

VPABASE: A PROTOTYPE OF DATABASE FOR STORAGE AND PROCESSING OF VPA-INPUT DATA

by

Per Sparre, Ole Folmer and Clara Ulrich

Danish Institute for Fisheries Research,
Charlottenlund Castle, DK 2920 Charlottenlund, Denmark.
email: pjs@dfu.min.dk, ofo@dfu.min.dk and clu@dfu.min.dk

Abstract

The paper describes a prototype of database for storage and processing of Working Group input for fish stock assessment. Input data are total catches by fleet and country combined with age distribution samples (when available). The processing of data applies the available age distribution samples to convert the national catches (landings and discards) into age compositions and sum the national data to total stock age composition, which is the input to the VPA. The database will maintain records on the processing of data, as well as records on misallocated and unreported catches, explaining differences between official data and so-called "WG-data". The database is implemented in ACCES/EXCEL, using a user-friendly interface, including a user's manual. Data may be entered through userforms or transferred from other databases through a file with a specified exchange format. The database was developed as a part of the EU funded project EMAS (Evaluation of Market Sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval, CFP Study Project 98/075).

List of contents

| | | |
|------|---|----|
| 1. | Introduction | 1 |
| 2. | Inputs to stock assessment | 2 |
| 2.1. | Catch at age | 3 |
| 2.2. | Weight at age | 3 |
| 3. | Estimation of catch and weight at age | 4 |
| 3.1. | Biological sampling | 4 |
| 3.2. | Sampled vs. unsampled strata | 4 |
| 3.3. | Estimating catch at age in unsampled strata | 5 |
| 3.4. | SOP-correction..... | 6 |
| 4. | National data and ICES Working group data | 6 |
| 4.1. | Official statistics and working group data | 6 |
| 4.2. | Discards and landings | 7 |
| 4.3. | Misreported and unallocated catch..... | 7 |
| 5. | (Prototype-) software | 8 |
| 5.1. | Tables and relations | 9 |
| 5.2. | User-interface..... | 9 |
| 5.3. | Manual data entry..... | 10 |
| 5.4. | Automatic data Input..... | 11 |
| 5.5. | Processing of data | 11 |
| 5.6. | Reports in Access format | 12 |
| 5.7. | Reports in Excel format | 12 |
| 6. | Data-storage and data-exchange | 13 |
| 7. | Discussion and conclusions | 14 |
| 8. | References | 15 |
| 9. | LIST OF Symbols..... | 15 |
| 10. | Tables..... | 16 |
| 11. | Figures | 21 |

List of Figures.

| | | |
|------------|---|----|
| Figure 1. | Illustration of the gap-filling-process..... | 21 |
| Figure 2. | The basic hierarchy of tables in VPABase..... | 21 |
| Figure 3. | Data Tables and relations of the ACCESS implementation of VPABase..... | 22 |
| Figure 4. | Data-tables, Look-Up-tables and relations of the ACCESS implementation of VPABase , | 23 |
| Figure 5. | Main menu of VPABase. | 24 |
| Figure 6. | Data entry form of VPABase..... | 25 |
| Figure 7. | Menu for data processing | 26 |
| Figure 8. | Data entry form for the selection of age compositions to be used to fill-in unsampled catches | 26 |
| Figure 9. | EXCEL worksheet used by the ICES Herring WG and the Mackerel, Horse mackerel, Sardine and Anchovy WG..... | 27 |
| Figure 10. | Example of report in EXCEL format..... | 27 |

1. Introduction

Biological advice on many commercially exploited stocks in the ICES area relies to a large extent on commercial catch-at-age. The quality of the biological advice is dependent on the quality of the underlying data. However, for most stocks there is hardly any insight into measures of uncertainty associated with the catch-at-age data. Furthermore, for the stocks exploited by various countries, the sampling and raising procedures to estimate national age-composition of catches might differ from one country to another. The methods used for combining national catch statistics into international catch at age are not uniform. The historical disaggregated national catch data are not always accessible, and thus the consistency of data series is not warranted. Thus the whole procedure of estimating international catch at age would gain in efficiency and transparency if a uniform method of aggregation was developed and applied, and if all data used in the estimation were gathered and centralised in one single database.

The objective of this paper is to present a prototype of such a database for the storage and the processing of ICES Working Group input for fish stock assessment. Input data are total catches by fleet (and country), area and time period combined with age distribution samples (when available). It was developed as a part of the EU funded project EMAS (Evaluation of Market Sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval, CFP Study Project 98/075; Pastoors *et al.*, 2001). This project intended to evaluate the adequacy of the international market sampling effort for a limited number of commercially exploited fish stocks in the ICES area, to quantify the uncertainties in estimated catch in numbers and mean weight at age, to advise on appropriate sampling levels and methods and to define protocols and develop prototype software to store and aggregate national sampling data.

The data processing uses the available age distribution samples to convert the national catches (landings and discards) into age compositions, and sums the national data to total stock age composition, which are the main input to the VPA (Virtual Population Analysis). The processing of data implies filling in gaps in data, caused by some countries not supplying full sets of input data. The gaps are filled in by extrapolating samples data to non-sampled catches data (**Figure 1**).

The database is a relational database implemented in ACCESS, chosen for its simplicity of use and because it is a standard software commonly used in fisheries institutes. Although ACCESS might not be designed for storing and manipulating large-scale databases, it was considered a suitable tool for the present application. It allows having a user-friendly interface. Data may be entered through user-forms or transferred from other databases through a file with a specified exchange format. Reports from the database are made in EXCEL sheets. A user's manual is available.

A number of software tools were already developed for data storage and combination of national catch information to the international level:

- (a) In the ICES Herring Assessment Working Group for the Area South of 62° N (HAWG) national data are submitted on agreed spreadsheet forms. The software called SALLOC is used to raise national data to the international level (Patterson 1998). This Fortran program provides a log listing the actual fill-ins used to cover all landings with missing biological information, but this list is not given in the WG report. Also, no list is available in the WG report giving the extrapolations made of each national data set before submitting the data to ICES. The Working Group provides a well-documented overview of the sampling coverage from each country in the report.
- (b) The Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) keeps a database system to store and aggregate data on the flatfish species in the North Sea. The database system (FISHBASE) consists of a user interface programmed in Pascal and a specific datafile format. Data are stored in separate datafiles by country, gear, area, sex (if available) and year. Different levels of aggregations may be used in the datafiles ranging from files, which have catch-numbers, weights at age and lengths at age by quarter, sex, gear, area and

country to files, which only have total landings, by country and year. Raising is performed from the lowest level aggregation upwards. The strong points of Fishbase are that it forces the user to correctly specify the input data and that data can be exchanged in separate datafiles, so that national institutes can submit their data in those files. Also it allows a historical access to the underlying data and other arrangements may be implemented to raise the data. Negative points are that no manual is available, that the program does not allow combination of areas, that the input format is very difficult to generate with any other software than Fishbase itself and that there is no specification of reallocation rules.

- (c) A series of basic programs are used to collect national age compositions and weights at age for roundfish stocks in the North Sea and west of Scotland. The programs were developed in the 1980s in the Marine Laboratory (Aberdeen) and are operated from the dos-prompt by invoking QBASIC. All data is stored in a common format. An exchange format is used for exchange of data. Strong points of the Aberdeen programs are that a rigid file structure is defined and that the raising procedure is carried through to generating the input files for stock assessment. Negative points are that there is a whole suite of programs, which should be used, in a specific order and that there are no explicit reallocation rules.

In the software presented here, the raising of catch-at-age distributions to non-sampled data is done automatically, with the same algorithm as in Patterson (1998). But improvements of the current work are threefold: Firstly, this algorithm is combined directly with the database, simplifying the manipulation and processing of data. Secondly, it is aimed to be a more generic and flexible tool that could be used for a number of different fisheries. Thirdly, it includes a higher level of data-disaggregation to the fleet level. This disaggregation is useful for improving the estimation of total catches at age, by taking into account the differences of exploitation pattern arising from differences in fishing practices.

The database is thus intended to output the inputs to the VPA. We consider here only the usual single-species VPA used in most ICES Working Groups. Multi-species VPA, although used for the assessment of some specific stocks, is not considered at this stage of the database elaboration. We do not consider the VPA tuning process, which involves further specifications of time series of effort and catches for some commercial and surveys tuning fleets. However, it is worth noting that as the database is implemented at the fleet level so that future improvements could easily be provided to store tuning data as well in the database.

In this paper we firstly summarise briefly the main inputs to be used in the VPA, and how they can be calculated from sampling data. We then summarise how and why these inputs are modified during working groups before being used in assessment. This aims to present the full set of parameters that should be included in the database. The database itself is presented afterwards.

2. INPUTS TO STOCK ASSESSMENT

The VPA compute stock numbers and fishing mortality coefficients by age and year, from data of total catches in number and estimates of natural mortality coefficients, by age and year (although natural mortality is most often considered as constant over years). Biomass is thus calculated by multiplying stocks numbers by weight at age.

To facilitate the notation, we introduce the “*Stratum*”. By a *stratum* is meant a combination of a stock St , a country Ct , a fleet Fl , an area Ar and a quarter q in a year y . We use the symbol “ Ω ” for a stratum:

$$\Omega = (St, y, Ct, Fl, Ar, q).$$

The symbol “ S ”, represent a set of strata: $S = \{\Omega_1, \Omega_2, \dots, \Omega_m\}$, where m is the number of elements in the set S . Appendix 1 contains a complete list of symbols used in the formal description.

2.1. Catch at age

The input to single species VPA without tuning, are the numbers caught (landed and discarded) of stock St by age group a and during quarter q in year y , from the total area of distribution of the stock, caught by all fleets exploiting the stock.

$$C_{VPA}(St, y, q, a)$$

We use here the quarter as the time step by default, as it is commonly used in the ICES area (although assessments are afterwards conducted on an annual basis). Any other division of the year, for which data are available, could be used. The age index is comprised in the range $[a_{Min}(St), a_{Max}(St)]$, where $a_{Min}(St)$ (≥ 0) is the youngest age group considered in the VPA and $a_{Max}(St)$ is the oldest age group considered in the VPA. The oldest age group may or may not be a plus-group.

The catch is taken by a number of fleets in a number of sub-areas of the total stock area, so that

$$C_{VPA}(St, y, q, a) = \sum_{Ct=1}^{Ct_{Max}} \sum_{Fl=1}^{Fl_{Max}(Ct)} \sum_{Ar=1}^{Ar_{Max}} Catch(St, y, q, a, Ct, Fl, Ar) = \sum_{Ct=1}^{Ct_{Max}} \sum_{Fl=1}^{Fl_{Max}(Ct)} \sum_{Ar=1}^{Ar_{Max}} Catch(\Omega, a) \quad (1)$$

where $Catch(\Omega, a) = Catch$ (Landings + discards) in numbers of age group a in a stratum Ω .

$Fl = 1, 2, \dots, Fl_{Max}(Ct)$, $Ar = 1, 2, \dots, Ar_{Max}$, where $Fl_{Max}(Ct) =$ Number of fleets in country Ct and $Ar_{Max} =$ Number of areas.

