

ICES SGMID Report 2006

ICES Advisory Committee on Ecosystems

ACE:02

Report of the Study Group on Management of Integrated (SGMID)

8-10 May 2006

ICES Headquarters



International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

**International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer**

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2006. Report of the Study Group on Management of Integrated (SGMID), 8-10 May 2006, ICES Headquarters. ACE:02. 24 pp.

For permission to reproduce material from this publication, please apply to 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.

© 2006 International Council for the Exploration of the Sea.

Contents

Executive Summary	1
1 Introduction	1
1.1 Terms of Reference (ToR).....	1
1.2 Participants	2
1.3 Structure of the report.....	2
1.4 Update on progress of SGMID issues.....	2
1.4.1 The new ICES data policy (ToR a).....	2
1.4.2 Database developments (ToR b).....	3
1.4.3 Progress on integration of dispersed data (ToR b).....	5
1.4.4 New international projects on data integration (ToR b).....	6
2 Implementation of the new ICES data policy 2006 (ToR a).....	7
2.1 Communication of the new policy.....	7
2.2 Review of ICES expert group's data policies (ToR e).....	8
2.3 Technical amendments to ICES databases	9
3 The ICES Data Centre User Survey (ToR c).....	10
4 Quality assurance, development of quality flags (ToR d)	11
5 Preparation of the ASC 2006 theme session on data integration (ToR f)	12
6 The future of SGMID (ToR g)	12
7 Recommendations.....	15
8 References	15
Annex 1: List of SGMID participants.....	17
Annex 2: Example "Conditions of Use" for data for UK National Tide Gauge Data from BODC	19
Annex 3: Detailed results of the ICES Data Centre User Survey	20

Executive Summary

The ICES Study Group on Management of Integrated Data met for the third and last time. The group convened at the ICES secretariat at the same time as the annual meeting of the Working Group on Marine Data Management (WGMDM), and half a day was devoted to a joint session on detailed planning of future needs for data groups in the ICES environment, and the organisation of the co-sponsored ASC theme session M on data integration.

SGMID's main task for this year was to review the implementation of the new ICES data policy. The new policy, drafted at SGMID's last meeting, was developed further by a number of ICES expert groups and unanimously adopted by the ICES Council in October 2005. The data policy was published in March, and SGMID discussed extensively how to communicate the new approach to data sources and data users. The group also suggested to the ICES Data Centre which technical amendments would be required for a timely implementation. An extension of the ICES in-house databases to hold information on data sources and data quality flags (including "known issues" with historic data) was thought to be of special importance in this context. Another task was the review of the outcome of the ICES data user questionnaire from the ASC 2005.

SGMID concluded that, while the group had reached its end-of-life as a study group, there would still be a need for such a user-oriented group dealing with strategic planning of ICES data issues. After discussion with WGMDM, the group made a number of suggestions on possible approaches. A final resolution requires a meeting between the chairs of the SGMID and WGMDM, the Consultative Committee Chair, and the ICES General Secretary.

1 Introduction

1.1 Terms of Reference (ToR)

The Study Group on Management of Integrated Data (SGMID) (Co-chairs: Peter Wiebe, USA, and Christopher Zimmermann, Germany), met in conjunction with WGMDM at ICES HQ from 8-10 May 2006 to:

- a) review the implementation of the new ICES data policy;
- b) review the status of data integration of external datasets currently not available to ICES, but needed for producing the advice relevant to an ecosystem based management;
- c) evaluate the outcome of the ICES data user questionnaire;
- d) investigate the need for quality assurance processes and indicators, and any existing standards. If no standards exist, propose a plan for quality process and/or indicator development;
- e) review data policies of individual ICES expert groups and flag differences between those and the new ICES policy, while taking account of new developments;
- f) together with WGMDM, complete the planning of the ASC 2006 theme session on data management and integration (including the formation of a subgroup for the selection of presentations to meet later in the year);
- g) prepare for merging WGMDM and SGMID activities into a new group of data managers, users and scientists called the Working Group on ICES Data and Information Management (WGDIM).

1.2 Participants

Peter H. Wiebe (co-chair) USA
 Christopher Zimmermann (co-chair) Germany
 Larry Atkinson USA
 Richard Ayers UK/England
 Steve Flatman UK/England
 Ole Folmer Denmark
 Martina Hennessy Ireland
 Lesley J. Rickards UK/England
 Yolanda Sarminaga Spain/Basque Country
 Reiner Schlitzer Germany

part-time participation from the ICES Secretariat:

Julie Gillin (Data Centre Manager)
 Hans M. Jensen
 Henrik Kjems-Nielsen
 Lena Larsen

1.3 Structure of the report

This report is split into seven chapters, addressing all ToR but combining some of them into chapters where necessary. The introduction deals with standard items and reviews progress on various issues since the last SGMID meeting. The second chapter deals with all aspects of the implementation of the new ICES data policy, emphasising the need for improved communication. Chapter 3 gives the evaluation of the ICES Data Centre user survey, chapter 4 reports on a discussion on quality assurance and the implementation of quality flags, and chapter 5 lists briefly what has been done in preparation of SGMID/WGMDM's co-sponsored theme session at the next annual science conference. Chapter 6 reiterates how the expertise of the group could be preserved after the Study Group's lifetime has formally ended, and discusses if there would be still a need for such a user-oriented group in the ICES community. The report concludes with recommendations and the annex.

1.4 Update on progress of SGMID issues

1.4.1 The new ICES data policy (ToR a)

The group was informed on the implementation of the new ICES data policy by the Bureau. At the 12 May 2005 meeting of the Bureau (BWGDDP), which took place after the SGMID meeting in Portugal, a reduced policy statement was approved that had the essential elements of the one produced and recommended by SGMID, but was reduced to fit on one page plus explanatory notes. The ICES council adopted the policy in October, but it was not made public immediately, in part because there was uncertainty about how to transition from the old policy to the new one. The policy was published in February 2006. The problem largely

concerned how ICES should deal with the opening-up of old data held at ICES under rules of restricted access. There was consensus in SGMID that there was a need to get legacy data sets on line with open access as soon as possible and it was important to give contributors a positive spin on doing this.

There was also discussion about the ability of ICES to persuade traditional data contributors to continue to submit data given the new policy. Some kinds of data would not likely be a problem, but fisheries data were singled out as a concern and ICES was dealing with the issue on a state by state basis. It was pointed out that it also was important to approach groups working up the aggregated data and making the reports and assessments to make their data accessible. Some essential data used by the committees are not kept at ICES, but need to be linked to the ICES data centre (e.g. acoustic and mackerel egg survey data). The issue of how quality assurance of linked data could be attained was raised. It was recognized that there needs to be a “frozen” copy of the data used in formulating assessments. It was also recognized that the assessment working groups and survey planning groups need to be informed about the new policy. More discussion about this topic is in section 2.1 below.

