

ICES WGDIM REPORT 2010

SCICOM STEERING GROUP ON HUMAN INTERACTIONS ON ECOSYSTEMS

ICES CM 2010/SCICOM:05

REF. SCICOM, ACOM

Report of the Working Group on Data and Information Management (WGDIM)

25–27 May 2010

Copenhagen, Denmark



ICES

International Council for
the Exploration of the Sea

CIEM

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. 2010. Report of the Working Group on Data and Information Management (WGDIM), 25-27 May 2010, Copenhagen, Denmark. ICES CM 2010/SCICOM:05. 39 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.

© 2010 International Council for the Exploration of the Sea

Contents

Executive summary	3
1 Opening of the meeting.....	4
2 Adoption of the agenda.....	4
3 Data accessibility (ToR a).....	4
4 Quality, interoperability and transparency (ToR b)	5
5 ICES Data Strategy (ToR c).....	8
6 VMS (ToR d).....	8
7 Assess progress with, and update, user engagement plan (ToR e)	10
8 DATRAS (ToR e)	11
9 Other Items	11
Annex 1: List of participants.....	13
Annex 2: Agenda.....	14
Annex 3: ToRs 2009/2010	17
Annex 4: WGDIM proposed terms of reference for the next meeting	18
Annex 5: Recommendations	20
Annex 6: Proposed ICES Data Strategy.....	21
Annex 7: WGDIM Data Strategy Working Document.....	22
Annex 8: User Engagement Strategy.....	26
Annex 9: Egg/Larval Database Action Plan	27
Annex 10: DUAP Report	29
Annex 11: Protocol/Process for development of data products (DRAFT)	37

Executive summary

WGDIM's transition to a group providing strategic and tactical advice is now complete, the group this year has focussed on issues that impact the wider ICES community along with specific tactical advice on data management issues.

The group has proposed a new ICES data strategy and recommended the adoption of a generic data quality flagging system that can be applied across all of ICES' data holdings.

WGDIM is working towards becoming more involved in the continuing development and maintenance of international standards with greater interaction with GBIF and MarineXML.

In 2009 WGDIM developed its mission statement:

"To provide ICES with advice on all aspects of data management including data policy, data strategy, data quality, technical issues and user-oriented guidance."

This mission has guided the work of the group over the last year and continues to hold true for the coming year.

1 Opening of the meeting

The 2010 meeting of the Working Group on Data and Information Management (WGDIM) was held at ICES Headquarters, Copenhagen from 25 to 27 May 2010. The meeting was attended by 25 scientists (Annex 1) and members of the ICES Data Centre. The agenda (Annex 2) was adopted. The terms of reference for the meeting are given in Annex 3.

The working group thanks ICES (Vivian Piil) for making the local arrangements and support during the meeting

2 Adoption of the agenda

The agenda for the meeting was agreed and amended during the meeting to facilitate discussions.

3 Data accessibility (ToR a)

Year of the stomach

The project to rescue the 'Year of the Stomach' data has been completed for North Sea data. The project was considerably more complex and took longer than anticipated. There were a number of issues resulting from trying to reconcile 2 data sets into a single data set whilst avoiding duplication of data. The completed dataset holds in excess of 250,000 individual stomach records.

The completed database can be found at – <http://ecosystemdata.ices.dk/stomachdata>

Before the next WGSAM meeting the Data Centre will evaluate the newly received version of the Baltic stomach data, with the aim of making it provisionally available online.

WGDIM recommends; WGSAM (the group that requested the work)

- a) Provide feedback to WGDIM and the Data Centre on the North Sea element and
- b) Consider the Baltic dataset (in its current state); whether there is enough value in undertaking further work and if so what resources are needed (i.e. volunteer experts from the Baltic States along with IT experts from the Data Centre).

Multi-disciplinary data

A presentation was given showing the functionality of the EcoSystemData system <http://ecosystemdata.ices.dk>. This system allows users to overlay a variety of datasets from multiple disciplines, e.g. fishery survey and oceanographic data, via a user friendly GIS style interface. This system is the first step in allowing users to investigate the possibilities of incorporating additional information into advice and scientific work. It is hoped that highlighting these possibilities will lead to increased awareness of the data sets available from ICES and their interactions.

Egg / Larval Database

Fish egg and fish larvae data have been collected in the ICES community for a long time for use in stock assessments. The collection of the data is usually organised by

international survey expert groups. The data are stored in national institutes and, in most cases, an expert group keeps a copy of the combined data to be able to supply the assessment working groups with the information needed. In this way, the information is available for the main purpose, but it is not accessible to the wider marine community.

Currently, there are a number of data sets which are not readily available to the wider marine community but coordinated by ICES groups and used in ICES stock assessments. These are:

- Mackerel and horse mackerel eggs Northeast Atlantic (WGMEGS)
- Cod and plaice eggs North Sea (WGEGGS),
- Herring larvae North Sea, herring larvae Western Baltic (both PGIPS)
- Herring larvae North Sea (MIK-IBTSWG).

Since these data sets are very similarly structured, and already worked up and quality controlled, they are considered a good starting point for the development of an ichthyoplankton database within ICES. It was therefore decided at WGDIM 2010 to create an action plan to collate ichthyoplankton data at ICES (Annex 9).

By developing such a database structure, the national data will be secured at an international data clearing house, it will be possible to provide an overview of available fish egg and larvae survey data collected and to provide a unified portal for scientific access to the ichthyoplankton survey data.

WGDIM recommends: The action plan is completed in conjunction with the Data Centre and expert groups and then taken forward

4 Quality, interoperability and transparency (ToR b)

Quality

A WGDIM subgroup began work on a proposal for a generic data quality control flag system that could be applied across all ICES' data holdings before the meeting and presented a working paper to an enlarged sub-group during the meeting. The enlarged sub-group took the working document forward and has proposed a conceptual quality control flag system that needs further discussion and refinement before it can be considered as a candidate for adoption by ICES. It should be noted that the proposed flags are for application to values within detail level data, i.e. individual records of measurements.

Within ICES data holdings there are 3 types of quality control flags:

- Contributor's quality control flags – Flags generated by the data contributor, these should be kept in the ICES database in their original form. Ideally ICES should also hold information regarding the QC/QA checks that were undertaken to produce these flags.
- Objective quality control flags – Flags that are assigned based on automatic quality control tests.
- Subjective quality control flags – Flags assigned by data managers or experts based on visual inspection of the data. They can be assigned from within ICES or from an advisory or science group.

There are a number of 'standard' quality flag schemes in use on an international basis, often a numeric scheme with an attached text description. In the tables below 2

such schemes are shown for comparison, Ocean Data View (ODV) and SeaDataNet (SDN). It was agreed that for some applications and some users, including ICES Data Centre, the important part of the scheme is not the underlying number but the text description attached to it, however, a scalable numeric system has advantages for users; they can set their own boundaries for inclusion or exclusion of 'questionable' data and, in addition, perform whole-data set quality assessments. The particular problem facing an ICES quality flag scheme is that it must be applicable across data from a wide range of scientific disciplines.

WGDIM recommends considering the ICES text descriptions below as a basis for a quality flag system. The underlying numbering/key system can be implemented in whatever form is most appropriate provided that conversion routines are available for the most commonly used international schemes.

Quality flag systems can be applied at varying levels within a data set from each detailed data record through to flags that apply to the data set as a whole, the discussion and suggestion below relate specifically to detailed level data.

Objective quality control flags

ICES DESCRIPTION	ODV	SDN
Passed quality analysis/checks carried out by ICES	0	1
Did not pass quality analysis/checks carried out by ICES	8	4
Quality analysis not undertaken by ICES	1	0

Subjective quality control flags

DESCRIPTION	ODV	SDN
Passed quality analysis/checks carried out by ICES	0	1
Did not pass quality analysis/checks carried out by ICES	8	4
Quality analysis not undertaken by ICES	1	0
Result of quality analysis undertaken by ICES is questionable	4	2

It is *recommended* that a list of the analysis/checks that were carried out for each dataset is attached to the dataset at the time the flags were assigned.

WGDIM discussed quality flagging in ICES databases on the data level and recommends that assessment and science groups discuss, review and provide advice on the quality flagging descriptions as proposed by WGDIM

Objective quality control flags (Flags that are assigned based on automatic quality control tests).