The area can be ICES sub-division, ICES division or ICES area. In a later version of the database, the area concept could be generalized to include ICES rectangles, and/or specific areas defined by working groups. Fleets are given homogeneous subsets of boats, whose definition is made externally by the Working Group (e.g. by fishing gear and/or boat length class...)

In general the dot “•” is used to indicate summation over one index, e.g. when data are not disaggregated by quarter but are available only on an annual basis :

$$Catch(St, y, \bullet, a, Ct, Fl, Ar) = \sum_{q=1}^4 Catch(St, y, q, a, Ct, Fl, Ar)$$

2.2. Weight at age

Estimated stock numbers are converted into stock and spawning stock biomasses by using weight and maturity coefficients by age and year. Predictions of yield are made from mean fishing and natural mortality at age and mean weight at age. The weight at age is thus a major parameter of the assessment procedure. It relates to a specific time, e.g. the beginning, middle, or mean over the time period. In subsequent equations the weight at age is noted :

$$W_c(St, y, q, a, Ct, Fl, Ar) = W_c(\Omega, a) = \text{Mean body weight in Catch of age group } a \text{ in a stratum.}$$

Number of fish caught is calculated as :

$$Catch(\Omega, a) = \frac{Yield(\Omega, a)}{W_c(\Omega, a)} \quad (2)$$

Where $Yield(\Omega, a)$ is the catch in weight of stock St and age a , in a stratum. The total yield $Yield(\Omega)$ must be available for all strata, to make the following estimation feasible..

3. ESTIMATION OF CATCH AND WEIGHT AT AGE

3.1. Biological sampling

Catch and weight at age are estimated from biological sampling data, which provide detailed information on length and/or age composition of the landings in a stratum. Data include the number of otoliths read and the number of fish measured by age. The associated coefficients of variation may be available. Sampling data by age and stratum in the database are thus :

$na(\Omega, a)$ = Number of age measurements (otoliths);
 $nl(\Omega, a)$ = number of length measurements;
 $CV_C(\Omega, a)$ = Coefficient of variation of the estimate of $Catch(\Omega, a)$;
 $CV_{w_C}(\Omega, a)$ = Coefficient of variation of the estimate of $w_C(\Omega, a)$

3.2. Sampled vs. unsampled strata

A *sampled stratum* is a stratum in which biological sampling was conducted. The catch numbers at age and the weights at age could be directly estimated. The total yield is also available. An *un-sampled stratum* is a stratum where only the total yield is known, and no sampling data are available. The estimation of age composition and weight-at-age for un-sampled strata has thus to be done using information from the sampled strata.

The set of all strata may be split into two, the sampled strata ($S^{sampled}$) and the un-sampled strata, ($S^{un-sampled}$). The number of sampled and unsampled strata are noted m_S and m_U respectively. A sampled and unsampled stratum is noted Ω^S and Ω^U respectively. Thus, the set of all strata, can be partitioned as (“ \cup ” indicates the union of two sets):

$$S^{All} = S^{sampled} \cup S^{un-sampled} \quad S^{Sampled} = \left\{ \Omega_1^S, \Omega_2^S, \dots, \Omega_{m_S}^S \right\} \quad S^{Un-Sampled} = \left\{ \Omega_1^U, \Omega_2^U, \dots, \Omega_{m_U}^U \right\}$$

In a sampled stratum Ω^S , the total number caught of all age groups is:

$$Catch(\Omega^S, \bullet) = \sum_{a=a_{Min}(st)}^{a_{Max}(st)} Catch(\Omega^S, a) \quad (3)$$

The proportion of each age group in the total catch number in the stratum is:

$$Pr_{Catch}(\Omega^S, a) = Catch(\Omega^S, a) / Catch(\Omega^S, \bullet) \quad (4)$$

The mean weight of all fish in the catch is

$$W_C^{Mean}(\Omega^S) = \sum_{a=a_{Min}(st)}^{a_{Max}(st)} Pr_{Catch}(\Omega^S, a) * W_C(\Omega^S, a) \quad (5)$$

The computed total weight of the catch (SOP, Sum Of Products) is:

$$SOP_{Catch}(\Omega^S) = \sum_{a=a_{Min}(st)}^{a_{Max}(st)} Catch(\Omega^S, a) * W_C(\Omega^S, a) = W_C^{Mean}(\Omega^S) * Catch(\Omega^S, \bullet) \quad (6)$$

3.3. Estimating catch at age in unsampled strata

The age composition within an unsampled stratum can be estimated using the age distribution of other strata, which can be assumed to have a comparable pattern.

Let Ω_i^U be an un-sampled stratum, and let Ω_j^S be a sampled stratum. If we assume that weight-at-age and relative catch-at-age are the same for Ω_i^U and Ω_j^S , then we can estimate the total number caught in the unsampled stratum as:

$$\text{Catch}(\Omega_i^U, \bullet) = \text{Yield}(\Omega_i^U) / W_C^{\text{Mean}}(\Omega_j^S) \quad (7)$$

and

$$\text{Catch}(\Omega_i^U, a) = \text{Catch}(\Omega_i^U, \bullet) * \text{Pr}_{\text{catch}}(\Omega_j^S, a) \quad (8)$$

It may also be preferable sometimes to use a subset of various sampled strata to estimate catch distribution in an unsampled stratum, instead of only one as described in the previous equations. The estimated distribution is thus a weighted mean of the age distributions in sampled strata. The weighting coefficients might be uniform (equal weight for all strata), or linked to the importance of each stratum (total yield, number of fish measured/aged, inverse of variance... etc...).

The subset of sampled strata used for the estimation is denoted S_k ($S_k \subset S^{\text{Sampled}}$) and has k elements, $S_k = \{\Omega_1^S, \Omega_2^S, \dots, \Omega_k^S\}$.

we define: $\text{SW}_{\text{Mt}}(S_k, a) = \{v_{\text{Mt}}(\Omega_1^S, a), v_{\text{Mt}}(\Omega_2^S, a), \dots, v_{\text{Mt}}(\Omega_k^S, a)\}$

where SW is a set of weights ($v_{\text{Mt}}(\Omega, a)$) associated with the set of sampled strata and age group a , selected by a weighting method Mt .

The weighted mean proportion at age in the unsampled strata is then:

$$\text{Pr}_{\text{Catch}}^{\text{Mean}}(\Omega_i^U, a, S_k, \text{Mt}) = \frac{\sum_{\Omega^S \in S_k} \text{Pr}_{\text{Catch}}(\Omega^S, a) * v_{\text{Mt}}(\Omega^S, a)}{\sum_{a=a_{\text{Min}}(\text{St})}^{a_{\text{Max}}(\text{St})} \sum_{\Omega^S \in S_k} \text{Pr}_{\text{Catch}}(\Omega^S, a) * v_{\text{Mt}}(\Omega^S, a)} \quad (9)$$

The calculation of a weighted mean proportion at age is only possible when the weighting factors are available for all strata in S_k . The methods $\text{Mt} = \text{equal weight}$ (all $v = 1.0$) or $\text{Mt} = \text{yield}$ are applicable for any set of strata. The method $\text{Mt} = \text{number of otoliths}$ requires having an associated number of otolith measurements for all strata.

Note that the procedure to calculate weighted mean proportions at age may also be applied to sampled strata. The otolith samples taken from a catch stratum are usually used to estimate the age composition of that stratum. However, if the sample is very small or for some other reason suspicious, it may be better to use alternative otolith samples, or to combine the suspicious sample with samples of other strata.

The weighted mean body weight in the catch is then :

$$W_C^{\text{Mean}}(\Omega_i^U, a, S_k, Mt) = \frac{\sum_{\Omega^S \in S_k} W_C(\Omega^S, a) * v_{Mt}(\Omega^S, a)}{\sum_{\Omega^S \in S_k} v_{Mt}(\Omega^S, a)} \quad (10)$$

Then we can calculate the total number of individuals caught (using Eq. 7) and then the number of individuals per age group (using Eq. 8).

A list of possible weighting options for the calculations of proportions at age and mean weights in the catch is presented in **Table 1**. The list of options for weighting factors may be extended in future versions of VPABase. Obvious extensions of the list includes characteristics of ALKs (Age length keys), such as number of length measurements and number of ALKs.

3.4.SOP-correction

The SOP is the sum of the products of catch numbers at age times the weights at age:

$$SOP_{\text{Catch}}(\Omega) = \sum_{a=a_{\text{Min}}(St)}^{a_{\text{Max}}(St)} \text{Catch}(\Omega, a) * w_C(\Omega, a) \quad (11)$$

Ideally, the SOP should be equal to the Yield(Ω) which was given as input. If that is not the case, the equality can be achieved by applying a “SOP-correction”, either to the catch or to the weight

$$\text{Yield}(\Omega) = \sum_{a=a_{\text{Min}}(St)}^{a_{\text{Max}}(St)} \text{Catch}(\Omega, a) * \text{Correction}(\Omega) * W_C(\Omega, a) \quad (12)$$

$$\text{where } \text{Correction}(\Omega) = \frac{\text{Yield}(\Omega)}{SOP_{\text{Catch}}(\Omega)}$$

The SOP correction was discussed by Lewy & Lassen (1997), who concluded that the SOP is an unbiased estimate of the total catch, and that SOP corrections should not be made. We consider four options in the case of SOP discrepancies:

- 1) Correct the catch-at-age,
- 2) Correct the weight-at-age,
- 3) Correct the input Yield (replace the value of input Yield by the value of SOP),
- 4) Do nothing.

According to Lewy and Lassen, option 3 is recommended. In the following, it is assumed that any SOP-problem has been resolved before the data are entered into the database. If however, there is a discrepancy, the value Yield will be automatically changed to the SOP (Option 3).

4. NATIONAL DATA AND ICES WORKING GROUP DATA

The annual yield by country (summed over fleets) and ICES area are reported by member governments to ICES as the Official statistics. These statistics are however sometimes modified by members of assessment working groups. This section discusses the discrepancies between official statistics and the so-called “working group data”, as all the data included in the “working group data” have to be included in the database.

4.1. Official statistics and working group data

Let Landings $W_{\text{Official}}(St, Ct, Ar)$ be the official statistics, i.e. the weight of the annual landings of stock St by country Ct in ICES area (sub-area or division) Ar . Members of ICES working groups deliver the so-

called *working group data*, which may be more disaggregated than the official data and which may also be different. The question is whether the official data are consistent with the working group data, that is

$$\text{Landings } W_{\text{Official}}(\text{St}, \text{Ct}, \text{Ar}) = ? = \sum_{\text{Fl}=1}^{\text{Fl}_{\text{Max}}(\text{St})} \sum_{\text{Ar} \in \text{ICES-Div.}} \sum_{\text{q}=1}^4 \text{Yield}(\text{St}, \text{Fl}, \text{Ar}, \text{q}) \quad (13)$$

Reasons for discrepancies might include the non-integration of discards data, as well as corrections made on misreported data.