1.4.2 Database developments (ToR b)

The ICES DOME database

H. Mose Jensen lead a discussion about data flow to ICES through DAD, DATSU, and DOME. Data can come in any format, but must be translated into a native format and then go into the DOME database using the ERT3.2 format. The issue of documenting the submitter of the data and the data owner was raised and discussed. At present, this is not a mandatory field in the DAD (Database on Accessions and Documentation).

The screening of the data through DATSU was described. Project development funds have been obtained from the EU to improve DATSU (specifically an outlier analysis). Currently only one version of the program is available on the ICES web page. It is used for screening data within the trawl survey databases and InterCatch.

Once data are accepted, they go into the DOME database using software that is being developed partly with funding from OSPAR. There is an OSPAR-Dome steering group made up of representatives from CEFAS, OSPAR, and ICES. There are several DOME modules with system operators and data users. The complex diagram of the DOME database was presented, indicating the links between station data and water column versus sediment versus biota and then the codes used internally. A firewall separates the data users from the rest of the system. A data user's web page is planned in Phase 3 of the DOME development. All the old system's databases are still running, so external users could get access to the data. The initial phase of this development project involves getting the following data into DOME:

- Contaminants and biological effects of contaminants including diseases in biota
- Contaminants and biological effects of contaminants in sediment
- Contaminants and biological effects of contaminants in seawater
- Water bottle data types

Also included in Phase 1 is the development of a standard (simple) output. Phase 2 (2006-2007) involves getting more data in (biological community data, CTD, and underway data) and implementing an enhanced output. Phase 3 (tentative 2007) concerns implementing the web interface, providing data editing facilities, and including fisheries data. One problem that ICES said they were dealing with was the development of the concept of a station as the focal point for a collection of marine data.

The group discussed that from a user perspective it would be good to put fisheries data up

first. It was pointed out that there is an inter-sessional group under OSPAR that is contributing to the development of the web interface. The need for collaboration on development of the web interface at a very early time was stressed, and a “raw” full system at an early time would be necessary to get user feedback. The need for such a system was appreciated by the Data Centre.

It was questioned how this system will accommodate externally linked data base systems. The discussion turned to SeaDataNet because this group is focussing on solutions to this problem. This project was funded and ICES is a participant in the project.

There appear to be two efforts at making a new software structure that accommodates the different data sets. It was noted that the other item needed is a validation check, i.e. finding duplicate data and other such things. The data centre informed that when migrating data to DOME, duplicate data sets will be eliminated, but there will be additional data checks.

SGMID suggested that when ICES is taking on board external support (as with DOME; partly developed for OSPAR), it should make sure it gets a lot out of it. ICES intends to get all of the data into the new database and then it can release resources to other activities. It was suggested that individual data inventories should be examined to see what needs to go forward, but the approach the data centre decided to take is to get the data into DOME as fast as possible to relieve pressure. The data migration has not started, but it will start in 2006. The Bureau tasked the Consultative Committee to review all databases currently held at ICES and to recommend which databases are not needed any more, if this is required. The evaluation is not yet finalised. The plan was to move forward without waiting for Consultative Committee to make recommendations. SGMID’s evaluation of the data survey has importance here, as science should drive the data needs for specific databases and not the other way around. Such an evaluation of data inventories is also made in other data centres, like in that of the Irish Marine institute.

New developments in the ICES DATRAS database

The next version of DATRAS was described by L. Larsen. ICES is storing all the trawl survey data from five surveys (Baltic BITS, North Sea IBTS, and Beam Trawl Survey, west of Scotland bottom trawl, French survey). In addition, the Spanish and Portuguese will be submitting data. The trawl surveys are internationally coordinated multiple ship efforts in the North Sea (from 1965 to present), and in the Baltic (from 1991), but for other areas, they are single ship surveys. The data from these surveys are not yet easily accessible by the public, not even in the aggregated form. DATRAS was first developed in 2001 and included the DATSU screening utility development. A lot of emphasis has been on delivering aggregated data for assessment working group use. There is a web page available with password protection. But it is not working that well. The EU has funded 50% of the improvements in DATRAS including the calculating of variance and upgrading and improving access. By 1 November 2006, the variance computation needs to be in place for the five surveys. An Expert group with representatives from survey working groups and statisticians will be coming during the time SGMID met to provide advice on how to do this. There is a need to improve existing functionalities of data calculations and a need to be able to deal with sub-areas. There is a further need to account for data that are missing (this is not currently fixable), to improve provisions for downloading of data, to make it possible to web upload other data now into the database, and to show how products are calculated. Additional items include adding computation of weight by age and length for CPUE, GIS mapping of data using FishMap, and improving data checking (working groups want graphs to view outliers). Flag functionality is needed to show when an exchange format has changed. The data contributor details and acknowledgement will be done through the DAD (which will be used for all databases at ICES). DATRAS and DOME might be joined eventually.

Two comments were made. The integration of five additional surveys did not spread effort to ICES network. It would be ideal to have the ICES community to create their own databases and link to ICES, instead bringing the data into ICES. The EU is co-funding of DATRAS to get variance, but sooner or later EU may eliminate some of their surveys. ICES should be prepared by building the new DATRAS system to accommodate anything new in a seamless way. The data centre manager informed that this is planned. For example, the DAD system is being used for the metadata for all of the databases.

The issue of Fishmap at CEFAS iSEA was raised. This is accessed through the ICES web page. The method of presenting DATRAS data on a map is user friendly and a good example of how survey data can be presented. The raw data however can not be downloaded yet. There are restrictions for how the data can be accessed.

The new ICES InterCatch database

H. Kjems-Nielsen led a discussion about the Intercatch database status.

For national commercial catch and catch at age data, a web based system is being developed for uploading data from national systems and then editing the data by allocating sampled age- or length distributions to unsampled catches. It will then aggregate the national data to an international catch, and allow to download the output. The reason for the development is to integrate and replace different systems. This will ease data handling and assure quality control. The system builds on DIFRES' FishFrame and ICES' Datsu. Aggregation to the international level is needed as immediate input to the fish stock assessments.

Current development includes producing better data overviews, building an export file including allocation details, and catch by country, tuning fleet output, and providing pivot tables to view data on a variety of scales. The system will be tested in 2006 by working groups to see if they fulfill the need i.e. functionality and output. There is a need for fleet definitions and mandatory checks determined by the stock coordinators. Two kinds of tests are now started: final acceptance testing and calculation testing. Thus far, no major show stopper has yet turned up. Among others, CEFAS had offered to test the system on a shelf region. The Data Centre recognised a need for a training course and a manual in order to extend the use of InterCatch.

