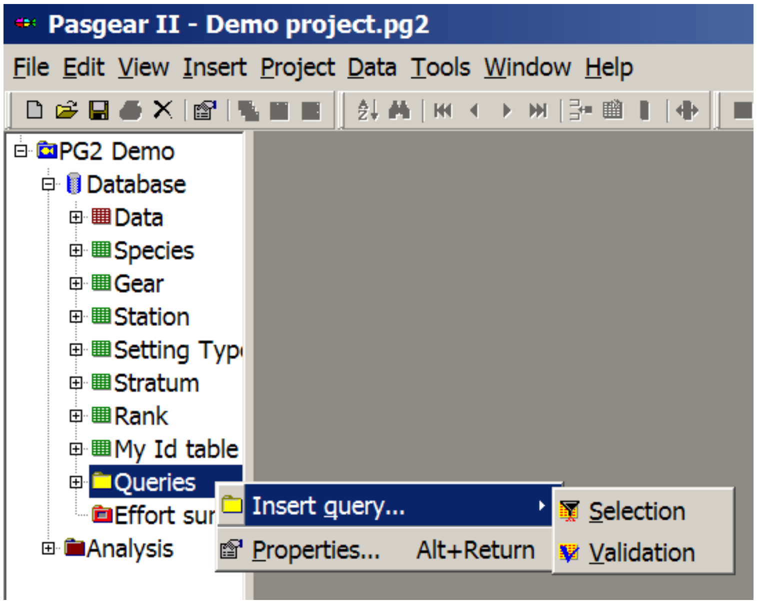


Rec No	Rank Id	Name	Label
1	0	No ranking	0
2	10	length is assumed wrong and corrected	10
3	20	weight is estimated from weight-length relationship	20
4	30	weight is estimated from mean weight in mesh size of species	30

The Rank field in Pasgear is typically used to give validity to the specific records. When the **Rank** field contains values different from 0 the record will not be used in certain of the calculation procedures, such as length-weight relationship, mean weight, condition factor etc. See Find, check, and correct records for possible sources of errors and the suggested codes used when correcting them.

The **Rank Id** (integer) should correspond with the ones you have used in the DATA table. Here you can now enter the meaning (**Name**) of the Id, i.e. how is this particular record ranked in terms of validity and usage. The **label** is used in the Analysis tables and diagrams

Queries



Queries (or selection filters) can be applied to the data or Id tables or to any analysis. Right click 'Queries' under database in project tree and choose query type from the sub menu. Available queries are:

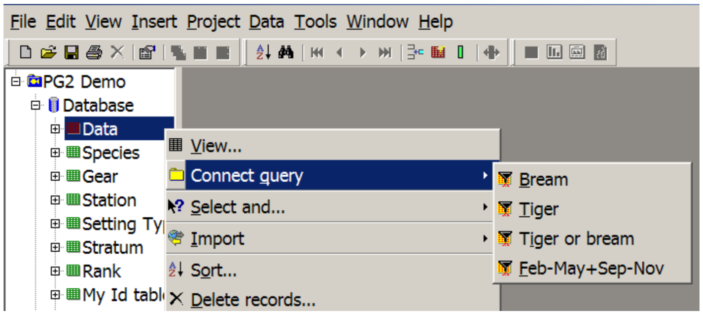
1. Selection
2. Validation

Once a query has been created, defined and optionally named it will be stored in the *pg2 file of the project and available on all subsequent entries. Queries can also be cut, copied, pasted, edited, deleted or – if in text mode - be added to Expression library.

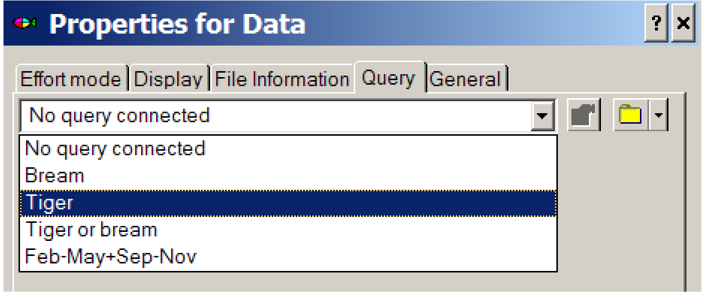
Connecting a query

You can connect a query to a table or analysis in two ways:

1. Right click on the table or the analysis, Choose **Connect query** ►and select one of the available queries that have been created:

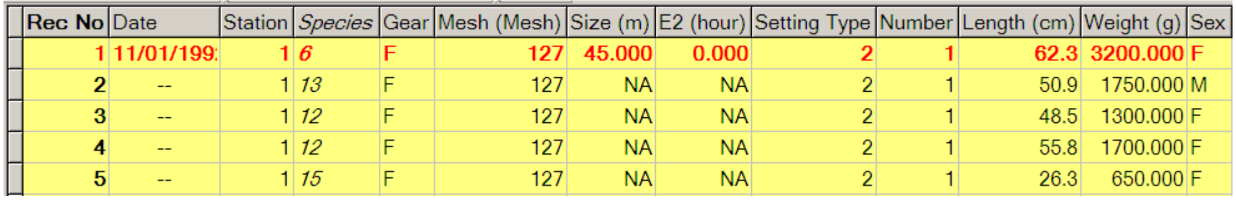


2. Go into properties [Alt + Return], choose the Query tab and select an available query or create a new by clicking 📁



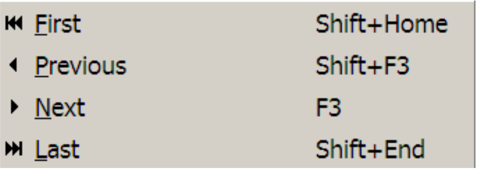
To disconnect a query then choose **No query connected** in the combo box, or Right click on the table or the analysis ► '**Disconnect [Query name]**'.

When a query is connected to a table the all matching records (those satisfying the selection criteria, see below) will be displayed in a different color (default = red but can be changed in Project properties/Tables layout).



Rec No	Date	Station	Species	Gear	Mesh (Mesh)	Size (m)	E2 (hour)	Setting Type	Number	Length (cm)	Weight (g)	Sex
1	11/01/199	1	6	F	127	45.000	0.000	2	1	62.3	3200.000	F
2	--	1	13	F	127	NA	NA	2	1	50.9	1750.000	M
3	--	1	12	F	127	NA	NA	2	1	48.5	1300.000	F
4	--	1	12	F	127	NA	NA	2	1	55.8	1700.000	F
5	--	1	15	F	127	NA	NA	2	1	26.3	650.000	F

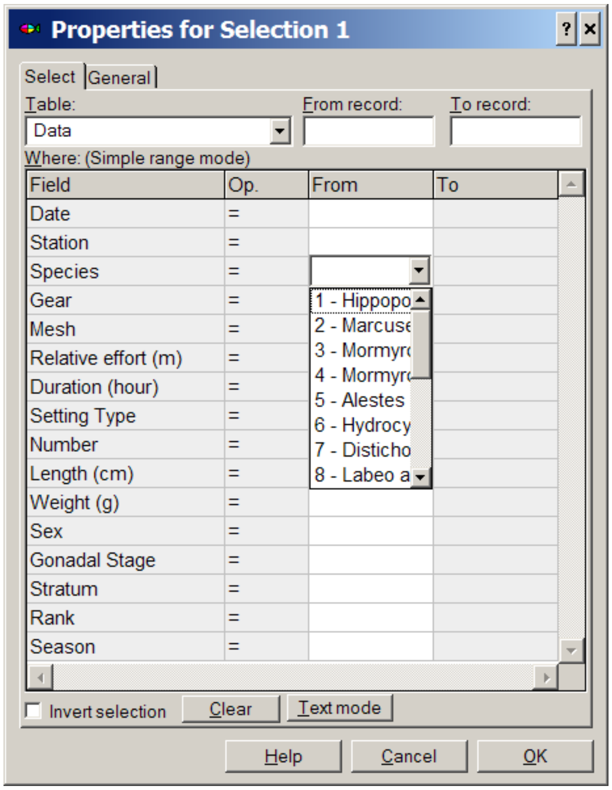
All matching records can be found successively by using the ⏮ ◀ ▶ ⏭ on the tool bar or using accelerator keys:



⏮ First	Shift+Home
◀ Previous	Shift+F3
▶ Next	F3
⏭ Last	Shift+End

Selection query

Simple range mode



Field	Op.	From	To
Date	=		
Station	=		
Species	=		
Gear	=		
Mesh	=		
Relative effort (m)	=		
Duration (hour)	=		
Setting Type	=		
Number	=		
Length (cm)	=		
Weight (g)	=		
Sex	=		
Gonadal Stage	=		
Stratum	=		
Rank	=		
Season	=		

For those fields you want to select on then enter field filter criteria (Operator: ‘**=**’ (equal) or ‘**<>**’ (different from), and the **From** value, **To** value) on the table fields. If you do not enter any values in the **From** or **To** columns, the field is not considered in the filter. If you only enter a value in either the **From** or in the **To** column, the other (empty) column will be considered to take the either the maximum or the minimum value of the field respectively. Each field filtered are combined with logical **And** operator (all filter criteria must be satisfied at the same time for the record to be accepted). For those fields which are Id’s and have corresponding Id Tables, a combo-box will appear from which you can choose your Ids.

The ‘From record’ and ‘To record’ fields define the part of the database that the query shall refer to. All uses of a query are restricted to only apply within the defined record range. If you do not enter any values in the **From** or **To** record fields, the query will apply to the whole table/file.

Invert selection. At the bottom of the query dialogue there is a check box for inverting the selection. When this is checked it simply means that all records that returned true (i.e. satisfies the selection criteria) now returns false and vice versa.

Advanced text mode

Queries can be constructed either by defining simple ranges combined with logical **And** operators, or by pressing Text mode where you write the query in text (like using SQL). When in text mode you have unlimited possibilities of constructing a query using a range of operators such as **or**, **not**, **in, if** etc. (see Expressions).
To help building a query expression press to enter the Expression builder.

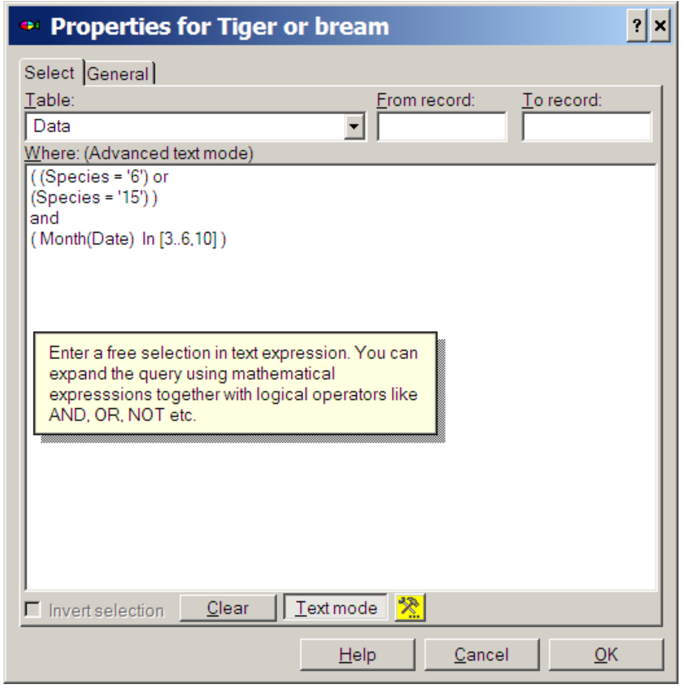


Table: Data

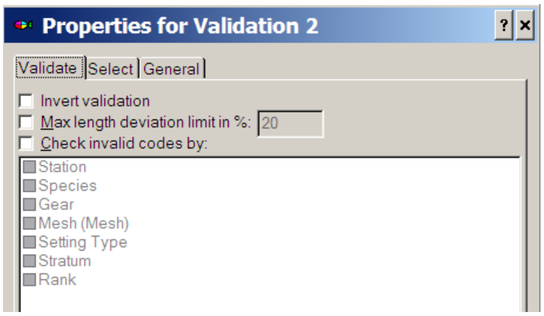
```
( (Species = '6') or
(Species = '15') )
and
( Month(Date)  In [3..6,10] )
```

Enter a free selection in text expression. You can expand the query using mathematical expresssions together with logical operators like AND, OR, NOT etc.

☐ Invert selection

In the example above the Query will only return those records containing either species '6' or '15' and only for the months March (3) to June (6) and October (10).

Validation query



☐ Invert validation
☐ Max length deviation limit in %: 20
☐ Check invalid codes by:
▨ Station
▨ Species
▨ Gear
▨ Mesh (Mesh)
▨ Setting Type
▨ Stratum
▨ Rank

A validation query is used to validate and check records. It consists of 2 parts

- A selection query (as above)
- A specification of the validation to be performed:
    - Length-weight validation
    - Validation of the Id's used in the Data table

Length-weight validation

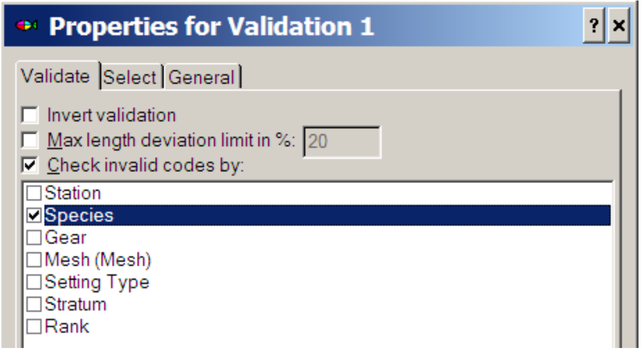
Check the box: [V] **Max length deviation in %** and set the percentage limit (default = 20%).
If the Species Id table (see Data tables) contains the coefficients of the species specific length-weight relationship (either entered manually or estimated form Pasgear, see Find, check, and correct records) then each record can be checked for deviations in the recorded lengths and weight values, where the deviation criteria is more than the specified percentage limit between the **observed** length, and the **expected** length (as calculated form the observed weight).

Length-weights can also be checked and directly edited by right click in the Data table and choose ► **Select and..** ► **Check length/weight**.

Validation of the Id's used in Data

This is used to check if the entered Id's (codes) of species, gears, mesh sizes, stations, settings etc, in the data file match with the one specified in the Id tables. All Id's found with no corresponding 'definition' in the Id table will be selected.

Check the box: [V] **Check invalid codes by**: and select the Id tables you want to check against



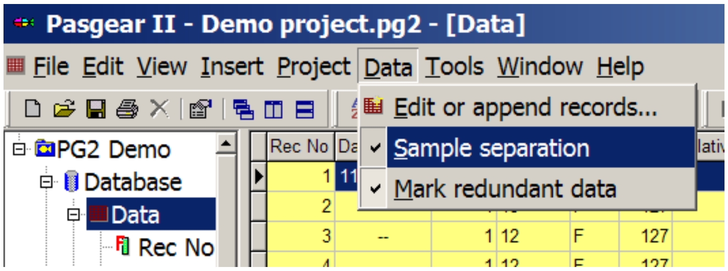
☐ Invert validation
☐ Max length deviation limit in %: 20
☑ Check invalid codes by:

- ☐ Station
- ☑ Species
- ☐ Gear
- ☐ Mesh (Mesh)
- ☐ Setting Type
- ☐ Stratum
- ☐ Rank

Effort summary

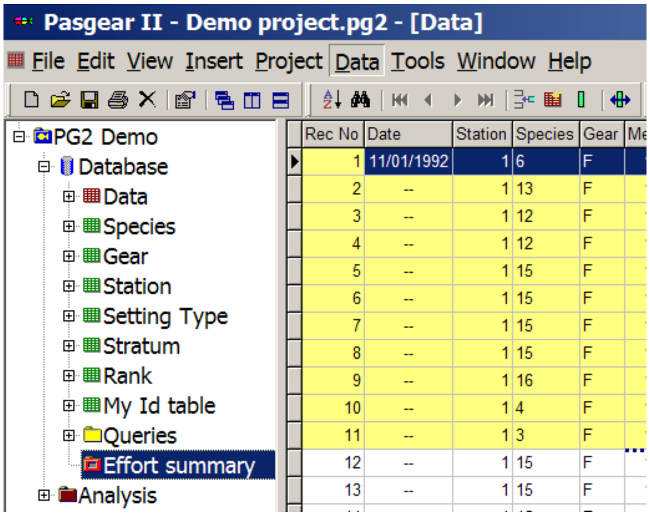
Chronology and the units of effort (settings or samples)

Chronology is important when analysing the data. This means that is each date should preferably follow chronologically and all fish caught in the same setting, gear, or sample, should follow each other (see also ‘sort data’ below). The reason is that the calculation and tabulating programmes will automatically assume that when any of the physical fields that define a primary sample unit (see Overview/2 stage sampling design in one record) such as either the **date** or the **station** or the **mesh/hook/trap size** or the **setting type** changes in the record series, then it is considered as a **new** setting or a new sample! You can see the separation of settings/samples in colours in the Data table by checking on ‘Sample separation’.



Check chronology and the number of settings/samples.

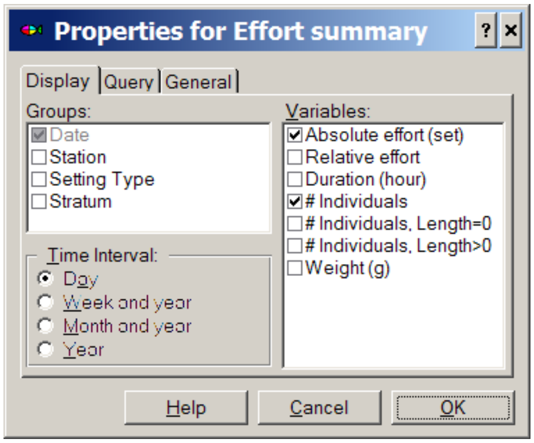
In order to check if all the settings or samples are correctly entered you should run the module **Effort summary** under Database in the tree view.



Rec No	Date	Station	Species	Gear	Me
1	11/01/1992	1	6	F	
2	--	1	13	F	
3	--	1	12	F	
4	--	1	12	F	
5	--	1	15	F	
6	--	1	15	F	
7	--	1	15	F	
8	--	1	15	F	
9	--	1	16	F	
10	--	1	4	F	
11	--	1	3	F	
12	--	1	15	F	
13	--	1	15	F	

This will give you a short summary of the date, record number where that day starts, the number of settings of each mesh size and optionally the total number of fish caught in each mesh size, relative effort etc.

Effort summary properties: Effort Summary can be run on various levels of group resolution and on various time intervals (date, week, month, and year). It can also give you a summary of selected variables within each defined group.



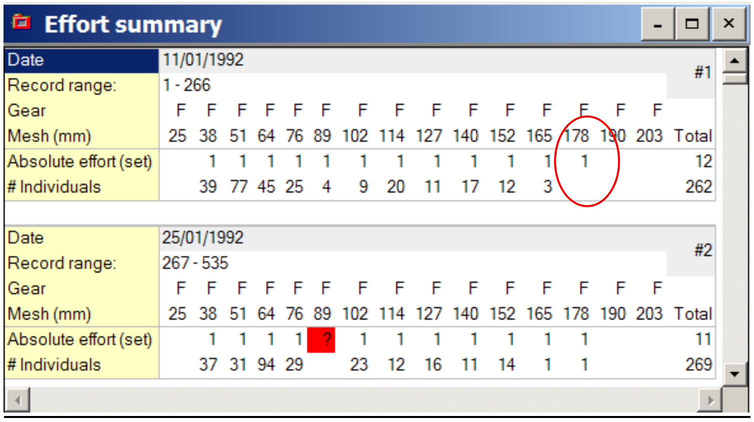
Groups:
☑ Date
☐ Station
☐ Setting Type
☐ Stratum

Time Interval:
● Day
○ Week and year
○ Month and year
○ Year

Variables:
☑ Absolute effort (set)
☐ Relative effort
☐ Duration (hour)
☑ # Individuals
☐ # Individuals, Length=0
☐ # Individuals, Length>0
☐ Weight (g)

The 'empty' setting.

For proper calculation of frequency of occurrence and catch per unit effort, it is very important that all 'empty settings' are included in the data base. An **empty setting** is a setting or hauls in which one or some of the mesh panels, lines or traps caught no fish. It is entered as a single record where all the physical fields are filled, but where the biological fields: the species (most importantly), number, length, weight and maturity, are all set to empty values, i.e. 0 or X (or the values that are defined as being empty, see column properties). An empty setting will be displayed as a unit of absolute effort but with no catch (#Individuals = 0), see mesh size 178 mm in the first date below, which is an empty setting.

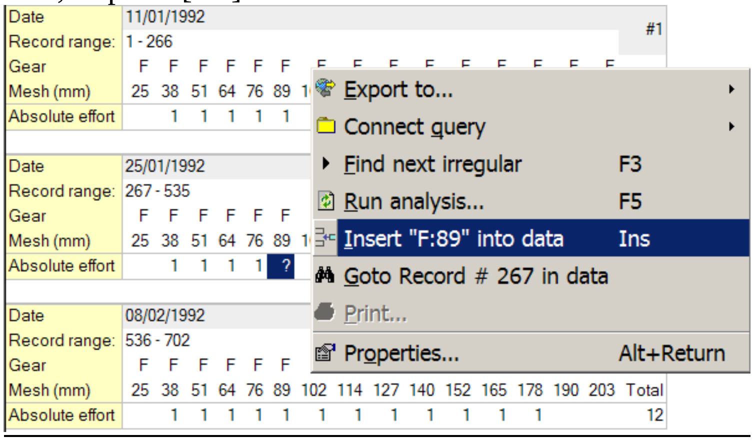


Date	11/01/1992															#1
Record range:	1 - 266															
Gear	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	
Mesh (mm)	25	38	51	64	76	89	102	114	127	140	152	165	178	190	203	Total
Absolute effort (set)		1	1	1	1	1	1	1	1	1	1	1	1			12
# Individuals		39	77	45	25	4	9	20	11	17	12	3				262

Date	25/01/1992															#2
Record range:	267 - 535															
Gear	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	
Mesh (mm)	25	38	51	64	76	89	102	114	127	140	152	165	178	190	203	Total
Absolute effort (set)		1	1	1	1	?	1	1	1	1	1	1	1			11
# Individuals		37	31	94	29		23	12	16	11	14	1	1			269

Missing settings

If a setting is ‘missing’, i.e. the mesh has not been registered in a date, but was found on the previous or the next date, it will be displayed in red with a question mark ?. If the missing setting is simply an empty setting then Pasgear 2 can automatically insert this into the data table by clicking on the ?, right click and choose ‘Insert [gear]:[mesh] into data, or press [Ins]



Date	11/01/1992						#1
Record range:	1 - 266						
Gear	F	F	F	F	F	F	
Mesh (mm)	25	38	51	64	76	89	
Absolute effort		1	1	1	1	1	

Date	25/01/1992					
Record range:	267 - 535					
Gear	F	F	F	F	F	F
Mesh (mm)	25	38	51	64	76	89
Absolute effort		1	1	1	1	?

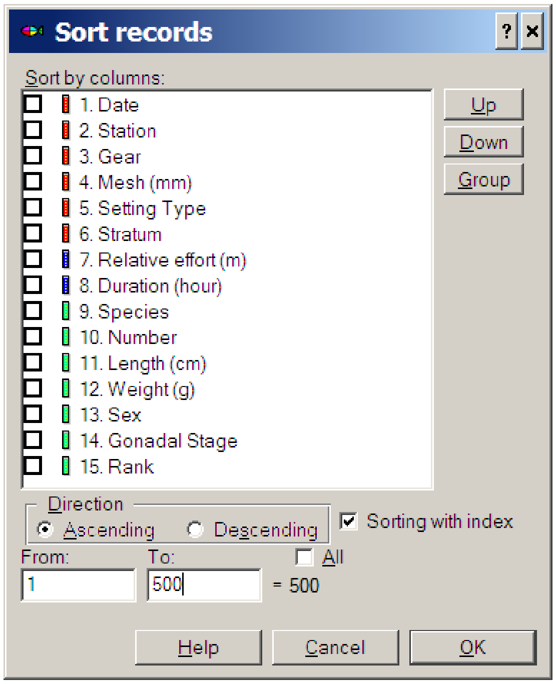
Date	08/02/1992														
Record range:	536 - 702														
Gear	F	F	F	F	F	F									
Mesh (mm)	25	38	51	64	76	89	102	114	127	140	152	165	178	190	203	Total
Absolute effort		1	1	1	1	1	1	1	1	1	1	1	1			12

Wrong absolute effort

If the data are not grouped (sorted) by the primary physical sample separation fields, then the number of primary sample units (PSU) and the absolute effort might be wrong. Pasgear assumes that the number of gear samples within a date (station, setting type etc.) should be constant. Thus if the same gear and mesh size is found at several places within a grouped time interval, and the absolute effort is not constant then these are shown in pink 2. If the reason is that the data have not been properly sorted then sort the data.

Sort the data

If the data have been entered (or imported) haphazardly and the dates and other physical fields are out of order, you can run the sorting module which will sort the records. Right click on table and choose **Sort** or press A↓Z**. NB.** Sorting can be done on any table.



Direction: ● Ascending ○ Descending

☑ Sorting with index

From: 1 To: 500 ☐ All = 500

The sorting will be done hierarchically down to the last item checked. You can define any sorting hierarchy by highlighting a field and move it up or down the list by the [Up] or [**D**own] buttons. By choosing [Group] the hierarchy will set up automatically by first PSU fields (red) the other physical fields (Blue) and lastly biological fields (green).

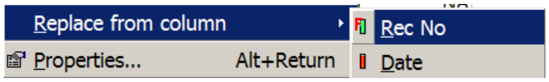
Check the **Sorting with index** in order to keep groups of data that do not need sorting, i.e. the catch in a mesh size, in the original order they were entered (this will facilitate the retrieval of records under cleaning and correction procedures as they will have the same order as the original sheets).

Lastly define the record range of the sorting operation in the From – To edit boxes or choose all by checking ☐ All.

Once the data are correctly entered, sorted chronologically, and all the settings are correctly displayed in the effort summary, you can start to 'clean' the data base for recording or punching mistakes as best as possible (see Find, check and correct records).

Hint: If you want to create an index so that you always can retrieve the original data order, irrespective of later sorting then do the following:

1) Add new column of type integer, and name it e.g. ‘Original Index’
2) Right click ‘Original Index’, choose **Replace from column ► Rec No**.

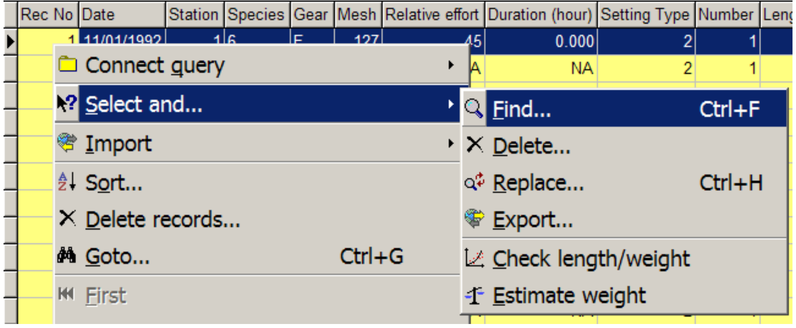


Find, check, and correct records

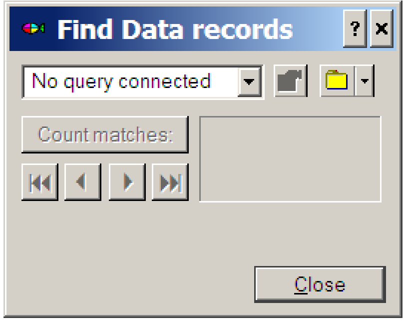
Find any specific records

By creating a query and entering the search criteria (any query combination) you can find one or several of the codes and values in a table (see Queries).

