Advisory Committee on Fisheries Management ICES CM 1999/ACFM:9

REPORT OF THE

STUDY GROUP ON FUTURE REQUIREMENTS FOR FISHERIES ASSESSMENT DATA AND SOFTWARE

ICES Headquarters
19 October 1998

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Participants

A. Aglen	Norway
R Cook (Chairman)	UK
H. Dooley	ICES Secretariat
S. Flatman	UK
G. Hopwood	ICES Secretariat
J-J. Maguire	Canada
B. Mesnil	France
W. Panhorst	ICES Secretariat
K. Patterson	UK
L. Pedersen	ICES Secretariat
S. Reeves	UK
H. Sparholt	ICES Secretariat

1.2 Terms of Reference

The **Study Group on Future Requirements for Fisheries Assessment Data and Software** [SGFADS] (Chair: Dr R.M. Cook, UK) will meet at ICES Headquarters on 19 October 1998 at national expense to:

a) consider and finalise inter-sessional work on the specification of assessment software programming guidelines, data file structures and working group data handling, protocols;

b) review progress on Internet access to working group data, results and reports. Identify future needs and develop a plan for their implementation;

c) consider the present status of IFAP and devise a timetable for phasing out Secretariat technical support to IFAP,

d) identify software needed within the Secretariat required by assessment working groups to expedite the production of reports. This software should concentrate on the production of standard graphics and tables needed in reports;

e) review progress with the Castle Room LAN and make recommendations for its technical support and appropriate working procedures for its use by working groups and advisory committees.

SGFADS will report to ACFM at its October 1998 meeting.

2 CASTLE ROOM LAN AND OTHER COMPUTER HARDWARE ISSUES

2.1 Technical support

The Castle Room network was installed in May 1998 to ACFM's specification. The network was designed to meet ACFM's requirements to create a paperless environment for their meetings, as well as to overcome restrictions then existing in the Secretariat's network with regard to access to the Internet. Since its installation, the network was used extensively during the May ACFM meeting, but subsequently has been used intermittently. The principal use has been to provide visitors to the Castle Room with access to the Internet via their laptops but has been used by some groups, such as SGSEL, as a working environment. Following their meeting, ACFM reported that they were satisfied with the performance of the network, but were concerned about the lack of support from the Secretariat, and the lack of a seamless access to the Secretariat's network.

Subsequent to the installation of the Castle Room network, new technology on the Secretariat's network server has been installed. The consequence of this is that the original concerns the Secretariat had with regard to virus-free access to its network have now been eliminated. In connection with this development, and prompted by the clear advantages offered by the Castle Room network, it is the intention to provide a hub in each meeting room, which will be connected to the Secretariat's network. This will allow all working group visitors seamless access to the Secretariat's network, and to the Internet, via their own laptop or workstation.

The advantages of this development over the existing Castle Room network are clear. In particular the question of the need to provide Secretariat support becomes irrelevant as this solution can be supported by the Secretariat's IT staff without any additional training or resources being required. It also solves the problem of seamless integration with the Secretariat's network at a stroke. Delays in the current implementation of the NT network mean that it is not possible to provide this capability for the forthcoming ACFM meeting, but it is expected to be in place in good time for the May 1999 ACFM meeting. It is also expected that those Assessment groups meeting in advance of the May meeting will be able to make use of these facilities. By that time all users will be working in a Windows95/NT/Unix environment.

2.2 Software needs

One of the most important uses of the LAN is to enable online editing of reports during plenary sessions. Ideally, members of the meeting will have laptops at their desk and be able to view the document as it is being reviewed and edited. At present this is not possible as the person editing the text has to keep saving the document for others to view. An interim solution to this problem is to project the file on a screen in the Castle room, but this is difficult to view from a distance, especially for tired old eyes, an affliction affecting many ICES scientists. A better solution would be to purchase software to enable simultaneous viewing of documents at the desk.

Action: The Secretariat will investigate possible software for file viewing and report to the chair of ACFM

2.3 Facility implications for ICES

The success of the Castle room LAN and decision to expand the availability of the resource to all meeting rooms has implications for ICES policy in the provision of computing hardware to working groups and meeting room facilities in general. The provision of a network connection means that working group members will increasingly make use of their own computers during meetings and since this will give direct access to the ICES network, there will be less need for the Secretariat to provide large numbers of machines. The Secretariat may therefore be able to save money by providing fewer stand alone computers and printers and the cost of upgrading may be similarly reduced. It might be desirable for the Secretariat to provide a few laptops for those working group members without their own machines so that desk access to the LAN is possible.

