

Theme session R

Environmental and fisheries data management, access, and integration

Co-Conveners: Christopher Zimmermann (Germany), Helge Sagen (Norway), and Peter H. Wiebe (USA).

Management of data and information within the ICES community has become a very important component in the quest to provide more comprehensive and time-sensitive advice to ecosystem and fishery managers. There are significant difficulties in integrating diverse fishery, oceanographic, and other marine environmental data, and the tools to enable fishery and environmental assessments needed to respond to the requirements for ecosystem-based management initiatives are still in a remedial state of development. There are a growing number of databases and data sources requiring new approaches to enable efficient access to the data. In addition, with the advent of new technologies and sensors to acquire oceanographic and fisheries data, new approaches are needed to acquire, document, manage, and distribute them to meet the enhanced requirements for data access and synthesis.

This theme session provided the opportunity to update the community on new approaches and endeavours by inviting database specialists, distributed data specialists, visualization specialists, end-users and others to present and/or demonstrate.

- technical solutions for data integration.
- novel ways to merge/integrate/distribute disparate data.
- experience with use and visualization of integrated data.
- data quality assurance and indicators.
- improved methods for displaying complex data sets.
- examples of successful applications.
- experience with (historic) data rescue.

The theme session initially attracted more than 45 contributions and thus was the most “popular” session during this year’s science conference. There were ultimately 12 posters and 25 oral talks in the session. Participants in the session ranged from about 50 in the beginning to about 47 at the end; most sub-sessions had around between 35 and 45 participants.

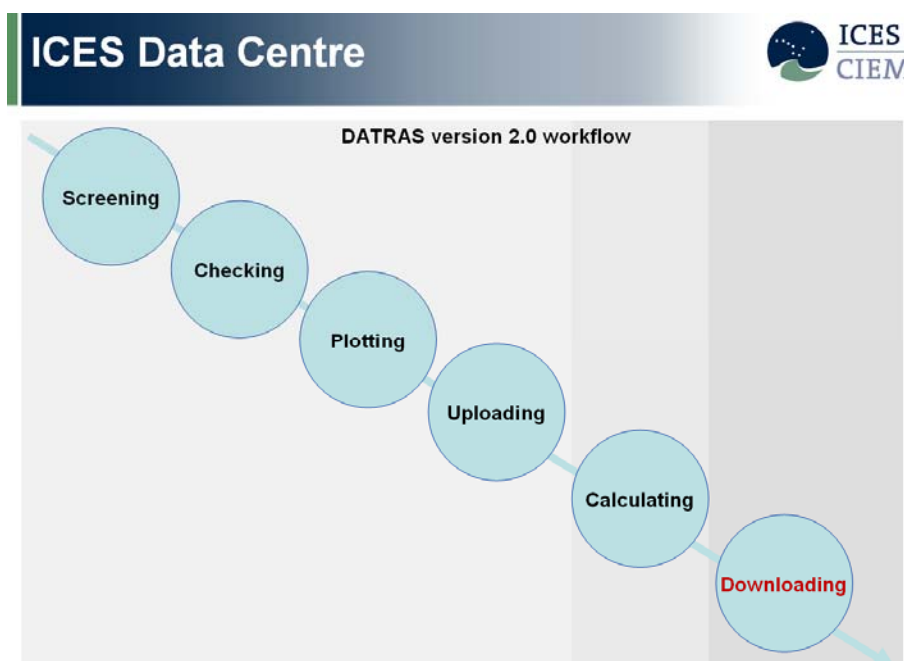
Theme session R was structured with 4 major sections.

There were four papers section “(1) Data acquisition”. These papers highlighted the advances that have taken place recently in the technology that now exists to assist investigators collect data and metadata at sea, conduct preliminary processing, prepare visualizations of the data for real-time display and decision making at sea, and prepare reports. The systems described also had the capability for real-time transmission of the data to shore for institutional purposes. In addition, there was a description of a federation of nations in the PICES community that are now able to share their metadata for Pacific wide marine data discovery through a central server providing a clearinghouse with standard protocols and formats.

There were 11 papers presented in section “(2) Data Management”. An overview talk on recent developments concerning data standards highlighted the main outcomes of a workshop held in 2008. Conclusions were reached on three topics, date and time, position, and country codes, all of which have ISO standards to be recommended for usage within the data management community. Three other important categories including quality control and flags, controlled vocabularies and ontology's, and discovery metadata will need attention in the future.