Alternatively press [Ctrl + F] or right click► **Select and..** ► **Find**

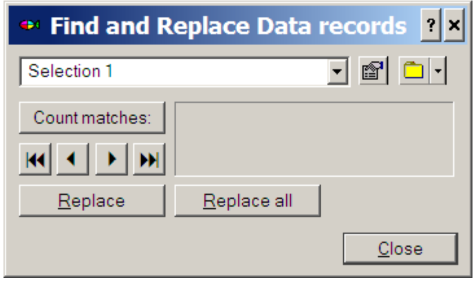


Here you can connect an already existing query or create a new by clicking

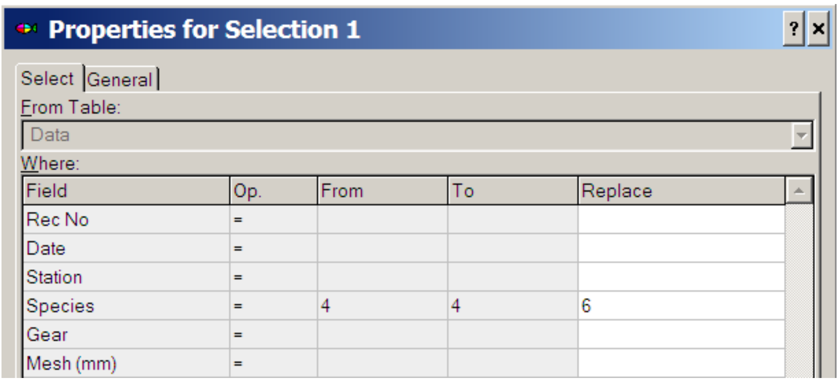


Find and replace any specific records

Press [Ctrl + H] or right click on table ►**Select and**.. ►**Replace**



The connected query will now be extended with a **Replace** column where you can specify the updates you want to make, e.g. changing species Id’s or any other values.



From Table: Data

Where:

Field	Op.	From	To	Replace
Rec No	=			
Date	=			
Station	=			
Species	=	4	4	6
Gear	=			
Mesh (mm)	=			

You can use [expressions](#) in the replace column by writing these or using the [expression builder](#)

Check codes (Id's)

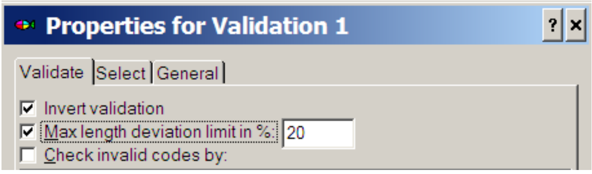
See [Queries/Validation of the Id's used in Data](#)

Check observed length versus weight

If you have individual fish measurements, you should check if the entered length and weight data of each fish and species are corresponding. This is best done by first running the [length-weight relationship](#). If there are inconsistencies between observed (recorded) length and weight then the r^2 in the length-weight relationship will be relatively low (r^2 is acceptable within roughly 0.95-1.00) and the diagram of length versus weight will show many [outliers](#).

First you should try to get the best set of estimated length-weight coefficients of a given species. This is done iteratively:

1. Run the length-weight Analysis (see [Analysis/Length-weight relationship](#))
2. Transfer coefficients to species table (see [Analysis/Length-weight relationship](#))
3. Connect a [Validation query](#) and check the '**Max length deviation in %**'
4. Invert the query (see [Queries](#)). This will return all records with less than the specified deviation

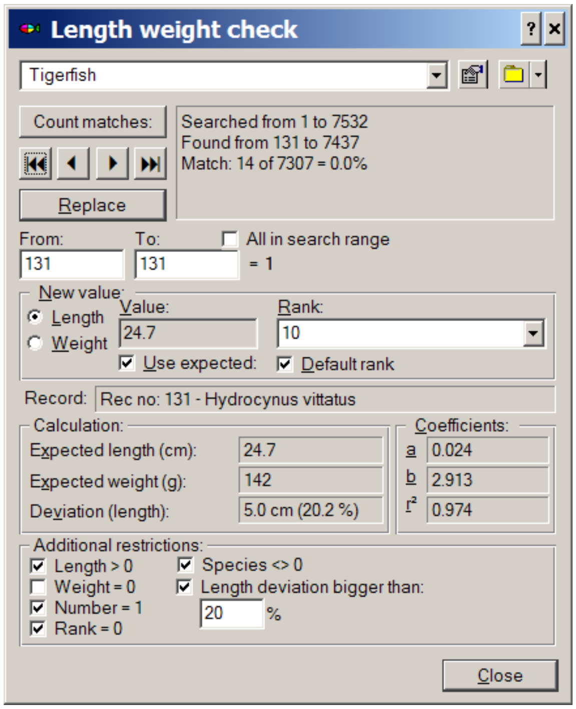


☑ Invert validation
☑ Max length deviation limit in %: 20
☐ Check invalid codes by:

5. Proceed with step 1 and 2 until there are no more outliers or r^2 is acceptable.

Now with a good set of length-weight coefficients in the species table you are able to go back to the original records to check the outliers for punching, or recording mistakes.

Right click in data table ►**Select** and.. ► **Check length/weight**:



Tigerfish

Searched from 1 to 7532
Found from 131 to 7437
Match: 14 of 7307 = 0.0%

From: 131
To: 131
☐ All in search range
= 1

New value:
● Length
○ Weight
Value: 24.7
Rank: 10
☑ Use expected:
☑ Default rank

Record: Rec no: 131 - Hydrocynus vittatus

Calculation:
Expected length (cm): 24.7
Expected weight (g): 142
Deviation (length): 5.0 cm (20.2 %)

Coefficients:
a 0.024
b 2.913
r^2 0.974

Additional restrictions:
☑ Length > 0
☐ Weight = 0
☑ Number = 1
☑ Rank = 0
☑ Species <> 0
☑ Length deviation bigger than:
20 %

Here you can now find each ‘outlier’ successively and decide whether the records need correction (e.g. punching mistake). Note that higher deviations are acceptable, and expected, between observed and expected weight as weight is a much more variable parameter than length.

- You can choose whether the length or the weight value should be updated under **New value**
- You can replace using the expected (calculated) value, check **Use expected**, or check this off and enter any new value manually
- Choose the validity code you want for the Rank field (see below) or use default.
- When ready press **[Replace]**

Validity codes in Rank field.

Punching mistakes should be relatively easy to disclose and correct with the above procedure, but some times it may not be possible to find the source of an error, particularly if the mistake is done in the original data sheets or during the sampling itself. In such case you can either decide to delete the record altogether or to change it the most sensible way. Remember that if you do so you must (!) enter a correction/validity code in the Rank field, see Table below.

Rank codes (suggested)

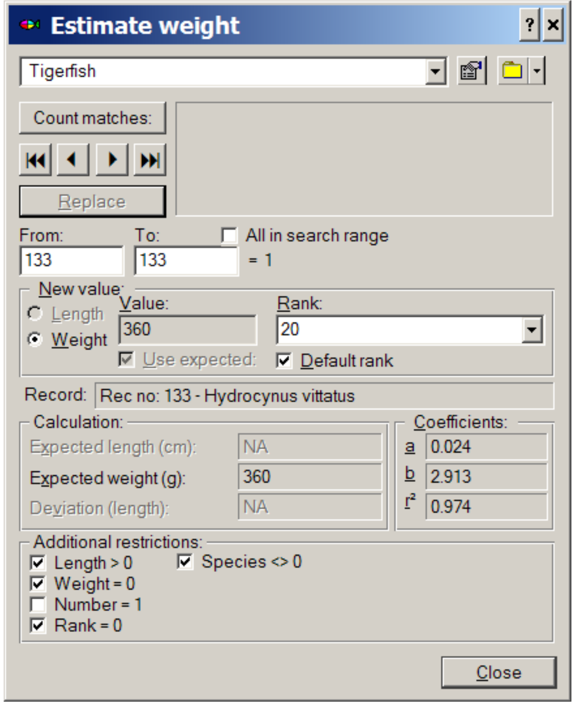
Id	Comments
10	when length is assumed wrong and corrected
20	when weight is estimated from weight-length relationship
30	when weight is estimated from mean weight in mesh size of species
40	when weight is measured as mean weight in catch of species

When the Rank field contains values different from 0 the record will not be used in certain of the calculation procedures, such as length-weight relationship, mean weight, condition factor etc. If deleting a possible mistaken weight from a record, but maintaining the length, the weight can be re-estimated and entered by using the Estimate weights option (see next section). This option should only be used at the very end of a file cleaning, when all the length-weight coefficients entered in the species table are as precise as possible!!

Checking and cleaning records is a tedious job, but very important if you want good and reliable results of your data analysis. Remember that in all computer programs 'rubbish in is equal to rubbish out'!

Estimate missing weights

If the records have length measurements, but weight measurements are missing (or wrongly measured) and set to 0, and the length-weight coefficients of the species are entered into the Species table, then weights can be automatically estimated and inserted in the data base by the module **Estimate weight**: Right click on data ►**Select and..** ►**Estimate weight**



Tigerfish

From: 133 To: 133 ☐ All in search range = 1

New value: ○ Length ● Weight

Value: 360

Rank: 20

☑ Use expected: ☑ Default rank

Record: Rec no: 133 - Hydrocynus vittatus

Calculation:

Expected length (cm): NA

Expected weight (g): 360

Deviation (length): NA

Coefficients:

a 0.024

b 2.913

r² 0.974

Additional restrictions:

☑ Length > 0 ☑ Species <> 0

☑ Weight = 0

☐ Number = 1

☑ Rank = 0

In such case the Rank value is set by default to 20 or user defined.

Pasgear 2 calculates missing weights in two different ways:

1) if the individual lengths are recorded in mm and each record contains one fish only then the weight is estimated from the normal length-weight relationship

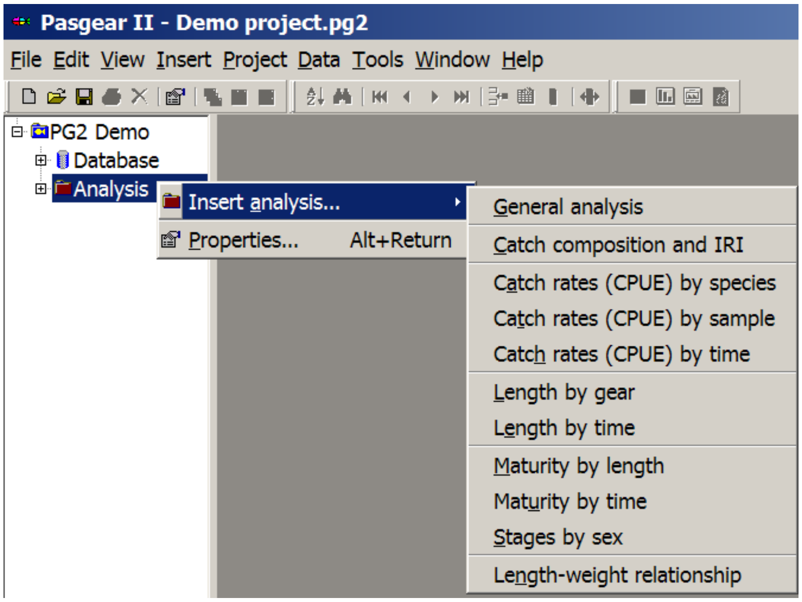
$$\text{weight} = a \cdot \text{lenght}^{b} \quad \text{(Where weight is in grams and length in cm)}$$

2) if the individual lengths are recorded in whole cm and the record represents a length frequency (see project properties/definitions), then the weight is estimated as the mean weight in the length class i (L_1-L_2) and calculated from Beyer (1987)

$$\overline{\text{weight}}_{i} = \frac{1}{(L_2 - L_1)} \cdot \frac{a}{(b+1)} \cdot (L_2^{b+1} - L_1^{b+1})$$

where *a* and *b* are the length-weight coefficients, L_1 is lower interval limit (which should be the recorded length in the data base) and L_2 is the higher interval limit, so L_2-L_1 is the length class interval in cm (default interval is 1 cm but other whole cm length class intervals can be specified).

Analysis



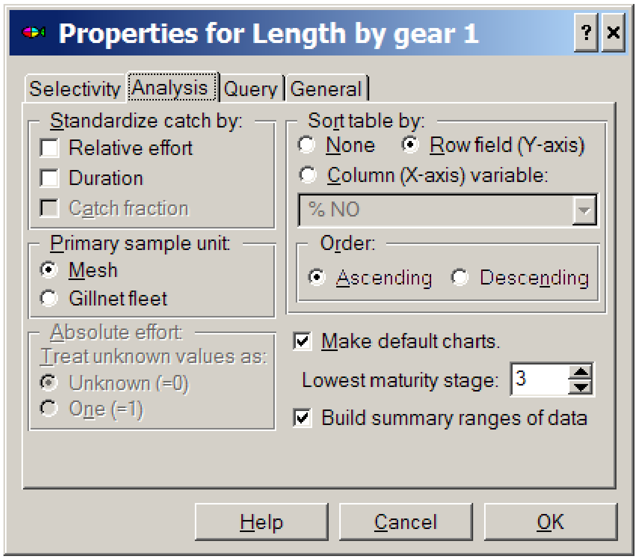
To make an analysis either:

- Choose Insert on the main menu or
- Right click on Analysis in project tree and insert one of the following analysis types:
  1. General analysis
  2. Catch composition and IRI

3. Catch rates (CPUE) by species
4. Catch rates (CPUE) by sample
5. Catch rates (CPUE) by time
6. Length by gear
7. Length by time
8. Maturity by time
9. Stages by sex
10. Length-Weight relationship

A **general analysis** is built from scratch (see general analysis below), while the remaining 9 analyses are predefined setups (macros), but which all can be modified.

Analysis properties


Properties for Length by gear 1

Standardize catch by:
☐ Relative effort
☐ Duration
☐ Catch fraction

Primary sample unit:
● Mesh
○ Gillnet fleet

Absolute effort:
Treat unknown values as:
● Unknown (=0)
○ One (=1)

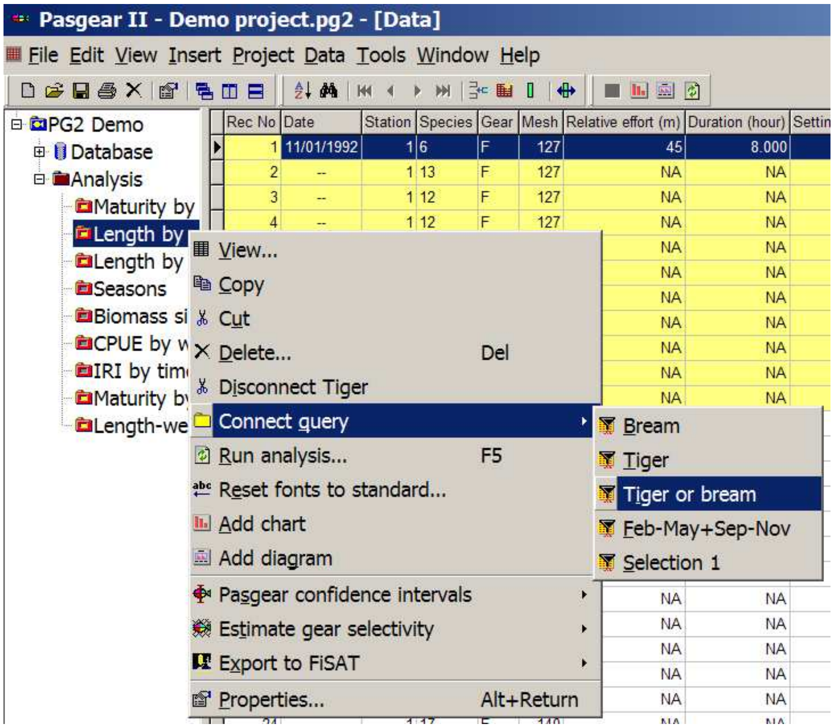
Sort table by:
○ None ● Row field (Y-axis)
○ Column (X-axis) variable:
% NO

Order:
● Ascending ○ Descending

☑ Make default charts.
Lowest maturity stage: 3
☑ Build summary ranges of data

Under the properties of an analysis, the analysis can be

- Named and given comments (Tab: Properties/General)
- Connected with a query (Tab: Properties/Query or right click and Connect and already existing query)
- Standardized to Standard effort
- If absolute effort is not default (one per sample) then there is an option for treating unknown values
- Result table can be sorted by row field groups or column variables
- Lowest maturity stage can be set for maturity analyses
- ☑ Make default charts. should be checked if you want the predefined graphs (add hoc charts) to be produced
- ☑ Build summary ranges of data should be checked if you want a summary of the ranges found in the data table
- If the analysis is a length grouped analysis it can be connected to a gear selectivity file (Tab: Properties/Selectivity)

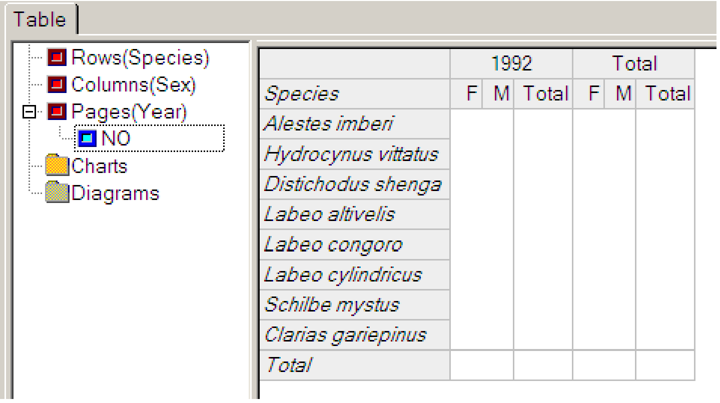


All analyses can also be copied or cut and pasted (within or among pasgear programs), and viewed or deleted.

According to the grouping setup and the chosen variables the results can also be exported to

- Estimate confidence intervals (Pasgear tool)
- Estimate gear selectivity (Pasgear tool)
- FiSAT (FAO ICLARM Stock Assessment Tools, Gayanilo et al. 1996)

General analysis

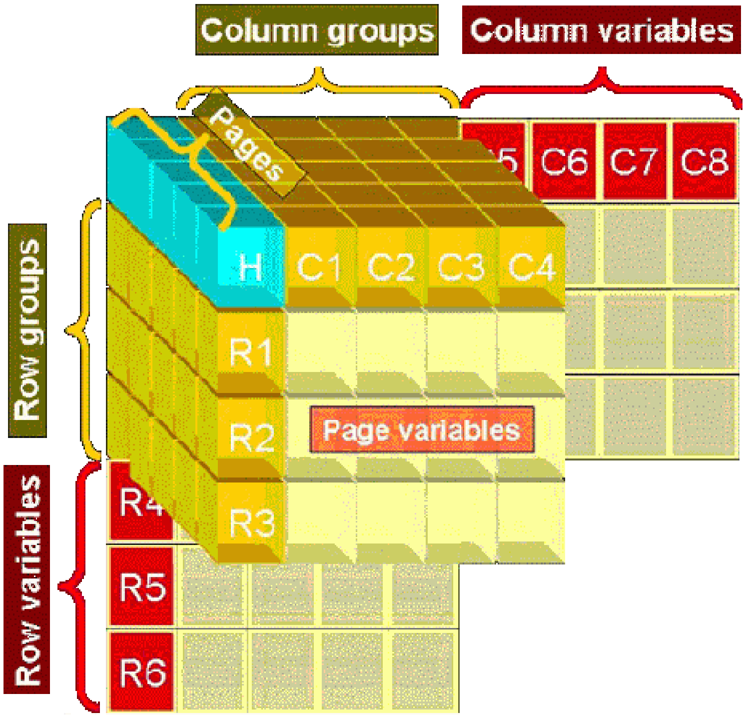


	1992			Total		
Species	F	M	Total	F	M	Total
Alestes imberi						
Hydrocynus vittatus						
Distichodus shenga						
Labeo altivelis						
Labeo congoro						
Labeo cylindricus						
Schilbe mystus						
Clarias gariepinus						
Total						

The general analysis (and all predefined macros) consists of 3 grouping dimension (rows, columns and pages, like the MS Excel pivot table), and a number of chosen variables or statistics can be added to each dimension by Add variable. By pressing the F5 button or refresh, the analysis spools over the data table and will aggregate groups (se below) and calculated variables up in a report table. A matrix table (X,Y) can be created by grouping rows and columns and adding one page-variable. An array of matrices can be made by grouping by pages (e.g. if you want the result of an analysis repeated for each year in the data then group pages by Year).

Grouping dimensions

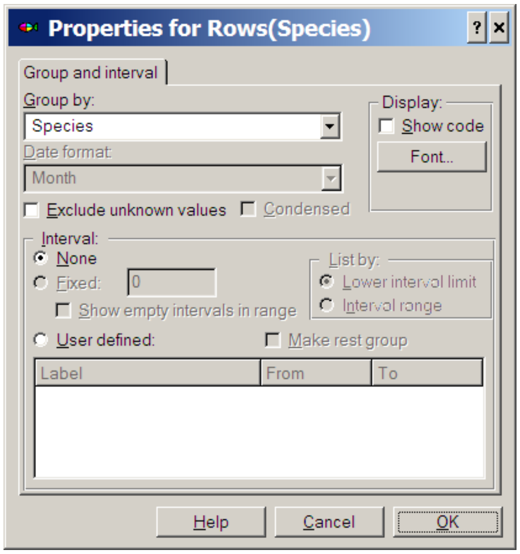
All data fields can be grouped by either Rows (Y), Columns (X) or Pages (Z), and any number of variables can be added to rows columns and pages. See for example the 'Length by time' analysis, where all three dimensions are used:



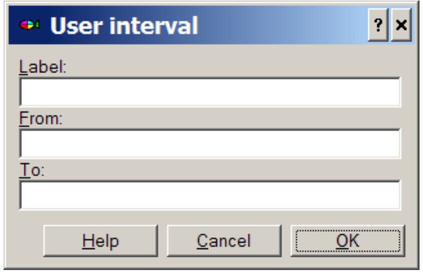
Technically, the page dimension is the 3D dimension of a cube, but for the display each page will be sliced off the cube and located next to each other (as 'pages'). If you are using the page dimension, the last page will always be the total off all the other pages summed.

To group in one of the dimensions, then right click i.e. Rows and choose Properties at the menu:

The property page for grouping dimension (in this case grouped by species):



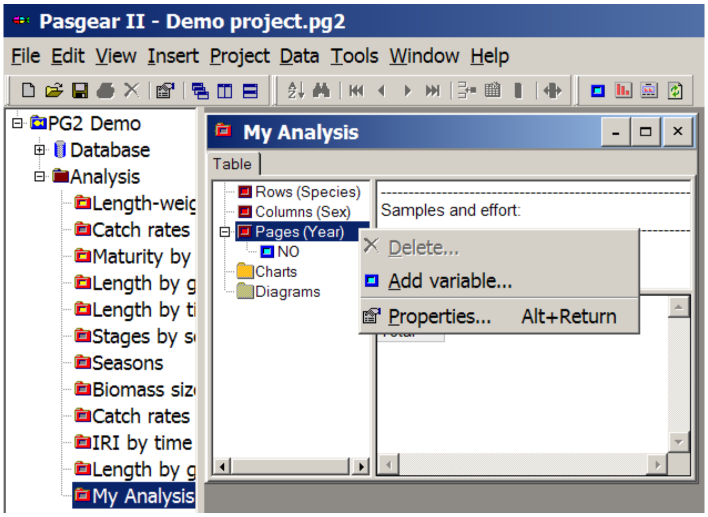
The grouping can be done by each individual element found (Interval = None), or by user defined fixed intervals of any positive integer >0. Groups can also be user defined by checking (●) User defined and then right click in the pane and choose [add]:



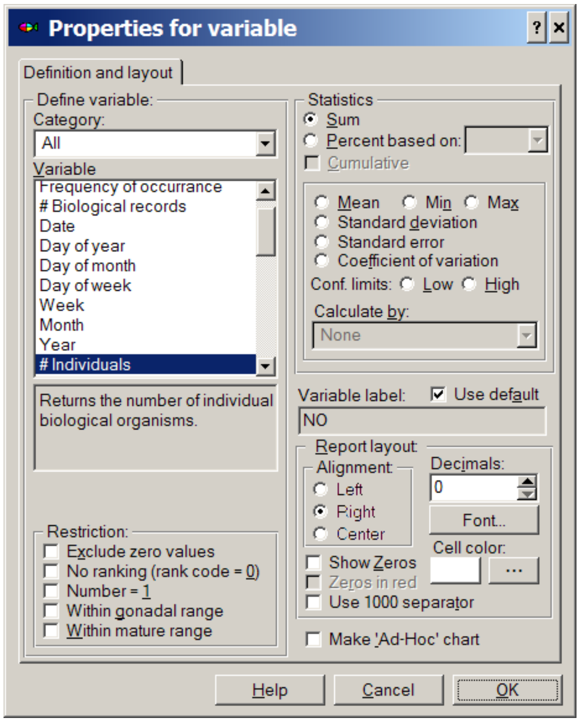
You can then give the group a label and select ranges From … to …

You can also define your own groupings by adding a calculated field in the data table and use an expression to make a new set of groups. See example.

Variables



Variables can be added to rows, columns or matrix (pages) by right clicking on the dimension and choose Add variable, or click on the ◘ speed button on the main menu . The property page for a variable is rather big and can best be studied by examining the ‘hints’ on the dialog or the status bar.



Category: All

Variable: # Individuals

Returns the number of individual biological organisms.

Statistics: ● Sum

Variable label: ☑ Use default

NO

Alignment: ● Right

Decimals: 0

The list of available variables can be categorised and listed (open the combo box Category) by:

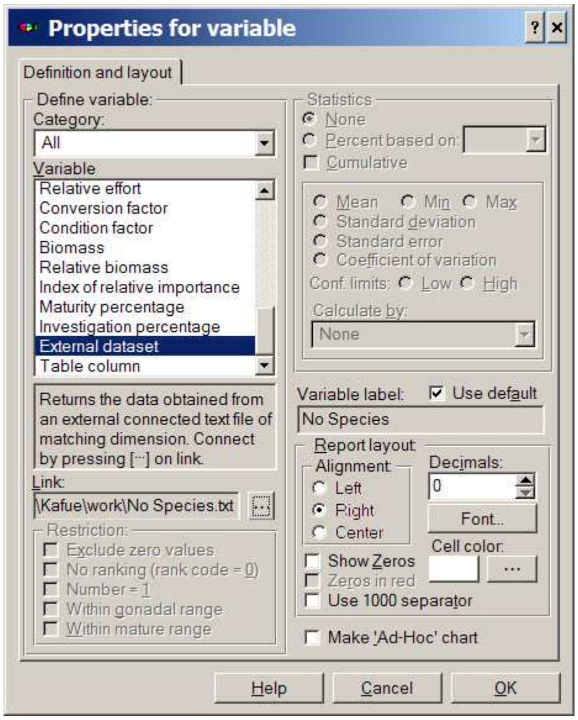
- All
- Basic (number, length, weight, frequency of occurrence etc. taken from the Data table)
- Effort (absolute, relative, duration or conversion factor for standardizing samples)
- Date (various ways of grouping or expressing dates)
- Derived (special such as Index of relative importance, Fulton’s condition factor, percent mature, Shannon’s diversity index etc.)
- External (reads the variables from an external file of rows and columns)

The info pane under the combo box with listed variables will tell what each chosen variable will return.