Action; The Secretariat should review its policy on the provision of computing hardware to working groups to find the most cost effective solution in the light of the LAN. This need to be done in the context of a broader policy which considers meeting room facilities in general.

3 STATUS AND FUTURE OF IFAP

3.1 Developments in 1997/1998

The main activities undertaken on IFAP in the past year have been:

- Making the final assessment run done in IFAP available on the Web
- Implementation of IFAP at IMR Bergen (with full cost recovery)
- Interfacing a new version of ICA with IFAP
- Simplification of the menu system
- Improvements in data security and back up
- Routine maintenance

Some of the activities undertaken in the last year reflect the fact that the decision to retain IFAP primarily as a data management package taken at last years meeting has not been fully implemented. This decision now needs to be implemented fully and is elaborated below.

3.2 Data handling by IFAP and the Secretariat

In 1997 the Study Group decided that IFAP should be used only as a secure data management system and that assessment tools would be used outside IFAP. The Group re-iterated this decision which means that the Secretariat may seek simpler and more cost-effective software solutions to this much reduced requirement for IFAP. It is expected that IFAP will not be used directly by working group members but would be used by the Secretariat. It is anticipated that the system will operate as follows:

1. Working groups will provide the Secretariat with data updates for IFAP before their meeting.

2. The Secretariat will be responsible for entering the data into IFAP either manually or electronically from Lowestoft format data input files.

3. IFAP will perform quality control checks on the data and retain appropriate data backups

4. The Secretariat will export the working group source data to a standard working directory for use by working groups

5. At the end of the working group meeting any modification made by working groups to the working data files will be reviewed and IFAP data updated ONLY IF APPROPRIATE. The ICES stock assessment scientist will assist in this process.

For the present the analytical tools within IFAP will remain available but this will end when IFAP is moved to a new environment. This means that it will be possible to save money by limiting the SAS license to a single user since the only direct users of IFAP will be the Secretariat. Analytical tools equivalent to those within IFAP will be made available to assessment working groups (see section 5.1).

The subject of data security was discussed by the group. Recent changes have been made to improve security (Annex 1). Although a backup is always made, at present it is still possible to read in a completely new data set to replace an old set without any quality checks or warnings of possible inappropriate modifications. What is needed is a report by IFAP on a comparison of the old data with the new data. The group also felt that the present system of identifying stock files by their directory only was potentially error prone since the data files by stock all have the same name. This is a hangover from an older system and needs to be changed. The move to a new environment will assist in this as it permits long file names which can be used to clearly identify the stock.

Action: The Secretariat will incorporate additional quality checks and data comparisons in IFAP and keep data security under review in discussion with ACFM. This must be done in close consultation with the ICES Oceanography Secretary, the chair of SGFADS and the chair of ACFM.

4 REVIEW OF PROGRESS IN TASKS IDENTIFIED IN 1997 SGFADS REPORT

4.1 File structures

Two working documents were prepared intersessionally on ASCII file structures which may be used to exchange data between assessment programs (WD1 and WD2). These papers were briefly reviewed. The technical details of these file structures still need to be resolved but this is a relatively simple problem. It is agreed that the Lowestoft input file format is the most useful example format to follow for typical VPA type analytical tools. The Lowestoft input file format is suitable for both for input and output data. There are also a number of other commonly used analytical programs which have established a de facto standard for other tools such as stock summary plots and catch forecasts. These file structures need to be fully described and made available to potential program developers. It is proposed to hold an ad hoc meeting of appropriate experts to finalise this issue.

4.2 Programming guidelines

A working document (WD3) was reviewed which elaborates some of the proposals on software development guidelines described in the 1997 SGFADS report. These guidelines are not formal programming standards as would apply to professional programmers but simply a set of common sense attributes which programs should have, such as protection against file overwriting, the maintenance of a log file to record source data files and program options etc. There is still a need to finalise the working document and it is proposed that this should be done by the same group identified in section 4.1.

4.3 Guide to working groups

ACFM recognises the need for assessment working to adopt improved quality control and assurance procedures. SGFADS has addressed part of this issue by considering assessment source data handling. Two working papers (WD4 and WD5) were reviewed which describe how assessment working groups should maintain and document their source data. Essentially what is required is a systematic way of storing data and recording how the basic data are worked up to provide inputs to assessments. It was agreed that these two documents would be circulated to assessment working

groups for consideration with a request for the working group to provide an outline of how their data is currently handled.