ICES DESCRIPTION
Passed quality analysis/checks carried out by ICES
Did not pass quality analysis/checks carried out by ICES
Quality analysis not undertaken by ICES

Subjective quality control flags (Flags assigned by data managers or experts based on visual inspection of the data)

DESCRIPTION
Passed quality analysis/checks carried out by ICES
Did not pass quality analysis/checks carried out by ICES
Quality analysis not undertaken by ICES
Result of quality analysis undertaken by ICES is questionable

Apart from the objective and subjective ICES quality flagging, the contributor's quality control flags will be kept.

The science groups are specifically asked to provide advice on whether it is appropriate to flag their data using the quality control flagging system suggested.

The assessment groups are specifically asked to provide advice on whether the flagging scheme provides sufficient input to the advisory process to enable data inclusion/exclusion based on its quality.

There is no proposal at this stage to identify the objective tests required to assess quality, these should be agreed by the relevant science / advisory group once a flagging system has been agreed.

Interoperability

R Lowry (BODC) presented a history (since 1990) and overview of MarineXML, its problems, current issues and benefits. The group supported the work of the MarineXML Steering Group (presently R Lowry) and endorsed the need for it to continue and expand. However, this is not a role for 1 person and it was agreed that the WGDIM co-chairs and R Lowry would discuss potential routes to progress this.

V. Chavan gave a presentation on the Global Biodiversity Information Facility (GBIF). The mission of the Global Biodiversity Information Facility (GBIF) is to facilitate free and open access to biodiversity data worldwide via the Internet to underpin sustainable development. Priorities, with an emphasis on promoting participation and working through partners, include mobilising biodiversity data, developing protocols and standards to ensure scientific integrity and interoperability, building an informatics architecture to allow the interlinking of diverse data types from disparate sources, promoting capacity building and catalysing development of analytical tools for improved decision-making.

GBIF currently has 54 countries and 44 international organisations as participants.

Transparency

Discussion on many of the topics during the meeting had elements focussed on the need for a data user to be able to understand and document the provenance of the data used. There is also the need to be able to re-analyse a data set as it existed at a given point in time (e.g. when a stock assessment group met). These needs are often very important where advice is being given on the basis of analysis of raw data. There are arrangements in place within ICES to provide this facility, but, it is a mainly manual system with inherent problems reconciling the needs of users wanting to process current data with those wishing to use a fixed data set from some time ago. A method for applying a consistent approach to data audit trails and version control needs to be investigated with a view to implementation within the ICES data management systems.

5 ICES Data Strategy (ToR c)

The ICES Data Strategy is due for renewal during 2010. A WGDIM subgroup worked on a proposal for a replacement strategy before the meeting and presented a working paper to an enlarged sub-group during the meeting (Annex 7). The enlarged sub-group took the working document forward and WGDIM has proposed a new data strategy for adoption (Annex 6).

The data strategy proposed is based on analysis of the ICES science plan, the ICES Strategy documents and input from a wide range of scientists within the ICES community and representatives from the ICES Data Centre.

The strategy has 3 main themes; "Support for the advisory and science groups", "Leading best practice in data management" and "ICES as a data resource".

The Data Strategy should be a 'living' document and be subject to regular review along with updates on progress in each of the themes.

WGDIM recommends that the proposed Data Strategy is adopted and implemented.

6 VMS (ToR d)

In 2009 WGDIM was asked to provide advice on the developing need for and usage of VMS (Vessel Monitoring System) data.

Definition of VMS from www.fao.org:

A fishing vessel monitoring system (VMS) is a programme of fisheries surveillance, in which equipment that is installed on fishing vessels provides information about the vessels' position and activity. This is different from traditional monitoring methods, such as using surface and aerial patrols, on-board observers, logbooks or dockside interviews.

VMS programme is comprised of several components. Each participating vessels must carry a VMS unit. This shipboard electronic equipment is installed permanently on board a fishing vessel and assigned a unique identifier. Most shipboard VMS equipment types use satellite communications systems that have an integrated Global Positioning System (GPS). The system calculates the unit's position and sends a data report to shore side users. The standard data report includes the VMS unit's unique identifier, date, time and position in latitude and longitude.

WGDIM felt that it did not have a wide enough base of experience or skills to fully address the task and proposed a short term study group to take the task forward. The proposal was accepted and the group (SGVMS), chaired by Dr Heino Fock will have its first meeting in Hamburg during September 2010.

A presentation was given by P. Kunzlik (Scotland) giving a lay-man's summary of the current legislation applicable to those countries covered by the EU Data Collection Framework and highlighting the differing approaches being taken by individual countries regarding access to VMS data:

WGDIM circulated a series of four questions to ICES member countries within the northeast Atlantic area. For those that are also member states of the EU, the questions were framed in terms of the obligations of the EU data collection framework (DCF -

Council Regulation (EC) No 199/2008). For non-EU states and autonomous states, the questions were similar, but made no reference to EU regulations:

Questions to EU Member States (addressed to relevant countries' DCF National Correspondents)

- (i) Do you have a formal policy regarding the provision of VMS data to 'end users', including ICES, as defined under EC 199/2008? If so, can the terms of the policy be made available to ICES (WGDIM)?
- (ii) If you have no formal policy, can you confirm whether you would provide VMS data to ICES, if requested, with explicit reference to Chapter IV of EC 199/2008?
- (iii) In either case, do you consider the requirement to provide anonymised VMS data under EC 199/2008 either to over-ride or to be subordinate to other legislation governing, for example, data protection, freedom of information, human rights (e.g. regarding intrusive surveillance) and commercial confidentiality?
- (iv) Assuming VMS data can be shared by your country, do you have any established protocols for ensuring the anonymity of 'natural persons' or 'legal entities' (i.e. for processing data to the 'detailed' level as defined under Article 2(g) of EC 199/2008) and do have a standardised format for exchanging the data at the detailed and aggregated levels?

Questions to non-EU and autonomous states (addressed to an ICES delegate of from relevant countries)

- (i) Do you have a formal policy regarding the provision of VMS data to end users such as ICES? If so, can the terms of the policy be made available to ICES (WGDIM)?
- (ii) If you have no formal policy, can you confirm whether you would provide VMS data to ICES, if requested for specific purposes, and indicate any conditions under which they could be provided?
- (iii) In either case, is your capacity to provide VMS data subordinate to other legislation governing, for example, data protection, freedom of information, human rights (e.g. regarding intrusive surveillance) and commercial confidentiality?
- (iv) Assuming VMS data can be shared by your country, do you have any established protocols for ensuring the anonymity of individuals or individual legal entities and do have a standard format for exchanging the data?

The response rates were:

- EU Member States: 15 requests, 6 replies
- Non-EU states: 3 requests, 2 replies
- Autonomous states: 2 requests, 2 replies

Briefly, the responses indicate that:

- With only a couple of exceptions, there seem to be no 'formal' policies governing the provision of VMS data to ICES. Replies instead made reference to the need to follow "the legal requirements";
- It was commonly stated that access would have to be negotiated on a case by case basis;
- A lay-man's interpretation of the different responses is that data protection principles and regulations will create 'shades of grey' about access to VMS

data even where things look black and white to non-lawyers – even in the case of the DCF that obligates EU Member States to provide VMS data to appropriate end users. It will be different for the non-EU states, but it is not clear whether that means it would be a lighter or darker shade of grey.

The following bullet points quote selectively from the responses:

EU Member States:

- “data should be requested by EC” (ie not by ICES or other users)
- “only for data collected within 3 years of the data request”
- “access would have to be negotiated (with the fisheries department)”
- “will respond to a fully-reasoned formal enquiry”

Non-EU states and autonomous states:

- “only on formal application”
- “requires description of requestor’s data access/security policy”
- “subject to NEAFC and NAFO provisions on secure and confidential treatment of data”
- “would need to be discussed with the Ministry”

ICES is holding a VMS Study Group in September 2010 with the following terms of reference:

Provide expert advice regarding VMS data, with particular reference to:

- i) ICES strategic position regarding VMS data, the level of involvement required in the short, medium and long term;
- ii) Storage and management of the data;
- iii) Access to raw data and data products;
- iv) Data products;
- v) Tools and methods for analysis;
- vi) Quality assurance, quality control and quality flags.