External Data sets

This is a powerful feature that enables the user to import external data to the report table. You can also export a report table to e.g Excel, do your own calculations and then re-import the new variable to the table or for use in the plots.

Variables can be imported from outside of Pasgear by connecting an external dataset from a file in text format. Choose the variable ‘External dataset’ and link the file by pressing […] on the Link edit line:



Define variable:

Category: All

Variable:
- Relative effort
- Conversion factor
- Condition factor
- Biomass
- Relative biomass
- Index of relative importance
- Maturity percentage
- Investigation percentage
- External dataset
- Table column

Returns the data obtained from an external connected text file of matching dimension. Connect by pressing [...] on link.

Link: \Kafue\work\No Species.txt

Restriction:
- ☐ Exclude zero values
- ☐ No ranking (rank code = 0)
- ☐ Number = 1
- ☐ Within gonadal range
- ☐ Within mature range

Statistics
- ● None
- ○ Percent based on:
- ☐ Cumulative
- ○ Mean ○ Min ○ Max
- ○ Standard deviation
- ○ Standard error
- ○ Coefficient of variation
- Conf. limits: ○ Low ○ High
- Calculate by: None

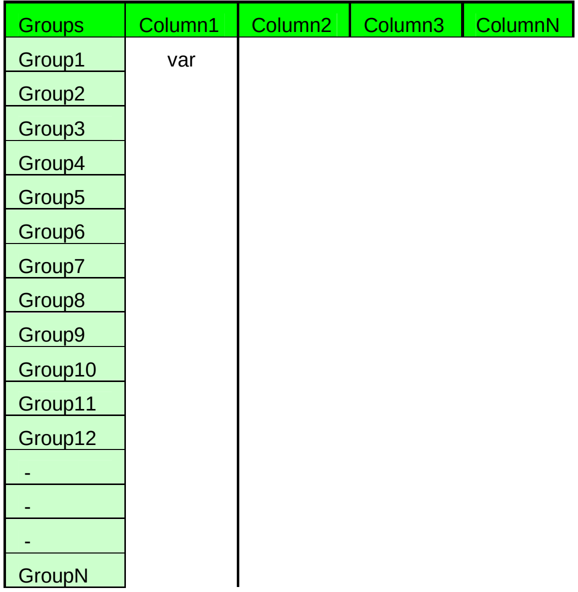
Variable label: ☑ Use default

No Species

Report layout
- Alignment: ○ Left ● Right ○ Center
- Decimals: 0
- Cell color:
- ☐ Show Zeros
- ☐ Zeros in red
- ☐ Use 1000 separator

☐ Make 'Ad-Hoc' chart

External data sets have the following format:



Groups	Column1	Column2	Column3	ColumnN
Group1	var			
Group2				
Group3				
Group4				
Group5				
Group6				
Group7				
Group8				
Group9				
Group10				
Group11				
Group12				
-				
-				
-				
GroupN				

If the variable is a row or column variable only the first column in the external set will be used (read). However if the variable is a page variable (matrix) all rows and all columns will be added to the table. Group names (rows) and column headers with same value (or name) as found in the data table will be grouped with these, whereas if the external data set contains groups or column headers that are new these will be added as new groups to the table.

Variable restrictions

Depending on the chosen variable they can be directly restricted (without connecting a query) when reading from the Data such as:

- Exclude zero values
- No ranking (only choose records with rank code = 0)

- Number = 1 (only choose records where number field = 1)
- Within gonadal range (as defined in Project properties)
- Within mature range (only records with gonadal stages defined as mature)

Variable statistics

Depending on the chosen variable they can be expressed in various statistical terms such as

- Count
- Sum
- Percent (based on totals in either table, page, rows or columns)
- Min or Max value observed
- Mean and standard spread statistics (based on various chosen denominators in ‘Calculate by’)
  - Standard deviation
  - Standard error of mean
  - Coefficient of variation
  - 95% confidence intervals

Standard deviation

In case where the estimated mean (e.g. CPUE) is the ratio of two variables, then SD’s are calculated from the Taylor series approximation by the following formula (Cochran 1977, Krebs 1989)

$$SD = \frac{1}{\bar{y}}\sqrt{\frac{\sum x^2 - 2\hat{R}\sum xy + \hat{R}^2\sum y^2}{n-1}} \quad \text{where} \quad \hat{R} = \frac{\bar{x}}{\bar{y}}$$

However, if the denominator (*y*) is equal to the sample size (*n*) then

- Each $y = 1$
- $\sum y^2 = y = n$
- $\sum xy = \sum x$

And the above ratio formula reduces to the normal standard deviation formula of

$$SD = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$$

Standard error is calculated as $SE = \frac{SD}{\sqrt{N}}$

Coefficient of variation is calculated as $CV = \frac{SD}{\bar{x}} \cdot 100\%$

95% Confidence intervals are calculated as $CI_{95\%} = \bar{x} \pm t_{n-1} \cdot SE$ where *t* is from Students t-table with (n-1) degrees of freedom and α = 0.025.

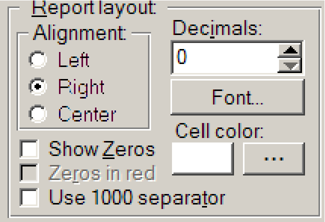
Variable layout

The chosen variable will always be given a default label as header in the Analysis table. However, if checking off ‘Use default’ the user can specify the given label.

Variables can be given a font, aligned, given a specific number of decimals, or optimally show zero values. The cell color in the Analysis table can also be given a specific background color in the report table.

Conditional formatting

If you press the […] buton under Report layout

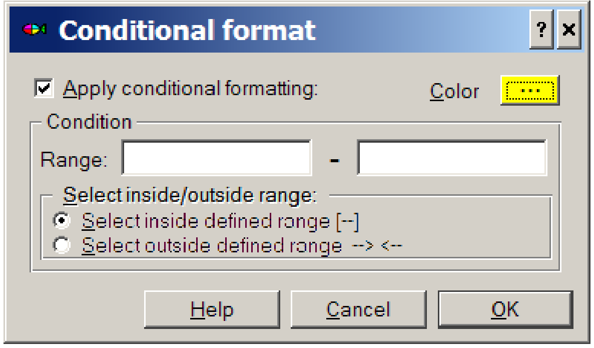


Alignment: ○ Left ● Right ○ Center

Decimals: 0

☐ Show Zeros
☐ Zeros in red
☐ Use 1000 separator

You can define a selected value range (inside or outside) of the variable which will be given a specified color format in the report table



☑ Apply conditional formatting:

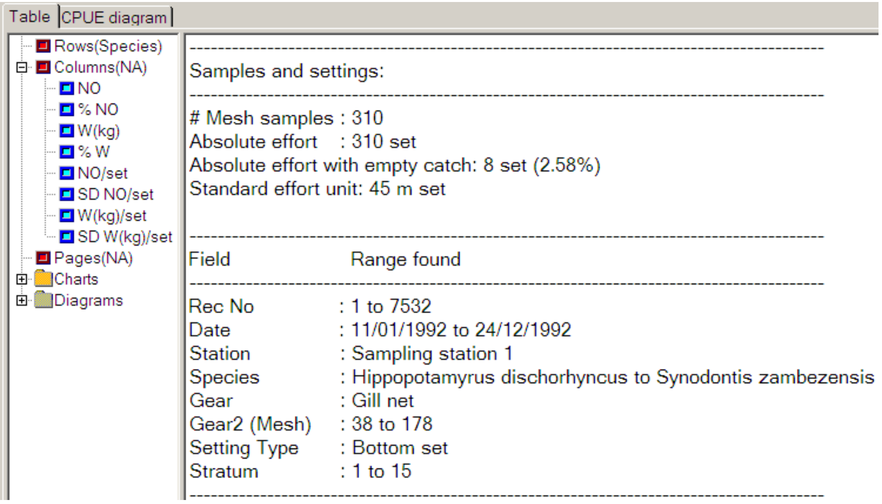
Range: -

Select inside/outside range:
● Select inside defined range [--]
○ Select outside defined range --> <--

Making 'Ad hoc' chart

When checking this option Pasgear will automatically set up a diagram with a plot of the chosen variable. This plot can be modified (see Diagrams and Charts).

Analysis results



```
--------------------------------------------------------------------------------
Samples and settings:
--------------------------------------------------------------------------------
# Mesh samples : 310
Absolute effort   : 310 set
Absolute effort with empty catch: 8 set (2.58%)
Standard effort unit: 45 m set

--------------------------------------------------------------------------------
Field              Range found
--------------------------------------------------------------------------------
Rec No             : 1 to 7532
Date               : 11/01/1992 to 24/12/1992
Station            : Sampling station 1
Species            : Hippopotamyrus dischorhyncus to Synodontis zambezensis
Gear               : Gill net
Gear2 (Mesh)       : 38 to 178
Setting Type       : Bottom set
Stratum            : 1 to 15
--------------------------------------------------------------------------------
```

When the analysis has been built by groupings and variables, then press F5, or the refresh icon , to run it.

The result will be shown in the ‘Table’ tab, which has 2 panes:

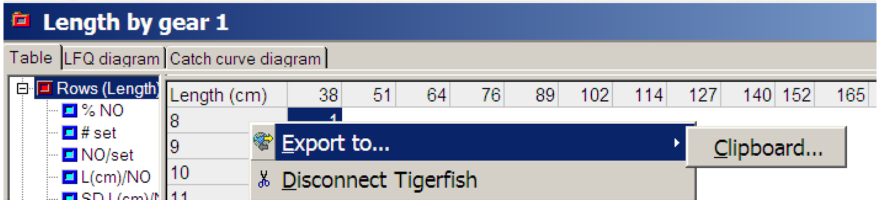
The upper is a summary pane giving:

- The number of primary sample units (Mesh samples) and absolute effort, and if the Analysis is standardized then also the Standard effort unit.
- The ranges found over the major physical fields (Records, date, station, etc.) and the species. If some of these ranges have been restricted by a connected query this will be displayed too.

The lower is the results table with the chosen variables and statistics built, layout and sorted into rows, columns (and pages). By right click in this pane you have various options such as e.g. Export the table to clipboard (can e.g. be pasted directly into MS Excel, Word etc).

Export results

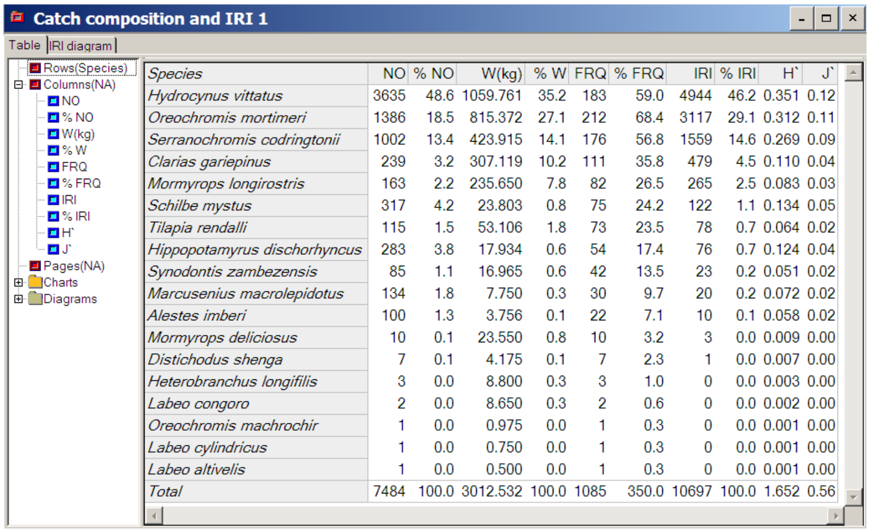
You can export the Pasgear report table via the clipboard by right click on the Table **►Export to… ►Clipboard…**



The content of the clipboard can then be pasted in other applications (e.g. Excel, Word etc.).

Predefined analyses

Catch composition and IRI



Species	NO	% NO	W(kg)	% W	FRQ	% FRQ	IRI	% IRI	H`	J`
Hydrocynus vittatus	3635	48.6	1059.761	35.2	183	59.0	4944	46.2	0.351	0.12
Oreochromis mortimeri	1386	18.5	815.372	27.1	212	68.4	3117	29.1	0.312	0.11
Serranochromis codringtonii	1002	13.4	423.915	14.1	176	56.8	1559	14.6	0.269	0.09
Clarias gariepinus	239	3.2	307.119	10.2	111	35.8	479	4.5	0.110	0.04
Mormyrops longirostris	163	2.2	235.650	7.8	82	26.5	265	2.5	0.083	0.03
Schilbe mystus	317	4.2	23.803	0.8	75	24.2	122	1.1	0.134	0.05
Tilapia rendalli	115	1.5	53.106	1.8	73	23.5	78	0.7	0.064	0.02
Hippopotamyrus dischorhyncus	283	3.8	17.934	0.6	54	17.4	76	0.7	0.124	0.04
Synodontis zambezensis	85	1.1	16.965	0.6	42	13.5	23	0.2	0.051	0.02
Marcusenius macrolepidotus	134	1.8	7.750	0.3	30	9.7	20	0.2	0.072	0.02
Alestes imberi	100	1.3	3.756	0.1	22	7.1	10	0.1	0.058	0.02
Mormyrops deliciosus	10	0.1	23.550	0.8	10	3.2	3	0.0	0.009	0.00
Distichodus shenga	7	0.1	4.175	0.1	7	2.3	1	0.0	0.007	0.00
Heterobranchus longifilis	3	0.0	8.800	0.3	3	1.0	0	0.0	0.003	0.00
Labeo congoro	2	0.0	8.650	0.3	2	0.6	0	0.0	0.002	0.00
Oreochromis machrochir	1	0.0	0.975	0.0	1	0.3	0	0.0	0.001	0.00
Labeo cylindricus	1	0.0	0.750	0.0	1	0.3	0	0.0	0.001	0.00
Labeo altivelis	1	0.0	0.500	0.0	1	0.3	0	0.0	0.001	0.00
Total	7484	100.0	3012.532	100.0	1085	350.0	10697	100.0	1.652	0.56

This analysis gives total catch composition in numbers and weight (kg), as well as frequency of occurrence (FRQ) in the fleet or mesh size settings (i.e. whether the species was present or not irrespective of the abundance). Each of these values is also given in percentage of total (not that the percentage frequency of occurrence %FRQ does not add to 100 as the total is the total number of settings). As a measure of relative abundance or commonness of each species (*i*) in the catch composition an index of relative importance (%*IRI*, Kolding 1989) is used

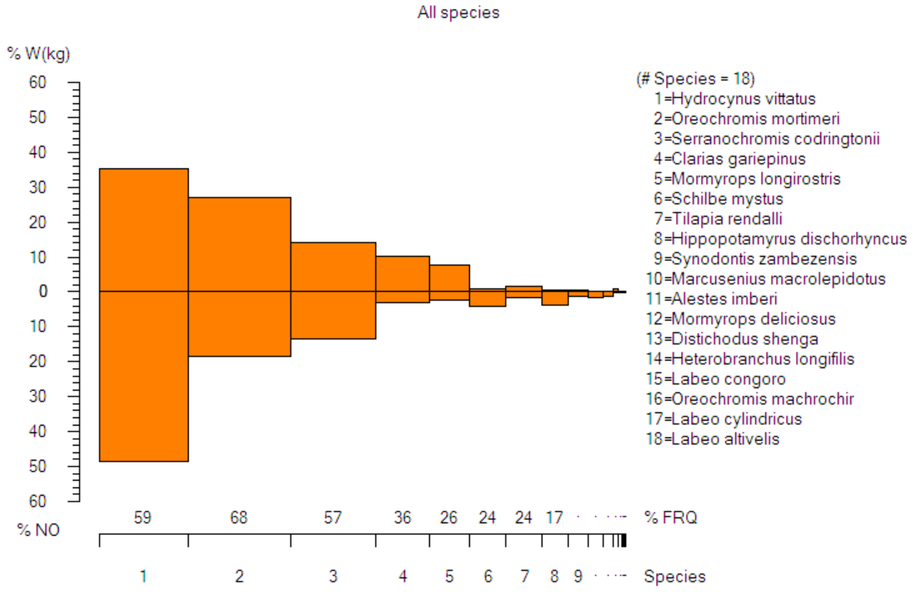
$$\%IRI_i = \frac{(\%W_i + \%N_i)\cdot\%F_i}{\sum_{j=1}^{S}(\%W_j + \%N_j)\cdot\%F_j}\cdot 100$$

where $\%W_i$ and $\%N_i$ is percentage weight and number of each species of total catch, $\%F_i$ is percentage frequency of occurrence of each species in total number of settings, and S is total number of species. This index was originally graphical and combines and show simultaneously the relative numeric abundance (N), the average size (W) and the commonness (F) of a species, IRI = (%N + %W)* %F (Pinkas *et al.* 1971, see also Caddy & Sharp 1986), displayed as a rectangle (see figure). %*IRI* gives the relative area of this rectangle in percentage to all the other species present. After running the programme there is an option for illustrating the IRI graphically instead of in percentages (see figure).

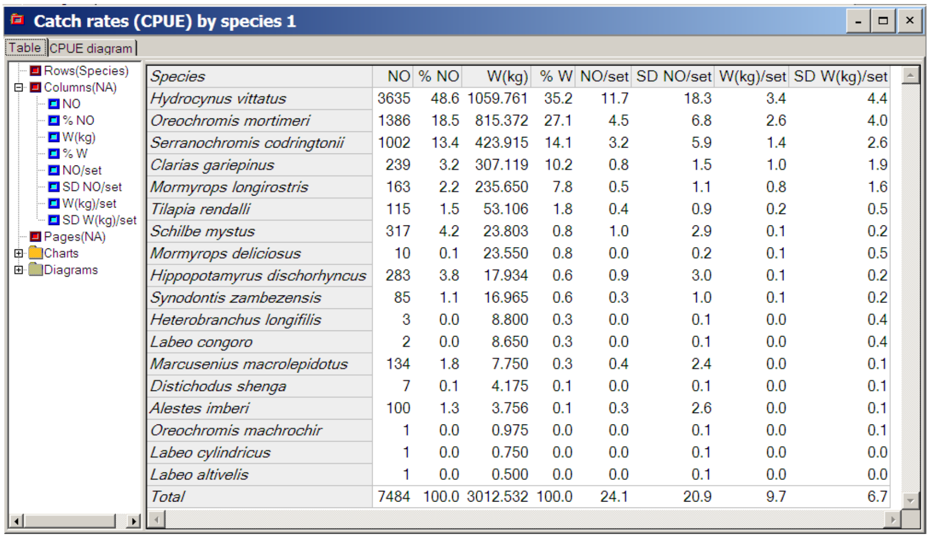
Shannon's diversity index. The IRI programme also calculates Shannon's diversity index (H') and the relative evenness (J') defined as:

$$H' = -\sum_{i=1}^{S} P_i \cdot Ln(P_i) \text{ and } J' = H'/H_{\max} \text{ where } H_{\max} = Ln(S)$$

where P_i is the relative abundance, i.e. the number of individuals for each species divided by the total number of individuals for all species (S) in each sample (Begon *et al.* 1990, p. 617).



Catch rates (CPUE) by species



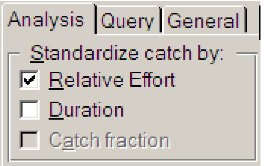
Species	NO	% NO	W(kg)	% W	NO/set	SD NO/set	W(kg)/set	SD W(kg)/set
Hydrocynus vittatus	3635	48.6	1059.761	35.2	11.7	18.3	3.4	4.4
Oreochromis mortimeri	1386	18.5	815.372	27.1	4.5	6.8	2.6	4.0
Serranochromis codringtonii	1002	13.4	423.915	14.1	3.2	5.9	1.4	2.6
Clarias gariepinus	239	3.2	307.119	10.2	0.8	1.5	1.0	1.9
Mormyrops longirostris	163	2.2	235.650	7.8	0.5	1.1	0.8	1.6
Tilapia rendalli	115	1.5	53.106	1.8	0.4	0.9	0.2	0.5
Schilbe mystus	317	4.2	23.803	0.8	1.0	2.9	0.1	0.2
Mormyrops deliciosus	10	0.1	23.550	0.8	0.0	0.2	0.1	0.5
Hippopotamyrus dischorhyncus	283	3.8	17.934	0.6	0.9	3.0	0.1	0.2
Synodontis zambezensis	85	1.1	16.965	0.6	0.3	1.0	0.1	0.2
Heterobranchus longifilis	3	0.0	8.800	0.3	0.0	0.1	0.0	0.4
Labeo congoro	2	0.0	8.650	0.3	0.0	0.1	0.0	0.4
Marcusenius macrolepidotus	134	1.8	7.750	0.3	0.4	2.4	0.0	0.1
Distichodus shenga	7	0.1	4.175	0.1	0.0	0.1	0.0	0.1
Alestes imberi	100	1.3	3.756	0.1	0.3	2.6	0.0	0.1
Oreochromis machrochir	1	0.0	0.975	0.0	0.0	0.1	0.0	0.1
Labeo cylindricus	1	0.0	0.750	0.0	0.0	0.1	0.0	0.0
Labeo altivelis	1	0.0	0.500	0.0	0.0	0.1	0.0	0.0
Total	7484	100.0	3012.532	100.0	24.1	20.9	9.7	6.7

Catch per unit effort (CPUE) in PASGEAR is calculated as:

$$CPUE = \frac{1}{y}\sum_{i=1}^{n} W_i \cdot \frac{SU}{U_i} \cdot \frac{ST}{T_i} \text{ ,} \qquad \text{where}$$

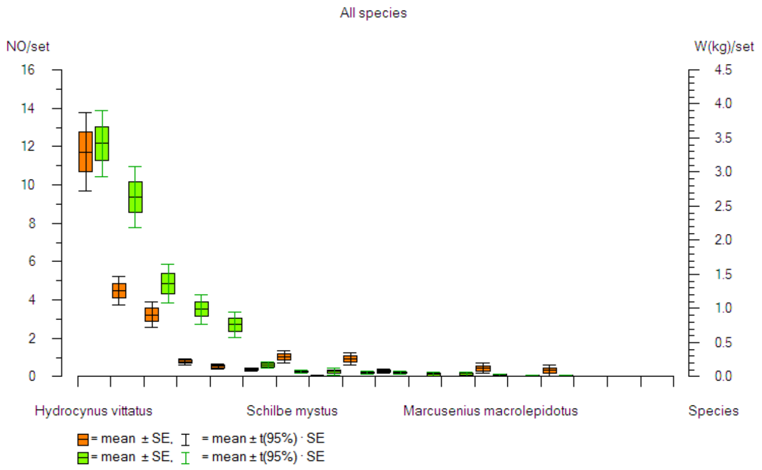
y = absolute effort, e.g. number of net panel (or fleet) settings and
n = number of samples (NB, if effort is not a variable then $y = n$).
W_i = catch (in weight or numbers) in set$_i$ or sample$_i$,
SU = standard relative effort unit (size) of a net panel (defined in the data table properties/Effort mode),
U_i = actual relative effort unit (size) of net$_i$ (this can be given in the Relative **effort** field in the Data Table
and/or defined in the Gear Table, see Effort definition and sample raising modes)
ST = standard time unit (hours or minutes) of a setting (defined in the data table properties/Effort mode),
T_i = actual time unit of setting$_i$ (this can be given in the **duration** field in the Data Table).

Standardization of the CPUE, i.e. expressing the result per standard effort units (gear size or standard time), is optionally achieved by activating one or both of the terms (SU/U_i and ST/T_i) in the above equation when setting the Analysis properties under Analysis/Standardize catch by:



Standardize catch by:
☑ Relative Effort
☐ Duration
☐ Catch fraction

For changing the standard units see Effort definition and sample raising modes.

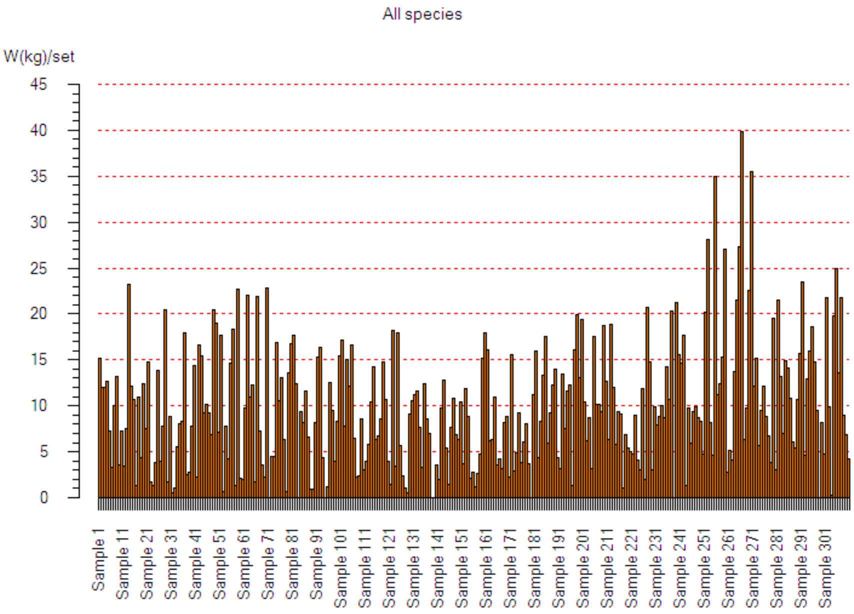


Catch rates (CPUE) by sample

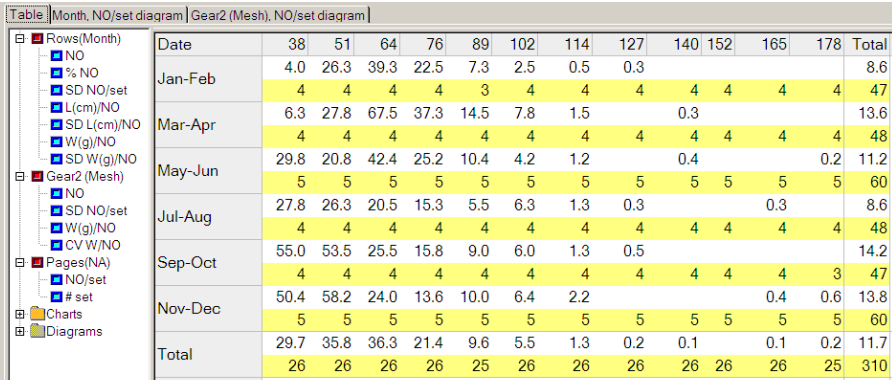
Sample	Hippopotamyrus dischorhyncus	Marcusenius macrolepidotus	Mormyrops
Sample 1	0.0	0.0	
Sample 2	0.0	0.0	
Sample 3	0.0	0.0	
Sample 4	0.0	0.0	
Sample 5	0.0	0.0	
Sample 6	0.0	0.0	
Sample 7	0.0	0.0	
Sample 8	0.2	0.0	
Sample 9	0.0	0.0	
Sample 10	0.7	0.5	
Sample 11	0.3	0.4	
Sample 12	0.0	0.0	
Sample 13	1.6	0.3	
Sample 14	0.3	0.0	

This macro gives CPUE by each individual sample (rows) and can be variously grouped (columns). The main purpose is for exporting the results to the tool "Estimate confidence intervals" in order to see the frequency distribution of the catch rates, and/or for various ways of calculating alternative confidence intervals. There are 4 options for calculating confidence intervals:

1) Standard parametric based on the arithmetic sample mean
2) The Pennington estimator based on the log-normal Delta distribution
3) Bootstrap on the arithmetic sample mean
4) Bootstrap on the Pennington estimator



Catch rates (CPUE) by time



Date	38	51	64	76	89	102	114	127	140	152	165	178	Total
Jan-Feb	4.0	26.3	39.3	22.5	7.3	2.5	0.5	0.3					8.6
	4	4	4	4	3	4	4	4	4	4	4	4	47
Mar-Apr	6.3	27.8	67.5	37.3	14.5	7.8	1.5		0.3				13.6
	4	4	4	4	4	4	4	4	4	4	4	4	48
May-Jun	29.8	20.8	42.4	25.2	10.4	4.2	1.2		0.4			0.2	11.2
	5	5	5	5	5	5	5	5	5	5	5	5	60
Jul-Aug	27.8	26.3	20.5	15.3	5.5	6.3	1.3	0.3			0.3		8.6
	4	4	4	4	4	4	4	4	4	4	4	4	48
Sep-Oct	55.0	53.5	25.5	15.8	9.0	6.0	1.3	0.5					14.2
	4	4	4	4	4	4	4	4	4	4	4	3	47
Nov-Dec	50.4	58.2	24.0	13.6	10.0	6.4	2.2				0.4	0.6	13.8
	5	5	5	5	5	5	5	5	5	5	5	5	60
Total	29.7	35.8	36.3	21.4	9.6	5.5	1.3	0.2	0.1		0.1	0.2	11.7
	26	26	26	26	25	26	26	26	26	26	26	25	310

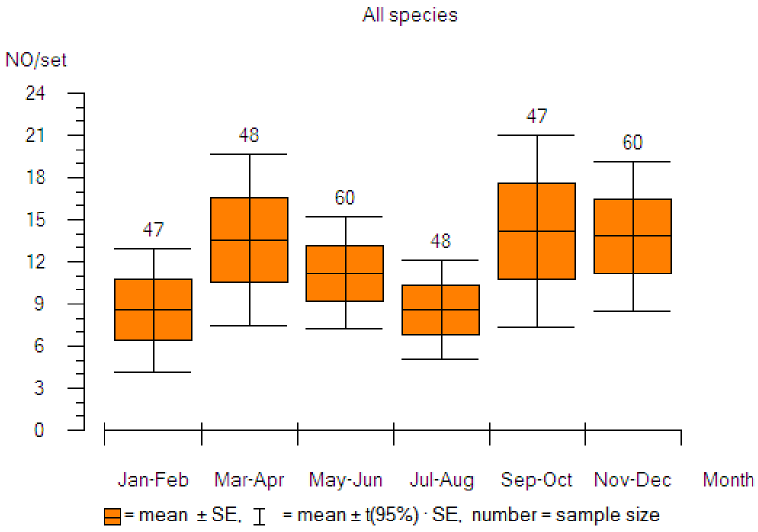
This macro is used for calculating and exploring changes in standardized CPUE by different time intervals and within these time intervals by mesh-size or any other groupings. It can also be used to explore changes in mean weight, length, or condition factor against the other variables.