1.4.3 Progress on integration of dispersed data (ToR b)

SGMID elaborated on this issue at its last meeting. Apart from the development of a centralised fisheries data database (**InterCatch**, see above), there is very little progress with respect to data integration. There have been contacts namely between the FRS Marinelab Aberdeen on the development of a web-accessible database on the ICES Mackerel and Horse Mackerel Egg Surveys, specifically to explore how the ICES Data Centre could help with that development; and with DIFRES Charlottenlund on possibilities to link the **FishFrame Acoustics** database to ICES inventories. FishFrameAcoustics now holds raw and aggregated data from the North Sea and Baltic acoustic surveys and is significantly further developed in 2005 and 2006.

A presentation of **the Planning Group for the North Sea Pilot Project NORSEPP** (PGNSP) work again highlighted that data already available in the ICES system is still insufficiently distributed. Even worse, it is often not known that these data exist and could easily be used by other expert groups. This knowledge gap should be closed by the ICES databases as soon as possible.

The chair of PGNSP, Hein Rune Skjoldal, presented the group's work and EuroGOOS during a joint session of the two groups. PGNSP suggests that there are clear effects of physics on fish stocks, and their growth and development. Thus there is a need for data on ocean climate

to be fed into assessments. A study group began in 2005 to produce updates on a quarterly basis that can be accessed on the ICES home page.

During the winter this year (2006), there was a negative NAO anomaly that was strongest by March. It will have an effect on the circulation in the North Sea. The group is using a mathematical model to look at fluxes across various transect lines across the North Sea to see the response to the NAO change. They are using the period 1955 to 2004 as a long-term reference for comparison of the fluxes. The model predicts reduced flows in the North Sea relative to the long-term. Results show dynamic changes in the flow field that are likely to influence the biological fields.

PGNSP has sought hydrographic data obtained regularly during fisheries surveys (mainly IBTS) to use in the work, but has had great difficulty receiving them on a timely basis. PGNSP emphasized the need for a speed up with data distribution. This should be possible as some of the fish data from the same IBTS cruises were available within 3-weeks in order for them to be used in the fishery assessments. The planning group would like to include more real-time data in the reports. The basic message was that there is a need for more real-time data to make comparisons between model output and field observational data. These real data would be needed in the models for data assimilation. The data centre mentioned that real-time data are already used in the Baltic.

There was some discussion about how to get timely data from contributors. Clearly the strongest argument for needing real-time data is for data assimilation in models. The SGMID supported the need for IBTS hydrographic data (CTD) to be delivered in the same time frame as the fish data. The issue of North Sea Ecosystem prediction in relation to fish stock assessment was touched on. Some concern was voiced that the quarterly reports were not being reviewed by those making assessments, and because of this environmental concerns have not been taken into account. The question was raised which ICES expert groups use which data, and it became apparent that ICES does not have a list providing this information to assist in coordination of data and information exchange. SGMID therefore recommends that ICES produce such a list.

Zooplankton are at present not part of the PGNSP report, so the which effect the changes that PGNSP was reporting might have on zooplankton, especially *Calanus finmarchicus*, cannot be answered at present. In the light of the recruitment failure some planktivorous North Sea fish stocks are currently experiencing, zooplankton data would be very useful for those fish stock assessments.

1.4.4 New international projects on data integration (ToR b)

SeaDataNet is an EU-funded project to develop a pan-European marine data management infrastructure, which brings together 49 major institutes and marine data centres from 35 countries bordering the Baltic, the Black Sea, the Mediterranean and the North-East Atlantic, including international organizations, including ICES and IOC/IODE. SeaDataNet will be capable of managing the wide variety and large volumes of data collected by the various observing systems – both *in situ* and remotely sensed – and of serving the many user needs. SeaDataNet aims to develop this infrastructure as a virtual data centre that will incorporate and enhance the existing infrastructures of the participating countries. It is coordinated by IFREMER (France) with the technical coordination by MARIS (Netherlands). The project has approximately 9 million Euros over 5 years and began on 1 April 2006.

SeaDataNet is divided into three types of inter-related activities: networking, trans-national access and joint research.

- Networking activities cover the project management, the technical system organization and monitoring, and joint activities of the data centres including

training and implementation of common protocols for data formatting and checking for quality, development and maintenance of common directories, promotion, and organization of user feedback and forum.

- Trans-national access activities will provide access to the national marine data and information services on-line from 40 data platforms. A first group of 11 *in-situ* and satellite Data Centres will initiate the data, meta-data and product dissemination through a unique portal by adopting and implementing the data communication standards and interfacing software tools, as resulting from the joint research activities. The remaining data centres will follow, with the target of interconnecting all the participating data centres by the end of the project.

The joint research activities will develop the distributed infrastructure and cover two main areas of activity: (i) development of the technical components of the system, including standardised protocols for communication and data quality checking, complying with international standards and procedures; developing software tools for connecting the data centres and for data search, selection, retrieval, formatting and checking; portable data management software (further development of Ocean Data View (ODV)) to equip the data centres when necessary and to support training; and, common geostatistics software for computing or checking the regional gridded fields and statistics; and (ii) development of regional data products over five zones (Mediterranean, Black Sea, Baltic Sea, Northern and Polar Seas the Atlantic, and Global Ocean).

2 Implementation of the new ICES data policy 2006 (ToR a)

2.1 Communication of the new policy

As noted in section 1.4.1 above, the new ICES data policy was officially approved by the ICES council in autumn 2005, but not publicly announced until February 2006. The problem largely concerns how one deals retro-respectively with opening up of old data for unrestricted public access that were held at ICES under restrictions, and that long-term data contributors might be scared of and stop delivering data to ICES under the new policy. A discussion ensued about the need to get those and other data sets on line with open access as soon as possible and the need to provide contributors with positive reasons for doing this.

The issue of incentives for individual data sources to contribute data was considered an important point. ICES should think about ways to get the community to abide to the new rules by providing “carrots”. So far, only Finland had reacted to the publication of the new data policy and imposed an access restriction to their data for 5 years. Not much resistance from other Baltic states is expected with regard to physical oceanographic data, but ICES will have trouble with the fisheries data. The ICES’ trawl survey database manager said that ICES had to approach the individual states, while others thought it was important to approach the groups working up the aggregated data and making the reports and assessments. In this context, the freedom of information act in EU was mentioned enabling anyone to get the data upon request. Most data sources within the EU would have significant workload problems when they would be have to produce the data to external requesters. It would be much better to have the data submitted and openly accessible at ICES and then a requester could be directed to the ICES web site to get the data.

The issue of acknowledgment of who submitted or owned the data was discussed. It was suggested making a publication that includes the contributors of the data along with data products (that need to be developed). The IOC publication of world ocean database updates could act as an example. Coverage of the data, statistics, etc should be displayed. SGMID was informed that the consultative committee were not in favour of GIS products being produced

by ICES. There were apparently two separate issues: 1) integration of the data and 2) producing some products. Apparently, there are requests that ICES produce more “products”, but no specifically desired products have been described. The data centre was trying to consolidate the various databases at ICES (into DOME). The hardest problem is near duplicated data. Amongst others, OSPAR is spurring this integration. SGMID highlighted again that a lot of data are not in-house at ICES and need to be linked to the ICES Centre (see Sect. 1.4.3 above), some of them still need to be put online. The idea was raised to deal with this within an ICES Workshop. Quality assurance could be a critical issue if such distributed databases are used to make reports and assessments. If someone comes back 5-years after an assessment is made and asks for the data, access to data originally used has to be assured. A need for a frozen copy of the data used in formulating the assessment was therefore highlighted.