Although the Study Group aims are solely directed to holding and processing the VMS data that may be supplied to ICES, and to provide a strategic overview of their use, the various responses to the questions, above, suggest that items (ii) & (iii) may raise legal issues that are outside the competence of scientists to deal with. Consequently, the SG may benefit from a broader-based membership to include, for example, representatives from national fisheries control and enforcement agencies.

WGDIM will review the first report of SGVMS and provide feedback.

7 Assess progress with, and update, user engagement plan (ToR e)

The user engagement plan was evaluated during the meeting, progress on some items was negligible, it was felt that the plan was over-ambitious with an excessive time-scale. A new user engagement strategy was written (Annex 8), this strategy has achievable goals and a 1 year lifetime. WGDIM will review and update the engagement plan on an annual basis.

Essential elements of the plan are that the users are in 3 distinct groups:

Data contributors: essentially as this group is already engaged and submitting data the emphasis should be on improving the ICES / Data Submitter interface and actively seeking new data streams to incorporate into the ICES Data Portfolio.

Internal ICES Users: There are 3 threads suggested for increasing user engagement within this group.

- An effective process for the requesting, development and release of new data products
- Improved traceability of data products and data audit trails
- Increasing the user group's awareness of the data sets, data products and systems available from ICES

Wider marine and maritime research communities: This is a much more difficult group to engage as they have a vast range of needs and scientific insight. The initial steps should be to increase awareness of the data portfolio and from there increase its access and exploration.

A theme that runs through a lot of the discussion was 'raising awareness', it has often been noted that scientists are focussed on their area of specialism and do not realise the breadth of the data holdings available to them.

The WGDIM proposal for a 2010 theme session was unfortunately unsuccessful, the group agreed to review the proposal and resubmit for ASC 2011.

8 DATRAS (ToR e)

At WGDIM 2009, in response to requests from users, a Datras User Advisory Panel (DUAP) was created to provide a central discussion and contact point for users of the DATRAS system. DUAP uses SharePoint and forum facilities provided by ICES, the group is co-ordinated and moderated by I de Boois (Netherlands).

DUAP is proving to be a useful forum for discussion regarding issues/development of the DATRAS system and as awareness of the group expands it will become an integral part of the DATRAS project. A report was presented regarding activities in the forum (Annex 10).

9 Other Items

A request was received from WKROUND (via ICES secretariat, not directly aimed at WGDIM), reproduced below

Linkage of assessments to environmental and ecosystem conditions

The ToR asked that the panel to "consider the possible inclusion of environmental drivers for stock dynamics in the assessments and outlook." Some of the stock annexes describe changes that could be the result of environmental or ecosystem drivers. These include the changes over time in the body weight-at-age of saithe and changes in the spatial distribution of recruitment to northern hake. In some cases there was speculation about possible causes of these observed changes. Several of the newer assessment approaches are amenable to inclusion of environmental time-series as data on temporal changes in important model factors, such as natural mortality, catchability, and growth. However, none of the assessments proposed linking such ob-

served stock changes to environmental drivers in order to improve the accuracy or precision of the assessment. The panel briefly considered possible steps to improve the degree to which environmental factors are directly considered in the assessments. One step is to commission specific work to elucidate the factors and develop relevant time-series. Another is to conduct management strategy evaluations to determine the needed precision of such relationships in order to realize improvements in the assessment.

After some discussion both inside and outside the meeting the issue was narrowed down to the following generic description that is applicable across many groups/disciplines:

- a) An EG saying "we need some data on something but do not know where to get it"
- b) A survey/data groups saying "we've got all sorts of data, let us know what it is you need exactly".

The group considered this to be another example of a lack of awareness of the ICES data portfolio and more specifically EcoSystemData.

WGDIM recommends:

ICES Secretariat staff receive training in the use and promotion of EcoSystemData and the ICES data portfolio and are encouraged to actively promote these to the expert and advisory groups.

A small (A4 3-fold) leaflet is produced highlighting EcoSystemData and the ICES data portfolio, this leaflet to be despatched to each expert and advisory group chair for distribution at their meetings.

A protocol is developed for users (initially, internal ICES users i.e. expert and advisory groups) to request development of new data products. A draft of this protocol was developed by WGDIM (Annex 11). This protocol should be further developed inter-sessionally between WGDIM, the Data Centre, advisory and science groups.

Annex 1: List of participants

NAME	EMAIL	STATUS	REPRESENTING
Daphne R. Johnson	Daphne.Johnson@noaa.gov	Member	US
Garry J. Dawson	garry.dawson@ukho.gov.uk	Member	GB
Helge Sagen	helge.sagen@imr.no	Chair	NO
Henrik Degel	hd@aqua.dtu.dk	Member	DK
Jan Szaron	jan.szaron@smhi.se	Member	SE
Jens Heilmann	jh@aqua.dtu.dk	Member	DK
Lena Szymanek	lena@mir.gdynia.pl	Member	PL
Marcin Wichorowski	wichor@iopan.gda.pl	Member	PL
Mia Devolder	M.Devolder@mumm.ac.be	Member	BE
Neil Holdsworth	neilh@ices.dk	ICES Secretariat	ICES
Nils Håkansson	nils.hakansson@fiskeriverket.se	Member	SE
Peter Wiebe	pwiebe@whoi.edu	Member	US
Phil Kunzlick	p.kunzlick@marlab.ac.uk	Member	GB
Richard Ayers	r.a.ayers@cefas.co.uk	Chair	GB
Riitta Olsonen	riitta.olsonen@ymparisto.fi	Member	FI
Taco de Bruin	bruin@nioz.nl	Member	NL
Fabrizio Manco	fabrizio.manco@cefas.co.uk	Chair- invited	GB
Vishwas Chavan	vchavan@gbif.org	Chair- invited	GBIF
Hannah R. Freeman	haee@bodc.ac.uk	Chair- invited	GB
Sjur R. lid	sjur.ringheim.lid@imr.no	Chair- invited	NO/IMR
Pekka Alenius	pekka.alenius@fmi.fi	Member	FI
Friedrich W. Nast	friedrich.nast@bsh.de	Member	DE
Christopher Zimmermann	christopher.zimmermann@vti.bund.de	Member	DE
Ingeborg de Boois	ingeborg.deboois@wur.nl	Member	NL

Annex 2: Agenda

Agenda for WGDIM 2010 in Copenhagen (Denmark) 25–27 May.

Tuesday 25 May – Rapporteur: G Dawson

0900 – 0930	Opening greetings and aims for the day Welcome by ICES representative	[Co-chairs]
	Local arrangements	[V. Piil]
0930 - 1030	Review meeting schedule and items for discussion Round table introductions Appoint rapporteurs for meeting	
	Elections – a note to think on	[H. Sagen]
	Review actions from last year’s WGDIM meeting	[R. Ayers]
1030 -1100	Coffee break	
1100 - 1115	WGDIM’s changing role – Presentation to SCICOM at ASC	[R Ayers]
1115 – 1200	Marine XML (collaborative working to produce oceanographic domain profiling of the emerging usage metadata standards under the ICES/IODE ‘MarineXML’ banner with the objective of creating a set of standards for submission to the IODE/JCOMM standards process)	[R Lowry]
1200 -1300	Data Centre Update, Status, Upcoming Activities Stomach data latest update Highlights from Working group link-ups Datras User Advisory Panel Update DATRAS – Update MODEG	[N. Holdsworth] [Carlos] [I Deboois] [N Holdsworth]
1300 - 1430	Lunch	
1430 - 1500	GBIF – An introduction, Key points and Highlights	[V Chavan]
1515 – 1535	Update on Oceanographic data system in the ICES Data Centre	[Data Centre]
1535 – 1600	GIS developments at ICES	[Data Centre]
1600 - 1630	Coffee break	
1630 - 1700	Data Accessibility & Data Portal, status, developments and future plans	[C Zimmerman]

1700 - 1730 Egg / Larval database
[N. Holdsworth]/[C Zimmerman]

1730 - 1745 Summary of Day 1 [Co-chairs]