In the diagrams the values can optionally be seen with 95% confidence limits. The confidence limits are calculated from

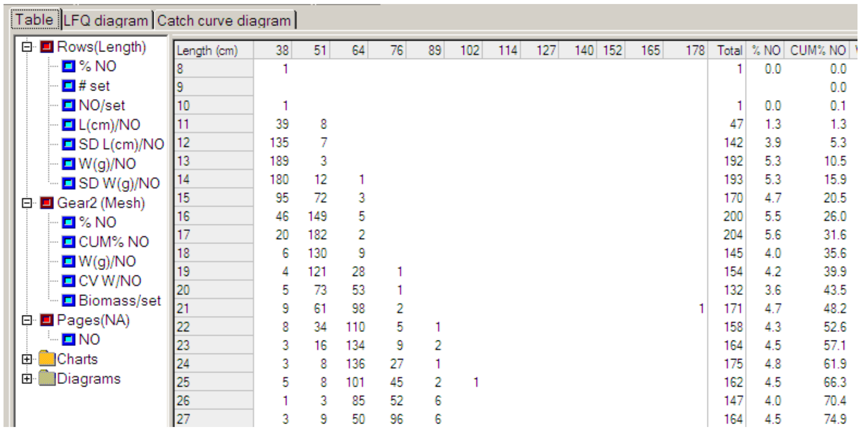
$$\overline{CPUE} \pm t_{(n-1)} \cdot SE(CPUE)$$

where *t* is Students t-value (at the 95% confidence limit) and *n* = sample size (number of observations). The standard error (SE) is the SD divided by the square root of *n*, and SD's are calculated the standard way, or from the Taylor series approximation depending on effort is a variable or not. If *n* is small then it may be better to calculate confidence limits based on bootstrapping as explained under Catch rates (CPUE) by sample.

This macro is particularly useful for data exploration and getting a quick overview of patterns or regularities in your experimental fishing or sampling design. It can also be used to calculate stratified mean catch per unit effort, with variance, depending on your sampling design.



Length by gear

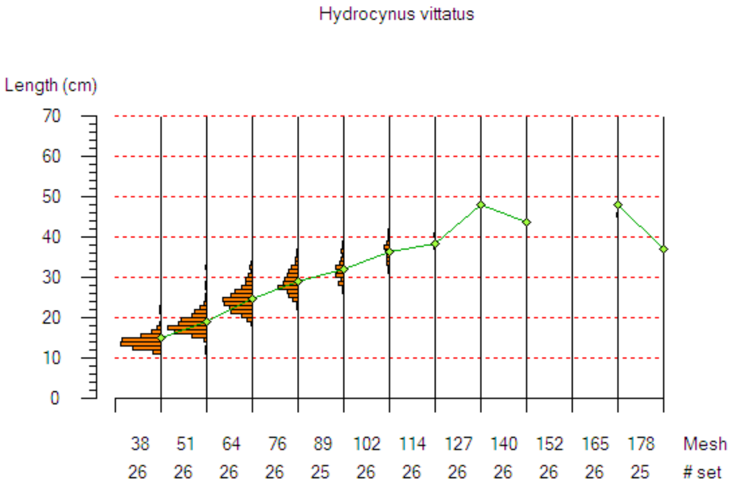


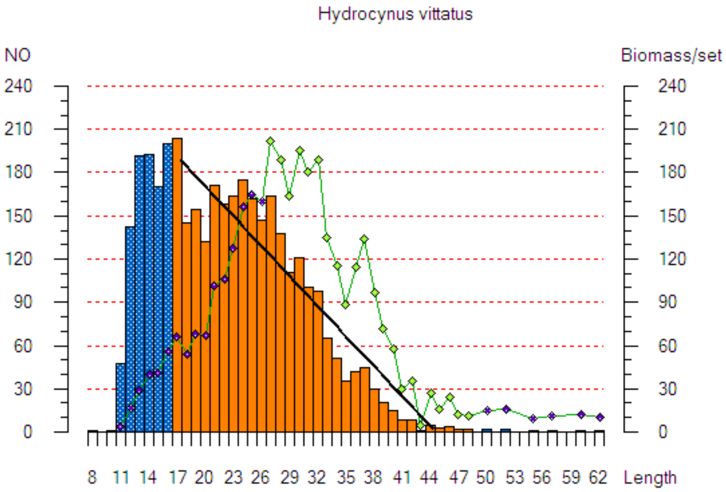
Length (cm)	38	51	64	76	89	102	114	127	140	152	165	178	Total	% NO	CUM% NO
8	1												1	0.0	0.0
9															0.0
10	1												1	0.0	0.1
11	39	8											47	1.3	1.3
12	135	7											142	3.9	5.3
13	189	3											192	5.3	10.5
14	180	12	1										193	5.3	15.9
15	95	72	3										170	4.7	20.5
16	46	149	5										200	5.5	26.0
17	20	182	2										204	5.6	31.6
18	6	130	9										145	4.0	35.6
19	4	121	28	1									154	4.2	39.9
20	5	73	53	1									132	3.6	43.5
21	9	61	98	2								1	171	4.7	48.2
22	8	34	110	5	1								158	4.3	52.6
23	3	16	134	9	2								164	4.5	57.1
24	3	8	136	27	1								175	4.8	61.9
25	5	8	101	45	2	1							162	4.5	66.3
26	1	3	85	52	6								147	4.0	70.4
27	3	9	50	96	6								164	4.5	74.9

This macro gives length frequencies by mesh (or gear) size in the nets or sampling gear (or any other chosen column grouping). Mean weight per length group with CV is displayed. Mean retention lengths (cm) and weight (g) for each mesh size with standard deviations.

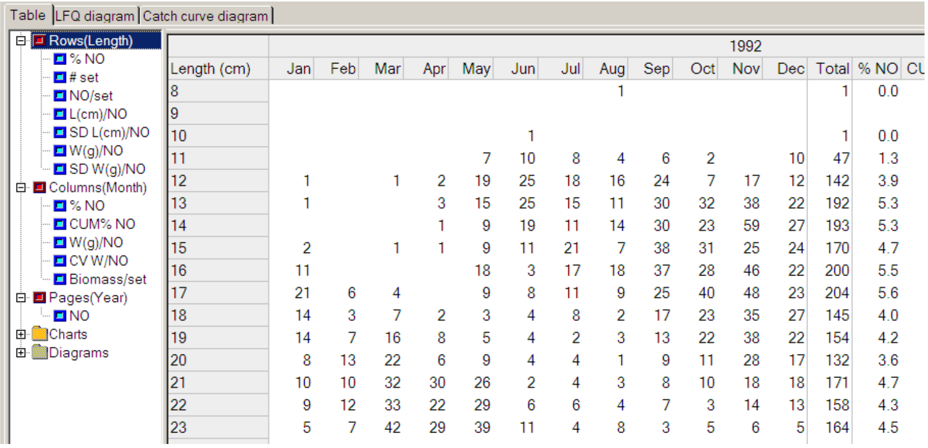
This macro is the standard programme for exporting the results into the tool "Estimate gear selectivity", in order to correct the length frequencies for gear selectivity (see Correction for gear selectivity).

The catch-curve diagram can be manipulated into various transformations such as ln-transformation of the length frequencies or calculation of relative biomass (CPUE in grams/set) in order to plot biomass-size distributions.





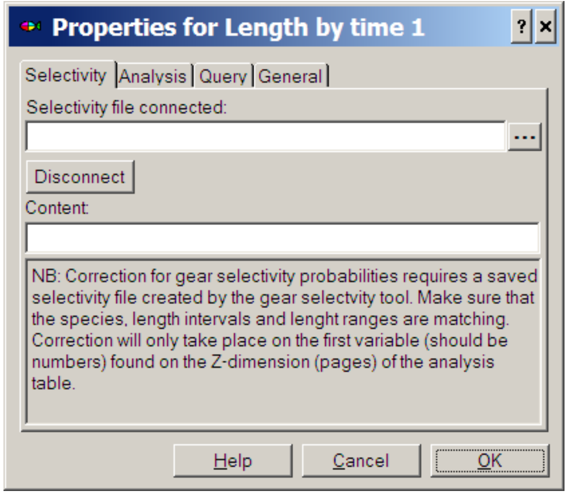
Length by time



	1992														
Length (cm)	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total	% NO	CU
8								1					1	0.0	
9															
10						1							1	0.0	
11					7	10	8	4	6	2		10	47	1.3	
12	1		1	2	19	25	18	16	24	7	17	12	142	3.9	
13	1			3	15	25	15	11	30	32	38	22	192	5.3	
14				1	9	19	11	14	30	23	59	27	193	5.3	
15	2		1	1	9	11	21	7	38	31	25	24	170	4.7	
16	11				18	3	17	18	37	28	46	22	200	5.5	
17	21	6	4		9	8	11	9	25	40	48	23	204	5.6	
18	14	3	7	2	3	4	8	2	17	23	35	27	145	4.0	
19	14	7	16	8	5	4	2	3	13	22	38	22	154	4.2	
20	8	13	22	6	9	4	4	1	9	11	28	17	132	3.6	
21	10	10	32	30	26	2	4	3	8	10	18	18	171	4.7	
22	9	12	33	22	29	6	6	4	7	3	14	13	158	4.3	
23	5	7	42	29	39	11	4	8	3	5	6	5	164	4.5	

The macro gives length frequencies by various chosen time intervals. If you set the pages (years) to NA the results will be pooled by time interval over all the years in the data base.

Correction for gear selectivity. If you have estimated gear selectivity of the particular species and stored the relative probabilities on file, you can optionally correct the observed length frequencies to estimated ‘true’ frequencies when examining the monthly catches or total catch curve or before exporting in FiSAT format. Choose the properties for the analysis, and then the ‘selectivity’ tab, and connect the appropriate selectivity file (*.sel) where species and length intervals must match with the format of the analysis:

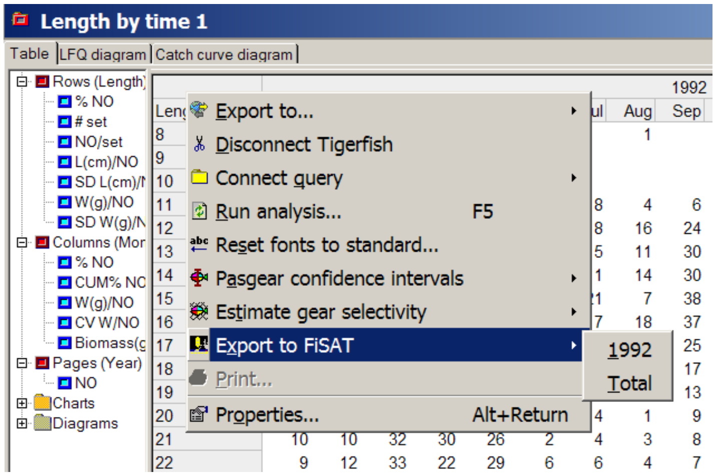


Selectivity file connected:

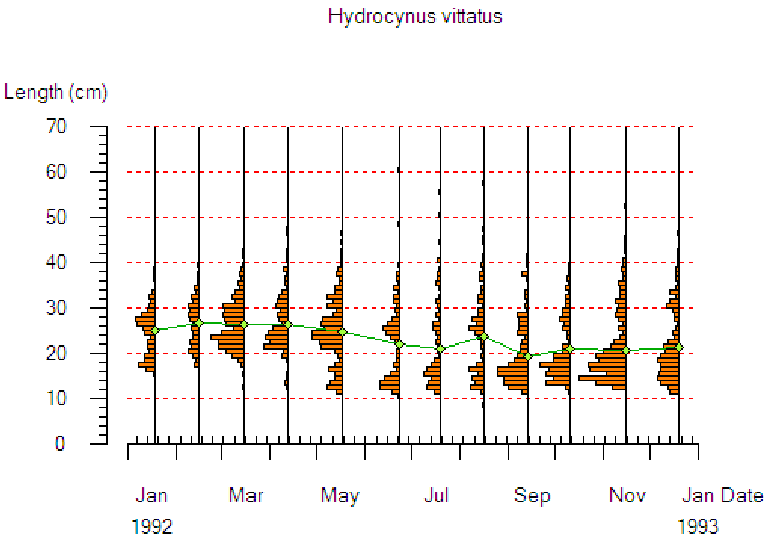
Content:

NB: Correction for gear selectivity probabilities requires a saved selectivity file created by the gear selectvity tool. Make sure that the species, length intervals and lenght ranges are matching. Correction will only take place on the first variable (should be numbers) found on the Z-dimension (pages) of the analysis table.

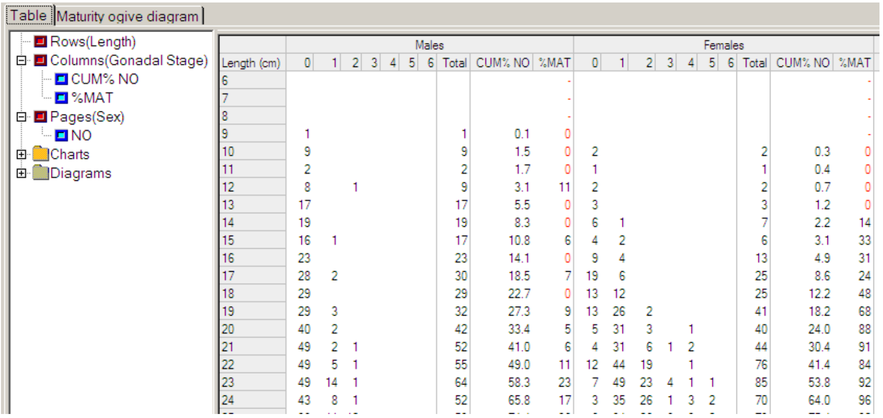
FiSAT export. This macro has also export facilities into FiSAT format (Gayanilo et al 1996). Right click on table and choose Export to FiSAT:



In the default chart the length frequencies are plotted on a time axis, either year by year or all years combined, to visually examine the various dynamics and/or modal progression which can be interpreted as growth or size-specific migration over time.

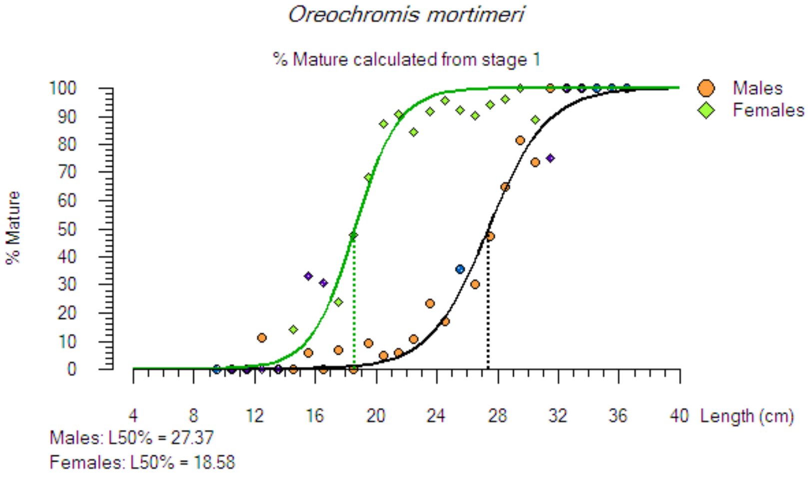


Maturity by length



	Males										Females									
Length (cm)	0	1	2	3	4	5	6	Total	CUM% NO	%MAT	0	1	2	3	4	5	6	Total	CUM% NO	%MAT
6										-										-
7										-										-
8										-										-
9	1							1	0.1	0										-
10	9							9	1.5	0	2							2	0.3	0
11	2							2	1.7	0	1							1	0.4	0
12	8		1					9	3.1	11	2							2	0.7	0
13	17							17	5.5	0	3							3	1.2	0
14	19							19	8.3	0	6	1						7	2.2	14
15	16	1						17	10.8	6	4	2						6	3.1	33
16	23							23	14.1	0	9	4						13	4.9	31
17	28	2						30	18.5	7	19	6						25	8.6	24
18	29							29	22.7	0	13	12						25	12.2	48
19	29	3						32	27.3	9	13	26	2					41	18.2	68
20	40	2						42	33.4	5	5	31	3		1			40	24.0	88
21	49	2	1					52	41.0	6	4	31	6	1	2			44	30.4	91
22	49	5	1					55	49.0	11	12	44	19		1			76	41.4	84
23	49	14	1					64	58.3	23	7	49	23	4	1	1		85	53.8	92
24	43	8	1					52	65.8	17	3	35	26	1	3	2		70	64.0	96

This macro gives distribution of gonadal stages by length and sex, as well as the percentage mature (%M) depending on the chosen stage for individuals considered mature.



Length at maturity ogive. A length-at-maturity ogive can be estimated by fitting a logistic model (Gunderson *et al.* 1980):

$$\mathrm{P}_m = \frac{1}{1+e^{(-aL+b)}} \cdot 100\text{, where}$$

P_m is percent mature (=%M) at length L, and a and b are fitted constants.

The model constants are estimated (see trends) by an iterative numerical search of the minimum sum of squares, $(=\sum(\textit{observed - predicted})^2)$. The non-linear search algorithm (Fletchers method) is adapted to Pascal from a QBasic program (FLET) supplied with Hilborn & Walters (1992).

$L_{50\%}$ is defined as the length where 50% of the fish are mature and are calculated (Rickey 1995) from the estimated constants a and b:

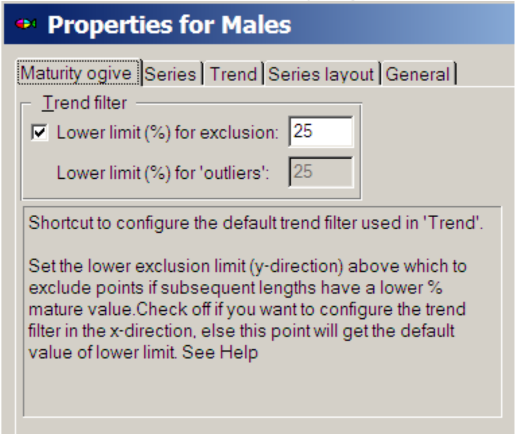
$$L_{50\%} = \frac{a}{b}$$

There are several options available when fitting the logistic model on each series (Males or females) which can be accessed by clicking on the series and choose properties, or simply double click:

- Lower limit (%) for exclusion (default = 25%):
- Lower limit (%) for 'outliers' (default = 25%):
- Minimum number of observations per point (default = 5):
- Weigh fitting with number of observations (default = 0):

These options should be explored in order to see what fit gives the best results.

Set lower limit (%) for exclusion. Double click on the series in the Diagram control pane and choose the 'Maturity ogive' tab:



☑ Lower limit (%) for exclusion: 25

Lower limit (%) for 'outliers': 25

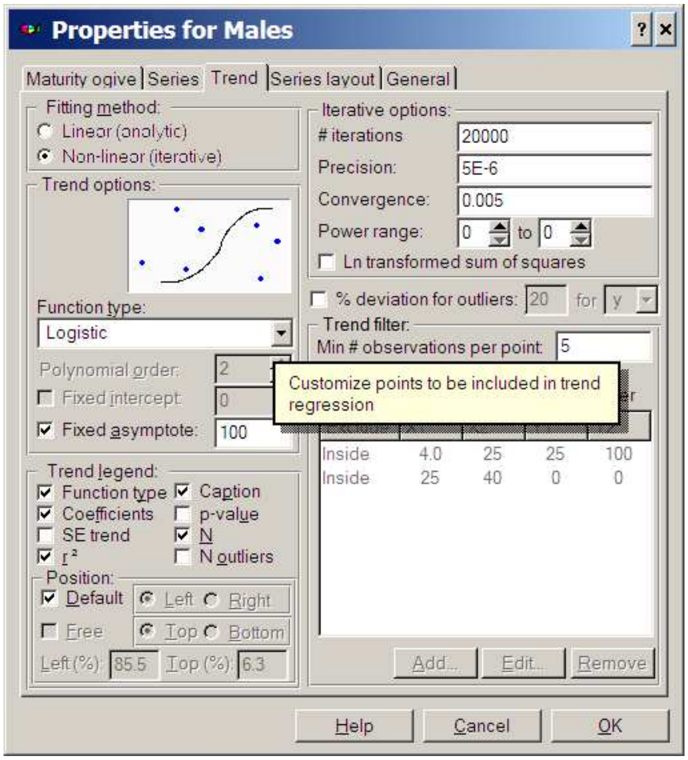
Shortcut to configure the default trend filter used in 'Trend'.

Set the lower exclusion limit (y-direction) above which to exclude points if subsequent lengths have a lower % mature value.Check off if you want to configure the trend filter in the x-direction, else this point will get the default value of lower limit. See Help

This tab is used as a shortcut to configure the optional default trend filter when estimating a maturity ogive. Since the value of % mature is based on the number of observed specimens, and this number can be low in the small or large length classes, there is sometimes 'noise' in the lower part of the logistic curve where a a high percentage of mature individuals can be caused one or two erroneous entries. Many small positive maturity values, intercepted by 0-values, may also be interpreted as 'noise'. However, if used they may tend to lower the slope of the ogive and thus over-estimate the $L_{50\%}$. By applying the default trend filter these will be excluded in order to improve the curve fit. The algorithm works as follows: When setting the Lower limit (%) for exclusion to e.g. 25%, then all values above this will be excluded **until** the point on the x-axis where the **subsequent** value is higher than the limit given by Lower limit (%) for 'outliers' = when $Y_i \geq$ lower limit **and** $Y_i < Y_{i+1}$. Normally the value for lower limit for exclusion = the value for lower limit for outliers (by default), but you can configure the two values separately by checking off the checkbox. The lower limit for exclusion (y-direction) will be kept, but the extent in the x-direction will be defined by the first point where the **subsequent** x-class (length) will have a higher % mature value than the new (normally lower) limit given for 'outliers'.

You can see the applied trend filter by checking '☐ Show excluded region' and/or the data not used by checking '☐ Data excluded from trend' on the Diagram control pane under the series.

Minimum number of observations per point. Double click on the series in the Diagram control pane and choose the Trend tab:



Fitting method:
○ Linear (analytic)
● Non-linear (iterative)

Trend options:
Function type: Logistic
Polynomial order: 2
☐ Fixed intercept: 0
☑ Fixed asymptote: 100

Trend legend:
☑ Function type ☑ Caption
☑ Coefficients ☐ p-value
☐ SE trend ☑ N
☑ r² ☐ N outliers

Position:
☑ Default ● Left ○ Right
☐ Free ● Top ○ Bottom
Left (%): 85.5 Top (%): 6.3

Iterative options:
iterations: 20000
Precision: 5E-6
Convergence: 0.005
Power range: 0 to 0
☐ Ln transformed sum of squares

☐ % deviation for outliers: 20 for y

Trend filter:
Min # observations per point: 5

Customize points to be included in trend regression

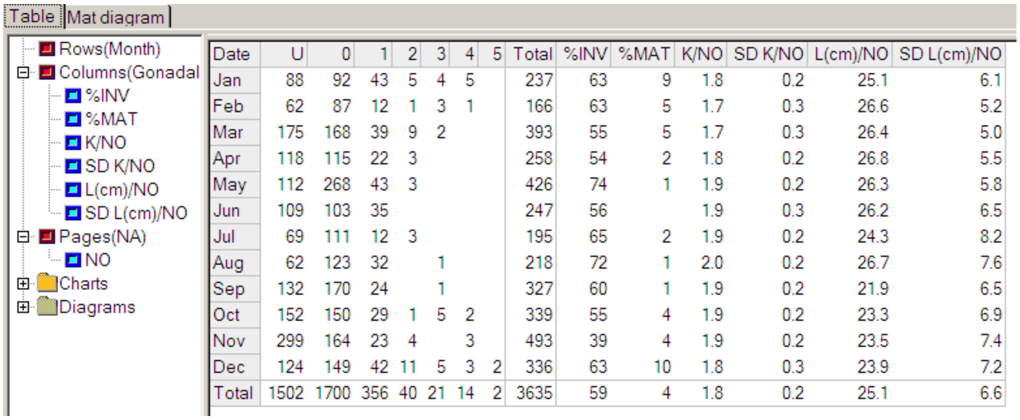
Inside 4.0 25 25 100
Inside 25 40 0 0

This filter is used to define the minimum number of observations required in a length group for the data-point to be included in the curve fit. Here you can also deactivate the default trend filter described above by checking off '□ Use trend filter'.