The following fields are used by the Working Groups (ICES, 2000), and are integrated in the VPABase :

| | |
|-------------------------------|---|
| Official Catch | Catches as reported by the official statistics to ICES |
| Unallocated Catch | Adjustments to the official catches made for any special knowledge about the fishery, such as under- or over-reporting for which there is firm external evidence. (can be negative) |
| Area misreported Catch | Used only to adjust official catches which have been reported from the wrong area. (can be negative). For any country the sum of all the area misreported catches should be zero |
| Discarded Catch | Catch which is discarded |
| WG Catch | The sum of the 4 categories above |

4.2. Discards and landings

Estimated catches (in numbers) and yield (in weight) is composed of landings and discards:

$$\begin{aligned} \text{Catch}(\Omega, a) &= \text{Landings } N(\Omega, a) + \text{Discard } N(\Omega, a), \\ \text{Yield}(\Omega) &= \text{Landings } W(\Omega) + \text{Discard } W(\Omega) \end{aligned}$$

The weight at age of discards and landings are usually different, the discard weight-at-age being smaller than the landings weight-at-age:

$$\begin{aligned} W_D(\Omega, a) &= \text{average body weight of discards.} \\ W_L(\Omega, a) &= \text{average body weight of landings.} \end{aligned}$$

The body weights relate to the body weight-at-age in the catch by:

$$W_C(\Omega, a) = \frac{W_L(\Omega, a) * \text{Landings } N(\Omega, a) + w_D(\Omega, a) * \text{Discards } N(\Omega, a)}{\text{Landings } N(\Omega, a) + \text{Discards } N(\Omega, a)} \quad (14)$$

whereby:

$$\text{Landings } W(\Omega) = \sum_{a=a_{\text{Min}}(st)}^{a_{\text{Max}}(St)} w_L(\Omega, a) * \text{Landings } N(\Omega, a)$$

and:

$$\text{Discards } W(\Omega) = \sum_{a=a_{\text{Min}}(st)}^{a_{\text{Max}}(St)} w_D(\Omega, a) * \text{Discards } N(\Omega, a)$$

4.3. Misreported and unallocated catch

When including the terms for unallocated and misreported landings as defined in the text table above, the following equations should always hold.

$$\begin{aligned}
\text{Catch}W_{WG}(St, Ct, Ar) = & \text{Landings}W_{Official}(St, Ct, Ar) + \\
& \text{Unallocated}W(St, Ct, Ar) + \\
& \text{Misreported}W(St, Ct, Ar) + \\
& \text{Discard}W(St, Ct, Ar)
\end{aligned} \tag{15}$$

where

$$\sum_{Ar \in ICES-Div.} \text{Misreported}(St, Ct, Ar) = 0$$

and

$$\text{Catch}W_{WG}(St, Ct, Ar) = \sum_{Fl \in Fleets(Ct)} \sum_{Ar \in ICES-Div.} \sum_{q=1}^4 \text{Yield}(St, Fl, Ar, q) \tag{16}$$

This equation must be met by all input data of each country. If it is not met, the Yield(St, Fl, Ar, q) will be corrected by the database system, using the country and area specific correction factor:

$$\text{Correction}WG(St, Ct, Ar) = \frac{\text{Landings}W_{Working\ Group}(St, Ct, Ar)}{\sum_{Fl \in Fleets(Ct)} \sum_{Ar \in ICES-Div.} \sum_{q=1}^4 \text{Yield}_{BeforeCorrection}(St, Fl, Ar, q)} \tag{17}$$

where $\text{Yield}_{BeforeCorrection}(St, Fl, Ar, q)$ and $\text{Yield}_{AfterCorrection}(St, Fl, Ar, q)$ stand for Yield before and after correction, respectively.

$$\text{Yield}_{AfterCorrection}(St, Fl, Ar, q) = \text{Yield}_{BeforeCorrection}(St, Fl, Ar, q) * \text{Correction}WG(St, Ct, Ar) \tag{18}$$

However, basically, working group members should supply the data, so that no correction would be necessary. The program will always compute the correction factor, and the user should check the value. The database will keep the original value (before correction), for the purpose of documentation.

5. (PROTOTYPE-) SOFTWARE

A prototype database system was developed under the working title of “VPABase”. The prototype version of VPABase was implemented in Microsoft ACCES and EXCEL (version 2000). The term “Prototype” means a “Working model in the early stage of the design process”. The prototype does run and executes some of the tasks specified in the foregoing sections. The prototype, however, is far from the stage where it is safe to distribute to ICES WGs. It has not yet been conscientiously tested, and so far no feedback has been received from the potential users, the ICES WGs. Thus, an unknown number of man-months of development, testing and consultations with the clients are still pending, before public distribution should be made.

A high degree of user-friendliness is aimed at in the design of user-interface and the output-options of VPABase. This feature has been attempted from the early stage of the development. Primary versions of a user’s manual and a reference manual of the VPABase has been prepared. This Section summarizes the two manuals.

The database can read input data from Excel-formats and also produce reports in Excel format. The example of EXCEL work sheet used in the present version of VPABase, has been adopted from the spreadsheets used by the Herring WG, and the Mackerel, Horse Mackerel, Sardine and Anchovy WG (ICES, 2000).

Input and output can also be handled out with plain ASCII text files (comma separated).

The VPABase is constructed as a relational database according to the standard database-design rules (normalization rules).

5.1. Tables and relations

The tables and the relations of VPABase consist of: “Input-data-tables”, “Look-up tables” and “Auxiliary tables”. The look-up tables are used to select standard input values, whenever possible, as that will reduce the risk of mistakes on data entry.

The principal data-tables of VPABase, contains the numbers at age and weight at-age distributions and the yields (weight of catches). The three principal data-tables of VPABase are:

1. Country record
2. Fleet record
3. Age composition record

In addition there is the auxiliary data tables, containing information on the data-processing:

4. All combinations of catch and samples for estimation of age composition.
5. Selected samples for estimation of age composition.

The tables are linked in a hierarchy of “one-to-many” relationships. Table 2 shows the fields in the input data-tables and the auxiliary tables. The first fields in the tables, e.g. “Key_Country_Catch” in the table “T_Country_Catch”) are the keys linking the tables in the hierarchy.

Table 2 indicates the relationships between the tables, which is also illustrated by Figure 3, which by arrows indicates the relations of the three principal data-tables, and the auxiliary tables.

Figure 4 shows the complete set of look-up tables together with the three principal input data tables.

A complete description of all fields, keys and relations of VPABase is rather voluminous. The reader interested in further details is referred to the reference manual of VPABase.

The auxiliary tables “T_Use_of_Samples” and “T_Use_Of_samples_Saved” contain the input, which specifies the processing of data. It contains information on which sampled records were used to estimate age-compositions of the un-sampled records. They also contains information on the methodology of data processing as well as the time of (last) processing. This part of the database is perhaps the most complicated part of it, as the number of choices for combination of sample and processing method is usually large. A major objective of VPABase is to store the information on data-processing in an unambiguous way.

5.2. User-interface

The purpose of the user interface is to guide the user through three steps:

1. Input data;
2. Make decisions on estimation of age-compositions for un-sampled strata by selecting age-compositions among samples strata;
3. Generate output tables.

The user interface implemented through user-forms that are of two types:

1. Menus (selection of options)
2. Data entry form;

The five first items of the main menu (Figure 5) guides the user through the different steps to produce the aggregated input data for VPA. Options 1 and 2 are alternative options, namely (1) to “Enter data from the keyboard” (see next section) or (2) “Enter data from an EXCEL worksheet” (see Section 5.4). Item 1 in the main menu ends with “.....” which indicates that this option leads to another user-form (see Figure 6). Item 3 activates the user-form for data processing. Items 5 and 6 deal with the reports, that is the output. Output can be in ACCESS-format (Item 4) or in EXCEL-format (Item 5).

In later versions of VPABase, the main menu will contain an option for entry of tuning data (survey data, commercial fishery effort data or commercial fishery CPUE).

5.3. Manual data entry

The first item in the main menu opens a data input form (Figure 6) where data can be filled in manually. One country record might contain several fleet records and one fleet record might contain several sampling records. The form in Figure 6 is actually composed of three forms, which correspond to levels of the hierarchy of tables (see Figure 6 or Table 2)

First level – Annual catch by country: The top part of the data-entry-form is related to the country catch record and the record selector in the bottom of the form tells the user which country catch record is current. The record selector is the field called “Record:” and the arrow buttons gives the user a way to go the first record, move one step back, move one step forward, to the last record and to a new record.

Second level - Catch by fleet and time period: The second part of the data-entry-form contains the fleet record, with fleet related information. The fleet information in question relates to the country record, in which the fleet record is embedded. Ideally, the sum of the fleet records over time periods of a country should sum up to the value of the country record.

Third level – Age composition sample: The third part of the form contains the age composition data (samples), relating to the fleet record in which it is embedded. The sample is automatically given a unique “sample number”. This number cannot be changed by the user and is only used to identify a sample for the later processing of samples.

The data entered through look-up tables, appear as “list-boxes” in the user-form. When clicking on the right hand side arrow-button, a list of options is displayed, from which you can select the input. (The example given here is taken from Figure 6, showing the entry of “Stock”). Only input from the list is accepted. Only users with special “privileges” can change the content of the list box. This approach should reduce the risk of mistyping and ambiguities in data.

In later versions of VPABase, the data entry form will contain a number of additional facilities, such as production of graphs and tables as well as facilities for “finding” records and for deleting records.

Later versions of VPABase will also be able to handle catch data by statistical rectangle.

Entry of tuning data in future versions of VPABase, will probably be done in a separate userform, allowing for entry of catch and effort or CPUE.

5.4. Automatic data Input

The second item on the main menu lets the user import data from a predefined EXCEL format or a text file (see example of EXCEL input in Figure 9). The present version of VPABase can read the EXCEL sheets currently used by the Herring WG and the Mackerel, Horse mackerel, Sardine and Anchovy WG (ICES, 2000).

In later version of VPABase, input data may also be transferred in ACCESS table formats from one replicate of VPABase to the master version of VPABase, which operation should be managed by an international data-coordinator.

5.5. Processing of data

The third item in the main-menu, “Select samples from un-sampled catch, and process data” opens a submenu (Figure 7) for data processing.

If a fleet-record has an age composition sample attached to it (is a “sampled catch”), the user does not need to take further action as far as processing is concerned. In that case the VPABase will make the estimation of the age composition of the total catch from that sample. The user need only take action in the two cases:

- 1) When no age composition samples is available,
- 2) When the user wants to use additional age composition samples, for a sampled catch.

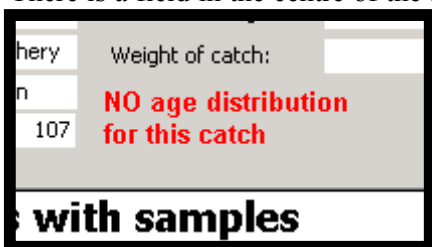
The action to take is to select one or more samples to represent the unsampled catch. If more than one sample is selected, the user will also have to specify which method should be used to pool the selected samples (how the weighted average should be calculated).

Before any actual selection of samples can be made, the user has to create a table to hold the tables of potential selections (the data-table “T_Use_Of_Samples, See Figure 3).