Wednesday 26 May – Rapporteur: S Lid + Sub Groups

0900-0915 Opening notes, aims for the day [Co-chairs]

0915 – 0930 VMS [P Kunzlik]

0930 - 1000 Define deliverables, chairs and rapporteurs for sub-group work
(General initial discussions and allocation to groups)

1000 - 1530 Sub group work :
Groups to break at appropriate points, decided by each group
Group 1 – Data Strategy – North Sea Room
Group 2 – Quality Flags- Atlantic Room

1545 – 1615 Coffee

1615 – 1700 Plenary Report and discussion Data Strategy Group

1530-1540 Group photo

1700 - 1745 Plenary Report and discussion Quality Flags Group

1745 - 1800 Summary of Day 2 [Co-chairs]

Thursday 27 February – Rapporteur: R Ayers

0900-0915 Opening notes, aims for the day [Co-chairs]
Election of Co-Chairs

0915 – 1000 User Engagement plan – Review, update and develop actions
for coming year.
Theme session 2011
Updated user engagement plan

1000 – 1015 Request from WGNSSK

1015 – 1045 Coffee

1045 – 1115 Actions and Recommendations for 2010/2011

1115 – 1200 ToRs 2011 (suggestions below)
SGVMS
Data Strategy
Data flags

Engagement plan

Data policy

Marine spatial service

Egg/larval database

1200- 1215	Report writing and agreement	[Co-chairs]
1215 – 1230	Next meeting	
1230 - 1245	Closure	[H Sagen]

Annex 3: ToRs 2009/2010

- a) **Data accessibility** – Provide advice on the functionality of the new integrated ICES data portal throughout the year. Deliverable: Review of data portal, recommendations for enhancements. (**Responsible focal point:** Chris Zimmerman)
- b) **Quality, interoperability and transparency** - Identify and resolve issues related to the use of quality flags in ICES Data Management, specifically in the areas of Biological and Chemical data, with particular reference to existing international quality flag systems or those that are envisioned as needed (e.g. SeaDataNet, EMODNET, HELCOM and OSPAR etc...) Deliverable: Guideline for development/adoption of a quality flag system suitable for application across ICES data holdings, including actions to harmonise across existing systems. (**Responsible focal point:** Gaynor Evans)
- c) **ICES Data Strategy** – Draft the 2011–2015 Data Strategy in line with ICES Strategic Plan 2009 – 2013. Identify emerging technologies that ICES Data Centre, WGDIM members and the wider ICES Community should be aware of (GIS metadata). Deliverable: Draft data strategy document, Report on emerging technologies and possible applications with ICES. (**Responsible focal point:** Richard Ayers / Helge Sagen)
- d) **VMS** – Review actions resulting from recommendations of WGDIM 2009, review progress of DG-MARE/2008/10 Lot 2 project and other developments in the VMS arena. Deliverable: Report on actions, Lot 2 progress, proposals for further work; (**Responsible focal point:** Phil Kunzlik)
- e) **Assess progress with, and update, user engagement plan** – Review the success of the ICES Data Centre Live at the ICES ASC, undertake planning of agreed workshops (e.g. mackerel egg) and propose new workshops or activities. Deliverable: Outcomes from ASC 2009, updated user plan with follow up actions, detailed work plan for workshops, further proposals. (**Responsible focal point:** Pekka Alenius)
- f) **DATRAS** – Provide feedback, guidance and advice on the ICES DATRAS system. Specifically to include liaison with data submitters and data consumers. This work will be completed inter-sessionally with progress reported at WGDIM annual meeting.

WGDIM will report by 1 July 2010 for the attention of SCICOM and ACOM.

Annex 4: WGDIM proposed terms of reference for the next meeting

The **Working Group on Data and Information Management** (Co-Chairs: H Sagen, Norway and I de Boois, Netherlands) will meet in Copenhagen, Denmark from 24–26 May 2011 to:

- a) **Review outputs/products of offspring groups:**
 - SGVMS – Study Group on VMS data
 - DUAP - DATRAS User Advisory Panel
- b) **Review ICES Strategic documents and recommend updates/amendments**
 - Data Policy
 - User Engagement Plan
 - Data Strategy
- c) **Data Quality Flags;** Review development of the data-point level quality flags recommended at WGDIM 2010. Develop a recommendation regarding data-set level quality flags
- d) **Support ICES Data Centre with feedback and advice on existing products, current developments and potential new products.**
 - Develop recommendation for implementation of data and data product version control
 - DATRAS
 - Regional databases
 - EcoSystemData
 - Egg / Larval database
 - Develop recommendation for implementation of a system for numbering/control of documents published by ICES that are not covered under existing arrangements (e.g. survey manuals)
- e) **Review progress on recommendations from WGDIM 2010**

WGDIM will report by 1 July for the attention of SCICOM and ACOM.

Supporting Information

Priority	The Group provides ICES with solicited and unsolicited advice on all aspects of data management including technical, data policy and data strategy and user oriented guidance. This Group flies the flag for ICES in setting standards for global databases. It also provides an important interface for oceanographic, environmental, and fisheries data management in ICES, and promotes good data management practice.
Scientific justification	<p>a) There are 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 are 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. The integrated ICES data portal will be reviewed with suggestions for improvements and enhancements.</p> <p>b) More effort will be put into making it possible to track the data used to make an assessments. If the external data are being used to formulate advice, it is often difficult to later re-establish the data sets and thus the basis for the advice. Thus the group should provide advice as to how improve this reporting. To maximize interoperability data quality must be known. It is important to evaluate the appropriateness of use of data for specific applications on the basis on data quality. Coordinate work with relevant working groups or projects like SeaDataNet, ECOOP, etc. on standards for metadata, data/data structures and vocabularies. As there is limited resource available it is essential to avoid duplication of work on data management. It is thus important to engage in collaboration with international bodies and programmes especially to ensure the inheritance after the 4th International Polar year.</p> <p>c) It is essential to ensure needs of users are met: there are a wide range of users for ICES data and products from HELCOM/OSPAR to WGs/SGs and individual scientists. WGDIM should have a key role to act as a mediator between Users/WGs and the Data Centre to prioritise activities, to ensure appropriate experts are available and to give reasons for priorities. Proper engagement with users will allow data submission problems to be resolved and integrated data products and thus advice can be provided in an appropriate form.</p>
Resource requiremen	None
Participants	The Group is expected to be attended by some 20–30 members and guests with half of the members from each of the two categories , data managers and data users
Secretariat facilities	Meeting facilities.
Financial	The Head of Data Centre should attend these meetings together with other employees at the Data Centre.
Linkages to advisory committees	ACOM
Linkages to other committees or groups	
Linkages to other organizations	There are linkages with relevant international bodies and programmes like PICES, IOC/IODE, GOOS, SeaDatanet, IPY, etc., with emphasis on IOC and its Working Committee on International Oceanographic Data and Information Exchange (IODE).

Annex 5: Recommendations

RECOMMENDATION	FOR FOLLOW UP BY:
The Egg/larval database action plan is completed in conjunction with the Data Centre and expert groups and then taken forward	Data Centre WGMEGS, WGEGGS,PGIPS,MIK-IBTSWG WGDIM
Provide feedback to WGDIM and the Data Centre on the North Sea element of the Year of the Stomach Database	WGSAM
Consider the Baltic dataset (in its current state); whether there is enough value in undertaking further work and if so what resources are needed (i.e. volunteer experts from the Baltic States along with IT experts from the Data Centre).	WGSAM
Consider the proposed ICES data quality flag system, in relation to data extracted from ICES' data systems by the scientific /advisory group, advise on appropriateness and usefulness of the flagging scheme along with recommendations for amendment if necessary	DATRAS related expert groups WGBIFS IBTSWG, WGBEAM WGMEGS WGEGGS Assessment groups relying on Data Centre output WGNSSK, HAWG Science/Advisory groups using ecosystem data: WGSAM, WGOOFE MCWG WGMS WGBEC WGPDMO WGZE WGPME BEWG WGIAB
The proposed Data Strategy is adopted and implemented	ACOM / SCICOM
ICES Secretariat staff receive training in the use and promotion of EcoSystemData and the ICES data portfolio and are encouraged to actively promote these to the expert and advisory groups.	Data Centre
A small (A4 3-fold) leaflet is produced highlighting EcoSystemData and the ICES data portfolio, this leaflet to be despatched to each expert and advisory group chair for distribution at their meetings.	Data Centre / PubCom
A protocol is developed for users (initially, internal ICES users i.e. expert and advisory groups) to request development of new data products. A draft of this protocol was developed by WGDIM (Annex 11). This protocol should be further developed intersessionally between WGDIM, the data centre, advisory and expert groups	Data Centre

