

REPORT OF THE
STUDY GROUP ON FUTURE REQUIREMENTS FOR
FISHERIES ASSESSMENT DATA AND SOFTWARE

ICES Headquarters
21 October 1997

This report is not to be quoted without prior consultation with the General Secretary. The document is a report of an expert group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

International Council for the Exploration of the Sea

Conseil International pour l'Exploration de la Mer

Palægade 2-4 DK-1261 Copenhagen K Denmark

...

...

...

...

...

TABLE OF CONTENTS

Section	Page
1 PARTICIPANTS.....	1
2 TERMS OF REFERENCE.....	1
3 INTRODUCTION.....	1
4 REVIEW OF EXISTING SOFTWARE.....	2
4.1 IFAP.....	2
4.2 Ad Hoc Software.....	2
5 OUTLINE OF A NEW SYSTEM.....	3
5.1 Introduction.....	3
6 IMPLEMENTATION PLAN.....	6
7 REFERENCES.....	6

1 PARTICIPANTS

Eero Aro	Finland
Roger Bailey	ICES
Robin Cook (Chairman)	UK
Poul Degnbol	Denmark
Harry Dooley	ICES
Steve Flatman	UK
Jean-Jacques Maguire	Canada
Garry Hopwood	ICES
Benoit Mesnil	France
Leif Pedersen	ICES
Henrik Sparholt	Denmark
Knut Sunnanå	Norway
Morten Vinther	Denmark

2 TERMS OF REFERENCE

At the 1997 Annual Science Conference it was decided that a study group on the future requirements for fisheries assessment data and software will be established under the chairmanship of Dr R. Cook (UK) and will meet at ICES Headquarters on the 21 October 1997 to:

- a) review the existing fisheries assessment data and software available in the ICES Secretariat;
- b) define future data software needs for assessment working groups, ACFM and the ICES Secretariat;
- c) make an implementation plan for meeting these future needs;
- d) advise on the requirements for making data accessible to outside users.

The Group will report to ACFM at its meeting from 22–31 October 1997.

3 INTRODUCTION

Since the development of IFAP many new analytical methods have been developed which are in common usage by assessment working groups. It has proved time consuming or impossible to incorporate these tools into IFAP and as a result most of these new methods are used outside the IFAP framework. This means there has been a lack of adequate quality control in the implementation and usage of the software as well as inefficiency in the exchange of data between programs. As a result there is considerable scope to improve the productivity of assessment working group meetings and reduce vulnerability to errors by trying to bring software development within an organised framework. Most of the meeting was devoted to identifying such a framework which would meet many of the immediate and medium term needs of assessment working groups, ACFM and the Secretariat. It was also recognised that there was a need also to consider a longer term more strategic view of how ICES assessment needs might develop. This was not discussed in detail at the meeting but a sub-group has been tasked to produce a policy paper in advance of the May 1998 ACFM meeting.

Most of the meeting was devoted to identifying a the conceptual framework within which new ICES assessment software would be developed and forms the bulk of this report. More specific details and specifications will be handled by sub-groups who will report back early in 1998 (see section 5).

4 REVIEW OF EXISTING SOFTWARE

4.1 IFAP

The ICES Fisheries Assessment Package is now about five years old. It was developed at a time when most assessments followed a routine procedure which generally consisted of:

- a) data preparation
- b) virtual population analysis (VPA) or some equivalent step
- c) age structured catch forecast
- d) yield per recruit analysis

Data preparation has remained predominantly in the hands of working groups. IFAP acts as a data management system once source data have been processed into a format ready for VPA. Working groups enter their assessment data into IFAP and use the package to perform various analyses. IFAP offers a number of data management facilities including editing, and file transfer. In general the data kept in the IFAP system should be the reference set used by working groups and ACFM.

At present IFAP offers two principal assessment tools. These are the Lowestoft VPA suite (Darby and Flatman, 1994) and Integrated Catch Analysis (ICA) (Patterson and Melvin, 1996). These tools are part of the IFAP package and runs using the IFAP system are fully documented and traceable. The results from these analyses are made available to the IFAP catch forecast and yield per recruit programs but the exchange of data between the VPA step and forecast requires user intervention which is error prone.

In addition to analytical programs, IFAP offers certain standard plotting facilities to produce routine outputs required for both working group and ACFM reports, such a long term stock trend summaries and yield per recruit curves. This is a particularly important function in improving the efficiency of report production by the secretariat but is seriously undermined by the increasing use of software external to IFAP. However, certain standard plots, such as those for Fmed on the stock-recruitment graph are not linked to the appropriate analytical tools which often results in errors in these plots. Some working groups have already written their own software to overcome this problem.