Communication of the new policy includes an extensive discussion about pros and cons within ICES expert groups. This discussion hasn't been taking place in most groups (partly because the implementation was much faster than expected). It appears that the opportunity has been missed earlier this year (e.g. at the annual meeting of the assessment working group chairs, AMAWGC), to discuss the policy with those groups for which the new policy has an immediate relevance.

A subgroup dealt with two issues closely related with the acceptance of the new data policy. The group felt that the present introduction to the new data policy on the ICES website would miss some essential elements (and was factually incorrect in some respects), so an amended introduction was produced and uploaded by the data centre during SGMID's meeting. The Data Centre asked SGMID to develop a document explaining the new data policy to data contributors. This document should be sent out to the present data sources, with the aim of facilitating a discussion how “historic” data could be made available under the new policy. After an extensive discussion, the group finalised such a letter asking specifically for the definition of a period of exclusive use. The letter should be sent out soon and feedback is expected until the Annual Science Conference in September. If there is no feedback, it is assumed that old data already physically held at ICES will also be available under the rules of the new data policy.

2.2 Review of ICES expert group's data policies (ToR e)

The only reactions to the new ICES data policy were received from the Baltic International Trawl Survey WG (BITS-WG) and the International Bottom Trawl Survey WG (IBTS WG). The first concluded that their data should be restricted mainly for quality assurance purposes for two years and then publicly available. In contrast, while acknowledging the benefits the new policy would have for the expert group's work, the IBTSWG expressed serious concerns about an open access to their data. The main reason for this was the fear of misinterpretations or just interpretations deviating from their present view. The IBTSWG therefore requests the Data Centre to impose a complex system of access restrictions to their data.

SGMID discussed this part of IBTSWG's report extensively and concluded that such access restrictions would be in clear conflict with the new data policy. There is apparently a need to discuss the policy openly with that group, and to try to convince them that the new policy also has enormous benefits for the data source. On the other hand, SGMID highlights that the only restriction to be imposed by the data source under the new policy is a temporal one, and any restriction for more than two years (i.e. a period needed for quality control) would need a thorough justification. It also has to be considered that different interpretations of the same dataset are a key element of scientific progress, so these are not per se negative. To accommodate some of IBTSWG's concerns, it is suggested that that group should produce an exhaustive list of “known issues” of their dataset(s) which is submitted to data requesters every time data are downloaded. Also, if some of the data from the very early period of the

survey are thought so full of errors that they are useless to the public, there might be a possibility not to grant unrestricted access to this specific data until they are critically reviewed. The letter from the ICES secretariat to data sources drafted by SGMID (see above) is thought to facilitate the needed discussion.

2.3 Technical amendments to ICES databases

In order to effectively implement the new ICES Data Policy, some changes across a variety of ICES data systems are required. Additional information must be captured as accessions are logged within the DAD system. This information is currently entered by Data Managers within ICES Data Centre. In many cases they will be aware of the extra information required, but some dialogue with the data submitters may also be required. For example:

- The Acknowledgement associated with each accession needs to be made mandatory (i.e. from now onwards this field must be completed). If not supplied then a default “Acknowledgement information not available” should be recorded in DAD.
- For every file submitted a Restriction Type plus Release Date needs to be identified. The Restriction Types are:
 - i) Publicly available – defaults to today’s date
 - ii) Available to ICES expert groups
 - iii) Not to be released – applies to raw commercial fisheries data

A standard list should be provided for Release Dates – e.g. today, 1 year, 2 years, 5 years.

- Conditions of use:
 - iv) Generic ICES conditions need to be formulated for supply with all data. An example of this type of conditions as used by BODC is attached in Annex 2.
 - v) Specific conditions if any apply to each accession are to be supplied by the submitter.
- Problems with the Data – any problems or issues with the data file should be recorded, so that these can be provided to future data users as necessary.

Once an entry has been completed in DAD, the information could be emailed back to the submitter to confirm the details are correct.

More meta-information must be supplied along with the data downloaded/extracted by web-users, in accordance with the ICES Data Policy. SGMID proposes that an information or metadata file to be supplied with every data extract. This file should include the following:

- General ICES conditions of use/disclaimer.
- Specific conditions that apply to the particular dataset, as supplied by the submitter.
- Any acknowledgements to be used.
- Definitions of Quality Flags used.

Before an extract is downloaded, the user should be prompted that this metadata file is available and it should be read. For example: “An information file describing conditions of use is supplied with the data. Please confirm that you will read and accept the conditions as described in this file.”

SGMID suggests that people extracting data on the website be encouraged to provide information on their intended use of the data, in order to help ICES Data Centre identify how to improve the services available.

Working Groups should be encouraged to submit aggregated data via the DAD system so that these data are also publicly available.

The ICES Data Centre should review the release conditions that apply to the historic archives they hold. All oceanographic data >10 years old are already freely available. Other conditions should be first compiled and then reviewed with the relevant institutes to assess if the new data policy can be applied.

3 The ICES Data Centre User Survey (ToR c)

SGMID was asked to evaluate the outcome of the ICES Data Centre User Survey. The questionnaire was prepared prior to the Annual Science Conference in Aberdeen, Sept. 2005, handed out to all registered ASC attendees along with the “welcome package”, and participants were encouraged on various occasions to fill in the survey. 140 filled-in questionnaires were received. The detailed evaluation per question is listed in Annex 3.

The questionnaire was also handed out to members of WGOH, but the very few returns (6) were kept separate for this evaluation.