Annex 6: Proposed ICES Data Strategy

ICES will be a leader in marine data and information management, providing best practices, data mobilisation and services for its advisory and science groups and the wider marine and maritime research communities

A service for ICES advisory and science groups

- ICES Data Centre will provide the advisory and science groups with data and products tailored for their needs
- ICES will ensure that data services are closely linked to the advisory and science group products
- ICES data will form an integral part of new services envisaged under the SCICOM and ACOM leadership, such as marine spatial service, climate change, biodiversity and information

A leader for best practice in the management of marine data

- ICES Data Centre will be recognised as a regional centre of excellence in the stewardship and mobilisation of marine data and information
- ICES will implement the international standards for marine data and information
- ICES will continue to contribute to the development and adoption of new standards for meta-data and spatial data through regional and international cooperation
- ICES will endeavour to embed these standards in the wider marine network through training, promotion, dissemination and online materials

A regional resource and marine data and information node

- ICES data will be utilised and made available to a greater number of potential end users through project participation and new co-operations with emerging infrastructures
- ICES will be seen as a resource for the marine and maritime research communities
- ICES will have tools that enable the majority of users to access all the data online
- ICES will be a trusted source of marine data and information of a known quality
- ICES will continue to act as a steward for marine data sets that could otherwise be lost to the marine community
- ICES will continuously review, update and apply the ICES Data Policy
- ICES Data Centre will continue to identify new or underutilised streams of data that can be incorporated into the ICES data portfolio and encourage the network of ICES scientists to contribute to this process

Annex 7: WGDIM Data Strategy Working Document

WGDIM Working Document

ICES Data Strategy 2010 onwards

Background

The data collated and managed by ICES provides the building blocks upon which the work to bring the ICES Science Plan and ICES Strategy to fruition is based. The activities to manage, support, extend and disseminate the data are driven by the ICES Data Strategy.

The Data Strategy is scheduled for revision in 2010, as part of its strategic remit WGDIM, in conjunction with the data centre, has started a review of the ICES data strategy with a view to providing significant input to the updated strategy.

The Science Plan and Strategy were dissected and data-centric activities identified; these activities lead naturally to what must be the core values of the ICES Data Strategy.

Relevant elements from the science plan;

Overarching themes;

- Ecosystem approach to Management (joining the data up as well as the advice)
- Socio-economics

Scope of ICES science activities

The success of ICES science in future depends on strengthening the links between environmental science, physical and biological oceanography, fishery science, and socio-economic sciences, and in developing integrated programmes. At the same time it is important to ensure that the individual disciplines are able to advance and flourish. The Science Plan will facilitate this endeavour, thus ensuring that fishery and environmental sciences remain strong, dynamic, and adaptive to the research and advisory needs of ICES.

Fish life-history information in support of EAM

Challenges are twofold: (i) monitor the status of populations and ecosystems with indicators and appropriate statistical techniques, and (ii) achieve predictability of population distributions, connectivities, and recruitment levels using complex fish-ecosystem coupled models and simulations.

To achieve these objectives, information on population life history and life cycles, as well as the relative influence of environment and genetics on these traits need to be integrated and coupled with information regarding environmental and ecosystem conditions. More precisely, life cycle spatial organization of fish populations need to be documented and modelled using cross mapping and their coupling to habitat requirements and vulnerability to anthropogenic activities.

Habitat characterization is becoming increasingly important. It will also be necessary to characterize biological functions of growth, reproduction, and feeding that rely on the quality of habitats. Understanding of fish physiology, behaviour, and their genetic basis is essential to coupling at such small scales. Also, at a regional scale, connectivity of larval transport between populations, fish behaviour, and movements need be understood and modelled. A spatial setting which incorporates operational oceanographic products would provide essential structure to all these studies.

The role of coastal-zone habitat in population dynamics of commercially exploited Species

Sustaining ecosystem goods and services, while meeting growing societal needs, requires ecosystem-based marine spatial planning. This topic will focus on processes linking habitat to spatial patterns at the population and community levels.

Role of top predators (marine mammals, seabirds, and large pelagics) in marine ecosystems

Top predators may have an important role in the functioning of marine ecosystems (e.g. in “top-down” controlled systems). In recent decades, there has been both a systematic removal of larger fish and in some areas increases in marine mammals and seabirds. This topic will include comparative analyses of ecosystem dynamics in response to changes in abundance and relative composition of top predators.

Sensitive ecosystems (deep-sea coral, seamounts, Arctic) as well as rare and data-poor Species

Sensitive habitats or ecosystems need to be identified and mapped as a basis for their conservation and management. This can include further development of habitat classification systems. There are likely to be a large number of new species that are as yet unknown to science in these special environments, especially in relation to rarer structures such as hydrothermal vents and cold seeps.

Integration of surveys and observational technologies into operational ecosystem surveys will be an essential part of any attempt to develop our knowledge of these individual ecosystems. An ecosystem monitoring programme will be developed for the ICES area based on existing time-series activities of member countries, emerging survey methodologies, and enhanced coordination (plankton, acoustics, optics, trawling); with the aim of providing indicators in support of advisory needs of integrated management and ecosystem status reporting. Elements will include remote observations (satellite and aircraft observations), observations from buoys, gliders, moorings, and tracking of biota.

Coordination of deep-sea monitoring will be addressed, as well as optimizing the use of vessels involved in “ICES surveys”. A permanent network of ICES fixed stations

using similar protocols will be proposed. Protocols for diverse components of the programme will be developed and data sharing and availability enhanced. Population and community level impacts of contaminants, eutrophication, and habitat changes in the coastal zone.

A synthesis of knowledge will be undertaken on the impacts of diverse land-based and marine activities, with the aim of characterizing the status of regional coastal-zone ecosystems and causal relationships. The rich datasets for the coastal zone that are available to the ICES community suggest that advances in ecological understanding will emerge from such a synthesis, as well as the identification of gaps in knowledge and monitoring needs.

Operational modelling combining oceanography, ecosystem, and population processes

There is a need to facilitate the availability and dissemination of long-term, high quality data required to advance the scientific understanding of the North Atlantic. It is therefore important to provide analysis, forecasts, and model-based products describing the marine conditions and to give a reliable description of the actual marine conditions including physical and ecosystem variables.

Enhanced research coordination in the North Atlantic

Existing ICES attributes and infrastructure which particularly underpin ICES leadership potential in this respect include: • Extensive oceanographic and marine biological databases; • Websites to serve as a public interface with constituents and internally for project coordination and data exchange.

Relevant elements from the strategic plan

ICES recognizes that there will be an increasing demand for greater involvement by, and transparency to, interested groups. In recent years, the precarious state of an increasing number of fish stocks and ecosystems means that the scientific information and advice ICES provides also needs to be more accessible to groups with direct and indirect interests and to the broader public.

The ICES mission includes: “Developing and maintaining accessible marine databases”.

ICES must continue to add value to scientific efforts by being a leader in mobilizing scientific resources to collect and manage data, conduct experiments, perform analyses, build models, and disseminate information.

Goal 4: Develop a comprehensive strategy to manage and disseminate marine data for the ICES Area in support of the Science and Advisory programmes. Currently, ICES has two pillars of “science” and “advice”. There is a need to consider a third “data” pillar. Marine policy is looking increasingly to performance measures and indicators for marine management, and data are an essential element of this need. ICES must develop a strategy for managing its data, and whether it should become a regional data centre and how it will be resourced. Outcome/deliverable = A documented plan accepted by customers and stakeholders

Proposal for the new data strategy

ICES will be a leader in marine data and information management, providing best practices, data mobilisation and services for its advisory and science groups and the wider marine community.