Weigh fitting with number of observations. Choose the Trend tab/Iterative options/Power range.

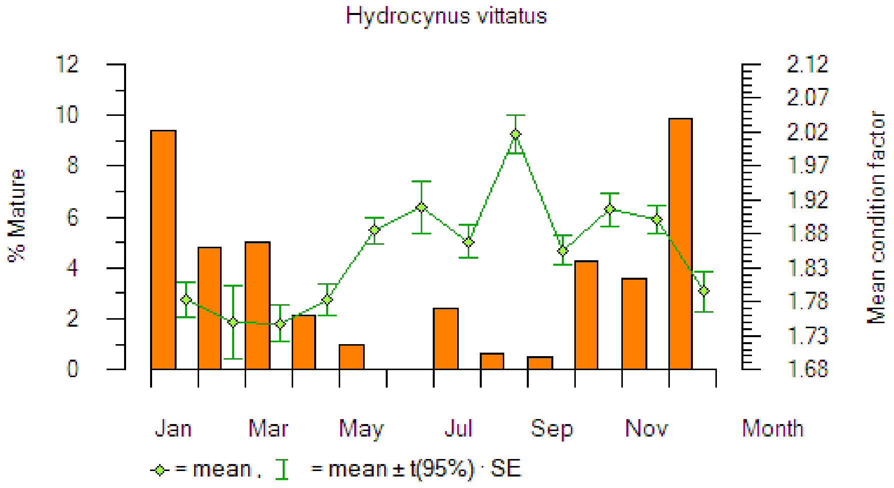
This option is used to give more weight to length (x-class) intervals with a high number of observations. The weighing is done by multiplying the sum of squares with the number of observations raised to the power of the entered weight value (which can range from 0 to 10), i.e. SumSQ=$\sum(\textit{observed} - \textit{predicted})^2 \cdot \#\text{obs}^{\text{weight}}$. Thus with weight value=0 there is no weighing. If you choose a range of power values, the fitted trend for each value will be displayed simoultaneously.

Maturity by time

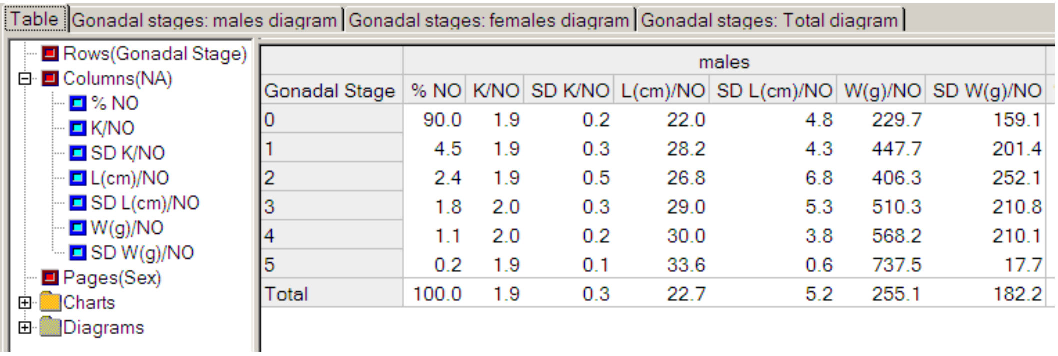


Date	U	0	1	2	3	4	5	Total	%INV	%MAT	K/NO	SD K/NO	L(cm)/NO	SD L(cm)/NO
Jan	88	92	43	5	4	5		237	63	9	1.8	0.2	25.1	6.1
Feb	62	87	12	1	3	1		166	63	5	1.7	0.3	26.6	5.2
Mar	175	168	39	9	2			393	55	5	1.7	0.3	26.4	5.0
Apr	118	115	22	3				258	54	2	1.8	0.2	26.8	5.5
May	112	268	43	3				426	74	1	1.9	0.2	26.3	5.8
Jun	109	103	35					247	56		1.9	0.3	26.2	6.5
Jul	69	111	12	3				195	65	2	1.9	0.2	24.3	8.2
Aug	62	123	32		1			218	72	1	2.0	0.2	26.7	7.6
Sep	132	170	24		1			327	60	1	1.9	0.2	21.9	6.5
Oct	152	150	29	1	5	2		339	55	4	1.9	0.2	23.3	6.9
Nov	299	164	23	4		3		493	39	4	1.9	0.2	23.5	7.4
Dec	124	149	42	11	5	3	2	336	63	10	1.8	0.3	23.9	7.2
Total	1502	1700	356	40	21	14	2	3635	59	4	1.8	0.2	25.1	6.6

This macro gives time interval frequencies of maturity stages for each specified species (or range of species) and percentage of 'gonadal active' (defined by an optional stage) of total investigated fish. Mean Fulton's condition factor (K) and mean length of total fish caught with standard deviations are also given. This macro can be used to determine the maximum spawning peaks and choosing the monthly sampling interval for calculating length-at-maturity.

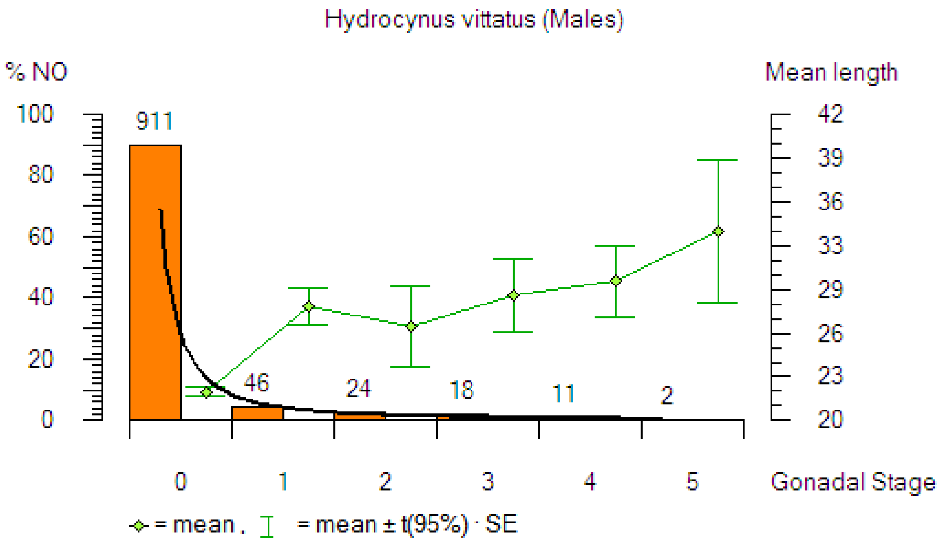


Stages by sex

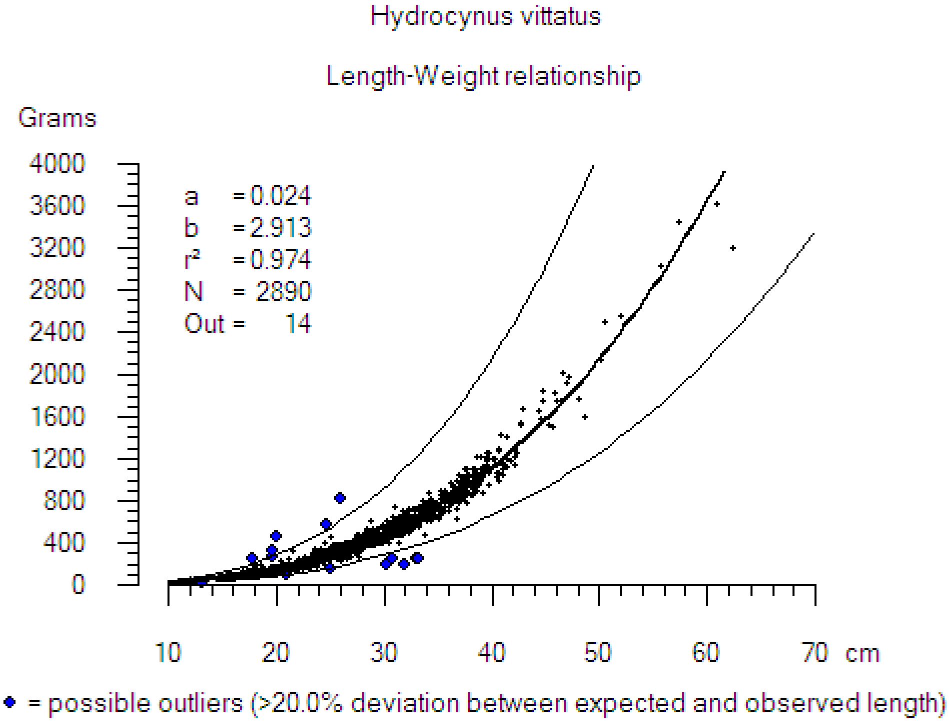


	males						
Gonadal Stage	% NO	K/NO	SD K/NO	L(cm)/NO	SD L(cm)/NO	W(g)/NO	SD W(g)/NO
0	90.0	1.9	0.2	22.0	4.8	229.7	159.1
1	4.5	1.9	0.3	28.2	4.3	447.7	201.4
2	2.4	1.9	0.5	26.8	6.8	406.3	252.1
3	1.8	2.0	0.3	29.0	5.3	510.3	210.8
4	1.1	2.0	0.2	30.0	3.8	568.2	210.1
5	0.2	1.9	0.1	33.6	0.6	737.5	17.7
Total	100.0	1.9	0.3	22.7	5.2	255.1	182.2

This macro gives the number of gonadal stages by sex, as well as the mean condition factor, mean length, and mean weight.



Length-Weight relationship

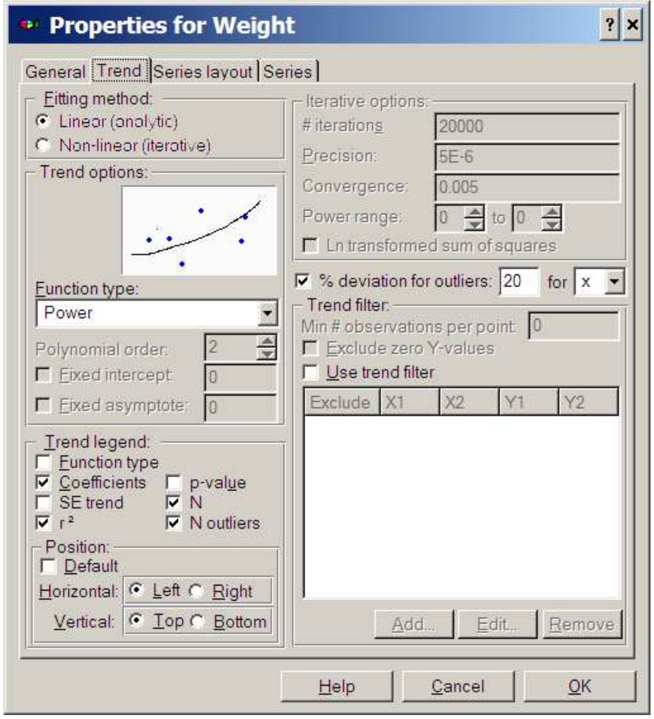


This macro is used to calculated length (cm)-weight (g) coefficients from the relationship

$$\text{weight} = a \cdot \text{length}^{b}$$

of the chosen species or range of species (in query), and show this in a diagram plot of the length-weight relationship, with possible outliers indicated. I also give the Condition factor as a function of length. Note that only records with Rank-code=0 are included by default in the length-weight regression, as different rank codes normally are set when either the length or the weight has been corrected under cleaning.

The length-weight relationship can be regressed both linearly by logarithmic transformation of the power function (default) or iteratively by selecting the fitting method under Series/Properties/trend. The default deviation for ‘outliers’ is 20% for x, i.e. when there is more than 20% deviation between the observed and the expected length (x) for a given weight (y).



Fitting method:
● Linear (analytic)
○ Non-linear (iterative)

Trend options:
Function type: Power
Polynomial order: 2
☐ Fixed intercept: 0
☐ Fixed asymptote: 0

Trend legend:
☐ Function type
☑ Coefficients ☐ p-value
☐ SE trend ☑ N
☑ r² ☑ N outliers

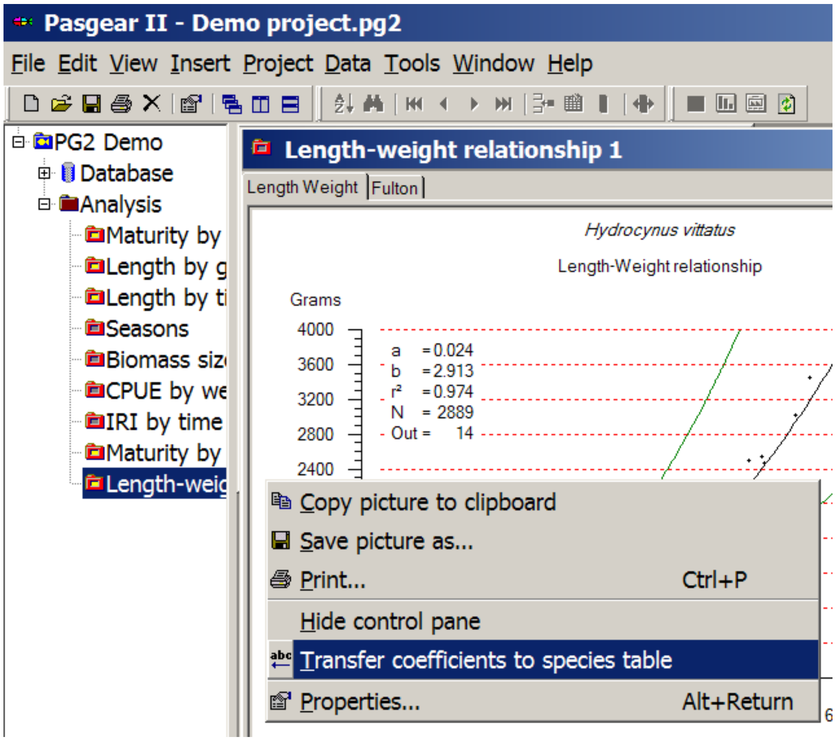
Position:
☐ Default
Horizontal: ● Left ○ Right
Vertical: ● Top ○ Bottom

Iterative options:
iterations: 20000
Precision: 5E-6
Convergence: 0.005
Power range: 0 to 0
☐ Ln transformed sum of squares

☑ % deviation for outliers: 20 for x

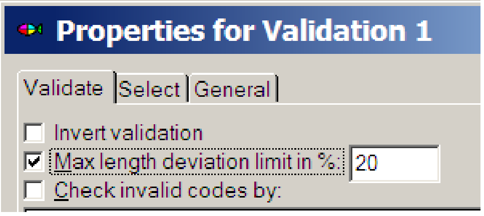
Trend filter:
Min # observations per point: 0
☐ Exclude zero Y-values
☐ Use trend filter

The coefficients (a and b, as well as the coefficient of determination (r^2) and the sample size N), if acceptable, can optionally be written automatically into the Species table for the respective species (right click on the Analysis and choose ► ‘Transfer coefficients to species table’).

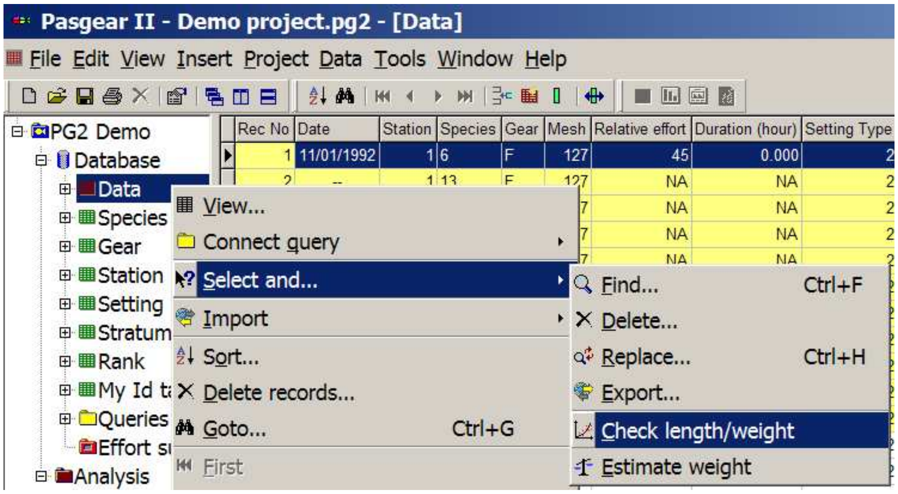


a = 0.024
b = 2.913
r^2 = 0.974
N = 2889
Out = 14

The length-weight coefficients are used when running validation queries on observed length versus weight in the database (in order to find the ‘outliers’), or when running the modules ‘Check length/weight’ or ‘Estimate weight’ on the data table (see also Find, clean and correct records, and missing weights).



☐ Invert validation
☑ Max length deviation limit in %: 20
☐ Check invalid codes by:



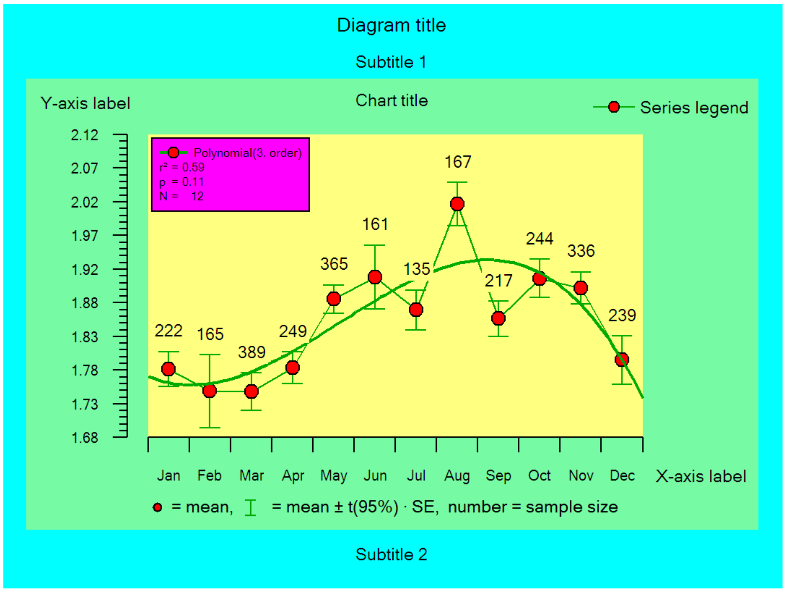
Rec No	Date	Station	Species	Gear	Mesh	Relative effort	Duration (hour)	Setting Type
1	11/01/1992	1	6	F	127	45	0.000	2
2	--	1	13	F	127	NA	NA	2
					7	NA	NA	2
					7	NA	NA	2
					7	NA	NA	2
					7	NA	NA	

Diagrams and charts

There are almost endless possibilities for designing your diagrams and charts in terms of layout and orientation. Below only a brief description is given on how to access the different property pages and their configuration.

Basic concepts

- A **diagram** is a frame (shown in cyan in the example below) that can display one to several charts, or other diagrams, as well as title, two subtitles and a free positioned legend.
- A **chart** consist of two parts:
  - ➢ The chart area (in green in example below) consisting of a title, axes, legends and a plot area.
  - ➢ The plot area (in yellow in example below) which plot one to several series on 2 or 3 axes.
- An **axis** (X,Y or Z) can hold one to several series and can be formatted in various ways.
- A **series** is a set of data connected to an axis that can be plotted with various layouts.
- A **trend** is a curve that can be fitted to a series and added to the plot with an optional legend
- A **graph element** is any item in a diagram or chart that has graphic properties such as color, line, style, fills, and for some fonts.



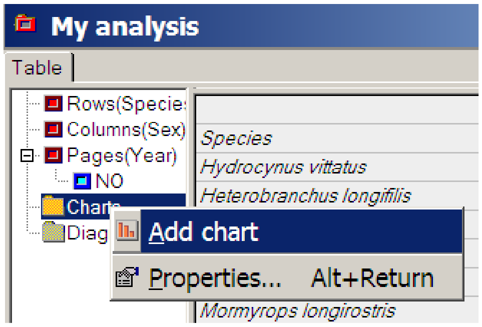
Making a chart

A diagram and chart can be created either automatically or by the user. When charts are created automatically (default charts) they are 'ad hoc' or non-persistent charts, which mean that they will be re-created on every run of the analysis. These charts can, however, be modified and saved as persistent if desirable (right click in the diagram and choose [Keep diagram] from the pop-up menu). If an 'ad hoc' diagram or chart has been made persistent, and the analysis still is asked to make default charts, then these will be added to the persistent charts.

Automatic charts are created in two ways:

1. All predefined analyses will make automatic default charts if the ☑ Make default charts. on the analysis property frame is checked.
2. Any single variable can also be rendered graphically if the ☐ Make 'Ad-Hoc' chart on the variable property frame is checked.

User created charts are created by either click on the cart icon on the toolbar, or right click on the Charts folder in the Analysis tree view ► **Add chart**



This will open the Chart property page/Chart type setup:

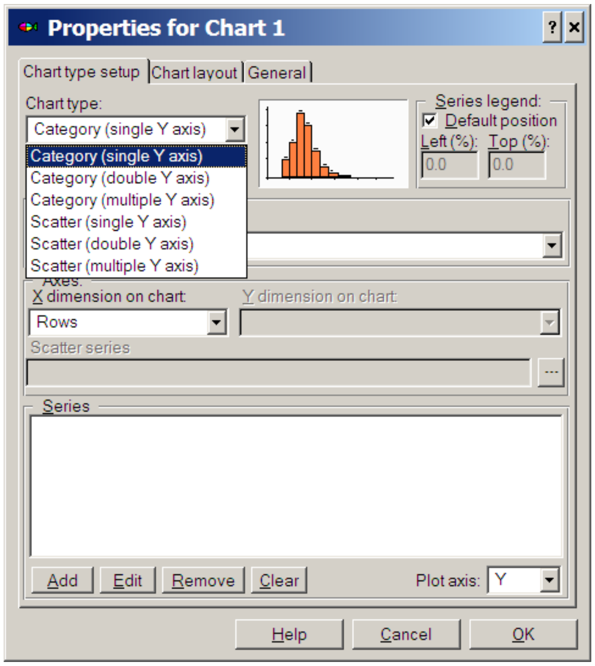


Chart type: Category (single Y axis)

☑ Default position

Left (%): 0.0

Top (%): 0.0

X dimension on chart: Rows

Plot axis: Y

You can choose between 6 types of charts:

1. Category (single Y-axis)
2. Category (double Y-axis)
3. Category (multiple Y-axis)
4. Scatter (single Y-axis)
5. Scatter (double Y-axis)
6. Scatter (multiple Y-axis)

The preview next to the combo-box will give you an idea of how each are looking.

Then you must choose your X-axis from the table dimension (rows, columns, or pages).

If you have a multiple Y-axis chart you must also choose your primary Y-axis from one of the remaining table dimensions. In this case the data will be plotted against a series of ‘Z-axes’ which is lying parallel to the X-axis, but has no scale (see The Diagram tree view and options for display for how to get the Z-scale on top of the plot area).

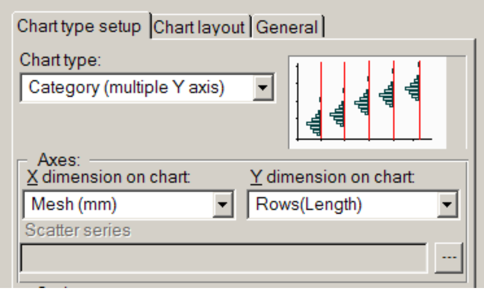


Chart type: Category (multiple Y axis)

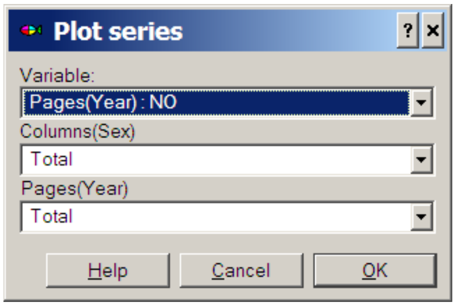
X dimension on chart: Mesh (mm)

Y dimension on chart: Rows(Length)

If you have a category plot you can now choose your series, and if you have a scatter plot you can choose the X-values and the Y-values (= series):

1. Category plot

Add your series by clicking [Add] under the series pane:

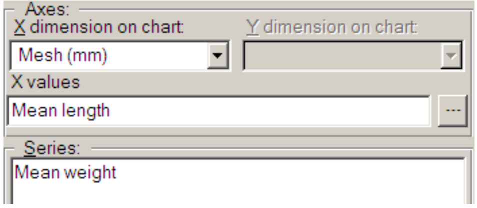


Variable: Pages(Year) : NO
Columns(Sex): Total
Pages(Year): Total

Choose the variable you want as a series, and where on the 'Y-dimension' of the table you find it. If your X-dimension is rows, then the available Y-dimension is columns and vice versa. If you have pages then also specify which page to choose from (default = Total).

2. Scatter plot

First choose your X-values by pressing […]



X dimension on chart: Mesh (mm)
Y dimension on chart:
X values: Mean length
Mean weight

Then choose your Y-values by clicking [Add] under the series pane:

You can add as many series as you have on any chart type. If you have double Y-axis on either the category- or scatter plot then assign each added series to one of the two axes by choosing plot axis:

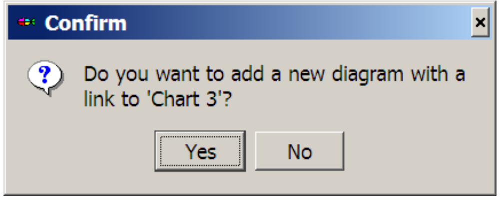
Plot axis: Y1

If you have multiple Y-axes then you can assign the series to either the primary Y-axis or the additional Z-axis.

Hint! You can go into any of the chart properties in the diagrams of the different predefined analyses (see Analysis) to see how they have been configured. All diagrams in Pasgear 2 are built from the same modules.

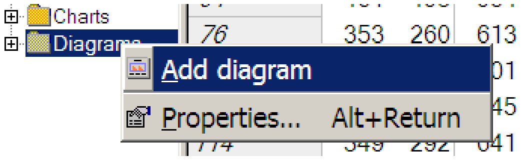
Linking a chart to a diagram

When you have created and designed a new chart and pressed [OK] you will be prompted:

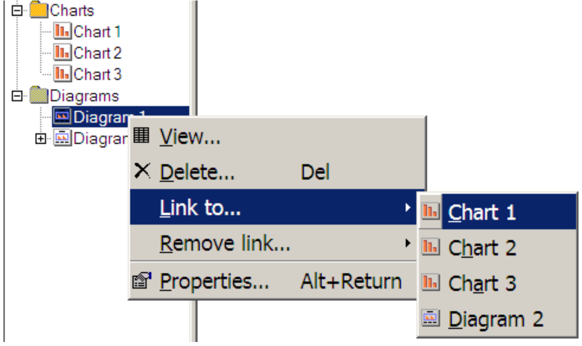


Do you want to add a new diagram with a link to 'Chart 3'?