An important feature of IFAP is that it has been developed in SAS. At the time of IFAP's conception, SAS was commonly used on the computer systems of ICES member countries. This is less so at present with many institutes opting for alternative statistical or data handling products. SAS is considered expensive by some institutes and few are prepared to support SAS simply to make IFAP available on their systems. This acts as a barrier to the transfer of IFAP to ICES institutes and to date IFAP is only available at ICES headquarters. It should be noted, however, that the Institute of Marine Research in Bergen is attempting to install IFAP on its own computer system. It is also possible to access IFAP at the Secretariat over the Internet.

4.2 Ad Hoc Software

Almost as soon as it was developed, a number of working groups found that the facilities within IFAP could not handle all of their needs. For example, the catch forecast in many assessments makes use of recruitment indices. The program RCT3 rapidly became a standard tool for predicting year class strength from survey data but it has remained outside the IFAP system meaning that the forecast step in assessments had to be manhandled between the VPA and forecast programs. This creates working delays and increases vulnerability to human error.

The forecasting stage of assessment has proved difficult for IFAP, primarily because the basic VPA tools were conceived on a single fleet basis. Data within IFAP are not stored in fleet disaggregated form which means partitioning out fleet or discard fishing mortalities has to be done externally, once

again introducing ample opportunity for human error. This frequently occurs and has been a particular problem for fleet disaggregated North Sea herring forecasts where ad hoc spreadsheets have been used. The catch forecast remains the most error prone step in assessments due to the need for manual preparation of much of the data in the heat of the meeting when mistakes can often be made. This is dangerous because much of the credibility of ICES advice is judged on the validity of catch prognoses which underpin TAC management systems.

It is inevitable that as science progresses, new methodology will emerge. Recent demands for medium term projections have lead to a number of different programs and spreadsheets being used with varying degrees of competence and utility. This software is demand led and has always outstripped the speed with which new techniques can be incorporated into IFAP and is therefore, likely to continue. Other new methods, such as Bayesian analysis and Time Series Analysis are already entering the working group arena and it is unlikely these can be incorporated into IFAP within a desirable time frame.

The focus by ACFM over the years on standard age based VPA assessments has limited the implementation of many useful and important methods. For example, there are no length based methods routinely available at ICES although the *Nephrops* Working Group makes extensive use of these using their own software. There are also simpler assessment methods, such as dynamic stock production models, which are occasionally used but which if more readily available would make an important improvement to assessments which do not easily fit into the VPA straight-jacket. ICES needs to have these tools available to its working groups.

Finally, it should not be forgotten that ICES scientists invested a substantial amount of development time into multispecies VPA. These programs are not easily accessible, particularly by untrained users, but are likely to be important tools for working groups dealing with species interactions. IFAP is not a suitable package in which to incorporate MSVPA but a framework does need to be developed in which such programs can be accessed.

5 OUTLINE OF A NEW SYSTEM

5.1 Introduction

It is proposed that in a new system, analytical programs would not be integrated into the assessment database system. This will allow more flexibility to use an appropriate technique for a particular stock. Most analytical assessment software is developed by scientists in national laboratories and this is where most new methods will emerge. It is important that ICES provides a framework which can take advantage of these developments in an orderly way which will provide a degree of quality control and assist efficiency by ensuring that data exchange between disparate programs can be standardised. It is proposed that ICES should aim to maintain a library of "accepted" routines. These routines would conform to minimum programming standards, have to pass certain quality checks and have adequate documentation. Such a library would contain the reference versions of programs in use and as such would minimise the use of multiple versions of the same program.

Basic data handling

At present the data contained in IFAP are aggregate data which are related to total catch. These are essentially the input data to catch at age analysis. This is likely to continue to be the case in the immediate future. IFAP provides an appropriate secure data management system for these type of data and it is envisaged that this element should remain the main data management tool for working groups. In the future it is to be expected that more disaggregated data will be required, containing information on fleets on a quarterly basis for example. When the use of assessment techniques that use disaggregated data becomes more widespread and the supply/restrictive use issues of these data are resolved a more comprehensive database will need to be developed. To satisfy ownership and control concerns the underlying database engine may have to be a distributed database with the physical location and access control being located within member institutes. However the initial desire is only to allow freer access to existing data and results from IFAP and the other assessment packages in use. Essentially the ICES web server will become the new information management system for assessment results.

There are no immediate plans to create a data management system within the Secretariat or the member institutes. In the meantime, however, there is a need to prepare a guide for working groups on the handling of the disaggregated data which are worked up for entry into IFAP. Some working groups already retain their data on a fleet/quarter basis with clearly identified rules for aggregating the data. Other groups do not retain this information which makes post hoc fleet disaggregations impossible and a considerable amount of fleet information has been lost over the years. Working groups should be encouraged to handle their basic data in a more systematic way.

Data Interface

The interface between the database and the analysis programmes will be via well documented file formats. The current format is based on the Lowestoft data format. However it was felt that, in addition, a simpler transfer format with less data structure that can be more easily transferred to a relational database will be required. The input data, results and report summary will be made available on the ICES web server initially in the Lowestoft formats. The files will also have a specific naming convention for each class of assessment. It will be necessary to develop a user guide for file formats for those classes of data in common use.