As a result of the above mission statement, the onward strategy should develop around three focal points:

i) A service for ICES advisory and science groups

- ICES Data Centre will provide the advisory and science groups with data tailored for their needs
- ICES will ensure that data services are closely linked to the advisory and science group products
- ICES Data will form an integral part of new services envisaged under the SCICOM and ACOM leadership, such as marine spatial service and information feeds of advisory content

ii) A leader for best practice in the management of marine data

- ICES Data Centre will be recognised as a regional centre of excellence in the stewardship and mobilisation of marine data and information
- ICES will implement the international standards for marine data and information
- ICES will contribute to the development and adoption of new standards for meta-data and spatial data through regional and international cooperation
- ICES will endeavour to embed these standards in the wider marine network through training, promotion, dissemination and online materials

iii) A regional resource and marine data and information node

- ICES Data will be utilised and made available to a greater number of potential end users through project participation and new co-operations with emerging infrastructures
- ICES will be seen as a resource for the marine community
- ICES will have tools that enable the majority of users to access all the data online
- ICES will be a trusted source of marine data and information of a known quality
- ICES will continue to act as a steward for marine data sets that could otherwise be lost to the marine community
- ICES will continuously review, update and apply the ICES Data Policy
- ICES Data Centre will continue to identify new streams of data that can be incorporated into the ICES Data portfolio and encourage the network of ICES scientists to contribute to this process

Annex 8: User Engagement Strategy

AIM: ICES services are tailored to the needs of the users with resulting user engagement			
<p>Data contributors</p> <ul style="list-style-type: none"> • Current data contributors are engaged and easily identifiable: for this group streamlining communication is main point of attention • Potential data contributors (new streams of data): make clear ICES is open to new data sets: <ul style="list-style-type: none"> ○ give publicity of the ICES data strategy ○ procedure on accepting and incorporating new data sets (dealing with requests) • Investigate data-tracking (e.g. which data are published where and by whom?), automatic or by a user feedback mechanisms 	<p>Internal ICES users (advisory and science groups)</p> <ul style="list-style-type: none"> • Procedure on how to handle a request for new standard output from data available at ICES (see below-to decide on by WGDIM/Data centre) • Possibility for traceability data used for standard output (see data strategy box 1 bullet 2) • 'you downloaded this standard output, you might also be interested in....' • Making advisory and science groups aware of new data developments by advertising leaflets • Promoting and training in the ICES Data Centre products to the ICES Secretariat 	<p>Use by the wider marine and maritime research communities</p> <ul style="list-style-type: none"> • Increase data visibility <ul style="list-style-type: none"> ○ possibility to explore the portfolio ○ clear website ○ data-portal • 'you downloaded this data-set, you might also be interested in....' • Continue 'Data Centre Live' at ASC <p>Longer term:</p> <ul style="list-style-type: none"> • Develop standard products • Guidelines 	
<p>-Delivering products and services requested by users</p> <p>-Putting an ICES 'brand' on the products increasing awareness of the source</p>			

Annex 9: Egg/Larval Database Action Plan

Adding more biological data to ICES: eggs and larvae

27 May 2010

Neil Holdsworth, Ingeborg de Boois, Mark Dickey-Collas, Christopher Zimmermann

Background to action

Fish egg and fish larvae data have been collected in the ICES community for a long time for use in stock assessments. The collection of the data is usually organised by international survey expert groups. The data are stored in the national institutes and in most cases, an expert group keeps a copy of the combined data to be able to supply the assessment working groups with the information needed. In this way, the information is available for the main purpose, but it is not accessible to the wider marine community.

Currently, there are a number of data sets which are not readily available to the wider marine community but coordinated by ICES groups and used in ICES stock assessments. These are:

- mackerel and horse mackerel eggs Northeast Atlantic (WGMEGS);
- cod and plaice eggs North Sea (WGEGBS);
- herring larvae North Sea, herring larvae Western Baltic (both PGIPS);
- herring larvae North Sea (MIK-IBTSWG).

WGDIM 2010 decision on action

Since these data sets are very similarly structured, and already worked up and quality controlled, they are considered a good starting point for the development of a ichthyoplankton database structure in ICES. It was therefore decided at WGDIM 2010 to create an action plan to collate ichthyoplankton data at ICES.

By developing such a database structure, the national data will be secured at an international data clearing house, it will be possible to provide an overview of available fish egg and larvae survey data collected and to provide a unified portal for scientific access to the ichthyoplankton survey data.

	EGG DATA		LARVAE DATA		
	Mackerel/horse mackerel	Cod/plaice	Herring (RHLS) Western Baltic	Herring (IHLS) North Sea	Herring (MIK)
Data stored at national level	X	X	X	X	X
Data stored in combined data set	X	X	X	X	X
Contact person for combined data set	Chair WGMEGS and Doug Beare	Chair WGEGBS	vTI-OSF Rostock (Christopher Zimmermann)	vTI-SF Hamburg (Joachim Groeger)	DTU Aqua (Peter Munk)

The table above gives an overview of the targeted egg and larvae data, the stock it is related to and which ICES group or institute is responsible for the dataset.

Action plan

Priority of incorporating data sets

Since the five data sets will overlap to some extent in the type of data collected, it is recommended to start with one of the sets to create the main structure. The database structure will need to be extensible, in the sense that the addition of variables will not affect the overall structure or new values for existing variables.

The order of adding data to ICES Data Centre, will be as follows:

A. Data used for ICES assessments:

1. Cod and plaice eggs
2. Larvae as collected in the MIK survey (IBTS Q1)
3. Mackerel and horse mackerel eggs
4. Larvae as collected in the herring larvae surveys (Western Baltic and North Sea) and Norwegian spring spawning herring larvae (with Norwegian approval)

B. Other ichthyoplankton surveys in the ICES area

5. Canadian herring larvae
6. Eastern Baltic herring larvae

Proposed Time schedule

DEADLINE	TASK	PRODUCT
2010		
Q2	Create action plan	
Q3	Contact chairs of groups under (A) to inform about the action plan, including a data request	
Q4	Work session during WGEGBS (9-12 November)	Database structure for WGEGBS data
2011		
Q1	Further phases to be planned as the 2010 plan develops	
Q2		
Q3		
Q4		

Annex 10: DUAP Report

Establishment of a DATRAS User Group

Introduction

IBTSWG and WGBEAM recommended the establishment of a DATRAS User Group to evaluate the functionality of the DATRAS database, to provide feedback by data submitters and data users, to suggest updates of the system where needed, and to prioritize future developments.

In October 2009, the Dattras User Advisory Panel (DUAP) was established as a group under WGDIM. Main task for DUAP is to provide feedback, guidance and advice on the ICES DATRAS system, specifically to include liaison with data submitters and data consumers.

Membership and coordination

Membership of the group is open for all DATRAS users (upload and download, all surveys present or planned to be in DATRAS). The group members discuss via <http://groupnet.ices.dk/duap/default.aspx>. DUAP is coordinated by Ingeborg de Boois, Netherlands (ingeborg.deboois@wur.nl). The work of DUAP will be completed inter-sessionally with progress reported at WGDIM annual meeting. The coordinator reports to WGDIM.

Statistics 2010

The table below lists the topics by owner, including the number of replies and the date of last update. In the following section the status of the discussions is described.

Statistics for discussions on the DUAP SharePoint (status on 25/05/2010).

DISCUSSION NUMBER	CREATED BY	SUBJECT	REPLIES	LAST UPDATED
1	Ingeborg de Boois	Automatic upload: different length codes per haul and species	1	12/04/2010 10:36
2	Brian Harley	Code to calculate indices	0	31/03/2010 12:15
3	Matthew Parker-Humphreys	Replicating Dattras Q1 North Sea indices	3	12/04/2010 09:04
4	Ingeborg de Boois	CA records without age information	2	31/03/2010 16:02
5	Ingeborg de Boois	Measuring length: which length did we measure?	0	23/03/2010 11:53

DISCUSSION NUMBER	CREATED BY	SUBJECT	REPLIES	LAST UPDATED
6	Ingeborg de Boois	Maturity scales: old and new	3	23/03/2010 11:28
7	Rainer Oeberst	BITS, data required for stock indices and additional stock information	1	09/03/2010 14:18
8	Finlay Burns, posted by Ingeborg de Boois	IBTS data query: CA Records on IBTS data files	3	22/01/2010 13:12

Topics discussed and progress

The topics discussed are listed in the paragraphs below, as is the status.