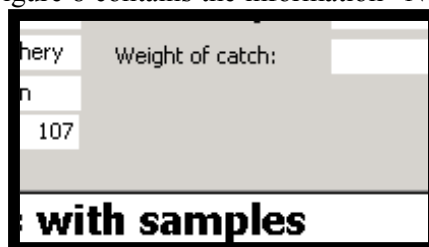
After having created the table of potetial samples the user can move to the next item on the processing-menu and start making selection for un-sampled strata (Item 2: “Select samples for unsampled catch”). The userform for selction of samples is shown in Figure 8.

The upper left part of this form contains a number of fields (headed “Catch record”) showing the particulars of the catch record to be processed. These fields are shown only for information. The fields are “locked”, so that the data cannot be entered or modified through this form.

There is a field in the centre of the form, which in the example of Figure 8 contains the information “No age distribution for this catch” (written in red). This means that the catch of the current record is un-sampled, and you need to select samples, for estimation of age composition. If there is a an age composition sample, the field will be empty.



A screenshot of a software form. At the top left, there are fields for 'Hery' and 'Weight of catch:'. Below these, there is a field containing the number '107'. In the center of the form, the text 'NO age distribution for this catch' is displayed in red. At the bottom of the form, the text 'with samples' is visible.



A screenshot of a software form, similar to the one above. It shows the same fields for 'Hery', 'Weight of catch:', and the number '107'. However, the central field for age distribution is empty. At the bottom of the form, the text 'with samples' is visible.

The form has a check field in the right column to indicate which sample is selected for a given catch presented in the top portion of the form. The user can change the selection by removing or adding check marks on the form.

In the hypothetical example of Figure 8 ,two samples (no. 138 and 139) originating from “United States industrial fishery” is used to represent Danish industrial fishery.

In the upper right hand corner of the form it is possible to change weighting method, for pooling of samples. In the example of Figure 8, the samples have been weighted by the “Sampled catch”. The weighting methods only applies if more than one sample is selected for a given catch. The default method is to give equal weight to each sample, but if additional information, i.e. numbers of otoliths, is present this can be used when the samples are combined.

After the selection has been made, the user moves back to the processing menu where she/he can save the selection to a permanent table (data table “T_Use_Of_Samples_Saved”, see Figure 3) so the selection scheme is stored for later documentation of the process.

Default for the creation of the table is that if a catch record has a sample, this sample will be used for this catch (but there is the option not to use the default: Option 2 above).

5.6.Reports in Access format

The ACCESS system can create reports (or “output-tables) on its own. The ACCESS-format tables, however, are not very flexible with respect of additional processing, and change of style (see example in Table 4). Tables in EXCEL format is believed to be more attractive for members of ICES working groups.

Reports in Access format are inventory reports of the database. They contain information on the number of records and other features of the input data and the national sampling programs. Prototypes of the inventory reports are given in Table 4 (country records) and Table 5 (fleet records).

5.7.Reports in Excel format

These are the standard tables suggested for (potential) inclusion in the ICES Assessment working group reports. Ideally, the assessment-software should be able to take its input directly from these tables. The present proto-type version of VPABase, actually contains very few table. Most of the tables listed here cannot be produced by VPABase as the system is now. The authors of the present paper, await the reaction and the wishes from the possible “costumers” of VPABase, before they will start the work of designing output tables. So what follows is what the present authors expect will become the requirements from ICES WGs, but this still has to be tested by consultations.

The Excel reports of VPABase can be grouped as:

- 1) Input data, as provided by WG-members and governments.
 - a. Country records (before processing)
 - b. Fleet records (before processing)
- 2) Processing: Tables showing the methods and input
- 3) Results of processing
 - a. Country records (after processing)
 - b. Fleet records (after processing)
- 4) VPA input (Landings, Discards and Catch)

In principle, the database can display the data with any sorting or any grouping on the user choice. Some tables, however, are expected to be in demand and for these there will be pre-prepared routines to produce them.

These reports will comprise the traditional tables for official statistics of ICES working group reports, as well as the “non-official” data. The example in Table 6 shows a table for all country-records of a given

ICES area. The system will be able to rearrange the tables on the users choice, for example to give a report by country with each row representing an ICES area.

Table 7 shows a proto-type of fleet-record table. There may or may not be age/weight-composition data available for a stratum. If age/weight-composition data are not available, then total catch (or landings) may be available, in which case only the last line in the table will be filled. The fleet records have unique samples numbers (*Sample No. y*), which are used for identification of the fleet-records when used for estimation of non-sampled strata.

An important requirement for the database system is to document the age composition estimation procedure. To document the process of using age compositions from other strata, means to record from where the “borrowed” information originates, and which method of weighting was applied. Table 8 shows a report which for each fleet record contains a log of its processing. Each row in the table represents a stratum. The first column gives the unique index of the stratum. The following 5 columns gives the particulars of the sampled stratum. The columns *Landing Samp.* and *Discard Samp.* Indicates if these catches were covered by the sampling program. The following columns contain the selection of “borrowed” samples used for estimation of landings- and discards age compositions. The first line contains the methods of weighting the data, if average values of several strata were used.

For the processed fleet records, the output will be a table similar to the output table for the sampled strata (Table 7). The table will contain information on the processing, extracted from Table 9 The results for the country, is either the sum over fleets or the estimation based on “borrowed” catch-distributions (Table 10).

The input to the stock assessment process, is the summation over country-records (Table 11). The only EXCEL report produced by the proto-type version of VPABase is the VPA-input as shown in Figure 10.

6. DATA-STORAGE AND DATA-EXCHANGE

Data are stored in tables and relations in the relational database, as described above. It is tradition in ICES to store data in so-called *flat-files*, that is a data-storage where each data record can be read independently, and where no relations between data are specified.

Data stored as flat-files usually takes up much more space than data stored as tables with relations. Should a flat-file be required, for example for analysis by the SAS, the database can easily produce such a flat-file.

Data can be submitted from countries to the central database in three ways.

1. Entry of data in replicate database
2. Entry of data in EXCEL format (Using a fixed standard set-up)
3. Entry of data from screen through the data entry forms

The data that is presented in Excel format, will be imported by the data base and added to the data already on the system. Alternatively, the database could have been implemented on the internet, and data could be entered directly in the central database. This solution is rather straight forward to implement.

The supply of data in a traditional “flat-file” format (e-g- SAS files) will be accommodated in the future. The proto-types of the Excel data-exchange formats corresponds to the tables of VPABase: Country file (Table 12) and Fleet/age file (Table 13)

The fleet-table and the age-distribution table of VPABase have been combined into one exchange file. The fields in this table may not all be filled in. For example, if catch-at-age and weight-at-age are given, the landings (or discards) are given as the sum of products, and these cells do not have to be filled in. If they are, the SOP will have priority over the landings (discards) in tonnes.

This version of VPABase is not an “online-system”, and when a record is updated, it is not immediately updated in the international network. Each institute of participating in the ICES assessment working

group will get a replica of the VPABase. A replica is simply a copy of VPABase, for which the design cannot be modified. Before the annual WG meeting, the master VPABase will be synchronised with all the replicas. To “synchronise” two replicas of a database, means to exchange all new records and updates of records. This task could be carried out by a data coordinator.

7. DISCUSSION AND CONCLUSIONS

A general conclusion from the whole project EMAS (Pastoors et al., 2001), was that the sampling procedures in the different countries have very different characteristics, adapted to the local circumstances, the availability of sampling options and the behaviour of the fishery. There is no obvious method for standardization of sampling procedures available. The implications of this uncertainty on stock assessment results are relatively minor when considering only the total catch at age matrix, but are larger when using commercial CPUE to tune the assessment model.

The storage of all basic data used in the estimation of catch at age is of great importance to keep track of the procedures used and record the consistency of sampling data. This is not yet the case in ICES, where different working groups use different methods and software, and where catch data are exchanged only at the national level of aggregation and with different format. The database presented here should therefore improve the homogeneity, standardisation, centralisation, transparency and flexibility of data used in assessment. It is generic and thus suitable for a great variety of fisheries. It allows accounting for various levels of aggregation and various scales of basic data. It gathers both the database and the program performing all calculations until the final VPA-input within one single software, with a great level of specification provided by the user. It has been designed with a great user-friendliness and easiness to use. It is not a rigid structure, as all user can have access to the programming code and, possibly, improve it for his own use.

However, VPABase is, at this stage of implementation, only a prototype, that requires both improvements and testing before being considered as possible standard software in ICES. The first improvement would be the ability to output not only the matrices of catches and weight at age, but also the tuning files. That would require extending the fleet record level to account for effort data, and also integrating survey data.

The second improvement would be to develop VPABase into a fully-integrated software that would perform the whole assessment procedure, instead of producing only input files to be used in existing software. That would avoid all problems due to format specifications and changes from one software to another, as it is currently the case in ICES working groups. The usual software used in fish stock assessment could be rewritten in a module in VPABase (e.g. in Visual Basic), which would become a full assessment package. Here again, the advantage of using ACCESS is the possibility for the user to see and modify the programming code, and thus to use the database and assessment programs with high flexibility and transparency.

The third improvement would be to add a plat-form independent module to the database, and to improve its access and use. Although ACCESS is widely used by a number of institutions, problems arise when using another operating system, or when updating the database (then a replica has to be sent to each individual partner each time). The best solution would be to provide an Internet access to VPABase, allowing using it from any kind of remote computer.

Finally, the database should be extensively tested by a number of working groups and fisheries research institutes before being validated as a standard tool. So far, the database has been tested only at DIFRES (Danish Institute for Fisheries Research). We believe that this tool could be of great help within ICES working groups, and we hope that it will encounter interest enough to be tested and used.

8. REFERENCES

ICES, 2000 a). Report of the working group on the assessment of mackerel, horse mackerel, sardine and anchovy. ICES C.M. 2000 / ACFM:5.

ICES 2000 b). Report of the Herring Assessment Working Group for the Area South of 60°N. ICES CM 2000/ACFM:10.

Lewy, P., Lassen, H., 1997. Should total landings be used to correct estimated catch in numbers or mean weight-at-age? ICES J. Mar. Sci., 54: 462-470.

Pastoor, M.A., O'Brien, C.M., Flatman, S., Darby, C.D., Maxwell, D., Simmonds, E.J., Degel, H., Vinther, M., Sparre, P., Vanderperren, E., 2001. Evaluation of MARKET Sampling strategies for a number of commercially exploited stocks in the North Sea and development of procedures for consistent data storage and retrieval (EMAS). Final Report of the EU study no. 98/075. RIVO (The Netherlands), CEFAS (UK), DFU (Denmark), SOAEFD (Scotland), CLO-DZ (Belgium).

Patterson, K. R. (1998). A Programme For Calculating Total International Catch-at-age and Weight at Age.