A number of general conclusions can be drawn from the answers, also considering the apparent limitations of the survey:

1. Respondents don't necessarily represent the ICES community as a whole, only those that are inclined to attend the Annual Science Conference. Further, it is likely only those who had some interest in data management issues took the time to provide answers. It is also difficult to say whether or not there is a bias toward responses by fisheries biologists. This would have to be determined by appropriate questions in future questionnaires.
2. Given that a significant fraction of individuals responding to the questionnaire were interested in data, the ICES databases and web site may be under utilized. A web-statistics program should be setup to try to log the access of the various ICES databases and put together statistics that can be compared to the survey responses to question 11 – “indicate the types, sources, and forms of data which you use”. Also, there are apparently members of this community that have the skills to technically help ICES develop the databases or web-based access. ICES should consider developing a technical users group to take advantage of their expertise.
3. The question (8) on the type of workstation used apparently stemmed from InterCatch development and the need to know to what extent non-Windows users make up the ICES community. This question appears to be irrelevant, as the term “PC” is not defined (a Mac is also a PC, while the appropriate alternate category would have been “mainframe”). What really matters is the operating system and browser being used. From this survey, it would appear that about 20% use a non-windows operating system and more than 30% use a non-Microsoft browser. Therefore, the software developed by ICES should strive for platform independence.
4. The majority of respondents sought data associated with fisheries surveys or catches in ICES databases. Also sought was oceanographic (CTD) and to a lesser extent was biological community data. Other categories including hydro-chemistry and contaminants data were specified much less frequently. It was, however, unclear in the question whether ICES referred to data actually in the ICES database or data collected under the ICES banner (like acoustic survey data). The answers on question (19) on how often requests have been sent to ICES databases indicates that ICES does not hold much data of interest to the respondents, the data are not readily accessible, or respondents are not aware of what data are available. Alternatively, the ICES database might serve a specific subset of the ICES community: those making resource management decisions. Most respondents use/are interested in trend data,

and in gridded or GIS outputs. The development of tools to provide these kinds of data on the ICES web page appears to be useful to the community.

5. The possible answers provided to question (20) concerning ICES data amount, quality, products, web tools, and response-time can be viewed as allowing three positive answers (excellent, good, satisfactory) and one negative answer (not satisfactory). Most of the respondents were split between good and satisfactory, which indicates that they were ambivalent about the ICES data. However, the way the question was worded makes it difficult to interpret whether the majority has a rather negative or a rather positive view on these issues.
6. Many of the questions had a set selection of answers with a single line to respond to another check mark. Future surveys should provide more free form space on specific questions to encourage people to provide opinion.
7. The results of the questionnaire should be made available to the people who responded and the rest of the ICES community as a first step in improving the communication link between ICES and its constituents. SGMID suggests to the Data Centre to produce a poster to provide the results of the survey in the data management theme session.
8. A *SurveyMonkey.com* type application may be useful for an online survey [<http://www.surveymonkey.com/>].

4 Quality assurance, development of quality flags (ToR d)

The SGMID discussed data quality flags at the 2005 meeting and concluded that the scientific value of integrated datasets critically depends on the quality and precision of the data. For scientists to perform analysis and provide sound advice they need to be aware of and have the ability to select data based on known levels of quality. The precision level required may vary for different analyses, for example a fisheries biologist using salinity data combined with stock abundance may be prepared to accept a lower level of data quality than a physical oceanographer modelling water masses.

Some data requests may require a simple good data / bad data selection whereas other requests may define a much deeper level, for example precision of instruments used for the data collection or the analysis method undertaken.

The present situation is that there is no consistency within the various databases maintained by ICES or more widely in the marine community. Currently no quality flags are stored within the ICES oceanographic database, although comments on data quality and any changes made to the data are noted and stored in text files. Some information may also appear in the ROSCOP/Cruise Summary Report (CSR) database. Any changes to the data are also discussed with the data supplier. A preliminary review of data quality flagging schemes in use within the marine community shows small variations on a theme for oceanographic data, but more complicated and detailed schemes do also exist. The most straightforward solution would involve mapping the various originators' schemes to a simple scheme – perhaps comprising the following:

- Not checked
- Good/correct value
- Doubtful/suspect value
- Bad value

Those organisations that provide data to ICES will not (and should not) change their data quality flagging systems. ICES will (and already does) store a copy of the original data

received (obviously including the supplier's data quality flags). From here there are two possibilities: either the original flags can be included in the ICES Data Centre databases (DOME) and also be mapped to a more simple set of quality flags, or as part of the conversion of data to the DOME system the simple set of flags only will be used. A definition of the system of quality flags used would be supplied by ICES in a meta-file linked to any data extracted.

The EU SeaDataNet project, as part of its standards development activity, includes the development of an agreed quality flag scale protocol. This is an early deliverable, with a draft due at the end of month 3 (e.g. end of June), with a validated version by the end of month 8. ICES, as a partner in SeaDataNet, will contribute to this. However, it should be noted that the agreed SeaDataNet scheme is being developed primarily for time-series and profile data may not be appropriate to all data types, in particular for fisheries data. Due to the often hierarchical nature of fisheries data (both scientific survey and commercial catch data) a flag at a higher level of data may not necessarily apply to all of its child data. This is particularly apparent for data that are a result of subjective scientific observations (such as sexual maturity stage or age) where only a small subset of the observations can be subject to quality control because of the labour intensive nature of the work.

5 Preparation of the ASC 2006 theme session on data integration (ToR f)

SGMID and WGMDM will co-sponsor a theme session on "Environmental and Fisheries Data Management, Access, and Integration" at the September 2006 Annual Science Conference in Maastricht. ICES staff informed both groups that 34 oral and 15 poster presentations were submitted to this theme session M. Session M appears to be the most attractive of the 18 theme sessions, accounting for 10% of the total contributions of 390 oral and 63 poster presentations received for the ASC 2006. The groups were told that ICES would decide upon the slots and total time allocation for the theme sessions shortly. It is expected that each oral presentation would have 15 mins allocated; the number of posters per session is likely not to be limited.

A steering committee for the detailed planning of the theme session was formed, consisting of the three chairs of SGMID and WGMDM (Helge Sagen, Peter Wiebe, Christopher Zimmermann) plus Lesley Rickards (UK, SGMID/WGMDM) and Robert Gelfeld (USA, WGMDM). The group will meet by correspondence after the ICES secretariat has determined the time available for oral presentations in this theme session, and will select the presentations and structure the session.

6 The future of SGMID (ToR g)

The SGMID discussed its future perspectives extensively. The group has met now officially on three occasions and as a Study Group has reached an ending point. This leaves a variety of options that were already listed in last years report:

- Dissolve the group without replacement – this would mean that the expertise of the group and its capacity to support and give advice to the ICES Data Centre would be lost.
- Merge the group with WGMDM – this has the drawback that both groups focus on very different aspects: While WGMDM is a rather technical, hands-on group under the Oceanography Committee, SGMID is user-driven and addresses more conceptual and policy related questions under ACE. There might be scope for a

change of the name of one of the groups (possibly the newer one, SGMID) to emphasize this difference. Another drawback is the likely size of a merged group, which will be difficult to handle and lose efficiency. Efficiency and impact are, however, key factors to attract scientists from the user side to such a group.

- Transfer SGMID into a working group that could give advice to the ICES secretariat as required. This group could meet regularly, but not necessarily annually, and do most of the work by correspondence. This would have the drawback that technical issues could not be addressed by the new group, and there would be scope for duplicate work with two distinct data groups within the ICES system. Both groups would then have to have very distinct ToR and names, which make the different perspective of both easily perceivable.

For options 2 and 3, it should be explored if another parent committee (currently ACE for SGMID and OCC for WGMDM) would be more appropriate and would provide a wider dissemination of its work.