Automatic upload: different length codes per haul and species

Subject: automatic upload: different length codes per haul and species

Hi everyone,

I'm trying to do an automatic upload for NS IBTS prior to 2004. Generally, it looks fine. However, I start having problems when trying to upload fish measured by 1 cm class and fish measured by 5 cm class in the same haul. I agree that it is not recommended to upload different length codes for the same species in the same haul, but I would propose to make an exception for this since prior to 2004 most countries used the 1 cm class up to 59 cm and a 5 cm class from 60 cm onwards.

I'm sure the old uploading system was able to deal with it.

If anyone has objections to or more support for this proposal, let me know!

Cheers, Ingeborg

If you don't have the larger fish in 1cm length groups, the system is going to have to let you upload in 5cm groups (although I assume it's not straightforward to do so). Is this for all species?

Brian

Status: Proposal to be followed up by ICES Data Centre

Code to calculate indices

Hi

Does anyone have the code (not just the formula as we have that), that DATRAS uses to calculate the various indices, in particular the one they use for the North Sea? Ce-

fas are trying to replicate the indices but are having a few issues. Any help would be great

Cheers, Brian

Status: pending (see 1.4.3)

Replicating Dattras Q1 North Sea indices

Hi Everyone

I have just been given access to this Sharepoint site as a little while ago I attempted to replicate the IBTSQ1 Dattras North Sea indices for a project I was working on, but had a couple of small differences. I have heard on the grapevine that others are attempting to do the same, so I thought it would be a good time to say hello, and see if together we can get to the bottom of this.

I am creating my indices in R and I am using the calculations described in the 'ICES Dattras Report 2006'. I am starting with the "CPUE per length per haul" and "SMALK" data, and then raising these to get the number of ages at each haul. I then create my indices from these data, and I am pretty sure my "numbers at age per haul" match those available from Dattras.

I have attached an Excel spreadsheet, which shows how close I was to matching the North Sea Q1 Cod Dattras indices. The spreadsheet isn't well annotated, but should be reasonably obvious. What is confusing me most is that I match most of the time, but in some years there are differences. I don't know why this is, but having written this in R the code for each year is the same (in a loop)???

Does anyone have any ideas which could explain the differences? I had very much hoped to get hold of the Dattras code that is used to create the indices and therefore see where any difference are, but I have asked a couple of times, and not got the code, so I guess this isn't possible?!?!?

I do know that there are certain "features" that are not explained in any text I have found... i.e. in one year there is no ageing info in Round Fish Area 8, so Dattras uses data from Round Fish Area 7. I also follow this process in my code, but I suspect there may be more fixes like this involved that we do not know about..... Without seeing the code!

Anyway this is where I am at. Has anyone got any further.... Or more to the point does anyone want to???

Thanks for looking.

Matt Parker-Humphreys (Cefas)

Dear Matt,

I think that uploading your spread sheet to the background and working documents section of the sharepoint may work and be easier than trying to attach it to the discussion board. To me it looks obvious that the small "features" pointed out by you, should be documented and clear at least to the assessment WGs and the users of the indices, otherwise all the quality control and data checking becomes really difficult.

Francisco Velasco

Hi all of you,

I tried to reproduce the indices as calculated in cpue per age per area based on cpue per age per haul and discovered night hauls were included. I'll put my SAS code (sorry Matt, did not shift to R yet) in the working docs as the input and the output.

So the small differences might be due to (implicit?) selection criteria not described in the report you use for the index calculation -as you already discovered for area 7 and 8 issues.

Cheers

Ingeborg

OK cannot find a way of actually attaching my Excel spreadsheet (keeps failing).

Will keep trying and hopefully it will appear soon!

Matt

Status: pending.

CA records without age information

Dear all,

During some surveys biological data (length, weight, maturity) without age information are collected. This information can not be uploaded to DATRAS at the moment. And what's not in, cannot come out....

Discussion:

Would you like to be able to upload/download biological data without age information?

If yes, why? If no, why not?

Best wishes, Ingeborg

Dear All,

Unknown age as usable in BITS which is marked by space or -9. These dataset are marked with warning during the screening procedures but the data can be uploaded.

Such type of data makes sense because they can be used for special analyses which do not include the age.

Rainer

I agree, the data without age is still important in the length weight relationship for example which in turn is required for the final stock weight calculations and so on. This is required to convert numbers at age to weight at age as commercial catches are given and managed in weight not number.

Dave

Status: to be followed up by the uploading institutes (missing yearclass=-9)

Measuring length: which length did we measure?

Hi all,

For all not participating in IBTSWG these days, yesterday we discussed the following topic:

The type of length measurement is not added in DATRAS. In most cases, for fish total length is measured. However, sometimes other length measurement types are used, mainly for deepwater fish species. A complete list for the measurement types for deepwater species can be found in the PGNEACS 2008 report, Table 2. (ICES, 2008b).

Additionally, there is a wish for measurement types for Cephalopods (mantle length, mantle width) and Crustaceans (carapace length, carapace width).

It is crucial to know which measurement method is used when creating length frequency diagrams. Length measurement information should be available on a record level in all HL and CA records and is additional to the class increment.

Input on more types of measurement is welcome, as is the need for adding this variable to DATRAS.

Cheers, Ingeborg

Status: Proposal to be followed up by ICES Data Centre (see also PGNEACS report 2008, IBTSWG report 2010)

Maturity scales: old and new

Dear all,

There are continuously developments in maturity stages for fish. In 2008, a new common scale for gadoids was proposed (and adopted?), two weeks ago, the same was done for sole, plaice, dab and flounder. However, even if institutes adopt a new common scale as a 'reporting' scale, they might still use their own scale for staging. Would it be interesting to flag all maturity records with two flags: (a) the reported maturity scale and (b) the scale used for staging?

And to the DATRAS team: would that be possible?

Cheers, Ingeborg

Hi Ingeborg,

It is possible by adding one extra field of maturity stage, but again it is a question of changing reporting format for CA record. These need to discuss in WG meeting.

Cheers, Vaishav

Hi,

For the BITS we prepared tables in the manual which describe the relation between the national codes and the agreed ICES code for cod, flatfish, herring and sprat. There is the agreement that in the DATRAS database only the international code is used.

That means that the national code in the national database will be transferred to the ICES code during the preparation of the data in the DATRAS exchange format. This seems to be a usable practice.

Best regards, Rainer

Hi,

IBTSWG suggested on this topic to use 61, 62, 63, 64, 65 and 66 for the new maturity scales. For future new maturity scales the same code can be applied if a 6 point scale is used. It is recommended to refer to maturity staging workshop reports in the DATRAS maturity reference tables to have more information on the description of the stages. Is there any problem with this?

Cheers, Ingeborg

Status: Done by ICES Data Centre for all surveys (61 etc. as coding for the 6 point maturity scales)

BITS, data required for stock indices and additional stock information

Dear All,

Based on the data of BITS stored in the DATRAS database it is possible to download different aggregation levels of stock indices. Unfortunately, it is not possible to repeat the calculations based on the data of the exchange format without additional data which are not available on the ICES website. That are in the case of the Baltic Sea (BITS):

- conversion factors which are used to transform the CPUE values of the different gears into units of the standard gear TVL
- areas of the depth layer of the different ICES subdivisions which are used as weighting factors for estimating the stock parameters
- description of the procedures for aggregating the data

From my point of view it is necessary that all information which are used for estimating stock indices must be available at the website. This is necessary because the two ways of calculation of stock indices based on the DATRAS database and based on the own calculation produce different estimates in some cases.

It is necessary that the procedures used in the DATRAS system for calculating aggregated data will be checked by members of the groups.

A second point is the intensive use of the available data. The data of BITS can be used for estimating the weight at age of the stock and the maturity ogive, etc. Unfortunately, this is not done until now by the Datras system to standardized the calculation process although German prepared the algorithms many years ago.

I propose that the survey working group prepare algorithms for estimating as much as possible standard result based on the survey to improve the standardization of the outcome of the surveys.