9. LIST OF SYMBOLS

| | |
|--|--|
| a | Age group |
| $a_{Max}(St)$ | Oldest age group of stock St, which may be a plus group |
| $a_{Min}(St)$ | Youngest age group of stock St (≥ 0) |
| Ar | Area index, $Ar = 1, 2, \dots, Ar_{Max}$ |
| Ar_{Max} | Ar_{Max} = Number of areas. |
| $Catch(\Omega, \bullet)$ | The total number caught of all age groups in stratum Ω . |
| $Catch(\Omega, a)$ | Catch (Landings + discards) in numbers of age a in stratum Ω |
| Ct | Country |
| $CV_C(\Omega, a)$ | Coefficient of variation of the estimate of $Catch(\Omega, a)$ |
| $CV_{wC}(\Omega, a)$ | Coefficient of variation of the estimate of $W_C(\Omega, a)$ |
| $C_{VPA}(St, y, q, a)$ | Input to VPA (Working group data) |
| $DiscardN(\Omega, a)$ | Numbers discarded in stratum Ω , age group a (Working group data) |
| $DiscardW(\Omega, a)$ | Yield discarded in stratum Ω (Working group data) |
| F1 | Fleet index, $F1 = 1, 2, \dots, F1_{Max}$ |
| $F1_{Max}(Ct)$ | $F1_{Max}$ = Number of fleets of country Ct |
| $LandingsN(\Omega, a)$ | Numbers landed in stratum Ω , age group a (Working group data) |
| $LandingsW(\Omega, a)$ | Yield landed in stratum Ω , age group a (Working group data) |
| $LandingsW_{Official}(St, Ct, Ar)$ | Weight of annual landings by country Ct. Official statistics delivered from governments to ICES. |
| m_S | Number of sampled strata |
| M_t | Method of weighting sampled strata to estimate unsampled distributions |
| m_U | Number of un-sampled strata |
| $n_a(\Omega, a)$ | Number of age measurements (otoliths) |
| $n_l(\Omega, a)$ | Number of length measurements |
| $Pr_{Catch}(\Omega, a)$ | The proportion made of age group in the total number |
| $Pr_{Catch}^{Mean}(\Omega_1^U, a, S_1, M_t)$ | Weighted mean, using the set of samples S, and the weighting method M_t : |
| q | Quarter (or index of any other division of the year) |
| S | $S = \{\Omega_1, \Omega_2, \dots, \Omega_m\}$, set of m strata. |
| S^{All} | Set of all samples. $S^{All} = S^{sampled} \cup S^{un-sampled}$ |
| $S^{Sampled}$ | Set of sampled strata. |

| | |
|------------------------------|--|
| $S^{\text{un-Sampled}}$ | Set of un-sampled strata. |
| $SOP_{\text{Catch}}(\Omega)$ | Computed weight of the catch in a stratum Ω (Sum Of Products) |
| St | Stock (a sub-set of a species) |
| $SW_{Mt}(S_k)$ | Set of weights associated with the set of sampled strata, selected by method Mt. $SW_{Mt}(S_k) = \{v_{Mt}(\Omega_1^S), v_{Mt}(\Omega_2^S), \dots, v_{Mt}(\Omega_k^S)\}$ |
| $W_C(\Omega, a)$ | Mean body weight in Catch of age group a in stratum Ω . |
| $w_C^{\text{Mean}}(\Omega)$ | The mean weight of all fish in the catch in stratum |
| y | Year |
| Yield(Ω) | Yield, Catch in weight in stratum Ω , (Working group data) |
| Ω | Stratum: $\Omega = (St, Ct, Fl, Ar, y, q)$. combination of Stock St, Country Ct, Fleet Fl, Area Ar, year y and quarter q. |
| Ω^S | Sampled stratum . |
| Ω^U | Un-sampled stratum. |
| $v_{Mt}(\Omega, a)$ | Weighting factor for weighted mean of age group a, using method Mt |

10. TABLES

Table 1. List of options for weighting the calculation of mean proportions at age and mean weight in the catch.

| Mt | Method | $V_{Mt}(\Omega, a)$ | Comment |
|-----------|------------------------------|--|---|
| 1 | Equal weight | 1.0 | No weighting |
| 2 | Catch | Catch(Ω, a) | Numbers caught (Cannot be used if "own" sample is used) |
| 3 | Landings | LandingsN(Ω, a) | Numbers landed |
| 4 | Discards | DiscardsN(Ω, a) | Numbers discarded |
| 5 | Yield (Landings or discards) | Yield(Ω) (LandingsW(Ω) or DiscardsW(Ω)) | Weight of catch (Landings or discards) |
| 6 | Number of otoliths | $n_C(\Omega, a)$ | Problematic (see below) |
| 7 | Inverse CV (Age) | $1/CV_C(\Omega, a)$ | Only for catch-at-age (e.g. from Bootstrap) |
| 8 | Inverse CV(weight) | $1/CV_{wC}(\Omega, a)$ | Only for weight-at-age (e.g. from Bootstrap) |

Table 2. Fields in input-data-tables of VPABase.

| COUNTRY CATCH RECORD (Name in ACCESS: T_Country_Catch , see Fig. 3) | |
|---|--|
| Name of Field | Comment |
| Key_Country_Catch | Unique key of country record |
| Country | Look-up, name of country |
| Species | Look-up (A species can be separated into "Stocks") |
| Year | ..., 1999,2000,2001,.... |
| ICES Sub division | Look-up, name of ICES subdivision |
| ICES Division | Look-up, Not filled in if above filled in |
| ICES Area | Look-up, Not filled in if above filled in |
| Official Landings | As reported to ICES by governments |
| Unreported landings | WG members estimate of unreported landings |
| Misallocated landings | WG members estimate of landings |
| Discards | WG members estimate of discards |
| Sampled Catch | The weight of the catch covered by samples |

(Continued...)

Table 3.(Continued) Fields in input-data-tables of VPABase.

FLEET CATCH RECORD (Name in ACCESS: **T_Fleet_Catch**, see Fig. 3)

| Name of Field | Comment |
|--------------------------|---|
| <i>Key_Catch</i> | Unique key of fleet record. Also Pointer to sample (if existing) which was taken from this catch. |
| <i>Key_Country_Catch</i> | Pointer to country record (one-to-many) |
| Region (of country) | Part of country where home ports of fleets are (optionally) |
| Month | Look-up |
| Quarter | Look-up, Not filled in if month filled in |
| Fleet | Look-up, Sub set of fishing harvesting sector of country |
| Stock | Look-up, Subset of "Species" |
| Use of Catch | Look-up, Discards/Landings |
| Date or recording | The date the record was entered (or modified) in the database |
| Weight of Catch | Weight of landings or discards (catch = landings + discards is not input) |
| Number of samples | Number of age composition samples (=0 if no samples available) |
| Method of weighing | Index for weighing method. Default value = 1 for equal weight for each sample, updated when new selection is made |
| Otoliths | Sum of number of otoliths, updated when new selection is made |
| length_measurements | Sum of numbers of length measurements, updated when new selection is made |

AGE COMPOSITION RECORD (SAMPLE) (Name in ACCESS: **T_Age_Distribution**, see Fig. 3)

| Name of Field | Comment |
|-----------------------------|---|
| <i>Key_Age_Distribution</i> | Unique key of age sample record |
| <i>Key_Catch</i> | Pointer to fleet record (one-to-many) |
| Age group | 0,1,.... |
| Coefficient of variation | Variation of estimate of this age group (e.g. From bootstrapping) |
| Number of otoliths | Number of otoliths in age composition sample in this age group |
| Total number in catch | In discards or landings (raised numbers) |
| Mean weight | Mean weight of age group during time period in question |

POTENTIAL USE OF SAMPLES FOR ESTIMATION OF AGE COMPOSITION *Auxiliary data-table, to contain information on the potential selection of samples to represent fleet records (catches). Contains all combinations of fleet-record and Age-composition-Record, whether selected or not.*

Name in ACCESS: **T_Use_of_Samles**,(see Fig. 3)

| Name of Field | Comment |
|----------------------|---|
| <i>Key_Catch_Age</i> | Pointer to age composition record (sample) |
| <i>Key_Catch</i> | Pointer to fleet record |
| Sample selected | Yes/No Yes: If a sample is selected to represent a catch, that is if the record "Key_Catch_Age" is selected to represent the fleet record "Key_Catch" |
| No sample | Yes/No. No: If no sample (age composition record) has been selected for the fleet Record "Key_Catch" |

ACTUAL USE OF SAMPLES FOR ESTIMATION OF AGE COMPOSITION

Data-table, containing information on the actual selection of samples to represent fleet records (catches). Contains only combinations of fleet-record and Age-composition-Record, which are selected for estimation of age composition.

Name in ACCESS: **T_Use_of_Samles_Saved**,(see Fig. 3)

| Name of Field | Comment |
|----------------------|---|
| <i>Key_Catch_Age</i> | Pointer to age composition record (sample) |
| <i>Key_Catch</i> | Pointer to fleet record |
| Sample selected | Yes/No Yes: If a sample is selected to represent a catch, that is if the record "Key_Catch_Age" is selected to represent the fleet record "Key_Catch" |
| Weighting method | Look-up. Method used to compute weighted average of two or more samples. |
| Time for selection | The time and date the selection was made (in the database) |

Table 4. Inventory of Country records (Example of report in ACCESS-format).

Inventory of Country Records

| Country: Denmark | | | | | | | | |
|-------------------------|----------------------|--------------------------|-----------------------|---------------------|------------------|-----------------------|-----------------------------|---------------------------------|
| Division: | Sub-Division: | Official landning | Mis-Allocated: | Un-reported: | Discards: | No. of fleets: | No of sampled fleets | Procent of catch sampled |
| IVa | IVa East | 3379 | 0 | 0 | 0 | 4 | 0 | 100% |
| IVa | IVa East | 14154 | 3506 | 0 | 0 | 4 | 1 | 166% |
| IVa | IVa Wes | 652 | 0 | 0 | 0 | 4 | 0 | 100% |
| IVa | IVa Wes | 22152 | 3506 | 0 | 0 | 4 | 0 | 116% |
| IVb | All | 3906 | 0 | 0 | 0 | 4 | 0 | 30% |
| IVc | All | 10648 | 3506 | 0 | 0 | 4 | 0 | 133% |

| Country: Netherlands | | | | | | | | |
|-----------------------------|----------------------|--------------------------|-----------------------|---------------------|------------------|-----------------------|-----------------------------|---------------------------------|
| Division: | Sub-Division: | Official landning | Mis-Allocated: | Un-reported: | Discards: | No. of fleets: | No of sampled fleets | Procent of catch sampled |
| IVa | IVa Wes | 1000 | 100 | 100 | 200 | 1 | 1 | 1400% |

Table 5. Inventory of fleet records

| | | LANDINGS | | | DISCARDS | | |
|-----------|---------|-----------------|--------------------|--------------------|-----------------|--------------------|--------------------|
| | | Total yield | Number of otoliths | Coeff.of Vari-ance | Total discards | Number of otoliths | Coeff.of Vari-ance |
| Country A | Fleet 1 | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | Fleet 2 | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | ... | ... | ... | ... | ... | ... | ... |
| | Fleet x | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| Country B | Fleet 1 | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | Fleet 2 | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | ... | ... | ... | ... | ... | ... | ... |
| | Fleet x | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | | | | | | | |
| Country X | Fleet 1 | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | Fleet 2 | Yes/No | n | Yes/No | Yes/No | n | Yes/No |
| | ... | ... | ... | ... | ... | ... | ... |
| | Fleet x | Yes/No | n | Yes/No | Yes/No | n | Yes/No |