If you confirm a new diagram will automatically be created and linked with the chart.
If you press [No] the new chart will not be visible before you have linked it to an existing diagram.
If there are no diagrams then add one by click on the Diagram icon on the toolbar, or right click on Diagrams folder ► Add diagram:

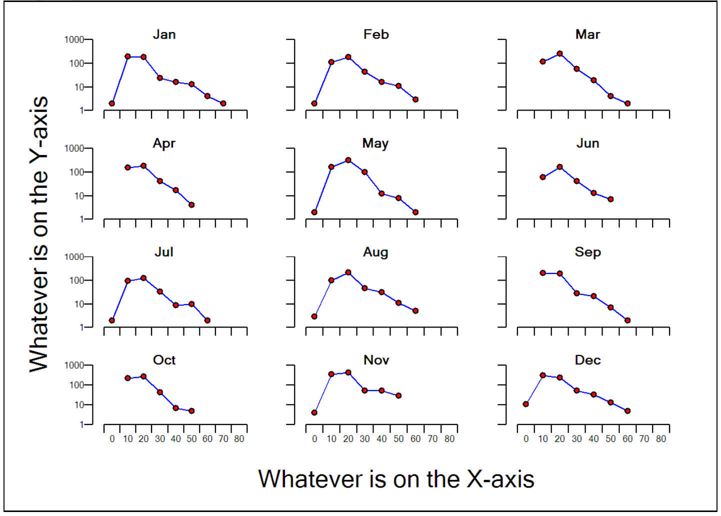


If a diagram exists then link the chart by right click the Diagram ► Link to.. ► choose object.



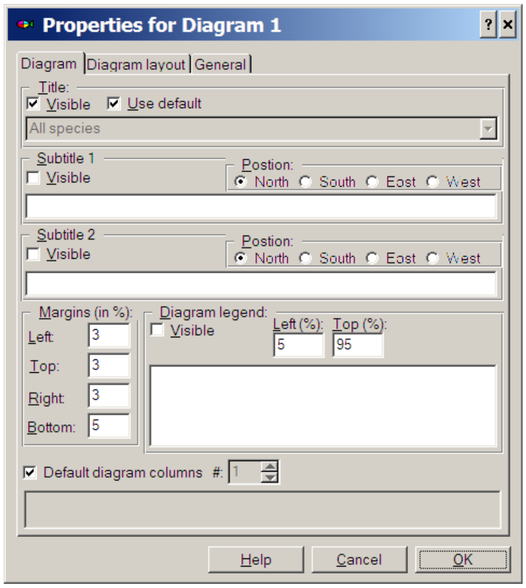
A diagram can also be linked to another diagram, and any links can be removed by right click the Diagram ► Remove link.. ► choose object.

Multiple charts in one diagram. You can link as many charts or diagrams into one diagram as you like and format the layout by defining the number of columns. If your analysis is grouped by pages then each chart will be given a default title corresponding to the page group.



Properties for diagrams

Diagrams, charts, axes, and series all have the property Tabs '**Layout**' and '**General**' in common. On the **General** frame you can, like elsewhere in Pasgear 2 give the **Caption** and add **comments**. The graph **Layout** frame is, besides the graph elements included, identical for and will be described below under Layout of Diagrams, Charts, and series.



Title: ☑ Visible ☑ Use default
All species

Subtitle 1: ☐ Visible
Postion: ● North ○ South ○ East ○ West

Subtitle 2: ☐ Visible
Postion: ● North ○ South ○ East ○ West

Margins (in %):
Left: 3
Top: 3
Right: 3
Bottom: 5

Diagram legend: ☐ Visible
Left (%): 5
Top (%): 95

☑ Default diagram columns #: 1

The diagram properties has

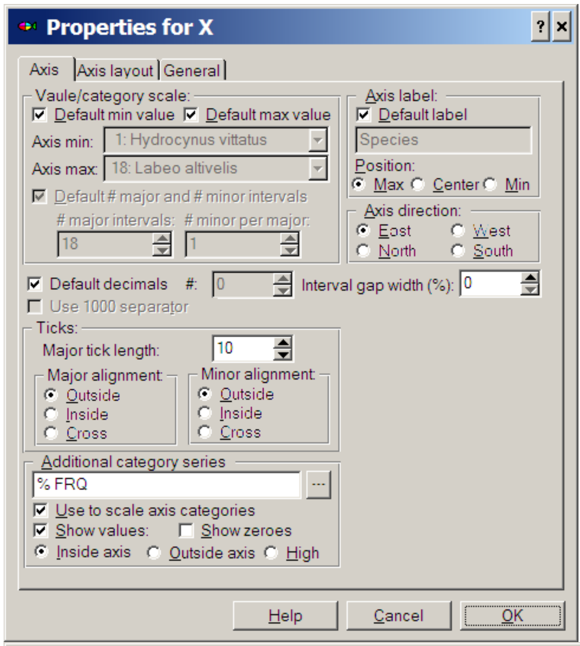
- Title
- Subtitle 1 + position (north, south , east , west)
- Subtitle 2 + position (north, south , east , west)
- Margins (in percent of area)
- A legend + free position
- A number of columns

The number of columns defines the relative position of the 1..N linked objects. If you use **default** the number of columns will be round(square root(# objects)). Thus two charts will be positioned next to each other. Four will be positioned two by two etc. If you want all object on top of each other then choose 1 column. If you want all next to each other then # columns = # objects.

Properties for charts

Is described above under Making a chart.

Properties for axes



Vaule/category scale:
☑ Default min value ☑ Default max value
Axis min: 1: Hydrocynus vittatus
Axis max: 18: Labeo altivelis
☑ Default # major and # minor intervals
major intervals: 18 # minor per major: 1

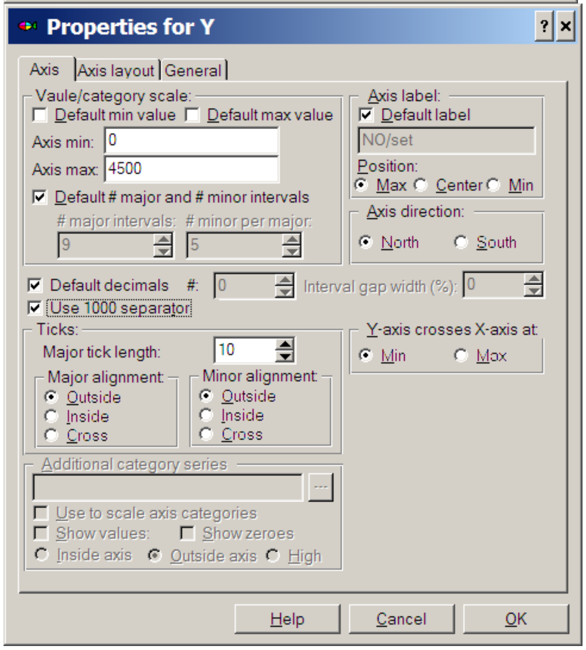
Axis label:
☑ Default label
Species
Position: ● Max ○ Center ○ Min

Axis direction:
● East ○ West
○ North ○ South

☑ Default decimals #: 0 Interval gap width (%): 0
☐ Use 1000 separator

Ticks:
Major tick length: 10
Major alignment: ● Outside ○ Inside ○ Cross
Minor alignment: ● Outside ○ Inside ○ Cross

Additional category series
% FRQ
☑ Use to scale axis categories
☑ Show values: ☐ Show zeroes
● Inside axis ○ Outside axis ○ High



Vaule/category scale:
☐ Default min value ☐ Default max value
Axis min: 0
Axis max: 4500
☑ Default # major and # minor intervals
major intervals: 9 # minor per major: 5

Axis label:
☑ Default label
NO/set
Position: ● Max ○ Center ○ Min

Axis direction:
● North ○ South

☑ Default decimals #: 0 Interval gap width (%): 0
☑ Use 1000 separator

Ticks:
Major tick length: 10
Major alignment: ● Outside ○ Inside ○ Cross
Minor alignment: ● Outside ○ Inside ○ Cross

Y-axis crosses X-axis at:
● Min ○ Max

Additional category series
☐ Use to scale axis categories
☐ Show values: ☐ Show zeroes
○ Inside axis ● Outside axis ○ High

Here you can set:

- The range (Min and max value) of both categories and values (if plot type is scatter or the axis is continuous)
- The axis label and position (Max, Min or Center),
- The axis direction (if on the X-axis you can turn this in all directions (East, West, South, North), if on Y then North or South, and if you have multiple Y-axes, then the data will be plotted against a Z-axis that can be turned East or West).
- The number of major intervals (if not a category axis)
- The number of minor intervals per major (if not a category axis)
- The axis values format (number of decimals or using 1000 separator, the format of the latter is defined in the Regional and Language options of the PC).
- If on the Y-axis you can set its position on the X-axis (crosses at Min or max)

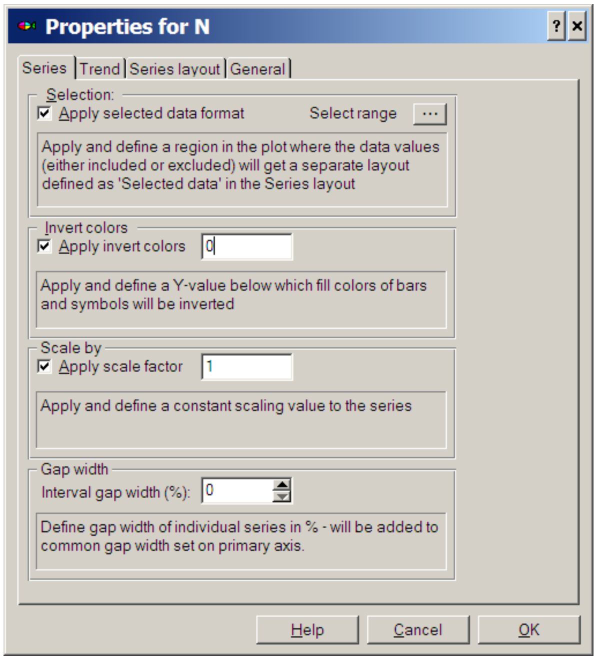
- If the axis is a category you can change the interval gap width, i.e. narrowing the displayed elements. This operation can also be performed on each series individually.
- The length of tick marks and their position on the axis (Inside, Outside or Cross).
- If on the X-axis (and category) you can add an additional series to be displayed, and even use its values to scale the X-categories (such as the IRI plot).

By changing the axis direction and position you can turn your chart in nearly all directions, and if you have double Y-axis these can be pointing in opposite directions (one North, the other South).

Properties for series

Series have two special property frames (tabs) besides 'Series layout' and 'General':

- Series
- Trend (see below under Add trend)



Selection:
☑ Apply selected data format — Select range ...
Apply and define a region in the plot where the data values (either included or excluded) will get a separate layout defined as 'Selected data' in the Series layout

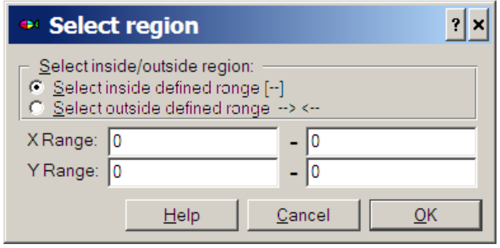
Invert colors
☑ Apply invert colors: 0
Apply and define a Y-value below which fill colors of bars and symbols will be inverted

Scale by
☑ Apply scale factor: 1
Apply and define a constant scaling value to the series

Gap width
Interval gap width (%): 0
Define gap width of individual series in % - will be added to common gap width set on primary axis.

Under Series you can do:
1) **Selection**: Define and apply a region in the series that you may want to give a special layout (color, marker etc.).

Click on […] and define the selected region:



Select inside/outside region:
● Select inside defined range [--]
○ Select outside defined range --> <--
X Range: 0 - 0
Y Range: 0 - 0

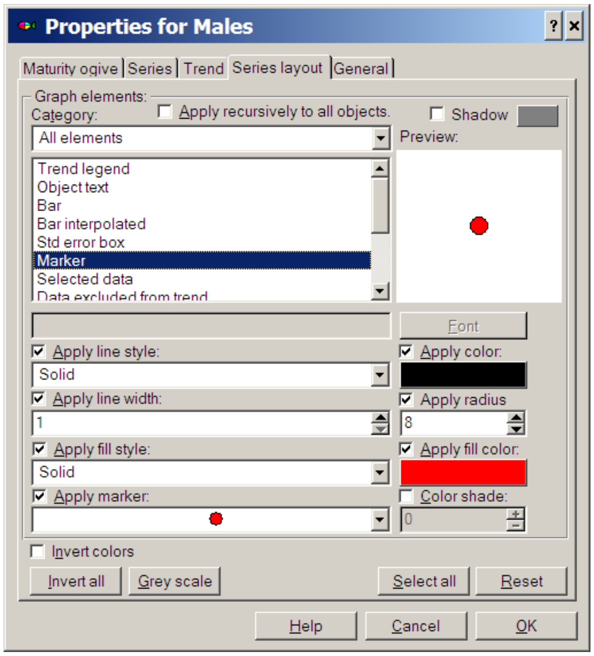
This region will have the graph element name 'Selected data' in the Series layout (see Layout of Diagrams, Charts and Series below).

2) **Invert colors**: Define a constant Y-value below which the colors of filled elements (bars and symbols) will be automatically inverted.

3) **Scale by**: Scale or raise (= multiply) all values by a constant factor.

4) **Gap width**: Give each individual series a gap width in the X-direction. This gapwidth will be added to the common gap width defined on the axis-property.

Layout of Diagrams, Charts, Axes and Series



Properties for Males

Graph elements:

Category: All elements

☐ Apply recursively to all objects.

☐ Shadow

Preview:

Trend legend
Object text
Bar
Bar interpolated
Std error box
Marker
Selected data
Data excluded from trend

☑ Apply line style: Solid

☑ Apply color:

☑ Apply line width: 1

☑ Apply radius 8

☑ Apply fill style: Solid

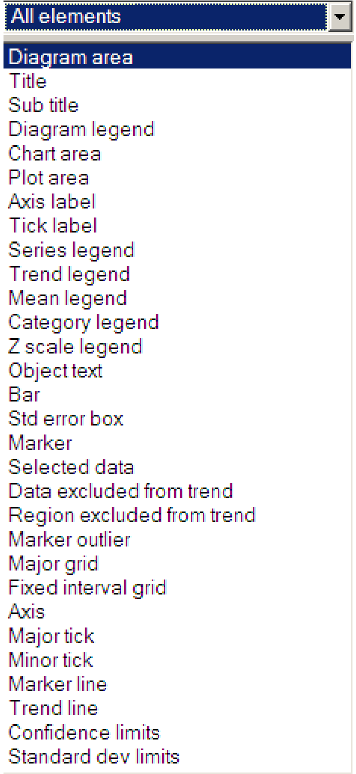
☑ Apply fill color:

☑ Apply marker: ●

☐ Color shade: 0

☐ Invert colors

There are 30 different graph elements that can all be given their own layout in terms of font, size, color, line, style, fill etc. depending on the graph element type.



All elements

Diagram area
Title
Sub title
Diagram legend
Chart area
Plot area
Axis label
Tick label
Series legend
Trend legend
Mean legend
Category legend
Z scale legend
Object text
Bar
Std error box
Marker
Selected data
Data excluded from trend
Region excluded from trend
Marker outlier
Major grid
Fixed interval grid
Axis
Major tick
Minor tick
Marker line
Trend line
Confidence limits
Standard dev limits

The various elements can be categorized and filtered under **Category** into

- All elements
- Font elements
- Line elements
- Fill elements

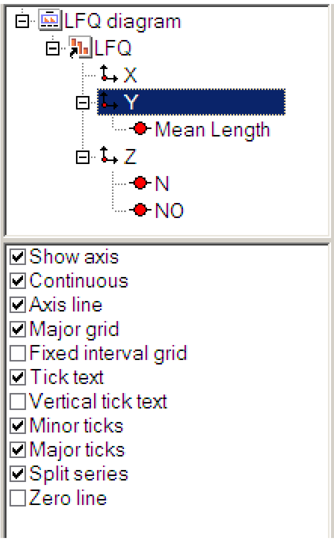
Depending on layout property page you are on (Diagram, Charts, or Series) you can access all elements from the hierarchical level you have selected and below, and apply your changes on all objects from your hierarchical level and below by checking:
☐ Apply recursively to all objects.

Depending on the element type you can change the available layout attributes (if some attributes are not available, such as Font on a marker symbol, then these will be grayed out on the property frame. You should select only the style attributes (line color, width, fill color, fill style etc.) you want to change by checking ☐ Apply

You can invert the colors on either selected element by checking '□ Invert colors', or all colors in the layout by pressing the Invert all button.
Similarly, you can convert all colors to grey scale (monochrome) by pressing the Grey scale button.

You can reset the layout to the specified default (see Project properties) by selecting all elements and press the Reset button. You can reset everything to default by entering at the highest hierarchical level (Diagram), check ☐ Apply recursively to all objects. and press Reset.

The Diagram tree view and options for display



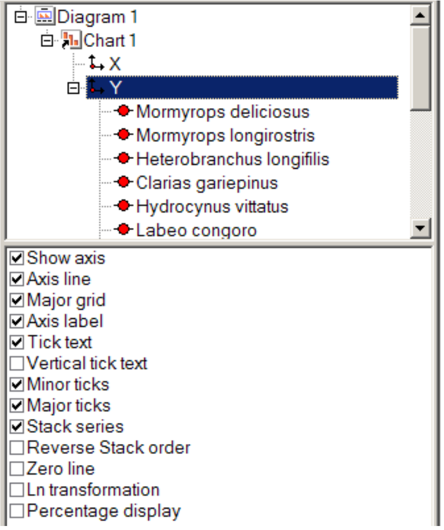
LFQ diagram
LFQ
X
Y
Mean Length
Z
N
NO

☑ Show axis
☑ Continuous
☑ Axis line
☑ Major grid
☐ Fixed interval grid
☑ Tick text
☐ Vertical tick text
☑ Minor ticks
☑ Major ticks
☑ Split series
☐ Zero line

On the diagram tree view you will see all the objects associated with the diagram:
The Chart (or charts), their respective axes (X,Y and Z if multiple Y-axes), and the series assigned to the respective axes. By clicking on these you get a number of options for display depending on the object. By double clicking you get into the properties frames of each object.

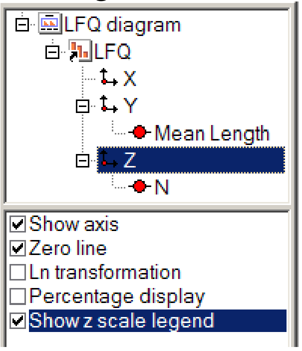
On the axes you can change configuration (grid, text, ticks etc,) or even make them invisible (Show axis).
If there are several series on the same axis these can be stacked, and on the opposite axis they can be split.

If the axis is continuous (i.e. not a category) the series values can be ln transformed or given in percent. If the values are ln-transformed there will be an additional option for converting the axes to log-scale display.



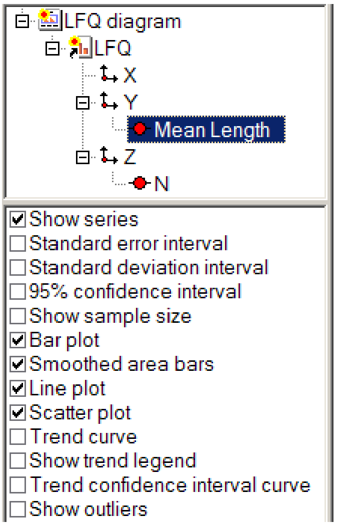
☑ Show axis
☑ Axis line
☑ Major grid
☑ Axis label
☑ Tick text
☐ Vertical tick text
☑ Minor ticks
☑ Major ticks
☑ Stack series
☐ Reverse Stack order
☐ Zero line
☐ Ln transformation
☐ Percentage display

If the axis is a Z-axis (in multiple Y-axes plots) then you can get a scale indication by checking ‘show z-scale legend’)



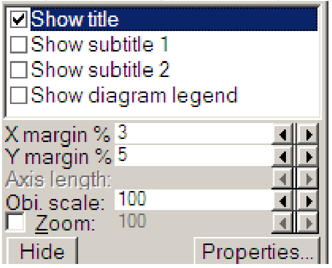
☑ Show axis
☑ Zero line
☐ Ln transformation
☐ Percentage display
☑ Show z scale legend

On the series you can remove these visually from the plot or change their display (bar, line, scatter). If bars these can be smoothed like an area plot.
If the series is a mean value of many observations then it can be shown with error bars (SE, SD or 95% CI) and the sample size (number of observations for each point). You can also fit and add an optional trend (see Add trend below).



☑ Show series
☐ Standard error interval
☐ Standard deviation interval
☐ 95% confidence interval
☐ Show sample size
☑ Bar plot
☑ Smoothed area bars
☑ Line plot
☑ Scatter plot
☐ Trend curve
☐ Show trend legend
☐ Trend confidence interval curve
☐ Show outliers

Below the tree view you have options for zooming, change the legend positions, change axes length or diagram margins

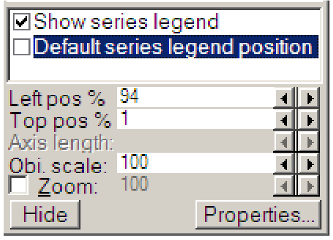


☑ Show title
☐ Show subtitle 1
☐ Show subtitle 2
☐ Show diagram legend

X margin % 3
Y margin % 5
Axis length:
Obj. scale: 100
☐ Zoom: 100

The **Objects** zoom will simultaneously change the relative size of selected objects in the diagram, and is a short cut for changing font size, line width, symbol size etc.

Changing legend positions: When clicking an object in the diagram which has a legend (Diagram, chart, or series if a trend is applied) and the legend is displayed, you check can off '□ Default position' and move the legend by changing the values in the 'Left pos' and 'Top pos' fields:



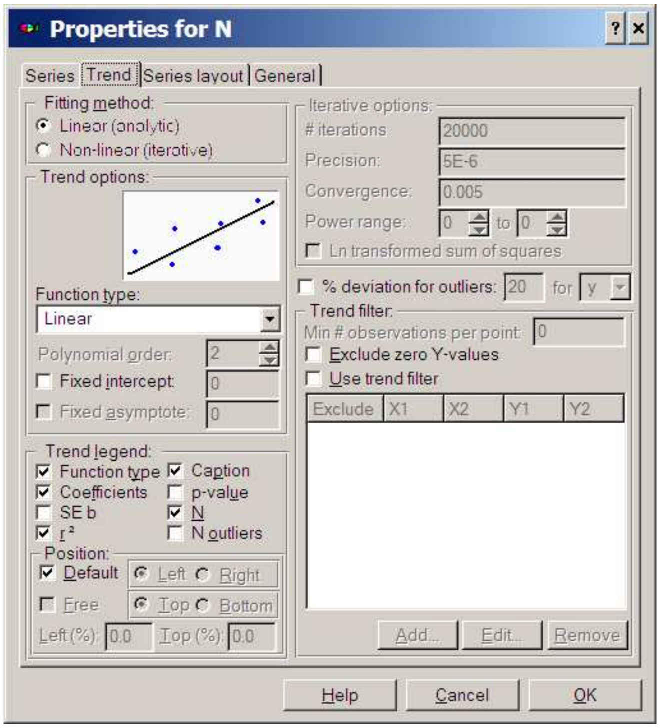
☑ Show series legend
☐ Default series legend position

Left pos % 94
Top pos % 1
Axis length:
Obj. scale: 100
☐ Zoom: 100

The **Zoom** is like an ordinary zoom where you can zoom in and out of the whole diagram You can also use **Hide** to remove the tree view. It can be reinstalled by right click in the diagram ► Show control pane.

Add Trend

On the series property page there is a frame for configuration of a fitted trend.



Fitting method:
● Linear (analytic)
○ Non-linear (iterative)

Trend options:
Function type: Linear
Polynomial order: 2
☐ Fixed intercept 0
☐ Fixed asymptote: 0

Trend legend:
☑ Function type ☑ Caption
☑ Coefficients ☐ p-value
☐ SE b ☑ N
☑ r² ☐ N outliers

Position:
☑ Default ● Left ○ Right
☐ Free ● Top ○ Bottom
Left (%): 0.0 Top (%): 0.0

Iterative options:
iterations 20000
Precision: 5E-6
Convergence: 0.005
Power range: 0 to 0
☐ Ln transformed sum of squares

☐ % deviation for outliers: 20 for y

Trend filter:
Min # observations per point 0
☐ Exclude zero Y-values
☐ Use trend filter

There are two different fitting methods

1. Linear (analytic)
2. Non-linear (iterative)

The linear is based on linear transformation of the data (depending on the trend type) and least squares regression.
The non-linear are based on an iterative numerical search of the minimum sum of squares, ($=\sum(observed - predicted)^2$). The non-linear search algorithm (Fletchers method) is adapted to Pascal from a QBasic program (FLET) supplied with Hilborn & Walters (1992).

Trend type

Depending on the chosen fitting method there are number of trends available.

For the linear method there are:

1. Linear
2. Logarithmic
3. Power
4. Exponential
5. Polynomial
6. Cubic Spline
7. Moving average

For the non-linear method you have the same (except the 2 smoothers: spline and moving avarage) plus in addition

8. Logistic
9. Standard normal
10. Skewed normal

- For the polynomial you can set the polynomial order to anything less than the number of points-2

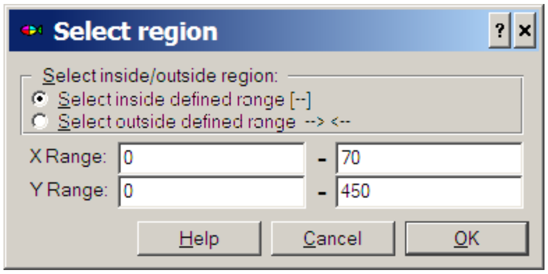
- For the linear, exponential and the polynomial (up to order of 2) you can fix and set the intercept
- For the Logistic you can fix and set the asymptote
- For the moving average you can set the Period (# elements included), and whether the curve should be centered within the period or at the end.

There are a number of options available (for the experienced user) for the iterative method. If the series is a mean or a percentage and the number of observations behind each point is not the same, then you can give weight to the fit by the number of observations by using the **power** option. The weighing is done by multiplying the sum of squares with the number of observations raised to the power of the entered value (which can range from 0 to 10 i.e. $\mathrm{SumSQ}=\sum(\mathit{observed} - \mathit{predicted})^2\cdot\#\mathrm{obs}^{\mathrm{power}}$. Thus with power value = 0 there is no weighing. Note that you can see the effect of different power values by giving a range. A curve will then be drawn for each power value in the range.

Trend filter

You can customize the number of data points included in the trend by applying a trend filter. There are 3 options:

- If the series is a mean or a percentage you can exclude all point where the underlying number of observation is below a given limit by setting **Min # observations per point [ ]**.
- You can simply **exclude all 0 Y-values**
- Or you can select a number of regions to include or exclude by **Use trend filter** and **Add**



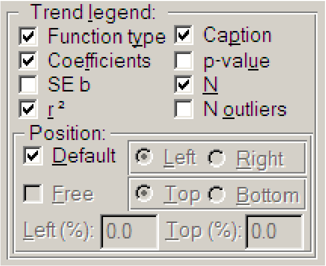
Select inside/outside region:
● Select inside defined range [--]
○ Select outside defined range --> <--
X Range: 0 - 70
Y Range: 0 - 450

Several of the predefined analyses in Pasgear is using a default trend filter (such as Catch-curve diagram and Maturity diagram). In such case there will be a check mark for **Use default** but this can be overridden or changed if desired.

If a trend filter is used there will be an option on the series display for ☐ Show excluded region.