Action : The chair of ACFM will write to assessment working group chairs circulating WD4 and WD5 and request a description of current data handling procedures.

4.4 Internet access to data and reports

It is ICES intention to make working group assessments and reports and the ACFM report available on the Web. Progress has already been made and is described in WD6. It is possible to access working group assessments that have been documented in IFAP but not others. IFAP assessments account for the large majority of assessments at present, however. The ACFM report is also available, but only as a large number of files of varying formats (Word, Excel etc). Furthermore the file naming is such that identification of relevant sections of the report is not obvious. These problems have been recognised for some time and are being addressed. It will be easier to address the ACFM report problems because the report format and inputs are more standard. Two factors will aid the production of a web version of the report. These are:

a) The move to NT will enable meaningful file names to be given to report sections so that users will be able to more easily identify required sections, and

b) The development of graphics software to produce the standard figures and tables from standard results files in a suitable electronic form (eg GIF or JPEG).

It is expected that NT will be the operating environment early in 1999 which will solve (a). With regard to (b), the Secretariat were provided with an example program used by working groups to plot stock trends from a standard input file. The Secretariat will use this as a basis for developing appropriate software which can be used both by the Secretariat and working groups to produce electronic copies of figures ready for web publishing.

5 FUTURE WORK

5.1 Assessment tools for working groups

The decision to restrict IFAP to data handling means that analytical tools currently embedded within or interfaced with IFAP will have to be provided in a windows directory. It was decided to set up a directory of standard programs in a read only directory which will contain equivalent analytical tools to IFAP. Initially these programs will be ICA, and the Lowestoft VPA suite. A yield per recruit and catch forecast program will also have to be provided. The two most likely candidates for this are WGFRAN4 and REFPOINT. In addition, INSENS which acts as an interface between the VPA suite and other programs will be provided. The choice of programs, which will require the agreement of ACFM, will be dealt with by the ad hoc group referred to earlier.

At present not all the programs noted above conform to an ideal standard in terms of user attributes, testing and documentation. However, it is necessary to provide the essential tools and it is important to establish a standard set of software so that multiple versions of the same program are not in use at meeting, a common occurrence in some working groups at present. Moreover, while these programs are deficient in certain respects (run documentation for example), they have the advantage of requiring less user intervention in preparing input data which makes certain errors less likely. These programs also have utilities for making standard plots which can save a considerable amount of time. It is important that these programs are tested in an NT environment before being loaded onto the ICES system.

It is recognised that in the long run it will be necessary to upgrade the above software to conform to the guidelines described in WD3. It is expected also that new software will conform to these standards and ultimately lead to the establishment of an agreed library of ICES approved assessment programs. The details of the acceptance protocol by which ICES includes programs in the library will be dealt with by the ad hoc group.

5.2 Non-assessment software needs for working groups

The increasing use of personal computers by working groups has meant that a large amount of the initial report production is done without Secretariat assistance. In particular standard software packages such as Excel are used to produce figures. The conversion of data from raw output from analytical program into report quality figures in this way is proving very time consuming and is inclined to be error prone. Clearly it is the appropriate task for assessment scientists to produce analytical tools. It is less appropriate for these individuals to produce presentation software,

especially for standard graphics needed in reports. The study group agreed that the Secretariat should develop presentation software for tables and figures. Such software would obtain input from analytical programs in an agreed ASCII file format and would produce graphics and tables suited to Secretariat needs in production of electronic reports.

Action: The Secretariat should liaise with ACFM and develop a suite of programs to produce standard tables and figures for ICES reports.

5.3 The need for a final workshop

During the course of the study group meeting it became clear that a workshop is required to finalise some significant technical details related to the assessment software which will run outside ICES. These matters include final file definitions to exchange data between programs, final programming guidelines, the establishment of a list of "standard" assessment programs to be made available at the Secretariat to perform the analytical work presently done within IFAP. There is an important constraint in the timing of such a meeting. IFAP must be moved to a new environment within the next few months because the present UNIX system will be phased out. In the process of this transfer, there is little point in undertaking substantial modifications to IFAP if it will no longer be the principal analytical tool and there are potential cost savings to be made for the reasons given in section 3.2. This means that new analytical tools must be made available during the course of 1999. If the proposed study group was established at the 1999 ASC it would mean that the most likely time a meeting could be held would be early in 2000. It is proposed, therefore, that a workshop be held in 1999 to address the outstanding issues as soon as possible. Proposed terms of reference are given in Annex 2.