Table 6. Report (proto-type): Country records. Official and non-official catch data.

| Country records of ICES Division No. X | | | | | |
|---|--------------------------|-------------------|----------------------|-----------------|--------------|
| | Official Landings | Unreported | Mis-allocated | Discards | Total |
| Country 1 | | | | | |
| Country 2 | | | | | |
| | | ... | | | |
| Country X | | | | | |
| Total | | | | | |

Table 7. Report (proto-type): Fleet record (sample). Catch-at-age, weight-at-age and additional information on sampling (number of otoliths and coefficients of variation).

| Country A | | Fleet a | | ICES Division No. x | | (sample No. y) | | | |
|--|---------------|-------------------|--------------------|--------------------------|-----------------|-------------------|--------------------|--------------------------|---------------------|
| Age gr. | LANDINGS | | | | DISCARDS | | | | TOTAL Number caught |
| | Number landed | Body wgt Landings | Number of otoliths | Coefficient of variation | Number Discards | Body wgt Discards | Number of otoliths | Coefficient of variation | |
| 0 | | | | | | | | | |
| 1 | | | | | | | | | |
| ... | | | | | ... | | | | |
| Max a | | | | | | | | | |
| TOTAL | SOP | Mean Wgt | Total | - | SOP | Mean Wgt | Total | - | SOP |
| If data above are not available, then total catc, landings or discards (in weight) may be available: | | | | | | | | | |
| Yield | | Not avail. | Not avail. | Not avail. | | Not avail. | Not avail. | Not avail. | |

Table 8. Report (proto-type): Processing of fleets records or country records with incomplete samples.

| | | | | | | Fleet record No. X | | | Fleet record No. Y | |
|---------------------|-------|-------|-----------|-------------|-------------|---------------------|----------------------|-------|---------------------|---------------------|
| Method of weighting | | | | | | Method no. | Method no. | | Method no. | Method no. |
| Sam-ple | Ct | Fleet | ICES-Div. | Land, Samp. | Disc. Samp. | Landings Samp. Used | Discards Sampl. Used | | Landings Samp. Used | Discard Sampl. Used |
| 1 | | | | Yes/No | Yes/No | Yes/No | Yes/No | | Yes/No | Yes/No |
| 2 | | | | Yes/No | Yes/No | Yes/No | Yes/No | | Yes/No | Yes/No |
| | | | | | | | | | | |
| Z | | | | Yes/No | Yes/No | Yes/No | Yes/No | | Yes/No | Yes/No |

Table 9. Report (proto-type): Results of processing of fleet-records and country-records.

| Country A | | Fleet a | | ICES Division No. x | | |
|------------|---------------------|----------------------|------|---------------------|-------------------|---------------------|
| Age gr. | LANDINGS | | | DISCARDS | | TOTAL Number caught |
| | Number landed | Body wgt Landings | | Number Discards | Body wgt Discards | |
| 0 | | | | | | |
| 1 | | | | | | |
| ... | | | | ... | | |
| Max a | | | | | | |
| TOTAL | SOP | Mean Wgt | | SOP | Mean Wgt | SOP |
| Estimation | Method no. | Method no. | | | | |
| Sample | Landings Samp. Used | Discards Sampl. Used | | | | |
| 1 | Yes | Yes | | | | |
| 2 | Yes | No | | | | |
| | | | | | | |
| Z | Yes | Yes | | | | |

Table 10. Report (proto-type): Country total of fleet records.

| Country A All Fleets All ICES Divisions | | | | | | | |
|---|---------------|-------------------|--------------------|---------------|-------------------|--------------------|---------------------|
| Age gr. | LANDINGS | | | DISCARDS | | | TOTAL Number caught |
| | Number landed | Body wgt Landings | Number of otoliths | Number landed | Body wgt Discards | Number of otoliths | |
| 0 | Σ | Σ | Σ | Σ | Σ | Σ | Σ |
| 1 | Σ | Σ | Σ | Σ | Σ | Σ | Σ |
| ... | | | | ... | | | |
| Max a | Σ | Σ | Σ | Σ | Σ | Σ | Σ |
| TOTAL | SOP | Mean Wgt | - | SOP | Mean Wgt | Total | SOP |

Table 11. Input table to stock assessment (last column).

| Age gr. | Number Landed | Number Discarded | TOTAL INPUT TO VPA |
|---------|---------------|------------------|--------------------|
| 0 | Σ | Σ | Σ |
| 1 | Σ | Σ | Σ |
| ... | | ... | |
| Max a | Σ | Σ | Σ |

Table 12. Proto-type of Country record for data exchange.

COUNTRY RECORD

| | |
|-------------------|---|
| Country | Denmark |
| Area | IV (Given only if Div or sub-div not filled in) |
| Division | IVc (Given only if sub-div not filled in) |
| Sub-division | IVc East |
| Year | 2001 |
| Official Landings | 9000 |
| Unreported | 7000 |
| Misallocated | 3000 |
| Discards | 5000 |

Table 13. Proto-type of fleet record for data exchange.

FLEET RECORD

| Country | Denmark | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---------------|---|--------------------|-------------------------------------|----------------------------|-------------------------------------|----------------------------|------------|-----|---|------|----|------------|------|----|------|---|------------|------|----|------|---|------------|-----|---|------|----|---------------|------|----|------|----|---------------|-----|---|------|----|---------------|---|---|--|--|---------------|---|---|--|--|
| Fleet | Trawlers (Can be all Fleets, or set of fleets) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Area | IV (Given only if Div or sub-div not filled in) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Division | IVc (Given only if sub-div not filled in) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Sub-division | IVc East | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Year | 2001 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Time period | Q1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Land. (tons) | 29135 (when SOP available, this field need not be filled in) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Disc. (tons) | 38692 (when SOP available, this field need not be filled in) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table border="1"> <thead> <tr> <th></th> <th>Total number landed/ discarded</th> <th>Number of Otoliths</th> <th>Average Body Weight (during period)</th> <th>Coefficient of Variation %</th> </tr> </thead> <tbody> <tr> <td>Land Age 1</td> <td>564</td> <td>6</td> <td>1.10</td> <td>12</td> </tr> <tr> <td>Land Age 2</td> <td>1128</td> <td>12</td> <td>1.50</td> <td>5</td> </tr> <tr> <td>Land Age 3</td> <td>1034</td> <td>11</td> <td>2.00</td> <td>8</td> </tr> <tr> <td>Land Age 4</td> <td>282</td> <td>3</td> <td>2.20</td> <td>15</td> </tr> <tr> <td>Discard Age 1</td> <td>2883</td> <td>12</td> <td>0.95</td> <td>16</td> </tr> <tr> <td>Discard Age 2</td> <td>512</td> <td>8</td> <td>1.15</td> <td>25</td> </tr> <tr> <td>Discard Age 3</td> <td>0</td> <td>0</td> <td></td> <td></td> </tr> <tr> <td>Discard Age 4</td> <td>0</td> <td>0</td> <td></td> <td></td> </tr> </tbody> </table> | | Total number landed/ discarded | Number of Otoliths | Average Body Weight (during period) | Coefficient of Variation % | Land Age 1 | 564 | 6 | 1.10 | 12 | Land Age 2 | 1128 | 12 | 1.50 | 5 | Land Age 3 | 1034 | 11 | 2.00 | 8 | Land Age 4 | 282 | 3 | 2.20 | 15 | Discard Age 1 | 2883 | 12 | 0.95 | 16 | Discard Age 2 | 512 | 8 | 1.15 | 25 | Discard Age 3 | 0 | 0 | | | Discard Age 4 | 0 | 0 | | |
| | Total number landed/ discarded | Number of Otoliths | Average Body Weight (during period) | Coefficient of Variation % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Land Age 1 | 564 | 6 | 1.10 | 12 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Land Age 2 | 1128 | 12 | 1.50 | 5 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Land Age 3 | 1034 | 11 | 2.00 | 8 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Land Age 4 | 282 | 3 | 2.20 | 15 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Discard Age 1 | 2883 | 12 | 0.95 | 16 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Discard Age 2 | 512 | 8 | 1.15 | 25 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Discard Age 3 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Discard Age 4 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

11. FIGURES

Figure 1. Illustration of the gap-filling-process.

The “borrowing” of samples is illustrated in the figure by the broken lines from a sample to a catch-age-composition.