Trend legend

You can customize the amount of information you want in the trend legend



Trend legend:
☑ Function type ☑ Caption
☑ Coefficients ☐ p-value
☐ SE b ☑ N
☑ r² ☐ N outliers
Position:
☑ Default ● Left ○ Right
☐ Free ● Top ○ Bottom
Left (%): 0.0 Top (%): 0.0

- **Function type** writes the chosen trend type
- **Caption** includes the series caption
- **Coefficients** give the estimated coefficients, where the number depends on the trend type
- **SE (b or trend)** gives the Standard error of the slope (if linear) or else the trend
- **r^2** gives the coefficient of determination (= regression sum of squares / total sum of squares)
- **p-value** gives the probability that the trend is different from 0 (F = regression mean squares / residual mean squares)
- **N** gives the number of points in the regression
- **N-outliers** give the number of points where there are more than a specified percentage deviation between the observed value and the expected value. The level of deviation is given in **% deviation for outliers** (default = 20).

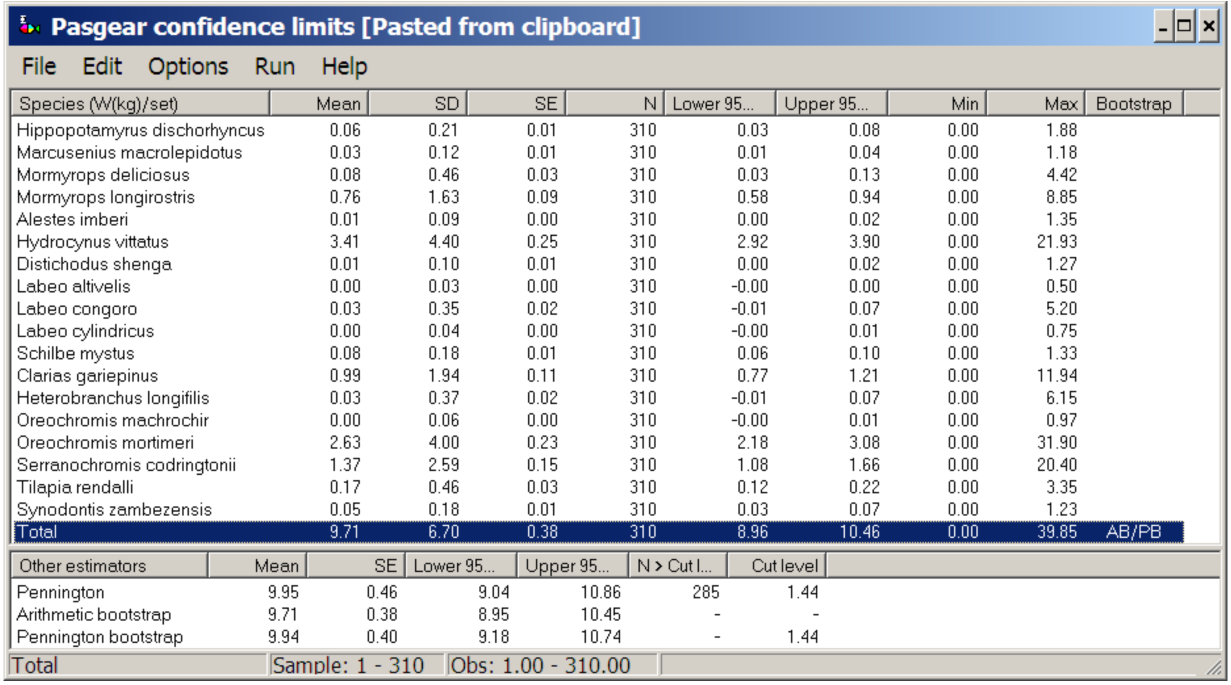
For the two smoothing trends (cubic spline and moving average), coefficients, r^2 and p-values are not available.

Finally you can give the position of the trend legend. Either chooses the default which is under the series legend, or in one of the corners of the plot area, or a ‘Free’ position.

Tools

Confidence limits

This stand alone tool (conflim.exe) of the Pasgear II package is meant of calculating alternative confidence intervals than standard parametric from a dataset, such as Pennington's estimator or bootstrap (see options for confidence intervals).



Species (W(kg)/set)	Mean	SD	SE	N	Lower 95...	Upper 95...	Min	Max	Bootstrap
Hippopotamyrus dischorhyncus	0.06	0.21	0.01	310	0.03	0.08	0.00	1.88	
Marcusenius macrolepidotus	0.03	0.12	0.01	310	0.01	0.04	0.00	1.18	
Mormyrops deliciosus	0.08	0.46	0.03	310	0.03	0.13	0.00	4.42	
Mormyrops longirostris	0.76	1.63	0.09	310	0.58	0.94	0.00	8.85	
Alestes imberi	0.01	0.09	0.00	310	0.00	0.02	0.00	1.35	
Hydrocynus vittatus	3.41	4.40	0.25	310	2.92	3.90	0.00	21.93	
Distichodus shenga	0.01	0.10	0.01	310	0.00	0.02	0.00	1.27	
Labeo altivelis	0.00	0.03	0.00	310	-0.00	0.00	0.00	0.50	
Labeo congoro	0.03	0.35	0.02	310	-0.01	0.07	0.00	5.20	
Labeo cylindricus	0.00	0.04	0.00	310	-0.00	0.01	0.00	0.75	
Schilbe mystus	0.08	0.18	0.01	310	0.06	0.10	0.00	1.33	
Clarias gariepinus	0.99	1.94	0.11	310	0.77	1.21	0.00	11.94	
Heterobranchus longifilis	0.03	0.37	0.02	310	-0.01	0.07	0.00	6.15	
Oreochromis machrochir	0.00	0.06	0.00	310	-0.00	0.01	0.00	0.97	
Oreochromis mortimeri	2.63	4.00	0.23	310	2.18	3.08	0.00	31.90	
Serranochromis codringtonii	1.37	2.59	0.15	310	1.08	1.66	0.00	20.40	
Tilapia rendalli	0.17	0.46	0.03	310	0.12	0.22	0.00	3.35	
Synodontis zambezensis	0.05	0.18	0.01	310	0.03	0.07	0.00	1.23	
Total	9.71	6.70	0.38	310	8.96	10.46	0.00	39.85	AB/PB

Other estimators	Mean	SE	Lower 95...	Upper 95...	N > Cut l...	Cut level
Pennington	9.95	0.46	9.04	10.86	285	1.44
Arithmetic bootstrap	9.71	0.38	8.95	10.45	-	-
Pennington bootstrap	9.94	0.40	9.18	10.74	-	1.44

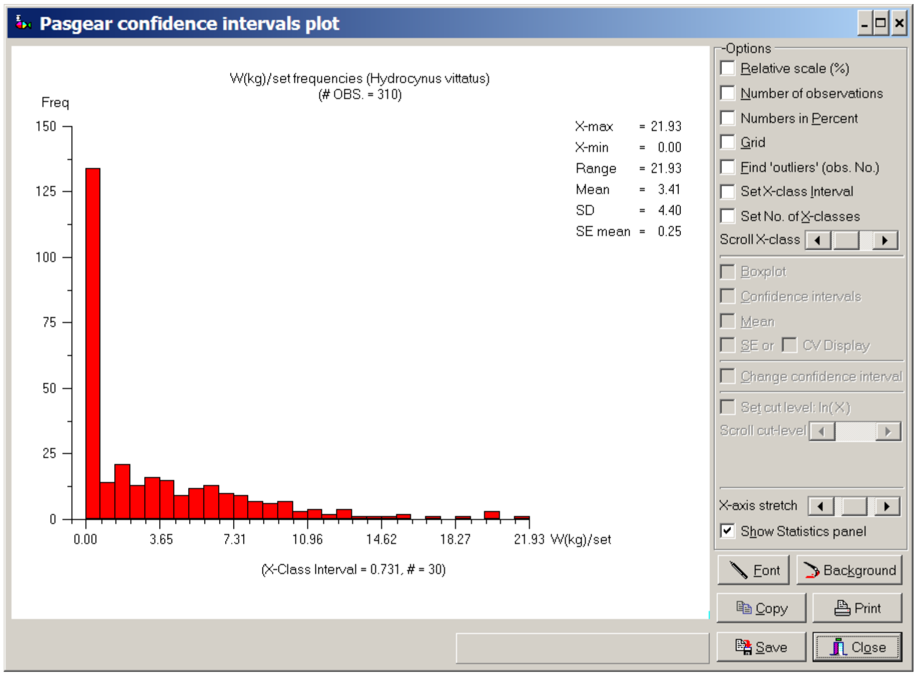
Total | Sample: 1 - 310 | Obs: 1.00 - 310.00

The program will read a general matrix from the clipboard or a text file of the following format (separators can be <tab> or comma <,>):

Line 1:	General header line					
Line2:	Group Label, e.g. species		value unit, e.g. kg/set			
Line 3:	SampleId_1	SampleId_2	Groupname_1	Groupname_2	….	Groupname_n
Line4:	Val	val	val	val	val	val
etc	Val	val	val	val	val	val

The two first columns <Sample Id_1> and <Sample Id_2> are for grouping and sorting. They can contain e.g. the station number and the depth, or any other grouping value. If no information then just fill out with dummies e.g. numbers 1..n. The remaining columns (from 3 and onwards) are the actual data values for each group. The length of each group column does not need to be the same, so for example if the groups are e.g. years and the data are samples (and there is an unequal sample size for each year) then just fill out fields with no information with a dash '-' or something which is not a number.

The frequency distribution



```
X-max    = 21.93
X-min    =  0.00
Range    = 21.93
Mean     =  3.41
SD       =  4.40
SE mean  =  0.25
```

- ☐ Relative scale (%)
- ☐ Number of observations
- ☐ Numbers in Percent
- ☐ Grid
- ☐ Find 'outliers' (obs. No.)
- ☐ Set X-class Interval
- ☐ Set No. of X-classes
- ☐ Boxplot
- ☐ Confidence intervals
- ☐ Mean
- ☐ SE or ☐ CV Display
- ☐ Change confidence interval
- ☐ Set cut level: ln(X)
- ☑ Show Statistics panel

Due to the behaviour of fish and often clumped (contagious) distribution, the frequency distributions of catch rates (CPUE) in most fisheries are highly skewed to the right (log normal or negative binomial). This has consequences for using parametric statistical tests when sample sizes are small, or if there are occasionally very large catches, as parametrical statistical tests assume that data are normally distributed.

In order to examine the actual frequency distribution of a particular data set, there is an option for viewing the frequency distribution graphically (see figure above). This option shows both the original value and the logarithmic transformations ($x = \ln(x+1)$ or the Delta distribution ($x = \ln(x)$, if $x>0$), (see below under the Pennington estimator).

Options for calculating confidence intervals

There are four options for calculating confidence intervals:

1. standard parametric based on arithmetic the sample mean
2. the Pennington estimator based on the log-normal Delta distribution
3. bootstrap on the arithmetic sample mean
4. bootstrap on the Pennington estimator

Each of the methods will be elaborated below

Estimates based on the sample mean

The estimate of the standard error for the mean CPUE is given by

$SE(CPUE) = \sqrt{S^2/n}$, where S^2 is the variance of the individual observations.

If the sample size is "large" enough, then the Central Limit Theorem states that there is a 95% chance that the true mean lies in the interval (see Cochran, 1977, pp. 39-44)

$\overline{CPUE} \pm t_{(n-1)} \cdot SE(CPUE)$,   where

t is from Students t-table with (n-1) degrees of freedom and $\alpha = 0.025$.

Estimates of the mean based on log-normal theory - The Pennington estimator

Since catch data from fisheries surveys usually have a large variance (much higher than the mean) and are highly skewed to the right, the sample sizes are typically not large enough so that estimates based on the sample mean and standard SE is a valid 95% confidence interval. In fact, the confidence associated with the parametric confidence interval is usually much lower than 95% (McConnaughey and Conquest 1992; Conquest *et al*. 1996; Pennington 1996). A major problem to the degree of skewness is due to the high proportion of zero catches often observed, and the occasional occurrence of very large catches. Development of confidence intervals is complicated by the asymmetric distribution, and the occurrence of zero catches confounds an effective normalization transformation. Logarithmic transformation will stabilize the variance but data will still not be normally distributed and interpretation of re-transformed means is difficult.

One way to generate more precise estimates of the mean and more accurate confidence statements for skewed catch data is to base the estimators on the log-normal Delta distribution (Pennington 1983, 1996; Conquest *et al.*, 1996), in which catches are divided into zero and non-zero units, followed by transformation of the non-zero values to natural logarithms. When it is found that the transformed non-zero data are approximated by a log-normal distribution (*i.e.* the logged values are normally distributed), then a more efficient estimator of mean CPUE is given by (Pennington, 1983, 1996)

$$c = \frac{m}{n}\exp(\bar{x})\,G_m(s_x^2/2) \quad , \qquad \text{where}$$

m is the number of sample values greater than 0,

$\bar{x}$ and s_x^2 are the mean and variance, respectively, of the logged values of non-zero catches,

$G_m(f)$ is an infinite series function of m and f [for example $f = s_x^2/2$ in above equation] which is used to correct for bias in re-transformation from log to arithmetic scale and is defined by

$$G_m(f) = 1 + \frac{m-1}{m}f + \sum_{j=2}^{\infty}\frac{(m-1)^{2j-1}f^{\,j}}{m^j(m+1)(m+3)\cdots(m+2j-3)\,j!}$$

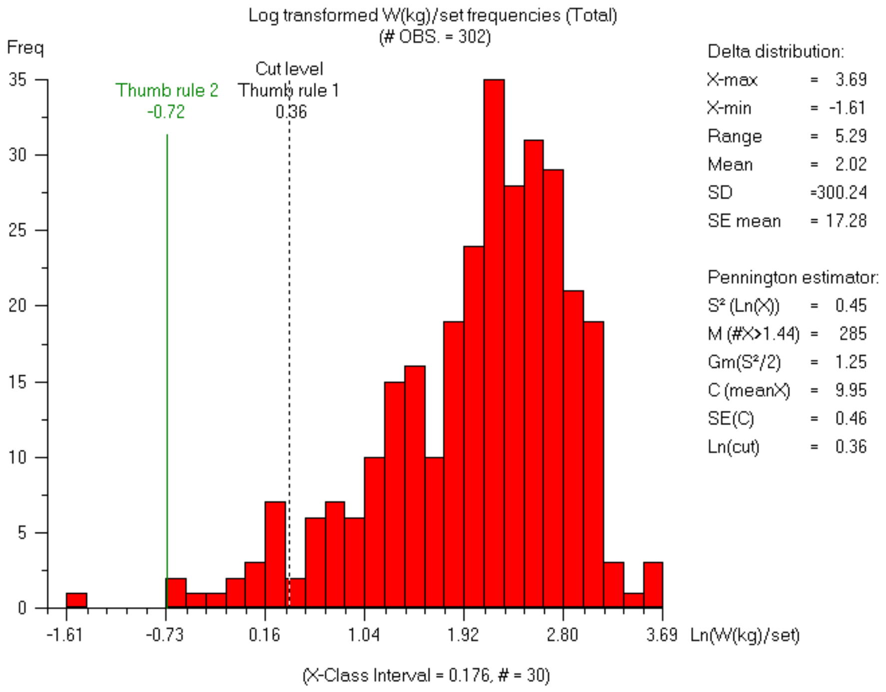
The variance of c is given by

$$\operatorname{var}(c) = \frac{m}{n}\exp(2\bar{x})\left\{\frac{m}{n}G_m^2(S_x^2/2) - \frac{(m-1)}{n-1}G_m\left(\frac{m-2}{m-1}S_x^2\right)\right\}$$

and the standard error is

$$SE = \sqrt{\operatorname{var}(c)}$$

The modified Pennington estimator



In contrast to estimates based on the arithmetic sample mean, which are highly sensitive to a single or a few isolated high catch rates that may account for more than 50% of the total catch, Pennington's estimator is sensitive to low catch rates which contribute little to the total catch, but when log-transformed may give large negative values resulting in a distribution becoming skewed to the left. In such a case a more precise estimator of mean CPUE, $\hat{\mu}$, is given by (modified by Michael Pennington from Pennington 1983, 1996)

$$\hat{\mu} = \frac{(n-m)}{n}\bar{y}' + \frac{m}{n}\exp(\bar{x})G_m(s_x^2/2)$$

where

m is the number of sample values greater than a defined 'cut-level' (rather than 0),
$\bar{y}'$ denotes the arithmetic mean of the non-transformed values less than the cut-level,
$\bar{x}$ and s_x^2 are the mean and variance, respectively, of the logged values of catches greater than the cut-level.

The variance of $\hat{\mu}$ is given by

$$\mathrm{var}(\hat{\mu}) = \mathrm{var}(c) + \left(\frac{n-m-1}{n(n-1)}\right)s'^2 + \left(\frac{m(n-m)}{n^2(n-1)}\right)\bar{y}'^2 - 2\left(\frac{n-m}{n(n-1)}\right)\bar{y}'\times c \; ,$$

where

s'^2 is the variance of the non-transformed values less than the cut-level, and
c and var(c) are the equations above but with m bigger than the cut-level instead of 0.

Setting the cut level:

There is no single objective criterion upon which to define a cut-level bigger than zero. Basically the logged Delta distribution should be viewed (in the plot module, see above figure) in order to determine if it is skewed to the left and/or contains isolated small catches. As a 'rule of thump' (Pennington pers. com.) the cut-level should be set as either:

- Thump rule 1 (R1) $= (2\overline{x} - x_{\max})$, where $\overline{x}_i$ and $x_{\max}$ are the mean and the largest value, respectively, of the log-transformed values greater than 0.

- Thump rule 2 (R2) = y% of $\overline{x}$, where default y = 5%, but the fraction (%) of the mean value can be changed by the user

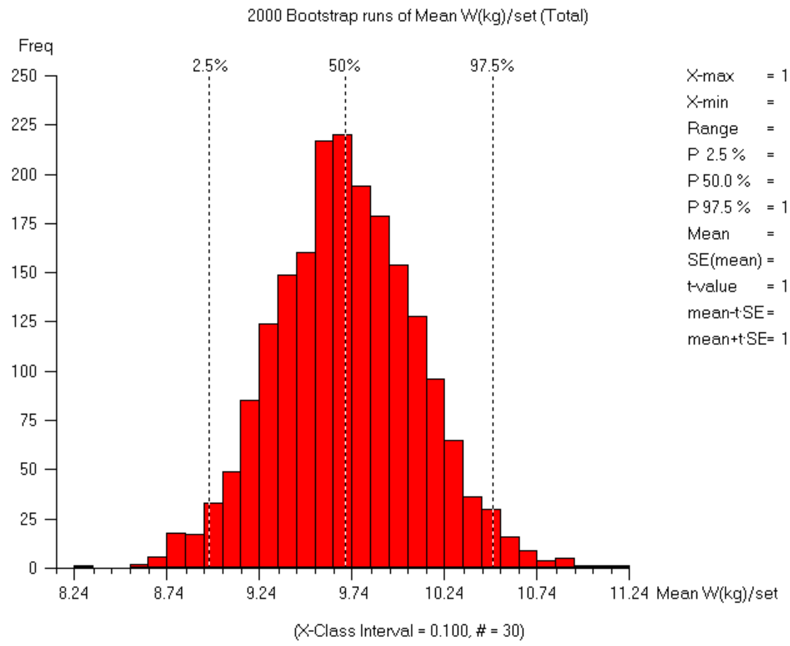
In the plotting module of conflim.exe there is an option for setting the cut-level bigger than 0 by pressing [T]. You can then either set your own cut-level by entering a number or automatically apply one of the two the 'rules of thump' by entering 'R1' or 'R2'. It is always advisable to view the Delta distribution to see the actual cut-level. When viewing the Delta distribution it is also possible to change the cut-level up and down (by toggling the arrow keys [→] and [←]) to see the effect on the Pennington's modified estimator.

Bootstrap confidence intervals

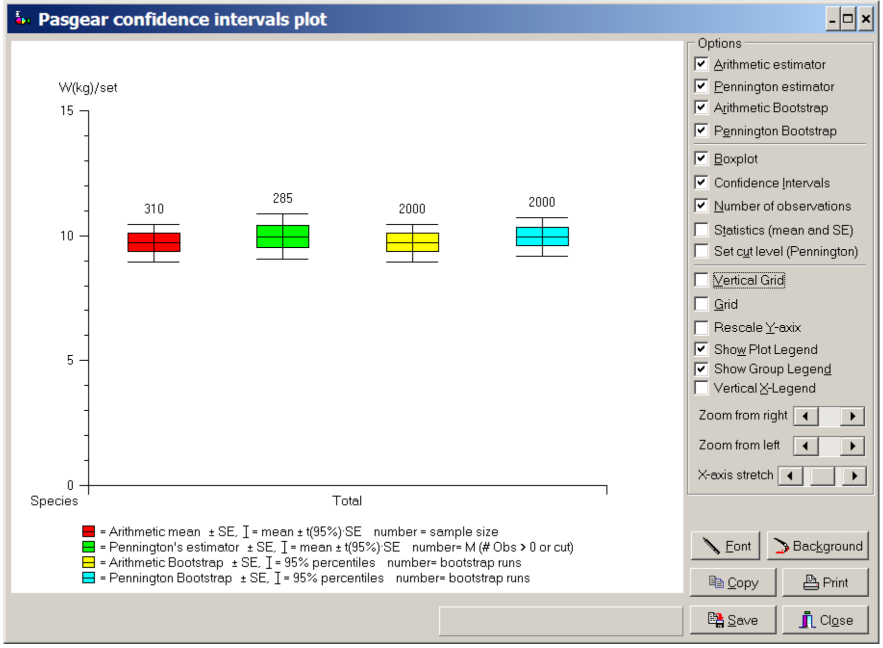
The bootstrap (Efron and Tibshirani 1986, 1993) is a computer-based re-sampling method for assigning measures of accuracy to statistical estimates. The properties of a sample statistic (e.g. mean CPUE) are determined from the sample information itself and thus are free from the usual distributional assumptions in parametric statistics. The basic idea is to construct a sequence of *new samples* (individual observations) which is obtained by randomly sampling N times with replacement from the original data points (i.e. the original individual CPUE estimates where N is the size of the original sample, i.e. the total number of individual CPUE estimates = settings). By generating a large number of independent bootstrap sample series (typically 2000), each of size N, then the frequency distribution of the mean of the bootstrap samples will approximate a normal distribution (Fig. 6). Standard 95% confidence limits can then be obtained by taking the 2.5% and 97.5% percentiles (P 2.5%, P 97.5%), or by calculating the standard error SE (and multiplying with 1.96 from the *t*-distribution when df. > 120 and α = 0.025). By calculating the

$$\overline{CPUE} \pm 1.96 \cdot SE$$

this will give the lower and upper CL (at the 95% confidence level). Both upper and lower CL's and P 2.5% and P 97.5% (or other optional confidence levels) are output in PASGEAR.



Bootstrap confidence intervals can be constructed on both the sample mean and on Pennington’s estimator and all four estimators can be displayed simultaneously:



Options
☑ Arithmetic estimator
☑ Pennington estimator
☑ Arithmetic Bootstrap
☑ Pennington Bootstrap
☑ Boxplot
☑ Confidence Intervals
☑ Number of observations
☐ Statistics (mean and SE)
☐ Set cut level (Pennington)
☐ Vertical Grid
☐ Grid
☐ Rescale Y-axix
☑ Show Plot Legend
☑ Show Group Legend
☐ Vertical X-Legend

🟥 = Arithmetic mean ± SE, I = mean ± t(95%)·SE number = sample size
🟩 = Pennington's estimator ± SE, I = mean ± t(95%)·SE number= M (# Obs > 0 or cut)
🟨 = Arithmetic Bootstrap ± SE, I = 95% percentiles number= bootstrap runs
🟦 = Pennington Bootstrap ± SE, I = 95% percentiles number= bootstrap runs

Gear selectivity

Indirect estimation of gear selectivity curves

The fish retained in a gear is usually only an unknown proportion of the various size classes available in the fished population. Selectivity is a quantitative expression of this proportion and represented as a probability of capture of a certain size of fish in a certain size of mesh. In PASGEAR 2 gillnet, hook, and trap selectivity is indirectly estimated from comparative data of observed catch frequencies across a series of mesh or hook sizes. The general statistical model (SELECT) is described in Millar (1992), and the specific application on gillnets and hooks is described in Millar and Holst (1997) and Millar and Fryer (1999).

For a given length class, *j*, the number of fish, Y_{ji} , that encounter gillnet *i* are assumed to be observations of independent Poisson variables

$$Y_{ji} = Po(p_i\lambda_j)$$

where the expected count, $p_i\lambda_j$, is the product of the abundance of length class *j* fish, and the relative fishing intensity of gillnet *i*. Fishing intensity can also be considered as a combination of fishing effort and fishing power (Millar 1992).

We denote the relative selectivity (catch or retention probability) of length class *j* fish in gillnet *i* by s_i (*j*). The number of length *j* fish caught in gillnet *i* is then Poisson distributed (Millar and Holst 1997)

$$N_{ji} = Po(p_i\lambda_j\ s_i\ (j))$$

Without loss of generality it can be assumed that the selection curves s_i (·) for each net have unit height because any differences in fishing powers is modeled through the relative fishing intensities p_i. This is the full general model.

In practice the researcher will have to make assumptions about the form of p_i, λ_j, and s_i (·). Options to be considered include (Millar and Holst 1997):

1. If the nets are fished with equal effort, should the relative fishing intensities p_i be assumed equal or some function of mesh size? (see Hamley 1975). In PASGEAR p_i is simply considered equal with standardized effort (i.e. number of settings of standardized panel area and time set).

2. Is it reasonable to postulate a form of the population length distribution by specifying λ_j? In PASGEAR the form of the population length distribution is not assumed.

3. Is the selection curve s_i (·) normal, log-normal, gamma, or perhaps bimodal shaped? Does the principle of geometric similarity apply? (i.e. length of maximum retention and spread of selection curve are both proportional to mesh size, Baranov 1948). In PASGEAR the user can explore this feature and make assumptions about the selection curve.

There are 5 choices of selection curves (all are of unit height):

A) Normal location shift: $\exp\left(-\frac{(L_j - \mathrm{k} \cdot \mathrm{m}_i)^2}{2\sigma^2}\right)$

where only the modes (maximum retention length) is changing with mesh size m_i. Spread is constant.

B) Normal scale shift: $$\exp\left(-\frac{(L_j - k_1 \cdot m_i)^2}{2(k_2 \cdot m_i)^2}\right)$$

where both the modes and the spreads of the selection curves are increasing with mesh size (i.e. the principle of geometric similarity).

The following three models all include asymmetrical retention modes (i.e. skewed distributions). The bimodal curve is appropriate if the fish are caught by different mechanisms, e.g. both wedged by the gills and entangled in the mesh sizes

C) Log Normal: $$\frac{1}{L_j}\exp\left[\mu_1 + \log\left(\frac{m_i}{m_1}\right) - \frac{\sigma^2}{2} - \frac{\left(\log(L_j) - \mu_1 - \log\left(\frac{m_i}{m_1}\right)\right)^2}{2\sigma^2}\right]$$

D) Gamma: $$\left(\frac{L_j}{(\alpha-1)k \cdot m_i}\right)^{\alpha-1} \cdot \exp\left(\alpha - 1 - \frac{L_j}{k \cdot m_i}\right)$$

E) Bi-modal: $$\exp\left(-\frac{(L_j - k_1 \cdot m_i)^2}{2(k_2 \cdot m_i)^2}\right) + w \cdot \exp\left(-\frac{(L_j - k_3 \cdot m_i)^2}{2(k_4 \cdot m_i)^2}\right)$$

where
μ_i = mean size (length) of fish caught in mesh size $i = k_1 \cdot m_i$
σ_i = standard deviation of the size of fish in mesh $i = k_2 \cdot m_i$ or $\alpha \cdot m_i$
L_j = mean size of fish in size (length) class j

In all cases of the above considerations and models, one can use the Poisson distribution of N_{ji} to apply maximum likelihood for purposes of statistical inferences and estimation fits. The code and implementation of a general non-linear maximum likelihood optimizer, and the five optional selection curves, into PASGEAR was done by René Holst, ConStat, Denmark.