SGMID discussed these options openly. There also was a joint session of both groups held on the last day of the meeting with the aim of exploring possibilities to merge the groups. Members of both groups were made familiar with the history, tasks achieved and possible future perspectives. Duplicity of work and of data was reviewed briefly. The option was mentioned that with SeaDataNet coming (see above) and dealing with a variety of technical issues, these should perhaps be transferred to them for development and study. It was re-emphasized that SGMID's lifetime is over, that it has been set up to do a specific task which was delivered and had, especially in relation to its size, significant impact to the system. It could therefore be dissolved without much pain. There are, however, thought to be benefits to keep the integrity and expertise of the group as a number of issues would require additional work, and as more issues to deal with appear every year (see below), ranging from questions where data to be integrated resides physically, through questions how to deal with international agreements that will impact ICES from a top-down level. Also, having a new policy now gives rise to the need for clear objectives to head for in terms of product development. To attract data, there is a need for ICES to show what benefit there is in way of products that can be provided. So the next group could be charged to work on defining this visibility and products. This could be used as carrot to data submitters. It is also thought that a new mandate for a data advisory group might be emerging that stems from ecosystem based management needs and assessments.

So the fate of SGMID was still uncertain in the minds of SGMID members, while WGMDM members mostly appeared to welcome a merger as this is seen as a chance to modernise the group without losing WGMDM's integrity. Most SGMID members indicated that it would not be very attractive to them to continue in WGMDM in the present scenario, when they would only make up a quarter of the participants but tasked with the majority of the ToRs. By the end of the meeting, it became clear that this issue would have to be discussed between the three chairs of the groups, the chair of the Consultative Committee and maybe the chair of OCC, ACE and the General Secretary.

Postscriptum: *The discussion on the future of ICES expert groups on data issues then took place at the ASC in September 2006. CONC-, OCC-, WGMDM- and SGMID chairs reached the agreement that both groups are to be dissolved and a new group is formed (WGDIM), initially and for only one year under the chairmanship of the three chairs of the dissolved groups, maintaining the balance between technology, environmental users and fisheries users backgrounds. This new group would have a new membership and should not be seen just as a successor of WGMDM, and it was also made clear that the participation of members with a technical and a user background should be balanced. The ToRs for the new group were revised in this light.*

Future issues for a data advisory group in ICES from SGMID's perspective

ICES is now charged to provide ecosystem based advice and develop ecosystem management concepts, especially with respect to fishery assessments. To do this, requires the most effective use of data. The new data policy serves to enhance this possibility.

1) Data availability: There are major gaps in the ecosystem assessments apparently caused by lack of data. However, more data are likely available for use than currently perceived either inside the ICES system or externally. Thus, groups developing the advice may not be aware of the existence of relevant data sets either because of a lack of communication or the fact that data not being delivered on a timely basis. In addition, those environmental assessments that are now being produced by some ICES working groups are not being effectively utilized by other groups making assessments where environmental data should be considered (NORSEPP, WGRED).

Within the ICES working group reports, trends and other kinds of information are present in figures and tables, but not available electronically. This makes it difficult to be useful in assessments.

Conclusions: A) Communication between ICES expert groups needs to be improved. B) Data contributors need to be encouraged to submit data when they are useful, not when they are completely quality controlled. [this leads to quality flags and data updates]

Action for conclusion A: A group (or the data centre?) should be charged to produce a data availability and usage map.

Action for conclusion B: There is a need for proposed technical solutions to assist the ICES Data Centre to handle this issue.

2) Transparency: Many of the data that are being used to make the environmental assessments do not reside within the ICES and little effort is being expended to track the data used to make the assessments. If the external data are being used to formulate advice, it is often difficult some time later to re-establish the data sets and thus the basis for the advice.

3) Products based on ICES data holdings: Trend plots and gridded data products are desired by ICES Annual meeting attendees (as derived from the user survey). These would serve as an incentive to the data contributors.

4) Monitoring of Data Policy Implementation: A lot of communication appears to be needed between data users and data sources to make the latter supporting the new policy. In addition, the ICES secretariat is expected to require support for decisions how deal with some data requests in the implementation phase.

Action for 2, 3, and 4: Identify areas of concern and give guidance to the ICES Data Center for scientific approaches and technical solutions.

5) The data centre needs to develop strategies that enables it to be a focal point for data storage and distribution to the ICES community. The strategies should be user driven.

Action for 5: Ultimately ICES should develop the means for user feedback about the ICES data centre and its effectiveness, for example through development of an online user survey. There needs to be an independent group evaluate the survey information.

7 Recommendations

SGMID recommends:

- to the ICES data centre, to develop the technical aspects required for a full implementation of the new ICES data policy as outlined in Sec. 22.3 of this report. ICES data bases should *inter alia* hold information on the data source (including contact details) and known issues about/limitations of the data to guide potential users and avoid misinterpretation.
- to the ICES secretariat, to produce an extensive list of the data which expert groups produce, process, or require to assist in the coordination of data exchange.
- to data contributors, to produce a data specific list of “known issues” with their data, which is to be submitted to a data requester every time data are downloaded.
- to the ICES Consultative Committee, to create a new expert group advising the ICES data centre and dealing with the issues outlined in the section 6 above.

In addition, the following recommendations made in last years report are still valid and are repeated herein.

- to the ICES secretariat to continue with the exploration of possibilities to acquire data not currently accessible to the ICES community, but needed for the production of advice and status reports
- to the ICES Data Centre and WGMDM, that the DiGIR system (Distributed Generic Information Retrieval), used by the Ocean Biogeographical Information System (OBIS) and GBIF, should be further investigated and possibly installed at the ICES secretariat.
- that database developments within ICES DC should include provision for storage of data quality indicator(s) and that the exchange files supporting these developments include the necessary data structure. When data that contains data quality indicators are submitted ICES should retain this information within the ICES database and make it available to the data users.
- to the ICES expert groups and the Data Centre, to conduct appropriate quality control procedures that can not be performed by individual data submitters. This includes consistency checks and comparisons between repeat cruises and the analysis of cross-over data. Problems should be reported back to the data submitter. Any correction or calibration change should be recorded in the associated meta data log.
- that ICES maintain (or develop further) their standard definitions and codes. However, in order to promote the ease of use and acceptance of a database also for inexperienced users, codes should have to be used by data submitters and requesters only when unavoidable.
- that sufficient funding is made available to ICES and national data centres to rescue data currently not available for ICES expert groups and the public. These include data available only non-electronically in historic ICES reports.

8 References

- ICES 2004: Report of the Study Group for Management of Integrated Data. ICES CM 2004/ACE:05
- ICES 2005: Report of the Bureau Working Group on the Data Development Project. ICES CM2005/June 2005/Bur. Doc. 1419.