A number of systems were presented that were designed to solve local challenges in dealing with fishery data. All systems have given priority to using open source solutions. Most problems are solvable using such software. However, the data in these fishery management systems are for the most part still not open to the public access, but some are available through cooperation.

There were six papers scheduled in “(3) Data application” section, but one was not presented. The recovery of some very old fisheries data was reported for the Adriatic Sea illustrating the value of reaching back in time to put the current status of the fish stocks in perspective. In contrast, a real-time data acquisition, data transmission, and automatic processing system was described to enable the environmental properties of the water in which fish were being reared in very large quantities in Norway to be assessed and reports of the environmental quality to be returned to the fishers. An application of fisheries data from Canada stored in OBIS was described that was designed to evaluate the various possible processing methods to generate stock indices and their effectiveness relative to a known temporal pattern in the stocks. A few of the shortcomings in using data currently stored in ICES DATRAS were highlighted that involved taxonomic mis-identifications and recommendations were made for adjustments to the data to avoid the resulting biases. Finally, the ICES and WODC physical data were used to



ICES CM 2008/R:41. Database Trawl Surveys. Vasihav Suni

generate a new high resolution climatology for the North-West European shelf seas. This data set will reside on the ICES database system and be publicly available to all interested users.

The last section consisted of four papers dealing with “(4) Data visualization”, but only three were presented during the session. A management tool for fishery data designed for decision makers was described in which the data were presented using a GIS interface. The interface lets the users build their own maps by selecting the underlying data. Information can be presented in three different ways, geo-statistically, statistically, and symbolically. GIS tools were described in which a common interface was used for visualization of data. Open Source tools were used if found appropriate. Visualization and analysis of particle tracking model output for fish and fish larvae showed possibilities in Open Source tools like GeoModeler, GeoTools, and Visualization Toolkit.

In addition to the presentations, at the end of the first day, there was a 40+ minute “**Interactive session on a multi-annual workplan of the ICES Data Centre and the Working Group on Data and Information Management**” with an introduction by Neil Holdsworth, ICES Head of Data Centre. This was a novel approach during an ASC and aimed at learning from data users how they think ICES data holdings could better serve their needs. It is expected that a series of concise workshops to deliver a clearly defined product attract users to participate in ICES data groups. Ideas for topical workshops to be conducted over the next years include:

- publish survey data on web through an interface to the ICES DC. The first example would be the Mackerel Egg Survey data that is collected during an international and ICES-coordinated multiship effort, and that is quality controlled, but so far not accessible
- create standard products according to clients needs (OSPAR as an example)
- develop ODBC interfaces for all data held under ICES auspices, aiming at making the production of products by users themselves easier; also create a set of standard views to ease access to data subsets for scientists not familiar with SQL
- harmonise quality flags, develop a standard at least for ICES coordinated surveys and fisheries data
- retrieve historic, non-digital data (metadata first) – or during a workshop, develop a strategy to retrieve these data
- give training courses in SQL and data handling for non-data administrators

The **discussion at the end of the theme session R** highlighted that all of the presentations, in some form or another presented a visualisation of data – most noticeably with various flavours of GIS. This demonstrates that end users, regardless of purpose, demand more than a simple delivery of data to their computer. It was also noted from data acquisition to application that the users demand speed of delivery and more information without compromises on quality.

Many of the presentations demonstrated methods to empower the end user, through web services, web features, data discovery, and other means. The question is whether these advances are still at the expert user level and how scientists with little interest in the technicalities of data can exploit these services. Open source technology was prevalent in many of the data management and web solutions presented; this was seen as a positive step towards the re-use of systems for similar or comparative services.

As the marine data community expands its horizons, there is a continual and urgent need to address the number of standards, best practices, and interoperability procedures we are faced with and work towards minimizing the creation of new ones without first examining what is already available or what might fit the emerging needs.

There were examples both on the national and international level of data clearing houses, virtual data nodes that harvest the descriptions of datasets and make this information available to other users and provide the means to discover data in a quick and easy manner. Hand in hand with these developments, there was a timely reminder that data quality issues and data re-use must be considered in the context of how the data were collected and prepared – however, the trend towards ‘democratic’ data use is growing and this discussion will develop further.

At the conclusion of the theme session, the question of holding another theme session was raised. The audience was strongly in favour of holding another session within a two year time frame because substantial advances in the area of management of marine data were anticipated. A proposal for a theme session in collaboration with PICES was prepared and submitted to ConC.