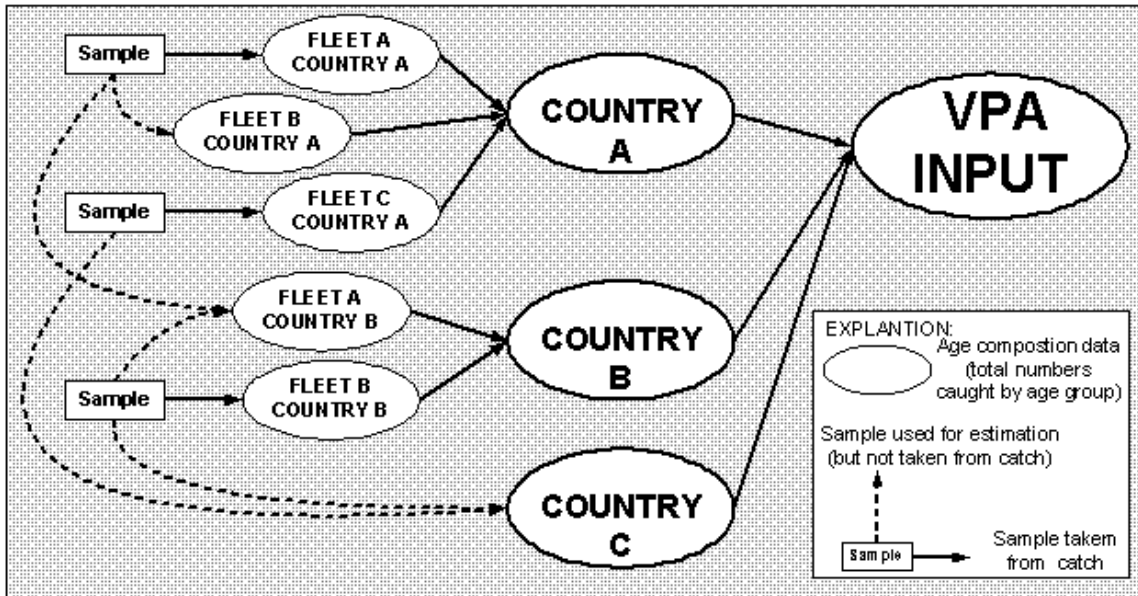


Figure 2. The basic hierarchy of tables in VPABase

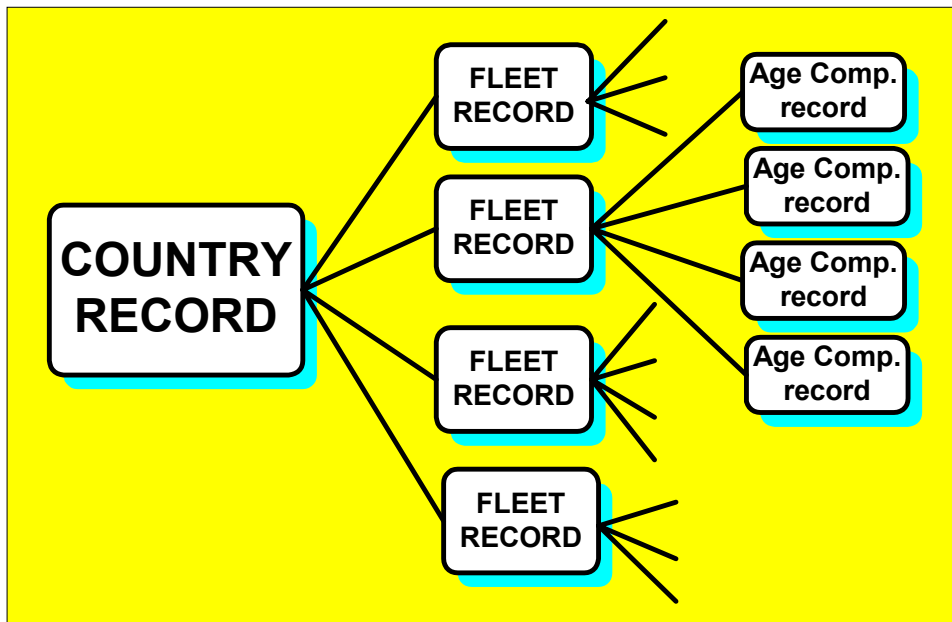
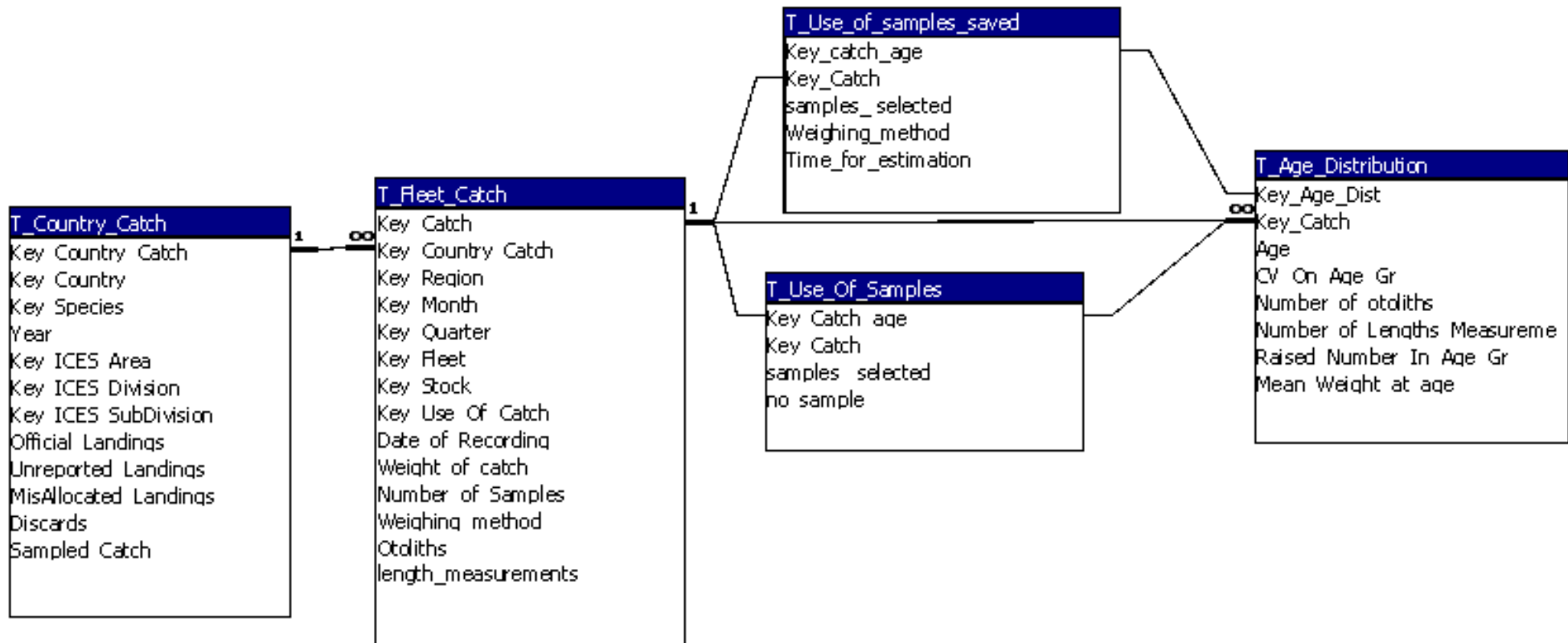


Figure 3. Data Tables and relations of the ACCESS implementation of VPABase

(Compare Table 1)

Note that the two tables “T_Use_Of_Samples” and “T_Use_Of_Samples_Saved”, are the tables containing information on the data-processing, whereas the other tables contain “observations” from fishery.



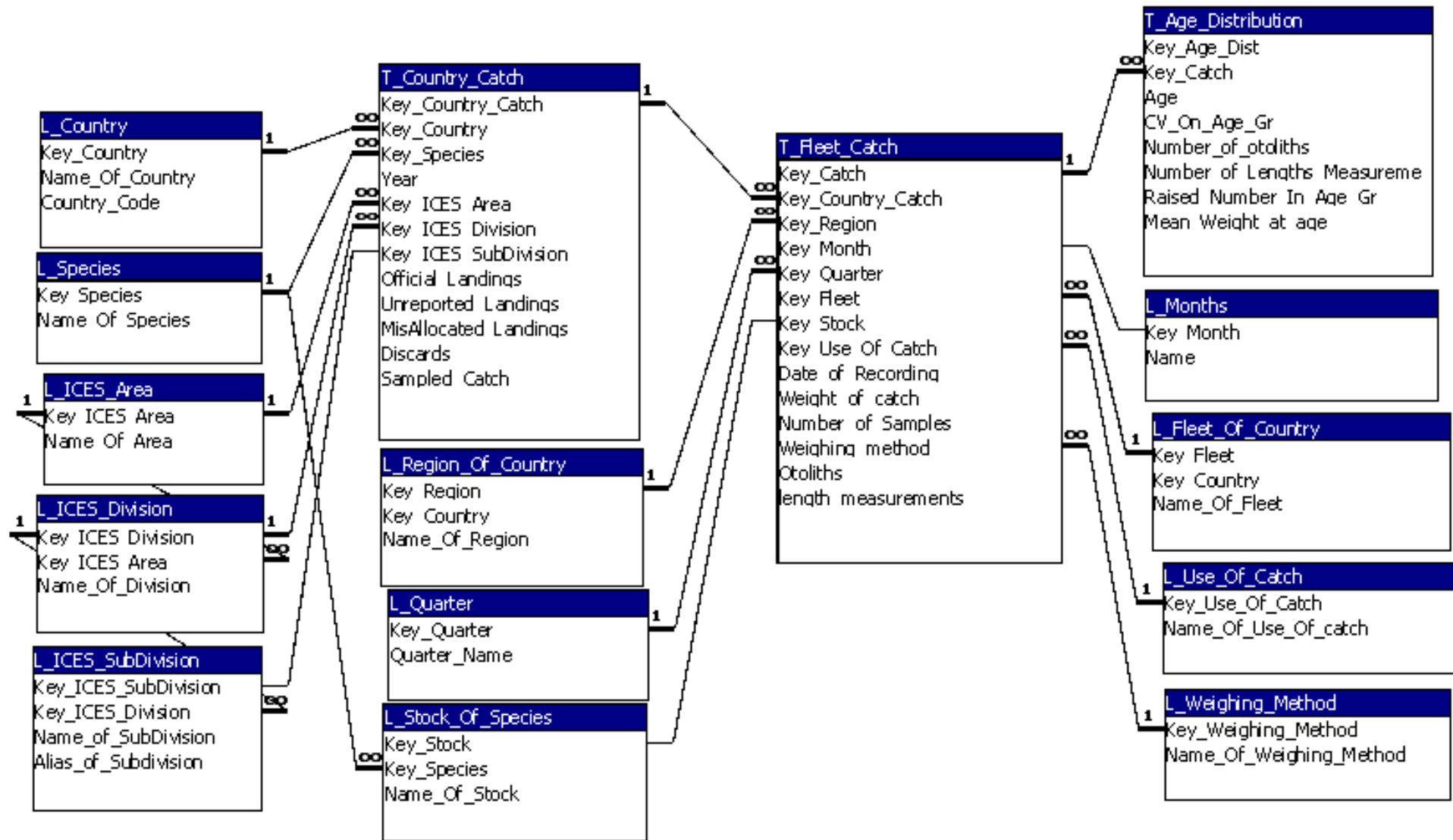


Figure 4. Data-tables, Look-Up-tables and relations of the ACCESS implementation of VPABase , (omitting here the two auxiliary tables “T_Use_Of_Samples” and “T_Use_Of_Samples_Saved”).

Figure 5. Main menu of VPABase.

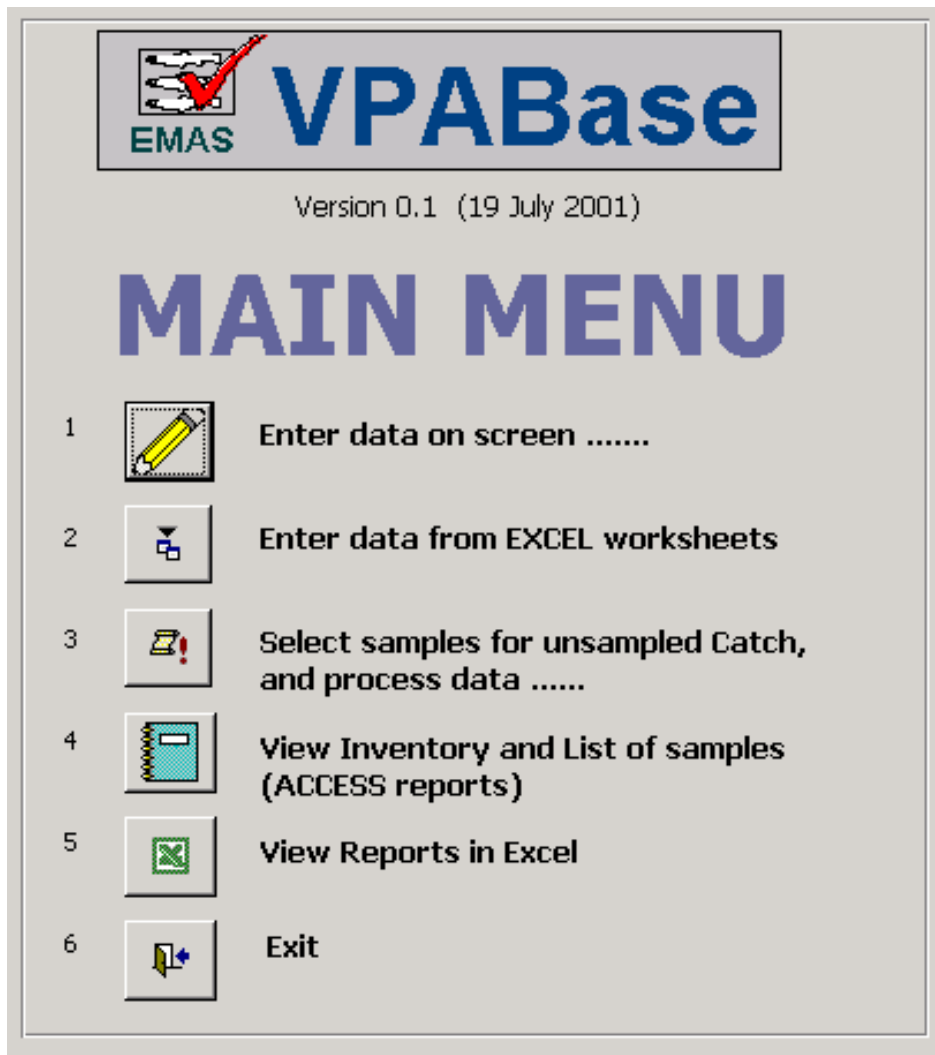


Figure 6. Data entry form of VPABase.