In order to explore the possible shapes of the selection curve $S_i(\cdot)$ and whether the principle of geometric similarity seems applicable, there is an option for estimating the mean, standard deviation and degree of skewness for each of the observed catch frequencies in mesh size i (Fig. 10).You can choose between two models or function types: The standard normal (function type 2):

$$E(n_{ij}) = \frac{n_i}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{(L_j - m_i)^2}{2\sigma_i^2}\right)$$

and a skew-normal function (Helser *et al.* 1991,1994), for species that also have some degree of entanglement (function type 1 and default):

$$E(n_{ij}) = \frac{n_i}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{(L_j - m_i)^2}{2\sigma_i^2}\right)\cdot\left\{1 - ½q_i \cdot \sigma_i^{2/3}\left[\frac{(L_j - m_i)}{\sigma_i} - \frac{(L_j - m_i)^3}{\sigma_i^3}\right]\right\}$$

where new symbols are

q_i = skewness coefficient of the distribution of fish in mesh i,
(when $q_i = 0$ the model reduces to the standard normal distribution)
n_{ij} = catch of fish of size class j in mesh i
n_i = total catch of fish in mesh i ($= \sum_j n_{ij}$)

The model parameters (μ_i, σ_i, q_i) are estimated by an iterative numerical search of the minimum sum of squares, ($=\sum(observed - predicted)^2$) between the expected catch based on the model, $E(n_{ij})$, and the observed catch (n_{ij}). The non-linear search algorithm (Fletchers method) is adapted to Pascal from a QBasic program (FLET) supplied with Hilborn and Walters (1992).

It is also possible to evaluate the statistical fit of the chosen selection curve by an examination of the residual plot or by the model deviance (squared sum of residuals). In the residual plot it is expected that the residuals are randomly distributed around 0 (the red line), and have a value smaller than ± 2 (white lines). Residuals with an absolute value bigger than 2 are represented in yellow. Under assumption that the deviance is χ^2 distributed (which is not always the case) a p-value is given with degrees of freedom (# L_{ji}- # parameters). The p-values will in general always be small, so the best measure of goodness of fit is looking at the deviance.

The total relative selectivity (catch or retention probability) of the gear is computed as:

$$S_{ij} = \sum \frac{s_i(j)}{\max_j}$$

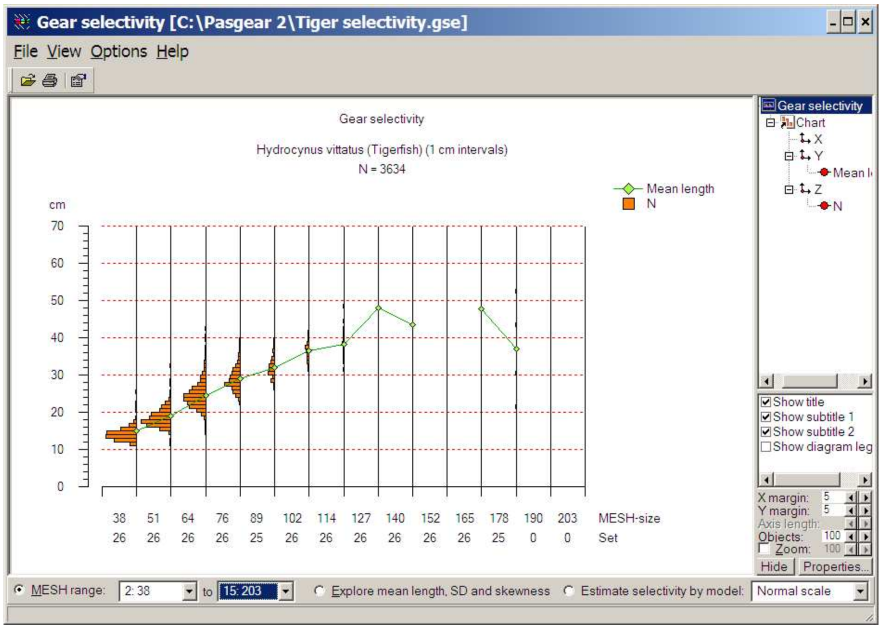
where each estimated mesh specific selectivity curve s_i (j) is weighed by the number of settings (effort) of that net. The total relative selectivity is used to estimate the corrected catch frequency in each size class (N_j) from the observed catches, computed as

$$N_{ij} = \sum \frac{n_{ij}}{S_{ij}} \quad \text{where } S_{ij} = 1 \text{ if } S_{ij} \le 0.1 \text{ (or optionally chosen value)}$$

The total relative selectivity by length groups (S_{ij}) can be stored on a file (ASCII) and used for selectivity corrections in the Analyses. The default name of the ASCII file with selectivity probabilities will consist of a combination of the species Id, the length interval used and the length unit, with the extension *.sel so that several selectivity files can be stored and accessed by the Analyses programs.

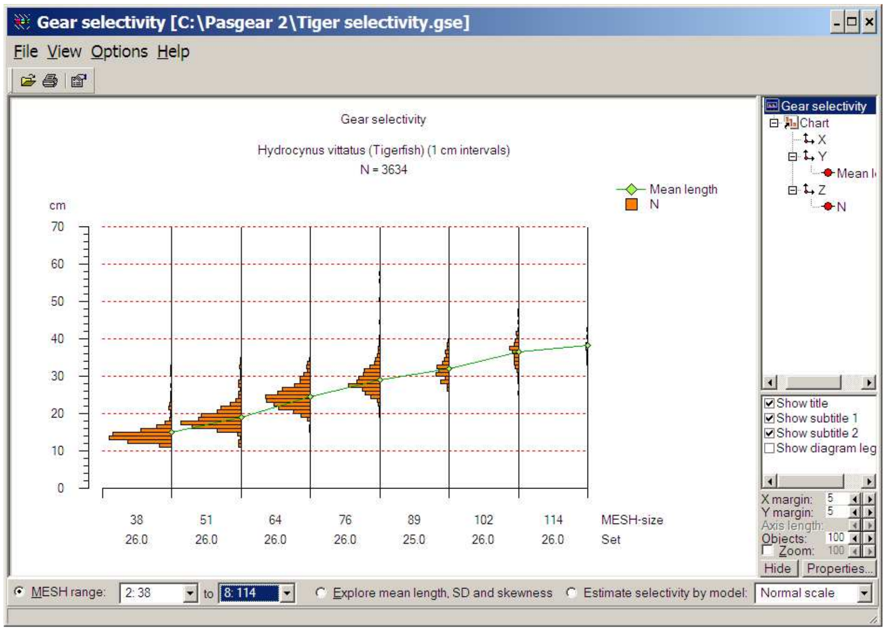
How to use Gear selectivity

1. First you must specify the range of mesh sizes you want to include in the analysis. This is best done by a visual examination of the first diagram



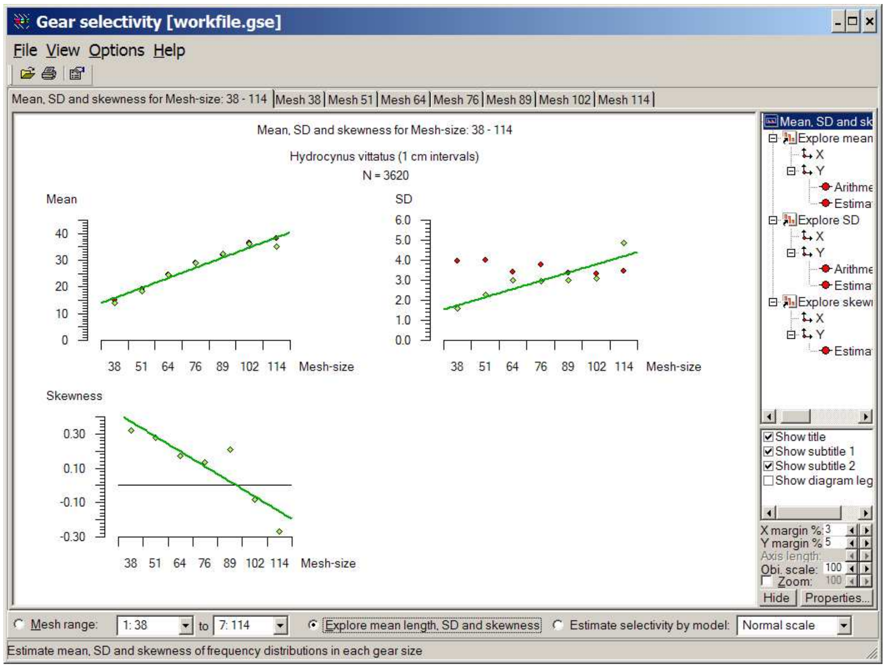
● MESH range: 2: 38 to 15: 203 ○ Explore mean length, SD and skewness ○ Estimate selectivity by model: Normal scale

Select the mesh size range where the mean length is reasonably ‘linear’ and the number of fish is not too small



● MESH range: 2: 38 to 8: 114 ○ Explore mean length, SD and skewness ○ Estimate selectivity by model: Normal scale

2. Then choose Explore mean length, SD and skewness



☑ Show title
☑ Show subtitle 1
☑ Show subtitle 2
☐ Show diagram leg

X margin %: 3
Y margin % 5
Axis length:
Obj. scale: 100
☐ Zoom: 100

○ Mesh range: 1: 38 to 7: 114 ● Explore mean length, SD and skewness ○ Estimate selectivity by model: Normal scale

Estimate mean, SD and skewness of frequency distributions in each gear size

This plot will assist in deciding which selectivity model to choose

- Normal location: where only the modes (maximum retention length) is changing with mesh size. Spread (SD) is constant.

- Normal scale: where both the modes and the spreads of the selection curves are increasing with mesh size (i.e. the principle of geometric similarity).

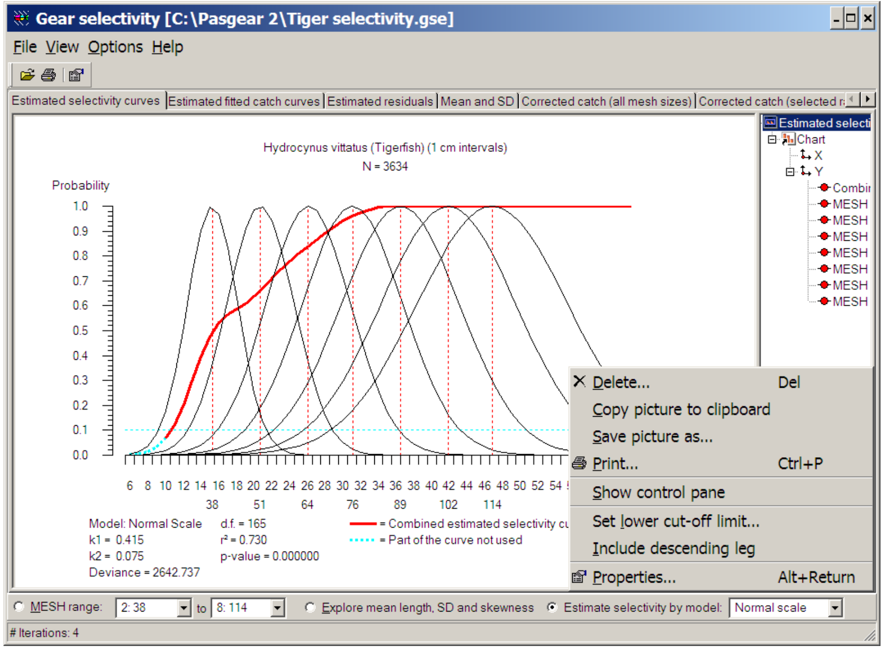
Or the following three models which all include asymmetrical retention modes (i.e. skewed distributions).

- Log Normal
- Gamma
- Bi-modal: The bimodal curve is appropriate if the fish are caught by different mechanisms, e.g. both wedged by the gills and entangled in the mesh sizes

3. When the model has been chosen then click estimate selectivity

After successful iteration the results will be displayed in a series of diagrams

The estimated selectivity curves for each mesh size. In this plot you can set a lower cut-off limit (default = 0.1) and whether to include or not the descending leg of the overall curve (default = not)

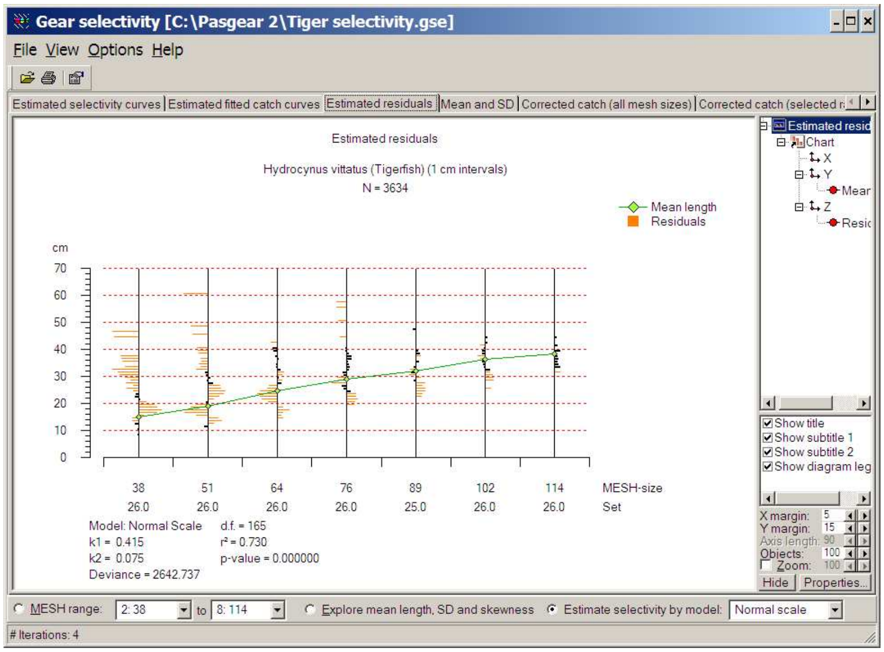


Model: Normal Scale
k1 = 0.415
k2 = 0.075
Deviance = 2642.737
d.f. = 165
r² = 0.730
p-value = 0.000000

○ MESH range: 2: 38 to 8: 114 ○ Explore mean length, SD and skewness ● Estimate selectivity by model: Normal scale

Iterations: 4

The residuals plot by mesh size



Model: Normal Scale d.f. = 165
k1 = 0.415 r² = 0.730
k2 = 0.075 p-value = 0.000000
Deviance = 2642.737

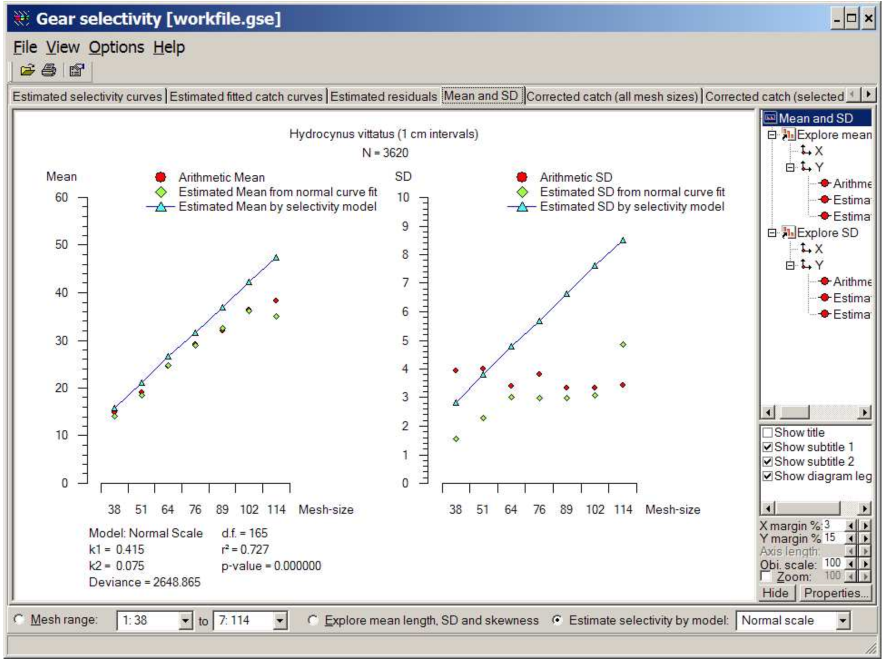
☑ Show title
☑ Show subtitle 1
☑ Show subtitle 2
☑ Show diagram leg

X margin: 5
Y margin: 15
Axis length: 90
Objects: 100
☐ Zoom: 100

○ MESH range: 2: 38 to 8: 114 ○ Explore mean length, SD and skewness ● Estimate selectivity by model: Normal scale

Iterations: 4

The estimated mean and spread to the observations



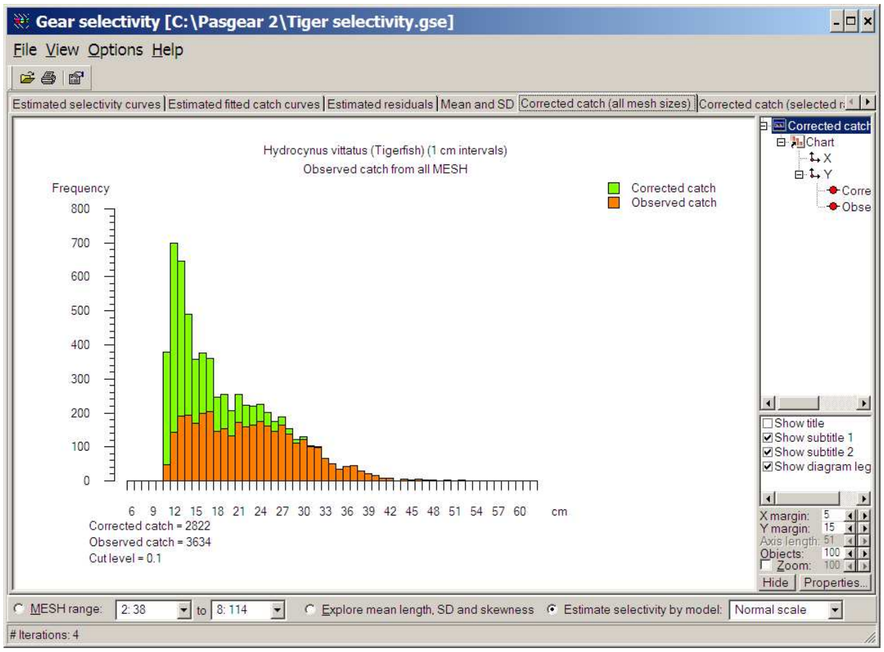
Model: Normal Scale d.f. = 165
k1 = 0.415 r² = 0.727
k2 = 0.075 p-value = 0.000000
Deviance = 2648.865

☐ Show title
☑ Show subtitle 1
☑ Show subtitle 2
☑ Show diagram leg

X margin % 3
Y margin % 15
Axis length:
Obj. scale: 100
☐ Zoom: 100

○ Mesh range: 1: 38 to 7: 114 ○ Explore mean length, SD and skewness ● Estimate selectivity by model: Normal scale

And the corrected catch superimposed on the observed



Corrected catch = 2822
Observed catch = 3634
Cut level = 0.1

☐ Show title
☑ Show subtitle 1
☑ Show subtitle 2
☑ Show diagram leg

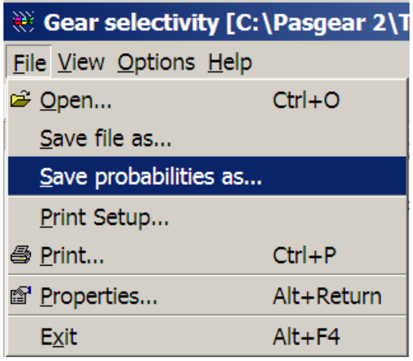
X margin: 5
Y margin: 15
Axis length: 51
Objects: 100
☐ Zoom: 100

○ MESH range: 2: 38 to 8: 114
○ Explore mean length, SD and skewness
● Estimate selectivity by model: Normal scale

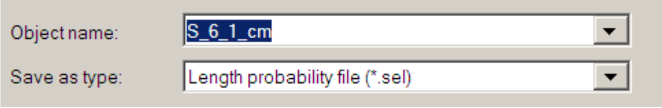
Iterations: 4

Save estimated gear selectivity

Once you are satisfied you can save the results to a file in order to correct for estimated selectivity in your analyses or before exporting to FiSAT.



The default file name will consist of a combination of the species Id, the length interval used and the length unit, with the extension *.sel. This will secure that selectivity curves are not applied on wrong species or with different length intervals or units in the analysis.



Object name: S_6_1_cm
Save as type: Length probability file (*.sel)

References

Bagenal, T. (ed.) 1978. *Methods for assessment of fish production in fresh waters*, IBP handbook No 3, 3d. edition. Blackwell Sci. Publ., Oxford, London, Edinburgh, Melbourne. 365 pp.

Baranov, F. 1948. Theory and assessment of fishing gear. *In* Theory of fishing with gillnets. Chap. 7. Pishchepromizdat, Moscow. (Translation from Russian by Ontario Dept. Of Lands, Maple, Ont. 45 pp).

Begon, M., Harper, J.L. and Townsend, C.R. 1990. *Ecology: Individuals, Populations and Communities*. 2nd. ed. Blackwell Scient. Publ., 945 pp.

Beyer, J.E. 1987. On length-weight relationships. Part I: Computing the mean weight of the fish in a given length class. *Fishbyte, April 1987*:11-13.

Cochran, W.G. 1977. *Sampling Techniques*, 3'rd ed. Wiley, New York.

Conquest, L., Burr, R., Donnelly, J., Chavarria, J. and Gallucci, V., 1996. Sampling methods for stock assessment for small-scale fisheries in developing countries. In: V. F. Gallucci, S. B. Salia, D. J. Gustafson and B. J. Rothschild (Editors), *Stock Assessment: Quantitative Methods and Applications for Small Scale Fisheries*, p. 179-225. CRC Press, New York, NY.

Degerman, E., Nyberg, P. and Appelberg, M. 1989. Estimating the number of species and relative abundance of fish in oligotrophic Swedish lakes using multi-mesh gillnets. *Nordic J. Freshw. Res.* **64**:91-100.

Caddy, J.F. and G.D. Sharp 1986. An ecological framework for marine fishery investigations. *FAO Fish. Tech. Pap. No 283*. FAO, Rome. 151 p.

Efron, B. and Tibshirani, R.J. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science* **2**:54-77.

Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability 57. Chapmann & Hall, London. 436 pp.

Fjälling, A. and Fürst, M. 1991. Introduction of standardized test-fishing methods to Zambian water. *Fisheries Development Series* **53**, ISSN 0280-5375, 41 p.

Gayanilo, F.C., Jr., Soriano, M. and Pauly, D. 1989. A draft guide to the Compleat ELEFAN. *ICLARM Software 2*, ISSN 0116-6964, ICLARM, Manila, 70 p.

Gayanilo, F.C., Jr., Sparre, P. and Pauly, D. 1996. The FAO-ICLARM Stock Assessment Tools (FiSAT) User's Guide. *FAO Computerized Information Series (Fisheries)*. No. 8. Rome, FAO, 266 p.

Gunderson, D.R., Callahan, P. and Goiney, B. 1980. Maturation and fecundity of four species of *Sebastes*. *Mar. Fish. Rev.*, **42**(3-4):74-79.

Hamley, J. M. 1975. Review of gillnet selectivity. *J. Fish. Bd. Can.*, **32**:1943-1969.

Helser, T.E., Geaghan, J. & Condrey, R.E. 1991. A new method of estimating gillnet selectivity, with an example for spotted seatrout, *Cynosion nebulosus*. *Can. J. Fish. Aquat. Sci.*, **48**:487-492.

Helser, T.E., Geaghan, J. and Condrey, R.E. 1994. Estimating size composition and associated variances of a fish population from gillnet selectivity, with an example for spotted seatrout (*Cynosion nebulosus*). *Fish. Res.*, **19**:65-86.

Hilborn, R. and Walters, C.J. 1992. *Quantitative Fisheries Stock Assessment - Choice, Dynamics and Uncertainty*. Chapman and Hall, New York. 570 p.

Kolding, J. 1989. The fish resources of Lake Turkana and their environment - Thesis for the Cand. Scient degree in Fisheries Biology and Final Report of KEN 043 Trial Fishery 1986-1987. University of Bergen, 262 pp.

Krebs, C.J. 1989. *Ecological Methodology*. Harper Collins, New York. 654 p.

Labrosse, P., M. Kulbicki and J. Ferraris. 2002. Underwater Visual Fish Census: proper use and implementation. Noumea, New Caledonia: Secretariat of the Pacific Community. 54 p. ISBN 982-203-878-X. http://www.spc.int

McConnaughey, R. A. and Conquest, L. L., 1992. Trawl survey estimation using a comparative approach based on lognormal theory. *Fish. Bull.*, **91**:107-118.

Millar, R.B. 1992. Estimating the size-selectivity of fishing gear by conditioning on the total catch. *Journal of the Americam Statistical Association* **87**:962-968.

Millar, R.B. and Holst, R. 1997. Estimation of gillnet and hook selectivity using log-linear models. *ICES Journal of Mar. Sci*., **54**:471-477

Millar, R.B. and Fryer, R.J. 1999. Estimating the size-selection curves of towed gears, traps, nets and hook. *Rev. Fish Biol. and Fisheries*, **9**:89-116.

Pauly, D. and David, N. 1981. ELEFAN I, a basic program for the objective extraction of growth parameters from length-frequency data. *Meeresforschung/Rep. Mar. Res.* **28(4)**:205-211.

Pennington, M., 1983. Efficient estimators of abundance, for fish and plankton surveys. *Biometrics*, **39**:281-286.

Pennington, M., 1996. Estimating the mean and variance from highly skewed marine data. *Fish. Bull*. **94**:498-505.

Pinkas, L., Oliphant, M.S. and I.L.K. Iverson 1971. Food habits of albacore, bluefin tuna and bonito in Californian waters. *Fish Bull. Calif. Dep. Fish game*, **152**:1-105.

Rickey, M.H. 1995. Maturity, spawning, and seasonal movement of Arrowtooth flounder, *Atherestes stomias*, off Washington. *Fish. Bull*., **93**(1):127-138.

Sparre, P. 1987. Computer programs for fish stock assessment, length-based fish stock assessment for Apple II computers. *Fao Fish. Tech. Pap. 101*, *suppl. 2*, Fao, Rome, 218 p.

Sparre, P. and Venema, S. 1998. Introduction to tropical fish stock assessment. FAO Fish. Tech. Pap. -T306/1Rev.2, 407 pp. http://www.fao.org/docrep/w5449e/w5449e00.htm

Strømme, T. 1992. NAN-SIS: Software for fishery survey data logging and analysis - User's manual. *FAO Computerized Information Series (Fisheries) No. 4*. FAO, Rome, 103 pp.