ICES 2005: Report of the Study Group for Management of Integrated Data. ICES CM 2005/ACE:03

ICES 2006: Report of the International Bottom Trawl Survey Group (IBTSWG). ICES CM 2006/RMC:03

ICES 2006: Baltic International Fish Survey Working Group (WGBIFS). ICES CM 2006/LRC:7

Annex 1: List of SGMID participants

NAME	ADDRESS	TELEPHONE	FAX	E-MAIL
Larry Atkinson	Dept. of Ocean, Earth & Atmospheric Sciences Old Dominion University, Norfolk VA 23529 USA	+1 757 683 3472		latkinso@odu.edu
Richard Ayers	CEFAS Lowestoft Laboratory Lowestoft Suffolk NR33 0HT United Kingdom	+44 1502 524224	+44 1502 513865	r.a.ayers@cefas.co.uk
Steve Flatman	CEFAS Lowestoft Laboratory Lowestoft Suffolk NR33 0HT United Kingdom	+44 1502 562244	+44 1502 513865	s.flatman@cefas.co.uk
Martina Hennessy	Marine Institute Rinville Oranmore Co. Galway Ireland	+353 91 387200	+353 91 387201	Martina.Hennessy@marine.ie
Todd D. O'Brien	NOAA - NMFS - Science & Technology F/ST7 - Room 12503 1315 east West Highway Silver Spring, MD 20910-3282 USA			tobrien@nodc.noaa.gov
Lesley Rickards	British Oceanographic Data Centre Joseph Proudman Building 6 Brownlow Street Liverpool L3 5DA UK	+44 151 653 1514	+44 151 652 3950	ljr@bodc.ac.uk
Yolanda Sarminaga	AZTI, Spain			ysagarminaga@pas.azti.es
Reiner Schlitzer	Alfred-Wegener-Institute for Polar and Marine Research Postfach 12 0161, D-27515 Bremerhaven Germany			rschlitzer@awi-bremerhaven.de
Peter Wiebe (Co-Chair)	Woods Hole Oceanographic Institution Woods Hole, MA 02543-1050	+1 508 289 2313	+1 508 457 2169	pwiebe@whoi.edu

NAME	ADDRESS	TELEPHONE	FAX	E-MAIL
	USA			
Christopher Zimmermann (Co-chair)	Fed. Res. Centre Fisheries Institute for Baltic Sea Fisheries Alter Hafen Süd 2 D-18069 Rostock, Germany	+49 381 81161-15	+49 381 81161-99	christopher.zimmermann@ior.bfa-fisch.de or czimmermann@clupea.de
Ole Folmer	Danish Institute for Fishery Research (DIFRES) Charlottenlund Slot DK-2920 Charlottenlund Denmark	+45 33963469	+45 33963333	ofodifres.dk
Hans Mose Jensen	ICES Secretariat H. C. Andersens Blvd. 44- 46 1553 Copenhagen Denmark	+45 33386736		hans.jensen@ices.dk
Henrik Kjems-Nielsen	ICES Secretariat H. C. Andersens Blvd. 44- 46 1553 Copenhagen Denmark	+45 33386736		henrikkn@ices.dk
Julie Gillin	ICES Secretariat H. C. Andersens Blvd. 44- 46 1553 Copenhagen Denmark	+45 3338 6712 Mobile: +45 22444910	+45 3393 4215	julie@ices.dk
Lena Larsen	ICES Secretariat H. C. Andersens Blvd. 44- 46 1553 Copenhagen Denmark			lena@ices.dk

Annex 2: Example "Conditions of Use" for data for UK National Tide Gauge Data from BODC

Preamble

It is BODC's policy to encourage the use of its data holdings. Oceanographic data are expensive to collect and, in particular, resources are required to maintain them and to make them available. The Environment Agency, the funders of the UK National Tide Gauge Network, and BODC are committed to providing these data free of charge. However, it is important that users of the data co-operate in making known their intended use of the data and in reporting back on the outcome of that use. Such information is required by BODC as evidence to:

- a) justify the continued maintenance of the data.
- b) demonstrate the scientific and economic value of making the data available.

Conditions of Use

1. Users must not pass on the data to third parties but instead direct interested parties to the NTSLF (National Tidal and Sea level Facility) web site.
2. BODC, NERC and the Environment Agency give no warranty as to the accuracy of the data or as to the suitability of the data for their intended use by the recipient and shall not be responsible for any errors or omissions, or for the use of, or results obtained from the use of this information.
3. Publications resulting from the use of the data shall include acknowledgement to BODC and the Environment Agency (Suggested text "The data were supplied by the British Oceanographic Data Centre as part of the function of the National Tidal & Sea Level facility, hosted by the Proudman Oceanographic Laboratory and funded by the Environment Agency and the Natural Environment Research Council").
4. The recipient shall:
 - a) Inform BODC of the outputs to which data have contributed e.g. references to scientific publications or reports.
 - b) Offer to BODC, if requested and without charge, a copy of any value added data set or product derived from the data.
 - c) Inform BODC of any inaccuracies or errors detected in the data.
5. BODC is entitled to terminate the licence at any time if the conditions of use are not observed and may revoke privileges in future if there is unauthorised redistribution of the data.
6. All intellectual property rights in the data and information supplied to you whether owned by NERC or third parties (e.g. Environment Agency) will continue to be owned by the respective parties.

Disclaimer

While BODC believes the information to be reliable, human or mechanical error remains a possibility. Therefore, the NERC does not guarantee the accuracy, completeness, timeliness, or correct sequencing of the information. Neither BODC, NERC nor any of the sources of the information shall be responsible for any errors or omissions, or for the use of or results obtained from the use of this information.

Annex 3: Detailed results of the ICES Data Centre User Survey

This annex gives the analysis of the answers by question. A summary can be found in the report section 3.

1. What is your position?

Of 140 respondents, 106 selected Scientist, 14 selected student, 20 selected either, administrators, IT, or other.

Conclusion: The majority of respondents were potential users of the ICES database and data services, or submitters of data to ICES.

2. What is your organization?

Of 124 respondents, 80 selected 'National Government' organization, 41 selected academic, and 9 selected Inter-governmental, Non-governmental, private, or other. The majority of respondents were from governmental or academic organizations.

3. In which country do you work?

UK 22, USA 14, Germany 12, France 11, Spain 10, Norway 8, Netherlands 6, Finland 5, Canada 4, Ireland 4, Sweden 4, Denmark 2, Belgium 2, Portugal 2, Iceland 2, Poland, Italy, Morocco, Holland, Faroe Islands, Greenland, Greece, Latvia, Malaysia. Twenty-four countries were represented by the respondents, mostly from Europe or North America. A second polling of hydrographers (WGOH) had respondents from Netherlands, UK 2, Germany, Spain, and Ireland.

4. What computer skills do you have? (check all that apply)

Basic users were slightly less than half of the respondents and knowledgeable or expert users were more than half of the respondents. Virtually all respondents have basic computer skills and are likely to be able to use web based access to the ICES database.