Rainer

Rainer and others,

I agree on your last point, in the sense that at least the survey planning/working groups have to check the DATRAS output. I know that there haven't been problems with the beam trawl survey data, too and I must admit that I don't know the status exactly at this very moment. I will put it on the WGBEAM agenda, I hope chairs of IBTSWG and WGBIFS will do the same.

As for the algorithms: maybe Vaishav/Anna can give us any direction on the way we can deliver the algorithms to you so you will be able to handle those easily.

Cheers, Ingeborg

Status: to be followed up ICES Data Centre (example and directions on the algorithms should be sent to ICES Data Centre by the planning WG's (IBTSWG, WGBIFS, WGBEAM, et al.)),

IBTS data query: CA Records on IBTS data files

Hello fellow DUAP'ers,

I suppose this query is for Vaishav and any other IBTS people. I am currently in the process of appending Individual live fish weights data onto our DATRAS transfer files for the 2009 bottom trawl surveys and have noticed that there is only one field at the end of the CA record format for individual fish weight (IndWgt). At the moment the plan is to add in the live or whole weight into this field but Ken and I both thought it might be worthwhile checking whether in fact there is a plan to expand the field number to incorporate gutted weight also. Otherwise I will only include whole weight.

Happy New Year and thanks in advance, Finlay

Hi Finlay,

There is a flexibility in DATRAS reporting format that we can add new fields. It needs to be agreed by working groups.

If there is a requirement of expanding weight column then it can be done.

Currently there is no such requirement.

Cheers, Vaishav

Hi Finlay and all,

I don't think that gutted wt is a survey-typical quantity to be stored in DATRAS. Of course, you can gather a lot of information from single fish during a survey, but as long as there is no demand from a survey planning group, we shouldn't inflate the database and occupy the DATRAS development team.

Regards, Uli

Hi Finlay and all,

I agree with Uli, we can expand the fields in DATRAS to include all kind of data, but if there's no request or general agreement on its necessity, I think that the appropriate place to store that information is in the database of the Institute collecting and using the information, otherwise we'll end with a database full of NA data. So, for me, the way to proceed would be to consider if that information is collected in more surveys, and discuss its use, and eventually propose the inclusion of a new field.

Best, Fran

Status: since there seems to be no direct need for this field in DATRAS, no follow-up is needed.

Evaluation of the system

Using the sharepoint

It was a choice to let the DUAP be organised by correspondence via the sharepoint site. Since membership is open to anyone within the ICES community there is possibly a large selection of people with different expertises. Compared to a mailing list, the sharepoint has the advantage that people not involved from the beginning.

Solving problems

The discussion board on the groupnet seems to work. Most data suppliers are member of the group and so, expertise from different surveys (IBTS, BITS and beam trawl) comes together.

In general, the DUAP sharepoint discussions can be divided in two categories:

Discussions to provide feedback to the ICES Data Centre and/or the survey working groups uploading data in DATRAS (see sections X.4.1, X.4.5, X.4.6, X.4.7)

Discussions to provide information on problems that occur with uploading, downloading or processing survey data (see sections X.4.3, X.4.4, X.4.8)

In both instances the problem is solved by the group members and expertise is used in an optimal way.

Challenges for the future

The representation of the data-submitters is good, mainly since there is a clear list of DATRAS data-submitters. It is to be expected that the peak of discussions for this group will take place in the period where data submission is highest. At this moment, it is not necessary to undertake any action to get data submitters more involved.

As for the representation of data-users progress can be made in the participation. Up to now, only the members of survey working groups were asked to join the DUAP group, which results in a set of experts for the DUAP group. However, it is more difficult to approach people downloading data not being expected on the particular surveys. Options to increase DUAP discussions might be:

Put a note on the DATRAS downloading site to create awareness on the existence of the group

Active role for ICES Data Centre in starting discussions on DATRAS issues

.....

Annex 11: Protocol/Process for development of data products (DRAFT)

When an ICES advisory or science group has a wish for standard output from ICES Data Centre, there is a responsibility for the group as well as for the ICES Data Centre to come to tailor-made output. Knowing that the experts on the data use are situated within the advisory and science groups, and ICES Data Centre has expertise on creating standard output from the data sets store at ICES, a dialogue is necessary. For this, WGDIM proposes the following procedure. Basically option A and B are the same, although the order of actions varies.

Option A

An advisory or science group requests standard output from a known data set

- 1) the advisory or science group calculates the requested output, based on a selection of the dataset
- 2) the group sends a request to ICES Data Centre to include the output as a standard ICES Data Centre output. Together with the request, the group delivers:
 - a. a copy of the raw dataset used and its origin
 - b. the selection criteria used to create the dataset (like selected years, quarters, aggregation levels of specific variables, etc.)
 - c. the algorithm used for the calculation, including specific information on excluded data or additional selection criteria
 - d. the outcome of the calculation
- 3) ICES Data Centre creates the output based on the information provided under 2a, b, c with its own programmes and compares the outcome to 2d. The results of the check will be archived by the ICES Data Centre.
- 4) ICES Data Centre sends the outcome of the comparison to the chair of the advisory or science group.
- 5) when the outcome of the ICES Data Centre algorithm is identical to the outcome of the WG, the algorithm can be used for creating standard output on the ICES website. When the outcome is not identical, a discussion between the advisory or science group and ICES Data Centre is necessary on potential sources of errors.

Example A:

IBTSWG requests international North Sea index calculations for red mullet in the IBTS Q1 survey

- 1) IBTSWG extracts the North Sea Q1 red mullet raw data for a given year from DATRAS, and calculates the international index in accordance with the standard IBTS index calculation
- 2) IBTSWG sends a request to ICES Data Centre to include the output as a standard ICES Data Centre output. Together with this request, IBTSWG delivers:
 - a. a copy of the raw data set used

- b. information on the selection criteria: species=red mullet, year=1990, survey=IBTS, quarter=1
 - c. the algorithm of the calculation, including the additional criteria used, i.e. only day hauls are taken into account, first average numbers per hour per statistical rectangle are calculated and then numbers per hour per roundfish area, sexes are aggregated, length measurements of the age readings are transformed from mm class to cm class to the cm below (so: 25.9 cm → 25 cm),
 - d. the table with the numbers per hour per age group for the North Sea
- 3) ICES Data Centre creates the output based on the information provided under 2a, b, c with its existing IBTS index calculating programmes and compares the outcome to 2d.
 - 4) ICES Data Centre sends the outcome of the comparison to the chair of IBTSWG.
 - 5) when the outcome of the ICES Data Centre algorithm is identical to the outcome of the IBTSWG, the index for red mullet will be standard output on the ICES website.
 - 6) When the outcome is not identical, a discussion between the IBTSWG and ICES Data Centre is necessary on potential sources of errors.

Option B

An advisory or science group requests output from information outside its expertise area

- 1) the advisory or science group sends a request to ICES Data Centre to deliver output from a data set outside its expertise area. Together with the request, the group delivers:
 - a. a specific description of the data requested, e.g. bottom sea water temperature in the Skagerrak-Kattegat 1990-2009
 - b. the aggregation level needed
 - c. a specified format of the output (maps, tables,) requested
- 2) ICES Data Centre gets into contact with the experts on the data requested and forwards the request of the expert or advisory group to the experts
- 3) follow **option A**, where 'advisory or science group' should be read 'the experts'
- 4) ICES Data Centre sends the outcome of option A to the advisory or science group that requested the data and publishes the output as standard output on the website

Example B

WGNSSK requests water temperatures in Skagerrak/Kattegat

- 1) WGNSSK sends a request to ICES Data Centre to deliver water temperatures for Skagerrak/Kattegat. Together with the request, the group delivers:
 - a. a specific description of the data requested:

- bottom water temperatures in the Skagerrak-Kattegat
 - period 1990-2009
- b. the aggregation level needed: data aggregated per year, quarter for the complete area
- c. Excel table including the columns:
- average bottom temperatures
 - year
 - quarter
 - number of observations used
- 2) ICES Data Centre contacts hydrographic experts (e.g. present in WGOH) and forwards the request
- 3) Follow option A, read WGOH/hydrographic experts instead of IBTSWG
- 4) ICES Data Centre sends the Excel file to WGNSSK and publishes the output as standard output on the website