catch
VPABase

Data entry

ANNUAL CATCH BY COUNTRY

Country:

Year:

Species:

ICES Sub Area:

ICES Division:

ICES Sub Division:

Official Landings: (tons)

Unreported Landings:

MisAllocated Landings:

Discards:

Sampled Catch:

Save and Exit

Add new record

CATCH BY FLEET AND TIME PERIOD

Quarter:

Fleet:

Stock:

Use Of Catch:

Number of samples:

Weight of Catch:

Date of Recording:

AGE COMPOSITION SAMPLE

| W-rings | Catch No.('000) | Body Weight (kg) | No. Lengths M. | No otoliths | CV (%) | Sample No. |
|---------|-----------------|------------------|----------------|-------------|--------|------------|
| 0 | 0 | 0 | 0 | 0 | 0 | 98 |
| 1 | 8800 | 0,015 | 0 | 0 | 0 | |
| 2 | 2930 | 0,044 | 0 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 0 | 0 | |
| 4 | 0 | 0 | 0 | 0 | 0 | |
| 5 | 0 | 0 | 0 | 0 | 0 | |
| 6 | 0 | 0 | 0 | 0 | 0 | |

Record: 1 of 4

Record: 1 of 12

Figure 7. Menu for data processing.

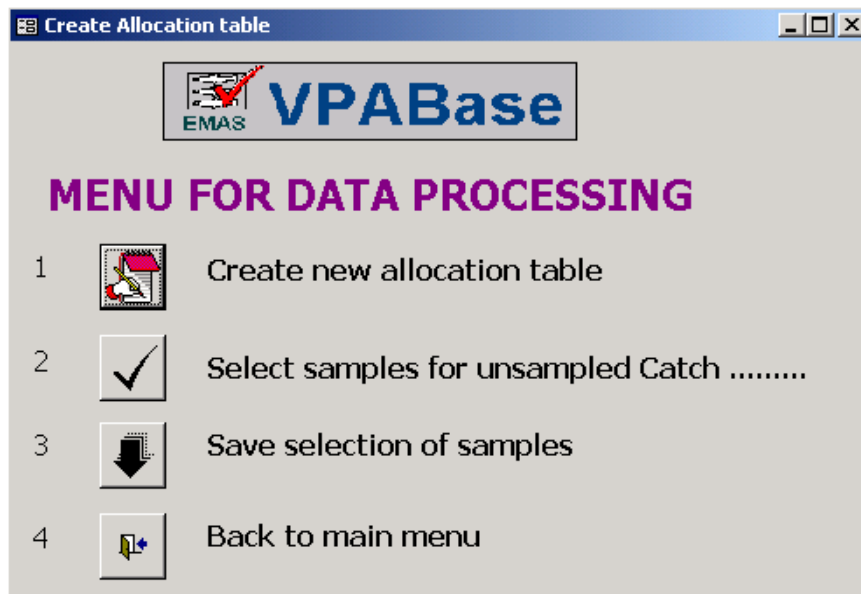


Figure 8. Data entry form for the selection of age compositions to be used to fill-in unsampled catches (hypothetical example).

Catch record:

Country: Denmark ICES Subdivision: IVa East
 Species: Herring ICES Division: IVa
 Year: 2000 ICES Area: IV
 Quarter: 2 Official Landings: 3379
 Fleet: Industrial fishery Weight of catch: 0
 Use of catch: Reduction
 Catch no. 107

NO age distribution for this catch

Exit Selection form

Equal weight of samples
 Weighted by number of otoliths
 Weighted by sampled catch
 Weighted by sample size
 Weighted by inverse CV (age)
 Weighted by inverse CV (weight)

| No. | Country: | Area: | Div. | Sub_Div. | Qrt. | Species | Fleet | Use of catch | Number of measurements | | | select |
|-----|--------------|-------|------|----------|------|---------|----------------|--------------|------------------------|--------|----------|-------------------------------------|
| | | | | | | | | | Samples | Lenght | Otoliths | |
| 126 | Denmark | IV | IVa | IVa East | 1 | Herring | Human cons | H.C. | 1 | 0 | 0 | <input type="checkbox"/> |
| 130 | Denmark | IV | IVa | IVa Wes | 1 | Herring | Human cons | H.C. | 0 | 0 | 0 | <input type="checkbox"/> |
| 133 | Denmark | IV | IVa | IVa Wes | 4 | Herring | Human cons | H.C. | 0 | 0 | 0 | <input type="checkbox"/> |
| 135 | Norway | IV | IVa | IVa East | 2 | Herring | A | Reduction | 25 | 0 | 0 | <input type="checkbox"/> |
| 136 | Norway | IV | IVa | IVa East | 3 | Herring | A | Reduction | 6 | 0 | 0 | <input type="checkbox"/> |
| 137 | Norway | IV | IVa | IVa East | 4 | Herring | A | Reduction | 0 | 0 | 0 | <input type="checkbox"/> |
| 138 | United Stat: | IV | IVa | IVa Wes | 1 | Herring | Industrial fi: | Reduction | 6 | 329 | 59 | <input checked="" type="checkbox"/> |
| 139 | United Stat: | IV | IVa | IVa Wes | 2 | Herring | Industrial fi: | Reduction | 3 | 6788 | 237 | <input checked="" type="checkbox"/> |
| 140 | United Stat: | IV | IVa | IVa Wes | 3 | Herring | A | Reduction | 0 | 0 | 0 | <input type="checkbox"/> |
| 141 | United Stat: | IV | IVa | IVa Wes | 4 | Herring | A | Reduction | 1 | 0 | 0 | <input type="checkbox"/> |

Record: 6 of 37

Figure 9. EXCEL worksheet used by the ICES Herring WG and the Mackerel, Horse mackerel, Sardine and Anchovy WG.

(The figure shows only a part of the worksheet)

| CATCH NUMBERS, LENGTH AND WEIGHT AT | | | | | | | | | | | | | | |
|--|--|-----------------------|------------------|----------------------------------|-----------------------|----------------------------------|------------------|-----------------------|------------------|------------------|-----------------------|------------------|------------------|-----------------------|
| Catch data (WG estimate) should include discards and unre WG for the Assessment of Herring in the Area south of 62° TONNES SHOULD ALWAYS BE FILLED IN (even if no numbe | | | | | | | | | | | | | | |
| Country: Denmark | | Division: IVaW | | Fleet: Industrial fishery | | Revised (date): 3-mar-'00 | | | | | | | | |
| Species: Herring | | Year: 2000 | | | | | | | | | *only t gear is | | | |
| | | Quarter 1 2000 | | | Quarter 2 2000 | | | Quarter 3 2000 | | | Quarter 4 2000 | | | Quar |
| Winter-Rings | | Numbers at age (*000) | Mean Length (cm) | Mean Weight (kg) | Numbers at age (*000) | Mean Length (cm) | Mean Weight (kg) | Numbers at age (*000) | Mean Length (cm) | Mean Weight (kg) | Numbers at age (*000) | Mean Length (cm) | Mean Weight (kg) | Numbers at age (*000) |
| 0 | | | | | | | | 37 | 11,5 | 0,014 | 651 | 11,5 | 0,014 | |
| 1 | | | | | | | | 1607 | 21,5 | 0,064 | 28 171 | 21,5 | 0,064 | |
| 2 | | | | | | | | 166 | 25,0 | 0,116 | 2 909 | 25,0 | 0,116 | |
| 3 | | 159 | 26,5 | 0,114 | | | | 238 | 26,2 | 0,133 | 4 172 | 26,2 | 0,133 | |
| 4 | | 635 | 28,0 | 0,140 | | | | 52 | 26,5 | 0,140 | 919 | 26,5 | 0,140 | |
| 5 | | | | | | | | 41 | 25,5 | 0,112 | 727 | 25,5 | 0,112 | |
| 6 | | | | | | | | | | | | | | |
| 7 | | | | | | | | | | | | | | |
| 8 | | 159 | 30,5 | 0,190 | | | | 41 | 31,5 | 0,206 | 727 | 31,5 | 0,206 | |
| Plus-group 9+ | | | | | | | | | | | | | | |
| Total/Mean --> | | 954 | 28,2 | 0,144 | | na | na | 2 182 | 22,5 | 0,080 | 38 276 | 22,5 | 0,080 | 0 |
| Catch | | 137 | (t) | | Catch | na | (t) | Catch | 175 | (t) | Catch | 3 067 | (t) | Catch |
| SoP | | 137 | (t) | | SoP | na | (t) | SoP | 175 | (t) | SoP | 3 064 | (t) | SoP |
| SoP | | 100% | (%) | | SoP | na | (%) | SoP | 100% | (%) | SoP | 100% | (%) | SoP |
| No. aged | | 2 | (n) | | No. aged | 0 | (n) | No. aged | 0 | (n) | No. aged | 13 | (n) | No. aged |

New sheets for other fishing areas may be copied by <EDIT> <MOVE OR COPY SHEET>.
 Then rename the sheet according the new fishing area by double clicking on it. (Rename also this CANUM sheet at

Figure 10. Example of report in EXCEL format.

| Winter-rings | Numbers (000) | Mean weight |
|--------------|---------------|-------------|
| 0 | 10175 | 4,56 |
| 1 | 2116 | 10,54 |
| 2 | 723 | 14,36 |
| 3 | 4 | 22,88 |
| 4 | 0 | 0,00 |
| 5 | 0 | 0,00 |
| 6 | 0 | 0,00 |
| 7 | 0 | 0,00 |
| 8 | 0 | 0,00 |
| 9 | 0 | 0,00 |
| Total | 13019 | 6,08 |