Conclusion: There are people in this community that have the skills to help ICES develop the technical development of the databases or web-based access. ICES should consider developing a technical users group to take advantage of their expertise.

5. How do you participate in ICES?

Out of 140 respondents, about 2/3 attend working or study groups, and they are active members of the ICES community. About the same number acknowledge reading ICES publications.

6. How often do you visit ICES website?

Out of 125 respondents to this question about 2/3 visit the ICES web site at least monthly. A significant fraction (1/3) rarely or never visit the web site.

7. Your relation to ICES data? (check all that apply)

Out of 140 respondents about 1/3 'contribute' data, about 2/3 'use' data, and a significant fraction do 'neither'.

8. What type of workstation do you use?

Out of 134 respondents, the vast majority use PC, as opposed to Macs or others.

Conclusion: This question is of no importance compared to the following that addresses the operating system.

9. What operating system do you use?

Out of 149 respondents, about 20% use an operating system other than Windows.

Conclusion - This is a significant fraction suggesting access to the ICES web site and database(s) should accommodate these users as well as Windows users.

10. Which browser do you use?

Out of 150 respondents, about 34% use a browser other than Internet Explorer. Most of these use a variant of the open source code based Netscape/Mozilla/Firefox.

Recommendation: The ICES web interface needs to accommodate nearly all of these users.

11. Please indicate the types, sources, and forms of data which you use (check all that apply). Of the types of ICES data specified in the questionnaire, which ones were most selected or specified as being used (rank order)?

Table 1. Rank order of data use specified by respondents with ICES as the data source.

QUESTION NO	QUESTION TEXT	TOTAL
11,13	Fisheries trawl survey	42
11,12	Commercial catch	31
11,01	CTD (high resolution profile)	25
11,14	Acoustic fishery survey	19
11,08	Biological Community	15
11,10	Habitat and Bioregions (e.g. Spawning areas, bottom types)	11
11,02	Hydrochemistry (bottle)	9
11,04	Contaminants in Sediment	5
11,05	Contaminants in Biota	5
11,06	Biological effects including fish disease	5
11,07	Intercalibration results	5
11,15	Satellite data	4
11,09	Continuous plankton recorder	3
11,03	Contaminants in Seawater	2
11,16	Other	2
11,11	Alien Species	1

Conclusion: The majority of respondents sought data associated with fisheries surveys or catches. Also sought was oceanographic (CTD) and to a lesser extent was biological community data. Other categories including hydrochemistry and contaminants data were specified much less frequently.

Notes: it was unclear in the question whether ICES referred to data actually held in the ICES database or data collected under the ICES banner. So, for example, while there are no acoustic survey data held in the ICES Data Centre, some respondents indicated ICES as source for acoustic survey data. They apparently did this because

the data were collected as part of ICES surveys even if the data were held elsewhere. In contrast, it is not clear why four respondents said they used ICES satellite data: None is available on the ICES web site nor gathered under ICES initiatives.

12. Do you use trend data?

Out of 76 respondents, about 70% said they use trend data.

Conclusion: This implies that most use time-series data sets or are interested in time-series patterns - a trend up or down. Therefore if ICES produced trend based products they might be found useful to the community.

13. Do you use gridded or mapped or GIS products?

Out of 114 respondents, about 70% said yes.

Conclusion: ICES produced gridded or mapped products might be found useful to the community. More individuals said they used these products than trend data products.

14. Time period(s) of interest (check all that apply)

Most preferred to have access to all time periods, although only small number wanted only data from prior to 1980. In addition, a minority only wanted recent data.

15. Eco-regions of interest (check all that apply)

Of the 330 regions specified, 85% selected the Northeast Atlantic (ICES regions), 9% selected Northwest Atlantic, 3% selected Mediterranean, and 3% from other world ocean areas. Most of the respondents were ICES centric as might be expected considering the participants.

16. Preferred data exchange format(s) (check all that apply)

Spreadsheet/comma-separated file format was strongly preferred over the other choices. Few selected the ICES formats and few selected the XML format.

Conclusion: This suggests that in order for individuals to download data from the ICES databases, simple formats as well as more sophisticated self-describing formats will be needed. If self-describing formats are going to be encouraged in the future considerable education must be part of the plan.

17.1 Preferred database type for storing your data:

Nearly half the respondents preferred to have their data on a local database and the others had no preference, did not know, or favoured a distributed system.

Conclusion: it is becoming irrelevant for the user whether data are stored locally or in a distributed system. SGMID commented already last year that most users would not even recognise the difference between the two systems when downloading data.

17.2. Preferred database for retrieving data from other sources:

About 32% of the respondents preferred to access data on a centralized database. Most had no preference, did not know, or favoured a distributed system.

Conclusion: it is becoming irrelevant where data are obtained.

18. Preferred data medium (check all that apply)

The vast majority of respondents preferred to download data via the internet. CD ROM was a distant, but important second method.

19. How many times have you downloaded or requested data from ICES in the past year?

About half did not download data in the past year from ICES. Most of the others had done it less than 5 times. Only 7 said that they downloaded data more than 5 times.

Conclusion: ICES does not hold much data of interest to the respondents, the data are not readily accessible, or respondents are not aware of what data are available.

Alternate conclusion: The ICES database serves a specific subset of the ICES community: those making resource management decisions.

20.1 Amount of data available via ICES website.

The majority of respondents thought the data were either good or satisfactory, but not excellent. Only half of the respondents to the questionnaire answered this question.

20.2 Quality of ICES data.

The majority of respondents thought the data were either good or satisfactory, but not excellent. Only half of the respondents to the questionnaire answered this question.

20.3 Data products

The majority of respondents thought the data products were either good or satisfactory, but not excellent. Only half of the respondents to the questionnaire answered this question.

20.4 User-friendliness of web tools.

The majority of respondents thought the web tools were either good or satisfactory, but not excellent. Only half of the respondents to the questionnaire answered this question.

20.5 Response time for manual data requests.

The majority of respondents thought the time for manual data request were either good or satisfactory, but not excellent. Less than half of the respondents to the questionnaire answered this question.

20.6 If you are at all unsatisfied, please explain why

Eighteen people chose to provide some explanation.

Overall, academics and governmental respondents had about the same proportions in responding to the 20 series questions.

Conclusion for all parts of question 20: A majority of the respondents chose the 2nd or 3rd of four choices for satisfaction. This suggests a considerable degree of dissatisfaction with the database.

21. Which ICES tools do you use? (check all that apply)

Of 101 respondents, about equal number of them use Accessions (23%), Integrated Inventory (20%), and Oceanographic data maps (26%). Fewer used the ROSCOP, reference codes, and data screening tools.

22. Are you familiar with the ICES data policy?

Sixty-two % were not familiar with the ICES data policy. When the survey was conducted, it was the old policy that they would have expressed knowledge about.