6 POTENTIAL PROBLEMS WITH THE NEW ASSESSMENT TOOLS

An important feature of IFAP is that it documents the final run. This provides a means of tracking the data used in an assessment and hence the appropriate files needed for standard tables and figures. This applies to the VPA, yield per recruit and catch forecast. Other analysis programs already operate outside IFAP and do not enjoy this documentation facility. With the proposed new arrangements there will be no automatic documentation procedures for final runs. This means there is an increased danger of confusion in the identification of final runs and also difficulties in finding the relevant data files needed by the Secretariat to finalise reports and inputs for the ACFM report. The problem can be partially overcome by improving certain aspects of the analytical software such as the creation of a log file. In addition, working groups will need to adopt better working procedures and chairs will need to ensure more pro-actively that the files used in the final assessment are archived in a form suitable for the Secretariat. It is also suggested that the ICES stock assessment scientist takes and active part during the assessment meeting to ensure this is done. This will help working groups and also save the Secretariat time in the long run. The Secretariat is also encouraged to record problems they encounter in clearing up assessment files and working group reports at the end of meeting. These problems should be reported to ACFM so that solutions can be found and efficiency improved.

Action: The chair of ACFM will alert working groups to the problem of quality control with the revised software arrangements and request that improved working procedures are adopted.

Action: The Secretariat will record housekeeping problems arising from working group meetings when finalising their reports and preparing the ACFM report. These problems will be drawn to the attention of ACFM.

7 WORKING DOCUMENTS

WD1. Contribution to the discussion on standard output files format. B. Mesnil
WD2. Standard formats for exchange files. S Flatman.
WD3. Programming Guidelines. R. Cook
WD4. Code of practice for data handling by assessment working groups. S Reeves, F. vn Beek, H. Sparholt, M Vinther.
WD5. A programme for calculating total international catch-at-age and weight at age. K. Patterson
WD6. Publication of data, results and reports on the World Wide Web. G. Hopwood.
WD7. Fisheries assessment software and data storage. L. Pedersen.

ANNEX 1

Data security in IFAP

In the present system data security is achieved by **logging** and **automatic backup**.

Logging: all data updating, data import and execution of analysis programs are recorded by *user code*, *date*, *time* and an *activity code*.

Automatic backup: no system data set is changed without first being copied to a backup data set (in this context a data set corresponds to a Lowestoft flat file, e.g. CANUM, CATON, etc.). In practice, this is handled by having three kinds of data (in IFAP called *data status*): *system data*, *user data* and *backup data*. There is full read access to all data sets, but only *user data sets* can be updated. *System data sets* can be overwritten, but before this happens an automatic *backup data set* is made. If a user wants to update (manually) a *system data set*, the following procedure is to be used:

1. 1.Copy the existing *system data set* to a (new) *user data set*.

2. 2.Update the *user data set*.

3. 3.Copy the updated *user data set* to the *system data set* (this will overwrite the old *system data set*, but first an automatic *backup data set* is made).

4. Delete the *user data set*.

Data may also be copied from ASCII files in the Lowestoft directly to a *system data set*, but then the existing *system data set* is first backed up automatically (corresponding to step 3 above). In addition to the automatic backups, all data updating, data set copying and data imports are logged, as described above. The system has been designed to allow all IFAP users to update the data and, at the same time, protected against data loss.

It is no problem to expand the security with various measures against data inconsistencies and/or inadvertent data changes when a system data set is imported in total. This may be implemented as **automatic warnings** when one of the following conditions arises:

1. Data for back years are changed.
2. The age span is changed
3. Unlikely data values.
4. Unlikely data changes.
5. Automatic SOP checks.

To be of any value, the checking in items iii. and iv. should, probably, be dependent on the species, the data type (CANUM, WECA, etc.) and the age group. It might be implemented by checking the new data against lower and upper limits set in look up tables based on these criteria.

ANNEX 2

Terms of reference for a proposed workshop

A **workshop on standard assessment tools for working groups** (Chair: Mr S Reeves, UK) will meet at Aberdeen, UK on dates to be decided in 1999, at national expense to:

a) Prepare a preliminary list of analytical software to be used by assessment working groups which will replace the analysis tools currently performed by IFAP.

b) Identify any additional software, currently in use, which might be usefully included in the standard set.

c) Document the files to be used by these programs to exchange data.

d) Agree a set of programming guidelines for assessment software developers and acceptance protocol for such programs to be included in an ICES assessment software library.

The ICES Secretariat should be represented at the meeting.