Software standards

To avoid anarchy with a plethora of techniques and versions of ad-hoc programs a software library will be created of certified assessment/analysis packages. To obtain certification and hence enable use of a particular package by a WG the following steps must be followed:

- a) The method must be documented or a suitable reference cited to enable the independent development of the method in another programming language or on another computer platform.
- b) Executable versions of the programme must be supplied with enough documentation to allow the installation and operation of the package by others.
- c) It must be possible to read and write (via an auxiliary a re-formatting programme if necessary) the standard assessment data file formats and use the standard naming conventions if they are appropriate for the particular technique. This is not only to allow others to easily use the package with their data or to reproduce results but to enable the automation of the publication of the results both on the web and in the ACFM report. If the package uses input data or produces results that are unrelated in content to the defined standard assessment data file formats their structure and content must be properly documented and contain a minimum set of details such as package version, input data, run date/time and operator.
- d) Supply a set of benchmark datasets and outputs.
- e) Supply source code for non-commercial packages.

f) A commitment to supply the library with new versions of packages as they are developed.

These certification requirements also apply spreadsheet programs, macros and conversion utilities, if they are required to enable a third party to reproduce the results of the working group.

Secretariat Needs

The principal requirement of the secretariat is access to data for standard plots and tables for reports, particularly the ACFM report. The system proposed will make this task more difficult because the results of assessments will not automatically be retained within IFAP. However, this problem already exists to some degree. It will be necessary for working groups to leave the required files for the use of the secretariat in a locatable directory on the ICES computer system. It will be the responsibility of the WG Chair to ensure that this task is completed. Provided that file standardisation is adhered to, it will be possible for the Secretariat to develop the appropriate software for report production. Given that there will no longer be a requirement for the Secretariat to develop IFAP, there will be additional resources for this need.

Data Accessibility

An important need, not fully realised at present, is for both the input data and results of assessments to be readily available to legitimate users for review and additional analyses. Many working groups and ACFM require additional work to be done after working group meetings for research purposes, assessment checking and additional advisory requests from customers. While IFAP can be accessed remotely, a simpler system would be to make the data available over the Internet. It is proposed that this should be done using the world wide web.

The ICES web server will allow retrieval by the following methods:

- (a) individual stock
- (b) working group stock
- (c) all stocks
- (d) shopping bag style - grab selected stocks

Initially all the data will have access restricted by password protection but it is expected such protection will become limited to particular stocks or data files after publication of the ACFM report. Display of the standard output files will also be available in the form of graphical plots. The ultimate aim is to make the assessment part of the ACFM report available digitally on the web independent of the type of assessment. Until the full specification of the standard file formats and the certified software library is created only those assessments that are conducted in IFAP will be available. It may also be appropriate to WG and ACFM reports available on CD-rom.

Training

Many working group members are unfamiliar with the assessment tools available within IFAP and the ad hoc software in common use. With the proposed new more flexible system, many more methods will become available at ICES. There is, therefore a need to run periodic training workshops to assist the dissemination of up to date techniques. This would help working groups to operate or determine the appropriateness of a specific assessment packages for a particular stock.

6 IMPLEMENTATION PLAN

The Study Group identified four main tasks which are required in order to implement the proposed new system. These are:

- a) Code of Practice for data handling by assessment working groups;
- b) A Programming standards guide for potential developers of assessment software;
- c) A definition of formats for file exchange between assessment programs;
- d) A strategic overview of how ICES assessment needs should be developed in the longer term.

It was agreed that these four tasks would be undertaken by correspondence and with a view to producing a draft of these documents for consideration by ACFM at its May 1998 meeting. Responsibility for producing the drafts were assigned as follows:

- a) Henrik Sparholt, Stuart Reeves, Frans van Beek, Morten Winther;
- b)& c) Robin Cook, Steve Flatman, Benoit Mesnil, Dankert Skagen;
- d) Jean-Jacques Maguire, Poul Degnbol.

Not all those identified were at the meeting and will have to be contacted. In the event that these are unable to participate, alternatives will be found. It is proposed that each sub-group will produce a first draft of the required product by the end of February 1998. These drafts will then be refined by the whole Study Group by correspondence in advance of the May 1998 ACFM meeting. It is recommended that the Study Group re-convene at ICES Headquarters for one day prior to the ACFM meeting to finalise the drafts. This could be done at national expense. Most of the members of the group will be present in Copenhagen at that time and will not incur significant extra costs.

7 REFERENCES

Darby, C. D., and Flatman, S. (1994). Virtual Population Analysis: version 3.1 (Windows/Dos) user Guide. MAFF Information Technology Series no.1, Directorate of Fisheries, Lowestoft.

Patterson, K. A., and Melvin, G. D. 1996. Integrated Catch at Age Analysis Version 1.2. Scottish Fisheries Research Report. No. 58.